



**UNA SEGMENTACIÓN DE CLIENTES QUE ADQUIRIERON
CRÉDITOS DE LIBRE INVERSIÓN EN UNA ENTIDAD
FINANCIERA COLOMBIANA I SEMESTRE 2015**

**OSCAR TUMBIA
JAIRO MARTINEZ
OMAR BELTRAN**

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA**

BOGOTÁ D.C.

2015



**SEGMENTACIÓN DE CLIENTES QUE ADQUIRIERON
CRÉDITOS DE LIBRE INVERSIÓN EN UNA ENTIDAD
FINANCIERA COLOMBIANA I SEMESTRE 2015**

**OSCAR TUMBIA
JAIRO MARTINEZ
OMAR BELTRAN**

**ASESOR:
CECY ALEXANDRA PENAGOS ESPINOSA**

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
TRABAJO DE GRADO PARA OPTAR EL TÍTULO DE ESPECIALISTA EN
ESTADÍSTICA APLICADA
BOGOTÁ D.C.
2015**

Nota de Aceptación

Firma del presidente del jurado

Firma del Jurado

Firma del Jurado

Bogotá, D.C 29 de julio de 2016

Las Directivas de la Universidad de
Los Libertadores, los jurados calificadores y el cuerpo
Docente no son responsables por los
criterios e ideas expuestas En el presente documento.
Estos corresponde únicamente a los autores

TABLA DE CONTENIDO

1 INTRODUCCIÓN	7
1.1 PROBLEMÁTICA.....	7
1.1 PREGUNTA O FORMULACIÓN	7
1.2 OBJETIVOS	8
1.2.1 Objetivo General	8
1.2.2 Objetivos Específicos.....	8
2 MARCO DE REFERENCIA	9
2.1 Mercado de créditos de libre inversión en Colombia.....	9
2.2 Análisis de correspondencias múltiples (ACM).....	11
2.3 Análisis cluster y dendograma	12
3 MARCO METODOLÓGICO.....	13
3.1 Obtención de datos de estudio.....	13
3.2 Variables de estudio	13
3.3 Unidades De Análisis	15
3.4 Procedimiento	15
4 RESULTADOS.....	17
4.1 ESTADÍSTICA DESCRIPTIVA DE LAS VARIABLES.....	17
4.2 DIAGRAMAS DE DISPERSIÓN	20
4.3 ANÁLISIS DE CORRESPONDENCIAS	22
5 DISCUSIÓN.	26
5.1 ANÁLISIS DE CLUSTER.	26
6 CONCLUSIONES.	31
7 RECOMENDACIONES	32
8 GLOSARIO	33
9 REFERENCIAS.....	35

LISTA DE TABLAS

Tabla 1: Clasificación de las entidades financieras de acuerdo al nivel de colocación.	10
Tabla 2: Ejemplo tabla de contingencia	11
Tabla 3: Ejemplo tabla de correspondencias.....	11
Tabla 4: Variables cuantitativas	14
Tabla 5: Transformación de variables cuantitativas en cualitativas.....	15
Tabla 6: Variables discretas de la base	15
Tabla 7: Resumen estadístico monto desembolsado y tasa de interés	17
Tabla 8: Porcentaje de la opinión general por calificación	18
Tabla 9: Composición clientes según su tipo	19
Tabla 10: Composición clientes por ocupación	20
Tabla 11: Tabla de Autovalores	23
Tabla 12: Tabla de contribuciones por factor	24
Tabla 13: Cantidad de individuos por cluster	25
Tabla 14: Descripción - cluster 1	27
Tabla 15: Descripción – cluster 2	28
Tabla 16: Descripción - cluster3	29
Tabla 17: Factor map	30

LISTA DE FIGURAS

Ilustración 1. Colocación de créditos de libre inversión	9
Ilustración 2: Boxplot monto del desembolso y tasa de interés	17
Ilustración 3: Porcentaje satisfacción en transformación cualitativa:	18
Ilustración 4: Composición clientes por segmento	19
Ilustración 5: Diagrama de dispersión monto desembolsado y tasa	20
Ilustración 6: Calificación general frente al desembolso	21
Ilustración 7: Tabla factorial dimensión 1 y dimensión 2.....	23
Ilustración 8: Dendograma.....	25

1 INTRODUCCIÓN

1.1 PROBLEMATICA.

Las entidades financieras elaboran estrategias para colocar en el mercado sus productos, entre ellos los créditos de libre inversión. En la actualidad existen más de 15 bancos en el país que ofertan este tipo de créditos; sumándose a ellos las cajas de compensación, fondos de empleados entre otros. Esto evidencia la alta competencia del mercado de créditos.

La alta competencia hace que los clientes sean más exigentes y analicen con mayor detalle las opciones que el mercado les ofrece.

Por lo cual resulta imprescindible conocer con mayor detalle quienes son estos clientes y cuáles son las características que tienen en común para poder conformar segmentos más específicos. De esta manera se podría hacer una demanda diferencial para cada uno de estos segmentos.

1.1 PREGUNTA O FORMULACIÓN

El presente estudio tiene como propósito identificar los segmentos más relevantes de los clientes a los que se les desembolsó créditos de libre inversión. En este sentido se formula la pregunta ¿En qué clusters se clasifican los clientes que solicitan créditos de libre inversión en la entidad financiera colombiana?

1.2 OBJETIVOS

1.2.1 Objetivo General

Identificar los segmentos más representativos de los clientes a los que se les desembolsó créditos de libre inversión en el primer semestre de 2015 en una entidad financiera colombiana.

1.2.2 Objetivos Específicos

- Determinar el número de clusters óptimo para hacer la segmentación.
- Definir factores de agrupamiento para los segmentos.
- Identificar las características de cada segmento de mercado.

2 MARCO DE REFERENCIA

2.1 Mercado de créditos de libre inversión en Colombia

En los últimos 9 años la colocación de créditos de libre inversión que han desembolsado las entidades financieras en Colombia ha tenido un incremento del 417% (Superintendencia Financiera de Colombia, 2015). En este periodo de tiempo el mercado se ha distribuido de la siguiente manera:

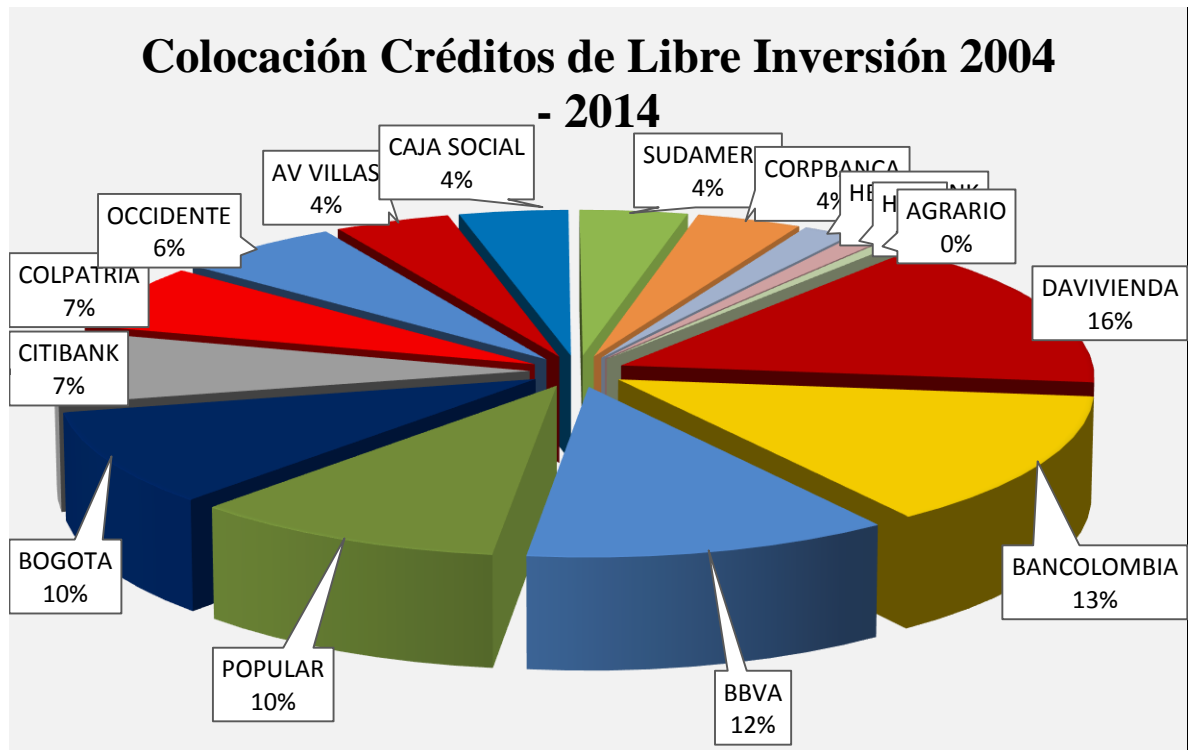


Ilustración 1. Colocación de créditos de libre inversión

Con estos datos clasificamos los Bancos según su colocación de créditos en tres grupos: Alta Colocación, Media Colocación y Baja Colocación.

Alta Colocación 41%	Media Colocación 40%	Baja Colocación 19%
<ul style="list-style-type: none"> ➤ DAVIVIENDA ➤ BANCOLOMBIA ➤ BBVA 	<ul style="list-style-type: none"> ➤ POPULAR ➤ BOGOTA ➤ CITIBANK ➤ COLPATRIA ➤ OCCIDENTE 	<ul style="list-style-type: none"> ➤ AV VILLAS ➤ CAJA SOCIAL ➤ SUDAMERIS ➤ CORPBANCA ➤ HELM BANK ➤ HSBC ➤ AGRARIO

Tabla 1: Clasificación de las entidades financieras de acuerdo al nivel de colocación.

Las empresas día a día toman decisiones estratégicas con la finalidad de mantener y ganar participación en el mercado (Schnaars, 1993):

- Bajar los precios, en este caso las Tasas de Interés
- Mejorando los productos, en este caso brindar beneficios adicionales
- Focalizando su demanda a mercados masivos
- Aumentar la publicidad
- Atraer clientes de la competencia directa (hacer compra de cartera de otras entidades)
- Brindar un servicio de Calidad

Y en el sector financiero estas específicas

- Exonerar de costos los productos
- Entregar regalos
- Permitir que los clientes acumulen puntos

2.2 Análisis de correspondencias múltiples (ACM)

Las correspondencias múltiples son una técnica descriptiva, la cual resume una gran cantidad de datos en un número reducido de dimensiones o factores (Rencher, 1934). Este análisis consiste en la extracción de relaciones entre categorías y define similitudes o disimilitudes entre ellas, lo que permitirá su agrupamiento si se detecta que se corresponden.

Existen dos tipos de Correspondencias:

- Análisis de correspondencias Simple (ACS)

Los datos a analizar tienen forma de tabla de contingencia y evalúa las relaciones existentes entre las dos variables.

	Enfermos	Sanos	
Expuestos (vacunados)	0	24	24
No expuestos (no vacunados)	4	27	31
	4	52	56

Tabla 2: Ejemplo tabla de contingencia

- Análisis de correspondencias Múltiple (ACM)

Se determina ACM cuando se trabaja con más de dos variables nominales

Tabla de correspondencias						
Productos	Características					
	Brillo	Duración	Olor	Comodidad	Limpeza	Margen activo
Producto A	68	60	67	45	35	275
Producto B	12	13	13	15	12	65
Producto C	95	84	94	63	49	385
Producto D	30	32	32	36	28	158
Margen activo	205	189	206	159	124	683

Tabla 3: Ejemplo tabla de correspondencias

El ACM reduce los datos analizado en un número de dimensiones que permite analizar la información de una manera más sencilla, y permitiendo hallar los conjuntos similares de individuos.

2.3 Análisis cluster y dendograma

En el análisis de clúster lo que se busca es reducir las dimensiones necesarias para interpretar el conjunto de datos, nuestro objetivo principal al utilizar esta herramienta es identificar grupos de individuos los cuales deben cumplir una serie de características importantes; lo primero es que los individuos dentro de cada grupo deben ser lo mas semejantes posibles entre si y lo segundo es que cada grupo debe ser lo más diferente posible entre los demás grupos, es decir homogéneos dentro cada grupo y heterogéneos entre cada grupo. Para el desarrollo de la técnica el software estadístico utiliza una métrica para medir las distancias entre los individuos entre sí, esta distancia se conoce como la distancia euclidea que quiere decir la distancia mínima entre los individuos. Después se calcula en dendograma.

Después de conseguir el objetivo de agrupar los individuos más semejantes entre si y que los grupos sean lo más diferentes entre si se procede con la organización de la información por medio de métodos jerárquicos o método de partición, para este caso se va a trabajar el dendograma que es un método jerárquico. En donde se utilizó la técnica del centroide que consiste en que cada clúster se reduce hasta un vector llamado centroide que contiene las medias de cada grupo en cada una de las variables medidas, los resultados obtenidos se muestran mediante un gráfico llamado DENDOGRAMA, el cual nos muestra de una forma gráfica la cantidad de clúster que debemos formar

3 MARCO METODOLÓGICO.

3.1 Obtención de datos de estudio

El banco toma la base de datos con los clientes a los que se les desembolsaron los créditos, la cual cuenta con la caracterización demográfica de los clientes (segmento, ocupación, tipo de cliente, entre otras).

A través del contact center se contacta a los clientes mencionados anteriormente y se le da información del producto que contrató, canales de atención, y se les hace una encuesta de satisfacción de la prestación del servicio.

La data se presenta para su análisis mensualmente y para la realización del estudio se consolidó el primer semestre.

3.2 Variables de estudio

Para el estudio se tuvieron en cuenta las siguientes variables:

Monto desembolsado: Es la cantidad de dinero que el Banco le presta a los clientes. Es una variable continua y se mide en pesos

Tasa: Es la tasa de interés a la que el Banco concede los créditos. Es una variable continua y se mide como porcentaje y se expresa en valores efectivos anuales.

Nivel de satisfacción general: Es la respuesta que dan los usuarios cuando en la encuesta telefónica se les pregunta por la satisfacción en el servicio que presta el Banco al momento de tomar el producto. Es una variable discreta ordinal que puede tomar valores entre uno y diez.

Segmento: Es una clasificación que el Banco hace de los clientes según sus ingresos, si gana desde 1 SMLV hasta 6 SMLV se considera un cliente masivo y si gana más de 6 SMLV se le considera un cliente preferente. Es una variable discreta.

Tipo de cliente: Clasificación de los clientes del Banco según su antigüedad y origen. Es una variable discreta que puede tomar los siguientes valores: antiguo, nuevo, antiguo OVN, nuevo OVN y nuevo sin oficina. Los clientes antiguo ovn y nuevo ovn son aquellos que ingresan por convenios de nómina. Los clientes nuevos sin oficina son los que adquirieron productos a través de internet

Ocupación: Es una variable cualitativa nominal que indica el origen de los recursos de los clientes. Se pueden clasificar como desempleados, empleados públicos, empleados privados, estudiantes, independientes, pensionados, transportadores, rentistas de capital, amas de casa y personas sin clasificar.

Variables Cuantitativas

	TIPO DE VARIABLE	UNIDAD
Monto desembolsado	Continua	Pesos
Tasa	Continua	Porcentaje
Nivel de satisfacción general	Ordinal	Numérica

Tabla 4: Variables cuantitativas

Estas variables cuantitativas se transformaron en cualitativas

	TIPO DE VARIABLE	VALORES	NOMBRE EN LA BASE
1 Monto desembolsado	Ordinal	<ul style="list-style-type: none"> ○ Entre 1 y 3 millones ○ Entre 3 y 5 millones ○ Entre 5 y 10 millones ○ Entre 10 y 20 millones ○ Entre 20 y 50 millones ○ Entre 50 y 100 millones ○ Más de 100 millones 	MONTO
2 Tasa	Ordinal	<ul style="list-style-type: none"> ○ Muy Baja ○ Media Baja ○ Media Alta ○ Alta ○ Limite Usura 	TASA

3	Nivel de satisfacción general	Ordinal	<input type="radio"/> Insatisfecho <input type="radio"/> Neutral <input type="radio"/> Satisfecho	SATISFACCION
----------	--------------------------------------	---------	---	---------------------

Tabla 5: Transformación de variables cuantitativas en cualitativas

Adicional se tienen estas variables cualitativas:

4	Segmento	Ordinal	<input type="radio"/> Masivo <input type="radio"/> Preferente	SEGMENTO
5	Tipo de cliente	Ordinal	<input type="radio"/> Antiguo <input type="radio"/> Antiguo OVN <input type="radio"/> Nuevo <input type="radio"/> Nuevo OVN <input type="radio"/> Nuevo Sin Oficina	TIPO DE CLIENTE
6	Ocupación	Ordinal	<input type="radio"/> Desempleado <input type="radio"/> Empleado Privado <input type="radio"/> Empleado Público <input type="radio"/> Hogar <input type="radio"/> Independiente <input type="radio"/> Pensionado <input type="radio"/> Por Definir <input type="radio"/> Rentista de Capital <input type="radio"/> Transportador	OCUPACIÓN

Tabla 6: Variables discretas de la base

3.3 Unidades De Análisis

Se estudian 3.648 registros de clientes de una entidad bancaria colombiana que relacionan detalles de caracterización de los clientes del Banco, los montos y tasas que se les desembolsaron con los resultados de una encuesta de satisfacción que practica el Banco a los clientes que adquieren un producto en la entidad.

3.4 Procedimiento

El enfoque del trabajo es exploratorio descriptivo. Para hacer la segmentación se toma la base y se comienza haciendo un análisis de correspondencias múltiples que permita transformar los datos a una presentación que facilite su análisis.

Luego se hace un análisis de conglomerados para formar agrupaciones entre los individuos más similares. El método utilizado es jerárquico aglomerativo; es decir que se parte de tener a todos los individuos por separado y se van juntando sucesivamente los individuos más cercanos hasta que todos los individuos queden agrupados en un único grupo. Se elabora entonces un dendograma que permite hallar saltos de varianza más importantes y en dicho punto se hace el corte del dendograma y se identifica así el número óptimo de clusters.

Acto seguido se describe la composición de los clusters con las tablas cla/mod mod/cla. La primera permite examinar la proporción de individuos que cumplen con la característica señalada que hacen parte del cluster y la segunda el peso relativo de los individuos con dicha característica dentro del cluster. A continuación se describe a un individuo de cada cluster y se presentan las conclusiones del estudio.

Los datos se procesan con el software estadístico R.

4 RESULTADOS.

4.1 ESTADÍSTICA DESCRIPTIVA DE LAS VARIABLES

En primer lugar se realiza una representación gráfica de las variables continuas utilizando

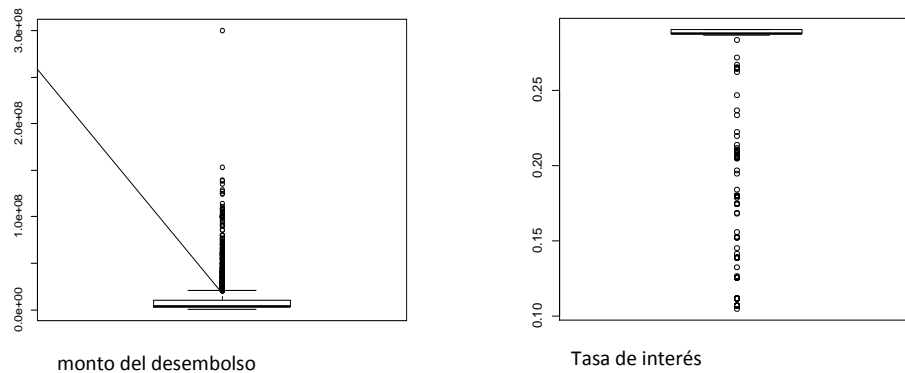


Ilustración 2: Boxplot monto del desembolso y tasa de interés

diagramas de caja y bigotes y se incluye una tabla que resume sus valores:

ESTIMADOR	MONTO DESEMBOLSO	TASA
Mínimo	707.500	10,51%
Primer cuartil	3.093.000	28,82%
Mediana	4.000.000	28,82%
Media	9.629.000	26,69%
Tercer cuartil	10.040.000	29,06%
Máximo	152.800.000	29,06%

Tabla 7: Resumen estadístico monto desembolsado y tasa de interés

Al observar el monto desembolsado podemos notar que la mayoría de las observaciones se agrupan entre \$3.093.000 (primer cuartil) y \$10.040.000 (tercer cuartil). Los datos atípicos son los desembolsos que superan los \$20.478.505 (Figura 1); y sumados, representan el 51,74% del total de los desembolsos efectuados por el banco durante el periodo. El máximo desembolso registrado asciende a \$152.800.000.

Por su parte, la mayoría de las observaciones en tasa de interés se agrupan entre el 28,82% y el 29,06%, presentando una alta presencia de datos atípicos por debajo del 28,35% que toma un valor mínimo de 10,51% (Figura 2)

Para analizar las variables discretas se utilizan tablas y gráficos. Primero se observa la opinión general sobre la atención al cliente:

CALIFICACIÓN	PORCENTAJE
1	0,31%
2	0,10%
3	0,21%
4	0,10%
5	1,03%
6	0,21%
7	2,99%
8	7,22%
9	8,35%
10	79,07%
NS/NR	0,41%

Tabla 8: Porcentaje de la opinión general por calificación

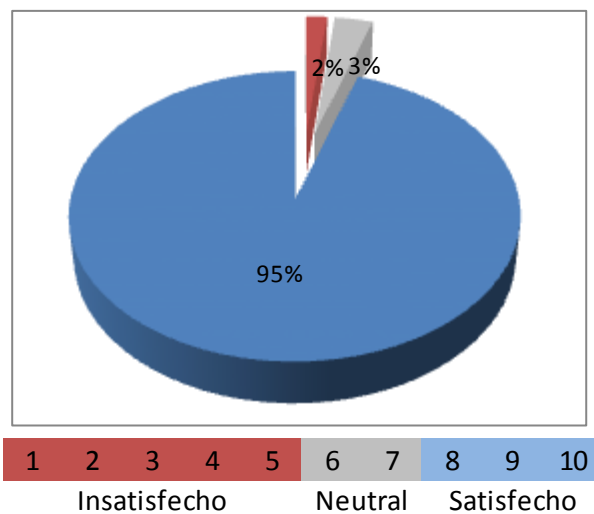


Ilustración 3: Porcentaje satisfacción en transformación cualitativa:

Al analizar la opinión general de la atención, notamos que la gran mayoría de los datos se agrupan en la calificación 10; representando el 79,07% de las observaciones.

Se observa también la composición de los clientes de acuerdo a su segmento. Notamos que el 81% se clasifican como masivo, y el 19% como preferente:

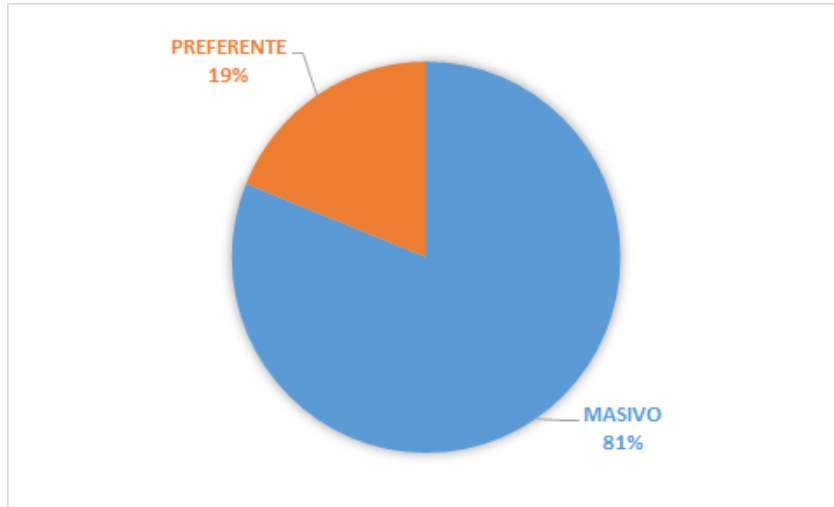


Ilustración 4: Composición clientes por segmento

Por el tipo de cliente, la mayor parte son clientes antiguos (51,65%) y nuevos (37,53%). El 10,31% corresponde a clientes antiguos y nuevos de nómina.

TIPO DE CLIENTE

ANTIGUO	51,65%
ANTIGUO OVN	10,10%
NUEVO	37,53%
NUEVO OVN	0,21%
NUEVO SIN OFICINA	0,52%

Tabla 9: Composición clientes según su tipo

Por ocupación, el 87,64% de los clientes son empleados privados, el 4,82% son pensionados, el 3,56% son empleados independientes y el 2,25% son rentistas de capital.

OCUPACIÓN

DESEMPLEADO	0,33%
EMPLEADO PRIVADO	87,64%

EMPLEADO PÚBLICO	0,30%
ESTUDIANTE	0,05%
HOGAR	0,08%
INDEPENDIENTE	3,56%
PENSIONADO	4,82%
POR DEFINIR	0,14%
RENTISTA DE CAPITAL	2,25%
TRANSPORTADOR	0,82%

Tabla 10: Composición clientes por ocupación

4.2 DIAGRAMAS DE DISPERSIÓN

Se hizo la representación gráfica de las variables buscando observar la correlación de las mismas. Al dibujar la tasa de interés frente al monto desembolsado se puede notar una

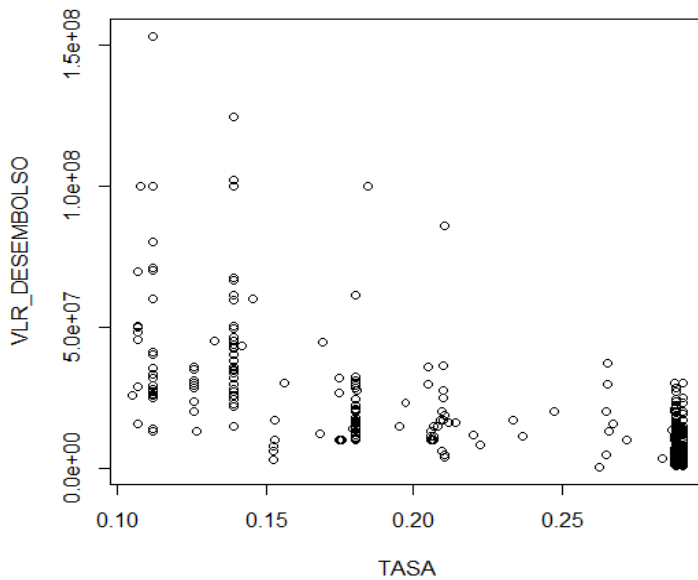


Ilustración 5: Diagrama de dispersión monto desembolsado y tasa

relación inversa.

También se graficó una serie de gráficos de caja y bigotes que muestran la relación entre la calificación y el valor de desembolso. Se puede notar que la relación entre monto desembolsado y calificación de la atención no muestran una relación al observar cajas y bigotes aunque si se puede percibir una relación directa al observar los datos atípicos.

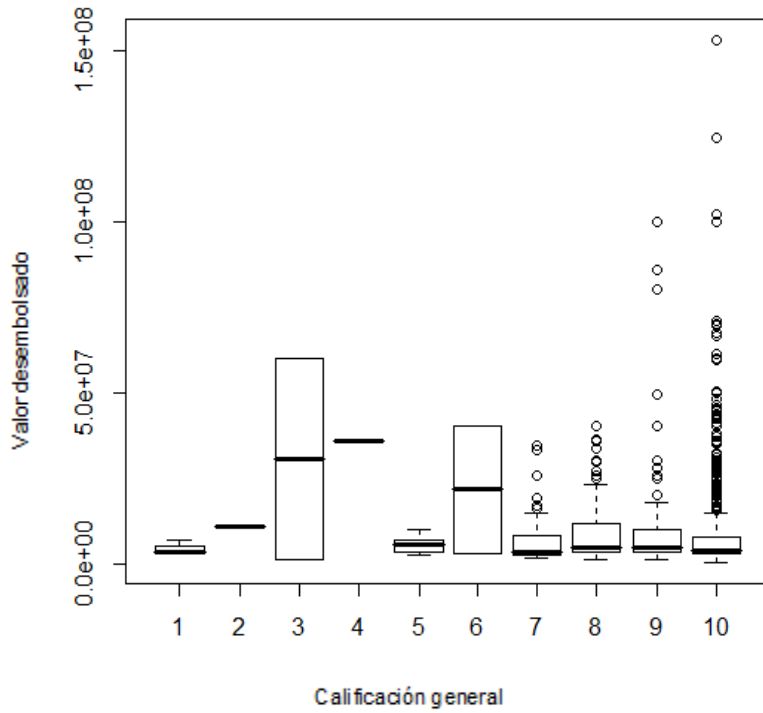


Ilustración 6: Calificación general frente al desembolso

4.3 ANÁLISIS DE CORRESPONDENCIAS

En el análisis de clúster lo que se busca es reducir las dimensiones necesarias para interpretar el conjunto de datos, nuestro objetivo principal al utilizar esta herramienta es identificar grupos de individuos los cuales deben cumplir una serie de características importantes; lo primero es que los individuos dentro de cada grupo deben ser lo mas semejantes posibles entre si y lo segundo es que cada grupo debe ser lo más diferente posible entre los demás grupos, es decir homogéneos dentro cada grupo y heterogéneos entre cada grupo. Para el desarrollo de la técnica el software estadístico utiliza una métrica para medir las distancias entre los individuos entre sí, esta distancia se conoce como la distancia euclidea que quiere decir la distancia mínima entre los individuos. Después se calcula en dendograma.

Después de conseguir el objetivo de agrupar los individuos más semejantes entre si y que los grupos sean lo más diferentes entre si se procede con la organización de la información por medio de métodos jerárquicos o método de partición, para este caso se va a trabajar el dendograma que es un método jerárquico. En donde se utilizó la técnica del centroide que consiste en que cada clúster se reduce hasta un vector llamado centroide que contiene las medias de cada grupo en cada una de las variables medidas, los resultados obtenidos se muestran mediante un gráfico llamado DENDOGRAMA, el cual nos muestra de una forma gráfica la cantidad de clúster que debemos formar.

Para comenzar se busca una forma de resumir la información en un número reducido de dimensiones para facilitar su análisis. El número óptimo de las mismas se determina examinando la tabla de autovalores e identificando puntos con saltos significativos de varianza. Se puede observar que los mayores saltos de varianza se encuentran entre las tres primeras dimensiones.

	eigenvalue	percentage of variance	cumulative percentage of variance
dim 1	0.44453352	10.2584659	10.25847
dim 2	0.24306131	5.6091072	15.86757
dim 3	0.20041929	4.6250605	20.49263
dim 4	0.19401337	4.4772315	24.96987
dim 5	0.18535963	4.2775298	29.24739
dim 6	0.18107902	4.1787466	33.42614
dim 7	0.17758735	4.0981696	37.52431
dim 8	0.17542021	4.0481587	41.57247
dim 9	0.17001743	3.9234792	45.49595
dim 10	0.16874742	3.8941713	49.39012
dim 11	0.16692618	3.8521427	53.24226
dim 12	0.16665837	3.8459624	57.08823
dim 13	0.16615410	3.8343253	60.92255
dim 14	0.16452202	3.7966619	64.71921
dim 15	0.16187538	3.7355856	68.45480
dim 16	0.16135262	3.7235219	72.17832
dim 17	0.15842534	3.6559693	75.83429
dim 18	0.15538626	3.5858367	79.42013
dim 19	0.15449816	3.5653422	82.98547
dim 20	0.14973213	3.4553569	86.44083
dim 21	0.14536586	3.3545967	89.79542
dim 22	0.14006070	3.2321699	93.02759
dim 23	0.12880140	2.9723400	95.99993
dim 24	0.09872949	2.2783728	98.27830
dim 25	0.04701044	1.0848564	99.36316
dim 26	0.02759635	0.6368388	100.00000

Tabla 11: Tabla de Autovalores

En el plano factorial de las dos primeras dimensiones tenemos una primera aproximación a las relaciones entre las variables. Podemos ver que la tasa y el monto están explicadas por

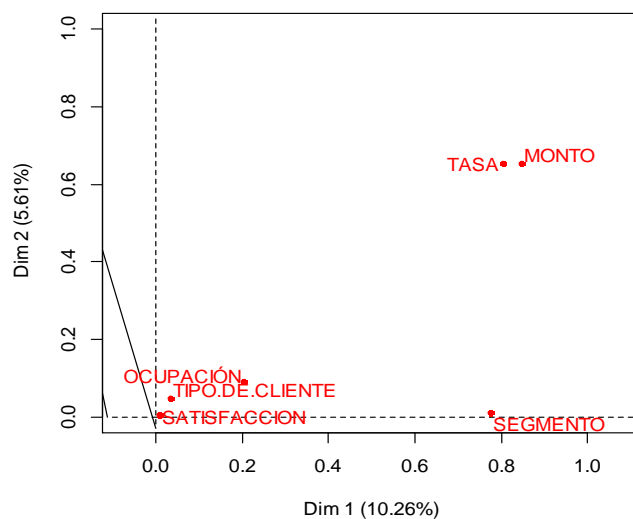


Ilustración 7: Tabla factorial dimensión 1 y dimensión 2

las dos dimensiones y el segmento por la primera dimensión.

Lo anterior se puede corroborar con la tabla de contribuciones de cada factor. Se puede notar que el segmento, la tasa y el monto están mejor representadas en la primer dimensión, mientras que la ocupación lo está en la tercera dimensión.

	Dim 1	Dim 2	Dim 3
TIPO.DE.CLIENTE	0.033512940	0.047115294	0.1738431654
SEGMENTO	0.773934508	0.009761668	0.0008319508
OCUPACIÓN	0.203743538	0.089941675	0.3318535521
TASA	0.802366006	0.654194737	0.4437069296
SATISFACCION	0.008287086	0.003697780	0.0261384071
MONTO	0.845357058	0.653656715	0.2261417254

Tabla 12: Tabla de contribuciones por factor

Por lo tanto se tiene que:

- La primer dimensión representa las variables segmento, tasa y monto
- La segunda dimensión representa las variables tasa y monto
- La tercera dimensión representa las variables ocupación, tasa y tipo de cliente.

Con esta información se busca hacer una clasificación de los clientes dentro de segmentos de mercado. Para lograrlo se construye un dendograma que señale donde ocurren los saltos más importantes. Se puede notar que se forman 3 cluster

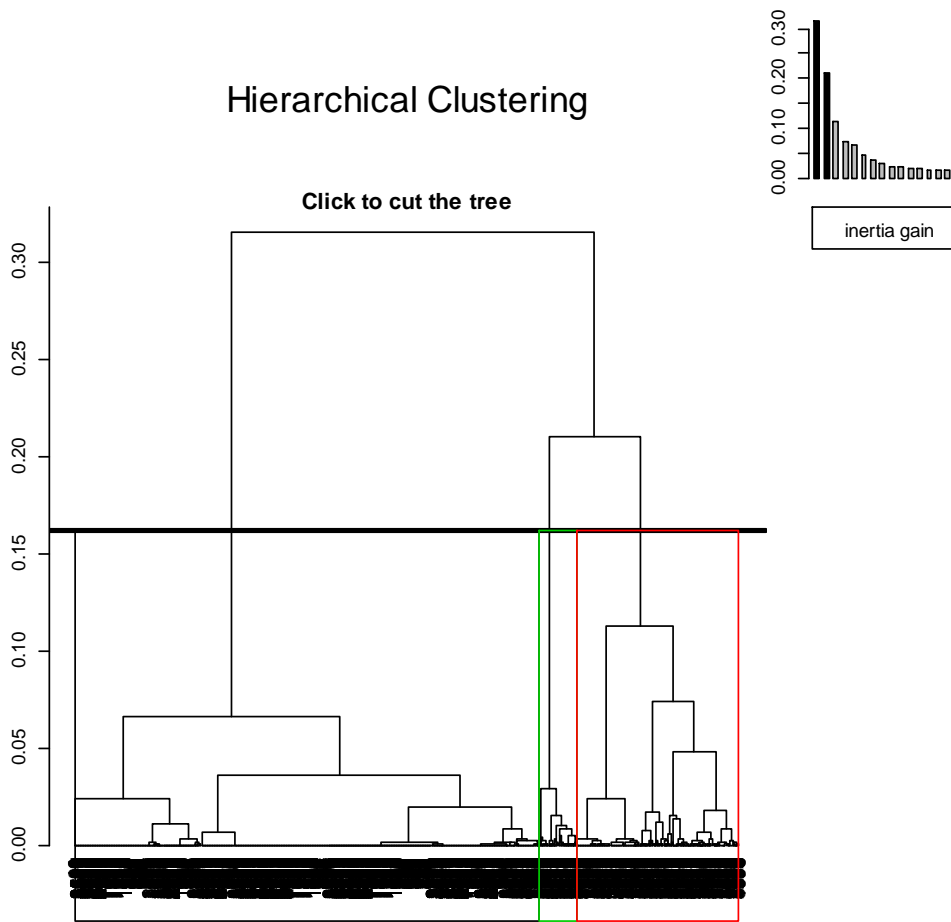


Ilustración 8: Dendograma

La cantidad de individuos en cada cluster se detalla a continuación:

Cluster	Tamaño
Cluster 1	2841
Cluster 2	612
Cluster 3	195

Tabla 13: Cantidad de individuos por cluster

5 DISCUSIÓN.

5.1 ANÁLISIS DE CLUSTER.

CLUSTER 1.

Teniendo en cuenta la información que nos muestra el análisis clúster, para este caso se observa que se encuentran clientes que han solicitado prestamos entre 3 y 5 millones con una participación del (47,8%), seguido de los clientes que han solicitado prestamos entre 1 y 3 millones (28,26%), y en este grupo también se logra ubicar a los clientes que han solicitado prestamos entre 5 y 10 millones de pesos con una participación del (17,45%). Adicionalmente se puede observar que la tasa de interés para estos clientes, se encuentra ubicada en el límite de usura (Tasa de interés muy alta) con una participación del (99,68%), también se observa que la mayoría de los clientes ubicados en este clúster son clientes antiguos ovn (11,08%), son empleados privados (92,99%), pertenecientes al segmento masivo (98,41%) y que según la encuesta de satisfacción el (95,98%) se encuentran satisfechos con el servicio prestado.

	Cla/Mod	Mod/Cla
OCUPACIÓN=Estudiante	0.0000000	0.0000000
OCUPACIÓN=Desempleado	0.0000000	0.0000000
MONTO=Más_de_100_millones	0.0000000	0.0000000
MONTO=Entre_50_y_100_millones	0.0000000	0.0000000
TASA=Media_Baja	0.5102041	0.03519887
TASA= Muy_Bajas	0.6666667	0.03519887
TASA=Media_Alta	1.6393443	0.17599437
TASA=Alta	3.5714286	0.07039775
MONTO=Entre_20_y_50_millones	4.0697674	0.49278423
SEGMENTO=PREFERENTE	8.2872928	1.58394931
OCUPACIÓN=Independiente	13.8461538	0.63357973
MONTO=Entre_10_y_20_millones	35.7723577	6.19500176
OCUPACIÓN=Rentista de Capital	36.5853659	1.05596621
TIPO.DE.CLIENTE=NUEVO SIN OFICINA	41.6666667	0.17599437
OCUPACIÓN=Transportador	43.3333333	0.45758536
SATISFACCION=Insatisfecho	65.0000000	1.37275607
SATISFACCION=Neutral	67.5675676	2.63991552
OCUPACIÓN=Pensionado	70.4545455	4.36466033
TIPO.DE.CLIENTE=NUEVO	75.1523358	39.07074974
SATISFACCION=Satisfecho	78.4296808	95.98732841
OCUPACIÓN=Empleado Privado	82.6399750	92.99542415
SEGMENTO=MASIVO	90.0483092	98.41605069
TIPO.DE.CLIENTE=ANTIGUO OVN	90.7780980	11.08764520
MONTO=Entre_5_y_10_millones	92.1933086	17.45864132
TASA=Limite_Usura	96.2937776	99.68321014
MONTO=Entre_1_y_3_millones	99.5043371	28.26469553
MONTO=Entre_3_y_5_millones	99.5581738	47.58887716

Tabla 14: Descripción - cluster 1

CLUSTER 2.

	Cla/Mod	Mod/Cla
MONTO=Entre_50_y_100_millones	0.0000000	0.0000000
MONTO=Entre_3_y_5_millones	0.4418262	0.9803922
MONTO=Entre_1_y_3_millones	0.4956629	0.6535948
TASA= Muy_Bajas	0.6666667	0.1633987
TASA=Limite_Usura	3.6042163	17.3202614
MONTO=Entre_5_y_10_millones	7.4349442	6.5359477
TIPO.DE.CLIENTE=ANTIGUO_OVN	7.4927954	4.2483660
SEGMENTO=MASIVO	9.8872786	50.1633987
OCUPACIÓN=Empleado Privado	13.0434783	68.1372549
SATISFACCION=Satisfecho	16.4797239	93.6274510
TIPO.DE.CLIENTE=NUEVO	19.2281652	46.4052288
SATISFACCION=Insatisfecho	28.3333333	2.7777778
MONTO=Más_de_100_millones	47.6190476	1.6339869
OCUPACIÓN=Transportador	50.0000000	2.4509804
OCUPACIÓN=Rentista de Capital	56.0975610	7.5163399
SEGMENTO=PREFERENTE	56.1694291	49.8366013
MONTO=Entre_10_y_20_millones	61.1788618	49.1830065
MONTO=Entre_20_y_50_millones	72.9651163	41.0130719
OCUPACIÓN=Independiente	73.8461538	15.6862745
TASA=Media_Baja	81.1224490	25.9803922
TASA=Media_Alta	95.7377049	47.7124183
TASA=Alta	96.4285714	8.8235294
OCUPACIÓN=Estudiante	100.0000000	0.3267974

Tabla 15: Descripción – cluster 2

Teniendo en cuenta la información que nos muestra el clúster 2, para este caso se observa que se encuentran clientes que han solicitado préstamos entre 10 y 20 millones (49.18%), también los clientes que han solicitado préstamos entre 20 y 50 millones (41.01%), con una tasa de interés media alta (47.71%), las personas de este se ubican en el segmento preferente (49.83%) y en el segmento masivo (50.16%), de los cuales son empleados privados (68.13%) y adicionalmente se encuentran satisfechos en un (93.62%).

CLUSTER 3.

	Cla/Mod	Mod/Cla
TASA=Alta	0.00000000	0.000000
MONTO=Entre_1_y_3_millones	0.00000000	0.000000
MONTO=Entre_3_y_5_millones	0.00000000	0.000000
SEGMENTO=MASIVO	0.06441224	1.025641
TASA=Limite_Usura	0.10200612	1.538462
MONTO=Entre_5_y_10_millones	0.37174721	1.025641
TIPO.DE.CLIENTE=ANTIGUO OVN	1.72910663	3.076923
TASA=Media_Alta	2.62295082	4.102564
MONTO=Entre_10_y_20_millones	3.04878049	7.692308
OCUPACIÓN=Empleado Privado	4.31654676	70.769231
SATISFACCION=Satisfecho	5.09059534	90.769231
OCUPACIÓN=Independiente	12.30769231	8.205128
SATISFACCION=Neutral	12.61261261	7.179487
OCUPACIÓN=Pensionado	13.63636364	12.307692
TASA=Media_Baja	18.36734694	18.461538
MONTO=Entre_20_y_50_millones	22.96511628	40.512821
SEGMENTO=PREFERENTE	35.54327808	98.974359
TIPO.DE.CLIENTE=NUEVO SIN OFICINA	50.00000000	3.076923
MONTO=Más_de_100_millones	52.38095238	5.641026
OCUPACIÓN=Desempleado	75.00000000	4.615385
TASA= Muy_Bajas	98.66666667	75.897436
MONTO=Entre_50_y_100_millones	100.00000000	45.128205

Tabla 16: Descripción - cluster3

En el clúster 3 se observan que para las personas que tuvieron préstamos entre 50 y 100 millones (45,12%), préstamos entre 20 y 50 millones (40.51%) a tasas de interés muy bajas (75,89%), la mayoría de estos clientes pertenecen al segmento preferente (98.97%), son empleados privados (70.76%), seguido de pensionados con un (12.30%) y adicionalmente se encuentran satisfechos en un (90.76%).

6 CONCLUSIONES.

- En cuanto al número óptimo de clusters y por tanto de segmentos óptimo es tres.
- El cluster 1 son clientes que solicitan créditos entre 1 y 10 millones de pesos, a los que se les cobra tasas de interés cercanas a la usura, pertenecen al segmento masivo.
- El cluster 2 se caracteriza porque son clientes que solicitan créditos entre 10 y 50 millones de pesos, que pertenecen tanto al segmento masivo como al preferente y son empleados privados.
- En el cluster 3 se encuentran clientes que solicitan préstamos entre 20 y 100 millones de pesos, a tasas de interés muy bajas y que pertenecen al segmento preferente.
- Los principales factores de agrupamiento son la tasa de interés, el monto desembolsado y el segmento dentro del cual el banco clasifica a sus clientes.
- En cuanto al nivel de satisfacción se puede percibir que a pesar de que es un factor relevante para el banco, los clientes no lo ven como un elemento determinante a la hora de adquirir un nuevo producto de crédito para el banco.
- El segmento al que pertenece cada cliente presenta una alta correlación con la variable tasa de interés ya que los clientes que se encuentran en el segmento preferente presentan tasas de interés bajas pero montos altos, y los que pertenecen al segmento masivo presentan tasa de interés consideradas altas pero montos desembolsados bajos.
- Mediante este estudio se logró identificar un nuevo grupo que se encuentra conformado en el clúster 2, ya que en él participan personas que tienen tasas de interés intermedias y que pertenecen al segmento preferente y masivo.

7 RECOMENDACIONES

- Actualmente los clientes se segmentan de acuerdo al nivel de ingresos. Se propone que el modelo sea utilizado para una nueva segmentación de acuerdo al monto desembolsado para focalizar demanda de otros productos que el banco pueda ofrecerle a los clientes
- Este estudio solo se practicó en créditos de libre inversión. Se propone aplicarlo a todos los productos del banco.
- Proponer una nueva investigación que incluya no solo a los clientes actuales que se les aprobó el producto, sino también a los clientes que se les negó el crédito. De tal manera que se logre identificar a la totalidad de los clientes.

8 GLOSARIO

- **Segmento:** Grupo homogéneo y diferenciado de clientes a los que se les dirige una política comercial determinada. En el banco estudiado existe una segmentación previa donde se diferencia entre clientes con ingresos entre 1 salario mínimo mensual legal vigente hasta 6 salarios mínimos mensuales legales vigentes. Sin embargo, en el presente trabajo se busca identificar otros posibles segmentos que permitan clasificar de una forma más precisa a los clientes del banco. En este orden de ideas, cuando se utilice la palabra segmento como variable se estará hablando de la clasificación previa del banco; y cuando se utilice en otro contexto se hablará del proceso de segmentación que se pretende hacer con este trabajo.
- **Ovn:** Oferta Valor de Nómina. Cliente que tiene cuenta de nómina con el banco y productos preaprobados.
- **Crédito de libre inversión:** Producto financiero que se basa en prestarle un monto de dinero a un cliente que puede destinar libremente
- **Cluster:**
- **Dendograma:** Método de representación gráfica que muestra la relación entre los datos de una forma jerárquica estilo diagrama de árbol.
- **Desembolsar:** Entregar una cantidad determinada de dinero
- **Colocación:** Acción de poner a circular dinero en la economía a través de la asignación de créditos.
- **Desempleado:** Persona que se encuentra en situación de paro forzoso y está buscando empleo
- **Empleado privado:** Persona que se encuentra vinculada laboralmente a una empresa u organización privada, es decir, que no se encuentra vinculada al estado, y de esta actividad percibe más del 80% de sus ingresos.
- **Empleado público:** Persona que se encuentra vinculada laboralmente a alguna entidad del estado, y de esta actividad percibe más del 80% de sus ingresos.

- **Estudiante:** Persona cuya ocupación principal es cursar estudios en un establecimiento de enseñanza.
- **Hogar:** Persona que no realiza ningún tipo de actividad económica, no está buscando trabajo y se dedica principalmente a la administración del hogar
- **Independiente:** Persona que realiza una actividad comercial, industrial o de servicios por cuenta propia, de la cual proviene más del 20% de sus ingresos. No incluye a personas que derivan su ingreso por la prestación de servicios de transporte.
- **Pensionado:** Persona que percibe un ingreso proveniente de una pensión, sea porque ha cumplido su ciclo laboral o porque tiene algún tipo de incapacidad que le impide trabajar.
- **Rentista de capital:** Persona natural que obtiene la mayor parte de sus ingresos por los rendimientos financieros provenientes de alguna inversión
- **Transportador:** Trabajador independiente que devenga la mayor parte de sus ingresos por la prestación de servicios de transporte
- **Usura:** Máxima tasa de interés permitida por el gobierno colombiano para el otorgamiento de créditos.

9 REFERENCIAS.

- Díaz, L., & Morales, M. (2012). *Análisis Estadístico de Datos Multivariados*. Bogotá, Colombia: Universidad Nacional de Colombia.
- Dutka, A. (1998). *Manual de AMA Para la Satisfacción del Cliente*. 1994: Granica.
- Langrand, C. (2009). *Análisis de Datos: Métodos y Ejemplos*. Bogotá: Escuela Colombiana de Ingeniería.
- Lebart, L., Morineau, A., & Fénelon, J. (1985). *Tratamiento Estadístico de Datos*. Barcelona: Marcombo-Boixareu Editores.
- Prado, C. (1992). *Análisis de la Aplicación del Método de Ward de Clasificación Jerárquica al Caso de Variables Cualitativas*. Bogotá: Universidad Nacional de Colombia.
- Rencher, A. (1934). *Methods of Multivariate Analysis*. Provo: Wiley-Interscience.
- Schnaars, S. P. (1993). *Estrategias de marketing: un enfoque orientado al consumidor*. Ediciones Díaz de Santos.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Willey and Sons.
- Superintendencia Financiera de Colombia. (Julio de 2015). *Indicadores Gerenciales - NIIF*.
Obtenido de
<https://www.superfinanciera.gov.co/jsp/loader.jsf?lServicio=Publicaciones&lTipo=publicaciones&lFuncion=loadContenidoPublicacion&id=10084493>
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. New York: Springer-Verlag.