

**MINERIA DE DATOS APLICADA AL SEGMENTO RETAIL DE LOS GASES
INDUSTRIALES**

**MURCIA MARTINEZ SANDRA PATRICIA C.C 52.058.342
QUINTERO MURILLO SANDRA PATRICIA C.C 52.106.561**

**UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD INGENIERIA DE MERCADOS
ESPECIALIZACION EN GERENCIA DE MERCADEO ESTRATEGICO
BOGOTA
2.013**

**MINERIA DE DATOS APLICADA AL SEGMENTO RETAIL DE LOS GASES
INDUSTRIALES**

MURCIA MARTINEZ SANDRA PATRICIA C.C 52.058.342

QUINTERO MURILLO SANDRA PATRICIA C.C 52.106.561

**TRABAJO DE GRADO EN MINERIA DE DATOS, PRESENTADO PARA OPTAR
AL TITULO DE ESPECIALIZACION EN GERENCIA DE MERCADEO
ESTRATEGICO**

ASESOR: LIZA PINZON – MBA

**UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD INGENIERIA DE MERCADOS
ESPECIALIZACION EN GERENCIA DE MERCADEO ESTRATEGICO**

BOGOTA

2.013

Nota de aceptación

Presidente del Jurado

Jurado: Liza Pinzón

Jurado: Humberto Martínez

Bogotá, 24 de Octubre del 2.013

Dedicamos este trabajo con todo nuestro amor y cariño a nuestros esposos que siempre nos han apoyado incondicionalmente, a nuestros amados hijos fuente de motivación e inspiración, y a las personas que de una u otra forma compartieron con nosotras, dándonos su conocimiento, alegrías y experiencias que nos hacen crecer y valorar estos momentos.

AGRADECIMIENTOS

El presente trabajo fue realizado bajo la supervisión de la profesora Liza Pinzón, a quien nos gustaría expresar nuestros más profundos agradecimientos, por la paciencia y dedicación.

A Dios, por darnos esta oportunidad, por permitirnos disfrutar cada día con las personas que más queremos, por guiarnos por el buen camino en cada paso que damos, por habernos dado nuestras estupendas familias que son nuestro soporte y compañía.

Agradecer a nuestras familias por el apoyo, comprensión, alegrías y fortalezas para salir adelante, por su amor y los maravillosos seres que son.

Para nuestros compañeros de grupo por compartir sus conocimientos, por su disposición a entregar información y experiencias profesionales.

CONTENIDO

	Pág.
INTRODUCCIÓN.	13
1. COMPRENSIÓN DEL NEGOCIO.	14
1.1 DEFINICIÓN DEL MERCADO.	14
1.2 ANÁLISIS DEL MERCADO.	14
1.3 SEGMENTACIÓN POR CATEGORIAS.	16
1.4 SEGMENTACIÓN GEOGRÁFICA.	17
1.5 ANÁLISIS DEL ENTORNO.	18
1.6 COMPAÑIAS LÍDERES.	19
1.6.1 Empresa 1.	19
1.6.2 Empresa 2.	20
1.6.3 Empresa 3.	20
1.6.4 Empresa 4.	20
1.7 PLANTEAMIENTO DEL PROBLEMA.	21
2. DETERMINACIÓN DE OBJETIVO.	23
3. PREPOCESAMIENTO DE LA BASE DE DATOS.	24
3.1 BASE DE DATOS PRE PROCESADA.	24
3.1.1 Primera base de datos pre procesada.	24
3.1.2 Segunda base de datos pre procesada.	25
3.1.3 Tercera base de datos pre procesada y escogida.	27
4. MINERIA DE DATOS.	32
4.1 DETERMINACION DE LOS OBJETIVOS.	32
4.2 REPROCESAMIENTO DE DATOS.	32

4.3	DETERMINACION DEL MODELO.	32
4.4	ANÁLISIS DE LOS RESULTADOS.	33
5.	MÉTODOS DE LAS APLICACIONES EN MINERÍA DE DATOS.	34
5.1	CLASIFICACIÓN DE LAS TÉCNICAS DE DATA MINING.	34
5.1.1	Métodos supervisados.	34
5.1.1.1	Predicción.	34
5.1.1.2	Clasificación.	35
5.2	Métodos no supervisados.	35
5.2.1	Asociación.	35
5.2.2	Segmentación.	35
6.	METODOLOGÍA.	36
6.1	INGRESO DE LA BASE DE DATOS SOFTWARE “RAPID MINER”.	37
6.2	SELECCIÓN DEL K.	50
6.3	CLUSTER.	51
6.3.1:	Clúster 0.	51
6.3.1.1	Descripción del clúster.	52
6.3.1.2	Beneficios cliente.	54
6.3.1.3	Beneficios compañía.	54
6.3.2	Clúster 1.	55
6.3.2.1	Descripción del clúster.	56
6.3.2.2	Beneficios cliente.	58
6.3.2.3	Beneficios compañía.	59
6.3.3	Clúster 2.	59
6.3.3.1	Descripción del clúster.	60
6.3.3.2	Beneficios compañía.	62
6.3.3.3	Beneficios para el cliente.	62

7.	CONCLUSIONES.	64
	RECOMENDACIONES.	65
	INFOGRAFÍA.	66

LISTA DE GRÁFICOS

	Pág.
Gráfico 1. Gases Industriales América del sur valor del mercado 2007-2011.	15
Gráfico 2. Gases Industriales América del Sur, valor del mercado.	16
Gráfico 3. Segmentación por categorías Gases Industriales.	17
Gráfico 4. Gases Industriales segmentación del Mercado por Geografía.	18
Gráfico 5. K-Means Primera BD Pre procesada.	25
Gráfico 6. Resultado de Rapid Miner para la segunda BD pre procesada.	27
Gráfico 7. Diagrama de los Métodos de la Minería de datos.	34
Gráfico 8. Método, Técnica y Algoritmo.	36
Gráfico 9. Ingreso al Software del Rapid Miner.	37
Gráfico 10. Bienvenida al Software de Rapid Miner.	38
Gráfico 51. Paso 1 Minería de datos.	38
Gráfico 12. Paso 2 Minería de Datos.	39
Gráfico 13. Paso 3 Minería de Datos.	40
Gráfico 14. Paso 4 Minería de Datos.	40
Gráfico 65. Paso 5 Minería de Datos.	41
Gráfico 16. Paso 6 Minería de Datos.	42
Gráfico 17. Paso 7 Minería de Datos.	42
Gráfico 187. Paso 8 Minería de Datos.	43
Gráfico 19. Paso 9 Minería de Datos.	44
Gráfico 20. K=2.	45
Gráfico 21. Clúster K=2.	45
Gráfico 22. K=3.	46

Gráfico 83. Clúster K=3.	46
Gráfico 24.K=4.	47
Gráfico 25. Clúster K=4.	47
Gráfico 96. K=5.	48
Gráfico 27. Clúster K=5.	48
Gráfico 28. K=6.	49
Gráfico 29. Clúster K=6.	49
Gráfico 30. K= 3 Seleccionado.	50
Gráfico 31. Clúster 0.	52
Gráfico 32. Representación del Clúster 0.	52
Gráfico 103. Cluster1.	55
Gráfico 34. Representación del Clúster 1.	56
Gráfico 35. Clúster 2.	60
Gráfico 116. Representación del Clúster 2.	60

LISTA DE TABLAS

	Pág.
Tabla 1. Primera Base de Datos Pre Procesada.	24
Tabla 2. Segunda Base de Datos Pre Procesada.	26
Tabla 3. Tercera Base Pre Procesada.	28
Tabla 4. Sucursales.	28
Tabla 5. Familias.	29
Tabla 6. Base de Datos por Sucursal y Familias.	29
Tabla 7. Pre Procesamiento de la BD, Porcentajes.	30
Tabla 8. Pre Procesamiento de la BD, se Reemplazan las comas por puntos.	31
Tabla 9. Composición de los Clúster más representativa.	51
Tabla 10. Estrategias de Fidelización Clúster 0 = Mantenimiento Industrial.	54
Tabla 11. Estrategias de Fidelización Clúster 1 = Reventa Industrial.	58
Tabla 12. Estrategias de Fidelización Clúster 2 = Medicinal.	62

LISTA DE IMÁGENES

	Pág.
Imagen 1. Fidelización de Clientes.	23
Imagen 2. Minería de Datos.	33
Imagen 3. Clúster 0.	51
Imagen 4. Reventa industria.	55
Imagen 5. Clúster medicinal.	59

INTRODUCCIÓN

Nuestra Compañía es una empresa del sector químico, con una amplia experiencia en el segmento Industrial y Medicinal en Colombia. La empresa se ha especializado en la producción y comercialización de gases de alta calidad para satisfacer las necesidades de los mercados de consumo.

Se cuenta con bases de datos de gran volumen de información, donde se puede encontrar el comportamiento de nuestros clientes, mediante las técnicas de Data Mining, que han demostrado que es posible analizar los datos almacenados por medios automáticos o semiautomáticos, con el fin de descubrir información valiosa para la gestión operativa y estratégica de la empresa, realizando análisis estadísticos y obteniendo patrones de comportamiento.

El presente trabajo contiene las estrategias que se utilizara para fidelizar a nuestros clientes, se enfocara en la bases de datos del segmento Maxi marketing por cada sucursal a nivel nacional, con el fin de poder identificar la tendencia y comportamiento de consumo de los clientes, conocer sus necesidades y poderlas satisfacer, buscando como resultado final el aumento en las ventas.

Lo primero que se realizara es procesar la base de datos de Maxi Marketing en el software de data mining, se utilizara el método no supervisado, con técnica de segmentación por tendencia de consumo y el algoritmo k-means, después se describirá cada clúster con sus estrategias de fidelización y beneficios para la compañía y el cliente, finalizando con las conclusiones y recomendaciones.

Como guía y soporte del trabajo, se tendrá en cuenta todos los conocimientos que se han adquirido durante el seminario y diferentes libros especializados sobre el tema.

1 COMPRENSIÓN DEL NEGOCIO

La compañía es una empresa que hace parte del sector Químico, que cuenta con más de 26.000 colaboradores en más de 50 países, su portafolio de productos se fundamenta en la transformación del aire en productos para la industria, transformación de subproductos en Dióxido de carbono (CO₂), Hielo seco, entre otros, así como asesoramiento en aplicaciones, dimensionamiento, construcción de redes, venta de equipos y consumibles.

Es una de las compañías más importantes de gases del aire a nivel mundial, se encuentra a la vanguardia en el desarrollo de productos y servicios para el sector industrial y medicinal.

1.1 DEFINICIÓN DEL MERCADO

El Mercado de gases incluye, los gases comprimidos, el suministro en el lugar de las grandes industrias, los gases de la salud y los gases utilizados en producción electrónica y otros.

1.2 ANÁLISIS DEL MERCADO

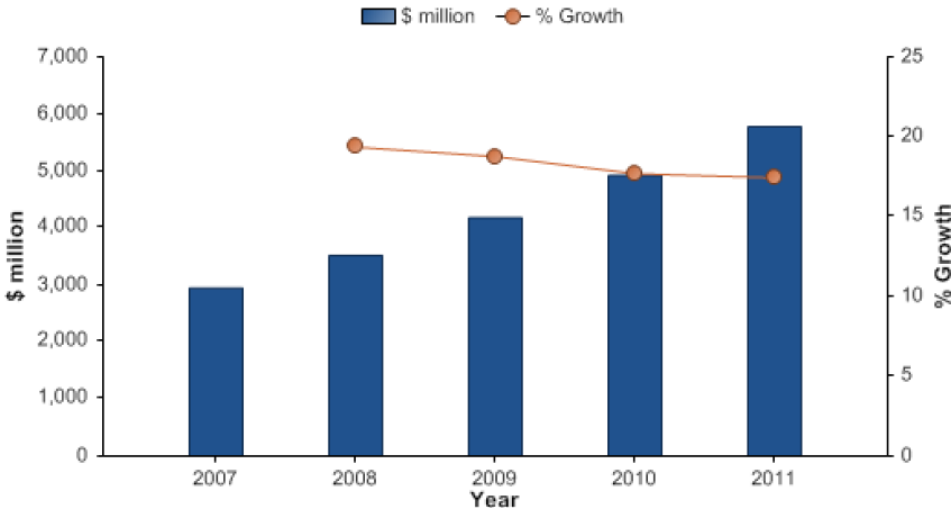
Para el propósito de este trabajo, se centrara en el mercado de gases industriales de América del Sur, el cual ha experimentado un dinámico crecimiento de dos dígitos en los últimos años, esta tendencia se pronostica que continuará, aunque a un ritmo ligeramente desacelerado.

El mercado de gases de América del Sur tuvo un ingreso total de US\$ 5,777.3 millones en el 2011, lo que representa una tasa de crecimiento anual del 18,3 % entre 2007 y 2011.

Se prevé que el comportamiento del mercado de una desaceleración, con un crecimiento del 14,4% para el período de cinco años 2011 - 2016, a un valor de \$ 11,312.6 millones a finales de 2016. Comparativamente.

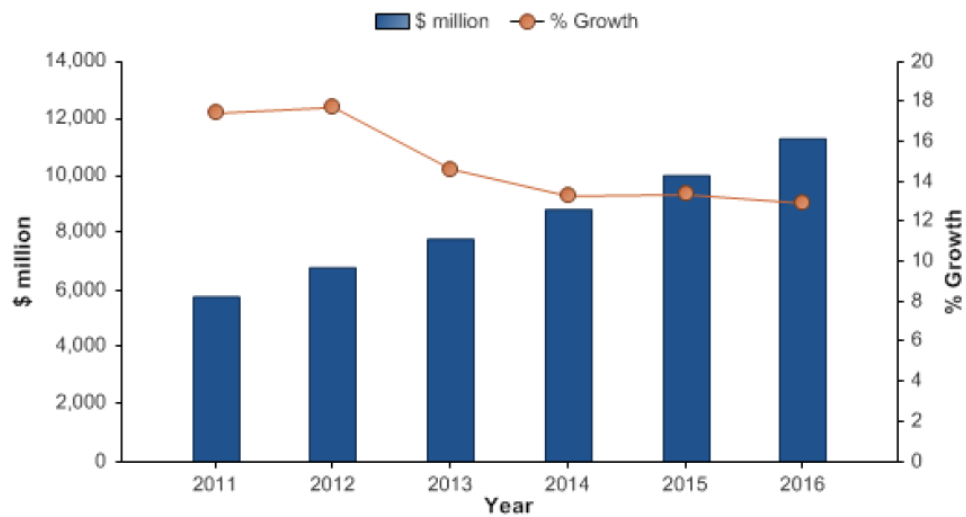
Los mercados de Argentina y Chile crecerán de 16,2% y 9,9%, respectivamente, durante el mismo período, para alcanzar valores respectivos de US\$2,014.6 millones y US\$ 398.700.000 en 2016.

Gráfico 12. *Gases Industriales en América del sur valor del mercado US\$ Millones, 2007-2011



Fuente: Market Line, 2012

Gráfico 13. Gases Industriales en América del Sur, valor del mercado Forecast: US\$ Millones, 2011-2016



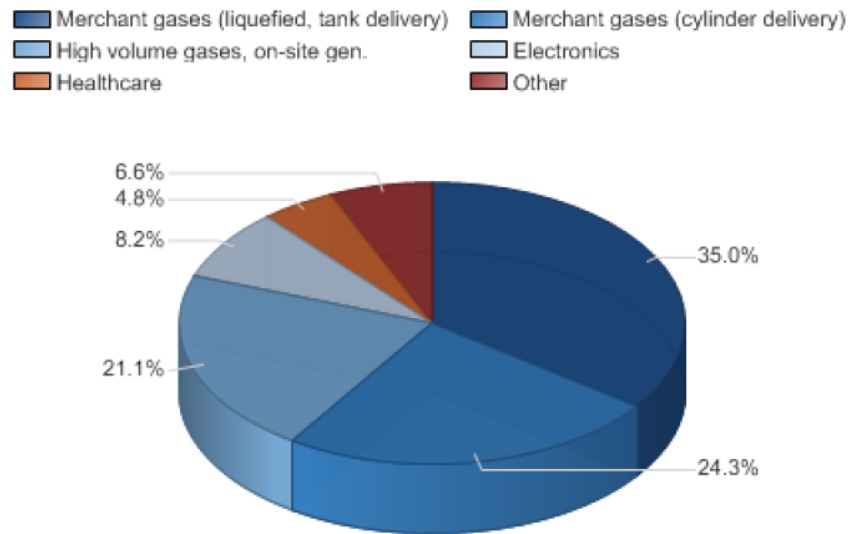
Fuente: Market Line, 2012

1.3 SEGMENTACIÓN POR CATEGORIAS

Los gases líquidos, presentan ingresos totales de US\$\$ 2,024.3 millones, equivalente al 35 % del valor total del mercado.

Los gases comprimidos, donde se centrara la investigación, presentan ingresos de US\$ 1,403.9 millones en 2011, lo que equivale al 24,3 % del valor total del mercado.

Gráfico 14. Segmentación por categorías Gases Industriales en América del sur, % de participación, por valor del 2011



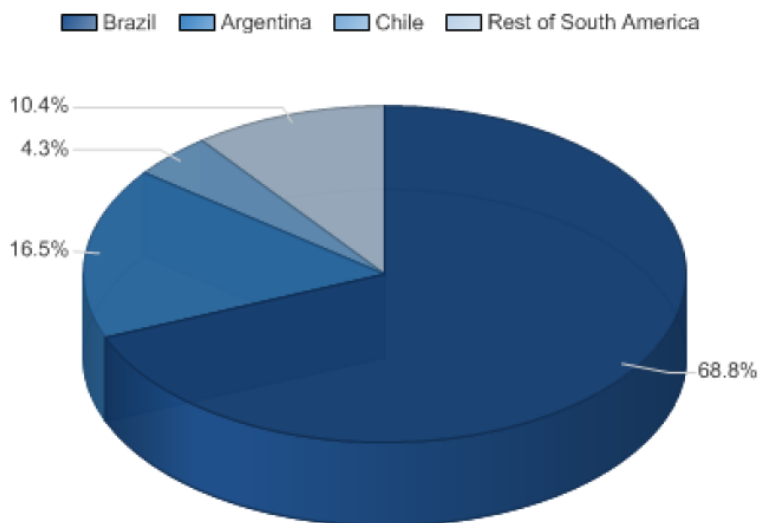
Fuente: Market Line, 2012

1.4 SEGMENTACIÓN GEOGRÁFICA

Brasil representa el 68,8% del valor de mercado de gases industriales de América del Sur.

Argentina representa otro 16,5% del mercado de América del Sur.

Gráfico 15. Gases Industriales segmentación del Mercado por Geografía: US\$ Millones, 2011



Fuente: Market Line, 2012

1.5 ANÁLISIS DEL ENTORNO

El mercado de gases industriales se caracteriza por la presencia de empresas internacionales grandes y diversificadas. Lo que traduce “una importante barrera a la entrada” y “la necesidad de inversión inicial fuerte para establecer plantas de producción de alto volumen”, esta condición reduce la amenaza de nuevas empresas que se instalen en este mercado.

Suplentes en el mercado de gases industriales son pocos, ya que en muchos casos los compradores de gases, compran en base a sus necesidades específicas, propiedades físicas y químicas, lo que reduce el riesgo de alternativas. Los elevados costos fijos y barreras de salida intensifican la rivalidad en el mercado.

A medida que estos gases son esenciales para muchos procesos de fabricación - por ejemplo, la industria del metal utiliza el oxígeno, el nitrógeno y argón, la industria del automóvil utiliza acetileno, propano, mezclas de gases combustibles, y el oxígeno, mientras que el nitrógeno líquido es vital para el reciclaje de

plásticos, envases y desechos de neumáticos: ejerce actores de la industria en una posición muy fuerte.

Por otro lado, el poder de compra se incrementa por el hecho de que los gases industriales son productos diferenciados sólo por factores, como la pureza y la marca, que es de poca importancia para los compradores, en general, el poder de compra se valora como moderado. La falta de diferenciación en el producto, hace que los compradores se vuelvan propensos a la conmutación debido a la poca fidelidad a la marca, esta situación hace aumentar la rivalidad entre los agentes del mercado, pero en sí mismo debería favorecer nuevos entrantes.

Los agentes del mercado también deben invertir en tuberías y recipientes para el transporte de gases industriales, de igual forma también hay un marco normativo importante en el mercado del gas industrial, que es restrictiva a la entrada de jugadores. Los nuevos operadores deben prestar especial atención a regulaciones ambientales estrictas, que se aplican a la producción, almacenamiento y transporte de gases industriales.

Los gases industriales son muy específicas en su composición, lo que significa la existencia de muy pocos sustitutos.

Los compradores adquieren los gases en función de sus propiedades físicas y químicas particulares, e incluso cuando un gas sustituto (es decir, uno que es ligeramente diferente en la composición, pero ofrece propiedades similares a la del usuario final) puede ser utilizado, es normalmente fabricados por las mismas empresas, como proveedores de este mercado son limitadas.

Los principales jugadores tienen operaciones diversificadas, lo que significa que no son exclusivamente dependientes de la venta de gases industriales.

1.6 COMPAÑIAS LÍDERES

1.6.1 Empresa 1: Se especializa en el suministro de gases industriales y medicinales y servicios asociados. La empresa suministra oxígeno, nitrógeno, hidrógeno y otros gases a una amplia gama de industrias, incluyendo el acero, la refinación de petróleo, productos químicos, vidrio, electrónica, papel, soldadura, metalurgia, procesamiento de alimentos, la salud y la industria aeroespacial.

Empresa 1 está presente en más de 80 países de Europa, América, Asia Pacífico y África y el Medio Oriente.

La compañía opera a través de tres segmentos de negocio: gas y los servicios, la ingeniería y la construcción, y otras actividades.

1.6.2 Empresa 2: Empresa 2 se enfoca en sector de tecnología, la energía, la industria y clientes del sector sanitario a nivel mundial con su cartera de productos, servicios y soluciones que incluyen gases atmosféricos, gases de proceso y especiales, materiales de alto rendimiento, equipos y servicios.

La compañía es uno de los mayores proveedores de hidrógeno y helio y cuenta con posiciones de liderazgo en los mercados como los materiales semiconductores, hidrógeno de refinería, licuefacción de gas natural, y recubrimientos avanzados y adhesivos.

Empresa 2 opera a través de tres segmentos de negocios: gases, la electrónica y el rendimiento de materiales y equipo y la energía.

1.6.3 Empresa 3: Empresa 3 es un productor y proveedor de gases industriales. El grupo también está involucrado en actividades de construcción industrial, el grupo opera principalmente a través de tres divisiones operativas de negocio: gases, ingeniería, y otras actividades. Las divisiones de gas y la ingeniería operan como Empresa 3 Gas e Ingeniería Empresa 3, respectivamente.

La división de negocios de Empresa 3 Gas es un proveedor de gases industriales. La división ofrece gases comprimidos y licuados, y productos químicos para diversas industrias. Empresa 3 Gas se centra en la producción, venta y distribución de gases de aplicaciones en el sector de la energía, la producción de acero, procesos químicos, la protección del medio ambiente, la soldadura, el procesamiento de alimentos, producción de vidrio, y la electrónica. También se dedica a la producción de gases de soldadura de protección, dióxido de carbono, monóxido de carbono, gases inertes, gases medicinales, gases purificados, y los gases de prueba. Esta división también está invirtiendo en la expansión de su negocio de la salud, que incluye gases medicinales.

1.6.4 Empresa 4: Empresa 4, es uno de los más grandes proveedores de gases industriales en Norte y Sudamérica. Empresa 4 sirve aproximadamente 25 industrias, incluyendo la salud y el petróleo, fabricación de chips de computadora y bebidas, carbonatación, la fibra óptica y la fabricación de acero, y la aeroespacial, química, y tratamiento de aguas.

La compañía opera en América del Norte, América del Sur, Europa y Asia.

Empresa 4 organiza sus operaciones en cuatro segmentos reportables geográficamente: América del Norte, América del Sur, Europa y Asia. Además, Empresa 4 opera su negocio mundial de tecnologías de la superficie, a través de su subsidiaria de propiedad total, Empresa 4 SurfaceTechnologies, lo que representa el quinto segmento reportable.

Las operaciones de Empresa 4 que constituyen las dos principales líneas de productos son: gases industriales y tecnologías de superficie. Los gases industriales en la fabricación y distribución de gases atmosféricos (oxígeno, nitrógeno, argón, y gases nobles) y gases de proceso (dióxido de carbono, helio, hidrógeno, gases electrónicos, gases especiales, y acetileno).

1.7 PLANTEAMIENTO DEL PROBLEMA

La empresa en Suramérica, segmenta sus clientes según su consumo, giro de activos, rotación de cartera y rentabilidad, en clientes Oro, Plata, Premium y **Maxi Marketing**, específicamente para esta investigación se enfocara el Segmento Maxi Marketingen Países Hispánicos, esta situación se da, debido a que la compañía ha centralizado la atención de estos clientes desde Colombia a través de una célula de negocios cuya promoción actual se realiza por Marketing directo, mediante tele marketing, este segmento está compuesto por clientes con consumos esporádicos, clientes retail, de todas las industrias incluyendo el sector medicinal. En la actualidad el segmento cuenta con más de 12.000 clientes a nivel regional.

Para la Corporación este tipo de clientes se transforma en uno de los mas importante debido a la rentabilidad que estos generan, lo que lo convierte en el segmento más competido en el mercado por la fuerte rivalidad no solamente de compañías multinacionales sino de pequeños distribuidores que compiten con servicio y portafolios complementarios logrando la Fidelización de los clientes.

Es importante resaltar algunas variables relevantes que dificultan la Fidelización de los clientes de nuestro segmento en investigación, y que según el análisis del

mercado, donde se expuso la gran inversión en activos que las compañías deben realizar para atender el sector de la industria, y la poca fidelidad del sector, para los clientes de segmento Maxi marketing la inversión en Cilindros es muy fuerte con el agravante de la pérdida de los mismos, por el cambio constante de proveedor y la falta de control, lo que se traduce en el deterioro del capital

Esta realidad hace que la compañía como estrategia decida fidelidad y captar este tipo de clientes, para lograr este objetivo, se utilizó Data Mining, como técnica, lo que va a permitir la exploración de las bases de datos, con el objetivo de lograr segmentar los clientes por tendencia de consumo e identificar sus necesidades para generar nuevas oportunidades.

2 DETERMINACIÓN DE OBJETIVO

Objetivo: Segmentar por sucursales, el tipo de consumo de nuestro portafolio de productos, para identificar tendencia y comportamiento de consumo y así poder diseñar estrategias de Fidelización a los clientes de nuestras sucursales que llevaran al aumento en ventas.

Imagen 1. Fidelización de Clientes



Fuente: Google images, 2013

3 PREPROCESAMIENTO DE LA BASE DE DATOS

Inicialmente se trabajo una base de datos de una sucursal donde se evidenciaron dificultades para el análisis, una tendencia marcada a un solo Clúster representativo, independiente de lo K- means atribuidos, y no permitía generar estrategias.

3.1 BASE DE DATOS PRE PROCESADA

3.1.1 Primera Base de Datos Pre procesada

Base de datos Sucursal, procesamiento de 125 clientes con 126 Productos, excluyendo servicios.

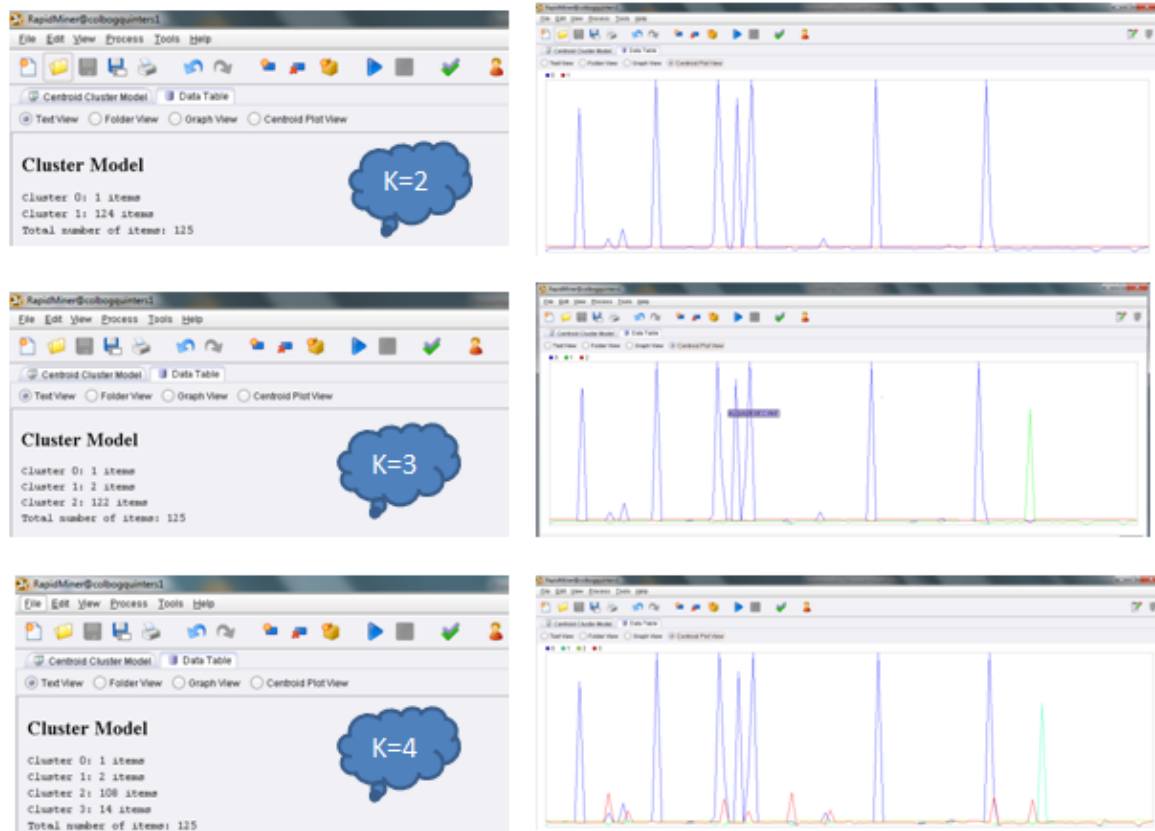
Tabla 10. Primera Base de Datos Pre Procesada

ID	ACETILENO	ACETILENO A_A	AIRE SINTETICO 4_7	AIRE SINTETICO 4_7 FID	ALAMBRE COBREADO PROSTAR 0_9MM	ALAMBRE COBREADO PROSTAR 1_2MM	ANHIDRIDO CARB_EN CILINDR	ANHIDRIDO CARB_TERM OS KGRS	ANHIDRIDO CARBONICO 4_0	ARGON 4_8	ARGON 5_0 ANALITICO
7658	0	0	0	0	0	0	0	0	0	0	0
10074	0	0	0	0	0	0	0	0	0	0	0
13598	0	0	0	0	0	0	0	0	0	0	0
13905	0	0	0	0	0	0	0	0	0	0	0
16544	0	0	0	0	0	0	0	0	0	0	0
100368	0	0	0	0	0	0	0	0	0	0	0
100450	0	0	0	0	0	0	0	0	0	0	0
100620	0	0	0	0	0	0	0	0	0	0	0
100705	0	0	0	0	0	0	0	0	0	0	0
100997	0	0	0	0	0	0	0	0	0	0	0
101421	0	0	0	0	0	0	0	0	0	0	0
101586	0	0	0	0	0	0	0	0	0	0	0
101749	0	0	0	0	0	0	0	0	0	0	0
101868	0	0	0	0	0	0	0	0	0	0	0
102205	0.582259169	0	0	0	0	0	0	0	0	0	0
102339	0	0	0	0	0	0	0	0	0	0	0
103518	0	0	0	0	0	0	0	0	0	0	0
105184	0	0	0	0	0	0	0	0.862562243	0	0	0
105378	0	0	0	0	0	0	0	0	0	0	0
105595	0	0	0	0	0	0	0	0	0	0	0
105729	0	0	0	0	0	0	0	0	0	0	0
105862	0	0	0	0	0	0	0	0	0	0	0

Fuente: Las autoras 2013

Gráfico 5. K-Means Primera BD Pre procesada

Modelos de Clúster de la base 1:



Fuente: Las autoras 2013

3.1.2 Segunda Base de Datos Pre procesada: Base de datos a Nivel Nacional, procesamiento de 9 Sucursales y 219 Productos del portafolio, excluyendo servicios, luego de pre procesado la base de datos.

Tabla 11. Segunda Base de Datos Pre Procesada

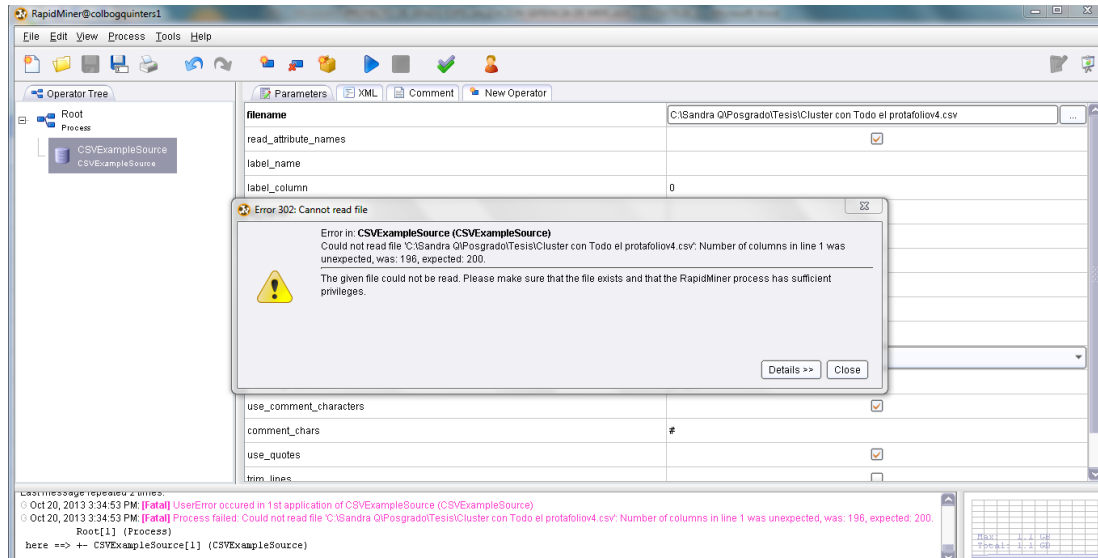


SUCURSAL	ACETILENO25 ABSORCIONA TOMIC	ACETILENO CILINDRO ARTICULA R	ACETILENO LINDROST D	AGUADESTIL ADA400ML	AIRE47FID	AIRECOMPRI MIDOSECO	AIREMEDICINA L COMPRIMIDO STD	AIRE SINTETICO 47	ALAMBRE TUBURA RE71T11EN00
BARRANQUILLA	0.000812730695	0	0.10101002654	0	0.012709520	0	0.0122645343639	0	0.000429943633607
BOGOTA	0.001743876879	0.003330725	0.01265429184	0	0.007998402	0	0.0002050260112	0.00287935917134	0
BUCARAMANGA	0	0.004399340	0.02501061884	0	0.000614500	0.00033381575	0.0010948554842	0	0
CALI	0.014283848026	0.004434414	0.05347995792	0	0	0	0.0015699350196	0	0
CARTAGENA	0	0.002126731	0.12285397620	0	0	0	0	0	0
IBAGUE	0.013027226529	0.000499232	0.03079936227	0	0.007539702	0.00078165082	0.0055918498844	0	0
MEDELLIN	0.014905652316	0.013313937	0.02610131518	0.00001923168	0.009257928	0	0	0	0
PEREIRA	0.005845529951	0.000817258	0.04990575782	0	0.003370908	0	0	0	0
SOGAMOSO	0	0	0.02639861068	0	0	0	0.0039534060973	0	0

Fuente: Las autoras 2013

Al momento de usar el software y subir la base de datos el resultado fue errorError in: CSVExampleSource (CSVExampleSource) No se pudo leer el archivo 'C: \ Sandra Q \ Posgrado \ Tesis \ Clúster con Todo el protafoliov4.csv': El número de columnas en la línea 1 fue inesperado, fue: 196, que se espera: 200. El archivo especificado no se pudo leer. Por favor, asegúrese de que el archivo existe y que el proceso RapidMiner tiene suficientes privilegios.

Gráfico 6. Resultado de Rapid Miner para la segunda BD pre procesada



Fuente: Las autoras 2013

Esto permite concluir que el software no tiene los suficientes privilegios para procesar la base de datos con el número de productos suministrado.

3.1.3 Tercera Base de Datos Pre procesada y escogida: Base de datos de 9 Sucursales, procesamiento de 11 familias con 199 Productos.

Se procedió a realizar la consulta de la información en el sistema de la compañía, donde se tomaron los siguientes datos para generar el reporte:

- Sucursales que manejan Marximarketing a nivel nacional.
- Ventas en pesos de los productos con fechas desde Enero al mes de Septiembre del 2013
- Volúmenes y metros cúbicos de los meses ya mencionados.

Tabla 12. Tercera Base Pre Procesada

ID	Productos	Ene-2013 (Ventas en Unidades Totales)	Ene-2013 (Ventas en Pesos)	Ene-2013 (Ventas en m3)
21-00147	- SUCURSAL BARRANQUILLA	1.906	27.663.363	1.450
1104203	+ ACARREO ACETILENO A.A			
1901034	+ ACARREO ACETILENO CILINDRO STD	0	50.862	0
1105103	+ ACARREO AIRE MEDICINAL COMPRIM			
1901118	+ ACARREO ARGON CILINDRO STD	0	383.728	0
1302101	+ ACARREO CO2 CILINDRO STD	0	11.000	0
1901033	+ ACARREO CO2 PARTICULAR			
1302106	+ ACARREO DIOXIDO DE CARBONO 4.0			
1106301	+ ACARREO FLY BALLOON			
1102401	+ ACARREO NITROGENO GASEOSO	0	10.500	0
1102501	+ ACARREO NITROGENO GASEOSO 4.6			
1102201	+ ACARREO NITROGENO LIQ.EN TERMO	0	15.000	0
1901442	+ ACARREO NITROGENO LIQUIDO EN T			
1901364	+ ACARREO NITRÓGENO ESTÁNDAR 300			
1901107	+ ACARREO OXIGENO CILINDRO INDUS	0	484.277	0
1101501	+ ACARREO OXIGENO GRADO ABO			
1101201	+ ACARREO OXIGENO LIQUIDO EN TER	0	21.400	0
1901406	+ ACARREO STARGOLD TUB	0	121.432	0
104203	+ ACETILENO 2.5 ABSORCION ATOMIC			
901034	+ ACETILENO CILINDRO STD	62	2.137.642	0
105302	+ AIRE 4.7 FID	14	385.799	14

Fuente: Las autoras 2013

Se procede a organizar la base de datos por las nueve (9) sucursales:

Tabla 13. Sucursales

BARRANQUILLA
BOGOTA
BUCARAMANGA
CALI
CARTAGENA
IBAGUE
MEDELLIN
PEREIRA
SOGAMOSO

Fuente: Las autoras 2013

Tabla 14. Familias

Los 199 productos se ingresaron a las 11 familias:

Carbuero
Cilindros GNV
Consumibles
Equipos Medicinales
Equipos y Accesorios Ind
Gases Comprimidos
Gases Especiales
Gases Licuados
Gases Medicinales
Prueba Hidrostatica
Soldadura

Fuente: Las autoras 2013

Tabla 15. Base de Datos por Sucursal y Familias

Quedando de la siguiente forma:

SUCURSAL	FAMILIA	PRODUCTO	Ventas 2013
BARRANQUILLA	Gases Especiales	ACETILENO2.5ABSORCIONATOMIC	214.986
BARRANQUILLA	Gases Comprimidos	ACETILENOCILINDROSTD	26.719.480
BARRANQUILLA	Gases Especiales	AIRE4.7FID	3.361.961
BARRANQUILLA	Gases Medicinales	AIREMEDICINALCOMPRIMIDOSTD	3.244.252
BARRANQUILLA	Consumibles	ALAMBRETUBURARE71T11EN0.0	113.730
BARRANQUILLA	Gases Comprimidos	ARGONCILINDROSTD	42.380.882
BARRANQUILLA	Gases Comprimidos	ARGONGASEOSOCILINDROPARTICU	174.000
BARRANQUILLA	Equipos y Accesorios Ind	CARRODOBLEPARAEQUIOPPU	280.000
BARRANQUILLA	Equipos y Accesorios Ind	CILINDROALUMINIO682LTSCGA	165.000
BARRANQUILLA	Cilindros GNV	CILINDROGNV100.324MAT	870.000
BARRANQUILLA	Cilindros GNV	CILINDROGNV40.267MAT	13.400.000
BARRANQUILLA	Cilindros GNV	CILINDROGNV93.390MAT	2.700.000
BOGOTA	Gases Comprimidos	ACETILENOCILINDROPARTICULAR	1.657.029
BOGOTA	Gases Especiales	ACETILENO2.5ABSORCIONATOMIC	867.575
BOGOTA	Gases Comprimidos	ACETILENOCILINDROSTD	6.295.483
BOGOTA	Gases Especiales	AIRE4.7FID	3.979.188
BOGOTA	Gases Medicinales	AIREMEDICINALCOMPRIMIDOSTD	102.000
BOGOTA	Gases Especiales	AIRESINTETICO4.7	1.432.475
BOGOTA	Gases Especiales	ARGON4.8	630.060
BOGOTA	Gases Especiales	ARGON5.0	15.311.978

BOGOTA	Gases Comprimidos	ARGONCILINDROSTD	13.365.294
BOGOTA	Gases Comprimidos	ARGONGASEOSOCILINDROPARTICU	215.385
BOGOTA	Equipos y Accesorios Ind	ASIENTOREGULADOROXIGENOPGR	24.200
BOGOTA	Gases Especiales	BAR90BAJORANGO	1.134.000
BOGOTA	Gases Especiales	BAR90MEDIORANGO	2.030.000
BOGOTA	Consumibles	BOQUILLADECORTEHARRISPLANA	18.000
BOGOTA	Consumibles	CABEZALCA2460REF03020093	43.395

Fuente: Las autoras 2013

Se procede a organizar la información por sucursales y productos, donde se determinan los porcentajes de cada familia en la venta por sucursal.

Tabla 16. Pre Procesamiento de la BD, Porcentajes

SUCURSAL	Carbuero	Cilindros GNV	Consumibles	Equipos Medicinales	Equipos y Accesorios Ind	Gases Comprimidos	Gases Especiales	Gases Licuados	Gases Medicinales	Prueba Hidrostatica	Soldadura
BARRANQUILLA	0	0.064440172	0.000606543	0.001184757	0.00246065	0.634024	0.064035137	0.089819	0.042964621	0.000341757	0.1001238
BOGOTA	0.004976	0.417559378	0.002867074	0	0.048831023	0.273902	0.195263849	0.010459	0.001537859	0.001235488	0.0433677
BUCARAMANGA	0.016127	0.358406711	0.000253945	0	0.009570565	0.522939	0.035055189	0.023666	0.003044725	0.004297178	0.0266392
CALI	0	0.230831297	0.000519695	0	0.131586832	0.404281	0.229282865	0	0.002978789	0.000519695	0
CARTAGENA	0	0	0.010909687	0	0	0.651816	0.11551511	0	0.002784302	0	0.2189752
IBAGUE	0	0	0.003725335	0.064057152	0.030954993	0.54183	0.136039874	0.057244	0.042974441	0.005292642	0.1178818
MEDELLIN	0.017325	0	5.25215E-05	0	0.075769527	0.475646	0.231337365	0.136577	0.002283267	0.000556567	0.0604531
PEREIRA	0	0	0.000658711	0.216000782	0.030790039	0.377673	0.142715558	0.138655	0.021746086	0.012391297	0.0593687
SOGAMOSO	0.002987	0.09005375	0.012505314	0	0.005628359	0.457127	0.076456188	0.037203	0.008324707	0.002770885	0.306943

Fuente: Las autoras 2013

Se reemplazan los campos vacíos por cero (0), y las comas se reemplazaron por puntos, guardando el archivo en formato CVS.

Tabla 17. Pre Procesamiento de la BD, se Reemplazan las comas por puntos

SUCURSAL	Carburo	Cilindros_GNV	Consumibles	Equipos_Medicinales	Equipos_y_Accesorios_Ind	Gases_Comprimidos	Gases_Especiales	Gases_Licuidos
BARRANQUILLA	0 0 0644401720184255	0 000606542644460996	0 00118475743487029	0 00246065005703829	0 634023617190139	0 0640351368676787	0 0898189369767713	
BOGOTA	0 00497597120593846	0 417559377681357	0 00286707418912508	0 0 0488310229299227	0 273902294821014	0 195263849294958	0 0104593949711987	
BUCARAMANGA	0 0161273968906502	0 358406711103542	0 000253944789049843	0 0 0095705650651085	0 522939237008328	0 0350551887132742	0 0236658852627456	
CALI	0 0 230831297037507	0 000519695227851798	0 0 131586831692075	0 0 404280826601591	0 229282865106122	0 0 651815718619547	0 115515110314306	
CARTAGENA	0	0 0 0109096865237198	0 0640571524355531	0 0309549926574042	0 541829893118315	0 136039874468021	0 0572438690863618	
IBAGUE	0	0 0 00372533495449222	0 0 075769526843739	0 476646051456381	0 231337365360231	0 136577115183203		
MEDELLIN	0 0173245101609213	0 0 0000525215192335749	0 216000781570137	0 030790039091164	0 377673404898594	0 142715557610149	0 138655431740243	
PEREIRA	0	0 0 000658711105250046	0 0 00562835939454231	0 457127158597054	0 0764561881923878	0 0372032650996067		
SOGAMOSO	0 002987359986334	0 090053750312677	0 0125053140371699					

Fuente: Las autoras 2013

Con este paso terminamos de reprocesar la base de datos, y podemos iniciar la metodología escogida de Data Mining, en este caso manejaremos el Software Rapid Miner.

4 MINERIA DE DATOS

Se puede definir la minería de datos como el descubrimiento automático de patrones o modelos interesantes y no obvios escondidos en una base de datos, los cuales tienen un gran potencial para contribuir en los aspectos principales del negocio que se traducen en estrategias o tácticas de una empresa.

Otra definición es: el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos, la minería de datos surge para intentar ayudar a comprender el contenido de las bases.

La minería de datos, desde un punto de vista estrecho, comprende, como sistema de extracción de relaciones, los métodos basados en la computadora, requiriendo poca involucración y ayuda por parte del analista en la obtención de información relevante. Se incluye aquí los algoritmos de redes neuronales artificiales, árboles de decisión, inducción de reglas, lógica difusa, análisis de links y los algoritmos genéticos.

En el momento que el usuario les atribuye algún significado especial pasan a convertirse en información, aunque en la minería de datos cada caso concreto puede ser radicalmente distinto al anterior, el proceso común a todos ellos se suele componer de cuatro etapas principales:

4.1 DETERMINACION DE LOS OBJETIVOS

Trata de la delimitación de los objetivos que el cliente desea bajo la orientación del especialista en la minería de datos.

4.2 REPROCESAMIENTO DE DATOS

Se refiere a la selección, la limpieza, el enriquecimiento, la reducción y la transformación de las bases de datos. Esta etapa consume generalmente alrededor del setenta por ciento del tiempo total de un proyecto de la minería de datos.

4.3 DETERMINACION DEL MODELO

Se comienza realizando unos análisis estadísticos de los datos, y después se lleva a cabo una visualización gráfica de los mismos para tener una primera aproximación. Según los objetivos planteados y la tarea que debe llevarse a cabo,

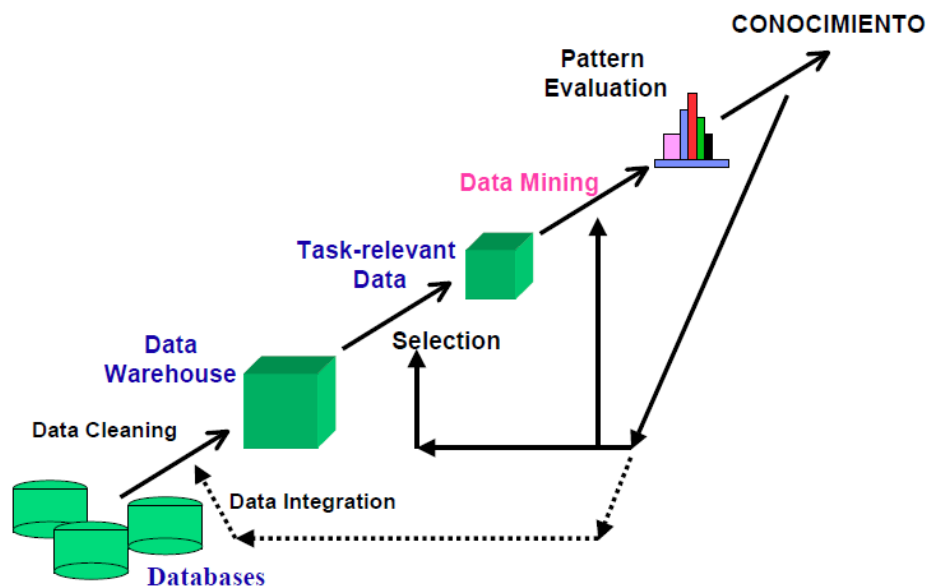
pueden utilizarse algoritmos desarrollados en diferentes áreas de la Inteligencia Artificial.

4.4 ANÁLISIS DE LOS RESULTADOS

Verifica si los resultados obtenidos son coherentes y los coteja con los obtenidos por los análisis estadísticos y de visualización gráfica. El cliente determina si son novedosos y si le aportan un nuevo conocimiento que le permita considerar sus decisiones.

En resumen, la minería de datos se presenta como una tecnología emergente, con varias ventajas: por un lado, resulta un buen punto de encuentro entre los investigadores y las personas de negocios; por otro, ahorra grandes cantidades de dinero a una empresa y abre nuevas oportunidades de negocios.

Imagen 2. Minería de Datos



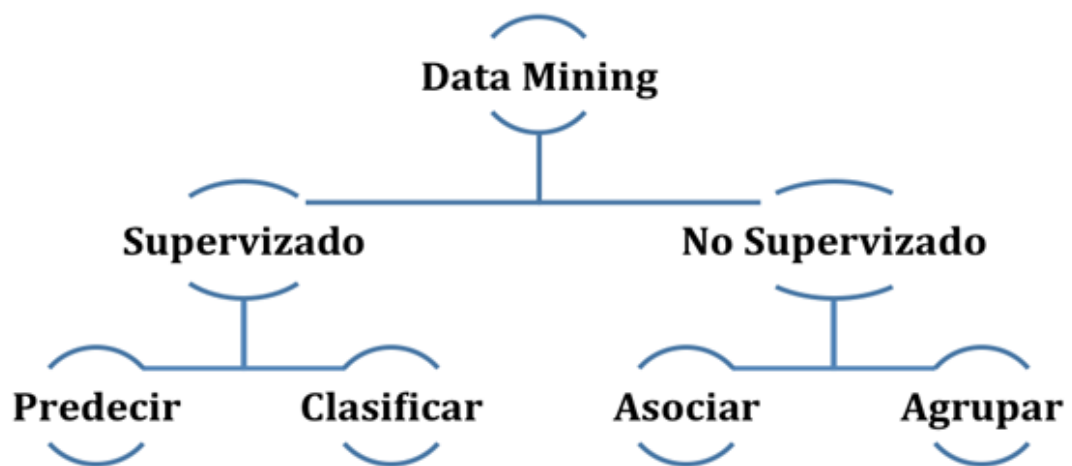
Fuente: Jose M. Gutiérrez, Extracción conocimiento en grandes bases de datos, 2010

5 Métodos de las Aplicaciones en Minería de Datos

Supervisados: Predictivos

No Supervisados: Descubrimiento del Conocimiento

Gráfico 7. Diagrama de los Métodos de la Minería de datos



Fuente: Universidad Sergio Arboleda, Semillero de investigación, 2013

5.1 CLASIFICACIÓN DE LAS TÉCNICAS DE DATA MINING

Las técnicas de Data Mining se pueden clasificar en Predicción, Clasificación, Clustering y Asociación

5.1.1 MÉTODOS SUPERVISADOS

5.1.1.1 Predicción: El análisis de predicción está relacionado con la técnica de regresión. La idea principal del análisis de predicción es descubrir las relaciones

entre variables dependientes e independientes y las relaciones entre variables independientes.

5.1.1.2 Clasificación: Los métodos tienen la intención de aprender diferentes funciones que clasifiquen los datos dentro de un conjunto predefinido de clases. Dado un nuevo de clases predefinidas, un número de atributos y un conjunto de datos de aprendizaje o entrenamiento, los métodos de clasificación pueden automáticamente predecir la clase de los datos previamente no clasificados.

Las claves más problemáticas relacionadas con la clasificación son las evaluación de los errores de clasificación y la potencia de predicción.

Las técnicas matemáticas más usadas para la clasificación son los arboles de decisión binarios, las redes neuronales, programación lineal y estadística.

5.2 MÉTODOS NO SUPERVISADOS

5.2.1 Asociación: La relación entre un ítem de una transacción y otro ítem en la misma transacción es utilizada para predecir patrones.

Cuando el factor de confianza y al factor de soporte están representados por las variables lingüísticas alto y bajo, la regla de asociación se puede escribir en forma de lógica difusa, como: “cuando el factor de soporte es bajo, X implica Y es alto”. Usan los algoritmos de reglas de asociación y árboles de decisión

5.2.2 Segmentación: El análisis de clúster toma datos sin agrupar y mediante el uso de técnicas automáticas realiza la agrupación de estos. El clustering no está supervisado y no requiere un set de datos de aprendizaje.

Usan los algoritmos de clustering y de sequence clustering.

Para nuestra investigación la técnica a utilizar según nuestro objetivo será Segmentación y/o Clustering.

6 METODOLOGÍA

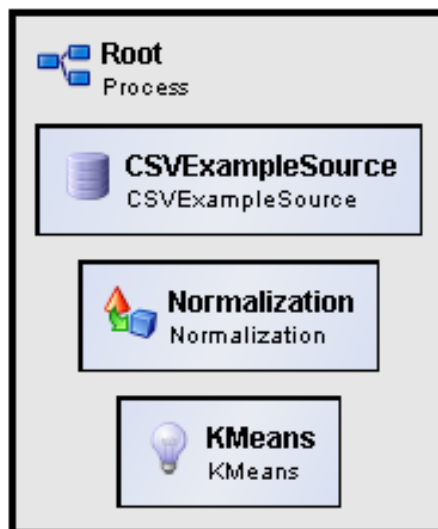
El método, técnica y algoritmo con el que se trabajó la base de datos fue el siguiente:

MÉTODO: No supervisado.

TÉCNICA: Segmentación por tendencia de consumo.

ALGORITMO: KMeans

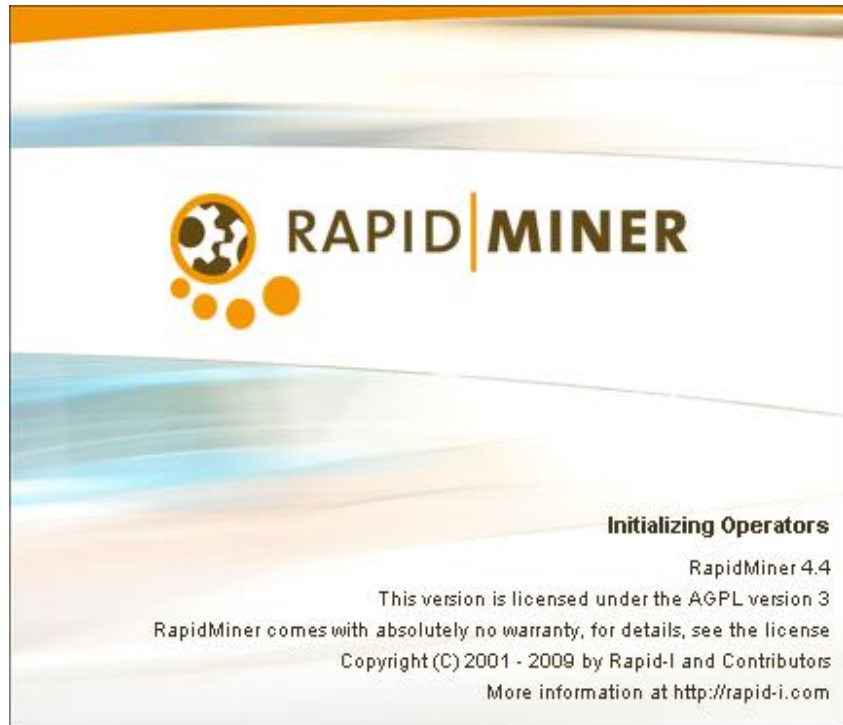
Gráfico 8. Método, Técnica y Algoritmo



Fuente: Las autoras 2013

6.1 INGRESO DE LA BASE DE DATOS EN EL SOFTWARE “RAPID MINER”

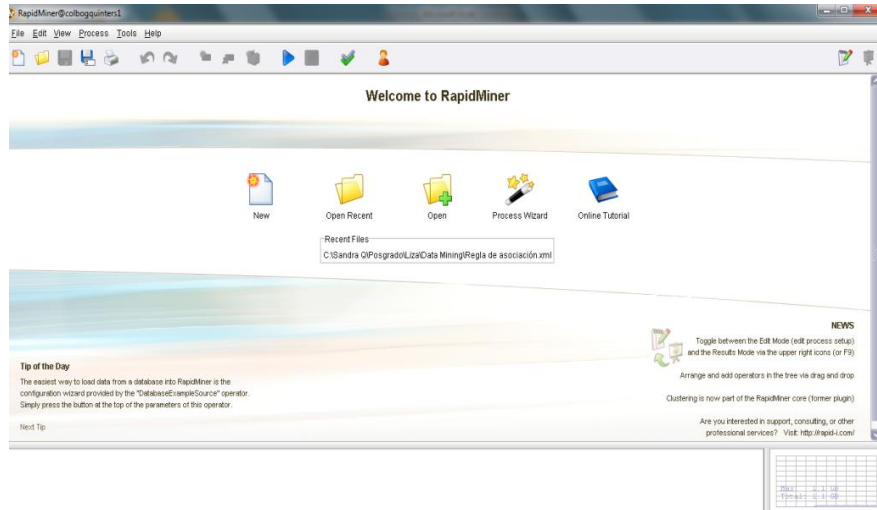
Grafico 9. Ingreso al Software del Rapid Miner



Fuente: Las autoras 2013

Gráfico 10. Bienvenida al Software de Rapid Miner

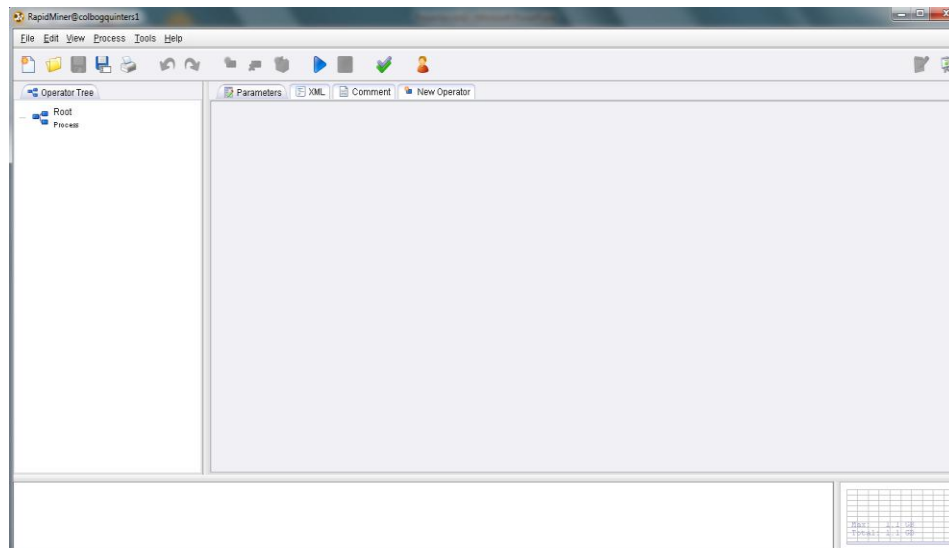
Se ingresa a New:



Fuente: Las autoras 2013

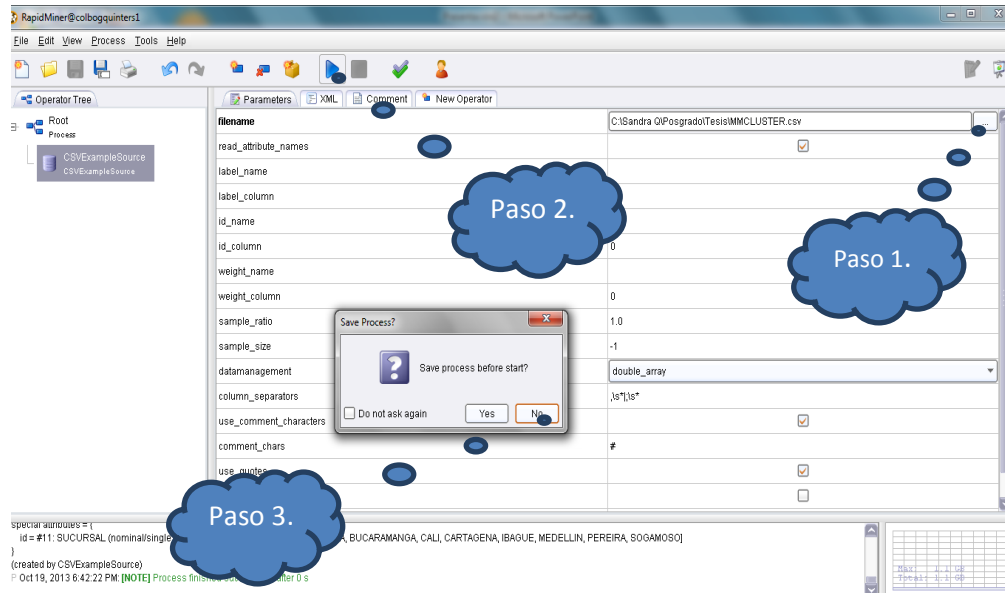
Gráfico 161. Paso 1 Minería de datos

Ubicación en Root - Process



Fuente: Las autoras 2013

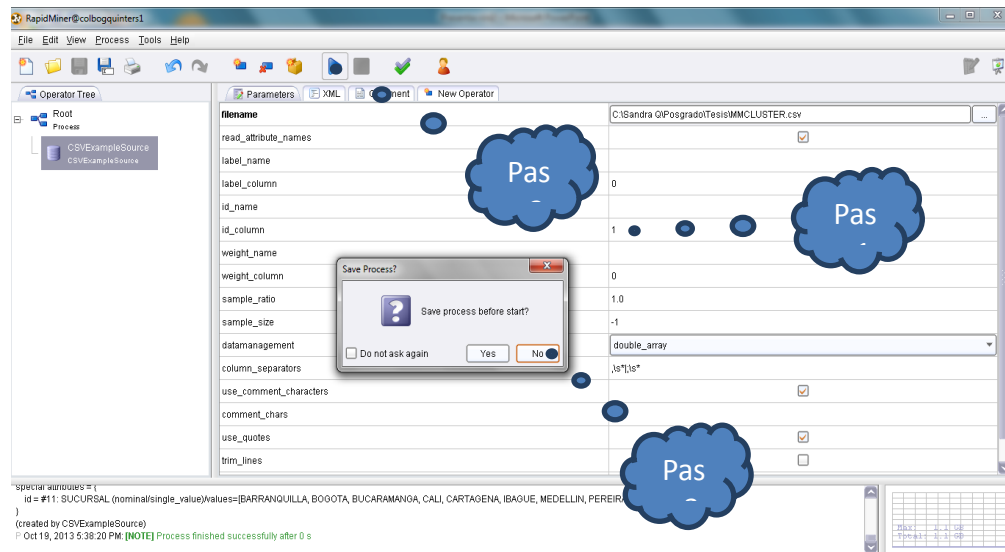
Gráfico 13. Paso 3 Minería de Datos



Fuente: Las autoras 2013

Gráfico 14. Paso 4 Minería de Datos

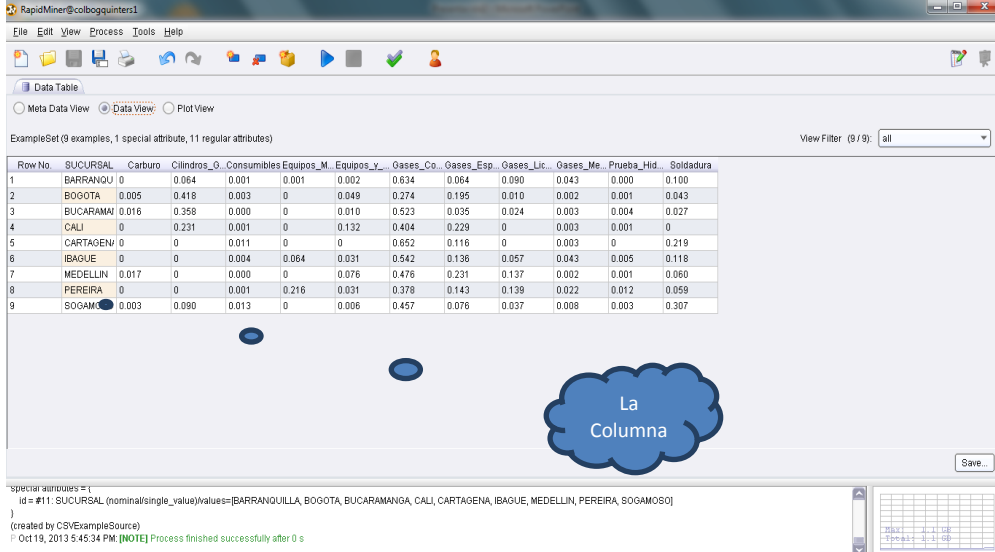
Se fija la columna definida, en este caso Sucursales



Fuente: Las autoras 2013

Gráfico 175. Paso 5 Minería de Datos

Se puede observar que después de procesar la información la Columna definida queda fija.

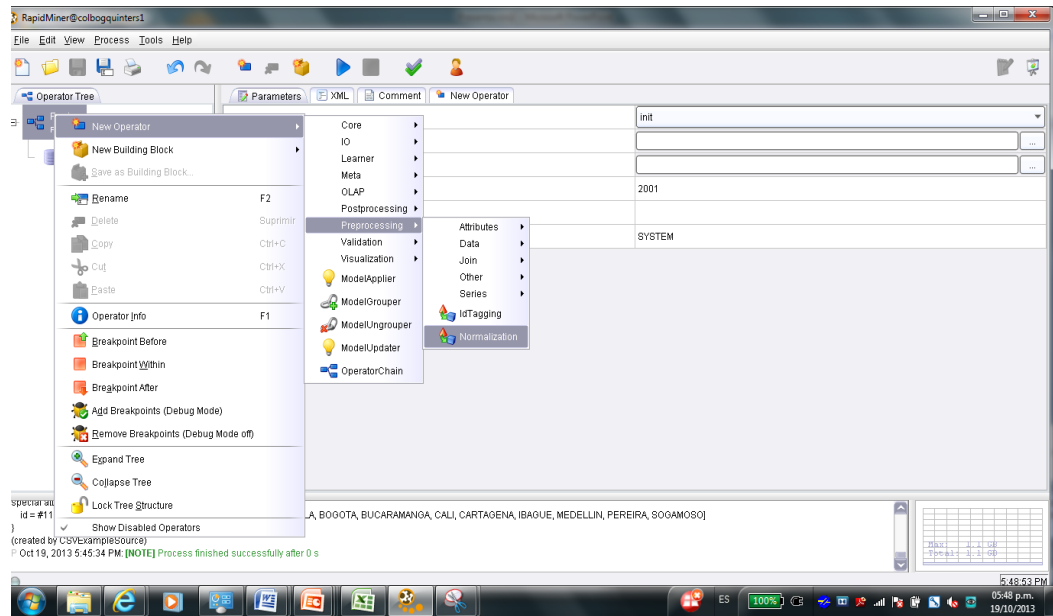


Row No.	SUCURSAL	Carburo	Cilindros_G.	Consumibles	Equipos_M.	Equipos_Y.	Gases_Co.	Gases_Esp.	Gases_Lic.	Gases_Me.	Prueba_Hid.	Soldadura
1	BARRANQU	0	0.084	0.001	0.001	0.002	0.634	0.064	0.090	0.043	0.000	0.100
2	BOGOTA	0.005	0.418	0.003	0	0.049	0.274	0.195	0.010	0.002	0.001	0.043
3	BUARAMAN	0.016	0.358	0.000	0	0.010	0.523	0.035	0.024	0.003	0.004	0.027
4	CALI	0	0.231	0.001	0	0.132	0.404	0.229	0	0.003	0.001	0
5	CARTAGENI	0	0	0.011	0	0	0.652	0.116	0	0.003	0	0.219
6	IBAGUE	0	0	0.004	0.004	0.031	0.542	0.136	0.057	0.043	0.005	0.118
7	MEDELLIN	0.017	0	0.000	0	0.076	0.476	0.231	0.137	0.002	0.001	0.060
8	PEREIRA	0	0	0.001	0.216	0.031	0.378	0.143	0.139	0.022	0.012	0.059
9	SOGAMOSO	0.003	0.090	0.013	0	0.006	0.457	0.076	0.037	0.008	0.003	0.307

Fuente: Las autoras 2013

Se procede a normalizar la información, Este operador realiza una normalización. Esto se puede hacer entre un definido valor mínimo y máximo del usuario o por un z-transformación, es decir, en la media 0 y varianza 1. O por una transformación proporcional como proporción de la suma total del atributo respectivo.

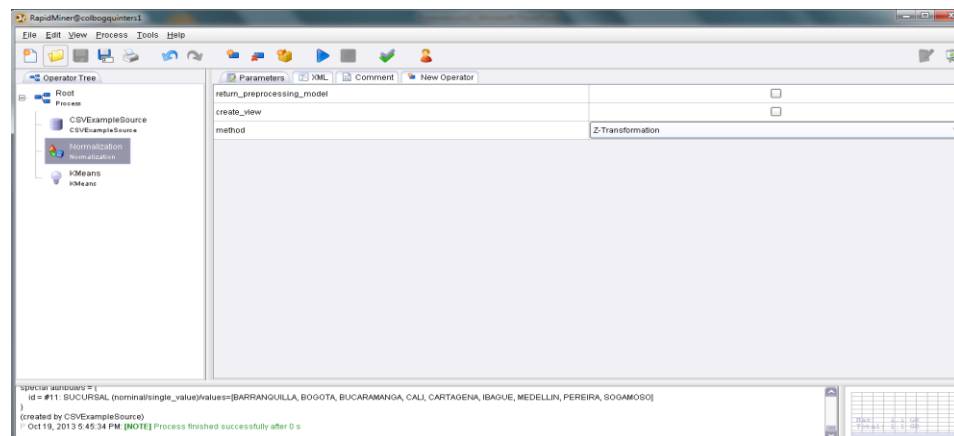
Gráfico 16. Paso 6 Minería de Datos



Fuente: Las autoras 2013

Gráfico 17. Paso 7 Minería de Datos

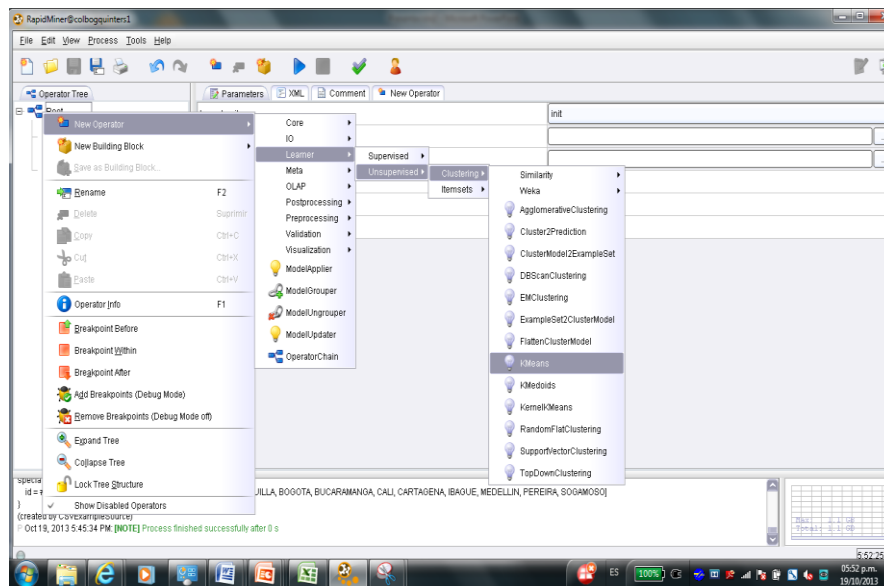
La Normalización se realiza por el método z-transformación



Fuente: Las autoras 2013

Se generan los K-means según el procedimiento anexo, Este operador representa una implementación de K-medias. Este operador creará un atributo de grupo si no está presente todavía, se establece como Máximo un K=6, lo que se traduce en dividir la base de datos en 6 grupos.

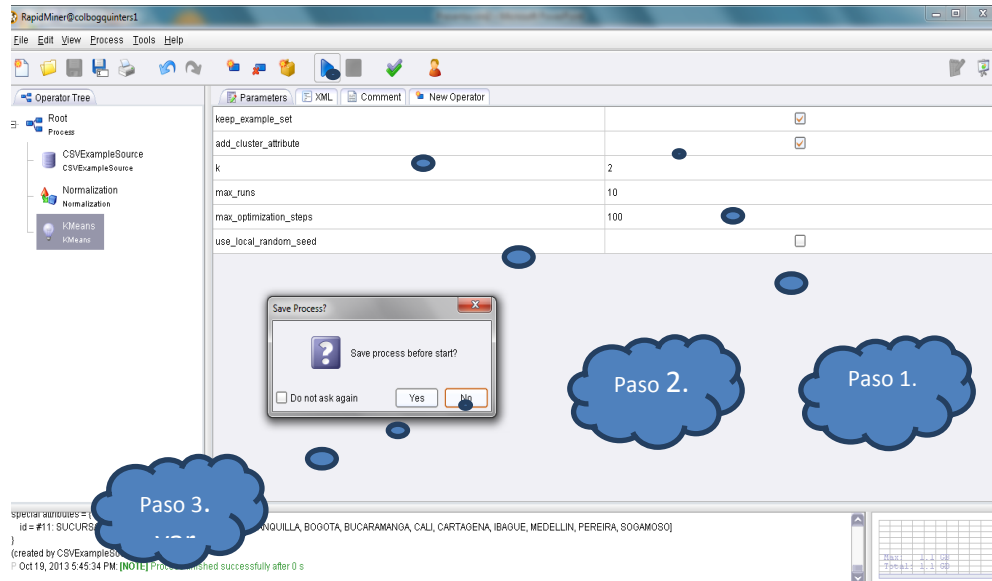
Gráfico 1818. Paso 8 Minería de Datos



Fuente: Las autoras 2013

Se selecciona el número de Clúster óptimo (K) a ser analizado, según la distribución de los mismos y los atributos.

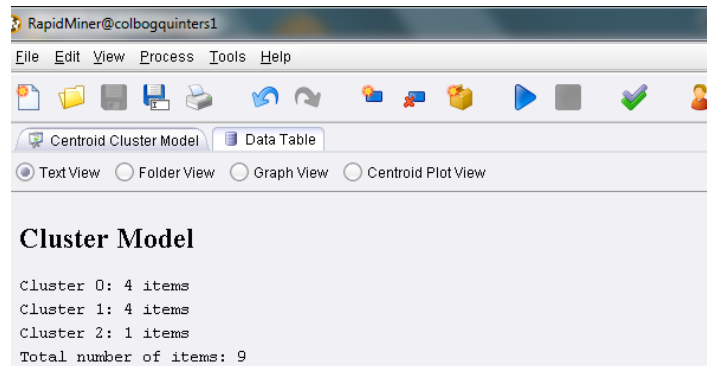
Gráfico 19. Paso 9 Minería de Datos



Fuente: Las autoras 2013

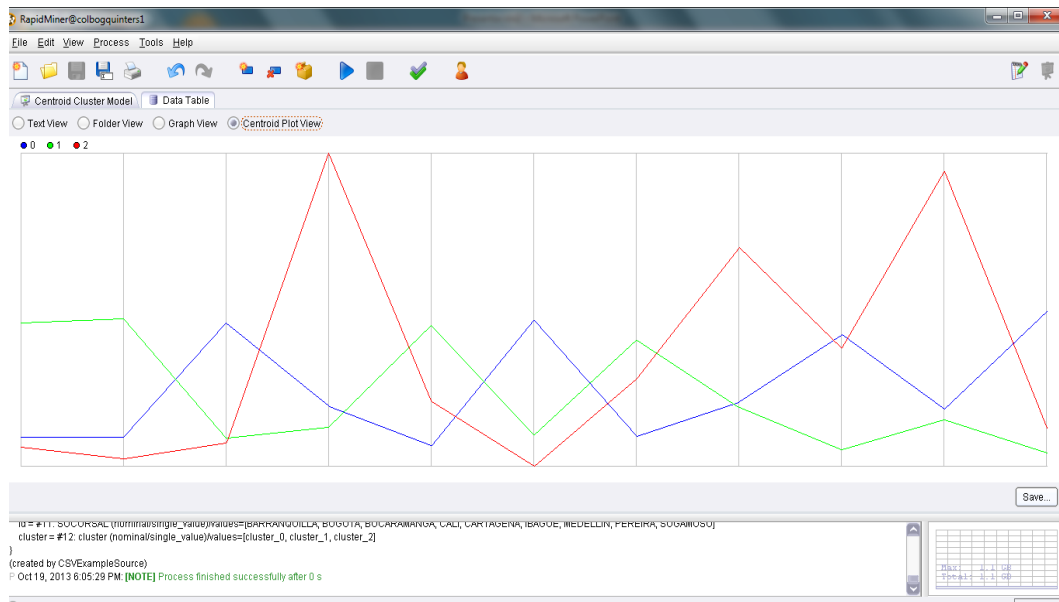
Se observan las gráficas accediendo a Centroid Plot View.

Gráfico 22. K=3



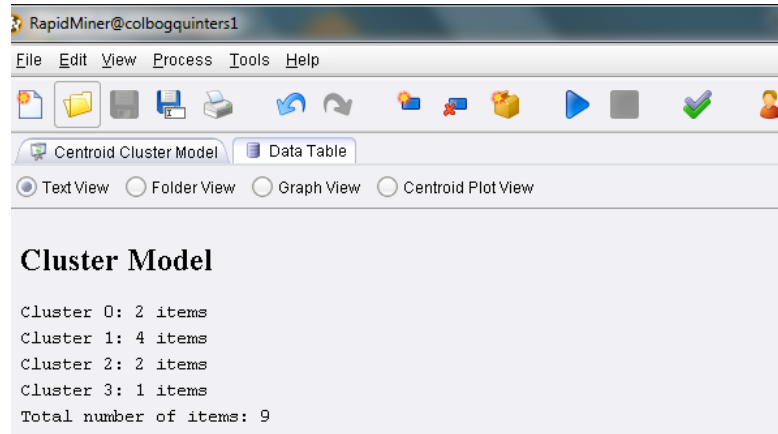
Fuente: Las autoras 2013

Gráfico 193. Cluster K=3



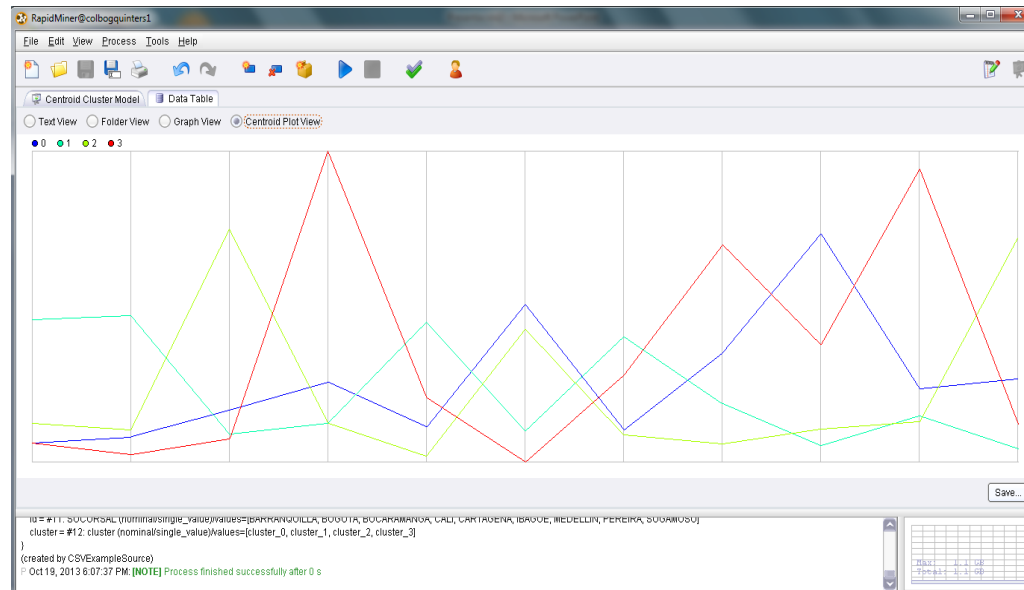
Fuente: Las autoras 2013

Gráfico 24.K=4



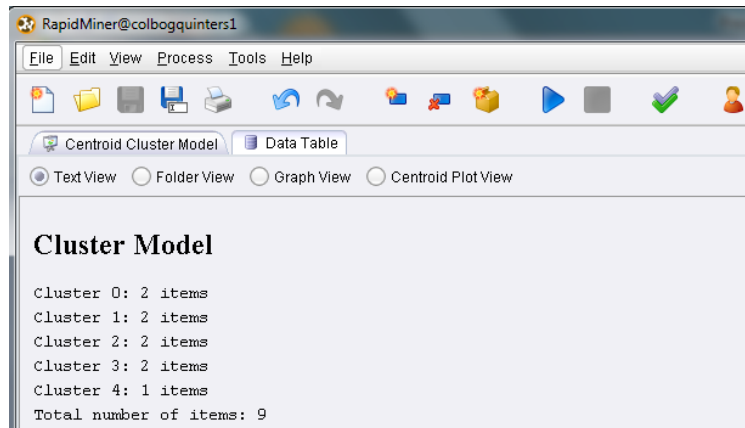
Fuente: Las autoras 2013

Gráfico 25. Cluster K=4



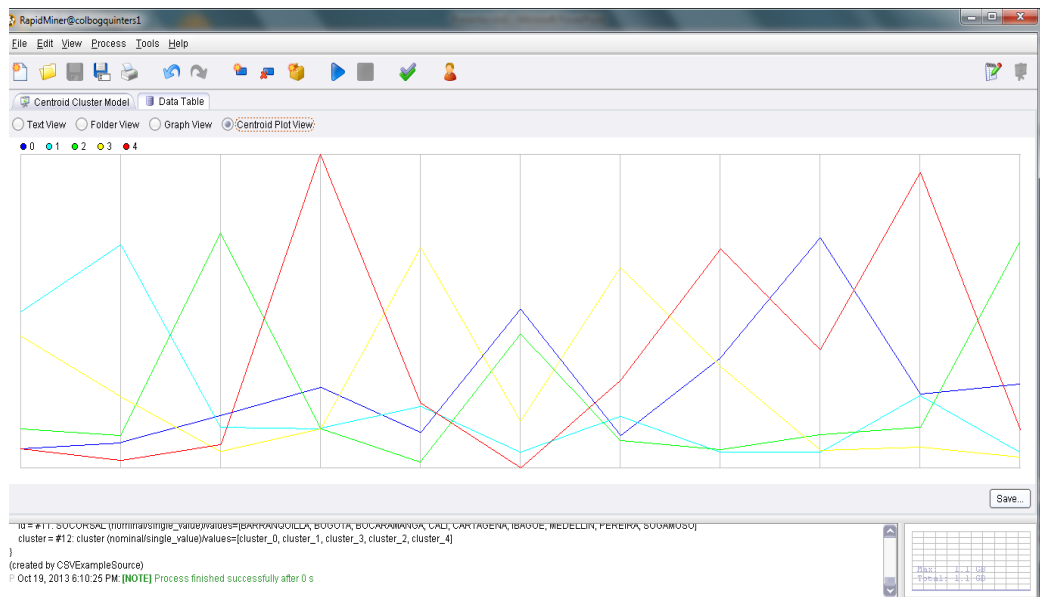
Fuente: Las autoras 2013

Gráfico 206. K=5



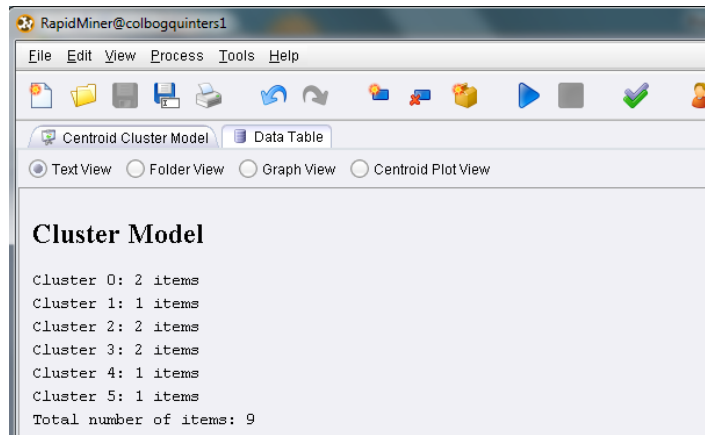
Fuente: Las autoras 2013

Gráfico 27. Cluster K=5



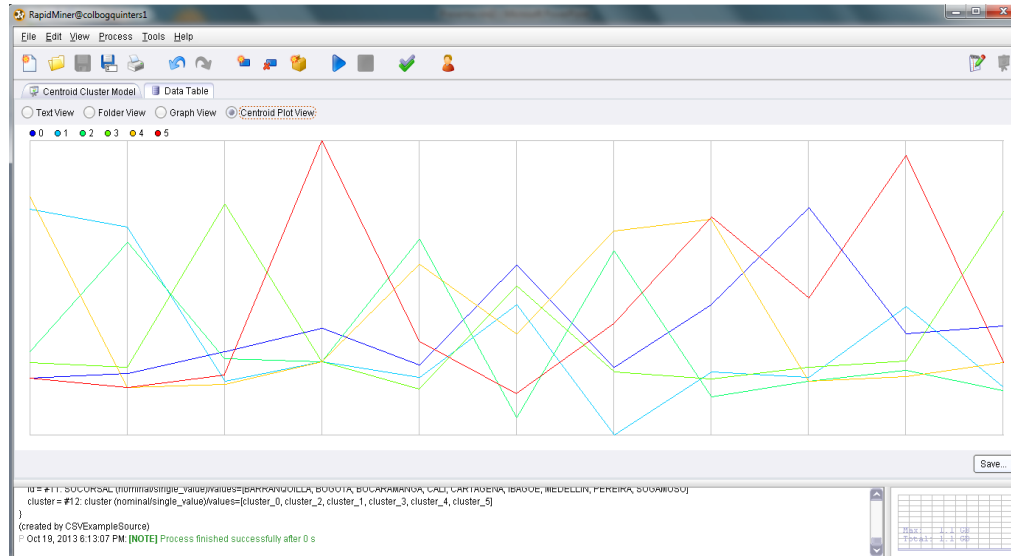
Fuente: Las autoras 2013

Gráfico 28. K=6



Fuente: Las autoras 2013

Gráfico 29. Cluster K=6

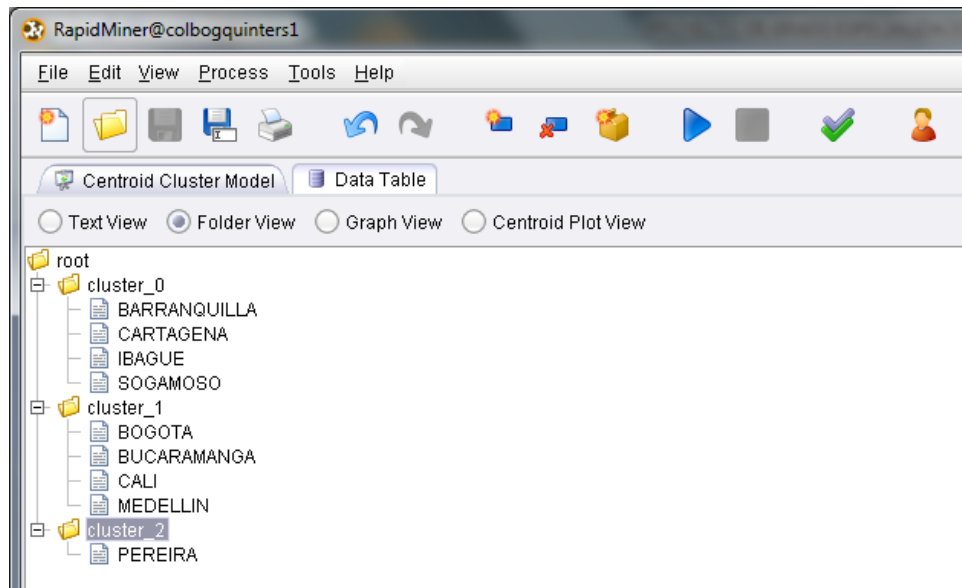


Fuente: Las autoras 2013

6.2 SELECCIÓN DEL K

Se determina que el K-means, más representativo para la investigación y objetivo es el K=3, por la segmentación realizada, se analizan los tres clúster determinados por la técnica:

Gráfico 30. K= 3 Seleccionado



Fuente: Las autoras 2013

Tabla 18. Composición de los Cluster más representativa

Familias	Cluster 0	Cluster 1	Cluster 2
Carburo		x	
GNV		x	
Consumibles	x		
Equipos Med			x
Equipos Ind		x	
Gases Ind	x		
Gases Esp			
Gases Lic			x
Gases Med	x		
Prueba H			x
Soldadura	x		

Fuente: Las autoras 2013

6.3 CLUSTER

6.3.1: Clúster 0.

Imagen 3. Mantenimiento Industrial

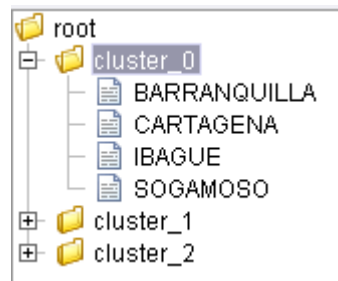


Fuente: Google images, 2013

Nombre: Mantenimiento Industrial

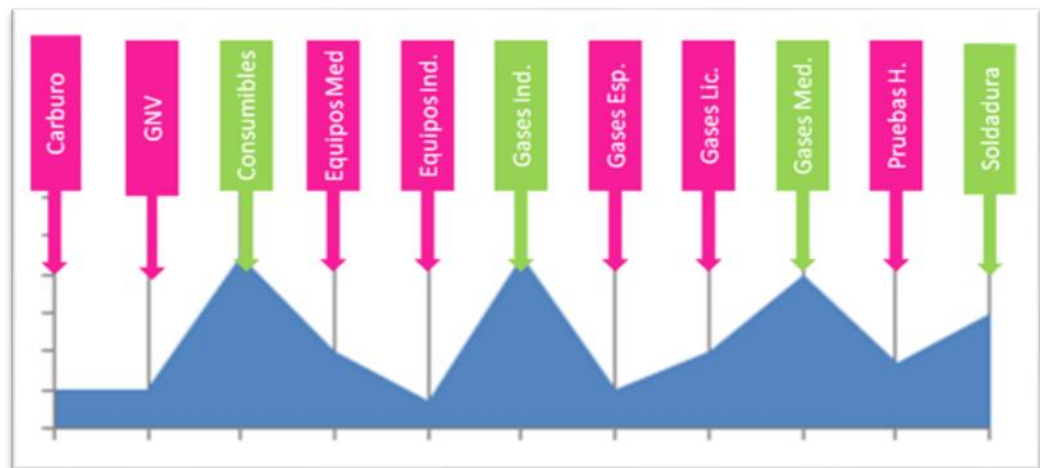
Composición: Compuesto por cuatro sucursales que equivalen al 44% del País.

Gráfico 31. Clúster 0



Fuente: Las autoras 2013

Gráfico 32. Representación del Clúster 0



Fuente: Las autoras 2013

6.3.1.1 Descripción del Clúster: Clientes del sector industrial, dedicados a procesos de mantenimiento y reparación en las diferentes industrias, donde sus

compras más representativas son los gases industriales, los consumibles para soldadura y la soldadura.

Son clientes que compran según sus necesidades inmediatas de consumo, y el volumen de trabajo que se presente en el momento, requieren de oportunidad en la entrega, debido a que su gran volumen de trabajo no es planeado, y es realizado por lo general a clientes internos de las diferentes áreas de la compañía que requieren su servicio y a clientes externos que requieren el servicio de mantenimiento y/o reparación y lo tercerizan según sus requerimientos específicos.

Adicional a los gases y la soldadura que adquieren, requieren de otras soluciones como equipos, accesorios, y acompañamiento en el desarrollo de aplicaciones para hacer más eficiente su trabajo y presentar menores pérdidas en los procesos.

Están Ubicados en las ciudades de Barranquilla, Cartagena, Ibagué y Sogamoso, las dos primeras con un gran desarrollo en la actualidad por la entrada del TLC, Ibagué y Sogamoso, por el asentamiento de industrias del cemento, acero y otras, que requieren servicios de mantenimiento constantemente y prefieren estos se realicen a través de contratistas.

Este segmento está dirigido en su gran mayoría por profesionales de la industria metalúrgica, metalmecánica, soldadores, e ingenieros, que buscan precio y oportunidad en la entrega, así como asignación de activos para poder desarrollar los diferentes frentes de trabajo que se les presenten, requieren también el desarrollo de un gran portafolio que les permita poder conseguir todos los equipos y accesorios, incluyendo los de seguridad, para la correcta ejecución de sus tareas.

Tabla 10. Estrategias de Fidelización Clúster 0 = Mantenimiento Industrial

Estrategia	Fidelización al Segmento Mantenimiento Industrial
Nombre	Tu Tarjeta
Descripción	Estrategia de fidelización, a través de Programa para otorgar puntos por la compra únicamente de Hardgoods (equipos, accesorios y seguridad industrial). Los cuales se pueden cambiar por Hardgoods, Según el número de puntos requerido por producto, los gases no acumulan puntos ni da premios
Como	Los puntos son acumulados por el sistema, por cada compra, la compañía otorga una tarjeta con código de barras que identifica el comprador, puede ser utilizada a nivel nacional.
Comunicación	Promoción de ventas – Programa de continuidad
Nombre	Combos Industriales
Descripción	Estrategia de On-Pack, que incluyan gases, soldadura y consumibles, según la aplicación a desarrollar, donde se otorgara un descuento del 10% en el valor real de la compra.
Como	Desarrollo de aplicaciones en clientes, que nos permitan conocer sus necesidades y crear los combos especializados.
Comunicación	Venta personalizada y especializada – Programa de incentivos
Nombre	Demo-Van
Descripción	Estrategia de fidelización con servicio pos-venta, asesor especializado que cuenta con una Demo - Van dotada de todas las herramientas necesarias para entregar soluciones en corte, soldadura y pruebas, con demostraciones a domicilio, asesorías y capacitación, con el fin de mejorar la productividad de los procesos en los clientes y el consumo de nuestros gases, equipos, accesorios y otros.
Como	Venta personalizada y especializada - Servicio pos venta
Comunicación	Promoción de ventas - demostraciones

Fuente: Las autoras 2013

6.3.1.2 Beneficios Cliente: Oportunidad en el servicio, capacitación en las diferentes aplicaciones desarrolladas en el segmento, lo que contribuirá a la productividad de sus procesos, soporte técnico permanente a domicilio para desarrollo de nuevas aplicaciones, descuentos permanentes en los productos de mayor consumo, acumulación de puntos que le permitirá adquirir gratis accesorios, consumibles y equipos necesarios para el desarrollo de su portafolio de servicios.

6.3.1.3 Beneficios Compañía: Fidelizar los clientes pertenecientes a este clúster, mayor rentabilidad, conocimiento del cliente que nos permitirá una mejor rotación

de mercancías y manejo de inventarios, así como un buen control de activos, lo que permitirá la reasignación de los mismos para garantizar el suministro oportuno y el aumento de las ventas de este clúster disminuyendo la inversión en capital.

6.3.2 Clúster 1.

Imagen 4. Reventa industrial

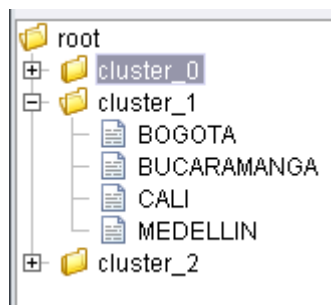


Fuente: Google images, 2013

Nombre: Reventa Industrial

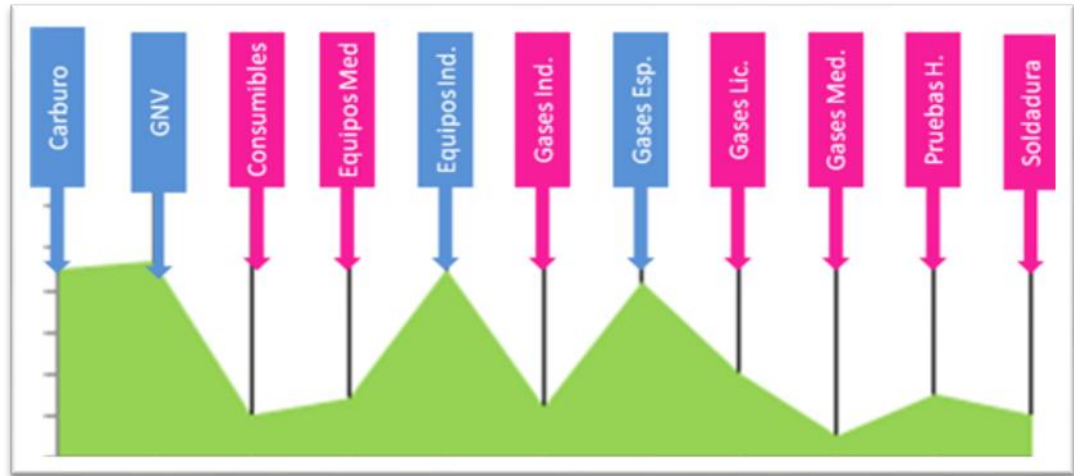
Composición: Compuesto por cuatro (4) sucursales que equivalen al 44% de las sucursales del País.

Gráfico 213. Cluster1



Fuente: Las autoras 2013

Gráfico 34. Representación del Clúster 1



Fuente: Las autoras 2013

6.3.2.1 Descripción del Clúster: Clientes del sector industrial, dedicados a procesos de transformación tales como. los que convierten los vehículos a GNV, los que sustituyen el acetileno por el carburo de calcio, generando una reacción exotérmica con el agua que produce cal apagada y acetileno, su uso es muy extendido y vendido en el sector de mantenimiento para utilizarlo en soldaduras autógena, la venta de equipos industriales al sector, así como consumidores de Gases Especiales determinados por aplicaciones diferenciales los cuales son de alto costo y muy alta rentabilidad, en este tipo de clientes se destacan los laboratorios, industria metrologías, así como los sectores de investigación y desarrollo.

Las ventas más representativas de este clúster están en cilindros de GNV (importados por nuestra compañía y fabricados por nuestras filiales en el mundo), carburo de calcio (sustituto del acetileno en procesos de soldadura, importado por nuestra compañía), y los gases especiales, gases con un mayor porcentaje de pureza en su composición y desarrollados según las especificaciones de la aplicación, la gran mayoría de estos gases son importados y fabricados por nuestras filiales en el mundo, así como la venta de equipos industriales que son adquiridos por intermedio de nuestros proveedores.

Como podemos observar en este clúster se ubican todos los productos que complementan nuestro portafolio de gases del aire, es decir productos que comercializamos y que son fabricados ya sea por nuestras filiales en el mundo y/o terceros, estos productos son los que generan menor rentabilidad a la compañía, pero permiten poder entregar un portafolio completo de productos y servicios que complementan los gases del aire.

Tabla 11. Estrategias de Fidelización Clúster 1 = Reventa Industrial

Estrategia	Fidelización al Segmento Reventa Industrial
Nombre	Big Box
Descripción	Desarrollar tiendas (Walk In) ubicadas en sectores estratégicos de la industria en cada sucursal, para la venta de equipos, accesorios y otros, que nos permitan generar venta de gases a los consumidores de este clúster que están comprando otros productos de nuestro portafolio, sin incluir los gases, buscando captar nuevos clientes y retener los ya existentes.
Como	Creación de una y/o varias plazas en cada sucursal Venta Cruzada
Comunicación	Marketing Mix: Promoción de Ventas – Exhibición -Publicidad
Nombre	Tarjeta Hardgoods
Descripción	Crear la tarjeta de crédito que represente la marca, la cual otorgara descuentos por compras de productos determinados, según promoción realizada mensualmente, y por intermedio de la cual se acumularan puntos, que podrán ser cambiados por hardgoods según catalogo de puntos.
Como	Compras con la tarjeta de crédito, que generan acumulación de puntos y descuentos.
Comunicación	Promoción de Ventas – Programa de Continuidad
Nombre	Producto Sustituto
Descripción	Crear un Brochure con el portafolio de corte y soldadura, donde se resalten los beneficios tanto económicos como de seguridad del uso del acetileno, con el fin de desmotivar el consumo de carburo, enfocándonos en la productividad de los procesos y la seguridad del cliente final
Como	Creación de portafolio especializado
Comunicación	Relaciones Publicas - Eventos - Revista Especializada del Sector Promoción de ventas -Demostraciones Marketing Directo – Entrega del catalogo

Fuente: Las autoras 2013

6.3.2.2 Beneficios Cliente: Contar con un sitio establecido con un portafolio completo para poder realizar las compras según sus necesidades, donde podrá además acumular puntos por sus compras y cambiar por otros productos

necesarios en sus actividades, sustituir productos para una mayor productividad y seguridad en los procesos.

6.3.2.3 Beneficios Compañía: Transformar este clúster, con compras no solo de equipos y accesorios, sino lograr la inclusión del portafolio de gases necesarios en todos y cada uno de sus procesos, mejorando la rentabilidad del clúster, la Fidelización de los clientes ya existentes y la captación de nuevos clientes.

6.3.3 Clúster 2.

Imagen 5. Clúster medicinal

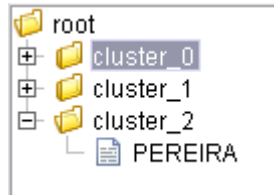


Fuente: Google images, 2013

Nombre: Medicinal

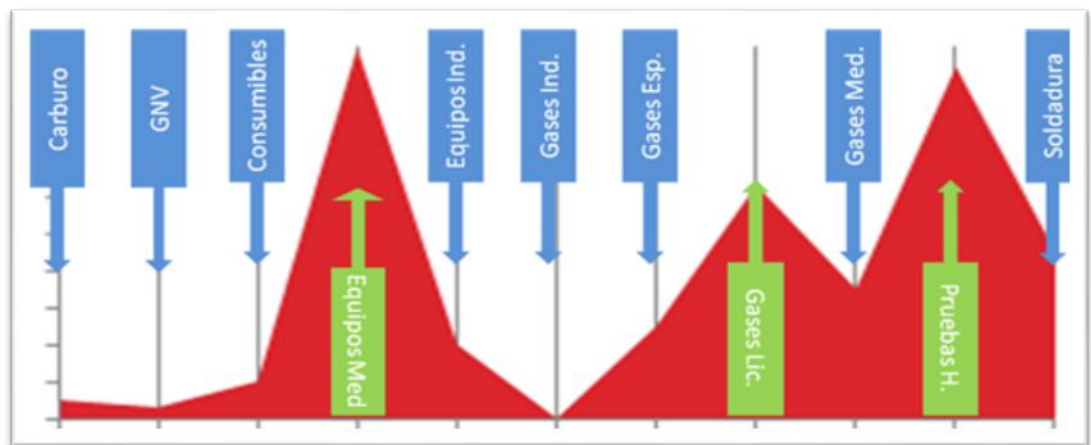
Composición: Compuesto por una sucursal que equivalen al 12% del País.

Gráfico 35. Clúster 2



Fuente: Las autoras 2013

Gráfico 226. Representación del Clúster 2



Fuente: Las autoras 2013

6.3.3.1 Descripción del Clúster: Clientes del sector de la salud, empresas privadas o estatales, donde la mayoría de sus compras son en equipos médicos, pruebas hidrostáticas y gases licuados.

Son empresas que consumen estos productos para la atención de sus pacientes como son las clínicas, centros médicos y de estética.

En este Clúster se evidencia que la mayoría de sus compras van dirigidas al sector de medicinal, donde son ventas representativas ya que están en las principales ciudad de Colombia como son Bogotá, Bucaramanga, Cali y Medellín, donde hay la mayor tecnología en atención en salud.

Este segmento está dirigido en su gran mayoría por profesionales de la salud donde el costo médico le impacta demasiado y buscan el mejor precio, con oportunidad y calidad del producto.

Tabla 12. Estrategias de Fidelización Clúster 2 = Medicinal

Estrategia	Fidelización al Segmento Medicinal
Nombre	Promoción de la semana
Descripción	Catalogo de novedades y productos disponibles, con descuentos que se aplicaran dentro de la semana vigente.
Como	Creación de portafolios especializados de productos y servicios según aplicación.
Comunicación	Marketing directo: Telemarketing - correo electrónico - catálogos
Nombre	Evento Medicinal
Descripción	Se realizaran eventos cada seis meses con el fin de que conozcan los productos ,y por compras en este evento tendrán descuentos del 20% sobre el total de la compra.
Como	Feria semestral
Comunicación	Relaciones publicas: Eventos – Revista en la empresa Ventas personales : Presentaciones de ventas Marketing directo : Catálogos – Correo electrónico Publicidad: Folletos – Afiches Promoción de ventas: Exhibiciones – Demostraciones
Nombre	Cupón Promocional
Descripción	Cupón de descuentos del 10% , por compras superiores a \$ 200.000, donde pueden hacer efectivo el descuento en los 3 meses siguientes
Como	Creación de portafolio de productos y servicios
Comunicación	Promoción de ventas: Cupones Publicidad: Folletos Marketing directo: Correo directo – correo electrónico – Tele marketing

Fuente: Las autoras 2013

6.3.3.2 Beneficios para la compañía: Crecimiento en ventas con mayor rentabilidad, rotación de inventario, producción a gran escala, conocimiento del cliente donde entenderemos las necesidades específicas que tiene, con el fin de poderlas cubrir y que no queden desabastecidos.

6.3.3.3 Beneficios para el cliente: Portafolio de productos para la salud de sus pacientes, garantía de calidad y experiencia, asesoría continua por parte de los

profesionales especializados en cada producto, oportunidad del servicio, soporte técnico permanente en cualquier momento las 24 horas y 365 días del año, descuentos especiales en los productos de mayor consumo.

7 CONCLUSIONES

La Minería de Datos, brinda herramientas prácticas para analizar las bases de datos de los clientes, permite descubrir patrones no obvios, que se pueden traducir o transformar en estrategias.

Los resultados obtenidos durante el desarrollo del trabajo demuestran que una correcta selección y pre-procesamiento de la base de datos, redundan en los resultados óptimos de la segmentación a realizar.

La selección del número de k-mens, clúster o segmentos, es determinante en la definición de las diferentes estrategias, ya que se debe buscar el número óptimo que permita conseguir la mejor distribución y los atributos que representen la tendencia de consumo de nuestros clientes por sucursal a nivel nacional.

La selección del número clúster, define el perfil y la tendencia de consumo de los clientes en las diferentes sucursales a nivel nacional, lo que permite desarrollar estrategias de fidelización que buscan prestar atención a la calidad del servicio que se entrega, definir programas de promoción y abrir canales de comunicación que permitan que el cliente sea una parte activa de nuestro negocio, mejorando la rentabilidad y aportando un mayor sentido a la negociación, creando vínculos prolongados y sostenidos en el tiempo.

Los canales de comunicación que se utilizaron para las estrategias descritas en cada uno de los clúster son los siguientes: Marketing directo, por medio de catálogos, correo directo, telemarketing; Ventas directas, donde se realizaron presentaciones de ventas, reuniones de ventas lo cual permite de forma directa hacerles conocer los productos; Relaciones publicas donde la empresa tiene participación en seminarios, revistas y publicaciones; Promoción de ventas por medio de cupones, sorteos, premios, programas de continuidad; Publicidad por medio de folletos, afiches, anuncios.

RECOMENDACIONES

Es muy importante contar con una base de datos que entregue información veraz, completa y procesable para conocer toda la información de los clientes, para que al momento de enfrentarse a procesar una base de datos la información suministrada cumpla con las especificaciones requeridas para iniciar la técnica de minería de datos.

Durante el desarrollo del trabajo se evidencia que el capítulo más importante es el pre-procesamiento de la base de datos, que determinara el correcto uso del método, la técnica y el algoritmo.

La selección correcta del número de clúster, permite generar estrategias orientadas a la fidelización y retención de nuestros clientes de forma efectiva.

La correcta Mezcla de comunicaciones de marketing va a permitir, informar, convencer y recordar, el portafolio de productos con los clientes.

Extender a mediano plazo el Maxi Marketing a las sucursales donde no hay presencia.

Elaborar frecuentemente planes de mercadeo, con el fin de identificar las nuevas necesidades de los clientes.

Analizar los resultados obtenidos para poder tomar decisiones inmediatas.

INFOGRAFÍA

data mining. (s.f.). Obtenido de data Mining: http://www.sinnexus.com/business_intelligence/datamining.aspx

Fidelizacion de Clientes. (s.f.). Obtenido de Fidelizacion de Clientes: https://www.google.com.co/search?q=rentabilidad&source=Inms&tbm=isch&sa=X&ei=6z1jUqOFKYjs8QSZ64HAAQ&sqi=2&ved=0CAcQ_AUoAQ&biw=1366&bih=651&dpr=1#q=fidelizacion+en+clientes&tbm=isch

Gases Comprimidos. (s.f.). Obtenido de https://www.google.com.co/search?q=gases+comprimidos&hl=es&qscrl=1&rlz=1T4LENP_esCO536CO536&source=Inms&tbm=isch&sa=X&ei=UiZkUryuBY_Q9gTwwYGwAQ&sqi=2&ved=0CAcQ_AUoAQ&biw=1366&bih=618#facrc=_&imgdii=_&imgrc=nfjNxxw9NU9YNpM%3A%3BB5vdYKa4vpzD3M%3Bhttp%253A%25

José M. Gutiérrez. (2100). *Data Mining.Extracción de Conocimiento en Grandes Bases de Datos.* Obtenido de http://personales.unican.es/gutierjm/docs/trans_DataMining.pdf

Market Line. (2012). *Industrial Gases in South America.* Estados Unidos: Market line.

Universidad sergio Arboleda. (s.f.). Semillero de investigacion.