



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL
Facultad de Ciencias Naturales y Matemáticas

**“Construcción de segmentación de clientes de un grupo
farmacéutico a través de modelos multivariados en base a
registros de ventas.”**

PROYECTO DE MATERIA INTEGRADORA

Previo a la obtención del Título de:
INGENIERA EN ESTADÍSTICA INFORMÁTICA

Presentado por:
Karen Daniela Erazo Guizado
Guayaquil – Ecuador
2015

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL FACULTAD DE
CIENCIAS NATURALES Y MATEMÁTICAS

INGENIERÍA EN ESTADÍSTICA E INFORMÁTICA

INFORME DEL PROFESOR DE LA MATERIA INTEGRADORA

Habiendo sido nombrado PROFESOR DE LA MATERIA
INTEGRADORA del señor,

**JUAN RAFAEL
CAMPOVERDE LÓPEZ**

Con el tema “DESARROLLO DE UN MODELO ESTADÍSTICO
PARA MEDIR LA POSIBILIDAD DE MOROSIDAD DE CLIENTES
EN UNA EMPRESA DE RETAIL”, previa a la obtención del título de
INGENIERO EN ESTADÍSTICA INFORMÁTICA, me permito
informar que he leído el contenido del proyecto integrador, luego de
lo cual indico que estoy de acuerdo en que el mismo se lo ha
desarrollado conforme a los lineamientos de la Unidad de Titulación
Especial de la ESPOL.

ING. SONNIA PAOLA REYES RAMOS
PROFESOR DE LA MATERIA INTEGRADORA

Guayaquil, 18 de Febrero del 2016



AGRADECIMIENTO

A mis padres y hermanos por haber sido mi motor de lucha, y a pesar de los inconvenientes que se presentan en el día a día siempre fueron constantes en su apoyo.

A cada uno de mis profesores por haber compartido sus conocimientos y hacer de mi vida universitaria una experiencia llena de buenos momentos y aprendizajes. En especial a Paola Reyes quien a más de haber sido maestra y tutora siempre estuvo ahí con sus consejos y apoyo incondicional.

A mi pareja y cómplice Ronald Vallejo quien siempre supo estar ahí dándome ánimos en cada objetivo propuesto.

A mis amigos y compañeros, por ese cariño y lucha que se vivió dentro y fuera de esta hermosa travesía llamada universidad.

DEDICATORIA

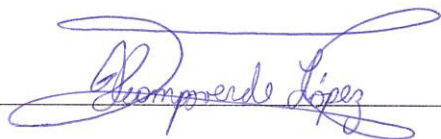
Este logro se lo dedico a mi mamá, ya que ella es quien con paciencia y amor me ha sabido guiar en cada uno de mis pasos, por ser un ejemplo admirable de mujer a seguir, por su confianza y apoyo infinito en cada una de mis decisiones.

A mi familia, por cada uno de sus consejos y ánimos de superación.

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de este Trabajo Final de Graduación, nos corresponde exclusivamente; y el patrimonio intelectual del mismo a la Escuela Superior Politécnica del Litoral".

(Reglamento de Graduación de la ESPOL)



Juan Rafael Campoverde López



ÍNDICE GENERAL

ÍNDICE GENERAL	I
ÍNDICE DE GRÁFICOS	II
ÍNDICE DE TABLAS	III
1 Descripción del problema y fuente de datos	5
1.1 Descripción del problema	5
1.2 Objetivos	8
1.2.1 Objetivo General.....	8
1.2.2 Objetivos Específicos	8
1.3 Origen de datos y Descripción de Variable a Utilizar.	9
1.4 Justificación.....	10
2 Metodología.....	12
2.1 Estadística Descriptiva	13
2.2 Estadística Inferencial.....	17
2.2.1 Prueba de Normalidad Kolmogórov – Smirnov	17
2.2.2 Contraste de Hipótesis	18
2.3 Metodología Especializada - Análisis Cluster.....	20
3 Análisis de Datos	23
3.1 Análisis Descriptivo	23
3.1.1 Cantidad de Clientes	23
3.1.2 Venta Neta	29
3.1.2.1 Análisis de la Venta Neta General.....	29
3.1.2.2 Análisis de la Venta Neta por cliente	34
3.2 Análisis Inferencial de los Datos	34
3.2.1 Prueba de Normalidad.....	35
3.2.2 Contraste Hipótesis – Diferencia de Medias	36
3.2.2.1 Comparación de Medias de Venta Neta por Segmento.....	36
4 Análisis Especializado: Análisis Clúster.	40
4.1 Análisis clúster.....	40
4.1.1 Aplicación del Análisis.	41
CONCLUSIONES	48
RECOMENDACIONES	49
REFERENCIAS BIBLIOGRÁFICA	50
5 Bibliografía	50

ÍNDICE DE GRÁFICOS

Gráfico 1: Tipo de Clientes vs Número de Clientes-B	24
Gráfico 2: Tipo de Clientes vs Número de Clientes – P	24
Gráfico 3: Tipo de Clientes vs Número de Clientes – C	25
Gráfico 4: Tipo de Clientes vs Número de Clientes – M	25
Gráfico 5: Tipo de Clientes vs Número de Clientes – V	26
Gráfico 6: Tipo de Clientes vs Número de Clientes – A	26
Gráfico 7: Tipo de Clientes vs Número de Clientes – D	27
Gráfico 8: Tipo de Clientes vs Número de Clientes – E	27
Gráfico 9: Tipo de Clientes vs Número de Clientes – J	28
Gráfico 10: Tipo de Clientes vs Número de Clientes – T	28
Gráfico 11: Tipo de Clientes vs Número de Clientes –	29
Gráfico 12: Ventas Netas Generales	30
Gráfico 13: Ventas Netas vs Clientes Totales	30
Gráfico 14: Porcentaje de las ventas netas mensuales del 2014.	31
Gráfico 15: Porcentaje de ventas y número de clientes por segmento.	32
Gráfico 16: Promedios de ventas por cliente y segmento	33
Gráfico 17: Prueba de Kolmogorov-Smirnov	36
Gráfico 18: Grupos de clientes con corte en Tabla 21: N° de clientes con ventas mayor	42

Gráfico 19: Grupos de clientes con corte en mayor sus ventas mayor igual a 1 millón.

Gráfico 20: Grupos de clientes con corte en mayor sus ventas mayor igual a 500 mil.

Gráfico 21: Grupos de clientes con corte en mayor sus ventas mayor igual a 100 mil.

Gráfico 22: Grupos de clientes con corte en mayor sus ventas mayor igual a 50 mil.

Gráfico 23: Grupos de clientes con corte en mayor sus ventas mayor igual a 10 mil.

Tabla 22: N° de clientes con ventas igual a 1 millón. 43

Tabla 23: N° de clientes con ventas igual a 500 mil. 45

Tabla 24: N° de clientes con ventas igual a 100 mil. 45

Tabla 25: N° de clientes con ventas igual a 50 mil. 46

Tabla 26: N° de clientes con ventas igual a 10 mil 47

ÍNDICE DE TABLAS

Tabla 1: Descripción de Clientes del Segmento B	24
Tabla 2: Descripción de Clientes del Segmento P	24
Tabla 3: Descripción de Clientes del Segmento C	25
Tabla 4: Descripción de Clientes del Segmento M	25
Tabla 5: Descripción de Clientes del Segmento V	26
Tabla 6: Descripción de Clientes del Segmento A	26
Tabla 7: Descripción de Clientes del Segmento D	27
Tabla 8: Descripción de Clientes del Segmento E	27
Tabla 9: Descripción de Clientes del Segmento J	28
Tabla 10: Descripción de Clientes del Segmento T	28

Tabla 11: Descripción de Clientes del Segmento I	29
Tabla 12: Comportamiento Mensual de las Ventas Netas en porcentajes.....	31
Tabla 13: Porcentaje de ventas y número de clientes por segmento.....	32
Tabla 14: Promedios de ventas por cliente y segmento.....	33
Tabla 15: Análisis descriptivos de las Ventas por cliente.	34
Tabla 16: Prueba de Kolmogorov-Smirnov para una muestra	35
Tabla 17: Estadísticos para prueba de Hipótesis J vs V.	37
Tabla 18: Prueba de muestras independientes J vs V	37
Tabla 19: Estadísticos para prueba de Hipótesis B vs V.....	38
Tabla 20: Prueba de muestras independientes B vs V.....	39
Gráfico 18: Grupos de clientes con corte en	Tabla 21: N° de clientes con ventas mayor 42
Gráfico 19: Grupos de clientes con corte en mayor sus ventas mayor igual a 1 millón.	Tabla 22: N° de clientes con ventas igual a 1 millón. 43
Gráfico 20: Grupos de clientes con corte en mayor sus ventas mayor igual a 500 mil.	Tabla 23: N° de clientes con ventas igual a 500 mil. 45
Gráfico 21: Grupos de clientes con corte en mayor sus ventas mayor igual a 100 mil.	Tabla 24: N° de clientes con ventas igual a 100 mil. 45
Gráfico 22: Grupos de clientes con corte en mayor sus ventas mayor igual a 50 mil.	Tabla 25: N° de clientes con ventas igual a 50 mil. 46
Gráfico 23: Grupos de clientes con corte en mayor sus ventas mayor igual a 10 mil.	Tabla 26: N° de clientes con ventas igual a 10 mil 47

CAPÍTULO I

1 Descripción del problema y fuente de datos

1.1 Descripción del problema

En vista de los cambios que han surgido por la implementación del “Reglamento para la fijación de precios de medicamentos de uso y consumo humano”, establecido en el decreto n°400 del Segundo Suplemento del Registro Oficial 299, realizado el 29 de Julio de 2014 en el gobierno actual.

Este reglamento posee un régimen regulado de precios en alrededor de 5000 medicamentos denominados “estratégicos”, los cuales son esenciales para la vida ya que sirven para tratar enfermedades raras y catastróficas, y se encuentran representando el 54% del consumo. Mientras que el 46% del consumo lo cubren los medicamentos denominados “no estratégicos”, los precios de estos se encuentran liberados, es decir el reglamento establece que los precios se ajustarán a lo que regule el mercado.

El gobierno señala que estos cambios fueron establecidos considerando que en Ecuador se consumen alrededor de 1.500 millones de dólares en medicamentos, y que la demanda no se genera en el consumidor sino que es inducida por los médicos o los farmacéuticos.

En base a estos acontecimientos el mercado empresarial ha tenido que adaptarse a cada una de estas regularizaciones, con el fin de buscar que estas no influyan de manera perjudicial para su negocio.

El mercado que se encuentran dentro de esta situación es el farmacéutico. Analizando los datos de una empresa perteneciente a este mercado, se han notado señales de alerta donde se cree conveniente evaluar algunos problemas que están afectando directamente en sus resultados.

Uno de los objetivos generales de toda empresa es la búsqueda e implementación en la optimización de recursos y correcto manejo de sus procesos, realizando mejoras en su productividad día a día.

Ante estos sucesos, la empresa se ha planteado crecer considerablemente en sus ventas, ya que el nivel de retorno será menor en ciertos casos, de este modo se busca poder conservar el nivel de ingresos.

Claramente se pretende mejorar la efectividad en sus procesos comerciales, por lo cual es considerable que uno de los procesos claves para que este objetivo se cumpla es realizar una correcta segmentación de sus clientes actuales.

La empresa maneja varios tipos de Modelos de negocio, el Retail y la Distribución a otras empresas.

Dentro del modelo de negocio de Distribución existe una división de los clientes por segmentos entre los cuales se han definido:

- 1) V (Cadenas Vip; Compromiso Comercial vip).

- 2) B (Clientes Locales; Farmacias, Boticas, Droguerías; Matriz; Ópticas B; Clientes Varios B).
- 3) E (Clientes Especiales; Clientes Objetivos; Farmacias Frontera).
- 4) J (Grupos Farmacéuticos).
- 5) D (Distribuidores).
- 6) M (Comunitarias).
- 7) P (Locales; Clientes PS).
- 8) I (Clientes Inactivos; Legal; Pre-Legal).
- 9) T (Distribuidores de Consumo).

Dependiendo el segmento al que se pertenece difiere el modelo de servicio, las condiciones de compra, etc.

Cuando se realiza segmentación de clientes se debe definir de manera concisa los criterios o parámetros que han de ser usados para la misma, los cuales permitan diseñar propuestas de valor y el tipo de relación que guarda cada segmento, y estos deben estar alineados con los objetivos planteados por la empresa.

En este caso no se conoce con exactitud cuáles fueron los parámetros que se tomaron en consideración para la segmentación actual, y la manera en que se iba segmentando simplemente fue consensuada en base a conveniencia de la empresa a medida que iba surgiendo cada tipo de cliente

Salta a la vista que la segmentación actual se guía más por tipo de cliente (mayorista - minorista), que por su comportamiento de venta que es lo que en realidad reflejaría una mejor segmentación en cuanto a sus clientes.

Es por esto que se pretende realizar una revisión de esta segmentación, para verificar si aplica luego de algunos años de implementación, enfocándonos en el comportamiento de compra de cada cliente.

1.2 Objetivos

1.2.1 Objetivo General

El objetivo del proyecto es aplicar técnicas estadísticas que en base a criterios representativos considerados, denoten mejores resultados y sean de gran ayuda para la productividad de la empresa, resolviendo algunos problemas que generan costos, los cuales disminuyen el nivel de retorno de la venta.

1.2.2 Objetivos Específicos

- Realizar un análisis de la Segmentación de los Clientes de Distribución actual.
- Analizar la información de compra que posee la empresa para captar a los clientes que sean significativos en el proceso de venta actual.
- En base al análisis de la variable propuesta, crear una nueva agrupación que se adapte más al entorno.

- Realizar una comparación entre la segmentación de clientes actual vs la segmentación generada, para así plantear propuestas de mejora en el proceso de venta de la empresa.

1.3 Origen de datos y Descripción de Variable a Utilizar.

Para el desarrollo del presente proyecto se necesita conocer y analizar los datos que se utilizarán y servirán como base para el alcance de los objetivos planteados.

Los datos utilizados para el análisis respectivo corresponden a un grupo distribuidor farmacéutico radicado en la ciudad de Guayaquil.

Estos datos son de una base correspondiente a la **venta por clientes**, la cual guarda en sus registros dicha empresa.

Se ha obtenido datos de las ventas de sus productos según la caracterización y segmentación establecida por la empresa, datos entre los cuales se conoce:

a) Tipo de cliente, b) Segmento al que corresponde, c) Identificación del cliente, d) Sucursal y e) Los valores de las ventas realizadas por cada sucursal.

Para el análisis posterior se utilizarán los datos de las ventas registradas en un periodo mensual desde Julio del año 2013 a Junio del presente año (dos años).

La variable a utilizar en este proyecto está denominada como:

- **Venta Neta**, en esta variable se reflejan las ventas totales en dólares de los productos que posee la empresa, dentro de los registros de ventas realizadas existen valores de devoluciones correspondientes a las mismas venta

1.4 Justificación

Se trata de resolver problemas que directamente influyen en el valor de retorno de la empresa, en este caso es un año con muchos cambios que afectan al sector Farmacéutico en el cual se debe hacer una análisis de todos los rubros en los cuales se podría estar perdiendo dinero.

Según datos históricos de la empresa, el canal de Distribución representa el 57% de las ventas, motivo por el cual se pretende realizar una segmentación de clientes basándonos en una nueva variable, la cuál será definida de acuerdo a información significativa que se presente en los datos de las ventas que posee actualmente la empresa.

El porcentaje restante se ubica en el canal de Retail, canal que no se tomará en cuenta para nuestro análisis ya que estos son rubros que representan los puntos de venta propios de la empresa.

La segmentación de clientes en curso se realizó aproximadamente hace 5 años atrás, la cual con todos los cambios que ha habido en el mercado se pretende validar la vigencia de la misma, caso contrario proponer cambios basados en la nueva segmentación.

Esto serviría para verificar si se le está dando una mejor condición de compra o algún tipo de descuento o beneficio a clientes que no son los adecuados o viceversa.

Y así realizar una comparación de que tan rentable resultaría mantener la segmentación actual vs una nueva propuesta, tomando en cuenta que uno de los

objetivos principales en el campo empresarial es mantener una constante mejora continua en sus procesos.

CAPÍTULO II

2 Metodología

Es importante conocer la caracterización establecida por la empresa para cada segmento y tipo de cliente y así poder realizar un correcto análisis de las variables tomadas en consideración para la segmentación establecida actualmente.

Esto servirá como base fundamental para la posterior toma de decisiones y así poder aplicar las diferentes técnicas estadísticas que se consideren necesarias para el respectivo análisis, ya que estas técnicas nos permitirán analizar las variables inmersas en los datos, establecer inferencias y relaciones sobre las mismas, con esto se puede observar de una manera más clara la posición actual de la empresa frente a la segmentación que posee.

Recordando que el objetivo es establecer una nueva segmentación en la que los valores establecidos en las ventas sean los que clasifiquen a los distintos clientes en su correcto segmento, para que con esto la empresa pueda tener un panorama claro de cuáles son sus clientes importantes y poder establecer beneficios y lineamientos para cada tipo de cliente, sin temor a que una mala clasificación le represente pérdidas monetarias.

2.1 Estadística Descriptiva

Para este proyecto se ha contemplado realizar un **análisis estadístico descriptivo**, en el cuál mediante gráficos de barras que es la manera más común de representar gráficamente la distribución de frecuencia de los datos¹, se pretende observar el estado actual de cada segmento.

Se hace un análisis previo de cómo se comportan los datos de las ventas registradas tanto por mes como por segmento.

Se realiza un **análisis de media** en relación a las ventas mensuales por tipo de segmento, para así tener una visualización clara de cuanto es promedio mensual aproximado en relación a las ventas registradas en cada periodo de tiempo.

También se quiere mostrar la cantidad de clientes que pertenecen a cada grupo o segmento definido por la empresa. Hay que tomar en consideración que los datos a analizar pertenecen a un grupo farmacéutico, por lo tanto cada segmento contiene datos de varias empresas dedicadas a este sector empresarial.

Los segmentos previamente establecidos por la empresa que se van a analizar son los siguientes:

^[1]Luis Rodríguez Ojeda, PROBABILIDAD Y ESTADÍSTICA BÁSICA PARA INGENIEROS Con el soporte de MATLAB para cálculos y gráficos estadísticos, (Escuela Superior Politécnica del Litoral, ICM, Guayaquil-Ecuador 2007).

1)

SEGMENTO	TIPO DE CLIENTE
B	CLIENTES LOCALES
	CLIENTES VARIOS
	FARMACIAS, BOTICAS,DROGUERÍAS
	MATRIZ
	OPTICAS B

2)

SEGMENTO	TIPO DE CLIENTE
P	CLIENTES LOCALES
	OPTICA PS

SEGMENTO B

***Clientes Locales:** Son las diferentes farmacias pertenecientes a personas naturales en las ciudades de Guayaquil, Quito y Cuenca.

***Clientes Varios:** Son los llamados clientes consumo, entre ellos se encuentran almacenes, bazares, farmatiendas, minimarkets, pañaleras, tiendas, entre otros.

***Farmacias, Boticas, Droguerías:** Registradas bajo el nombre de CIA LTDA.

***Matriz:** Farmacias correspondientes a la Televenta.

***Ópticas B:** Ópticas correspondientes a este segmento.

3)

SEGMENTO	TIPO DE CLIENTE
C	FRANQUICIAS

4)

SEGMENTO	TIPO DE CLIENTE
M	COMUNITARIAS

SEGMENTO C

***Franquicias:** Diferentes franquicias registradas en el grupo farmacéutico.

SEGMENTO M

***Comunitarias:** Se refieren a las diferentes farmacias comunitarias.

5)

SEGMENTO	TIPO DE CLIENTE
V	CADENAS VIP
	COMPROMISO COMERCIAL VIP

SEGMENTO V

Cadenas Vip: Cadenas de farmacias que aportan un significativo volumen de compra al grupo farmacéutico.

Compromiso Comercial Vip: Cadenas de boticas y farmacias con los que se mantiene un compromiso comercial.

SEGMENTO A

Instituciones Privadas: Clínicas, Dispensarios médicos, Empresas, Fundaciones, Laboratorios clínicos, médicos, otros.

Instituciones Públicas: Distribuidores públicos, instituciones públicas, municipios.

Prestadores Ext. Nivel 1, 2,3: Prestadores externos que brindan servicios privados.

6)

SEGMENTO	TIPO DE CLIENTE
A	INSTITUCIONES PRIVADAS
	INSTITUCIONES PÚBLICAS G.
	PRESTADORES EXT. NIVEL 1
	PRESTADORES EXT. NIVEL 2
	PRESTADORES EXT. NIVEL 3

7)

SEGMENTO	TIPO DE CLIENTE
D	DISTRIBUIDORES
	MAYORISTAS FARMACÉUTICOS

8)

SEGMENTO	TIPO DE CLIENTE
E	CLIENTES ESPECIALES
	CLIENTES OBJETIVOS
	FARMACIAS FRONTERA

SEGMENTO D

***Distribuidores:** Distintas distribuidoras farmacéuticas y Multiventas.

***Mayoristas Farmacéuticos:** Mayoristas farmacéuticas varios.

SEGMENTO E

***Farmacias Frontera:** Farmacias que lindan con la frontera de Ecuador.

9)

SEGMENTO	TIPO DE CLIENTE
J	GRUPO FARCOMED
	GRUPO FARMAENLACE
	GRUPO FARMASOL
	GRUPO ORELLANA
	GRUPO SUMELAB
	QUIFATEX

10)

SEGMENTO	TIPO DE CLIENTE
T	AUTOSERVICIOS
	DISTRIBUIDORES
	MAYORISTAS
	PAÑALERAS T

SEGMENTO J

Grupos FARCOMED, FARMAENLACE, FARMASOL, ORELLANA, SUMELAB: Farmacias de la competencia que compran ciertos productos al grupo analizado.

QUIFATEX: Distribuidor que pertenece a la competencia.

SEGMENTO T

Autoservicios: gasolineras, oki doki, minimarkets, productos etc.

Distribuidores: Empresas distribuidoras de consumo.

Mayoristas: Mayoristas de consumo.

Pañaleras T: Pañaleras correspondientes a este segmento.

11)

SEGMENTO	TIPO DE CLIENTE
I	CLIENTES INACTIVOS
	LEGAL
	PRELEGAL

SEGMENTO I

Clientes Inactivos: Son todos los clientes inactivos de compra correspondientes a los diferentes segmentos.

Legal: Clientes que se encuentran atravesando un juicio por falta de pago.

Prelegal: Clientes con algún compromiso de pago establecido para evitar juicio.

2.2 Estadística Inferencial

La realización de un análisis estadístico inferencial es contemplado en este proyecto ya que en esta parte de la estadística es en donde diferentes métodos y procedimientos influyen sobre una determinada muestra de la población y través de deducciones o afirmaciones a probar, nos dan como resultado diferentes propiedades de la población objetivo analizada.

2.2.1 Prueba de Normalidad Kolmogórov – Smirnov

La prueba de bondad de ajuste de Kolmogórov-Smirnov es una prueba no paramétrica, también se la conoce como prueba K-S, ésta es utilizada para medir el grado de relación o bondad de ajuste entre una muestra de datos y una distribución de probabilidad previamente establecida.

Es decir, se la utiliza para saber si los datos pertenecientes a una determinada población provienen de alguna distribución que se tenga como referencia. Esto se lleva a cabo comparando la distribución acumulada de la distribución previamente establecida con la distribución empírica de la muestra de datos escogida y se realiza un cálculo de las diferencias que pudieran existir.

Las hipótesis a contrastar en esta prueba son las siguientes:

H_0 : Los datos de la muestra provienen de una población que sigue una distribución M.

H_1 : Los datos de la muestra no provienen de una población que sigue una distribución M.

2.2.2 Contraste de Hipótesis

Una manera excepcional de realizar inferencias es mediante **prueba de hipótesis**, la cual “consiste en suponer algún valor para el parámetro de interés y usar los datos de la muestra para aceptar o rechazar esta afirmación”². Es decir que a través de alguna afirmación propuesta como una hipótesis, nos permite realizar contrastes ante alguna evidencia que se tiene de la información registrada.

^[2] Luis Rodríguez Ojeda, PROBABILIDAD Y ESTADÍSTICA BÁSICA PARA INGENIEROS Con el soporte de MATLAB para cálculos y gráficos estadísticos, (Escuela Superior Politécnica del Litoral, ICM, Guayaquil-Ecuador 2007).

En este caso se realizará una prueba de hipótesis de contraste de medias, definida como:³

$$\begin{array}{c} H_0 = \mu_1 \neq \mu_2 \neq \mu_3 \\ \text{Vs} \\ H_1 = \mu_1 = \mu_2 = \mu_3 \end{array}$$

Los pasos que generalmente se contemplan para la realización de una prueba de hipótesis son:

- a. Formulación de la hipótesis de investigación y obtención de los datos
- b. Formulación de la hipótesis alternativa y la hipótesis nula
- c. Selección de la prueba de hipótesis más adecuada
- d. Determinación del nivel de significancia
- e. Determinación del tamaño de la muestra
- f. Determinación de la distribución muestral de la prueba estadística para H_0

Para nuestro caso la realización de una prueba de hipótesis en relación a las medias resulta novedoso, ya que se puede establecer un análisis en cuanto a la igualdad de las medias de las ventas por tipo de segmento, como se parte de la premisa que la segmentación actual se encuentra mal hecha en base a la falta de consideración de los parámetros para la misma, se busca probar que no existe una diferencia significativa entre las medias de cada segmento.

^[3] Cazau, P. *Fundamentos de estadística*.

2.3 Metodología Especializada - Análisis Cluster.

Con la finalidad de establecer una nueva segmentación se propone hacer uso de una técnica estadística multivariada denominada Análisis de Conglomerados (Cluster).

El análisis de conglomerados (cluster) Es una técnica multivariante que busca agrupar elementos (o variables) tratando de lograr la máxima homogeneidad en cada grupo y la mayor diferencias entre los grupos.⁴

En el análisis de conglomerados se busca patrones en un conjunto o grupos de datos mediante la agrupación de las observaciones.

El objetivo es encontrar un agrupamiento óptimo para que las observaciones u objetos dentro de cada grupo sean similares, pero los grupos sean diferentes entre sí.

Se espera encontrar las agrupaciones naturales en los datos, agrupaciones que tengan sentido para el análisis establecido.

Para agrupar las observaciones en conglomerados, muchas técnicas comienzan con similitudes entre todos los pares de observaciones. En muchos casos, las similitudes se basan en alguna medida de distancia.

Otros métodos de conglomerados utilizan una elección preliminar para centros de conglomerados o una comparación de la variabilidad intra e inter- cluster. También es posible agrupar las variables, en cuyo caso la similitud podría haber una correlación.

^[4] Manuel Terrádez Gurrea, <http://www.uoc.edu/in3/emath/docs/Cluster.pdf>

El análisis de conglomerados también se ha referido como la clasificación, reconocimiento de patrones (específicamente, aprendizaje no supervisado), y la taxonomía numérica.

Las técnicas de análisis de conglomerados se han aplicado ampliamente a los datos en muchos campos, como la medicina, la psiquiatría, la sociología, la criminología, la antropología, la arqueología, la geología, la geografía, la teledetección, estudios de mercado, economía e ingeniería.⁵

Centrándose en gran medida de las variables cuantitativas la matriz de datos se puede escribir como:

$$Y = \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{pmatrix} = (y_{(1)}, y_{(2)}, \dots, y_{(p)}),$$

Donde $Y(i)$ es una fila (vector observación) y $Y(j)$ es una columna (que corresponde a una variable) .

Por lo general, deseamos agrupar las n $Y'(i)$ ' s (filas) dentro de g conglomerados .

También es posible que desee agrupar las columnas $Y(j)$, $j = 1, 2, \dots, p$.

Dos enfoques comunes a la agrupación de los vectores de observación son la agrupación jerárquica y la partición

^[5] Methods of Multivariate Analysis, Second Edition, Alvin C. Rencher, Brigham Young University, WILEY-INTERSCIENCE A JOHN WILEY & SONS, INC. PUBLICATION 2002.

En la agrupación jerárquica se suele comenzar con n grupos, uno para cada observación, y se termina con un solo grupo que contenga las n observaciones. En cada paso, una observación o un grupo (cluster) de observaciones es absorbido dentro de otro grupo (cluster).

También podemos revertir este proceso, es decir, comenzar con un solo grupo (cluster) que contenga todas las n observaciones y terminar con n grupos de un solo elemento para cada observación.

En la partición, simplemente dividimos las observaciones dentro de g grupos (clusters). Esto se puede hacer comenzando con una partición inicial o con centros de los conglomerados y luego la reasignación de las observaciones de acuerdo con algún criterio de optimalidad.⁶

El objetivo de utilizar esta técnica en nuestro proyecto es para ver el comportamiento de las ventas a través de tiempo y en base a este parámetro segmentar o separar por grupos a los diferentes clientes que se encuentran formando parte de la base de datos de la empresa, así se podrá observar cuáles son los clientes que se agrupan de acuerdo a su registro de ventas, y no necesariamente pueden pertenecer al segmento previamente establecido por la empresa.

^[6] Methods of Multivariate Analysis, Second Edition, Alvin C. Rencher, Brigham Young University, WILEY-INTERSCIENCE A JOHN WILEY & SONS, INC. PUBLICATION 2002.

CAPÍTULO III

3 Análisis de Datos

3.1 Análisis Descriptivo

Se analiza el comportamiento de la variable ventas netas, a través de esto se podrá comprender de una manera clara el comportamiento de las ventas en el tiempo establecido para el sector empresarial estudiado.

3.1.1 Cantidad de Clientes

La base de datos se encuentra estructurada de tal manera que se cuenta con una segmentación o agrupación general previamente establecida, dentro de cada segmento se tiene diferentes agrupaciones de acuerdo al tipo de cliente y por último se posee la especificación del cliente.

A través de gráficos de barras se puede observar la presencia en cuanto a la cantidad de clientes por segmentos dentro de toda la base de datos mencionada.

Se presenta a detalle una tabla para cada segmento, con el nombre del segmento, tipo de cliente y frecuencia al que pertenecen los distintos clientes. Con su respectivo gráfico de barras.

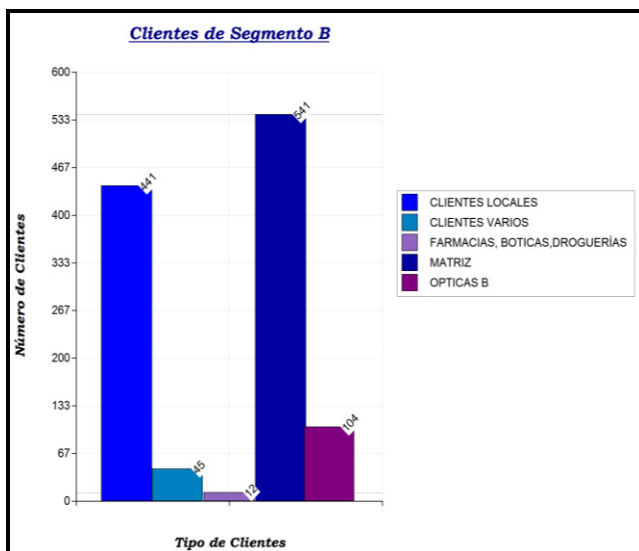


Gráfico 1: Tipo de Clientes vs Número de Clientes-B

Tabla 1: Descripción de Clientes del Segmento B

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
B	CLIENTES LOCALES	441
	CLIENTES VARIOS	45
	FARMACIAS, BOTICAS, DROGUERÍAS	12
	MATRIZ	541
	OPTICAS B	104

En este segmento podemos notar que el subsegmento “Matriz” posee el mayor número de clientes, a diferencia del resto.

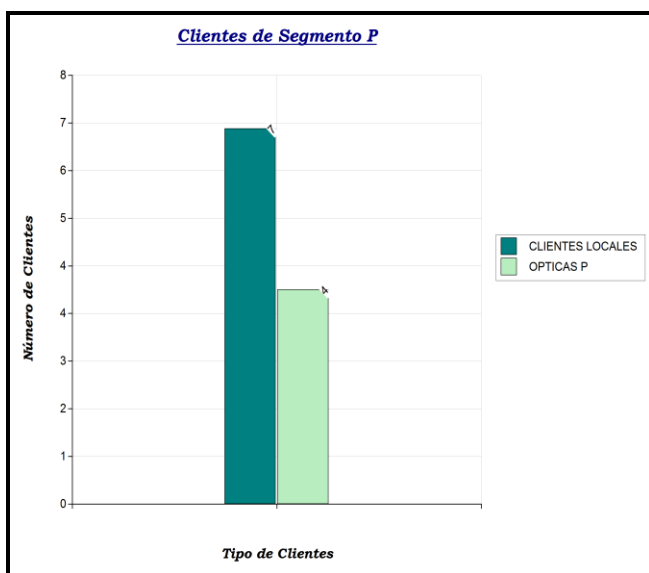


Gráfico 2: Tipo de Clientes vs Número de Clientes – P

Tabla 2: Descripción de Clientes del Segmento P

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
P	CLIENTES LOCALES	7
	OPTICA PS	4

En el segmento P el número de clientes varía por pocas unidades

Tabla 3: Descripción de Clientes del Segmento C

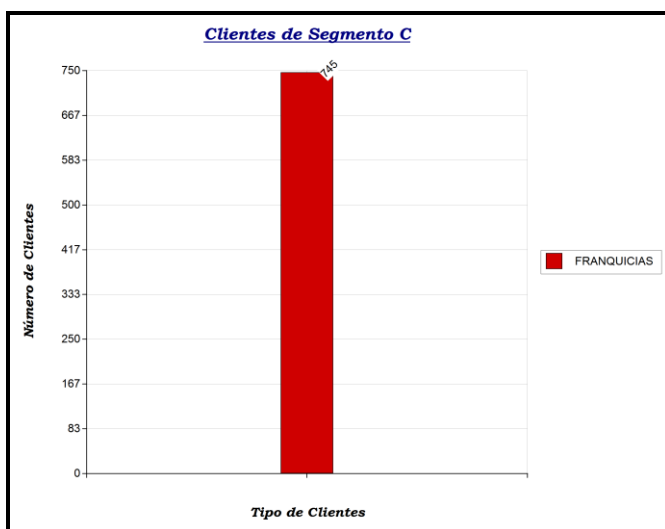


Gráfico 3: Tipo de Clientes vs Número de Clientes – C

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
C	FRANQUICIAS	745

En este segmento se tiene un solo tipo de cliente, en donde se ubican todas las franquicias.

Tabla 4: Descripción de Clientes del Segmento M

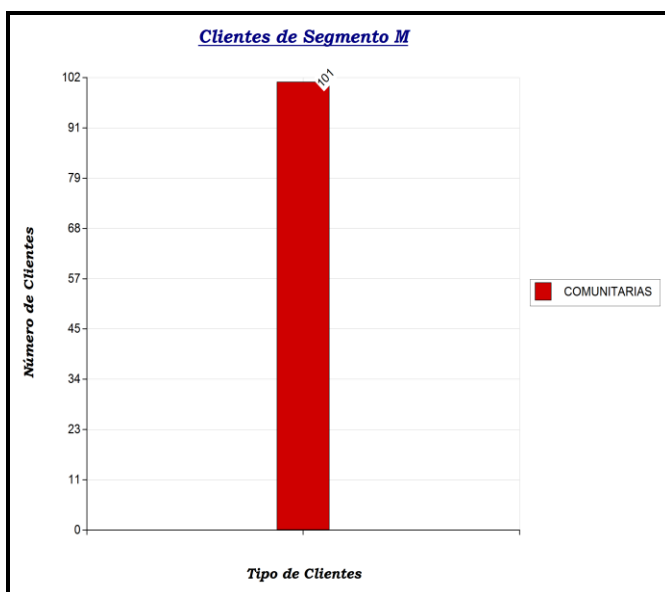


Gráfico 4: Tipo de Clientes vs Número de Clientes – M

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
M	COMUNITARIAS	101

En este segmento se tiene un solo tipo de cliente, en donde se ubican todos los establecimientos denominados comunitarias.

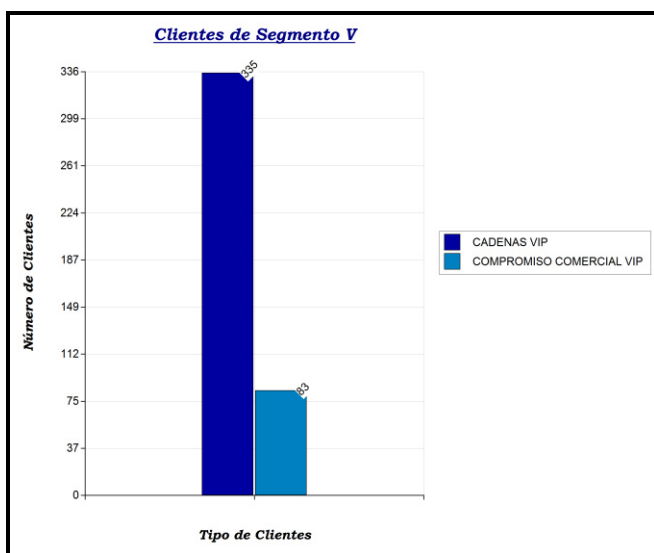


Gráfico 5: Tipo de Clientes vs Número de Clientes – V

Tabla 5: Descripción de Clientes del Segmento V

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
V	CADENAS VIP	335
	COMPROMISO COMERCIAL VIP	83

Este segmento es considerado como clientes potenciales para la empresa, como se puede observar los clientes denominados como “Cadenas Vip” son los que más clientes poseen.

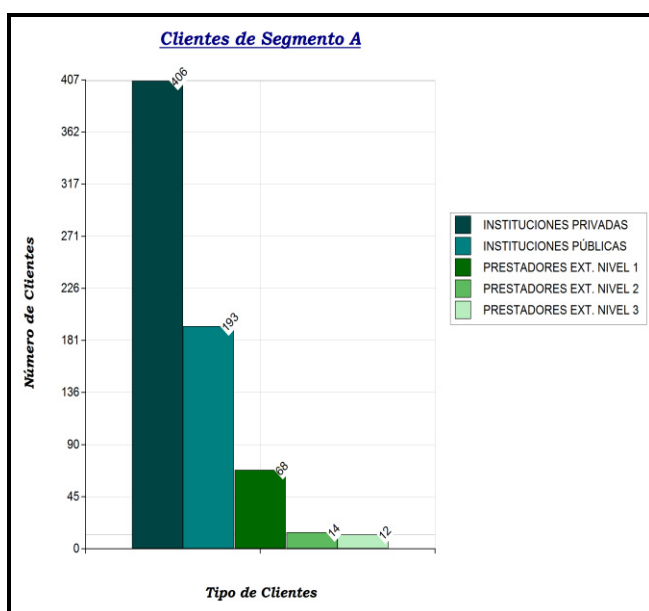


Gráfico 6: Tipo de Clientes vs Número de Clientes – A

Tabla 6: Descripción de Clientes del Segmento A

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
A	INSTITUCIONES PRIVADAS	406
	INSTITUCIONES PÚBLICAS G.	193
	PRESTADORES EXT. NIVEL 1	68
	PRESTADORES EXT. NIVEL 2	14
	PRESTADORES EXT. NIVEL 3	12

En este segmento se encuentran los clientes cuyos productos son destinados para instituciones de salud, ya sean públicas o privadas, el mayor número de clientes lo poseen las instituciones privadas.

Tabla 7: Descripción de Clientes del Segmento D

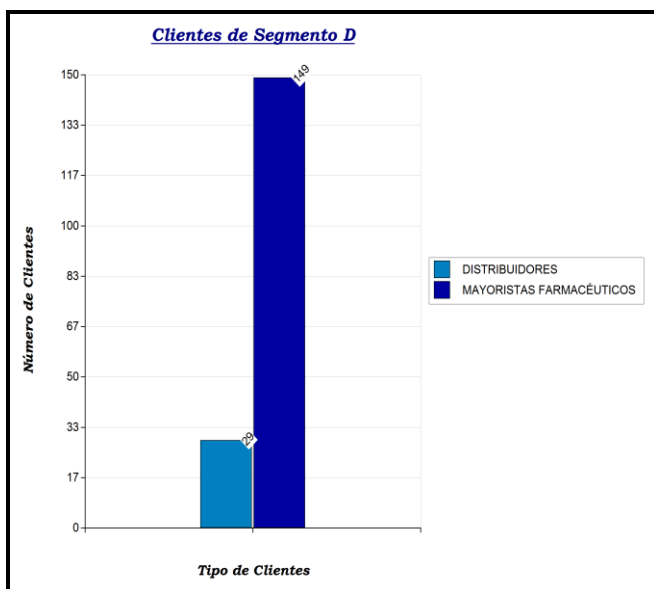


Gráfico 7: Tipo de Clientes vs Número de Clientes – D

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
D	DISTRIBUIDORES	29
	MAYORISTAS FARMACÉUTICOS	149

El segmento D poseen clientes que realizan ventas al por mayor, los clientes que más consumo registran en las ventas de la empresa son los mayoristas farmacéuticos.

Tabla 8: Descripción de Clientes del Segmento E

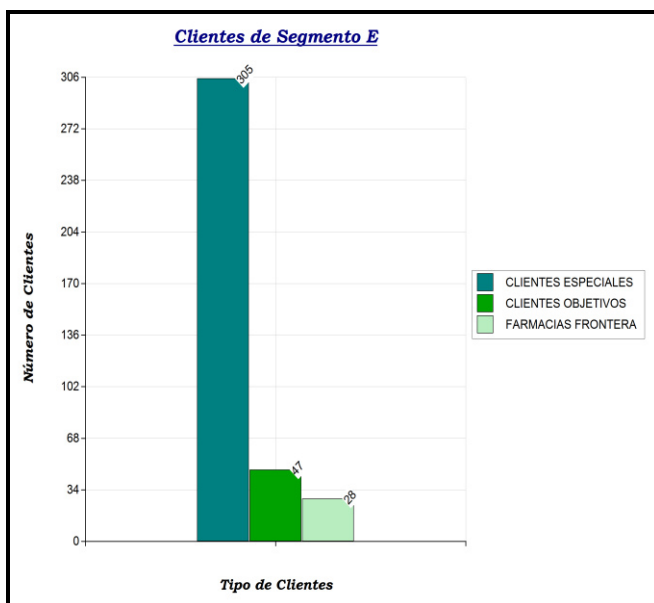


Gráfico 8: Tipo de Clientes vs Número de Clientes – E

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
E	CLIENTES ESPECIALES	305
	CLIENTES OBJETIVOS	47
	FARMACIAS FRONTERA	28

Clientes que lindan con la frontera del Ecuador, la mayoría de clientes dentro de este segmento se encuentran en los clientes de tipo “Especiales”

Tabla 9: Descripción de Clientes del Segmento J

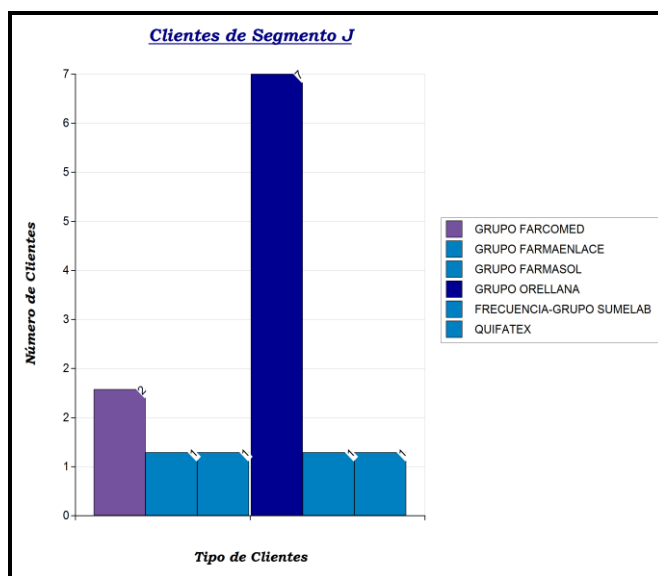


Gráfico 9: Tipo de Clientes vs Número de Clientes – J

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
J	GRUPO FARCOMED	2
	GRUPO FARMAENLACE	1
	GRUPO FARMASOL	1
	GRUPO ORELLANA	7
	GRUPO SUMELAB	1
	QUIFATEX	1

En este segmento se encuentran los clientes pertenecientes a la competencia, como se puede observar la mayoría de los clientes pertenecen al grupo Orellana.

Tabla 10: Descripción de Clientes del Segmento T

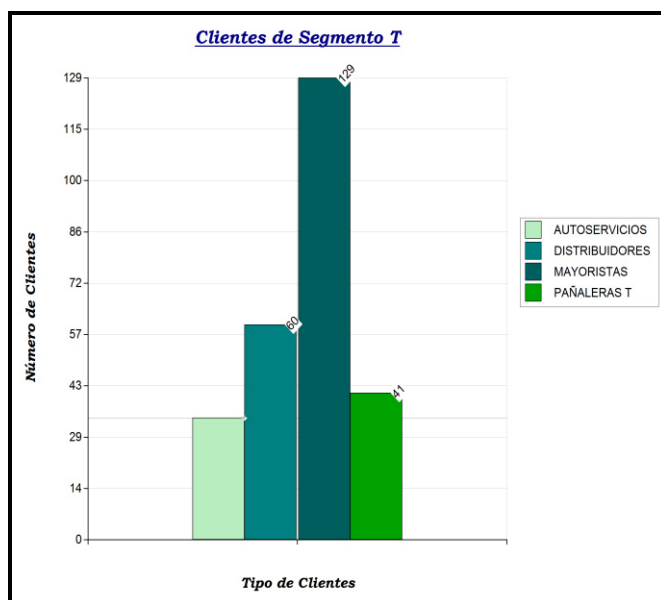


Gráfico 10: Tipo de Clientes vs Número de Clientes – T

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
T	AUTOSERVICIOS	34
	DISTRIBUIDORES	60
	MAYORISTAS	129
	PAÑALERAS T	41

El mayor número de clientes se encuentran en el tipo de clientes denominado “Mayoristas”.

Tabla 11: Descripción de Clientes del Segmento I

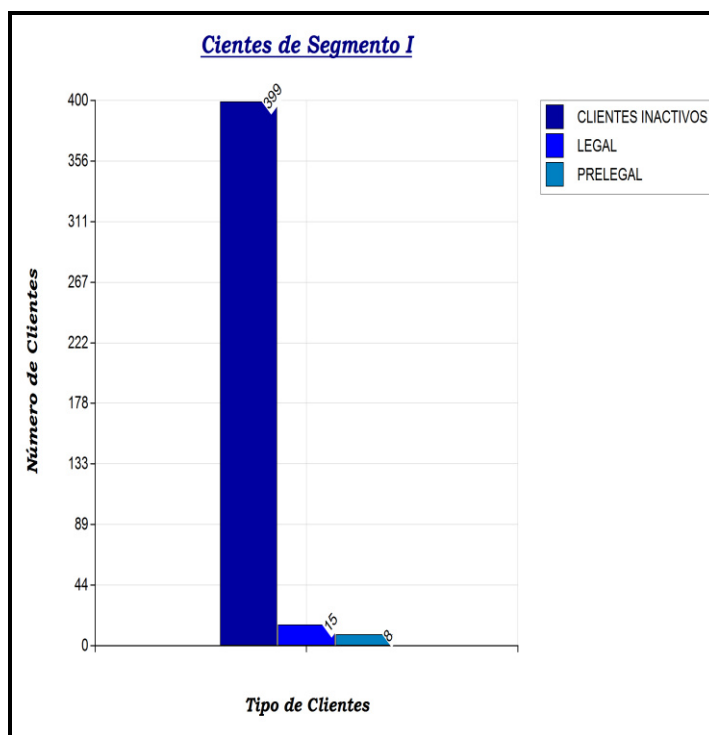


Gráfico 11: Tipo de Clientes vs Número de Clientes –

SEGMENTO	TIPO DE CLIENTE	FRECUENCIA
I	CLIENTES INACTIVOS	399
	LEGAL	15
	PRELEGAL	8

Los clientes de este segmento son denominados inactivos ya que no mantienen algún comportamiento de compra, y deudores ya que algunos de ellos mantienen deudas o juicios legales con la empresa, el mayor número de clientes en este segmento lo reflejan los clientes inactivos.

3.1.2 Venta Neta

Una vez descritos la cantidad de clientes con sus respectivos tipos de clientes, se procede a realizar un análisis general de las ventas netas para visualizar el estado de los datos actualmente. La variable analizada es Venta Neta en dólares.

3.1.2.1 Análisis de la Venta Neta General

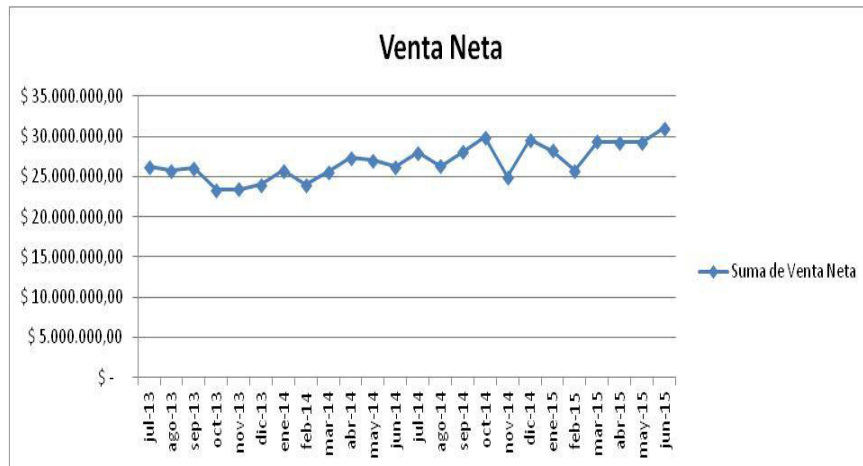


Gráfico 12: Ventas Netas Generales

En el gráfico 12 podemos observar el comportamiento de las ventas totales por mes de toda la variable Venta Neta. Las ventas generales para esta empresa se mantienen por encima de los \$25.000.000,00

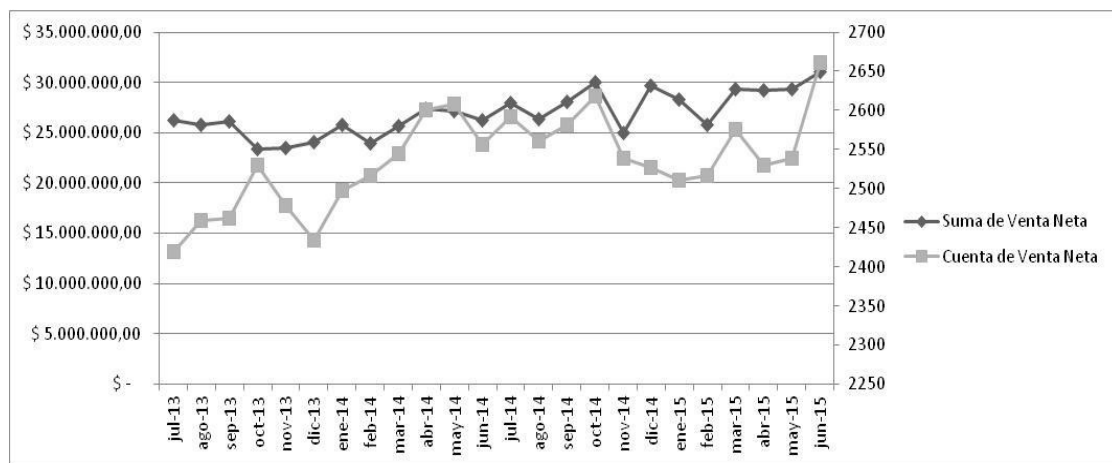


Gráfico 13: Ventas Netas vs Clientes Totales

En el gráfico 13 se observa el comportamiento de las ventas totales en dólares vs la cantidad de clientes a los que se le han realizado las ventas mensualmente. Se puede

notar como en el mes de Julio del 2013 a pesar de poseer pocos clientes en relación al resto de meses se mantuvo una venta sobre los \$25.000.000.00

Tabla 12: Comportamiento Mensual de las Ventas Netas en porcentajes.

Año/Mes	1	2	3	4	5	6	7	8	9	10	11	12
2013	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	17,65 %	17,27 %	17,50 %	15,68 %	15,74 %	16,15 %
2014	7,97%	7,43%	7,93%	8,49%	8,39%	8,12%	8,65%	8,15%	8,70%	9,28%	7,73%	9,18%
2015	16,33 %	14,90 %	16,97 %	16,89 %	16,96 %	17,95 %	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Total general	8,37%	7,72%	8,52%	8,78%	8,75%	8,88%	8,41%	8,07%	8,40%	8,27%	7,51%	8,33%

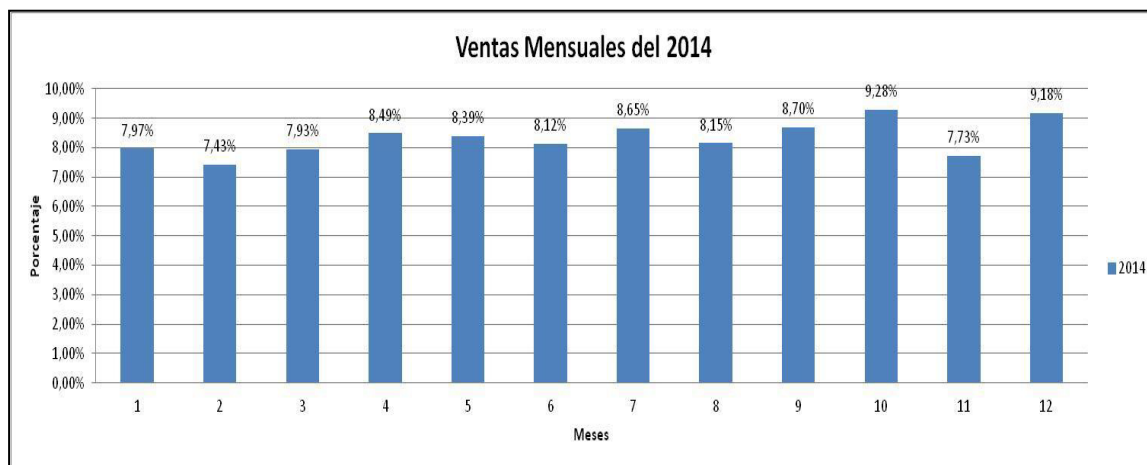


Gráfico 14: Porcentaje de las ventas netas mensuales del 2014.

En el gráfico 14 se presenta el flujo de las ventas totales durante el año 2014, se puede notar que en el mes de Octubre el porcentaje de las ventas alcanza un 9,8% siendo este el mes más productivo durante aquel año.

A continuación se muestra la participación de las ventas por segmento y número de clientes por segmento

Tabla 13: Porcentaje de ventas y número de clientes por segmento.

Segmentos	Suma de Venta Neta	Cuenta de Venta Neta (núm. de clientes)
A	5,21%	10,94%
B	7,81%	30,85%
C	54,17%	26,80%
D	2,48%	4,33%
E	6,09%	10,54%
J	7,41%	0,38%
M	2,38%	3,35%
P	0,00%	0,02%
T	1,72%	3,56%
V	12,74%	9,23%
Total general	100,00%	100,00%

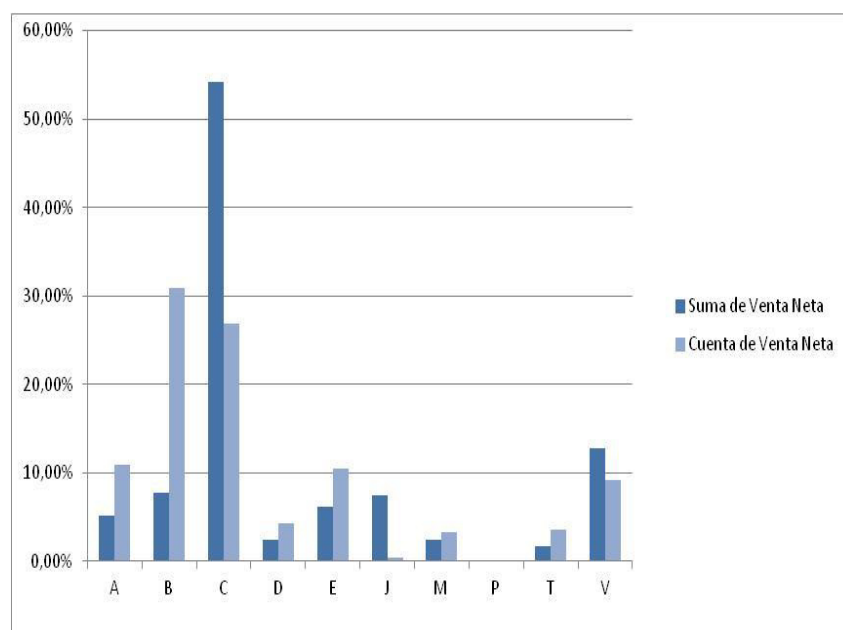


Gráfico 15: Porcentaje de ventas y número de clientes por segmento.

Como se puede observar El segmento que mayor ventas representa es el segmento C con un 54,17%, el cual corresponde a Franquicias, sin embargo el Segmento B es el que posee un mayor porcentaje de clientes 30,85%, en relación a los demás segmentos.

También se consideró analizar el comportamiento de las ventas en promedio por cliente de cada segmento, donde se obtuvo los siguientes resultados:

Tabla 14: Promedios de ventas por cliente y segmento.

Segmentos	Promedio de Venta Neta
A	5044,507094
B	2684,58875
C	21426,71679
D	6075,130585
E	6123,777906
J	206871,6453
M	7554,376616
P	317,326
T	5102,965588
V	14629,87501
Total general	10600,94766



Gráfico 16: Promedios de ventas por cliente y segmento

El gráfico 16 muestra como las ventas promedio por cliente significativas se encuentran en el segmento J con un valor de \$206871,64, el segundo segmento significativo es el C con \$21426,71679 y el tercer segmento que sobresale es el V con \$14629,87501.

3.1.2.2 Análisis de la Venta Neta por cliente

Tabla 15: Análisis descriptivos de las Ventas por cliente.

Venta Neta		Estadístico	Error típ.
Media		10600,947661	97,1741068
I. de conf. para la media al 95%	Lím. inf.	10410,486184	
	Lím. Sup.	10791,409139	
Media recortada al 5%		7654,714086	
Mediana		3987,760000	
Varianza		5,746E8	
Desv. típ.		23971,6921076	
Mínimo		-125876,7000	
Máximo		726806,0500	
Rango		852682,7500	
Amplitud intercuartil		11936,6500	
Asimetría		11,557	,010
Curtosis		204,296	,020

En base a los resultados se observa que existen valores negativos, esto se debe a las devoluciones en dólares realizadas por los clientes. La venta Máxima por cliente realizada es de \$726.806,05. En general el 50% de las ventas por cliente son valores menores iguales a \$10.600,94. También se puede observar que reduciendo los valores atípicos, las ventas en promedio son de \$ 7.654,71.

3.2 Análisis Inferencial de los Datos

3.2.1 Prueba de Normalidad

Se desea observar si los datos utilizados de las Ventas netas en este proyecto siguen una distribución normal, por lo tanto se realizó una prueba de normalidad para ver el comportamiento de los datos.

Tabla 16: Prueba de Kolmogorov-Smirnov para una muestra

		Venta Neta
N		60855
Parámetros normales ^{a,b}	Media	10600,947661
	Desviación típica	23971,6921076
Diferencias más extremas	Absoluta	,319
	Positiva	,211
	Negativa	-,319
Z de Kolmogorov-Smirnov		78,630
Sig. asintót. (bilateral)		,000
a. La distribución de contraste es la Normal.		
b. Se han calculado a partir de los datos.		

La prueba de Kolmogorov- Smirnov nos dice que nuestros datos no siguen una distribución normal, ya que según el nivel de significancia de la prueba debe ser mayor 0,05 para definir normalidad en los datos, y como podemos observar para nuestro caso el valor es menor a 0,05.

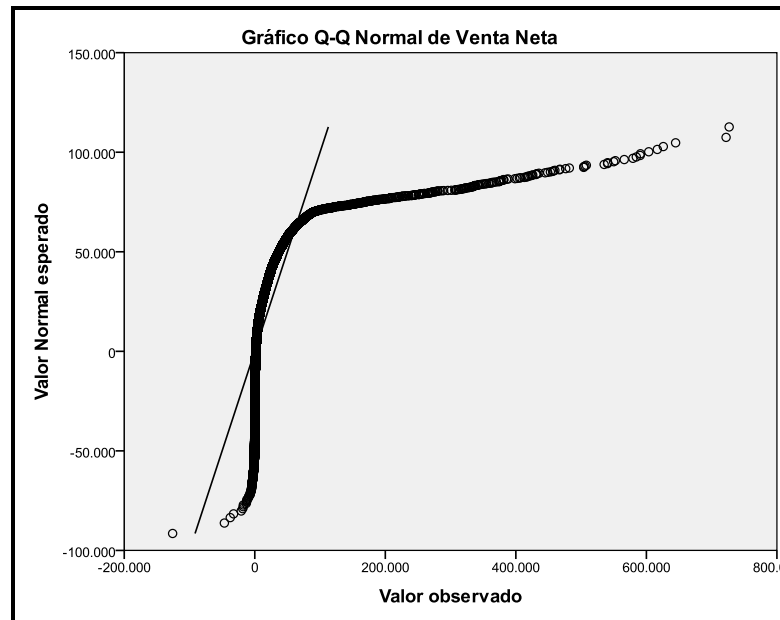


Gráfico 17: Prueba de Kolmogorov-Smirnov

Como podemos observar en el gráfico el comportamiento de los datos no se asemejan a una distribución normal.

3.2.2 Contraste Hipótesis – Diferencia de Medias

3.2.2.1 Comparación de Medias de Venta Neta por Segmento

a) Se define la hipótesis con respecto a las Ventas de los Clientes más representativos para la empresa como :

$$H_0 = \mu_{\text{SegmentoJ}} \neq \mu_{\text{SegmentoV}}$$

Vs

$$H_1 = \mu_{\text{SegmentoJ}} = \mu_{\text{SegmentoV}}$$

Se pretende demostrar que aunque el segmento V este establecido como los clientes de más importancia para la empresa, la medias de sus ventas en relación al Segmento J que básicamente son empresas de la competencia presentan una diferencia significativa.

Tabla 17: Estadísticos para prueba de Hipótesis J vs V.

Estadísticos de grupo				
SEGMENTO GENERAL	N	Media	Desviación típ.	Error típ. de la media
Venta Neta J	231	206871,645281	152644,1417334	10043,2471779
Venta Neta V	5616	14629,875011	42087,6785489	561,6185220

Tabla 18: Prueba de muestras independientes J vs V

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	Gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
									Inferior Superior
Venta Neta	Varianzas iguales	1608,286	,000	55,959	5845	,000	192241,7702707	3435,3883255	185507,1382978 198976,4022436
	Varianzas distintas			19,112	231,441	,000	192241,7702707	10058,9377790	172422,9780504 212060,5624910

Según los resultados obtenidos podemos notar las medias de los segmentos escogidos no son iguales, ya que debido al nivel de significancia de la prueba notamos que existe evidencia estadística significativa para aceptar H_0 .

Es decir, tal como se estableció la hipótesis, las medias de los dos grupos resultaron diferentes. La media de las ventas del grupo V reflejan un menor valor en comparación a las del grupo J, a pesar de que el Segmento V posee más clientes en comparación al segmento J.

b) Se plantea otra hipótesis donde se toma en consideración el segmento B, debido a que este es uno de los más representativos en cuanto a número de clientes.

Por lo tanto se establece que:

$$H_1 = \mu_{\text{Segmento B}} = \mu_{\text{Segmento V}}$$

Vs

$$H_0 = \mu_{\text{Segmento B}} \neq \mu_{\text{Segmento V}}$$

Se quiere probar que la media de las ventas del segmento B son iguales en comparación a las medias de las ventas de segmento V.

Tabla 19: Estadísticos para prueba de Hipótesis B vs V.

Estadísticos de grupo				
SEGMENTO GENERAL	N	Media	Desviación típ.	Error típ. de la media
Venta Neta B	18771	2684,588750	4348,3883255	31,7383698
V	5616	14629,875011	42087,6785489	561,6185220

Una vez realizada la prueba, tenemos como resultados que la media de las ventas del segmento B es menor en relación a la media del segmento V, a pesar de que el número de clientes en B es mayor que en V. Esto nos indica que el comportamiento de las ventas entre estos segmentos varía y por lo tanto se rechaza la hipótesis nula planteada, ya que no existe suficiente evidencia estadística para determinar que las medias son iguales.

Tabla 20: Prueba de muestras independientes B vs V.

Prueba de muestras independientes									
	Prueba de L. para la igualdad de varianzas		Prueba T para la igualdad de medias						
	F	Sig.	T	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
								Inferior	Superior
Varianzas Venta iguales	1965,239	,000	-38,211	24385	,000	-11945,2862610	312,6106922	-12558,0221785	-11332,5503435
Neta Varianzas distintas			-21,236	5650,905	,000	-11945,2862610	562,5146117	-13048,0308362	-10842,5416858

CAPÍTULO IV

4 Análisis Especializado: Análisis Clúster.

4.1 Análisis clúster

En el análisis de conglomerados se busca patrones en un conjunto o grupos de datos mediante la agrupación de las observaciones.

El objetivo es encontrar un agrupamiento óptimo para que las observaciones u objetos dentro de cada grupo sean similares, pero los grupos sean diferentes entre sí.

Se espera encontrar las agrupaciones naturales en los datos, agrupaciones que tengan sentido para el análisis establecido.

Una vez realizado el respectivo análisis Descriptivo e Inferencial de los datos se procede a aplicar la técnica de Análisis Clúster, con la finalidad de establecer nuevas agrupaciones, se espera que sean diferentes a los segmentos previamente establecidos por la empresa, tomando como base de referencia para esta agrupación la variable de las ventas netas analizada previamente

4.1.1 Aplicación del Análisis.

Para establecer este análisis se ha procedido a realizar una sumatoria total de las ventas netas mensuales por cliente, para así obtener una sola variable condensada con los valores de las ventas que se daban para cada año, mes a mes de cada cliente.

Una vez establecido este parámetro se procedió a calcular la matriz de distancias de la variable mencionada anteriormente.

En este caso de análisis se utilizará el **enfoque de agrupación jerárquica**, el cual se lo ha detallado con anterioridad.

Para este caso puntual se ha hecho uso de la función `h clust` con la ayuda del software R.

Dando como resultado un dendograma donde nos indica, que la mayoría de las ventas netas para cada tipo de cliente en sus diferentes segmentos se encuentran por debajo del millón de dólares, tal como se pudo visualizar en los análisis anteriores.

Por lo tanto para tener una idea clara de cómo se debería separar a los clientes por segmentos se procede a realizar puntos de corte en 5 millones, 1 millón, 500 mil, 100 mil, 50 mil, 10mil y se observan las diferentes agrupaciones que van tomando los clientes para cada nuevo segmento.

Realizando el primer punto de corte obtenemos:

GRUPO 1 (Clientes cuyo valor de ventas sean mayores a 5 millones)

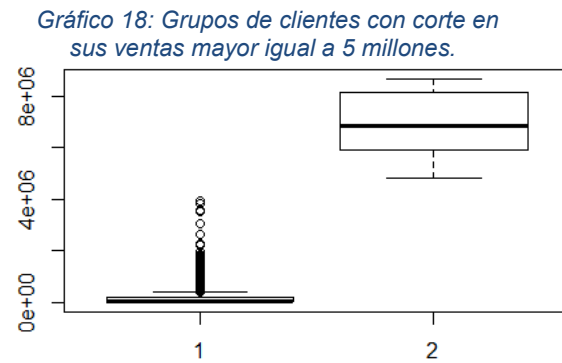


Tabla 21: N° de clientes con ventas mayor igual a 5 millones.

1	2
3937	7

Se muestran dos grupos donde, el segundo grupo contiene a los clientes con ventas mayores e iguales a 5 millones. En total son 7 clientes.

Venta Neta	Segmento	Cod cliente
\$ 3.610.894,00	V	c2015
\$ 2.489.120,00	V	c2066
\$ 2.846.186,00	J	c3668
\$ 4.378.156,00	J	c3670
\$ 4.567.318,00	J	c3674
\$ 4.074.419,00	J	c3678
\$ 3.286.176,00	J	c3679

Como podemos observar los clientes que pasan a formar parte de este nuevo segmento formaban parte de los segmentos V y J, e acuerdo al código del cliente que tienen establecidos se puede ubicar a los clientes pertenecientes a cada segmento anterior.

GRUPO 2 (Clientes cuyo valor de ventas sean mayores a 1 millón)

Gráfico 19: Grupos de clientes con corte en sus ventas mayor igual a 1 millón.

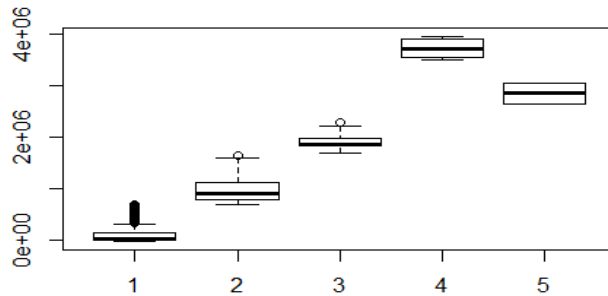


Tabla 22: N° de clientes con ventas mayor igual a 1 millón.

1	2	3	4	5
3750	170	11	4	2

Se muestran cinco grupos donde, el segundo, tercer, cuarto y quinto grupo contiene a los clientes con ventas mayores e iguales a 1 millón. En total son 187 clientes que pasan a formar parte de este nuevo grupo.

Se puede notar que los clientes de esta nueva agrupación eran parte de la segmentación de inicial de los grupos C, V, D, A, E, J, T.

Segmento	Cod. Cliente
B	c139
B	c525
B	c943
C	c1167
C	c1168
C	c1169
C	c1173
C	c1177
C	c1180
C	c1181
C	c1182
C	c1187
C	c1189
C	c1193
C	c1194
C	c1202
C	c1203
C	c1204
C	c1205
C	c1206
C	c1207
C	c1208
C	c1209
C	c1210
C	c1211
C	c1212
C	c1213
C	c1214
C	c1215
C	c1216
C	c1217
C	c1218
C	c1219
C	c1220
C	c1221
C	c1222
C	c1223
C	c1224
C	c1226
C	c1231
C	c1235
C	c1237
C	c1246
C	c1248
C	c1250
C	c1252
C	c1265
C	c1271
C	c1272
C	c1275
C	c1279
C	c1285
C	c1291
C	c1293
C	c1297
C	c1299
C	c1307

C	c1307
C	c1314
C	c1316
C	c1318
C	c1319
C	c1321
C	c1322
C	c1361
C	c1569
C	c1572
C	c1574
C	c1577
C	c1578
C	c1580
C	c1581
C	c1583
C	c1584
C	c1589
C	c1590
C	c1594
C	c1595
C	c1596
C	c1597
C	c1598
C	c1599
C	c1602
C	c1603
C	c1605
C	c1607
C	c1608
C	c1609
C	c1611
C	c1612
C	c1624
C	c1635
C	c1636
C	c1643
C	c1644
C	c1649
C	c1662
C	c1663
C	c1665
C	c1667
C	c1669
C	c1675
C	c1677
C	c1678
C	c1679
C	c1680
C	c1681
C	c1683
C	c1685
C	c1686
C	c1689
C	c1690
C	c1701
C	c1714
C	c1716
C	c1718
C	c1724
C	c1725
C	c1735
C	c1749
C	c1760
C	c1761
C	c1764
C	c1767
C	c1768
C	c1770
C	c1771
C	c1772
C	c1773
C	c1794
C	c1797
C	c1803
C	c1806
C	c1810
C	c1815
C	c1817

GRUPO 3 (Clientes cuyo valor de ventas sean mayores e iguales a 500 mil)

Gráfico 20: Grupos de clientes con corte en sus ventas mayor igual a 500 mil.

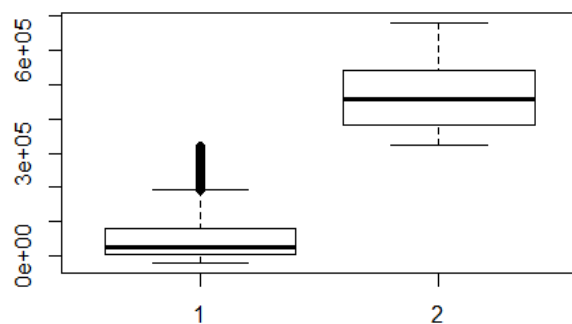


Tabla 23: N° de clientes con ventas mayor igual a 500 mil.

1	2
337	413

Se muestran dos grupos donde, el segundo grupo contiene a los clientes con ventas mayores e iguales a 500 mil dólares. En total son 413 clientes que pasan a formar parte de este nuevo grupo.

GRUPO 4 (Clientes cuyo valor de ventas sean mayores e iguales a 100 mil)

Gráfico 21: Grupos de clientes con corte en sus ventas mayor igual a 100 mil.

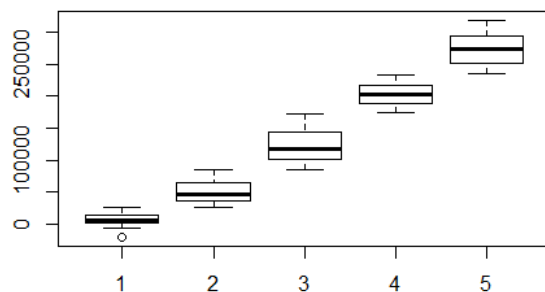


Tabla 24: N° de clientes con ventas mayor igual a 100 mil.

1	2	3	4	5
1727	816	417	173	204

Se muestran 5 grupos en donde a partir del grupo 3 hasta el 5 poseen clientes con ventas mayores e iguales 100 mil dólares. En total son 794 clientes que pasan a formar parte de este nuevo grupo.

GRUPO 5 (Clientes cuyo valor de ventas sean mayores e iguales a 50 mil)

Gráfico 22: Grupos de clientes con corte en sus ventas mayor igual a 50 mil.

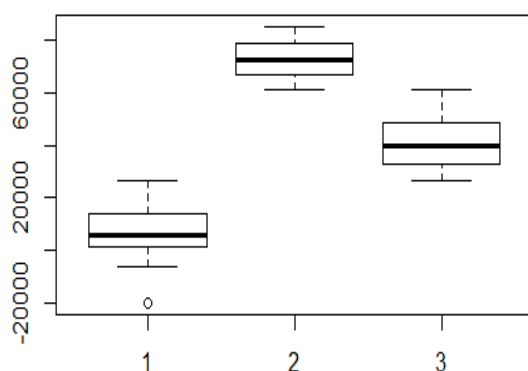


Tabla 25: N° de clientes con ventas mayor igual a 50 mil.

1	2	3
1727	235	581

Se muestran 2 grupos en donde los grupos dos y tres poseen clientes con ventas mayores e iguales 50 mil dólares. En total son 816 clientes que pasan a formar parte de este nuevo grupo.

GRUPO 6 (Clientes cuyo valor de ventas sean mayores e iguales a 10 mil)

Gráfico 23: Grupos de clientes con corte en sus ventas mayor igual a 10 mil.

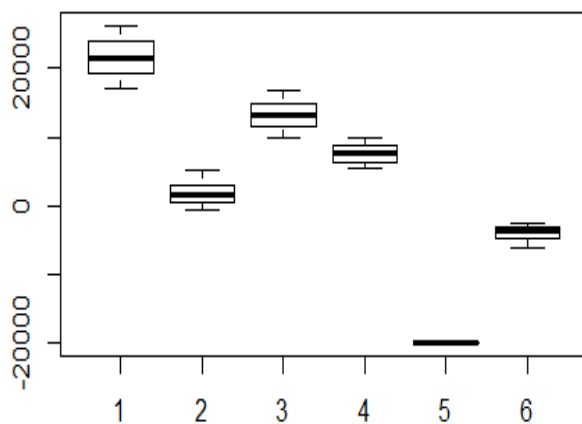


Tabla 26: N° de clientes con ventas mayor igual a 10 mil

1	2	3	4	5	6
304	841	318	260	1	3

Se muestran 6 grupos en donde los grupos uno y tres poseen clientes con ventas mayores e iguales 10 mil dólares. En total son 622 clientes que pasan a formar parte de este nuevo grupo.

CONCLUSIONES

- Se realizó el respectivo análisis de la segmentación establecida por la empresa, en el cual se pudo notar que la cantidad de cantidad de clientes en los diferentes segmentos, no influyen de manera directa con el comportamiento de compra de cada uno de ellos.
- Con el análisis clúster se obtuvieron 6 grupos separados en base a relación en la cantidad de ventas que guardan cada uno de los clientes.
- Con esta nueva segmentación se tiene una visualización clara de cuáles son los clientes que aportan de manera significativa con sus compras a la empresa, basados netamente en el comportamiento de compra de cada uno de ellos.

RECOMENDACIONES

En base a los resultados obtenidos durante el proceso de extracción de los datos, depuración de los mismos y análisis se plantean las siguientes recomendaciones:

- Los clientes que la empresa considera como clientes Vip, no tienen un comportamiento significativo en cuanto a sus ventas, por lo tanto se debería considerar, establecer nuevos beneficios para este tipo de clientes.
- Tomar en consideración la propuesta planteada, pues se basa en el comportamiento de compra que posee cada cliente, y con esto la empresa puede designar de manera segura los beneficios que poseen cada uno de los clientes, de acuerdo al grado de importancia que se establezca.

REFERENCIAS BIBLIOGRÁFICA

5 Bibliografía

Luis Rodríguez Ojeda, PROBABILIDAD Y ESTADÍSTICA BÁSICA PARA INGENIEROS Con el soporte de MATLAB para cálculos y gráficos estadísticos, (Escuela Superior Politécnica del Litoral, ICM, Guayaquil-Ecuador 2007).

Cazau, P. *Fundamentos de estadística*.

Manuel Terrádez Gurrea, <http://www.uoc.edu/in3/emath/docs/Cluster.pdf>.

Methods of Multivariate Analysis, Second Edition, Alvin C. Rencher, Brigham Young University, WILEY-INTERSCIENCE A JOHN WILEY & SONS, INC. PUBLICATION 2002.

Microsoft. (2015). *SQL Server 2008*. Recuperado el 28 de Julio de 2015, de <https://technet.microsoft.com/es-es/library/ms159640%28v=sql.100%29.aspx>