

**UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA**

**INSTITUTO DE INGENIERÍA**

**MAESTRÍA Y DOCTORADO EN CIENCIAS E INGENIERÍA**



**“ANÁLISIS DE LA RELACIÓN ACTITUD HUMANA – EMISIÓN DE GASES  
REFRIGERANTES A PARTIR DE TÉCNICAS DE CLASIFICACIÓN Y  
REGRESIÓN.”**

**TESIS PARA OBTENER EL GRADO DE:**

**DOCTOR EN INGENIERÍA**

**PRESENTA**

**ELIA IVETTE COTA RIVERA**

**DIRECTOR**

**DR. ALEJANDRO ADOLFO LAMBERT ARISTA**

**CODIRECTOR**

**DR. RAFAEL VILLA ANGULO**

**Mexicali, B. C.**

**Septiembre 2016**

# Contenido

---

Agradecimientos.....	i
Listado de Figuras.....	ii
Listado de Tablas.....	vi
Capítulo I Introducción.....	1
1.1 Antecedentes.....	1
1.2 Objetivo General.....	3
1.3 Objetivos Específicos.....	3
1.4 Fundamentación.....	3
Capítulo 2 Materiales y Métodos.....	6
2.1 Agrupaciones utilizando Análisis de Componentes Principales.....	6
2.2 Las redes neuronales en estudios ambientales.....	6
2.3 Árboles de Regresión (Classification And Regression Trees).....	8
2.3.1 Ventajas y desventajas del algoritmo CART (Salinas & Valcárcel, 2000).....	10
2.3.2. Pasos del análisis CART (Salinas & Valcárcel, 2000).....	11
2.3.3 Reglas de parada (Salinas & Valcárcel, 2000).....	12
Capítulo 3 Desarrollo del Estudio.....	13
3.1 Descripción del área de estudio.....	13
3.2 Muestra.....	14
3.3 Instrumento.....	14
3.3.1 Qué es una actitud y cómo evaluarla.....	14

3.3.2 Construcción de una escala de actitudes tipo Likert.....	16
<b>3.3 Análisis estadístico .....</b>	<b>19</b>
<b>3.4 Análisis por componentes principales .....</b>	<b>20</b>
<b>3.5 Análisis por Redes neuronales .....</b>	<b>20</b>
<b>3.5.1 Predicción de emisión (cantidad en kg) mediante RNA.....</b>	<b>20</b>
3.5.2 Predicción de emisión (SI/NO) mediante RNA .....	21
<b>3.6 Análisis mediante arboles de regresión .....</b>	<b>22</b>
3.6.1 CART: Classification And Regression Trees .....	22
<b>Capítulo 4 Resultados .....</b>	<b>23</b>
<b>4.1 Instrumento.....</b>	<b>23</b>
4.6.1 Análisis descriptivo de la información.....	24
<b>4.2 Análisis de variables obtenidas de encuestas .....</b>	<b>26</b>
4.2.1 Análisis cuantitativo de la emisión.....	47
4.2.2 Caracterización de familias según su comportamiento ambiental .....	59
<b>4.3 Análisis de Componentes Principales .....</b>	<b>64</b>
<b>4.4 Redes Neuronales .....</b>	<b>79</b>
4.4.1 Desempeño de la RNA con todas las variables de la encuesta para estimar la emisión .....	80
<b>4.5 Árboles de Regresión.....</b>	<b>87</b>
<b>4.6 Regresión Lineal.....</b>	<b>90</b>
<b>Capítulo 5 Conclusiones .....</b>	<b>96</b>
<b>Referencias .....</b>	<b>97</b>



## **Agradecimientos**

---

Quiero expresar mi agradecimiento a todas las personas que contribuyeron para finalizar una etapa más, en el largo camino de la educación y por compartir sus conocimientos para el buen desarrollo de este trabajo.

A CONACYT por el financiamiento para la realización de esta investigación. Al Dr. Alejandro Adolfo Lamber Arista, por todos los consejos, experiencia y tiempo dedicado a las correcciones que me hizo, que sin duda sirvieron para enriquecer esta investigación y lograr terminar esta meta.

Al comité revisor por su valiosa aportación al desarrollo de esta investigación, por su tiempo dedicado en la revisión de este trabajo. Por sus recomendaciones, asesorías y consejos para mejorar la investigación.

En especial al Dr. Rafael Villa Angulo, por su invaluable aportación al compartir conmigo su tiempo, sus conocimientos, su buen ánimo, quien siempre tuvo una actitud positiva de apoyo y confianza en esta investigación y en mí, y que sin duda alguna todo esto fue el mayor punto de apoyo para sacar adelante esta investigación y reforzar los resultados obtenidos.

## Listado de Figuras

---

Figura 1 Ubicación geográfica georreferenciada de la ciudad de Mexicali (imagen obtenida de <a href="http://www.bajacalifornia.gob.mx">http://www.bajacalifornia.gob.mx</a> ).....	13
Figura 2 Dimensiones de la actitud .....	15
Figura 3 Histograma de la variable "mantenimiento" .....	26
Figura 4 Histograma de la variable fraccionamiento.....	27
Figura 5 Histograma de la variable "familia" .....	28
Figura 6 Histograma de la variable "escolaridad" .....	29
Figura 7 Histograma de la variable "número de habitantes" .....	30
Figura 8 Histograma de la variable "ingreso" .....	31
Figura 9 Histograma "ítem 1" .....	32
Figura 10 Histograma "ítem 2" .....	33
Figura 11 Histograma "ítem 3" .....	34
Figura 12 Histograma "ítem 4" .....	35
Figura 13 Histograma "ítem 5" .....	36
Figura 14 Histograma "ítem 6" .....	37
Figura 15 Histograma "ítem 7" .....	38
Figura 16 Histograma "ítem 8" .....	39
Figura 17 Histograma "ítem 9" .....	40
Figura 18 Histograma ítem 10.....	41
Figura 19 Histograma ítem 11.....	42
Figura 20 Histograma ítem 12.....	43
Figura 21 Histograma ítem 13.....	44
Figura 22 Histograma ítem 14.....	45

Figura 23 Histograma ítem 15.....	46
Figura 24 Histograma ítem 16.....	47
Figura 25 Emisión de refrigerante por estrato socioeconómico .....	49
Figura 26 Emisión media por Escolaridad .....	50
Figura 27 Emisión media por tipo de familia.....	51
Figura 28 Relación Ingreso mantenimiento .....	52
Figura 29 Relación escolaridad mantenimiento.....	53
Figura 30 Relación tipo de familia y mantenimiento .....	54
Figura 31 Número de habitantes por tipo de familia .....	55
Figura 32 Escolaridad por tipo de familia .....	56
Figura 33 Ingreso por tipo de familia .....	57
Figura 34 Relación entre ingreso y número de habitantes .....	58
Figura 35 Relación entre el grado de estudios y el ingreso .....	59
Figura 36 Caracterización de la muestra por servicio de mantenimiento al aparato de A/C .....	60
Figura 37 Respuestas por grado de estudios.....	62
Figura 38 Respuestas por ingreso .....	63
Figura 39 Respuestas por tipo de familia .....	64
Figura 40 Análisis de Componentes Principales. La figura muestra una clara diferenciación entre familias con fugas clasificadas como Bajas (0, 1, 2 y 3 kg), Medias (4, 5, 6, y 7 kg) y Altas (8, 9 y 10 kg).....	65
Figura 41 Componentes Principales para la Variable Tipo de Familia.....	66
Figura 42 Componentes Principales para la Variable Escolaridad .....	67
Figura 43 Componentes Principales Para la variable Ingreso .....	67
Figura 44 Componentes principales para la pregunta 1.- La contaminación no afecta personalmente mi vida .....	68

Figura 45 Componentes Principales para la pregunta 2.- Antes de comprar los productos que necesito los comparo para seleccionar el que contenga menor cantidad de empaques o envoltorios .....	69
Figura 46 Componentes Principales para la pregunta 3 Incluso si el transporte público fuera más eficiente yo preferiría llevar mi auto al trabajo .....	70
Figura 47 Componentes Principales para la pregunta 4 El gobierno tendrá que introducir duras medidas para frenar la contaminación ya que poca gente las regulara por si misma .....	71
Figura 48 Componentes Principales para la pregunta 5.- No deberíamos preocuparnos por matar demasiados animales de caza porque a la larga las cosas se equilibrarán.....	72
Figura 49 Componentes Principales para la pregunta 6.- Estaría dispuesto a hacer sacrificios personales para reducir el ritmo de la contaminación aunque los resultados inmediatos no puedan parecer significativos.....	72
Figura 50 Componentes Principales para la pregunta 7.- Si el aparato de refrigeración presenta fuga considero necesario mandarla reparar aunque ya se vaya a acabar la temporada de verano.....	73
Figura 51 Componentes Principales para la pregunta 8.- Se debe utilizar el aparato de aire acondicionado solo el tiempo necesario para adquirir un confort térmico u no mantenerlo encendido todo el día.....	74
Figura 52 Componentes Principales para la pregunta 9.- Prefiero comprar un aparato de refrigeración usado ya que su costo es menor.....	74
Figura 53 Componentes Principales para la pregunta 10.- El gas que fuga mi aparato de refrigeración es muy poco como para que impacte en el ambiente .....	75
Figura 54 Componentes Principales para la pregunta 11.-Es una costumbre familiar tener al menos un día de convivencia en familia.....	76
Figura 55 Componentes Principales para la pregunta 12.- La vida diaria es muy atareada y no hay tiempo para actividades familiares. ....	76

Figura 56 Componentes Principales para la pregunta 13.- Cada cuatro meses aproximadamente le doy servicio a mi automóvil para mantenerlo en buenas condiciones.....	77
Figura 57 Componentes Principales para la pregunta 14.- Es preferible tener una televisión en cada recamara para evitar discusiones sobre la programación.....	78
Figura 58 Componentes Principales para la pregunta 15.-Cuando voy a comprar adquiero solo lo que necesito.....	78
Figura 59 Componentes Principales para la pregunta 16.- Para que la ropa sucia no se acumule el fin de semana es preferible lavar diario cantidades pequeñas.....	79
Figura 60 Configuración RNA utilizada .....	80
Figura 61 RNA para estación de la emisión .....	83
Figura 62 Datos reales contra datos predichos por la RNA.....	85
Figura 63 Regresión de entrenamiento, validación y prueba contra datos reales.....	86
Figura 64 Progreso del desempeño de la red .....	87
Figura 65 Árbol generado por el algoritmo CART.....	89
Figura 66 Reglas de decisión para el algoritmo CART .....	90
Figura 67 Distribución de la muestra en las variables seleccionadas del modelo.....	92

## Listado de Tablas

---

Tabla 1 Composición general de la encuesta.....	18
Tabla 2 Formato de la escala tipo Likert.....	19
Tabla 3 Composición de la muestra.....	47
Tabla 4 Matriz de datos estandarizados para la RNA.....	81
Tabla 5 Parámetros de entrenamiento de la RNA.....	82
Tabla 6 tabla general de validación cruzada .....	84
Tabla 7 Coeficiente de relación R para el modelo de datos.....	85
Tabla 8 Fragmento de los datos utilizados para el árbol de decisión.....	88
Tabla 9 Resultados de la regresión lineal múltiple. Los variables de AFE1 a AFE6 corresponden al área afectiva, las variables de COG1 a COG4 corresponden al área cognitiva y las variables de CON1 a CON6 corresponden al área conductual.....	91

## Capítulo I Introducción

---

### 1.1 Antecedentes

Mexicali es una de las ciudades del país en donde se registran las mayores temperaturas durante los meses de mayo a septiembre, con valores promedio por arriba de los 40°C y máximos de hasta 49°C en julio y agosto. En estas condiciones, el uso de aparatos de refrigeración y aire acondicionado es esencial para el confort de sus habitantes, lo que ha conducido a que la actividad de mantenimiento y servicio de estos dispositivos sea de gran importancia (Lambert, 2009)

Los refrigerantes son fluidos que por sus propiedades termodinámicas resultan adecuados para transportar calor, por lo que son usados como sustancias de trabajo en los aparatos de refrigeración y aire acondicionado. Los refrigerantes también contribuyen de forma importante en el efecto invernadero ya que su potencial de calentamiento global es muy alto (varias centenas mayor que el del CO<sub>2</sub>), por lo que en ciudades con clima extremo como Mexicali es muy importante establecer mecanismos que, por un lado, permitan cuantificar las emisiones de estos gases y, por otro lado, ayuden a establecer programas de entrenamiento técnico, y conciencia social para disminuir las fugas o emisiones intencionales. (Villanueva, 2004)

El origen de muchos problemas ambientales puede relacionarse con pautas del comportamiento humano y de la organización social. Puede decirse que los problemas ambientales tienen soluciones comportamentales (Berenguer & Corraliza, 2000)

De todas las soluciones que residen en el comportamiento humano, el estudio de las actitudes y de los procesos de cambio de actitudes ambientales es el que ha generado un buen número de investigaciones y trabajos (Aragonés & Américo, 1991) A continuación mencionaremos algunos de ellos

J. Aragonés y M. Amérigo (Aragonés & Amérigo, 1991) en su artículo Un estudio empírico sobre las actitudes ambientales realizan una validación de la versión castellana de la escala *Environmental Concern Scale* en la cual entre otras cosas se estudia las relaciones entre las actitudes hacia la preocupación por el medio ambiente , la energía nuclear y la ideología política .

A. González y M. Amerito (González & Amérigo, 1999) elaboraron un estudio en el campo de las actitudes ambientales con el fin de indagar en las actitudes pro ambientales y en los constructos de egocentrismo y antropocentrismo y a su vez evaluar la relación que establecen con la conducta ecológica.

Barrigón, Vílches et al (Barrigón, Vílchez, & Gómez, 2002), elaboraron una encuesta para evaluar la molestia producida por el ruido en cualquier entorno urbano, mediante la cual se pretende conocer la importancia relativa de la contaminación acústica así como evaluar las molestias, perturbaciones o efectos que el ruido genera en nuestra vida cotidiana pudiendo diferenciar las principales fuentes que lo causan.

Kalantari, Shabanali (Kalantari & Shabanali, 2007) Realizaron un estudio sobre los factores individuales y sociales que afectan el comportamiento ambiental de los residentes en la ciudad de Teherán, Irán, para esto se creó un marco conceptual para examinar las relaciones entre los factores personales, las actitudes hacia el medio ambiente y el comportamiento ambiental.

Campos, Pasquali et al (Campos, Pascali, & Peinado, 2008) Elaboraron un instrumento de medición de actitudes pro ambientales en escolares venezolanos, instrumento que fue aplicado en el marco de evaluación del programa de reciclaje escolar de dos centros educativos.

## **1.2 Objetivo General**

El objetivo de este estudio es encontrar una técnica para el análisis de emisiones de gases refrigerantes, que asocie información tanto de emisión como de actitud humana, y sea capaz de clasificar a los hogares de la ciudad de Mexicali Baja California.

## **1.3 Objetivos Específicos**

- Crear gráficas del comportamiento de variables de actitud humana y asociarlas con las emisiones de refrigerantes en la ciudad de Mexicali Baja California.
- Analizar cualitativa y cuantitativamente a los usuarios en base a la cantidad de emisiones de refrigerantes presentada en su vivienda.
- Establecer patrones de asociación entre las emisiones y los hábitos socioeconómicos y culturales de las familias evaluadas.
- Evaluar el impacto ambiental asociado a las emisiones de refrigerantes del sector doméstico de la ciudad de Mexicali B.C.

## **1.4 Fundamentación**

En la sociedad moderna los avances de la ciencia y la tecnología han permitido que el hombre pueda satisfacer con facilidad sus necesidades, pero esto ha traído como consecuencia la alteración de las características naturales del medio. El medio ambiente se ha visto modificado por la presencia de sustancias contaminantes que perjudican la vida de las personas y los organismos (Lozano, 2004).

Uno de los mayores problemas ambientales en la actualidad es la alta concentración de gases de efecto invernadero, el exceso de ese tipo de gases hace que la liberación al espacio de energía calórica por irradiación sea menor que la que tendría si se mantuvieran en su nivel

histórico. La consecuencia de este efecto es un cambio climático en el orbe, con el que la temperatura aumenta paulatinamente.

Dentro de los GEI se encuentran los gases refrigerantes que son el objeto principal de nuestra investigación. Los refrigerantes son fluidos que por sus propiedades termodinámicas resultan adecuados para transportar calor.

Los clorofluorcarbonos (CFC) tienen moléculas muy estables con mucha probabilidad de llegar a la estratósfera y actúan como gases invernadero en la tropósfera. Los hidroclorofluorcarbonos (HCFC), sustituyen a los CFC, se descomponen fundamentalmente en la tropósfera y su tiempo de vida es mucho más corto, no obstante, el potencial de calentamiento global de estos gases es muy alto. Pueden ser considerados como una solución temporal, pero no como la solución ideal, por lo tanto todos los CFC plantean problemas de efecto invernadero y han sido incluidos en el Protocolo de Kioto, de modo que su utilización se encuentra también sujeta a restricciones.

La realización de este proyecto es de gran importancia ya que en una ciudad como Mexicali, en la que las altas temperaturas hacen necesario el uso de los aparatos de refrigeración por un periodo largo durante el año, con un abundante consumo de gases refrigerantes es necesario caracterizar los factores que influyen directamente y de manera importante en las emisiones por fugas a partir de los aparatos ya mencionados. De esta manera, se pretende entender la tendencia de las emisiones de gases refrigerantes para hacer estimaciones futuras bajo distintos escenarios, así como crear una herramienta útil para la toma de decisiones, la investigación científica y la identificación de actividades que mitiguen los impactos generados por las emisiones en nuestra ciudad así como la contribución de ésta al clima global.

Como se mencionó anteriormente uno de los objetivos de este estudio es desarrollar perfiles domésticos de una muestra de habitantes en la ciudad de Mexicali y establecer si existe una

relación significativa entre estas características y las emisiones de refrigerantes registradas en los domicilios de las personas encuestadas. Para ello se elaboró una escala multidimensional basada en un esquema teórico previo, a fin de determinar en qué grado el cambio de actitud y conocimiento referente a este problema puede ser determinante para el control o la reducción de emisiones. Esta escala atiende tanto a la actitud hacia problemas ambientales, es decir la predisposición general, como a las conductas de hábito y los conocimientos generales sobre la influencia de los refrigerantes en el medio ambiente.

## Capítulo 2 Materiales y Métodos

---

### 2.1 Agrupaciones utilizando Análisis de Componentes Principales

De manera formal, el Análisis de Componentes Principales (PCA) está definido como una transformación lineal ortogonal, que transforma los datos a un Nuevo Sistema de coordenadas tal, que la mayor varianza para cualquier proyección de los datos yace en la primer coordenada (llamada el primer componente principal), la segunda mayor varianza en la segunda coordenada y así sucesivamente. PCA es en teoría la transformación óptima para un conjunto de datos dado, en términos de mínimos cuadrados. El procedimiento para obtener los componentes principales puede resumirse de la siguiente manera.

Dado un vector  $X^T$  de  $n$  dimensiones,  $X = [x_1, x_2, \dots, x_n]^T$ , del cual se vector de medias  $M$  y covarianza  $C$  son descritos por

$$M = E(X) = [x_1, x_2, \dots, x_n]^T \quad (1)$$

$$C = E[(X-M)(X-M)^T] \quad (2)$$

Calcular los eigenvalores  $\lambda_1, \lambda_2, \dots, \lambda_n$ , y los eigenvectores  $P_1, P_2, \dots, P_n$ ; y ordenarlos de acuerdo a su magnitud

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

Seleccionar  $d$  eigenvectores para representar  $n$  variables,  $d < n$ . Entonces  $P_1, P_2, \dots, P_d$  son llamados componentes principales.

### 2.2 Las redes neuronales en estudios ambientales

Los métodos determinísticos y estocásticos son usados en modelos ambientales, Los modelos determinísticos requieren datos extensos de mecanismos de reacción, *chemical kinetics*, transporte, y parámetros meteorológicos. Las Redes neuronales artificiales han sido

introducidas recientemente como alternativas a los métodos estadísticos convencionales para modelado de contaminantes con las siguientes ventajas: RNA no realizan previas asunciones concernientes a la distribución de los datos; pueden modelar funciones no lineales; pueden ser re entrenados para una mejor generalización cuando estén disponibles datos nuevos o no vistos (Holger & Graeme, 2000).

La unidad de una red neuronal artificial es un procesador elemental llamado neurona que posee la capacidad de calcular, en general, una suma ponderada de sus entradas y luego le aplica una función de activación para obtener una señal que será transmitida a la próxima neurona. Estas neuronas artificiales se agrupan en capas o niveles y poseen un alto grado de conectividad entre ellas, conectividad que es ponderada por los pesos. A través de un algoritmo de aprendizaje supervisado o no supervisado, las RNA ajustan su arquitectura y parámetros de manera de poder minimizar alguna función de error que indique el grado de ajuste a los datos y la capacidad de generalización de las RNA.

F. Inal (Inal, 2010) desarrolló un modelo de red neuronal para predecir el nivel máximo de ozono en Estambul utilizando una red neuronal multicapa de tipo perceptrón utilizando variables meteorológicas y de contaminación y encontró que no existe diferencia significativa entre el modelo de regresión de la red neuronal y las técnicas de regresión para la predicción de las concentraciones de ozono en Estambul. Yi y Prybutok (Yi & Prybutok, 1996) desarrollaron una red neuronal para predecir el nivel diario máximo de ozono usando variables meteorológicas y de contaminación, el cual fue comparado con dos modelos estadísticos tradicionales, de regresión y de box Jenkins ARIMA, los resultados mostraron que el modelo de red neuronal fue superior al de regresión y a box-jenkims ARIMA.

Wang y Lu (Wang & Lu, 2003) desarrollaron un modelo de red neuronal que combina *adaptive radial basis function* con características estadísticas de ozono y lo usaron para

predecir el nivel diario máximo de concentración de ozono en Hong Kong durante 1999 y 2000, dicha simulación demostró efectividad y confiabilidad.

Actualmente el uso de redes neuronales, particularmente redes multicapa que puede ser entrenada para virtualmente cualquier función se ha vuelto popular en estudios ambientales y han producido resultados prometedores tales como en Coman et al (Coman, Ionescu, & Candau, 2008), Chattopadhyay y Chattopadhyay (Chattopadhyay & Chattopadhyay, 2009), Chairez et al. (Chairez, Fuentes, Poznyac, & Franco, 2010), Gardner and Dorling (Gardner & Dorling, 1998), Salcedo et al (Salcedo, Camacho, & Perez, 2011), Kolehmainen et al (Kolehmainen, Martikainen, & Ruuskanen, 2001).

### **2.3 Árboles de Regresión (Classification And Regression Trees)**

Los árboles de clasificación y de regresión, conocidos como algoritmos CART, del inglés classification and regression trees; (Breiman, Friedman, Olshen, & Stone, 1984) constituyen una aproximación multivariada no paramétrica que permite identificar y dimensionar las variables ambientales de mayor impacto en los rendimientos, sin las restricciones que imponen las técnicas de análisis basadas en los modelos lineales clásicos (Lobell, Ortiz, & Asner, 2005). Los modelos CART particionan los datos en forma recursiva de modo tal de conformar subconjuntos cada vez más homogéneos en base a criterios de partición de las variables explicativas. Cada árbol se obtiene a partir de la clasificación de un nodo parental o raíz que contiene la totalidad de los datos, mediante un algoritmo de partición especificado en función de un criterio de partición referido al tamaño del nodo formado o a la variabilidad contenida en los datos del nodo (De'Ath, 2002). Luego se separan dos subconjuntos disjuntos cuya unión comprende el nodo parental (De'Ath & Fabricius, 2000) . Si estos nodos se ven afectados por una nueva partición, se los llama nodos internos; si por el contrario, los datos del nodo tienen

suficiente homogeneidad o bien el tamaño del mismo es suficiente, éste nodo no vuelve a particionarse y recibe el nombre de nodo terminal (Rosales, Bruno, & Balzarini, 2010).

La compañía SPSS (Inc., 1998) menciona que "una de las técnicas más populares en Data Mining es la segmentación basada en árboles. Esta clase de métodos incluyen al CHAID, árboles de regresión y clasificación (CART), y al QUEST. Estos métodos empiezan con todos los registros y los dividen en dos o más subgrupos que tienden a ser más homogéneos que la muestra original. Cada uno de estos subgrupos se divide nuevamente en subgrupos más pequeños y así sucesivamente hasta encontrar una buena solución o cuando no hayan más individuos para continuar"

Bonsón (Bonson, 1999), menciona que "la clasificación supervisada es una de las tareas que más frecuentemente son llevadas a cabo por los denominados Sistemas Inteligentes. Por lo tanto, un gran número de paradigmas desarrollados ya sea por la Estadística (regresión logística, análisis discriminante, K-NN) o bien por la Inteligencia Artificial (redes neuronales, inducción de reglas, árboles de clasificación, redes bayesianas) son capaces de realizar las tareas propias de la clasificación"

Según Johnson (Johnson, 2000) ha sugerido otro enfoque para el análisis discriminante. Puede valer la pena considerar este método si se tiene un gran número de variables discretas. Este procedimiento, del cual puede disponerse en el software CART, desarrolla una serie de preguntas a las que se puede responder sí o no. Las respuestas a estas preguntas llevan al investigador a través de las ramas de un árbol, hasta que llega a un lugar en el que se puede hacer una clasificación. En cada nodo del árbol, el procedimiento encuentra la mejor pregunta para hacer en esa etapa en particular, con el fin de ayudar al investigador a tomar la mejor decisión. Al final de cada rama se toma decisión en cuanto a dónde debe clasificarse una observación particular."

### 2.3.1 Ventajas y desventajas del algoritmo CART (Salinas & Valcárcel, 2000)

- Es un método bastante flexible.
- Permite considerar los costes de clasificación errónea en el proceso de desarrollo del árbol.
- Permite asignar la distribución de probabilidades a priori en los problemas de clasificación.
- CART no hace asunciones sobre la distribución de ningún tipo, ya sea de las variables predictoras como de la variable criterio.
- Las variables pueden ser de diferentes tipos como continuas, discretas, categóricas o nominales
- CART no está afectado por valores extremos, colinealidad, heterocedasticidad que afecten los procedimientos establecidos. Los valores outliers pueden ser aislados en un nodo y no tienen ningún efecto en la división.
- CART puede detectar y revelar interacciones en un conjunto de datos
- CART es invariante con relación a la transformación de las variables independientes. Es decir, las transformaciones no afectan el árbol producido.
- CART con el conjunto de variables predictoras o independientes, produce resultados usando sólo las variables más importantes.

No obstante, CART presenta algunos inconvenientes. Entre sus principales desventajas se tienen (Salinas & Valcárcel, 2000):

- Al ser un algoritmo binario, tiende a generar árboles de muchos niveles. Por ello, el árbol resultante puede que no presente los resultados de manera eficiente, sobre todo si la misma variable ha sido utilizada para la división de varios niveles consecutivos.
- También tiende a seleccionar las variables que pueden permitir un mayor número de divisiones en el proceso de crecimiento del árbol. Por tanto, las conclusiones extraídas de las estructuras de árbol pueden resultar poco fiables.
- CART es un método muy complejo. Los cálculos pueden ser muy lentos cuando se utilizan conjuntos de datos voluminosos. CART utiliza la división por substitutos para tratar los valores perdidos.
- CART no está basado en un modelo probabilístico. No existen intervalos de confianza asociados con las predicciones derivadas usando el algoritmo CART para clasificar a un conjunto de datos.

### 2.3.2. Pasos del análisis CART (Salinas & Valcárcel, 2000)

1. Para llevar a cabo un análisis CART, comenzando por el nodo  $t=1$ , buscar la división  $s^*$  de entre todos los candidatos posibles  $S$  que dé lugar a la mayor reducción de impureza.

$$\Phi(s^*, 1) = \max(\phi_{s \in S}(s, 1)) \quad (3)$$

Luego dividir el nodo 1 ( $t=1$ ) en dos nodos,  $t=2$  y  $t=3$ , utilizando la división  $s^*$

2. Repetir el proceso de búsqueda de divisiones para uno de los nodos  $t=2$  y  $t=3$ , y así sucesivamente.
3. Continuar el proceso de desarrollo del árbol hasta alcanzar al menos una de las reglas de parada

### 2.3.3 Reglas de parada (Salinas & Valcárcel, 2000)

Se detiene el proceso de desarrollo del árbol cuando se cumple una de las diversas reglas de parada disponibles. Un nodo no se dividirá si se cumplen alguna de las siguientes condiciones:

- Todos los casos de un nodo tienen valores idénticos en todos los predictores.
- El nodo se vuelve puro; esto es, todos sus casos tienen el mismo valor en la variable criterio.
- La profundidad del árbol ha alcanzado el valor máximo preestablecido.
- El número de casos que constituyen el nodo es menor que el tamaño mínimo preestablecido para un nodo parental.
- La división del nodo tiene como resultado un nodo filial cuyo número de casos es menor que el tamaño mínimo preestablecido para un nodo filial.
- La reducción máxima de la impureza es menor que el valor preestablecido  $\beta$

## Capítulo 3 Desarrollo del Estudio

---

### 3.1 Descripción del área de estudio

La ciudad de Mexicali (ver Figura 1 ) se localiza a lo largo de la frontera internacional a los 32° 40' de latitud norte y 115° 28' de longitud oeste, a una altura promedio sobre el nivel del mar de 10 m. La región forma parte de la cuenca del río Colorado (Castro, Contreras, & González, 1999). Colinda al norte con Estados Unidos de América; al este con el estado de Sonora y el Golfo de California; al sur con el municipio de Ensenada; y al oeste con los municipios de Ensenada y Tecate (INEGI, 2011)]. Cuenta con una población de 855,962 habitantes (INEGI, 2005). Ocupa en la actualidad una extensión de aproximadamente 11,372 hectáreas a las que se agregan 795 hectáreas periféricas con carácter rural. En el año de 1996 contaba con una reserva territorial de 504 hectáreas para su crecimiento urbano (Zuk, Guadalupe, Cervantes, & Rojas, 2001).



Figura 1 Ubicación geográfica georreferenciada de la ciudad de Mexicali (imagen obtenida de <http://www.bajacalifornia.gob.mx>).

### **3.2 Muestra**

Se realizó un total de 321 encuestas en los meses de septiembre 2011 de abril de 2012, a la población de Mexicali arriba de los 18 años de edad y a quien se consideró como representante de la familia entrevistada, el cuestionario fue aplicado de dos maneras, se creó un instrumento digital que se aplicó por medio electrónico y se aplicaron encuestas personales de manera aleatoria en distintos puntos de la ciudad.

### **3.3 Instrumento**

Para realizar este estudio fue necesario generar un instrumento que permita medir el grado de conocimiento y conciencia ambiental de una muestra de población y determinar la influencia de estos factores en las emisiones de refrigerantes con el fin de predecir en qué grado el cambio de actitud y conocimiento referente a este problema puede ser determinante para el control o la reducción de emisiones.

#### **3.3.1 Qué es una actitud y cómo evaluarla**

Actitud es un estado de disposición psicológica, adquirida y organizada a través de la propia experiencia que incita al individuo a reaccionar de una manera característica frente a determinadas personas, objetos o situaciones (Morales, 2006).

Las actitudes no son susceptibles de observación directa sino que han de ser inferidas de las expresiones verbales o de la conducta observada. Esta medición indirecta se realiza por medio de escalas en las que, partiendo de una serie de afirmaciones, proposiciones o juicios sobre los que los individuos manifiestan su opinión, se deducen o infieren las actitudes (Fernandez de Pinedo, 2007).

En el 2003, Fernández y colaboradores (Fernandez & Hueto, 2003) mencionan que para evaluar actitudes se deben considerar 3 elementos (Figura 2)

- Los conocimientos o creencias sobre el tema; llamado el aspecto cognitivo de la actitud
- La disposición (favorable o desfavorable), actuar en una dirección determinada; conocido como aspecto afectivo de la actitud.
- La conducta de hecho ante una situación determinada; denominada el aspecto conativo de la actitud.

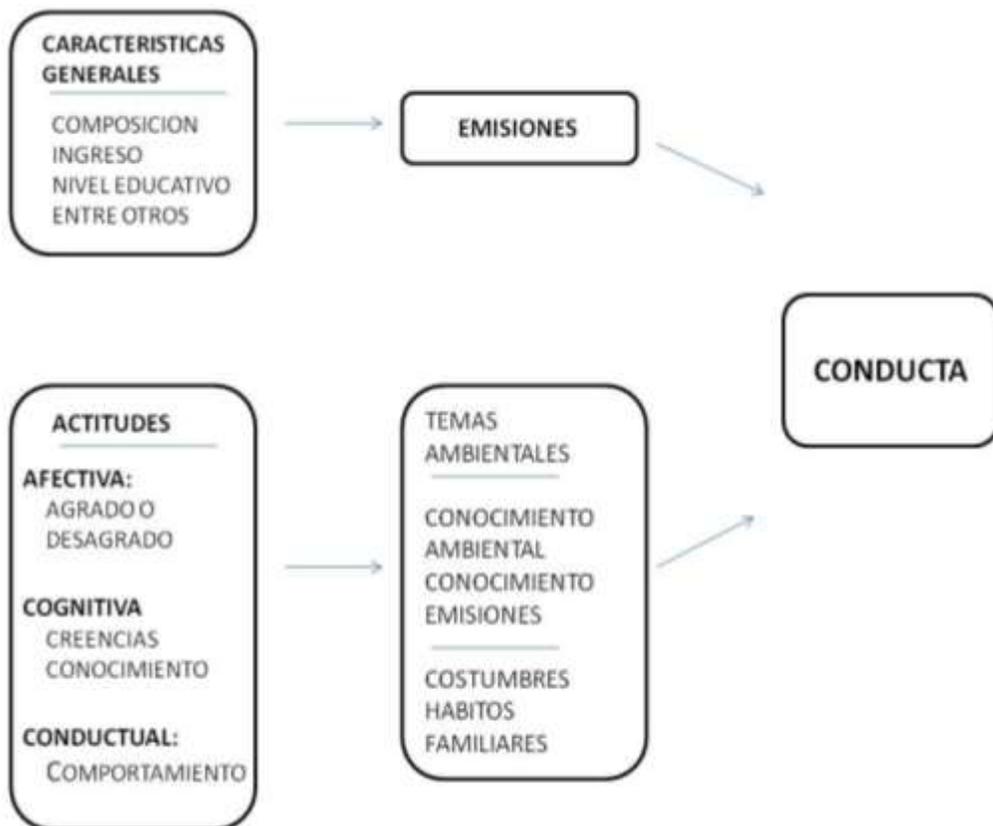


Figura 2 Dimensiones de la actitud

### **3.3.2 Construcción de una escala de actitudes tipo Likert.**

Definimos una escala como una serie de ítems o frases que han sido cuidadosamente seleccionados, de forma que constituyan un criterio válido, fiable y preciso para medir de alguna forma los fenómenos sociales. En nuestro caso, este fenómeno será una actitud cuya intensidad queremos medir. (Fernandez de Pinedo, 2007).

La construcción de una escala aditiva tipo Likert consta de los siguientes pasos:

1. Se recoge una larga serie de ítems relacionados con la actitud que queremos medir y se seleccionan, aquellos que expresan una posición claramente favorable o desfavorable.
2. Se selecciona un grupo de sujetos similar a aquél al que piensa aplicarse la escala. Estos responden, eligiendo en cada ítem la alternativa que mejor describa su posición personal
3. Las respuestas a cada ítem reciben puntuaciones más altas cuanto más favorables son a la actitud, dándose a cada sujeto la suma total de las puntuaciones obtenidas.
4. Para asegurar la precisión de la escala, se selecciona al 25% de los sujetos con puntuación más alta y al 25% con puntuación más baja, y se seleccionan los ítems que discriminan a los sujetos de estos dos grupos, es decir, aquellos con mayor diferencia de puntuaciones medias entre ambos grupos.
5. Para asegurar la fiabilidad por consistencia interna, se halla la correlación entre la puntuación total y la puntuación de cada ítem para todos los individuos, seleccionándose los ítems con coeficiente más alto.

6. Con los criterios anteriores de precisión y fiabilidad se selecciona el número de ítems deseado para la escala. Para asegurar la validez del contenido, aproximadamente la mitad de los ítems deben expresar posición favorable y desfavorable la otra mitad.

En ciertos casos, para obligar a los individuos a definir su posición de forma más clara, se puede suprimir la respuesta "indeciso". El número de ítems de una escala Likert suele oscilar entre 15 y 30.

Para la recopilación de la información correspondiente al perfil de familia se utilizó como procedimiento de investigación: la encuesta, instrumento que permite obtener y elaborar datos de manera rápida y eficaz.

- Para conseguir los objetivos propuestos se diseñó un cuestionario auto administrado que contiene varios apartados:
- El primer apartado estará compuesto por datos generales de la familia tales como, estructura de la familia, número de miembros, escolaridad (que será un estimado del promedio de los padres) y nivel socioeconómico.
- En el segundo apartado se incluirán ítems relacionados con la parte afectiva, es decir, de agrado o desagrado respecto al tema estudiado y en general a temas sobre el medio ambiente.
- El tercer apartado contendrá ítems del área cognitiva es decir, qué tanto conoce el sujeto entrevistado sobre la problemática que resulta de la falta de control en las emisiones de los gases refrigerantes de los aparatos de aire acondicionado de su hogar, y en qué medida está comprometido a colaborar con el buen uso y mantenimiento adecuado de dichos aparatos.

- En el cuarto apartado se evalúa el aspecto conductual y de costumbres de la familia, con el fin de establecer una relación entre los hábitos y la cantidad de emisiones que aporta esa familia al ambiente (Tabla 1).

**Tabla 1 Composición general de la encuesta**

<b>Apartado</b>	<b>Contenido</b>	<b>Preguntas</b>
Primer apartado	Información general	A-D
Segundo apartado	Área afectiva	1-6
Tercer apartado	Área cognitiva	7-10
Cuarto apartado	Área conductual	11-16

En este cuestionario se incluyó parte de la Environmental Concern Scale (ECS) de Weigel y Weigel (Weigel & Weigel, 1978). cuya adaptación ya fue utilizada en trabajos anteriores como el de Aragonés y Amérigo (Aragonés & Amérigo, 1991) en el que se tienen 16 ítems de tipo Likert de 5 puntos, desde muy en acuerdo hasta muy en desacuerdo. Esta escala es una medida tradicional de preocupación ambiental y versa desde la conservación y contaminación del medio ambiente, temas sustantivos e indicativos para medir este constructo según Van Lier y Dunlap (Van Lier & Dunlap, 1981).

Para este estudio y con el fin de reducir la extensión del instrumento para comodidad del sujeto entrevistado varios ítems fueron ligeramente modificados con el fin de facilitar la comprensión del entrevistado así como orientar las preguntas al tema principal de este estudio.

En los apartados se incluyen ítems que evalúan también el grado de apatía de las personas encuestadas hacia el medio ambiente expresando falta de interés general hacia temas ambientales y creencia de que estos temas se han exagerado por los medios de comunicación. Todos ítems de estos apartados a excepción del primero tienen un formato tipo Likert de 5 puntos que van desde totalmente de acuerdo a totalmente en desacuerdo (Tabla 2).

**Tabla 2 Formato de la escala tipo Likert**

Pregunta	TD	D	I	A	TA
<b>1.- Cuando voy a comprar solo adquiero las cosas que realmente necesito</b>	1	2	3	4	5
<b>2.- Antes de comprar los productos que necesito los comparo para seleccionar el que tenga menos empaques y envolturas.</b>	1	2	3	4	5
<b>3.- Al comprar bebidas prefiero consumir en envases no retornables</b>	5	4	3	2	1
<b>4.- Compro cosas porque me gustan</b>	5	4	3	2	1

TD= Totalmente desacuerdo D= Desacuerdo I= Indeciso A= Acuerdo TA= Totalmente de acuerdo

Como paso previo a su distribución entre la población, y para corregir algunas posibles lagunas y defectos en las preguntas elaboradas, se llevó a cabo una prueba piloto con 30 personas seleccionadas de manera aleatoria, de esa forma se consiguió depurar la encuesta y corregir algunas dudas que surgían entre los encuestados ante el planteamiento inicial de algunas cuestiones.

La confiabilidad general del cuestionario fue medida a través del índice de consistencia interna Alfa de Chronbach ( $\alpha$ ) que arrojó un valor de .609 para el instrumento, demostrando ser éste un instrumento de medición confiable (Moreno, Corraliza, & Ruiz, 2005)

### **3.3 Análisis estadístico**

Como primer paso se realizó un análisis con estadística descriptiva. Esta se empleó para describir, analizar y representar la muestra de datos obtenidos, utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos.

Para este análisis se usó la información de la base de datos obtenida de las encuestas; los gráficos elaborados fueron para las variables mantenimiento, fraccionamiento, tipo de familia,

escolaridad y para cada uno de los ítems que componen la escala de actitud y conciencia ambiental.

### **3.4 Análisis por componentes principales**

Se realizó un análisis utilizando el método de componentes principales con el fin de establecer la estratificación de los datos y determinar la presencia de cúmulos en relación a las variables de la encuesta con el fin de determinar si se pueden clasificar las respuestas de las personas encuestadas ya sea por estatus socioeconómico, tipo de familia, escolaridad, número de habitantes en la vivienda.

Para realizar este análisis se utilizó el lenguaje R, el análisis se realizó a toda la encuesta y se obtuvo el primer componente principal, una vez hecho esto se ordenó de menor a mayor según ese primer componente.

### **3.5 Análisis por Redes neuronales**

Para este análisis se utilizó JavaNeural Network Simulator (JavaNNS) el cual es un simulador de redes neuronales desarrollado en Wilhelm-Schickard-Institute for Computer Science (WSI) en Tübingen, Alemania.

#### **3.5.1 Predicción de emisión (cantidad en kg) mediante RNA**

El objetivo de este análisis fue probar el gado de predicción que puede hacer una RN, tomando como entrada las características de una persona encuestada y sus respuestas a la encuesta Likert. Para este procedimiento se de la matriz de datos aquellos registros que tuvieran un “NO” como respuesta a la pregunta: “se realizó mantenimiento al equipo de aire acondicionado?”.

Una vez completa la matriz de datos a utilizar se escalaron los datos de entrada entre 0 y 1 y se generaron salidas binarias de la siguiente manera:

0 0 0 1 para las emisiones con valor de 0

0 0 1 0 para emisiones mayores que 0 y menores que 3kg

0 1 0 0 para las emisiones desde 3kg hasta 5kg

1 0 0 0 para las emisiones mayores de 5 kg

Se creó una red neuronal de 20 capas de entrada, 4 de salida y 300 capas internas, y se utilizó la función de aprendizaje backpropagation.

Una vez configurada la red se procedió al aprendizaje con los datos escalados, el conjunto de datos se dividió en 3 archivos: entrenamiento prueba y validación los cuales contenían el 75, 15 y 15% de los datos.

Una vez entrenada la red y validada se realizó un proceso de validación cruzada (crossvalidation) de 10 repeticiones para poder estimar un error promedio y de esta manera validar el modelo generado.

### **3.5.2 Predicción de emisión (SI/NO) mediante RNA**

El objetivo de este análisis fue predecir si habrá fuga (emisión) en alguno de los aparatos de A/C dependiendo de las características de la persona encuestada y sus respuestas en la encuesta likert, Para este procedimiento se eliminó de la matriz de datos aquellos registros que tuvieran un “NO” como respuesta a la pregunta: “se realizó mantenimiento al equipo de aire acondicionado?”, siendo un total de 281 registros utilizados para la predicción.

Una vez completa la matriz de datos a utilizar se escalaron los datos de entrada entre 0 y 1 y se generaron salidas binarias de la siguiente kg manera:

0 1 para cuando no hubo emisión

1 0 para cuando si hubo emisión

Se creó una red neuronal de 20 capas de entrada, 4 de salida y 300 capas internas, y se utilizó la función de aprendizaje backpropagation.

Una vez configurada la red se procedió al aprendizaje con los datos escalados, el conjunto de datos se dividió en 3 archivos: entrenamiento prueba y validación los cuales contenían el 75, 15 y 15% de los datos.

Una vez entrenada la red y validada se realizó un proceso de validación cruzada (crossvalidation) de 10 repeticiones para poder estimar un error promedio y de esta manera validar el modelo generado.

### **3.6 Análisis mediante arboles de regresión**

#### **3.6.1 CART: Classification And Regression Trees**

El algoritmo CART (árboles de clasificación y regresión) se aplica en situaciones donde se tienen un conjunto de datos de individuos en los que se han medido variables predictoras o independientes y una variable de clasificación o de criterio que define el grupo al que cada individuo pertenece. Luego, se quiere encontrar un conjunto de reglas de decisión que permitan explicar la clasificación existente y utilizar estas reglas para clasificar a un nuevo individuo (Salinas & Valcárcel, 2000).

## Capítulo 4 Resultados

---

### 4.1 Instrumento

El resultado de esta etapa de la investigación fue la creación del instrumento por medio del cual se lleva a cabo la encuesta.

El cuestionario resultante consta de 4 preguntas de opción múltiple para determinar las características generales de la familia, y 16 ítems a ser evaluados en escala de Likert, donde 50% de los ítems denotan afirmaciones con carácter negativo y el otro 50% de carácter positivo, y la puntuación de las respuestas se realiza de la siguiente manera:

Ejemplo 1.

Pregunta:

#### 1.- La contaminación no afecta personalmente mi vida

Respuestas a elegir:

TA    A    I    D    TD

Ejemplo 2.

Pregunta:

#### 2.- Cuando voy a comprar adquiero únicamente lo que necesito

Respuestas a elegir:

TA    A    I    D    TD

**Donde:**

**TA** = Totalmente de Acuerdo

**A** = de Acuerdo

**I** = Indiferente

**D** = Desacuerdo

**TD** = Totalmente en Desacuerdo

En los dos ejemplos reproducidos, la valoración de las respuestas sería 1, 2, 3, 4, 5, en el primero y 5, 4, 3, 2, 1, en el segundo.

Un individuo que conteste a estos ítems en las dos respuestas marcadas con el círculo tendría una puntuación de 2 puntos en el primer ítem y de 1 punto en el segundo, su suma, 3, sería la posición de este individuo. Si la escala estuviera formada por estos dos ítems solamente, la puntuación de los individuos iría desde 2, actitud más negativa hacia la formación, hasta 10, actitud más positiva.

El individuo que obtiene una puntuación de 3, en principio parece que tiene una actitud baja o en contra del fenómeno que se mide, pero para poder decir esto hay que esperar a compararlo con la puntuación obtenida en la misma escala, por otros individuos.

#### **4.6.1 Análisis descriptivo de la información**

En total se aplicó la encuesta a 321 hogares, ubicados en 137 colonias, cuya última reparación de su sistema de aire acondicionado fue en el verano de 2011 (junio-agosto). La **Error! Reference source not found.** presenta la distribución de los hogares estudiados, en función de las principales características consideradas en el estudio. De las 321 familias entrevistadas 281 de estas si realizaron servicio a sus aparatos de aire acondicionado y 40 no lo hicieron, en esos 281 hogares en los que si se realizó el servicio a su aparato de aire acondicionado 106 no presentaron fugas en sus aparatos y 175 si presentaron, es decir, 175 hogares emitieron

gas refrigerante a la atmosfera. Otra categoría a evaluar es el tipo de familia; como resultado de esto: 219 familias fueron del tipo nuclear, es decir, que la familia consta de ambos padres y los hijos, 55 familias fueron del tipo monoparental, es decir que la estructura familiar solo cuenta con uno de los padres de familia, y 46 familias resultaron ser de tipo extendida, es decir la estructura familiar consta de uno o ambos padres, hijos y personas adicionales al núcleo familiar, como pueden ser abuelos, tíos y primos, entre otros. En cuanto a la escolaridad de los hogares encuestados, del total de la muestra se tomó el nivel de escolaridad más alto que tuviera el o los padres de familia, de esto se obtuvo que 12 personas tienen como máximo grado de estudios el nivel primaria, en 31 familias secundaria, 48 preparatoria, 171 licenciatura y 59 posgrado. Otra clasificación que se utilizó para la encuesta es el nivel socioeconómico, cuya distribución es la siguiente: 57 familias tienen un nivel socioeconómico alto, 165 nivel medio y 57 bajo nivel de ingresos (la correspondencia de nivel socioeconómico se tomó de acuerdo a sus nivel de ingreso en salarios mínimos, como sigue: 5 o menos salarios mínimos mensuales corresponde a nivel bajo, entre 5 y 10 salarios mínimos mensuales corresponde a nivel medio y arriba de 10 salarios mínimos mensuales corresponde a nivel alto).

Con el propósito de hacer una estimación general de la cantidad de gas refrigerante emitido a la atmosfera por los hogares de Mexicali realizamos el siguiente análisis: Partiendo de 321 familias encuestadas, en 137 colonias, podemos estimar la cantidad de familias por colonia dividiendo ambas cantidades, lo que nos arroja de un total de 2.34 familias por colonia. Ahora, la emisión máxima por familia encuestada fue de 10 kg y la mínima es 0 kg, que nos da como resultado un promedio de 1.89 kg de gas refrigerante emitidos a la atmosfera por familia en una temporada. Tomando en cuenta que el número de viviendas en Mexicali es de 218,912 (en el año 2011) y suponiendo que la distribución de la muestra es homogénea podemos estimar que aproximadamente 413,743 kg de gas refrigerante es emitido a la atmosfera por las familias de Mexicali, en una temporada de verano.

#### 4.2 Análisis de variables obtenidas de encuestas

El gráfico que se presenta a continuación (ver Figura 3 **Error! Reference source not found.**) nos indica el número de personas del total de la muestra que dieron servicio de mantenimiento a su aparato de refrigeración y a las personas que no lo hicieron, podemos observar que de un total de 321 casos, el 87.5 % sí realizó el mantenimiento a su aparato de A/C, y tan solo el 12.5% no lo realizo.

En el histograma presenta una curva asimétrica con un sesgo positivo, dando como resultado una distribución leptocúrtica de 3.2.

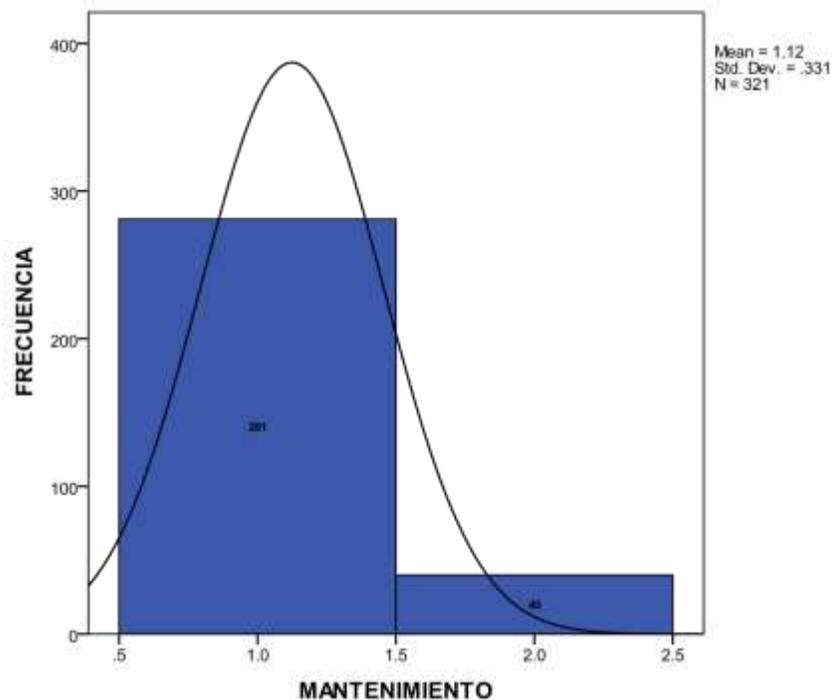
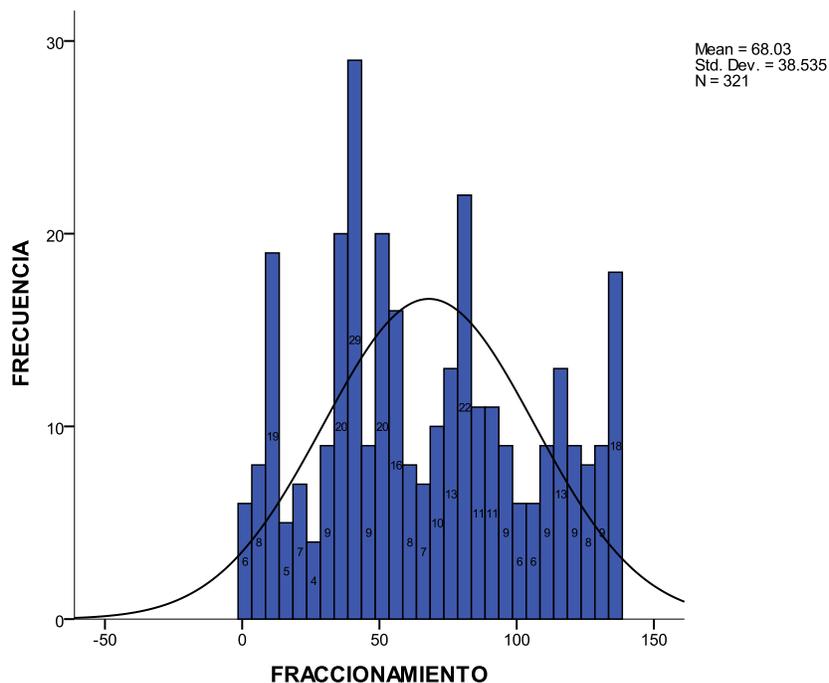


Figura 3 Histograma de la variable “mantenimiento”

La **Error! Reference source not found.** nos indica el número de fraccionamientos distintos donde viven las familias encuestadas, del total de la muestra encuestada fueron 137 colonias en total 321 casos, la colonia que observamos con mayor frecuencia fue la colonia Independencia con un total del 6.2% del total de la muestra, seguida por la colonia prohogar que abarca un 4% de la población encuestada.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de -1.37



**Figura 4 Histograma de la variable fraccionamiento**

La Figura 4 **Error! Reference source not found.** nos indica cómo está distribuida nuestra muestra según el tipo de familias que la componen, como se mencionó anteriormente los tipos de familia manejados en este estudio son, nuclear (familias en las que viven dos padres y los

hijos), monoparental (familias en la que solo existe un padre como cabeza de la familia) y extendida (familias en las que aparte de la familia principal radican otros familiares u otras personas) en las familias encuestadas, del total de la muestra encuestada el 68% pertenece al tipo de familia nuclear, el 55% al tipo monoparental y el 47% a tipo extendido.

En el histograma presenta una curva asimétrica con un sesgo positivo, dando como resultado una distribución leptocúrtica de -0.048.

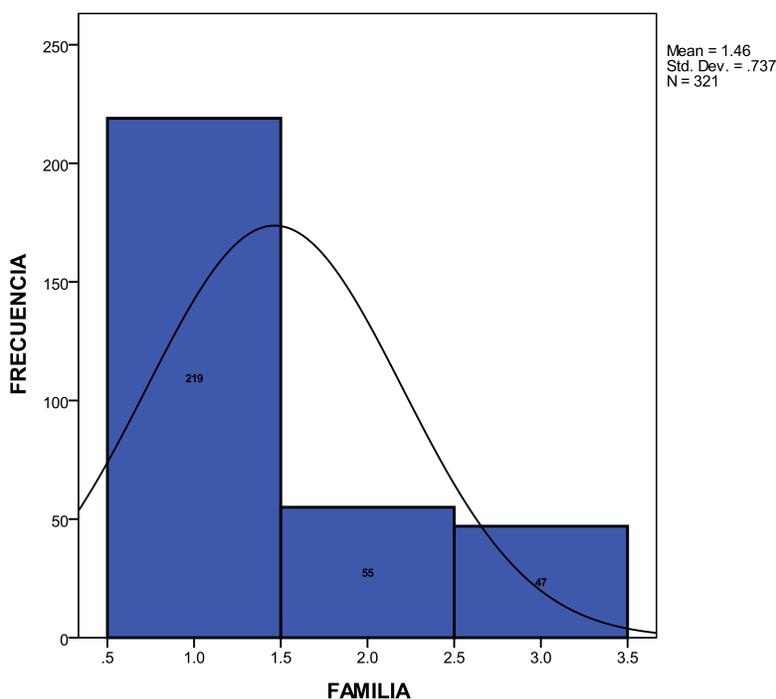


Figura 5 Histograma de la variable "familia"

El gráfico siguiente (ver Figura 5 **Error! Reference source not found.**) nos indica la distribución de la muestra según el grado de escolaridad mayor en la familia encuestada del total de 321 casos, el 53% de la población de familias encuestadas tiene nivel de licenciatura, el 18% nivel de posgrado, el 15% nivel de preparatoria, y el 3.7% nivel primaria

En el histograma presenta una curva asimétrica con un sesgo positivo, dando como resultado una distribución leptocúrtica de .631.

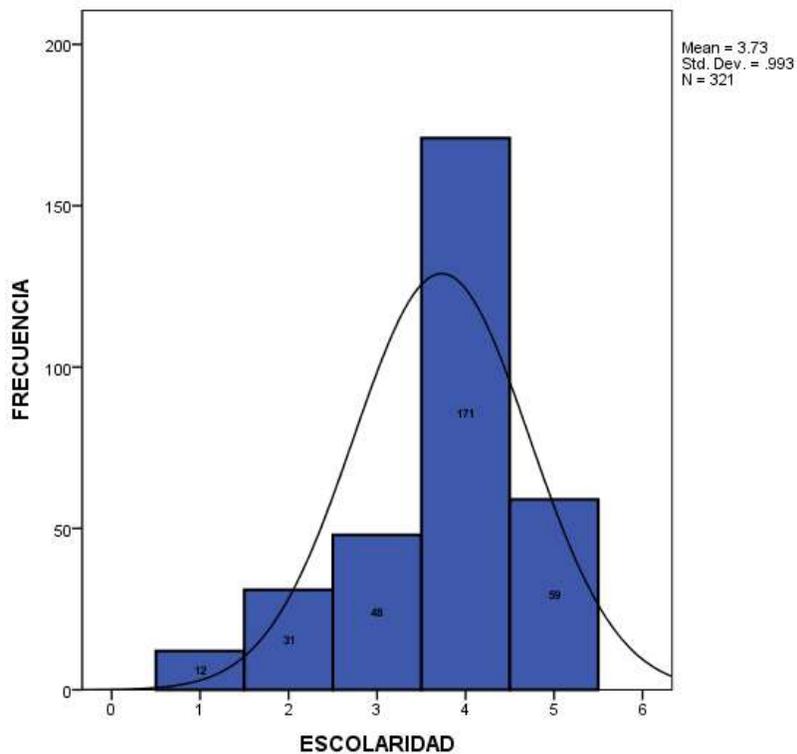


Figura 6 Histograma de la variable "escolaridad"

La Figura 7 **Error! Reference source not found.** nos indica cómo está distribuida nuestra muestra según el número de habitantes que integran el grupo de familias encuestadas, el grafico muestra que el 28.7% de la población encuestada tiene un total de 4 habitantes por familia, el 22% cuenta con un total de 3 personas en la familia, el 17.1% lo integran familias de 5 miembros, el 15% de la población lo integran familias de 2 personas, seguida por las familias

de un solo integrante que fue el 6.9% del total de la muestra, a esto le siguen las familias de 6,7,8 y 12 integrantes los cuales resultaron 4,3.4,1.6, y 3 por ciento respectivamente.

En el histograma presenta una curva asimétrica con un sesgo positivo, dando como resultado una distribución leptocúrtica de 2.118.

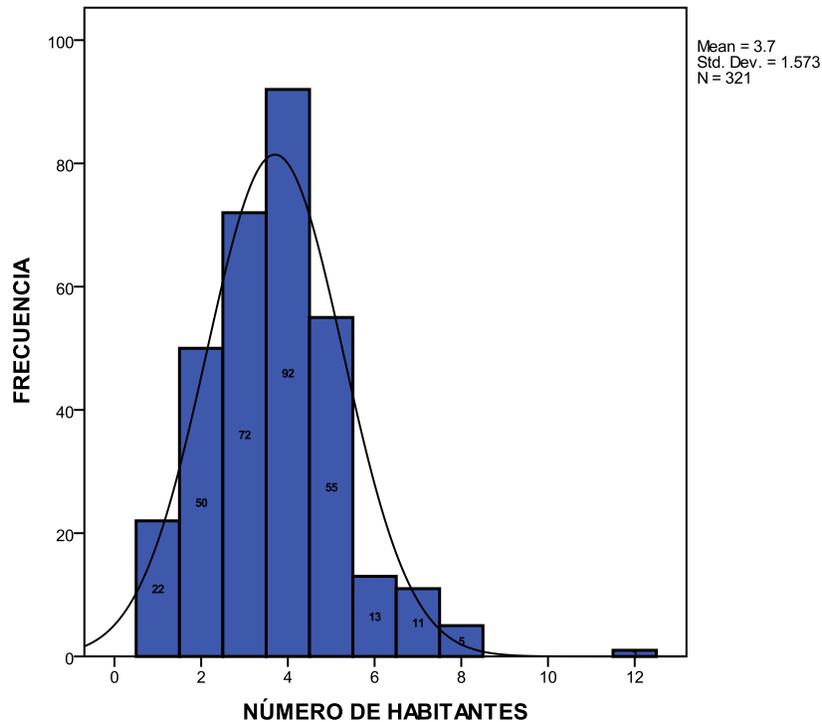


Figura 7 Histograma de la variable "número de habitantes"

La Figura 7 **Error! Reference source not found.** nos indica cómo está distribuida nuestra muestra según su nivel de ingreso económico el grafico muestra que el 51% de las familias encuestadas pertenecen al nivel medio, el 30% al nivel bajo, y el 17% al nivel alto.

En el histograma presenta una curva simétrica, dando como resultado una distribución leptocúrtica de 0.136.

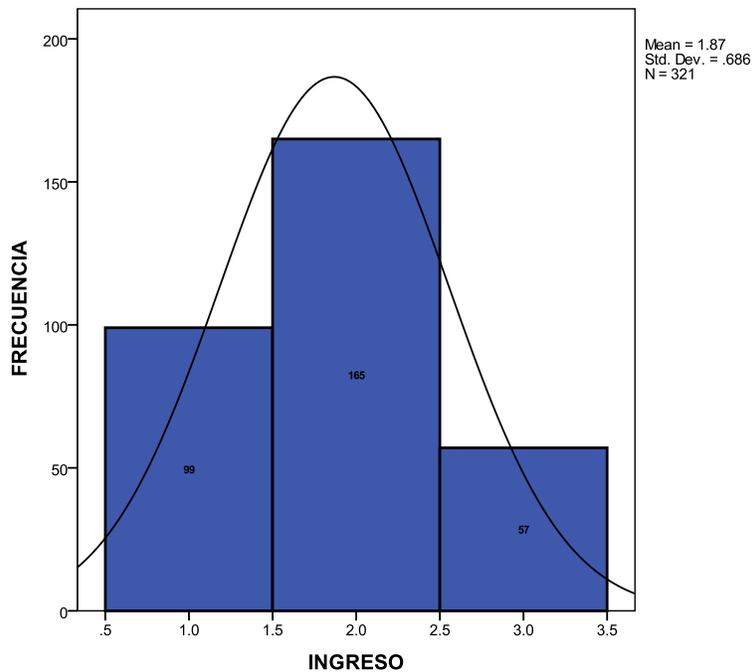


Figura 8 Histograma de la variable "ingreso"

Los gráficos que se presentan a continuación son las distribuciones de las respuestas de la población a los distintos ítems de la escala likert.

La Figura 9 **Error! Reference source not found.** indica la distribución de las respuestas de la muestra según la primer afirmación que se les presentó, la contaminación no afecta personalmente mi vida, a esto la población encuestada respondió de la siguiente manera, el 64% afirma estar en total desacuerdo con la propuesta antes mencionada, el 21.5%, está en desacuerdo con dicha propuesta, un 3.7% y 8.4% están de acuerdo y totalmente de acuerdo con la afirmación y un 1.9% se declara indiferente.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de 2.17.

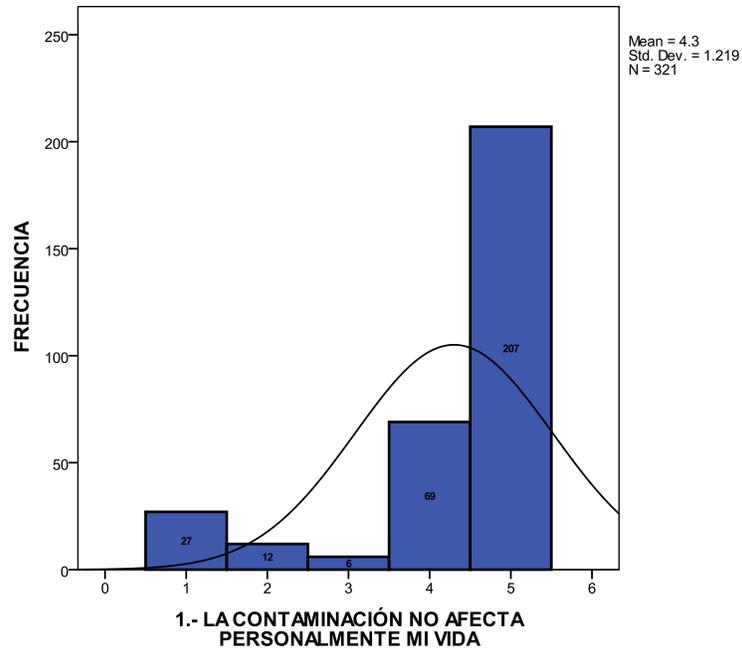


Figura 9 Histograma "item 1"

La Figura 10 **Error! Reference source not found.** indica la distribución de las respuestas de la muestra según la segunda afirmación propuesta por la escala, Antes de comprar los productos que necesito los comparo para seleccionar el que contenga menor cantidad de empaques o envoltorios, a esto la mitad de población encuestada se mostró indiferente esto es 50.8% , el 17% afirma estar de acuerdo con la propuesta antes mencionada, el 14.6%, está en desacuerdo con dicha propuesta, y el 8.4% de las familias están totalmente de acuerdo, de igual forma el 8.4% de las familias están totalmente en desacuerdo.

En el histograma presenta una curva simétrica, dando como resultado una distribución leptocúrtica de 0.062.

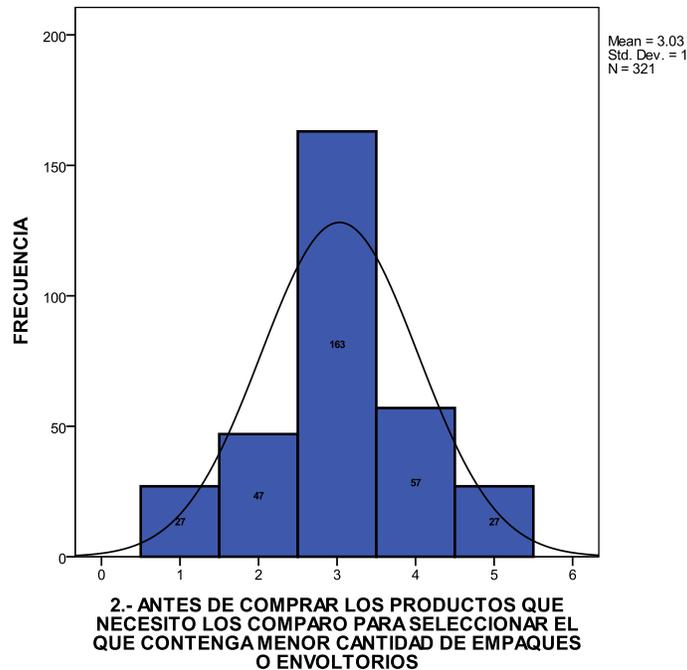


Figura 10 Histograma "item 2"

El siguiente gráfico (ver Figura 11 **Error! Reference source not found.**) indica la distribución de las respuestas de la muestra según la tercer afirmación propuesta por la escala, Incluso aunque el transporte público fuera más eficiente yo preferiría llevar mi auto al trabajo, en este ítem se observó la muestra dividida de manera muy equitativa en sus respuestas, el 27.1% está de acuerdo con esta afirmación, el 24.9% está en desacuerdo, el 18.7 % está totalmente de acuerdo, el 14.5% está totalmente en desacuerdo y el 14.6% se declara indiferente, a pesar el al similitud de los porcentajes aquí presentados ligeramente podemos apreciar una ligera tendencia a la actitud ambiental negativa por parte de las personas encuestadas.

En el histograma presenta una curva simétrica, dando como resultado una distribución mesocúrtica de -2.285.

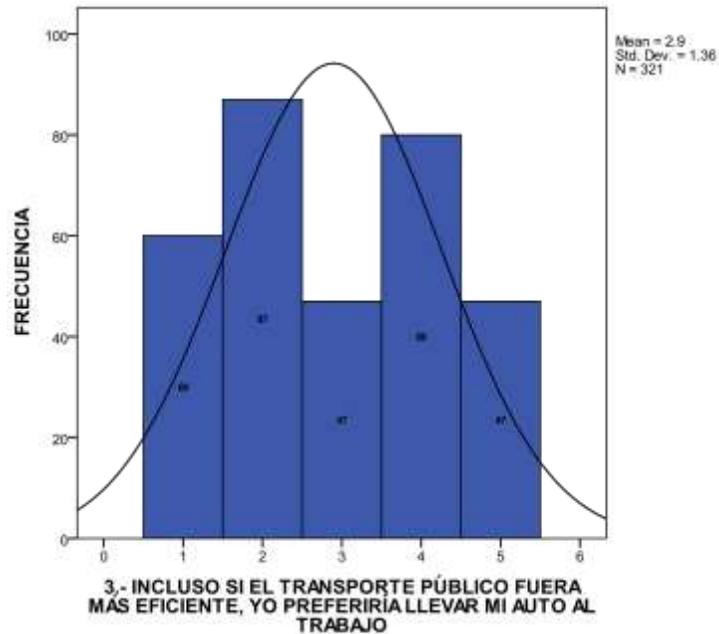


Figura 11 Histograma "item 3"

El siguiente grafico (ver Figura 12 **Error! Reference source not found.**) indica la distribución de las respuestas de la muestra según la cuarta afirmación propuesta en la escala, el gobierno tendrá que introducir duras medidas para frenar la contaminación ya que poca gente las regulara por sí misma, en este ítem se observó que la mayor parte de la muestra se declara a favor de establecer sanciones con el propósito de regular los hábitos y costumbres ambientales, el 46.2 % y el 42.1% se declararon en total acuerdo y de acuerdo, a la propuesta antes mencionada y tan solo el 5%, 4.7% y 2.2% respondieron estar en totalmente en desacuerdo, en desacuerdo o indiferentes respectivamente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 2.657.

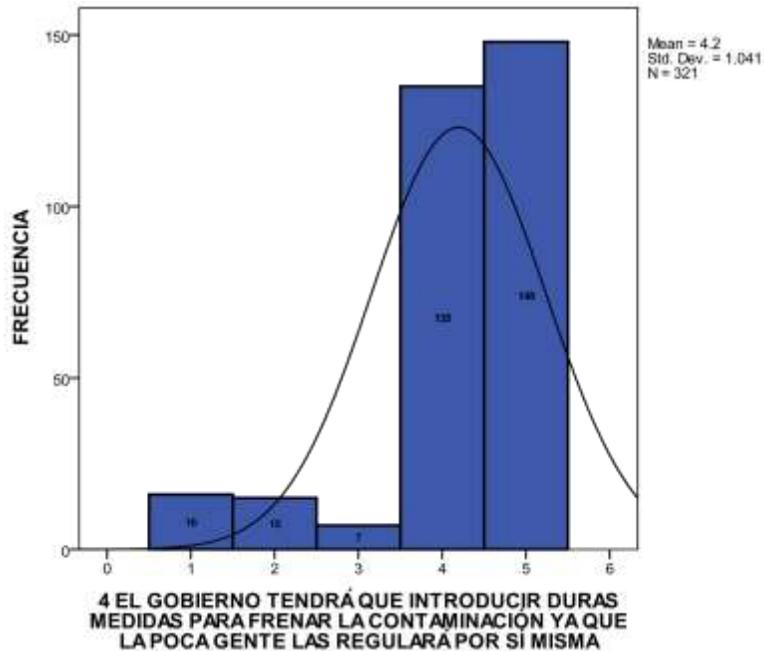


Figura 12 Histograma "item 4"

La Figura 13 **Error! Reference source not found.** muestra la distribución de las respuestas de la población encuestada en la quinta afirmación propuesta en la escala, no deberíamos preocuparnos por matar demasiados animales de caza porque a la larga las cosas se equilibrarán, en esta afirmación se observó que la mayor parte de la muestra se declara a en contra de la propuesta, donde el 65.7% y el 19.3% se mostraron en total desacuerdo y en desacuerdo respectivamente, el 5% y el 3.7% se mostraron totalmente de acuerdo y de acuerdo respectivamente y el 6.2% de la muestra se mostro indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 2.762.

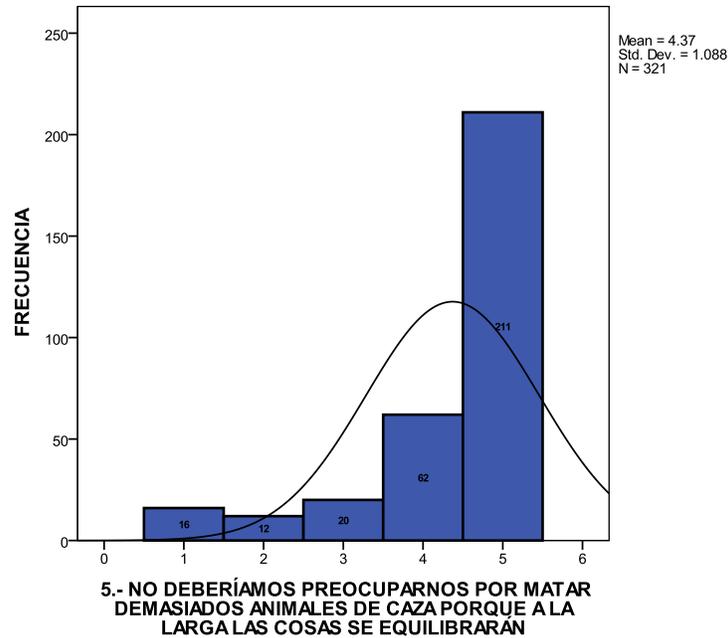


Figura 13 Histograma "ítem 5"

La Figura 14 **Error! Reference source not found.** muestra la distribución de las respuestas de la población encuestada en la sexta afirmación propuesta en la escala, estaría dispuesto a hacer sacrificios personales para reducir el ritmo de la contaminación aunque los resultados inmediatos no puedan parecer significativos, en esta afirmación se observó que la mayor parte de la muestra se declara en favor de la propuesta, donde el 53.3% y el 31.5% se mostraron de acuerdo y totalmente de acuerdo respectivamente, el 2.8% y el 4.4% se mostraron totalmente en desacuerdo y en desacuerdo respectivamente y el 8.1% de la muestra se mostró indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 2.431

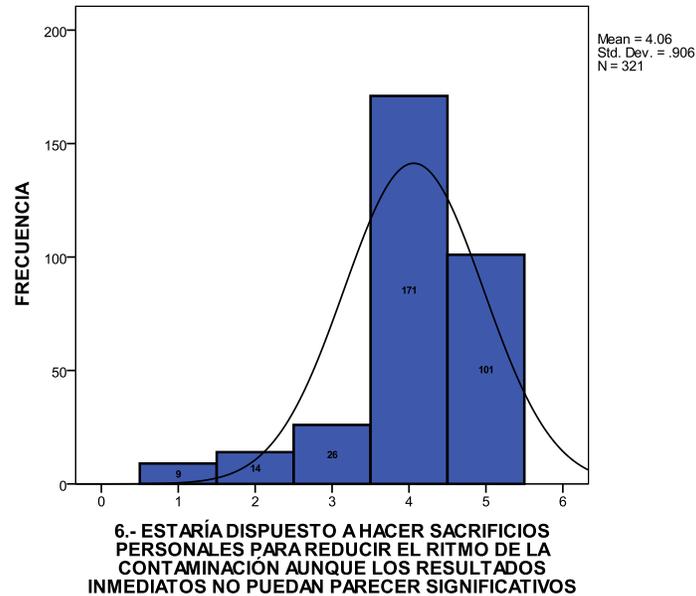


Figura 14 Histograma "item 6"

La Figura 15 **Error! Reference source not found.** muestra la distribución de las respuestas de la población encuestada en la séptima afirmación propuesta en la escala, si el aparato de refrigeración presenta fuga considero necesario mandarla reparar aunque ya se vaya a acabar la temporada de verano, en esta afirmación se observó que la mayor parte de la muestra se declara en inconformidad con la propuesta, donde el 44.2% y el 34.3% se mostraron en desacuerdo y totalmente de desacuerdo respectivamente, el 2.8% y el 4.4% se mostraron totalmente en desacuerdo y en desacuerdo respectivamente y el 10% de la muestra se mostró indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 0.737

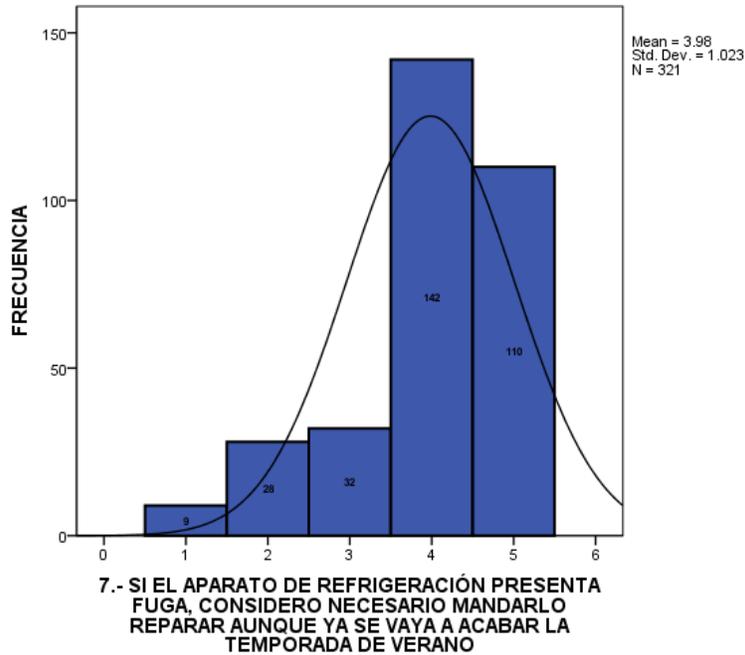


Figura 15 Histograma "ítem 7"

La figura que se muestra a continuación (ver Figura 16 **Error! Reference source not found.**) indica la distribución de las respuestas de la muestra según el octavo ítem de la escala. Se debe utilizar el aparato de aire acondicionado solo el tiempo necesario para adquirir un confort térmico y no mantenerlo encendido todo el día, en este ítem se observó una marcada tendencia a estar de acuerdo con la afirmación, el 38.6% está de acuerdo con esta afirmación, el 35.5% está totalmente de acuerdo, el 11.2% está totalmente de acuerdo, el 6.9% está totalmente en desacuerdo y el 7.8% se declara indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de -0.004.

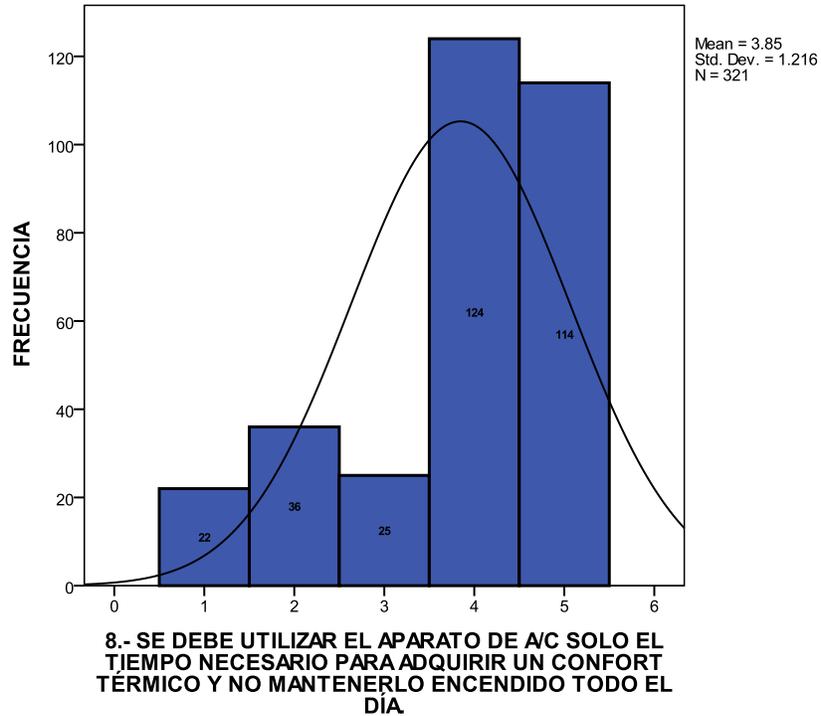


Figura 16 Histograma "ítem 8"

La Figura 17 **Error! Reference source not found.** indica la distribución de las respuestas de la muestra según el ítem número nueve de la escala, prefiero comprar un aparato de refrigeración usado ya que su costo es menor, en este ítem se observó una marcada tendencia a estar en contra de la afirmación, el 54.5% está de totalmente en desacuerdo con esta afirmación, el 32.7% está simplemente en des acuerdo, el 3.7 % está de acuerdo, el .6% está totalmente en acuerdo y el 8.4% se mantiene indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 1.874.

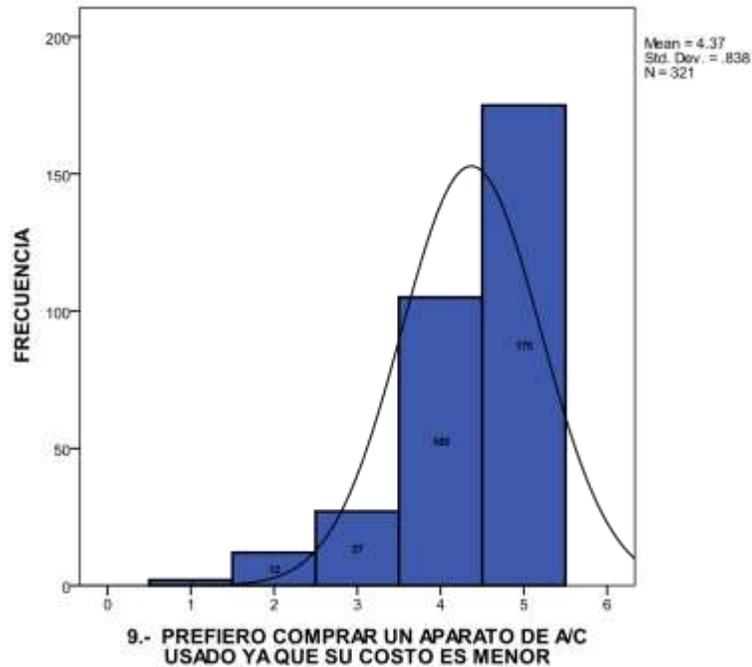
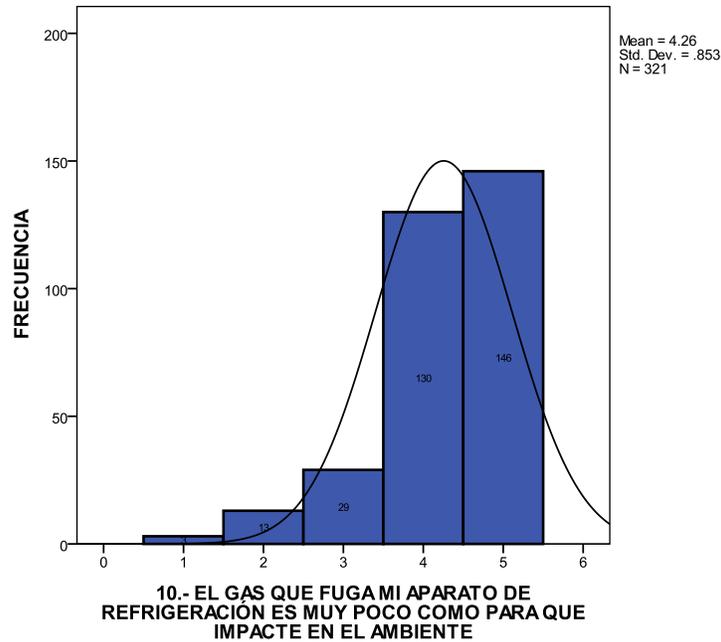


Figura 17 Histograma "ítem 9"

El siguiente gráfico (Figura 18 **Error! Reference source not found.**) indica la distribución de las respuestas de la muestra según el ítem número diez de la escala, el gas que fuga mi aparato de refrigeración es muy poco como para que impacte en el ambiente, en este ítem se observó una marcada tendencia a estar en contra de la afirmación, el 45.5% está de totalmente en desacuerdo con esta afirmación, el 40.5% está simplemente en desacuerdo, el 4 % está de acuerdo, el .9% está totalmente en acuerdo y el 9% se mantiene indiferente.

En el histograma presenta una curva asimétrica con sesgo negativo, dando como resultado una distribución leptocurtica de 1.719.



**Figura 18 Histograma ítem 10**

El gráfico que se presenta a continuación (ver Figura 19 **Error! Reference source not found.**) nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 11 de la escala, es una costumbre familiar tener al menos un día de convivencia en familia, podemos observar que del total de los casos, el 44.2 % y el 34.3% respondieron estar de acuerdo y totalmente de acuerdo respectivamente, el 5.3% y el 2.2% se mostraron el desacuerdo y totalmente en desacuerdo, y el 14% se declaró indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de 1.079

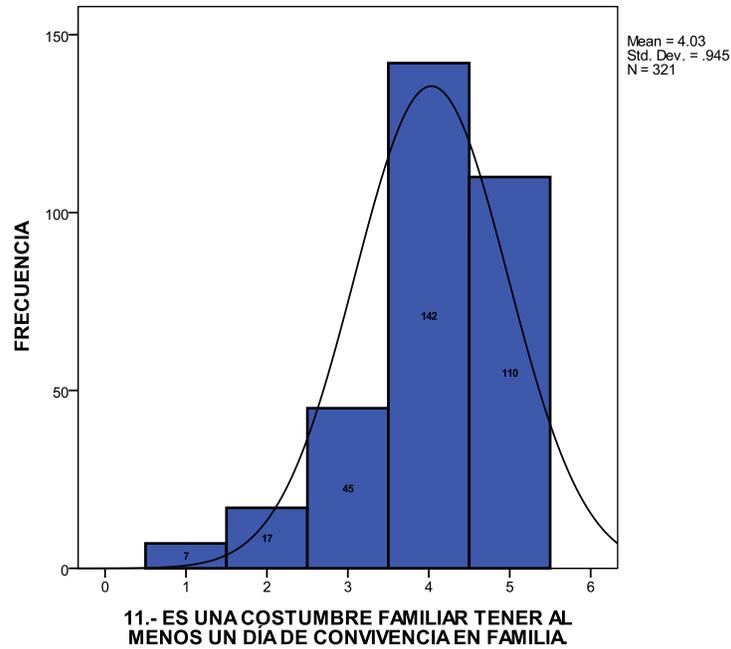


Figura 19 Histograma ítem 11

El gráfico que se presenta a continuación (ver Figura 20 **Error! Reference source not found.**) nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 12 de la escala, la vida diaria es muy atareada y no hay tiempo para actividades familiares, de esta figura podemos observar que del total de los casos, el 36.8 % y el 39.9% respondieron estar en total desacuerdo y en desacuerdo y respectivamente, el 12.5% y el 2.5% se mostraron el desacuerdo y totalmente en desacuerdo, y el 8.4% se declaró indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de .111

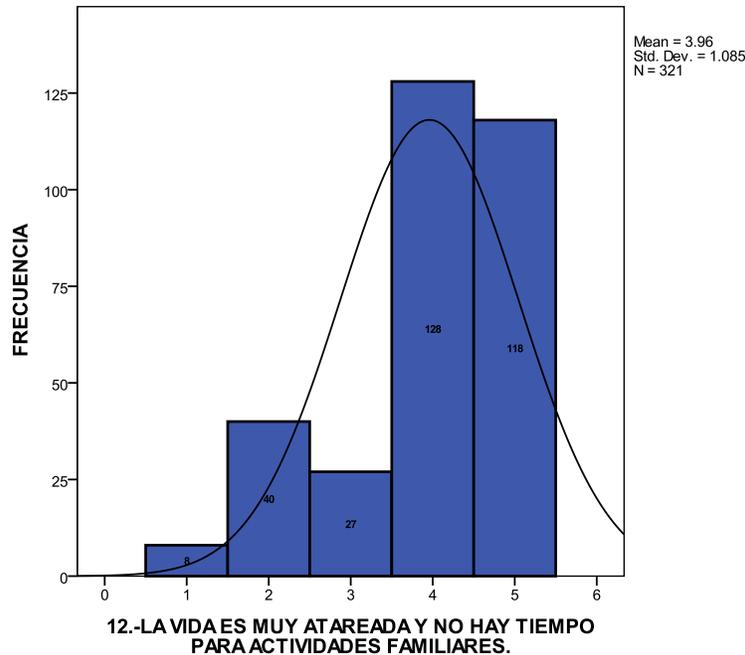


Figura 20 Histograma ítem 12

El gráfico que se presenta a continuación (ver Figura 21 **Error! Reference source not found.**) nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 13 de la escala, Cada cuatro meses aproximadamente le doy servicio a mi automóvil para mantenerlo en buenas condiciones, de esta figura podemos observar que los porcentajes por respuesta son aproximados, sin embargo en acumulado las repuestas con sesgo positivo resultaron mayores, el 42.7 % y el 22.4% respondieron estar de acuerdo y totalmente de acuerdo respectivamente, el 17.8% y el 4.4% se mostraron el desacuerdo y totalmente en desacuerdo, y el 12.8% se declaro indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de  $-0.586$ .

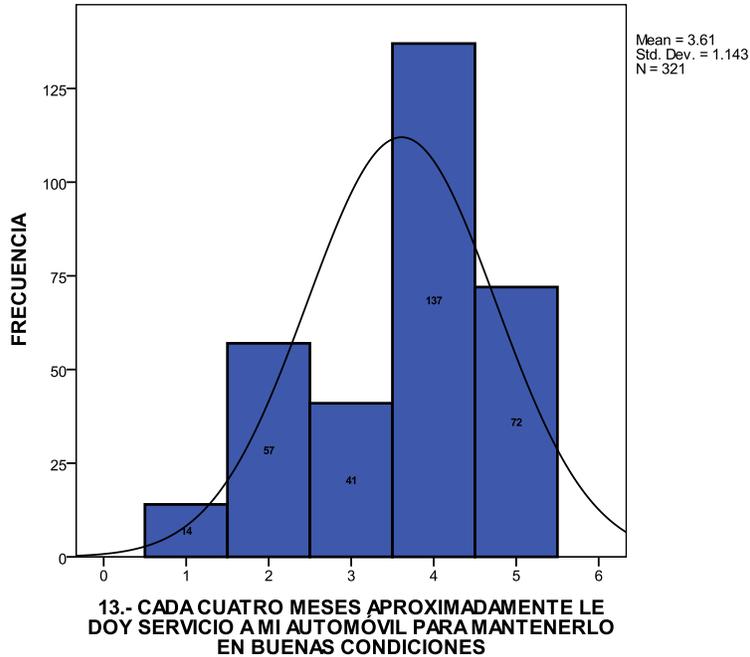


Figura 21 Histograma ítem 13

La figura que se presenta a continuación (ver Figura 22Error! Reference source not found.) nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 14 de la escala, es preferible tener una televisión en cada recámara para evitar discusiones sobre la programación., de esta figura podemos observar que los porcentajes por respuesta son aproximados, con un alto porcentaje de indiferencia, el 28.7 % y el 17.1% respondieron estar en desacuerdo y totalmente en desacuerdo respectivamente, el 23.4% y el 10.6% se mostraron de acuerdo y totalmente de acuerdo, y el 20.2% se declaró indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de -1.089.

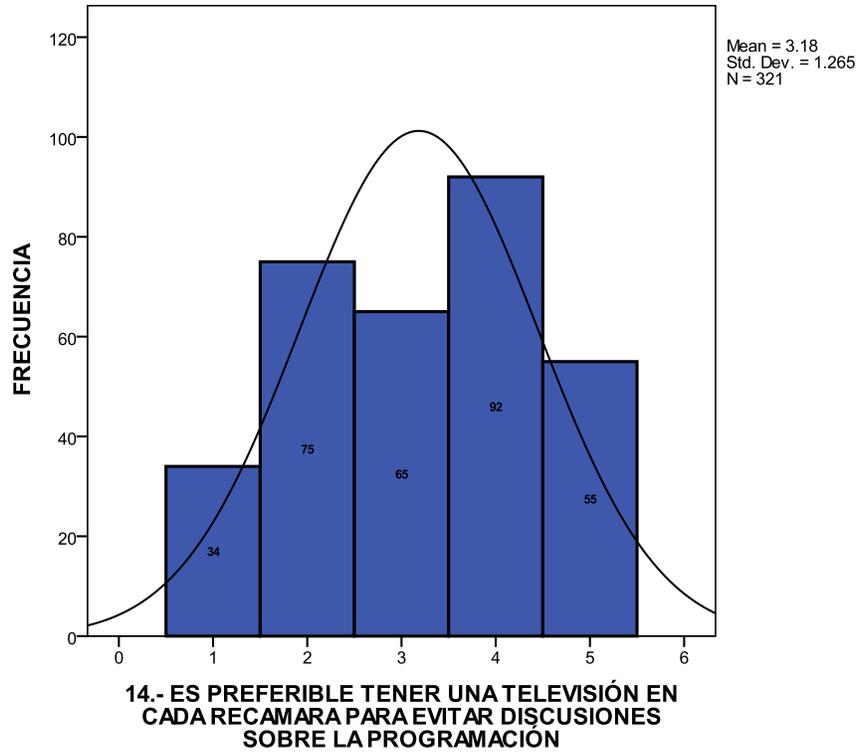


Figura 22 Histograma ítem 14

La gráfica que se presenta a continuación (ver Figura 23 **Error! Reference source not found.**) nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 15 de la escala, cuando voy a comprar adquiero solo lo que necesito., de esta figura podemos observar que el 44.5 % y el 22.4% respondieron estar de acuerdo y totalmente de acuerdo respectivamente, el 11.8.4% y el 4.4% se mostraron en desacuerdo y totalmente en desacuerdo, y el 16.8% se declaró indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de  $-.057$ .

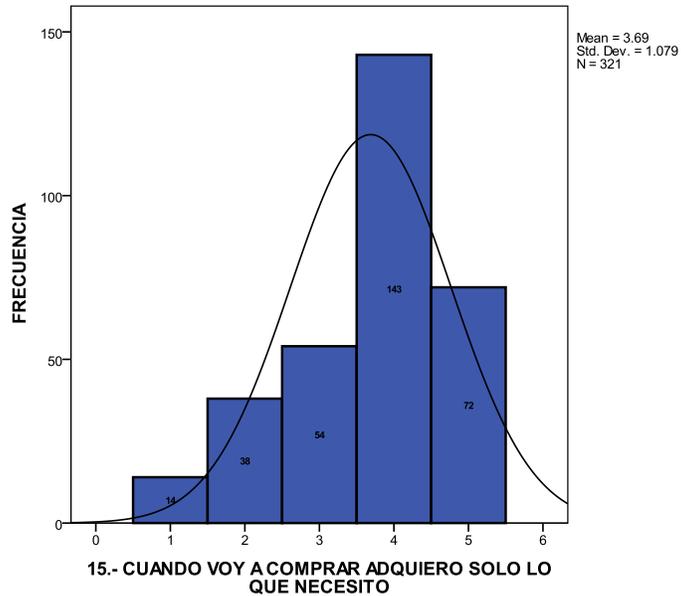


Figura 23 Histograma ítem 15

La Figura 24 **Error! Reference source not found.** que nos indica las respuestas de las diferentes familias encuestadas en relación al ítem número 16 de la escala, para que la ropa sucia no se acumule el fin de semana es preferible lavar diario cantidades pequeñas., de esta figura podemos observar que el 35.8 % y el 34% respondieron estar en desacuerdo y totalmente en desacuerdo respectivamente, el 11.2% y el 5.9% se mostraron de acuerdo y totalmente de acuerdo, y el 13.1% se declaró indiferente ante esta afirmación.

En el histograma presenta una curva asimétrica con un sesgo negativo, dando como resultado una distribución leptocúrtica de  $-0.178$ .

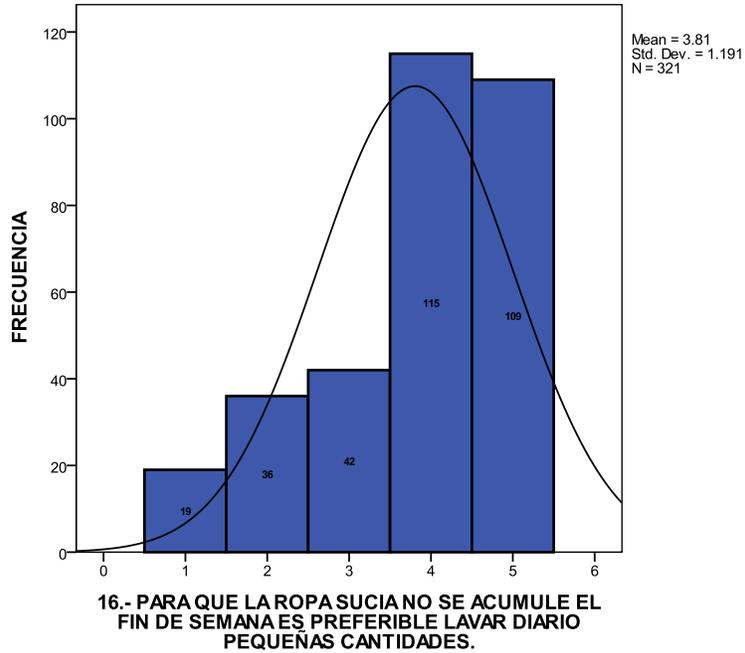


Figura 24 Histograma ítem 16

#### 4.2.1 Análisis cuantitativo de la emisión

En este capítulo se realiza un análisis estadístico de la información obtenida, para identificar diversos escenarios del comportamiento de los datos y encontrar variables dependientes e independientes existentes, con el propósito de encontrar perfiles de emisión ya sea por familia, por ingreso, por escolaridad e identificar las actitudes ambientales de las familias seleccionadas para la investigación.

La Tabla 3 refleja la distribución de las 321 personas entrevistadas, en función de las principales características consideradas en el trabajo

Tabla 3 Composición de la muestra

Composición de la muestra n=321		
Personas que dieron servicio	Si	281

<b>a su aparato de aire acondicionado</b>	No	40
<b>Tipo de familia</b>	Nuclear	219
	Monoparental	55
	Extendida	46
<b>Escolaridad</b>	Primaria	12
	Secundaria	31
	Preparatoria	48
	Licenciatura	171
	Posgrado	59
<b>Ingreso</b>	Alto	57
	Medio	165
	Bajo	57
<b>Emisiones de las personas que sí dieron servicio a su aparato de aire acondicionado</b>	no fugaron	106
	si fugaron	175

En la Figura 25 **Error! Reference source not found.** se presenta la emisión promedio en los diferentes estratos sociales que se contemplan en este estudio, cabe mencionar que el nivel socioeconómico alto contribuyó con un total de 167 kg de refrigerante al ambiente, el nivel medio 301.75 kg y el nivel bajo 139.75, sin embargo se observa que el nivel social alto tiene mayor emisión promedio, aproximadamente 3 kilos, lo cual representa cerca de el doble de emisión promedio relación a los otros dos niveles.

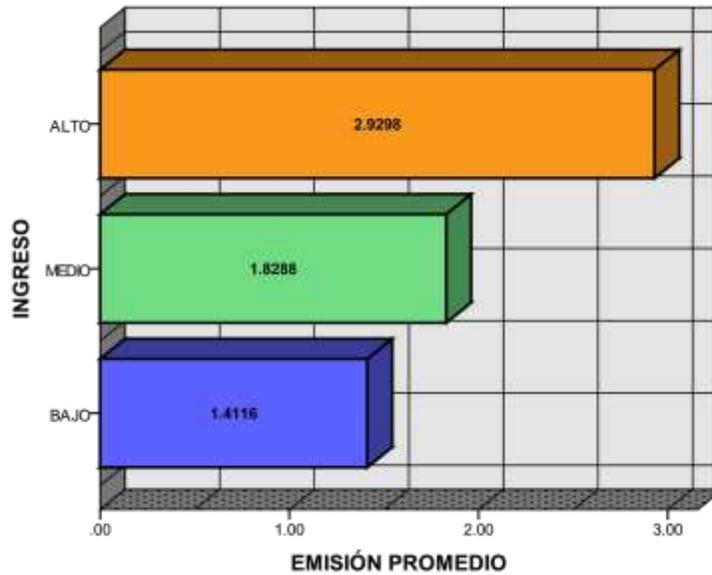


Figura 25 Emisión de refrigerante por estrato socioeconómico

En lo que se refiere al grado de escolaridad de las familias encuestadas, la Figura 26 **Error! Reference source not found.** nos muestra el promedio de emisión por nivel de estudio, el cual nos indica que el nivel de escolaridad que tiene el promedio más alto es el nivel primaria con 2.54 kg de promedio, seguido por el nivel de licenciatura con 2 kg y posgrado con 1.7kg, de los cuales el total de emisión para los diferentes niveles de escolaridad es 30.5, 40.75, 80.25, 352.5, 104.5 para primaria, secundaria, preparatoria, licenciatura y posgrado respectivamente. A la vista de estos datos podemos suponer que entre más bajo el nivel de estudios, mayor es la contribución de refrigerante al ambiente.

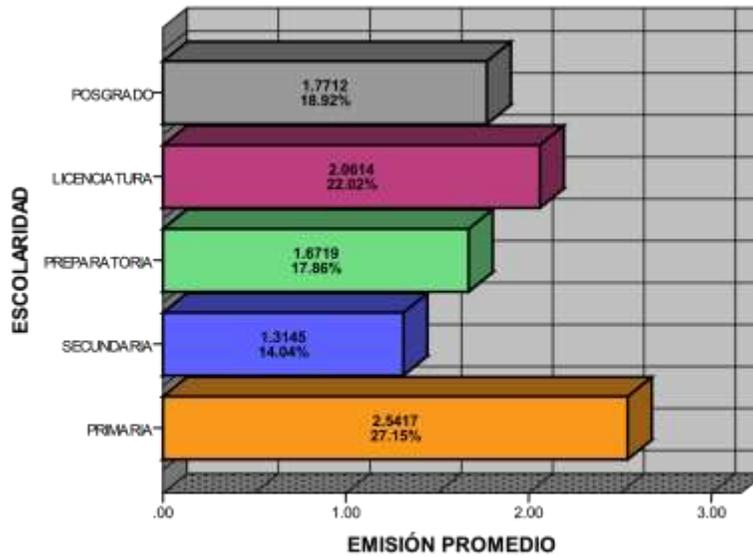


Figura 26 Emisión media por Escolaridad

La Figura 27 **Error! Reference source not found.** muestra la emisión promedio según el tipo de familia, en cuanto a esto los totales de emisión por tipo de familia fueron nuclear 411 kg, monoparental 93.5 y extendida 104, sin embargo, como era de esperarse el tipo de familia extendida tiene mayor emisión promedio con una media de 2.21kg, seguida por el nivel nuclear y por último el monoparental con 1.8 y 1.7 kg cada uno. Comparativamente, cabe destacar que el 50 % las familias encuestadas de tipo monoparental, está compuesto por hogares de un solo habitante, y un alto porcentaje de las familias de tipo extendido (60%) tienen entre 8 y 12 habitantes.

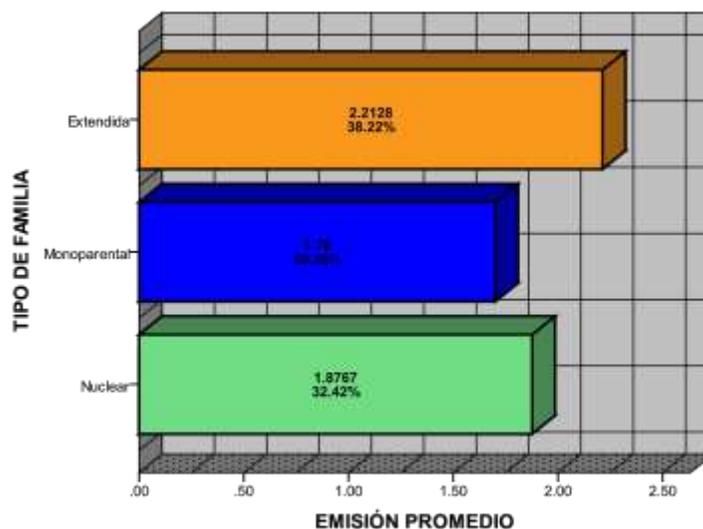


Figura 27 Emisión media por tipo de familia

En lo referente a si se da, o no servicio a aparato de aire acondicionado como se puede apreciar en la **Error! Reference source not found.** del total de la población encuestada perteneciente el nivel bajo de ingresos, el 68% no le dio servicio a su aparato de refrigeración ese verano, mientras que arriba del 40% de la población de nivel medio de igual forma no le realizó servicio a su aparato de aire acondicionado, y solo el 20% de la población de ingresos altos no lo hicieron, es decir de el total de las personas que no realizaron el servicio correspondiente a su o a sus aparatos de aire acondicionado, el 50% tienen un ingreso económico bajo, el 38% ingreso medio y el 5% ingreso alto (Figura 28)

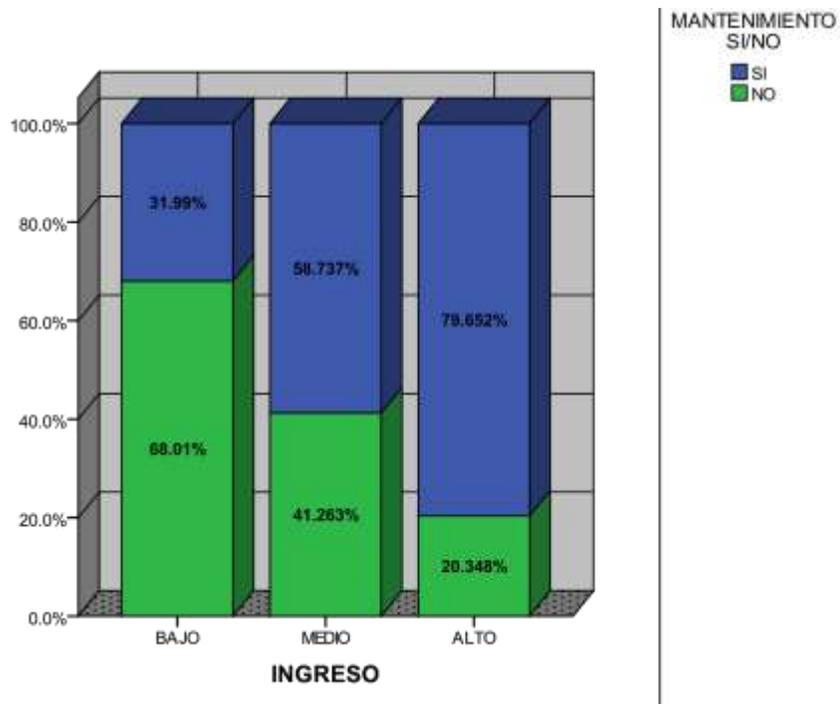
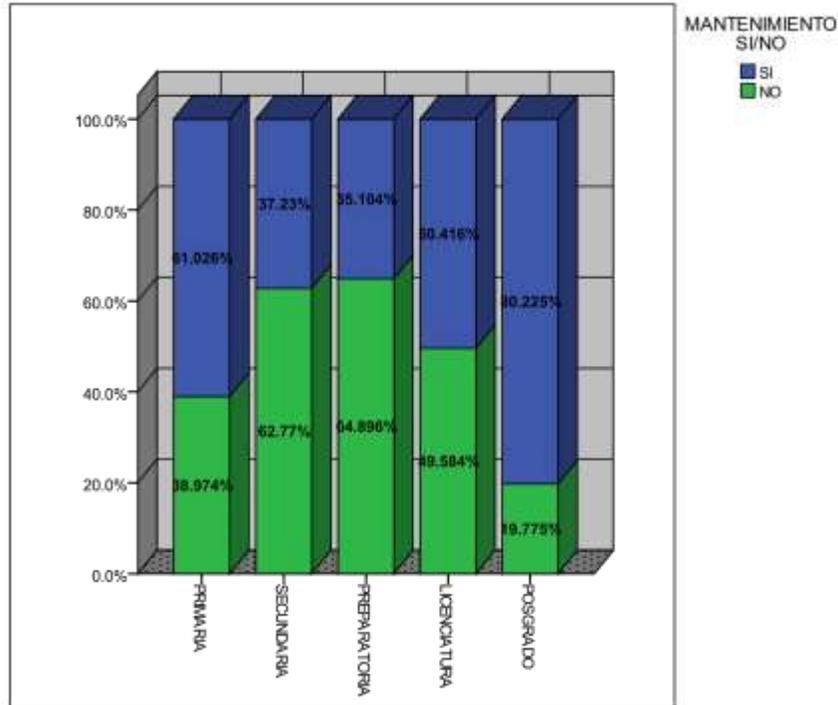


Figura 28 Relación Ingreso mantenimiento

En lo referente a la escolaridad de las personas de la muestra que realizaron el servicio de mantenimiento a su aparato de A/C podemos observar en la **Error! Reference source not found.** que para las familias donde la escolaridad promedio es secundaria, preparatoria e incluso licenciatura, el porcentaje de familias que no atendieron el servicio de A/C fluctúa entre el 50% y 65% lo cual indica que más de la mitad de las personas pertenecientes a la muestra para cada nivel no realizaron el mantenimiento sugerido al aparato de refrigeración, de la misma del total de familias con un nivel superior a la licenciatura solo el 20% de estas no realizó el servicio (Figura 29).



**Figura 29 Relación escolaridad mantenimiento**

Por su parte cuando se les pregunta a las personas encuestadas si conocen la importancia de dar servicio a sus aparatos de refrigeración o si están conscientes del impacto que provoca la emisión del refrigerante contenido en su aparato de refrigeración, la mayoría de las personas encuestadas dice tener conciencia de esto, comparativamente casi la mitad de cada uno de los tipos de familia de la muestra no realizaron el servicio los porcentajes varían entre el 40% y el 60%. A la vista de estos datos puede concluirse que independientemente del tipo de familia al que pertenecen no influye en la determinación de darle o no el servicio al aparato de refrigeración, ver Figura 30 **Error! Reference source not found.**

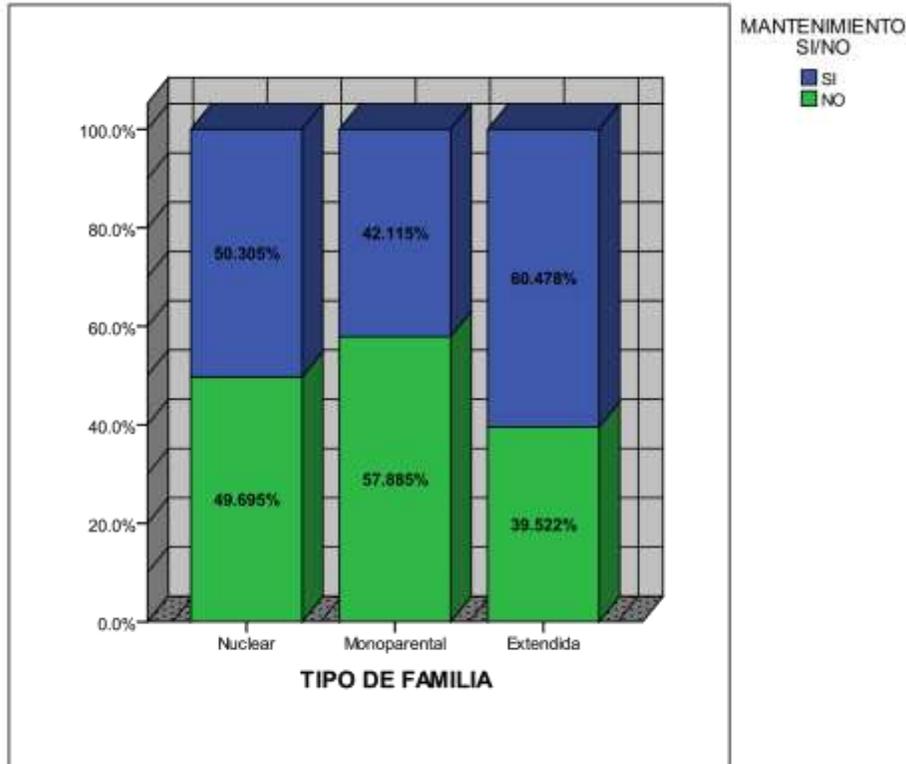


Figura 30 Relación tipo de familia y mantenimiento

La Figura 31 **Error! Reference source not found.** presenta los valores del indicador tipo de familia y el número de habitantes registrado por cada tipo. Del cual podemos destacar que el tipo de familia extendido (aparte de padres e hijos viven primos hermanos etc.) arriba del 30% de estas familias tiene más de 12 habitantes, el grafico nos indica que el tipo de familia nuclear, es decir que ya sea padre o madre está encargado unitariamente de la familia, está constituido principalmente (arriba del 50% del total) por un solo habitante en esa casa. En lo que se refiere al número de habitantes que hay en las familias de tipo nuclear los porcentajes están más uniformes arriba del 10% cada categoría con excepción de cuando la casa cuenta con un solo habitante es únicamente el 6% aproximadamente.

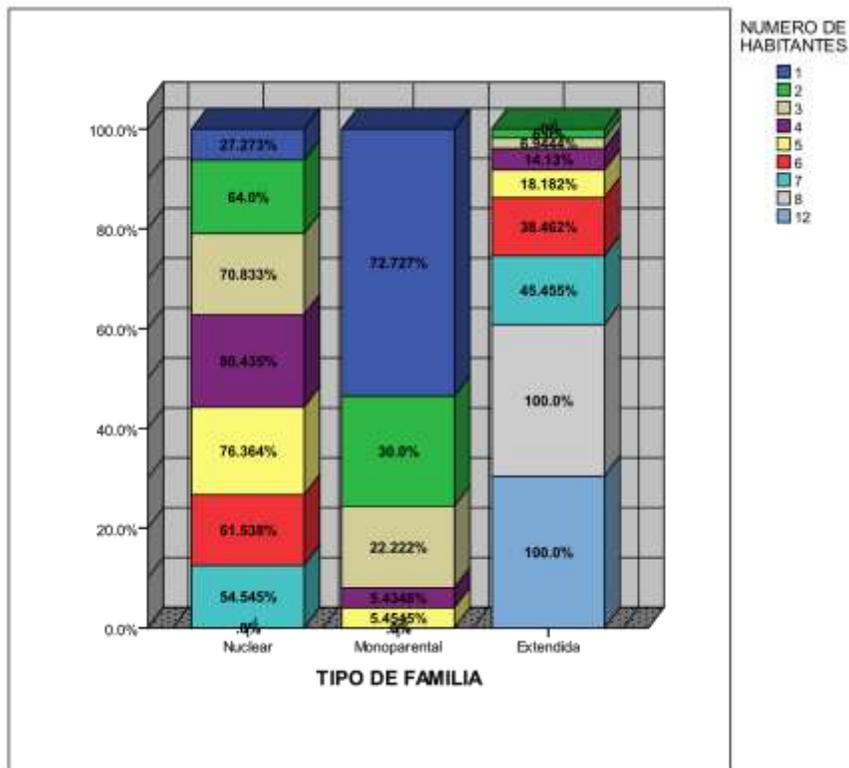


Figura 31 Número de habitantes por tipo de familia

En cuanto al grado de estudios se refiere, la Figura 32 **Error! Reference source not found.** refleja la escolaridad promedio que presentan los distintos tipos de familia, podemos observar que la escolaridad que predomina en el tipo de familia monoparental es el nivel primaria (50.98%), en el tipo de familia monoparental el grado de estudios predominante es preparatoria (30%) y en el tipo nuclear es posgrado (27%).

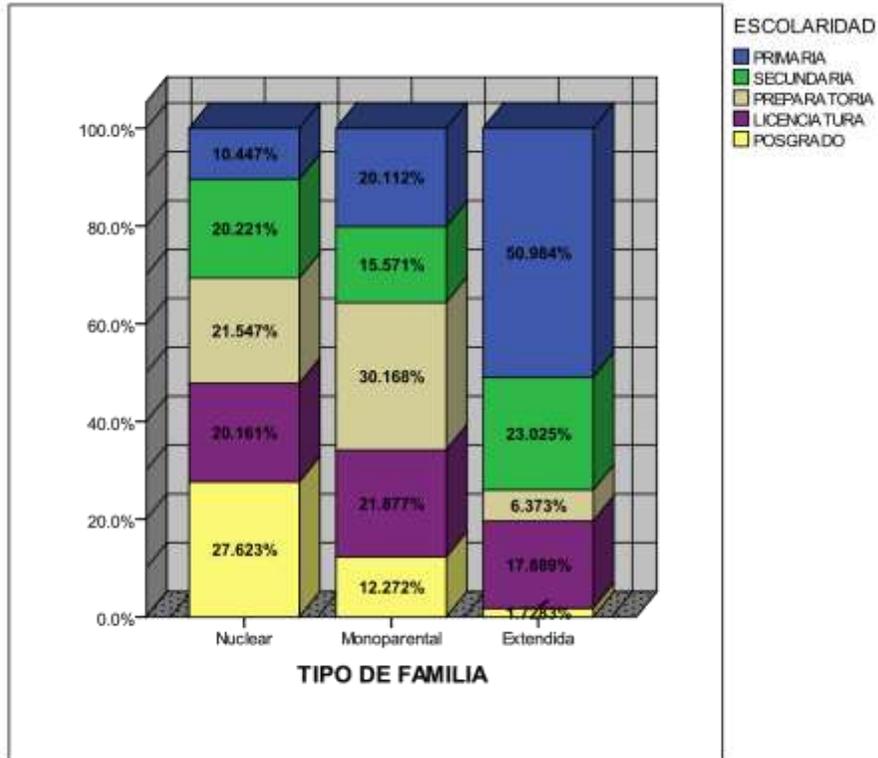


Figura 32 Escolaridad por tipo de familia

En relación al ingreso que presenta cada tipo de familia el predominante es el nivel medio para los tres tipos con 52%,49% y 46% para nuclear monoparental y extendida respectivamente, seguido por el nivel bajo con 27%,40% y 34% respectivamente (Figura 33).

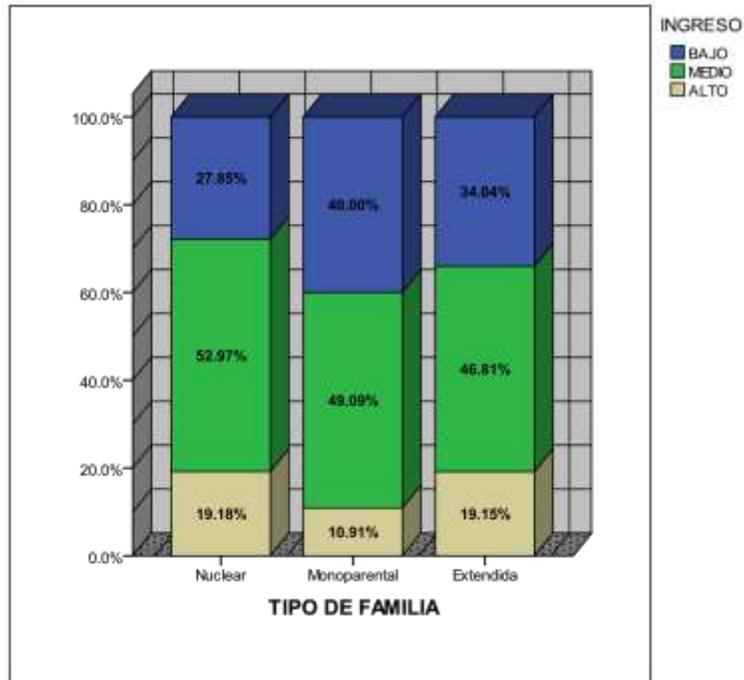


Figura 33 Ingreso por tipo de familia

La Figura 34 **Error! Reference source not found. Error! Reference source not found.** muestra el ingreso que percibe la familia en relación con la escolaridad y el ingreso percibido en relación con el número de personas que habitan en el hogar, podemos observar que entre más bajo el nivel de estudios el ingreso tiende a bajar, ya que casi el 70% de las familias que tienen un nivel de estudios básico también tiene un nivel económico bajo, sin embargo las estas mismas familias de nivel social bajo son las que presentan mayor número de habitantes por hogar

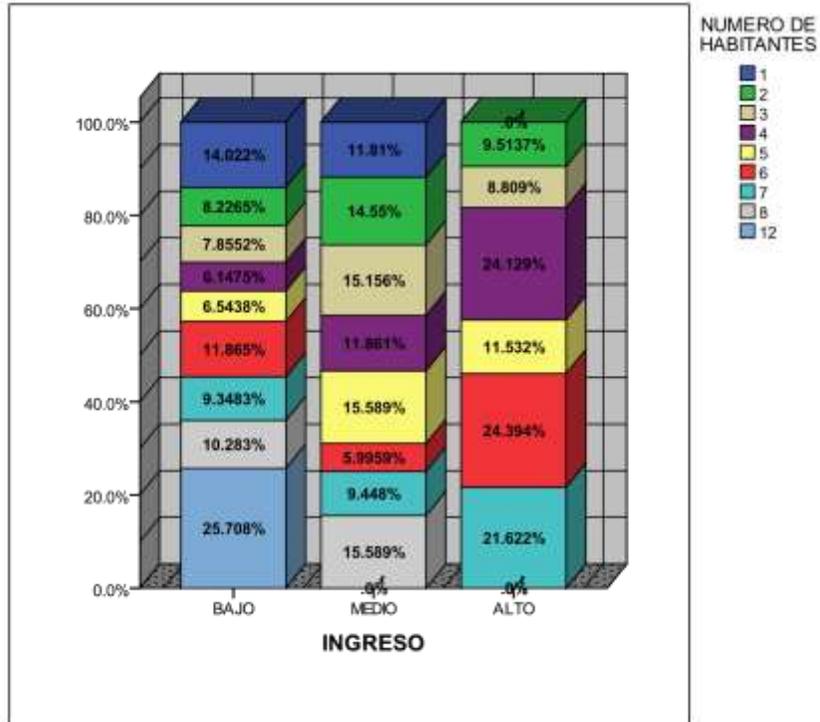


Figura 34 Relación entre ingreso y número de habitantes

El gráfico siguiente (ver Figura 35 **Error! Reference source not found.**) muestra la relación entre el ingreso y la escolaridad máxima de la familia encuestada, es importante mencionar que entre mas alto el grado de estudios mayor es el porcentaje de familias con ingreso alto

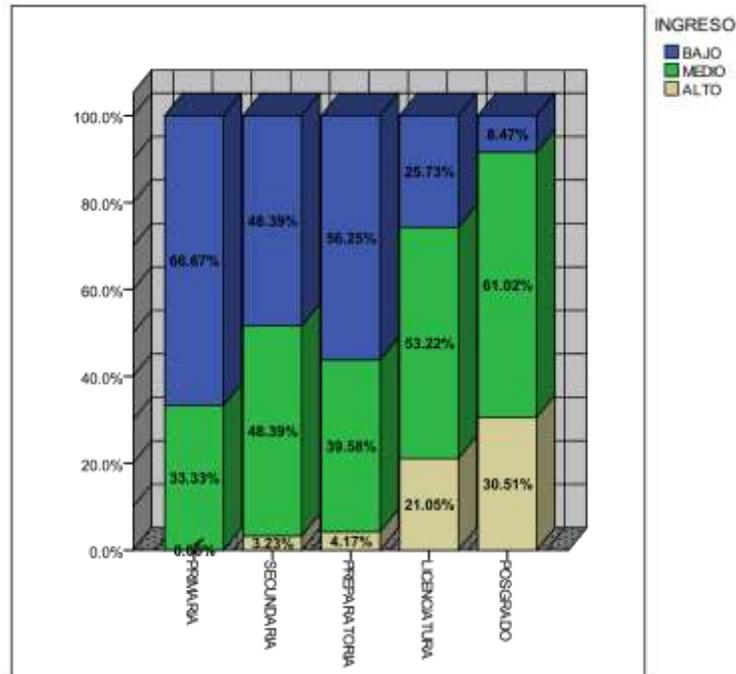


Figura 35 Relación entre el grado de estudios y el ingreso

#### 4.2.2 Caracterización de familias según su comportamiento ambiental

Los gráficos que se presentan a continuación, muestran la respuesta promedio de las diferentes variables de clasificación que se tomaron en cuenta en este estudio, como se mencionó anteriormente, las respuestas de la encuesta estaban dadas en escala likert con niveles de total desagrado hasta total agrado incluyendo la opción de indiferente. A la luz de esto, cada respuesta estaba valuada con valores del 1 al 5 donde entre más alta sea la puntuación

En el siguiente gráfico (ver Figura 36 **Error! Reference source not found.**) se representan para los diferentes ítems los valores que toman en la escala ordinal las variables cualitativas que han sido seleccionada en este caso podemos apreciar las respuestas a la escala tanto de las personas que realizaron servicio de mantenimiento al aparato de aire acondicionado como las que no, la puntuación más alta es 5 y se observa la similitud entre las respuestas de los dos

grupo, con ligeras variantes en el ítem 12 que cuestiona el tiempo para las actividades familiares se puede señalar que el grupo que sí realizó el servicio de mantenimiento a su aparato de refrigeración obtuvo una mayor puntuación, otro ítem en el que se observa diferencia en el promedio de respuestas en el ítem 9 el cual afirma la preferencia por comprar un aparato de aire acondicionado usado, donde nuevamente el grupo que sí realizó el mantenimiento, obtuvo una ligera mayor puntuación de manera inversa al responder el ítem 8 el cual afirma que es preferible utilizar el aparato de aire acondicionado solo el tiempo necesario para adquirir un confort térmico el grupo que no dio mantenimiento a su aparato de aire acondicionado obtuvo mayor puntuación, lo que significa que este grupo en promedio tiende encender el aparato de A/C únicamente el tiempo necesario, de igual manera difirieron en el ítem 2 en que se menciona el hábito de comprar los productos que tengan menor cantidad de empaques podemos observar que el grupo que sí realizó mantenimiento tuvo menor puntuación.

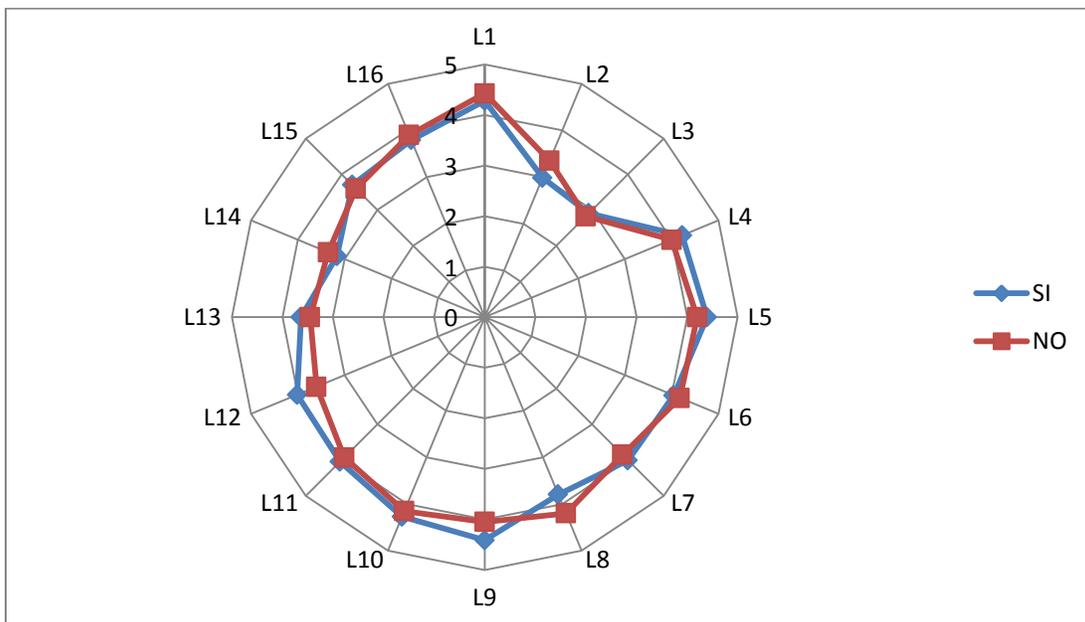


Figura 36 Caracterización de la muestra por servicio de mantenimiento al aparato de A/C

El siguiente gráfico (ver Figura 37 **Error! Reference source not found.**) muestra las respuestas a los diferentes ítems de la escala según el grado de estudio de la persona encuestada, es importante observar la que a pesar de que en algunos ítems el promedio de respuesta es similar, hay otros en los que sí es muy marcada la diferencia de respuesta promedio en los diferentes grados de escolaridad, por ejemplo en el ítem 14 el cual afirma sobre la preferencia de contar con un aparato televisor en cada recámara podemos ver que entre mas alto el grado de estudio la puntuación de respuesta es mayor, sin embargo el mayor grado de estudios tabulado en la escala, que es el posgrado se muestra prácticamente indiferente a la situación, y el menor grado de estudios tabulado que es primaria tuvo una puntuación promedio de 2 lo cual indica que en general está de acuerdo con esta actividad. Otro caso es el del ítem 16 el cual presenta una afirmación relacionada con el hecho de utilizar la máquina de lavado diario para que la ropa no se acumule en fin de semana, en este ítem todos los grupos respondieron no estar de acuerdo con esa idea a excepción del grupo con escolaridad primaria que se mostró indiferente ante la propuesta, otra cuestión importante radica en los ítems 2 y 3 en los cuales toda la población resultó con promedios de respuesta muy bajos y en el caso del ítem 2 la población en general mostro indiferencia ante la afirmación propuesta, donde el ítem 3 expone la preferencia por la persona encuestada por utilizar su automóvil incluso habiendo la posibilidad de un transporte público más eficiente, en este caso el sector primaria se muestra indiferente y los demás simplemente prefieren el transporte personal.

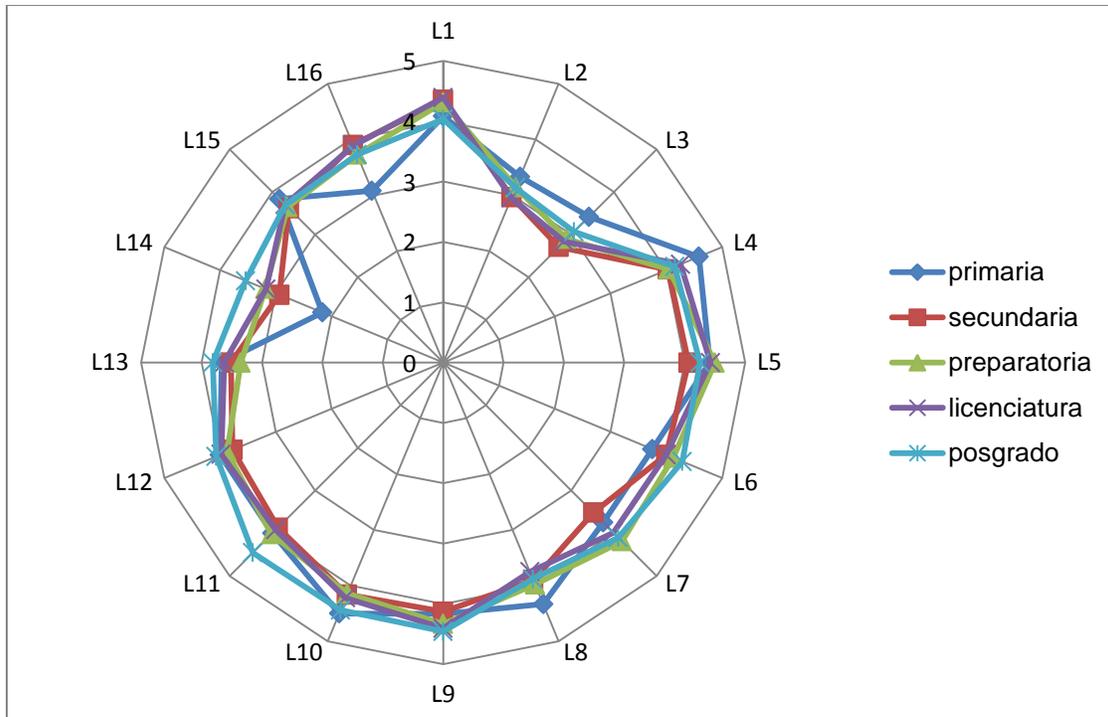


Figura 37 Respuestas por grado de estudios

La Figura 38 **Error! Reference source not found.** muestra las respuestas al cuestionario según el nivel de ingreso socioeconómico en esta grafica podemos observar la semejanza de las repuestas en los distintos niveles, y también se puede apreciar que en general el promedio de respuesta por grupo es alto, recordando que la mayor puntuación de respuestas es 5, los tres grupos muestran un alto grado de responsabilidad y conciencia ambiental según la escala.

No obstante, es importante mencionar que en algunos ítems de la escala, las respuestas de los tres diferentes grupos por tienden a la actitud indiferente en los ítems 2, 3 y 14, de los cuales como se menciono anterior mente el ítem número dos afirma que la persona en cuestión al comprar productos los selecciona si tiene la menor cantidad de empaques, el ítem 3 se refiere a transportarse en automóvil propio incluso si el transporte público fuera adecuado, y el ítem numero 13 se refiere a la importancia que le da la persona encuestada da a mantener aparatos de televisión en las recamaras de los miembros de la familia.

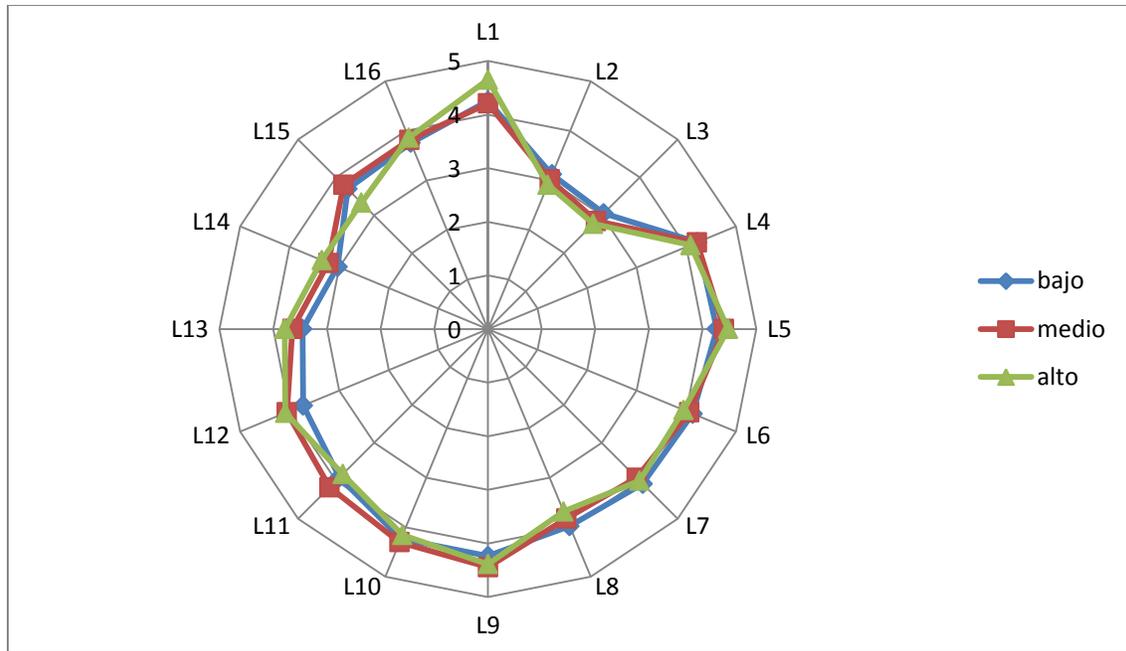


Figura 38 Respuestas por ingreso

El gráfico que se muestra a continuación (ver Figura 39 **Error! Reference source not found.**) muestra las respuestas a la escala aplicada según el tipo de familia al que el individuo pertenece, podemos apreciar en la figura que la tendencia de respuestas es muy similar para la mayoría de los ítems, nuevamente nos encontramos la reacción indiferente para los ítems 2,3 y 14.

Se observa una diferencia en el promedio de respuestas al ítem 16, que afirma la preferencia por lavar la ropa varias veces a la semana para que esta no se acumule el fin de semana, como es de esperarse entre menos miembros integren la familia tienden a lavar menos veces en la semana, por lo tanto el grupo correspondiente a familia extendida tiene puntuación más baja, lo que significa que lava más veces en la semana.

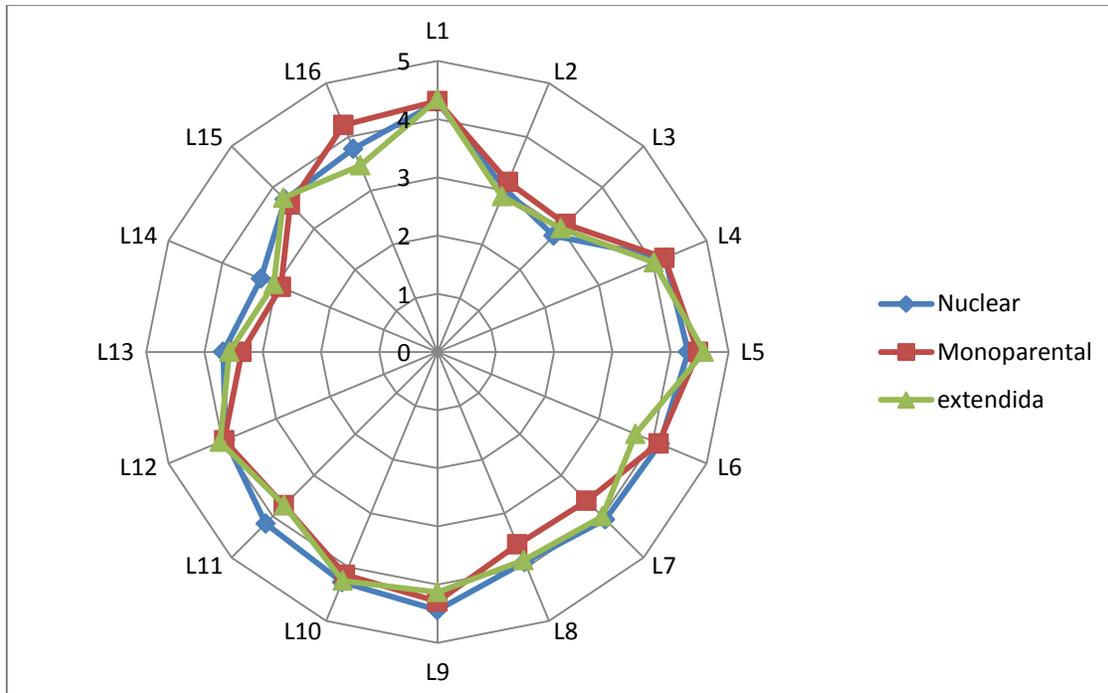


Figura 39 Respuestas por tipo de familia

### 4.3 Análisis de Componentes Principales

Con la información capturada de la encuesta, para cada familia se generó un vector numérico de información de 21 dimensiones. Las primeras cuatro dimensiones corresponden a un valor que representa las características generales de la familia, las siguientes dieciséis corresponden a un valor que representa la respuesta de cada uno de los ítems de actitudes afectivas, cognitiva y conductual, y la última dimensión corresponde a la cantidad de gas refrigerante emitido. Los vectores resultantes fueron utilizados para realizar un Análisis de Componentes Principales (PCA) y ver si existen patrones estructurales y tendencias numéricas en los datos, así como vislumbrar posibles cúmulos de información. Para realizar este análisis se utilizó el programa estadístico R.

La figura 40 muestra una gráfica del Componente Principal 1 contra el Componente principal 2 (PC1 vs PC2). Para verificar si existe diferenciación basada en la cantidad de fugas de gas refrigerante, se dividió el rango de valores de gas registrado en las encuestas en tres intervalos como sigue: los valores de fugas de 0, 1, 2 y 3 kg fueron asignados como cantidad de fugas baja y se les etiquetó con la letra **B**, los valores de fugas de 4, 5, 6 y 7 kg fueron asignados como cantidad de fugas media y se les etiquetó con una **M**, y los valores de fugas de 8, 9, y 10 kg fueron asignados como cantidad de fugas alta y se les etiquetó con una **A**.

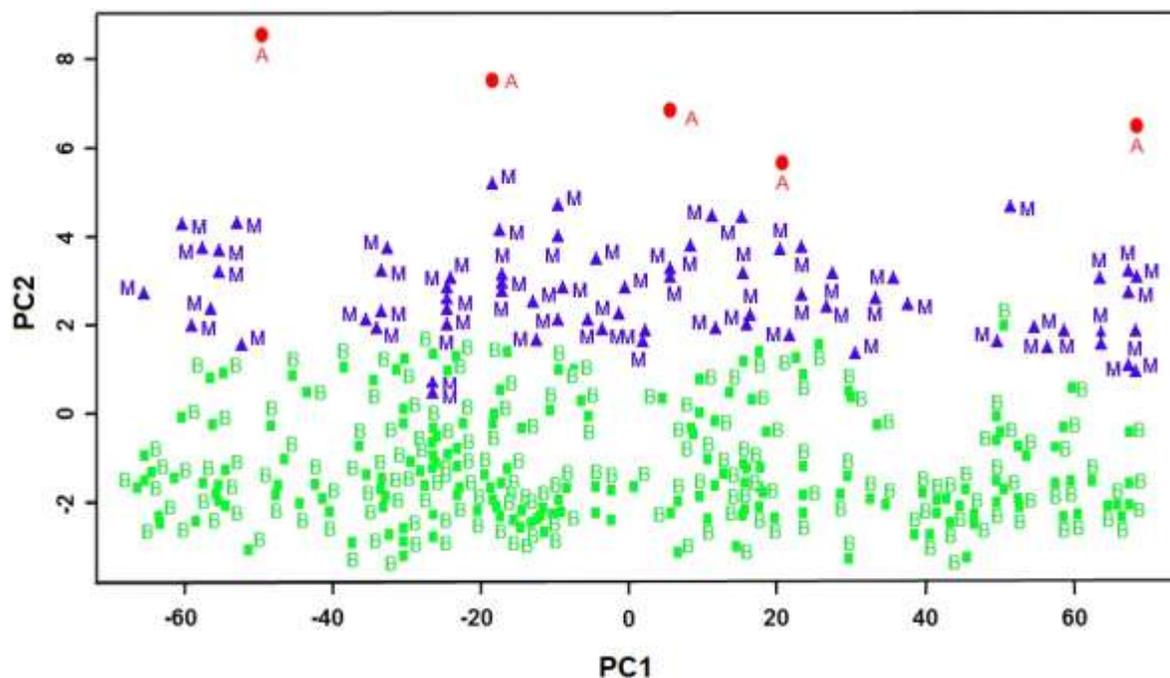


Figura 40 Análisis de Componentes Principales. La figura muestra una clara diferenciación entre familias con fugas clasificadas como Bajas (0, 1, 2 y 3 kg), Medias (4, 5, 6, y 7 kg) y Altas (8, 9 y 10 kg).

En la Figura 40 se observa una clara diferenciación entre las familias que tuvieron los tres tipos de fugas, si observamos los datos desde el eje PC2. Este componente principal, aun cuando no es el que captura la mayor variabilidad de los datos, es el que logra una clara separación entre los vectores, basado en los tres intervalos de las fugas de refrigerante. Como se observa en la gráfica, las familias con una cantidad de fugas clasificada como alta (A) resultaron con una

carga mayor a 5.5 en PC2, las familias con una cantidad de fugas clasificada como media (M) resultaron con una carga entre 1.5 y 5.4, con algunas familias apareciendo entrelazadas con familias con fugas consideradas bajas (B), las cuales resultaron con carga entre 0 y 1.8.

La Figura 41 muestra la gráfica de componentes principales para la variable Tipo de familia, como se puede observar en la figura se la estratificación de la variable en la muestra.

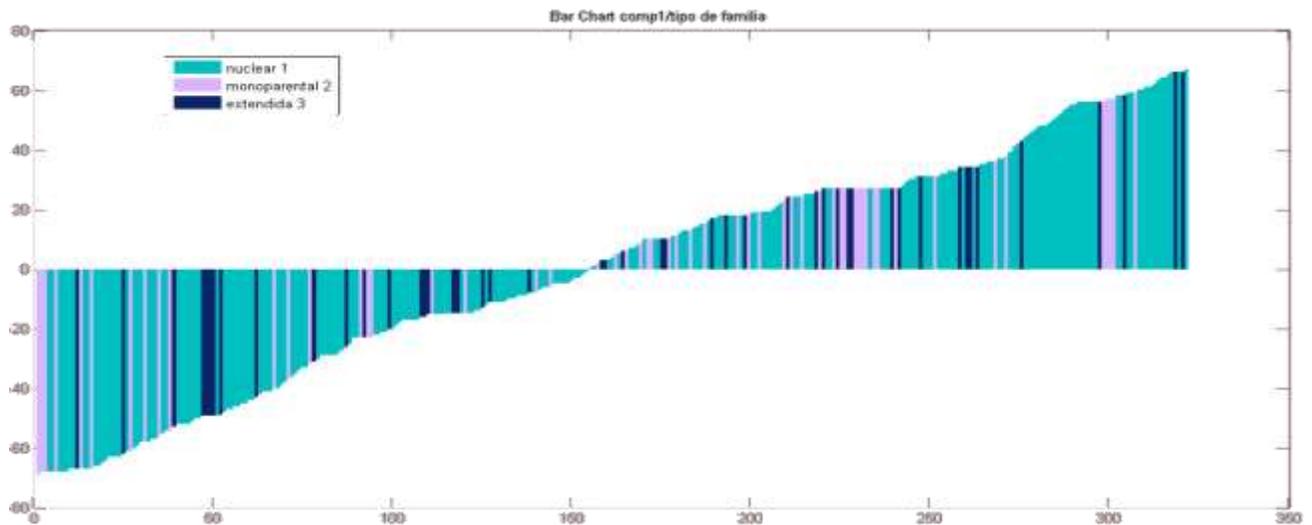


Figura 41 Componentes Principales para la Variable Tipo de Familia

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la variable Escolaridad, como se puede observar en la figura se la estratificación de la variable en la muestra.

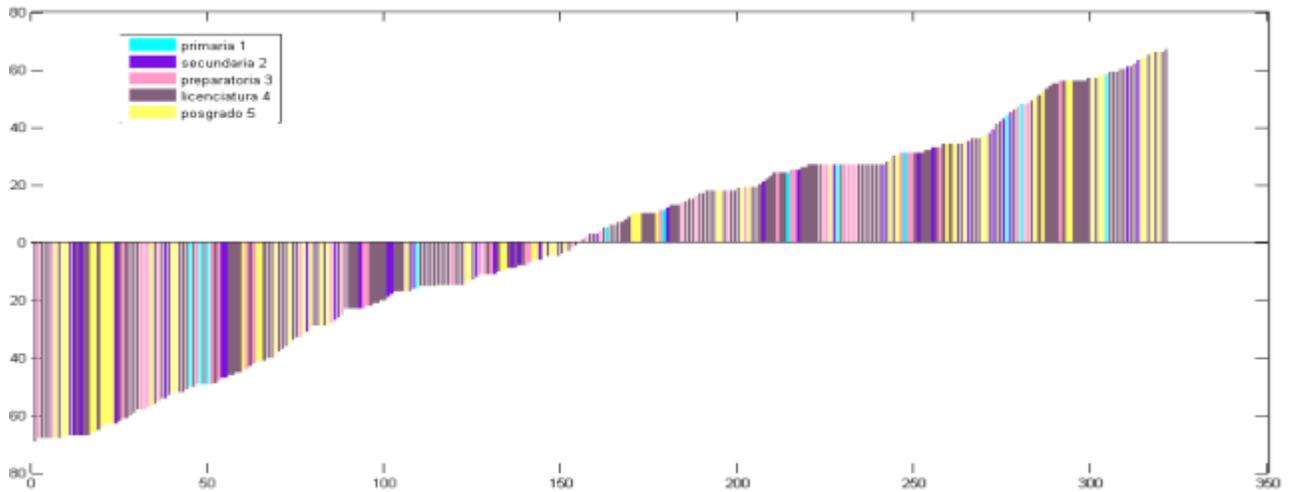


Figura 42 Componentes Principales para la Variable Escolaridad

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la variable Ingreso, como se puede observar en la figura se la estratificación de la variable en la muestra.

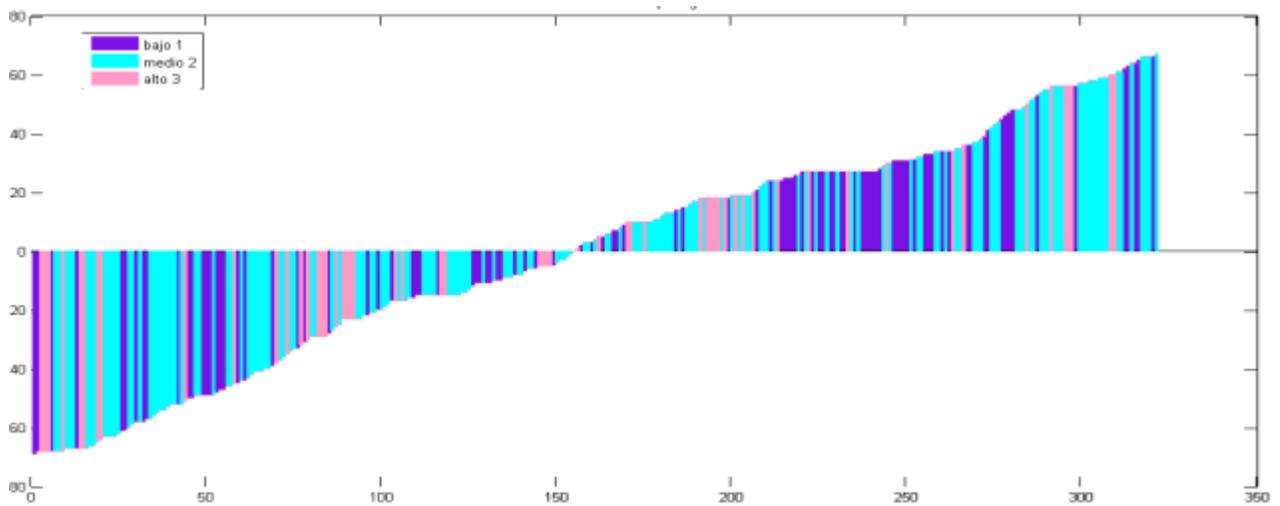
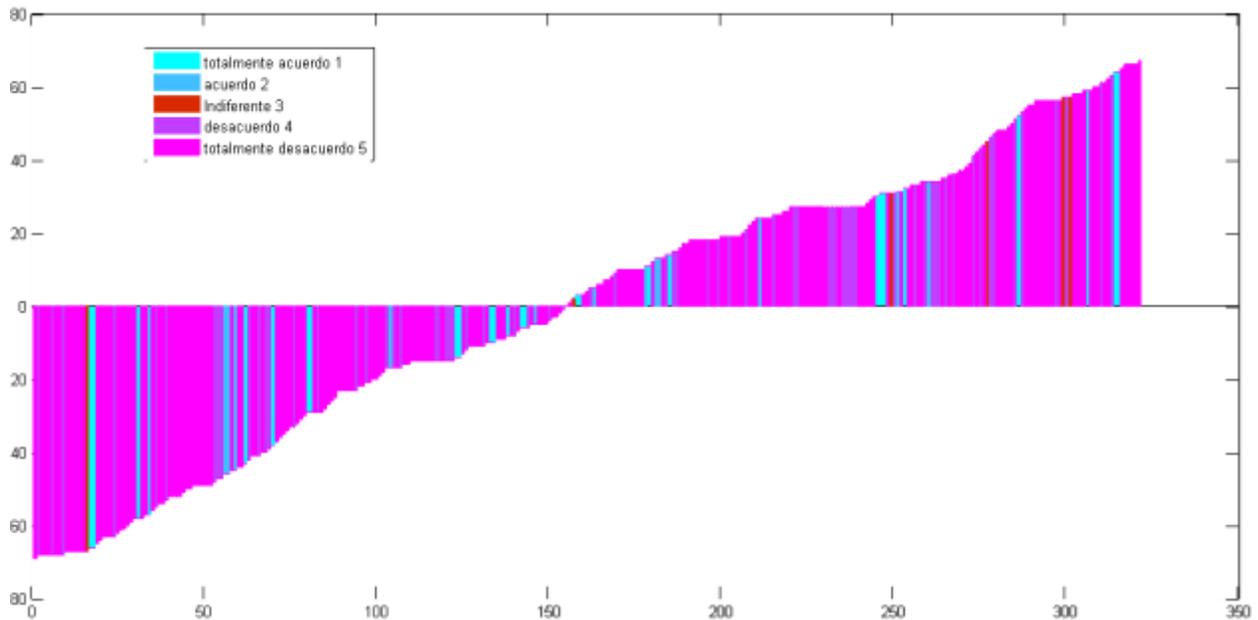


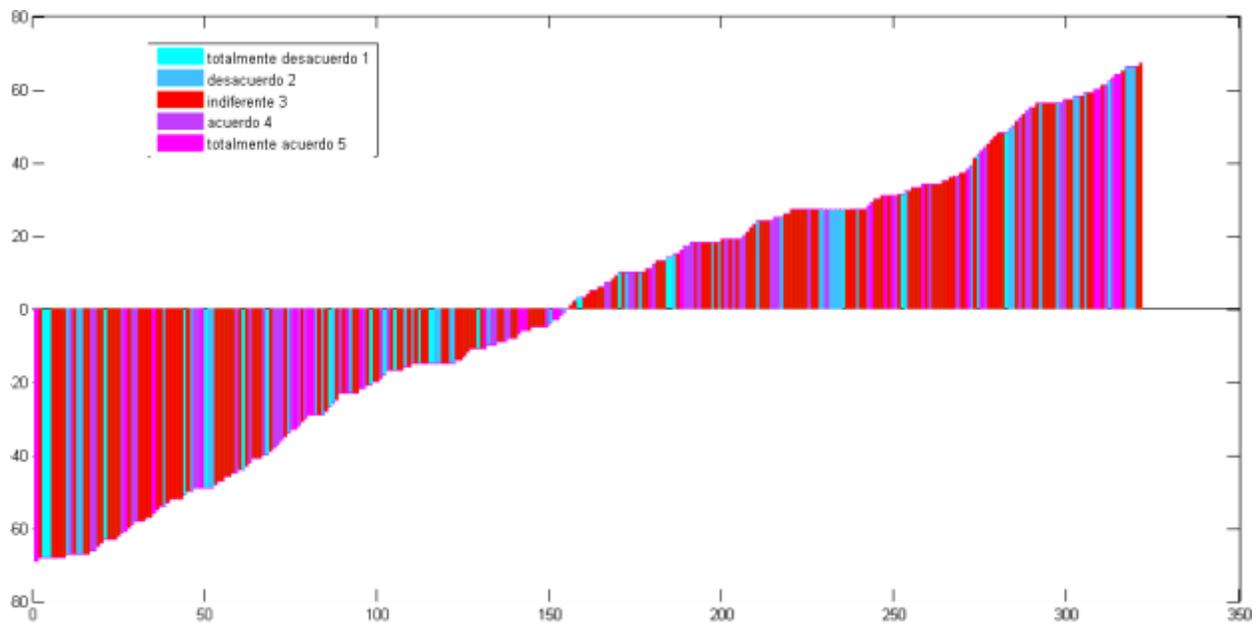
Figura 43 Componentes Principales Para la variable Ingreso

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la primer pregunta de la encuesta en escala de likert, la cual es parte del área afectiva y tiene un sesgo negativo, como se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



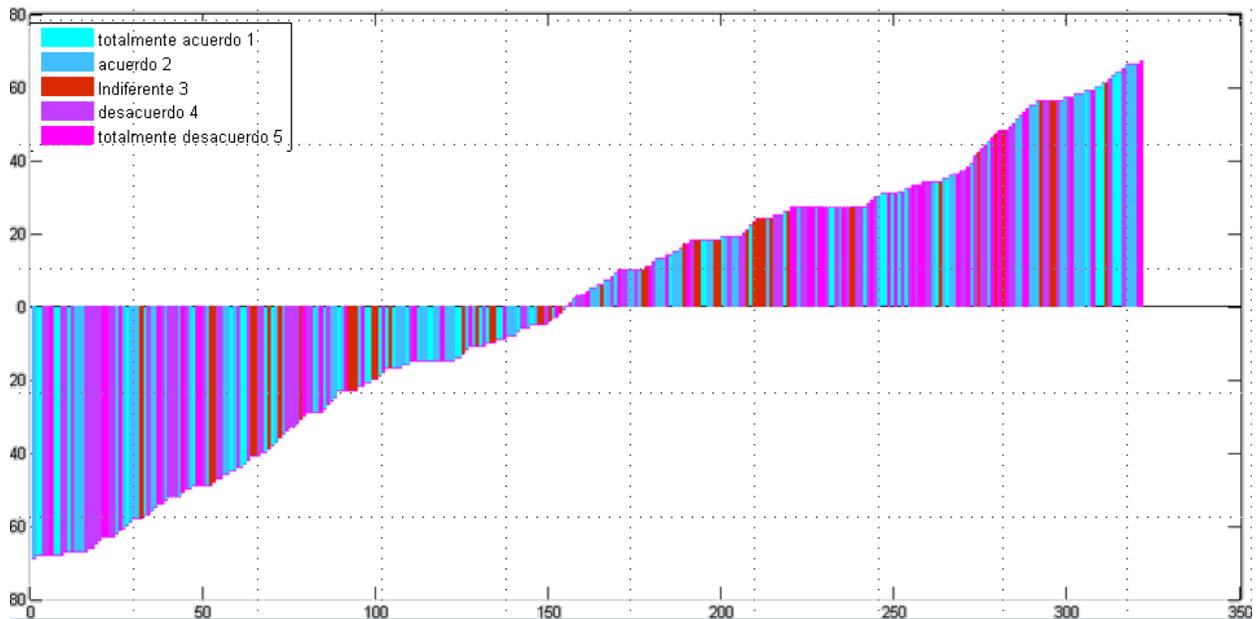
**Figura 44 Componentes principales para la pregunta 1.- La contaminación no afecta personalmente mi vida**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta dos de la encuesta en escala de likert, la cual es parte del área afectiva y tiene un sesgo positivo, de igual forma se puede observar no se presentan cúmulos o agrupaciones en las respuestas, cabe mencionar el alto grado de indiferencia que la muestra encuestada refleja hacia esta pregunta



**Figura 45 Componentes Principales para la pregunta 2.- Antes de comprar los productos que necesito los comparo para seleccionar el que contenga menor cantidad de empaques o envoltorios**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta tres de la encuesta en escala de likert, la cuál es parte del área afectiva y tiene un sesgo negativo, de igual forma se puede observar no se presentan cúmulos o agrupaciones en las respuestas, sin embargo es importante mencionar alto grado de indiferencia que la muestra encuestada refleja hacia esta pregunta



**Figura 46 Componentes Principales para la pregunta 3 Incluso si el transporte público fuera más eficiente yo preferiría llevar mi auto al trabajo**

La

muestra la gráfica de componentes principales para la pregunta cuatro de la encuesta en escala de likert, la cual es parte del área afectiva y tiene un sesgo positivo, de igual forma se puede observar no se presentan cúmulos o agrupaciones en las respuestas, sin embargo como se menciona anteriormente la población encuestada se muestra muy de acuerdo en establecer sanciones legales como una solución a los problemas ambientales.

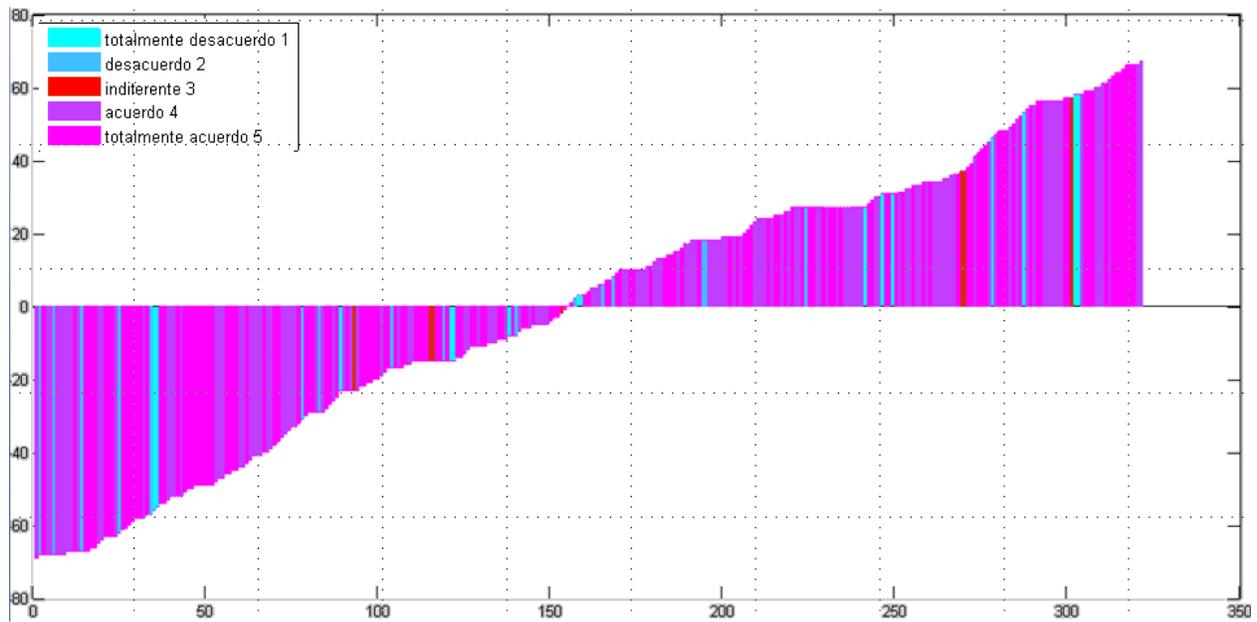
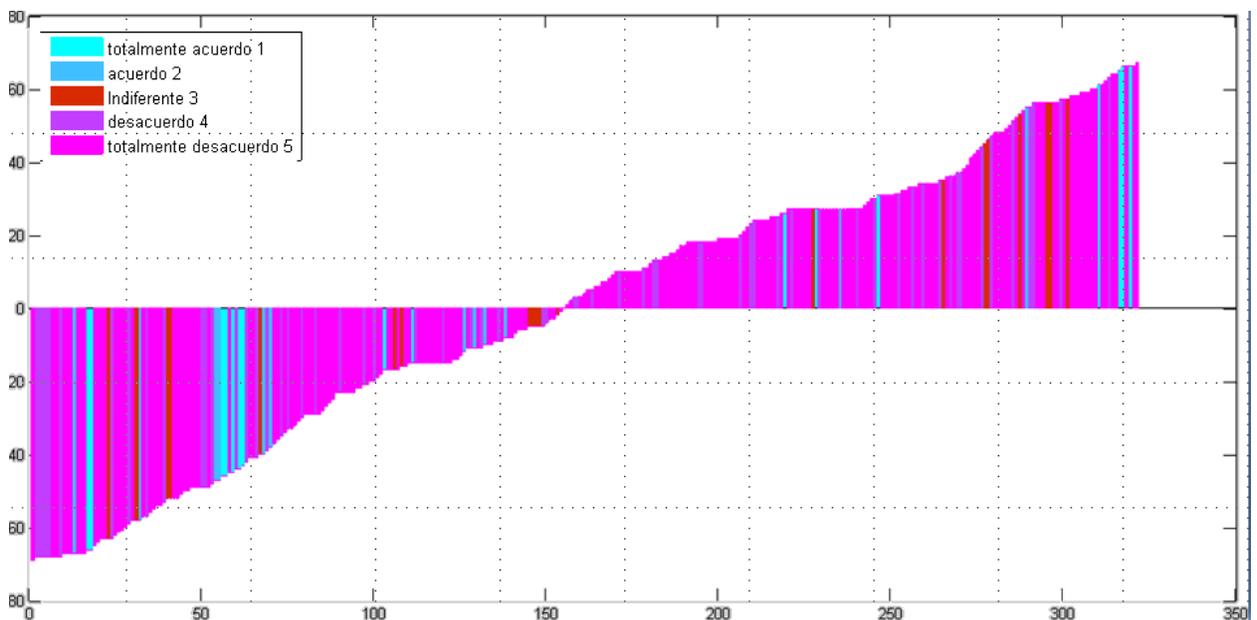


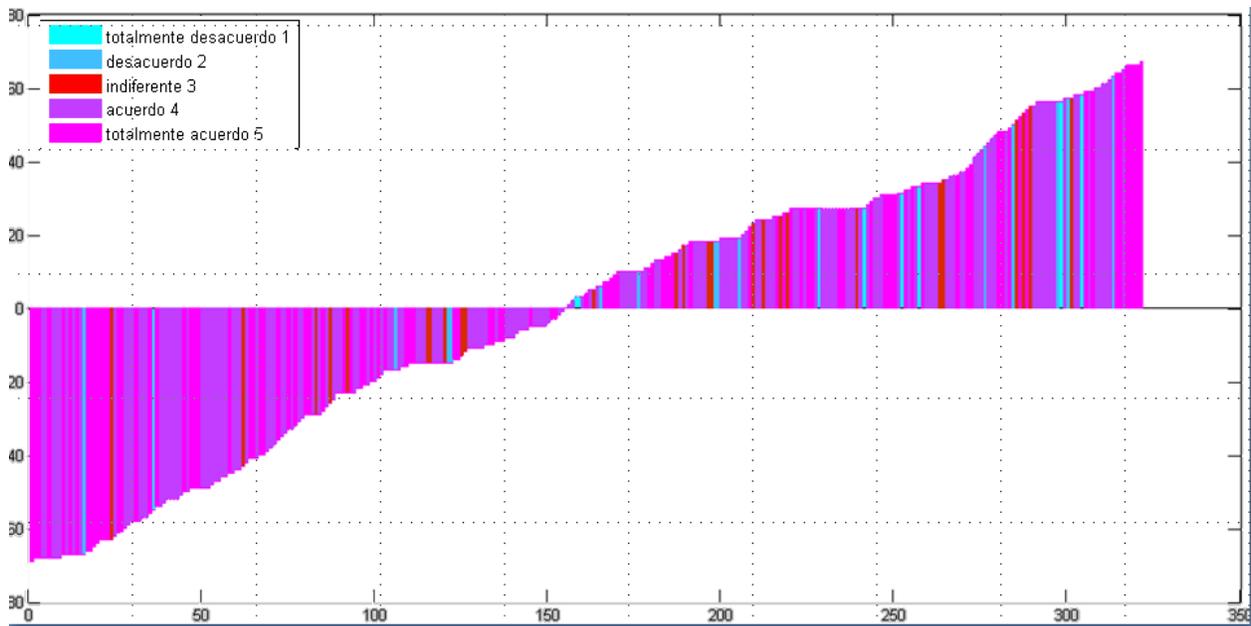
Figura 47 Componentes Principales para la pregunta 4 El gobierno tendrá que introducir duras medidas para frenar la contaminación ya que poca gente las regulara por si misma

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta cinco de la encuesta en escala de likert, la cual es parte del área afectiva y tiene un sesgo negativo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



**Figura 48 Componentes Principales para la pregunta 5.- No deberíamos preocuparnos por matar demasiados animales de caza porque a la larga las cosas se equilibrarán**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta seis de la encuesta en escala de likert, la cual es parte del área afectiva y tiene un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas, sin embargo se observa que gran cantidad de la muestra está dispuesta a realizar sacrificios personales con tal de reducir el ritmo de la contaminación.



**Figura 49 Componentes Principales para la pregunta 6.- Estaría dispuesto a hacer sacrificios personales para reducir el ritmo de la contaminación aunque los resultados inmediatos no puedan parecer significativos**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta siete de la encuesta en escala de likert, la cual es parte del área cognitiva y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

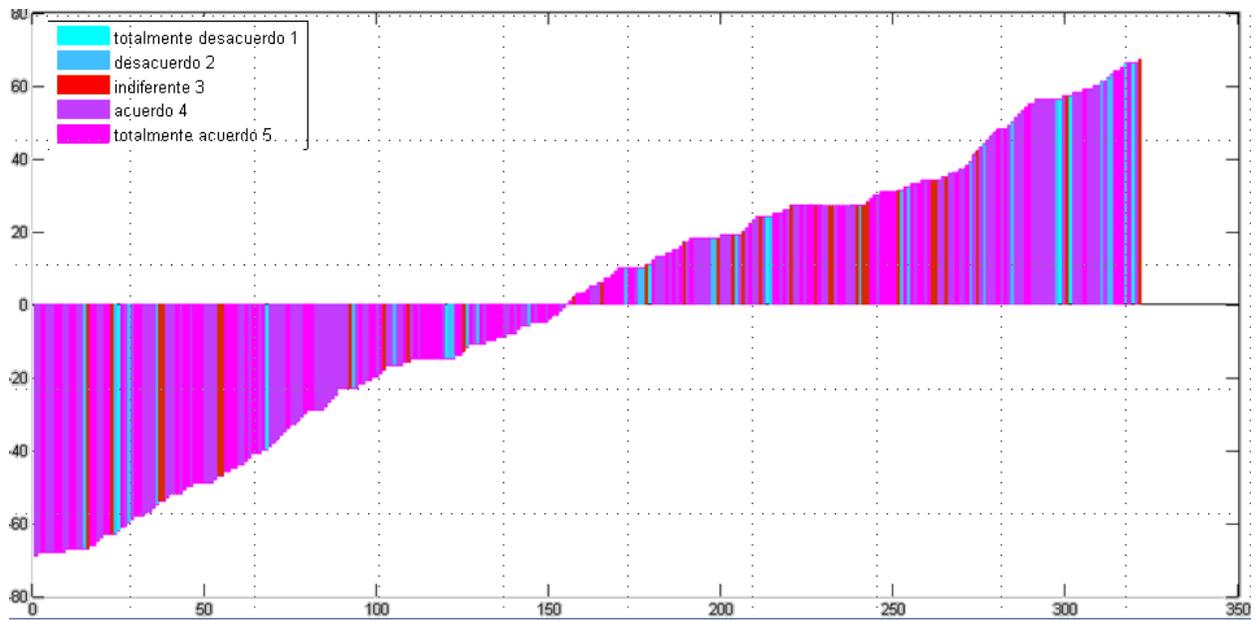


Figura 50 Componentes Principales para la pregunta 7.- Si el aparato de refrigeración presenta fuga considero necesario mandarla reparar aunque ya se vaya a acabar la temporada de verano

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 8 de la encuesta en escala de likert, la cual es parte del área cognitiva y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

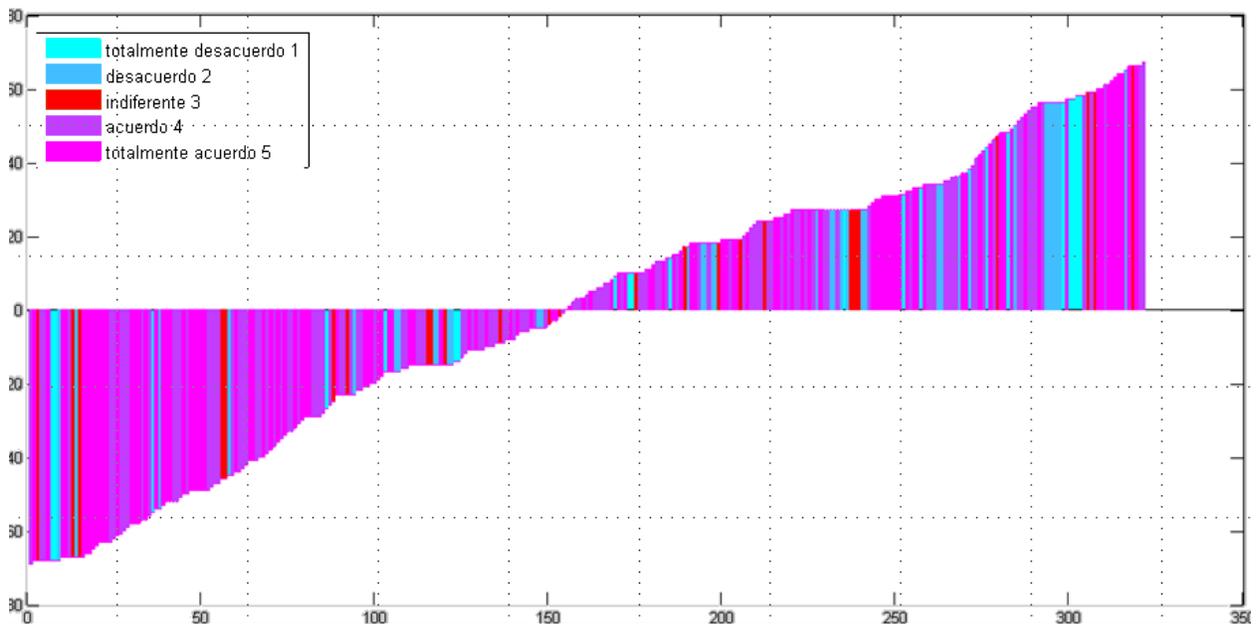


Figura 51 Componentes Principales para la pregunta 8.- Se debe utilizar el aparato de aire acondicionado solo el tiempo necesario para adquirir un confort térmico u no mantenerlo encendido todo el día

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 9 de la encuesta en escala de likert, la cual es parte del área cognitiva y presenta un sesgo negativo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

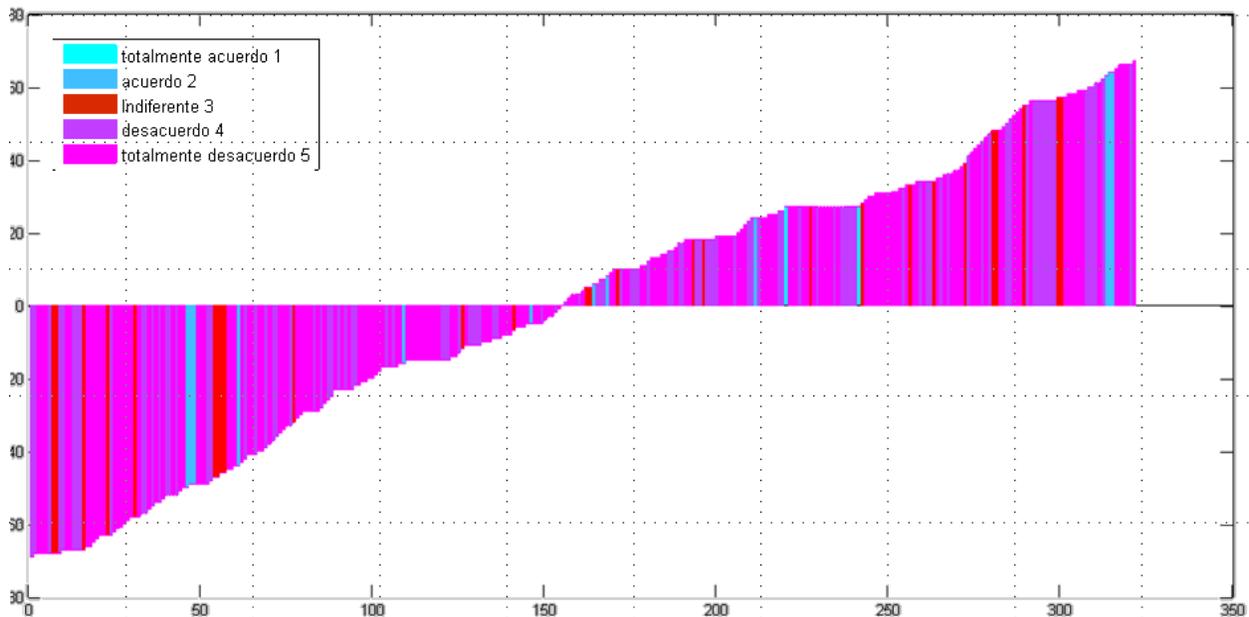


Figura 52 Componentes Principales para la pregunta 9.- Prefiero comprar un aparato de refrigeración usado ya que su costo es menor.

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 10 de la encuesta en escala de likert, la cual es parte del área cognitiva y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

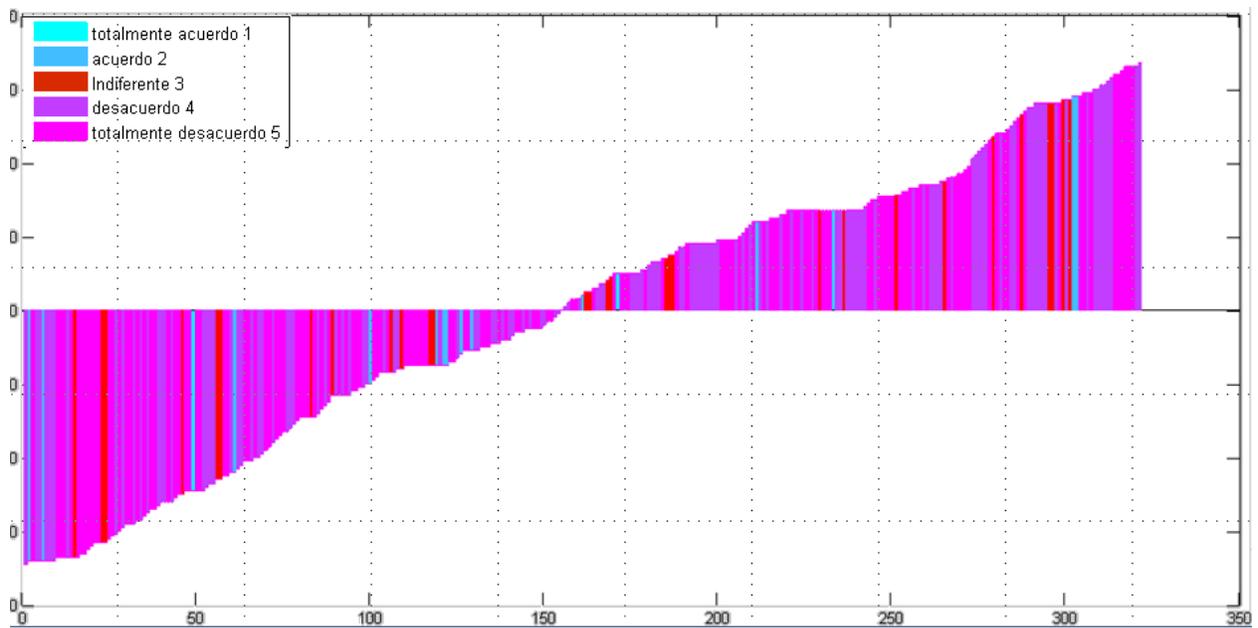
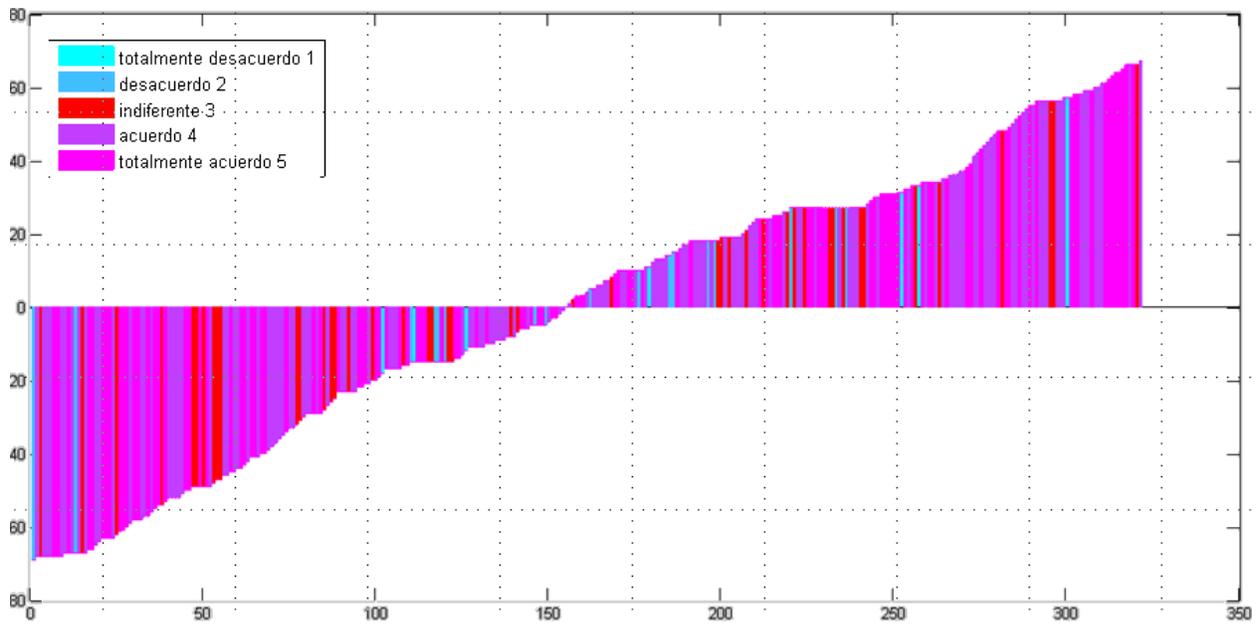


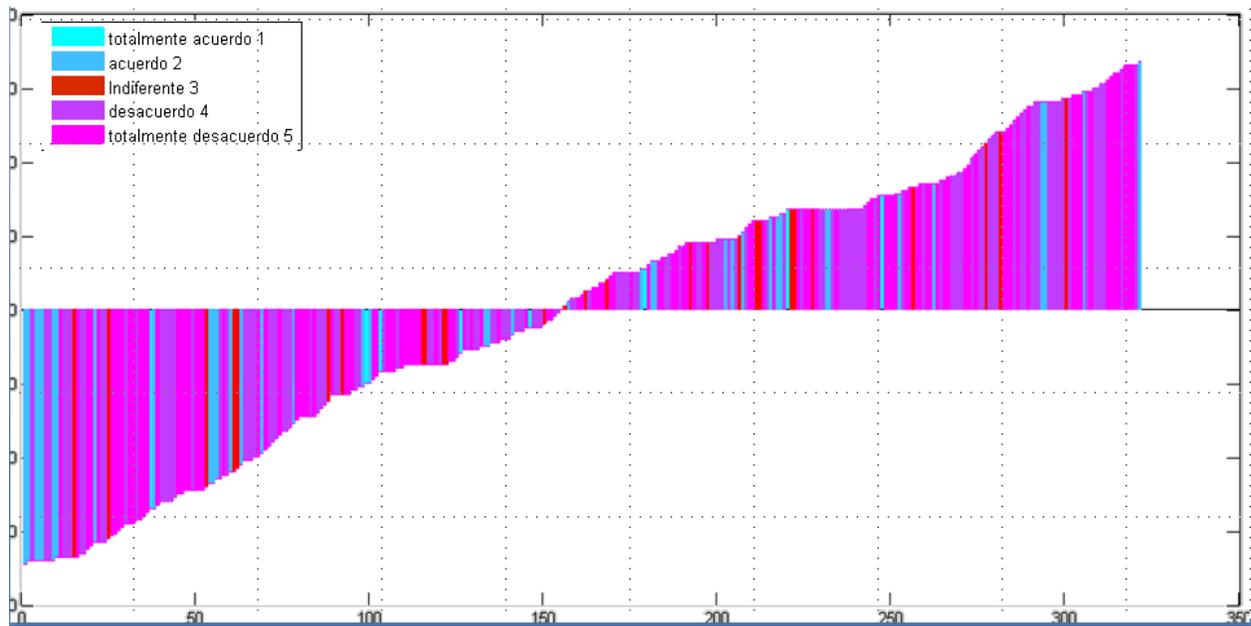
Figura 53 Componentes Principales para la pregunta 10.- El gas que fuga mi aparato de refrigeración es muy poco como para que impacte en el ambiente

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 11 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



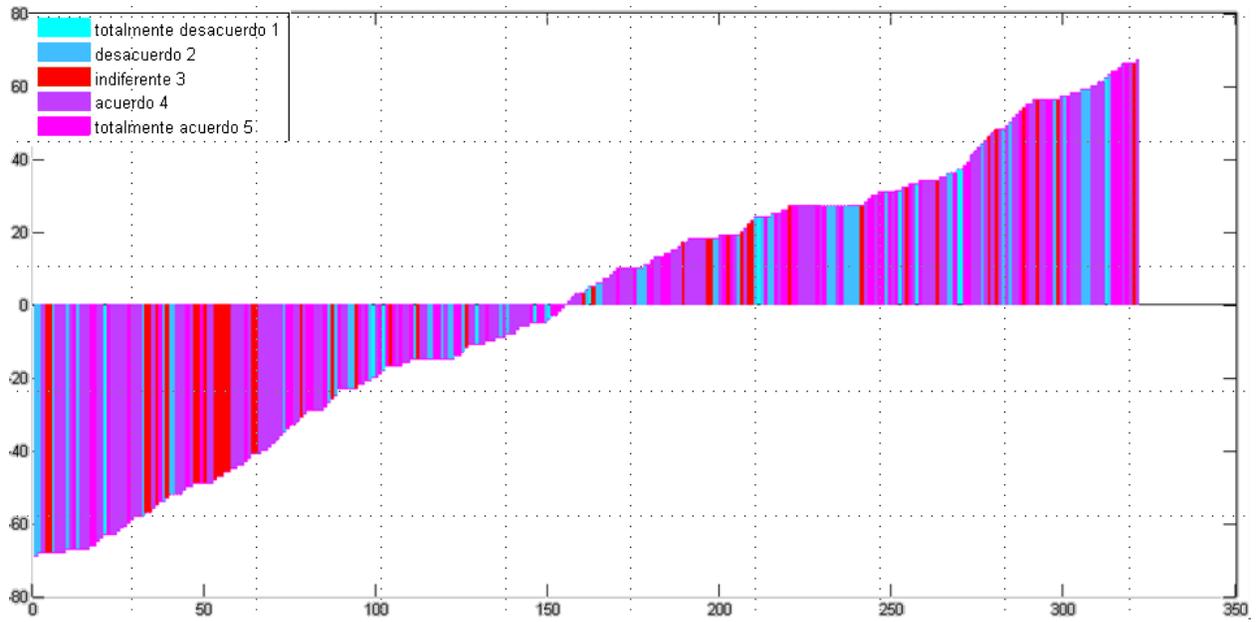
**Figura 54 Componentes Principales para la pregunta 11.-Es una costumbre familiar tener al menos un día de convivencia en familia**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 12 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo negativo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



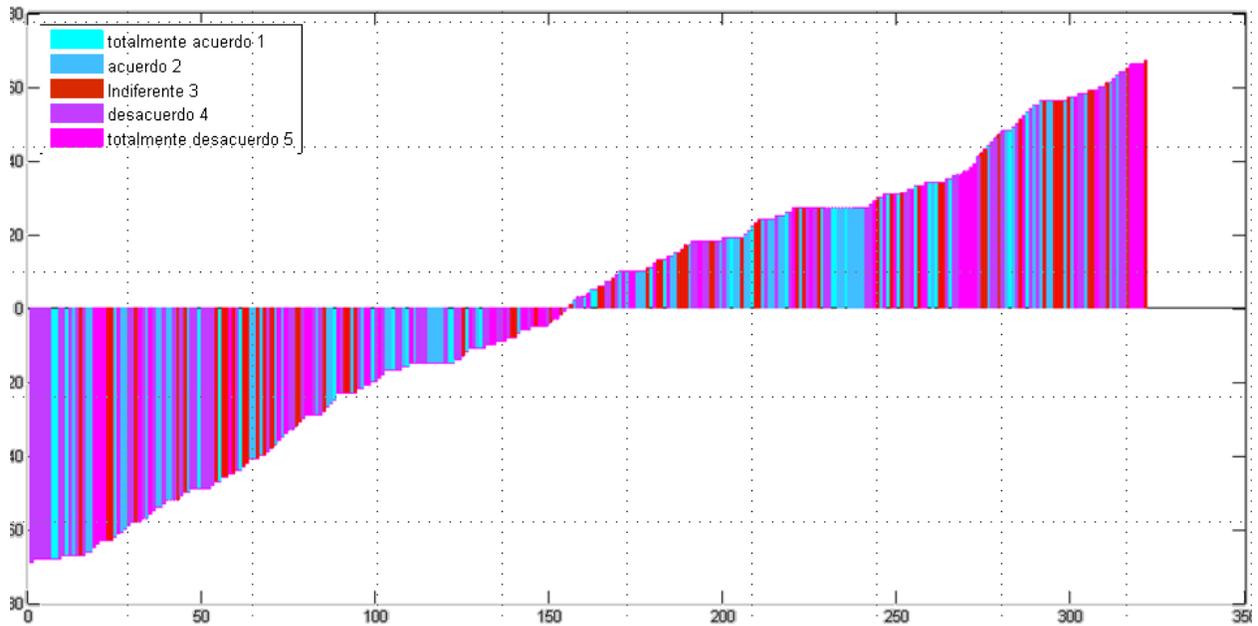
**Figura 55 Componentes Principales para la pregunta 12.- La vida diaria es muy atareada y no hay tiempo para actividades familiares.**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 13 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



**Figura 56 Componentes Principales para la pregunta 13.- Cada cuatro meses aproximadamente le doy servicio a mi automóvil para mantenerlo en buenas condiciones**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 14 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo negativo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.



**Figura 57 Componentes Principales para la pregunta 14.- Es preferible tener una televisión en cada recámara para evitar discusiones sobre la programación.**

La **Error! Reference source not found.** muestra la gráfica de componentes principales para la pregunta 15 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo positivo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

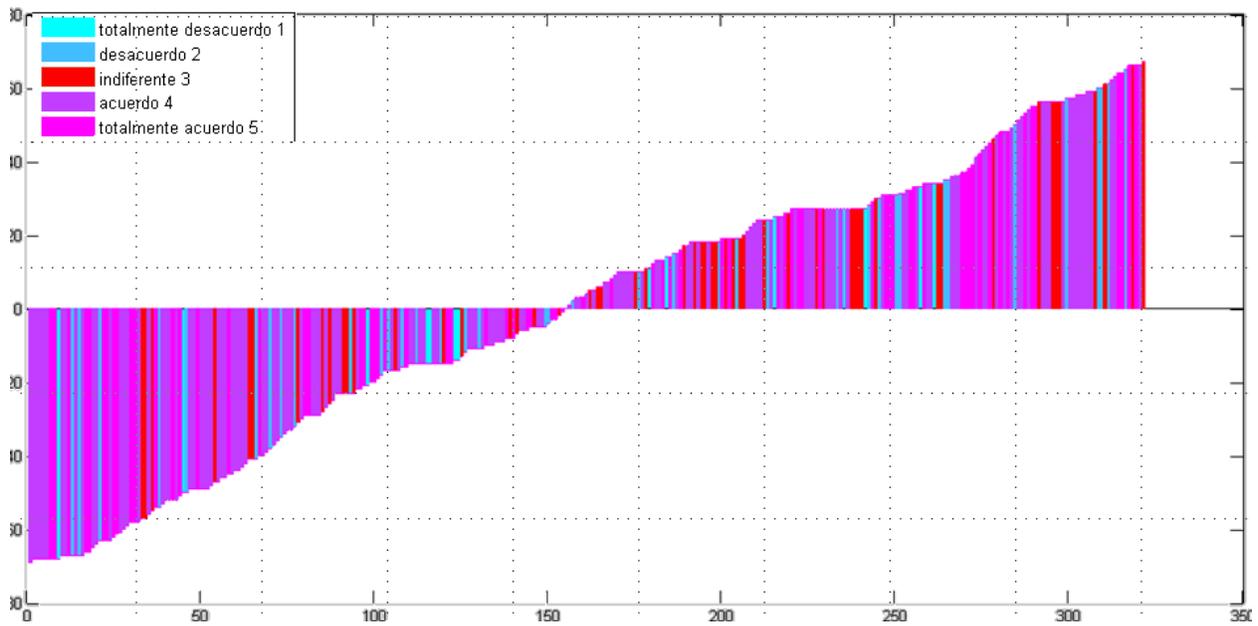


Figura 58 Componentes Principales para la pregunta 15.-Cuando voy a comprar adquiero solo lo que necesito.

La **Error! Reference source not found.59** muestra la gráfica de componentes principales para la pregunta 16 de la encuesta en escala de likert, la cual es parte del área conductual y presenta un sesgo negativo, se puede observar no se presentan cúmulos o agrupaciones en las respuestas.

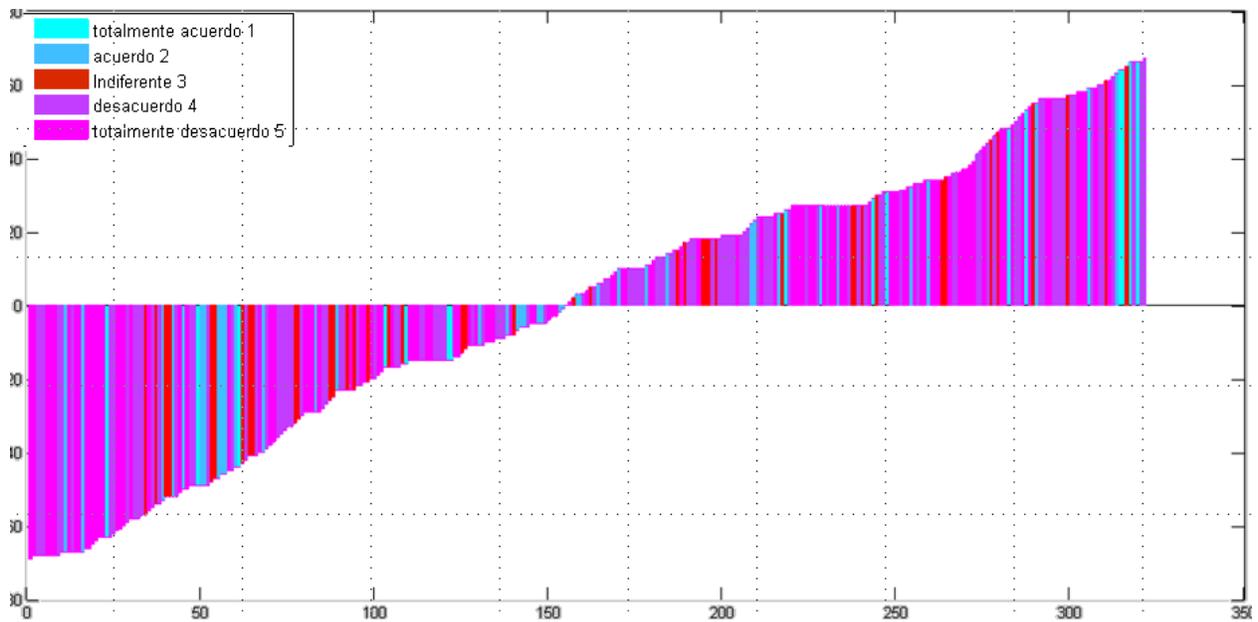


Figura 59 Componentes Principales para la pregunta 16.- Para que la ropa sucia no se acumule el fin de semana es preferible lavar diario cantidades pequeñas.

#### 4.4 Redes Neuronales

Existen diversas arquitecturas de RNA (Cios & Pedrycz, 2007) sin embargo, en este proyecto se utilizó la arquitectura del perceptrón multicapa; también llamada Arquitectura feedforward-backpropagation, por el algoritmo de entrenamiento que estas utilizan, ver **Error! Reference source not found.** Aprovechando las bondades de este tipo de redes, este estudio se enfoca en utilizar una red neuronal artificial para aproximar emisión en de refrigerante de una familia dadas ciertas condiciones y determinar si habrá emisión o no.

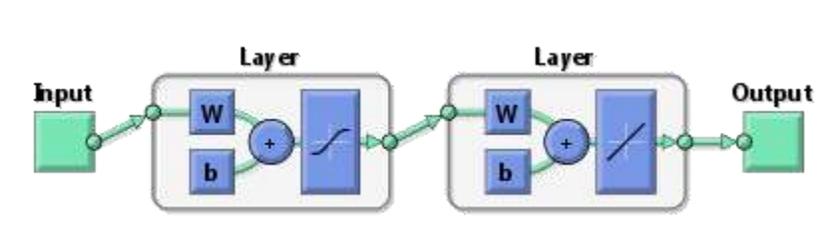


Figura 60 Configuración RNA utilizada

#### 4.4.1 Desempeño de la RNA con todas las variables de la encuesta para estimar la emisión

La medida de desempeño para establecer el comportamiento de la RNA fue el Error Cuadrático Medio (MSE) definido en (4). En la Tabla 4 podemos ver los resultados medios –calculados a partir de las 10 ejecuciones para el modelo. En esta se puede observar un buen desempeño entre los error de predicción de *training*, *validation* y de *test*.

$$MSE = \frac{1}{N} \sum_i (d_i - y_i)^2 \quad (4)$$

Como se mencionó en el capítulo tercero de esta investigación, para realizar este análisis el primer paso fue construir los conjuntos de datos y dejarlos en un formato apropiado para su procesamiento. Para facilitar el análisis de las visualizaciones generadas por los mapas auto-organizativos se creó una matriz de datos en formato CSV en la que se manejan variables de la encuesta además de las variables descriptivas. En la Tabla 4 se muestra un segmento de la matriz final que se proceso por medio de las redes neuronales.

**Tabla 4 Matriz de datos estandarizados para la RNA**

0.13	0.63	0.5	0.38	0.63	0.38	0.13	0.63	0.38	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	10
0.13	0.5	0.5	0.25	0.63	0.38	0.13	0.5	0.63	0.5	0.5	0.63	0.5	0.5	0.38	0.5	0.5	0.5	0.63	0.5	10
0.13	0.63	0.63	0.38	0.63	0.25	0.5	0.63	0.63	0.25	0.25	0.25	0.5	0.5	0.5	0.63	0.25	0.5	0.25	0.5	10
0.13	0.5	0.5	0.25	0.63	0.38	0.38	0.63	0.5	0.5	0.5	0.5	0.63	0.25	0.63	0.25	0.5	0.5	0.63	0.5	8
0.13	0.5	0.88	0.38	0.63	0.13	0.5	0.5	0.5	0.5	0.5	0.5	0.63	0.5	0.5	0.25	0.38	0.5	0.5	0.5	8
0.13	0.5	0.88	0.38	0.63	0.13	0.5	0.5	0.5	0.5	0.5	0.5	0.63	0.5	0.5	0.25	0.38	0.5	0.5	0.5	8
0.25	0.63	0.13	0.25	0.63	0.38	0.38	0.5	0.63	0.5	0.25	0.5	0.5	0.5	0.63	0.5	0.25	0.25	0.5	0.5	7
0.13	0.5	0.38	0.25	0.63	0.5	0.63	0.5	0.63	0.63	0.63	0.5	0.63	0.63	0.63	0.5	0.5	0.38	0.5	0.5	7
0.13	0.5	0.63	0.25	0.63	0.38	0.25	0.63	0.63	0.5	0.25	0.63	0.63	0.63	0.5	0.25	0.5	0.25	0.25	0.63	7
0.13	0.5	0.5	0.38	0.63	0.13	0.63	0.63	0.63	0.5	0.63	0.5	0.63	0.63	0.5	0.5	0.5	0.5	0.63	0.63	7
0.13	0.38	0.38	0.25	0.63	0.38	0.63	0.25	0.63	0.5	0.5	0.63	0.63	0.25	0.63	0.63	0.25	0.25	0.25	0.5	6.5
0.13	0.5	0.63	0.25	0.63	0.5	0.38	0.63	0.63	0.5	0.5	0.63	0.63	0.5	0.63	0.5	0.5	0.25	0.25	0.63	6.5
0.38	0.13	0.25	0.25	0.63	0.38	0.25	0.5	0.63	0.13	0.5	0.5	0.63	0.63	0.5	0.63	0.5	0.25	0.5	0.63	6
0.13	0.5	0.63	0.25	0.63	0.13	0.38	0.5	0.25	0.5	0.25	0.63	0.5	0.25	0.63	0.63	0.13	0.25	0.5	0.63	6
0.13	0.5	0.63	0.25	0.63	0.38	0.25	0.63	0.63	0.5	0.63	0.63	0.63	0.63	0.38	0.5	0.5	0.38	0.38	0.63	6
0.13	0.5	0.5	0.38	0.25	0.38	0.25	0.5	0.38	0.5	0.5	0.5	0.25	0.5	0.25	0.13	0.13	0.38	0.38	0.63	5
0.13	0.5	0.38	0.13	0.63	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.25	0.5	0.25	0.5	0.25	0.5	5
0.25	0.5	0.38	0.13	0.63	0.5	0.25	0.5	0.63	0.13	0.13	0.13	0.5	0.63	0.5	0.5	0.38	0.25	0.5	0.5	5
0.38	0.5	1	0.13	0.63	0.38	0.38	0.63	0.63	0.63	0.63	0.63	0.63	0.5	0.38	0.13	0.13	0.13	0.63	0.63	5
0.13	0.25	0.5	0.25	0.5	0.38	0.5	0.5	0.5	0.38	0.25	0.5	0.5	0.38	0.5	0.5	0.5	0.38	0.5	0.5	5
0.13	0.38	0.5	0.13	0.5	0.13	0.63	0.63	0.63	0.63	0.63	0.5	0.63	0.63	0.5	0.63	0.63	0.63	0.13	0.25	5
0.13	0.63	0.38	0.38	0.63	0.38	0.25	0.5	0.5	0.5	0.5	0.5	0.63	0.63	0.5	0.5	0.25	0.25	0.25	0.5	5
0.25	0.38	0.25	0.25	0.63	0.5	0.5	0.5	0.63	0.5	0.5	0.5	0.63	0.63	0.63	0.5	0.63	0.5	0.5	0.5	5
0.38	0.5	0.75	0.13	0.63	0.38	0.13	0.63	0.63	0.63	0.63	0.63	0.5	0.5	0.5	0.25	0.63	0.13	0.63	0.13	5
0.13	0.25	0.38	0.25	0.63	0.13	0.5	0.63	0.63	0.5	0.5	0.13	0.63	0.63	0.5	0.5	0.25	0.25	0.5	0.5	5
0.13	0.5	0.25	0.38	0.63	0.38	0.25	0.63	0.63	0.5	0.25	0.38	0.5	0.38	0.38	0.38	0.5	0.38	0.25	0.63	5
0.13	0.5	0.63	0.25	0.63	0.13	0.13	0.63	0.5	0.5	0.5	0.5	0.63	0.63	0.63	0.63	0.38	0.5	0.25	0.63	5
0.13	0.25	0.25	0.38	0.63	0.25	0.25	0.25	0.63	0.63	0.5	0.25	0.5	0.63	0.5	0.5	0.5	0.5	0.5	0.63	5

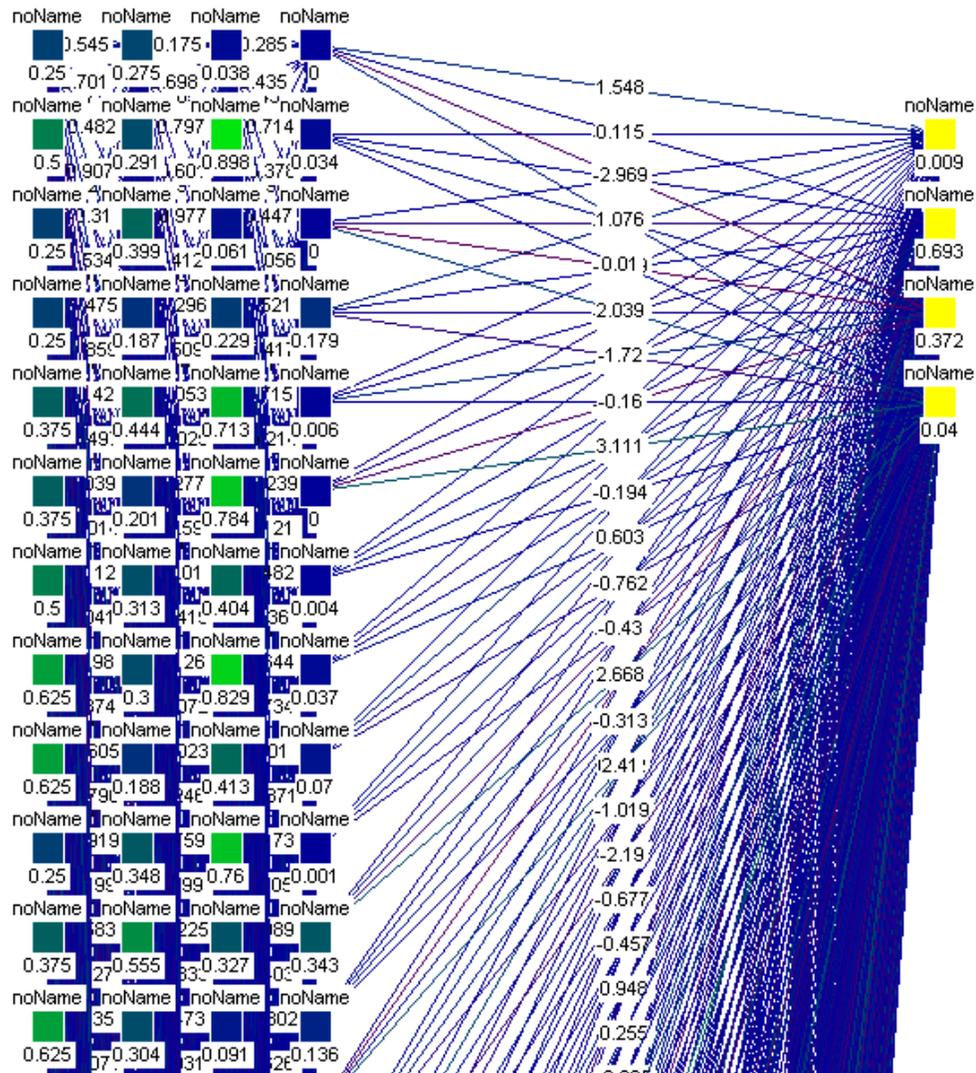
Dicho lo anterior, se procedió a ingresar este conjunto de datos tratados a la red neuronal para que en base a las características de cada caso pueda estimar la emisión de refrigerante de la familia encuestada. Se realizó una separación aleatoria de datos para crear dos subconjuntos, uno para entrenamiento con el 70% de los datos, otro grupo prueba con el 15% y otro conjunto para validación con el 15% restante. Los parámetros utilizados para el entrenamiento de la RNA se enlistan en la siguiente tabla (

Tabla 5):

Tabla 5 Parámetros de entrenamiento de la RNA

Nombre	Formula	
<i>Función de Activación interna</i>		
Hyperbolic Tangent Sigmon	$\theta(r) = \frac{2}{1 + e^{-r}} - 1$	(5)
<i>Función de Activación de Salida</i>		
Linear Transfer	$\theta(i) = \frac{Y_i}{X_i}$	(6)
<i>Actualización de Pesos</i>		
Gradient descent weight	$\delta(z) = \theta(\beta \cdot (\sum_i w_{ij} \cdot \alpha_i))$	(6)
<i>Taza de Aprendizaje</i>		
Bias learning	0.01	
Performance measure	Error Cuadrático Medio (ECM)	

Una vez establecido lo anterior se realizó el entrenamiento de una red neuronal **Error! Reference source not found.** La red recibe como entrada los valores de las variables para cada registro de información dado, y arroja una salida binaria que indica el nivel de emisión estimado por la RNA.



**Figura 61 RNA para estación de la emisión**

Se hizo un análisis de correlación simple (Lutz & Woodhouse, 1997) entre el valor de la predicción de la red neuronal y el valor observado en los datos, para el modelo. En la Tabla 6 se la prueba de validación cruzada que se realizó en 10 ejecuciones para la red la cual muestra el error cuadrático medio al punto de cruce y la precisión promedio de estas ejecuciones la cual es 20.72%

Tabla 6 tabla general de validación cruzada

Validación	topología	no ciclos al punto	error cuadrático al punto de cruce	error de precisión sobre 43
V1	20-300-4	262	31	5
V2	20-300-4	270	30	11
V3	20-300-4	260	26	6
V4	20-300-4	250	30	13
V5	20-300-4	260	32	13
V6	20-300-4	250	35	13
V7	20-300-4	270	35	12
V8	20-300-4	252	32	7
V9	20-300-4	245	30	10
V10	20-300-4	280	28	8
		259.9	30.9	9.8
				22.79

La **Error! Reference source not found.** muestra los datos observados contra los datos predichos para el modelo de datos propuesto. Para la construcción de esta gráfica, se tomo una ejecución promedio de las 10 ejecuciones aleatorias realizadas, como podemos observar hay una diferencia significativa entre los datos reales y los datos predichos.

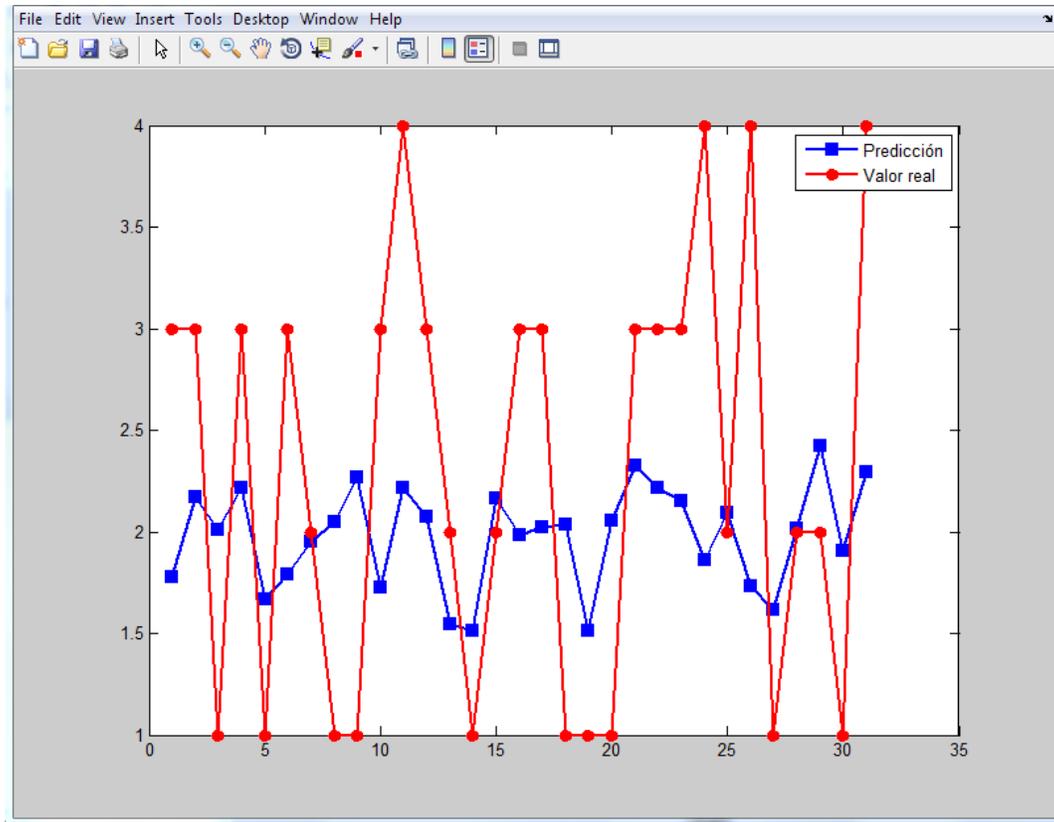


Figura 62 Datos reales contra datos predichos por la RNA

La Tabla 7 ver **Error! Reference source not found.** nos muestra el desempeño basado en el coeficiente de correlación R para los datos de entrenamiento, validación, prueba y regresión del modelo de RNA propuesto en la metodología. Estas tablas observan una ejecución obtenida de las 10 ejecuciones realizadas

Tabla 7 Coeficiente de relación R para el modelo de datos

	Entrenamiento	Validación	Prueba	Regresión
<b>Ejecución promedio</b>	0.32759	0.2634	0.52169	0.34454

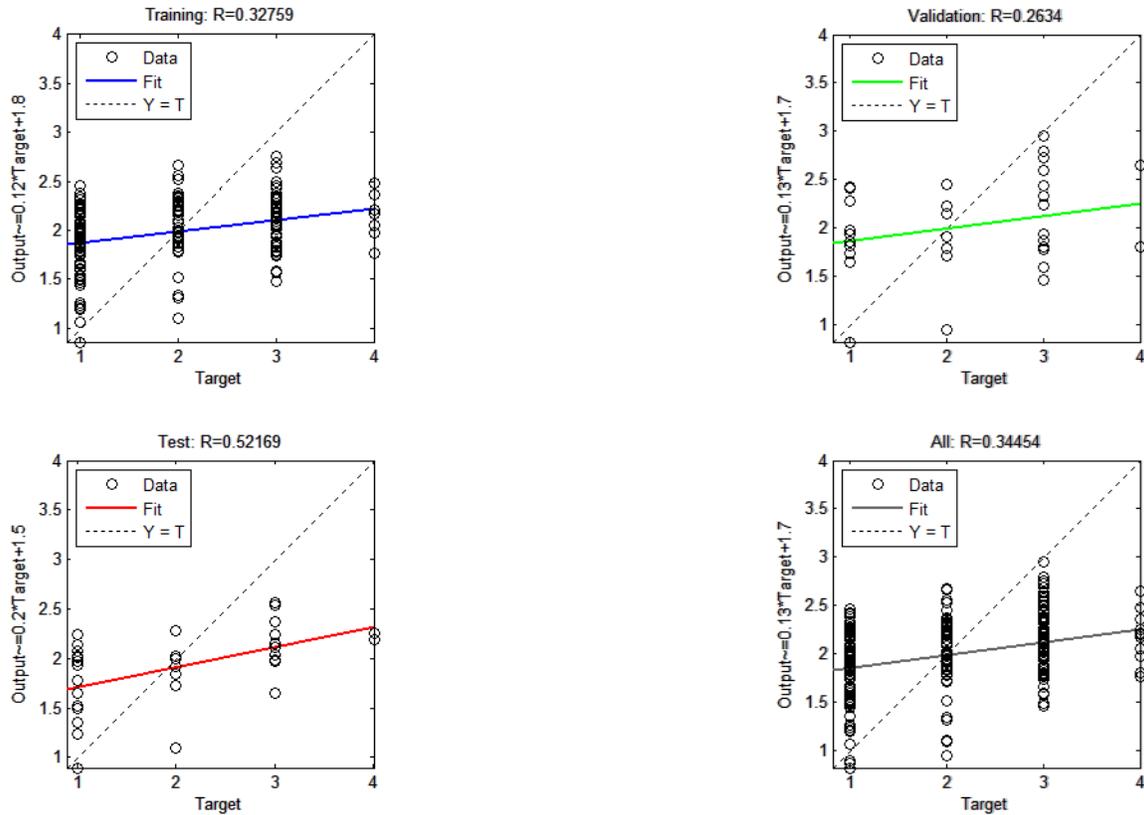


Figura 63 Regresión de entrenamiento, validación y prueba contra datos reales

La **Error! Reference source not found.** muestra el progreso del desempeño de la red, podemos observar que la medida de desempeño alcanzó su mínimo en la tercer iteración, la figura indica que no hay mayor problema con el entrenamiento, las curvas de validación y prueba son muy similares ya que si la curva de prueba aumentara significativamente sobre la curva de validación eso significaría un sobre entrenamiento.

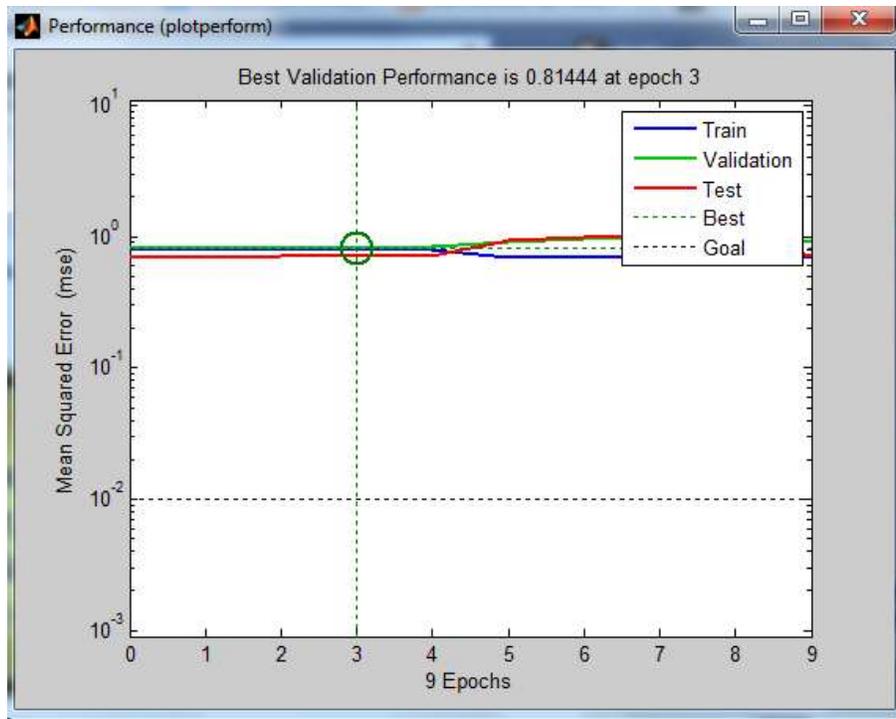


Figura 64 Progreso del desempeño de la red

#### 4.5 Árboles de Regresión

Para la aplicación del algoritmo CART utilizaremos la matriz de datos en formato CSV (comma separated values) con todos los registros que Si dieron mantenimiento al aparato de aire acondicionado, y todas las variables correspondientes al a encuesta likert mas las variables de clasificación, el objetivo será predecir la cantidad de gas emitido al medio ambiente basándonos en las variables antes mencionadas.

Se incluyó una etapa de reducción de dimensionalidad, seleccionando aquellos atributos que arrojaran mejor aproximación, esto se realizó mediante el conocido método *Backward Search*, búsqueda hacia atrás, que consiste en considerar todas las variables incluidas en el modelo e ir las eliminando una a una según su capacidad explicativa e ir probando la efectividad del

modelo con el nuevo set de variables hasta encontrar aquel que presente el mejor valor de desempeño (Lavalle, 2006).

Una vez terminado el proceso de la selección de atributos se obtuvo un nuevo conjunto de datos con las variables tipo de familia, y escolaridad, Tabla 8 muestra un fragmento de la tabla de datos utilizada para este análisis.

**Tabla 8 Fragmento de los datos utilizados para el árbol de decisión**

<b>tipo familia</b>	<b>escolaridad</b>	<b>emisión</b>
1	5	alto
1	4	alto
1	5	alto
1	4	alto
1	4	alto
1	4	alto
2	5	alto
1	4	alto
1	4	alto
1	4	alto
1	3	alto
1	4	alto
3	1	medio
1	4	medio
2	4	medio
3	4	medio
1	2	medio

1	3	medio
1	5	medio
2	3	medio
3	4	medio
1	2	medio
1	4	medio

La **Error! Reference source not found.** muestra el árbol generado por el algoritmo CART, el cual se genera mediante las reglas de decisión que se observan en la

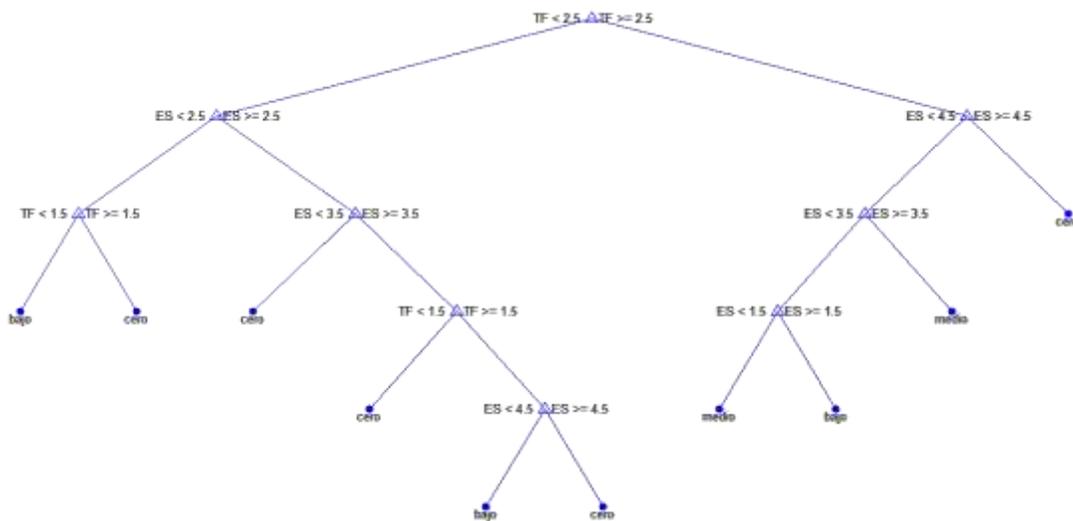


Figura 65 Árbol generado por el algoritmo CART

```
Decision tree for classification
1 if TF<2.5 then node 2 else node 3
2 if ES<2.5 then node 4 else node 5
3 if ES<4.5 then node 6 else node 7
4 if TF<1.5 then node 8 else node 9
5 if ES<3.5 then node 10 else node 11
6 if ES<3.5 then node 12 else node 13
7 class = cero
8 class = bajo
9 class = cero
10 class = cero
11 if TF<1.5 then node 14 else node 15
12 if ES<1.5 then node 16 else node 17
13 class = medio
14 class = cero
15 if ES<4.5 then node 18 else node 19
16 class = medio
17 class = bajo
18 class = bajo
19 class = cero
```

Figura 66 Reglas de decisión para el algoritmo CART

Una vez establecidas las reglas para la clasificación se procede a comparar los datos reales contra los datos estimados por el árbol de esta manera se obtuvo que el este método obtiene una precisión de 43%, la cual es significativamente baja para determinar la estimación sin embargo presenta un modelo más sencillo y preciso que la RNA presentada anteriormente.

## 4.6 Regresión Lineal

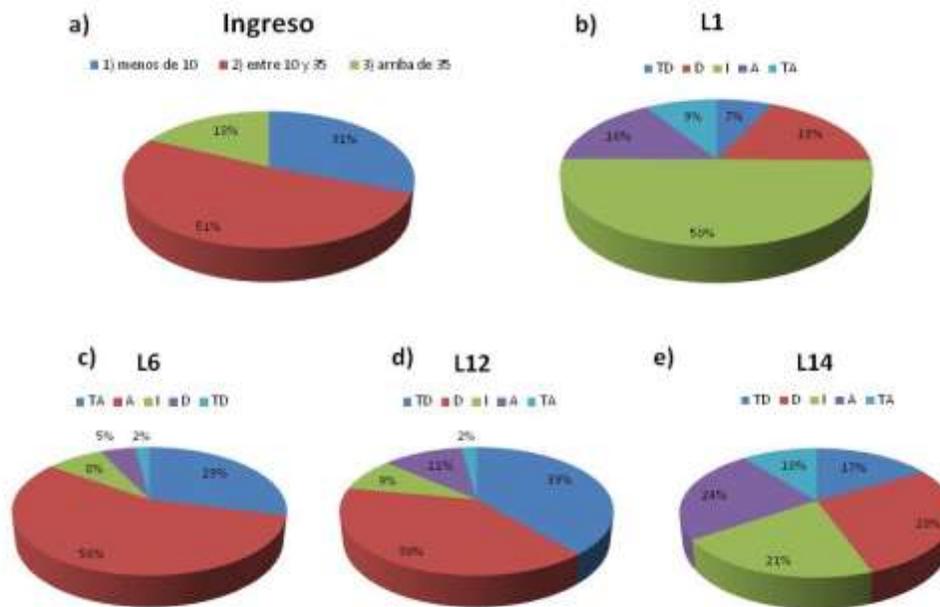
Para evaluar el efecto de cada variable sobre la cantidad de gas emitido, se aplicó regresión lineal múltiple al conjunto de vectores generado; tomando la cantidad de emisiones como proyectando, o variable dependiente,  $Y$ , y los factores conductuales, cognitivos, afectivos y generales como los proyectores,  $X$ 's, o variables independientes. El análisis se realizó con el software  $R$ . Una vez ajustado el modelo de regresión, se realizó una prueba de significancia estadística  $F$  para probar la calidad de ajuste de todas las variables al modelo, esta prueba produjo un valor  $p$  de 0.001377 por lo que en conjunto todas las variables consideradas de manera simultánea tienen un buen ajuste, por lo tanto el modelo es adecuado. Igualmente se aplicó una prueba  $t$  para probar la influencia de cada una de las variables en forma independiente. Los resultados de esta prueba se muestran en la tabla 9.

**Tabla 9 Resultados de la regresión lineal múltiple. Los variables de AFE1 a AFE6 corresponden al área afectiva, las variables de COG1 a COG4 corresponden al área cognitiva y las variables de CON1 a CON6 corresponden al área conductual**

Variable	Estimación	Error estándar	Estadística $t$	Valor $p$
<b>(Intersección)</b>	2.0830486	1.6013364	1.301	0.194477
<b>Fraccionamiento</b>	-0.0002778	0.0036257	-0.077	0.938994
<b>Tipo de familia</b>	0.0900632	0.189449	0.475	0.634906
<b>Escolaridad</b>	-0.1296093	0.1442021	-0.899	0.369592
<b>No. de habitantes</b>	0.1494119	0.0939862	1.59	0.113117
<b>Ingreso familiar</b>	0.5192687	0.2125722	2.443	0.015243
<b>AFE1</b>	0.3783191	0.1123451	3.367	0.000874
<b>AFE2</b>	0.026678	0.1477662	0.181	0.856868
<b>AFE3</b>	-0.0948748	0.1032238	-0.919	0.358889
<b>AFE4</b>	0.0992615	0.1388228	0.715	0.475239
<b>AFE5</b>	0.0548028	0.1282888	0.427	0.669601
<b>AFE6</b>	-0.3323593	0.1794436	-1.852	0.06514
<b>COG1</b>	-0.1017786	0.1422808	-0.715	0.475046
<b>COG2</b>	0.1101344	0.1169166	0.942	0.347075

<b>COG3</b>	-	0.1790721	-0.099	0.920873
	0.0178054			
<b>COG4</b>	-	0.1713526	-1.329	0.185059
	0.2277058			
<b>CON1</b>	0.119888	0.1685435	0.711	0.477528
<b>CON2</b>	-	0.1398132	-2.254	0.025001
	0.3152072			
<b>CON3</b>	-	0.1339889	-1.046	0.296441
	0.1401801			
<b>CON4</b>	0.2489166	0.1130414	2.202	0.028547
<b>CON5</b>	-0.076923	0.1280346	-0.601	0.5485
<b>CON6</b>	-	0.1155202	-0.12	0.904311
	0.0139011			

En las columnas se presenta la variable analizada, la estimación de su efecto sobre sobre la cantidad de fuga, el error estándar de la estimación, la estadística *t* y el valor *p* obtenido de la prueba estadística. El primer renglón presenta el valor de la intersección del modelo de regresión, los siguientes cuatro corresponden a las variables familiares, las siguientes seis a las variables afectivas, las siguientes cuatro a las variables cognitivas y las últimas seis a las variables conductuales. Podemos observar que cuatro variables presentan una influencia estadísticamente significativa (Ingreso familiar, AFE1, CON2 y CON4, con una valor  $p < 0.05$ ), y una presenta tendencia clara a la significancia estadística (AFE6, con un valor  $p = 0.06514$ ). Estas cinco variables fueron seleccionadas para los análisis subsecuentes debido a que son las que definen principalmente el resultado.



**Figura 67 Distribución de la muestra en las variables seleccionadas del modelo**

La figura 72 presenta la distribución de la muestra con respecto a cantidad de coincidencias en las respuestas obtenidas en la encuesta, para las cinco variables que resultaron más significantes. En la imagen 72a se observa la variable de *Ingreso Familiar*, la cual resultó con un coeficiente de 0.51927. Esta pertenece al segmento de caracterización familiar en cuestionario. La estratificación de las familias en este rubro fue la siguiente: 53% tienen un nivel socioeconómico medio, 27% tiene un nivel socioeconómico bajo y 20 % tienen un nivel socioeconómico alto. En esta figura observamos que la clase media es la que predomina en la muestra, con más del 50%, por lo que podemos ver que la clase media tiene la mayor influencia y define el comportamiento general de la emisión de gas en la ciudad Mexicali.

En la imagen 72b se observa la distribución de las respuestas a la variable de *Afectación Personal* (AFE1), la cual resultó con la mayor significancia estadística. El enunciado de esta pregunta en el cuestionario es “*la contaminación no afecta mi vida personal*” y el encuestado elige una de las siguientes cinco posibles respuestas: Totalmente en Desacuerdo (TD), Desacuerdo (D), Indiferente (I), De Acuerdo (A) y Totalmente de Acuerdo (TA). Esta variable resultó con un coeficiente de regresión de 0.37832 y pertenece al área afectiva del cuestionario. Como podemos observar en la imagen, el 9% de la población encuestada respondió que está totalmente de acuerdo con esta premisa, el 16% que simplemente está de acuerdo, el 50% se mostró indiferente, el 18% estuvo en desacuerdo y el 7% totalmente en desacuerdo. Este resultado es consistente con los resultados encontrados por Mapfre, 2011 (MAPFRE, 2011), en la ciudad de Madrid, España, la cual afirma que el 60% de los españoles se muestran indiferentes ante el cambio climático. Mapfre menciona que una posible explicación del por qué esta pregunta resulta ser una de las que presentan mayor influencia en la cantidad de emisión es el alto grado de indiferencia que presenta la población hacia el tema de la contaminación. Afirma que se aprecian dificultades teóricas y operativas en lo concerniente al tema de concientización sobre problemas que afectan nuestra salud o calidad de vida, y enfatiza la necesidad de implementar programas de concientización ambiental.

En la imagen 72c se observa la distribución de las respuestas a la variable de *Disposición Personal* (AFE6), la cual resultó con un coeficiente de regresión de -0.33236. El enunciado de esta pregunta en el cuestionario es: “*Estoy dispuesto a hacer sacrificios personales para reducir el ritmo de la contaminación aunque los resultados inmediatos no puedan parecer significativos*”. Esta pertenece al área afectiva del cuestionario, y las respuestas presentan la siguiente distribución: 2% están Totalmente en Desacuerdo, 5% están en Desacuerdo, 8% se muestran Indiferentes, 56% están De acuerdo y 26 % están Totalmente de Acuerdo. Este resultado muestra una alta disposición, por parte de los encuestados, a hacer sacrificios

personales para combatir la contaminación, sin embargo la variable anterior (AFE1) denota un alto grado de indiferencia; de esto se puede asumir que la sociedad está dispuesta a poner de su parte siempre y cuando existan guías o programas que le orienten e impulsen a hacerlo, lo cual refuerza la idea de que es necesaria la implementación de planes y programas de concientización y mejoramiento del medio ambiente. En general las personas encuestadas muestran disponibilidad para realizar acciones que disminuyan la contaminación y esa buena disposición puede ser la razón por la cual esta variable resulta estar entre las más significativas con respecto a la emisión de refrigerante.

En la imagen 72d se muestra la distribución de las respuestas a la variable *Actividad Familiar Personal* (CON2). Esta variable pertenece al área conductual del cuestionario, cuya pregunta es: “*La vida diaria es muy atareada y no hay tiempo para actividades familiares*”. Esta variable resultó con un coeficiente de regresión de -0.31521 y sus respuestas se distribuyen como sigue: 39% de las personas está Totalmente en Desacuerdo, 39% está en Desacuerdo, 24% se muestra Indiferente, 11% está de Acuerdo y 2% está Totalmente de Acuerdo. Esta pregunta está orientada a determinar hábitos en la vida cotidiana de las personas, la cual discurre a partir de prácticas privadas y públicas que remiten a patrones y estilos de vida que manifiestan nuestras maneras de pensar, sentir y de actuar. Las respuestas de la población encuestada reflejan una inclinación a la estabilidad familiar, hecho que puede ser aprovechado al momento de la elaboración de programas de concientización, enfatizando que el mejoramiento ambiental trae consigo un bienestar familiar general.

Por último, la imagen 72e muestra la distribución de las respuestas a la variable *Comodidad Personal* (CON4). Esta variable resultó con un coeficiente de regresión de 0.24892, pertenece al área conductual del cuestionario y el enunciado es la siguiente: “*Es preferible tener un televisor en cada recámara para evitar discusiones sobre la programación*”. Las respuestas se distribuyen de la siguiente manera: 17% están Totalmente en Desacuerdo, 28% están en

Desacuerdo, 21% se muestran Indiferente, 24% están de Acuerdo y 10 % están Totalmente de Acuerdo. El resultado muestra un patrón de comportamiento familiar en apoyo a su economía, al compartir el aparato de televisión, lo cual es consistente con la idea de hacer sacrificios personales para apoyar el bienestar familiar general, y que influye de manera importante en el patrón general de la emisión de refrigerantes.

## Capítulo 5 Conclusiones

---

En este estudio se muestrearon 321 familias de la ciudad de Mexicali Baja, California, México, midiendo la cantidad de gas refrigerante fugado por sus aparatos de refrigeración en el periodo comprendido entre septiembre de 2011 y abril de 2012, y se aplicó una encuesta diseñada para medir aspectos estructurales, afectivos, cognitivos y conductuales de las familias. Los datos fueron analizados utilizando técnicas de detección de cúmulos de información y regresión lineal con el fin de encontrar patrones numéricos que relacionen los hábitos socioeconómicos y culturales de las familias con las fugas en sus aparatos de aire acondicionado.

Con los datos de fugas recabados, y asumiendo un comportamiento homogéneo en la ciudad de Mexicali, estimamos que los hogares emiten aproximadamente 413,714 kg de refrigerante a la atmosfera en una temporada. El análisis de cúmulos de información permitió detectar diferenciación entre las familias, al clasificar la cantidad de gas fugado como Baja, Media y Alta. El análisis con regresión lineal múltiple mostró cinco variables con influencia estadísticamente significativa en la cantidad de fugas de gas. Un resultado relevante es que las variables involucradas de una forma estadísticamente significativa pertenecen al rubro de estructura socioeconómica, al de dimensión afectiva y al de aspectos conductuales, mientras que las variables del rubro de aspectos cognitivos no muestra una significancia estadística, sugiriendo no ser relevantes. Consecuentemente, de los resultados del estudio podemos inferir que: mientras gran parte de la población encuestada piensa que la contaminación no afecta su vida personal, al mismo tiempo muestra una gran disposición para hacer sacrificios personales y reducir el ritmo de la contaminación. Los hábitos personales, que reflejan el estilo de vida de los encuestados, muestran inclinación a la estabilidad familiar, tanto en lo económico como en el ambiente de convivencia, hecho que puede ser aprovechado al momento de la elaboración de

programas de concientización, enfatizando que el mejoramiento ambiental trae consigo un bienestar familiar general.

El uso de cuestionarios con la escala de tipo Likert, resulta ser un instrumento muy útil para evaluación, ya que mediante él se puede definir la actitud de cierto grupo de personas con respecto a un tema en específico, como en nuestro caso la actitud humana de diversas familias de la ciudad de Mexicali, además de reconocer diferencias entre distintos tipos de familias. Podemos decir que se trata de una técnica que permite obtener información de un modo rápido, y se puede aplicar tanto para reconocer la situación inicial, como para averiguar los resultados de actividades de sensibilización ante la problemática presentada. Es importante señalar que puede obtenerse la misma puntuación con distintas combinaciones de ítems, lo que demuestra que el mismo puntaje puede significar resultados distintos, para esto es importante un buen análisis de resultados así como la presencia de ítems de validación dentro del mismo cuestionario. Como aspectos positivos de este tipo de escalas es que son de fácil construcción y aplicación además de que se pueden incluir ítems no relacionados directamente con el tema pero que a su vez manifiesten mucho del perfil de la familia encuestada.

## Referencias

---

Aragonés, J., & Amérigo, M. (1991). estudio empírico sobre las actitudes ambientales. *Revista de psicología social*, 3(2), 223-240.

Barrigón, J., Vílchez, R., & Gómez, V. (2002). Presentación de una encuesta para la realización de estudios sociales sobre el impacto del ruido urbano. *Revista acústica*, 27-33.

Berenguer, J., & Corraliza, J. (2000). Preocupación ambiental y comportamientos ecológicos. *Psicothema*, 12, 325-329.

Bonson, E. (1999). *Tecnologías inteligentes para la gestión empresarial*. ra-ma.

- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees*. Pacific Grove, Cal: Wadsworth & Brooks.
- Campos, M., Pascali, C., & Peinado, S. (2008). Evaluación psicométrica de un instrumento de medición de actitudes pro-ambientales en escolares venezolanos. *Paradigma*, 135-156.
- Castro, H., Contreras, M., & González, A. (1999). *Programa para mejorar la calidad del aire de Mexicali 2000-2005*. Mexicali.
- Chairez, I., Fuentes, R., Poznyac, T., & Franco, M. (2010). Numerical modeling of the benzene reaction with ozone in gas phase using differential neural networks. *Catalysis Today*, 159-165.
- Chattopadhyay, G., & Chattopadhyay, S. (2009). Autoregressive forecast of monthly total ozone concentration: A neurocomputing approach. *Computers & Geosciences*, 1925-1932.
- Cios, K., & Pedrycz, W. (2007). *Data Mining a Knowledge Discovery Approach*. Springer Science+Business Media.
- Coman, A., Ionescu, A., & Candau, Y. (2008). Hourly ozone prediction for a 24-h horizon using neural networks. *Environmental Modelling & Software*, 1407-1421.
- De'Ath, G. (2002). Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology*, 1105-1117.
- De'Ath, G., & Fabricius, K. (2000). Combining Field Surveys, Remote Sensing, and Regression Trees to Understand Yield Variations in an Irrigated Wheat Landscape. *Ecology*, 3178-3192.
- Fernandez de Pinedo, I. (2007). *Construcción de una escala de actitudes tipo Likert*. Barcelona, España: Centro de Investigación y Asistencia Técnica.
- Fernandez, R., & Hueto, A. (2003). ¿Qué miden las escalas de actitudes? Análisis de un ejemplo para conocer la actitud hacia los residuos urbanos. *Ecosistemas*.
- Gardner, M., & Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*, 2627-2636.

- González, A., & Américo, M. (1999). Actitudes hacia el medio ambiente y conducta ecológica. *Psicothema*, 11(1), 13-25.
- Holger, M., & Graeme, D. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 101-124.
- Inal, F. (2010). Artificial Neural Network Prediction of Tropospheric Ozone Concentrations in Istanbul, Turkey. *CLEAN*, 897-908.
- Inc., S. (1998). *Data Mining with Confidence*. SPSS.
- INEGI. (2005). *Principales Resultados por Localidad, XIII Censo General de Población y Vivienda*. Mexicali, Baja California: Press Release.
- INEGI. (2011). *Perspectiva estadística Baja California*. Mexicali, Baja California: Press Release.
- Johnson, D. (2000). *Métodos multivariados aplicados al análisis de datos*. International Thomson.
- Kalantari, K., & Shabanali, H. (2007). Investigating factors Affecting Environmental Behavior of Urban Residents: A Case Study in Tehran City. *American Journal of Environmental Sciences*, 67-74.
- Kolehmainen, M., Martikainen, H., & Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment*, 815-825.
- Lambert, A. (2009). Cuantificación de Emisiones de Gases Refrigerantes en Baja California. // *Simposio Iberoamericano de Residuos Barranquilla*.
- Lavalle, S. (2006). *Planing Algorithms*. Cambridge: Cambridge University Press.
- Lobell, D., Ortiz, J. I., & Asner, G. (2005). Combining Field Surveys, Remote Sensing, and Regression Trees to Understand Yield Variations in an Irrigated Wheat Landscape. *ACSESS*, 241-249.
- Lozano, G. (2004). *Generación de desechos contaminantes en la familia y su diferenciación por estrato socioeconómico*. Mexicali, Baja California: Instituto de Ingeniería UABC.

- Lutz, R., & Woodhouse, R. (1997). Requirements analysis using forward and backward search. *Annals of Software Engineering*, 459-475.
- Morales, P. (2006). *Medición de actitudes en psicología y educación*. Madrid, España: Comillas.
- Moreno, M., Corraliza, J., & Ruiz, J. (2005). Escala de actitudes ambientales hacia problemas específicos. *Psicothema*, 502-508.
- Rosales, S., Bruno, C., & Balzarini, M. (2010). Identificación de relaciones entre rendimientos y variables ambientales vía árboles de clasificación y regresión (CART). *Interciencia*, 876-07.
- Salcedo, S., Camacho, J., & Perez, A. (2011). Improving the prediction of average total ozone in column over the Iberian Peninsula using neural networks banks. *Neurocomputing*, 1492-1496.
- Salinas, J., & Valcárcel, V. (2000). Aplicación del algoritmo CART como una Técnica de clasificación y predicción. *Anales Científicos UNALM*, 148-165.
- Van Lier, K., & Dunlap, R. (1981). Environment Concern. Does it make a difference how its measured? *Environment and Behavior*, 651-676.
- Villanueva, R. (2004). *Refrigerantes para aire acondicionado y refrigeracion*. España: Club Universitario.
- Wang, W., & Lu, W. (2003). Prediction of maximum daily ozone level using combined neural network and statistical characteristics. *Environment International*, 555-562.
- Weigel, R., & Weigel, J. (1978). Environmental concern. The development of measure. *Environment and Behavior*, 3-15.
- Yi, J., & Prybutok, V. (1996). A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Pollution*, 349-357.
- Zuk, M., Guadalupe, M., Cervantes, T., & Rojas, L. (2001). *Tercer almanaque de datos y tendencias de calidad del aire en nueve ciudades mexicanas*. INE-Semarnat.