PARSHVANATH CHARITABLE TRUST'S
A. P. SHAH INSTITUTE OF TECHNOLOGY
Department of Information Technology
(NBA Accredited)

Uber Data Analysis

Muskan Rao 20104096
Sanskruti Rawal 20104020
Divyal Chheda 21204005

Project Guide

Prof. Geetanjali Kalme

Contents

1. Introduction

- Problem Identified :

1. **Pricing optimization:** Uber's pricing strategy is crucial to its success. Data analysis can help Uber identify pricing patterns and trends to optimize fares and increase revenue.
2. **Rider satisfaction:** Uber's success is also heavily dependent on rider satisfaction. Data analysis can help identify areas where riders are dissatisfied and suggest improvements to enhance the overall customer experience.
3. **Driver performance:** Uber needs to ensure that its drivers are providing a high level of service to riders. Data analysis can help identify areas where drivers are underperforming and suggest training or other measures to improve driver performance.

1. Introduction

- Solution Proposed :

1. The solutions proposed by Uber data analysis can vary depending on the specific insights gained from the analysis.
2. Here are a few examples of how Uber data analysis can inform solutions :
    - Optimal Pricing Strategy
    - Efficient Routing
    - Driver Allocation
    - Surge pricing
    - Driver Incentives

2. Objectives

The objectives of Uber data analysis are to gain insights and make data driven decisions that improve the overall performance of the Uber platform. Here are some specific objectives of Uber data analysis:

- To improve rider experience by analyzing rider behavior, feedback, and other data.
- To Increase driver earnings by analyzing driver earnings data.
- To use statistics and visualisations to identify trends in datasets.
- To Optimize pricing by analyzing demand patterns and other data.

3. Scope

- The scope of Uber exploratory data analysis (EDA) is to investigate and understand the data generated by Uber rides.
- This involves analyzing data on source, destination, trip details, and other variables that impact the overall performance of the Uber platform.
- The goal of EDA is to uncover patterns and insights in the data that can inform future decisions and strategies.
- Data quality and completeness: EDA can involve an examination of the quality and completeness of the data. This includes identifying missing data, outliers, and errors that can impact the analysis.

4. Literature Survey

Sr No.	Title	Author(s)	Year	Outcomes	Methodology	Result
1	UBER DATA ANALYSIS USING GGPLOT	Mrunal Patil, Vidya Kumari, Adarsh Patil, Laxmikant Ahire and Asst.Prof. Umakant Mandawkar from B.Tech, Computer Science and Engineering, Sandip University, Nashik, India	2021	Uber analyze historical data for say, last 3 or 4 weeks and identifies pockets within the city that witness extremely high demand	Estimated Time of Arrival can be reduced with increase in the number of Uber drivers which successively will make Uber more liked by the customers and hence, the company will get more revenue and drivers will also be profited.	At the end of all procedure we get to see different graphs giving us insights.

4. Literature Survey

Sr No.	Title	Author(s)	Year	Outcomes	Methodology	Result
2	Uber Related Data Analysis using Machine Learning	Rishi Srinivas, B. Ankayarkanni, R. Sathya Bama Krishna	2021	The paper explains the working of an Uber dataset, which contains data produced by Uber for New York City. Uber is defined as a P2P platform.	Using the information, the paper explains the use of the k-means clustering algorithm on the set of data and classify the various parts of New York City. Since the industry is booming and expected to grow shortly.	Effective taxi dispatching will facilitate each driver and passenger to reduce the wait time to seek out one another. The model is employed to predict the demand on points of the city.

5. Proposed System

We use the NYC dataset containing data on over 4.5 million trips in NYC from Apr. to Sept. 2014. The files are separated by month and the following columns:

- Date
- Time Stamp
- Destination
- Source
- Price

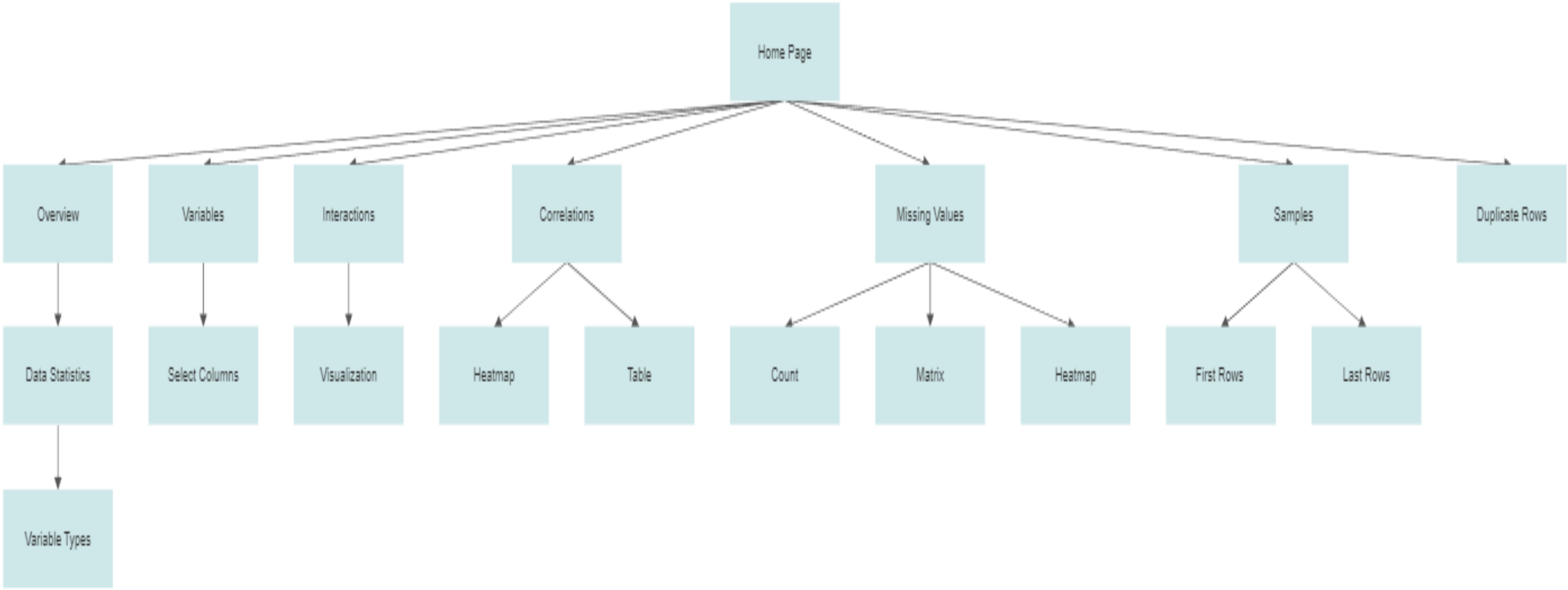
There are some of the questions this dataset can be used to answer, we'll choose the following:

- Uber trips and distribution
- Time when Uber trips occur regularly
- Price range of all rides

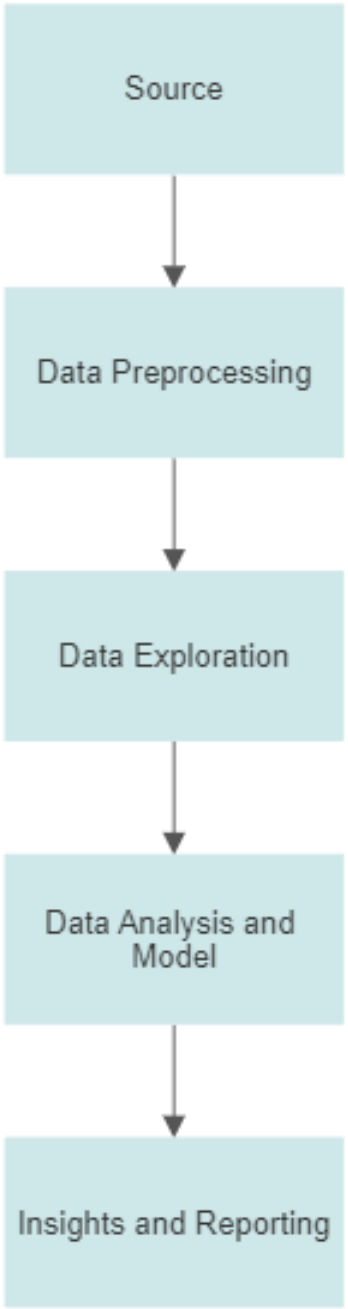
7. Outcome of Project

1. We have seen how the demand follows unique patterns during the day and at some stage in the week.
2. We have also seen how time affected client trips.
3. Uber exploratory data analysis (EDA) shows us insights and knowledge that can inform business decisions and strategies.
4. We have seen how to find accuracy using multiple algorithms and relations between different data attributes.

8. Block Diagram



9. DFD Diagram



10. Technology Stack

- Python 3
- Google colab, VS Code
- Streamlit
- Packages Used – numpy, pandas, sklearn, scipy, matplotlib, seaborn etc.

Suggestions in Review - 1

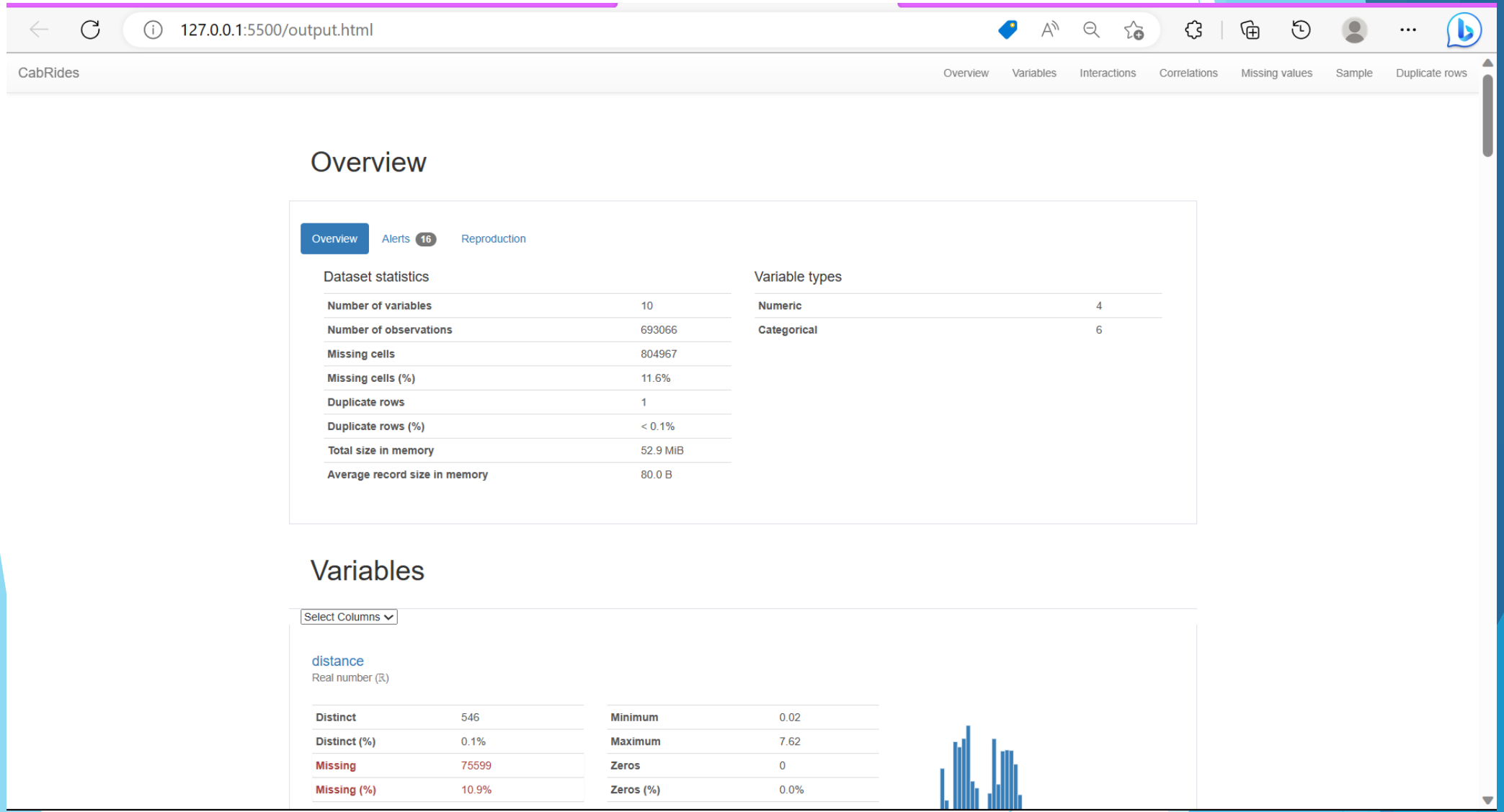
1. Change in Objectives in PPT & Report
2. GUI for the project
3. Use case/DFD Diagram

Result and Discussion

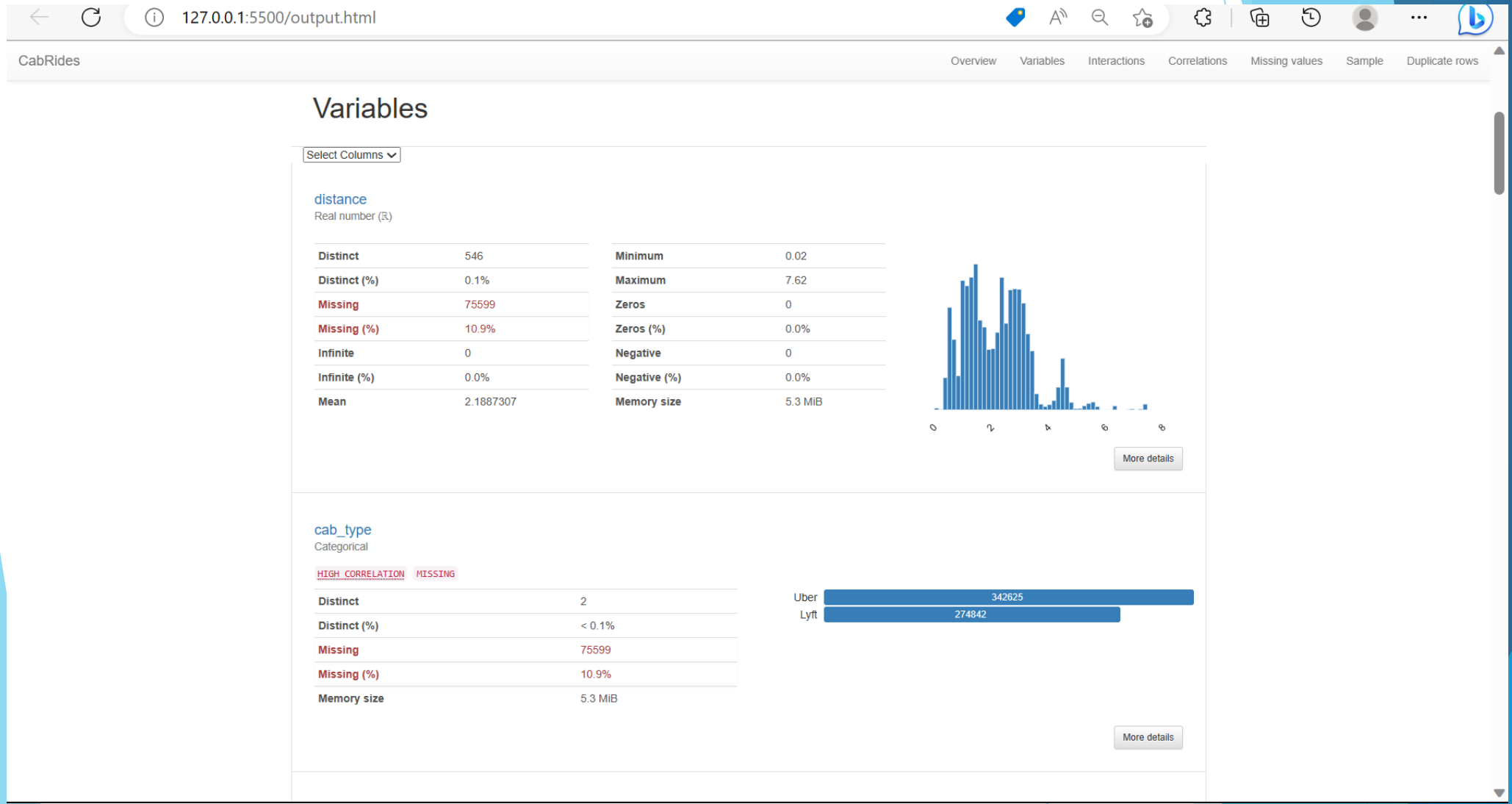
The result of Uber data analysis can provide valuable insights into the operations of the taxi service and inform solutions to improve efficiency and customer satisfaction. Here are some possible results and discussions that may arise from Uber data analysis:

- Demand Patterns
- Route Optimization
- Surge pricing
- Driver Incentives
- Customer Behavior

Result and Discussion



Result and Discussion



Result and Discussion

127.0.0.1:5500/output.html

CabRides

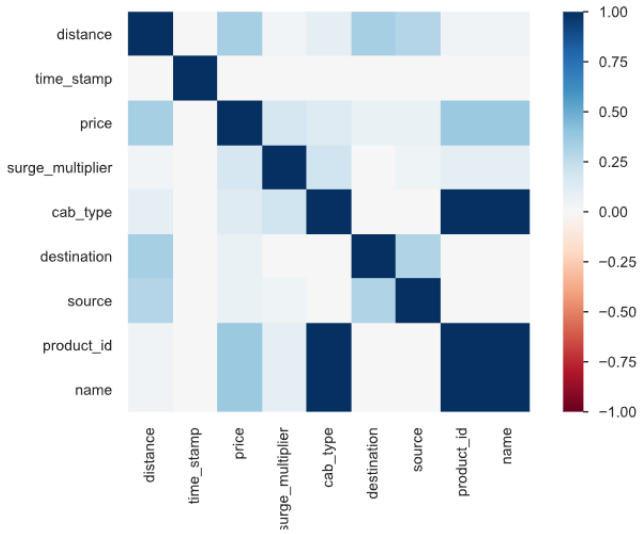
Overview Variables Interactions Correlations Missing values Sample Duplicate rows

distance

Correlations

Auto

Heatmap Table



Missing values

Conclusion and Future Scope

1. Uber data analysis is a powerful tool that can inform solutions to improve the efficiency and customer experience of the taxi service. By analyzing data on demand patterns, route optimization, surge pricing, driver incentives, and customer behavior, Uber can optimize its operations and provide a better experience for both passengers and drivers.
2. The Future Scope of Uber Data Analysis :
    - Artificial Intelligence
    - Personalization
    - Safety & Security
    - Environmental sustainability
    - Integration with other services

References

1. https://jespublication.com/upload/2021-V12I759.pdf
2. https://ieeexplore.ieee.org/document/94323
3. https://www.analyticsvidhya.com/blog/2021/10/end-to-end-predictive-analysis-on-ubers-data/
4. https://github.com/Unnati0104/Uber-Data-Analysis.git

Thank You...!!