# Port Authority Project: Traffic on Bridges and Tunnels

## Executive Summary

The Port Authority of New York and New Jersey aims to understand and improve traffic management across critical infrastructure such as the Holland Tunnel, George Washington Bridge, Lincoln Tunnel, and other outer bridges. Given the operational focus on maintaining facilities, robust analysis of available data is often secondary, but vital. This project aims to analyze traffic patterns, busiest periods, and the impact of seasonality, weather, and holidays on traffic flow and speeds.

### Goals of the Project:

1. Identify factors influencing the usage of bridges and tunnels, including weather, seasonality, holidays, and other events.  
2. Assess the volume and behavior of toll violators (individuals who evade toll payments).  
3. Determine peak traffic times across the year and assess the impact of traffic patterns on speeds and operations.

### Findings:

Factors Affecting Traffic Usage:

Weather: Precipitation and wind speed significantly influence traffic flow.  
Holidays and Events: Traffic peaks during certain holiday periods and NYC events.  
Seasonality: Summer months typically see increased traffic volumes.

Toll Violators:

Analysis shows a measurable percentage of toll violators. Further investigation is required to identify patterns and improve enforcement measures.

Busiest Times:

Traffic peaks were observed during weekday rush hours and holiday seasons. Specific months such as July and December show significantly higher activity.

### Model and Dashboard:

A Time Series Forecasting model was employed using Azure AutoML to predict traffic volumes and busiest times. The dashboard created in Power BI provides visual insights into:  
- Traffic trends across months, weeks, and hours.  
- Impact of weather and seasonality.  
- Toll violator trends and peak periods.  
- Holiday and event-related traffic patterns.  
This analysis allows the Port Authority to optimize resource allocation, enhance enforcement, and mitigate negative effects such as congestion and delays.

## Main Section

### Data Integration and Preparation

The analysis was conducted using historical data from traffic, weather, and operational logs. Key steps included:

1. Data Cleaning: Missing values for traffic speeds, toll payments, and weather conditions were filled using interpolation and averages.  
2. Feature Engineering: Derived fields for seasonality (Month, Day of Week, Holiday) and weather conditions (Temperature, Precipitation, Wind Speed).  
3. Data Consolidation: Combined datasets from multiple sources into a unified dataset to analyze relationships.

### Modeling and Tools

Model:  
A Time Series Forecasting model using Azure AutoML to predict traffic patterns.  
Model captured seasonality, trends, and external factors like holidays and weather.

Tools:  
Power BI: Visualization and dashboarding.  
Python: Data preprocessing and feature engineering.  
Excel: Initial data cleaning and merging.

Rationale: The Time Series Forecasting model was chosen for its ability to capture temporal patterns and provide accurate predictions, aligning with the Port Authority's need for operational insights.

### Visuals and Insights

Traffic Trends Heatmap:  
Shows hourly and monthly traffic volumes.  
Highlights peak hours (e.g., weekday rush hours) and peak months (e.g., July, December).

Weather Impact Analysis:  
Scatter plots showing relationships between weather (precipitation, temperature) and traffic volume.

Toll Violator Trends:  
Bar charts illustrating monthly toll violation rates.  
Provides insights into enforcement effectiveness and busiest months for violators.

Seasonality and Holidays:  
Line charts showing holiday-related traffic peaks.  
Annotations for major NYC events.

Speed Analysis:  
Line charts correlating traffic volume with average speeds to highlight congestion periods.

### Recommendations

1. Resource Optimization:  
Allocate resources during identified peak periods (e.g., holidays, weekday rush hours).  
2. Enforcement Enhancements:  
Improve toll collection measures during high-violation periods.  
3. Traffic Mitigation:  
Develop strategies to reduce congestion during severe weather and high-traffic events.

## Appendix

Screenshots of the Power BI dashboard.  
DAX formulas and Python scripts used for the analysis.  
Key datasets:  
1. Final consolidated dataset.  
2. Forecasted data.