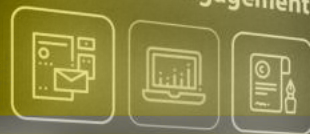


1^{era} Edición



Sponsored Listings

Shares and Engagement



SEM

SEARCH

MARKET
VALUE

Segmentación Efectiva de Clientes utilizando R

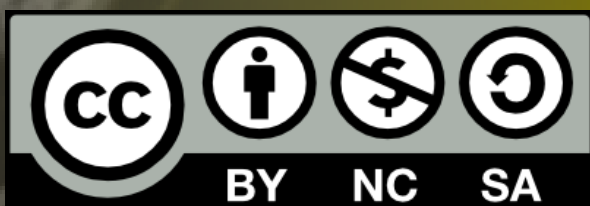
Técnicas para Administración y Marketing Avanzado





AUTORES:

- Correa Peralta Mirella Azucena
- Almeida Salazar Byrone Antonio
- Espinoza Bravo Maria Gabriela
- Cabezas Reyes Andrea Tatiana
- Castillo Villegas Kevin Gabriel



SEGMENTACIÓN EFECTIVA DE CLIENTES UTILIZANDO R: Técnicas para Administración y Marketing Avanzado



PRIMERA EDICIÓN, ENERO 2024

***Segmentación Efectiva de Clientes utilizando R: Técnicas para
Administración y Marketing Avanzado***

ISBN: 978-9942-7173-2-0

DOI: <https://doi.org/10.62131/978-9942-7173-2-0>

Editado por:

Sello editorial: © Editorial Investigativa Latinoamericana (SciELA)
Quevedo, Los Rios, Ecuador

E-mail: admin@editorial-sciela.org

Código Postal: 120303

*Este libro se sometió a arbitraje bajo el sistema de doble ciego (peer review)
y antiplágio*

Dirección editorial: Lic. Alexander Fernando Haro, MSl.

Revisor (1): Ing. Gabriela Lizbeth Merino Lema, Mg.

Revisor (2): Eco. José Julián Coronel Reyes, Mg.

Revisor (3): Ing. Fabian Alberto Gallardo Gonzaga, Mg.

Revisor (4): Ing. Johnny Javier Triviño Sánchez, Mg.

Reservados todos los derechos. Está prohibido, bajo las sanciones penales y el resarcimiento civil previstos en las leyes, reproducir, registrar o transmitir esta publicación, íntegra o parcialmente, por cualquier sistema de recuperación y por cualquier medio, sea mecánico, electrónico, magnético, electroóptico, por fotocopia o por cualquiera otro, sin la autorización previa por escrito a la Editorial Investigativa Latinoamericana (SciELA).

| AUTORES |





Mirella Azucena Correa Peralta



<https://orcid.org/0000-0003-1426-0244>



mcorreap@unemi.edu.ec



Universidad Estatal de Milagro;
Universidad Andina Simón Bolívar -
Bolivia

Cursando doctorado en Administración de Empresas en la Universidad Andina Simón Bolívar sede Bolivia. Dos maestrías, una en Gerencia de Tecnologías y otra en Administración de Empresas. Ingeniera en Sistemas computacionales y Licenciada en Ciencias de la Educación. Experiencia profesional en administración de proyectos informáticos en la Universidad Estatal de Milagro. Además, dedicada a consultorias relacionados con TI, administración, educación. Docente de pregrado y postgrado.



Byrone Antonio Almeida Salazar



<https://orcid.org/0000-0002-4412-1125>



balmeidas@unemi.edu.ec



Universidad Estatal de Milagro

Magíster en Automatización y Control Industrial (ESPOL). Magíster en Administración y Dirección de Empresas (UTEG). Diplomado Superior en Currículo por Competencias (UTA). Graduado de Ingeniero en Electricidad Especialización Potencia (ESPOL). Profesional con experiencia en el área de planificación de proyectos y mantenimiento de redes eléctricas del Sistema Eléctrico Público durante 29 años, lo que le ha permitido ser actualmente el Líder de Mantenimiento en CNEL-EP MILAGRO. Así mismo desempeñar el rol de docente a nivel Pregrado en la Facultad de Ciencias de la Ingeniería (UNEMI) desde el año 2002.



Maria Gabriela Espinoza Bravo



<https://orcid.org/0000-0002-7377-8302>



mespinozab4@unemi.edu.ec



Universidad Estatal de Milagro

Profesional altamente capacitada, Magíster en Sistemas de Información Mención con Inteligencia de Negocios y Analítica de Datos Masivos, otorgado por la prestigiosa Universidad Estatal de Milagro. Complementando su formación, cuenta con una sólida base académica como Ingeniera en Computación e Informática, graduada con distinción de la Universidad Agraria del Ecuador. Con una sólida trayectoria académica y un enfoque especializado en el ámbito de la informática, actualmente desempeña el rol de docente en la Facultad de Ciencias e Ingeniería de la Universidad Estatal de Milagro.



Andrea Tatiana Cabezas Reyes



<https://orcid.org/0009-0006-8209-0629>



acabezasr@unemi.edu.ec



Universidad Estatal de Milagro

Máster en Gestión de los Recursos Humanos en las organizaciones en la Universidad Autónoma de Barcelona, Psicóloga mención Industrial, su experiencia profesional enfocada al área de Talento Humano (OBRASPRIVA S.A. y otras entidades) y trabajos freelance (Ministerio del Interior zona 5 y Distrito de Educación 09D17), adicionalmente como docente a nivel de Pregrado en Psicología, Economía e Investigación (Universidad Agraria del Ecuador, Universidad Estatal de Milagro) desde el año 2015.



Kevin Gabriel Castillo Villegas



<https://orcid.org/0009-0004-1688-5613>



kcastillov@unemi.edu.ec



Universidad Estatal de Milagro

Máster en Sistemas de Información con
mención en Inteligencia de Negocios
y Analítica de Datos e Ingeniero
en Sistemas Computacionales, con
experiencia en docencia, vinculación
e investigación en educación superior.





Bar Chart



1000

500

0

Bar Chart

| ÍNDICE |

INTRODUCCIÓN

tecnología + matemáticas + marketing = efectividad

Contexto de la Segmentación de Clientes.....	14
Importancia para la Administración y el Marketing	15
Objetivos del Libro.....	16

CAPÍTULO I.

Fundamentos de la Segmentación de Clientes

Definición de Segmentación de Clientes	19
Importancia en Estrategias de Marketing	20
Objetivos de la Segmentación de Clientes	21
Mejora de la Personalización.....	21
Caso argumentativo.....	22
Optimización de Recursos.....	24
Caso argumentativo.....	25
Valoración de los conocimientos	26

CAPÍTULO II.

Tipos de Segmentación

Segmentación Demográfica.....	31
Ejemplo - Segmentación Demográfica.....	32
Segmentación Psicográfica.....	33
Ejemplo - Segmentación Psicográfica.....	34
Segmentación Conductual	34
Ejemplo - Segmentación Conductual	36
Segmentación Geográfica	37
Ejemplo - Segmentación Geográfica	38
Valoración de los conocimientos	39

CAPÍTULO III.

Acercamiento a R Y RSTUDIO

¿Qué es R?	45
¿Qué es RStudio?	46
Instalación de R.....	48
Instalación de RStudio.....	49
Entorno de RStudio.....	50
Tipos de Variables en RStudio	52
Valoración de los conocimientos	53

CAPÍTULO IV.

Segmentación No Jerárquica

¿Qué es Segmentación No Jerárquica?.....	57
Ventajas de la Segmentación No Jerárquica.....	57
Desventajas de la Segmentación No Jerárquica.....	59
Método K-means.....	60
Acercamiento a la metodología	61
Valoración de los conocimientos	68

CAPÍTULO V.

Segmentación No Jerárquica aplicada al entorno

Detalles del caso	73
Pasos Propuestos	74
Cargamos la base de datos.....	75
Suavización de la base de datos.....	75
Conocer el k adecuado	76
Elaboración de gráficos clúster	77
Unificación de clúster	80
Descarga con clúster	80
Resultados analizados	81
Estrategias para Cada Clúster	82
Clúster 1: "Clientes Frecuentes"	82
Clúster 2: "Compradores Confortables"	83
Clúster 3: "Seniors Fieles"	84

CAPÍTULO VI.

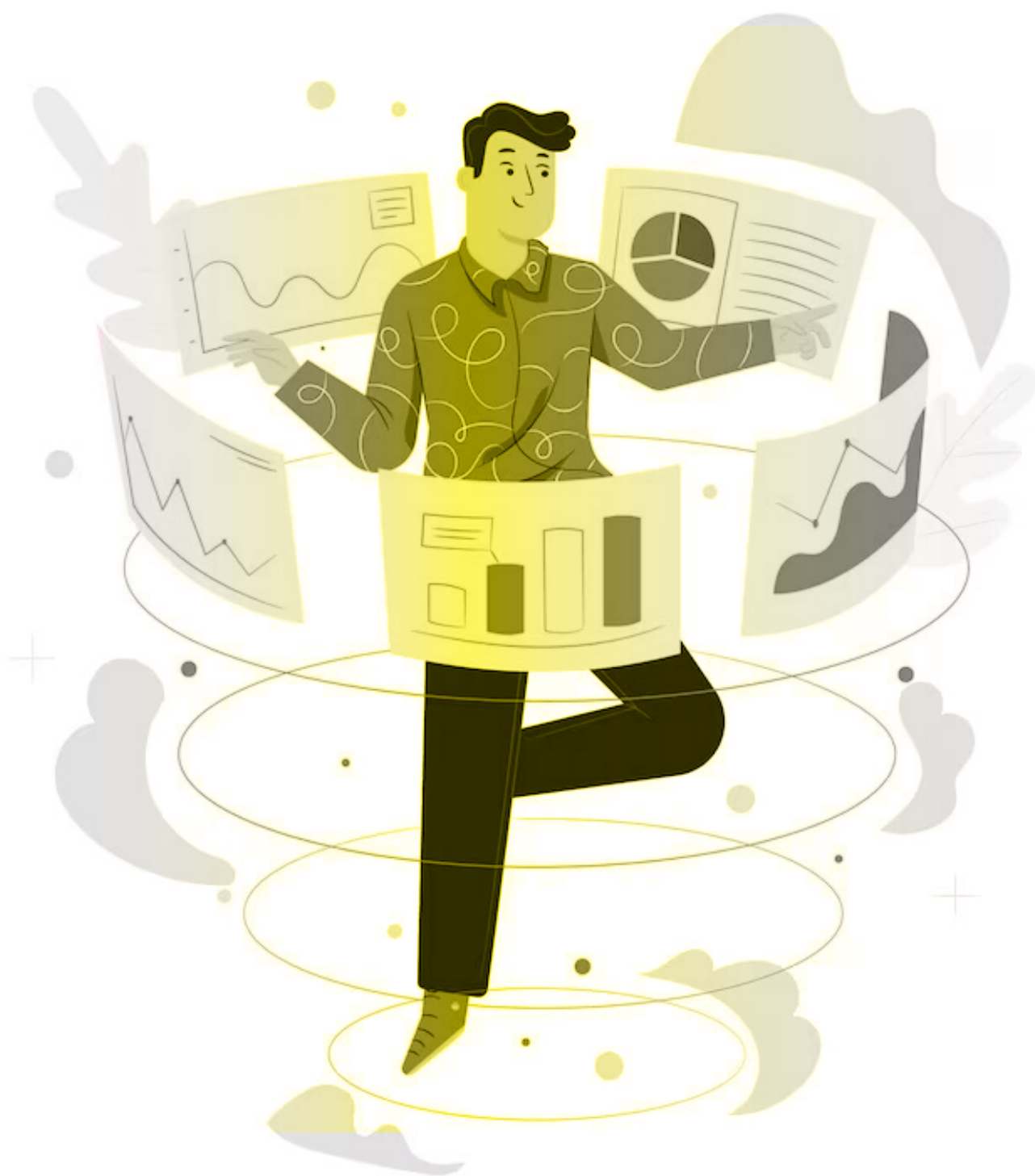
Segmentación Jerárquica

Segmentación Jerárquica	87
Ventajas de la Segmentación Jerárquica	88
Desventajas de la Segmentación Jerárquica	89
Medidas de distancia: Métricas	90
Medidas de distancia: Criterio de enlace	92
Acercamiento a la metodología	93

CAPÍTULO VII.

Segmentación Jerárquica aplicada al entorno

Detalles del caso	99
Pasos propuestos	100
Cargamos la base de datos	101
Suavización de la base de datos	101
Calculamos la distancia Euclidiana	102
Generamos del dendrograma	102
Dendrograma con marcas	103
Integración del Clúster	104
Exportar base con Clúster	104
Resultados analizados	105
Estrategias para Cada Clúster	105
Referencias	109



**ESTA PÁGINA HA SIDO INTENCIONALMENTE DEJADA EN
BLANCO**



INTRODUCCIÓN

TECNOLOGÍA + MATEMÁTICAS + MARKETING = EFECTIVIDAD

CONTEXTO DE LA SEGMENTACIÓN DE CLIENTES

La segmentación de clientes constituye un elemento fundamental en el ámbito del marketing contemporáneo, ofreciendo un enfoque estratégico para comprender y abordar las complejidades inherentes a la diversidad de la base de clientes. Este proceso implica la división sistemática de la clientela en grupos homogéneos con características y comportamientos similares, con el propósito de mejorar la eficacia de las estrategias de marketing y optimizar la satisfacción del cliente¹.

En el argumento de la segmentación de clientes, es esencial considerar la relevancia de criterios demográficos, psicográficos y conductuales para identificar patrones discernibles dentro de la audiencia. Los factores demográficos, como la edad, el género y la ubicación geográfica, ofrecen una base sólida para la diferenciación inicial, mientras que los aspectos psicográficos, como valores, estilos de vida y actitudes, proporcionan una comprensión más profunda de las motivaciones subyacentes de los consumidores.

Además, la segmentación basada en comportamientos se centra en la respuesta del cliente a productos o servicios específicos, considerando factores como la lealtad, la frecuencia de compra y la preferencia de

marca. Este enfoque permite una adaptación más precisa de las estrategias de marketing a las necesidades individuales de los segmentos identificados.

Una vez llevada a cabo la segmentación, se despliegan estrategias de marketing diferenciadas para cada grupo, maximizando así la eficiencia de los recursos y mejorando la personalización de las interacciones con los clientes. Este enfoque estratégico no solo facilita la captación y retención de clientes, sino que también contribuye a la construcción de relaciones a largo plazo, fundamentales en la actual dinámica competitiva del mercado.

IMPORTANCIA PARA LA ADMINISTRACIÓN Y EL MARKETING

La segmentación de clientes se erige como un pilar fundamental para la administración y el marketing contemporáneos, desempeñando un papel crucial en la optimización de recursos y el fortalecimiento de la eficacia estratégica. La comprensión detallada de las distintas características, comportamientos y preferencias que definen a grupos homogéneos dentro de la base de clientes permite a las organizaciones diseñar estrategias más precisas y adaptadas a las necesidades específicas de cada segmento.

Desde una perspectiva administrativa, la segmentación de clientes proporciona a las empresas una herramienta estratégica para asignar eficientemente recursos limitados. La identificación y priorización de segmentos de mercado estratégicamente relevantes permiten a las organizaciones dirigir sus esfuerzos y recursos hacia aquellos sectores que ofrecen el mayor potencial de retorno de inversión. Esta asignación precisa de recursos contribuye a una gestión más eficiente de presupuestos y a la maximización de resultados en términos de captación, retención y fidelización de clientes².

En el ámbito del marketing, la segmentación de clientes desencadena un impacto significativo al permitir una personalización más profunda y significativa de las estrategias. Al comprender las motivaciones, preferencias y comportamientos específicos de cada segmento, las organi-

zaciones pueden adaptar sus mensajes, productos y servicios de manera más efectiva, generando conexiones más sólidas y significativas con sus audiencias. Esta personalización no solo mejora la relevancia de las iniciativas de marketing, sino que también fortalece la lealtad del cliente y contribuye al desarrollo de una ventaja competitiva sostenible³.

Conjuntamente, la segmentación de clientes facilita una comprensión más profunda de la dinámica del mercado y de las tendencias emergentes. Al analizar las variaciones en los comportamientos y necesidades de diferentes segmentos, las organizaciones pueden anticipar cambios en la demanda y ajustar proactivamente sus estrategias para mantenerse a la vanguardia en un entorno empresarial dinámico.

OBJETIVOS DEL LIBRO

El libro Segmentación Efectiva de Clientes utilizando R: Técnicas para Administración y Marketing Avanzado se propone abordar de manera integral los desafíos contemporáneos que enfrentan las organizaciones en la gestión estratégica de su base de clientes. Con un enfoque centrado en el uso de la plataforma R, reconocida por su versatilidad y poder analítico, este texto aspira a lograr los siguientes objetivos:

- **Profundizar en las Bases Conceptuales de la Segmentación de Clientes:** El libro busca proporcionar a los lectores una comprensión sólida de los fundamentos teóricos que sustentan la segmentación de clientes. Explorará los conceptos demográficos, psicográficos y conductuales que constituyen la base de esta práctica, permitiendo a los profesionales y académicos consolidar sus conocimientos en el tema.
- **Integrar Herramientas Analíticas Avanzadas:** A través del uso de la plataforma R, el libro pretende introducir a los lectores en técnicas analíticas avanzadas para la segmentación de clientes. Desde métodos tradicionales hasta enfoques más innovadores, se proporcionarán ejemplos prácticos y casos de estudio que permitirán a los lectores aplicar eficazmente estas herramientas en entornos empresariales y de marketing sofisticados.

- **Aplicar en Contextos Empresariales Reales:** Uno de los objetivos clave es conectar la teoría con la práctica. El libro se esforzará por ilustrar la aplicación de técnicas de segmentación de clientes utilizando R en situaciones empresariales del mundo real. Esto incluirá ejemplos específicos de diversos sectores, destacando la relevancia y la adaptabilidad de las metodologías propuestas.



- **Facilitar la Toma de Decisiones Estratégicas:** Al destacar la importancia de la segmentación de clientes en la toma de decisiones estratégicas, el libro se propone equipar a los lectores con las habilidades necesarias para utilizar la información derivada de la segmentación en el diseño y la implementación de estrategias de marketing efectivas y en la gestión general de la relación con el cliente.



- **Estimular la Innovación en Marketing:** Al explorar técnicas avanzadas y metodologías innovadoras, el libro busca inspirar la creatividad y la innovación en las prácticas de marketing. Se pretende fomentar el desarrollo de enfoques personalizados y vanguardistas para la segmentación de clientes, promoviendo así la diferenciación competitiva y el éxito a largo plazo en el mercado actual.

CAPÍTULO I

FUNDAMENTOS DE LA SEGMENTACIÓN
DE CLIENTES



CAPÍTULO I.

FUNDAMENTOS DE LA SEGMENTACIÓN DE CLIENTES

DEFINICIÓN DE SEGMENTACIÓN DE CLIENTES

La segmentación de clientes es un proceso estratégico en el ámbito del marketing que implica la subdivisión de la base de clientes en grupos más pequeños y homogéneos. Esta práctica busca identificar patrones, características y comportamientos compartidos dentro de estos segmentos, permitiendo a las organizaciones personalizar sus estrategias y adaptarse de manera más precisa a las necesidades y preferencias específicas de cada grupo⁴.

La segmentación se basa en criterios diversos, como características demográficas, psicográficas y conductuales, con el objetivo de entender mejor a la audiencia y mejorar la eficacia de las acciones de marketing. Al comprender las distintas dinámicas presentes en cada segmento, las empresas pueden diseñar estrategias más efectivas para la captación, retención y fidelización de clientes, contribuyendo así al logro de objetivos comerciales más específicos y a la maximización del valor del cliente⁵.

Este proceso de segmentación no solo permite a las organizaciones adaptar sus estrategias a las características únicas de cada grupo de clientes, sino que también facilita la asignación eficiente de recursos

y la personalización de mensajes y ofertas. Al analizar factores como la edad, el comportamiento de compra o los intereses compartidos, las empresas pueden crear experiencias más relevantes y significativas para sus clientes, estableciendo conexiones más sólidas y duraderas⁶.

La segmentación de clientes, por lo tanto, va más allá de una simple categorización; es una herramienta estratégica fundamental para comprender la diversidad de la base de clientes y responder de manera efectiva a sus expectativas en un entorno empresarial dinámico. Este enfoque holístico no solo beneficia la relación con el cliente, sino que también impulsa la rentabilidad y la competitividad de las organizaciones al alinear sus estrategias con las necesidades específicas de cada segmento, creando así un camino sólido hacia el éxito empresarial a largo plazo.

IMPORTANCIA EN ESTRATEGIAS DE MARKETING

La segmentación de clientes desempeña un papel crítico en el diseño de estrategias de marketing, ya que proporciona a las empresas una comprensión más profunda y matizada de su audiencia. Al dividir el mercado en segmentos más pequeños y homogéneos, las organizaciones pueden personalizar sus enfoques y mensajes de manera más efectiva. Esta personalización permite que las campañas de marketing sean más relevantes para cada segmento, lo que, a su vez, aumenta la probabilidad de atraer la atención de los clientes y generar respuestas positivas⁷.

La optimización de recursos es otra ventaja clave de la segmentación de clientes en el ámbito del marketing. Al dirigir estratégicamente los esfuerzos hacia segmentos específicos, las empresas pueden asignar sus presupuestos y recursos de manera más eficiente. Esto evita el desperdicio de recursos en audiencias que pueden no estar interesadas en los productos o servicios ofrecidos, permitiendo un uso más efectivo de los recursos disponibles⁸.

La mejora de la experiencia del cliente es un tercer aspecto crucial. La segmentación no solo se trata de enviar mensajes personalizados; también implica adaptar la experiencia del cliente en todos los puntos

de contacto. Desde el diseño de sitios web hasta la atención al cliente, la segmentación permite ajustar cada interacción para satisfacer las expectativas y preferencias específicas de cada grupo de clientes, contribuyendo a una experiencia general más satisfactoria⁹.

Al mismo tiempo, la segmentación de clientes mejora la efectividad de las campañas de marketing al permitir una alineación más precisa entre la oferta de la empresa y las necesidades del cliente. Al comprender las características distintivas de cada segmento, las empresas pueden adaptar sus mensajes de manera que resuenen con los valores y deseos específicos de esos grupos, resultando en tasas de conversión más altas y una rentabilidad mejorada¹⁰.

La capacidad de adaptar productos y servicios es otro beneficio clave de la segmentación. Al conocer las demandas únicas de diferentes segmentos, las empresas pueden ajustar sus ofertas para satisfacer mejor esas necesidades específicas. Este enfoque no solo mejora la aceptación en el mercado, sino que también posiciona a la empresa como receptiva y centrada en el cliente¹¹.

Finalmente, la segmentación de clientes contribuye significativamente a la fidelización del cliente. La personalización y la atención individualizada resultantes de la segmentación hacen que los clientes se sientan valorados y comprendidos. Este sentimiento de conexión y atención específica aumenta la lealtad del cliente, reduciendo la probabilidad de que busquen alternativas en la competencia¹².

OBJETIVOS DE LA SEGMENTACIÓN DE CLIENTES

MEJORA DE LA PERSONALIZACIÓN

La mejora de la personalización es uno de los objetivos clave de la segmentación de clientes en estrategias de marketing. Este objetivo se centra en la idea fundamental de que los clientes no son homogéneos; tienen necesidades, preferencias y comportamientos distintivos. Al segmentar la base de clientes, las empresas pueden entender y adaptar sus mensajes y ofertas de manera más específica, proporcionando una

experiencia más personalizada para cada grupo¹³.

La segmentación permite a las empresas crear perfiles detallados de diferentes segmentos de clientes, identificando sus características demográficas, comportamientos de compra, preferencias y valores. Con esta información detallada, las empresas pueden diseñar mensajes de marketing que resuenen con cada segmento de manera única. Desde campañas de correo electrónico hasta contenido en redes sociales, la mejora de la personalización garantiza que cada interacción con el cliente sea relevante y significativa.

Asimismo, la mejora de la personalización se extiende más allá de los mensajes; también afecta a la oferta de productos y servicios. Al comprender las preferencias específicas de cada segmento, las empresas pueden adaptar sus productos y servicios para satisfacer las necesidades particulares de esos grupos. Esta adaptación no solo fortalece la conexión emocional con los clientes, sino que también aumenta la probabilidad de satisfacción y lealtad a largo plazo.

En un entorno empresarial cada vez más competitivo, la mejora de la personalización a través de la segmentación se ha vuelto esencial para destacar entre la multitud. Los clientes valoran experiencias personalizadas que se alinean con sus preferencias y estilo de vida, y las empresas que logran proporcionar estas experiencias tienden a ganar la preferencia del cliente. En última instancia, la mejora de la personalización a través de la segmentación no solo impulsa el compromiso del cliente, sino que también contribuye a la construcción de relaciones sólidas y a largo plazo con la audiencia¹⁴.

CASO ARGUMENTATIVO

Imaginemos una empresa de comercio electrónico especializada en tecnología avanzada, desde dispositivos inteligentes hasta soluciones de domótica. Utilizando la segmentación de clientes para mejorar la personalización, la empresa ha identificado tres segmentos clave: Entusiastas de la Tecnología, Profesionales de Negocios Conectados y Familias Digitales.

Para los Entusiastas de la Tecnología, la empresa personaliza la experiencia ofreciendo contenido exclusivo sobre los últimos avances tecnológicos, acceso prioritario a preventas de productos y descuentos específicos para productos de alta gama. Además, se les proporciona un programa de recompensas que incluye beneficios exclusivos, como invitaciones a eventos de lanzamiento y acceso a una línea de atención al cliente VIP especializada.

En cuanto a los Profesionales de Negocios Conectados, la empresa se centra en soluciones tecnológicas diseñadas para aumentar la productividad. Los mensajes personalizados incluyen recomendaciones de dispositivos compatibles con herramientas de productividad empresarial, ofertas especiales para paquetes empresariales y acceso a seminarios web exclusivos sobre la integración efectiva de la tecnología en el entorno empresarial. Además, se les ofrece un servicio de atención al cliente dedicado para abordar sus necesidades empresariales específicas.

Para las Familias Digitales, la personalización se centra en soluciones que mejoren la vida cotidiana. La empresa ofrece recomendaciones de productos adaptados a las necesidades familiares, como dispositivos de seguridad para el hogar, asistentes de voz con funciones familiares y productos que facilitan la gestión de la vida digital de los niños. Además, se envían ofertas exclusivas para paquetes familiares y acceso a contenido educativo sobre el uso seguro y beneficios de la tecnología para los niños.

Esta estrategia avanzada de mejora de la personalización va más allá de simplemente dirigirse a los clientes con sus nombres. Se trata de comprender las necesidades y preferencias únicas de cada segmento, ofreciendo experiencias personalizadas que van desde recomendaciones de productos hasta servicios exclusivos. Al hacerlo, la empresa no solo fomenta la lealtad del cliente, sino que también se destaca en un mercado competitivo al proporcionar un valor diferenciado y adaptado a cada grupo de clientes.



OPTIMIZACIÓN DE RECURSOS

Uno de los objetivos clave de la segmentación de clientes es la optimización de recursos en el ámbito empresarial. Este objetivo se enfoca en dirigir eficientemente los recursos disponibles, como presupuestos de marketing y esfuerzos operativos, hacia los segmentos de clientes más estratégicos y relevantes¹⁵.

La segmentación permite a las empresas identificar y comprender las características distintivas de diferentes grupos de clientes. Al conocer las necesidades, comportamientos y preferencias específicas de cada segmento, las organizaciones pueden adaptar sus estrategias de manera más precisa. Esto resulta en la asignación de recursos de marketing de manera estratégica, evitando gastos innecesarios en audiencias que podrían no generar un retorno significativo¹⁶.

La asignación eficiente de presupuestos es fundamental para maximizar el retorno de inversión (ROI) de las actividades de marketing. Al concentrar los esfuerzos en los segmentos que tienen más probabilidades de responder positivamente a las campañas, las empresas pueden aumentar la efectividad de sus estrategias sin incurrir en costos excesivos. Esta optimización se

traduce en una mejor eficiencia en términos de costos por adquisición de clientes y, en última instancia, contribuye a la rentabilidad general de la empresa.

Además, la optimización de recursos no se limita solo al ámbito financiero. También se refiere a la asignación adecuada de tiempo y esfuerzo del equipo de marketing y ventas. Al comprender los segmentos prioritarios, las empresas pueden enfocar sus recursos humanos en actividades que generen el mayor impacto, como la personalización de mensajes, la atención específica al cliente y el desarrollo de productos adaptados a las necesidades del segmento¹⁷.

CASO ARGUMENTATIVO

Supongamos que la empresa de servicios financieros tiene un presupuesto total de marketing de \$1 millón. Utilizando la segmentación de clientes, la asignación de estos recursos se realiza de la siguiente manera:

Inversionistas Sofisticados: Este segmento representa el 20% de la base de clientes. La empresa destina el 40% del presupuesto de marketing a este grupo, ya que son más propensos a realizar inversiones significativas.

- Cálculo: $\$1 \text{ millón} * 0.20 \text{ (porcentaje de clientes)} * 0.40 \text{ (porcentaje de presupuesto)} = \$80,000$.

Cientes de Banca Personal: Este segmento abarca el 50% de los clientes. Dada su importancia para la empresa, se asigna el 30% del presupuesto para asegurar una atención y servicio personalizados.

- Cálculo: $\$1 \text{ millón} * 0.50 * 0.30 = \$150,000$.

Usuarios de Servicios Básicos: Representan el 30% de la base de clientes. La empresa decide destinar el 30% del presupuesto, ya que la optimización de recursos aquí se centra en la eficiencia y la

digitalización.

- Cálculo: \$1 millón * 0.30 * 0.30 = \$90,000.

Estos cálculos reflejan una asignación estratégica de recursos basada en la segmentación de clientes. La empresa ha optimizado su presupuesto de marketing para adaptarse a las necesidades y comportamientos específicos de cada grupo, maximizando así el impacto de sus esfuerzos en la satisfacción del cliente y la rentabilidad. Este enfoque calculado garantiza que cada segmento reciba la atención y los recursos necesarios para impulsar el éxito general de la estrategia de marketing.

VALORACIÓN DE LOS CONOCIMIENTOS

¿Cuál es el objetivo principal de la segmentación de clientes?

- A. Identificar patrones, características y comportamientos compartidos dentro de los segmentos.
- B. Adaptar las estrategias de marketing a las necesidades y preferencias específicas de cada segmento.
- C. Asignar eficientemente los recursos y personalizar los mensajes y ofertas.

Respuesta: B.

¿Cuáles son los criterios más comunes para la segmentación de clientes?

- A. Características demográficas, psicográficas y conductuales.
- B. Ingresos, nivel educativo y ocupación.
- C. Edad, género y ubicación geográfica.

Respuesta: A.

¿Cuáles son los beneficios de la segmentación de clientes para las empresas?

- A. Mejora la efectividad de las campañas de marketing, optimiza los recursos y mejora la experiencia del cliente.
- B. Aumenta la rentabilidad, posiciona a la empresa como receptiva al cliente y contribuye a la fidelización del cliente.
- C. Todas las anteriores.

Respuesta: C.

¿Cuáles son los objetivos específicos de la segmentación de clientes en estrategias de marketing?

- A. Mejorar la personalización, optimizar recursos y adaptar productos y servicios.
- B. Aumentar la efectividad de las campañas de marketing, mejorar la experiencia del cliente y contribuir a la fidelización del cliente.
- C. Todas las anteriores.

Respuesta: C.

¿Cómo se puede mejorar la personalización a través de la segmentación de clientes?

- A. Creando perfiles detallados de diferentes segmentos de clientes, identificando sus características demográficas, comportamientos de compra, preferencias y valores.
- B. Diseñando mensajes de marketing que resuenen con cada segmento de manera única.
- C. Adaptando los productos y servicios para satisfacer las necesidades particulares de cada grupo.

Respuesta: A.

¿Por qué es importante la mejora de la personalización en un entorno

empresarial cada vez más competitivo?

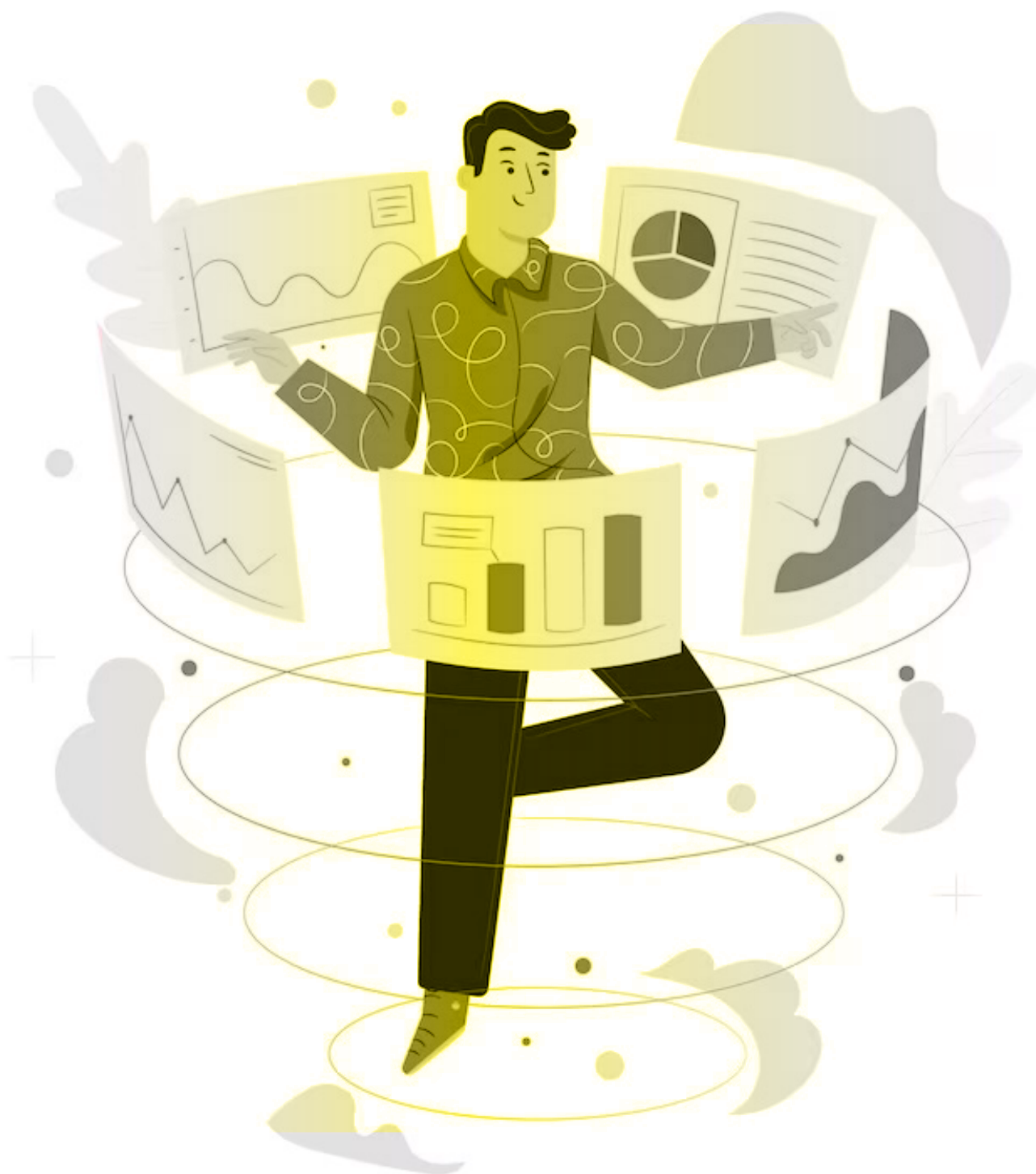
- A. Porque los clientes valoran experiencias personalizadas que se alinean con sus preferencias y estilo de vida.
- B. Porque las empresas que logran proporcionar estas experiencias tienden a ganar la preferencia del cliente.
- C. Todas las anteriores.

Respuesta: C.

¿Cuáles son los beneficios específicos de la mejora de la personalización a través de la segmentación de clientes?

- A. Impulsa el compromiso del cliente y contribuye a la construcción de relaciones sólidas y a largo plazo con la audiencia.
- B. Ayuda a las empresas a destacar entre la multitud y ganar la preferencia del cliente.
- C. Todas las anteriores.

Respuesta: C.



**ESTA PÁGINA HA SIDO INTENCIONALMENTE DEJADA EN
BLANCO**

CAPÍTULO II

TIPOS DE SEGMENTACIÓN



CAPÍTULO II.

TIPOS DE SEGMENTACIÓN

SEGMENTACIÓN DEMOGRÁFICA

La segmentación demográfica es una estrategia de marketing que implica dividir el mercado en grupos basados en características demográficas específicas de la población. Estas características pueden incluir variables como la edad, el género, el nivel educativo, el ingreso, la ocupación, la ubicación geográfica y el estado civil. La segmentación demográfica se utiliza para comprender mejor las necesidades y comportamientos de grupos específicos de consumidores, permitiendo a las empresas adaptar sus estrategias de marketing de manera más efectiva¹⁸. Se especifican algunos aspectos clave de la segmentación demográfica¹⁹:

- **Edad:** Dividir la audiencia según la edad puede ayudar a adaptar productos, mensajes y canales de marketing. Por ejemplo, las estrategias para jóvenes adultos pueden diferir significativamente de las diseñadas para personas mayores.
- **Género:** La segmentación por género se centra en las diferencias entre hombres y mujeres. Algunos productos y mensajes pueden ser más efectivos cuando se personalizan para satisfacer las preferencias y necesidades específicas de cada género.

- **Ingresos y Nivel Educativo:** La segmentación según el ingreso y el nivel educativo puede ser crucial para adaptar estrategias de precios y mensajes. Los consumidores con diferentes niveles de ingresos o niveles educativos pueden tener prioridades y expectativas distintas.
- **Ocupación:** Las diferencias en las ocupaciones de las personas pueden influir en sus decisiones de compra y comportamientos. La segmentación demográfica por ocupación permite a las empresas adaptar sus productos y mensajes para satisfacer las necesidades específicas de diferentes profesiones o industrias.
- **Estado Civil y Tamaño de la Familia:** Las necesidades y preferencias de personas solteras pueden diferir de las de familias con hijos. La segmentación demográfica basada en el estado civil y el tamaño de la familia permite a las empresas adaptar sus productos y mensajes para abordar estas diferencias.

EJEMPLO - SEGMENTACIÓN DEMOGRÁFICA

Considere el ejemplo de una empresa de seguros que ofrece seguros de vida, seguros de salud y seguros de automóviles. La empresa podría segmentar a sus clientes por edad, sexo, ingresos y estado civil para crear las siguientes estrategias de marketing y ventas:

- **Jóvenes adultos:** La empresa podría centrarse en crear conciencia sobre la importancia de los seguros de vida y de salud entre los jóvenes adultos. La empresa podría ofrecer descuentos especiales para jóvenes adultos que contraten un seguro de vida o de salud.
- **Adultos jóvenes:** La empresa podría centrarse en crear productos y servicios que satisfagan las necesidades específicas de los adultos jóvenes, como seguros de vida para parejas jóvenes o seguros de salud con cobertura para personas solteras que viven solas.

- **Adultos maduros:** La empresa podría centrarse en crear productos y servicios que ayuden a los adultos maduros a prepararse para la jubilación, como seguros de vida con cobertura para enfermedades graves o seguros de salud con cobertura para personas mayores.
- **Adultos mayores:** La empresa podría centrarse en crear productos y servicios que ayuden a los adultos mayores a mantenerse saludables y seguros, como seguros de vida con cobertura para cuidados prolongados o seguros de salud con cobertura para personas con enfermedades crónicas.

SEGMENTACIÓN PSICOGRÁFICA

La segmentación psicográfica es una estrategia de marketing que se centra en dividir a los consumidores en grupos según sus características psicológicas, actitudes, valores, intereses y estilos de vida. A diferencia de la segmentación demográfica, que se basa en datos como la edad o el ingreso, la segmentación psicográfica busca comprender las motivaciones más profundas que influyen en el comportamiento del consumidor^{20,21}.

- **Valores y Actitudes:** Esta dimensión se enfoca en los principios fundamentales que guían las decisiones y acciones de los individuos. Comprender los valores y actitudes permite a las empresas alinear sus mensajes y productos con las creencias centrales de su audiencia.
- **Estilos de Vida:** La segmentación psicográfica examina los patrones de vida, actividades y comportamientos cotidianos. Esto incluye aspectos como la rutina diaria, las preferencias de ocio y las interacciones sociales, proporcionando una visión más completa del consumidor.
- **Intereses y Hobbies:** Se refiere a las actividades que atraen y apasionan a los individuos. Conocer los intereses y hobbies permite a las

empresas desarrollar productos y campañas que se alineen con las áreas que más interesan a su audiencia.

- **Personalidad del Consumidor:** La segmentación psicográfica también tiene en cuenta las diferencias en la personalidad de los consumidores. Algunas personas pueden ser extrovertidas y aventureras, mientras que otras pueden ser más introvertidas y centradas en la estabilidad.

EJEMPLO - SEGMENTACIÓN PSICOGRÁFICA

Considere el ejemplo de una empresa de ropa que ofrece ropa casual y formal para hombres y mujeres. La empresa podría segmentar a sus clientes por estilo de vida y personalidad para crear las siguientes estrategias de marketing y ventas:

- **Clientes activos:** La empresa podría centrarse en crear ropa que sea cómoda y funcional para actividades al aire libre. La empresa podría asociarse con atletas o celebridades que representen un estilo de vida activo.
- **Clientes contemplativos:** La empresa podría centrarse en crear ropa que sea elegante y sofisticada. La empresa podría asociarse con artistas o diseñadores que representen un estilo de vida contemplativo.
- **Clientes trabajadores:** La empresa podría centrarse en crear ropa que sea profesional y presentable. La empresa podría asociarse con empresas o organizaciones que representen un estilo de vida trabajador.

SEGMENTACIÓN CONDUCTUAL

La segmentación conductual es una estrategia de marketing que se basa en el comportamiento de los consumidores en relación con un producto o servicio. Este enfoque se centra en cómo los clientes interactúan

con un producto, sus patrones de uso, lealtad a la marca y decisiones de compra²². Se pormenorizan aspectos clave:

- **Patrones de Uso:** Se clasifica a los clientes según la frecuencia y la cantidad de uso de un producto. Por ejemplo, en la industria de telecomunicaciones, algunos clientes pueden usar intensivamente los servicios de datos móviles, mientras que otros pueden usarlos de manera más ocasional. Esto permite adaptar estrategias de marketing y ofrecer paquetes o promociones específicos para cada segmento.
- **Lealtad a la Marca:** La lealtad del cliente es un aspecto crucial. Algunos consumidores pueden ser leales a una marca específica, mientras que otros pueden probar diferentes opciones. La segmentación conductual permite identificar y retener a los clientes leales mediante programas de fidelización, ofertas exclusivas o beneficios especiales.
- **Ciclo de Vida del Producto:** Se analiza en qué etapa del ciclo de vida del producto se encuentra el consumidor. Por ejemplo, en la industria de la tecnología, algunos clientes adoptan rápidamente nuevos productos, mientras que otros prefieren esperar hasta que se vuelvan más establecidos. La estrategia de marketing se adapta según estas preferencias.
- **Comportamiento de Compra:** Se considera cómo los consumidores toman decisiones de compra. Algunos pueden ser compradores impulsivos, mientras que otros pueden investigar y comparar antes de tomar una decisión. Las estrategias de marketing pueden ajustarse para satisfacer las necesidades de ambos grupos.
- **Respuestas a la Publicidad y Promociones:** Los consumidores pueden reaccionar de manera diferente a la publicidad y promociones. Algunos pueden ser más receptivos a ofertas especiales, descuentos o eventos promocionales. La segmentación conductual permite dirigir campañas publicitarias específicas para maximizar el impacto.
- **Nivel de Conciencia de la Marca:** Se evalúa el grado en que los consu-

midores están familiarizados con la marca y sus productos. La estrategia de marketing puede centrarse en construir conciencia de marca para aquellos que no están familiarizados o en reforzar la relación con los clientes que ya conocen la marca.

EJEMPLO - SEGMENTACIÓN CONDUCTUAL

Considere el ejemplo de una empresa de telecomunicaciones que ofrece servicios de telefonía, internet y televisión. La empresa podría segmentar a sus clientes por uso del producto, comportamiento de compra y valor del cliente para crear las siguientes estrategias de marketing y ventas:

- **Usuarios activos:** La empresa podría centrarse en ofrecer descuentos y promociones especiales a los usuarios activos para mantener su lealtad.
- **Usuarios ocasionales:** La empresa podría centrarse en crear campañas de marketing que incentiven a los usuarios ocasionales a utilizar sus servicios con más frecuencia.
- **Usuarios inactivos:** La empresa podría centrarse en recuperar a los usuarios inactivos ofreciendo descuentos y promociones especiales para que vuelvan a utilizar sus servicios.
- **Clientes de alta lealtad:** La empresa podría centrarse en ofrecer servicios y beneficios exclusivos a los clientes de alta lealtad para mantener su lealtad.
- **Clientes de media lealtad:** La empresa podría centrarse en ofrecer servicios y beneficios personalizados a los clientes de media lealtad para aumentar su lealtad.
- **Clientes de bajo valor:** La empresa podría centrarse en identificar las causas del bajo valor de estos clientes y tomar medidas para mejorar su valor.

SEGMENTACIÓN GEOGRÁFICA

La segmentación geográfica es una estrategia de marketing que divide a un mercado en segmentos basados en la ubicación geográfica de los consumidores. Esta forma de segmentación reconoce que las necesidades, comportamientos y preferencias de los consumidores pueden variar según la región, el país, el clima o la densidad poblacional^{23,24}. Se especifica:

- **Regiones Geográficas:** La segmentación puede basarse en áreas geográficas más amplias, como continentes, países, estados o provincias. Cada región puede tener características distintivas que afectan las preferencias del consumidor y la forma en que se abordan los mensajes de marketing.
- **Densidad Poblacional:** La densidad de la población, ya sea urbana, suburbana o rural, puede influir en las necesidades de los consumidores. Por ejemplo, productos y servicios diseñados para áreas urbanas densamente pobladas pueden diferir de aquellos dirigidos a comunidades rurales más dispersas.
- **Clima y Geografía Física:** Las condiciones climáticas y la geografía física pueden afectar las preferencias de productos y la demanda del mercado. Artículos como ropa, productos para el hogar y alimentos pueden variar significativamente según el clima de una región.
- **Tamaño de la Ciudad o Localidad:** El tamaño de la ciudad o localidad puede influir en el comportamiento del consumidor. Por ejemplo, las estrategias de marketing para áreas metropolitanas pueden enfocarse en la conveniencia y la eficiencia, mientras que en localidades más pequeñas se pueden destacar aspectos comunitarios y de cercanía.
- **Fronteras Políticas:** Las fronteras políticas, como las divisiones entre países o estados, pueden afectar las regulaciones, los impuestos y otros aspectos comerciales. La segmentación geográfica puede ayudar a adaptar estrategias para cumplir con los requisitos específicos de cada área.

EJEMPLO - SEGMENTACIÓN GEOGRÁFICA

Considere el ejemplo de una empresa de alimentos que ofrece una variedad de productos, incluidos refrescos, bocadillos y alimentos congelados. La empresa podría segmentar a sus clientes por ubicación y clima para crear las siguientes estrategias de marketing y ventas:

- **Clientes de áreas urbanas:** La empresa podría centrarse en crear campañas de marketing que se dirijan a los clientes de áreas urbanas. Por ejemplo, la empresa podría crear anuncios que se transmitan en televisión o radio en áreas urbanas.
- **Clientes de áreas suburbanas:** La empresa podría centrarse en crear campañas de marketing que se dirijan a los clientes de áreas suburbanas. Por ejemplo, la empresa podría crear anuncios que se publiquen en periódicos o revistas locales en áreas suburbanas.
- **Clientes de áreas rurales:** La empresa podría centrarse en crear campañas de marketing que se dirijan a los clientes de áreas rurales. Por ejemplo, la empresa podría crear anuncios que se distribuyan a través de correo directo en áreas rurales.
- **Clientes de clima cálido:** La empresa podría centrarse en crear productos y servicios que satisfagan las necesidades de los clientes de clima cálido. Por ejemplo, la empresa podría crear una línea de refrescos sin calorías que sean adecuados para el clima cálido.
- **Clientes de clima templado:** La empresa podría centrarse en crear productos y servicios que satisfagan las necesidades de los clientes de clima templado. Por ejemplo, la empresa podría crear una línea de bocadillos saludables que sean adecuados para el clima templado.
- **Clientes de clima frío:** La empresa podría centrarse en crear pro-

- ductos y servicios que satisfagan las necesidades de los clientes de clima frío. Por ejemplo, la empresa podría crear una línea de alimentos congelados que sean adecuados para el clima frío.

VALORACIÓN DE LOS CONOCIMIENTOS

¿Cuál es el objetivo principal de la segmentación demográfica?

- A. Comprender mejor las necesidades y comportamientos de los consumidores.
- B. Adaptar productos, mensajes y canales de marketing a grupos específicos de consumidores.
- C. Maximizar el retorno de la inversión (ROI) de las estrategias de marketing.

Respuesta: B.

¿Cuáles son las características demográficas más comunes utilizadas para la segmentación demográfica?

- A. Edad, género, ingreso y ocupación.
- B. Edad, género, nivel de sentimentalismo y estado civil.
- C. Edad, género, ocupación y lugar favorito.

Respuesta: A.

¿Cómo puede la segmentación demográfica ayudar a las empresas a adaptar sus productos y mensajes?

- A. Al identificar las necesidades y preferencias específicas de cada segmento.
- B. Al utilizar diferentes canales de marketing para llegar a diferentes segmentos.
- C. Al ofrecer descuentos y promociones personalizados.

Respuesta: A.

¿Cómo puede la segmentación demográfica ayudar a las empresas a maximizar el retorno de la inversión (ROI) de sus estrategias de marketing?

- A. Al evitar gastos innecesarios en audiencias que no son relevantes.
- B. Al personalizar los mensajes y la atención al cliente para satisfacer las necesidades específicas de cada segmento.
- C. Al desarrollar productos y servicios adaptados a las necesidades del segmento.

Respuesta: A.

¿Cuál es el objetivo principal de la segmentación psicográfica?

- A. Comprender mejor las motivaciones que influyen en el comportamiento del consumidor.
- B. Adaptar productos, mensajes y canales de marketing a grupos específicos de consumidores.
- C. Maximizar el retorno de la inversión (ROI) de las estrategias de marketing.

Respuesta: A.

¿Cuáles son las dimensiones más comunes utilizadas para la segmentación psicográfica?

- A. Valores y actitudes, estilos de vida, intereses y hobbies.
- B. Edad, género, nivel educativo y ocupación.
- C. Ubicación geográfica, nivel de ingresos y estado civil.

Respuesta: A.

¿Cómo puede la segmentación psicográfica ayudar a las empresas a comprender mejor las motivaciones que influyen en el comportamiento

del consumidor?

- A. Al identificar las necesidades y preferencias específicas de cada segmento.
- B. Al utilizar diferentes canales de marketing para llegar a diferentes segmentos.
- C. Al ofrecer descuentos y promociones personalizados.

Respuesta: A.

¿Cuál es la diferencia entre la segmentación demográfica y la segmentación psicográfica?

- A. La segmentación demográfica se basa en datos observables, como la edad o el ingreso, mientras que la segmentación psicográfica se basa en datos psicológicos, como los valores o la personalidad.
- B. La segmentación demográfica se centra en dividir a los consumidores en grupos homogéneos, mientras que la segmentación psicográfica se centra en identificar las diferencias entre los consumidores.
- C. La segmentación demográfica es más eficaz para comprender las necesidades y motivaciones de los consumidores, mientras que la segmentación psicográfica es más eficaz para adaptar productos y mensajes.

Respuesta: A.

¿Cuál es la dimensión de la segmentación psicográfica que se enfoca en los principios fundamentales que guían las decisiones y acciones de los individuos?

- A. Estilos de vida
- B. Valores y actitudes
- C. Intereses y hobbies

Respuesta: B.

¿Cuál es la dimensión de la segmentación psicográfica que se refiere a las actividades que atraen y apasionan a los individuos?

- A. Valores y actitudes
- B. Estilos de vida
- C. Intereses y hobbies

Respuesta: C.

¿Cuál es la dimensión de la segmentación psicográfica que también tiene en cuenta las diferencias en la personalidad de los consumidores?

- A. Valores y actitudes
- B. Estilos de vida
- C. Personalidad del consumidor

Respuesta: C.

¿Cuál es el objetivo principal de la segmentación conductual?

- A. Comprender mejor las motivaciones que influyen en el comportamiento del consumidor.
- B. Adaptar productos, mensajes y canales de marketing a grupos específicos de consumidores.
- C. Maximizar el retorno de la inversión (ROI) de las estrategias de marketing.

Respuesta: B.

¿Cuáles son las dimensiones más comunes utilizadas para la segmentación conductual?

- A. Patrones de uso, lealtad a la marca, ciclo de vida del producto y comportamiento de compra.
- B. Ubicación geográfica, nivel de ingresos y estado civil.
- C. Edad, género y nivel educativo.

Respuesta: A.

¿Cómo puede la segmentación conductual ayudar a las empresas a comprender mejor las motivaciones que influyen en el comportamiento del consumidor?

- A. Al analizar los patrones de uso, la lealtad a la marca y el ciclo de vida del producto.
- B. Al analizar los valores y actitudes, los estilos de vida e intereses y hobbies.
- C. Al analizar la edad, el género y el nivel educativo.

Respuesta: A.

¿Cuál es la dimensión de la segmentación conductual que se enfoca en cómo los clientes interactúan con un producto?

- A. Ciclo de vida del producto
- B. Lealtad a la marca
- C. Patrones de uso

Respuesta: C.

¿Cuáles son las dimensiones más comunes utilizadas para la segmentación geográfica?

- A. Regiones geográficas, densidad poblacional, clima y geografía física, tamaño de la ciudad o localidad y fronteras políticas.
- B. Ubicación geográfica, nivel de ingresos y estado civil.
- C. Edad, género y nivel educativo.

Respuesta: A.

CAPÍTULO III

ACERCAMIENTO A R Y RSTUDIO



CAPÍTULO III.

ACERCAMIENTO A R Y RSTUDIO

¿QUÉ ES R?



R es un lenguaje de programación y un entorno de software diseñado específicamente para estadísticas y análisis de datos. Fue creado por Ross Ihaka y Robert Gentleman en la Universidad de Auckland, Nueva Zelanda, y su desarrollo continuo está a cargo de la comunidad de usuarios de R en todo el mundo^{25,26}.

Las características principales de R incluyen:

- **Estadísticas y Análisis de Datos:** R proporciona una amplia gama de funciones estadísticas y métodos de análisis de datos. Es utilizado por científicos de datos, estadísticos, analistas y profesionales en campos relacionados para realizar análisis exploratorios de datos, modelado estadístico, visualización de datos y más.
- **Entorno de Programación:** R es un lenguaje de programación interpretado y orientado a objetos. Permite a los usuarios escribir scripts y funciones para realizar tareas específicas de análisis de datos.

- **Gráficos y Visualización:** R ofrece capacidades avanzadas para la creación de gráficos y visualización de datos. Esto incluye la capacidad de generar una variedad de gráficos estadísticos, como histogramas, gráficos de dispersión, diagramas de caja, entre otros.
- **Paquetes y Extensiones:** Una de las fortalezas clave de R es su capacidad para extender sus funcionalidades mediante paquetes. Hay miles de paquetes disponibles que proporcionan funciones adicionales y herramientas especializadas para tareas específicas.
- **Comunidad Activa:** R tiene una comunidad activa de usuarios y desarrolladores que contribuyen con paquetes, tutoriales y soluciones a problemas. Esto hace que sea un entorno dinámico y en constante evolución.
- **Open Source:** R es un software de código abierto, lo que significa que es gratuito y su código fuente está disponible para que los usuarios lo modifiquen y lo distribuyan.

¿QUÉ ES RSTUDIO?

RStudio es un entorno de desarrollo integrado (IDE) diseñado específicamente para trabajar con el lenguaje de programación R. Es una interfaz gráfica que facilita la escritura, ejecución y depuración de código R. RStudio ofrece una serie de características y herramientas que hacen que el proceso de análisis de datos y desarrollo en R sea más eficiente y efectivo^{27,28}. Se detallan elementos clave:



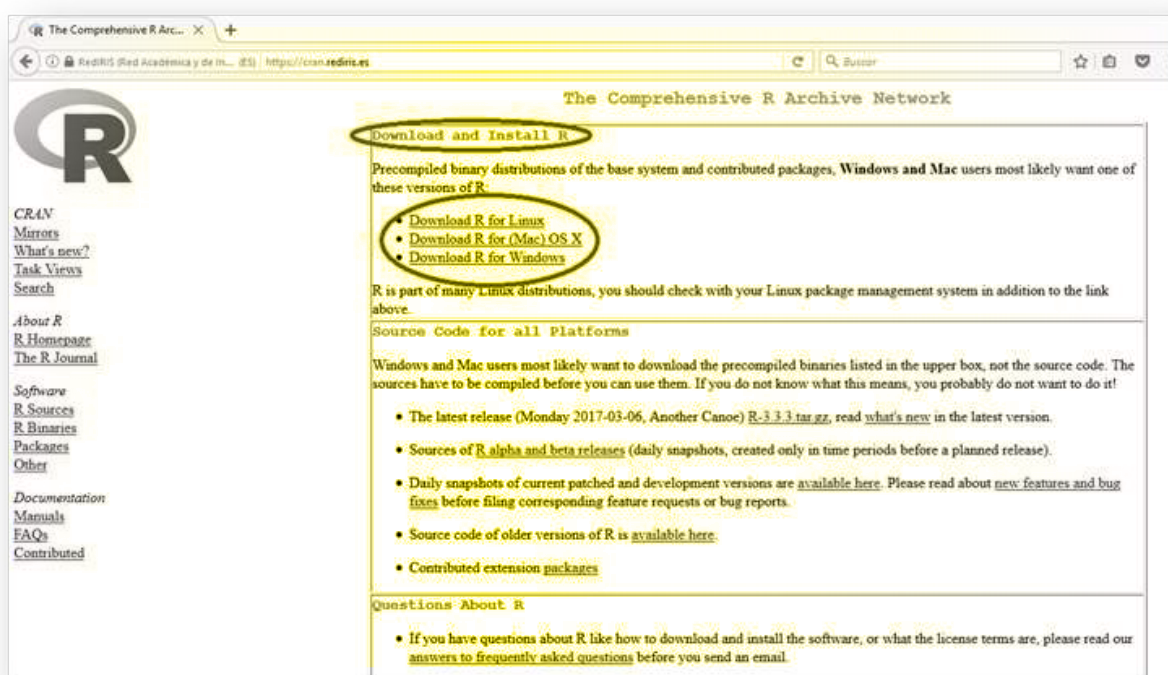
- **Consola de R:** RStudio incluye una consola interactiva de R, lo que permite a los usuarios ejecutar comandos R de manera directa y ver los resultados de inmediato. Esto facilita la experimentación y la exploración de datos.
- **Editor de Script:** Proporciona un editor de script integrado que resalta la sintaxis y facilita la escritura de código R. El editor incluye funciones como la autocompletación de código y la revisión de código.
- **Ambiente de Trabajo Integrado:** RStudio organiza el espacio de trabajo en paneles, que incluyen el script en el que estás trabajando, la consola, el historial de comandos, los objetos en el entorno y la salida de gráficos.
- **Herramientas de Visualización:** Facilita la creación y visualización de gráficos de manera interactiva. Puedes ver los gráficos directamente en la ventana de gráficos de RStudio y ajustarlos según tus necesidades.
- **Historial y Control de Versiones:** RStudio ofrece un historial de comandos ejecutados, lo que facilita la revisión de sesiones anteriores. También es compatible con sistemas de control de versiones como Git.
- **Integración con R Markdown:** RStudio permite la creación de documentos dinámicos utilizando R Markdown. Estos documentos pueden incluir código R, resultados y texto formateado en un solo archivo, lo que es útil para la generación de informes reproducibles.
- **Soporte para Paquetes:** RStudio facilita la instalación y gestión de paquetes de R. También proporciona herramientas para explorar y documentar paquetes.
- **Desarrollo Web:** RStudio incluye características para el desarrollo web con Shiny, un marco de desarrollo para construir aplicaciones web interactivas en R.

INSTALACIÓN DE R

Para proceder con la instalación nos dirigimos a la página web de R project: <http://www.r-project.org>.

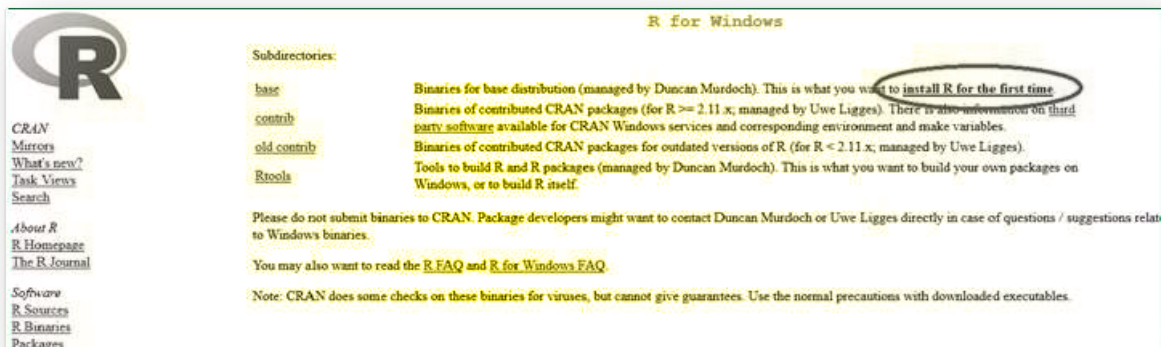


Para descargar damos clic en CRAN y pinchamos sobre el enlace más próximo a nuestra ubicación. Seleccionemos la URL (<https://cloud.r-project.org>).



Ahora, en función de nuestro caso, seleccionamos el sistema operativo correspondiente.

Supongamos que tenemos Windows al hacer clic sobre Download R for Windows iremos a la página que se reproduce más abajo. Hacer clic sobre install R for the first time.

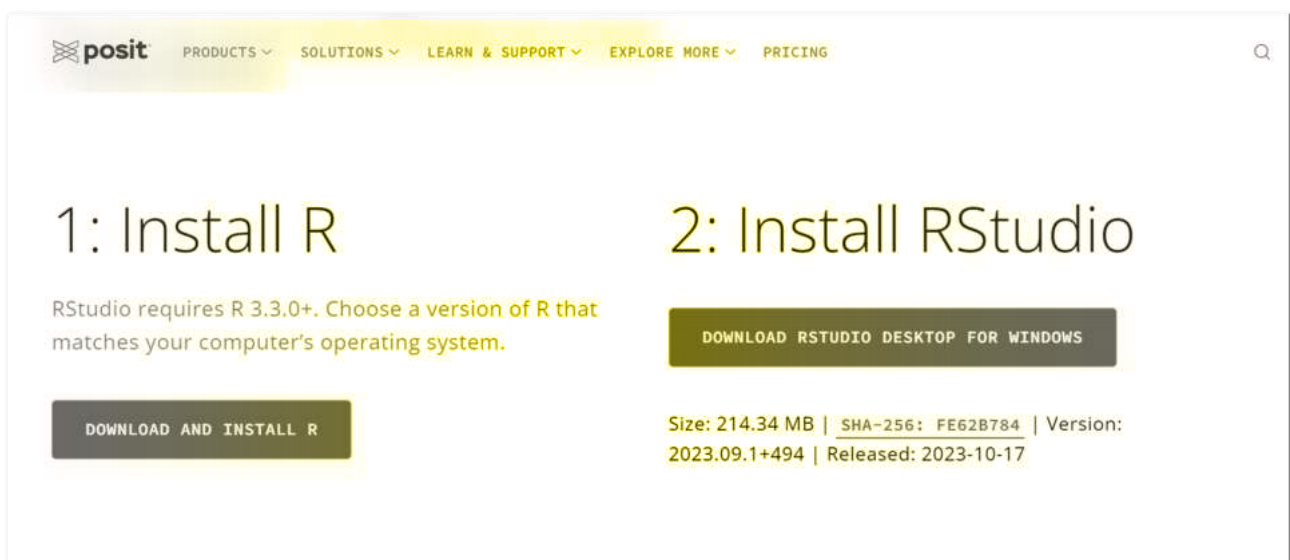


En la siguiente ventana, hacer clic sobre [Download R-4.3.2 for Windows](#) y guardar el archivo de instalación.

Ejecutar el archivo descargado para proceder a la instalación de R.

INSTALACIÓN DE RSTUDIO

Abre de nuevo tu navegador y nos direccionamos al sitio oficial de RStudio (<https://posit.co/download/rstudio-desktop/>)

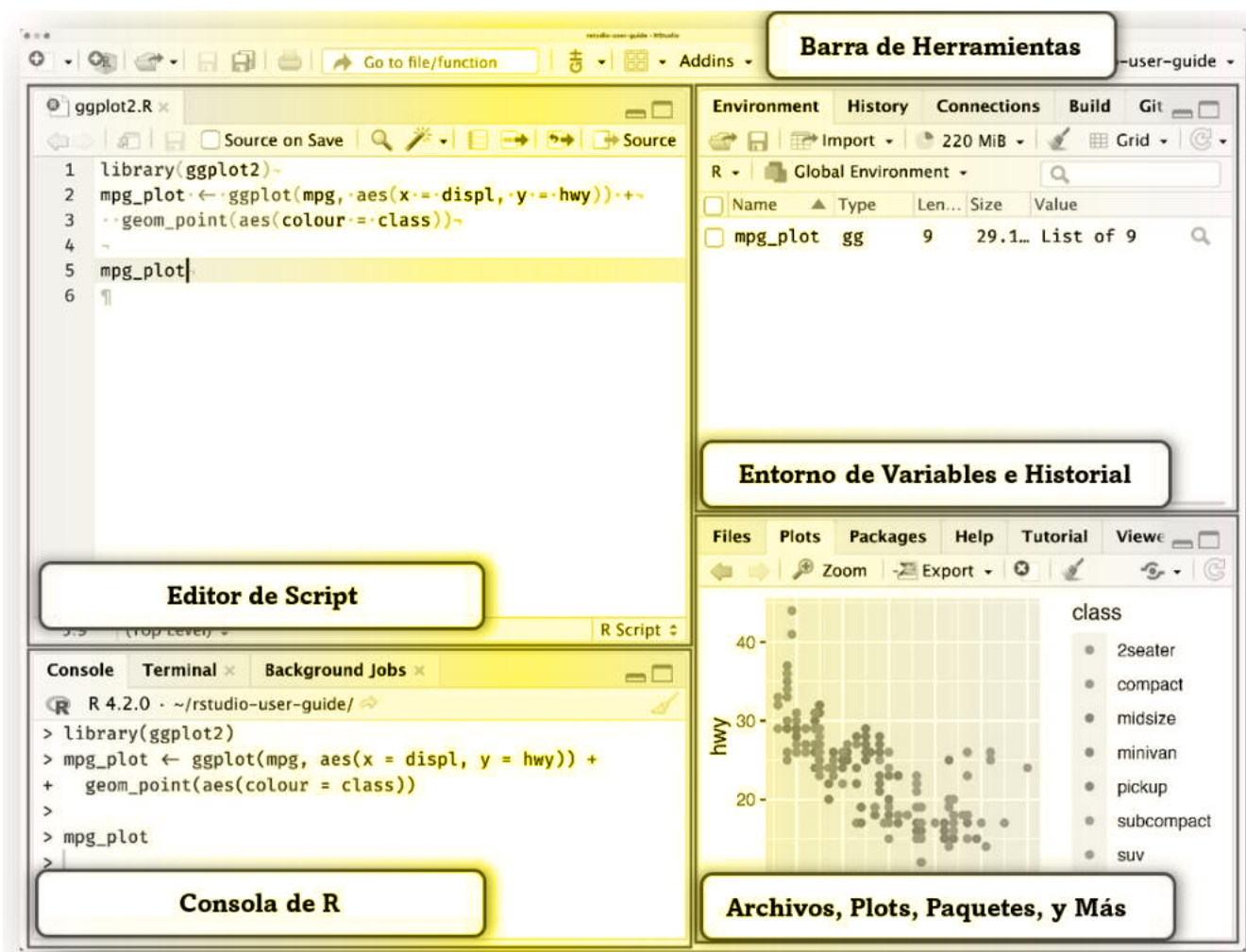


Procedemos a guardar



Una vez descargado ejecutamos y seguimos todos los pasos solicitados

ENTORNO DE RSTUDIO



- **Editor de Script:** Este panel es donde puedes escribir y editar tus scripts de R. Soporta resaltado de sintaxis, autocompletado y otras características para facilitar la escritura del código.
- **Consola de R:** La consola de R en la parte izquierda inferior de la interfaz es donde puedes ingresar comandos R directamente y ver la salida correspondiente. También puedes ejecutar líneas o bloques de código desde el script directamente en la consola.
- **Entorno e Historial:** En la parte superior derecha, hay pestañas para ver el entorno actual (objetos y variables que has definido) y el historial de comandos ejecutados. Puedes explorar y gestionar objetos desde el entorno.
- **Pestañas de Gráficos y Ayuda:** RStudio tiene pestañas específicas para ver gráficos generados y obtener ayuda. Cuando produces gráficos, aparecerán en la pestaña de Plots. También puedes acceder a la documentación y ayuda de funciones específicas en la pestaña de Help.
- **Archivos, Plots, Paquetes, y Más:** En la pestaña de Files, puedes explorar y gestionar tus archivos. La pestaña de Plots muestra gráficos generados, y la pestaña de Packages te permite instalar, cargar y gestionar paquetes de R.
- **Barra de Herramientas:** La barra superior contiene herramientas y opciones adicionales. Puedes encontrar botones para ejecutar, detener y depurar código, así como acceder a otras funciones y configuraciones.
- **Visor de Ayuda:** Cuando buscas ayuda para una función específica, el visor de ayuda aparece en la esquina derecha inferior. Proporciona información sobre la función seleccionada, sus argumentos y ejemplos de uso.
- **Terminal y Git (opcional):** RStudio también incluye una terminal y herramientas de integración con Git para el control de versiones. Estas

funciones son útiles si estás trabajando en proyectos colaborativos o si utilizas Git para el seguimiento de cambios en tu código.

- **Marcadores de Punto y Navegación:** Puedes colocar marcadores de punto en tu script para facilitar la navegación y referencia. Además, hay atajos de teclado útiles para la navegación rápida por el código.

TIPOS DE VARIABLES EN RSTUDIO

- **Numéricas (Numeric):** Representan números reales y se utilizan para operaciones matemáticas.
- **Caracteres (Character):** Almacenan cadenas de texto.
- **Lógicas (Logical):** Representan valores lógicos o booleanos, es decir, verdadero (TRUE) o falso (FALSE).
- **Factores (Factor):** Se utilizan para representar variables categóricas o de tipo factor.
- **Enteras (Integer):** Representan números enteros.
- **Fecha y Hora (Date y POSIXct):** Se utilizan para almacenar información de fechas y horas.
- **Listas (List):** Pueden contener elementos de diferentes tipos, incluyendo otras listas.
- **Matrices (Matrix):** Son arreglos bidimensionales que contienen elementos del mismo tipo.
- **Data Frames (Data Frame):** Estructuras tabulares que pueden contener diferentes tipos de variables.



VALORACIÓN DE LOS CONOCIMIENTOS

¿Cuál es el objetivo principal de R?

- A. Proporcionar una amplia gama de funciones estadísticas y métodos de análisis de datos.
- B. Ser un lenguaje de programación interpretado y orientado a objetos.
- C. Ofrecer capacidades avanzadas para la creación de gráficos y visualización de datos.

Respuesta: A.

¿Qué es RStudio?

- A. Un lenguaje de programación interpretado y orientado a objetos.
- B. Un entorno de desarrollo integrado (IDE) diseñado específicamente para trabajar con el lenguaje de programación R.
- C. Una biblioteca de funciones estadísticas y métodos de análisis de datos.

Respuesta: B.

¿Cuáles son los elementos clave de RStudio?

- A. Consola de R, editor de script, entorno de trabajo integrado, herramientas de visualización, historial y control de versiones, integración con R Markdown, soporte para paquetes, desarrollo web.
- B. Consola de R, editor de script, entorno de variables e historial, barra de herramientas.
- C. Consola de R, editor de script, herramientas de visualización, historial y control de versiones.

Respuesta: A.

¿Cómo se instala R?

- A. Descargando el archivo de instalación desde el sitio web del proyecto R y ejecutándolo.
- B. Descargando el archivo de instalación desde el sitio web de RStudio y ejecutándolo.
- C. Ingresando el comando `install.packages("R")` en la consola de R.

Respuesta: A.

¿Cómo se instala RStudio?

- A. Descargando el archivo de instalación desde el sitio web del proyecto R y ejecutándolo.
- B. Descargando el archivo de instalación desde el sitio web de RStudio y ejecutándolo.
- C. Ingresando el comando `install.packages("RStudio")` en la consola de R.

Respuesta: B.

¿Cuáles son los tipos de variables en RStudio?

- A. Numéricas, caracteres, lógicas, factores, enteras, formulas y hora.
- B. Numéricas, caracteres, lógicas, factores, enteras, fecha, hora, listas, matrices y data frames.
- C. Numéricas, caracteres, lógicas, factores, enteras, caracteres dobles y hora, listas y matrices.

Respuesta: B.

¿Qué tipo de variable se utiliza para representar números reales?

- A. Numérica
- B. Caracter
- C. Lógica

Respuesta: A.

¿Qué tipo de variable se utiliza para representar cadenas de texto?

- A. Numérica
- B. Caracter
- C. Lógica

Respuesta: B.

¿Qué tipo de variable se utiliza para representar valores lógicos o booleanos?

- A. Numérica
- B. Caracter
- C. Lógica

Respuesta: C.

¿Qué tipo de variable se utiliza para representar números enteros?

- A. Numérica
- B. Caracter
- C. Lógica

Respuesta: A.

¿Qué tipo de variable se utiliza para almacenar información de fechas y horas?

- A. Numérica
- B. Caracter
- C. Fecha y hora

Respuesta: C.

CAPÍTULO IV

SEGMENTACIÓN NO JERÁRQUICA



CAPÍTULO IV.

SEGMENTACIÓN NO JERÁRQUICA

¿QUÉ ES SEGMENTACIÓN NO JERÁRQUICA?

La segmentación no jerárquica implica dividir un conjunto de datos en grupos distintos, donde cada grupo consiste en observaciones que comparten características similares. Este enfoque no establece una jerarquía entre los grupos. De especifican Características de la Segmentación No Jerárquica^{29,30}:

- **Número de Grupos:** Se debe especificar previamente el número de grupos deseados.
- **Criterio de Similitud:** Se utiliza un criterio de similitud para agrupar las observaciones. Pueden ser métodos de distancia, como la distancia euclidiana, o enfoques basados en densidad.
- **Sin Estructura de Árbol:** A diferencia de la segmentación jerárquica, donde los grupos se organizan en una estructura de árbol, la segmentación no jerárquica no impone una jerarquía entre los grupos.

VENTAJAS DE LA SEGMENTACIÓN NO JERÁRQUICA

La segmentación no jerárquica ofrece varias ventajas en comparación con métodos de segmentación jerárquica. Se detallan algunas de las principales ventajas:

- **Flexibilidad en el Número de Grupos:** La segmentación no jerárquica no requiere especificar previamente el número de grupos. Esto proporciona flexibilidad al adaptarse a la estructura de los datos sin imponer restricciones en la cantidad de grupos.
- **Adaptabilidad a Formas y Densidades Irregulares:** Los métodos no jerárquicos, como DBSCAN (Density-Based Spatial Clustering of Applications with Noise), son capaces de adaptarse a formas y densidades irregulares en los datos. Pueden identificar grupos de diferentes formas y tamaños, lo que los hace más flexibles en ciertos contextos.
- **Manejo de Ruido y Outliers:** Algunos métodos no jerárquicos son más robustos al manejar ruido y outliers. DBSCAN, por ejemplo, puede identificar puntos que no pertenecen a ningún grupo como "ruido", lo que lo hace útil en conjuntos de datos con presencia de datos atípicos.
- **Eficiencia Computacional:** En general, los métodos no jerárquicos pueden ser computacionalmente más eficientes, especialmente en conjuntos de datos grandes, en comparación con algunos métodos jerárquicos. Esto se debe a que no construyen una estructura de árbol que crece con el tamaño del conjunto de datos.
- **Interpretación Directa de Resultados:** Los resultados de la segmentación no jerárquica suelen ser más fáciles de interpretar directamente. Cada observación se asigna a un grupo específico sin la complejidad de la estructura jerárquica, lo que puede facilitar la comprensión de los resultados.
- **Detección de Grupos con Formas Arbitrarias:** Algunos métodos no jerárquicos son capaces de detectar grupos con formas y tamaños arbitrarios, lo que los hace adecuados para datos que no cumplen

con las suposiciones de forma esférica o elipsoide.

DESVENTAJAS DE LA SEGMENTACIÓN NO JERÁRQUICA

A pesar de sus ventajas, la segmentación no jerárquica también presenta algunas desventajas que deben considerarse al seleccionar un método de agrupamiento. Algunas son las siguientes:

- **Sensibilidad a Parámetros:** Algunos algoritmos de segmentación no jerárquica requieren la configuración de parámetros, como la distancia umbral en DBSCAN. La elección de estos parámetros puede afectar significativamente los resultados y puede ser difícil determinar los valores óptimos.
- **Dependencia de la Inicialización:** Algunos algoritmos, especialmente aquellos basados en la asignación de centroides iniciales, pueden ser sensibles a la inicialización. Dependiendo de la posición inicial de los centroides, los resultados pueden variar, lo que puede hacer que sea necesario realizar múltiples ejecuciones con diferentes inicializaciones.
- **Problemas con Grupos de Densidades Variables:** Algunos métodos no jerárquicos pueden tener dificultades para identificar grupos con densidades variables o formas complejas. Algunos grupos pueden fusionarse o fragmentarse, y la detección de grupos pequeños en áreas de baja densidad puede ser un desafío.
- **Dificultad para Manejar Conjuntos de Datos de Alta Dimensión:** En conjuntos de datos de alta dimensión, la distancia euclidiana puede perder su significado, y algunos algoritmos de segmentación no jerárquica pueden tener dificultades para identificar patrones significativos en dimensiones superiores.
- **Dependencia de la Elección del Método:** La elección del método de segmentación no jerárquica adecuado depende de la naturaleza de los datos y del tipo de patrones que se buscan. No existe un método

universalmente superior, y la elección incorrecta del método puede conducir a resultados subóptimos.

- **Limitaciones en la Interpretación de Resultados:** Aunque los resultados de la segmentación no jerárquica son más directos que los de la jerárquica, la interpretación aún puede ser desafiante, especialmente si se utilizan métodos más complejos. Comprender la interpretación de los grupos puede requerir un análisis adicional.
- **Problemas con Grupos de Formas No Convencionales:** Algunos algoritmos pueden tener dificultades para identificar grupos con formas no convencionales o agrupamientos que no se ajustan bien a las suposiciones del algoritmo utilizado.

MÉTODO K-MEANS

El método K-means es un algoritmo de agrupamiento o clustering que se utiliza para dividir un conjunto de datos en grupos homogéneos o clusters. El objetivo es asignar cada dato a un grupo de manera que la varianza intracluster sea mínima. El algoritmo sigue estos pasos básicos³¹⁻³⁴:

- **Inicialización:** Selecciona el número de clusters k que deseas obtener. Inicializa k centroides de manera aleatoria. Cada centroide representa el centro de un cluster.
- **Asignación de datos a clusters:** Asigna cada dato al centroide más cercano. Puedes utilizar medidas de distancia como la distancia euclidiana.



- **Actualización de centroides:** Calcula el nuevo centroide para cada cluster tomando el promedio de los datos asignados a ese cluster.
- **Repetición:** Repite los pasos 2 y 3 hasta que los centroides no cambien significativamente o hasta que se alcance un número máximo de iteraciones.

El algoritmo converge hacia una solución, pero el resultado puede depender de la inicialización de los centroides. Por lo tanto, a veces se ejecuta el algoritmo varias veces con diferentes inicializaciones y se selecciona el mejor resultado. El valor de k (número de clusters) a menudo se determina mediante técnicas como el método del codo, la silueta, o la validación cruzada.

ACERCAMIENTO A LA METODOLOGÍA

mediante técnicas como el método del codo, la silueta, o la validación cruzada

R

```
#Instalamos librerías necesarias
install.packages("tidyverse")
install.packages("cluster")
install.packages("factoextra")
install.packages("tidyr")
install.packages("NbClust")
```

R

```
#Cargamos la base de datos prototipo
("USArrests")
head(USArrests)
#Visualizamos la data
USArrests
```

```
> USArrests
```

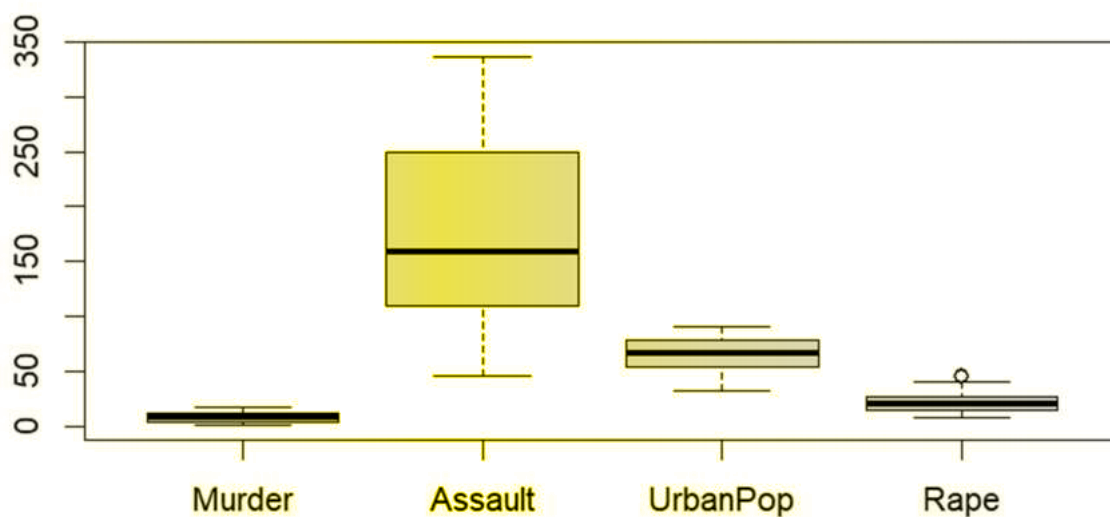
	Murder	Assault	UrbanPop	Rape
Alabama	13.2	236	58	21.2
Alaska	10.0	263	48	44.5
Arizona	8.1	294	80	31.0
Arkansas	8.8	190	50	19.5
California	9.0	276	91	40.6
Colorado	7.9	204	78	38.7
Connecticut	3.3	110	77	11.1
Delaware	5.9	238	72	15.8
Florida	15.4	335	80	31.9
Georgia	17.4	211	60	25.8
Hawaii	5.3	46	83	20.2

R

#Analizamos outliers (conocido como Valor atípico)

```
boxplot(USArrests)
```

```
boxplot.stats(USArrests$Rape)
```



En la variable “Rape” podemos ver un punto, eso es un dato atípico

R

#Verificamos valores nulos o vacios

```
summary(is.na(USArrests))
```

```
> summary(is.na(USArrests))
Murder      Assault      UrbanPop      Rape
Mode :logical Mode :logical Mode :logical Mode :logical
FALSE:50    FALSE:50    FALSE:50    FALSE:50
```

Observamos con el FALSE que no tenemos datos atípicos en ninguna de las variables presentes.

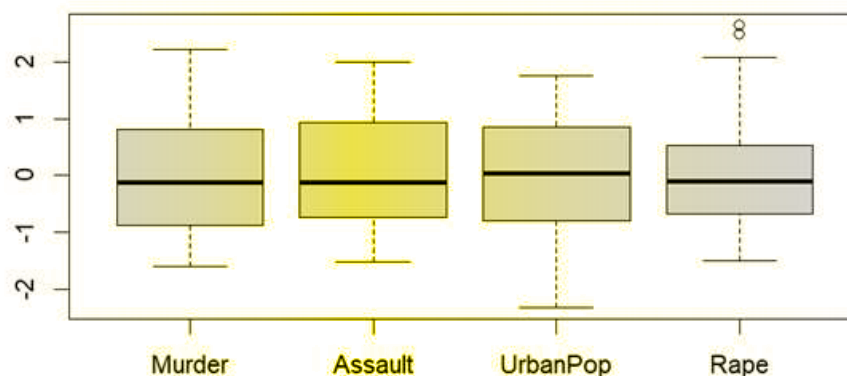
R

```
#Estandarizamos las variables (implica ajustar las variables para que tengan
una media de cero y una desviación estándar de uno)
#Restamos su promedio y dividimos su desviación estándar
datos <- scale(USArrests)
head(datos)
boxplot(datos))
```

```
> head(datos)
      Murder      Assault      UrbanPop      Rape
Alabama  1.24256408  0.7828393 -0.5209066 -0.003416473
Alaska   0.50786248  1.1068225 -1.2117642  2.484202941
Arizona  0.07163341  1.4788032  0.9989801  1.042878388
Arkansas 0.23234938  0.2308680 -1.0735927 -0.184916602
California 0.27826823 1.2628144  1.7589234  2.067820292
Colorado 0.02571456  0.3988593  0.8608085  1.864967207
```

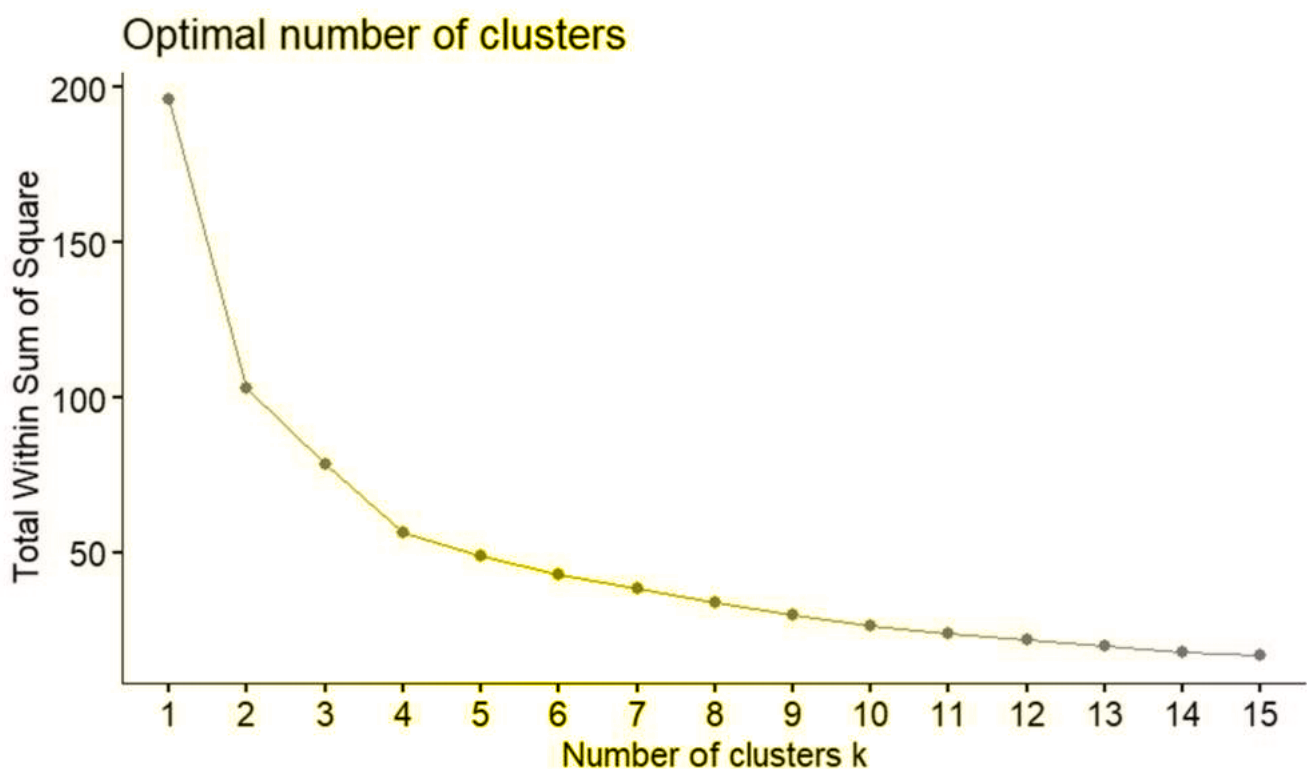
R

```
#Creamos un nuevo diagrama de caja
boxplot(datos))
```



R

```
##Aplicamos el Método del codo  
library(factoextra)  
fviz_nbclust(x = datos, FUNcluster = kmeans, method = "wss", k.max = 15,  
             diss = get_dist(datos, method = "euclidean"), nstart = 50)  
  
k <- kmeans(datos, centers = 4, nstart = 25)  
k  
#Revisamos el tamaño del Cluster  
k$size
```



Basado en el “Optimal number of cluster” se selecciona el k adecuado cuando observamos el punto de quiebre más significativo y, posterior al mismo, las demás siguen un comportamiento ordinario y aplanado. En este ejemplo el k estaría entre 2 y 4.

R

```
#centers = (ubicamos el k adecuado o podemos plotear hasta encontrar el mejor)
```

```
k <- kmeans(datos, centers = 4, nstart = 25)
```

```
k
```

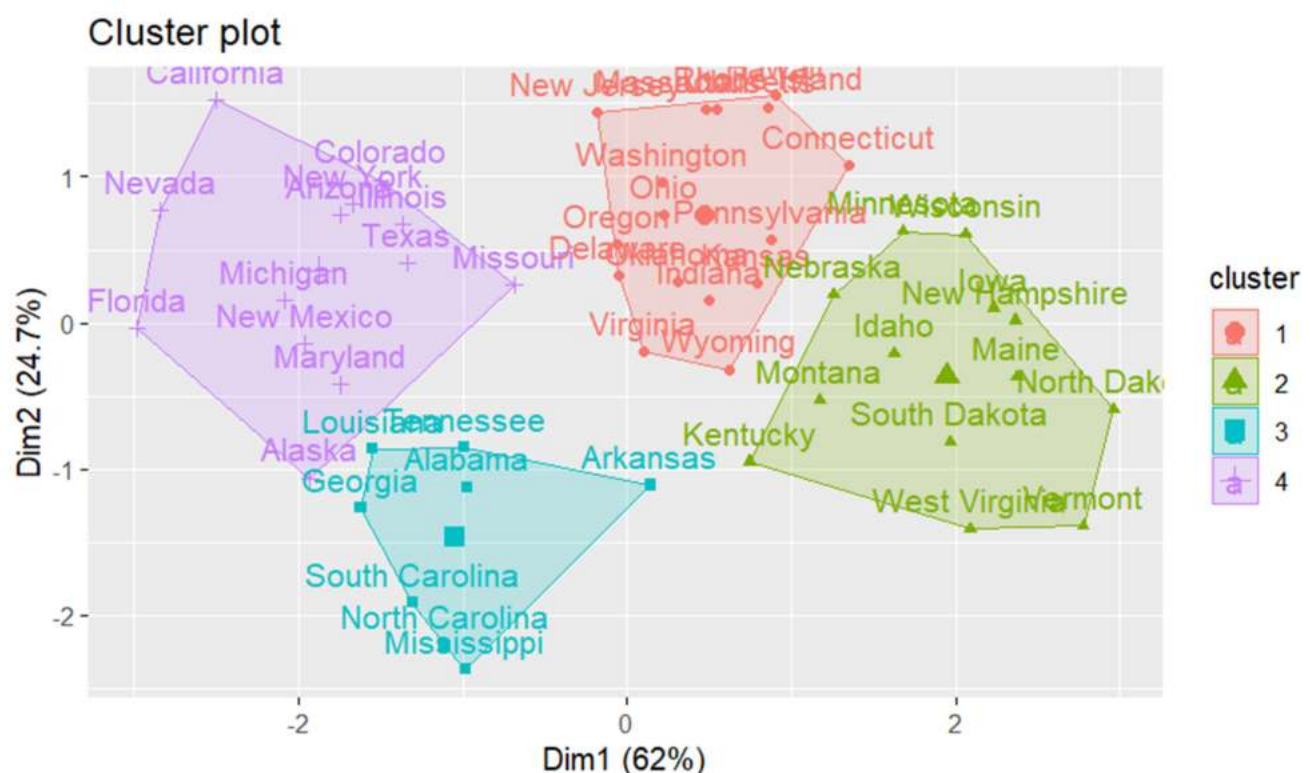
```
#Revisamos el tamaño del Cluster
```

```
k$size
```

R

```
#Piloteamos los cluster
```

```
fviz_cluster(k, data = datos)
```



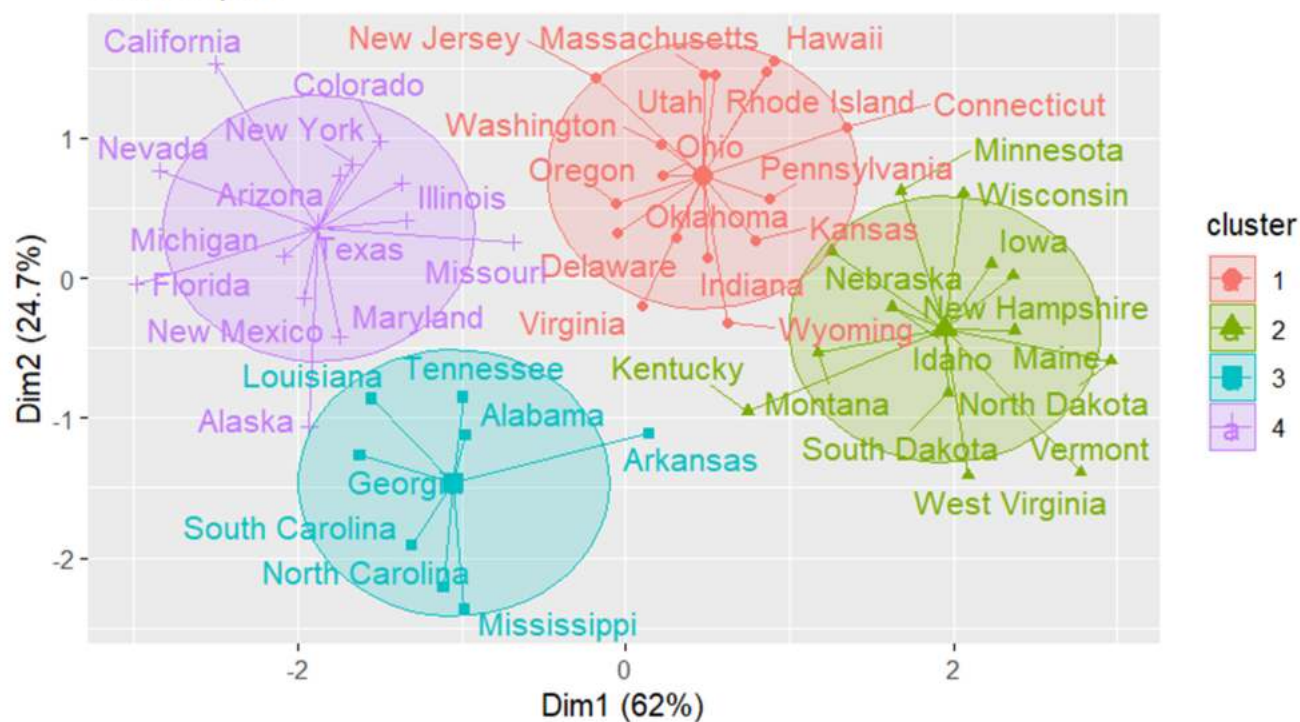
R

```
#Seleccionamos ellipse.type= Por defecto "t" asume una distribución t multivariante, y "norm" asume una distribución normal multivariante. "euclid" dibuja un círculo con el radio igual a nivel, que representa la distancia euclidiana desde el centro.
```

R

```
fviz_cluster(k, data = datos, ellipse.type = "euclid", repel = TRUE, star.plot = TRUE)
```

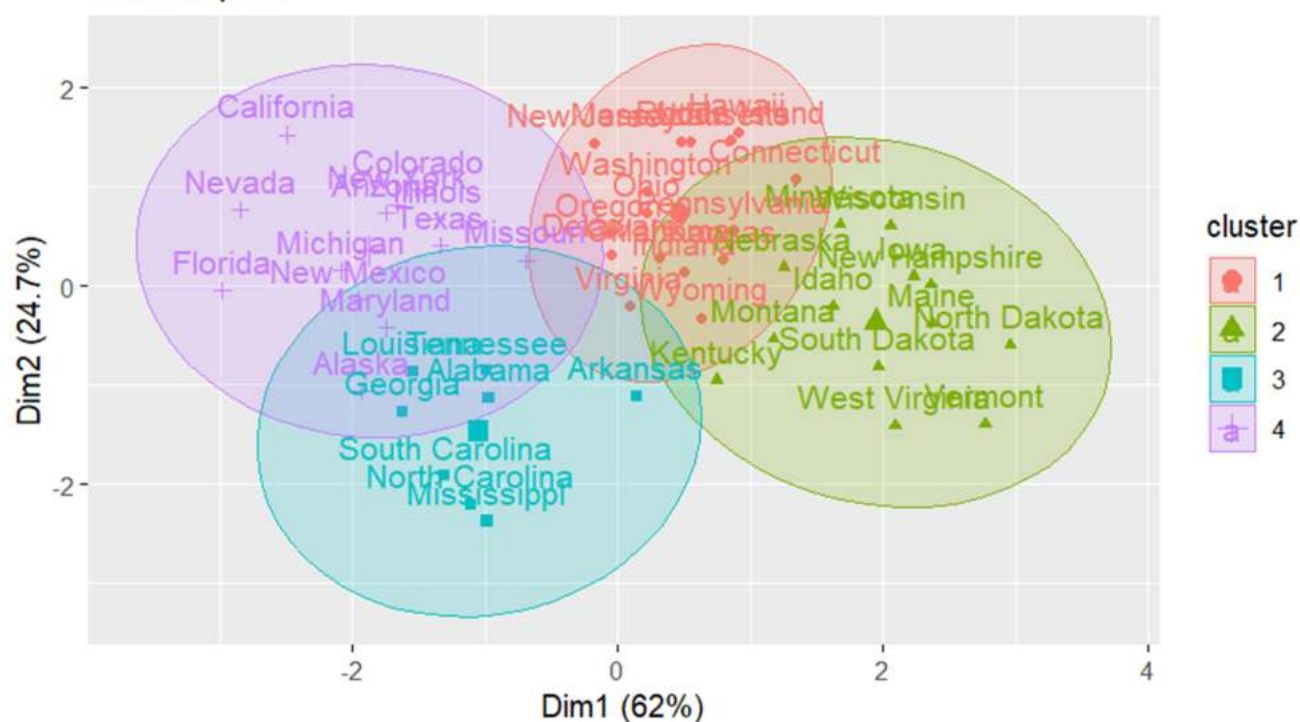
Cluster plot



R

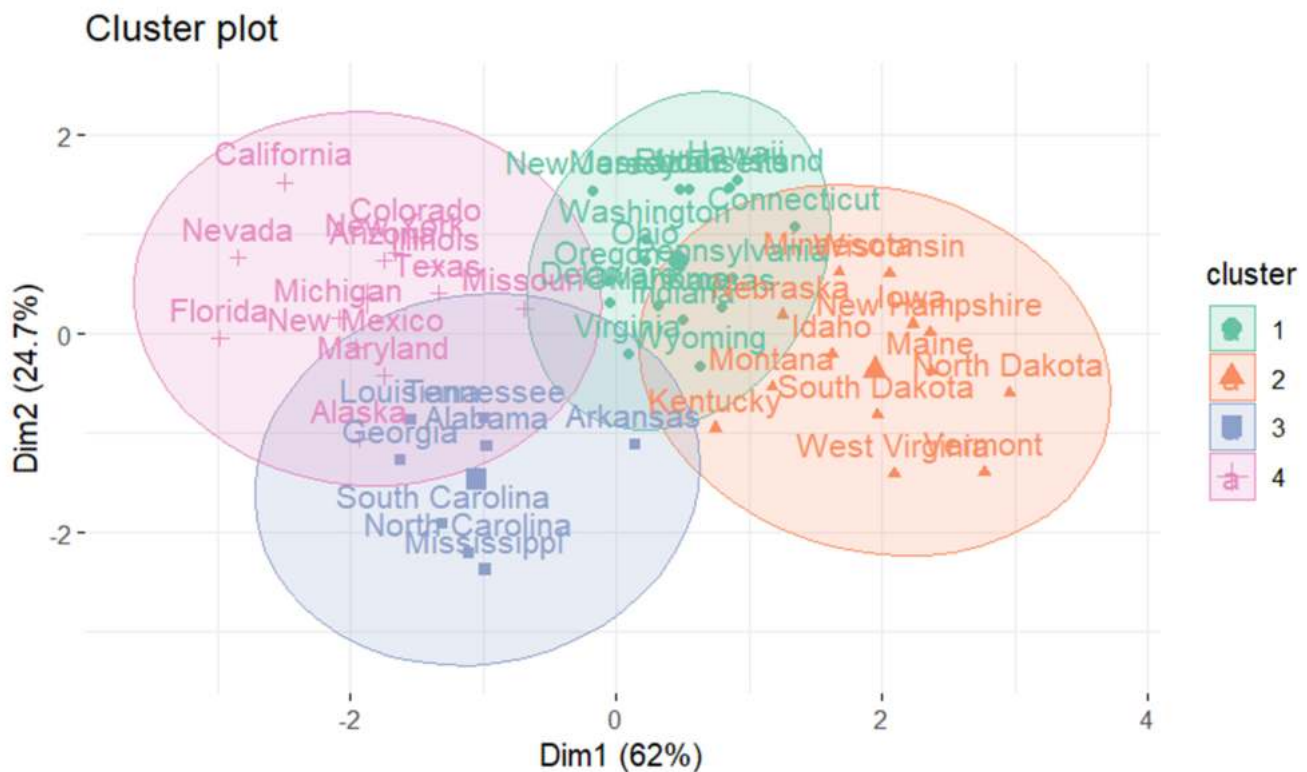
```
fviz_cluster(k, data = datos, ellipse.type = "norm")
```

Cluster plot



R

```
fviz_cluster(k, data = datos, ellipse.type = "norm", palette = "Set2", ggtheme = theme_minimal())
```

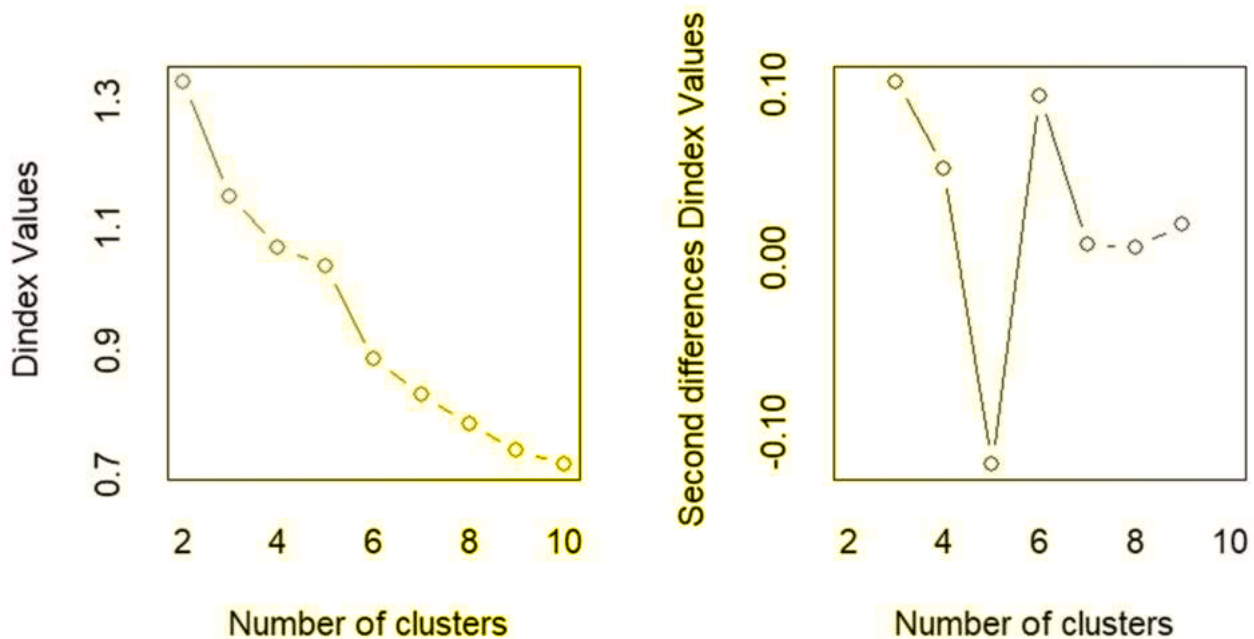


En el caso que se haga dificultoso seleccionar el clúster k adecuado podemos implementar otro código con una nueva librería; se detalla:

R

```
library(NbClust)
resnumclust<-NbClust(datos, distance = "euclidean", min.nc=2, max.nc=10,
method = "kmeans", index = "alllong")
fviz_nbclust(resnumclust)
```

```
*****
* Among all indices:
* 13 proposed 2 as the best number of clusters
* 2 proposed 3 as the best number of clusters
* 1 proposed 4 as the best number of clusters
* 1 proposed 5 as the best number of clusters
* 7 proposed 6 as the best number of clusters
* 1 proposed 9 as the best number of clusters
* 3 proposed 10 as the best number of clusters
***** conclusion *****
* According to the majority rule, the best number of clusters is 2
```



Con un gran conjunto de interacciones, el código ratifica que el k mejor adaptado a los datos es 2.

VALORACIÓN DE LOS CONOCIMIENTOS

¿Cuál es una de las ventajas de la segmentación no jerárquica?

- A. Es más flexible en el número de grupos.
- B. Es más robusta al manejar ruido y outliers.
- C. Es más eficiente computacionalmente.

Respuesta: A.

¿Cuál es una de las desventajas de la segmentación no jerárquica?

- A. Es sensible a parámetros.
- B. Es dependiente de la inicialización.
- C. Puede tener problemas con grupos de densidades variables.

Respuesta: C.

¿Cuál es el objetivo del método K-means?

- A. Dividir un conjunto de datos en grupos homogéneos.
- B. Minimizar la varianza intracluster.
- C. Asignar cada dato a un grupo de manera que la distancia entre el dato y el centroide del grupo sea mínima.

Respuesta: B.

¿Cuál es el paso inicial del método K-means?

- A. Seleccionar el número de clusters k .
- B. Inicializar k centroides de manera aleatoria.
- C. Asignar cada dato al centroide más cercano.

Respuesta: A.

¿Cuál es el paso final del método K-means?

- A. Repetir los pasos 2 y 3 hasta que los centroides no cambien significativamente o hasta que se alcance un número máximo de iteraciones.
- B. Calcular el nuevo centroide para cada cluster tomando el promedio de los datos asignados a ese cluster.
- C. Asigna cada dato al centroide más cercano.

Respuesta: A.

¿Cuál es el método más común para determinar el valor de k en el método K-means?

- A. El método del codo
- B. La silueta
- C. La validación cruzada

Respuesta: A.

¿Cuál de los siguientes métodos no jerárquicos es más adecuado para identificar grupos con formas y densidades irregulares?

- A. K-means
- B. DBSCAN
- C. Aglomeración jerárquica

Respuesta: B.

¿Cuál de los siguientes métodos no jerárquicos es más adecuado para manejar ruido y outliers?

- A. K-means
- B. DBSCAN
- C. Aglomeración jerárquica

Respuesta: B.

¿Cuál de los siguientes métodos no jerárquicos es más eficiente computacionalmente?

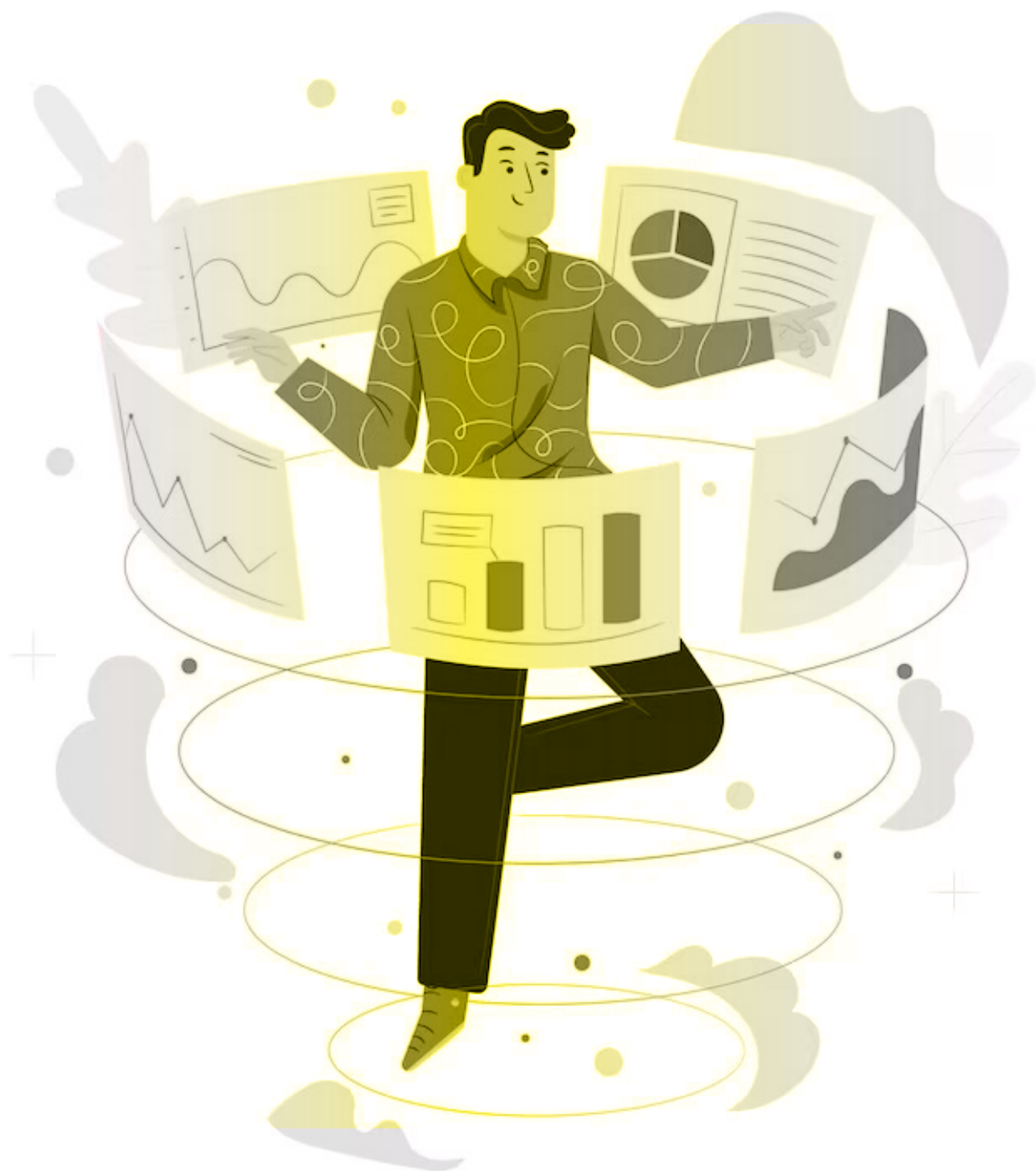
- A. K-means
- B. DBSCAN
- C. Aglomeración jerárquica

Respuesta: A.

¿Cuál de los siguientes métodos no jerárquicos es más fácil de interpretar?

- A. K-means
- B. DBSCAN
- C. Aglomeración jerárquica

Respuesta: A.



**ESTA PÁGINA HA SIDO INTENCIONALMENTE DEJADA EN
BLANCO**

CAPÍTULO V

SEGMENTACIÓN NO JERÁRQUICA APLICADA
AL ENTORNO



CAPÍTULO V.

SEGMENTACIÓN NO JERÁRQUICA APLICADA AL ENTORNO

DETALLES DEL CASO

Empresa de **Consumo Masivo**: Caso para Análisis de Clustering

Contexto: Somos una empresa local de consumo masivo que se enfrenta a un mercado diverso y dinámico. Con el objetivo de comprender mejor las características y comportamientos de nuestros clientes, hemos decidido realizar un análisis de clustering. La información clave que hemos recopilado sobre cada cliente incluye:

- **Edad**: La edad de cada cliente, un indicador importante para entender las preferencias y comportamientos de compra.
- **Línea de Crédito**: El monto máximo de crédito anual asignado a cada cliente, lo que puede indicar su capacidad de compra.
- **Km**: La distancia en kilómetros que los clientes recorren para llegar a nuestra tienda más cercana. Esto podría relacionarse con la conveniencia y accesibilidad de nuestras ubicaciones.
- **Consumo**: El gasto semanal de cada cliente en productos de consumo masivo, una medida clave de su contribución a nuestras ventas.

Objetivo del Análisis de Clustering: El objetivo principal de este análisis de clustering es identificar grupos homogéneos de clientes que comparten características similares en términos de edad, capacidad de compra, distancia de desplazamiento y patrones de consumo. Al entender estas segmentaciones, esperamos adaptar nuestras estrategias de marketing, oferta de productos y servicios para satisfacer mejor las necesidades específicas de cada grupo, mejorando así la satisfacción del cliente y aumentando la eficacia de nuestras iniciativas comerciales.

PASOS PROPUESTOS

- **Exploración de Datos:** Analizar la distribución y estadísticas descriptivas de cada variable para comprender la variabilidad en los datos.
- **Preprocesamiento de Datos:** Realizar cualquier transformación necesaria en los datos, como la normalización o estandarización, para garantizar que todas las variables tengan un peso equitativo en el análisis.
- **Selección del Algoritmo de Clustering:** Considerar algoritmos de clustering como k-means, hierarchical clustering o DBSCAN según la naturaleza de los datos y los objetivos del análisis.
- **Determinación del Número de Clústeres:** Utilizar métodos como el codo (elbow method) o el silhouette score para identificar el número óptimo de clústeres.
- **Aplicación del Algoritmo de Clustering:** Implementar el algoritmo seleccionado en los datos y asignar a cada cliente a un clúster específico.
- **Interpretación de Resultados:** Analizar los perfiles de cada clúster, identificar las características distintivas y entender las implicaciones para la estrategia comercial.
- **Desarrollo de Estrategias Personalizadas:** Utilizar la información obtenida para adaptar estrategias de marketing, promociones y servi-

cios a las necesidades y preferencias específicas de cada segmento de clientes.

CARGAMOS LA BASE DE DATOS

Cargamos un archivo CSV (Comma Separated Values), en este caso en particular se usa como delimitador o separador (;) pero puede ser distinto dependiendo del caso ejecutado, o bien, un archivo en formato texto o Excel.

R

```
library(readr)
dd <- read_delim("PRACTICA.csv", delim = ";",
                 escape_double = FALSE, trim_ws = TRUE)
View(dd)
```

	ID	Edad	Linea	Km	consumo
1	1	52	10800	12.55	24
2	2	43	9600	13.22	24
3	3	66	10800	22.42	24
4	4	30	10800	35.28	24
5	5	40	10800	35.28	24
6	6	51	11280	35.28	24
7	7	37	18120	25.19	25
8	8	43	17280	22.42	26
9	9	50	17040	28.63	26

SUAVIZACIÓN DE LA BASE DE DATOS

R

#Escalamos la base de datos y visualizamos

dd<-scale(dd[, -1])

View(dd)

	Edad	Linea	Km	consumo
1	-0.59515405	2.3941462508	0.19316205	-1.13644255
2	-0.79078540	2.3812037537	-0.76224988	-1.22039287
3	1.49158038	2.4847437304	-0.91050345	-0.52780269
4	-0.79078540	-1.0226729773	0.19316205	-1.15743013
5	-0.98641676	2.4718012333	0.64959089	-1.15743013
6	0.18737136	2.3941462508	-0.84598569	-1.43026868
7	1.42636993	2.5494562157	-0.91050345	-1.08397359
8	0.25258181	2.5106287245	-0.76224988	-1.16792392
9	-0.39952270	2.4976862274	-0.86451739	-0.71669092
10	1.68721173	2.4847437304	-0.91050345	-1.02101085

CONOCER EL K ADECUADO

R

library(NbClust)

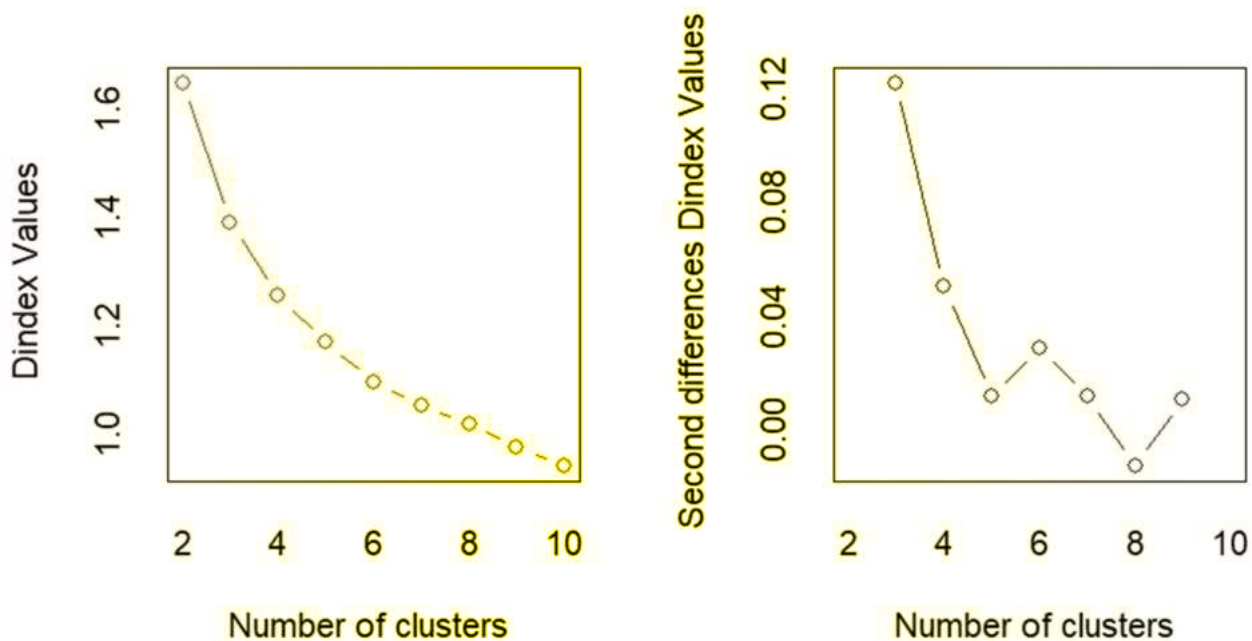
```
resnumclust_1<-NbClust(dd, distance = "euclidean", min.nc=2, max.nc=10,
method = "kmeans", index = "alllong")
```

fviz_nbclust(resnumclust_1)

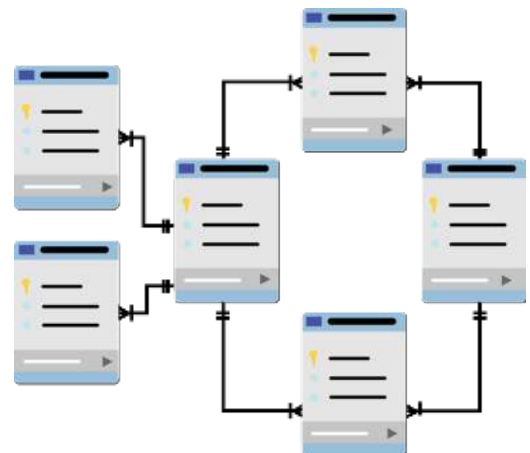
```
*****
* Among all indices:
* 5 proposed 2 as the best number of clusters
* 11 proposed 3 as the best number of clusters
* 3 proposed 4 as the best number of clusters
* 4 proposed 6 as the best number of clusters
* 4 proposed 10 as the best number of clusters

      ***** Conclusion *****

* According to the majority rule, the best number of clusters is 3
```



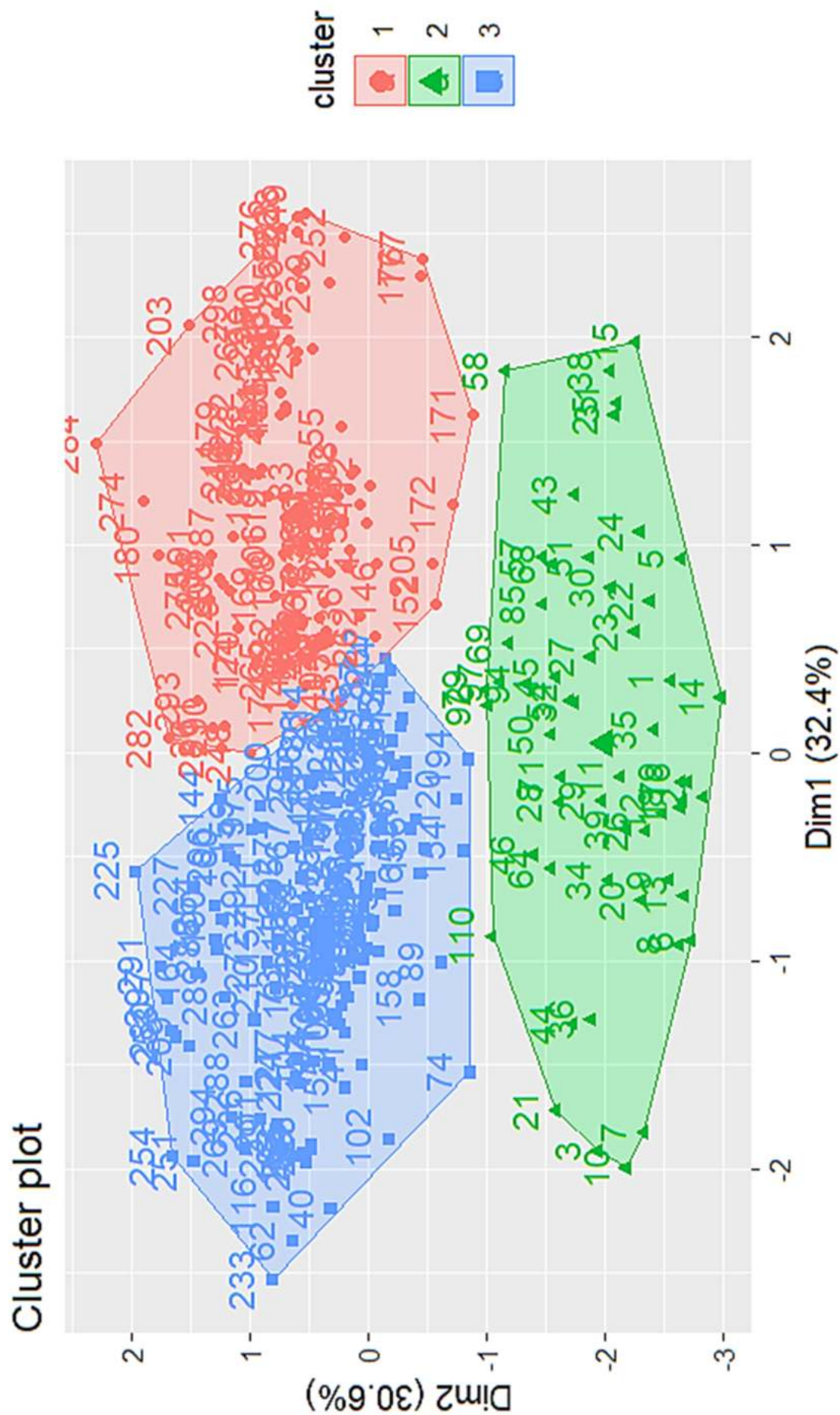
Con una extensa cantidad de interacciones, el código reafirma que el valor óptimo para el parámetro k , más adecuado para los datos, es 3. Este resultado, obtenido en el contexto de un análisis de segmentación no jerárquica, sugiere que la estructura subyacente en los datos se mejoraría y se representaría de manera más precisa mediante la partición en tres grupos distintos.



ELABORACIÓN DE GRÁFICOS CLÚSTER

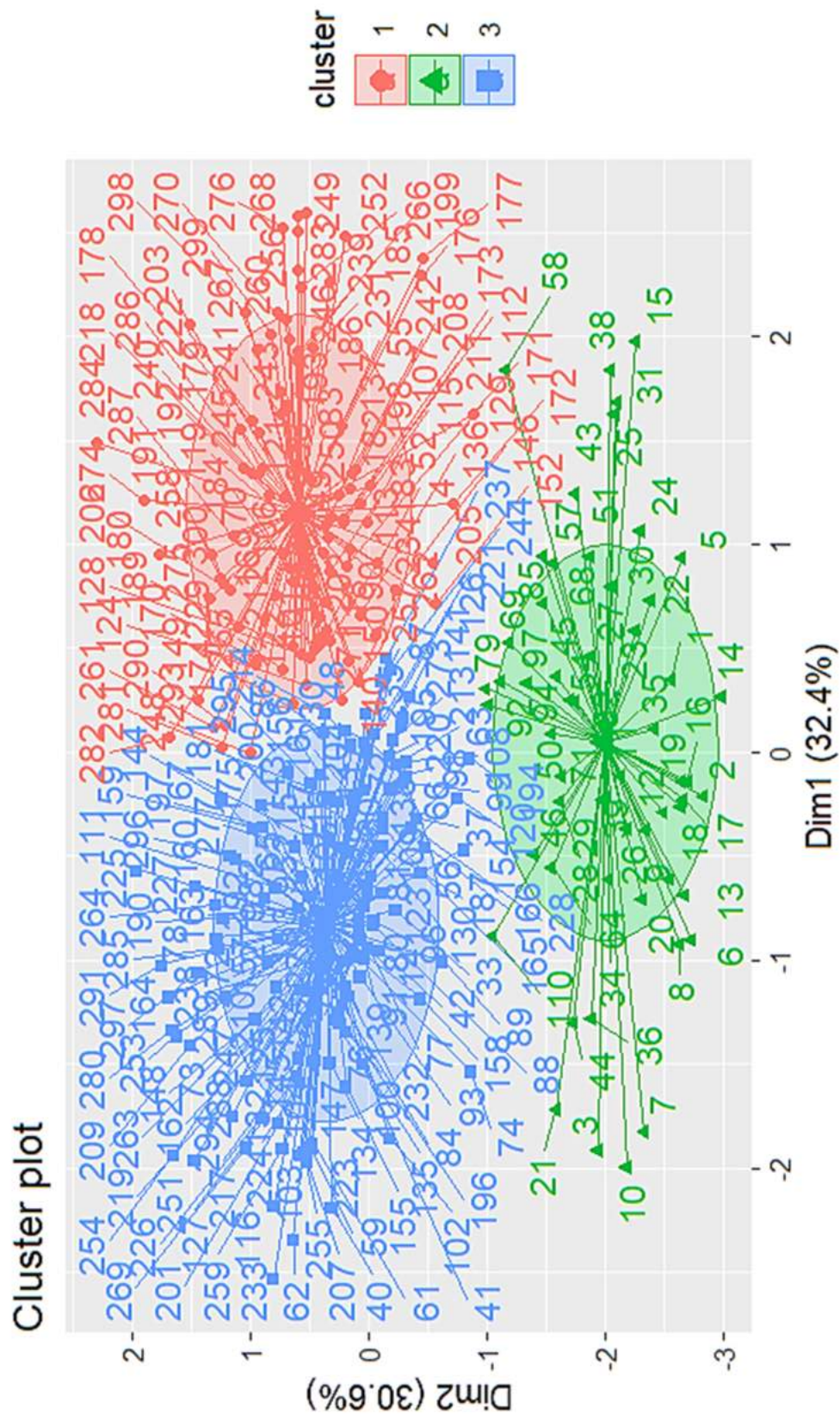
R

```
#Considerar que anteriormente el k calculado mediante código es de 3
k3 <- kmeans(dd, centers = 3, nstart = 25)
library(factoextra)
fviz_cluster(k3, data = dd)
```



R

```
fviz_cluster(k3, data = dd, ellipse.type = "euclid", repel = TRUE, star.plot = TRUE)
```



UNIFICACIÓN DE CLÚSTER

R

```
#Cargamos nuevamente la base de datos con otro nombre
ddx <- read_delim("PRACTICA.csv", delim = ";",
                  escape_double = FALSE, trim_ws = TRUE)
#Añadimos el número de clúster a cada dato
datos_con_clusters <- cbind(ddx, cluster = k3$cluster)
#Visualizamos
View(datos_con_clusters)
```

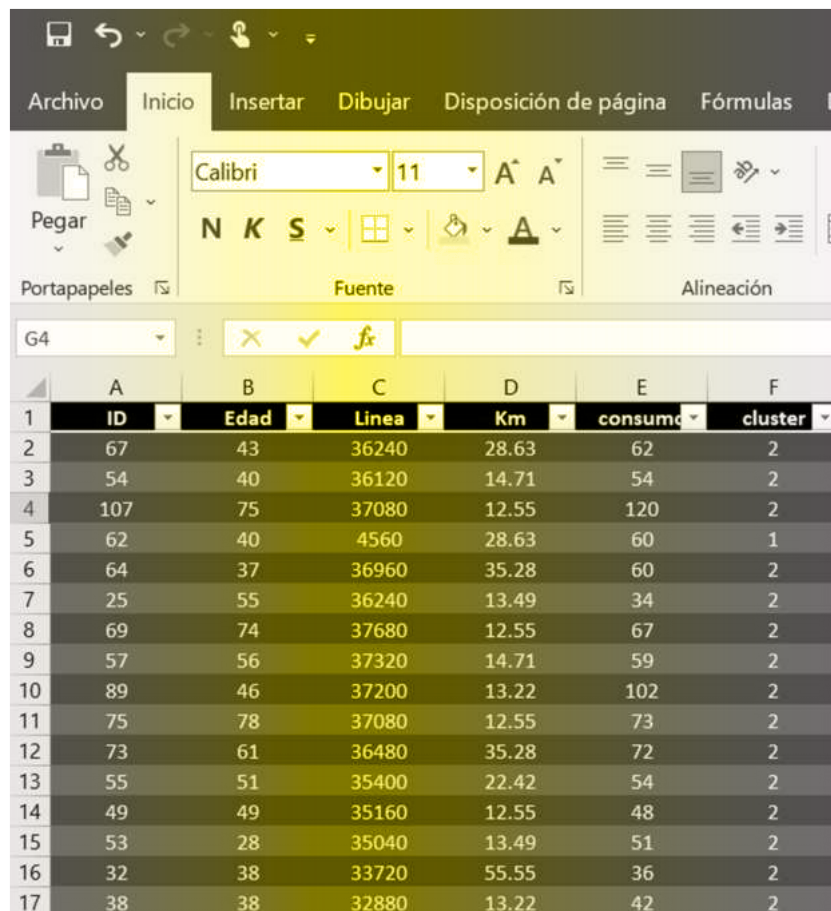
	ID	Edad	Linea	Km	consumo	cluster
1	67	43	36240	28.63	62	2
2	54	40	36120	14.71	54	2
3	107	75	37080	12.55	120	2
4	62	40	4560	28.63	60	1
5	64	37	36960	35.28	60	2
6	25	55	36240	13.49	34	2
7	69	74	37680	12.55	67	2
8	57	56	37320	14.71	59	2

DESCARGA CON CLÚSTER

R

```
#Descargamos la base de datos con los clúster definidos
write.csv(datos_con_clusters, "datos_con_clusters.csv", row.names = FALSE)
```

Una vez con el archivo en nuestro PC donde este definido nuestro directorio podemos tenerlo disponible en Excel para generar diversos análisis sobre la base.



	A	B	C	D	E	F
	ID	Edad	Linea	Km	consumo	cluster
1	67	43	36240	28.63	62	2
2	54	40	36120	14.71	54	2
3	107	75	37080	12.55	120	2
4	62	40	4560	28.63	60	1
5	64	37	36960	35.28	60	2
6	25	55	36240	13.49	34	2
7	69	74	37680	12.55	67	2
8	57	56	37320	14.71	59	2
9	89	46	37200	13.22	102	2
10	75	78	37080	12.55	73	2
11	73	61	36480	35.28	72	2
12	55	51	35400	22.42	54	2
13	49	49	35160	12.55	48	2
14	53	28	35040	13.49	51	2
15	32	38	33720	55.55	36	2
16	38	38	32880	13.22	42	2
17						

Si por alguna razón no sabes donde esta alojado tu archivo usamos el siguiente código para conocer nuestro directorio en R.

R

```
getwd()
```

```
> getwd()
[1] "C:/Users/HOME/Documents"
```

RESULTADOS ANALIZADOS

Desarrollamos tablas dinámicas para analizar cada clúster, esta actividad también podría ser desarrollada en el panel de RStudio:

Clúster	Promedio de Edad	Promedio de Línea	Promedio de Km	Promedio de consumo
1	44	\$ 8.945,35	40,91	\$ 189,48
2	48	\$ 28.660,36	24,62	\$ 49,47
3	59	\$ 12.033,33	15,68	\$ 202,99
Total general	52,13	\$ 14.042,00	25,82	\$ 170,30

- **El Clúster 1** representa un grupo de clientes con una edad promedio de 44 años, que cuentan con una línea de crédito moderada y recorren una distancia significativa para llegar a nuestras tiendas. Estos clientes tienen un nivel de consumo semanal relativamente alto, indicando un comportamiento de compra robusto.
- **El Clúster 2** está compuesto por clientes ligeramente más mayores, con una edad promedio de 48 años. Este grupo tiene una línea de crédito considerablemente alta, pero recorre distancias moderadas para llegar a nuestras tiendas. Aunque tienen una capacidad de compra significativa, el nivel de consumo semanal es más bajo en comparación con el Clúster 1.
- **El Clúster 3** está caracterizado por clientes de mayor edad, con una edad promedio de 59 años. A pesar de tener una línea de crédito más baja que el Clúster 2, estos clientes recorren distancias más cortas y tienen el mayor nivel de consumo semanal entre los clústeres. Esto podría indicar que, a pesar de una línea de crédito menor, están dispuestos a realizar compras frecuentes.

ESTRATEGIAS PARA CADA CLÚSTER

CLÚSTER 1: "CLIENTES FRECUENTES"

Perfil:

- Edad promedio de 44 años.
- Línea de crédito moderada.
- Recorren distancias significativas (40.91 km).

- Alto consumo semanal (\$189.48).

Estrategias:

1. Programas de **Lealtad Exclusivos**: Ofrecer programas de lealtad que recompensen la frecuencia de compra.
2. Promociones **Personalizadas**: Enviar ofertas personalizadas basadas en patrones de compra anteriores.
3. Servicio de **Entrega a Domicilio**: Facilitar el acceso a productos mediante servicios de entrega, dada la distancia recorrida.
4. Eventos **Especiales**: Organizar eventos exclusivos o ventas especiales para este grupo.

CLÚSTER 2: "COMPRADORES CONFORTABLES"

Perfil:

- Edad promedio de 48 años.
- Línea de crédito alta (\$28,660.36).
- Distancias moderadas (24.62 km).
- Consumo semanal más bajo (\$49.47).

Estrategias:

1. Ofertas **Exclusivas de Alto Valor**: Crear ofertas que aprovechen la alta línea de crédito.
2. Programas de **Recompensas Financieras**: Brindar recompensas financieras o descuentos por compras frecuentes.
3. Experiencia de **Compra Mejorada**: Mejorar la experiencia de compra

en la tienda para fomentar la lealtad.

4. Asesoramiento **Personalizado**: Ofrecer asesoramiento personalizado para maximizar el valor de la compra.

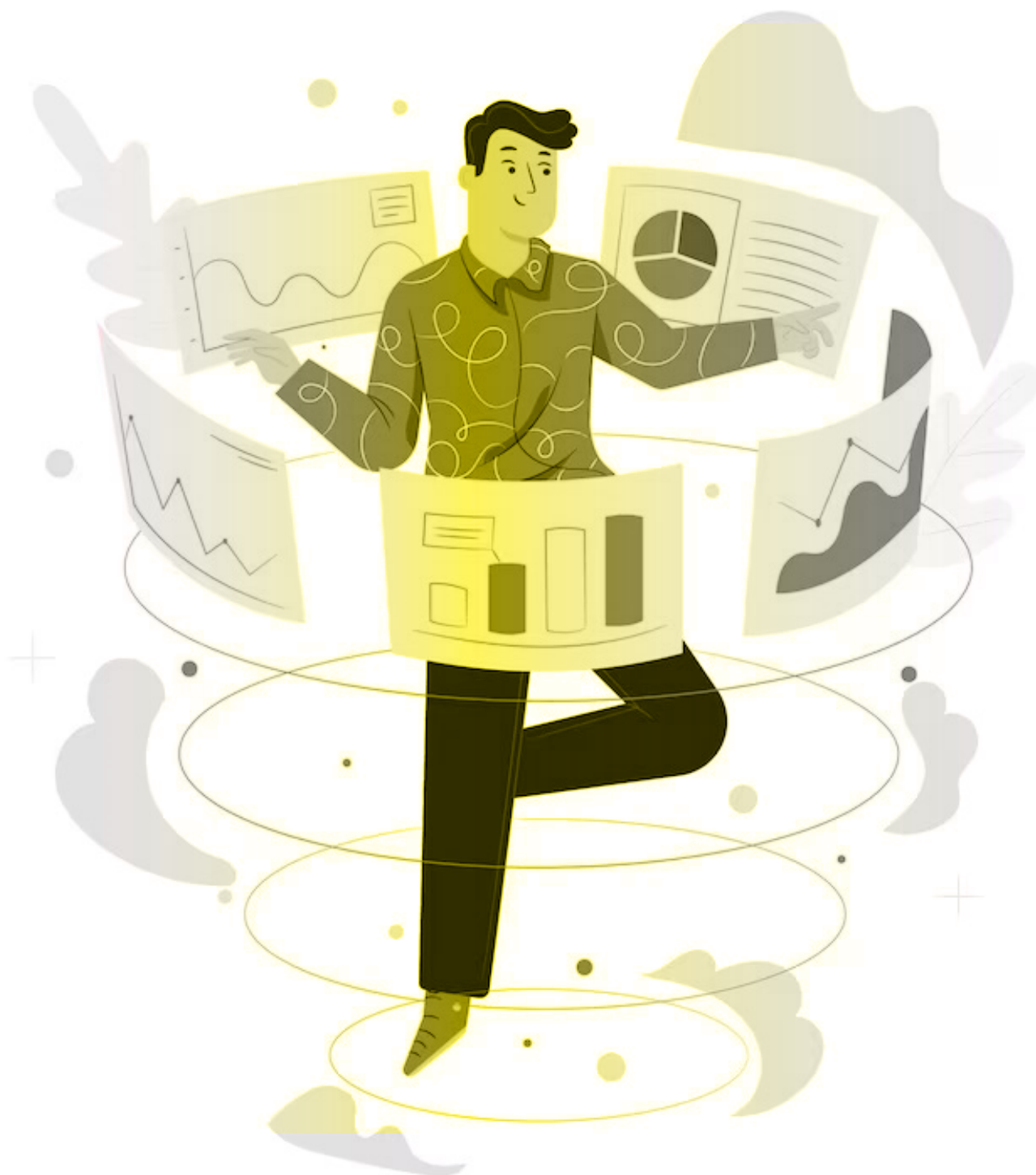
CLÚSTER 3: "SENIORS FIELES"

Perfil:

- Edad promedio de 59 años.
- Línea de crédito más baja.
- Distancias cortas (15.68 km).
- Mayor consumo semanal (\$202.99).

Estrategias:

1. Descuentos **Especiales para Mayores**: Implementar descuentos específicos para clientes mayores.
2. Programas de **Entrega a Domicilio Preferenciales**: Facilitar la entrega a domicilio para aquellos que prefieren no desplazarse largas distancias.
3. Eventos **Sociales y Comunitarios**: Organizar eventos sociales para construir una comunidad y fomentar la lealtad.
4. Promociones **Centradas en Salud**: Ofrecer productos y promociones que se centren en la salud y bienestar para esta demografía.



**ESTA PÁGINA HA SIDO INTENCIONALMENTE DEJADA EN
BLANCO**

CAPÍTULO VI

SEGMENTACIÓN JERÁRQUICA



CAPÍTULO VI.

SEGMENTACIÓN JERÁRQUICA

SEGMENTACIÓN JERÁRQUICA

La segmentación jerárquica es un enfoque de agrupamiento de datos que organiza los elementos en un árbol jerárquico. A diferencia de otros métodos de agrupamiento, como K-means, la segmentación jerárquica no requiere especificar de antemano el número de clústeres. En cambio, construye una jerarquía de clústeres en la que los elementos pueden ser parte de varios niveles de agrupación^{35–37}

Hay dos enfoques principales para la segmentación jerárquica: aglomerativo y divisivo.

- **Aglomerativo:** Comienza considerando cada punto de datos como un clúster individual y luego fusiona iterativamente los clústeres más cercanos hasta que todos los puntos de datos están en un solo clúster. En cada paso, los clústeres más cercanos se fusionan para formar un nuevo clúster.
- **Divisivo:** Comienza con todos los datos en un solo clúster y, en cada paso, divide un clúster en dos hasta que cada punto de datos esté en su propio clúster.

Ambos enfoques resultan en una jerarquía de clústeres que se puede visualizar como un dendrograma, que es un árbol que muestra cómo los clústeres se agrupan a diferentes niveles de similitud.

El proceso de agrupamiento en la segmentación jerárquica se basa en medir la similitud o disimilitud entre los elementos. La elección de la medida de distancia (o similitud) y el criterio para fusionar o dividir clústeres pueden variar según la aplicación y los datos.

VENTAJAS DE LA SEGMENTACIÓN JERÁRQUICA

La segmentación jerárquica presenta varias ventajas que la hacen atractiva en ciertos contextos:

- Sin necesidad de especificar el número de clústeres de antemano: A diferencia de algunos métodos de agrupamiento, como K-means, que requieren que se especifique el número de clústeres de antemano, la segmentación jerárquica no tiene este requisito. El dendrograma resultante permite explorar diferentes niveles de granularidad en la agrupación de datos.
- Visualización intuitiva: La jerarquía de clústeres se puede representar fácilmente mediante un dendrograma, que es una representación visual intuitiva de cómo los elementos se agrupan a diferentes niveles de similitud. Esto facilita la interpretación de la estructura de los datos y la identificación de patrones.
- Flexibilidad en la interpretación: La estructura jerárquica permite una interpretación más flexible de los datos, ya que se pueden examinar clústeres a diferentes niveles de detalle. Esto



es útil para explorar la organización de datos complejos.

- **Robustez frente a ruido:** La segmentación jerárquica puede ser robusta frente al ruido en los datos, ya que los errores o valores atípicos en un nivel específico del dendrograma pueden tener un impacto limitado en los niveles superiores.
- **No sensibilidad a la forma de los clústeres:** A diferencia de algunos métodos que asumen formas específicas para los clústeres (como K-means, que asume clústeres esféricos), la segmentación jerárquica no hace suposiciones sobre la forma de los clústeres y puede adaptarse a estructuras de datos más complejas.
- **Identificación de subgrupos:** Al explorar diferentes niveles del dendrograma, es posible identificar subgrupos y patrones más específicos dentro de los datos, lo que puede ser útil en aplicaciones donde la identificación de subestructuras es importante.

DESVENTAJAS DE LA SEGMENTACIÓN JERÁRQUICA

Aunque la segmentación jerárquica tiene varias ventajas, también presenta algunas desventajas que deben considerarse al elegir esta técnica de agrupamiento:

- **Complejidad computacional:** La segmentación jerárquica puede ser computacionalmente costosa, especialmente para conjuntos de datos grandes. La complejidad computacional es del orden de $O(n^3)$, lo que puede hacer que el método sea menos eficiente en comparación con otros enfoques, especialmente en conjuntos de datos extensos.
- **Sensibilidad a los ruidos y valores atípicos:** La jerarquía de clústeres puede



ser sensible a valores atípicos o ruido en los datos. En algunos casos, la presencia de valores atípicos puede afectar significativamente la estructura jerárquica resultante.

- **Irreversibilidad de las fusiones:** En el enfoque aglomerativo, una vez que se han fusionado dos clústeres, la decisión es irreversible. Esto significa que, si dos clústeres se han fusionado en un nivel incorrecto, no se puede corregir en los niveles superiores del dendrograma.
- **No siempre produce clústeres de forma esférica:** A diferencia de métodos como K-means, que asumen clústeres esféricos, la segmentación jerárquica no impone restricciones sobre la forma de los clústeres. Sin embargo, esto también puede ser una desventaja en situaciones donde se asume una forma específica para los clústeres.
- **Dependencia de la medida de similitud/distancia:** La elección de la medida de similitud o distancia puede afectar significativamente los resultados de la segmentación jerárquica. Diferentes medidas pueden conducir a agrupaciones distintas, y elegir la medida adecuada puede ser un desafío.
- **Dificultad en la identificación del número óptimo de clústeres:** Aunque la segmentación jerárquica no requiere especificar el número de clústeres de antemano, la elección del nivel de corte en el dendrograma para obtener clústeres finales puede ser subjetiva y no siempre está claro cuál es el número óptimo de clústeres.

MEDIDAS DE DISTANCIA: MÉTRICAS

Las medidas de distancia son herramientas esenciales en muchos campos, incluida la estadística, el aprendizaje automático, la minería de datos y la agrupación. Algunas métricas de distancia comunes utilizadas para medir la similitud o disimilitud entre puntos en un espacio métrico son:

- **Distancia Euclidiana:** Representa la longitud del camino más corto

entre dos puntos en un espacio bidimensional. Es la medida de "línea recta" entre dos puntos³⁸.

- **Distancia Manhattan (o distancia de la ciudad):** Se calcula sumando las diferencias absolutas entre las coordenadas de dos puntos en un plano. Refleja la distancia a lo largo de las calles de una ciudad, donde solo se permite moverse horizontal o verticalmente³⁹.
- **Distancia de Minkowski:** Es una generalización de las distancias Euclidiana y Manhattan. Dependiendo del valor del parámetro p , puede adoptar diferentes formas y propiedades⁴⁰.
- **Distancia de Chebyshev:** Mide la diferencia máxima entre las coordenadas de dos puntos. Representa la distancia máxima a lo largo de cualquier dimensión⁴¹.
- **Distancia de Hamming (para datos categóricos):** Utilizada para comparar secuencias de datos categóricos de igual longitud. Cuenta la cantidad de posiciones en las que los elementos correspondientes son diferentes⁴².

La elección de la mejor medida de distancia depende del contexto específico del problema y de las características de los datos. No hay una medida de distancia universalmente "mejor" o "más usada", ya que diferentes métricas son apropiadas para diferentes tipos de datos y escenarios.

En muchos casos, la distancia Euclidiana es una opción común, especialmente en algoritmos de agrupamiento como K-means. Es intuitiva y fácil de entender, y funciona bien cuando los datos se distribuyen de manera aproximadamente esférica en el espacio.

La distancia de Manhattan puede ser preferida cuando la métrica de "ciudad" es más relevante para el problema, por ejemplo, cuando la distancia a lo largo de dimensiones específicas es más significativa que la distancia diagonal.

La elección de la métrica también puede depender del tipo de datos que estás manejando. Por ejemplo, para datos categóricos, la distancia de Hamming es más apropiada, ya que tiene en cuenta las discrepancias en las categorías.

MEDIDAS DE DISTANCIA: CRITERIO DE ENLACE

El criterio de enlace, también conocido como regla de fusión, es un componente importante en algoritmos de agrupamiento jerárquico aglomerativo. Este criterio determina cómo se calcula la distancia entre dos clústeres existentes cuando se decide fusionarlos en un paso específico del proceso de agrupamiento. Se detallan algunos criterios de enlace comunes:

- **Enlace simple (o mínimo):** La distancia entre dos clústeres se define como la distancia mínima entre cualquier par de puntos, uno de cada clúster. Es decir, se toma la distancia más corta entre los elementos de los dos clústeres.
- **Enlace completo (o máximo):** La distancia entre dos clústeres se define como la distancia máxima entre cualquier par de puntos, uno de cada clúster. Es decir, se toma la distancia más larga entre los elementos de los dos clústeres.
- **Enlace promedio:** La distancia entre dos clústeres se define como el promedio de las distancias entre todos los pares de puntos, uno de cada clúster.
- **Enlace del centroide:** La distancia entre dos clústeres se define como la distancia entre sus centroides, que es el punto medio calculado a lo largo de todas las dimensiones.
- **Enlace de Ward:** Minimiza la varianza dentro de los clústeres. Al fusionar dos clústeres, se selecciona la fusión que minimiza el aumento total de la varianza dentro de los clústeres.

La elección del criterio de enlace puede afectar significativamente la estructura del dendrograma resultante y, por lo tanto, la interpretación de los clústeres. La selección del criterio de enlace depende de la naturaleza de los datos y los objetivos del análisis. Algunos criterios pueden ser más adecuados para ciertos tipos de datos o para capturar ciertos patrones en los datos. En la práctica, es común probar varios criterios de enlace y comparar los resultados para determinar cuál se ajusta mejor a los objetivos del análisis.

ACERCAMIENTO A LA METODOLOGÍA

R

```
#Instalamos y corremos las librerías necesarias
install.packages("maps")
install.packages("mapdata")
install.packages("ggplot2")
install.packages("dplyr")

library(maps)
library(mapdata)
library(ggplot2)
library(dplyr)
```

R

```
#Cargamos la data
data("USArrests")
head(USArrests)
```

	Murder	Assault	UrbanPop	Rape
Alabama	13.2	236	58	21.2
Alaska	10.0	263	48	44.5
Arizona	8.1	294	80	31.0
Arkansas	8.8	190	50	19.5
California	9.0	276	91	40.6
Colorado	7.9	204	78	38.7

R

```
#Suavizamos los datos escalándolos
datos <- scale(USArrests)
head(datos)
```

	Murder	Assault	UrbanPop	Rape
Alabama	1.24256408	0.7828393	-0.5209066	-0.003416473
Alaska	0.50786248	1.1068225	-1.2117642	2.484202941
Arizona	0.07163341	1.4788032	0.9989801	1.042878388
Arkansas	0.23234938	0.2308680	-1.0735927	-0.184916602
California	0.27826823	1.2628144	1.7589234	2.067820292
Colorado	0.02571456	0.3988593	0.8608085	1.864967207

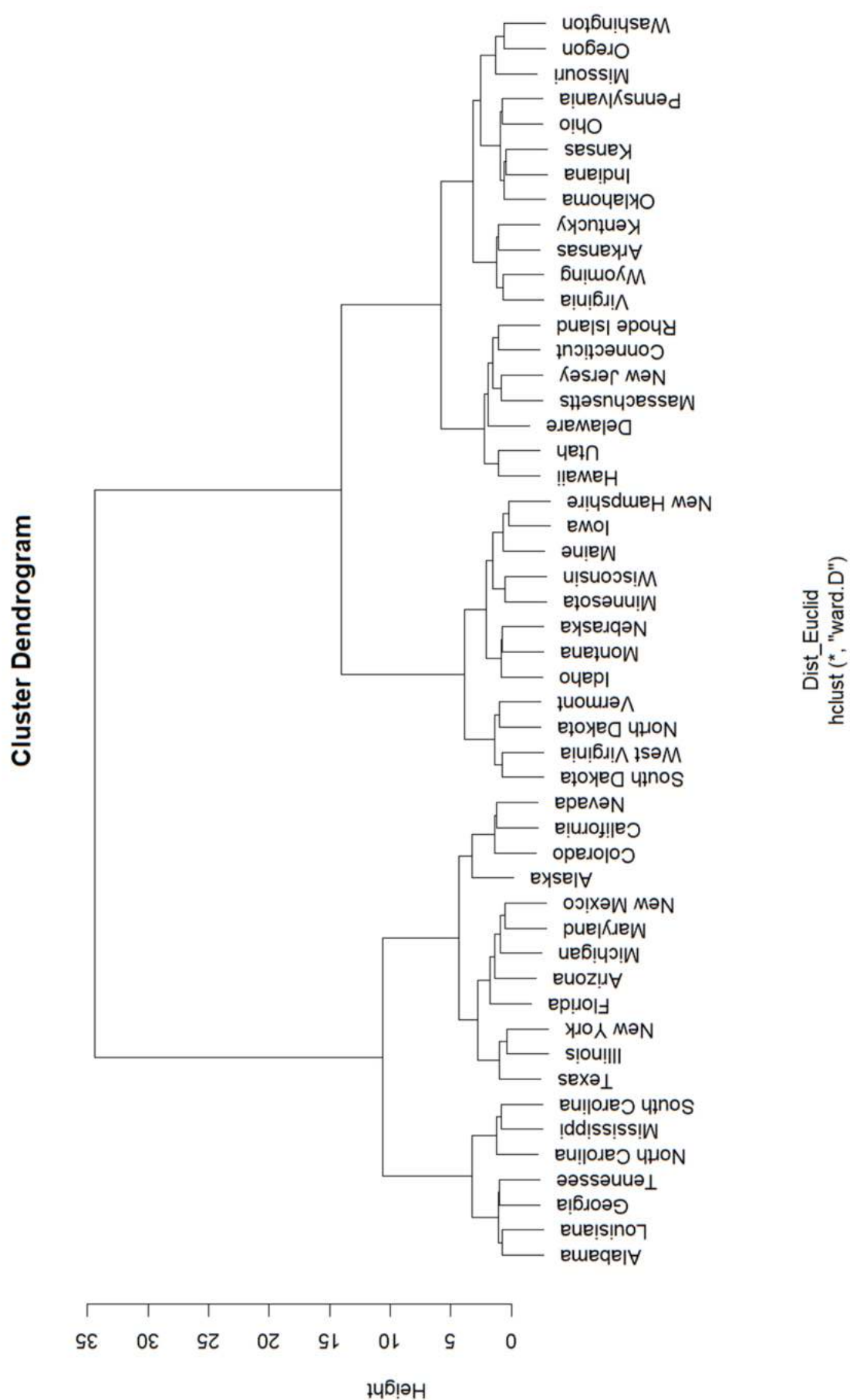
R

```
#Calculamos la distancia euclidiana
Dist_Euclid<-dist(datos,method='euclidean')
Dist_Euclid

#Calculamos un clúster bajo el enlace de Ward
Clus_Jerarquico<-hclust(Dist_Euclid, method = 'ward.D')

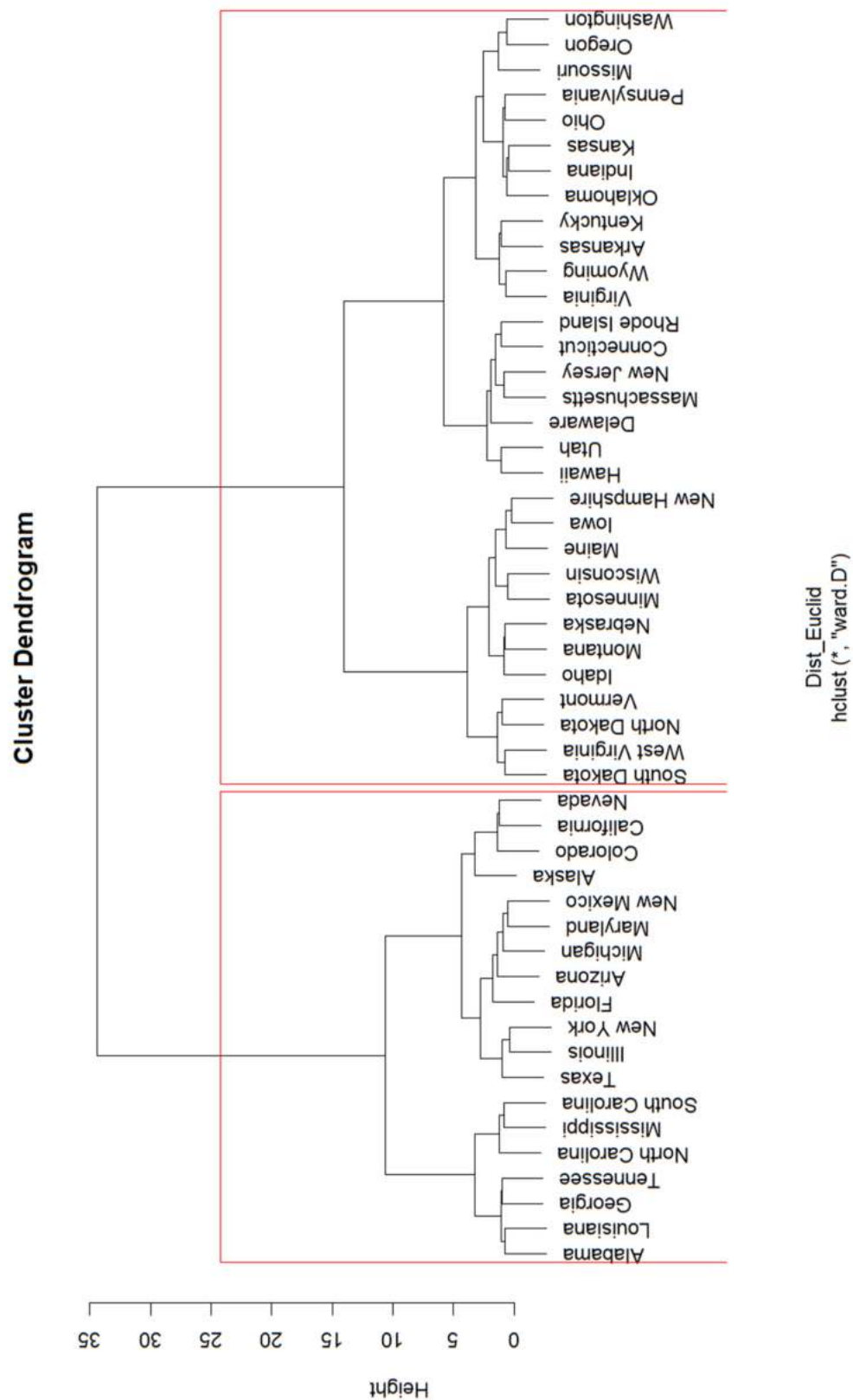
#Ploteamos
plot(Clus_Jerarquico)
```





R

```
#Ploteamos generando divisiones con R
rect.hclust(Clus_Jerarquico,k=2, border="red")
```



R

#Unificamos el clúster en la base de datos

datos_Clusters<-cbind(datos,cutree(Clus_Jerarquico,k=3))

	Murder	Assault	UrbanPop	Rape	V5
Alabama	1.24256408	0.78283935	-0.52090661	-0.003416473	1
Alaska	0.50786248	1.10682252	-1.21176419	2.484202941	1
Arizona	0.07163341	1.47880321	0.99898006	1.042878388	1
Arkansas	0.23234938	0.23086801	-1.07359268	-0.184916602	2
California	0.27826823	1.26281442	1.75892340	2.067820292	1
Colorado	0.02571456	0.39885929	0.86080854	1.864967207	1
Connecticut	-1.03041900	-0.72908214	0.79172279	-1.081740768	2

R

#Unificamos el clúster en la base de datos

datos_Clusters<-cbind(datos,cutree(Clus_Jerarquico,k=3))



CAPÍTULO VII

SEGMENTACIÓN JERÁRQUICA APLICADA AL
ENTORNO



CAPÍTULO VII.

SEGMENTACIÓN JERÁRQUICA APLICADA AL ENTORNO

DETALLES DEL CASO

Entidad Financiera **Intermediaria (Cooperativa)**: Caso para Análisis de Clustering

Contexto: El caso se centra en 30 agencias de una cooperativa de ahorro y crédito (COAC) del Ecuador. El objetivo es asignar correctamente créditos de consumo a los clientes de estas agencias.

Las cooperativas de ahorro y crédito son entidades financieras que ofrecen una variedad de productos y servicios financieros a sus socios. Los créditos de consumo son uno de los productos más populares que ofrecen.

Para asignar correctamente los créditos de consumo, las cooperativas deben tener en cuenta una variedad de factores, como la capacidad de pago del cliente, el riesgo del préstamo y las características del cliente.

Variables

Las variables que se utilizarán para el análisis de clustering son las si-

guientes:

- **Promedio de línea de crédito:** El promedio de la cantidad de crédito que ha sido aprobado por cada agencia.
- **Proporción de trabajo fijo:** La proporción de clientes de cada agencia que tienen un trabajo fijo.
- **Nivel de riesgo:** Una medida del riesgo impago de cada préstamo.
- **Efectividad de la agencia:** La proporción de ventas de créditos de consumo de cada agencia sobre el número de clientes que asisten a la agencia.

Objetivo del análisis de clustering: El objetivo del análisis de clustering es agrupar las 30 agencias en grupos homogéneos. Cada grupo representará un tipo de agencia con características similares.

PASOS PROPUESTOS

Los pasos propuestos para el análisis de clustering son los siguientes:

- **Preparación de los datos:** Los datos se deben limpiar y preparar para el análisis de clustering. Esto incluye eliminar valores atípicos y estandarizar las variables.
- **Selección del método de clustering:** Se debe seleccionar un método de clustering apropiado para el conjunto de datos. En este caso, se utilizará el clustering jerárquico.
- **Ejecución del análisis de clustering:** El análisis de clustering se ejecutará utilizando el método seleccionado.
- **Interpretación de los resultados:** Los resultados del análisis de clustering se interpretarán para



identificar los grupos de agencias homogéneos.

CARGAMOS LA BASE DE DATOS

R

```
library(readr)
Caso <- read_delim("Caso.csv", delim = ";",
  escape_double = FALSE, trim_ws = TRUE)
View(Caso)
```

	AGENCIA	Linea_Credit	Nivel_trab_Fijo	Nivel_Rsgo_Impago	Efectividad
1	Guayaquil	10557648	0,696	0,18	0,252
2	Quito	11332572	0,54	0,18	0,24
3	Cuenca	9755292	0,672	0,108	0,228
4	Santo_Domingo	12369696	0,636	0,132	0,204
5	Durán	856344	0,504	0,156	0,18
6	Machala	10544448	0,564	0,228	0,168
7	Manta	1116252	0,624	0,216	0,168
8	Portoviejo	7549092	0,384	0,216	0,168
9	Loja	12917232	0,72	0,168	0,168
10	Quevedo	9273144	0,624	0,18	0,156

SUAVIZACIÓN DE LA BASE DE DATOS

R

```
#Verificamos si tenemos datos vacíos
which(is.na(Caso))

#Escalamos la data
Caso_Std <- scale(Caso[,c(-1,-5)])
```

CALCULAMOS LA DISTANCIA EUCLIDIANA

R

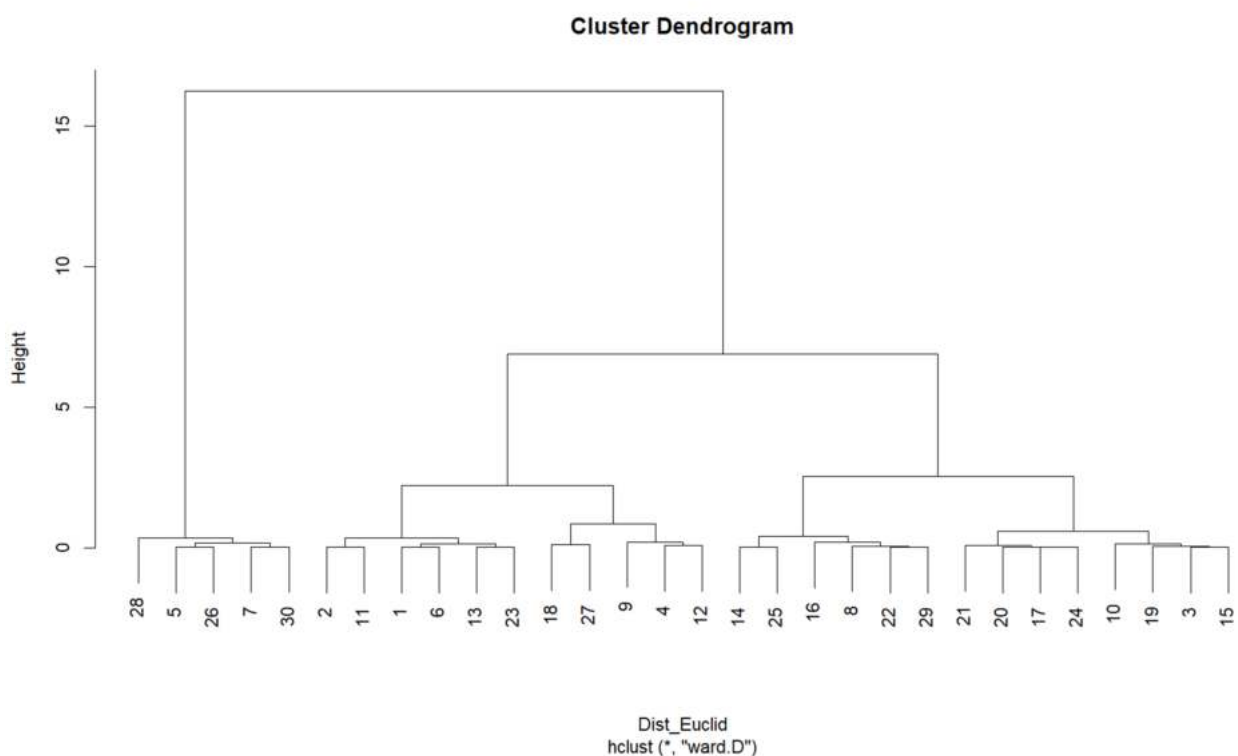
```
Dist_Euclid<-dist(Caso_Std,method='euclidean')
Dist_Euclid
```

```
25
26 1.519612438
27 1.793630628 3.313243065
28 1.708511996 0.188899558 3.502142623
29 0.129915373 1.649527811 1.663715255 1.838427369
30 1.402911540 0.116700898 3.196542168 0.305600456 1.532826913
```

GENERAMOS DEL DENDROGRAMA

R

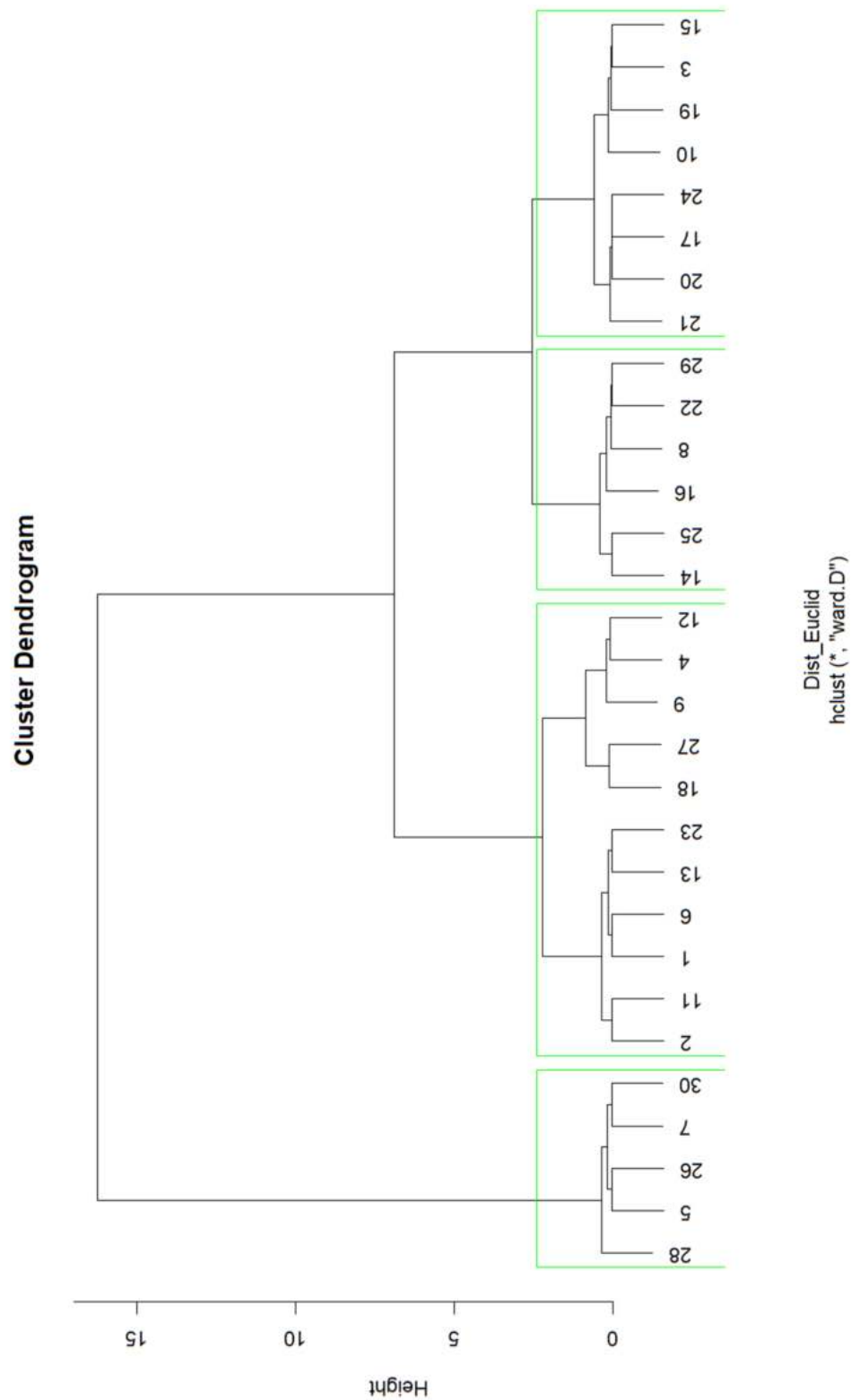
```
Clus_Jerarquico<-hclust(Dist_Euclid, method = 'ward.D')
plot(Clus_Jerarquico)
```



DENDROGRAMA CON MARCAS

R

```
rect.hclust(Clus_Jerarquico,k=4, border="green")
```



INTEGRACIÓN DEL CLÚSTER

R

```
datos_Clusters<-cbind(Caso,cutree(Clus_Jerarquico,k=3))
View(datos_Clusters)
```

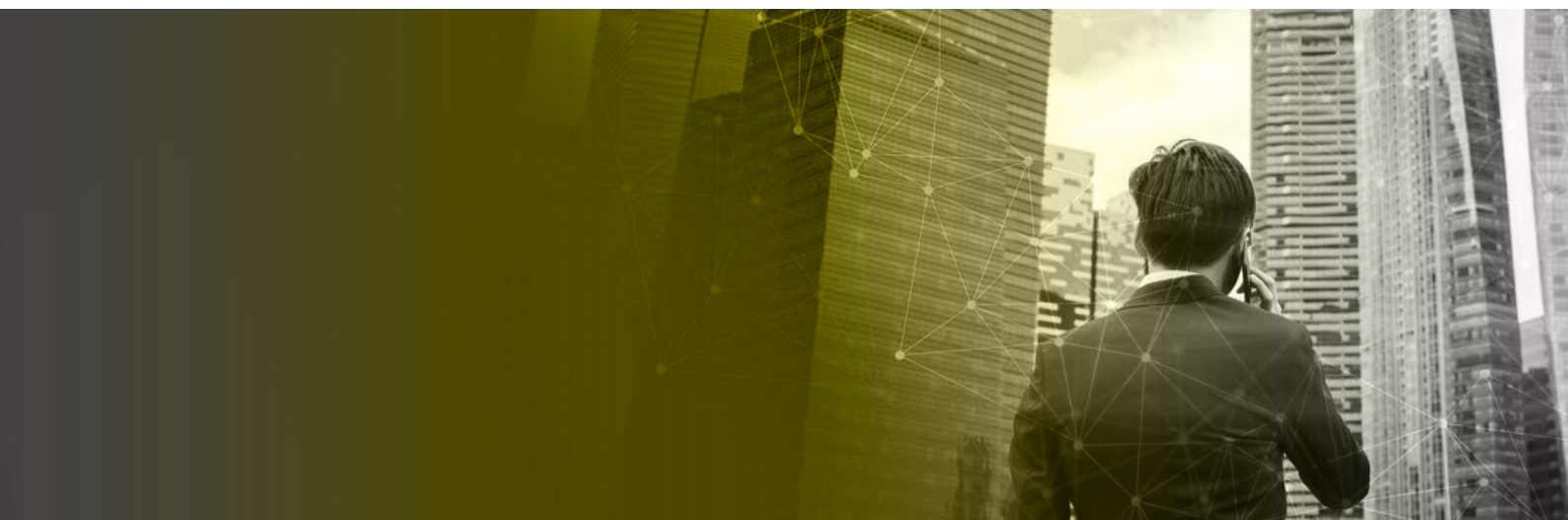
	AGENCIA	Linea_Credit	Nivel_trab_Fijo	Nivel_Rsgo_Impago	Efectividad	cutree(Clus_Jerarquico, k = 3)
1	Guayaquil	10557648	0,696	0,18	0,252	1
2	Quito	11332572	0,54	0,18	0,24	1
3	Cuenca	9755292	0,672	0,108	0,228	2
4	Santo_Domingo	12369696	0,636	0,132	0,204	1
5	Durán	856344	0,504	0,156	0,18	3
6	Machala	10544448	0,564	0,228	0,168	1
7	Manta	1116252	0,624	0,216	0,168	3
8	Portoviejo	7549092	0,384	0,216	0,168	2
9	Loja	12917232	0,72	0,168	0,168	1
10	Quevedo	9273144	0,624	0,18	0,156	2

EXPORTAR BASE CON CLÚSTER

R

```
write.csv(datos_Clusters, "Cooperativas.csv", row.names = FALSE)
```

	A	B	C	D	E	F
1	AGENCIA	Linea_Credit	Nivel_trab_Fijo	Nivel_Rsgo_Impago	Efectividad	cutree(Clus_Jerarquico, k = 3)
2	Guayaquil	10557648	0,696	0,18	0,252	1
3	Quito	11332572	0,54	0,18	0,24	1
4	Cuenca	9755292	0,672	0,108	0,228	2
5	Santo_Domingo	12369696	0,636	0,132	0,204	1
6	Durán	856344	0,504	0,156	0,18	3
7	Machala	10544448	0,564	0,228	0,168	1
8	Manta	1116252	0,624	0,216	0,168	3
9	Portoviejo	7549092	0,384	0,216	0,168	2
10	Loja	12917232	0,72	0,168	0,168	1
11	Quevedo	9273144	0,624	0,18	0,156	2



RESULTADOS ANALIZADOS

Clúster	Promedio de Línea_ Credit	Promedio de Nivel_ trab_Fijo	Promedio de Nivel_ Rsgo_Impago	Promedio de Efectividad
1	\$ 11.915.605,09	59,45%	18,65%	15,49%
2	\$ 8.402.542,29	54,43%	18,94%	11,91%
3	\$ 794.492,80	55,44%	18,88%	9,22%
Total general	\$ 8.422.657,07	56,44%	18,83%	12,78%

Clúster 1 Las cooperativas pueden asignar créditos de consumo de mayor monto a clientes con altos ingresos y con un alto nivel de estabilidad laboral.

Clúster 2 Las cooperativas pueden asignar créditos de consumo de menor monto a clientes con ingresos medios y con un alto nivel de estabilidad laboral.

Clúster 3 Las cooperativas pueden asignar créditos de consumo de menor monto a clientes con ingresos medios y con un alto nivel de estabilidad laboral.

ESTRATEGIAS PARA CADA CLÚSTER

Clúster 1: Este clúster representa a las agencias que ofrecen créditos de consumo de mayor monto, a clientes con altos ingresos y con un alto nivel de estabilidad laboral. El nivel de riesgo de los préstamos de este clúster es moderado.

Las siguientes son algunas estrategias específicas que las cooperativas pueden implementar para este clúster:

- Ofrecer productos y servicios financieros personalizados para clientes de alto nivel. Esto podría incluir créditos de consumo a plazos más largos, con tasas de interés más bajas.
- Ofrecer servicios de asesoría financiera para ayudar a los clientes a gestionar sus finanzas. Esto podría ayudar a los clientes a asegurar-

se de que pueden pagar sus préstamos de manera responsable.

- Promover los créditos de consumo de este clúster a través de canales de marketing dirigidos a clientes de alto nivel. Esto podría incluir publicidad en medios de comunicación de alta gama, como revistas y periódicos.

Clúster 2: Este clúster representa a las agencias que ofrecen créditos de consumo de menor monto, a clientes con ingresos medios y con un alto nivel de estabilidad laboral. El nivel de riesgo de los préstamos de este clúster es moderado.

Las siguientes son algunas estrategias específicas que las cooperativas pueden implementar para este clúster:

- Ofrecer productos y servicios financieros accesibles para clientes de ingresos medios. Esto podría incluir créditos de consumo con montos más bajos y tasas de interés más altas.
- Ofrecer programas de educación financiera para ayudar a los clientes a entender sus finanzas. Esto podría ayudar a los clientes a tomar decisiones financieras informadas.
- Promover los créditos de consumo de este clúster a través de canales de marketing dirigidos a clientes de ingresos medios. Esto podría incluir publicidad en medios de comunicación de masas, como la televisión y la radio.

Clúster 3: Este clúster representa a las agencias que ofrecen créditos de consumo de menor monto, a clientes con ingresos medios y con un alto nivel de estabilidad laboral. El nivel de riesgo de los préstamos de este clúster es bajo.

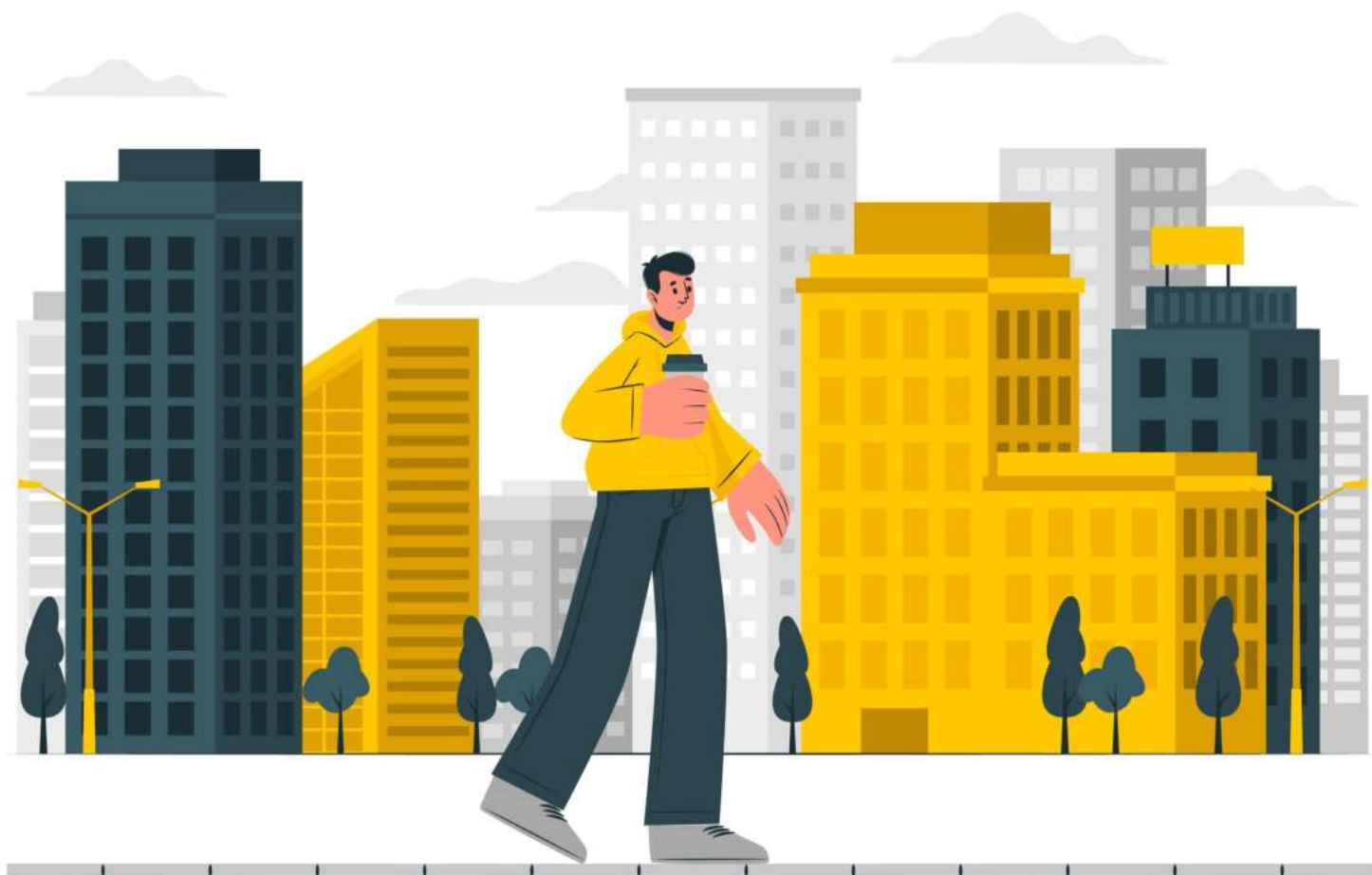
Las siguientes son algunas estrategias específicas que las cooperativas pueden implementar para este clúster:

- Ofrecer productos y servicios financieros innovadores para clien-

tes de ingresos medios. Esto podría incluir créditos de consumo con montos variables o con tasas de interés que se ajustan a los ingresos del cliente.

- Ofrecer servicios de atención al cliente personalizados para ayudar a los clientes a resolver sus problemas. Esto podría ayudar a los clientes a sentirse cómodos con su cooperativa.
- Promover los créditos de consumo de este clúster a través de canales de marketing dirigidos a clientes de ingresos medios. Esto podría incluir publicidad en medios de comunicación de nicho, como las redes sociales.

Es importante tener en cuenta que estas son solo algunas sugerencias generales. Las cooperativas deben adaptar sus estrategias a las necesidades específicas de sus clientes y de su mercado.





REFERENCIAS

1. López ÁJC, Caicedo CG, Oviedo PCJ. Análisis multivariado para segmentación de clientes basada en RFM. *Tecnura*. 2017;21(54):41-51. doi:10.14483/22487638.12957
2. Kotler P, Armstrong G, Escalona Garcia RT. Fundamentos de marketing. Published online 2003.
3. Gouvêa MA, Niño FM. Marketing estratégico: segmentación y posicionamiento de las agencias de viajes de San Pablo, Brasil. *Estudios y Perspectivas en Turismo*. 2004;13(3/4):240-257.
4. Águeda ET, Antonio MJ Juan, otros. Fundamentos de marketing. ESIC; 2022.
5. Morales Cajahuaranga HR, Saravia Lagos CE. Prevalencia de la segmentación de mercado como técnica óptima para el posicionamiento empresarial. Revisión sistemática. Repositorio Institucional - UCV. Published online 2021. Accessed November 28, 2023. <https://repositorio.ucv.edu.pe/handle/20.500.12692/82867>
6. Morillo Moreno MC, Morillo Moreno M del C, Coa Ravelo MDV. Diversidad de usuarios en las agencias bancarias: Un análisis de clases latentes, municipio Libertador, estado Mérida, Venezuela. *Fórum Empresarial*. 2015;20(1):29-58.
7. Picón Prado E, Varela Mallou J. Una evaluación empírica del procedimiento de segmentación conjunto de clases latentes. *Metodología de las ciencias del comportamiento*. 2004;5(Extra 1):491-498.
8. Gordon Ospina DH. Modelo de segmentación de clientes del Canal Trade Evedistribuciones S.A.S. Published online June 21, 2012. Accessed November 28, 2023. <http://repositorio.ucp.edu.co/handle/10785/1060>
9. Aragón Gallegos ADC, Cerquin Silva SI, Ecurra Yactayo RO, Roncalla Viena AL. Segmentación de clientes para mejorar la experiencia de compra de productos electrónicos en Falabella. Published online

2023. Accessed November 28, 2023. <https://repositorio.esan.edu.pe/handle/20.500.12640/3380>

10. Beerli-Palacio A, Martín-Santana JD, Porta M. El marketing como herramienta para incrementar la eficacia de los planes de salud pública. Informe SESPAS 2008. Gaceta Sanitaria. 2008;22:27-36. doi:10.1016/S0213-9111(08)76072-2

11. Sarango AH. EL MARKETING DIGITAL: UN MEDIO DE DIGITALIZACIÓN DE LAS PYMES EN ECUADOR EN TIEMPOS DE PANDEMIA: Manuscrito recibido: 23/12/2020 Aceptado: 17/05/2021. Investigación y Desarrollo. 2021;14(1):121-133. doi:10.31243/id.v14.2021.1277

12. Alcaide JC. Fidelización de clientes 2a. ESIC Editorial; 2015.

13. Pareti S, Córdova A, Fuentes JL. Tecnologías de la información y comunicación (CRM y ERP) como herramientas de segmentación y fidelización de clientes: El caso de los emprendimientos hoteleros de cerro alegre y cerro concepción, Valparaíso Chile. RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao. 2020;2020(E31):342-356.

14. Porter ME. Ventaja Competitiva: Creación y sostenimiento de un desempeño superior. Grupo Editorial Patria; 2016.

15. Gladys R, Choque Pilco RG. Segmentación de mercados y posicionamiento: elementos claves para el éxito de una estrategia de Marketing. Published online November 30, 2019.

16. Verdesoto GJZ. La gestión de relación con los clientes (CRM) en el desarrollo comercial de las empresas. Espíritu Emprendedor TES. 2020;4(3):49-64. doi:10.33970/eetes.v4.n3.2020.204

17. Muñoz Muñoz A, Sepúlveda Muñoz V. Análisis comparativo de estrategias de posicionamiento para la Empresa Migisa Traders, Puebla. Published online 2022. Accessed November 28, 2023. <http://repository.unilibre.edu.co/handle/10901/26421>

18. Yankelovich D, Meer D. Redescubriendo la segmentación de mercado. *Harvard Business Review*. 2006;84(2):98-108.
19. Rengifo Sanchez EP, Sangama Rivera J. Segmentación demográfica y su relación con la cultura del ahorro en los comerciantes de los mercados de Tarapoto 2021. Repositorio - UNSM. Published online December 29, 2022. Accessed November 28, 2023. <http://repositorio.unsm.edu.pe/handle/11458/5017>
20. Jeria Cánovas A, Wall Opazo C. Segmentación psicográfica : una aplicación para Chile. Published online January 2005. Accessed November 28, 2023. <https://repositorio.uchile.cl/handle/2250/142085>
21. Matute G, Bohórquez K, Carbajal L, Díaz C, Espinoza A, Jiménez C. Segmentación psicográfica de la pequeña y microempresa. Universidad ESAN; 2008. Accessed November 28, 2023. <https://repositorio.esan.edu.pe///handle/20.500.12640/151>
22. Hernández Oquendo B, López Avila L, Cuadreado Vega A, Ochoa Alvear C. Plan de mejoramiento de la empresa Distribuciones Quimicord S.A.S. Published online 2017. doi:10.57799/11227/2985
23. Ciribeli JP, Miquelito S. La segmentación del mercado por el criterio psicográfico: un ensayo teórico sobre los principales enfoques psicográficos y su relación con los criterios de comportamiento. *Visión de futuro*. 2015;19(1):0-0.
24. Espinosa-Zúñiga JJ. Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería, investigación y tecnología*. 2020;21(1). doi:10.22201/fi.25940732e.2020.21n1.008
25. López-Roldán P, Fachelli S. Software para el análisis de datos : SPSS, R y SPAD. *Metodología de la investigación social cuantitativa*. Published online 2015:cap. III.1.
26. Fernández Lizana ML. Ventajas de R como herramienta para el Análisis y Visualización de datos en Ciencias Sociales. *Revis-*

ta Científica de la UCSA. 2020;7(2):97-111. doi:10.18004/ucsa/2409-8752/2020.007.02.097

27. Verzani J. Getting Started with RStudio. O'Reilly Media, Inc.; 2011.

28. Corporación Centro de Investigación en Palma de Aceite C, Vargas LE, Mesa-Fúquen E. Introducción al análisis de datos con RStudio. Cenipalma; 2021. Accessed November 28, 2023. <https://repositorio.fedepalma.org/handle/123456789/141281>

29. Huerta Bustos MA. Segmentación de Clientes de Tienda Retail Especialista. Published online 2011. Accessed November 28, 2023. <https://repositorio.uchile.cl/handle/2250/104296>

30. Ramírez Soplin ML. Identificación de perfiles de clientes crediticios aplicando técnicas de segmentación y regresión logística multinomial. Universidad Nacional Agraria La Molina. Published online 2014. Accessed November 28, 2023. <http://repositorio.lamolina.edu.pe/handle/20.500.12996/2280>

31. Likas A, Vlassis N, J. Verbeek J. The global k-means clustering algorithm. Pattern Recognition. 2003;36(2):451-461. doi:10.1016/S0031-3203(02)00060-2

32. Na S, Xumin L, Yong G. Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm. In: 2010 Third International Symposium on Intelligent Information Technology and Security Informatics. ; 2010:63-67. doi:10.1109/IITSI.2010.74

33. Hamerly G, Elkan C. Learning the k in k-means. In: Advances in Neural Information Processing Systems. Vol 16. MIT Press; 2003. Accessed November 28, 2023. <https://proceedings.neurips.cc/paper/2003/hash/234833147b97bb6aed53a8f4f1c7a7d8-Abstract.html>

34. Ahmed M, Seraj R, Islam SMS. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. Electronics. 2020;9(8):1295. doi:10.3390/electronics9081295

35. Coslado Aristizabal FJ. Desarrollo hardware de un algoritmo para segmentación jerárquica de imágenes foveales. <http://purl.org/dc/dcmitype/Text>. Universidad de Málaga; 2004. Accessed November 28, 2023. <https://dialnet.unirioja.es/servlet/tesis?codigo=235256>
36. Herrera Conislla DM. Técnica de segmentación jerárquica Chaid de clientes para otorgamiento de créditos financieros. Repositorio de Tesis - UNMSM. Published online 2016. Accessed November 28, 2023. <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/6120>
37. Rivera E de JZ. Segmentación jerárquica en redes: aplicaciones. <http://purl.org/dc/dcmitype/Text>. Universidad Complutense de Madrid; 2013. Accessed November 28, 2023. <https://produccioncientifica.ucm.es/documentos/5d1df61929995204f76613d8>
38. Hernandez F, Millward A, México B. El Concepto de Distancia y su Aplicación en Estadística Multivariada. Published online November 28, 2023.
39. Demey J, Pla L, Vicente Villardon JL, Di Rienzo J, Casanoves F. MEDIDAS DE DISTANCIA Y SIMILITUD. In: ; 2011:47-59.
40. Díaz JMP. Geometría del espacio-tiempo: los diagramas de Minkowski. Pre-Impresos Estudiantes. 2020;(17). Accessed November 28, 2023. <https://revistas.pedagogica.edu.co/index.php/Pl/article/view/15617>
41. Saavedra Molina JF saavedra.juan@outlook.com. Identificación de nuevos factores discriminantes basados en distancia entre aminoácidos en zona de interacción entre proteínas. Published online 2017. Accessed November 28, 2023. <http://repobib.ubiobio.cl/jspui/handle/123456789/3177>
42. Salvador J. Optimización y Procesamiento Inteligente de Grandes Volúmenes de Datos Categóricos. Published online 2019. Accessed November 28, 2023. <http://rua.ua.es/dspace/handle/10045/95548>

ISBN: 978-9942-7173-2-0



9 789942 717320