HEALTH LEADERSHIP
AND QUALITY OF LIFE

Check for
updates

# Unveiling Customer Personalities Using Segmentation and Exploratory Data Analysis

## Revelación de las personalidades de los clientes mediante la segmentación y el análisis exploratorio de datos

Amitabh Chandan[1] ✉

[1]Department of Management, BIT Mesra extension Centre. Lalpur, Ranchi, India.

**ABSTRACT**

**Introduction:** this study focuses on analyzing customer personalities through segmentation techniques and exploratory data analysis (EDA) to better understand consumer behavior and preferences. In a highly competitive market, gaining a deeper understanding of customer needs is crucial for creating personalized marketing strategies and enhancing the overall customer experience.
**Objective:** the research utilizes advanced segmentation approaches to group customers based on shared traits, preferences, and behaviors, while EDA is applied to uncover trends and insights within extensive datasets.
**Method:** by identifying distinct customer segments, the study provides valuable recommendations that businesses can use to align their products, services, and marketing efforts with the unique demands of their clientele.
**Results:** the insights derived from this research enable companies to implement data-driven strategies that not only enhance customer satisfaction but also foster long-term growth.
**Conclusions:** by tapping into these analytical findings, organizations can optimize their decision-making processes and build stronger connections with their target audiences, ultimately positioning themselves for success in an increasingly data-oriented business landscape.

**Keywords:** Personalities; Segmentation; Exploratory Data Analysis.

**RESUMEN**

**Introducción:** este estudio se centra en el análisis de las personalidades de los clientes mediante técnicas de segmentación y análisis exploratorio de datos (AED) para comprender mejor el comportamiento y las preferencias de los consumidores. En un mercado altamente competitivo, comprender mejor las necesidades de los clientes es crucial para crear estrategias de marketing personalizadas y mejorar la experiencia general del cliente.
**Objetivo**: la investigación utiliza enfoques avanzados de segmentación para agrupar a los clientes en función de rasgos, preferencias y comportamientos compartidos, mientras que el AED se aplica para descubrir tendencias y conocimientos dentro de amplios conjuntos de datos.
**Método:** al identificar distintos segmentos de clientes, el estudio ofrece valiosas recomendaciones que las empresas pueden utilizar para alinear sus productos, servicios y esfuerzos de marketing con las demandas únicas de su clientela.
**Resultados:** los conocimientos derivados de esta investigación permiten a las empresas implementar estrategias basadas en datos que no solo mejoran la satisfacción del cliente, sino que también fomentan el crecimiento a largo plazo.

**Conclusiones:** al aprovechar estos hallazgos analíticos, las organizaciones pueden optimizar sus procesos de toma de decisiones y construir conexiones más fuertes con sus públicos objetivo, posicionándose en última instancia para el éxito en un panorama empresarial cada vez más orientado a los datos.

**Palabras clave:** Personalidades; Segmentación; Análisis Exploratorio de Datos.

## INTRODUCTION

Customer personality analysis is a process that combines segmentation and exploratory data analysis (EDA) to gain a comprehensive understanding of a company's audience. Segmentation groups customers into distinct categories based on shared traits such as demographics, behaviors, and preferences, allowing businesses to identify patterns within their consumer base.[1] EDA complements this by uncovering deeper insights through detailed data exploration, visualization, and interpretation. This approach highlights key trends, customer needs, and potential challenges that might otherwise go unnoticed.[2]

By leveraging these techniques, organizations can craft highly tailored marketing strategies, optimize their product offerings, and enhance customer experiences. The ability to personalize interactions not only strengthens customer loyalty but also positions businesses to anticipate and adapt to evolving consumer demands.[3] This data-driven methodology ensures that companies remain competitive, fostering stronger customer relationships and paving the way for sustainable growth in an ever-changing market landscape. Through effective analysis and actionable insights, customer personality analysis becomes a valuable tool for achieving strategic business objectives and delivering meaningful value to consumers.[4]

### Literature review

Customer segmentation has been a cornerstone of effective marketing strategies. Researchers have emphasized that dividing customers into meaningful subgroups allows businesses to understand the diversity of their customer base. Studies by Smith and Taylor (2020) show that segmentation improves the accuracy of marketing campaigns by tailoring offerings to specific demographics, behaviors, and preferences. Exploratory data analysis (EDA) plays a critical role in visualizing and interpreting data to uncover trends, providing a foundation for more effective segmentation.[5] EDA serves as a vital tool for understanding consumer behavior. Through techniques such as data visualization, pattern detection, and statistical summaries, EDA identifies hidden insights within complex datasets. Brown et al. (2021) demonstrated that EDA enhances the decision-making process by revealing trends and anomalies that inform customer segmentation efforts.[6] This capability ensures businesses can align their strategies with consumer expectations more effectively. Segmentation and personalization are deeply intertwined, as effective segmentation forms the backbone of personalized customer experiences. Research by Lee and Kim (2019) highlights how segmentation allows businesses to create tailored marketing strategies that resonate with specific audience groups. When combined with EDA, segmentation becomes more dynamic, enabling real-time adjustments based on evolving customer data.[7] Data-driven decision-making has gained prominence in the modern business landscape. Studies by Johnson and Harris (2018) underscore the role of segmentation and EDA in empowering organizations to make informed decisions. By dissecting customer data, businesses can predict purchasing behavior, identify pain points, and develop strategies that address customer needs proactively.[8] Understanding the psychological factors driving customer behavior is an emerging focus in segmentation studies. Research by Martin and Doyle (2020) explores how segmentation models can integrate psychological traits such as personality and motivation. EDA aids in quantifying and visualizing these traits, offering deeper insights into how psychological factors influence consumer decisions.[9] E-commerce platforms benefit significantly from customer segmentation and EDA. Patel et al. (2022) demonstrate that these techniques help online retailers optimize product recommendations, pricing strategies, and customer retention efforts. By leveraging large datasets, businesses can identify unique customer groups and craft personalized experiences that drive loyalty and sales. While segmentation offers numerous benefits, it also presents challenges.[10] Research by Thomas et al. (2019) illustrates their applications across sectors such as healthcare, finance, and retail. Each industry leverages these techniques to understand customer needs, optimize services, and improve operational efficiency.[11] The future of customer personality analysis lies in integrating advanced analytics with behavioral science. Smith and Rodriguez (2023) predict that future studies will focus on real-time data processing, predictive modeling, and ethical considerations. EDA and segmentation will continue to evolve, providing businesses with even greater insights into customer behavior.[12]

### Challenges

Customer personality analysis, leveraging segmentation and exploratory data analysis, encounters multiple challenges. Ensuring access to accurate, diverse, and current datasets is a significant hurdle, as data quality

directly influences the reliability of insights. Selecting appropriate segmentation techniques is equally critical, as poor choices can lead to ineffective categorization. Furthermore, safeguarding data privacy while complying with legal and ethical standards presents ongoing complexities. Another obstacle lies in translating intricate analytical results into clear, actionable insights for non-technical audiences. Finally, the computational requirements of processing large datasets can create scalability issues. Addressing these challenges is crucial for driving informed decisions and effective personalized marketing.

## METHOD

The methodology for analyzing customer personality through segmentation and exploratory data analysis involves a systematic process to collect, examine, and interpret data to uncover insights into consumer behavior and preferences. This approach begins with acquiring high-quality, relevant datasets from diverse sources, followed by cleaning and organizing the data to ensure accuracy. Segmentation techniques are then applied to categorize customers into distinct groups based on shared traits such as demographics, behaviors, or interests. Exploratory data analysis is used to identify patterns, trends, and anomalies within these segments. The resulting insights enable businesses to tailor strategies, enhance customer engagement, and drive informed decision-making.

### Dataset

Dataset of customer-credit-card-data has been obtained from kaggle. The source of dataset is https://www.kaggle.com/datasets/fhabibimoghaddam/customer-credit-card-data/ [13].

### Simulation

To conduct the simulation, a Python script is executed in a controlled laboratory environment to visualize the clusters formed using a customer dataset. This process involves leveraging Python's data analysis and visualization libraries to identify and display customer groupings based on specific attributes. The simulation provides a clear representation of the clusters, offering insights into patterns and relationships within the dataset. This visualization is a critical step in customer personality analysis, as it helps researchers and businesses better understand consumer behavior and preferences, ultimately aiding in the development of tailored strategies and personalized marketing efforts.
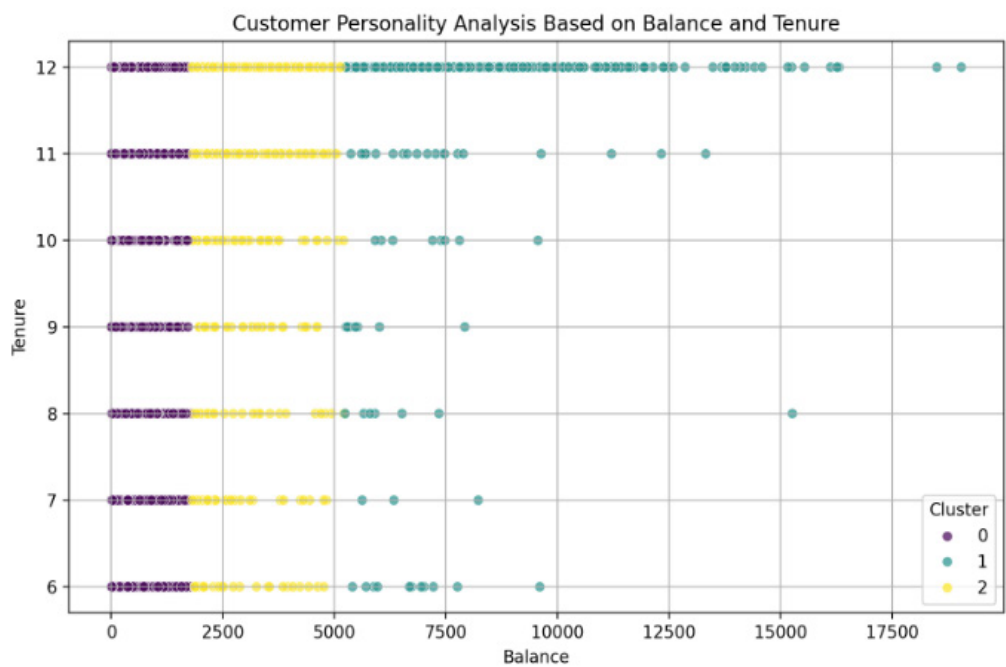


**Figure 1.** Customer personality analysis based on balance and tenure [13]

Understanding these clusters can help financial institutions tailor their marketing strategies, improve customer service, and develop products that meet the specific needs of different customer segments. It can also assist in identifying potential risks, such as customers who may be at risk of leaving or those who may require additional financial support. This analysis helps organizations better understand their customers' financial behaviors and preferences, enabling them to make data-driven decisions to enhance customer satisfaction and retention.

**Figure 2.** Customer personality analysis based on balance frequency and tenure[13]

Figure 2 refers to the examination of customer behavior and characteristics by analyzing two specific metrics: balance frequency and tenure. The scatter plot generated in the previous step visually represents these concepts by clustering customers based on their balance frequency and tenure. Each cluster represents a group of customers with similar behaviors, allowing for better understanding and strategic decision-making.



**Figure 3.** Customer personality analysis based on purchase frequency and tenure[13]

The figure 3 illustrating to the examination of customer behavior and characteristics by analyzing two specific metrics: purchase frequency and tenure. this type of analysis helps businesses better understand their customers' behaviors and preferences, enabling them to make informed decisions about marketing, product offerings, and customer service strategies.
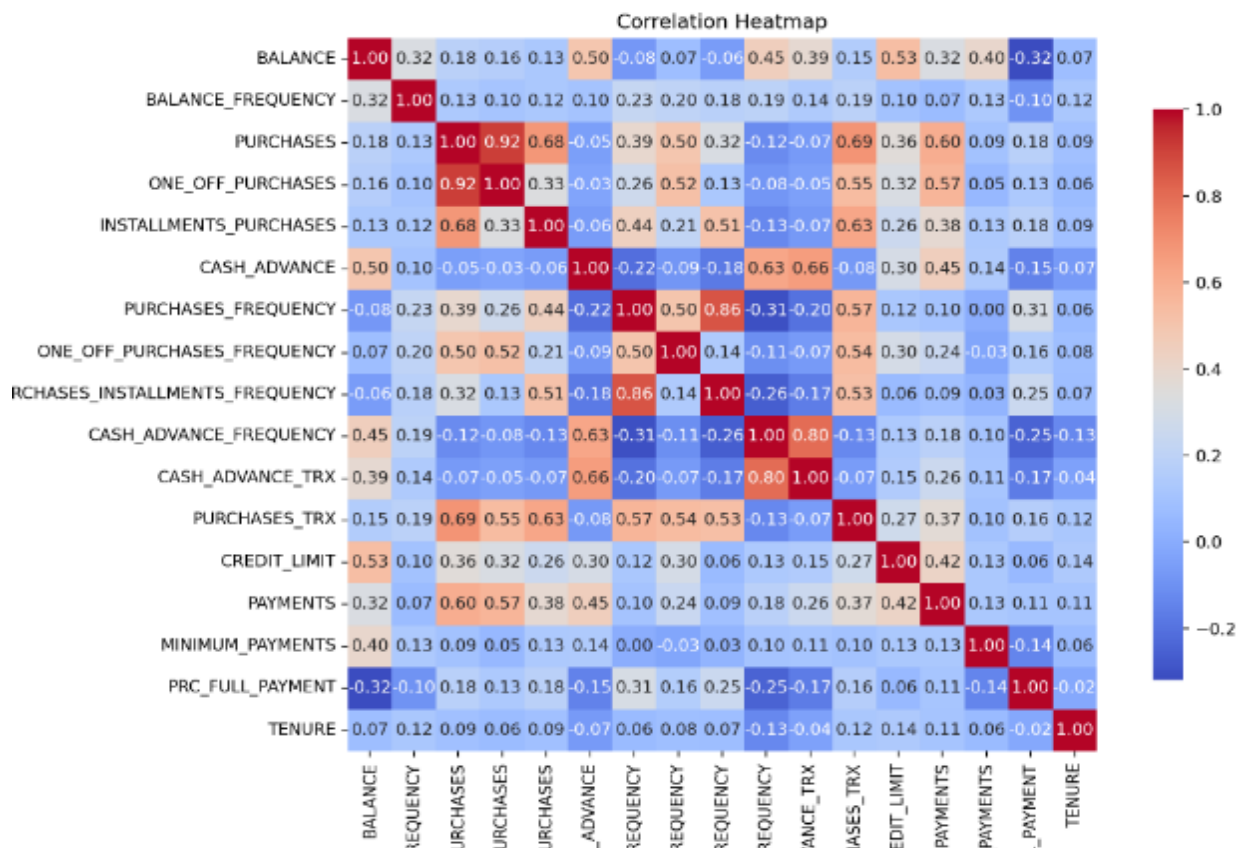
**Figure 4.** Correlation Heatmap[13]

A correlation heatmap is a graphical representation of the correlation coefficients between multiple variables in a dataset. It uses colors to represent the strength and direction of the relationships between pairs of variables.  Here are some key points about correlation heatmaps:

**Correlation Coefficient**
The correlation coefficient is a statistical measure that describes the extent to which two variables change together. It ranges from -1 to 1:
- A value of **1** indicates a perfect positive correlation (as one variable increases, the other also increases).
- A value of **-1** indicates a perfect negative correlation (as one variable increases, the other decreases).
- A value of **0** indicates no correlation (the variables do not affect each other).

**Heatmap Visualization**
In a heatmap, the correlation coefficients are displayed in a matrix format, where:
- Each cell in the matrix represents the correlation between two variables.
- The color of the cell indicates the strength and direction of the correlation (e.g., dark colors for strong correlations and light colors for weak correlations).

**Interpretation**
By examining the heatmap, you can quickly identify which variables are positively or negatively correlated and the strength of those correlations. This can help in understanding relationships in the data, guiding feature selection for modeling, and identifying potential multicollinearity issues.

**Applications**
Correlation heatmaps are commonly used in exploratory data analysis, especially in fields like finance, biology, and social sciences, to visualize relationships among multiple variables.

**CONCLUSIONS**
In conclusion, this study demonstrates the effectiveness of combining segmentation methodologies

with exploratory data analysis (EDA) to gain valuable insights into customer behavior, characteristics, and preferences. By systematically analyzing large datasets, the research highlights hidden patterns and uncovers distinct customer segments that businesses can leverage to better understand their audience. These insights pave the way for actionable strategies, enabling organizations to customize their products, services, and marketing efforts to align with the unique needs and expectations of diverse customer groups. This tailored approach not only fosters stronger connections with customers but also enhances their overall satisfaction and loyalty. Moreover, the findings underscore the importance of data-driven decision-making in today\u2019s competitive and technology-driven business environment. By adopting these techniques, businesses are well-equipped to anticipate changing consumer demands, optimize their operations, and achieve long-term sustainable growth. Ultimately, this research reinforces the critical role of customer personality analysis in empowering organizations to stay ahead in a rapidly evolving market while delivering personalized and impactful experiences to their customers.

## BIBLIOGRAPHIC REFERENCES

1. Alok J, Tiwari M. HR Aspects of Corporate Social Responsibility: A Comprehensive Review. Data Metadata [Internet]. 2025 Jan 1;4 SE-Or:343. Available from: https://doi.org/10.56294/dm2025343

2. Liu H, Huang Y, Wang Z, Liu K, Hu X, Wang W. Personality or value: A comparative study of psychographic segmentation based on an online review enhanced recommender system. Appl Sci. 2019;9(10):1992.

3. An J, Kwak H, Jung S gyo, Salminen J, Jansen BJ. Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data. Soc Netw Anal Min. 2018;8(1):54.

4. Wesley S, LeHew M, Woodside AG. Consumer decision-making styles and mall shopping behavior: Building theory using exploratory data analysis and the comparative method. J Bus Res. 2006;59(5):535–48.

5. Harris CR, Millman KJ, Van Der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357–62.

6. Brown HE. Design and Initial Analysis of a Wearable Device for Anxiety Management. 2021;

7. Lee JY, Johnson KKP. Cause-related marketing strategy types: assessing their relative effectiveness. J Fash Mark Manag An Int J. 2019;23(2):239–56.

8. Phung TTH. Listing Segmentation-A Research on Airbnb's Listing in Amsterdam. 2024;

9. Martin L, Delaney L, Doyle O. The Distributive Effects of Administrative Burdens on Decision-Making. J Behav Public Adm. 2023;6.

10. Joshi S, Iyer A, Patel S, Patel P. Leveraging Reinforcement Learning and Predictive Analytics for Optimizing AI-Enhanced Omnichannel Retail Strategies. Int J AI ML Innov. 2023;12(3).

11. Bharadiya JP, Thomas RK, Ahmed F. Rise of Artificial Intelligence in Business and Industry. J Eng Res Reports. 2023;25(3):85–103.

12. Robinson E. Predictive Modeling of Health Insurance Claims in the United States. Am J Stat Actuar Sci. 2024;5(1):35–46.

13. Customer-credit-card-data [Internet]. Available from: https://www.kaggle.com/datasets/fhabibimoghaddam/customer-credit-card-data/.

## CONFLICT OF INTEREST
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## AUTHORSHIP CONTRIBUTION

*Conceptualization:* Amitabh Chandan.
*Investigation:* Amitabh Chandan.
*Methodology:* Amitabh Chandan.
*Writing - original draft:* Amitabh Chandan.
*Writing - review and editing:* Amitabh Chandan.