Uber Trip Analysis – Project Report

Phase 1: Problem Definition
The primary objective of this project was to analyze Uber trip data to identify booking patterns, revenue trends, and trip efficiency, enabling data-driven decision-making. The analysis aimed to provide stakeholders with actionable insights on customer behavior, operational efficiency, and revenue generation.

Key Business Requirements

1. Understand total bookings, revenue, and distance metrics.
2. Analyze trip patterns by time, day, and location.
3. Identify most common pickup and drop-off points.
4. Detect trends in vehicle preferences and payment methods.
5. Provide a detailed data drill-through capability for granular analysis.

Phase 2: Data Collection and Preparation
Three data sources were provided:

- **Trip Details Table**: Transaction-level ride data including pickup/drop-off times, passenger count, trip distance, fare amount, payment type, surge fees, and vehicle type.
- **Location Table**: Maps numerical LocationIDs to actual location names and cities.
- **Calendar Table**: Created as a supporting dimension to enable analysis by day, date, and weekday name.

Key Data Preparation Steps

- Ensured correct data types (e.g., datetime for pickup and drop-off times, numeric for trip distance and fare).
- Created calculated columns such as **Trip Type (Day/Night)** using pickup hour logic.
- Verified relationships between tables:
    - Trip Details to Location Table (via Pickup LocationID and Drop-off LocationID).
    - Trip Details to Calendar Table (via Pickup Date).
- Added a **Dynamic Measure Table** (disconnected) to facilitate KPI switching in visuals.

Phase 3: Data Modeling
A star schema was designed with **Trip Details** as the fact table, connected to dimension tables.

- **Trip Details (Fact Table)**: Core transactional data.
- **Location Table (Dimension)**: Lookup for pickup and drop-off locations.
- **Calendar Table (Dimension)**: Provides temporal hierarchy for day, date, and weekday analysis.
- **Dynamic Measure (Disconnected Dimension)**: Used to control KPI selection through DAX.

The inactive relationship between Trip Details[DOLocationID] and Location Table[LocationID] was modeled to allow drop-off analysis using USERELATIONSHIP in measures.

Phase 4: DAX Calculations
Multiple measures were developed to support analysis.

Booking and Revenue Measures

- Total Bookings = COUNT of Trip IDs
- Total Booking Value = SUM of fare amount
- Average Booking Value = DIVIDE(Total Booking Value, Total Bookings)

Distance and Time Measures

- Total Trip Distance = SUM of trip distance
- Average Trip Distance = AVERAGE of trip distance
- Average Trip Time = AVERAGE of trip duration

Location Analysis Measures

- Most Frequent Pickup Point: Used SUMMARIZE and TOPN to identify the location with maximum bookings.
- Most Frequent Drop-off Point: Used inactive relationship with USERELATIONSHIP to calculate most common drop-off.
- Farthest Trip: Retrieved pickup and drop-off locations for the trip with the maximum distance.
- Top 5 Locations by Bookings: Ranked locations based on booking frequency.

Trip Type Measure

- Trip (Day/Night): Categorized trips into Day Trip or Night Trip based on pickup hour.

Dynamic Measure Implementation

- A disconnected table allowed users to switch between KPIs (Total Bookings, Total Booking Value, Total Trip Distance).
- DAX SWITCH statements enabled visuals to update dynamically based on selection.

Phase 5: Dashboard Development in Power BI

Dashboard 1: Overview Analysis

- KPIs: Total Bookings, Total Booking Value, Average Booking Value, Total Trip Distance, Average Trip Distance, Average Trip Time.
- Visuals:
    - Donut charts for bookings by payment type and trip type (Day/Night).
    - Matrix for vehicle type analysis with KPIs and conditional formatting.
    - Line chart for bookings by day.
    - Location insights including frequent pickup/drop-off points, farthest trip, and top locations.

Dashboard 2: Time Analysis

- Visuals:
    - Area chart of bookings by 10-minute pickup intervals.
    - Line chart of bookings by day name.
    - Heatmap (matrix) of bookings across hours of the day and days of the week.
- Dynamic Title: Implemented using SELECTEDVALUE to adjust chart titles based on selected KPI.

Dashboard 3: Details Tab

- Provided a grid view of trip-level data (Trip ID, Pickup Date, Payment Type, Passengers, Distance, Value, Pickup and Drop-off Locations).
- Enabled drill-through functionality from other dashboards for detailed exploration.
- Added bookmarks for switching between filtered drill-through view and full dataset view.

Phase 6: Enhancements

- Dynamic Titles: Adjusted chart headings based on KPI selection for improved readability.
- Bookmarks: Implemented a “Data Details” bookmark to show metric definitions and context.

- Clear Filters Button: Added to reset all slicers at once.
- Export Option: Configured to allow download of raw data for external analysis.
- Conditional Formatting: Applied in vehicle type matrix to highlight high and low performance.

Phase 7: Insights and Outcomes

- Booking Trends: Identified variations in daily, weekly, and hourly bookings.
- Revenue Analysis: Provided clarity on revenue generation by payment type and trip type.
- Location Insights: Pinpointed the most common pickup and drop-off points, including long-distance trips.
- Operational Efficiency: Average trip times and distances highlighted customer travel patterns.
- Customer Preferences: Vehicle type and payment type breakdowns revealed behavioral insights.

Phase 8: Conclusion

The Uber Trip Analysis Dashboard successfully integrated transactional data with contextual dimensions, enabling stakeholders to explore booking patterns, revenue trends, and trip efficiency. By leveraging Power BI features such as dynamic measures, drill-through, bookmarks, and conditional formatting, the project delivered both high-level insights and granular detail.

This analysis can assist decision-makers in optimizing pricing strategies, allocating resources more effectively, and improving customer satisfaction through better understanding of ride demand and patterns.