Namma Yatri Methodology Report

1. Problem Statement

Namma Yatri aims to optimise its ride-hailing service in Bengaluru by uncovering actionable insights through data analysis. The objective is to:

- Identify peak demand periods.
- Understand customer and driver cancellation behaviour.
- Optimise payment methods.
- Identify high-performing geographical zones.

2. Data Preparation

2.1 Data Sources

The dataset comprises five tables:

- **Assembly:** Geographical zones in Bengaluru.
- **Duration:** Time periods of the day.
- **Payment:** Payment methods used.
- **Trip Details:** Customer and driver interactions.
- **Trips:** Journey details including fares and locations.

2.2 Data Import in Tableau

- Imported `nammayatri.xlsx` into Tableau.

2.3 Data Joining Strategy

Base Table	Join Column	Joined Table	Join Type
Trips	tripid	Trip Details	Inner
Trips	faremethod	Payment	Left
Trips	duration	Duration	Left
Trips	loc_to	Assembly	Left

2.4 Data Cleaning Steps

- Removed rows with **null or zero fares.**
- Standardised the `method` column in **Payment Table** using Calculated Fields.
- Filtered out **duplicate trip IDs**.
- Created calculated fields for cancellation rates, trip duration, and OTP status.

3. Exploratory Data Analysis (EDA)

3.1 Classification of Variables

- **Categorical:** Assembly, duration, method, loc_from, loc_to, faremethod, driverid, custid, otp_entered.
- **Numerical:** fare, distance, searches, searches_got_quotes.

3.2 Ride Demand Over Time

- Peak demand at **1 PM - 2 PM** with 195 searches.
- Additional peaks: **11 AM - 12 PM**, **5 PM - 6 PM**.
- **Visual:** Line chart (Group Identifier: DataTrio_CharuAshAkanksha)

3.3 Revenue by Time Periods

- Revenue highest during **12-2 PM**, **0-1 AM**, **4-5 AM**, **6-7 AM**, **11 PM-12 AM**.
- **Visual:** Pie chart (Group Identifier: DataTrio_CharuAshAkanksha)

3.4 Trip Hour vs Revenue

- Positive revenue peaks in midday and late night.
- **Visual:** Dual axis graph (Group Identifier: DataTrio_CharuAshAkanksha)

3.5 Payment Method Preferences

- **Most Used:** Credit Card (15,648 rides)
- Followed by UPI and Debit Card.
- **Visual:** Bar chart (Group Identifier: DataTrio_CharuAshAkanksha)

3.6 Zone-wise Performance

- **Top by Rides:** Mahadevapura, Ramanagaram, Gandhi Nagar.
- **Top by Revenue:** loc 6, 18, 17.
- **Visual:** Heatmap (Group Identifier: DataTrio_CharuAshAkanksha)

3.7 Ride Time Periods Across Zones

- Example: Zone 53 busiest late night, Zone 52 in afternoons.
- **Visual:** Stacked bar (Group Identifier: DataTrio_CharuAshAkanksha)

3.8 Top 5 Zones by Trip Volume

1. Ramanagaram
2. Yeshwantpur
3. Dasarahalli
4. Bangalore South
5. Rajarajeshwarinagar

3.9 Cancellations & Conversions

- Customer cancellation: ~12%
- Driver cancellation: ~8%
- Conversions higher in evenings.
- **Visual:** Line plot (Group Identifier: DataTrio_CharuAshAkanksha)

3.10 Parameters & Filters

- **Parameter:** Minimum Fare filter to focus on high-value trips.
- Insight: High-fare trips contribute significantly to total revenue.

4. Key Recommendations

4.1 Operational Efficiency

- Dynamic driver allocation in Mahadevapura, Ramanagaram, Gandhi Nagar.
- Promote Credit Card payments via offers.
- Reduce cancellations with training and incentives.

4.2 Marketing & Business Strategy

- Targeted promotions for high-fare riders.
- Localised marketing in underperforming zones.
- Loyalty programs and dynamic pricing strategies.

5. Dashboards Attached (Group Identifier: DataTrio_CharuAshAkanksha)

1. **Demand Trends**
2. **Revenue Breakdown**
3. **Payment Preferences**
4. **Zone-wise Heatmaps**
5. **Cancellations & Conversion Analysis**