



**Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

**Facultad de Ciencias Matemáticas**

**Escuela Profesional de Estadística**

**Segmentación de clientes de una empresa de casa de  
apuestas deportivas del Perú en el año 2023**

**TRABAJO DE SUFICIENCIA PROFESIONAL**

Para optar el Título Profesional de Licenciado en Estadística

**AUTOR**

Arturo Emerson LOPEZ VEGA

**ASESOR**

Mg. Emerson Damián NORABUENA FIGUEROA

Lima, Perú

2023



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

López, A. (2023). *Segmentación de clientes de una empresa de casa de apuestas deportivas del Perú en el año 2023*. [Trabajo de Suficiencia Profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Matemática]. Repositorio institucional Cybertesis UNMSM.

---

## Metadatos complementarios

<b>Datos de autor</b>	
Nombres y apellidos	Arturo Emerson Lopez Vega
Tipo de documento de identidad	DNI
Número de documento de identidad	47060042
URL de ORCID	<a href="https://orcid.org/0000-0002-5201-1233">https://orcid.org/0000-0002-5201-1233</a>
<b>Datos de asesor</b>	
Nombres y apellidos	Emerson Damián Norabuena Figueroa
Tipo de documento de identidad	DNI
Número de documento de identidad	45259683
URL de ORCID	<a href="https://orcid.org/0000-0003-2909-7080">https://orcid.org/0000-0003-2909-7080</a>
<b>Datos del jurado</b>	
<b>Presidente del jurado</b>	
Nombres y apellidos	Zoraida Judith Huamán Gutiérrez
Tipo de documento	DNI
Número de documento de identidad	09890094
<b>Miembro del jurado 1</b>	
Nombres y apellidos	Hugo Marino Rodríguez Orellana
Tipo de documento	DNI
Número de documento de identidad	40162362
<b>Datos de investigación</b>	
Línea de investigación	A.3.1.2. Análisis de Datos y Modelamiento de Problemas Sociales.

Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Lima Dirección: Av. Germán Amézaga s/n. Ciudad Latitud: -12.05642315 Longitud: -77.0843326901621
Año o rango de años en que se realizó la investigación	2023
URL de disciplinas OCDE	Estadísticas, Probabilidad <a href="https://purl.org/pe-repo/ocde/ford#1.01.00">https://purl.org/pe-repo/ocde/ford#1.01.00</a>



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú. Decana de América  
FACULTAD DE CIENCIAS MATEMÁTICAS  
ESCUELA PROFESIONAL DE ESTADÍSTICA

**ACTA DE SUSTENTACIÓN DEL TRABAJO DE SUFICIENCIA PROFESIONAL  
PARA LA OBTENCIÓN DEL TÍTULO PROFESIONAL DE LICENCIADO EN  
ESTADÍSTICA  
(PROGRAMA DE TITULACIÓN PROFESIONAL 2023)**

En la UNMSM – Ciudad Universitaria – Facultad de Ciencias Matemáticas, siendo las *11:30* horas del sábado 21 de octubre del 2023, se reunieron los docentes designados como Miembros del Jurado Evaluador (PROGRAMA DE TITULACIÓN PROFESIONAL 2023); Dra. Zoraida Judith Huamán Gutiérrez (PRESIDENTE), Mg. Hugo Marino Rodríguez Orellana (MIEMBRO) y el Mg. Emerson Damián Norabuena Figueroa (MIEMBRO ASESOR), para la sustentación del Trabajo de Suficiencia Profesional titulado: “**SEGMENTACIÓN DE CLIENTES DE UNA CASA DE APUESTAS DEPORTIVAS DEL PERÚ EN EL AÑO 2023**”, presentado por el señor **Bachiller ARTURO EMERSON LOPEZ VEGA**, para optar el Título Profesional de Licenciado en Estadística.

Luego de la exposición del Trabajo de Suficiencia Profesional, la Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

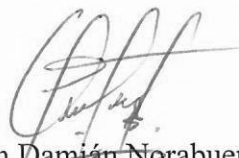
Realizada la evaluación correspondiente por los Miembros del Jurado Evaluador, el expositor mereció la aprobación ..... *Bueno* ....., con un calificativo promedio de *Dieciséis (16)*

A continuación, los Miembros del Jurado Evaluador dan manifiesto que el participante **Bachiller ARTURO EMERSON LOPEZ VEGA**, en vista de haber aprobado la sustentación de su Trabajo de Suficiencia Profesional, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las *12:00* horas se levantó la sesión firmando para constancia la presente Acta.

  
Dra. Zoraida Judith Huamán Gutiérrez  
PRESIDENTE

  
Mg. Hugo Marino Rodríguez Orellana  
MIEMBRO

  
Mg. Emerson Damián Norabuena Figueroa  
MIEMBRO ASESOR

## CERTIFICADO DE SIMILITUD

Yo, Emerson Damián Norabuena Figueroa en mi condición de asesor acreditado con Resolución Decanal N° 001622-2023-D-FCM/UNMSM del Trabajo de Suficiencia Profesional, cuyo título es “SEGMENTACIÓN DE CLIENTES DE UNA EMPRESA DE CASA DE APUESTAS DEPORTIVAS DEL PERÚ EN EL AÑO 2023”, presentado por el bachiller ARTURO EMERSON LOPEZ VEGA, para optar el título de Licenciado en Estadística.

Certifico que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante el software de similitud textual, el documento evaluado cuenta con el porcentaje de **10%** de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional**.

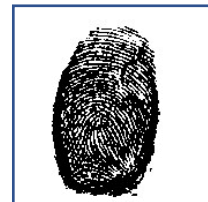
Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título correspondiente.



---

DNI: 45259683

Mg. Emerson Damián Norabuena Figueroa



Huella Digital

## Resumen

El trabajo se realiza para una empresa de casa de apuestas en la categoría de apuestas deportivas es muy importante contar con opciones diversas de publicidad que le brinde información para la respectiva toma de decisiones con las personas interesadas a dicho rubro, más aún que nos encontramos inmersos con diversidad de negocios con las mismas características, por lo expuesto es necesario conocer las características demandantes de los clientes. Para ello se desarrollará una segmentación en los usuarios con la finalidad de asociarlas por sus características particulares de conducta empleando la técnica de análisis de conglomerados como K medias. Cabe señalar que se emplea estas técnicas visto que la información recopilada tiene contiene datos atípicos, como también, está segmentada por faces, por lo cual se emplea indicadores de validación clúster cuya finalidad es realizar una buena segmentación en los usuarios y así tomar buenas decisiones.

**Palabras clave:** Segmentación de clientes, análisis de conglomerados, K medias, casas de apuesta, apuestas deportivas.



## **Abstract**

The work is carried out for a betting house company in the sports betting category. It is very important to have diverse advertising options that provide information for the respective decision-making with people interested in said field, even more so that we are immersed with a diversity of businesses with the same characteristics, therefore it is necessary to know the demanding characteristics of the clients. To do this, a segmentation of users will be developed in order to associate them by their particular behavioral characteristics using the cluster analysis technique such as K means. It should be noted that these techniques are used since the information collected contains atypical data, as well as, it is segmented by faces, which is why cluster validation indicators are used whose purpose is to carry out a good segmentation of users and thus make good decisions.

Keywords: Customer segmentation, cluster analysis, K means, bookmakers, sports betting.

# Índice

I.	Introducción .....	3
II.	Descripción de la Actividad .....	4
2.1.	Datos de la empresa o institución.....	4
2.1.1.	<i>Nombre de la institución</i> .....	4
2.1.2.	<i>Periodo de duración del TSP</i> .....	4
2.1.3.	<i>Razón Social</i> .....	4
2.1.4.	<i>Dirección</i> .....	4
2.1.5.	<i>Correo electrónico</i> .....	4
2.1.6.	<i>Organigrama de la empresa</i> .....	5
2.2.	Finalidad y Objetivos de la empresa .....	6
2.2.1.	<i>Finalidad</i> .....	6
2.2.2.	<i>Objetivos de la empresa</i> .....	6
2.3.	Descripción de la actividad.....	6
2.3.1.	<i>Organigrama del área</i> .....	7
2.3.2.	<i>Finalidad del trabajo</i> .....	8
2.3.3.	<i>Objetivos del Trabajo</i> .....	8
2.3.4.	<i>Problemática</i> .....	8
2.3.5.	<i>Metodología y procedimientos</i> .....	9
III.	Marco teórico .....	10
3.1.	Antecedentes .....	10
3.1.1.	<i>Antecedentes Nacionales</i> .....	10
3.1.2.	<i>Antecedentes Internacionales</i> .....	12
3.2.	Análisis Clúster.....	13
3.2.1.	<i>Función de distancia</i> .....	14
3.2.2.	<i>Distancia desde Mahalanobis</i> .....	14
3.2.3.	<i>Metodos jerárquicos y no jerarquicoso</i> .....	16
3.2.4.	<i>Promedios K</i> .....	17
3.2.5.	<i>Criterio de convergencia</i> .....	18
3.2.6.	<i>Prueba de pseudo F</i> .....	19
IV.	Resultados.....	19
4.1.	Entendimiento del negocio.....	19
4.2.	Entendimiento de la data .....	19
4.3.	Preparación de los datos .....	20
4.4.	Modelado y validación del modelo clúster .....	23
4.5.	Validación .....	24

<b>V. Conclusiones</b> .....	28
<b>VI. Recomendaciones</b> .....	29
<b>VII. Bibliografía</b> .....	30

## Índice de Tablas

<b>Tabla 1</b> Descripción de la Variable .....	37
<b>Tabla 2</b> indicadores de validación de clústeres .....	31
<b>Tabla 3</b> Segmentación de clientes por porcentajes .....	32

## Índice de Figuras

<b>Figura 1</b> Métodos jerárquicos y Métodos no jerárquicos.....	21
<b>Figura 2</b> Histórico de Ventas .....	25
<b>Figura 3</b> Matriz de Correlaciones .....	27
<b>Figura 4</b> Distancia de Mahalanobis .....	28
<b>Figura 5</b> Gráfico de Elbow Method .....	29
<b>Figura 6</b> Gráfico de Promedio de silueta para diferentes valores de K .....	30
<b>Figura 7</b> Gráfico de modelo de clusterización seleccionado con cinco clústeres .....	32





## **I. Introducción**

Este trabajo que se denomina suficiencia profesional esta elabora para una empresa en peruana que se dedica en la implementación de casas de apuestas y brinda una alternativa de generar ingresos a los clientes, en todo lo referente a apuestas deportivas, carrera de caballos, casinos virtuales, carrera de galgos, entre otros fuegos deportivos virtuales. Se encuentra en el mercado desde 2015, dedicada a la implementación y comercialización de juegos virtuales y apuestas deportivas, enfocados mayormente en casas de apuesta, la empresa cuenta con dos modos de venta online y retail.

Como en el año 2020, el contexto del covid 19 afecto gravemente a la mayoría de las empresas incluidas a las empresas del rubro de apuestas deportivas, donde las ventas sufrieron una caída considerable, presentando una reactivación económica muy lenta en la empresa. Cabe resaltar que, las ventas para el 2023 se están registrando mucho mejor de lo esperado, dado que se vino implantado mejoras para las facilidades de accesos en los clientes. Especialmente en las casas de apuestas retail y online, cabe señalar que, los perfiles de los clientes son muy varían como consumidor.

La empresa tiene como consolidar usuarios satisfechos con lo que ofrece la compañía; es por ello, con un objetivo claro de brindar mejores utilidades a sus asociados consumidores, se plantea ofrecer facilidades de pago y acceso a las apuestas deportivas. Mediante la técnica de clúster se quiere agrupar a los clientes por sus atributos de consumo, con la finalidad de aplicar una mejor técnica de marketing.



## **II. Descripción de la Actividad**

La empresa dedicada a la implementación de casas de apuestas con softwares de apuestas deportivas retail y online siempre con la mejor tecnología y ofreciendo asesoramiento profesional a todos nuestros clientes, acompañando y apoyando en todas las fases desde la planificación, construcción e impulso y crecimiento de sus propios locales de apuestas deportivas.

Ante el incremento de aperturas de casas de apuestas deportivas y la alta presencia de competencia de negocios del mismo rubro en el mercado, la empresa decidió en identificar las características comunes que cumplen los clientes potenciales, para así poder ofrecer los productos de la empresa y consolidar como clientes para la empresa.

### **2.1. Datos de la empresa o institución**

#### ***2.1.1. Nombre de la institución***

KINGTECH CORPORATION S.A.

#### ***2.1.2. Periodo de duración del TSP***

Del 1 de enero al 31 de mayo

#### ***2.1.3. Razón Social***

Derechos reservados

#### ***2.1.4. Dirección***

No aplica por términos de confidencialidad

#### ***2.1.5. Correo electrónico***

No aplica por términos de confidencialidad

2.1.6. Organigrama de la empresa

Figura 1

Esquema organizacional



## **2.1. Finalidad y Objetivos de la empresa**

### **2.1.1. Finalidad**

La tiene como finalidad de trabajar siempre con las mejores tecnologías y ofrecer asesoramiento profesional a todos nuestros clientes en la implementación de sus locales propios.

### **2.1.2. Objetivos de la empresa**

#### **a) Objetivo**

Ser una de las empresas líderes en el mercado de apuestas deportivas, empleando innovación tecnológica y lograr gratitud nacional y mundial.

#### **b) Misión**

Acompañar y apoyar a sus clientes en todas las fases, desde la planificación, construcción e impulso y crecimiento de tu propio local de apuestas deportivas.

#### **c) Visión**

Obtener el reconocimiento como la mejor empresa en el rubro de casas de apuestas deportivas, acompañada con la innovación tecnológica y ganar prestigio nacional y mundial.

## **2.2. Descripción de la actividad**

La industria de casas de apuesta deportiva está conformada con un área que se dedica al planeamiento y marketing, con el objetivo de:

- Elaboración presupuestal para marketing basadas en los resultados históricos de ventas.
- Determinar estrategias de publicidad y venta con la finalidad de mejorar las ventas para así cumplir las expectativas trazadas.
- Elaboración de las promociones y campañas publicitarias.

- Capacitación y charla para dar a conocer los productos de la empresa.
- Realizar estudios de mercado con la finalidad de conocer a los clientes.
- Analizar permanentemente el comportamiento de las sedes de las casas de apuestas que abren y cierran.
- Reportar a gerencia los puntos necesarios para la mejora de la empresa por áreas.

### 2.2.1. Organigrama del área

**Figura 2**  
*Esquema organizacional*



Fuente: Elaboración propia

### **2.2.2. Finalidad del trabajo**

Este presente trabajo de suficiencia profesional sostiene como fin, agrupar la cartera de clientes inmerso ámbito mercantil, y mundial cuyo fin es que la empresa pueda tomar mejor las decisiones en las estrategias en cuanto a la captación de clientes.

### **2.2.3. Objetivos del Trabajo**

#### **a) Objetivo general**

El objetivo de este trabajo es segmentar la cartera de clientes que está inmerso en el ámbito mercantil, nacional e internacional para lograr tomaras mejores decisiones estratégicas de negocio.

#### **b) Objetivos específicos**

- Identificar los atributos en base a la información recaudada en la apertura de locales a los clientes consolidados.
- Análisis exploratorio a los clientes que logran abrir su propio local.
- Segmentación a la cartera de clientes dispuestos por la gerencia de marketing y comercio.
- Identificar perfiles de consumo para cada clúster.

### **2.2.4. Problemática**

La empresa es importante porque se dedica a implementar locales de casas de apuestas deportivas retail y online de todos los clientes.

La problemática que se presenta en la compañía es que su cartera de clientes donde se visualiza el comportamiento de cada uno de los clientes donde cada perfil es distinto, cabe resaltar que las decisiones empleadas en el desarrollo de la aplicación del marketing no lograron consolidar a un cliente, ya conociendo el detalle expuesto se requiere dar a conocer el perfil de los posibles clientes.

### **2.2.5. Metodología y procedimientos**

#### **a) Metodología**

Según Hernández S. R. (2014), este trabajo de suficiencia profesional tiende a ser un enfoque cuantitativo, dado que los cálculos de realiza la segmentación, según el nivel es de tipo descriptivo por que se conocerá los indicadores de comportamiento de cada cliente, cuyo diseño para esta investigación sería no experimental ya que no hay presencia de manipulación en las variables y observación a los clientes.

#### **b) Población y unidad de análisis**

La población está conformada ventas realizadas por los asociados de la empresa de casas de apuestas deportivas durante el periodo de enero a mayo del 2023, la unidad de Análís son cada uno de los asociados que cuentan con una casa de apuestas deportivas durante el periodo de enero a mayo deñ 2023.

#### **c) Técnica e instrumento de recolección de data**

El instrumento para la recolección de informacion es usar la data de todos los asociados registrados en la base de datos de la plataforma durante el periodo de enero a mayo del 2023 ya que nos muestra toda la información recaudada durante este periodo.

#### **d) Plan de procesamiento y análisis estadístico de datos.**

Se analizaron los datos, empleado el software estadístico R-studio (versión 4.3.1), Python (versión 3.10.0) y Microsoft Excel 2019, sistematizando el análisis de la siguiente mearcan.

- Comprender el negocio.
- Entender la data
- Depurar la data
- Modelación y aprobación del mismo

- lectura del modelo

### **III. Marco teórico**

#### **3.1. Antecedentes**

##### **3.1.1. Antecedentes Nacionales**

Santos A. (2022) cuya tesis “Marketing digital y segmentación del mercado en la empresa Enología Perú EIRL, Ica, 2022”, cuyo objetivo es la determinación del grado de la relación que hay entre Marketing digital con la segmentación mercantil, el método usado es de aplicada de tipo transversal, y por su alcance descriptivo correlacional con diseño no experimental. La población está conformado por 20 colaboradores que trabajaron en la empresa durante ese periodo y la unidad de análisis es cada uno de los trabajadores que trabajaron en esa empresa durante ese periodo, recolectando la información mediante la encuesta virtual, y posteriormente se realizó el análisis con ayuda del programa SPSS 25 obteniendo como resultado una correlación de  $r=0.746$ , así mismo el 5% tienen un nivel deficiente de manejo de marketing, 30% regular, 65% eficiente y 15% nivel deficiente, concluyendo que hay variables Marketing digital con segmentación del mercado están altamente relacionados.

Soto P. (2019) en su tesis que lleva como título “Segmentación y decisión de compra en la empresa TECREP S.A.C, San Isidro, Lima 2019”, con el objetivo de identificar si existe o no relación entre las variables segmentación con decisión para comprar, la metodología usada es de tipo aplicativa con el nivel correlacional con un enfoque cuantitativa no experimental, hipotético deductiva, cuya población es esta conformado por 30 vendedores por menor y por mayor que realizaron sus ventas en Lima, que venden productos tecnológicos del mes de enero hasta abril del año 2019, en la colección de la información se aplicó el cuestionario que posterior se realizó el análisis usando el paquete estadístico SPSS 25, obteniendo como resultado regresión de Spearman de 0.814 concluyendo que existe correlación positiva alta entre las dos variables, por lo tanto TECREP S.A.C está considerado como potencial opciones de compra.

Mendoza J., Saldaña F., y Vivanco S. (2021), en su trabajo de investigación que lleva por título “Identificación de obras urbanas para la Ciudad de Lima a través del uso de herramientas basadas en Machine Learning”. Empleando el método supervisado en Machine Learning para lograr una efectiva clasificación, recopilando la información de las redes sociales donde los ciudadanos realizaron demandas en las redes sociales, las demandas que más exigidas son obras acordes a la necesidad de los ciudadanos. Teniendo como resultado, presencia de patrones que llevan a decidir las prioridades en obras que debe realizar el Municipio. Concluyendo que la Máquina que sirve de soporte Vectorial tiene una exactitud del 78% clasificando el análisis de sentimientos, priorizando la implementación de parques y lugares comunes para reuniones sociales.

(Córdova & Torres, 2018). Rn en su tesis “Aplicación de minería de datos para pronosticar el riesgo de morosidad de los estudiantes de la Universidad Autónoma del Perú”, cuyo objetivo es poder implementar una herramienta tecnológica para que pueda identificar a un estudiante que tienen problemas con los pagos de su pensión. Empleando el tipo de investigación no experimental, la población de estudio está conformado por todos los alumnos que de la Universidad Autónoma del Perú, la unidad de análisis es cada



uno de los alumnos de dicha universidad, para analizar se empleó el uso del paquete estadístico spss 25, obteniendo como resultado que para detectar un alumno con problemas de pagos de pensión, se demora 2 mnts; la predicción del mismo, 3 mnts; para su respectivo reporte, 4 mnts; en cuanto a la detección de las cualidades, 5 mnts, concluyendo que el modelo es válido para poder tomar buenas decisiones con respeto a los alumnos con problemas de pagos de pensión.

Norabuena A. (2022) realizó un trabajo titulado “segmentación de obras paa una asignación correcta de equipos de trabajo en una empresa siderúrgica en el primer semestre del 2020” , tiene como objetivo lograr clasificar de manera correcta todas las obras que realizo la empresa durante el rango de tiempo mencionado por la empresa dedicada a la siderurgia, se empleó el uso de Machine Learnin para realizar dicha agrupación, obteniendo como resultado que se obtuvo 5 clusteres que están correctamente definidos, es decir nadie de los que conforman el grupo se encuentra debajo del cero eso significa que nadie de los integrantes de cada grupo pertenezca a otro grupo, llegando a la conclusión que se visualizaron las falencias que presenta la empresa y mejorara en cuanto el sistema de entrega de los materiales con anticipación, rapidez y en perfectas condiciones.

### ***3.1.2. Antecedentes Internacionales***

Nova, T., Caro, J., y Cruz, C., (2021) en su investigación titulada “Diseño de un modelo de evaluación de proyectos a partir de herramientas de Machine Learning o aprendizaje automatizado”, el objetivo es realizar un modelo para dar solución, usando como técnica, Machine Learning, a la problemática en cuanto al ejercicio de obras civiles en Colombia. el modelo se propuso a base de información históricos que contiene datos de los diez años últimos de obras civiles en el Colombia, con una muestra de un conjunto de datos de ensayo y de testeo de (95-5), obteniendo como resultado fracaso en los proyectos que comenzaron tarde su ejecución y las que demoraron en su entrega al tiempo estimado de. Como resultado obtuvo una medición exacta de 0.8, concluyendo que

al desarrollar este modelo con anterioridad se pudo haber prevenido aproximadamente 500 millones en obras por parte del estado colombiana.

### **3.2. Análisis Clúster**

Aggarwal, et al. (2016) el Análisis Clúster o también llamado análisis de conglomerados o incluso Clustering, esta técnica trata con Machine Learning, su aplicación es muy común ya que se estandariza en cualquier rubro industrial. Se emplea en el procesamiento de datos para segmentar. Según Hair et al. (2019) El dendograma obtenido se inspeccionó visualmente para determinar grupos obvios presente.

Básicamente, trata de formar grupos en un conjunto, las cuales están basadas en las distancias estadísticas entre observaciones. Los grupos se forman según las afinidades o similitudes que muestran las observaciones generando la segmentación. La distancia facilita conseguir una matriz de distancia de los (n) individuos con sus respectivas características (k) antes de que estas estén clasificadas. Para un gran número de variables, no es posible identificar visualmente los grupos, debido a la limitación gráfica del espacio tridimensional, sin embargo, se pueden utilizar otros criterios de agrupación, como medidas de distancia y otras medidas de similitud. en general.

- ✓ Medidas de distancia: medida de separación entre dos puntos, utilizada para variables métricas.
- ✓ Medidas correlacionales: permiten trabajar con variables categóricas.
- ✓ Medidas de asociación: cuando los individuos se agrupan en base a individuos cercanos entre si deben pertenecer a un mismo grupo.

### 3.2.1. Función de distancia

El análisis de conglomerados normalmente se basa en una función de similitud, una función que recibe dos objetos y devuelve la distancia entre ellos. Para ser determinados por una métrica de calidad, los grupos deben ser muy homogéneos internos y heterogéneos externos. Los individuos también pueden llamarse objetos, instancias, t-tuplas o registros. Cada individuo representa una entrada de datos que puede estar constituida por un vector de atributos que son campos numéricos o categóricos.

Los métodos de agrupamiento asumen que todas las relaciones relevantes entre objetos pueden describirse mediante una matriz que contiene una medida de disimilitud o proximidad entre cada par de objetos.

### 3.2.2. Distancia desde Mahalanobis

Según Aldas y Uriel (2018), La distancia de Mahalanobis es una métrica que difiere de la distancia euclidiana por tener en cuenta la correlación entre los conjuntos de datos, además, no es necesario que las variables tengan la misma unidad de medida. Su fórmula para la distancia entre dos vectores ( $X_i, X_j$ ) que tienen una matriz de covarianza  $\Sigma$  — matriz simétrica que resume la covarianza entre  $p$  variables viene dada por:

$$d(x_i, x_j) = \sqrt{(x_i, x_j)^t \Sigma^{-1} (x_i, x_j)}$$

El uso de la distancia de Mahalanobis corrige algunas de las limitaciones de la distancia euclidiana, ya que automáticamente tiene en cuenta la escala de los ejes de coordenadas. Como punto negativo, las matrices de covarianza pueden ser difíciles de determinar y la memoria y el tiempo de cálculo crecen cuadráticamente con el número de características. La métrica de Mahalanobis también se utiliza para medir la distancia entre

un elemento  $X$  y un grupo de elementos cuya media viene dada por  $\mu = (\mu_1, \mu_2, \dots, \mu_n)$  y que tiene una matriz de covarianzas dada por  $\Sigma$  (matriz elevada a la 2 que tiene la varianza y covarianza asociada a las respectivas variables de se investiga). En este caso, la distancia viene dada por la fórmula:

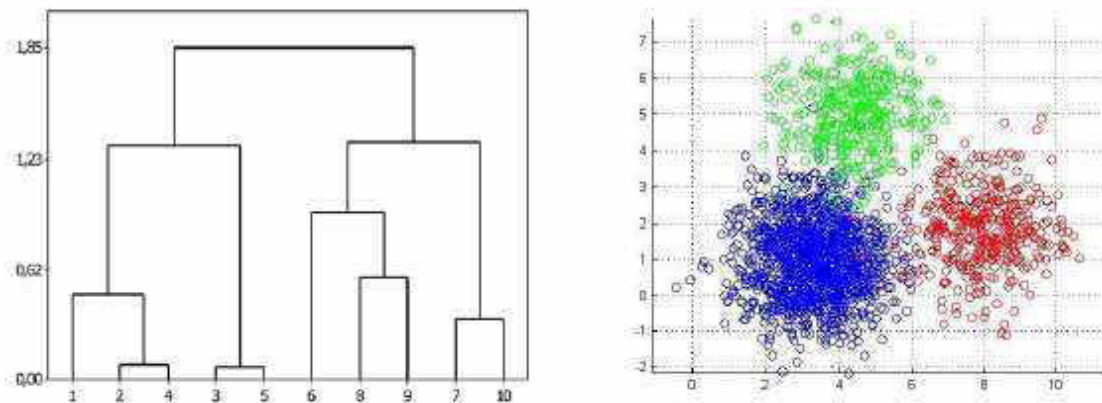
$$d = \sqrt{(x_i, x_j)^t \Sigma^{-1} (x_i, x_j)}$$

Conceptualmente, es como si la pertinencia de un elemento se evaluara no sólo por su distancia al centro (media) del grupo, sino también por su variabilidad, determinando así la distancia del individuo  $x_i$  en términos de una comparación con la desviación estándar del conglomerado. Cuanto mayor sea el valor de esta métrica, mayor será el número de desviaciones estándar que un individuo está lejos del centro del grupo, y menor su probabilidad de pertenecer a él.

### 3.2.3. Métodos jerárquicos y no jerárquicos

Según Amat Rodrigo (2017) los métodos jerárquicos y no jerárquicos se refieren principalmente a la forma en que se dividen u organizan los datos. Los métodos jerárquicos construyen una jerarquía de particiones, mientras que los métodos no jerárquicos construyen una partición de los datos.

**Figura 1:** Métodos jerárquicos y Métodos no jerárquicos



Según Amat Rodrigo (2017) Los algoritmos jerárquicos crean una jerarquía de relaciones entre individuos utilizando una medida de distancia. Los algoritmos no jerárquicos separan a los individuos en grupos según las características que poseen. Como EFD solicitó tres grupos de clientes, se utilizará un método no jerárquico.

Según Amat Rodrigo (2017) Los métodos no jerárquicos apuntan a encontrar directamente una partición de los individuos en grupos, de modo que la partición cumpla con dos requisitos básicos: similitud interna, también llamada “cohesión”, y aislamiento de los grupos formados, también llamada “separación” de grupos. Para encontrar la mejor partición de orden  $k$ , se deben emplear algunos criterios de calidad de partición. Computacionalmente, es imposible crear todas las posibles particiones de orden  $k$  y, tras conocer estas particiones, elegir

la más adecuada. Por lo tanto, se necesitan procesos que investiguen una parte de las posibles particiones para encontrar la partición casi óptima. Hay varias formas en las que los métodos no jerárquicos difieren de los jerárquicos. Primero, los denominados métodos que no son jerárquicos requieren que el estudioso especifique previamente la cantidad de  $k$  grupos deseados, en contraposición a las técnicas jerárquicas. En cada etapa de agrupación, se pueden formar nuevos grupos dividiendo o fusionando grupos ya combinados en pasos anteriores, cuando se utiliza un método no jerárquico. Es común que los algoritmos computacionales utilizados en los métodos no jerárquicos sean del tipo iterativo y, en comparación con los métodos jerárquicos, existe una mayor capacidad para analizar grupo de datos con un gran número de observaciones.

Según Amat Rodrigo (2017) Los métodos K- Means utilizados para particiones sin superposición de datos, el Fuzzy c-Means que es un método de agrupación superpuesta (en este caso los individuos pertenecen a todos los grupos con diferentes grados de pertinencia) y las redes neuronales artificiales aplicadas al análisis Cluster, que es un método que tiene la capacidad de manejar datos inexactos, incompletos o completamente nuevos, son ejemplos de métodos no jerárquicos. En este estudio, los sujetos deben pertenecer a un solo grupo y los datos son precisos, entonces se utilizará el método de K-Means.

#### 3.2.4. Promedios $K$

(Aggarwal, C., y Reddy, C.,2016) El método K – Means es uno de los más conocidos y utilizados para problemas prácticos. Y se puede resumir en cuatro pasos como se muestra a continuación.



En pocas palabras cada elemento de la muestra se asigna al grupo cuyo centroeide (vector de medias de la muestra) es el más cercano al vector de valores observados para el elemento respectivo. El proceso consta de cuatro pasos:

i. El investigador elige  $k$  centroides, también llamados "semillas" o "prototipos", para iniciar el proceso de partición.

ii. Luego, cada elemento del conjunto de datos se comparará con cada centroide, utilizando una medida de distancia. El elemento se asigna al grupo cuya distancia es la más pequeña.

iii. Después de aplicar el paso (ii) a todos los  $n$  elementos de la muestra, se calculan los valores del centroide para cada grupo formado. Luego, se repite el paso (ii), teniendo en cuenta los nuevos centroides de los grupos.

IV. Los pasos (ii) y (iii) deben repetirse hasta que todos los individuos de la muestra estén bien distribuidos en sus respectivos grupos. En otras palabras, hasta que no sea necesaria más reasignación de individuos, dado que cada individuo estará en el grupo cuya distancia al centroide del grupo será la más pequeña.

La elección de las semillas del agrupamiento inicial influye en el agrupamiento final, por lo que el investigador debe ser consciente de la elección de las semillas. La mayoría de software estadísticos tienen como estándar elegir las semillas iniciales mediante el uso de las  $k$  primeras observaciones de la base de datos. Sin embargo, el investigador puede especificar qué semillas quiere usar para iniciar el algoritmo. Este procedimiento puede traer buenos resultados cuando los  $k$  primeros elementos muestran discrepancia entre sí y no se recomienda cuando son similares entre sí.

### ***3.2.5. Criterio de convergencia***

El valor por defecto del criterio de convergencia utilizado para saber cuántas iteraciones, cuántas veces se recalcularán los centroides y se reasignarán los individuos, es 0.0001. Las iteraciones se detienen cuando la máxima variación relativa en las semillas del grupo es menor o igual al criterio de convergencia. El cambio relativo en una semilla en el grupo es la distancia entre la semilla antigua y la semilla nueva dividida por un factor de escala, que en este caso es la distancia mínima entre las semillas iniciales

### 3.2.6. Prueba de pseudo – F

Según Amat Rodrigo (2017) La pseudo-F se basa en la varianza entre grupos en cada nivel de agregación. Por lo tanto, un valor de prueba alto es deseable e implicaría rechazar la hipótesis de homogeneidad entre los grupos creados. El estadístico pseudo-F está dado por:

$$F = \frac{\frac{R^2}{c-1}}{\frac{1-R^2}{n-c}}$$

dónde  $R^2$ , es el coeficiente de determinación global observado  $c$  es el número de grupos y  $n$  el número de observaciones.

## IV. Resultados

El trabajo de investigación se emplea la metodología CRISP-DM cuyos pasos a tener en cuenta se mencionan a continuación.

### 4.1. Entendimiento del negocio

La industria de casas de apuestas deportivas cuenta con información de venta en detalle de cada clientes. Como todas las empresas su objetivo es generar mayor rentabilidad para una mejor sostenibilidad, en las sedes de venta. El estudio se dirige a los asociados que cuentan con una casa de apuestas deportivas. Para poder llegar al objetivo se ve la necesidad de segmentar la cartera de asociados, cuya finalidad es conocer mejor a nuestros posibles clientes y poder decidir mejor en las propuestas de marketing y publicidad.

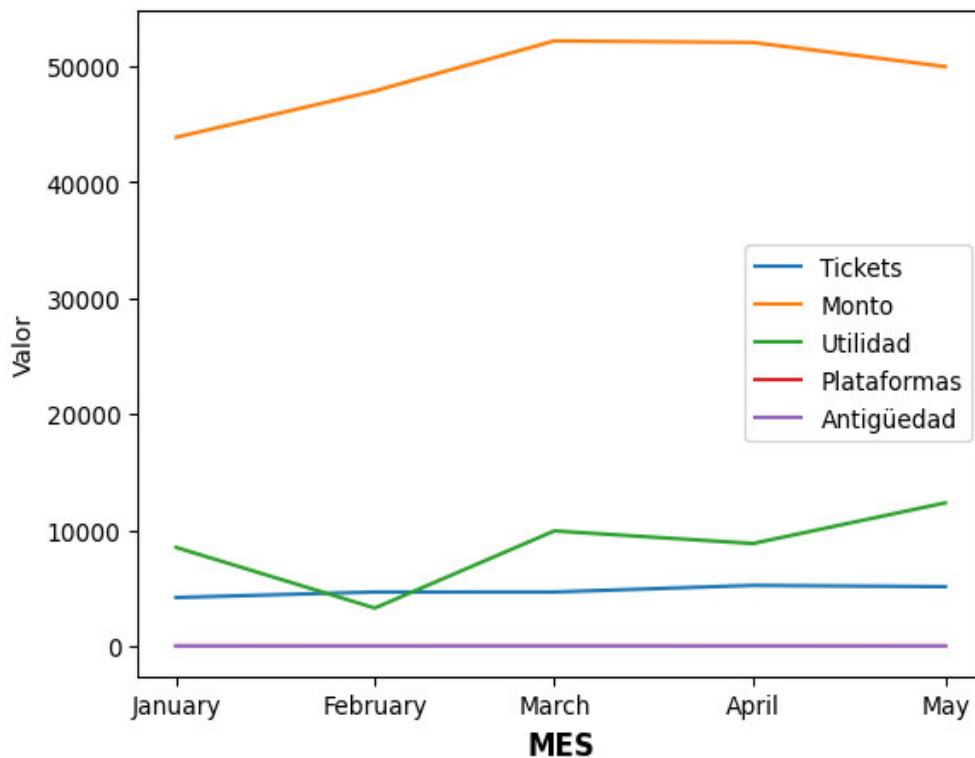
### 4.2. Entendimiento de la data

En esta etapa se emplea los datos registrados en la plataforma durante el periodo de enero a mayo del 2023 de las ventas de todas las sedes a nivel nacional e internacional, cuyo propósito es obtener una base de datos que nos proporcionan las ventas realizadas en cada



una de las sedes de las casas de apuestas como número de tickets, cantidad ingresada y utilidades. En la siguiente figura podemos observar los históricos de las ventas.

**Figura 2**  
*Histórico de Ventas*



fuelle: Elaboración propia

### 4.3. Preparación de los datos

Con el objetivo trasado debemos tener en cuenta sigilosamente el tratamiento de la data para su respectiva depuración, la data estaba conformado por 19 variables, pero para esta investigación solo se tomó en cuenta algunos de estas variables, es decir solo las variables cuantitativas. No consideramos las variables de tipo cualitativas ya que los algoritmos a utilizar son netamente para variables numéricas.

De toda la información que se tenía solo se consideró 3 variables que muestras lo necesario para desarrollar la investigación.

Las variables en estudio de mencionan a continuación:

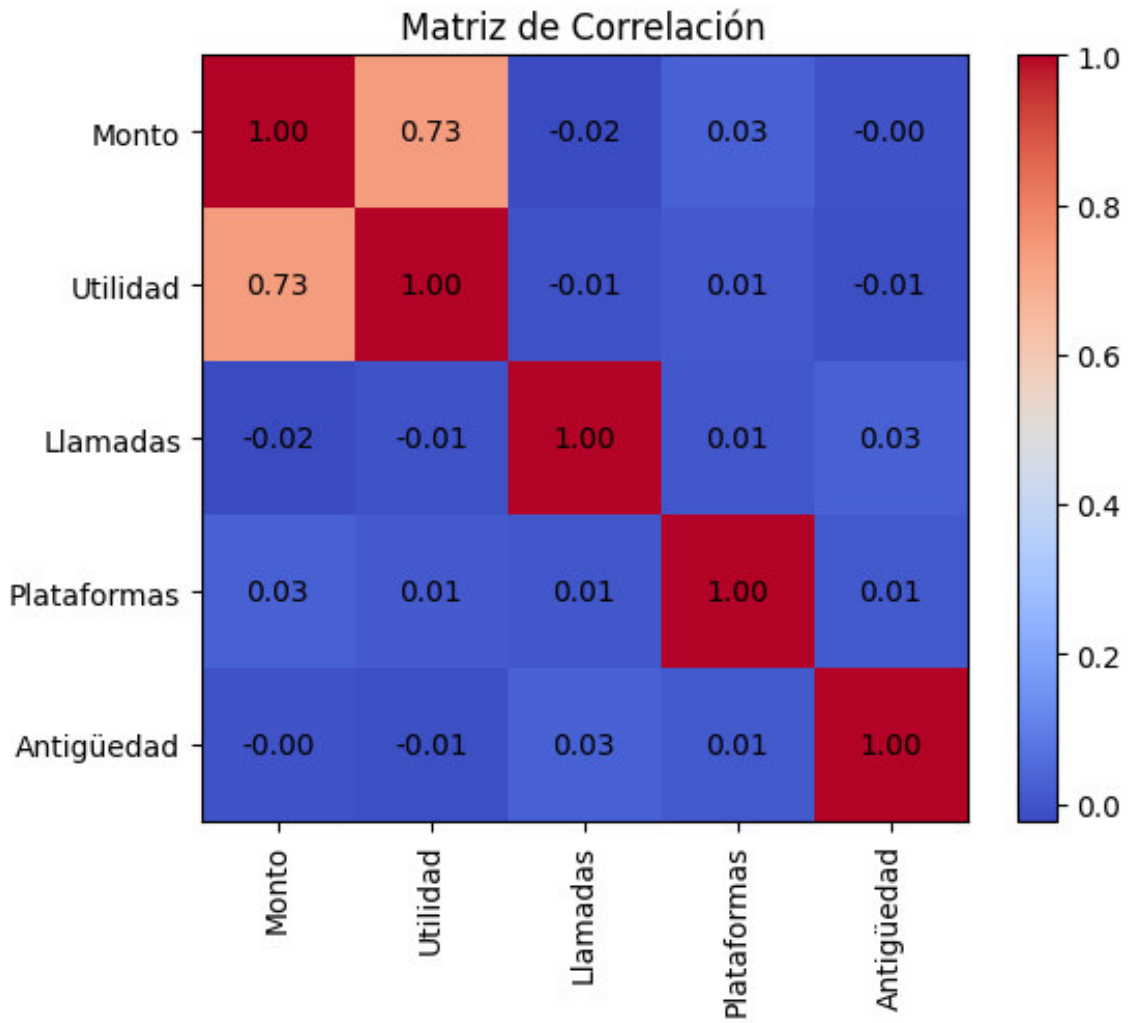
**Tabla 1***Descripción de la Variables*

Variable	Definición	Tipo de Variable
Número total de tickets	Numero de tickets vendidos en el periodo de enero a mayo 2023.	Cuantitativa discreta
Monto total ingresado	Monto total vendidos en el periodo de enero a mayo 2023.	Cuantitativa continua
Utilidad neta	Monto neto de las utilidades durante el periodo de enero a mayo 2023	Cuantitativa continua
Plataformas	Cantidad de plataformas durante el periodo de enero a mayo 2023	Cuantitativa continua
Llamadas	Número de llamadas durante el periodo de enero a mayo 2023	Cuantitativa continua
Tiempo	Tiempo de funcionamiento durante el periodo de enero a mayo 2023	Cuantitativa continua

*Nota.* Elaboración propia.

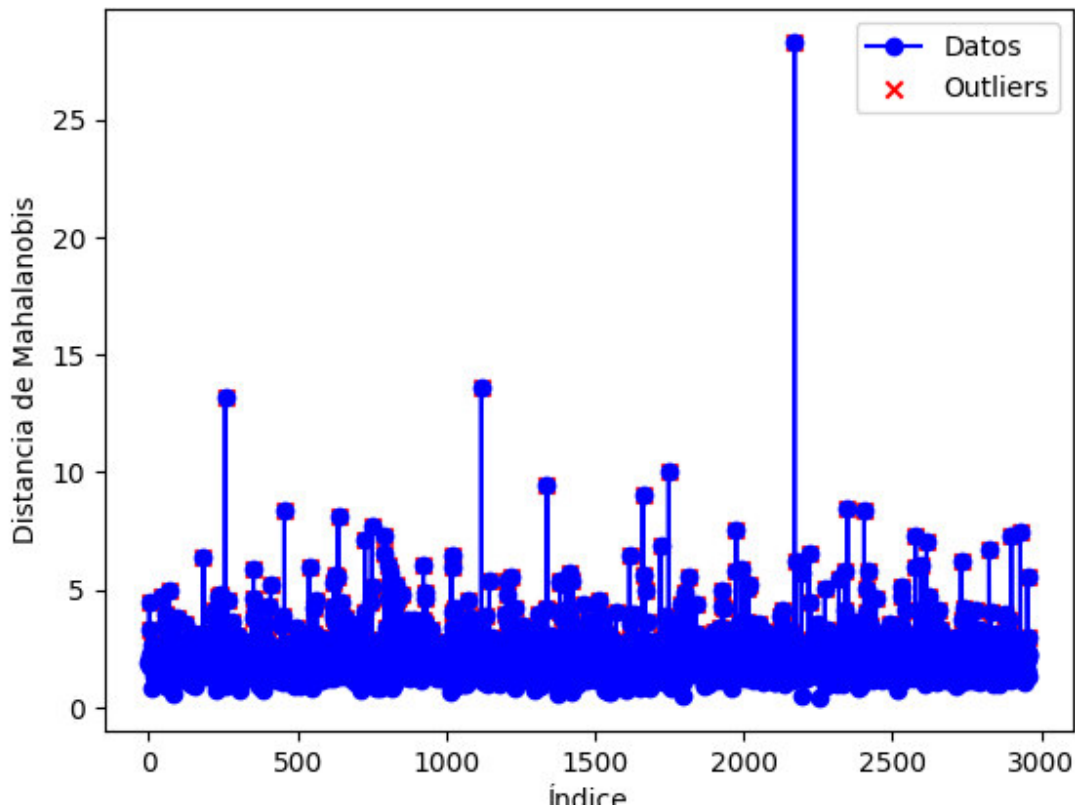
Una vez definida las variables a emplear en la investigación se procede a un análisis exploratorio de los datos, verificado la correlación que existe entre las variables tomadas en cuenta, para no presentar inconvenientes al momento de identificar las similitudes, donde visualizamos en la Figura 2 que existe una baja correlación en la mayoría de las variables, es decir que con este resultado podemos proceder al modelado ya que no presentaremos inconvenientes de sobredimensionar en modelo.

**Figura 3**  
*Matriz de Correlaciones*



Nota: Elaboración propia

**Figura 4**  
*Distancia de Mahalanobis*



Nota: Elaboración propia

#### 4.4. Modelado y validación del modelo clúster

En esta etapa se pondrá en práctica el método de clúster para poder solucionar los problemas que se visualizaron en la aplicación de la primera metodología, escogiendo cuidadosamente que método será el apropiado ya que en nuestra data se observan muchos datos atípicos, la cual se puede visualizar en la gráfica número, se pone en evidencia los datos alejados que serían no considerados en alguno de los grupos de los asociados que presentan caracteres particulares parecidos.

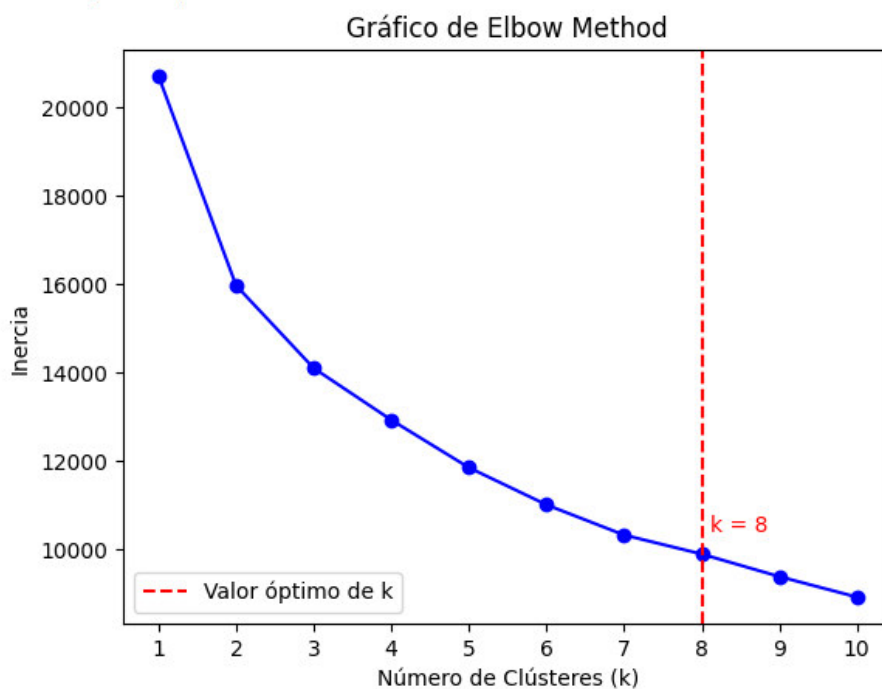
En esta oportunidad la técnica que se tendrá en cuenta es la de k medias o también llamada k means, este método se usa cuando las variables a ser usadas son de tipo numéricas, cabe señalar que dicha técnica está basada en observar los promedios de los asociados que conformaran un clúster.

Como un aporte importante de esta técnica es la determinación de la cantidad de clúster que se formaran. Ya que por anteriores decisiones tomadas por el área comercial y los

históricos de las cuales se afianzan para poder determinar la posible cantidad de clúster, lo que se debe realiza, es tener en cuenta el criterio óptimo para lograr el número exacto de clústeres. Es por eso que se emplea la técnica de WSS o también llamada método de Elbow que se encarga de calcular la cantidad de grupos, así como también disminuye la suma de cuadrados de lo observado.

**Figura 5**

*Gráfico de Elbow Method*



Nota: Elaboración propia

El grafico nos indica que el tamaño adecuado de la cantidad de clústeres es 8, la cual tendremos en cuenta al momento de agrupar a los asociados

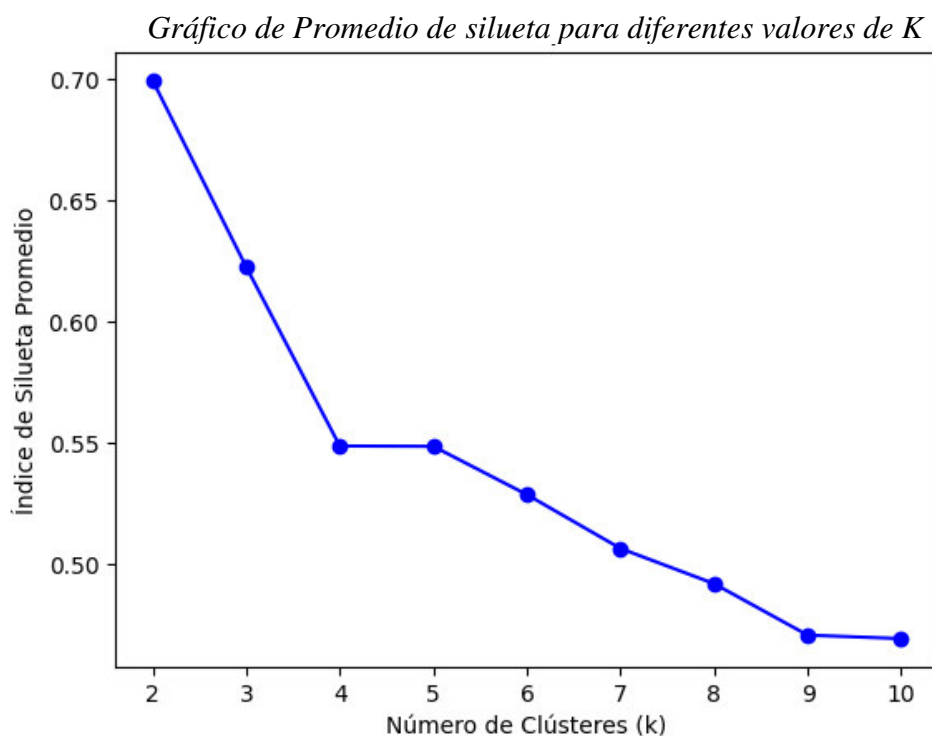
#### 4.5. Evaluación

En esta fase se realizará la validación y aprobación de los grupos o clústeres cuando se desarrolló al momento de realizar el modelo. Hay muchas maneras de evaluar para identificar la consolidación de los grupos, es este trabajo se considerará las validaciones internas y externas Amat, J. (2017).

#### 4.5.1 Validación interna

La alternativa que se tiene en consideración para confirmar la cantidad correcta de los clústeres se empleará el promedio de Índices de Siluetas, los cuales nos mostrarán el promedio de las diferentes magnitudes de clústeres, siendo considerado mejor el que presenta el más alto.

Figura 6



Nota: Elaboración propia

En la gráfica se muestra la evolución de los promedios de cada índice de las siluetas, en donde nos indica que el que posee el tamaño más grande será que le mejor formación de clúster tendrá, en este caso será el clúster 2.

La Puntuación de Siluetas nos indica el grado de confiabilidad en la que están conformados los clústeres dándonos a conocer la calidad de los asociados que lo conforman. También en la gráfica 5 podemos observar que los clústeres conformados esta debidamente definidas ya que ninguno de los individuos se encuentra por debajo del cero, es decir que la agrupación de los individuos es correcta y no existe la posibilidad de que los miembros de

cada conglomerado estén en otros conglomerados.

**Tabla 2**  
*indicadores de validación de clústeres*

Variable	Definición
Medida de inercia	2.345
Índice Dunn	0.1521
Índice Davies-Bouldin	0.102849

Existe otra manera de validar la correcta agrupación de clústeres son los índices de Dunn y Davies-Bouldin y medidas de inercia. En la table numero 2 la medida de inercia obtenida es 2.345, representa la baja distancias de los asociados que se encuentran en un mismo clúster. Por otros lados los índices de Dunn y Davies-Bouldin se acercan uno del otro, quieres decir que existe una distancia moderada entre clústeres, es así que confirmamos que los clústeres consolidados están debidamente agrupados.

#### ***4.5.2 Validación Externa***

La otra manera de validar es a base de la comparación de los resultados que se obtuvieron en las primeras fases, con el fin de establecer los indicadores que miden dentro de los clústeres, como están conformados y su respectivo porcentaje. Como podemos observar en la tabla numero 3 visualiza la cantidad de grupos y el porcentaje que les corresponde a cada clúster.

En la figura 6 podemos observar como se forman los clústeres en triple dimensión, en la que visualizamos lo evidente con respectos a las separaciones entre clústeres, donde la distancia entre cada individuo que forma parte del clúster es reducida.

**Tabla 3**

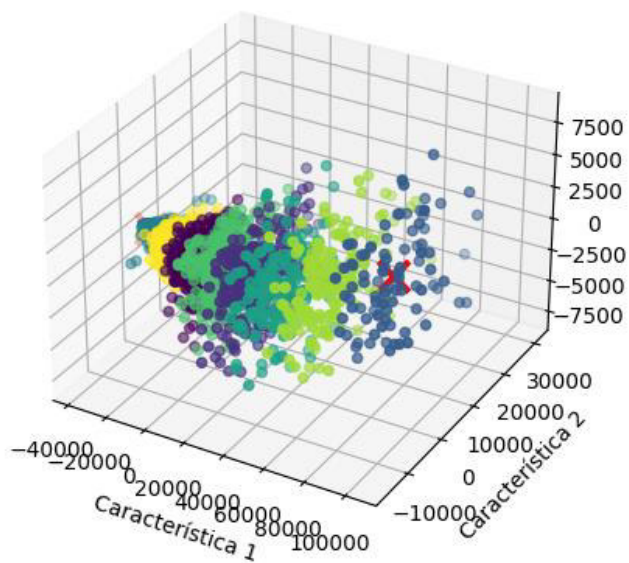
*Segmentación de clientes por porcentajes*

Cluster	K-MEANS	porcentaje
1	389	14%
2	181	7%
3	528	20%
4	277	10%
5	90	3%
6	572	21%
7	506	19%
8	153	6%
total	2696	100%

Nota: Elaboración propia

**Figura 7**

*Gráfico de modelo de clusterización seleccionado con cinco clústeres*



Nota: Elaboración propia

Ya que el modelo definido para esta segmentación nos permite conocer el perfil de los clústeres que se crearon con antecedentes del área comercial y marketing, lo cual nos permite establecer por parte de los especialistas en el ámbito a realizar un diseño para la captación de los nuevos asociados.



**Tabla 4***Análisis descriptivo del número de tickets vendidos por clúster*

<b>Cluster</b>	<b>Media</b>
Segmento 1	7221
Segmento 2	6785
Segmento 3	5850
Segmento 4	4655
Segmento 5	3785
Segmento 6	2785
Segmento 7	1785
Segmento 8	1003

Según la tabla 4, la cantidad de tickets vendidos durante el periodo de enero a mayo del 2023, se observó que la cantidad media más alta de la venta de tickets se encuentra en el segmento numero 1 con un total de 7221 tickets, seguido por el segmento 2 y 3 con una media total de 6785 y 5850 respectivamente, así mismos el segmento 8 presenta la menos cantidad de tickets con una media total de 1002.

**Tabla 5***Análisis descriptivo de los montos vendidos por clúster*

<b>Cluster</b>	<b>Media</b>
Segmento 1	S/43,846.20
Segmento 2	S/47,845.44
Segmento 3	S/72,161.67
Segmento 4	S/62,023.47
Segmento 5	S/49,942.39
Segmento 6	S/50,163.83
Segmento 7	S/33,785.00
Segmento 8	S/21,003.00

Según la tabla 5 del monto vendidos durante el periodo de enero a mayo del 2023, se observó que la cantidad media más alta del monto vendido se encuentra en el segmento numero 3 con un total de S/72,161.67, seguido por el segmento 4 y 6 con una media total de S/62,023.47 y S/50,163.83 respectivamente, así mismos el segmento 8 presenta la menos cantidad de ventas con una media total de S/21,003.00.

**Tabla 6***Análisis descriptivo de la cantidad de plataformas por clúster*

<b>Cluster</b>	<b>Media</b>
Segmento 1	5
Segmento 2	5
Segmento 3	7
Segmento 4	7
Segmento 5	4
Segmento 6	6
Segmento 7	2
Segmento 8	2

Según la tabla 6 de la cantidad de plataformas con las que cuenta cada asociado durante el periodo de enero a mayo del 2023, se observó que la cantidad más alta se encuentra en el segmento 4 y 5 con un total de 7 plataformas, seguido por el segmento 6 con 6 plataformas, el segmento 1 y 2 con un total de 5 plataformas, el segmento 5 con 4 plataformas y los segmentos 7 y 8 con 2 plataformas.

**Tabla 7***Análisis descriptivo de la antigüedad de años por clúster*

<b>Cluster</b>	<b>Media</b>
Segmento 1	3
Segmento 2	4
Segmento 3	5
Segmento 4	5
Segmento 5	4
Segmento 6	4
Segmento 7	1
Segmento 8	1

Según la tabla 7 que muestra la antigüedad en años con las que cuenta cada asociado durante el periodo de enero a mayo del 2023, se observó que el segmento 3 y 4 son los que presentan 5 años, los segmentos 2, 5, 6 que presentan 4 años de antigüedad, el segmento 1 con 3 años de antigüedad y los segmentos 7 y 8 presentan 1 año de antigüedad.

Clúster 1 que representa el 14%, está conformado por los asociados que registran un promedio de 7,221 tickets vendidos, tiene un promedio de S/43,846.20 en el monto vendido, poseen 5 plataformas de apuestas deportiva y tienen 3 años de antigüedad. El clúster 2 que representa el 7 %, está conformado por los asociados que registran un promedio de 6,785 tickets vendidos, tiene un promedio de S/47,845.44 en el monto vendido, poseen 5 plataformas de apuestas deportiva y tienen 4 años de antigüedad. El clúster 3 que representa el 20%, está conformado por los asociados que registran un promedio de 5,850 tickets vendidos, tiene un promedio de S/72161.67 en el monto vendido, poseen 7 plataformas de apuestas deportiva y tienen 5 años de antigüedad. El Clúster 4 que representa el 10%, está conformado por los asociados que registran un promedio de 4,655 tickets vendidos, tiene un promedio de S/62,023.47 en el monto vendido, poseen 7 plataformas de apuestas deportiva y tienen 5 años de antigüedad. El clúster 5 que representa el 3%, está conformado por los asociados que registran un promedio de 3,785 tickets vendidos, tiene un promedio de S/49942.39 en el monto vendido, poseen 4 plataformas de apuestas deportiva y tienen 4 años de antigüedad. El clúster 6 que representa el 21%, está conformado por los asociados que registran un promedio de 2785 tickets vendidos, tiene un promedio de S/50,163.83 en el monto vendido, poseen 6 plataformas de apuestas deportiva y tienen 5 años. El clúster 7 que representa el 19%, está conformado por los asociados que registran un promedio de 1,785 tickets vendidos, tiene un promedio de S/33,785.00 en el monto vendido, poseen 2 plataformas de apuestas deportiva y tienen 2 años de antigüedad. El clúster 8 que representa el 6 %, está conformado por los asociados que registran un promedio de 1,003 tickets vendidos, tiene un promedio de S/21,003.00 en el monto vendido, poseen 2 plataformas de apuestas deportiva y tienen 1 años de antigüedad.

## V. Conclusiones

Con la investigación realizada que consiste en realizar un modelo de cauterización para poder agrupar a los asociados de la empresa con la técnica de K-medias que es una técnica que no es sensible a los outlets, como también se empleó los indicadores Silueta, Davies Bouldin, Dunn y Conectividad las cuales no sirvieron para la validación de los clústeres. la empresa presentaba dificultades en la planificación de la parte comercial y marketing por la falta de conocimiento en el comportamiento de los asociados, con el trabajo realizado se llegaron a concluir lo siguiente:

- Nos permitió conocer la problemática que pasaba la empresa, de distinguir los perfiles de cada asociado.
- La cauterización muestra con claridad como los asociados se distinguen con la posesión de sus locales de casas de apuestas deportivas en zonas estratégicas, como también en la obtención del monto en las utilidades y el tiempo que permanece como asociado.
- Se logro la obtención de la segmentación de la cartera de asociados, de manera efectiva, logrando obtener 8 clústeres identificadas por sus características propias, ya que esto conlleva a las áreas encargadas planificar mejor las propuestas futuras en cuanto a las captaciones de nuevos asociados.

## **VI. Recomendaciones**

Cabe resaltar que la realización de este modelo de segmentación nos resultó favorable, hay puntos donde que se podría mejorar, ya que se visualizaron una alta cantidad de outlet las cuales quedaron fuera de la segmentación, es por ello que se podrían optar por otras técnicas de segmentación para logara trabajar con el universo completo de los registros obtenidos por la empresa en cuanto a sus asociados.

## II. Bibliografía

- Martínez, G. (2017). Una vision global del aprendizaje automatico. Aprendizaje automático, 14
- Data Science Process Alliance. (2020). Metodología CRISP-DM y la minería de datos actual. Recuperado de: <https://www.datascience-pm.com/blog/>
- Clemente Rivera, B. (2021). Segmentación de lectores digitales registrados de un sitio web informativo con el algoritmo de análisis cluster k-means. [Tesis de pregrado, Universidad Nacional Agraria La Molina] Archivo digital.  
<https://hdl.handle.net/20.500.12996/4845>
- Espinoza Zúñiga, J. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería Investigación y Tecnología*, 21 (01), 1-13. <https://doi.org/10.22201/ifi.25940732e.2020.21n1.008>
- Freitas, A. (2002). Data mining and knowledge discovery with evolutionary algorithms. Springer Science & Business Media
- Gallardo, J. (2020). *Introducción al Análisis Clúster*. Universidad de Granada, España.  
<https://www.ugr.es/~gallardo/pdf/cluster-g.pdf>
- García, Humberto & Gosling, Marlusa (2020, marzo). Cluster Analysis in Practice: Dealing with Outliers in Managerial Research. *Revista de Administração Contemporânea* (Vol.25, No. 1). <https://doi.org/10.1590/1982-7849rac2021200081>
- Gonçalves, G. (2017). *APLICAÇÃO DE ANÁLISE DE CONGLOMERADOS PARA A SEGMENTAÇÃO DE CLIENTES*. Universidade Federal de Uberlândia Faculdade de Matemática, Brasil.  
<https://repositorio.ufu.br/handle/123456789/25010>

Norabuena, A. (2022). *Segmentación de obras para la asignación correcta de equipos de trabajo en una empresa siderúrgica en el primer semestre del 2022.*

Universidad Nacional Mayor de san Marcos, Perú.

<https://cybertesis.unmsm.edu.pe/handle/20.500.12672/18624>

Landeo, J. (2023). *Propuestas de Mejora en la Segmentacion Comercial Para LAND EXPORT CORPORATION SAC.* Universidad del Pacifico, Perú.

[https://repositorio.up.edu.pe/bitstream/handle/11354/3911/Landeo%2C%20Jose Trabajo%20de%20suficiencia%20profesional\\_Administracion\\_2023.pdf?sequence=1&isAllowed=y.](https://repositorio.up.edu.pe/bitstream/handle/11354/3911/Landeo%2C%20Jose%20Trabajo%20de%20suficiencia%20profesional_Administracion_2023.pdf?sequence=1&isAllowed=y)

Santos, A. (2022). *Marketing digital y segmentación del Mercado en la Empresa Enología Perú EIRL, Ica, 2022.* Universidad Cesar Vallejo, Perú.

<https://repositorio.ucv.edu.pe/>