The University of New Haven - Port Authority of New York and NJ Data Analytics Project

University of New Haven

Project Requirements

1. **Determine when the weekday departures will be past 125,000 passenger departures and 3,900 bus departures**. Setup your forecasting visuals by year and month. Try to determine the factors that lead to those spikes. For example, a holiday, a taxi strike, a UN week or anything else.

2. **Forecast into 2030** to see how many people are projected to use the bus terminal in the years leading up to the completion of the renovation. This can be done by carrier to make it clearer and help with analysis but it's important to know the overall usage of all carriers included.

3. Develop three forecasting models, train them, test them to see which one works best. Include in your final project submission the model/tool that Port Authority can use with future data to make accurate predictions. Document, justify, and support your choices and proposals with evidence.

Factors you must include in your model:

Population. Explore if the amount of people using the busses followed the natural growth in the population of the surrounding area/the areas that these buses go to.

Rain and Snow

For the rain and snow the big question would be if we see a decrease/increase in users through the terminal when its raining/snowing, cold weather is unavoidable in the winter/fall months but there are also events that make it so that there may be more usage then anyway.

Temperatures
While I would still be interested in seeing if there's any correlation with temperatures, we often wonder at the agency how much the weather impacts the usage of our facilities so more detailed analysis in that vein would be helpful.

Notes:
In the data, departures happen during the weekdays only, 6AM-10 PM. You do have the data for bus departures and passenger departures.

You may notice in reviewing the data that one bus carrier drops off in 2021/2022 and that is correct, they ran out of business after covid and got replaced, if it makes it easier to analyze the carriers can be grouped in whatever way you see fit although internally, we grouped it based off of (inner city, commuter and NJ transit). Try to take into consideration past snowstorms, UN events, strikes, gas prices etc.