1**.INTRODUCTIO**

1.1 Overview

Uber is a multinational transportation network company that operates a ride-hailing platform. It was founded in 2009 by Garrett Camp and Travis Kalanick and is based in San Francisco, California. Uber provides a convenient way for individuals to request rides from drivers who use their own personal vehicles.

Uber Driver Analysis refers to the Analyzing the number of trips taken by Uber drivers can provide insights into their overall activity and the demand for rides in specific areas. Daily, Weekly, or Monthly Analysis: Uber's data can be analyzed on a daily,

weekly, monthly basis to understand the trends and patterns of trip volumes. This analysis can help identify peak hours or days of high demand and optimize driver availability during those times. Trips can be analyzed based on geographic regions or specific cities to identify areas with higher demand. This analysis can help Uber drivers decide where to focus their driving efforts for maximum efficiency and profitability. The Major of our project is to use data Analyzing techniques to find unknown patterns in the Uber Drives dataset. The research is carried out on Uber drives data collected from the year 2016.

## 1.2 Purpose

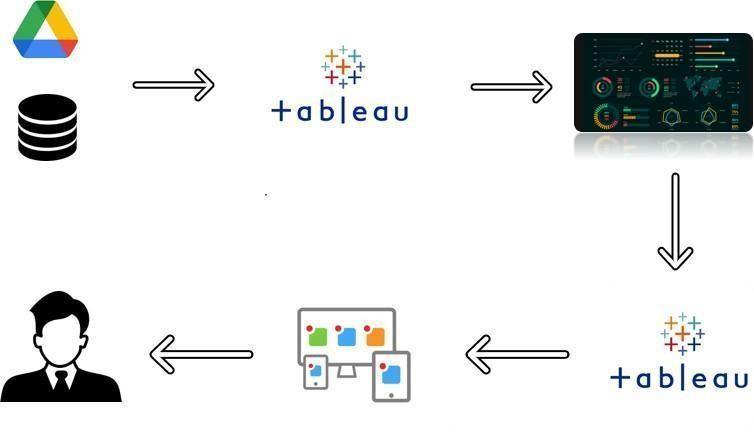
## Takeaways from Uber Data Analysis Project using Machine Learning

We looked at two different types of Uber ride datasets – a personal ride history for a single person and a two-month record of all the Uber rides in Boston, MA. We compared similar graphs in the EDA process of both these datasets to generate real-world insights into the behavior of the Uber riders and trips in the city of Boston. Furthermore, we used different ML models to perform a price prediction of the Uber ride based on a fixed number of features from the second dataset. We also perform feature selection to reduce the number of features and find the optimal amount to improve model performance to a certain degree. We find that Gradient Boosting Machine (GBM) works best in this dataset, yielding 0.95.

Finally, we noted some crucial issues faced by scientists during EDA and data analysis and listed the challenges of working with the Uber datasets.

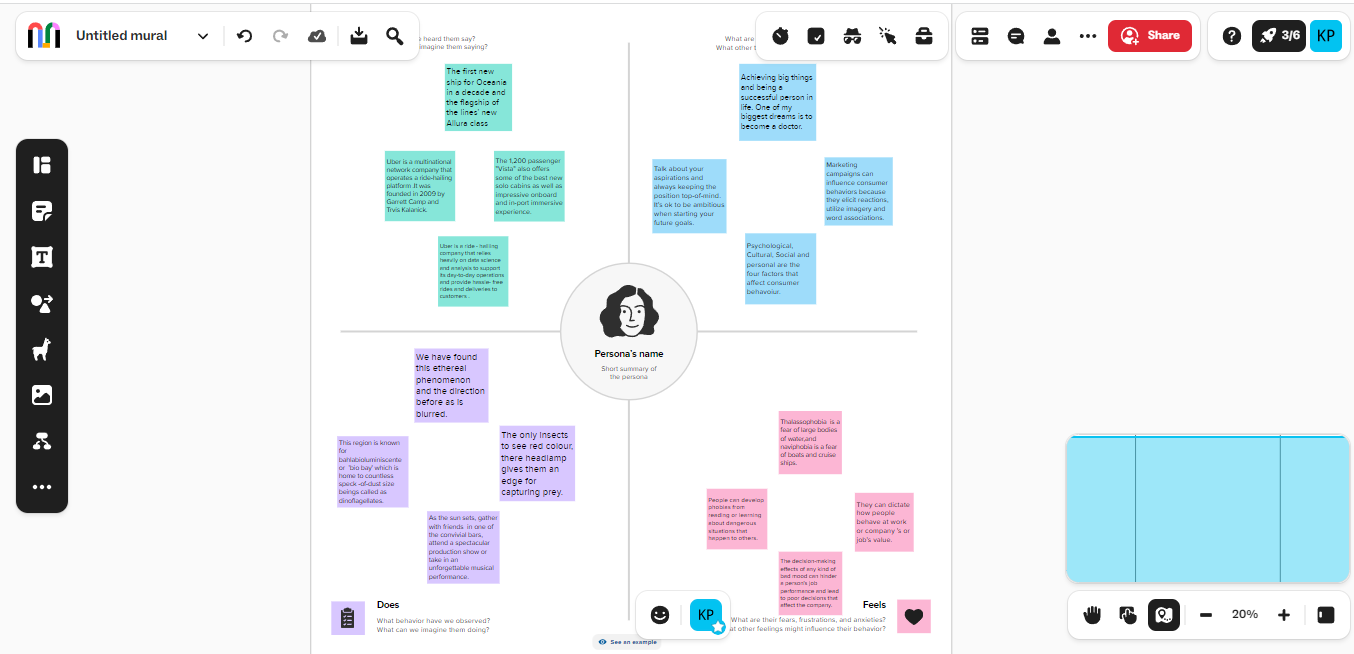
Technical Architecture

**Technical Architecture:**

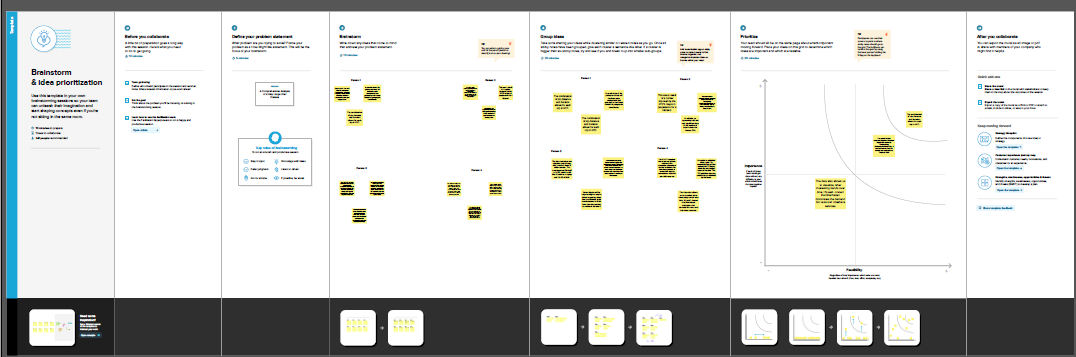


2. Problem Definition & Design Thinking

2.1 Empathy Map



2.2 Ideation & Brainstorming Map



3. RESULT

**RECOMMENDATION:** Do not start your analysis without completing the Business Problem Definition, since it determines your analysis’ focus and quality. Besides that, this process will help you to think about new possibilities/questions while trying to answer the previous ones set.

**NOTE:** In order to organize better my analysis, I will create an additional data frame, **removing all trips with status CANCELED and DRIVER\_CANCELED**, since they should be disregarded from some questions.

4. ADVANTAGES & DISADVANTEAGES

## Safety and Flexibility For Drivers

Safety is an important advantage for drivers working with Uber and other e-hail services. The riders using the service have registered their identities and their credit card numbers on the app. They are not random strangers on the street.

Because the transaction is cashless, a driver doesn't risk unpaid fares or need to carry cash for change.

Rude, aggressive, and disruptive passengers are weeded out because drivers rate their customers. Consistently low ratings or reports of unsafe behavior toward drivers can cause the deactivation of an account.5

Unlike yellow cab taxi drivers who work 12-hour shifts or black car drivers who are scheduled by dispatchers, Uber drivers enjoy considerable freedom and flexibility. Drivers log in and out of the system anytime they choose and pick their own hours.6

Drivers avoid expensive taxi rental leases by using their own vehicles. They also pay their own fuel and maintenance costs. All else being equal, this may mean more profit for drivers.

Drivers are also spared any office politics because the app renders dispatchers irrelevant.

## Negative Impact of Price Competition

Price competition can be destructive for any industry. Increasingly, [Uber, Lyft,](https://www.investopedia.com/articles/personal-finance/010715/key-differences-between-uber-and-lyft.asp)and other e-hail services are engaged in an intense battle to provide the cheapest service. They are directly competing with each other, and with traditional taxi and car services for both customers and drivers.

With competition from other ride-sharing services and the continuous hiring of new drivers, average earnings are being pushed downward. This means that drivers have to work longer hours to earn an income comparable to what they would have earned a year or two ago.

5.APPLICATIONS

Exploring data is certainly one of the most important stages in Data Science processes. Despite its simplicity, it can be a powerful tool to put you ahead on data and business context, as well as to determine crucial treatments before creating machine learning models.

To turn things a little bit more interesting, I’ve decided to have some fun with Python on my personal Uber rides data and see which insights I could extract.

In this post, I will guide you through the following steps:

1. Problem Definition

2. Data Discovery

3. Data Preparation

4. Data Analysis & Storytelling

6.CONCLUSION

Exploratory Data Analysis is **not a trivial task**! It requires lots of work and patience, however, it is surely a **powerful tool if correctly applied to your business context**.

This post briefly demonstrated some tips and steps to make analysis easier and undoubtedly highlighted the **crucial importance of a well-defined business problem**, guiding all coding efforts to a specific objective, and also highlighting important insights. This business case also tried to reflect a **practical application of python in daily business activities**, showing how fun, valuable, and interesting it could become.

Thank you so much for getting here! I hope you liked it! 😃

7.FUTURE SCOPE

# Feature Engineering: Creating new features

Based on <fare\_amount> and <distance\_miles> features I've created a new feature called <amount\_km>, which would help us understand **how much is paid by kilometer ridden**.

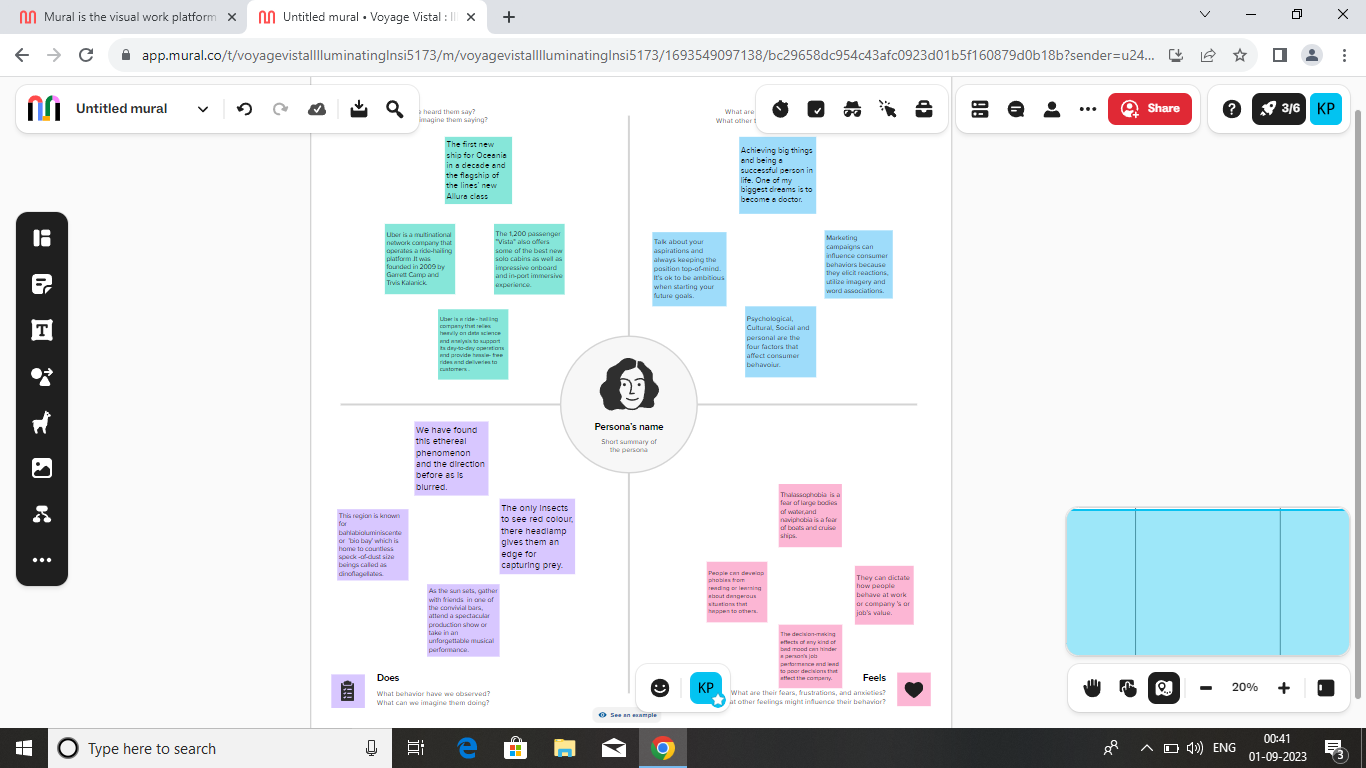
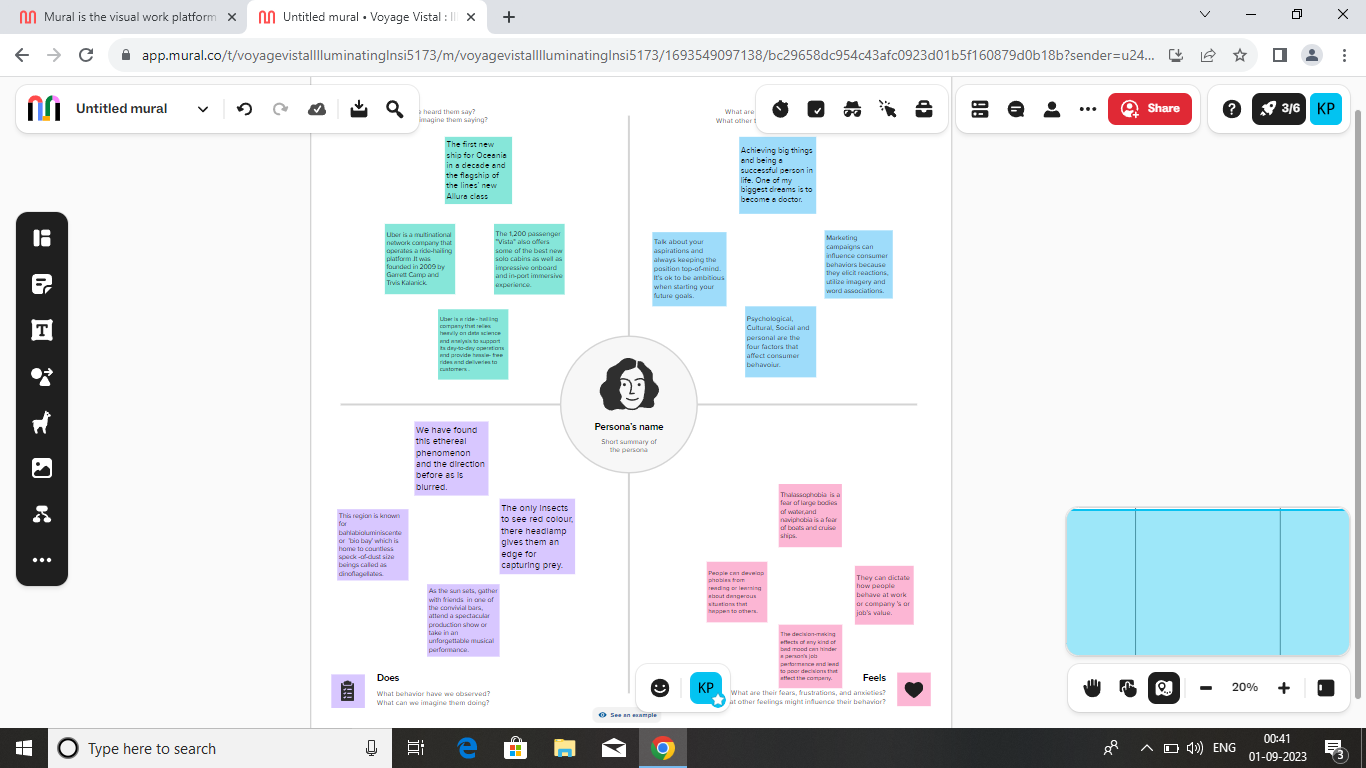
Delta time between <request\_time> and <begin\_time> will let us know **how much time (in minutes) I usually waited for Uber cars to arrive at my destination**. In this case, it was calculated on a minutes base.

8. APPENDIX

<https://github.com/bdu1621me1291/Voyage-Vista-Illuminating_NM2023TMID25634-Insights-from-Uber-Expeditionary-Analysis/upload/main>

THANKG YOU

2.Prob

2lem Definition and Design Thinking

2.1 Empathy Map