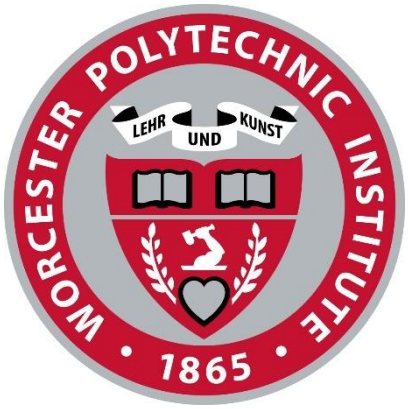
Uber Ride Analytics

Driving A Successful Business Intelligence Approach



Final Project Report

MIS-584 Business Intelligence

Group 01
Rama Raju Kanumuri
Veronica Madepogu
Shahbaz Shah Khan
Keerthan Tamil Selvan
Sai Ajethaa Manickka Subramanian

Professor Daniel Treku

Fall 2024

Table of Contents

(*All the sections below are hyperlinked to the respective pages*)

Section 1: Executive Summary and Project Overview 02

- Problem Statement 03
- Project Relevance 03
- Project Goals 04
- Problems Addressed 05
- Project Deliverables 07

Section 2: Introduction 06

- Company Background 06
- Data Warehouse 08
- Integrated Information Systems 08
- Specific BI Tools used 09
- BI Maturity Model 10
- Summary of Case Studies Presented 10

Section 3: Proposed BI Solution 12

- Data Warehouse Structure (Star Schema) 13
- Role of Proposed BI Solution 16

Section 4: Three Use Cases / Prototypes 20

- Protype 01 20
- Protype 02 26
- Social Media Analytics 34

Section 5: Implementation and Recommendations 38

- Managerial Aspects 38
- Technical Aspects 38

- Ethical Aspects ……. 39
- Phases of Implementation ……. 40
- Kotter's Model ……. 41

Section 6: Summary and Conclusion …… 43

References and Datasets used …… 44

Executive Summary

Uber, as a leading ride-hailing platform, faces ongoing challenges in balancing operational efficiency, customer satisfaction, and market dynamics. Trips between cities and airports represent a significant revenue opportunity but are plagued by issues such as driver unavailability, ride cancellations, and mismatched supply and demand. These challenges not only impact Uber's revenue but also its reputation as a reliable service provider.

Problem Statement:

Uber's primary challenges include frequent driver unavailability, especially at airports during evening hours, and high ride cancellation rates, particularly for city pickups in the morning. These issues are compounded by persistent supply-demand mismatches, which result in significant revenue leakage and unmet customer expectations. As less than 50% of requests are fulfilled during most timeframes, Uber's reliability is at risk, potentially driving customers toward competitors.

This project leverages advanced data analytics and interactive dashboards to identify root causes and provide actionable solutions for these issues. By analyzing ride data, the initiative aims to uncover key patterns and trends, enabling Uber to optimize its operations, enhance driver performance, and improve the customer experience. The project outcomes focus on addressing critical inefficiencies while empowering stakeholders with tools to make data-driven decisions in real time.

The **Airport-City Trips Dashboard** highlights specific challenges for trips between cities and airports, such as high ride cancellations during morning city pickups and driver unavailability during evening airport pickups. The **Operational Performance Dashboard** provides a broader perspective on overall revenue trends, driver contributions, and ride efficiency metrics. Together, these dashboards deliver actionable insights to tackle operational challenges and improve decision-making across Uber's service areas.

Project Relevance

This project is vital to Uber's strategic goals. It directly addresses operational inefficiencies, enhances customer satisfaction, and aligns with Uber's need for dynamic market adaptability. By solving issues such as cancellations, driver unavailability, and supply-demand gaps, the project empowers Uber to:

- Improve operational efficiency by optimizing resource allocation and driver productivity.
- Enhance reliability and customer satisfaction by addressing key pain points in service delivery.
- Maximize revenue potential by leveraging data-driven recommendations such as dynamic pricing, driver incentives, and pre-booking options.
- Maintain competitiveness in the rapidly evolving ride-hailing industry.

By addressing these challenges with a data-driven approach, this project ensures Uber's ability to meet customer expectations, improve service reliability, and capitalize on untapped revenue opportunities.

Goals of the Project:

1. **Boost Operational Efficiency**:
    - Analyze driver distribution and trip efficiency to minimize idle time and enhance productivity.
    - Address gaps in driver availability during high-demand times, such as mornings in city areas and evenings at airports.
2. **Enhance Customer Satisfaction**:
    - Identify and mitigate causes of ride cancellations and delays.
    - Propose solutions to ensure timely and reliable service delivery for passengers.
3. **Understand Market Dynamics**:
    - Examine demand patterns across locations and timeframes to predict customer needs.
    - Optimize pricing strategies and implement incentives to align with geographic and temporal trends.

Problems Addressed:

1. **Driver Unavailability**:
    - Nonavailability of cars, especially at airports during evening hours, leads to missed revenue opportunities and customer dissatisfaction.
2. **High Ride Cancellation Rates**:
    - Cancellations, particularly for city pickups in the morning, frustrate customers and undermine trust in Uber's services.
3. **Supply-Demand Mismatch**:
    - Persistent gaps in supply and demand throughout the day reduce service efficiency and overall revenue.
4. **Revenue Leakage**:
    - Unserved requests and inefficiencies result in significant missed revenue opportunities, especially for high-value airport-city trips.

Project Deliverables:

1. **Trend Analysis Report**:
    - Detailed findings on demand patterns, cancellation trends, and driver performance.
    - Insights into peak demand times and high-traffic areas to inform operational strategies.
2. **Interactive Dashboards**:
    - The **Airport-City Trips Dashboard** visualizes challenges and trends for trips between cities and airports, such as cancellations and nonavailability.
    - The **Operational Performance Dashboard** provides insights into revenue optimization, trip efficiency, and driver contributions.
    - Both dashboards feature dynamic visuals such as heatmaps, bar charts, and time-series trends for real-time monitoring.

3. **Strategic Recommendations**:
    - Targeted strategies, including:
        - Driver incentives for critical times and locations.
        - Penalties for unnecessary cancellations to improve reliability.
        - Promotion of pre-booking options to enhance availability.
        - Geographic service area adjustments to address supply-demand gaps.
4. **Customer and Driver Experience Insights**:
    - Analysis of factors impacting satisfaction, such as wait times, cancellation rates, and availability issues.
    - Recommendations to enhance the overall experience for both passengers and drivers.

INTRODUCTION

Company Background

Name: Uber Technologies, Inc.

Industry: Transportation and Technology

Uber Technologies, Inc., founded in 2009 by Garrett Camp and Travis Kalanick, is a pioneer in the ride-hailing industry and has become one of the most recognizable technology companies globally. The company revolutionized urban mobility by connecting riders with drivers through its innovative mobile application, enabling convenient and efficient transportation services.

Products and Services:

Uber offers a diverse portfolio of products and services designed to cater to a wide range of customer needs:

- **Ride-Hailing Services**:

  Uber's flagship service allows customers to book rides across various categories, including:

  - **UberX**: Affordable everyday rides.
  - **Uber Comfort**: Rides in more spacious, premium vehicles.
  - **Uber Black**: Luxury rides with professional drivers.
  - **UberPool**: Shared rides for cost savings (available in select markets).
- **Food Delivery**:
  - **Uber Eats**: A leading online food delivery platform connecting customers with restaurants and couriers.
- **Freight Services**:
  - **Uber Freight**: A logistics platform that matches carriers with shippers, streamlining freight transportation.
- **Micro-Mobility Options**:
  - Bike and scooter rentals available in select urban markets to promote eco-friendly transportation alternatives.
- **Advanced Mobility Solutions**:
  - Initiatives in autonomous vehicle development and partnerships with aviation companies to explore urban air mobility.

Geographic Market Area:

Uber operates in over 900 metropolitan areas across six continents, offering localized services tailored to market-specific demands. Its global presence makes it a significant player in urban transportation, with a strong foothold in major cities worldwide, including New York, London, Mumbai, and Sydney.

Company Size:

As of 2023, Uber employs approximately **30,400 people** globally. In addition to its employees, Uber has a vast network of independent drivers who partner with the company to provide services. This hybrid workforce model enables Uber to scale its operations efficiently.

Business Model:

Uber operates a platform-based business model, generating revenue through commissions on rides, delivery services, and freight transactions. The company's ability to dynamically adjust pricing based on demand (surge pricing) has been a key factor in optimizing revenue.

Structure:
Uber's organizational structure is divided into several business units to ensure focus and efficiency:

- *Rides*: Core ride-hailing services.
- *Uber Eats*: Food delivery services.
- *Freight*: Logistics and freight matching services.
- *Advanced Technology Group*: Focused on research and development, including autonomous vehicle technology. The company employs a centralized yet flexible management approach, with regional teams adapting strategies to local market conditions.

Scope of the Project:

This project focuses on the **Data Analytics Department** within the **Rides Business Unit**. The objective is to analyze operational inefficiencies related to airport-city trips and provide actionable insights through advanced Business Intelligence tools. The project aims to improve Uber's operational performance, customer satisfaction, and revenue optimization in this critical service area.

Market Position and Competitive Landscape:

Uber is a market leader in the ride-hailing industry, competing with companies like Lyft (North America), Bolt (Europe), and Grab (Southeast Asia). Its innovative approach, driven by advanced analytics and technology, has enabled Uber to maintain its competitive edge despite fierce market competition.

Commitment to Innovation:

Uber is deeply committed to innovation, with significant investments in:

- *Autonomous Vehicles*: Collaborations with automotive companies to develop self-driving technology.
- *Green Initiatives*: Transitioning to electric vehicles and promoting sustainable transportation.
- *AI and Data Analytics*: Leveraging AI and BI tools to enhance operations and customer experiences.

---

1. Use of MS Excel

While Uber may use MS Excel for small-scale or ad-hoc analysis, the scale of its operations demands more robust tools. Excel is not the primary tool for BI or analytics due to the complexity, volume, and velocity of data generated by Uber's global operations. Instead, Uber has transitioned to sophisticated platforms that offer better scalability, automation, and real-time capabilities.

---

2. Data Warehouses

Uber employs state-of-the-art data warehousing solutions to handle the vast amounts of data generated by its operations. Key components include:

- **Presto**:
  - An open-source SQL engine optimized for querying massive datasets across Uber's distributed data architecture.
  - Supports Uber's Extract-Transform-Load (ETL) processes and enables federated queries for seamless access to multiple data sources.
- **Apache Pinot**:
  - A real-time analytics platform used for low-latency queries.
  - Supports Uber's operational dashboards by providing instant access to metrics such as order status, driver availability, and cancellation trends. These data warehousing solutions are vital for managing and analyzing ride data, driver metrics, customer feedback, and external datasets like weather and traffic.

---

3. Integrated Information Systems

Uber's information systems are highly integrated to ensure seamless data flow across business functions:

- Integration with external APIs, such as:
  - **Google Maps** for real-time traffic updates.
  - **OpenWeather** for weather data affecting surge pricing and demand forecasts.
- Internally, Uber's systems enable real-time synchronization of ride data, payment details, and customer feedback to provide a holistic view of operations.

Although Uber has not disclosed its use of ERP systems, its operational and transactional systems demonstrate the characteristics of a highly integrated ERP-like environment. This integration allows Uber to streamline operations and enable data-driven decision-making at all levels.

---

4. Dedicated BI Team

Uber has a specialized Business Intelligence (BI) team responsible for turning raw data into actionable insights. This team collaborates across departments, including operations, marketing, customer service, and driver management, to align KPIs with strategic objectives. Their key responsibilities include:

- **Dashboard Development**:

  Designing interactive dashboards for stakeholders, enabling real-time monitoring of KPIs such as cancellations, revenue, and driver utilization.

- **Advanced Analytics**:

  Developing predictive models to forecast demand and optimize resource allocation.

- **Stakeholder Engagement**:

  Ensuring that insights are actionable and aligned with the needs of operations, customer service, and other business units.

---

5. Specific BI Tools

Uber employs a combination of proprietary and third-party BI tools to meet its analytical needs:

- **In-House Tools**:

    - **Michelangelo**: Uber's machine learning platform used for predictive and prescriptive analytics.
    - **Gairos**: A data visualization tool designed for custom dashboard development.
- **Third-Party Tools**:
    - **Tableau**: Used for dynamic and interactive dashboard creation.
    - **Microsoft Power BI**: Provides user-friendly, customizable visualizations for cross-functional teams.

---

6. Development or Purchase of BI Systems

Uber has heavily invested in developing in-house BI systems tailored to its unique operational requirements. These proprietary platforms allow Uber to:

- Scale its analytics across regions and service lines.
- Customize insights for various teams and stakeholders.

While in-house systems form the core of Uber's BI infrastructure, third-party tools like Tableau are used to enhance data visualization and usability.

---

7. BI Maturity Model

Uber is at an **advanced stage** of the BI maturity model, characterized by:

- **Scalable Infrastructure**: Sophisticated systems such as Presto and Apache Pinot enable Uber to process vast datasets in real-time.
- **Predictive and Prescriptive Analytics**: Machine learning models are used to forecast demand, optimize pricing, and improve driver allocation.
- **Real-Time Monitoring**: Operational dashboards provide real-time insights into key metrics, supporting agile decision-making.
- **Comprehensive Integration**: External data sources, such as traffic and weather, are seamlessly integrated with internal systems to provide actionable insights.

Summary of Case Studies Presented

As part of the initial presentation, two case studies provided critical insights into the effective use of BI in transportation and ride-hailing:

1. **Case Study 1: Gett – Leveraging BI for Corporate Transportation**

- **Overview**: Gett, a UK-based ride-hailing service, emphasized efficiency and reliability for corporate clients.
- **Key BI Practices**:
  - Developed user-centric dashboards to monitor KPIs such as ride completion rates and driver availability.
  - Used real-time analytics to identify operational inefficiencies and improve service delivery.
- **Relevance to Uber**: Demonstrates the importance of real-time monitoring and customized dashboards in addressing operational challenges.

2. **Case Study 2: Lyft – Comprehensive BI Strategy**
   - **Overview**: Lyft implemented a robust BI strategy to improve operational efficiency and customer satisfaction.
   - **Key BI Practices**:
     - Integrated customer feedback into dashboards to align service improvements with customer needs.
     - Used predictive analytics to optimize resource allocation during peak demand periods.
   - **Relevance to Uber**: Highlights the critical role of BI in enhancing customer satisfaction and managing demand fluctuations.

Proposed BI Solution

What is Business Intelligence?

Business Intelligence (BI) refers to the processes, technologies, and tools that organizations use to collect, analyze, and visualize data to make informed decisions. BI enables businesses to transform raw data into actionable insights, driving efficiency, innovation, and strategic growth. The ultimate goal of BI is to improve decision-making by presenting historical, current, and predictive data in an accessible and interactive format.

Main Components of a BI Solution

A robust BI solution consists of several interconnected components, as outlined in the BI Framework (Figure 1). These components include:

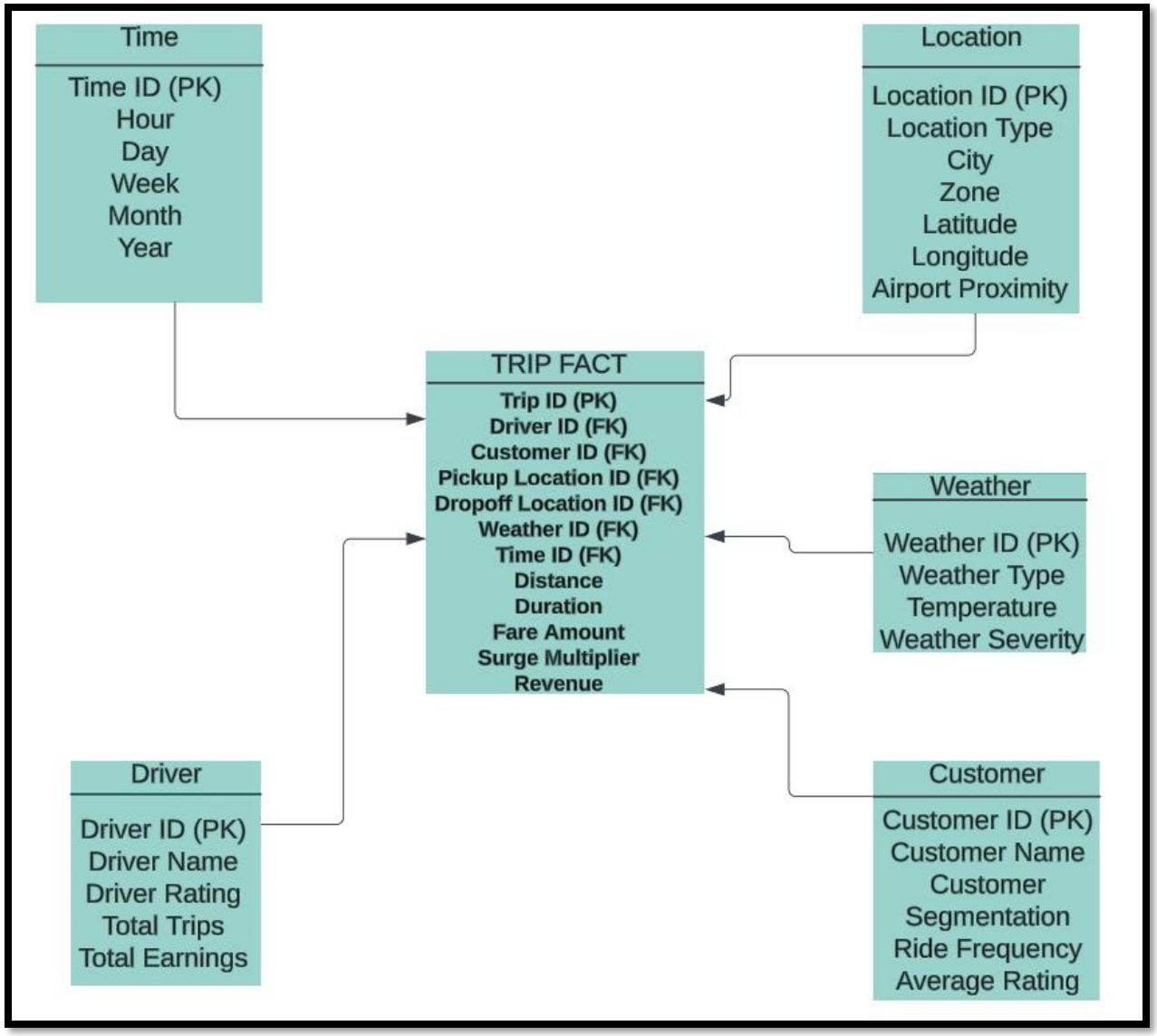
1. **Data Sources**:
    - **Operational Data**: Internal data collected from Uber's ride transactions, driver performance, customer feedback, and real-time GPS tracking.
    - **External Data**: Information from third-party APIs, such as Google Maps (traffic), OpenWeather (weather conditions), and public transportation data.
2. **ETL Process (Extraction, Transformation, and Loading)**:
    - **Role**: Extracts raw data from multiple sources, transforms it into a standardized format, and loads it into a data warehouse.
    - **Value for Uber**: Ensures the integration of ride data, external traffic/weather inputs, and financial records into a unified repository for analysis.
3. **Data Warehouse**:
    - **Role**: A central repository that stores integrated, cleaned, and organized data.
    - **Value for Uber**: Enables advanced analytics, historical trend analysis, and support for scalable operations. For Uber, tools like Apache Pinot and Presto are key data warehouse components.
4. **Data Mart**:
    - **Role**: Smaller, department-specific subsets of the data warehouse designed for targeted analysis.
    - **Value for Uber**: Tailors insights for specific teams, such as operations, marketing, and customer service. For instance, a data mart for the operations team can focus on supply-demand gaps and ride efficiency.
5. **Query and Reporting**:
    - **Role**: Provides a mechanism to retrieve specific insights from the data warehouse using SQL queries and reporting tools.

- **Value for Uber**: Enables on-demand data retrieval for ride cancellations, revenue trends, and peak demand patterns.

6. **Data Visualization**:
   - **Role**: Transforms data into interactive dashboards and charts.
   - **Value for Uber**: Improves stakeholder understanding by presenting insights visually. Dashboards like the "Airport-City Trips" and "Operational Performance" dashboards offer real-time updates on critical KPIs.
7. **Monitoring and Alerting**:
   - **Role**: Tracks metrics in real-time and triggers alerts for deviations from predefined thresholds.
   - **Value for Uber**: Supports real-time incident response, such as alerting the operations team when driver availability drops below a critical level in high-demand areas.
8. **Data Analytics**:
   - **Role**: Includes descriptive, diagnostic, predictive, and prescriptive analytics.
   - **Value for Uber**: Predicts demand spikes during events or adverse weather conditions, recommends dynamic pricing strategies, and evaluates driver performance.

Data Warehouse (Star Schema)

A **star schema** is an effective data warehouse design for supporting Uber's BI needs. It organizes data into a central fact table surrounded by dimension tables, enabling efficient querying and analytics for operational and strategic purposes. Below is a tailored **star schema** for Uber, specifically addressing the airport-city trip analysis and operational performance needs:



Fact Table

Name: Trip_Fact

Dimension Tables

1. **Time Dimension**
   - **Table Name**: Time_Dimension
   - **Purpose**: Provides temporal attributes for analysis, such as peak hours and seasonal trends.

2. **Location Dimension**
   - **Table Name**: Location_Dimension

- **Purpose**: Tracks pickup and drop-off locations for geographic insights.

3. **Driver Dimension**
   - **Table Name**: Driver_Dimension
   - **Purpose**: Provides detailed information about drivers.

4. **Customer Dimension**
   - **Table Name**: Customer_Dimension
   - **Purpose**: Stores rider-specific information for segmentation and analysis.

5. **Weather Dimension**
   - **Table Name**: Weather_Dimension
   - **Purpose**: Tracks weather conditions during the trip.

Relevance of the Star Schema

The star schema supports Uber's BI solution by:

1. **Facilitating Efficient Queries**: The schema is optimized for OLAP (Online Analytical Processing), enabling quick analysis of ride data, trends, and KPIs.
2. **Providing Granular Insights**: Allows multidimensional analysis across time, location, driver, and customer dimensions.
3. **Supporting Dashboard Development**: Feeds real-time dashboards with relevant data for monitoring operations, revenue, and customer satisfaction.
4. **Scalability**: The schema can handle large datasets while maintaining simplicity for end-users.

Monitoring, Data Analytics, and Visualization

1. **Monitoring**:
   - **Operational Dashboards**: Display real-time metrics such as driver availability, ride cancellations, and surge pricing rates. For example:
     - A **supply-demand gap dashboard** can alert operations when the demand exceeds available drivers in specific locations, allowing for immediate corrective action.
   - **Strategic Dashboards**: Track long-term trends such as monthly revenue growth and customer satisfaction rates. These dashboards can guide high-level decision-making.
2. **Data Analytics**:
   - **Descriptive Analytics**: Provides historical insights into ride demand patterns, revenue performance, and customer behavior.

- **Predictive Analytics**: Forecasts future demand based on historical data, weather conditions, and event schedules. For instance, predicting higher demand near airports during holiday seasons.
- **Prescriptive Analytics**: Recommends strategies such as adjusting driver incentives during peak hours or introducing targeted promotions to balance supply and demand.

3. **Data Visualization**:
   - **Examples**:
     - Heatmaps displaying ride density across city zones.
     - Bar charts showing driver performance by revenue contribution.
     - Line graphs tracking cancellations and unavailability trends over time.
   - **Value**: Interactive dashboards improve data accessibility for stakeholders, enabling quicker decision-making and more informed strategies.

Role of the Proposed BI Solution

The proposed Business Intelligence (BI) solution plays a critical role in addressing Uber's operational inefficiencies, enhancing decision-making, and driving value across the organization. Below are the specific roles that the BI solution serves, tailored to Uber's needs:

1. Operational Optimization

- **Problem Addressed**: Persistent supply-demand mismatches, high driver unavailability, and inefficient resource allocation.
- **Role of BI**:
  - Monitors real-time metrics like driver availability, trip efficiency, and cancellation rates.
  - Enables predictive analytics to forecast demand spikes (e.g., during events or weather changes).
  - Recommends optimal driver allocation and dynamic pricing strategies to align supply with demand.
- **Impact**: Ensures operational efficiency by minimizing idle time for drivers and reducing customer wait times.

2. Enhanced Decision-Making

- **Problem Addressed**: Lack of actionable insights for stakeholders to make timely and informed decisions.
- **Role of BI**:
  - Integrates data from diverse sources (e.g., ride transactions, external APIs for weather/traffic) into a centralized system.

    - Provides role-specific dashboards for teams (operations, marketing, customer service) to access targeted insights.
    - Facilitates real-time decision-making with up-to-date information.
- **Impact**: Improves responsiveness to operational challenges and empowers strategic planning.

3. Customer Experience Improvement

- **Problem Addressed**: High cancellation rates, long wait times, and inconsistent service quality.
- **Role of BI**:
    - Analyzes customer feedback and satisfaction scores to identify pain points in the rider experience.
    - Tracks key metrics like average wait times, ride ratings, and service reliability.
    - Suggests interventions, such as targeted promotions for dissatisfied customers or pre-booking options for high-demand periods.
- **Impact**: Enhances customer satisfaction and retention by delivering a more reliable and efficient service.

4. Revenue Optimization

- **Problem Addressed**: Revenue leakage due to unfulfilled ride requests and inefficient pricing models.
- **Role of BI**:
    - Tracks revenue per trip, revenue per driver, and surge pricing effectiveness through dashboards.
    - Identifies high-revenue opportunities, such as peak hours and underutilized geographic zones.
    - Supports dynamic pricing adjustments to maximize earnings during demand surges.
- **Impact**: Increases profitability by reducing revenue losses and optimizing pricing strategies.

5. Strategic Planning and Growth

- **Problem Addressed**: Limited long-term insights into market trends and business expansion opportunities.
- **Role of BI**:
    - Analyzes historical data to identify growth opportunities in under-served markets or underutilized services.
    - Tracks market dynamics, such as customer demographics, demand trends, and competitor activity.
    - Supports planning for new initiatives, like expanding micro-mobility services or enhancing airport-city ride offerings.

- **Impact**: Drives informed decisions on business expansion and long-term investments.

6. Real-Time Monitoring and Alerting

- **Problem Addressed**: Delays in identifying and addressing critical issues such as driver shortages or service disruptions.
- **Role of BI**:
    - Real-time dashboards monitor KPIs such as supply-demand gaps and service outages.
    - Automated alerting systems notify stakeholders of deviations from performance thresholds (e.g., driver availability below critical levels).
- **Impact**: Enables immediate corrective actions to mitigate operational disruptions.

7. Departmental Collaboration and Accountability

- **Problem Addressed**: Disjointed data silos and lack of visibility across departments.
- **Role of BI**:
    - Provides integrated data views to foster collaboration among operations, finance, marketing, and customer service teams.
    - Tracks department-specific KPIs to ensure accountability and alignment with organizational goals.
- **Impact**: Creates a unified framework for decision-making and performance evaluation across departments.

8. Competitive Advantage

- **Problem Addressed**: Intense competition in the ride-hailing industry with players like Lyft and Bolt.
- **Role of BI**:
    - Leverages advanced analytics to provide deeper insights into customer behavior, operational efficiency, and market trends.
    - Supports innovation by identifying areas for differentiation, such as targeted customer campaigns or sustainability initiatives.
- **Impact**: Strengthens Uber's market position and ensures sustained competitiveness.

Conclusion

The proposed BI solution acts as a cornerstone for Uber's operational efficiency, customer satisfaction, and revenue growth. By integrating real-time data monitoring, predictive analytics, and actionable insights, the BI solution not only addresses Uber's immediate challenges but also positions the company for long-term strategic success in the dynamic ride-hailing industry.

Value of the Proposed BI Solution for Uber

The proposed BI solution generates significant value across Uber's departments:

- **Operations**: Enhances driver allocation and reduces idle time, ensuring better alignment of supply with demand.
- **Finance**: Tracks revenue metrics in real-time, aiding in financial forecasting and surge pricing optimization.
- **Customer Service**: Analyzes customer feedback to identify areas for service improvement.
- **Marketing**: Supports data-driven campaigns by analyzing customer segments and ride patterns.

By implementing this comprehensive BI solution, Uber can strengthen its decision-making, increase operational efficiency, and improve customer satisfaction. In the next section, three use cases will demonstrate how dashboards and analytics techniques can be applied to achieve these objectives.

Three Use Cases or Prototypes

This section demonstrates how Uber can leverage Business Intelligence (BI) tools to solve critical operational and strategic challenges through three use cases. Two use cases involve dashboard prototypes: one focusing on **operational efficiency** (Airport-City Trips Dashboard) and the other on **strategic performance monitoring** (Operational Performance Dashboard). The third use case highlights **social media analytics**, providing insights into customer sentiment and public perception of Uber's services. Together, these prototypes illustrate the power of BI in addressing Uber's challenges and enhancing decision-making.

Use Case 1: Operational Dashboard – Airport-City Trips

(Protype 01)

Using Data to Analyze Uber Trips Between Airport and City

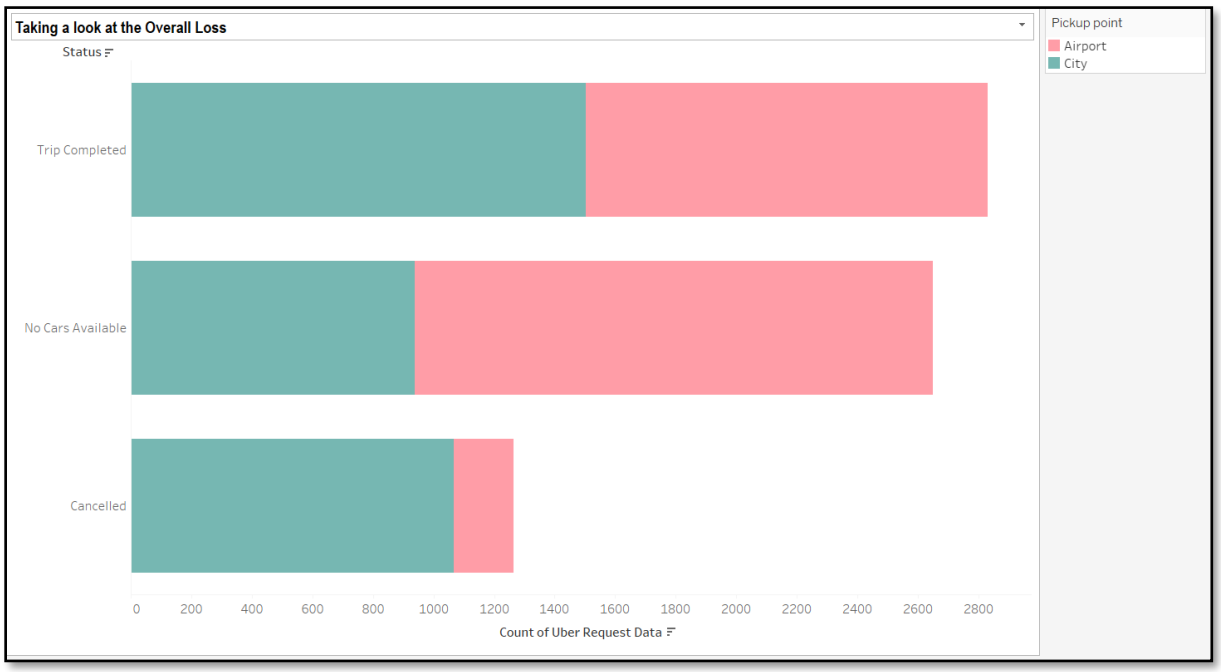
Uber and its Problem

Uber is a popular alternative to traditional taxis, connecting drivers with riders through a simple app. A significant demand for car services exists for trips to and from airports, as people prefer not to leave their cars there. This presents a huge revenue opportunity for Uber, especially with airport rides. However, Uber struggles with a key issue: frequent unavailability of drivers or cancellations for airport trips. As a result, Uber misses out on potential revenue and risks losing customers due to its unreliability. Analyzing Uber's data can help identify the root causes of this problem. By pinpointing when and where these issues occur, Uber can implement targeted solutions to improve reliability and increase revenue.

The Dataset

The dataset used for analysis is a masked data set which is similar to what data analysts at Uber handle. The source of the dataset is a company called upGrad. It contains 6,745 observations of trips between city and airport during regular weekdays of a week in July 2016. So, it is used as a sample set representing Uber trips between city and airport when it is not a holiday or weekend. The data looks at trip requests and tells us whether a request was accepted and completed, accepted and then cancelled by the driver, or was unaccepted due to no cars being available.

Taking a look at the Overall Loss

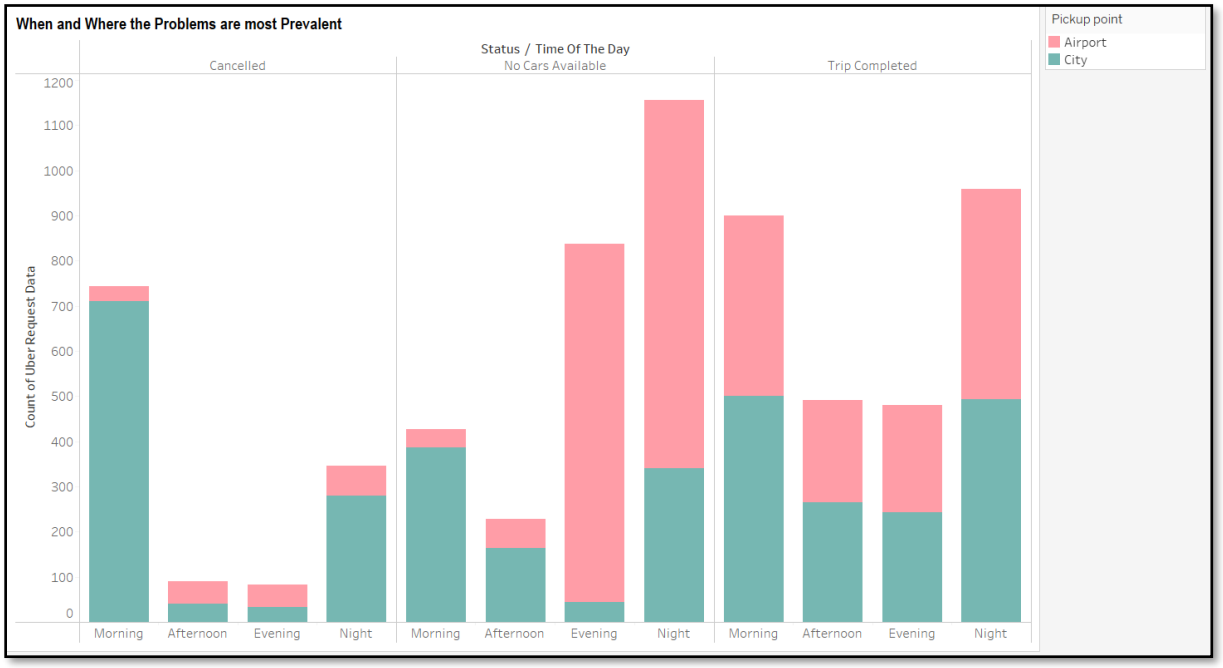


From the bar graph, we see that out of all requested Uber trips, more trips don't end up happening than do end up happening. Uber could be completing over 100% more trips in this category of trips than it does. This means that Uber is missing out on a lot of revenue. This also means that most of the time, Uber can't be relied upon for trips to and from the airport, which is detrimental to Uber's reputation.

From comparing the amount in the "cancelled" category to the amount in "no cars available" category, we see that the bigger issue is the nonavailability of cars. This means that Uber has to focus its efforts into getting more drivers and getting its current drivers to be available more often.

When looking at the "cancelled" category, we see that cancellations by drivers is mainly an issue when the pickup point is the city area. This can be for a number of reasons, although we don't have the data on why the drivers cancel. When looking at the category of "no cars available", we see that non-availability of cars is a significant problem in both the airport and city pickup points. But in terms of count, it is a bigger problem by the airport pickup point.
In conclusion, both cancellations and nonavailability of cars are quite common. Nonavailability of cars is Uber's top issue, particularly by the airport. When analyzing requests with city as the pickup point, cancellations and nonavailability are both big issues. When analyzing requests with airport as the pickup point, the problem of cancellations pales in comparison to the problem of nonavailability.

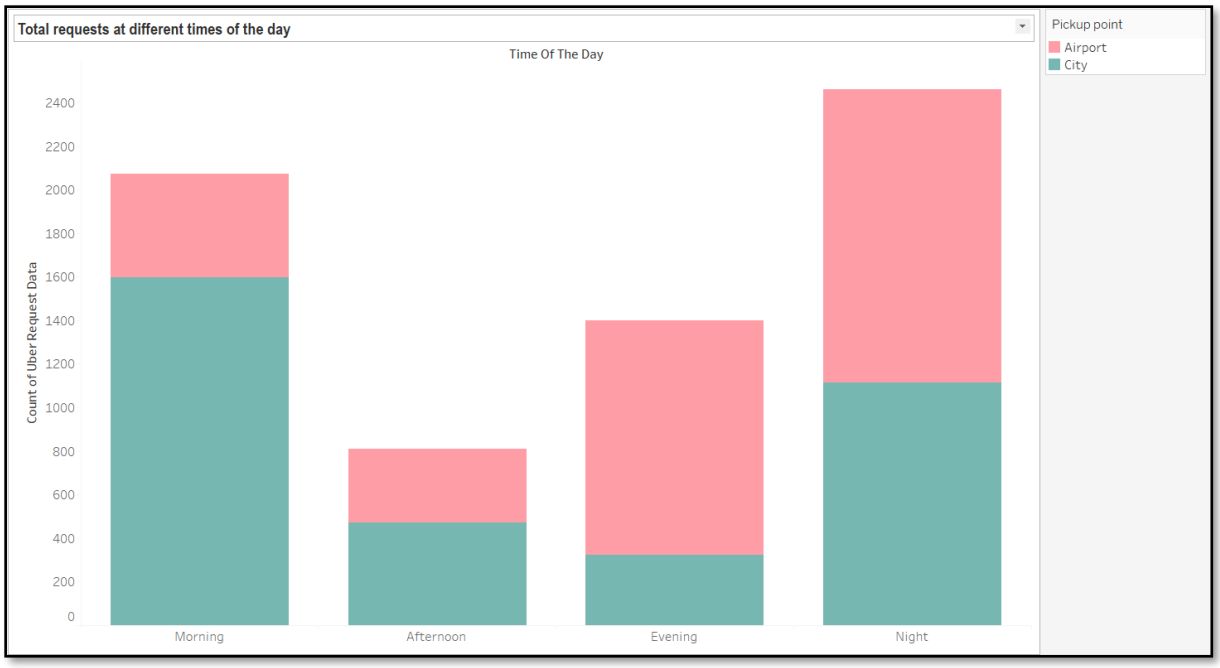
When and Where the Problems are most Prevalent



Here are two bar graphs displaying a closer look at the categories "cancelled" and "no cars available". These graphs incorporate time and place so that we know both where and when the issues are happening most.

We see here that when the day is divided into six timeframes of equal length, for the city, the main timeframe during which requests aren't completed is 4:00 AM-11:59 AM. In particular, cancellations are more common than nonavailabilities, but they both have pretty high counts, which is a finding consistent with our previous analysis on the more generic bar chart. For the airport, the main time of day during which requests aren't completed is 4:00 PM-11:59 PM. Nonavailabilities has a significantly higher count than cancellations, which makes sense based on the previous generic analysis.

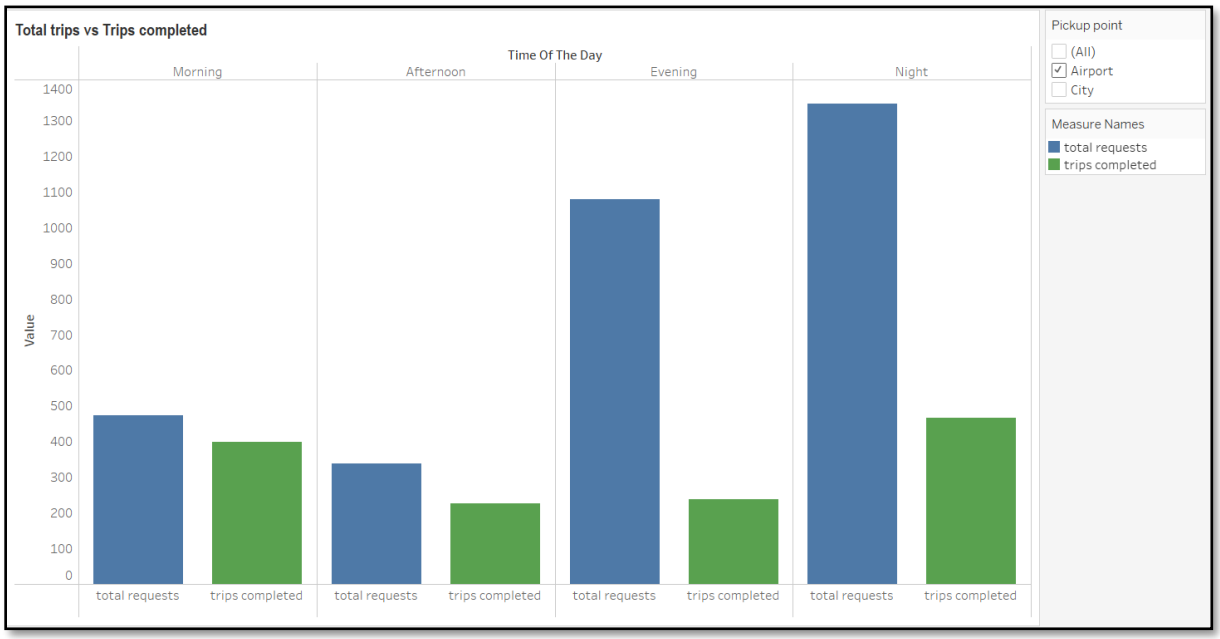
We see that Uber would greatly benefit from more drivers completing trips in the city area during the early mornings until noon and in the airport area during the evenings after 4:00 PM.



We also see, from examining the total requests, including completed trips, that those particular times of day in those particular locations have the highest amounts of requests. Therefore, we can conclude that in general, many drivers are needed during mornings in the city and evenings in the airports.

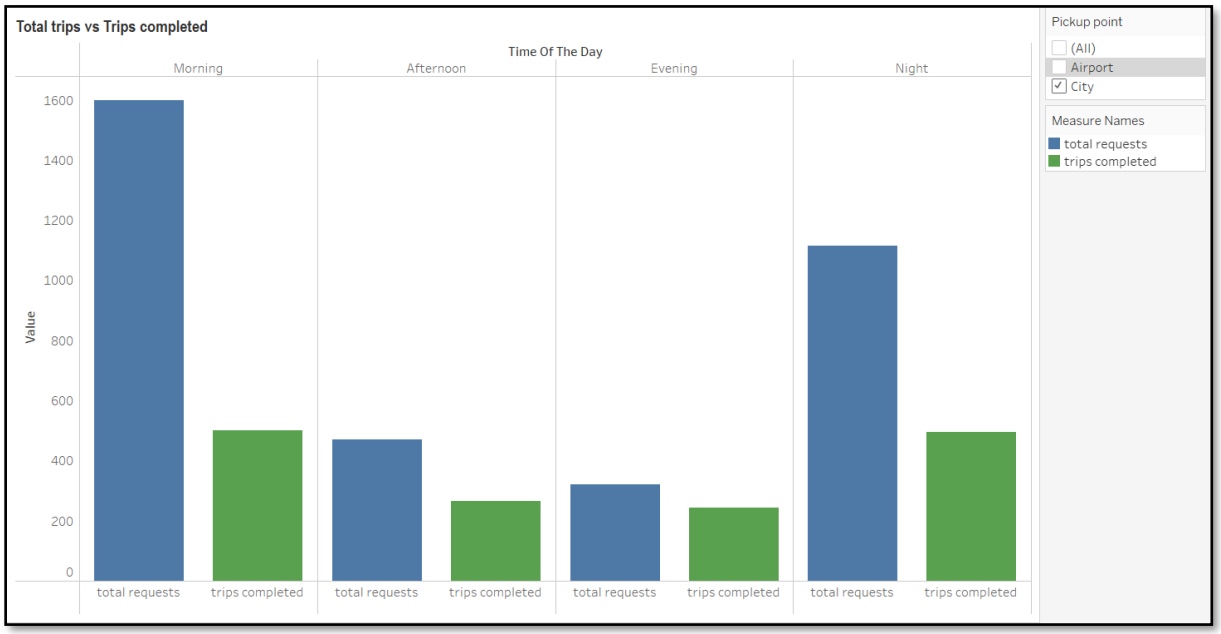
Here is a visualization of the "supply-demand" gap:

Pickup point: airport



Supply V Demand

Pickup point: City



In both charts we see that throughout the day, besides for the 12:00 AM - 3:59 AM timeframe, a pretty similar amount of requests are completed during each time interval. We see that the supply barely correlates with the demand, as a similar amount of requests are completed when the demand is huge as when the demand is small.

It would be beneficial for Uber to emphasize when and where the demand is high to its current drivers and potential new drivers. The supply-demand gap should be more consistently small throughout the day so that a greater percentage of requests in each timeframe throughout the day will be fulfilled. This way, Uber would be closer to gaining as much revenue as possible, and it would be known as more reliable.

Uber's Reliability

As noted earlier, it is very important for Uber to have a reputation for being reliable. If Uber is known to be unreliable for trips between airport and city, people may turn to other car service options before turning to Uber. Additionally, if people are let down by Uber often enough, they may not try Uber again in the future for other types of trips.

Through displaying the proportions of request statuses per timeframe, these pie charts illustrate Uber's reliability for trips between city and airport. We can see that 12:00 PM - 3:59 PM is the only timeframe during which more than 50% percent of requests are completed. This means that for the majority of the day, Uber is pretty unreliable, as less than 50% of requests get completed. Most people end up unsatisfied with Uber during most of the day for this category of trips. It would be greatly beneficial for Uber to increase its reliability. Having frustrated customers is never good for the growth of a company, as the company can lose its customers and not gain new ones.

Using the Analysis for Company Benefit and Growth

This data and analysis can be used by Uber for internal discussion, and also can be used by companies similar to Uber, as it tells companies when drivers are most needed at the airports and when drivers are most needed to bring people to airports from cities. The data and analysis can also provide insights for drivers. (As noted earlier, this analysis applies to weekdays.)
Uber can provides incentives for drivers to fulfill requests in the cities in the morning hours and airports in the evening hours, like bonuses or salary boosts. Uber can also impose penalties on drivers for accepting and then cancelling requests, unless the cancellation is done for a very legitimate reason. Cancellations can frustrate customers even more than no cars being available. It is more of a let-down for a customer to think they have a ride and then be cancelled on, than for them to know right away that there are no cars available.

It is possible that by now, the year 2021, Uber has implemented new policies regarding cancellation. Perhaps there are already, or if not there should be, policies about drivers getting penalties after cancelling a certain percentage of rides they accept. Drivers do need the right to cancel, since they can encounter a customer with a very low rating or a customer that they can't find and get in touch with. The point is that they shouldn't do it often enough that it can hurt Uber's reputation and revenue.

Uber should also try to attract new drivers and create incentives for them to join the company, and particularly incentives to accept trips between airport and city. Uber can attract new drivers by showing that they have so many requests that aren't being fulfilled. People can be attracted to join the company if they know that they're basically guaranteed to get driving jobs.

A solution to improve Uber's reputation for reliability is to promote bookings in advance. Uber actually introduced a feature in 2016 that allows for scheduling rides in advance. When people know that scheduling a car in advance is a choice that they can choose, they can see Uber as more dependable. If they choose to do book in advance, the frustrating scenario of no cars being available, which can turn people away from relying on Uber in general, can be avoided.

Further Analysis Ideas

It would be interesting if there was data on why drivers cancel. It would be beneficial to analyze why drivers cancel; and based on the reasons, come up with solutions to stop these cancellations from happening.

It would also be interesting to analyze how weekend and holiday data differ from weekday data. I would want to know if weekends and holidays have the same or different problems as weekdays and how the magnitudes of the problems compare.

Additionally, I would like to point out that now it is 2021. This analysis is mainly beneficial for the Uber company in 2016 and soon after. I think it would be interesting to analyze Uber's data since then to see if trends have shifted, and if those trends shifted after the introduction of ideas like those I mentioned and/or other changes in the company.

Use Case 2: Strategic Dashboard – Operational Performance

COMPRAHEBSIVE RIDE INSIGHTS DASHBOARD of UBER

Objective: Provide a high-level view of Uber's operational performance, focusing on revenue optimization, driver productivity, and trip efficiency.

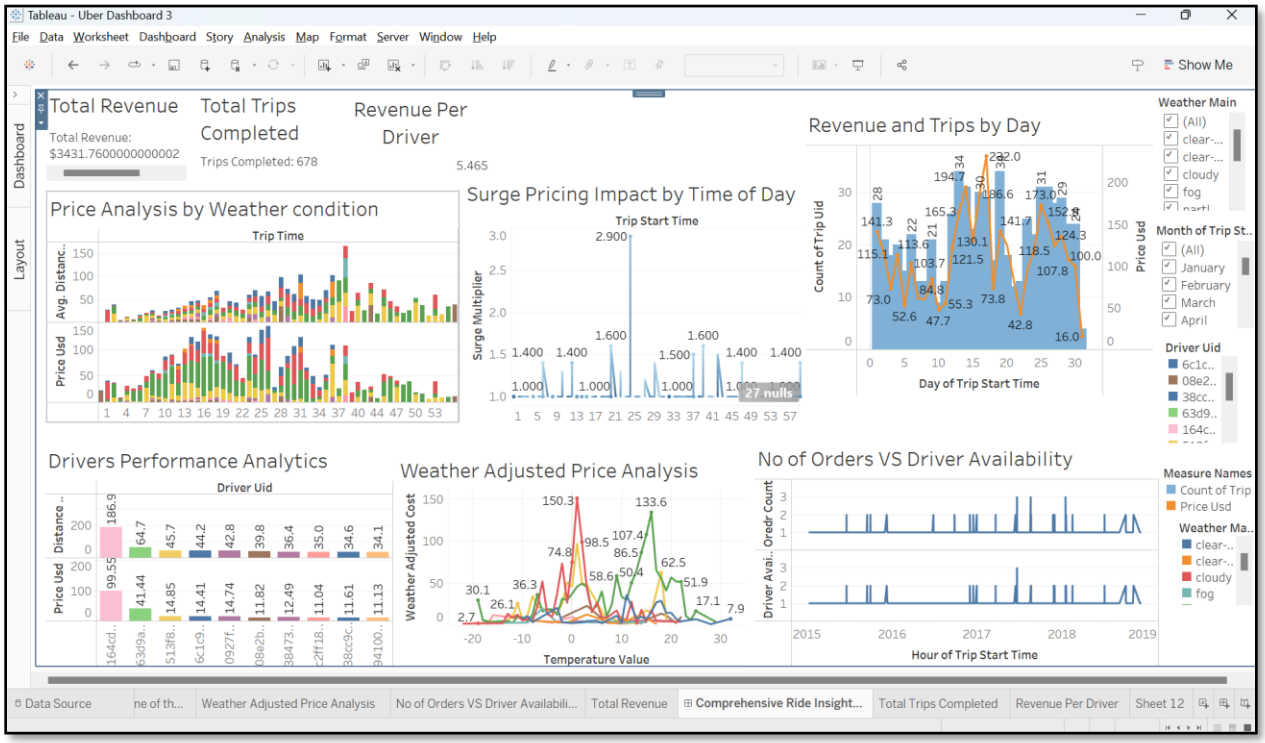
Target Audience:

- **Senior Management**: To track strategic goals such as profitability and market share.
- **Financial Analysts**: To monitor revenue trends and optimize pricing models.

Key Metrics (KPIs):

1. **Total Revenue**: Consolidated earnings across all ride categories.
2. **Revenue per Driver**: Average earnings per active driver.
3. **Trip Efficiency**: Average cost per kilometer for completed rides.
4. **Peak Hour Revenue Contribution**: Percentage of revenue generated during peak demand periods.
5. **Driver Utilization**: Percentage of driver time spent on active rides versus idle time.

The Reference of Dashboard that we created→



The above dashboard of Uber empowers to make data-driven decisions that enhance revenue, optimize operations, and improve both driver and customer experiences. This dashboard includes the below KPI's and managerial questions that is answered by each of the charts.

Purpose of Dashboard About Uber

This dashboard provides a holistic view of Uber's operational and financial performance. It emphasizes the following key insights:

1. **Revenue Trends**:
    - Revenue peaks during specific times (e.g., mornings, evenings, or adverse weather conditions), suggesting opportunities for strategic pricing and promotions.
2. **Driver Management**:
    - Identifying top-performing drivers and aligning availability with demand ensures operational efficiency and customer satisfaction.
3. **Impact of External Factors**:
    - Weather and temperature significantly affect trip costs and duration, necessitating dynamic adjustments to pricing strategies.
4. **Operational Efficiency**:
    - Analyzing trip efficiency by comparing distance, time, and pricing highlights areas for improvement in driver routing and fare structure.

    This dashboard empowers Uber to make data-driven decisions that enhance revenue, optimize operations, and improve both driver and customer experiences.

The charts are→

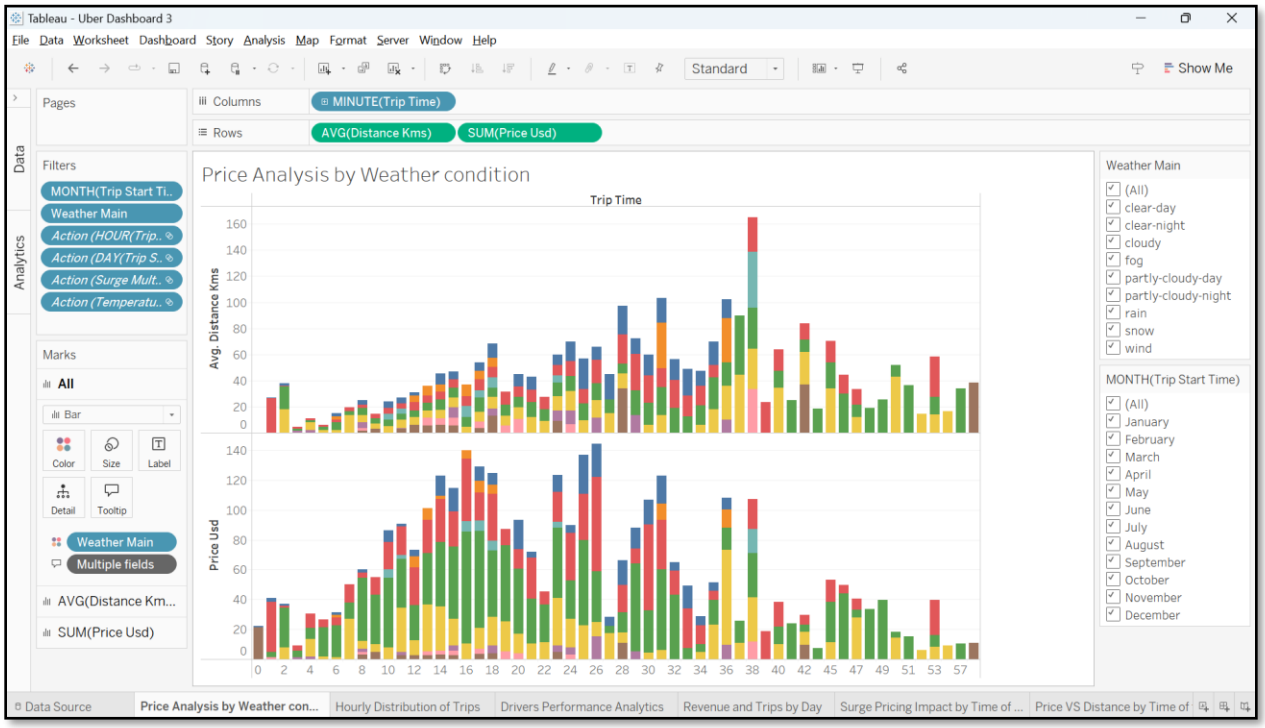
1. **Price Analysis by Weather condition :-**
   This chart highlights how different weather conditions (e.g., cloudy, rainy, foggy) affect ride prices. It also provides contextual details about trip duration and distance.

Managerial questions that are answered with the help of this sheet is→

1. **Which weather conditions result in higher pricing?**
   By sorting weather main by average price (price USD), this chart identifies weather patterns that lead to increased pricing (e.g., rainy days may see higher costs due to demand or driving difficulty)
2. **How does trip duration vary by weather?**
    i. Tooltip details for Weather provide insights into whether certain weather conditions result in longer trips.

- **Insights for Uber**:

- Pricing may need to be adjusted dynamically based on weather to maintain customer satisfaction and ensure driver incentives.



2. Number of Orders vs. Driver Availability

- A dual-axis line chart comparing the number of trips (Order Count) against available drivers over time.

The Managerial Questions that are answered by this char tis:

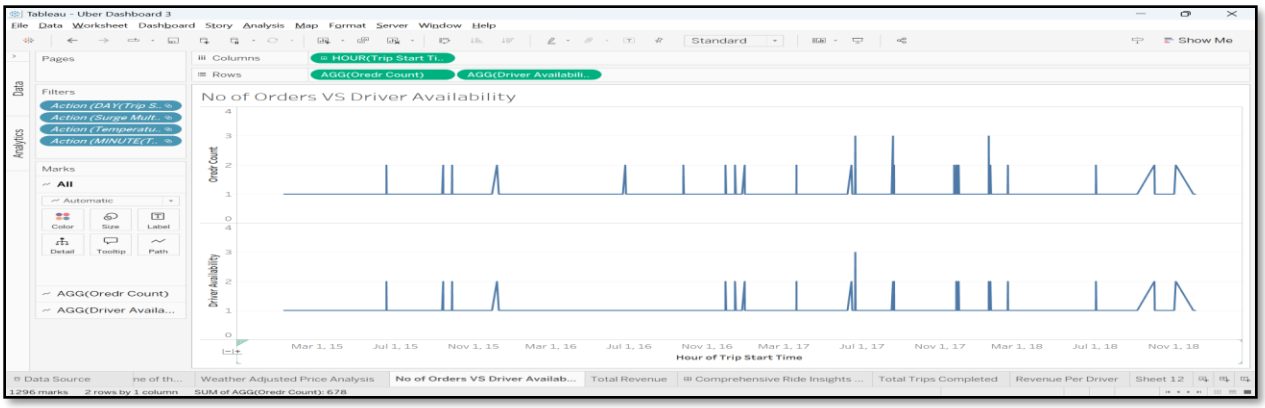
➔ **Are there periods where demand exceeds supply?**

- Peaks in Order Count without a corresponding increase in Driver Availability highlight supply-demand gaps.

➔ **How should Uber adjust driver incentives or hiring?**

ii. Insights from this chart can guide decisions on surge pricing, driver bonuses, or recruitment efforts.

- **Insights for Uber**:
  - Aligning driver availability with demand can minimize missed trips and customer dissatisfaction during peak periods.



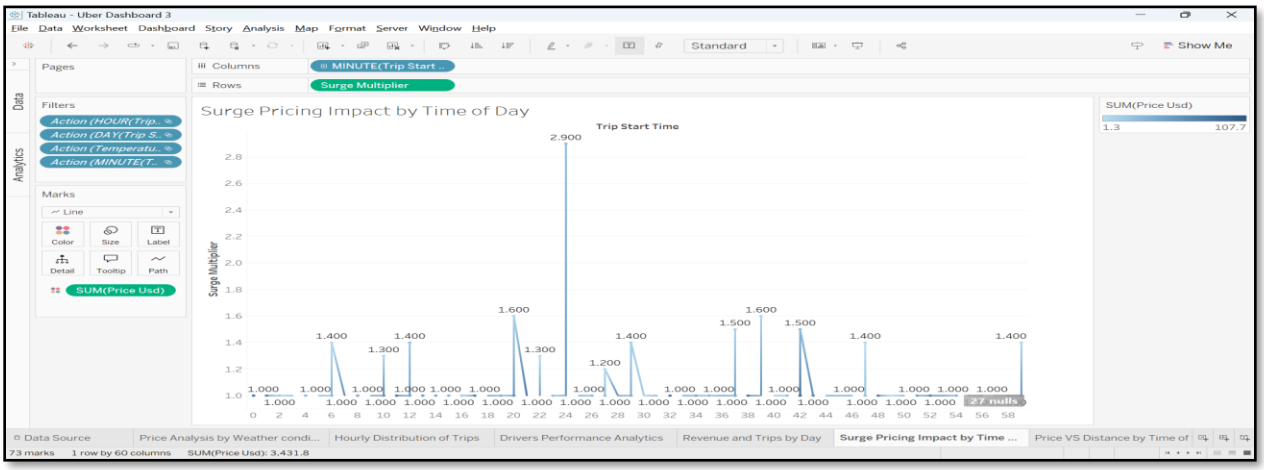
3. Surge Pricing Impact by Time of Day

A line chart showing how surge multipliers fluctuate throughout the day, with price levels encoded by color.

- **How It Answers Managerial Questions**:

1. **What are the most profitable times of day?**
    - High surge multipliers indicate peak hours, where pricing and revenue generation are highest.
2. **How does pricing behavior vary throughout the day?**
    - Patterns in surge multipliers and prices help Uber identify optimal times to encourage driver participation.

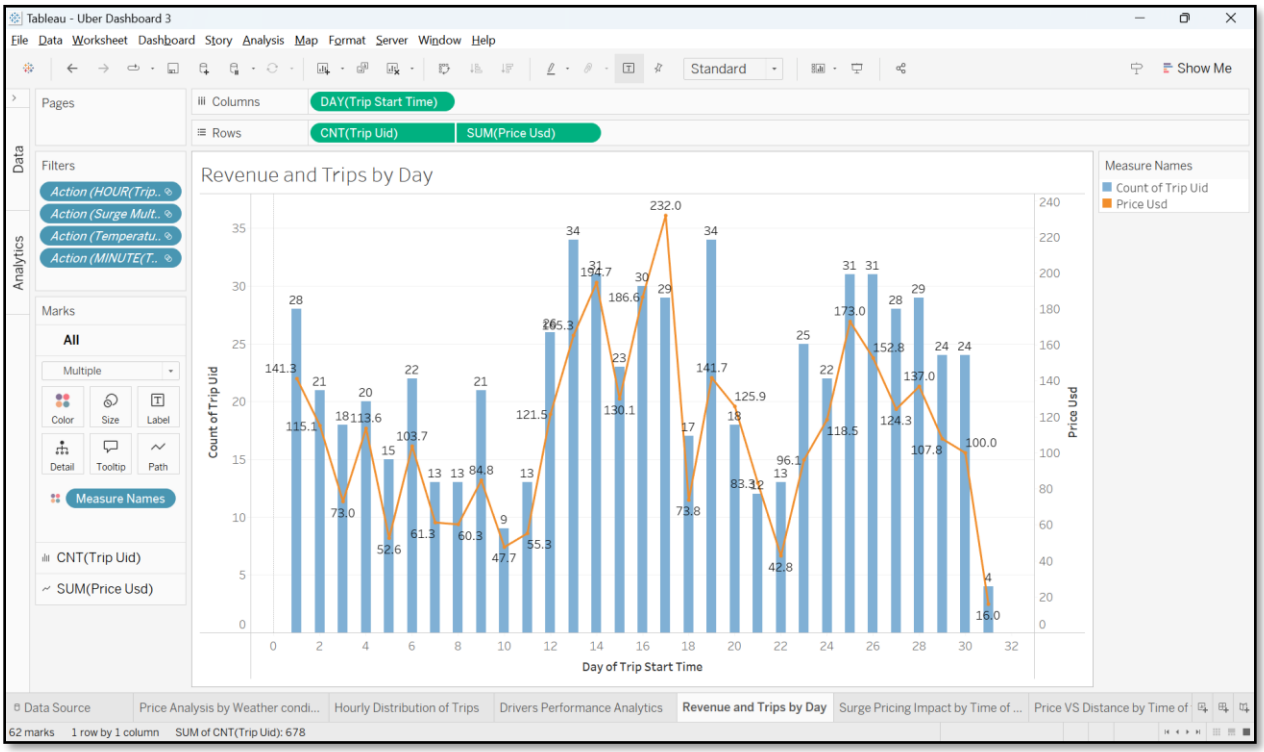
- **Insights for Uber**:
    - Surge pricing is crucial during early mornings and evenings, indicating high customer demand and revenue potential.



4. Revenue and Trips by Day

A dual-axis chart correlating total revenue (price_usd) and trip counts (trip_uid) on a daily basis.

- **How It Answers Managerial Questions**:
  1. **What are the most profitable days?**
     - Peaks in revenue highlight specific days that drive business performance.
  2. **How does trip volume correlate with revenue?**
     - Comparing trip counts and revenue clarifies whether increased trips or higher pricing contributes more to revenue growth.
- **Insights for Uber**:
  - Strategic promotional campaigns can be planned for days with lower revenue to boost trip volumes.

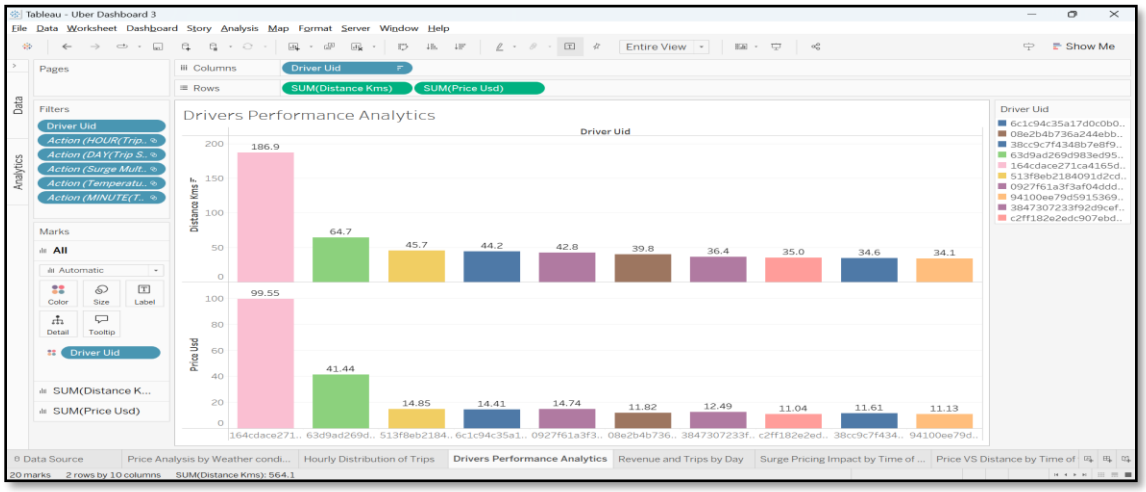


5. Driver Performance Analysis (Revenue and Distance)

A bar chart ranking drivers based on their revenue and distance contributions.

- **How It Answers Managerial Questions**:
  1. **Which drivers contribute the most to revenue?**
     - Top-performing drivers are easily identifiable for potential rewards or incentives.
  2. **What is the relationship between distance traveled and revenue generated?**
     - Comparing distance and revenue per driver helps evaluate efficiency and identify areas for improvement.
- **Insights for Uber**:

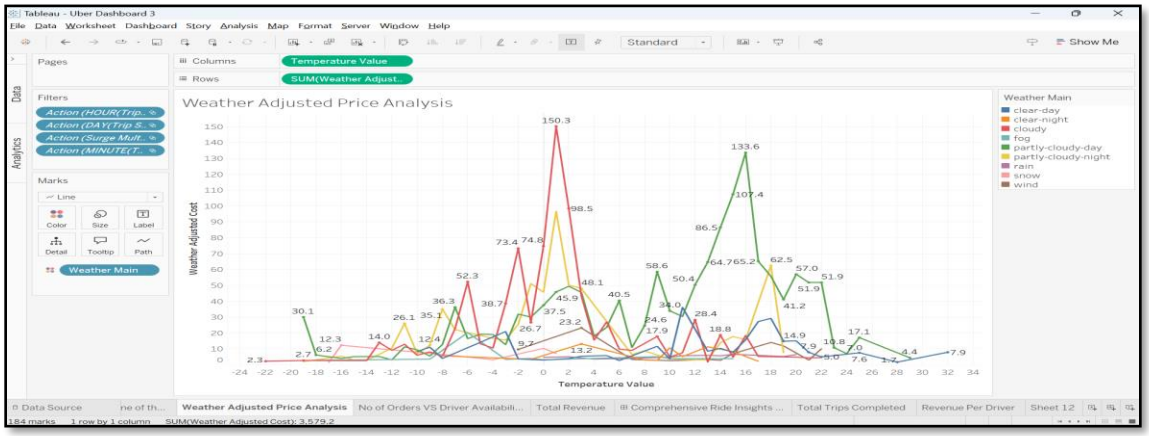
- Recognizing and rewarding top-performing drivers can boost morale and encourage higher productivity.



6. Weather-Adjusted Price Analysis

A line chart analyzing how weather conditions (e.g., temperature) affect adjusted trip prices.

- **How It Answers Managerial Questions**:
  1. **How does temperature affect ride pricing?**
     - Spikes in Weather-Adjusted Cost during extreme temperature ranges (cold or hot) suggest demand fluctuations or operational challenges.
  2. **Does weather severity influence trip efficiency?**
     - Understanding these patterns helps optimize operations during challenging weather conditions.
- **Insights for Uber**:
  - Seasonal or temperature-based pricing strategies could improve profitability while balancing customer satisfaction.



7. Total Revenue, Total Trips Completed, and Revenue per Driver

Summary KPIs at the top of the dashboard:

- Total revenue generated.
- Total trips completed.
- Average revenue per driver.

- **How It Answers Managerial Questions**:
  1. **How is Uber performing overall?**
     - These KPIs provide a snapshot of operational and financial health.
  2. **Are drivers generating sufficient income?**
     - Revenue per Driver ensures fair compensation and operational efficiency.
- **Insights for Uber**:
  - Monitoring these metrics regularly can identify trends and areas needing intervention.

Managerial Questions Answered by the Dashboard

1. **What factors influence ride pricing and revenue?**
   - Charts like "Enhanced Price Analysis by Weather Condition" and "Weather-Adjusted Price Analysis" show how external factors like weather and time of day affect revenue.
2. **How can Uber optimize driver availability?**
   - The "Number of Orders vs. Driver Availability" chart helps align driver supply with customer demand.
3. **When should Uber focus on surge pricing?**
   - "Surge Pricing Impact by Time of Day" pinpoints peak hours for surge pricing to maximize revenue.
4. **Which drivers are the most efficient?**
   - "Driver Performance Analysis" evaluates revenue and distance contributions for each driver.
5. **What are the revenue trends by day?**

"Revenue and Trips by Day" provides a clear picture of daily operational performance.

Use Case 3: Social Media Analytics

Objective: Analyze customer sentiment and public perception of Uber services using data from Twitter and online reviews.

Target Audience:

- **Marketing Team**: To improve customer engagement strategies.
- **Customer Support Team**: To address recurring service-related issues highlighted by customers.

Data Sources:

- **Twitter API**: Collect tweets mentioning Uber within the last six months.
- **Online Reviews**: Extract reviews from app stores (Google Play, Apple App Store).

Key Insights Derived:

1. **Sentiment Analysis**: Categorize tweets and reviews as positive, neutral, or negative.
2. **Recurring Themes**: Identify keywords or topics frequently mentioned (e.g., cancellations, driver behavior, pricing).
3. **Geographic Insights**: Determine regions with the highest concentration of negative or positive feedback.

Analytics Techniques:

- **Natural Language Processing (NLP)**: Used to process text data and extract sentiment scores.
- **Topic Modeling**: Identifies recurring themes, such as complaints about driver cancellations or praise for reliable service.
- **Trend Analysis**: Tracks how sentiment changes over time, particularly after major events like a price hike.

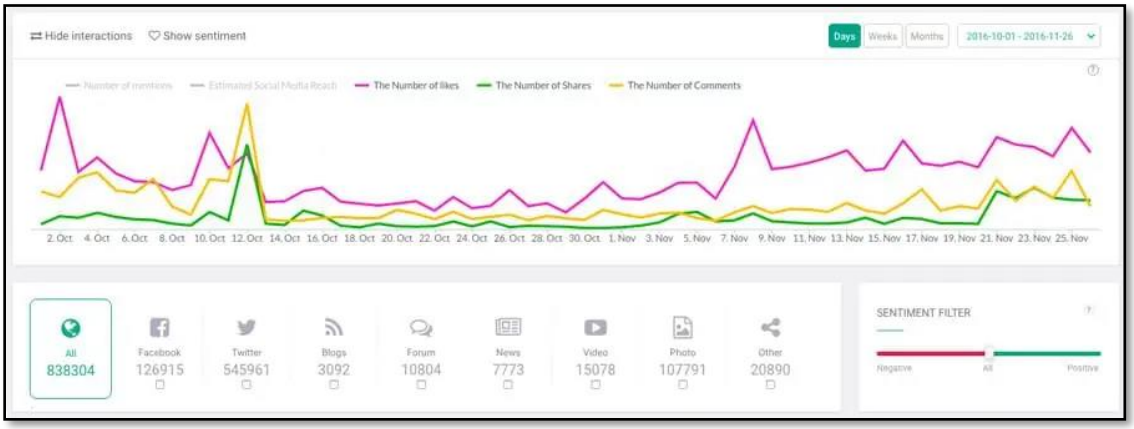
Implications for Uber:

- Negative sentiment spikes during periods of high cancellation rates, particularly near airports.
- Customers often praise drivers for politeness but criticize delays caused by unavailability.
- Insights can guide targeted campaigns, such as offering discounts to customers in regions with low sentiment scores.

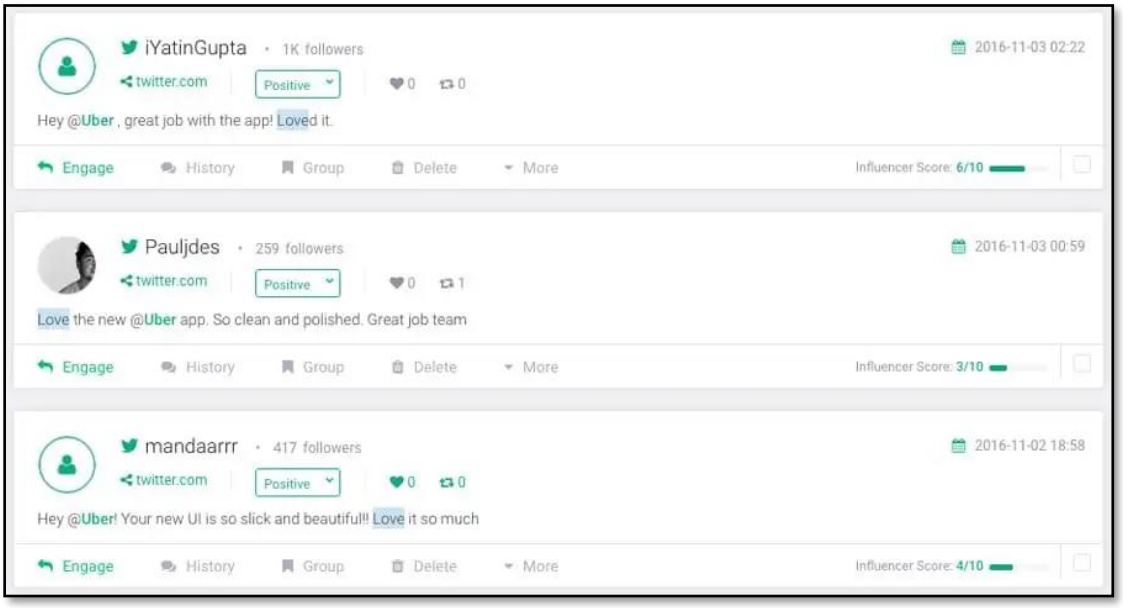
Text Analysis (Sentiment Analysis)

On November 2, 2016, Uber released their new rider app, which was greeted with enthusiasm by users worldwide. Within 2 days of the rollout, the amount of online discussion about Uber increased by 100%, from approximately 8-9k mentions to 16-22k in just the first three weeks following the initial release.

The volume of discussions around Uber didn't increase during the next week or so, but remained high for an extended period of time. Engagement also increased significantly in the days following the introduction of the new app, reaching a peak between November 7 and November 11. As you can see on the graph below, the number of likes (marked in pink) increased by 90% in comparison to peaks from the end of October. Similarly, the number of shares (marked in green) rocketed following the app rollout - they increased by 250% in comparison to the end of October.



While many riders expressed their love for the brand-new app, Uber was also on the lookout for the not-so-happy ones so they could use their feedback in the future.



How Uber uses Social Listening in Product Development

Uber keeps an eye on their driver and rider opinions because this allows them to continuously optimize their product.

Thanks to the real-time feedback that Uber collects with social listening, they know which features are frequently requested by their customers. Simultaneously, Uber can also use this feedback to decide which features have been unsuccessful and require further modification.

Uber uses social media monitoring to discover how their current customer base feels about the change. They also continue to monitor customer satisfaction from using Uber on a regular basis.

Social Listening ROI

The global rollout of the new rider app started in early November 2016 and will take several weeks, so Uber will only be able to sum up this particular project in terms of sales growth after it's been completed. Thanks to social listening, they can estimate social ROI, which, as stated previously, equals:

- 100% Growth in discussion volume
- 90% Growth in number of likes
- 250% Increase in number of shares

Conclusion

These three use cases illustrate how Uber can leverage BI to address its most pressing challenges:

1. **The Operational Dashboard** improves real-time decision-making for airport-city trips by addressing inefficiencies and boosting revenue potential.
2. **The Strategic Dashboard** provides a high-level view of operational performance, enabling senior management to align strategic goals with market dynamics.
3. **Social Media Analytics** bridges the gap between internal operations and customer perceptions, offering actionable insights to enhance the customer experience.

By combining visual analytics with advanced data analytics, Uber can optimize its operations, improve customer satisfaction, and maintain its competitive edge in the ride-hailing market.

IMPLEMENTATION AND RECOMMENDATIONS

The implementation of the proposed BI solution requires careful consideration of managerial, technical, and ethical aspects to ensure success. This section outlines the step-by-step approach for Uber to integrate BI into its operations while addressing challenges such as data preparation, dashboard development, and ethical data usage. The focus is on delivering actionable insights that align with Uber's strategic goals, operational efficiency, and customer expectations.

---

1. Managerial Aspects

Managerial oversight is critical to the success of the BI implementation. Uber must ensure clear communication, stakeholder involvement, and effective training programs during each phase of implementation.

- **Stakeholder Engagement**: Operations, finance, and marketing teams must collaborate to define KPIs and ensure the dashboards address real business needs.
- **User Training**: Training sessions will equip stakeholders with the knowledge to use BI tools, interpret insights, and make data-driven decisions.
- **Feedback Mechanism**: Post-implementation, feedback from dashboard users should guide iterative improvements to the BI system.

---

2. Technical Aspects

The technical implementation focuses on integrating Uber's existing infrastructure with the proposed BI solution to enable real-time analytics and predictive insights.

Data Preparation and Transformation:

- **Data Sources**: Internal data from ride transactions, driver records, and customer feedback; external data from traffic and weather APIs.
- **ETL Processes**: Automated pipelines will extract, transform, and load (ETL) data into a structured data warehouse designed using the proposed star schema.
- **Data Quality Assurance**:
    - Validation rules to check data completeness, accuracy, and consistency.
    - Regular audits to identify and rectify anomalies in real-time and historical data.

Dashboard Development:

- **Tools**: Tableau for dashboards, Apache Pinot for real-time analytics, and Python for advanced analytics such as sentiment analysis.
- **Key Functionalities**:
    - **Interactive Filters**: Allow users to segment data by time, location, or ride type.
    - **Drill-Down Options**: Provide detailed insights into driver or customer performance metrics.
    - **Alerts and Notifications**: Enable proactive responses to operational inefficiencies.

Analytics and Insights:

- **Predictive Models**: Forecast demand patterns during peak times and suggest optimal driver allocation.
- **Sentiment Analysis**: Utilize Python-based NLP techniques to analyze social media sentiment and customer reviews.

---

3. Ethical Aspects

Ethics play a crucial role in the responsible implementation of BI systems. Uber must address the following ethical considerations:

- **Data Privacy**:
    - Customer and driver data must be anonymized before analysis to protect individual identities.
    - Implement robust data encryption protocols to secure sensitive information, such as customer locations and payment details.
- **Transparency**:
    - Clearly communicate how data is collected, processed, and used to internal stakeholders and customers.
    - Provide options for customers to opt out of non-essential data processing (e.g., sentiment analysis).
- **Bias Mitigation**:
    - Ensure algorithms used in predictive analytics and driver allocation are free from biases that could disadvantage specific demographics or geographic regions.
    - Regular audits of predictive models and dashboards to identify and address potential biases.
- **Compliance**:
    - Adhere to regulations such as GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act) for data handling and usage.
    - Engage with legal teams to ensure full compliance with regional data protection laws.

---

4. Implementation Phases

Phase 1: Planning and Requirement Gathering

- **Activities**: Conduct workshops, define KPIs, and finalize the scope of dashboards and analytics.
- **Output**: A detailed BI implementation roadmap.

Phase 2: Data Integration and Infrastructure Setup

- **Activities**: Set up data pipelines, integrate APIs, and configure data warehouses with the star schema.
- **Output**: A functional data infrastructure supporting real-time and historical analytics.

Phase 3: Dashboard Development

- **Activities**: Develop interactive dashboards with operational and strategic KPIs.
- **Output**: Two Tableau dashboards (Airport-City Trips and Operational Performance).

Phase 4: Deployment and User Training

- **Activities**: Deploy dashboards, conduct training, and establish a feedback loop.
- **Output**: Fully implemented BI system and trained users.

---

5. Completed Deliverables

- **Airport-City Trips Dashboard**:
  - Interactive filters, cancellation heatmaps, and trip performance metrics.
- **Operational Performance Dashboard**:
  - Revenue trends, driver productivity insights, and peak-hour contribution analysis.
- **Social Media Sentiment Analysis**:
  - Insights into customer sentiment trends and recurring service-related themes.

---

6. Evaluation Metrics

- **Operational Metrics**:
  - Reduction in cancellation rates.
  - Increase in driver availability during peak hours.
- **Ethical Metrics**:
  - Full compliance with data protection regulations.

    - Reduction in bias within predictive models.
- **User Adoption Metrics**:
    - Percentage of stakeholders using dashboards.
    - Positive feedback from training sessions.

Kotter's Model

Step	Action	Outcome
1. Establish a Sense of Urgency	We must address critical issues such as high cancellation rates, driver unavailability, revenue leakage, and declining customer sentiment. These challenges directly impact our revenue and reputation.	Creates momentum across teams by making it clear why implementing a BI solution is critical right now.
2. Form a Powerful Guiding Coalition	We will build a cross-functional team with key stakeholders from operations, customer service, finance, marketing, and BI specialists.	Ensures everyone is aligned, engaged, and committed to the success of the BI solution.
3. Create a Vision for Change	Together, we'll define a clear vision for how BI will optimize operations, enhance customer satisfaction, and drive revenue growth through data-driven insights.	Provides us with a shared goal and ensures everyone understands how BI supports our success.
4. Communicate the Vision	We'll use presentations, workshops, and success stories to show how BI dashboards and analytics can solve our operational challenges.	Gains alignment and excitement from all levels of the organization about our BI transformation.
5. Empower Others to Act on the Vision	By providing hands-on training and resources, we'll ensure that our managers and teams can effectively use dashboards and interpret analytics results.	Empowers our teams to make faster, smarter decisions with confidence.
6. Create Short-Term Wins	Our first win will come from the Airport-City Trips Dashboard, which will immediately improve driver allocation and reduce cancellations.	Builds trust and confidence in the BI solution by demonstrating early successes.
7. Consolidate Gains and Produce More Change	We'll expand the solution to include the Operational Performance Dashboard and Social Media Analytics to address additional needs.	Reinforces the BI solution as a valuable tool for both operational and strategic decision-making.
8. Anchor the Changes in	We'll embed BI-driven insights into our daily operations, performance reviews, and strategic planning processes.	Makes BI a natural and integral part of how we operate and make decisions.

Step	Action	Outcome
Organizational Culture		

Recommendations

The proposed BI implementation plan for Uber addresses managerial, technical, and ethical aspects comprehensively. By integrating dashboards and analytics tools while adhering to ethical guidelines, the solution ensures data-driven decision-making that aligns with operational efficiency, customer satisfaction, and regulatory compliance. This approach positions Uber to achieve long-term success and maintain its competitive edge in the ride-hailing industry.

- Begin with dashboards solving **immediate issues** (cancellations, driver availability).
- Scale to predictive analytics and **sentiment analysis**.
- Establish a **governance framework** for ethical and effective data usage.

Summary and Conclusion

In this project, we demonstrated how Uber can leverage Business Intelligence (BI) to address critical operational and strategic challenges. The analysis focused on optimizing airport-city trips, improving operational performance, and gaining customer insights through sentiment analysis. Three use cases were developed to showcase the potential of BI: an **Airport-City Trips Dashboard**, an **Operational Performance Dashboard**, and **Social Media Analytics**.

The **Airport-City Trips Dashboard** provided actionable insights to reduce cancellation rates and improve driver availability during peak hours, ensuring better resource allocation and increased revenue. The **Operational Performance Dashboard** offered a high-level view of revenue trends, trip efficiency, and driver productivity, enabling Uber's leadership to make informed strategic decisions. Lastly, **Social Media Analytics** highlighted customer sentiment trends, offering a deeper understanding of public perceptions and guiding customer-centric improvements.

To implement these solutions, a phased approach was outlined, including planning, data integration, dashboard development, and training. By leveraging existing tools like Tableau and Apache Pinot, combined with predictive analytics and NLP techniques, Uber can establish a robust BI ecosystem. The proposed solution is scalable, user-friendly, and designed to align with Uber's goals of operational efficiency, customer satisfaction, and revenue optimization.

We recommend that Uber adopt the proposed BI solution to address its immediate challenges and achieve long-term competitiveness. Key recommendations include:

1. **Implement Real-Time Dashboards**: Ensure that operations and strategy teams have access to real-time insights for quick decision-making.
2. **Leverage Predictive Analytics**: Use demand forecasting and driver allocation models to optimize supply-demand alignment.
3. **Monitor Customer Sentiment**: Regularly analyze social media feedback to improve service reliability and reputation.
4. **Invest in User Training**: Conduct extensive training programs to ensure stakeholders can effectively utilize BI tools.

By embracing the proposed BI solution, Uber can transform its operations, enhance customer satisfaction, and secure its leadership position in the dynamic ride-hailing industry.

References of Datasets Used

a) Prototype 1 dataset:

https://wpi0my.sharepoint.com/:x:/g/personal/vmadepogu_wpi_edu/ESjhZAnLgStEnNmo0rqBhBEBluFdXnoUSBPrDY87iKjDUg?e=3G6vHa

b) Prototype 2 dataset:

uber_dataset 1.xlsx

https://github.com/Priyachakraborty/Uber_data_analysis/blob/master/uber_dataset.csv

c) Social Media Analytics dataset:

https://wpi0my.sharepoint.com/:x:/g/personal/vmadepogu_wpi_edu/EaLeXtxbwJFDhy4PQ1dDe-kB59I7xvo7zUzeHJLKYSm3HQ?e=BZIPz6