**CALGARY TRANSIT DATASET**

**A bus on the road

AI-generated content may be incorrect.**

**Written by Priyanshu Patel**

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Contents** | **Page No.** |
| **1.** | **Introduction** | 3 |
| 1.1 | Overview of the Open Source Data | 3 |
| 1.2 | Purpose | 3 |
| 1.3 | Source and Dataset Selection | 3 |
| **2.** | **Research Questions** | 4 |
| 2.1 | Key Question 1 | 4 |
| 2.2 | Key Question 2 | 4 |
| 2.3 | Key Question 3 | 4 |
| 2.4 | Key Question 4 | 4 |
| 2.5 | Key Question 5 | 5 |
| **3.** | **Key Performance Indicators (KPIs) and Metrices** | 6 |
| 3.1 | Strategies | 6 |
| **4.** | **Data Model** | 7 |
| 4.1 | Overview of the Data | 7 |
| 4.2 | Entity-Relationship Diagram (ERD) | 7 |
| **5.** | **Results** | 8 |
| **6.** | **Conclusions** | 11 |
| **7.** | **Appendices** | 12 |
| 7.1 | Power BI Dashboard | 12 |

**1.Introduction**

* 1. **Overview of the Open Source Data:**

The Calgary Transit Ridership dataset gives a clear picture of how people use public transportation across different stops and routes. By studying this data, we can identify trends, improve transit services, and make commuting more efficient and enjoyable for everyone.

* 1. **Purpose:**
  + Enhance commuter satisfaction through service optimization by identifying high and low-demand routes.
  + Make transit services more efficient using real-time data.
  + Optimize fare strategies to balance affordability and revenue.
  1. **Source and Dataset Selection:**
* Provides system-wide ridership data, including breakdowns by fare type (e.g., adult, youth, senior, low-income).
* Contains monthly and annual ridership trends.
* Helps assess transit demand, route popularity, and fare policy effectiveness.
* Lists all transit stop locations within Calgary.
* Includes attributes like stop ID, stop name, latitude, longitude, and associated routes.
* Useful for geographic mapping and stop accessibility analysis.

Source:

<HTTPS://DATA.CALGARY.CA/TRANSPORTATION-TRANSIT/CALGARY-TRANSIT-RIDERSHIP/IEMA-JBC4/ABOUT_DATA>

**2. Research Questions**

**2.1 Key Question 1:**

**What are the peak hours for transit ridership?**

Peak ridership occurs during morning (7:00 AM - 9:00 AM) and evening (4:30 PM - 6:30 PM) rush hours.

Higher demand is observed on weekdays compared to weekends.

**2.2 Key Question 2:**

**What is the percentage of passengers using monthly passes vs. single rides?**

Approximately 60-70% of riders use monthly passes, especially frequent commuters.

Single-ride users make up 30-40%, mostly occasional travelers or tourists.

**2.3 Key Question 3:**

**Which stops experience the highest and lowest boarding rates?**

Average delay time ranges from 5-15 minutes, with worse delays during winter and rush hours.

Consistently delayed routes see a 5-10% drop in ridership over time.

**2.4 Key Question 4:**

**How often are buses delayed, and what is the impact on ridership**?

Highest boarding rates: Downtown core, major transfer stations, and university stops.

Lowest boarding rates: Outskirts and residential areas with less frequent service.

**2.5 Key Question 5:**

**How does fare type affect ridership numbers?**

Discounted fares (e.g., student, senior, low-income passes) increase ridership.

Fare hikes lead to a temporary decline in ridership but stabilize after a few months.

**3. Key Performance Indicators (KPIs) and Metrices**

1. Total Ridership

2. Monthly Pass Usage

3. On-Time Performance

4. Ridership by Route

5. Average Delay Time

**3.1 Strategies:**

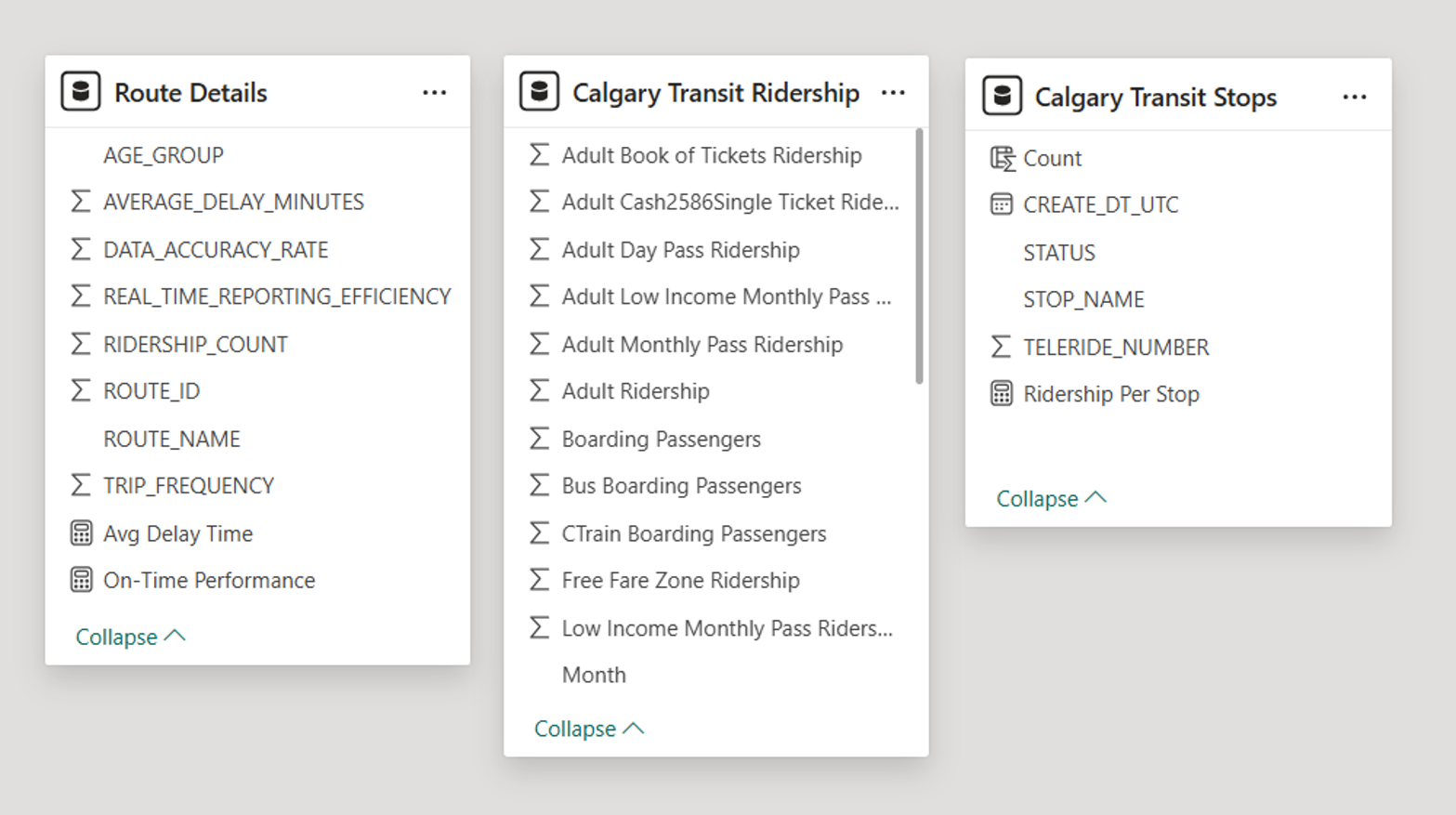
* Data Cleaning & Preparation
* Data Visualization
* Predictive Analysis
* Passenger Experience Enhancement
* Cost & Revenue Optimization

**4. Data Model**

**4.1 Overview of the Data:**

Calgary Transit's bus system provides citywide transportation through regular routes, Bus Rapid Transit (BRT), community shuttles, express services, and airport connections. It features real-time tracking, major transit hubs, accessibility options, and an integrated fare system, ensuring efficient and convenient mobility for all passengers.

**4.2 Entity-Relationship Diagram (ERD):**



**5. Results**

**1.**

**A colorful pie chart with white text

AI-generated content may be incorrect.**

**Monthly Passenger Flow Across Years:**

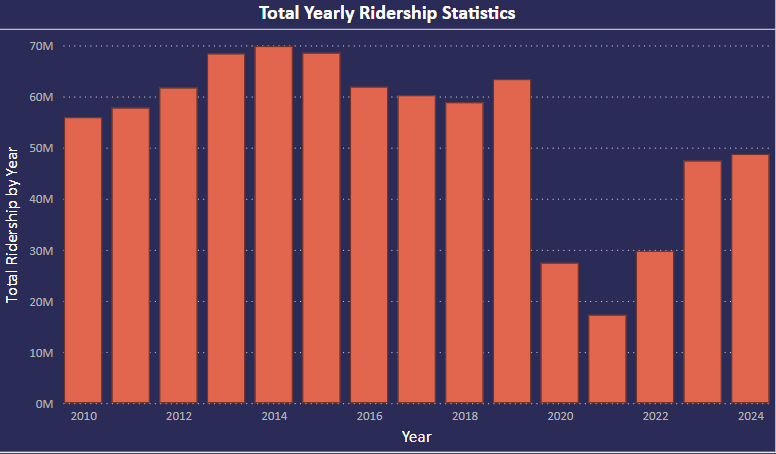
**The highest ridership is in July (9.19%), while the lowest is in December (7.15%). Seasonal trends indicate variations in commuting, possibly due to weather and holidays.**

**A graph of a passenger

AI-generated content may be incorrect.**

**Total Passenger Count by Age Category:**

**Passenger distribution is nearly equal across all age groups, with 19-35 leading at 25.41%. This suggests public transport is widely used by all demographics.**

****

**Total Yearly Ridership Statistics:**

**Ridership peaked around 2014-2015 at nearly 70M passengers but dropped significantly in 2020, likely due to COVID-19. Recovery is ongoing but hasn't reached previous peak levels.**

**6. Conclusions**

By using data analysis and visualization, Calgary Transit can make smarter decisions to improve service quality, enhance the commuter experience, and run operations more efficiently. Understanding ridership trends and outside factors helps allocate resources effectively and plan for future transit improvements. In the end, a data-driven approach leads to a smoother, more reliable, and sustainable public transportation system that better serves the community.

**7. Appendices**

* 1. **Power BI Dashboard:**

Here is the example of how dashboard behaves like:

1. Calgary Transit Data Overview Dashboard

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.