Table of Contents

1. Dataset Overview

This synthetic dataset is designed for Strategic Marketing MSc students to analyze customer acquisition, retention, and campaign success for PickMe Rides.

Files Included:
- dim_customer.csv: 5,000 rows of customer profiles and acquisition data
- dim_driver.csv: 800 rows of driver attributes
- dim_campaign.csv: 8 rows of campaign metadata
- fact_campaign_spend.csv: 649 rows of daily spend, impressions, clicks
- fact_rides.csv: 86,911 rows of ride transactions
- dim_date.csv: 365 rows of calendar dates

All monetary values are in LKR (Sri Lankan Rupees). Attribution uses a last-click model.

2. Power BI Data Model Setup

Relationships:
- fact_rides[customer_id] -> dim_customer[customer_id]
- fact_rides[driver_id] -> dim_driver[driver_id]
- fact_rides[attributed_campaign_id] -> dim_campaign[campaign_id]
- dim_customer[acquisition_campaign_id] -> dim_campaign[campaign_id]
- fact_campaign_spend[campaign_id] -> dim_campaign[campaign_id]

Date Intelligence:
- Create Request Date column in fact_rides using DATEVALUE
- Relate dim_date[date] to fact_rides[Request Date] and fact_campaign_spend[date]
- Mark dim_date as a Date table

3. KPI Definitions

Acquisition:
- New Customers: Count of customers with is_first_ride = TRUE
- CAC: Spend divided by New Customers
- CVR: New Customers divided by Clicks

Retention:
- Active Customers (30d): Customers with a completed ride in last 30 days
- Churned (60d): Customers whose last ride was more than 60 days ago
- Reactivation: Churned customers who returned in last 30 days
- Cohort Retention: Fraction of a signup cohort active in month N

Campaign Success:
- ROAS: Attributed Revenue divided by Spend
- CPR: Spend divided by Completed Rides
- CTR: Clicks divided by Impressions

4. DAX Measures (Plain Text)

Examples of DAX Measures:

Request Date: Create a column using DATE(YEAR, MONTH, DAY) from request_datetime

New Customers: Count of customer_id where is_first_ride is TRUE

Active Customers (30d): Count of distinct customer_id with completed rides in last 30 days

Churned Customers (60d): Count of customers whose last completed ride date is more than 60 days ago

Retention Rate (Monthly): Calculate cohort start month and current month, then divide active customers in current month by total cohort customers

CAC: Spend divided by count of new customers using USERELATIONSHIP to align dates

ROAS: Net Fare from attributed rides divided by campaign spend

Customer Segment: Categorize customers as Prospect, Active, At Risk, or Churned based on last ride date

5. Dashboard Design Suggestions

Executive Summary:
- KPI Cards: Completed Rides, Net Fare, Gross Margin, New Customers, CAC, ROAS
- Line Chart: Completed Rides and Net Fare over time
- Sparklines: Summary by channel or campaign

Acquisition and Funnel:
- Stacked Column: New Customers by channel
- Funnel Chart: Impressions to Clicks to New Customers
- Scatter Plot: ROAS vs CAC with bubble size as Spend

Retention and Cohorts:
- Matrix Heatmap: Cohort Month vs Active Month
- Line Chart: Active Customers (30d) trend

Campaign Effectiveness:
- Waterfall Chart: Fare to Margin breakdown
- Column Chart: Attributed Revenue and Spend by Campaign

Customer Lifecycle:
- Table: Customer Segment with counts and Net Fare

Geography and Ops:
- Map: Pickup city with Completed Rides
- Bar Chart: Payment method vs Net Fare

6. Validation and Troubleshooting Tips

Common Issues:
- MAX function error: Use MAXX or create a Request Date column
- STARTOFMONTH error: Use DATE(YEAR, MONTH, 1) instead of scalar in STARTOFMONTH
- Blank visuals: Ensure date relationships are active and slicers are aligned
- Incorrect totals: Use USERELATIONSHIP to align date paths

Sanity Checks:
- Missing Dates in fact_rides or fact_campaign_spend: Compare against dim_date
- Relationship setup: Single-direction, active relationships from dim_date to facts
- Avoid bi-directional filters on date tables