

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE ECONOMÍA Y RELACIONES INTERNACIONALES
Programa de Doctorado en Ciencias Económicas



Impacto de las Capacidades de Analíticos de Big Data en el proceso de toma de decisiones y el desempeño financiero: el caso de las empresas en Tijuana, Baja California

Para obtener el grado de:
Doctora en Ciencias Económicas

Presenta:
Sara Cristina Valenzuela Valenzuela

Director de tesis doctoral
Dr. Natanael Ramírez Ángulo

Tijuana, Baja California, noviembre de 2023

Dedicatoria

A la memoria de mi madre, Altagracia, que siempre vivirá en mi corazón.

Te amo, siempre.

Sara Cristina Valenzuela Valenzuela

Agradecimientos

Agradezco la oportunidad brindada por la Universidad Autónoma de Baja California y la Facultad de Economía y Relaciones Internacionales para realizar éste doctorado. Asimismo, el apoyo económico otorgado por el Consejo Nacional de Humanidades, Ciencia y Tecnología.

Al Dr. Natanael Ramírez Ángulo por su acompañamiento como director de tesis, por sus oportunas sugerencias, apoyo y recomendaciones para cumplir en tiempo y forma. Asimismo a mis lectores, el Dr. José Gabriel Aguilar Barceló y el Dr. Duniesky Feitó Madrigal, por su tiempo y oportuna retroalimentación, así como a mis lectores externos, la Dra. Deisy Milena Sorzano Rodríguez y el Dr. Arturo Serrano Santoyo, por compartir su experiencia y perspectiva.

A mis compañeros de posgrado por acompañarnos en ésta travesía y por hacerla menos solitaria. A la Dra. Ana Isabel Acosta, coordinadora del doctorado, a mis profesores, al personal administrativo y a todas las personas de la universidad y facultad que de alguna forma contribuyeron a la culminación del posgrado.

A la Hamburg University of Applied Sciences, especialmente al Dr. Walter Leal, a la Queen Mary University London, Durham University, University of Leeds, Instituto Tecnológico y de Estudios Superiores de Monterrey, Johannes Kepler Universität, así como a todas las instituciones, y personas que las conforman, que han contribuido a mi desarrollo académico, profesional y personal.

Agradezco especial e infinitamente a mi mamá, por todo su esfuerzo, dedicación, amor y cuidados, siempre incondicionales. Por inculcarme la importancia del estudio y el aprendizaje, por estar siempre presente, por creer en mí y porque aunque ya no estés conmigo, tu amor ha trascendido, pues eres y seguirás siendo parte de lo que soy.

A mis hermanas, Nubia y Pilar, a mis tías, abuelas, amigas, y a todas las mujeres que han sido parte de mi vida, que me inspiran y son un gran ejemplo. A las personas que me han acompañado en la vida, especialmente a Rafael, que me alientan a seguir, y comparten su tiempo, apoyo y conocimiento. Éste trabajo es un cúmulo de todas las experiencias, vivencias, amor, cariño y apoyo que he recibido a lo largo de mi vida de personas realmente valiosas, con quiénes comparto y a quiénes dedico el resultado de un gran viaje de aprendizaje.

Resumen

El crecimiento sin precedentes en el volumen, variedad y velocidad de los datos generados por las empresas y su entorno en los últimos años, ha llevado a las empresas a explotar los datos para extraer conocimiento y aplicarlo a sus procesos de toma de decisiones, permitiendo así generar ventajas competitivas. En la presente investigación se aborda el desarrollo de las capacidades de analíticos de Big Data en las empresas y su impacto en el desempeño proceso de toma de decisiones y el desempeño financiero de las mismas, caracterizando a las capacidades de analíticos de Big Data como un índice formativo multidimensional (calidad de los datos, tamaño de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de las herramientas).

Buscando aportar a la literatura local y regional, la presente investigación se centra en las empresas de Tijuana, Baja California, independientemente de su tamaño y de la industria a la que pertenecen. Derivado de la inexistencia de fuentes de información pertinente, fue necesario realizar trabajo de campo, diseñando un instrumento de recolección de datos en concordancia con los lineamientos de la literatura internacional, el cual fue aplicado a 139 directivos de empresas en la ciudad de Tijuana. Posteriormente, se utilizó un modelo de ecuaciones estructurales con mínimos cuadrados parciales, también en concordancia con la literatura de referencia, para analizar las relaciones entre las variables latentes: capacidades de analíticos de Big Data, desempeño del proceso de toma de decisiones y desempeño financiero de las empresas.

Aplicando la metodología descrita, se encontró que a mayor desarrollo de las capacidades de analíticos de Big Data, las empresas muestran un mejor desempeño de su proceso de toma de decisiones y un mejor desempeño financiero, sin embargo, no se encontró evidencia de un efecto mediador del desempeño del proceso de toma de decisiones entre las capacidades de analíticos de Big Data y el desempeño financiero.

Palabras clave: capacidades de analíticos de Big Data, toma de decisiones, desempeño financiero, calidad de los datos, tamaño de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de las herramientas

Abstract

The unprecedented growth in the volume, variety and velocity of data generated by companies and their environment in recent years has led companies to exploit data to extract knowledge and apply it to their decision-making processes, allowing them to generate competitive advantages. The present research addresses the development of Big Data analytics capabilities in companies and their impact on their decision-making process performance and financial performance, characterizing Big Data analytics capabilities as a multidimensional formative index (data quality, data size, analytical skills, domain knowledge and tool sophistication).

Seeking to contribute to local and regional literature, the present research focuses on companies in Tijuana, Baja California, regardless of their size or industry to which they belong. Due to the lack of relevant sources of information, it was necessary to conduct field work, designing a data collection instrument in accordance with the guidelines of international literature, which was applied to 139 company managers in the city of Tijuana. Subsequently, a structural equation model with partial least squares was used, also in accordance with the reference literature, to analyze the relationships between the latent variables: Big Data analytics capabilities, performance of the decision-making process and financial performance of the companies.

Applying the described methodology, it was found that the higher the development of Big Data analytics capabilities, the better the performance of their decision-making process and the better the financial performance, however, no evidence was found of a mediating effect of the performance of the decision-making process between Big Data analytics capabilities and financial performance.

Keywords: Big Data analytics capabilities, decision making, financial performance, data quality, data size, analytical skills, domain knowledge, tool sophistication.

Índice de contenidos

Introducción	10
Capítulo I.	16
Estado del arte sobre la relación entre Capacidades de Analíticos de Big Data y el desempeño de las empresas	16
¿Qué son los Analíticos de Big Data?	16
Relación entre ABD y el desempeño de las firmas: evidencia empírica	24
Capítulo II.	54
Marco Teórico	54
Teoría Visión Basada en los Recursos	54
Teoría de las Capacidades Dinámicas	56
Teoría de la Visión Basada en los Recursos de Tecnologías de la Información	59
Big Data	61
Analíticos de Big Data	64
Capacidades de ABD	66
Recursos tangibles	68
Recursos humanos e intangibles	69
Teoría de los Efectos de TI avanzadas en el Diseño Organizacional, Inteligencia y Toma de Decisiones	71
Desempeño de la firma	73
Entorno económico en Tijuana, Baja California	76
Hipótesis de investigación	79
Capítulo III.	86
Diseño metodológico	86
Modelo PLS-SEM	86
Criterios de aplicación de un modelo PLS-SEM	87
Algoritmo PLS-SEM y modelos de trayectoria	88
Evaluación de un modelo PLS-SEM	92
Evaluación de modelos de medición reflectivos	96
Evaluación de modelos de medición formativos	97
Evaluación del modelo estructural	99
Modelos de segundo orden	102
Tamaño de la muestra y Bootstrapping	104
Datos	106
Diseño del instrumento de recolección de datos	108
Muestra y recolección de datos	115
Modelo empírico	122

Capítulo IV.	129
Análisis y resultados	129
Caracterización de las empresas de la muestra	129
Evaluación del modelo PLS-SEM	140
Evaluación del modelo de medición para constructos de primer orden	141
Evaluación de constructos reflectivos de primer orden	145
Evaluación de constructos formativos de primer orden	148
Evaluación del modelo de medición para constructos de segundo orden	155
Evaluación del modelo estructural	164
Evaluación del efecto mediador del DPTD	170
Impacto de las CABD en el desempeño del proceso de toma de decisiones y financiero de las empresas en Tijuana, Baja California	171
Capítulo V.	173
Conclusiones y perspectivas	173
Conclusiones	173
Perspectivas finales: estudiando el impacto del uso de datos para la toma de decisiones	176
Capítulo VI.	179
Referencias bibliográficas	179
Anexos	186
Anexo 1. Código R (RStudio) utilizado para la aplicación del modelo PLS-SEM	186

Índice de Tablas

Tabla 1. Medición del impacto de las capacidades de ABD en el desempeño de la empresa.....	50
Tabla 2. La teoría RBV y la TCD como marco teórico en la literatura revisada.....	58
Tabla 3. Operacionalización de Analíticos de Datos en la literatura revisada.....	67
Tabla 4. Operacionalización de Desempeño en la literatura revisada.....	76
Tabla 5. Metodologías utilizadas en el análisis de las capacidades de analíticos de datos.....	87
Tabla 6. Criterios para la aplicación de un modelo PLS-SEM.....	88
Tabla 7. Criterios de evaluación del modelo PLS-SEM.....	95
Tabla 8. Tamaños mínimos de muestra para diferentes niveles de coeficiente de trayectoria (p_{min}) y un poder de 80%.....	105
Tabla 9. Instrumento de recolección de datos.....	109
Tabla 10. Estratificación de empresas de acuerdo al número de trabajadores.....	112

Tabla 11. Sectores de actividad económica (SCIAN, 2018).....	113
Tabla 12. Modificaciones a instrumento de recolección de datos.....	116
Tabla 13. Características de la muestra - Respondientes.....	120
Tabla 14. Características de la muestra - Empresas.....	122
Tabla 15. Estadística descriptiva de los ítems del instrumento de recolección de datos.....	138
Tabla 16. Fiabilidad de los indicadores de constructos reflectivos de primer orden.....	145
Tabla 17. Fiabilidad de consistencia interna y validez de convergencia de constructos reflectivos de primer orden.....	146
Tabla 18. Validez discriminante de constructos reflectivos de primer orden (HTMT).....	148
Tabla 19. Resultados de modelos PLS-SEM para análisis de redundancia de constructos formativos de primer orden.....	145
Tabla 20. Evaluación de colinealidad para los constructos formativos de primer orden.....	150
Tabla 21. Significancia del peso de los indicadores en constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	151
Tabla 22. Cargas y significancia estadística de los indicadores en constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	154
Tabla 23. Resultados de modelos PLS-SEM para análisis de redundancia de constructos formativos de segundo orden.....	160
Tabla 24. Evaluación de colinealidad para los constructos formativos de segundo orden.....	160
Tabla 25. Significancia del peso de los indicadores en constructos formativos de segundo orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	161
Tabla 26. Cargas y significancia estadística de los indicadores en constructos formativos de segundo orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	163
Tabla 27. Evaluación de colinealidad de constructos endógenos del modelo estructural.....	167
Tabla 28. Significancia de los coeficientes de trayectoria del modelo estructural (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	162
Tabla 29. Significancia de los coeficientes de trayectoria totales del modelo estructural (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	169
Tabla 30. Evaluación del poder explicativo del modelo.....	169
Tabla 31. Tamaño del efecto f^2 de los constructos exógenos del modelo.....	170
Tabla 32. Evaluación del efecto mediador del DPTD entre las CABD y el DF.....	170

Índice de Figuras

Figura 1. Objetivos y preguntas de investigación	14
Figura 2. Historia y evolución del Big Data.....	19
Figura 3. Revolución tecnológica versus desarrollo de datos.....	21
Figura 4. Marco de evaluación de los recursos de la firma.....	55
Figura 5. Teoría de las Capacidades Dinámicas y la creación de ventajas competitivas.....	57
Figura 6. Clasificación de los recursos de TI.....	60
Figura 7. Definición de Big Data de acuerdo a las características de los datos.....	63
Figura 8. Definición y proceso de Analíticos de Big Data.....	65
Figura 9. Caracterización de las Capacidades de Analíticos de Big Data.....	68
Figura 10. Desempeño del proceso de toma de decisiones de la empresa.....	73
Figura 11. Valor estratégico de los ABD.....	75
Figura 12. Marco teórico de investigación.....	83
Figura 13. Objetivos, preguntas e hipótesis de investigación.....	85
Figura 14. Modelo de investigación propuesto.....	85
Figura 15. Etapas del algoritmo básico PLS-SEM.....	89
Figura 16. Ejemplo de un modelo de trayectorias simple.....	91
Figura 17. Evaluación de un modelo PLS-SEM.....	93
Figura 18. Proceso de evaluación de un modelo PLS-SEM con constructos reflectivos.....	96
Figura 19. Proceso de evaluación de un modelo PLS-SEM con constructos formativos.....	98
Figura 20. Proceso de decisión para incluir o eliminar un indicador formativo.....	99
Figura 21. Proceso de evaluación del modelo estructural.....	100
Figura 22. Modelo de segundo orden tipo II.....	103
Figura 23. Evaluación de resultados del modelo de medición para modelos con constructos de segundo orden (formativos) - Enfoque de dos etapas (<i>two-stage</i>).....	104
Figura 24. Proceso de recolección de datos.....	119
Figura 25. Modelo de investigación propuesto.....	124
Figura 26. Caracterización de la muestra - Respondiente.....	131
Figura 27. Caracterización de la muestra - Empresa.....	133
Figura 28. Utilización de analíticos de datos para la toma de decisiones.....	134

Figura 29. Indicadores de desempeño financiero de la empresa.....	135
Figura 30. Uso de analíticos de datos e indicadores de desempeño de la empresa.....	137
Figura 31. Modelo de investigación propuesto.....	141
Figura 32. Modelo de medición para evaluar los constructos de primer orden.....	143
Figura 33. Gráfico de resultados del modelo de medición de constructos de primer orden.....	144
Figura 34. Fiabilidad de consistencia interna y validez de convergencia de constructos reflectivos de primer orden.....	147
Figura 35. Modelos PLS-SEM para análisis de redundancia de constructos formativos de primer orden.....	149
Figura 36. Pesos y prueba de significancia estadística de indicadores para constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	153
Figura 37. Modelo de investigación propuesto.....	156
Figura 38. Resultados del modelo PLS-SEM.....	157
Figura 39. Modelos PLS-SEM para análisis de redundancia de constructos formativos de segundo orden.....	159
Figura 40. Pesos y prueba de significancia estadística de indicadores para constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras <i>bootstrapped</i>).....	162
Figura 41. Modelo de investigación propuesto.....	165
Figura 42. Resultados del modelo PLS-SEM.....	166
Figura 43. Hipótesis, modelo estructural empírico propuesto y resultados.....	172
Figura 44. Objetivos, preguntas e hipótesis de investigación y resultados.....	176

Introducción

En los últimos años la importancia del uso de datos para la toma de decisiones ha crecido de forma exponencial en las empresas. Específicamente, se ha encontrado que aquellas firmas que adoptan un modelo con toma de decisiones basadas en datos presentan una mejora en medidas objetivas sobre desempeño financiero y operacional (McAfee y Brynjolfsson, 2012). El crecimiento sin precedentes en el volumen, variedad y velocidad de los datos en la última década es un fenómeno al que comúnmente se le denomina “Big Data” y los analíticos derivados de estos datos han transformado la forma en la que las firmas compiten (Müller et al., 2018).

La posibilidad de que las firmas generen una ventaja competitiva derivada del uso de Analíticos de Big Data (ABD) ha sido mayor con el paso del tiempo, pues los costos de almacenaje, memoria, procesamiento, ancho de banda, entre otros, han disminuido con los avances tecnológicos, permitiendo que más empresas adopten ABD en sus procesos de toma de decisiones de una forma menos costosa (Ferraris et al., 2019). ABD es el proceso necesario para comprender el conglomerado de datos con el fin de extraer y generar información y conocimientos útiles (Chen et al., 2012), que, visto de forma holística y soportado con soluciones específicas para ABD, dichos conocimientos se vuelven prácticos para ofrecer valor sostenido, medir el rendimiento y establecer ventajas competitivas (Fosso Wamba et al., 2015). Dado el potencial y la popularidad de las soluciones de ABD, han surgido cuestionamientos sobre si el fenómeno Big Data contribuye a la creación de valor y de ventajas competitivas, y cómo contribuye a ello, así como los factores que influyen o determinan los efectos del Big Data (Sena et al., 2019). En sentido estricto, ABD se define como un campo de toma de decisiones que consiste en Big Data, el uso de varias herramientas y técnicas estadísticas, aprendizaje automático, inteligencia artificial, aprendizaje profundo e incluso minería de datos para obtener información potencial de enormes conjuntos de datos, mejorando el desempeño del mercado y desempeño financiero de una empresa (Gupta et al., 2018).

Por otra parte, con sustento en la teoría visión basada en los recursos, las capacidades de ABD (CABD), para efectos de la presente investigación, se caracterizan como la capacidad de la empresa para desarrollar, desplegar y combinar recursos de analíticos de datos (calidad de los datos, tamaño de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de

las herramientas) que permiten extraer el conocimiento de dichos datos y permearlo a todos los niveles de la organización, incorporándolo a su proceso de toma de decisiones, buscando construir una ventaja competitiva (Ferraris et al., 2019; Ghasemaghahi et al., 2018; y Bharadwaj, 2000). Específicamente, los recursos tangibles, grandeza de los datos y sofisticación de las herramientas, se desarrollan a través de la incorporación y uso de herramientas de tecnologías de la información que permitan a la empresa recolectar y procesar un gran volumen de datos, precisamente para extraer conocimiento de los mismos (Ghasemaghahi et al., 2018; Detlor et al., 2013; y Jeble et al., 2018). Mientras que los recursos humanos e intangibles, calidad de los datos, dominio del conocimiento y habilidades técnicas, se desarrollan mediante la recolección de datos exactos, objetivos, precisos, consistentes y accesibles, así como recursos humanos con un entendimiento profundo de los procedimientos, hechos y procesos de la empresa, su industria, el entorno socioeconómico y del mercado en el que opera, y con las habilidades técnicas necesarias para transformar los datos en conocimiento accionable (Ibid, 2018, 2013, 2018).

Diferentes estudios han encontrado una relación positiva entre la implementación de CABD y la creación de valor en las empresas (Côte-Real et al., 2017; Chen et al., 2012), el desempeño financiero (Müller et al., 2018; Torres et al., 2018; Bogdan y Borza, 2019; Behl, 2019; Maroufkhani et al., 2019), y el desempeño del proceso de toma de decisiones (Ghasemaghahi et al., 2018). No obstante, algunos estudios también advierten a los tomadores de decisiones sobre el hecho de que el Big Data no es una panacea (Akter et al., 2016), “un análisis crítico de conjuntos de datos mal entendidos no genera conocimiento” (Matthias et al., 2017). Asimismo, se ha encontrado que la mayoría de las firmas no puede aprovechar al máximo el uso de herramientas de ABD debido a una variedad de razones, como tener datos de baja calidad, no usar herramientas de análisis de datos adecuadas y el lago de habilidades analíticas disponibles (Colas et al., 2014).

Lo anterior sugiere que es necesario investigar sobre las CABD y su impacto en el proceso de toma de decisiones y el desempeño financiero de las firmas. Para el análisis sobre las competencias y CABD y su efecto en el proceso de toma de decisiones de la firma se toma el modelo desarrollado por Ghasemaghahi, et., al. (2018), el cuál categoriza la capacidad de analíticos de datos como “sofisticación de herramientas (Infraestructura de Tecnologías de la Información o TI), habilidades analíticas y dominio del conocimiento que tienen los empleados

(recursos humanos de TI) y calidad y tamaño de datos (intangibles habilitados para TI)” y la utiliza para medir el impacto que tiene en el desempeño de la toma de decisiones a través de la calidad y eficiencia de las mismas. Mientras que para el impacto de las CABD en el desempeño financiero, se utiliza el trabajo de Ferraris et al. (2019), dónde prueba de forma empírica si las CABD (como CABD desde una perspectiva tecnológica y gerencial) tienen un impacto positivo en el desempeño de las empresas.

De forma más específica, la mayoría de la literatura revisada se concentra en el estudio de las CABD y su impacto en el desempeño financiero en economías desarrolladas, como Europa y América del Norte (Canadá y Estados Unidos de América), con un número muy limitado de estudios en economías emergentes (Anwar et al., 2018), por lo que la presente investigación busca contribuir a la literatura existente al proporcionar evidencia del impacto positivo y beneficios del desarrollo de las CABD en el desempeño del proceso de toma de decisiones y financiero de las empresas en México. Particularmente, se busca comprender el impacto de las CABD en el proceso de toma de decisiones y el desempeño financiero de las empresas en Tijuana, Baja California. Las CABD conducen a la mejora del desempeño financiero, ya que generan nuevos conocimientos, respaldan la toma de decisiones y mejoran la productividad de los procesos de la firma, pues conducen a una “planificación más eficiente (p. ej., variación en el cronograma y costos, utilización de la capacidad), proceso de fabricación (p. ej., tiempo de inactividad del proceso, eficiencia de la máquina, reducción de desperdicios) y garantía de calidad (p. ej., unidades defectuosas, unidades rechazadas)” (Popovič et al., 2018).

El dinamismo empresarial de la ciudad de Tijuana, Baja California, derivado de la ubicación geográfica de la ciudad y de la inversión extranjera que propició principalmente el crecimiento de la industria maquiladora de exportación buscando aprovechar ventajas de costos salariales y proximidad geográfica (Mendoza, 2010), una industria que históricamente ha utilizado analíticos de datos para la toma de decisiones, mediante la implementación de metodología Lean Six Sigma para mantener su competitividad (Castro, J. 2011). Asimismo, la presencia de empresas transnacionales, aunado a nuevas tecnologías, la revolución en las comunicaciones y la internacionalización de capitales, ha diseminado dicho dinamismo a otros sectores de actividad económica, reflejado en la modernización de empresas de otras industrias, que se han visto influenciados de forma positiva, permitiendo la adopción del uso de datos para

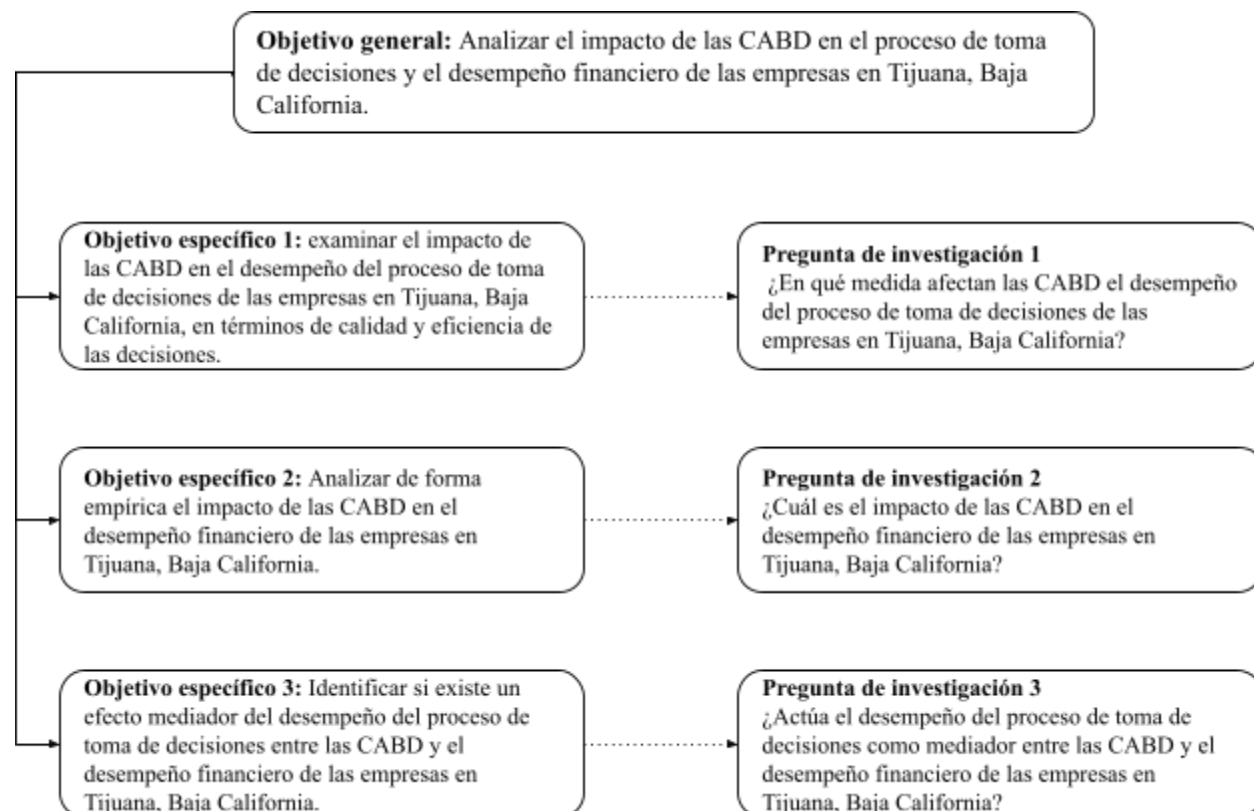
la toma de decisiones en empresas de todos los tamaños y de diferentes sectores (Puyana y Romero, 2006; Martínez Cuero, 2017).

Por otra parte, la investigación sobre el impacto de CABD en la región es inexistente, o bien, se centra en el análisis de la inteligencia de negocios como factor fundamental en la competitividad de empresas del sector de tecnologías de información y comunicación (Ahumada, 2016), quién encontró que el conocimiento es el activo de mayor valor en las empresas; el entorno de negocios es fundamental para la competitividad; la innovación, los sistemas de información y los procesos de toma de decisiones son parte de la inteligencia de negocios que requieren las empresas, y que la inteligencia de negocios desarrolla la competitividad a partir de la gestión del conocimiento. Lo que nos lleva a cuestionarnos si lo anterior se cumple cuando hablamos de CABD y de otros sectores económicos, tanto en lo relacionado con su impacto en el desempeño del proceso de toma de decisiones, como en lo referente a competitividad medida a través de desempeño económico. Aunado a la necesidad de literatura sobre el tema en la región, no se encontraron fuentes de datos relacionadas con la adopción del uso de datos para la toma de decisiones en empresas, independientemente de la industria o el tamaño, volviendo fundamental la realización de trabajo de campo para recabar información pertinente.

Entonces tenemos que el objetivo general de la presente investigación se centra en analizar el impacto de las CABD en el desempeño del proceso de toma de decisiones y en el desempeño financiero de las empresas en la ciudad de Tijuana, Baja California, México, pertenecientes a diversas industrias y de todos los tamaños. Mientras que los objetivos específicos son: 1) examinar el impacto de las CABD en el desempeño del proceso de toma de decisiones de las empresas en Tijuana, Baja California, en términos de calidad y eficiencia de las decisiones; 2) analizar de forma empírica el impacto de las CABD en el desempeño financiero de las empresas en Tijuana, Baja California; y 3) identificar si existe un efecto mediador del desempeño del proceso de toma de decisiones entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California. De lo anterior se derivan las preguntas de investigación, que son: 1) ¿En qué medida afectan las CABD el desempeño de la toma de decisiones de las empresas en Tijuana, Baja California?; 2) ¿Cuál es el impacto de las CABD en el desempeño financiero de las empresas en Tijuana, Baja California?; y 3) ¿Actúa el desempeño del proceso

de toma de decisiones como mediador entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California?; las cuales se muestran en la figura 1.

Figura 1. Objetivos y preguntas de investigación



Fuente: elaboración propia.

Para alcanzar los objetivos propuestos y responder a las preguntas de investigación, y dada la naturaleza y relaciones que se buscan probar, se propone un estudio de investigación utilizando la metodología de Modelo de Ecuaciones Estructurales con Mínimos Cuadrados Parciales (PLS-SEM por sus siglas en inglés), el cual se utiliza en el análisis y modelación de relaciones complejas entre variables latentes o constructos y sus indicadores (Sarstedt et al., 2018 y Hair et al., 2021). Para ello se realizó el diseño de un instrumento de levantamiento de información con las preguntas o indicadores que miden o conforman las variables o constructos latentes, posteriormente se realizó el levantamiento de la información de agosto a noviembre de 2022, con tomadores de decisiones en empresas de Tijuana, Baja California, y una vez recolectados se procedió al análisis de los datos para probar las hipótesis de investigación derivadas de los objetivos y preguntas descritos.

Considerando lo anterior, la investigación se estructura de la siguiente forma: en el capítulo I se revisa el estado del arte sobre la relación entre las CABD y el desempeño de diferentes áreas en las empresas, describiendo conceptos a utilizar durante la investigación; en el capítulo II se plantea el marco teórico de la investigación, describiendo las diferentes teorías y enfoques a utilizar, así como las hipótesis de investigación; en el capítulo III se detalla el proceso metodológico para la aplicación de un modelo PLS-SEM, desde los criterios de evaluación hasta el instrumento de levantamiento de información y el modelo teórico/empírico propuesto; en el capítulo IV se presenta el análisis y resultados del modelo empírico, incluyendo la caracterización de la muestra; y finalmente, en el capítulo V se presentan las conclusiones, perspectivas y limitación de la investigación.

Capítulo I.

Estado del arte sobre la relación entre Capacidades de Analíticos de Big Data y el desempeño de las empresas

Las firmas son cada vez más conscientes del potencial de recolectar, procesar y analizar datos para generar valor y mejorar su desempeño, sin embargo, no existe un consenso sobre cómo se puede generar dicho valor, pues con los avances tecnológicos cada vez se generan más datos y se complica la tarea de identificar cuáles analíticos y capacidades son necesarios para obtener una ventaja competitiva. En principio, la definición de Big Data se concentra en las características de los datos (Ferraris et al., 2019), mientras que la de analíticos está orientada a la información o conocimientos que podemos extraer de dichos datos (Cao et al., 2015).

Teniendo en cuenta lo anterior, en el presente capítulo se enfoca en tratar de definir qué son los Analíticos de Big Data (ABD) y cómo se relacionan con el desempeño de las firmas. Inicialmente, se esboza la historia y evolución de los analíticos de datos, desde el análisis estadístico hasta sus mutaciones propiciadas por los avances tecnológicos y el Big Data para llegar a los ABD, para posteriormente enfocarnos en el estudio de la relación entre ABD y el desempeño de las firmas.

¿Qué son los Analíticos de Big Data?

El uso de analíticos para la toma de decisiones es casi tan antiguo como el uso de la estadística, desde que la humanidad recopila datos, la necesidad de descubrir información útil y perspectivas en ellos ha estado presente (Foote, 2021). La definición literal de analítico de acuerdo con la Real Academia de la Lengua Española (RAE, 2023) es “perteneciente o relativo al análisis”, sin embargo, definiciones más actuales son más específicas y lo relacionan con el análisis estadístico de los datos, por ejemplo la definición de analítico por *Oxford Languages* de Google es “el análisis computacional sistemático de datos o estadísticos” o “información resultante del análisis sistemático de datos o estadísticos”.

Como podemos ver los analíticos están entrelazados con la historia de la estadística, pues los analíticos vienen de los datos y el análisis de los datos se basa en la estadística, la cuál se

especula ha sido utilizada desde la época de los antiguos egipcios para cálculos relacionados con la construcción de las pirámides (Foote, 2021). Los analíticos surgen precisamente del análisis de los datos para extraer conocimiento de los mismos: una vez que se recolectan los datos, es necesario procesarlos para intentar descubrir información útil (Smallcombe, 2021).

A lo largo de la historia, el análisis estadístico de los datos ha jugado un rol importante en la toma de decisiones. Por ejemplo, los gobiernos utilizan el análisis de los datos recopilados en los censos de población para planear actividades que van desde inversión en infraestructura y educación, hasta la recolección de impuestos (Smallcombe, 2021). Mientras que por el lado de las firmas, existen registros de casos sobre el uso de analíticos de datos para la toma de decisiones desde el siglo XIX, por ejemplo en 1878 Frederick Winslow Taylor inició con estudios sobre el análisis de la gestión del tiempo y remuneración del trabajo, mientras que Henry Ford midió la velocidad de las cadenas de ensamblaje para reducir los tiempos de producción e incrementar el volumen de la misma (Foote, 2021).

La tasa de creación de datos, su almacenamiento y procesamiento se ha disparado en la última década, gracias a la generalización del uso del internet, el uso de redes sociales y de tecnología en los negocios, siendo necesario implementar nuevas herramientas y estrategias para compartir, minar y visualizar los datos en aras de extraer información útil para tomar decisiones en diversas áreas de la empresa como comunicación, planeación, investigación y desarrollo entre otras (APICS, 2015).

Precisamente esos avances tecnológicos que permiten crear, almacenar y procesar los datos de forma cada vez más eficiente son los que han permitido que los analíticos de datos evolucionen a ABD, puesto que ahora tenemos muchos más datos que analizar. Tenemos entonces qué identificar cómo pasamos de los datos a los macrodatos o Big Data, por lo que a continuación se describen algunos de los eventos históricos clave para llegar al concepto de Big Data y posteriormente al de ABD.

La RAE define el Big Data como el “conjunto de técnicas que permiten analizar, procesar y gestionar conjuntos de datos extremadamente grandes que pueden ser analizados informáticamente para revelar patrones, tendencias y asociaciones, especialmente en relación con la conducta humana y las interacciones de los usuarios” (RAE, 2023). Sin embargo, mucho antes

de los datos extremadamente grandes, el primer análisis de datos utilizando técnicas estadísticas del que se tiene registro se remonta al siglo XVII con el estudio de John Graunt sobre la peste bubónica (Foote, 2017). Mientras que un par de siglos después, en 1865, Richard Millar Devens acuñó el término Inteligencia de Negocios utilizando como ejemplo a un banquero que utiliza información de su entorno para convertirla en utilidad (Phillips, 2021), a la par de los trabajos de Frederick Winslow Taylor y Henry Ford que utilizaron el conocimiento extraído del análisis de la información para la toma de decisiones en los negocios.

No obstante, cómo podemos observar en la figura 2, la historia del término Big Data está entrelazada con la evolución de la tecnología moderna, pues por primera vez la cantidad de datos y la capacidad de procesarlos se volvió un problema con el Censo de 1880 en EUA, el cuál se resolvió con la invención de la Máquina Tabuladora realizada por Herman Hollerith y que permitió reducir el tiempo de registro y análisis de los datos de diez años a tres meses. Asimismo, en 1911 fundó la Empresa de Informática-Tabulación-Registro (posteriormente *International Business Machines Corporation* o IBM), precursora en tecnologías para el análisis de grandes cantidades de datos (Phillips, 2021). Otro gran pilar histórico para la evolución del Big Data es el desarrollo de la cinta magnética por Fritz Pfleumer en 1928, la cuál sirvió como base para el almacenaje de información en video cassettes y carretes de películas (Foote, 2017).

Para 1943, los avances tecnológicos permitían realizar operaciones booleanas y contables para analizar grandes volúmenes de datos, pues derivado de la necesidad de intervenir y descifrar códigos alemanes en la segunda guerra mundial, el Reino Unido desarrolló Colossus, la primera computadora técnica (Foote, 2017). Mientras que para 1959, el término *Machine Learning*, ahora tan en tendencia, era acuñado por Arthur Samuel, un programador de IBM, y solamente unos años después, en 1965 el gobierno de EUA ya tenía planes de construir el primer edificio que serviría como centro de datos para almacenar millones de devoluciones de impuestos y huellas dactilares en cinta magnética. Posteriormente, en 1969 la Red de Agencias de Proyectos de Investigación Avanzada (ARPANET por sus siglas en inglés) creó la primera red de área amplia, que incluía los protocolos TCI/IP, que constituyen las bases para el internet como lo conocemos actualmente y que resulta fundamental para la creación, recolección y procesamiento de datos (Phillips, 2021).

Figura 2. Historia y evolución del Big Data



Fuente: elaboración propia con base en Phillips, 2021 y Foote, 2017.

Los eventos anteriores constituyen las bases para el desarrollo del Big Data, sin embargo con la era del internet en la década de los noventa, es cuando realmente se desarrollaron las tecnologías que permitieron su avance. Empezando con el concepto de *World Wide Web* acuñado por Tim Berners-Lee y Robert Cailliau, mientras trabajaban en la Organización Europea para la Investigación Nuclear (CERN por sus siglas en francés) en 1989 y el desarrollo de los comandos HTML, URL y HTTP en 1990, lo que constituye el internet como lo conocemos ahora y que tenía como objetivo compartir información utilizando un sistema de hipertextos (Phillips, 2021 y Foote, 2017).

Tan sólo seis años después, el almacenamiento de información digital ya resultaba ser más rentable que almacenar datos en papel, de acuerdo con un artículo publicado por R.J.T. Morris y B.J. Truskowski, mientras que en 1997 se registró el dominio google.com, empresa pionera en el desarrollo e innovación de tecnologías para el Big Data, y para 1998 Carlo Strozzi desarrolla NoSQL, una base de datos relacional de código abierto que permite almacenar y recuperar datos que no están modelados de forma tabular, permitiendo almacenar y analizar aún muchos más datos, dando pie a que por primera vez en 1999 se intentara cuantificar la cantidad de información digital disponible en el mundo (Foote, 2017).

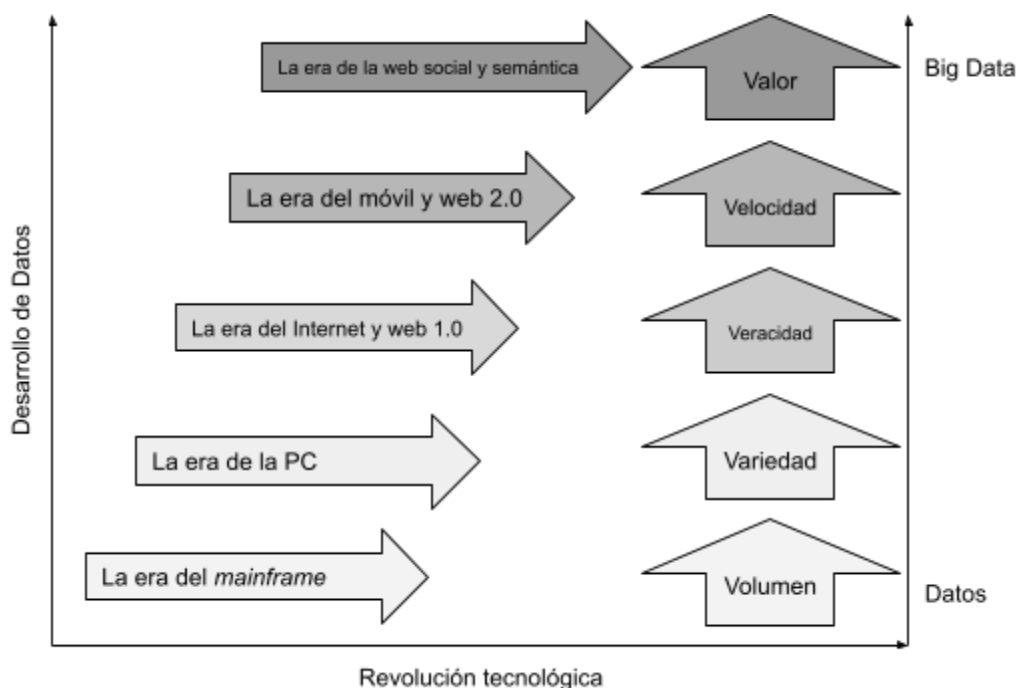
Para el siglo XXI, se desarrolla el Big Data cómo lo conocemos hasta ahora y es definido en 2001 por Doug Laney con las tres primeras V's (volumen, variedad y velocidad), las cuáles, como ya se mencionó, en la actualidad ascienden hasta 10, dependiendo del autor. La invención del marco de trabajo *Apache Hadoop* por Doug Cutting y Mike Cafarella en 2005, así como el inicio de la computación en la nube a través de *Amazon Web Services* en 2006, representan hitos en la generalización del uso de analíticos de datos para la toma de decisiones, llegando a convertirse en la prioridad número uno de los Directores de Información de las grandes empresas, de acuerdo con un reporte de Gartner (Phillips, 2021 y Foote, 2017).

En 2012, bajo la administración de Obama, EUA aprobaba la iniciativa *Big Data Research and Development*, comprometiendo 200 millones de dólares para la misma. Muy de acuerdo con el valor de mercado global del Big Data (10 billones de dólares en 2013) y con la creación acelerada de datos propiciada por el incremento en el uso de teléfonos inteligentes y,

por consecuencia de internet, pues, de acuerdo con un reporte de IBM de 2016, el 90% del total de datos existentes se habían generado en los últimos dos años (Ibid, 2021 y 2017)

En la actualidad, las firmas reconocen el potencial del Big Data para crear una ventaja competitiva a través del análisis de la información disponible, lo que ha llevado a una rápida aceptación y adaptación con el paso del tiempo y de la mano con el desarrollo de nuevas tecnologías (Anwar, et al., 2018). En la figura 3 podemos observar que las revoluciones tecnológicas han sido clave para el desarrollo y adaptación del Big Data, específicamente con el desarrollo del internet. Por ejemplo, durante la era del *mainframe* (computadora central) era posible almacenar y procesar grandes volúmenes de datos, al popularizarse el uso de las computadoras personales, se comenzaron a producir diferentes tipos de datos, agregando una dimensión más al concepto de Big Data, mientras que las de veracidad, velocidad y valor se desarrollaron con el uso generalizado del internet, los teléfonos móviles y las redes sociales (Ibid, 2018).

Figura 3. Revolución tecnológica versus desarrollo de datos



Fuente: elaboración propia con base en Anwar et al., 2018.

En todas partes del mundo, industrias y organizaciones recopilan más datos que nunca, buscando estrategias de negocio inteligentes que les permitan aprovechar la revolución del Big

Data, incorporándolas a sus modelos de negocio, como es el caso de Google, Amazon, Harrah's, Capital One y Netflix, o de Gartner, IBM y McKinsey & Co., quienes ofrecen sus servicios de consultoría con servicios para ayudar a las empresas a capitalizar el Big Data y generar valor, ventaja competitiva e incluso mejorar su desempeño (Wamba et al., 2018).

Las aplicaciones del Big Data varían dependiendo la industria y la empresa. En la industria manufacturera se utiliza para mantenimiento predictivo, mejorar la eficiencia operacional y la optimización de la producción; en el comercio al por menor se puede aplicar para el desarrollo de productos, mejorar la experiencia del cliente y crear lealtad de los mismos, estrategia de precios y optimización; en la industria de la salud el Big Data permite realizar investigación genómica, mejorar la experiencia del paciente y resultados de tratamientos, fraude en reclamos de seguros médicos o analíticos sobre facturación; o en la industria de servicios financieros, el Big Data es muy útil en temas de seguridad como identificar fraudes e implementar políticas anti lavado de dinero (Oracle, 2018).

Por ejemplo, Amazon utiliza el Big Data para analizar la información recopilada de sus clientes no solamente para publicidad, sino que también la utiliza para mejorar el servicio al cliente; American Express, utiliza Big Data al analizar transacciones históricas para predecir el comportamiento de sus clientes y fortalecer la lealtad de los mismos; Capital One se enfoca en utilizar el Big Data para garantizar el éxito de la oferta de productos a sus clientes al analizar los hábitos de consumo y demográficos permitiéndoles crear estrategias de mercadotecnia dirigidas y relevantes; General Electric utiliza datos de sensores en maquinaria como turbinas de gas y de jets para identificar formas de mejorar los procesos y confiabilidad de las turbinas, así como su eficiencia; reportes de Big Data, análisis, experimentación y *Machine Learning* permiten a Miniclip monitorear y mejorar la experiencia de los usuarios de sus juegos digitales para retenerlos y apoyar el crecimiento de su negocio (O'Neill, 2022).

En el caso de Netflix, utilizan Big Data para analizar los datos sobre hábitos de visualización de sus consumidores y comisionar contenido original de acuerdo a sabiendas de que tendrá buena aceptación en ciertas regiones; Next Big Sound utiliza el Big Data para combinar datos de distintas plataformas de streaming y de redes sociales para predecir el próximo éxito en la industria de la música; mientras que Starbucks, también utiliza el Big Data

con fuentes de información externas, como tráfico, área demográfica, comportamiento de los clientes, para determinar el éxito potencial de una nueva locación (O'Neill, 2022).

Lo anterior no significa que empresas de menor tamaño (PyMEs) no puedan beneficiarse del uso de Big Data, puesto que si removemos la idea del volumen o cantidad de datos que implica la definición de Big Data, los beneficios y aplicaciones descritos anteriormente se pueden aplicar a PyMEs, esto gracias al declive en los costos de las tecnologías y herramientas de innovación necesarias, las PyMEs pueden aprovechar los datos generados al interior de la empresa, o bien utilizar fuentes de datos externas, para obtener información valiosa (Çelebi, 2020).

Sen et al., (2016) sugieren que utilizar Big Data para generar conocimiento que permita crear una ventaja competitiva aplica en la misma medida para grandes empresas que para PyMEs, siempre que se eliminen las barreras del costo de las herramientas. Incluso, aseguran que en algunos casos las PyMEs pueden crear innovación y permear una estrategia de datos en su organización de una manera más rápida, puesto que tienen una estructura de TI más flexible, menos problemas heredados y una naturaleza de rápida adaptación a los cambios debido a su tamaño, implementando el Big Data para identificar correlaciones, riesgos, oportunidades, así como mantenimiento predictivo, predicción de la demanda, optimización de procesos, planeación predictiva de inventarios, segmentación de mercados, entre otras aplicaciones (Sen, et al., 2016).

De igual forma, Iqbal, et al., (2018) reconocen el valor que el Big Data puede traer a las empresas al ayudarlas a crecer incorporando la información y el conocimiento extraído de los datos al proceso de toma de decisiones, independientemente de su tamaño, pero identificando algunos retos en la adopción de Big Data para las PyMEs, como la conciencia sobre la utilidad y definición de Big Data, falta de talento especializado en el área, acceso a herramientas, entre otros, los cuales se han ido mitigando con la competitividad de los costos de la implementación de Big Data.

Es precisamente el conocimiento que se extrae de los datos, independientemente de su tamaño, lo que constituye un analítico, pues a pesar del potencial del Big Data para transformar los negocios al permitir comprender y enfrentar los retos y oportunidades del entorno en el que

operan (Singh y El-Kassar, 2019), la habilidad para analizar los datos, extraer conocimiento y efectivamente extraer valor de ellos, representa uno de los principales retos para las empresas (Niebel et al., 2019). Esta creciente necesidad de las firmas por tomar mejores decisiones, más informadas y más rápidas ha llevado a complementar los recursos tecnológicos y de datos del Big Data con otros recursos clave de la empresa para construir ABD que le permitan crear una ventaja competitiva (Ghasemaghaei et al., 2017).

En ese sentido los ABD se pueden definir cómo una forma de analíticos avanzados que implica todos los procesos y tecnologías para examinar los datos, o Big Data, a efectos de descubrir información (por ejemplo patrones, correlaciones, tendencias, observaciones atípicas o preferencias de los consumidores) que pueda resultar útiles para la toma de decisiones de negocio de las firmas, ya sea sobre su operación, desempeño o su entorno (Chai et al., 2021).

Relación entre ABD y el desempeño de las firmas: evidencia empírica

Como se ha revisado, a pesar de las diferentes perspectivas utilizadas para definir los ABD por los diferentes autores, la mayoría coincide en que, independientemente del tipo de analíticos, los procesos y tecnologías para examinar los datos o las capacidades de la empresa, para que los ABD puedan generar una ventaja competitiva, es necesario su integración en el proceso de toma de decisiones de la firma (Bogdan y Borza, 2019).

En éste apartado se revisan algunos de los principales artículos científicos publicados que analizan el impacto del uso de datos para la toma de decisiones, ya sea en el desempeño de la firma, el proceso de toma de decisiones o, de forma más general en la creación de una ventaja competitiva para las empresas que hacen uso. La revisión se enfoca principalmente en revisar las teorías que soportan el marco de la investigación, su metodología, variables a utilizar e hipótesis, así como los principales hallazgos y conclusiones de los autores.

Cao et al. (2015), encontraron que los analíticos de datos, a través de un efecto mediador dado por un entorno basado en datos, tiene un impacto positivo en la capacidad de procesamiento de información, que a su vez tiene un efecto positivo en la efectividad del proceso de toma de decisiones. El análisis está basado en la visión del procesamiento de la información, la cual sugiere que la tarea clave de las organizaciones es manejar la incertidumbre (como tareas

complejas) y el cambio del entorno en el que operan a través del despliegue de mecanismos para el procesamiento de la información, y la teoría de la contingencia, la cual define la capacidad de adaptarse de una firma como el grado en el que las necesidades, demandas, metas, objetivos y/o estructuras de un componente son consistentes con las necesidades, demandas, metas, objetivos y/o estructuras de otro componente.

La metodología de Cao et al. (2015) consistió en desarrollar los constructos formativos basándose en la literatura existente de analíticos de datos y de TI utilizando un Modelo de Ecuaciones Estructurales con Mínimos Cuadrados Parciales (*Partial Least Squares Structural Equation Model* o PLS-SEM por sus siglas en inglés), componiendo el modelo los siguientes constructos: analíticos de datos, entorno basado en datos, capacidad de procesamiento de datos, toma de decisiones basada en datos y efectividad del proceso de toma de decisiones. Cada uno de los constructos se compone a su vez de indicadores o ítems con respuestas que utilizan una escala de cinco puntos de Likert¹ y que conformaron el instrumento de levantamiento de datos, que fue aplicado de forma electrónica a gerentes de empresas encontrados en la base de datos FAME (*Forecasting Analysis and Modeling Environment* por sus siglas en inglés) de Reino Unido. Los hallazgos proporcionan información sobre cómo se pueden utilizar los analíticos de datos en los negocios para facilitar la toma de decisiones basada en datos y cómo se deben implementar para hacer más efectivo el proceso de toma de decisiones.

Por su parte, Chen et al. (2015) se basan en la teoría de las capacidades dinámicas para conceptualizar las capacidades de ABD como una capacidad única de procesamiento de datos que permite a la organización crear una ventaja competitiva y, por lo tanto crear valor, así como el marco teórico tecnología-organización-ambiente (TOE por sus siglas en inglés) para identificar los factores que influyen el uso de analíticos de Big Data. El análisis utiliza la metodología PLS-SEM para evaluar el modelo de investigación propuesto, encontrando que el uso de ABD en la organización afecta la creación de valor en la misma y que está moderado por el dinamismo del ambiente, que los factores tecnológicos tienen una influencia directa en el uso organizacional de dichos analíticos, y que los factores organizacionales y tecnológicos en

¹ La escala autoreportada de Likert es uno de los métodos de medición más utilizados en las ciencias sociales pues provee una forma conveniente de medir constructos no observables, que va desde definir el constructo y las preguntas o ítems hasta la creación de la escala (Jebb et al., 2021).

conjunto afectan de forma indirecta el uso de ABD cuando existe apoyo de los tomadores de decisiones en la organización.

Específicamente, Chen et al. (2015) proponen que el uso de ABD se compone de dimensiones tecnológicas como los beneficios esperados y las capacidades tecnológicas, pero que también depende del soporte de la alta dirección de la empresa, el cual se deriva de la preparación organizacional y de la presión competitiva. A su vez, el uso de ABD tiene un impacto positivo en la productividad de los activos y en el crecimiento del negocio, siendo en ambos casos el dinamismo del ambiente una variable mediadora. Los datos para probar el modelo fueron recolectados utilizando un cuestionario aplicado de forma digital a profesionales en el área de cadena de suministros de empresas en todo el mundo. Los resultados, como ya se mencionó, muestran que el uso de ABD a nivel organizacional impacta significativamente de forma positiva la creación de valor en la cadena de suministros de dos formas: productividad de los activos y crecimiento del negocio; dichos efectos son moderados por el dinamismo del entorno; los factores tecnológicos tienen un impacto directo en el uso de ABD a nivel organizacional; y factores organizacionales y del entorno tienen una influencia indirecta en el uso de ABD a nivel organizacional a través del soporte de la alta dirección en la firma (Chen et al., 2015).

En la literatura revisada también se encontraron enfoques que van más allá de las capacidades de la empresa y analizan la creación de valor y, por lo tanto, el desempeño de la firma en un ambiente de Big Data. Por ejemplo, Ren et al. (2016), utilizaron la teoría RVB y literatura sobre el éxito de los sistemas de información para ligar la calidad de los sistemas, la calidad de la información con la creación de valor y el desempeño de la firma. Partiendo de Barney (2014) caracterizan a la competencia de ABD de la firma y argumentan que la calidad de los sistemas de información en un entorno de ABD tienen un impacto significativo en el valor de negocio, el cual a su vez influencia el desempeño de la firma. De igual forma, utilizan la teoría del éxito de los SI de DeLone y McLean (1992) para caracterizar a la calidad de los sistemas, puesto que se ha identificado que tanto la calidad de los sistemas como de la información son factores importantes para promover el valor y desempeño derivado de las TI.

Partiendo de la metodología PLS-SEM y de las teorías mencionadas, Ren et al. (2016) caracterizan la calidad de los sistemas como un constructo de segundo orden compuesto de seis dimensiones: confiabilidad, adaptabilidad, integración, accesibilidad, tiempo de respuesta y privacidad; la calidad de la información se define con cuatro dimensiones: completitud, divisa, formato y precisión; el valor de negocios que crean los ABD se divide en transaccional, estratégico o transformacional; mientras que el desempeño de la firma se mide en desempeño financiero y de mercado.

Al igual que otros trabajos, los datos para probar el modelo propuesto por Ren et al. (2016) fueron recolectados a través de un cuestionario con preguntas o ítems múltiples para la medición de cada uno de los constructos o dimensiones, utilizando una escala de siete puntos de Likert para las respuestas. Dicho instrumento fue enviado de forma digital a gerentes de TI y analistas de negocio de empresas Chinas y aplicando la metodología PLS-SEM a las respuestas obtenidas, se encontró que la calidad de los sistemas utilizados para ABD tienen una mayor influencia en el desempeño financiero de la empresa que la calidad de los datos, a pesar de que ambos influyen de forma positiva. Estos resultados implican que las empresas que ya utilizan ABD en su proceso de toma de decisiones, pueden mejorar la calidad de los sistemas y la información empleados para generar los analíticos de forma constante, permitiéndoles mejorar el valor de su negocio y, a su vez, incrementar el desempeño de la empresa (Ren et al., 2016).

Al igual que Ren et al. (2016), otros estudios utilizan como fundamento la teoría basada en los recursos o RBV para el desarrollo de las capacidades de ABD en una firma. Akter et al. (2016), complementan la teoría basada en los recursos con la visión de enredo del socio materialismo para construir un modelo que analice si efectivamente las capacidades de ABD mejoran el desempeño de la firma. Asimismo, incorpora al alineamiento de las capacidades de ABD con la estrategia de negocios de la firma como una variable mediadora entre las capacidades de ABD y el desempeño de la empresa.

Para la recolección de los datos, Akter et al. (2016) iniciaron con la revisión de la literatura pertinente para identificar las dimensiones de las capacidades de ABD, para posteriormente realizar dos estudios Delphi² para explorar las subdimensiones de las dimensiones primarias identificadas en la revisión de la literatura. Se identificaron 11 subdimensiones, las cuáles se utilizaron posteriormente para diseñar un cuestionario, en el que todas las preguntas se miden utilizando una escala de siete puntos de Likert para aplicarlos a profesionales en las áreas de analíticos de negocios, de Big Data o de TI en EUA.

Utilizando la metodología PLS con Análisis Factorial Confirmatorio, Akter et al. (2016) encontraron que las capacidades de ABD, compuestas por la capacidad de administración, de tecnología y de talento, tienen un impacto positivo en el desempeño de la firma. Tanto el modelo como los hallazgos, contribuyen a la literatura al utilizar la teoría basada en los recursos y la visión de enredo del socio materialismo para proponer un modelo integral que esboza las capacidades de ABD como una integración, no solamente de datos de calidad, sino de la capacidad de habilidades administrativas, capacidad de tecnología y capacidad del talento, y que juntas tienen un impacto positivo en el desempeño de la firma (Akter et al., 2016).

Ahondando en la misma dirección, Gupta y George (2016) parte de que la inversión para construir capacidades de ABD por sí misma no es suficiente para crear una ventaja competitiva, por lo que se apoya en la teoría RBV para validar de forma empírica que el impacto de las capacidades de ABD es efectivamente positivo en el desempeño de la firma. Específicamente, argumenta que la capacidad de ABD en una firma se compone de tres dimensiones: recursos tangibles, habilidades humanas y recursos intangibles, todo esto teniendo como base la literatura de capacidades de TI, la cuál a su vez parte de la teoría basada en los recursos, adicionando a la literatura que define las capacidades de ABD como una capacidad no solamente técnica.

El modelo propuesto por Gupta y George (2016) caracteriza a las capacidades de Big Data como un constructo de tercer orden, que se compone a su vez de tres dimensiones o

² De acuerdo con Okoli y Pawlowski (2004), a pesar de las diferentes variaciones en el método Delphi, la definición que captura las características comunes es la de Linstone y Turoff: “Delphi se puede caracterizar como un método para estructurar un proceso de comunicación grupal de modo que el proceso sea efectivo para permitir que un grupo de individuos, como un todo, se ocupe de un problema complejo. Para lograr esta “comunicación estructurada” se proporciona: alguna retroalimentación de las contribuciones individuales de información y conocimiento; alguna evaluación del juicio o punto de vista del grupo; alguna oportunidad para que las personas revisen sus puntos de vista; y cierto grado de anonimato para las respuestas individuales”.

constructos de segundo orden (recursos tangibles, habilidades humanas y recurso intangibles), los cuales a su vez se derivan de constructos de primer orden. En una primera versión del modelo, dichas capacidades de Big Data se utilizan como variable dependiente junto con el desempeño de mercado de la firma, mientras que en la segunda versión, el desempeño operacional de la firma constituye la única variable dependiente. Al igual que la mayoría de la literatura revisada, los datos se recolectan a través de un cuestionario con diferentes ítems o preguntas con una escala de Likert, los que a su vez conforman los constructos de primer orden y se utiliza la metodología PLS-SEM para identificar los principales aspectos que influyen en la formación de una capacidad de ABD en una firma, validando la relación entre las capacidades de ABD y el desempeño de las empresas (Gupta y George, 2016).

Por su parte, Côte-Real et al. (2017), se apoyan en la teoría de la visión del conocimiento y de las capacidades dinámicas para proponer un modelo conceptual y probar de forma empírica que los ABD pueden generar valor en varias etapas de la cadena productiva de la firma y, por lo tanto, generar una ventaja competitiva, puesto que utilizarlos para generar conocimientos y agilizar el proceso de toma de decisiones en la firma puede ayudar a construir una capacidad dinámica. Asimismo, se prueba que la agilidad actúa como mediador parcial entre los activos de conocimiento y el desempeño de la firma.

La evidencia empírica del trabajo de Côte-Real et al. (2017) proviene de información recolectada de un cuestionario en línea aplicado a directivos de empresas europeas y para su análisis se utilizó la metodología PLS-SEM, encontrando que la inversión en ABD puede ser estratégica para dichas firmas, puesto que puede aumentar su agilidad organizacional ayudándoles a sobrevivir en mercados competitivos, siempre que la agilidad se derive del conocimiento emanado de los ABD y su implementación en los procesos de la firma, puesto que ayuda a crear una ventaja competitiva que impacta en el desempeño financiero y estratégico de la firma.

Siguiendo con el mismo enfoque, Ghasemaghaei et al. (2017) buscan identificar las condiciones bajo las cuales las capacidades de TI pueden llevar a incrementar la agilidad en la firma. Específicamente, la capacidad del uso de analíticos de datos para la toma de decisiones dentro de la firma y cómo se utiliza para mejorar la agilidad y conducir a la creación de una capacidad dinámica. Al igual que otros estudios, se basa en la teoría de las capacidades

dinámicas que a su vez se deriva de la teoría visión basada en los recursos, las cuales han sido los cimientos para probar que los recursos de TI conducen a la creación de una ventaja competitiva en las firmas. En ese sentido, se toma a los ABD y a la agilidad generada en la firma por su uso en el proceso de toma de decisiones como un recurso de TI que conduce a la generación de una ventaja competitiva a través de un modelo que incorpora también la teoría de contingencia y el concepto de ajuste, en el cual sugiere que los resultados de la organización (e.g. agilidad, competitividad) son consecuencia del ajuste de dos o más factores (e.g. habilidades de los empleados, recursos de TI).

La hipótesis general de Ghasemaghaei et al. (2017) sugiere que el uso de analíticos de datos incrementa la agilidad de la firma, teniendo como mediadores los factores de ajuste personas-herramientas, datos-herramientas y tareas-herramientas y se pone a prueba utilizando la metodología PLS-SEM encontrando que la efectividad del uso de herramientas de analíticos de datos requiere de ajustes continuos en otros recursos de la organización como las personas, los datos y las tareas para las que se utilizan los datos, solamente cuando se utilizan analíticos de datos en combinación con estos recursos es que se puede realmente incrementar la agilidad de la firma y, por lo tanto, generar una ventaja competitiva.

La mayoría de la literatura reciente se enfoca en identificar la relación entre los ABD y el desempeño de la firma, teniendo como base la teoría visión basada en los recursos, el éxito de los recursos de sistemas de información y el valor de las TI, conceptualizando a las capacidades de ABD como una capacidad dinámica de la firma. Wamba et al. (2017) proponen justamente un modelo de capacidades de ABD dónde examinan los efectos directos de dichas capacidades en el desempeño de la firma, así como los efectos mediadores de las capacidades dinámicas orientadas a procesos en la relación de las capacidades de ABD con el desempeño financiero y de mercado de la firma.

Wamba et al. (2017) se basan en la teoría visión basada en los recursos, así como en la teoría del socio materialismo para caracterizar a las capacidades de ABD como un constructo de tercer orden compuesto a su vez por tres constructos de segundo orden: la capacidad de infraestructura de ABD, la capacidad de administración y la capacidad del personal. El modelo propone un efecto positivo de las capacidades de ABD en el desempeño de la firma (financiero y de mercado) y un efecto positivo en las capacidades dinámicas orientadas a procesos, con un

efecto mediador positivo indirecto de estas últimas en el desempeño de la firma. Para probar el modelo se utilizó la metodología PLS-SEM basada en información obtenida a través de cuestionarios con preguntas de opción múltiple utilizando una escala de Likert de siete puntos, que fueron aplicados a gerentes de TI y analistas de negocios de empresas Chinas. Los resultados confirman los efectos directos e indirectos de las capacidades de ABD en el desempeño financiero y de mercado de las empresas, así como el fuerte papel como mediador que tienen las capacidades dinámicas orientadas a procesos (Wamba et al., 2017).

Mientras que el trabajo de Mandal (2018) se basa en el modelo propuesto por Wamba et al. (2017) para investigar el impacto de las capacidades de gestión de análisis de Big Data en la preparación, el estado de alerta y la agilidad de la cadena de suministro. Al igual que Wamba et al. (2017), el marco teórico de Mandal (2018) reside en la teoría de las capacidades dinámicas, una extensión de la teoría RVB, bajo el cual proponen que la planeación, toma de decisiones respecto a la inversión, coordinación y control de los ABD, o capacidades de gestión de los ABD, influyen de forma positiva en la preparación, estado de alerta y agilidad de la cadena de suministro de la firma.

Con base en literatura existente y utilizando la metodología PLS en vez de PLS-SEM, Mandal (2018) define las capacidades de gestión de los ABD como cuatro constructos de primer orden (planeación, toma de decisiones respecto a la inversión, coordinación y control), al igual que la preparación, estado de alerta y agilidad de la cadena de suministro. Cada uno de los constructos se mide con preguntas o ítems que utilizan una escala de Likert de siete puntos y que se plasmaron en un cuestionario digital administrado a profesionales de TI con mínimo dos años de experiencia en el campo de analíticos relacionados con la gestión de cadena de suministros, obteniendo un total de 249 respuestas. Para probar las hipótesis, como ya se mencionó, se utilizó la metodología PLS en vez de PLS-SEM, puesto que el objetivo del estudio es revisar la validez de las trayectorias propuestas a través de los coeficientes, así como un enfoque basado en componentes (Mandal, 2018).

Teniendo en cuenta lo anterior, los resultados de Mandal (2018) sugieren que la planeación de ABD influye positivamente la preparación, estado de alerta y agilidad de la cadena de suministro, la coordinación de ABD influye positivamente en la preparación, estado de alerta y agilidad de la cadena de suministro, el control de ABD influye positivamente la

preparación, estado de alerta y agilidad de la cadena de suministro, sin encontrar evidencia de que la toma de decisiones respecto a la inversión en ABD influya de forma positiva en la preparación, estado de alerta y agilidad de la cadena de suministro. En otras palabras, el estudio contribuye a la teoría de las capacidades dinámicas enfocándose en la cadena de suministro, encontrando que la planeación, coordinación y control, o capacidades de gestión, de los ABD son habilitadores prominentes en la preparación y agilidad de la cadena de suministro y son críticos para el desarrollo de agilidad en la misma.

Como hemos revisado, gran parte de los artículos sobre ABD, se centran en su caracterización como una capacidad de la firma que tiene un impacto positivo en el desempeño de la misma. Sin embargo, Dubey et al. (2018) toman ésta caracterización para conceptualizar un modelo enmarcado en la teoría visión basada en los recursos contingentes, la cual se deriva a su vez de la teoría RBV, y la aplican a probar si las capacidades de ABD y predictivos tienen una influencia significativa en la visibilidad y coordinación de las cadenas de suministro humanitarias en Organizaciones No Gubernamentales (ONGs).

El argumento central de Dubey et al. (2018) es que las ONGs, al igual que las firmas y dadas condiciones específicas, pueden crear ventajas competitivas a través de la agrupación de recursos estratégicos que son valiosos, raros, inimitables y no sustituibles. Específicamente, la capacidad de ABD y predictivos se considera como una capacidad organizacional que mejora la visibilidad de la ONG, una capacidad deseada en las cadenas de suministro humanitarias pues apalanca el riesgo de pobre coordinación debido a la información asimétrica, así como el rol mediador de la confianza rápida o “*swift trust*” entre las capacidades de ABD y predictivos y la visibilidad y coordinación entre los diferentes actores de las cadenas de suministro humanitarias. En el modelo, el constructo de capacidades de ABD y predictivos son un constructo reflectivo de segundo orden y se conforma por siete constructos reflectivos de primer orden: recursos básicos, datos, tecnología, habilidades tecnológicas, habilidades gerenciales, cultura organizacional y aprendizaje organizacional; mientras que las capacidades de ABD y predictivos llevan a los constructos reflectivos de tercer orden visibilidad y coordinación, de acuerdo con la lógica de la visión basada en los recursos (Dubey et al., 2018).

Para probar el modelo, Dubey et al. (2018) utilizaron un instrumento de recolección de datos con ítems o preguntas medidas con una escala Likert de cinco puntos que componían cada

uno de los constructos y fue aplicado a informantes clave que tenían posiciones gerenciales en ONG internacionales o que habían colaborado con una ONG internacional en el último año. Los resultados demuestran que, efectivamente hay un efecto positivo de la capacidad de ABD y predictivos, pues contribuyen a mejorar la confianza rápida, la visibilidad y la coordinación en las cadenas de suministro humanitarias.

En otro estudio, Ghasemaghaei et al. (2018), se concentran en el potencial de la capacidad de ABD para mejorar el proceso de toma de decisiones en las empresas. Para ello utilizan como marco teórico la teoría de la visión basada en los recursos y cómo las empresas construyen una competencia a través de su habilidad para integrar, ensamblar y desplegar recursos valiosos. En ese sentido, la competencia de analíticos de datos es concebida y operacionalizada como un índice formativo multidimensional que se compone de los constructos de primer orden: calidad de los datos, grandeza de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de las herramientas, cumpliendo con el marco de la RBV y el de los recursos claves basados en TI; mientras que el desempeño del proceso de toma de decisiones se operacionaliza a través de dos constructos relacionados a la calidad y la eficiencia de las decisiones (Ghasemaghaei et al., 2018).

El modelo propuesto por Ghasemaghaei et al. (2018) sugiere que cuanto mayor sea el nivel de competencia de análisis de datos de la organización, mayor será el rendimiento de la toma de decisiones de la empresa, y la metodología para probarlo es PLS-SEM, lo cuál también es consistente con otros trabajos en la literatura existente sobre capacidades de ABD. Mientras que la información para probar el modelo se recopiló utilizando un cuestionario aplicado a gerentes de TI y analistas de datos de organizaciones en EUA, el cual se compone de diferentes ítems o preguntas que conforman los constructos de primer orden (e.g. calidad de los datos, habilidad técnicas) medidos con una escala de Likert de siete puntos (Ghasemaghaei et al., 2018).

Los resultados de Ghasemaghaei et al. (2018) sugieren que todas las dimensiones que conforman la capacidad de analíticos de datos en la firma mejoran de forma significativa la calidad de las decisiones y que todas, excepto la grandeza de los datos, también contribuyen a incrementar la eficiencia de las decisiones en una firma, validando así la conceptualización y operacionalización de las capacidades de ABD en la firma como un constructo de segundo orden

y cada uno de sus componentes en términos de probar el impacto positivo que tiene en el desempeño del proceso de toma de decisiones de la firma.

Müller et al. (2018), van un paso más allá al trabajar en un modelo para identificar los efectos de los ABD en el desempeño de la firma basándose en medidas objetivas de los activos de ABD, en vez de caracterizarlos como una capacidad. Para ello, utilizan el enfoque “automatizar/informar”, el cuál define de forma separada los recursos de TI que se utilizan para automatizar las operaciones de la firma al incrementar la continuidad y control de los procesos de trabajo, de los recursos que se utilizan para informar a los tomadores de decisiones al crear información que mejora la comprensión de los procesos de trabajo de la firma (Müller et al., 2018).

En ese sentido, Müller et al. (2018) analizan el impacto de la inversión en activos de TI orientados a informar como variable independiente y el desempeño o productividad de la firma como la variable dependiente, utilizando un modelo de Mínimos Cuadrados Ordinarios con datos de panel de 814 firmas de todo el mundo. A través de vendedores de software, obtuvieron una base de datos longitudinal sobre los activos de ABD de sus clientes, los cuales clasificaron en tres categorías: tecnología fundamental de bases de datos, soluciones de minería de datos y *Machine Learning*, y herramientas de visualización y presentación de datos (Müller et al., 2018). Por otra parte, fusionaron dichos datos con información financiera para empresas públicas que cotizan la bolsa de valores de EUA, utilizando la base de datos *Compustat Global Fundamentals Annual Database*, para llegar a un panel de datos balanceado de 814 firmas para el periodo 2008-2014 con un total de 5,698 observaciones (Müller et al., 2018).

Los resultados de Müller et al. (2018) muestran que los activos de ABD se asocian con una mejora promedio del 3 al 7% en la productividad de la firma. Sin embargo, todavía se puede observar una diferencia sustancial en los retornos de los activos de ABD cuando se considera la industria en la que opera la firma, siendo las firmas que operan en industrias intensivas en el uso de TI o en industrias altamente competitivas las que pueden extraer más claramente valor de los activos de ABD, sin que se pudiera detectar una mejora medible de la productividad para empresas fuera de esos grupos (Müller et al., 2018).

Regresando a la teoría visión basada en los recursos, pero manteniendo la inversión en ABD como variable independiente, Raguseo y Vitari (2018) demuestran que el valor comercial logrado a partir de las inversiones en ABD conduce a ventajas en términos de rendimiento financiero de una firma. Utilizando la teoría visión basada en los recursos, caracterizan los activos de información inimitables y no sustitutos, es decir, la información extraída utilizando TI para construir ABD es la creación de valor, mientras que el desempeño financiero de la firma, lo operacionalizan en términos de retorno de la inversión en recursos de ABD (e.g. recursos de TI, recursos humanos especializados en ABD).

Teniendo en cuenta lo anterior, Raguseo y Vitari (2018) proponen que el valor de negocio de las soluciones de ABD tienen un efecto positivo en el desempeño financiero de una firma, mientras que el desempeño de mercado y la satisfacción de los clientes tienen un efecto mediador entre ambos, dando una visión más global al desempeño de la firma, pues no solamente se concentran en el desempeño financiero, sino que también incluyen al mercado y a los clientes en el modelo. La metodología utilizada es SEM basado en varianza y caracterizan al valor de negocio de las soluciones de ABD como un constructo de segundo orden conformado por cuatro dimensiones (valor de negocio transaccional, estratégico, transformacional e informacional), mientras que el desempeño financiero se mide con el incremento de las ventas, la retención de clientes y la rentabilidad (Raguseo y Vitari, 2018).

Los datos utilizados en el modelo de Raguseo y Vitari (2018) provienen de la administración de un cuestionario, con preguntas medidas con una escala de siete puntos de Likert, a Directores de Información de empresas francesas, contactados a partir de información de la base de datos *Bureau Van Dijk*, recolectando 76 cuestionarios válidos para probar el modelo. Finalmente, encontraron que los resultados de la aplicación del modelo muestran que el valor de negocio de las soluciones de ABD son capaces de explicar el 62.4% de las variaciones en la satisfacción de los clientes, 71.9% de la variación en el desempeño de mercado y 78.6% de la variación del desempeño financiero de la firma. Sin embargo, solamente el valor de negocio de las soluciones de ABD y la satisfacción de los clientes tienen un efecto en el desempeño financiero, directo e indirecto respectivamente, dejando fuera el efecto mediador del desempeño de mercado. De forma general, Raguseo y Vitari (2018) concluyen que las firmas deberían orientar sus esfuerzos a diseñar soluciones de ABD holísticas para que generen valor

transaccional, transformacional, estratégico e informacional, lo que eventualmente se traduce en una mejora en el desempeño financiero de la firma.

Por su parte, Jeble et al. (2018) dan un paso más allá al concentrarse en un tipo específico de ABD, los analíticos predictivos, y su impacto en la sustentabilidad de la cadena de suministros de la firma. Para ellos toman como base la teoría RBV, la cual, como ya revisamos, argumenta que una firma puede derivar su ventaja competitiva de los recursos y capacidades que posee y que pueden ser valiosos, raros, imperfectamente imitables y no sustituibles (recursos de capital físico, recursos de capital humano y recursos de capital organizacional). Sin embargo, Jeble et al. (2018) profundizan hasta llegar a la teoría visión basada en los recursos contingente, pues consideran que la RBV tiene una naturaleza algo estática, mientras que la RBV contingente sostiene que la habilidad de la firma para que sus recursos y capacidades generen una ventaja competitiva depende de variables internas y externas.

Asimismo, de las definiciones de Big Data, Jeble et al. (2018) se centran en la que clasifica a los analíticos derivados del Big Data en descriptivos, predictivos y prescriptivos, tomando a los analíticos predictivos como objeto de estudio central y enmarcándolos en la RVB contingente como la capacidad de analíticos de Big Data y predictivos de la firma. Mientras que por otro lado, se centran en el impacto que tiene dicha capacidad en la sustentabilidad de la cadena de suministro de la firma, a través del desempeño ambiental, social y económico de la firma, así como el efecto mediador de la complejidad de la base de suministro (Jeble et al., 2018).

Específicamente, Jeble et al. (2018) utilizan la metodología de PLS-SEM para operacionalizar la capacidad de analíticos de Big Data y predictivos como un constructo de segundo orden que se compone de los recursos tangibles, habilidades técnicas, habilidades administrativas, aprendizaje organizacional y una cultura basada en datos. Para la recolección de los datos se diseñó un cuestionario como instrumento de recolección de datos, en el que se operacionalizan los constructos con base en literatura ya existente y modificaciones pertinentes de acuerdo a la definición de dichos constructos, utilizando ítems o preguntas medidos con una escala de Likert de cinco puntos (Jeble et al., 2018). Dicho instrumento se aplicó vía correo electrónico a gerentes senior de empresas manufactureras de la industria automotriz en India, utilizando la base de datos *Confederation of Indian Industry*. Los resultados de Jeble et al. (2018)

revelan que la capacidad de analíticos de Big Data y predictivos tiene un impacto positivo en el desempeño ambiental, en el desempeño social y en el desempeño económico de la firma, sin embargo, no se encontró evidencia del efecto mediador en dicho impacto de la complejidad de la base de suministro.

Otro estudio que aborda el impacto de las capacidades de ABD en el desempeño de la firma desde el enfoque de la teoría RBV es el de Anwar et al., (2018). En él profundizan en una corriente más reciente derivada de la teoría RBV, la teoría de las capacidades dinámicas, al igual que Chen et al. (2015), Côte-Real et al. (2017), Ghasemaghahi et al. (2017) y Wamba et al. (2017). Asimismo, Anwar et al., (2018) incorporan el concepto de ventaja competitiva como una de las variables dependientes en el modelo y dividen las capacidades de ABD en dos categorías: capacidades técnicas de ABD y capacidades personales de ABD.

De forma específica, el modelo operacionaliza las capacidades tecnológicas de ABD como un constructo formado por la compatibilidad, conectividad y la modularidad, las capacidades personales de ABD es también un constructo de segundo orden conformado por las dimensiones conocimiento técnico, administración del conocimiento tecnológico y conocimiento del negocio, mientras que el desempeño de la firma se conforma del desempeño financiero (retorno en activos, retorno en inversión y retorno en activos propios) y el no financiero (satisfacción de los clientes, satisfacción de los empleados y desempeño general) (Anwar et al., 2018). Anwar et al., (2018) proponen que hay una relación significativamente positiva entre las capacidades técnicas de ABD y el desempeño y ventaja competitiva de la firma, al igual que una relación significativamente positiva entre las capacidades personales de ABD y el desempeño y ventaja competitiva de la firma, así como una relación significativamente positiva entre la ventaja competitiva y el desempeño de la firma, mientras que la ventaja competitiva juega un rol mediador entre las capacidades técnicas y personales de ABD y el desempeño de la firma.

Al igual que otros trabajos, la recolección de los datos para el modelo empírico de Anwar et al., (2018) se llevó a cabo a través de un cuestionario con los ítems que conforman cada una de las dimensiones de los constructos y medidos con una escala de Likert, los cuales fueron aplicados a ejecutivos, dueños y profesionales de TI en empresas Chinas. La metodología es SME con énfasis en el Análisis Factorial Confirmatorio y los resultados concluyen que la relación positiva entre las capacidades técnicas y personales de ABD con el desempeño y ventaja

competitiva de la firma se confirma para el caso de estudio, probando que las capacidades técnicas se complementan con las capacidades del personal que utiliza los recursos tecnológicos para extraer información valiosa de los datos, para después ponerla en acción y mejorar el desempeño de la firma, contribuyendo a la creación de una ventaja competitiva (Anwar et al., 2018).

Por su parte, Wamba et al. (2018) profundizan en la definición de Big Data y se enfocan en una de sus dimensiones: la calidad de la información de los ABD y su relación con el valor de negocio, la satisfacción de los usuarios y el desempeño de la firma. Particularmente, proponen que la percepción de calidad de la información tiene un impacto significativamente positivo en la percepción de valor de negocio en los ABD y en la percepción de satisfacción de los usuarios de ABD, mientras que a su vez el valor de negocio percibido tiene un impacto significativamente positivo en la satisfacción percibida de los usuarios de ABD y que ambas (valor de negocio y satisfacción percibida de los usuarios de ABD) tienen un impacto significativamente positivo en el desempeño de la firma (Wamba et al., 2018).

Basándose en el marco de evaluación-respuesta emocional-afrontamiento, Wamba et al. (2018) proponen una teoría sobre cómo la dinámica de la calidad de la información ayuda a lograr el valor empresarial, la satisfacción del usuario y el desempeño de la empresa con la estrategia y la implementación de Big Data. Mientras que para probar el modelo utilizan la metodología PLS-SEM, definiendo el constructo calidad de la información como un constructo de segundo orden conformado por la integridad, precisión, formato y actualidad de la información producida por los ABD; el constructo valor de negocio se define como la el valor transaccional, estratégico y transformacional de los ABD; la satisfacción de los usuarios se refiere a cómo los usuarios se sienten respecto al uso de ABD; y el constructo desempeño de la firma hace referencia a la habilidad de la empresa para ganar y retener clientes, incrementar sus ventas, rentabilidad y el retorno de la inversión (Wamba et al., 2018).

La evidencia empírica para el trabajo de Wamba et al. (2018) fue recolectada mediante un cuestionario basado en literatura existente sobre cada uno de las dimensiones de los constructos, adaptándolos al contexto de analíticos de negocio y utilizando una escala de siete puntos de Likert para medir cada uno de los ítems o preguntas que conforman cada dimensión o constructo de primer orden. Dicho instrumento de recolección fue administrado a profesionales en las áreas

de TI, analíticos y análisis de negocio de empresas en EUA y Francia, obteniendo un total de 302 cuestionarios válidos, con un total de once constructos o variables latentes. Los resultados del modelo estructural muestran que la calidad de la información tiene un impacto significativamente positivo en el valor de negocio percibido, la calidad de la información y el valor de negocio tienen un efecto moderado y significativo en la satisfacción percibida de los usuarios de ABD, mientras que la satisfacción percibida de los usuarios y el valor de negocio percibido tienen un impacto significativamente positivo en el desempeño de la firma (Wamba et al., 2018).

Uno de los estudios que contribuye a la literatura sobre los beneficios de la implementación en el desempeño de la firma utilizando reportes financieros en vez de cifras autorreportadas a través de un instrumento de recolección de datos, es el de Huang et al. (2018). Para ello se basan en literatura sobre el valor de negocios de la TI y del Big Data desde una perspectiva de negocios o económica. Derivado de la teoría RBV, Huang et al. (2018) analizan el impacto de la implementación de Big Data, en caso de que exista, en el desempeño operacional (desempeño financiero y productividad) y valor de mercado, así como la adopción temprana en la implementación de Big Data y su impacto positivo en el desempeño operacional y valor de mercado de la firma.

Específicamente, Huang et al. (2018) proponen que las firmas con implementaciones de Big Data están asociadas con un mejor desempeño financiero, mayores niveles de productividad y con mayores valores de mercado, asimismo, que las firmas con adopciones tempranas en la implementación de Big Data están relacionadas con un mejor desempeño financiero, mayores niveles de productividad y mayores valores de mercado. Los datos relacionados con el desempeño de la firma utilizados para el análisis, como se mencionó, provienen de la base de datos *Compustat*, por lo que en vez de utilizar cifras auto reportadas sobre la rentabilidad, operaciones, estructura de capital, valor de mercado y productividad, se calcularon de acuerdo a la información financiera que cada firma en la base de datos *Compustat* (Huang et al., 2018). Por otra parte, los datos sobre la implementación de Big Data se obtuvo realizando búsquedas en *Google* con 72 palabras claves que indicaran la adopción de Big Data entre 2010 a 2014, por ejemplo, artículos dónde la empresa mencionara nuevos sistemas, capacidades nuevas o mejoradas, inteligencia de negocio y *hardware* relacionado a Big Data, excluyendo en ambos

casos a empresas de la industria de TI, quedando una muestra de 386 firmas: 174 con anuncios de implementación de Big Data y 212 sin anuncios de implementación (Huang et al., 2018).

La metodología aplicada por Huang et al. (2018) para el análisis de los datos es igual al modelo de Hitt et al. (2002) para medir los beneficios de un sistema ERP, utilizando mínimos cuadrados ordinarios y controlando por la variable industria, así como otros dos modelos, uno para la función de producción (Cobb-Douglas) y otro para el valor de mercado, separando la base de datos de firmas que adoptaron Big Data en 2010 y 2011 (primeros usuarios). Los resultados encontrados sugieren que la implementación de Big Data está relacionada de forma positiva con el desempeño financiero de la firma y, que si bien la implementación de Big Data no da beneficios inmediatos en la mejora de la productividad, la percepción del mercado es positiva y puede traer beneficios futuros, lo que resulta en una asociación positiva, sin embargo, no se encontró evidencia de que el efecto fuese más fuerte para las firmas que adoptaron Big Data como primeros usuarios (Huang et al., 2018).

Explorando una nueva perspectiva, Singh y El-Kassar, (2019) toman la teoría de las capacidades dinámicas para proponer y probar empíricamente un modelo de investigación que examina el impacto de las capacidades sostenibles impulsadas por el compromiso corporativo que resulta de la integración de tecnologías de Big Data, gestión de una cadena de suministro ecológica y prácticas ecológicas de gestión de recursos humanos en el desempeño operacional y ambiental de la firma. Asimismo, buscan conocer la influencia que tienen las prácticas ecológicas de recursos humanos en la integración de tecnologías de Big Data con los procesos organizacionales para incrementar las relaciones entre prácticas ecológicas internas y externas de gestión de cadena de suministros, lo que a su vez tiene influencia en el desempeño sostenible de una empresa (Singh y El-Kassar, 2019).

La metodología utilizada por Singh y El-Kassar (2019) para probar las hipótesis propuestas en su trabajo es PLS-SEM, para lo cual se diseñó un cuestionario de medición para cada uno de los constructos con base en literatura existente, teniendo como población objetivo a empleados y gerentes de empresas en Reino de Arabia Saudita, Emiratos Árabes Unidos, Egipto y Líbano, que han implementado Big Data y que han adoptado prácticas ecológicas de gestión de cadena de suministros. Los constructos con los que operacionalizan las variables son los siguientes: compromiso corporativo, aceptación de Big Data, rutinización de Big Data,

asimilación de Big Data, colaboración de cadena de suministro ecológica, prácticas ecológicas de recursos humanos, innovación ecológica y desarrollo sostenible (Singh y El-Kassar, 2019).

Como primer paso de la metodología PLS-SEM, Singh y El-Kassar (2019) validaron el modelo exterior para estimar los coeficientes de fiabilidad de los constructos junto con su validez convergente y discriminante. Como segundo paso se ajustó el modelo interno para estimar la ruta de los coeficientes, para después utilizar *bootstrapping* para medir la significancia de los resultados obtenidos. Finalmente, se realizó un análisis de clúster para analizar el efecto mediador de las prácticas y capacitación en recursos humanos ecológicos, agrupando a las empresas encuestadas de acuerdo a sus prácticas en recursos humanos (Singh y El-Kassar, 2019).

Los resultados de Singh y El-Kassar (2019) muestran que el compromiso corporativo tiene un impacto positivo en la aceptación, rutinización y asimilación de Big Data, asimismo, influencia de forma positiva las prácticas ecológicas internas y externas de gestión de cadena de suministros. Por otra parte, también encontraron que la asimilación del Big Data a través de la aceptación y rutinización tiene un impacto positivo en las prácticas ecológicas internas de cadena de suministros y en el desempeño ambiental y organizacional de la firma, y que las prácticas ecológicas internas y externas de gestión de cadena de suministros influencia de forma positiva el desempeño ambiental y organizacional de la firma (Singh y El-Kassar, 2019). Por último, Singh y El-Kassar (2019) también encontraron que las prácticas y entrenamiento de recursos humanos ecológicos influyen la aceptación del Big Data en la rutinización, la colaboración ecológica de la cadena de suministros en la gestión ecológica interna de la cadena de suministros, y la gestión ecológica interna de la cadena de suministros en el desempeño ambiental y organizacional de la firma.

Al igual que Jeble et al. (2018), Gupta et al. (2019) se enfocan en un tipo específico de analíticos de Big Data, los predictivos, para investigar cómo los Servicios de Planeación de Recursos Empresariales basados en la nube (CERP por sus siglas en inglés) y los analíticos de Big Data y predictivos impactan el desempeño de la firma. Al igual que la mayoría de la literatura revisada, Gupta et al. (2019) parten de la teoría RVB y el argumento de que las firmas pueden construir capacidades al ser más eficientes que sus competidores para definir y desplegar recursos internos y externos. Específicamente, enmarcan el modelo en la visión de las capacidades dinámicas, definida como el proceso distintivo que permite a los recursos

relacionarse con el desempeño financiero de la firma, enfocándose en el aprendizaje y capacidades de cambio, conceptualizando a los analíticos de Big Data y predictivos y a los CERP como una de esas capacidades (Gupta et al., 2019).

Considerando lo anterior, Gupta et al. (2019) definen las capacidades de analíticos de Big Data y predictivos como un constructo de segundo orden que se compone de tres dimensiones: datos, habilidades gerenciales y habilidades técnicas; los servicios CERP como un constructo de segundo orden conformado por factores organizacionales, factores humanos y factores tecnológicos; al desempeño de la firma lo dividen en desempeño de mercado y desempeño operacional; y por último incluyen en su modelo a la cultura organizacional como un variable de efecto moderador entre los analíticos de Big Data y predictivos, los servicios CERP y el desempeño de la firma, clasificando a la cultura organizacional entre una cultura orientada al control y una cultura flexible.

La metodología utilizada para probar empíricamente el modelo de Gupta et al. (2019) es PLS-SEM, iniciando con el diseño de un instrumento de recolección de datos basado en literatura existente para operacionalizar los constructos incluidos en el modelo. Dicho instrumento se administró de forma digital a personas en puestos directivos de firmas que operan con servicios CERP y con capacidades de analíticos de Big Data y predictivos (Gupta et al., 2019). Los resultados de Gupta et al. (2019) sugieren que los servicios CERP se relacionan de forma positiva con los analíticos de Big Data y predictivos y que estos a su vez tienen un impacto positivo en el desempeño operacional y de mercado de una organización, al igual que los servicios CERP. Respecto al efecto moderador de la cultura organizacional, se encontró que tanto una orientación de control como una flexible moderan positivamente la relación entre los analíticos de Big Data y predictivos y el desempeño de mercado de la firma (Gupta et al., 2019).

En línea con otros trabajos y partiendo de la teoría RBV, Aydiner et al. (2019) proponen un modelo para examinar los efectos de la adopción de analíticos de negocio en el desempeño de los procesos de negocio, así como el rol mediador de éste último en la relación entre la adopción de analíticos de negocio y el desempeño de la firma. Para ello, definen los componentes del modelo de analíticos de negocio como las aplicaciones de adquisición y procesamiento de datos, analíticos descriptivos, analíticos predictivos y analíticos prescriptivos; y los procesos de negocio incluyen la flexibilidad financiera y no financiera, confiabilidad, capacidad de respuesta e

instrumentos de costo o activos (Aydiner et al., 2019). En ese sentido, Aydiner et al. (2019) sostienen que la adopción de analíticos de negocio influye positivamente el desempeño de los procesos de negocio, que el desempeño de los procesos de negocio influyen positivamente el desempeño de la firma y que a su vez juegan un rol mediador en la influencia de adopción de analíticos de negocio en el desempeño de la firma.

Para probar el modelo empíricamente, Aydiner et al. (2019) diseñaron un instrumento de recolección de datos basado en literatura previa, utilizando una escala de Likert de cinco puntos para medir cada uno de los ítems que conforman a los constructos del modelo: aplicaciones de adquisición y procesamiento de datos, analíticos descriptivos, predictivos y prescriptivos como constructos de primer orden que conforman la adopción de analíticos de datos, desempeño de los procesos de negocio y desempeño de la firma. Dicho instrumento fue aplicado a directivos o gerentes de firmas medianas y grandes que cotizan en la Unión de Cámaras y Bolsa de Valores de Turquía (Aydiner et al., 2019). La metodología utilizada fue SEM, encontrando que, como se propuso en las hipótesis, la adopción de analíticos de negocio tiene una influencia positiva en el desempeño de la firma, que el desempeño de los procesos de negocio influye positivamente el desempeño de la firma y que también juega un rol mediador en la relación entre adopción de analíticos de negocio y el desempeño de la firma (Aydiner et al., 2019).

Profundizando más en el tema, Ferraris et al. (2019) aseguran que el objetivo principal del proceso de recolección y análisis de Big Data es desarrollar ideas procesables y nuevos conocimientos para establecer ventajas competitivas, poniendo a la administración del conocimiento como una variable mediadora entre las capacidades de ABD y el desempeño de la firma. Al igual que gran parte de la literatura revisada, Ferraris et al. (2019) parten de la teoría de los recursos de TI y la teoría RBV, pero incorporan la administración del conocimiento, la cual incluye el proceso de la firma para adquirir nuevo conocimiento y transformarlo de tal forma que sea fácil de utilizar y acceder para poder aplicarlo dentro de la organización.

El argumento principal de Ferraris et al. (2019) es que las firmas con mayor experiencia en la administración de diferentes tipos de conocimiento tienden a ser más innovadoras pues aprovechan y explotan mejor las capacidades internas, como las capacidades de ABD, mejorando el efecto positivo de éstos en el desempeño de la firma. Teniendo en cuenta lo anterior, Ferraris et al. (2019) proponen que entre mayor sean las capacidades de ABD de la firma, mayor será el

desempeño de la firma, que a mayor orientación a la administración del conocimiento, mayor será el desempeño de la firma y que, a su vez la gestión del conocimiento actúa como una variable mediadora entre las capacidades de ABD de la firma y su desempeño.

Utilizando la metodología de SME y con base en literatura ya existente, Ferraris et al. (2019) operacionalizan las capacidades de ABD como un constructo de segundo orden conformado por capacidades de administración de ABD y capacidades técnicas de ABD; la gestión del conocimiento también se operacionaliza como un constructo de segundo orden, compuesto de las dimensiones adquisición de conocimiento, disseminación del conocimiento y sensibilidad al conocimiento; mientras que el desempeño de la firma lo miden con relación a la retención de clientes, incremento de ventas, rentabilidad y retorno de la inversión.

La recolección de los datos para probar el modelo se realizó a través de un instrumento de recolección electrónico con ítems o preguntas para medir los constructos utilizando una escala de Likert de siete puntos, el cual fue administrado a directores ejecutivos de empresas italianas miembros de la Asociación Nacional de Analíticos de Big Data (Ferraris et al., 2019). Los resultados de Ferraris et al. (2019) muestran que las firmas que desarrollan más capacidades de ABD que otras, tanto de gestión como técnicas, incrementaron su desempeño, y que una orientación hacia la gestión del conocimiento juega un rol importante en la amplificación del efecto de dichas capacidades de ABD en el desempeño de la firma.

Otra de las vertientes en la literatura revisada examina el vínculo entre los ABD y la innovación de las empresas (Niebel et al., 2019). Partiendo del argumento de que el uso de ABD puede no garantizar efectos positivos en el desempeño de la firma, sino que el conocimiento obtenido de la información analizada para construir lo ABD es el que facilita la creación de nuevos productos y servicios, la mejora de los existentes o incluso el desarrollo de nuevos modelos de negocio, Niebel et al. (2019) se basan en la literatura sobre los efectos de la productividad de las inversiones en tecnologías de la información y la comunicación, así como los efectos complementarios entre dichas inversiones y el capital intangible o capital basado en el conocimiento, para argumentar que existe una relación positiva entre los ABD y la innovación de producto en las empresas, la cual depende de factores potenciales complementarios.

El marco teórico para el modelo de Niebel et al. (2019) se basa en la función de producción del conocimiento introducida por Griliches (1979), la cual postula que un proceso de transformación que vincula diversos insumos asociados con la acumulación de conocimiento, como inversiones en investigación y desarrollo o en capital humano, con la producción innovadora de las empresas. Niebel et al. (2019) proponen dos modelos, uno que tiene como variable dependiente la propensión latente de la empresa para lograr innovaciones de producto operacionalizada como una variable *dummy* de las innovaciones de producto, y en otro como la participación porcentual de los nuevos productos en la facturación de la firma. Mientras que ambos modelos incluyen como variables independientes el uso de Big Data operacionalizada como una variable *dummy* para indicar si la empresa utiliza tecnologías de Big Data, así como un vector con las siguientes variables de control: porcentaje de inversión en investigación y desarrollo, las intensidades de las tecnologías de la información y la comunicación operacionalizada como el porcentaje de empleados que utilizan predominantemente una computadora, una variable binaria para indicar si la empresa utiliza o no software empresarial, el porcentaje de empleados altamente calificados, el porcentaje de empleados medianamente calificados, el porcentaje de empleados que ha participado en un entrenamiento específico de TI, la estructura de la edad de la fuerza laboral (porcentaje de trabajadores menores de 30 años y porcentaje de trabajadores mayores de 50 años), los años de la empresa, el tamaño de la empresa (número de empleados), inversión, si la firma es exportadora o no, si es parte de una multinacional, si es parte de un grupo nacional, si pertenece a la región que solía conformar Alemania del este, así como la industria a la que pertenece la firma (Niebel et al., 2019).

El análisis econométrico de Niebel et al. (2019) se realizó con datos de la encuesta sobre Tecnologías de la Información y Comunicación del centro de Investigación Económica Europeo (ZEW por sus siglas en alemán) en su edición 2015, la cuál es una encuesta aplicada a empresas manufactureras y de servicios localizadas en Alemania y que tienen 5 empleados o más, utilizando modelos *probit*, dividiendo la muestra entre empresas de servicios y manufactureras y con la variable *dummy* para innovación en productos como dependiente, estimaciones Tobit y Logit Fraccional con la misma división de la muestra, pero utilizando la participación porcentual de los nuevos productos en la facturación, así como regresiones *probit* dividiendo la muestra entre habilidades (generales de capital humano y específicas de TI) altas o bajas. Los resultados muestran que el uso de ABD está asociado con una mayor propensidad a innovar, así como una

mayor intensidad en la innovación y que las asociaciones son de magnitudes similares, independientemente si la empresa es manufacturera o de servicios (Niebel et al., 2019).

En literatura más reciente, el enfoque se ha dado en áreas específicas de la empresa, o bien, en alguna industria o tipo de empresa, como es el caso de Behl (2020), quién trata de entender cómo las capacidades de ABD en *startups* pertenecientes a la industria de la tecnología les permite ganar una ventaja competitiva y mejorar su desempeño. Particularmente, Behl (2020) investiga el rol que juegan las capacidades de ABD en la mejora del desempeño teniendo como contexto a *startups* de tecnología en India y China, asimismo, si hay la cultura organizacional de dichas firmas juega un rol moderador entre las capacidades de ABD y el desempeño de la firma. Teniendo como base la teoría de las capacidades dinámicas, la cuál se deriva de la teoría RBV, Behl (2020) se propone explicar cómo las capacidades de ABD ayudan a la firma a ganar una ventaja competitiva, la que posteriormente impacta de forma positiva en el desempeño de la firma, utilizando la innovación y la cultura organizacional como variables de control moderadoras en dicha relación.

Las principales hipótesis propuestas por Behl (2020) son que existe un impacto significativamente positivo de las capacidades de ABD en el desempeño de empresas *startups*, que existe un impacto significativamente positivo de las capacidades de ABD en la ventaja competitiva de las empresas *startups* y que a su vez esa ventaja competitiva tiene un impacto positivo en el desempeño de dichas empresas, así como un efecto moderador positivo de la cultura organizacional en ambas relaciones. El modelo se prueba empíricamente utilizando la metodología PLS-SEM utilizando cinco constructos: desempeño de la firma, capacidades de ABD, ventaja competitiva, cultura organizacional e innovación, y dos variables de control: edad de la firma y país de origen (Behl, 2020).

Dichos constructos se midieron utilizando un instrumento de recolección de datos con ítems o preguntas con una escala de Likert de cinco puntos basadas en literatura existente, el cual fue aplicado a los miembros fundadores de *startups* en India y China (Behl, 2020). Los datos recolectados se utilizaron para probar el modelo teórico de Behl (2020), encontrando que efectivamente, las capacidades de ABD tienen un efecto positivo y significativo en el desempeño y ventaja competitiva de empresa *startups*, con un impacto positivo de la cultura organizacional como moderador en dichas relaciones y del país de origen en el éxito de la firma, sin embargo, el

efecto de la innovación resultó ser no significativo en el modelo propuesto, así como la edad de la firma.

Por otra parte, Maroufkhani et al. (2020) se concentran en estudiar los factores tecnológicos, organizacionales y del entorno que influyen la adopción de ABD en las pequeñas y medianas empresas (PyMEs) y los efectos que su adopción tiene en el desempeño financiero y de mercado de las firmas. Como en otros estudios revisados, Maroufkhani et al. (2020) parten de la teoría RBV y cómo contribuyen a que las firmas construyan ventajas competitivas, agregando el marco teórico TOE al igual que Chen et al. (2015), y la teoría de difusión de la innovación para proponer que los factores tecnológicos de los ABD, ventaja relativa, compatibilidad, complejidad, incertidumbre e inseguridad, probabilidad y observabilidad impactan el proceso de adopción de ABD, que los factores organizacionales, soporte de los altos directivos y la preparación organizacional también tienen un efecto en la adopción de ABD, así como los factores del entorno, presión competitiva, soporte externo de vendedores y el alto grado de regulación gubernamental, mientras que, a su vez, la adopción de ABD tiene un impacto positivo en el valor de mercado y desempeño financiero de las PyMEs.

Siguiendo la metodología PLS-SEM y con base en literatura existente, los datos utilizados por Maroufkhani et al. (2020) fueron recolectados a través de un cuestionario con las medidas o ítems de cada uno de los constructos en una escala de Likert de cinco puntos, el cual fue administrado a dueños o gerentes de PyMEs de diferentes sectores de la industria manufacturera de Teherán, Irán. En el instrumento de recolección de datos y en el modelo, la adopción de ABD fue operacionalizada como un constructo de segundo orden que se refiere a cuatro valores de negocio (valor estratégico, valor transaccional, valor transformacional y valor informacional), los factores tecnológicos se componen de seis constructos (ventaja relativa, compatibilidad, complejidad, incertidumbre e inseguridad, probabilidad y observabilidad), los factores organizacionales se componen de los constructos soporte de los altos directivos y preparación organizacional, los factores del entorno de los constructos presión competitiva, soporte externo y regulación gubernamental, mientras que las dimensiones del constructo desempeño de la firma se miden con la percepción de los respondientes respecto al desempeño financiero y de mercado de la firma (Maroufkhani et al., 2020).

Maroufkhani et al. (2020) encontraron que de los factores tecnológicos, la ventaja relativa y compatibilidad no tienen efecto en la adopción de ABD para las PyMEs, lo cual es inconsistente con literatura sobre adopción de ABD en firmas grandes, mientras que la complejidad, la incertidumbre e inseguridad tienen un efecto negativo en la adopción de ABD en las PyMEs, y la probabilidad y observabilidad tienen un efecto significativamente positivo; dentro de los factores del entorno, solamente el soporte externo juega un rol sustancial en la adopción de ABD entre las PyMEs; por último que la adopción de ABD en las PyMEs si tiene un efecto sustancial en su desempeño financiero y de mercado.

Otros autores se centran en el papel moderador que juegan el tamaño de la firma y la concentración de la industria en la relación entre soluciones de ABD y la rentabilidad de la firma. Tal es el caso de Raguseo et al. (2020), quiénes al igual que estudios previos, parten de la teoría RVB para conceptualizar a los ABD como un proceso holístico que reúne recurso y capacidades que son valiosos, raros e inimitables para alcanzar ventajas competitivas, así como en la teoría de la organización industrial para soportar el rol moderador del tamaño de la firma y concentración de la industria en el desempeño de la empresa. Bajo la premisa de hallazgos previos, Raguseo et al. (2020) utilizan el desempeño financiero de la firma como variable dependiente para medir la ventaja competitiva de la firma en su modelo, también asignan un rol mediador al tamaño de la firma pues se ha encontrado que las firmas más grandes se benefician menos que las firmas pequeñas de la inversión en TI, así como un rol mediador a la concentración de la industria en la que operan las firmas pues de acuerdo con la evidencia, a menor concentración existe menor presión competitiva.

Teniendo en cuenta lo anterior, Raguseo et al. (2020) proponen que el efecto positivo de las soluciones de ABD en la rentabilidad es menor en firmas pequeñas que en grandes y que a mayor nivel de concentración en la industria en la que opera la firma, menor es la contribución de las soluciones de ABD a la rentabilidad. Para probar el modelo empírico, Raguseo et al. (2020) utilizaron una base de datos de panel para 176 empresas francesas para el periodo 2004-2016, obtenidos de la base de datos *Bureau Van Dijk DIANE* para las variables financieras, mientras que el resto de las variables se recolectaron a través de un cuestionario aplicado a los directores de información de las empresas con ítems o preguntas para medir cada una de las variables incluidas en el modelo.

La variable dependiente fue la rentabilidad, medida a través del retorno en ventas, retorno de activos y retorno sobre recursos propios, mientras que las variables independientes son ABD, concentración de la industria y tamaño de la firma, incluyendo como variables de control la edad de la firma, el año analizado (variable *dummy* de 2005 a 2016), el riesgo asumido por la firma, el desempeño de años anteriores y el dinamismo de la industria (Raguseo et al., 2020). Los resultados del modelo de datos de panel con efectos fijos aplicado por Raguseo et al. (2020) sugieren que las empresas más grandes pueden esperar mayor retorno de las inversiones en soluciones de ABD que las firmas pequeñas y que a mayor concentración del mercado, menor es la contribución de los ABD a la rentabilidad de la firma, es decir, las firmas más grandes y que operan en ambientes altamente competitivos se pueden beneficiar más de las inversiones en soluciones de ABD.

Como hemos revisado en la literatura sobre ABD, dependiendo del enfoque, la operacionalización de las capacidades de ABD se compone de diferentes constructos o dimensiones. En el caso de Hussain et al. (2020), partiendo de la teoría RBV de Barney (1991) y las aportaciones de Grant (1996) que clasifican a los recursos organizacionales en tangibles, intangibles y humanos, así como literatura relevante sobre el tema (e.g. Gupta y George (2016), Ghasemaghaei et al. (2018), Jaruphatirum y Zahedi (2007) y Wang et al. (2012)), operacionalizan la competencia de ABD como un constructo de segundo orden compuesto de recursos tangibles (datos, tecnología y recursos básicos), recursos humanos (habilidades técnicas de análisis, habilidades de gestión y dominio del conocimiento) y recursos intangibles (grandeza de los datos, calidad de los datos y cultura orientada a los datos). Sin embargo, al igual que otros autores, proponen que la competencia de ABD de una organización se asocia de forma positiva con el desempeño de la firma, así como con el desempeño del proceso de toma de decisiones, incorporando a éste último y a la capacidad de innovación como mediadores de asociación positiva entre la competencia de ABD y el desempeño de la firma (Hussain et al., 2020).

La metodología utilizada por Hussein et al. (2020) consistió en la recolección de datos a través de un cuestionario para medir cada una de las dimensiones de la competencia de ABD mediante ítems o preguntas con una escala de Likert de siete puntos, aplicado a empleados de TI o directivos de empresas en el sector telecomunicaciones o de banca en Pakistán. Posteriormente se realizó un análisis de regresión, utilizando como variable dependiente el desempeño de la

firma y como variable independiente la competencia de ABD, encontrando que el desempeño financiero y del proceso de toma de decisiones de las firmas se incrementa de forma significativa cuando se desarrolla un mayor nivel de competencia de ABD (Hussain et al., 2020). Por último, realizando un análisis mediador, Hussain et al. (2020) también encontraron que el desempeño del proceso de toma de decisiones y la competencia de innovación actúan como mediadores positivos entre la competencia de ABD y el desempeño financiero de la empresa.

A continuación se presenta un cuadro resumen con el estado del arte sobre la relación entre los ABD y el desempeño de las firmas, organizado por autor, muestra y unidad de análisis, método utilizado, así como la operacionalización de las prácticas de ABD y del desempeño de la empresa, así como los hallazgos de cada estudio revisado.

Tabla 1. Medición del impacto de las capacidades de ABD en el desempeño de la empresa

Autor(es)	Muestra y unidad de análisis	Método	Operacionalización de prácticas de ABD	Operacionalización el desempeño	Hallazgos
Cao et al., (2015)	740 respuestas	Mínimos Cuadrados Parciales con Modelo de Ecuaciones Estructurales (PLS-SEM por sus siglas en inglés)	Toma de decisiones basada en datos	Efectividad de la toma de decisiones	Los analíticos de negocio tienen un impacto positivo en la afectividad de la toma de decisiones.
Chen et al., (2015)	Muestra de 161	PLS-SEM	Uso de ABD	Crecimiento empresarial de la productividad de los activos	Uso de ABD impacta positivamente en activos y en crecimiento.
Ren et al., (2015)	287 cuestionarios	PLS-SEM	Calidad del sistema de ABD, Calidad de la información de los ABD.	Desempeño financiero; Desempeño de mercado	Valor de mercado de los ABD en desempeño de la firma.
Akter et al., (2016)	152 respuestas	Modelado de ruta de PLS y Análisis Factorial Confirmatorio (AFC)	Capacidad de gestión de ABD; Capacidad de tecnología ABD; Capacidad de talento de ABD.	Retención de clientes, crecimiento de ventas, rentabilidad, retorno de la inversión	Capacidad de ABD en el desempeño de la empresa.
Gupta y George (2016)	340 respuestas	AFE utilizando análisis de componentes principales y rotación Varimax.	Capacidad ABD: recursos tangibles; recursos humanos; recursos intangibles.	Desempeño operacional y Desempeño de mercado	Relación positiva entre CABD y desempeño operacional y de mercado
Côrte-Real et al., (2017)	175 encuestas	MCP	Alineación estratégica entre TI y negocio, rol estratégico de ABD, uso de ABD.	Desempeño estratégico y financiero	El uso de ABD tiene un impacto positivo significativo en el uso sostenido de ABD.

Autor(es)	Muestra y unidad de análisis	Método	Operacionalización de prácticas de ABD	Operacionalización el desempeño	Hallazgos
Ghasemaghaei et al., (2017)	215 encuestas	SmartPLS versión 3.0 con bootstrapping utilizando 500 re-muestras.	Uso de herramientas de datos y ABD que encajan, herramientas de personas que encajan, las tareas se adaptan.	Agilidad de la firma	El modelo propuesto se encontró no significativo.
Wamba et al., (2017)	Muestra de 297	PLS-SEM	Flexibilidad de la infraestructura de ABD; capacidades de gestión de ABD, capacidad de experiencia del personal de ABD.	El desempeño de la firma se divide en desempeño financiero y de mercado	La capacidad de ABD tiene un efecto positivo significativo en el desempeño financiero.
Mandal (2018)	249 cuestionarios	PLS en SmartPLS 2.0.M3	Capacidades de administración de ABD.	Preparación, estado de alerta y agilidad de la cadena de suministro	La planeación, coordinación y control de ABD tiene un impacto positivo. La decisión de inversión en ABD no tiene un efecto significativo.
Dubey et al., (2017)	205 cuestionarios	Regresión de Mínimos Cuadrados Ordinarios (MCO)	Capacidad ABD: recursos tangibles; recursos humanos; recursos intangibles.	La "confianza rápida" afecta la visibilidad y la coordinación.	Se ha identificado que la confianza rápida no tiene un impacto significativo en las relaciones entre la BDPA y la visibilidad y coordinación.
Ghasemaghaei et al., (2018)	151 respuestas	SmartPLS versión 2.0 y bootstrapping con 500 re-muestras	Capacidad de ABD	Calidad de las decisiones y Eficiencia de las decisiones	Una relación significativa y positiva entre la capacidad de ABD y el desempeño del proceso de toma de decisiones de la firma.
Müller et al., (2018)	Panel de 814 firmas para el periodo 2008 - 2014	MCO; estimador de efectos fijos y efectos fijos con mínimos cuadrados de dos etapas	Activos de ABD: mano de obra, capital, materiales	Productividad de la firma (ventas)	Los activos de ABD están asociados con un incremento en la productividad de la firma de 3-7% en promedio.
Raguseo y Vitari (2018)	76 cuestionarios	SEM basado en varianza	Soluciones ABD de valor empresarial.	Desempeño financiero, desempeño de mercado y satisfacción de los clientes	Las soluciones de ABD explican 62.4% de la varianza en satisfacción del cliente, 71.9% de la varianza en desempeño de mercado y 78.6% de la varianza del desempeño económico.
Jebble et al., (2018)	205 respuestas	PLS-SEM	Capacidades de ABD	Desempeño ambiental, social y económico	Impacto positivo de CABD en el desempeño ambiental, desempeño social y desempeño económico.
Anwar et al., (2018)	312 firmas	SEM y AFC	Capacidades tecnológicas de ABD; Capacidades del personal de ABD.	Ventaja competitiva y Desempeño de la firma	BDTC y BDPC tienen un impacto positivo significativo en el DF Y la VC

Autor(es)	Muestra y unidad de análisis	Método	Operacionalización de prácticas de ABD	Operacionalización el desempeño	Hallazgos
Wamba et al., (2018)	302 analistas de negocios	PLS-SEM y algoritmo REBUSPLS	Calidad de la información: integridad, actualidad, formato y precisión.	El desempeño de la firma se construye a través del valor de negocio y la satisfacción de los clientes.	La calidad de la información tiene un impacto positivo significativo en el desempeño de la firma.
Huang et al., (2018)	386 encuestas	Modelos de regresión de MCO	Implementaciones de ABD.	Desempeño financiero, productividad, valuación de mercado.	La implementación de ABD afecta positivamente el desempeño financiero y la valuación de mercado pero no la productividad.
Singh y El-Kassar (2019)	215 cuestionarios	PLS-SEM en Smart PLS 3	Compromiso corporativo, aceptación de Big Data, asimilación de Big Data.	Desempeño ambiental y desempeño organizacional	El coeficiente de ruta para el impacto de la asimilación de Big Data en DO es 0.253.
Gupta et al., (2019)	231 cuestionarios	PLS-SEM con Warp PLS versión 6.0.	Capacidad ABD: datos, habilidades gerenciales, habilidades técnicas.	Desempeño de mercado y Desempeño operacional	BDPA tiene un efecto positivo en DM. BDPA tiene un efecto positivo en DO.
Aydiner et al., (2019)	204 respuestas	AFC y SEM utilizando AMOS	Adopción de analíticos de negocio	Desempeño de los procesos de negocio y Desempeño de la firma	La adopción de analíticos de negocios influyen de forma positiva el DPN y el DFA.
Ferraris et al., (2019)	88 encuestas utilizables	SEM	Gestión de ABD y capacidades tecnológicas de ABD.	Desempeño financiero	El desarrollo de capacidades de ABD mejoró el desempeño financiero de la firma.
Niebel et al., (2019)	2706 firmas	Modelo Probit, Modelo Tobit model, Modelo Logit fraccional.	ABD es información digitalizada y tecnologías analíticas.	El porcentaje de ventas que se debió a nuevos productos / servicios.	Los ABD incrementan la tendencia a innovar en 6.7%.
Behl (2020)	514 respuestas	PLS-SEM	Capacidades de ABD: datos y habilidades técnicas y no técnicas	Ganancias promedio, retorno sobre ventas, crecimiento de cuota de mercado y crecimiento de ventas.	Relación positiva entre CABD y Desempeño de la firma
Maroufkhani et al., (2020)	171 respuestas	PLS-SEM	Factores, tecnológicos, organizacionales y del ambiente componen la capacidad de adopción de ABD	Desempeño financiero y de mercado	En las PyMEs la adopción de ABD tiene un efecto sustancial en su desempeño financiero y de mercado
Raguseo et al., (2020)	176 empresas	Regresión de datos de panel con efectos fijos	Uso de soluciones de ABD	Rentabilidad: retorno sobre ventas, retorno sobre activos y retorno sobre capital	Las empresas más grandes pueden esperar mayores retornos de las inversiones en soluciones de ABD en ambientes altamente competitivos

Autor(es)	Muestra y unidad de análisis	Método	Operacionalización de prácticas de ABD	Operacionalización el desempeño	Hallazgos
Hussain et al., (2020)	300 empresas	Análisis de regresión con mediación dual	Competencia de ABD: recursos tangibles, humanos e intangibles	Desempeño financiero: rentabilidad	Efecto positivo entre la CABD y el desempeño financiero con mediación positiva por el proceso de toma de decisiones y la competencia de innovación

Fuente: elaboración propia con base en Cao et al., (2015), Chen et al., (2015), Ren et al., (2015), Akter et al., (2016), Gupta y George (2016), Côte-Real et al., (2017), Ghasemaghahi et al., (2017), Wamba et al., (2017), Mandal (2018), Dubey et al., (2017), Ghasemaghahi et al., (2018), Müller et al., (2018), Raguseo y Vitari (2018), Jeble et al., (2018), Anwar et al., (2018), Wamba et al., (2018), Huang et al., (2018), Singh y El-Kassar (2019), Gupta et al., (2019), Aydiner et al., (2019), Ferraris et al., (2019), Niebel et al., (2019), Behl (2020), Maroufkhani et al., (2020), Raguseo et al., (2020) y Hussain et al., (2020).

Capítulo II.

Marco Teórico

Teoría Visión Basada en los Recursos

La teoría visión basada en recursos (*Resource Based View* o RBV por sus siglas en inglés) propuesta por Barney (1991), parte de dos supuestos principales. El primero sugiere que las firmas pueden controlar recursos estratégicos heterogéneos incluso si operan dentro de la misma industria, mientras que el segundo asume que dichos recursos estratégicos pueden no ser perfectamente móviles entre firmas, por lo que la heterogeneidad entre las firmas puede ser de largo plazo. Estos supuestos son la base para estudiar las implicaciones que tienen los recursos estratégicos heterogéneos como fuente de una ventaja competitiva sostenida en la firma.

Los recursos estratégicos de la firma incluyen todos los activos, capacidades, procesos organizacionales, atributos, información, conocimientos, etc. que puede controlar para implementar estrategias que mejoren su efectividad y eficiencia y los clasifica en tres categorías: recursos de capital físico, recursos de capital humano y recursos de capital organizacional. En ese sentido, la teoría RBV argumenta que la competencia entre las empresas se desarrolla sobre la base de recursos que son raros por naturaleza, difíciles de imitar, valiosos cuando se explotan y están debidamente organizados (Barney, 1991).

Por otra parte, en lo que respecta a la ventaja competitiva, Barney (1991) argumenta que una firma tiene una ventaja competitiva sobre sus competidores cuando está implementando una estrategia de creación de valor que no está siendo implementada de forma simultánea por algún otro competidor actual o potencial. Asimismo, que dicha ventaja competitiva es sostenida cuando los beneficios de la estrategia de creación de valor implementada por una firma no pueden ser replicados por sus competidores.

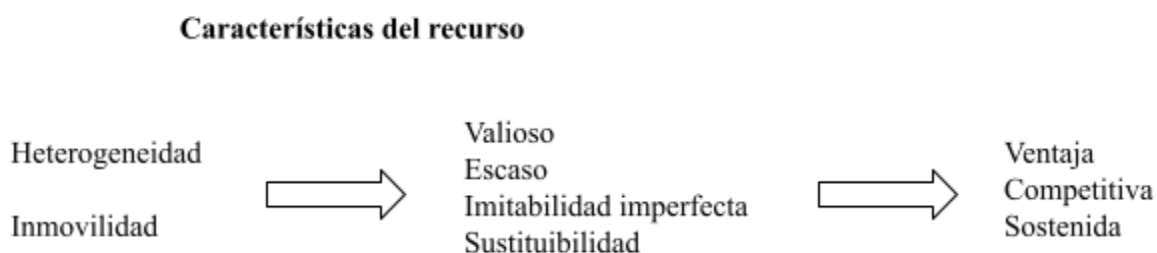
Considerando lo anterior, Barney (1991) menciona que los sistemas de procesamiento de información pueden ser fuente de una ventaja competitiva sostenida para la firma, dependiendo del tipo de sistema de procesamiento de información. Es decir, el capital físico, como computadoras, por sí solo no pueden ser una fuente de ventaja competitiva sostenida, mientras que cuando el sistema de procesamiento de información está arraigado profundamente en el

proceso de toma de decisiones formal e informal de la firma puede tener el potencial de crear una ventaja competitiva sostenida. Esto debido a que la interacción entre los sistemas de procesamiento de información y los tomadores de decisiones en la firma, así como el tipo de datos y uso que se le da a la información obtenida de dichos sistemas es un proceso a menudo complejo y específico de cada firma, por lo que se vuelve difícil de imitar para los competidores, aún cuando pertenezcan a la misma industria.

La heterogeneidad de los recursos es uno de los supuestos clave en la teoría RBV, pues indica que algunas firmas tienen mayor capacidad que otras para lograr ciertas funciones utilizando sus recursos únicos, mientras que el supuesto de la inexistencia de perfecta movilidad de los recursos facilita la diferenciación de dichos recursos (Peteraf y Barney, 2003). Bajo éstos supuestos, las capacidades de una firma representan un tipo especial de recursos que tienen por objetivo incrementar la productividad de otros recursos de la firma (Morgan et al., 2009).

Desde su publicación en 1991, el trabajo de Barney que se resume en la figura 4, ha servido como base para un sinnúmero de investigaciones en diversas áreas de la ciencia, ya sea utilizando solamente la teoría RBV como marco teórico, alguna de sus diferentes extensiones, o complementándola con otras teorías. Particularmente en lo que respecta al impacto del uso de ABD como un recurso heterogéneo y valioso que puede llevar a crear una ventaja competitiva para la firma al construir una capacidad con dichos analíticos.

Figura 4. Marco de evaluación de los recursos de la firma



Fuente: elaboración propia con base en Barney (1991).

Por ejemplo, Gupta y George (2016) utilizan la teoría RBV para validar que las inversiones para construir capacidades de ABD por sí solas no son suficientes para crear una ventaja competitiva y si efectivamente las capacidades de ABD tienen un impacto positivo en el

desempeño de la firma. Mientras que Aydiner et al. (2019) toman la teoría RBV como marco teórico para analizar el efecto de adoptar analíticos de negocio en el desempeño de los procesos de negocio y a su vez en el desempeño de la firma.

Por otra parte, Raguseo et al. (2018) la utilizan para conocer el efecto de las inversiones en ABD en el desempeño de la firma y también para conocer si el tamaño de la firma y la industria en la que opera juegan un rol mediador en la relación entre el uso de soluciones de ABD y el desempeño de la firma Raguseo et al. (2020). Asimismo, Hussain et al. (2020) también enmarcan su trabajo sobre la asociación entre la competencia de ABD en una firma con su desempeño financiero y de su proceso de toma de decisiones con la capacidad de innovación como un mediador en la teoría RBV.

Teoría de las Capacidades Dinámicas

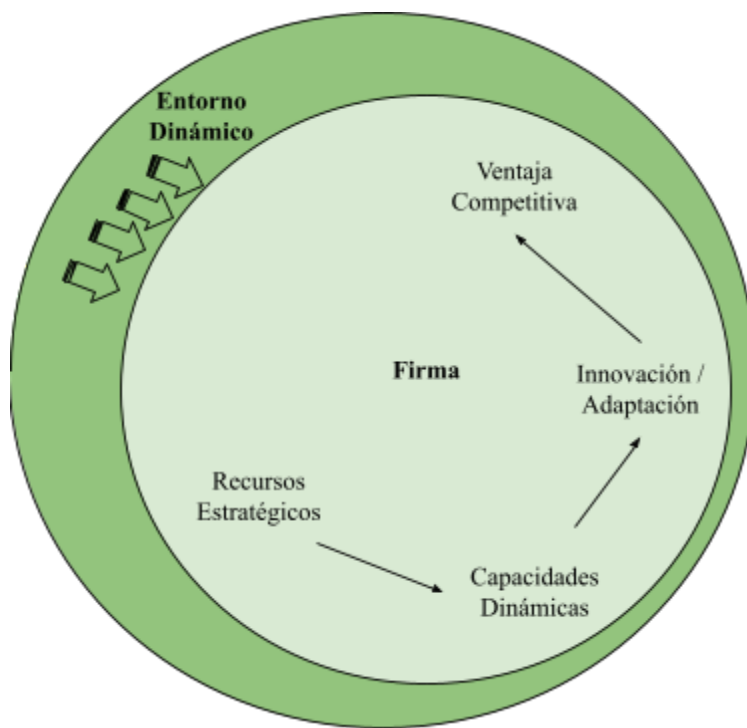
Uno de los principales enfoques de la teoría RBV es la teoría de las capacidades dinámicas (TCD), las que pueden ser definidas como procesos organizacionales únicos (específicamente los procesos que integran, reconfiguran, ganan y despliegan recursos) utilizados para crear una ventaja competitiva en un entorno dinámico (Eisenhardt y Martin, 2000). La TCD sugiere que las capacidades dinámicas permiten a las organizaciones alterar y modificar sus recursos para responder de forma eficiente a cambios rápidos en la demanda del mercado e incorporar habilidades y conocimiento a sus procesos para ganar una ventaja competitiva y crear valor (Helfat y Peteraf, 2009).

Teece et al. (1997) extendieron la teoría RBV sugiriendo que las firmas necesitan adaptarse a los cambios e incertidumbre propios del ambiente de negocios o industria en la que operan, realizando cambios o modificaciones pertinentes a sus recursos, capacidades y rutinas o procesos, refiriéndose a dicha capacidad de adaptación a la dinámica de cambios como capacidades dinámicas de la firma. La visión de la TCD intenta revelar las características de las capacidades de la firma que son necesarias para crear una combinación de recursos que relacionen procesos de negocios, como el desarrollo de nuevos productos, la asignación de los recursos y procesos de creación de conocimiento (Eisenhardt y Martin, 2000).

En resumen, la TCD surgió como respuesta a la naturaleza estática de la RBV e intenta mostrar cómo los recursos de una empresa se desarrollan e integran en un mercado que sufre

cambios rápidos y constantes (Teece et al., 1997), y se podría definir como el proceso distintivo que habilita los recursos y se enfoca en el aprendizaje y las capacidades de cambio para relacionarlas con el desempeño de la firma y la creación de una ventaja competitiva incluso en un entorno dinámico cómo se plasma en la figura 5. En ese sentido, la TCD resulta un marco teórico apropiado para caracterizar a los ABD y los recursos que los habilitan como una capacidad dinámica, puesto que recursos como los datos, activos de TI, habilidades gerenciales, habilidades técnicas, entre otros pueden ayudar a construir la capacidad de ABD y ésta a su vez puede tener un impacto en el desempeño de la organización (Gupta et al., 2019).

Figura 5. Teoría de las Capacidades Dinámicas y la creación de ventajas competitivas



Fuente: elaboración propia con base en Teece et al. (1997).

Teniendo en cuenta lo anterior, y al igual que la teoría RBV, la TCD ha sido ampliamente utilizada para estudiar el impacto del uso de ABD en las organizaciones. En la tabla 2 se enlistan los autores que han utilizado la teoría RBV como marco teórico de trabajos relacionados con el impacto de ABD en el desempeño de las empresas, así como una combinación de la teoría RBV y la TCD. Específicamente, se toma como marco teórico, en ocasiones complementado por alguna otra teoría, para caracterizar a las capacidades de ABD de la firma como una capacidad

dinámica. Uno de esos trabajos es el de Chen et al. (2015), quién utiliza la TCD y el marco teórico tecnología-organización-entorno (TOE *framework* por sus siglas en inglés) para investigar sobre los factores que influyen el uso de ABD en las organizaciones.

Tabla 2. La teoría RBV y la TCD como marco teórico en la literatura revisada

Teoría Visión Basada en los Recursos	Teoría RBV - Teoría de las Capacidades Dinámicas
Chen et al., (2015) Ren et al., (2016) Akter et al., (2016) Gupta y George, (2016) Côte-Real et al., (2017) Ghasemaghaei et al., (2017) Wamba et al., (2017) Mandal (2018) Dubey et al., (2018) Ghasemaghaei et al., (2018) Raguseo y Vitari (2018) Jeble et al., (2018) Anwar et al., (2018) Huang et al., (2018) Singh y El-Kassar (2019) Gupta et al., (2019) Aydiner et al., (2019) Ferraris et al., (2019) Behl (2020) Maroufkhani et al., (2020) Raguseo et al., (2020) Hussain et al., (2020)	Chen et al., (2015) Gupta y George, (2016) Côte-Real et al., (2017) Ghasemaghaei et al., (2017) Wamba et al., (2017) Mandal (2018) Anwar et al., (2018) Huang et al., (2018) Singh y El-Kassar (2019) Gupta et al., (2019) Aydiner et al., (2019) Behl (2020) Maroufkhani et al., (2020)

Fuente: elaboración propia con base en Chen et al., (2015), Ren et al., (2015), Akter et al., (2016), Gupta y George (2016), Côte-Real et al., (2017), Ghasemaghaei et al., (2017), Wamba et al., (2017), Mandal (2018), Dubey et al., (2017), Ghasemaghaei et al., (2018), Raguseo y Vitari (2018), Jeble et al., (2018), Anwar et al., (2018), Huang et al., (2018), Singh y El-Kassar (2019), Gupta et al., (2019), Aydiner et al., (2019), Ferraris et al., (2019), Behl (2020), Maroufkhani et al., (2020), Raguseo et al., (2020) y Hussain et al., (2020).

Côte-Real et al. (2017) también enmarcan su investigación con la TCD, incorporando la teoría visión del conocimiento para proponer un modelo que pruebe el que los ABD pueden crear valor en la cadena productiva de la firma y generar una ventaja competitiva. Mientras que Ghasemaghaei et al. (2017) utilizan la TCD para caracterizar las capacidades de TI como una capacidad dinámica de la firma que puede llevar a incrementar la agilidad de la firma, particularmente, la capacidad de TI referente a la capacidad de analíticos de datos para la toma de decisiones.

Algunos investigadores se enfocan en las capacidades dinámicas que emanan de algún área particular de la firma. Como Mandal (2018), quién utiliza la TCD para investigar sobre la influencia positiva de las capacidades de gestión de ABD (planeación, toma de decisiones respecto a inversión, coordinación y control) de la firma en la cadena de suministros de la misma. O Singh y El-Kassar (2019), que también enmarcan su investigación en la TCD para analizar el impacto de la integración de Big Data, gestión ecológica de la cadena de suministros y gestión ecológica de recursos humanos, caracterizadas como una capacidad dinámica, en el desempeño operacional y ambiental de la firma.

Mientras que algunos, además de caracterizar las capacidades de ABD como una capacidad dinámica, como Anwar et al. (2018), las desglosan en capacidades técnicas y personales para medir su relación con la ventaja competitiva y el desempeño de la firma. Otros se enfocan en las herramientas utilizadas para construir capacidades de ABD, como Gupta et al. (2019), quiénes utilizan el marco de la TCD para medir el impacto de los Servicios de Planeación de Recursos Empresariales basados en la nube (*Cloud ERP* por sus siglas en inglés) y los ABD en el desempeño de la firma. O bien, se enfocan en algún tipo de firma particular, como Behl (2020), quién enmarca las capacidades de ABD como una capacidad dinámica para estudiar si generan ventaja competitiva y mejoran el desempeño de firmas *startups* que operan en la industria de TI.

Teoría de la Visión Basada en los Recursos de Tecnologías de la Información

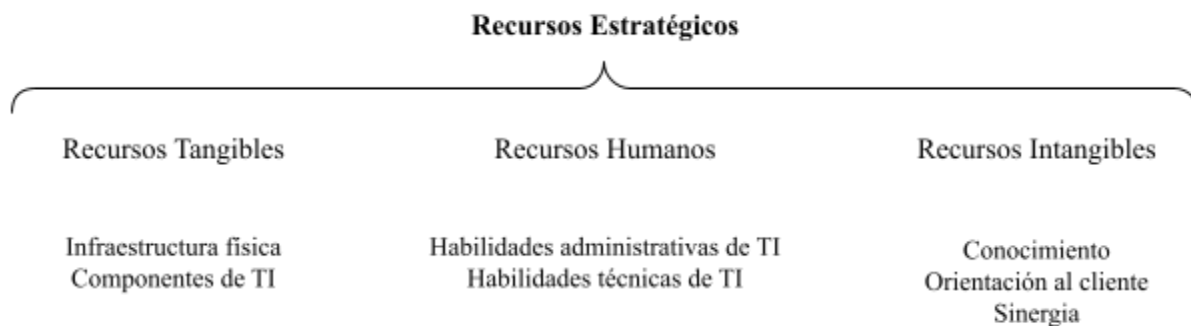
Aunque la definición de los recursos en la teoría RBV generaliza para incluir activos, conocimiento, capacidades y procesos organizacionales, Grant (1991) hace una distinción entre recursos y capacidades, clasificando además los recursos en tangibles, intangibles y basados en el personal. Los recursos tangibles se refieren al capital financiero y los activos físicos de la firma como planta y equipo; los recursos intangibles comprenden activos como la reputación, imagen de la marca y calidad de los productos o servicios; mientras que los recursos basados en el personal incluyen dimensiones como la cultura organizacional, entrenamiento a los empleados, lealtad, entre otros (Grant, 1991).

Bajo la perspectiva de la teoría RBV y tomando en cuenta la clasificación de los recursos propuesta por Grant (1991), investigadores del área de Sistemas de la Información (SI) han

identificado diversos recursos de Tecnologías de la Información (TI) que pueden ser fuentes potenciales de una ventaja competitiva para las firmas (Ferraris et al., 2019). Por ejemplo, las habilidades gerenciales de TI como un recurso basado en el personal de la firma, los sistemas inter organizacionales como recurso tangible o una fuerte relación entre las unidades de TI y de administración como un recurso intangible (Bharadwaj, 2000).

Teniendo en cuenta lo anterior, Bharadwaj (2000) desarrolla la noción tradicional de capacidades organizacionales como una función de las TI, definiendo la capacidad de TI de una firma como su “habilidad para movilizar y desplegar los recursos basados en TI en combinación o de forma co-presente con otros recursos o capacidades”. Por otra parte, utilizando la clasificación de Grant (1991) para los recursos de la firma, Bharadwaj (2000) clasifica los recursos clave o estratégicos basados en TI en: 1) recursos tangibles, que comprenden la infraestructura física de componentes de TI; 2) recursos humanos de TI, que comprenden las habilidades administrativas y técnicas de TI; y 3) los recursos intangibles habilitados por TI como recursos de conocimiento, orientación al cliente y sinergia (figura 6).

Figura 6. Clasificación de los recursos de TI



Fuente: elaboración propia con base en Bharadwaj (2000).

Bharadwaj (2000) va más allá al definir las dimensiones críticas de recursos humanos de TI que incluyen: habilidades técnicas, como el diseño y análisis de sistemas, programación y competencias en tecnologías emergentes, y habilidades administrativas de TI, que se refiere a la destreza para interactuar y coordinar efectivamente con los usuarios y concluyendo que la RBV de TI sugiere que las firmas pueden y en efecto se diferencian entre sí con base en sus recursos de TI. En esa misma línea, Fosso Wamba et al. (2019) aseguran que una capacidad de ABD

puede operacionalizarse como la combinación de datos y habilidades (tanto técnicas como blandas o *soft skills*) para hacer predicciones significativas y oportunas del mercado que, a su vez, ayudarán a diseñar una mejor estrategia frente a sus competidores y hace especial énfasis en la importancia de los datos y de las habilidades técnicas y administrativas. Asimismo, se ha demostrado empíricamente que la capacidad de ABD permite a las firmas mejorar su desempeño (Gupta y George, 2016).

Uno de los principales retos para obtener todos los beneficios de los ABD es que los administradores necesitan alinear las capacidades y cultura organizacional existentes a través de toda la organización, por lo que la incorporación de la administración del conocimiento complementa a la extensión de la teoría RBV sobre los recursos claves basados en TI para investigar el impacto de las capacidades de ABD en el desempeño de la firma (Ferraris et al., 2019). Al igual que Niebel et al. (2019), quiénes se basan en la función de productividad del conocimiento y en los efectos de la productividad de las inversiones en TI y comunicación para investigar la relación entre los ABD y la innovación de producto en la firma.

También es necesario que los empleados de todos los niveles puedan comprender los beneficios de los ABD y tomar decisiones basadas en datos, combinando varios tipos de recursos para orientarlos hacia la creación de una ventaja competitiva (Shah et al., 2012). En ese sentido, Ghasemaghaei et al., (2018) utilizan el marco teórico de los recursos clave basados en TI para medir el impacto de las capacidades de ABD en el proceso de toma de decisiones de la firma, encontrando una relación positiva.

Big Data

El uso de Big Data en una firma tiene el potencial de generar conocimiento perspicaz procesando datos generados por la empresa, el mercado u otras fuentes, el cual aunado a con las habilidades humanas promete ser otra gran revolución en la computación y crear oportunidades para las firmas (APICS, 2015).

A pesar de que no existe consenso sobre la definición o características de Big Data, puesto que la definición puede variar dependiendo del enfoque y de los autores, el término fue inicialmente acuñado para reflejar la “grandeza” o tamaño voluminoso de datos generados como

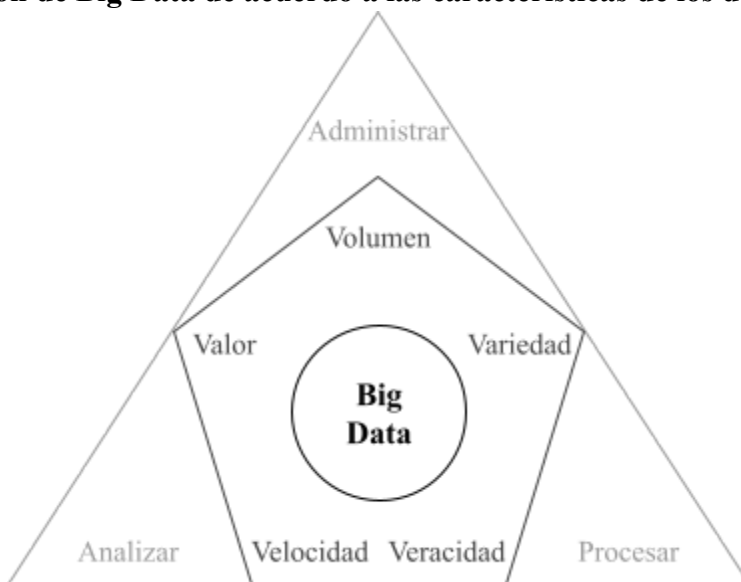
resultado de utilizar nuevas formas de tecnología, como redes sociales, etiquetas de identificación de radiofrecuencia, teléfonos inteligentes y sensores (Gupta y George, 2016). Dicha definición evolucionó para incorporar variedad (diferentes tipos de datos) y velocidad (la velocidad a la que se generan los datos) como una característica adicional al volumen del Big Data, posteriormente veracidad (desorden de los datos) y valor (perspectivas previamente desconocidas), llegando en algunos casos a incluir hasta 10 V's como características (Gupta y George, 2016).

En ese mismo sentido, Taylor (2022) define a los datos como “las cantidades, caracteres o símbolos sobre los que una computadora realiza operaciones, que pueden almacenarse y transmitirse en forma de señales eléctricas y grabarse en medios de grabación magnéticos, ópticos o mecánicos. Teniendo en cuenta lo anterior “Big Data es una colección de datos que es enorme en volumen, pero que crece exponencialmente con el tiempo... datos con un tamaño y una complejidad tan grandes que ninguna de las herramientas tradicionales de gestión de datos puede almacenarlos y procesarlos de manera eficiente... también es un dato pero con un tamaño enorme” (Taylor, 2022).

Gartner Inc. (s.f.) lo define como “activos de información de gran volumen, alta velocidad y/o gran variedad que exigen formas rentables e innovadoras de procesamiento de información que permitan una mejor comprensión, toma de decisiones y automatización de procesos” concentrándose en las características de los datos. Aydiner et al. (2019) se enfocan en los tipos de análisis que involucra el Big Data, como descriptivo, predictivo y prescriptivo, mientras que otros autores lo definen como enormes cantidades de datos estructurados y no estructurados y que son accesibles en tiempo real (Einav y Levin, 2013; O’leary, 2013).

El diccionario de APICS en su 14.^a edición (2015) define Big Data como “una colección de datos y tecnología que accede, integra e informa todos los datos disponibles mediante el filtrado, la correlación y la generación de informes de información que no se podía obtener con tecnologías de datos anteriores. Describe el procesamiento de datos más allá de la escala humana”.

Figura 7. Definición de Big Data de acuerdo a las características de los datos



Fuente: elaboración con base en Ferraris et al. (2019).

Ferraris et al. (2019) también se concentran en las características de los datos y lo definen como el enfoque para administrar, procesar y analizar cinco dimensiones de datos (figura 7), siendo estas:

1. **Volumen:** la cantidad de datos generados crece de forma exponencial.
2. **Variedad:** dado que las fuentes de los datos son diversas, pues provienen de plataformas tecnológicas diversas y en diferentes formatos.
3. **Veracidad:** los datos recolectados deben ser de calidad y provenir de fuentes confiables, evitando incluir ruido, información incompleta o desactualizada.
4. **Velocidad:** hoy en día, los datos se obtienen en tiempo real o muy cercano a tiempo real, por lo que la velocidad se vuelve más importante que el volumen, dada la competitividad actual del entorno económico, pues permiten a una empresa ser mucho más ágil y rápida en el proceso de toma de decisiones.
5. **Valor:** la extracción de beneficios económicos de la información disponible es de vital importancia y, por lo general, el valor se relaciona con la habilidad de la organización para tomar mejores decisiones.

En ese sentido, los datos representan un activo, en función del cual se persigue el valor económico y se requieren herramientas de última generación y conocimientos especializados

para ser procesados, por lo que los beneficios de implementar una estrategia de Big Data son nuevos descubrimientos, mejores sistemas de toma de decisiones y automatización (Bogdan y Borza, 2019). Sin embargo, no se debe confundir Big Data con analíticos, puesto que los ABD tienen por objetivo el drenar inteligencia de los datos y traducirlo en una ventaja competitiva para la firma (Ferraris et al., 2019), mientras que Big Data se refiere a los activos de disponer de los datos generados por la operación de la organización.

Analíticos de Big Data

El concepto de analíticos de datos fue desarrollado desde mediados de la década de 1950 y se refiere al “uso extensivo de datos, análisis estadístico y cuantitativo, modelos explicativos y predictivos, así como una administración basada en hechos para impulsar decisiones y acciones” (Cao et al., 2015). Ghasemaghaei et al. (2017) definen analíticos de datos como las TI y procesos que respaldan los informes, análisis estadísticos y minería de datos, y los dividen en 3 categorías: descriptivos o analíticos de datos para entender qué ocurrió en el pasado, predictivos o analíticos para comprender qué va a ocurrir en el futuro, y prescriptivos o analíticos utilizados para estimular resultados de posibles acciones.

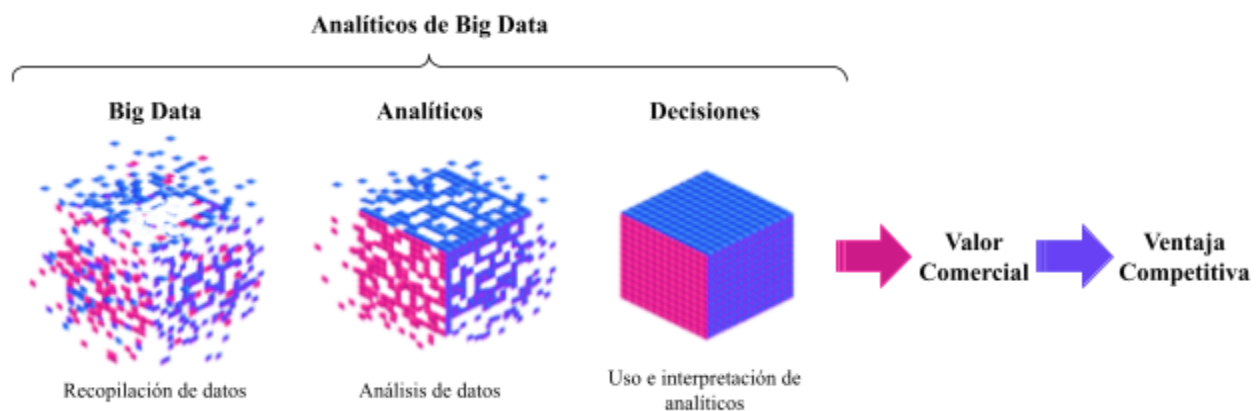
Los ABD utilizan dichas metodologías y clasificación, valiéndose de técnicas estadísticas como la regresión, análisis de factores o minería de datos para generar analíticos con bases de datos de gran volumen y con diferentes estructuras (Jeble et al., 2018). Mandal (2018) resume los ABD como una combinación de dos conceptos clave: Big Data o la capacidad para procesar datos con las 3 dimensiones (volumen, variedad y velocidad), y la habilidad para sacar conclusiones significativas de los datos mediante la aplicación de estadísticas, matemáticas, econometría, simulaciones, optimizaciones u otras técnicas para ayudar a las empresas a tomar mejores decisiones.

Sin embargo, para llegar a obtener resultados válidos y relevantes de los ABD, los encargados de extraer conocimiento de la información o ABD en la firma deben tener una comprensión detallada de los datos disponibles y conocer qué es lo que están buscando en ellos, por lo que la preparación de los datos, o primer paso para construir ABD, incluye crear perfiles, limpieza, validación y transformación de bases de datos (Botelho, 2022). Una vez que se

preparan las bases de datos para el análisis, se aplica alguna de las técnicas ya mencionadas para sacar conclusiones significativas de dichos datos que soporten el proceso de toma de decisiones.

Gupta et al. (2019) coinciden en que los ABD es un campo interdisciplinario que combina no solamente técnicas estadísticas como regresión o análisis de series de tiempo, sino que también incorpora herramientas computacionales como la minería de datos y el *Machine Learning*, para definirlo como “un proceso sistemático de análisis descriptivo para explicar las reglas de los datos, análisis predictivo para esbozar perspectivas futuras y análisis prescriptivo para optimizar o simular los resultados de las decisiones organizacionales”.

Figura 8. Definición y proceso de Analíticos de Big Data



Fuente: elaboración propia con base en Wamba et al., (2015).

Teniendo en cuenta que los datos son la materia prima, Davenport y Harris (2007) definen los ABD como “el uso extensivo de datos, análisis estadístico y cuantitativo, modelos explicativos y predictivos, y gestión basada en hechos para impulsar decisiones y acciones”. Lo que implica que los resultados obtenidos a través de ese proceso sean utilizados por los directivos de las empresas para tomar decisiones informadas y basadas en datos, llevando a un nuevo estilo de administración. Considerando lo anterior, ABD se define como un proceso holístico que involucra la recopilación, análisis, uso e interpretación de datos para varias divisiones funcionales, con miras a obtener información procesable, crear valor comercial y establecer ventajas competitivas (Wamba et al., 2015), el cual se puede observar en la figura 8.

Capacidades de ABD

Como resultado de lo expuesto anteriormente, los ABD se están convirtiendo cada vez más en un componente clave de los procesos de toma de decisiones en diferentes tipos de negocios debido a un nuevo enfoque proactivo y con visión de futuro (Hagel, 2015). Por lo que algunos autores argumentan que para obtener un resultado positivo, una organización tiene que desarrollar una capacidad de ABD, definiendo a capacidad como una combinación única de recursos, lo que dificulta, si no imposible, que los competidores la imiten y que en literatura sobre administración, esto se llamaría obtener una ventaja competitiva (Bogdan y Borza, 2019).

La teoría RBV es una de las teorías de la literatura sobre administración más utilizadas para enfocar las capacidades de ABD, la cual, como ya revisamos, argumenta que la competencia entre las empresas se desarrolla sobre la base de recursos que son raros por naturaleza, difíciles de imitar, valiosos cuando se explotan y están debidamente organizados (Barney, 1991).

En ese sentido, las capacidades de ABD en una firma, son una capacidad importante que le lleva a construir una ventaja competitiva sustentable en el ambiente de Big Data (Ferraris et al., 2019). Sin embargo, el valor agregado que se puede utilizar para construir esa ventaja competitiva emana de la capacidad de la firma para extraer conocimiento de los datos y diseminarlo en todos los niveles de la organización, haciéndolo comprensible para todos los empleados e incorporando los datos en su proceso de toma de decisiones (Shah et al., 2012).

Bajo la premisa de que, para construir capacidades que lleven a la firma a crear una ventaja competitiva, es necesario la combinación de recursos en una forma única y difícil de imitar por los competidores, la literatura sobre el tema a menudo complementa la teoría RBV con algún otro enfoque. Específicamente, se ha caracterizado a los ABD como una capacidad dinámica que combina recursos de TI y habilidades técnicas y de gestión, apoyados en la TCD y en la teoría de recursos de TI (Mandal, 2018; Singh y El-Kassar, 2019; Anwar et al., 2018; Anwar et al., 2018; Gupta et al., 2019; Behl, 2020; Ferraris et al., 2019; Ghasemaghahi et al., 2018; Bharadwaj, 2000; Fosso Wamba et al., 2019; Niebel et al., 2019). En la tabla 3 se puede observar de forma detallada cómo algunos autores operacionalizan a los analíticos de datos para su análisis.

Tabla 3. Operacionalización de Analíticos de Datos en la literatura revisada

Variable	Fuente
Capacidades de ABD	Akter et al., (2016) Anwar et al., (2018) Chen et al., (2015) Côrte-Real et al., (2019) Dubey et al., (2017) Ferraris et al., (2019) Ghasemaghahi et al., (2018) Gupta y George, (2016) Gupta et al., (2018) Huang et al., (2018) Jeble et al., (2018) Mandal, (2018) Müller et al., (2018) Nebel et al., (2019) Raguseo y Vitari, (2018) Ren et al., (2016) Singh y El-Kassar, (2019) Wamba et al., (2017) Behl, (2020) Maroufkhani et al., (2020) Raguseo et al., (2020) Hussain et al., (2020)
Analíticos de negocio	Aydiner et al., (2019) Ghasemaghahi et al., (2017)
Toma de decisiones	Cao et al., (2015) Hosoya and Kamioka, (2018) Wamba et al., (2018)

Fuente: elaboración propia con base en Cao et al., (2015), Chen et al., (2015), Ren et al., (2015), Akter et al., (2016), Gupta y George (2016), Côrte-Real et al., (2017), Ghasemaghahi et al., (2017), Wamba et al., (2017), Mandal (2018), Dubey et al., (2017), Ghasemaghahi et al., (2018), Müller et al., (2018), Raguseo y Vitari (2018), Jeble et al., (2018), Anwar et al., (2018), Wamba et al., (2018), Huang et al., (2018), Hosoya and Kamioka, (2018), Singh y El-Kassar (2019), Gupta et al., (2019), Aydiner et al., (2019), Ferraris et al., (2019), Nebel et al., (2019), Behl (2020), Maroufkhani et al., (2020), Raguseo et al., (2020) y Hussain et al., (2020).

Por ejemplo, se plantea que uno de los principales retos para obtener todos los beneficios de los ABD es que los administradores necesitan alinear las capacidades y cultura organizacional existentes a través de toda la organización (Ferraris et al., 2019), por lo que es necesario que los empleados de todos los niveles puedan comprender los beneficios de los ABD y tomar decisiones basadas en datos (Shah et al., 2012), lo que lleva a combinar recursos técnicos de TI, con recursos humanos y recursos intangibles para construir una ventaja competitiva.

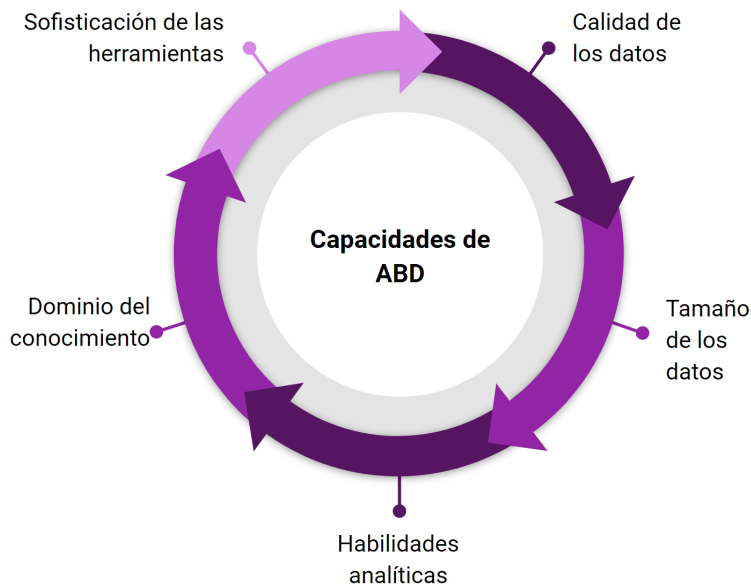
En ese sentido, Ferraris et al. (2019) definen las capacidades de ABD como “la habilidad de una firma para desplegar y combinar recursos de analíticos de datos para realizar análisis de datos rigurosos y accionables”. Por lo que las capacidades de ABD de una firma se crean con la interacción del tamaño y calidad de los datos, habilidades analíticas, dominio del conocimiento y la sofisticación de las herramientas (Ferraris et al., 2019).

Por otra parte, Ghasemaghahi et al. (2018) definen las capacidades de ABD como “la habilidad de una firma para combinar recursos de analíticos de datos para un análisis de datos riguroso y orientado a la acción”, identificando a dichas capacidades como una interacción entre

la calidad y tamaño de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de las herramientas. En la misma línea, Gupta y George (2016) argumentan que para que una firma construya una capacidad de ABD es necesaria una combinación única de recursos tangibles, humanos e intangibles, puesto que esa combinación es lo que vuelve a la capacidad de ABD única, permitiendo que lleve a la firma a la creación de una ventaja competitiva.

Adoptando la teoría RBV, su extensión TCD, el enfoque de Bharadwaj (2000) sobre los recursos basados en TI y las dimensiones propuestas por Ghasemaghaei et al. (2018), se define a las capacidades de ABD como la habilidad de la empresa para desplegar y combinar recursos de ABD para un análisis de datos riguroso y orientado a la toma de decisiones. En ese sentido se combinan recursos tangibles, intangibles y humanos en la caracterización de las capacidades de ABD, los cuales se ilustran en la figura 9 y se describen a continuación.

Figura 9. Caracterización de las Capacidades de Analíticos de Big Data



Fuente: elaboración propia con base en Ferraris et al. (2019).

Recursos tangibles

De acuerdo con la teoría RBV (Barney, 1991) y la teoría de los recursos claves de TI (Bharadwaj, 2000), los recursos tangibles de TI incluyen infraestructura de TI, redes, conectividad, fuentes de datos, entre otras, por lo que las inversiones en herramientas para

recopilar, almacenar y procesar los datos, así como los datos mismos son considerados recursos tangibles (Jeble et al., 2018). Gupta y George (2016) también mencionan a la tecnología, disponibilidad de datos, tiempo y dinero como los recursos básicos para la construcción de una capacidad de ABD.

En ese sentido, se incluye a la sofisticación de las herramientas de Big Data y a la grandeza de los datos recopilados, almacenados y procesados con ellas como las dimensiones que incorporan a los recursos tangibles en la caracterización de las capacidades de ABD para efectos de nuestro análisis.

Grandeza de los datos: como ya se mencionó, una de las características del Big Data es precisamente el volumen de los datos, pues estos crecen de forma exponencial con la digitalización derivada de una creciente adopción de la tecnología en todos los procesos de negocio. Considerando lo anterior, así como el resto de las dimensiones del Big Data, la grandeza de los datos se define utilizando tres de las cinco dimensiones mencionadas, es decir datos de gran volumen, variedad y velocidad (Ghasemaghaei et al., 2018).

Sofisticación de las herramientas: es una dimensión que tiene por objetivo capturar el nivel de *expertise* tecnológico al interior de la empresa, es decir, la complejidad y maduración de las herramientas de TI con las que cuenta para realizar el análisis de los datos y extraer información de los mismos (Ghasemaghaei et al., 2018). Por lo que, recurriendo nuevamente al marco de referencia de recursos clave basados en TI de Bharadwaj (2000), la sofisticación de las herramientas es un recurso clave basado en TI y se puede considerar como una dimensión de las capacidades de ABD, puesto que puede llevar a la empresa a construir una ventaja competitiva en el sentido del tipo y sofisticación del análisis a incorporar en el proceso de toma de decisiones. Por ejemplo, las empresas más grandes cuentan con más recursos que las pequeñas (Chen et al., 2014), por lo que pueden tener más desarrolladas las capacidades de ABD, al contar con herramientas de TI más sofisticadas que les permitan extraer un mayor valor de los datos.

Recursos humanos e intangibles

Como ya revisamos, las herramientas y datos por sí solos no generan capacidades de ABD ni llevan a la creación de una ventaja competitiva, puesto que es necesario crear una combinación única de recursos al interior de la firma, considerando el entorno dinámico para poder lograrlo

(Gupta y George, 2016). Uno de esos recursos únicos se compone de talento humano con habilidades técnicas y dominio del conocimiento sobre el negocio de la firma, así como liderazgo y una cultura orientada hacia la toma de decisiones basada en datos que permee en todos los niveles de la organización (Barney, 1991; Gupta y George, 2016; Jeble et al., 2018).

La TCD enfatiza en la necesidad de construir capacidades dinámicas que permitan a la firma adaptarse a cambios políticos, económicos, sociales, tecnológicos, entre otros, para poder desarrollar una ventaja competitiva (Teece et al., 1997). Dichos recursos son los que no aparecen en los estados financieros de la empresa, sino que se refieren principalmente a la cultura organizacional, puesto que ésta difiere de una firma a otra, volviéndola un recurso único que se desarrolla a través del aprendizaje continuo, no solamente de la organización, sino también del entorno, para convertirse así en un recurso intangible clave que contribuye a la construcción de una capacidad de ABD (Jeble et al., 2018). Teniendo en cuenta lo anterior, las dimensiones dominio del conocimiento y habilidades técnicas, así como la calidad de los datos se incluyen en la caracterización de las capacidades de ABD en nuestra propuesta de modelo.

Dominio del conocimiento y habilidades técnicas: se refiere al conocimiento y profundo entendimiento de los procedimientos, hechos y procesos involucrados en una cierta firma o industria que posee el individuo, los cuáles habilitan al analista para identificar los atributos clave y con ello poder resolver los problemas de interés de la empresa de una forma más efectiva (Ghasemaghaei et al., 2018), por lo tanto también se considera una dimensión de las capacidades de ABD. De acuerdo con la teoría RBV, para que los recursos de la firma se vuelvan difíciles de imitar por los competidores, se necesita prestar especial atención a las competencias de los individuos, pues estos son un factor crítico para mejorar el proceso de toma de decisiones a través del uso de datos. Aún si se cuenta con las dos primeras dimensiones para construir las capacidades de ABD (grandeza y calidad de los datos), es necesario que dichos datos sean transformados en conocimiento, tarea en la que los recursos humanos, sus competencias, habilidades y capacidades se vuelven críticos, pues son ellos los que transforman los datos en conocimiento a utilizar para mejorar el desempeño de la empresa.

Calidad de los datos: de acuerdo con Detlor et al. (2013), se define como la calidad de los hechos en bruto que reflejan las características de una entidad o evento. El marco de referencia para definir la calidad de los datos a utilizar es el de Wang et al. (1996), el cual diferencia entre

cuatro categorías de calidad de los datos: intrínseca, contextual, representacional y accesibilidad, así como las dimensiones subyacentes de cada categoría. La categoría intrínseca se refiere a la exactitud innata de los datos, independientemente del contexto en el que están siendo utilizados, como por ejemplo la objetividad y precisión de los mismos. Por otra parte, la categoría contextual hace referencia justamente al contexto en el cuál están siendo utilizados los datos, pues éste puede variar dependiendo de la tarea que se lleve a cabo, como integridad, oportunidad y relevancia de los datos. La categoría representacional se refiere al grado en que los datos se presentan de forma clara, por ejemplo, consistencia representacional y facilidad de comprensión. Y, por último, la categoría de accesibilidad está relacionada con la facilidad con la que se tiene acceso a los datos que se buscan. En ese sentido, y considerando también el marco de referencia de recursos clave basados en TI de Bharadwaj (2000), la calidad de los datos es considerada como una dimensión de las capacidades de ABD, puesto que es un recurso intangible clave basado en TI, al que únicamente puede extraérsele valor, o información útil para la toma de decisiones basada en datos, si se cuenta con datos de calidad a la hora de iniciar el análisis.

Teoría de los Efectos de TI avanzadas en el Diseño Organizacional, Inteligencia y Toma de Decisiones

Aún cuando se tiene una capacidad de ABD en la firma, cumpliendo con las definiciones anteriores, el reto sigue siendo permear una cultura basada en datos a los empleados de todos los niveles de la organización, desde tomadores de decisiones hasta operativos, puesto que las dimensiones mencionadas (e.g. calidad de los datos, dominio del conocimiento), dependen de la confianza que se tenga en la capacidad de ABD y el involucramiento de la información o conocimiento obtenido de los datos en el proceso de toma de decisiones (Barton y Court, 2012).

La distinción entre SI utilizados para almacenar datos transaccionales en las firmas (e.g. *Customer Relationship Management*, *Enterprise Resource Planning* o CRM y ERP por sus siglas en inglés), de los sistemas utilizados para analizar dicha información (soluciones de minería de datos o tableros de control), se originó desde los años 70's y se ha mantenido hasta hoy, argumentando que los primeros dan soporte a tareas estructuradas y rutinarias relacionadas con la operación, mientras que los segundos se utilizan para dar soporte a la toma de decisiones en actividades de administración y estratégicas (Gupta y George, 2016).

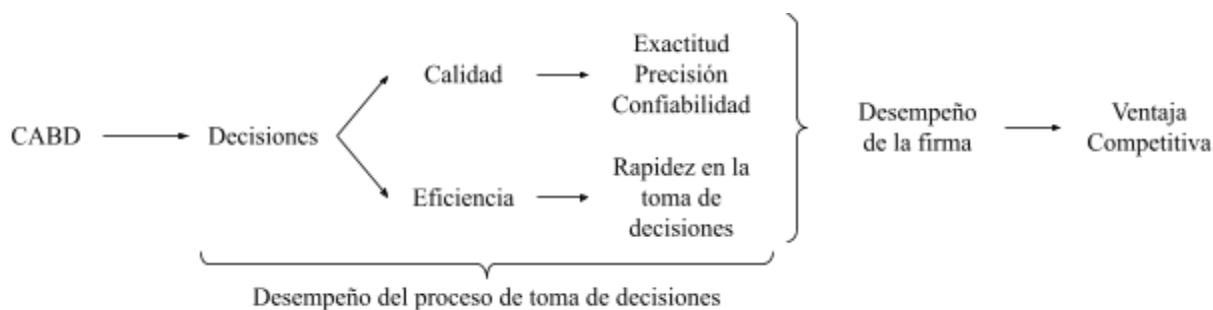
Siguiendo con la misma diferenciación, Huber (1990) desarrolló una teoría para medir los efectos que tiene la implementación de comunicaciones asistidas por computadora y tecnologías para asistir la toma de decisiones en el diseño organizacional, la inteligencia y el proceso de toma de decisiones de la firma. Dicha teoría propone que, una vez adoptada, la nueva tecnología acorta significativamente el proceso de toma de decisiones y mejora la calidad de las decisiones en la firma, puesto que la tecnología altera profundamente la forma en que se toman las decisiones (Huber, 1990; Ghasemaghaei et al., 2018).

Complementando la teoría de Huber (1990) con la teoría RBV de TI, se ha encontrado que las capacidades de TI llevan a las empresas a mejorar su proceso de toma de decisiones, puesto que propician la inteligencia de la empresa al disponer de información oportuna, comprensiva y precisa, que ayuda a identificar oportunidades y problemas, permitiendo tomar decisiones más efectivas y de calidad, argumentando que las herramientas de analíticos de datos se están convirtiendo en un componente crítico del proceso de toma de decisiones de las empresas (Ghasemaghaei et al., 2018; Wamba et al., 2017).

Teniendo en cuenta lo anterior, el proceso de toma de decisiones de la organización se vuelve clave a la hora de construir una ventaja competitiva derivada de la capacidad de ABD, puesto que las dimensiones descritas (grandeza de los datos, sofisticación de las herramientas, calidad de los datos y dominio del conocimiento y habilidades técnicas) tienen un mayor potencial para generar valor a la firma cuando efectivamente se toman decisiones que lleven a lograr sus objetivos, basadas en el conocimiento extraído de los ABD (Ferraris et al., 2019).

Por otra parte, algunos investigadores sugieren que las inversiones en Big Data no reflejan retornos debido a la resistencia de las empresas no toman decisiones basadas en la inteligencia o conocimiento extraído de los datos, enfatizando la importancia de implementar una cultura organizacional orientada a la toma de decisiones basadas en datos, en vez del instinto de los tomadores de decisiones (Gupta y George, 2016).

Figura 10. Desempeño del proceso de toma de decisiones de la empresa



Fuente: elaboración propia con base en Huber (1990), Jarupathirun (2007), Ferraris et al. (2019), Ghasemaghaei et al. (2017) y Ghasemaghaei et al. (2018).

Con base en lo descrito, para efectos de nuestro análisis el desempeño del proceso de toma de decisiones de la firma se refiere a la evaluación de los tomadores de decisiones en la firma sobre la calidad y eficiencia de su proceso de toma de decisiones (figura 10), entendiendo por calidad de las decisiones que tan exactas, precisas y confiables son los resultados de las decisiones tomadas, mientras que la eficiencia de las decisiones se refiere a qué tan rápido se llegan a tomar las decisiones (Huber, 1990, Jarupathirun, 2007; Ghasemaghaei et al., 2017; y Ghasemaghaei et al., 2018).

Desempeño de la firma

Cuando hablamos de desempeño de la firma, podemos hablar de desempeño financiero y de mercado (Raguseo y Vitari, 2018). El desempeño financiero se refiere al crecimiento de las ganancias y rentabilidad, mientras que el desempeño de mercado se refiere a mejorar la posición de la firma frente a sus competidores en un cierto mercado, lo cual no necesariamente implica un crecimiento en ganancias o en rentabilidad (Tippins y Sohi, 2003). La productividad multifactorial es una de las medidas más comunes para evaluar la productividad de la firma, relacionando una medida de producción como ventas o valor agregado, con algún factor de producción, como capital o trabajo, siendo la función Cobb-Douglas la más utilizada para medir dicha relación, puesto que en adición a los factores tradicionales, se pueden incorporar factores como los activos de TI (Müller et al., 2018). Dado que la interpretación de los coeficientes resultantes de la función Cobb-Douglas puede ser como los efectos marginales de los factores de producción en la productividad de la firma, varios estudios han concluido que los efectos de los

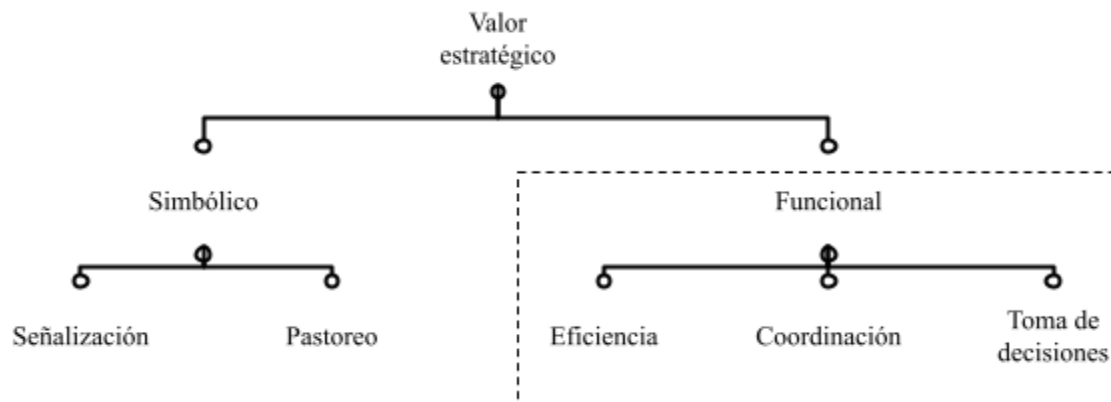
activos de TI en la productividad de la firma no son positivos, llevando a lo que se conoce como “la paradoja de la productividad de TI” (Gupta y George, 2016).

Sin embargo, se ha encontrado que a pesar de ésta paradoja, el efecto que tienen las TI, como ABD, que contribuyen a mejorar la recolección de datos, su análisis para la extracción de información y conocimiento, así como la utilización del conocimiento resultante para la toma de decisiones, es un efecto positivo en la calidad de la toma de decisiones, lo cual a su vez conduce a un incremento en la productividad de la firma (Brynjolfsson et al., 2011).

Tomando en cuenta lo anterior, hay cinco formas generales en las que el uso de ABD puede generar valor: a través de hacer que la información sea transparente y accesible a toda la organización para que se utilice; con la creación y almacenaje de más datos se puede recopilar y analizar más información para identificar patrones y variabilidad en diversas áreas, permitiendo buscar formas de mejorar el desempeño; experimentación controlada para mejorar la toma de decisiones administrativas; los pronósticos son cada vez más precisos que permiten anticiparse a las demandas del mercado, o bien, segmentar el mercado objetivo; segmentación de clientes que permite identificar productos o servicios personalizados; uso de análisis sofisticado para mejorar la toma de decisiones; y mejorar el desarrollo de nuevos productos y servicios (McKinsey, s.f.).

El valor estratégico que crean los ABD puede ser funcional o simbólico: el valor funcional (p. ej. cuota de mercado, desempeño financiero) se refiere a mejoras en el desempeño como resultado directo de la adopción de ABD, mientras que el valor simbólico (p. ej. imagen y reputación positivas, mitigación de presiones del entorno) se deriva en gran medida a través del “efecto de señal” de la inversión en ABD (Bogdan y Borza, 2019). En la figura 11 se presenta el valor estratégico de los ABD, de acuerdo con Bogdan y Borza (2019), seleccionando el enfoque de ésta investigación dentro del espectro funcional, utilizando la eficiencia orientada al desempeño de la firma, así como la toma de decisiones.

Figura 11. Valor estratégico de los ABD



Fuente: elaboración propia con base en Bogdan y Borza (2019).

Dentro de la literatura revisada como referencia, se encontró que la gran mayoría de los trabajos utilizan la rentabilidad de la firma como variable dependiente y que operacionaliza el desempeño de la firma respecto a las capacidades de ABD. En la tabla 4 se puede observar que además de las ventas y la rentabilidad, algunos autores utilizan el desempeño del proceso de toma de decisiones, los procesos de negocio o la agilidad de la firma como variable dependiente, sin embargo predomina el enfoque hacia el desempeño general de la firma con variables como ventas, clientes, retorno de la inversión, rentabilidad o productividad.

Tabla 4. Operacionalización de Desempeño en la literatura revisada

Variable	Fuente
Clientes	Akter et al., (2016) Gupta y George, (2016) Gupta et al., (2018) Raguseo y Vitari, (2018) Wamba et al., (2018) Behl, (2020) Maroufkhani et al., (2020)
Ventas	Akter et al., (2016) Huang et al., (2018) Müller et al., (2018) Niebel et al., (2019) Raguseo y Vitari, (2018) Ren et al., (2016) Wamba et al., (2017) Behl, (2020) Maroufkhani et al., (2020) Raguseo et al., (2020)
Rentabilidad	Akter et al., (2016) Anwar et al., (2018) Hosoya and Kamioka, (2018) Huang et al., (2018) Jebble et al., (2018) Côrte-Real et al., (2019) Ferraris et al., (2019) Raguseo y Vitari, (2018) Ren et al., (2016) Wamba et al., (2017) Behl, (2020) Maroufkhani et al., (2020) Raguseo et al., (2020) Hussain et al., (2020)
Retorno sobre inversión	Akter et al., (2016) Raguseo et al., (2020)
Procesos de negocio	Aydiner et al., (2019) Dubey et al., (2017) Singh y El-Kassar, (2019)
Toma de decisiones	Cao et al., (2015) Ghasemaghaei et al., (2018)
Agilidad de la firma	Ghasemaghaei et al., (2017) Hosoya and Kamioka, (2018) Mandal, (2018)
Productividad	Chen et al., (2015)

Fuente: elaboración propia con base en Cao et al., (2015), Chen et al., (2015), Ren et al., (2015), Akter et al., (2016), Gupta y George (2016), Côrte-Real et al., (2017), Ghasemaghaei et al., (2017), Wamba et al., (2017), Mandal (2018), Dubey et al., (2017), Ghasemaghaei et al., (2018), Müller et al., (2018), Raguseo y Vitari (2018), Jebble et al., (2018), Anwar et al., (2018), Wamba et al., (2018), Huang et al., (2018), Hosoya and Kamioka, (2018), Singh y El-Kassar (2019), Gupta et al., (2019), Aydiner et al., (2019), Ferraris et al., (2019), Niebel et al., (2019), Behl (2020), Maroufkhani et al., (2020), Raguseo et al., (2020) y Hussain et al., (2020).

Entorno económico en Tijuana, Baja California

Como ya describimos, existe evidencia sobre la relación positiva entre las CABD, el desempeño del proceso de toma de decisiones y financiero de la firma, a través de la creación de una ventaja competitiva, sin embargo hay pocos estudios al respecto, siendo éstos relativamente nuevos, raros, además de que se concentran en países de Norte América y Europa (Anwar et al., 2018).

Asimismo, la mayoría de los trabajos sobre el impacto del uso de datos para la toma de decisiones en el desempeño financiero de la firma, también se concentran en empresas grandes o bien, de una determinada industria (Anwar et al., 2018). Lo anterior nos lleva a proponer como objeto de estudio y como horizonte temporal y espacial a las empresas de todos los tamaños en la ciudad de Tijuana, Baja California, México. Buscando contribuir a la literatura sobre el impacto de las CABD en el desempeño del proceso de toma de decisiones y financiero de la firma en países emergentes, tomando cómo base la teoría RBV, TCD y la teoría de los efectos de TI avanzadas en el diseño organizacional, inteligencia y toma de decisiones.

Baja California se localiza al noroeste de México, teniendo como frontera norte al estado de California de E.U.A. y se compone por cinco municipios: Ensenada, Mexicali, Tecate, Tijuana, Playas de Rosarito, San Quintín y San Felipe (INEGI, 2023). En término de Producto Interno Bruto (PIB), Baja California contribuye con el 3.8% al PIB nacional a precios corrientes, ubicándose en el séptimo lugar en 2021, representando una variación de 8.2% anual, respecto de 2020 (INEGI, 2022). Al interior, Baja California se caracteriza por una importante actividad manufacturera, las cuáles representaron un 32% del PIB estatal en 2021 (INEGI, 2022), mientras que el municipio con mayor participación en términos de PIB y de población es Tijuana, que contribuye con el 60.8% del PIB estatal y concentra el 57.3% de la población total de Baja California (Citibanamex, 2022).

El crecimiento del municipio de Tijuana, en términos de población, inició en la década de los 60's con el Programa Nacional Fronterizo (PRONAF), un programa enfocado a promover la economía local y a atraer el turismo internacional, mientras que el crecimiento económico se le atribuye al programa Maquiladora establecido en 1964, cuyo objetivo era aprovechar el enorme mercado de consumo en E.U.A. y los costos competitivos del trabajo en México, así como menores costos de transportación desde la zona fronteriza del país, permitiendo a plantas ensambladoras para exportación nacionales y extranjeras ubicarse en toda la franja de la frontera norte (Ganster y Collins, 2017).

Como ya se mencionó, al igual que Baja California, Tijuana se caracteriza por una importante actividad manufacturera, aunque las actividades de construcción y comercio al por mayor y al pormenor, así como los servicios, también son actividades con gran relevancia en la región, pues durante 2021, la construcción representó el 10.2%, el comercio al por mayor y al

por menor 21% y los servicios el 23.5% del PIB estatal, situación reflejada en la ciudad de Tijuana (INEGI, 2022). Ésta composición del panorama económico en Tijuana se deriva de la intensidad de las actividades manufactureras en las últimas décadas, teniendo un boom a partir de 1980 con el establecimiento del Tratado de Libre Comercio de Norte América (TLCAN) (Mendoza y Dupeyron, 2020). El crecimiento del sector manufacturero ha resultado un detonante para otros sectores de la economía en la región, teniendo un efecto dominó al generar un mayor ingreso, impactando de forma positiva en el crecimiento de los servicios y el comercio, aunado a la dinámica generada por la interacción con la ciudad de San Diego, California (Mendoza y Dupeyron, 2020).

Por otra parte, el crecimiento y dinamismo de la región se incrementó no solamente en la industria manufacturera, sino que se dispersó a empresas de otras industrias, apoyado de nuevas tecnologías en producción y diseño, así como de la revolución en las comunicaciones, permitiendo el óptimo aprovechamiento de las economías de escala (Puyana y Romero, 2006). La internacionalización del capital a través de la industria maquiladora, hace posible el surgimiento de vínculos entre territorios con nuevas formas sociales y de producción definidos a partir del establecimiento de empresas transnacionales, generando un dinamismo en la economía regional, influyendo tanto de forma positiva como de forma negativa en el resto de los sectores de actividad económica (Martínez Cuero, 2017). Por ejemplo, el incremento en actividades relacionadas con la logística, turismo, alimentos y bebidas, así como la industria inmobiliaria, industrial y comercial, son sectores que se ven influenciados de forma positiva por el establecimiento de empresas transnacionales en la región fronteriza, particularmente en Baja California, mostrando un crecimiento de 3.9% en el indicador trimestral de la actividad económica estatal para 2022 (INEGI, 2023).

Teniendo en cuenta lo anterior, para efectos de ésta investigación, se considera cómo objeto de estudio a empresas localizadas en la ciudad de Tijuana, Baja California, independientemente del sector de actividad económica o el tamaño de las mismas, esto a efectos de cumplir con lineamientos metodológicos de muestra pues es necesario realizar trabajo de campo para recolectar información relacionada con el desarrollo de las CABD y sus efectos en el desempeño del proceso de toma de decisiones y financiero de las empresas. Específicamente, se busca conocer, en primera instancia si las empresas de Tijuana utilizan analíticos de datos para la

toma de decisiones, y en segunda si el desarrollo de las CABD tiene un impacto positivo en el desempeño de su proceso de toma de decisiones y desempeño financiero, buscando contribuir así a la literatura sobre el impacto de las CABD en el desempeño del proceso de toma de decisiones y financiero de las firmas en mercados emergentes.

Hipótesis de investigación

Teniendo en cuenta el estado del arte sobre la relación entre CABD y el desempeño de las firmas, así como el marco teórico descrito en párrafos anteriores, en éste apartado se busca resumir y definir las teorías y conceptos a utilizar para plantear las hipótesis propuestas que buscan dar respuesta a las preguntas y objetivos de investigación. Como se ha dicho antes, las preguntas de la presente investigación son: 1) ¿en qué medida afectan las CABD el desempeño del proceso de toma de decisiones de las empresas en Tijuana, Baja California?; 2) ¿Cuál es el impacto de las CABD en el desempeño financiero de las empresas en Tijuana, Baja California?; y 3) ¿actúa el desempeño del proceso de toma de decisiones como mediador entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California?

Mientras que el objetivo general de investigación es analizar el impacto de las CABD en el proceso de toma de decisiones y el desempeño financiero de las empresas en Tijuana, Baja California, mientras que los objetivos específicos son 1) examinar el impacto de las CABD en el desempeño del proceso de toma de decisiones de las firmas en Tijuana, Baja California, en términos de calidad y eficiencia de las decisiones; 2) analizar de forma empírica el impacto de las CABD en el desempeño financiero de las firmas en Tijuana, Baja California; y 3) identificar si existe un efecto mediador del desempeño del proceso de toma de decisiones entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California.

La teoría RBV propuesta por Barney (1991) sugiere que las firmas pueden controlar recursos estratégicos heterogéneos que pueden ser fuentes de una ventaja competitiva sostenida en la firma. Los recursos estratégicos son aquellos recursos raros, difíciles de imitar, valiosos y que debidamente organizados permiten a la firma desarrollar una ventaja competitiva, y se clasifican en tres categorías: recursos de capital físico, humano y organizacional (Barney, 1991). La competencia entre las empresas se desarrolla sobre la base de dichos recursos, mientras que la ventaja competitiva se logra cuando una firma implementa una estrategia de creación de valor

que no es implementada simultáneamente por sus competidores y cuando los beneficios de dicha estrategia no pueden ser replicados por ellos (Barney, 1991), logrando esto a través de recursos especiales orientados a incrementar la productividad de otros recursos, lo que se conoce como capacidades de la firma (Morgan et al., 2009).

En ese sentido, los sistemas de procesamiento de información pueden ser fuentes de ventaja competitiva sostenida, ya que la interacción entre los sistemas y los tomadores de decisiones es un proceso complejo y específico de cada firma, difícil de imitar por los competidores (Barney, 1991). Teniendo en cuenta lo anterior, la teoría RBV, complementada con algún otro enfoque, ha sido ampliamente utilizada como marco teórico en investigaciones en diversas áreas de la ciencia, particularmente en lo que respecta al impacto del uso de ABD como un recurso heterogéneo y valioso que puede crear una ventaja competitiva para la firma al construir una capacidad con dichos analíticos.

Uno de los enfoques más utilizados para complementar la teoría RBV en el estudio del impacto del uso de ABD en la firma, es el de la TCD, la cuales se definen como procesos organizacionales únicos utilizados para crear una ventaja competitiva en un entorno dinámico (Eisenhardt y Martin, 2000). La TCD surgió como respuesta o extensión de la teoría RBV, intentando demostrar cómo los recursos de la firma se deben desarrollar e integrar en un entorno con cambios constante y rápidos (Teece et al., 1997), por lo que resulta un marco teórico apropiado para el estudio de los ABD y los recursos que las habilitan (datos, activos de TI, habilidades gerenciales, entre otros) como una capacidad dinámica que puede tener un impacto en el desempeño de la firma (Gupta et al., 2019), siendo utilizada por varios autores dentro de éste campo de investigación (Chen et al., 2015; Côte-Real et al., 2017; Ghasemaghahi et al., 2017; Mandal, 2018; Singh y El-Kassar, 2019; Anwar et al., 2018; Behl, 2020; y Maroufkhani et al., 2020).

De forma más específica, la teoría RBV y la TCD han sido utilizadas en el área de SI, identificando varios recursos de TI que pueden ser fuentes potenciales de ventaja competitiva para las firmas, clasificándolos en tangibles (infraestructura física), humanos (habilidades administrativas y técnicas) e intangibles (conocimiento, orientación al cliente y sinergia) (Grant, 1991; Bharadwaj, 2000; y Ferraris et al., 2019). Teniendo en cuenta lo anterior, la teoría RBV de TI ha sido utilizada como marco teórico para la investigación del impacto de las capacidades de

ABD en el desempeño de la firma (Ferraris et al., 2019), así como en el proceso de toma de decisiones de la misma (Ghasemaghaei et al., 2018).

Por otra parte, la definición de Big Data para efectos del presente trabajo, se define cómo el enfoque para administrar, procesar y analizar las cinco dimensiones de datos: volumen, variedad, veracidad, velocidad y valor (Ferraris et al., 2019). Big Data hace referencia al proceso de disponer de los datos generados por la firma y recolectados de su entorno, los cuáles representan un activo del cual se puede extraer conocimientos para toma de decisiones que puede derivar en la creación de una ventaja competitiva y una mejora en el desempeño de la firma, a través de los ABD (Bogdan y Borza, 2019 y Ferraris et al., 2019). Dicho proceso culmina en los analíticos o ABD, un campo interdisciplinario que utiliza los datos como un recurso del cuál se drena inteligencia de negocios, y se define como un proceso holístico de recopilar, analizar, utilizar e interpretar los datos de la firma para generar conocimiento accionable que permita crear valor comercial y establecer ventajas competitivas (Wamba et al., 2015 y Ferraris et al., 2019).

Considerando las definiciones anteriores, para efecto de la presente investigación se tomará como base la teoría RBV (Barney, 1991), su extensión TCD (Teece et al., 1997), el enfoque de Bharadwaj (2000) sobre la teoría RBV y los recursos basados en TI y las dimensiones de ABD propuestas por Ghasemaghaei et al. (2018) y Ferraris et al. (2019), para definir a las CABD como una habilidad de la firma para desplegar y combinar recursos de ABD para un análisis riguroso y orientado a la toma de decisiones, caracterizándose como una combinación de recursos tangibles (grandeza de los datos y sofisticación de las herramientas) y recursos humanos e intangibles (dominio del conocimiento, habilidades técnicas y calidad de los datos). Éstas cinco dimensiones caracterizadas como CABD tienen, a su vez, un impacto tanto en el desempeño del proceso de toma de decisiones, como en el desempeño financiero de la firma, contribuyendo finalmente en la creación de ventajas competitivas para la firma (Huber, 1990, Jarupathirun, 2007; Ghasemaghaei et al., 2017; Ferraris et al., 2019; y Ghasemaghaei et al., 2018).

En lo que respecta al impacto que tienen las CABD en el proceso de toma de decisiones de la firma, Ghasemaghaei et al. (2018) y Wamba et al. (2017) complementan la teoría de los efectos de TI avanzadas en el diseño organizacional, inteligencia y toma de decisiones (Huber, 1990) con la teoría RBV de TI y la TCD, para probar que las capacidades de TI, en este caso las

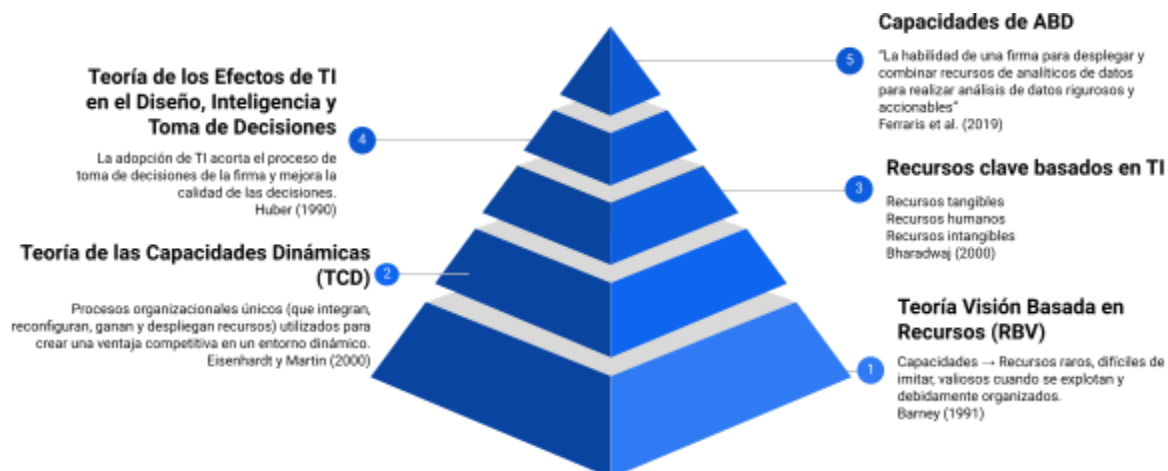
CABD, llevan a un mejor desempeño del proceso de toma de decisiones, en el sentido de contribuir a desarrollar capacidades e inteligencia de negocios pues se dispone de información transformada en conocimiento que permite tomar decisiones efectivas y eficientes. Teniendo en cuenta lo anterior, evaluamos el proceso de toma de decisiones de la firma mediante la percepción que los tomadores de decisiones tienen referente a la calidad y eficiencia del proceso, entendiendo por calidad de las decisiones lo que respecta a la exactitud, precisión y confiabilidad de los resultados obtenidos, mientras que eficiencia se evalúa midiendo la rapidez con la que se llega a tomar una decisión (Huber, 1990, Jarupathirun, 2007; Ghasemaghaei et al., 2017; y Ghasemaghaei et al., 2018).

El desarrollo de las CABD en la firma para una mejor recolección y análisis de los datos, contribuye a la generación de información y conocimiento, que, cuando es utilizado para la toma de decisiones, puede tener un efecto positivo en la calidad y efectividad de las mismas, conduciendo a un incremento de la productividad de la firma, actuando el proceso de toma de decisiones como un mediador entre las CABD y el desempeño financiero (Brynjolfsson et al., 2011). El trabajo de Ferraris et al. (2019) hace énfasis en la importancia del conocimiento y su gestión para fortalecer las capacidades de ABD de la firma y el efecto que tiene sobre el desempeño financiero de la firma, caracterizando a éste último como retención de clientes, rentabilidad y retorno de la inversión.

Adicionalmente, con sustento en la teoría RBV, se ha demostrado que las capacidades de ABD tienen un efecto positivo en el desempeño de las firmas (Bogdan y Borza, 2019). Por ejemplo, Côte-Real et al. (2019) encontraron que la utilización de ABD tiene un impacto positivo en el desempeño financiero de la organización, siendo el contribuidor más crítico a esa generación de valor el uso constante de los ABD para la toma de decisiones. Asimismo, Akter et al. (2016) demuestran que la capacidad de ABD tiene un efecto positivo en el desempeño de la firma, el cual se mide con cuatro variables: retención de clientes, crecimiento de ventas, rentabilidad y retorno de la inversión, al igual que Anwar et al. (2018), quienes adicional a la relación positiva entre capacidades de ABD y desempeño de la firma, también encuentran un desempeño positivo en la ventaja competitiva de las firmas que utilizan ABD en su proceso de toma de decisiones.

De forma general, el modelo de investigación propuesto se basa en la teoría RBV de Barney (1991) para caracterizar a las CABD como una capacidad de la firma, pues se trata de desarrollar la recolección y análisis de los datos, así como el conocimiento extraído de los mismos, como recursos raros, difíciles de imitar, valiosos cuando se explotan y organizarlos debidamente para generar una ventaja competitiva en la firma. Además se complementa con la TCD (Eisenhardt y Martin, 2000), para enmarcar dichas capacidades en el entorno dinámico actual y en la propuesta de los recursos clave basados en TI de Bharadwaj (2000) para incorporar dentro de los recursos a desarrollar como CABD a los recursos tangibles, humanos e intangibles de la firma que se involucran en el proceso de desplegar y combinarlos para realizar análisis de datos rigurosos y accionables (Ferraris et al., 2019). Por otra parte, el modelo se basa en la teoría de los efectos de TI en el diseño, inteligencia y toma de decisiones de Huber (1990) para explorar el efecto de las CABD en el proceso de toma de decisiones de la firma, así como en los hallazgos de diversos autores sobre el efecto positivo de las CABD en el desempeño financiero de la firma medido a través de la rentabilidad y ventas (Akter et al., 2016; Anwar et al., 2018; Hosoya and Kamioka, 2018; Huang et al., 2018; Jeble et al., 2018; Côte-Real et al., 2019; Ferraris et al., 2019; Ren et al., 2016; Wamba et al., 2017; Behl, 2020; Maroufkhani et al., 2020; Raguseo et al., 2020). En la siguiente figura se puede apreciar un resumen de las teorías base para el modelo de investigación propuesto.

Figura 12. Marco teórico de investigación



Fuente: elaboración propia con base en Barney, 2019; Bharadwaj, 2000; Huber, 1990; Eisenhardt y Martin, 2000; Ferraris et al. 2019.

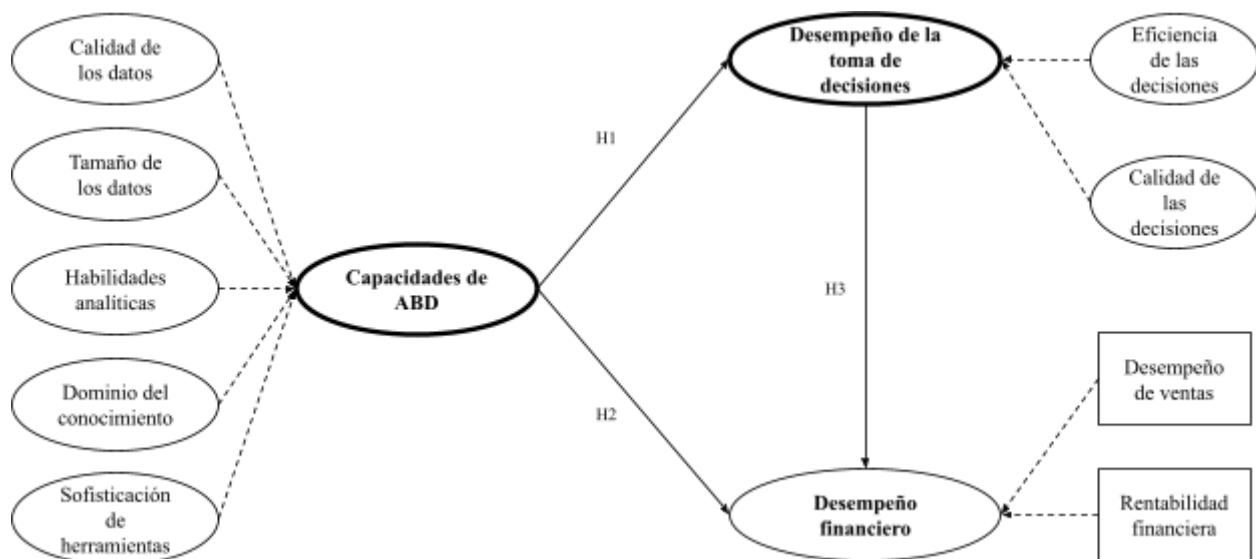
Teniendo en cuenta lo anterior, las CABD de la firma se componen de recursos tangibles (grandeza de los datos y sofisticación de las herramientas de TI), y recursos humanos e intangibles (dominio del conocimiento, habilidades técnicas y calidad de los datos), el desempeño del proceso de toma de decisiones se mide a través de la percepción de los tomadores de decisiones de la firma sobre la calidad y eficiencia de las decisiones tomadas, mientras que el desempeño financiero de la firma se mide a través de la rentabilidad y las ventas. En ese sentido, se considera al desempeño del proceso de toma de decisiones como la variable dependiente para buscar responder a la pregunta de investigación sobre el efecto de las CABD en el desempeño del proceso de toma de decisiones de la firma, teniendo como primera hipótesis que a mayores CABD, mejor será el desempeño de la toma de decisiones de la firma. Continuando con ese razonamiento, se considera también como variable dependiente al desempeño de la firma se mide la rentabilidad financiera y el desempeño de las ventas para responder a la pregunta de investigación relacionada con el impacto de las CABD en el desempeño de la firma, proponiendo como segunda hipótesis que el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la firma. Asimismo, se plantea como tercera hipótesis que el desempeño del proceso de toma de decisiones de la firma tiene un efecto mediador entre las CABD y el desempeño financiero, buscando responder a la pregunta de investigación relacionada con el efecto mediador del desempeño del proceso de toma de decisiones en la organización. En la figura 13 se plantea un resumen con los objetivos, preguntas e hipótesis de investigación, mientras que en la figura 14 se ilustra el modelo de investigación propuesto.

Figura 13. Objetivos, preguntas e hipótesis de investigación

Objetivo general: analizar el impacto de las CABD en el proceso de toma de decisiones y el desempeño financiero de las empresas en Tijuana, Baja California			
Objetivos específicos	Preguntas de investigación	Hipótesis	
1 Examinar el impacto de las CABD en el desempeño del proceso de toma de decisiones de las empresas en Tijuana, Baja California, en términos de calidad y eficiencia de las decisiones	¿En qué medida afectan las CABD el desempeño del proceso de toma de decisiones de las empresas en Tijuana, Baja California?	A mayores CABD, mejor será el desempeño del proceso de toma de decisiones de la empresa	H1
2 Analizar de forma empírica el impacto de las CABD en el desempeño financiero de las empresas en Tijuana, Baja California	¿Cuál es el impacto de las CABD en el desempeño financiero de las empresas en Tijuana, Baja California?	El desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la empresa	H2
3 Identificar si existe un efecto mediador del desempeño del proceso de toma de decisiones entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California	¿Actúa el desempeño del proceso de toma de decisiones como mediador entre las CABD y el desempeño financiero de las empresas en Tijuana, Baja California?	El desempeño del proceso de toma de decisiones de la empresa tiene un efecto mediador entre las CABD y el desempeño financiero	H3

Fuente: elaboración propia.

Figura 14. Modelo de investigación propuesto



Fuente: elaboración propia.

Capítulo III.

Diseño metodológico

Con el objetivo de responder a las preguntas de investigación y probar las hipótesis propuestas, en ésta sección se plantea un modelo de investigación basado en la metodología de los Modelos de Ecuaciones Estructurales estimados a través de Mínimos Cuadrados Parciales (*Partial Least Squares Structural Equation Model* o PLS-SEM por sus siglas en inglés). Teniendo en cuenta lo anterior, éste capítulo se centra en explicar cómo se respondieron las preguntas de investigación, describiendo los métodos y criterios de aplicación y aceptación de un modelo PLS-SEM, el modelo empírico propuesto para probar las hipótesis, así como el proceso de recolección de los datos, diseño del instrumento de recolección, y se organiza en tres secciones: en la primera nos enfocamos en el proceso de aplicación de los modelos PLS-SEM, desde las consideraciones previas, hasta los criterios de aceptación de las variables; posteriormente, en la segunda sección, describimos el modelo empírico propuesto teniendo en cuenta la metodología PLS-SEM; y, finalmente, en la tercera sección, describimos el proceso metodológico de recolección de datos y definición de la muestra a utilizar en el modelo empírico.

Modelo PLS-SEM

De acuerdo con la literatura pertinente, la metodología más apropiada para analizar el impacto de los analíticos de datos en el desempeño de la firma con base en la teoría RBV y TCD, son los Modelos de Ecuaciones Estructurales con Mínimos Cuadrados Parciales (*Partial Least Squares Structural Equation Model* o PLS-SEM por sus siglas en inglés). En la tabla 5 podemos observar que la mayoría de las investigaciones recientes sobre el impacto de los analíticos de datos en el desempeño de la firma, utilizan los Modelos de Ecuaciones Estructurales en combinación con la metodología de Mínimos Cuadrados Parciales o PLS-SEM. Teniendo esto en cuenta, se seleccionó la metodología PLS-SEM para analizar el impacto de las CABD en el desempeño del proceso de toma de decisiones y financiero de las empresas en Tijuana, Baja California, objeto de ésta investigación.

Tabla 5. Metodologías utilizadas en el análisis de las capacidades de analíticos de datos

Método	Fuente
Mínimos Cuadrados Parciales (MCP)	Akter et al., (2016) Côrte-Real et al., (2019) Mandal, (2018) Maroufkhani et al., (2020)
Análisis Factorial Confirmatorio (AFC)	Akter et al., (2016) Anwar et al., (2018) Aydiner et al., (2019)
Modelo de Ecuaciones Estructurales (MEE)	Anwar et al., (2018) Aydiner et al., (2019) Ferraris et al., (2019) Hosoya and Kamioka, (2018) Raguseo y Vitari, (2018)
Modelo de Ecuaciones Estructurales con Mínimos Cuadrados Parciales (PLS-SEM)	Cao et al., (2015) Chen et al., (2015) Ghasemaghaei et al., (2018) Ghasemaghaei et al., (2017) Gupta et al., (2018) Behl, (2020) Jeble et al., (2018) Ren et al., (2016) Singh, (2019) Wamba et al., (2017) Wamba et al., (2018) Hussain et al., (2021)
Mínimos Cuadrados Ordinarios (MCO)	Dubey et al., (2017) Huang et al., (2018) Müller et al., (2018)
Análisis Factorial Exploratorio (AFE)	Gupta y George, (2016)
Modelo Probit y Logit	Niebel et al., (2019)

Fuente: elaboración propia con base en Cao et al., (2015), Chen et al., (2015), Ren et al., (2015), Akter et al., (2016), Gupta y George (2016), Côrte-Real et al., (2017), Ghasemaghaei et al., (2017), Wamba et al., (2017), Mandal (2018), Dubey et al., (2017), Ghasemaghaei et al., (2018), Müller et al., (2018), Raguseo y Vitari (2018), Jeble et al., (2018), Anwar et al., (2018), Wamba et al., (2018), Huang et al., (2018), Hosoya and Kamioka, (2018), Singh y El-Kassar (2019), Gupta et al., (2019), Aydiner et al., (2019), Ferraris et al., (2019), Niebel et al., (2019), Behl (2020), Maroufkhani et al., (2020), Raguseo et al., (2020) y Hussain et al., (2020).

Criterios de aplicación de un modelo PLS-SEM

Los modelos PLS-SEM son una metodología utilizada para analizar relaciones complejas entre constructos e indicadores (Hair et al., 2023). Éste tipo de modelo se utiliza cuando el objetivo de la investigación es la predicción o el desarrollo de teoría y es, conceptual y prácticamente, similar al análisis de regresión múltiple, pues su objetivo principal es maximizar la varianza explicada en los constructos dependiente, así como evaluar la calidad de los datos con base en las características de medición del modelo (Hair et al., 2011). En la tabla 6 se enlistan los criterios bajo los cuáles se considera pertinente utilizar un modelo PLS-SEM. En ese sentido y

teniendo en cuenta el marco teórico de la presente investigación, la metodología a aplicar es un modelo PLS-SEM puesto que el objetivo de investigación es una extensión de teoría existente, se manejan constructos formativos y con relaciones complejas entre los indicadores, adicional a disponer de un tamaño de muestra relativamente bajo y la utilización de puntuaciones para construir las variables latentes o no observables.

Tabla 6. Criterios para la aplicación de un modelo PLS-SEM

Objetivos de la investigación
El objetivo de la investigación es exploratorio o una extensión de una teoría estructural existente.
Especificación del modelo de medición
Si los constructos formativos son parte del modelo estructural.
Modelo estructural
Cuando el modelo estructural es complejo, es decir, tiene muchos constructos y muchos indicadores.
Características de los datos y algoritmo
Cuando el tamaño de la muestra es relativamente bajo.
Si los datos no cumplen, en alguna medida, con los supuestos de normalidad.
El tamaño mínimo de la muestra debe ser mayor que: 10 veces más grande que el número de indicadores formativos utilizados para medir un constructo; o 10 veces más grande que el número de trayectorias estructurales dirigidas a un constructo latente en particular dentro del modelo estructural.
Evaluación del modelo
Si se van a utilizar las puntuaciones de las variables latentes en análisis subsecuentes.

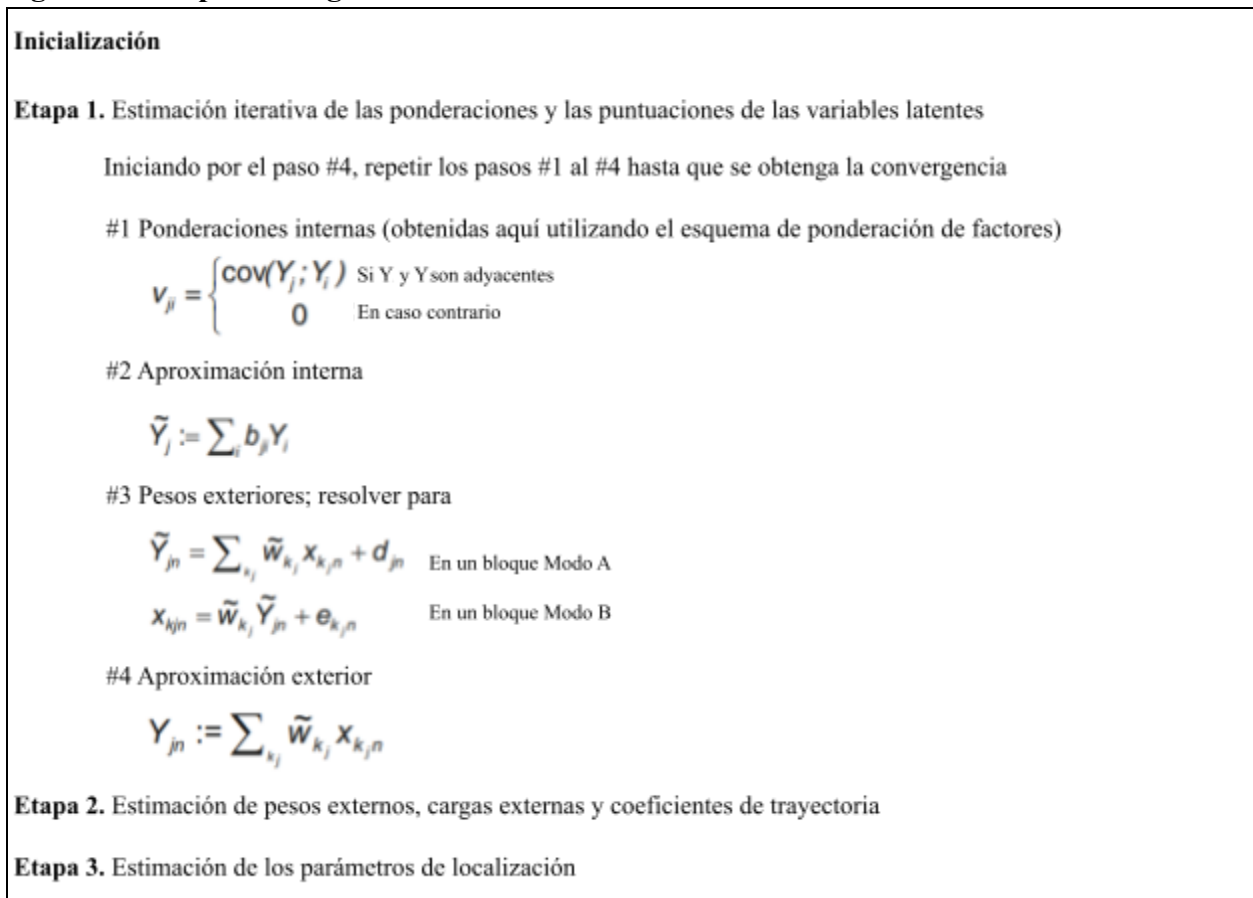
Fuente: elaboración propia con base en Hair et al. (2011).

Algoritmo PLS-SEM y modelos de trayectoria

El algoritmo PLS-SEM fue desarrollado por Wold en 1982 y modificado por Lohmöller en 1989, y es básicamente una secuencia de regresiones en términos de vectores de pesos, que obtenidos en convergencia, satisfacen ecuaciones punto fijo (Ringle et al., 2022). La estimación de modelos PLS-SEM sigue un proceso de tres etapas, al igual que la familia de algoritmos alternantes de mínimos cuadrados, las cuáles se describen en la figura 15 (Sarstedt et al., 2021). La fase de inicialización del algoritmo es en la que se establecen las puntuaciones preliminares de las variables latentes, utilizando para su cálculo ponderaciones unitarias para todos los indicadores del modelo de medición (Sarstedt et al., 2021). Posteriormente, en la etapa uno, el algoritmo determina de forma iterativa las ponderaciones internas (o coeficientes de trayectoria)

y las puntuaciones de las variables latentes, a través de cuatro pasos: 1) utilizando las puntuaciones preliminares de la fase de inicialización, se determinan las ponderaciones internas entre las variables dependientes e independientes del modelo estructural; 2) se calculan *proxies* para las variables latentes usando la suma ponderada de las puntuaciones de sus variables latentes adyacentes, para obtener así las aproximaciones del modelo interior; 3) posteriormente, se calculan nuevas ponderaciones externas para todos los indicadores del modelo de medición, las cuáles indican la fuerza de la relación entre cada variable latente y sus indicadores, diferenciando entre variables latentes o constructos reflectivos (Modo A o ponderaciones de correlación) y constructos formativos (Modo B o ponderaciones de regresión); y 4) se actualizan las puntuaciones de las variables latentes, combinando linealmente los pesos exteriores del paso 3 y los indicadores; el proceso se repite de forma iterativa hasta que los pesos obtenidos en el paso 3 cambian marginalmente de una iteración a otra o cuando el número máximo de iteraciones termina, las cuales son especificadas a la hora de correr el algoritmo (Ringle et al., 2022).

Figura 15. Etapas del algoritmo básico PLS-SEM



Fuente: elaboración propia con base en Sarstedt et al. (2021).

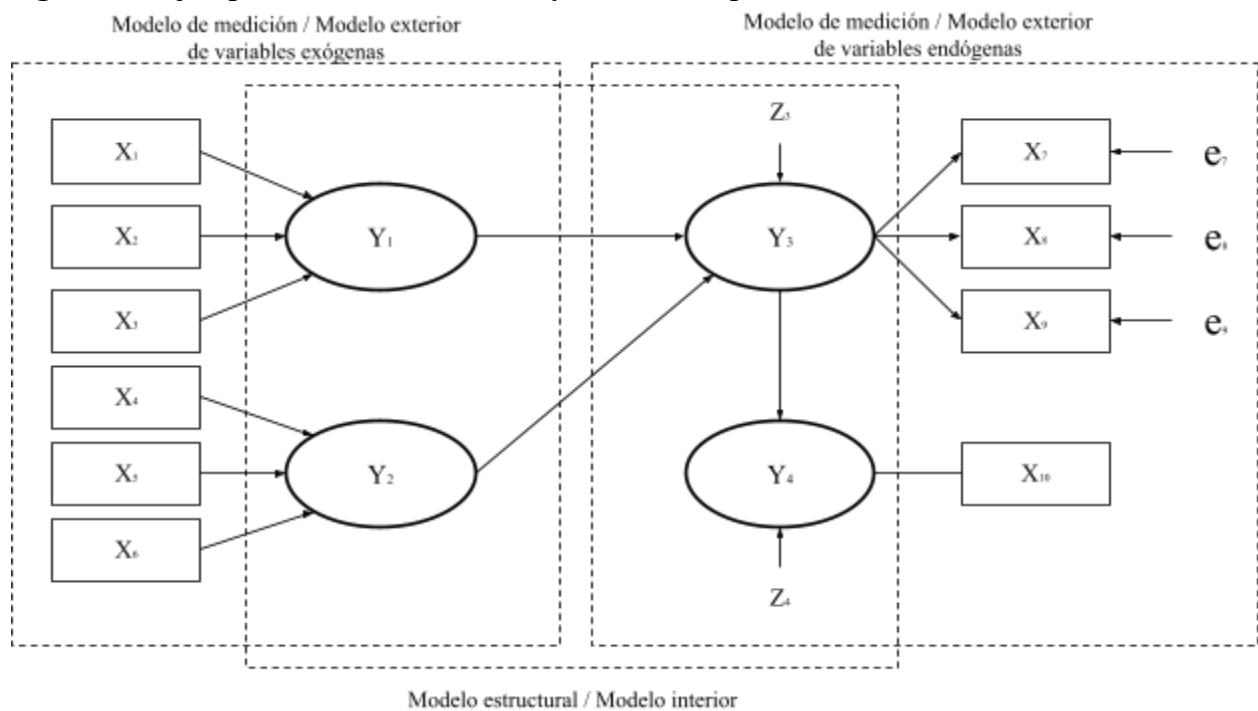
Un modelo PLS-SEM estima las cargas (*loadings*) de las variables indicadores para los constructos exógenos con base en las predicciones de los constructos endógenos, y no en la varianza compartida entre las variables indicadoras del mismo constructo, por lo que las cargas o *loadings* son, de alguna forma, su contribución a la trayectoria (*path*) de los coeficientes, los cuáles corresponden a los coeficientes beta en un análisis de regresión (Hair et al., 2011). Hair et al. (2011) describen dos componentes del algoritmo de un modelo de ecuaciones estructurales: el modelo de medición o modelo exterior y el modelo estructural o modelo interior. Los modelos de medición incluyen las relaciones unidireccionales predictivas entre cada constructo latente y sus indicadores asociados observados y no permite relaciones múltiples, por lo que las variables indicadoras son asociadas únicamente con un constructo latente. Como ya se mencionó, se va a trabajar con un modelo formativo, lo que significa que se asumen que los indicadores formativos son causantes de un constructo latente, por lo que cambios en los indicadores determinan los cambios en el constructo latente (Diamantopoulos y Winklhofer, 2001).

La evaluación del algoritmo PLS-SEM se realiza en dos etapas, de acuerdo con los componentes, modelo de medición o exterior y modelo estructural o interior, dónde dependiendo de la naturaleza de los constructos (formativo o reflectivo) se evalúan los indicadores que conforman las variables latentes o constructos (Hair et al., 2011). Un constructo formativo es aquel que indica una relación de causalidad entre los indicadores o variables observadas (preguntas o ítems) y el constructo o variable latente, mientras que un constructo reflectivo es aquel que indica el supuesto de que el constructo o variable latente causa la medida de los indicadores o variables observadas, los cuales, dada su naturaleza, tienen un término de error asociado con cada indicador, que no es el caso de los constructos formativos (Hair et al., 2023). Es importante mencionar que también existen constructos que se miden utilizando un solo indicador o ítem, por lo que la relación entre el constructo y el indicador es sin dirección (Hair et al., 2023).

La ilustración de las hipótesis y la relación de las variables en un modelo PLS-SEM se realiza utilizando modelos de trayectorias, que son diagramas en los que los constructos o variables latentes se representan utilizando círculos o elipses, mientras que los indicadores o variables manifiestas, los que contienen las variables observadas o datos brutos, se representan con rectángulos, mientras que las relaciones entre ellos se hace a través de flechas de una sola

punta indicando causalidad (Hair et al., 2023). En la figura 16 se representa un ejemplo de un modelo de trayectoria en el que se tienen cuatro constructos o variables latentes (Y_1 a Y_4), así como diez ítems o preguntas (X_1 a X_{10}), así como flechas que indican la relación entre ellos, para diferenciar entre un constructo formativo y uno reflectivo. En la misma figura podemos observar los dos componentes del modelo PLS-SEM, el modelo estructural o modelo interior (Y_1 a Y_4), que muestra la relación entre los constructos, y un modelo de medición o modelo exterior, que muestra la relación entre los constructos y los indicadores que lo componen (X_1 a X_{10}), diferenciando también entre estos últimos, para asociarlos a los constructos exógenos o endógenos (Hair et al., 2023). Asimismo, se observan los términos de error (e_7 a e_9), que representan la varianza no explicada al estimar el modelo y se asocian únicamente con los indicadores de constructos reflectivos (Hair et al., 2023).

Figura 16. Ejemplo de un modelo de trayectorias simple

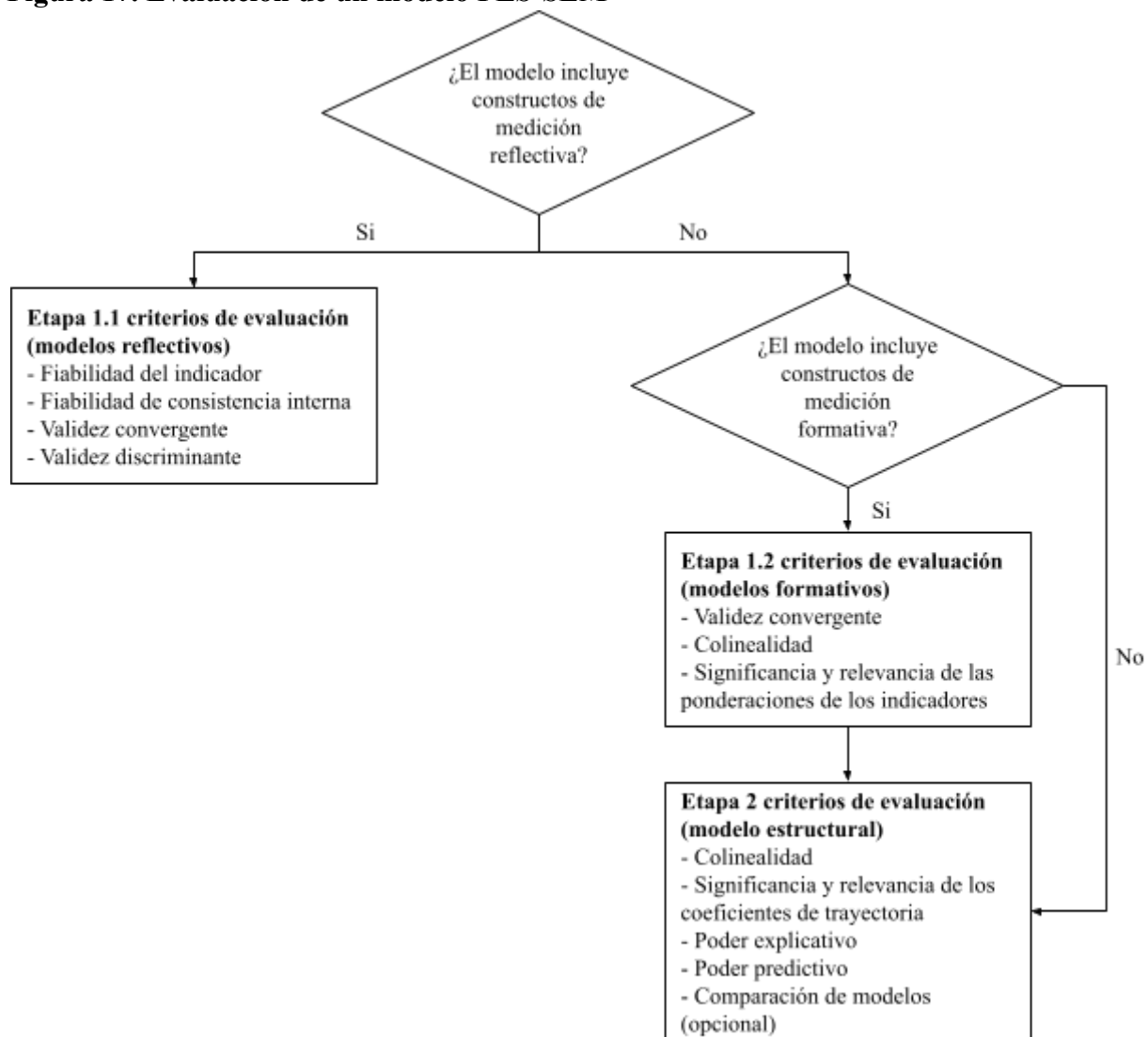


Fuente: elaboración propia con base en Hair et al. (2023).

Evaluación de un modelo PLS-SEM

Teniendo en cuenta lo anterior, la evaluación de un modelo PLS-SEM debe considerar la examinación de los constructos reflectivos y formativos dentro de la primera etapa para justificar la calidad de las mediciones o indicadores, y posteriormente evaluar las relaciones estructurales del modelo entre las variables latentes que representan las hipótesis propuestas (Hair et al., 2023). En la figura 17 se observan los lineamientos propuestos por Hair et al. (2023) para la evaluación del modelo PLS-SEM, considerando el tipo de constructos que se incluyen en el mismo: en caso de ser constructos reflectivos, primero se evalúa la fiabilidad de indicador, fiabilidad de consistencia interna, así como la validez convergente y discriminante; mientras que si son constructos formativos, primero se evalúa la validez convergente, colinealidad y significancia y relevancia de las ponderaciones de los indicadores; evaluando en ambos casos, como etapa dos, el modelo estructural a través de la colinealidad, significancia y relevancia de los coeficientes de trayectoria y el poder explicativo y predictivo del modelo.

Figura 17. Evaluación de un modelo PLS-SEM



Fuente: elaboración propia con base en Hair et al. (2023).

Como se mencionó anteriormente, la aplicación de un modelo PLS-SEM consta de la especificación del modelo de medición y del modelo estructural, por lo que su evaluación se divide consiste en evaluar de forma individual cada uno de los modelos. El primer paso consiste en revisar la confiabilidad y validez de las medidas, de acuerdo con ciertos criterios que dependen de si la especificación del modelo de medición es formativa o reflectiva (Hair et al., 2019). La lógica detrás del primer paso es que si no se tiene confianza de que las medidas representan los constructos de interés, no hay razón para evaluar las relaciones estructurales entre ellas. Una vez evaluadas las medidas de los constructos se procede con el segundo paso, en el que se evalúan las estimaciones del modelo estructural. En la tabla 7 se describen los criterios a considerar para la evaluación de los modelos de medición formativos y reflectivos, así como del modelo estructural, los cuales se describen en detalle en los párrafos subsecuentes.

Tabla 7. Criterios de evaluación del modelo PLS-SEM

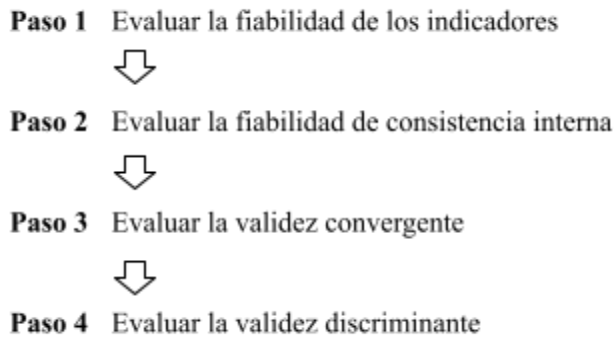
Modelo de medición reflectivo	
Cargas de indicadores reflectivos	≥ 0.708
Fiabilidad de consistencia interna	El alfa de Cronbach es el límite inferior, la fiabilidad compuesta es el límite superior de la fiabilidad de la consistencia interna. ρ_A suele situarse entre estos límites y puede servir como una buena representación de la fiabilidad de la consistencia interna de un constructo, suponiendo que el modelo factorial sea correcto Mínimo 0.70 (o 0.60 en investigación exploratoria) Máximo 0.95 para evitar redundancia de indicadores, lo que podría comprometer la validez del contenido Recomendado 0.70-0.90
Validez convergente	$AVE \geq 0.50$
Validez discriminante	Para constructos conceptualmente similares: $HTMT < 0.90$ Para constructos conceptualmente diferentes: $HTMT < 0.85$
Modelo de medición formativo	
Validez convergente (análisis de redundancia)	≥ 0.708 correlación
Colinealidad (Factor de Inflación de la Varianza o VIF)	Problemas de colinealidad probables (crítico) cuando $VIF \geq 5$ Posibles problemas de colinealidad cuando $VIF \geq 3-5$ Idealmente $VIF < 3$
Significancia estadística de los pesos	$p\text{-value} < 0.05$ o el intervalo de confianza al 95% no incluye cero
Relevancia de los indicadores con un peso significativo	Los pesos significativos más grandes son más relevantes (contribuyen más)
Relevancia de los indicadores con un peso no significativo	Las cargas ≥ 0.50 que son estadísticamente significativas son consideradas relevantes
Modelo estructural	
Colinealidad (Factor de Inflación de la Varianza o VIF)	Problemas de colinealidad probables (crítico) cuando $VIF \geq 5$ Posibles problemas de colinealidad cuando $VIF \geq 3-5$ Idealmente $VIF < 3$
Valor de R ²	Ajuste del modelo 0.75 → sustancial 0.50 → moderado 0.25 → débil > 0.90 → típicamente indica sobreajuste
Valor de Q ²	Precisión predictiva del modo de trayectoria PLS > 0 → pequeño > 0.25 → mediano > 0.50 → grande
Poder predictivo (PLSpredict)	Fijar K=10, asumiendo que el tamaño de la muestra es lo suficientemente grande $Q^2_{\text{predict}} > 0$ indica que el desempeño del modelo supera a la referencia más <i>naïve</i> (las medias de los indicadores de la muestra de análisis)

Fuente: elaboración propia con base en Hair et al. (2019).

Evaluación de modelos de medición reflectivos

Cuando el modelo PLS-SEM incluye constructos reflectivos, es necesario evaluar en términos de fiabilidad y validez, lo que incluye como primer paso la evaluación la fiabilidad de las medidas, a nivel de indicadores (fiabilidad de los indicadores) y a nivel de los constructos (fiabilidad de consistencia interna), para posteriormente evaluar la validez de convergencia de cada una de las medidas utilizando la varianza media extraída (AVE por sus siglas en inglés) y evaluar la validez discriminante comparándola con las medidas de otros constructos en el mismo modelo utilizando la razón heterotrazo-monotrazo (HTMT por sus siglas en inglés) (Hair et al., 2021).

Figura 18. Proceso de evaluación de un modelo PLS-SEM con constructos reflectivos



Fuente: elaboración propia con base en Hair et al. (2021).

En la figura 18 se puede observar el proceso de evaluación de un modelo con constructos reflectivos. En el primer paso, la evaluación de la fiabilidad de los indicadores, lo que se busca es conocer qué parte de la varianza de cada indicador se explica por su constructo, lo cuál es un indicativo de la fiabilidad del indicador y se obtiene al elevar al cuadrado la carga de indicador; en el segundo paso, la fiabilidad de consistencia interna, nos indica el grado en que los indicadores que miden el mismo constructo se asocian entre sí, medido a través del alfa de Cronbach (ecuación (1)) y el coeficiente de fiabilidad compuesta rho (ecuación (2)); la validez de convergencia se evalúa como tercer paso y nos dice en qué medida converge el constructo para explicar la varianza de los indicadores utilizando la varianza media extraída o AVE (ecuación (3)) para todos los indicadores en cada constructo; y por último, se evalúa la validez discriminante, que mide hasta qué punto un constructo es empíricamente distinto de otros constructos del modelo estructural y se recomienda la razón HTMT (ecuación (4)) para

evaluarlo, pues se define como el valor medio de las correlaciones de los indicadores entre constructos, en relación con la media geométrica de las correlaciones medias de los indicadores que miden el mismo constructo (Hair et al., 2021).

$$Cronbach's \alpha = \frac{K \cdot \bar{r}}{[1+(K-1) \cdot \bar{r}]} \quad (1)$$

Dónde:

K = número de indicadores del constructo

$\bar{r}$ = el coeficiente medio de correlación de indicadores no redundantes

$$\rho_c = \frac{\left(\sum_{k=1}^K l_k \right)^2}{\left(\sum_{k=1}^K l_k \right)^2 + \sum_{k=1}^K var(e_k)} \quad (2)$$

Dónde:

l_k = carga externa estandarizada de la variable indicadora k de un constructo específico medido con K indicadores

e_k = error de medida de la variable indicadora k

$var(e_k)$ = varianza del error de medida de la variable indicadora k

$$AVE = \frac{\left(\sum_{k=1}^K l_k^2 \right)}{K} \quad (3)$$

$$HTMT_{ij} = \frac{1}{K_i K_j} \sum_{g=1}^{K_i} \sum_{h=1}^{K_j} r_{ig,jh} \div \left(\frac{2}{K_i(K_i-1)} \cdot \sum_{g=1}^{K_i-1} \sum_{h=g+1}^{K_i} r_{ig,ih} \cdot \frac{2}{K_j(K_j-1)} \cdot \sum_{g=1}^{K_j-1} \sum_{h=g+1}^{K_j} r_{jg,jh} \right)^{\frac{1}{2}} \quad (4)$$

Dónde:

$r_{ig,jh}$ = correlación de los indicadores (dentro y a través de los modelos de medición de las variables latentes Y_i y Y_j)

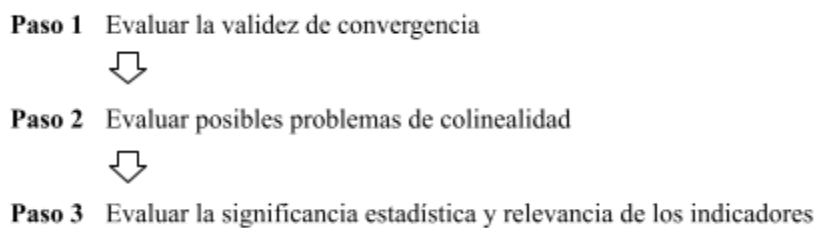
Evaluación de modelos de medición formativos

En caso de que el modelo incluya constructos de naturaleza formativa, el proceso de evaluación del modelo se compone de tres pasos: evaluar la validez de convergencia, en la que se evalúa el grado en que los constructos especificados como formativos se correlacionan con otras variables del mismo concepto medidos de forma reflectiva y se utiliza un análisis de redundancia de correlación; como paso dos se evalúa la colinealidad de los indicadores, puesto que una correlación alta incrementa el error estándar de los pesos de los indicadores, lo que puede error de tipo II (falsos negativos), y se utiliza el factor de inflación de la varianza (VIF por sus siglas en inglés) para evaluar la colinealidad de los indicadores (ecuación (5)); por último, en el paso tres se evalúa la significancia estadística y relevancia de los pesos de los indicadores, los cuales

resultan de un análisis de regresión entre el constructo formativo y los indicadores que lo componen y se evalúan asumiendo el intervalo de confianza y revisando los *t-values* utilizando como referencia una prueba de dos colas de una distribución normal (Hair et al., 2021). En la figura 19 se puede observar el proceso de evaluación de un modelo PLS-SEM con constructos formativos.

$$VIF_k = \frac{1}{1-R_k^2} \quad (5)$$

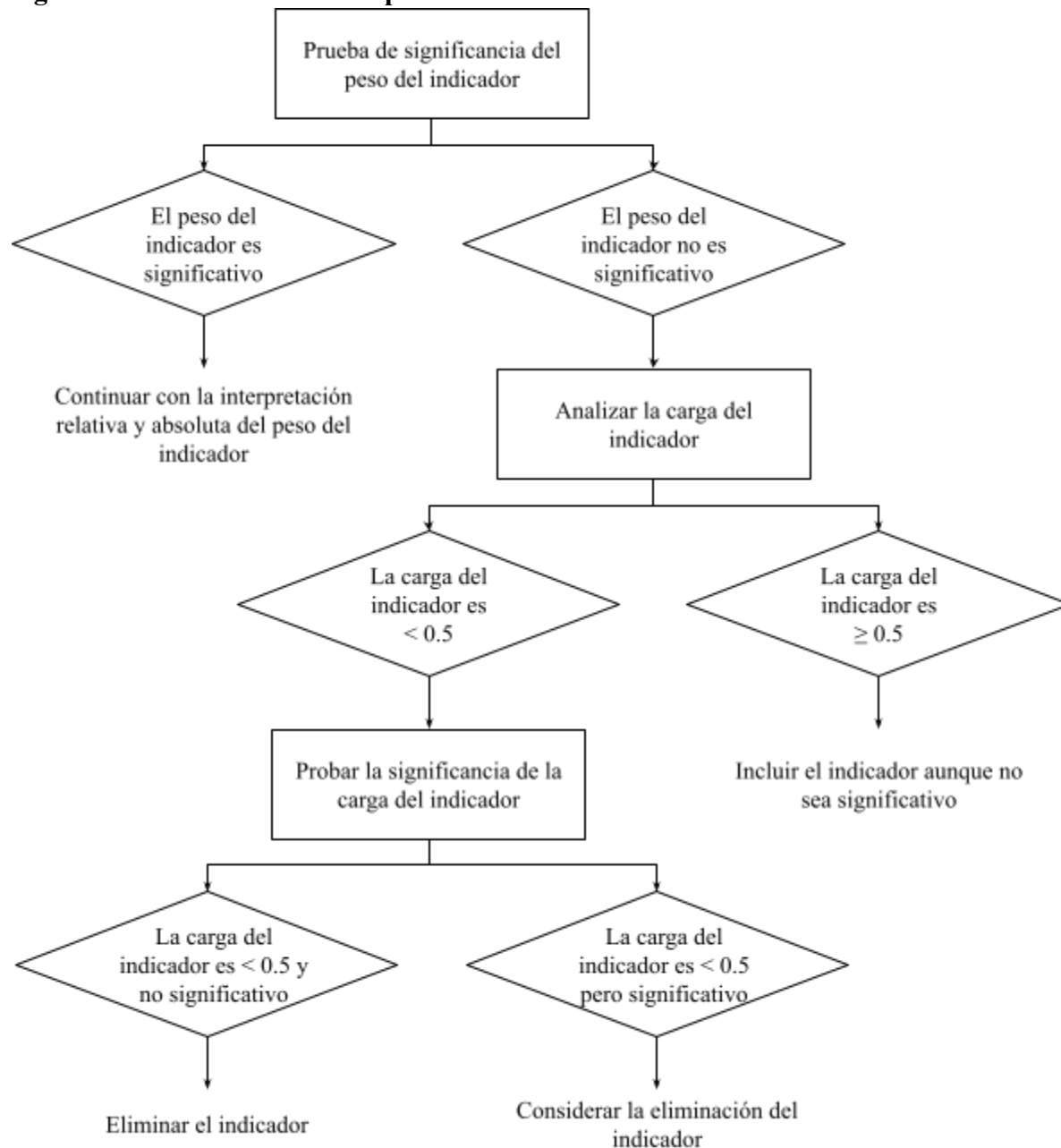
Figura 19. Proceso de evaluación de un modelo PLS-SEM con constructos formativos



Fuente: elaboración propia con base en Hair et al. (2021).

Si en el último paso alguno de los pesos de un indicador es no significativo, se recomienda considerar la contribución absoluta del indicador al constructo, antes de eliminarlo, lo cual se puede evaluar con la carga de indicador formativo, la cual debe ser estadísticamente significativa para poder considerar que contribuye de forma relativa al constructo que pertenece (Hair et al., 2021). En la figura 20 se ilustra el proceso de decisión para incluir o eliminar un indicador formativo, el cual consiste en probar si el peso del indicador es significativo, si lo es se toma el peso como su tamaño absoluto y relativo, en caso contrario, se debe analizar la carga del indicador, en caso de que ésta resulte menor a 0.5 se procede a eliminar el indicador, mientras que en el caso contrario se debe evaluar su significancia estadística y en caso de resultar no significativo se debe eliminar el indicador, mientras que si la carga del indicador es significativa solamente se debe considerar la inclusión del mismo con el debido soporte teórico (Hair et al., 2021).

Figura 20. Proceso de decisión para incluir o eliminar un indicador formativo



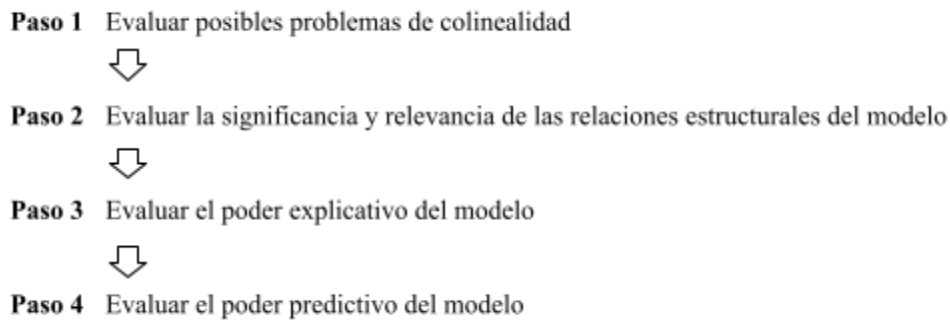
Fuente: elaboración propia con base en Hair et al. (2021).

Evaluación del modelo estructural

Una vez que se evalúan los modelos de medición para los constructos reflectivos y formativos, se procede a evaluar el modelo estructural. El proceso de evaluación del modelo estructural consiste de cuatro pasos: evaluar para identificar posibles problemas de colinealidad, evaluar la significancia y relevancia de las relaciones estructurales del modelo, evaluar el poder

explicativo del modelo, evaluar el poder predictivo del modelo y como un paso adicional opcional, realizar comparaciones entre modelo (Hair et al., 2021). En la figura 21 se ilustra el proceso de evaluación del modelo estructural hasta el paso cuatro, pues son los pasos pertinentes a la presente investigación.

Figura 21. Proceso de evaluación del modelo estructural



Fuente: elaboración propia con base en Hair et al. (2021).

En el primer paso, dado que el algoritmo PLS-SEM calcula los coeficientes para las relaciones entre los constructos estimando una serie de ecuaciones de regresión, es necesario evaluar posibles problemas de colinealidad que puedan afectar al modelo (Hair et al., 2021). Esto se realiza con un proceso similar al que se utiliza para evaluar los constructos formativos, con la diferencia de que las puntuaciones de los constructos exógenos en cada regresión del modelo estructural se utilizan para calcular los valores del factor de inflación de varianza o VIF (ecuación (5)), teniendo como referencia los mismos valores que en constructos formativos: valores de VIF mayores a 5 indican probables problemas de colinealidad, aunque también puede ocurrir con valores entre 3 y 5, lo cual se puede corregir creando constructos de segundo orden, es decir, constructos que se componen de otros constructos o constructos de primer orden (Hair et al., 2021).

Posteriormente, como segundo paso se debe evaluar la significancia y relevancia de las relaciones del modelo estructural o de los coeficientes de trayectoria (Hair et al., 2021). La significancia de las relaciones estructurales del modelo se evalúa a través de la significancia de los coeficientes de trayectoria, y se utilizan los errores estándar resultantes del proceso de *bootstrapping* como base para el cálculo de valores *t* de los coeficientes de trayectoria o intervalos de confianza alternos, donde, por ejemplo, un coeficiente es significativo a un nivel de

5% si el valor cero no cae en el intervalo de confianza de 95% (Hair et al., 2021). Mientras que la relevancia de los coeficientes de trayectoria, que normalmente se encuentran entre -1 y +1, dónde -1 indica una fuerte relación negativa, y cuando son cercanos a +1 indica una fuerte relación positiva, dejando fuera a coeficientes mayores que +/- 1 puesto que es un indicativo de problemas de colinealidad (Hair et al., 2021). Por ejemplo, un coeficiente de trayectoria de 0.5 indica que cuando el constructo predictor se incrementa en una unidad de desviación estándar, el constructo endógeno va a incrementar en 0.5 unidades de desviaciones estándar (Hair et al., 2021).

Por otra parte, el poder explicativo del modelo se mide a través del coeficiente de determinación o R^2 de los constructos endógenos, el cuál representa la varianza explicada en cada uno de los constructos endógenos y toma valores entre 0 y 1, dónde un valor más alto indica un mayor poder explicativo del modelo, teniendo siempre en cuenta el contexto teórico del modelo (Hair et al., 2021). Adicionalmente, se evalúan los cambios en el valor de R^2 cuando se omite alguno de los constructos exógenos del modelo, con el objetivo de medir si el constructo omitido tiene un impacto sustantivo en el constructo endógeno y se utiliza f^2 , o efecto de tamaño (ecuación (6)), dónde se estima el modelo incluyendo y excluyendo alguno de los constructos (Hair et al., 2021). Otra forma de medir la precisión predictiva del modelo de trayectorias es a través del valor Q^2 , el cual se basa en un procedimiento llamado *blindfolding*, que consiste en eliminar datos de la matriz de datos original para reemplazarlos con la media de la estimación del modelo de parámetros, combinando métodos de predicción dentro y fuera de la muestra, y el cuál indica mayor precisión predictiva del modelo de trayectorias entre menor sea la diferencia entre el valor original y el valor reemplazado (Hair et al., 2021).

$$f^2 = \frac{R^2_{incluido} - R^2_{excluido}}{1 - R^2_{incluido}} \quad (6)$$

Por último, para evaluar el poder predictivo del modelo sin utilizar la muestra empleada para estimar el modelo original, se utiliza el procedimiento PLSpredict, el cual involucra estimar el modelo en un análisis de muestra o entrenamiento para realizar predicciones fuera de la muestra, es decir una porción de la muestra que se reserva para realizar pruebas de predicción del modelo, para lo cual ejecuta una validación cruzada de k -pliegues o subgrupos de la muestra

total, dividiendo el conjunto total de datos aleatoriamente en k subconjuntos de datos de igual tamaño (Hair et al., 2021). Por ejemplo, en una estimación PLSpredict con $k = 5$, el proceso de validación cruzada divide el total de la muestra en 5 subgrupos de igual tamaño, para después combinar 4 de esos subgrupos en un sólo análisis que se utiliza para intentar predecir el subgrupo 5, repitiendo el proceso 5 veces, entrenando el modelo con 4 subgrupos y dejando a un subgrupo diferente como muestra para las pruebas de predicción (Hair et al., 2021). Para evaluar el modelo, se evalúa el grado de error de predicción del modelo, utilizando el error cuadrático medio (RMSE por sus siglas en inglés), la cual mide la raíz cuadrada de la media de las diferencias al cuadrado entre las predicciones del modelo y las observaciones reales, utilizando el modelo y un modelo de regresión lineal (ML), si para todos los indicadores el análisis PLS-SEM tiene menor RMSE comparado con el ML *naïve*, el modelo tiene un poder de predicción alto, si la mayoría o el mismo número de indicadores tiene menor RMSE en PLS-SEM comparado con el ML, hay un poder de predicción medio, mientras que si solamente una minoría de los indicadores arroja menor RMSE en PLS-SEM comparado con el ML, el poder predictivo del modelo es bajo.

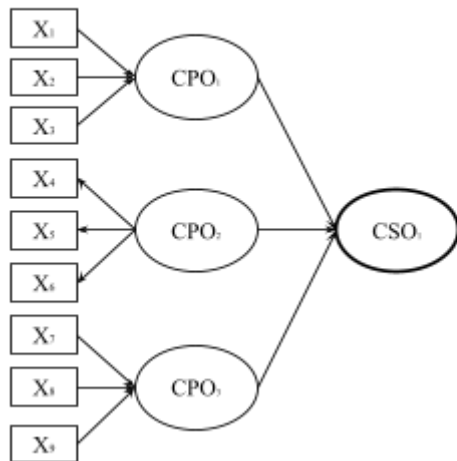
Modelos de segundo orden

Los modelos de segundo orden o modelos de componentes jerárquicos, son el marco de referencia para caracterizar constructos en una dimensión más abstracta (constructo o componente de segundo orden o de orden superior) y sus dimensiones más concretas (constructos o componentes de primer orden o de orden inferior), lo que permite reducir el número de relaciones en el modelo estructural, para lograr la parsimonia del modelo al resumir los constructos independientes en un constructo de orden superior, volviendo obsoletas las relaciones de los componentes de orden inferior con los constructos dependientes (Sarstedt et al., 2019). Para la utilización de constructos de segundo orden es necesario contar con el soporte teórico necesario para la especificación y medición de modelo, iniciando con la especificación del modelo de medida de los constructos de primer orden, para posteriormente definir la relación entre el constructo de segundo orden y los constructos de primer orden, pudiendo ser de carácter reflectivo o formativo; asimismo, se debe considerar el enfoque para identificar el constructo de segundo orden: enfoque de indicadores repetidos o el enfoque de dos etapas; y, por último, los componentes del constructo de segundo orden o constructos de primer orden deben ser evaluados

en términos de fiabilidad y validez de cada uno de los indicadores que lo conforman para posteriormente evaluar la validez discriminante del constructo de segundo orden completo (Sarstedt et al., 2019). En la figura 22 se muestra un ejemplo de un modelo con un constructo de segundo orden.

Figura 22. Modelo de segundo orden tipo II

Constructo de segundo orden tipo II: formativo-formativo

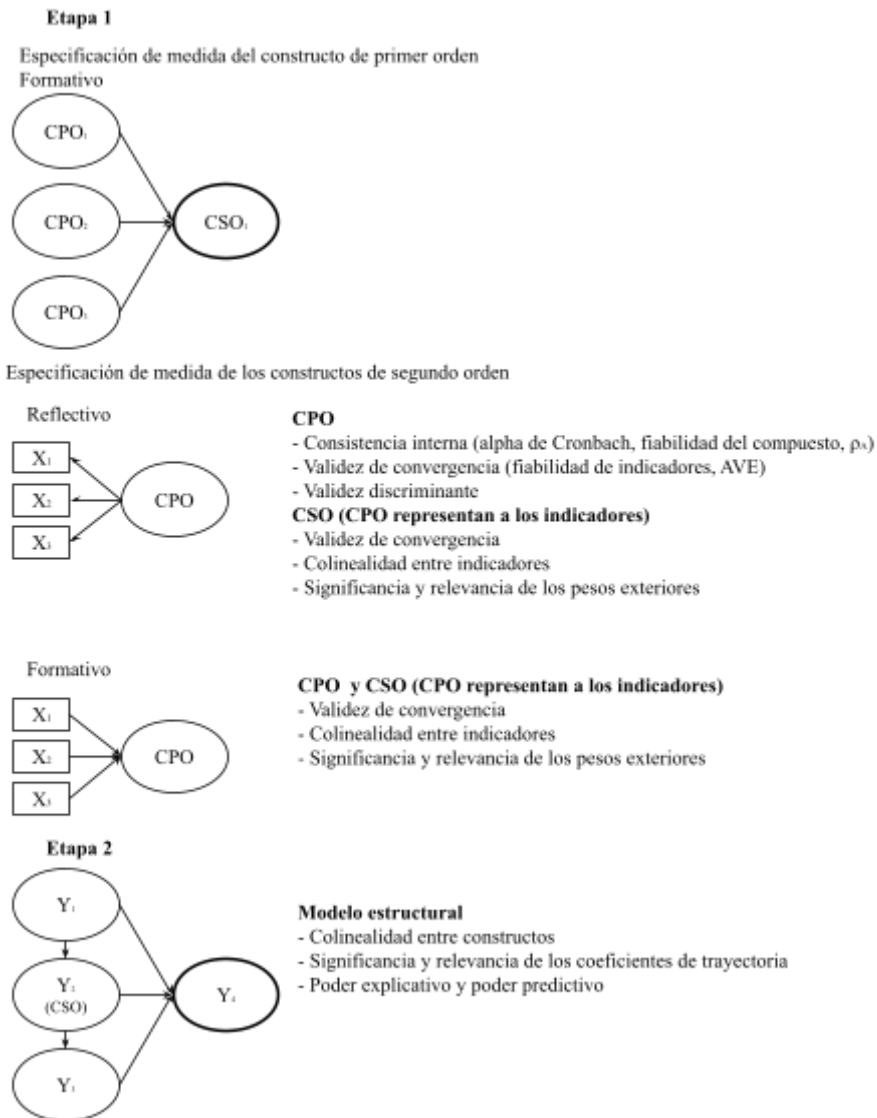


Fuente: elaboración propia con base en Sarstedt et al. (2019).

Teniendo en cuenta lo anterior y considerando el marco teórico de la presente investigación, las CABD se caracterizan como un constructo formativo de segundo orden compuesto por cinco constructos de primer orden: calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas. En ese sentido, se selecciona el enfoque de dos etapas (incrustado) para la estimación de las CABD, el que consiste en estimar primeramente los coeficientes de trayectoria de los constructos de primer orden agregando las puntuaciones como nuevas variables a la base de datos, para posteriormente utilizarlas como indicadores en el modelo de medición de los constructos de segundo orden (Sarstedt et al., 2019). Por otra parte, para la validación de los modelos con constructos de segundo orden, a pesar de que aplican los mismos criterios que en un análisis PLS-SEM convencional, se necesita evaluar también, como ya lo mencionamos, los modelos de medida de los constructos de primer orden que componen al constructo de segundo orden, así como el modelo de medida del constructo de segundo orden, antes de evaluar el modelo estructural (Sarstedt et al., 2019). En la figura 23 se muestra el proceso y criterios de evaluación de modelos con constructos de segundo orden u orden superior formativos, así como los criterios a evaluar

dependiendo de los constructos de primer orden, lo cual difiere si éstos son formativos o reflectivos.

Figura 23. Evaluación de resultados del modelo de medición para modelos con constructos de segundo orden (formativos) - Enfoque de dos etapas (*two-stage*)



Fuente: elaboración propia con base en Sarstedt et al. (2019).

Tamaño de la muestra y *Bootstrapping*

De acuerdo con los criterios para la aplicación del PLS-SEM definidos en el marco de referencia de Hair et al. (2011), las características de la muestra y de los datos establecen que el tamaño mínimo de la muestra debe ser mayor que alguno de los siguientes: 10 veces más grande que el

número de indicadores formativos utilizados para medir un constructo; ó 10 veces más grande que el número de trayectorias estructurales dirigidas a un constructo latente en particular dentro del modelo estructural. En el caso del modelo planteado, se plantea a las CABD como un constructo latente compuesto de 5 trayectorias estructurales o dimensiones (calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas), lo cual le convierte en el constructo con un mayor número de trayectorias y la referencia para determinar el tamaño de la muestra de investigación, por lo que las respuestas válidas al instrumento de recopilación de datos deben ser igual o mayor a 50, para cumplir con los requisitos de aplicación de un modelo PLS-SEM. Por otra parte, Ghasemaghaei (2018) utiliza como referencia el criterio de Roldán y Sánchez-Franco (2012), basado en el número de predictores incluidos en el modelo y esperando obtener un poder de 0.80 con un alpha de 0.05, se tiene que en el caso del modelo propuesto, el número mayor de predictores es de 5, por lo que para un efecto medio, se considera que 91 es un tamaño de muestra aceptable.

Otro criterio es el que proponen Hair et al. (2021), utilizando los niveles mínimos de los coeficientes de trayectoria y un poder de 80% (o 0.80), en los diferentes niveles de significancia (tabla 8), donde a menor nivel del valor de coeficientes de trayectoria, mayor debe ser el tamaño mínimo de la muestra, dependiendo también del nivel de significancia. Teniendo en cuenta lo anterior, tomamos como referencia el valor mínimo medio de los coeficientes de trayectoria, así como un nivel de significancia del 5%, para determinar que el tamaño mínimo de la muestra para aplicar el modelo PLS-SEM es de 69 observaciones válidas.

Tabla 8. Tamaños mínimos de muestra para diferentes niveles de coeficientes de trayectoria (p_{min}) y un poder de 80%

p_{min}	Nivel de significancia		
	1%	5%	10%
0.05-0.1	1004	619	451
0.11-0.2	251	155	113
0.21-0.3	112	69	51
0.31-0.4	63	39	29
0.41-0.5	41	25	19

Fuente: elaboración propia con base en Hair et al. (2021).

Adicional a determinar el tamaño mínimo de la muestra, es importante mencionar que la metodología PLS-SEM no supone que los datos están distribuidos normalmente, por lo que se

utiliza *bootstrapping* no paramétrico, un proceso que involucra muestreo aleatorio repetido con reemplazo de la muestra original para crear una muestra *bootstrap* para obtener los errores estándar necesarios para las pruebas de hipótesis, asumiendo que la distribución de la muestra es una representación razonable de la distribución prevista de la población, permitiendo probar la significancia de los coeficientes estimados en el modelo PLS-SEM (Hair et al., 2011). Tomando nuevamente la definición del procedimiento de *bootstrapping* descrita por Hair et al. (2011), una vez obtenida la muestra de cuestionarios válidos, se procederá a crear 10,000 muestras de *bootstrap* extrayendo casos al azar con reemplazo de la muestra original. Cada una de las muestras *bootstrap* debe tener el mismo número de casos que la muestra original, para que el algoritmo PLS-SEM pueda estimar los resultados de cada una de las muestras. Las estimaciones de los parámetros repetidos por *bootstrap* se utilizan para construir una distribución de muestreo empírica para cada parámetro del modelo, tomando la desviación estándar de dichas distribuciones como una aproximación para el error estándar empírico del parámetro. Los coeficientes de trayectoria obtenidos del modelo forman una distribución *bootstrap*, la cual representa una aproximación de la distribución de muestreo, mientras que los resultados PLS-SEM de las muestras *bootstrap* nos dan como resultado el error estándar para uno de los coeficientes de trayectoria del modelo, lo que nos permite realizar una prueba *t* de Student para medir la significancia de las relaciones de trayectoria del modelo. La hipótesis nula establece que el coeficiente es igual a cero, contra la hipótesis alterna que establece que el coeficiente es diferente de cero y su efectividad depende de la representatividad de la muestra en términos de la población objetivo.

Datos

Los Modelos de Ecuaciones Estructurales son utilizados para medir modelos de relaciones complejas teniendo en cuenta el error de medición inherente a los indicadores utilizando un método de análisis multivariante de datos de segunda generación, siendo las variables constructos compuestos de una o más variables en forma de indicador, es decir, variables compuestas por variables (Hair et al., 2021). Considerando lo anterior, para utilizar dicha metodología es necesario definir las variables a utilizar, las relaciones entre ellas, así como las variables que van a componer dichas variables, con el debido soporte teórico (Hair et al., 2021),

por lo que la definición de las variables descrita en el capítulo de marco teórico puede ser considerado como el primer paso de nuestra metodología.

El siguiente paso consiste en definir las variables o indicadores que van a componer a las variables principales del modelo, utilizando como referencia la literatura revisada, plasmándolo en un instrumento de recolección de datos con formato de cuestionario y utilizando la escala Likert³ de siete puntos para las respuestas. Retomando las definiciones del marco teórico sobre nuestras variables, tenemos que en las tres hipótesis planteadas se propone a las cinco dimensiones de las capacidades de ABD (calidad de los datos, grandeza de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas), como esenciales, pues cada una de ellas contribuye a crear y fortalecer dichas capacidades, de acuerdo con el marco de referencia de recursos clave basados en TI de Bharadwaj (2000).

En ese sentido, se define y se operacionaliza el concepto de capacidades de ABD como un índice multidimensional y formativo. Esto significa que las capacidades de ABD son un constructo de orden superior (o segundo orden), que está representado o constituido por sus dimensiones (es decir, constructos de orden inferior (o primer orden)) (Ghasemaghaei et al., 2019 y Pavlou y El Sawy, 2006). Dichos constructos se modelan como reflexivos y formativos respectivamente (Becker et al., 2012). Mientras que en el primer caso, el concepto general se “manifiesta” por sus dimensiones, en el último caso, el constructo formativo es un constructo agregado que es un compuesto de sus dimensiones (Edwards, 2001). En otras palabras, las dimensiones se combinan para producir el constructo (Edwards, 2001; Wetzels et al., 2009), mientras que las dimensiones se componen de las preguntas o ítems del instrumento de recolección de datos.

A continuación se describe el proceso para diseñar el instrumento de recolección de datos siguiendo los lineamientos de la literatura pertinente revisada, así como las etapas de prueba de dicho instrumento, selección de la muestra y aplicación del mismo.

³ La escala autoreportada de Likert es uno de los métodos de medición más utilizados en las ciencias sociales pues provee una forma conveniente de medir constructos no observables, que va desde definir el constructo y las preguntas o ítems hasta la creación de la escala (Jebb et al., 2021). En el caso del instrumento de recolección planteado, las preguntas se responden a través de siete opciones o una escala de Likert de siete puntos que son: 1) Totalmente en desacuerdo, 2) En desacuerdo, 3) Algo en desacuerdo, 4) Ni de acuerdo ni en desacuerdo, 5) Algo de acuerdo, 6) De acuerdo y 7) Totalmente de acuerdo.

Diseño del instrumento de recolección de datos

Tras una revisión exhaustiva de fuentes públicas y privadas sobre información pertinente al uso de analíticos de datos para la toma de decisiones en las empresas, de cualquier tamaño e industria, en Tijuana, Baja California, se concluyó pertinente el diseño y aplicación de un instrumento de recolección de datos, aunado a la especificidad de los lineamientos para la aplicación de la metodología encontrados en la literatura existente sobre el tema. En ese sentido, como ya se mencionó, el primer paso de acuerdo con las recomendaciones de Moore y Benbasat (1991), fue revisar literatura pertinente al impacto de los analíticos de datos en el proceso de toma de decisiones y el desempeño financiero de la firma, para identificar los ítems o preguntas con los que se puede formar cada uno de los constructos, incluyendo preguntas sobre las características de la muestra que nos pueden servir como estadística descriptiva de la misma. En ese sentido, después de una revisión exhaustiva de los trabajos presentados en el estado del arte y marco teórico del presente documento, se tomó como referencia a Ghasemaghaei et al. (2018), Côte Real et al. (2017), Ferraris et al. (2019), Jarupathirun (2007), Tippins y Sochi (2003), Wang et al. (1996) y Bassellier y Benbasat (2004) para diseñar un instrumento de recolección de datos.

Recapitulando secciones anteriores, tenemos que las CABD de la firma son la variable dependiente en el modelo y se componen de cinco dimensiones: calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas; lo cuales a su vez se componen de seis, tres, tres, cuatro y seis preguntas o ítems, respectivamente. A su vez, el desempeño del proceso de toma de decisiones se caracteriza como un constructo que se compone de la calidad y la eficiencia de las decisiones, dimensiones que se conforman por seis y dos ítems, respectivamente. Por último, el desempeño financiero de la empresa es un constructo de primer orden integrado por dos preguntas, una relacionada con la rentabilidad de la firma medida a través de la utilidad neta, y otra relacionada con la rentabilidad medida a través de las ventas de la firma.

Tabla 9. Instrumento de recolección de datos

Variable	Constructo	Pregunta o indicador	Fuente
Características de la muestra	Tomador de decisiones	Seleccione el grupo de edad al que pertenece	Ghasemaghaei et al. (2018) y Côté Real et al. (2017)
		Sexo	
	Organización	¿Cuál es el puesto que desempeña en la organización a la que pertenece?	
		Nombre comercial de la empresa	
		Años de operación de la empresa	
		Tamaño de la organización a la que pertenece (número de empleados)	
		Industria en la que opera la empresa	
		¿En qué porcentaje se incrementó la rentabilidad financiera (utilidad neta) de la empresa respecto al año pasado?	
		¿Cuál es el porcentaje de incremento en las ventas de la empresa respecto al año pasado?	
		¿La empresa utiliza Analíticos de Datos para la toma de decisiones?	
		En caso de no utilizar Analíticos de Datos, ¿cuál es la razón?	
		En caso de sí utilizar Analíticos de Datos, ¿desde hace cuántos años se utilizan?	
		En su organización ¿se han reducido los costos con el uso de Analíticos de Datos para la toma de decisiones?	
		En su organización ¿se ha incrementado el margen de utilidad con el uso de Analíticos de Datos para la toma de decisiones?	
		En su organización ¿considera que las Capacidades de Analíticos de Big Data se encuentran desarrolladas?	
Desempeño de la organización	Desempeño financiero de la organización	En su organización ¿se ha incrementado el desempeño financiero con el uso de Analíticos de Datos para la toma de decisiones?	Ferraris et al. (2019)
		La utilidad neta de la empresa mejoró desde la implementación de Analíticos de Datos en el proceso de toma de decisiones.	
		¿En qué porcentaje se incrementó la utilidad neta desde la implementación de Analíticos de Datos para la toma de decisiones?	
		Las ventas de la empresa incrementaron desde la implementación de Análisis de Datos en el proceso de toma de decisiones.	
		¿Cuál es el porcentaje de incremento de las ventas desde la implementación de Analíticos de Datos para la toma de decisiones?	
Desempeño del proceso de toma de decisiones	Desempeño del proceso de toma de decisiones	En nuestra organización existe un buen desempeño del proceso de toma de decisiones.	Jarupathirun (2007)

Variable	Constructo	Pregunta o indicador	Fuente
Capacidades de ABD	Calidad de las decisiones	En nuestra organización, los resultados de las decisiones son a menudo exactos.	Jarupathirun (2007)
		En nuestra organización, los resultados de las decisiones son a menudo correctos.	
		En nuestra organización, los resultados de las decisiones son a menudo precisos.	
		En nuestra organización, los resultados de las decisiones son a menudo perfectos.	
		En nuestra organización, los resultados de las decisiones son a menudo sin errores.	
		En nuestra organización, los resultados de las decisiones son a menudo confiables.	
	Eficiencia de las decisiones	En nuestra organización, el tiempo para llegar a decisiones es rápido.	
		En nuestra organización, la velocidad para llegar a decisiones es alta.	
	Habilidades Analíticas	Nuestros usuarios de analíticos de datos son expertos en lo que respecta a utilizar las herramientas de análisis de datos.	Tippins y Sochi (2003)
		Nuestros usuarios de analíticos de datos poseen un alto grado de experiencia en análisis de datos.	
		Nuestros usuarios de analíticos de datos son hábiles en el uso de herramientas de análisis de datos.	
Capacidades de ABD	Calidad de los datos	En nuestra organización, la calidad de los datos utilizados para construir Analíticos de Datos es muy buena.	Wang et al. (1996)
		En nuestra organización, los datos utilizados en analíticos de datos son confiables.	
		En nuestra organización, los datos utilizados en analíticos de datos tienen el nivel apropiado de detalle.	
		En nuestra organización, los datos utilizados en analíticos de datos son seguros.	
		En nuestra organización, los datos utilizados en analíticos de datos son oportunos.	
		En nuestra organización, los datos utilizados en analíticos de datos son relevantes para la tarea en cuestión.	
		En nuestra organización, los datos utilizados en analíticos de datos son precisos.	
	Conocimiento del dominio	En nuestra organización, hay un alto nivel de conocimiento del entorno externo (e.g. gobierno, competidores, proveedores y clientes).	
		En nuestra organización, hay un alto nivel de conocimiento de las metas y objetivos organizacionales.	Bassellier y Benbasat (2004)
		En nuestra organización, hay un alto nivel de conocimiento de las capacidades básicas de la organización.	
		En nuestra organización, hay un alto nivel de conocimiento de los factores clave que deben ir bien para que la organización tenga éxito.	

Variable	Constructo	Pregunta o indicador	Fuente
	Tamaño de los datos	Nuestra organización recopila y analiza grandes volúmenes de datos.	Ghasemaghaei et al. (2018)
		Nuestra organización procesa grandes volúmenes de datos.	
		Nuestra organización procesa datos en tiempo real.	
		Nuestra organización procesa diferentes tipos de datos.	
	Sofisticación de las herramientas	En nuestra organización, se utilizan las herramientas de análisis de datos apropiadas de acuerdo a nuestras necesidades.	Ghasemaghaei et al. (2018)
		En nuestra organización, utilizamos herramientas que brindan capacidades de procesamiento y recuperación de información.	
		En nuestra organización, utilizamos herramientas que realizan modelado y simulación.	
		En nuestra organización, utilizamos herramientas que realizan analíticos de lenguaje natural (extracción de información de fuentes no estructuradas como redes sociales).	
		En nuestra organización, utilizamos herramientas que proporcionan información en tiempo real.	
		En nuestra organización, utilizamos herramientas que identifican problemas.	
		En nuestra organización, utilizamos herramientas que evalúan diferentes alternativas.	

Fuente: elaboración propia con base en Ghasemaghaei et al. (2018), Côté Real et al. (2017), Ferraris et al. (2019), Jarupathirun (2007), Tippins y Sochi (2003), Wang et al. (1996) y Bassellier y Benbasat (2004).

Como se puede observar en la tabla anterior, cada una de los constructos se toma de la literatura existente: el desempeño de la organización se mide como la rentabilidad o utilidad neta y con las ventas, contando con dos ítems para cada constructo (Ferraris et al., 2019); el desempeño de la toma de decisiones se mide como un constructo formativo de segundo orden utilizando seis y dos preguntas con escalas reflectivas para medir la calidad y eficiencia de las decisiones, respectivamente, los cuáles se toma de Jarupathirun (2007); mientras que para la variable de capacidades de ABD, la calidad de los datos se mide como un constructo formativo que consta de seis preguntas (Wang et al., 1996), dominio del conocimiento y habilidades analíticas se miden como constructos reflectivos utilizando tres preguntas para dominio del conocimiento tomando como base a Tippins y Sohi (2003) y cuatro preguntas para habilidades analíticas con base en Jarupathirun (2007).

La primera parte del instrumento se centra en las características de la muestra. Específicamente, las características de la organización, tales como tamaño, industria en la que opera, el uso de ABD, tiempo de su implementación y si ésta ha contribuido o no a la reducción de costos o incremento de rentabilidad y ventas (Ghasemaghaei, 2018). En el caso del tamaño de la industria se utilizó la clasificación para empresas del sector industria de la Secretaría de Economía publicada en el Diario Oficial de la Federación (DOF) el 30 de junio de 2009, la cuál sigue vigente y se muestra en la tabla 10.

Tabla 10. Estratificación de empresas de acuerdo al número de trabajadores

Tamaño	Sector	Rango de número de trabajadores
Micro	Todas	Hasta 10
Pequeña	Industria y servicios	Desde 11 hasta 50
Mediana	Industria	Desde 51 hasta 250
Grande	Industria	Desde 251 en adelante

Fuente: elaboración propia con información del DOF (2009), recuperado de https://www.economia.gob.mx/files/marco_normativo/A539.pdf el 22 de abril de 2022.

Mientras que para la industria en la que opera la firma, se utilizó la clasificación del Sistema de Clasificación Industrial de América del Norte en su versión 2018 (SCIAN, 2018) publicada por el INEGI, el cual tiene por objetivo “proporcionar un marco único, consistente y actualizado para la recopilación, análisis y presentación de las estadísticas de tipo económico, que refleje la estructura de la economía mexicana” (INEGI, 2023). Las opciones de respuesta a la pregunta sobre la industria en la que opera la firma son los sectores de actividad económica enlistados en la tabla 11.

Tabla 11. Sectores de actividad económica (SCIAN, 2018)

Sector	Nombre
11	Agricultura, cría y explotación de animales, aprovechamiento forestal, pesca y caza
21	Minería
22	Generación, transmisión, distribución y comercialización de energía eléctrica, suministro de agua y de gas natural por ductos al consumidor final
23	Construcción
31-33	Industrias manufactureras
43	Comercio al por mayor
46	Comercio al por menor
48-49	Transportes, correos y almacenamiento
51	Información en medios masivos
52	Servicios financieros y de seguros
53	Servicios inmobiliarios y de alquiler de bienes muebles e intangibles
54	Servicios profesionales, científicos y técnicos
55	Corporativos
56	Servicios de apoyo a los negocios y manejo de residuos, y servicios de remediación
61	Servicios educativos
62	Servicios de salud y de asistencia social
71	Servicios de esparcimiento culturales y deportivos, y otros servicios recreativos
72	Servicios de alojamiento temporal y de preparación de alimentos y bebidas
81	Otros servicios excepto actividades gubernamentales

Fuente: elaboración propia con información del SCIAN 2018 (INEGI, 2023).

Asimismo, el primer apartado del instrumento de recolección de datos, también se centra en las características sobre el tomador de decisiones de la empresa que responde el cuestionario, recopilando información sobre el grupo etario al que pertenece, sexo y puesto que tiene dentro de la organización, buscando que el perfil del respondiente sea que ocupe una posición que conozca la operación de la organización, las estrategias, capacidades de TI y datos, mercado en el que compiten, así como acceso a información financiera de la misma (Côte-Real et al., 2017).

En la siguiente sección, nos enfocamos en las dos variables dependientes: desempeño financiero de la firma y desempeño del proceso de toma de decisiones. En ese sentido, las preguntas van orientadas a la mejoría de la utilidad neta y de las ventas, desde la implementación de analíticos de datos para la toma de decisiones para medir justamente el desempeño financiero,

mientras que la evaluación del desempeño del proceso de toma de decisiones se centra en la percepción del respondiente sobre la calidad y eficiencia de las decisiones. En todos los casos, las preguntas son planteadas con opción múltiple, utilizando una escala de Likert de siete puntos que va de totalmente en desacuerdo a totalmente de acuerdo, con el objetivo de estandarizar las respuestas que no se van a utilizar como estadística descriptiva (Côte-Real et al., 2017).

En la última sección nos concentramos en recopilar información sobre las CABD, es decir, nuestra variable independiente, la cual cómo ya mencionamos se compone de cinco dimensiones, los que a su vez se componen de ítems o preguntas. Para identificar las habilidades analíticas de la firma, las preguntas se orientan en conocer el nivel de *expertise* de los usuarios de las herramientas de análisis de datos, su experiencia y sus habilidades. En el caso de la calidad de los datos, las preguntas se enfocan en la confiabilidad, detalle, oportunidad, relevancia y precisión de los datos utilizados para construir analíticos de datos para la toma de decisiones. En lo que respecta al conocimiento de dominio, el instrumento se enfoca en conocer el nivel de conocimiento de los usuarios de datos y miembros de la organización sobre el negocio de la misma, sobre el entorno en el que opera la firma, las metas y objetivos organizacionales, las capacidades básicas de la empresa y los factores clave para lograr el éxito. Para incorporar otra de las dimensiones de las CABD, en el sentido de manejar un volumen grande de datos para que sean considerados analíticos de *Big Data*, las preguntas se orientan justamente a la grandeza de los datos, el momento en el que se procesan y los diferentes tipos de datos que se utilizan. Por último, se debe considerar la dimensión relacionada con la sofisticación de las herramientas, investigando sobre el tipo de herramientas que se disponen para el análisis de los datos recopilados, por ejemplo capacidades de procesamiento, tipos de modelos, su orientación a la resolución de problemas o a la evaluación de diferentes alternativas. Al igual que con el resto de las variables y para efectos de estandarizar las respuestas, todas las preguntas se podían responder mediante una escala de Likert de siete puntos que va de totalmente en desacuerdo a totalmente de acuerdo.

Una vez diseñado el instrumento de recolección de información sobre cada una de las variables a utilizar en el modelo con base en la literatura revisada, se procedió a proporcionar una primera versión a investigadores y académicos en la línea de investigación, expertos en el uso de datos para la toma de decisiones, así como profesionales en el diseño de cuestionarios y

encuestas para la recolección de datos con el objetivo de recibir retroalimentación, siguiendo con los lineamientos de Brislin (1970). Después de una serie de revisiones a cada una de las preguntas para evaluar su contenido, alcance y propósito, se modificaron algunas secciones del cuestionario, principalmente las relacionadas con las características de la muestra, y estuvieron orientadas a incluir más preguntas para extraer estadística descriptiva sobre la muestra y no en modificar la composición de los constructos o dimensiones. En la tabla 9 se muestra la versión final del instrumento de recolección de datos utilizada tanto en la prueba piloto como en la recopilación de los datos a utilizar en el modelo empírico. La versión final del cuestionario fue desarrollada en *Google Forms*, un software utilizado para la administración de formularios, cuestionarios y encuestas, cuyos resultados se despliegan en una hoja de cálculo, proporcionando también un resumen de los resultados en gráficos, facilitando su análisis, la cuál forma parte de la plataforma web gratuita *Google Workspace* (Google, s.f.).

Muestra y recolección de datos

Como ya se mencionó, el objetivo de investigación es analizar el impacto de las CABD en el proceso de toma de decisiones y el desempeño financiero de las empresas en Tijuana, Baja California, independientemente de la industria a la que pertenecen o el tamaño de las mismas, siendo éstas últimas el objeto de estudio, mientras que la ciudad de Tijuana, nuestro horizonte espacial. En ese sentido, con el objetivo de obtener información de las firmas en la ciudad de Tijuana, el instrumento se debe aplicar a tomadores de decisiones en dichas empresas, siempre que sean tomadores de decisiones con un conocimiento profundo sobre la estrategia de la organización, los recursos de analíticos de datos de los que dispone, así como los resultados financieros de la misma. Teniendo en cuenta lo anterior, el instrumento se aplicó a tomadores de decisiones de empresas en la ciudad de Tijuana, Baja California de diversas industrias o sectores de actividad económica con los siguientes puestos: director general, director administrativo, director de TI, dirección de servicio al cliente, dirección comercial, dirección de finanzas, gerente de producción, gerente de operaciones, presidencia o vicepresidencia y, en general, puestos a nivel directivo con la capacidad de responder en su totalidad el instrumento de recolección de datos.

Una vez definido el objeto de estudios y el perfil de los respondientes, se procedió a realizar una prueba piloto para garantizar que las preguntas fuesen claras para los respondientes y que no hubiese problemas con la redacción o la escala de medida utilizada. Después de la prueba piloto, se recibió retroalimentación respecto a algunos conceptos técnicos utilizados en la redacción de varias preguntas, así como poca claridad respecto a las secciones del formulario electrónico. Teniendo en cuenta lo anterior, se agregó una introducción sobre el objetivo del cuestionario para dar formalidad al documento, asimismo, se dividió el formulario en secciones para facilitar la comprensión del respondiente sobre el objetivo de las preguntas, y finalmente, se agregaron definiciones clave en algunas preguntas en las que los conceptos no eran del todo comprensibles. En la tabla 12 se enlistan los cambios realizados al instrumento de recolección de datos, después de recibir retroalimentación de la prueba piloto.

Tabla 12. Modificaciones a instrumento de recolección de datos

Sección	Pregunta	Modificaciones
		Investigación académica sobre el uso de datos para la toma de decisiones
		Se le invita a participar en esta investigación completando la siguiente encuesta.
		El objetivo de esta investigación es conocer el impacto de utilizar analíticos de datos para la toma de decisiones en el desempeño financiero de las empresas en Baja California. El siguiente cuestionario requerirá aproximadamente de 5 a 10 minutos para completarlo.
1	Introducción	Gracias por tomarse su tiempo para ayudarme a responder esta encuesta. Bajo ninguna circunstancia está obligado a responder ninguna de las preguntas, sin embargo, hacerlo me será de gran ayuda para completar mi investigación y mejorar la comprensión de este enfoque de investigación.
		Los datos recopilados se mantendrán confidenciales y se utilizarán únicamente con fines académicos sin mencionar nombres y utilizando resúmenes estadísticos para su presentación. Asimismo, se compartirá la presentación de los resultados con todos los participantes una vez concluida.
		Doctorando en Economía Sara Valenzuela correoelectrónico@uabc.edu.mx Universidad Autónoma de Baja California - Facultad de Economía y Relaciones Internacionales Tutor de tesis: Dr. Natanael Ramírez Ángulo
2	Introducción	Datos del respondiente Instrucciones: favor de responder a cada una de las preguntas seleccionando alguna de las opciones o cómo se indique
3	Introducción	Información sobre la empresa

Sección	Pregunta	Modificaciones
		Instrucciones: favor de responder a cada una de las preguntas seleccionando alguna de las opciones o cómo se indique
	1	Nombre comercial de la empresa Información confidencial a utilizar únicamente por el investigador
	2	Años de operación de la empresa Indique el número de años que tiene en operación la empresa
	3	Tamaño de la organización a la que pertenece (número de empleados) Seleccione el rango que corresponda al número de empleados de la empresa
	4	Industria en la que opera la empresa Seleccione la opción que enliste la industria en la que se clasifica la empresa de acuerdo con Sistema de Clasificación Industrial de América del Norte (SCIAN, 2018)
	7	¿La empresa utiliza Analíticos de Datos para la toma de decisiones? Entiéndase por Analíticos de Datos todos aquellos indicadores contruidos a partir del análisis de datos generados dentro o fuera de la organización
	12	En su organización ¿considera que las Capacidades de Analíticos de Big Data se encuentran desarrolladas? Entiéndase por Capacidades de Analíticos de Big Data como los recursos de analíticos de datos para realizar un análisis de datos riguroso y accionable
4	Introducción	Desempeño de la organización Instrucciones: seleccionar la opción que mejor aplique en el caso de su empresa
5	Introducción	Toma de decisiones Instrucciones: seleccionar la opción que mejor aplique en el caso de su empresa
6	Introducción	Habilidades analíticas y calidad de los datos Instrucciones: seleccionar la opción que mejor aplique en el caso de su empresa

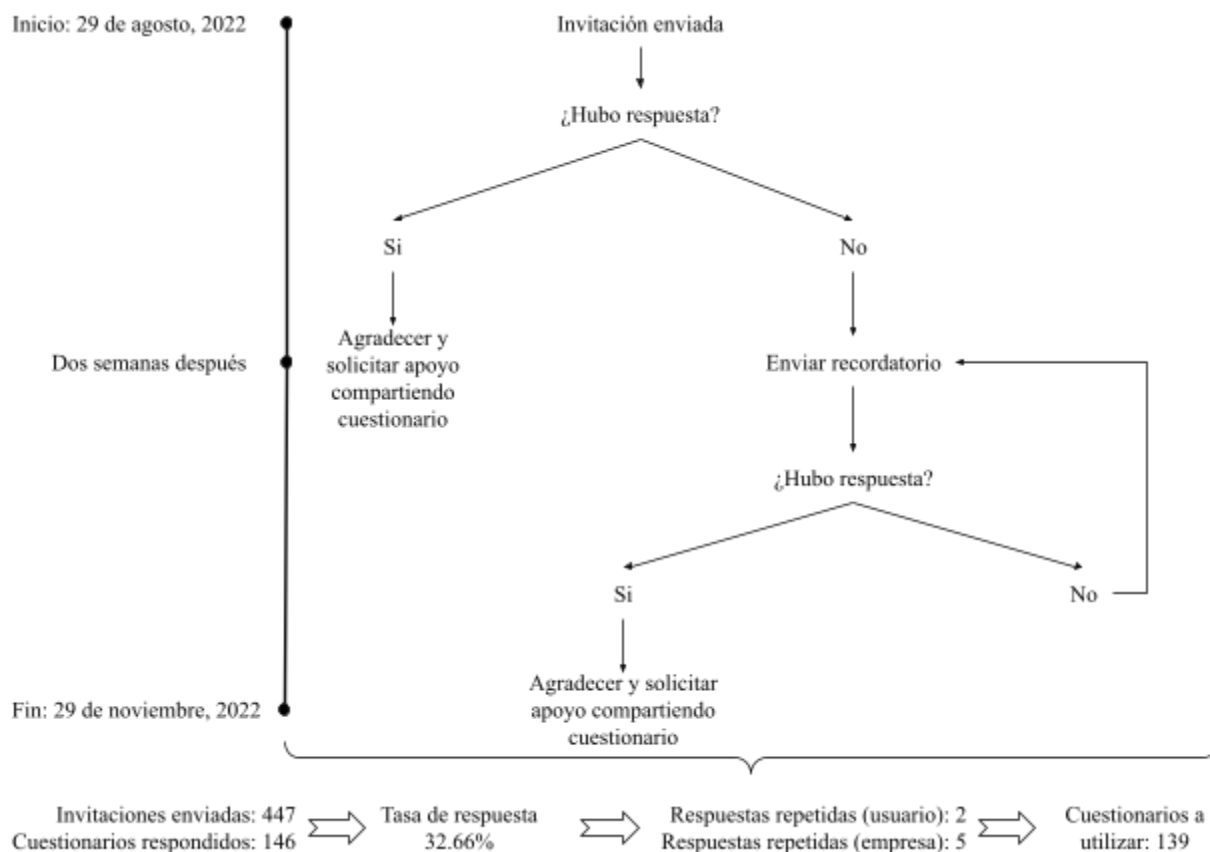
Fuente: elaboración propia.

Una vez modificado el *Google Forms* de acuerdo a la retroalimentación recibida en la prueba piloto, se procedió a definir la muestra a la cual se aplicaría el cuestionario. Como ya mencionamos, la población se define como tomadores de decisiones en empresas de Tijuana, Baja California, con un perfil de amplio conocimiento sobre la estrategia y operación de la empresa. Dados los retos enfrentados para acceder a una muestra de la población objetivo que satisficiera los criterios mencionados, puesto que los tomadores de decisiones con ese perfil son difíciles de contactar, se decidió aplicar una técnica de muestreo de bola de nieve modificada, en la que las semillas iniciales para el muestreo son los contactos directos del investigador pertenecientes a la población objetivo, lo cual tiene la ventaja de dirigir los esfuerzos con tomadores de decisiones con el perfil adecuado (Torres et al., 2018; Thompson, 2002; Goodman, 1961). Esto no significa que la selección de la muestra se realizó con algún tipo de sesgo, puesto que de tener disponible toda la información de contacto de la población, el perfil de los

tomadores de decisiones seguiría actuando cómo un filtro natural a la población, debido a las especificaciones del perfil y a la naturaleza de la información solicitada.

Teniendo en cuenta lo anterior, se utilizó la red social *LinkedIn* como población para realizar el muestreo y enviar solicitud de respuesta al instrumento de recolección de datos, siendo el periodo de aplicación del 29 de agosto al 29 de noviembre de 2022. *LinkedIn* es “la red más grande de profesionales y tiene como misión crear oportunidades económicas para cada uno de los miembros de su red global de fuerza de trabajo” (LinkedIn, s.f.). Como contacto directo, se enviaron 447 invitaciones personalizadas con la liga al formulario de *Google* a través del servicio de mensajería de *LinkedIn*, dónde en caso de obtener respuesta al cuestionario, se les solicitaba compartir con otros miembros de su red de profesionales tomadores de decisiones (Torres et al., 2018), y en caso de no obtener respuesta, se enviaba un amable recordatorio cada dos semanas. De las 447 solicitudes directas, se obtuvo respuesta a 146 cuestionarios, representando una tasa de respuesta del 32.66%, las cuáles fueron filtradas para eliminar respondientes y empresas repetidas, quedando un total de 139 cuestionarios válidos. El proceso de obtención de los datos se puede observar en la figura 24.

Figura 24. Proceso de recolección de datos



Fuente: elaboración propia.

Sin embargo, teniendo en cuenta los requerimientos descritos anteriormente en lo que respecta al tamaño de la muestra, los 139 cuestionarios válidos se dividieron entre los respondientes que afirmaron sí utilizar analíticos de datos para la toma de decisiones y los que no, quedando un total de 95 cuestionarios a utilizar como muestra para la aplicación del modelo PLS-SEM, es decir un 74.82% de los cuestionarios válidos. De acuerdo con el criterio planteado por Hair et al. (2021), se toma como referencia el valor mínimo medio de los coeficientes de trayectoria, así como un nivel de significancia del 5%, para determinar que el tamaño mínimo de la muestra para aplicar el modelo PLS-SEM es de 69 observaciones válidas, por lo que con los 95 cuestionarios a utilizar se cumple con los lineamientos estándar para la aplicación de la metodología PLS-SEM. En las siguientes tablas se desglosan las características de los cuestionarios válidos entre los respondientes que pertenecen a empresas que sí utilizan analíticos de datos para la toma de decisiones, esto es, las características de la muestra a utilizar.

Del total de respuestas válidas obtenidas y que sí utilizan analíticos de datos para la toma de decisiones, el 80% de los tomadores de decisiones es de sexo masculino, mientras que el 20% restante es de sexo femenino. De igual forma, la mayoría de los tomadores de decisiones que respondieron el instrumento de recolección de datos desempeñan puestos de Dirección General, Gerencia de Operaciones y otros puestos directivos, representando un 64.22% del total de la muestra. Mientras que en lo que respecta al grupo etario de los respondientes, el 35.79% se ubica dentro del rango de 41 a 50 años, y un 25.26% en el grupo de 31 a 40 años de edad. En la tabla 13 se resumen las características de la muestra correspondientes a los tomadores de decisiones o respondientes del cuestionario utilizado para recopilar los datos.

Tabla 13. Características de la muestra - Respondientes

Características del respondiente	Porcentaje (%)
Puesto que desempeña	
Dirección General	26.32%
Gerencia de Operaciones	10.53%
Otros puestos directivos	27.37%
Otros puestos	11.58%
Puestos directivos en ingeniería	7.37%
Otros puestos de gerencia	8.42%
Puestos directivos en TI	4.21%
Puestos de supervisión	4.21%
Sexo	
Femenino	20.00%
Masculino	80.00%
Grupo etario	
20 - 30 años	18.95%
31 - 40 años	25.26%
41 - 50 años	35.79%
51 - 65 años	18.95%
Más de 65 años	1.05%

Fuente: elaboración propia.

Por otra parte, la mayoría de las empresas pertenecientes a la muestra tienen entre 1.5 y 5 años de operación (23.16%), seguida por las empresas en el grupo de 6 a 10 años de operación con 14.74% del total, siendo en general una de las características que presenta una distribución más homogénea con porcentajes similares en cada categoría. Dadas las características económicas de la ciudad de Tijuana, Baja California, era de esperarse que la industria en la que

operan las empresas de la muestra fuese en su mayoría la correspondiente a las industrias manufactureras, representando el 44.21% de la muestra, seguidas por las empresas pertenecientes a industrias de servicios (otros servicios, servicios profesionales, de salud y financieros) con un 22.11%. En ese mismo sentido, más de la mitad de las empresas (65.26%) son empresas consideradas grandes, teniendo más de 250 empleados, lo cual es consistente con la concentración de empresas manufactureras en la región. En la tabla 14 se describen en detalle las características de las empresas pertenecientes a la muestra a utilizar en la presente investigación.

Tabla 14. Características de la muestra - Empresas

Características de la empresa	Porcentaje (%)
Años de operación	
1 año	10.53%
1.5 a 5 años	23.16%
6 a 10 años	13.68%
11 a 15 años	7.37%
16 a 20 años	14.74%
21 a 25 años	7.37%
26 a 30 años	10.53%
31 a 100 años	12.63%
Tamaño de la empresa (número de empleados)	
0 - 10	8.42%
11 - 50	10.53%
51 - 250	15.79%
251 o más	65.26%
Industria	
Industrias manufactureras	44.21%
Otros servicios excepto actividades gubernamentales	7.37%
Servicios profesionales, científicos y técnicos	6.32%
Otras industrias manufactureras	5.26%
Servicios de salud y de asistencia social	4.21%
Servicios financieros y de seguros	4.21%
Comercio al por mayor	3.16%
Comercio al por menor	3.16%
Construcción	3.16%
Corporativos	3.16%
Fabricación de equipo de computación, comunicación, medición y de otros equipos, componentes y accesorios electrónicos	3.16%
Transportes, correos y almacenamiento	3.16%
Fabricación de productos metálicos	2.11%
Información en medios masivos	2.11%
Otras industrias	5.25%

Fuente: elaboración propia.

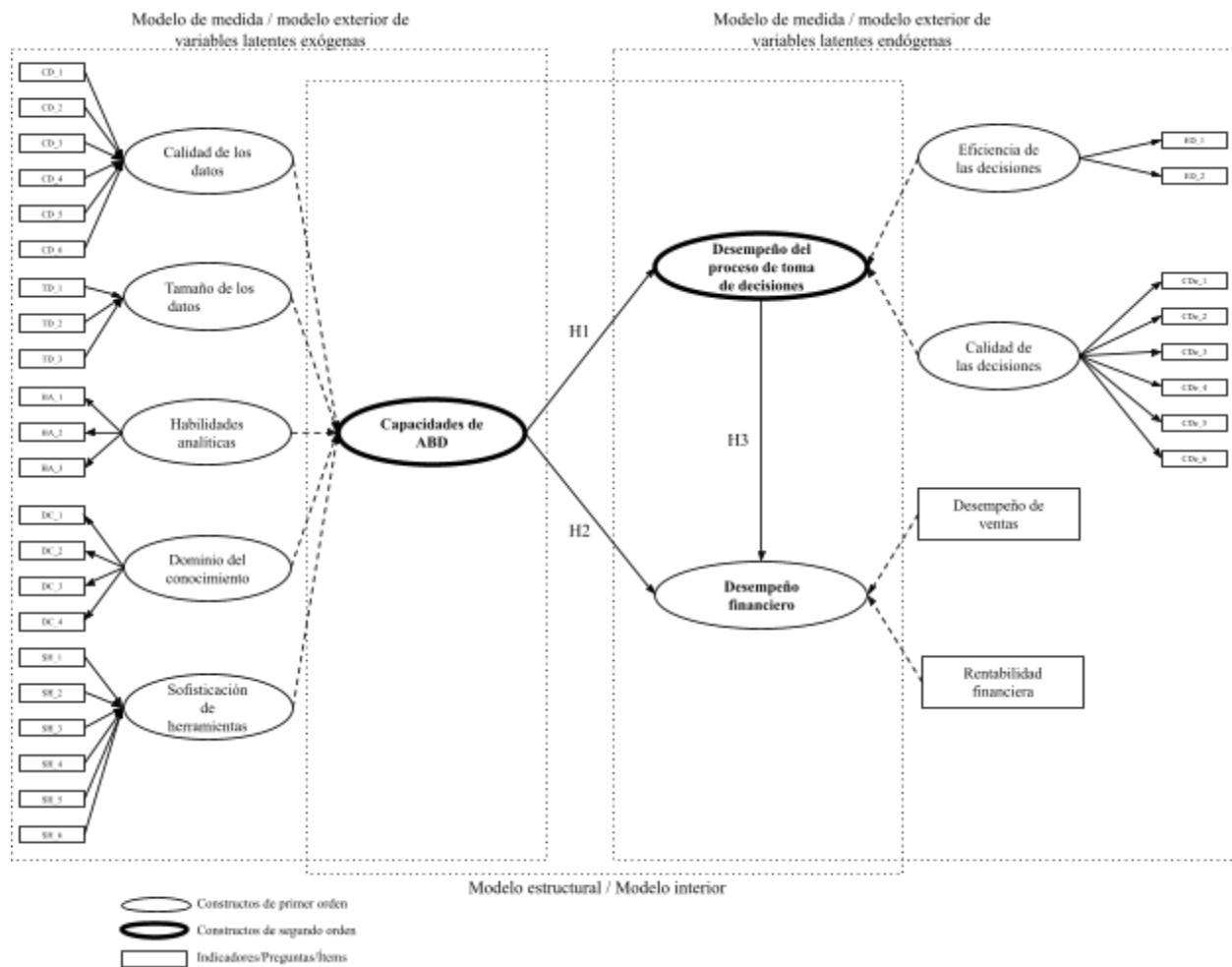
Modelo empírico

Como hemos revisado, las hipótesis propuestas son: 1) a mayores CABD, mejor será el desempeño del proceso de toma de decisiones de la firma; 2) el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la firma; y 3) el desempeño del proceso de toma

de decisiones de la firma tiene un efecto mediador entre las CABD y el desempeño financiero. Asimismo, se ha mencionado que la metodología de análisis será un modelo PLS-SEM, una técnica estadística que se utiliza para analizar relaciones complejas entre variables latentes o no observables (Hair et al., 2017). Las variables latentes en un modelo PLS-SEM son constructos que no se pueden observar directamente, por lo que se requiere inferirlas utilizando otras variables que sí pueden ser observadas, por ejemplo las CABD de una firma no se pueden medir por sí solas, sino que se tiene que utilizar sustento teórico para definir sus componentes y recabar información sobre dichos componentes, como se hizo utilizando el instrumento de recolección de datos (Hair et al., 2017).

Considerando lo anterior, las CABD se operacionalizan en el modelo empírico PLS-SEM como un constructo formativo de segundo orden compuesto por cinco constructos de primer orden: calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas; mientras que el desempeño del proceso de toma de decisiones se operacionalizan como un constructo formativo de segundo orden conformado por dos constructos de primer orden: calidad de las decisiones y eficiencia de las decisiones; por último, el desempeño financiero de la empresa se operacionaliza como un constructo de primer orden. En la figura 25 se ilustra el modelo de medición y el modelo estructural propuestos, en el cuál los indicadores formativos se representan por flechas con una sola punta dirigida hacia dentro del constructo latente de las variables indicadoras. El modelo estructural, o modelo interior, muestra las relaciones (trayectorias o *paths*) entre los constructos latentes o variables no observadas, permitiendo solamente relaciones recursivas y no bucles causales, por lo que las trayectorias estructurales entre los constructos latentes solamente pueden ir en una sola dirección, diferenciando entre constructos: exógenos, que son constructos latentes que no tienen ninguna trayectoria estructural apuntando hacia ellos; y endógenos o constructos latentes que son explicados por otros constructos a través de relaciones estructurales del modelo.

Figura 25. Modelo de investigación propuesto



Fuente: elaboración propia.

En ese sentido, las CABD constituyen nuestra variable latente dependiente, mientras que el desempeño del proceso de toma de decisiones y el desempeño financiero de la firma son las variables latentes dependientes del modelo. Como variables latentes o no observables, es necesario definirlas y recabar datos para poder estimarlas. Teniendo esto en cuenta y como ya se ha mencionado, las CABD se componen de cinco dimensiones o variables observables: calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento del dominio y sofisticación de las herramientas. En la ecuación (7) se describe la estimación de las CABD un constructo formativo de segundo orden, mientras que en las ecuaciones (8), (9), (10), (11) y (12) describen la estimación de las variables observables, las cuáles se estimaron con las preguntas o ítems del instrumento de recolección de datos y su estimación está sujeta a si son constructos

formativos o reflectivos, en éste caso, los constructos de primer orden reflectivos son las habilidades analíticas y el conocimiento del dominio, por lo que los coeficientes corresponden a las cargas o pesos de los factores, contrario a los constructos de primer orden formativos dónde los coeficientes corresponden a la ponderación, pesos, factor compuesto o coeficiente de regresión de cada uno de los indicadores que lo componen.

$$CABD = \beta_1 CD + \beta_2 TD + \beta_3 HA + \beta_4 DC + \beta_5 SH + \varepsilon_1 \quad (7)$$

Dónde:

CABD = Capacidades de Analíticos de *Big Data*

CD = Calidad de los Datos

TD = Tamaño de los Datos

HA = Habilidades Analíticas

DC = Conocimiento del Dominio

SH = Sofisticación de las Herramientas

β_i = Coeficiente de ponderación o factor compuesto

ε_1 = Término de error

$$CD = \beta_6 CD_1 + \beta_7 CD_2 + \beta_8 CD_3 + \beta_9 CD_4 + \beta_{10} CD_5 + \beta_{11} CD_6 + \varepsilon_2$$

(8)

Dónde:

CD = Calidad de los Datos

CD_j = Preguntas o ítems

β_i = Coeficiente de ponderación o factor compuesto

ε_2 = Término de error

$$TD = \beta_{12} TD_1 + \beta_{13} TD_2 + \beta_{14} TD_3 + \varepsilon_3 \quad (9)$$

Dónde:

TD = Tamaño de los Datos

TD_j = Preguntas o ítems

β_i = Coeficiente de ponderación o factor compuesto

ε_3 = Término de error

$$HA = \lambda_1 HA_1 + \lambda_2 HA_2 + \lambda_3 HA_3 + \varepsilon_4 \quad (10)$$

Dónde:

HA = Habilidades Analíticas

HA_j = Preguntas o ítems

λ_i = Coeficiente de cargas o pesos de los factores

ε_4 = Término de error

$$DC = \lambda_4 DC_1 + \lambda_5 DC_2 + \lambda_6 DC_3 + \lambda_7 DC_4 + \varepsilon_5 \quad (11)$$

Dónde:

DC = Conocimiento del Dominio

DC_j = Preguntas o ítems

λ_i = Coeficiente de cargas o pesos de los factores

ε_5 = Término de error

$$SH = \beta_{15} SH_1 + \beta_{16} HA_2 + \beta_{17} HA_3 + \beta_{18} HA_4 + \beta_{19} HA_5 + \beta_{20} HA_6 + \varepsilon_6 \quad (12)$$

Dónde:

SH = Sofisticación de las Herramientas

SH_j = Preguntas o ítems

β_i = Coeficiente de ponderación o factor compuesto

ε_6 = Término de error

En lo que respecta a las variables dependientes, en primera instancia se define al desempeño del proceso de toma de decisiones de la empresa como una variable latente o no observable que se compone de la calidad y la eficiencia de las decisiones (ecuación (13)), variables observables que se componen, a su vez de las preguntas del instrumento de recolección, como se especifica en las ecuaciones (14) y (15).

$$DPTD = \beta_{21} CDe + \beta_{22} ED + \varepsilon_7 \quad (13)$$

Dónde:

DPTD = Desempeño del Proceso de Toma de Decisiones

CDe = Calidad de las Decisiones

ED = Eficiencia de las Decisiones

β_i = Coeficiente de ponderación o factor compuesto

ε_7 = Término de error

$$CDe = \lambda_8 CDe_1 + \lambda_9 CDe_2 + \lambda_{10} CDe_3 + \lambda_{11} CDe_4 + \lambda_{12} CDe_5 + \lambda_{13} CDe_6 + \varepsilon_8 \quad (14)$$

Dónde:

CDe = Calidad de las Decisiones

CDe_j = Preguntas o ítems

λ_i = Coeficiente de cargas o pesos de los factores

ε_8 = Término de error

$$ED = \lambda_{14} ED_1 + \lambda_{15} ED_2 + \varepsilon_9 \quad (15)$$

Dónde:

ED = Eficiencia de las Decisiones

ED_j = Preguntas o ítems

λ_i = Coeficiente de cargas o pesos de los factores

ε_9 = Término de error

Finalmente, el desempeño financiero de la empresa, la segunda variable dependiente del modelo, a pesar de también ser una variable latente o no observables, ésta corresponde con un constructo de primer orden, por lo que sus componentes son las preguntas o indicadores en sí, por lo que el desempeño financiero de la firma se define de acuerdo con la ecuación (16).

$$DF = \beta_{23} UN + \beta_{24} Vtas + \varepsilon_{10} \quad (16)$$

Dónde:

DF = Desempeño Financiero

UN = Utilidad Neta

Vtas = Ventas

β_i = Coeficiente de ponderación o factor compuesto

ε_{10} = Término de error

Una vez definidas cada una de las variables, se procede a definir las ecuaciones para estimar las relaciones del modelo interior, es decir las relaciones causales entre las variables propuestas. Teniendo esto en cuenta, las CABD actúan como variables independientes y tienen un impacto en el desempeño del proceso de toma de decisiones y el desempeño financiero de la firma, ambas variables dependientes, mientras que el desempeño del proceso de toma de decisiones también actúa como mediador entre las CABD y el desempeño financiero de la empresa. En las ecuaciones (17) y (18) se plantea la relación directa entre las CABD y las variables dependientes, mientras que en la ecuación (19) se plantea el rol mediador del

desempeño del proceso de toma de decisiones entre las CABD y el desempeño financiero de la firma.

$$DPTD = \delta_1 CABD + \varepsilon_{11} \quad (17)$$

Dónde:

DPTD = Desempeño del Proceso de Toma de Decisiones

CABD = Capacidades de Analíticos de *Big Data*

δ_1 = Coeficiente de regresión, ponderación o factor compuesto

ε_{11} = Término de error

$$DF = \delta_2 CABD + \varepsilon_{12} \quad (18)$$

Dónde:

DF = Desempeño Financiero

CABD = Capacidades de Analíticos de *Big Data*

δ_2 = Coeficiente de regresión, ponderación o factor compuesto

ε_{12} = Término de error

$$DF = \delta_3 CABD + \delta_4 DPTD + \varepsilon_{13} \quad (19)$$

Dónde:

DF = Desempeño Financiero

CABD = Capacidades de Analíticos de *Big Data*

DPTD = Desempeño del Proceso de Toma de Decisiones

δ_i = Coeficiente de regresión, ponderación o factor compuesto

ε_{13} = Término de error

Capítulo IV.

Análisis y resultados

El presente capítulo se centra en la aplicación del diseño metodológico descrito previamente para conocer el impacto de las CABD en el proceso de toma de decisiones y desempeño financiero de las empresas en Tijuana, Baja California y se divide en cuatro secciones. En la primera se caracteriza la muestra utilizada a través de la estadística descriptiva, haciendo un análisis del total de cuestionarios válidos, para comparar las diferencias entre las empresas que utilizan analíticos de datos para la toma de decisiones y las que no. Posteriormente se evalúa el modelo de medición, en primera instancia para los constructos de primer orden y posteriormente para los constructos de segundo orden utilizando el enfoque de dos etapas, de acuerdo con los lineamientos descritos en el diseño metodológico y utilizando el paquete SEMinR⁴ del software RStudio⁵ para realizar todos los cálculos del modelo PLS-SEM. En la tercera sección, se describe la evaluación del modelo estructural, es decir, la evaluación de los resultados del modelo en general, mientras que en la última sección se procede a discutir dichos resultados empíricos, considerando las preguntas e hipótesis de investigación.

Caracterización de las empresas de la muestra

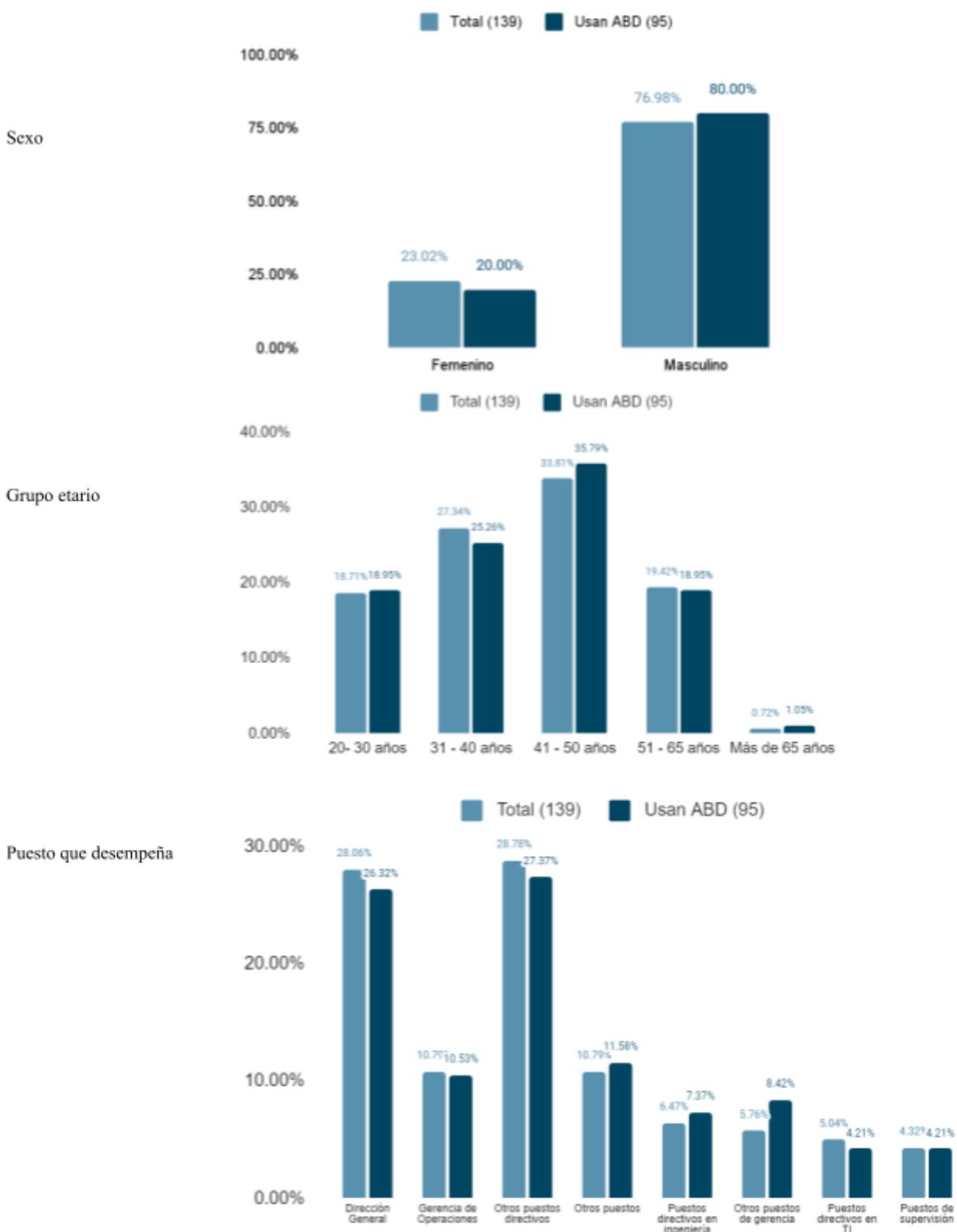
Como ya se ha mencionado, de un total de 447 cuestionarios solicitados, se recibieron un total de 146 respuestas válidas, dando una tasa de respuesta de 32.66%, las cuales se convirtieron en 139 cuestionarios válidos a utilizar, debido a la eliminación de respuestas repetidas, ya fuese por empresa o por respondiente. En la figura 26 se presentan las características de la muestra de cuestionarios válidos a utilizar, enfocada en el respondiente, así como una caracterización

⁴ El paquete SEMinR, el cual permite a los usuarios crear y modificar fácilmente modelos de ecuaciones estructurales (SEM). Permite la estimación mediante SEM basado en la covarianza (CBSEM, como LISREL/Lavaan) o modelos de trayectoria de mínimos cuadrados parciales (PLS-SEM, como SmartPLS/semPLS) y fue creado por Soumya Ray, Nicholas Danks y André Calero Valdez, con colaboración de James Uanhoro, Arturo Heynar Cano Bejar y Johannes Nakayama (Ray et al., s.f.).

⁵ R es un programa de software libre y de código abierto para el análisis estadístico, basado en el lenguaje S, mientras que RStudio es un IDE (entorno de desarrollo integrado) gratuito y de código abierto para R. Su interfaz está organizada de forma que el usuario pueda ver claramente gráficos, tablas de datos, código R y resultados, todo al mismo tiempo. También ofrece una función similar a Import-Wizard que permite a los usuarios importar archivos CSV, Excel, SAS (*.sas7bdat), SPSS (*.sav) y Stata (*.dta) a R sin tener que escribir el código para hacerlo (Kent State University, s.f.).

diferenciando entre los cuestionarios en los que el respondiente afirmó que la empresa a la que pertenece utiliza analíticos de datos en su proceso de toma de decisiones. Se puede observar que las características para los respondientes de los cuestionarios válidos y la muestra utilizada son muy similares y solamente se observan ligeras diferencias en sexo, cambiando en 3.02%; mientras que en grupo etario se mantiene la predominancia del grupo de 41 a 50 años; asimismo, se mantiene predominancia en los grupos de dirección general y otros puestos directivos en la pregunta sobre el puesto que desempeña el respondiente.

Figura 26. Caracterización de la muestra - Respondiente

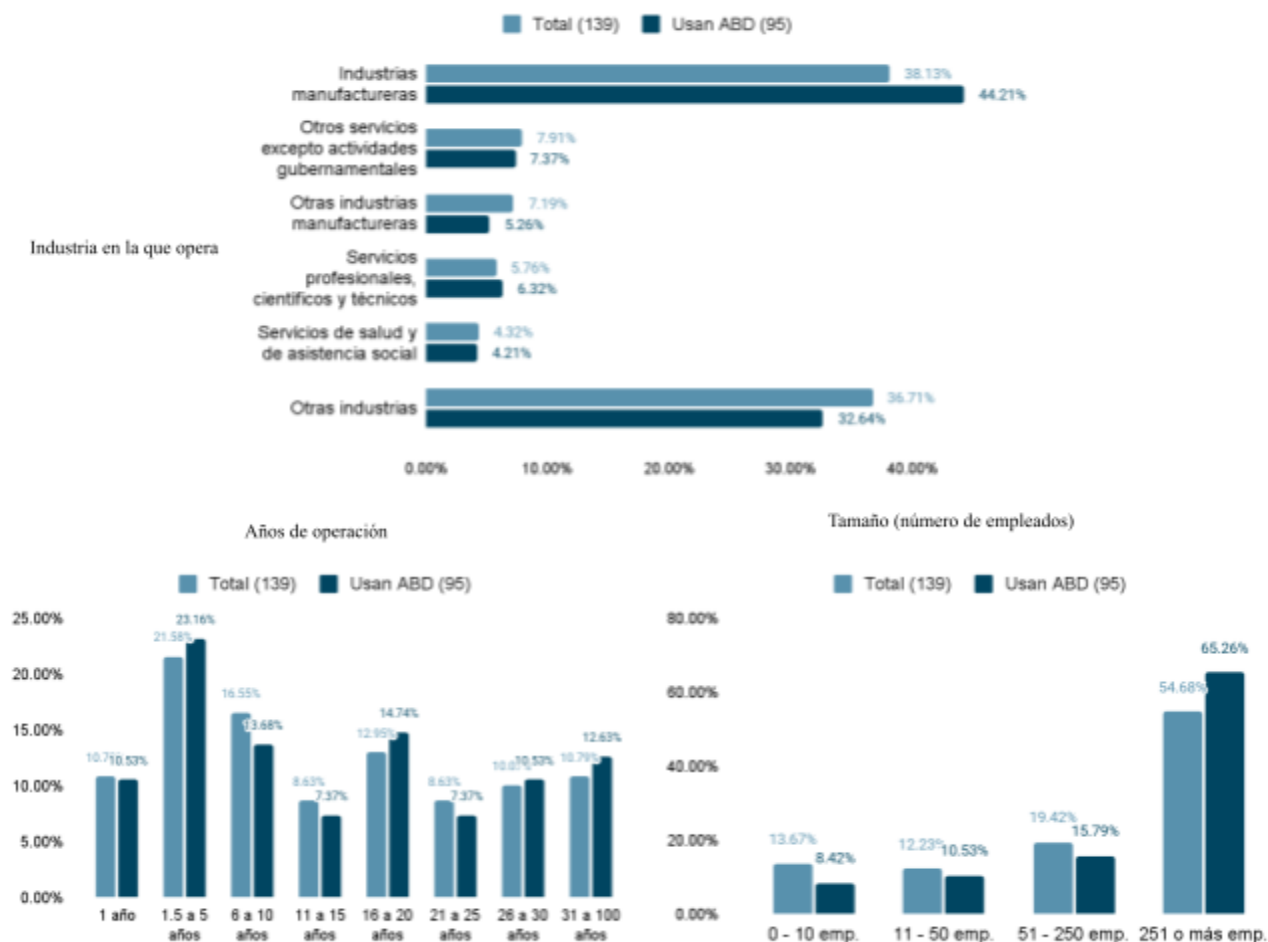


Fuente: elaboración propia.

Es importante mencionar que la categoría de otros puestos directivos incluye dirección comercial, dirección de administración, dirección de finanzas, dirección de planeación, dirección de proyectos, dirección de recursos humanos, dirección de servicio al cliente, vicepresidencia, socio de operaciones y subgerencia de operaciones; la categoría de otros puestos incluye administración, administración de proyectos, analista de datos, consultoría, puestos de contabilidad y de mercadotecnia; la categoría de puestos directivos de ingeniería incluye puestos directivos en áreas calidad, manufactura, mejora continua y sustentabilidad; la categoría otros puestos de gerencia incluye gerencias en áreas de calidad, cuentas, logística y producción; la categoría de puestos directivos en TI incluye dirección de seguridad en tecnologías de la información, dirección de TI y especialista en TI; por último, la categoría de puestos de supervisión incluye supervisión en las áreas de compras, finanzas, operaciones y producción, por lo que, a pesar de no tener un puesto directivo, son quienes recolectan los datos e implementan las decisiones tomadas con los resultados de su análisis.

Por otra parte, las características de la empresa en la que se desempeña el respondiente se presentan en la figura 27, en la que podemos observar que al igual que con los respondientes, las características de los cuestionarios válidos es muy similar entre las empresas que utilizan y no utilizan analíticos de datos en su proceso de toma de decisiones, por ejemplo, en ambos casos predominan las empresas que se ubican en el rango de 1.5 a 5 años de operación, sin embargo, para las empresas que sí utilizan analíticos de datos, la categoría siguiente es la de 16 a 20 años; algo similar ocurre con el tamaño de la empresa, en ambos casos más de la mitad de las empresas a las que pertenecen los respondientes tienen 251 empleados o más; por último, en lo que respecta a la industria en la que operan las empresas, a pesar de que también existe una concentración en las industrias manufactureras, el porcentaje cambia de 38.13% a 44.21% para las empresas que utilizan analíticos de datos.

Figura 27. Caracterización de la muestra - Empresa



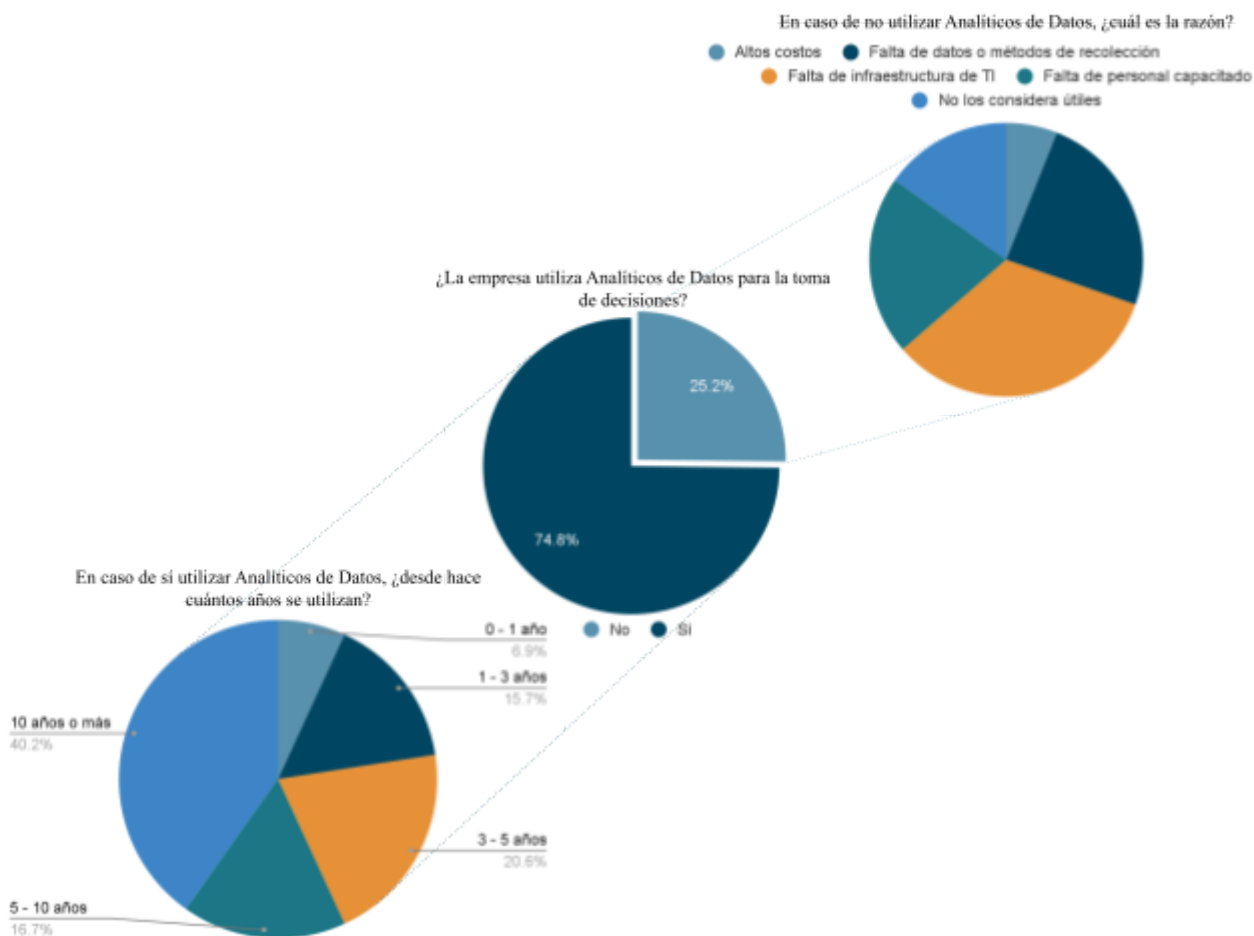
Fuente: elaboración propia.

Una de las principales preguntas del instrumento de recolección de datos hacía referencia a si la empresa a la que pertenece el respondiente utiliza analíticos de datos para la toma de decisiones. La importancia reside en que fue justamente la pregunta que permitió segmentar las respuestas válidas entre respuestas a utilizar en la muestra del modelo empírico o no, así como la descripción de la muestra representada en las figuras anteriores. Posteriormente, se hicieron preguntas sobre el tiempo de utilización de los analíticos de datos para la toma de decisiones, en caso de que sí los utilizaran, o bien, la razón por la que no los utilizan.

En la figura 28 se puede observar el desglose de las respuestas obtenidas, donde se puede observar que, como ya se mencionó, del total de cuestionarios válidos, el 74.8% de los respondientes afirmó que la empresa en la que labora sí utilizan analíticos de datos para la toma de decisiones y de esas empresas el 40.2% tiene utilizándolos 10 años o más, mientras que la

principal razón mencionada por las empresas que no utilizan analíticos de datos para la toma de decisiones es la falta de infraestructura de TI con 33.3%, seguido por falta de datos o métodos de recolección con 24.2% y falta de personal capacitado con 21.2%.

Figura 28. Utilización de analíticos de datos para la toma de decisiones

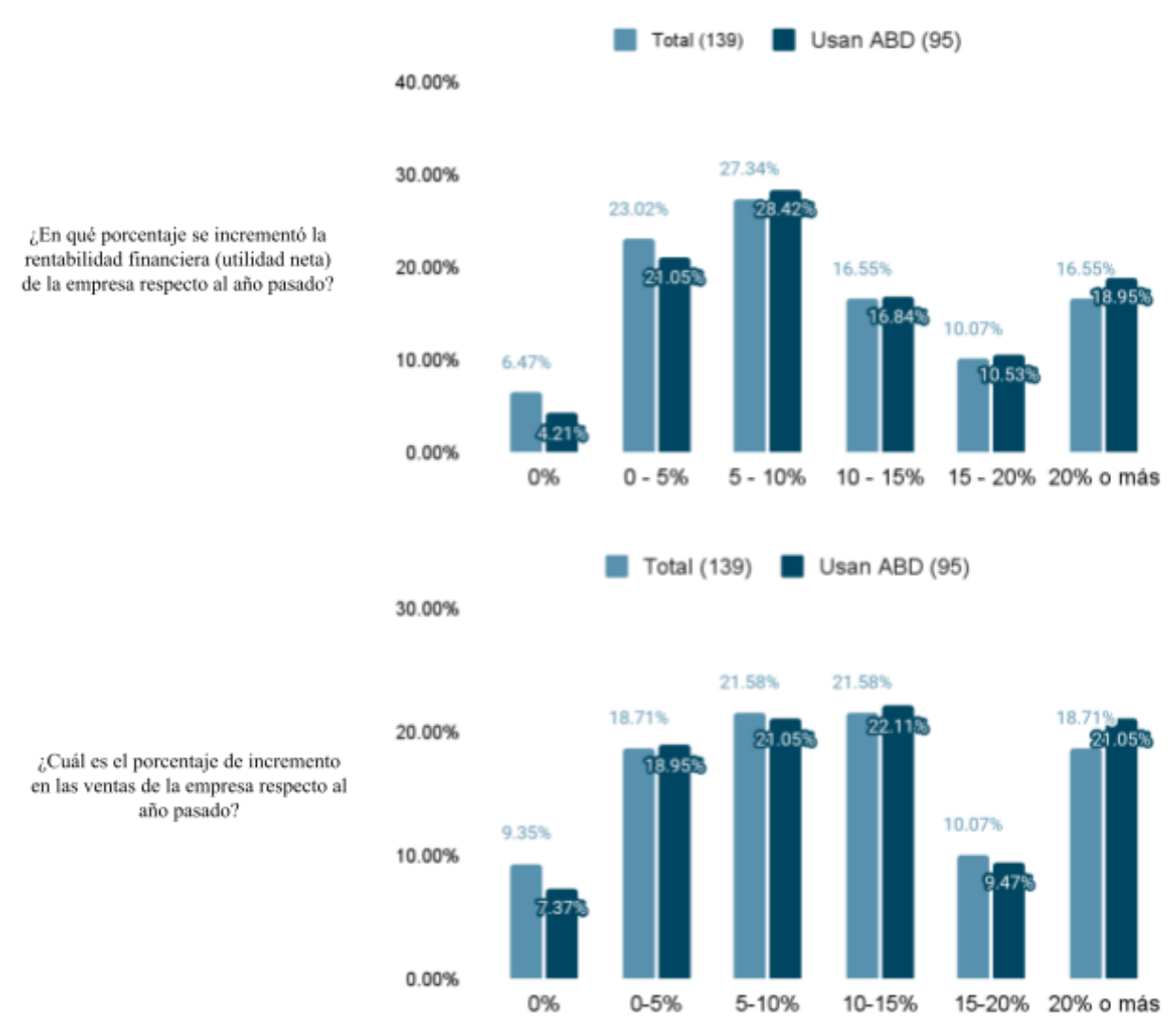


Fuente: elaboración propia.

Asimismo, dentro del apartado sobre las características de la empresa, se les preguntó a los respondientes el porcentaje de incremento en la utilidad neta en la empresa respecto al año pasado, así como el porcentaje de incremento en las ventas respecto al año pasado, independientemente del uso de analíticos de datos para la toma de decisiones. Esto con el objetivo de conocer si la utilización de analíticos de datos para la toma de decisiones tiene algún impacto en dichos indicadores, para posteriormente indagar más a detalle sobre la. En la figura 29 se puede observar que en lo que respecta al incremento en la utilidad neta, cuando se segmenta a las empresas que sí utilizan analíticos de datos para la toma de decisiones los

porcentajes de incremento son mayores en alrededor del 2%, mientras que el incremento en las ventas exhibe el mismo comportamiento, siendo más apreciado en la categoría de incremento en ventas mayor al 20% respecto al año anterior, la cual pasa del 18.71% al 21.05% para empresas que sí utilizan analíticos de datos para la toma de decisiones.

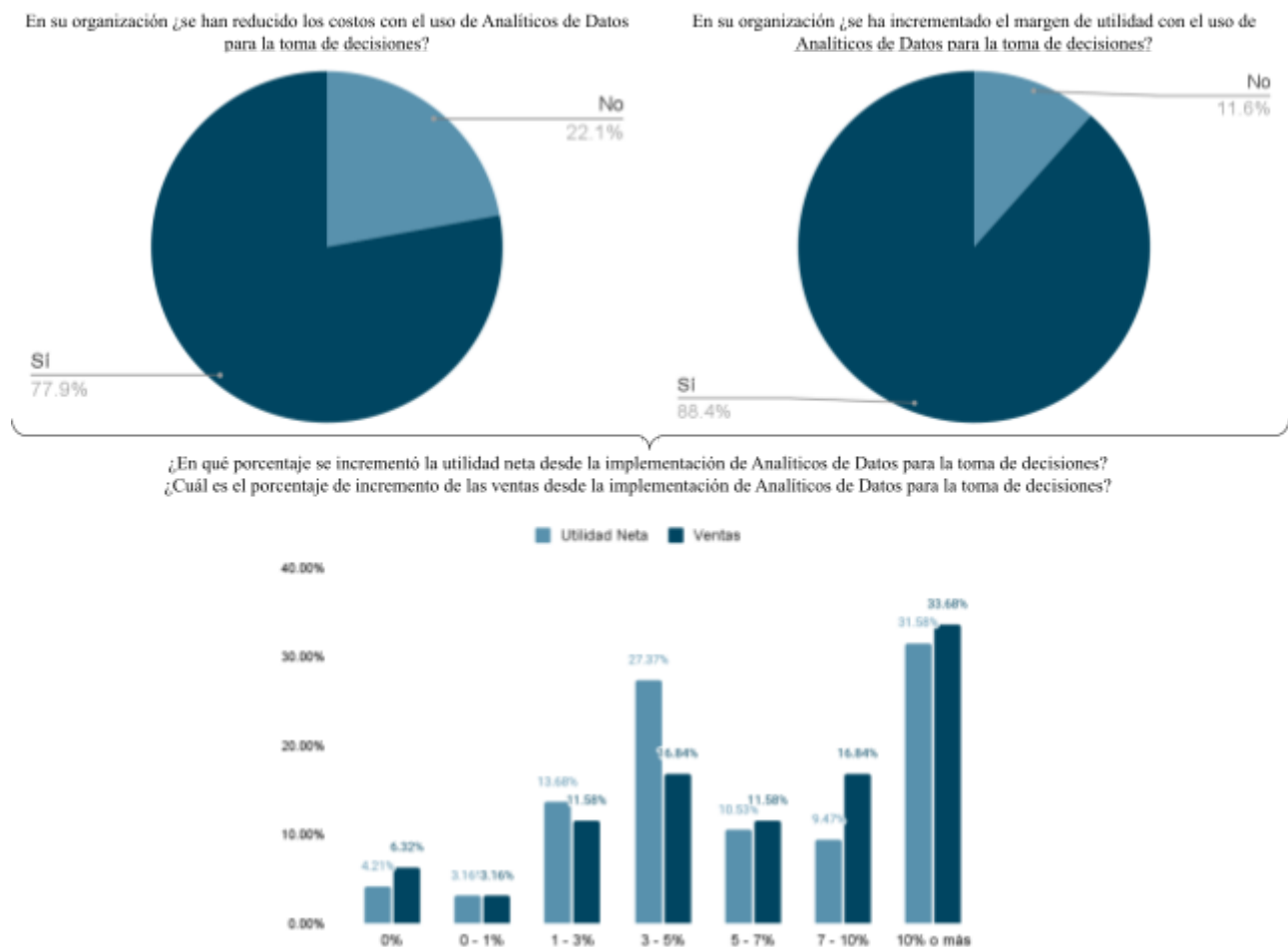
Figura 29. Indicadores de desempeño financiero de la empresa



Fuente: elaboración propia.

Por otra parte, continuando con las empresas que sí utilizan analíticos de datos para la toma de decisiones, se cuestionó sobre la relación entre el incremento en el margen de utilidad y la reducción de costos con el uso de analíticos de datos. En la figura 30 se observa la relación entre el uso de analíticos de datos para la toma de decisiones y la reducción de costos en la empresa, pues de las empresas que utilizan analíticos de datos para la toma de decisiones, el 77.9% respondieron que sí se han reducido los costos con su uso. De igual forma, el 88.4% de las empresas que utiliza analíticos de datos para la toma de decisiones afirmó experimentar un incremento en su margen de utilidad derivado de su utilización. En ese mismo sentido, se cuestionó sobre el tamaño del impacto del uso de analíticos de datos para la toma de decisiones, medido a través del porcentaje de incremento en la utilidad neta y en las ventas, encontrando que en más del 30% de las empresas el incremento en utilidad neta y en ventas fue superior al 10% desde la implementación del uso de analíticos de datos para la toma de decisiones, seguido por el 27.37% en utilidad neta de las empresas que experimentó un incremento de entre 3% y 5%, y un 16.84% de empresas que afirmó experimentar un crecimiento de 3-5% y de 7-10% en sus ventas desde la implementación del uso de analíticos de datos para la toma de decisiones, respectivamente.

Figura 30. Uso de analíticos de datos e indicadores de desempeño de la empresa



Fuente: elaboración propia.

Finalmente, como caracterización de la muestra, tenemos las estadísticas descriptivas de cada una de las preguntas o ítems incluidos en el instrumento de recolección de datos, incluyendo la media, mínimo, máximo, así como la desviación estándar. Es importante mencionar que las respuestas del instrumento de recolección de datos estaban codificadas de acuerdo con una escala Likert de siete puntos: 1 para totalmente en desacuerdo, 2 para en desacuerdo, 3 algo de acuerdo, 4 para ni de acuerdo ni en desacuerdo, 5 para algo de acuerdo, 6 para de acuerdo y 7 para totalmente de acuerdo.

Tabla 15. Estadística descriptiva de los ítems del instrumento de recolección de datos

Ítem o pregunta	Media	Mín	Max	Desv Std
En su organización ¿considera que las Capacidades de Analíticos de Big Data se encuentran desarrolladas?	5.04	1	7	1.74
En nuestra organización, la calidad de los datos utilizados para construir Analíticos de Datos es muy buena.	5.62	2	7	1.03
En nuestra organización, los datos utilizados para construir Analíticos de Datos son confiables.	5.55	1	7	1.13
En nuestra organización, los datos utilizados para construir Analíticos de Datos tienen el nivel apropiado de detalle.	5.38	1	7	1.21
En nuestra organización, los datos utilizados para construir Analíticos de Datos son seguros.	5.57	1	7	1.23
En nuestra organización, los datos utilizados para construir Analíticos de Datos son oportunos.	5.43	2	7	1.09
En nuestra organización, los datos utilizados para construir Analíticos de Datos son relevantes para la tarea en cuestión.	5.81	2	7	0.95
En nuestra organización, los datos utilizados para construir Analíticos de Datos son precisos.	5.42	2	7	1.18
Nuestra organización recopila y analiza grandes volúmenes de datos.	5.71	2	7	1.12
Nuestra organización procesa grandes volúmenes de datos.	5.53	1	7	1.55
Nuestra organización procesa datos en tiempo real.	5.44	1	7	1.37
Nuestra organización procesa diferentes tipos de datos.	5.99	1	7	1.01
Nuestros usuarios de Analíticos de Datos son expertos en lo que respecta a utilizar las herramientas de análisis de datos.	5.04	1	7	1.39
Nuestros usuarios de Analíticos de Datos poseen un alto grado de experiencia en análisis de datos.	4.97	1	7	1.33
Nuestros usuarios de Analíticos de Datos son hábiles en el uso de herramientas de análisis de datos.	5.14	1	7	1.36
En nuestra organización, hay un alto nivel de conocimiento del entorno externo (e.g. gobierno, competidores, proveedores y clientes).	5.60	1	7	1.35
En nuestra organización, hay un alto nivel de conocimiento de las metas y objetivos organizacionales.	5.92	1	7	1.35
En nuestra organización, hay un alto nivel de conocimiento de las capacidades básicas de la organización.	5.81	1	7	1.06
En nuestra organización, hay un alto nivel de conocimiento de los factores clave para que la organización tenga éxito.	5.78	1	7	1.20
En nuestra organización, se utilizan las herramientas de análisis de datos apropiadas de acuerdo a nuestras necesidades.	5.21	2	7	1.05

Ítem o pregunta	Media	Mín	Max	Desv Std
En nuestra organización, utilizamos herramientas que brindan capacidades de procesamiento y recuperación de información.	5.78	2	7	1.00
En nuestra organización, utilizamos herramientas que realizan modelado y simulación.	5.05	2	7	1.53
En nuestra organización, utilizamos herramientas que realizan analíticos de lenguaje natural (extracción de información de fuentes no estructuradas como redes sociales).	4.46	1	7	1.70
En nuestra organización, utilizamos herramientas que proporcionan información en tiempo real.	5.44	1	7	1.32
En nuestra organización, utilizamos herramientas que identifican problemas.	5.41	1	7	1.32
En nuestra organización, utilizamos herramientas que evalúan diferentes alternativas.	5.32	1	7	1.30
En nuestra organización existe un buen desempeño del proceso de toma de decisiones.	5.08	2	7	1.07
En nuestra organización, los resultados de las decisiones son a menudo exactos.	5.06	2	7	1.32
En nuestra organización, los resultados de las decisiones son a menudo correctos.	5.47	2	7	1.07
En nuestra organización, los resultados de las decisiones son a menudo precisos.	5.32	2	7	1.10
En nuestra organización, los resultados de las decisiones son a menudo perfectos.	4.23	1	7	1.59
En nuestra organización, los resultados de las decisiones son a menudo sin errores.	4.35	1	7	1.64
En nuestra organización, los resultados de las decisiones son a menudo confiables.	5.42	1	7	1.18
En nuestra organización, el tiempo para llegar a decisiones es rápido.	5.17	1	7	1.33
En nuestra organización, la velocidad para llegar a decisiones es alta.	5.14	1	7	1.39
En su organización ¿se ha incrementado el desempeño financiero con el uso de Analíticos de Datos para la toma de decisiones?	4.02	1	7	1.67
La utilidad neta de la empresa mejoró desde la implementación de Analíticos de Datos en el proceso de toma de decisiones.	5.65	1	7	1.54
Las ventas de la empresa incrementaron desde la implementación de Análisis de Datos en el proceso de toma de decisiones.	5.54	1	7	1.44

Fuente: elaboración propia en RStudio.

Evaluación del modelo PLS-SEM

Como se mencionó anteriormente, el modelo de investigación propuesto para responder a las preguntas de investigación hace referencia a tres hipótesis: a mayores CABD, mejor será el desempeño del proceso de toma de decisiones de la firma; el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la firma; y por último, el desempeño del proceso de toma de decisiones tiene un efecto mediador entre las CABD y el desempeño financiero de la firma. En ese sentido, se propone un modelo de investigación empírico utilizando la metodología PLS-SEM, en el cual se operacionalizan las CABD como un constructo formativo de segundo orden, el desempeño del proceso de toma de decisiones se operacionaliza como un constructo formativo de segundo orden y el desempeño financiero se operacionaliza como un constructo formativo de primer orden.

En la figura 31 se puede observar el modelo de investigación propuesto, desglosando los constructos de primer orden, segundo orden, así como las preguntas o ítems que componen a cada uno de ellos. En ese sentido tenemos que las CABD se componen de cinco constructos de primer orden: calidad de los datos (formativo), tamaño de los datos (formativo), habilidades analíticas (reflectivo), conocimiento de dominio (reflectivo) y sofisticación de las herramientas (formativo). Asimismo, el desempeño del proceso de toma de decisiones de la firma se compone de dos constructos de primer orden reflectivos: calidad de las decisiones y eficiencia de las decisiones. Por último, el desempeño financiero de la empresa es un constructo de primer orden formativo compuesto por ítems relacionados con el desempeño de las ventas y la rentabilidad financiera.

La estimación de cada uno de los constructos, ya sea de primer o segundo orden, se describen en el apartado anterior en las ecuaciones (7) a (19), diferenciando entre el modelo de medición, dónde se evalúa la validez de los ítems o preguntas para los constructos de primer orden, y el modelo estructural dónde se evalúan las relaciones estructurales propuestas entre variables dependientes e independientes, como se puede también observar en la figura 31. A continuación se exponen y se discuten los resultados de la aplicación del modelo de investigación propuesto, primeramente del modelo de medición para constructos de primer orden y posteriormente se evalúa el modelo de medición para los constructos de segundo orden.

Figura 31. Modelo de investigación propuesto



Fuente: elaboración propia.

De acuerdo con Hair et al. (2021), cuando se tienen modelos de investigación PLS-SEM con constructos de segundo orden (o *higher order constructs*), es necesario evaluar en primera instancia la fiabilidad y validez de los constructos de primer orden (o *lower order constructs*), independientemente de si son constructos formativos o reflectivos. Una vez evaluado el modelo de medición de los constructos de primer orden, el segundo paso es evaluar la validez de los constructos de segundo orden, o bien, el modelo de medición de los constructos de segundo orden (Hair et al., 2021).

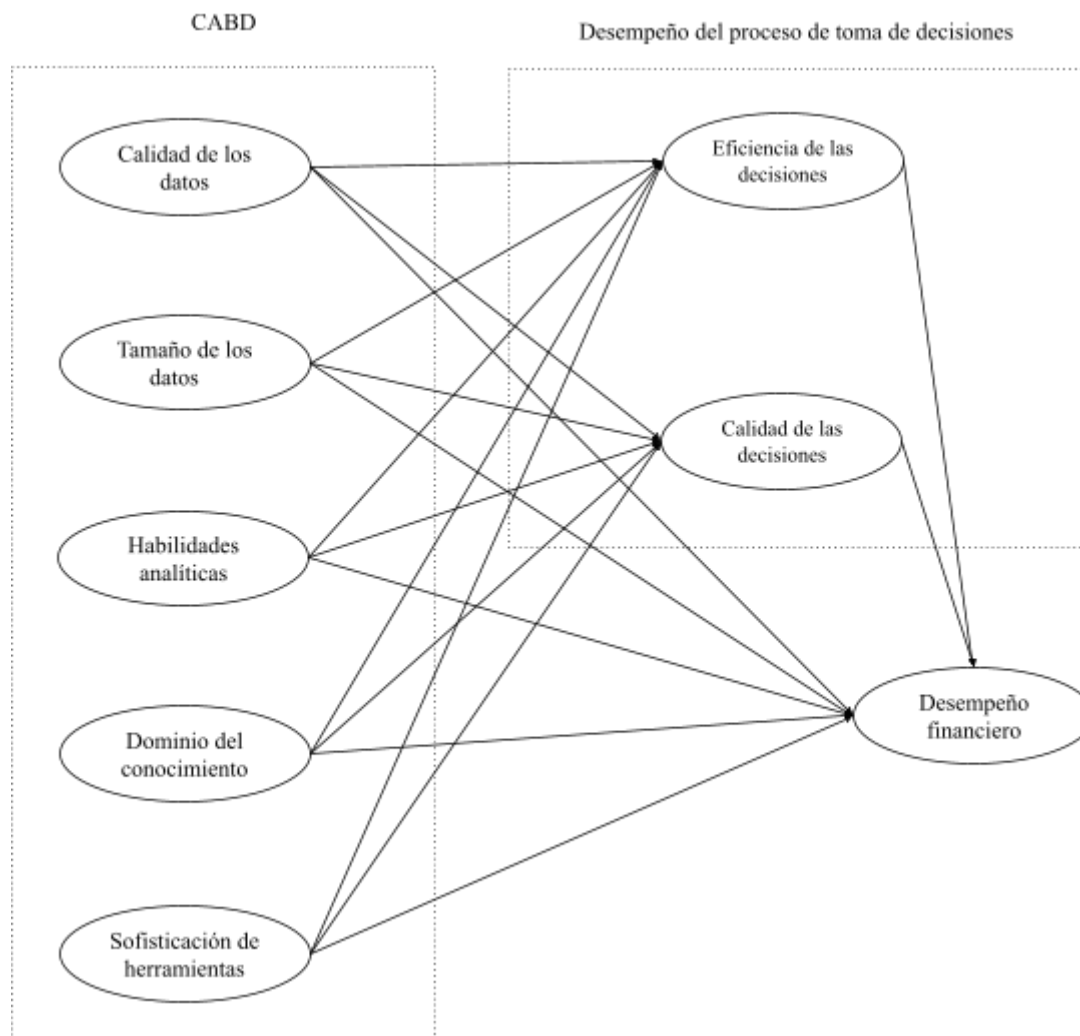
Evaluación del modelo de medición para constructos de primer orden

Teniendo en cuenta lo anterior, procedemos a evaluar el modelo de medición para los constructos de primer orden reflectivos del modelo de investigación propuesto (habilidades analíticas,

conocimiento de dominio, calidad de las decisiones y eficiencia de las decisiones), incluyendo su fiabilidad, validez convergente y validez discriminante. Una vez evaluados los constructos de primer orden reflectivos se procede a evaluar los constructos de primer orden formativos (calidad de los datos, tamaño de los datos, sofisticación de las herramientas y desempeño financiero), verificando su validez de convergencia, colinealidad, significancia estadística y relevancia de los indicadores.

Para poder evaluar los constructos de primer orden, independientemente si son de carácter reflectivo o formativos, es necesario correr un modelo PLS-SEM en donde las relaciones estructurales sean entre todos los constructos de primer orden (Hair et al., 2021). Como se puede observar en la figura 32, cada uno de los constructos de primer orden se relaciona con los constructos que pertenecen a constructos de segundo orden de acuerdo con el modelo estructural de investigación propuesto, por ejemplo, los cinco constructos de primer orden que componen las CABD (calidad de los datos, tamaño de los datos, habilidades analíticas, conocimiento de dominio y sofisticación de las herramientas), se relacionan directamente con los dos constructos de primer orden que componen al constructo de segundo orden desempeño del proceso de toma de decisiones de la empresa, mientras que todos éstos se relaciona a su vez con el constructo de primer orden desempeño financiero.

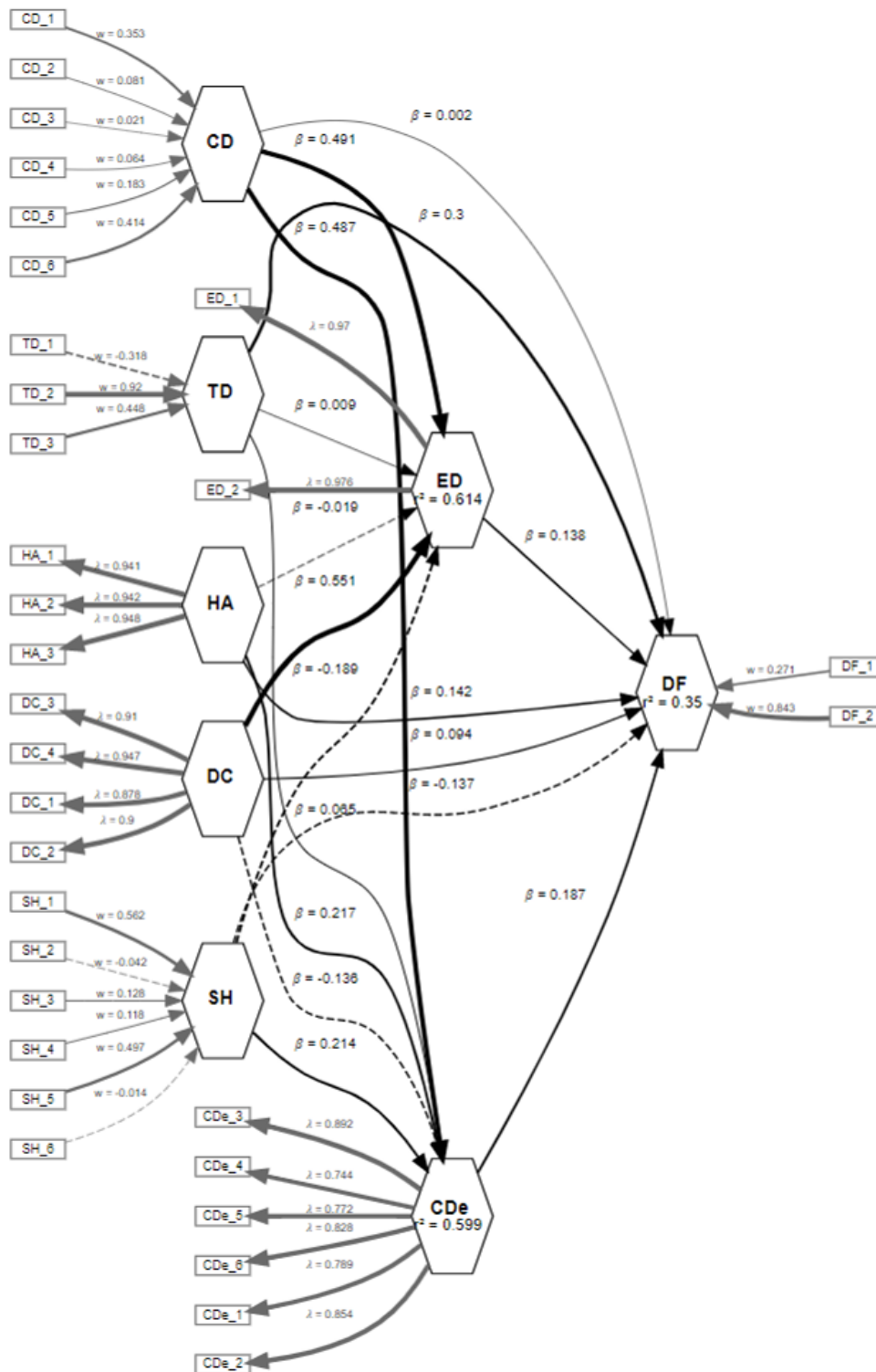
Figura 32. Modelo de medición para evaluar los constructos de primer orden



Fuente: elaboración propia.

En la figura 33 podemos observar los resultados de correr el modelo de medición para los constructos de primer orden, en el que podemos observar los coeficientes de carga o *loadings* de cada uno de los indicadores o preguntas de los constructos reflectivos (habilidades analíticas, dominio del conocimiento, calidad de las decisiones y eficiencia de las decisiones), así como los pesos o *weights* de cada indicador o pregunta que conforman a los constructos formativos (calidad de los datos, tamaño de los datos, sofisticación de las herramientas y desempeño financiero).

Figura 33. Gráfico de resultados del modelo de medición de constructos de primer orden



Fuente: elaboración propia en RStudio. (HA: habilidades analíticas; DC: dominio del conocimiento; CDe: calidad de las decisiones; ED: eficiencia de las decisiones; CD: calidad de los datos; TD: tamaño de los datos; SH: sofisticación de las herramientas; y DF: desempeño financiero.)

Evaluación de constructos reflectivos de primer orden

Como ya se mencionó en la metodología, el proceso para evaluar constructos reflectivos en un modelo PLS-SEM consiste en evaluar la fiabilidad de los indicadores, la fiabilidad de consistencia interna, así como la validez convergente y discriminante de los mismos. En ese sentido, en la tabla 16 se muestran las cargas y cargas al cuadrado de cada uno de los indicadores de los constructos de primer orden del modelo empírico propuesto, lo que nos permite evaluar la fiabilidad de los indicadores. De acuerdo con el umbral de aceptación para las cargas de cada indicador (Hair et al., 2021), todos los indicadores de los constructos reflectivos de primer orden muestran cargas superiores a 0.708, mientras que la varianza explicada por cada indicador, o carga al cuadrado, es superior a 0.5, indicando también la fiabilidad de cada uno de los indicadores de los constructos de primer orden del modelo de medición.

Tabla 16. Fiabilidad de los indicadores de constructos reflectivos de primer orden

Indicador	Cargas				Cargas al cuadrado			
	HA	DC	CDe	ED	HA	DC	CDe	ED
HA_1	0.941	0.000	0.000	0.000	0.886	0.000	0.000	0.000
HA_2	0.942	0.000	0.000	0.000	0.887	0.000	0.000	0.000
HA_3	0.948	0.000	0.000	0.000	0.898	0.000	0.000	0.000
DC_1	0.000	0.879	0.000	0.000	0.000	0.772	0.000	0.000
DC_2	0.000	0.900	0.000	0.000	0.000	0.810	0.000	0.000
DC_3	0.000	0.910	0.000	0.000	0.000	0.829	0.000	0.000
DC_4	0.000	0.947	0.000	0.000	0.000	0.897	0.000	0.000
CDe_1	0.000	0.000	0.789	0.000	0.000	0.000	0.623	0.000
CDe_2	0.000	0.000	0.854	0.000	0.000	0.000	0.730	0.000
CDe_3	0.000	0.000	0.892	0.000	0.000	0.000	0.795	0.000
CDe_4	0.000	0.000	0.744	0.000	0.000	0.000	0.554	0.000
CDe_5	0.000	0.000	0.772	0.000	0.000	0.000	0.596	0.000
CDe_6	0.000	0.000	0.828	0.000	0.000	0.000	0.685	0.000
ED_1	0.000	0.000	0.000	0.970	0.000	0.000	0.000	0.941
ED_2	0.000	0.000	0.000	0.977	0.000	0.000	0.000	0.954

Fuente: elaboración propia en RStudio.

HA: habilidades analíticas.

DC: dominio del conocimiento.

CDe: calidad de las decisiones.

ED: eficiencia de las decisiones.

Como segundo paso, se debe evaluar la fiabilidad de consistencia interna, es decir el grado en que los indicadores que miden al mismo constructo se asocian entre sí, y se mide utilizando el alfa de Cronbach y el coeficiente de fiabilidad compuesta o rho. El criterio de medición consiste en utilizar el alfa de Cronbach como límite inferior y la confiabilidad compuesta de rho_c es el límite superior, mientras que el coeficiente de confiabilidad rho generalmente se encuentra entre estos límites y sirve como una buena representación de la confiabilidad de consistencia interna de un constructo, siendo un indicador confiable cuando dicho valor se encuentra entre 0.8 y 0.95 (Hair et al., 2021). Asimismo, se debe evaluar la validez convergente de cada constructo reflectivo para saber en qué medida converge el constructo para explicar la varianza de los indicadores que le componen, lo cual se mide utilizando la varianza media extraída (AVE por sus siglas en inglés) (Hair et al., 2021).

En la tabla 17 y en la figura 34 podemos observar que los constructos reflectivos de primer orden de nuestro modelo de medición PLS-SEM se encuentran dentro de los parámetros establecidos en la literatura referente de modelos PLS-SEM (Hair et al., 2021), puesto que en todos los casos el rhoA se encuentra dentro del límite inferior (alpha) y el límite superior (rhoC), lo que indica que en todos los casos los constructos reflectivos muestran confiabilidad de consistencia interna. Por otra parte, en lo que respecta a la validez de convergencia, todos los constructos reflectivos de primer orden (habilidades analíticas, dominio del conocimiento, calidad de las decisiones y eficiencia de las decisiones) exhiben una varianza media extraída (AVE) superior a 0.5.

Tabla 17. Fiabilidad de consistencia interna y validez de convergencia de constructos reflectivos de primer orden

Constructo	alpha	rhoC	AVE	rhoA
HA	0.938	0.961	0.890	0.941
DC	0.930	0.950	0.827	0.934
CDe	0.899	0.922	0.664	0.916
ED	0.944	0.973	0.947	0.955

Fuente: elaboración propia en RStudio.
 HA: habilidades analíticas.
 DC: dominio del conocimiento.
 CDe: calidad de las decisiones.
 ED: eficiencia de las decisiones.

Figura 34. Fiabilidad de consistencia interna y validez de convergencia de constructos reflectivos de primer orden



Fuente: elaboración propia en RStudio.
 HA: habilidades analíticas.
 DC: dominio del conocimiento.
 CDe: calidad de las decisiones.
 ED: eficiencia de las decisiones.

Finalmente, se debe evaluar la validez discriminante de los constructos, es decir, medir hasta qué punto un constructo es empíricamente distinto de otros constructos del modelo estructural (Hair et al., 2021). Para ello se utiliza la razón heterotrazo-monotrazo (HTMT por sus siglas en inglés), se define como el valor medio de las correlaciones de los indicadores entre constructos, en relación con la media geométrica de las correlaciones medias de los indicadores que miden el mismo constructo, y debe ser menor a 0.9 para constructos conceptualmente similares y menor a 0.85 para constructos conceptualmente diferentes (Hair et al., 2021). En la

tabla 18 podemos observar que en todos los constructos la razón heterotrazo-monotrazo (HTMT) es menor a 0.85, inclusive en constructos de primer orden que son similares, como es el caso de la calidad de las decisiones y la eficiencia de las decisiones.

Tabla 18. Validez discriminante de constructos reflectivos de primer orden (HTMT).

	HA	DC	CDe	ED
HA				
DC	0.790			
CDe	0.624	0.481		
ED	0.642	0.742	0.532	

Fuente: elaboración propia en RStudio.

HA: habilidades analíticas.

DC: dominio del conocimiento.

CDe: calidad de las decisiones.

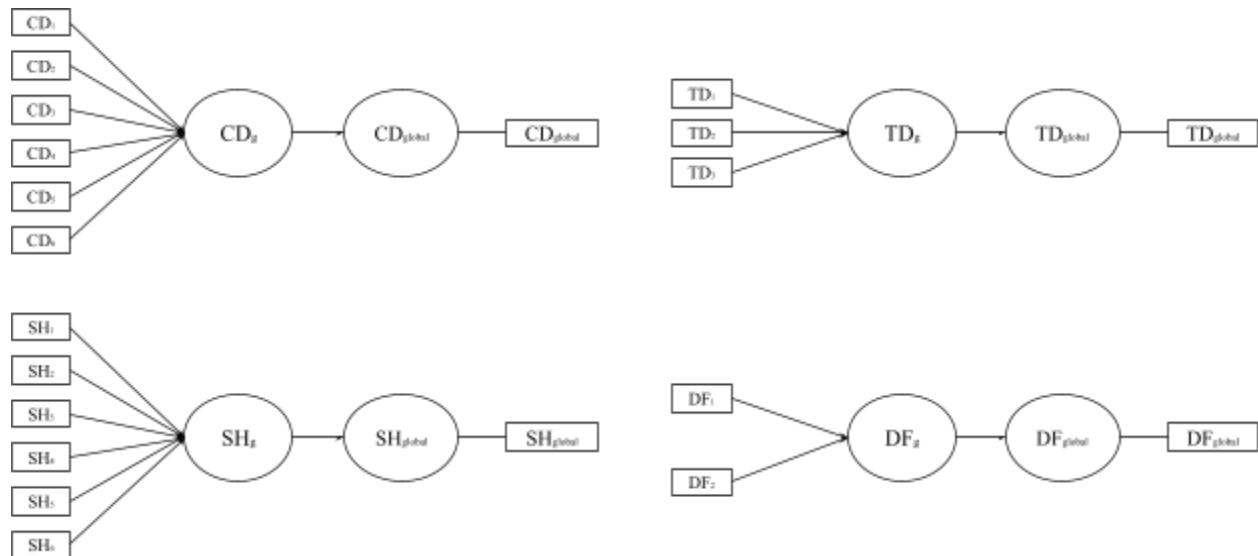
ED: eficiencia de las decisiones.

De acuerdo con lo anterior, se puede concluir que los constructos reflectivos de primer orden, habilidades analíticas, dominio del conocimiento, calidad de las decisiones y eficiencia de las decisiones, cumplen con los criterios de evaluación establecidos en la literatura pertinente relacionada con modelos PLS-SEM, por lo que se consideran constructos válidos a incluir en el modelo empírico PLS-SEM a utilizar para probar las hipótesis de investigación planteadas.

Evaluación de constructos formativos de primer orden

En el caso de los constructos formativos de primer orden, como primer paso se debe evaluar la validez de convergencia, utilizando un análisis de redundancia incluyendo un ítem o pregunta que mida de forma global el constructo formativo como alternativa a cada una de los ítems o preguntas que lo conforman (Hair et al., 2021). En la literatura revisada se sugiere que la correlación del constructo formativo medido a través de indicadores reflectivos (el caso de un solo indicador o pregunta se considera reflectivo) debe ser superior a 0.708, lo que implica que dicho constructo explica en más del 50% de la varianza de la medida alterna o indicador global para el constructo (Hair et al., 2021). El análisis de redundancia involucra el diseño de un modelo de medida y estructural para cada uno de los constructos formativos incluidos en el modelo (calidad de los datos, tamaño de los datos, sofisticación de las herramientas y desempeño financiero), es decir, modelos PLS-SEM para poder evaluar la validez de convergencia, los cuáles se muestran en la figura 35.

Figura 35. Modelos PLS-SEM para análisis de redundancia de constructos formativos de primer orden



Fuente: elaboración propia.

CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

DF: desempeño financiero.

En la tabla 19 se muestran los resultados de cada uno de los modelos PLS-SEM definidos para evaluar la validez de convergencia de los constructos formativos de primer orden, haciendo énfasis en los coeficientes de trayectoria o *paths* para cada uno de los modelos. Como podemos observar, en todos los constructos formativos de primer orden se observa un coeficiente de trayectoria superior a 0.708, indicando que el constructo explica más del 50% de la varianza del indicador formativo o medida global alterna para el constructo (Cheah et al., 2018 y Hair et al., 2021). Específicamente, tenemos que para el constructo calidad de los datos, el coeficiente de trayectoria es 0.980, para desempeño financiero es de 0.864, para sofisticación de las herramientas es de 0.971 y para tamaño de los datos es de 0.975, proporcionando soporte de la validez de convergencia para todos los constructos formativos de primer orden.

Tabla 19. Resultados de modelos PLS-SEM para análisis de redundancia de constructos formativos de primer orden

CD_global	
R ²	0.961
AdjR ²	0.960
CDg	0.980
DF_global	
R ²	0.746
AdjR ²	0.743
DFg	0.864
SH_global	
R ²	0.944
AdjR ²	0.943
SHg	0.971
TD_global	
R ²	0.950
AdjR ²	0.950
TDg	0.975

Fuente: elaboración propia en RStudio.

CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

DF: desempeño financiero.

Posteriormente, se debe evaluar la colinealidad de los constructos formativos de primer orden que se incluyen en el modelo. La colinealidad se refiere a que dos o más indicadores en un constructo formativo del modelo de medición están altamente correlacionados y se mide a través del factor de inflación de la varianza (VIF por sus siglas en inglés), a mayor VIF, mayor el nivel de colinealidad, particularmente, valores superiores a 5 (Hair et al., 2021). En la tabla 20 se muestran los resultados para los constructos formativos de primer orden a incluir en el modelo, y dónde podemos observar que en todos los casos el factor de inflación de la varianza (VIF) de cada uno de los indicadores respecto al indicador global del constructo es inferior a 5, indicando que no existen problemas de colinealidad.

Tabla 20. Evaluación de colinealidad para los constructos formativos de primer orden

CDg :					
CD_1	CD_2	CD_3	CD_4	CD_5	CD_6
3.926	4.331	3.932	2.059	2.326	4.612
DFg :					
DF_1	DF_2				
1.286	1.286				
SHg :					
SH_1	SH_2	SH_3	SH_4	SH_5	SH_6
2.06	2.202	1.218	1.745	2.62	2.692
TDg :					
TD_1	TD_2	TD_3			
1.9	1.516	1.566			

Fuente: elaboración propia en RStudio.

CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

DF: desempeño financiero.

Finalmente, se debe evaluar el peso de los indicadores en términos de su significancia y relevancia. Primeramente, se debe evaluar la significancia estadística de los indicadores realizando un proceso de *Bootstrapping* a cada modelo PLS-SEM especificado por constructo formativo de primer orden (Hair et al., 2021). En la tabla 21 se muestra la significancia del peso de los indicadores para cada modelo PLS-SEM con 10,000 muestras *bootstrap* incluyendo el peso de los indicadores, así como la significancia estadística de cada uno. Como se puede observar, con excepción de los indicadores CD_4 en el constructo calidad de los datos y el indicador SH_2, SH_4 y SH_6 en el constructo sofisticación de las herramientas, todos los indicadores resultan ser significativos en un intervalo de confianza del 95% (prueba de dos colas). Mientras que en la figura 36 se muestran los resultados de forma gráfica, indicando la relación de cada indicador con su constructo formativo de primer orden y sus pesos, así como su significancia estadística.

Tabla 21. Significancia del peso de los indicadores en constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CD_1 -> CDg	0.402	0.397	0.058	6.911	0.281	0.513
CD_2 -> CDg	-0.118	-0.118	0.054	-2.186	-0.222	-0.010
CD_3 -> CDg	0.450	0.452	0.058	7.717	0.336	0.566
CD_4 -> CDg	0.004	0.013	0.039	0.095	-0.059	0.100
CD_5 -> CDg	0.239	0.236	0.056	4.251	0.125	0.350
CD_6 -> CDg	0.121	0.118	0.056	2.144	0.005	0.228
CD_global -> CD_global	1.000	1.000	0.000	.	1.000	1.000
TD_1 -> TDg	0.470	0.470	0.043	10.831	0.383	0.552
TD_2 -> TDg	0.424	0.426	0.050	8.431	0.330	0.527
TD_3 -> TDg	0.304	0.302	0.053	5.728	0.198	0.404
TD_global -> TD_global	1.000	1.000	0.000	.	1.000	1.000
SH_1 -> SHg	0.325	0.327	0.045	7.165	0.239	0.417
SH_2 -> SHg	0.064	0.065	0.043	1.507	-0.019	0.150
SH_3 -> SHg	0.596	0.600	0.051	11.805	0.507	0.705
SH_4 -> SHg	-0.099	-0.103	0.056	-1.791	-0.214	0.002
SH_5 -> SHg	0.375	0.379	0.048	7.741	0.289	0.480
SH_6 -> SHg	0.073	0.071	0.035	2.073	-0.006	0.134
SH_global -> SH_global	1.000	1.000	0.000	.	1.000	1.000
DF_1 -> DFg	0.630	0.621	0.086	7.346	0.433	0.770
DF_2 -> DFg	0.534	0.541	0.069	7.786	0.408	0.677
DF_global -> DF_global	1.000	1.000	0.000	.	1.000	1.000

Fuente: elaboración propia en RStudio.

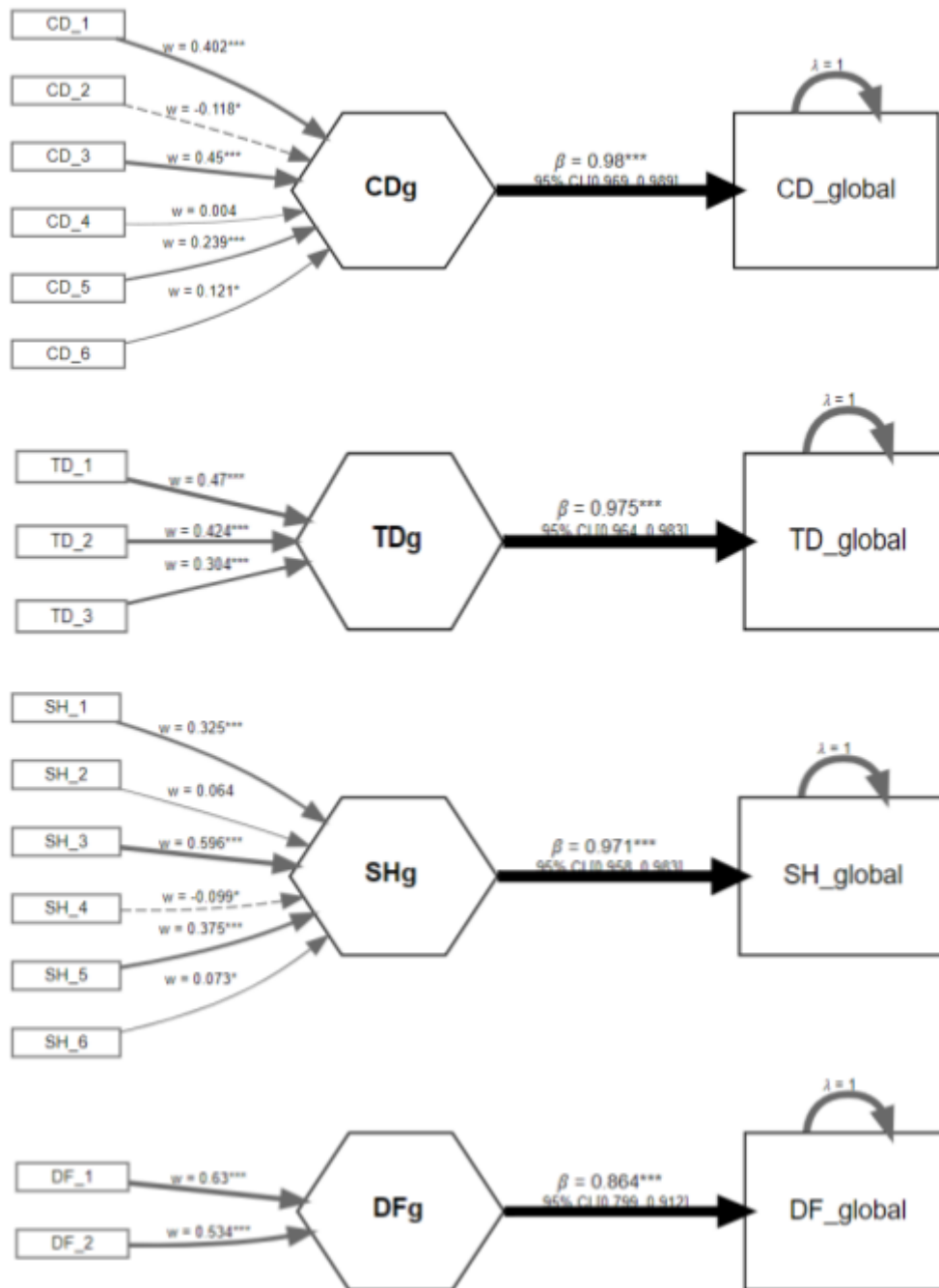
CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

DF: desempeño financiero.

Figura 36. Pesos y prueba de significancia estadística de indicadores para constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*).



Fuente: elaboración propia en RStudio.

CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

DF: desempeño financiero.

Los indicadores que resultan no significativos estadísticamente deben ser considerados para eliminarse del modelo, sin embargo, previo a esto se debe tener en cuenta la contribución

absoluta del indicador, lo cuál se mide a través de las cargas de cada indicador: si la carga resulta superior al 0.5, y es estadísticamente significativa, el indicador se debe mantener en el modelo (Hair et al., 2021). En la tabla 22 se muestran las cargas de los indicadores de cada constructo formativo, utilizando los modelos PLS-SEM con *bootstrap*, y podemos observar que las cargas de todos los indicadores son superiores a 0.5 y son estadísticamente significativas, confirmando que todos los indicadores se pueden mantener en el constructo formativo al que pertenecen.

Tabla 22. Cargas y significancia estadística de los indicadores en constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CD_1 -> CDg	0.919	0.916	0.025	36.207	0.856	0.957
CD_2 -> CDg	0.819	0.814	0.053	15.354	0.692	0.900
CD_3 -> CDg	0.948	0.945	0.019	48.962	0.900	0.975
CD_4 -> CDg	0.681	0.680	0.106	6.437	0.455	0.862
CD_5 -> CDg	0.810	0.810	0.056	14.522	0.693	0.908
CD_6 -> CDg	0.864	0.858	0.043	20.122	0.758	0.925
CD_global -> CD_global	1.000	1.000	0.000	.	1.000	1.000
TD_1 -> TDg	0.892	0.890	0.026	34.171	0.833	0.934
TD_2 -> TDg	0.823	0.823	0.041	20.065	0.733	0.893
TD_3 -> TDg	0.763	0.769	0.057	13.418	0.653	0.876
TD_global -> TD_global	1.000	1.000	0.000	.	1.000	1.000
SH_1 -> SHg	0.629	0.625	0.072	8.766	0.475	0.754
SH_2 -> SHg	0.680	0.677	0.058	11.731	0.552	0.780
SH_3 -> SHg	0.777	0.777	0.039	20.005	0.696	0.848
SH_4 -> SHg	0.507	0.506	0.091	5.569	0.314	0.669
SH_5 -> SHg	0.758	0.753	0.053	14.240	0.638	0.845
SH_6 -> SHg	0.744	0.740	0.052	14.301	0.628	0.831
SH_global -> SH_global	1.000	1.000	0.000	.	1.000	1.000
DF_1 -> DFg	0.882	0.881	0.038	23.399	0.799	0.947
DF_2 -> DFg	0.832	0.832	0.072	11.630	0.672	0.951
DF_global -> DF_global	1.000	1.000	0.000	.	1.000	1.000

Fuente: elaboración propia en RStudio.

CD: calidad de los datos.

TD: tamaño de los datos.

SH: sofisticación de las herramientas.

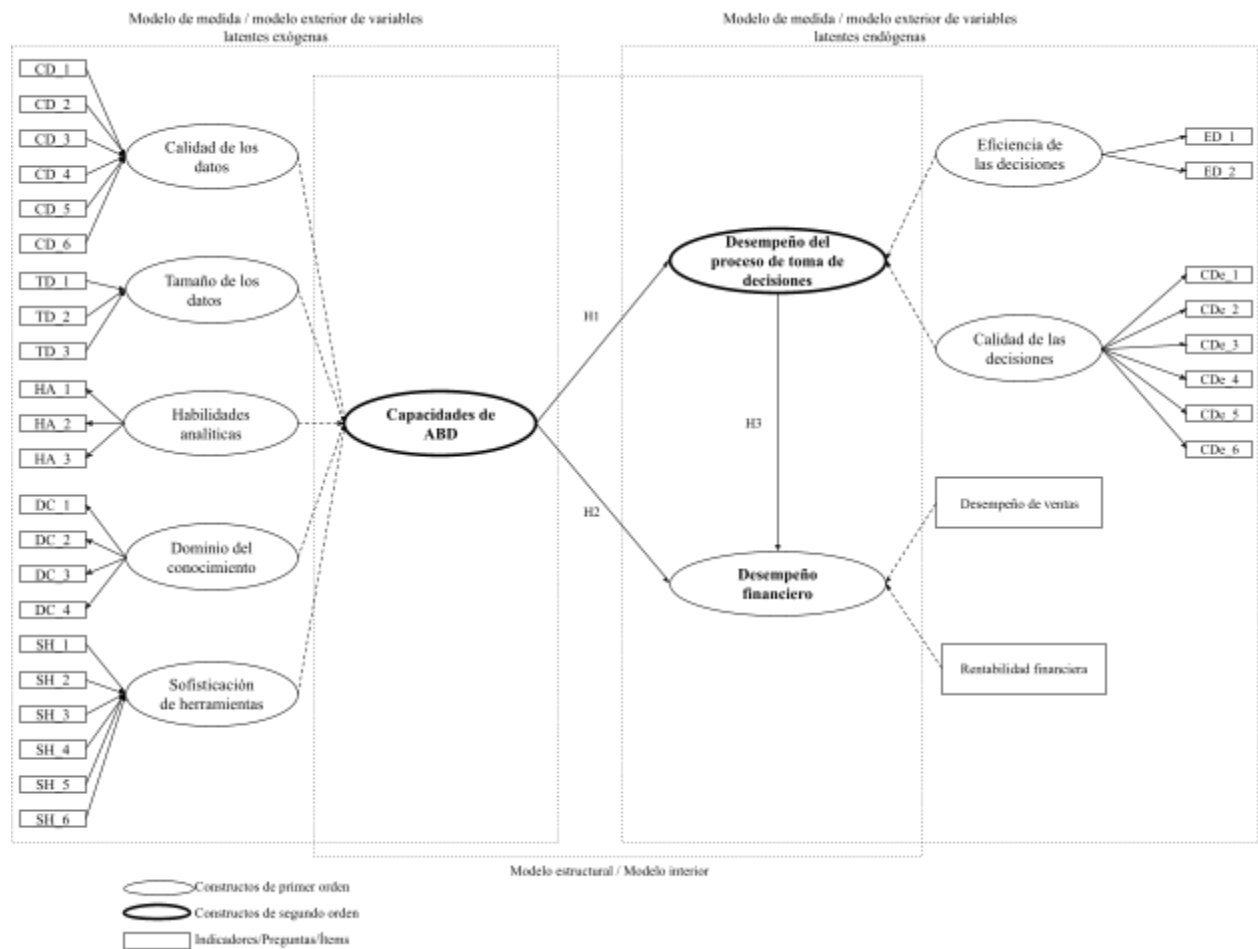
DF: desempeño financiero.

Teniendo en cuenta los resultados obtenidos, podemos incluir en modelo de medición que incluye constructos de segundo orden todos los constructos formativos de primer orden: calidad de los datos, tamaño de los datos, sofisticación de las herramientas y desempeño financiero. Esto pues de acuerdo con los lineamientos de la literatura revisada sobre la aplicación de modelos PLS-SEM, concluyendo que todos los indicadores de los constructos formativos de primer orden son válidos y se deben incluir en el modelo empírico PLS-SEM propuesto para probar las hipótesis de investigación.

Evaluación del modelo de medición para constructos de segundo orden

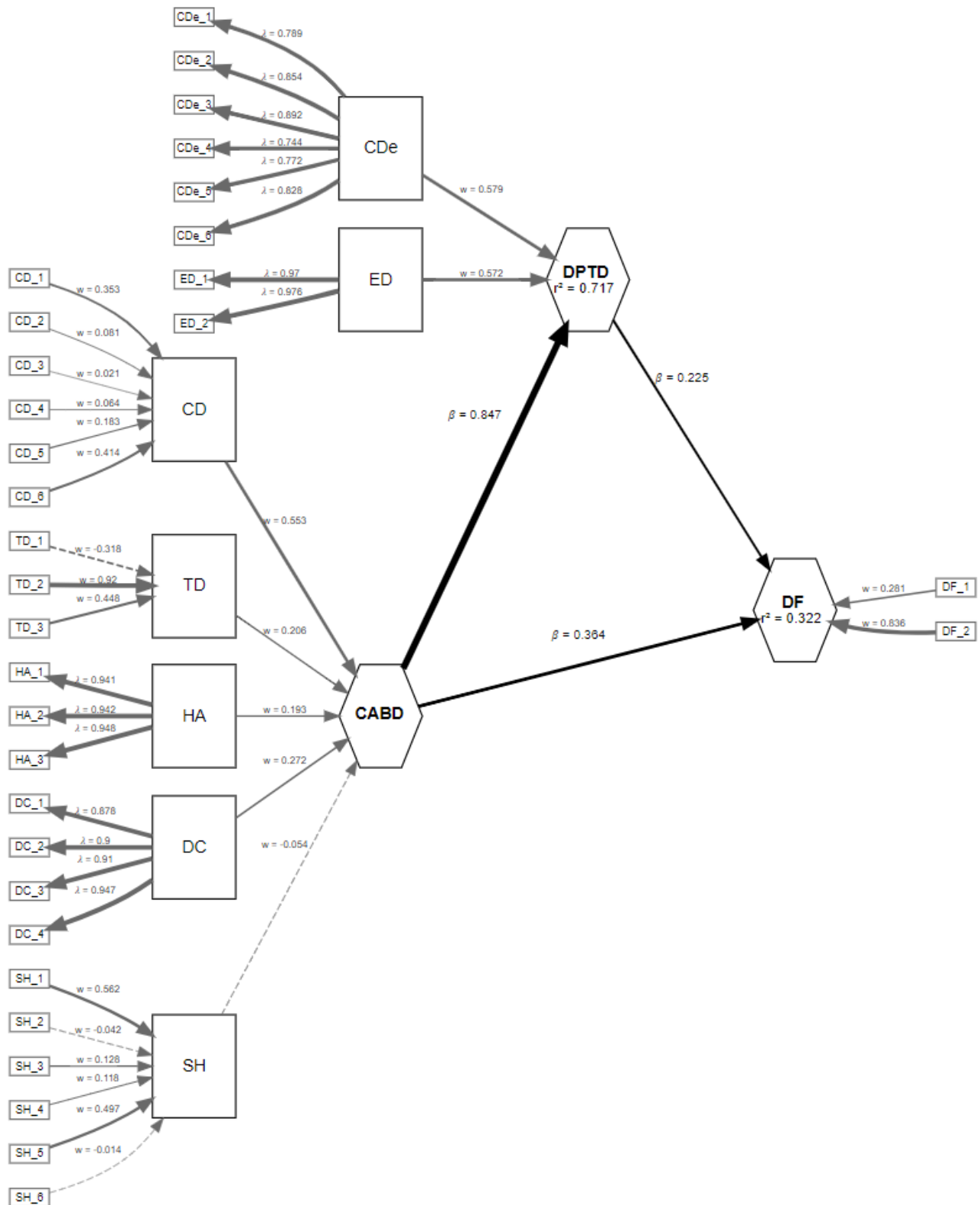
Una vez evaluada la confiabilidad y validez de todos los constructos de primer orden, independientemente de si son de naturaleza reflectiva o formativa, se procede a evaluar la confiabilidad y validez de los constructos de segundo orden, incluyendo los constructos de primer orden involucrados en la especificación del modelo. En la figura 37 tenemos el modelo de investigación propuesto, donde se especifica la naturaleza y orden de cada constructo, así como el modelo de medición y estructural de forma general. Mientras que en la figura 38 se muestran los resultados del modelo PLS-SEM que incluye los constructos de primer orden (formativos y reflectivos) ya evaluados, así como los constructos de segundo orden. En ella podemos observar los resultados que se discuten en los siguientes párrafos.

Figura 37. Modelo de investigación propuesto.



Fuente: elaboración propia.

Figura 38. Resultados del modelo PLS-SEM

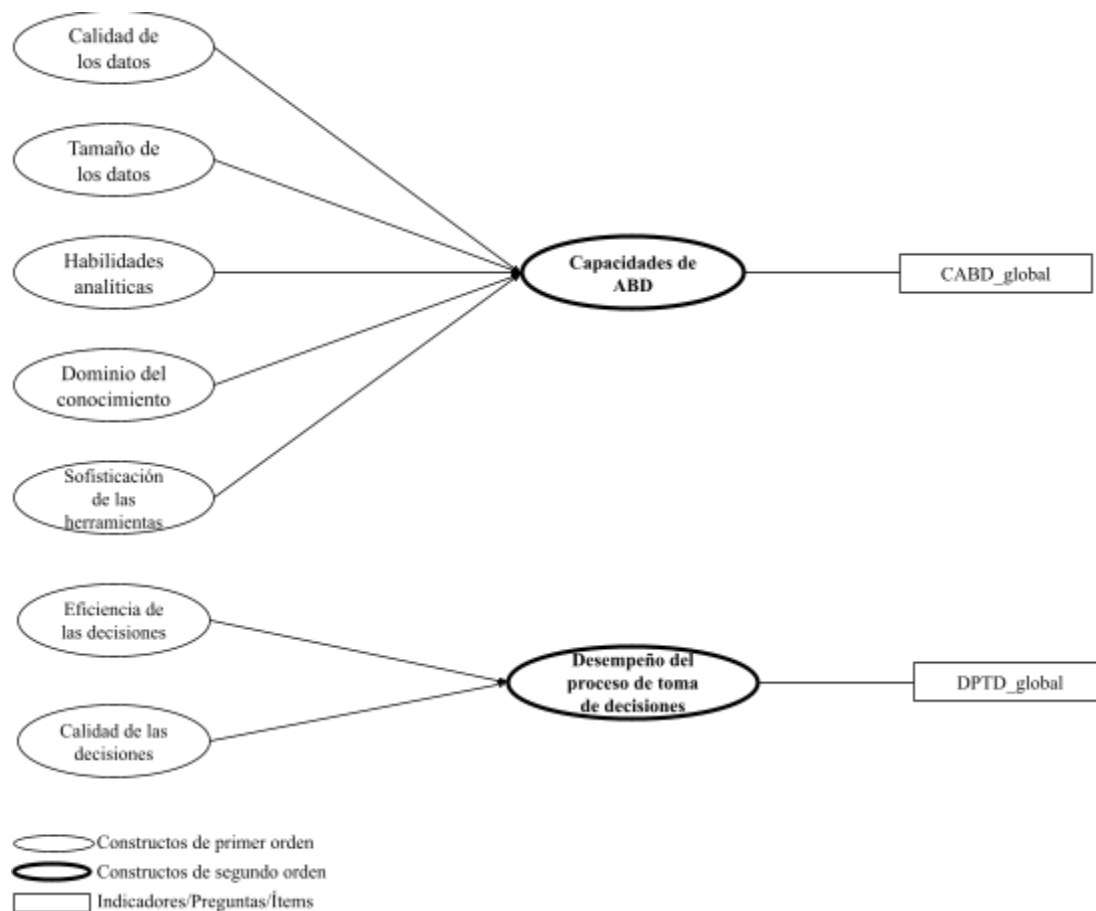


Fuente: elaboración propia en RStudio. (CABD: capacidades de analíticos de Big Data; DPTD: desempeño del proceso de toma de decisiones; CD: calidad de los datos; TD: tamaño de los datos; HA: habilidades analíticas; DC:

dominio del conocimiento; SH: sofisticación de las herramientas; CDe: calidad de las decisiones; ED: eficiencia de las decisiones; y DF: desempeño financiero).

Al igual que con los constructos formativos de primer orden, el primer paso consiste en evaluar la validez de convergencia para los constructos formativos de segundo orden, en éste caso, CABD y DPTD, puesto que DF es un constructo de primer orden y ya fue evaluado. Para evaluarlos, es necesario realizar un análisis de redundancia, dónde ahora los constructos de primer orden son tratados como los indicadores o ítems que se evalúan contra el indicador o pregunta que mide de forma global el constructo formativo de segundo orden (Hair et al., 2021). Como se ha mencionado, la literatura sugiere que la correlación del constructo formativo de segundo orden medido a través de indicadores reflectivos (el caso de un solo indicador o pregunta se considera reflectivo) debe ser superior a 0.708, lo que implica que dicho constructo explica en más del 50% de la varianza de la medida alterna o indicador global para el constructo (Hair et al., 2021). El análisis de redundancia involucra el diseño de un modelo de medida y estructural para cada uno de los constructos formativos incluidos en el modelo (calidad de los datos, tamaño de los datos, sofisticación de las herramientas y desempeño financiero), es decir, modelos PLS-SEM para poder evaluar la validez de convergencia, los cuáles se muestran en la figura 39.

Figura 39. Modelos PLS-SEM para análisis de redundancia de constructos formativos de segundo orden



Fuente: elaboración propia.

Una vez ejecutados los modelos PLS-SEM para los constructos formativos de segundo orden, capacidades de analíticos de Big Data (CABD) y desempeño del proceso de toma de decisiones (DPTD), se revisan los coeficientes de trayectoria o *paths* para evaluar si los constructos formativos de primer orden explican más del 50% de la varianza del indicador formativo o medida global del constructo, es decir exhiben un coeficiente de trayectoria superior a 0.708 (Cheah et al., 2018 y Hair et al., 2021). Los resultados de ambos modelos se muestran en la tabla 23 y se puede ver que en ambos casos se soporta la validez de convergencia para los constructos formativos de segundo orden, específicamente, en el caso de las CABD se obtuvo un coeficiente de 0.937, mientras que el desempeño del proceso de toma de decisiones muestra un coeficiente de 0.963, ambos por encima del umbral de 0.708, indicando validez de convergencia en ambos constructos.

Tabla 23. Resultados de modelos PLS-SEM para análisis de redundancia de constructos formativos de segundo orden

CABD_global	
R ²	0.877
AdjR ²	0.876
CABDg	0.937
DPTD_global	
R ²	0.927
AdjR ²	0.926
DPTDg	0.963

Fuente: elaboración propia en RStudio.

CABD_global: constructo reflectivo global las capacidades de analíticos de Big Data.

CABDg: constructo formativo de las capacidades de analíticos de Big Data.

DPTD_global: constructo reflectivo global el desempeño del proceso de toma de decisiones.

DPTDg: constructo formativo del desempeño del proceso de toma de decisiones.

Al ser constructos de naturaleza formativa, se debe evaluar la colinealidad de dichos constructos, independientemente si son de primer o segundo orden. Como ya se mencionó, la colinealidad se refiere a la correlación entre los indicadores que conforman a un constructo y se mide a través del VIF, el cual debe ser menor a 5 para poder rechazar cualquier problema de colinealidad (Hair et al., 2021). En la tabla 24 tenemos que el factor de inflación de la varianza (VIF) es menor a 5 en todos los indicadores (o constructos de primer orden) que conforman a los constructos de segundo orden CABD y DPTD, indicando la inexistencia de problemas de colinealidad.

Tabla 24. Evaluación de colinealidad para los constructos formativos de segundo orden

CABD :				
CD	TD	HA	DC	SH
2.676	2.622	2.644	2.46	2.989
DPTD :				
CDe	ED			
1.324	1.324			

Fuente: elaboración propia en RStudio.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

CD: calidad de los datos.

TD: tamaño de los datos.

HA: habilidades analíticas.

DC: dominio del conocimiento.

SH: sofisticación de las herramientas.

CDe: calidad de las decisiones.

ED: eficiencia de las decisiones.

Como último paso de la evaluación de los constructos formativos de segundo orden se debe evaluar el peso de los indicadores, o constructos de primer orden, en términos de significancia y relevancia, iniciando con un proceso de *Bootstrapping* a cada modelo PLS-SEM especificado para los constructos formativos de segundo orden, de acuerdo con lo establecido en la literatura revisada (Hair et al., 2021). Se aplicó un *Bootstrapping* de 10,000 muestras para cada modelo, cuyos resultados se observan en la tabla 25, donde se muestran los pesos de cada indicador, así como su significancia estadística, teniendo que solamente el indicador o constructo de primer orden habilidades analíticas (HA) es estadísticamente no significativo en un intervalo de confianza del 95% (prueba de dos colas). En la figura 40 se muestran los resultados de forma gráfica, indicando la relación de cada indicador con su constructo formativo de primer orden y sus pesos, así como su significancia estadística.

Tabla 25. Significancia del peso de los indicadores en constructos formativos de segundo orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CD -> CABDg	0.393	0.384	0.075	5.265	0.234	0.530
TD -> CABDg	0.251	0.242	0.060	4.206	0.125	0.359
HA -> CABDg	0.105	0.108	0.062	1.705	-0.012	0.232
DC -> CABDg	-0.124	-0.128	0.058	-2.122	-0.244	-0.014
SH -> CABDg	0.468	0.478	0.068	6.859	0.347	0.615
CABD_global -> CABD_global	1.000	1.000	0.000	NA	1.000	1.000
CDe -> DPTDg	0.786	0.788	0.034	23.463	0.720	0.851
ED -> DPTDg	0.341	0.341	0.034	10.135	0.272	0.405
DPTD_global -> DPTD_global	1.000	1.000	0.000	NA	1.000	1.000

Fuente: elaboración propia en RStudio.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

CD: calidad de los datos.

TD: tamaño de los datos.

HA: habilidades analíticas.

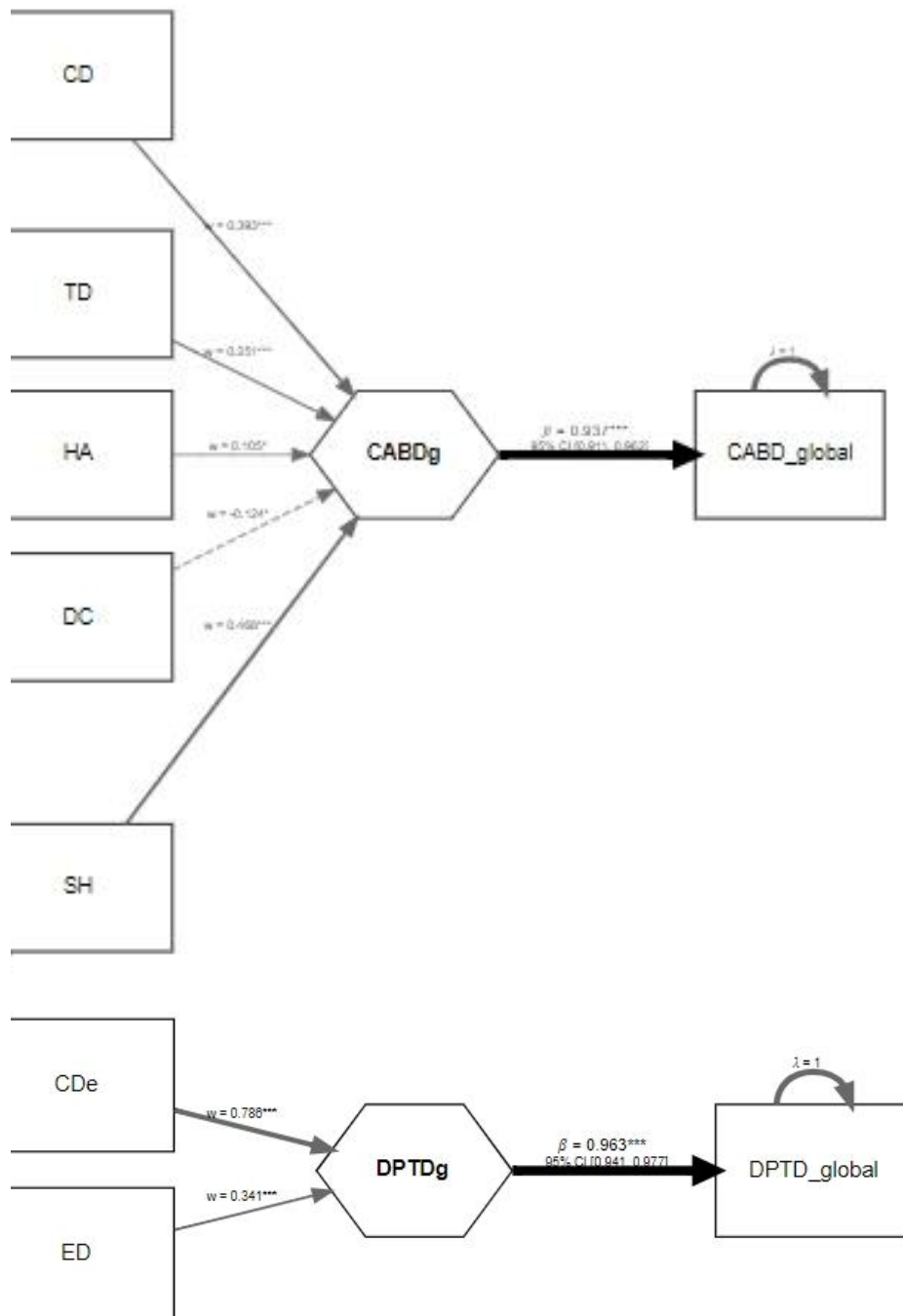
DC: dominio del conocimiento.

SH: sofisticación de las herramientas.

CDe: calidad de las decisiones.

ED: eficiencia de las decisiones.

Figura 40. Pesos y prueba de significancia estadística de indicadores para constructos formativos de primer orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*).



Fuente: elaboración propia en RStudio. (CABD: capacidades de analíticos de Big Data; DPTD: desempeño del proceso de toma de decisiones; CD: calidad de los datos; TD: tamaño de los datos; HA: habilidades analíticas; DC: dominio del conocimiento; SH: sofisticación de las herramientas; CDe: calidad de las decisiones; ED: eficiencia de las decisiones; y DF: desempeño financiero).

A pesar de que el indicador o constructo de primer orden habilidades analíticas (HA) resulta no significativo estadísticamente, es importante revisar la contribución absoluta del constructo antes de considerar eliminarlo del modelo. La contribución absoluta de los indicadores, o constructos de primer orden en éste caso, debe ser superior a 0.5 y ser significativo estadísticamente para que se puedan mantener en el modelo especificado (Hair et al., 2021). En la tabla 26 se puede confirmar que, a pesar de que el peso del indicador o constructo de primer orden habilidades analíticas (HA) resultó ser no significativo, la carga es superior a 0.5 es estadísticamente significativa, por lo que el constructo de primer orden si contribuye al constructo de segundo orden de forma absoluta, por lo que se puede mantener en el modelo.

Tabla 26. Cargas y significancia estadística de los indicadores en constructos formativos de segundo orden (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CD -> CABDg	0.892	0.898	0.034	26.470	0.820	0.951
TD -> CABDg	0.871	0.872	0.031	28.494	0.803	0.923
HA -> CABDg	0.667	0.667	0.117	5.714	0.410	0.848
DC -> CABDg	0.577	0.586	0.152	3.790	0.269	0.820
SH -> CABDg	0.923	0.926	0.023	39.594	0.874	0.965
CABD_global -> CABD_global	1.000	1.000	0.000	NA	1.000	1.000
CDe -> DPTDg	0.955	0.955	0.012	76.949	0.928	0.976
ED -> DPTDg	0.730	0.728	0.066	11.081	0.587	0.843
DPTD_global -> DPTD_global	1.000	1.000	0.000	NA	1.000	1.000

Fuente: elaboración propia en RStudio.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

CD: calidad de los datos.

TD: tamaño de los datos.

HA: habilidades analíticas.

DC: dominio del conocimiento.

SH: sofisticación de las herramientas.

CDe: calidad de las decisiones.

ED: eficiencia de las decisiones.

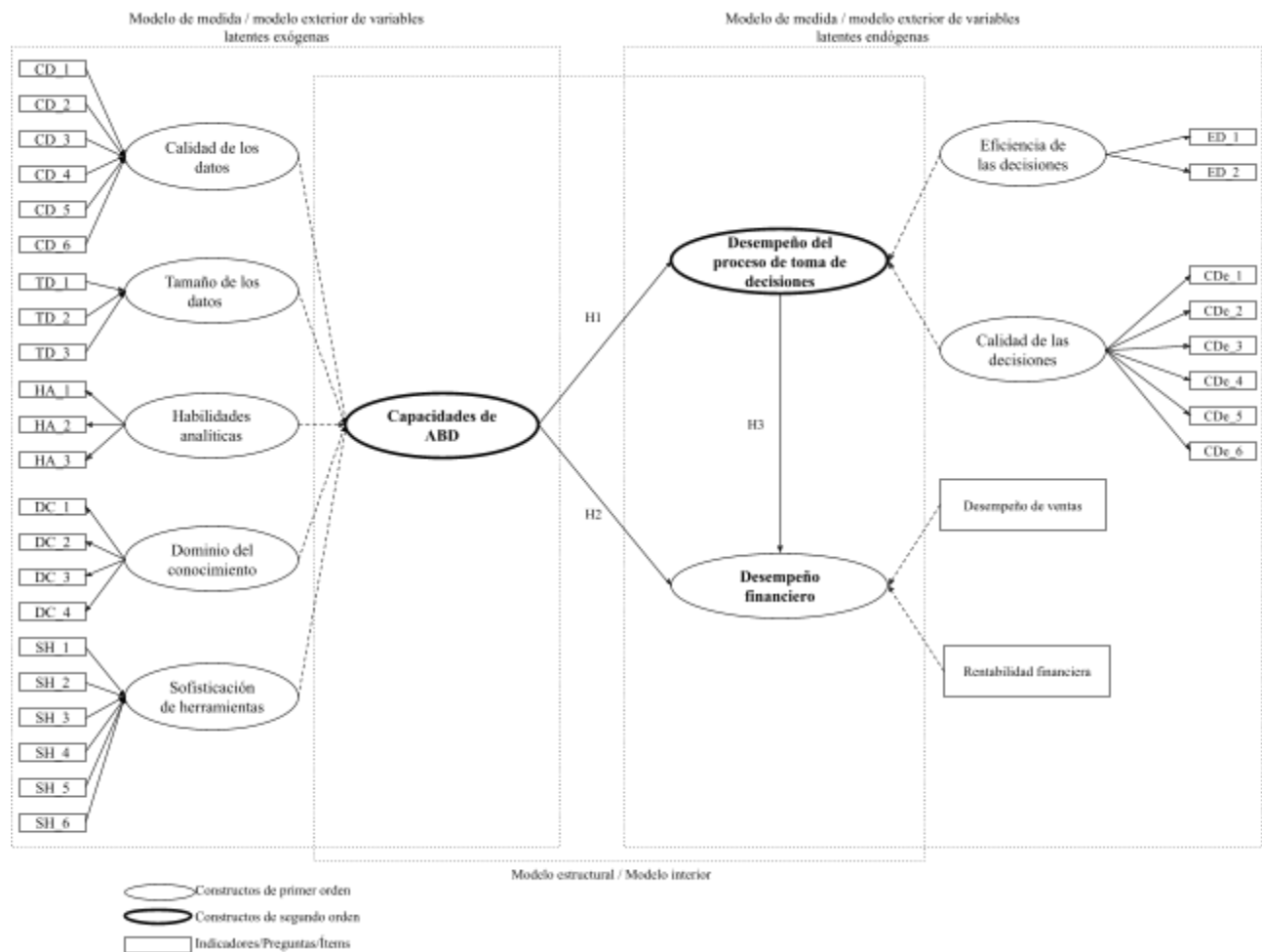
Considerando los resultados obtenidos, los constructos formativos de segundo orden, capacidades de analíticos de Big Data (CABD) y desempeño del proceso de toma de decisiones

(DPTD) pueden ser incluidos en el modelo estructural PLS-SEM a utilizar para responder las preguntas de investigación planteadas y probar las hipótesis propuestas.

Evaluación del modelo estructural

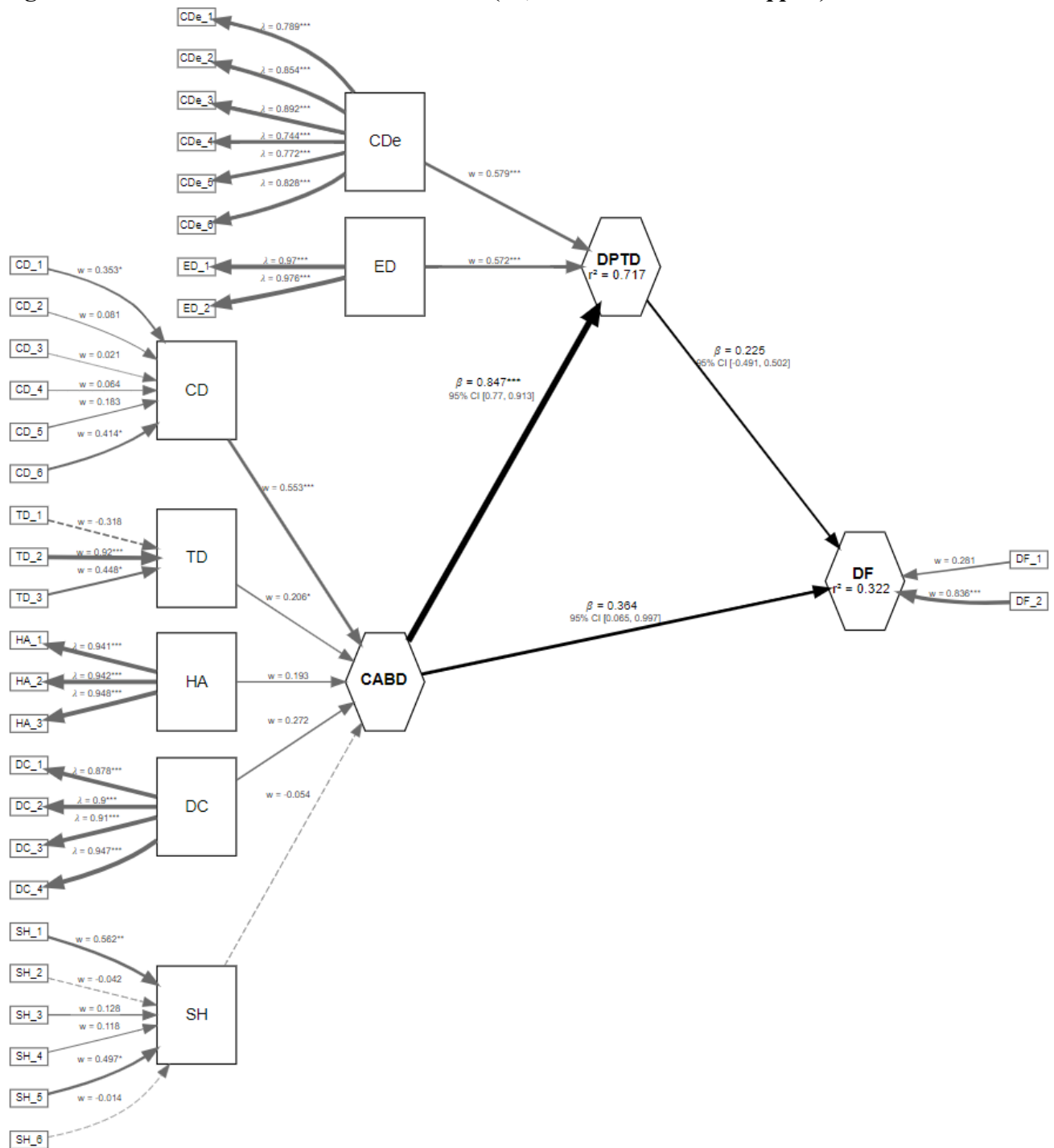
Una vez evaluados los constructos de primer y segundo orden, se procede a evaluar el modelo PLS-SEM propuesto para probar las hipótesis de investigación. Específicamente, probar si las capacidades de analíticos de Big Data (CABD) tienen un impacto positivo en el desempeño de toma de decisiones (DPTD) y el desempeño financiero (DF) de la empresa, así como el efecto mediador del DPTD en el DF. En la primera parte evaluaremos el modelo estructural para evaluar la relación entre las CABD y el DPTD, así como la relación entre las CABD y el DF, para posteriormente evaluar un modelo donde el DPTD actúe como mediador entre las CABD y el DF de la empresa. En la figura 41 se muestra el modelo de investigación propuesta, donde podemos ver que hemos evaluado los modelos de medición de los constructos de primer y segundo orden o modelos exteriores, por lo que ahora se procede a evaluar el modelo interior o modelo estructural.

Figura 41. Modelo de investigación propuesto.



Fuente: elaboración propia.

Figura 42. Resultados del modelo PLS-SEM (10,000 muestras *bootstrapped*)



Fuente: elaboración propia en RStudio. (CABD: capacidades de analíticos de Big Data; DPTD: desempeño del proceso de toma de decisiones; CD: calidad de los datos; TD: tamaño de los datos; HA: habilidades analíticas; DC: dominio del conocimiento; SH: sofisticación de las herramientas; CDe: calidad de las decisiones; ED: eficiencia de las decisiones; y DF: desempeño financiero).

En la figura 42 se muestran los resultados del modelo PLS-SEM propuesto obtenidos utilizando el paquete SEMinR del software RStudio para realizar todos los cálculos del modelo PLS-SEM. En él se pueden observar todas las relaciones estructurales, así como los resultados entre las relaciones de constructos de primer y segundo orden, las cuáles fueron evaluadas de forma separada en las secciones anteriores. De igual forma podemos observar que efectivamente existe un impacto de las CABD en el DPTD de las empresas, así como entre las CABD y el DF, sin embargo, el impacto del DPTD en el DF (con y sin efecto mediador como veremos más adelante) resulta ser no significativo.

Una vez que se evalúan los modelos de medición de los constructos de primer y segundo orden, se debe evaluar el modelo estructural por posibles problemas de colinealidad, significancia y relevancia de las relaciones estructurales del modelo, así como el poder explicativo y predictivo del mismo (Hair et al., 2021). En el primer paso, evaluamos si las relaciones estructurales del modelo presentan algún problema de colinealidad, puesto que la estimación de los coeficientes de trayectoria del modelo se basan en regresiones de Mínimos Cuadrados Ordinarios (MCO) para cada uno de las variables latentes o constructos endógenos con su variable latente exógena o constructo predictivo correspondiente (Hair et al., 2021). Al igual que en el modelo de medición, la colinealidad se presenta en un modelo cuando los valores del VIF es superior a 5, excepto que éste caso el VIF se calcula con los coeficientes de cada regresión MCO planteada en el modelo estructural, por lo que solamente aplica para los constructos endógenos (Hair et al., 2021). En la tabla 27 podemos ver que en manos casos el VIF es menor que 5, por lo que el modelo estructural no presenta problemas de colinealidad entre el constructo endógeno desempeño financiero (DF) y los constructos predictivos capacidades de analíticos de Big Data (CABD) y desempeño del proceso de toma de decisiones (DPTD).

Tabla 27. Evaluación de colinealidad de constructos endógenos del modelo estructural

DF :	
CABD	DPTD
3.531	3.531

Fuente: elaboración propia en RStudio.

DF: desempeño financiero.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

Posteriormente, se debe evaluar la relevancia y significancia de las relaciones estructurales del modelo, las cuáles se presentan en la tabla 28, dónde podemos observar que la relación entre el DPTD y el DF no resulta ser estadísticamente significativa, sugiriendo que el desempeño del proceso de toma de decisiones de las empresas no tiene tanto impacto en el desempeño financiero de las mismas, para el caso de estudio, es decir, el caso de las empresas en Tijuana, Baja California. Asimismo, podemos ver que las CABD sí muestran un gran impacto en el desempeño financiero de las empresas, y uno menor en el desempeño del proceso de toma de decisiones.

Tabla 28. Significancia de los coeficientes de trayectoria del modelo estructural (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CABD -> DPTD	0.847	0.855	0.037	23.051	0.770	0.913
CABD -> DF	0.364	0.529	0.237	1.540	0.065	0.997
DPTD -> DF	0.225	0.082	0.252	0.893	-0.491	0.502

Fuente: elaboración propia en RStudio.

DF: desempeño financiero.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

A pesar de la no significancia del coeficiente de trayectoria entre el DPTD y el DF, se debe considerar evaluar los efectos totales para tener una idea del impacto de los dos constructos exógenos (CABD y DPTD) en el constructo endógeno (DF) (Sarstedt et al., 2021). Para ello debemos revisar los coeficientes de trayectoria totales del modelo estructural con *Bootstrap*, los cuales se muestran en la tabla 29 y nos muestran que las CABD tienen un fuerte impacto en el DPTD, seguido por el impacto de las CABD en el DF de la empresa, lo que sugiere que el desarrollar las CABD si tiene un impacto positivo en la calidad y eficiencia de las decisiones, así como en el desempeño financiero de las empresas en Tijuana, Baja California, mientras que el efecto del DPTD tiene un impacto no significativo en el DF de las mismas.

Tabla 29. Significancia de los coeficientes de trayectoria totales del modelo estructural (Modelos PLS-SEM con 10,000 muestras *bootstrapped*)

	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
CABD -> DPTD	0.847	0.855	0.037	23.051	0.770	0.913
CABD -> DF	0.555	0.598	0.109	5.077	0.363	0.769
DPTD -> DF	0.225	0.082	0.252	0.893	-0.491	0.502

Fuente: elaboración propia en RStudio.

DF: desempeño financiero.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

El siguiente paso consiste en evaluar el poder explicativo del modelo y se realiza a través del valor de R^2 de los constructos o variables endógenas (DF y DPTD) y el tamaño del efecto f^2 de los constructos exógenos o predictivos (CABD y DPTD) (Hair et al., 2021). En la tabla 30 podemos observar los resultados de las R^2 para los constructos endógenos del modelo y tenemos que el DPTD muestra un valor moderado y el DF un efecto débil de acuerdo con los parámetros establecidos en la literatura pertinente revisada (Sarstedt et al., 2021). Esto puede ser resultado del planteamiento del modelo teórico en lo que respecta a los indicadores que conforman el constructo del DF de la empresa, puesto que al ser un constructo de primer orden, se compone solamente de dos indicadores o preguntas.

Tabla 30. Evaluación del poder explicativo del modelo

	DPTD	DF
R^2	0.717	0.322
Adj R^2	0.714	0.307
CABD	0.847	0.364
DPTD	.	0.225

Fuente: elaboración propia en RStudio.

DF: desempeño financiero.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

Posteriormente se evaluó el tamaño del efecto f^2 de los constructos exógenos o predictivos (CABD y DPTD), cuyos resultados se muestran en la tabla 31, teniendo a los constructos que fungen como endógenos en las columnas y a los constructos exógenos en las filas. Teniendo en cuenta lo anterior, podemos observar que el efecto de las CABD como variable exógena o predictora en el DPTD y en el DF es grande y mediano, respectivamente,

mientras que el efecto del DPTD como variable exógena en el DF de la empresa también es mediano.

Tabla 31. Tamaño del efecto f^2 de los constructos exógenos del modelo

	CABD	DPTD	DF
CABD	0.000	2.531	0.048
DPTD	0.000	0.000	0.018
DF	0.000	0.000	0.000

Fuente: elaboración propia en RStudio.

DF: desempeño financiero.

CABD: capacidades de analíticos de Big Data.

DPTD: desempeño del proceso de toma de decisiones.

Evaluación del efecto mediador del DPTD

Una de las hipótesis de investigación es que el DPTD de la empresa, además del impacto que tienen las CABD en él y éste a su vez en el DF, también existe un efecto mediador del DPTD entre las CABD y el DF. Para evaluar el efecto mediador del DPTD se debe correr un modelo especificando dicha relación estructural, pero manteniendo el modelo de medición especificado para poder medir así la significancia del efecto específico que se genera entre las CABD y el DF cuando el DPTD. En la tabla 32 se muestran los resultados del modelo con efecto mediador del DPTD, dónde podemos observar que el efecto mediador resulta ser no significativo, por lo que podemos concluir que no existe un efecto mediador del DPTD entre las CABD y el DF de las empresas en Tijuana, Baja California.

Tabla 32. Evaluación del efecto mediador del DPTD entre las CABD y el DF

Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
0.190	0.069	0.215	0.884	-0.423	0.429

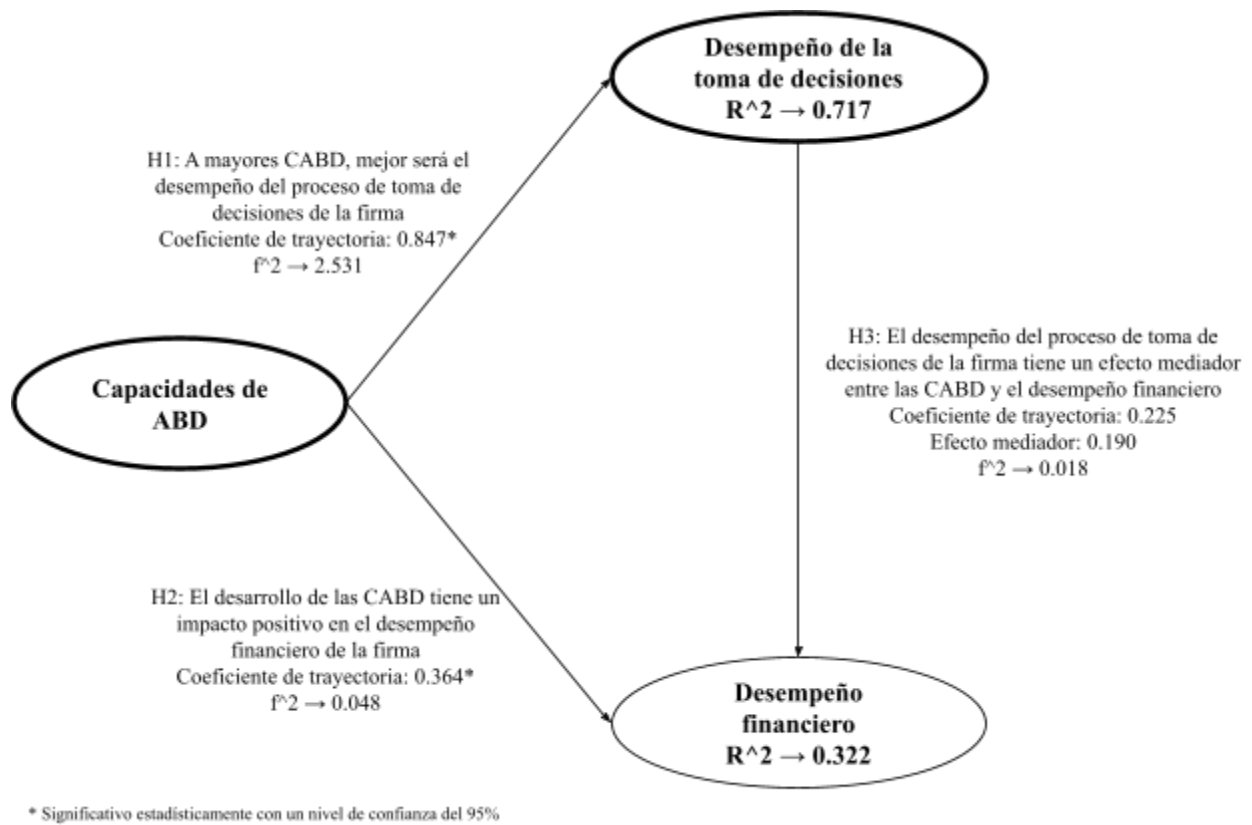
Fuente: elaboración propia en RStudio.

Impacto de las CABD en el desempeño del proceso de toma de decisiones y financiero de las empresas en Tijuana, Baja California

Recapitulando, las preguntas de investigación se refieren a la medida en que las CABD afectan el desempeño del proceso de toma de decisiones y el desempeño financiero de la empresa, así como el efecto mediador del desempeño del proceso de toma de decisiones de la empresa entre las CABD y el desempeño financiero. En ese sentido, las hipótesis planteadas son que a mayores las CABD, mejor será el desempeño del proceso de toma de decisiones de la empresa (H1), el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la empresa (H2) y que el desempeño del proceso de toma de decisiones tiene un efecto mediador entre las CABD y el desempeño financiero de la empresa (H3).

En la figura 43 podemos observar el modelo planteado incluyendo las hipótesis y resultados obtenidos, concluyendo que: la hipótesis 1, a mayores CABD, mejor será el desempeño del proceso de toma de decisiones (DPTD) es aceptada; que la hipótesis 2, el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la empresa (DF), también se acepta; y que, la hipótesis 3, el DPTD tiene un efecto mediador entre las CABD y el DF, no es soportada con el modelo y datos utilizados en la presente investigación. Los resultados, al igual que estudios previos (Cao et al., 2015; Ghasemaghaei et al., 2018; Singh y El-Kassar, 2019; Gupta y George, 2019; Aydiner et al., 2019; y Hussain et al., 2020), muestran que para las empresas estudiadas en Tijuana, Baja California, las CABD si tienen un efecto positivo el DPTD, medido en términos de la calidad y efectividad de las decisiones. De igual forma, las empresas de Tijuana, Baja California que desarrollan sus CABD observan un impacto positivo en su DF, lo cual es consistente con estudios como los de Ren et al. (2015), Akter et al. (2016), Wamba et al. (2017), Müller et al. (2018), Raguseo y Vitari (2018), Jeble et al. (2018), Anwar et al. (2018), Wamba et al. (2018), Huang et al. (2018), Gupta et al. (2019), Aydiner et al. (2019), Ferraris et al. (2019), Behl (2020), Maroufkhani et al. (2020) y Hussain et al. (2020). Finalmente, que la hipótesis relacionada al efecto mediador que tiene el DPTD entre las CABD y el DF de la firma, a pesar de la evidencia empírica encontrada en otros estudios (Hussain et al. 2020 y Gupta y George, 2016), no es el caso para el horizonte temporal y espacial objeto del presente estudio.

Figura 43. Hipótesis, modelo estructural empírico propuesto y resultados



Fuente: elaboración propia con resultados obtenidos en RStudio.

Capítulo V.

Conclusiones y perspectivas

Conclusiones

Desde que la humanidad recopila datos, la necesidad de descubrir información útil y perspectivas ha estado presente (Foote, 2021), surgiendo con dicha necesidad los analíticos, como una estrategia para extraer conocimiento de los datos, resultando de su procesamiento para intentar descubrir información útil (Smallcombe, 2021). Con la tasa de crecimiento de la generación y capacidades de almacenamiento y procesamiento de datos, surge una nueva necesidad: la de implementar nuevas herramientas y estrategias para extraer información útil, o conocimiento, para tomar decisiones en las diferentes áreas de la empresa (APICS, 2015), llevando a las empresas a desarrollar lo que en éste estudio definimos como desarrollo de las capacidades de analíticos de Big Data (CABD).

Las CABD, enmarcadas en la teoría RBV (Barney, 1991), se definen como la habilidad de la firma para desarrollar y combinar recursos que les permitan justamente analizar los datos recolectados para extraer información útil, o conocimiento, que pueda ser implementada en el proceso de toma de decisiones de la firma, orientado siempre a lograr los objetivos de negocio (Ferraris et al., 2019 y Ghasemaghaei et al., 2018), permitiendo a las empresas construir una ventaja competitiva a partir de las CABD (Gupta y George, 2016). Teniendo en cuenta lo anterior, se ha vuelto cada vez más relevante y actual el estudiar el impacto que tiene el desarrollo e inversión de las CABD en diversos indicadores (e.g. desempeño financiero, desempeño de mercado, desempeño del proceso de toma de decisiones, entre otros), buscando encontrar una relación entre los diferentes recursos que se combinan para desarrollar las CABD, su inversión, versus el impacto que éstas tienen, encontrando, en la mayoría de los casos, que la relación es positiva (Akter et al., 2016; Anwar et al., 2018; Chen et al., 2015; Côte-Real et al., 2019; Dubey et al., 2017; Ferraris et al., 2019; Ghasemaghaei et al., 2018; Gupta y George, 2016; Gupta et al., 2018; Huang et al., 2018; y Jeble et al., 2018).

A pesar de la popularidad y rápida adopción de los ABD en muchas empresas a nivel mundial, buscando el desarrollo de las CABD para construir una ventaja competitiva, la mayoría

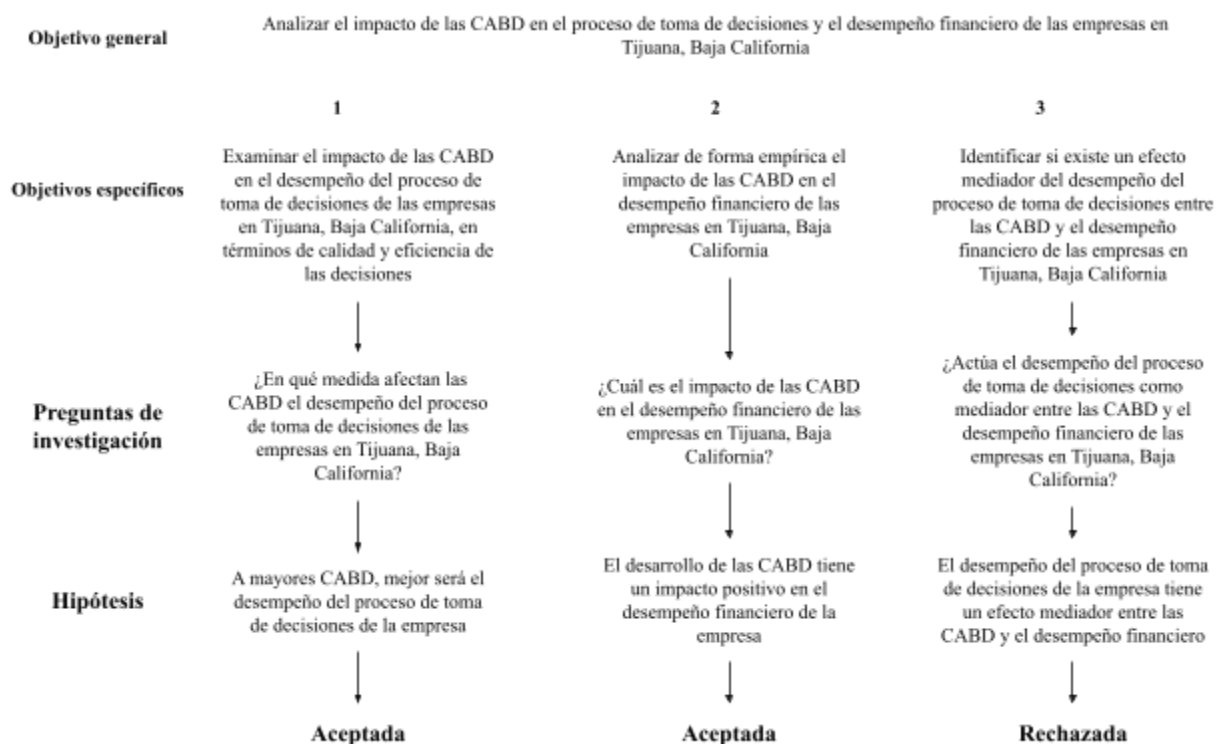
de la literatura se ha concentrado en el estudio de su impacto en diversos indicadores de las empresas localizadas en países desarrollados, encontrando nula evidencia de que el impacto positivo del desarrollo de las CABD en el desempeño financiero o de mercado, por ejemplo, sea el caso en las empresas de países no desarrollados. Sin embargo, en un mundo globalizado, dónde la generación, explotación y almacenamiento de datos no es exclusiva de empresas en países desarrollados, puesto que justamente el desarrollo de las CABD requiere de herramientas que no necesariamente necesitan estar físicamente en la empresa, haciendo que estén al alcance de empresas de todos tamaños y localizadas casi en cualquier lugar. De ahí la relevancia de estudiar si efectivamente, el desarrollar las CABD tiene un impacto positivo en el DPTD y el DF de las empresas localizadas en Tijuana, Baja California, México. No solamente porque estén localizadas en un país no desarrollado, sino porque justamente por la localización geográfica, tienen un más fácil acceso a herramientas que les permiten identificar, implementar y desarrollar recursos que les permitan desarrollar las CABD para tomar decisiones basadas en datos que les puedan llevar a crear una ventaja competitiva.

En ese sentido, teniendo en cuenta la literatura existente pertinente (Ferraris et al., 2019 y Ghasemaghaei et al., 2018), las CABD se caracterizaron y operacionalizaron como una combinación de recursos humanos, tangibles e intangibles. Específicamente, las CABD se compone de cinco dimensiones: calidad de los datos, tamaño de los datos, habilidades analíticas, dominio del conocimiento y sofisticación de las herramientas. Por otra parte, el DPTD es operacionalizado a través de dos dimensiones: la calidad y eficiencia de las decisiones tomadas (Huber, 1990; Jarupathirun, 2007; Ferraris et al., 2019; Ghasemaghaei et al., 2017; y Ghasemaghaei et al., 2018). Por último, el desempeño financiero de la empresa (DF) se caracterizó como la variable dependiente y se operacionaliza con información sobre el incremento en las ventas y en la utilidad neta de la empresa (Akter et al., 2016; Huang et al., 2018; Müller et al., 2018; Niebel et al., 2019; Raguseo y Vitari, 2018 y 2020; Ren et al., 2016; Wamba et al., 2017; Behl, 2020; Maroufkhani et al., 2020; y Hussain et al., 2020). Tomando en cuenta el objetivo de investigación sobre analizar el impacto de las CABD en el DPTD y el DF de las empresas en Tijuana, Baja California, el modelo teórico plantea una relación positiva entre las CABD y el DPTD y DF de las empresas, con un efecto mediador del DPTD entre las CABD y El DF.

De acuerdo con la mayoría de la literatura consultada sobre la relación entre las CABD y otras variables en las empresas, la metodología a seguir es un modelo de ecuaciones estructurales estimadas a través de Mínimos Cuadrados Parciales, o lo que se conoce como PLS-SEM (*Partial Least Squares Structural Equation Model* o PLS-SEM por sus siglas en inglés) (Cao et al., 2015; Chen et al., 2015; Ghasemaghaei et al., 2018; Ghasemaghaei et al., 2017; Gupta et al., 2018; Behl, 2020; Jeble et al., 2018; Ren et al., 2016; Singh, 2019; Wamba et al., 2017; Wamba et al., 2018; y Hussain et al., 2021). Bajo éste marco teórico, en el modelo propuesto las CABD se operacionalizan como un constructo formativo de segundo orden, compuesto a su vez por tres constructos formativos de primer orden (calidad de los datos, tamaño de los datos y sofisticación de las herramientas), y dos constructos reflectivos de primer orden (habilidades analíticas y dominio del conocimiento); el DPTD se operacionaliza a su vez como un constructo formativo de segundo orden compuesto por dos constructos reflectivos de primer orden (calidad y eficiencia de las decisiones); y el DF se operacionaliza como un constructo formativo de primer orden medido a través de dos indicadores: incremento en ventas e incremento en utilidad neta. Las relaciones entre las variables descritas conforman el modelo estructural, mientras que los constructos de primer orden y sus indicadores conforman el modelo de medición (Sarstedt et al., 2021 y Hair et al., 2021).

Una vez evaluado el modelo de medición, tanto para constructos formativos y reflectivos de primer y segundo orden, así como el modelo estructural con y sin efecto mediador del DPTD entre las CABD y el DF, se puede concluir que, de forma general, se cumplió con el objetivo de investigación. Particularmente, se logró alcanzar los tres objetivos específicos y preguntas de investigación planteados: examinar el impacto de las CABD en el DPTD de las empresas en Tijuana, Baja California; analizar de forma empírica el impacto de las CABD en el DF de las empresas en Tijuana, Baja California; e identificar si existe un efecto mediador del DPTD entre las CABD y el DF de las empresas en Tijuana, Baja California.

Figura 44. Objetivos, preguntas e hipótesis de investigación y resultados



Fuente: elaboración propia.

Derivado de dichos objetivos y de las respectivas preguntas de investigación, una vez evaluado el modelo PLS-SEM a partir del modelo teórico, se concluye que: la hipótesis 1, a mayores CABD, mejor será el desempeño del proceso de toma de decisiones de la empresa, es aceptada; la hipótesis 2, el desarrollo de las CABD tiene un impacto positivo en el desempeño financiero de la empresa, es aceptada; mientras que la hipótesis 3, el DPTD de la empresa, tiene un efecto mediador entre las CABD y el DF, es rechazada. En la figura 44 se pueden observar los resultados de la evaluación del modelo, relacionado con los objetivos, preguntas e hipótesis de investigación.

Perspectivas finales: estudiando el impacto del uso de datos para la toma de decisiones

Considerando que el estudio del impacto del uso de datos para la toma de decisiones y el desarrollo de las CABD, es un tema de reciente estudio y que la literatura disponible sobre su estudio es relativamente novedosa, desde conceptos teóricos como Big Data, hasta la operacionalización de las variables, las metodologías empleadas para su análisis y las áreas de la

ciencia que se concentran en dicha tarea, la expectativa sobre limitaciones a la hora de emprender un proyecto de investigación relacionado con el tema eran relacionadas con temas de recolección de información, su disponibilidad, la correcta definición y caracterización de conceptos con el debido soporte teórico, así como con la implementación de un modelo cuantitativo apropiado. En ese sentido, a continuación se describen las limitaciones y alcances de la investigación, así como sus aportaciones y recomendaciones que pudieran servir para futuros trabajos relacionados con el impacto de las CABD en el DPTD y DF de las empresas, un tema que, a pesar de la relevancia y notoriedad que ha cobrado en los últimos años, sigue existiendo poca evidencia empírica sobre los beneficios esperados.

Durante el proceso de investigación, las principales limitaciones enfrentadas estuvieron relacionadas con el levantamiento de datos, diseño y tamaño de la base de datos sobre el uso de ABD para la toma de decisiones. Por ejemplo, ese a que el diseño del instrumento de levantamiento de datos, desde constructos y preguntas hasta el modelo teórico, se basó en la metodología indicada en la literatura de referencia, algunos conceptos y terminología se tuvieron que adaptar para evitar problemas de comprensión derivados de la traducción, por lo que en futuras investigaciones relacionadas se recomienda realizar pruebas piloto más extensas y con un panel de expertos en el área y tomadores de decisiones que puedan aportar una perspectiva diferente sobre la estructura y formulación de las preguntas, o bien, ampliar el marco teórico de referencia para robustecer la validez de los constructos y el modelo de medición. Adicionalmente, dada la extensión del instrumento para que pudiese abarcar todos los conceptos teóricos involucrados, así como el uso de la escala Likert de siete puntos para las respuestas, fue complicado el obtener respuestas válidas, lo que se podría solucionar a través de un incentivo a los tomadores de decisiones que participen como respondientes del instrumento de levantamiento de datos.

En ese sentido, el tamaño de la muestra, independientemente de sí cumplir con los lineamientos teóricos para la aplicación de un modelo PLS-SEM, fue pequeño, lo que limitó el poder realizar un análisis más profundo que permitiera obtener resultados más robustos. Ésta limitante se puede solucionar, ampliando el horizonte temporal y espacial objeto de la investigación, o inclusive considerar una metodología alterna que permita el uso de fuentes de información ya existentes, por ejemplo, la Encuesta Nacional sobre Productividad y

Competitividad de las Micro, Pequeñas y Medianas Empresas publicada por INEGI o información sobre empresas públicas. Esto también podría permitir utilizar un análisis de datos de panel, teniendo la posibilidad de indagar más sobre la adopción, desarrollo e impacto de las CABD en las empresas, comparando los diferentes resultados, o desempeño del proceso de toma de decisiones y financiero de las empresas.

A pesar de las limitaciones y mejoras aplicables a la investigación, se vuelve mucho más relevante continuar con el estudio del impacto de las CABD en diferentes áreas de la empresa, puesto que, como se evidenció en los resultados de la presente investigación, existe una relación positiva entre el desarrollo y combinación de ciertos recursos para incentivar el uso de analíticos de datos en el proceso de toma de decisiones que puede llevar a la empresa a crear una ventaja competitiva e incrementar su desempeño financiero y crecimiento. Particularmente, en países no desarrollados, donde existe una brecha de conocimiento respecto a países más desarrollados o con una adopción más temprana de nuevas tecnologías y recursos que fomentan el desarrollo de las CABD, lo que llevaría a tener evidencia empírica sobre los beneficios, incentivando su adopción, promoviendo el crecimiento de las empresas y, por consecuencia, el desarrollo económico y social de los países.

Capítulo VI.

Referencias bibliográficas

Abusweilema, M. A., y Abualoush, S. (2019). “The Impact of Knowledge Management Process and Business Intelligence on Organizational Performance”. *Management Science Letters*, 9(12), 2143–2156.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., y Childe, S. J. (2016). “How to improve firm performance using big data analytics capability and business strategy alignment?”. *International Journal of Production Economics*, 182, 113–131.

Aydiner, A. S., Tatoglu, E., Bayraktar, E., Zaim, S., y Delen, D. (2019). “Business Analytics and Firm Performance: The Mediating Role of Business Process Performance”. *Journal of Business Research*, 96, 228-237.

Barney, J.B. (1991), “The Resource Based View of Strategy: Origins, Implications, and Prospects”. *Journal of management*, Vol. 17 No. 1, pp. 97-211.

Barton, D. y Court, D. (2012), “Making advanced analytics work for you”, *Harvard Business Review*, Vol. 90 No. 10, pp. 78-83.

Behl, A., Dutta, P., Lessmann, S., Dwivedi, Y.K. y Kar, S. (2019), “A Conceptual Framework for the Adoption of Big Data Analytics by E-Commerce Start-Ups: A Case-Based Approach”. *Information Systems and E-Business Management*, Vol. 17 No. 2, pp. 285-318.

Bharadwaj, A.S. (2000), “A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation”. *MIS Quarterly*, Vol. 24 No. 1, pp. 169-196.

Bogdan, M., y Borza, A. (2019). “Big Data Analytics and Organizational Performance: A Meta-Analysis Study”. *Management and Economics Review*, 4(2), 147–162.

Braganza, A., Brooks, L., Nepelski, D., Ali, M., Moro, R., 2017. “Resource Management in Big Data Initiatives: Processes and Dynamic Capabilities”. *Journal of Business Research*. 70, 328–337.

Castro, Juan (2011). “La Metodología Lean Six Sigma en la Manufactura de Productos Médicos Como Estrategia Competitiva en Tijuana B.C.: Caso de Estudio” [Tesis de maestría no publicada]. Universidad Autónoma de Baja California.

Côrte-Real, N., Oliveira, T., y Ruivo, P. (2017). “Assessing Business Value of Big Data Analytics in European Firms”. *Journal of Business Research*, 70, 379-390.

Davenport, T. H., y Harris, J. G. (2007). “Competing on analytics: The New Science of Winning”. Harvard Business Press.

Einav, L. y Levin, J.D. (2013), “The Data Revolution and Economic Analysis”. *National Bureau of Economic Research*, Vol. 364 No. 6210, pp. 715-722.

Ferraris, A., Mazzoleni, A., Devalle, A., y Couturier, J. (2019). “Big Data Analytics Capabilities and Knowledge Management: Impact on Firm Performance”. *Management Decision*, 57(8), 1923–1936.

Ferguson, G., Mathur, S., y Shah, B., (2005). “Evolving from information to insight”. *MIT Sloan Management Review*, 46 (2), 51–58.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., y Gnanzou, D. (2015). “How ‘Big Data’ Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study”. *International Journal of Production Economics*, 165, 234–246.

Ganster, P., y Collins, K. (2017). “Binational Cooperation and Twinning: A View from the US–Mexican Border, San Diego, California, and Tijuana, Baja California”. *Journal of Borderlands Studies*, 32(4), 497–511.

Ghasemaghaei, M., Ebrahimi, S., y Hassanein, K. (2018). “Data Analytics Competency for Improving Firm Decision Making Performance”. *The Journal of Strategic Information Systems*, 27(1), 101-113.

Ghasemaghaei, M., Hassanein, K., y Turel, O. (2017). “Increasing Firm Agility Through the Use of Data Analytics: The Role Of Fit”. *Decision Support Systems*, 101, 95-105.

- Ghasemaghaei, M., y Turel, O. (2018). "Data Analytics Competency for Improving Firm Decision Making Performance". *The Journal of Strategic Information Systems*, 27(1), 101-113.
- Goodman, L. A. (1961). "Snowball Sampling". *The Annals of Mathematical Statistics*, 32(1), 148–170.
- Grant, R.M. (1991), "The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation". *California Management Review*, vol. 33, no. 3: 114-135.
- Gupta, M. y George, J.F. (2016), "Toward the Development of a Big Data Analytics Capability", *Information & Management*, Vol. 53 No. 8, pp. 1049-1064.
- Gupta, S., Kar, A.K., Baabdullah, A. y Al-Khowaiter, W.A.A. (2018), "Big Data with Cognitive Computing: A Review for the Future", *International Journal of Information Management*, Vol.42, pp.78-89.
- Hagel, J. (2015), "Bringing Analytics to Life". *Journal of Accountancy*, Vol. 219 No. 2, pp. 24-25.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., y Sarstedt, M. (2017). "A primer on partial least squares structural equation modeling (PLS-SEM)". Thousand Oaks, CA. Sage Publications.
- Hair, J. F., Risher, J. J., Sarstedt, M., y Ringle, C. M. (2019). "When to use and how to report the results of PLS-SEM". *European Business Review*, 31(1), 2–24.
- Hair, J. F., Ringle, C. M., y Sarstedt, M. (2011). "PLS-SEM: Indeed a silver bullet". *The Journal of Marketing Theory and Practice*, 2011, 15.
- Hair, J. F., Sarstedt, Marko, Hult, G. Tomas M., Ringle, C. M., Danks, Nicholas P. y Ray, Soumya (2021). "Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R". A Workbook. Tercera Edición. Springer Open Access 2021.
- Huber, G.P., (1990). "A Theory of The Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision Making". *Academic of Management Review*, 15 (1), 47–71.

Instituto Nacional de Estadística y Geografía (INEGI), (2023). Sistema de Clasificación Industrial de América del Norte (SCIAN, 2018). Recuperado de [https://www.inegi.org.mx/app/scian/#:~:text=2018%20\(SCIAN%202018\)-,Sistema%20de%20Clasificaci%C3%B3n%20Industrial%20de%20Am%C3%A9rica%20del%20Norte%202018%20](https://www.inegi.org.mx/app/scian/#:~:text=2018%20(SCIAN%202018)-,Sistema%20de%20Clasificaci%C3%B3n%20Industrial%20de%20Am%C3%A9rica%20del%20Norte%202018%20) el 15 de mayo de 2013.

Instituto Nacional de Estadística y Geografía (INEGI) (2022). Producto Interno Bruto por Entidad Federativa Baja California 2021. Comunicado de Prensa Núm. 739/22. 7 de diciembre de 2022.

Instituto Nacional de Estadística y Geografía (INEGI) (2022). Directorio Nacional de Unidades Económicas (DENUE), XVIII edición. Recuperado de la página de consulta <https://www.inegi.org.mx/app/mapa/denue/Default.aspx?idee=6463810> el 24 de mayo de 2022.

Instituto Nacional de Estadística y Geografía (INEGI) (2023). Cuéntame de México. Recuperado de https://cuentame.inegi.org.mx/monografias/informacion/bc/territorio/div_municipal.aspx?tema=me&e=02 el 2 de mayo de 2023.

Jyoti, J., y Rani, A. (2017). “High Performance Work System And Organizational Performance: Role of Knowledge Management”. *Personnel Review*, 46(8), 1770-1795.

Kent State University (s.f.). “Statistical & Qualitative Data Analysis Software: About R And Rstudio”. University Libraries. Recuperado de [https://libguides.library.kent.edu/statconsulting/r#:~:text=RStudio%20is%20a%20free%2C%20open,CSV%2C%20Excel%2C%20SAS%20\(*](https://libguides.library.kent.edu/statconsulting/r#:~:text=RStudio%20is%20a%20free%2C%20open,CSV%2C%20Excel%2C%20SAS%20(*) el 12 de agosto de 2023.

Liu, H., Ke, W., Wei, K. K., y Hua, Z. (2013). “The Impact Of IT Capabilities On Firm Performance: The Mediating Roles Of Absorptive Capacity And Supply Chain Agility”. *Decision Support Systems*, 54(3), 1452–1462.

Lu, Y., Ramamurthy, K. (Ram), (2011). “Understanding the Link Between Information Technology Capability and Organizational Agility: An Empirical Examination”. *MIS Quarterly*, 35 (4), 931–954.

Masa'deh, R., Obeidat, B., y Tarhini, A. (2016). "A Jordanian Empirical Study Of The Associations Among Trans-Formational Leadership, Transactional Leadership, Knowledge Sharing, Job Performance, And Firm Performance: A Structural Equation Modeling Approach". *Journal of Management Development*, 35(5), 681-705.

Matthias, O., Ian Fouweather, I. Gregory, Andy Vernon, I. Gregory, I. Gregory, y Vernon, A.. (2017). "Making Sense of Big Data– Can it Transform Operations Management?". *International Journal of Operations & Production Management*, 37.

Mendoza, J. (2010). "El mercado laboral en la frontera norte de México: estructura y políticas de empleo. Introducción Desde la instrumentación del Tratado de Libre Comercio de América del Norte". *Estudios Fronterizos*, 11(21), 9–42.

Mendoza, J. E., y Dupeyron, B. (2020). "Economic Integration, Emerging Fields and Cross-border Governance: The Case of San Diego–Tijuana". *Journal of Borderlands Studies*, 35(1), 55–74.

O'Leary, D.E. (2013), "'Big Data', the 'Internet of Things' and the 'Internet of Signs'". *Intelligent Systems in Accounting, Finance and Management*, 20: 53-65.

Oracle (2018). "22 Big Data Use Cases You Want to Know". Oracle 2018. Recuperado de https://www.oracle.com/a/ocom/docs/dc/em/top-22-use-cases-for-big-datav2.pdf?source=%3Aow%3Awp%3Acpo%3A%3A&intcmp=:ow:lp:pt::RC_WWMK180502P00084:LPD100713386 el 21 de noviembre de 2022.

Pavlou, P.A., El Sawy, O.A., (2006). "From IT Leveraging Competence to Competitive Advantage in Turbulent Environments: The Case of New Product Development". *Information Systems Research*, 17 (3), 198–227.

Peteraf, M.A., Barney, J.B., (2003). "Unraveling the resource-based tangle". *Managerial and Decisions Economics* 24, 309–323.

Phillips, Andrés (1 de abril de 2021), "A History and Timeline of Big Data". Integrate.io 2021. Recuperado de <https://www.techtarget.com/whatis/feature/A-history-and-timeline-of-big-data> el 6 de marzo de 2023.

Popovič, A., Hackney, R., Tassabehji, R., y Castelli, M. (2018). “The Impact of Big Data Analytics on Firms' High Value Business Performance”. *Information Systems Frontiers*, 20(2), 209–222.

Puyana, A., & Romero, J. (2006). “Hacia una Evaluación de los Efectos Multiplicadores de la Actividad Maquiladora”. *Estudios Sociológicos*, XXIV(70), 65-97.

Ringle, C. M., Wende, Sven, y Becker, Jan-Michael (2022). *Oststeinbek: SmartPLS*. Recuperado de <https://www.smartpls.com/documentation/algorithms-and-techniques/higher-order> el 16 de mayo de 2023.

Roldán, J. L., y Sánchez-Franco, M. J. (2012). Variance-Based Structural Equation Modeling: Guidelines for Using Partial Least Squares in Information Systems Research. En M. Mora, O. Gelman, A. Steenkamp, y M. Raisinghani (Ed.), *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems* (pp. 193-221).

Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M., y Ringle, C. M. (2019). “How to specify, estimate, and validate higher-order constructs in PLS-SEM”. *Australasian Marketing Journal (AMJ)* 27, (pp. 197-211).

Sen, D., Ozturk, M., y Vayvay, O. (2016). “An Overview of Big Data for Growth in SMEs”. *Procedia - Social and Behavioral Sciences*, 235 (October), 159–167.

Shah, S., Horne, A. y Capellá, J. (2012), “Good Data Won't Guarantee Good Decisions”. *Harvard Business Review*, Vol. 90 No. 4, pp. 23-25.

Smallcombe, Mark (7 de diciembre 2021), “A Brief History of Data Analysis”. *Dataiversity* 2021. Recuperado de <https://www.integrate.io/blog/a-brief-history-of-data-analysis/> el 6 de marzo de 2023.

Taylor, David (1 de octubre de 2022). “What is Big Data?”. *Guru99*. Recuperado de <https://www.guru99.com/what-is-big-data.html> el 21 de noviembre de 2022.

Thompson, S.K. (2002). “Sampling”. New York, NY, EUA. Wiley.

Torres, R., Sidorova, A., y Jones, M. C. (2018). "Enabling Firm Performance Through Business Intelligence and Analytics: A Dynamic Capabilities Perspective". *Information & Management*, 55(7), 822-839.

Wesley Chai, Mark Labbe y Craig Stedman (Diciembre 2021), "Big Data Analytics". TechTarget 2021. Recuperado de <https://www.techtarget.com/searchbusinessanalytics/definition/big-data-analytics> el 13 de marzo de 2023.

Wesley Chai, Mark Labbe y Craig Stedman (Diciembre 2021), "Big Data". TechTarget 2021. Recuperado de <https://www.techtarget.com/searchdatamanagement/definition/big-data> el 14 de octubre de 2022.

Anexos

Anexo 1. Código R (RStudio) utilizado para la aplicación del modelo PLS-SEM

```
#### Resultados ####
```

```
## Instalación de paquetes y librería  
install.packages("semnr")  
library(semnr)
```

```
## Cargar e inspeccionar la base de datos  
CABD_Data <- CABD_DB  
head(CABD_Data)
```

```
## Especificación general de los constructos e indicadores del modelo
```

```
# Definición de los constructos
```

```
# Constructos de primer orden
```

```
# HA: Habilidades analíticas del personal que trabaja con datos en la empresa para  
transformarlos en conocimiento
```

```
# CD: Calidad de los datos se refiere a la calidad de los hechos en bruto que reflejan las  
características de una
```

```
# entidad o evento
```

```
# TD: Grandeza de los datos se refiere a datos de gran volumen, variedad y velocidad
```

```
# DC: Dominio del conocimiento sobre procedimientos, hechos y procesos involucrados en la  
empresa y la industria
```

```
# en la que opera
```

```
# SH: Expertise tecnológico al interior de la empresa, es decir, la complejidad y maduración de  
las herramientas
```

```
# de TI con las que cuenta para realizar el análisis de los datos y extraer información de los  
mismos
```

```
# CDe: Calidad de las decisiones se refiere a que tan exactas, precisas y confiables son los  
resultados de las
```

```
# decisiones tomadas
```

```
# ED: Eficiencia de las decisiones se refiere a qué tan rápido se llegan a tomar las decisiones
```

```
# DF: Desempeño financiero de la empresa se refiere al crecimiento de la utilidad neta y ventas
```

```
# Constructos de segundo orden (Higher Order Constructs)
```

```
# CABD: Habilidades de la empresa para desplegar y combinar recursos de ABD para un  
análisis de datos riguroso y
```

```
# orientado a la toma de decisiones
```

DPTD: Evaluación de los tomadores de decisiones en la firma sobre la calidad y eficiencia de su proceso de toma de decisiones

Especificación del Modelo

#	Constructo	ID	Tipo de constructo	Pregunta / ítem
#	DF	DF_global	1er orden	En su organización ¿se ha incrementado el desempeño financiero con el uso de Analíticos de Datos para la toma de decisiones?
#		DF_1	Ítem	La utilidad neta de la empresa mejoró desde la implementación de Analíticos de Datos en el proceso de toma de decisiones.
#		DF_2	Ítem	Las ventas de la empresa incrementaron desde la implementación de Análisis de Datos en el proceso de toma de decisiones.
#	CABD	CABD_global	2do orden	En su organización ¿considera que las Capacidades de Analíticos de Big Data se encuentran desarrolladas?
#		HA_1	1er orden / ítem	Nuestros usuarios de analíticos de datos son expertos en lo que respecta a utilizar las herramientas de análisis de datos.
#		HA_2	1er orden / ítem	Nuestros usuarios de analíticos de datos poseen un alto grado de experiencia en análisis de datos.
#		HA_3	1er orden / ítem	Nuestros usuarios de analíticos de datos son hábiles en en uso de herramientas de análisis de datos.
#	CD	CD_global	1er orden	En nuestra organización, la calidad de los datos utilizados para construir Analíticos de Datos es muy buena.
#		CD_1	Ítem	En nuestra organización, los datos utilizados en analíticos de datos son confiables.
#		CD_2	Ítem	En nuestra organización, los datos utilizados en analíticos de datos tienen el nivel apropiado de detalle.
#		CD_3	Ítem	En nuestra organización, los datos utilizados en analíticos de datos son seguros.
#		CD_4	Ítem	En nuestra organización, los datos utilizados en analíticos de datos son oportunos.
#		CD_5	Ítem	En nuestra organización, los datos utilizados en analíticos de datos son relevantes para la tarea en cuestión.
#		CD_6	Ítem	En nuestra organización, los datos utilizados en analíticos de datos son precisos.
#	TD	TD_global	1er orden	Nuestra organización recopila y analiza grandes volúmenes de datos.
#		TD_1	Ítem	Nuestra organización procesa grandes volúmenes de datos.
#		TD_2	Ítem	Nuestra organización procesa datos en tiempo real.
#		TD_3	Ítem	Nuestra organización procesa diferentes tipos de datos.
#	DC	DC_1	1er orden / ítem	En nuestra organización, hay un alto nivel de conocimiento del entorno externo (e.g. gobierno, competidores, proveedores y clientes).
#		DC_2	1er orden / ítem	En nuestra organización, hay un alto nivel de conocimiento de las metas y objetivos organizacionales.
#		DC_3	1er orden / ítem	En nuestra organización, hay un alto nivel de conocimiento de las capacidades básicas de la organización.

DC_4 1er orden / ítem En nuestra organización, hay un alto nivel de conocimiento de los factores clave que deben ir bien para que la organización tenga éxito.

SH SH_global 1er orden En nuestra organización, se utilizan las herramientas de análisis de datos apropiadas de acuerdo a nuestras necesidades.

SH_1 Ítem En nuestra organización, utilizamos herramientas que brindan capacidades de procesamiento y recuperación de información.

SH_2 Ítem En nuestra organización, utilizamos herramientas que realizan modelado y simulación.

SH_3 Ítem En nuestra organización, utilizamos herramientas que realizan analíticos de lenguaje natural (extracción de información de fuentes no estructuradas como redes sociales).

SH_4 Ítem En nuestra organización, utilizamos herramientas que proporcionan información en tiempo real.

SH_5 Ítem En nuestra organización, utilizamos herramientas que identifican problemas.

SH_6 Ítem En nuestra organización, utilizamos herramientas que evalúan diferentes alternativas.

DPTD DPTD_global 2do orden En nuestra organización existe un buen desempeño del proceso de toma de decisiones.

CDe CDe_1 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo exactos.

CDe_2 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo correctos.

CDe_3 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo precisos.

CDe_4 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo perfectos.

CDe_5 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo sin errores.

CDe_6 1er orden / ítem En nuestra organización, los resultados de las decisiones son a menudo confiables.

ED ED_1 1er orden / ítem En nuestra organización, el tiempo para llegar a decisiones es rápido.

ED_2 1er orden / ítem En nuestra organización, la velocidad para llegar a decisiones es alta.

Trayectorias (Paths)

CD -> CABD
 # TD -> CABD
 # HA -> CABD
 # DC -> CABD
 # SH -> CABD
 # CDe -> DPTD
 # ED -> CPTD
 # Vtas-> DF

```

# CABD-> DPTD
# CABD-> DF
# DPTD-> DF (Efecto Mediador)

# Tipo de constructo (Reflectivo vs Formativo)

# CD -> Formativo
# TD -> Formativo
# HA -> Reflectivo
# DC -> Reflectivo
# SH -> Formativo
# CDe -> Reflectivo
# ED -> Reflectivo
# CABD-> Formativo
# DF -> Formativo
# DPTD-> Formativo

## Modelo de Medición

# Evaluación de constructos de primer orden
# (CD, TD, HA, DC, SH, CDe, ED y DF)
# El primer paso es evaluar la confiabilidad y validez de los constructos de primer orden
independientemente si son
# Reflectivos o Formativos quitando a los constructos de segundo orden de las trayectorias

# Especificación del Modelo de Medición

CABD_LOC_MM <- constructs(
  composite("CD", multi_items("CD_", 1:6), weights = mode_B),
  composite("TD", multi_items("TD_", 1:3), weights = mode_B),
  composite("HA", multi_items("HA_", 1:3)),
  composite("DC", multi_items("DC_", 1:4)),
  composite("SH", multi_items("SH_", 1:6), weights = mode_B),
  composite("CDe", multi_items("CDe_", 1:6)),
  composite("ED", multi_items("ED_", 1:2)),
  composite("DF", multi_items("DF_", 1:2), weights = mode_B)
)

# Nota: los constructos CD, TD y SH son constructos formativos, por lo que se estiman con
# mode_B (regression weights), mientras que HA, DC, CDe y ED son cosntructos reflectivos y
se estiman con mode_A
# (correlation_weights) dentro de la función composite()

# Especificación del Modelo Estructural

# Trayectorias (Paths)

```

```

# CD -> CDe
# TD -> CDe
# HA -> CDe
# DC -> CDe
# SH -> CDe
# CD -> ED
# TD -> ED
# HA -> ED
# DC -> ED
# SH -> ED
# CD -> DF
# TD -> DF
# HA -> DF
# DC -> DF
# SH -> DF
# CDe -> DF
# ED -> DF

CABD_LOC_SM <- relationships(
  paths(from = c("CD", "TD", "HA", "DC", "SH"), to = c("CDe", "ED")),
  paths(from = c("CD", "TD", "HA", "DC", "SH"), to = c("DF")),
  paths(from = c("CDe", "ED"), to = c("DF"))
)

# Estimación del Modelo

CABD_LOC_Model <- estimate_pls(
  data = CABD_Data,
  measurement_model = CABD_LOC_MM,
  structural_model = CABD_LOC_SM,
  inner_weights = path_weighting,
  missing = mean_replacement,
  missing_value = "-99"
)

# Graficamos el Modelo

plot(CABD_LOC_Model)

# Resumen de resultados

summary_CABD_LOC <- summary(CABD_LOC_Model)
View(summary_CABD_LOC)
summary_CABD_LOC
write.csv(x = summary_CABD_LOC$descriptives, file = "descriptive_statistics.csv")

```

```

# Revisamos las iteraciones del Algoritmo para converger
summary_CABD_LOC$iterations
# 18 iteraciones
write.csv(x = summary_CABD_LOC$iterations, file = "CABD_LOC_iterations.csv")

## Evaluación del modelo de medición para los constructos de primer orden (LOC)

# Evaluación de LOC reflectivos → LOC = HA, DC, CDe y ED

# Crear cargas de los indicadores
summary_CABD_LOC$loadings
write.csv(x = summary_CABD_LOC$loadings, file = "summary_CABD_LOC_loadings.csv")

# Crear cargas de los indicadores al cuadrado
summary_CABD_LOC$loadings^2
write.csv(x = summary_CABD_LOC$loadings^2, file = "CABD_LOC_loadings2.csv")

# Evaluar la fiabilidad y validez convergente de los constructos
summary_CABD_LOC$reliability
plot(summary_CABD_LOC$reliability)
write.csv(x = summary_CABD_LOC$reliability, file = "CABD_LOC_reliability.csv")

# Evaluar la validez discriminante (HTMT y Fornell & Larcker)
summary_CABD_LOC$validity$htmt
write.csv(x = summary_CABD_LOC$validity$htmt, file = "CABD_LOC_HTMT.csv")
summary_CABD_LOC$validity$fl_criteria
write.csv(x = summary_CABD_LOC$validity$fl_criteria, file = "CABD_LOC_FL.csv")

## Evaluar LOC formativos → CD, TD, SH y DF

# Crear el modelo de medición para cada una de los constructos formativos de primer orden

simple_CD <- constructs(
  composite("CDg", multi_items("CD_", 1:6), weights = mode_B),
  composite("CD_global", single_item("CD_global"))
)

simple_TD <- constructs(
  composite("TDg", multi_items("TD_", 1:3), weights = mode_B),
  composite("TD_global", single_item("TD_global"))
)

simple_SH <- constructs(
  composite("SHg", multi_items("SH_", 1:6), weights = mode_B),
  composite("SH_global", single_item("SH_global"))
)

```



```

)

simple_DF <- constructs(
  composite("DFg", multi_items("DF_", 1:2), weights = mode_B),
  composite("DF_global", single_item("DF_global"))
)

# Crear el modelo estructural para cada uno de los constructos formativos de primer orden

simple_CD_sm <- relationships(
  paths(from = c("CDg"), to = c("CD_global"))
)

simple_TD_sm <- relationships(
  paths(from = c("TDg"), to = c("TD_global"))
)

simple_SH_sm <- relationships(
  paths(from = c("SHg"), to = c("SH_global"))
)

simple_DF_sm <- relationships(
  paths(from = c("DFg"), to = c("DF_global"))
)

# Estimar el modelo para cada uno de los constructos formativos de primer orden

CD_redundancy_pls_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_CD,
  structural_model = simple_CD_sm
)

TD_redundancy_pls_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_TD,
  structural_model = simple_TD_sm
)

SH_redundancy_pls_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_SH,
  structural_model = simple_SH_sm
)

DF_redundancy_pls_model <- estimate_pls(

```

```

data = CABD_Data,
measurement_model = simple_DF,
structural_model = simple_DF_sm
)

# Creamos el resumen de los resultados de cada modelo PLS-SEM de los constructos formativos
de primer orden

sum_red_CD_model <- summary(CD_redundancy_pls_model)

sum_red_TD_model <- summary(TD_redundancy_pls_model)

sum_red_SH_model <- summary(SH_redundancy_pls_model)

sum_red_DF_model <- summary(DF_redundancy_pls_model)

# Obtener los coeficientes de trayectoria para cada uno de los modelos estimados

sum_red_CD_model$paths
sum_red_DF_model$paths
sum_red_SH_model$paths
sum_red_TD_model$paths

# Graficar los resultados de cada modelo

plot(CD_redundancy_pls_model)
plot(TD_redundancy_pls_model)
plot(SH_redundancy_pls_model)
plot(DF_redundancy_pls_model)

# Evaluar colinealidad en los constructos formativos de primer orden extrayendo el VIF del
resumen

sum_red_CD_model$validity$vif_items
sum_red_DF_model$validity$vif_items
sum_red_SH_model$validity$vif_items
sum_red_TD_model$validity$vif_items

# Evaluación de la significancia y relevancia de los indicadores que conforman los constructos
# Como primer paso se deben correr cada uno de los modelos PLS-SEM con un bootstrapping de
10,000 muestras

CD_boot_model <- bootstrap_model(
  semnr_model = CD_redundancy_pls_model,
  nboot = 10000,
  seed = 123

```

```

)

TD_boot_model <- bootstrap_model(
  seminr_model = TD_redundancy_pls_model,
  nboot = 10000,
  seed = 123
)

SH_boot_model <- bootstrap_model(
  seminr_model = SH_redundancy_pls_model,
  nboot = 10000,
  seed = 123
)

DF_boot_model <- bootstrap_model(
  seminr_model = DF_redundancy_pls_model,
  nboot = 10000,
  seed = 123
)

# Se grafican los resultados

plot(CD_boot_model)
plot(TD_boot_model)
plot(SH_boot_model)
plot(DF_boot_model)

# Se crean un resumen para cada resultado de los modelos bootstrapped con un valor crítico de
significancia de 5%

sum_CD_boot <- summary(CD_boot_model, alpha = 0.05)
sum_TD_boot <- summary(TD_boot_model, alpha = 0.05)
sum_SH_boot <- summary(SH_boot_model, alpha = 0.05)
sum_DF_boot <- summary(DF_boot_model, alpha = 0.05)

# Extraer los resultados

sum_CD_boot$bootstrapped_weights
sum_TD_boot$bootstrapped_weights
sum_SH_boot$bootstrapped_weights
sum_DF_boot$bootstrapped_weights

# Revisar la importancia de los indicadores con las cargas de los modelos PLS-SEM con
bootstrap

sum_CD_boot$bootstrapped_loadings

```

```

sum_TD_boot$bootstrapped_loadings
sum_SH_boot$bootstrapped_loadings
sum_DF_boot$bootstrapped_loadings

## Evaluar los constructos de segundo orden

# Crear el modelo de medición para los constructos de segundo orden o del modelo general

simple_CABD_mm <- constructs(
  composite("CD", multi_items("CD_", 1:6), weights = mode_B),
  composite("TD", multi_items("TD_", 1:3), weights = mode_B),
  composite("HA", multi_items("HA_", 1:3)),
  composite("DC", multi_items("DC_", 1:4)),
  composite("SH", multi_items("SH_", 1:6), weights = mode_B),
  higher_composite("CABD", c("CD", "TD", "HA", "DC", "SH"), method = two_stage, weights
= mode_B),
  composite("CDe", multi_items("CDe_", 1:6)),
  composite("ED", multi_items("ED_", 1:2)),
  higher_composite("DPTD", c("CDe", "ED"), method = two_stage, weights = mode_B),
  composite("DF", multi_items("DF_", 1:2), weights = mode_B))

# Crear el modelo estructural
# Paths
# CABD -> DPTD
# CABD -> DF
# DPTD -> DF

simple_CABD_sm <- relationships(
  paths(from = "CABD", to = "DPTD"),
  paths(from = "CABD", to = "DF"),
  paths(from = "DPTD", to = "DF")
)

# Estimar el modelo

simple_CABD_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_CABD_mm,
  structural_model = simple_CABD_sm
)

# Graficar el modelo

plot(simple_CABD_model)

# Crear el resumen del modelo

```

```

summary_simple_CABD_model <- summary(simple_CABD_model)

# Extraer las cargas de los indicadores o constructos de primer orden

summary_simple_CABD_model$iterations
summary_simple_CABD_model$loadings

# Análisis de redundancia de los constructos formativos de segundo orden

# Crear el modelo de medición para cada una de los constructos formativos de segundo orden

simple_RA_CABD_mm <- constructs(
  composite("CD", multi_items("CD_", 1:6), weights = mode_B),
  composite("TD", multi_items("TD_", 1:3), weights = mode_B),
  composite("HA", multi_items("HA_", 1:3)),
  composite("DC", multi_items("DC_", 1:4)),
  composite("SH", multi_items("SH_", 1:6), weights = mode_B),
  higher_composite("CABDg", c("CD", "TD", "HA", "DC", "SH"), method = two_stage, weights
= mode_B),
  composite("CABD_global", single_item("CABD_global"))
)

simple_RA_DPTD_mm <- constructs(
  composite("CDe", multi_items("CDe_", 1:6)),
  composite("ED", multi_items("ED_", 1:2)),
  higher_composite("DPTDg", c("CDe", "ED"), method = two_stage, weights = mode_B),
  composite("DPTD_global", single_item("DPTD_global"))
)

# Crear el modelo estructural para cada uno de los constructos formativos de segundo orden

simple_RA_CABD_sm <- relationships(
  paths(from = c("CABDg"), to = c("CABD_global"))
)

simple_RA_DPTD_sm <- relationships(
  paths(from = c("DPTDg"), to = c("DPTD_global"))
)

# Estimar el modelo para cada uno de los constructos formativos de segundo orden

CABD_redundancy_pls_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_RA_CABD_mm,
  structural_model = simple_RA_CABD_sm
)

```

```

)

DPTD_redundancy_pls_model <- estimate_pls(
  data = CABD_Data,
  measurement_model = simple_RA_DPTD_mm,
  structural_model = simple_RA_DPTD_sm
)

# Creamos el resumen de los resultados de cada modelo PLS-SEM de los constructos formativos
# de segundo orden

sum_red_CABD_model <- summary(CABD_redundancy_pls_model)

sum_red_DPTD_model <- summary(DPTD_redundancy_pls_model)

# Obtener los coeficientes de trayectoria para cada uno de los modelos estimados

sum_red_CABD_model$paths
sum_red_DPTD_model$paths

# Graficar los resultados de cada modelo

plot(CABD_redundancy_pls_model)
plot(DPTD_redundancy_pls_model)

# Evaluar colinealidad en los constructos formativos de segundo orden extrayendo el VIF del
# resumen

sum_red_CABD_model$validity$vif_items
sum_red_DPTD_model$validity$vif_items

# Evaluación de la significancia y relevancia de los indicadores que conforman los constructos
# Como primer paso se deben correr cada uno de los modelos PLS-SEM con un bootstrapping de
# 10,000 muestras

CABD_boot_model <- bootstrap_model(
  seminr_model = CABD_redundancy_pls_model,
  nboot = 10000,
  seed = 123
)

DPTD_boot_model <- bootstrap_model(
  seminr_model = DPTD_redundancy_pls_model,
  nboot = 10000,
  seed = 123
)

```

)

Se grafican los resultados

```
plot(CABD_boot_model)
plot(DPTD_boot_model)
```

Se crean un resumen para cada resultado de los modelos bootstrapped con un valor crítico de significancia de 5%

```
sum_CABD_boot <- summary(CABD_boot_model, alpha = 0.05)
sum_DPTD_boot <- summary(DPTD_boot_model, alpha = 0.05)
```

Extraer los resultados

```
sum_CABD_boot$bootstrapped_weights
sum_DPTD_boot$bootstrapped_weights
write.csv(x = sum_CABD_boot$bootstrapped_weights, file = "sum_CABD_boot_weights.csv")
write.csv(x = sum_DPTD_boot$bootstrapped_weights, file = "sum_DPTD_boot_weights.csv")
```

Revisar la importancia de los indicadores con las cargas de los modelos PLS-SEM con bootstrap

```
sum_CABD_boot$bootstrapped_loadings
sum_DPTD_boot$bootstrapped_loadings
write.csv(x = sum_CABD_boot$bootstrapped_loadings, file = "sum_CABD_boot_loadings.csv")
write.csv(x = sum_DPTD_boot$bootstrapped_loadings, file = "sum_DPTD_boot_loadings.csv")
```

Evaluación del modelo estructural sin efecto mediador de DPTD (incluye HOC)

Especificar el modelo de medición

```
CABD_MM <- constructs(
  composite("CD", multi_items("CD_", 1:6), weights = mode_B),
  composite("TD", multi_items("TD_", 1:3), weights = mode_B),
  composite("HA", multi_items("HA_", 1:3)),
  composite("DC", multi_items("DC_", 1:4)),
  composite("SH", multi_items("SH_", 1:6), weights = mode_B),
  higher_composite("CABD", c("CD", "TD", "HA", "DC", "SH"), method = two_stage, weights
= mode_B),
  composite("CDe", multi_items("CDe_", 1:6)),
  composite("ED", multi_items("ED_", 1:2)),
  higher_composite("DPTD", c("CDe", "ED"), method = two_stage, weights = mode_B),
  composite("DF", multi_items("DF_", 1:2), weights = mode_B)
)
```

```

# Especificar el modelo estructural

# Paths

CABD_SM <- relationships(
  paths(from = "CABD", to = "DPTD"),
  paths(from = "CABD", to = "DF"),
  paths(from = "DPTD", to = "DF")
)

# Estimar el modelo

CABD_Model <- estimate_pls(
  data = CABD_DB,
  measurement_model = CABD_MM,
  structural_model = CABD_SM,
  missing = mean_replacement,
  missing_value = "-99"
)

# Graficar el modelo

plot(CABD_Model)

# Guardar resumen del modelo

summary_CABD_Model <- summary(CABD_Model)

# Revisar iteraciones del algoritmo para converger

summary_CABD_Model$iterations

# Realizar proceso de Bootstrap al modelo estructural

CABD_Model_Boot <- bootstrap_model(
  semnr_model = CABD_Model,
  nboot = 10000,
  cores = parallel::detectCores(),
  seed = 123
)

# Graficar el modelo

plot(CABD_Model_Boot)

```



```
# Guardar el resumen del modelo con Bootstrap con un nivel de significancia de 0.05 (prueba de dos colas)
```

```
summary_CABD_Model_Boot <- summary(CABD_Model_Boot, alpha = 0.05)
```

```
# Extraer el VIF para evaluar colinealidad
```

```
summary_CABD_Model$vif_antecedents
```

```
# Evaluar la relevancia y significancia de las relaciones estructurales del modelo
```

```
summary_CABD_Model_Boot$bootstrapped_paths
```

```
# Inspeccionar los efectos totales de los coeficientes de trayectoria del modelo con Bootstrap
```

```
summary_CABD_Model_Boot$bootstrapped_total_paths
```

```
write.csv(x = summary_CABD_Model_Boot$bootstrapped_total_paths, file =  
"summary_CABD_Model_Boot_bootstrapped_total_paths.csv")
```