Project Progress Report 1

Team members:

Giacomo Bizzotto

Rishitha Thatipally

Irfan Shaik

Errolla Vivek

Shraddha Shakya

Sampath Sai Raghav Allaboyina

A ) Business Goals

The Port Authority of New York and New Jersey is looking to enhance its planning and operational efficiency by forecasting passenger volume at the Port Authority Bus Terminal from 2025 to 2030. Given the dynamic nature of commuter behavior, particularly post-COVID, accurate forecasting will help in making informed infrastructure and operational decisions. Below are the specific business goals in detail:

1) Key Predictive Factors: Identify the most influential factors affecting passenger numbers.

2) Passenger Projections: Forecast bus terminal passenger volume from 2025 to 2030 to determine staging facility needs.

3) Carrier-Specific Projections: Project passenger volumes for individual bus carriers.

4) Comparison to 2019: Analyze current usage compared to pre-pandemic levels (2019).

5) Peak Usage Times: Determine the busiest times for the staging facilities (weekly, monthly, and annually).

B) Tools

* R
* Power BI
* Excel
* SQL

C) Algorithms and Models

* Regression Analysis
* Time Series Forecasting

D) Rationale For Choosing These Tools and Models

The chosen tools and models form a robust framework for passenger forecasting. R, with its specialized forecasting packages, is ideal for building predictive models like ARIMA, Prophet, and ETS. These time series models are crucial for capturing trends, seasonality, and the impact of holidays, respectively, leading to more accurate projections. Power BI then transforms these forecasts into interactive dashboards, making it easy to visualize trends, peak usage, and year-over-year comparisons, providing actionable insights. Excel plays a supporting role, enabling efficient data preparation before it's fed into R or Power BI. SQL serves as the backbone for data storage, cleaning, and retrieval, ensuring efficient handling of large datasets before analysis in R and visualization in Power BI.

Beyond forecasting, regression analysis is incorporated to explore the influence of external factors like economic conditions and fuel prices on passenger numbers. This holistic approach, combining time series models with regression analysis, allows for a more comprehensive understanding of passenger trends and more reliable predictions. The combination of R, Power BI, SQL, and Excel provides a powerful and versatile toolkit for data analysis, visualization, and forecasting.