

UNIVERSIDAD AUTONOMA DE BAJA CALIFORNIA

Facultad de Contaduría y Administración

Facultad de Ciencias Químicas e Ingeniería

Maestría en Tecnologías de la Información y la Comunicación



Modelo de almacén de datos aplicado a la rastreabilidad de productos frescos del campo: “Caso Espárragos”

PARA OBTENER EL GRADO DE

MAESTRO EN TECNOLOGIAS DE LA INFORMACION Y LA COMUNICACIÓN

Presenta:

Aldo Jorge Atondo Inzunza

Bajo la Dirección de:

MTIC Ricardo Javier Ibarra Ibarra

Codirección:

Dra. Esperanza Manrique Rojas

Tijuana, Baja California, México

Marzo 2013

Por su invaluable apoyo dedico a mi esposa Mary Monreal y a mi hija Camila, les doy las gracias por todo el amor brindado durante la maestría, que sin su ayuda no hubiera podido alcanzar una de mis metas, gracias por ser parte fundamental de mi vida; prometo recompensarles en más amor por su ayuda; las amo.

Agradecimientos

A mi familia

Agradezco de antemano a mi esposa y a mi hija la paciencia y amor brindado durante el curso de la maestría ya que sin él no hubiera sido posible la culminación de un escalón más durante mi desarrollo personal.

A mis padres

Agradezco a dios por haberme dado la oportunidad de tener unos padres como los que tuve, ya que sin el amor otorgado por ellos durante mi vida, el aliento de salir adelante y los consejos recibidos de siempre estar en el camino correcto, que con trabajo y dedicación no hay meta que se resista.

A mis hermanos

Gracias les doy por haberme dado ánimo de seguir siempre adelante en la superación personal, gracias por ser mis hermanos.

A mis maestros y amigos

Les agradezco por haber sido parte de este episodio de mi vida, gracias por todos los consejos recibidos, Dra. Margarita Ramírez gracias por todos sus consejos y buena vibra que siempre me dió para alcanzar esta meta; MTIC Ricardo Javier Ibarra Ibarra gracias también por todo su apoyo y asesoría brindada.

Resumen

El presente trabajo nace de la necesidad de cumplir con los requisitos legales de parte del gobierno de los Estados Unidos de América; está pidiendo que cualquier ente que realice exportaciones hacia su país cumpla con rastreabilidad, cada producto tiene que ser rastreable, es decir; que se le pueda identificar su origen, su proceso y ruta que siguió hasta llegar al mercado final; todo esto con el fin de que en cualquier momento que surja una contingencia de salud se le pueda seguir un rastro exacto y poder aislar su origen del resto del mercado; ya en la marcha del estudio nos dimos cuenta de que el modelo de almacén de datos que se propone les ayudará en muchas otras áreas dentro de las empresas involucradas en el proceso de comercialización, para este caso en productos frescos del campo.

Por parte del empacador, se podrá identificar exactamente hacia donde se fueron los productos originados de un lote productor, en que cantidades, en que presentaciones, rentabilidad neta de cada cosecha por lote, por rancho productor; se puede decir que la información que puede estar proporcionando un modelo de almacén de datos como el aquí presentado puede ayudar significativamente en la toma de decisiones a la gerencia de la empresa.

Índice

Índice De Figuras	X
INTRODUCCIÓN	1
CAPITULO I.....	2
1 EL PROBLEMA.....	2
1.1 Planteamiento del Problema.	2
1.2 Experiencias recientes.	3
1.3 Trazabilidad.....	3
1.4 Importancia de la trazabilidad.	4
1.5 Impacto de la trazabilidad.	5
1.6 Estándares Mundiales de la Cadena de Abastecimiento.	5
1.7 Experiencias en México.	6
1.7.1 Caso México.	6
1.7.2 Caso Baja California.	7
1.7.3 Exportación de Tomate.	9
1.7.4 Caso Sinaloa.	9
1.7.5 Guerra del Tomate.	11
1.8 Planteamiento del Problema.	12
1.9 Objetivos.	13
1.9.1 Objetivo del trabajo.....	13
1.9.2 Objetivos específicos.	13
1.10 Justificación.	13
1.11 Alcance	14
2 MARCO TEORICO	15
2.1 Antecedentes Históricos.....	15
2.2 Conceptos Básicos	16
2.2.1 Almacén de datos	16
2.2.2 Metadatos.	17

2.2.3 Datos detallados actuales. -----	17
2.2.4 Datos detallados históricos.-----	17
2.2.5 Datos ligeramente resumidos. -----	18
2.2.6 Datos muy resumidos. -----	18
2.2.7 Mercado de datos (Data Mart)-----	21
2.2.8 Características de un Mercado de Datos -----	21
2.2.9 Top-Down. -----	22
2.2.10 Bottom-Up. -----	22
2.2.11 OLTP (Online Transaction Proccesing). -----	23
2.2.12 OLAP (<i>Online Analitical Proccesing</i>).- -----	23
2.2.13 Tabla de Hechos-----	23
2.2.14 Tabla de Dimensiones.- -----	24
2.2.15 Cubo de Datos.- -----	24
2.2.17 Atributos.- -----	25
2.2.18 Jerarquías.- -----	25
2.2.19 Esquema Estrella.- -----	27
2.2.19 Esquema Copo de Nieve.- -----	29
2.2.20 Esquema Constelación.- -----	30
2.2.21 Inteligencia de Negocios (Bussines Intelligence, BI).- -----	30
2.3 Características de Almacén de datos. -----	31
2.4 Objetivos de un Almacén de datos.-----	32
2.5 Forma de Valorar un almacén de datos.-----	32
2.6 Rastreabilidad -----	33
2.7 Modelo de Rastreabilidad de la cadena de abastecimiento de productos frescos. -----	34
2.8 Mantenimiento de un registro exacto y a tiempo.-----	35
2.9 Composición del lote. -----	35
2.10 Comunicación electrónica de datos de Rastreabilidad. -----	35
2.11 Modelo de rastreabilidad -----	36
2.12 Números Mundiales de Artículo Comercial EAN.UCC -----	37
2.13 Códigos de barra GS1-----	38
2.14 Números Mundiales de localización -----	38
2.15 Definiciones de las unidades del consumidor, comerciales y logísticas -----	39

2.16 Definición de una unidad del consumidor.	39
.....	39
2.17 Definición de unidad comercial.	39
2.18 Definición de unidades logísticas.	40
2.19 Definición de Información de Atributo.	40
2.20 Alertas de Inocuidad en los Estados Unidos de América.	41
Propuesta de Solución.....	42
Herramientas de Rastreabilidad GS1.....	42

3 Metodología..... 44

3.1 Planeación y administración del proyecto.	45
3.2 Definir Requerimientos de Información.	45
3.2.1 Por parte del Distribuidor.....	46
3.2.2 Empacador.....	46
3.2.3 Por Lote y Campo.....	47
3.2.4 Por Linea de Producción.	47
3.2.5 Por Manifiesto.	47
3.2.6 Por Transporte.	47
3.2.7 Por distribuidor.	47
3.3 Diseño de Arquitectura.....	47
3.4 Trazabilidad a través de un portal WEB	48
3.5 Información Histórica	48
3.6 Equipo e Infraestructura	48
3.7 Selección Del Software.	49
3.7.1 Pentaho BI Platform.	49
3.7.2 Arquitectura.	50
3.7.3 Pentaho Reporting.	50
3.7.4 Pentaho Analysis.	52
3.7.5 Pentaho Dashboards.	52
3.7.6 Pentaho Data Integration.	53
3.7.7 Propiedades básicas.	53
3.7.8 Data Mining.....	54
3.8 METODOLOGIA APLICADA A LA TRAZABILIDAD	55

3.8.1.- Definición de Requerimientos.	55
3.8.2 Producción Total Por Lote y campo.	55
3.8.3 Producción por Línea de Empaque.	55
3.8.3.1 Conformación de una caja (Producto).	56
3.8.3.2 Conformación de Pallets	56
3.8.4 Manifiesto de Embarque.	57
3.9 Manifiestos de Embarque con el Distribuidor.	57
3.10 Grupo de Trabajo.	58
3.11 Requerimientos técnicos.	58
3.11.1 Portal de Internet.	58
3.12 Herramientas para análisis.	58
3.12.1 En el empacador.	59
3.12.2 En el distribuidor.	59
3.13 Información histórica.	59
3.14 Acceso a la información de forma oportuna.	60
3.15 Integración con otros sistemas.	60
3.16 Equipo e Infraestructura.	61
3.17 Modelado del Almacén de datos.	61
3.18 Modelo Dimensional.	62
3.18.1 Dimensiones.	62
3.18.2 Medidas (hechos).	62
3.19 Modelo Estrella.	62
3.20 Diseñar el data warehouse.	63
3.21 Implementación incremental para data warehouse.	63
3.21.1 Definición.	63
3.21.2 Modelado.	64
3.21.3 Construcción.	64
3.21.4 Producción.	65
4 Desarrollo	66
4.1 Distribuidor.	66
4.2 Empacador.	67
4.3 Productor.	68

4.4 Diagrama general. -----	69
4.5 La herramienta de planeación creada por Kimball. -----	70
4.6 Proceso que determina la rastreabilidad. -----	71
4.7 Proceso Envíos Realizados por empacador. -----	73
4.8 Proceso pallets realizados por empacador. -----	74
4.9 Proceso de productos empacados. -----	75
4.10 Proceso de Producción por Lote. -----	76
4.11 Proceso por Línea de Empaque. -----	77
4.12 Proceso de manifiestos recibidos por distribuidor. -----	78
4.13 Proceso de ventas a clientes minoristas por distribuidor. -----	79
4.14 Tabla de Hechos: Eventos -----	81
4.15 Tabla de Hechos: Envíos Realizados.-----	82
4.16 Tabla de Hechos: Pallets Realizados. -----	83
4.17 Tabla de Hechos: Cajas Empacadas. -----	84
4.18 Tabla de Hechos: Producción por línea de empaque. -----	85
4.19 Tabla de Hechos: Manifiestos Recibidos por distribuidor.-----	86
4.20 Tabla de Hechos: Ventas del Distribuidor. -----	87
4.21 Tabla de Hechos: Producción por Lote. -----	88
4.22 Tabla de Dimensiones: Fecha.-----	89
4.23 Tabla de Dimensiones: Empacador.-----	90
4.24 Tabla de Dimensiones: Distribuidores. -----	91
4.25 Tabla de Dimensiones: Transporte. -----	92
4.26 Tabla de Dimensiones: Zona Geográfica. -----	93
4.27 Tabla de Dimensiones: Productos. -----	94
4.28 Tabla de Dimensiones: Líneas de Empaque. -----	95
4.29 Tabla de Dimensiones: Lote y Campos. -----	96
4.30 Tabla de Dimensiones: Empleados. -----	96
4.31 Tabla de Dimensiones: Minorista.-----	98
4.32 Tabla de Dimensiones: Envíos. -----	99
4.33 Tabla de Dimensiones: Pallets. -----	99
4.34 Tabla de Dimensiones: Cajas.-----	100
4.35 Desarrollo del Modelo físico de datos. -----	101

4.35.1 Base de Datos	101
Conclusión	111
Líneas Futuras	112
Referencias.....	116

Índice De Figuras

<i>Figure 1 Estructura de un almacén de datos.</i>	<i>17</i>
<i>Figure 2 Estructura lógica de un almacén de datos.</i>	<i>18</i>
<i>Figure 3 Diferencias en los modelos de W. Inmmon y R. Kimball.</i>	<i>20</i>
<i>Figure 4 Arquitectura Top-Down</i>	<i>22</i>
<i>Figure 5 Arquitectura Top-Down.....</i>	<i>22</i>
<i>Figure 6 Arquitectura Bottom-Up</i>	<i>23</i>
<i>Figure 7 Ventas totales por Producto.</i>	<i>24</i>
<i>Figure 8 Ventas totales por Mes.</i>	<i>25</i>
<i>Figure 9 Ventas Totales.</i>	<i>26</i>
<i>Figure 10 Ventas totales por año.</i>	<i>26</i>
<i>Figure 11 Ventas Totales por Producto por año.</i>	<i>26</i>
<i>Figure 12 Cubo de Ventas Totales (Regiones, Productos, Tiempo).</i>	<i>27</i>
<i>Figure 13 Esquema Estrella</i>	<i>28</i>
<i>Figure 14 Esquema Copo de Nieve.</i>	<i>29</i>
<i>Figure 15 Esquema de Constelación.</i>	<i>30</i>
<i>Figure 16 Ejemplo de un archivo EDI.</i>	<i>36</i>
<i>Figure 17 Modelo de rastreabilidad de productos frescos.</i>	<i>37</i>
<i>Figure 18 Número mundial de artículo comercial.</i>	<i>37</i>
<i>Figure 19 Unidades del consumidor.</i>	<i>39</i>
<i>Figure 20 Unidad Comercial.</i>	<i>40</i>
<i>Figure 21 Unidad Logística.</i>	<i>40</i>
<i>Figure 22 Diagrama del ciclo de vida de Kimball.</i>	<i>44</i>
<i>Figure 23 Diagrama general.</i>	<i>69</i>
<i>Figure 24 Herramienta de planeación para la rastreabilidad de productos frescos.</i>	<i>70</i>
<i>Figure 25 Eventos.</i>	<i>71</i>
<i>Figure 26 Modelo de Envíos realizados por empacador.</i>	<i>73</i>
<i>Figure 27 Modelo de pallets realizados por empacador.</i>	<i>74</i>
<i>Figure 28 Modelo de productos empacados.</i>	<i>75</i>
<i>Figure 29 Modelo de producción de Lote-Campo.</i>	<i>76</i>
<i>Figure 30 Modelo de producción por línea de empaque.</i>	<i>77</i>

Figure 31 Modelo manifiestos recibidos por distribuidor.	78
Figure 32 Modelo de ventas a Clientes minoristas por distribuidor.	79
Figure 33 Tabla de hechos, Fact_Eventos.	81
Figure 34 Tabla de Hechos, FACT_Envios_realizados.	82
Figure 35 Tabla de Hechos, FACT_Pallets_realizados.	83
Figure 36 Tabla de hechos, FACT_Cajas_empacadas.	84
Figure 37 Tabla de Hechos, FACT_LineasdeEmpaque.	85
Figure 38 Tabla de Hechos, FACT_Recibos_Distribuidor.	86
Figure 39 Tabla de Hechos, FACT_Distribuidor_Ventas.	87
Figure 40 Tabla de Hechos, FACT_Produccion_Lote.	88
Figure 41 Tabla de Dimensiones, DIM_FECHA.	89
Figure 42 Tabla de Dimensiones, DIM_Empacador.	90
Figure 43 Tabla de dimensiones, DIM_Distribuidores.	91
Figure 44 Tabla de Dimensiones, DIM_Transporte.	93
Figure 45 Tabla de Dimensiones, DIM_Zona_Geografica.	93
Figure 46 Tabla de Dimensiones, DIM_Productos.	95
Figure 47 Tabla de Dimensiones, DIM_LineasDeEmpaque.	95
Figure 48 Tabla de Dimensiones, DIM_Lote_Campo.	96
Figure 49 Tabla de Dimensiones, DIM_Empleados.	97
Figure 50 Tabla de Dimensiones, DIM_Clientes.	98
Figure 51 Tabla de Dimensiones, DIM_Envios.	99
Figure 52 Tabla de Dimensiones, DIM_Pallets.	100
Figure 53 Tabla de Dimensiones, Dim_Cajas.	100
Figure 54 Modelado dimensional, Para la rastreabilidad de productos frescos.	102
Figure 55 - Estadísticas de producción principal: Para pago de productividad.	105
Figure 56 - Estadísticas de ranchos productores por Lote.	106
Figure 57 - Estadísticas por Línea de producción.	107
Figure 58 – Modelo de Reporte: Producción de cajas (producto) por lote.	109
Figure 59 Modelo de Reporte: Cantidad de cajas enviadas a minorista por distribuidor.	110

INTRODUCCIÓN

La rápida evolución de las funciones en la cadena de valor, los nuevos canales de distribución, los patrones de demanda cambiantes y las mayores expectativas en el área de servicios han contribuido a incrementar la importancia de las tecnologías de la información en los procesos comerciales.

La seguridad y confianza que el consumidor pueda tener al momento de comprar se ha convertido en una de las cuestiones más críticas y prioritarias de la cadena de abastecimiento alimentaria. A pesar de los esfuerzos de los integrantes de la cadena alimentaria de productos frescos, los problemas de seguridad alimentaria no pueden excluirse por completo. Sin embargo, un sistema de trazabilidad efectivo y de costo-eficiente puede indicar con precisión tal problema, dentro de una región específica, lugar de embalaje, grupo de productores, productor o incluso un área, más que dentro de todo un grupo de producción. Al reducir el potencial campo de acción de un problema, se actúa de conformidad a los requisitos legales, y se puede reducir el impacto económico negativo sobre los integrantes de la cadena de abastecimiento, los que no siempre son responsables de que todo esto suceda.

Desde una perspectiva de salud pública, al mejorar la velocidad y exactitud del rastreo y localización de los artículos alimentarios implicados en alguna contingencia pública, se puede ayudar a disminuir los riesgos relacionados a la seguridad alimentaria. La rápida y efectiva trazabilidad puede, también, minimizar el gasto innecesario de recursos privados y públicos, y aliviar las preocupaciones del consumidor. Más aún, el rastreo de artículos alimentarios en cuestión, puede ayudar a los servicios de salud pública y a los operarios industriales a determinar las causas potenciales de un problema, con lo que proporcionarían la información necesaria, para identificar y minimizar los riesgos a la salud.

CAPITULO I

1 EL PROBLEMA

1.1 Planteamiento del Problema.

El temor en Estados Unidos de un ataque bioterrorista a través de sus fronteras, y el creciente índice de brotes de enfermedades tales como Salmonella, ligadas de manera directa o indirecta a productos del campo tanto de EEUU como del extranjero, han estimulado que las dependencias gubernamentales de nuestro vecino país del norte actúe creando legislaciones que afectan de forma directa el comercio internacional y que redefinen el concepto de rastreabilidad de productos.

La ley “Bioterrorism Act” promulgada en el año 2002, empezó a hacer que todos los entes económicos integrados en cadenas de abasto de cualquier tipo de productos llevaran sistemas de rastreabilidad o registros de sus movimientos por lo menos de un paso hacia atrás y otro hacia adelante en la cadena de suministro.

El FDA (*Food and Drug Administration*) y el USDA (*United States Department of Agriculture*) propusieron un cambio legislativo que incluye la autorización de aumentar en el escrutinio e inspecciones a productos de importación; lo cual el congreso aprobó el requerimiento de rastreabilidad con el objetivo de tener mejor visibilidad, rapidez para determinar el origen y consecuentemente contener el tamaño en casos de brotes futuros.

A raíz de estos sucesos, la industria privada en Estados Unidos (*Produce Marketing Association PMA* , *Canadá Produce Marketing Association CPMA* y *United Fresh*), unieron esfuerzos para la creación y puesta en práctica del “*Produce Traceability Initiative*” (PTI) que tiene como objetivo el impulsar un estándar común para la rastreabilidad de la cadena de abasto en la industria.

Los recientes brotes de enfermedades provocados por los virus de Salmonella ligados a alimentos y el impacto económico a las diferentes industrias así como por cierre de fronteras a productos hortofrutícolas mexicanos, han causado que la cadena de abasto de alimentos este bajo investigación constante; a continuación se hace una remembranza de los acontecimientos en los que México ha sido perjudicado.

1.2 Experiencias recientes.

Las experiencias con productos mexicanos han sido desastrosas, como se puede apreciar en la tabla 1, impactos recientes en los problemas sanitarios de productos frescos.

Tabla 1, Grandes impactos en la economía mexicana. (CAADES, 2010)

PRODUCTO	IMPACTO
Fresas Mexicanas	En Baja California, a finales de los 80's un distribuidor de San Diego, California fue enviado a la cárcel.
Cebollín Mexicano	En los 90's por Hepatitis 'A', el valle de Mexicali, en México tardo un año en regresar al mercado.
Melón Cantaloupe	A principios del 2000 por Salmonella, México perdió materialmente el mercado; de 205 mdd que México exportaba en el 2000, en al año 2009 solo exportó 14mdd.

1.3 Trazabilidad.

El termino de trazabilidad en su definición de la Organización internacional para la estandarización (ISO):

“Es la propiedad del resultado de una medida o del valor de un estándar donde este puede estar relacionado con referencias especificadas, usualmente estándares nacionales o internacionales, a través de una cadena continua de comparaciones todas con incertidumbres especificadas.”

Según el comité de seguridad alimentaria de AECOC:

“Se entiende como trazabilidad aquellos procedimientos preestablecidos y

Autosuficientes que permiten conocer el histórico, la ubicación y la trayectoria de un producto o lote de productos a lo largo de la cadena de suministros en un momento dado, a través de unas herramientas determinadas.”

La trazabilidad es la habilidad de conocer en cualquier momento de consulta la posición exacta de un producto dentro de la cadena de abasto, dando a conocer el historial detallado de su ubicación física actual, de su origen, su destino. (www.sigweb.cl/biblioteca/Trazabilidad.pdf, 2011)

La Trazabilidad Interna y la Trazabilidad Externa

A la hora de tener que entender la Trazabilidad de un producto que se mueve a través de su cadena de suministro o de su cadena logística, el concepto de trazabilidad se divide en dos partes bien diferenciadas:

- **La Trazabilidad Interna**, que no es más que poder obtener la traza que va dejando un producto por todos los procesos internos de una compañía, con sus manipulaciones, su composición, la maquinaria utilizada, su turno, su temperatura, su lote, etc., es decir, todos los indicios que hacen o pueden hacer variar el producto para el consumidor final.
- **La Trazabilidad Externa**, que no es más que poder externalizar los datos de la traza interna y añadirle algunos indicios más si fuera necesario, como una rotura del embalaje, un cambio en la cadena de temperatura, etc.

1.4 Importancia de la trazabilidad.

La trazabilidad es muy importante, particularmente por el impacto que los incidentes y las preocupaciones por la seguridad alimentaria, tienen sobre los consumidores, compañías, grupos de productos, gobiernos y comercio. Debido a la diversidad de prácticas internacionales en la cadena de abastecimiento de productos, es de suma prioridad que los productores, empacadores, importadores / exportadores y proveedores logísticos, trabajen con sus socios en la distribución y la venta al por menor, a fin de desarrollar tecnologías y estándares que permitan la identificación, a fin de acompañar al producto, desde el campo hasta el minorista. Acordada según procedimientos y estándares para la identificación y documentación del producto, la

trazabilidad permitirá una indagación rápida y automatizada de las unidades comerciales del producto individual, desde el campo hasta el minorista.

1.5 Impacto de la trazabilidad.

La implementación de los sistemas de trazabilidad públicos y privados, por medio de la captura automatizada de datos, procesamiento electrónico de datos y comunicaciones electrónicas, puede mejorar significativamente la exactitud y velocidad del acceso a la información sobre la producción y destino del alimento. Estos pueden disminuir el riesgo y la inseguridad de la cadena de abastecimiento y entre las partes comercializadoras. Todo esto, sin embargo, requiere una visión general de la cadena de abastecimiento alimentaria, la que solamente se puede lograr al desarrollar estándares comerciales internacionales.

1.6 Estándares Mundiales de la Cadena de Abastecimiento.

El motivo por el cual se usan estándares comerciales aceptados internacionalmente, es superar las barreras del comercio que crean los estándares específicos nacionales, de la industria y de la empresa, cuando se les usa en lugar de estándares internacionales para multi-industria. Comercializar, rastrear y localizar mercaderías es más costoso debido a la necesidad de cumplir con todos los requerimientos legales y con los requisitos de comunicación de cada país importador o empresa. La clave para elaborar sistemas de trazabilidad eficientes y de costo-efectivo, es satisfacer los requisitos legales y del cliente, al aplicar un estándar mundial.

El Sistema de normas aplicadas por *Global System, Global Standard, Global Solution* (GS1) permite el manejo eficiente de la cadena de abastecimiento y del comercio internacional, al proporcionar herramientas estándar que permiten a todos los integrantes de la cadena de abastecimiento de productos, comunicarse en un mismo idioma comercial mundial. Los conceptos clave relacionados a la aplicación del sistema GS1, pueden resumirse en tres áreas:

- 1.- Automatización de los procesos comerciales por medio de la captura automatizada de datos (ADC) y del procesamiento electrónico de datos (EDP).

2.- Comunicación de la información del modo más rápido y más exacto, por medio de mensajes electrónicos estándar que automáticamente actualizan las aplicaciones de la computadora, con datos provenientes de los pares que intervienen en el comercio.

3.- Compresión de tiempo, lo que ofrece oportunidades estratégicas para mejorar la satisfacción al cliente, no solo por medio de la eficiente Rastreabilidad del producto, sino también por medio de procesos comerciales de re-ingeniería, a través de la cadena de abastecimiento.

En México se han tenido diversas experiencias relacionadas con la inocuidad alimentaria, en donde diversos productos se vieron envueltos en brotes de enfermedades relacionados al campo, a continuación se mencionan algunos casos más sonados.

1.7 Experiencias en México.

A continuación se hace una remembranza de acontecimientos ocurridos y que afectaron a México, acusándole de ser el causal de problemas de inocuidad.

1.7.1 Caso México.

<http://www.grooultimasnoticias.com/Archivo%20Historico/2008/Junio/Ultimas%20Noticias%2019.06.08%20%28544%29.pdf>

En el pasado reciente (2008) Estados Unidos acusó a México de que sus exportaciones de tomate, fresa, melón y cilantro estaban contaminadas de salmonella, lo cual no se pudo comprobar; El investigador Antonio Segura Miranda de la Universidad Autónoma de Chapingo consideró que detrás de la campaña de supuesta salmonella orquestada contra el tomate mexicano “podría estar un nuevo boicot de Estados Unidos”; es difícil pensar que Estados Unidos, que cuenta con una ley “muy dura de bioterrorismo”, desconozca la ruta que siguió un producto supuestamente contaminado, ya que ellos incluso pueden penalizar a los responsables de esa cadena porque los considerarían bioterroristas.

Además en caso de que se llegara a comprobar que el tomate contaminado salió de México sería a través de la rastreabilidad; México salió libre de esos pretextos sanitarios, los productos mexicanos han sufrido un daño importante en sus exportaciones, por lo que el gobierno debe hacer una férrea defensa con bases

científicas para demostrar que el tomate está libre de salmonella; Lo que México a través de las empresas productoras, de transporte puede hacer es que se le siga la rastreabilidad a la producción para conocer incluso de qué huerto salió la bacteria, dado que en la actualidad los campos agrícolas tienen sus zonas de empaque.

1.7.2 Caso Baja California.

12 Junio de 2008

Fuente: Ensenada.net

B.C., no es fuente salmonella tomates

Es imposible que Baja California sea la entidad de donde se derivó un brote de salmonella en Estados Unidos, dado que la alerta se emitió, 10 días antes que Baja California inició la exportación de su producción tomatera.

En un comunicado de Prensa Genaro López Bojórquez Delegado de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA), en Baja California, asegura que el brote fue detectado el 16 de abril de 2008; mientras la producción del tomate cultivado en Baja California, comenzó a circular en los mercados comerciales a partir del 26 de abril de 2008; esto representa 10 días después de que Estados Unidos emitiera la alerta sanitaria tras la detección en los Estados de Texas y Nuevo México, por lo que existe la certeza de que el tomate de Baja California no fue el agente causal de dicho brote.

Esta situación puede ser comprobada por el Departamento de Comercio de Estados Unidos (USDOC), ya que el “Acuerdo de Suspensión de *Dumping* para Tomates Mexicanos”, obliga a los productores mexicanos a marcar cada caja que exporta; esta contiene una etiqueta con el nombre del productor y el número de identificación asignado por el mismo Departamento de Comercio, por lo que una irregularidad, hubiera sido reclamada y publicada desde el primer momento.

Agregó que México aporta el 85% del tomate que se consume en Estados Unidos, debido a la calidad con la que se produce, la que está certificada por empresas internacionales; en el 2007, México exportó a Estados Unidos 949 mil toneladas de tomate fresco, para mantenerse como la principal nación que comercializa el producto

en este país; esto hace muy poco probable que el brote de salmonella detectado en Estados Unidos, proceda de tomates producidos en México en particular del Estado de Baja California.

El funcionario explicó que el tomate que se produce en México, y en especial en los campos de Baja California, son tomates que guardan las reglas de inocuidad alimentaria de las denominadas “Buenas Prácticas de Agricultura y Manufactura”.

Aunado a esto se tiene la certificación de los campos, los trabajadores, las fuentes de agua y los equipos que operan y están en contacto con la fruta.

López Bojórquez comentó que SAGARPA y el Gobierno del Estado a través de la Secretaría de Fomento Agropecuario, y ante la presencia del Director General de Sanidad Vegetal, Javier Trujillo y el Director de Inocuidad SAGARPA-SENASICA Juan de los Ríos, se estableció un acuerdo con los productores de tomate de la entidad, para que aporten toda la información requerida para poder concluir satisfactoriamente el estudio de análisis de riesgo; que compruebe que el tomate de Baja California no fue el agente causal del problema que se generó en el vecino país del Norte.

El Delegado aclaró que la elaboración de este análisis junto con el dato de las fechas en que dio inicio la producción de tomate en Baja California, permitirán que la Administración de Drogas y Alimentos (FDA) enliste a México, y en particular a Baja California, entre los países o regiones que producen tomates sanos y seguros.

Agregó, que la Secretaría ha enviado a un técnico mexicano para que participe junto con la FDA, en las investigaciones sanitarias correspondientes en Estados Unidos.

Declaró que actualmente la frontera no está cerrada para los tomates mexicanos, puesto que México puede exportar hacia los mercados de Estados Unidos; y en Baja California, de acuerdo a información proporcionada por las autoridades de la Secretaría de Salud (ISESALUD), no se tiene confirmado ningún reporte de brote de salmonella asociado con el consumo de tomates.

López Bojórquez añadió que el cultivo del tomate reviste una gran importancia económica para el Estado de Baja California, ya que durante el ciclo Otoño-Invierno

2007-2008 se logró establecer una superficie de 3,371 hectáreas, que en promedio producen 50 toneladas por hectárea, una derrama económica por el orden de los 1,200 millones de pesos, y más de 3000 empleos directos en la entidad.

Sin embargo, existen también invernaderos con tecnología de punta, donde la producción de tomate, es de más de 220 toneladas por hectárea uno de los rendimientos más altos del mundo en ambiente controlado. En la tabla 2 se pueden ver los números de producción en el caso de tomate en el año 2008.

1.7.3 Exportación de Tomate.

Tomate Mexicano en Estados Unidos: 85% de la producción mexicana

Tabla 2: Tomate Baja California 2008

Baja California	Inició Exportaciones	26/04/2008
Producción	50 Toneladas	Superficie Sembrada
Campo Abierto		3 371 Hectáreas
En Invernadero	Hasta 220 Toneladas	
Empleos Generados	3000	
Derrama Económica	1200 mdp	
Alerta Sanitaria	Departamento de comercio de Estados Unidos de América (USDOC)	16/04/2008

1.7.4 Caso Sinaloa.

12 Junio de 2008

Fuente: El Universal

La campaña 'negra', para ayudar a labriegos de EU

CULIACÁN, Sin. — La guerra comercial del tomate rojo no es nueva. Desde hace más de 100 años, Estados Unidos ha mostrado interés de cerrar la frontera norte al fruto mexicano o por lo menos de obstaculizar su distribución para beneficiar a productores de esa nación.

Ya sea por supuesto *dumping* (competencia desleal en la oferta de productos en mercados extranjeros), la intención de fijar aranceles, la restricción de exportar según el diámetro de la fruta y ahora por presunta infección de salmonelosis.

En una de las tantas defensas contra esta cerrazón, el sr. José Carlos de Saracho Calderón (Expresidente de la Unión Nacional de productores de hortalizas) advirtió que la medida obligaría a más de 200 mil jornaleros mexicanos a decidir entre dos alternativas: dedicarse al cultivo ilícito o cruzar de forma ilegal el río Bravo.

Por su parte el Sr. Armando Borboa López, presidente de la Asociación de Agricultores del Río Sinaloa Poniente comentó que esta nueva alerta es parte de una campaña negativa con fines proteccionistas para favorecer a los productores de aquel país que compiten con la producción enviada de Sinaloa.

De Saracho Calderón indicó que la competencia comercial entre agricultores de Sinaloa, en México, y de Florida, en Estados Unidos, se vuelve más tensa, sobre todo porque los productores de ese país no sólo están expuestos a las inclemencias climatológicas, como ciclones y heladas, sino a las alzas del petróleo y sus derivados. Subrayó que los horticultores estadounidenses no aceptan que las cosechas mexicanas cubran su déficit en esta clase de alimentos, sobre todo en época de invierno, cuando las condiciones climatológicas extremas lo reclaman.

Además, señaló, en Florida los costos de la tierra de uso agrícola se elevan por la expansión urbana, así como por los combustibles, fertilizantes, transporte, plásticos y la escasez de mano de obra en sus surcos; la política migratoria golpea sus costos de producción. En su opinión, el uso de la mano de obra ilegal en los campos agrícolas en Estados Unidos se va a complicar y en consecuencia los inversionistas tendrán que elevar los salarios para atraer a obreros y con esto se elevarán los costos de sus tomates y otras hortalizas.

A raíz de que en 1907 se declaró a los horticultores mexicanos proveedores confiables de los mercados estadounidenses, sus homólogos de Florida han recurrido a instrumentos económicos, financieros, legales, políticos y propagandísticos para reducir su participación en los abastos de invierno.

1.7.5 Guerra del Tomate.

12 Junio de 2008

Fuente: El Universal

Productores se alistan contra guerra del tomate

Productores de tomate rojo sinaloense se alistan para librar una guerra comercial en caso que Estados Unidos determine boicotear las importaciones de este cultivo; por lo pronto, ya existe allá una alerta sobre el consumo de las variedades mexicanas Roma-Saladette y Bola.

Las exportaciones mexicanas de tomate —Sinaloa aporta casi la mitad—, que el año pasado significaron ingresos por 572 millones de dólares, sufren una nueva ofensiva en los mercados de Estados Unidos, ante la insistente versión de que pueden estar infestados por la bacteria de la salmonella.

En un comunicado conjunto, las secretarías de Agricultura y de Salud exigieron a las autoridades estadounidenses que se concluya, a la brevedad, la investigación que ha iniciado y determine el origen del brote de salmonelosis en 17 entidades de esa nación, en donde hay reportes de un total de 167 pacientes, de los cuales 27 requirieron hospitalización; los productores sinaloenses que aportan 45% de las ventas de este fruto agropecuario temen un boicot. El gerente de la Comisión de Investigación y Defensa de las Hortalizas, Mario Haroldo Robles Escalante, expuso que aunque las autoridades estadounidenses no determinan todavía el origen de la contaminación por salmonella en el fruto, algunas informaciones se lo atribuyen al tomate mexicano.

Dado que la bacteria que causa estos brotes fue identificada como *Salmonella Saintpaul*, que es poco común, ésta puede venir de diversas fuentes: agua, insectos, suelo y excremento de animales domésticos, entre otros; las informaciones que actualmente se manejan en medios de comunicación de Estados Unidos apuntan a que, en este caso, la bacteria proviene de las variedades de tomate rojo Roma y Bola de origen mexicano. “Pero llevan una intención distinta a la alerta emitida, máxime que en Sinaloa no se tiene ningún reporte de casos de salmonelosis por consumos de tomate”, aseguró.

Anunció que de darse una acción formal encaminada a parar el ingreso de esta verdura a Estados Unidos, los productores sinaloenses se alistan para su defensa, consistente ésta en contar con registros de control de siembras, recolección de los frutos, empaque y traslado en camiones refrigerados a la frontera norte.

En la temporada 2006-2007, de las 753 mil 603 toneladas de tomate rojo —en todas sus variedades— exportadas por México, el estado de Sinaloa participó con 316 mil 828 toneladas, en su mayoría Roma y Bola; subrayó que llama la atención que el 7 de junio la Administración de Drogas y Alimentos (FDA, por sus siglas en inglés) estadounidense publicó una lista de zonas domésticas y de países donde crece el tomate que no están asociados al brote de la salmonelosis.

De ese listado se desprende que cosechas en proceso de recolección de Canadá, Guatemala, República Dominicana, Puerto Rico e Israel no están bajo investigación, ya que aún no llegan a sus mercados. No obstante, no sucede lo mismo con el producto de los estados de Baja California, Jalisco, Michoacán, San Luis Potosí y Coahuila que aún no exportan. Dijo que a diferencia de sus competidores del estado de Florida, en México las siembras de tomate rojo se realizan en vara, no a ras de tierra. Además, se adoptan controles de inocuidad y certificación por parte de autoridades mexicanas y empresas extranjeras.

En Sinaloa, desde hace tiempo se siembran nuevas variedades resistentes a la vulnerabilidad climática y plagas, y su maduración en la planta le da mayor sabor, consistencia y vida en anaquel, enumeró el especialista.

1.8 Planteamiento del Problema.

El no identificar la rastreabilidad de los posibles productos contaminados por alguna bacteria, que pone en peligro la salud de millones de personas, se tiene el riesgo de acusar indebidamente a un país, una región en especial de ser el causal de un problema sanitario; lo cual le causa grandes pérdidas al país, pérdida de empleos.

La inocuidad es un tema muy sensible que preocupa a los productores agrícolas cuando aparecen brotes de contaminación. Afecta al Comercio, la imagen del país, la imagen del producto, existe la sospecha de si fue accidental o provocado, se pierde la

confianza en el producto y en la fuente de abasto; la percepción del público consumidor sobre el origen del producto; existe la tendencia a responsabilizar a la fuente primaria de producción; los casos de contaminación son el tema preferido de los medios de comunicación, se abusa de la información, causando pánico entre los consumidores, se interrumpe el abasto, obliga al retiro de todo el producto en la cadena. Mientras más dure el tiempo en la rastreabilidad, más incertidumbre y más pérdidas, el que siempre sale perdiendo es el productor debido a que se le detiene su producción sin haber certeza de que la misma sea la causante de tal eventualidad; para fines de este caso práctico se revisó el proceso en una empresa que comercializa espárragos.

1.9 Objetivos.

1.9.1 Objetivo del trabajo.

Realizar un modelo de Almacén de Datos que ayude Identificar a través de la rastreabilidad los productos, su origen, y ubicación dentro de la cadena de abastecimiento de una manera rápida; que ayuden a determinar el origen exacto de un producto que se ha identificado como posible causante de una contingencia de salud.

1.9.2 Objetivos específicos.

Realizar un modelado de *Data Warehouse* con los datos generados desde el productor hasta el minorista en la cadena de suministro de productos frescos.

Modelar reportes que ayuden a identificar la producción por lotes, por líneas de producción, por empacador.

1.10 Justificación.

Lograr una cuantificación económica de los factores de valor no es fácil ni natural a diferencia de los factores de costos, agregar valor económico a los factores de valor resulta ser en extremo complejo y subjetivo. Una alternativa es hacer una valoración desde la perspectiva de costos evitables, relacionados con los “costos de no disponer en la organización de información apropiada”, tanto a un nivel técnico como de procesos empresariales (en especial, para el proceso de Toma de Decisiones).

En este tipo de proyectos es difícil estimar de antemano con exactitud los beneficios económicos, aunque si el valor que introduce en la organización que se implementa,

pero se puede mostrar en base a estadísticas realizadas el beneficio que se obtendrá al mediano y largo plazo.

Realizar un almacén de datos es una estrategia de largo plazo. Al querer implementar un almacén de datos, se debe evaluar el costo y el valor considerando un período de tiempo razonable para obtener beneficios. El retorno sobre la inversión de un almacén de datos, se comienza a percibir bastante más tarde del tiempo en el cual se realizó la inversión inicial. Si se calcula costo/valor desde una perspectiva de corto plazo, los costos serán significativamente más altos en proporción al valor; pero sin embargo en el modelo almacén de datos que se presenta no se necesitará mucho tiempo para mirar los resultados ya que en un corto tiempo se podrá ver la rastreabilidad de todos los productos comercializados en esta cadena de abasto.

1.11 Alcance

El presente caso tiene como objetivo modelar el almacén de datos que ayude a determinar el origen de un producto y hacia donde se distribuyeron tales productos cuyo origen sea el mismo, cantidades, fechas y entes participantes en la cadena de suministro.

Capítulo II

2 MARCO TEORICO

2.1 Antecedentes Históricos.

La humanidad siempre ha tenido la necesidad de proveerse alimentos sanos y confiables; a través de la historia han existido situaciones en la que la ingesta de alimentos contaminados por alguna bacteria han ocasionado muertes, lo cual ha hecho que la sociedad sea más cuidadosa al momento de elegir dónde comprar los productos frescos del campo; ya sea en los mercados de su mayor confianza, y a su vez los mercados se abastecen de los distribuidores que mayor seguridad y confianza le dan en la inocuidad de los mismos, haciendo de todo esto que todos los socios comerciales involucrados en la cadena de abasto cumplan con la normatividad vigente en la inocuidad alimentaria; siendo esto de suma importancia.

Las ultimas regulaciones federales en los Estados Unidos de América han sido muy específicos al momento de comercializar con los productos que ingresen a su territorio, le han pedido a cada socio comercial que cuando menos lleven un registro antes y después de la compra de un producto; con esta regulación se pretende identificar de manera exacta todos los entes involucrados en la cadena de suministro de productos frescos; si en dado caso ocurriera una emergencia sanitaria donde se viera involucrado un producto fresco del campo, se pueda tener la certeza y seguridad de poder saber que ruta siguió dicho producto desde el campo hasta el consumidor final.

Aprovechando el auge tecnológico actual y las herramientas que existen para llevar un registro detallado de todos los movimientos que ha hecho un producto a través de la cadena de suministro, donde ya no existe la limitación de los dispositivos de almacenamiento, he propuesto unir todos los datos que se generan en todos los

eslabones de la cadena en un solo lugar, donde sea posible de una forma rápida y eficaz; identificar todos los eventos generados por un producto desde el momento en que se generó hasta llegar al consumidor. Pero no solo eso podrá ser identificado, al reunir gran cantidad de datos se podrá extraer otra información relevante para cada una de las partes involucradas.

Para tal fin se ha modelado un almacén de datos, el cual la importancia y relevancia del mismo radica en que los datos están organizados de tal manera que permite el análisis y consulta de los mismos de una manera rápida.

Las tecnologías de la información aplicadas en el desarrollo de soluciones informáticas ayudan a la sociedad actual a proveerse de información valiosa de forma rápida, misma que ayuda a tomar las mejores decisiones a la gerencia de una empresa o negocio.

2.2 Conceptos Básicos

A continuación para contextualizar el ambiente donde se desarrolló el modelado de la solución.

2.2.1 Almacén de datos

Ha sido definido como la integración de datos que pueden provenir de diferentes fuentes, en diferentes formatos y que han sido consolidados en un solo lugar con el propósito de que puedan ser analizados, son datos no volátiles y que están en un medio físico diferente al de su origen, están organizados de tal manera que permiten por medio de técnicas OLAP ver los datos como en un cubo; algunos especialistas en la materia lo definen como los datos que giran alrededor de un tema específico y que están organizados de tal forma que le ayudan a la gerencia en la toma de decisiones y que dichos datos son una copia de las transacciones y están organizados en datamarts. (Bill Inmon, Kimball), en la figura 1 se puede ver en forma general la estructura de un almacén de datos.

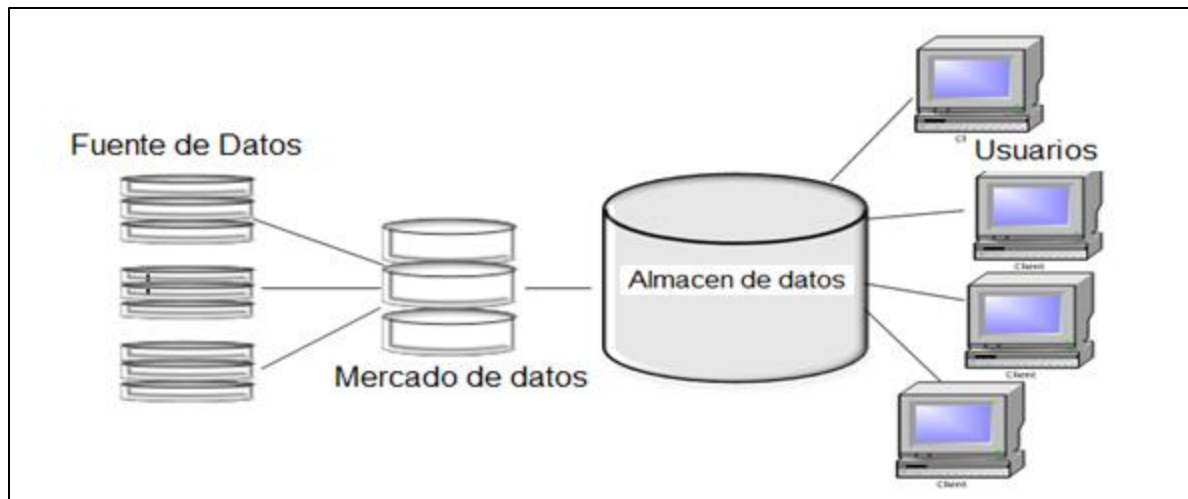


Figure 1 Estructura de un almacén de datos.

La estructura lógica de un Almacén de Datos está compuesta por los siguientes niveles:

2.2.2 Metadatos.

Describen la estructura de los datos contenidos en el almacén.

- Están ubicados en una dimensión distinta al resto de niveles.

2.2.3 Datos detallados actuales.

Se obtienen directamente del procesamiento de los datos.

- Ellos forman el nivel más bajo de detalle.
- Ocupan mucho espacio.
- Se almacenan en disco, para facilitar el acceso.

2.2.4 Datos detallados históricos.

Se obtienen directamente del procesamiento de los datos; pero con datos correspondientes al pasado.

- Se suelen almacenar en un dispositivo externo, debido a que su acceso es poco frecuente.

2.2.5 Datos ligeramente resumidos.

Conforman el primer nivel de agregación de los datos detallados actuales.

- Estos dan respuesta a las consultas habituales.
- Se almacenan en disco.

2.2.6 Datos muy resumidos.

Pertenecen al nivel más alto de agregación.

- Estos dan respuesta a las consultas que se realizan muy a menudo y por lo tanto se deben obtener muy rápidamente.
- Suelen estar separados del Almacén de datos, regularmente en *mercados de Datos (Data Marts)*.

En la figura 2 se puede apreciar la estructura lógica de un almacén de datos.

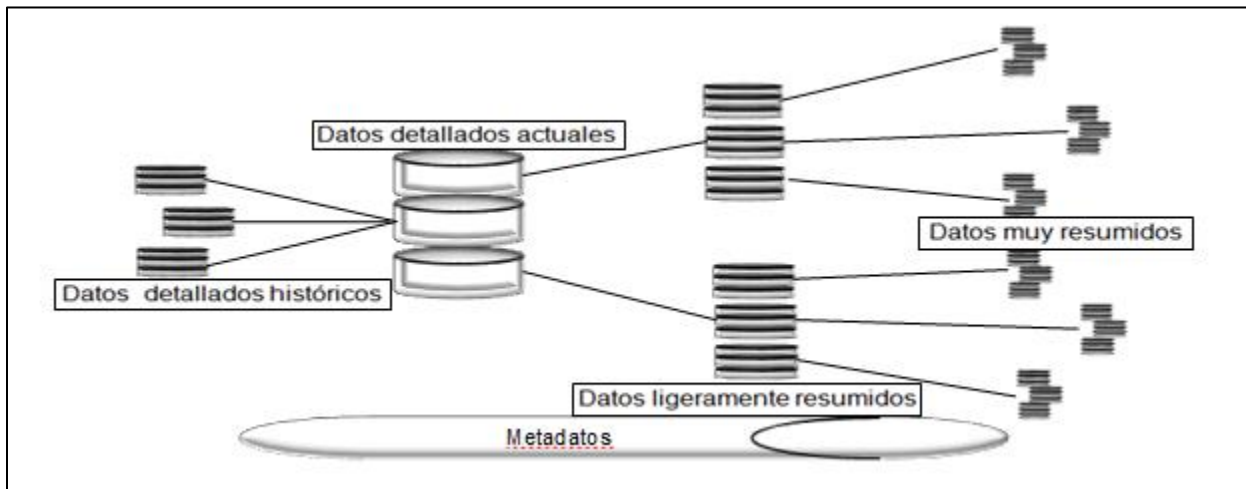


Figure 2 Estructura lógica de un almacén de datos.

Para desarrollar un almacén de datos sobresalen dos metodologías muy utilizadas actualmente; por un lado está la de William Inmon y por otro lado está la de Ralph Kimball.

Ralph Kimball propone un desarrollo de almacén de datos que parte de lo particular a lo general en la literatura en inglés lo encontramos como “*The Bottom Up Approach*”, de abajo hacia arriba; Kimball define al almacén de datos como una colección virtual de

“data marts” los cuales una vez unidos en un arreglo formarán un almacén de datos y por su parte William Inmon propone su desarrollo partiendo de lo general a lo particular, el ve al almacén de datos como un todo, primero se desarrolla el almacén de datos y a partir de ahí los *data mart* (“*The Top Down Approach*”).

A continuación en la figura 3 se pueden apreciar las principales características de las dos metodologías.

Características	Ralph Kimball	William Inmon
Visión General	Parte de lo particular a lo general.	Parte de lo General a lo particular
Filosofía	Parte de la prioridad que se le dan a los procesos del negocio. Se desarrollan data marts a partir de los procesos del negocio seleccionado utilizando los modelos dimensionales.	Parte del modelo de datos de toda la empresa. Desarrolla el Almacén de datos basándose en la base de datos empresarial; los data marts se desarrollan con datos obtenidos del almacén de datos.
Data marts	Se construye el data mart con los datos al menor nivel de detalle en base a cada proceso de negocio. Se construye a partir de los datos que se extraen de los sistemas transaccionales. Existen relación entre todos los data marts Los data mart guardan información histórica.	El data mart guarda todos los datos agregados que se relaciona a la unidad de negocio. El data mart es construido desde la extracción de datos del almacén de datos. Los data mart no tienen relación entre sí. Contienen un histórico de los datos limitado.

Figure 3 Diferencias en los modelos de W. Inmon y R. Kimball.

Debido a las características de la empresa donde se pretende implementar el almacén de datos contiene varias características que hacen elegir la metodología de Ralph Kimball es la más adecuada, partiendo del desarrollo de *data marts* harán del almacén más viable empezar con los procesos de negocio más relevantes y de ahí replicar la experiencia hacia los demás procesos hasta cubrir la totalidad del almacén de datos y llegar hasta obtener la rastreabilidad de todos y cada uno de los productos comercializados por todos los entes de la cadena de suministro.

Dado que la metodología propuesta por Kimball está basada en el modelo multidimensional; es necesario hacer una descripción de los conceptos utilizados por el mismo; si un almacén de datos está compuesto por la unión de varios data marts, mismos que son conformados desde la extracción de los datos desde los sistemas transaccionales y están organizados de tal manera que sea posible la consulta de información de forma rápida, para ello los datos son transformados y cargados al sistema multidimensional.

2.2.7 Mercado de datos (Data Mart)

Un mercado de datos es un conjunto de datos históricos que no participa en las operaciones diarias de la organización; se desarrollan en un área específica de la organización, detallando así a un proceso de negocio; visto desde el punto de vista de kimball, una data mart es un componente del almacén de datos.

2.2.8 Características de un Mercado de Datos

- Sirve como una fuente información para la inteligencia de negocios.
- No se diseñan al igual que una base de datos operacional.
- Se construyen para que tengan velocidad de acceso.
- Sigue siendo una base de datos relacionales, pero con menos uniones de tablas de datos.
- Es aceptable tener datos repetidos (sin normalizar) por el bien de la velocidad.
- Las reglas de normalización son sustituidas por un método diferente de diseño organizado en torno a los "hechos". Estos nuevos enfoques de diseño se llaman estrellas, copos de nieve y de constelación.

Se puede decir que un data mart es un almacén de datos con un alcance restringido a un proceso de negocio y se puede aplicar para su desarrollo las arquitecturas “Bottom Up” y “Top Down”.

2.2.9 Top-Down.

Primero se define el almacén de datos y luego se desarrollan, se construyen y cargan los mercados de datos a partir del mismo. En la figura 4 se puede apreciar los detalles de esta arquitectura.

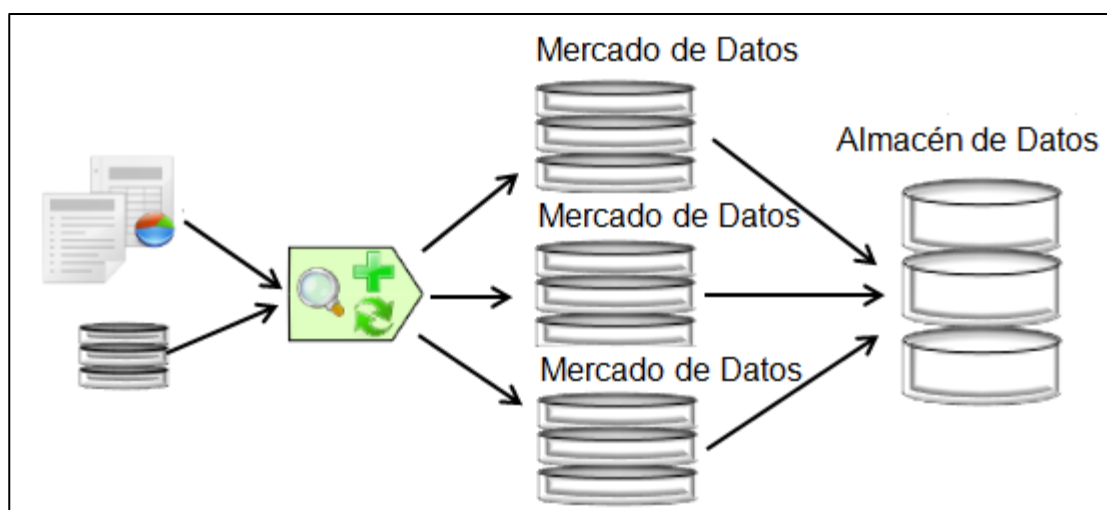


Figure 4 Arquitectura Top-Down

Aquí el almacén de datos se alimenta de los datos que son extraídos directamente de las bases de datos transaccionales a través de procesos especializados y desde aquí se alimentan los diferentes mercados de datos, recibiendo cada uno de ellos sus datos correspondientes dependiendo del proceso de que se trate; su ventaja radica en que no incurre en complicadas sincronizaciones de los hechos, y su desventaja es la inversión que se necesita para su construcción tanto en dinero como en tiempo.

2.2.10 Bottom-Up.

La característica de esta arquitectura, se definen previamente los mercados de datos y luego se integran en un almacén de datos centralizado. En la figura 5 se aprecia dicha arquitectura:

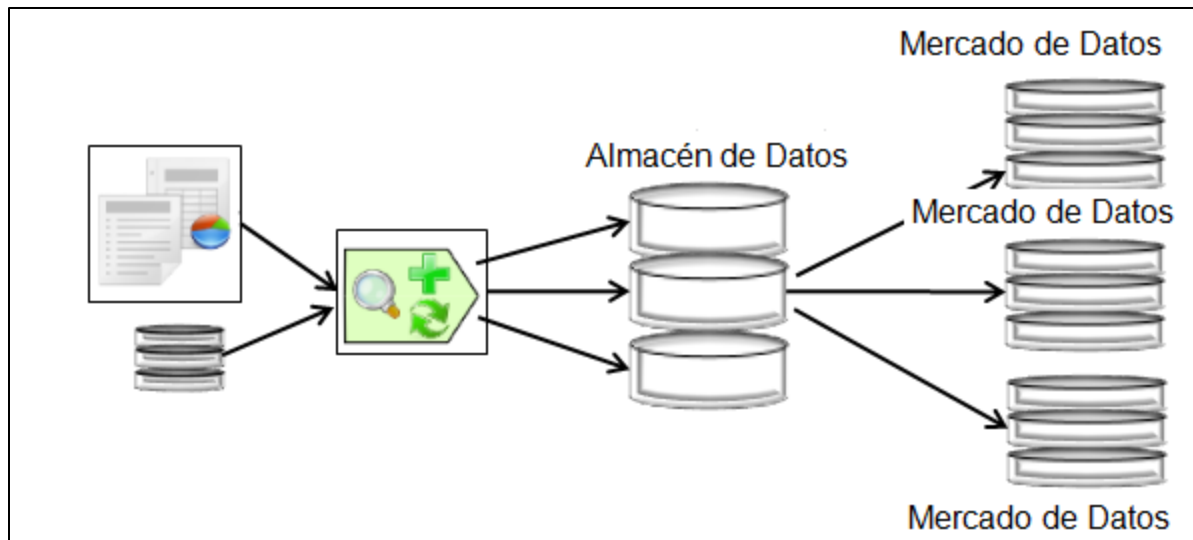


Figure 5 Arquitectura **Bottom-Up**

2.2.11 OLTP (**Online Transaction Proccesing**). -

Son las bases de datos operativas (también conocidas como bases de datos transaccionales); son diseñadas para el trabajo diario, procesan un gran volumen de transacciones y pueden entregar respuestas a las consultas de los usuarios, sus datos se actualizan en tiempo real.

2.2.12 OLAP (**Online Analitical Proccesing**).-

Es una solución muy utilizada en el campo de la Inteligencia de Negocios (*Business Intelligence*), consiste en consultas a estructuras multidimensionales (o **Cubos OLAP**) que contienen datos resumidos de grandes Bases de Datos o Sistemas Transaccionales (**OLTP**); Los sistemas OLAP presentan datos utilizando hechos, dimensiones, agregados, atributos y jerarquías.

Base de Datos Multidimensional o Cubo de Datos.- Su estructura está basada en tabla de hechos y tabla de dimensiones, está orientada a consultas complejas y alto rendimiento.

2.2.13 **Tabla de Hechos**.-

Contiene los datos capturados en un punto en el tiempo y constituyen la base de todo lo que se hace con la inteligencia de negocios. Están representados como una cantidad numérica; La información que representa esta cantidad se utiliza para

apoyar o evaluar la toma de decisiones y el desempeño de la organización; un hecho también es identificado como una medida.

2.2.14 Tabla de Dimensiones.-

Representa cada una de las posibles clasificaciones que se pueden aplicar a filas o columnas; definen como están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio. Representan los ejes del cubo de datos, y los aspectos de interés, mediante los cuales el usuario podrá filtrar y manipular la información almacenada en la tabla de hechos.

2.2.15 Cubo de Datos.-

El Cubo es una estructura que contiene un valor para una o más medidas para cada combinación única de los miembros de todas sus dimensiones. El cubo también puede contener valores agregados formados por las jerarquías de dimensión. En la figura 6 se ilustra un ejemplo de un cubo de datos y como es que se conforma, tomando en cuenta a las ventas totales del producto de *Espárragos para el distribuidor “Legumbres del norte” en el mes de Abril*. *Espárragos* es miembro de la dimensión de Productos. “Legumbres del norte” es miembro de la dimensión de Distribuidores. Abril es un miembro de la dimensión de Meses. Las ventas totales en esta intersección son \$19,350. (Larson, 2008)

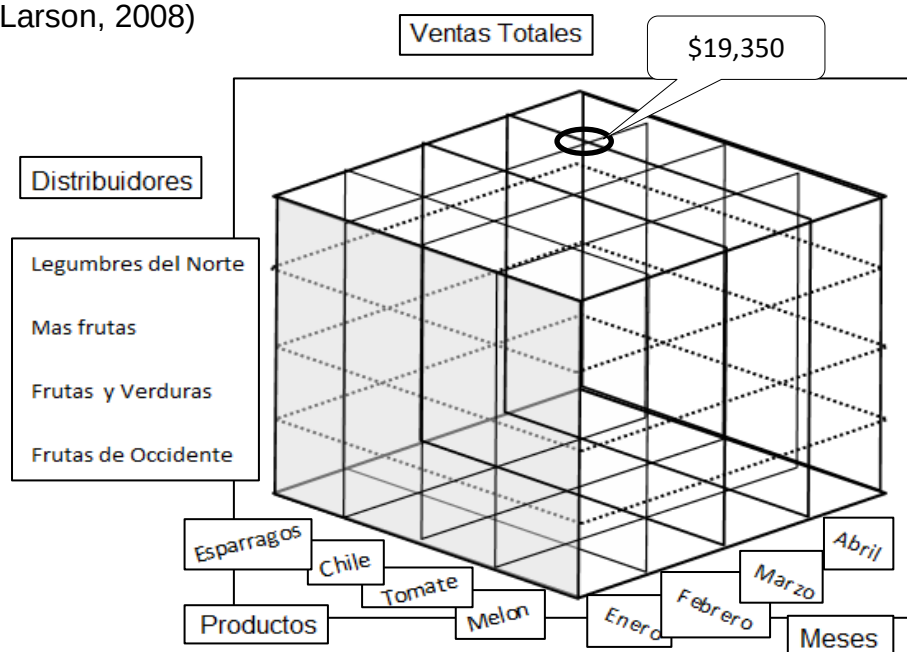


Figure 6 Ventas totales por Producto.

2.2.16 Agregado.-

Es un valor formado por la combinación de valores de una dimensión o un conjunto de medidas para crear un valor único. Esto suele hacerse mediante la suma de los valores, junto con el agregado suma, pero otros cálculos de agregación también se pueden utilizar; en la figura 7 se puede apreciar el ejemplo de un valor agregado; Para determinar el total de las ventas de abril para todos los distribuidores, se tienen que agregar todas las ventas totales de todos los distribuidores a través de todos los productos que se venden en abril. (Larson, 2008)

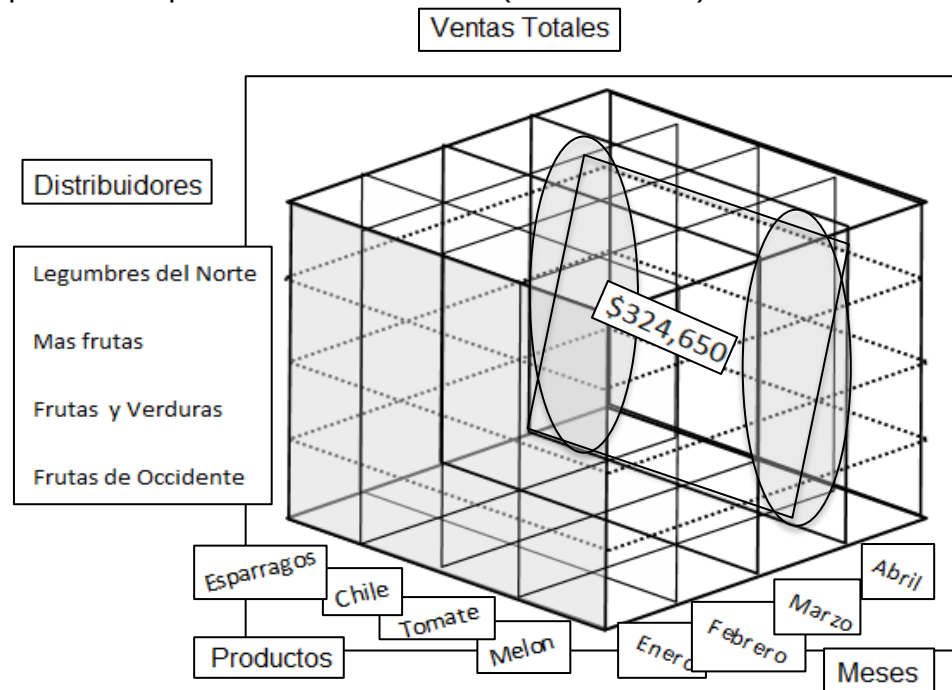


Figure 7 Ventas totales por Mes.

2.2.17 Atributos.-

Los atributos se utilizan para describir con más detalle a los miembros de una dimensión. Y pueden ser usados como criterio de análisis como parte de su producción de inteligencia de negocios.

2.2.18 Jerarquías.-

Una jerarquía es un atributo organizado de tal forma que le permite al usuario moverse en diferentes niveles de granularidad dentro de las medidas; lo cual le ayuda al usuario a analizar los datos desde un nivel general a detalle o desde un nivel detalle a un nivel general.

Revisando un poco más a detalle la conformación de un cubo de datos, en geometría un punto no tiene dimensiones, si se solicita el importe de ventas totales como tal, nos referimos solo a un punto, véase la figura 8 de Ventas Totales.



Figure 8 Ventas Totales.

Si se forma una línea en el tiempo, difundiendo las ventas totales por año (utilizando la dimensión de tiempo) y como en geometría una línea es un objeto unidimensional (véase la figura 9)

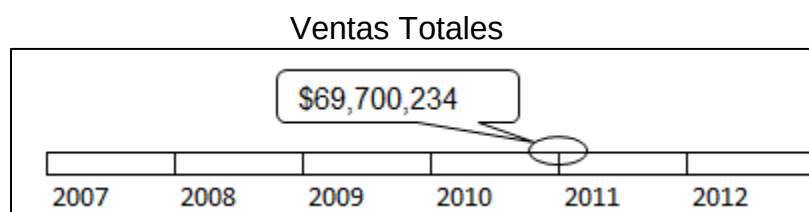


Figure 9 Ventas totales por año.

Ahora se tiene una línea formada por las medidas de las ventas totales en diversos puntos de la línea, un punto por cada año de la dimensión. En geometría, una línea es un objeto unidimensional. Por lo tanto, se añade una nueva dimensión a la medida: 2007, 2008, 2009, 2010, 2011 y 2012 cada uno dice que los miembros pertenecen a la dimensión de año.

Una vez más se extiende el total de ventas, esta vez para cada tipo de producto. Como se aprecia en la figura 10. (Larson, 2008)

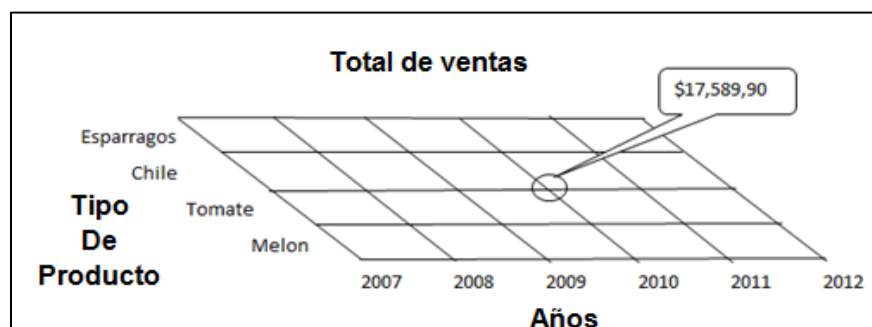


Figure 10 Ventas Totales por Producto por año.

Las medidas de las ventas totales están dispuestas en un cuadrado. Debido a que un cuadrado es un objeto de dos dimensiones, se ha añade otra dimensión al melón, *tomate*, *chile*, y *espárragos* de cada miembro de la dimensión del tipo de producto.

Se difunde el total de las ventas, una vez más, esta vez por la zona de venta. La medida de las ventas totales se ha convertido en un cubo, que es un objeto tridimensional. Se puede ver que cada vez que añade un nuevo criterio para extender la medida, se aumenta la dimensión de la medida, por lo tanto, el nombre de dimensión. (Véase la figura 11)

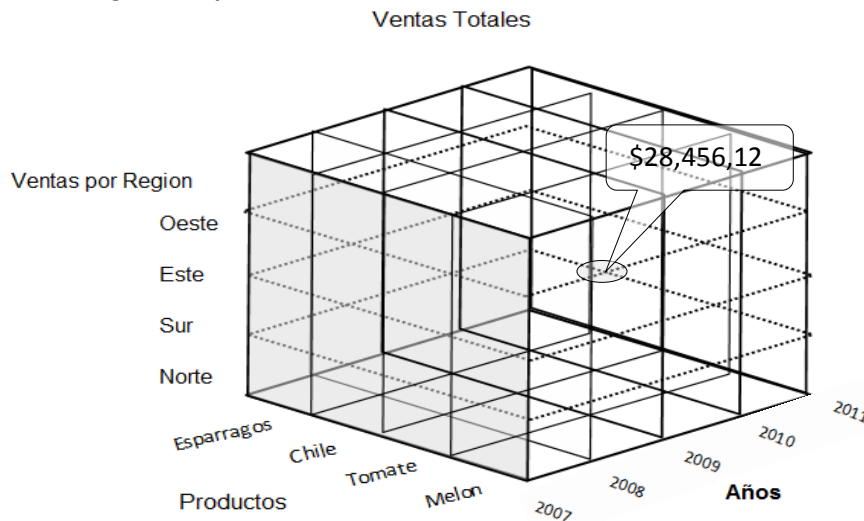


Figure 11 *Cubo de Ventas Totales (Regiones, Productos, Tiempo).*

Los almacenes de datos se pueden desarrollar utilizando las siguientes arquitecturas de diseño. (Ralph Kimball M. R., 2007)

2.2.19 Esquema Estrella.-

Es el esquema más simple para desarrollar un almacén de datos, consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a esta, a través de sus respectivas claves. En la figura 12 se puede apreciar un esquema en estrella estándar.

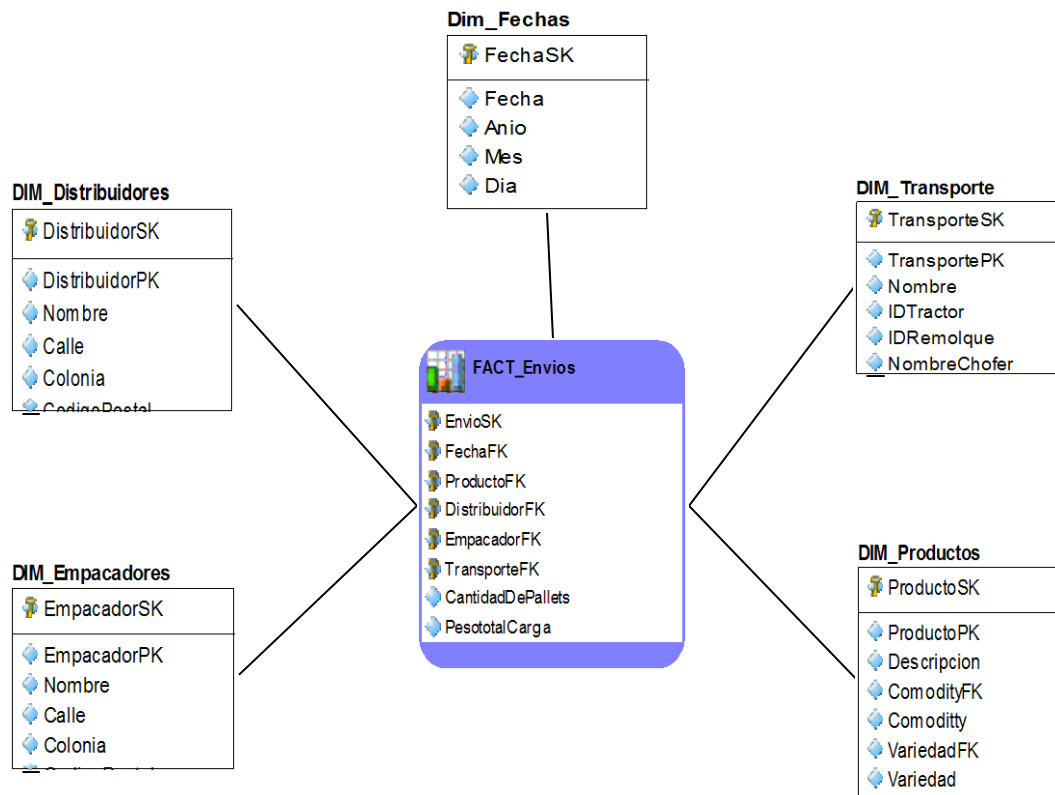


Figure 12 Esquema Estrella

El centro de la estrella está formado por la tabla de hechos. La tabla de hechos tiene una columna para las medidas y una columna para cada dimensión que contiene la clave externa de un miembro de esa dimensión. La clave principal de esta tabla se crea mediante la concatenación de todos los campos de clave externa. Esto se conoce como una clave compuesta; Las dimensiones se almacenan en tablas de dimensiones. La tabla de dimensiones tiene una columna para el identificador único de un miembro de la dimensión, por lo general un número entero o un valor corto tipo carácter. La tabla de dimensiones tiene otra columna para la descripción de los miembros. (Larson, 2008)

A continuación alguna característica de este modelo, que ayudan a comprender sus ventajas:

- Posee los mejores tiempos de respuesta.
- Su diseño es fácilmente modificable.
- Existe paralelismo entre su diseño y la forma en que los usuarios visualizan y manipulan los datos.

- Simplifica el análisis.
- Facilita la interacción con herramientas de consulta y análisis.

2.2.19 Esquema Copo de Nieve.-

Este esquema representa una extensión del modelo en estrella cuando las tablas de dimensiones se organizan en jerarquías de dimensiones; existe una tabla de hechos central que está relacionada con una o más tablas de dimensiones, quienes a su vez pueden estar relacionadas o no con una o más tablas de dimensiones. En la figura 13 se puede apreciar un ejemplo.

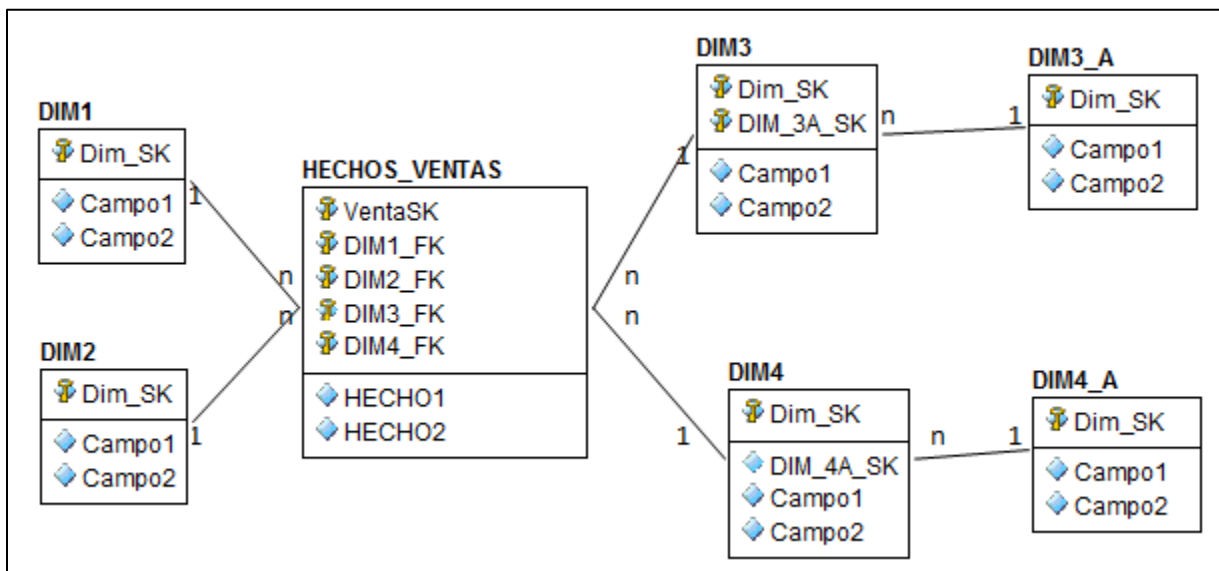


Figure 13 Esquema Copo de Nieve.

Se pueden definir las siguientes características del modelo copo de nieve:

- Posee mayor complejidad en su estructura.
- Hace una mejor utilización del espacio.
- Es muy útil en tablas de dimensiones de muchos registros.
- Las tablas de dimensiones están normalizadas, por lo que requiere menos esfuerzo de diseño.
- Puede desarrollar clases de jerarquías fuera de las tablas de dimensiones, lo cual permite realizar análisis de lo general a lo detallado y viceversa.

2.2.20 Esquema Constelación.-

Este modelo está compuesto por una serie de esquemas en estrella, y tal como se puede apreciar en la figura 14, está formado por una tabla de hechos principal (“HECHOS_VENTAS”) y por una o más tablas de hechos auxiliares (“HECHOS_DETALLE_VENTAS”), las cuales pueden ser sumalizaciones de la principal. Dichas tablas se ubican en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones; puede ser que las diferentes tablas de hechos compartan las mismas tablas de dimensiones, ya que, las tablas de hechos auxiliares pueden vincularse con solo algunas de las tablas de dimensiones asignadas a la tabla de hechos principal, y también pueden hacerlo con nuevas tablas de dimensiones.

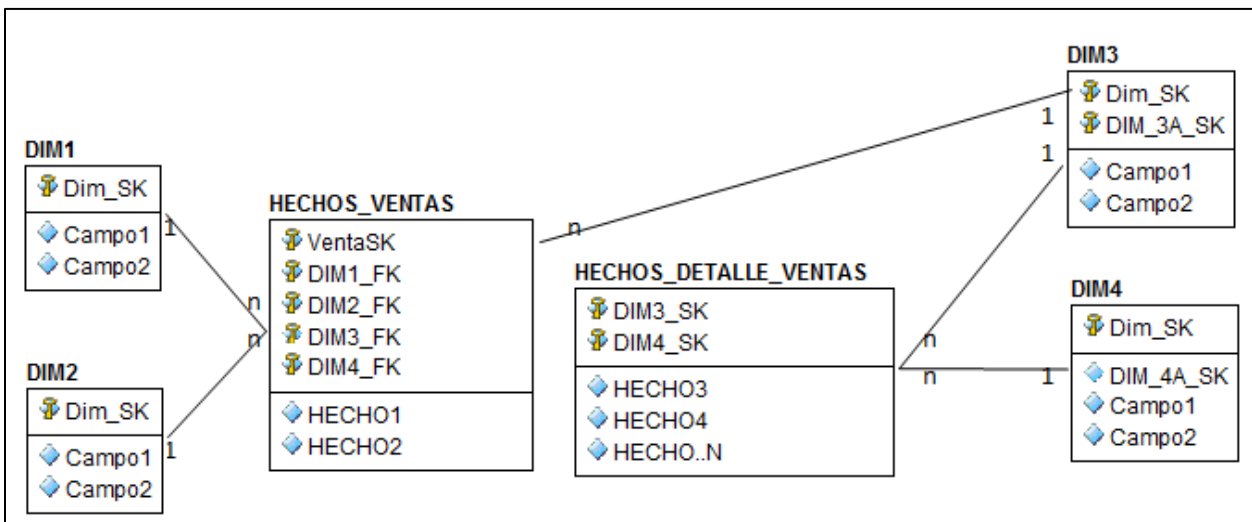


Figure 6 Esquema de Constelación.

Algunas de las características del esquema de constelación se mencionan a continuación.

- Permite tener más de una tabla de hechos, por lo cual se podrán analizar más aspectos claves del negocio con un mínimo esfuerzo adicional de diseño.
- Contribuye a la reutilización de las tablas de dimensiones, ya que una misma tabla de dimensión puede utilizarse para varias tablas de hechos.

2.2.21 Inteligencia de Negocios (Business Intelligence, BI).-

Se puede definir como el proceso de convertir información en conocimiento, mismo que ayuda a tomar decisiones.

2.3 Características de Almacén de datos.

Un almacén de datos se caracteriza por ser:

- **Integrado:** los datos almacenados en el almacén de datos deben integrarse en una estructura consistente. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- **Temático:** Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del almacén de datos. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
- **Histórico:** el tiempo es parte implícita de la información contenida en un almacén de datos. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el almacén de datos sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el almacén de datos se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- **No volátil:** el almacén de información de un almacén de datos existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del almacén de datos la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Otra característica del almacén de datos es que contiene metadatos, es decir, datos sobre los datos. Los metadatos permiten saber la procedencia de la información, su periodicidad de refresco, su fiabilidad, forma de cálculo... etc.; ellos serán los que permitan simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales y sus objetivos son:

2.4 Objetivos de un Almacén de datos.

- Dar soporte al usuario final, ayudándole a acceder al almacén de datos con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de *Business Intelligence* como DSS, EIS o CMI.
- Dar soporte a los responsables técnicos del almacén de datos en aspectos de auditoría, gestión de la información histórica, administración del almacén de datos, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos... etc.
- Se destaca que para comprender íntegramente el concepto de almacén de datos, es importante entender cuál es el proceso de construcción del mismo, denominado ETL (Extracción, Transformación y Carga), a partir de los sistemas operaciones de una compañía:
 - Extracción: obtención de información de las distintas fuentes tanto internas como externas.
 - Transformación: filtrado, limpieza, depuración, homogeneización y agrupación de la información.
 - Carga: organización y actualización de los datos y los metadatos en la base de datos.

El éxito de almacén de datos no está en su construcción, sino en usarlo para mejorar procesos empresariales, operaciones y decisiones. Posicionar un almacén de datos para que sea usado efectivamente. [Sperley, 1999]. Y la forma de valorarlo queda descrito en tres dimensiones [Inmmon & Hackathorn,1994]:

2.5 Forma de Valorar un almacén de datos.

1- Al Mejorar la Entrega de Información: información que está completa, que sea correcta, consistente, oportuna y accesible. Información que la gente necesita, en el tiempo que la necesita y en el formato que la necesita.

2- Facilitar el Proceso de Toma de Decisiones: con un mayor soporte de información se obtienen decisiones más rápidas; así también, la gente de negocios adquiere mayor confianza en sus propias decisiones y las del resto, y logra un mayor entendimiento de los impactos de sus decisiones.

3- Impacto Positivo sobre los Procesos Empresariales: cuando a la gente accede a una mejor calidad de información, la empresa puede mejorar:

- Al eliminar los retardos de los procesos empresariales que resultan de información incorrecta, inconsistente y/o no existente.
- Al Integrar y optimizar procesos empresariales a través del uso compartido e integrado de las fuentes de información.
- Al Eliminar la producción y el procesamiento de datos que no son usados ni necesarios, producto de aplicaciones mal diseñados o ya no utilizados.

En el entorno de este caso práctico, donde las autoridades de EEUU están exigiendo rastreabilidad a todos los productos que se comercializan en su territorio, se puede modelar un almacén de datos que ayude a la industria a obtener una rastreabilidad de todos los productos de una cadena de abastecimiento, para ello iniciamos declarando los principales conceptos que se estarán mencionando.

2.6 Rastreabilidad

El *EuroHandelsinstitut* (EHI), la *European Association of Fresh Produce Importers* (CIMO), el *Euro Retailer Produce Working Group* (EUREP), la *European Union of the Fruit and Vegetable Wholesale, Import and Export Trade* (EUCOFEL) y la *Southern Hemisphere Association of Fresh Fruit Exporters* (SHAFPE), reconocen la necesidad de adoptar un estándar común de identificación, comunicación y rastreabilidad.

Bajo la coordinación de "*EAN International*", han diseñado el "*Fresh Produce Traceability Project*" (Proyecto de Rastreabilidad del Producto Fresco - FPTP). El grupo a cargo de este proyecto desarrolló las "*Fresh Produce Traceability Guidelines*" (Normas de rastreabilidad para el Producto Fresco - Normas FPT), para permitir la identificación eficiente del origen de los defectos, así como la identificación y separación del producto defectuoso.

Las normas proporcionan a las partes de la cadena de abastecimiento de productos frescos, un sistema mundial de rastreabilidad para productos, sujeto a una nomenclatura de código de uso, comenzando con los dígitos 07 y 08. Ellas suministran una guía a los productores de productos frescos, empaques, proveedores logísticos, exportadores/importadores y distribuidores, así como también a sus clientes y proveedores, buscando introducir los estándares GS1, para implementar eficientemente un sistema de rastreabilidad, acordado internacionalmente.

2.7 Modelo de Rastreabilidad de la cadena de abastecimiento de productos frescos.

El objetivo de este modelo es explicar la rastreabilidad dentro de la cadena de abastecimiento de productos frescos, examinando los flujos físicos y de información, y determinando los estándares GS1 relevantes que van a ser desplegados; la rastreabilidad exige un método verificable para identificar productores, campos y productos, en todas sus configuraciones de embalaje y transporte/almacenamiento, en todas las etapas de la cadena de abastecimiento. Los números de identificación deben aplicarse y registrarse con exactitud, garantizando un vínculo entre ellos.

El rastreo y localización de alimento ha generado un interés y debate significativos. Es importante distinguir entre las exigencias legales, las tecnologías requeridas para proporcionar una capacidad de rastreo y localización, y los estándares GS1. El sistema GS1 permite el eficiente manejo de seguridad alimentaria, pero es responsabilidad de las empresas individuales y cadenas de abastecimiento, tomar ventaja, voluntariamente, de las capacidades que éste ofrece.

Dentro de la discusión de las capacidades de rastreabilidad suministradas por los estándares GS1, es importante aclarar la diferencia entre los términos rastreo y localización. Se menciona a continuación la definición acordada para cada término en el modelo FPT:

- El rastreo del Producto es la capacidad de seguir el camino de una unidad específica de un producto a través de la cadena de abastecimiento, mientras éste se mueve entre las organizaciones. Los productos son rastreados

rutinariamente en busca de su obsolescencia, manejo de inventario y propósitos logísticos. Dentro del contexto de las normas FPT, los intereses comunes se enfocan en el rastreo del producto, desde el productor hasta el minorista, en unidades logísticas no modificadas.

- La localización del producto es la capacidad de identificar el origen de una unidad particular y/o de un lote de productos ubicados dentro de la cadena de abastecimiento, de acuerdo a referencia de registros que tuvieron lugar contracorrente en dicha cadena. Los productos se localizan con propósitos tales como la anulación de los mismos o investigación de reclamos. Dentro del contexto del FPT, el interés común se centra en la localización del producto en unidades comerciales no modificadas, desde el minorista al productor.

Enlaces entre sucesivas configuraciones de unidad comercial y de unidad logística.

Los números de identificación deben aplicarse y registrarse con exactitud, garantizando un enlace entre las sucesivas configuraciones de empaque y transporte/almacenamiento. Es responsabilidad de cada empresa, manejar los enlaces entre lo que los proveedores les envían, los procesos de alteración del producto y lo que ellos embarcan a sus clientes.

2.8 Mantenimiento de un registro exacto y a tiempo.

Algunos datos deben transmitirse sistemáticamente entre las partes de la cadena de abastecimiento, mientras otros datos deben ser registrados únicamente. Depende de las partes comercializadoras decidir qué datos transmitir sistemáticamente.

2.9 Composición del lote.

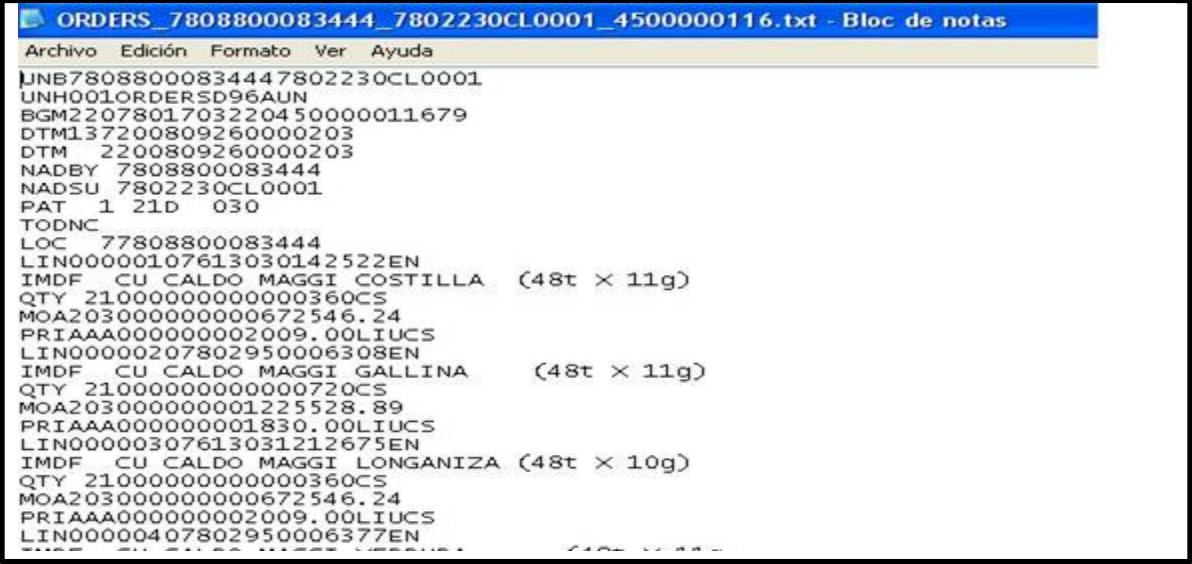
La eficiencia de cualquier sistema de rastreabilidad depende del vínculo más débil dentro de la cadena de abastecimiento. La composición del lote es un punto crítico en este proceso. Determina la exactitud de todo sistema de rastreabilidad. Cuanto más homogéneos son los lotes, más exacto será el sistema de rastreabilidad.

2.10 Comunicación electrónica de datos de Rastreabilidad.

Los datos de rastreabilidad pueden transmitirse a través de medios electrónicos, tales como los mensajes *EANCOM EDI*, y relacionarse al número de identificación de las

unidades logísticas - el "Serial Shipping Container Code" (Código Seriado del Contenedor de Embarque - SSCC).

En la figura 15, se puede apreciar cómo puede estar desarrollado un archivo EANCOM EDI.



```
ORDERS_7808800083444_7802230CL0001_4500000116.txt - Bloc de notas
Archivo Edición Formato Ver Ayuda
UNB78088000834447802230CL0001
UNH001ORDERSD96AUN
BGM2207801703220450000011679
DTM137200809260000203
DTM 2200809260000203
NADBY 7808800083444
NADSU 7802230CL0001
PAT 1 21D 030
TODNC
LOC 77808800083444
LIN00000107613030142522EN
IMDF CU CALDO MAGGI COSTILLA (48t X 11g)
QTY 21000000000000360CS
MOA2030000000000672546.24
PRIAAA0000000002009.00LIUCS
LIN00000207802950006308EN
IMDF CU CALDO MAGGI GALLINA (48t X 11g)
QTY 21000000000000720CS
MOA2030000000001225528.89
PRIAAA0000000001830.00LIUCS
LIN00000307613031212675EN
IMDF CU CALDO MAGGI LONGANIZA (48t X 10g)
QTY 21000000000000360CS
MOA2030000000000672546.24
PRIAAA0000000002009.00LIUCS
LIN00000407802950006377EN
```

Figure 15 Ejemplo de un archivo EDI.

2.11 Modelo de rastreabilidad

El modelo representa los: Flujos físicos (flecha) dentro de la cadena de abastecimiento. Solo se toman en cuenta los pasos donde tiene lugar una transformación; como se muestra en la figura 16.

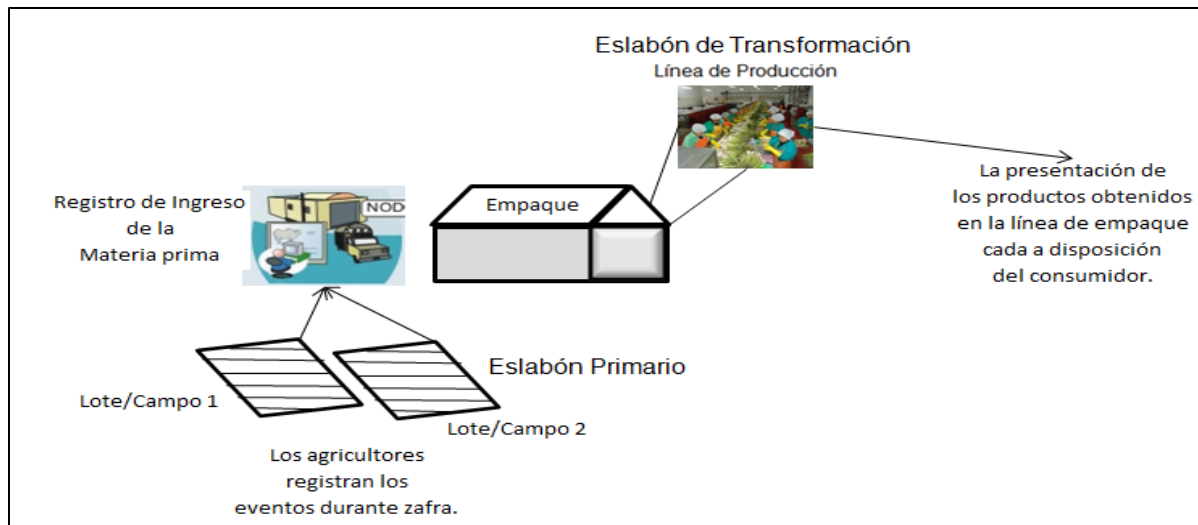


Figure 167 Modelo de rastreabilidad de productos frescos.

2.12 Números Mundiales de Artículo Comercial EAN.UCC

En el más simple nivel, la numeración del artículo es lo que el nombre sugiere - un sistema para identificar artículos, dándole a cada uno un único número. La numeración puede aplicarse en cualquier etapa de la producción y distribución. Se usa para identificar productos y servicios. Mientras el aspecto más visible de la numeración del artículo es el código de barra, ésta es solo una representación de un número, leído por una máquina. El número es el elemento más importante en el sistema EAN.UCC, porque identifica el artículo al cual ha sido asignado.

En la figura 17 se puede apreciar un claro ejemplo de ello en forma de código de barra

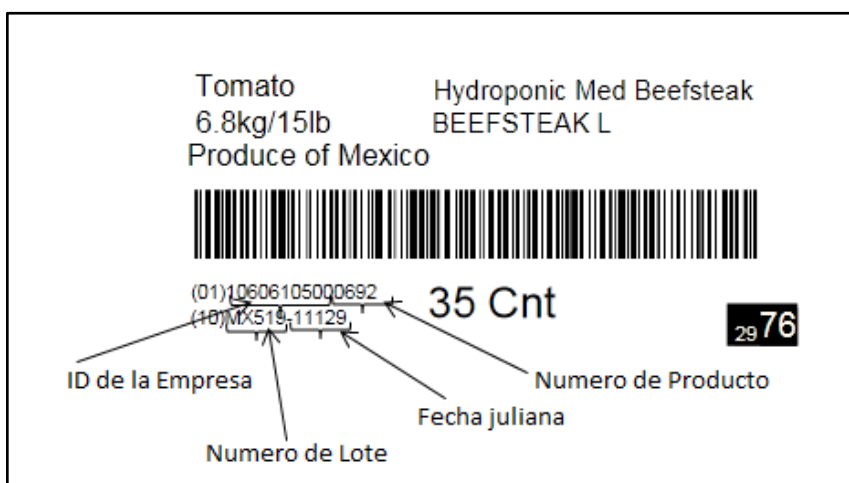


Figure 178 Número mundial de artículo comercial.

El sistema de numeración GS1 proporciona una unidad mundial y supera problemas de confusión, duplicaciones y mala interpretación, porque todos los usuarios del sistema GS1 siguen las mismas reglas de código. Un número GS1 puede reconocerse, no solo por las empresas comerciales locales, sino también, por otras operando en el exterior. Cada número GS1 es único a lo ancho del mundo, por lo tanto no hay posibilidad de confusión. El sistema de numeración GS1 también proporciona a los artículos la habilidad para transportar, dentro de la convención numérica, información extra o de atributo de los mismos.

2.13 Códigos de barra GS1

Los códigos de barra son transportadores de datos. El código de barra se usa en el sistema GS1 para representar los números EAN/UCC. En términos simples, un código de barra consiste en una serie de barras y espacios, paralelos y adyacentes. Se usan modelos de ancho predeterminado, para representar datos actuales en el código de barra. Estos datos pueden ser el número de artículo o información de atributo, relacionados al artículo. Un lector de código de barra (*scanner*), se mueve a través del código de barra, de un lado al otro. Al hacer esto, un equipo de lectura analiza el modelo de ancho de las barras y espacios, y se recuperan los datos originales de tiempo real, en forma exacta y rápida

Los códigos de barra GS1 permiten la captura automática de datos, lo que significa una solución comercial clave, dentro de una eficiente cadena de abastecimiento. La numeración GS1 y el sistema de codificación de barra, permiten la entrada rápida y exacta de los datos, a los sistemas de computadora, automatizando el flujo de información dentro de los procesos comerciales. También permiten la captura mejorada de datos y la transferencia de información, al tiempo que se reducen los costos.

2.14 Números Mundiales de localización

Un Número Mundial de Localización (GLN) es un código numérico que identifica cualquier entidad legal, funcional o física, dentro de un comercio u organización. A cada localización se le asigna un número. La estructura de numeración EAN.UCC-13 se usa para la identificación de localizaciones.

2.15 Definiciones de las unidades del consumidor, comerciales y logísticas

Para su utilidad, encontrará a continuación las definiciones de unidades del consumidor, comerciales y logísticas. Ellas permiten una comprensión en común, de los términos utilizados en la próxima sección, las normas de rastreabilidad para Productos Frescos:

2.16 Definición de una unidad del consumidor.

Un artículo minorista, o más comúnmente conocido como unidad del consumidor, es todo artículo para la venta al consumidor final, a través de un punto de venta minorista. Cualquier artículo que pueda considerarse artículo minorista (unidad del consumidor), y artículo no-minorista (unidad comercial), se numera y codifica con barras, de acuerdo a las reglas aplicables a los artículos minoristas. La norma FPT no apunta a la identificación, codificación y rastreabilidad de artículos minoristas (pre-embalados o sueltos). En la figura 18 se pueden apreciar diferentes frutas definidas como una unidad del consumidor.

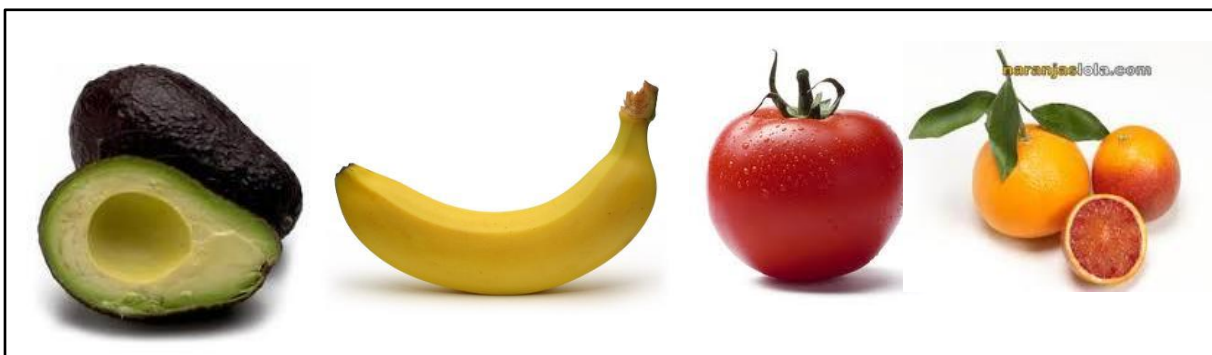


Figure 18 Unidades del consumidor.

2.17 Definición de unidad comercial.

Un artículo no-minorista o más comúnmente conocido como unidad comercial, y a partir de ahora mencionado en este documento como una unidad comercial, es todo artículo o agrupación estándar de artículos, establecidos para facilitar las operaciones de manipuleo, almacenamiento, preparación del pedido, embarque, etc. Una unidad comercial puede ser un cajón, caja, bin, pallet o cualquier tipo similar de embalaje, creado con el propósito de ser embarcado o manipulado. En la figura 19 se aprecia un ejemplo de lo que viene siendo una unidad comercial.



Figure 199 Unidad Comercial.

2.18 Definición de unidades logísticas.

Una unidad logística es un artículo de cualquier composición, establecido para transportar y/o almacenar, que necesita ser manejado a través de la cadena de abastecimiento. Las unidades logísticas, que requieren ser rastreadas y localizadas individualmente, a través de cualquier cadena de abastecimiento, necesitan un número único de identificación; como se puede apreciar en la figura 20.



Figure 20 Unidad Logística.

2.19 Definición de Información de Atributo.

La información de atributo es toda información variable requerida sobre la unidad comercial o la identificación de la unidad logística, tal como el número de lote. En el sistema GS1, esta información se expresa por medio de los "*EAN.UCC Application Identifiers*" (Identificadores de Aplicación EAN.UCC - AI). La información de atributo está codificada por barras en la simbología de código de barra UCC/EAN-128.

2.20 Alertas de Inocuidad en los Estados Unidos de América.

Para saber cómo se trabajan en alertas de inocuidad alimentaria, se remite a una organización en los Estados Unidos de América, quien es la que toma la iniciativa de la investigación.

La forma como se opera al momento de surgir una alerta fitosanitaria, la determina **la CDC**; La **CDC** es parte del Servicio de Salud Pública de los Estados Unidos, con la misión de utilizar la mejor información científica disponible para vigilar, investigar, controlar y prevenir los problemas de salud pública. Utilizando herramientas de epidemiología y las ciencias del laboratorio, la CDC proporciona una evaluación científica a las amenazas para la salud pública. La CDC trabaja en estrecha asociación con los departamentos de salud pública para determinar la frecuencia de determinadas enfermedades y lleva a cabo vigilancia nacional de las mismas. La CDC proporciona una consulta experta, epidemiológica y microbiológica a los departamentos de salud y otras entidades federales sobre toda una serie de cuestiones de salud pública, entre ellas la enfermedad transmitida por los alimentos, y estaciona a epidemiólogos en departamentos de salud de los estados para ayudar con la vigilancia y la investigación de muchos problemas. La CDC, en colaboración con los funcionarios de salud pública del estado, también puede enviar un equipo al terreno para realizar investigaciones de campo de emergencia en brotes grandes o inusitados. Los investigadores de la CDC formulan nuevos métodos para identificar, caracterizar y determinar la huella de los microbios que ocasionan la enfermedad. Traducen la investigación de laboratorio a métodos prácticos de campo que puedan utilizar las autoridades de salud pública en los estados y los condados.

La CDC no es una entidad reglamentaria. La reglamentación gubernamental de la seguridad alimentaria la llevan a cabo la Administración de Alimentos y Fármacos (FDA), el Departamento de Agricultura de los Estados Unidos (USDA), el Servicio Nacional de Pesquerías Marinas y otra entidades reglamentarias. La CDC mantiene contacto regular con estas entidades.

Cuando aparecen nuevas amenazas de salud pública, la CDC determina cuáles son y cómo pueden controlarse mediante una rápida investigación de laboratorio y de campo de carácter científico. La CDC comparte los resultados de estas investigaciones con los estados, con las entidades federales reglamentarias y con las propias industrias. Aunque no reglamentan la seguridad de los alimentos, la CDC evalúa la eficacia de las iniciativas actuales de prevención. Proporcionando una evaluación científica de cuáles son los problemas, cómo pueden controlarse y dónde se hallan las lagunas de conocimiento.

Propuesta de Solución

Herramientas de Rastreabilidad GS1

La automatización aumenta significativamente la productividad y reduce los gastos asociados al papeleo y la administración. También elimina los errores inevitables que surgen del ingreso y procesamiento manual de datos. La comunicación electrónica permite a las empresas mejorar su manejo y controlar el ciclo comercial, lo que redundará en un manejo logístico perfeccionado. La compresión de tiempo incrementa la sensibilidad al servicio al cliente, y ayuda a restablecer la confianza del cliente en caso que se retire un producto.

La introducción de los estándares GS1 puede mejorar la eficiencia en registrar e intercambiar información entre los integrantes de la cadena de abastecimiento. Cuando se los usa juntamente con bases de datos que contienen registros exactos y puntuales, los estándares GS1 proporcionan a los participantes de la cadena de abastecimiento la capacidad técnica para ver el origen de un producto, en sus propias localizaciones y a través de toda la cadena de abastecimiento.

Los estándares GS1 transportan datos, los que permiten a los integrantes de la cadena, rastrear y localizar productos. La aplicación de estos estándares requiere que los productores, empacadores, importadores/exportadores, transportadores, distribuidores y minoristas, guarden registros de los números seriados de las unidades logísticas (SSCC), de los números de identificación (GTIN) y asignen información sobre las unidades comerciales y los números de localización de su origen (GLN). El hecho

de guardar registro, les permite a los productores y empacadores, suministrar los datos de rastreabilidad necesarios a los importadores/exportadores y distribuidores, como así también a sus clientes. La habilidad para convertir esta capacidad en beneficios prácticos, requiere acuerdos bilaterales, para compartir la información de inventario correspondiente.

Para ello se hará un modelado de un almacén de datos donde se reúnan todos los datos de los entes participantes en la cadena de abastecimiento de los productos del campo; para este caso práctico se modeló el almacén de datos en base a los datos que proporcionan el empacador y distribuidor en la cadena de abasto.

Capítulo III

3 Metodología

La metodología para el desarrollo y llevar a buen término el proyecto, la que propone Ralph Kimball en su ciclo de vida almacenes de datos; es la que se sigue en este caso, ya que se enfoca a diseñar la base de datos dimensional, la cual resguardara los datos que llegan a generar información y a su vez con la aplicación de herramientas de inteligencia de negocios generan conocimiento para la toma de decisiones, dicho modelo está basado en tablas de hechos y en tablas de dimensiones. En la figura 23 se puede ver el diagrama del ciclo de vida de Ralph Kimball.

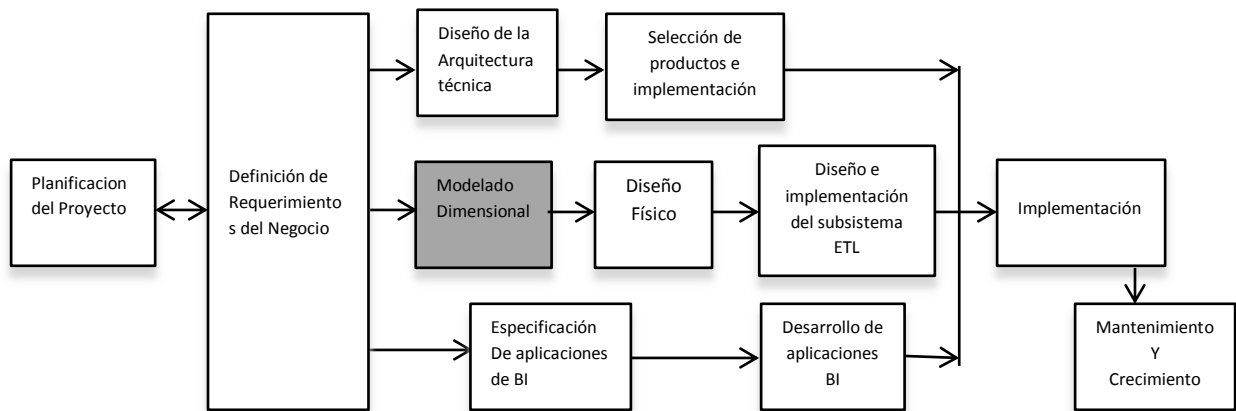


Figure 21 Diagrama del ciclo de vida de Kimball.

3.1 Planeación y administración del proyecto.

La parte que detonó la chispa para iniciar este proyecto, son las regulaciones por parte del gobierno de los Estados Unidos de América, el cual necesita que esté bien definida e identificable la cadena de suministro, la iniciativa por parte de los particulares por realizar un sistema de rastreabilidad que cumpla con los requerimientos solicitados ha hecho la puesta en marcha del desarrollo del proyecto de rastreabilidad.

El Identificar los principales usos a los datos generados por los sistemas transaccionales; a través de entrevistas abiertas con las personas que toman las decisiones más importantes. Se encontró la forma de unir los datos en un solo repositorio para sacar la mayor ventaja de ellos mediante un modelo de almacén de datos; para ello fue necesario lo siguiente:

- Definir los requerimientos del Negocio
- Identificar los requerimientos de Información
- Definir los requerimientos técnicos
- Realizar el diseño multidimensional

3.2 Definir Requerimientos de Información.

Durante las pláticas realizadas con la persona encargada de ser el enlace entre la empresa y en este caso el aspirante de grado; se hizo saber lo siguiente:

Que existe la necesidad de tener reportes confiables de la producción por lotes, campos, por fecha, de líneas de producción en todos los productos; ¿en qué carga (Manifiesto de envío) se enviaron?, ¿Quién lo transportó?, ¿Cuándo?, ¿A quién se le envió?, ¿Cuántas tarimas?, ¿Cuántas cajas se le han enviado en una fecha determinada?, en un rango de fechas?

Saber identificar que lote, que campo han sido los más productivos en determinado producto, así mismo que líneas de empaque, que operadores son los más rentables en cuanto a calidad y cantidad en los productos. Solo por mencionar algunos de los requerimientos de información que se necesitan obtener.

3.2.1 Por parte del Distribuidor

Identificar que Empaque (Proveedor) le envía más y de mejor calidad de productos.

Cantidad de envíos hechos por empacador, cajas, tarimas han recibido por una fecha determinada, rango de fechas, etc.

A su vez existe la necesidad de saber a qué cliente se le vende que producto, cuantas cajas, cuantas tarimas, quien es el proveedor de dichos productos, de que lote y que campo provienen; todo ello con el fin de aplicarle rastreabilidad a los productos frescos.

Para que en el momento en que pueda existir una contingencia sanitaria en alguno de los productos se pueda identificar a que clientes se les ha enviado, de que lote de producción y de que campo proviene dicho producto; y así poder comunicar a los clientes a quienes se les pudiera haber vendido un producto contaminado y no crear un pánico innecesario en la población.

Con ello se obtienen grandes ahorros tanto en tiempo y dinero a todos los integrantes de la cadena de suministro así como a la entidad gubernamental encargada de sancionar a los que pudiesen haber iniciado dicho problema.

3.2.2 Empacador

Con el sistema actual; con los datos contenidos en la base de datos se pueden contestar muchas preguntas hasta cierto punto; pero una vez que se haya hecho el modelo de almacén de datos se podrá tener una visión universal de los datos que se generan.

- Podrá tener, saber, conocer la cantidad de cualquier producto generado por un lote, campo en específico en un determinado momento.
- Podrá identificar qué línea de producción fue la que procesó más cantidad y de mejor calidad un producto en específico, de que lote y que campo era, todo ello aplicado en una fecha determinada o rango de fecha.
- Podrá identificar en que envío (Manifiesto) se fue, hasta que distribuidor se fue, quien lo transporta y quien fue su cliente final.

3.2.3 Por Lote y Campo.

- Se podrá saber con exactitud de cuanto ha sido el total de su producción, por fecha en especial, por temporada, etc.
- De qué calidad han sido sus productos
- La cantidad de productos empacados

Solo por mencionar algunos.

3.2.4 Por Línea de Producción.

- Cantidad de productos empacados por fecha
- Calidad de los productos empacados
- Operador con mayor producción

3.2.5 Por Manifiesto.

- Cantidad de envíos realizados por distribuidor, por fechas, etc.
- Cantidad pallets y cajas que han sido enviados.

3.2.6 Por Transporte.

- Cantidad Envíos realizados
- Cantidad de Pallets y cajas transportados

3.2.7 Por distribuidor.

- Cantidad de Cargas recibidas, filtradas tanto por empacador, por producto, por fecha, etc.
- Cantidad de Pallets recibidos, cantidad de cajas recibidas por producto, etc.
- Cantidad de pallets vendidos a sus clientes.
- De que producto, de que lote y campo provienen, etc.

3.3 Diseño de Arquitectura

Equipo de trabajo de planeación e implementación de Arquitectura

El equipo de trabajo se conforma solo por el aspirante de grado, para fines de diseñar la arquitectura completa del almacén de datos es necesario involucrar a personas que juegan roles diferentes dentro del proceso de trazabilidad.

Requerimientos técnicos basados en las necesidades de los usuarios.

- 1.- Trazabilidad a través de un portal Web
- 2.- Herramientas de análisis de información
- 3.- Información histórica

3.4 Trazabilidad a través de un portal WEB

El portal web podría contemplar la opción de rastrear un producto desde su cualquier lugar; para fines de esta tesis (caso práctico) solo se propone llevar a cabo el portal, el objetivo es solo el modelo del almacén de datos y como obtener un óptimo servicio del mismo a través de un portal web.

3.5 Información Histórica

Es muy importante saber cómo ha sido la producción a través del tiempo en los diversos lotes, campos; así como los envíos realizados a cada distribuidor, desde la perspectiva del emparador-productor, de igual forma desde la perspectiva del distribuidor los productos recibidos por emparador-productor tanto en total de envíos recibidos, cantidad de tarimas, cantidad de cajas, filtradas por producto, por fecha, por empresa de transporte, solo por mencionar algunas necesidades.

3.6 Equipo e Infraestructura

Toda vez que se haya identificado los requerimientos técnicos, ahora le toca el turno al hardware y software necesario para llevar a cabo el proyecto; solo se realizará un análisis del mismo más no se contempla la adquisición de él; solo para fines de propuesta se hará un análisis de manera general.

- 1.- Servidor Para procesos de Extracción – Transformación – Carga (ETL por sus siglas ingles)
- 2.- Servidor para el manejador de base de datos
- 3.- Renta de un Servidor WEB

Que contenga: Windows Server, MS SQL Server Enterprise Edition

- 4.- Servidor para las herramientas de presentación

5.- Dispositivo de Almacenamiento masivo en red (SAN por sus siglas en inglés) para los respaldos de datos.

3.7 Selección Del Software.

Con la evolución del software libre, existen herramientas para aplicar Inteligencia de negocios de una forma eficiente, eficaz y con una curva de aprendizaje lo suficientemente aceptable como para su elección, por tal motivo se eligió Pentaho y toda vez que tiene las herramientas necesarias para llevar a cabo una implementación exitosa.

3.7.1 Pentaho BI Platform.

La plataforma de Pentaho BI provee de servicios críticos incluyendo programación, seguridad, integración automatización y flujo de trabajo. Proporcionando habilidades a los usuarios finales de Pentaho y proveyendo un lugar central para administrar y mantener el despliegue de la empresa BI.

- Integración con procesos de negocio
- Administra y programa reportes
- Administra seguridad de usuarios

3.7.1.1 Componentes.

- Plataforma 100% J2EE, asegurando la escalabilidad, integración y portabilidad.
- Servidor: puede correr en servidores compatibles con J2EE como JBOSS AS, *WebSphere*, *Tomcat*, *WebLogic* y Oracle AS.
- Base de datos: vía JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird.
- Sistema operativo: no hay dependencia. Lenguaje interpretado.
- Lenguaje de programación: Java, Javascript, JSP, XSL (XSLT/XPath/XSL-FO).
- Interfaz de desarrollo: Java SWT, Eclipse, Web-based.
- Repositorio de datos basado en XML.
- Todos los componentes están expuestos vía Web Services para facilitar la integración con Arquitecturas Orientadas a Servicios (SOA).

3.7.2 Arquitectura.

El siguiente esquema nos muestra la arquitectura estructurada de las diferentes componentes que forman parte de pentaho:

La solución Business Intelligence OpenSource Pentaho pretende ser una alternativa a las soluciones propietarias tradicionales más completas: Business Objects, Cognos, Microstrategy, Microsoft, etc... por lo que incluye todos aquellos componentes que se pueden encontrar en las soluciones BI propietarias más avanzadas:

- Reporting
- Análisis
- Dashboards
- Workflow
- Data Mining
- ETL
- Single Sign-On. Ldap
- Auditoría de uso y rendimiento
- Planificador
- Notificador
- Seguridad. Perfiles

El proyecto Pentaho BI abarca las siguientes aéreas de reporte:

3.7.3 Pentaho Reporting.

Es una solución proporcionada por pentaho e integrada en su suite para el desarrollo de informes.



Existen tres productos con diferentes enfoques y dirigidos a diferentes tipos de usuarios.

1. Pentaho Report Designer: Editor basado en eclipse con prestaciones profesionales y de calidad y con capacidad de personalización de informes a las necesidades de negocio destinado a desarrolladores.

Incluye Asistentes para facilitar la configuración de propiedades.

Está estructurado de forma que los desarrolladores pueden acceder a sus prestaciones de forma rápida:

Incluye un editor de consultas para facilitar la confección de los datos que serán utilizados en un informe.

2. Pentaho Report Design Wizard: Herramienta de diseño de informes, que facilita el trabajo y permite a los usuarios obtener resultados de forma inmediata. Está destinada a usuarios con menos conocimientos técnicos.

A través de pasos sencillos permite:

- Conectarse a todo tipo de bases relacionales
- Integrar el resultado dentro del portal pentaho
- Posibilidad de montar codificación semafórica

3. Web ad-hoc Reporting: Es el similar a la herramienta anterior pero vía web.

Extiende la capacidad de los usuarios finales para la creación de informes a partir de plantillas pre configuradas y siguiendo un asistente de creación.

Las características generales son:

- Proporciona funcionalidad crítica para usuarios finales como:
 - Acceso vía web
 - Informes parametizados
 - *Scheduling*
 - Suscripciones
 - Distribución (*bursting*)
- Proporciona claras ventajas a especialistas en informes:
 - Acceso a fuentes de datos heterogéneos: relacional (vía jdbc), OLAP, XML, transformaciones de pentaho data integration.
 - Capacidad de integración en aplicaciones o portales: jsp, portlet, web service.

- Definición modular de informes (distinción entre presentación y consulta)
- Diseño de informes flexible
 - Entorno de diseño gráfico
 - Capacidad de uso de templates
 - Acceso a datos relacionales, OLAP y XML
- Desarrollado para:
 - Ser embebible
 - Ser fácil de extender
 - No consumir muchos recursos
 - 100% Java: portabilidad, escalabilidad e integración
- Multiplataforma (tanto a nivel de cliente como servidor): Mac, linux/Unix y Windows

3.7.4 Pentaho Analysis.

Ayuda a operar con máxima efectividad para ganar intuición y entender lo necesario para tomar optimas decisiones.

Las características generales son:

- Vista dimensional de datos (por ventas, por periodo)
- Navegar y explorar
 - Análisis Ad Hoc
 - Drill-down
 - Seleccionar un especifico miembro para el análisis
- Interactuar con alto rendimiento
 - Tecnología optimizada para rápida respuesta interactiva

3.7.5 Pentaho Dashboards.

Provee inmediata perspicacia en un rendimiento individual, departamental o empresarial. Para elegir métricas claves en una atractiva e intuitiva interfaz visual,

Pentaho Dashboards es para los usuarios que necesitan información crítica y que quieran entender y mejorar el rendimiento organizacional.

- Identificación de unas Métricas Clave (KPI's, Key Performance Indicators)

- Monitoreo/Métricas
- Investiga detalles subyacentes
 - Drill a reportes de soporte
- Seguimiento de excepciones
 - Alertas basadas en reglas del negocio

3.7.6 Pentaho Data Integration.

Muchas organizaciones tienen información disponible en aplicaciones y base de datos separados. Pentaho Data Integration abre, limpia e integra esta valiosa información y la pone en manos del usuario. Provee una consistencia, una sola versión de todos los recursos de información, que es uno de los más grandes desafíos para las organizaciones TI hoy en día. Pentaho Data Integration permite una poderosa ETL (Extracción, Transformación y Carga).

El uso de *kettle* permite evitar grandes cargas de trabajo manual frecuentemente difícil de mantener y de desplegar.



3.7.7 Propiedades básicas.

A parte de ser open source y sin costes de licencia, las características básicas de esta herramienta son:

- Entorno gráfico de desarrollo
- Uso de tecnologías estándar: Java, XML, JavaScript
- Fácil de instalar y configurar
- Multiplataforma: windows, macintosh, linux
- Basado en dos tipos de objetos: Transformaciones (colección de pasos en un proceso ETL) y trabajos (colección de transformaciones)
- Incluye cuatro herramientas:
 - **Spoon:** para diseñar transformaciones ETL usando el entorno gráfico
 - **PAN:** para ejecutar transformaciones diseñadas con spoon

- **CHEF:** para crear trabajos
- **Kitchen:** para ejecutar trabajos

3.7.8 Data Mining.

Es el proceso de correr datos en algoritmos completamente sofisticados, relevando significantes patrones y correlaciones que pueden estar escondidos. Esto puede ser usado para ayudar a entender lo mejor para el negocio y explotar el rendimiento de este en un futuro prediciendo completamente en el análisis.

Se caracteriza por:

- Descubrir patrones ocultos y correlacionales en los datos
- Prevenir eventos futuros basados en patrones históricos
- Contar con la tecnología de:
 - Poderoso motor de Data Mining
 - Herramientas de Diseño Grafico
 - Seguridad y conformidad
 - Servicios Web, Repositorios y definiciones basadas en XML
 - Rendimiento y escalabilidad

3.8 METODOLOGIA APLICADA A LA TRAZABILIDAD

3.8.1.- Definición de Requerimientos.

En los inicios del desarrollo del sistema OLTP (sistema actual que genera muchos datos y que no se le ha explotado como tal), se contempló otorgar trazabilidad a los productos generados, el eslabón que quedaba suelto por enlazar a la cadena de suministro era los datos generados por el distribuidor.

Para ello al aspirante de grado se le otorga una lista de reportes adicionales a la trazabilidad que la gerencia necesita saber; muchos de los cuales se responden de los datos contenidos en la base de datos actual.

Dichos reportes necesitan ser diseñados de tal manera que los usuarios puedan hacer uso de ellos; actualmente el sistema operacional genera los datos necesarios, pero no otorgan reporte alguno salvo las etiquetas que se imprimen y van adheridas a los productos así mismo el manifiesto de envío y una factura proforma.

A partir de los datos generados por el sistema actual el usuario podrá tener acceso a la información necesaria para la toma de decisiones un claro ejemplo a saber sería lo siguiente:

3.8.2 Producción Total Por Lote y campo.

Para fines estadísticos es necesario saber la cantidad de cajas producidas, el total en peso, por fecha determinada, por temporada, la cantidad de productos empacados y la calidad de los mismos.

3.8.3 Producción por Línea de Empaque.

Para cubrir la buena administración del piso de producción; la gerencia tiene la necesidad saber:

Que línea de empaque tiene mayor productividad.

Que línea de empaque le trabaja más a que lote-campo

En qué línea de empaque que productos se producen más, con qué calidad, cantidad de rechazos que tiene y con qué frecuencia.

Por cada orden de trabajo que el empacador cumple, se registran varios datos, mismos que son necesarios para cumplir con los requerimientos de trazabilidad.

Desde el momento del empaque de cada producto se le está identificando de que lote-campo es la cosecha, así mismo en la descripción de cada producto tiene asociada ciertas características como son : Peso, color, tamaño, calidad, variedad, puede ser también la cantidad de piezas contenidas en una caja; todas ellas unidas al Lote-campo y fecha de empaque así como al identificador único que se le otorga al empacador por parte del GS1 y el identificador otorgado a cada producto por el mismo empacador se obtiene el GTIN por sus siglas en inglés (*Global Trade International Number*).

3.8.3.1 Conformación de una caja (Producto).

Para los fines de la trazabilidad cada producto queda conformado de la siguiente manera:

GTIN + Lote + FechaEmpaque

3.8.3.2 Conformación de Pallets

Para fines del buen manejo de la información los pallets se pueden llegar a conformar de la siguiente manera:

Estándar

El total de cajas son de un solo producto, mismo lote, campo y misma fecha de empaque.

Mixtos

El total de cajas contenidas pueden ser de la siguiente manera:

Mismo Producto:

- Pero con diferente lote-campo y fecha de empaque.
- Diferentes Productos:
- Diferente producto, lote-campo y fecha de empaque.

Identificador de Pallet (SSCC por sus siglas en ingles *Serial Shipping Container Code*) se conforma de la siguiente manera:

Numero de GS1 (empacador) + Numero consecutivo de producción.

3.8.4 Manifiesto de Embarque.

Para cumplimentar el siguiente paso de la trazabilidad al momento del embarque se registran los siguientes datos:

3.8.4.1 De la empresa transportista.

- Nombre de la empresa
- Datos del tractor
- Datos de la caja (remolque)
- Datos del Chofer

3.8.4.2 Del Empacador.

- Identificador del embarque
- Cantidad Pallets a enviar
- Detalle de cada pallet (contenido del mismo)
- Fecha de envío
- Datos de la agencia aduanal, etc.

3.8.4.3 Del Distribuidor (Cliente).

- Datos del distribuidor

Los datos anteriormente mencionados están contenidos en el reporte de manifiesto de embarque, por lo tanto actualmente se lleva un registro en la base de datos operacionales; mismos que son pieza fundamental en la trazabilidad de productos frescos.

3.9 Manifiestos de Embarque con el Distribuidor.

El distribuidor registra en su base de datos:

Los manifiestos recibidos por empacador, así mismo que productos, cantidad de pallets, cantidad de cajas, fechas de arribo, fechas de envío, empresa transportista que los llevó, etc.

Los pallets enviados a sus clientes, cantidad de cajas, que productos, de que lotes-campo provienen.

3.10 Grupo de Trabajo.

Debido a que el presente trabajo es para efectos de tesis, el grupo de trabajo se conforma solo por el aspirante de grado; una vez que se pretenda llevar a cabo, será necesario conjuntar un equipo de trabajo mucho mayor que cubran diferentes roles para diseñar la arquitectura completa del almacén de datos e Inteligencia de negocios.

3.11 Requerimientos técnicos.

Basándose en la necesidad de la gerencia y para lograr una trazabilidad se puede determinar que necesita:

- 1.- Portal de Internet
- 2.-Herramientas para análisis
- 3.-Informacion histórica
- 4.-Acceso a la información de forma oportuna y veraz.

3.11.1 Portal de Internet.

El portal de internet es una aplicación web que además de otorgar trazabilidad, se podrán ver los inventarios de producción como son:

- Inventario General
- Productos en cuarto frio
- Productos en transito
- Productos en distribuidor

Se determina así para que las partes involucradas tengan acceso a la misma información en un tiempo más cercano al real desde su navegador de internet.

3.12 Herramientas para análisis.

La importancia de la información necesaria para la toma de decisiones debe de estar lo más accesible posible, de una forma amigable, de modo que los usuarios puedan disponer de la información para realizar una acción.

Actualmente cierta información es obtenida a través de reportes hechos con Microsoft Excel; tales como:

3.12.1 En el emparador.

- Producción por lote y campo
- Unidades producidas por línea de empaque
- Unidades producidas por emparador
- Cantidad de cajas recibidas por lote y campo
- Cantidad en peso recibidas por lote y campo
- Cantidad de pallets enviados por manifiesto
- Cantidad de cajas enviadas por manifiesto
- Número de envíos realizados

3.12.2 En el distribuidor.

- Cantidad de manifiestos recibidos por emparador
- Cantidad de pallets recibidos por emparador
- Cantidad de cajas recibidas por emparador
- Cantidad de cajas por producto
- Cantidad de ventas realizadas a sus clientes
- Cantidad de pallets vendidos por cliente, etc.

Se puede contar con esta información pero no con el tiempo y la veracidad que se requiere para la buena toma de decisiones.

Lo que no se puede obtener al momento es la trazabilidad; uniendo las bases de datos se obtendrá la ruta completa que siguió un producto desde su cosecha hasta el cliente minorista.

3.13 Información histórica.

La importancia de conocer como ha sido la producción por lote-campo, por línea de producción, por producto, por operador en una fecha determinada o rangos de fecha.

Que productos y en qué cantidades se le ha enviado a un distribuidor, en que fechas, quien lo ha transportado; todo ello con la finalidad de que en cualquier contingencia que

ocurra se pueda identificar las rutas que han seguido, quienes han sido las partes involucradas hasta cierto modo en la cadena de suministro.

3.14 Acceso a la información de forma oportuna.

La parte más importante es tener un rápido acceso a la información de trazabilidad en cualquier contingencia sanitaria que se pueda llegar a presentar, obteniendo información real en forma oportuna; además de esto, es importante que los usuarios no se vean afectados en tiempo al momento de consultar cualquier información.

Se necesita modelar la base de datos dimensional que contendrá el almacén de datos, donde estarán los datos de los sistemas operacionales.

Migrar la información a la base de datos dimensional. (Para fines de esta tesis no se contempla llevarla a término)

Las bases de datos de los diversos empacadores están de forma local, y la base de datos dimensional se propone que este ubicado en un server en la nube, al cual tendrían acceso vía conexión a Internet.

El distribuidor interactúa con su base de datos en forma remota, ya que su sistema es un portal web donde realiza sus transacciones.

Para que los usuarios tengan acceso a la información que necesitan se migrarán los datos necesarios a la base de datos dimensional.

Tanto de la perspectiva del distribuidor como del empacador-productor; todo para contar con la trazabilidad de los productos frescos, así como proporcionar información relevante a los usuarios que toman decisiones.

Cabe hacer mención que cierta información necesaria puede ser obtenida desde sus propias bases de datos, solo que no está organizada para que ellos puedan acceder de manera fácil y rápida.

3.15 Integración con otros sistemas.

No se contempla para fines de este proyecto; ya que solo será el modelado del almacén de datos, pero es importante que se haga mención porque las T.I. Están

cambiando a un ritmo vertiginoso y las herramientas que hoy nos ayudan, el día de mañana podrían quedar obsoletas.

3.16 Equipo e Infraestructura.

Es necesario revisar que infraestructura se va a utilizar. Para el caso de esta tesis se va a proponer mas no se contempla la adquisición del mismo.

Entre las buenas prácticas se recomienda separar el sistema de almacén de datos / inteligencia de negocios en diferentes servidores; esto con la finalidad de coadyuvar con la carga de todos los procesos, para ello se propone lo siguiente:

Un Servidor para la extracción, transformación y carga de los datos

Proceso ETL (*Extraction, Transformation and Load*) por sus siglas en ingles.

Un Servidor para manejar/administrar la base de datos dimensional

Un Servidor para la presentación de la información

Herramientas de presentación

Un dispositivo de almacenamiento masivo (SAN: Storage Área Network)

3.17 Modelado del Almacén de datos.

Se identificaron a las partes integrantes de la cadena de suministro y su relación con los eslabones tanto que la preceden así como la que está en el nivel próximo.

Productor / Empaque / Logística – Distribuidor / Minorista

- Determinó que sistemas transaccionales están asociados en la cadena de suministro.
- Se definió el origen de los datos operacionales. El modelo de datos fuente.
- Se definió el modelo de datos del *data warehouse*. El modelo de datos conceptual *data warehouse*.

En base a lo analizado se eligió utilizar el siguiente modelo para dar respuesta al caso práctico:

3.18 Modelo Dimensional.

El modelado dimensional se basa en HECHOS (*Facts*) y es una alternativa al modelado relacional. Sus principales ventajas son:

- Enfocado en el negocio y sus actividades.
- Permite búsquedas a gran velocidad.

El modelo dimensional divide el mundo de los datos en dos grandes tipos: las medidas y las descripciones del entorno de estas medidas (dimensiones). Las medidas, que generalmente son numéricas, se almacenan en las tablas de hechos y las descripciones de los entornos que son textuales se almacenan en las tablas de dimensiones.

3.18.1 Dimensiones.

- Son calificadores que dan sentido a los hechos.
- Son las diferentes perspectivas por la que se puede ver los hechos de negocio.

Ejemplo de las dimensiones: Tiempo, Producto, Ubicación, Además las dimensiones pueden ser jerárquicas, cada elemento de la dimensión puede ser modelado como una entidad.

Ej.: Producto – Tipo Producto – Fabricante

3.18.2 Medidas (hechos).

- Son datos cuantitativos acerca de un tema de negocio (acerca de los hechos).
- Generalmente son numéricos.

Ejemplo: Cantidad, monto, días, entre otros.

3.19 Modelo Estrella.

Se modeló el almacén de datos a través del Modelo en Estrella, que es un tipo de Modelo Dimensional, y por lo tanto, tiene como elementos una tabla de Hechos y tablas de Dimensiones. Lo que caracteriza al Modelo en Estrella, y lo diferencia de otros modelos dimensionales como el Modelo de Copo-de-Nieve, es cómo se lleva a cabo la relación entre las tablas de Dimensiones, más en concreto, como se construye la

relación entre los distintos niveles dentro de una Dimensión. En el Modelo en Estrella, cada dimensión utiliza únicamente una tabla.

3.20 Diseñar el data warehouse.

Definir la Arquitectura tecnológica *data warehouse*.

Solo como información referente al proceso especializado ETL (Extracción, transformación y carga) para fines de este caso práctico solo se especifica el siguiente proceso a seguir; En donde la extracción es la selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico *data warehouse*; la transformación es el proceso para transformar, sumarizar y realizar otros cambios en los datos operacionales para reunir los objetivos de orientación a los temas a explotar y finalmente la carga es la inserción sistemática de datos en el componente de almacenamiento físico *data warehouse*.

Definir las interfaces de Usuario para el acceso y explotación de los datos.

3.21 Implementación incremental para data warehouse.

3.21.1 Definición.

La fase de definición es el primer paso de la implementación, y consta del planteamiento de los objetivos del proyecto, para confirmar, documentar, y priorizar el nivel de los requerimientos del negocio para una adecuada implementación de la solución.

Los requerimientos son establecidos y priorizados en las reuniones de definición de requerimientos, de las cuales se obtiene como resultado modelos de negocio de alta definición que nos permitirán capturar en el menor tiempo posible el alcance de la solución.

Es importante en esta fase definir el plan inicial y los miembros que integrarán el equipo de proyecto.

3.21.2 Modelado.

El éxito de la fase de modelamiento, consiste en comprobar que el equipo de desarrollo ha entendido los procesos de negocio y los requerimientos establecidos en la fase de definición. Por lo tanto las reuniones periódicas con los líderes de usuarios finales y el desarrollo de un prototipo funcional arquitectónico, como medio de comunicación entre ellos, serán necesario para la confirmación de que los desarrolladores han comprendido los requerimientos del negocio.

El prototipo arquitectónico también será utilizado, para validar el modelo de la arquitectura y así reducir el riesgo de desarrollo. Aquí será necesaria la creación de un ambiente de desarrollo para soportar estas actividades.

Adicionalmente se deberá definir el equipo de pruebas y administración de la solución para una revisión del modelo arquitectónico. En lo posible este equipo deberá estar compuesto por miembros de los equipos de usuarios que dominan el tema y tienen el compromiso de sacar adelante el proyecto.

En las fases de construcción y producción; que aunque no se llevaran a cabo para fines de esta tesis se hará mención en qué consisten y que es lo que se pretende lograr.

3.21.3 Construcción.

La fase de construcción tiene como objetivo obtener una solución que optimice los requerimientos del negocio. En esta fase, se utiliza un proceso de intercambio para refinar los componentes desarrollados para la adquisición de información y acceso a la misma, hasta que se logre alcanzar el fin propuesto en todos los requerimientos del negocio.

Una vez planeadas las iteraciones deben ser completadas por cada partición y deberán estar integradas dentro de los componentes de administración. Por último la solución completa es probada. Los líderes de usuarios proveerán de los detalles mientras duren estas iteraciones; Finalmente es requerido en esta fase una prueba de desempeño y la creación de los materiales de documentación como soporte de la solución.

3.21.4 Producción.

Esta fase comprende desde la instalación del sistema hasta que finalmente este se encuentre funcionando en el ambiente de producción. Es requerida una última validación y carga de la información inicial, para comenzar a administrar el crecimiento y soporte de la solución; el soporte de la aplicación es dado de manera transitoria por el equipo de desarrollo del cliente o también llamado equipo de soporte en post-producción.

CAPITULO IV

4 Desarrollo

Basándose en que para fines de este caso práctico, las partes que conforman la cadena de abasto de productos frescos del campo son el empacador y distribuidor, mismos que están obligados a registrar un paso atrás y un paso adelante en la compra venta de sus activos, se logra tener datos suficientes para llevar a cabo una rastreabilidad confiable; desde el campo hasta el minorista.

Como se van registrando los datos en cada socio comercial en la cadena de abastecimiento, mismos datos que generaran información que ayudan en el proceso de rastreabilidad; a continuación se describe brevemente desde el registro de datos por parte del distribuidor y empacador.

4.1 Distribuidor.

Lleva un registro detallado de sus movimientos de envío-recepción de sus productos, tales como:

De que empacadores recibe productos.

- Cantidad de productos que recibe.
- Cantidad de pallets recibidos.
- Cantidad de cajas recibidas.
- Que empresa transportista lo ha hecho.

Hacia que clientes le envía; de la misma forma registrando lo siguiente:

- Cantidad de productos enviados.
- Cantidad de pallets enviados.

- Cantidad de cajas enviadas.
- Que empresa transportista hace la operación de traslado.

4.2 Empacador.

Los registros que lleva a cabo el empacador es de lo más importante, ya que aquí es donde se le da la identidad a cada producto, y donde un producto cumple con una serie de atributos que lo hacen único, un producto puede estar compuesto de una serie de características tales como: Variedad, tamaño, color, peso, marca comercial por mencionar algunas.

El empacador como parte fundamental de la cadena de abasto, está obligado a llevar un registro detallado acerca del origen de los productos que ahí se procesan; así mismo lleva un registro hacia donde envía sus productos, y quienes los transporta; algunas de los datos más relevantes que se registran son los siguientes:

Origen de cada cosecha:

- Lote
- Cantidad en cajas
- Peso total de la cosecha

Se le otorga un identificador a la cosecha recibida, misma que sirve al momento de ser asignada a una línea de empaque, ayudando así a determinar la cantidad de cosecha procesada por línea de empaque, así como la cantidad de productos con origen en un lote en específico; en este proceso de la producción nace la identidad de cada producto, misma que ayuda a identificarlo en los procesos subsiguientes de la cadena de abasto.

Una vez que sale del proceso de empaque cada caja (producto) es asignada a un pallet (tarima) al cual se le asigna un número único de identificación a través de la cadena de abasto llamado SSCC (*serial shipping container code*); cuando el empacador cumple con la cantidad de cajas solicitadas y que ya fueron asignadas a un pallet es cuando se hace el envío de las mismas hacia el distribuidor; recapitulando:

Proceso caja – Proceso Pallet – Proceso Envío



GTIN + Lote + Fecha



SSCC



Manifiesto de Envío

Para efectos de ir cerrando el círculo la empresa prevé contar con registros que sean incluidos en los datos de la cadena de abastos de productos frescos lo siguiente:

4.3 Productor.

Que el productor proporcione datos acerca de sus procesos de zafra como lo son:

Trabajos previos a la siembra, labranza, fertilización, aplicación de insecticidas (si los hubiera), riegos (tipo de agua usada), etc.; que lleve registros de todos sus procesos de producción en su base de datos que es lo que está produciendo, donde lo está produciendo, variedades, como lo siembra, cuidados del mismo, etc., también registrar sus tiempos de cosecha, en que cantidades, de donde lo cosecha, a quien le está vendiendo, quien se lo empaqueta, etc., cabe hacer notar que aquí al “producto” del productor para fines del sistema se le denomina *Comodity*.

La norma de rastreabilidad está pidiendo que se registre al menos un eslabón al frente y otro hacia atrás en la cadena de suministro, esto con el único fin de poder darle un seguimiento preciso a cada producto procesado y sea de fácil localización la ruta que siguió el mismo desde su origen hasta el destino.

Los registros con los que se cuentan para fines de llevar a cabo la rastreabilidad en este caso práctico son los proporcionados por el empaquetador y distribuidor, si ya sabemos que están obligados a llevar un registro antes y después, se puede decir que se cuentan con datos de su origen (productor), empaque (transformación), distribuidor y minorista, estos datos son de gran ayuda para encontrar el origen y traslado de un producto hasta el minorista; en la figura 22 se puede ver un diagrama en forma general de la cadena de suministro de productos frescos del campo.

4.4 Diagrama general.

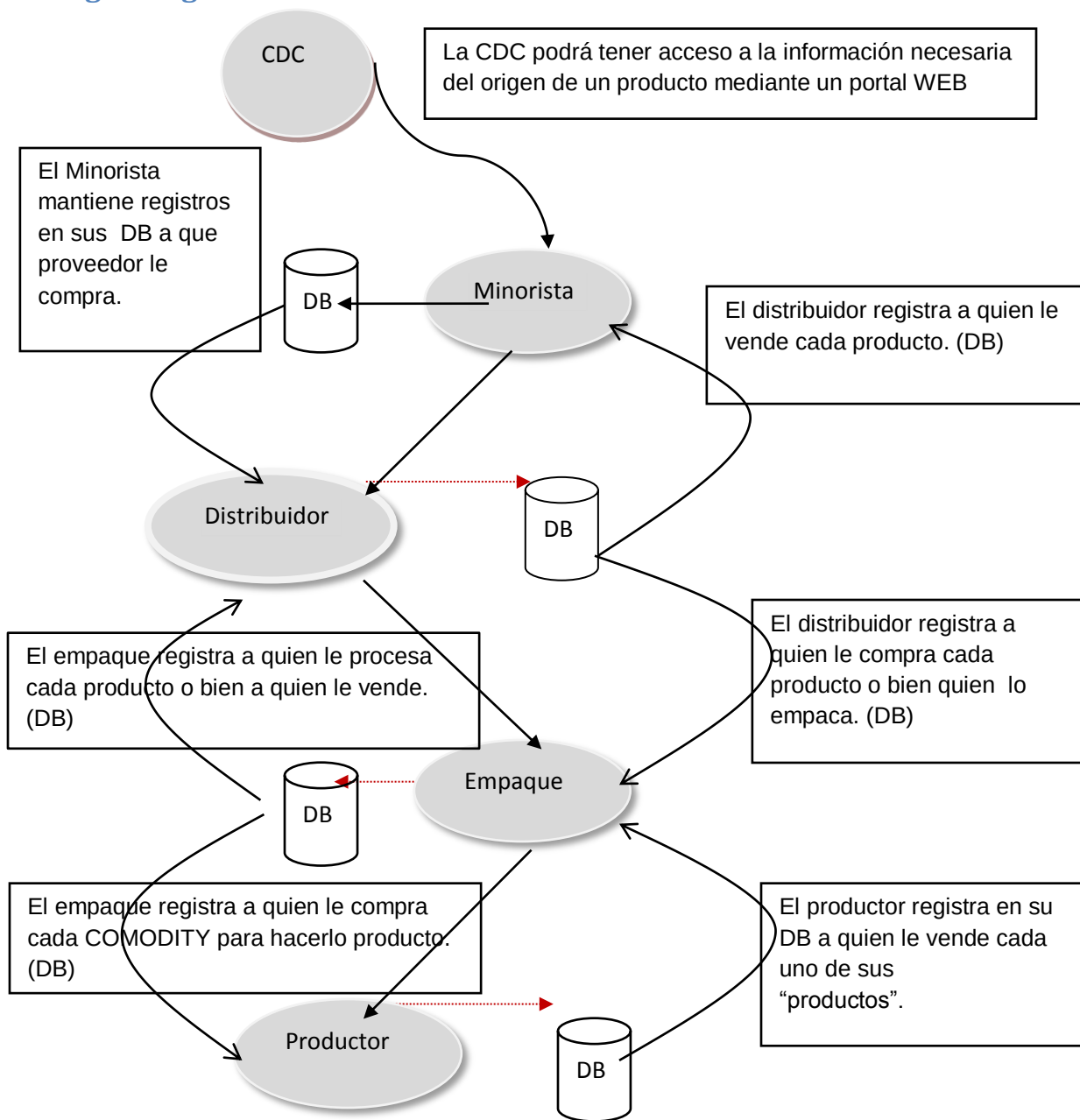


Figure 10 Diagrama general.

4.5 La herramienta de planeación creada por Kimball.

Ofrece una perspectiva general, independientemente de las preferencias de la base de datos o la tecnología, al mismo tiempo que hace que la identificación de los esfuerzos de desarrollo sea razonablemente manejable.

Cada línea de la herramienta de planeación corresponde a un proceso de negocio de la organización; cada columna corresponde a una dimensión del negocio.

Las celdas de la matriz son sombreadas para indicar cuales columnas están relacionadas con cada línea.

Como ejemplo se muestra el proceso de negocio de todos los envíos realizados con todas y cada una de las columnas, en este caso todas las dimensiones con las cuales están relacionadas.

Las medidas (tabla de hechos): Envíos realizados

Con las dimensiones de: Fecha, Empacador, Distribuidor, Transporte, Zona geográfica, Productos. (Véase la figura 23 de herramienta de planeación para la rastreabilidad de productos frescos)

Proceso de Negocio / Evento	DIMENSIONES										
	Fecha	Empacador	Distribuidor	Transporte	Zona Geografica	Producto	Envios	Pallets	Cajas	Línea de Empaque	Minorista
EVENTOS	X	X	X	X	X	X	X	X	X		X
ENVIOS REALIZADOS	X	X	X	X	X	X					
PALLETS ARMADOS	X	X	X		X	X		X			
PRODUCCION CAJAS	X	X	X		X	X			X	X	X
PRODUCCION LINEA DE EMPAQUE	X	X	X		X	X			X	X	X
PRODUCCION LOTE-CAMPO	X	X			X						
RECIBOS DE MANIFIESTOS	X	X	X	X	X	X	X	X	X		
VENTAS CLIENTES-DISTRIBUIDOR	X		X	X	X	X	X	X	X		X

Figure 11 Herramienta de planeación para la rastreabilidad de productos frescos.

4.6 Proceso que determina la rastreabilidad.

Después de haber definido con la herramienta de planeación los procesos de negocio a modelar se procede a diseñar los procesos a diagramas de alto nivel con el fin de que sean más explícitos y que las personas que se integren al equipo de desarrollo posteriormente puedan ver de forma general el proceso y las partes que lo conforman.

Cada línea de la herramienta de planeación corresponde a un proceso de negocio, en este caso se empezará con el diagrama de Eventos, que es el que engloba de manera general y resuelve la rastreabilidad de productos frescos del campo. En la figura 24 se aprecia el proceso de negocio “Eventos” con todas las dimensiones que tienen relación a dicho evento; en este caso corresponde a la tabla de hechos “Eventos”

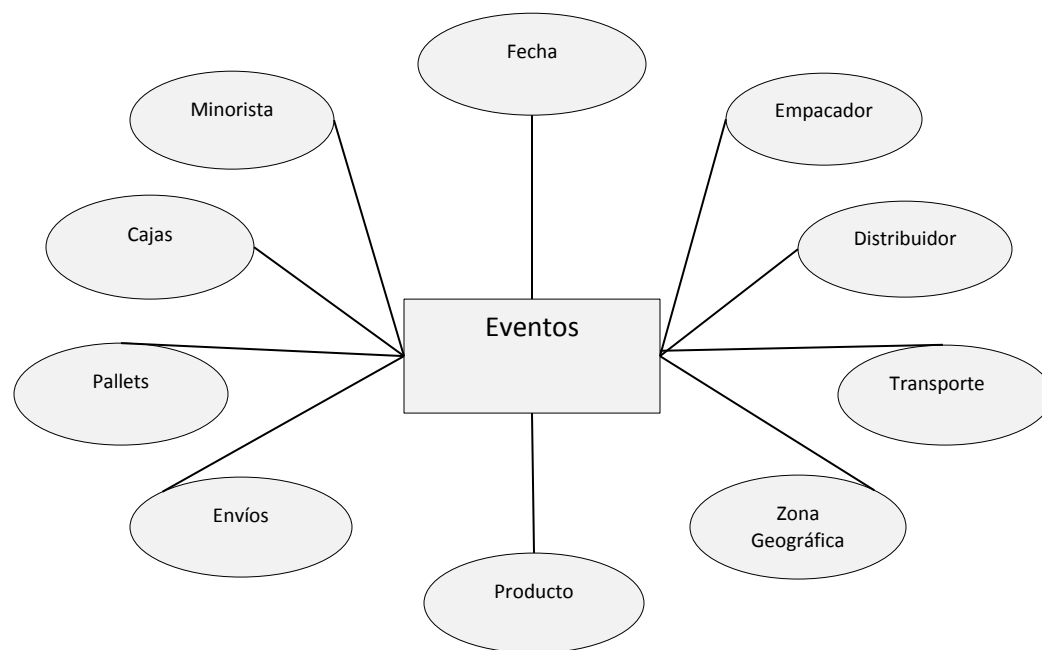


Figure 12 Eventos.

Con este diagrama de alto nivel lo que se puede apreciar es que puede resolver el problema de rastreabilidad tanto desde la parte del empacador hacia el minorista y puede decir con exactitud desde que Lote se originó un producto y ¿hacia Qué distribuidores?, ¿A Que Minoristas fueron enviados dichos productos?; a continuación se enlistan unas posibles preguntas de las muchas que se pueden llegar a contestar.

Para responder la pregunta o preguntas que se estarán haciendo a partir de que pudiera ser detectado un problema de inocuidad.

- ¿Quién o Quienes comercializaron productos con origen del Lote en cuestión?
- ¿Quién o Quienes lo distribuyeron?
- ¿Quién o Quiénes pudieron empacar productos con origen del Lote en cuestión?
- ¿Qué productos tuvieron origen del Lote en cuestión?
- ¿Cuántos productos fueron?
- ¿Cuándo fueron procesados tales productos?

Solo por mencionar algunas de las preguntas a contestar con la llevada a cabo del modelo anteriormente mencionado

4.7 Proceso Envíos Realizados por empacador.

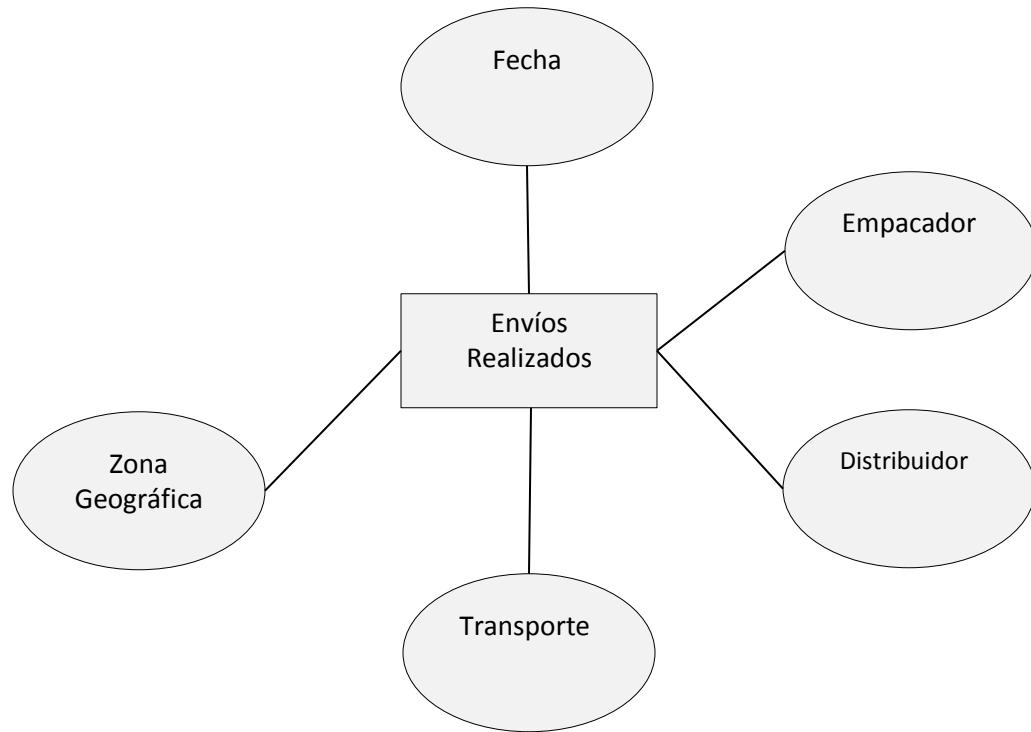


Figure 13 Modelo de Envíos realizados por empacador.

El modelo de la figura 25, proporciona información requerida por la administración, tal como:

- Cantidad de envíos realizados por empacador
- Cantidad de pallets enviados por embarque.
- Total de pallets enviados.
- Cantidad de cajas enviadas por embarque
- Total de cajas enviadas
- Peso total por embarque
- Peso total enviado.

Por mencionar alguna información que podrá salir de ahí.

4.8 Proceso pallets realizados por empacador.

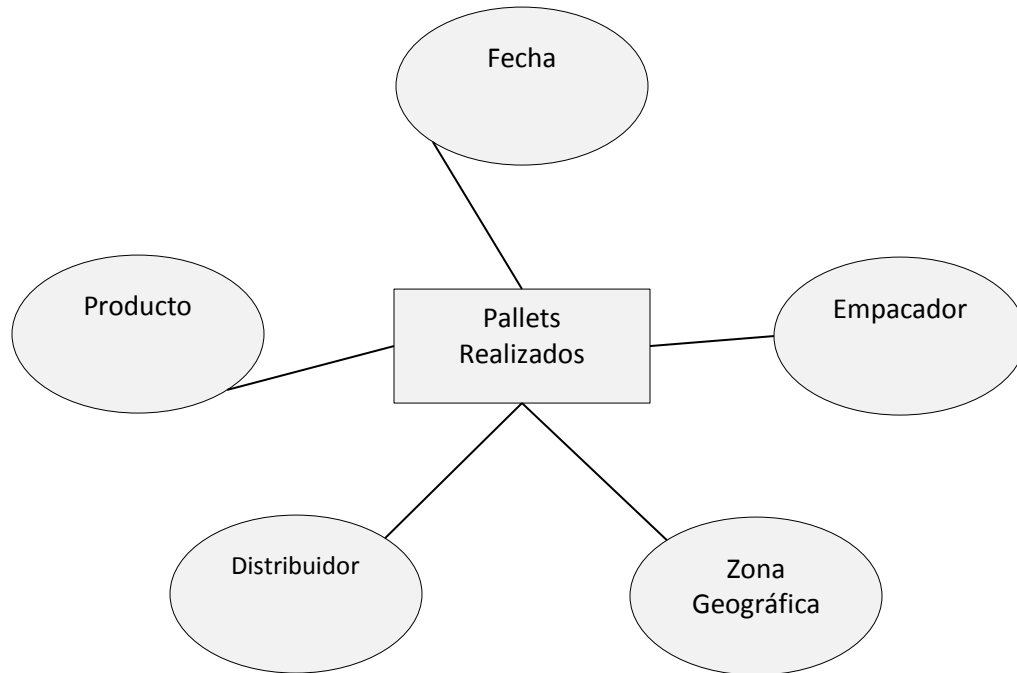


Figure 14 Modelo de pallets realizados por empacador.

En la figura 26 se puede apreciar el modelo de pallet realizados, genera información en la parte del empacador, donde les proporciona información necesaria para la toma de decisiones; tal como:

- Cantidad de pallets realizados
- Peso total de pallets
- Cantidad de pallets realizados por empacador
- Cantidad de pallets realizados por distribuidor

4.9 Proceso de productos empacados.

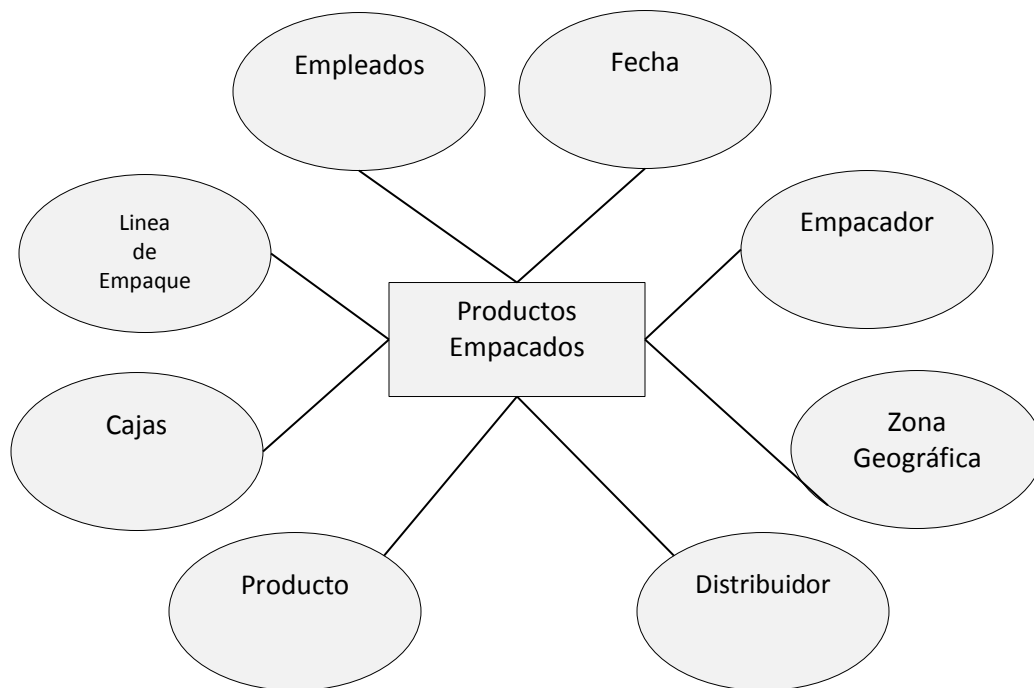


Figure 15 Modelo de productos empacados.

En el modelo de la figura 27 de productos empacados, se refiere a las cajas que se empacan en determinado momento, que es lo que viene siendo un producto, mismo que agrupa ciertas características, el modelo nos arroja información también importante en la toma de decisiones como lo es:

- Cantidad de cajas producidas
- Cantidad de cajas producidas por producto
- Cantidad de cajas producidas por distribuidor
- Cantidad de cajas producidas por línea de empaque
- Cantidad de cajas por producto por línea de empaque
- Peso total Procesado por línea de empaque

Solo por mencionar alguna información que podrá ser proporcionada por dicho modelo.

4.10 Proceso de Producción por Lote.

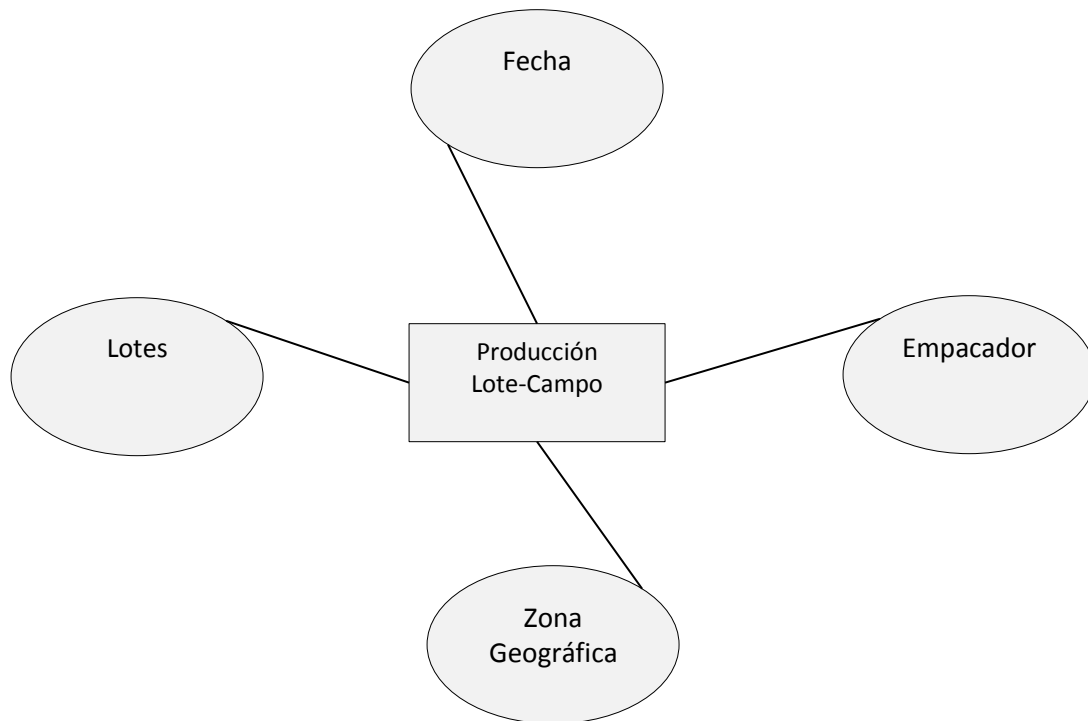


Figure 16 Modelo de producción de Lote-Campo.

EL modelo representado en la figura 28, ayuda a comprender el flujo de datos registrados al momento de recibir la materia prima desde los lotes del productor; de aquí partimos para contestar algunas de las preguntas en los requerimientos de información como son:

- Total en peso de la producción
- Total de producción enviada por empacador

Así mismo desde el registro puntual de las cosechas por los lotes, la asignación de la producción de los Lotes a las líneas de empaque para que de origen a un producto, que es donde se liga el Lote de origen con el Producto o productos resultantes, los datos subsecuentes que se registren oportunamente le darán al sistema el poder de la rastreabilidad.

4.11 Proceso por Línea de Empaque.

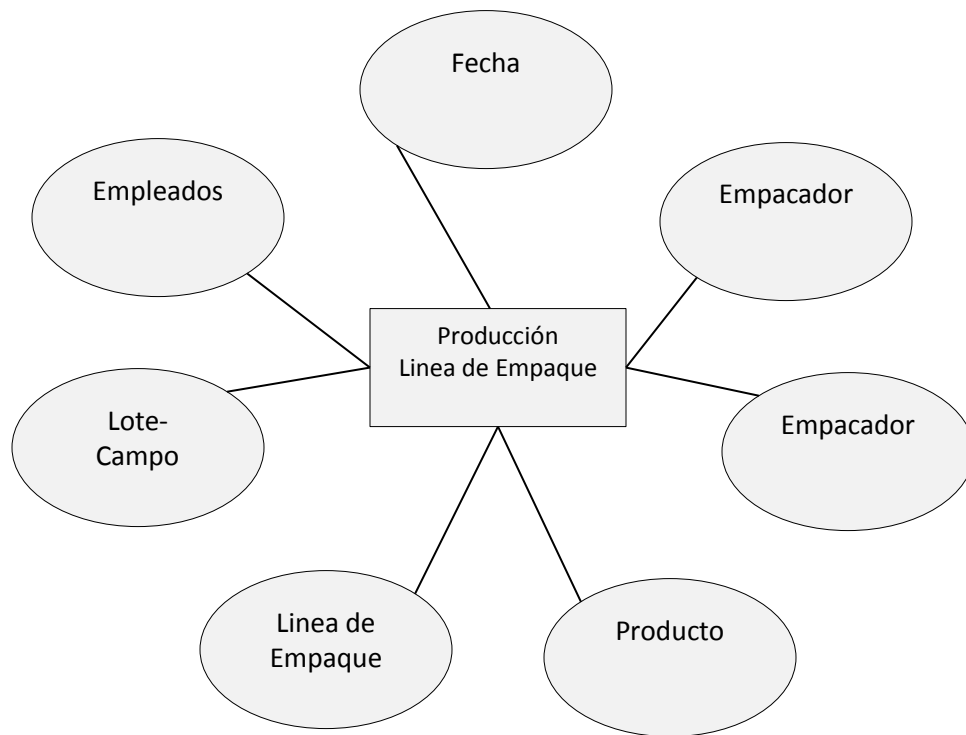


Figure 17 *Modelo de producción por línea de empaque.*

El modelo de la figura 29 nos ayuda a responder preguntas tales como:

- Producción total por línea de empaque
- Producción por empleado
- Número total de productos producidos por todas las líneas de empaque
- Total de cajas producidas por lote-campo
- Total de productos producidos por lote

4.12 Proceso de manifiestos recibidos por distribuidor.

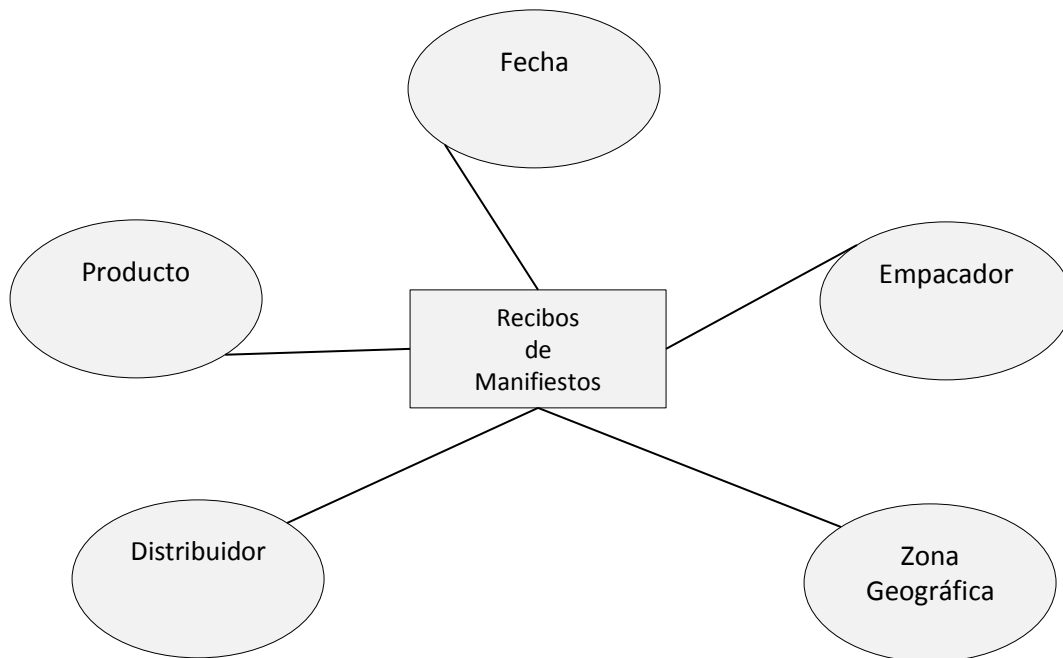


Figure 18 Modelo manifiestos recibidos por distribuidor.

El modelo de la figura 30 se refiere a los datos que se procesan por el distribuidor al recibir los diferentes envíos realizados por empacador, el cual le proporciona información al mismo distribuidor que es de suma importancia para llevar a cabo la trazabilidad.

La diferente información que podrá ser proporcionada es la siguiente:

- Total de manifiestos recibidos
- Total de manifiestos recibidos por empacador
- Total de pallets recibidos
- Total de pallets recibidos por empacador
- Total de cajas recibidas
- Total de cajas recibidas por empacador
- Total de pallets recibidos por producto
- Total de cajas recibidas por producto
- Total en peso de las cargas recibidas

Solo por mencionar algunos ya que no incluyen todas las posibles combinaciones de información que pudiera necesitar en algún momento dado el distribuidor.

4.13 Proceso de ventas a clientes minoristas por distribuidor.

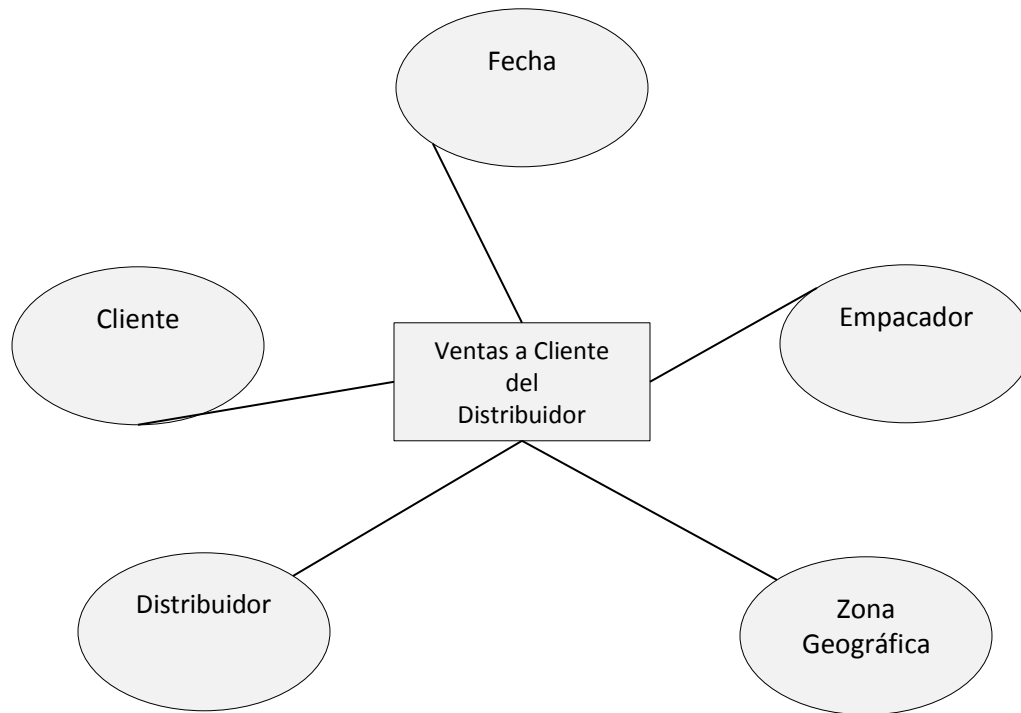


Figure 19 Modelo de ventas a Clientes minoristas por distribuidor.

El modelo de la figura 31 nos muestra el flujo de la información a seguir cuando se quiere encontrar información del origen de cada producto, vital en la trazabilidad de cada uno de ellos.

- Las ventas por cliente
- El total de pallets vendidos por cliente
- El número total de cajas por producto vendidas cliente
- Cantidad de cajas vendidas por Lote filtrados por cliente
- Cantidad de pallets vendidos por Lote filtrados por cliente

En Seguida se hace una presentación en detalle de las tablas de hechos.

Para las tablas de hechos y dimensiones se utilizan los siguientes sufijos en los campos

SK: Llave subrogada

PK: Llave primaria

FK: Llave foránea

4.14 Tabla de Hechos: Eventos

Tabla:	FACT_EVENTOS
Tipo de tabla:	Hechos
Alias:	Tabla de eventos
Descripción:	Registros de eventos

Nombre del Atributo	Descripción	Nombre Alternativo	Valores ejemplo
EventoSK	Llave subrogada de evento		1,2,3
CajaFK	Llave foránea de cajas	CajaID	1,2,3,4,5
ProductoFK	Llave foránea de producto	ProductoID	1020,1030
FechaFK	Llave foránea de fechas	FechaID	1,2,3
DistribuidorFK	Llave foránea de distribuidor	Distribuidor ID	1,2,3,4
EmpacadorFK	Llave foránea del empacador	Empacador ID	1,2,3
TransporteFK	Llave foránea del transporte	Transporte ID	0,1,2,3,4
PalletFK	Llave foránea del pallet	Pallet ID	0,1,2,3
EnvioFK	Llave foránea del envio	Envio ID	0,1,2,3,4
MinoristaFK	Llave foránea del minorista	Cliente ID	0,1,2,3
Zonageografica	Llave foránea de la zona geográfica	ZonaGeograficaID	1,2,3,4
PesodeCaja	Peso de la caja del evento	Peso	1,2,5,10

Figure 20 Tabla de hechos, Fact_Eventos.

4.15 Tabla de Hechos: Envíos Realizados.

Tabla:	FACT_Envios_realizados
Tipo de tabla:	Hechos (Facts)
Alias:	Envíos realizados
Descripción:	Cantidad de envíos realizados por empacador

Nombre del Atributo	Descripción	Nombre Alternativo	Valores ejemplo
SalidaSK	Llave subrogada	ID Salida	1,2,3,4
ManifiestoFK	Llave foránea de Manifiestos	ID Manifiesto	1,2,3,4,
EmpacadorFK	Llave foránea del Empacador	ID del empacador	1,2,3,4
DistribuidorFK	Llave foránea del Distribuidor	ID del distribuidor	1,2,3,4
TransporteFK	Llave foránea del Transportista	ID del transportista	1,2,3,4
ZonaGeograficaFK	Llave foránea de la zona geográfica	ID zona geográfica	1,2,3,4
FechaFK	Llave foránea de la fecha de envío	ID de fechas	1,2,3,4
ProductoFK	Llave foránea de Productos	ID del Producto	12,20,500,1050
TotalPalletsCarga	Cantidad de pallets en un envío realizado	Cantidad de pallet	18,19,20,22
PesoTotalCarga	Peso total de los pallets en el envío realizado	Peso total Envío	10500, 19200

Figure 21 Tabla de Hechos, FACT_Envios_realizados.

4.16 Tabla de Hechos: Pallets Realizados.

Tabla:	FACT_Pallets_realizados
Tipo de tabla:	Hechos (Facts)
Alias:	Pallets realizados
Descripción:	Cantidad de pallets realizados por empacador

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
PalletSK	Identificador único del pallet	Etiquetapallet	01,02,20,120,999
DateFK	Llave foránea de la dimensión de fechas	ID de la fecha	1,2,3,4,5,11,100
EmpacadorFK	Llave foránea del empacador	ID del Empacador	1,2,3,4
DistribuidorFK	Llave foránea del Distribuidor	ID del Distribuidor	1,2,3,4
ProductoFK	Llave foránea del producto	ID del Producto	1, 10, 20, 200, 1020
CantidadCajas	Total de cajas por pallet	Cajas por pallet	20, 60, 100
PesoTotalPallet	Peso total de pallets	Peso del pallet	800, 900, 1020

Figure 22 Tabla de Hechos, FACT_Pallets_realizados.

4.17 Tabla de Hechos: Cajas Empacadas.

Tabla:	Fact_Cajas_Empacadas
Tipo de tabla:	Hechos (Facts)
Alias:	Cajas empacadas
Descripción:	Cajas producidas

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
CajaSK			
EtiquetaCaja	Identificador de la caja, según el lote de origen y fecha producida	Identificador de Caja	01500123020LOT135
DateFK	Llave foránea de la dimensión de fechas	ID de la fecha	1,10,21,100
EmpacadorFK	Llave foránea del Empacador	ID del empacador	1,2,3,4
DistribuidorFK	Llave foránea del Distribuidor	ID del distribuidor	1,2,3,4
ProductoFK	Llave foránea del Producto	ID del producto	1,10,500,1020
LineadeEmpaqueFK	Llave foránea de la línea de empaque	ID de la línea de empaque	1,2,3,4
PesodeCaja	Peso total de cajas por producto	TotalPesoporProducto	200,000

Figure 23 Tabla de hechos, FACT_Cajas_empacadas.

4.18 Tabla de Hechos: Producción por línea de empaque.

Tabla:	Fact_LineasDeEmpaque
Tipo de tabla:	Hechos (Facts)
Alias:	LineaDeEmpaque
Descripción:	Producción por Líneas de Empaque

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
LineaEmpaquePK	Llave foránea de la línea de empaque	ID Línea Empaque	1,2,3
DateFK	Llave foránea de la dimensión de fechas	ID de la fecha	1,2,3,4,5
EmpacadorFK	Llave foránea del empacador	ID del Empacador	1,2,3,4
DistribuidorFK	Llave foránea del distribuidor	ID del Distribuidor	1,2,3,4,5
ProductoFK	Llave foránea del Producto	ID del Producto	1,2,40,500,1020
Lote-campoFK	Llave foránea de los lotes-campos	ID del Lote-campo	1,2,3
QtyProductosLinea	Cantidad de productos producidos por línea	ProductosxLinea	20,500,1000

Figure 24 Tabla de Hechos, FACT_LineasdeEmpaque.

4.19 Tabla de Hechos: Manifiestos Recibidos por distribuidor.

Tabla:	FACT_Recibos_Distribuidor
Tipo de tabla:	Hechos (Facts)
Alias:	Cargas recibidas por el distribuidor
Descripción:	Manifiestos recibidos por el distribuidor

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
ReciboIDSK	Llave Subrogada		1,2,3,4
FechaFK	Llave foránea de la fecha	FechaID	1,2,3,
EnvioFK	Llave foránea de envíos	EnvioID	HD1001
DistribuidorFK	Llave foránea del Distribuidor	DistribuidorID	1,2,3
EmpacadorFK	Llave foránea del empacador	EmpacadorID	1,2,3
ZonageograficaFK		ZonaGeograficaID	1,2,3,4
TransporteFK	Llave foránea del transporte	TransporteID	1,2,3,4
ProductoFK	Llave foránea del producto	ProductoID	1,10,500,2900
PalletFK	Llave foránea del pallet	Pallet ID	1,2,3
CajaFK	Llave foránea de la caja	Caja ID	1,2,3
PesoCaja	Peso de la caja	Peso	1,5,10

Figure 25 Tabla de Hechos, FACT_Recibos_Distribuidor.

4.20 Tabla de Hechos: Ventas del Distribuidor.

Tabla:	Fact_Distribuidor_Ventas
Tipo de tabla:	Hechos (Facts)
Alias:	Ventas Distribuidor
Descripción:	Registro de las ventas del Distribuidor

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
VentaSK	Identificador de la venta	ID Venta	1,2,3
EnviosFK	Numero de envio	ID del Envio	1,2,3
FechaFK	Llave foránea de la fecha	Identificador de la fecha	1,2,3
MinoristaFK	Llave foránea del Cliente	Identificador del cliente	1,2,3
DistribuidorFK	Llave foránea del distribuidor	Identificador del distribuidor	1,2,3
ProductoFK	Llave foránea del producto	Identificador del Producto	1,30,500,1016
TransporteFK	Llave foránea de transporte	TransporteID	1,2,3,4
ZonaGeograficaFK	Llave foránea de la zona geográfica	Zona geográfica ID	1,2,3
PalletsFK	Llave foránea del pallet	Pallet ID	1,2,3
CajaFK	Llave foránea de la caja	Caja ID	1,2,3,4
PesoCaja	Peso de la caja	Peso	1,5,10

Figure 26 Tabla de Hechos, FACT_Distribuidor_Ventas.

4.21 Tabla de Hechos: Producción por Lote.

Tabla:	FACT_Produccion_Lote
Tipo de tabla:	Hechos (Facts)
Alias:	Producción por Lote
Descripción:	Registro de la producción por lote

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
CosechaSK	Llave subrogada de la cosecha	ID de la cosecha	1,2,3
LotePK	Llave foránea del Lote de producción	ID del Lote	1,2,3
ProductorFK	Llave foránea del Productor	ID del productor	1,2,3
EmpacadorFK	Llave foránea del empacador	ID del empacador	1,2,3
DateFK	Llave foránea de la fecha	ID de la fecha	1,2,3,4
TotalCajas	Número total de cajas	CajasTotal	70,90
PesoTotal	Peso total de la cosecha	TotalenPeso	800,1000

Figure 27 Tabla de Hechos, FACT_Produccion_Lote.

El otro de tipo de tablas utilizadas en el modelo dimensional son las tablas de dimensiones, las cuales ayudan a relacionar los hechos con varios registros de las dimensiones, por ejemplo una compañía de transporte puede ser utilizada por varios empacadores para transportar sus productos, o bien la cantidad de envíos realizados por empacadores a un distribuidor.

*

*

Para cerrar el modelado dimensional

4.22 Tabla de Dimensiones: Fecha.

Tabla:	Dim_FECHA
Tipo de tabla:	Dimensión
Alias:	Fechas
Descripción:	Registro de Fechas

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
FechaPK			1,2,3
Hora	Hora del evento	Horaevento	15:00
Fecha	Fecha Completa	FechaCompleta	26 de Abril de 2012
Día	Día del mes	Diaevento	26
Semana	Número de semana	Semanaevento	4
Mes	Nombre del Mes	Número de Mes	Abril
Año	Numero de Año	Número de Año	2012

Figure 28 *Tabla de Dimensiones, DIM_FECHA.*

4.23 Tabla de Dimensiones: Empacador.

Tabla:	DIM_Empacador
Tipo de Tabla:	Dimensión
Alias:	Empacadores
Descripción:	Catálogo de empacadores

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
EmpacadorSK	Llave primaria subrogada		1,2,3,4
EmpacadorPK	Llave primaria del empacador	EmpacadorID	1,2,3,4,5
Nombre	Nombre o razón social del Empacador	Nombre del Empacador	San Clemente
Calle	Nombre de la calle	Nombre de la calle donde está ubicado el empaque	C. Jacintos
Colonia	Nombre de la colonia	Nombre de la colonia donde está ubicado el empaque	Col. San Gabriel
CodigoPostal	Numero de código postal	Código postal	267104

Figure 29 *Tabla de Dimensiones, DIM_Empacador.*

4.24 Tabla de Dimensiones: Distribuidores.

Tabla:	DIM-Distribuidores
Tipo de Tabla:	Dimensiones
Alias:	Empacadores
Descripción:	Catálogo de Distribuidores

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
DistribuidorSK	Llave primaria subrogada		1,2,3,4
DistribuidorPK	Llave primaria del empacador	Identificador del Distribuidor	1,2,3,4
Nombre	Nombre del Distribuidor	Nombre Distribuidor	GDM
Calle	Nombre de la calle donde está ubicado el distribuidor	Calle del Distribuidor	Magnolias
Colonia	Nombre de la colonia donde está ubicado el distribuidor	Nombre de la colonia	Saint Peter Valley
CodigoPostal	Número del código postal donde está ubicado el distribuidor	Número Código Postal	09780
GS1	Numero único GS1, identifica como miembro de la organización GS1	PrefijoGS1	61030505

Figure 30 Tabla de dimensiones, DIM_Distribuidores.

4.25 Tabla de Dimensiones: Transporte.

Tabla:	DIM_Transporte
Tipo de Tabla:	Dimensiones
Alias:	Transportistas
Descripción:	Catálogo de transportistas

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
TransporteSK	Llave primaria subrogada		1,2,3,
TransportePK	Llave primaria del transportista	Identificador de la empresa De transporte	1,2,3,4
Nombre	Nombre de la empresa Transportista	Nombre del transporte	Hermanos Pérez
IDTractor	Identificador del tractor	Numero de tractor	20,30,190
IDRemolque	Identificador del remolque	Numero de remolque	1,2,3,40
NombreChofer	Nombre del chofer	Nombre del chofer	Jorge Inzunza
Calle	Nombre de la calle donde está ubicada la empresa	Nombre de la calle	Valparaíso 500
Ciudad	Nombre de la ciudad donde está ubicada la empresa	Nombre de la ciudad	Juventino Rosas
CodigoPostal	Número del código postal donde está ubicada la empresa	Código postal	12345

Municipio	Nombre del Municipio	Nombre del municipio	Angostura
Estado	Donde está ubicada la empresa	Nombre del estado	Sinaloa
País	Nombre del estado donde está ubicada la empresa De transporte	Nombre del país	México

Figure 31 Tabla de Dimensiones, DIM_Transporte.

4.26 Tabla de Dimensiones: Zona Geográfica.

Tabla:	DIM_Zona_Geografica
Tipo de Tabla:	Dimensión
Alias:	Zona Geográfica
Descripción:	Datos de la zona geográfica donde está ubicado el empacador

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
ZonaGeograficaSK	Llave subrogada		1,2,3
ZonaGeograficaPK	Llave Primaria	Zgeográfica	1,2,3,4
Ciudad	Nombre de la ciudad	ZGCiudad	Salamanca
Estado	Nombre del estado	ZGEstado	Guanajuato
País	Nombre del país	ZGPaís	México
Región	Nombre de la región	ZGRegión	El Bajío
Población	Nombre de la población	Rancho	Salamanca

Figure 32 Tabla de Dimensiones, DIM_Zona_Geografica.

4.27 Tabla de Dimensiones: Productos.

Tabla:	DIM_Productos
Tipo de Tabla:	Dimensiones
Alias:	Productos
Descripción:	Catálogo de productos

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
ProductoSK	Llave primaria subrogada		1,2,3,4,
ProductoPK	Llave primaria de los productos	Identificador del producto	1,2,3
Descripción	Descripción del producto	identificador	Tomato Campary
Comodity	Tipo de producto	ProductoFresco	Tomato
Variedad	Variedad del producto	Identificador de Variedad	Tomate Bola
Calidad	Calidad del producto	CalidadProducto	Jumbo
Tamaño	Tamaño del producto	TamanoProducto	XL
Envase	Tipo de envase del producto	EnvaseProducto	Cartón
Color	Color del producto	ColorProducto	Rojo
Conteo	Cantidad de piezas contenidas	CantidadUnidades	15,16,18
Metodocrecimiento	Como es el método de crecimiento del	ComoSeCultiva	Cielo Abierto

EtiquetaProducto	Comodity	MarcaProducto	GDM
	Marca comercial del producto		
Peso	Cantidad en peso del producto	PesoProducto	11

Figure 33 Tabla de Dimensiones, DIM_Productos.

4.28 Tabla de Dimensiones: Líneas de Empaque.

Tabla:	DIM_LineaDeEmpaque
Tipo de Tabla:	Dimensiones
Alias:	Línea de Empaques
Descripción:	Catálogo de Líneas de Empaque

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
LineaEmpaqueSK	Llave primaria subrogada		1,2,3
LineaEmpaquePK	Llave primaria de la línea de empaque	ID LineaEmpaque	1,2,3,4
DescripcionLinea	Descripción de la línea	NombredeLinea	Linea1, Linea2

Figure 34 Tabla de Dimensiones, DIM_LineasDeEmpaque.

4.29 Tabla de Dimensiones: Lote y Campos.

Tabla:	DIM_Lote_Campo
Tipo de Tabla:	Dimensiones
Alias:	LoteCampo
Descripción:	Catálogo de lotes y campos

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
LoteCampoSK	Llave primaria Subrogada		1,2,3,4
LoteCampoPK	Llave primaria de LoteCampos	ID Lote_Campo	1,2,3,4
NombreDeLote	Descripción de Lote y Campo	NumeroDeLote	SJLote1

Figure 35 Tabla de Dimensiones, DIM_Lote_Campo.

4.30 Tabla de Dimensiones: Empleados.

Tabla:	DIM_Empleados
Tipo de Tabla:	Dimensiones
Alias:	Empleados
Descripción:	Catálogo de Empleados

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
---------------------	-------------	--------------------	-----------------

EmpleadoSK	Llave Primaria Subrogada		1,2,3
EmpleadoPK	Llave primaria del Empleado	ID del Empleado	1,2,3
Nombre	Nombre del Empleado	NombredeEmpleado	Jacinto Cenobio
FechaNacimiento	Fecha de Nacimiento	Nacimiento	26/04/1971
Calle	Nombre de la calle donde vive el empleado	CalleNombre	Gardenias
Colonia	Nombre de la colonia donde vive el empleado	ColoniaNombre	Francisco I. Madero
Ciudad	Nombre de la ciudad donde vive el empleado	CiudadNombre	Tijuana
Municipio	Nombre del Municipio donde vive el empleado	MunicipioNombre	Tijuana
Estado	Nombre del estado donde vive el empleado	EstadoNombre	Baja California
País	Nombre del empleado donde vive el empleado	PaisNombre	México

Figure 36 Tabla de Dimensiones, DIM_Empleados.

4.31 Tabla de Dimensiones: Minorista.

Tabla:	DIM_Minorista
Tipo de Tabla:	Dimensiones
Alias:	Cientes
Descripción:	Catálogo de Clientes del Distribuidor

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
ClienteSK	Llave primaria subrogada		1,2,3
ClientePK	Llave primaria del Cliente	ID Cliente	1,2,3
Nombre	Nombre del Cliente	NombreCliente	SAMSClub
Calle	Calle donde está ubicado el cliente	CalleCliente	Main
Colonia	Colonia donde está ubicado el cliente	ColoniaCliente	Chulavista
Ciudad	Ciudad donde está ubicado el cliente	CiudadCliente	Chulavista
ZipCode	Código postal del cliente	ZIPCliente	91909
Estado	Estado donde vive el cliente	EstadoCliente	California
País	País donde está ubicado el cliente	PaisCliente	USA

Figure 37 Tabla de Dimensiones, DIM_Clientes.

4.32 Tabla de Dimensiones: Envíos.

Tabla:	DIM_Envíos
Tipo de Tabla:	Dimensiones
Alias:	Envíos
Descripción:	Registros de los manifiestos de carga hacia el distribuidor

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
EnvioSK	Llave subrogada		1,2,3
EnvioPK	Llave primaria	NumeroEnvio	1,2,3,4
Manifiesto	Número de control del envío	Identificación del envío	MAN1001
NoAsignado	Para indicar si una caja o pallet han sido asignados a un Envio	NAsignado	NA

Figure 38 Tabla de Dimensiones, DIM_Envíos.

4.33 Tabla de Dimensiones: Pallets.

Tabla:	DIM_Pallets
Tipo de Tabla:	Dimensiones
Alias:	Tarimas
Descripción:	Pallets procesados (Armados,Enviados,Recibidos)

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
---------------------	-------------	--------------------	-----------------

PalletSK	Llave Subrogada		1,2,3
PalletPK	Llave primaria	PalletID	1,2,3
SSCC	Identificador único del pallet	EtiquetaPallet	12345670002501
NoAsignado	Indica si un pallet ha sido enviado, recibido	SinAsignar	NA

Figure 39 Tabla de Dimensiones, DIM_Pallets.

4.34 Tabla de Dimensiones: Cajas.

Tabla:	DIM_Cajas
Tipo de Tabla:	Dimenciones
Alias:	CajaProductos
Descripción:	Cajas procesadas (empacadas,enviadas,recibidas)

Nombre del Atributo	Descripción	Nombre alternativo	Valores ejemplo
CajaSK	Llave subrogada		1,2,3
CajaBarCode	Identificador de caja	EtiquetaCaja	123456-SJ01-201267
LineadeEmpaque	Línea de producción	LineaID	Linea1
Nombredeempleado	Nombre del empleado que empaco una caja	NumEmpleado	Juan Cárdenas
LoteParcela	Identificador del Lote donde se cosecha el commodity	NumLote	SJ01

Figure 40 Tabla de Dimensiones, Dim_Cajas.

4.35 Desarrollo del Modelo físico de datos.

4.35.1 Base de Datos

Se creará la base de datos del modelo dimensional DWHtrazabilidad

Creación de las tablas de dimensión y la tabla de hechos.

Para la creación de las tablas en el modelo lógico y físico se estará usando el ERStudio donde se podrá emigrar a la SMDB que se requiera por parte de la administración general de la empresa.

De manera que el siguiente modelado dimensional, trata de dar respuesta a la rastreabilidad de productos frescos mediante la unión de los datos necesarios para tal fin, migrando los mismos desde las bases de datos operacionales a través de ciertos procesos especializados (Extracción, transformación y carga).

Cada atributo contenido en la tabla de dimensión juega un papel importante al momento de que la gerencia (el que toma las decisiones) realice un reporte a partir del modelado siguiente una vez que haya sido implementado; si sabemos que el atributo “Lote” esta contenido dentro de la dimensión “Cajas” y que cada atributo nos sirve para detallar cada evento transcurrido en la cadena de abasto de productos frescos registrado en la tabla de hechos “Eventos”, donde se congregan todos los eventos realizados desde la creación de una caja (donde se le liga el Lote de origen) hasta los envíos por parte del distribuidor hacia los minoristas; en la figura 55 se aprecia el modelado dimensional que resuelve la rastreabilidad de productos frescos del campo.

4.36 Modelo de almacén de datos.

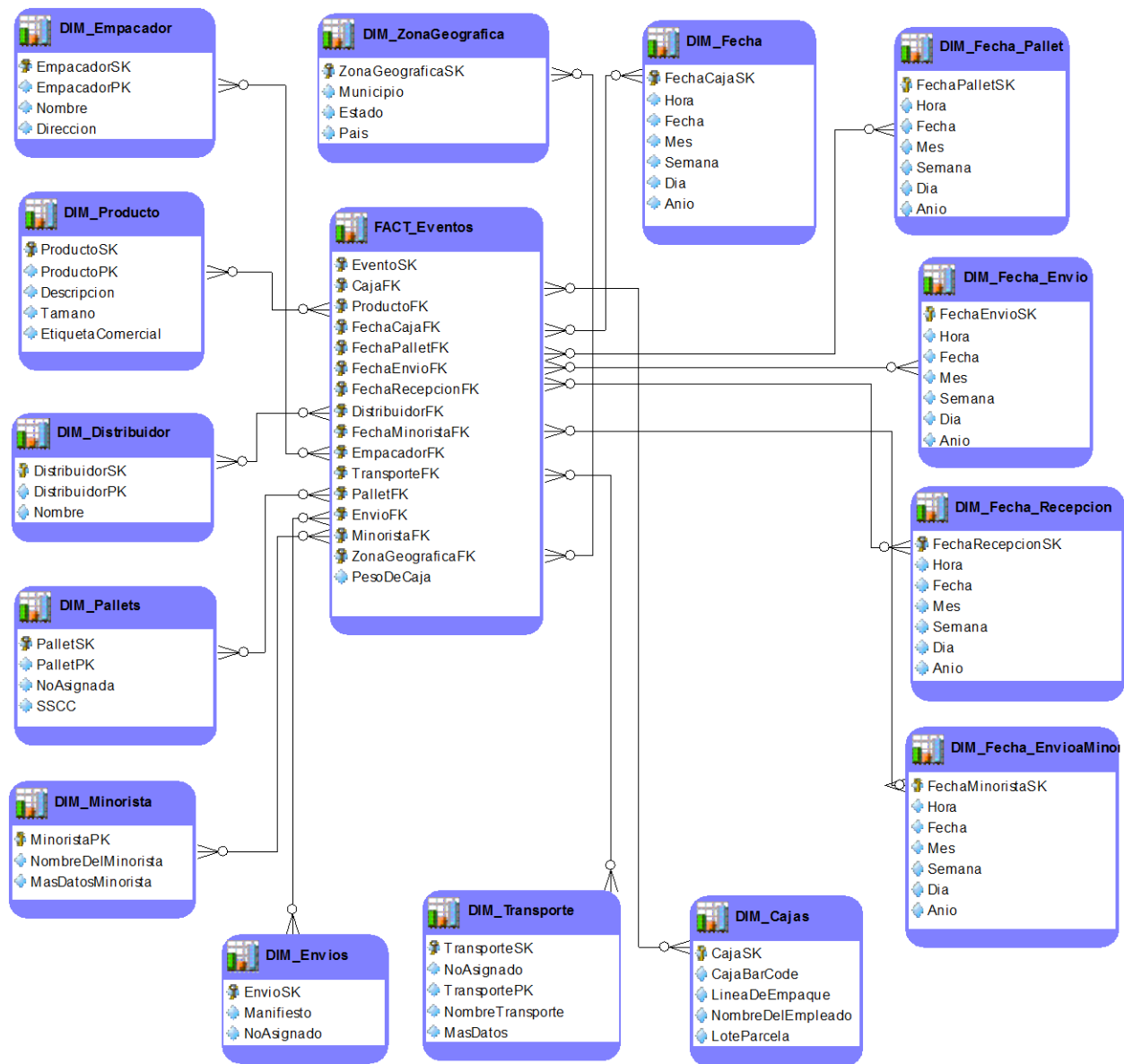


Figure 41 Modelado dimensional, Para la rastreabilidad de productos frescos.

Se hace la sugerencia de tener una cantidad considerable de reportes prediseñados que cumplan con los requerimientos más usados por la gerencia para su administración así como para toma de decisiones, se en lista a continuación una serie de información a recopilar por medio de dichos reportes:

Reportes:

1. RASTREO

- Por lote de producción
- Por lote de origen
- Fecha de producción/envío
- Números de manifiesto (s) y datos del envío(s)
/transportista/caja/tractor/chofer
- Tamaño (en cajas) del lote en cuestión
- Atributos de productos involucrados
 - Tamaños
 - Marcas
 - Numero de cajas por tamaño y marca
- Destinos/transportista/chofer/unidad/fechas y horas

2. INVENTARIO

- Producto
- Atributos
- Marca
- Lote
- Fecha

3. PRODUCCIÓN

Total por:

- Total por lote de origen
- Recepción
- Producción
- Ligar Recepción/Producción
- Producto/Atributos
- Por producto, numérico y porcentual
- Por marca, numérico y porcentual
- Comparativo Cajas y tarimas producidas vs enviadas, numérico y porcentual
- Productividad de empleado

Por tal motivo anexo unos posibles reportes que podrían estar solicitando en el paquete básico de reportes, ya que dada su naturaleza y con el tipo de información que arrojan, lo que es a la gerencia; los ve como indispensables, salvo que a los mismos les desee agregar un dato adicional, los estarían haciendo vía reportes ad-hoc con la herramienta de Pentaho; mismos que se pueden ver en las figuras 55, 56 y 57.

Estadísticas de Producción Empaque Principal									
Fecha Empaque	Para control de pago a destajo :								
02/22/2012									
Empacador	Código de Producto								Empacador
	217 1	217 5	217 8	220 2	220 6	220 8	222 4	222 6	Total general
AIDE CASTRO VALENZUELA	2				3	5	7	2	19
Alvarado Zaragoza Alba Gilberta		1	2	2	5	6			18
AVILA PALMA JESUS JOEL	1		3	5	6	2	4		21
Bobadilla Arredondo Guadalupe Adela	1		2		4		3	1	11
CAMEZ CHAVARRIA ROSALIA		3		3	6				14
Campos Castaño María Luisa		5	3		3	5			20
Domínguez Morales Ana Selene	4	5		8	4		3	4	31
FRASQUILLO CRUZ NORA GRISELDA		3	4			3	1		11
GARCIA MARTINEZ SOFIA		2	3	4	4	3	5		26
HERIBERTO CAMARGO	6		5		5				17
HERRERA FLORES LIZBETH ADRIANA		1		3		3	7	5	26
Jorge Luis Lozano Torres			5	3	8				18
Laguna Pórtela Luz Belén				2					3
LOPEZ RABAGO MARIA DE LOS ANGELEZ	1	4		3	4	5	4	2	28
MARIA JIMENEZ			5	3	8		1		17
Márquez Molina Yuridia Elizabeth		2	3	6		4	5	3	37
Méndez Morales Rosalba		4	5		2	3	4	4	42
MENDOZA SOQUI MARIA LUISA									1
Mijares Pacheco Silvia Raquel	4		2	4		3	3	4	21
Orduña Espinoza Estela		4	6	2		7	5		26
Rodríguez Moza Delia Guadalupe	3		3	5					11
SABORI ECHEVERRIA MARTHA PATRICIA		5		4	5	2	8		27
VALDEZ SOTO JORGE GUADALUPE	1	3	5		2		3		15
VELAZQUEZ MARTINEZ MARIA CRISTINA		6	8	9		3	5	7	45
VERAZ SANCHEZ CLARISA			1						2
VILLARREAL MARRUFO MARIA INES	1	3	5	2	7		3	2	30
ZAVALA ROBLEDO MARIA GRISELDA					1	1			4
OCAMPO ENCIZO LILIA MERCEDES		1	2		3	3	4	4	30
MOLINA BARRON CECILIA	3	5		7		4			20
Total general / Producto	24	56	67	68	66	49	64	36	591

Figure 42 - Estadísticas de producción principal: Para pago de productividad.

Fecha Cosecha	Estadísticas de Espárragos Recibidos en Empaque Principal							
02/21/2012	Entregas de Ranchos Productores por Lote						Total	Total
Rancho Productor	Cajas	Peso	Cajas	Peso	Cajas	Peso	Cajas	Peso
	LT501		LT502		LT503			
El Rebelde	336	2901	1596	13734	72	6400	2004	23035
Lote	JE100		JE101		JE102			
Juan Escutia	144	7356					144	7356
Lote	LB100		LB101		LB102			
La Barras			360	3276	144	1301	504	4577
Lote	LM100		LM101		LM102			
La Macarena	480	9980	222	2000			702	11980
Lote	MD100		MD101		MD102			
María Dolores	72	6500	72	67			144	6567
Lote	SJ2100		SJ2101		SJ2102			
San Joaquín 2			72	6300			72	6300
Lote	SS100		SS101		SS102			
Santa Susana	72	6800	1174	10717			1246	17517
Lote	VC100		VC101		VC102			
Villa Centauro	712	12489	720	6600			1432	19089
Total general	1816	46026	4216	42694	216	7701	6248	96421

Figure 43 - Estadísticas de ranchos productores por Lote.

Fecha	Empaque Principal						Total
Empaque							Empacado
2/21/2012							Por
	Estadísticas por Líneas de Producción						
	Líneas de Producción						
Código de Producto	L-1	L-2	L-3	L-4	L-5	L-6	Producto
2161	5		4			18	27
2164		6			2	5	13
2171		239		36			275
2175		21			19		40
2178	1		27	1		23	52
2195	11		45	4		14	74
2200	19						19
2201	52		6		16		74
2202		44		10		56	110
2204		11			2		13
2227		312		159			471
2228		26	45	21		14	106
2229		234		104		79	417
2230		8		147			155
2330		11	8	9		32	60
2331			23	2			25
2358		1	8	3		8	20
2365		3		11			14
2389			21		3		24
2433		70		45		25	140
2457	1		1				2
Total Por Línea	120	1065	225	646	64	398	

Figure 44 - Estadísticas por Línea de producción.

Aplicando un poco lo que es la herramienta de BI que proporciona Pentaho en la elaboración de informes que de algún modo no hayan sido contemplados en los reportes prediseñados, puede ser dada su naturaleza de que puedan ser poco comunes y se pueda realizar ad-hoc.

Si la granularidad de los eventos son las cajas, no importa la cantidad de líneas contenidas en un envío, mientras se pueda dar respuesta de manera exacta a la solicitud que haga la gerencia en el almacén de datos, y menos ahora que no hay impedimento en cuanto al límite de almacenamiento con los nuevos dispositivos existentes y con la evolución que existe al momento y que con los precios existentes son más accesibles.

Por mencionar un ejemplo, un envío regular contiene de 1800 a 2000 cajas, mismas que se transforman a líneas de registros en la tabla de hechos, y si se habla de que el orden de registros contenidos en una tabla de hechos rondan los millones de ellos; por cada evento de envíos realizado incluyendo desde la elaboración hasta el envío a minorista es mínimo el espacio requerido para tal evento.

En la siguiente figura (58) se puede apreciar un modelo de análisis de información con la herramienta de BI Pentaho, en el cual se muestra la cantidad de cajas producidas por Lote, por Producto, por Mes y Semana.

Lo que hace de Pentaho BI, una herramienta fácil de usar para los usuarios que toman las decisiones sin necesidad de mucho entrenamiento, lo cual incluye también gráficos para un mejor entendimiento de los datos contenidos.

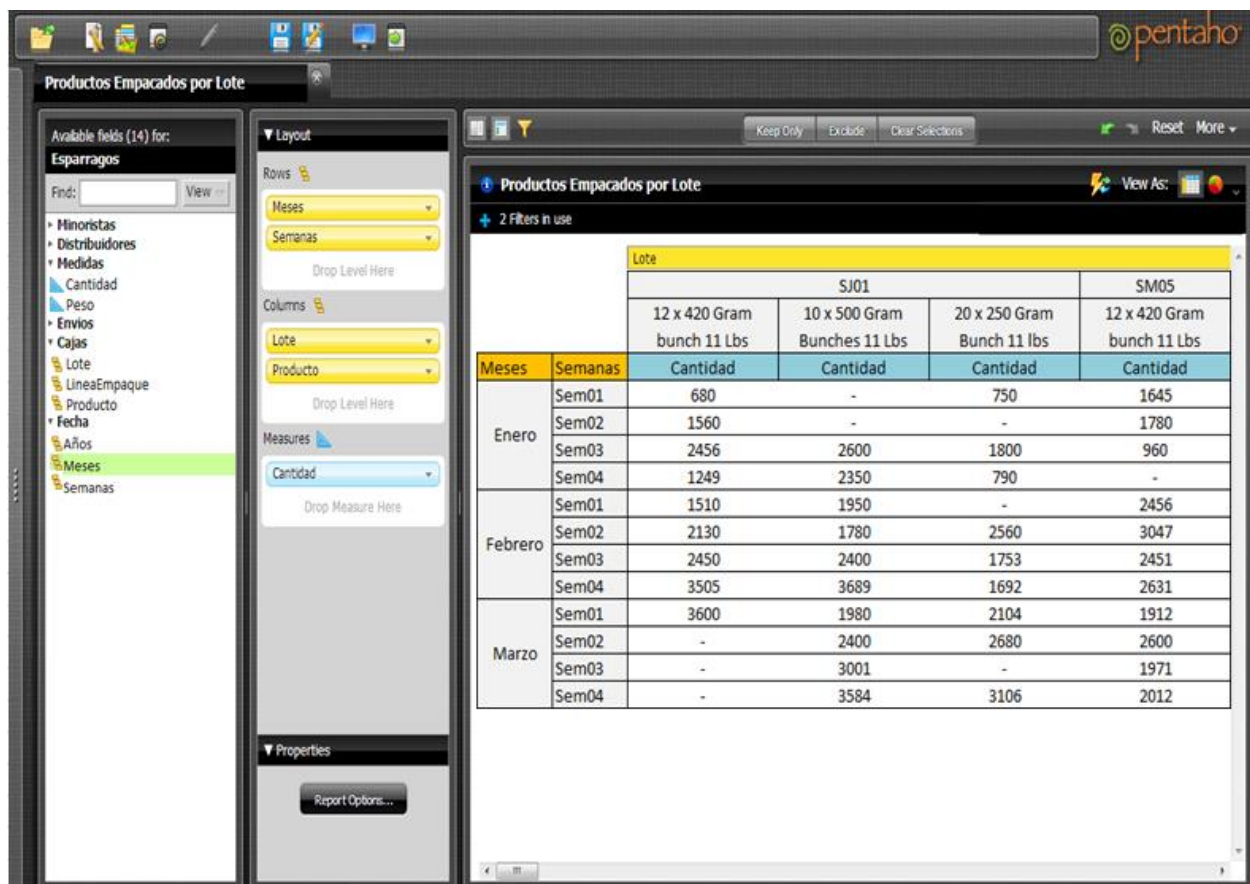


Figure 45 – Modelo de Reporte: Producción de cajas (producto) por lote.

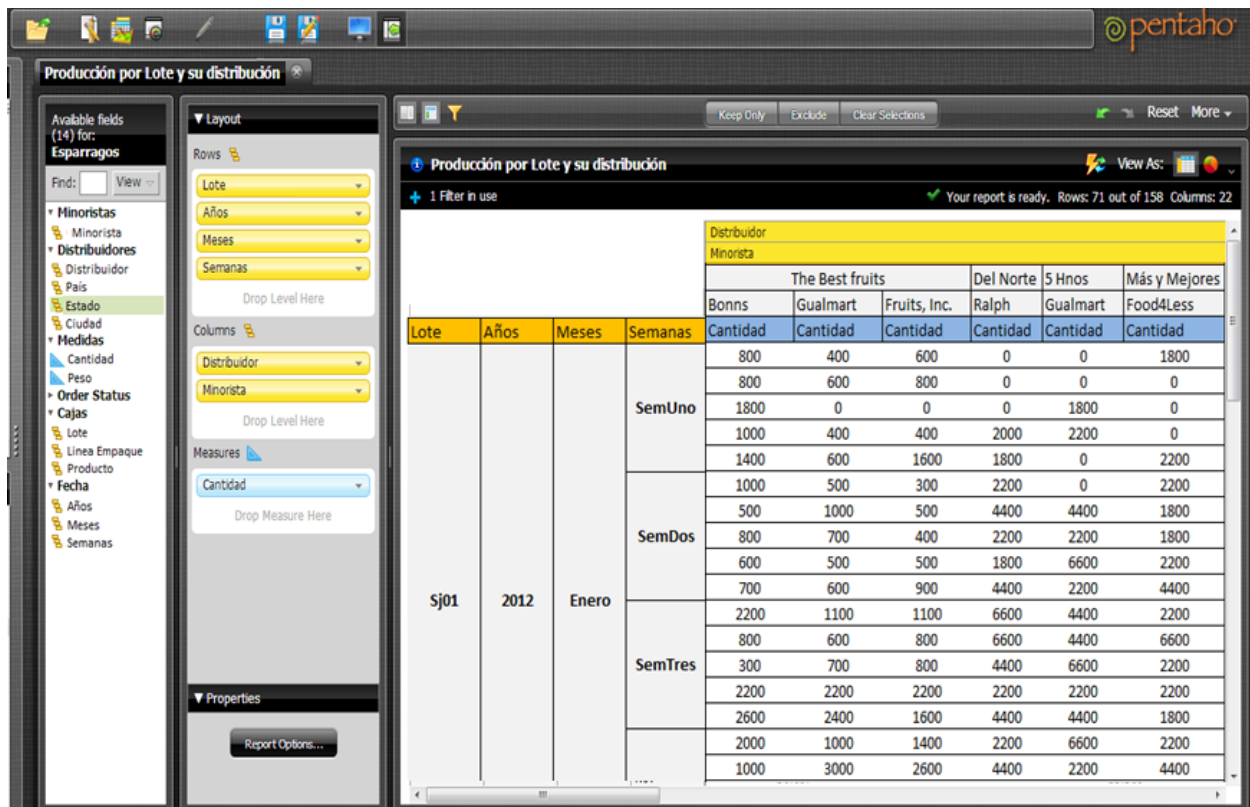


Figure 46 Modelo de Reporte: Cantidad de cajas enviadas a minorista por distribuidor.

Las habilidades del personal que solicita información al almacén de datos vía herramientas inteligentes como lo es pentaho determinará la calidad de la misma; toda empresa deberá contar con capital humano capacitado para llevar a cabo dicha tarea.

Así mismo deberán coexistir reportes prediseñados con información necesaria que se extrae comúnmente, como lo son algunos reportes mostrados anteriormente.

Conclusión

La propuesta de solución de un almacén de datos , aplicando Inteligencia de Negocios (BI), para el caso práctico que aquí se presenta proveerá una respuesta a cada una de las inquietudes presentadas por la administración de la empresa solicitante, la tarea que espera después del modelado presentado no es una tarea sencilla ya que conlleva la cooperación de un gran equipo de trabajo para el desarrollo y la implementación del mismo; sin embargo los beneficios que se pueden obtener con dicha implementación redundará en la confianza que se podrá proporcionar a todos los integrantes que conforman la cadena de suministro , desde el productor hasta el minorista.

Ya que se estaría proporcionando información de gran confiabilidad en tiempo y en forma; por ser un área de gran sensibilidad; el proporcionar confianza al consumidor final de que cada producto que se consume tiene un origen confiable y que en caso de que se pudiese contaminar, se podría identificar rápidamente su origen y que ruta siguió hasta el consumidor final; la inteligencia de negocios es una ventaja competitiva que muchas empresas están aplicando a su propia área de negocio.

El modelar un almacén de datos hace conocer a gran detalle los procesos de negocios y las bases de datos que conforman una empresa, al ser esta una posible solución a la rastreabilidad de productos frescos, deja la posibilidad para que otros procesos se puedan integrar en un futuro.

Líneas Futuras

El modelar una base de datos es una tarea que apenas es el comienzo de una gran tarea que se avecina, ya que al ser desarrollada e implementada dará como resultado la acumulación de una gran cantidad de datos que se podrá además de la inteligencia de negocios también aplicar una herramienta adicional como lo es la minería de datos, que sirve para buscar patrones de comportamiento del mercado, de productos así como también realizar pronósticos de consumo de nuestro clientes, nuestras líneas de producción, de nuestros productores; la empresa tendrá su activo más importante listo para sacar la mayor ventaja del mismo; sus datos organizados de tal manera que le aportarán un plus con respecto a la competencia; se tendrá que seguir invirtiendo en la adecuación de los procesos cambiantes que se originan de la mejora continua.

Glosario

Almacén de Datos. (Del inglés Data warehouse) es una colección de datos que están integrados, no volátiles, variables en el tiempo y que pueden tener diferentes fuentes, están organizados de tal manera que ayudan en la toma de decisiones.

BI. Abreviación de Business Intelligence, Su traducción es Inteligencia de Negocios; y consiste en obtener conocimiento a partir de la información que se recabe de las diferentes bases de datos a través de diversas herramientas.

Bus matrix. Es una herramienta de planeación de proyectos de almacenes de datos.

Business Intelligence. Ver BI

Cubo Multidimensional. También conocido como hipercubo, representa o convierte los datos planos que se encuentran en filas y columnas, en una matriz de N dimensiones.

Data warehouse. Véase Almacén de datos.

Datamarts. Es un almacenamiento de datos especializado en un área específica de negocio; la unión de varios Datamarts conforma un almacén de datos.

Dimensión. Forma parte de un cubo de datos, por medio de sus atributos se le da una perspectiva a los datos contenidos en la tabla de hechos.

DW/BI. Es un sistema que integra un almacén de datos y herramientas de inteligencia de negocios.

Granularidad. Se define como el nivel de detalle de una tabla, si se habla de una jerarquía; la granularidad inicia por la parte más alta de la jerarquía, la granularidad mínima el nivel más bajo.

Llave foránea. Clave que representa a la llave primaria de una tabla externa.

Llave primaria. Clave que representa o identifica a una fila en una tabla.

Llave subrogada. Llave primaria de una tabla que no representa o no tiene significado para el proceso de negocio.

Metadatos. Datos que describen a otros datos

OLAP. Acrónimo del Ingles On-line Analytical Processing (Procesamiento analítico en línea); es una solución que se utiliza en la inteligencia de negocios y su propósito fundamental es agilizar la consulta de grandes volúmenes de datos.

OLTP. Acrónimo del Ingles On-line Transaction Processing (Procesamiento de transacciones en línea), es un sistema que se utiliza para la entrada, procesamiento y recuperación de datos.

Portal BI. Es un sitio de intranet o portal de internet que puede estar enfocado a proporcionar reportes de inteligencia de negocios.

Trazabilidad. Es la habilidad de conocer la posición exacta de un producto a través de la cadena de suministro.

Pallet. Es una estructura de agrupación de carga, fabricada generalmente con madera. La funcionalidad del pallet es el transportar carga, y regularmente tiene forma rectangular o cuadrada.

GTIN. Es el acrónimo de Global Trade Ítem Number (Número mundial de un artículo comercial) Se utiliza para identificar de manera única a cualquier producto.

SSCC. Acrónimo de Serial Shipping Container Code () Es un número de 18 dígitos usado para identificar unidades logísticas.

FDA. Acrónimo de Food and Drug Administration (agencia de drogas y alimentos). Es la agencia del gobierno de los Estados Unidos responsable de la regulación de alimentos (tanto para personas como para animales), suplementos alimenticios, medicamentos (humanos y veterinarios), cosméticos, aparatos médicos (humanos y animales), productos biológicos y derivados sanguíneos.

GS1. Es una organización privada global dedicada a la elaboración y aplicación de normas mundiales y soluciones para mejorar la eficiencia y visibilidad de las cadenas de abastecimiento, la oferta y la demanda a nivel mundial y en todos los sectores.

SAGARPA. Acrónimo de Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación; entre sus funciones están: Organizar y fomentar las investigaciones agrícolas, ganaderas, avícolas, apícolas y silvícolas, estableciendo institutos experimentales, laboratorios, estaciones de cría, semilleros y viveros, vinculándose a las instituciones de educación superior de las localidades que correspondan.

ISESALUD. Acrónimo de Instituto De Servicios De Salud Pública. Organismo gubernamental Mexicano que vela por la salud de la población.

Dumping. En el contexto de las leyes del comercio internacional se define como la práctica en donde una empresa establece un precio inferior para los bienes exportados

que para los costos de producción que tiene la empresa del país adonde se importan esos bienes, sacando competitividad a la empresa local.

Rastreabilidad. Véase trazabilidad.

Lote. En el contexto de este caso práctico se utiliza para identificar a la parcela donde se cultiva la materia prima que da origen a un producto.

CDC. Acrónimo de Centers for Disease Control and Prevention (Centro para el control y prevención de enfermedades). Es una agencia del Departamento de Salud y Servicios Humanos de los Estados Unidos cuya responsabilidad a nivel nacional radica en el desarrollo y la aplicación de la prevención y control de enfermedades, salud ambiental y la realización de actividades de educación y promoción de la salud.

Commodity. Se le llama así a la material prima que da origen a un producto Agrícola, en el contexto de este caso práctico; viene siendo lo que se cosecha en un lote.

Referencias

Apoyo para la elaboración de trabajo terminal: Libro

Bowersox Donald J.; Closs David J. , COOPER M. BIXBY,(2007) **Administración y logística de la cadena de suministros 2 Ed.**

Ralph Kimball; Margy Ross, **The Data Warehouse Toolkit** (Second Edition)(2007)

Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker; **The Data Warehouse LifeCycle Toolkit** (Second Edition)(2007)

Catherine M. Ricardo ,(2009), **Bases de datos.**

Brian Larson,(2008), Delivering **Business Intelligence with Microsoft SQL Server 2008**: 2008 (2nd Revised Edition)

Confederación de Asociaciones Agrícolas del Estado de Sinaloa (CAADES),

Inocuidad en la horticultura, recuperado Diciembre 2010, web site

http://s3.esoft.com.mx/esofthands/include/upload_files/4/Archivos/i20100611174810ForInocuidad2010.pdf

Centro de Investigación en Alimentación y Desarrollo, A.C., Dr. Jorge H. Siller

Cepeda Ponencia sobre horticultura, recuperado Diciembre 2010, web site

http://www.uaaan.mx/academic/Horticultura/Memhort03/Ponencia_10.pdf

Logic Electronics, Rastreabilidad cárnica, recuperado Noviembre 2010, web site

<http://www.logicelectronic.com/productos/FRIGOLOGIC/Descripci%C3%B3n%20Rastreabilidad%20Alimentos%20Carne.html>

GS1 México, Asociación Mexicana de Estándares para el Comercio electrónico,
recuperado Noviembre 2010, Web site <http://www.gs1mexico.org/site/>

GS1 Paraguay, Manual del usuario, recuperado Noviembre 2010, web site
<http://www.gs1py.org/docs/ManualUsuario2009.pdf>

Sinergia e Inteligencia de Negocio S.L., recuperado Diciembre 2010, web site
http://www.sinnexus.com/business_intelligence/datawarehouse.aspx

Universidad del Cauca, Sistemas para Management y Sistemas para Soporte a las
Decisiones, recuperado Febrero 2011, web site
<http://fccea.unicauca.edu.co/old/datawarehouse.htm>

Ultimas Noticias, Quintana Roo
<http://www.qrooultimasnoticias.com/Archivo%20Historico/2008/Junio/Ultimas%20Noticias%2019.06.08%20%28544%29.pdf>; Recuperado Dic. 2010

<http://www.sigweb.cl> , El portal de los expertos en prevención de riesgos de Chile;
recuperado Noviembre 10 2011.

<http://www.sigweb.cl/biblioteca/Trazabilidad.pdf>

<http://inmoncif.com/registration/whitepapers/TEXTUAL%20BUSINESS%20INTELLIGENCE.pdf>

Agile Data Warehouse Design
Lawrence Corr, Jim Stagnitto, Decision Press, *November 2011*

Mastering Data Warehouse Design;

Relational and Dimensional Techniques.

Claudia Imhoff, Nicholas Gallemmo, Jonathan G. Geiger.

Wiley Publishing, Inc. 2003

Modern Software Engineering Methodologies Meet Data Warehouse Design: 4WD

13th International Conference on Data Warehousing and Knowledge Discovery

(DaWaK'11), August 29, 2011

Matteo Golfarelli, Stefano Rizzi, Elisa Turricchia