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

Project 2: The Port Authority Tunnels and Bridges

**PROJECT POINT-OF-CONTACT**
*Armando Guzman, Program Manager, arguzman@panynj.gov*

## PROJECT DESCRIPTION Fall 2025

Here at The Port Authority, we pride ourselves in trying to keep the region moving through our facilities. Over the past few years and especially after Covid we have been implementing efforts to help detail overall health and areas for improvement within our facilities. Some efforts take priority and with that we may not always be able to do robust analysis on all our data since keeping everything in good repair takes priority and that's where we could use some help.

For our facilities at the Holland Tunnel, George Washington Bridge, Lincoln Tunnel, and the rest of our outer bridges, we want to know what our busiest times are throughout the years and how our traffic and speeds are affected by things like seasonality, weather, and holidays. We are concerned about the traffic and patterns that may arise in the data that can help us figure out what affects our facilities most and if there's anything we can do to help mitigate any negative effects.

**Project 2 Goals**: **The Port Authority Traffic on Bridges and Tunnels**
1. What are the top five factors that affect the usage of bridges and terminals by drivers?
2. How many toll violators are there? Provide the numbers by time interval such as year, month, week. Also, provide the results by facility such as tunnel, bridge etc.

3. What are the busiest times throughout the year and how the traffic and speeds are affected by factors like seasonality, toll violators, type of car, holidays, and other NYC related events? Report by time interval such as year, month, week.
4. Report on congestion due to pricing in 2025 with respect to previous years. Report by facility. For example, do you see a shift in traffic patterns so that drivers try to avoid tolls and congestion in other facilities are happening?
5. Forecast the usage of facilities by year beyond 2025 to the max year you can. Use AutoML to develop a prediction model. Use the best model suggested by the AutoML Azure platform.
6. The model chosen in the previous step by AutoML, you need to develop it using Python code with Jupyter Notebooks and the Anaconda platform.
7. As a group, create a GitHub portfolio site and name it: "Capstone Project 6900_01_Group X Fall 2025". All group members should have access to this folder. In this Github site add:

   a. The final recommendations document for the project detailed in the Final Project Report.
   b. The Power BI analytics dashboard you have created to answer all the analytics questions in this project. You can add more visualizations at will. The Power BI dashboard must have the data embedded without any links to outside Excel or text files.
   c. A word document named "AutoML model Resources". Here you will include all the characteristics and metrics of the AutoML model.
   d. A word document named "Python Model Resources" Here you will include all the code and other resources you have used to develop the Python Model. Make sure you do not just include code but also a description of the Python model used.
   e. The data file you have used to develop the models in steps c and d.