**Project Overview**

This project analyzes Uber fare data to uncover patterns related to time, distance, pricing, and location. The objective was to clean, explore,

and enhance the dataset using Python, and visualize the results in an interactive Power BI dashboard.

**Dataset Description**

- **Source:** [Uber Fares Dataset on Kaggle](https://www.kaggle.com/datasets/yasserh/uber-fares-dataset)

**Tools Used**

* Python (Pandas, Matplotlib, Seaborn)
* Data Cleaning, EDA, Feature Engineering s
* Power BI Desktop
* Data Visualization and Dashboard Design
* GitHub
* Documentation and Version Control

**Steps and Methodology**

. **Data Cleaning**

* Removed null and invalid entries
* Filtered outliers (e.g., negative fares, zero distances)
* Converted pickup\_datetime to datetime format

. **Feature Engineering**

* Extracted hour, day, month, and weekday
* Created peak/off-peak hour flags
* Exported the enhanced dataset to `Cleaned\_UberData.csv

**Data Visualization**

* Imported CSV into Power BI
* Created visuals: bar charts, histograms, scatter plots, line charts, maps
* Built an interactive dashboard with slicers and drilldowns

-

**Dashboard Highlights**

* Fare Distribution: Box plot and histogram
* -based Ride Analysis: Hourly, daily, monthly views
* Fare vs Distance Correlation
* Time Series Trends
* Geographic Ride Heatmap
* Interactive Filters: Peak hours, date ranges, weekdays

**Key Insights**

* Peak ride hours are early mornings and late afternoons
* Fare generally increases with distance
* Weekdays, especially Fridays, see the most rides
* Evening rides are often more expensive
* Certain areas show higher ride demand

**Deliverables**

* Cleaned\_UberData.csv
* – Cleaned and enhanced dataset
* UberFaresDashboard.pbix
* – Power BI Dashboard File /screenshots/
* – Folder with key analysis and dashboard screenshots
* pptx– Final report (this document)
* README.md` – Project documentation (you are reading it)

**Recommendations**

* Adjust pricing strategies during peak hours
* Deploy more drivers in high-demand time blocks
* Use geographic insights for targeted marketing
* Offer discounts during off-peak hours to balance demand

--