



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

**VICERRECTORADO DE INVESTIGACIÓN
INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA
CENTRO DE POSGRADOS**

**PROGRAMA DE MAESTRÍA EN GESTIÓN DE SISTEMAS DE
INFORMACIÓN E INTELIGENCIA DE NEGOCIOS**

**TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO
DE MAGISTER EN GESTIÓN EN SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS**

**TEMA “OPTIMIZACIÓN DEL PROCESO DE ADQUISICIÓN DE
ARTÍCULOS PARA TIENDAS DEPORTIVAS EN ECUADOR A TRAVÉS
DE TÉCNICAS DE DATA MINING PARA EL DESCUBRIMIENTO DE
PATRONES EN LA VENTA DE PRODUCTOS”**

AUTOR: DÍAZ BALDEÓN ALICIA ELIZABETH

DIRECTOR: MOLINA BUSTAMANTE MARCO

SANGOLQUÍ

2017

CERTIFICADO

ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN
INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

PROGRAMA DE MAESTRÍA EN GESTIÓN DE SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS

CERTIFICACIÓN

Certifico que el trabajo de titulación, “OPTIMIZACIÓN DEL PROCESO DE ADQUISICIÓN DE ARTÍCULOS PARA TIENDAS DEPORTIVAS EN ECUADOR A TRAVÉS DE TÉCNICAS DE DATA MINING PARA EL DESCUBRIMIENTO DE PATRONES EN LA VENTA DE PRODUCTOS” realizado por la señora ALICIA ELIZABETH DÍAZ BALDEÓN, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a la señora ALICIA ELIZABETH DÍAZ BALDEÓN para que lo sustente públicamente.

Sangolquí, septiembre del 2017

Una firma manuscrita en tinta azul que parece leerse 'M. Molina B.' sobre una línea horizontal punteada.

MARCO MOLINA BUSTAMANTE
DIRECTOR

AUTORÍA DE RESPONSABILIDAD



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN
INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

PROGRAMA DE MAESTRÍA EN GESTIÓN DE SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS

AUTORÍA DE RESPONSABILIDAD

Yo, **ALICIA ELIZABETH DÍAZ BALDEÓN**, con cédula de identidad N° **1003230917**, declaro que este trabajo de titulación **“OPTIMIZACIÓN DEL PROCESO DE ADQUISICIÓN DE ARTÍCULOS PARA TIENDAS DEPORTIVAS EN ECUADOR A TRAVÉS DE TÉCNICAS DE DATA MINING PARA EL DESCUBRIMIENTO DE PATRONES EN LA VENTA DE PRODUCTOS”** ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, septiembre del 2017

Una firma manuscrita en tinta azul que parece decir 'Alicia Díaz Baldeón', escrita sobre una línea horizontal punteada.

ALICIA ELIZABETH DÍAZ BALDEÓN
C.C 1003230917

AUTORIZACIÓN BIBLIOTECA VIRTUAL



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN
INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

PROGRAMA DE MAESTRÍA EN GESTIÓN DE SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS

AUTORIZACIÓN

Yo, **ALICIA ELIZABETH DÍAZ BALDEÓN**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación “**OPTIMIZACIÓN DEL PROCESO DE ADQUISICIÓN DE ARTÍCULOS PARA TIENDAS DEPORTIVAS EN ECUADOR A TRAVÉS DE TÉCNICAS DE DATA MINING PARA EL DESCUBRIMIENTO DE PATRONES EN LA VENTA DE PRODUCTOS**” cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, septiembre del 2017

Una firma manuscrita en tinta azul que parece decir 'Alicia Díaz Baldeón', escrita sobre una línea horizontal punteada.

ALICIA ELIZABETH DÍAZ BALDEÓN
C.C 1003230917

DEDICATORIA

Amor esta meta te la dedico a ti, a la luz de mis ojos,
al por quien de mi felicidad, a mi fuerza, a mi apoyo,
al pilar fundamental de mi vida, a mi mejor amigo y mi amor eterno.

A mis padres por ser mi aliento, por impulsarme
y acompañarme en cada paso, por su gran ejemplo y su amor incondicional.

Alicia

AGRADECIMIENTO

A Dios, porque ha guiado cada paso en mi vida,
y me ha permitido alcanzar cada sueño.

A mi esposo por todo tu apoyo en cada momento,
porque nunca me dejaste rendirme y siempre has tenido las palabras perfectas
para impulsarme, motivarme
y demostrarme tu amor y que siempre caminaremos juntos.

A mis padres porque transforman mis metas y mis sueños en los suyos,

A mis mejores amigos, mis hermanos.

Alicia

ÍNDICE

CARÁTULA	II
CERTIFICADO	II
AUTORÍA DE RESPONSABILIDAD.....	III
AUTORIZACIÓN BIIBLIOTECA VIRTUAL	IV
DEDICATORIA.....	V
AGRADECIMIENTO	VI
ÍNDICE	VII
ÍNDICE DE TABLAS.....	XII
ÍNDICE DE FIGURAS.....	XIII
RESUMEN.....	XV
ABSTRACT.....	XVI
CAPÍTULO I	
INTRODUCCIÓN	1
CAPÍTULO 2	
EL PROBLEMA	5

2.1 ANTECEDENTES	5
2.2 PLANTEAMIENTO DEL PROBLEMA.....	6
2.3 FORMULACIÓN DEL PROBLEMA.....	7
2.4 JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	7
2.5 OBJETIVOS DE LA INVESTIGACIÓN	7
□ OBJETIVO GENERAL	7
□ OBJETIVOS ESPECÍFICOS.....	8
2.6 PROPOSICIÓN	8

CAPÍTULO 3

MARCO REFERENCIAL..... 9

3.1 MARCO TEÓRICO	9
3.1.1 MARKETING.....	9
3.1.1.1 Conceptos de Marketing.....	12
3.1.1.2 Mercado y Segmentación de mercados.....	15
3.1.1.2.1 Mercado	15
3.1.1.2.2 Segmentación de Mercados	16
3.1.1.3 Comportamiento del Consumidor.....	23
3.1.1.3.1 Proceso de Compra	26
3.1.1.3.2 Factores influyen en el estilo de vida.....	27
3.1.1.3.3 Comportamiento de compra del Consumidor Organizacional.....	30
3.1.1.4 Estrategia de producto.....	30
3.1.1.4.1 Clasificación de los productos	31
3.1.1.4.2 Posicionamiento de un producto.....	32
3.1.1.4.3 Línea y mezcla de productos.....	33
3.1.1.4.3.1 Decisiones sobre la Línea de producto	33
3.1.1.4.3.2 Factores que influyen en los cambios de la mezcla de productos.....	34
3.1.1.4.3.3 Portafolio de Productos.....	34

3.1.1.4.3.4 Ciclo de vida del producto	36
3.1.1.5 Estrategia de Precio.....	37
3.1.1.6 Estrategia de Distribución.....	38
3.1.1.6.1 Intermediarios	39
3.1.1.6.2 Distribución física del producto.....	40
3.1.1.6.3 Elementos del sistema de distribución física	40
3.1.1.6.4 Objetivos Secundarios de la distribución física.....	41
3.1.1.7 Ventas en Retail	42
3.1.1.7.1 Estrategia de la venta de retail	45
3.1.1.7.2 Planeación venta de retail	47
3.1.2 INTELIGENCIA DE NEGOCIOS	48
3.1.2.1 Desarrollo de Inteligencia de Negocios	51
3.1.2.2 Data Warehouse	54
3.1.3 MINERÍA DE DATOS	56
3.1.3.1 Tipos de Datos.....	58
3.1.3.2 Proceso de descubrimiento de conocimiento en bases de datos	59
3.1.3.3 Aplicaciones.....	61
3.1.3.4 Sistemas y herramientas de minería de datos.....	63
3.1.3.5 Tipos de modelos	64
3.1.3.5.1 Reglas de Asociación y Dependencia	65
3.1.3.5.1.1 Reglas de asociación.....	68
3.1.3.5.1.2 Reglas de dependencias	69
3.1.3.5.2 Árboles de Decisión.....	71
3.1.3.5.3 Agrupación de datos o Clustering.....	73
3.1.3.5.4 Clasificación de Datos	74
3.2 TRABAJOS RELACIONADOS	74

CAPÍTULO 4

MARCO METODOLÓGICO 86

4.1 METODOLOGÍA.....	86
4.1.1 COMPRENSIÓN DEL NEGOCIO	87
4.1.2 COMPRENSIÓN DE LOS DATOS	88
4.1.3 PREPARACIÓN DE LOS DATOS	90
4.1.4 MÉTODOS Y HERRAMIENTAS	90
4.1.4.1 Tarea de Visualización.....	91
4.1.4.1 Tarea de Clasificación.....	92
4.1.4.2 Tarea de Clustering	93
4.1.4.3 Tarea de Asociación	95
4.1.5 RESULTADOS.....	96

CAPÍTULO 5

RESULTADOS 97

5.1 TAREA DE VISUALIZACIÓN.....	97
5.2 TAREA DE ASOCIACIÓN.....	127

CAPÍTULO 6

CONCLUSIONES Y RECOMENDACIONES..... 139

6.1 CONCLUSIONES.....	139
6.2 RECOMENDACIONES	142

CAPÍTULO 7

BIBLIOGRAFÍA..... 143

ÍNDICE DE TABLAS

Tabla 1 Índices de venta y liquidación de la empresa	6
Tabla 2 Cesta de la compra / ejemplo.....	66
Tabla 3 Ejemplo de dependencias.....	70
Tabla 4 Campos y descripción de los datos	89
Tabla 5 Análisis de cruces de visualización	92

ÍNDICE DE FIGURAS

Figura 1 Conceptos centrales de Marketing.....	10
Figura 2 Criterios para segmentar los mercados.....	18
Figura 3 Bases para la segmentación de mercado y variables seleccionadas	19
Figura 4 Clasificación de tiendas minoristas	45
Figura 5 Arquitectura general de soluciones de Business Intelligence	48
Figura 6 Estructura de Data Warehouse	56
Figura 7 Proceso de KDD Fuente	59
Figura 8 Metodología aplicada	86
Figura 9 Resultados tarea de Clasificación.....	94
Figura 10 Resultados Tarea de Asociación.....	95
Figura 11 Contraste Forma de pago / Mes.....	97
Figura 12 Contraste Forma de pago / Región.....	98
Figura 13 Contraste Grupo / Segmentación.....	98
Figura 14 Contraste Actividad / Segmentación	99
Figura 15 Contraste Tienda / Segmentación.....	100
Figura 16 Contraste Tienda / Año.....	100
Figura 17 Contraste Grupo / Género edad	101
Figura 18 Contraste Grupo / Región.....	101
Figura 19 Contraste Grupo / Tienda	102
Figura 20 Contraste Actividad / Grupo.....	103
Figura 21 Contraste Género edad / Actividad.....	103
Figura 22 Contraste Región / Actividad	104
Figura 23 Contraste Tienda / Actividad.....	105
Figura 24 Contraste Línea / Actividad.....	106
Figura 25 Contraste Género edad / Línea	107
Figura 26 Contraste Región / Línea	107
Figura 27 Contraste Tienda / Línea	108
Figura 28 Contraste Tienda / Género edad	110

Figura 29 Contraste Región / Mes	111
Figura 30 Contraste Género edad / Línea	112
Figura 31 Contraste Grupo / Línea	112
Figura 32 Contraste Grupo / Género edad, Tienda: City Mall	113
Figura 33 Contraste Género edad / Talla, Tienda: City Mall	114
Figura 34 Contraste Grupo / Género edad, Tienda: Condado Shopping	115
Figura 35 Contraste Género edad / Talla, Tienda: Condado Shopping	115
Figura 36 Contraste Grupo / Género edad, Tienda: Mall del Río.....	116
Figura 37 Contraste Género edad / Talla, Tienda: Mall del Río.....	117
Figura 38 Contraste Grupo / Género edad, Tienda: Mall el Jardín.....	117
Figura 39 Contraste Género edad / Talla, Tienda: Mall el Jardín.....	118
Figura 40 Contraste Grupo / Género edad, Tienda: Scala Shopping.....	119
Figura 41 Contraste Género edad / Talla, Tienda: Scala Shopping	119
Figura 42 Contraste Grupo / Género edad, Tienda: Mall del Sol	120
Figura 43 Contraste Género edad / Talla, Tienda: Mall del Sol	121
Figura 44 Contraste Grupo / Género edad, Tienda: Village Plaza	121
Figura 45 Contraste Género edad / Talla, Tienda: Village Plaza	122
Figura 46 Reglas de Asociación City Mall	129
Figura 47 Reglas de Asociación Condado Shopping.....	130
Figura 48 Reglas de Asociación Mall del Río	132
Figura 49 Reglas de Asociación Mall el Jardín	134
Figura 50 Reglas de Asociación Scala Shopping	135
Figura 51 Reglas de Asociación Mall del Sol.....	136
Figura 52 Reglas de Asociación Village Plaza	138

RESUMEN

El proceso acelerado de globalización ha conllevado una rápida expansión de lo que se ha denominado *Business Intelligence* (BI) cuyo propósito es facilitar el acceso al análisis de la información para la toma de decisiones, a partir de un proceso de transformación de los datos empresariales en conocimiento. En el presente estudio, se incursiona en el uso del BI para resolver un problema frecuente de las empresas comercializadoras, altos niveles de liquidación, se aborda el problema de una empresa multinacional con sede local. En el estudio se utilizaron técnicas de *data mining*, las que permiten hacer un análisis profundo de los datos y obtener conocimiento. El problema de liquidación ocurren debido a que para reposición de los artículos no se toman en cuenta variables importantes como son: el mix de producto, rotación de ítems, perfil de consumidor, etc. En el estudio se desarrolló la siguiente metodología, que parte del entendimiento del funcionamiento de la empresa y de sus necesidades, el entendimiento de los datos, consistentes en las ventas de 4 años divididas en 7 tiendas a nivel nacional, depuración y transformación de los datos, para su posterior uso en el siguiente paso que fue la aplicación de las técnicas de *data mining*. Al final de este trabajo se logró encontrar el mix de producto adecuado para cada tienda, y las necesidades de productos acorde con la tienda, finalmente, se determinó el perfil del consumidor por zona, a la vez como consecuencia, se descubrió información importante para focalización de campañas y visual merchandising.

PALABRAS CLAVE:

- Inteligencia de Negocios
- Minería de Datos
- Patrones de Consumo

ABSTRACT

The accelerated process of globalization has led to a rapid expansion of what has been called Business Intelligence (BI), whose purpose is to facilitate access to the analysis of information for decision making, starting from a process of transformation of business data in knowledge. In the present study, it into the use of BI to solve a frequent problem of trading companies, high levels of settlement, it addresses the problem of a multinational company with local headquarters. In the study, data mining techniques were used, which allow an in-depth analysis of the data and obtain knowledge. The liquidation problem occurs due to the fact that, for the replacement of the items, important variables such as product mix, item rotation, consumer profile, etc. are not taken into account. The study developed the following methodology, based on the understanding of the operation of the company and its needs, the understanding of the data, consisting of sales of 4 years divided into 7 stores nationwide, debugging and transformation of data, for later use in the next step was the application of data mining techniques. At the end of this work, it was possible to find the right product mix for each store, and the product needs according to the store, finally, the profile of the consumer was determined by zone, and as a consequence, important information was discovered for targeting campaigns and visual merchandising.

KEY WORD:

- Business Intelligence
- Data Mining
- Consumer Profile

CAPÍTULO I

INTRODUCCIÓN

El siglo XXI, es un siglo de grandes retos para las empresas, ya que fenómenos como la globalización que es un proceso dinámico de libertad e integración mundial de los mercados de trabajo, bienes, servicios, tecnología y capitales; han exigido la competitividad, porque ahora nos encontramos con mercados que actúan, son más críticos, tienen a su alcance herramientas de gran poder sin complicaciones en su uso, es un mercado que tiene más accesos a la información que no está dispuesto a pagar dinero por cualquier producto, la globalización se orienta al cliente, olvida los monopolios y oligopolios, se extiende para abrir los mercados y competir con todo el mundo; no es necesario estar ahí físicamente para lograrlo, tampoco es necesario tener un corporativo en Wall Street, en Dubái o en Londres, porque ahora se puede operar a una empresa desde un computador en la comodidad de un hogar o a través de un GPS en cualquier momento.

El futuro de las ventas en retail se desempeña cada vez más en un entorno competitivo, el cual incluirá costos menores y márgenes más bajos, el uso de nueva tecnología, tiendas con menos cantidad de inventario, un enfoque hacia el valor del cliente a largo plazo y un mercado global, en el siglo XXI, la supervivencia podría depender de ser ahorrativos y eficientes en todos los aspectos de las operaciones, incluyendo la superación de la competencia global en la utilización de tecnología para reducir costos y mejorar de manera evidente los niveles del servicio del consumidor, se puede concluir que la tecnología será la herramienta usada para ayudar a los empresarios a cumplir los objetivos.

Las empresas deben pensar en grande, obligarse a ser productivos, con el objetivo de desarrollar ventajas competitivas para conseguir una participación de mercado sustentable y de crecimiento, por lo cual deben analizar cada uno de sus procesos y optimizarlos, varios autores determinan que uno de los secretos de las grandes empresas es el “Customer Relation ship Management (CRM)”, o “Estrategia de negocios centrada en el cliente”; ya que no existe empresa que no trabaje para sus clientes.

Vivimos en una época en que la información es la clave para obtener una ventaja competitiva en el mundo de los negocios, jugando un papel relevante ya que a partir de ella se genera conocimiento, y así la toma de decisiones pueda ser oportuna, consistente, rápida y fácil, por eso se dice que algo peor que no tener información disponible es tener mucha información y no saber qué hacer con ella. La Inteligencia de Negocios o Business Intelligence (BI) es una posible solución al problema, ya que por medio de dicha información se podrían generar escenarios, pronósticos y reportes que apoyen a la toma de decisiones, siendo este uno de sus mayores beneficios, ya que trae con estas decisiones el crecimiento o pérdidas en un negocio, es la información en el lugar y momento justo.

Existe una gran variedad de software de BI con aplicaciones que son utilizadas en diferentes áreas de la empresa, tales como, ventas, marketing, finanzas, etc.

Uno de los componentes de BI, es Minería de Datos o Data Mining, una técnica que consiste en la extracción de información que reside de manera implícita en los datos, con el objetivo de encontrar patrones repetitivos que no sean evidentes, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto, que aporte información y conocimiento para la toma de decisiones. (Hernández, J; Ramírez, M; Ferri, C; 2008)

El desarrollo de data mining en los últimos años ha sido acelerado por los varios beneficios que presenta, y la diversidad en los campos de aplicación; gobierno, económico, financiero, marketing, medico, militar, comunicaciones, etc.

En el caso del proceso de adquisición de productos, es un proceso complejo que genera una gran cantidad de datos, que se transforma en información, ya que involucran varios factores de análisis, como son: ubicación geográfica, condiciones climáticas, situación socio cultural, ingreso socio económico, tendencias actuales, ventas pasadas, historial de compra y descripción de características de producto, variables de tiempo, etc. El proceso de adquisición de artículos para tiendas deportivas en Ecuador, necesita de información clara y oportuna, existen problemas como

el desconocimiento del mix de productos de mayor rotación, y tampoco se ha determinado el perfil del consumidor por cada zona, es decir no se está enfocando este proceso en las necesidades del cliente ni en el seguimiento de compra del mismo.

El presente trabajo se ha motivado en las empresas dedicadas a la importación de artículos deportivos en el Ecuador generen una ventaja competitiva en el mercado mediante la optimización en el proceso de adquisición de artículos para tiendas deportivas a nivel país a través de técnicas de data mining para el descubrimiento de patrones en la venta de productos, este estudio se realizará en las ciudades de Quito, Guayaquil y Cuenca, dentro de estas ciudades se sectorizará para mejores resultados del estudio.

Por lo tanto, el objetivo es; optimizar el proceso de adquisición de artículos en tiendas deportivas de Ecuador, descubriendo el mix de mayor rotación de producto y el perfil del consumidor por zona a través de técnicas de data mining.

Se cree que efectivizar el proceso de adquisición de productos traerá como beneficios, un cargue en punto de venta óptimo, una rotación de inventario rápida siendo máximo de liquidación del 25%, así se podrá optimizar recursos de varios tipos como son: financieros, tiempo, espacio, operacionales.

Como se mencionó en los párrafos anteriores el cliente debe ser el foco de toda empresa, se piensa que con la optimización del proceso de compra también se dará un mejor servicio y se generará satisfacción en el consumidor, así se le entregará el mix de producto que busca, en lugar y momento que lo necesita, a la vez tendrá una experiencia de costo beneficio positiva, lo que trae consigo fidelización, recompra, repercutiendo en el nivel de ventas y aumentando la participación de mercado, es decir que los ingresos de la compañía crecerán y serán sostenibles en el tiempo al igual que el poder de marca.

Determinar el perfil del consumidor por zona, es un beneficio que no solo repercute en llegar al cliente exacto, en el momento correcto y con la oferta adecuada, también permite que la

comunicación y acciones de marketing y publicidad tengan el enfoque correcto en el segmento elegido, ya que se sabrá con exactitud a quien dirigir campañas, en qué lugares, en que momentos, con cuáles medios, el producto y las herramientas más efectivas.

En contexto se cree que se conseguirá ventaja competitiva para la empresa, para este estudio se levantará información de las ventas en las ciudades mencionadas, y se aplicará la metodología de data mining para encontrar los patrones de consumo.

El presente trabajo está compuesto por 6 capítulos; el capítulo 1 es la introducción al tema de estudio y presentación del estudio realizado; el capítulo 2 contiene el planteamiento del problema, donde expondremos la problemática sobre la base de datos de la industria del comercio y se plantean los objetivos y la hipótesis. En el capítulo 3 se aborda el marco teórico que trata tres puntos claves: marketing, inteligencia de negocios y data mining. En este capítulo también haremos referencia a trabajos ya realizados en el ámbito que nos ocupa. El capítulo 4 trata de la metodología utilizada, explicando cada paso que se realizó y los resultados obtenidos. En el capítulo 5 se exponen los resultados encontrados y su análisis, finalmente, el capítulo 6 detalla las conclusiones y recomendaciones, que responden a los objetivos planteados.

CAPÍTULO 2

EL PROBLEMA

En el presente capítulo se detalla el problema partiendo desde sus antecedentes, el planteamiento y la formulación del problema, con ello se justificará la investigación y se plantearán los objetivos y la proposición.

2.1 Antecedentes

Business Intelligence es un concepto y una necesidad perenne que está mostrando importantes avances. Disponer de un efectivo y ágil sistema de información gerencial que procese los datos y los convierta en información se ha transformado en una necesidad dentro de las organizaciones. La brecha entre los datos y el empleo de nuevas y eficientes tecnologías para la toma de mejores decisiones se está cerrando hacia su utilización universal.

El emprendimiento de un proyecto de Business Intelligence en una empresa será exitoso si los funcionarios de toda área o departamento pueden acceder e interactuar con su información. Con las nuevas tendencias de Business Intelligence esto será posible, a la vez esta interacción y colaboración de las distintas áreas se refleja en el éxito de proyectos de Data mining.

Las implementaciones de sistemas tradicionales se han encontrado con importantes inconvenientes, tales como: el manejo de grandes volúmenes de información y la existencia de altos tiempos de respuesta, integración de diversas fuentes de datos, alta dependencia de los expertos en tecnología. En definitiva, estas limitaciones se han traducido en altos costos (software y hardware, tiempo de consultoría, mantenimiento) y extenso tiempo en su implementación.

Data mining, como herramienta de la inteligencia de negocios, ha ganado espacio y prácticamente se está tornando indispensable en los proyectos relacionados en este dominio.

El desarrollo de data mining en Ecuador, como en el resto de países de América Latina es reducido, por ejemplo, en Estados Unidos su sistema de sanidad usa técnicas de minería de datos,

nuestros países se apoyan aún en la estadística tradicional para resolver problemas de información en sus sistemas sanitarios.

Otro de los problemas suscitados con el desarrollo de data mining, es que las empresas son muy celosas de sus datos y de los resultados en data mining, dificultando el avance en el área de aplicación de data mining en Business Intelligence.

2.2 Planteamiento del problema

Data mining está en pleno auge y aún mucha gente no es consciente de la importancia que tiene, en Ecuador no ha sido desarrollado, y un claro ejemplo se puede identificar en la industria de comercialización de bienes, donde encontramos constantemente liquidaciones de productos en distintas formas, ya que no se mantiene el stock adecuado de acuerdo a los niveles de compra, relacionado directamente con el mix de producto en las tiendas, se tiene un sobre stock y se debe liquidar el inventario para poder renovarlos, al igual la tienda no tiene relacionados sus productos en la mayoría de ocasiones, lo que no permite tener un ticket promedio alto.

De acuerdo a la información proporcionada por una empresa dedicada a la comercialización de productos deportivos para mantener un stock adecuado se debe vender en 6 meses el 75% de la mercadería con el propósito de liquidar el 25%, pero los índices de liquidación que esta compañía ha tenido en su canal de retail son los siguientes:

Tabla 1
Índices de venta y liquidación de la empresa

Temporada	Detalle	% de venta	% de liquidación
SS16	Enero-Junio 2016	43%	57%
FW16	Julio-Diciembre 2016	50%	50%
SS17	Enero-Junio 2017	42%	58%

Adicional a esta información se nos compartió que en la temporada julio a diciembre 2016 la liquidación fue de más de 70.000 unidades y de enero a junio del 2017 fue de 60.000, en esta última liquidación el 77% de esta mercadería entró con descuentos superiores al 40% durante 2 meses y lograron limpiar el 85% del inventario.

2.3 Formulación del Problema

- ¿Cuáles son los mix de productos de mayor rotación en cada tienda?
- ¿Cuál es el perfil del consumidor por tienda y ciudad?

El problema en la adquisición de artículos en las tiendas deportivas en Ecuador, es que se desconoce cuál es el mix de mayor rotación entre ellos por tienda, a la vez no se tiene determinado el perfil del consumidor por cada zona.

2.4 Justificación de la investigación

El problema de desconocimiento del perfil del consumidor ocasiona ofrecer al cliente un mix de producto que no es atractivo para concluir en una venta, ya que en todo el proceso desde la compra, visual merchandising, tallas, colores, rango de precios, campañas, son inadecuados, y estos mismos atributos nombrados son factores determinantes en la decisión de compra, por lo que el no mantener un sistema que permita manejar los datos de ventas y generar información para la toma de decisiones en el proceso de compra repercute directamente en las ventas, en las ganancias y en el posicionamiento en el cliente.

El generar un perfil del consumidor por cada tienda, ya que cada tienda ocupa una zona diferente tanto en lo geográfico, como en social y cultural, permitirá el enfoque en la compra sobre datos reales en proporciones, y al conocer el mix de producto adecuado permitirá colocar en punto de venta de forma adecuada, lo que generará deseo de compra.

2.5 Objetivos de la investigación

- **Objetivo General**

Optimizar el proceso de adquisición de artículos en tiendas deportivas de Ecuador, determinando el mix de mayor rotación de los productos por cada tienda y al relacionarlo con el perfil del consumidor por zona a través de técnicas de data mining.

- **Objetivos Específicos**

- Descubrir el mix de productos de mayor rotación en cada tienda, para identificar que líneas de ítems deben ser comprados juntos y en que proporciones.
- Definir el perfil del consumidor por tienda y ciudad, con el fin de formar clusters por región de acuerdo a los comportamientos encontrados.
- Definir los meses de mayor venta en el año por tienda para programar de forma efectiva los ingresos mensuales.

2.6 Proposición

El descubrimiento de los patrones de ventas acorde al perfil del consumidor, nos permitirá encontrar el mix adecuado para cada punto de venta.

CAPÍTULO 3

MARCO REFERENCIAL

En el presente capítulo se desarrolla el marco teórico y trabajos relacionados. El marco teórico estará basado en tres pilares; 1) Marketing en esta sección hablaremos de los conceptos básicos de esta ciencia y de los temas relacionados con este trabajo como son producto, mercado, segmentación, canales de distribución y proceso de compra; 2) Inteligencia de negocios, de la misma forma trataremos sus conceptos y aplicaciones; 3) Data Mining donde se detalla sus diferentes técnicas. En este capítulo también se trata de trabajos relacionados con el tema de estudio, sus aplicaciones y resultados, de los cuales se está tomando como base para el desarrollo de este estudio.

3.1 Marco Teórico

3.1.1 Marketing

Al hablar de marketing debemos entender que su fin es entablar y gestionar relaciones rentables con los clientes mediante una oferta de valor superior frente a la competencia, es orientar a una empresa totalmente hacia el cliente, es decir se gestiona relaciones rentables con ellos.

Por lo cual las empresas que entienden el concepto de marketing y lo colocan como el corazón de las mismas, comparten la pasión por atender y satisfacer las necesidades de los clientes en mercados objetivos bien definidos, relaciones duraderas, a través de un valor para ellos y de una satisfacción superior.

A continuación, se presentan algunas definiciones de marketing:

El marketing, más que cualquier otra función empresarial, se ocupa de los clientes. La creación de relaciones con clientes basadas en la satisfacción y en el valor para el cliente es lo que conforma el núcleo del marketing moderno (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

La American Marketing Association ha definido el marketing como “el proceso de planear y ejecutar el concepto, el precio, la promoción y la distribución de ideas, bienes y servicios para

crear intercambios que satisfagan objetivos individuales y organizacionales” (Hasty, R; Reardon, J; 1998)

“La gestión del marketing se define como el arte y la ciencia de seleccionar mercados objetivos y de crear relaciones rentables con sus agentes. La gestión de marketing debe encontrar el modo de hacer frente a los diferentes estados de la demanda” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

De acuerdo a los conceptos referenciados sobre marketing, se lo puede definir como el proceso mediante el cual grupos e individuos obtienen lo que necesitan y desean a través de la creación de productos y del intercambio con otros, se desarrollan actividades tales como el desarrollo de productos, la investigación, la comunicación, la distribución, la fijación de precios y el servicio son actividades fundamentales del marketing.

Entender el marketing como la satisfacción de las necesidades del cliente, es el punto más importante y de partida, viéndolo desde forma, todo empieza en la identificación de las necesidades de los clientes, luego se desarrollan productos que ofrezcan un valor superior, seguido se fijan los precios, se continua con la distribución de los productos y finalmente se promocionan de forma eficaz, así lo muestra la figura 1.

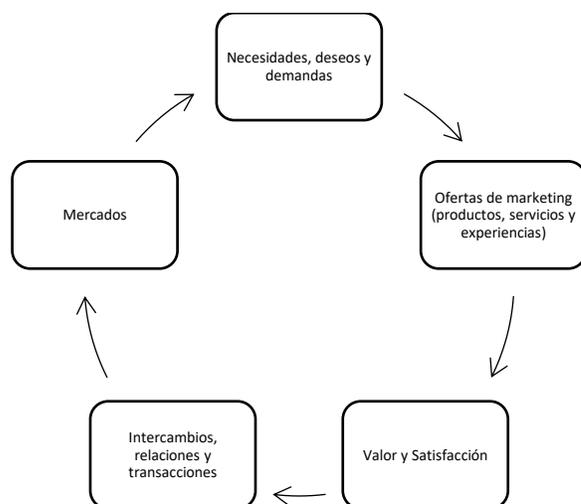


Figura 1 Conceptos centrales de Marketing
(Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Para entender a cabalidad el concepto de Marketing y el proceso expuesto en la figura 1, se debe conceptualizar necesidad y deseo.

- **Necesidad**

“Las necesidades humanas son estados de carencia. Estas incluyen las necesidades *físicas* de alimentación, vestido, calor y seguridad; las necesidades *sociales* de afecto y de pertenencia a un grupo; y las necesidades *individuales* de conocimiento y de expresión personal” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

- **Deseo**

Los deseos son las formas que adoptan las necesidades humanas una vez determinadas por la cultura y la personalidad del individuo. Los deseos vienen determinados por la sociedad a la que pertenece y se describen como los objetos que satisfacen esas necesidades. Cuando los deseos vienen acompañados por una capacidad de adquisición concreta se convierte en demandas (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Las empresas de marketing más exitosas dedican importantes esfuerzos a identificar y comprender las necesidades, los deseos y las demandas de sus clientes. Llevan a cabo investigaciones de mercado y analizan un sinnúmero de informaciones provenientes de ventas, devoluciones y servicios.

Las empresas abordan las necesidades de los consumidores mediante una propuesta de valor, es decir, prometen a los consumidores un conjunto de beneficios que satisfará sus necesidades. La propuesta de valor se materializa mediante una oferta de marketing, una combinación de productos, servicios, información o experiencias que se ofrecen en un mercado para satisfacer una necesidad o un deseo.

Los consumidores escogen entre tantas ofertas de marketing en base a sus percepciones sobre el valor y sobre la satisfacción que le ofrecen los diversos productos y servicios. El valor para el cliente es la diferencia entre los valores que el cliente recibe por poseer y utilizar un producto y el coste de obtención de ese producto. Los consumidores se forman estas expectativas basándose en experiencias de compra pasadas, en opiniones de amigos, en la información y las promesas del vendedor y de la competencia.

El nivel de *satisfacción del cliente* tras una compra depende de la medida en que los resultados del producto cumplan las expectativas del consumidor. La satisfacción del cliente ejerce una influencia primordial en el comportamiento de compra futuro. Los clientes satisfechos volverán a comprar y contarán a otros sus experiencias positivas. Los clientes insatisfechos cambiarán a la competencia y menospreciarán el producto delante de otros. El valor para el cliente y el nivel de satisfacción son pilares fundamentales para el desarrollo y la gestión de las relaciones con el cliente (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

3.1.1.1 Conceptos de Marketing

- Intercambio, transacciones y relaciones

De acuerdo a (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004) se puede definir al intercambio como acto de obtener un objeto deseado de otra persona, ofreciéndole algo a cambio y mientras que una transacción consiste en un canje de valores entre dos partes.

Según (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004) “Un mercado es el conjunto de compradores reales y potenciales de un producto. Estos compradores comparten una necesidad o un deseo particular que puede satisfacerse mediante una relación de intercambio”

En base a los conceptos leídos se entiende que el tamaño de un mercado depende del número de personas que comparten la misma necesidad, que disponen de recursos necesarios que ofrezcan a cambio y que están dispuestas a entregar dichos recursos para obtener lo que desean, así se menciona en (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004) que los especialistas del marketing trabajan para comprender las necesidades y los deseos de mercados específicos, para escoger los mercados en los que pueden ser de utilidad.

- **Gestión de relaciones con el cliente**

La gestión de relaciones con el cliente (CRM) se ha definido de forma limitada como la acción de gestionar la base de datos de clientes.

“La gestión de relaciones con el cliente se define hoy como el proceso global de construcción y conservación de relaciones rentables con los clientes, mediante la entrega de un valor superior y de una mayor satisfacción” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Las empresas que emplean la gestión de relaciones con el cliente las utilizan para conservar a los clientes que tienen y desarrollar relaciones rentables y duraderas con ellos. Las empresas modernas se enfrentan a unas nuevas realidades del marketing: los cambios demográficos, una competencia más sofisticada y el exceso de producción en muchos sectores hacen que haya menos clientes disponibles. Muchas empresas luchan en la actualidad por la cuota de mercados planos o decadentes. Por consiguiente, los costos de captación de nuevos clientes están aumentando. De hecho, cuesta más, entre cinco y diez veces más de media, atraer a un cliente nuevo que conservar uno antiguo satisfecho.

La clave para construir relaciones duraderas con los clientes consiste en crear un valor superior para éstos y una mayor satisfacción. Los clientes satisfechos tienen más posibilidades de convertirse en clientes fieles, y los clientes fieles tienen más posibilidades de proporcionar a la empresa una mayor cuota en el mercado.

Los pilares de la creación de relaciones son el valor para el cliente y su satisfacción.

Valor del cliente. Un consumidor compra productos de la empresa que ofrece el mayor valor percibido (evaluación que hace el consumidor de la diferencia entre otros los beneficios y todos los costos de una oferta de marketing respecto a las de la competencia). Normalmente, los clientes no evalúan los valores y los costos del producto de forma objetiva y precisa, sino que se mueven por el valor percibido” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Satisfacción del cliente. La satisfacción del cliente depende de los resultados obtenidos con un producto en relación con las expectativas del comprador. Éstos volverán a comprar y contarán a otros sus buenas experiencias con el producto. La clave es que las expectativas de los consumidores se correspondan con los resultados de la empresa” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Después de las distintas definiciones y del análisis de la información encontrada se puede decir que los clientes satisfechos producen beneficios para la empresa, en primer lugar, son menos sensibles respecto al precio, además, hablan positivamente de la empresa y de sus productos, y se

mantienen fieles durante periodos más largos, es así que una manera de aumentar las ventas es la venta cruzada, a esta (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004) la define como “conseguir más negocios a partir de los consumidores de un producto mediante la venta de ofertas adicionales”.

Se debe recalcar la importancia que tiene no solo el hecho de captar clientes, sino el hecho de conservarlos y desarrollar su valor en el tiempo. La gestión de relaciones con el cliente tiene un enfoque a largo plazo. Hoy en día, las empresas inteligentes no solo buscan conseguir clientes, quieren “poseerlos” de por vida, y crear un capital cliente.

- **La gestión de la información de marketing**

Con el fin de generar un valor y una satisfacción mayor para el cliente, las empresas necesitan información casi en cada momento, información sobre competidores, distribuidores, y otros agentes y fuerzas del mercado.

Cada vez más, las empresas ven la información no solo como un input para el marketing, sino también como un activo estratégico importante y como una herramienta de marketing fundamental. La información de una empresa puede constituir su principal ventaja competitiva, los competidores pueden copiar en el equipamiento, los productos y los procedimientos los unos de los otros, pero no pueden duplicar la información y la capacidad intelectual de la empresa.

En el entorno actual cambiante, la dirección de marketing tiene que actualizar la información puntualmente y tomar decisiones de alta calidad, uno de los principales problemas que se dan son el no poder manejar la información que se tiene.

“En esta sobrecogedora era de la información, resulta muy sencillo acabar enterrado, cargado y quemado por el exceso de la información” (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roch, I; 2004)

Los conceptos generales de marketing en el presente trabajo nos permiten identificar las necesidades de las empresas comerciales del manejo de la información y lo fundamental que es

para esta ciencia del marketing la satisfacción al cliente y que esta se basa en las percepciones del mismo de acuerdo a lo que encuentra en el momento de recibir un producto o servicio.

3.1.1.2 Mercado y Segmentación de mercados

3.1.1.2.1 Mercado

La definición de mercado según (Fischer, L; Espejo, J; 2007) “Es el lugar donde se reúnen oferentes y demandantes, y es en el mercado donde se determinan los precios de los bienes y servicios a través del comportamiento de la oferta y la demanda”, así también menciona que está formado por los consumidores reales y potenciales de un producto o servicio, es así que posee los siguientes elementos:

- La presencia de uno o varios individuos con necesidades y deseos por satisfacer.
- La presencia de un producto que pueda satisfacer esas necesidades.
- La presencia de personas que ponen los productos a disposición de los individuos con necesidades, a cambio de una remuneración.

También se puede hablar de mercados reales y mercados potenciales. El primero se refiere a las personas que normalmente adquieren el producto, y en segundo a quienes podrían comprarlo.

- **Tipos de mercados**

Según Fischer, L; Espejo, J; 2007, la clasificación es la siguiente:

- **Mercado internacional.** Es aquel que comercializa bienes y servicios en el extranjero.
- **Mercado nacional.** Efectúa intercambio de bienes y servicios en todo el territorio nacional.
- **Mercado regional.** Cubre zonas geográficas determinadas libremente, que no coinciden de manera necesaria con los límites políticos.
- **Mercado de intercambio comercial al mayoreo.** Es aquel que se desarrolla en áreas donde las empresas trabajan al mayoreo dentro de una ciudad.
- **Mercado metropolitano.** Cubre un área dentro y alrededor de una ciudad relativamente grande.

- **Mercado local.** Puede desarrollarse en una tienda establecida o en modernos centros comerciales dentro de un área metropolitana.
- **Mercado del consumidor.** En este tipo de mercado los bienes y servicios son rentados o comprados por individuos para su uso personal, no para ser comercializados.
- **Mercado del productor o industrial.** Está formado por individuos y organizaciones que adquieren productos, materias primas y servicios para la producción de otros bienes y servicios, el producto final sufre transformación
- **Mercado del revendedor.** Está conformado por individuos y organizaciones que obtienen utilidades al revender o rentar bienes y servicios, este mercado se denomina también de distribuidores o comercial, y está conformado por mayoristas, agentes, corredores, etcétera, el producto final no sufre transformación.

En el estudio se vuelve imprescindible el conocer en cual mercado estamos enfocándonos, para entender cómo funciona y que necesidades tiene la empresa para que la información y conocimiento que se obtenga sea de acuerdo al entorno del mismo.

3.1.1.2.2 Segmentación de Mercados

De acuerdo a (Schiffman, L; Lazar, L; 1997) la segmentación del mercado y su diversidad están totalmente relacionadas, sin la diversidad en el mercado, compuesto de diferentes personas con diferentes antecedentes y grupos étnicos de origen, diferentes intereses, necesidades y requerimientos, es necesaria una segmentación de mercado.

“La segmentación de mercado puede ser definida como el proceso de dividir el mercado potencial en diferentes subconjuntos de consumidores con necesidades comunes o características, y seleccionar como objetivo a uno o más segmentos para encauzarlo con una mezcla específica de mercadotecnia” (Schiffman, L; Lazar, L; 1997)

En base al análisis realizado sobre la segmentación del mercado se puede decir que es una estrategia atractiva, posible, y potencialmente muy rentable. Las condiciones necesarias para una satisfactoria segmentación son: contar con una población lo suficientemente grande o un grupo de personas con suficiente dinero para gastar (poder adquisitivo general) y que tengan la suficiente diversidad para comenzar a dividirla en grupos de dimensiones adecuadas, de acuerdo con sus características demográficas, psicológicas u otras variables estratégicas.

Según (Schiffman, L; Lazar, L; 1997) las razones principales para que exista una segmentación de mercados son:

- Que el mercado sea muy amplio, ya que esta condición no permite que una sola empresa pueda proporcionar todos los bienes y servicios para satisfacer las necesidades de los consumidores.
- Importancia de delimitar el mercado, ya que dentro de él se presentan distintos tipos de consumidores con diferentes necesidades y deseos.
- Agrupar a los consumidores que posean las mismas características, ya que no todos buscan la misma calidad y cantidad en los productos, ni tienen los mismos intereses y necesidades de compra.
- Por la forma de operación de las empresas se subdivide el mercado total en segmentos pequeños y homogéneos, ya que si no se realiza este proceso es imposible llegar a todos los consumidores.

De acuerdo a este análisis podemos concluir que la segmentación de mercados es el proceso mediante el cual se identifica o se toma a un grupo de compradores homogéneos, es decir, se divide el mercado en varios submercados o segmentos de acuerdo con los diferentes deseos de compra y requerimientos de los consumidores. Los elementos de cada submercado son similares en cuanto a preferencias, necesidades y comportamientos, en base a esto antes de segmentar un mercado deben realizarse estudios acerca de éste para determinar que parte es potencial y dirigirse a ella con un producto específico.

Según (Fischer, L; Espejo, J; 2007) existe un gran número de factores que influyen en la segmentación de mercados, estos se combinan para obtener un conocimiento profundo del mercado y así obtener un perfil más exacto, en la figura 3 se describen los principales criterios para segmentar los mercados.

Criterios para segmentar los mercados			
Geográficos	Demográficos	Psicográficos	Posición del usuario
Regional	Edad	Estilo de vida	No usuarios
Urbana	Sexo	Personalidad	Ex usuarios
Rural	Ocupación	Beneficio del producto	Usuarios potenciales
Suburbana	Educación	Motivos de compra	Usuarios primera vez
Interurbana	Profesión	Conocimiento del producto	Usuarios regulares
Clima	Nacionalidad	Uso del producto	Tasa de uso que se divide en:
	Estado Civil		- Usuario leve
	Tamaño de la familia		- Usuario mediano
	Ingresos		- Usuario fuerte
	Ciclo de vida familiar		- Posición de lealtad
	Religión		Etapas de disposición:
	Clase Social		- Sin noticias del producto
	Características físicas		- Conocimiento bueno
	Actividades		- Conocimiento regular
			- Conocimiento nulo
			- Deseoso y con intención de comprar

Figura 2 Criterios para segmentar los mercados

Fuente (Fischer, L; Espejo, J; 2007)

- **Bases para la segmentación**

Los autores revisados concuerdan que el primer paso en el desarrollo de una estrategia de segmentación es seleccionar las bases sobre las cuales se segmentará el mercado, en la figura 4 se describen las bases de segmentación, existen ocho categorías principales de características del consumidor que proporcionan las bases más populares para la segmentación del mercado.

BASES DE SEGMENTACIÓN	VARIABLES DE SEGMENTACIÓN SELECCIONADAS
Segmentación geográfica.	
Región	Norte, sur, este, oeste.
Tamaño de la ciudad	Áreas metropolitanas más importantes, pequeñas ciudades, pueblos
Densidad de la población	Urbana, suburbana, fuera del área urbana, rural
Clima	Templado, caliente, húmedo.
Segmentación demográfica	
Edad	Menores de 11 años, de 12 a 17, de 18 a 34, de 35 a 49, de 50 a 64, de 65 a 74, 75 +
Sexo	Masculino, femenino.
Estado civil	Soltero, casado, divorciado, unión libre, viudo.
Ingreso	Inferior a \$ 15,000, de 15,000 a 24,999, de 25,000 a 39,999, de 40,000 a 64,999, 65,000 +.
Educación	Preparatoria incompleta, graduado de preparatoria, universidad incompleta, graduado de universidad, posgrado.
Ocupación	Profesional, trabajadoras manuales, oficinistas, agrícolas.
Segmentación psicológica/psicográfica.	
Necesidades-motivación	Casa, seguridad física y moral, afecto, sensación de autoestima
Tipo de personalidad	Extrovertida, introvertida, agresivos, quejumbrosos
Percepción	Bajo riesgo, riesgo moderado, alto riesgo
Aprendizaje-involucramiento	Poco involucramiento, gran involucramiento.
Actitudes	Actitud positiva, actitud negativa.
Psicográfica (estilo de vida)	Libertinos, rectos, conservadores, buscadores de status.
Segmentación sociocultural.	
Cultura	Estadounidense, italiana, china, mexicana
Subcultura	
Religión	Judía, católica, protestante, otra
Raza/origen étnico	Afro estadounidense, caucásico, oriental, hispánico
Clase social	Baja, media, superior
Estilo de vida familiar	Soltero, recién casado, nidos vacíos.
Segmentación relacionada con el uso	
Intensidad de uso	Usuarios pesados o fuertes, usuarios medianos, usuarios ligeros, no usuarios
Estado de conciencia	No conscientes, conscientes, interesados, entusiastas
Lealtad a la marca	Ninguna, alguna, fuerte
Segmentación por situación de uso	
Tiempo	libre, trabajo, apremiante, mañana, noche
Objetivo	Uso personal, regalo, bocado, diversión, logro
Ubicación	Hogar, trabajo, casa de un amigo, en la tienda
Persona	Para sí mismo, amigos, jefe, colegas
Segmentación de beneficios	Comodidad, prestigio, economía, valor-del-dinero
Segmentación híbrida	
Perfiles	
Demográfico/psicográfico	Combinación de características demográficas y psicográficas
Geo demográficas	Suburbanos jóvenes, propiedades de alcurnia
VALS 2	Actualizador, realizado, creyente, triunfador, competidor, experimentador, hacedor, luchador.

Figura 3 Bases para la segmentación de mercado y variables seleccionadas

Fuente (Schiffman, L; Lazar, L; 1997)

Para poder definir las bases de segmentación tomamos de fuente a los autores; (Schiffman, L; Lazar, L; 1997) y (Fischer, L; Espejo, J; 2007), es así que tenemos los siguientes textos:

- **Segmentación geográfica**

En la segmentación geográfica, el mercado se separa por ubicación física, la teoría que respalda esta estrategia es que la gente que vive en el mismo lugar tiene necesidades y deseos similares, y que tales necesidades y deseos son diferentes de las de personas que viven en otros sitios. Algunas diferencias de consumo por regiones pueden estar relacionadas con el clima, es relativamente fácil encontrar diferencias con base geográfica para muchos productos, además, los segmentos geográficos pueden ser fácilmente alcanzados a través de los medios de comunicación locales, incluidos periódicos, TV, radio y ediciones regionales de revistas.

- **Segmentación demográfica**

Las características demográficas, como edad, sexo, estado civil, ingresos, ocupación, educación, son empleadas como base para la segmentación de mercados.

La demografía se refiere a las estadísticas vitales y mensurables de la población. La demografía ayuda a identificar el mercado objetivo, en tanto que las características psicológicas y socioculturales ayudan a describir a las personas del grupo que es lo que piensan y cuáles son sus sentimientos (Schiffman, L; Lazar, L; 1997)

Es importante resaltar que la información demográfica es más accesible y menos costosa para identificar el mercado objetivo, los datos demográficos son fáciles de medir que otras variables de segmentación.

- **Segmentación psicológica/psicográfica**

Las características psicológicas se refieren a los aspectos y cualidades naturales o adquiridas del consumidor individual. Las estrategias de segmentación de consumidores con frecuencia se basan en variables psicológicas específicas, por ejemplo, los consumidores pueden segmentarse de acuerdo con sus necesidades, y motivaciones, personalidad, percepciones, aprendizaje, nivel de involucramiento y actitudes.

La investigación psicográfica, también conocida como análisis del estado de vida, es usada para la promoción de productos. Esta variable es utilizada en investigaciones para descubrir perspectivas y desarrollar perfiles de los consumidores a quienes desean canalizarse.

- **Segmentación sociocultural.**

Esta segmentación tiene éxito en los mercados de consumo, donde se han dividido en segmentos, sobre la base de etapa en el ciclo de vida de la familia, clase social, valores culturales, membrecías subculturales, y afiliación transcultural.

- **Segmentación relacionada con el uso**

Esta segmentación es popular y efectiva, divide a los consumidores en categorías en términos de las características de uso del producto, servicio o marca, como tasa de uso, status de conciencia, y grado de lealtad a la marca.

“La segmentación por tasa de uso establece diferencias entre usuarios pesados, usuarios medios, usuarios ligeros y no usuarios de un producto, servicio o marca específica” (Schiffman, L; Lazar, L; 1997)

El status de conciencia incluye la noción de conciencia del consumidor, nivel de interés, o disposición del comprador, algunas veces el criterio que se utiliza como base para la segmentación es la lealtad a la marca, manera que puede distinguirse los esfuerzos promocionales a la gente que comparte esa misma característica en la población más grande.

- **Segmentación por situación de uso**

Se reconoce que la situación u ocasión a menudo determinan lo que los consumidores compran o consumen, en ocasiones se incluye la idea de la conveniencia de ciertos productos en ciertas situaciones; otros tratan de romper con hábitos del consumidor. La presión del tiempo también es un componente de la situación del uso.

- **Segmentación de beneficios**

Los cambios en los estilos de vida desempeñan un rol importante en la determinación de los beneficios de producto para la elección del consumidor, es así que se encuentra con esta segmentación oportunidades para desarrollar nuevos productos y servicios.

“La segmentación de beneficios se usa sólo para encontrar un nicho para nuevos productos, sino también para repositionar productos establecidos” (Schiffman, L; Lazar, L; 1997)

- **Segmentación Híbrida**

A las distintas bases de segmentación se las puede combinar, en el mundo de marketing tienen un índice de uso alto, encontramos: perfiles psicográficos/demográficos, Segmentación geodemográfica.

• **Criterios de segmentación**

Existen varios criterios para la segmentación de mercados, de esta manera se podrá concluir si este segmento se debe analizar o no.

- **Identificación**

Para dividir el mercado en segmentos separados sobre la base de una necesidad o característica común que sea relevante para el producto o servicio, se debe identificar esta característica relevante.

Algunas variables de segmentación, como la geográfica (ubicación) o las demográficas (edad, sexo, ocupación, raza) son relativamente fáciles de identificar rápidamente de acuerdo a la experticia del analista. Otras, como la educación, el ingreso o el estado civil pueden determinarse con cuestionarios, pero hay otras características, como beneficios buscados o estilo de vida, que son más difíciles de identificar. (Schiffman, L; Lazar, L; 1997)

- **Suficiencia**

Para que un segmento de mercado pueda considerarse un objetivo, debe tener suficiente número de personas para que justifique la preparación de un producto o campaña promocional ajustados a sus necesidades o intereses específicos. Para estimar el tamaño de cada segmento en consideración, es frecuente que se utilicen datos demográficos secundarios o desarrollen una encuesta probabilística cuyos resultados puedan ser proyectados al mercado total.

- **Estabilidad**

Un enfoque en segmentos del consumidor que sean relativamente estables en términos de factores y necesidades demográficas y psicológicas, y que es probable que crezcan aún más en el curso del tiempo, se evitan segmentos inestables que son imprevisibles en el seguimiento de modas.

- **Accesibilidad**

Es un análisis para que el segmento meta al que se dirijan sea accesible de acuerdo a los recursos que se posean, a la vez también existe una búsqueda constante sobre nuevos medios de difusión que permita alcanzar mercados objetivos con un mínimo de competencia y desperdicio de circulación.

El tema de segmentación de mercado es una de las bases para el desarrollo de este estudio ya que el mercado, está compuesto de una diversidad amplia de personas y todas con distintas procedencias, cultura, hábitos, intereses, necesidades, es por ello que se vuelve necesario dividir el mercado potencial en diferentes subconjuntos de consumidores con necesidades comunes o características, y los esfuerzos de la empresa encaminarlos a ellos agrupados para crear clusters de mercado y entender a cabalidad su comportamiento relacionado.

3.1.1.3 Comportamiento del Consumidor

Al comportamiento del consumidor podemos definirlo como “Actos, procesos y relaciones sociales sostenidas por individuos, grupos y organizaciones para la obtención, uso y experiencia consecuente con productos, servicios y otros recursos” (Fischer, L; Espejo, J; 2007)

El término comportamiento del consumidor se refiere a la conducta que los consumidores tienen cuando buscan, compran, usan, evalúan y desechan productos y servicios que esperan que satisfagan sus necesidades. El estudio del comportamiento del consumidor es el estudio de cómo los individuos toman decisiones para gastar sus recursos disponibles (tiempo, dinero, esfuerzos) en asuntos relacionados con el consumo. Esto incluye el estudio de qué es lo que compran, por qué lo compran, cuándo lo compran, dónde lo compran, con qué frecuencia lo compran, y que tan frecuentemente lo usan (Schiffman, L; Lazar, L; 1997)

Es así que los actos, procesos y relaciones sociales a los que se hace referencia en los conceptos anteriores incluyen diversas actividades del consumidor como son el conocimiento de una necesidad, las compras por comparación entre tiendas, el razonamiento de la información con que se cuenta para analizar los beneficios y riesgos del producto deseado, o buscar el consejo sobre esta búsqueda.

En el proceso de compra, el consumidor reúne información acerca de las alternativas, la procesa, aprende acerca de los productos disponibles y determina qué alternativa es la más acorde con las necesidades percibidas. Se debe resaltar que el estilo de vida, al ser un aspecto significativo del comportamiento del consumidor, “Significa cómo vive una persona, qué productos compra, cómo los utiliza, la forma como el individuo se ve a sí mismo, sus emociones, percepciones y deseos como resultado de factores externos que afectan directamente sus decisiones” (Schiffman, L; Lazar, L; 1997)

Podemos concluir que el estudio del comportamiento del consumidor, se centra en cómo y por qué los consumidores toman decisiones para comprar bienes y servicios, a la vez también considera los usos que los consumidores hacen de los bienes que compran y su evaluación posterior, ya que existe interés en la forma que los individuos se deshacen de lo que alguna vez fue un producto recién comprado y así vincularlo con la fabricación del producto y la frecuencia con que los consumidores adquieren el reemplazo.

Existen diferentes tipos de consumidores, como los personales y los consumidores organizacionales. El consumidor personal compra bienes y servicios para uso propio, es decir que los bienes se compran para uso final de los individuos, mientras que los consumidores organizacionales son organismos lucrativos y no lucrativos.

Las empresas orientadas a la mercadotecnia en el libro de (Schiffman, L; Lazar, L; 1997) se dedicaron a fabricar sólo lo que habían investigado que los consumidores habrían de comprar. Las necesidades y deseos del consumidor se convirtieron en el punto primario de la empresa, por lo cual una compañía debe determinar las necesidades y deseos de mercados objetivos específicos, y entregar mejor que los competidores la satisfacción deseada. El mercadólogo debe fabricar lo que puede vender.

Para identificar necesidades insatisfechas del consumidor, las compañías tuvieron que realizar extensas investigaciones de mercado al respecto, descubrieron que los consumidores eran individuos muy complejos, sujetos a una diversidad de necesidades psicológicas y sociales muy independientes de sus necesidades de supervivencia, descubrieron que las necesidades y prioridades de diferentes segmentos de consumidores diferían profundamente. Y también descubrieron que, con el fin de diseñar nuevos productos y estrategias de mercadotecnia que permitieran la satisfacción de necesidades del consumidor.

La investigación del consumidor es la metodología empleada para estudiar el comportamiento del consumidor, enfoque positivista y el interpretivista. Los positivistas tienden a ser objetivos y empíricos, a buscar causas para el comportamiento, y a realizar estudios de investigación que pueden generalizarse a poblaciones más grandes. Los interpretivistas, tiende a ser cualitativa, y basada en pequeñas muestras. Aunque tienden a visualizar cada situación de consumo como única e irrepetible, los interpretivistas tratan de hallar patrones comunes de valores, significados y comportamiento operativos en todas las situaciones de consumo. (Schiffman, L; Lazar, L; 1997)

Es importante para este estudio abordar el tema de comportamiento del consumidor ya que es análisis a la conducta de los consumidores cuando toman decisiones para usar sus recursos en productos y servicios, uno de los objetivos planteados es encontrar el perfil del consumidor y esta base teórica nos explica el detalle al que debemos llegar, es decir que conocer de qué es lo que compran, por qué lo compran, cuándo lo compran, dónde lo compran, con qué frecuencia lo compran.

3.1.1.3.1 Proceso de Compra

Para (Fischer, L; Espejo, J; 2007) el “Proceso de compra es el conjunto de etapas por las que pasa un consumidor para poder adquirir algo, estas etapas cambian gradualmente según sean los gustos y las necesidades que el consumidor tenga”

- **Pasos del proceso de compra**

Los siguientes pasos del proceso de compra son tomados de (Fischer, L; Espejo, J; 2007)

- Necesidad sentida

El punto de partida está en identificar la necesidad del consumidor, la cual representa un estado de tensión, es decir, que la persona busca satisfacer o disminuir dicha tensión, lo cual logrará presumiblemente con un objeto o actividad.

- Actividad previa a la compra

“La necesidad genera la acción y hace que el individuo se vuelva más sensible a estímulos relacionados con la misma; la reacción del individuo dependerá de la intensidad de dicha necesidad” (Fischer, L; Espejo, J; 2007)

La persona no satisface muchas de las necesidades de manera inmediata, sino que espera adquirir y acumular experiencia. La duración y la clase de actividad anterior a la compra varían de acuerdo con el tipo de producto, con la necesidad que experimenta el comprador y con su personalidad. Esta parte es el resultado de los comportamientos aprendidos. La idea primordial es que el individuo percibe, busca para conocer el producto, compara con lo que le gusta, lo prefiere y llega a la convicción que lo conduce a la acción.

- Decisión de compra

La decisión de compra constituye un conjunto de decisiones en donde intervienen variables como el producto, la marca, el estilo, la cantidad, el lugar, la tienda o el vendedor, la fecha, el precio y la forma de pago, originándose con ello una gran gama de combinaciones que finaliza con la decisión.

- Sentimientos posteriores a la compra

El producto puede satisfacer las necesidades actuales, pero también crear incertidumbre o sentimientos negativos, a los que se denomina disonancia cognoscitiva, y que se deben principalmente a dos razones: 1) La incertidumbre originada en la fase anterior a la compra; por ejemplo, dificultad de elegir entre posibles opciones. 2) Información desfavorable sobre la elección efectuada o insatisfacción en el uso del producto. (Fischer, L; Espejo, J; 2007)

- **Participantes en el sistema de Compra**

En la decisión de compra intervienen diferentes personas o elementos que participan activamente en el proceso, desde su inicio hasta la culminación de la compra, se los clasifica de acuerdo con la función que desempeñan en la decisión:

- Influyentes; personas o elementos que explícita o implícitamente ejercen alguna influencia en la decisión.
- Decisor; Es quien determina, en último término, parte de la decisión de compra o su totalidad; si se compra, qué debe adquirirse, cómo, cuándo y dónde conviene hacerlo.
- Comprador; Persona que realiza, de hecho, la compra o adquisición del bien.
- Usuario; Quien o quienes consumen o utilizan el producto o servicio.
- Evaluador; Persona que da una opinión de valor respecto de la decisión de compra.

Según (Fischer, L; Espejo, J; 2007) esta clasificación se da ya que la persona que compra un producto no siempre es la usuaria, o la única usuaria, del producto. Ni es el comprador necesariamente la persona que toma la decisión de comprar.

3.1.1.3.2 Factores influyen en el estilo de vida

La mercadotecnia se interesa en comportamiento humano, en la comprensión de la conducta humana; en saber por qué los individuos actúan de cierta manera en relación con el consumo, y conocer cuáles influencias internas y externas.

- **Factores de Influencia Externos**

- **Cultura**

“La representación de factores tales como los conocimientos, las creencias, los valores, el arte, la moral, las leyes, las costumbres y los hábitos adquiridos por el hombre como miembro de una sociedad” (Fischer, L; Espejo, J; 2007)

- **Aspectos demográficos**

Estos aspectos demográficos influyen en el estilo de vida, particularmente en lo relacionado a ingreso, edad, situación geográfica.

- **Estratos sociales o niveles socioeconómicos**

Los estratos o niveles socioeconómicos son divisiones relativamente permanentes y homogéneas dentro de una sociedad, en la que los individuos comparten estilos de vida y conductas similares.

- **Grupos de referencia y convivencia**

Un grupo de convivencia es aquel cuyos valores son utilizados por otros individuos como base para su conducta en un momento determinado, el grupo de referencia es donde la interrelación de los miembros es aislada o nula. Estos grupos se dividen en aspiracionales y disociativos. Un grupo aspiracional es aquel al que deseamos pertenecer; grupo disociativos, generalmente por tratarse de metas ya superadas. Tres grandes grupos de consumidores han destacado: Generación Baby Boom, Generación X y Generación Y. (Fischer, L; Espejo, J; 2007)

- **La familia**

Sin duda el comportamiento de consumo dentro de la familia ha ido cambiando y cada vez más, todos los miembros de la familia participan activamente en las decisiones de compra de productos y servicios para el hogar.

- **Factores Internos de influencia**

- **Percepción**

La percepción son aquellas actividades mediante las cuales un individuo adquiere y da significado a los estímulos. El estímulo aparece dentro del campo de los receptores sensoriales y genera una acción. La mercadotecnia necesita propiciar estos estímulos para que el individuo adquiera un producto (Fischer, L; Espejo, J; 2007)

De acuerdo a (Fischer, L; Espejo, J; 2007) el primer paso de la percepción es buscar la atención hacia el objetivo; éste puede ser un anuncio publicitario, un envase, el precio atractivo, los colores de una etiqueta, la marca, etcétera; todos estos elementos deben ser atractivos para que el consumidor fije su atención en ellos y los perciba; los estímulos son percibidos por cada individuo en forma diferente de acuerdo con los factores externos.

- **Aprendizaje**

Gran parte del comportamiento humano es aprendido. A través del aprendizaje adquirimos la mayoría de nuestras actitudes, valores, costumbres, gustos, conductas, sentimientos, preferencias, deseos y significados. La cultura, el nivel socioeconómico, la estructura familiar, los amigos y las instituciones proporcionan experiencias de aprendizaje que influyen de manera importante en el estilo de vida de los individuos.

- **Motivación**

El término motivación se refiere al comportamiento suscitado por necesidades y dirigido hacia la obtención de un fin.

Motivación es un comportamiento o una actitud del consumidor para conseguir un bien o un servicio, Abraham Maslow desarrollo una jerarquía de motivos que va de los básicos a los avanzados, señalando que un ser humano debe satisfacer sus necesidades básicas para posteriormente complacer sus necesidades más superfluas. La escala considera las necesidades, Fisiológicas, de seguridad, de pertenencia, de estima, de autorrealización. (Fischer, L; Espejo, J; 2007)

3.1.1.3.3 Comportamiento de compra del Consumidor Organizacional

El comportamiento de compra de las organizaciones puede definirse como

El proceso de toma de decisiones mediante el cual las empresas y otras organizaciones formales, como las instituciones públicas o privadas, establecen la necesidad de comprar productos y servicios e identifican, evalúan y eligen entre las distintas ofertas y suministradores relacionados con los productos o servicios que ellos a su vez venden a sus clientes (Fischer, L; Espejo, J; 2007)

Las empresas se colocan metas a lograr como:

- Mantener la continuidad del abastecimiento.
- Hacerlo con la inversión mínima en existencia.
- Evitar duplicidades, desperdicios e inutilización de los materiales.
- Mantener los niveles de calidad en los materiales.
- Procurar materiales al precio más bajo compatible con la calidad.
- Mantener la posición competitiva de la empresa.

Se aborda el tema de comportamiento organizacional, ya que el proceso que se propone optimizar es parte de una empresa y de su desarrollo.

3.1.1.4 Estrategia de producto

La estrategia de producto es una de las más importantes dentro de la mercadotecnia, ya que los productos fracasarán si no satisfacen los deseos, necesidades y expectativas de los consumidores.

“Un producto se puede considerar como el conjunto de beneficios y servicios que ofrece un comerciante en el mercado”. (Fischer, L; Espejo, J; 2007)

“Un producto engloba atributos tangibles e intangibles (embalaje, color, precio, prestigio del fabricante y del vendedor, etc.) que el comprador acepta como algo que ofrece satisfacción a sus deseos o necesidades” (William J. Stanton)

3.1.1.4.1 Clasificación de los productos

Según (Fischer, L; Espejo, J; 2007) podemos clasificar a los productos:

- **Productos de consumo**

Son aquellos que están destinados a ser utilizados y adquiridos por los consumidores, de acuerdo con sus deseos y necesidades, y se pueden utilizar sin proceso industrial adicional; es decir, son adquiridos en última instancia por el consumidor en su forma original para ser consumidos o utilizados en el hogar (Fischer, L; Espejo, J; 2007)

Los productos de consumo pueden clasificarse en los siguientes subgrupos según (Fischer, L; Espejo, J; 2007):

- **Duraderos y no duraderos;** Los productos duraderos son artículos tangibles y de uso cotidiano (por ejemplo, televisores, automóviles, refrigeradores, estéreos, licuadoras, etcétera). Los no duraderos son los que tienen poca vida (por ejemplo, alimentos)
- **De conveniencia o habituales;** Son los que el consumidor compra con cierta regularidad, sin planificación (por ejemplo, cigarrillos, dulces, pastas dentales, etcétera).
- **De elección o compra;** Son aquellos cuyos atributos se comparan en el proceso de selección y de compra (por ejemplo, prendas de vestir, perfumes, relojes, etcétera)
- **Especiales o de especialidad;** Son artículos con características únicas o de identificación; el consumidor está dispuesto a entregar más dinero o esfuerzo para adquirirlos, y por su mente no pasa la idea de aceptar otro artículo (por ejemplo, un regalo de cumpleaños, seguros médicos, servicios funerarios, ataúdes, etcétera).
- **No buscados;** Son los artículos por los que el consumidor no hace ningún esfuerzo de compra, quizá porque no tiene presente la necesidad o deseo.

- **Productos industriales**

“Son bienes o servicios utilizados en la producción de otros artículos, es decir, no se venden a los consumidores finales” (Fischer, L; Espejo, J; 2007)

Los bienes industriales abarcan suministros, accesorios, servicios e incluso fábricas o equipo, y se clasifican en según Fischer, L; Espejo, J:

- Instalaciones (plantas industriales, terrenos)
- Equipos (herramientas)
- Materiales de operación (aceites, papelería, focos)
- Servicios (despachos fiscales y contables, agencias de publicidad, bancos, etcétera)
- Materiales de fabricación

3.1.1.4.2 Posicionamiento de un producto

Recopilando varios autores entre ellos (Fischer, L; Espejo, J; 2007) y (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roche, I; 2004) el término posicionamiento se refiere a un programa general de mercadotecnia que influye en la percepción mental (sensaciones, opiniones, impresiones y asociaciones) que los consumidores tienen de una marca, producto, grupo de productos o empresa, en relación con la competencia.

Hoy en día las empresas se preocupan por proporcionar satisfacciones que lleguen a posicionarse fuertemente en la mente del consumidor, los mercadólogos disponen de diferentes estrategias para alcanzar y reforzar el posicionamiento deseado para su empresa, marca, producto o grupo de productos, entre las estrategias más comunes según (Schiffman, L; Lazar, L; 1997) son:

- Posicionamiento con base en los atributos
- Posicionamiento con base en los beneficios
- Posicionamiento con base en las ocasiones de uso
- Posicionamiento con base en los usuarios
- Posicionamiento comparativo
- Estrategia de reposicionamiento
- Posicionamiento en contra

3.1.1.4.3 Línea y mezcla de productos

“La línea de productos es el grupo de productos estrechamente relacionados porque satisfacen una necesidad o se usan conjuntamente” (Fischer, L; Espejo, J; 2007).

Después de un análisis del concepto podemos decir que una línea de productos es una agrupación de productos que se relacionan entre sí por algunas de sus características, ya sea por su razón de uso, por su distribución, por su segmentación o su precio.

En cada empresa o giro de negocio se debe determinar las líneas de producto, como ejemplos de líneas de productos existen:

- Línea blanca (refrigeradores, estufas, alacenas, etcétera).
- Línea electrónica (televisores, planchas, radios, consolas, estéreos, tostadores, etcétera).
- Línea de cosméticos (lápices labiales, sombras, rubores, esmaltes, tintes, etcétera).
- Línea textil (ropa de niño, adulto, etc.)

“La mezcla de productos es la lista de todos los productos que una empresa ofrece al consumidor, la estructura de la mezcla tiene dos dimensiones: amplitud y profundidad” (Fischer, L; Espejo, J; 2007), es así que la amplitud se mide por el número de productos que ofrece la empresa en una línea; a esto se le conoce también como variedad, en cambio a la profundidad es el surtido de tamaños, colores, modelos, precios y calidad que ofrece una línea.

3.1.1.4.3.1 Decisiones sobre la Línea de producto

Existen las estrategias de línea que son utilizadas tanto por fabricantes como por intermediarios; es así que tenemos según (Fischer, L; Espejo, J; 2007):

- **Decisiones sobre la profundidad de la línea**

Aquí se puede tomar dos decisiones:

- Podar la línea; hacen referencia a la eliminación de productos de esa línea

- Extender la línea; se puede extender o alargar la profundidad de una de sus líneas de dos formas: completándola o ampliándola.

- **Decisiones sobre modernizar la línea**

(Fischer, L; Espejo, J; 2007) describe que en primer lugar se debe realizar un análisis de la longitud de la línea, para saber si es la adecuada, y se puede tomar decisiones de modernización para aumentar los beneficios. La empresa puede optar por modernizar la línea de forma gradual o simultánea.

- **Decisiones sobre crear una nueva línea**

La nueva línea puede estar o no relacionada con los productos actuales de la empresa, y ser comercializada con la misma marca o con una marca distinta.

3.1.1.4.3.2 Factores que influyen en los cambios de la mezcla de productos

Según (Fischer, L; Espejo, J; 2007) tenemos:

- **Población de consumidores y usuarios industriales.** Un sector de la población puede inducir a la empresa a cambiar su mezcla de productos, pues los gustos y las necesidades van cambiando.
- **Poder de compra.** Cuando cambia el poder de compra es necesario modificar la mezcla de producto, ya que los segmentos de mercado se amplían o reducen.
- **Comportamiento del consumidor.** Hay varias situaciones referentes a la motivación, actitudes, preferencias y hábitos de compra del consumidor, con base en lo cual el mercadólogo decide aumentar o contraer su mezcla de productos, es decir, el comportamiento cambia al pasar de un segmento de mercado a otro.

3.1.1.4.3.3 Portafolio de Productos

Según (Fischer, L; Espejo, J; 2007):

La amplitud se refiere al número de líneas de productos, que maneja una empresa; la extensión es el número total de productos que integran el portafolio; la profundidad se

relaciona con el número de variantes o versiones de producto que una empresa ofrece en cada una de las líneas de producto, y la consistencia es el punto relativo a qué tan relacionadas se encuentran las líneas en cuanto a uso final, requerimientos de producción, sistemas de distribución, proveeduría, etcétera. (Fischer, L; Espejo, J; 2007)

Después de un análisis de los conceptos de (Fischer, L; Espejo, J; 2007) y de (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roche, I; 2004) podemos hablar de que un portafolio de productos es un conjunto de todos los productos agrupados en líneas que una organización ofrece a su mercado, este conjunto posee cuatro características fundamentales: amplitud, extensión, profundidad y consistencia.

El análisis de rentabilidad por producto tiene como fin identificar los productos cuyas ganancias son más atractivas y a los que están siendo subsidiados. El análisis de competitividad tiene como fin comparar integralmente los productos entre empresas; el modelo comúnmente utilizado para este fin se conoce como benchmarking.

La decisión de crear nuevas líneas de productos o ampliar las existentes es una acción orientada a lograr los siguientes objetivos según (Fischer, L; Espejo, J; 2007):

- Atender de manera más personalizada algún segmento del mercado.
- Aprovechar la capacidad instalada.
- Mejorar la participación de mercado.
- Bloquear la entrada de la competencia.

Siempre se debe tener presente que el análisis del portafolio de productos tiene cuatro objetivos centrales según (Schiffman, L; Lazar, L; 1997):

- Mantener un equilibrio en la obtención y asignación de recursos;
- Estar en posibilidad de competir exitosamente en multi mercados con multi productos
- Proyecta a la empresa hacia el futuro, porque si solo se tiene un producto y éste se muere, con él se muere la empresa
- Con el análisis, identificar claramente dónde se debe invertir, mantener, cosechar o desinvertir los recursos.

3.1.1.4.3.4 Ciclo de vida del producto

Todo producto tiene un ciclo de vida que se divide en cuatro etapas:

- **Etapa de introducción**

Es el momento en que se lanza el producto bajo una producción y programa de comercialización.

“La introducción se caracteriza por un aumento lento de ventas y altos gastos promocionales debido a la necesidad de informar a los consumidores potenciales sobre el producto, de estimular la prueba del producto y de lograr su distribución en las tiendas detallistas” (Fischer, L; Espejo, J; 2007)

- **Etapa de crecimiento**

En esta etapa, el producto es aceptado en el mercado y se constata un aumento en la curva de las ventas y de los beneficios, también existe un aumento de la competencia, un manejo de la calidad de productos, mejoramiento en canales y promociones de uso de producto, la estrategia en la etapa de crecimiento consiste en sostener el índice de mercado.

- **Etapa de madurez**

“Las tácticas de la mercadotecnia y la imagen de la marca son bien conocidas durante esta etapa, además de la lealtad de los clientes y la participación de mercado; el producto se estabiliza y disminuye el margen de utilidad debido a que los precios se acercan más a los costos” (Fischer, L; Espejo, J; 2007)

- **Etapa de declinación**

En la última etapa, el paso de la moda es inevitable porque los nuevos productos empezarán un ciclo de vida para sustituir a los viejos. Abordar el tema de producto es importante en este estudio, ya que la identificación de las características del producto son la base del mismo, a la vez uno de los objetivos a obtener es el mix de producto por lo cual es relevante el análisis de qué tipo de producto estudiamos, como se forma un mix de producto, las características de la línea de producto y las decisiones y factores que se involucran. También identificar el ciclo de vida de un

producto y las estrategias de posicionamiento para saber cuál es la opción rentable para la selección del mix o estrategia a realizar.

3.1.1.5 Estrategia de Precio

El dinero es la medida social del valor, existen dos tipos de valor según Fischer, L; Espejo, J; 2007:

- **Valor de uso;** El valor de una cosa depende de la utilidad específica que represente para el individuo. Este valor es subjetivo a individual, haciendo abstracción de toda idea de intercambio.
- **Valor de cambio;** El valor de una cosa depende de la importancia que los demás le adjudiquen; este tipo de valor, al ser intercambiado, satisface las necesidades del que lo posee en forma indirecta; también es subjetivo, y por lo general está estrictamente ligado a la idea de intercambio.

El precio de un producto es sólo una oferta para probar el pulso del mercado. Si los clientes aceptan la oferta, el precio asignado es correcto; si la rechazan, deben cambiarse el precio con rapidez o bien retirar el producto del mercado.

La fijación de precios es probablemente la más compleja y difícil de las tareas, una función clave de la mercadotecnia y, sin lugar a dudas, el tema al que más tiempo le dedica el mercadólogo.

Según la teoría económica, el precio, el valor y la utilidad son conceptos relacionados. La utilidad es el atributo de un artículo mediante el cual se satisfacen deseos. El valor es la expresión cuantitativa del poder que tiene un producto de atraer otros productos a cambio. El precio es el valor expresado en moneda. (Fischer, L; Espejo, J; 2007)

En resumen, el precio es la cantidad de dinero que se necesita para adquirir en intercambio la combinación de un producto y los servicios que lo acompañan. La clave para determinar el precio de un producto es entender el valor que los consumidores perciben de él, este valor es el resultado de las percepciones de los consumidores acerca de la satisfacción total que el producto proporciona, partiendo de todos los beneficios.

El conjunto de satisfacciones del producto incluye, además de las características tangibles, las intangibles; por ejemplo, la imagen de la empresa, del distribuidor, la garantía y la marca. El

precio del producto es con frecuencia el elemento más sobresaliente que induce a la compra; es el primer atributo que busca el consumidor potencial.

Existe una gran importancia del precio en la economía, ya que la fijación de precios equilibrados da como resultado una economía sana y de acuerdo al perfil del consumidor o el perfil del sitio de su distribución es un factor decisivo en la compra.

3.1.1.6 Estrategia de Distribución

Acorde al concepto de (Fischer, L; Espejo, J; 2007) “El canal de distribución lo constituye un grupo de intermediarios relacionados entre sí que hacen llegar los productos y servicios de los fabricantes a los consumidores y usuarios finales”

Las decisiones sobre los canales de distribución dan a los productos los beneficios del lugar y al consumidor los beneficios de tiempo.

El beneficio del lugar se refiere al hecho de llevar un producto cerca del consumidor con la finalidad de que éste no tenga que recorrer grandes distancias para obtenerlo y satisfacer así una necesidad. El beneficio de lugar se puede ver desde dos puntos de vista: el primero considera los productos cuya compra se favorece cuando están muy cerca del consumidor, el cual no está dispuesto a realizar un gran esfuerzo por obtenerlos. El segundo punto de vista considera los productos exclusivos, los cuales deben encontrarse sólo en ciertos lugares para no perder su carácter de exclusividad; en este caso, el consumidor está dispuesto a realizar algún esfuerzo, en mayor o menor grado, para obtenerlo, según el producto de que se trate (Fischer, L; Espejo, J; 2007)

Los factores que influyen en el diseño de los canales de distribución son:

- Características de los clientes.
- Características de los productos.
- Características de los intermediarios.
- Características de la competencia.
- Características de la empresa.
- Características ambientales.

3.1.1.6.1 Intermediarios

Los intermediarios son grupos independientes que se encargan de transferir el producto del fabricante al consumidor, obteniendo una utilidad por el trabajo realizado.

En el tema planteado de estudio, nos enfocamos en los minoristas o detallistas. “Los minoristas o detallistas son aquellos comerciantes cuyas actividades se relacionan con la venta de bienes y/o servicios a los consumidores finales” (Fischer, L; Espejo, J; 2007)

A los minoristas/detallistas se los puede clasificar en:

- a) **Tipo de tienda.** Se refiere al esfuerzo que el consumidor realiza para hacer su compra. El tipo de tiempo se clasifica de acuerdo a como el consumidor la percibe o la imagen que ésta proyecta:
 - Tiendas de servicio rápido
 - Tiendas comerciales
 - Tiendas especializadas
 - Plazas comerciales
- b) **Forma de propiedad.** Los minoristas pueden operar en forma independiente o formar parte de cadenas.
 - Minorista independiente
 - Tiendas en cadena
 - Organización por cooperativas
 - Franquicia
- c) **Líneas de productos,** se clasifica por la variedad y el surtido de productos que ofrecen al público.
 - Minoristas de mercancías en general.
 - Minoristas de líneas limitadas.
 - Minoristas de líneas especiales.
 - Clubes de compra.

3.1.1.6.2 Distribución física del producto

A la distribución física podemos mencionar con un indicador del éxito o el fracaso en los negocios, esta incluye la integración de una serie de actividades que permiten ofrecer un nivel de servicio que satisfaga las necesidades del consumidor.

“El objetivo principal de la distribución física es incrementar la satisfacción de los clientes y mejorar su nivel de vida, tomándose las medidas necesarias para que los productos adecuados estén disponibles en el lugar y en el tiempo precisos para el consumidor” (Fischer, L; Espejo, J; 2007)

Un sistema de distribución física eficaz contribuye al beneficio que los compradores esperan.

3.1.1.6.3 Elementos del sistema de distribución física

Según (Fischer, L; Espejo, J; 2007) los elementos del sistema de distribución física son y se los detalla así:

- **Pronóstico de la demanda;** Para determinar un sistema de distribución física eficiente deben identificarse los deseos y las necesidades del mercado y de los intermediarios, se realiza un pronóstico de la demanda así se puede controlar debidamente el flujo de productos por medio de un sistema logístico.
- **Pronóstico de pedidos;** Las necesidades del siguiente miembro del canal pueden determinarse presentando un pedido de productos.
- **Manejo del inventario;** El pronóstico de la demanda puede haber divergencia entre la demanda real y la que se previó, es así que los productos que se almacenan deben tratar de compensar la inexactitud de dichas divergencias.

El inventario de productos debe resultar en una forma de protección ante la incapacidad de satisfacer directamente la demanda partiendo de una línea de producción. Esta solución puede ser costosa, ya que implica costos de espacio para almacenamiento, seguros, costos de manejo de materiales y costos asociados. Éstos se reducen limitando el tamaño del inventario, aunque se corre el riesgo de perder ventas por no disponer del producto. En cambio, prever perfectamente la demanda del producto y desarrollar el inventario en forma apropiada para satisfacer dicha demanda, sería ideal. (Fischer, L; Espejo, J; 2007)

- **Cantidad de pedido;** Esta decisión repercute directamente en la frecuencia con que debe pedirse el producto, uno de los dos elementos principales para determinar la cantidad de pedido son los costos de tramitación; el segundo factor que debe tomarse en cuenta es el costo de inventario. Éste aumenta de acuerdo con el volumen de existencias, los costos de mantenimiento de inventarios se dividen en tres categorías:
 - Gastos por espacio y almacenamiento
 - Costos de capital
 - Depreciación y envejecimiento

- **Almacenamiento;** Si se mantiene un inventario, se requiere de un lugar para guardar los productos, hay varios aspectos a tomar en cuenta como tamaño, cantidad y ubicación de las instalaciones para almacenarlos. Las funciones del almacén son las siguientes:
 - Recibir mercancías
 - Identificar las mercancías
 - Clasificar mercancías
 - Enviar las mercancías al almacenaje
 - Conservar mercancías
 - Retirar, seleccionar o escoger mercancías
 - Ordenar el embarque
 - Despachar el embarque

3.1.1.6.4 Objetivos Secundarios de la distribución física

Según (Hasty, R; Reardon, J; 1998) los objetivos secundarios que se deben lograr con la distribución son:

- Disponibilidad de productos
- Exactitud para llenar los pedidos
- Servicio seguro
- Servicio rápido
- Disponibilidad de servicio

El tema de distribución física dentro del estudio es importante ya que trata de la cantidad, variedad, pasos para que la mercadería llegue al punto de venta, estas decisiones en la distribución dan a los productos los beneficios del lugar y al consumidor los beneficios de tiempo, el objetivo de optimizar el proceso de compra tiene sus fundamentos en la distribución cuando hablamos de sus elementos: pronóstico de la demanda, pronóstico de pedidos, manejo del inventario; cantidad de pedido y almacenamiento; esto a la vez para la empresa es un punto delicado ya que es manejo de recursos de todo tipo como financiero, humano, tiempo, etc.

3.1.1.7 Ventas en Retail

La venta al detalle o retail está estrechamente entrelazada con los consumidores y es así que puede obtener de primera fuente las necesidades de estos, en (Hasty, R; Reardon, J; 1998) nombra que “la idea principal es primero la satisfacción del cliente, ya que esta será la piedra angular para tener éxito en la venta en retail durante el siglo XXI”.

“La venta de retail, al menudeo o al por menor, se define como, las actividades involucradas en la venta de bienes y servicios a los consumidores para uso personal, familiar o doméstico” (Hasty, R; Reardon, J; 1998)

Es importante mencionar la mezcla de marketing, esta es: producto, precio, plaza y promoción, el elemento producto incluye todas las facetas de suministrar bienes y servicios, incluyendo marca, empaque y diseño de producto. El precio está diseñado para recuperar los gastos de hacer el negocio, al igual que el costo de los bienes y permitir una ganancia para los propietarios o accionistas de la firma. La promoción incluye un amplio rango de actividades de publicidad, venta personal, promoción de ventas y relaciones públicas. La distribución incluye las actividades de localización, logística, cambio de estructura de propiedad, transporte y almacenamiento de bienes.

De acuerdo a los conceptos estudiados podemos concluir que la venta de retail es el enlace en el sistema de distribución de marketing que pone bienes y servicios a disposición de los consumidores.

Las ventas a compradores organizacionales se llaman transacciones mayoristas, y las ventas a consumidores finales se llaman transacciones minoristas, según (Hasty, R; Reardon, J; 1998) la diferencia entre transacciones mayoristas y minoristas se centran en cinco características y las describo así:

- **Motivos de compra;** una compra organizacional tiene motivos de utilidad económica o fines sociales, mientras que los compradores finales no son ni fabricantes ni revendedores, y no tienen la rentabilidad a largo plazo dentro de su motivo de compra.
- **Descuentos comerciales;** a los compradores mayoristas se les ofrece un descuento comercial, que puede consistir en la reducción del precio por prestar servicios, y realizan funciones de venta y almacenamiento.
- **Requerimientos de almacenamiento;** para poder tener el surtido necesario se debe utilizar recursos en almacenamiento, los compradores organizacionales a diferencia de los finales no compran para un solo trabajo.
- **Impuestos de ventas;** la mayor parte de países tienen impuestos de ventas o valor agregado, y casi siempre estos impuestos de ventas se aplican solo a las ventas de retail.
- **Precios diferenciales;** una firma que vende al por mayor está limitada en los precios diferenciales que puede cobrar a las firmas competidoras.

De acuerdo a estas diferencias podemos notar que la venta de retail beneficia a los consumidores, ya que los minoristas hacen más fácil la compra. Los minoristas realizan la función de selección al evaluar todos los diferentes productos disponibles y ofrecerlos en una distribución óptima para escoger. Como resultado, los consumidores se benefician de tener un proceso de compras más eficiente y pueden sacar ventaja de mayores niveles de compras en un mismo lugar.

El almacenamiento significa que el minorista adquiere y conserva segura un amplio inventario de productos, de manera que estén a disposición del cliente cuando éste los necesite. La función de almacenamiento es también una de las más riesgosas para el minorista, ya que cuando el minorista tiene productos en inventario, hay riesgos de deterioro físico, incendio, robo, inundación, cambio de moda e inclusive, obsolescencia tecnológica.

Los inventarios también tienen un costo debido a que ocupan espacio sobre el cual el minorista tiene que pagar impuestos y pólizas de seguro.

Un aspecto adicional a considerar es el costo de oportunidad del inventario si el dinero que representa se hubiera invertido en otro activo que produjera un retorno mayor.

Los minoristas brindan muchos servicios a los clientes, en (Hasty, R; Reardon, J; 1998) describe varios como: trabajar en horarios adaptados, prestar servicios de despacho, crédito, modificaciones, privilegios comerciales, zonas de parqueo, explicaciones sobre el uso del producto y prestar ayuda por medio de vendedores amables y conocedores del producto, el minorista debe decidir cuál es el nivel de servicio necesario o apropiado, según el precio que el cliente quiere pagar.

Parte de la utilidad bruta que se paga al minorista se debe a que el sistema de venta de retail realiza la tarea de comprar en grandes cantidades y vender en las pequeñas cantidades que el cliente necesita, la variedad de servicios que un minorista específico ofrece también afectará dicho porcentaje.

“La venta de retail trae consigo varios beneficios económicos, la importancia del marketing y de la venta al detal proviene del hecho de que para una economía es necesario brindar cuatro beneficios básicos: forma, tiempo, lugar y posesión” (Hasty, R; Reardon, J; 1998)

- *Beneficio de forma;* se crea mediante la fabricación y otros medios de producción y cuando los bienes están clasificados en unidades más pequeñas.
- *Beneficio de tiempo;* este se crea cuando el minorista almacena bienes y los pone a disposición del consumidor cuando éste quiere hacer una compra.
- *Beneficio de lugar;* se crea mediante la transferencia de bienes desde el sitio de producción hasta el lugar en donde los consumidores pueden tener fácil acceso a ellos.
- *Beneficio de posesión;* se presenta cuando el minorista ayuda al consumidor a adquirir un bien o servicio mediante su compra, alquiler o intercambio.

Una de las características particulares de la venta de retail en una tienda es que el comprador sale a buscar al vendedor, el minorista espera que los clientes vayan a su tienda, es así que cuando

el cliente llega por lo general espera una satisfacción instantánea, esperan que los bienes estén a su disposición para llevarlos a casa al momento en que los compran. Cuando los clientes van hasta donde el vendedor, esperan servicio, un lugar apropiado y una selección de productos a un precio que puedan pagar.

A las tiendas minoristas se las puede clasificar según (Hasty, R; Reardon, J; 1998) como se detalla en la figura 5.

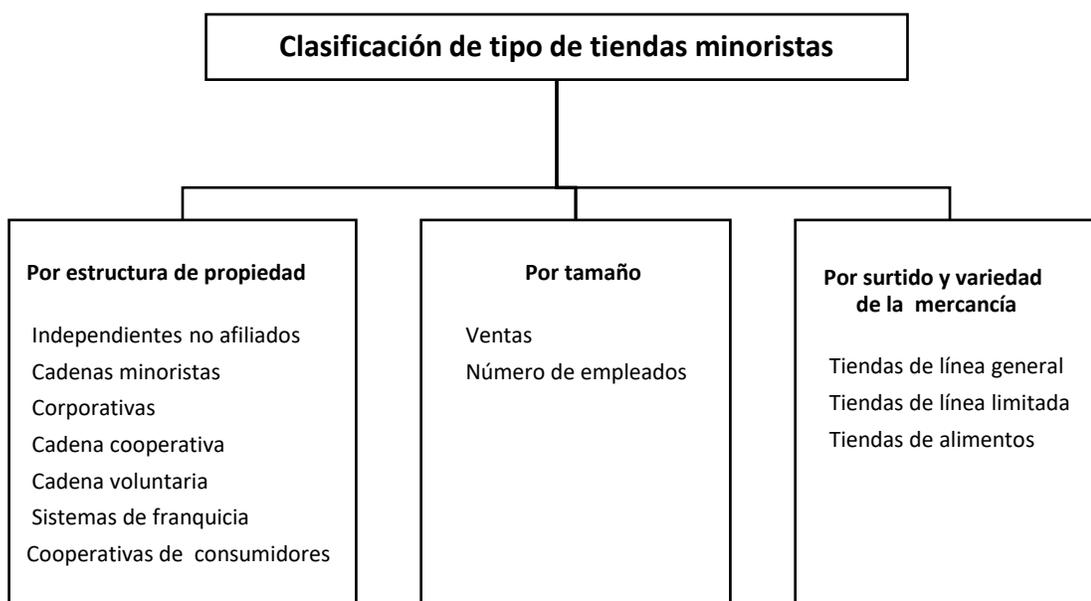


Figura 4 Clasificación de tiendas minoristas
Fuente (Hasty, R; Reardon, J; 1998)

La tecnología será la herramienta usada para ayudar a los minoristas a cumplir a cabalidad con el imperativo continuo de aumentar las ventas por metro cuadrado. Uno de los principales objetivos con el desarrollo tecnológico es que las tiendas tengan menos productos en inventario en el sitio de venta, pero seguirán prestando un servicio efectivo a los clientes.

3.1.1.7.1 Estrategia de la venta de retail

En (Fischer, L; Espejo, J; 2007) habla que la estrategia principal para la venta de retail es la creación de ventaja competitiva, el desarrollo de una estrategia define el campo de acción de la

tienda en relación con sus competidores. En el libro (Hasty, R; Reardon, J; 1998) encontramos cinco dimensiones principales de una estrategia de venta de retail:

- **Ubicación;** Para la mayoría de clientes la competencia es buena ya que permite tener una mejor oferta de productos y servicios a precios más bajos.
- **Mercancía;** Planear la mezcla de mercancía es uno de los aspectos más importantes de las operaciones de la tienda, pues esta representa el rango total de productos que el minorista ofrece a los clientes potenciales. La variedad en la mercancía se refiere al número de tipos o clases de productos diferentes con que cuenta un minorista, el surtido se refiere a la cantidad de artículos de una línea de producto o a cuantos estilos y marcas diferentes tiene un minorista en cada línea de producto, las tiendas tienen que elegir entre contar con un gran surtido o una gran variedad, ya que tratar de tener ambas cosas aumentan los costos de inventarios lo que puede llegar a convertirlas en poco competitivas.
- **Precio;** La posición de manejar márgenes altos y rotación alta es casi imposible, ya que cuando los precios aumentan para mejorar los márgenes, los clientes buscan a la competencia; también es imposible mantener márgenes bajos y rotación baja porque no se tendrá suficiente dinero.
- **Servicio;** Determinar el nivel de servicio que es apropiado para cada tienda, incluye la calidad de los servicios que prestan los vendedores y la cantidad de servicios asociados que brinda la tienda; por ejemplo, la envoltura de regalos y servicio de entrega.
- **Comunicaciones;** Esta es la estrategia general para alcanzar a los consumidores potenciales, “existen cuatro dimensiones estratégicas que pueden examinarse en la política de comunicación: alcance, frecuencia, contenido y personalización” (Fischer, L; Espejo, J; 2007)

Estas dimensiones se apoyan para lograr el posicionamiento estratégico en:

- a) **Operaciones;** actividades que garanticen que el personal brinde servicios consistentes con la estrategia de venta y que la apariencia de la tienda se mantenga para brindar una atmósfera atrayente.

- b) **Compras/logística;** estas pueden ser utilizadas como una ventaja competitiva; hablamos de niveles de inventario en la tienda, garantía que los bienes sean aquellos que el consumidor quiere, las tiendas se definen por los productos que deciden tener en existencia.
- c) **Investigación del mercado;** el contacto con las expectativas, los deseos, las necesidades, los gastos y el comportamiento del consumidor.
- d) **Financiamiento**
- e) **Tecnología;** el uso de la tecnología ha penetrado tanto en la venta al detalle que con frecuencia se considera una decisión estratégica

3.1.1.7.2 Planeación venta de retail

La venta de retail es un campo dinámico y desafiante, es por eso la importancia de una planificación, según (Hasty, R; Reardon, J; 1998) las ventajas que da la planeación son:

- Ayuda al gerente a manejar el futuro de la manera más efectiva.
- Ayuda a los gerentes a dirigir esfuerzos hacia los principales objetivos del negocio.
- Ayuda a unificar y coordinar las actividades de cada persona implicada.
- Minimiza los costos y ayuda a un gerente a solucionar los problemas en una forma creativa.
- Permite al gerente evaluar, medir el progreso y el desempeño.

Sin planeación, no existe un curso predeterminado de acción, y sin éste, los gerentes no pueden saber qué hacer, dónde hacerlo y por qué hacerlo.

La venta de retail forma parte importante de este estudio, ya que es el lugar en donde se basa el mismo, retail de artículos deportivos, en este proceso de venta la empresa beneficia a los consumidores, ya que realizan la función de selección al evaluar todos los diferentes productos disponibles y ofrecerlos en una distribución óptima para escoger, es por eso la importancia de una planificación y este se basa en el entendimiento del funcionamiento de un retail; así se podrá analizar los resultados obtenidos y visualizar desde el punto que la tecnología se encuentra íntimamente relacionada por ser la herramienta usada en general para cumplir el objetivo de aumentar las ventas por metro cuadrado, hay que considerar que una meta como empresa y globalmente es que las tiendas tengan menos productos en inventario en el sitio de venta, pero seguirán prestando un servicio efectivo a los clientes, es decir el objetivo principal del estudio.

Por otro lado, tratar el tema de la clasificación de las tiendas de retail, se torna indispensable cuando la empresa de la que está realizando este estudio es una franquicia de una empresa multinacional con operación en todos los países del mundo, de ahí la importancia de este tema.

3.1.2 Inteligencia de Negocios

La inteligencia de negocios (Business Intelligence, BI de sus siglas en inglés) es un conjunto de tecnologías y procesos que permiten a personas de todos los niveles en una organización, tener acceso a datos y su análisis.

“La inteligencia de negocios permite a las personas de todos los niveles de una organización tener acceso, interactuar y analizar información para administrar el negocio, mejorar el rendimiento, descubrir oportunidades y operar eficientemente” (Howson, C; 2009)

La Inteligencia de Negocios BI es una herramienta con la cual diferentes tipos de organizaciones, pueden soportar la toma de decisiones basadas en información precisa y oportuna; garantizando la generación del conocimiento necesario que permita escoger la alternativa que sea más conveniente para el éxito de la empresa (Rosado, A; Rico, Dewar; 2010)

Por lo tanto, se puede decir que una compañía utiliza inteligencia de negocios cuando ponga en manos de los usuarios herramientas para llegar a la información y hacerla útil.

En la figura 5, se visualiza la arquitectura general en BI, donde se inicia con el nivel operativo transaccional, sigue Data Warehouse, y luego se pasa a la capa de análisis, explotación y finalmente la llegada a los usuarios finales.

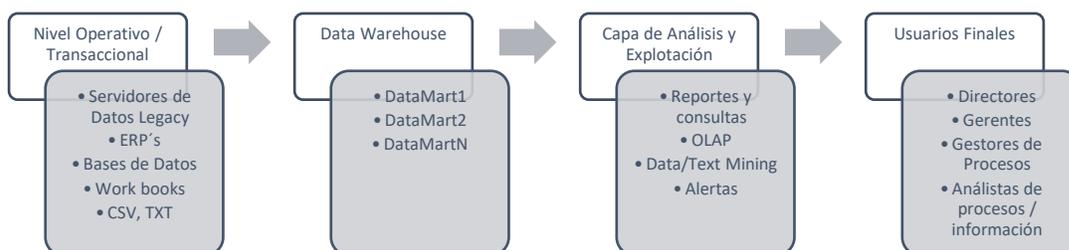


Figura 5 Arquitectura general de soluciones de Business Intelligence
Fuente (Latinos BI)

De acuerdo a (Howson, C; 2009) la inteligencia de negocios atraviesa todas las funciones e industrias, en párrafos posteriores se hablará de las distintas aplicaciones de BI, es así que se puede asegurar que BI llega a todas las personas en una compañía y más allá, a clientes y proveedores. Sin embargo, como se mencionó anteriormente, la inteligencia de negocios solamente puede proporcionar valor comercial cuando es manejada con eficacia por las personas, es decir que un mejor acceso a la información no afecta el desempeño de la compañía; la diferencia está en que hacen las compañías con la información.

El poder competitivo de una empresa se puede basar en la calidad y cantidad de la información que sea capaz de usar en la toma de decisiones; mediante la implementación de Inteligencia de Negocios se proporcionan las herramientas necesarias para aprovechar los datos almacenados en las bases de datos de los sistemas transaccionales para utilizar la información como respaldo a las decisiones, reduciendo el efecto negativo que puede traer consigo una mala determinación.

BI aporte en varias áreas:

- **BI para administración y control**

“En su sentido más básico la inteligencia de negocios proporciona información a los gerentes para saber qué sucede en el comercio. Careciendo de inteligencia de negocios, los gerentes pueden hablar de cómo “vuelan a ciegas”, sin perspectiva, hasta la publicación de los números financieros trimestrales” (Howson, C; 2009)

Con inteligencia de negocios, la información es accesible de manera oportuna y flexible, para proporcionar una perspectiva de:

- Como las ventas tienen seguimiento en varias regiones y por varias líneas de producto.
- Si los gastos concuerdan con un plan o salen del presupuesto.
- Si las capacidades de depósito están en niveles óptimos.
- Si los trámites de ventas están donde deben estar.

La inteligencia de negocios permite además a los usuarios, investigar detalles subyacentes, para determinar por qué las métricas están fuera de la meta y tomar medidas para mejorar la situación.

- **BI para mejorar el desempeño comercial**

Usada eficazmente la BI, permite a las organizaciones mejorar su desempeño, este se mide por un número de indicadores financieros tales como ingresos, margen, rentabilidad, costo de servicio y demás. En negocios, las ganancias de desempeño pueden lograrse mejorando la velocidad de respuesta para campañas en particular, al identificar características de los clientes más atentos, eliminar campañas ineficaces ahorra millones de dólares a las compañías anualmente.

“La inteligencia de negocios permite a las compañías aumentar sus ganancias, con ventas adicionales de productos a clientes existentes” (Howson, C; 2009)

Una señal clave de inteligencia de negocios exitosa, es el grado en que impacta el desempeño comercial, objetivo de este trabajo, pero para medir el impacto comercial de BI no es tarea sencilla ya las mejoras en el desempeño son atribuibles a factores más allá de inteligencia de negocios.

- **BI de operaciones**

Mientras la inteligencia de negocios va avanzando dentro de una empresa va jugando un papel más crítico en las funciones diarias, por lo cual el acceso a la información se vuelve una labor fundamental. Se encontró los siguientes ejemplos en (Rosado, A; Rico, Dewar; 2010) y (Howson, C; 2009):

- Hospitales y salas de emergencia utilizarían inteligencia de negocios para determinar los niveles óptimos de personal durante los periodos de máxima demanda.
- Walt Disney World's Magic Kingdom utiliza inteligencia de negocios, con el fin de minimizar el tiempo de espera en las colas para sus distintos juegos, es por lo cual, que genera Fast Pass gratuitos mediante su aplicación, las herramientas de inteligencia de

negocios monitorean tiempos de espera en los juegos más populares, para balancear el número de boletos emitidos en ciertos periodos durante el día.

- **BI para mejorar procesos**

Las operaciones de un negocio pueden estar conformadas por procesos individuales. “La BI puede conformar decisiones que los individuos toman en cada paso de un proceso. También puede usarse para ayudar a modernizar un proceso, al medir cuánto tardan los subprocesos e identificar áreas para mejoría” (Howson, C; 2009)

BI puede analizar entradas, tiempo y salidas para cada paso del proceso, es así que puede ayudar a identificar embotellamientos del proceso. Se encontró el siguiente ejemplo en (Rosado, A; Rico, Dewar; 2010)

- En Norway Post, los trabajadores del correo monitorean el número de paquetes y cartas clasificadas por hora, cualquier cambio de estas métricas puede llevar a una revisión del proceso para optimizar el flujo de trabajo.

- **BI para mejorar servicio al cliente**

La inteligencia de negocios puede ayudar a las compañías a brindar altos niveles de servicio al cliente, proporcionando procesamiento de pedidos, aprobación de préstamos, manejo de problemas y demás. Se encontró el siguiente ejemplo en (Rosado, A; Rico, Dewar; 2010):

- Continental Airlines utiliza inteligencia de negocios para monitorear que tan llenas están las cabinas de la clase de negocios y asegurarse que sus clientes más valiosos obtengan ascensos gratuitos cuando el cupo lo permite.

3.1.2.1 Desarrollo de Inteligencia de Negocios

Desde una perspectiva comercial, el mercado de inteligencia de negocios es un mercado de \$20 mil millones, según la firma IDC. Entonces, el índice de crecimiento en los últimos años ha sido dentro de un rango de 11%, considerando que los precios de unidades para muchos componentes de inteligencia de negocios han bajado y, en contraste con otros mercados de tecnología informática, cuyo crecimiento ha disminuido (Howson, C; 2009)

El terreno del comercio ha cambiado dramáticamente en los últimos 20 años en base a la tecnología por el desarrollo global que se ha tenido, el cambio rápido de la tecnología ha sido un impulsor de este ritmo acelerado en los negocios; y de esta manera han posibilitado la llegada de la inteligencia de negocios a todas las industrias, podemos ver empresas que operan y compiten 24/7.

La riqueza de información al alcance de consumidores y negocios, pone gran presión sobre los precios y hace de la masa de clientes una constante amenaza a través de las industrias. El ritmo del cambio es rápido, las compañías compiten en tiempo de mercado, innovaciones de productos, todo a un ritmo frenético, por ello los negocios no pueden darse el lujo de ignorar lo que está sucediendo, interna y externamente, así como a los niveles de detalle que deben llegar, es así que el papel de la inteligencia de negocios se ha elevado como una tecnología informática que da a las compañías una ventaja competitiva.

Cuando la inteligencia de negocios es desplegada eficazmente, dicha información se convierte en un valor estratégico que debe explotarse.

Según (Howson, C; 2009) la inteligencia de negocios puede ser:

- **BI basado en Web;** la inteligencia de negocios basada en Web ha permitido que las herramientas se desplieguen a través de intranets y extranets corporativos, hacia miles de empleados y clientes externos en cuestión de horas.
- **BI suites y juego de herramientas.** Las herramientas para inteligencia de negocios tienen múltiples componentes de entrada, tales como herramientas de interrogatorios comerciales, OLAP y tableros.

A la inteligencia de negocios en (Howson, C; 2009) se lo nombra como un catalizador para el cambio, es por ello que cualquiera que quiera preservar el status quo, no aceptará una iniciativa de BI, es por eso que se puede comparar la implementación de Inteligencia de Negocios en una empresa con una batalla en la cual se da situaciones como:

- Luchas de poder entre IT y el negocio, cuando cualquiera de los dos pierde áreas de control o no están de acuerdo en el alcance y acercamiento.
- Empleos eliminados cuando los desarrolladores de reportes personalizados ya no son necesarios.
- Gerentes de mercadotecnia despedidos, cuando los indicadores muestran la mala ejecución de campañas.
- Software y tecnología que no siempre funciona como se espera ya que no son de fácil entendimiento para los usuarios.

Implementar herramientas de BI dentro de la organización permite soportar las decisiones que se toman; al nivel interno ayuda en la gestión del personal y del lado externo produce ventajas sobre sus competidores.

Existen ocasiones en las cuales no se pueden lograr todos los beneficios que tiene BI; debido al proceso que lleva consigo implementar un proyecto de estas características, se puede cometer errores en la definición del planteamiento de las necesidades de conocimiento de la empresa; el no determinar la magnitud de los problemas de información a solucionar generalmente repercute en el desempeño del proyecto, de aquí la importancia del análisis de BI en este estudio.

Para implementar una arquitectura de BI, se analizan según (Rosado, A; Rico, Dewar; 2010) cuatro valores básicos, pero dependiendo de cada organización puede incluir los que vayan en relación con su propia estrategia:

- Adaptabilidad continua
- Trabajo conjunto
- Jerarquías Flexibles
- Personas antes que procesos

Según (Rosado, A; Rico, Dewar; 2010) la tecnología de bases de datos ha respondido a los avances con una nueva arquitectura: el almacén de datos (*data warehouse*), este se trata de un repositorio de fuentes heterogéneas de datos, integrados y organizados bajo un esquema unificado para facilitar su análisis y dar soporte a la toma de decisiones. Esta tecnología incluye operaciones

de procesamiento analítico en línea (*On-Line Analytical Processing, OLAP*), es decir, técnicas de análisis como pueden ser el resumen, la consolidación o la agregación, así como la posibilidad de ver la información desde distintas perspectivas.

Sin embargo, a pesar de que las herramientas OLAP soportan cierto análisis descriptivo y de sumariación que permite transformar los datos en otros datos agregados o cruzados de manera sofisticada, no generan reglas, patrones, pautas, es decir, conocimiento que pueda ser aplicado a otros datos.

Todos estos problemas y limitaciones de las aproximaciones clásicas han hecho surgir la necesidad de una nueva generación de herramientas y técnicas para soportar la extracción de conocimiento útil desde la información disponible, y que se engloban bajo la denominación de minería de datos, de la misma forma podemos mencionar que la estadística es en cierto modo, la madre de la minería de datos. (Rosado, A; Rico, Dewar; 2010)

3.1.2.2 Data Warehouse

(Rosado, A; Rico, Dewar; 2010) conceptualiza a Data Warehouse como “El proceso de extraer datos de distintas aplicaciones (internas y externas), para que una vez depurados y especialmente estructurados sean almacenados en un depósito de datos consolidado para el análisis del negocio”.

Data Warehouse requiere una combinación de metodologías, técnicas, hardware y los componentes de software que proporcionen en conjunto una infraestructura para soportar el proceso de información. La estructura que se defina debe reflejar las necesidades y características del negocio, sus departamentos, equipos de trabajo y directivos, esto permitirá responder a interrogantes generados al tratar de tomar las decisiones y con el tiempo se va convirtiendo en la memoria corporativa; describiendo el pasado y el presente de la empresa. Data Warehouse desglosa, resume, ordena y compara, pero no descubre, ni predice.

Implementar una solución de este tipo, ocasiona un costo que no todas las organizaciones están dispuestas a pagar, pero deben tener claro que el retorno de la inversión del proyecto equiparando inversión contra beneficios.

Data warehouse, al ser un depósito de datos consolidado para el análisis del negocio, necesita tomar datos de distintas fuentes, internas y externas, y como las características de las empresas son diferentes la cantidad de registros almacenados en algunas de ellas puede llegar a ser de proporciones exponenciales; es así que (Rosado, A; Rico, Dewar; 2010) define que se necesita de procesos que optimicen los tiempos de extracción, transformación y transferencia de los datos del sistema de información a la fuente de datos.

La estructura de data Warehouse tiene ciertas distinciones. Hay niveles diferentes de esquematización y detalle que delimitan el data warehouse. La estructura de data warehouse se muestra en la Figura 6, donde se muestran los diferentes componentes conceptualizados por (Rosado, A; Rico, Dewar; 2010):

- Detalle de datos actuales; reflejan las ocurrencias más recientes, las cuales son de gran interés, es voluminoso.
- Detalle de datos antiguos; la data antigua es aquella que se almacena sobre alguna forma de almacenamiento masivo, no es frecuentemente accesible y se almacena a un nivel de detalle, consistente con los datos detallados actuales.
- Datos ligeramente resumidos; la data ligeramente resumida es aquella que proviene desde un bajo nivel de detalle encontrado al nivel de detalle actual.
- Datos completamente resumidos; estos datos son compactos y fácilmente accesibles.
- Meta data; juega un rol especial y muy importante en data warehouse y es usada como un directorio para ayudar al analista a ubicar los contenidos de data warehouse, una guía para el mapping de datos de cómo se transforma, del ambiente operacional al de data warehouse. Una guía de los algoritmos usados para la esquematización entre el detalle de datos actual, con los datos ligeramente resumidos y éstos, con los datos completamente resumidos, etc.

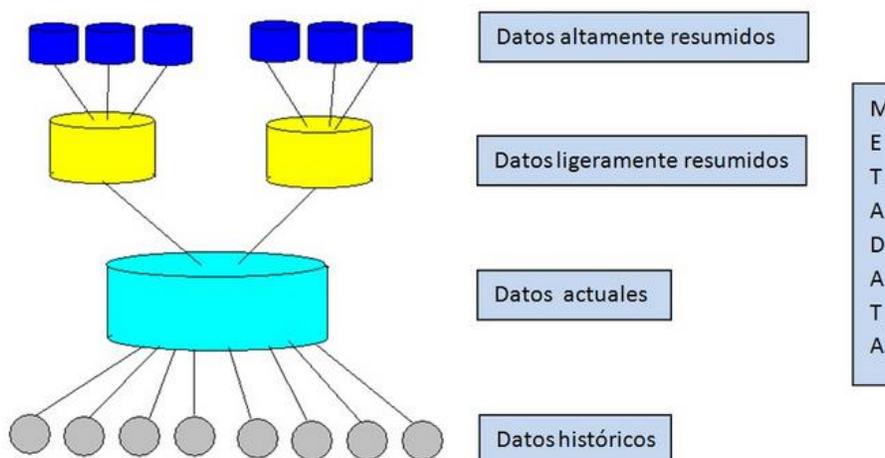


Figura 6 Estructura de Data Warehouse
Fuente (Latinos BI)

De acuerdo a (Rosado, A; Rico, Dewar; 2010) “El uso de Data Warehouse es tan amplio que llega a diferentes tipos de organizaciones y distintos temas de interés”, puede ser implementado con conceptos administrativos, en esta área ayuda en la identificación de elementos de cambio que definan una nueva manera de hacer negocios, en donde la competencia debe estar orientada a trabajar no sólo de forma aislada, sino en colaboración con los diversos grupos de interés o actores de la industria, buscando referencias diferenciadoras para alcanzar el éxito.

3.1.3 Minería de Datos

La minería de datos se crea por la aparición de nuevas necesidades y, especialmente, por el reconocimiento de un nuevo potencial: el valor, hasta ahora sub utilizado de la gran cantidad de datos almacenados informáticamente en los sistemas de información de instituciones, empresas, gobiernos y particulares. (Hernández, J; Ramírez, M; Ferri, C; 2008)

Los datos pasan de ser un producto a ser una materia prima que hay que explotar para obtener el verdadero producto elaborado, el conocimiento; este realmente es valioso ya que ayudan para la toma de decisiones sobre el ámbito en que se han recopilado o extraído los datos.

El aumento del volumen y la variedad de información que se encuentra informatizada en base de datos digitales y otras fuentes ha crecido con el pasar de los años y sigue aumentando hasta

hoy en día, por lo cual gran parte de la información es histórica, esta información es útil para explicar el pasado, entender el presente y predecir la información futura.

El método tradicional de convertir los datos en conocimiento consiste en un análisis e interpretación realizada de manera manual, este es impracticable en dominios donde el volumen de los datos crece exponencialmente, ya que la enorme abundancia de datos no permite que la capacidad humana los comprenda completamente sin la ayuda de herramientas potentes. (Riqueleme, J; Ruiz, R; Gilbert K; 2006)

Muchas decisiones importantes se realizan, no sobre la base de la gran cantidad de datos disponibles, sino siguiendo la propia intuición del usuario y en base a su conocimiento empírico, es así que (Hernández, J; Ramírez, M; Ferri, C; 2008) define como el principal objetivo de la minería de datos; es eliminar el empirismo a la hora de tomar decisiones.

Según (Hernández, J; Ramírez, M; Ferri, C; 2008) la minería de datos se distingue de las aproximaciones antes descritas en BI, porque no obtiene información extensional, datos, sino conocimiento y, además, el conocimiento no es, generalmente, una parametrización de ningún modelo preestablecido o intuitivo por el usuario, sino que es un modelo novedoso y original, extraído completamente por la herramienta.

El resultado de la minería de datos son conjuntos de reglas, ecuaciones, árboles de decisión, redes neuronales, grafos probabilístico, los cuales pueden usarse para, por ejemplo, responder a cuestiones como ¿existe un grupo de clientes que se comporta de manera diferenciada?, ¿qué secuencia de tratamientos puede ser más efectiva para este nuevo síndrome?, ¿existen asociaciones entre los factores de riesgo para realizar un seguro de automóvil?, ¿cómo califico automáticamente los mensajes de correo entre más o menos susceptible de ser spam?” (Hernández, J; Ramírez, M; Ferri, C; 2008)

(Witten; Frank; 2000) definen la minería de datos como “El proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. Es decir, la tarea fundamental de la minería de datos es encontrar modelos inteligibles a partir de los datos” por lo cual para que este proceso sea efectivo debería ser automático o semi-automático y el uso de los patrones descubiertos debería ayudar a tomar decisiones más seguras que reporten, por tanto, algún beneficio a la organización.

La minería de datos, permite la gestión en tiempo real de manera eficaz, es una herramienta aplicable a cualquier tipo de empresa. Una amplia gama de compañías puede tener aplicaciones exitosas con ella.

Hay beneficios asociados a la minería de datos, (Hernández, J; Ramírez, M; Ferri, C; 2008) describe varios: incremento de los resultados como consecuencia del aumento de la cuota de mercado; fidelización de la clientela dada una mejor respuesta a sus requerimientos; mejora del rendimiento; reducción del factor riesgo; optimización de las estrategias y toma de decisiones y optimización de la gestión, maximizando rentabilidades.

(Hernández, J; Ramírez, M; Ferri, C; 2008) define que son dos los retos de la minería de datos:

- Trabajar con grandes volúmenes de datos, procedentes mayoritariamente de sistemas de información, con los problemas que ello conlleva.
- Usar técnicas adecuadas para analizar los datos y extraer conocimiento novedoso y útil.

En muchos casos la utilidad del conocimiento minado está íntimamente relacionada con la comprensibilidad del modelo inferido, podríamos decir que el objetivo de la minería de datos es convertir datos en conocimiento.

3.1.3.1 Tipos de Datos

La minería de datos, puede aplicarse a cualquier tipo de información, siendo las técnicas de minería diferentes para cada una de ellas.

Desde el punto de vista de las técnicas de minería de datos más habituales nos interesa distinguir sólo entre dos tipos de datos, numéricos y categóricos.

- Los atributos numéricos contienen valores enteros o reales. Por ejemplo, atributos como el salario o la edad son numéricos.
- Los atributos categóricos o nominales toman valores en un conjunto finito y preestablecido de categorías. Por ejemplo, atributos como el sexo (H, M), el nombre del departamento (Gestión, Marketing, Ventas) son categóricos.

Esta diferenciación es importante en el estudio, porque en los siguientes capítulos para cada herramienta se deberá tomar la decisión de que datos nos darán los resultados buscados y como trabajar en ellos.

3.1.3.2 Proceso de descubrimiento de conocimiento en bases de datos

Todo apunta a que existe la necesidad de metodologías de análisis inteligente de datos, las cuales permiten descubrir conocimiento útil a partir de los datos. El término KDD (iniciales de Knowledge Discovery in Databases).

“KDD es el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos” (Hernández, J; Ramírez, M; Ferri, C; 2008).

El conocimiento extraído es válido, novedoso, potencialmente útil y comprensible, la figura 8 muestra el proceso de KDD, estos sistemas según (Hernández, J; Ramírez, M; Ferri, C; 2008) permiten la selección, limpieza, transformación y proyección de los datos; analizar los datos para extraer patrones y modelos adecuados; evaluar e interpretar los patrones para convertirlos en conocimiento; consolidar el conocimiento resolviendo posibles conflictos con conocimiento previamente extraído; y hacer el conocimiento disponible para su uso.

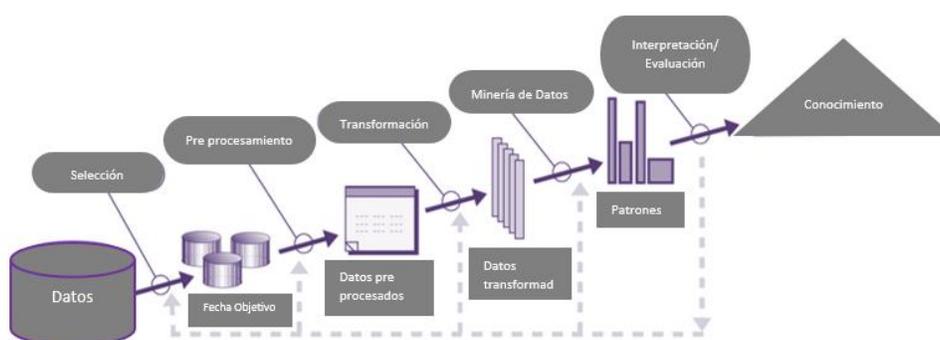


Figura 7 Proceso de KDD

Fuente (Hernández, J; Ramírez, M; Ferri, C; 2008)

En realidad, los términos minería de datos y KDD son a menudo confundidos como sinónimos. En general se acepta que la Minería de Datos es un paso particular en el proceso consistiendo en la aplicación de algoritmos específicos para extraer patrones, modelos de los datos.

Otros pasos en el proceso KDD, son la preparación de los datos, la selección y limpieza de los mismos, la incorporación de conocimiento previo, y la propia interpretación de los resultados de minería.

“Las tareas de la fase de minería de datos pueden ser descriptivas, (descubrir patrones interesantes o relaciones describiendo los datos), o predictivas (clasificar nuevos datos basándose en los anteriormente disponibles)” (Hernández, J; Ramírez, M; Ferri, C; 2008)

La minería de datos debe trabajar mano a mano con los almacenes de datos, sobre todo en los casos de volúmenes de datos muy grandes o de inter-relaciones entre los datos complejas, es decir, que no puedan ser expresadas en una tabla plana. El KDD se centra en el proceso global de descubrir conocimiento de grandes volúmenes de datos, incluyendo el almacenaje y acceso a tales datos, escalado de algoritmos a bases de datos masivas, interpretación y visualización de resultados, y el modelado y soporte de la interacción general hombre máquina.

El proceso KDD es interactivo e iterativo conteniendo los siguientes pasos según (Riqueleme, J; Ruiz, R; Gilbert K; 2006):

- Comprender el dominio de aplicación: este paso incluye el conocimiento relevante previo y las metas de la aplicación.
- Extraer la base de datos objetivo: recogida de los datos, evaluar la calidad de los datos y utilizar análisis exploratorio de los datos para familiarizarse con ellos.
- Preparar los datos: incluye limpieza, transformación, integración y reducción de datos de dimensionalidad y numerosidad. Se intenta mejorar la calidad de los datos a la vez que disminuir el tiempo requerido por el algoritmo de aprendizaje aplicado posteriormente.
- Minería de datos: como se ha señalado anteriormente, este es la fase fundamental del proceso, está constituido por una o más de las siguientes funciones, clasificación, regresión, clustering, resumen, recuperación de imágenes, extracción de reglas, etc.

- Interpretación: explicar los patrones descubiertos, así como la posibilidad de visualizarlos.
- Utilizar el conocimiento descubierto: hacer uso del modelo creado

3.1.3.3 Aplicaciones

La aplicación de la Minería de Datos, además de permitir el descubrimiento del conocimiento en general, como se mencionó se utiliza en varios campos. Entre las principales aplicaciones de la minería de datos relacionadas al tema de estudio, son las siguientes descritas en (Hernández, J; Ramírez, M; Ferri, C; 2008):

- Aplicaciones financieras y banca
 - Obtención de patrones de uso fraudulento de tarjetas de crédito.
 - Determinación del gasto en tarjeta de crédito por grupos.
 - Cálculo de correlaciones entre indicadores financieros.
 - Identificación de reglas de mercado de valores a partir de históricos.
 - Análisis de riesgos en créditos.
- Análisis de mercado, distribución y, en general, comercio
 - Análisis de la cesta de la compra (compras conjuntas, secuenciales, ventas cruzadas, señuelos, etc.).
 - Evaluación de campañas publicitarias.
 - Análisis de la fidelidad de los clientes. Reducción de fuga.
 - Segmentación de clientes.
 - Estimación de stocks, de costes, de ventas, etc.
 - Reposición de productos.
- Seguros y salud privada
 - Determinación de los clientes que podrían ser potencialmente caros.
 - Análisis de procedimientos médicos solicitados conjuntamente.
 - Predicción de qué clientes contratan nuevas pólizas.
 - Identificación de patrones de comportamiento para clientes con riesgo.

- Identificación de comportamiento fraudulento.
- Predicción de los clientes que podrían ampliar su póliza para incluir procedimientos extras (dentales, ópticos).

- Procesos industriales
 - Extracción de modelos sobre el comportamiento de compuestos.
 - Detección de piezas con trabas. Modelos de calidad.
 - Predicción de fallos y accidentes.
 - Estimaciones de composiciones óptimas en mezclas.
 - Extracción de modelos de costes.
 - Extracción de modelos de producción.

- Medicina
 - Identificación de patologías. Diagnóstico de enfermedades.
 - Detección de pacientes con riesgo de sufrir una patología concreta.
 - Gestión hospitalaria y asistencial. Predicciones temporales de los centros asistenciales para el mejor uso de recursos, consultas, salas y habitaciones.
 - Recomendación priorizada de fármacos para una misma patología.
 - Efectividad de los tratamientos.

- Biología, bioingeniería y otras ciencias
 - Predecir si un compuesto químico causa cáncer.
 - Clasificación de cuerpos celestes.
 - Predicción de recorrido y distribución de inundaciones.

- Telecomunicaciones
 - Establecimiento de patrones de llamadas.
 - Detección de fraude.

- Otras áreas

- Correo electrónico y agendas personales: clasificación y distribución automática de correo, detección de correo spam, gestión de avisos, análisis del empleo del tiempo.
- Recursos humanos: selección de empleados.
- Web: análisis del comportamiento de los usuarios, detección de fraude en el comercio electrónico, análisis de los blogs de un servidor web.
- Turismo: determinar las características socioeconómicas de los turistas en un determinado destino o paquete turístico, identificar patrones de reservas, etc.
- Política: diseño de campañas políticas, estudios de tendencias de grupos, etc.
- Comercio y banca: segmentación de clientes, previsión de ventas, análisis de riesgo.
- Recuperación de información no numérica: minería de texto, minería web, búsqueda e identificación de imagen, video, voz y texto de bases de datos multimedia.
- Ciencias Sociales: Estudio de los flujos de la opinión pública. Planificación de ciudades: identificar barrios con conflicto en función de valores socio demográfico.

3.1.3.4 Sistemas y herramientas de minería de datos

En base a los conceptos analizados podemos mencionar que la diversidad de disciplinas donde se puede aplicar minería de datos se da por una gran variedad de sistemas de minería de datos, cada uno de ellos posee unas características apropiadas para realizar determinadas tareas o para analizar cierto tipo de datos, (Hernández, J; Ramírez, M; Ferri, C; 2008) describe:

- Tipo de base de datos minada: sistemas textuales, multimedia, espaciales o web.
- Tipo de conocimiento minado: sistemas que buscan regularidades en los datos, patrones, frente a los que analizan las irregularidades, excepciones.
- Tipo de funcionalidad y de técnica: pueden clasificar basándose en su funcionalidad, clasificación, agrupamiento, o por los métodos de análisis de los datos empleados, técnicas estadísticas, redes neuronales, etc.
- Tipo de aplicación: dos clases de sistemas: los sistemas de propósito general y los sistemas específicos.

Existen cientos de productos de minería de datos y de compañías de consultoría, en base a (Hernández, J; Ramírez, M; Ferri, C; 2008) pueden resaltarse por su mayor expansión las

siguientes: SAS con SAS Script y SAS Enterprise Miner; SPSS y el paquete de minería Clementine; IBM con IntelligentMiner; Microsoft incluye características de minería de datos en el MS SQL Server; otras compañías son Oracle, Angoss y Kxen. En la línea del software libre Weka es un producto con mayor orientación a las técnicas provenientes de la IA, pero de fuerte impacto.

También tenemos a RapidMiner que según KDnuggets, en el 2009 ocupó el segundo lugar en herramientas analíticas y de minería de datos utilizadas para proyectos reales y en el 2010 se ubicó en el primer lugar.

3.1.3.5 Tipos de modelos

La minería de datos según (Riqueleme, J; Ruiz, R; Gilbert K; 2006) “Tiene como objetivo analizar los datos para extraer conocimiento, este conocimiento puede ser en forma de relaciones, patrones o reglas que se infieren de los datos y se desconocen; estas relaciones o resúmenes constituyen el modelo de los datos analizados”.

(Riqueleme, J; Ruiz, R; Gilbert K; 2006) asegura que los modelos pueden ser de dos tipos: predictivos y descriptivos. Los modelos predictivos pretenden estimar valores futuros o desconocidos de variables de interés, que se denominan variables objetivo o dependientes, usando otras variables o campos de la base de datos, a las que se refiere como variables independientes o predictivas. Los modelos descriptivos, en cambio, identifican patrones que explican o resumen los datos, es decir, sirven para explorar las propiedades de los datos examinados, no para predecir nuevos datos.

Por ejemplo, una agencia de viaje desea identificar grupos de personas con unos mismos gustos, con el objeto de organizar diferentes ofertas para cada grupo y poder así remitirles esta información; para ellos analiza los viajes que han realizado sus clientes e infiere un modelo descriptivo que caracteriza estos grupos, este ejemplo se aplica al tema de estudio cuando queremos formar perfiles del consumidor por cada tienda, para acorde a esas necesidades específicas ofertar el producto.

Algunas tareas de minería de datos que producen modelos predictivos son la clasificación y la regresión, y las que dan lugar a modelos descriptivos son el agrupamiento, las reglas de asociación y el análisis correlacional.

En base a los conceptos encontrados se concluye que la minería de datos es la construcción de un modelo que ajustado a unos datos proporciona un conocimiento, es así que (Hernández, J; Ramírez, M; Ferri, C; 2008) distingue dos pasos en una tarea de minería de datos, por un lado, la elección del modelo y por otro el ajuste final de éste a los datos.

La elección del modelo se determinada por dos condicionantes: el tipo de los datos y el objetivo que se quiera obtener, también es importante la elección del nivel de comprensibilidad que se quiera obtener del modelo final, ya que hay modelos fáciles de explicar al usuario como por ejemplo las reglas de asociación y otros con claras dificultades como las redes neuronales.

3.1.3.5.1 Reglas de Asociación y Dependencia

Estas reglas expresan patrones de comportamiento entre los datos en función de la aparición conjunta de valores de dos o más atributos, la característica principal de estas reglas es que tratan con atributos nominales, a diferencia de los análisis correlacionales que lo hacen con los numéricos. (Riqueleme, J; Ruiz, R; Gilbert K; 2006)

Las reglas de asociación son una manera popular de expresar patrones de datos de una base de datos, de igual forma para conocer el comportamiento general de la base datos y de la información que pueda asistir en la toma de decisiones. Por ejemplo, en un supermercado podemos conocer qué productos suelen comprarse conjuntamente.

Las reglas de asociación “surgieron inicialmente para afrontar el análisis de las cestas de la compra de los comercios. Las diferentes cestas de la compra se pueden expresar formando una base de datos de una sola tabla” (Hernández, J; Ramírez, M; Ferri, C; 2008)

Se toma el siguiente ejemplo de (Hernández, J; Ramírez, M; Ferri, C; 2008), se utilizará las filas de la Tabla 2, que se refieren a una cesta de un supermercado, mientras que las columnas son cada uno de los productos en venta en el supermercado. La tabla contiene solo valores binarios.

Un 1 en la posición (ij) indica que la cesta i incorpora el producto j , mientras que un 0 indica que el cliente no ha adquirido el producto:

“Una regla de asociación es una proposición probabilística sobre la ocurrencia de ciertos estados en una base de datos” (Hernández, J; Ramírez, M; Ferri, C; 2008)

Tabla 2
Cesta de la compra / ejemplo

	Vino el cabezón	Gaseosa “Chispa”	Vino “Tío Paco”	Horchata “Xufer”	Biscochos “Goloso”	Galletas “Trigo”	Chocolate “La vaca”
T1	1	1	0	0	0	1	0
T2	0	1	1	0	0	0	0
T3	0	0	0	1	1	1	0
T4	1	1	0	1	1	1	1
T5	0	0	0	0	0	1	0
T6	1	0	0	0	0	1	1
T7	0	1	1	1	1	0	0
T8	0	0	0	1	1	1	1
T9	1	1	0	0	1	0	1
T10	0	1	0	0	1	0	0

Fuente (Hernández, J; Ramírez, M; Ferri, C; 2008)

Una regla de asociación tomando el ejemplo de (Hernández, J; Ramírez, M; Ferri, C; 2008):

SI biscochos “Goloso” **Y** horchata “Xufer” **ENTONCES** galletas “Trigo

De acuerdo a (Hernández, J; Ramírez, M; Ferri, C; 2008), una regla de asociación puede ser vista como reglas de la forma SI α ENTONCES β , donde α y β son dos conjuntos de ítems disjuntos. Otra forma muy utilizada de expresar una regla de asociación es $\beta \leftarrow \alpha$, o también $\alpha \rightarrow \beta$, el conjunto α recibe el nombre de predecesor de la regla, y a β se le denomina sucesor o consecuente.

Dada una regla de asociación, se puede trabajar con dos medidas para conocer la calidad de la regla: cobertura y confianza, es decir el interés de una regla de asociación está dado por su soporte y su confianza.

La cobertura, también denominada soporte de una regla se define como el número de instancias que la regla predice correctamente, se la suele expresar en porcentaje; es decir el soporte es la frecuencia de aparición del ítem en la colección de la combinación de productos involucrados en la regla. (Hernández, J; Ramírez, M; Ferri, C; 2008)

“La confianza, también llamada precisión mide el porcentaje de veces que la regla se cumple cuando se puede aplicar” (Hernández, J; Ramírez, M; Ferri, C; 2008),

Es así que podemos considerar que una regla es interesante si su soporte y su confianza son mayores o iguales que ciertos umbrales de mínimo soporte y mínima confianza especificados. Existe una clasificación de familias de reglas de asociación basándose en los siguientes criterios (Hernández, J; Ramírez, M; Ferri, C; 2008):

- **Tipos de los valores utilizados en las reglas:** podemos tener reglas que trabajan con atributos binarios que indican la presencia o ausencia de un ítem, este es el caso típico de la cesta de compra.
- **Dimensiones de los datos:** la regla *SI Comprar (vino “El cabezón”) ENTONCES Comprar (gaseosa “Chispa”)*, se refiere a tan sólo una dimensión, que corresponde al hecho de aparecer o no en la cesta de la compra. Podemos incrementar las dimensiones de una regla incluyendo por ejemplo, la dimensión tiempo, o la dimensión cliente. Un ejemplo de regla multi-dimensional sería *SI Comprar (vino “El cabezón”), Cliente (Juan), Tiempo (Marzo) ENTONCES Comprar (gaseosa “Chispa”)*.
- **Niveles de abstracción:** algunos sistemas o algoritmos permiten incorporar a las reglas diferentes niveles de abstracción representados por conceptos que aglutinan otros conceptos o ítems. Este tipo de reglas se conocen como reglas multi-nivel. Un ejemplo sería: *SI Comprar (vino) ENTONCES Comprar (gaseosa)*.
- **Instantáneas o secuenciales:** depende de si se considera relaciones en un instante de tiempo (por ejemplo, una determinada compra) o en una secuencia o serie (varias compras o visitas a una página web).

“Las reglas de asociación y dependencia se caracterizan precisamente por el hecho de que se expresan en forma de reglas de tipo “SI... ENTONCES...” y que los algoritmos tienen como objetivo primordial la eficiencia” (Hernández, J; Ramírez, M; Ferri, C; 2008)

3.1.3.5.1.1 Reglas de asociación

Las reglas de asociación se basan en su confianza y cobertura, es así que los algoritmos de aprendizaje trabajan en la búsqueda de reglas que cumplan unos requisitos mínimos en estas medidas.

Un algoritmo de aprendizaje de reglas de asociación muy simple y popular es el algoritmo *A priori*. El funcionamiento de este algoritmo se basa en la búsqueda de los conjuntos de ítems con determinada cobertura. Para ello, en primer lugar, se construyen simplemente los conjuntos formados por sólo un ítem que superan la cobertura mínima. Este conjunto de conjuntos se utiliza para construir el conjunto de conjuntos de dos ítems con la cobertura requerida. (Hernández, J; Ramírez, M; Ferri, C; 2008)

De acuerdo a (Hernández, J; Ramírez, M; Ferri, C; 2008), una vez se han seleccionado los conjuntos de ítems que cumplen con la cobertura mínima, el siguiente paso consiste en extraer de estos conjuntos de reglas las que tengan un nivel de confianza mínimo, es por esto que el aprendizaje de reglas de asociación se divide normalmente en dos fases: la extracción de los conjuntos de ítems que cumplan con la cobertura requerida desde datos, y la generación de las reglas a partir de estos conjuntos.

Según (Hernández, J; Ramírez, M; Ferri, C; 2008), la siguiente fase consiste en la creación de reglas a partir de los conjuntos de ítems frecuentes. Si sólo se busca reglas de asociación con un ítem el proceso es sencillo: de un conjunto de ítems de tamaño i , se crean i reglas colocando siempre un único ítem diferente en la parte derecha. Por ejemplo, si tenemos el conjunto de ítems *horchata “Xufer” Y bizcochos “Goloso” Y galletas “Trigo”* se construyen las reglas:

- **Si** bizcochos “Goloso” **Y** horchata “Xufer” **Entonces** galletas “Trigo” Cb=4, Cf = $\frac{3}{4}$
- **Si** bizcochos “Goloso” **Y** galletas “Trigo” **Entonces** horchata “Xufer” Cb=3, Cf = $\frac{3}{3}$
- **Si** galletas “Trigo” **Y** horchata “Xufer” **Entonces** bizcochos “Goloso” Cb=3, Cf = $\frac{3}{3}$

El cálculo de la confianza de las reglas se puede realizar, “dividiendo la cobertura del conjunto de ítems por la cobertura de la parte derecha, este último dato se calculó en la construcción de los conjuntos de ítems” (Hernández, J; Ramírez, M; Ferri, C; 2008),

Si se desea utilizar campos con más de un ítem se puede utilizar la siguiente propiedad, por ejemplo:

- **Si** bizcochos “Goloso” **Y** horchata “Xufer” **Entonces** galletas “Trigo” **Y** vino “Tío Paco”

Esta regla puede dividirse en dos reglas con un ítem en la parte derecha:

- **Si** bizcochos “Goloso” **Y** horchata “Xufer” **Entonces** vino “Tío Paco”
- **Si** bizcochos “Goloso” **Y** horchata “Xufer” **Entonces** galletas “Trigo”

La cobertura de ambas reglas será idéntica a la regla anterior, por tener la misma parte izquierda. En cuanto a la confianza, se puede asegurar que será igual o mayor a la regla con dos ítems en la parte derecha, ya que el número de registros que cumplan la parte izquierda en los que aparezca vino “Tío Paco” será mayor o igual al número de registros que cumplan la parte izquierda en los que aparezcan *vino “Tío Paco” y galletas “Trigo”*.

3.1.3.5.1.2 Reglas de dependencias

El aprendizaje de reglas de asociación es herramienta relevante en la búsqueda de patrones dentro de bases de datos, pero tiene limitaciones por estar demasiado orientada al problema de la extracción de patrones.

Según (Hernández, J; Ramírez, M; Ferri, C; 2008) una limitación es que no permite considerar reglas que contemplen la ausencia de un ítem:

Si gaseosa “Chispa” = 1 **Y** vino “El cabezón” = 1 **Entonces** vino “Tío Paco” = 0

Esta regla tiene una cobertura igual a 3, y una confianza del 100 por 100, por lo que se podría considerar una regla aceptable. Esta limitación se puede subsanar si se modifica el marco de aprendizaje de reglas de asociación, la idea sería considerar como ítem cada posible valor diferente de un determinado atributo.

Por ejemplo, (Hernández, J; Ramírez, M; Ferri, C; 2008) *vino “Tío Paco” = 1*, sería un ítem, pero *vino “Tío Paco” = 0* también lo sería. De la misma manera, esta idea también se puede

extender para atributos nominales que tengan más de dos valores, por ejemplo, podríamos tener el ítem *país = Italia*, o el ítem *país = Portugal*, o el ítem *país = Alemania*, etc.

Es por esto que varias reglas de asociación pueden parecerse a las reglas de clasificación, pero se diferencian en que no hay una clase definida, ya que cualquier atributo puede ser tomado.

Una regla de dependencia se define como cualquier conjunto de variables atributo que sea dependiente. El método propuesto consiste en medir la significancia de las diferentes dependencias mediante el test ji-cuadrado usado en la estadística clásica para las correlaciones, el test ji-cuadrado nos permite establecer estadísticamente si dos atributos son dependientes o no. (Hernández, J; Ramírez, M; Ferri, C; 2008)

- **Interés de dependencias**

Una manera de comprobar la dependencia interna de los valores de los atributos es mediante las probabilidades de cada uno de los valores. “El interés entre dos eventos x e y se define de manera similar al cálculo de la dependencia entre dos ítems $(xy) = p(xy) / p(x) p(y)$ ” (Hernández, J; Ramírez, M; Ferri, C; 2008)

En el libro de (Hernández, J; Ramírez, M; Ferri, C; 2008), encontramos que, dado un atributo, se considera un evento como cada uno de los posibles valores (o ítems) que puede tomar. Por lo tanto, para una dependencia entre dos atributos binarios, se construye una tabla 2x2 que contenga el interés de cada una de las combinaciones. Un valor de interés superior a 1 significa dependencia positiva, mientras un valor inferior a 1 indica dependencia negativa.

Por ejemplo, en la tabla 3 para el caso de los atributos “*Pan Blanco*” y “*Pan Integral*” tenemos la siguiente tabla de interés:

Tabla 3
Ejemplo de dependencias

	Pan Integral	\neg (Pan Integral)
Pan Blanco	0,833	1,071
\neg (Pan Blanco)	1,667	0,714

Fuente (Hernández, J; Ramírez, M; Ferri, C; 2008)

Es así que de esta tabla se extraer, que existe una fuerte dependencia positiva, entre comprar pan integral y no comprar pan blanco, y una ligera dependencia negativa entre no comprar pan blanco y no comprar pan integral.

Utilizando el marco de aprendizaje de reglas de asociación encontrado en (Hernández, J; Ramírez, M; Ferri, C; 2008), basado en cobertura/confianza, fijando una cobertura mínima del 25 por ciento y una confianza del 50 por ciento, se dan tres reglas con el ejemplo anterior:

- **SI** “Pan Blanco” **ENTONCES** \neg (Pan Integral”) Cb = 80, Cf = 0,75
- **SI** “Pan Integral” **ENTONCES** “Pan Blanco” Cb = 30, Cf = 0,67
- **SI** \neg (Pan integral”) **ENTONCES** “Pan Blanco” Cb = 70, Cf = 0,86

En este caso, se observa que las mejores reglas en cobertura y confianza coinciden con la relación dominante, sin embargo, esto no siempre se cumple.

“Se puede demostrar que existe una correspondencia directa entre el valor de una celda de la tabla de interés con el valor de esa celda utilizado para calcular el valor ji-cuadrado de la relación” (Silverstein et al. 1977)

3.1.3.5.2 Árboles de Decisión

Un árbol de decisión es un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que la decisión final a tomar se puede determinar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta algunas hojas. Los árboles de decisión se utilizan desde hace siglos, y son especialmente apropiados para expresar procedimientos médicos, legales, comerciales, estratégicos, matemáticos, lógicos, etc. (Hernández, J; Ramírez, M; Ferri, C; 2008)

Una de las grandes ventajas de los árboles de decisión es que las opciones posibles a partir de una determinada condición son excluyentes. Esto permite analizar una situación y siguiendo el árbol de decisión llegar a una sola acción o decisión a tomar.

Para entender un árbol de decisión solo se debe realizar las preguntas y seguir las respuestas hasta alguna de las hojas del árbol, catalogadas con un “no” o un “si”. Este árbol de decisión en concreto funciona como “clasificador”, es decir, dado un nuevo individuo nos clasifica en una de las dos clases posibles: “no” o “si”.

La tarea de aprendizaje para la cual los árboles de decisión se adecuan mejor es la clasificación, de hecho, clasificar es “determinar de entre varias clases a qué clase pertenece un objeto; la estructura de condición y ramificación de un árbol de decisión es idónea para este problema” (Hernández, J; Ramírez, M; Ferri, C; 2008).

La característica más importante del problema de la clasificación es que se asume que las clases son disjuntas, es decir, una sustancia es de la clase *a* o de la clase *b*, pero no puede ser al mismo tiempo de las clases *a* y *b*.

En este sentido, la clasificación se diferencia de la categorización, donde se permiten más de una clase, etiqueta o categoría para cada instancia. Debido a que la clasificación trata con clases o etiquetas disjuntas, un árbol de decisión conducirá un ejemplo hasta una y sólo una hoja, asignando, por tanto, una única clase.

“La propiedad exhaustiva permite el esquema básico de los primeros algoritmos de aprendizaje de árboles de decisión; el espacio de instancias se iba *partiendo* de arriba abajo, utilizando cada vez una partición, es decir, un conjunto de condiciones excluyentes y exhaustivas. Estos algoritmos se llaman algoritmos de *partición* o algoritmos de, divide y vencerás” (Hernández, J; Ramírez, M; Ferri, C; 2008)

Otra característica importante de los algoritmos de aprendizaje de árboles de decisión es que, una vez elegida la partición dicha partición, no se podía cambiar, aunque más tarde se pensara que había sido una mala elección, por tanto, uno de los aspectos más importantes en los sistemas de aprendizaje de árboles de decisión es el denominado criterio de partición, ya que una mala elección de la partición generará un peor árbol.

Partiendo de esta afirmación encontrada en (Hernández, J; Ramírez, M; Ferri, C; 2008), los dos puntos importantes para que el algoritmo funcione correctamente son:

- Particiones a considerar
- Criterio de selección de particiones

- **Particiones posibles**

“Las particiones son un conjunto de condiciones exhaustivas y excluyentes. Lógicamente, cuantos más tipos de condiciones se permitan, más posibilidades existen de encontrar patrones detrás de los datos” (Hernández, J; Ramírez, M; Ferri, C; 2008), del mismo autor encontramos el siguiente ejemplo:

“Cuantas más particiones existan más expresivos podrán ser los árboles de decisión generados y probablemente más precisos, también debemos considerar que mientras más particiones haya, la complejidad del algoritmo será mayor” (Hernández, J; Ramírez, M; Ferri, C; 2008)

En el libro (Hernández, J; Ramírez, M; Ferri, C; 2008) encontramos y se definen así las particiones:

- Particiones nominales
- Particiones numéricas: si un atributo es numérico y continuo, se debe intentar obtener particiones de intervalos, son de la forma ($x_i \leq a$, $x_i > a$).

3.1.3.5.3 Agrupación de datos o Clustering

“La agrupación o el clustering consisten en agrupar un conjunto de datos, sin tener clases predefinidas, basándose en la similitud de los valores de los atributos de los distintos datos” (Hernández, J; Ramírez, M; Ferri, C; 2008)

Esta agrupación, a diferencia de la clasificación, se realiza de forma no supervisada, ya que no se conoce de antemano las clases del conjunto de datos de entrenamiento.

El clustering identifica clusters, o regiones densamente pobladas, de acuerdo a alguna medida de distancia, en un gran conjunto de datos multidimensional.

“El objetivo del clustering es maximizar la similitud de las instancias en cada cluster y minimizar la similitud entre clusters” (Hernández, J; Ramírez, M; Ferri, C; 2008)

“K-Means es un método particional de clustering donde se construye una partición de una base de datos D de n objetos en un conjunto de k grupos, buscando optimizar el criterio de particionamiento elegido. En K-Means cada grupo está representado por su centro. K-Means intenta formar k grupos, con k predeterminado antes del inicio del proceso. Asume que los atributos de los objetos forman un vector espacial. El objetivo que se intenta alcanzar es minimizar la varianza total intra-grupo o la función de error cuadrático” (Hernández, J; Ramírez, M; Ferri, C; 2008)

3.1.3.5.4 Clasificación de Datos

“La clasificación se utiliza para clasificar un conjunto de datos basado en los valores de sus atributos” (Hernández, J; Ramírez, M; Ferri, C; 2008)

Por ejemplo, se podría clasificar a distintas personas para el otorgamiento de un préstamo en riesgo bajo, medio y alto, teniendo en cuenta información histórica de las mismas.

La clasificación encuentra las propiedades comunes entre un conjunto de objetos y los clasifica en diferentes clases, de acuerdo a un modelo de clasificación. Para construir este modelo, se utiliza un conjunto de entrenamiento, en el que cada instancia consiste en un conjunto de atributos y el valor de la clase a la cual pertenece.

“Los algoritmos mayormente utilizados para las tareas de clasificación son los algoritmos de inducción. En la actualidad existen numerosos enfoques de algoritmos de inducción y variedad en cada enfoque” (Hernández, J; Ramírez, M; Ferri, C; 2008)

“ID3 es un sistema típico de construcción de árboles de decisión, el cual adopta una estrategia de arriba hacia abajo e inspecciona solo una parte del espacio de búsqueda. ID3 garantiza que será encontrado un árbol simple, pero no necesariamente el más simple” (Hernández, J; Ramírez, M; Ferri, C; 2008)

3.2 Trabajos Relacionados

Business Intelligence ha tenido un desarrollo y crecimiento importante en la última década, es así que ha generado el interés de los empresarios en este tipo de herramientas. A nivel mundial existen varios estudios realizados con Data mining, en la búsqueda de encontrar patrones en los distintos comportamientos.

Uno de los casos más exitosos de Data Mining es el desarrollado por el Banco Santander, cuando adopta solución de Data Mining de SAS y construye modelos de crédito para anticipar riesgos de no cumplimiento de las obligaciones, (Garrido, L; Latorre, J; 2000)

En el trabajo (Garrido, L; Latorre, J; 2000), asegura que el banco para optimizar los modelos de crédito ofrecidos a sus clientes y asegurar la reducción de los niveles de falta de cumplimiento de las obligaciones, el banco está invirtiendo en recursos eficaces de tecnología.

Este banco adoptó la solución de Data Mining de SAS, que forma parte del CRM Analítico de SAS, para explotar y obtener conocimientos de su base de datos. La institución realizó una evaluación para medir los beneficios y el desempeño de la solución, comparándola con otras del mercado, antes de adoptarla. Después de esta etapa, se concluyó que la solución de SAS procesó un gran volumen de datos en un corto espacio de tiempo, lo que demostró su desempeño superior.

Un punto que se deberá destacar es que, en la mayoría de las tareas ejecutadas a través del recurso ofrecido por el SAS/Enterprise Miner, obtuvo un desempeño del 70% superior en comparación con la solución ofrecida por los competidores, considerando también el tiempo de procesamiento de las actividades. Además de ello, otro ítem que ha tenido influencia en la adopción de la solución de Data Mining fue la calidad de los servicios ofrecidos por el soporte técnico de SAS. (Garrido, L; Latorre, J; 2000)

La solución está siendo usada por el Banco Santander para procesar datos de más de 1 millón de clientes, además de trabajar con un gran volumen de informaciones, la solución SAS consigue actuar con diversas variables, a la vez en (Garrido, L; Latorre, J; 2000) describe que una gran ventaja de la utilización de la solución es el hecho de ser una solución tecnológica completa que ofrece diversas metodologías de análisis que otras no presentan.

“En el área de Marketing La solución SAS también está siendo usada para la excelencia de los flujos de la relación con clientes y prospectos, incluyendo la adquisición, activación, retención y recuperación de clientes” (Garrido, L; Latorre, J; 2000)

Un estudio en la Industria de automóviles, aplicando la minería de datos, es “Aplicación de data mining para control de calidad de cajas de cambio” (Bustamante, R; 2015), el objetivo de

este trabajo es el desarrollo de un sistema de diagnóstico automático para detección de cajas de cambio defectuosas.

Descripción; Durante las pruebas de control de calidad de los productos fabricados industrialmente, se inspeccionan diferentes parámetros para detectar fallos del material, fallos de producción, fallos de ensamblaje o ruidos extraños durante el funcionamiento. En el caso de las cajas de cambio, todas las unidades producidas por un fabricante de automóviles son evaluadas acústicamente por especialistas humanos. Las pruebas acústicas de estos componentes permiten seleccionar los componentes válidos e identificar los componentes defectuosos por sus patrones acústicos característicos. (Bustamante, R; 2015)

(Bustamante, R; 2015) menciona que, debido a la compleja estructura de una caja de cambio, el elevado número de componentes móviles y la influencia de diferentes estados operativos, es necesario analizar un gran número de características acústicas, extraídas de las señales medidas. Para ayudar y, finalmente, automatizar la decisión del inspector humano, se ha desarrollado un sistema clasificador borroso. El principal problema del diseño del clasificador es la cantidad de características a examinar. En estas condiciones, una aproximación basada en una implementación directa, del conocimiento de los expertos sería poco eficiente respecto al tiempo necesario y requeriría además entender muy bien el proceso.

Pese a los inconvenientes, con esta opción pueden obtenerse soluciones válidas, pero los problemas de complejidad del diseño hacen que este método resulte inviable, en general, y que se opte por alternativas más eficientes. Una aproximación alternativa consiste en aplicar técnicas de data mining para generar de forma automática el clasificador a partir de los datos extraídos de una serie de ejemplos. Analizando las señales de una muestra de objetos, que incluye ejemplos tanto de componentes en buen estado como defectuosos, es posible extraer una base de reglas de diagnóstico, sin necesidad de conocimiento del proceso.

En este caso, limitando las combinaciones a un máximo de 6 expresiones, el número total de reglas posibles es del orden de 10^7 , imposible de comprobar íntegramente. Aplicando métodos de minería de datos, se han analizado las características de 1060 ejemplos (1000 correspondientes a objetos en buen estado y 60 a casos defectuosos) para generar una base de varios miles de reglas con la que se clasifican correctamente todos los ejemplos. (Bustamante, R; 2015)

Mediante diferentes técnicas, la base de reglas puede reducirse para obtener un sistema de diagnóstico de menor tamaño y más comprensible, es así que (Bustamante, R; 2015) encuentra un clasificador que consta de un centenar de reglas, manteniendo el porcentaje de diagnósticos correctos en el 100% para los casos utilizados como ejemplos y con un porcentaje de aciertos del 92% cuando se utiliza para clasificar datos nuevos.

Otra aplicación importante de la minería de datos, es en el campo de la medicina, como describe el artículo de Aplicación de minería de datos para el diagnóstico de accidentes cerebro vasculares agudos (ACVAs), (DAEDALUS; 2015); el objetivo de este estudio es el desarrollo de un sistema de soporte a la decisión para el diagnóstico de las causas de accidentes cerebro vasculares agudos.

Descripción, El ACVA es una enfermedad frecuente y grave que amenaza la vida de miles de personas cada año. Las causas de esta dolencia pueden ser muy diversas y para el tratamiento es importante disponer de una estimación rápida y precisa del origen del problema. Para determinarlo, se cuenta con información acerca de unas 170 variables que describen el estado del paciente, como la edad, la presión sanguínea, la presencia de arritmias, el grado de coma en la escala de Glasgow. (DAEDALUS; 2015)

En el artículo (DAEDALUS; 2015) permite entender las complicaciones y dificultad de este diagnóstico médico y lo que a su vez plantea la necesidad de un sistema de ayuda a la toma de decisiones, que proporcione una primera indicación que permita aplicar un tratamiento de forma inmediata.

El sistema de soporte a la decisión desarrollado en este proyecto (DAEDALUS; 2015) consiste básicamente en un clasificador que divide los ataques en una serie de clases, en función de sus causas.

“En este caso, se definen cinco categorías, asociadas a las causas más comunes de un ACVA, en concreto arteriosclerosis, embolia, trombosis y derrame cerebral, agrupándose en una quinta clase los de causa desconocida o que no responden a ninguna de las anteriores” (DAEDALUS; 2015)

En este estudio (DAEDALUS; 2015) para el desarrollo del sistema, se dispone de los valores que toman las variables de interés para un conjunto de casos de muestra. Aplicando

técnicas de minería de datos puede extraerse la información contenida en dichos ejemplos y definir un modelo que represente las relaciones entre las variables observadas y el diagnóstico. Una de las técnicas más utilizadas para el desarrollo de sistemas de diagnóstico médico son los árboles de decisión.

Se trata de clasificadores que se basan en algoritmos inductivos para extraer conocimiento a partir de un conjunto de casos y definir las relaciones que aparecen en el proceso que se intenta simular. El conocimiento se representa en forma de árboles, en los que cada nodo representa una condición sobre una variable. Después de una serie de evaluaciones en los nodos intermedios se alcanza una conclusión definitiva al llegar a las hojas o nodos finales, cada una de las cuales se asocia a una de las clases consideradas por el sistema. (DAEDALUS; 2015)

La representación de la información en forma de árboles puede resultar muy intuitiva, y a la vez pudiendo inmediatamente extraer reglas, lo que facilita la comprensión y evaluación del modelo por parte de los expertos, es así que gracias a este estudio (DAEDALUS; 2015) podemos concluir que este tipo de representación es ideal para modelar procesos de diagnóstico, en los que no sólo interesa la conclusión final sino el proceso de razonamiento seguido.

Esta es una de las razones porque la minería de datos tiene un auge en aplicaciones médicas, la precisión que obtenga de los modelos dependerá del esquema de inducción aplicado para obtener la base de reglas.

Al evaluar casos nuevos, no utilizados para la construcción del modelo, la tasa de acierto de los expertos es del 66%, aproximadamente. Los resultados obtenidos por el sistema con dichos datos son de hasta un 75% de diagnósticos correctos y aún pueden ser sensiblemente mejorados con un ajuste final basado en las indicaciones de los expertos. (DAEDALUS; 2015)

Un caso de éxito en el campo de Gestión- Marketing y bases de datos, es: Aplicación de minería de datos para segmentación de clientes de entidades financieras, con el objetivo de segmentación de los clientes de un banco mediante un modelo basado en agrupamiento (clustering) (Valcarcel, V; 2004)

Descripción; el concepto de segmentación de clientes tiene su origen en el campo del marketing y la investigación de mercado. Describe la división analítica de un mercado, clasificando los potenciales clientes conforme a distintos criterios, formando grupos o segmentos de clientes homogéneos internamente y heterogéneos externamente, a cada uno de los cuales pueden dirigirse acciones específicas de marketing, con lo que mejora la eficacia de estas actividades. (Valcarcel, V; 2004)

Estas técnicas son especialmente valiosas para las compañías de servicios financieros, en las que la relación con el cliente es un factor clave para el éxito. De acuerdo a (Valcarcel, V; 2004) la filosofía de las técnicas de segmentación de mercado consiste básicamente en caracterizar los objetos mediante una serie de propiedades, como edad, sexo, estatus familiar, educación, profesión, ingresos, patrimonio, etc., y organizarlos de acuerdo con los valores que dichas propiedades toman en cada caso. Así, los objetos se agrupan de modo que los que pertenecen a un mismo grupo (cluster) son posiblemente similares y, viceversa, los que pertenecen a clases distintas deberían ser posiblemente diferentes, siempre en cuanto a los valores que toman las características observadas.

El análisis de agrupamientos (clustering) tradicional, ampliamente utilizado para la segmentación de clientes en investigaciones de mercado, incluye diversos métodos de búsqueda de estructuras en conjuntos de datos. Con estos métodos, cada cliente se asigna rígidamente y completamente a un determinado grupo y los grupos están separados entre sí por umbrales estrictos y precisos. “Esta aproximación no resulta realista y supone una limitación importante de este tipo de técnicas” (Valcarcel, V; 2004)

A diferencia de la rígida visión resultante de las técnicas tradicionales, la minería de datos sugiere soluciones, los métodos de agrupamiento borroso, en las que a cada objeto se le asocian grados de pertenencia continuos a las diferentes clases. Los modelos basados en agrupamiento borroso permiten reflejar situaciones en las que un cliente comparte características de varios grupos, imposibles de representar con los modelos tradicionales, que asignan este tipo de clientes a un determinado grupo igual que los que no ofrecen ningún género de dudas. La perspectiva gradual resultante de los modelos borrosos proporciona una visión más exacta y realista de la posición de cada cliente en el espacio de clases, lo que permite al responsable de marketing ajustar las acciones a cada usuario en particular de una forma más individualizada y, por tanto, más eficaz. (Valcarcel, V; 2004)

Este trabajo encontró además de las ventajas desde el punto de vista estático, el agrupamiento borroso presenta también ventajas para el análisis dinámico. Los métodos tradicionales únicamente reflejan las transiciones de un segmento a otro, pero no explican los cambios, puesto que no son capaces de representar la trayectoria seguida por el cliente. Al considerar valores graduales, los métodos de borrosos permiten observar cómo evoluciona un cliente con el tiempo y detectar inmediatamente cómo cambian los requisitos de los usuarios.

Un estudio relevante para el trabajo de tesis desarrollado, es el publicado por (DAEDALUS; 2017), titulado “Aplicación de minería de datos para predicción de ventas” este es un tema puramente de gestión, con el objetivo del desarrollo de un modelo para predecir las ventas de un producto en un determinado mes, basándose en datos sobre las ventas en meses previos.

En el estudio (DAEDALUS; 2017), trata que uno de los campos de aplicación tradicionales de la minería de datos es la predicción de la evolución en el futuro de una variable a partir de datos históricos sobre su comportamiento en el pasado.

Las técnicas de minería de datos constituyen una alternativa útil y eficaz a las aproximaciones matemáticas tradicionales, especialmente en el caso de variaciones muy irregulares, complicadas de modelar con los métodos clásicos. La empresa Bayer mantiene un registro histórico de diferentes datos, entre ellos las cifras de ventas. Basándose únicamente en los datos de ventas de uno de sus productos, sin indicadores adicionales, pretende desarrollar un modelo del comportamiento de dicho producto en el mercado que le permita predecir las ventas del mismo con cierta anticipación. En concreto, se dispone de las cifras de los últimos 56 meses. Con esta información, los resultados obtenidos aplicando métodos de predicción tradicionales no son suficientemente precisos. Así, tomando como estimación para un mes el importe correspondiente al mes anterior, se obtiene un error del 25,6%, y del 14% si se usa el de ese mes en el año anterior. La calidad de la predicción mejora utilizando medias móviles, pero el error es aún del 11,8%. (DAEDALUS; 2017)

Debido a los altos porcentajes de error obtenidos descritos en el artículo (DAEDALUS; 2017), ninguno de estos métodos satisface las necesidades de la compañía. Para mejorar la precisión del modelo y conseguir la exactitud necesaria en las predicciones en este mismo artículo se describe que el camino a seguir es la aplicación de técnicas de minería de datos, es así que se encuentra en primer lugar analizado las características básicas de las ventas.

Se encontraron a simple vista que se presenta una tendencia creciente en el tiempo que puede modelarse con medias móviles. También se observan oscilaciones estacionales, aunque estas regularidades no aparecen en todos los meses. Por ejemplo, si bien los valores de las ventas son siempre bajos en agosto, los de mayo presentan grandes variaciones. Esto puede significar que la serie incluye varios factores de influencia con distintos periodos. Estas observaciones se ven confirmadas por el análisis del espectro de frecuencia, que muestra varios máximos. (DAEDALUS; 2017)

En el artículo (DAEDALUS; 2017) detalla las conclusiones preliminares, donde se sugiere la conveniencia de incluir en el modelo información no sólo de los valores de ventas en los meses previos sino también sobre la tendencia de la serie y sobre la temporada en cuestión, datos todos ellos contenidos en la propia serie.

Es así que el sistema desarrollado, basado en una red neuronal, predice las ventas en un determinado mes partiendo únicamente de características extraídas de la serie de ventas en función del tiempo, sin indicadores adicionales.

Los resultados encontrados fueron “Los datos de entrada con los que se han obtenido los mejores resultados son: Valores de las ventas en los tres meses anteriores. Ventas del mes a predecir en el año anterior. Valor medio de las ventas durante los últimos 12 meses. Identificador del mes” (DAEDALUS; 2017)

Las predicciones del modelo obtenido aplicando minería de datos son más precisas, que las expuesta en experimentos pasados, técnicas clásicas, con un error medio del 6,4%.

Se encontró el estudio “Aplicación de minería de datos en detección y control del fraude en el uso de tarjetas de crédito, este tiene como objetivo el análisis de los atributos característicos de las transacciones fraudulentas y desarrollo de un sistema para su identificación y detección” (Bustamante, R; 2017), este estudio es un claro referente de la aplicación de minería de datos en el sector financiero.

Descripción; el uso fraudulento de tarjetas de crédito supone un coste de miles de millones de dólares anuales para el sistema bancario y la economía mundial. Pese a las numerosas medidas ensayadas para combatirlo, la cantidad y sofisticación de este tipo de delitos aumenta cada año, superándose sistemáticamente las medidas anti-fraude. Generalmente, los bancos emisores disponen de sistemas que realizan algún tipo de comprobación de las transacciones, utilizando sencillas reglas si—entonces. El problema de estos sistemas es que, aunque intuitivamente se sepa que ciertas reglas detectan el uso irregular de una tarjeta, normalmente resulta imposible expresarlas con validez empírica (Bustamante, R; 2017)

Dado el antecedente en (Bustamante, R; 2017) concluye que el banco a menudo enfrenta el dilema de identificar erróneamente una tarjeta como fraudulenta cuando en realidad no es el caso, lo que implica el riesgo potencial de deteriorar la relación con el cliente.

El sistema desarrollado en este proyecto (Bustamante, R; 2017) se basa en la hipótesis de que un usuario no autorizado utiliza una tarjeta de forma cualitativa y cuantitativamente diferente de como la ha utilizado anteriormente el usuario legítimo, factores como la frecuencia de empleo de la tarjeta, el tipo y situación de los comercios en que suele utilizarse, hasta qué punto el usuario respeta su límite de crédito, forman en conjunto una huella que puede identificar de forma unívoca al usuario legítimo. Una ruptura de estos patrones puede utilizarse como indicador para detectar si otra persona está utilizando la tarjeta de forma fraudulenta.

Es vital que los sistemas bancarios sean capaces de reconocer dichas violaciones de los esquemas típicos lo más pronto posible. Esencialmente, el infractor es un usuario de los servicios del banco, si bien un usuario indeseable. Aplicando técnicas de minería de datos, puede diferenciarse claramente su comportamiento del de los clientes normales. En este proyecto (Bustamante, R; 2017) el análisis mediante métodos de clustering de una serie de datos sobre el titular de la tarjeta, el comercio y la transacción ha permitido: Definir las características que, combinadas, caracterizan los diversos tipos de fraude. Diferenciar el uso fraudulento del normal.

En este estudio la conclusión obtenida por (Bustamante, R; 2017) de los resultados obtenidos es la propuesta de una serie de recomendaciones y pautas de supervisión del uso de las tarjetas en tiempo real, entre ellas el análisis de comercios (para detectar aquéllos sospechosos de colaborar con los autores del fraude); la necesidad de consolidar la información sobre transacciones fraudulentas, lo que proporcionaría una visión más exacta del problema; y el análisis

de vulnerabilidad, que permitiría revisar más exhaustivamente las transacciones cuando el riesgo es máximo puesto que está demostrado que las tarjetas son particularmente vulnerables en ciertas condiciones.

Es así que análisis basados en técnicas de minería de datos, son muy útiles en problemas como en este caso de detección precoz de fraudes en áreas como banca, financiación, telecomunicaciones, en las que identificar los problemas y reaccionar rápidamente, incluso en tiempo real, es vital para el éxito de la organización.

El desarrollo de Data mining en el Ecuador está empezando, y con el paso de los años está tomando más fuerza, dentro del análisis de proyectos donde a través de técnicas de data mining se descubran patrones de comportamiento, como ejemplos generales encontramos:

El estudio realizado por Raúl Muñoz Álvarez, Minería de datos aplicada al manejo de la relación del Cliente en una Entidad Bancaria (CRM), encontrado en el repositorio digital de la Escuela Politécnica Nacional, nos habla sobre la actualidad del entorno financiero, como está marcado por el avance de la tecnología y los efectos que ésta proporciona en la relación entre una Entidad Bancaria y sus clientes.

“Negocios como la entrega de servicios y colocación de productos financieros, que tradicionalmente eran realizados en las oficinas o agencias de servicios, hoy en día son ofertados mediante los nuevos canales de contacto tecnológicos, llegando a convertirse en el principal vínculo entre el cliente y la entidad Bancaria” (MUÑOZ, A; 2013)

De acuerdo a esta afirmación en este trabajo se analiza los patrones de comportamiento de un cliente inmersos en los datos recolectados en los canales transaccionales de una Entidad Bancaria, mediante la utilización de técnicas de Minería de Datos y el CRM, con la finalidad de determinar perfiles de clientes, que permitan un mejor conocimiento de la relación existente entre cliente y su Entidad Bancaria.

Los resultados obtenidos en este estudio son que las entidades bancarias miran a los canales transaccionales y de servicios como el principal vínculo para relacionarse con sus clientes.

Para las entidades los canales se han convertido en oportunidades para captar, satisfacer y conocer a sus clientes.

El marketing dentro de una entidad bancaria, es el encargado de generar las estrategias necesarias para llevar de manera adecuada y sostenible la relación que vincula a los clientes hacia los productos y servicios. El CRM como parte del marketing, tiene como meta el mejorar los procesos de comunicación hacia el cliente correcto, proveyéndolo del producto o servicio correcto, a través del canal correcto y en el tiempo correcto.

Existe una tesis, sobre Desarrollo de un módulo de inteligencia de negocios para el análisis de la entrega del bono de desarrollo humano desde el año 2009, en el Ecuador, realizada por Bustamante Bustamante, Jorge, y Santillán Haro, Karen, basándose que el Ministerio de Inclusión Económica y Social maneja procesos como el pago del Bono de Desarrollo Humano en ventanilla, pago del Bono de Desarrollo humano en cuenta, Crédito de desarrollo Humano y Corresponsabilidad, son procesos que generan un cantidad de datos considerables que son almacenados en bases de datos transaccionales e históricas, el análisis de estos datos es realizado manualmente y de manera periódica por lo que automatizarlo reducirá la cantidad de tiempo y personal involucrado en este proceso.

El proceso de pago del Bono de Desarrollo humano en ventanilla fue la base utilizada para el desarrollo e implantación del módulo de inteligencia de negocios, basándose en una problemática de que el porcentaje de personas que reciben esta indemnización es mayor comparado con los demás procesos, debido a que parte de las bases de datos involucradas son transaccionales se procuró integrar esta información en una base de datos independiente que no afecte el correcto funcionamiento de los procesos que se maneja dentro de las áreas involucradas y que alimente al modelo dimensional que concluirán en la generación de reportes e indicadores de gestión, que permitirán corroborar si el dinero invertido en este proceso está enfocado en cubrir las necesidades de las personas que lo requieren (Bustamante, J; Santillán, H; 2009)

En este estudio se encontró que el Ministerio de Inclusión Económica y Social tiene varios procesos relacionados con el pago del bono de desarrollo humano, es así que la parte inicial se realizó el primer datamart sobre pago en ventanilla del bono de desarrollo humano y pensiones, a la vez con este se plantea que el siguiente trabajo es la implementación de procesos básicos que sean necesarios para el desarrollo de la información.

También en (Bustamante, J; Santillán, H; 2009) concluyen que la información obtenida del análisis, es inconsistente por lo cual especifican los requerimientos mínimos para tener un correcto funcionamiento y rendimiento en un periodo de tiempo determinado. Esto se debe a que el servidor será utilizado constantemente para la generación de reportes que serán la base para formulación de informes, que permitirán toma de decisiones adecuadas.

Se ha realizado una búsqueda amplia de estudios que hayan aplicado minería de datos en el campo de comercialización de artículos en el país y no hemos obtenido hallazgos de aplicaciones en este campo, pero los estudios encontrados describiendo los problemas planteados en distintas áreas como son financiera, industria automovilística, gubernamental, médica, comercial, servicio al cliente, ventas y a la vez el uso de distintas herramientas de minería de datos nos permite tener una visión de la importancia y el avance de esta temática.

De igual forma la importancia de un dominio en la problemática que se quiere resolver, ya que no todas las herramientas llegan a obtener el objetivo, pero si se identifica de forma correcta el beneficio y funcionalidad de ellas se pueden obtener la solución se problemas y entregar a los mandos medios y directivos herramientas eficientes mediante información para la toma de decisiones.

CAPÍTULO 4

MARCO METODOLÓGICO

En el presente capítulo explica de manera ordenada y secuencial los pasos requeridos para poder alcanzar los objetivos propuestos, en primer lugar, se entiende el negocio y las necesidades que con el trabajo se desea cubrir, luego se detalla las fuentes de los datos, el manejo y depuración de los mismos, el análisis de las tareas de data mining que se utilizaron y los resultados que arrojaron al utilizar estas herramientas.

4.1 Metodología

La figura 8 muestra esquemáticamente la metodología utilizada en el desarrollo de este trabajo.

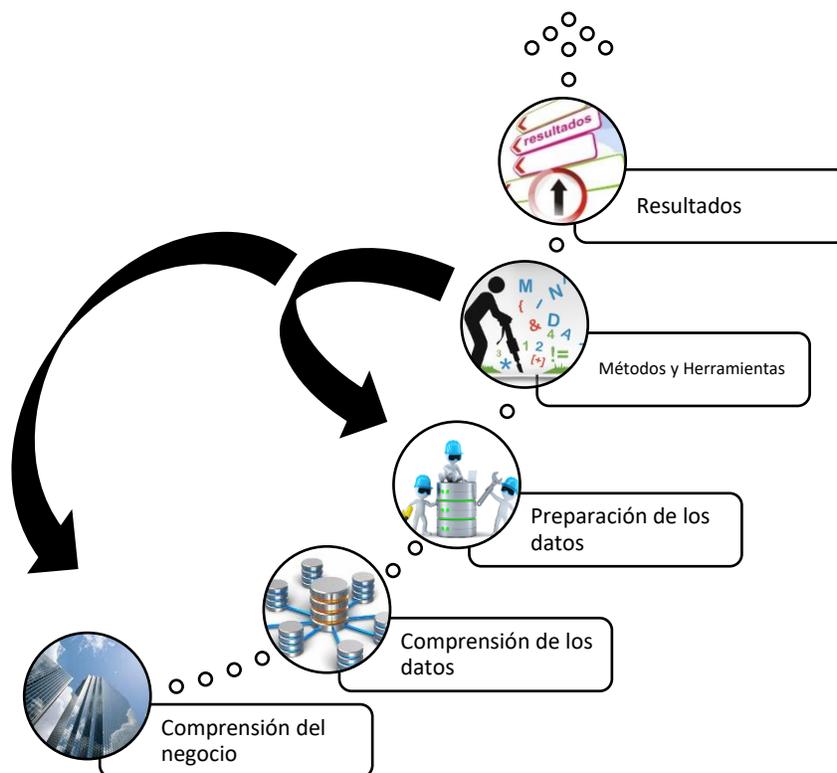


Figura 8 Metodología aplicada

4.1.1 Comprensión del negocio

La empresa que ha proporcionado los datos está en la industria del comercio en artículos deportivos, presenta venta al menudeo es decir dedicada al retail, basándose en las ventas detallistas a los clientes.

Las empresas comercializadoras de estos artículos deportivos realizan compras de la mercadería que ingresará al país durante seis meses, y la compra se realiza con 6 meses antes del primer ingreso de esta al país, a la vez dividen esta compra de seis meses en ingresos mensuales y colecciones, al igual que por tiendas en los casos más sofisticados, esta compra la realizan de acuerdo a la percepción o creencia de ventas.

Estos ítems deportivos están clasificados por grupos y a la vez por categorías o actividades, por ejemplo, fútbol, correr, entrenar, estilo deportivo, etc. Y los grupos son zapatos, ropa y accesorios, a la vez existen tipo de producto por su tecnología y precio divididos en tres segmentos, top, enhanced y core.

Hay que comprender que al realizar la compra de un artículo hay que tomar la decisión en: cantidad total del ítem para cada tienda, la curva a manejarse es decir las cantidades por cada talla, ejemplo, (curva total de 6 en una camiseta; 1 XS, 2 S, 2 M, 1 L), el mes de ingreso de este ítem, cuál será el conjunto de ítems a ingresar en ese mes.

Analizamos esta industria a nivel nacional, centrándonos en las ciudades de Quito, Guayaquil y Cuenca, específicamente nos referiremos a los siguientes puntos de venta:

- Mall el Jardín (Quito)
- Condado Shopping (Quito)
- Scala Shopping (Quito)
- City Mall (Guayaquil)
- Mall del Sol (Guayaquil)
- Village Plaza (Guayaquil)
- Mall del Río (Cuenca)

Para las empresas de este sector se analiza que, de su compra total, el 75% del stock como mínimo debe venderse a full price, y el otro 25% puede ser vendido con algún tipo de descuento en liquidación, el mismo que dependerá de la rotación que haya tenido.

El ciclo normal de un producto en punto de venta debe ser de 90 días, a partir de este momento se presenta las acciones de liquidación.

Los resultados obtenidos de las últimas 3 compras efectuadas por esta empresa se muestran en la tabla 1, donde se encuentra en la temporada de enero a junio 2016 el 57% de índice de liquidación, de julio a diciembre el 50% y de enero a junio un 58%.

Al entender el giro del negocio y el objetivo de querer liquidar el mínimo porcentaje, un análisis del comportamiento del consumidor en las decisiones de compra de porque un artículo de cierta actividad, cuales son las tallas de mayor rotación, que ítems se venden solos y cuales salen en conjuntos, cuales son las características primordiales en la decisión de compra, y entendiendo la variabilidad del mercado este análisis se lo debe llevar a nivel de tienda por las diferencias que se sustentan en la segmentación descrita en el capítulo 3, nos permite la comprensión de que manejar patrones del comportamiento del consumidor y el mix adecuado de producto por cada tienda es la herramienta más valiosa para que las compras que se efectúen no sean realizadas empíricamente si no con un sustento en conocimiento de la situación real, de esta forma se podrá minimizar los índices de liquidación que existen, al ofrecer al cliente la mercadería adecuada y necesaria en los puntos de venta, de igual forma los esfuerzos de marketing estarán dirigidos correctamente y no se desperdiciará recursos temporales, financieros ni humanos.

4.1.2 Comprensión de los datos

Los datos que utilizaremos son proporcionados por una empresa dedicada a la comercialización de productos deportivos, estos datos sobre los que trabajaremos y buscaremos los patrones de comportamiento están situadas en 3 provincias del país, con un total de 7 tiendas.

- Provincia Pichincha, ciudad de Quito (3 tiendas, Scala, Jardín y Condado)
- Provincia Guayas, ciudad de Guayaquil (3 tiendas, Village, Sol, City)
- Provincia Azuay, ciudad de Cuenca (1 tienda, Mall del Río)

- **Estructura de los datos**

Las bases fueron obtenidas del sistema que maneja esta empresa SAP BI descargadas en Excel, la información que se obtuvo es desde el año 2013 al 2016.

La estructura de estos datos se describe a continuación en la tabla 4:

Tabla 4
Campos y descripción de los datos

Campo	Descripción
Código Internacional	Código de identificación de cada producto
Número de Factura	Número de Factura en la que se realiza la venta del producto
Documento de Identificación	Número de Cédula o pasaporte del cliente
Forma de pago	Forma en que se realiza el pago de la factura
Segmentación	Clasificación de la marca para los productos: top, enhanced, core
Familia	Colección a la que pertenece un producto
Color	Color del producto
Grupo	Clasificación del producto por su tipo: calzado, ropa, accesorios
Actividad	Clasificación por actividad a la que pertenece el producto, ejm Training, Fútbol
Línea	Descripción del producto, ejm: short, calentador, gorra
Código MCU	Código de identificación para cada punto de venta (tienda)
Región	Región a la que pertenece el punto de venta
Género Edad	Identificación del género y edad del producto
Mes	Mes de venta del producto
Fecha	Fecha de venta del producto
Año	Año de venta del producto
Descripción sku	Nombre del producto
Talla	Talla del producto
Cantidad	Cantidad de venta
Stock	Cantidad con la que se termina el mes en punto de venta
Costo unitario	Costo unitario del producto
Pvp	Precio de venta del producto
Venta	Cantidad en dinero por la venta del producto

4.1.3 Preparación de los datos

Para poder realizar el análisis, se realizó un proceso de limpieza de datos, dado que los datos obtenidos no estaban en el mismo formato, existían cantidades nulas, referencias con el mismo significado, pero en su escritura diferente, se procedió a realizar esta limpieza y depuración.

Dentro de los datos obtenidos se realizó una depuración principalmente en los campos de identificación del cliente y forma de pago.

Para que el proceso de carga fuera efectivo y los resultados obtenidos tengan un sentido más amplio se procedió a crear rangos y agregaciones en varios campos por la numerosidad de los datos, estos fueron en color del producto, rangos en las tallas, rangos en el precio de venta al público y concepto del producto. Finalmente se comprobó que todos los datos dentro de un mismo campo tuvieran el mismo formato.

4.1.4 Métodos y Herramientas

- **Weka**

Para el desarrollo del proyecto se ha utilizado un framework de data mining denominado Weka, cuyo nombre proviene de un ave endémica de Nueva Zelanda. Esta gallinácea en peligro de extinción es famosa por su curiosidad y agresividad.

Este framework con el nombre de un ave, es una extensa colección de algoritmos de aprendizaje automático desarrollados por la universidad de Waikato ubicada en Nueva Zelanda, implementados en Java, útiles para ser aplicados sobre datos mediante las interfaces que ofrece o para embeberlos dentro de cualquier aplicación, a su vez Weka contiene las herramientas necesarias para realizar transformaciones sobre los datos, tareas de clasificación, clustering, asociación y visualización. Weka está diseñado como una herramienta orientada a la extensibilidad por lo que permite añadir nuevas funcionalidades.

4.1.4.1 Tarea de Visualización

La tarea de visualización como su nombre lo dice nos permite visualizar los cruces entre variables de interés, y a la vez obtener información de estos. Para poder utilizar con eficiencia esta tarea se debieron retirar los siguientes atributos:

- Cod_Factura
- Id Cliente
- Familia
- Color
- Fecha
- Costo Unitario
- PVP
- Venta
- Venta Impuesto
- Descripción del producto
- Cantidad

Esta depuración se realiza por la numerosidad de estos atributos, ya que no permitirían tener un análisis visual de resultados ni porcentajes concluyentes, también se procede a unir todos los años analizados, y en el campo de Genero_ Edad, se ha realizado una agregación adicional.

Para realizar los cruces de visualización se analizaron conforme a la tabla 5.

Esta tabla se analizó en conjunto con los expertos de la empresa acorde a las necesidades y a los objetivos del trabajo.

A la vez también se realizó una base adicional donde se segmentaba por Tienda, para poder analizar grupo, género y talla.

Es así que mediante la tarea de visualización se pudo obtener resultados relevantes, y encontrar diferentes comportamientos de ventas por tiendas y generales.

Tabla 5
Análisis de cruces de visualización

	Forma de pago	Segmentación	Grupo	Actividad	Línea	Tienda	Región	Genero	Mes	Año
Forma de pago	X									
Segmentación	NO	X								
Grupo	NO	SI	X							
Actividad	NO	SI	SI	X						
Línea	NO	NO	SI	SI	X					
Tienda	NO	SI	SI	SI	SI	X				
Región	SI	NO	SI	SI	SI	NO	X			
Genero	NO	NO	SI	SI	SI	SI	NO	X		
Mes	SI	NO	NO	NO	NO	NO	SI	SI	X	
Año	NO	NO	NO	NO	NO	SI	NO	NO	NO	X

4.1.4.1 Tarea de Clasificación

En la tarea de clasificación se dirigió hacia la creación de árboles de decisión (J48) que nos permitan mediante la secuencia de decisiones de los clientes en el proceso de compra, distinguir atributos o comportamientos para cada tienda o sectorizarlos por regiones, de esta forma se plantearon dos tipos de análisis:

- Análisis 1

El análisis se realizó con los siguientes campos, y solo considerando el grupo zapatos.

- Talla
- Genero edad
- Línea
- Región
- Tienda

La clase escogida para este análisis fue: Tiendas

El modelo nos da como resultado un 42% de confianza, pero en el árbol obtenido no se pudo llegar a un resultado, ya que existe una diversificación de datos que no permite determinar una tienda final partiendo de tallas o líneas de venta.

- Análisis 2

Los campos a utilizarse en este análisis son:

- Talla
- Genero edad
- Línea
- Grupo
- Región

Para poder realizar el análisis, se realizó un proceso de limpieza de datos, porque los datos obtenidos no estaban en el mismo formato, existían cantidades nulas, referencias con el mismo significado, pero en su escritura diferente, se procedió a realizar esta limpieza y depuración.

Para que el proceso de carga fuera efectivo y los resultados obtenidos tengan un sentido más amplio se procedió a crear rangos y agregaciones en varios campos por la numerosidad de los datos.

La clase escogida en este análisis fue la Región.

Con este modelo obtuvimos un 59,41% de aciertos, pero al igual que el análisis anterior el árbol obtenido no es consistente en sus resultados ni permiten tener una lectura para realizar una clasificación y llegar a concluir una región.

4.1.4.2 Tarea de Clustering

En esta tarea el objetivo fue identificar las características comunes entre las ventas que se dan de cada tienda, para que no haya discrepancia y se pueda obtener resultados coherentes se realizó un análisis a cada grupo, es decir para zapatos, ropa, y accesorios.

Las variables que se consideraron fueron:

- Color
- Línea
- Tienda

Al realizar los cluster de mercado no se pudo obtener clusters identificados dado que las características de venta de las tiendas son similares entre sí, y esta herramienta realizaba discriminaciones que dan resultados no concluyentes.

Ejemplo de Resultado 1: se realizó el análisis con la clase objetivo de tiendas, y con el grupo de zapatos, como se puede visualizar en las figuras 9.

```

=== Run information ===

Scheme:      weka.clusterers.EM -I 100 -N 7 -X 10 -max 7 -ll-cv 1.0E-
Relation:    1 grupo zap-weka.filters.unsupervised.attribute.Remove-R
Instances:   77540
Attributes:  3
             COLOR
             LINEA

Ignored:
             COD_TIENDA

Test mode:   Classes to clusters evaluation on training data

=== Clustering model (full training set) ===

EM
==

Number of clusters: 7
Number of iterations performed: 0

Attribute      Cluster
              0      1      2      3      4      5      6
              (0.4) (0.23) (0.05) (0.17) (0.07) (0.03) (0.04)
=====
COLOR
Black          24244      1  1578      1      1      1      1
Blue           1      1 1231 13304      1      1      1
White          1 18161      1      1      1      1      1
Grey           1      1  277      1 5489      1      1
Brown         1632      1      1      1      1      1      1
Purple        1275      1  154      1      1      1      1
Green         1218      1      1      1      1      1      1
Pink          1774      1  229      1      1      1      1
Red            1      1  48      1      1 2577      1
Yellow         1      1  31      1      1      1 3447
Orange        187      1      6      1      1      1      1
Silver         537      1      1      1      1      1      1
Blue/Purple   103      1      1      1      1      1      1
Wheat          8      1      1      1      1      1      1
Petrol         3      1      1      1      1      1      1
Turquesa      51      1      1      1      1      1      1
[total]       31037 18176 3562 13319 5504 2592 3462

[total]       31037 18176 3562 13319 5504 2592 3462
LINEA
Zapato        31022 17749      1 13304 5489 2577 3447
Sandalia      1  413 3547      1      1      1      1
[total]       31023 18162 3548 13305 5490 2578 3448

Time taken to build model (full training data) : 0.76 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      31021 ( 40%)
1      18160 ( 23%)
2      3546 (  5%)
3      13303 ( 17%)
4      5488 (  7%)
5      2576 (  3%)
6      3446 (  4%)

Log likelihood: -2.03671

Class attribute: COD_TIENDA
Classes to Clusters:

   0   1   2   3   4   5   6 <-- assigned to cluster
5841 3597 555 2604 1194 514 649 | 221JAR1
7103 3915 1050 3119 1171 575 744 | 221SOL1
4460 2512 396 1747 824 390 459 | 221CUE1
5901 3526 365 2522 936 461 504 | 221CON1
2355 1390 177 1016 508 174 310 | 221SCA1
3245 1889 605 1399 551 267 396 | 221CITI
2116 1331 398 896 304 195 384 | 221VILL

Cluster 0 <-- 221SOL1
Cluster 1 <-- 221CON1
Cluster 2 <-- 221CITI
Cluster 3 <-- 221JAR1
Cluster 4 <-- 221CUE1
Cluster 5 <-- 221SCA1
Cluster 6 <-- 221VILL

```

Figura 9 Resultados tarea de Clasificación

4.1.4.3 Tarea de Asociación

Para el uso de esta herramienta se realizó una segmentación en los datos a nivel de años y tiendas; por cada tienda se obtuvo una base de datos, la cual se ingresó al programa con el objetivo de encontrar las principales asociaciones de compra entre líneas de artículos.

A la vez los datos de trabajo para esta herramienta fueron transformados a nominales, esto se realizó a nivel de documento de compra.

En esta tarea se puso obtener resultados concluyentes, un ejemplo de estos resultados se muestra en la figura 10.

The screenshot shows the Weka Explorer interface with the 'Associate' tab selected. The 'Apriori' algorithm is running, and the 'Associate' output window displays a list of 62 association rules. Each rule is formatted as: `Item1=SI Item2=SI => Item3=SI Item4=SI <conf:(0.47)> lift:(1.14) lev:(0.01) [98] conv:(1.11)`. The rules represent various combinations of clothing items like 'Medias', 'Zapato', 'T-Shirt', 'GORRA', 'Pantalon', 'Chompa', 'Sandalia', 'BILLETERA', 'MOCHILA', 'Agua', and 'BOLSO'.

```

20. MEDIAS=SI 1693 ==> Zapato=SI 795 <conf:(0.47)> lift:(1.14) lev:(0.01) [98] conv:(1.11)
21. MEDIAS=SI 1693 ==> T-Shirt=NO Zapato=SI 636 <conf:(0.38)> lift:(1.05) lev:(0) [32] conv:(1.03)
22. Chompa=SI 774 ==> Zapato=SI 188 <conf:(0.24)> lift:(0.59) lev:(-0.01) [-130] conv:(0.78)
23. Polo=SI 633 ==> Zapato=SI 153 <conf:(0.24)> lift:(0.59) lev:(-0.01) [-107] conv:(0.77)
24. Pantalon=SI 891 ==> Zapato=SI 209 <conf:(0.23)> lift:(0.57) lev:(-0.01) [-157] conv:(0.77)
25. Short=SI 1597 ==> Zapato=SI 351 <conf:(0.22)> lift:(0.53) lev:(-0.02) [-306] conv:(0.75)
26. Short=SI T-Shirt=NO 1004 ==> Zapato=SI 185 <conf:(0.18)> lift:(0.45) lev:(-0.02) [-228] conv:(0.72)
27. GORRA=SI 2039 ==> Zapato=SI 305 <conf:(0.15)> lift:(0.36) lev:(-0.04) [-534] conv:(0.69)
28. T-Shirt=NO 11313 ==> GORRA=SI 1685 <conf:(0.15)> lift:(1.1) lev:(0.01) [158] conv:(1.02)
29. Short=SI 1597 ==> MEDIAS=SI 211 <conf:(0.13)> lift:(1.18) lev:(0) [32] conv:(1.02)
30. Zapato=SI 6217 ==> MEDIAS=SI 795 <conf:(0.13)> lift:(1.14) lev:(0.01) [98] conv:(1.02)
31. MEDIAS=SI 1693 ==> Short=SI 211 <conf:(0.12)> lift:(1.18) lev:(0) [32] conv:(1.02)
32. T-Shirt=NO Zapato=SI 5384 ==> MEDIAS=SI 636 <conf:(0.12)> lift:(1.05) lev:(0) [32] conv:(1.01)
33. GORRA=SI T-Shirt=NO 1685 ==> Zapato=SI 199 <conf:(0.12)> lift:(0.29) lev:(-0.03) [-494] conv:(0.67)
34. T-Shirt=NO 11313 ==> MEDIAS=SI 1334 <conf:(0.12)> lift:(1.05) lev:(0) [66] conv:(1.01)
35. Short=SI 1597 ==> T-Shirt=NO Zapato=SI 185 <conf:(0.12)> lift:(0.33) lev:(-0.03) [-384] conv:(0.73)
36. MEDIAS=SI 1693 ==> GORRA=SI 180 <conf:(0.11)> lift:(0.79) lev:(-0) [-48] conv:(0.97)
37. Zapato=SI 6217 ==> MEDIAS=SI T-Shirt=NO 636 <conf:(0.1)> lift:(1.16) lev:(0.01) [87] conv:(1.02)
38. GORRA=SI 2039 ==> T-Shirt=NO Zapato=SI 199 <conf:(0.1)> lift:(0.27) lev:(-0.03) [-527] conv:(0.71)
39. T-Shirt=NO 11313 ==> Short=SI 1004 <conf:(0.09)> lift:(0.84) lev:(-0.01) [-191] conv:(0.98)
40. GORRA=SI 2039 ==> MEDIAS=SI 180 <conf:(0.09)> lift:(0.79) lev:(-0) [-48] conv:(0.97)
41. Zapato=SI 6217 ==> Short=SI 351 <conf:(0.06)> lift:(0.53) lev:(-0.02) [-306] conv:(0.95)
42. T-Shirt=NO 11313 ==> MEDIAS=SI Zapato=SI 636 <conf:(0.06)> lift:(1.07) lev:(0) [40] conv:(1)
43. T-Shirt=NO 11313 ==> Chompa=SI 619 <conf:(0.05)> lift:(1.07) lev:(0) [39] conv:(1)
44. Zapato=SI 6217 ==> GORRA=SI 305 <conf:(0.05)> lift:(0.36) lev:(-0.04) [-534] conv:(0.91)
45. T-Shirt=NO 11313 ==> Sandalia=SI 536 <conf:(0.05)> lift:(1.14) lev:(0) [66] conv:(1.01)
46. T-Shirt=NO 11313 ==> Pantalon=SI 507 <conf:(0.04)> lift:(0.76) lev:(-0.01) [-160] conv:(0.99)
47. T-Shirt=NO 11313 ==> Polo=SI 483 <conf:(0.04)> lift:(1.02) lev:(0) [9] conv:(1)
48. T-Shirt=NO 11313 ==> Bvd=SI 451 <conf:(0.04)> lift:(1) lev:(-0) [-1] conv:(1)
49. T-Shirt=NO Zapato=SI 5384 ==> GORRA=SI 199 <conf:(0.04)> lift:(0.27) lev:(-0.03) [-527] conv:(0.9)
50. T-Shirt=NO 11313 ==> FELOTA=SI 392 <conf:(0.03)> lift:(1.17) lev:(0) [57] conv:(1.01)
51. T-Shirt=NO Zapato=SI 5384 ==> Short=SI 185 <conf:(0.03)> lift:(0.33) lev:(-0.03) [-384] conv:(0.93)
52. Zapato=SI 6217 ==> Pantalon=SI 209 <conf:(0.03)> lift:(0.57) lev:(-0.01) [-157] conv:(0.97)
53. Zapato=SI 6217 ==> GORRA=SI T-Shirt=NO 199 <conf:(0.03)> lift:(0.29) lev:(-0.03) [-494] conv:(0.92)
54. Zapato=SI 6217 ==> Chompa=SI 188 <conf:(0.03)> lift:(0.59) lev:(-0.01) [-130] conv:(0.98)
55. Zapato=SI 6217 ==> Short=SI T-Shirt=NO 185 <conf:(0.03)> lift:(0.45) lev:(-0.02) [-228] conv:(0.96)
56. Zapato=SI 6217 ==> Polo=SI 153 <conf:(0.02)> lift:(0.59) lev:(-0.01) [-107] conv:(0.98)
57. T-Shirt=NO 11313 ==> BILLETERA=SI 260 <conf:(0.02)> lift:(1.17) lev:(0) [38] conv:(1)
58. T-Shirt=NO 11313 ==> MOCHILA=SI 223 <conf:(0.02)> lift:(1.17) lev:(0) [32] conv:(1)
59. T-Shirt=NO 11313 ==> GORRA=SI Zapato=SI 199 <conf:(0.02)> lift:(0.87) lev:(-0) [-29] conv:(1)
60. T-Shirt=NO 11313 ==> Short=SI Zapato=SI 185 <conf:(0.02)> lift:(0.7) lev:(-0.01) [-77] conv:(0.99)
61. T-Shirt=NO 11313 ==> Agua=SI 179 <conf:(0.02)> lift:(1.05) lev:(0) [9] conv:(1)
62. T-Shirt=NO 11313 ==> BOLSO=SI 166 <conf:(0.01)> lift:(1.12) lev:(0) [17] conv:(1)

```

Figura 10 Resultados Tarea de Asociación

4.1.5 Resultados

Después de haber realizado las pruebas en las distintas tareas de data mining, las que nos arrojaron resultados analizables y coherentes a los objetivos del trabajo y en base al funcionamiento y necesidades de las empresas comercializadoras de artículos deportivos son, la tarea de visualización, manejando las contrastaciones seleccionadas y analizando los puntos de cruce, nos permitirá encontrar el perfil del consumidor y analizar el comportamiento de ventas; la tarea de asociación da como resultado reglas de compras que permiten conocer de igual forma un perfil del consumidor, los productos con mayor rotación y a la vez el mix de producto, encontrar las líneas y combinaciones de mayor venta a la vez permite encontrar un visual merchandising para las tiendas con el objetivo de maximizar los tickets promedio y las cantidades de venta.

- **Contrastación 2: Forma de pago / Región**

Se visualiza claramente que en la región costa existe una preferencia en el pago en efectivo, y las transacciones dadas por tarjeta de crédito son casi nulas, mientras que en la sierra el uso de la tarjeta de crédito y efectivo tiene casi la misma preferencia.

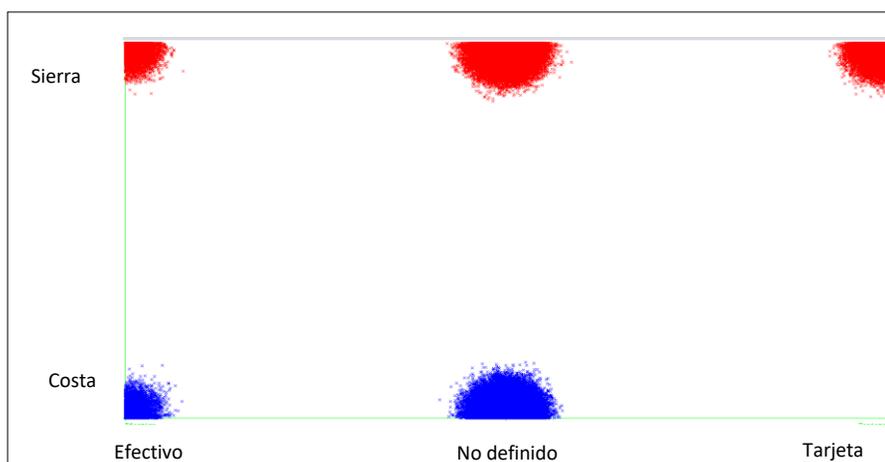


Figura 12 Contraste Forma de pago / Región

- **Contrastación 3: Grupo / Segmentación**

En este análisis se puede visualizar que el segmento enhanced es el de mayor presencia en los tres grupos, en el caso de accesorios tenemos una presencia reducida del segmento Top, mientras que en zapatos esta parte top junto con la core tienen similar presencia.

Nos da un claro análisis que las compras para las tiendas deben estar enfocadas en un porcentaje mayor en enhanced, luego en core y finalmente una pequeña proporción en top.

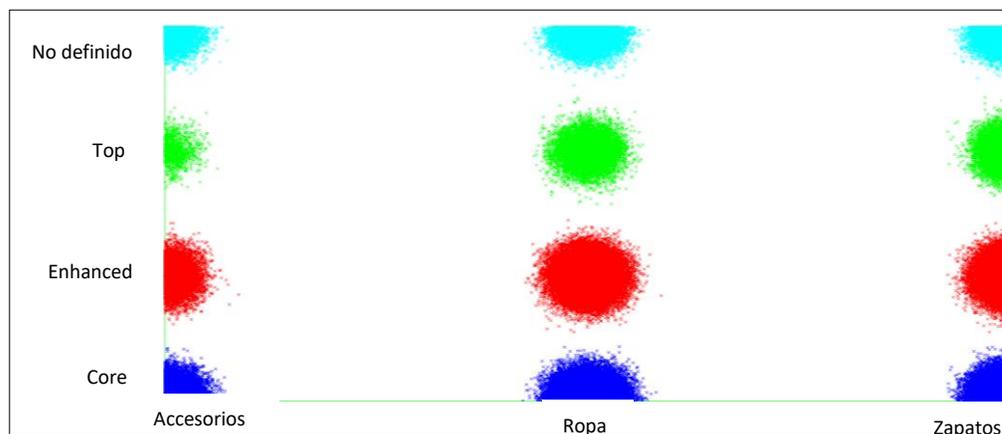


Figura 13 Contraste Grupo / Segmentación

- Contrastación 4: Actividad / Segmentación

En el resultado de este contraste de actividad y segmentación se denota que existen actividades con una representatividad casi nula como son voley, neo, fútbol sala, squash y atletismo, mientras que training es la categoría con mayor presencia, al analizarlo por segmento, en la parte top running tiene una venta importante y natación es la que ocupa el nivel más bajo, en la parte media de la pirámide es decir enhanced encontramos que la menos representativa es deporte de aventura, y en la parte core tenemos una representatividad similar en todas las actividades.

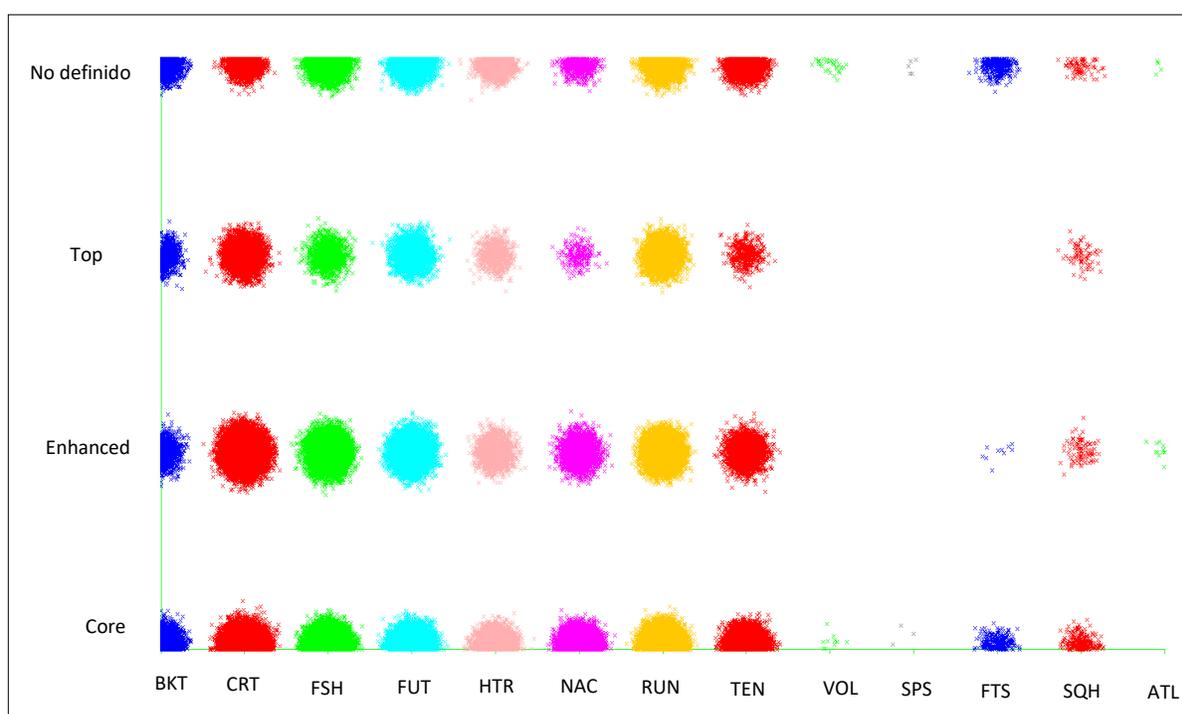


Figura 14 Contraste Actividad / Segmentación

- Contrastación 5: Tienda / Segmentación

Se puede visualizar que el segmento top es el de menor venta en todas las tiendas, pero esta diferencia se denota con más fuerza en Village Plaza y Scala Shopping.

Como habíamos analizado en gráficas anteriores la mayor presencia de venta se da en el segmento enhanced, seguido de core y finalmente top.

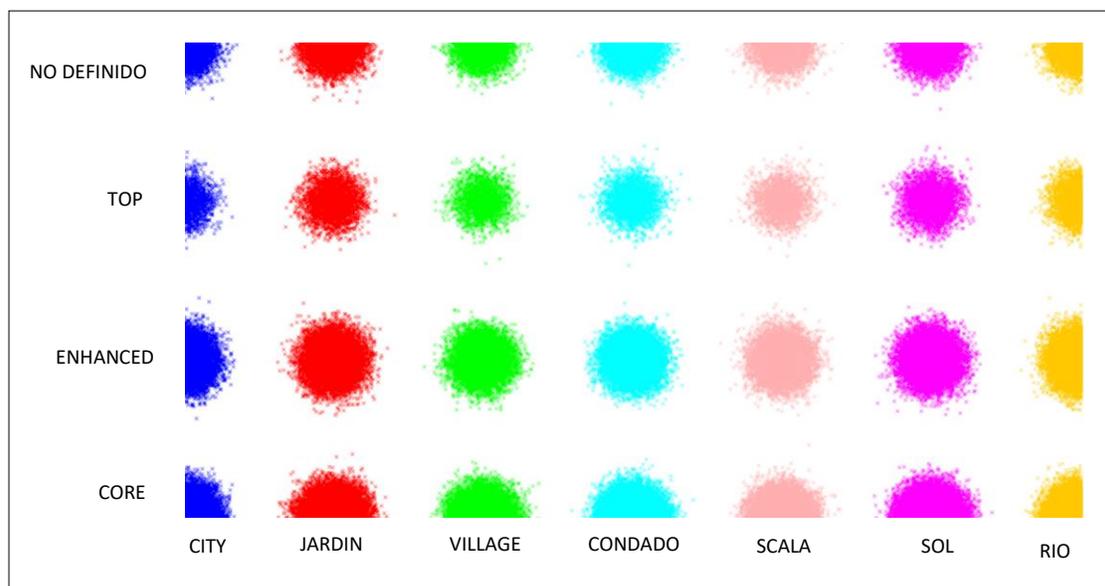


Figura 15 Contraste Tienda / Segmentación

- Contratación 6: Tienda / Año

En la visualización de este cruce se puede concluir que las ventas de 2013 a 2014 crecieron, pero de 2014 a 2015 tuvieron un decrecimiento, este comportamiento general se da en todas las tiendas, mientras que el año 2016 fue más alto en venta que 2014, también se encontró que la mayor cantidad de venta en los distintos años se da en Mall del Sol y Mall el Jardín, también existe una presencia fuerte en Mall del Río, mientras que la tienda con menores niveles de ventas es Village Plaza y City Mall.

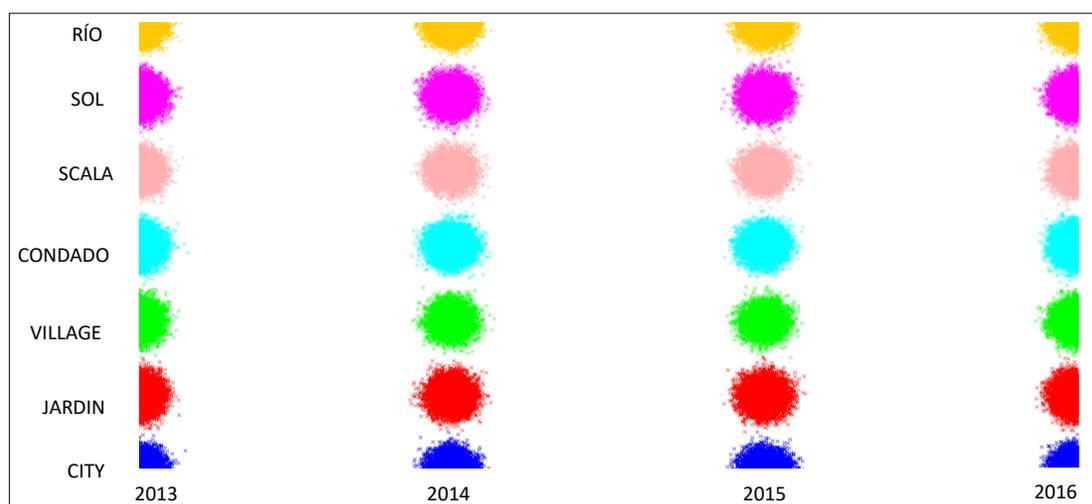


Figura 16 Contraste Tienda / Año

- **Contrastación 7: Grupo / Género edad**

Este análisis nos permite tener una visión global sobre el comportamiento general de cada género en las tiendas, en el caso de unisex adultos encontramos que su presencia ponderable es solo en el grupo de accesorios, mientras que en infantes vemos que la mayor fuerza toma en zapatos seguido de ropa y por último en accesorios, cuando analizamos kids y hombre adulto vemos un comportamiento parecido una presencia casi igual en el grupo de ropa y zapatos, Finalmente al revisar los datos de mujer existe una clara preferencia por el grupo de ropa.

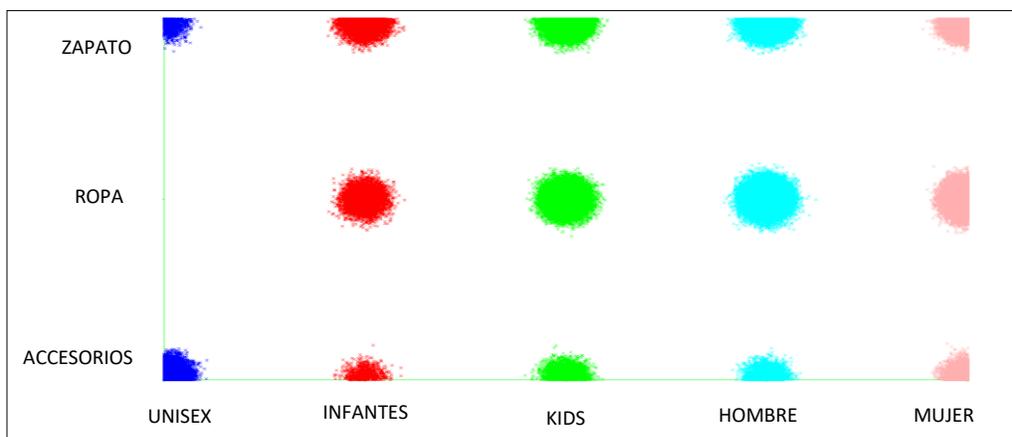


Figura 17 Contraste Grupo / Género edad

- **Contrastación 8: Grupo / Región**

Se puede concluir que el grupo de mayor venta en unidades es ropa y su mayor presencia se encuentra en la sierra, mientras que en el grupo zapatos encontramos una presencia similar en ambas regiones al igual que accesorios.

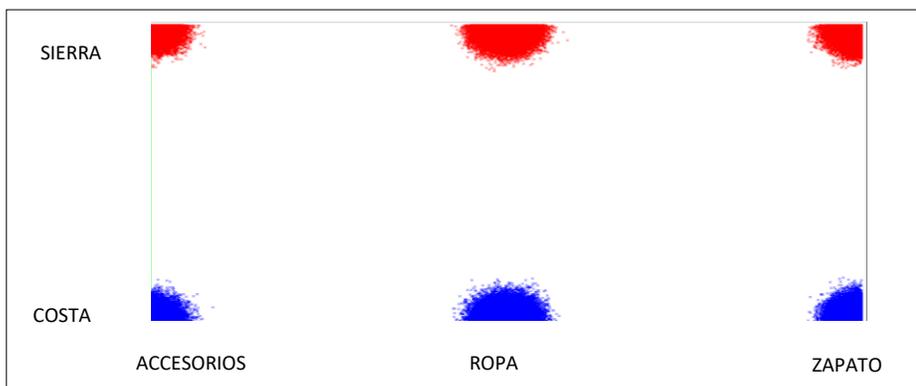


Figura 18 Contraste Grupo / Región

- **Contrastación 9: Grupo / Tienda**

Se puede visualizar que en los tres grupos las tiendas con mayor presencia son Mall el Jardín y Mall del Sol, presentando mayores ventas, pero se denotan también que las tiendas de City Mall y Village Plaza tienen una presencia mayoritaria en ventas en ropa, seguida de calzado y por una diferencia notable en último lugar accesorios. La venta mayoritaria por grupo en todas las tiendas es en ropa, pero al analizarlo en dólares zapatos es el grupo con mayor representación.

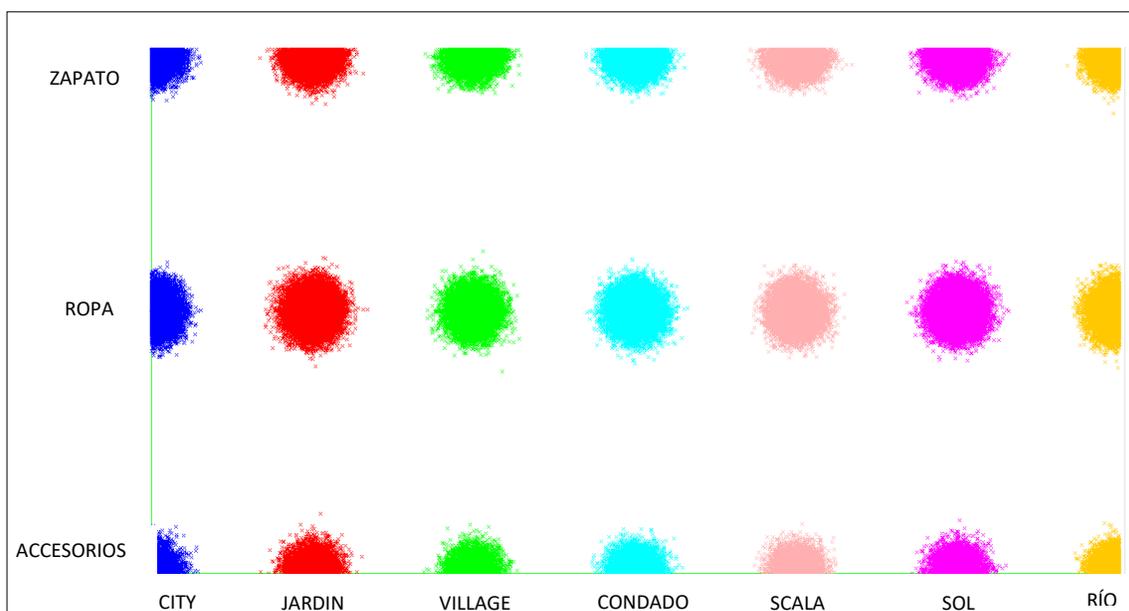


Figura 19 Contraste Grupo / Tienda

- **Contrastación 10: Actividad / Grupo**

En el grupo de zapatos podemos analizar que la presencia de vóley, sps y atletismo es casi nula, mientras que originals y running tienen una mayor participación.

En el grupo de ropa la categoría con mayor aportación es training y la de menor demanda es outdoors.

En accesorios se replica que la categoría con mayor participación es training y resalta fútbol, mientras que vóley y fútbol sala tienen las ventas más bajas.

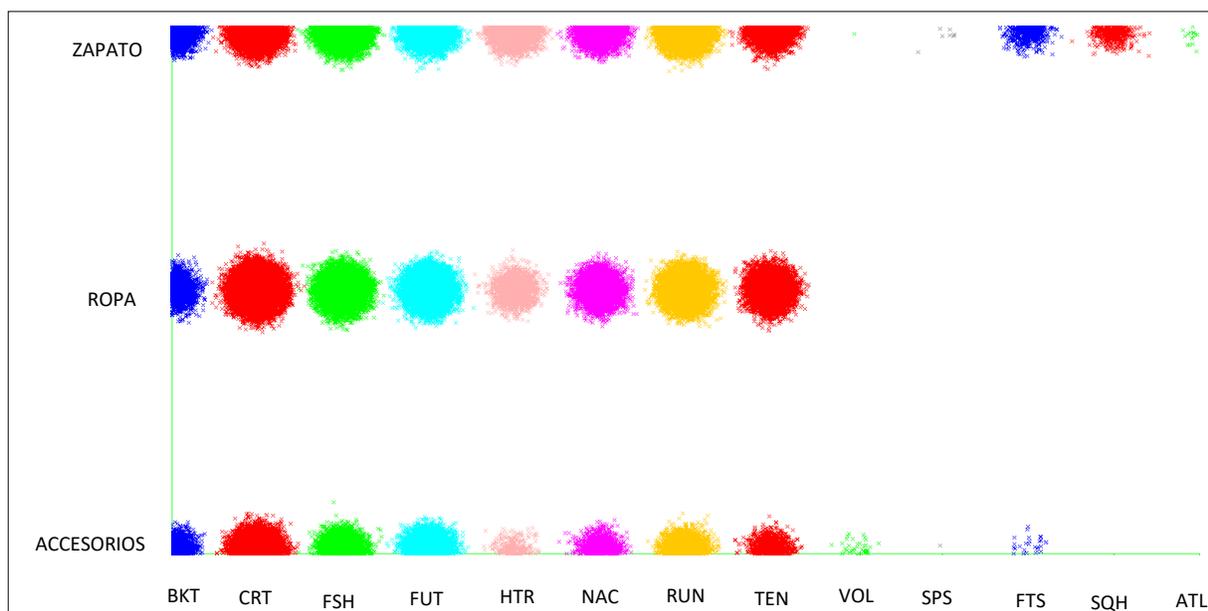


Figura 20 Contraste Actividad / Grupo

- Contrastación 11: Género edad / Actividad

Al analizar por género, resalta que en mujer existe una venta baja en basket, fútbol y squash, mientras que training, originals y natación tiene una participación equilibrada en general.

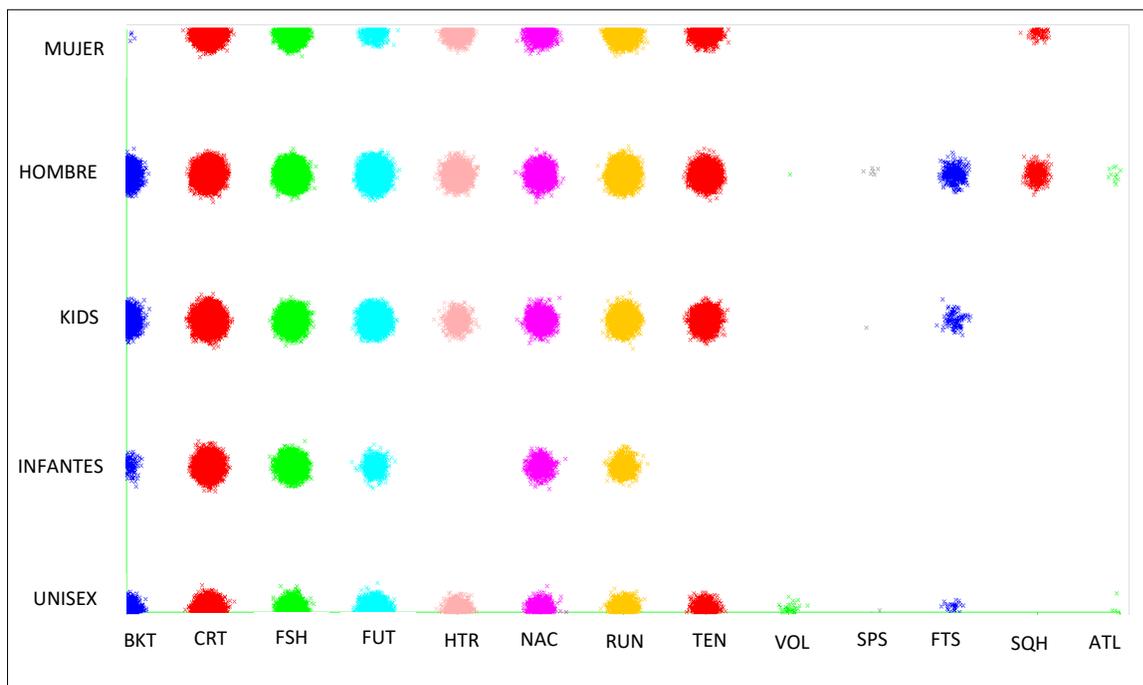


Figura 21 Contraste Género edad / Actividad

- **Contrastación 12: Región / Actividad**

En este análisis visual se concluye que las principales categorías tienen una venta similar tanto en sierra y costa, resaltando únicamente que basket y outdoors tienen una venta mayor en la sierra, mientras que las categorías con menor participación como vóley es más fuerte en la sierra y fútbol sala y squash en la costa.

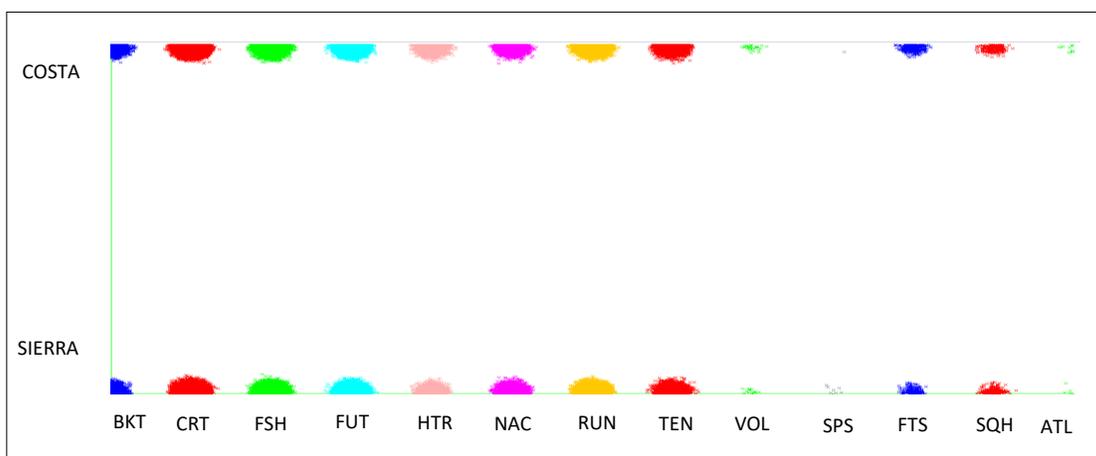


Figura 22 Contraste Región / Actividad

- **Contrastación 13: Tienda / Actividad**

Al analizar el comportamiento de las actividades en las tiendas resalta que Mall el Jardín, Mall del Sol y Mall del Río tienen una participación similar en todos los deportes, con presencia mayoritaria en training y running y menor en fútbol sala, squash y nula en atletismo, sps y vóley.

Actividades como outdoors tiene una menor participación en la tienda de Scala Shopping y Village Plaza. El deporte de vóley, sps, y atletismo es casi nula en todas las tiendas su mínima presencia solo se denota en Mall el Jardín.

City Mall, presenta su mayor venta en actividad en training con un 33%, seguida de fútbol con un 23%, y en tercer lugar hay ventas iguales de fashion y running, también existe una pequeña participación de natación, tenis y outdoors.

La tienda del Condado Shopping, su mayor venta por categoría es 30% en training, 24% en fútbol, 20% fashion y 14% en running.

Mall del Río tiene las ventas más altas en training con un 33%, fútbol con 23%, el tercer lugar lo ocupa running con un 17% y fashion con 16%, también tenemos una pequeña participación de 4% en outdoors, 3% en natación y tenis y 2% en basket.

En la tienda Mall el Jardín el comportamiento de venta a nivel de categoría es training con un 35%, seguida de fútbol con un 22%, luego en similares condiciones con el 17% fashion y 16% running, en esta tienda también hay presencia de tenis, natación, outdoors y basket.

Scala Shopping la mayor participación por categoría la tiene training con un 33%, seguida de fútbol con un 26%, luego se encuentra running con el 16% y fashion con un 14%, tenis también sobresale con un 5%, 3% en natación, 3% en outdoors y 2% en basket.

Mall del Sol, el comportamiento de venta a nivel de categoría es de un 31% en training, 23% en fútbol, 17% en running y 15% en fashion.

Village Plaza el comportamiento de venta a nivel de categoría es de 33% training, 26% fútbol, 15% running, 11% fashion 6% tenis, 5% natación.

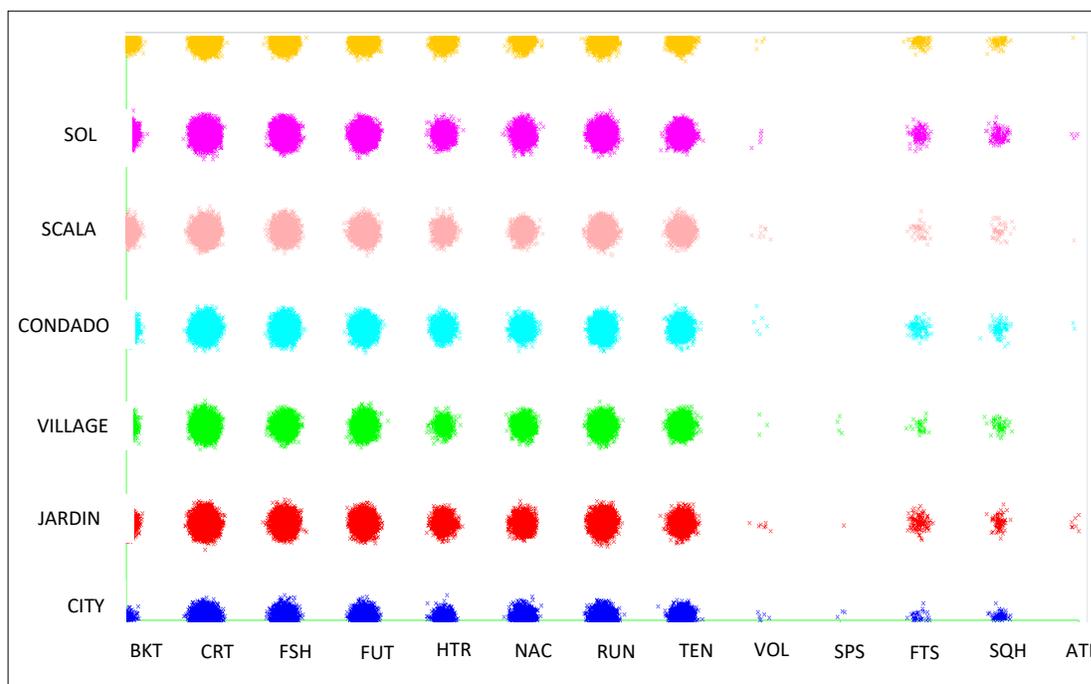


Figura 23 Contraste Tienda / Actividad

- **Contrastación 14: Línea / Actividad**

En actividades como atletismo, squash, sps solo tienen una oferta de producto en la línea de zapatos, es por eso que en gráficas anteriores tenemos una participación baja en ventas. Se visualiza que las líneas de mayor venta en todas las categorías son t-shirts, shorts y medias.

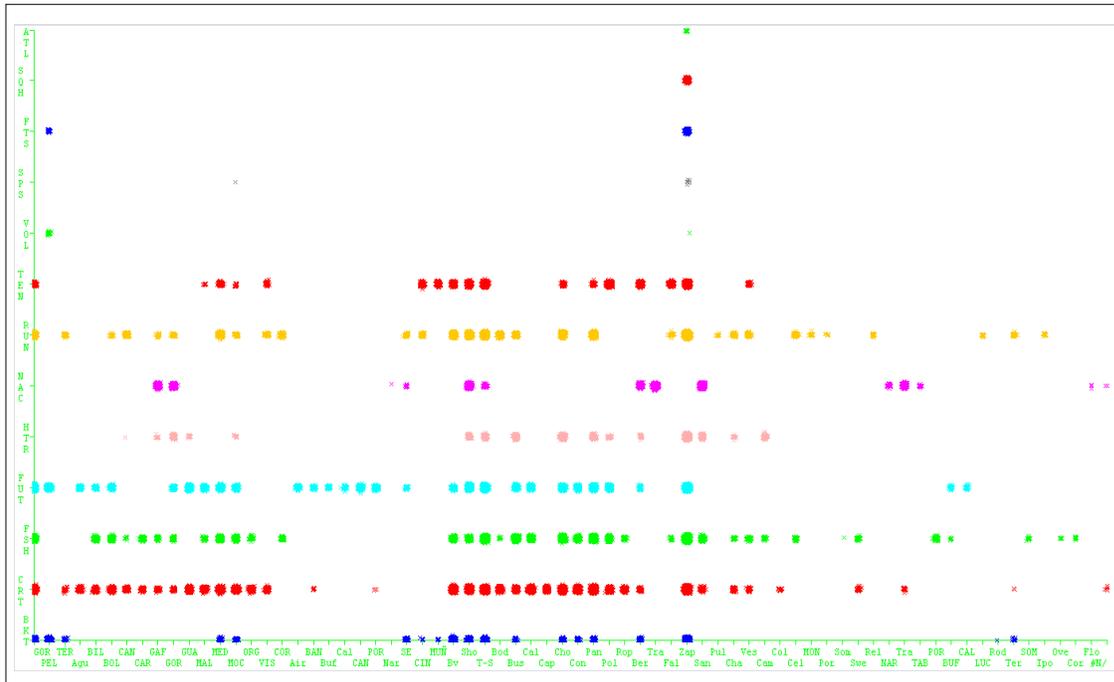


Figura 24 Contraste Línea / Actividad

- **Contrastación 15: Genero edad / Línea**

Por la gran cantidad de puntos en esta gráfica la visualización no permite obtener un análisis visual.

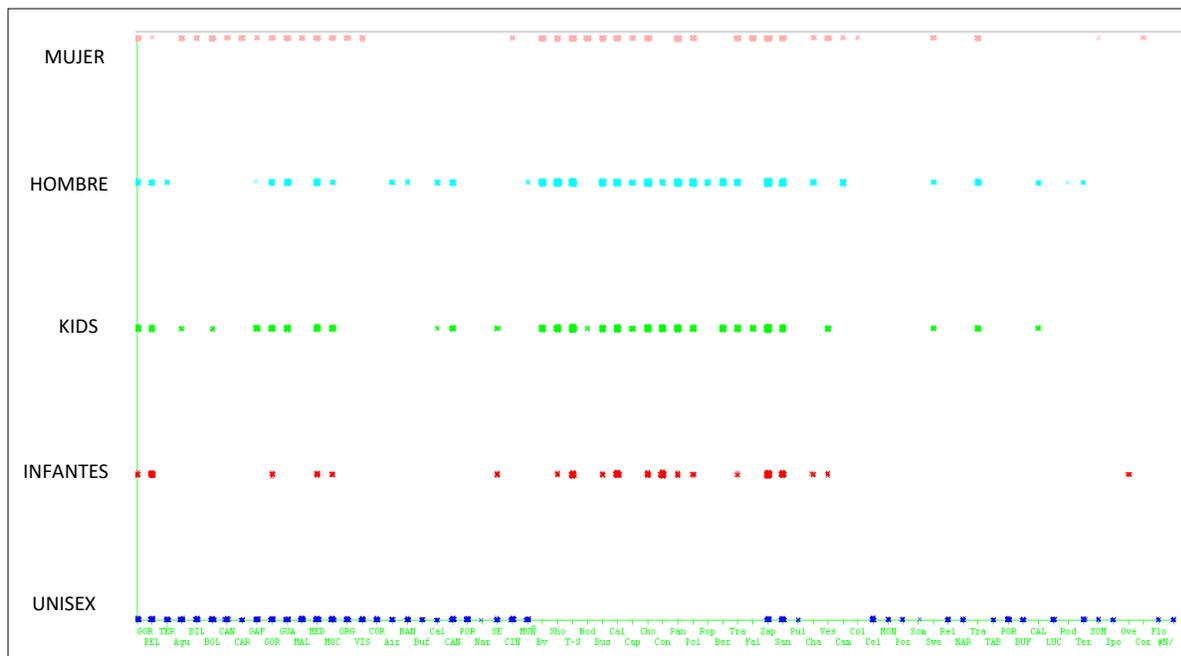


Figura 25 Contraste Género edad / Línea

- Contrastación 16: Región / Línea

Por la amplia cantidad de líneas no se puede obtener un claro análisis visual, pero resaltan puntos como, los organizadores, correas, pulseras, chalecos, calentadores y buzos tienen una mayor venta en la sierra que en la costa, y artículos como bufandas y gorros solo tienen venta en la sierra.

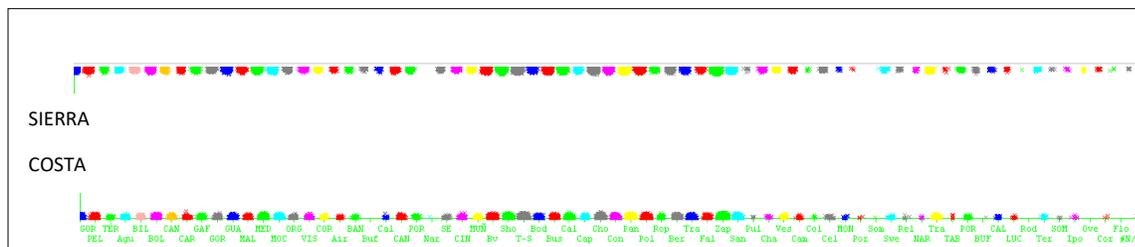


Figura 26 Contraste Región / Línea

- **Contrastación 17: Tienda / Línea**

La gráfica de visualización permite denotar varios comportamientos de ventas como, la línea sandalias tiene una venta mayoritaria en las tiendas de la costa y en la sierra es una presencia mínima siendo casi nula en Scala Shopping y Condado Shopping, como la línea de gorros y bufandas tiene presencia en la sierra y mínima o nula en la costa, mientras que líneas como botellas de agua tiene una presencia minoritaria en cuenca.

Líneas como relojes tienen participación nula en Condado Shopping, City Mall y Mall del Sol, y camisetas, shorts, calentadores, pantalones, no se puede visualizar un comportamiento de ventas mayores o menores en una tienda.

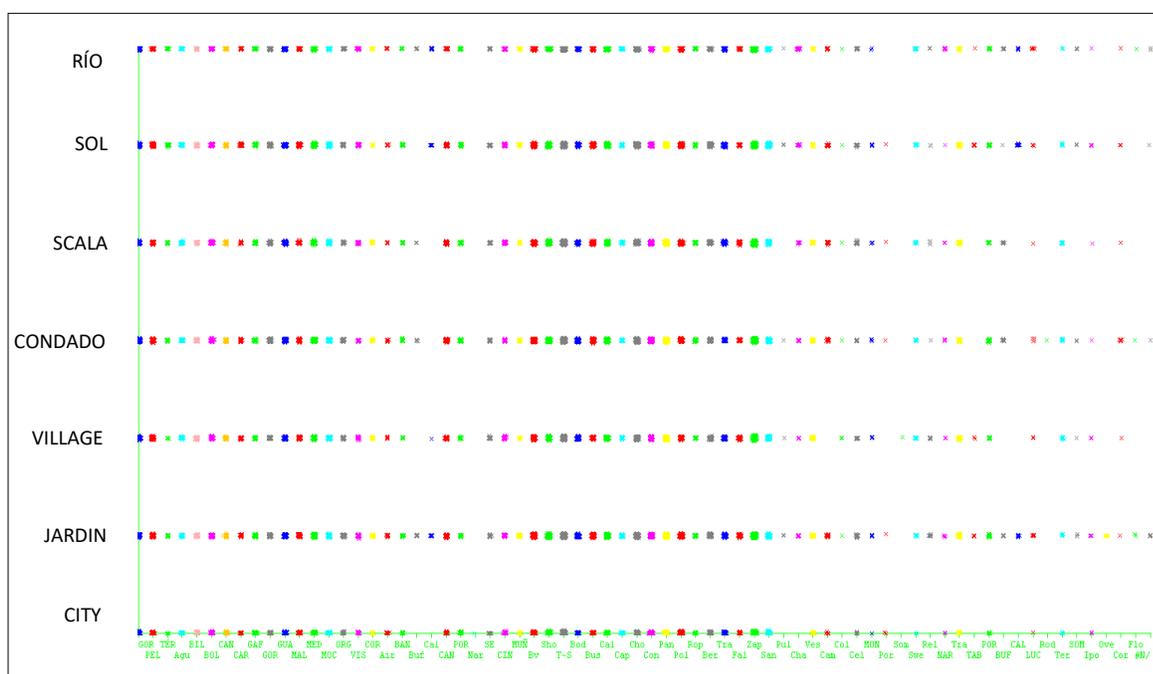


Figura 27 Contraste Tienda / Línea

- **Contrastación 18: Tienda / Género edad**

Encontramos que en el género mujer la tienda que proporciona menos nivel de ventas es City Mall, mientras que kids tenemos a Village Plaza y Scala Shopping con menor presencia.

Los géneros que representan menor ventas en todas las tiendas son unisex adulto e infantiles, resaltando infantiles como el género más pequeño en Village Plaza y unisex en Mall del Río.

En City Mall el género con mayor representación es hombre con un 49%, seguido de mujer con 14% y kids de 13%, finalmente infantiles es una pequeña venta del 4%.

En Condado Shopping la venta predominante es hombre adulto, seguido de mujer con 16% y kids de 15% e infantiles apenas es el 4%.

En Mall del Río la venta a nivel de género se comporta un 43% en hombre, 18% en mujer, un 16% en kids y un 5% en infantiles.

En Mall el Jardín la venta a nivel de género se comporta con un 42% en venta de hombre adulto, el 19% mujer, 16% kids y el 5% infantiles.

En Scala Shopping al analizar la venta por género el 40% es hombre adulto, el 18% es mujer el 16% kids y el 5% infantiles.

En la tienda de Mall del Sol, la venta de hombre es 48% casi la mitad, seguida de 16% de mujer, 14% en kids y 4% en infantiles.

En Village Plaza la venta a nivel de género es de 41% en hombre, 18% mujer, 17% kids y 4% infantiles, siendo muy similares los porcentajes entre mujer y kids.

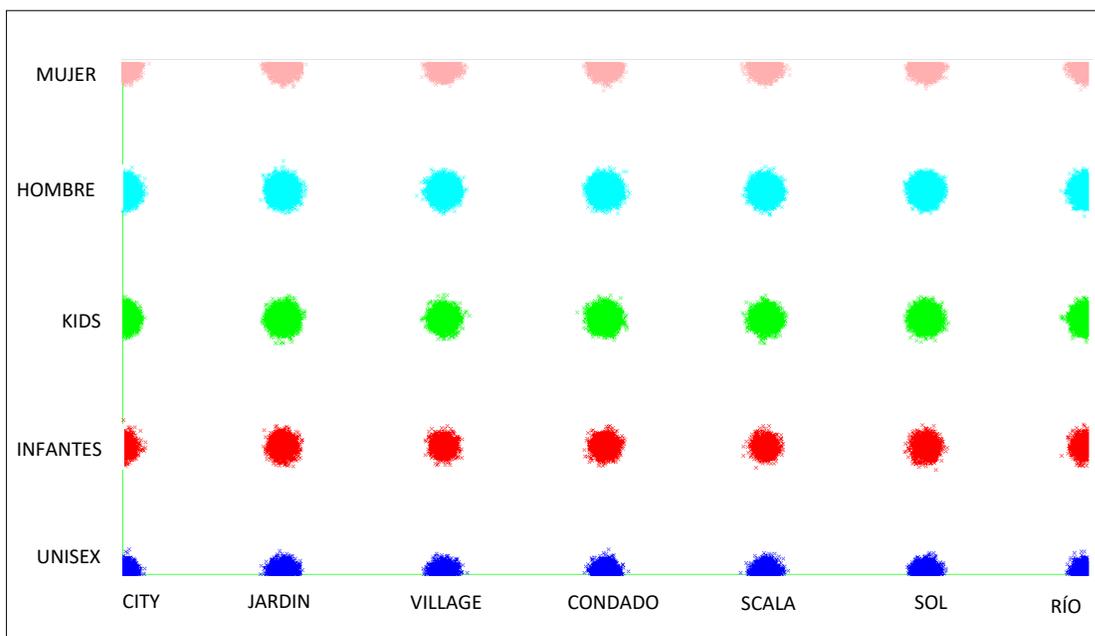


Figura 28 Contraste Tienda / Género edad

- **Contrastación 19: Región / Mes**

Al realizar este análisis encontramos que en meses como mayo, noviembre, septiembre y agosto la venta es mayor en la región sierra, mientras que en los meses de marzo, abril y junio las ventas mayores se dan en la costa.

Lo que nos lleva a relacionar que dentro del período de vacaciones las ventas en las tiendas deportivas aumentan.



Figura 29 Contraste Región / Mes

- **Contrastación 20: Género edad / mes**

Se denota un crecimiento en las ventas en el género de mujer en el mes de mayo y el de menor en enero, en el de hombre se da la mayor venta en el mes de junio y la menor en octubre.

También denotamos claramente que la menor proporción de ventas en todos los meses recae en los infantes, siendo el género menor vendido, repuntando solamente en noviembre y diciembre.

El género kids tiene una constante de ventas con pequeños incrementos en los meses de abril y agosto, estos coinciden con el regreso a clases, viendo una oportunidad de crecimiento en estos meses en las ventas.

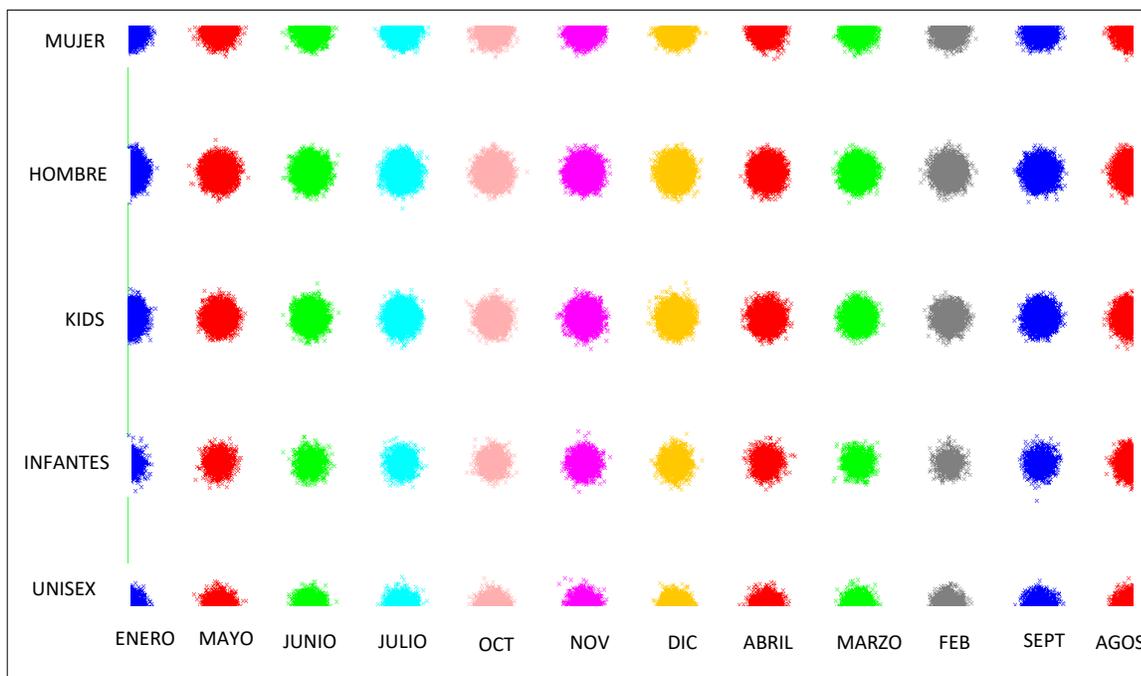


Figura 30 Contraste Género edad / Línea

- Contrastación 21: Grupo / Línea

Aquí podemos realizar el análisis a nivel de grupo, ya que en contrastes anteriores por la gran cantidad de líneas no hemos podido tener un análisis a profundidad. El grupo zapatos, contamos solo con dos líneas, siendo la de menor venta sandalias. El grupo de ropa tenemos una presencia mayoritaria en t-shirts, shorts, pantalones y polos, y de una venta casi nula y no representativa para profundizar en las compras de overoles y sacos. En accesorios tenemos una amplia gama, pero los más representativos en ventas son las gorras, medias y mochilas.

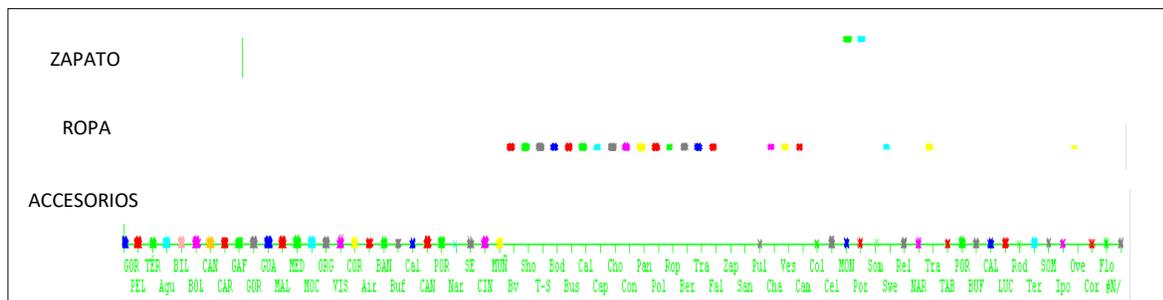


Figura 31 Contraste Grupo / Línea

Para realizar un análisis a nivel de tienda, se cruzan las variantes de género edad y talla, los resultados fueron:

- City Mall (Guayaquil)

Grupo / Género: En el grupo de zapatos tiene la mayor venta en hombre por una diferencia considerable de 62% de venta vs al segundo lugar que es compartido entre kids y mujer con el 15% y finalmente en 7% en infantes. En ropa tiene una venta en hombre adulto de 65% y el segundo lugar un empate de 17% en kids y mujer y finalmente 2% en infante. En accesorios la mayor venta en todas las tiendas es en unisex adulto.

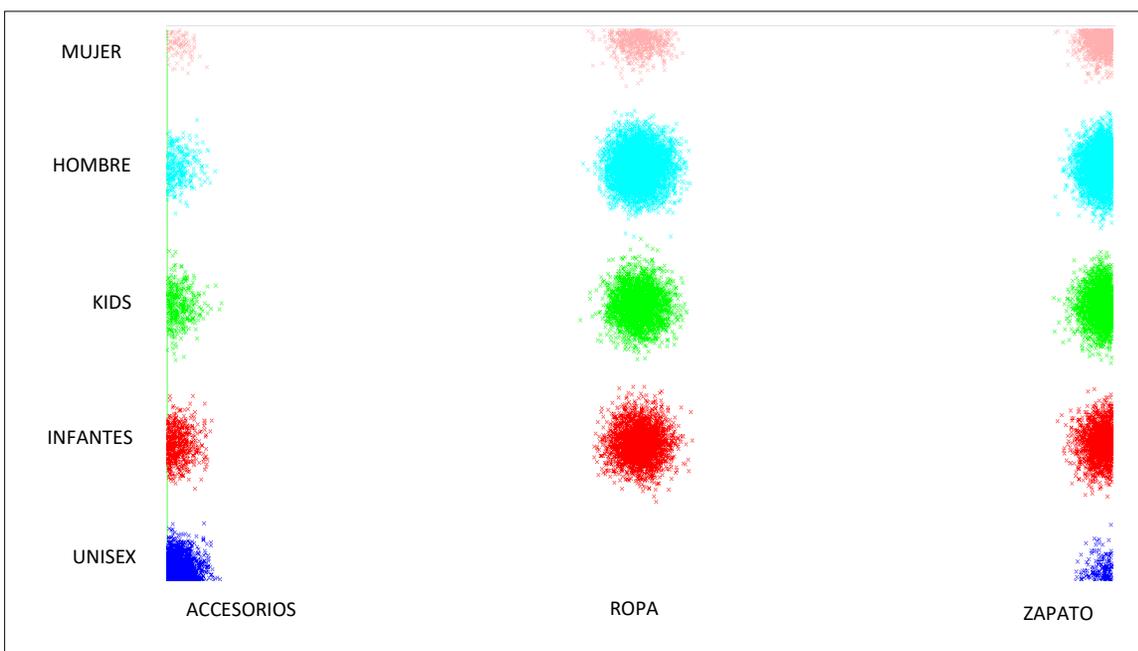


Figura 32 Contraste Grupo / Género edad, Tienda: City Mall

Género/Talla: Al analizar City Mall a nivel de talla, en zapatos hombre las tallas más vendidas con casi igual participación son 8.0-8.5 con 33% y 9.0-9.5 con 32%, la siguiente talla tiene una participación 17% en la talla de 7.0-7.5, en zapatos mujer la talla más vendida es 6.0-6.5 con 35% y 31% en 7.0-7.5 y la siguiente talla ocupa el 15% es 8.0-8.5, En kids zapatos, existen varias tallas con igual porcentaje de ventas 14% de 2.0 a 4.5, e infantes al igual existe un rango de tallas más vendidas que son de 7k a 9.5k.

En City Mall en la venta de ropa, en hombre adulto la talla más vendida es M con un 41%, mientras que S es 28% y L es 24%, en mujer la talla más vendida es S con un 40%, M el 30% y

XS 22%, En kids M vende el 30%, L el 25% S, el 22% y XL 11%, y en infantiles 24% en 3T, 23% 2T, 20% en 4T y 18% en 18 meses.

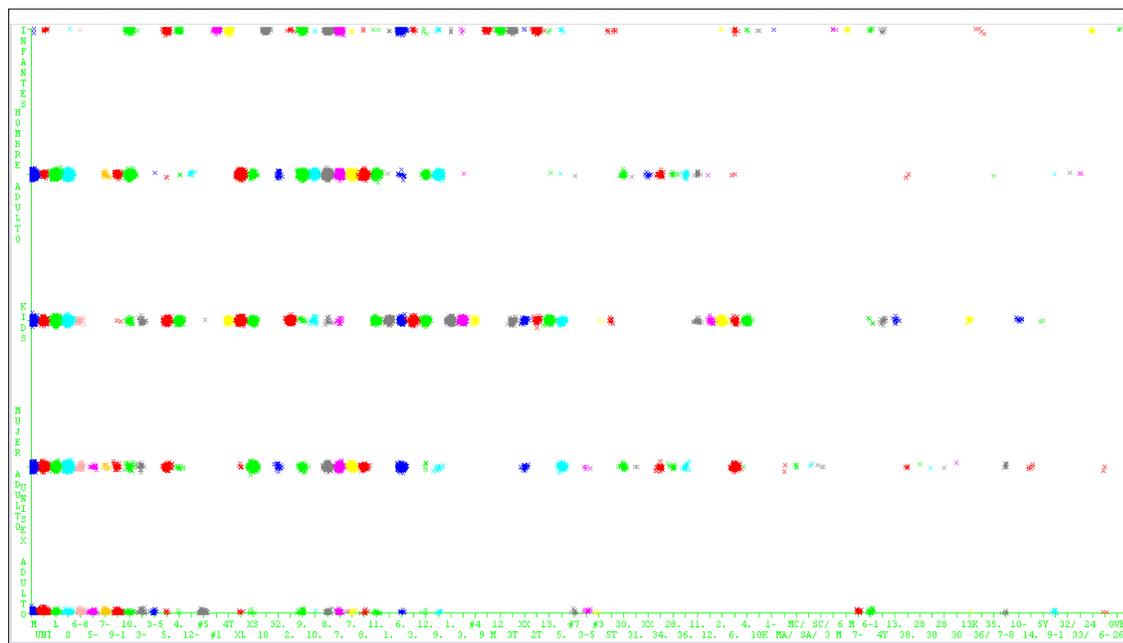


Figura 33 Contraste Género edad / Talla, Tienda: City Mall

- Condado Shopping (Quito)

Grupo / Género: al analizar la venta de zapatos el género con mayor participación es hombre adulto con 58%, mientras que kids ocupa el 19%, mujer el tercer lugar con un 16% y finalmente infantiles con un 19%, de igual manera al analizar ropa tenemos que el género con mayor venta es hombre adulto, seguido de mujer adulto con el 22% y kids con un 18%, e infantiles la menor con un 3%.

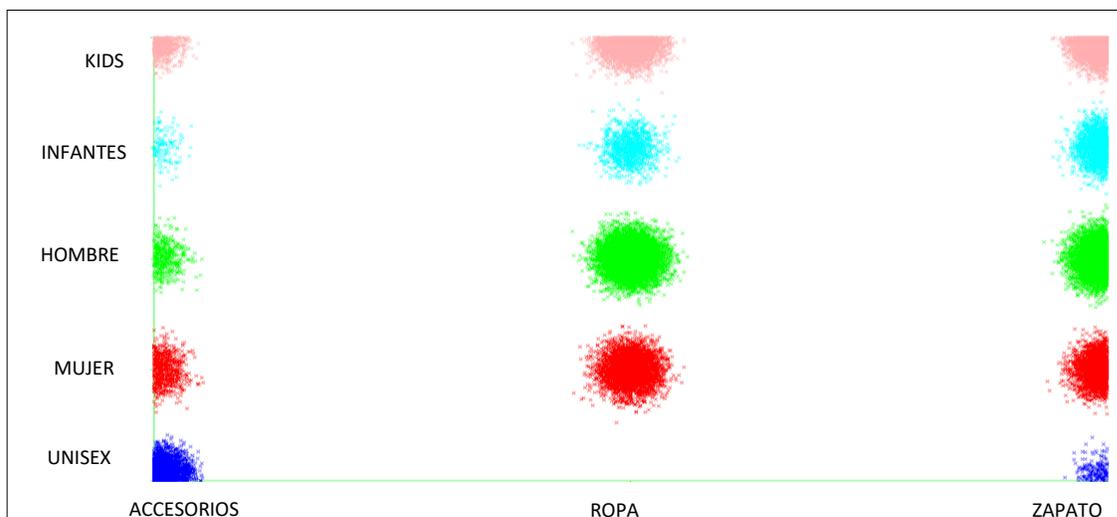


Figura 34 Contraste Grupo / Género edad, Tienda: Condado Shopping

Género / Talla: La tienda Condado Shopping al analizar en zapatos en hombre adulto las tallas más vendidas son 8.0-8.5 con un 36%, mientras que 9.0-9.5 ocupa un 31% y 7.0-7.5 un 18%, en mujer tenemos que el 37% compran de 6.0-6.5, el 31% de 7.0-7.5 y el 17% de 5.0-5.5, mientras que en kids existe un rango de tallas más vendidas que es desde la 1.0 a la 4.5, y en infantes es de 7.0 a 8.5 tienen mayor venta.

Al analizar el comportamiento de venta de ropa en esta tienda, tenemos que las tallas con mayor venta en hombre es M con un 40%, S 31%, L 21%, mientras que en mujer la talla más vendida es S con un 40%, le sigue M con un 30%, XS con un 21%, en el caso de kids la talla más solicitada es M con un 30%, L 25%, S 23%, y finalmente infantes las tallas con mejores ventas son 2T, 3T y 4T.

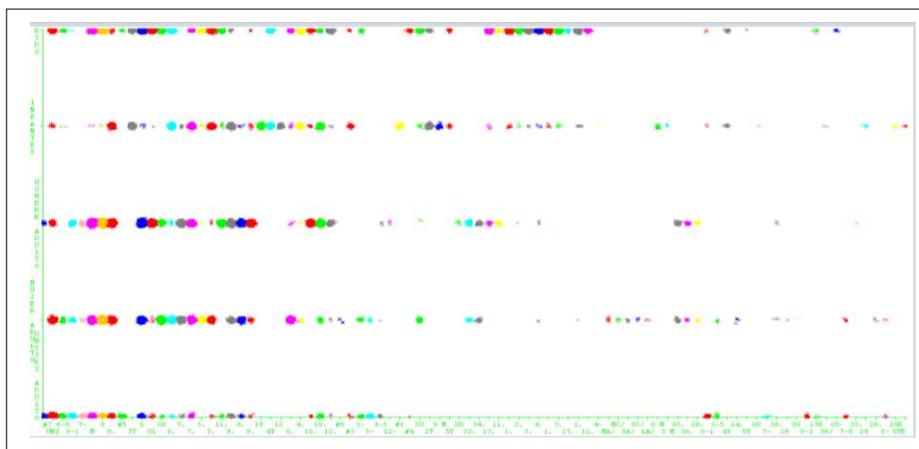


Figura 35 Contraste Género edad / Talla, Tienda: Condado Shopping

- Mall del Río (Cuenca)

Grupo / Género: La tienda Mall del Río en zapatos tiene una mayor venta en hombre adulto con un 56%, seguido de kids con un 20%, mujer representa un 16%, existiendo una diferencia notable entre estos dos géneros, y finalmente infantiles con un 7%. En el caso de ropa, la mayor veta se repite enfocada a hombre con un 54%, pero en este caso el segundo lugar lo ocupa mujer con un 26% y kids es un 18%, al igual que en zapatos la diferencia es amplia, y finalmente el 3% está representado por infantiles.

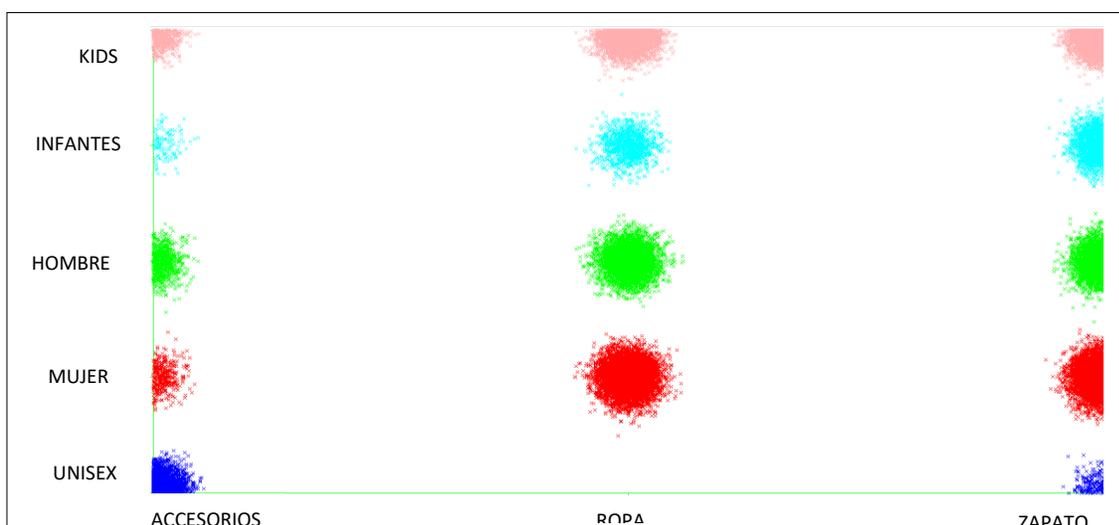


Figura 36 Contraste Grupo / Género edad, Tienda: Mall del Río

Género / Talla: Mall del Río a nivel de zapatos la venta mayor en talla de adulto es de 8.0-8.5 con un 34% mientras que la talla de 9.0-9.5 es del 32% y 7.0-7.5 es de 17%, en mujer las tallas con mayor veta son de 6.0-6.5 con un 35% y de 7.0-7.5 con un 32%, en kids el rango de tallas de 1.0 a 4.5 tiene un dominio en las ventas y en infantiles de 7k a 9.5k.

En el caso de ropa en Mall del Río tenemos que en hombre el 41% compra la talla M, el 32% la talla S, 21% L, y 5% apenas XL, en mujer encontramos que el 40% compra S, el 30% M, el 20% XS Y EL 8% XL, en el caso de kids tenemos que en M es la talla más vendida y en infante 2T, 3T y 4T.

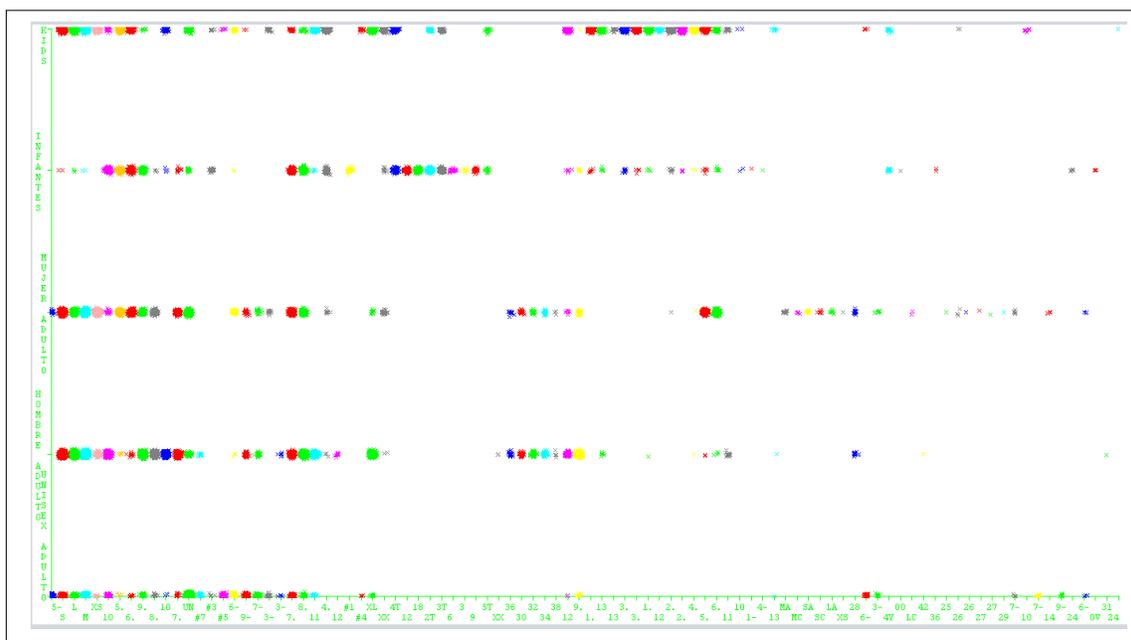


Figura 37 Contraste Género edad / Talla, Tienda: Mall del Río

- Mall el Jardín (Quito)

Grupo / Género: En la tienda Mall el Jardín la venta de zapatos tiene mayor participación hombre adulto con 52%, 21% en kids, 17% en mujer, y 9% infantes, mientras que en ropa sigue la mayor participación en hombre adulto, pero en el segundo lugar ocupa mujer con 25%, y en tercer lugar kids con un 18%, hay una diferencia muy notable, y finalmente infantes con un 3%.

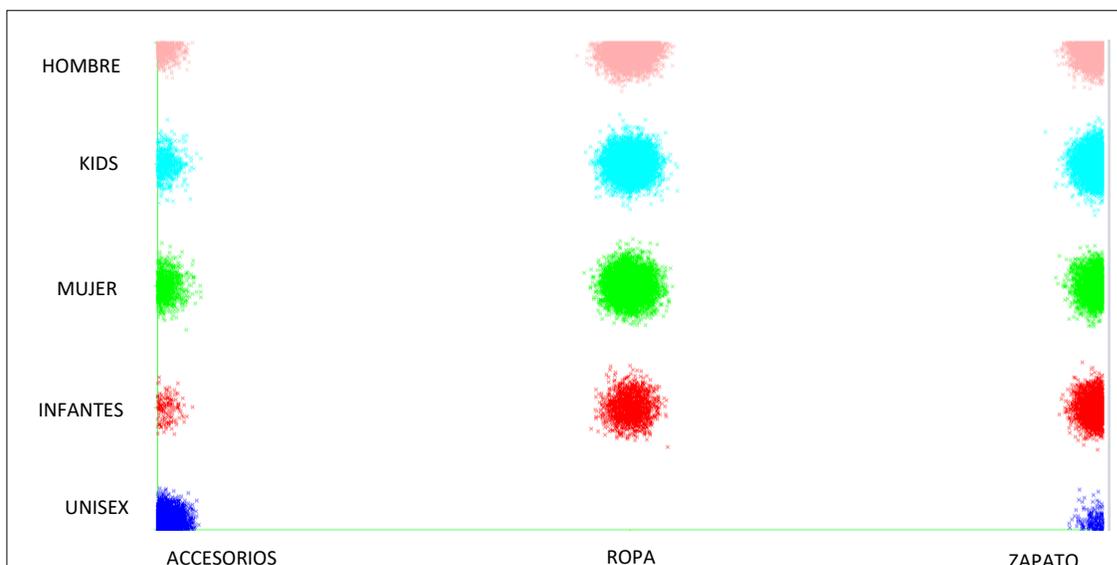


Figura 38 Contraste Grupo / Género edad, Tienda: Mall el Jardín

Género / Talla: En la tienda Mall el Jardín, al analizar las tallas vendidas en zapatos encontramos que en hombre adulto la más vendida es 8.0-8.5 con el 35%, seguida de 9.0-9.5 con el 31% y luego con una diferencia alta esta 7.0-7.5 con un 17%, en el caso de mujer tenemos que la talla más vendida es 6.0-6.5 con el 36%, y con el 31% 7.0-7.5, mientras que 5.0-5.5 es el 16%, en el caso de niños tenemos que el rango con mayor demanda es de 1.0-4.5, y en infantes de 7.0 a 8.5. En el caso de ropa en esta tienda tenemos que en hombre adulto tiene una venta de 40% en talla M, 31% en S, 22% en L, y 6% en XL, en mujer el 40% está en S, el 30% en M, el 20% en XS, y el 8% en L, en el caso de kids tenemos que el 30% es venta en M, y tanto S como L es el 23%, en infantes las tallas más vendidas son 2T, 3T y 4T.

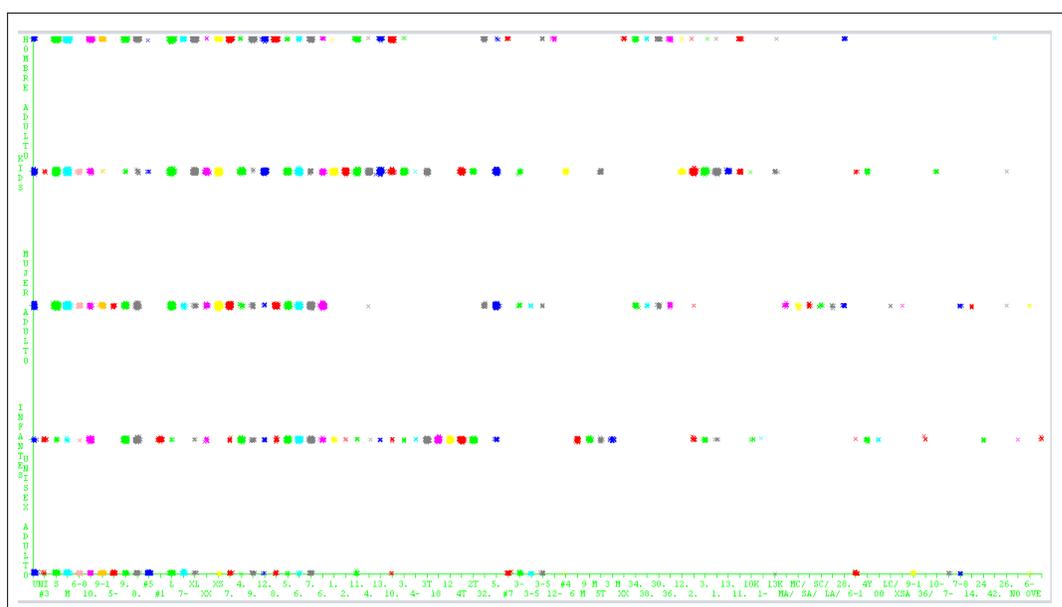


Figura 39 Contraste Género edad / Talla, Tienda: Mall el Jardín

- Scala Shopping (Quito)

Grupo / Género: En la tienda de Scala Shopping al analizar la venta de zapatos el 53% está en hombre, el 22% en kids, el 17% en mujer y el 7% en infantes, mientras que en ropa se mantiene en primer lugar hombre con 52%, pero en segundo lugar esta mujer con 19%, y kids en tercero con 15%, mientras que infantes representa apenas el 3%.

- Mall del Sol (Guayaquil)

Grupo / Género: En la tienda Mall del Sol al analizar la venta de zapatos encontramos que el 61% pertenecen a hombre adulto, en segundo lugar existe un empate entre kids y mujer con un 16% y finalmente tenemos a infantiles con un 7%, en el caso de ropa la predominancia del 61% se mantiene en hombre, pero el segundo lugar lo ocupa mujer con un 20%, seguido de kids con un 17%, y finalmente infante con una presencia pequeña de 2%.

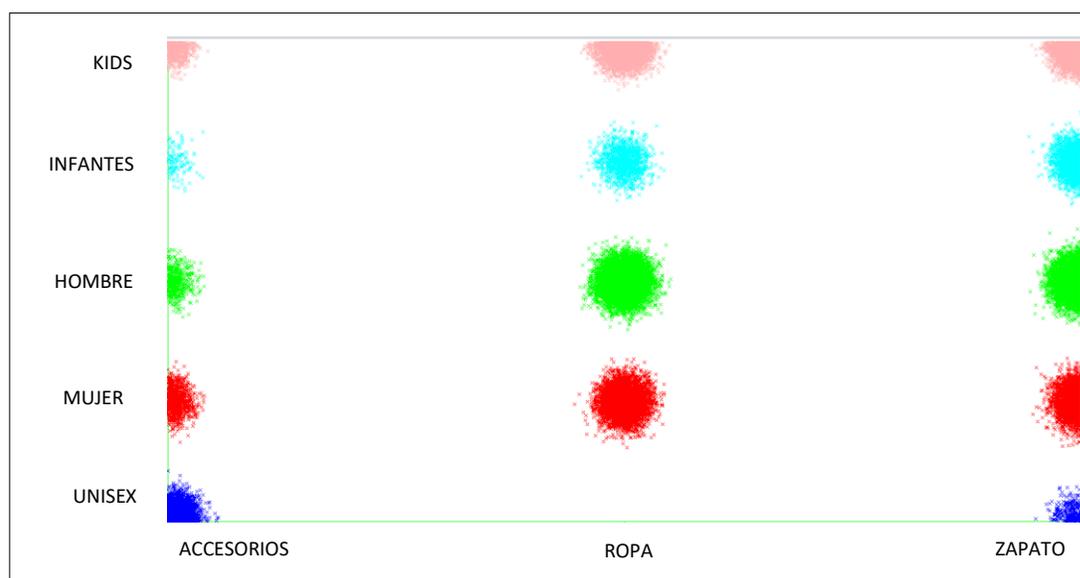


Figura 42 Contraste Grupo / Género edad, Tienda: Mall del Sol

Género / Talla: En la tienda de Mall del Sol analizar la venta por tallas encontramos que en zapatos, hombre tiene una venta muy similar de 33% en 8.0-8.5 y 9.0-9.5 con un 32%, el tercer lugar lo ocupa 7.0-7.5 con un 16% y 10.0-10.5 con un 14%, en mujer los resultados también se dan en dos rangos de tallas muy parejos como lo son el 6.0-6.5 con el 34% y el 7.0-7.5 con un 31%, en tercer lugar tenemos 5.0-5.5 con un 16% y 8.0-8.5 con el 15%, en el caso de niños existe el rango de tallas más vendido que va de 1.0 a 4.5, y en infantiles con una diferencia mínima está el rango de 7k a 8.5k con el rango de 9k.

En el caso de ropa tenemos que en hombre la talla más vendida es M con un 41%, S 28% L 24% y XL 7%, en mujer el 40% es S, el 29% es M, el 21% es XS y el 8% es L, En kids tenemos que M ocupa un 29%, L 25% y S 23%, y en infantiles la talla 2T, 3T, y 4T son las de mayor venta.

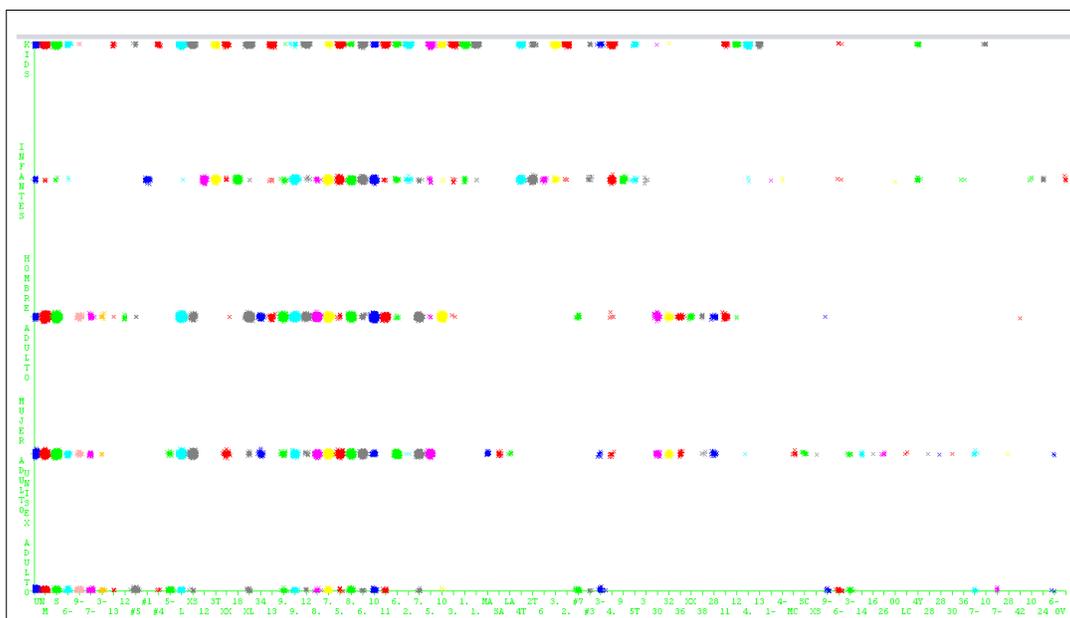


Figura 43 Contraste Género edad / Talla, Tienda: Mall del Sol

- Village Plaza (Guayaquil)

Grupo / Género: En la tienda Village Plaza a nivel de grupo, tenemos que en zapatos el 53% es venta de hombre, 21% en kids, 16% mujer y 9% infantes, en el caso de ropa el 54% ocupa hombre, mostrando en esta tienda una predominancia total, el segundo lugar cambia de posición y con el 24% lo ocupa mujer, mientras que kids ocupa el 20% y el 2% en infantes.

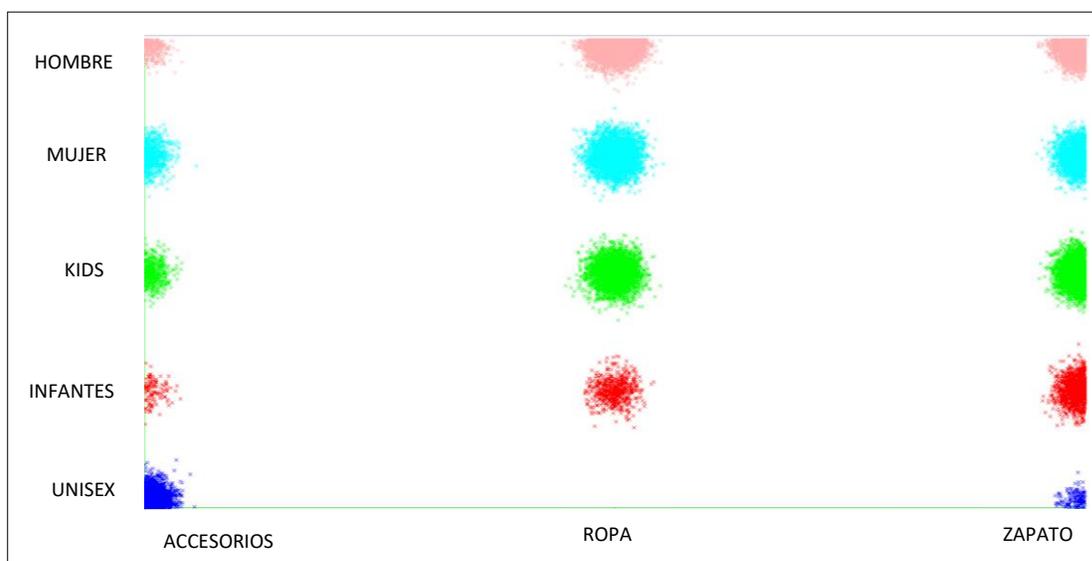


Figura 44 Contraste Grupo / Género edad, Tienda: Village Plaza

Género / Talla: En la tienda Village Plaza al analizar las ventas a nivel de tallas encontramos que en zapatos, en hombre existe una venta por igual del 33% en 9.0-9.5 y 8.0-8.5, lo que no había sucedido en otras tiendas, en esta tienda existe una venta de tallas grandes, en tercer lugar esta 10.0-10.5 con un 15% y 7.0-7.5 con un 14%, en el caso de mujer tenemos que 6.0-6.5 está en 34% y 7.0-7.5 es un 32%, mientras que el 17% es 8.0-8.5 y el 13% es 5.05-5.5, en kids tenemos que existe un rango de 1.0 a 4.5 de tallas más solicitadas, y en infantiles de 7k a 9.5k.

En ropa tenemos en hombre que el 42% la venta es M, el 27% es S, el 23% es L, y el 6% es XL, en mujer el 41% es S, 29% M, 22% XS y 7% L, kids, se mantiene el 30% en M, 25% en L, y el 23% en S, mientras que infantiles la talla más vendida es 3T, con un 23% mientras que 4T y 2T son el 21%.

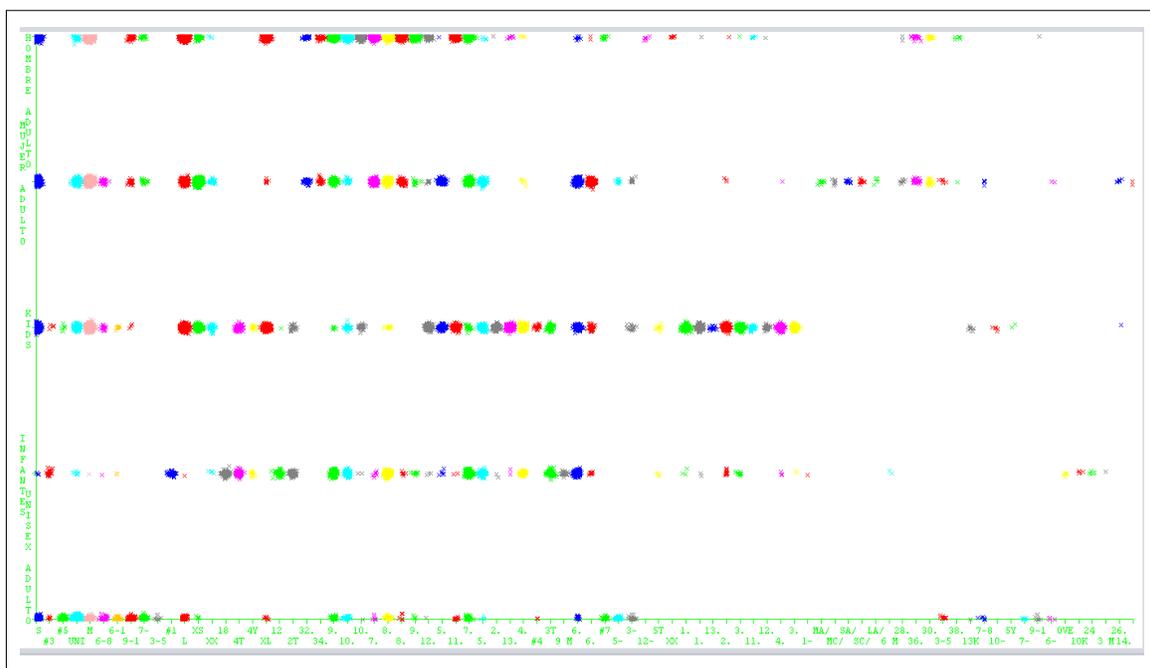


Figura 45 Contraste Género edad / Talla, Tienda: Village Plaza

Después de este análisis encontramos un perfil a nivel de tienda:

- City Mall (Guayaquil)

La categoría con mayor venta en unidades es training, seguida de fútbol, fashion y running; mientras que el género con mayor aportación es hombre, y mostrando una superioridad de kids sobre las ventas de mujer.

El grupo más significativo en la tienda es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre adulto son las mayores con 62%, mujeres y kids se encuentran en un 15% y las menores son las de infantes con un 7%, en ropa se replica este comportamiento y en accesorios las ventas en más del 80% son unisex.

Al analizar a nivel de tallas encontramos en zapatos, la talla con mayor venta en hombre es 8.0-8.5, en mujer 6.0-6.5, en kids 4.0-4.5 y finalmente en infantes es de 8k-8.5k.

En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 años.

- Condado Shopping (Quito)

La categoría con mayor venta en unidades es training, seguida de fútbol, fashion y running, existiendo una gran diferencia entre las ventas estas dos últimas, al analizar las ventas por géneros el más representativo es hombre seguido de mujer, luego kids y finalmente infantes, cabe mencionar que la diferencia entre las ventas de kids y mujer es mínima.

El grupo más significativo en la tienda es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre adulto son las mayores con 58%, kids un 19%, mujer 16% e infantes con un 6%, en ropa el comportamiento en hombre se mantiene siendo número uno y con mayor preferencia, pero en segundo lugar tenemos a las ventas de mujer, y luego las de niños, por lo cual podemos ver que es una tienda que tiene una venta dividida para estos dos géneros.

Al analizar a nivel de tallas encontramos en zapatos, la talla con mayor venta en hombre es 8.0-8.5, en mujer 6.0-6.5, en kids 4.0-4.5 y finalmente en infantes es de 7k-7.5k.

En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 años.

- Mall del Río (Cuenca)

La categoría con mayor venta en unidades es training, seguida de fútbol, running y fashion; mientras que el género con mayor aportación es hombre, seguido de mujer y finalmente al género de kids con infantes.

El grupo más significativo en la tienda es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre adulto son las mayores, luego hay una preferencia con una diferencia notable para kids y luego para mujer y finalmente infantes, en ropa cambia este comportamiento ya que la diferencia entre el segundo lugar ocupado por mujeres y el tercer lugar de kids, es de 8 puntos.

Al analizar a nivel de tallas encontramos en zapatos, la talla con mayor venta en hombre es 8.0-8.5, en mujer 6.0-6.5, en kids 3.0-3.5 y finalmente en infantes es de 8k-8.5k.

En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 4 años.

- Mall del Jardín (Quito)

La categoría con mayor venta en unidades es training, fútbol, fashion y running; la diferencia entre estas dos últimas categorías es mínima, al analizar por género con mayor aportación es hombre, mujer, kids e infantes.

El grupo más significativo en la tienda es ropa con un 47% seguido de zapatos y por último de accesorios, pero estos porcentajes se mueven al analizarlo a nivel de dólares por el precio promedio de zapatos es más alto que ropa. Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre adulto son las mayores,

seguidas de kids, luego mujeres y por último infantes, en ropa tenemos un comportamiento diferente en el según y tercer lugar, ya que en este el segundo lugar corresponde a mujer y el tercero con una diferencia significativa a kids y por último infantes.

Al analizar a nivel de tallas encontramos, en zapatos, la talla con mayor venta en hombre es 8.0-8.5, en mujer 6.0-6.5, en kids 3.0-3.5 y finalmente en infantes es de 8k-8.5k.

En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 años.

- Scala Shopping (Quito)

La categoría con mayor venta en unidades es training, fútbol, running y fashion, mientras que el género con mayor aportación es hombre, mujer, kids e infantes.

El grupo con mayor venta es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre adulto, seguido de kids, mujeres e infantes, en ropa el comportamiento es de hombres, mujeres, kids e infantes.

Al analizar a nivel de tallas encontramos, en zapatos, la talla con mayor venta en hombre es 8.0-8.5, en mujer 6.0-6.5, en kids 4.0-4.5 y finalmente en infantes es de 8k-8.5k. En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 años.

- Mall del Sol (Guayaquil)

La categoría con mayor venta en unidades es training, seguida de fútbol, running y fashion; mientras que el género con mayor aportación es hombre adulto, mujer, siendo el género adulto el de mayor predominancia.

El grupo más significativo en la tienda es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre son las mayores, mujeres y kids se encuentran en un 15% y las menores son las de infantes con un 7%, en ropa el comportamiento cambia, el número uno es hombre adulto, seguido de mujer y kids, mientras que en accesorios las ventas en más del 80% son unisex.

Al analizar a nivel de tallas encontramos, en zapatos, la talla con mayor venta en hombre es 8.0-8.5 pero con una venta muy similar en 9.0 a 9.5; en mujer 6.0-6.5, en kids 3.0-3.5 y finalmente en infantes es de 8k-8.5k. En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 y 2 años.

- Village Plaza (Guayaquil)

La categoría con mayor venta en unidades es training, seguida de fútbol, running y fashion; mientras que el género con mayor venta es hombre, mujer y kids.

El grupo más significativo en la tienda es ropa, seguido de zapatos y por último de accesorios, pero al momento de analizarlo las ventas en dólares los zapatos tienen venta mayor por su precio promedio.

Al analizar las ventas a nivel de grupo y relacionarlo con género encontramos que en el grupo zapatos las ventas de hombre, seguido de kids, mujeres e infantes. En el caso de ropa tenemos un comportamiento de las ventas más altas en hombre adulto, mujer, kids e infantes.

Al analizar a nivel de tallas encontramos, en zapatos, la talla con mayor venta en hombre es 9.0-9.5, seguida muy de cerca el 8.0-8.5, un comportamiento que estas tallas grandes solo se da en la región costa, en mujer 6.0-6.5, al igual muy de cerca esta las tallas 7.0-7.5, en kids hay un empate desde la talla 1.0 a 4.5 y finalmente en infantes también el rango es de 7k-9.5k. En el grupo de ropa el comportamiento se da en hombre la talla con mayores ventas es M, en mujer la S, en niño M, y en infante en 3 años.

5.2 Tarea de Asociación

Los resultados encontrados en la tarea de asociación son importantes en la toma de decisiones para la realización de una compra efectiva, ya que por el giro de negocio y la estructura para realizar el proceso de compra se necesita conocer con precisión los outfits y mix de producto a ingresar, involucrando a su vez puntos estratégicos como visual merchandising en cada tienda.

Es así que a nivel de todas las tiendas podemos concluir que la zona de zapatos, es un área caliente en la tienda, es decir un área con niveles de venta altos, pero la venta de un zapato es casi excluyente de la venta de ropa, del 2% al 10% de las personas que compran un zapato buscan asociar su compra con ropa. Por lo cual cuando se realice la compra de estos ítems no es fundamental una asociación con los colores, tecnologías y beneficios de las prendas de vestir y zapatos.

En el caso de artículos de vestir, tenemos un comportamiento diferente y aquí las asociaciones juegan un papel importante en el cierre de una venta y en el incremento de los volúmenes de compra.

Los resultados encontrados en la tarea de asociación al realizar el cruce general de las líneas que se venden en estas tiendas encontramos

- **City Mall (Guayaquil)**

- En el año 2013, encontramos que el 92% de las personas que no compran zapatos adquieran un pantalón y t-shirt, en el 2014 este porcentaje se convierte a 90% mientras que en el 2013 cuando compran un pantalón existe un 26% de probabilidad de comprar una t-shirt en el 2014 es del 22%, y la condición de cuando se compra una t-shirt se complemente la compra con un pantalón en ambos años es del 6%.
- En el año 2013 y 2014 el comportamiento es que el 91% de las personas que no adquieren un zapato realiza una compra de short y t-shirt, en el año 2013 el 22% de las personas que compran un short adquieren también una t-shirt, en el 2014 este porcentaje disminuye a 19%; y cuando compran una t-shirt existe un 9% de probabilidad de comprar un short; en el 2013 y de 7% en el 2014.

- Por los resultados hallados en 2013, 2014 y 2015 podemos asegurar que el comportamiento de compra de esta tienda existe una preferencia de las t-shirt sobre los bvd, ya que el 95% de las personas que no adquieren un bvd, adquieren una combinación sea en short y una camiseta, o una camiseta y un pantalón.
- Encontramos que el 96% de las personas que no adquieren un Bvd, dentro de su compra se encuentra una t-shirt y un zapato; cuando solo adquieren la t-shirt el 5% también compra el zapato, y cuando compran el zapato solo el 4% adquiere la t-shirt.
- En el año 2016 el comportamiento de compra de la combinación con t-shirt sigue siendo la más alta en la tienda con un 94%, y a la vez que la camiseta es el complemento escogido para los pantalones y los shorts; mientras que cuando adquieren una t-shirt la probabilidad que la complemente es menos del 5%.
- En City Mall, las asociaciones más altas al comprar son con las t-shirt, esta prenda es el complemento en 24% para pantalones y 22% para shorts. En esta tienda las t-shirt tienen una predominancia sobre los bvd.

```

8. MEDIAS=SI Zapato=SI 795 ==> T-Shirt=NO 636 <conf:(0.8)> lift:(1.07) lev:(0) [40] conv:(1.25)
9. Chompa=SI 774 ==> T-Shirt=NO 619 <conf:(0.8)> lift:(1.07) lev:(0) [39] conv:(1.25)
10. Agua=SI 227 ==> T-Shirt=NO 179 <conf:(0.79)> lift:(1.05) lev:(0) [9] conv:(1.16)
11. MEDIAS=SI 1693 ==> T-Shirt=NO 1334 <conf:(0.79)> lift:(1.05) lev:(0) [66] conv:(1.18)
12. Polo=SI 633 ==> T-Shirt=NO 483 <conf:(0.76)> lift:(1.02) lev:(0) [9] conv:(1.05)
13. Bvd=SI 605 ==> T-Shirt=NO 451 <conf:(0.75)> lift:(1) lev:(-0) [-1] conv:(0.98)
14. GORRA=SI Zapato=SI 305 ==> T-Shirt=NO 199 <conf:(0.65)> lift:(0.87) lev:(-0) [-29] conv:(0.72)
15. Short=SI 1597 ==> T-Shirt=NO 1004 <conf:(0.63)> lift:(0.84) lev:(-0.01) [-191] conv:(0.68)
16. Pantalon=SI 891 ==> T-Shirt=NO 507 <conf:(0.57)> lift:(0.76) lev:(-0.01) [-160] conv:(0.58)
17. Short=SI Zapato=SI 351 ==> T-Shirt=NO 185 <conf:(0.53)> lift:(0.7) lev:(-0.01) [-77] conv:(0.53)
18. MEDIAS=SI T-Shirt=NO 1334 ==> Zapato=SI 636 <conf:(0.48)> lift:(1.16) lev:(0.01) [87] conv:(1.12)
19. T-Shirt=NO 11313 ==> Zapato=SI 5384 <conf:(0.48)> lift:(1.16) lev:(0.05) [728] conv:(1.12)
20. MEDIAS=SI 1693 ==> Zapato=SI 795 <conf:(0.47)> lift:(1.14) lev:(0.01) [98] conv:(1.11)
21. MEDIAS=SI 1693 ==> T-Shirt=NO Zapato=SI 636 <conf:(0.38)> lift:(1.05) lev:(0) [32] conv:(1.03)
22. Chompa=SI 774 ==> Zapato=SI 188 <conf:(0.24)> lift:(0.59) lev:(-0.01) [-130] conv:(0.78)
23. Polo=SI 633 ==> Zapato=SI 153 <conf:(0.24)> lift:(0.59) lev:(-0.01) [-107] conv:(0.77)
24. Pantalon=SI 891 ==> Zapato=SI 209 <conf:(0.23)> lift:(0.57) lev:(-0.01) [-157] conv:(0.77)
25. Short=SI 1597 ==> Zapato=SI 351 <conf:(0.22)> lift:(0.53) lev:(-0.02) [-306] conv:(0.75)
26. Short=SI T-Shirt=NO 1004 ==> Zapato=SI 185 <conf:(0.18)> lift:(0.45) lev:(-0.02) [-228] conv:(0.72)
27. GORRA=SI 2039 ==> Zapato=SI 305 <conf:(0.15)> lift:(0.36) lev:(-0.04) [-534] conv:(0.69)
28. T-Shirt=NO 11313 ==> GORRA=SI 1685 <conf:(0.15)> lift:(1.1) lev:(0.01) [158] conv:(1.02)
29. Short=SI 1597 ==> MEDIAS=SI 211 <conf:(0.13)> lift:(1.18) lev:(0) [32] conv:(1.02)
30. Zapato=SI 6217 ==> MEDIAS=SI 795 <conf:(0.13)> lift:(1.14) lev:(0.01) [98] conv:(1.02)
31. MEDIAS=SI 1693 ==> Short=SI 211 <conf:(0.12)> lift:(1.18) lev:(0) [32] conv:(1.02)
32. T-Shirt=NO Zapato=SI 5384 ==> MEDIAS=SI 636 <conf:(0.12)> lift:(1.05) lev:(0) [32] conv:(1.01)
33. GORRA=SI T-Shirt=NO 1685 ==> Zapato=SI 199 <conf:(0.12)> lift:(0.29) lev:(-0.03) [-494] conv:(0.67)
34. T-Shirt=NO 11313 ==> MEDIAS=SI 1334 <conf:(0.12)> lift:(1.05) lev:(0) [66] conv:(1.01)
35. Short=SI 1597 ==> T-Shirt=NO Zapato=SI 185 <conf:(0.12)> lift:(0.33) lev:(-0.03) [-384] conv:(0.73)
36. MEDIAS=SI 1693 ==> GORRA=SI 180 <conf:(0.11)> lift:(0.79) lev:(-0) [-48] conv:(0.97)
37. Zapato=SI 6217 ==> MEDIAS=SI T-Shirt=NO 636 <conf:(0.1)> lift:(1.16) lev:(0.01) [87] conv:(1.02)

```

```

38. GORRA=SI 2039 ==> T-Shirt=NO Zapato=SI 199 <conf:(0.1)> lift:(0.27) lev:(-0.03) [-527] conv:(0.71)
39. T-Shirt=NO 11313 ==> Short=SI 1004 <conf:(0.09)> lift:(0.84) lev:(-0.01) [-191] conv:(0.98)
40. GORRA=SI 2039 ==> MEDIAS=SI 180 <conf:(0.09)> lift:(0.79) lev:(-0) [-48] conv:(0.97)
41. Zapato=SI 6217 ==> Short=SI 351 <conf:(0.06)> lift:(0.53) lev:(-0.02) [-306] conv:(0.95)
42. T-Shirt=NO 11313 ==> MEDIAS=SI Zapato=SI 636 <conf:(0.06)> lift:(1.07) lev:(0) [40] conv:(1)
43. T-Shirt=NO 11313 ==> Chompa=SI 619 <conf:(0.05)> lift:(1.07) lev:(0) [39] conv:(1)
44. Zapato=SI 6217 ==> GORRA=SI 305 <conf:(0.05)> lift:(0.36) lev:(-0.04) [-534] conv:(0.91)
45. T-Shirt=NO 11313 ==> Sandalia=SI 536 <conf:(0.05)> lift:(1.14) lev:(0) [66] conv:(1.01)
46. T-Shirt=NO 11313 ==> Pantalón=SI 507 <conf:(0.04)> lift:(0.76) lev:(-0.01) [-160] conv:(0.99)
47. T-Shirt=NO 11313 ==> Polo=SI 483 <conf:(0.04)> lift:(1.02) lev:(0) [9] conv:(1)
48. T-Shirt=NO 11313 ==> Bvd=SI 451 <conf:(0.04)> lift:(1) lev:(-0) [-1] conv:(1)
49. T-Shirt=NO Zapato=SI 5384 ==> GORRA=SI 199 <conf:(0.04)> lift:(0.27) lev:(-0.03) [-527] conv:(0.9)
50. T-Shirt=NO 11313 ==> PELOTA=SI 392 <conf:(0.03)> lift:(1.17) lev:(0) [57] conv:(1.01)
51. T-Shirt=NO Zapato=SI 5384 ==> Short=SI 185 <conf:(0.03)> lift:(0.33) lev:(-0.03) [-384] conv:(0.93)
52. Zapato=SI 6217 ==> Pantalón=SI 209 <conf:(0.03)> lift:(0.57) lev:(-0.01) [-157] conv:(0.97)
53. Zapato=SI 6217 ==> GORRA=SI T-Shirt=NO 199 <conf:(0.03)> lift:(0.29) lev:(-0.03) [-494] conv:(0.92)
54. Zapato=SI 6217 ==> Chompa=SI 188 <conf:(0.03)> lift:(0.59) lev:(-0.01) [-130] conv:(0.98)
55. Zapato=SI 6217 ==> Short=SI T-Shirt=NO 185 <conf:(0.03)> lift:(0.45) lev:(-0.02) [-228] conv:(0.96)
56. Zapato=SI 6217 ==> Polo=SI 153 <conf:(0.02)> lift:(0.59) lev:(-0.01) [-107] conv:(0.98)
57. T-Shirt=NO 11313 ==> BILLETERA=SI 260 <conf:(0.02)> lift:(1.17) lev:(0) [38] conv:(1)
58. T-Shirt=NO 11313 ==> MOCHILA=SI 223 <conf:(0.02)> lift:(1.17) lev:(0) [32] conv:(1)
59. T-Shirt=NO 11313 ==> GORRA=SI Zapato=SI 199 <conf:(0.02)> lift:(0.87) lev:(-0) [-29] conv:(1)
60. T-Shirt=NO 11313 ==> Short=SI Zapato=SI 185 <conf:(0.02)> lift:(0.7) lev:(-0.01) [-77] conv:(0.99)
61. T-Shirt=NO 11313 ==> Agua=SI 179 <conf:(0.02)> lift:(1.05) lev:(0) [9] conv:(1)
62. T-Shirt=NO 11313 ==> BOLSO=SI 166 <conf:(0.01)> lift:(1.12) lev:(0) [17] conv:(1)

```

Figura 46 Reglas de Asociación City Mall

- **Condado Shopping (Quito)**

- En el año 2013 existe una posibilidad de que el 95% de las personas que no compran un zapato, compren un pantalón y t-shirt, mientras que en el 2014 esta disminuye a 91%, lo mismo sucede en la asociación de compra de pantalón y complemente con una t-shirt cambia de 19% a 18%, y de las personas que adquieran una t-shirt y complementen con un pantalón en ambos años del 8%.
- Cuando una persona adquiere un short en esta tienda existe un 22% que compre una t-shirt, en el 2013 mientras que el 2014 este porcentaje baja a 20%; y en el caso de si compra una t-shirt también compre un short es del 8% en 2013 y 6% en 2014.
- En el año 2013 y 2015 encontramos un comportamiento en esta tienda de que un 17% de las personas que compran un pantalón compren una chompa y de un 13% de las que compran una chompa compren un pantalón.
- En esta tienda se encontró que la combinación de preferencia es entre pantalón y t-shirt, y las prendas que menos se complementan son los bvd y los shorts, alcanzando un 96% y 97% en estas combinaciones.

- El 95% de las personas no llevan ni el short ni el bvd en una misma compra.
- Al analizar la compra de pantalones y t-shirt es de 19% y de t-shirt y complemento pantalón es del 8%.
- Condado, en esta tienda existen varias combinaciones exitosas, pero la de mayor fuerza es la de pantalón y t-shirt, y también adicional a las relaciones encontradas en todas las tiendas esta de chompa y pantalón, es por eso que adicional a la t-shirt que en todas las plazas presenta una mayor asociación, los pantalones también presentan un alto nivel de complemento.

```

1. PELOTA=SI 1225 ==> Zapato=NO 1225 <conf:(1)> lift:(1.64) lev:(0.01) [479] conv:(479.31)
2. MOCHILA=SI 472 ==> Zapato=NO 472 <conf:(1)> lift:(1.64) lev:(0.01) [184] conv:(184.68)
3. BILLETERA=SI 414 ==> Zapato=NO 414 <conf:(1)> lift:(1.64) lev:(0) [161] conv:(161.99)
4. Sandalia=SI 404 ==> Zapato=NO 403 <conf:(1)> lift:(1.64) lev:(0) [157] conv:(79.04)
5. GORRA=SI 2763 ==> Zapato=NO 2743 <conf:(0.99)> lift:(1.63) lev:(0.03) [1061] conv:(51.48)
6. BOLSO=SI 369 ==> Zapato=NO 366 <conf:(0.99)> lift:(1.63) lev:(0) [141] conv:(36.09)
7. MEDIAS=SI 1497 ==> Zapato=NO 1474 <conf:(0.98)> lift:(1.62) lev:(0.02) [562] conv:(24.41)
8. Conjunto=SI 377 ==> Zapato=NO 368 <conf:(0.98)> lift:(1.6) lev:(0) [138] conv:(14.75)
9. Calentador=SI 813 ==> Zapato=NO 789 <conf:(0.97)> lift:(1.59) lev:(0.01) [294] conv:(12.72)
10. Pantalón=SI 2065 ==> Zapato=NO 1988 <conf:(0.96)> lift:(1.58) lev:(0.02) [730] conv:(10.36)
11. Pantalón=SI T-Shirt=SI 386 ==> Zapato=NO 368 <conf:(0.95)> lift:(1.57) lev:(0) [133] conv:(7.95)
12. Polo=SI 579 ==> Zapato=NO 550 <conf:(0.95)> lift:(1.56) lev:(0.01) [197] conv:(7.55)
13. T-Shirt=SI 4905 ==> Zapato=NO 4632 <conf:(0.94)> lift:(1.55) lev:(0.05) [1646] conv:(7)
14. Chompa=SI 2531 ==> Zapato=NO 2389 <conf:(0.94)> lift:(1.55) lev:(0.03) [848] conv:(6.93)
15. Bvd=SI 602 ==> Zapato=NO 568 <conf:(0.94)> lift:(1.55) lev:(0.01) [201] conv:(6.73)
16. Buso=SI 386 ==> Zapato=NO 363 <conf:(0.94)> lift:(1.54) lev:(0) [128] conv:(6.29)
17. Short=SI 1680 ==> Zapato=NO 1571 <conf:(0.94)> lift:(1.54) lev:(0.02) [548] conv:(5.98)
18. Zapato=NO 20463 ==> T-Shirt=SI 4632 <conf:(0.23)> lift:(1.55) lev:(0.05) [1646] conv:(1.71)
19. Short=SI 1680 ==> T-Shirt=SI 368 <conf:(0.22)> lift:(1.5) lev:(0) [122] conv:(1.09)
20. Pantalón=SI 2065 ==> T-Shirt=SI 386 <conf:(0.19)> lift:(1.28) lev:(0) [84] conv:(1.05)
21. Pantalón=SI Zapato=NO 1988 ==> T-Shirt=SI 368 <conf:(0.19)> lift:(1.27) lev:(0) [77] conv:(1.05)
22. Pantalón=SI 2065 ==> T-Shirt=SI Zapato=NO 368 <conf:(0.18)> lift:(1.29) lev:(0) [83] conv:(1.05)
23. Pantalón=SI 2065 ==> Chompa=SI 341 <conf:(0.17)> lift:(2.19) lev:(0.01) [185] conv:(1.11)
24. Chompa=SI 2531 ==> Pantalón=SI 341 <conf:(0.13)> lift:(2.19) lev:(0.01) [185] conv:(1.08)
25. Zapato=NO 20463 ==> GORRA=SI 2743 <conf:(0.13)> lift:(1.63) lev:(0.03) [1061] conv:(1.06)
26. Zapato=NO 20463 ==> Chompa=SI 2389 <conf:(0.12)> lift:(1.55) lev:(0.03) [848] conv:(1.05)
27. Zapato=NO 20463 ==> Pantalón=SI 1988 <conf:(0.1)> lift:(1.58) lev:(0.02) [730] conv:(1.04)
28. T-Shirt=SI Zapato=NO 4632 ==> Pantalón=SI 368 <conf:(0.08)> lift:(1.29) lev:(0) [83] conv:(1.02)
29. T-Shirt=SI 4905 ==> Pantalón=SI 386 <conf:(0.08)> lift:(1.28) lev:(0) [84] conv:(1.02)
30. Zapato=NO 20463 ==> Short=SI 1571 <conf:(0.08)> lift:(1.54) lev:(0.02) [548] conv:(1.03)
31. T-Shirt=SI 4905 ==> Short=SI 368 <conf:(0.08)> lift:(1.5) lev:(0) [122] conv:(1.03)
32. T-Shirt=SI 4905 ==> Pantalón=SI Zapato=NO 368 <conf:(0.08)> lift:(1.27) lev:(0) [77] conv:(1.02)
33. Zapato=NO 20463 ==> MEDIAS=SI 1474 <conf:(0.07)> lift:(1.62) lev:(0.02) [562] conv:(1.03)
34. Zapato=NO 20463 ==> PELOTA=SI 1225 <conf:(0.06)> lift:(1.64) lev:(0.01) [479] conv:(1.02)
35. Zapato=NO 20463 ==> Calentador=SI 789 <conf:(0.04)> lift:(1.59) lev:(0.01) [294] conv:(1.01)
36. Zapato=NO 20463 ==> Bvd=SI 568 <conf:(0.03)> lift:(1.55) lev:(0.01) [201] conv:(1.01)
37. Zapato=NO 20463 ==> Polo=SI 550 <conf:(0.03)> lift:(1.56) lev:(0.01) [197] conv:(1.01)
38. Zapato=NO 20463 ==> MOCHILA=SI 472 <conf:(0.02)> lift:(1.64) lev:(0.01) [184] conv:(1.01)
39. Zapato=NO 20463 ==> BILLETERA=SI 414 <conf:(0.02)> lift:(1.64) lev:(0) [161] conv:(1.01)
40. Zapato=NO 20463 ==> Sandalia=SI 403 <conf:(0.02)> lift:(1.64) lev:(0) [157] conv:(1.01)
41. Zapato=NO 20463 ==> Conjunto=SI 368 <conf:(0.02)> lift:(1.6) lev:(0) [138] conv:(1.01)
42. Zapato=NO 20463 ==> Pantalón=SI T-Shirt=SI 368 <conf:(0.02)> lift:(1.57) lev:(0) [133] conv:(1.01)
43. Zapato=NO 20463 ==> BOLSO=SI 366 <conf:(0.02)> lift:(1.63) lev:(0) [141] conv:(1.01)
44. Zapato=NO 20463 ==> Buso=SI 363 <conf:(0.02)> lift:(1.54) lev:(0) [128] conv:(1.01)

```

Figura 47 Reglas de Asociación Condado Shopping

- **Mall del Río (Cuenca)**

- En esta tienda existe un comportamiento de que el 91% de las personas que no adquieren un zapato, compran una chompa y un pantalón, mientras que solo el 18% de los que compran un pantalón compran una chompa, y de los que compran una chompa existe un 15% de probabilidad que lleven un pantalón.
- Cuando una persona no compra un zapato existe un 89% de probabilidad que lleve un conjunto de short y t-shirt, mientras que si solo lleva el short existe un 29% de posibilidad que lleve la t-shirt, y cuando lleven una t-shirt existe el 8% de que lleven un short.
- Cuando una persona compra un pantalón existe un 18% que lleven una t-shirt, mientras que existe apenas un 8% de las que compran una t-shirt compran un pantalón.
- Esta tienda presenta un comportamiento de compra preferente a la combinación de chompa y pantalón, es así que el 95% de las personas que no adquieren un zapato, realizan la combinación mencionada, superando este porcentaje a la combinación de pantalón - t-shirt; y short - t-shirt.
- El 2015 nos permite encontrar que en esta tienda existe un 17% de personas que compran una chompa llevan también un pantalón y el 13% que lleva un pantalón complementa la adquisición con una chompa.
- La prenda con menor combinación y rotación en esta tienda es el bvd, es así que el 100% de las personas que no llevan un bvd llevan en su compra un pantalón y t-shirt, y un 97% una chompa y un pantalón.
- La prenda con mayores combinaciones son las t-shirts que son combinadas en un 21% con los shorts, y 31% con los pantalones.
- Otra prenda que presenta un relevante número de combinaciones en el 2015 y 2016 son los pantalones, ya que realizan una asociación con la compra de chompas en un 14% en el 2015 y de 19% en el 2016; mientras que también lo complementan en la compra de t-shirts; en el 2015 del 9% y en el 2016 del 7%.

```

16. Chompa=SI Pantalon=SI 313 ==> Zapato=NO 286 <conf:(0.91)> lift:(1.48) lev:(0) [92] conv:(4.26)
17. Short=SI T-Shirt=SI 307 ==> Zapato=NO 274 <conf:(0.89)> lift:(1.44) lev:(0) [84] conv:(3.44)
18. Pantalon=SI T-Shirt=SI 327 ==> Zapato=NO 288 <conf:(0.88)> lift:(1.42) lev:(0) [85] conv:(3.12)
19. Zapato=NO 15809 ==> T-Shirt=SI 3664 <conf:(0.23)> lift:(1.51) lev:(0.05) [1234] conv:(1.1)
20. Short=SI 1384 ==> T-Shirt=SI 307 <conf:(0.22)> lift:(1.44) lev:(0) [94] conv:(1.09)
21. Short=SI Zapato=NO 1274 ==> T-Shirt=SI 274 <conf:(0.22)> lift:(1.4) lev:(0) [78] conv:(1.08)
22. Short=SI 1384 ==> T-Shirt=SI Zapato=NO 274 <conf:(0.2)> lift:(1.38) lev:(0) [75] conv:(1.07)
23. Pantalon=SI 1781 ==> T-Shirt=SI 327 <conf:(0.18)> lift:(1.19) lev:(0) [53] conv:(1.04)
24. Pantalon=SI 1781 ==> Chompa=SI 313 <conf:(0.18)> lift:(2.21) lev:(0.01) [171] conv:(1.12)
25. Pantalon=SI Zapato=NO 1674 ==> T-Shirt=SI 288 <conf:(0.17)> lift:(1.12) lev:(0) [30] conv:(1.02)
26. Pantalon=SI Zapato=NO 1674 ==> Chompa=SI 286 <conf:(0.17)> lift:(2.15) lev:(0.01) [152] conv:(1.11)
27. Pantalon=SI 1781 ==> T-Shirt=SI Zapato=NO 288 <conf:(0.16)> lift:(1.13) lev:(0) [32] conv:(1.02)
28. Pantalon=SI 1781 ==> Chompa=SI Zapato=NO 286 <conf:(0.16)> lift:(2.2) lev:(0.01) [155] conv:(1.1)
29. Chompa=SI 2034 ==> Pantalon=SI 313 <conf:(0.15)> lift:(2.21) lev:(0.01) [171] conv:(1.1)
30. Chompa=SI Zapato=NO 1867 ==> Pantalon=SI 286 <conf:(0.15)> lift:(2.2) lev:(0.01) [155] conv:(1.1)
31. Chompa=SI 2034 ==> Pantalon=SI Zapato=NO 286 <conf:(0.14)> lift:(2.15) lev:(0.01) [152] conv:(1.09)
32. Zapato=NO 15809 ==> Chompa=SI 1867 <conf:(0.12)> lift:(1.48) lev:(0.02) [608] conv:(1.04)
33. Zapato=NO 15809 ==> Pantalon=SI 1674 <conf:(0.11)> lift:(1.52) lev:(0.02) [571] conv:(1.04)
34. Zapato=NO 15809 ==> GORRA=SI 1532 <conf:(0.1)> lift:(1.61) lev:(0.02) [578] conv:(1.04)
35. T-Shirt=SI 3926 ==> Pantalon=SI 327 <conf:(0.08)> lift:(1.19) lev:(0) [53] conv:(1.01)
36. Zapato=NO 15809 ==> MEDIAS=SI 1278 <conf:(0.08)> lift:(1.58) lev:(0.02) [466] conv:(1.03)
37. Zapato=NO 15809 ==> Short=SI 1274 <conf:(0.08)> lift:(1.49) lev:(0.02) [417] conv:(1.03)
38. T-Shirt=SI Zapato=NO 3664 ==> Pantalon=SI 288 <conf:(0.08)> lift:(1.13) lev:(0) [32] conv:(1.01)
39. T-Shirt=SI 3926 ==> Short=SI 307 <conf:(0.08)> lift:(1.44) lev:(0) [94] conv:(1.03)
40. T-Shirt=SI Zapato=NO 3664 ==> Short=SI 274 <conf:(0.07)> lift:(1.38) lev:(0) [75] conv:(1.02)
41. Zapato=NO 15809 ==> PELOTA=SI 1179 <conf:(0.07)> lift:(1.62) lev:(0.02) [449] conv:(1.03)

42. T-Shirt=SI 3926 ==> Pantalon=SI Zapato=NO 288 <conf:(0.07)> lift:(1.12) lev:(0) [30] conv:(1.01)
43. T-Shirt=SI 3926 ==> Short=SI Zapato=NO 274 <conf:(0.07)> lift:(1.4) lev:(0) [78] conv:(1.02)
44. Zapato=NO 15809 ==> Calentador=SI 831 <conf:(0.05)> lift:(1.55) lev:(0.01) [293] conv:(1.02)
45. Zapato=NO 15809 ==> Bvd=SI 501 <conf:(0.03)> lift:(1.48) lev:(0.01) [162] conv:(1.01)
46. Zapato=NO 15809 ==> Sandalia=SI 442 <conf:(0.03)> lift:(1.61) lev:(0.01) [167] conv:(1.01)
47. Zapato=NO 15809 ==> Polo=SI 379 <conf:(0.02)> lift:(1.53) lev:(0.01) [131] conv:(1.01)
48. Zapato=NO 15809 ==> Buso=SI 359 <conf:(0.02)> lift:(1.53) lev:(0) [124] conv:(1.01)
49. Zapato=NO 15809 ==> Agua=SI 339 <conf:(0.02)> lift:(1.62) lev:(0.01) [129] conv:(1.01)
50. Zapato=NO 15809 ==> Conjunto=SI 306 <conf:(0.02)> lift:(1.56) lev:(0) [109] conv:(1.01)
51. Zapato=NO 15809 ==> Pantalon=SI T-Shirt=SI 288 <conf:(0.02)> lift:(1.42) lev:(0) [85] conv:(1.01)
52. Zapato=NO 15809 ==> Chompa=SI Pantalon=SI 286 <conf:(0.02)> lift:(1.48) lev:(0) [92] conv:(1.01)
53. Zapato=NO 15809 ==> MOCHILA=SI 277 <conf:(0.02)> lift:(1.61) lev:(0) [104] conv:(1.01)
54. Zapato=NO 15809 ==> Short=SI T-Shirt=SI 274 <conf:(0.02)> lift:(1.44) lev:(0) [84] conv:(1.01)

```

Figura 48 Reglas de Asociación Mall del Río

- **Mall del Jardín (Quito)**

- En esta tienda las personas tienen un comportamiento de que el 89% de las personas que no compran zapatos, compran un pantalón y t-shirt, mientras que el 19% que compra un pantalón compra una t-shirt, y el 8% que compra t-shirt compra un pantalón.
- El 85% de las personas que no compran un zapato compran short y t-shirt, el 24% que compra un short compra t-shirt y el 9% que compra t-shirt compra un short.

- El 15% de las personas que compran una chompa compra un pantalón.
- Uno de los artículos menos vendidos y el que queda fuera de las asociaciones en esta tienda son las pelotas, es decir es un ítem que se vende sin asociación. En esta tienda las personas que compran un t-shirt lo complementan en un 10% con pantalones, 7% con shorts y 6% con zapatos, es decir es el ítem con mayor preferencia de venta y de combinación es así que el 23% que compra shorts, el 21% que compra pantalones y el 3% que compra zapatos complementa su compra con t-shirts. complementar a otros artículos.
- Al analizar las asociaciones de esta tienda encontramos que existe una preferencia de los pantalones sobre los shorts, es así que el 93% de las personas que no compran el short combinan el pantalón y t-shirt, un 21%, y se refleja que en una compra sola el 8% que no compra short compra pantalones.
- Al analizar la combinación de una de las prendas más vendidas como son las t-shirt encontramos que las personas complementan esta prenda en un 5% con los zapatos.

```

1. PELOTA=SI 1033 ==> Zapato=NO 1033 <conf:(1)> lift:(1.53) lev:(0.01) [359] conv:(359.27)
2. Sandalia=SI 587 ==> Zapato=NO 587 <conf:(1)> lift:(1.53) lev:(0.01) [204] conv:(204.15)
3. MOCHILA=SI 501 ==> Zapato=NO 500 <conf:(1)> lift:(1.53) lev:(0) [173] conv:(87.12)
4. GORRA=SI 2986 ==> Zapato=NO 2972 <conf:(1)> lift:(1.53) lev:(0.03) [1024] conv:(69.23)
5. BOLSO=SI 509 ==> Zapato=NO 504 <conf:(0.99)> lift:(1.52) lev:(0) [172] conv:(29.5)
6. Agua=SI 489 ==> Zapato=NO 484 <conf:(0.99)> lift:(1.52) lev:(0) [165] conv:(28.34)
7. MEDIAS=SI 2111 ==> Zapato=NO 2075 <conf:(0.98)> lift:(1.51) lev:(0.02) [698] conv:(19.84)
8. Calentador=SI 1135 ==> Zapato=NO 1093 <conf:(0.96)> lift:(1.48) lev:(0.01) [352] conv:(9.18)
9. Conjunto=SI 407 ==> Zapato=NO 389 <conf:(0.96)> lift:(1.47) lev:(0) [123] conv:(7.45)
10. Polo=SI 733 ==> Zapato=NO 691 <conf:(0.94)> lift:(1.45) lev:(0.01) [212] conv:(5.93)
11. Pantalón=SI 2861 ==> Zapato=NO 2688 <conf:(0.94)> lift:(1.44) lev:(0.02) [822] conv:(5.72)
12. T-Shirt=SI 6683 ==> Zapato=NO 6198 <conf:(0.93)> lift:(1.42) lev:(0.05) [1839] conv:(4.78)
13. Chompa=SI 2765 ==> Zapato=NO 2543 <conf:(0.92)> lift:(1.41) lev:(0.02) [739] conv:(4.31)
14. Short=SI 2546 ==> Zapato=NO 2340 <conf:(0.92)> lift:(1.41) lev:(0.02) [679] conv:(4.28)
15. Busto=SI 516 ==> Zapato=NO 472 <conf:(0.91)> lift:(1.4) lev:(0) [135] conv:(3.99)
16. Bvd=SI 822 ==> Zapato=NO 741 <conf:(0.9)> lift:(1.38) lev:(0.01) [204] conv:(3.49)
17. Pantalón=SI T-Shirt=SI 555 ==> Zapato=NO 492 <conf:(0.89)> lift:(1.36) lev:(0) [130] conv:(3.02)
18. Short=SI T-Shirt=SI 616 ==> Zapato=NO 521 <conf:(0.85)> lift:(1.3) lev:(0) [119] conv:(2.23)
19. Zapato=NO 25146 ==> T-Shirt=SI 6198 <conf:(0.25)> lift:(1.42) lev:(0.05) [1839] conv:(1.1)
20. Short=SI 2546 ==> T-Shirt=SI 616 <conf:(0.24)> lift:(1.4) lev:(0) [174] conv:(1.09)
21. Short=SI Zapato=NO 2340 ==> T-Shirt=SI 521 <conf:(0.22)> lift:(1.28) lev:(0) [115] conv:(1.06)
22. Short=SI 2546 ==> T-Shirt=SI Zapato=NO 521 <conf:(0.2)> lift:(1.27) lev:(0) [111] conv:(1.05)
23. Pantalón=SI 2861 ==> T-Shirt=SI 555 <conf:(0.19)> lift:(1.12) lev:(0) [59] conv:(1.03)
24. Pantalón=SI Zapato=NO 2688 ==> T-Shirt=SI 492 <conf:(0.18)> lift:(1.06) lev:(0) [26] conv:(1.01)
25. Pantalón=SI 2861 ==> T-Shirt=SI Zapato=NO 492 <conf:(0.17)> lift:(1.07) lev:(0) [32] conv:(1.01)
26. Chompa=SI 2765 ==> Pantalón=SI 424 <conf:(0.15)> lift:(2.07) lev:(0.01) [218] conv:(1.09)
27. Pantalón=SI 2861 ==> Chompa=SI 424 <conf:(0.15)> lift:(2.07) lev:(0.01) [218] conv:(1.09)
28. Zapato=NO 25146 ==> GORRA=SI 2972 <conf:(0.12)> lift:(1.53) lev:(0.03) [1024] conv:(1.05)
29. Zapato=NO 25146 ==> Pantalón=SI 2688 <conf:(0.11)> lift:(1.44) lev:(0.02) [822] conv:(1.04)
30. Zapato=NO 25146 ==> Chompa=SI 2543 <conf:(0.1)> lift:(1.41) lev:(0.02) [739] conv:(1.03)
31. Zapato=NO 25146 ==> Short=SI 2340 <conf:(0.09)> lift:(1.41) lev:(0.02) [679] conv:(1.03)
32. T-Shirt=SI 6683 ==> Short=SI 616 <conf:(0.09)> lift:(1.4) lev:(0) [174] conv:(1.03)
33. T-Shirt=SI Zapato=NO 6198 ==> Short=SI 521 <conf:(0.08)> lift:(1.27) lev:(0) [111] conv:(1.02)
34. T-Shirt=SI 6683 ==> Pantalón=SI 555 <conf:(0.08)> lift:(1.12) lev:(0) [59] conv:(1.01)
35. Zapato=NO 25146 ==> MEDIAS=SI 2075 <conf:(0.08)> lift:(1.51) lev:(0.02) [698] conv:(1.03)
36. T-Shirt=SI Zapato=NO 6198 ==> Pantalón=SI 492 <conf:(0.08)> lift:(1.07) lev:(0) [32] conv:(1.01)
37. T-Shirt=SI 6683 ==> Short=SI Zapato=NO 521 <conf:(0.08)> lift:(1.28) lev:(0) [115] conv:(1.02)
38. T-Shirt=SI 6683 ==> Pantalón=SI Zapato=NO 492 <conf:(0.07)> lift:(1.06) lev:(0) [26] conv:(1)
39. Zapato=NO 25146 ==> Calentador=SI 1093 <conf:(0.04)> lift:(1.48) lev:(0.01) [352] conv:(1.01)
40. Zapato=NO 25146 ==> PELOTA=SI 1033 <conf:(0.04)> lift:(1.53) lev:(0.01) [359] conv:(1.01)
41. Zapato=NO 25146 ==> Bvd=SI 741 <conf:(0.03)> lift:(1.38) lev:(0.01) [204] conv:(1.01)
42. Zapato=NO 25146 ==> Polo=SI 691 <conf:(0.03)> lift:(1.45) lev:(0.01) [212] conv:(1.01)
43. Zapato=NO 25146 ==> Sandalia=SI 587 <conf:(0.02)> lift:(1.53) lev:(0.01) [204] conv:(1.01)
44. Zapato=NO 25146 ==> Short=SI T-Shirt=SI 521 <conf:(0.02)> lift:(1.3) lev:(0) [119] conv:(1)

```

Figura 49 Reglas de Asociación Mall el Jardín

- **Scala Shopping (Quito)**

- El comportamiento en esta tienda presenta que el 92% de las personas que no compran zapatos, se deciden por pantalón y t-shirt en el 2013 y en el 2014 este porcentaje crece a 94% el mismo que se mantiene en 2015 y 2016, mientras que el mismo comportamiento de no comprar zapatos, y escoger una combinación de short y t-shirt es de 87% en el 2013 y en el 2014 cambia a 64%, es decir la preferencia aumento hacia los pantalones.
- En las preferencias de esta tienda tenemos que las personas que compran shorts un 22% lo complementa con t-shirt y de las que compran t-shirt un 8% lo complementan con short.

- Las asociaciones también nos muestran que las personas que compran pantalones un 18% lo complementa con t-shirt y de las que compran t-shirt un 8% lo complementan con pantalones.

1.	PELOTA=SI 853 ==> Zapato=NO 853	<conf:(1)> lift:(1.44) lev:(0.01) [261] conv:(261.27)
2.	MOCHILA=SI 224 ==> Zapato=NO 224	<conf:(1)> lift:(1.44) lev:(0) [68] conv:(68.61)
3.	Agua=SI 198 ==> Zapato=NO 197	<conf:(0.99)> lift:(1.43) lev:(0) [59] conv:(30.32)
4.	Sandalia=SI 180 ==> Zapato=NO 179	<conf:(0.99)> lift:(1.43) lev:(0) [54] conv:(27.57)
5.	GORRA=SI 1217 ==> Zapato=NO 1209	<conf:(0.99)> lift:(1.43) lev:(0.02) [364] conv:(41.42)
6.	Calentador=SI 492 ==> Zapato=NO 487	<conf:(0.99)> lift:(1.43) lev:(0.01) [145] conv:(14.5)
7.	MEDIAS=SI 1051 ==> Zapato=NO 1032	<conf:(0.98)> lift:(1.42) lev:(0.02) [302] conv:(16.1)
8.	Polo=SI 290 ==> Zapato=NO 282	<conf:(0.97)> lift:(1.4) lev:(0) [80] conv:(9.87)
9.	Pantalon=SI 1392 ==> Zapato=NO 1326	<conf:(0.95)> lift:(1.37) lev:(0.02) [360] conv:(6.36)
10.	T-Shirt=SI 3246 ==> Zapato=NO 3076	<conf:(0.95)> lift:(1.37) lev:(0.05) [824] conv:(5.81)
11.	Buso=SI 225 ==> Zapato=NO 212	<conf:(0.94)> lift:(1.36) lev:(0) [55] conv:(4.92)
12.	Bvd=SI 418 ==> Zapato=NO 393	<conf:(0.94)> lift:(1.36) lev:(0.01) [103] conv:(4.92)
13.	Chompa=SI 1192 ==> Zapato=NO 1119	<conf:(0.94)> lift:(1.35) lev:(0.02) [292] conv:(4.93)
14.	Short=SI 1412 ==> Zapato=NO 1316	<conf:(0.93)> lift:(1.34) lev:(0.02) [336] conv:(4.46)
15.	Pantalon=SI T-Shirt=SI 260 ==> Zapato=NO 239	<conf:(0.92)> lift:(1.33) lev:(0) [58] conv:(3.62)
16.	Short=SI T-Shirt=SI 316 ==> Zapato=NO 276	<conf:(0.87)> lift:(1.26) lev:(0) [56] conv:(2.36)
17.	Zapato=NO 12103 ==> T-Shirt=SI 3076	<conf:(0.25)> lift:(1.37) lev:(0.05) [824] conv:(1.09)
18.	Short=SI 1412 ==> T-Shirt=SI 316	<conf:(0.22)> lift:(1.2) lev:(0) [53] conv:(1.05)
19.	Short=SI Zapato=NO 1316 ==> T-Shirt=SI 276	<conf:(0.21)> lift:(1.13) lev:(0) [31] conv:(1.03)
20.	Short=SI 1412 ==> T-Shirt=SI Zapato=NO 276	<conf:(0.2)> lift:(1.11) lev:(0) [27] conv:(1.02)
21.	Pantalon=SI 1392 ==> T-Shirt=SI 260	<conf:(0.19)> lift:(1) lev:(0) [1] conv:(1)
22.	Pantalon=SI Zapato=NO 1326 ==> T-Shirt=SI 239	<conf:(0.18)> lift:(0.97) lev:(-0) [-7] conv:(0.99)
23.	Pantalon=SI 1392 ==> T-Shirt=SI Zapato=NO 239	<conf:(0.17)> lift:(0.97) lev:(-0) [-6] conv:(0.99)
24.	Zapato=NO 12103 ==> Pantalon=SI 1326	<conf:(0.11)> lift:(1.37) lev:(0.02) [360] conv:(1.03)
25.	Zapato=NO 12103 ==> Short=SI 1316	<conf:(0.11)> lift:(1.34) lev:(0.02) [336] conv:(1.03)
26.	Zapato=NO 12103 ==> GORRA=SI 1209	<conf:(0.1)> lift:(1.43) lev:(0.02) [364] conv:(1.03)
27.	T-Shirt=SI 3246 ==> Short=SI 316	<conf:(0.1)> lift:(1.2) lev:(0) [53] conv:(1.02)
28.	Zapato=NO 12103 ==> Chompa=SI 1119	<conf:(0.09)> lift:(1.35) lev:(0.02) [292] conv:(1.03)
29.	T-Shirt=SI Zapato=NO 3076 ==> Short=SI 276	<conf:(0.09)> lift:(1.11) lev:(0) [27] conv:(1.01)
30.	Zapato=NO 12103 ==> MEDIAS=SI 1032	<conf:(0.09)> lift:(1.42) lev:(0.02) [302] conv:(1.03)
31.	T-Shirt=SI 3246 ==> Short=SI Zapato=NO 276	<conf:(0.09)> lift:(1.13) lev:(0) [31] conv:(1.01)
32.	T-Shirt=SI 3246 ==> Pantalon=SI 260	<conf:(0.08)> lift:(1) lev:(0) [1] conv:(1)
33.	T-Shirt=SI Zapato=NO 3076 ==> Pantalon=SI 239	<conf:(0.08)> lift:(0.97) lev:(-0) [-6] conv:(1)
34.	T-Shirt=SI 3246 ==> Pantalon=SI Zapato=NO 239	<conf:(0.07)> lift:(0.97) lev:(-0) [-7] conv:(1)
35.	Zapato=NO 12103 ==> PELOTA=SI 853	<conf:(0.07)> lift:(1.44) lev:(0.01) [261] conv:(1.02)
36.	Zapato=NO 12103 ==> Calentador=SI 487	<conf:(0.04)> lift:(1.43) lev:(0.01) [145] conv:(1.01)
37.	Zapato=NO 12103 ==> Bvd=SI 393	<conf:(0.03)> lift:(1.36) lev:(0.01) [103] conv:(1.01)
38.	Zapato=NO 12103 ==> Polo=SI 282	<conf:(0.02)> lift:(1.4) lev:(0) [80] conv:(1.01)
39.	Zapato=NO 12103 ==> Short=SI T-Shirt=SI 276	<conf:(0.02)> lift:(1.26) lev:(0) [56] conv:(1)
40.	Zapato=NO 12103 ==> Pantalon=SI T-Shirt=SI 239	<conf:(0.02)> lift:(1.33) lev:(0) [58] conv:(1)
41.	Zapato=NO 12103 ==> MOCHILA=SI 224	<conf:(0.02)> lift:(1.44) lev:(0) [68] conv:(1.01)
42.	Zapato=NO 12103 ==> Buso=SI 212	<conf:(0.02)> lift:(1.36) lev:(0) [55] conv:(1)
43.	Zapato=NO 12103 ==> Agua=SI 197	<conf:(0.02)> lift:(1.43) lev:(0) [59] conv:(1)
44.	Zapato=NO 12103 ==> Sandalia=SI 179	<conf:(0.01)> lift:(1.43) lev:(0) [54] conv:(1)

Figura 50 Reglas de Asociación Scala Shopping

- **Mall del Sol (Guayaquil)**

- El 91% de las personas que llevan un pantalón y t-shirt no llevan zapatos, mientras que el 21% de los que llevan un pantalón también complementan su compra con t-shirt; y los que compran t-shirt el 6% lleva también un pantalón.

- El 87% de las personas que compran un short y t-shirt no llevan zapatos, y el 22% que lleva un short lleva una camiseta también, mientras que solo el 10% que lleva una camiseta también lleva un short.
- Esta tienda presenta una asociación relevante del 97% al relacionar que los clientes que no compran bvd, realizan un complemento entre t-shirt y zapatos, un 6% y en la compra de zapatos un 3% lo complementan con t-shirt.
- Uno de los ítems con menores ventas relacionales son los bvd, ya que los descartan en conjuntos como t-shirt y zapato, short y t-shirt, pantalón y t-shirt, demostrando una preferencia por las t-shirt frente a los bvd.

```

1. PELOTA=SI 970 ==> Zapato=NO 970 <conf:(1)> lift:(1.53) lev:(0.01) [336] conv:(336.38)
2. MOCHILA=SI 463 ==> Zapato=NO 462 <conf:(1)> lift:(1.53) lev:(0) [159] conv:(80.28)
3. GORRA=SI 3721 ==> Zapato=NO 3709 <conf:(1)> lift:(1.53) lev:(0.03) [1278] conv:(99.26)
4. Sandalia=SI 1131 ==> Zapato=NO 1124 <conf:(0.99)> lift:(1.52) lev:(0.01) [385] conv:(49.03)
5. Agua=SI 479 ==> Zapato=NO 476 <conf:(0.99)> lift:(1.52) lev:(0) [163] conv:(41.53)
6. MEDIAS=SI 3074 ==> Zapato=NO 3027 <conf:(0.98)> lift:(1.51) lev:(0.02) [1019] conv:(22.21)
7. Bvd=SI 1526 ==> Zapato=NO 1443 <conf:(0.95)> lift:(1.45) lev:(0.01) [446] conv:(6.3)
8. Pantalón=SI 2369 ==> Zapato=NO 2237 <conf:(0.94)> lift:(1.45) lev:(0.02) [689] conv:(6.18)
9. Polo=SI 1320 ==> Zapato=NO 1244 <conf:(0.94)> lift:(1.44) lev:(0.01) [381] conv:(5.94)
10. Chompa=SI 1877 ==> Zapato=NO 1766 <conf:(0.94)> lift:(1.44) lev:(0.01) [539] conv:(5.81)
11. T-Shirt=SI 8373 ==> Zapato=NO 7861 <conf:(0.94)> lift:(1.44) lev:(0.05) [2391] conv:(5.66)
12. Short=SI 3674 ==> Zapato=NO 3426 <conf:(0.93)> lift:(1.43) lev:(0.02) [1026] conv:(5.12)
13. Pantalón=SI T-Shirt=SI 506 ==> Zapato=NO 459 <conf:(0.91)> lift:(1.39) lev:(0) [128] conv:(3.66)
14. Short=SI T-Shirt=SI 819 ==> Zapato=NO 713 <conf:(0.87)> lift:(1.33) lev:(0) [178] conv:(2.65)
15. Zapato=NO 29080 ==> T-Shirt=SI 7861 <conf:(0.27)> lift:(1.44) lev:(0.05) [2391] conv:(1.11)
16. Short=SI 3674 ==> T-Shirt=SI 819 <conf:(0.22)> lift:(1.19) lev:(0) [127] conv:(1.04)
17. Pantalón=SI 2369 ==> T-Shirt=SI 506 <conf:(0.21)> lift:(1.14) lev:(0) [60] conv:(1.03)
18. Short=SI Zapato=NO 3426 ==> T-Shirt=SI 713 <conf:(0.21)> lift:(1.11) lev:(0) [68] conv:(1.02)
19. Pantalón=SI Zapato=NO 2237 ==> T-Shirt=SI 459 <conf:(0.21)> lift:(1.09) lev:(0) [38] conv:(1.02)
20. Short=SI 3674 ==> T-Shirt=SI Zapato=NO 713 <conf:(0.19)> lift:(1.1) lev:(0) [64] conv:(1.02)
21. Pantalón=SI 2369 ==> T-Shirt=SI Zapato=NO 459 <conf:(0.19)> lift:(1.1) lev:(0) [40] conv:(1.02)
22. Zapato=NO 29080 ==> GORRA=SI 3709 <conf:(0.13)> lift:(1.53) lev:(0.03) [1278] conv:(1.05)
23. Zapato=NO 29080 ==> Short=SI 3426 <conf:(0.12)> lift:(1.43) lev:(0.02) [1026] conv:(1.04)
24. Zapato=NO 29080 ==> MEDIAS=SI 3027 <conf:(0.1)> lift:(1.51) lev:(0.02) [1019] conv:(1.04)
25. T-Shirt=SI 8373 ==> Short=SI 819 <conf:(0.1)> lift:(1.19) lev:(0) [127] conv:(1.02)
26. T-Shirt=SI Zapato=NO 7861 ==> Short=SI 713 <conf:(0.09)> lift:(1.1) lev:(0) [64] conv:(1.01)
27. T-Shirt=SI 8373 ==> Short=SI Zapato=NO 713 <conf:(0.09)> lift:(1.11) lev:(0) [68] conv:(1.01)
28. Zapato=NO 29080 ==> Pantalón=SI 2237 <conf:(0.08)> lift:(1.45) lev:(0.02) [689] conv:(1.03)
29. Zapato=NO 29080 ==> Chompa=SI 1766 <conf:(0.06)> lift:(1.44) lev:(0.01) [539] conv:(1.02)
30. T-Shirt=SI 8373 ==> Pantalón=SI 506 <conf:(0.06)> lift:(1.14) lev:(0) [60] conv:(1.01)
31. T-Shirt=SI Zapato=NO 7861 ==> Pantalón=SI 459 <conf:(0.06)> lift:(1.1) lev:(0) [40] conv:(1.01)
32. T-Shirt=SI 8373 ==> Pantalón=SI Zapato=NO 459 <conf:(0.05)> lift:(1.09) lev:(0) [38] conv:(1)
33. Zapato=NO 29080 ==> Bvd=SI 1443 <conf:(0.05)> lift:(1.45) lev:(0.01) [446] conv:(1.02)
34. Zapato=NO 29080 ==> Polo=SI 1244 <conf:(0.04)> lift:(1.44) lev:(0.01) [381] conv:(1.01)
35. Zapato=NO 29080 ==> Sandalia=SI 1124 <conf:(0.04)> lift:(1.52) lev:(0.01) [385] conv:(1.01)
36. Zapato=NO 29080 ==> PELOTA=SI 970 <conf:(0.03)> lift:(1.53) lev:(0.01) [336] conv:(1.01)
37. Zapato=NO 29080 ==> Short=SI T-Shirt=SI 713 <conf:(0.02)> lift:(1.33) lev:(0) [178] conv:(1.01)
38. Zapato=NO 29080 ==> Agua=SI 476 <conf:(0.02)> lift:(1.52) lev:(0) [163] conv:(1.01)
39. Zapato=NO 29080 ==> MOCHILA=SI 462 <conf:(0.02)> lift:(1.53) lev:(0) [159] conv:(1.01)
40. Zapato=NO 29080 ==> Pantalón=SI T-Shirt=SI 459 <conf:(0.02)> lift:(1.39) lev:(0) [128] conv:(1)

```

Figura 51 Reglas de Asociación Mall del Sol

- **Village Plaza (Guayaquil)**

- En esta tienda tenemos un comportamiento en la compra de un outfit de dos prendas, con las opciones de que el 97% que compra un short y una t-shirt no adquiere una chompa, mientras que el 96% que combina un pantalón y una t-shirt no complementa con la chompa.
- El 11% de las personas que compran una t-shirt la complementan con un short, y el 7% la complementa con un pantalón; mientras que el 25% que adquieren un short compra también una t-shirt, y el 23% de los que compran un pantalón complementan con una t-shirt.
- En esta tienda hay una relación cuando no existe compra de bvd; en 98% entre t-shirt y zapatos, 97% en shorts y t-shirts y 92% entre pantalón y t-shirt.
- El ítem con mayor relación son las t-shirt en esta tienda, cuando adquieren un short complementan con una t-shirt el 23%, cuando adquieren un pantalón lo complementan con la t-shirt el 22% y en el caso de los zapatos lo complementan con t-shirt el 4%.
- A la compra de t-shirts las complementan con el 9% en shorts, el 6% en zapatos y el 5% en pantalones.
- El comportamiento relacionado a zapatos en esta tienda es cuando no se compran el outfit a elección es short y t-shirt en un 91%, mientras que cuando se compran un short un 23% lo complementa con la t-shirt, y cuando compran t-shirt el 5% lo relaciona con un short.

```

1. MEDIAS=SI 1039 ==> Chompa=NO 1039 <conf:(1)> lift:(1.04) lev:(0) [40] conv:(40.55)
2. PELOTA=SI 769 ==> Chompa=NO 769 <conf:(1)> lift:(1.04) lev:(0) [30] conv:(30.01)
3. MOCHILA=SI 239 ==> Chompa=NO 239 <conf:(1)> lift:(1.04) lev:(0) [9] conv:(9.33)
4. BOLSO=SI 211 ==> Chompa=NO 211 <conf:(1)> lift:(1.04) lev:(0) [8] conv:(8.23)
5. Sandalia=SI 410 ==> Chompa=NO 409 <conf:(1)> lift:(1.04) lev:(0) [14] conv:(8)
6. GORRA=SI 1240 ==> Chompa=NO 1236 <conf:(1)> lift:(1.04) lev:(0) [44] conv:(9.68)
7. Agua=SI 272 ==> Chompa=NO 271 <conf:(1)> lift:(1.04) lev:(0) [9] conv:(5.31)
8. Zapato=SI 4653 ==> Chompa=NO 4624 <conf:(0.99)> lift:(1.03) lev:(0.01) [152] conv:(6.05)
9. T-Shirt=SI 3530 ==> Chompa=NO 3470 <conf:(0.98)> lift:(1.02) lev:(0) [77] conv:(2.26)
10. Bvd=SI 857 ==> Chompa=NO 841 <conf:(0.98)> lift:(1.02) lev:(0) [17] conv:(1.97)
11. Short=SI 1481 ==> Chompa=NO 1451 <conf:(0.98)> lift:(1.02) lev:(0) [27] conv:(1.86)
12. Polo=SI 416 ==> Chompa=NO 406 <conf:(0.98)> lift:(1.02) lev:(0) [6] conv:(1.48)
13. Short=SI T-Shirt=SI 376 ==> Chompa=NO 364 <conf:(0.97)> lift:(1.01) lev:(0) [2] conv:(1.13)
14. Pantalon=SI T-Shirt=SI 232 ==> Chompa=NO 222 <conf:(0.96)> lift:(1) lev:(-0) [0] conv:(0.82)
15. Pantalon=SI 1005 ==> Chompa=NO 961 <conf:(0.96)> lift:(1) lev:(-0) [-4] conv:(0.87)
16. Chompa=NO 15908 ==> Zapato=SI 4624 <conf:(0.29)> lift:(1.03) lev:(0.01) [152] conv:(1.01)
17. Short=SI 1481 ==> T-Shirt=SI 376 <conf:(0.25)> lift:(1.19) lev:(0) [60] conv:(1.05)
18. Chompa=NO Short=SI 1451 ==> T-Shirt=SI 364 <conf:(0.25)> lift:(1.18) lev:(0) [54] conv:(1.05)
19. Short=SI 1481 ==> Chompa=NO T-Shirt=SI 364 <conf:(0.25)> lift:(1.17) lev:(0) [53] conv:(1.05)
20. Chompa=NO Pantalon=SI 961 ==> T-Shirt=SI 222 <conf:(0.23)> lift:(1.08) lev:(0) [17] conv:(1.02)
21. Pantalon=SI 1005 ==> T-Shirt=SI 232 <conf:(0.23)> lift:(1.08) lev:(0) [17] conv:(1.02)
22. Pantalon=SI 1005 ==> Chompa=NO T-Shirt=SI 222 <conf:(0.22)> lift:(1.05) lev:(0) [11] conv:(1.01)
23. Chompa=NO 15908 ==> T-Shirt=SI 3470 <conf:(0.22)> lift:(1.02) lev:(0) [77] conv:(1.01)
24. T-Shirt=SI 3530 ==> Short=SI 376 <conf:(0.11)> lift:(1.19) lev:(0) [60] conv:(1.02)
25. Chompa=NO T-Shirt=SI 3470 ==> Short=SI 364 <conf:(0.1)> lift:(1.17) lev:(0) [53] conv:(1.02)
26. T-Shirt=SI 3530 ==> Chompa=NO Short=SI 364 <conf:(0.1)> lift:(1.18) lev:(0) [54] conv:(1.02)
27. Chompa=NO 15908 ==> Short=SI 1451 <conf:(0.09)> lift:(1.02) lev:(0) [27] conv:(1)
28. Chompa=NO 15908 ==> GORRA=SI 1236 <conf:(0.08)> lift:(1.04) lev:(0) [44] conv:(1)
29. T-Shirt=SI 3530 ==> Pantalon=SI 232 <conf:(0.07)> lift:(1.08) lev:(0) [17] conv:(1.01)
30. Chompa=NO 15908 ==> MEDIAS=SI 1039 <conf:(0.07)> lift:(1.04) lev:(0) [40] conv:(1)
31. Chompa=NO T-Shirt=SI 3470 ==> Pantalon=SI 222 <conf:(0.06)> lift:(1.05) lev:(0) [11] conv:(1)
32. T-Shirt=SI 3530 ==> Chompa=NO Pantalon=SI 222 <conf:(0.06)> lift:(1.08) lev:(0) [17] conv:(1)
33. Chompa=NO 15908 ==> Pantalon=SI 961 <conf:(0.06)> lift:(1) lev:(-0) [-4] conv:(1)
34. Chompa=NO 15908 ==> Bvd=SI 841 <conf:(0.05)> lift:(1.02) lev:(0) [17] conv:(1)
35. Chompa=NO 15908 ==> PELOTA=SI 769 <conf:(0.05)> lift:(1.04) lev:(0) [30] conv:(1)
36. Chompa=NO 15908 ==> Sandalia=SI 409 <conf:(0.03)> lift:(1.04) lev:(0) [14] conv:(1)
37. Chompa=NO 15908 ==> Polo=SI 406 <conf:(0.03)> lift:(1.02) lev:(0) [6] conv:(1)
38. Chompa=NO 15908 ==> Short=SI T-Shirt=SI 364 <conf:(0.02)> lift:(1.01) lev:(0) [2] conv:(1)
39. Chompa=NO 15908 ==> Agua=SI 271 <conf:(0.02)> lift:(1.04) lev:(0) [9] conv:(1)
40. Chompa=NO 15908 ==> MOCHILA=SI 239 <conf:(0.02)> lift:(1.04) lev:(0) [9] conv:(1)
41. Chompa=NO 15908 ==> Pantalon=SI T-Shirt=SI 222 <conf:(0.01)> lift:(1) lev:(-0) [0] conv:(1)
42. Chompa=NO 15908 ==> BOLSO=SI 211 <conf:(0.01)> lift:(1.04) lev:(0) [8] conv:(1)

```

Figura 52 Reglas de Asociación Village Plaza

CAPÍTULO 6

CONCLUSIONES Y RECOMENDACIONES

6.1 CONCLUSIONES

- Los resultados encontrados en la tarea de asociación son importantes en la toma de decisiones para la realización de una compra efectiva, ya que por el giro de negocio y la estructura para realizar el proceso de compra se necesita conocer con precisión los outfits y mix de producto a ingresar, involucrando a su vez puntos estratégicos como visual merchandising en cada tienda. Es así que a nivel de todas las tiendas podemos concluir que la zona de zapatos, es un área caliente en la tienda, es decir un área con niveles de venta altos, pero la venta de un zapato es casi excluyente de la venta de ropa, del 2% al 10% de las personas que compran un zapato buscan asociar su compra con ropa. Por lo cual cuando se realiza la compra de estos ítems, no es fundamental una asociación con los colores, tecnologías y beneficios de las prendas de vestir y zapatos. En el caso de artículos de vestir, tenemos un comportamiento diferente y aquí las asociaciones juegan un papel importante en el cierre de una venta y en el incremento de los volúmenes de compra.
- Encontramos que dentro de los perfiles de consumo independiente de la zona, existe una mayor demanda por la compra de la parte media de la pirámide es decir enhanced, indicando que las personas acuden a una tienda deportiva especializada a diferencia de una multimarca, en busca de productos con tecnología y para deportes específicos, a la vez están dispuestos a pagar un precio más alto por estos artículos por sus beneficios, pero debemos tomar en cuenta que este comportamiento es general, pero al momento de analizar solo la parte de la pirámide top, existe una diferencia de tiendas que tienen mayor demanda de producto de costo más alto como son Mall el Jardín y Mall del Sol, mientras que hacía más producto core de precio bajo encontramos a Village Plaza y Scala Shopping.

- Se encuentran resultados de niveles de venta por grupo, en unidades la venta mayoritaria está en ropa y su mayor presencia se encuentra en la sierra, mientras que en el grupo zapatos encontramos una presencia similar en ambas regiones al igual que accesorios, pero al analizar en montos de facturación encontramos una mayor venta en zapatos, ya que el precio promedio de estos es más alto que ropa y accesorios, esto nos lleva a concluir que en la sierra se arman outfit completos de pantalones, chompas y camisetas en la compra mientras que en la región costa, solo se enfocan en complementar el outfit de una camiseta con un short o un pantalón, por lo cual las profundidades de compra para las tiendas de la sierra debe ser mayor que en la costa.
- Existen actividades donde el enfoque desde marketing debe ser en un solo grupo para potencializarlo y no dividir esfuerzos, es el caso de running, se debe enfocar las campañas en zapatos que el retorno es 4 veces mayor al enfocarse sus ventas de ropa, mientras que viceversa en la actividad de training donde las campañas deben tener como foco ropa ya que su retorno es de 5 a 1 de los zapatos de esta categoría.
- Existe la categoría de fashion, que su participación varía de acuerdo a la tienda y en las mismas la participación que decrece para que esta aumente es categorías de deportes muy específicos como running, outdoors, tennis, se puede concluir que el perfil del consumidor de estas tiendas está encaminado a ropa deportiva pero como forma de vestir o estilo de vestimenta y no para uso puramente deportivo, estas son City Mall, Condado Shopping y Mall el Jardín.
- Encontramos que la venta por meses tiene distintas fluctuaciones pero descartado diciembre, existen meses con mayor demanda, mayo, noviembre, septiembre y agosto la venta es mayor en la región sierra, mientras que en los meses de marzo, abril y junio las ventas mayores se dan en la costa, lo que nos lleva a concluir que los meses de vacación y regreso a clases tiene una gran influencia en las ventas por lo cual el ingreso de colecciones debe ser mayor en estos meses por tienda

- Al analizar el mix de producto y las asociaciones que los clientes realizan encontramos que la base de las asociaciones en todas las tiendas son las t-shirts, y dependiendo de la zona, en sierra se la conjuga con pantalones y chompas, mientras que en la costa es con short y pantalones, por lo cual la diversificación de la oferta al público debe estar enfocado en las camisetas, y la proporción de venta es de 3 a 1, por lo cual se debe comprar 3 t-shirt por un pantalón.
- Existe una diferencia de compra a nivel de talla de acuerdo a la zona de la tienda, esta diferencia se denota mayormente en el grupo zapatos, en hombre existe el mismo porcentaje de ventas o más del tallaje de 8.0 en adelante, mientras que en la sierra las tallas más vendidas son de 9.0 para bajo, lo mismo sucede en mujer con los porcentajes de venta, en el caso de textiles las tallas más vendidas son iguales pero las proporciones son las que se diferencian, por lo que se concluye que se debe realizar una compra de tallaje por tienda y las tallas grandes son vendidas en la región costa.

6.2 RECOMENDACIONES

- Al realizar el análisis de ventas, encontramos que las ventas de 2013 a 2014 crecieron pero de 2014 a 2015 tuvieron un decrecimiento, este comportamiento general se da en todas las tiendas atado a un incremento en los precios lo que ocasionó una menor demanda, pero el año 2016 fue más alto en venta que 2014, se encontró grandes diferencias en las ventas entre las tiendas, por lo cual se recomienda realizar una compra personalizada en cantidad, por lo cual en tema de cantidad de stock tenemos que un cluster de mercado estaría formado por Mall de Sol, Mall el Jardín y Mall del Río; el segundo estaría formado por Scala Shopping y Condado Shopping, y el último por City Mall y Village Plaza.
- Se debe generar un sistema que permita la actualización mensual de la información por venta, para generar un movimiento de mercadería entre las tiendas con la finalidad de aprovechar las oportunidades comerciales de cada artículo, ya que se encontró diferentes comportamientos de venta por productos o grupos.
- La compra de producto para ingreso a cada tienda mensualmente debe estar acorde al perfil de consumo encontrado en este estudio para cada sector, es así que ingresar el mix de producto correcto permitirá aumentar las ventas y el potencial de mercado de cada tienda.
- Cada tienda debe poseer un visual merchandising diferente, este debe ser acorde al mix de producto encontrado para cada una de ellas y basándose en las asociaciones obtenidas en este estudio.
- Al analizar las ventas a nivel de género y grupo, en zapatos a pesar que el stock es mayor en el género de mujer, en Condado Shopping, Mall del Río, Mall el Jardín, Scala Shopping y Village Plaza las ventas son mayores en kids, por lo cual se debe potencializar el producto en este género.

CAPÍTULO 7

BIBLIOGRAFÍA

- (Hernández, J; Ramírez, M; Ferri, C; 2008); Introducción a la minería de datos, Madrid. España
- (Kotler, P; Armstrong, G.; Camara Ibañez, D; Cruz Roche, I; 2004); Marketing, 10ma edición, Madrid España
- (Fischer, L; Espejo, J; 2007); Mercadotecnia, 3era edición. Mexico D.F
- (Schiffman, L; Lazar, L; 1997); Comportamiento del Consumidor. 5ta edición. Mexico
- (Hasty, R; Reardon, J; 1998). Gerencia de Ventas al Detal. Bogota. Colombia
- (Howson, C; 2009); Business Intelligence. Mexico
- (Rosado, A; Rico, Dewar; 2010) Inteligencia de Negocios, 2010. Universidad Tecnológica de Pereira.
- (Riqueleme, J; Ruiz, R; Gilbert K; 2006) Minería de Datos.
- (Valenga, F; Fernández, E; Merlino, H; Rodríguez, D; Procopio, C; Britos, P; García, R); Minería de Datos Aplicada a la Detección de Patrones Delictivos en Argentina Universidad de Morón; Universidad de Buenos Aires; Argentina
- Latinos BI; Soluciones de BI; Bogotá Colombia.
- (Muñoz, A; 2013) Muñoz Álvarez, Raúl Andrés. (Julio 2013). 12. Minería de datos aplicada al manejo de la relación del Cliente en una Entidad Bancaria (CRM). Quito: Escuela Politécnica Nacional.
- (Bustamante, J; Santillán, H; 2009) Bustamante Bustamante, Jorge Andrés; Santillán Haro Karen (Agosto 2015). Desarrollo de un módulo de inteligencia de negocios para el análisis de la entrega del bono de desarrollo humano desde el año 2009. Quito: Escuela Politécnica Nacional.
- (Garrido, L; Latorre, J; 2000) Luis Garrido, José Ignacio Latorre. (2000). Aplicaciones empresariales de Data Mining. Universidad de Barcelona. España

- (Valcarcel, V; 2004) Valcarcel Asencios, Violeta. Data mining y el descubrimiento del Conocimiento. *Ind. data*. [online]. jul./dic. 2004, vol.7, no.2 [citado 29 Noviembre 2015], p.83-86. Disponible en la World Wide Web: revistas.concytec.gob.pe
- (DAEDALUS; 2017) Predicción de ventas (s.f.), Recuperado el 7 de Septiembre del 2017; http://www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/DAEDALUS-MD20-Prediccion_Ventas.pdf
- (DAEDALUS; 2015) Accidentes Cardiovasculares (s.f) Recuperado el 29 de Noviembre del 2015; [www.daedalus.es / fileadmin / daedalus / doc / MineríaDeDatos / DAEDALUS-MD19- Accidentes_Cardiovasculares.pdf](http://www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/DAEDALUS-MD19-Accidentes_Cardiovasculares.pdf)
- (Bustamante, R; 2017) Ejemplos de Data Mining – Tarjetas Crédito (Ricardo Bustamante) Instituto Tecnológico Superior de Lerdo, Recuperado el 7 de Septiembre de 2017; www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/DAEDALUS-MD12-Seguros_Credito.pdf
- (Bustamante, R; 2015) Ejemplos de Data Mining – Automóviles (Ricardo Bustamante) *Instituto Tecnológico Superior de Lerdo*, Recuperado el 29 de Noviembre de 2015 de <http://www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/DAEDALUS-MD14-Automovil.pdf>