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)

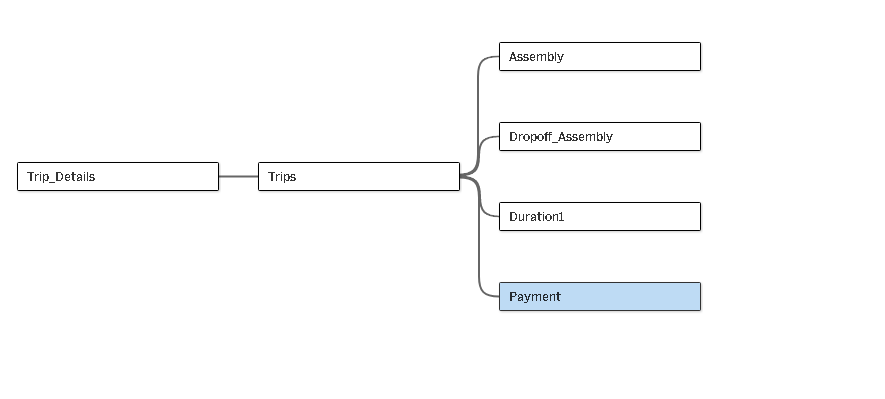
## Analysis and Visualisation

### 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)*



Trip Details and Trips: added tripid as the matching field

Trips and Assembly: Joined Trips.loc\_from to Assembly.ID.

Trips and Assembly: renamed it “Dropoff\_Assembly” , and joined Trips.loc\_to to

Dropoff\_Assembly.ID.

Trips and Duration: Joinef Trips.duration to Duration.id.

Trips and Payment: Joined Trips.faremethod to Payment.id.

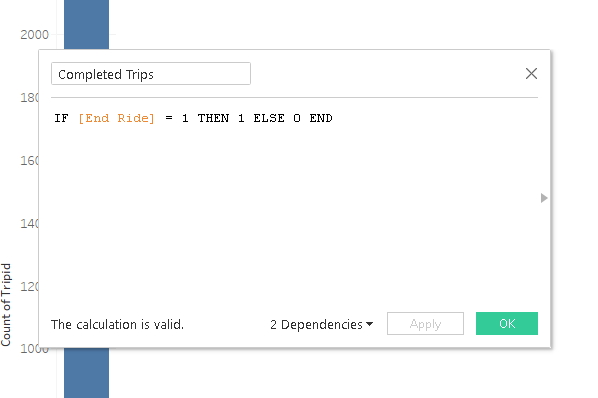
1. Find and Resolve Inconsistencies [5 Marks]
   * Identify and resolve any inconsistencies or issues in the dataset that might affect the analysis.
   * Clean the data to ensure it is structured properly for analysis, removing any irrelevant, duplicate, or erroneous entries.
   * While performing the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

*Solution:*

There are no null values found in the data.

Ther are no dublicates as well

created Completed Trips



### 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:*

**Field like** fare, distance, searches are numerical

Fields like Assembly, duration, method are groups

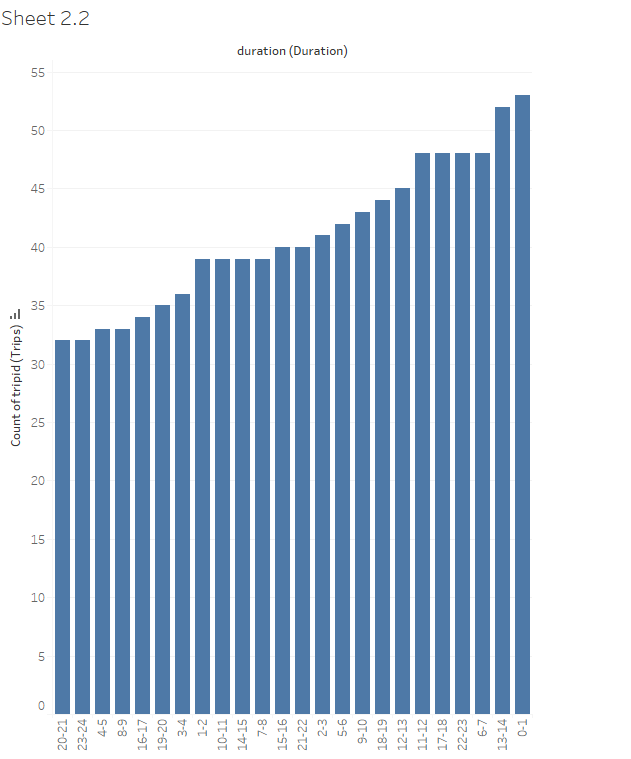
And end\_ride is catagorical

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:*

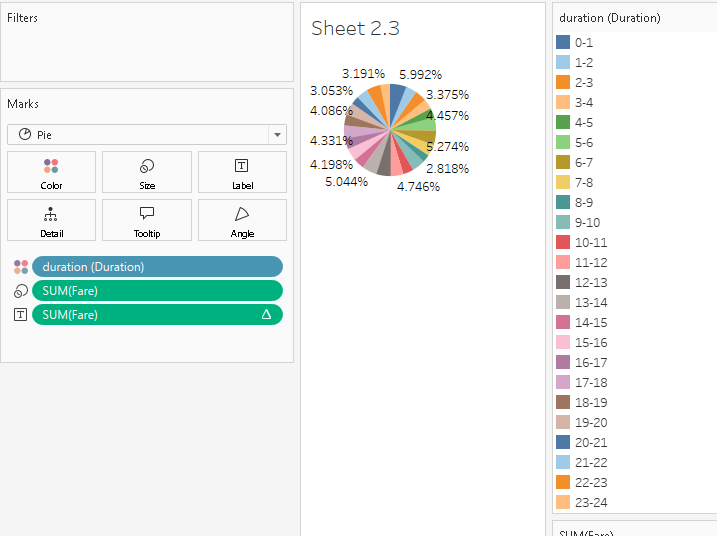
The tallest ones are the busiest hour’s data is suggesting 6-7,13-14,0-1 are busiest

Insight for Report: Peak demand is at 6-7,13-14,0-1 hours so more drivers might be needed then.



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:*

****

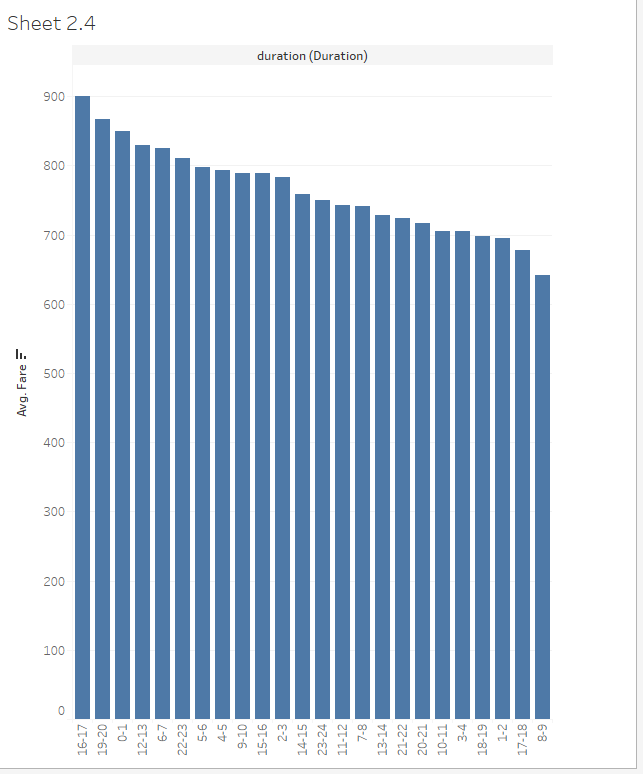
Hour 0-1 gives 5.99% of revenue, showing it’s a high-earning time.

2nd is 6-7

3rd is 13-14

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:*

****

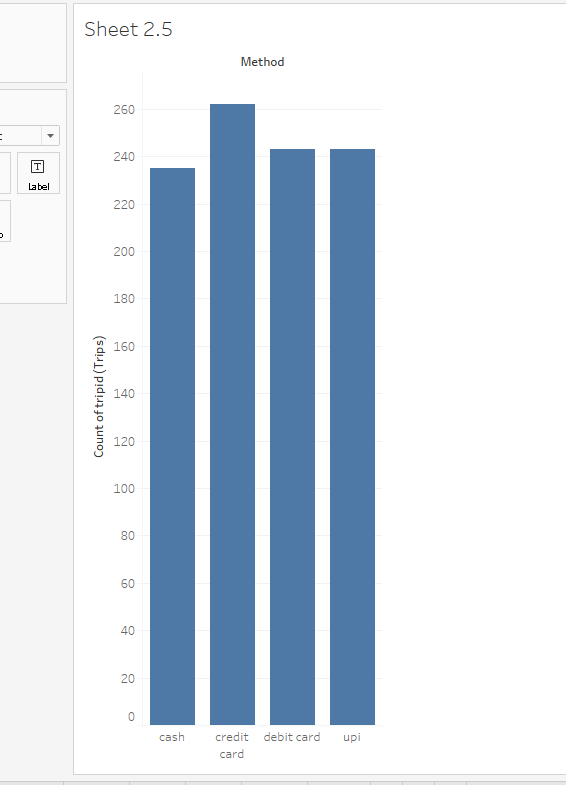
Hour [16-17] has higher fares, maybe from longer trips or surge pricing.

Fallowed by 19-20 ,0-1 and so on

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.

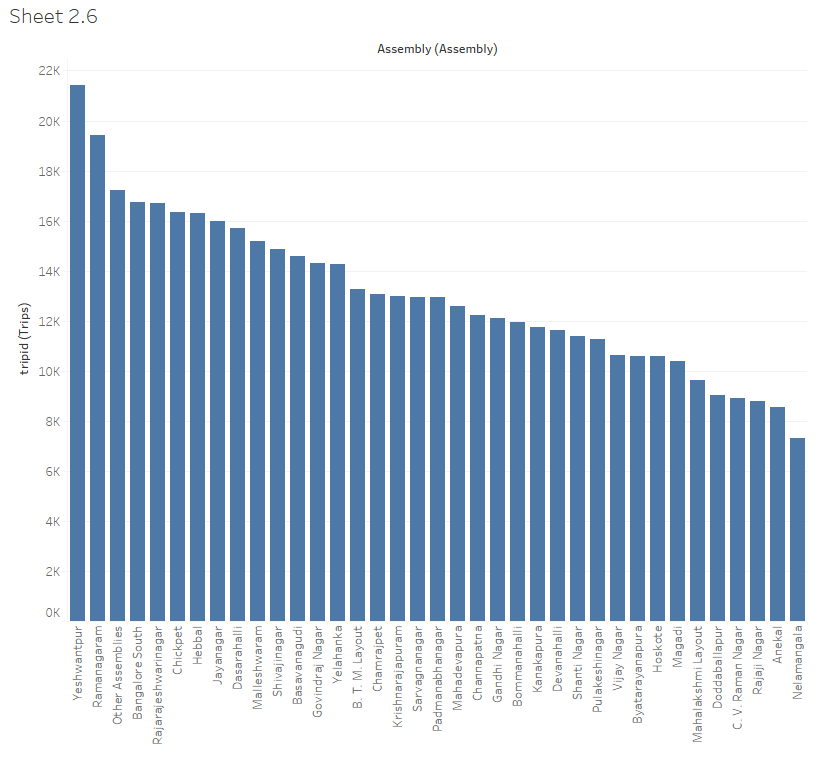
*Solution:*

**Most poplar one is credit card**

****

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:*



Zone yeshwantpur earns the most, showing it’s profitable

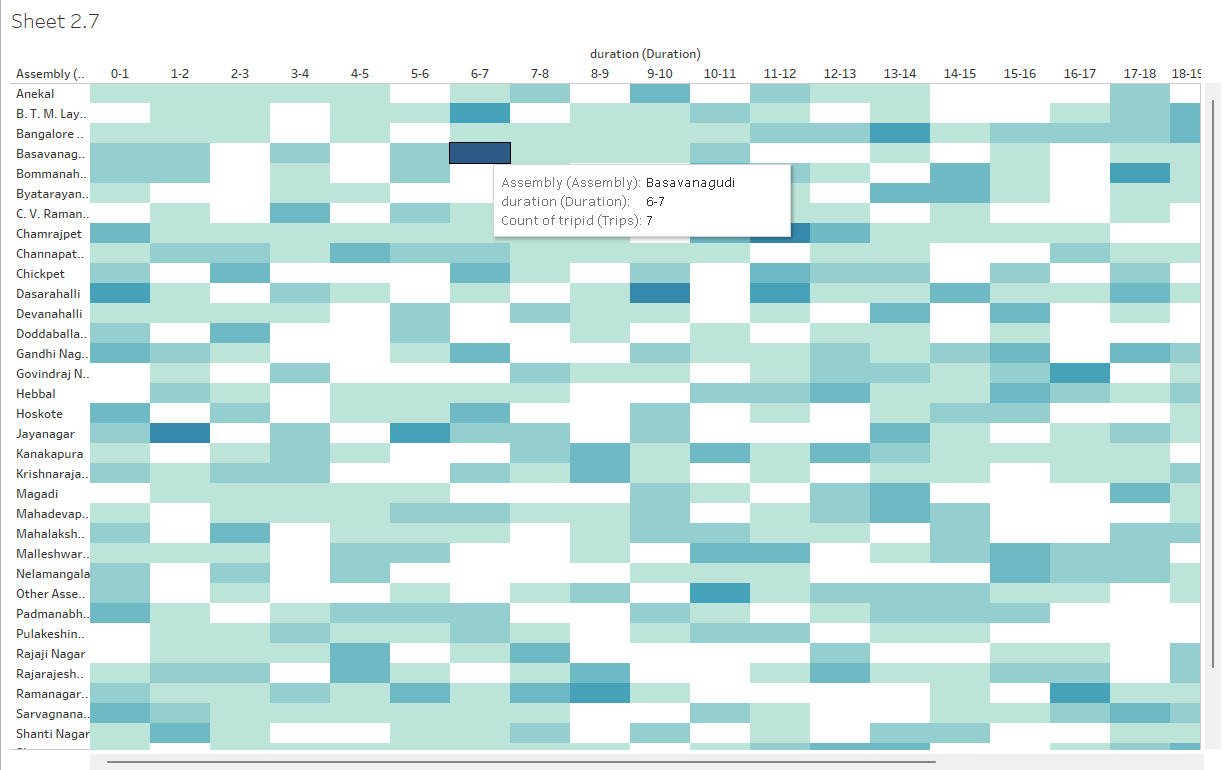
* + 2.6.2. Revenue: Identify pickup zones generating the highest revenue.  
    [3 marks]

*Solution:*

Zone yeshwantpur earns the most, showing it’s profitable

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

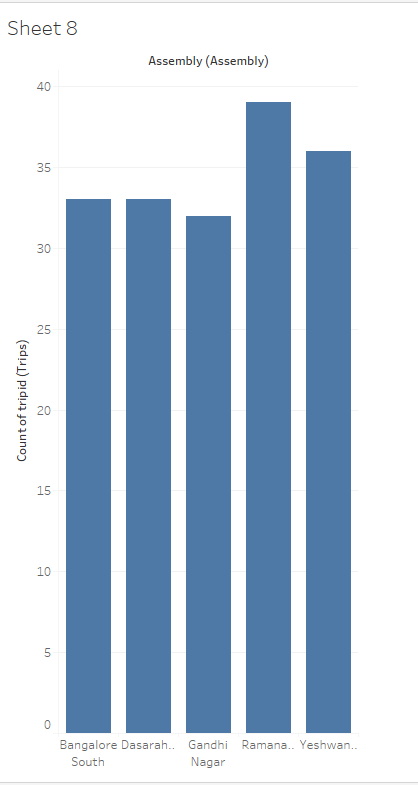
*Solution:*

****

Zone baswangudi is busy at [6-7], helping plan driver shifts

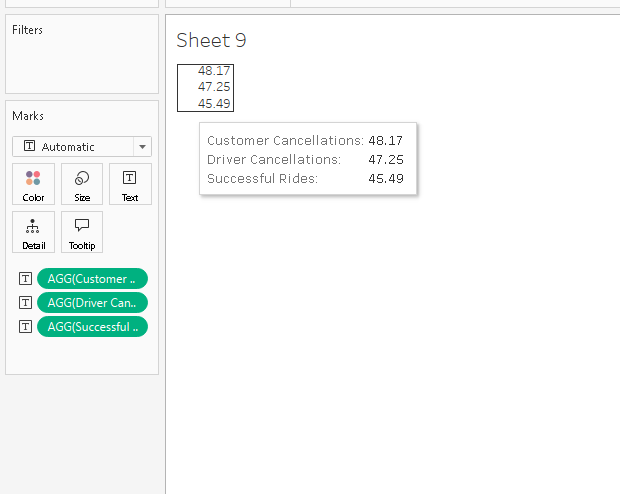
1. Top Zones with Highest Trip Volume [3 Marks]
   * Identify the top 5 pickup zones with the highest total number of completed trips.
   * Analyse factors contributing to the higher number of trips.

*Solution:*

****

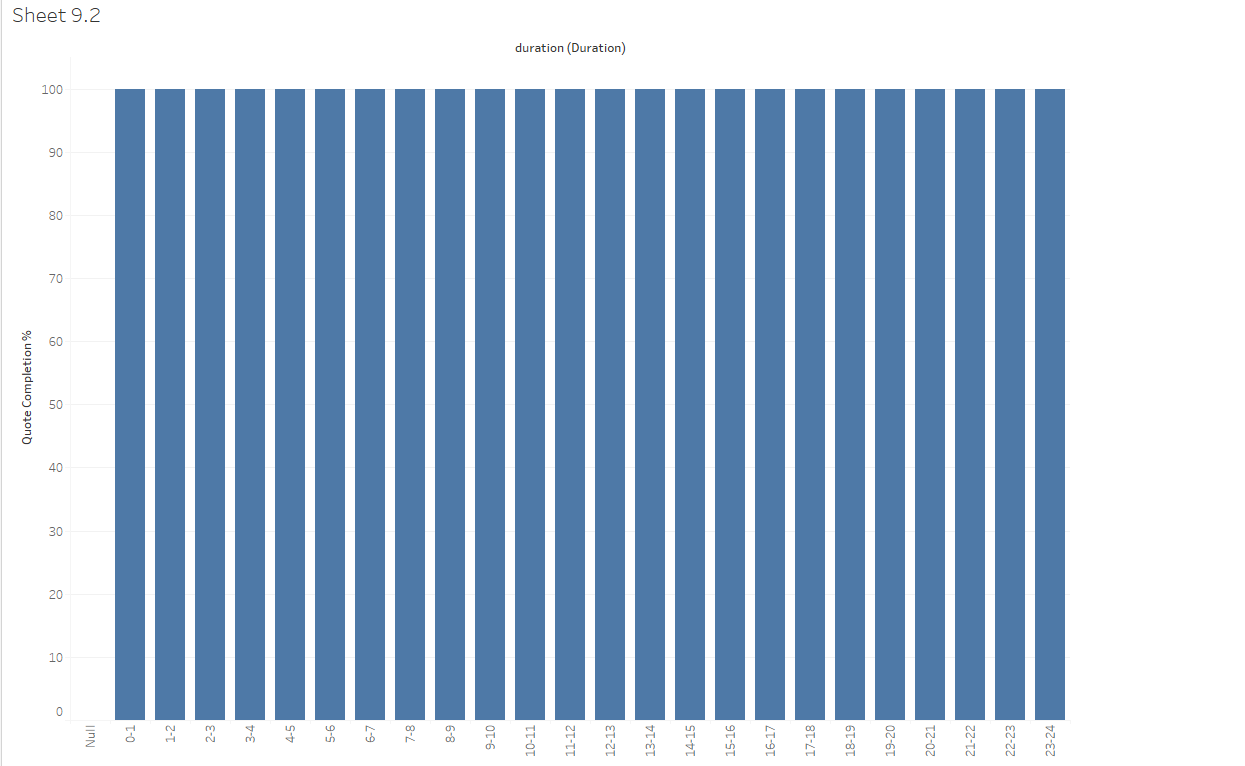
1. Basic Analytical Tasks [8 Marks]
   * 2.9.1   
     What are the percentages of cancellations and successful rides by both driver and customer? [3 marks]

*Solution:*

****

* + 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:*

****

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:*

### 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:*

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:*