IFT 533: Data Visualization & Reporting for IT

Project - Phase III: Dashboard Implementation

Course Project Group - 18

Team Members:

Kavya Bijja (1231777770)

Sreeja Vaddi (1232211840)

Saketh Vardhan Kalashikam (1232198359)

Eswar Reddy Kothapalli (1231807956)

Instructor: Asmaa Elbadrawy

Due Date: 12/01/2024

Table of Contents

[“Section 1: The Dashboard” 3](#_Toc183987571)

[“Section 2: The Dataset” 6](#_Toc183987572)

[“Section 3: Dashboard Users” 11](#_Toc183987573)

[“Section 4: Questions” 14](#_Toc183987574)

[“Section 5: Plots” 15](#_Toc183987575)

[“Section 6: Interactivity” 25](#_Toc183987576)

[“Extra Credit: Tableau Story Implementation” 31](#_Toc183987581)

[Extra Credit: Tableau Story Implementation Link: 32](#_Toc183987582)

[Dashboard Link of Mural: 32](#_Toc183987583)

[Final Dashboard Submission Link: 33](#_Toc183987584)

# **“Section 1: The Dashboard”**

**A screenshot of a computer

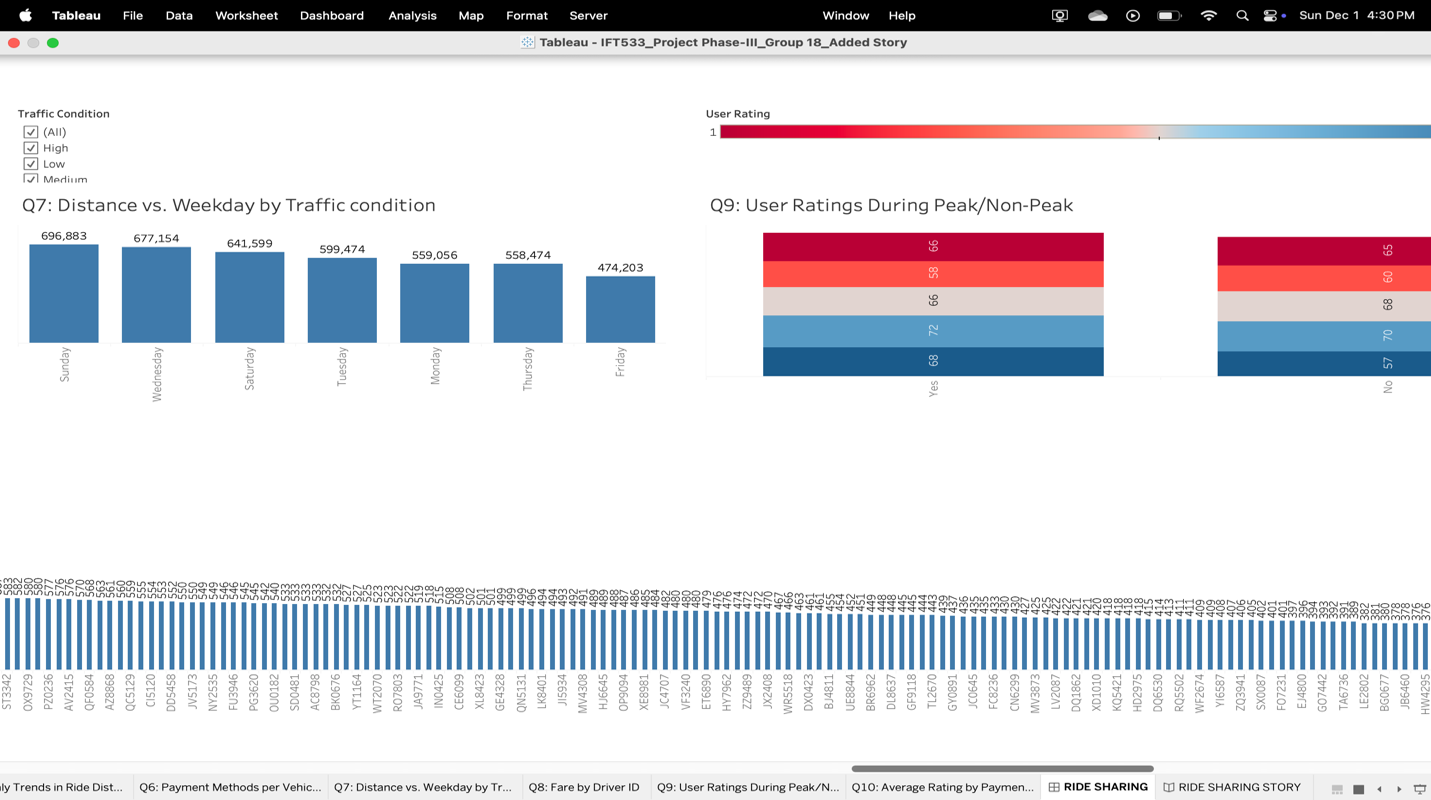
Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

****

**A screenshot of a computer

Description automatically generated**

**A screenshot of a graph

Description automatically generated**

Key trends, performance indicators, and user preferences are revealed by the dashboard’s analysis of ride-sharing data. For a variety of stakeholders, it offers actionable insights: marketing teams can adjust their plans based on user behavior, operations managers may optimize operator and vehicle distribution, and executives can keep an eye on business success. Financial analysts evaluate income and profitability, urban planners pinpoint high-demand regions for improved city planning, and customer service teams use ratings to increase customer happiness. Throughout the company, this all-inclusive tool facilitates data-driven, well-informed decision-making.

# **“Section 2: The Dataset”**

**Dataset Selected:** Ride Sharing Dataset

**A screenshot of a computer

Description automatically generated**

**Link of the Dataset:** <https://www.kaggle.com/datasets/zahraatouq/ride-sharing-dataset>

**Description OF Dataset:** The 650 records in the ride-sharing dataset, which has 18 columns, include ride details, user reviews, location coordinates, and fare information. It covers topics like vehicle type, traffic circumstances, and payment options using a combination of numerical and category data. The data covers a range of days and contains details on public holidays and peak hours.

**Summary:-**

**Rows:** 650

**Columns:** 18

**Attributes:** Ride ID, Request Time, Pickup Location, Latitude Pickup, Longitude Pickup, Dropoff Location, Latitude Dropoff, Longitude Dropoff, Ride Distance (in miles)

Fare Amount (in $), Payment Method, Driver ID, Vehicle Type, Traffic Condition, Peak Hours, Day of Week, Public Holiday, User Rating.

**Understanding the Attributes of Ride Sharing Dataset:**

1. **Ride ID**

Integer

Domain: Integer values from 1 to 91.

Unique identifier for each ride.

1. **Request Time:**

Interval

Domain: Date and Time values ranging from Jan 1, 2022 to Nov 16, 2024.

Time when the ride was requested

1. **Pickup Location:**

Categorical

Domain: String representation of latitude and longitude coordinates.

Coordinates of the pickup location

1. **Latitude Pickup:**

Ratio

Domain: Float values ranging from -90 to 90.

Latitude of the pickup location

1. **Longitude Pickup:**

Ratio

Domain: Float values ranging from -180 to 180.

Longitude of the pickup location

1. **Dropoff Location:**

Categorical

Domain: String representation of latitude and longitude coordinates.

Coordinates of the drop-off location

1. **Latitude Dropoff:**

Ratio

Domain: Float values ranging from -90 to 90.

Latitude of the pickup location

1. **Longitude** **Dropoff:**

Ratio

Domain: Float values ranging from -180 to 180.

Longitude of the drop-off location

1. **Ride Distance:**

Ratio

Domain: Positive float values, ranging from approximately 583 to 12,009 miles.

Distance of the ride in miles

1. **Fare Amount:**

Ratio

Domain: Positive float values, ranging from approximately $78 to $1220.

Cost of the ride in dollars

1. **Payment Method:**

Categorical

Domain: “Visa”, “Debit Card”, “Cash”, “PayPal”, “Apple Pay”.

Method used for payment.

1. **Driver ID:**

Categorical

Domain: Alphanumeric string values in the format of two letters followed by four digits.

Unique identifier for the driver

1. **Vehicle Type:**

Categorical

Domain: “SUV”, “Sedan”, “Bus”, “Motorcycle”

Type of vehicle used for the ride

1. **Traffic Condition:**

Ordinal

Domain: “Low”, “Medium”, “High”.

Traffic condition during the ride

1. **Peak Hours:**

Categorical

Domain: “Yes”, “No”.

Whether the ride occurred during peak hours

1. **Day of Week:**

Categorical

Domain: “Mon”, “Tue”, “Wed”, “Thu”, “Fri”, “Sat”, “Sun”.

Day of the week when the ride occurred

1. **Public Holiday:**

Categorical (Binary)

Domain: “Yes”, “No”.

Boolean values represented as “Yes” or “No”.

1. **User Rating:**

Ordinal

Domain: Integer values from 1 to 5.

Integer values from 1 to 5.

**“Data Pre-processing”**

The preparation procedures listed below were carried out to guarantee the dataset was clean and prepared for visualization:

**Request Time:** After being initially saved as a string with the format “MM/DD/YY HH:MM,” the Request Time column has been transformed into a datetime data type. As a result, temporal components like Year, Month, Day, and Hour may be extracted. Effective temporal filtering and aggregation are made possible by these modifications, enabling studies such as riding trends across time.

**Latitude and Longitude Columns:** To make mapping visualizations easier, the geographic columns-Latitude Pickup, Longitude Pickup, Latitude Dropoff, and Longitude Dropoff-have been given the proper geographic functions. To guarantee accuracy in spatial analysis, values in these columns were verified to make sure they fall within the permitted ranges (-90 to 90 for latitude and -180 to 180 for longitude).

**Missing Data:** We looked for missing data in several important categories, such as User Rating, Fare Amount (in dollars), and Ride Distance (in miles). The median was used to impute missing numerical values, while the mode was used to fill in categorical features such as Payment Method. To ensure data integrity and dependability, rows with significant missing data in latitude or longitude characteristics were eliminated.

This preprocessing ensures the dataset is complete, accurate, and ready for advanced visualization and analysis.

# **“Section 3: Dashboard Users”**

1. **Executives:**

**Role:** Monitor performance metrics and trends.

**Helpful For Dashboard:**

* Utilizing thorough ride-sharing data to inform strategic decision-making.
* Recognizing the patterns and trends that affect the course of the business.
* Evaluation of performance indicators to direct corporate strategy.
* For improved positioning, comprehend the competitive environment and market dynamics.
* Examining use and revenue data to make resource allocation easier.

1. **Operations Managers:**

**Role:** Optimize driver allocation and vehicle usage

**Helpful For Dashboard:**

* Keeping an eye on data in real time for effective resource management and operations.
* Using data analysis to locate and fix operational bottlenecks.
* Adjusting driver assignments as needed to account for traffic and demand.
* Monitoring important performance metrics to gauge operational effectiveness.
* Putting operational plans into action based on insights from data.

1. **Marketing Teams:**

**Role:** Analyze user preferences for targeted promotions

**Helpful For Dashboard:**

* Acquiring knowledge about consumer behavior and preferences in order to customize marketing tactics.
* Using ride patterns and user involvement to evaluate the success of marketing efforts.
* Determining prospective clientele groups for focused marketing campaigns.
* Monitoring customer reviews and ratings to enhance brand recognition.
* Examining fare patterns to modify price tactics in advertising campaigns.

1. **Customer Service:**

**Role:** Track user ratings and improve service quality

**Helpful For Dashboard:**

* Keeping an eye on customer reviews and ratings to enhance service and maintain quality.
* Use data analysis to find frequent service problems in order to improve user happiness.
* Delivering data-driven answers to complaints and questions from clients.
* Monitoring service patterns to have a deeper understanding of user expectations.
* Examining patterns and periods of high usage to improve user support.

1. **Financial Analysts**

**Role:** Evaluate revenue trends and profitability

**Helpful For Dashboard:**

* Evaluating financial performance by carefully examining ticket prices and expenditures.
* Using ride data and user behavior to forecast income patterns.
* Financial planning and budgeting informed by insights from operational data.
* ROI analysis is used to track how cost-effective marketing efforts are.
* Evaluating data to find financial possibilities and threats.

1. **Urban Planners**

**Role:** Understand transportation needs and patterns

**Helpful For Dashboard:**

* Leveraging ride-sharing data to guide the construction of urban transportation infrastructure.
* Examining ride trends to pinpoint high-demand locations for improved service positioning.
* Incorporating ride-sharing knowledge into public transportation and city planning plans.
* Evaluating the effects of ride-sharing trends on the environment in cities.
* Enhancing city transportation solutions in partnership with ride-sharing providers.

1. **Data Scientists**

**Role:** Explore correlations for predictive modeling

**Helpful For Dashboard:**

* Utilizing predictive analytics to project future trends and demand for rides.
* Constructing computer learning models to improve ride-sharing pricing and algorithms.
* Finding user demographics and behaviors through cluster analysis.
* Producing insights from huge datasets to aid in improvements and strategic choices.
* Enhancing user experience via customization and data-driven suggestions.

# **“Section 4: Questions”**

1. How does total ride distance vary across traffic conditions for selected days of the week?
2. What is the distribution of traffic conditions (Low, Medium, High) across different ride requests?
3. How do average fare amounts vary across traffic conditions?
4. How does ride distance vary across different vehicle types?
5. What are the monthly trends in total ride distances?
6. What payment methods are most popular across different vehicle types?
7. How does the total ride distance for each day of the week change based on selected traffic conditions?
8. How do fare amounts vary for different driver IDs?
9. What is the distribution of user ratings during peak hours compared to non-peak hours?
10. What is the average user rating for each payment method?

# **“Section 5: Plots”**

**Q1: Total Ride Distances (Public Holidays vs. Regular Days)**

1. **Proposed Plot:** Bar chart comparing ride distances for selected weekdays.
2. **Pre-attentive Attributes:** Use Day of week to differentiate holidays from regular days.
3. **Question Addressed:** How does total ride distance vary across traffic conditions for selected days of the week?
4. **How it Addresses:** With a day-of-week filter, the chart uses color bars to compare ride distances on holidays and ordinary days, making it easy to see how traffic conditions differ on particular weekdays.

**A screenshot of a computer

Description automatically generated**

**Q2: Traffic Conditions Distribution**

1. **Proposed Plot:** Bar chart for traffic conditions.
2. **Pre-attentive Attributes:** Use bar color to represent traffic conditions.
3. **Question Addressed:** What is the distribution of traffic conditions (Low, Medium, High) across different ride requests?
4. **How it Addresses:** The chart effectively analyzes traffic dynamics across trip requests by displaying the frequency of each traffic scenario using bar heights and color.

**A screenshot of a computer

Description automatically generated**

**Q3: Fare Amount vs. Traffic Condition**

1. **Proposed Plot:** Bar chart for fare amounts.
2. **Pre-attentive Attributes:** Use bar colors to indicate different traffic levels.
3. **Question Addressed:** How do average fare amounts vary across traffic conditions?
4. **How it Addresses:** For clear and useful information, the graphic illustrates how traffic circumstances affect fare pricing by highlighting changes in price amounts with different bar colors and heights.

A screenshot of a computer

Description automatically generated

**Q4: Ride Distance by Vehicle Type**

1. **Proposed Plot:** Bar chart for ride distances by vehicle type.
2. **Pre-attentive Attributes:** Bar color to represent each vehicle type.
3. **Question Addressed:** How does ride distance vary across different vehicle types?
4. **How it Addresses:** The graphic compares ride distances by vehicle type using color and bar heights, which makes it simple to spot trends and preferences in vehicle utilization.

A screenshot of a computer

Description automatically generated

**Q5: Monthly Trends in Ride Distances**

1. **Proposed Plot:** Line chart for monthly distances.
2. **Pre-attentive Attributes:** Different line styles and markers for visual clarity.
3. **Question Addressed:** What are the monthly trends in total ride distances?
4. **How it Addresses:** In a clear, sequential fashion, the line chart illustrates variations in the seasons, abnormalities, or growth patterns by displaying patterns and changes in ride distances over the course of several months.

A screenshot of a computer

Description automatically generated

**Q6: Payment Methods per Vehicle Type**

1. **Proposed Plot:** Stacked bar chart for payment methods by vehicle type.
2. **Pre-attentive Attributes:** Unique colors for each payment method in the stacks.
3. **Question Addressed:** What payment methods are most popular across different vehicle types?
4. **How it Addresses:** With stack heights and colors that make it simple to spot trends and preferences briefly, the chart graphically analyzes the prevalence of payment methods among car kinds.

A screenshot of a computer

Description automatically generated

**Q7: Distance vs. Weekday by Traffic condition**

1. **Proposed Plot:** Bar chart for distances and days of week for selected traffic level.
2. **Pre-attentive Attributes:** Use bar height to differentiate each week
3. **Question Addressed:** How does the total ride distance for each day of the week change based on selected traffic conditions?
4. **How it Addresses:** Weekday traffic trends can be easily identified because to the chart’s usage of bar heights to display differences in ride distances for each weekday under various traffic situations.

A screenshot of a computer

Description automatically generated

**Q8: Fare by Driver ID**

1. **Proposed Plot:** Bar chart for fares by driver ID.
2. **Pre-attentive Attributes:** Bar height to represent fare amounts.
3. **Question Addressed:** How do fare amounts vary for different driver IDs?
4. **How it Addresses:** By graphically comparing fee amounts across drivers and emphasizing any discrepancies or patterns, the chart makes it simpler to spot good performers or regions that require improvement.

A screenshot of a computer

Description automatically generated

**Q9: User Ratings During Peak/Non-Peak**

1. **Proposed Plot:** Stacked bar chart for user ratings.
2. **Pre-attentive Attributes:** Color-coded stacks for peak and non-peak times.
3. **Question Addressed:** What is the distribution of user ratings during peak hours compared to non-peak hours?
4. **How it Addresses:** The graphic illustrates variations in customer satisfaction by time of day by contrasting user rating distributions for peak and non-peak hours.

A screenshot of a computer

Description automatically generated

**Q10: Average Rating by Payment Method**

1. **Proposed Plot:** Bar chart for average ratings by payment method.
2. **Pre-attentive Attributes:** Bar length to indicate average ratings.
3. **Question Addressed:** What is the average user rating for each payment method?
4. **How it Addresses:** To find patterns in customer satisfaction based on payment preferences, the graphic contrasts the average ratings for each payment method.

A screenshot of a graph

Description automatically generated

# **“Section 6: Interactivity”**

The dashboard incorporates two key interactive controls that enhance user exploration and insights:

### **Control 1: Day of the Week Filter**

**Purpose:** Allows users to filter visualizations based on specific days of the week. This enables focused analysis of ride distances and other metrics for selected weekdays.

**Plots Connected:**

**Q1: Total Ride Distances (Public Holidays vs. Regular Days):**

Users can filter the bar chart to display ride distances on specific weekdays, comparing public holidays versus regular days.

**Value Range: Options:** Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday.

Multi-select functionality to analyze trends across multiple days simultaneously.

**Attribute Used:** Day of Week.

**Implementation Steps**

**In Tableau**:

**Day of the Week Filter**: Drag Day of Week to the Filters shelf, Enable Multiple Values (List), Display the filter on the dashboard as a dropdown or checkbox list for interactivity.

**“Before Applying Filter”**

**A screenshot of a computer

Description automatically generated**

**“After Applying Filter”**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

### **Control 2: Traffic Condition Filter**

**Purpose:** Enables users to filter data based on traffic conditions (Low, Medium, High). This helps to uncover patterns in ride distances under specific traffic levels.

**Plots Connected:**

**Q7: Distance vs. Weekday by Traffic Condition:** Updates the visualization to show total ride distances for selected weekdays, filtered by traffic conditions.

**Value Range:** **Options:** Low, Medium, High. Multi-select functionality to compare trends across multiple traffic levels.

**Attribute Used:** Traffic Condition.

**Implementation Steps**

**In Tableau**:

**Traffic Condition Filter**: Drag Traffic Condition to the Filters shelf, Enable Multiple Values (List), Display this filter as a dropdown or checkbox list.

Use Tableau's Dashboard Actions to link both filters to the connected plots.

**“Before Applying Filter”**

A screenshot of a computer

Description automatically generated

**“After Applying Filter”**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

### **Summary of Interactive Controls**

| **Control** | **Plots Connected** | **Value Range** | **Attribute Used** |
| --- | --- | --- | --- |
| Day of the Week Filter | Q1 | Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday | Day of Week |
| Traffic Condition Filter | Q7 | Low, Medium, High | Traffic Condition |

### **Benefits of Interactivity**

These interactive controls enhance the dashboard by allowing users to:

* Explore ride patterns for specific weekdays and traffic conditions.
* Gain granular insights into how ride distances differ based on holidays, regular days, and traffic levels.
* Compare multiple scenarios (e.g., weekday vs. weekend, low traffic vs. high traffic) dynamically.

These features ensure the dashboard remains flexible and user-centric, enabling stakeholders to derive actionable insights.

# **“Extra Credit: Tableau Story Implementation”**

**Tool/Feature:** Tableau Story Creation

**Implementation Plan:**

We’ll combine several dashboards and sheets into a narrative using Tableau’s “Story” function to effectively convey findings. The narrative will display a series of visuals that emphasize significant discoveries, patterns, or suggestions.

**Procedures:**

**1.⁠ ⁠Determine the Main Takeaways:** To incorporate into the narrative, pick the most significant data points or visualizations from the current Tableau workbook.

**2.⁠ ⁠**Use the “New Story” option from the Tableau workspace to add a narrative to Tableau.

**3.⁠ ⁠Create Story Points:** Put together the chosen visualizations as “story points” in a way that makes sense. For context, include descriptions or captions for every point.

**4.⁠ ⁠Personalize Appearance:** Make use of formatting tools to make sure the tale is aesthetically pleasing and appropriate for the intended audience.

**5.⁠ ⁠Interactive Elements:** Use interactive highlights or filters to provide the consumer a more engaging experience.

This feature will improve stakeholders understanding by showcasing data insights in an organized and powerful manner once it is put into use.

A screenshot of a computer

Description automatically generated

**Extra Credit: Tableau Story Implementation Link:** <https://public.tableau.com/shared/TWK4MSY9Z?:display_count=n&:origin=viz_share_link>

# **Dashboard Link of Mural:**

<https://app.mural.co/t/kbijjaasuedu3761/m/kbijjaasuedu3761/1731445272945/bd7da758cb9ea79d4084aa6565c3f52e2c835f73>

**Final Dashboard Submission Link:** <https://public.tableau.com/views/IFT533_ProjectPhase-III_Group18_Dashboard/RIDESHARING?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link>