

Project Proposal – Uber Trips Data Analysis

1. Project Title

Uber Trips Data Analysis (2009–2015)

2. Project Overview

This project analyzes a sample dataset of Uber trip records from 2009 to 2015.

The study focuses on discovering trip patterns, demand behavior, ride characteristics, and external factors affecting trip dynamics. The project also includes building a machine learning model to predict fare amount.

3. Project Objectives

- Explore trip behavior across different time periods.
- Identify peak demand hours, days, and months.
- Analyze fare structure and ride characteristics.
- Understand the impact of weather and traffic conditions.
- Build a predictive machine learning model for estimating fare amount.
- Present findings through a structured analytical dashboard.

4. Team Members

1. Basem Hamada Ebrahim
2. Mohamed Nabil Elawad
3. Mohannad Ashraf Abd Albadea
4. Khaled Ashraf Abd Shahed
5. Mahmoud Sobhy Hamed Atiaa

5. Project Scope

Included:

- Data cleaning and preprocessing
- Analytical exploration of trip patterns
- Visualizations (charts & KPIs only — maps excluded)
- Dashboard development using Power BI
- Machine learning model for fare prediction
- Insights and summary findings

Not Included:

- Tableau
- Geographic maps
- Real-time data integration
- API pipelines

6. Dataset Description

The dataset includes Uber trip records with:

- Date/Time
- Pickup & Dropoff Locations
- Distance
- Fare Amount
- Weather Condition
- Traffic Condition
- Airport Pickup/Dropoff Indicator
- Passenger Count

7. Methodology

1 — Data Cleaning

- Handling missing values
- Removing duplicates
- Standardizing formats

2 — Analytical Exploration

- Trip frequency patterns
- Fare behavior
- Distance patterns
- Traffic & weather influence
- Airport trip comparison

3 — Machine Learning Model

- Target: Fare Amount
- Models: Linear Regression, Random Forest, XGBoost
- Features: distance, time, passenger count, traffic, weather
- Evaluation: RMSE, MAE, R²

4 — Dashboard Development

- Built using Power BI with KPIs, charts, and filters

8. Key Features

- Clean and well-prepared dataset
- Professional analytical workflow
- Machine learning prediction model
- Interactive dashboard
- Actionable insights

9. Deliverables

- Cleaned dataset
- Analytical findings
- Machine learning model
- Power BI dashboard
- Final presentation
- Project documentation

10. Future Improvements

- Real-time updates
- Enhanced predictive models
- API integrations
- Fleet optimization
- Mobile dashboard

11. Timeline

Data Cleaning: 1 week

Analysis: 1 week

Machine Learning: 1 week

Dashboard: 1 week

Documentation: 3–4 days

12. Tools & Technologies

Python, SQL Server, Power BI, Excel, Jupyter Notebook

13. Conclusion

This project delivers insights into Uber trip behavior through data processing, analysis, and machine learning, supporting informed decision-making.

