

INFSCI 2415 Final Report: PRT Bus Information Visualization

1: Project Overview:

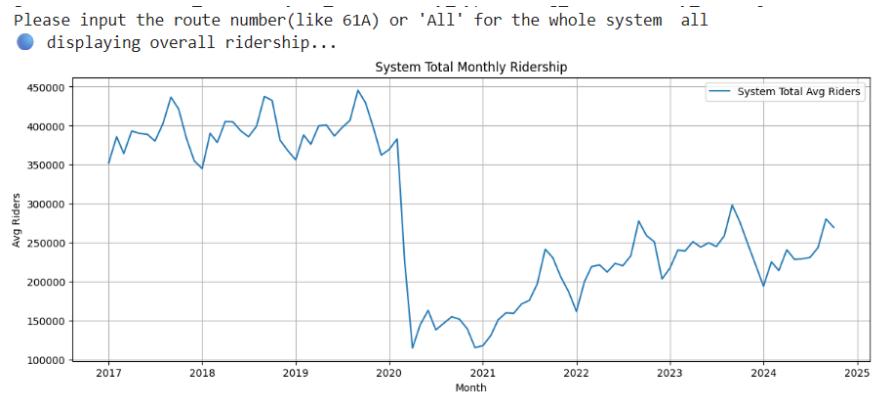
Background: Bus is one of the most common public transportation vehicles, and a reliable way to show the convenience of a city. A reliable bus ridership (ons and offs count) will be necessary for analysis from buses companies and city managers.

Target: Provide an interactive tool to search for specialized route passengers' ridership in months.

The result: An interactive board, route ridership visualize graph and a Python module structure

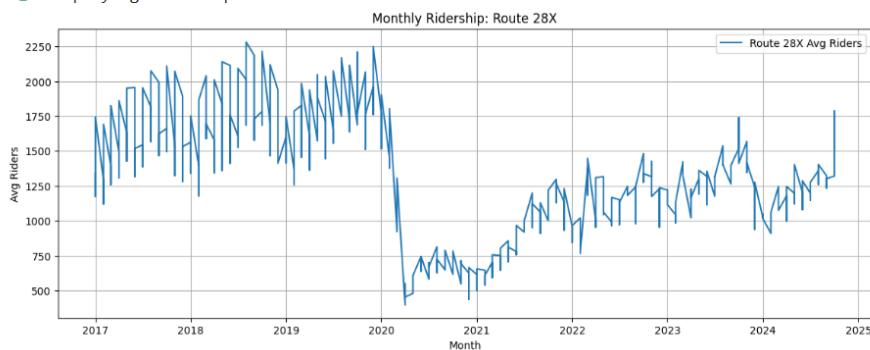
2: Example image:

Result of whole system ridership:



Result of single route ridership:

Ridership columns: ['_id', 'route', 'ridership_route_code', 'route_full_name', 'current_garage', 'mode', 'month_start', 'year_month', 'day_type', 'avg_riders', 'day_count']
Please input the route number (like 61A) or 'All' for the whole system 28X
● displaying ridership for 28X ...



Result of invalid route number:

```
Please input the route number(like 61A) or 'All' for the whole system 100
● displaying ridership for 100 ...
✗ The route 100 is DNE
```

3: The title can display any ridership of your chose route; the line means average riders of specialized route or overall data. X axis means the timeline and the Y axis is the quantity of riders.

4: Function

This system can be used to check riders data for routes and the passengers.

2017-01-01~2019-12-31: Ridership on the route increased steadily, reaching a peak above 2250 riders/month in August 2018.

2020-01-01~2021-04-01: COVID Shock, there's a huge collapse on ridership occurred, formed a lowest record in April 2020, only 15% of pre-COVID time.

Time from 2021-06-01: the ridership has been recovered since the date, but it's hard to reach peak in August 2018

Seasonal patterns: seasonal variation like holiday dips, overseas abroad, back to school time are observable

5: Data source:

Name: monthly_ridership and schedule_trip_counts (in .csv forms)

Variables Used: route, route_full_name, year_month, avg_riders, day_count, etc.

Source: Downloaded from PRT Open Data

Date range: 2017-01-01 to 2024-9-30

6.Methods

I. Data Cleaning

Filter by user-selected route.

Convert year_month to datetime objects.

Handle missing values or irregular points.

II. Visualization

Python (Matplotlib / Pandas) to produce time-series line plots.

Interactive input: user types route number or "All" to trigger display.

Optional panels: summary statistics, percent-change charts, or comparisons.

III. Analysis

Compute monthly averages, trends, and year-over-year comparisons.

Identify structural breaks (e.g., pandemic-induced changes).

7. Significance Statement

There's importance for Understanding public transit ridership patterns on transportation planning, budget allocation, and service optimization.

The project visualized highlighted long-term structural shifts in travel behavior, like COVID-19 pandemic impact and the subsequent recovery. Providing an interactive, route-level examination tool to offer insight into how different transit services are utilized over time and supports data-driven decision-making for both policymakers and the public.

Kaggle link:

<https://www.kaggle.com/code/kurotsukikurosawa/notebookcefba0d817/edit/run/284963468>

Github link:

<https://github.com/Hakuhi/PRT-bus-information>