



**Universidad ORT Uruguay**

**Diploma de Especialización en Analítica de  
Negocios**

**Obligatorio**

**Entregado como requisito para la aprobación del curso de Machine  
Learning no Supervisado.**

**Kary Francia - 269671**

**Guillermo Vázquez - 268211**

**Profesor: Damián Coltzau**

**28/11/2022**

## ÍNDICE

1.	Descripción del problema de negocio	3
2.	Objetivos	3
3.	Descripción de la base de datos	3
4.	Determinación del K-Óptimo	5
5.	Análisis con K-Means	6
6.	Análisis con K-Medoids	8
7.	Comparación de resultados	12
8.	Selección del modelo y descripción cualitativa de los clústeres	13
9.	Recomendaciones de potenciales estrategias	17

## **Obligatorio: Proyecto Corporativo**

### **1. Descripción del problema de negocio:**

El problema global planteado es que el banco XYZ acaba de desembarcar en el mercado local y planifica implementar una línea de créditos corporativos para diversas empresas. El banco necesita entender el potencial del mercado y detectar patrones que pudieran ser de utilidad para su negocio en cuanto hacia qué grupos debería invertir sus esfuerzos para promocionar su nueva línea de crédito minimizando los riesgos de tener una empresa morosa como cliente.

### **2. Objetivos:**

**Objetivo analítico:** Detectar diferentes grupos de empresas dentro de los elementos de la base de datos, utilizando algoritmos de agrupamiento para hallar una estructura en los datos de manera que las observaciones del mismo clúster sean más similares entre sí y que sean diferentes de las observaciones de los otros clústeres, a fin de formular recomendaciones de potenciales estrategias de negocio para el banco XYZ.

**Objetivo de negocio:** Identificar los grupos de clientes más rentables y solventes dentro de las 4000 empresas disponibles en el mercado local, para que el banco XYZ pueda definir una estrategia de otorgamientos de líneas de crédito corporativo según la performance financiera evaluada con 5 indicadores, encontrando los factores claves y evaluando las vulnerabilidades para minimizar el riesgo que una empresa no cumpla con los pagos del crédito a tiempo.

### **3. Descripción de la base de datos:**

La base de datos cuenta con información de 4000 empresas. Esta información se distribuye en 5 variables:

- ***Liquidez de corto plazo***

$WC.TA = \text{Capital de Trabajo} / \text{Total de Activos}$

- ***Rentabilidad históricas y actuales***

$RE.TA = \text{Ganancias Retenidas} / \text{Total de Activos}$

- ***EBITDA (capacidad de generar dinero)***

$EBIT.TA = \text{Ganancias antes de intereses e impuestos} / \text{Total de Activos}$

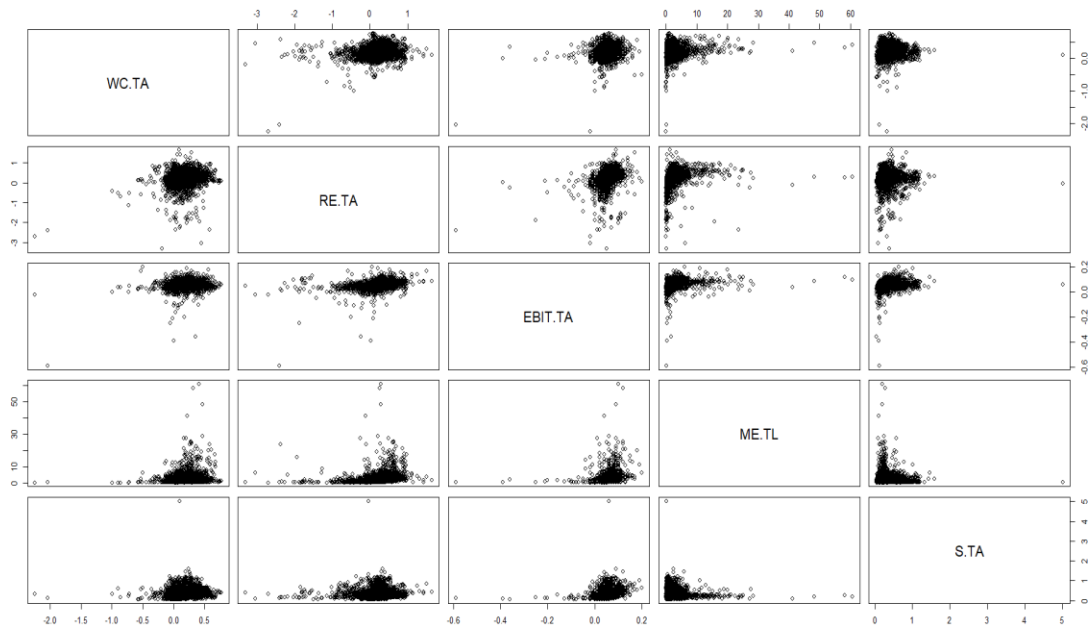
- ***Apalancamiento (capacidad de endeudamiento)***

$ME.TA = \text{Valor de Mercado del Equity} / \text{Total de pasivos}$

- ***Competitividad***

$S.TA = \text{Ventas} / \text{Total de activos} \# \text{cuanto mayor, más competitivo.}$

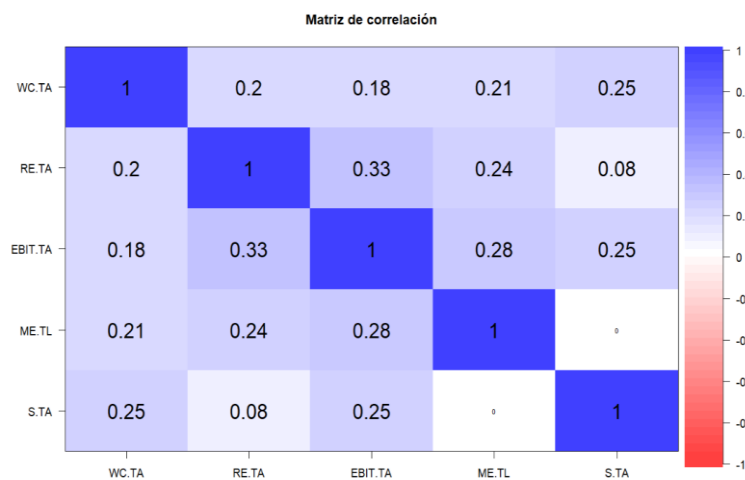
Al analizar la estructura de los datos, se observa rápidamente un gráfico de dispersión de toda la base para conocer más sobre la distribución de los mismos. Es notorio que el rango de la escala de la variable ME.TL es superior a las otras.



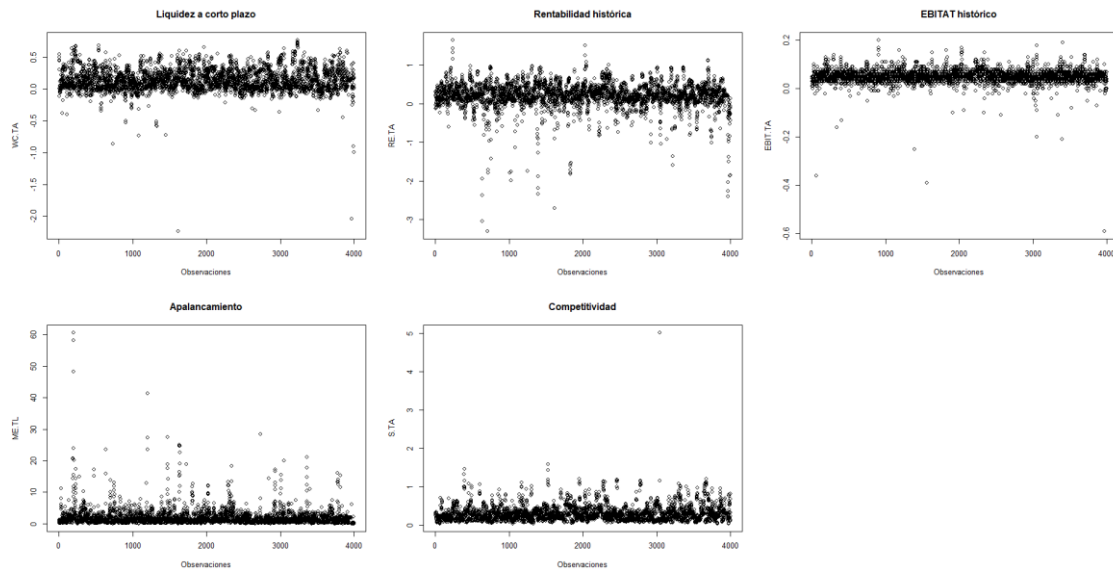
Luego, se observaron descriptivamente y cualitativamente las observaciones para detectar presencia de valores perdidos, existencia de outliers e inconsistencia entre el tipo de variable (cuantitativas o cualitativas). En este caso, no se detectaron valores perdidos, se detectó la presencia de outliers y también se corroboró que las variables se condigan con su naturaleza. Al ser indicadores y ratios financieros, las variables deberían ser numéricas. En este caso, todas las variables son numéricas por lo que no debemos de hacer ningún tratamiento en este sentido.

WC.TA			RE.TA			EBIT.T		ME.TL		S.TA		
Min.	-	2,24	Min.	-	3,31	Min.	-	0,59	Min.	0,02	Min.	0,04
1st Qu.		0,03	1st Qu.		0,09	1st Qu.		0,04	1st Qu.	0,62	1st Qu.	0,17
Median		0,12	Median		0,22	Median		0,05	Median	1,14	Median	0,26
Mean		0,14	Mean		0,21	Mean		0,05	Mean	1,95	Mean	0,30
3rd Qu.		0,24	3rd Qu.		0,37	3rd Qu.		0,07	3rd Qu.	2,24	3rd Qu.	0,37
Max.		0,77	Max.		1,64	Max.		0,20	Max.	60,61	Max.	5,01

Asimismo, las variables presentan muy baja correlación entre ellas, por lo que podríamos inferir que no hay multicolinealidad.



Por otro lado, realizamos gráficos de dispersión de los datos para cada variable, pudiendo reafirmar la presencia de outliers.

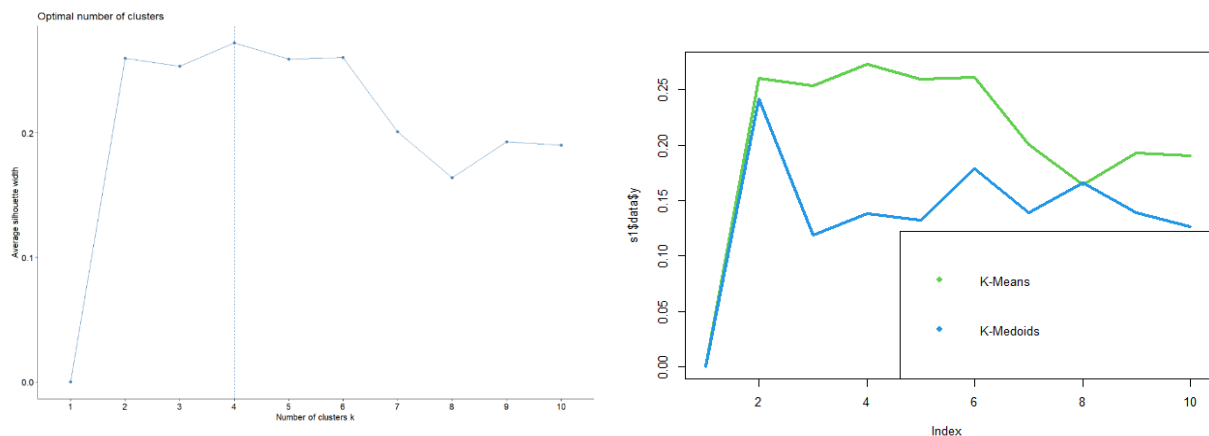


En cuanto al análisis de inspección de los datos con la desviación estándar, se detecta que la magnitud de la variable ME.TL difiere notablemente de las otras variables, por lo que se decide escalarlas antes de continuar con el análisis de clustering.

#### 4. Determinación del K-Óptimo:

Debido a que no se dispone de información adicional en la que orientarse, ni los conocimientos suficientes para intuir la cantidad de clústeres naturales; se estima el número de K óptimo mediante el método de silueta para el modelo de K-Means y para K-Medoids.

De acuerdo al siguiente gráfico comparativo, con K-Means a partir de 4 clústeres la reducción en la suma total de cuadrados internos parece estabilizarse, indicando que K = 4 es una buena opción. Por otro lado, con K-Medoids utilizando como medida de similitud la distancia de Manhattan, ya que es menos sensible a outliers que la euclídea, nos da que a partir de 2 clústeres la reducción en la suma total de diferencias internas parece estabilizarse, indicando que K = 2 sería una buena opción. Sin embargo, K-Means tiene una mayor silueta que K-Medoids, por lo que para una mejor descripción de los grupos se considerará para los siguientes análisis en todos los casos K=4.



## 5. Análisis con K-Means:

- *Base Pura:*

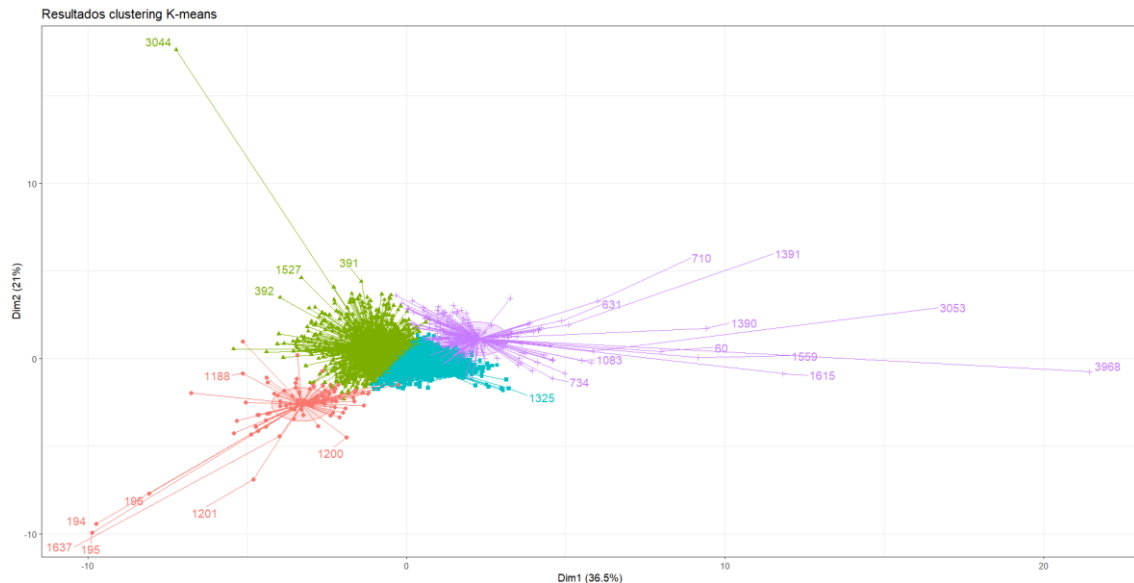
La distribución de los clústeres con la base incluida los outliers no es uniforme, siendo que:

Clúster 1 = 94 empresas

Clúster 2 = 1504 empresas

Clúster 3 = 2211 empresas

Clúster 4 = 191 empresas



Por otro lado, la proporción de la suma de cuadrados intra sobre el total es de  $W1 = 0.626$ , y los centros de cada clúster los siguientes:

	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA
1	0.9202885	0.8735129	0.9112742	4.7875766	-0.1980544
2	0.6613234	0.4642400	0.4316275	0.1058538	0.6897920
3	-0.4484124	-0.1115787	-0.2248544	-0.2456012	-0.4829385
4	-0.4696212	-2.7938567	-1.2443686	-0.3466609	0.2562676

- *Base Capeada:*

Se comprende que, una de las desventajas del algoritmo K-Means es que no es robusta a los outliers, por lo que los resultados anteriores no serían los más idóneos para describir a los clústeres. Por tal motivo, se decide realizar primero el análisis capeando a los outliers en el cuantil 99 antes de depurarlos, considerando que se obtiene una mejor información por la cantidad significativa que representa el hecho de retirar demasiados outliers.

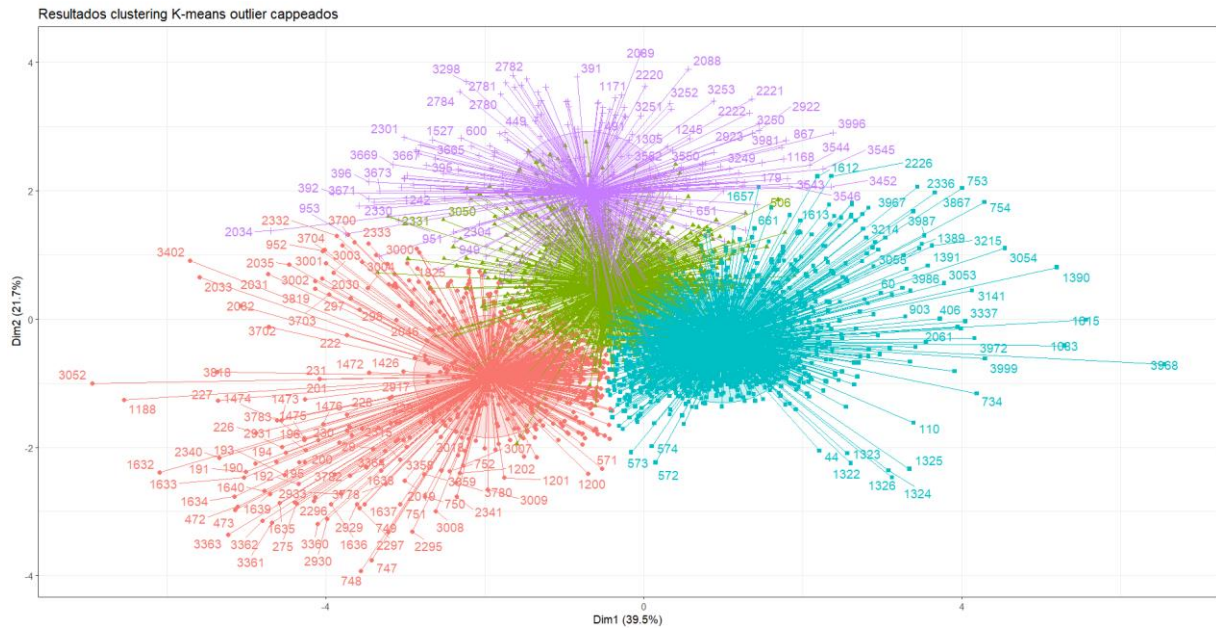
La distribución de los clústeres con la base capeada de los outliers es mucho más uniforme que el anterior, siendo que:

Clúster 1 = 625 empresas

Clúster 2 = 1065 empresas

Clúster 3 = 1967 empresas

Clúster 4 = 343 empresas



Por otro lado, la proporción de la suma de cuadrados intra sobre el total es de  $W2 = 0.573$ , minimizando aún más el  $w$ . Y los centros de cada clúster los siguientes:

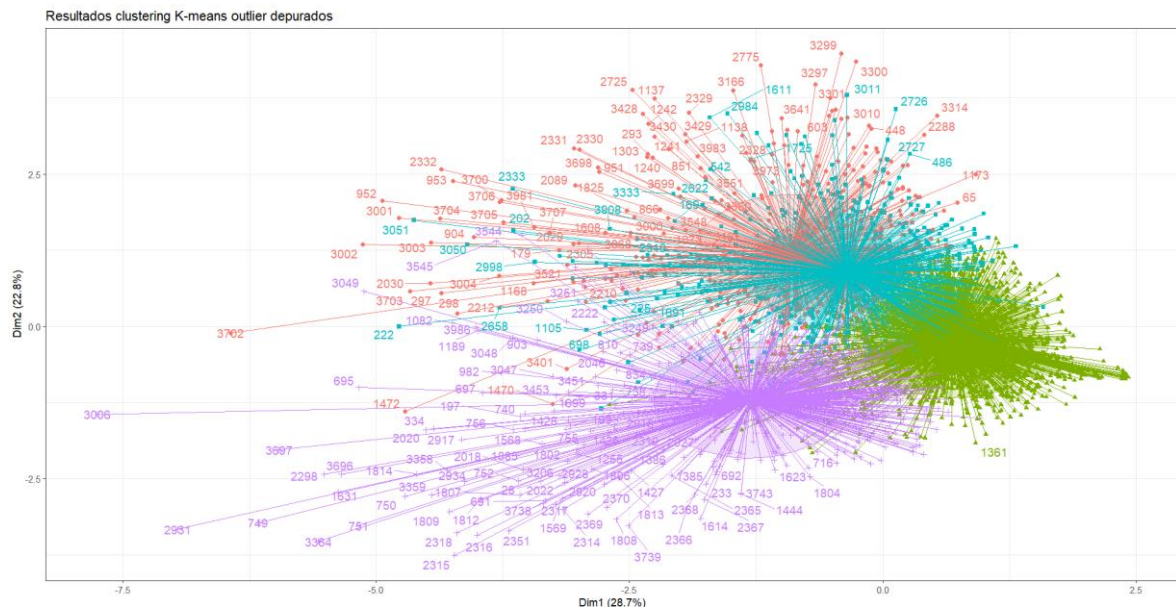
	WC.TA_cap	RE.TA_cap	EBIT.TA_cap	ME.TL_cap	S.TA_cap
1	0.1413283	0.9668462	0.953444126	1.0799713	0.1283034
2	0.9678164	0.1713878	-0.004374202	-0.1092570	0.1306476
3	-0.5779255	-0.3091430	-0.314438204	-0.2968631	-0.5067657
4	0.1247713	-0.2501313	0.290891835	-0.2507032	2.1537296

- *Base sin outliers:*

Para finalizar con el análisis del algoritmo K-Means, se decidió observar cómo se agrupaban los clústeres retirando los outliers de la base. Se tuvo como criterio retirar aquellas observaciones con más o menos 3 desviaciones estándar (Z-score), quedando una base de 3740 observaciones para analizar. Debido a que, si se iba por el criterio de boxplot depurando a las observaciones fuera de los bigotes, se obtenía una base con 3219 observaciones, retirando un 19.5% de la base, lo que no era permisible para el análisis.

La distribución de los clústeres con la base depurada de los outliers es también uniforme, siendo que:

Clúster 1 = 504 empresas  
 Clúster 2 = 2057 empresas  
 Clúster 3 = 606 empresas  
 Clúster 4 = 573 empresas



Los centros de cada clúster los siguientes:

	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA
1	0.7780751	0.4326995	1.0225272	0.4412353	1.3556444
2	0.4766353	0.3460717	0.3809255	0.3426526	0.4774028
3	1.5681668	0.4862994	0.5133959	0.3952456	0.5589250
4	0.5938058	1.4147049	0.7805405	0.6306374	0.4343795

Por otro lado, la proporción de la suma de cuadrados intra sobre el total es de  $W3 = 0.599$ , maximizando el W en comparación con la base capeada.

En la siguiente comparativa, podemos concluir que el análisis de K-Means con la base capeada es la mejor solución para este algoritmo porque minimiza la proporción de suma de cuadrado intra sobre el total.

	W1	W2	W3
K-Means	0.626	0.573	0.599

## 6. Análisis con K-Medoids:

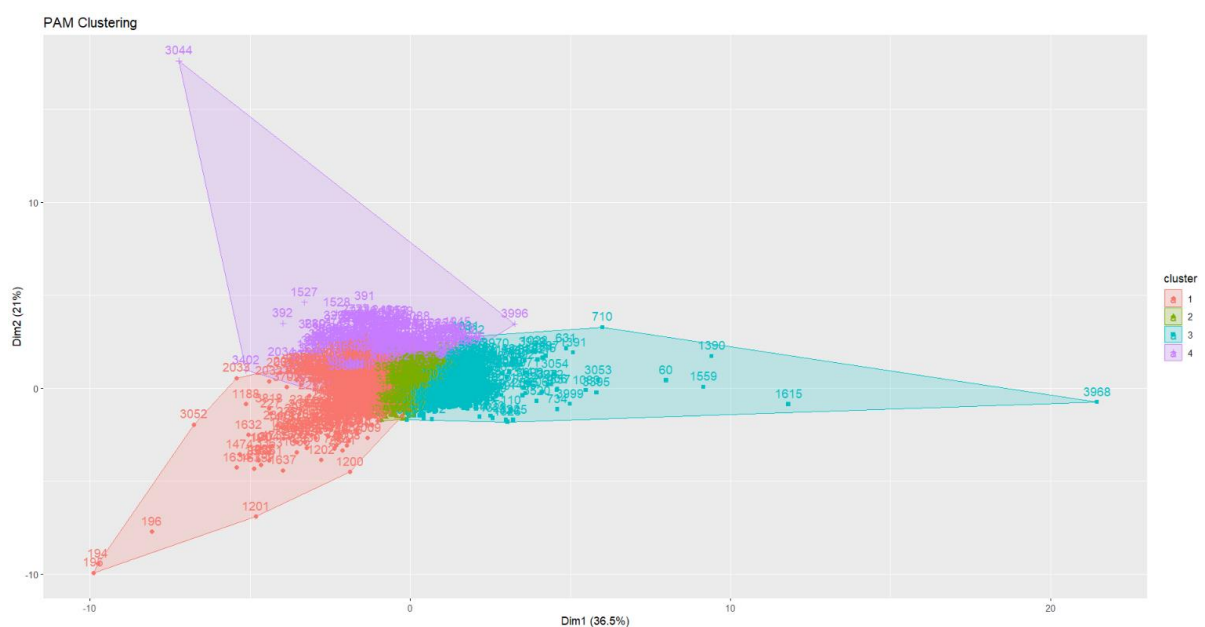
### ● Base Pura:

Luego de realizar la evaluación con el algoritmo PAM, se obtiene la siguiente información:

- Con el análisis de silueta se observa que la distribución de los datos se concentra más en el clúster 2 y clúster 3. El clúster 1 tiene promedio negativo por influencia de los outliers, y el promedio total de la silueta es de 0.205.

	cluster	size	ave.sil.width
1	1	871	-0.05
2	2	1387	0.26
3	3	1351	0.14
4	4	391	0.10





- Los centroides representativos de cada clúster son las siguientes observaciones:

Clúster 1: empresa 2716

Clúster 2: empresa 2138

Clúster 3: empresa 3439

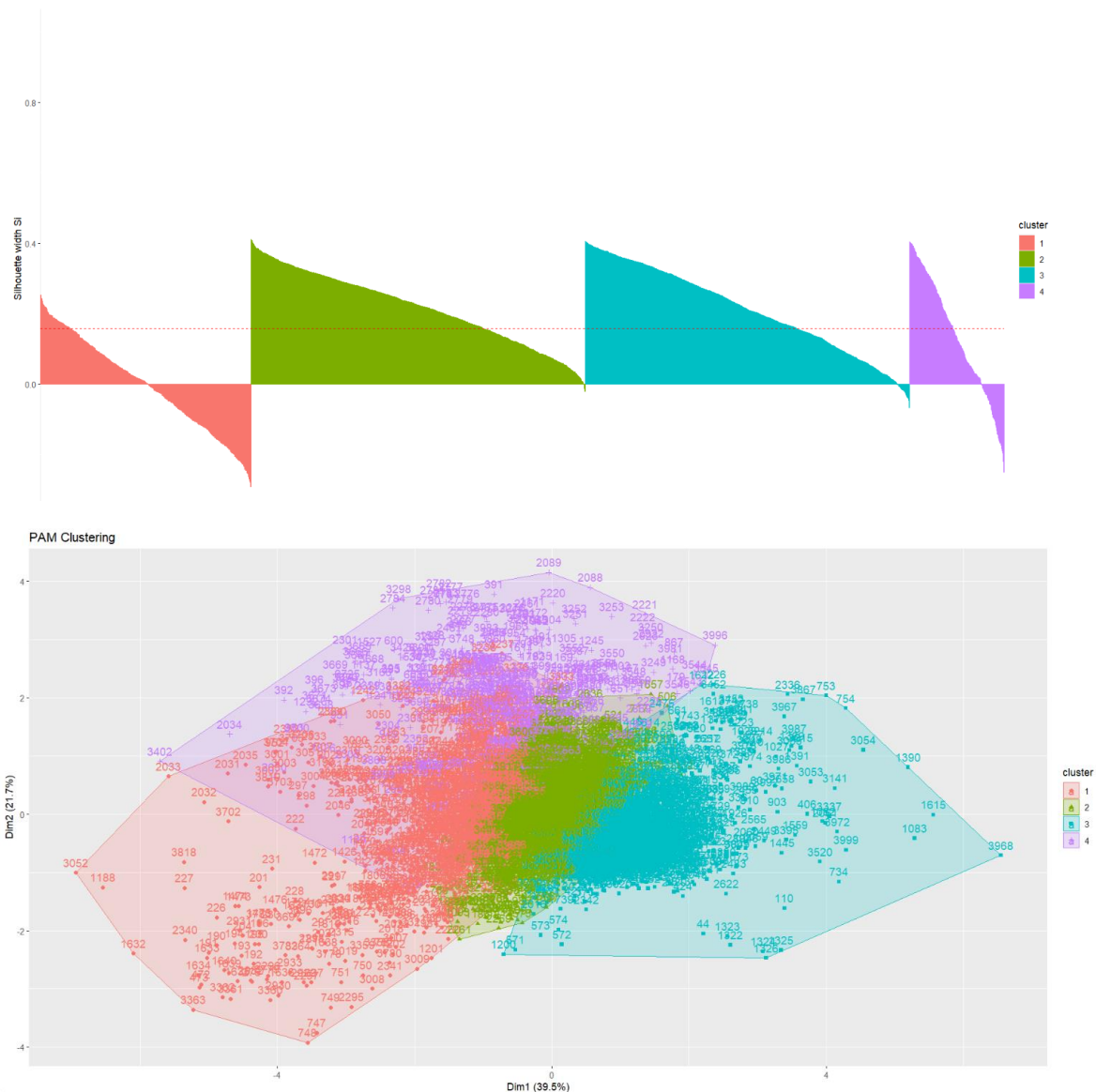
Clúster 4: empresa 65

	ID	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA
[1,]	2716	0.62926486	-0.03119712	-0.06116135	-0.1716265	-0.1148435
[2,]	2138	-0.60088644	-0.24169897	-0.39846370	-0.2818295	-0.6492537
[3,]	3439	-0.01510011	0.81081025	0.61344336	0.3660302	0.1766529
[4,]	65	0.27779306	0.02894626	0.27614100	-0.2217188	1.9742143

#### ● *Base Capeada:*

Se aplicó el algoritmo PAM también para la base capeada, obteniendo la siguiente información:

- Con el análisis de silueta se observa que la distribución de los datos es más uniforme. El clúster 1 mantiene el promedio negativo, pero en menor magnitud que en el caso anterior. Siendo el promedio total de la silueta de 0.222.



- Los centroides representativos de cada clúster son las siguientes observaciones:

Clúster 1: empresa 2716

Clúster 2: empresa 2138

Clúster 3: empresa 3439

Clúster 4: empresa 65

En comparación con los anteriores centroides, se puede ver que el clúster 4 mantiene el mismo. Sin embargo, los clústeres 1, 2 y 3 cambiaron de centroides, pero al comparar sus características se puede notar que algunos elementos se intercambiaron pasando del clúster 1 de la base completa al clúster 3 de la base capeada, el clúster 2 de la base completa al clúster 1 de la base capeada y el clúster 3 de la base completa al clúster 2 de la base capeada.

	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA	id	
2004	0.31	0.45	0.07	3.18	0.34	1	3
3439	0.14	0.48	0.07	3.05	0.34	3	

	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA	id	
2891	0.15	0.28	0.05	1.24	0.28	2	1
2716	0.25	0.20	0.05	1.44	0.28	1	

	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA	id	
267	0.03	0.11	0.04	0.85	0.14	3	2
2138	0.04	0.13	0.04	1.11	0.17	2	

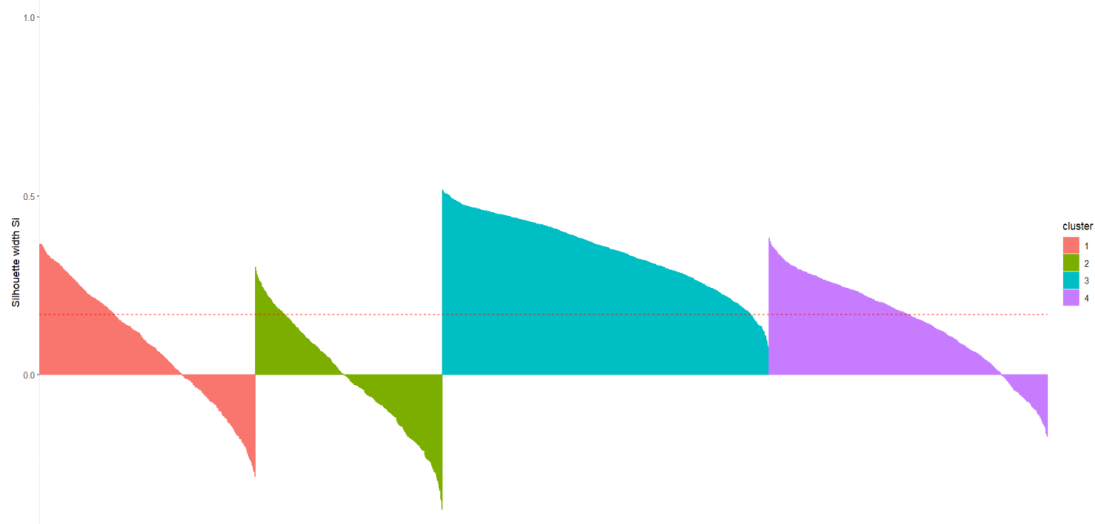
	WC.TA	RE.TA	EBIT.TA	ME.TL	S.TA	id	
65	0.190	0.220	0.060	1.290	0.710	4	
65	0.19	0.22	0.06	1.29	0.71	4	

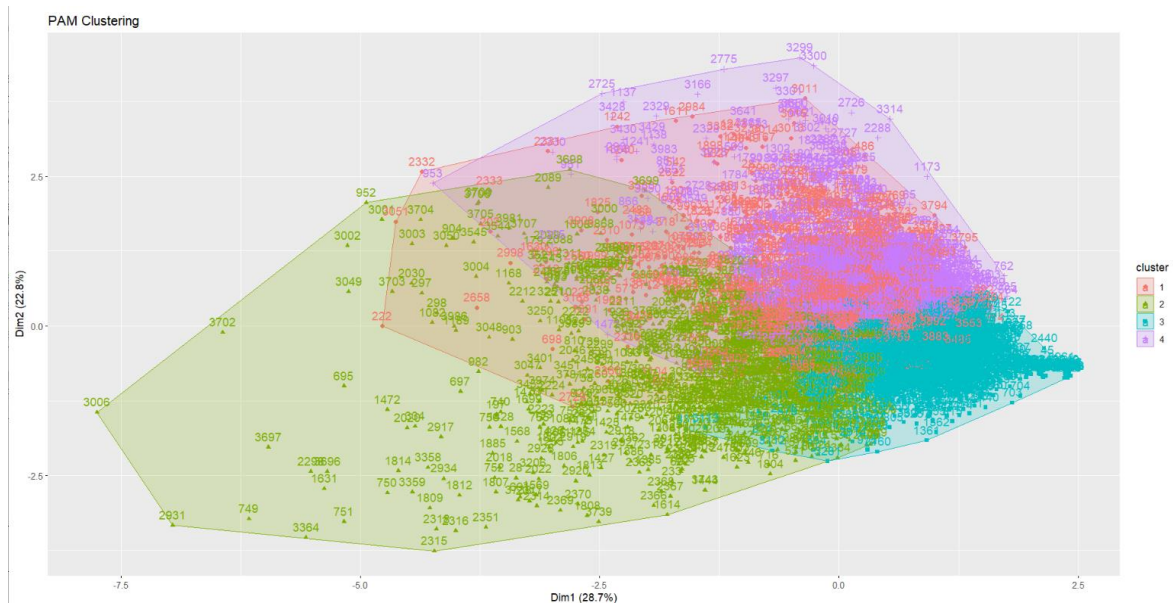
- *Base sin outliers:*

Se aplicó el algoritmo PAM también para la base depurada, obteniendo la siguiente información:

- Con el análisis de silueta se observa que la distribución de los datos no es más uniforme que el de la base capeada. En este caso, el clúster 2 tiene el promedio negativo. Siendo el promedio total de la silueta de 0.168, el menor de todos los promedios antes evaluados.

cluster	size	ave.sil.width
1	800	0.08
2	693	-0.01
3	1214	0.35
4	1033	0.15





## 7. Comparación de resultados:

Al comparar los resultados de K-Medoids y K-Means en la matriz de contingencia, se observa que presentan soluciones totalmente diferentes, tanto en la base con outliers, con outliers capeados y con outliers depurados. Particionando una con otra a los clústeres en todos los casos.

### Matriz de contingencia de la base con outliers

	kmedoids				
kmeans	1	2	3	4	Sum
1	6	0	87	1	94
2	560	0	602	342	1504
3	477	1477	242	15	2211
4	35	121	0	35	191
Sum	1078	1598	931	393	4000

### Matriz de contingencia de la base con outliers capeados

	kmedoids				
kmeans	1	2	3	4	Sum
1	8	2	604	11	625
2	821	0	187	57	1065
3	243	1592	128	4	1967
4	8	3	10	322	343
Sum	1080	1597	929	394	4000

### Matriz de contingencia de la base con outliers depurados

	kmedoids				
kmeans	1	2	3	4	Sum
1	30	114	3	357	504
2	182	48	1186	641	2057
3	564	12	0	30	606
4	24	519	25	5	573
Sum	800	693	1214	1033	3740

Por otro lado, se visualiza que las observaciones quedan mejor distribuidas con el algoritmo K-Medoids que con K-Means, tanto en la base con outliers como en los capeados por lo que se decidió compararlas también en la matriz de contingencia, observando que son cualitativamente parecidas.

kmedoidscap					
kmedoids	1	2	3	4	Sum
1	1077	0	0	1	1078
2	3	1595	0	0	1598
3	0	2	929	0	931
4	0	0	0	393	393
Sum	1080	1597	929	394	4000

Asimismo, se decidió correr K-Medoids con k=2, debido a que era el K óptimo recomendado por el método silueta. Se observa que el centroide se mantiene para el clúster 2 (observación 2138), por lo que entendemos que al utilizar K=4 lo que estamos logrando es determinar grupos más específicos de las empresas que serían los clientes potenciales del banco XYZ.

kmedoidscap4					
kmedoidscap2	1	2	3	4	Sum
1	867	2	817	376	2062
2	213	1595	112	18	1938
Sum	1080	1597	929	394	4000

## 8. Selección del modelo y descripción cualitativa de los clústeres:

En primer lugar, en cuanto a la descripción cualitativa de los clústeres, para poder decir que las soluciones son cualitativamente iguales, se debe cumplir que las filas y las columnas tienen que tener un elemento que no sea 0. Sin embargo, en este caso, podemos decir que todas las soluciones son cualitativamente diferentes, ya que además de no cumplirse con la condición mencionada, la distribución de observaciones de un algoritmo (PAM) con otro el otro (K-Means) difiere totalmente.

Comparando los resultados y analizando los indicadores de silueta y W para los distintos modelos y las diferentes casuísticas presentadas (base pura, base capeada y base sin outliers), se considera que utilizar **la base capeada parece ser la mejor solución** ya que es donde se minimiza aún más el W y tiene una mayor silueta, tal como se mencionó a lo largo del análisis:

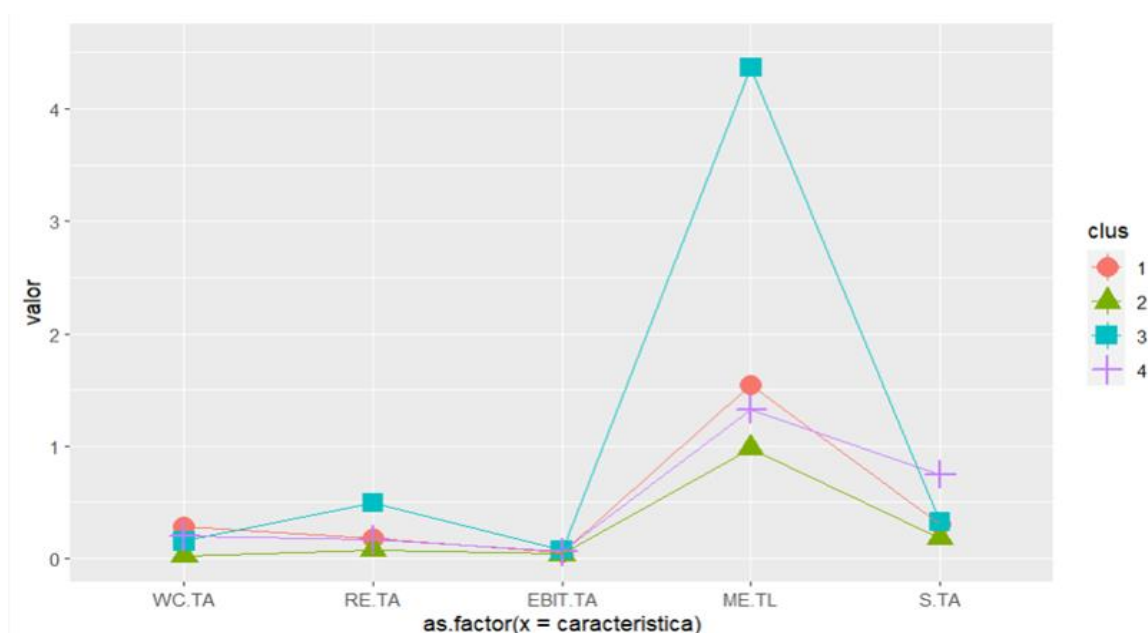
	Pura	Capeada	Sin outliers
	Silueta 1	Silueta 2	Silueta 3
K-Medoids	0,205	0,222	0,168

	Pura	Capeada	Sin outliers
	W1	W2	W3
K-Means	0,626	0,573	0,599

En adición a esto, si observamos y comparamos la distribución de las observaciones en cada clúster, vemos que para el caso de la base capeada se logra una mayor distribución de las observaciones, con lo cual se pueden extraer mejores conclusiones.

Si comparamos funcionalmente los algoritmos K-means y K-Medoids, este último asegura mayor estabilidad. Los centros de PAM sufren menor desplazamiento debido a los outliers capeados, no así con k-means. Esto significa que los centros son más representativos de la naturaleza del clúster. ***Es por esto, que utilizaremos el algoritmo PAM con la base capeada para extraer conclusiones cualitativas para realizar la mejor estrategia de negocio.***

A continuación, se logra agrupar a las empresas respecto a sus variables financieras, pudiendo identificarse 4 grupos de empresas, cada una con características diferentes.



- **Liquidez de corto plazo (WC.TA):**

Clúster	Tamaño	Centroide (id)	WC.TA	RE.TA	EBIT.TA	ME.TA	S.TA
1	1080	2716	0.25	0.20	0.05	1.44	0.28
2	1597	2138	0.04	0.13	0.04	1.11	0.17
3	929	3439	0.14	0.48	0.07	3.05	0.34
4	324	65	0.19	0.22	0.06	1.29	0.71

Clúster	Característica
1	Se caracterizan por evidenciar muy buena capacidad crediticia, ya que tienen la capacidad de pagar sus operaciones normales cubriendo además sus deudas a corto plazo.
2	Se caracterizan por evidenciar muy baja capacidad para asumir sus deudas financieras a corto plazo y para afrontar circunstancias desfavorables sobrevenidas.
3	Se caracterizan por evidenciar poca capacidad para asumir sus deudas financieras a corto plazo, pero podrían afrontar circunstancias desfavorables liquidando inmovilizados para atender dichas deudas de corto plazo.
4	Se caracterizan por ser empresas con activos circulantes superiores a sus pasivos circulantes, es decir que cuenta con dinero para realizar sus operaciones normales después de haber cubierto sus obligaciones a corto plazo.

- **Rentabilidad históricas y actuales (RE.TA):**

Clúster	Tamaño	Centroide (id)	WC.TA	RE.TA	EBIT.TA	ME.TA	S.TA
1	1080	2716	0.25	0.20	0.05	1.44	0.28
2	1597	2138	0.04	0.13	0.04	1.11	0.17
3	929	3439	0.14	0.48	0.07	3.05	0.34
4	324	65	0.19	0.22	0.06	1.29	0.71

Clúster	Característica
1	Se caracterizan por ser empresas que si bien tienen bajas ganancias retenidas, utilizan sus ganancias para crecer pero también dependen de financiamiento de deuda y capital.
2	Se caracterizan por ser empresas que no dependen de las ganancias retenidas para expandirse, sino que dependen de otros tipos de financiamiento como de deuda y de capital, reembolsando a los tenedores de la deuda con capital e intereses a lo largo del tiempo o emitiendo nuevas acciones y vendiendo fracciones de la empresa para recaudar fondos.
3	Se caracterizan por ser empresas con ganancias retenidas sólidas, lo que podría significar que estas empresas utilizarán sus ganancias en un futuro para expandirse o reinvertir en el negocio.
4	Se caracterizan por ser empresas con ganancias retenidas, la cuales utilizan para crecer a través de la reinversión.

- **EBITDA (capacidad de generar dinero):**

Debido a que la variable EBIT.TA usa los valores contables de los activos totales del balance general, podría ser que se subestime el valor de mercado real de los activos fijos de las empresas, por ende, el resultado de la ratio podría mostrar un rendimiento de relación más alto de lo que realmente debería ser.

Clúster	Tamaño	Centroide (id)	WC.TA	RE.TA	EBIT.TA	ME.TA	S.TA
1	1080	2716	0.25	0.20	0.05	1.44	0.28
2	1597	2138	0.04	0.13	0.04	1.11	0.17
3	929	3439	0.14	0.48	0.07	3.05	0.34
4	324	65	0.19	0.22	0.06	1.29	0.71

Teniendo en cuenta lo antes mencionado, se realiza la siguiente descripción:

Clúster	Característica
1	Se caracterizan por ser empresas que generan pocas ganancias empleando sus activos.
2	Se caracterizan por ser empresas con muy baja eficiencia para generar beneficios utilizando sus activos en comparación a sus ganancias.
3	Se caracterizan por ser empresas muy eficientes utilizando sus activos para generar ganancias. Sin embargo, puede que algunos generen deudas para comprar un activo, por lo que el indicador podría parecer favorable, aunque dichas empresas puedan tener problemas para realizar sus pagos de gastos por intereses.
4	Se caracterizan por ser empresas eficientes utilizando sus activos para generar ganancias.



- **Apalancamiento (capacidad de endeudamiento)**

Clúster	Tamaño	Centroide (id)	WC.TA	RE.TA	EBIT.TA	ME.TA	S.TA
1	1080	2716	0.25	0.20	0.05	1.44	0.28
2	1597	2138	0.04	0.13	0.04	1.11	0.17
3	929	3439	0.14	0.48	0.07	3.05	0.34
4	324	65	0.19	0.22	0.06	1.29	0.71

Clúster	Característica
1	Se caracterizan por ser empresas cuyos mecanismos principales para apalancarse es a través de la deuda para realizar inversiones con ellas, aumentando poco a poco el valor de su patrimonio en el mercado.
2	Se caracterizan por ser empresas con excesiva deuda (corto y largo plazo), por lo que el valor de su patrimonio ha sido absorbido por las pérdidas acumuladas, viéndose reflejado a posteriori en el valor del mercado.
3	Se caracterizan por ser empresas saludables, con deudas de mejor calidad en lo que al valor de su patrimonio se refiere, generando confianza en el mercado, por lo que sus acciones representan un alto valor.
4	Se caracterizan por ser empresas que para apalancarse obtienen excesiva deuda a corto plazo, por lo que su patrimonio presenta un bajo valor en el mercado.

- **Competitividad (S. TA):**

Habitualmente esta variable se utiliza para comparar empresas que pertenecen a un mismo sector o son similares entre sí, y en base a ello poder inferir cuál es la empresa o grupo de empresas más eficientes. En este caso de estudio, se desconoce si las 4000 empresas son similares o de distintos rubros, por lo que la ratio puede variar significativamente de un sector a otro. Por ello, no aportaría información útil comparar la rotación de activos de empresas que se encuentran en sectores diferentes.

Por lo general, cada sector tiene su propio valor de referencia S. TA, por ejemplo, el sector de supermercados o retail puede llegar a tener un ratio estimado de 10, mientras que la ratio del sector industrial está alrededor de 1.

Clúster	Tamaño	Centroide (id)	WC.TA	RE.TA	EBIT.TA	ME.TA	S.TA
1	1080	2716	0.25	0.20	0.05	1.44	0.28
2	1597	2138	0.04	0.13	0.04	1.11	0.17
3	929	3439	0.14	0.48	0.07	3.05	0.34
4	324	65	0.19	0.22	0.06	1.29	0.71

Entonces, para poder describir cualitativamente a cada clúster se asume que todas las empresas son parecidas o pertenecen a sectores similares.



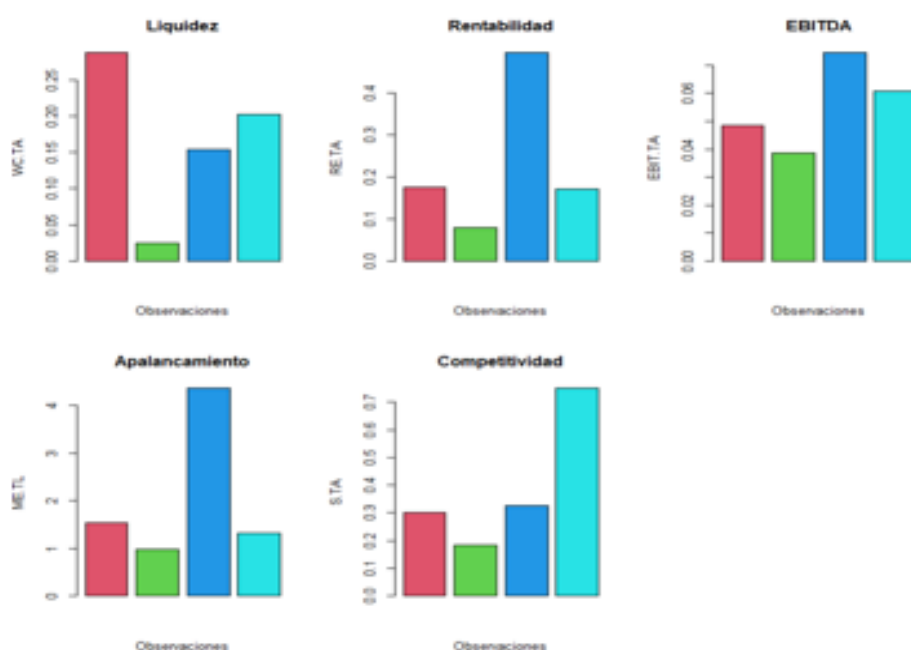
Clúster	Característica
1	Se caracterizan por ser empresas que evidencian una capacidad ociosa planificada o que su capacidad ociosa está latente por no utilizar adecuadamente sus activos.
2	Se caracterizan por ser empresas que tienen demasiado capital inmovilizado en su base de activos.
3	Se caracterizan por ser empresas que tienen pocos activos para enfrentar un aumento de ventas potenciales o que su base de activos está obsoleta y buscan realizar la inversión correspondiente para actualizarlos.
4	Se caracterizan por ser empresas más eficiente al momento de generar ingresos o ventas a partir de sus activos, están en constante actualización de los mismos.

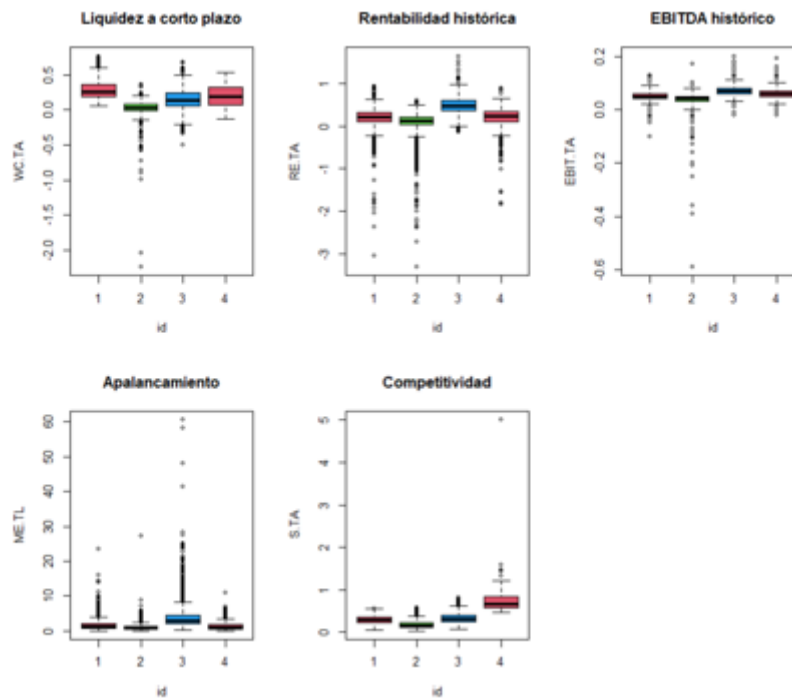
## 9. Recomendaciones de potenciales estrategias:

En primer lugar, se analizó la información cualitativa de cada clúster para entender la relación entre las variables y los clústeres definiendo patrones de comportamiento, a fin de recomendar una estrategia adecuada:



### INFORMACIÓN CUALITATIVA DE CADA CLUSTER





En función de los gráficos de barplot y boxplot, se extrajeron algunas conclusiones sobre cada clúster y sobre cada variable, clasificando cada combinación según los siguientes conceptos:

Clasificación			
Bajo	Medio	Alto	Muy Alto

CLASIFICACIÓN CLUSTER SEGÚN GRÁFICOS (BOXPLOT Y BARPLOT)					
CLUSTER	Liquidez	Rentabilidad	EBITDA	Apalancamiento	Competitividad
1	Muy Alto	Medio	Medio	Medio	Medio
2	Bajo	Bajo	Bajo	Bajo	Bajo
3	Medio	Muy Alto	Muy Alto	Muy Alto	Medio
4	Alto	Medio	Alto	Medio	Muy Alto

En este caso, todas las variables son importantes para la estrategia de negocio. Cuanto más altos los indicadores en cada una de las mismas, aporta positivamente a la clasificación de cada clúster.

Más allá de esto, nos pareció pertinente realizar una ponderación por cada variable según la importancia la hora de definir una estrategia en base al otorgamiento de líneas de crédito corporativa.

ASIGNACIÓN PONDERACIÓN Y PUNTAJE POR VARIABLE Y CLASIFICACIÓN					
CLASIFICACIÓN	22%	19%	19%	30%	10%
	Liquidez	Rentabilidad	EBITDA	Apalancamiento	Competitividad
Bajo	0,50	0,50	0,50	0,50	0,50
Medio	2,00	2,00	2,00	2,00	2,00
Alto	3,00	3,00	3,00	3,00	3,00
Muy Alto	5,00	5,00	5,00	5,00	5,00

En línea con lo anterior, cuanto más alta la clasificación en todas las variables, mejor será el puntaje final según la ponderación asignada. Realizamos las asignaciones de los puntajes correspondientes según las tablas anteriores y en función de cada ponderación:

SCORE FINAL						
	Liquidez	Rentabilidad	EBITDA	Apalancamiento	Competitividad	
CLUSTER	PUNTAJE					Total SCORE
1	1,10	0,38	0,38	0,60	0,20	2,66
2	0,11	0,10	0,10	0,15	0,05	0,50
3	0,44	0,95	0,95	1,50	0,20	4,04
4	0,66	0,38	0,57	0,60	0,50	2,71

Cuanto más alto el score, menos riesgo tiene asociado el clúster en función de todas las variables y las características de los mismos teniendo en cuenta el objetivo de negocio. En este caso, el clúster 3, es la mejor calificación (riesgo bajo). Para el primer puntaje asignado al clúster 3 en la variable liquidez, se tuvo en cuenta su clasificación (Medio - puntuación 2), la ponderación según la variable y la importante en el objetivo de negocio (22%). Si multiplicamos nos da una puntuación de 0,44. Este procedimiento se repitió para cada variable según las ponderaciones y calificaciones asignadas anteriormente, llegando al score final.

#### **La estrategia de otorgamiento de líneas de crédito la dividimos en 4:**

- 1 - Decisión de otorgar una línea de crédito (Si/No) - **En función del riesgo asociado determinado por el score.**
- 2 - Monto de la línea de crédito (Bajo/Medio/Alto) - **En función de la ratio de apalancamiento.**
- 3 - Plazo mínimo otorgado (Corto plazo / Mediano plazo / Largo plazo) - **En función de la liquidez.**
- 4 - Precio (Tasa efectiva anual) - Risk Based Price - **En función del riesgo asociado determinado por el score.**

ESTRATEGIA						
CLUSTER	Línea de crédito	Monto	Plazo	Score	Clasificación Riesgo (Score)	Precio (TEA) - RBP
1	Si	Medio	Corto Plazo	2,66	Medio	Media*
2	No	Bajo	Largo Plazo	0,50	Alto	Alta**
3	Si	Alto	Mediano plazo	4,04	Bajo	Baja***
4	Si	Medio	Mediano plazo	2,71	Medio	Media*

\* Tasa media del mercado en función del plazo y monto otorgado

\*\* Tasa promedio alta del mercado en función del plazo y monto otorgado

\*\*\* Tasa promedio baja del mercado en función del plazo y monto otorgado

En cuanto a la asignación de tasas según el tipo de riesgo asociado a cada clúster, asumimos que las empresas con bajo riesgo y haciendo referencia al método de asignación de precio (RBP), se puede reducir el precio ya que el riesgo asociado que presentan es bajo en relación al resto del mercado. Cuanto mayor riesgo, se asigna una tasa mayor ya que hay más probabilidad de incumplimiento y el negocio debe ser lo más rentable posible.

Para el caso de las empresas categorizadas en el clúster 2, la recomendación que le hacemos al banco XYZ es no otorgarle una línea de crédito ya que la probabilidad de incumplimiento es alta debido al score asociado. Más allá de esto, en caso de que, por una decisión estratégica/política, la decisión es afirmativa (se le quiere conceder una línea de crédito), le sugerimos hacerlo con las recomendaciones expuestas tanto en los montos de línea de crédito (bajo), financiado en plazos largos y con una TEA alta para maximizar lo mayor posible los beneficios.