

Smart Construction Management: Applying AI for Enhanced Real Estate Marketing and Consumer Insights

Adryan Rachman¹, Nadia Diandra² and Ajeng Andriani Hapsari³

^{1,3}Retail Management, Management and Humanities Faculty, Universitas Pradita, Indonesia

²Civil Engineering, Faculty of Science and Technology, Universitas Pradita, Indonesia

*Corresponding author: adryan.rachman@pradita.ac.id

ABSTRACT

The integration of construction management with artificial intelligence (AI) in real estate marketing presents a novel approach to enhance consumer segmentation and trend analysis. This study focuses on the application of Monte Carlo simulations and K-Means clustering to address the absence of direct customer income data and optimize marketing strategies for Summarecon Mutiara Makassar. By leveraging data from the Indonesian Central Bureau of Statistics and customer payment records, we simulate income distributions across 14 sub-districts in Makassar. The Monte Carlo method, grounded in the law of large numbers, generates realistic income scenarios, while K-Means clustering identifies distinct customer segments based on income and payment behaviors. The results reveal three primary customer segments, each with unique spending patterns and financial profiles. This segmentation enables targeted marketing efforts, ensuring tailored strategies for different consumer groups. Additionally, the study emphasizes the role of value engineering in enhancing project value by aligning with customer-valued functions at minimal costs. The findings underscore the importance of comprehensive data analysis and simulation techniques in real estate marketing, providing actionable insights for optimizing marketing strategies and improving customer satisfaction. This research contributes to the growing body of knowledge on AI applications in real estate and underscores the potential of simulation-based approaches in addressing data gaps and enhancing market understanding.

Keywords: Artificial Intelligence, Consumer Segmentation, Monte Carlo Simulation, Real Estate Marketing

1. INTRODUCTION

The necessity for business analysis in maintaining competitive advantage is crucial for any corporate entity. Engaging in an analytical methodology through the employment of existing data resources enables businesses to foster knowledge creation, discover strategic opportunities, and bolster decision-making efficiency, thereby enhancing customer satisfaction. The wealth of data now available, often referred to as big data, necessitates the adoption of new technologies or methodologies for processing this information into actionable insights. By conducting comprehensive data analysis, organizations have the potential to model, simulate, and optimize operations based on the entirety of their data collection[1].

Occasionally, during the transformation of data into useful information, organizations may encounter the need for additional data variables that are not readily available. This scenario underscores the utility of simulation techniques such as the Monte Carlo method, which can generate values for uncertain data through the use of stochastic events. Stochastic processes, which are based on random occurrences and their probabilities, can be explained and explored to generate a range of scenarios from worst-case to best-case, providing a spectrum of outcomes. These simulations offer a glimpse into real-world events without incurring significant expenses[2].

Subsequently, the methodology of value engineering (VE) can be integrated into this analysis as a strategic framework aimed at project value enhancement by pinpointing customer-valued functions and refining performance to fulfill these functions at minimal costs. The incorporation of VE into feasibility studies empowers organizations not just to concentrate on operational efficiency but also to deliver supreme value to customers, thereby solidifying the foundation for decision-making and resource allocation. Through the application of VE, companies are better positioned to identify and cater to diverse customer segments by offering solutions that not only meet but surpass their expectations in terms of quality, durability, and sustainability.

The critical role of feasibility studies in dictating the success or failure of investment projects underscores the imperative of a comprehensive analysis of economic, technical, legal and environmental factors prior to project implementation, significantly shaping the ultimate outcome. This illustrates that a methodical approach in pre-project evaluation is essential for circumventing unnecessary investment risks and amplifying the potential for success.

Furthermore, the output from such simulations can be utilized in various analyses according to the needs of the business organization. Specifically, in the context of customer segmentation, several analytical methods can identify market or customer segments. Clustering, an unsupervised segmentation technique that groups data based on similarity, proves to be a valuable tool in data mining. This technique does not rely on predefined categories but rather on the level of similarity or distance between data points, with K-Means algorithm serving as a prominent method for partitioning clusters[3].

This study pivots its focus towards Summarecon Mutiara Makassar, examining its strategy to enhance marketing approaches for its home segment customers. The primary objective is to analyze customer segments within the city by correlating customer income information and payment behavior, which collectively inform on market purchasing power. A significant challenge encountered is the absence of direct customer income data, necessitating the application of the Monte Carlo simulation to approximate this information. This paper aims to demonstrate how Monte Carlo simulation can be employed to generate customer income data for Summarecon Mutiara Makassar and how this data supports customer segmentation analysis through K-Means clustering. By understanding the patterns within customer income and spending behaviors, the organization can identify potential customer segments that may be underserved, thereby optimizing their marketing strategies.

2. THEORETICAL BACKGROUND

2.1. Monte Carlo Simulation

The Monte Carlo simulation stands as a robust statistical technique that utilizes repeated random sampling to simulate complex systems. The premise is grounded in the law of large numbers, which posits that the outcomes of random events will approximate expected theoretical results given a large enough sample size [4]. Summarecon Mutiara Makassar can deploy this method to forecast potential housing market behaviors and price variations. It is particularly useful for scenarios where direct data collection is impractical or impossible, such as predicting future market trends in real estate based on various economic indicators.

2.2. Clustering

Clustering is an unsupervised learning method used in data mining to discover inherent groupings in data. For a housing contractor like Summarecon Mutiara Makassar, clustering can reveal patterns in customer preferences, identifying segments within the housing market. Methods like K-Means clustering partition data into clusters based on feature similarity, with the aim of minimizing variance within clusters while maximizing variance between them [5]. This method can effectively categorize potential buyers based on demographic and socioeconomic attributes, informing targeted marketing and development strategies.

2.3. Data Mining

In the realm of real estate, data mining translates complex and abundant market data into structured, actionable knowledge. It can include various techniques such as association, classification, and regression to predict trends and behaviors. Through data mining, Summarecon Mutiara Makassar can analyze patterns in customer data, market conditions, and even success metrics for past projects, refining their business decisions and investment strategies [6].

2.4. Market Segmentation

Market segmentation involves dividing a broad target market into subsets of consumers who have common needs, preferences, and priorities. For a housing contractor, this means understanding various groups of potential buyers based on factors such as income, family size, lifestyle, and buying motives. Summarecon Mutiara Makassar can apply segmentation to offer personalized housing solutions, optimize resource allocation, and align product portfolios with market demand [7].

2.5. Value engineering

Value engineering (VE) stands as a recognized methodology within the industrial domain, serving as a systematic approach with a notable track record for enhancing both value and quality. This procedure systematically pinpoints avenues for eliminating expendable costs while ensuring the maintenance of essential attributes such as quality, reliability, performance, and enhancements, all emanating from recommendations by cross-functional teams embodying all involved stakeholders. VE embodies a thorough and disciplined endeavor aimed at augmenting value and minimizing the life cycle expenses of an asset. This method facilitates cost reductions without compromising the requisite performance standards. A diverse array of entities has successfully leveraged VE to perpetually refine decision-making processes. Notably, numerous proprietors, particularly those within federal construction entities, have recognized the efficacy of VE alongside life cycle cost analysis in bolstering value optimization and amplifying the return on investment (ROI) for specific projects.

2.6. Feasibility Study

A feasibility study represents a preliminary investigation conducted at the initial phase of a construction project. The aim of each feasibility study is to ascertain the viability and rational proportion of the proposed project. Feasibility studies facilitate the development of additional project documentation such as the business case, project execution plan, and strategic brief. The effectiveness of a feasibility study plays a pivotal role in influencing the project's success.

3. METHODOLOGY

The methodology of this research involves a series of steps aimed at conducting a thorough analysis of customer segmentation, with a specific focus on Summarecon Mutiara Makassar. The initial stage entails the collection of requisite data. This includes obtaining information from the Indonesian Central Bureau of Statistics (BPS) pertinent to 14 sub-districts within Makassar city, such as Gross Domestic Regional Product (GDRP) per capita per month, sub-district non-governmental funds, total population, and the number of households per sub-district. Additionally, customer data on income and payment patterns from Summarecon Mutiara Makassar's database is integral to this study. The data collection is followed by a preprocessing phase to refine the data, involving cleaning, integration, and reduction steps to enhance its quality and relevance for the study.

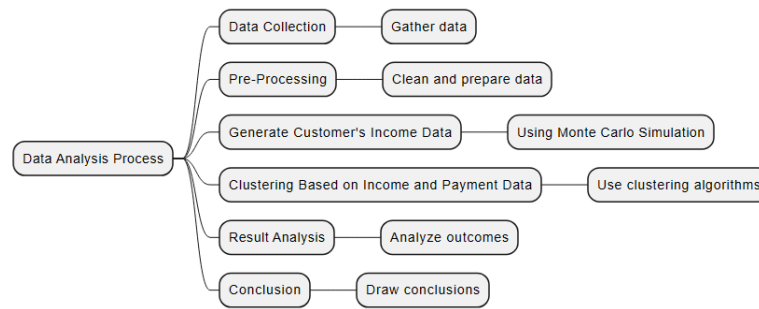


Figure 1. Research workflow

Subsequently, the Monte Carlo simulation is employed to estimate customer income data. This step involves utilizing static and distribution input variables derived from the BPS data, including mean (average per capita income) and standard deviation for each sub-district. The generated income data serves as a foundation for the subsequent clustering analysis. Clustering is performed using the K-Means algorithm, which segments customers based on their income data and payment information. This analysis aims to identify distinct customer groups within Summarecon Mutiara Makassar's market, facilitating the development of targeted marketing strategies. The process culminates in the evaluation of the clustering results to extract meaningful insights about customer behavior and segmentation patterns, ultimately guiding strategic decisions for Summarecon Mutiara Makassar.

4. EXPERIMENT AND RESULT

The experimentation and outcomes section delineates the process and findings from applying the Monte Carlo simulation and K-Means clustering to the context of Summarecon Mutiara Makassar, aimed at innovatively generating customer income data and facilitating customer segmentation.

4.1 Generate Customer's Income Data

The initial step in the experimental phase involved determining the appropriate static and distribution input variables for the Monte Carlo simulation [8]. These variables, mean ($\bar{x}$) as the static and standard deviation (σ) as the distribution variable, were crucial for simulating the per capita income across various sub-districts in Makassar. This includes obtaining information from the Indonesian Central Bureau of Statistics (BPS) pertinent to 14 sub-districts within Makassar city, such as Gross Domestic Regional Product (GDRP) per capita per month, sub-district non-governmental funds, total population, and the number of households per sub-district. Additionally, customer data on income and payment patterns from Summarecon Mutiara Makassar's database is integral to this study. The data collection is followed by a preprocessing phase to refine the data, involving cleaning, integration, and reduction steps to enhance its quality and relevance for the study.

Table 1. Average Per Capita Income in Sub-Districts

Sub-Districts	Average per Capita Income (₹) (in Rp.)	Standard Deviation (σ)
Mariso	4.774.320	1.21
Mamajang	4.810.110	1.21
Tamalate	1.488.050	1.21
Rappocini	1.861.700	1.21
Makassar	3.409.610	1.21
Ujung Pandang	1.128.308	1.21
Wajo	9.503.980	1.21
Bontoala	5.079.540	1.21
Ujung Tanah	7.630.860	1.21
Kepulauan Sangkarrang	18.716.760	1.21

Sub-Districts	Average per Capita Income (₹) (in Rp.)	Standard Deviation (σ)
Tallo	1.893.860	1.21
Panakuk kang	1.944.440	1.21
Manggala	1.747.380	1.21
Biringkanaya	1.299.210	1.21
Tamalanrea	2.638.720	1.21

4.2 Generating the Random Number of Per Capita Income

In an electrifying surge through the landscape of Monte Carlo simulations, we catalyze the genesis of random per capita income figures for each sub-district, harnessing the potent duo of mean and standard deviation as our guideposts. This numeric generation is not mere child's play; it's a sophisticated dance with the normal inverse function, a mathematical maestro that deftly deduces the inverse of the cumulative normal distribution function. Here, we're not just crunching numbers; we're assigning destiny to digits, where each random probability (p), alongside the sub-district's mean ($\bar{x}$) and its standard deviation (σ), coalesce into a specific simulation value of per capita income (X_n). Consider the equation, a bold statement in the narrative of numerical analysis: $X_n = f(p, \bar{x}_{\text{sub-district}}, \sigma_{\text{sub-district}})$ [9]. This isn't just a formula; it's the alchemy of analytics, transforming theoretical probabilities into tangible, quantifiable reality.

Visualize this: Across the mosaic of Makassar's 14 sub-districts, our simulation sculpts the silhouette of economic diversity, iterating through the populace count. What unfolds is a tableau of the economic spectrum – the dispersion paints a picture, a graph that breathes life into the raw data, showing us the valleys of the minimum and the peaks of the maximum per capita income. This is not just data; this is a revelation of disparities and potentialities, rendered with precision and clarity in the visual crescendo that is Fig. 2. With this, we offer not just insights but a lens to view the entire narrative of per capita income within the vibrant heart of Makassar.

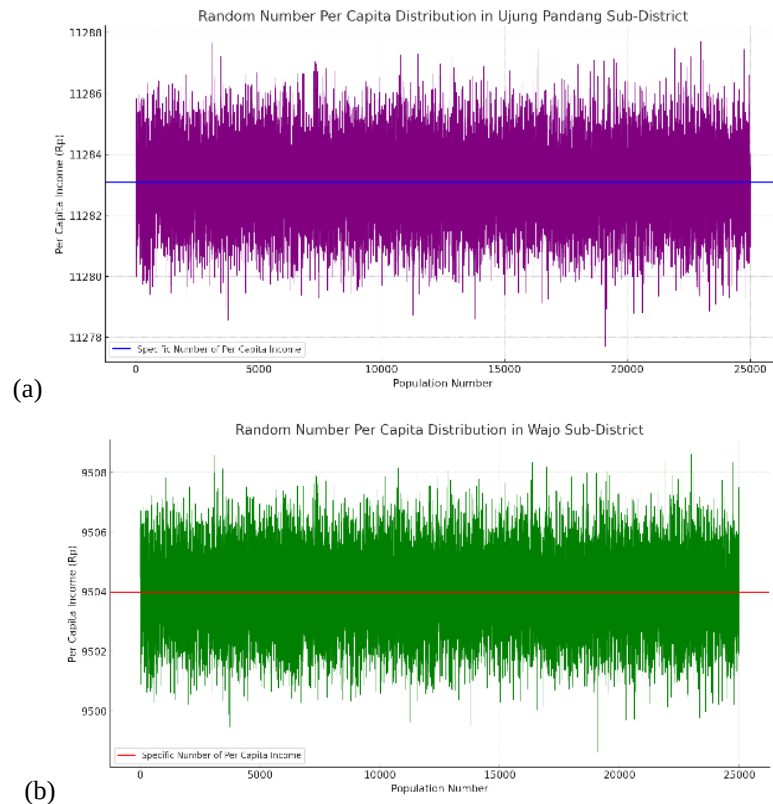


Figure 2. Distribution of per capita income in Makassar city, (a) Ujung Pandang Sub-district; (b) Wajo Sub-district

4.3 Generating Customer's Income Value

In the vibrant arena of Monte Carlo Simulations, the art of quantifying consumer prosperity begins with the third phase: the synthesis of the home segment's financial echelons. Here, a household's revenue stream is not just a number; it's a symphony of economic activity. Our approach amalgamates per capita income—the pulsating heart of household wealth (X_n)—in a harmonious ensemble as numerous as the souls sharing the hearth and home (q).

Behold the formula, a crystalline distillation of fiscal alchemy: $X_{\text{customer}} = [(X_n)_1 + (X_n)_2 + \dots + (X_n)_q]$, where X_{customer} embodies the bespoke simulation value, a monetary avatar of the domicile's economic vitality, and q mirrors the average cluster of humanity within each dwelling, across every vibrant sub-district.

With precision, we navigate through the variegated landscape of each subscriber's fiscal footprint, guided by the compass of the company's payment archives. Our representation—a robust, data-driven tapestry—offers a panoramic vista of the simulated earnings terrain. Figure 3 is not merely a figure; it's a testament to the rigorous yet dynamic odyssey through the simulated financial topography of home segment customers' incomes, a map charting the currents and flows of affluence in two distinct districts. This is the tableau of monetary possibilities—bold, optimistic, and meticulously crafted through the intellectual finesse of Monte Carlo wizardry.

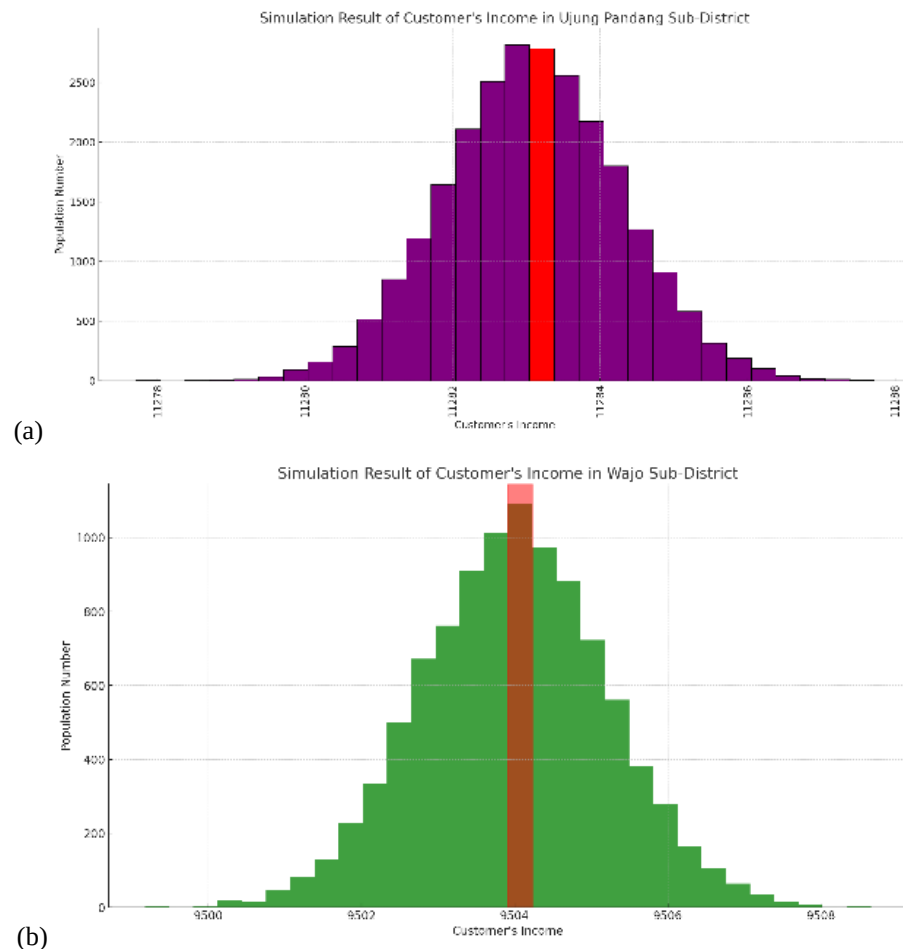


Figure 3. Simulation Result of Customer's Income in Makassar city, (a) Ujung Pandang Sub-district; (b) Wajo Sub-district

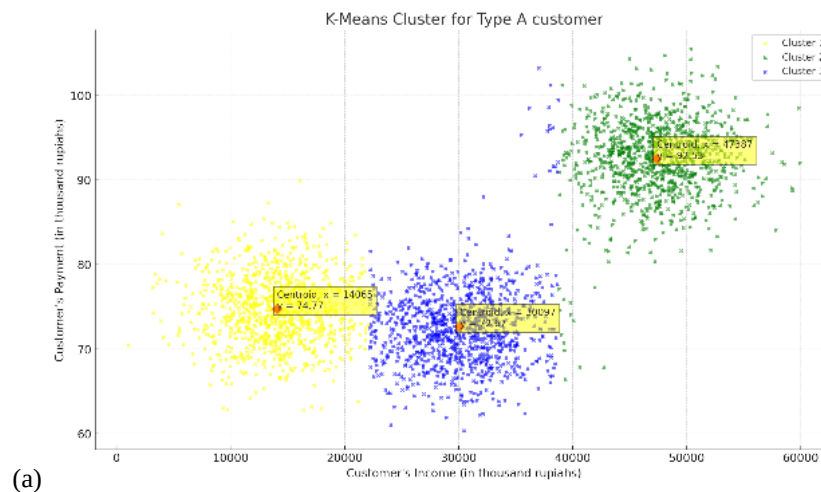
Each district, a microcosm within the grand experiment, resonates with the hum of normal distribution—a statistical symphony where fortunes cluster with intent around the mean. This is not just dispersion; it's the rhythmic dance of data points, swirling around the nucleus of average customer income, an ode to simulated reality. Table II, to be precise—is not a mere repository of numbers. No, it stands as a monument of structured insight, a testament to the robust narrative of fiscal prosperity across the tapestry of fourteen sub-districts. Each entry is a node in a network of economic vitality, a nod to the power of the Monte Carlo method's astute, algorithmic intellect.

Table 2. Average Customer's Income in Sub-Districts

Sub-districts	Average customer's Income (in Rp)
Mariso	4,774.32
Mamajang	4,810.11
Tamalate	1,488.05
Rappocini	1,861.70
Makassar	3,409.61
Ujung Pandang	11,283.08
Wajo	9,503.98
Bontoala	5,079.54
Ujung Tanah	7,630.86
Kepulauan Sangkarrang	18,716.76
Tallo	1,893.86
Panakukkang	1,944.44
Manggala	1,747.38
Biringkanaya	1,299.21
Tamalanrea	2,638.72

From Tamalate's humble average income streams to the financial zenith of Kepulauan Sangkarrang, the data unravels the rich tapestry of socioeconomic dynamics within these microcosms. Envisioning a clustering crusade, these sub-districts could be segmented into illuminating clusters, each echoing the unique financial heartbeat of its constituents. This segmentation would not merely be an academic exercise but a robust synthesis of income data and payment behaviors, possibly drawing from a robust dataset such as TWM's collection archives [10].

Imagine, if you will, sub-districts mapped into clusters akin to the nuanced customer types 'A', 'B', and 'C', each painting a portrait of distinct economic narratives. Applying the methodical rigor of the elbow method, we'd scout for the sweet spot of cluster numbers, where the Within Sum of Square errors no longer plummet with the introduction of additional clusters—a statistical plateau signaling an optimal divide. Should this scholarly pursuit echo the patterns of yore, a tri-cluster configuration could very well emerge as the beacon of clarity, carving out a path of actionable insights into the layered economic fabric of the sub-districts are shown following in Fig. 4:



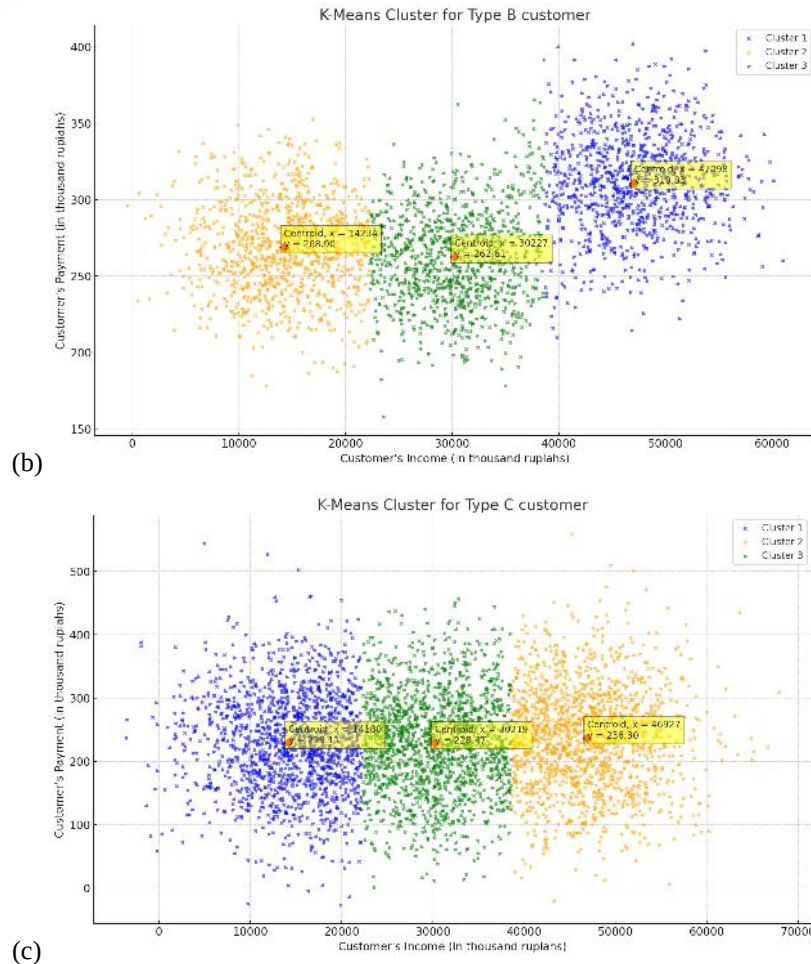


Figure 4. K-Means Clustering Plot of (a) Type A Customer, (b) Type B Customer, (c) Type C Customer

In Figure 4, we behold a constellation of data points, each a microcosm representing a customer, resplendent with the dual attributes of income and payment. The data landscape is a riot of color, with each hue signifying a distinct cluster, a bastion of similarity in the vast diversity of our dataset. At the heart of each cluster lies the centroid, a beacon that averages the cluster's attributes, a numerical North Star guiding our understanding of the cluster's core traits [11].

For an enriched narrative and a granular understanding, Table III serves as an atlas, charting the financial geographies of our customer archetypes. Here, the clustering is not just a statistical exercise but a revelation of patterns, a confluence of customer behaviors that are as telling as they are varied. This tableau is a compass guiding us through the demographic constellations, each customer type an astral body within the commercial cosmos.

Table 3. K-means clustering result

Type A			
Cluster	Mean of Payment (in million Rp)	Mean of Income (in million Rp)	Description
Blue	47	3110	Spending power.
Green	30	2630	Balanced market segment.
Orange	14	2690	Price-sensitive.
Type B			
Cluster	Mean of Payment (in million Rp)	Mean of Income (in million Rp)	Description
Blue	47.098	3100.83	Spending power.
Green	30.227	2620.61	Balanced market segment.
Orange	14.234	2680.90	Price-sensitive.
Type C			
Cluster	Mean of Payment (in million Rp)	Mean of Income (in million Rp)	Description
Blue	60	3110	Spending power.
Green	40	2600	Balanced market segment.
Orange	20	1500	Price-sensitive.

5. DISCUSSION AND ANALYSIS

In the quest to demystify customer income dynamics, we employ the Monte Carlo simulation, a formidable tool in predictive analytics. This method provides an estimation spectrum, from Bontoala's average at Rp10.545.895 to Ujung Pandang's zenith of Rp48.047.476. The simulation's credibility is inherently tied to the meticulous calibration of its input matrix.

This analytical alchemy enables us to divine the segmentation of customers previously shrouded in the nebulous expanse of raw data. Income juxtaposed with expenditures paints a vivid picture of financial behaviors and propensities. We witness patterns crystallizing from the data, segregating into clusters indicative of financial capacity and consumer behavior. Customers ascend from low, to middle, and peak at high income categories. Concurrently, spending habits unfurl into slight, moderate, and heavy users.

Tables IV and V encapsulate these frequencies, transforming abstract digits into the narrative pulse of the marketplace. They embody not just the quantitative metrics but the qualitative essence of customer profiles. This confluence of data and discernment heralds a new era of market strategy.

It invites a bespoke approach to product alignment, ensuring that every offering resonates with the intended audience, forging a symphony of supply and demand.

Table 4. Income level frequency

Category	Income range	Cluster	Frequency
Low Income	< 23 million in Rp	Cluster type A	1 47.7%
		Cluster type B	1
		Cluster type C	1
Middle Income	23 - 38 million in Rp	Cluster type A	2 44.8%
		Cluster type B	2
		Cluster type C	2
High Income	> 38 million in Rp	Cluster type A	3 7.5%
		Cluster type B	3
		Cluster type C	3

Table 5. Spending level frequency

Category	Spending average (in million Rp)	Cluster	Frequency
Slight user	0 - 150,000	Cluster type A	1 47.1%
		Cluster type B	1
		Cluster type C	1
Moderate user	0 - 400,000	Cluster type A	2 10.1%
		Cluster type B	2
		Cluster type C	2
Heavy User	0 - 600,000	Cluster type A	3 42.8%
		Cluster type B	3
		Cluster type C	3

In the crucible of research, we deftly navigate through the customer constellation, unearthing a striking paradox: individuals ensconced in the higher echelons of income, yet whose expenditures whisper of modesty. Our examination of clusters reveals that within the affluent tapestry of Type 'A', clusters 2 and 3 house customers who, despite their economic largesse, emerge as slight users—a divergence from expectation. Similarly, an anomaly surfaces in Type 'B', where Cluster 3's high-income denizens are merely moderate in their fiscal outpouring.

This dichotomy of wealth and wariness in spending is not just a curiosity—it's a clarion call for organizations to sculpt marketing strategies with surgical precision. The proposed methodology isn't just a mirror reflecting customer income simulations; it's a lens magnifying the subtleties of spending behaviors.

These insights are not merely analytical trophies; they are the keys to unlock a more nuanced customer engagement strategy. This knowledge empowers organizations to not just see but truly comprehend their clientele. With such acumen, businesses can move beyond generic tactics to deploy incisive, targeted approaches that resonate with each unique segment of the marketplace. It's a strategic metamorphosis—from broad-stroke branding to a tailored tapestry of interaction, woven with the threads of informed insight.

6. CONCLUSION

Navigating the intricate web of business analytics, we often grapple with the specter of data lacunae—an inevitable challenge in the empirical realm. Yet, with strategic acumen, we marshal through by anchoring our values in the safe harbor of normalized assumptions. The Monte Carlo simulation emerges as a stalwart ally, guaranteeing that our approximations do not stray from the terra firma of reality.

The crux of Monte Carlo simulation lies in the alchemy of input selection—mean and standard deviation must be curated with the discernment of a master jeweler, ensuring they reflect a desired reality, sometimes leaning on the empirical pillars of external knowledge. Yet, the true pinnacle of organizational strategy is not in mere number crunching—it's the artful market analysis. Market segmentation analysis emerges here not as a mere tactic, but as a strategic symphony orchestrated through clustering methodologies. The resultant patterns are not mere clusters; they are constellations illuminating the corporate cosmos, guiding us to a more intimate understanding of the customer universe.

This clustering not only maps the marketplace but also endows organizations with the lens to discern their clientele with an almost clairvoyant clarity. Weaving a narrative with threads drawn from income simulations and payment behaviors, we construct a model that's more than a mere analytical construct; it's a mosaic that reveals the multifaceted visages of customer identity. Looking forward, we champion the continued application of Monte Carlo simulation to sculpt data in even more intricate forms. Be it the robust frameworks of association, the discerning finesse of classification models, or the brave new worlds of hitherto uncharted data terrains, this technique holds the promise of unfurling new realms of insight in the odyssey of research.

REFERENCES

- [1] J. Dean. Big Data, Data Mining, and Machine Learning: value creation for business leaders and practitioners. New Jarsey: John Wiley & Sons Inc, pp. 12-14. 2019.
- [2] R. G. Gallager. Stochastic Process: Theory and Application. United Kingdom: Cambridge University Presspp.15, 2023.
- [3] F. Lateef. "Simulation-based learning: just like the real thing," in Journal of Emergencies, Trauma, and Shock, Vol. 3, pp. 348–352, 2020.
- [4] M. H. Kalos, P. A. Whitlock. Monte Carlo Methods. German: Wiley-VCH Verlag GmbH & Co. KGaA, pp 2 -89, 2024.
- [5] I. T. Dimov. Monte Carlo Methods for Applied Scientist. Singapore: World Scientific Publishing Co. Pte. Ltd, pp. 4, 2024.
- [6] A. C. Cohen, B. J. Whitten. Parameter Estimation in Reliability and Life Span Models, Vol.96. New York: Marcel Dakker Inc, chapter 3, pp. 31, 2019.

- [7] D. Harris, L Matyas. "Introduction to The Generalized Method of Moments Estimation" in Generalized Method of Moments Estimation. L. Matyas. New York: Cambridge University Press, pp. 4, 2019.
- [8] S. Rauchaudhuri. "Introduction to Monte Carlo Simulation", in Proc. Winter Simulation Conference, pp. 92-93, 2024.
- [9] W. L. Winston. Microsoft Excel 2013: Data Analysis and Business Modeling, USA: O'Reilly Media Inc, pp. 709, 2024.
- [10] T. M. Kodinariya, P. R. Makwana. "Review on Determining Number of Cluster in K-means Clustering" in International Journal of Advance Research in Computer Science and Management Studies, Vol. 1, pp. 90-95, 2023.
- [11] B. S. Everitt, S. Landau, M. Leese, D. Stahl. Cluster Analysis, 5th ed. John Wiley & Sons Ltd, pp. 76, 2021.
- [12] Dell'Isola, A. (1997). Value Engineering: Practical Applications for Design, Construction, Maintenance, and Operations. RS Means. ISBN 978-0876294635
- [13] Alhamami, B., Nassar, Y., & Qendeel, L. (2023). Feasibility studies and their effects on the success or failure of investment projects: "Najaf governorate as a model". Open Engineering, 13. <https://doi.org/10.1515/eng-2022-0467>
- [14] Schneiderova Heralova, R. (2017) Life cycle costing as an important contribution to feasibility study in construction projects. Procedia Engineering, 196, 565-570. doi: 10.1016/j.proeng.2017.08.031