

**DATA MINING Y PROGRAMAS DE FIDELIZACIÓN PARA LOS CLIENTES
DE LA CAJA DE COMPENSACIÓN**

ANA MARÍA DIAZ VARGAS

CC.53.177.403

ANDREA RONCANCIO

CC 52.787.711

**UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD DE INGENIERIA DE MERCADOS
ESPECIALIZACIÓN EN GERENCIA DE MERCADEO ESTRATÉGICO
BOGOTA D.C.
AÑO 2013**

**DATA MINING Y PROGRAMAS DE FIDELIZACIÓN PARA LOS CLIENTES
DE LA CAJA DE COMPENSACIÓN**

ANA MARÍA DIAZ VARGAS

CC.53.177.403

ANDREA RONCANCIO

CC 52.787.711

**Trabajo de investigación en data mining para optar por el título de
especialista en gerencia de mercadeo estratégico**

Director, Liza Pinzón

**UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD DE INGENIERIA DE MERCADOS
GERENCIA DE MERCADEO ESTRATÉGICO
BOGOTA D.C.
AÑO 2013**

Nota de aceptación

Presidente del jurado

Jurado

Jurado

Bogotá DC, 28 de Agosto de 2013

AGRADECIMIENTOS

El presente trabajo fue realizado con la supervisión académica de la docente Liza Pinzón a quien le expresamos nuestra más profunda gratitud por brindarnos la oportunidad de trabajar bajo su supervisión en nuestro proyecto de investigación. Gracias a su apoyo, direccionamiento y constancia se hizo posible la pronta culminación de nuestro trabajo. Sin duda alguna, además de ser nuestra directora de trabajo es una excelente persona, con la experticia en los temas trabajados.

Agradecemos a Dios y a nuestras familias por darnos todo su apoyo y colaboración durante este proceso.

TABLA DE CONTENIDO

INTRODUCCION	6
1. COMPRESION DEL NEGOCIO	6
1.1. CAJAS DE COMPENSACIÓN COLOMBIA.....	6
2. OBJETIVO GENERAL.....	14
2.1. OBJETIVOS ESPECIFICOS	14
3. METODOLOGÍA.....	15
3.1. RECOLECCION DE DATOS.....	15
3.1.1. Solicitud de información al Cliente	15
3.1.2. Entrega de información del Cliente.....	15
3.1.3. Pre procesamiento de la data	15
4. MODELO	17
4.1. MINERIA DE DATOS	17
4.2. METODO NO SUPERVISADO	18
4.3. ALGORITMO	18
4.4. ALGORITMO A PRIORI	19
4.5. ALGORITMO K-MEANS	19
4.6. RAPID MINER	20
4.7. MINERIA DE DATOS	20
4.7.1. Clúster o segmentación.....	21
4.7.2. Asociación	23
4.8. DESCRIPCION DE SEGMENTOS.....	24
4.7.1. CI24	
4.9. ASSOCIATION RULES.....	28
4.9.1. Análisis 1: [FUTBOL] --> [TURCO_Y_SAUNA] (confidence: 0.994)	29
4.9.2. Análisis 2: [TENIS_DE_CAMPO] --> [TURCO_Y_SAUNA] (confidence: 0.950)	29
4.9.3. Análisis 3: [NATACION] --> [TURCO_Y_SAUNA] (confidence: 0.946)	29
5. ESTRATEGIAS (CONCLUSIONES).....	30
ESTRATEGIA GENERAL.....	30
5.1. SEGMENTO 1: LA PENSIONADA.....	30
5.1.1. OBJETIVO	30
5.1.2. ESTRATEGIA	30
5.1.3. TACTICAS	31
5.1.4. PLAN DE MEDIOS.....	32
5.2. SEGMENTO 2 : EL YUPPIE.....	33
5.2.1. OBJETIVO	33
5.2.2. ESTRATEGIA	33

5.2.3.	TACTICAS.....	33
5.2.4.	PLAN DE MEDIOS.....	35
5.3.	SEGMENTO 3 : LA FAMILIA GONZALEZ.....	35
5.3.1.	OBJETIVO.....	35
5.3.2.	ESTRATEGIA.....	36
5.3.3.	TACTICAS.....	36
5.3.4.	PLAN DE MEDIOS.....	37
5.4.	SEGMENTO 4: EL ADOLECENTE.....	38
5.4.1.	OBJETIVO.....	38
5.4.2.	ESTRATEGIA.....	38
5.4.3.	TACTICAS.....	38
5.4.4.	PLAN DE MEDIOS.....	39
6.	RECOMENDACIONES.....	41
7.	BIBLIOGRAFÍA.....	42

INTRODUCCION

El siguiente trabajo tiene como objetivo comprender los modelos y sus técnicas para desarrollar la minería de datos la cual permite concentrarse en lo más importante de las bases de datos y entregar a la organización información de tipo prospectiva y proactiva que contribuya a la toma de decisiones o a la predicción de tendencias o comportamientos de los consumidores.

Posteriormente se realizaran y analizaran los pasos a seguir para la recolección de datos, pre procesamiento de la data, la minería de datos y los resultados y análisis para el desarrollo de estrategias de marketing y programas de fidelización.

A continuación, se toma una base de datos recibida de una caja de compensación familiar del municipio de Cundinamarca, de la cual no daremos a conocer su nombre por seguridad, para la elaboración del procesamiento de la minería de datos en la cual se seleccionó el modelo No Supervisado bajo las técnicas de Asociación y Agrupación haciendo uso de los algoritmos apriori y k- means respectivamente, con sus concernientes parámetros C y K, los cuales explicaremos en el transcurso de este trabajo.

Finalmente, se identificarán los diferentes segmentos de clientes y las correlaciones de los servicios de acuerdo a los resultados obtenidos a través del aplicativo Rapid Miner, lo cual permitirá plantear las diferentes estrategias para presentárselas a la caja de compensación familiar para la cual se desarrolla la siguiente investigación.

1. COMPRESION DEL NEGOCIO

1.1. CAJAS DE COMPENSACIÓN COLOMBIA

El objetivo principal de las cajas de compensación familiar es mejorar la calidad de vida de los trabajadores Colombianos y la población más vulnerable ofreciendo diferentes servicios en áreas relacionadas con capacitación, recreación y turismo, estas fueron creadas hace cerca de 50 años y desde entonces han ido ampliando sus servicios buscando ofrecer bienestar a sus afiliados.

Las cajas de compensación son las únicas instituciones en Colombia de gobierno bipartito (Trabajadores y empresarios), estas tienen enfoque social, son sin ánimo de lucro y están vigiladas por la superintendencia del subsidio familiar. (ASOCAJAS, 2012)

“La ley 789 del 2002 reorganizó el funcionamiento de las cajas de compensación donde determino alguna de las actividades de estas entidades:

- Administrar programas de subsidios en dinero o especie, salud, vivienda, entre otros.
- Ejecutar actividades relacionadas con sus servicios, la protección y la seguridad social.
- Realizar actividades de mercadeo, incluida la administración de farmacias con independencia financiera.

Sin embargo, el gobierno nacional determina los campos en los cuales las cajas de compensación familiar pueden invertir”. (ELEMPLEO.COM)

Actualmente, las empresas son las que deciden a que caja de compensación familiar afiliar a sus empleados, sin embargo, los afiliados y sus familiares no utilizan la caja de compensación

muchas veces por desconocimiento y falta de información sobre los programas y servicios ofrecidos por estas. (ELEMPLERO.COM), Entre los servicios ofrecidos por las cajas de compensación encontramos créditos, turismo, capacitaciones, actividades recreativas y deportivas, cursos que se enfocan en orientación a la competitividad, cursos para bebés en donde ellos pueden desarrollar su motricidad en las diferentes etapas, todas estas actividades y cursos están acompañados de seguimiento personalizado con profesionales expertos y especialistas en cada área.

Las actividades de bienestar de las cajas de compensación siempre trabajan en busca de una proyección y crecimiento del ser humano, adopción de hábitos saludables mediante procesos integrales de formación para mejorar el desempeño empresarial, familiar, académico, social y deportivo de la sociedad en general.

Para todos los servicios que prestan las cajas de compensación ya sea en recreación, deporte, cultura, capacitación, entrenamiento, salud, subsidios, vivienda, alojamiento en centros vacacionales, entre otros servicios, pueden ser accedidos por cualquier persona “Afiliado o No Afiliado”, lo único que difiere es: Para los afiliados existen unas categorías A, B o C, las cuales son establecidas de acuerdo al ingreso del cotizante; para los No Afiliados se les aplicará una tarifa para la toma de servicio; sin embargo se estipula que el servicio para uno o para el otro debe ser transparente y con excelente calidad.

“En cuanto a los servicios más importantes, se destacan la Recreación y el Turismo Social los cuales se han caracterizado e identificado ampliamente a las Cajas de Compensación Familiar entre todos los Colombianos. Por tal razón, no es de extrañar que sean los de más amplia cobertura y utilización dentro del mercado. Actualmente los centros vacacionales de las cajas de compensación cuentan con modernos centros vacacionales y centros recreativos que le permiten a la familia tanto del trabajador, beneficiario y no cotizante, disponer de las

mejores alternativas para un día de esparcimiento, un fin de semana de descanso o unas completas vacaciones, con todo al alcance de sus posibilidades y con precios accesibles.”.(SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR)

“ Para la población adulta, especialmente los Pensionados cuya mesada sea de \$884.000 pesos o menos, podrán acceder totalmente gratis a los servicios que sobre cultura, recreación y deporte ofrecen las cajas de compensación familiar. Según cifras del Ministerio de Trabajo, en el país existen 1'650.000 pensionados, y de esta cifra, de acuerdo con los reportes de la Asociación Nacional de Cajas de Compensación Familiar (Asocajas), solo 35.198 personas acceden a los beneficios de las cajas, a hora bien, el gobierno pide a los pensionados afiliarse a las cajas cumpliendo el requisito de acreditar una permanencia de 20 años en una o en varias cajas de compensación familiar. Este periodo, según la norma aprobada en el Congreso, puede ser "continuo o discontinuo".”(REDACCION POLITICA, 2013)

“Al igual, el gobierno ha pensado en otra población para su afiliación como los son los trabajadores del servicio doméstico, los cuales tendrán que ser afiliados por sus patronos a las cajas de compensación familiar, según lo dispuesto en el Decreto 721 expedido por el Ministerio del Trabajo, que explicó que con la afiliación este grupo de personas podrá acceder a beneficios como programas de educación, recreación y subsidio familiar. Sin embargo, la norma establece como requisito esencial laborar al menos 96 horas al mes. Además de las empleadas de servicio, se consideran trabajadores domésticos los jardineros, conductores particulares y cuidanderos, entre otros”.(REDACCIÓN ECONOMÍA Y NEGOCIOS, 2013)

Estos nuevos decretos ayudan a que aumente el número de afiliados a las cajas de compensación, y que desaparezca un poco más la informalidad del trabajo, de esta manera se les pueden ofrecer servicios a más personas mejorando así su calidad de vida y su bienestar

al poder acceder a actividades deportivas, créditos de vivienda, turismo, capacitaciones entre otros servicios.

A continuación, se podrá observar dos tablas con el comparativo de afiliados y su participación en las cajas de compensación a nivel nacional y en las diferentes regiones.

TABLA 1. Comparativo de afiliados a las cajas de compensación a nivel nacional a marzo de 2011, 2012 y 2013

**TRABAJADORES AFILIADOS POR REGIONES
COMPARATIVO A MARZO 2013/2012/2011**

GRUPO CAJAS				TRABAJADORES AFILIADOS (1)			Participación % sobre el total		
No.	REGION	Cod	Nombre Caja	A	A	A	A	A	A
				MARZO/2013	MARZO/2012	MARZO/2011	MARZO/2013	MARZO/2012	MARZO/2011
6	BOGOTA D.C.	18	C.C.F. AFIDRO - BOGOTA (2)	-	-	-	-	-	-
		21	CAFAM	688.167	698.410	633.550	10,14	10,62	10,31
		22	COLSUBSIDIO	956.442	897.533	816.845	14,09	13,65	13,29
		24	COMPENSAR	804.981	891.675	823.646	11,86	13,56	13,40
		26	COMFACUNDI	37.984	34.447	39.730	0,56	0,52	0,65
		SUBTOTAL				2.487.574	2.522.065	2.313.771	36,66
ORDEN NACIONAL		68	COMCAJA	7.929	7.854	10.292	0,12	0,12	0,17
SUBTOTAL				2.495.503	2.529.919	2.324.063	36,78	38,48	37,81
TOTAL REGIONES 1 A 5	*	Subtotal resto cajas		4.290.188	4.044.298	3.822.335	63,22	61,52	62,19
GRAN TOTAL				6.785.691	6.574.217	6.146.398	100,00	100,00	100,00

FUENTE: Informes Estadísticos Mensuales Marzo 2013/2012/2011 - Superintendencia del Subsidio Familiar-Oficina de Planeación-Grupo de Estadística.

(1) Incluye:
- Salarios
- Pensión
- Cajas de Compensación

(2) Caja fusionada con Colsubsidio

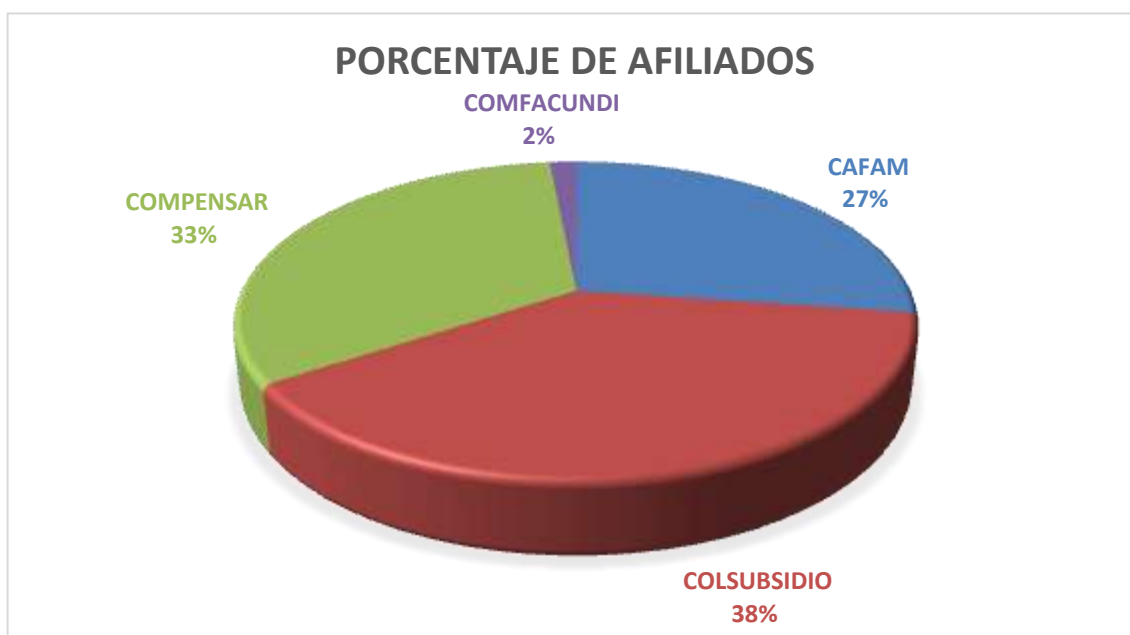
FINANCIACION
CONSOLIDADO PRIMER
TRIMESTRE DE 2013

TABLA 2 Participación de No de Afiliados a las cajas de compensación Familiar en el Departamento de Cundinamarca.

Nombre	Participación Afiliados	Número de Afiliados
CAFAM	27%	696.459
COLSUBSIDIO	39%	985.913
COMPENSAR	33%	839.087
COMFACUNDI	2%	38.704
	100%	2.560.163

FUENTE: SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR 2013

GRAFICA 1 Porcentaje de afiliados a las cajas de compensación familiar en el Departamento de Cundinamarca.



FUENTE: SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR 2013

Se puede observar en la anterior gráfica que la participación del número de afiliados las cajas de compensación en Cundinamarca están distribuidas así: Colsubsidio con un 38% teniendo

la mayor participación, Compensar con un 33%, Cafam con un 27% y Comfacundi con un 2% perdiendo su participación.

TABLA 3 Participación de las Cajas de Compensación Familiar en el mercado.

Nombre CAJA	PARTICIPACIÓN	EMPRESAS AFILIADAS
CAFAM	27%	30.667
COLSUBSIDIO	37%	41.321
COMPENSAR	34%	38.775
COMFACUNDI	2%	2.007
	100%	112.770

FUENTE: SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR 2013

GRAFICA 2 Participación de las cajas de compensación familiar en el mercado.



FUENTE: SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR 2013

Colsubsidio y Cafam fueron fundadas en 1957 siendo las más antiguas, Comfacundi fundada en 1974 y Compensar fue fundada en 1978 convirtiéndose en la caja más joven y con la segunda participación más alta de cajas de compensación en Cundinamarca.

TABLA 4 Comparativo de afiliados a las cajas de compensación a nivel regional a marzo de 2011, 2012 y 2013

**TRABAJADORES AFILIADOS POR REGIONES
COMPARATIVO A MARZO 2013/2012/2011**

GRUPO CAJAS		TRABAJADORES AFILIADOS (1)			Participación % sobre el total				
No.	REGION	Cod	Nombre Caja	A MARZO/2013	A MARZO/2012	A MARZO/2011	A MARZO/2013	A MARZO/2012	A MARZO/2011
1	COSTA ATLANTICA	5	CAJACOPI	34.298	32.110	31.378	0,51	0,49	0,51
		6	COMBARRANQUILLA	96.693	90.448	86.077	1,42	1,38	1,40
		7	COMFAMILIAR ATLANTICO	173.019	161.742	153.192	2,55	2,46	2,49
		8	COMFENALCO CARTAGENA	158.398	141.647	125.256	2,33	2,15	2,04
		9	COMFAMILIAR CARTAGENA	45.023	54.059	52.460	0,66	0,82	0,85
		15	COMFACESAR	79.443	84.230	83.098	1,17	1,28	1,35
		16	COMFACOR	81.735	59.244	55.826	1,20	0,90	0,91
		30	COMFAGUAJIRA	44.879	39.359	37.135	0,66	0,60	0,60
		33	CAJAMAG	83.351	74.996	73.638	1,23	1,14	1,20
		41	COMFASUCRE	42.213	37.469	36.480	0,62	0,57	0,59
		64	CAJASAI	13.424	13.079	12.404	0,20	0,20	0,20
		SUBTOTAL	852.476	788.383	746.944	12,56	11,99	12,15	
2	REGION OCCIDENTE	2	CAMACOL	18.731	21.208	32.810	0,28	0,32	0,53
		3	COMFENALCO ANTIOQUIA	317.170	316.326	294.370	4,67	4,81	4,79
		4	COMFAMA	769.410	698.331	657.428	11,34	10,62	10,70
		11	CONFAMILIARES CALDAS	109.863	107.345	101.173	1,62	1,63	1,65
		12	C.C.F. LA DORADA	-	-	-	-	-	-
		14	COMFACAUCA	73.609	69.162	63.045	1,08	1,05	1,03
		29	COMFACHOCO	15.724	14.861	9.427	0,23	0,23	0,15
		35	COMFAMILIAR NARIÑO	68.984	61.612	64.917	1,02	0,94	1,06
		42	COMFAMILIAR DEL QUINDIO (-	-	-	-	-	-
		43	COMFENALCO QUINDIO	57.239	54.217	53.354	0,84	0,82	0,87
		44	COMFAMILIAR RISARALDA	123.425	121.106	121.302	1,82	1,84	1,97
		51	COMFAMAR - BUENAVENTUR	-	-	-	-	-	-
		56	COMFENALCO VALLE	211.222	213.075	204.220	3,11	3,24	3,32
		57	COMFANDI	537.769	526.294	499.769	7,93	8,01	8,13
		59	COMFAMILIAR CARTAGO (4)	-	-	-	-	-	-
61	COMFAUNION PALMIRA (4)	-	-	-	-	-	-		
62	COMFAMILIAR TULUA (4)	-	-	-	-	-	-		
		SUBTOTAL	2.303.146	2.203.537	2.101.815	33,94	33,52	34,20	
3	REGION CENTRO ORIENTE	10	COMFABOY	119.816	115.203	106.167	1,77	1,75	1,73
		32	COMFAMILIAR HUILA	121.540	86.526	76.304	1,79	1,32	1,24
		36	COMFAORIENTE	58.523	58.807	58.268	0,86	0,89	0,95
		37	COMFANORTE	78.240	73.766	77.330	1,15	1,12	1,26
		38	CAFABA	40.218	38.511	51.000	0,59	0,59	0,83
		39	CAJASAN	171.938	168.918	162.775	2,53	2,57	2,65
		40	COMFENALCO SANTANDER	158.196	141.178	126.697	2,33	2,15	2,06
		46	CAFASUR	6.332	5.898	5.350	0,09	0,09	0,09
		47	COMFAMINORTE - HONDA (5)	-	-	-	-	-	-
		48	COMFATOLIMA	40.513	42.012	40.431	0,60	0,64	0,66
50	COMFENALCO TOLIMA	67.209	72.684	59.892	0,99	1,11	0,97		
		SUBTOTAL	862.525	803.503	764.214	12,71	12,22	12,43	
4	REGION ORINOQUIA	34	COFREM	102.204	88.396	78.415	1,51	1,34	1,28
		67	COMFIAR	12.658	12.600	11.493	0,19	0,19	0,19
		69	COMFACASANARE	114.947	106.199	79.466	1,69	1,62	1,29
			SUBTOTAL	229.809	207.195	169.374	3,39	3,15	2,76
5	REGION AMAZONIA	13	COMFACA	20.732	21.642	21.240	0,31	0,33	0,35
		63	COMFAPUTUMAYO	17.186	15.805	14.831	0,25	0,24	0,24
		65	CAFAMAZ	4.314	4.233	3.917	0,06	0,06	0,06
			SUBTOTAL	42.232	41.680	39.988	0,62	0,63	0,65
BOGOTA D.C. Y	*	Cinco cajas (ver siguiente c	2.495.503	2.529.919	2.324.063	36,78	38,48	37,81	
GRAN TOTAL			6.785.691	6.574.217	6.146.398	100,00	100,00	100,00	

FUENTE: Informes Estadísticos Mensuales Marzo 2013/2012/2011 - Superintendencia del Subsidio Familiar-Oficina de Planeación-Grupo de Estadística.

2. OBJETIVO GENERAL

Identificar los segmentos de clientes afiliados (trabajador¹ y beneficiario²) y no afiliados³ a la caja de compensación de la ciudad de Bogotá D.C., como la interrelación de los servicios deportivos, con el fin, de conocer su comportamiento de consumo.

2.1. OBJETIVOS ESPECIFICOS

- Identificar la correlación de los servicios complementarios, con el fin de establecer una estrategia de cross-selling con el portafolio de servicios deportivos de la caja de compensación.
- Desarrollar estrategias de promoción y programas de fidelización para captar y conservar los clientes, afiliados y no afiliados a la caja de compensación, con el fin de generar un posicionamiento de marca de la misma en Cundinamarca.

¹ Cotizante o persona que paga una cuota mensual o periódica para acceder a los servicios de una caja de compensación.

² Persona afiliada por el cotizante para que pueda acceder a los servicios de la misma caja de compensación.

³ Personas cotizantes y afiliadas a otra caja de compensación pero que pueden ser usuarios de ésta.

3. METODOLOGÍA

3.1. RECOLECCION DE DATOS

3.1.1. Solicitud de información al Cliente

Se solicita a la caja de compensación una base de datos de la unidad de negocios de deportes, la cual deberá contener los afiliados y no afiliados que en el transcurso del año 2013 hayan tomado los servicios. La base deberá tener los campos de Identificación (ID), datos socio demográficos, servicios tomados, y demás información importante para el desarrollo de esta investigación.

3.1.2. Entrega de información del Cliente

El cliente entrega una data en formato Excel con 300.000 usuarios, con los campos de Nombres, Apellidos, Edad, Categoría de Servicio, Línea de Servicio, Horario de toma del servicio; esta base fue extraída del Data Marts de la Organización.

3.1.3. Pre procesamiento de la data

En este paso se seleccionan, limpian y transforman los datos.

Recibida la base de datos en un archivo de Excel se procede a filtrar por ID y línea de Servicio, con el fin de eliminar los campos sin datos, al realizar este ejercicio se deja una data de 100.186 filas la cual se procede a convertir a formato del operador de la siguiente forma:

En el archivo de Excel se abren las siguientes hojas de cálculo:

Hoja de cálculo (1): En esta hoja se pega en valores la data de 100.186.

Hoja de cálculo (2): Se realiza una tabla dinámica donde se cruzó los ID con la variable Líneas de servicio.

Hoja de cálculo (3): En los campos desocupados se adiciona un cero (0).

Hoja de cálculo (4): Para hallar los porcentajes de cada campo, por fila se toma el valor de cada campo y se divide por el total general de la tabla dinámica por ID.

*Para trabajar clúster o segmentación: Se dejan los ceros y los números decimales igual.

*Para trabajar Asociación: Los números decimales se cambian por 1 y los ceros quedan igual.

Hoja de cálculo (5): Todos los datos se copian y se pegan en valores, a su vez, se remplazan las comas por puntos. En la columna final se concatena haciendo uso de la siguiente fórmula:
(=CONCATENAR(A1;" ";B1;" ";C1;" ";D1;" ";E1;" ";F1;" ";G1;" ";H1;" ";I1;" ";J1;" ";K1;" ";L1;" ";M1;" ";N1;" ";O1;" ";P1))

Hoja de cálculo (6): Se pegan en valores la columna que se concateno, se revisa que no haya quedado puntos.

Finalizado este esquema, la hoja de cálculo 6 se guarda como nuevo archivo en formato CSV (Delimitado por comas).

De esta forma el archivo ha quedado listo para iniciar el proceso de minería de datos.

4. MODELO

4.1. MINERIA DE DATOS

El data mining no es un software, pero si reúne las ventajas de varias áreas como la Estadística, Inteligencia Artificial y la Computación Gráfica, se presenta como una tecnología emergente que genera un punto de encuentro entre los investigadores y las personas de negocios ya que permite concentrarse en lo más importante de las datas y entregar información prospectiva, proactiva y conducida por un conocimiento. Los resultados de data Mining contribuye a la toma de decisiones o a la predicción de tendencias o comportamientos (VALLEJOS, 2006)

Las herramientas de data Mining pueden responder a preguntas de negocios basadas en la exploración de las grandes bases de datos en busca de patrones ocultos, encontrando información predecible que un experto no puede llegar a encontrar porque se encuentra fuera de sus expectativas. (Carne, 2007)

El data mining se apoya del proceso de conocimiento, el cual consiste en usar métodos de minería de datos (algoritmos) para identificar parámetros usando una base de datos junto con pre procesamientos y post-procesamientos, al igual, la minería de datos es una técnica que aplica la inteligencia artificial para encontrar patrones y relaciones dentro de los datos permitiendola creación de modelos.

La minería de datos produce cinco tipos de información:

- Asociaciones.
- Secuencias.
- Clasificaciones.
- Agrupamientos.
- Pronósticos

(VALLEJOS, 2006)

4.2. METODO NO SUPERVISADO

“En los métodos no supervisados no se intentan alcanzar un objetivo en particular, sino que el algoritmo busca patrones y estructuras ocultas, no conocidas de antemano, entre las variables. Se distingue del Aprendizaje supervisado por el hecho de que no hay un conocimiento a priori. El aprendizaje no supervisado típicamente trata los objetos de entrada como un conjunto de variables aleatorias, siendo construido un modelo de densidad para el conjunto de datos.

Los resultados deben ser luego analizados para darles una interpretación. Junto con los métodos supervisados, son métodos de la minería de datos”. (JOJOOA- Tecnología, Marketing y CRM)

4.3. ALGORITMO

El algoritmo de minería de datos es el mecanismo que crea un modelo de minería de datos.

Para crear un modelo, un algoritmo analiza primero un conjunto de datos y luego busca patrones y tendencias específicos. El algoritmo utiliza los resultados de este análisis para definir los parámetros del modelo de minería de datos, estos parámetros se aplican en todo el conjunto de datos para extraer patrones procesables y estadísticas detalladas.

- **Algoritmos de segmentación**, que dividen los datos en grupos, o clústeres, de elementos que tienen propiedades similares.
- **Algoritmos de asociación**, que buscan correlaciones entre diferentes atributos de un conjunto de datos. La aplicación más común de esta clase de algoritmo es la creación de reglas de asociación, que pueden usarse en un análisis de la cesta de compra.
(Microsoft SQL Server)

4.4. ALGORITMO A PRIORI

El algoritmo apriori se usa en minería de datos para encontrar reglas de asociación en un conjunto de datos. Este algoritmo se basa en el conocimiento previo o “a priori” de los conjuntos frecuentes, esto sirve para reducir el espacio de búsqueda y aumentar la eficiencia.
(Wikipedia, 2013)

4.5. ALGORITMO K-MEANS

En minería de datos, k -means es un método de agrupamiento, que tiene como objetivo la] n en k grupos en el que cada observación pertenece al grupo más cercano a la media, los grupos

del datos son recogidos en "bloques" (clusters) separados por características comunes.

(Wikipedia, 2013)

4.6. RAPID MINER

Rapid Miner es un software desarrollado por Rapid-I, empresa basada en Dortmund, Alemania, este programa permite el desarrollo de procesos de análisis de datos mediante operadores a través de un entorno gráfico. Se usa en investigación educación, capacitación, creación rápida de prototipos y en aplicaciones empresariales.

La versión inicial fue desarrollada por el departamento de inteligencia artificial de la Universidad de Dortmund en 2001. Se distribuye bajo licencia AGPL y está hospedado en SourceForge desde el 2004.

Rapid Miner proporciona más de 500 operadores orientados al análisis de datos, incluyendo los necesarios para realizar operaciones de entrada y salida, pre procesamiento de datos y visualización. (Wikipedia, 2013)

Para la minería de datos se hizo uso del programa Rapid Miner 4.4 en el cual se modeló las técnicas de Clúster y Asociación para la investigación a continuación.

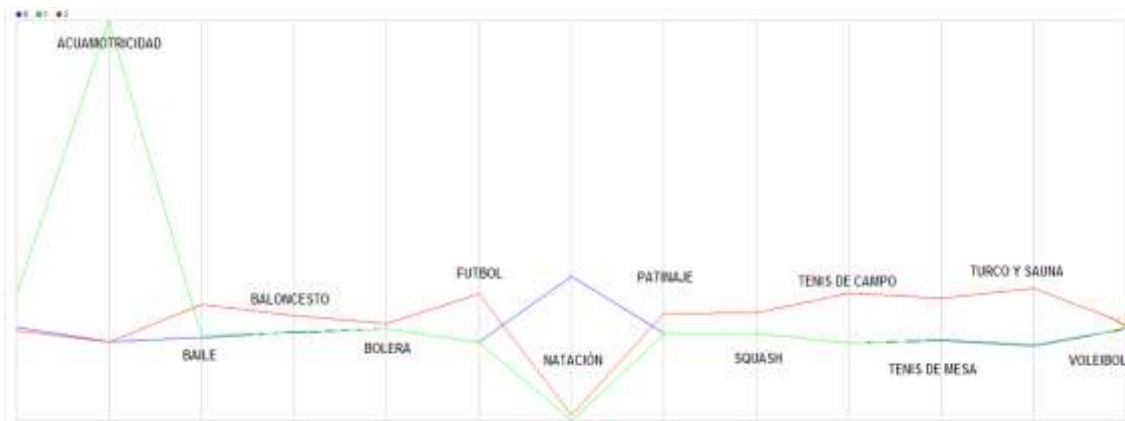
4.7. MINERIA DE DATOS

4.7.1. Clúster o segmentación

Se hace uso de la técnica de segmentación para agrupar en subconjuntos lo más homogéneos posibles a los clientes de la unidad de negocio de acuerdo a comportamientos homogéneos, con el objetivo de maximizar la similitud de individuos del clúster y maximizar las diferencias entre los mismos. (Vallejos)

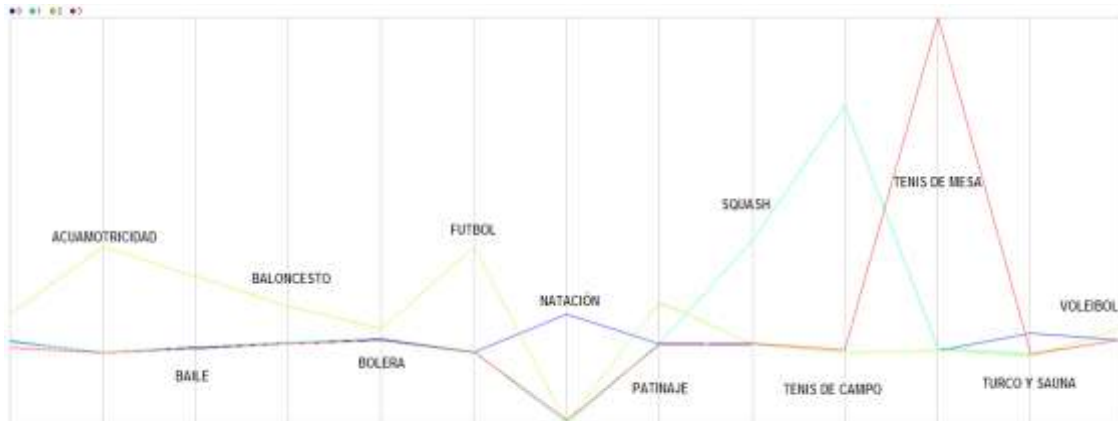
Para identificar el número de parámetro K, se realizó inicialmente el proceso con un K 3, en el cual no se identificó claramente las diferentes tendencias de segmentos, por lo anterior, se realizó nuevamente el ejercicio con K4, el cual arrojó con mayor claridad la tendencia del consumo de los servicios en cada uno de los diferentes segmentos.

GRAFICA 3 K3



FUENTE: ELABORACIÓN PROPIA

GRAFICA 4 K4



FUENTE: ELABORACIÓN PROPIA

Se estable un modelo de trabajo de la siguiente forma:

Método: No supervisado

Técnica: Agrupación

Algoritmo: K- Means

Parámetro: K4

4.6.1.1. Ruta de Proceso:

- 1) New
- 2) Root Process---New Operator----IO---Examples---CSVExampleSource---(Se carga la data en formato CSV (Delimitado por comas).
- 3) CSVExampleSource--- Id_Column: 1 ó id_name: ID
- 4) Root Process—New Operator---Preprocessing---Normalización
- 5) Root Process---Leamer---Unsupervised---Clustering---KMeans
- 6) K- Means---K4

7) Play

4.7.2. Asociación

Se hace uso de la técnica de asociación con el fin de encontrar las reglas que determinan la correlación entre los servicios de un portafolio de servicios, para este caso serán los servicios deportivos para los clientes de la caja de compensación. (Vallejos)

Se establece un modelo de trabajo de la siguiente forma:

Método: No supervisado

Técnica: Asociación

Algoritmo: Apriori

Parámetro: C8

4.6.2.1. Ruta de Proceso:

1) New

2) RootProcess---New Operator----IO---Examples---CSVExampleSource---(Se carga la data en formato CSV (Delimitado por comas).

3) Root Process---New Operator---Preprocessing---Atributes---Filter---Converter---Numerical---2Binomial

4) Root Process---New Operator---Leamer---unsupervised---Itemsets---Fpgrowth---AssociationRuleGenerator

5) AssociationRuleGenerator---min_confidence (establecer el nivel de confianza de 1 a 10)

6)Min_confidence: **C8**

7) Play

4.8. DESCRIPCION DE SEGMENTOS

4.7.1. CLUSTER MODEL

Cluster 0: 77689 items

Cluster 1: 5732 items

Cluster 2: 13153 items

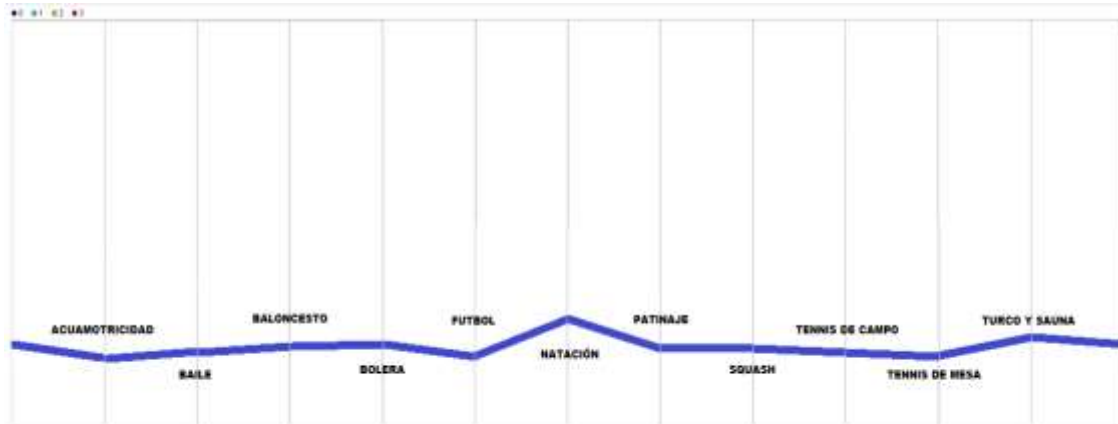
Cluster 3: 3612 items

Total number of items: 100186

4.7.1.1. SEGMENTO 1

Del total de la población 100.186, el segmento 1 corresponde a 77.689 personas equivalentes al 78%, las cuales tienen características homogéneas como son el uso de los servicios de la piscina para natación y de las zonas húmedas para turco y sauna.

GRAFICA 5 LA PENSIONADA



FUENTE: ELABORACIÓN PROPIA

LA PENSIONADA

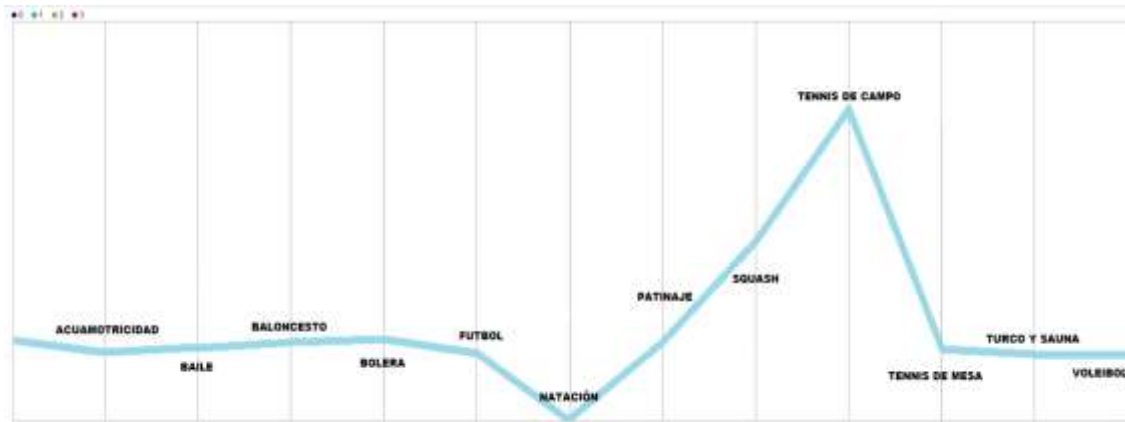
La pensionada vive en la Castellana, actualmente vive sola ya que hace poco tiempo quedo viuda, tiene 52 años y hace 2 años se pensiono de la empresa donde trabajaba como Gerente de Recursos Humanos, su ingreso mensual es superior a los 4SMLV.

Se ha apasionado por su bienestar, por eso le encanta ir entre semana a su caja de compensación más cercana y disfrutar de una rutina en las zonas húmedas, inicia por turco y sauna y finaliza con una hora de natación. Los fines de semana, le gusta compartir con sus sobrinos y hermanas en la bolera.

4.7.1.2. SEGMENTO 2

Del total de la población 100.186, el segmento 2 corresponde a 5.732 personas equivalentes al 6%, las cuales tienen tendencias de consumo similares como es el uso de los servicios la cancha de tenis y squash.

GRAFICA 6 EL YUPPIE



FUENTE: ELABORACIÓN PROPIA

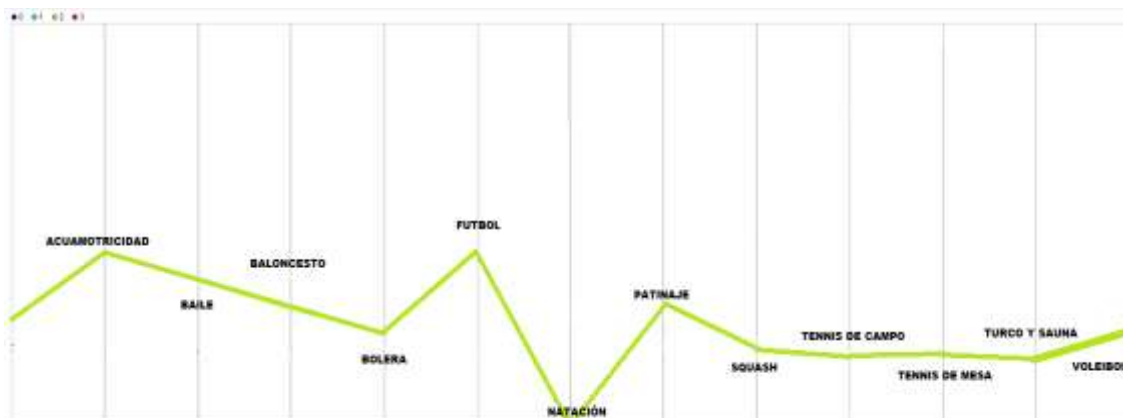
EL YUPPIE

Actualmente vive en aparta estudio en Chapinero, trabaja como jefe de investigación en una pequeña empresa donde tiene un salario superior a los 3SMLV. El mejor plan para él es salir de su trabajo y reunirse con sus amigos a jugar squash, los fines de semana le gusta invitar a sus amigos a salir de Bogotá disfrutar de las canchas de polvo de ladrillo y jugar un partido de tenis y de vez en cuando unas líneas de bolos.

4.7.1.3. SEGMENTO 3

Del total de la población 100.186, el segmento 3 corresponde a 13.153 personas equivalentes al 13%, las cuales tienen diversas tendencias de consumo en los servicios, por lo cual se estable un grupo familiar a quienes le gusta hacer uso de los siguientes servicios Acuamotricidad, Fútbol y Patinaje.

GRAFICA 7 LA FAMILIA GONZALEZ



FUENTE: ELABORACIÓN PROPIA

LA FAMILIA GONZALEZ

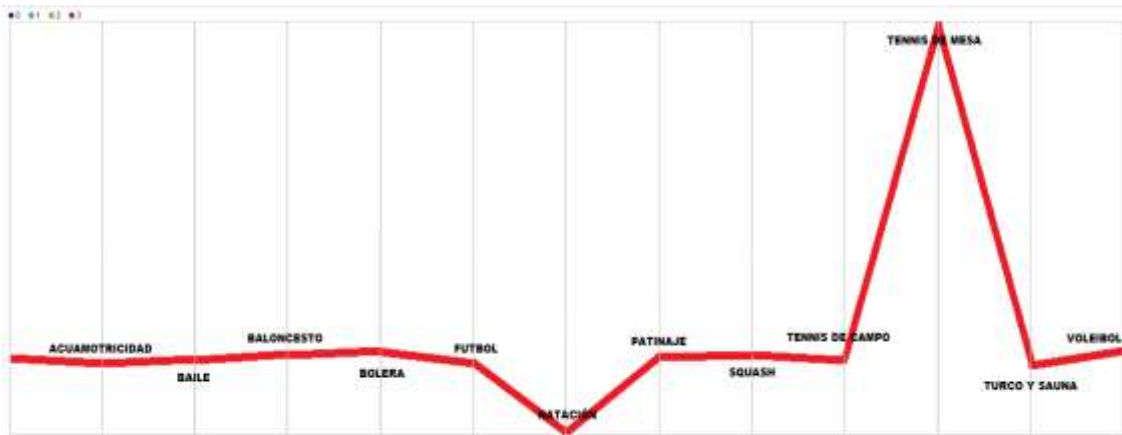
La familia González está compuesta por Ana y Gabriel esposos hace 17 años y viven en Villa Luz con sus tres hijos, Andrés Camilo de 18 meses, Juan David de 10 años y Mariana de 15 años, ambos son empleados, sus ingreso familiar mensual es mayor de los 2SMLV.

Para Ana y Gabriel es muy importante compartir con sus hijos una actividad familiar los fines de semana, por ello inscribieron a sus hijos en los diferentes cursos: Ana acompaña y Andrés Camilo ingresan a Acua-motricidad, Gabriel acompaña a Juan David a su clase de Fútbol y Mariana por ser la mayor, está en un curso de patinaje.

4.7.1.4. SEGMENTO 4

Del total de la población 100.186, el segmento 4 corresponde a 3.612 personas equivalentes al 4%, los cuales tienen una tendencia de consumo así los deportes de tenis de mesa, voleibol o bolos.

GRAFICA 8 EL ADOLECENTE



FUENTE: ELABORACIÓN PROPIA

EL ADOLECENTE

El adolescente tiene 16 años, vive actualmente con su mami quien trabaja en una plata de producción en San Carlos, ella tiene un ingreso mensual de 1SMLV.

Los fines de semana se encuentra con su papa quién lo lleva a la caja de compensación a disfrutar un día agradable y poder compartir un juego de tenis de mesa, si van acompañados de sus primas disfrutan de una línea de bolos o un pequeño partido de voleibol.

4.9. ASSOCIATION RULES

Inicialmente se realizo el proceso con un nivel de confianza de 0.5, buscando un mayor número de correlaciones, pero se encontraron las mismas al realizarlo con un nivel de confianza de 0.8.

NIVEL DE CONFIANZA 0.8

[NATAACION] --> [TURCO_Y_SAUNA] (confidence: 0.946)

[TENIS_DE_CAMPO] --> [TURCO_Y_SAUNA] (confidence: 0.950)

[FUTBOL] --> [TURCO_Y_SAUNA] (confidence: 0.994)

4.9.1. Análisis 1: [FUTBOL] --> [TURCO_Y_SAUNA] (confidence: 0.994)

Con un nivel de confianza del 8%, existe una probabilidad de que el 99.4% de que los afiliados y no afiliados a la caja, hagan uso del servicio de Turco y Sauna ya que han hecho uso inicialmente del servicio de la Cancha de Futbol.

4.9.2. Análisis 2: [TENIS_DE_CAMPO] --> [TURCO_Y_SAUNA] (confidence: 0.950)

Con un nivel de confianza del 8%, existe una probabilidad de que el 95% de que los afiliados y no afiliados a la caja, hagan uso del servicio de Turco y Sauna ya que han hecho uso inicialmente del servicio de la Tenis de campo.

4.9.3. Análisis 3: [NATAACION] --> [TURCO_Y_SAUNA] (confidence: 0.946)

Con un nivel de confianza del 8%, existe una probabilidad de que el 94.6% de que los afiliados y no afiliados a la caja, hagan uso del servicio de Turco y Sauna ya que han hecho uso inicialmente del servicio de la piscina.

5. ESTRATEGIAS (conclusiones)

ESTRATEGIA GENERAL

Establecer una estrategia de penetración de mercado (para clientes afiliados y no afiliados a la caja) apoyada de las diferentes herramientas de promoción.

SEGMENTO 1: LA PENSIONADA

5.1.1. OBJETIVO

Desarrollar las diferentes estrategias de promoción para generar un incremento en el número de usuarios en los segmentos de pensionados afiliado y no afiliados a la caja.

5.1.2. ESTRATEGIA

Captar y conservar los clientes afiliados y no afiliados del segmento de pensionados, generando una oferta de valor enfocadas al bienestar y servicios complementarios.

5.1.3. TACTICAS

- Tácticas de captación:

- Precios especiales: Se ofrecerá un paquete con precio especial el cual contenga un programa integral de bienestar (uso de zonas húmedas, gimnasio, clases adicionales como Pilates y yoga) el que podrá ser utilizado entre semana en la sede que prefiera.
- Acumulación de puntos: Se les entregara una tarjeta COMPENSAR con la cual acumulara puntos si durante 4 meses consecutivos utiliza los servicios de sauna y turco y piscina, los puntos podrán ser redimibles en cursos de cocina internacional y/o manualidades y si l utiliza estos servicios por 8 meses consecutivos sin redimir puntos tendrá derecho a un pasadía en un cetro vacacional (todo incluido).

- Tácticas de conservación:

- Trato preferencial y descuentos: Crear un club de personas pensionadas, las cuales tendrán derecho a un plan integral que incluye revisión medica mensual, descuentos especiales en los servicios de turco, sauna, piscina, cursos de manualidades, cocina, yoga, pilates y descuentos en cadenas convenio con caja.

5.1.4. PLAN DE MEDIOS

La comunicación para este segmento se realizara a través de los siguientes medios de comunicación:

- Telefónico: se llamara a los pensionados afiliados y no afiliados de las bases de datos para informarles acerca de los nuevos beneficios al hacer uso de los servicios de la caja.

- Personalizado: a las personas de este segmento que se acerquen a las diferentes sedes se les informara con personas especializadas estos beneficios y se les entregara un volante con la información.

- Virtual: Se enviara un mailing personalizado a todas las personas del segmento informando los beneficios de tomar los servicios de la caja.

Para divulgar este plan se hará uso de medios de comunicación masiva como son Radio y prensa.

5.2. SEGMENTO 2 : EL YUPPIE

5.2.1. OBJETIVO

Desarrollar las diferentes estrategias de promoción para generar un incremento en el número de usuarios en los segmentos de hombres y mujeres entre 20 y 40 años de nivel socio económico medio alto afiliado y no afiliados a la caja.

5.2.2. ESTRATEGIA

Captar y conservar los clientes afiliados y no afiliados del segmento hombres y mujeres entre 20 y 40 años de nivel socio económico medio alto, generando una oferta de valor enfocadas al bienestar y servicios complementarios.

5.2.3. TACTICAS

- Tácticas de captación:

Referidos: *Se entregará un descuento del 20% tanto al afiliado como no afiliado que entre semana tome el servicio de práctica libre de Tenis de campo o squash y una hora de zonas húmedas.

*Si el afiliado invita 2 No Afiliados, el afiliado tendrá un descuento del 50% sobre la tarifa de la categoría, el No afiliado tendrá un descuento del 20%, aplica para fines de semana.

*El No afiliado tendrá un descuento del 30% de descuento sobre el curso de tenis de campo, squash o bolos, si hace la inscripción con otra persona No afiliada.

- Trato preferencial: Tendrá a su disposición un profesional especializado para guiarlo en las actividades en tenis de campo, squash y bolos, al igual que la posibilidad de participar en torneos internos que realice la caja en estas tres disciplinas.

- **Tácticas de conservación:**

- Trato preferencial: Si toma 4 meses de práctica libre en alguna de las disciplinas de tenis de campo, squash o natación, podrá participar con un 20% de descuento en los torneos que realice la caja.
- Descuentos Especiales: Tendrá un descuento en tiendas convenio con la caja de compensación para compra de artículos deportivos de los deportes de: tenis de campo y squash.
- Membrecía: Se les otorgara una tarjeta elite donde Serán los primeros en ser informados de los cursos que se lleven a cabo en la caja para las 3 disciplinas, tenis de campo, bolos y squash, adicionalmente, obtener descuentos en la estadía de los hoteles de la caja y posibilidad de obtener créditos para libre inversión.

5.2.4. PLAN DE MEDIOS

La comunicación para este segmento se realizara a través de los siguientes medios de comunicación:

- Social media: Utilizaremos las redes sociales, microblogs y medios móviles, para informar acerca de los cursos, torneos y nuevos beneficios al hacer uso de los servicios de la caja
- Virtual: Se enviara un mailing personalizado a todas las personas del segmento informando los beneficios de tomar los servicios de la caja y poder realizar pagos en línea de los servicios que vaya a tomar.

5.3. SEGMENTO 3 : LA FAMILIA GONZALEZ

5.3.1. OBJETIVO

Desarrollar las diferentes estrategias de promoción para generar un incremento en el número de usuarios en los segmentos familiares afiliados y no afiliados a la caja.

5.3.2. ESTRATEGIA

Captar y conservar los clientes afiliados y no afiliados del segmento familiares, generando una oferta de valor enfocadas al bienestar y servicios complementarios.

5.3.3. TACTICAS

- **Tácticas de captación:**

- Descuentos: Los fines de semana habrá descuento por paquete familiar que incluye para niños menores a 4 años clases de matrogimnasia y acuamotricidad, para niños entre los 5 y los 12 años descuentos en práctica libres de natación, baile, patinaje, futbol y atracciones infantiles. Para pre adolescentes entre 13 y 15 años descuentos en práctica libre de natación, patinaje, futbol y para papá y mamá combos de turco, sauna, natación y /o terapia de relajación y como complemento una hora gratis de guardería (la hora del cuento).
- Tarjeta puntos: Se otorgara una tarjeta puntos, los cuales podrán redimir en compras en cadenas convenio con la caja.
- Invitaciones: Podrán vivir una experiencia inolvidable al compartir los diferentes escenarios y las actividades que en ellos se realizan (cuenteros,

pintucaritas, obras de teatro y actividades de acuerdo a la temporada) totalmente gratis.

- **Tácticas de conservación:**

- Descuentos: Los fines de semana habrá descuento por paquete familiar que incluye para niños menores a 4 años clases de matrogimnasia y acuamotricidad, para niños entre los 5 y los 12 años descuentos en práctica libres de natación, baile, patinaje, futbol y atracciones infantiles. Para pre adolescentes entre 13 y 15 años descuentos en práctica libre de natación, patinaje, futbol y para papá y mamá combos de turco, sauna, natación y /o terapia de relajación y como complemento una hora gratis de guardería (la hora del cuento).
- Tarjeta puntos: Se otorgara una tarjeta puntos, los cuales podrán redimir en compras en cadenas convenio con la caja

5.3.4. PLAN DE MEDIOS

La comunicación para este segmento se realizara a través de los siguientes medios de comunicación:

- Correo directo: Se enviara un correo directo a todas las personas del segmento, en el que se les enviara un llavero con la invitación a conocer y participar de los diferentes escenarios y servicios de la caja en sus diferentes sedes.

- Virtual: Se enviara un mailing personalizado a todas las personas del segmento informando los beneficios de tomar los servicios de la caja y poder realizar pagos en línea de los servicios que vaya a tomar.

5.4. SEGMENTO 4: EL ADOLECENTE

5.4.1. OBJETIVO

Desarrollar las diferentes estrategias de promoción para generar un incremento en el número de usuarios en los segmentos de adolescentes afiliados y no afiliados a la caja.

5.4.2. ESTRATEGIA

Captar y conservar los clientes afiliados y no afiliados del segmento familiares, generando una oferta de valor enfocadas al bienestar y servicios complementarios.

5.4.3. TACTICAS

- **Tácticas de captación:**

- Creación de eventos especiales y vínculos emotivos: invitación a exposición de deportes extremos, donde se rifaran bicicletas entre las personas de este segmento que adquirieran cursos libres de tenis de mesa, gimnasio y zonas húmedas.

- **Tácticas de conservación:**

- Descuentos: por la compra de un paquete que contenga tenis de mesa, voleibol, natación y turco y sauna entre semana tendrá un descuento preferencial y para los fines de semana tendrá media hora adicional gratis en su práctica de tenis de mesa, esta última oferta es condicionada a que vaya con un adulto.

5.4.4. PLAN DE MEDIOS

La comunicación para este segmento se realizara a través de los siguientes medios de comunicación:

- Virtual: Se enviara un mailing personalizado a todas las personas del segmento informando los beneficios de tomar los servicios de la caja y poder realizar pagos en línea de los servicios que vaya a tomar.

- Social media: Utilizaremos las redes sociales, microblogs y medios móviles, para informar acerca de los cursos, torneos y nuevos beneficios al hacer uso de los servicios de la caja.

6. RECOMENDACIONES

Se recomienda a la caja de compensación mejorar el proceso de captación de información el cual se puede ajustar realizando capacitación y entrenamiento en los diferentes software de toma de datos de los afiliados y no afiliados.

7. BIBLIOGRAFÍA

ASOCAJAS. (DICIEMBRE de 2012). *ESTADÍSTICAS: ASOCAJAS ASOCIACION NACIONAL DE CAJAS DE COMPENSACIÓN*. Obtenido de ASOCAJAS ASOCIACION NACIONAL DE CAJAS DE COMPENSACIÓN:
http://www.asocajas.org.co/index.php?option=com_wrapper&view=wrapper&Itemid=111

Carne, C. P. (2 de JUNIO de 2007). *MONOGRAFÍAS.COM*. Obtenido de MONOGRAFÍAS.COM:
<http://www.monografias.com/trabajos/datamining/datamining.shtml>

ELEMPLEO.COM. (s.f.). *INVESTIGACIÓN LABORAL- ELMPLEO.COM*. Recuperado el 17 de 08 de 2013, de ELEMPLEO.COM:
http://www.elempleo.com/colombia/investigacion_laboral/cajas-mnes-que-compensacin-n-familiar-----/6586557

JOJOOA- Tecnología, Marketing y CRM. (n.d.). *Diccionario de CRM y Marketing:JOJOOA- Tecnología, Marketing y CRM*. Retrieved 2013 йил 21-Agosto from JOJOOA- Tecnología, Marketing y CRM:
<https://sites.google.com/site/jojooa/inteligencia-artificial/definicion-de-metodos-no-supervisados-que-son-los-metodos-no-supervisados>

Microsoft SQL Server . (n.d.). *Library: Microsoft SQL Server* . Retrieved 2013 йил 21-AGOSTO from Microsoft SQL Server : [http://technet.microsoft.com/es-es/library/ms175595\(v=sql.105\).aspx](http://technet.microsoft.com/es-es/library/ms175595(v=sql.105).aspx)

REDACCIÓN ECONOMÍA Y NEGOCIOS. (17 de ABRIL de 2013).

PUBLICACION:ELTIEMPO.COM: SECCIÓN: ECONOMIA. Recuperado el 16 de AGOSTO de 2013, de ELTIEMPO.COM:

<http://www.eltiempo.com/archivo/documento/CMS-12745695>

REDACCION POLITICA. (20 de MARZO de 2013). *POLITICA: ELTIEMPO.COM.*

Recuperado el 17 de AGOSTO de 2013, de ELTIEMPO.COM:

http://www.eltiempo.com/politica/ARTICULO-WEB-NEW_NOTA_INTERIOR-12704665.html

SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR. (s.f.). *CAJAS DE COMPENSACIÓN*

FAMILIAR: SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR. Recuperado el 22 de 08 de 2013, de SUPERINTENDENCIA DEL SUBSIDIO FAMILIAR:

<http://www.ssf.gov.co/wps/portal/Espa%C3%B1ol/Cajas%20de%20Compensaci%C3%B3n%20Familiar!/ut/p/b1/vZTJjptAEIafZR5gRLHjIxiGtdn3i8VqGxuDjQ2Gpx8yl ygjJbkk7jpV6299-v8uFZZiMZZesvG4z->

[7H7pKdf_Qps8MtjpYlgwALzagijQjg4YIiiNWQbIKtjKvUKwBwBkyDSqvBO7GI Ungyb-9j7C4TlhTHN](http://www.ssf.gov.co/wps/portal/Espa%C3%B1ol/Cajas%20de%20Compensaci%C3%B3n%20Familiar!/ut/p/b1/vZTJjptAEIafZR5gRLHjIxiGtdn3i8VqGxuDjQ2Gpx8yl ygjJbkk7jpV6299-v8uFZZiMZZesvG4z-7H7pKdf_Qps8MtjpYlgwALzagijQjg4YIiiNWQbIKtjKvUKwBwBkyDSqvBO7GI Ungyb-9j7C4TlhTHN)

Wikipedia. (2013 йил 30-JULIO). *Wikipedia la enciclopedia libre.* Retrieved 2013 йил 21-

AGOSTO from Wikipedia la enciclopedia libre: <http://es.wikipedia.org/wiki/K-means>

Wikipedia. (2013 йил 21-MAYO). *Wikipedia la enciclopedia libre.* Retrieved 2013 йил 24-

AGOSTO from Wikipedia la enciclopedia libre:

http://es.wikipedia.org/wiki/Algoritmo_apriori

Wikipedia. (2013 йил 26-JUNIO). *Wikipedia la enciclopedia libre*. Retrieved 2013 йил 21-

AGOSTO from Wikipedia la enciclopedia libre:

<http://es.wikipedia.org/wiki/RapidMiner>