

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 08/13/23

Internship Batch: 109454990337545509761

Version: #677

Data intake by: Jason Sanchez

Data intake reviewer: Data Glacier Reviwer

Data storage location: https://github.com/Kuzma12/VC/blob/G2M-insight-for-Cab-Investment-firm_VC/Cab_Investment_Firm.ipynb

Tabular data details:

Cab_Data.csv

Total number of observations	359,392
Total number of files	4
Total number of features	7
Base format of the file	csv
Size of the data	20 MB

City.csv

Total number of observations	19
Total number of files	4
Total number of features	3
Base format of the file	csv
Size of the data	1 KB

Customer_ID.csv

Total number of observations	49,171
Total number of files	4
Total number of features	4
Base format of the file	csv
Size of the data	1 MB

Transaction.csv

Total number of observations	440,098
Total number of files	4
Total number of features	3
Base format of the file	csv
Size of the data	8 MB

Proposed Approach:

- The process of data identification begins by discerning the data related to two specific companies in the ride-hailing industry and ascertaining the prevalence of each company's services in various cities. This initial step enables us to attribute values to respective cities based on their popularity for each cab company. Subsequently, a comprehensive analysis of the income distribution among the users can be conducted by categorizing them into income ranges. This process often unveils patterns that aid in determining which demographic group utilizes the cab services most frequently. Furthermore, the data analysis can encompass a comparison of the fare charges across different cab services. By juxtaposing the actual charges with the prevailing market rates, a percentage variance can be calculated to understand whether the charges are either higher or lower than the standard fare. This evaluation contributes to assessing the competitiveness of each company's pricing strategy, potentially providing insights into their market positioning. These intricate steps collectively drive the decision-making process concerning the two cab companies. By meticulously following this approach, a comprehensive and informed judgment can be reached regarding the performance and potential strategies of the firms. The data-driven insights gained from these analyses can offer valuable recommendations, aiding in strategic business decisions that may enhance the efficiency, profitability, and overall performance of the respective cab companies in consideration.
- When approaching the data analysis, it is imperative to recognize the multifaceted nature of user behavior and the underlying factors that drive their choice of using a cab service. A thorough examination of the data should extend beyond mere user categorization and delve into the nuanced intricacies that influence the decision-making process. One key aspect that must not be overlooked is the purpose behind each cab ride. Different individuals utilize cab services for a myriad of reasons, ranging from daily commuting to occasional leisure trips, business meetings, travel, and more. The significance of this distinction lies in the potential variance it introduces to the user rate based on their specific needs. For instance, a user who frequently relies on cabs for daily commuting purposes might contribute more to the overall usage rate compared to a sporadic user who employs cab services only on occasion. To extract meaningful insights from the data, it is imperative to gain a comprehensive understanding of the underlying motives for using the cab service. This can be achieved through methods such as surveys or feedback forms that gather information about the primary purpose of the cab rides. By identifying whether the user's choice is driven by necessity, convenience, leisure, or any other factor, a more nuanced analysis can be performed. Moreover, the frequency of cab usage is also a critical variable that can significantly impact the results. For instance, a user who occasionally uses cabs may have a relatively lower contribution to the overall data, while a frequent user might wield a more substantial influence. Acknowledging such variations and their potential implications on the data interpretation is essential to draw accurate conclusions. By delving into the nuances of user behavior and considering their varied motivations, a comprehensive and holistic understanding of the data can be achieved. This approach ensures that the data analysis is not oversimplified and enables us to draw insights that are closer to the reality of user preferences and habits. Such a thorough

exploration of user behavior is indispensable in ensuring that the data is utilized to its fullest potential, leading to more informed and effective decision-making for the cab service provider.