

DOI: <https://doi.org/10.56124/refcale.v13i2.004>

Análisis De Clúster Para La Segmentación De Clientes En La Concesión De Créditos Otorgados Por La Corporación Financiera Nacional

Análisis De Clúster Para Créditos CFN

Autores:

Shirley Patricia Hidalgo Bailón¹

Luis Cedeño-Valarezo²

Dirección Para Correspondencia: pattyhidalgob24.05@gmail.com

Fecha de recepción: julio 14, 2025

Fecha de aceptación: julio 14, 2025

Resumen

La segmentación de clientes en instituciones financieras es clave para optimizar la personalización de productos y mejorar la inclusión financiera. Diversos estudios han aplicado métodos como análisis de clústeres, redes neuronales y aprendizaje automático para evaluar el comportamiento de los clientes y mejorar la gestión crediticia. En el Ecuador, existen factores entre ellos la ubicación geográfica y las particularidades socioeconómicas inciden en la distribución de los créditos. Este estudio con el objetivo de aplicar el análisis de clúster jerárquico para segmentar las provincias ecuatorianas según los créditos otorgados por la Corporación Financiera Nacional (CFN) en 2023. Se procesaron datos de 20 provincias, considerando variables como monto otorgado, tipo de crédito y factores socioeconómicos. Se aplicó la distancia de

¹ Licenciada en Servicios Generales. pattyhidalgob24.05@gmail.com

² Ingeniero en Informática. Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López. lcedeno@espam.edu.ec. <https://orcid.org/0000-0002-2403-9781>. Calceta, Ecuador.

Mahalanobis para mitigar valores atípicos y mejorar la calidad del agrupamiento. Los resultados identificaron tres clústeres principales, donde la mayoría de las provincias se agrupan en uno, mientras que las más representativas forman clústeres independientes. Sin embargo, métricas como el índice de Dunn y el coeficiente de silueta sugieren que el modelo no es óptimo. Esto resalta la necesidad de explorar enfoques alternativos para mejorar la precisión de la segmentación y optimizar la asignación de recursos financieros en el país.

Palabras Claves: Análisis Clúster Jerárquico; Corporación Financiera Nacional; Tipos de Créditos

Cluster Analysis For Customer Segmentation In The Granting Of Loans By The National Financial Corporation

Abstract

Customer segmentation in financial institutions is key to optimizing product customization and improving financial inclusion. Several studies have applied methods such as cluster analysis, neural networks, and machine learning to evaluate client behavior and improve credit management. In Ecuador, factors such as geographic location and socioeconomic characteristics affect the distribution of credit. This study uses hierarchical cluster analysis to segment Ecuadorian provinces according to loans granted by Corporación Financiera Nacional (CFN) in 2023. Data from 20 provinces were processed, considering variables such as amount granted, type of credit, and socioeconomic factors. Mahalanobis distance was applied to mitigate outliers and improve the quality of the clustering. The results identified three main clusters, where most of the provinces are grouped into one, while the most representative ones form independent clusters. However, metrics such as Dunn's index and the silhouette coefficient suggest that the model is not optimal. This highlights the need to explore alternative approaches to improve segmentation accuracy and optimize the allocation of financial resources in the country.

Keywords: Hierarchical Cluster Analysis, National Finance Corporation, Types of Loans.

Introducción:

La segmentación de clientes a partir de su naturaleza se rige como una herramienta importante para optimizar la personalización de productos financieros, en particular si se trata de instituciones de crédito. En ese mismo orden, la inclusión financiera y la eficiencia en la asignación de recursos son una prioridad, por ejemplo, comprender las características de los clientes considerando su naturaleza resulta fundamental en el diseño de estrategias efectivas que procuren el acceso equitativo al crédito.

Por otro lado, las últimas investigaciones en relación con la segmentación de los clientes en las diferentes instituciones financieras, con énfasis en la concesión de créditos, su aporte científico ha sido significativo, en ese sentido, se ha aprovechado el aprendizaje automático, y consecuentemente su análisis de datos y metodologías asociadas al análisis estadístico, con el objetivo de mejorar la comprensión del comportamiento de los clientes y la evaluación de riesgos.

En el mismo orden de la idea anterior, una ilustración se recoge por medio de un análisis de clúster, que se considerada como una técnica multivariante para detectar patrones, relaciones y similitudes en volúmenes grandes de datos. Adicional, permite asociar los clientes que ostentan características comunes y en consecuencia facilita las decisiones estratégicas. Además, en el área financiera, se aplican métodos como k-means, clúster jerárquico y DBSCAN, estos son utilizados en el análisis de los perfiles crediticios, con la intención de mejorar su gestión.

Lawrence et al (2010), los investigadores utilizan un conjunto de datos específico, se consideró variables como ingresos, actividad crediticia, edad, además los autores declaran que se simularon datos a partir de datos reales, por otro lado, los métodos aplicados por los investigadores fueron modelo discriminante multigrupo, acompañado de técnicas estadísticas, entre los resultados, la clasificación de tres segmentos diferentes y así mejorar significativamente el mercado de servicios financieros.

Borges et al (2014), los autores plantean en su investigación la relación existente entre los procesos de concesión de créditos y los riesgos de incumplimiento, en ese sentido, se utilizan métodos como análisis discriminante, redes neuronales y análisis de conglomerados K-

means, el primero le sirve para clasificar a los clientes según su historial crediticio, el segundo para clasificar el riesgo crediticio y el tercero para segmentar y comparar los datos.

Shokrgozar y Sobhani (2016) los autores analizan la segmentación de clientes bancarios mediante técnicas de minería de datos, entre sus aportes se recoge que, a partir de una evaluación de los segmentos de clientes, el sistema bancario considera otros factores más eficientes y cruciales para la toma de decisión y estimar de manera más precisa cada segmento de clientes, y en consecuencia tipos y montos de créditos más adecuados a sus características.

Goncarovs (2018), en su investigación resalta la división en la segmentación del cliente según sus características, en el procedimiento de análisis hace hincapié en la preparación adecuada de la data para una segmentación válida, en ese mismo sentido, entre sus conclusiones se rescata "un proceso de segmentación de clientes bien estructurado puede mejorar significativamente la capacidad de las instituciones financieras para entender a sus clientes".

Syakur et al. (2018), ellos analizan la integración del método K-Means y el método Elbow e identificar grupos de perfiles de clientes, en ese mismo orden de ideas, los autores sostienen que el método de agrupamiento K-Means genera muchos errores y malos resultados, para corregir dicha limitación se acompaña del método del coco (elbow), para una mejor asignación de los clusters.

El autor Ros (2022), en su investigación realiza un enfoque de agrupamiento evolutivo con la utilización de algoritmo genético y k-means, con la intención de segmentar a los clientes y en consecuencia identificar las características importantes de los clientes, y así predecir nuevos grupos de segmentos de clientes y entender sus deseos y necesidades para comprenderlos en un nivel más profundo.

Gaol (2023), estudia la segmentación de clientes en función del algoritmo K-Means que "se utiliza para segmentar a los clientes en diferentes grupos en función de sus puntuaciones de RFM. Este método es una técnica de agrupamiento no jerárquica que divide los datos en clústeres distintos, lo que facilita el análisis de los segmentos de clientes".

En el mismo orden de ideas del autor, su aporte a partir de su investigación con la agrupación en clústeres de k-means arrojan grupo de clientes bien definidos, pero con distinciones muy significativas entre sus clústeres, esto es

de gran ayuda para las instituciones financieras para centrarse en segmentos de clientes determinados para las decisiones de cobranzas.

Amutha & Khan (2023), estos autores para alcanzar la caracterización de su objeto de estudio, aplican un modelo de aprendizaje automático, con lenguaje de programación en Python, a través de la agrupación de clústeres con valores de K-medias, además, entre sus resultados se rescata la identificación del comportamiento de los clientes y sus preferencias.

Shaikh & Khedkar (2023), los autores a partir de un análisis exhaustivo sobre la segmentación de clientes del sector bancario por medio de métodos de segmentación convencional como K-medias, cabe mencionar que entre sus conclusiones realizan una crítica a los métodos convencionales utilizado, porque tienen dificultades para captar patrones presentes en los datos bancarios complejos.

Krishnan (2024), considera importante en el análisis de la segmentación del cliente el enfoque tradicional en base a aspectos demográficos, es decir, que la segmentación depende de estos aspectos, como la edad, los ingresos y la ubicación, en ese sentido, se comprende el perfil de los clientes, no obstante, se carecía de información detallada sobre su comportamiento.

Kumar (2024), el autor en su investigación hace uso de varios métodos que complementan el estudio, el análisis de conglomerados y el aprendizaje forman parte del estudio sobre la segmentación de clientes. Entre sus resultados, se destaca que los clientes necesitan servicios y estrategias de marketing más personalizados, a partir de tres grupos diferenciados el historial, preferencias bancarias y datos demográficos.

Créditos en el Ecuador

Reyes et al. (2023), el presente artículo aborda un estudio de los factores determinantes en el otorgamiento de créditos a la microempresa en el Ecuador, en el cual aplican un análisis de revisión bibliográfica, entre los factores claves encontrados se tiene el carácter, la capacidad el capital, condiciones y garantías conocida como las 5c.

Samaniego-Namicela & Armas Herrera (2024), la investigación realizada por los autores tiene como enfoque el estudio de los factores que influyen en el acceso al crédito, tomando como data las solicitudes de préstamos presentadas por las asociaciones de la Economía Popular y Solidaria de Ecuador, para ello se aplica un modelo de regresión logística, que entre sus conclusiones claves se encuentra que la ubicación rural influye en el acceso al financiamiento.

El presente artículo aborda el problema de la limitación en la segmentación efectiva del tipo de crédito asociado con las características de cada provincia del Ecuador e influenciadas por factores económicos, sociales y productivos. Por ejemplo, ciertas provincias reciben una mayor proporción de créditos debido a las características productivas dominantes y relacionadas con los diferentes tipos de crédito que oferta la Corporación Financiera Nacional.

Finalmente, se evidencia que la distribución de préstamos otorgados a las provincias del Ecuador por parte las instituciones financieras están influenciadas por una variedad de factores como las condiciones socioeconómicas, las características demográficas, entre otras. Por lo tanto, estos factores desencadenan desigualdades en el acceso a créditos en las regiones del Ecuador.

MATERIALES Y MÉTODOS

A partir de la presente metodología se busca dar respuesta a la pregunta ¿cuáles son los factores que influyen en la distribución de créditos otorgados en las diferentes provincias del Ecuador?, por lo tanto, esta sección describe el enfoque metodológico aplicado al análisis clúster, la intención es segmentar a los clientes por regiones en la concesión de créditos otorgados por la Corporación Financiera Nacional.

Diseño del Estudio

El enfoque metodológico es cuantitativo, de tipo descriptivo-predictivo, además el análisis clúster Mehmed Kantardzic (2020, p. 317) lo define como un conjunto de métodos de clasificación, en ese mismo sentido, es adecuado para la segmentación de regiones en el Ecuador, con la intención de identificar patrones y segmentos homogéneos en las diferentes provincias del país.

Recolección de la Data

Los datos se obtuvieron de fuente secundaria a través de datos abiertos de la Corporación Financiera Nacional para el año 2023. Se incluyen

variables monto otorgado, tipo de crédito, tipo de operación y provincias. Se excluyen otros registros para garantizar la calidad del análisis.

Preprocesamiento de la Data

Se aplica un procesamiento de la data en función de las provincias del Ecuador, se procede a totalizar los montos otorgados de 20 provincias de las diferentes regiones en relación al tipo de crédito (microcrédito, crédito corporativo, crédito empresarial y crédito pymes). A posteriori, los datos fueron normalizados para asegurar una misma escala para todas las variables.

En el mismo orden de la idea anterior, es importante definir la naturaleza de los tipos de crédito para la relación de patrones con los montos otorgados, el microcrédito son aquellos créditos de pequeña cuantía y corta duración regularmente destinados para emprendimiento y negocios familiares

El crédito corporativo, en cambio está destinado para grandes empresa o corporaciones con ventas brutas anuales superiores \$5.000.000,00; se caracterizan por tener un gran poder comercial en el mercado y regularmente operan de manera internacional.

Las empresas Pymes (pequeñas y medianas empresa) operan a nivel nacional apuestan por la innovación y el desarrollo industrial. Las ventas anuales brutas de las pequeñas oscilan entre \$100.001,00 a \$1.000.000,00 y las de las medianas empresas oscilan entre \$1.000.001,00; a \$5.000.000,00. En cambio, el crédito empresarial busca financiar sus actividades productivas de las empresas, independiente del tamaño de la empresa, sus ventas brutas y el número de empleados.

Técnicas de Análisis Clúster

Según Alboukadel Kassambara (2017) el clustering es considerado uno de los métodos importantes dentro de la minería de datos para descubrir conocimiento de datos multidimensionales. Además, su objetivo es identificar patrones o grupos de objetos con características similares que se extraen de un conjunto de datos.

En el mismo orden de la idea anterior, en la literatura también es conocido como aprendizaje automático no supervisado, considerado no supervisado porque guía por ideas a priori sobre que variables o muestras

pertenecen a qué conglomerados. Y aprendizaje porque el algoritmo que utiliza la máquina aprende a agrupar.

Análisis de Clustering Jerárquico (HCA)

Introducción

El análisis de clustering jerárquico (HCA, por sus siglas en inglés) se considera una técnica de agrupamiento que organiza observaciones en base a una estructura jerárquica en función a similitudes. En comparación con otros métodos como K-Means, el HCA no requiere de una especificación en el número de clúster, es a priori, en ese sentido, resulta útil para la exploración de los datos.

Fundamentos Teóricos del Clustering Jerárquico

Sea X una matriz de datos con I filas y J columnas en que cada fila representa una observación y cada columna una variable. En ese mismo orden, se define una métrica de distancia entre las observaciones con el objetivo de construir una matriz de disimilitud D , que incluyen elementos d_{ij} que representan la distancia entre las observaciones i y j .

Matriz de distancias D :

$$D = d_{ij} \text{ para } i, j = 1, \dots, I$$

Donde puede calcularse mediante la métrica Euclidiana, Manhattan u otras funciones de distancia.

Método de enlace: Establece cómo se combinan los clústeres en cada interacción del algoritmo, para su fin, utilizan diferentes enfoques, entre los más utilizados el WARD, cuya técnica minimiza la varianza intraclúster en cada combinación.

Además, el clúster jerárquico se caracteriza por se basa en la construcción de dendogramas, estos representan de manera gráfica la jerarquía de agrupación y de manera paralela determinar el número óptimo de clústeres, para ello utiliza métodos visuales o través de métricas como el coeficiente de cophenética e índice de Dunn.

Aplicación del Software

Para la aplicación del Análisis de Clustering Jerárquico en R se utiliza paqueterías (cluster, factoextra, NbClust y clusterCrit), entre sus cálculos se encuentran: la matriz de distancia, el método de algoritmo jerárquico y validación del modelo.

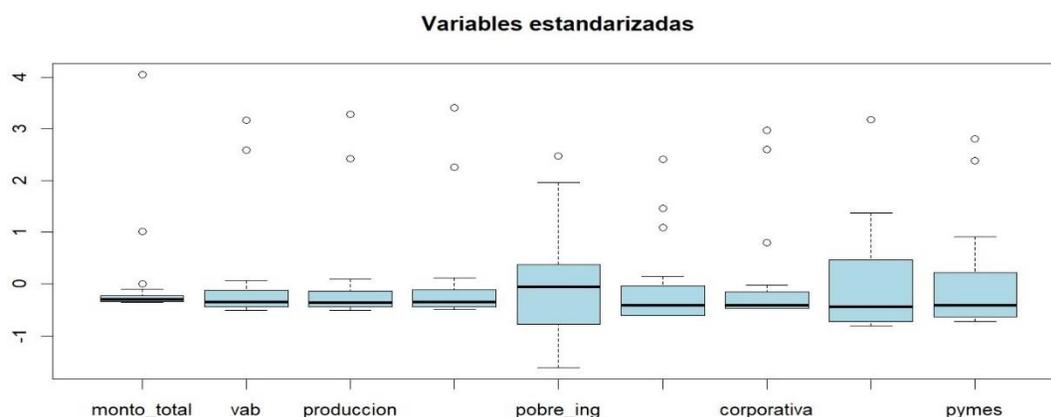
RESULTADOS Y DISCUSIÓN:

La data está compuesta por 9 variables y 20 observaciones que corresponden a las provincias del Ecuador, la misma es de corte transversal del año 2023 y corresponden a créditos otorgados por la Corporación Financiera Nacional CFN, cuya misión es "Contribuir a la transformación del modelo productivo nacional otorgando soluciones financieras, asegurando el desarrollo integral de oportunidades".

El objetivo del Análisis de Clúster Jerárquico es encontrar patrones asociados a la distribución de los créditos y las características inherentes a cada provincia y factores socioeconómicos: total de monto por provincia, valor agregado bruto, producción, consumo interno, tasa de pobreza por ingreso (%) y tipos de créditos otorgados por la CFN.

El primer gráfico (figura 1) generado corresponde a diagramas de caja, cuyo propósito es identificar la presencia de datos atípicos. La existencia de outliers sugiere la necesidad de aplicar técnicas para contrastar la varianza, lo cual se manifiesta en este caso de estudio. Como se observa, la presencia de puntos fuera de la caja y los bigotes evidencia la existencia de valores atípicos.

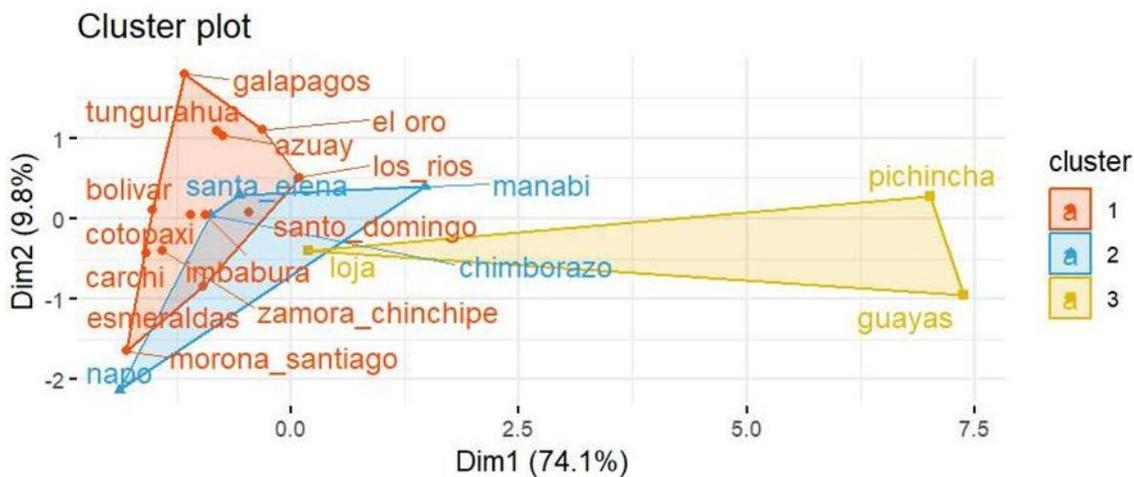
Figura 1. Gráficos para identificar outliers



Este gráfico identifica que todas las variables cuentan con datos atípicos, aunque los valores dispersos son pocos, no obstante, los outliers pueden afectar de manera significativa el análisis de clúster jerárquico por la distorsión en las distancias entre observaciones, en la alteración en la formación del dendograma y los métodos de enlace, en consecuencia, la limitación para explicar la interpretación de los clústeres.

Para mitigar los outliers, se aplica la técnica de distancia de Mahalanobis con el objetivo de mejorar la cohesión dentro de los clústeres. Al evaluar el modelo mediante métricas como el índice de Dunn y gráficamente a través de la silueta, los resultados obtenidos fueron un índice de Dunn de 0.063 y un valor de silueta de 0.01. En este contexto, los resultados sugieren que el modelo de clústeres no es óptimo (ver figura 2).

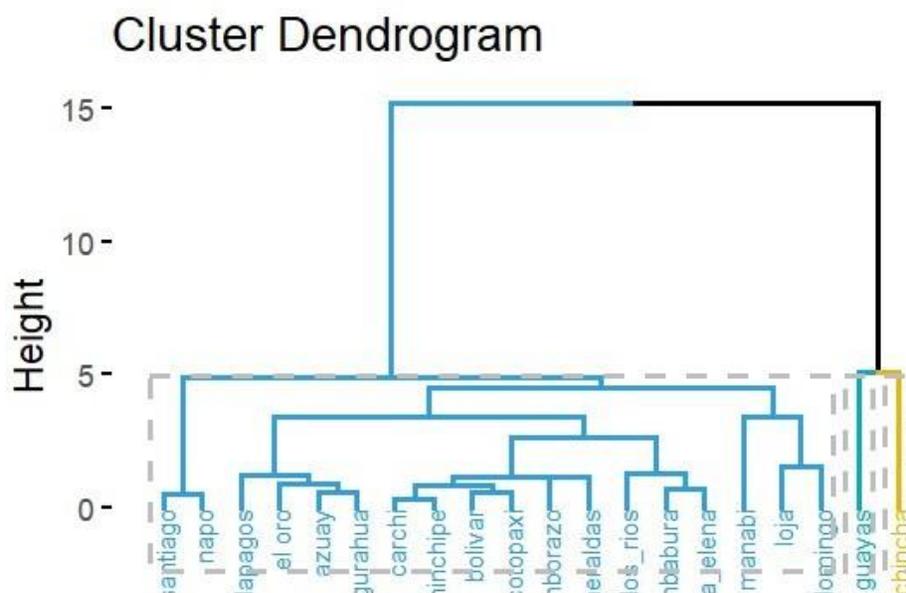
Figura 2. Análisis Clúster Jerárquico con Distancia Mahalanobis



El dendograma (figura 3) se muestra el número óptimo de clústeres que la data necesita para el clúster jerárquico, corresponde a la representación gráfica del árbol jerárquico que muestran la forma de agrupación de las observaciones en diferentes niveles de similitud, generada por la función

hclust(). Los dendrogramas pueden generarse en R con la aplicación de la función en base a plot (res.hc).

Figura 3. Dendrograma para el número de clúster óptimo



A partir de las características y de la naturaleza de las variables y en función del clúster dendrogram, su clasificación es de 3 clústeres.

Diferentes resultados se ofrecen en el Análisis de Clustering Jerárquico, en la tabla # 1 se observa la clasificación de los 3 grupos de clústeres con sus respectivas provincias. Donde en el primer clúster captura la mayor parte de las provincias del Ecuador, dejando a las provincias más representativas del país con clústeres independientes cada uno.

Tabla 1. Lista de las provincias en los diferentes clústeres

```

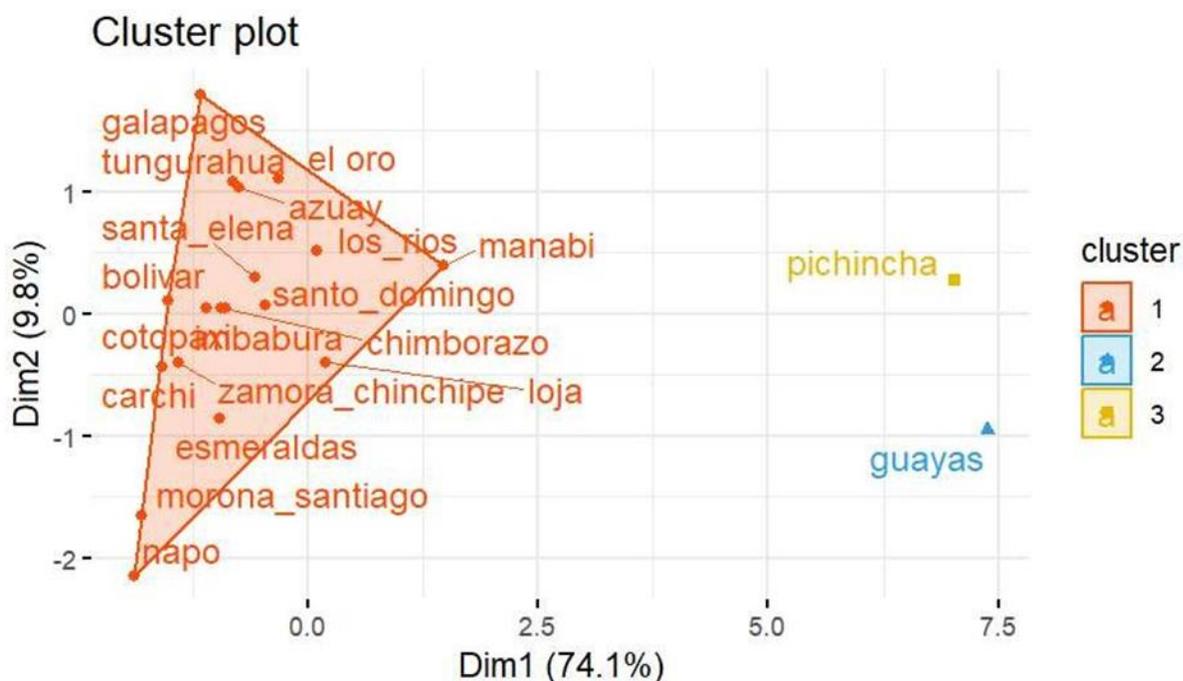
> rownames(df)[grp == 1]
 [1] "azuay"           "bolivar"         "carchi"          "chimborazo"
 [5] "cotopaxi"        "el oro"          "esmeraldas"     "galapagos"
 [9] "imbabura"        "loja"            "los_rios"       "manabi"
[13] "morona_santiago" "napo"            "santa_elena"    "santo_domingo"
[17] "tungurahua"      "zamora_chinchi"
> rownames(df)[grp == 2]
 [1] "guayas"
> rownames(df)[grp == 3]
 [1] "pichincha"
> |

```

El análisis clúster es una técnica estadística utilizada para identificar estructuras ocultas en una base de datos. En el contexto económico y financiero, como el presente estudio segmenta las regiones según la similitud y características socioeconómicas.

La presente investigación aplica el método de clúster jerárquico para el análisis la distribución de créditos otorgados por la CFN y variables socioeconómicas de las provincias del Ecuador.

Figura 4. Visualización del Análisis de Clúster Jerárquico



El dendrograma permite identificar tres grupos diferenciados de provincias, a partir de sus características económicas y financieras. El clúster 1 (color rojo) concentra la mayor parte de las provincias y comparten características socioeconómicas similares, como los niveles de producción, consumo interno, tasa de pobreza monetaria, consumo interno y acceso a los diferentes tipos de crédito y un nivel intermedio de acceso a créditos por parte de la CFN.

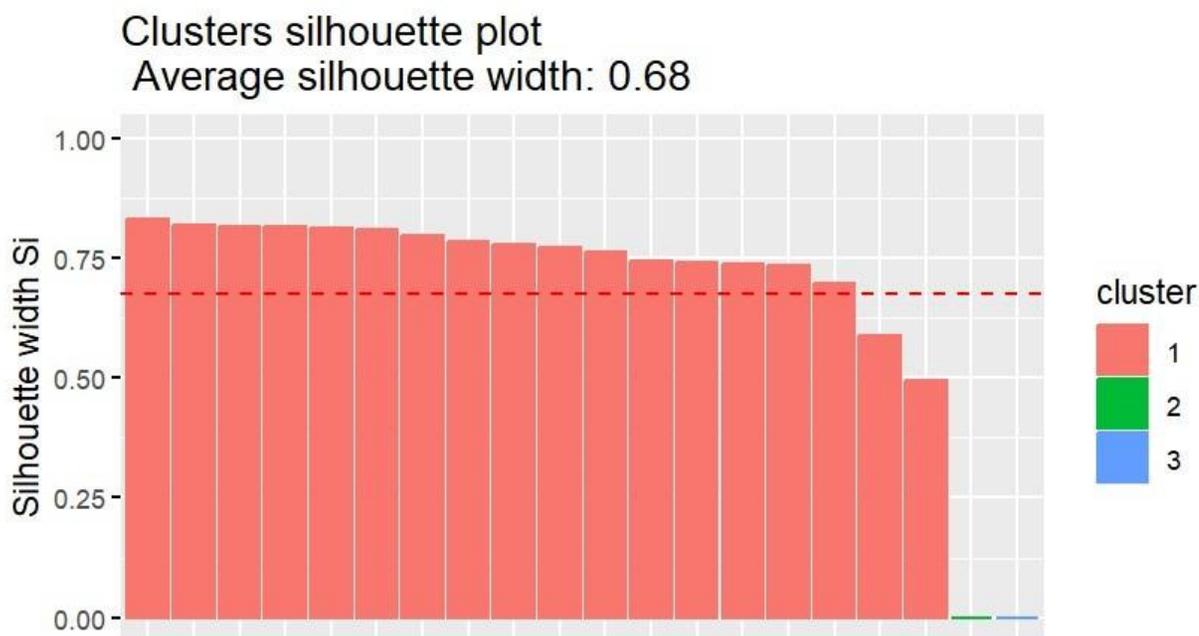
En el clúster 2 (color azul) contiene una sola provincia Guayas, este se encuentra alejado del resto de provincias (clúster 1), indicando diferencias significativas en variables socioeconómicas. Esto implica que Guayas ostenta características únicas, como niveles altos de crédito y una mayor actividad económica, pero también al nivel alto de pobreza por ingreso.

En el clúster 3 (color amarillo) se encuentra representado Pichincha, al encontrarse también alejado comparte similitudes con Guayas, pero de manera paralela presenta diferencias

que justifican su distancia con el resto de clúster. Las similitudes que comparte con Guayas son niveles altos de crédito, pero un nivel bajo de producción y pobreza monetaria.

Por otro lado, los ejes Dimm 1 (74.10%) y Dimm 2 (9.80%) capturan el 83.90% de la variabilidad de los datos, lo que indica que estos dos componentes resumen de manera efectiva la estructura del clúster. Guayas y Pichincha presentan un comportamiento diferenciado respecto al resto de las provincias, lo que sugiere que ambas se distinguen en términos de indicadores sociales y económicos. El clúster 1 agrupa economías homogéneas en cuanto a producción y acceso al crédito, mientras que Guayas se caracteriza por su industrialización y Pichincha por su alta empleabilidad en el sector público, dado que en esta provincia se encuentra la capital, Quito.

Figura 5. Cluster Silhouette plot



El gráfico evalúa la calidad del clúster jerárquico, el objetivo es medir las distancias, en otras palabras, que tan bien separadas o cohesionadas están las observaciones con respecto a los clústeres, como se observa en el gráfico el promedio es 0.68, es un indicador de un buen rendimiento, considere las provincias del primer clúster se encuentran cercano a 1, eso indica que las observaciones están bien agrupadas

dentro del clúster. Además, los clústeres 1 y 2 tienen valores de silueta bajos, entonces pueden no estar bien definidos.

Referencias Bibliográficas

Amutha, R., & Khan, A. A. (2023). Customer segmentation using machine learning techniques.

<https://doi.org/10.52783/tjjpt.v44.i3.653>

Banco Central del Ecuador. (s.f.). Cuentas Nacionales Anuales. https://contenido.bce.fin.ec/documentos/informacioneconomica/cuentasnacionales/ix_cuentasnacionalesanuales.html#

Borges, V. A., Lima, F. G., Junior, T. P., & Gaio, L. E. (2014). Impact of the segmentation for models of analysis in granting of credit: A comparison of results. *Business Management Review*, 3(10), 1–16. <https://repositorio.usp.br/item/002707621>

Corporación Financiera Nacional. (s.f.). Sitio web oficial. <https://www.cfn.fin.ec/>

Gaol, F. L. (2023). Customer segmentation of personal credit using recency, frequency, monetary (RFM) and K-means on financial industry. *International Journal of Advanced Computer Science and Applications*, 14(4). <https://doi.org/10.14569/ijacsa.2023.0140417>

Goncarovs, P. (2018). Using Data Analytics for Customers Segmentation: Experimental Study at a Financial Institution. *International Scientific Conference Information Technology and Management Science Riga Technical University*. <https://doi.org/10.1109/ITMS.2018.8552951>

Krishnan, G. U. (2024). Customer segmentation for credit card customers: A comprehensive guide. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(5), 286–296. <https://doi.org/10.32628/cseit241051015>

Kassambara, A. (2017). *Practical Guide To Cluster Analysis In R Unsupervised Machine Learning*. STHDA (Statistical Tools for High-throughput Data Analysis, www.sthda.com/english)

Kumar, L. (2024). A Study on Customer Segmentation for Banking Sector Through Cluster Analysis: Ethical Implications. *Deleted Journal*, 31(5s), 57–66. <https://doi.org/10.52783/cana.v31.999>

Lawrence, K. D., Pai, D. K., Klimberg, R. K., Kudbya, S., & Lawrence, S. M. (2010). *Segmenting Financial Services Market: An Empirical Study of Statistical and Non-parametric Methods* (pp. 1061–1066). Springer, Boston, MA. https://doi.org/10.1007/978-0-387-77117-5_68

Mehmed, K. (2020). *Data Mining Concepts, Models, Methods, and Algorithms Third Edition*.

IEEE Press Editorial Board.

Reyes, Y., Gallardo Solís, J. M., Toalombo Capuz, M. M., & Moscoso Jurado, D. E. (2023). Análisis del otorgamiento de créditos a la microempresa en Ecuador: estudio de los factores determinantes. *Salud, Ciencia y Tecnología*. <https://doi.org/10.56294/sctconf2023374>

Ros, L. (2022). Metaheuristic-based machine learning approach for customer segmentation. En Título del libro (pp. 101–133). Editorial. https://doi.org/10.1007/978-981-19-3888-7_4

Samaniego-Namicela, A., & Armas Herrera, R. (2024). Microcredit, Gender and Geography. Case of Banecador (pp. 218–228). Springer International Publishing. https://doi.org/10.1007/978-3-031-43733-5_20

Shaikh, A. S., & Khedkar, S. (2023). Clustering approach to high-dimensional data for banking customer segmentation. *International Journal of Advanced Research in Computer and Communication Engineering*, 12(8). <https://doi.org/10.17148/ijarcce.2023.12805>

Shokrgozar, N., & Sobhani, F. (2016). Segmentación de clientes de bancos basada en el descubrimiento de su relación transaccional mediante algoritmos de minería de datos. *Modelos y métodos matemáticos en ciencias aplicadas*, 10, 283. <https://doi.org/10.5539/MAS.V10N10P283>

Syakur, M., Khotimah, B., Rochman, E., & Satoto, B. (2018). Método de agrupamiento de integración K-Means y método Elbow para la identificación del mejor grupo de perfiles de clientes. *Serie de conferencias IOP: Ciencia e ingeniería de materiales*, 336. <https://doi.org/10.1088/1757-899X/336/1/012017>