International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

Enhancing Car Showroom Profitability through K-Means Clustering for Customer Segmentation

Muhammad Rizky Satya Bhaskara[1], Meditya Wasesa[2]
[1,2] Master of Business Administration, School of Business and Management Bandung Institute of Technology, Bandung, Indonesia, 40132

ABSTRACT: Indonesia is known to be one of Southeast Asia's largest markets for used cars. Used car showrooms in Indonesia are numerous and varied, some focus on one car brand, some focus on the lower middle class, some focus on the upper middle class, and some provide all types of cars. One of them is PQR Showroom, Even though PQR Showroom was able to generate such a great amount of sales, the profits generated by the PQR showroom are not proportional to the amount of capital invested.
To increase the profit, we proposed the solution using descriptive analytics and prescriptive analytics using K-Means. We also carry out simulations by comparing sales in existing years with the results of the descriptive and prescriptive analytics that have been made for the expected profit.
The results show that the simulation comparison with the data we have obtained from descriptive and prescriptive analysis gives the best-expected profit compared to the initial sales results at the PQR Showroom.
It shows that the data using descriptive and K-Means is great than before. The fastest cars that have been sold is LCGC and the most wanted car is MPY Toyota Avanza and the best profit that can generate is SUV Toyota Fortuner.

KEYWORDS: Descriptive Analytics, K-Means, Prescriptive Analytics, Sales, Used cars.

INTRODUCTION

Indonesia is known to be one of Southeast Asia's largest markets for used cars. According to Mendiratta (2021), the Indonesian used car industry is estimated to be worth USD 50.77 billion in 2021. This phenomenon happens because many people in the country prefer to buy used cars because they are less expensive. Many middle-class people want to own a car but cannot manage to buy a new one. Furthermore, many used cars on the market are relatively new, between three and seven years old, and in decent condition, making them highly feasible choices to look into (Mordor Intelligence, 2021). According to Mordor Intelligence (2021), the used car market in Indonesia is expected to rise to USD 74.47 billion by 2028, at a CAGR of around 5.74% throughout the forecast period (2023–2028). According to OLX Autos Indonesia (2021), there is a higher purchase intentation in Q4 2021, 46% of customers intend to purchase right now, compared to 22% at the same time last year. Used car showrooms in Indonesia are numerous and varied, some focus on one car brand, some focus on the lower middle class, some focus on the upper middle class, and some provide all types of cars. One that provide all types of cars i.e. PQR Showroom (pseudonym), is a local used car dealer which was established in 2010 in Bukittinggi, West Sumatera. PQR Showroom provides used cars ranging from city cars, family cars, and minivans from various car brands. PQR Showroom needs to take advantage of this economic improvement and enhance its performance. Even though according to Mordor Intelligence (2021), the used car market in Indonesia can continue to grow to reach USD 74.47 billion in 2028, PQR Showroom does not feel that way, the huge gap between sales and profit is a problem for the owener of PQR Showroom, sales always reach billions of rupiah per year but the profit earned by PQR Showroom is only hundreds of millions per year. According to Figure 1, the sales and profit of the PQR Showroom vary each year. From 2010 untill March 2023, the PQR Showroom generated sales of more than IDR 77 billion, with the highest sales in 2018 reaching IDR 10.238.000. But, the highest profit earned was in 2019 which amounts to IDR 592.000.000. Even though PQR Showroom was able to generate such a great amount of sales, the profits generated by PQR Showroom are not proportional to the amount of capital invested. In some certain years, the amount of profit generated is greater than the amount of profit generated in another year with more capital invested. For instance, in 2014, PQR Showroom was able to obtain a profit of as much as Rp. 292.000.000 by investing Rp. 5.685.000. Meanwhile, the showroom gained less in 2013, by investing 6.292.000, PQR Showroom only obtained Rp. 274.000.000 as their profit. Unproportional capital and profit relationships

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

also happened in other years. In 2016, PQR Showroom succeded gain a profit as much as Rp 405.000.000 by investing 5.691.000.000. On the other hand, PQR Showroom only generated Rp498.000.000 in 2015 by almost twice as much as it invested in 2016.

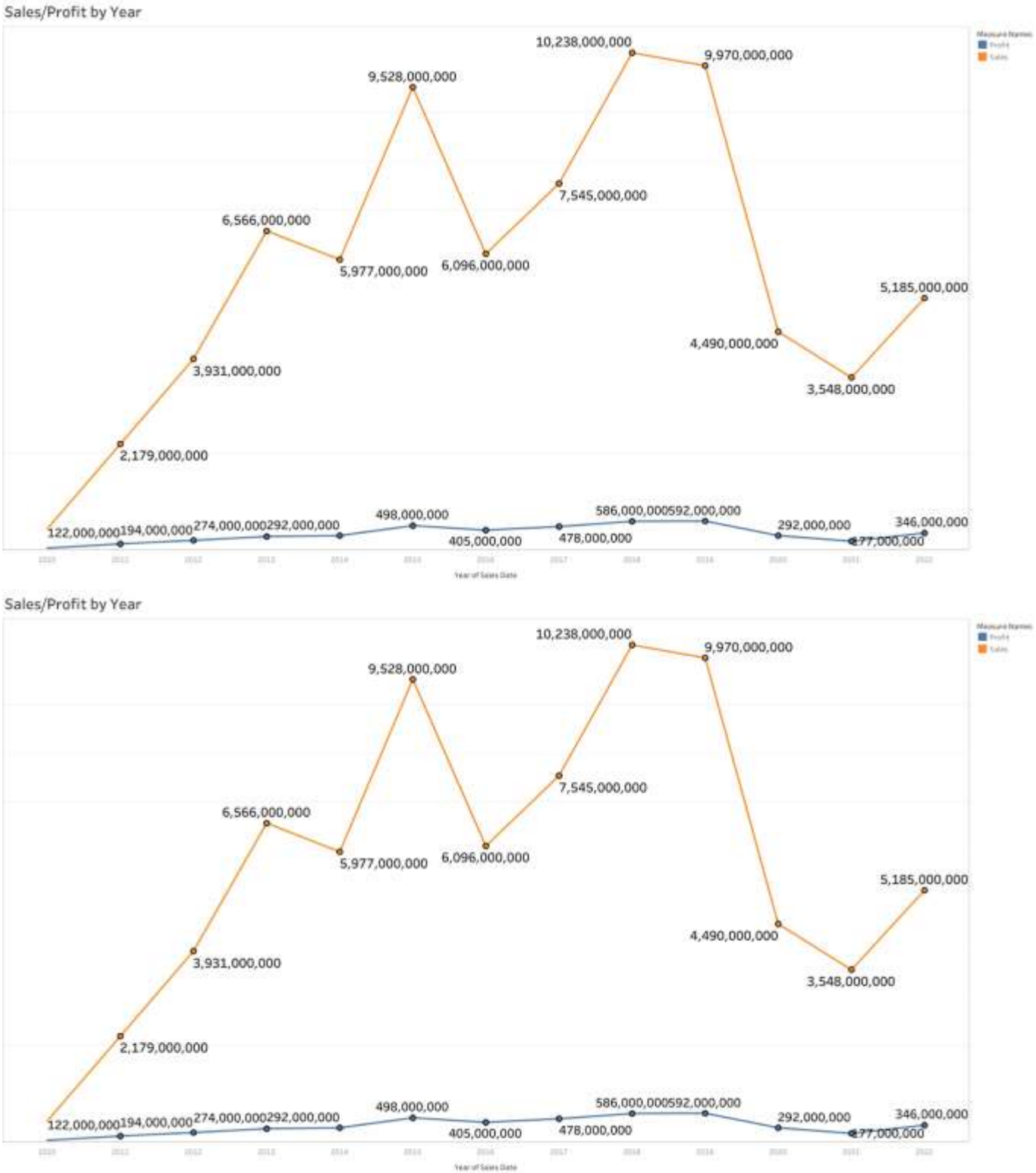

Figure 1: Sales and Profit of PQR Showroom

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

LITERATURE REVIEW

1. Business Analytics

There is currently no universal definition of business analytics. Experts in many domains have defined the word business analytics from various perspectives. According to Liu et al. (2023), business analytics can be defined into four categories: techniques, processes, practice, and management. First, from the perspectives of techniques, business analytics is defined as an application of data analytics (Duan & Xiong, 2015) or data science (Chen et al., 2012) in business fields that employ statistical and quantitative tools and techniques to analyze a large collection of data sources to support business decisions (Delen & Zolbanin, 2018). Second, from the processes perspective, business analytics is a collection of tools used to transform data into actionable insights via a mathematical, scientific, or intelligent process (Delen & Ram, 2018). Third, from the practical perspective, business analytics is described as a company's or organization's ability to collect, organize, and analyze data from various sources in order to improve understanding of business processes, operations, and systems (Kraus et al., 2019). The last one, according to the management perspective, business analytics is "a paradigm shifter of models, technologies, opportunities, and capabilities used to scrutinize a corporation's data and performance in order to transpire data-driven decision-making analytics for the company's future decision and investment plans" (Bayrak, 2015).

2. Descriptive Analytics

Descriptive analysis is a process of characterizing historical data. There are two main methods of descriptive analytics which are data analysis and data visualization. Data analysis commonly used statistical methods include mean, median, standard deviation, range, etc. to reveal hidden patterns in the data. Data visualization generates graphical images of data which helps decision-making (Ware, 2004). There are several tools that can be used to conduct descriptive analytics such as Tableau Software, Microsoft Power BI, Excel, FusionCharts, Sisense, etc. All of them are known to have the ability to provide visual representations of massive data sets. Data visualization is essential in any data-driven business. It converts data into visualizations that are easier to understand and to be used for making critical business decisions. Data visualization generates actionable insights that might have not been identified (Liu et al., 2023).

3. Prescriptive Analytics

Prescriptive analytics mainly refers to the utilization of operations research techniques such as mathematical programming models and intelligent optimization algorithms to provide recommendations on the optimal actions that a business should take (Liu et al., 2023). According to Lepenioti et al (2020), prescriptive analytics can be classified into six, namely probabilistic models, machine learning/data mining, mathematical programming, evolutionary computation, simulation, and logic-based models.

4. K-Means Clustering Analysis

Cluster Analysis is an unsupervised learning for object or data mining. K-means clustering analysis is a non-hierarchical cluster analysis method that seeks to partition an existing object into one or more clusters of objects based on their characteristics so that objects which have the same characteristics are grouped in the same cluster and objects that have different characteristics are grouped into another cluster (Ediyanto et al., 2013).

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

RESEARCH METHODOLOGY

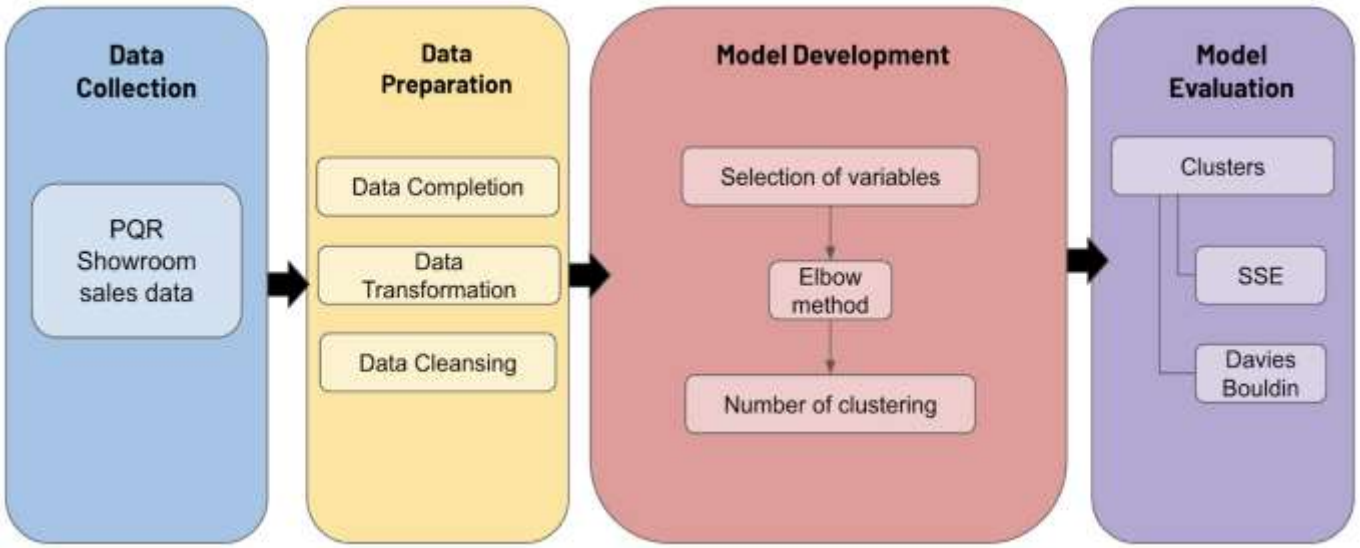

Figure 2: Research Design

The data collected for this research is PQR Showroom sales data from 2010 – 2023 (March). Preparation of the data starts by completing the data with some variables which include transmission, fuel type, cc, days, and age of the car. After that data transformation is carried out to transform categorical data (polynomials or binomials) to numerical data since the k-means algorithm can only process numerical data. After that data cleansing which includes fixing structural errors (typos, and inconsistent capitalization, checking for missing values, and removing duplicate data) is conducted. Two models are used in this research, which are descriptive analytics and prescriptive analytics. Model development of prescriptive analytics (k-means clustering), starts with the selection of variables. The selection of variables is based on the objective being set (for example: to find out the fastest sold car and its profit, thus variable days, car_type, and profit_per_car are used for k-means clustering analysis). After choosing the variables, the number of optimum clusters is determined by employing the elbow method. The best number of clusters is determined by assessing a graph of comparison between the number of clusters and its SSE (Madhulatha, as cited in Merliana et al., 2015) The point which forms an elbow in the graph will show the best number of clusters. After conducting the clustering, the model is evaluated using SSE and Davies Bouldin. SSE is defined as the sum of each point's squared Euclidean distance to its nearest centroid. Because this is a measure of error, the goal of k-means is to strive to minimize it (Arvai, 2020). Davies Bouldin is a metric used to assess clustering performance. Its primary idea is to analyze the distance between two clusters, which should be as large as feasible between clusters and as little as possible within a cluster. A lower one will mean that the clustering is better. (Xiao et al., 2017).

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

RESULTS

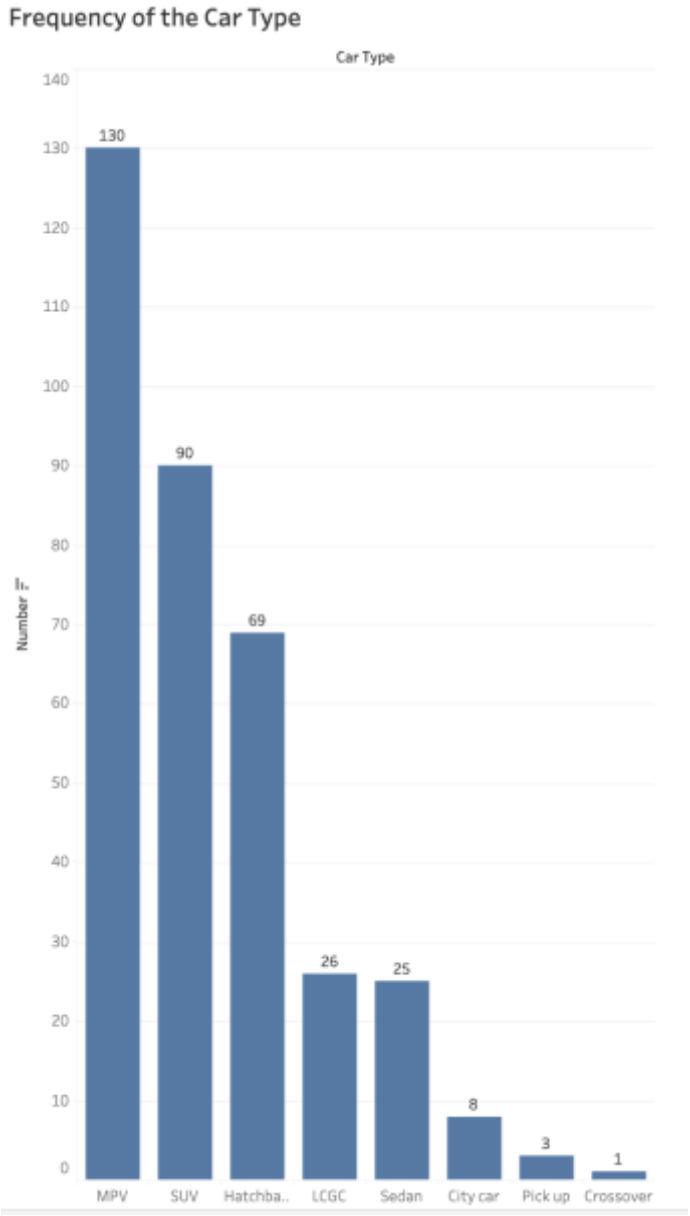

Figure 3: Frequency of the Car Type

According to Figure 3, MPV, SUV, and Hatchback are the three most sought-after car types and crossover is the least that customers bought in PQR Showroom. Hatchbacks came in third with sales of 69 cars, SUVs came in second with sales of 90 cars, and MPV cars were the most purchased by customers, with sales of 130 cars. By knowing the most purchased car types at the PQR Showroom, we can focus on that type of car and narrow down to determine which car brand is most in demand in the MPV segment.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

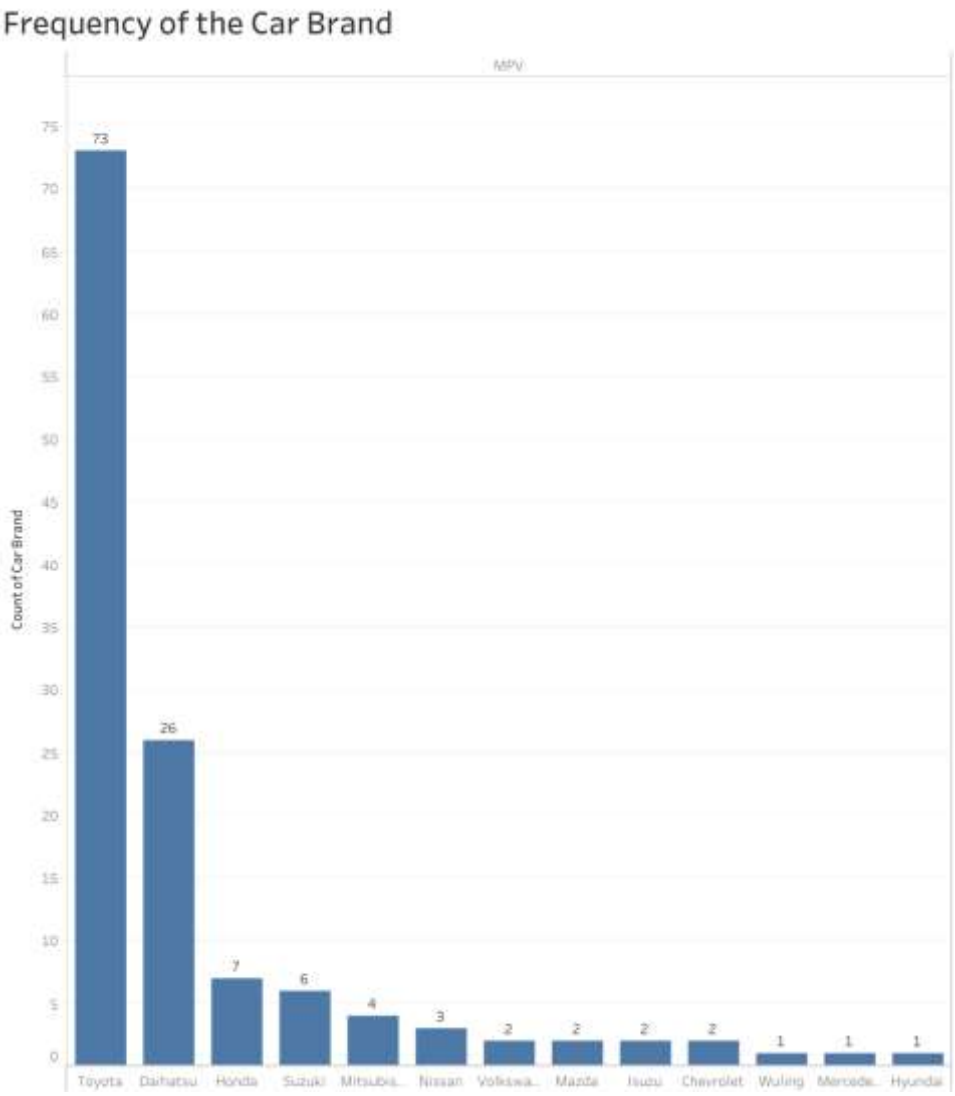

Figure 4: Frequency of the Car Brand

According to Figure 4, after being narrowed down from the type of car, it is known that in MPV segment, Hyundai is the least sold car brand which was only sold once, and the most sold brand is Toyota with 73 cars or 53.85% of the total car sales in the PQR Showroom. This shows that the brand that has existed since 1937 still dominates and is in great demand by the public, especially for PQR Showroom customers. After knowing that MPV-type cars with the Toyota brand dominate the market the most, we can narrow it down again to find out the specifics of the cars that have sold the most in the showroom.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

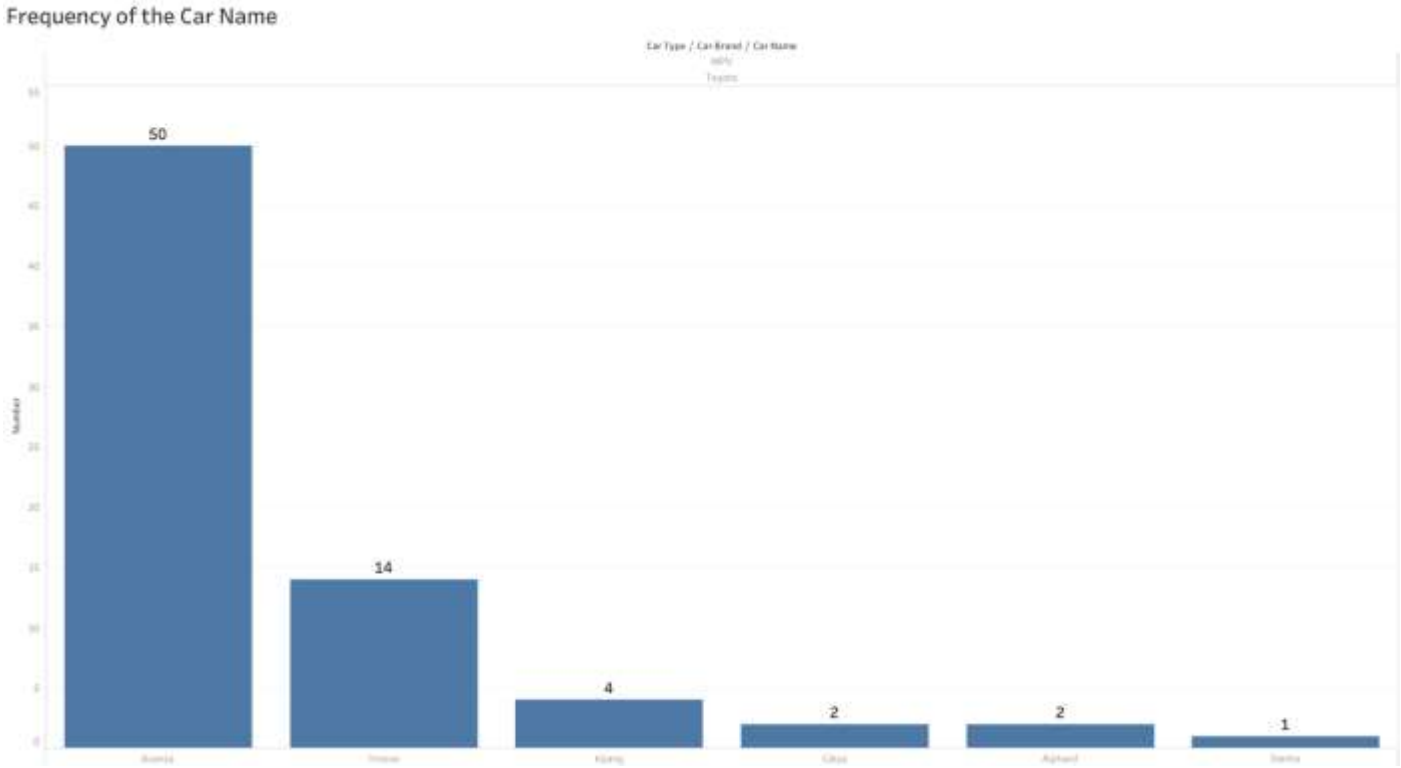

Figure 5: Frequency of the Car Name

According to Figure 5, after knowing the most cars based on the type and brand of car, then it is narrowed down again by knowing the specifics name of the car. After it was discovered that cars with the MPV type from the Toyota brand were the ones being bought the most, next we look at car sales from the Toyota brand. There are several car names from Toyota brand, namely Avanza, Innova, Kijang, Calya, Alphard, and Sienta. Sienta is the least sold car from the Toyota brand with only one car sold, while the car that was first launched in 2003, namely Avanza, still holds the position as the most car sold which is 50 cars. It can be concluded that the most sold car is the MPV type car, namely the Toyota Avanza, with sales reaching to 50 unit out of 352 cars that have been sold at the PQR Showroom.

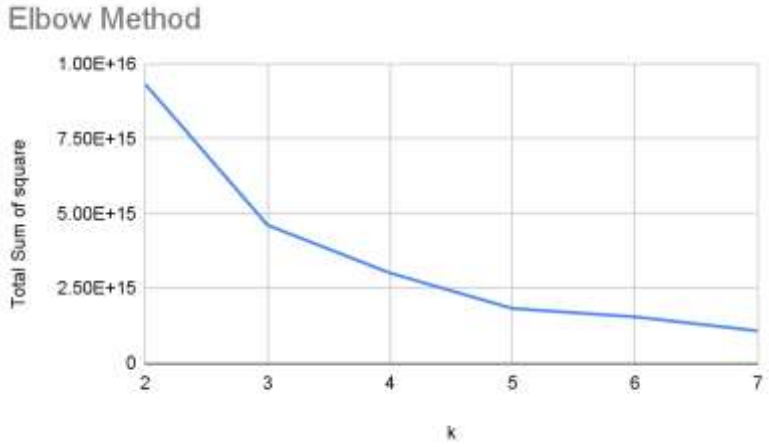

Figure 6: Elbow Method for Clustering Car Type, Days, and Profit per Car

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

K-means analysis is conducted to find out the most profitable car and the car that required a short time to be sold. Before conducting the k-means clustering analysis, the elbow method is carried out to determine the best number of clusters. According to Figure 6, the number of best clusters for clustering Car Type, Days, and Profit per Car is 3 (k=3), with the results of these elbows, we use three clusters for K-Means on car type, days, and profit per car.

k-Means - Centroid Table

Cluster	car_type	days	profit_per_car
Cluster 0	4.145	134.513	8880341.884
Cluster 1	7.588	177.039	21970588.243
Cluster 2	4.699	244.496	8278195.492

Figure 7: K-Means for Car Type, Days, and Profit per Car

The result of the k-means clustering analysis can be seen in Figure 7. According to Figure IV.2.5, cluster 0 is LCGC with an average time to be sold of 134.513 days and generate a profit of Rp 8,880,341. Cluster 1 is an SUV with average days to be sold of 177.0 days and generates a profit of Rp 21,970,588. As for cluster 2, it is MPV with average days to be sold 244.5 days and generating a profit of Rp 8,278,195. From the results of the K-means, it can be concluded that LCGC is the fastest car to be sold and MPV has the longest time to be sold. In addition, the most profitable car is the SUV.

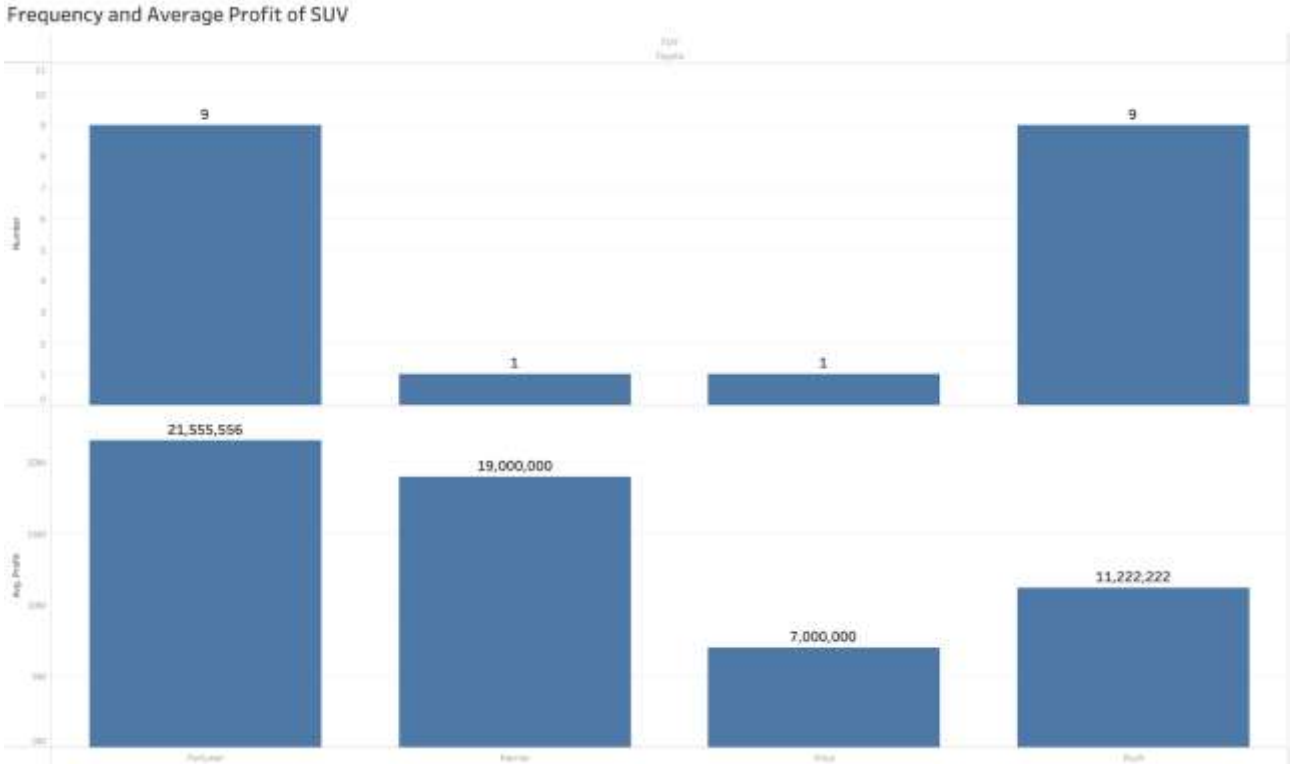

Figure 8: Frequency and Average Profit of SUV

Based on Figure 8, the Toyota Fortuner and Rush SUVs were the most sold with nine cars and the Toyota Harrier and Hilux SUVs were the least selling with one car. Even though the sales of the Toyota Fortuner and Rush are balanced, the profits generated are different. The Toyota Fortuner generates a higher average profit than all Toyota SUVs in the PQR Showroom with an average profit of Rp. 21,555,556.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

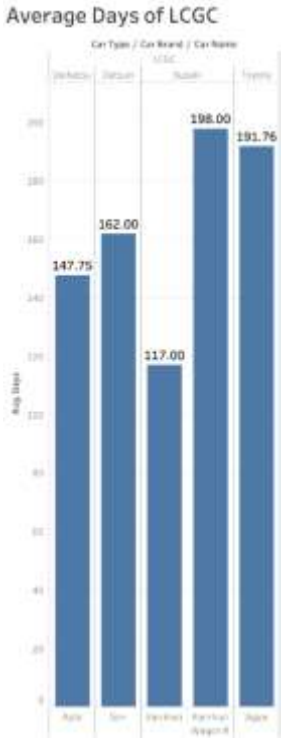

Figure 9: Average Days of LCGC

According to Figure 9, there are several LCGC type cars sold at the PQR Showroom, including the Daihatsu Ayla, Datsun Go+, Suzuki Karimun, Suzuki Karimun Wagon R, and Toyota Agya. The average days for cars sold for each LCGC car are also different, for Daihatsu Ayla it takes an average of 147.75 days to sell, Datsun Go+ takes 162 days to sell, Suzuki Karimun 117 days, Suzuki Karimun Wagon R 198 days, and Toyota Agya 191 .76 days. From these results, it can be seen that the car with the fastest and longest average days to besold is the Suzuki brand. The name of the Suzuki brand that differenciate the average days to be sold. It is known that the LCGC type car with the fastest average selling days is Suzuki Karimun.

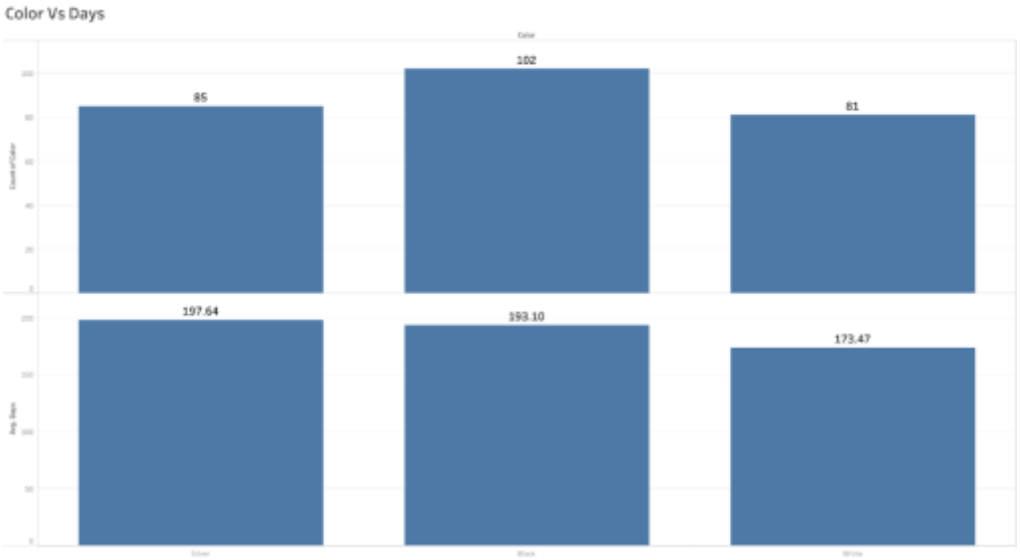

Figure 10: Frequency and Average Days of the Car Color

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

According to Figure 10, it can be seen that there are three colors that are most chosen by PQR Showroom customers, namely white, silver, and black. White became the third choice with sales of 81 cars, silver became the secondchoice with sales of 85 cars, and black became the firstchoice that sold the most which is 102 cars. However, the average days for cars sold the fastest were white with 173.47 days, followed by black with 193.10 days, and silver with 197.64 days. It can be concluded that the color that is most in demand by customers is black with and the color that sells the fastest is white with an average day of 173.47 days.

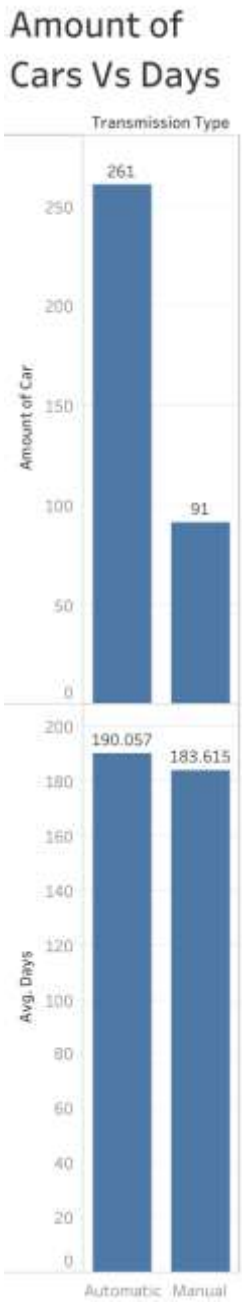

Figure 11: Brand and Transmission Vs Days

According to Figure 11, the automatic transmission is the transmission that sold the most in PQR Showroom, which amounted to 261 cars. Meanwhile, there are only 91 cars with manual transmissions that are sold. There is no big difference in average days for both transmissions to be sold, the difference is only 6 days.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

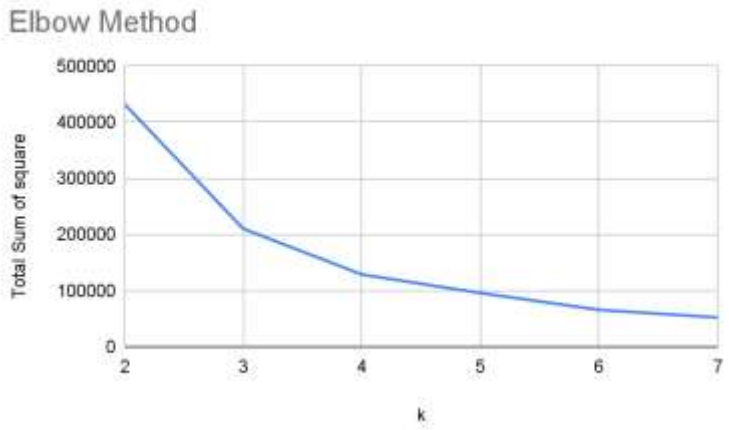

Figure 12: Elbow Method for Clustering Age of Car and Days

K-means analysis is conducted to find out the average age of the car that is sold fastest in the PQR Showroom. Before conducting the k-means clustering analysis, the elbow method is conducted to find out the best number of clusters for clustering Age of Car and Days. According to Figure 12, the best number of clustering using the Age of Car and Days variables on the elbow method is 3 (k=3).

k-Means - Centroid Table

Cluster	Age of Car (years)	days
Cluster 0	4.547	134.742
Cluster 1	25.000	143.000
Cluster 2	5.388	234.973

Figure 13: K-Means of Age of Car vs Days

According to Figure 13, we have three clusters, in cluster 0, the average car age is 4.55 years with the average days to be sold is 134.5 days. In cluster 1, the average car age is 25 years with the average of days to be sold is 143 days. And the last one, in cluster 2, the average of car age is 5.4 years with the average of days to be sold is 245 days. It can be concluded that cluster 0 represents the age of the car that sold the fastest the in PQR Showroom for 134.5 days or around 2 months.

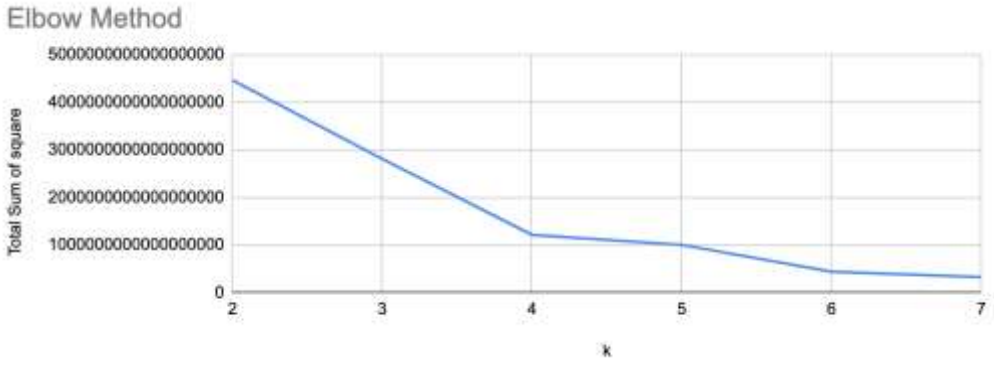

Figure 14: Elbow Method for Payment Method vs Selling Price

K-means analysis using Payment Method and Selling price as the variables is carried out to find out the trends of payment method in PQR Showroom. According to the Figure 14, the best number of clusters for these two variables is 4 (k=4).

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

k-Means – Centroid Table

Cluster	payment_method	selling_price
Cluster 0	3	171468296.629
Cluster 1	1.783	217937989.511
Cluster 2	2.000	1850000042.570
Cluster 3	2.059	711470604.607

Figure 15: K-Means of Payment Method vs Selling Price

The result of the k-means clustering analysis can be seen in Figure IV.2.13. According to Figure IV.2.14, cluster 0 is debit, and the average selling price is Rp 8,880,341. Cluster 1 is a credit card with an average car selling price of Rp 217,937,989. Cluster 2 is a credit card with the average car selling price is Rp 185,000,000. Cluster 3 is credit card 3 with an average car selling price Rp 711,470,604.

CONCLUSION AND RECOMMENDATION

PQR Showroom has been facing low profit from sales and failed to optimize the profit. According to Tree diagram, there are two main causes that posed PQR Showroom to have low profit which are: Data is not utilized as part of decision making and the inability to recruit marketing team due to budget limitation. Data analytics processes using descriptive analytics and k-means clustering analysis are conducted in this research in order to help PQR Showroom to improve its profit. Based on the result and business interpretation conducted, there are several solutions to improve PQR profit: Prioritize in providing SUV Toyota Fortuner; provide more units of LCGC Suzuki Karimun; keep providing sufficient units of MPV Avanza with higher profit margin; prioritize to provide a car with the age of <5 years, since this car sold fastest; providing a majority of neutral colours such as Black, silver, and white when stocking the inventory; and facilitate customers with leasing which provide low interest rate and small down payment to encourage customers to buy. Several scenarios are carried out to assess the proposed business solutions. According to the scenarios, proposed business solutions are expected to be able to improve PQR Showroom profit and result in efficiency of car inventories. PQR Showroom is suggested to use its sales data in the future as part of its business decision-making process, since data will be able to reveal trends and patterns that can benefit PQR Showroom for keeping sustainable business and further growth.

ACKNOWLEDGEMENT

The authors would like to thank Mr. Meditya Wasesa for his guidance, those who supported this research, and the editors who received and reviewed this paper.

REFERENCES

1. Mendiratta, A. (2021, January 21). *Digitalization Transforming the Used Car Industry in Indonesia*. Ken Research. Retrieved July 9, 2023, from https://www.kenresearch.com/interviews/digitalization-transforming-the-used-car-industry-in-indonesia/45.html
2. Mordor Intelligence. (2022, December). *Southeast Asia Used Car Market - Share, Analysis & Growth*. Mordor Intelligence. Retrieved July 9, 2023, from https://www.mordorintelligence.com/industry-reports/south-east-asia-used-car-market
3. OLX Autos Indonesia. (2021, December). *2021 Used Car Market Review & 2022 Outlook*. OLX News. Retrieved July 9, 2023, from https://news.olx.co.id/wp-content/uploads/2022/02/OLX-Whitepaper-final-Eng-revcopy-10.12.pdf
4. Liu, S., Liu, O., & Chen, J. (2023). Mathematics Journal. *A Review on Business Analytics: Definitions, Techniques, Applications, and Challenges*, *11*(899).
5. Duan, L., & Xiong, Y. (2015, March 19). Journal of Management Analytics. *Big Data Analytics and Business Analytics*, *2*(1), 1-21. http://www.tandfonline.com/loi/tjma20
6. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012, December). MIS Quarterly. *Business Intelligence and Analytics: From Big Data to Big Impact*, *36*(4), 1165-1188. https://www.jstor.org/stable/41703503
7. Delen, D., & Zolbanin, H. M. (2018, September). Journal of Business Research. *The Analytics Paradigm in Business Research*, *90*, 186-195.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 07 July 2023
DOI: 10.47191/ijcsrr/V6-i7-106, Impact Factor: 6.789
IJCSRR @ 2023
www.ijcsrr.org

8. Delen, D., & Ram, S. (2018). Journal of Business Analytics. *Research Challenges and Opportunities in Business Analytics*, *1*(1), 2-12.
9. Kraus, M., Feuerriegel, S., & Oztekin, A. (2019, September 11). European Journal of Operational Research. *Deep Learning in Business Analytics and Operations Research: Models, Applications, and Managerial Implications*, *239*(3), 628-641.
10. Bayrak, T. (2015). Procedia - Social Behavioral Sciences. *A Review of Business Analytics: A Business Enabler or Another Passing Fad*, *195*(2015), 230-239.
11. Ware, C. (2004). *Information Visualization: Perception for Design*. Morgan Kaufmann.
12. Lepenioti, K., Bousdekis, A., Apostolous, D., & Mentzas, G. (2020). International Journal of Management. *Prescriptive Analytics: Literature Review and Research Challenges*, *50*(2020), 57-70.
13. Ediyanto, Mara, M. N., & Satyahadewi, N. (2013). Buletin Ilmiah Mat, Stat, dan Terapannya. *Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis*, *2*(2), 133-136.
14. Merliana, N. P. E., Ernawati, & Santoso, A. J. (2015). PROSIDING SEMINAR NASIONAL MULTI DISIPLIN ILMU & CALL FOR PAPERS UNISBANK. *ANALISA PENENTUAN JUMLAH CLUSTER TERBAIK PADA METODE K-MEANS CLUSTERING*.
15. Arvai, K. (2020). *K-Means Clustering in Python: A Practical Guide – Real Python*. Real Python. Retrieved July 9, 2023, from https://realpython.com/k-means-clustering-python/
16. Xiao, J., Lu, J., & Li, X. (2017). Intelligent Data Analysis. *Davies Bouldin Index Based Hierarchical Initialization K-Means*, *21*, 1327-1388.

Cite this Article: Muhammad Rizky Satya Bhaskara, Meditya Wasesa (2023). Enhancing Car Showroom Profitability through K-Means Clustering for Customer Segmentation. International Journal of Current Science Research and Review, 6(7), 4904-4916