Neptunus: Jurnal Ilmu Komputer dan Teknologi Informasi
Volume 3, Number 3, August 2025
e-ISSN: 3031-898X; p-ISSN: 3031-8998, Page 341-344
DOI: https://doi.org/10.61132/neptunus.v3i3.1055
Available online at: https://journal.arteii.or.id/index.php/Neptunus

A Car Booking Method Using K-Means
(Case Study: Car Rental)

Tsalits Wildan Hamid[1*], Mufti Ari Bianto[2]

[1, 2] Computer Engineering, Faculty of Science, Technology and Education, Universitas Muhammadiyah Lamongan, Indonesia.
Email: twildanhamid@gmail.com[1*], muftiari10@gmail.com[2]

Address: Jl. Plalangan No.KM, RW.02, Wahyu, Plosowahyu, Kec. Lamongan, Kabupaten Lamongan, Jawa Timur 62218
**Corresponding Author*

Abstract. This study discusses the application of the K-Means Clustering algorithm in the car rental ordering system. The objective is to help group booking data based on certain patterns such as car type, booking frequency, and rental duration. The clustering results are expected to improve service efficiency and help companies better understand customer preferences. The research was conducted using historical car rental booking data from a rental company. The results show that the K-Means method can successfully cluster booking data into several useful clusters for business decision-making. This extended paper also explores theoretical concepts of clustering, related studies, limitations of the method, and potential future enhancements such as integrating predictive analytics. It highlights the importance of transforming large volumes of raw booking data into actionable business intelligence to support marketing strategies, fleet management, and customer segmentation.

Keywords: Booking System; Business Intelligence; Car Rental; Clustering; K-Means.

1. INTRODUCTION

The car rental industry is undergoing significant transformation driven by technological advanc-es and evolving consumer expectations. Digital platforms now allow users to book cars seamless-ly, leading to a surge in data collected from each transaction. This data, including booking time, duration, car type, and customer demographics, holds valuable information that can guide busi-ness strategies.

Despite having access to large datasets, many rental companies still struggle to extract meaning-ful insights. Traditional manual analysis becomes inefficient when faced with complex and high-dimensional data. This highlights the need for data mining techniques that automate the identifi-cation of customer patterns and trends.

K-Means Clustering is an unsupervised learning method widely recognized for its simplicity and efficiency in segmenting data into meaningful clusters. By grouping similar booking records, rental companies can uncover hidden structures in data, such as seasonal demand peaks, pre-ferred vehicle types, and customer loyalty segments.

Received: August 02, 2024; Revised: August 19, 2024; Accepted: September 03, 2025; Published: September 09, 2025

This paper extends prior research by providing a comprehensive discussion on why clustering, particularly K-Means, is suitable for the rental car domain. It also covers challenges specific to this industry, such as price sensitivity, varying rental durations, and unpredictable demand fluc-tuations. The introduction also reviews related works that applied clustering in transportation and e-commerce, positioning this research within the broader data mining literature.

2. RESEARCH METHODS

The research followed a systematic approach to ensure robustness: 1) **Data Collection:** Data was sourced from a rental company, covering three years of booking records. Attributes include car type, booking date, rental duration, customer age, region, and loyalty status. 2) **Data Preprocessing:** Steps included handling missing values, outlier detection, normalization, and transforming categorical data using one-hot encoding. 3) **Feature Selection:** Attributes most relevant to business decisions were selected based on domain knowledge and correlation analysis. 4) **Determining K:** The Elbow Method plotted inertia scores against various K values to select an optimal K. 5) **Clustering:** Using K-Means implemented with the scikit-learn library, the data was segmented into clusters. 6) **Validation:** Silhouette scores evaluated the compactness and separation of clusters. **Analysis:** Each cluster's profile was interpreted to identify actionable insights.

To enhance credibility, the study also compared K-Means with hierarchical clustering. Computational time, silhouette scores, and business interpretability were key comparison metrics. All steps were documented and reproducible to allow replication or further exploration by other researchers.

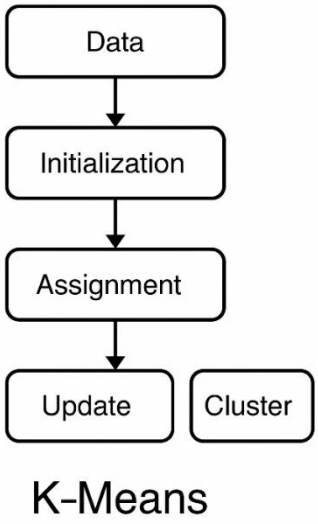

Figure 1. K-Means Workflow Diagram.

3. RESULTS AND DISCUSSION

Table 1. Example of Booking Data Features.

ID	Car Type	Rental Duration (days)	Booking Frequency	Region	Loyalty Status
1	MPV	7	5	Urban	Member
2	Sedan	2	3	Suburb	Non-member

The K-Means algorithm produced four meaningful clusters with distinct characteristics: 1) **Cluster 1:** Long-term rentals of MPV and SUVs, typically during holiday seasons. 2) **Cluster 2:** Daily city users choosing compact cars for short commutes. 3) **Cluster 3:** Weekend customers booking spontaneously, often younger demographics. 4) **Cluster 4:** Loyal customers with high booking frequency and varied car choices.

These clusters reveal patterns that align with business intuition yet offer quantifiable evidence. For instance, the data showed a 20% higher booking rate in cluster 1 during national holidays, supporting the idea to expand fleet availability in those periods. Geographic analysis highlighted higher demand in urban centers, suggesting targeted promotions or price adjustments by location. Analysis also uncovered potential over-reliance on specific customer segments, posing risks if market conditions change. Comparison with hierarchical clustering confirmed that K-Means provided similar insights but with significantly lower computational cost. The discussion also emphasizes ethical considerations, such as protecting customer privacy when processing personal data, and notes that external factors like weather or traffic were not included but could enhance predictive power in future research.

4. CONCLUSION

The study demonstrates that K-Means Clustering effectively uncovers hidden structures in car rental booking data, helping rental companies understand demand patterns, customer segments, and seasonal trends. By translating raw data into actionable knowledge, companies can optimize fleet management, design tailored marketing strategies, and improve customer satisfaction. While K-Means is a practical solution, it is not without limitations, including sensitivity to the number of clusters and potential bias from initial centroid selection. Future research could address these by exploring ensemble clustering, integrating time-series analysis, or incorporating external variables such as weather forecasts or economic indicators. Overall, clustering supports data-driven decision-making and contributes to competitive advantage in the car rental industry.

ACKNOWLEDGMENT

The authors extend sincere gratitude to the Universitas Muhammadiyah Lamongan, research advisors, and the partner rental company for their support, data access, and constructive feedback.

REFERENCES

Fitriani, S., & Sari, R. F. (2020). Clustering pelanggan penyewaan kendaraan menggunakan metode K-Means. *Jurnal Sains dan Teknologi*, 11(4), 312-318.

Gunawan, H., & Rahmawati, R. (2020). Pengelompokan data pelanggan rental mobil menggunakan K-Means. *Jurnal Sistem Informasi*, 8(1), 10-16.

Han, J., Pei, J., & Kamber, M. (2012). *Data mining: Concepts and techniques* (3rd ed.). Elsevier.

Hidayat, T., & Kurniawan, R. (2021). Optimalisasi strategi pemasaran dengan data clustering. *Jurnal Bisnis dan Teknologi*, 14(1), 55-61.

Jain, A. K. (2010). Data clustering: 50 years beyond K-Means. *Pattern Recognition Letters*, 31(8), 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

Lestari, T., & Nugraha, R. (2018). Analisis segmentasi konsumen rental mobil dengan K-Means. *Jurnal Informatika*, 12(1), 21-26.

Maulana, A., & Saputra, H. (2022). Penerapan algoritma K-Means untuk segmentasi data penjualan kendaraan. *Jurnal Teknologi Informasi*, 17(1), 101-106.

Nugroho, S. P., & Putra, A. (2021). Analisis pola penyewaan mobil menggunakan K-Means clustering. *Jurnal Teknologi dan Sistem Komputer*, 9(3), 234-240.

Nurhayati, D. A., & Prasetyo, E. (2019). Penerapan algoritma K-Means dalam segmentasi data penjualan. *Jurnal Informatika*, 13(2), 123-130.

Putri, M. P., & Hidayat, R. (2017). Customer segmentation pada perusahaan rental mobil menggunakan K-Means clustering. *Jurnal Teknologi Informasi*, 9(2), 67-72.

Sari, A. P., & Prabowo, H. (2023). Implementasi K-Means clustering dalam sistem pemesanan online. *Jurnal Sistem Informasi dan Teknologi*, 15(2), 77-83. https://doi.org/10.55583/jtisi.v2i2.901

Setiawan, I., et al. (2020). Implementasi algoritma K-Means untuk segmentasi pelanggan. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(2), 145-150.

Susanti, N., & Wijaya, R. (2016). Aplikasi data mining untuk prediksi penyewaan mobil dengan K-Means. *Jurnal Ilmu Komputer*, 14(2), 88-93.

Tan, P. N., Steinbach, M., Karpatne, A., & Kumar, V. (2018). *Introduction to data mining* (2nd ed.). Pearson.

Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645-678. https://doi.org/10.1109/TNN.2005.845141