



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**SEGMENTACIÓN ESTRATÉGICA BANCARIA**

**REPORTE DE TRABAJO PROFESIONAL**

**QUE PARA OBTENER EL TÍTULO DE:  
ACTUARIO**

**P R E S E N T A :**

**JOSÉ ADRIÁN SUÁREZ RANGEL**



**TUTORA**

**MAT. MARGARITA ELVIRA CHAVEZ CANO**

**2015**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno

Suárez  
Rangel  
José Adrián  
55 5688 8684  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
Número de Cuenta: 096534474

2. Datos del Tutor

Matemático  
Margarita Elvira  
Chávez  
Cano

3. Datos del sinodal 1

Actuario  
Jaime  
Vázquez  
Alamilla

4. Datos del sinodal 2

Actuario  
Francisco  
Sánchez  
Villareal

5. Datos del sinodal 3

Actuario  
Rosa Daniela  
Chávez  
Aguilar

6. Datos del sinodal 4

Actuario  
Ángel Manuel  
Godoy  
Aguilar

7. Datos del trabajo escrito

Segmentación estratégica Bancaria  
33 páginas  
2015

## Índice

1. Introducción .....	4
2. Historia y Origen del Problema .....	6
2.1. El grupo Internacional .....	6
2.2. Una sola forma de trabajo .....	6
3. Segmentación Estratégica .....	8
3.1. Tipos de Segmentación en la Banca multinacional .....	8
3.2. Usos de la segmentación en la Banca multinacional .....	9
3.3. Procesos de Segmentación más comunes en la banca multinacional .....	10
3.4. Segmento de clientes de Banca Personal .....	11
4. Modelo Estadístico de Segmentación. El Método de K-medias .....	13
4.1. Método K-medias .....	15
4.2. Propiedades del Algoritmo K-medias.....	16
4.3. Algoritmo K-medias (planteamiento Intuitivo).....	17
4.4. Método (desarrollo formal).....	17
5. Proceso SEMMA para el manejo de SAS en un proyecto de Minería de Datos .....	19
6. Caso Práctico.....	21
6.1. Modelo BBVA Perú.....	21
7. Factores críticos de éxito .....	31
8. Conclusiones .....	34
Glosario.....	35
Bibliografía.....	36

# Segmentación Estratégica Bancaria.

## 1. Introducción

En este reporte se expone el problema de segmentación de clientes del Banco BBVA desde su origen, el módulo estadístico utilizado que está dentro del Análisis Multivariado, el Análisis de Conglomerados y específicamente el método K-Medias, y lo que es propiamente la experiencia profesional en el banco BBVA, finalmente el análisis y conclusiones al respecto.

El Banco Bilbao Vizcaya Argentaria (BBVA) es un banco líder a nivel mundial que ha sabido sortear la crisis financiera actual teniendo rendimientos constantes en este periodo tan complicado.

Uno de los principales problemas identificado por la dirección es el de no contar con un modelo estándar de segmentación a nivel grupo que permita implantar modelos de gestión diferenciados a segmentos distintos.

El reto es grande considerando que actualmente se tienen alrededor de 15 modelos de segmentación distintos con criterios, segmentos y umbrales diferentes.

El enfoque estadístico de separar estos grupos está basado en el análisis de Conglomerados, que es una técnica cuyo proceso es formar grupos a partir de un grupo de elementos. Tales grupos deben estar compuestos por elementos lo más parecido posible dentro de cada grupo y además lo mas diferente que sea posible entre grupos.

En la sección 2 se planteará el problema que se tuvo al ser BBVA el resultado de una fusión de distintas entidades bancarias. Mientras que en la sección 3 se abordará el tema desde un punto de vista comercial, en la sección 4 se hará desde un punto de vista estadístico. En la sección 5 se revisará la metodología recomendada por SAS para la elaboración de modelos de Minería. La sección 6 presenta el caso práctico de la segmentación de Perú y los distintos escenarios

que se revisaron como alternativas de solución. En la sección 7 se enuncian, desde un punto de vista administrativo, las consideraciones que deben tenerse para que el proyecto tenga éxito. Finalmente en la sección 8 se hace una breve conclusión de todo el proceso de segmentación y de sus resultados.

## **2. Historia y Origen del Problema**

De acuerdo a la página del banco (<http://www.bbva.com>) la historia de BBVA comienza en 1857 en Bilbao, una ciudad situada al norte de España, cuando la Junta de Comercio promueve la creación de Banco de Bilbao como banco de emisión y descuento.

Durante el siglo XX el Banco de Bilbao continua con su crecimiento en España formando un consorcio con el banco Vizcaya.

BBV y Argentaria anuncian su proyecto de fusión el 19 de octubre de 1999, con el objetivo de seguir adquiriendo valor y un mayor potencial de crecimiento de beneficios.

### **2.1. El grupo internacional**

Es a partir de 1995, cuando el BBVA desarrolla una estrategia de expansión internacional que lo llevó a construir una gran franquicia en América Latina, realizando una importante inversión en capital, tecnología y recursos humanos.

En 2004, el Grupo inicia otra línea de expansión internacional con la entrada al mercado estadounidense y la adquisición de una cantidad considerable de acciones del principal banco de Turquía.

### **2.2. Una sola forma de trabajo**

Con esta fusión de distintas Instituciones Financieras con distintas formas de hacer banca, distintas visiones y tecnologías el reto que se planteó fue como establecer un mismo modelo de negocio para todos.

Bajo la visión de “Hacer poco pero hacerlo bien”, que es una práctica corporativa del BBVA, se crean tres iniciativas en las cuales se va a trabajar durante los siguientes 3 años a partir de 2012:

- Segmentación Estratégica.

- Atención diferenciada a cada segmento.
- Impulso a los distintos canales digitales.

Como primer punto debe atacarse el problema de la segmentación, que si bien cada institución cuenta con un modelo propio, se debe trabajar en estandarizar los segmentos, criterios y umbrales que definen a cada colectivo para extraer el máximo potencial de cada uno de ellos. Esta decisión fue tomada por el Comité de Dirección e informada a todo el grupo en el año 2012.



### **3. Segmentación Estratégica**

Según la consultora Cognodata, empresa de origen Español y que lidera el mercado de análisis de segmentación de instituciones financieras, el diferenciar correctamente a los segmentos y crear modelos de gestión diferenciada nos permitirá:

- a) Por un lado, la adaptación de la oferta de servicios y del modelo de relación a segmentos de clientes con necesidades financieras diferentes mejora el ritmo de captación, la vinculación y la fidelidad.
- b) Por otro lado, la identificación de clientes con diferencias importantes de valor y potencial, permite dirigir la capacidad comercial ahí donde existe un mayor retorno en valor de cliente.

La segmentación Estratégica debe cumplir las siguientes condiciones:

- Identificar segmentos con necesidades financieras diferenciadas.
- Identificar segmentos con rentabilidad, fidelidad y potencial diferente.
- Es necesario que a los segmentos se les pueda asociar alguna estrategia de negocio.
- El número de segmentos debe ser reducido y limitado por la complejidad de gestión.
- El tamaño de cada segmento y su valor deben justificar un tratamiento diferenciado.

### 3.1. Tipos de Segmentación en la Banca multinacional

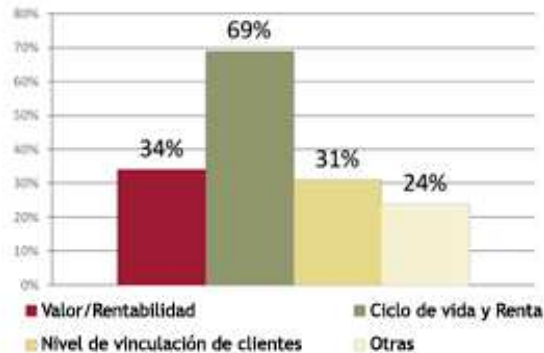
Según un estudio realizado por Cognodata en 36 entidades financieras multinacionales y españolas para finales del año 2011<sup>1</sup>, más del 90% de las entidades tiene una segmentación corporativa para clientes particulares.

1

¿Existe en su organización una segmentación corporativa para clientes particulares?



¿Qué criterios se han usado para realizar la segmentación?



El 60% de las entidades utilizan segmentaciones con un sólo criterio, mientras que el resto trabajan con una combinación. Entre estas combinaciones, la más habitual es ciclo de vida y renta de sus clientes.

### 3.2. Usos de la segmentación en la Banca multinacional

De acuerdo al mismo estudio antes mencionado en el 96% de los casos se utiliza para desarrollar y lanzar campañas, mientras que en el 85% se utiliza para desarrollar productos especializados, le siguen el 65% que las utiliza para definir responsables por segmentos y el 31% para desarrollar una red especializada en algún nicho como Banca Personal o Pymes.

<sup>1</sup> [http://www.cognodata.com/newsletter\\_sector\\_financiero5.asp](http://www.cognodata.com/newsletter_sector_financiero5.asp)

Lo que llama la atención es que el 37% de los bancos estudiados aun no utilizan dentro de sus planes de mercadotecnia a la segmentación. Como ya dijimos antes, este plan es la guía estratégica para desarrollar todo el potencial de la Segmentación Estratégica.

### **3.3. Procesos de Segmentación más comunes en la banca multinacional**

Para algunas empresas de bienes de consumo es de vital importancia contar con datos concretos de los clientes, que al no contar con ellos, recurren a construir segmentaciones complejas. En cambio, la abundancia de esta información en el sector financiero no se está aprovechando a pesar de la eficacia probada de las técnicas de segmentación y mercadotecnia desde hace décadas.

Actualmente los bancos, a diferencia de otras entidades, cuentan con información muy rica, conocen a sus clientes, su rentabilidad, su potencial para productos y servicios, fidelidad, vinculación, etc. Sin embargo esta información esta subutilizada o limitada, dejando de lado la información externa, de mercado que profundiza en tendencias, actitudes, elecciones de contratación, etc.

### 3.4. Segmento de clientes de Banca Personal

Este colectivo de clientes requiere de una atención especial que responda a las necesidades que tiene en términos de servicio, rentabilidad y sofisticación financiera. Para ellos será necesario incluir variables específicas como pueden ser:



Gracias a estas variables obtenemos los segmentos “Familias maduras, inversionistas sofisticados y multicanal con elevado saldo en otras entidades” o “Jóvenes hipotecados vinculados, con gusto por alta rentabilidad en mercados”.

La segmentación estratégica sólo es el primer paso de un camino difícil hacia la gestión por segmentos, en la que es necesario involucrar a los departamentos de gestión comercial y mercadotecnia. Si no se tiene en cuenta y se planea desde el principio la forma de implantar esta segmentación en la gestión real se llegará irremediablemente al fracaso.

Se deben considerar entonces los siguientes puntos:

1. Plan comercial de mercadotecnia diferenciado por segmento: campañas, objetivos, especialización.

2. Sistemática diferencial: despliegue por zonas y oficinas, estilo comercial y protocolo de atención por segmento.
3. Oferta de productos: planes especiales o paquetes de productos.
4. Nivel de servicio diferenciado: formarse en diferentes filas de acuerdo a segmento, distintos tiempos de solución, distintos canales y oficinas.
5. Estrategia segmentada de canales.
6. Política segmentada de precios y promociones.

La práctica más común de segmentación de clientes en el grupo BBVA es a través del estudio de conglomerados, específicamente en el método de K-medias. Esto es debido principalmente a la facilidad que se tiene de explicar este método a la dirección, la desconfianza por aplicar cosas que no entienden ha sido un problema en la implantación de modelos más complicados.

#### 4. Modelo Estadístico de Segmentación. El Método de K-medias

El análisis de conglomerados es una técnica que se usa en minería de datos<sup>2</sup>, agrupa los objetos en varios grupos o cúmulos de tal forma que los objetos dentro de cada grupo se parecen mucho, en este caso se dice que son similares, pero no se parecen a los objetos de otros grupos. Se habla de similitudes y disimilitudes entre los objetos, y estas se establecen de acuerdo a los atributos que describen a los mismos. Los algoritmos de agrupamiento se usan para organizar, para categorizar, para la comprensión de los datos y construcción de modelos, para detectar observaciones discordantes, etc.

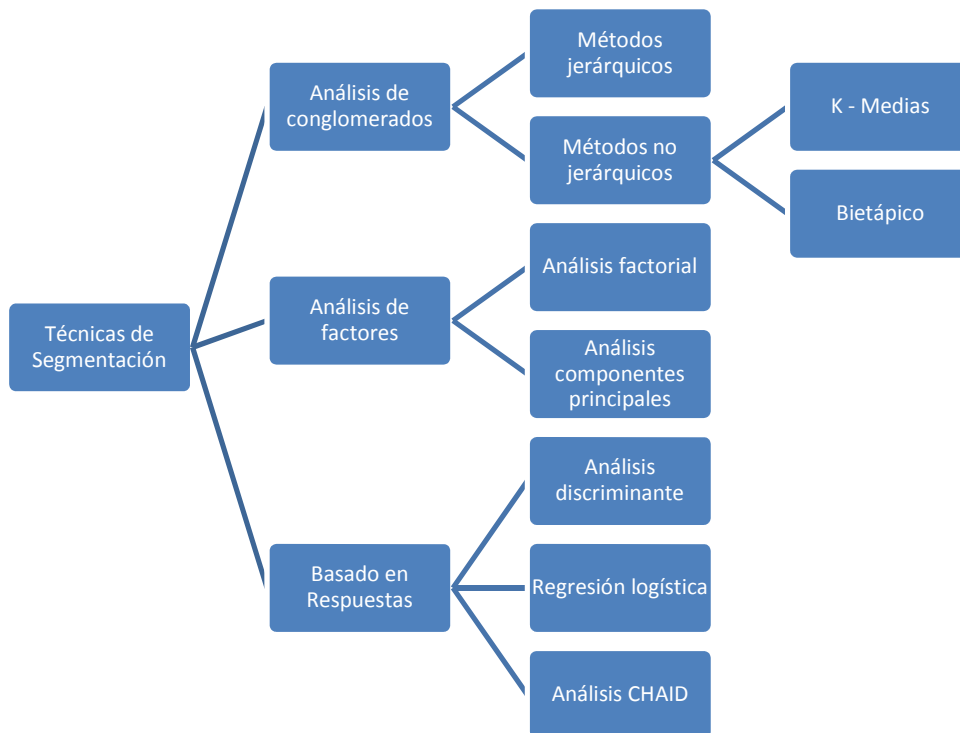
El análisis de conglomerados puede usarse como una herramienta autónoma de minería de datos para obtener conocimiento de la distribución de los datos, o como un paso previo de proceso para otros algoritmos de minería de datos que operan sobre los grupos detectados. Se han desarrollado muchos algoritmos de agrupamiento y están categorizados en diferentes aspectos tales como métodos de particionar, métodos jerárquicos, métodos basados en densidades y métodos basados en cuadros o redes. Además los datos pueden ser numéricos o categóricos. Las propiedades geométricas inherentes de los datos numéricos pueden ser explotadas para definir de forma natural la función distancia entre los puntos de datos, mientras que los datos categóricos pueden derivarse de datos cuantitativos o cualitativos donde las observaciones son tomadas directamente de conteos.

Entre las técnicas de agrupamiento de datos está el método de K-medias que es el método que se usa en la minería de datos. Esta técnica normalmente es fácil de explicar a los gerentes de segmentos y, al ser posible su interpretación gráfica, es de fácil implementación a las ofertas de los productos del banco.

---

<sup>2</sup> La minería de datos o exploración de datos es un campo de las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior.

La tarea de minería de datos real es el análisis automático o semi-automático de grandes cantidades de datos para extraer patrones interesantes hasta ahora desconocidos, como los grupos de registros de datos (análisis clúster), registros poco usuales (la detección de anomalías) y dependencias (minería por reglas de asociación). Estos patrones pueden entonces ser vistos como una especie de resumen de los datos de entrada, y pueden ser utilizados en el análisis adicional o, por ejemplo, en la máquina de aprendizaje y análisis predictivo.



Los algoritmos de formación de conglomerados se agrupan en dos categorías:

- Algoritmos de partición: Método de dividir el conjunto de observaciones en  $k$  conglomerados (clusters), en donde  $k$  lo define inicialmente el usuario.
- $f$  Algoritmos jerárquicos: Método que entrega una jerarquía de divisiones del conjunto de elementos en conglomerados.

Un método jerárquico aglomerativo parte con una situación en que cada observación forma un conglomerado y en sucesivos pasos se van uniendo, hasta que finalmente todas las situaciones están en un único conglomerado.

Un método jerárquico disociativo sigue el sentido inverso, parte de un gran conglomerado y en pasos sucesivos se va dividiendo hasta que cada observación queda en un conglomerado distinto.

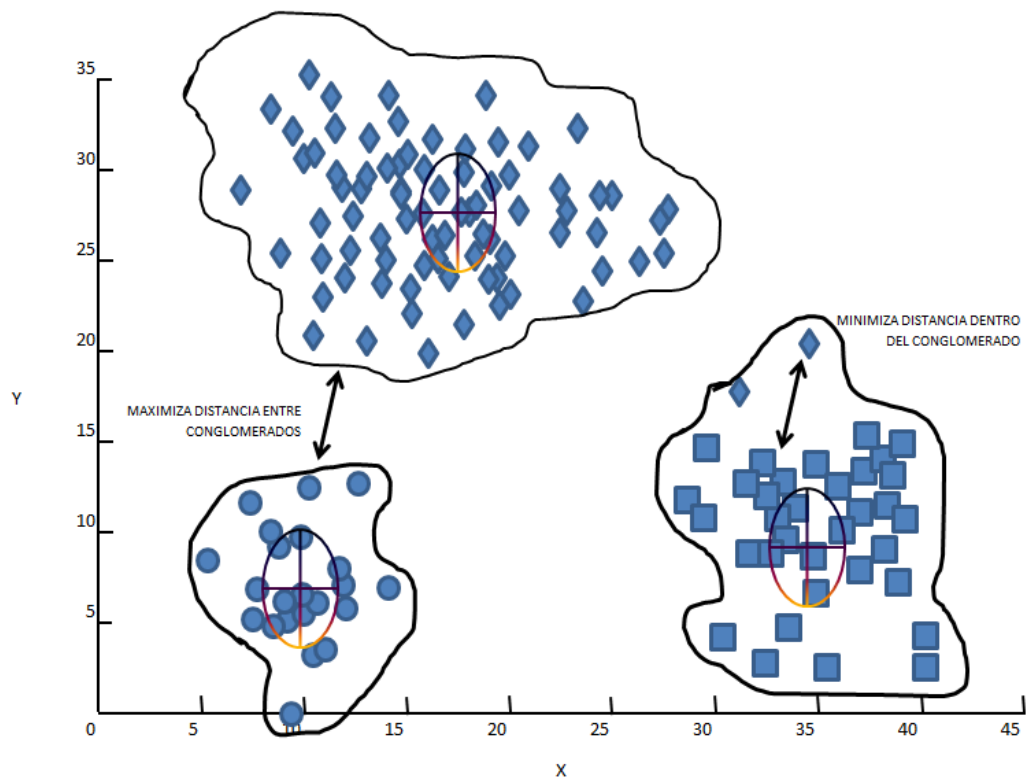
#### **4.1. Método K-medias**

- Es un método de partición que obtiene grupos mutuamente ajenos de forma esférica.
- Genera un número específico de grupos ajenos y no jerárquicos.
- Organiza los objetos en k-particiones donde cada partición representa un grupo.
- Se inicia con un conjunto de medias y se clasifican los casos basados en sus distancias con sus centros.
- A continuación se calculan las medias de los grupos nuevamente, usando los casos que están asignados a los grupos.
- Luego, se vuelven a clasificar todos los casos basados en el nuevo conjunto de medias de los grupos.
- Se repite este paso hasta que las medias de los grupos no cambian entre pasos sucesivos.
- Finalmente se calculan las medias de los grupos nuevamente y se asignan los casos a sus grupos permanentes.



## 4.2. Propiedades del Algoritmo K-medias

- Siempre hay  $K$  grupos. El número  $K$  se elije arbitrariamente.
- Siempre hay al menos un elemento en cada grupo.
- Los grupos son no- jerárquicos y no se traslapan.
- Todo miembro de un grupo está más cercano a su grupo que a cualquier otro grupo debido a que la cercanía no siempre involucra al centro de los grupos



Agrupación de observaciones por el metodo k-medias

### 4.3. Algoritmo K medias (planteamiento intuitivo)

- El conjunto de datos es particionado en K grupos y los datos son asignados aleatoriamente a los grupos dando como resultado grupos que tienen el mismo número de datos aproximadamente.
- Para cada dato:
  - o Calcular la distancia desde este dato a cada grupo.
    - Si el punto está más cercano a su propio grupo, dejarlo donde está.
    - Si el dato no está más cercano a su propio grupo, moverlo al grupo más cercano.
- Repetir el paso anterior hasta que una pasada completa a través de todos los datos de cómo resultado que no se mueve ningún dato de un grupo a otro. En este punto los grupos son estables y el proceso de agrupamiento termina.

### 4.4. Método (desarrollo formal)

El algoritmo requiere como entrada una matriz de M puntos en N dimensiones y una matriz de K centros de grupos iniciales en N dimensiones. El número de puntos en el grupo L se denota por  $NC(L)$ .  $D(I, L)$  es la distancia euclidiana entre el punto I y el grupo L. El procedimiento general es la búsqueda de una K-partición con una suma de cuadrados localmente óptima dentro de cada grupo moviendo los puntos de un grupo a otro.

Paso 1: Para cada punto I ( $I=1, 2, \dots, M$ ), se obtienen los centros de los grupos más cercanos y los segundo más cercanos,  $IC1(I)$  y  $IC2(I)$ , respectivamente. Se asigna punto I al grupo  $IC1(I)$ .

Paso 2: Se actualizan de los centros de los grupos como los promedios de los puntos contenidos en ellos.

Paso 3: En un principio, todos los grupos pertenecen al conjunto vivo.

Paso 4: Esta es la etapa de transferencia óptima (OPTRA): Considere cada punto I ( $I=1, 2, \dots, M$ ) en turno. Si el grupo L ( $L=1, 2, \dots, K$ ) se actualiza en la última etapa de transferencia rápida (QTRAN), entonces pertenece al conjunto vivo durante toda esta etapa. De lo contrario, en cada paso, que no está en conjunto vivo, si no se ha actualizado en los últimos M pasos de transferencia optima. Deje que el punto I este en el grupo LI. Si LI esta en el conjunto vivo, haga el Paso 4a; de lo contrario, hacer el paso 4b.

Paso 4a. Calcule el mínimo de la cantidad,  $R2 = [NC(L) * D(I, L)^2] / [NC(L)+1]$ , sobre todos los grupos L ( $L \neq LI, L=1, 2, \dots, K$ ). Sea L2 el grupo con la menor R2. Si este valor es mayor o igual a  $[NC(LI) * D(I, LI)^2] / [NC(LI)-1]$ , no es necesario

reasignar y L2 es la nueva IC2 (I). (Tenga en cuenta que el valor  $[ NC ( L1 ) * D ( I , LI )^2 ] / [ NC ( LI ) - 1 ]$  es fijo y seguirá siendo el mismo para el punto I hasta que se actualice el grupo LI). De lo contrario, el punto I es asignado a clúster L2 y LI es la nueva IC2 (I). Los centros de los grupos se actualizan para que sean las medias de los puntos asignados a ellos si la reasignación ha tenido lugar. Los dos grupos que están involucrados en la transferencia del punto I en este paso en particular se encuentran ahora en el conjunto vivo.

Paso 4b. Este paso es el mismo que el Paso 4a, excepto que el R2 mínimo se calcula sólo sobre los grupos en el conjunto vivo.

Paso 5: Detener si el conjunto vivo está vacío. De lo contrario, vaya al paso 6 después de un recorrido a través del conjunto de datos.

Paso 6: Esta es la etapa de transferencia rápida ( QTRAN ) : Considere cada punto I (  $1 = 1 , 2 , \dots , M$  ) en turno . Sean  $LI = IC1 ( I)$  y  $L2 = IC2 ( I)$ . No es necesario comprobar el punto I si ambos grupos LI y L2 no han cambiado en los últimos M pasos.

Calcule los valores de

$$RI = [ NC ( L1 ) * D ( I , L1 )^2 ] / [ NC ( L1 ) - 1 ] \text{ y}$$

$$R2 = [ NC ( L2 ) * D ( I , L2 )^2 ] / [ NC ( L2 ) + 1 ]$$

(Como se señaló anteriormente, RI se fija y seguirá siendo la misma hasta que se actualice el grupo LI.) Si RI es menor que R2, el punto I permanece en el grupo L1. De lo contrario, cambiar ICL (I) e IC2 (I) y actualizar los centros de los cúmulos LI y L2.

Los dos grupos también son destacados por su implicación en una transferencia en este paso.

Paso 7: Si no se llevó a cabo transferencia en los últimos M pasos, vaya al Paso 4 De lo contrario, vaya al Paso 6.

## 5. Proceso SEMMA para el manejo de SAS en un proyecto de Minería de Datos

El acrónimo SEMMA – sample (muestreo), explore (explorar), modify (modificar), model (modelar), assess (evaluar) - se refiere al proceso central de llevar a cabo la minería de datos. A partir de una muestra estadísticamente representativa de los datos, SEMMA hace que sean fácil de aplicar técnicas estadísticas y de visualización de exploración, selección y transformación de las variables de entrada más importantes, el modelado de las variables para predecir los resultados y confirmar la precisión del modelo.

Antes de examinar cada etapa del SEMMA, revisemos un malentendido común, que es referirse a SEMMA como metodología de minería de datos. SEMMA no es una metodología de minería de datos, sino más bien una organización lógica del sistema de herramienta funcional de SAS Enterprise Miner para la realización de las tareas principales de la minería de datos. Enterprise Miner se puede utilizar como parte de cualquier metodología de minería de datos iterativo adoptado por el cliente. Naturalmente pasos tales como la formulación de un negocio bien definido o problema de investigación y montaje de fuentes de datos representativos de calidad son fundamentales para el éxito global de cualquier proyecto de minería de datos. SEMMA se centra en los aspectos de desarrollo de modelos de minería de datos:

*Muestreo* (opcional) de sus datos mediante la extracción de una parte de un gran conjunto de datos suficientemente grande como para contener la información importante, pero lo suficientemente pequeño para manipular rápidamente.

*Explorar* sus datos mediante la búsqueda de tendencias y anomalías imprevistas a fin de obtener la comprensión y las ideas. Explorar ayuda a refinar el proceso de descubrimiento. Se deben de identificar los valores extremos que pueden representar en sí mismos un segmento.

*Modificar* sus datos mediante la creación, selección y transformación de las variables para concentrar el proceso de selección del modelo. Con base en sus descubrimientos en la fase de exploración, es posible que deba manipular sus datos para incluir información tal como la agrupación de los clientes y los subgrupos significativos, o para introducir nuevas variables.

*Modelar* los datos al permitir que el software busque automáticamente una combinación de datos que permite predecir de forma fiable el resultado deseado.

*Evaluar* sus datos mediante el análisis de la utilidad y fiabilidad de los resultados del proceso de minería de datos y estimar lo bien que funciona. Un medio común

de evaluación de un modelo es aplicarla a una parte de los datos por separado durante la etapa de muestreo. Si el modelo es válido, lo debe ser para esta muestra de prueba, así como para la muestra utilizada para construir el modelo.

Todos estos pasos fueron aplicados en el desarrollo del modelo.

## 6. Caso Práctico

Después de la fusión de distintas instituciones financieras en varios países y de tomar la decisión de realizar una sola segmentación para todo el grupo, el corporativo elaboró la propuesta de umbrales que definían a los segmentos basados en dos variables.

La primera variable es el nivel de ingresos de una persona. A esta variable se le ha llamado renta o nómina.

La segunda variable es el capital con el que un cliente cuenta y que esta invertido o depositado en cuentas de ahorro, vista, plazo fijo o cualquier otro instrumento en el que los recursos del cliente se encuentran en poder del banco. A esta variable se le ha llamado recursos o pasivo.

Después de recibir la propuesta en cada uno de los países se pidió ayuda al área de Inteligencia Comercial de América del Sur para ver si esta segmentación estaba alineada a las características de los clientes de BBVA en cada país.

Se revisó la forma en que el algoritmo K means asignaba a los clientes en distintos segmentos y, sobre todo los umbrales (importes mínimos) necesarios en cada variable para pertenecer al segmento.

### 6.1. Modelo BBVA Perú

El presente análisis ha sido hecho con los clientes de BBVA Perú.

Los umbrales que se proponen a nivel corporativo están dados como sigue:

Segmento	Perú	
	Monto Recursos*	Monto Renta*
Individuos de Muy Alto Valor (IMAV)	1MM USD = 2,8MM /s	
Adinerado	180K USD = 504k /s	5K USD = 14k /s
Medio alto	35K USD = 98k /s	2.5K USD = 7k /s
Masa	RESTO - NO RENT	RESTO - NO RENT
No Rentable	< 200 USD = 560 /s	< 400 USD = 1.1 /s

\*USD=Dolares Americanos

\*/s soles Peruanos

El segmento "No Rentable" se establece de acuerdo a cálculos de las áreas de negocio y financieras con los mínimos estimados para considerar que la gestión de los clientes en estos segmentos es rentable, de tal forma que estos individuos no participan en el ejercicio pues su segmento está dado por reglas de negocio y la estrategia a seguir con ellos ya está definida: incentivar el uso de medios de

transacción como son banco por teléfono, banco por internet o banca móvil que tienen los costos más bajos y dejar de atender por sucursales que tienen costos superiores.

Comenzaremos con el primer paso que se refiere a obtener una muestra de los datos, se tomó un 10% de la población para obtener velocidad en los procesos y en las consultas.

Análisis exploratorio de los datos.

Estadísticas Resumen.

Variable	Media	Desviación Estándar	Varianza	Mínimo	Máximo	Rango	N	N nulos	Percentil 99
PASIVO	8,568.43	144,582.2	20,904,013,426	0.010	108,118,792	108,118,792	1,722,327	263,981	13,390,337
NOMINA	713.188	1,914.17	3,664,045.92	0	622,556	622,556	19,57016	29,292	7,358

Se puede observar que los datos de ambas variables se distribuyen en rangos muy grandes, sin embargo el mayor porcentaje de ellos se sitúa en un intervalo menor.

Para efectos de poder analizar los datos de manera más simple y que no se generen segmentos aislados con pocos elementos que correspondan a los valores extremos, los eliminaremos del estudio y posteriormente los calificaremos con los valores que obtengamos del proceso para cada segmento para ambas variables.

Haremos coincidir el límite superior de la variable pasivo con el umbral sugerido para el segmento de individuos de muy alto valor que es de 2 Millones, 800 mil Soles. Por otro lado eliminaremos a los clientes que tienen una nómina mayor a 14 mil soles, que es el doble del percentil 99 para esta variable.

### Análisis de Correlación de las variables pasivo y nómina.

Coeficiente de correlación Pearson Prob >  r  bajo H0: Rho=0 Numero de Observaciones		
	NOMINA	PASIVO
NOMINA	1.00000	0.06841
NOMINA	1,957,016	<.0001
		1,722,325
PASIVO	0.06841	1.00000
PASIVO	<.0001	1,722,327
	1,722,325	

Se observa que la correlación entre las variables nómina y pasivo es baja, 6.8% (ver cuadro anterior). Esto da una idea de la conformación de los segmentos en donde no se esperaría encontrar un segmento con muchos miembros con ambas variables con valores medios, o altos.

También se puede observar que los rangos en los que se mueven ambas variables son distintos, por lo que para no influir con estos datos en el proceso de segmentación utilizaremos alguna transformación. Procederemos pues a estandarizar a la normal (0,1) ambas variables para que el proceso sea el responsable de evaluar la importancia de cada variable y no los datos mismos.

Revisaremos distintos modelos con diferente número de segmentos a fin de comparar entre ellos y poder seleccionar el que más sentido tenga desde el punto de vista de implementación comercial, esta puede ser la parte más delicada del proceso ya que aunque los distintos modelos son correctos estadísticamente, pueden no tener una manera de aplicarlos al negocio o puede no encontrarse la estrategia correcta para cada uno de ellos.



Modelo 1. K-medias 6 segmentos.

Estadísticas Resumen. Resultados

**Cluster 1**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	517,470.87	138,812.27	348,304.44	883,627.82	93
NOMINA	7,476.87	3,351.37	1,172.00	13,832.00	93

**Cluster 2**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	2,132,912.59	270,192.72	1,884,958.31	2,453,738.71	4
NOMINA	7,615.89	3,373.40	4,376.43	10,569.00	4

**Cluster 3**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	7,418.74	11,856.75	560.30	145,504.24	19,737
NOMINA	2,213.84	889.79	1,101.00	4,508.00	19,737

**Cluster 4**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	181,468.74	63,394.32	102,917.12	357,886.38	515
NOMINA	6,546.93	2,870.44	1,155.00	13,992.00	515

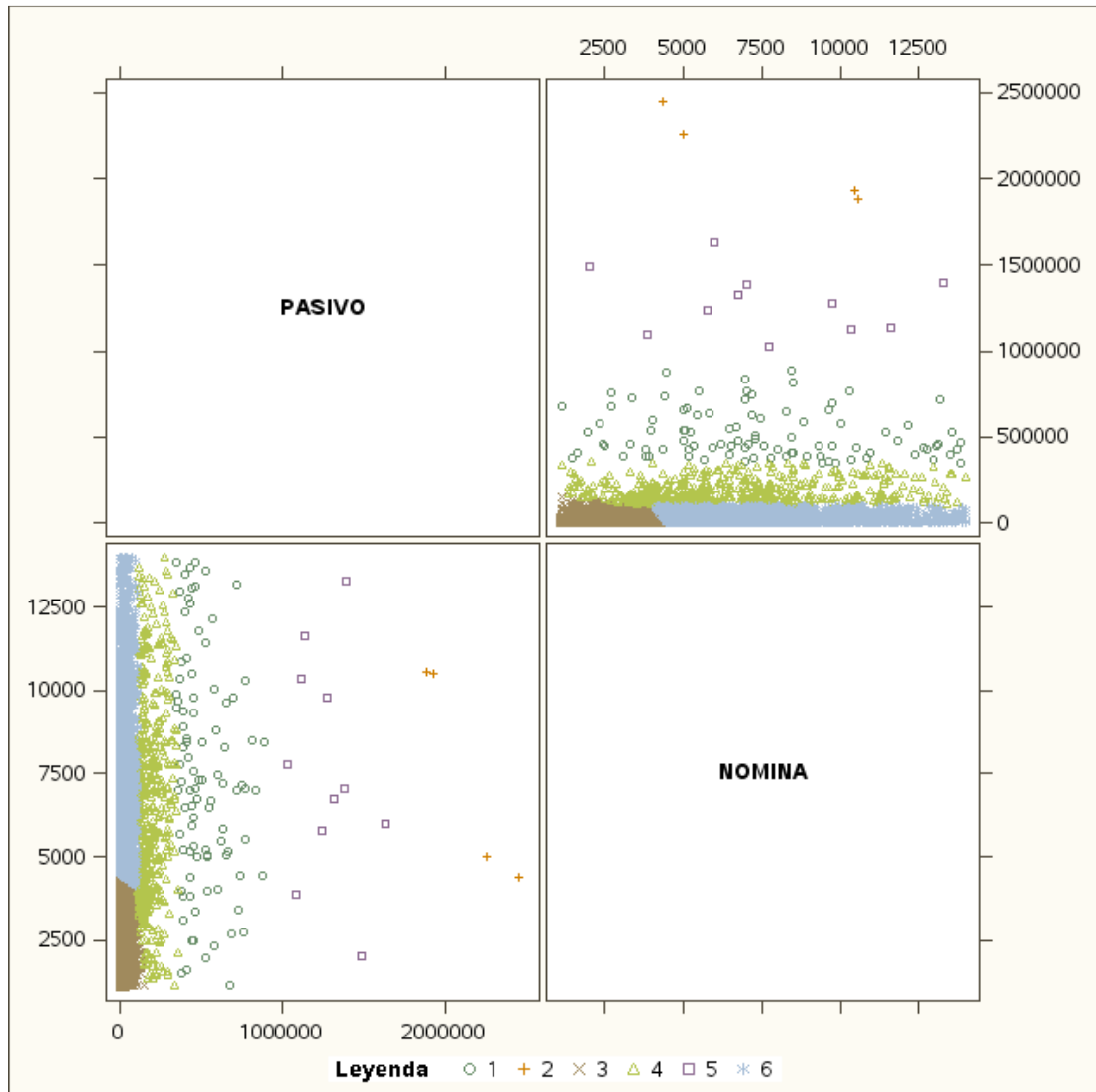
**Cluster 5**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	1,282,894.62	184,321.47	1,030,337.42	1,634,699.92	11
NOMINA	7,658.25	3,354.75	2,020.00	13,290.00	11

**Cluster 6**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	21815.14	23199.63	563.50	104545.14	4391
NOMINA	6675.96	2087.19	4131.00	14000.00	4391

## Matriz del trazado de dispersión



Matriz del trazado de dispersión de los segmentos obtenidos con método de K medias para 6 grupos, en donde cada color y figura representa un segmento.

Puede observarse en la gráfica de trazado, si vemos el tercer cuadrante (abajo izquierda), que para la variable nómina, en el eje vertical, se establece un sólo corte, mientras que para la de pasivo, en el eje horizontal, se establecen 4 cortes, el modelo le está dando mayor importancia a esta última variable, como es de esperar en estos casos de distribución de riquezas y rentas existen segmentos con pocos individuos y altos valores y segmentos con muchos individuos y valores bajos.

Modelo 2. K-medias 5 segmentos.

Estadísticas Resumen. Resultados.

#### Cluster 1

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	181,742.00	64,655.55	102,801.62	357886.38	523
NOMINA	6,566.82	2,873.41	1,155.00	13,992.00	523

#### Cluster 2

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	21,743.76	23,103.86	563.50	102,962.76	4,398
NOMINA	6,669.35	2,086.87	4,044.00	14,000.00	4,398

#### Cluster 3

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	7,404.31	11,803.58	560.30	145,504.24	19,737
NOMINA	2,212.58	888.53	1,101.00	4,503.00	19,737

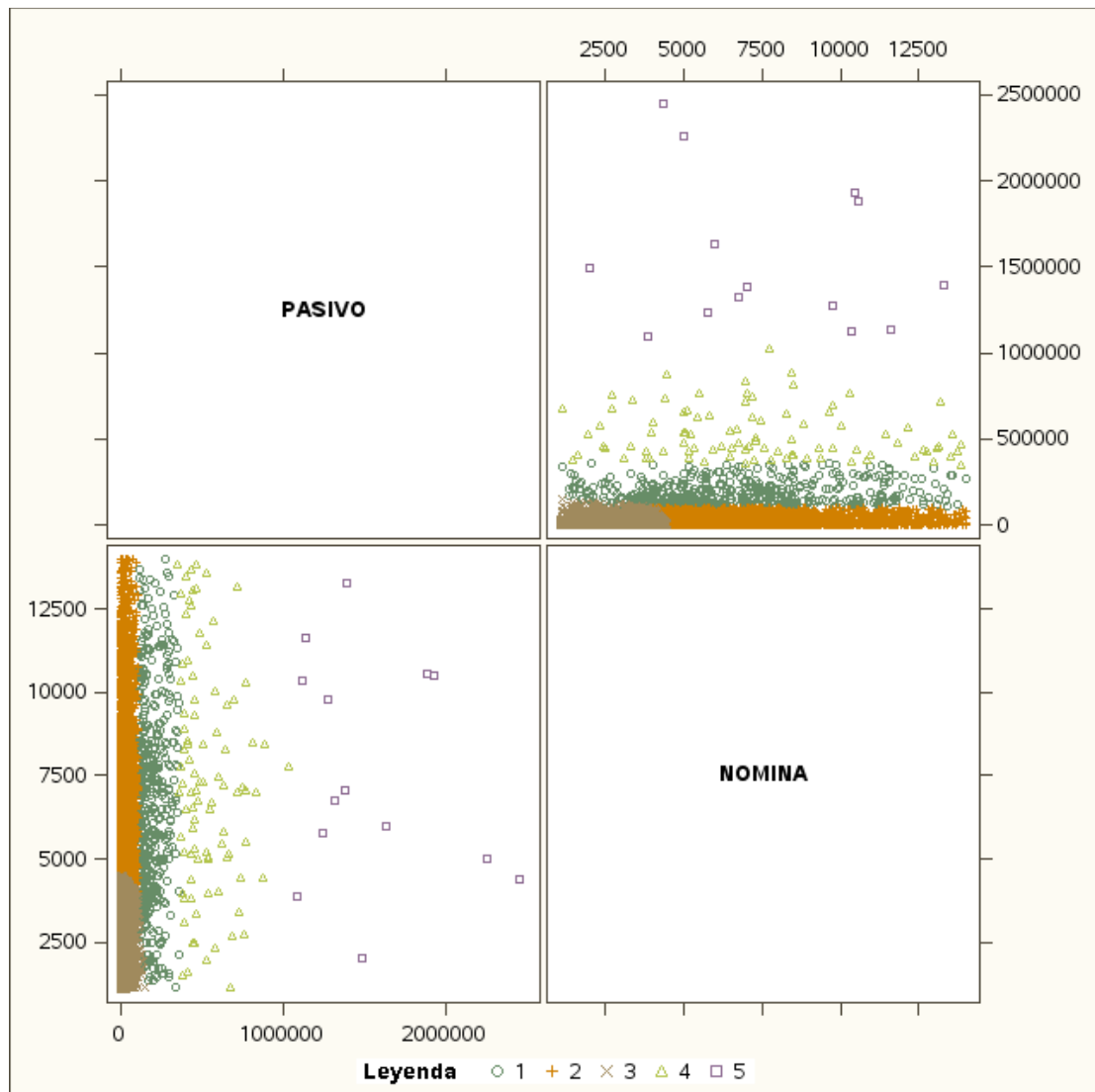
#### Cluster 4

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	528,520.68	146,959.84	348,304.44	1,030,337.42	91
NOMINA	7,407.91	3,364.05	1,172.00	13,832.00	91

#### Cluster 5

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	1,543,796.70	432,534.12	1,090,875.57	2,453,738.71	14
NOMINA	7,638.59	3,358.95	2,020.00	13,290.00	14

## Matriz del trazado de dispersión



Matriz del trazado de dispersión de los segmentos obtenidos con método de K medias para 5 grupos, en donde cada color y figura representa un segmento.

De la misma manera que en el caso anterior se observa que a la variable que el modelo le da más importancia es el pasivo, ahora tenemos un corte para nómina y tres para pasivo.

Modelo 3. K-medias 4 segmentos.

Estadísticas Resumen. Resultados.

**Cluster 1**

<b>Variable</b>	<b>Media</b>	<b>Desviación Estándar</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Observaciones</b>
PASIVO	1,509,566.08	437,376.86	1,030,337.42	2,453,738.71	15
NOMINA	7,646.95	3,236.93	2,020.00	13,290.00	15

**Cluster 2**

<b>Variable</b>	<b>Media</b>	<b>Desviación Estándar</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Observaciones</b>
PASIVO	341,446.88	152,599.68	183,917.20	883,627.82	279
NOMINA	7,256.05	3,068.69	1,155.00	13,992.00	279

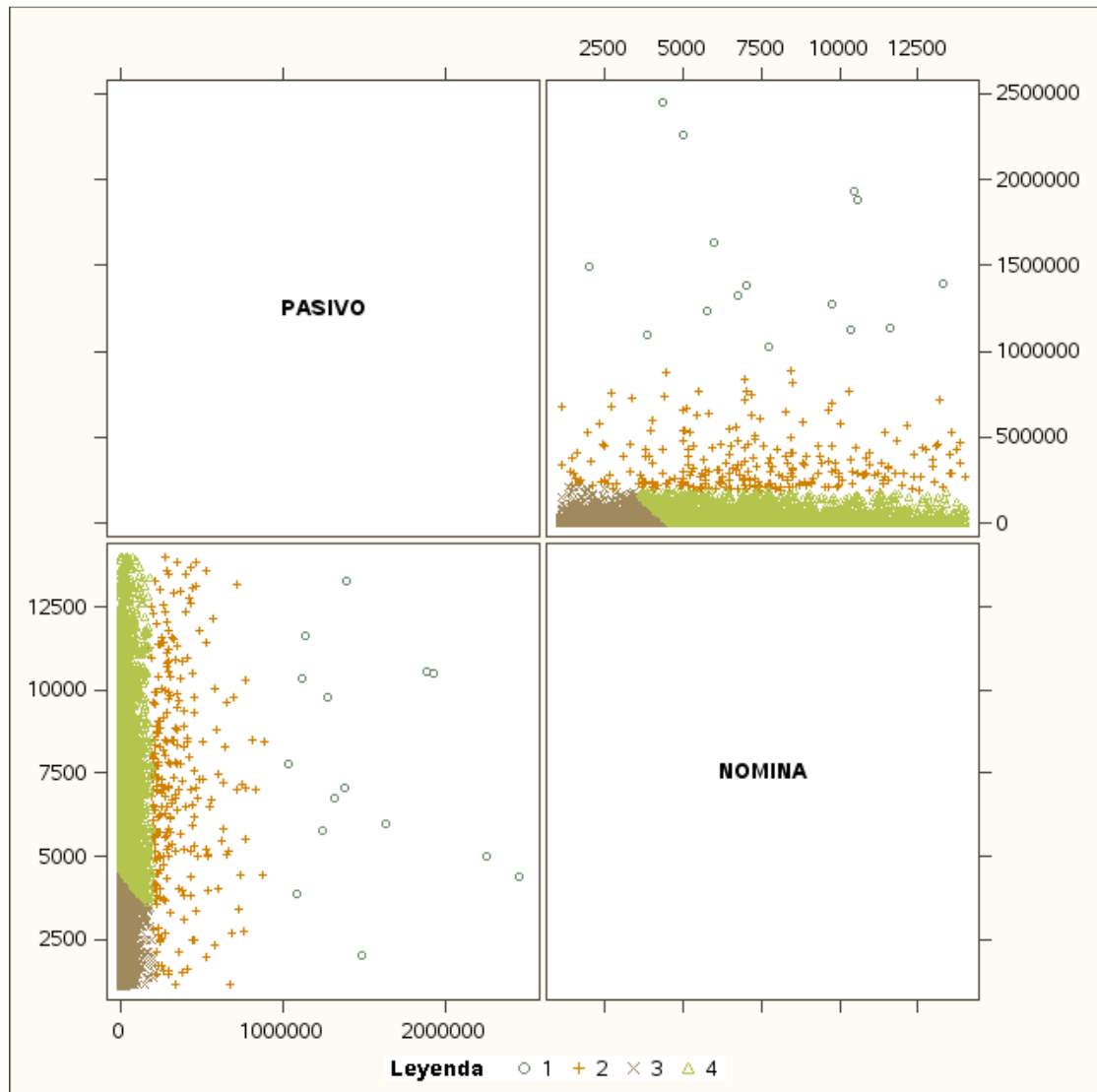
**Cluster 3**

<b>Variable</b>	<b>Media</b>	<b>Desviación Estándar</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Observaciones</b>
PASIVO	7,865.75	14,190.15	560.30	207,538.01	19,995
NOMINA	2,239.03	915.11	1,101.00	4,657.00	19,995

**Cluster 4**

<b>Variable</b>	<b>Media</b>	<b>Desviación Estándar</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Observaciones</b>
PASIVO	29,416.40	36,333.18	563.50	189,470.90	4,462
NOMINA	6,786.61	2,113.25	3,636.95	14,000.00	4,462

## Matriz del trazado de dispersión



Matriz del trazado de dispersión de los segmentos obtenidos con método de K medias para 4 grupos, en donde cada color y figura representa un segmento.

Tenemos ahora un corte en nómina y dos en pasivo.

El objetivo es encontrar dos cortes para cada variable y ver qué tanto difieren de los que está proponiendo el corporativo, es decir, trataremos de revisar qué tanto se parece el modelo global a la realidad de los clientes del banco de Perú.

Hagamos un ejercicio más, tomando un modelo que nos segmente sólo por nómina y otro sólo por pasivo y veamos los cortes que nos dividen de manera más heterogénea a los segmentos.

Modelo 3. K-medias 3 segmentos sólo Nómina.

**Cluster 1**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
NOMINA	9,076.85	1,814.37	6,922.00	14,000.00	1,819

**Cluster 2**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
NOMINA	4,613.82	960.14	3,306.00	6,915.00	6,162

**Cluster 3**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
NOMINA	1,923.03	599.60	1,101.00	3,305.00	16,770

Modelo 4. K-medias 3 segmentos sólo Pasivo.

**Cluster 1**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	1,509,566.08	437,376.86	1,030,337.42	2,453,738.71	15

**Cluster 2**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	11,440.29	20,383.22	560.31	166,659.48	24,405

**Cluster 3**

Variable	Media	Desviación Estándar	Mínimo	Máximo	Observaciones
PASIVO	315,996.78	152,049.09	167,266.09	883,627.82	331

En este último ejercicio obtenemos los cortes de aproximadamente 3,300/s y 6,900/s que comparado con los que se proponen que son de 2,500/s y 7,000/s son muy parecidos pero no iguales.

Para pasivo tenemos los cortes 165,000/s y 1,000,000 /s, contra los que propone el corporativo de 98,000/s y 504,000/s son casi un 100% más altos.

Sin embargo podríamos notar que en el segmento con pasivo mayor a 1,000,000/s se tienen sólo 15 individuos, el tomar el valor propuesto hará crecer ese segmento. Lo mismo ocurre al tomar el valor de 98.000/s para el otro segmento en lugar del que el modelo nos proporciona.

## **7. Factores críticos de éxito**

Dirección.

Siendo la segmentación estratégica uno de los objetivos corporativos se debe de comunicar directamente a todo el grupo de la decisión, las implicaciones que tiene y lo que se espera de cada una de las áreas involucradas. En este sentido se ha podido apreciar que, aunque en los niveles más altos la tarea fue efectiva, para una proporción grande de los empleados no se comunico de manera efectiva el objetivo y el plan de trabajo que se tenía para implementarlo.

Los resultados, sin embargo, han sido satisfactorios, aunque es posible que, cuando se baje a la operación de las oficinas, se tenga una resistencia mayor al cambio al no haberlo comunicado con anticipación.

Planeación.

Se realizó un plan en donde se estableció de manera general las fases que seguiría el proceso de segmentación, las áreas involucradas, insumos y todo lo necesario para llegar a ello.

Algo que fue determinante en el resultado y cumplimiento del plan fue el hecho de que la dirección estableció los tiempos de las actividades sin considerar las opiniones de la gente especialista en cada tema, sus estimaciones fueron en la mayoría de los casos insuficientes y como consecuencia se tuvieron retrasos.

Organización.

La ardua tarea de organización deberá considerar los siguientes aspectos para tener éxito:

Definición de segmentos de acuerdo a estudios socioeconómicos diferenciando, si cabe, áreas económicas que se adapten a la realidad de cada país sin olvidar la definición global de lo que cada segmento representa.

Identificar las fuentes de datos para la implementación del modelo de segmentación.



Contar con la tecnología necesaria para llevar a cabo la explotación de la información de 70 millones de clientes.

Tener el personal capacitado en el manejo de bases de datos para la correcta asignación de segmentos.

Transmitir a las distintas áreas de la necesidad y beneficios de implementar un mismo modelo para todo el grupo.

Implementar un calendario de fases para planteamiento, desarrollo, validación y análisis de resultados.

Dada la importancia del tema y la visión de la dirección de enfocarse en estas tres iniciativas deberemos realizar ajustes en las estructuras de los departamentos quitando prioridad a otras actividades sin descuidar la operación diaria y la atención a los clientes.

Control.

Las estrategias de control implementadas van desde un plan de trabajo con el desglose de actividades, fechas de entrega, responsables, motivos de retraso, insumos necesarios; pasando por planes de contingencia y juntas en distintos niveles de la organización, con diferentes niveles de detalle respecto al avance que se tiene en el proyecto.

Innovación Creatividad y Cambio

Debido a que las necesidades de los clientes están cambiando es necesario transformar las estructuras y la forma de operar de la organización, asimismo será necesario contar con las personas con el perfil acorde a esta nueva forma de trabajo. Se deberá capacitar y gestionar el cambio para tener el menor impacto en la productividad.

Forma de trabajo.

Para el logro de la implantación del proyecto se trabajó con un equipo multifuncional en el que cada departamento colaboró con los conocimientos y tareas de su especialidad, repartiendo el trabajo, las responsabilidades y decisiones, llevando a foros las opiniones de todos y tomando algunas decisiones en conjunto.

Motivación.

Se han alineado todos los sistemas de incentivos del grupo a las líneas estratégicas que se definieron por la presidencia y el comité de dirección siendo parte fundamental de los objetivos de cada gerencia y de cada puesto, directa o indirectamente.

## **8. Conclusiones**

Se analizaron distintos modelos de segmentación para estimar cortes en las variables de pasivo y nómina que separaran a los clientes en grupos heterogéneos entre segmentos y homogéneos dentro del segmento. Al combinar las dos variables en el segmento el modelo dio más peso a la variable de pasivo, aun después de estandarizar la variable, esto puede ser un efecto de que la cantidad de clientes con algún valor en esta última variable es mucho mayor a la cantidad de clientes con los valores de nómina informados.

Al analizar las variables por separado se observa que los valores, sobre todo de nómina, son muy parecidos a los propuestos corporativamente para la segmentación.

Se supone que estos valores se tomaron con estudios de personas en diferentes partes del mundo, que los valores de ambas variables estaban mejor informadas y que se consideró el tener segmentos muy pequeños para los cuales no valía la pena crear un segmento especial, de esta forma se ajustaron algunos cortes para hacer estos segmentos más grandes y los valores que están proponiendo parten entonces de un modelo estadístico con información más completa y con los ajustes necesarios.

Se concluye que los umbrales propuestos para la creación de los segmentos son correctos y se adaptan a la realidad de los clientes de Perú.

## **Glosario**

**Canal.** Medio por medio del cual un cliente realiza transacciones con el banco, puede ser por teléfono, por internet, a través del teléfono móvil, en la sucursal, en un cajero automático, etcétera.

**Canalidad.** Indica que canal o canales son los que un cliente utiliza para realizar sus transacciones con el banco.

**Multicanalidad.** Se refiere a la utilización de múltiples canales para realizar transacciones con el banco.

**Captación.** Proceso mediante el cual el sistema financiero recoge recursos del público ahorrador y los utiliza como fuente del Mercado Financiero. En el caso de la captación bancaria son todos los Recursos que la banca obtiene a través de sus instrumentos de captación (cuenta de cheques, cuenta de Ahorros, depósitos a Plazo fijo, etc.), que conforman los pasivos del sistema bancario e incluyen Recursos en Moneda nacional y extranjera.

**Vinculación.** Mientras mayor número de productos tiene un cliente se dice que está más vinculado, es decir, la relación entre el banco y el cliente se considera mayor mientras más productos tenga contratados.

**Fidelidad.** La fidelidad supone la existencia de una valoración positiva de la compañía y de un vínculo afectivo con la misma.

## Bibliografía

WEB BBVA, Información Corporativa.

<http://www.bbva.com/TLBB/tlbb/esp/informacion-corporativa/conozcanos/historia/el-comienzo/index.jsp> recuperado el 21 de enero de 2014.

WEB BBVA, El Grupo Internacional.

<http://www.bbva.com/TLBB/tlbb/esp/informacion-corporativa/conozcanos/historia/el-grupo-internacional/index.jsp> recuperado el 21 de enero de 2014.

Cognodata, Segmentación Estratégica.

<http://www.cognodata.com/news/banca/banca5/despertador.html> recuperado el 21 de enero de 2014

Metodología SAS para Proyectos de Minería.

<http://www.sas.com/offices/europe/uk/technologies/analytics/datamining/miner/semma.html> recuperado el 20 de marzo de 2014.

Robbins, S.P. (2009). Administración. México: Pearson Educación.

Candelas, E. (2012) Fundamentos de Administración. De la Facultad de Contaduría y Administración UNAM.

Hartigan, J. A.; Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society Series C (Applied Statistics)*.