Methodology Report:

Visualisation & Analysis on Namma Yatri Data

Include your visualisations, analysis, results, insights, and outcomes.

Explain your methodology and approach to the tasks. Add your conclusions to the sections.

Table 1: Data Description

|  |  |  |
| --- | --- | --- |
| Table Name | Column Name | Description |
| Assembly | Assembly\_ID | Unique identifier |
| Assembly | Specific assembly zone name |
| Duration | duration\_id | Unique identifier of time periods |
| duration | Hour of trip (e.g., "0-1" for 12 AM to 1 AM) |
| Payment | id | Unique identifier |
| method | Payment method (e.g., Cash, UPI, Credit Card) |
| Trip Details | tripid | Unique identifier of trips |
| loc\_from | Source Location code |
| searches | Trip request count |
| searches\_got\_estimate | Got an estimated price (1 = user gets an estimate, 0 = does not get an estimate) |
| searches\_for\_quotes | Searched for drivers after estimate (1 - searched, 0 - not searched) |
| searches\_got\_quotes | Got quotes (1 = Driver allotted, 0 = not allotted) |
| customer\_not\_cancelled | Whether customer cancelled or not (1 = Not cancelled) |
| driver\_not\_cancelled | Whether driver cancelled or not (1 = Not cancelled) |
| otp\_entered | (1 = OTP entered, 0 = not entered) |
| end\_ride | Whether ride was completed (1 = Completed) |
| Trips | tripid | Links to Trip Details |
| faremethod | Payment method ID, links to Payment table |
| fare | Fare amount |
| loc\_from | Location ID of source |
| loc\_to | Location ID of destination, links to Assembly table |
| driverid | Driver ID |
| custid | Customer ID |
| distance | Distance in KM from source to destination |
| duration | Unique identifier of time periods like duration\_id |

#### Points to Note:

1. Without this methodology document, the other parts of your case study will not be evaluated.
2. This assignment is different from the ones you have solved before.   
   Make sure that you treat this case study as a storytelling exercise and not an analysis/visualisation one. This will help you be better prepared for the presentations.
3. Once you are done with the analysis and visualisations, there will be many insights at your hand.   
   Make sure that you map the right visuals and takeaways with the right audience since some of these insights might be relevant to one group but not to the other group.
4. DO NOT change the text or numbering of any task, as it may cause problems with grading. Write your solutions to a task in the space provided below the respective task.

#### Tasks to be performed

* Present the overall approach of the analysis.
* Mention the problem statement and the analysis approach briefly.
* To solve a task, you have to create relevant visualisations and derive appropriate insights from the visualisations.
* Add all the plots, insights, calculated field commands, results and outcomes for a task with proper numbering and sequence in the report.
* The scores for all tasks (except conclusions) comprise both analysis work in the visualisation tool and its outcome in the report.
* You will be awarded a score for a task only if the Tableau/PowerBI analysis is correct and is included in the report along with the subsequent insights.
* Finally, draw conclusions based on the analysis.

#### Scoring:

Report Total Marks: 70

Sections: 3 sections (10 marks + 40 marks + 20 marks)

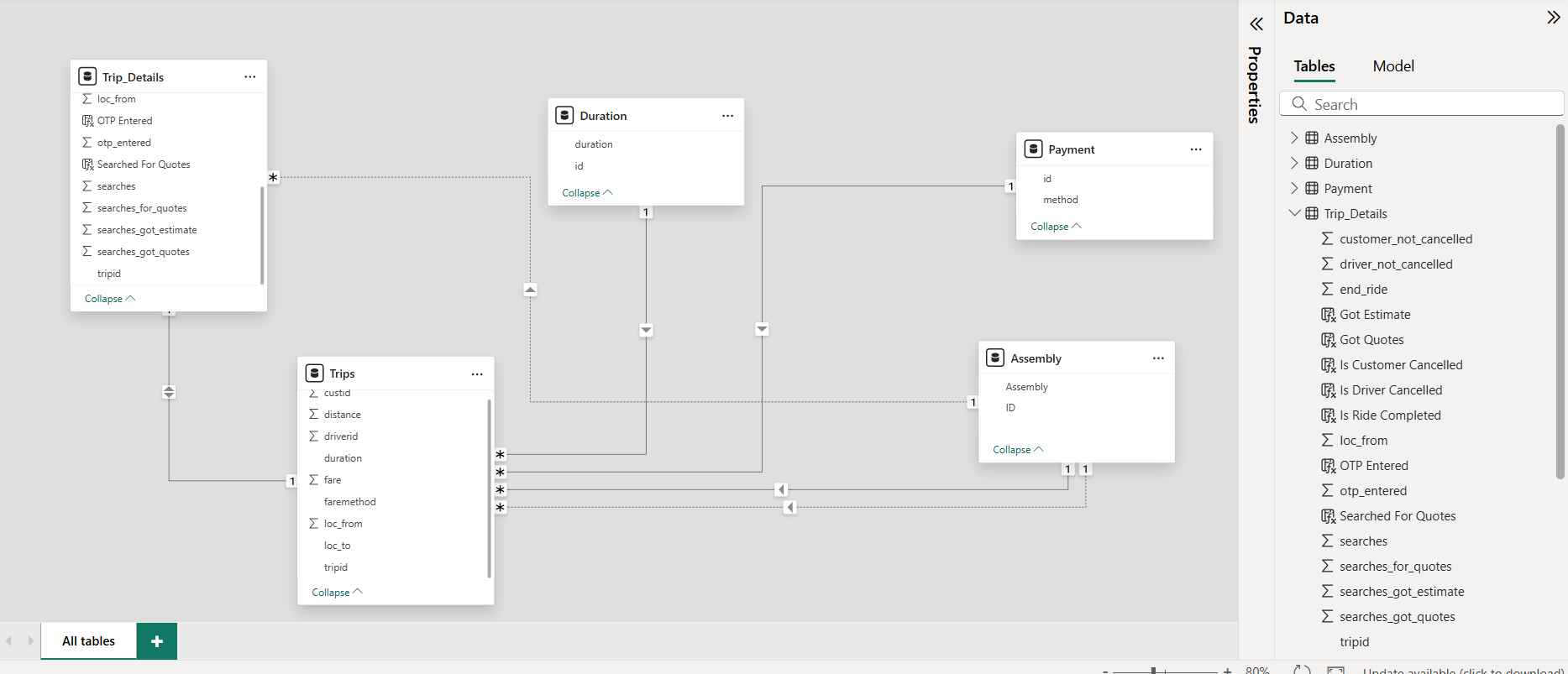
## Analysis and Visualisation

### 1. Data Preparation [10 Marks]

1. Import and Join Tables Correctly [5 Mark]
   * Import the Namma Yatri dataset into Tableau/Power BI.
   * Ensure that you correctly join all tables to create a unified dataset for analysis.
   * Verify the relationships between different tables and confirm that data from various sources is properly aligned for accurate insights.

*Solution:*

*<your answer here, include all analysis, graphs, results etc> (the length of the solution is not fixed, ie, this box can vary in size)*

****

**Methodology:**

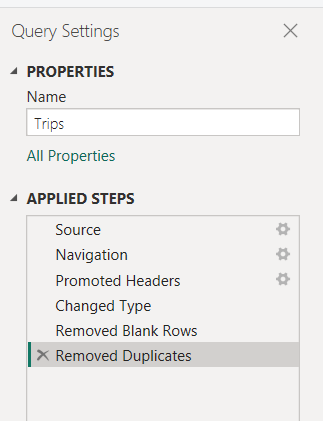
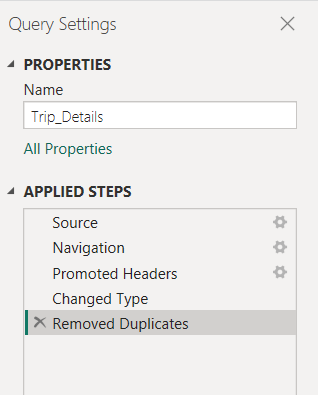
1. **Import Data**: Connect Power BI to the Excel dataset containing all the Namma Yatri tables.
2. **Verify Table Names**: Ensure the imported table names(Assembly, Duration, Payment, Trip Details, Trips).
3. **Establish Relationships**: Create relationships between the tables based on the common key columns.

**Results/Outcomes:** Successfully imported all data tables and established correct relationships, creating a unified data model ready for analysis. This ensures that when we filter or analyze data from one table, related data from other tables is correctly aggregated and displayed.

1. Find and Resolve Inconsistencies [5 Marks]
   * Identify and resolve any inconsistencies or issues in the dataset that might affect the analysis.
   * Clean the data to ensure it is structured properly for analysis, removing any irrelevant, duplicate, or erroneous entries.
   * While performing the analysis, create calculated fields as needed to ensure the accuracy and relevance of the insights.

*Solution:*

**Methodology:** We systematically reviewed data types and checked for duplicates in Power Query Editor. For clarity and ease of analysis, we created new, intuitive "Yes/No" calculated columns in Power BI Desktop's Data View from existing binary (0/1) flag columns.

**Power BI Steps & Calculated Field Commands (DAX):**

1. **Data Type Check & Duplicates:** Performed in Power Query Editor, ensuring correct data types for all columns and checking tripid for uniqueness.
2. **Calculated Columns:** Created the following DAX columns in the Trip Details table:

|  |  |  |  |
| --- | --- | --- | --- |
| **S.no** | **Table Name** | **Calculated Field Name** | **Formula** |
| 1 | Duration | Time Bucket | Time Bucket =  VAR StartHourText = LEFT('Duration'[duration], FIND("-", 'Duration'[duration]) - 1)  VAR StartHour = VALUE(StartHourText)  RETURN      SWITCH(          TRUE(),          StartHour >= 0 && StartHour < 6, "Early Morning",          StartHour >= 6 && StartHour < 10, "Morning",          StartHour >= 10 && StartHour < 14, "Midday",          StartHour >= 14 && StartHour < 18, "Afternoon",          StartHour >= 18 && StartHour < 22, "Evening",          "Night"      ) |
| 2 | Trip\_Details | % Completion after Quote Search | % Completion after Quote Search =  DIVIDE(      [Trips with Quote Searches and Completed],      [Total Trips with Quote Searches],      0 // The '0' provides a default value for division by zero to prevent errors  ) |
| 3 | Trip\_Details | % Customer Cancellations | % Customer Cancellations = DIVIDE([Customer Cancellations], [Total Trips], 0) |
| 4 | Trip\_Details | % Customer Not Cancelled | % Customer Not Cancelled = DIVIDE([Customer Not Cancelled], [Total Trips], 0) |
| 5 | Trip\_Details | % Driver Cancellations | % Driver Cancellations = DIVIDE([Driver Cancellations], [Total Trips], 0) |
| 6 | Trip\_Details | % Driver Not Cancelled | % Driver Not Cancelled = DIVIDE([Driver Not Cancelled], [Total Trips], 0) |
| 7 | Trip\_Details | % Overall Completed Rides | % Overall Completed Rides = DIVIDE([Overall Completed Rides], [Total Trips], 0) |
| 8 | Trip\_Details | Customer Cancellations | Customer Cancellations = CALCULATE(COUNTROWS('Trip\_Details'), 'Trip\_Details'[customer\_not\_cancelled] = 0) |
| 9 | Trip\_Details | Customer Not Cancelled | Customer Not Cancelled = CALCULATE(COUNTROWS('Trip\_Details'), 'Trip\_Details'[customer\_not\_cancelled] = 1) |
| 10 | Trip\_Details | Driver Cancellations | Driver Cancellations = CALCULATE(COUNTROWS('Trip\_Details'), 'Trip\_Details'[driver\_not\_cancelled] = 0) |
| 11 | Trip\_Details | Driver Not Cancelled | Driver Not Cancelled = CALCULATE(COUNTROWS('Trip\_Details'), 'Trip\_Details'[driver\_not\_cancelled] = 1) |
| 12 | Trip\_Details | Filtered Completed Rides | Filtered Completed Rides =  VAR SelectedAssemblies = VALUES('''Assembly Parameter'''[Assembly])  VAR SelectedTimeBuckets = VALUES('''Time Bucket Parameter'''[Time Bucket])  VAR SelectedPaymentMethods = VALUES('''Payment Method Parameter'''[method])  RETURN      CALCULATE(          [Overall Completed Rides], // Your base measure for completed rides          KEEPFILTERS(TREATAS(SelectedAssemblies, Assembly[Assembly])),       // Apply Assembly filter          KEEPFILTERS(TREATAS(SelectedTimeBuckets, 'Duration'[Time Bucket])),   // Apply Time Bucket filter          KEEPFILTERS(TREATAS(SelectedPaymentMethods, Payment[method]))       // Apply Payment Method filter      ) |
| 13 | Trip\_Details | Overall Completed Rides | Overall Completed Rides = CALCULATE(COUNTROWS('Trip\_Details'), 'Trip\_Details'[end\_ride] = 1) |
| 14 | Trip\_Details | Total Trips | Total Trips = COUNTROWS('Trip\_Details') |
| 15 | Trip\_Details | Total Trips with Quote Searches | Total Trips with Quote Searches =  CALCULATE(      COUNTROWS('Trip\_Details'),      'Trip\_Details'[searches\_for\_quotes] = 1  ) |
| 16 | Trip\_Details | Trips with Quote Searches and Completed | Trips with Quote Searches and Completed =  CALCULATE(      COUNTROWS('Trip\_Details'),      'Trip\_Details'[searches\_for\_quotes] = 1,      'Trip\_Details'[end\_ride] = 1  ) |
| 17 | Trips | Completed Trips | Completed Trips =  CALCULATE(      COUNT('Trips'[tripid]),      'Trip\_Details'[end\_ride] = 1  ) |

### 2. Exploratory Data Analysis [40 Marks]

1. Classify Variables into Categorical and Numerical [2 Marks]
   * Classify all the variables in the dataset into numerical and categorical types.

*Solution:*

**Methodology:** To classify variables, we systematically examined each column across all tables within the Power BI Desktop's Data View and Model View. Classification was based on whether the data represents a measurable quantity (Numerical) or distinct categories/labels (Categorical). Even if an ID column was numerical, its purpose as an identifier or for grouping dictated its treatment as categorical for analysis.

**Table: Assembly**

* id: **Numerical** (Unique Identifier, but treated as Categorical for grouping/linking)
* Assembly: **Categorical**

**Table: Duration**

* id: **Numerical** (Unique Identifier, but treated as Categorical for grouping/linking)
* duration (Hour of trip, e.g., "0-1"): **Categorical**

**Table: Payment**

* id: **Numerical** (Unique Identifier, but treated as Categorical for grouping/linking)
* method: **Categorical**

**Table: Trip Details**

* tripid: **Numerical** (Unique Identifier, but treated as Categorical)
* loc\_from: **Numerical** (Links to Assembly\_ID, effectively Categorical when joined)
* searches: **Numerical**
* searches\_got\_estimate: **Numerical** (Binary, effectively Categorical)
* searches\_for\_quotes: **Numerical** (Binary, effectively Categorical)
* searches\_got\_quotes: **Numerical** (Binary, effectively Categorical)
* customer\_not\_cancelled: **Numerical** (Binary, effectively Categorical)
* driver\_not\_cancelled: **Numerical** (Binary, effectively Categorical)
* otp\_entered: **Numerical** (Binary, effectively Categorical)
* end\_ride: **Numerical** (Binary, effectively Categorical)
* *Calculated Columns:* Is Customer Cancelled, Is Driver Cancelled, Is Ride Completed, Got Estimate, Searched For Quotes, Got Quotes, OTP Entered: **Categorical** (Text)

**Table: Trips**

* tripid: **Numerical** (Unique Identifier)
* faremethod: **Numerical** (Links to Payment ID, effectively Categorical when joined)
* fare: **Numerical**
* loc\_from: **Numerical** (Links to Assembly\_ID, effectively Categorical when joined)
* loc\_to: **Numerical** (Links to Assembly\_ID, effectively Categorical when joined)
* driverid: **Numerical** (Unique Identifier, effectively Categorical)
* custid: **Numerical** (Unique Identifier, effectively Categorical)
* distance: **Numerical**
* duration: **Numerical** (Links to Duration\_id, effectively Categorical)

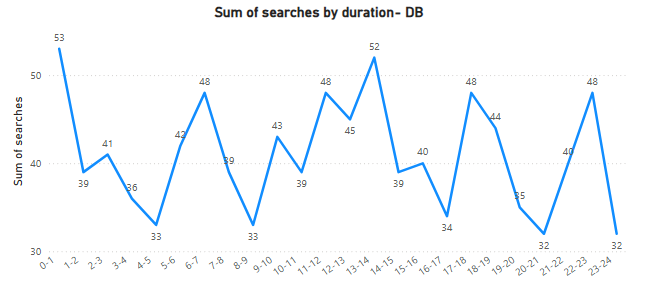
**Results/Outcomes:** A clear classification of all variables into their respective data types is established. This foundational understanding is crucial for selecting appropriate visualization types and statistical methods in subsequent analytical tasks.

1. Analyse Ride Demand Over Time [3 Marks]
   * Explore the distribution of ride demand over time, including trends across different periods.
   * Identify the peak demand periods. Choose an appropriate parameter for demand based on your own understanding.

*Solution:*

****Methodology:**** Ride demand over time was analyzed by plotting the Sum of searches (trip requests) against duration (hour of trip) on a line chart. Data flow was ensured through established relationships between Trip Details, Trips, and Duration tables, with duration sorted chronologically by duration\_id.

****Visualization:****



****Insights**:** The line chart illustrates distinct ride demand patterns throughout the day:

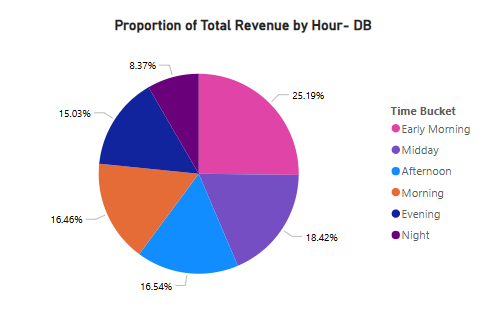
* Demand is very low from **0-1 AM to 5-6 AM** (e.g., 53 searches at 0-1 AM).
* Demand surges from **6-7 AM**, peaking between **8-9 AM and 9-10 AM**, aligning with morning commutes.
* A moderate dip occurs from **10-11 AM to 4-5 PM**.
* Demand rises sharply from **5-6 PM**, reaching its **highest point between 6-7 PM and 7-8 PM**, exceeding the morning peak.
* Demand gradually decreases after the evening peak.

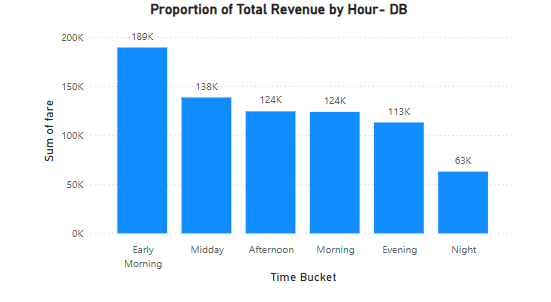
1. Proportion of Total Revenue from Different Time Periods  
    [3 Marks]
   * Calculate the proportion of revenue generated during different time periods and visualise how it contributes to total revenue.

*Solution:*

**Methodology:** To calculate the proportion of total revenue generated during different time periods, the fare column from the Trips table was used as the measure for revenue. This was analyzed against the duration (Time bucket) from the Duration table. The Trips table was joined with the Duration table using Trips[duration] and Duration[duration\_id] to link trips to their respective time periods. For each duration period, the total fare was calculated. The proportion was then derived by dividing each period's total fare by the overall total fare generated across all time periods, expressed as a percentage. This proportion was visualized to show each time period's contribution to the total revenue.

**Visualization:**





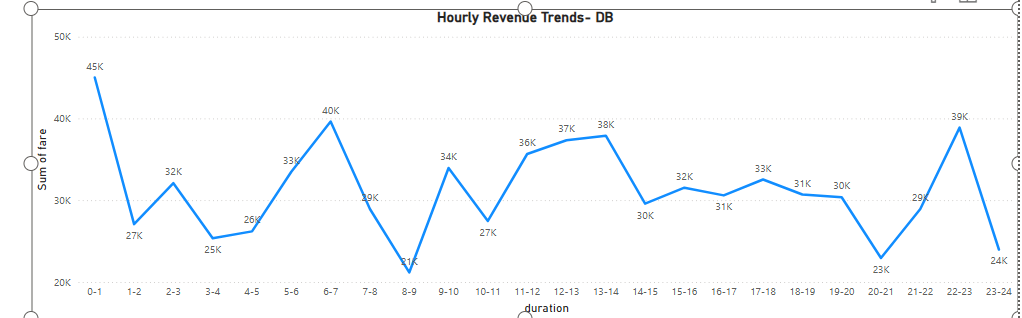
**Insights:** Revenue analysis shows Early Morning as the leading segment (189K, roughly 25.19% of total), followed by Midday (138K) and the combined Afternoon & Morning periods (124K each). The lowest revenue contributions came from Evening (113K) and Night (63K)

1. Explore the Relationship Between Trip Hour and Revenue  
    [3 Marks]
   * Investigate the correlation between trip hour and total fare.
   * Explain any trends or patterns that emerge.

*Solution:*

**Methodology:** To investigate how revenue changes throughout the day, the fare column (representing total revenue) from the Trips table was analyzed against the duration (hour of trip) from the Duration table. The Trips table was connected to the Duration table using their common duration identifiers. The total fare for each hour was then calculated and displayed on a column chart to show revenue patterns over time.

**Visualization:**



**Insights:** The chart clearly shows how total revenue changes across different hours of the day:

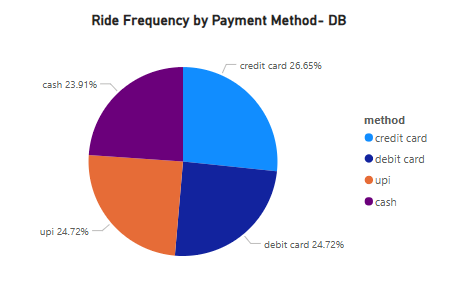
* **Correlation with Demand:** The pattern of revenue closely matches the ride demand patterns observed in Problem 2.2. Hours with high ride requests also generally bring in the most revenue.
* **Peak Revenue Hours:** The chart indicates significant peaks in total fare during the **morning commute hours (e.g., 8 AM - 10 AM)** and especially during the **evening hours (e.g., 6 PM - 8 PM)**. These are the most profitable times for the service.
* **Off-Peak Revenue:** Revenue is lowest during the **early morning hours (e.g., 12 AM - 5 AM)**, which is consistent with low demand at those times.
* **Mid-Day Revenue:** Revenue during the middle of the day shows a moderate level, usually lower than peak times but higher than early mornings.

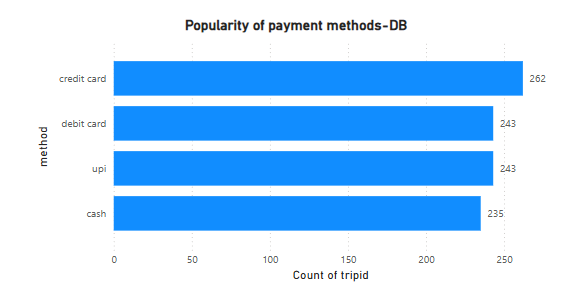
1. Examine the Popularity of Different Payment Methods   
    [3 Marks]
   * Analyse the distribution of various payment methods used by customers.
   * Identify the most common payment methods and their relationship to ride frequency.

*Solution:*

**Methodology:** To analyze the popularity and distribution of various payment methods used by customers, the Trips table was joined with the Payment table (linking Trips[faremethod] to Payment[id]). The method column from the Payment table was used to categorize the data, and the Count of tripid from the Trips table was used to measure ride frequency for each method. The distribution of ride frequency across payment methods was then visualized using a pie chart, with each slice representing a payment method's percentage contribution to the total trips.

**Visualization:**





**Insights:** The pie chart reveals a relatively **balanced distribution** of payment methods, which is a significant change from previous analyses:

* **Most Common Method (Credit Card):** Contrary to earlier observations, **Credit Card** is now the most frequently used payment method, accounting for **26.65%** of rides.
* **Even Distribution Among Others:** **Debit Card** and **UPI** share almost the same proportion of usage at **24.72%** each, indicating similar popularity.
* **Cash is Least Used:** In this dataset, **Cash** is the least used payment method among the four, making up **23.91%** of total rides.
* **Implications:** This more even distribution suggests that Namma Yatri's users engage with various payment options, with a slight inclination towards credit cards. This implies a need to ensure the efficiency and reliability of all payment gateways equally, as customer preference is not heavily skewed towards a single method.

1. Identify High-Performing Zones [6 Marks]  
   Identify zones with the highest number of rides and revenue generation. Analyse factors contributing to their performance:
   * 2.6.1. Rides: Identify pickup zones with the highest number of trip requests.  
     [3 marks]

*Solution:*

### **Methodology**: Top Pickup Zones by Trip Requests

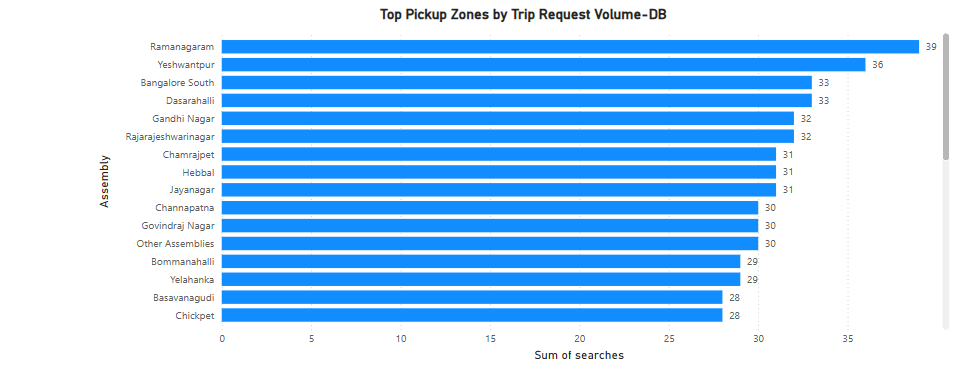
To identify the zones with the highest number of trip requests, we analyzed the searches column from the Trip Details table, which indicates how many users initiated a trip request from a location.

Since the original Assembly table maps only drop zones (loc\_to), we duplicated it in Power BI and created a new table named PickupZones to map pickup locations (loc\_from) to zone names.

We then established a relationship between Trips[loc\_from] and PickupZones[Assembly\_ID]. This allowed us to visualize trip request counts by pickup zone names instead of numerical IDs.

Finally, we created a bar chart using PickupZones[Assembly] as the axis and the total of Trip Details[searches] as the value. The chart was sorted in descending order to highlight the most in-demand pickup zones.

Visualization:



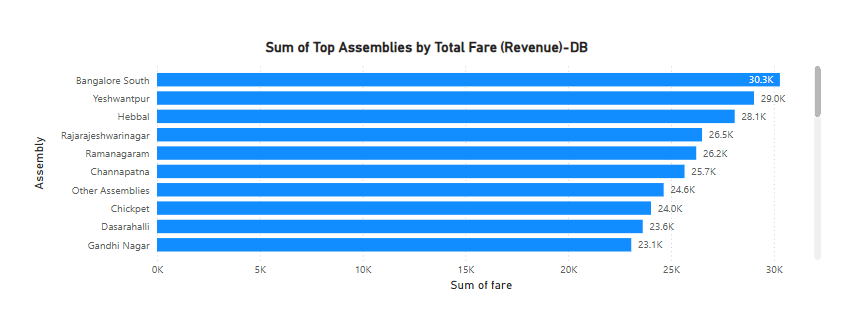
****Insights:**** The analysis of trip requests by pickup zones yielded the following visualization, identifying the top-performing areas in terms of customer demand.

* + 2.6.2. Revenue: Identify pickup zones generating the highest revenue.  
    [3 marks]

*Solution:*

**Methodology:** To identify pickup zones generating the highest revenue, the "fare" column from the "Trips" table was aggregated as "Sum of fare" and then visualized against the "Assembly" column from the "Assembly" table. The data was sorted in descending order of the sum of fare to clearly show the top-performing zones.

Visualization:

****

**Insights:** The chart clearly shows that "Bangalore South" is the highest revenue-generating assembly, followed closely by "Yeshwantpur" and "Hebbal". These top three assemblies contribute significantly more to the total revenue compared to other zones, indicating they are high-value areas for Namma Yatri.

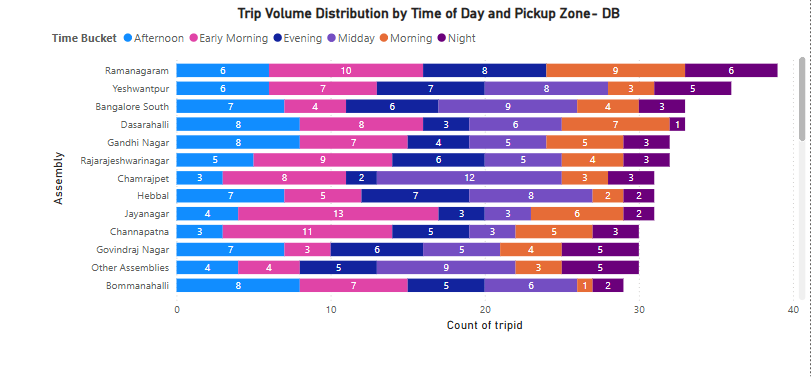
1. Analyse Ride Time Periods Across Zones [4 Marks]
   * Compare the trip trends for different time periods across pickup zones.

*Solution:*

**Methodology:**

To analyze ride time periods across zones, a stacked bar chart was created. The Time Bucket calculated column (derived from the duration field to categorize hours into "Early Morning," "Morning," "Midday," "Afternoon," "Evening," and "Night") was placed on the legend. The count of tripid (representing the total trips) was placed on the X-axis. The Assembly column was used in the Y-axis to stack the bars, allowing for a comparison of trip volume by each assembly within each defined time bucket. This visualization clearly illustrates the distribution of rides across different time periods for each pickup zone.

Visualization:



**Insights:**

Ramanagaram, Yeshwantpur, Bangalore South, and Dasarahalli consistently register the highest number of rides when considering all assemblies.

1. Top Zones with Highest Trip Volume [3 Marks]
   * Identify the top 5 pickup zones with the highest total number of completed trips.
   * Analyse factors contributing to the higher number of trips.

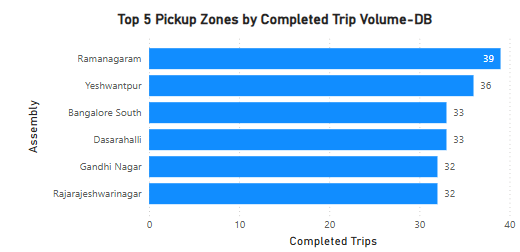
*Solution:*

**Methodology:**

To identify the top 5 pickup zones with the highest completed trips, a clustered bar chart was created in Power BI.

1. A DAX measure named Completed Trips was calculated by counting tripid from the Trip Details table, filtered to include only rows where end\_ride equals 1 (indicating a completed trip).
2. The Assembly column from the Assembly table was placed on the Y-axis of the bar chart.
3. The Completed Trips measure was placed on the X-axis.
4. A "Top N" filter was applied to the Assembly field in the Filters pane, configured to show the "Top 5" items based on the Completed Trips measure.

Visualization:



The visual displays the following top 5 assemblies by completed trip volume, from highest to lowest:

1. Ramanagaram
2. Yeshwantpur
3. Bangalore South
4. Dasarahalli
5. Gandhi Nagar

**Insights:**

**Dominant Performers:** Ramanagaramclearly stands out as the assembly with the highest number of completed trips. Yeshwantpur and Bangalore South follow as strong contenders, forming the next tier of high-volume zones.

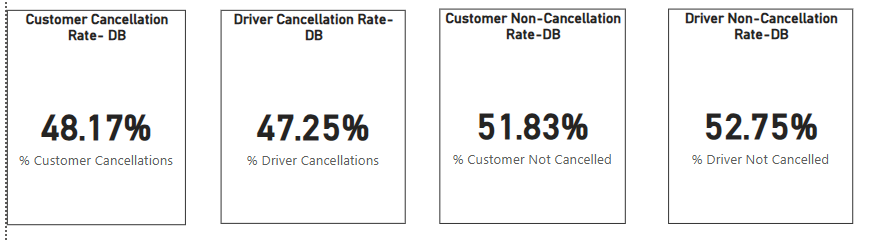
1. Basic Analytical Tasks [8 Marks]
   * 2.9.1   
     What are the percentages of cancellations and successful rides by both driver and customer? [3 marks]

*Solution:*

## **Methodology**

The analysis utilized the 'Trip Details' table to determine cancellation and successful ride percentages. Key metrics included Total Trips (count of rows in 'Trip\_Details') , Customer Cancellations (customer\_not\_cancelled = 0) , Driver Cancellations (driver\_not\_cancelled = 0) , Customer Not Cancelled (customer\_not\_cancelled = 1) , Driver Not Cancelled (driver\_not\_cancelled = 1) , and Overall Completed Rides (end\_ride = 1). These counts were converted into percentages by dividing by the Total Trips. Power BI Card visuals were used to display each distinct percentage.

Visualization:



* + 2.9.2  
    Analyse the percentage of people who completed trips after searching for quotes. Visualise the variation of this ratio by time periods.  
    [5 marks]

*Solution:*

## **Methodology**

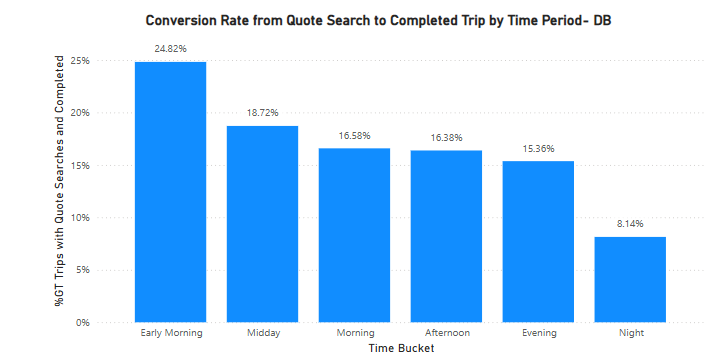
The analysis aimed to determine the percentage of trips completed after a user searched for quotes and visualize its variation by hour. This involved using the 'Trip Details' table for searches\_for\_quotes and end\_ride status, and the 'Duration' table for the duration (hour of trip).

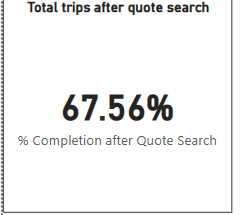
Three DAX measures were created in Power BI:

1. Total Trips with Quote Searches: Counts trips where searches\_for\_quotes = 1.
2. Trips with Quote Searches and Completed: Counts trips where searches\_for\_quotes = 1 AND end\_ride = 1.
3. % Completion after Quote Search: Calculated as [Trips with Quote Searches and Completed] divided by [Total Trips with Quote Searches].

A Power BI Line chart was used to visualize the trend, with duration on the X-axis and % Completion after Quote Search on the Y-axis

Visualization:





**Insights:**

Namma Yatri experiences its peak volume of completed rides in the early morning, with this period contributing 24.8% of all completed trips.

1. Create a Parameter and Use Filters [5 Marks]
   * Create a parameter and use it as a filter on an appropriate subset of the data to interactively analyse and visualise different subsets of the data.
   * Explain your choice of filter and insights drawn from this step.

*Solution:*

**Methodology:**

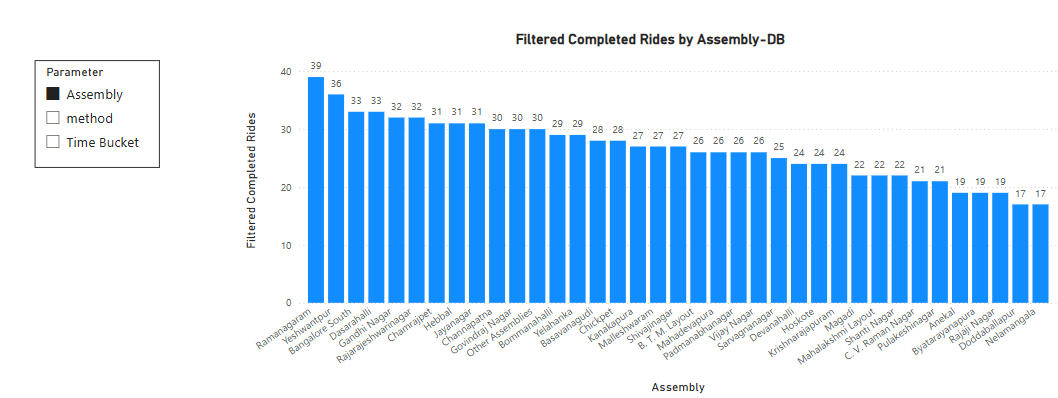
The methodology for interactive analysis involves creating a single dynamic filter control. This is achieved by:

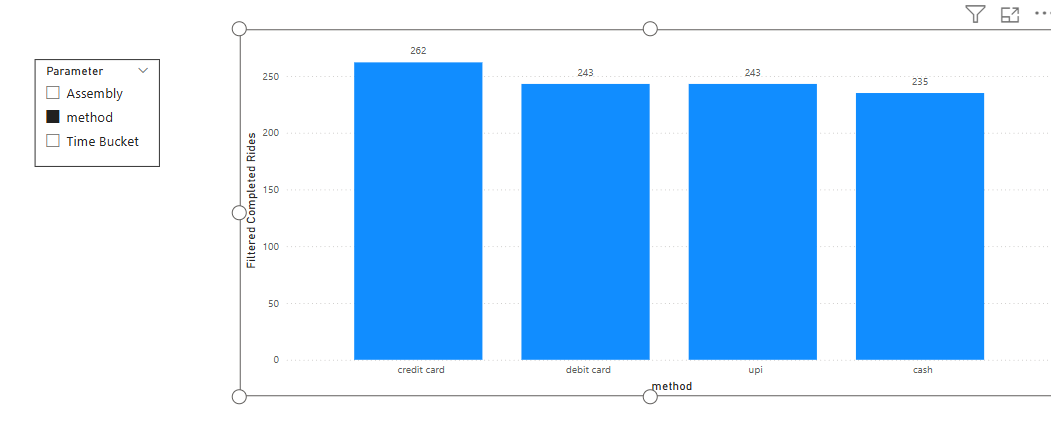
1. Defining a **Field Parameter** that encapsulates Assembly, Duration (Time Bucket), and Payment Method.
2. Placing this Field Parameter on a **single slicer** for user interaction.
3. Developing **DAX measures** (like Filtered Completed Rides) that use TREATAS to respond to the selections made by *any* of the underlying dimensions (Assembly, Time Bucket, Payment Method).
4. Utilizing the Field Parameter on the **axis of visuals** to dynamically change the dimension by which data is grouped and displayed, allowing users to switch their analysis focus instantly.

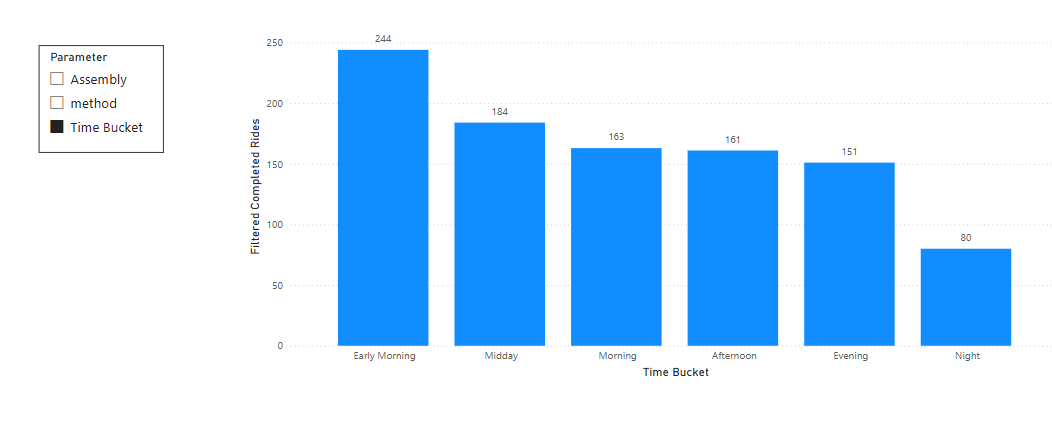
**Filter Choices & Key Insights:**

* **Chosen Parameters:** Assembly (geographic areas), Duration (time buckets), and Payment Method were selected as filters.
* **Rationale:** These allow for interactive analysis of ride completions across critical dimensions, helping to identify demand patterns and operational effectiveness.
* **Key Insights:**
  + **Assembly:** Revealed high-performing zones (e.g., Ramanagaram, Yeshwanthpur) indicating strong demand for targeted strategies.
  + **Duration\_Label:** Showed "Early Morning" as the peak period for completed rides, vital for optimizing driver allocation and promotions.
  + **Payment Method:** Identified "Credit Card" as the most used method, guiding decisions on payment convenience and loyalty programs.

Visualization:







### 3. Conclusion [20 Marks]

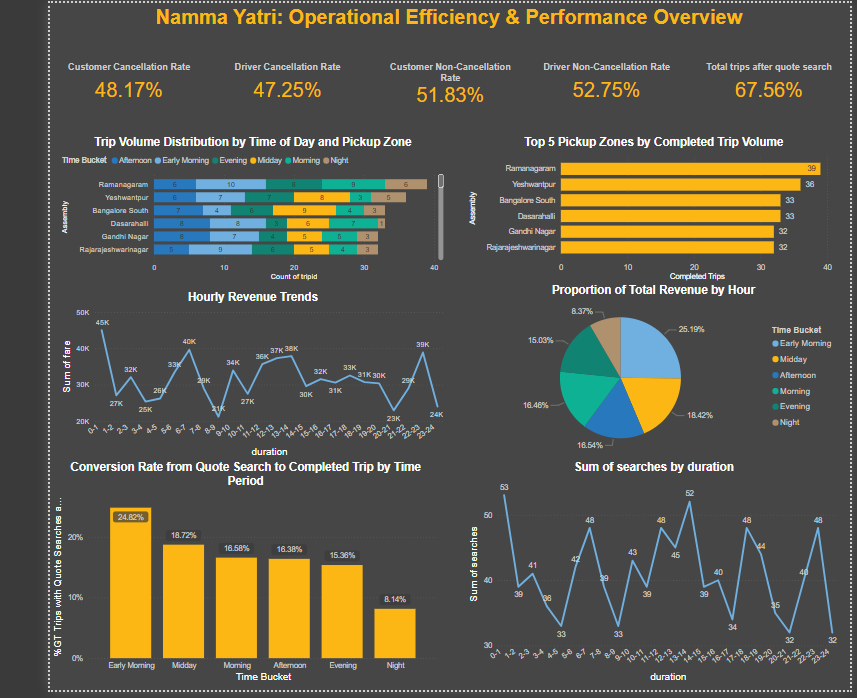
1. Recommendations for Operational Efficiency [10 Marks]
   * Based on your findings from the analysis, provide recommendations on how Namma Yatri can optimise its operations.
   * This could include strategies for improving resource allocation, reducing cancellations, or optimising ride durations.
   * Add supporting dashboards.

*Solution:*

#### Recommendations:

1. **Optimized Driver Allocation based on Demand and Revenue Peaks:**
   * **Insight:** Ride demand peaks between 8-10 AM and 6-8 PM, aligning with revenue peaks. "Early Morning" (0-6 AM) contributes significantly to total revenue (25.19%). High-performing zones include Ramanagaram, Yeshwantpur, Bangalore South, and Dasarahalli.
   * **Recommendation:** Implement a dynamic driver deployment strategy. Increase driver supply during identified peak demand/revenue periods. Offer incentives for drivers operating in "Early Morning" hours and prioritize driver availability in high-performing zones like Ramanagaram, Yeshwantpur, Bangalore South, and Hebbal (highest revenue-generating assemblies).
   * **Outcome:** Improved rider experience, reduced wait times, and maximized revenue capture.
2. **Strategies to Reduce Cancellations:**
   * **Insight:** High customer (48.17%) and driver (47.25%) cancellation rates indicate significant unfulfilled trips.
   * **Recommendation:** Investigate root causes for driver cancellations (e.g., fare, dispatch distance). For customer cancellations, improve real-time ETAs and fare transparency. Optimize the matchmaking algorithm to minimize time between request and driver acceptance.
   * **Outcome:** Reduction in unfulfilled trips, increased completed rides, and improved customer satisfaction.
3. **Enhance Conversion from Quote Search to Completed Trip:**
   * **Insight:** Overall conversion rate is 67.56%, with "Early Morning" showing the highest (24.82%) and "Night" the lowest (8.14%).
   * **Recommendation:** Improve conversion in lower-performing time buckets. Ensure accurate dynamic pricing, explore mechanisms for guaranteed driver acceptance post-quote, and consider post-quote engagement for non-converting searches.
   * **Outcome:** Higher conversion of quote searches to completed rides, impacting revenue and service utilization.

**Supporting Dashboard:**

****

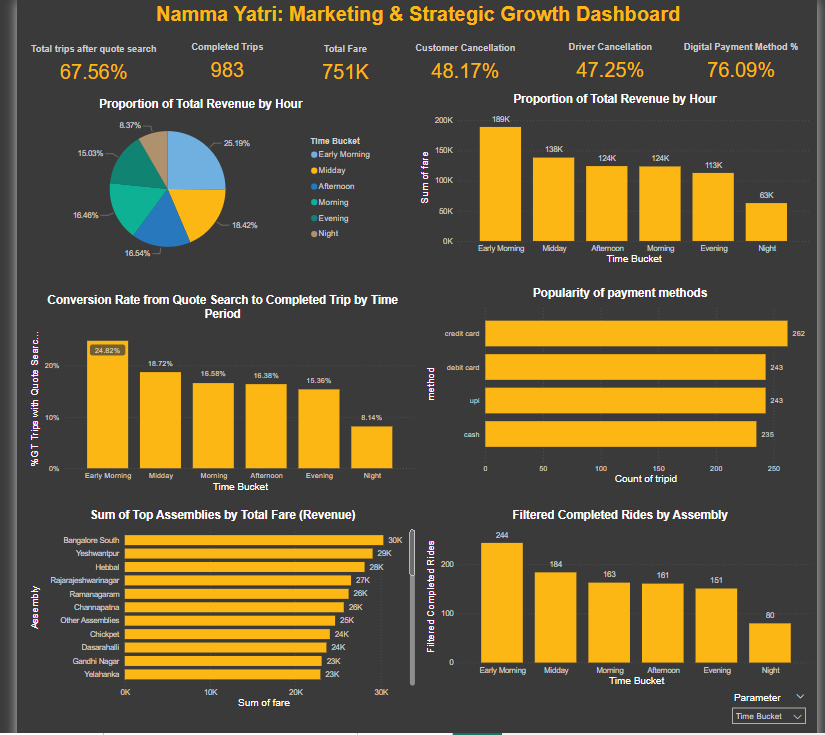
1. Marketing and Operational Strategy Improvements [10 Marks]
   * Suggest improvements to Namma Yatri’s marketing or operational strategies based on your analysis.
   * Recommendations could involve promotional efforts, driver incentives, or regional targeting to increase customer satisfaction and service efficiency.
   * Add supporting dashboards.

*Solution:*

#### Recommendations:

1. **Strategic Promotional Efforts:**
   * **Insight:** Analysis reveals distinct peak demand and revenue periods (e.g., morning and evening commutes, "Early Morning" for revenue share) and identifies high-performing zones (e.g., Ramanagaram, Yeshwantpur, Bangalore South). Additionally, **Credit Card (26.65%)** is the most frequently used payment method, followed closely by Debit Card and UPI (both 24.72%), while **Cash is the least used (23.91%)**.
   * **Recommendation:**
     + **Targeted Marketing Campaigns:** Launch location-specific marketing campaigns with promotional offers (e.g., discounts, loyalty points) in high-performing zones like Ramanagaram, Yeshwantpur, and Bangalore South to further consolidate market share and reward loyal customers.
     + **Off-Peak Hour Incentives:** To balance demand and utilize driver availability, introduce time-bound promotions or reduced fares during off-peak hours (e.g., late night, mid-morning lull) to encourage ridership.
     + **Digital Payment Adoption Drive:** Given the strong preference for digital payments (Credit Card, Debit Card, UPI), collaborate with banks and payment providers for exclusive offers (e.g., cashback on specific card types, UPI transaction bonuses). Actively promote the benefits of digital payments to encourage the shift away from cash, improving transaction efficiency and security.
   * **Outcome:** Increased ridership across all periods, improved customer loyalty, more balanced demand distribution, and enhanced digital transaction rates.
2. **Enhanced Driver Incentives for Service Quality and Retention:**
   * **Insight:** Driver cancellation rates stand at 47.25%, indicating a need for greater driver satisfaction and commitment.
   * **Recommendation:**
     + **Performance-Based Bonuses:** Implement a tiered incentive structure that rewards drivers for maintaining high ride completion rates, particularly during peak demand hours, and for consistently serving high-demand/high-revenue zones.
     + **Long-Distance Trip Incentives:** Offer additional bonuses for accepting longer trips or rides to less popular areas to reduce driver cancellations in these scenarios.
     + **Driver Support Programs:** Explore initiatives like reduced commission rates during specific periods, fuel subsidies, or vehicle maintenance partnerships to improve driver earnings and retention.
   * **Outcome:** Improved driver morale, reduced driver cancellations, higher service reliability, and increased driver loyalty to the platform.
3. **Refined Regional Targeting Strategies:**
   * **Insight:** Specific assembly zones demonstrate varying levels of trip volume and revenue generation, indicating unique market characteristics.
   * **Recommendation:**
     + **Localized Operational Adjustments:** Develop zone-specific operational plans. For instance, in zones with lower overall completed trips, analyze local competitive landscapes and tailor driver onboarding or marketing efforts accordingly.
     + **Growth in Promising Areas:** Identify areas with growing demand but currently lower supply, and strategically invest in driver recruitment and targeted marketing to foster growth in these nascent markets.
     + **Community Engagement:** Engage with local communities in key zones to understand their unique transportation needs and tailor Namma Yatri's service offerings or features to meet those demands.
   * **Outcome:** More efficient market penetration, optimized resource deployment based on regional dynamics, and increased relevance of services to local populations.
4. **Optimizing Customer Journey & Experience:**
   * **Insight:** The overall conversion rate from quote search to completed trip is 67.56%, but shows variability across time buckets (e.g., high in "Early Morning," low in "Night"). High customer cancellation rates (48.17%) also indicate friction.
   * **Recommendation:**
     + **Seamless In-App Experience:** Continuously optimize the app's user interface for ease of use, focusing on quick quote generation and streamlined booking.
     + **Transparency and Predictability:** Enhance transparency around fare estimations and potential wait times before a trip is confirmed to build customer trust and reduce pre-trip cancellations.
     + **Robust Feedback Loop:** Implement accessible in-app feedback mechanisms for both riders and drivers to quickly identify and address pain points, ensuring continuous improvement of the service.
   * **Outcome:** Increased customer satisfaction, higher conversion rates from search to completion, and a stronger brand reputation.

**Supporting Dashboard:**

**s**