

# **Predicting Urban Traffic Flow Using Multivariate Data Analysis**

## **Introduction**

Given the quick development and influx of residents and workers into NYC during the past decades, it is crucial to continue with a sustainable urban landscape, and an efficient management of traffic flow congestion in order to minimize carbon emissions, and improve overall city livability.

Traditional traffic management systems often rely on historical data and static models, which do not adequately address the dynamic nature of urban traffic. This project is aimed at leveraging data analysis techniques to predict urban traffic flow.

## **Problem Statement**

Urban traffic is influenced by factors such as the time of day, weather conditions, urban events, and infrastructure changes.

These variables makes traffic prediction challenging despite its crucial role in urban planning and management.

The question of this project is: *How can we accurately predict urban traffic flow in NYC by analyzing multivariate data?*

## **Importance of the Problem**

Efficient traffic flow in NYC affects everything from emergency response times to the daily commute of over a million people.

Inadequate traffic management results in pollution, wasted time, and heightened stress levels among not only the residents of NYC but also all of the surrounding neighbourhoods and cities.

## **Data Set and Analysis Techniques**

I will cross reference and use several comprehensive data sets found in [NYC Open Data](#). This data will include traffic volume and speed data from various sensors across the city as well as weather conditions, construction, and other infrastructure changes.

This data set will offer a detailed insight to urban traffic patterns.

To analyze this data, we will use techniques such as

- Linear Regression: To understand the relationship between traffic flow and various predictors.
- Machine Learning Models: Such as decision trees and neural networks, to predict traffic flow based on the engineered features.

## **Conclusion**

This proposal outlines an approach to tackle an important urban, social and environment issue through multivariate data analysis.