



CÁLCULO DEL VALOR DEL CLIENTE USANDO AHP, TÉCNICA RFM, Y
TÉCNICAS DE CLUSTERING PARA APOYAR LAS ESTRATEGÍAS DE
MARKETING EN UNA EMPRESA COMERCIALIZADORA DE LA CIUDAD DE
CALI

ANGÉLICA MARÍA PINEDA HERNÁNDEZ

UNIVERSIDAD DEL VALLE
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INDUSTRIAL Y ESTADÍSTICA
INGENIERÍA INDUSTRIAL
SANTIAGO DE CALI
2010



CÁLCULO DEL VALOR DEL CLIENTE USANDO AHP, TÉCNICA RFM, Y
TÉCNICAS DE CLUSTERING PARA APOYAR LAS ESTRATEGÍAS DE
MARKETING EN UNA EMPRESA COMERCIALIZADORA DE LA CIUDAD DE
CALI

ANGÉLICA MARÍA PINEDA HERNÁNDEZ

Trabajo de Grado para Optar por el Título de:
Ingeniero Industrial

Director
Juan José Bravo Bastidas,
MSc en Ingeniería de Sistemas,
Ingeniero Industrial

UNIVERSIDAD DEL VALLE
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INDUSTRIAL Y ESTADÍSTICA
INGENIERÍA INDUSTRIAL
SANTIAGO DE CALI
2010

NOTA DE ACEPTACIÓN

Presidente del Jurado: _____

Jurado: _____

Jurado: _____

Santiago de Cali, 2010

CONTENIDO

INTRODUCCIÓN	11
1. PLANTEAMIENTO DEL PROBLEMA.....	13
2. OBJETIVOS.....	15
2.1. OBJETIVOS GENERALES	15
2.2. OBJETIVOS ESPECIFICOS	15
3. JUSTIFICACIÓN.....	16
4. ESTADO DEL ARTE	18
5. MARCO TEÓRICO	21
5.1. LA IMPORTANCIA DE LAS ESTRATEGIAS DE MARKETING EN LAS ORGANIZACIONES.....	21
5.2. TÉCNICA RFM (RECIENTE, FRECUENTE, Y MONETARIO)	22
5.3. PROCESO JERÁRQUICO ANALÍTICO (AHP)	23
5.4. CLUSTERING	29
5.5 CLASIFICACIÓN DEL VALOR DEL CLIENTE (CLV)	37
6. CASO DE ESTUDIO.....	39
7. EXPERIENCIA COMPUTACIONAL	41
7.1. OBTENCIÓN DE LA BASE DE DATOS.....	42
7.2. DEPURACIÓN DE LA BASE DE DATOS	42
7.3. EXTRACCIÓN DE LAS VARIABLES RFM	43
7.4. CÁLCULO DE LA IMPORTANCIA RELATIVA DE LAS VARIABLES RFM	49

7.5. ANÁLISIS DE LAS VARIABLES RFM.....	56
7.6. NORMALIZACIÓN Y PONDERACIÓN DE LAS VARIABLES RFM	60
7.7. AGRUPACIÓN DE LOS CLIENTES EN CLÚSTERES	65
7.8. ANÁLISIS DE LOS GRUPOS RESULTANTES DEL K-MEANS	71
7.9. CÁLCULO DEL VALOR DEL CLIENTE (CLV).....	80
8. DISCUSIÓN Y ANÁLISIS DE RESULTADOS	84
9. CONSLUSIONES Y RECOMENDACIONES	88
BIBLIOGRAFÍA	89
ANEXOS	91

LISTA DE FIGURAS

<i>Figura No. 1, Registros de las Transacciones de los Clientes de la Base de Datos de la Empresa</i>	<i>39</i>
<i>Figura No. 2, Metodología Propuesta.....</i>	<i>41</i>
<i>Figura No. 3, Cálculo del Valor de la Variable R (Reciente).....</i>	<i>45</i>
<i>Figura No. 4, Cálculo del Valor de la Variable F (Frecuencia).....</i>	<i>46</i>
<i>Figura No. 5, Cálculo del Valor de la Variable M (Monetario).....</i>	<i>48</i>
<i>Figura No. 6, Comparación entre las Variables R y F</i>	<i>51</i>
<i>Figura No. 7, Comparación entre las Variables R y M</i>	<i>51</i>
<i>Figura No. 8, Comparación entre las Variables F y M.....</i>	<i>52</i>
<i>Figura No. 9, Inicializando el Cálculo del Coeficiente de Correlación en Minitab ..</i>	<i>59</i>
<i>Figura No. 10, Correlación en Minitab</i>	<i>59</i>
<i>Figura No. 11, Cálculo de X^L y X^S</i>	<i>62</i>
<i>Figura No. 12, Valores Normalizados de las Variables RFM (R_N, F_N, M_N)</i>	<i>63</i>
<i>Figura No. 13, Cálculo del Producto entre los Valores Normalizados de RFM y las Importancias Relativas de las Variables RFM, (R_N', F_N', M_N')</i>	<i>64</i>
<i>Figura No. 14, Ingreso de Valores de las Variables R_N' F_N' M_N' en Minitab</i>	<i>67</i>
<i>Figura No. 15, Inicializando la Agrupación de Clientes por medio de K – Means en Minitab.....</i>	<i>68</i>
<i>Figura No. 16, K – Means en Minitab</i>	<i>69</i>
<i>Figura No. 17, Columna de Membresía en Minitab</i>	<i>69</i>

<i>Figura No. 18, Columna de Membresía en la Hoja de Trabajo de Minitab una vez Obtenida la agrupación por K-mean.....</i>	<i>70</i>
<i>Figura No. 19, Inicializando el ANOVA (Análisis de Varianza) en Minitab</i>	<i>72</i>
<i>Figura No. 20, ANOVA (Análisis de Varianza) en Minitab</i>	<i>73</i>
<i>Figura No. 21, ANOVA para la Variable R_N' en Minitab</i>	<i>74</i>
<i>Figura No. 22, ANOVA para la Variable F_N' en Minitab.....</i>	<i>75</i>
<i>Figura No. 23, ANOVA para la Variable M_N' en Minitab.....</i>	<i>76</i>
<i>Figura No. 24, Cálculo del Valor Promedio de las Variables RFM de cada Grupo o Clúster en Excel</i>	<i>78</i>
<i>Figura No. 25, Cálculo del Valor Promedio de las Variables $R_N F_N M_N$ para cada Grupo de Clientes</i>	<i>81</i>
<i>Figura No. 26, Cálculo de la Evaluación Integrada para todos los Grupos</i>	<i>82</i>

LISTA DE TABLAS

<i>Tabla No. 1, Estructura de la Base de Datos</i>	<i>23</i>
<i>Tabla No. 2, Escala de Puntuación usada en la Comparación en Pares para el Cálculo de los Pesos Relativos de las Variables RFM</i>	<i>24</i>
<i>Tabla No. 3, Matriz de Comparación en Pares de las Variables RFM</i>	<i>25</i>
<i>Tabla No. 4, Matriz Inicializada de Comparación en Pares de las Variables RFM</i>	<i>25</i>
<i>Tabla No. 5, Matriz Completa de Comparación en Pares de las Variables RFM ..</i>	<i>25</i>
<i>Tabla No. 6, Índice de Consistencia Aleatoria.....</i>	<i>28</i>
<i>Tabla No. 7, Promedios de las Variables RFM para cada Grupo de Clientes</i>	<i>34</i>
<i>Tabla No. 8, Clasificación del Valor del Cliente (CLV)</i>	<i>38</i>
<i>Tabla No. 9, Matriz de Comparación en Pares.....</i>	<i>50</i>
<i>Tabla No. 10, Resumen del Resultado de la Encuesta</i>	<i>52</i>
<i>Tabla No. 11, Matriz de Comparación Pareada</i>	<i>52</i>
<i>Tabla No. 12, Matriz de Comparación en Pares Completa</i>	<i>53</i>
<i>Tabla No. 13, Sumatoria de las Columnas B, C y D.....</i>	<i>53</i>
<i>Tabla No. 14, Matriz de Normalización</i>	<i>54</i>
<i>Tabla No. 15, Matriz Normalizada Completa.....</i>	<i>54</i>
<i>Tabla No. 16, Cálculo del Índice de Consistencia (CI)</i>	<i>54</i>
<i>Tabla No. 17, Importancia Relativa de las Variables RFM</i>	<i>55</i>
<i>Tabla No. 18, Correlaciones entre las Variables R, F y M arrojadas por Minitab ..</i>	<i>58</i>
<i>Tabla No. 19, Valores Originales de las Variables RFM</i>	<i>61</i>

<i>Tabla No. 20, Posibles Combinaciones de las Entradas (RFM).....</i>	<i>65</i>
<i>Tabla No. 21, Medias de cada Variable $R_N'F_N'M_N'$ en los Clústeres</i>	<i>70</i>
<i>Tabla No. 22, Número de Clientes u Observaciones en cada Grupo arrojado por Minitab.....</i>	<i>71</i>
<i>Tabla No. 23, Resultados del K - Means entre 158 Clientes.....</i>	<i>78</i>
<i>Tabla No. 24, Valores de C_R^j, C_F^j, y C_M^j</i>	<i>82</i>
<i>Tabla No. 25, CLV (Valor del Cliente)</i>	<i>83</i>
<i>Tabla No. 26, Tipos de Clientes de acuerdo a la Clasificación de la Lealtad.....</i>	<i>87</i>

LISTA DE GRAFICAS

<i>Gráfica No. 1, Diferencia entre las Medias de las Variables R de los Ocho Grupos</i>	56
<i>Gráfica No. 2, Diferencia entre las Medias de las Variables F de los Ocho Grupos</i>	57
<i>Gráfica No. 3, Diferencia entre las Medias de las Variables M de los Ocho Grupos</i>	57

INTRODUCCIÓN

En la actualidad, la mayoría de los países, independientemente de su grado de desarrollo económico, reconocen la importancia del marketing, y es por esta razón que las empresas diseñan estrategias en torno a ello.

Muchas compañías han decidido que, para tener éxito, necesitan centrar su atención en las necesidades de sus clientes y llevar a cabo una gama más alta de marketing. El concepto de marketing insiste en la orientación hacia el cliente y en la coordinación de las actividades de marketing para lograr los objetivos de desempeño de la organización.¹

Las actividades de marketing coordinadas y con orientación hacia el cliente son los medios usados para lograr el fin que se busca, es decir, el logro de los objetivos de desempeño de la organización.

Teniendo en cuenta que los mercados son más agresivos y más competitivos, el reto principal de toda empresa que quiera sobrevivir en él, debe ser el mayor nivel de satisfacción para sus clientes, debido a que en las manos de estos clientes está el benéfico para la empresa. Existen clientes que son indispensables para la compañía, y es a éstos a quienes hay que direccionar las estrategias de marketing que se diseñen para este fin. Un cliente con tal poder se convierte en un cliente leal si se le ofrece productos o servicios que satisfagan sus necesidades.

Este estudio presenta la forma en que una empresa identifica los clientes más leales, quienes son los más rentables y ofrecen mayores beneficios para la compañía a diferencia de los demás, a través del cálculo del *Valor del Cliente*, lo cual puede servir de ayuda para que una empresa genere estrategias de marketing en torno a esto, logrando así la retención de este tipo de clientes durante un largo periodo de tiempo, o si es posible, durante toda la vida.

Para tal fin, se hará uso de un método que involucra datos tanto cuantitativos como cualitativos con la ayuda de diversos métodos tales como: la *Técnica RFM*, herramientas multicriterio como el *AHP* (Proceso Jerárquico Analítico), y herramientas estadísticas de *Clustering*.

¹ STANTON, William J., ETZEL. Michael J., y WALKER. Bruce J. Fundamentos de Marketing. Mc Graw Hill, Novena edición, 1993. P. 6.

Una base de datos de clientes de una empresa es el punto de partida de esta investigación, la cual deberá tratarse con el fin de extraer ciertas variables, las cuáles determinan ciertos grupos de clientes. Una vez se tengan estas variables (RFM), a través de la *Técnica RFM* (Reciente, Frecuencia, y Monetario), se proseguirá a determinar los pesos de las variables *RFM* mediante la herramienta multicriterio *AHP*. Las ponderaciones de estas variables permitirán que los clientes de una empresa sean agrupados mediante la *Técnica de Agrupación K-Means*, haciendo uso del software *Minitab 15*. Una vez se tengan los grupos de clientes o Clústeres resultantes, se aplica la metodología para el cálculo del *Valor del Cliente* o la *Lealtad del Cliente*, con el fin de identificar los clientes más leales.

En la actualidad, muchos planes de marketing se implementan sin resultados satisfactorios desde el punto de vista de las ventas. La hipótesis de esta investigación es que el método ya mencionado permite aumentar la eficiencia de estos planes.

Este estudio muestra la factibilidad del uso del método propuesto y para ello se ha respaldado de una amplia literatura relacionada con diversos casos aplicados al marketing. El empleo de modelos matemáticos que se desarrollan a largo de este estudio resultan de gran importancia, debido a los conceptos implícitos en ellos, a aquellas variables y restricciones que caracterizan al problema.

Como caso aplicativo, la metodología propuesta se implementará en una empresa comercializadora de la ciudad de Cali, lo cual generará un mayor beneficio en ésta tal como se podrá confrontar en los resultados obtenidos en el estudio. También proporcionará pautas para desarrollar una estrategia de marketing.

El desarrollo de estrategias encaminadas a la retención de los clientes, ha sido un asunto que requiere mejorar la gestión de la relación que la empresa tiene con sus clientes. La comunicación entre la empresa y cliente, debe ser excelente y ser la suficiente como para siempre saber lo que los clientes desean y necesitan, y así poder brindárselos.

En primera instancia se determinará el problema que se va a desarrollar dentro de este trabajo, el cual está relacionado con la poca capacidad que tienen algunas empresas para lograr la identificación y retención de los clientes más leales, los cuales garantizan un mayor beneficio para ésta. Luego se determinarán los objetivos a seguir; la justificación, la cual hace que este estudio tenga validez; y luego se mostrará un estado de arte, donde se recopilará toda la literatura revisada relacionada con el problema de estudio, y después se proseguirá con el caso de estudio, la experiencia computacional, discusión y análisis de resultados, y por último las conclusiones.

1. PLANTEAMIENTO DEL PROBLEMA

El que una empresa no tenga problemas graves y que no esté mal financieramente, no significa que no pueda mejorar. La implementación de una nueva metodología en una empresa puede lograr que ésta obtenga mejores beneficios de los que tenía antes de ser implementada.

Cualquier empresa que desee sobrevivir en un mercado competitivo debe tener en cuenta que el marketing es indispensable en este proceso. El marketing implica que la empresa se comprometa realmente con sus clientes, y es por eso, que es importante una buena comunicación con ellos. El estudio del comportamiento de los consumidores es de gran ayuda para las empresas que buscan satisfacer adecuadamente las necesidades de éstos, buscando así mismo alcanzar el éxito.

Desde la perspectiva del marketing, muchas empresas no logran obtener clientes rentables y leales para toda la vida. Para hacer que esto sea posible, es necesario que las empresas diseñen e implementen estrategias direccionadas a retener clientes, y que se centren en aquellos que le son más rentables.

El enfoque del marketing parte de la base de que la empresa tenga lo que los consumidores necesitan o quieren. Lo anterior dá como resultado la retención de clientes, pero no de cualquier cliente, hay que centrarse en los clientes rentables y leales, para favorecer su retención a largo plazo. Una vez se tiene claro los clientes a los cuales hay que enfocarse para brindarles atención prioritaria, se debe evaluar el impacto que esta medición tiene sobre las estrategias de marketing.

Muchas empresas, específicamente las mipymes no tienen bien definido cómo lograr la expansión o consecución de nuevos clientes a partir del uso de diversas herramientas que involucran datos tanto cualitativos como lo son las herramientas multicriterio, y otras técnicas empleadas por el marketing para determinar cuáles son los clientes más importantes para la empresa, ya que la relación entre cliente y proveedor en estos casos es más directa, lo cual hace más fácil manejar estas relaciones de modo amistoso, con amabilidad, y así lograr la satisfacción de los clientes.

El cliente debe ser lo más importante en una organización, y el estudio que se realice en torno a éste debe ser tomado como una actividad prioritaria, y debe

prestársele la suficiente atención posible. Los encargados de esto, los tomadores de decisión, se encuentran con el problema de que no saben qué tipo de clientes son los importantes, cuando se trata de una empresa que tiene muchos clientes.

En este sentido, la pregunta que el trabajo de grado desea responder es: ¿De qué manera la medición del *Valor del Cliente* o de la *Lealtad de los Clientes* puede favorecer el diseño de estrategias de marketing?

2. OBJETIVOS

2.1. OBJETIVOS GENERALES

Calcular el *Valor del Cliente (CLV)* o la *Lealtad del Cliente* en una empresa comercializadora, con el fin de brindar apoyo a la toma de decisiones en marketing asociado a la retención de clientes.

2.2. OBJETIVOS ESPECIFICOS

- Seleccionar la información específica de la literatura revisada, que esté relacionada con la retención de los clientes, la búsqueda de nuevos clientes, la satisfacción de éstos por parte de las empresas, y las técnicas usadas por el marketing para tal fin.
- Documentar y describir un caso de estudio, explorando el comportamiento de los clientes en la empresa elegida por medio de un análisis de las transacciones considerando las variables **RFM** (Reciente-Frecuencia de compra- Monto de la compra).
- Aplicar la medición **CLV** y concluir sobre las implicaciones de los resultados en las estrategias de retención.

3. JUSTIFICACIÓN

En la actualidad los mercados tienen que ser más competitivos ya que día a día se desarrollan nuevas metodologías dirigidas a retener y satisfacer a los clientes.

Las mejores relaciones entre cliente y empresa se logran gracias a una clara y concreta definición y clasificación de los clientes, la cual permite saber qué clientes se deben retener, ya que son significativamente importantes (leales) debido a los beneficios que brindan a lo largo de su relación, y mantener la lealtad por un largo tiempo.

La búsqueda de la eficiencia y la productividad de una compañía, de los sectores industriales están contribuyendo a adoptar metodologías de apoyo en la toma de decisiones, en general, y para el fomento de la competitividad, en particular, en escenarios donde intervienen múltiples variables o criterios de selección.²

Por consiguiente, esta búsqueda no es ajena en aquellos ambientes que son muy competitivos, en donde las empresas día a día ven la importancia de tener una relación más estrecha con sus clientes, y así poder maximizar sus ganancias. Uno de los métodos más empleados, son los métodos de decisión multicriterio, lejos de ser considerados elementos infalibles y certeros, cuya utilización permite encontrar una solución óptima u definitiva, son una base, sustentada en elementos científicos, que aporta mejoras distintivas para asumir una decisión.³

Teniendo en cuenta los medios informativos económicos, se puede concluir que las grandes empresas han trabajado en torno a la fidelización de los clientes y a la retención de unos nuevos, y han desarrollado diversas estrategias en torno a esto, pero las mipymes, como lo es la empresa que se presentará en el caso de estudio para la validación del método propuesto en este estudio, aún no se han desarrollado muy bien en este campo, y muy pocas se han interesado por mejorar la relación con sus clientes.

Dentro de este contexto se hace necesario medir el *Valor de los Clientes* o de la *Lealtad de los Clientes*, para así también poder estimar o determinar sus gustos y necesidades, de una posible próxima compra. Por esa razón es necesario

² BERUMEN. Sergio A., REDONDO LLAMAZARES. Francisco. La utilidad de los Métodos de Decisión Multicriterio (como el AHP) en un entorno de competitividad creciente. 2007. P. 67.

³ Ibid, p. 67.

implementar una metodología que permita medir la lealtad de cada cliente de una empresa, y a partir de esto recomendar la metodología propuesta como apoyo a las *estrategias de marketing*.

El fracaso de muchas empresas se debe a que no tienen en claro el comportamiento de sus clientes más leales. Al conocer mejor sus clientes, se hace un gran trabajo, se obtienen ideas novedosas, y lo más importante, se generan ideas que los satisfacen, lo cual hace que éstos elijan un producto o servicio determinado.

4. ESTADO DEL ARTE

La medición del *Valor del Cliente* (CLV) relacionada con las técnicas de *Clustering*, las herramientas multicriterio (AHP), y con la técnica *RFM*, ha sido objeto de estudio en el campo del marketing, aunque no ha sido muy conocida. Por tal motivo se documentará a continuación, trabajos relacionados con el cálculo del *Valor del Cliente*, ya sea por la metodología propuesta a continuación, o por otras; y también otras metodologías relacionadas con la satisfacción y retención de clientes, y la predicción del comportamiento de éstos.

En el año 2000 se publicó un artículo llamado “***Brand Choice Model Selection based on Consumers Multicriteria Preferences and Experts Knowledge***”⁴, el cual muestra que el comportamiento y las actitudes de los consumidores pueden ser modelados matemáticamente, y además demuestra que el *análisis multicriterio* es una poderosa herramienta para predecir patrones determinantes de los consumidores en el momento de escoger un producto o servicio ante una serie de alternativas. Los modelos que permiten predecir el comportamiento de los consumidores junto con la *Teoría de Utilidad* le permiten al analista explorar y describir un modelo de acuerdo a la distribución de las utilidades asignadas. De esta forma el decisor obtiene un instrumento adicional para modelar de manera más eficiente el ambiente del mercado. Este artículo presenta un método heurístico y una base de conocimiento para la selección del modelo “*Selección de Marca*”. El modelo desarrollado está basado en dos factores: (1) El comportamiento de los consumidores es medido por el tipo de distribución de utilidades asignadas; y (2) La observación de la aplicabilidad de los modelos específicos de la *Selección de Marca* sobre unos ciertos tipos de comportamientos de consumidores. Todos los comportamientos de los consumidores que suelen ser difíciles de observar, son modelados por un simple “si...entonces”.

“***A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clusterig Analysis***”⁵, es un artículo publicado en el año 2003. Este artículo propone que la mejor estrategia de marketing es capturar clientes que le sean rentables a la empresa, y así mismo mantener aquellos que son leales y que ya tienen un largo periodo de relación con la empresa. Se

4 MATSATSINI, Nikolaos F. Brand Choice model selection based on consumers’ multicriteria preferences and experts’ Knowledge. 2000.

5 SHIH, Ya-Yueh A method for customer lifetime value ranking – Combining the analytic hierarchy process and clustering analysis. 2003.

propone una técnica llamada, *RFM (Reciente, Frecuencia, y Monetario)* aplicada a una base de datos de clientes, para luego calcular la importancia relativa de las variables *RFM*, a través del *Proceso Jerárquico Analítico (AHP)*, con la cual se evalúa el *CLV (Valor del Cliente)*. Las técnicas de *Clustering* son empleadas para agrupar o segmentar los clientes de acuerdo a los valores ponderados de las variables *RFM*. La simple ponderación es usada para obtener la Evaluación del *CLV*, y también para segmentar los consumidores y poderlos de esta forma, comparar claramente. Este estudio proporciona una clasificación del *CLV*, y ayuda a que se realice una segmentación más efectiva. Resultados experimentales indican que el método propuesto produce una más razonable evaluación del *CLV* que los métodos que no consideran la importancia relativa de las variables *RFM*. Este estudio valida la metodología propuesta usando un algoritmo de árbol de decisión por las razones de acierto de la clasificación calculada y la construcción de un modelo para predecir qué otro consumidor potencial puede existir.

El artículo ***“Integrating AHP and Data Mining for Product Recommendation based on Customer Lifetime Value⁶”*** publicado en el año 2004 evalúa una metodología recomendada que combina las técnicas de *“Data Mining”* con la toma de decisiones de grupo. El *Procesos Jerárquico Analítico (AHP)* fue aplicado dentro de este estudio para calcular los pesos relativos de las variables *RFM* (Reciente, Frecuencia, y Monetario), los cuales son necesarios para calcular el *CLV (Valor del Cliente)*, y de esta forma predecir el comportamiento de los consumidores con la ayuda de una base de datos. Este estudio incluye una metodología nueva que combina el *AHP*, el *Clustering*, y una asociación de métodos a base de reglas. Los resultados experimentales muestran que la metodología propuesta de verdad puede proporcionar recomendaciones de una calidad más alta.

En el mismo año, se publicó el artículo, ***“Measuring Consumer for a Better Comprehension, Specification and Assessment of Product Semantics⁷”***, el cual presenta un acercamiento general para evaluar la semántica de Productos (es un cambio en el paradigma del diseño de las cosas, que sugiere enfocarse en que los productos se adapten a lo que se necesita, es decir, no enfocarse en cómo deban funcionar sino en lo que los productos significan para quiénes lo usan) por un camino sano. Este acercamiento está basado en pruebas de utilidad, e involucra varios métodos clásicos usados en el marketing y en la teoría de toma de decisiones. Esta investigación provee a los diseñadores un instrumento que ayuda a entender la necesidad de especificar la parte semántica. Dentro de este estudio se han realizado pruebas de utilidad aplicadas a la metodología del diseño de vasos tallados. Los conceptos de productos emergentes pueden ser evaluados de acuerdo a los requerimientos en una forma comparativa y cualitativa simple.

6 LIU, Duen-Ren. Integrating AHP and data mining for product recommendation based on customer lifetime value. Information & Management. 2004.

7 PETIOT, Jean-Francois. Measuring consumer for a better comprehension, specification and assessment of product semantics. 2004.

En conclusión, desde el año 2000 se encontraron una serie de artículos, los cuales muestran un gran interés por la predicción del comportamiento de los consumidores frente a un producto o servicio determinado. Este interés se pudo percibir por medio de la intención que expone cada artículo, al tratar de entender a sus clientes, a través de diferentes herramientas o modelos, tales como: herramientas multicriterio, Data Mining, Técnica RFM, Modelo Brand Choice y Semántica de Productos. Estos artículos sirvieron de guía para el desarrollo del estudio que se expondrá en este documento, aunque el más determinante fue el artículo llamado *“A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clusterig Analysis”*, en el cual está expuesta la metodología usada en este documento, la cual será aplicada a un caso de estudio particular.

5. MARCO TEÓRICO

5.1. LA IMPORTANCIA DE LAS ESTRATEGIAS DE MARKETING EN LAS ORGANIZACIONES

El éxito de muchas empresas depende, en gran parte, de la capacidad de considerar con anticipación lo que harán sus clientes. Por tal motivo, es importante conocer el mercado en el cual la empresa se encuentra, y también el mercado objetivo, para saber hacia qué población debe dirigir sus productos o servicios.

Un *mercado* consiste en personas u organizaciones con necesidades a satisfacer, dinero para gastar y el deseo de hacerlo. Un *mercado objetivo* comprende de un grupo de consumidores (personas u organizaciones) a quienes el vendedor dirige una *estrategia de marketing*.⁸

El marketing es un sistema total de actividades de negocios diseñado para planear, fijar precios, promocionar y distribuir productos que satisfacen necesidades a mercados objetivos para lograr las metas organizacionales. Y las *estrategias de marketing* son un plan de acción amplio mediante el cual una organización intenta lograr el objetivo del *marketing*.

El mercado total para la mayor parte de los productores es demasiado variado (demasiado heterogéneo) para poderlo considerar como una sola entidad uniforme, esta variabilidad se puede atribuir a las diferencias en los hábitos de compra, en las formas en que se usa el bien o servicio, en los motivos para comprar, o en otros factores. La *Segmentación del Mercado* toma en cuenta tales factores.

La *Segmentación del Mercado*⁹ es el proceso de dividir el mercado total heterogéneo para un bien o servicio en varios segmentos, cada uno de los cuales tiende a ser homogéneo en todos los aspectos importantes. Las empresas seleccionan uno o más de estos segmentos como *mercado objetivo*.

⁸STANTON, William J., ETZEL, Michael J., y WALKER, Bruce J. Fundamentos de Marketing. Mc Graw Hill, Novena edición, 2003. P.1993.

⁹Ibid, p. 92.

Al adaptar *estrategias de marketing* a los segmentos individuales, las empresas pueden hacer un mejor trabajo de marketing.

Una compañía puede segmentar su mercado en muchas formas diferentes y las bases para hacerlo varían de un producto a otro, y de la atención que ofrece la empresa. Sin embargo, este documento emplea como base de segmentación para los mercados de clientes, el comportamiento del cliente hacia el producto (o bases relacionadas con el producto), por medio de herramientas multicriterio y herramientas estadísticas para el cálculo del Valor del Cliente, o de la lealtad, como se mostrará en el desarrollo de la investigación.

5.2. TÉCNICA RFM (RECIENTE, FRECUENTE, Y MONETARIO)

Para una empresa, conocer sus clientes es un recurso vital para poder llevar a cabo las estrategias de marketing. Por esta razón es importante contar con una base de datos histórica en donde se especifiquen los registros de compras de los clientes, con el fin de conocerlos con más detalle. Estas transacciones le permiten a la compañía conocer las preferencias de sus clientes, sus niveles de ingreso y nivel de consumo, y sobre todo sus gustos. Esta información es valiosa para crear ventajas competitivas, y es por eso que se han desarrollado una serie de técnicas para el manejo de la base de datos de clientes.

Una base de datos de los clientes debe registrar los comportamientos de compra de éstos. Los factores que mejoran la efectividad de una base de datos de clientes está representado por tres variables: (**R**) La fecha de la última compra; (**F**) La frecuencia de compra; y (**M**) el valor del monto de la compra. A esto se le conoce cómo *Técnica RFM*.

A continuación se muestra el significado de cada una de las variables:

- **R (Reciente):** Es el periodo de tiempo desde la última compra realizada por un cliente. Un valor bajo de este indicador significa una alta probabilidad de que el cliente repita la compra.
- **F (Frecuencia):** Es el número de compras realizadas dentro de un periodo de tiempo específico. Una alta frecuencia indica un alto nivel de lealtad de un cliente.
- **M (Monetario):** Es la cantidad de dinero que gasta un cliente durante un

determinado periodo de tiempo.¹⁰

Para aplicar la *Técnica RFM* se debe contar con una base de datos de clientes de un determinado periodo, preferiblemente mayor o igual a dos años, tal como se realizó en el artículo “*A method for customer lifetime value ranking – Combining the analytic hierarchy process and clustering analysis*” mencionado anteriormente en la sección “*Marco Teórico*”, para el caso de estudio que éste aplico. A las transacciones realizadas por los clientes que están en la base de datos, se le debe extraer las variables RFM, en una tabla como se muestra a continuación: (Ver **Tabla No.1**).

Tabla No. 1, Estructura de la Base de Datos

Cliente No.	R (días)	F	M (pesos colombianos)

En la anterior tabla se recolectan las variables extraídas de la base de datos. El cliente puede ser referido como un código o como un nombre, o de las dos formas; la variable **R** se da en días, y éste valor son los días que ha transcurrido desde la última fecha de compra de un cliente específico a la fecha que se toma como punto de partida para la evaluación; la variable **F** está dada en veces por periodo de tiempo; y la variable **M** está dada por una unidad monetaria. Este cálculo se le hace a cada uno de los clientes que tiene la empresa.

Una vez se ha calculado los valores de las variables *RFM*, se emplea *AHP* (*Procesos Jerárquico Analítico*) para evaluar la ponderación de estas variables.

5.3. PROCESO JERÁRQUICO ANALÍTICO (AHP)

El *Proceso Jerárquico Analítico* (AHP) es quizás el método multicriterio más conocido y ampliamente utilizado en numerosos campos por diversas organizaciones, tanto del sector privado como del sector público, por ser de fácil entendimiento, de fácil interpretación, y capaz de producir resultados que

¹⁰ SHIH, Ya-Yueh y LIU, Chung-Yuan. *Hierarchy Process and Clustering Analysis*. Septiembre 24 de 2003. P.161.

coincidan con las experiencias.¹¹ El *AHP* ayuda a las organizaciones a tomar mejores decisiones en una gran variedad de circunstancias. El *AHP* consiste en los siguientes pasos, cuando se aplica a la ponderación de las variables *RFM* (Reciente, Frecuente, y Monetario):

Paso 1: *Comparaciones en Pares*. Este paso cuestiona a los evaluadores (tomadores de decisión) para hacer *Comparaciones en Pares* de los pesos relativos de las variables *RFM*.

El principal objetivo de las *Comparaciones en Pares* en el *AHP* es hacer una comparación de las variables *RFM*. Para compararlas se usa a menudo la siguiente escala: $\{1/9, 1/8, \dots, 1/2, 1, 2, \dots, 8, 9\}$. Esta escala fue introducida por Thomas Saaty, (Ver **Tabla No.2**), y es la escala usada para calcular los pesos relativos de las variables *RFM*.

Tabla No. 2, Escala de Puntuación usada en la Comparación en Pares para el Cálculo de los Pesos Relativos de las Variables *RFM*

Importancia Comparativa	Descripción	Explicación
1	De igual importancia.	Dos actividades contribuyen de igual manera al objetivo.
2	Intermedio entre igual y débil.	La experiencia y un juicio ligeramente a favor de una actividad sobre otra.
3	Importancia débil de una sobre otra.	Una actividad sobre otra.
4	Intermedio entre débil y fuerte.	La experiencia y un juicio fuertemente a favor de una actividad sobre otra.
5	Importancia Esencial o Fuerte.	Una actividad sobre otra.
6	Intermedio entre fuerte y demostrado.	Una actividad que está fuertemente a favor sobre otra, y esta dominancia es demostrada en la práctica.
7	Importancia demostrada.	
8	Intermedio entre demostrado y absoluto.	La evidencia a favor de una actividad sobre otra es de los más altos órdenes posibles de afirmación.
9	Importancia absoluta o extrema.	

Fuente: *A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis*, 24 de Septiembre de 2003, Ya-Yueh Shih, Chung-Yuan Liu, página 164.

¹¹ FORMAN, Ernest H., y GASS, Saul I. *The Analytic Hierarchy Process – An Exposition*. Octubre de 1999. P.485.

Para realizar las *Comparaciones en Pares* se realiza una Matriz de Comparación en Pares, en dónde a modo de ilustración se compararon las tres variables *RFM* (Ver **Tabla No.3**).

Tabla No. 3, Matriz de Comparación en Pares de las Variables RFM

	R	F	M
R			
F			
M			

El *AHP* compara parejas de opciones (por ejemplo, R vs. F, F vs. M, etc.). Como **F** vs. **F**, no es comparable, y así mismo para la comparación que se hace entre las mismas variables, a las respectivas celdas se asigna el valor de uno. Con esto se inicializa la Matriz de Comparación en Pares (Ver **Tabla No.4**).

Tabla No. 4, Matriz Inicializada de Comparación en Pares de las Variables RFM

	R	F	M
R	1		
F		1	
M			1

De acuerdo a la **Tabla No.2**, se prosigue a hacer las comparaciones en pares y así mismo llenar la Matriz de Comparación en Pares. Las comparaciones siempre empiezan “abajo” de la diagonal principal. Estos valores son dados por evaluadores al hacer juicios intuitivos.

En cuanto se llene la matriz por debajo de la diagonal, se prosigue a llenar las celdas ubicadas por arriba de la diagonal, que corresponden al valor inverso de la respuesta que ya se tiene del evaluador (ver **Tabla No.5**).

Tabla No. 5, Matriz Completa de Comparación en Pares de las Variables RFM

	R	F	M
R	1	1/X	1/Y
F	X	1	1/W
M	Y	W	1

Paso 2: Calcular los Pesos Relativos. Este paso determina el peso de cada variable, de acuerdo a la *Comparación en Pares*. Es necesario calcular el *Vector Propio* de la Matriz de Comparaciones en Pares (A) para generar los pesos relativos de los criterios, o de las variables.

Dada la matriz de Comparación en Pares, que es una matriz 3x3, según el caso, el número real λ es un *Valor Propio* (también conocidos como, valor característico, autovalor o eigenvalor) de la matriz A si existe un vector x distinto de cero tal que:

$$Ax = \lambda x \quad (1)$$

Todo vector x distinto de cero que satisfaga la formula (1) es un *Vector Propio* de la matriz, asociado con el valor propio de λ . Los *Vectores Propios* también son llamados como *Vectores Característicos*, *Autovectores*, *Vectores Latentes* o *Eigenvectores*.

Se prosigue al cálculo de la Matriz de Normalización para determinar los pesos relativos de las variables, y para ello, primero se calcula la sumatoria de cada una de las columnas de la Matriz de Comparación en Pares; luego, se calculan los nuevos valores de las celdas de la matriz, los cuáles resultan de la división de cada uno de los valores ubicados en la Matriz de Comparación en Pares original, entre el valor de la sumatoria de la columna correspondiente. Con estos datos, finalmente, se calcula el valor promedio de las filas de la matriz resultante, los cuales conforman el *Vector Propio*.

El *Vector Propio* indica los porcentajes de importancia para cada variable (W_R , W_F , y W_M), dada por el evaluador.

Para los grupos de clientes con un valor similar, basado en las ponderaciones de las variables *RFM*, se puede emplear el *Algoritmo K – Means Clustering*. Los grupos resultantes representan los segmentos del mercado los cuales deben ser abordados con diferentes estrategias de marketing.

Paso 3: Evaluar la Consistencia de la Matriz de Comparación en Pares. Los evaluadores quizás hacen juicios inconsistentes cuando hacen Comparaciones en Pares. Por ejemplo, “me gusta más las manzanas que las naranjas”, “me gusta más las naranjas que las bananas”, y “me gusta más las bananas que las manzanas”.¹² El grado de consistencia de una Matriz de Comparación en Pares es calculado por el *Índice de consistencia*.

¹²Ya-Yueh Shih, Chung-Yuan Liu. A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis. Página 164. 24 de Septiembre de 2003.

La consistencia de una Matriz de Comparación en Pares expresa el correcto juicio del decisor al momento de construir la matriz, ya que en el caso de que la matriz resulte inconsistente, el decisor deberá de replantear sus juicios.

Matemáticamente, una matriz $A_{n \times n}$ es consistente si $a_{ij} \cdot a_{jk} = a_{ik}$ ($\forall i, j, k = 1, 2, 3, \dots, n$); lo cual es indispensable que la matriz sea linealmente dependiente.

En el caso general, se conoce que si la matriz $A_{n \times n}$ es consistente, entonces genera una matriz $N_{n \times n}$ la cual está normalizada de elementos W_j ($\forall j = 1, 2, 3, \dots, n$) de modo tal que las columnas son idénticas.

$$N = \begin{bmatrix} w_1 & w_1 & \dots & w_1 \\ w_2 & w_2 & \dots & w_2 \\ \vdots & \vdots & \ddots & \vdots \\ w_n & w_n & \dots & w_n \end{bmatrix}$$

Ahora, se puede determinar la matriz A en función de N, de la siguiente manera:

$$A = \begin{bmatrix} 1 & w_1/w_2 & \dots & w_1/w_n \\ w_2/w_1 & 1 & \dots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \dots & 1 \end{bmatrix}$$

Dividiendo los elementos de la columna i entre W_j .

En base a la definición A, se desglosa:

$$\begin{bmatrix} 1 & w_1/w_2 & \dots & w_1/w_n \\ w_2/w_1 & 1 & \dots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \dots & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} nw_1 \\ nw_2 \\ \vdots \\ nw_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

Es decir:

$$AW = nW$$

Donde W es el vector propio de pesos relativos W_j que se aproxima con los n elementos de las filas en la matriz normalizada N.

Haciendo $\bar{W}$ el estimado calculado, se puede notar que:

$$A \bar{W} = \lambda_{\max} \bar{W}$$

Donde $\lambda_{\max} \geq n$, es decir, mientras más próximo este $\lambda_{\max}$ a n , la matriz de comparación por pares A , será más consistente.

Como resultado, el **AHP** calcula la razón de consistencia (**RC**), que se expresa como el cociente entre el índice de consistencia de la matriz A y el índice de consistencia aleatorio (**IA**).

$$RC = CI/IA \quad (2)$$

Donde **CI** es el *Índice de Consistencia* de la matriz A , calculándose de la siguiente manera¹³:

$$CI = (\lambda_{\max} - n) / (n - 1), \quad (3)$$

donde $\lambda_{\max}$ denota el máximo *Valor Propio* de la Matriz de Comparación en Pares, y n es el grado de la matriz. Esta matriz es absolutamente consistente si $\lambda_{\max} = n$, y **CI** = 0. El procedimiento estándar permite aceptar un conjunto de juicios cuando el **CI** no es más grande que el valor de 1/10 del *Índice de Consistencia* medio de las matrices generadas aleatoriamente de orden n .¹⁴

El *Índice de Consistencia Aleatorio (IA)* es el índice de consistencia de una matriz de comparación por pares, generada de forma aleatoria. Ver **Tabla No. 6**.

Tabla No. 6, Índice de Consistencia Aleatoria

Número de Elementos Comparados	Índice Aleatorio (IA)
3	0,58
4	0,89
5	1,11
6	1,24
7	1,32
8	1,40
9	1,45
10	1,49

Fuente: Criterios de Selección de Personal mediante el uso del proceso de análisis jerárquico. MAURTUA OLLAGUEZ, Diego Edher.. Lima – Perú, 2006. P. 11.

¹³ MAURTUA OLLAGUEZ, Diego Edher. Criterios de Selección de Personal mediante el uso del proceso de análisis jerárquico. Lima – Perú: 2006. P. 10.

¹⁴ MIROSLAW, Kwiesielewicz, UDEN, Eva Van. Inconsistent and Contradictory Judgements in Pairwise Comparison Method in the AHP. 2004. P. 714.

Una vez tenidos el índice de consistencia y el índice de consistencia aleatorio, ya se puede determinar la consistencia de una matriz, mediante el cálculo de la razón de consistencia (**RC**).

Si la razón de consistencia resulta menor o igual a 0,10, entonces la matriz de comparación por pares es consistente.

Si la razón de consistencia resulta estrictamente mayor que 0,10, entonces la matriz de comparación por pares es inconsistente, lo que llama a la reflexión del decisor, para que replantee sus juicios en base a las comparaciones de los criterios o variables que había hecho.¹⁵

5.4. CLUSTERING

Clustering es una técnica estadística cuyo objetivo es el de agrupar individuos que presentan datos similares. A los datos que difieren en detalles insignificantes se les da el mismo nombre y pueden ser tratados de la misma forma, y se espera que actúen de forma similar. La palabra *Clustering* es casi sinónimo de clasificación.¹⁶ Uno de los algoritmos más empleados es el *Algoritmo K - Means*.

Algoritmo K - Means

El *Algoritmo k-means Clustering* es una función que recibe como entrada un conjunto de datos, y un número g . Con esta entrada, el algoritmo genera una salida conformada por un número g de conjuntos de datos.¹⁷ Este algoritmo tiene como objetivo maximizar la varianza entre grupos mientras minimiza la varianza dentro de los grupos.¹⁸

Supóngase una muestra de a elementos con b variables. El objetivo es dividir esta muestra en un número de grupos prefijado, g . El *Algoritmo K-Medias* requiere las cuatros etapas siguientes:

¹⁵ MAURTUA OLLAGUEZ, Diego Edher. Criterios de Selección de Personal mediante el uso del proceso de análisis jerárquico. Lima – Perú: 2006. P. 11.

¹⁶ HARTIGAN, John A. Clustering Algorithms. Department of Statics Yale University. 1975. P. 1, 6, 9, y 10.

¹⁷ WILKIN, Gregory A., HUANG, Xiuzhen. A Practical Comparison of Two k-Means Clustering Algorithms. Agosto de 2007. P. 2.

¹⁸ SHIH, Ya-Yueh, y LIU, Chung-Yuan. A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis. Septiembre 24 de 2003. P. 161.

1. Seleccionar g puntos como centros de los grupos iniciales. Esto puede hacerse:

- a) Asignando aleatoriamente los objetos a los grupos y tomando los centros de los grupos así formados;
- b) Tomando como centros los g puntos más alejados entre sí;
- c) Construyendo unos grupos iniciales con información *a priori* y calculando sus centros, o bien seleccionando los centros *a priori*.

2. Calcular las distancias euclídeas de cada elemento a los centros de los g grupos, y asignar cada elemento al grupo de cuyo centro esté más próximo. La asignación se realiza secuencialmente y al introducir un nuevo elemento en un grupo se recalculan las coordenadas del nuevo centro del grupo.

3. Definir un criterio de optimalidad y comprobar si reasignando alguno de los elementos mejora el criterio.

4. Si no es posible mejorar el criterio de optimalidad, terminar el proceso.¹⁹

Implementación del Algoritmo

El criterio de homogeneidad o de optimalidad, que se utiliza en el algoritmo de *K-Means* es minimizar la *suma de cuadrados dentro de los grupos (SCDG)* para todas las variables, dada por:

$$SCDG = \sum_{g=1}^G \sum_{j=1}^p \sum_{i=1}^{n_g} (x_{ijg} - \bar{x}_{jg})^2 \quad (4)$$

Donde x_{ijk} es el valor de la variable j en el elemento i del grupo g y $\bar{x}_{jk}$ la media de esta variable en el grupo. Este criterio es equivalente a la suma ponderada de las varianzas de las variables en los grupos, ya que puede escribirse como:

¹⁹ PEÑA, Daniel. Análisis de Datos Multivariantes. España: Mac Graw-Hill Book Company, 2002. P. 221.

$$\min SCDG = \min \sum_{g=1}^G \sum_{j=1}^p n_g s_{jg}^2$$

Donde n_k es el número de elementos del grupo g y s_{jk}^2 es la varianza de la variable j en dicho grupo.

Las varianzas de las variables en los grupos son claramente una medida de la heterogeneidad de la clasificación y al minimizarlas obtendremos grupos más homogéneos. Un criterio alternativo de homogeneidad sería minimizar las distancias al cuadrado entre los puntos y sus centros del grupo. Si se mide la distancia con la norma euclídea, este criterio se escribe:

$$\min \sum_{g=1}^G \sum_{i=1}^{n_g} (\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)' (\mathbf{x}_{ig} - \bar{\mathbf{x}}_g) = \min \sum_{g=1}^G \sum_{i=1}^{n_g} d^2(i, g)$$

Donde $d^2(i, g)$ es el cuadrado de la distancia euclídea entre el elemento i del grupo g y su media de grupo. Es fácil comprobar que ambos criterios son idénticos. Como un escalar es igual a su traza, se puede escribir este último criterio como:

$$\min \sum_{g=1}^G \sum_{i=1}^{n_g} \text{tr} [d^2(i, g)] = \min \text{tr} \left[\sum_{g=1}^G \sum_{i=1}^{n_g} (\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)(\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)' \right]$$

y llamando $\mathbf{W}$ a la matriz de suma de cuadrados dentro de los grupos,

$$\mathbf{W} = \sum_{g=1}^G \sum_{i=1}^{n_g} (\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)(\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)' \quad (5)$$

Tenemos que:

$$\text{Min tr}(\mathbf{W}) = \min SCDG$$

Y ambos criterios coinciden. Este criterio se denomina *Criterio de la Traza*, y fue propuesto por Ward (1963).

La minimización de SCDG requería calcular (3) para todas las posibles particiones, labor claramente imposible, salvo para valores de n muy pequeños. El algoritmo de *K-Medias* busca la partición óptima con la restricción de que en cada iteración sólo se permite mover un elemento de un grupo a otro. El algoritmo funciona como sigue:

- (1) Partir de una asignación inicial.
- (2) Comprobar si moviendo algún elemento se reduce la $tr(\mathbf{W})$.
- (3) Si es posible reducir $tr(\mathbf{W})$ moviendo un elemento hacerlo, recalculando las variables de los dos grupos afectados por el cambio y volver a (2). Si no es posible reducir la $tr(\mathbf{W})$, terminar.

En consecuencia, el resultado del algoritmo puede depender de la asignación inicial y del orden de los elementos. Conviene siempre repetir el algoritmo con distintos valores iniciales y permutando los elementos de la muestra. El efecto del orden suele ser pequeño, pero conviene asegurarse en cada caso.

El *Criterio de la Traza* tiene dos propiedades importantes. La primera es que no es invariante ante cambios de escala. Cuando las variables vayan en unidades distintas conviene estandarizarlas, para evitar que el resultado del algoritmo de *K-Means* dependa de cambios irrelevantes en la escala de medida. Cuando vayan en las mismas unidades suele ser mejor no estandarizar, ya que una varianza mucho mayor que el resto puede ser debida precisamente a que existen dos grupos de observaciones en esa variable, que podemos ocultar al estandarizar.

La segunda propiedad del *Criterio de la Traza* es que minimizar la distancia euclídea produce grupos aproximadamente esféricos. Por otro lado, este criterio supone variables cuantitativas y, aunque puede aplicarse si existe un pequeño número de variables binarias, cuando muchas son atributos es mejor utilizar los métodos jerárquicos.²⁰

En consecuencia, el *Algoritmo K-Means Clustering* es usado para agrupar conjuntos de clientes con un valor similar (CLV).

²⁰ PEÑA, Daniel. *Análisis de Datos Multivariantes*. España: Mac Graw-Hill Book Company, 2002. P. 222.

Como se trata de agrupar los clientes, teniendo en cuenta las variables *RFM*, se debe llevar a cabo los siguientes pasos:

- Especificar el número de grupos (*g*).
- Normalizar los valores de las variables *RFM*. Esta normalización se hace de la siguiente forma:

“La Formula de Beneficio, $X' = (X - X^S) / (X^L - X^S)$, es usada para normalizar la variable F (Frecuencia), y M (Monetario), ya que F y M influyen positivamente al *Valor del Cliente* o a la lealtad. Notablemente, X' y X representan el valor de las variables F y M, normalizado y el original, respectivamente, mientras que X^L y X^S representa el valor más grande y pequeño de las variables F y M de todos los clientes, respectivamente.”²¹

“La Formula de Costo, $X' = (X^L - X) / (X^L - X^S)$, es usada para normalizar el valor de la variable R (Reciente), ya que R influye negativamente al valor del cliente o a la lealtad. Notablemente, X' y X representan el valor normalizado y original de la variable R, respectivamente. Mientras X^L y X^S representan el más grande y el más pequeño valor de R de todos los clientes, respectivamente.”²²

- Los valores normalizados de las variables *RFM* de cada cliente se multiplican por los pesos relativos de cada variable *RFM* (W_R , W_F , y W_M), que han sido determinados por el *AHP*²³.

Cuando la base de datos es muy grande, como lo es en este caso de estudio, el cálculo de los grupos se puede llevar a cabo por medio del uso de un software llamado, *Minitab*.²⁴

La **Tabla No. 7** muestra los grupos resultantes de la aplicación de *Minitab* en la columna “Grupo de Clientes”; el número de clientes que conforman cada grupo, en la columna “Número de Clientes”, y los promedios de las variables *RFM* para cada grupo de clientes. La última fila muestra el promedio total, que es la media general de los valores de las variables *RFM* de todos los clientes. Luego se compara el promedio de los valores de las variables *RFM* de cada grupo con los promedios totales calculados en la última fila.

²¹ SHIH, Ya-Yueh, y LIU, Chung-Yuan. A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis. Septiembre 24 de 2003. P. 165.

²² Ibid, p. 165.

²³ Ibid, p. 166.

²⁴ WASS, John A. Minitab 15: Not Just for Six Sigma. Enero de 2008. P. 27.

Si el valor promedio de cualquier variable *RFM*, de un cliente excede el promedio total de esa variable, entonces una flecha hacia arriba se mostrará; de lo contrario, una flecha hacia abajo se mostrará. La última columna de la **Tabla No. 7** muestra el patrón de las variables *RFM* para cada grupo, el cual indica el comportamiento de las variables de cada grupo de clientes de acuerdo al promedio.

Tabla No. 7, Promedios de las Variables RFM para cada Grupo de Clientes

Grupo de Clientes	Número de Clientes	Reciente (R)	Frecuencia (F)	Monetario (M)	Patrón
1					
2					
3					
4					
5					
6					
...					
...					
Promedio					

Cada grupo representa la segmentación del mercado.

De acuerdo al comportamiento del patrón se le denominará un nombre a cada grupo de clientes. Estos nombres dependen de los tipos de clientes existentes.

Tipos de Clientes

Los diversos tipos de clientes de una compañía pueden clasificarse de acuerdo a su comportamiento de compra:

➤ **Clientes Potencialmente Críticos:** Son clientes que quizás alguna vez pudieron haber disfrutado de una buena relación con la empresa. Las deserciones pueden llevarse a cabo en una empresa, si ésta hace caso omiso a estos clientes. Estos clientes no han realizado transacciones recientemente, posiblemente a causa de la reubicación de sus negocios o del sufrimiento de quiebra.

➤ **Clientes sin Valor:** Son aquellos clientes que solo se preocupan por los precios. Estos clientes generalmente solo hacen compras durante las rebajas. Las empresas pueden reducir los precios para atraer a estos clientes, pero podrían sufrir márgenes reducidos como resultado. Ya que pocas empresas obtienen beneficios de estos clientes, ellos son clasificados como clientes sin valor. Tales clientes fácilmente se van dónde otras empresas que ofrezcan un mejor precio. Una compañía puede sólo retener tales clientes ofreciendo precios bajos, pero los beneficios que se pueden obtener de estos clientes son muy limitados.

➤ **Clientes de Oro:** Estos clientes son los más leales quienes compran frecuentemente y contribuyen una gran cantidad monetaria. Los clientes de Oro son los más grandes contribuyentes para la rentabilidad de la empresa. Para este tipo de clientes se diseñan estrategias de marketing con el fin de mantener sus actividades por medio de la prestación de servicios especiales, y así poder asegurarlos.

➤ **Clientes Potenciales:** Estos clientes tiene el potencial para convertirse en clientes de oro. Son clientes que están empezando a desarrollar una relación más estrecha con la empresa, y que también están aumentando gradualmente su capacidad de compra. Los clientes potenciales deben ser conscientes de que si aumentan la frecuencia de compra, la empresa les dará un mejor servicio, similar al que se les brinda a los clientes de oro. Las empresas tratarán de establecer una relación a largo plazo con los clientes.²⁵

Para cada tipo de cliente, la empresa propone estrategias de marketing ya sea para la retención de los clientes, o la captura de nuevos clientes, aunque mantener a un cliente ya existente es más barato que el desarrollo de uno nuevo. La idea es centrarse en los clientes más leales.

Análisis de la Varianza

La comparación simultánea de varias medias poblacionales se conoce como **Análisis de la Varianza (ANOVA)**. El ANOVA es uno de los usos de la distribución *F* (distribución de probabilidad que se usa como estadística de prueba en varias situaciones. Se emplea para probar si dos muestras provienen de poblaciones que poseen varianzas iguales, y también se aplica cuando se trata de comparar simultáneamente varias medias poblacionales), en la que se comparan tres o más medias muestrales para determinar si provienen de poblaciones iguales. Para utilizar esta técnica, se supone lo siguiente:

²⁵ SHIH, Ya-Yueh, y LIU, Chung-Yuan. A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis. Septiembre 24 de 2003. P. 166.

- Las poblaciones tiene una distribución normal.
- Las poblaciones tienen desviaciones estándar (σ) iguales.
- Las muestras se seleccionan de manera independiente.

Cuando se cumplen estas condiciones, se utiliza F como estadístico de prueba.²⁶

En la prueba ANOVA, la estrategia fundamental consiste en calcular la varianza de la población (la desviación estándar elevada al cuadrado) de dos formas y luego encontrar la relación de esas dos estimaciones. Si la relación está próxima a uno, entonces ambas estimaciones, de forma lógica, son iguales, y se concluye que las medias de la población son iguales. Si la relación es muy distinta de uno, entonces se concluye que las medias de la población son las mismas. La distribución F indica que la relación es mucho mayor que uno para haber ocurrido por casualidad.

Para llevar a cabo el ANOVA, se deben llevar a cabo los siguientes pasos:

- *Establecer la hipótesis nula y la hipótesis alterna.* La hipótesis nula es que las puntuaciones medias son iguales para los n rangos.

$$H_0 = \mu_1 = \mu_2 = \dots = \mu_n$$

La hipótesis alternativa es que las puntuaciones medias no son iguales para los rangos.

H_1 : Las puntuaciones medias no son iguales

Si la hipótesis nula no se rechaza, se concluye que no hay diferencia entre las medias. Si se rechaza H_0 , se concluye que existe una diferencia en al menos un par de las medias, pero en ese punto no se sabe cuáles de los pares diferirán, ni por cuánto.

- *Seleccionar el Nivel de Significancia.* Se selecciona el nivel de significancia.
- *Formular la Regla de Decisión.*
- *Seleccionar la Muestra, Realizar los cálculos y Tomar una decisión.* Es conveniente resumir los cálculos de la estadística F en un cuadro **ANOVA**.

²⁶ MASON, Robert D., y LIND, Douglas A. Estadística para Administración y Economía. México: Alfaomega Grupo Editor, 1995. P. 340, 344.

Con la ayuda del software llamado *Minitab* se puede realizar los anteriores cálculos. Este software arroja un valor p , y se ubica bajo en el encabezado “P”, y éste es la probabilidad de hallar un valor F a la derecha del valor de F arrojado por Minitab con determinados grados de libertad en el numerador y determinados grados en el denominador, dado que H_0 sea verdadera. Así la probabilidad de cometer un error al rechazar un H_0 verdadero es del valor de $P * 100$ por ciento.

5.5 CLASIFICACIÓN DEL VALOR DEL CLIENTE (CLV)

El *Valor del Cliente* (CLV) es definido como el valor actual neto de los beneficios futuros derivados de los clientes, e identifica la contribución directa que un cliente hace durante un periodo a la rentabilidad de la empresa. El CLV puede conducir estrategias que incrementan la equidad de los clientes, y a guiar los esfuerzos para optimizar la experiencia individual del cliente.²⁷

Si la empresa cuenta con una base de datos de los clientes, ésta puede calcular el CLV para cada registro de compras, y así poder predecir el beneficio para el cliente individual, distribuir las promociones, y asignar los recursos para retener a los clientes.²⁸

Generalmente, el *Valor del Cliente* (CLV) es evaluado en términos de las variables RFM (Reciente, Frecuencia, y Monetario).

La clasificación del CLV es generada para ayudar a los practicantes del mercado a desarrollar estrategias más efectivas para retener clientes y por lo tanto, claramente identificar y comparar los segmentos del mercado. Esta clasificación de los grupos procede de la siguiente forma:

Los valores de las variables RFM de cada cliente son normalizadas, como se describió anteriormente. La **Tabla No.8**, muestra el promedio de los valores normalizados de las variables RFM, denotado como C_R^j , C_F^j , y C_M^j , respectivamente, para $j = 1$ hasta k (el número de grupos). C_R^j , C_F^j , y C_M^j son calculados promediando los valores de las variables RFM normalizados de los clientes en el grupo j .

²⁷ GREENBERG, Paul. Managing Social Customers for Profit. Agosto de 2009. P. 1.

²⁸ SHIH, Ya-Yueh Shih, y LIU, Chung-Yuan. A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis. Septiembre 24 de 2003. P. 160.

La Evaluación Integrada, C_I^j es calculada como la suma ponderada de C_R^j , C_F^j , y C_M^j ; esto es de la siguiente forma: $C_I^j = W_R * C_R^j + W_F * C_F^j + W_M * C_M^j$, donde W_R , W_F Y W_M son la importancia relativa de las variables *RFM*, como fue determinado por el *AHP*.

En cuanto se calcula la *Evaluación Integrada* se puede hacer la clasificación del *Valor del Cliente* (CLV) de los grupos de clientes. Entre más grande es el valor de la evaluación integrada, éste corresponderá a la más alta clasificación, es decir el número de 1, y así mismo para el más bajo valor de la evaluación integrada, éste corresponderá a la más baja clasificación, en este caso el número ocho. La más alta clasificación indica que los clientes que están en este grupo son los más leales.

En consecuencia, los tomadores de decisión pueden realizar diferentes estrategias para retener a los clientes de acuerdo a la clasificación de *CLV* o la lealtad.

Tabla No. 8, Clasificación del Valor del Cliente (CLV)

Segmentación de los Clientes	Reciente (R)	Frecuencia (F)	Monetario (M)	Evaluación Integrada	Clasificación de la Lealtad
1					
2					
3					
4					
5					
6					
7					
8					
Promedio Total					

6. CASO DE ESTUDIO

Un conjunto de datos fue recolectado a partir de una base de datos de transacciones realizadas por clientes durante dos años (desde el primero de noviembre de 2007 hasta el 31 de octubre de 2009), aproximadamente 3500 registros, con el fin de ilustrar la metodología que se va a exponer en este documento. El caso de estudio concierne a una empresa comercializadora de herramientas para la construcción, especialmente de discos, copas y brocas diamantadas de alta calidad, en la ciudad de Cali. Esta compañía en la actualidad produce alrededor de 246 productos, los cuales son comercializados dentro de la misma ciudad.

El propósito del documento es que la empresa pueda centrarse en ciertos grupos de clientes, en los que son más rentables. De esta forma la empresa podrá desarrollar estrategias de marketing en torno a éstos, con el fin de satisfacer sus necesidades y generar un incremento en su utilidad.

Código Cliente	Cantidad Monetaria (Pesos Colombianos)	Fecha
000009	51395,0	14/11/2007
000009	164635,0	14/11/2007
000009	297635,0	21/11/2007
000009	69825,0	14/11/2007
000009	65930,0	14/11/2007
000009	110722,0	03/12/2007
000009	144637,0	26/11/2007
000009	79990,0	14/11/2007
000009	79325,0	26/11/2007
...	...	...
...	...	...
...	...	...
000039	3611882,0	04/09/2009
000043	13417521,0	20/02/2009
000370	15519759,0	03/06/2009

Cada fila representa un registro de las transacciones que realizan los clientes.

Fecha: Es la fecha en que el cliente realizó la compra.

Código del Cliente: Código mediante el cual la empresa identifica a cada uno de sus clientes.

Cantidad Monetaria: Es el valor de la compra que realiza un cliente.

Figura No. 1, Registros de las Transacciones de los Clientes de la Base de Datos de la Empresa

Los datos recolectados son procesados para extraer las transacciones de los clientes, con el fin de extraer así mismo los valores de las variables **RFM** para cada cliente.

Los valores de las variables **RFM** de los 158 clientes que tiene la empresa durante los dos años de estudio como se podrá ver a continuación, fueron extraídos para poder llevar a cabo la metodología, tal como se explicará en el próximo capítulo. Las 3500 filas representan 3500 registros de transacciones en la base de datos. Cada fila registra un producto que un cliente compró. (**Ver Figura No.1**).

El interés de la empresa frente a la idea de poder calcular el *Valor del Cliente*, y así poder identificar los clientes más leales, se debe a que han percibido que el mercado de las herramientas para la construcción ha ido evolucionando, y las exigencias de los clientes son cada vez más grandes, dado que estamos en una época donde nuevos proyectos se están generando en la ciudad. La compañía vio a este estudio como una oportunidad de mejora, y consideró que esta clasificación de los clientes, puede ser la base para el desarrollo de nuevas estrategias de marketing direccionadas hacia éstos.

Actualmente, la empresa no cuenta con una técnica específica para el cálculo del *Valor del Cliente*, pero sí considera la importancia de la relación que se debe tener con sus clientes, y para ello, la compañía se acerca a éstos por medio de demostraciones amistosas y de amabilidad, que le den a entender al cliente que es muy importante para la empresa.

La lealtad de un cliente es percibida por la empresa como aquel cliente que siempre está presente a través del tiempo, sin importar nada más, es por esta razón, que este estudio le permitirá a la compañía saber con certeza la lealtad de éstos teniendo en cuenta diversos aspectos que caracterizan su comportamiento, y así poder aumentar las ganancias.

7. EXPERIENCIA COMPUTACIONAL

La metodología propuesta que se aplicó al caso de estudio se representa gráficamente de la siguiente forma:

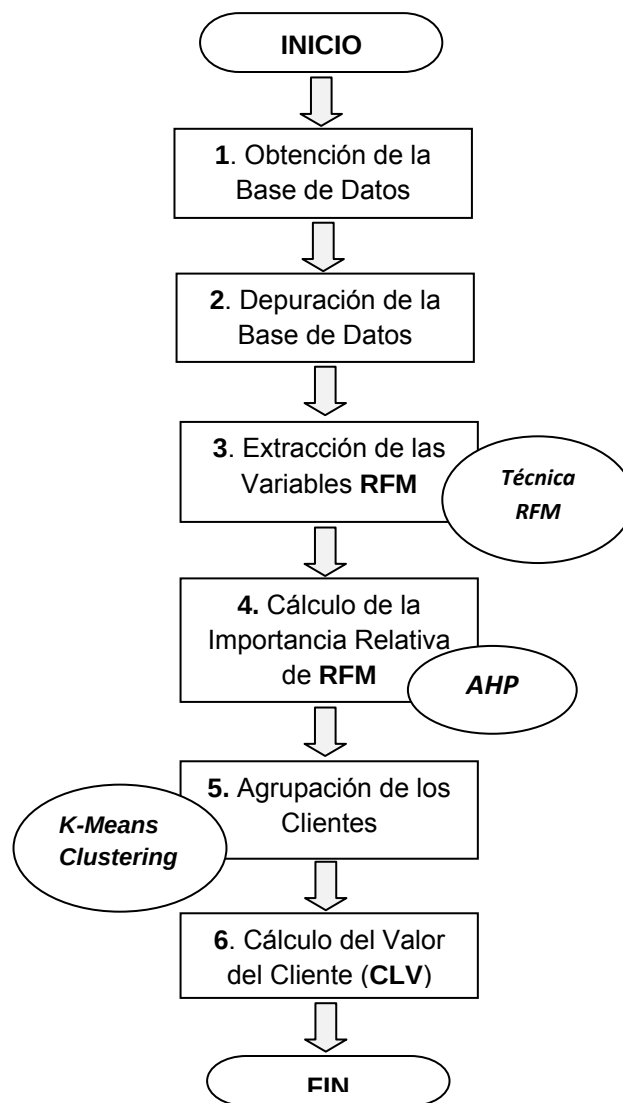


Figura No. 2, Metodología Propuesta

Los pasos definidos anteriormente en la figura se detallan a continuación.

7.1. OBTENCIÓN DE LA BASE DE DATOS

La base de datos obtenida a partir del software contable llamado *Unopymes*, que es el que usa la compañía, fue examinada con detalle, con el fin de eliminar los registros incoherentes o los que no son útiles para este caso de estudio, tales como los productos que fueron devueltos por los clientes con sus respectivas equivalencias monetarias, y las columnas con datos que no son necesarios para llevar a cabo ningún cálculo propuesto en la metodología. Los datos fueron exportados a la hoja de cálculo directamente desde el software.

7.2. DEPURACIÓN DE LA BASE DE DATOS

Después de tener la base de datos importada en una hoja de cálculo de EXCEL, se procedió a depurar los determinados datos innecesarios que se muestran a continuación:

- a)** Nombres de los clientes por cuestión de privacidad de información de la compañía.
- b)** Códigos y nombres de los productos que se venden por cada transacción, ya que en la metodología usada no se hace seguimiento a productos.
- c)** Cantidades de los productos comprados, ya que lo importante es el valor de la compra.
- d)** Números de facturas.
- e)** Las cantidades de productos que los clientes han devuelto durante estos dos años de estudio, ya que éstas son compras que los clientes alguna vez realizaron pero que luego decidieron devolver a la compañía por algún motivo específico. Es posible que las devoluciones hayan sido completas o parciales. La base de datos

contiene todos los registros que los clientes han realizado, tanto de sus compras como de las devoluciones. Por lo tanto fue necesario analizar cada registro que indicaba una devolución (cantidad monetaria negativa), identificándolos junto con los registros que indicaban la compra de estos productos que fueron devueltos posteriormente o en el mismo día. Se encontraron devoluciones que se realizaron hasta un mes después o más, de la fecha en que el cliente compró. Una vez se identificó el registro que indicaba devolución, se buscó el respectivo registro de su compra. Cuando el valor monetario de la compra era igual a la de la devolución, se eliminaron los dos registros, ya que para la compañía, este cliente en últimas no compro nada; y si el valor monetario de la devolución era menor al valor de la compra, se calculó lo que verdaderamente el cliente compró para la compañía, que fue el valor de la compra menos el valor de la devolución, y este dio un nuevo valor que fue puesto en un nuevo registro, cuya fecha de transacción es la fecha en la que se realizó la compra inicialmente. Esta depuración dio origen a una nueva base de datos más manejable.

Por lo tanto, se tiene una base de datos con información de los códigos de los clientes, la cantidad monetaria de cada transacción realizada por éstos, y las fechas en las que realizaron cada transacción.

Los datos originalmente están dados por años, cada hoja de cálculo es un año, en donde para el año 2007 están las ventas efectuadas en el mes de noviembre y en el mes de diciembre; para el año 2008 están las ventas efectuadas durante todos los meses de este año, y para el año 2009 están las ventas efectuadas durante los meses comprendidos entre el mes de enero y octubre. De esta forma se tiene un registro de 24 meses de transacciones.

Para un mejor manejo de los datos, éstos fueron reordenados en una nueva hoja de cálculo, en donde están reunidos los dos años de transacciones de compras de los clientes, en una sola tabla. Mediante la ayuda de los filtros de Excel, se acomodó la información en orden ascendente para que de esta forma se tuvieran las transacciones agrupadas por clientes, desde el Cliente 000001 hasta el Cliente 000642. Con esta información se prosiguió al cálculo de los valores de las variables **RFM**.

7.3. EXTRACCIÓN DE LAS VARIABLES RFM

El cálculo de los valores de las variables **RFM** se hace mediante los siguientes cálculos:

1. La variable **R** (Reciente), que es la fecha de la última compra expresada en días, se calcula así:

Teniendo en cuenta la columna de las fechas de las transacciones que corresponde a la Columna B de la **Figura No. 3, (a)**, se calculó la variable **R**. Se tomó como punto de referencia, una fecha base, el 31 de Octubre del año 2009, siendo ésta la fecha límite del periodo de tiempo de estudio. Como la variable **R** se debe calcular por clientes y no por transacciones (filas), se prosiguió a calcular la diferencia entre la fecha base y cada una de las fechas ubicada en la Columna B. Estos valores fueron ubicados en una nueva columna, en la Columna C, como se puede ver en la **Figura No. 3, (a)**. Luego para obtener los valores de la variable, se calcularon los valores mínimos de las diferencias obtenidas en la Columna C, por cliente. En la **Figura No.3, (b)** se muestra el valor de R para cada cliente.

2. La variable **F** (Frecuencia), que es la frecuencia de compra, se calcula como se muestra continuación:

Para el cálculo de la frecuencia de compra de un determinado cliente, se determinó las veces en que éste realizó una compra durante un periodo de tiempo de dos años. Teniendo en cuenta los datos de las transacciones de los dos años de estudio. Como la frecuencia se puede calcular a partir de las fechas, solo se extrajo la columna del Código de los Clientes, la Columna A, y la columna de la Fecha de Compra, la Columna B (ver **Figura No.4, (a)**). En la **Figura No. 4, (b)**, se muestran los valores de la variable **F** para cada cliente, y la forma en qué se calculó.

3. La variable **M** (Monetario), que es el monto o la cantidad monetaria de la compra, se calcula como se muestra a continuación.

La variable **M** se calcula mediante la sumatoria de la cantidad monetaria aportada por cada uno de los clientes durante los dos años de estudio. Esta variable al igual que las dos anteriores (**R** y **F**) se calculó por cliente. De la base de datos se extrajo, por cliente, la sumatoria de la cantidad monetaria (ver **Figura No.5**), la cual se convertiría en el valor de **M** para cada cliente.

	A	B	C
1	C. Cliente	Fecha	
2	000001	29/09/2008	397
3		24/11/2008	341
4		24/05/2008	525
5		06/09/2008	420
6		18/09/2008	408
7		02/02/2009	271
8		30/04/2009	184
9		06/10/2009	25
10		11/06/2009	142
11		29/05/2009	155
12		13/04/2009	201
13		20/09/2009	41
14	...	...	...
15	...	...	...
16	000642	05/05/2008	544
17		17/09/2008	409
18		27/06/2008	491
19		12/05/2008	537
20		07/04/2008	572
21		07/09/2008	419
22		28/08/2008	429
23		07/09/2009	54
24		07/08/2009	85
25		07/10/2009	24
26		07/10/2009	24
27		18/09/2009	43

= (31/10/2009) - B6.
Este valor es
expresado en días.

(a) **Primer Paso:** Cálculo de los días comprendidos entre la fecha base (31/10/2009), y la fecha de cada transacción. En la columna C se presentan todos estos valores.

Figura No. 3, Cálculo del Valor de la Variable R (Reciente)

	F	G
1	Código Cliente	R (Reciente)
2	000001	25
3	000002	26
4	000004	378
5	000008	19
6	000009	3
7	000012	222
8	...	...
9	...	...
10	000639	143
11	000640	26
12	000642	24

= MIN (C2:C13) → Valor Mínimo de los valores de la Columna C de la Figura No. 3, (a).
 = MIN (C16:C27) → Valor Mínimo de los valores de la Columna C de la Figura No. 3, (a).

(b) Segundo Paso: Cálculo de la variable **R** para cada cliente, a partir del cálculo del mínimo de los días ubicados en la columna C que están ordenados por clientes.

Figura No.3, Cálculo del Valor de la Variable R (Reciente)

	A	B
1	Código Cliente	Fecha de Compra
2	000009	14/11/2007
3	000009	14/11/2007
4	000009	21/11/2007
5	000009	14/11/2007
6	000009	14/11/2007
7	000009	03/12/2007
8	000009	26/11/2007
9	000009	14/11/2007
10	...	...
11	...	...
12	000063	11/10/2009
13	000039	04/09/2009
14	000043	20/02/2009
15	000370	03/06/2009

(a) Datos extraídos de la Base de Datos para el cálculo de la variable **F**. En esta tabla están todas las transacciones realizadas durante los dos años por los clientes. En la Columna A están los códigos de los clientes (estos códigos no están en clasificados por cliente), y en la Columna B están todas las fechas por transacción realizadas por un determinado cliente.

Figura No. 4, Cálculo del Valor de la Variable F (Frecuencia)

	E	F
1	Código Cliente	F (Frecuencia)
2	000001	12
3	000002	8
4	000004	11
5	000008	8
6	000009	202
7	000012	7
8	000013	4
9	000014	8
10	000016	20
11	...	...
12	...	...
13	000628	12
14	000629	6
15	000632	6
16	000637	6
17	000638	9
18	000639	7
19	000640	9
20	000642	12

Celdas de la **Figura No.4, (a)**

= CONTAR.SI (A2:A15; E9)

De esta forma se obtuvo la cantidad de veces en qué el Cliente X compró durante los dos años, es decir la F.

(b) Cálculo de la variable F por medio de la formula **CONTAR.SI** (rango; criterio), que cuenta las celdas en el rango que coinciden con la condición dada.

Figura No.4, Cálculo del Valor de la Variable F (Frecuencia)

De esta forma quedaron calculados los valores de las variables **RFM** (ver **Anexo No.2**). El cálculo de estas variables se le conoce como el **Método RFM**, el cual ayudará para el cálculo del **Valor del Cliente (CLV)** como se mostrará más adelante.

	A	B
	Código Cliente	Cantidad Monetaria (Pesos Colombianos)
1		
2	000001	97500
3		65000
4		3123292
5		8485564
6		6578685
7		402500
8		2663515
9		11854157
10		10149578
11		1967176
12		22509329
13		9652958
14	...	...
15	...	...
16	000642	12984643
17		7230592
18		3242812
19		11941300
20		11259576
21		5782260
22		9094369
23		38400
24		38400
25		56000
26		49500
27		5578525

(a) Base de datos por clientes, la cual se utilizó para el cálculo de la variable *M*.

Figura No. 5, Cálculo del Valor de la Variable *M* (Monetario)

	E	F
1	Código Cliente	M (Monetario)*
2	000001	77549254
3	000002	53910982
4	000004	12490939
5	000008	64935306
6	000009	43325347
7	000012	32276505
8	000013	23297511
9	000014	66615199
10	000016	91216061
11	000018	17000003
12	000023	3492000
13	...	...
14	...	...
15	000632	59774705
16	000637	16109914
17	000638	86037489
18	000639	65739523
19	000640	48425485
20	000642	67296377
21	*Cantidad dada en pesos colombianos.	

Celdas de la **Figura No.5, (a).**



= SUMA (B2:B13)

De esta forma se obtuvo la cantidad monetaria que cada cliente aportó a la compañía.

(b) Cálculo del valor de **M** por medio de la formula SUMA (número1; número2;...), que suma todos los números en un rango de celdas.

Figura No.5, Cálculo del Valor de la Variable M (Monetario)

7.4. CÁLCULO DE LA IMPORTANCIA RELATIVA DE LAS VARIABLES RFM

Una vez fueron calculados los valores de las variables **RFM**, se calculó la importancia relativa de estas variables, las cuales fueron denotadas de la siguiente forma: W_R , W_F , y W_M , donde:

W_R : Importancia relativa de la variable **R**.

W_F : Importancia relativa de la variable **F**.

W_M : Importancia relativa de la variable **M**.

Para ello, se empleó el **AHP** (Proceso Jerárquico Analítico), el cual le dará una ponderación a cada una de las variables **RFM**.

Para llevar a cabo el **AHP**, en primera instancia se realizó una encuesta con el fin de obtener los juicios intuitivos para hacer las Comparaciones en Pares de los pesos relativos o importancias relativas de las variables **RFM**. Esta encuesta no fue realizada de manera personalizada, sino que se realizó de manera grupal en una reunión en forma de mesa redonda, en la que, en un tablero se expuso el modelo de entrevista (ver **Anexo No.3**), y después de darles a conocer las preguntas se llegó a un consenso para dar un solo juicio por todos. La encuesta se les realizó a las siguientes personas:

- Gerente de Ventas de la Compañía.
- Dos vendedores de los productos que comercializa la compañía.
- Director de Cuenta de Publicidad (*Outsourcing*).

La entrevista fue diseñada teniendo en cuenta la Matriz de Comparación en Pares (Ver **Tabla No. 9**). La matriz se diseñó con el fin de comparar cada variable con las demás, y darle una puntuación a esta comparación de acuerdo a la **Tabla No.2**.

Tabla No. 9, Matriz de Comparación en Pares

	R (Reciente)	F (Frecuencia)	M (Monetario)
R (Reciente)			
F (Frecuente)			
M (Monetario)			

En primer lugar, se inicializó la matriz, asignando el valor de uno cada una de las celdas que componen la diagonal principal.

Se obtuvieron los siguientes resultados a partir de la encuesta:

- Para el grupo entrevistado, ante la primera pregunta, éstos contestaron que era más importante que un cliente les compre frecuentemente, a que éste le haya realizado una compra últimamente. Y le dieron una puntuación del valor de **cinco**, es decir que es de importancia fuerte que un cliente sea frecuente a que éste haya comprado recientemente. Esta comparación es representada en la **Figura No. 6**.

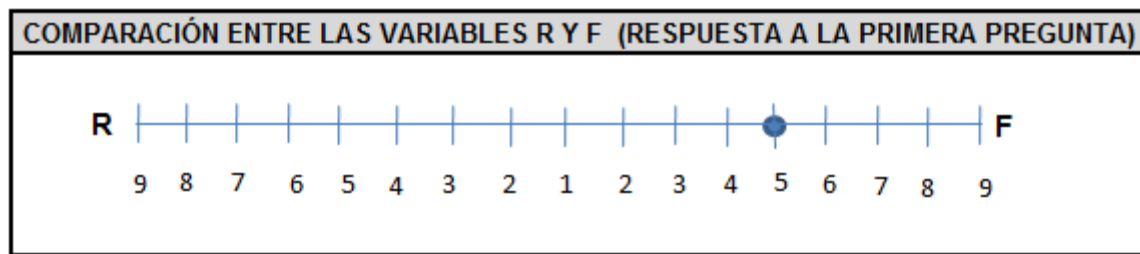


Figura No. 6, Comparación entre las Variables R y F

Con este valor de **cinco**, se llenó la Celda B24 (ver **Tabla No.11**).

➤ Ante la segunda pregunta, el grupo contestó que es más importante que un cliente haya realizado una compra últimamente, a que éste aporte una gran cantidad monetaria por sus compras, independientemente si han sido frecuentes o recientes, a la compañía. Y le dieron una puntuación del valor de **tres**, es decir, que es de débil importancia que un cliente haya efectuado una compra últimamente, a que éste haya aportado una gran cantidad de dinero por sus compras, independientemente si la realizaron hace poco o mucho tiempo, ni tampoco cuántas compras ha realizado. Esta comparación es representada en la **Figura No. 7**.

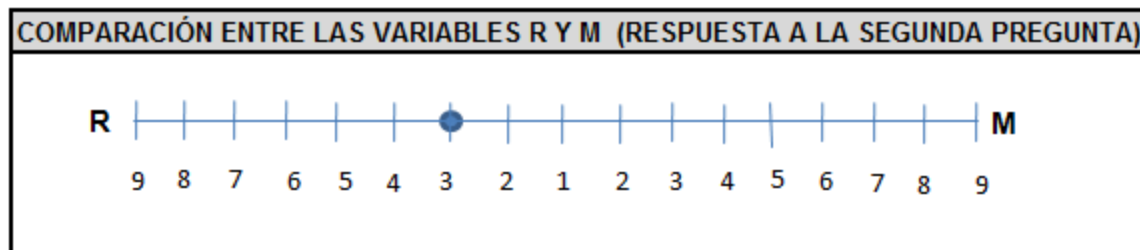


Figura No. 7, Comparación entre las Variables R y M

Con el valor de **1/3** se llenó la Celda B25 (ver **Tabla No. 11**).

➤ Para la tercera pregunta, la respuesta fue que es más importante que un cliente haya comprado frecuentemente en un determinado periodo de tiempo, a que éste haya aportado una gran cantidad monetaria por sus compras independientemente si han sido frecuentes o recientes, a la compañía. Les dieron una puntuación de **siete**, es decir que se trata de una importancia relevante de una variable sobre otra. Esta comparación es representada en la **Figura No. 8**.

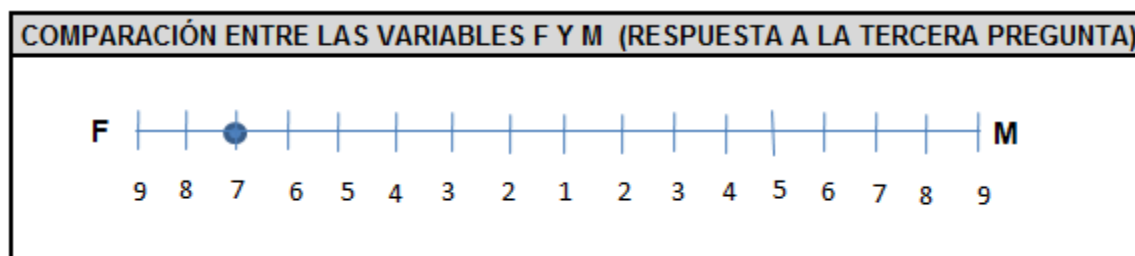


Figura No. 8, Comparación entre las Variables F y M

Con el valor de **1/7** se llenó la Celda C25 (ver **Tabla No. 11**).

En la **Tabla No. 10**, se muestran resumidas las puntuaciones dadas por los evaluadores, para cada pregunta.

Tabla No. 10, Resumen del Resultado de la Encuesta

Pregunta No.	Puntuación
1	5
2	1/3
3	1/7

Con el resultado de la encuesta se prosiguió a llenar las celdas ubicadas por debajo de la diagonal principal de la Matriz de Comparación en Pares (Ver **Tabla No. 11**), de la siguiente forma:

Tabla No. 11, Matriz de Comparación Pareada

	A	B	C	D
22		R (Reciente)	F (Frecuencia)	M (Monetario)
23	R (Reciente)	1		
24	F (Frecuente)	5	1	
25	M (Monetario)	1/3	1/7	1

La puntuación para la primera pregunta fue ubicada en la celda B24, la de la segunda pregunta se puso en la celda B25, y por último, la de la tercera pregunta, en C25.

Para llenar las demás celdas se hizo lo siguiente:

$$C23 = 1/B24 = 1/5$$

$$D23 = 1/B25 = 1/(1/3) = 3$$

$$D24 = 1/C25 = 1/(1/7) = 7$$

Con los anteriores datos se llenó por completo la Matriz de Comparación en Pares (ver **Tabla No.12**).

Tabla No. 12, Matriz de Comparación en Pares Completa

	A	B	C	D
22		R (Reciente)	F (Frecuencia)	M (Monetario)
23	R (Reciente)	1	1/5	3
24	F (Frecuente)	5	1	7
25	M (Monetario)	1/3	1/7	1

El paso siguiente fue realizar una Matriz de Normalización con el fin de obtener las importancias o pesos relativos de cada variable RFM. Para ello, primero se calculó la sumatoria de las celdas de las Columnas B, C y D, ubicándolas en una nueva fila (ver **Tabla No.13**).

Tabla No. 13, Sumatoria de las Columnas B, C y D

	A	B	C	D
35		R (Reciente)	F (Frecuencia)	M (Monetario)
36	R (Reciente)	1	1/5	3
37	F (Frecuente)	5	1	7
38	M (Monetario)	1/3	1/7	1
39	SUMATORIA	6,33	1,34	11,00

La Fila 39 de la **Tabla No.13** muestra las sumatorias de las Columnas B, C y D.

Luego se divide cada valor de las celdas de la matriz de la **Tabla No. 13**, entre la sumatoria de la columna respectiva. (Ver **Tabla No. 14**).

Tabla No. 14, Matriz de Normalización

	A	B	C	D
42		R (Reciente)	F (Frecuencia)	M (Monetario)
43	R (Reciente)	0,16	0,15	0,27
44	F (Frecuente)	0,79	0,74	0,64
45	M (Monetario)	0,05	0,11	0,09

Una vez fueron calculados los anteriores valores, se calcula el promedio de las filas ubicándolos en una nueva columna (ver **Tabla No. 15**).

Tabla No. 15, Matriz Normalizada Completa

	A	B	C	D	E
48		R (Reciente)	F (Frecuencia)	M (Monetario)	Promedio
49	R (Reciente)	0,16	0,15	0,27	0,193
50	F (Frecuente)	0,79	0,74	0,64	0,724
51	M (Monetario)	0,05	0,11	0,09	0,083

El promedio de cada fila de la **Tabla No. 15**, se ubicó en la Columna E.

Esta nueva columna es el *Vector Propio*, el cual indica las ponderaciones o importancia relativa, W_R , W_F , y W_M , de las variables R, F, y M, respectivamente.

Antes de continuar, es necesario evaluar la consistencia de la matriz dado que los evaluadores, en este caso el grupo entrevistado, pudieron quizás hacer juicios inconsistentes al hacer las comparaciones. Para ello, se calculó la *Razón de Consistencia (CR)*.

Dado que la *Razón de Consistencia* es CI/IA , se calculó en primera instancia, el *Índice de Consistencia (CI)* teniendo en cuenta la formula (3). Para este caso $n = 3$, que es el tamaño de la matriz. Ver la **Tabla No. 16**, para el cálculo del CI.

Tabla No. 16, Cálculo del Índice de Consistencia (CI)

	A	B
56	$\lambda_{max} =$	3,111
57	$n =$	3
58	$CI ((\lambda_{max} - n) / (n - 1)) =$	0,053

En la **Tabla No. 16**, el λ_{max} se calculó así, teniendo en cuenta los valores de las

tablas (**Tabla No. 13** y **Tabla No. 15**):

$$\lambda_{\max} = ((E49*B39) + (E50*C39) + (E51*D39)) = \mathbf{3,111}$$

siendo E49, E50, y E51, celdas de la **Tabla No. 15**; y B39, C39, y D39, celdas de **Tabla No. 13**.

Y el *Índice de Consistencia (CI)*, se calculó así, según la **Tabla No. 16**:

$$CI = ((B56-B57) / (B57-1)) = \mathbf{0,053}$$

Ya que el valor del *índice de Consistencia Aleatorio (IA)*, según la **Tabla No. 6**, para una matriz 3x3, es de 0,58, la *Razón de Consistencia* se calculó de la siguiente forma:

$$CR = CI/IA = (0,053) / (0,58) = \mathbf{0,091}$$

Como el **CR** tuvo un valor de **0,091**, y este es un valor menor que **0,10**, entonces los juicios realizados por el grupo de personas entrevistadas fueron coherentes, es decir los juicios fueron consistentes.

Sabiendo que los juicios son consistentes, se prosiguió a establecer las importancias relativas para cada variable **RFM**.

En la **Tabla No. 17**, se muestran los valores del *Vector Propio* expresados en porcentajes, los cuales corresponden a las importancias relativas de las variables **RFM**.

Tabla No. 17, Importancia Relativa de las Variables RFM

	A	B	C
66	Variable	Importancia Relativa	Valor de la Importancia Relativa
67	R (Reciente)	W_R	19,32%
68	F (Frecuencia)	W_F	72,35%
69	M (Monetario)	W_M	8,33%

De acuerdo a la **Técnica AHP**, los pesos relativos de las variables **RFM** son **19,32%**, **72,35%**, y **8,33%** respectivamente. Esto quiere decir que la variable **F** (Frecuencia) es la más importante, seguida por la variable **R** (Reciente), y por último, por la variable **M** (Monetario).

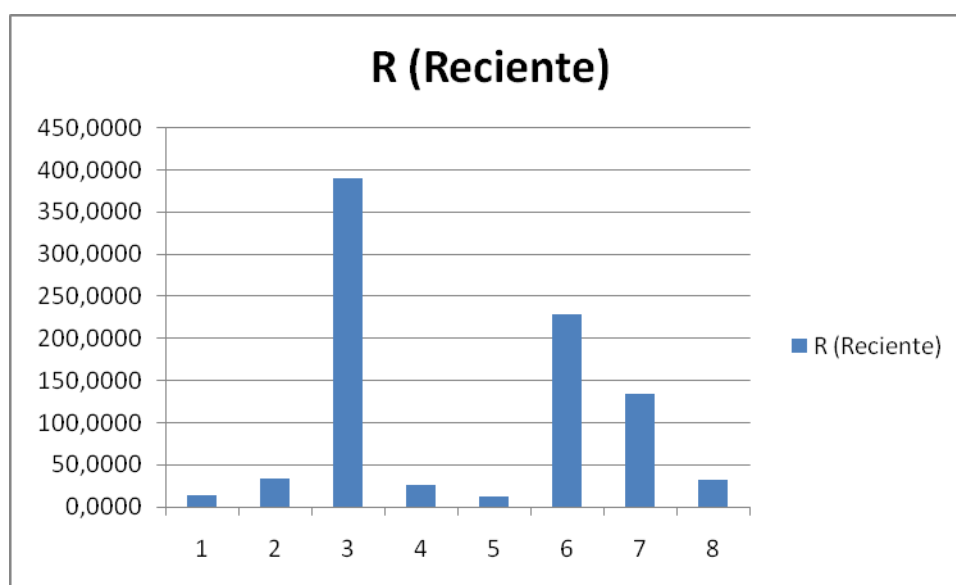
Estas ponderaciones serán aplicadas sobre las variables normalizadas ($R_N F_N M_N$),

como paso previo a la conformación de grupos.

7.5. ANÁLISIS DE LAS VARIABLES RFM

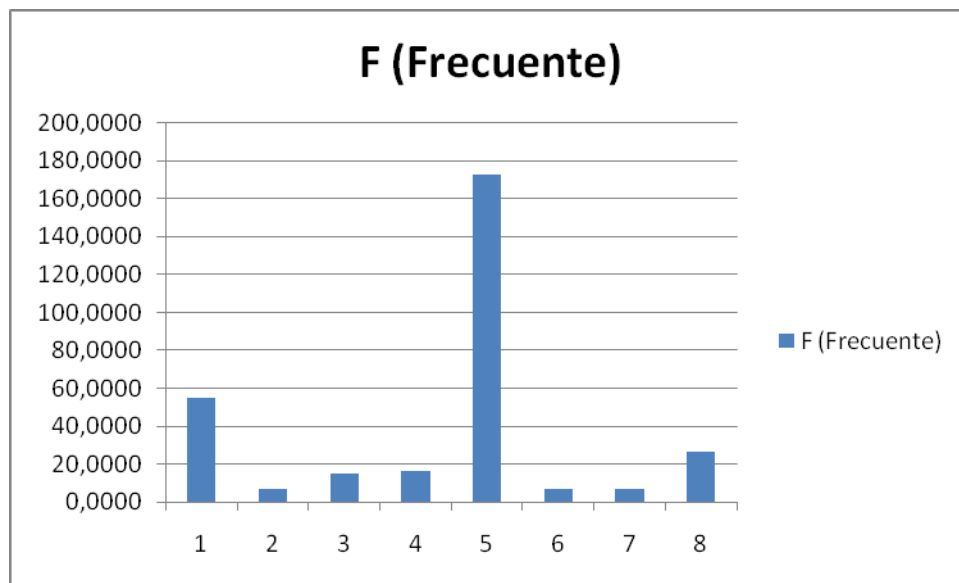
Con el fin de analizar las variables *RFM* puras generadas de la base de datos, se realizó un *Análisis Gráfico*, en dónde, por medio de los diagramas de columna, se graficaron las medias de cada variable por grupo. (Ver **Gráfica No.1**, **Gráfica No. 2**, y **Gráfica No.3**).

Gráfica No. 1, Diferencia entre las Medias de las Variables R de los Ocho Grupos



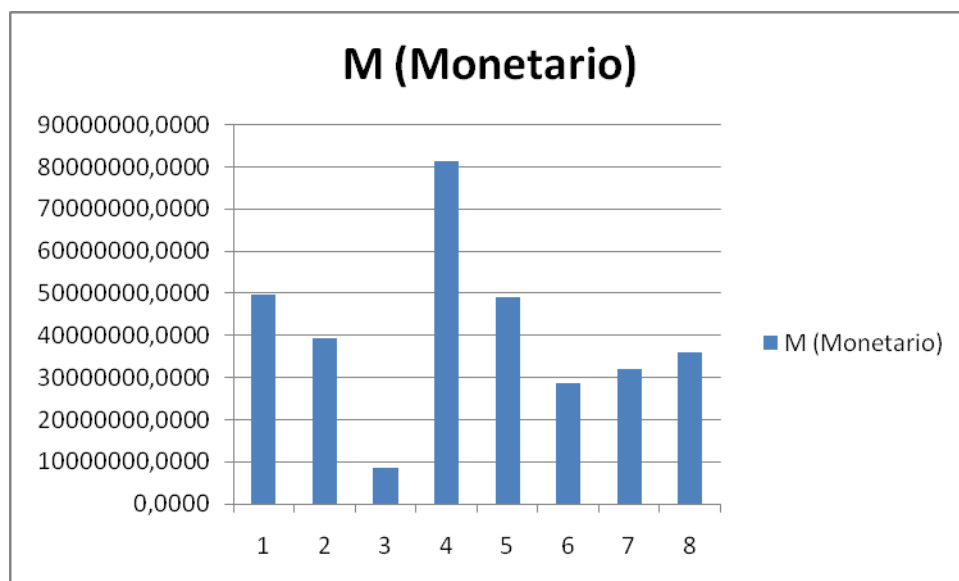
En la **Gráfica No. 1**, se puede deducir que en promedio, los clientes del *Grupo 3*, son los que han comprado más recientemente en comparación con los otros; seguido por los clientes del *Grupo 6*, con un promedio entre 200 y 250 días, y así sucesivamente para los demás grupos.

Gráfica No. 2, Diferencia entre las Medias de las Variables F de los Ocho Grupos



De acuerdo a la **Gráfica No.2**, se puede concluir que los clientes del *Grupo 5* son los que más hacen compras frecuentes, con un promedio entre 160 y 180 veces durante los dos años de periodo de tiempo de evaluación.

Gráfica No. 3, Diferencia entre las Medias de las Variables M de los Ocho Grupos



El grupo de clientes que aporta mayor cantidad de dinero a la empresa, son los clientes del *Grupo 4*, según la **Gráfica No.3**, aportando en promedio, 80 millones de pesos aproximadamente.

Por otro lado, fue necesario verificar si alguna variable *RFM*, es redundante, y para ello con la ayuda de *Minitab*, por medio del acceso a las siguientes opciones en la barra de herramientas: *Estadísticas* → *Estadísticas Básicas*, y luego, *Correlación*, se pudo llevar a cabo este análisis. (Ver **Figura No. 9**).

En la **Figura No. 10** se muestra una ventana en la que se seleccionaron las variables para el análisis, y se señaló la opción “*Presentar los Valores de p*”

Las correlaciones obtenidas fueron las siguientes (ver **Tabla No. 18**):

Tabla No. 18, Correlaciones entre las Variables R, F y M arrojadas por Minitab

		R	F
F	Correlación de Pearson	-0,202	
	Valor p	0,011	
M	Correlación de Pearson	-0,366	0,069
	Valor p	0,000	0,387

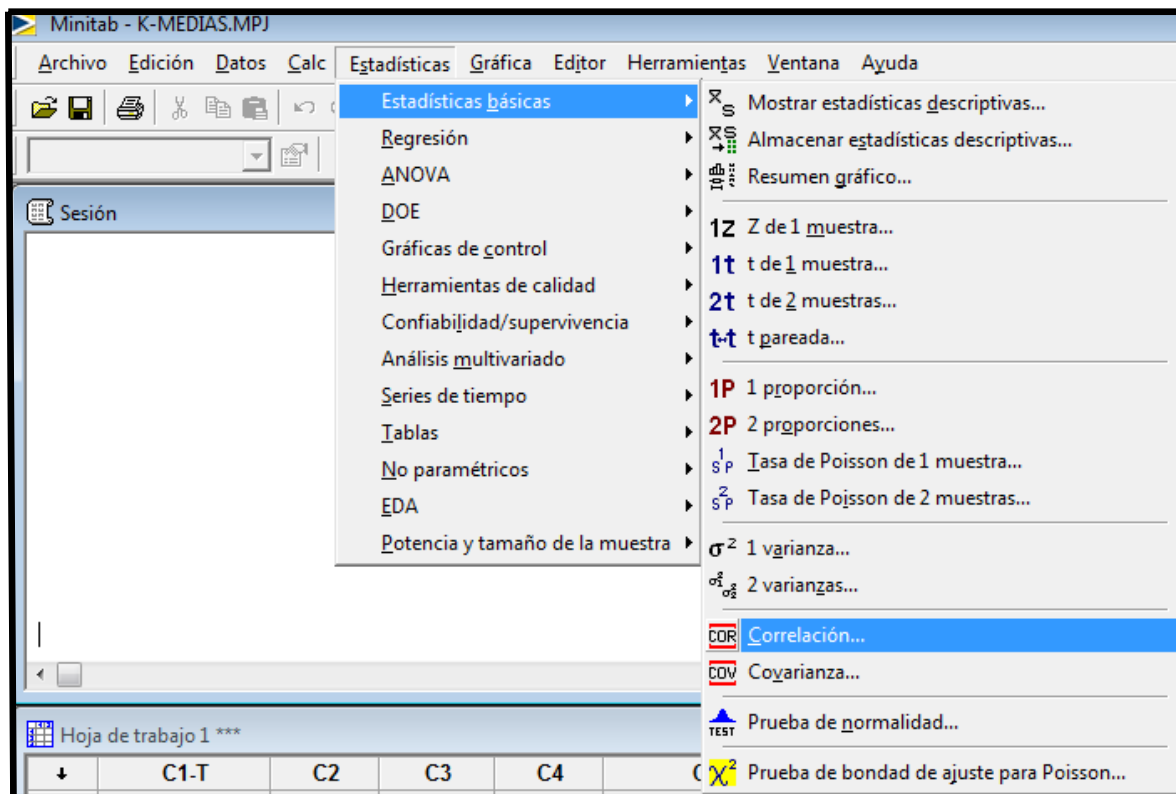


Figura No. 9, Inicializando el Cálculo del Coeficiente de Correlación en Minitab

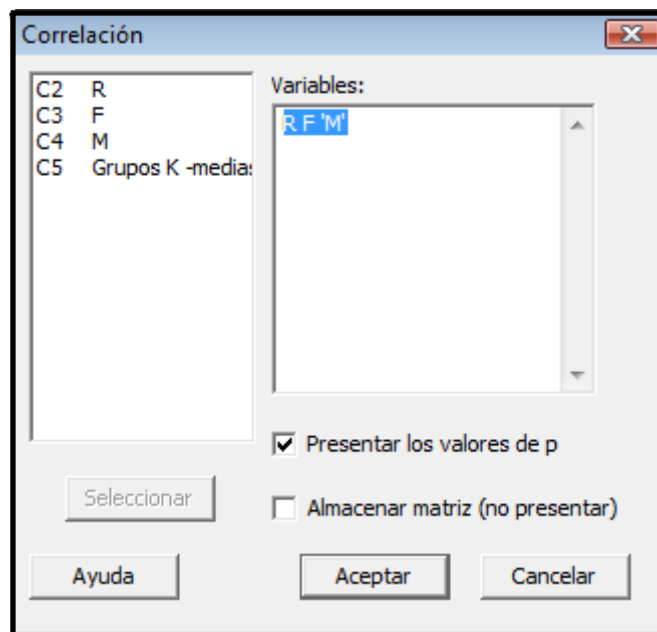


Figura No. 10, Correlación en Minitab

Minitab calculó el resultado utilizando la *Correlación de Pearson*. Las correlaciones se hicieron entre dos variables. Según la **Tabla No. 18**, el coeficiente de pearson entre las variables **R** y **F** es de, -0,202 con un valor p de 0,011; el coeficiente de pearson entre **R** y **M**, es de -0,366 con un valor p de 0,000; y el coeficiente de pearson entre las variables **F** y **M**, es de 0,069 con un valor de p de 0,387. Un valor de *p* pequeño significa que la probabilidad de que los resultados obtenidos se deban al azar es pequeña. Por lo tanto para los tres casos antes expuestos, según el valor de *p* arrojado por *Minitab*, la probabilidad de que los coeficientes de pearson calculados se deban al azar es pequeña.

Dado que el coeficiente de pearson para los tres casos, es un valor más cercano al valor de cero que al valor de uno, entonces se puede decir que la correlación entre estas dos variables no es tan fuerte. Por lo tanto ninguna de las tres variables *RFM* redundan en el análisis, y no es necesario omitir ninguna de ellas.

7.6. NORMALIZACIÓN Y PONDERACIÓN DE LAS VARIABLES RFM

Normalización de las Variables RFM

Como las variables *RFM* originales obtenidas de la base de datos son números grandes, lo que se pretende con la normalización es tener sólo referencias en la escala.

La normalización se llevó a cabo de la siguiente forma:

➤ Para normalizar las variables **F** y **M** del **Anexo No.2**, se empleó la siguiente formula, ya que **F** y **M** influyen positivamente al *Valor del Cliente*.

$$X' = (X - X^S) / (X^L - X^S),$$

donde: X' = Valor normalizado de la variable **F** o **M**.

X = Valor original de la variable **F** o **M**.

X^L = Valor más grande de la variable **F** o **M**.

X^S = Valor más pequeño de la variable **F** o **M**.

➤ Para normalizar la variable **R** del **Anexo No. 2**, se empleó la siguiente formula, ya que **R** influye negativamente al *Valor del Cliente*.

$$X' = (X^L - X) / (X^L - X^S),$$

donde: X' = Valor normalizado de la variable **R**.

X = Valor original de la variable **R**.

X^L = Valor más grande de la variable **R**.

X^S = Valor más pequeño de la variable **R**.

De acuerdo a los valores originales de las variables **RFM** calculados anteriormente (ver **Tabla No.19**), se inició la normalización de estas variables.

Tabla No. 19, Valores Originales de las Variables RFM

	A	B	C	D
	Código Cliente	R (Reciente)	F (Frecuente)	M (Monetario)
1				
2	000001	25	12	77549254
3	000002	26	8	53910982
4	000004	378	11	12490939
5	000008	19	8	64935306
6	000009	3	202	43325347
7	000012	222	7	32276505
8	000013	164	4	23297511
9	000014	24	8	66615199
10	000016	5	20	91216061
11	000018	3	24	17000003
12	...	...	...	...
13	...	...	...	...
14	000637	2	6	16109914
15	000638	1	9	86037489
16	000639	143	7	65739523
17	000640	26	9	48425485
18	000642	24	12	67296377

En la **Figura No. 11**, se calcularon los valores de X^L y X^S , presentes en las anteriores formulas ($X' = (X - X^S) / (X^L - X^S)$), y $X' = (X^L - X) / (X^L - X^S)$), considerando los datos de la **Tabla No. 19**.

	F	G	H	I	
1		R (Reciente)	F (Frecuente)	M (Monetario)	
2	X^L	563	202	219257706	=MAX(D2:D18)
3	X^S	1	2	1200840	=MIN(D2:D18)

Arrows from the table to formulas:
 From G2 (563) to =MAX(B2:B18)
 From G3 (1) to =MIN(B2:B18)
 From H2 (202) to =MAX(C2:C18)
 From H3 (2) to =MIN(C2:C18)

Figura No. 11, Cálculo de X^L y X^S

Por último se calcularon los valores normalizados (R_N , F_M , M_N) de las variables **RFM**, considerando la **Tabla No. 19** y la **Figura No. 11**.

En la **Figura No. 12**, se muestra cómo se realizó en cálculo de los valores normalizados de las variables **RFM**.

	N	O	P	Q
1	Valores Normalizados de las Variables RFM			
2	Código Cliente	R Normalizado	F Normalizado	M Normalizado
3	000001	0,9573	0,0500	0,3501
4	000002	0,9555	0,0300	0,2417
5	000004	0,3292	0,0450	0,0518
6	000008	0,9680	0,0300	0,2923
7	000009	0,9964	1,0000	0,1932
8	000012	0,6068	0,0250	0,1425
9	000013	0,7100	0,0100	0,1013
10	000014	0,9591	0,0300	0,3000
11	000016	0,9929	0,0900	0,4128
12	...	...	...	...
13	...	...	...	...
14	000632	0,9733	0,0200	0,2686
15	000637	0,9982	0,0200	0,0684
16	000638	1,0000	0,0350	0,3891
17	000639	0,7473	0,0250	0,2960
18	000640	0,9555	0,0350	0,2166
19	000642	0,9591	0,0500	0,3031

$=((D2-I3)/(I2-I3))$

$=((G2-B17)/(G2-G3))$ $=((C9-H3)/(H2-H3))$

Figura No. 12, Valores Normalizados de las Variables RFM (R_N , F_N , M_N)

Los valores normalizados de las variables **RFM** se reubicaron en una nueva tabla (ver **Anexo No. 4**).

Ponderación de las Variables RFM

Para la ponderación de las variables *RFM* se multiplica cada valor de las variables R_N , F_N , M_N , por sus respectivas importancias relativas. En el **Anexo No. 5**, se puede ver la tabla con los valores de las variables normalizadas y ponderadas (R_N' , F_N' , M_N'). En la **Figura No.13**, se muestra tal cálculo, considerando la **Tabla No. 17**, y la **Figura No. 12**.

A estos nuevos valores se les denotará de la siguiente forma: R_N' , F_N' , y M_N' , a los productos de las variables **RFM** normalizadas con sus respectivas importancias relativas.

	A	B	C	D
1	Valores R'F'M'			
2	Código Cliente	R _N '	F _N '	M _N '
3	000001	0,184936121	0,036175303	0,029168651
4	000002	0,184592374	0,021705182	0,02013773
5	000004	0,063593276	0,032557773	0,004313344
6	000008	0,186998606	0,021705182	0,024349536
7	000009	0,192498565	0,723506057	0,016093524
8	000012	0,117217876	0,018087651	0,011872352
9	000013	0,137155228	0,007235061	0,008441958
10	000014	0,185279869	0,021705182	0,024991333
11	000016	0,19181107	0,065115545	0,034390009
12	...	...	...	...
13	...	...	...	...
14	000632	0,188029848	0,014470121	0,022377946
15	000637	0,192842312	0,014470121	0,005695961
16	000638	0,19318606	0,025322712	0,032411553
17	000639	0,144373924	0,018087651	0,024656784
18	000640	0,184592374	0,025322712	0,018042015
19	000642	0,185279869	0,036175303	0,025251575

=(Q3*C69)

=(O19*C67) =(P17*C68)

Figura No. 13, Cálculo del Producto entre los Valores Normalizados de RFM y las Importancias Relativas de las Variables RFM, (R_N', F_N', M_N')

En la **Figura No.13** se muestra cómo se calculó los valores de R_N', F_N', y M_N', con las respectivas formulas empleadas en Excel. Las celdas O19, P17, y Q3 fueron extraídas de la **Figura No. 12**, y las celdas C67, C68, y C69, fueron extraídas de la **Tabla No. 17**.

7.7. AGRUPACIÓN DE LOS CLIENTES EN CLÚSTERES

En este estudio se empleó el método *K-means*, el cual es un algoritmo del *Clustering*.

Para llevar a cabo la agrupación de los clientes, en primer lugar, se determinó el número de grupos “*g*”, deseados.

El método propone agrupar los clientes de acuerdo a los valores de las variables **RFM**, por lo tanto se prosiguió a la clasificación de éstos en Clústeres, y fue necesario analizar el comportamiento de estas variables. Según el artículo “*A Method for Customer Lifetime Value Ranking – Combining the Analytic Hierarchy Process and Clustering Analysis*”, el análisis consistió en determinar si el valor promedio de la variable (R, F, o M) a analizar, en un Clúster, es menor o mayor que el promedio total de tal variable. Si el promedio de una variable **RFM** de un Clúster fue mayor que el promedio total, se le asignó a la variable **RFM** del Clúster, la flecha ↑, y de lo contrario, si éste es menor al promedio total, se asignó la flecha ↓.

Lo anterior explica la necesidad de haber tenido que determinar previamente el número de Clústeres, el cual debe ser **ocho**, ya que para cada Clúster se van a tener los valores ponderados de las variables **RFM**, y por lo tanto éstas van a tener 8 (2x2x2) posibles combinaciones, de acuerdo a las direcciones de las flechas asignadas (ver **Tabla No.20**).

Tabla No. 20, Posibles Combinaciones de las Entradas (RFM)

Grupo No.	Posibles Combinaciones
1	↑↓↓↓
2	↑↓↑↑
3	↑↑↑↓
4	↑↑↑↑
5	↓↑↑↑
6	↓↑↑↓
7	↓↓↓↑
8	↓↓↓↓

La asignación de las flechas a las variables *RFM* en cada Clúster, determinó el comportamiento característico de cada Clúster, y éstos fueron nombrados de acuerdo a la siguiente clasificación:

- *Clientes Potencialmente Críticos*
- *Clientes sin Valor*
- *Clientes de Oro*
- *Clientes Potenciales*

Las características de estos cuatro tipos de clientes fueron explicadas anteriormente, en la sección “*Marco Teórico*”.

De acuerdo al comportamiento característico de cada grupo de clientes dado por el *Patrón de Comportamiento*, el cual se mostrará al final de esta sección, y teniendo en cuenta la situación actual de la compañía en cuanto a su relación con los clientes, se puede concluir qué tipo de clientes serían los ocho grupos.

Dentro de los ocho grupos de clientes están los cuatro tipos de clientes (*Clientes Potencialmente Críticos*, *Clientes sin Valor*, *Clientes de Oro*, y *Clientes Potenciales*). La cantidad de grupos (*Grupo 1*, *Grupo 2*, *Grupo 3*, *Grupo 4*, *Grupo 5*, *Grupo 6*, *Grupo 7*, y *Grupo 8*) que conforman un tipo de cliente es indefinido, mientras que los grupos que lo conformen, presenten el mismo *Patrón*.

Partiendo de las variables *RFM* normalizadas y ponderadas (R_N' , F_N' , M_N') calculadas anteriormente, se ingresaron estos valores al software “*Minitab 15*”, con el fin de aplicar *K-Means*, para llevar a cabo la agrupación de los clientes en ocho grupos.

Los datos fueron ingresados al software de forma tal que cada columna correspondiera a una variable $R_N'F_N'M_N'$, (ver **Figura No. 14**). En esta figura se puede ver solo una parte de los datos. El nombre de las variables está ubicado en la parte superior. Hay 158 códigos de clientes, y 158 valores de las variables $R_N'F_N'M_N'$, para cada una.

Cientes a Agrupar

Variables R_N , F_N , y M_N

↓	C1-T	C2	C3	C4
	Código Cliente	R	F	M
1	000001	0,184936	0,036175	0,0291687
2	000002	0,184592	0,021705	0,0201377
3	000004	0,063593	0,032558	0,0043133
4	000008	0,186999	0,021705	0,0243495
5	000009	0,192499	0,723506	0,0160935
6	000012	0,117218	0,018088	0,0118724
7	000013	0,137155	0,007235	0,0084420
8	000014	0,185280	0,021705	0,0249913
9	000016	0,191811	0,065116	0,0343900
10	000018	0,192499	0,079586	0,0060360
11	000023	0,000000	0,094056	0,0008753
12	000024	0,163280	0,021705	0,0005311
13	000027	0,191467	0,057880	0,0467762
14	000028	0,191467	0,611363	0,0180272
15	000031	0,187342	0,260462	0,0035599
16	000034	0,193186	0,072351	0,0031254

Figura No. 14, Ingreso de Valores de las Variables R_N , F_N , M_N en Minitab

Luego se procedió a agrupar los clientes teniendo en cuenta estas variables, por medio del *Análisis Multivariado*. El método escogido fue el *K-Means*. Ver la **Figura No.15**, que muestra los pasos que se realizaron. El orden a seguir una vez se hayan ingresado los datos, ubicándonos en barra de herramientas, es: *Estadísticas* → *Análisis Multivariado* → *Conglomerados de K-medias*.

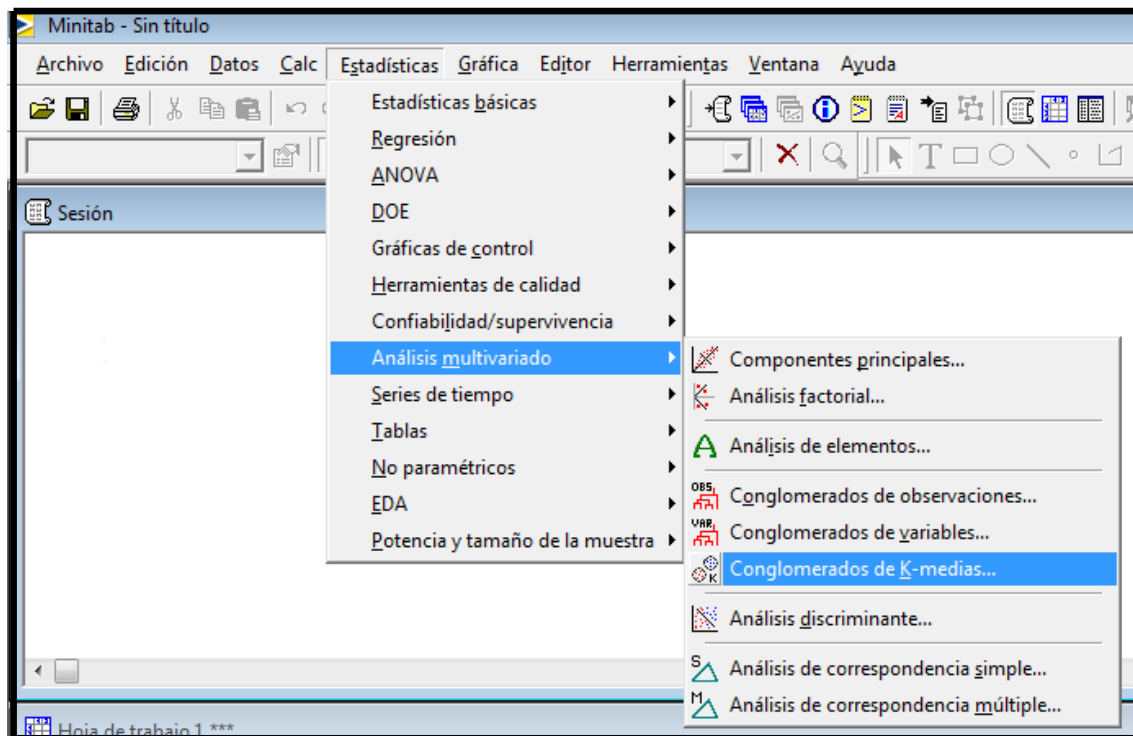


Figura No. 15, Inicializando la Agrupación de Clientes por medio de K – Means en Minitab

En la opción “Conglomerados de K-medias” se desglosó una ventana (ver **Figura No.16**) en la cual se seleccionaron las variables R_N' , F_N' , y M_N' , y el número de Conglomerados o Clústeres, el cual se determinó en el valor de ocho, **8**.

Al ingresar al botón *Almacenamiento* de la **Figura No. 16**, se desglosó una ventana (ver **Figura No.17**), en la cual se determina la columna de membresía en la opción “Columna de Membresía del Conglomerado”.

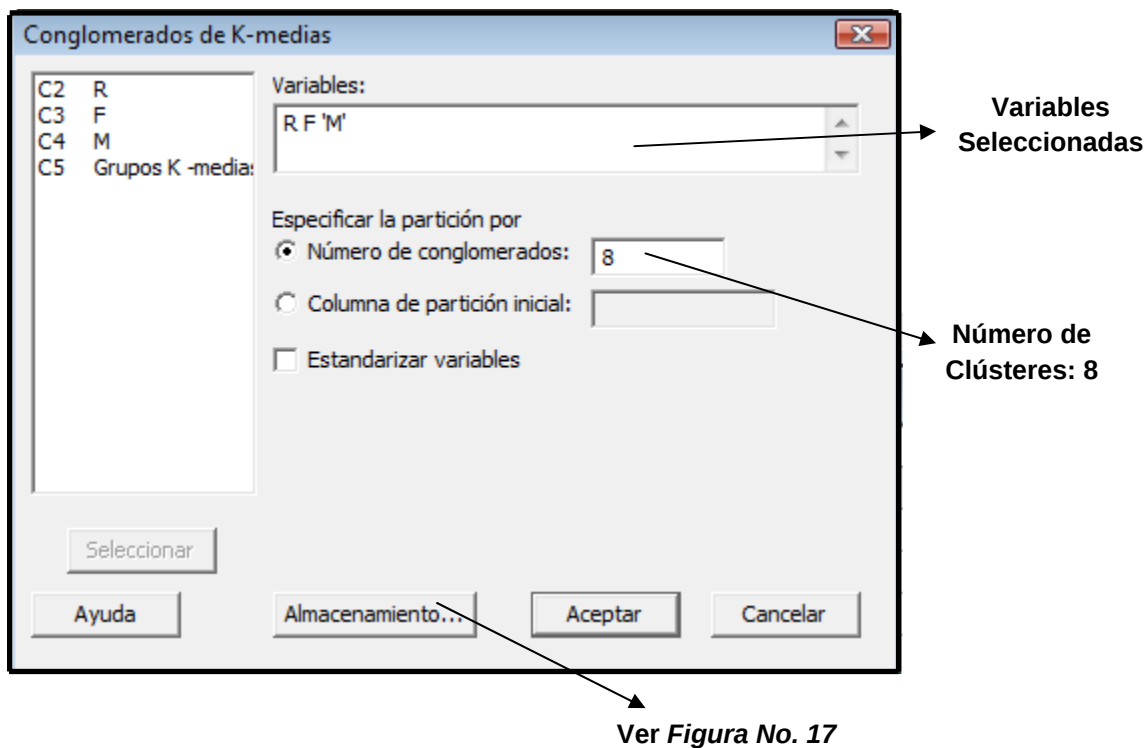


Figura No. 16, K – Means en Minitab

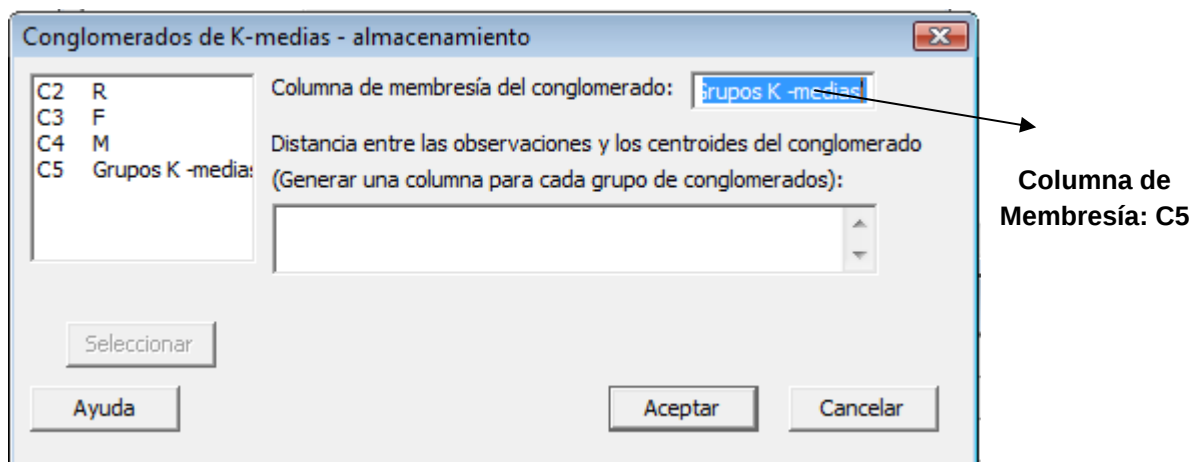


Figura No. 17, Columna de Membresía en Minitab

Hoja de trabajo 1 ***					
↓	C1-T	C2	C3	C4	C5
	Código Cliente	R	F	M	Grupos K - Medias
1	000001	0,184936	0,036175	0,0291687	4
2	000002	0,184592	0,021705	0,0201377	2
3	000004	0,063593	0,032558	0,0043133	3
4	000008	0,186999	0,021705	0,0243495	2
5	000009	0,192499	0,723506	0,0160935	5
6	000012	0,117218	0,018088	0,0118724	6
7	000013	0,137155	0,007235	0,0084420	7
8	000014	0,185280	0,021705	0,0249913	2
9	000016	0,191811	0,065116	0,0343900	4
10	000018	0,192499	0,079586	0,0060360	8
11	000023	0,000000	0,094056	0,0008753	3
12	000024	0,163280	0,021705	0,0005311	7
13	000027	0,191467	0,057880	0,0467762	4
14	000028	0,191467	0,611363	0,0180272	5
15	000031	0,187342	0,260462	0,0035599	1
16	000034	0,193186	0,072351	0,0031254	8

Columna de Membresía:
Agrupación de los Clientes en los ocho grupos

Figura No. 18, Columna de Membresía en la Hoja de Trabajo de Minitab una vez Obtenida la agrupación por K-mean

En la **Figura No. 18**, se muestra la columna de membresía (Columna C5), en la cual se determinó el número del grupo o clúster al cual pertenece cada cliente. En el **Anexo No.6**, se puede ver la agrupación que arrojó *Minitab* a partir de *K-medias*.

Como resultado del análisis multivariado que arrojó *Minitab*, se pudo observar que:

- Los centroides o medias de cada variable por grupo son (ver **Tabla No.21**):

Tabla No. 21, Medias de cada Variable R_N F_N M_N en los Clústeres

Variable	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
R'_N	0,1886	0,1819	0,0594	0,1847	0,1892	0,1148	0,1477	0,1825
F'_N	0,1922	0,0185	0,0457	0,0515	0,6174	0,0165	0,0177	0,0881
M'_N	0,0185	0,0145	0,0028	0,0307	0,0183	0,0104	0,0117	0,0132

➤ Los números de observaciones, es decir el número de clientes en un grupo o clúster, fueron (ver **Tabla No. 22**):

Tabla No. 22, Número de Clientes u Observaciones en cada Grupo arrojado por Minitab

Grupo No.	Número de Observaciones o Clientes
<i>Grupo 1</i>	7
<i>Grupo 2</i>	44
<i>Grupo 3</i>	11
<i>Grupo 4</i>	37
<i>Grupo 5</i>	6
<i>Grupo 6</i>	16
<i>Grupo 7</i>	20
<i>Grupo 8</i>	17

La **Tabla No.22**, muestra el número de observaciones, es decir, el número de clientes que conforman cada grupo. La sumatoria del número de observaciones debe ser igual al número total de clientes, es decir, igual a **158**.

7.8. ANÁLISIS DE LOS GRUPOS RESULTANTES DEL K-MEANS

Con el fin de analizar la diferencia entre promedios de las variables R_N', F_N', M_N' de los ocho grupos obtenidos por *K-Means*, y así saber si hay una discriminación significativa entre estos segmentos del mercado, se hizo un análisis de varianza por medio de las herramientas de *Minitab*. Este análisis no fue realizado para todas las variables de manera conjunta, en su lugar se llevo a cabo entre cada una de las variables de estudio, es decir se hizo el ANOVA para los valores de las variables R_N' de los ocho grupos, e igualmente para los valores de las variables F_N' y M_N' .

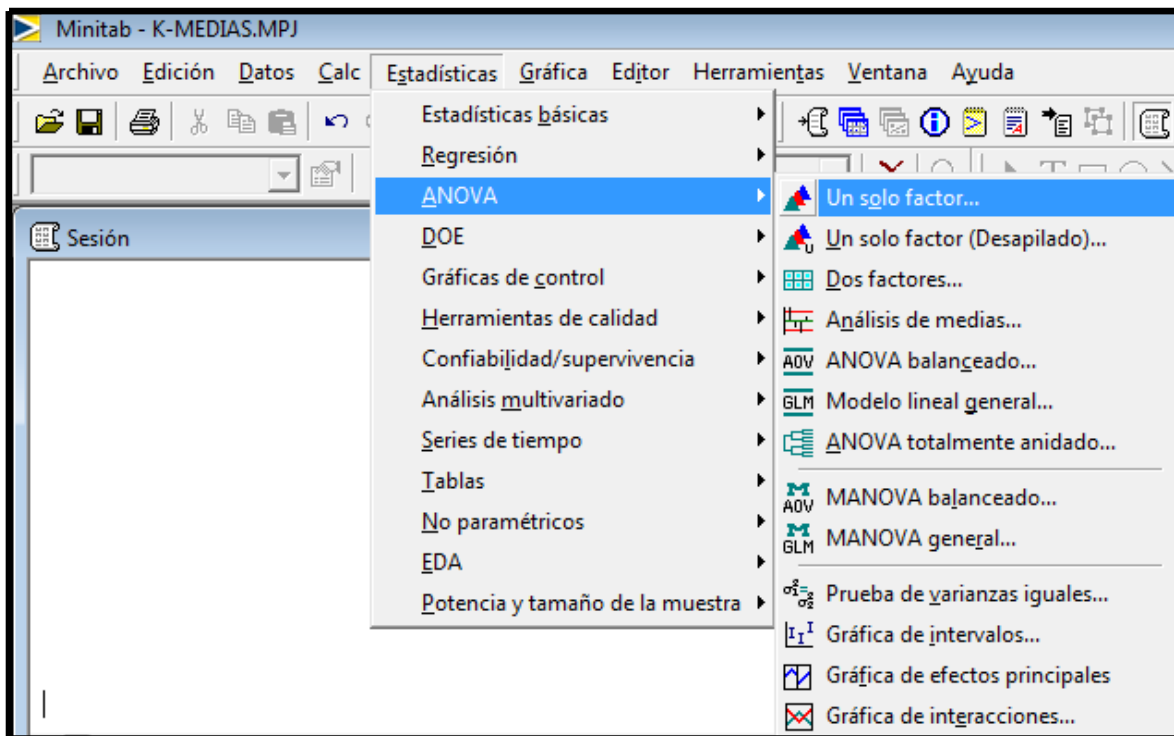


Figura No. 19, Inicializando el ANOVA (Análisis de Varianza) en Minitab

En la **Figura No.19**, por medio de las siguientes opciones de la barra de herramientas: *Estadísticas* → **ANOVA** → *Un solo factor*, se llevo a cabo el análisis de la varianza entre las tres variables. Es de un solo factor porque el resultado fue una sola clasificación, que es la de los clientes en ocho grupos.

Luego se desglosó la siguiente ventana (ver **Figura No. 20**):

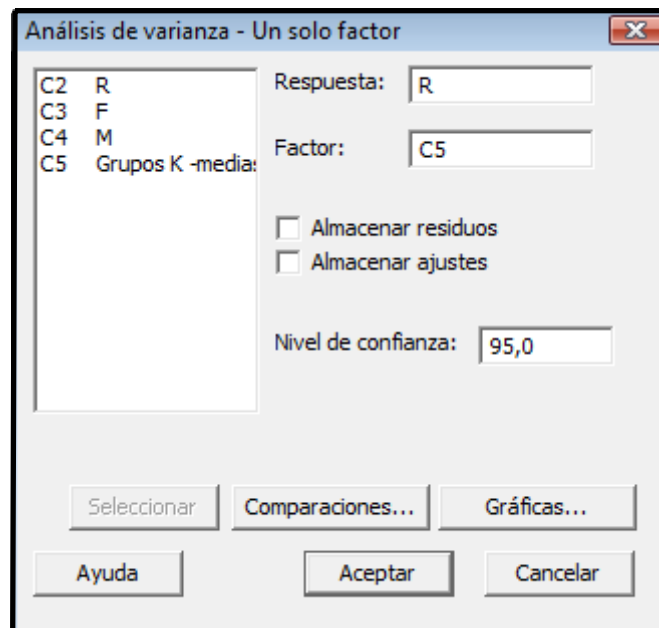


Figura No. 20, ANOVA (Análisis de Varianza) en Minitab

En la **Figura No.20** se puede observar que la *Respuesta* es la variable a analizar (en este caso a modo de ejemplo se muestra para R); que el *Factor* es el criterio de análisis, es decir la clasificación de los clientes en los ocho grupos; y que el nivel de confianza se determinó como 95,0%.

El ANOVA parte de una prueba de hipótesis, la cual supone verdadera una afirmación inicial y luego prueba esta afirmación mediante datos de muestra.

Las pruebas de hipótesis incluyen dos hipótesis: la hipótesis nula (señalada como H_0) y la hipótesis alternativa (señalada como H_1). La hipótesis nula es la afirmación inicial y con frecuencia se especifica sobre la base de investigaciones anteriores o de conocimiento común. La hipótesis alternativa es lo que se creyó que puede ser verdadero.²⁹ En este caso la hipótesis nula es que los promedios de las variables $R_N'F_N'M_N'$ son iguales para los ocho grupos; y la hipótesis alternativa es que los promedios de las variables $R_N'F_N'M_N'$ no son iguales para los ocho grupos de clientes.

$$H_0: \mu_R = \mu_F = \mu_M$$

H1: Los promedios de las variables $R_N'F_N'M_N'$ no son iguales.

²⁹ Meet Minitab 15. 2007.

El ANOVA unidireccional en *Minitab* se realizó para determinar si hay o no, diferencia significativa entre los grupos de clientes. Este ANOVA probó la igualdad entre las variables (variables $R_N'F_N'M_N'$ de un grupo con las variables $R_N'F_N'M_N'$ de otro grupo, respectivamente) de acuerdo a la clasificación de los clientes.

El proceso de decisiones para la prueba de hipótesis se puede basar en el valor de la probabilidad (valor p). Si el valor p es menor o igual que un nivel determinado de significancia (nivel α , que fue de 0,05), se rechaza la hipótesis nula (los promedios de las variables $R_N'F_N'M_N'$ son iguales) y se acepta la hipótesis alternativa (los promedios de las variables $R_N'F_N'M_N'$ no son iguales); y si por el contrario, el valor p es mayor que el nivel α , no se rechaza la hipótesis nula, y no se acepta la hipótesis alterna.

Al hacer el análisis de varianza (ANOVA) para cada variable, entre los ocho grupos se obtuvieron los resultados mostrados en las figuras: **Figura No. 21**, **Figura No.22**, y la **Figura No. 23**.

- Para la variable R_N' , se tienen los siguientes resultados:

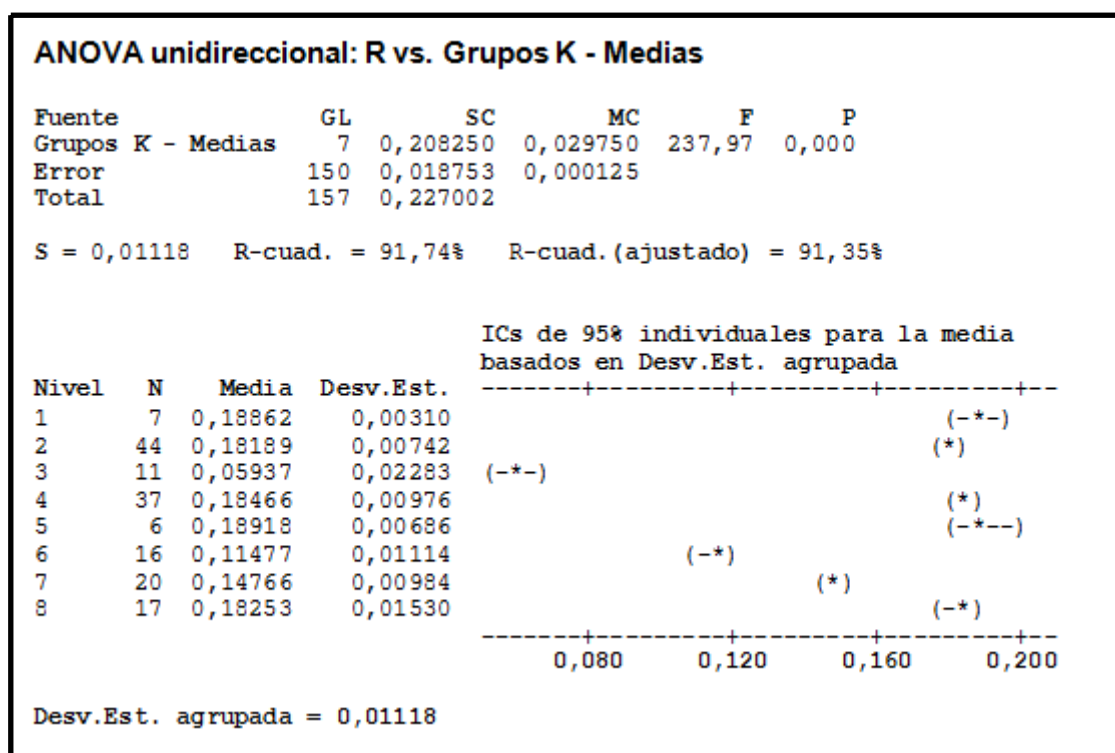


Figura No. 21, ANOVA para la Variable R_N' en Minitab

De la **Figura No. 21**, se tiene que el valor p el cual está encabezado como “P”, es de 0,000, y este valor es la probabilidad de hallar un valor F a la derecha de 237,97 con siete grados de libertad en el numerador y 150 en el denominador, dado que H_0 sea verdadera.

Como el valor de **P**, es de 0,000 según los resultados arrojados por *Minitab*, y este valor es menor que el nivel de significancia α (0,05), se aceptó la hipótesis alternativa, en la cual se determinó que los promedios de las variables R_N' no son iguales, y se rechazó la hipótesis nula en la cual se determinó que los promedios de las variables R_N son iguales. Por lo tanto, se concluyó que los promedios de las variables R_N' no son iguales para los ocho grupos.

En la **Figura No. 21**, se puede ver el número de grupos (*Nivel*), el número de Clientes que conforman cada grupo (N), la media de la variable R_N' para cada grupo, y la desviación estándar. La figura muestra el comportamiento de la variable R_N' para cada grupo de clientes.

➤ Para la Variable F_N' , se tienen los siguientes resultados:

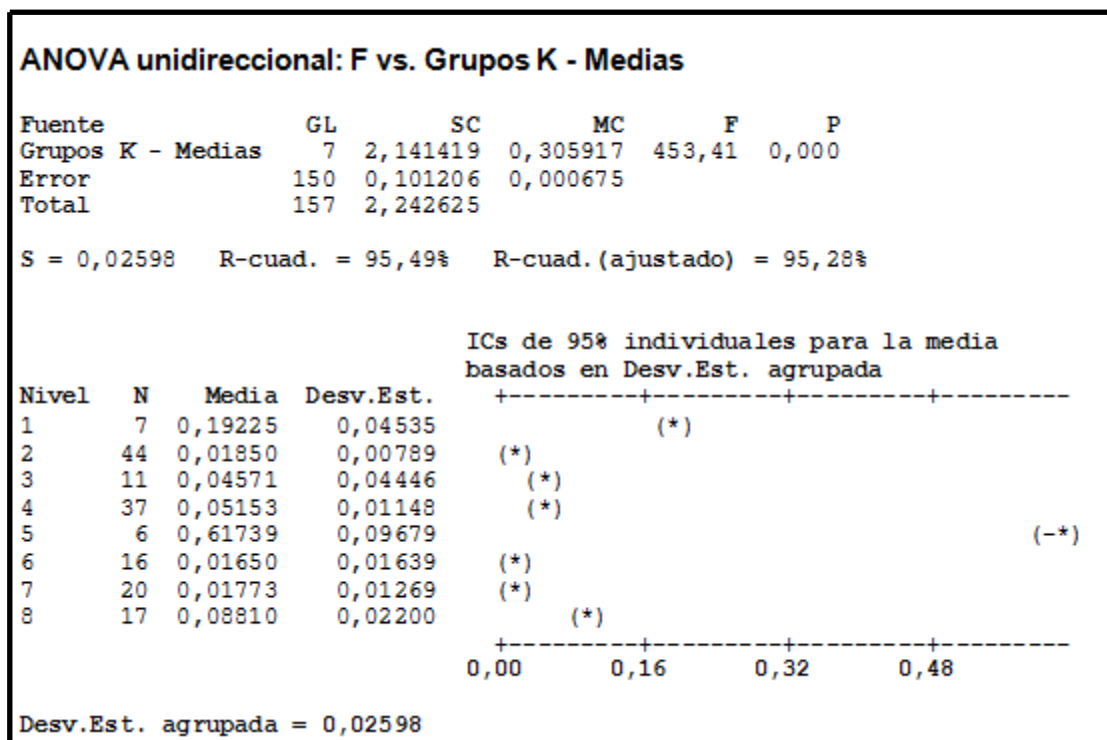


Figura No. 22, ANOVA para la Variable F_N' en Minitab

De la **Figura No. 22**, se tiene que el valor p es de 0,000, y este valor es la probabilidad de hallar un valor F a la derecha de 453,41 con siete grados de libertad en el numerador y 150 en el denominador, dado que H_0 sea verdadera.

Como el valor de P , es de 0,000, y este valor es menor que el nivel de significancia α (0,05), se aceptó la hipótesis alternativa, en la cual se determinó que los promedios de las variables F_N' no son iguales, y se rechazó la hipótesis nula en la cual se determinó que los promedios de las variables F_N' son iguales. Por lo tanto, se concluyó que los promedios de las variables F_N' no son iguales para los ocho grupos.

➤ Para la Variable M_N' , se tienen los siguientes resultados:

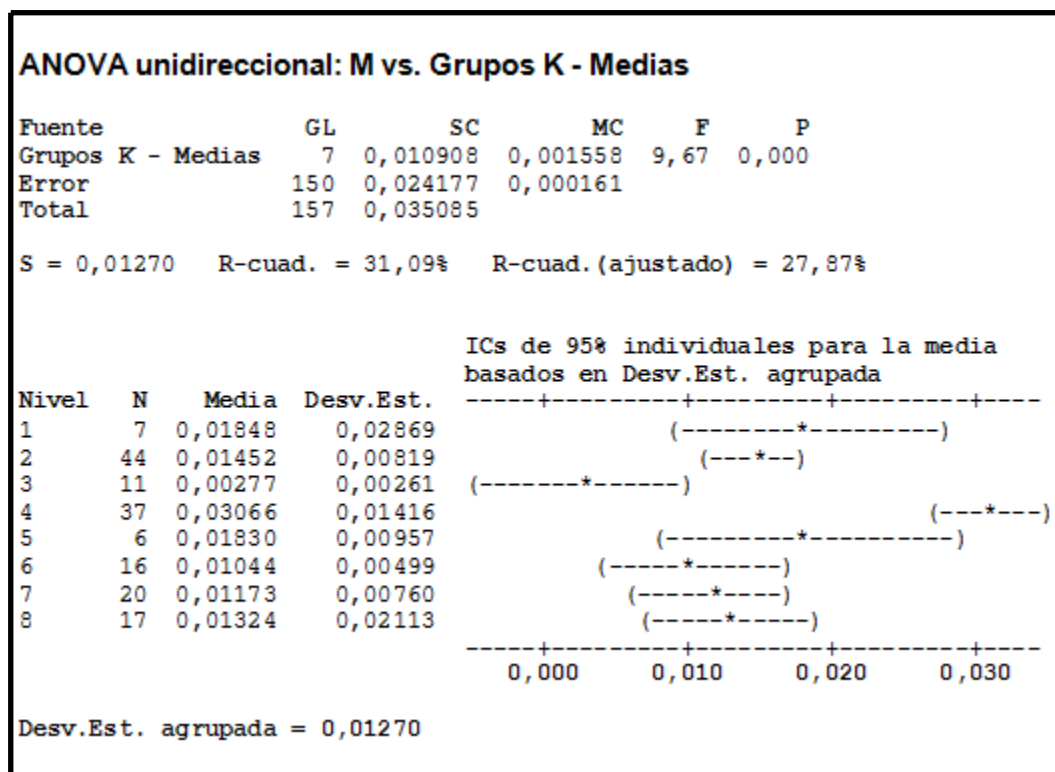


Figura No. 23, ANOVA para la Variable M_N' en Minitab

En la **Figura No. 23**, se tiene que el valor p es de 0,000, y este valor es la probabilidad de hallar un valor F a la derecha de 9,57 con siete grados de libertad en el numerador y 150 en el denominador, dado que H_0 sea verdadera.

Dado que el valor de P , es de 0,000, y este valor es menor que el nivel de significancia α (0,05), se aceptó la hipótesis alternativa, en la cual se determinó

que los promedios de las variables M_N' no son iguales, y se rechazó la hipótesis nula en la cual se determinó que los promedios de las variables M_N' son iguales. Por lo tanto, se concluyó que los promedios de las variables M_N' no son iguales para los ocho grupos.

En conclusión, del *Análisis de Varianza* arrojado por *Minitab*, se concluyó que sí existe una diferencia significativa entre las variables $R_N'F_N'M_N'$ de los ocho grupos.

Los Grupos o Clústeres de clientes ya fueron definidos, y se comprobó que la diferencia entre las variables $R_N'F_N'M_N'$ de los grupos era significativa, lo que dio a pensar que sí era razonable tener los ocho grupos. Luego se prosiguió a realizar un análisis del comportamiento de las variables de estudio por cada grupo de clientes.

El análisis a seguir consiste en:

- Listar los ocho grupos con sus respectivos valores de las variables $R_N'F_N'M_N'$ de cada cliente. (ver **Anexo No.6**).
- Calcular el promedio de cada una de las variables *RFM* por grupo.
- En una tabla anexar, el número de clientes por grupo; el promedio de cada una de las variables *RFM* de los grupos; y el patrón de comportamiento de las variables *RFM*. (Ver **Tabla No.23**). En la tabla se debe calcular el promedio total por cada variable.
- Calcular el patrón teniendo en cuenta si el promedio de una variable es inferior o superior al valor del promedio total calculado. Si el promedio de la variable es inferior al promedio total, entonces se asignó una flecha con dirección hacia abajo, y de lo contrario, si el promedio de la variable es superior al valor del promedio total, entonces se asignó una flecha con dirección hacia arriba. Este patrón se anexó como una última columna en la **Tabla No.23**.

	A	B	C	D	E
1	Grupo de Clientes	Código Clientes	R (Reciente)	F (Frecuente)	M (Monetario)
2	1	000031	18,0000	74,0000	10518920,0000
3	1	000058	19,0000	42,0000	219257706,0000
4	1	000060	29,0000	43,0000	21233764,0000
5	1	000271	14,0000	43,0000	18036328,0000
6	1	000318	12,0000	60,0000	19799005,0000
7	1	000319	1,0000	63,0000	29398311,0000
8	1	000481	7,0000	61,0000	28781107,0000
9	PROMEDIO		14,2857	55,1429	49575020,1429
10	...		...	...	...
11	...		...	...	...
12	8	000253	2,0000	23,0000	2641359,0000
13	8	000321	2,0000	23,0000	3098920,0000
14	8	000413	23,0000	22,0000	2986420,0000
15	8	000526	19,0000	35,0000	2919277,0000
16	PROMEDIO		32,0000	26,3529	35852231,2941

=PROMEDIO
(C2:C8)

Figura No. 24, Cálculo del Valor Promedio de las Variables RFM de cada Grupo o Clúster en Excel

En la **Figura No. 24** se muestra cómo se calcula el valor promedio de las variables RFM para cada grupo de clientes, lo cual fue necesario para seguir con el análisis.

Tabla No. 23, Resultados del K - Means entre 158 Clientes

Clúster de Clientes	Número de Clientes	R	F	M	Patrón
1	7	14,2857	55,1429	49575020,1429	R ↓ F ↑ M ↑
2	44	33,8636	7,1136	39213639,3636	R ↓ F ↓ M ↓
3	11	390,2727	14,6364	8449449,0909	R ↑ F ↓ M ↓
4	37	25,8108	16,2432	81450007,3514	R ↓ F ↓ M ↑
5	6	12,6667	172,6667	49102029,1667	R ↓ F ↑ M ↑
6	16	229,1250	6,5625	28538028,4375	R ↑ F ↓ M ↓
7	20	133,4500	6,9000	31906934,0500	R ↑ F ↓ M ↓
8	17	32,0000	26,3529	35852231,2941	R ↓ F ↓ M ↓
Promedio Total		108,9343	38,2023	40510917,3621	

Las observaciones del patrón para cada grupo fueron las siguientes:

- Para el **Grupo 1**, estos clientes en promedio, no han realizado compras recientemente; son frecuentes para comprar; y aportan grandes cantidades monetarias.
- Para el **Grupo 2**, estos clientes en promedio, no han comprado recientemente, ni son frecuentes en sus compras, y tampoco contribuyen grandes cantidades monetarias en éstas.
- Para el **Grupo 3**, los clientes en promedio, han realizado compras recientemente, pero no han realizado compras frecuentes, ni tampoco contribuyen monetariamente en sus compras.
- Para el **Grupo 4**, estos clientes en promedio no han comprado recientemente, ni son frecuentes a la hora de comprar, pero si aportan grandes cantidades monetarias en sus compras.
- Para el **Grupo 5**, estos clientes tienen un comportamiento similar al de los clientes del **Grupo 1**.
- Para el **Grupo 6**, estos clientes tienen un comportamiento similar al de los clientes del **Grupo 3**.
- Para el **Grupo 7**, estos clientes tienen un comportamiento similar al de los clientes del **Grupo 6**, y a los del **Grupo 3**.
- Para el **Grupo 8**, estos clientes tienen un comportamiento similar al **Grupo 2**.

Una vez se calculó el patrón de comportamiento de cada uno de los ocho grupos, es necesario aclarar ciertas suposiciones relacionadas con la situación actual de la empresa del caso de estudio, y con las relaciones existentes entre ésta y sus clientes, las cuales permitirán hacer una adecuada clasificación de los grupos de clientes en los cuatro tipos de clientes definidos (*Clientes Potencialmente Críticos*, *Clientes sin Valor*, *Clientes de Oro*, y *Clientes Potenciales*).

Suposiciones sobre los Clientes de la Empresa

1. Los clientes de la empresa son relativamente nuevos (aproximadamente de seis años atrás), ya que se trata de una compañía nueva en el sector de las herramientas de la construcción. Sus clientes son compañías que se

desarrollan en este sector, y tienen una gran capacidad monetaria. La mayoría de estos clientes tienen varios proveedores, lo que hace que la empresa del caso de estudio no sea su único proveedor, por lo cual ésta se está preparando para enfrentar retos competitivos.

2. Para la empresa, es más importante que los clientes compren frecuentemente a que, aporten grandes cantidades de dinero, o hayan comprado recientemente, como se pudo ver en los resultados de *AHP*.
3. Los clientes, en los últimos años, han comprado frecuentemente, ya que éstos han hecho parte de los proyectos de construcción que se han llevado a cabo en la ciudad.
4. Existen nuevos clientes que apenas están empezando a establecer una relación estrecha con la empresa. Estos clientes están aumentando gradualmente su capacidad de compra.

7.9. CÁLCULO DEL VALOR DEL CLIENTE (CLV)

Para el cálculo del *Valor del Cliente (CLV)*, el cual define qué tan leal es un cliente de acuerdo a una escala del 1 al 8 (teniendo en cuenta que son ocho grupos de clientes), el valor de uno corresponde al grupo de clientes más leal, y así mismo, el valor de ocho corresponde al grupo de clientes menos leal. Este cálculo se basó en una serie de cálculos, los cuales se describen a continuación:

En primer lugar se tuvo que calcular el promedio de las variables $R_N F_N M_N$ por cada grupo de clientes. Esto se realizó de la siguiente forma: (ver **Figura No.25**).

En la **Figura No. 25**, se calculó el promedio para cada variable $R_N F_N M_N$ de los grupos, tal como se puede ver en la celda **C9**, por ejemplo.

A estos promedios de las variables $R_N F_N M_N$, se les denominó de la siguiente forma C_R^j , C_F^j , y C_M^j , respectivamente, para $j = 1$ hasta k (el número de grupos = 8). Estos valores se reubicaron en la **Tabla No. 24**.

	A	B	C	D	E
1	Grupos K - Medias	Código Cliente	R _N	F _N	M _N
2	1	000031	0,9698	0,3600	0,0427
3	1	000058	0,9680	0,2000	1,0000
4	1	000060	0,9502	0,2050	0,0919
5	1	000271	0,9769	0,2050	0,0772
6	1	000318	0,9804	0,2900	0,0853
7	1	000319	1,0000	0,3050	0,1293
8	1	000481	0,9893	0,2950	0,1265
9	PROMEDIO		0,9764	0,2657	0,2218
10	...		...	...	...
11	...		...	...	...
12	8	000168	0,9804	0,0900	0,0048
13	8	000190	0,7331	0,1700	0,0094
14	8	000253	0,9982	0,1050	0,0066
15	8	000321	0,9982	0,1050	0,0087
16	8	000413	0,9609	0,1000	0,0082
17	8	000526	0,9680	0,1650	0,0079
18	PROMEDIO		0,9448	0,1218	0,1589

= PROMEDIO (C2:C8)

Figura No. 25, Cálculo del Valor Promedio de las Variables $R_N F_N M_N$ para cada Grupo de Clientes

Teniendo en cuenta las importancias relativas de las variables RFM calculadas anteriormente (ver **Tabla No. 17**), y la **Tabla No. 24**, se calculó la *Evaluación Integrada*, C^j .

La *Evaluación Integrada* fue calculada como la suma ponderada de C_R^j , C_F^j , y C_M^j , lo cual es: $C^j = W_R * C_R^j + W_F * C_F^j + W_M * C_M^j$, donde W_R , W_F Y W_M son la importancia relativa de las variables RFM .

Tabla No. 24, Valores de C_R^j , C_F^j , y C_M^j

	A	B	C	D
1	Clúster No.	C_R^j	C_F^j	C_M^j
2	1	0,9764	0,2657	0,2218
3	2	0,9415	0,0256	0,1743
4	3	0,3073	0,0632	0,0332
5	4	0,9559	0,0712	0,3680
6	5	0,9792	0,8533	0,2197
7	6	0,5941	0,0228	0,1254
8	7	0,7643	0,0245	0,1408
9	8	0,9448	0,1218	0,1589

	F	G
1	Clúster No.	Evaluación Integrada, C_I^j
2	1	0,3993
3	2	0,2149
4	3	0,1079
5	4	0,2668
6	5	0,8249
7	6	0,1417
8	7	0,1771
9	8	0,2839

$$= ((B2 * C67) + (C2 * C68) + (D2 * C69))$$

Las celdas B2, C2, y D2 pertenecen a la **Tabla No. 24**, y las celdas C67, C68, y C69, son de la **Tabla No. 17**.

Figura No. 26, Cálculo de la Evaluación Integrada para todos los Grupos

En la **Figura No. 26**, se presenta cómo se calcula el C_I^j , y para ello se hizo uso de los valores de C_R^j , C_F^j , y C_M^j de la **Tabla No. 24**, y las importancias relativas del **Tabla No. 17**.

Luego se recopiló los anteriores datos en la siguiente tabla (ver **Tabla No. 25**). En esta tabla se calculó la *Clasificación de la lealtad*, o el *Valor del Cliente*.

Tabla No. 25, CLV (Valor del Cliente)

Clúster No.	Número de Clientes	R	F	M	Evaluación Integrada	Clasificación de la Lealtad
1	7	0,9764	0,2657	0,2218	0,3993	2
2	44	0,9415	0,0256	0,1743	0,2149	5
3	11	0,3073	0,0632	0,0332	0,1079	8
4	37	0,9559	0,0712	0,3680	0,2668	4
5	6	0,9792	0,8533	0,2197	0,8249	1
6	16	0,5941	0,0228	0,1254	0,1417	7
7	20	0,7643	0,0245	0,1408	0,1771	6
8	17	0,9448	0,1218	0,1589	0,2839	3

De la anterior tabla (**Tabla No. 25**), se puede decir que los clientes más leales, son los que pertenecen al *Grupo 5*, con una clasificación de uno, y los menos leales, son los del *Grupo 3*, con una clasificación de ocho.

8. DISCUSIÓN Y ANÁLISIS DE RESULTADOS

El cálculo del *Valor del Cliente* es una herramienta muy útil para una compañía, en el momento de generar nuevas estrategias de marketing. Una vez se haya calculado este valor, y la agrupación de clientes, la cual permitirá posteriormente caracterizarlos de acuerdo al grupo al que pertenecen, y al valor de sus variables *RFM*, la empresa estará en capacidad de ubicar a un nuevo cliente en un grupo (uno de los ocho grupos ya conformados), teniendo en cuenta su comportamiento, y así mismo, saber qué trato ofrecerle a éste. También favorece a la retención de clientes ya existentes, dado que saben qué clientes son más leales (a quiénes hay que prestarles atención prioritaria); y la caracterización de los grupos permitirá a la empresa conocer sus clientes, y así brindarles un mejor servicio.

El comportamiento de las variables *RFM* refleja el comportamiento de los clientes durante los dos años de estudio. El análisis de estas variables según el ANOVA, es el siguiente, considerando la **Figura No. 21**, **Figura No. 22**, y la **Figura No. 23**):

➤ Para la variable **R** (ver **Figura No.21**):

Se pudo interpretar a partir de la figura, que los clientes del *Grupo 5* son los que más han comprado recientemente respecto a los demás grupos, seguido por los clientes del *Grupo 1*, el *Grupo 3*, el *Grupo 6*, y el *Grupo 8*, y por último por los del *Grupo 2*, el *Grupo 4*, y el *Grupo 7*.

➤ Para la variable **F** (ver **Figura No.22**):

La interpretación fue que los clientes que realizan compras más frecuentes son los del *Grupo 5*, seguidos por los demás siete grupos.

➤ Para la variable **M** (ver **Figura No. 23**):

Se interpretó que los clientes que contribuyen monetariamente más que los demás, son los Clientes del *Grupo 5*, luego los del *Grupo 1*, luego los del *Grupo 3*, después por los del *Grupo 6*, el *Grupo 7*, y el *Grupo 8*, y por último, por los del *Grupo 2*, y el *Grupo 4*.

Teniendo en cuenta que sí existe una diferencia significativa entre las variables *RFM* de cada grupo, se concluyó que sí existe una diferencia significativa entre los grupos de clientes.

Una vez fueron conformados los grupos, y se calculó la *Clasificación de la Lealtad* (ver **Tabla No.25**), se realizó el siguiente análisis teniendo en cuenta además, las suposiciones sobre los clientes de la empresa mencionadas anteriormente en la sección “ANÁLISIS DE LOS GRUPOS RESULTANTES DEL K-MEANS”.

La puntuación de uno y de dos dentro de la *Clasificación de la Lealtad* corresponde a los grupos de clientes más leales, siendo la puntuación de uno para el grupo más leal, y la puntuación de dos para el grupo más leal que le sigue. El *Grupo 5* tiene una puntuación dentro de la *Clasificación de la Lealtad*, de uno, y el *Grupo 1* tiene una puntuación de dos. A estos dos grupos se les denominó como *Clientes de Oro*. Los clientes del *Grupo 1* y del *Grupo 5*, presentan el patrón $R \downarrow F \uparrow M \uparrow$. Estos clientes son considerados como “*Clientes de Oro*”, ya que éstos contribuyen una gran cantidad monetaria a la empresa, y además compran frecuentemente. Los hace *Clientes de Oro*, ya que los promedios de la variable *F* para los dos grupos de clientes que conforman esta clasificación, son los más grandes valores de los promedios de las variables *F* de los ocho grupos (ver **Tabla No. 23**). Estos clientes presentan una frecuencia superior al promedio total, ya que se trata de una empresa que vende herramientas para la construcción, y éstas se venden cuando hay proyectos de construcción que requieran de estas herramientas, y como últimamente se han llevado a cabo mucho de éstos, entonces se explica el porqué de esta situación durante estos dos años de estudio. Además se trata de clientes que hacen grandes aportes a la rentabilidad de la empresa. Para este tipo de clientes se deben diseñar estrategias de marketing con el fin de que éstos siempre prefieran a esta compañía por encima de las demás, ofreciéndoles una atención personalizada.

Siguiendo con la misma lógica del orden de la *Clasificación de la Lealtad*, y de acuerdo a la similitud existente entre los grupos de clientes. Los clientes del *Grupo 8* pertenecen a los llamados “*Clientes Potenciales*” de la empresa, con una puntuación dentro de la *Clasificación de la Lealtad*, de tres. Este grupo tiene el patrón $R \downarrow F \downarrow M \downarrow$. Estos clientes han visitado recientemente la empresa para hacer una compra, y su actividad de compra indica que éstos tienen el potencial para convertirse en *Clientes de Oro*, ya que quizás, apenas están comenzando a establecer una relación estrecha con la empresa. Estos clientes son denominados como “*Clientes Potenciales*”. Aunque su contribución monetaria, la frecuencia de compra, y la variable reciente, estuvieron por debajo del promedio, estos clientes están empezando a aumentar gradualmente su capacidad de compra. Ofrecerles un buen servicio, e incentivos de compra, pueden ser ideas para apoyar las estrategias que se deben diseñar en torno a los *Clientes Potenciales*.

Una puntuación de cuatro y de cinco dentro de la *Clasificación de la Lealtad*, fue asignada al *Grupo 4* y *Grupo 2*, respectivamente. A estos clientes se les dio el nombre de “*Clientes Potencialmente Críticos*”. Estos grupos presentan dos patrones ($R \downarrow F \downarrow M \uparrow$ para el *Grupo 4*, y $R \downarrow F \downarrow M \downarrow$ para el *Grupo 2*). Este tipo de clientes son considerados indispensables para la compañía, ya que si éstos no existen, ocurriría una deserción, dado que una gran cantidad de clientes, 81 clientes de 158 clientes (ver **Tabla No. 26**), pertenecen a este tipo. Estos clientes aunque no sean muy frecuentes ni han comprado recientemente, son críticos para la compañía, dado que este grupo hace una gran contribución monetaria. Una estrategia para este tipo de clientes debería estar encaminada a la fidelización de éstos, ofreciéndoles siempre un buen servicio, y tentándolos a realizar compras más frecuentemente.

Por último, se tiene a los clientes menos leales, que son los de puntuación de seis, de siete y de ocho. El *Grupo 7* corresponde a una puntuación dentro de la *Clasificación de la Lealtad*, de seis; el *Grupo 6*, a una puntuación de siete; y el *Grupo 3*, a una puntuación de ocho. Estos tres grupos de clientes conforman a los llamados “*Clientes sin Valor*” de la compañía. Los grupos presentan el mismo patrón ($R \uparrow F \downarrow M \downarrow$). Estos clientes casi no hacen compras, y tampoco compran grandes cantidades en dinero. Generalmente estos clientes realizan compras en época de rebajas. Una estrategia de marketing encaminada a este tipo de clientes por parte de la empresa, puede centrarse en la reducción de los precios para atraer a estos clientes, sin causar disminuciones significativas en las ganancias de la empresa, y así retenerlos, ya que éstos fácilmente pueden cambiar de proveedor.

Una vez, se sabe qué tipo de clientes tiene, la empresa está en la capacidad de generar estrategias de marketing encaminadas a la retención de estos clientes, o a la adquisición de nuevos clientes. Aunque para la empresa, es más barato mantener un cliente existente que el desarrollo de uno nuevo.

Como se pudo ver, la *Clasificación de la Lealtad* debe responder a un orden, es decir que los grupos de los *Clientes de Oro* ocupen los primeros puestos de la clasificación, y que los grupos de los *Clientes Potenciales* ocupen los siguientes puestos, y así mismo, para los grupos de los *Clientes Potencialmente Críticos* quienes deben ocupar los siguientes puestos, y por último, los grupos de los *Clientes sin Valor*. (Ver **Tabla No. 26**). Esto se debe a que una clasificación baja corresponde a clientes más leales, que los que tienen una clasificación más alta.

Tabla No. 26, Tipos de Clientes de acuerdo a la Clasificación de la Lealtad

Clasificación de la Lealtad	Grupo No.	Número de Clientes	Tipo de Cliente
1	5	13	<i>Clientes de Oro</i>
2	1		
3	8	17	<i>Clientes Potenciales</i>
4	4	81	<i>Clientes Potencialmente Críticos</i>
5	2		
6	7	47	<i>Clientes sin Valor</i>
7	6		
8	3		

La *Clasificación de la Lealtad* le permitirá a la compañía saber cuáles son sus clientes más leales, y así, esta información servir como a poyo a las estrategias de marketing que la empresa pueda desarrollar en torno a la retención de clientes, y la búsqueda de nuevos.

9. CONSLUSIONES Y RECOMENDACIONES

- La *Técnica RFM* es una herramienta eficaz que permite conocer el comportamiento de los clientes a través de los valores de las variables RFM (Reciente, Frecuencia, y Monetario) obtenidos de una base de datos de la compañía.
- La aplicación del AHP (Proceso Jerárquico Analítico) para determinar la importancia relativa de las variables RFM es importante, ya que los pesos de estas variables quizás varían con las características de los productos y de la industria.
- La aplicación de la metodología propuesta a un caso de estudio, evidenció que una agrupación de los clientes, es una forma adecuada de caracterizar a estos e identificarlos, para así, si algún día la compañía obtiene un nuevo cliente, ésta sepa en qué grupo ubicarlo. De esta forma todos los clientes, en adelante, serán identificados fácilmente, y se sabrá que trato ofrecerles para garantizar su retención.
- El cálculo del *valor del Cliente* (CLV) o de la *Lealtad* servirá de apoyo al diseño de estrategias de marketing que la compañía pueda desarrollar con el fin de retener a sus clientes.
- La metodología propuesta aplicada a la empresa comercializadora de herramientas para la construcción en la ciudad de Cali, arrojó una clasificación de la lealtad, la cual permite saber cuáles son los clientes más leales, y cuáles son los menos leales.

BIBLIOGRAFÍA

BERUMEN, Sergio A. y REDONDO LLAMAZARES, Francisco. La utilidad de los Métodos de Decisión Multicriterio (como el AHP) en un entorno de competitividad creciente. 2007.

GREENBERG, Paul. Managing Social Customers for Profit. EBSCO, Agosto de 2009.

HARTIGAN, John A. Clustering Algorithms, a Wiley Publication in applied Statistics. Estados unidos de América, 1975.

LIU, Duen-Ren. Integrating AHP and data mining for product recommendation based on customer lifetime value. Science Direct, 2005.

MASON, Robert D., y LIND, Douglas A. Estadística para Administración y Economía. México: Alfaomega Grupo Editor, 1995.

MATSATSINI, Nikolaos F. Brand Choice model selection based on consumers' multicriteria preferences and experts' Knowledge. Science Direct, 2000.

Meet Minitab 15 for Windows. 2007.

MIROSLAW, Kwiesielewicz y UDEN, Eva Van. Inconsistent and Contradictory Judgements in Pairwise Comparison Method in the AHP. Science Direct, 2004.

PEÑA, Daniel. Análisis de Datos Multivariantes. España: Mac Graw-Hill Book Company, 2002.

PETIOT, Jean-Francois. Measuring consumer for a better comprehension, specification and assessment of product semantics. Science Direct, 2004.

SHIH, Ya-Yueh y LIU, Chung-Yuan. A method for customer lifetime value ranking-Combining the analytic hierarchy process and clustering analysis. EBSCO, Septiembre de 2003.

SHIH, Ya-Yueh y LIU, Chung-Yuan. Hierarchy Process and Clustering Analysis. EBSCO, 2003.

STANON, William J., ETZEL, Michael J. y WALKER, Bruce J. Fundamentos de Marketing. Mc Graw Hill, Novena edición, 1993.

WASS, John A. Minitab 15: Not Just for Six Sigma. EBSCO, Enero de 2008.

WILKIN, Gregory A. y HUANG, Xiuzhen. Practical Comparison of Two k-Means Clustering Algorithms. EBSCO, Agosto de 2007.

ANEXOS

Anexo No. 1, Código de Clientes de la Empresa del Caso de Estudio

Código Cliente	Código Cliente	Código Cliente
000001	000065	000158
000002	000066	000168
000004	000069	000169
000008	000072	000170
000009	000074	000171
000012	000075	000172
000013	000078	000173
000014	000079	000184
000016	000082	000190
000018	000083	000230
000023	000085	000241
000024	000088	000242
000027	000089	000246
000028	000090	000247
000031	000095	000253
000034	000096	000255
000035	000097	000256
000036	000098	000257
000037	000099	000267
000039	000100	000270
000040	000101	000271
000042	000102	000273
000043	000103	000281
000045	000105	000284
000050	000110	000287
000055	000111	000295
000056	000118	000299
000058	000139	000318
000060	000140	000319
000062	000141	000321
000063	000149	000331

Continuación Anexo No. 1, Código de Clientes de la Empresa del Caso de Estudio

Código Cliente	Código Cliente
000349	000527
000350	000533
000351	000534
000356	000539
000369	000540
000370	000559
000376	000561
000379	000564
000382	000566
000392	000575
000393	000585
000395	000586
000413	000588
000422	000589
000424	000600
000425	000601
000441	000607
000443	000617
000449	000619
000464	000620
000474	000621
000475	000622
000481	000623
000490	000625
000491	000626
000498	000628
000502	000629
000504	000632
000505	000637
000510	000638
000521	000639
000526	000640

Anexo No. 2, Valores de las Variables RFM (Reciente, Frecuente, y Monetario)

Código Cliente	R (Reciente)	F (Frecuente)	M (Monetario)
000001	25	12	77549254
000002	26	8	53910982
000004	378	11	12490939
000008	19	8	64935306
000009	3	202	43325347
000012	222	7	32276505
000013	164	4	23297511
000014	24	8	66615199
000016	5	20	91216061
000018	3	24	17000003
000023	563	28	3492000
000024	88	8	2590951
000027	6	18	123636729
000028	6	171	48386608
000031	18	74	10518920
000034	1	22	9381440
000035	78	3	21312341
000036	380	3	22528115
000037	8	8	4393460
000039	14	16	66272776
000040	358	32	4597940
000042	29	26	16519976
000043	6	37	154377865
000045	10	13	43788246
000050	104	14	36362917
000055	388	13	1200840
000056	37	5	54976059
000058	19	42	219257706
000060	29	43	21233764
000062	26	19	2203131
000063	20	21	183961426
000065	23	8	59341911
000066	60	21	15442397
000069	26	33	144093227
000072	19	31	7428579
000074	8	16	44325837
000075	200	8	43511905
000078	133	19	72490578
000079	4	155	50528168
000082	53	130	45019996
000083	28	12	8442903
000085	1	184	92661935
000088	44	19	107837227
000089	45	6	31115125

Continuación Anexo No. 2, Valores de las Variables RFM (Reciente, Frecuente, y Monetario)

Código Cliente	R (Reciente)	F (Frecuente)	M (Monetario)
000090	6	31	136459472
000095	155	5	65536559
000096	55	8	25438931
000097	138	6	23727817
000098	14	7	43867287
000099	6	15	57333276
000100	110	8	25892273
000101	23	24	80497435
000102	145	6	32381765
000103	224	4	31907744
000105	5	20	66710427
000110	211	4	35869898
000111	141	4	23711224
000118	63	4	20451496
000139	25	8	37039892
000140	136	21	8940698
000141	135	13	60574865
000149	93	15	57736896
000158	49	6	56323804
000168	12	20	2247634
000169	252	4	19392446
000170	9	194	14690121
000171	20	9	60981739
000172	59	11	67888203
000173	90	20	56813269
000184	326	23	2637600
000190	151	36	3250099
000230	368	5	18835250
000241	55	7	46782950
000242	26	17	51197022
000246	325	34	6098050
000247	141	15	62873718
000253	2	23	2641359
000255	17	25	166557720
000256	7	15	155232294
000257	28	13	83958237
000267	6	20	103977581
000270	263	5	33211705
000271	14	43	18036328
000273	174	6	22008155
000281	29	14	103110631
000284	18	7	43525564
000287	207	5	28472498
000295	2	14	96915258

Continuación Anexo No. 2, Valores de las Variables RFM (Reciente, Frecuente, y Monetario)

Código Cliente	R (Reciente)	F (Frecuente)	M (Monetario)
000299	40	10	42188859
000318	12	60	19799005
000319	1	63	29398311
000321	2	23	3098920
000331	435	4	6659473
000349	220	21	55196932
000350	237	13	42350240
000351	174	5	47268507
000356	60	8	31088878
000369	418	3	5558234
000370	23	15	79294427
000376	318	6	18933597
000379	66	20	120331689
000382	33	16	91875722
000392	13	12	95537802
000393	40	9	31455350
000395	70	4	11448273
000413	23	22	2986420
000422	89	7	47296783
000424	34	7	24367647
000425	354	5	8845499
000441	27	15	96844022
000443	23	18	79531519
000449	271	4	9194528
000464	118	6	7627259
000474	26	6	47968665
000475	58	4	21998721
000481	7	61	28781107
000490	52	8	95424212
000491	60	13	23932122
000498	27	14	85102511
000502	212	4	26270173
000504	165	5	19999451
000505	209	5	34727207
000510	212	3	4206322
000521	216	7	23577826
000526	19	35	2919277
000527	22	4	26779461
000533	53	13	59522969
000534	19	19	54469073
000539	140	5	23505649
000540	192	5	17508929
000559	16	12	116803912
000561	2	4	20756268

Continuación Anexo No. 2, Valores de las Variables RFM (Reciente, Frecuente, y Monetario)

Código Cliente	R (Reciente)	F (Frecuente)	M (Monetario)
000564	157	3	12569106
000566	60	7	27080539
000575	86	9	27119545
000585	5	17	1880000
000586	81	8	42042384
000588	5	6	33251249
000589	9	13	46523865
000600	102	2	8055103
000601	20	8	67434274
000607	2	19	66893180
000617	23	8	66911487
000619	44	6	20187021
000620	11	3	5624051
000621	37	6	16643390
000622	1	7	33878781
000623	4	16	48700785
000625	30	8	35306783
000626	28	16	32592349
000628	9	12	59829325
000629	23	6	21940972
000632	16	6	59774705
000637	2	6	16109914
000638	1	9	86037489
000639	143	7	65739523
000640	26	9	48425485
000642	24	12	67296377

Anexo No. 3, Encuesta AHP

ENTREVISTA

Fecha de la Entrevista:

Objetivo: Determinar el grado de importancia de algunos comportamientos de compra de los clientes hacia la compañía.

Conteste las siguientes preguntas de forma ordenada y lo más coherente posible.

Tenga en cuenta que:

- 1: De igual importancia.
- 2: Intermedio entre igual y débil.
- 3: Importancia débil de una sobre otra.
- 4: Intermedio entre débil y fuerte.
- 5: De importancia Fuerte.
- 6: Intermedio entre fuerte y demostrado.
- 7: De importancia demostrado.
- 8: Intermedio entre relevante y absoluto.
- 9: De importancia absoluta o extrema.

Cuestionario:

Conteste las siguientes respuestas imaginándose que usted es el dueño de la compañía.

- 1.** ¿Qué es más importante para usted, que un cliente le haya realizado una compra últimamente, o que un cliente le compre frecuentemente?

Respuesta:

Marque con una **X** qué tan importante es la primera opción sobre la segunda opción, de acuerdo a la escala mostrada anteriormente, en el siguiente espacio:

1	2	3	4	5	6	7	8	9

2. ¿Qué es más importante para usted, que un cliente le haya realizado una compra últimamente, o que éste aporte una gran cantidad monetaria por sus compras independientemente si han sido frecuentes o recientes, a la compañía?

Respuesta:

Marque con una **X** qué tan importante es la primera opción sobre la segunda opción, de acuerdo a la escala mostrada anteriormente, en el siguiente espacio:

1	2	3	4	5	6	7	8	9

3. ¿Qué es más importante para usted, que un cliente le haya comprado frecuentemente en un determinado periodo de tiempo, o que éste aporte una gran cantidad monetaria por sus compras independientemente si han sido frecuentes o recientes, a la compañía?

Respuesta:

Marque con una **X** qué tan importante es la primera opción sobre la segunda opción, de acuerdo a la escala mostrada anteriormente, en el siguiente espacio:

1	2	3	4	5	6	7	8	9

Muchas Gracias!!!

Anexo No. 4, Valores Normalizados de las Variables RFM ($R_N F_N M_N$)

Código Cliente	R Normalizado	F Normalizado	M Normalizado
000001	0,957295374	0,05	0,35013075
000002	0,955516014	0,03	0,241726587
000004	0,329181495	0,045	0,051775939
000008	0,96797153	0,03	0,292283693
000009	0,996441281	1	0,193181292
000012	0,606761566	0,025	0,142511747
000013	0,709964413	0,01	0,101334443
000014	0,959074733	0,03	0,299987614
000016	0,992882562	0,09	0,412806176
000018	0,996441281	0,11	0,072454325
000023	0	0,13	0,010507167
000024	0,84519573	0,03	0,006374993
000027	0,991103203	0,08	0,561486053
000028	0,991103203	0,845	0,216392031
000031	0,96975089	0,36	0,042732339
000034	1	0,1	0,037515902
000035	0,862989324	0,005	0,092230533
000036	0,325622776	0,005	0,097806024
000037	0,987544484	0,03	0,014641227
000039	0,976868327	0,07	0,298417276
000040	0,364768683	0,15	0,015578964
000042	0,950177936	0,12	0,07025294
000043	0,991103203	0,175	0,702463664
000045	0,983985765	0,055	0,195304128
000050	0,816725979	0,06	0,161251868
000055	0,3113879	0,055	0
000056	0,93594306	0,015	0,246610987
000058	0,96797153	0,2	1
000060	0,950177936	0,205	0,091870182
000062	0,955516014	0,085	0,004596466
000063	0,966192171	0,095	0,838132682
000065	0,960854093	0,03	0,266632609
000066	0,895017794	0,095	0,065311206
000069	0,955516014	0,155	0,655298728
000072	0,96797153	0,145	0,02856016
000074	0,987544484	0,07	0,197769498
000075	0,645907473	0,03	0,194036839
000078	0,765124555	0,085	0,326931866
000079	0,994661922	0,765	0,226213138
000082	0,90747331	0,64	0,200952884
000083	0,951957295	0,05	0,033211809
000085	1	0,91	0,419436896
000088	0,923487544	0,085	0,489030173
000089	0,921708185	0,02	0,137185705

Continuación Anexo No. 4, Valores Normalizados de las Variables RFM
($R_N F_N M_N$)

Código Cliente	R Normalizado	F Normalizado	M Normalizado
000090	0,991103203	0,145	0,620290636
000095	0,725978648	0,015	0,295041015
000096	0,903914591	0,03	0,111154909
000097	0,756227758	0,02	0,103307809
000098	0,976868327	0,025	0,195666607
000099	0,991103203	0,065	0,257421089
000100	0,806049822	0,03	0,113233917
000101	0,960854093	0,11	0,363650989
000102	0,743772242	0,02	0,142994465
000103	0,603202847	0,01	0,140820624
000105	0,992882562	0,09	0,300424326
000110	0,62633452	0,01	0,158990903
000111	0,75088968	0,01	0,103231714
000118	0,889679715	0,01	0,088282733
000139	0,957295374	0,03	0,164356448
000140	0,759786477	0,095	0,035494677
000141	0,761565836	0,055	0,272286886
000149	0,836298932	0,065	0,259272074
000158	0,914590747	0,02	0,252791692
000168	0,980427046	0,09	0,004800555
000169	0,553380783	0,01	0,083425972
000170	0,985765125	0,96	0,061861299
000171	0,966192171	0,035	0,274152794
000172	0,896797153	0,045	0,305825559
000173	0,841637011	0,09	0,255036358
000184	0,421708185	0,105	0,006588923
000190	0,733096085	0,17	0,009397819
000230	0,346975089	0,015	0,080870694
000241	0,903914591	0,025	0,20903772
000242	0,955516014	0,075	0,229280476
000246	0,423487544	0,16	0,022458408
000247	0,75088968	0,065	0,282829333
000253	0,998220641	0,105	0,006606162
000255	0,971530249	0,115	0,758319988
000256	0,989323843	0,065	0,706382041
000257	0,951957295	0,055	0,379522088
000267	0,991103203	0,09	0,471329992
000270	0,533807829	0,015	0,146800537
000271	0,976868327	0,205	0,07720687
000273	0,692170819	0,02	0,095421508
000281	0,950177936	0,06	0,467354195
000284	0,96975089	0,025	0,194099479
000287	0,633451957	0,015	0,125066725
000295	0,998220641	0,06	0,438942464

Continuación Anexo No. 4, Valores Normalizados de las Variables RFM
($R_N F_N M_N$)

Código Cliente	R Normalizado	F Normalizado	M Normalizado
000299	0,930604982	0,04	0,187969403
000318	0,980427046	0,29	0,085290435
000319	1	0,305	0,129312466
000321	0,998220641	0,105	0,008704518
000331	0,227758007	0,01	0,025033071
000349	0,610320285	0,095	0,247623902
000350	0,580071174	0,055	0,18870949
000351	0,692170819	0,015	0,211264464
000356	0,895017794	0,03	0,137065338
000369	0,258007117	0,005	0,019982833
000370	0,960854093	0,065	0,358134043
000376	0,43594306	0,02	0,081321709
000379	0,884341637	0,09	0,546329273
000382	0,943060498	0,07	0,415831355
000392	0,978647687	0,05	0,432625506
000393	0,930604982	0,035	0,138745964
000395	0,877224199	0,01	0,046994315
000413	0,960854093	0,1	0,008188598
000422	0,84341637	0,025	0,211394137
000424	0,941281139	0,025	0,106242043
000425	0,371886121	0,015	0,035058098
000441	0,953736655	0,065	0,438615778
000443	0,960854093	0,08	0,359221337
000449	0,519572954	0,01	0,036658731
000464	0,791814947	0,02	0,029471299
000474	0,955516014	0,02	0,214475361
000475	0,898576512	0,01	0,095378244
000481	0,989323843	0,295	0,126481993
000490	0,909252669	0,03	0,432104587
000491	0,895017794	0,055	0,104244743
000498	0,953736655	0,06	0,384769682
000502	0,62455516	0,01	0,114966951
000504	0,708185053	0,015	0,086209672
000505	0,629893238	0,015	0,153750568
000510	0,62455516	0,005	0,013783019
000521	0,617437722	0,025	0,102619956
000526	0,96797153	0,165	0,007880683
000527	0,962633452	0,01	0,117302525
000533	0,90747331	0,055	0,267462933
000534	0,96797153	0,085	0,24428597
000539	0,752669039	0,015	0,102288955
000540	0,660142349	0,015	0,074788239
000559	0,973309609	0,05	0,53015103
000561	0,998220641	0,01	0,089680405

**Continuación Anexo No. 4, Valores Normalizados de las Variables RFM
($R_N F_N M_N$)**

Código Cliente	R Normalizado	F Normalizado	M Normalizado
000564	0,722419929	0,005	0,05213441
000566	0,895017794	0,025	0,118683257
000575	0,848754448	0,035	0,118862137
000585	0,992882562	0,075	0,0031146
000586	0,857651246	0,03	0,187297675
000588	0,992882562	0,02	0,146981884
000589	0,985765125	0,055	0,207849566
000600	0,820284698	0	0,031433374
000601	0,966192171	0,03	0,303743859
000607	0,998220641	0,085	0,301262424
000617	0,960854093	0,03	0,301346379
000619	0,923487544	0,02	0,087069861
000620	0,982206406	0,005	0,020284667
000621	0,93594306	0,02	0,070818912
000622	1	0,025	0,149859721
000623	0,994661922	0,07	0,217832834
000625	0,948398577	0,03	0,15640848
000626	0,951957295	0,07	0,143960195
000628	0,985765125	0,05	0,26886787
000629	0,960854093	0,02	0,09511341
000632	0,973309609	0,02	0,268617384
000637	0,998220641	0,02	0,068372413
000638	1	0,035	0,389057453
000639	0,747330961	0,025	0,295971799
000640	0,955516014	0,035	0,216570319
000642	0,959074733	0,05	0,303111469

Anexo No. 5, Valores R_N F_N M_N

Código Cliente	R_N	F_N	M_N
000001	0,184936121	0,036175303	0,029168651
000002	0,184592374	0,021705182	0,02013773
000004	0,063593276	0,032557773	0,004313344
000008	0,186998606	0,021705182	0,024349536
000009	0,192498565	0,723506057	0,016093524
000012	0,117217876	0,018087651	0,011872352
000013	0,137155228	0,007235061	0,008441958
000014	0,185279869	0,021705182	0,024991333
000016	0,19181107	0,065115545	0,034390009
000018	0,192498565	0,079585666	0,006036016
000023	0	0,094055787	0,00087533
000024	0,163280033	0,021705182	0,000531087
000027	0,191467323	0,057880485	0,046776214
000028	0,191467323	0,611362618	0,018027162
000031	0,187342353	0,260462181	0,003559941
000034	0,19318606	0,072350606	0,00312537
000035	0,166717507	0,00361753	0,00768353
000036	0,062905781	0,00361753	0,008148013
000037	0,190779828	0,021705182	0,00121973
000039	0,188717343	0,050645424	0,024860511
000040	0,070468225	0,108525909	0,00129785
000042	0,183561132	0,086820727	0,005852624
000043	0,191467323	0,12661356	0,058520761
000045	0,190092333	0,039792833	0,016270373
000050	0,157780074	0,043410363	0,013433552
000055	0,060155802	0,039792833	0
000056	0,180811152	0,010852591	0,020544639
000058	0,186998606	0,144701211	0,083307883
000060	0,183561132	0,148318742	0,00765351
000062	0,184592374	0,061498015	0,000382922
000063	0,186654859	0,068733075	0,069823059
000065	0,185623616	0,021705182	0,022212598
000066	0,172904961	0,068733075	0,005440938
000069	0,184592374	0,112143439	0,05459155
000072	0,186998606	0,104908378	0,002379286
000074	0,190779828	0,050645424	0,016475758
000075	0,12478032	0,021705182	0,016164798
000078	0,147811398	0,061498015	0,027236002
000079	0,192154818	0,553482134	0,018845338
000082	0,175311193	0,463043877	0,016740959
000083	0,183904879	0,036175303	0,002766805
000085	0,19318606	0,658390512	0,0349424
000088	0,17840492	0,061498015	0,040740068
000089	0,178061173	0,014470121	0,011428651

Continuación Anexo No. 5, Valores R_N , F_N , M_N

Código Cliente	R_N	F_N	M_N
000090	0,191467323	0,104908378	0,0516751
000095	0,140248955	0,010852591	0,024579242
000096	0,174623698	0,021705182	0,00926008
000097	0,146092661	0,014470121	0,008606355
000098	0,188717343	0,018087651	0,016300571
000099	0,191467323	0,047027894	0,021445206
000100	0,155717589	0,021705182	0,009433278
000101	0,185623616	0,079585666	0,030294994
000102	0,143686429	0,014470121	0,011912566
000103	0,116530381	0,007235061	0,011731468
000105	0,19181107	0,065115545	0,025027715
000110	0,120999098	0,007235061	0,013245196
000111	0,145061419	0,007235061	0,008600016
000118	0,171873719	0,007235061	0,007354648
000139	0,184936121	0,021705182	0,013692188
000140	0,146780156	0,068733075	0,002956986
000141	0,147123903	0,039792833	0,022683644
000149	0,161561296	0,047027894	0,021599408
000158	0,176686183	0,014470121	0,021059541
000168	0,189404838	0,065115545	0,000399924
000169	0,106905453	0,007235061	0,006950041
000170	0,19043608	0,694565815	0,005153534
000171	0,186654859	0,025322712	0,022839089
000172	0,173248709	0,032557773	0,02547768
000173	0,162592538	0,065115545	0,021246539
000184	0,081468143	0,075968136	0,000548909
000190	0,141623944	0,12299603	0,000782912
000230	0,06703075	0,010852591	0,006737166
000241	0,174623698	0,018087651	0,01741449
000242	0,184592374	0,054262954	0,019100871
000246	0,08181189	0,115760969	0,001870962
000247	0,145061419	0,047027894	0,023561913
000253	0,192842312	0,075968136	0,000550345
000255	0,187686101	0,083203197	0,063174033
000256	0,191123575	0,047027894	0,058847192
000257	0,183904879	0,039792833	0,031617182
000267	0,191467323	0,065115545	0,039265504
000270	0,103124231	0,010852591	0,012229642
000271	0,188717343	0,148318742	0,006431941
000273	0,133717753	0,014470121	0,007949364
000281	0,183561132	0,043410363	0,038934289
000284	0,187342353	0,018087651	0,016170017
000287	0,122374088	0,010852591	0,010419044
000295	0,192842312	0,043410363	0,036567367

Continuación Anexo No. 5, Valores R_N F_N M_N

Código Cliente	R_N	F_N	M_N
000299	0,17977991	0,028940242	0,015659333
000318	0,189404838	0,209816757	0,007105366
000319	0,19318606	0,220669347	0,010772748
000321	0,192842312	0,075968136	0,000725155
000331	0,043999672	0,007235061	0,002085452
000349	0,117905371	0,068733075	0,020629023
000350	0,112061665	0,039792833	0,015720988
000351	0,133717753	0,010852591	0,017599995
000356	0,172904961	0,021705182	0,011418623
000369	0,049843378	0,00361753	0,001664728
000370	0,185623616	0,047027894	0,029835389
000376	0,084218122	0,014470121	0,006774739
000379	0,170842476	0,065115545	0,045513535
000382	0,182186142	0,050645424	0,03464203
000392	0,189061091	0,036175303	0,036041115
000393	0,17977991	0,025322712	0,011558632
000395	0,169467487	0,007235061	0,003914997
000413	0,185623616	0,072350606	0,000682175
000422	0,162936285	0,018087651	0,017610798
000424	0,181842394	0,018087651	0,0088508
000425	0,071843214	0,010852591	0,002920616
000441	0,184248627	0,047027894	0,036540152
000443	0,185623616	0,057880485	0,029925969
000449	0,100374252	0,007235061	0,003053961
000464	0,15296761	0,014470121	0,002455191
000474	0,184592374	0,014470121	0,017867488
000475	0,173592456	0,007235061	0,00794576
000481	0,191123575	0,213434287	0,010536947
000490	0,175654941	0,021705182	0,035997718
000491	0,172904961	0,039792833	0,008684409
000498	0,184248627	0,043410363	0,032054348
000502	0,120655351	0,007235061	0,009577653
000504	0,13681148	0,010852591	0,007181945
000505	0,121686593	0,010852591	0,012808634
000510	0,120655351	0,00361753	0,001148234
000521	0,119280361	0,018087651	0,008549051
000526	0,186998606	0,119378499	0,000656523
000527	0,185967364	0,007235061	0,009772225
000533	0,175311193	0,039792833	0,022281771
000534	0,186998606	0,061498015	0,020350947
000539	0,145405166	0,010852591	0,008521476
000540	0,127530299	0,010852591	0,00623045
000559	0,188029848	0,036175303	0,04416576
000561	0,192842312	0,007235061	0,007471085

Continuación Anexo No. 5, Valores R_N' F_N' M_N'

Código Cliente	R_N'	F_N'	M_N'
000564	0,13956146	0,00361753	0,004343207
000566	0,172904961	0,018087651	0,009887251
000575	0,163967528	0,025322712	0,009902153
000585	0,19181107	0,054262954	0,000259471
000586	0,165686265	0,021705182	0,015603373
000588	0,19181107	0,014470121	0,01224475
000589	0,19043608	0,039792833	0,017315507
000600	0,158467569	0	0,002618648
000601	0,186654859	0,021705182	0,025304258
000607	0,192842312	0,061498015	0,025097535
000617	0,185623616	0,021705182	0,025104529
000619	0,17840492	0,014470121	0,007253606
000620	0,189748586	0,00361753	0,001689873
000621	0,180811152	0,014470121	0,005899774
000622	0,19318606	0,018087651	0,012484496
000623	0,192154818	0,050645424	0,018147192
000625	0,183217384	0,021705182	0,013030059
000626	0,183904879	0,050645424	0,011993019
000628	0,19043608	0,036175303	0,022398813
000629	0,185623616	0,014470121	0,007923697
000632	0,188029848	0,014470121	0,022377946
000637	0,192842312	0,014470121	0,005695961
000638	0,19318606	0,025322712	0,032411553
000639	0,144373924	0,018087651	0,024656784
000640	0,184592374	0,025322712	0,018042015
000642	0,185279869	0,036175303	0,025251575

Anexo No.6, Agrupación de los 158 Clientes a partir de K – Medias de Minitab

Código Cliente	R _N	F _N	M _N	Grupos K - Medias
000031	0,187342353	0,26046218	0,003559941	1
000058	0,186998606	0,14470121	0,083307883	
000060	0,183561132	0,14831874	0,00765351	
000271	0,188717343	0,14831874	0,006431941	
000318	0,189404838	0,20981676	0,007105366	
000319	0,19318606	0,22066935	0,010772748	
000481	0,191123575	0,21343429	0,010536947	
PROMEDIO	0,18861913	0,1922459	0,018481191	
000002	0,184592374	0,02170518	0,02013773	2
000008	0,186998606	0,02170518	0,024349536	
000014	0,185279869	0,02170518	0,024991333	
000035	0,166717507	0,00361753	0,00768353	
000037	0,190779828	0,02170518	0,00121973	
000056	0,180811152	0,01085259	0,020544639	
000065	0,185623616	0,02170518	0,022212598	
000083	0,183904879	0,0361753	0,002766805	
000089	0,178061173	0,01447012	0,011428651	
000096	0,174623698	0,02170518	0,00926008	
000098	0,188717343	0,01808765	0,016300571	
000118	0,171873719	0,00723506	0,007354648	
000139	0,184936121	0,02170518	0,013692188	
000158	0,176686183	0,01447012	0,021059541	
000171	0,186654859	0,02532271	0,022839089	
000172	0,173248709	0,03255777	0,02547768	
000241	0,174623698	0,01808765	0,01741449	
000284	0,187342353	0,01808765	0,016170017	
000299	0,17977991	0,02894024	0,015659333	
000356	0,172904961	0,02170518	0,011418623	
000393	0,17977991	0,02532271	0,011558632	
000395	0,169467487	0,00723506	0,003914997	
000424	0,181842394	0,01808765	0,0088508	
000474	0,184592374	0,01447012	0,017867488	
000475	0,173592456	0,00723506	0,00794576	
000490	0,175654941	0,02170518	0,035997718	
000491	0,172904961	0,03979283	0,008684409	
000527	0,185967364	0,00723506	0,009772225	
000561	0,192842312	0,00723506	0,007471085	
000566	0,172904961	0,01808765	0,009887251	
000586	0,165686265	0,02170518	0,015603373	
000588	0,19181107	0,01447012	0,01224475	
000601	0,186654859	0,02170518	0,025304258	
000617	0,185623616	0,02170518	0,025104529	
000619	0,17840492	0,01447012	0,007253606	
000620	0,189748586	0,00361753	0,001689873	

Anexo No.6, Agrupación de los 158 Clientes a partir de K – Medias de Minitab

Código Cliente	R _N '	F _N '	M _N '	Grupos K - Medias
000621	0,180811152	0,014470121	0,005899774	2
000622	0,19318606	0,018087651	0,012484496	
000625	0,183217384	0,021705182	0,013030059	
000629	0,185623616	0,014470121	0,007923697	
000632	0,188029848	0,014470121	0,022377946	
000637	0,192842312	0,014470121	0,005695961	
000638	0,19318606	0,025322712	0,032411553	
000640	0,184592374	0,025322712	0,018042015	
PROMEDIO	0,181889269	0,018498734	0,014522661	
000004	0,063593276	0,032557773	0,004313344	3
000023	0	0,094055787	0,00087533	
000036	0,062905781	0,00361753	0,008148013	
000040	0,070468225	0,108525909	0,00129785	
000055	0,060155802	0,039792833	0	
000184	0,081468143	0,075968136	0,000548909	
000230	0,06703075	0,010852591	0,006737166	
000246	0,08181189	0,115760969	0,001870962	
000331	0,043999672	0,007235061	0,002085452	
000369	0,049843378	0,00361753	0,001664728	
000425	0,071843214	0,010852591	0,002920616	
PROMEDIO	0,059374557	0,045712428	0,002769306	
000001	0,184936121	0,036175303	0,029168651	4
000016	0,19181107	0,065115545	0,034390009	
000027	0,191467323	0,057880485	0,046776214	
000039	0,188717343	0,050645424	0,024860511	
000045	0,190092333	0,039792833	0,016270373	
000063	0,186654859	0,068733075	0,069823059	
000074	0,190779828	0,050645424	0,016475758	
000078	0,147811398	0,061498015	0,027236002	
000088	0,17840492	0,061498015	0,040740068	
000099	0,191467323	0,047027894	0,021445206	
000105	0,19181107	0,065115545	0,025027715	
000149	0,161561296	0,047027894	0,021599408	
000173	0,162592538	0,065115545	0,021246539	
000242	0,184592374	0,054262954	0,019100871	
000255	0,187686101	0,083203197	0,063174033	
000256	0,191123575	0,047027894	0,058847192	
000257	0,183904879	0,039792833	0,031617182	
000267	0,191467323	0,065115545	0,039265504	
000281	0,183561132	0,043410363	0,038934289	
000295	0,192842312	0,043410363	0,036567367	
000370	0,185623616	0,047027894	0,029835389	
000379	0,170842476	0,065115545	0,045513535	
000382	0,182186142	0,050645424	0,03464203	

Anexo No.6, Agrupación de los 158 Clientes a partir de K – Medias de Minitab

Código Cliente	R _N	F _N	M _N	Grupos K - Medias
000392	0,189061091	0,036175303	0,036041115	4
000441	0,184248627	0,047027894	0,036540152	
000443	0,185623616	0,057880485	0,029925969	
000498	0,184248627	0,043410363	0,032054348	
000533	0,175311193	0,039792833	0,022281771	
000534	0,186998606	0,061498015	0,020350947	
000559	0,188029848	0,036175303	0,04416576	
000585	0,19181107	0,054262954	0,000259471	
000589	0,19043608	0,039792833	0,017315507	
000607	0,192842312	0,061498015	0,025097535	
000623	0,192154818	0,050645424	0,018147192	
000626	0,183904879	0,050645424	0,011993019	
000628	0,19043608	0,036175303	0,022398813	
000642	0,185279869	0,036175303	0,025251575	
PROMEDIO	0,184657407	0,051525364	0,030658921	
000009	0,192498565	0,723506057	0,016093524	5
000028	0,191467323	0,611362618	0,018027162	
000079	0,192154818	0,553482134	0,018845338	
000082	0,175311193	0,463043877	0,016740959	
000085	0,19318606	0,658390512	0,0349424	
000170	0,19043608	0,694565815	0,005153534	
PROMEDIO	0,189175673	0,617391835	0,018300486	
000012	0,117217876	0,018087651	0,011872352	6
000075	0,12478032	0,021705182	0,016164798	
000103	0,116530381	0,007235061	0,011731468	
000110	0,120999098	0,007235061	0,013245196	
000169	0,106905453	0,007235061	0,006950041	
000270	0,103124231	0,010852591	0,012229642	
000287	0,122374088	0,010852591	0,010419044	
000349	0,117905371	0,068733075	0,020629023	
000350	0,112061665	0,039792833	0,015720988	
000376	0,084218122	0,014470121	0,006774739	
000449	0,100374252	0,007235061	0,003053961	
000502	0,120655351	0,007235061	0,009577653	
000505	0,121686593	0,010852591	0,012808634	
000510	0,120655351	0,00361753	0,001148234	
000521	0,119280361	0,018087651	0,008549051	
000540	0,127530299	0,010852591	0,00623045	
PROMEDIO	0,114768676	0,016504982	0,01044408	
000013	0,137155228	0,007235061	0,008441958	7
000024	0,163280033	0,021705182	0,000531087	
000050	0,157780074	0,043410363	0,013433552	
000095	0,140248955	0,010852591	0,024579242	
000097	0,146092661	0,014470121	0,008606355	

Anexo No.6, Agrupación de los 158 Clientes a partir de K – Medias de Minitab

Código Cliente	R _N '	F _N '	M _N '	Grupos K - Medias
000100	0,155717589	0,021705182	0,009433278	7
000102	0,143686429	0,014470121	0,011912566	
000111	0,145061419	0,007235061	0,008600016	
000141	0,147123903	0,039792833	0,022683644	
000247	0,145061419	0,047027894	0,023561913	
000273	0,133717753	0,014470121	0,007949364	
000351	0,133717753	0,010852591	0,017599995	
000422	0,162936285	0,018087651	0,017610798	
000464	0,15296761	0,014470121	0,002455191	
000504	0,13681148	0,010852591	0,007181945	
000539	0,145405166	0,010852591	0,008521476	
000564	0,13956146	0,00361753	0,004343207	
000575	0,163967528	0,025322712	0,009902153	
000600	0,158467569	0	0,002618648	
000639	0,144373924	0,018087651	0,024656784	
PROMEDIO	0,147656712	0,017725898	0,011731159	
000018	0,192498565	0,079585666	0,006036016	8
000034	0,19318606	0,072350606	0,00312537	
000042	0,183561132	0,086820727	0,005852624	
000043	0,191467323	0,12661356	0,058520761	
000062	0,184592374	0,061498015	0,000382922	
000066	0,172904961	0,068733075	0,005440938	
000069	0,184592374	0,112143439	0,05459155	
000072	0,186998606	0,104908378	0,002379286	
000090	0,191467323	0,104908378	0,0516751	
000101	0,185623616	0,079585666	0,030294994	
000140	0,146780156	0,068733075	0,002956986	
000168	0,189404838	0,065115545	0,000399924	
000190	0,141623944	0,12299603	0,000782912	
000253	0,192842312	0,075968136	0,000550345	
000321	0,192842312	0,075968136	0,000725155	
000413	0,185623616	0,072350606	0,000682175	
000526	0,186998606	0,119378499	0,000656523	
PROMEDIO	0,182529889	0,088097502	0,013238446	