University of New Haven

Pompea College of Business

Course: BANL-6430-03 – Database Management for Business Analytics
Instructor: Dr. Pindaro Demertzoglou

Project Title:

Port Authority Bus Terminal Passenger Prediction

Project Progress Report-1

Prepared By:
Aastha Kale
Ayesha Kabir
Meghana Lakshminarayana Swamy
Rajyalakshmi Nelakurthi
Venkata Sai Phanindra Namburu

Submission Date- February 24, 2025

Project Progress Report:

A. Business Goals:

The primary objective of this project is to provide the Port Authority of New York with accurate forecasts of bus terminal passenger flow from 2025 to 2030. These predictions will support operational planning, particularly for temporary staging facilities, ensuring efficient resource allocation and infrastructure development.

Key goals include:

- Passenger Volume Prediction: Estimating the number of passengers utilizing the bus terminal and temporary facilities.
- Identifying Key Factors: Analyzing weather conditions (temperature, rainfall, extreme events) and their impact on passenger movement.
- Comparing Travel Trends: Evaluating post-COVID travel patterns against pre-pandemic (2019) data to understand behavioral shifts.
- Predicting Peak Usage Periods: Determining the busiest times (weekly, monthly, yearly) to optimize scheduling and resource planning.
- Carrier-Specific Forecasting: Projecting passenger numbers for individual transportation carriers to improve coordination and service delivery.

By leveraging data-driven insights, this project aims to enhance operational efficiency, improve passenger experience, and support long-term infrastructure planning.

B. Tools & Technologies

To ensure accurate analysis and forecasting, we will utilize the following tools:

- **Python** (Pandas, NumPy, Scikit-learn, Matplotlib) – For data cleaning, statistical modeling, machine learning, and visualization.
- **R** – For advanced statistical computations and data exploration.
- **SQL** – For querying and managing historical passenger and weather data stored in databases.
- **Excel** – For preliminary data organization, basic trend analysis, and quick calculations.
- **Power BI** – For interactive dashboards and data visualization, making insights accessible and actionable.

This combination of tools ensures a robust, scalable, and interpretable analytical process.

C. Models & Algorithms

To generate accurate forecasts and meaningful insights, the following models and algorithms will be implemented:

- **Time Series Forecasting (LSTM, MLR):** Predicting passenger trends over time, considering seasonality and external factors.
- **Regression Analysis (Linear, Multiple Regression):** Identifying key factors (e.g., weather) influencing passenger numbers.
- **Classification Models (Decision Trees, Logistic Regression):** Categorizing trends and detecting anomalies in passenger flow.
- **Clustering (K-Means, Hierarchical Clustering):** Segmenting passengers based on travel behavior and usage patterns.

D. Rationale for Selection

- **Python & R:** Provide extensive libraries for data processing, predictive modeling, and statistical analysis.
- **SQL:** Enables efficient data retrieval from large historical datasets.
- **Power BI:** Converts complex data into intuitive visualizations for better decision-making.
- **Time Series Models:** Essential for forecasting future trends based on historical data.
- **Regression & Classification Models:** Identify influential factors and classify trends, enhancing predictive accuracy.
- **Clustering:** Helps in understanding passenger behavior and optimizing resource allocation.

By integrating these tools and methodologies, the Port Authority will gain actionable insights, enabling strategic planning and improved service delivery for the bus terminal's future operations.

Conclusion

This project will provide the Port Authority of New York with data-driven insights to optimize bus terminal operations from 2025 to 2030. By leveraging advanced forecasting models and analytical tools, we aim to enhance resource planning, predict passenger trends, and improve service efficiency. The combination of time series forecasting, regression analysis, and classification models will ensure accurate predictions, while visualization tools like Power BI will make the insights actionable.

With a strategic, data-focused approach, this project will support informed decision-making, helping the Port Authority improve passenger experience, streamline operations, and effectively manage future infrastructure needs.