**EXECUTIVE SUMMARY**

**PROBLEM STATEMENT**

Traffic congestion in the US has gotten worse over the past year. According to Inrix, the average driver in the US spent time 29% more than in 2021, equivalent to $134 more spent on fuel in a year. This problem is expected as US population has steadily increased by 300 million people every year. By modelling the population movements, the pattern can be analyzed to determine the needs for significant changes that require attention by the town ship planner. Modelling and forecasting were done for Cook County, IL, where Chicago as the second most congested city in the world is located.

**DATA**

Data was obtained from Kaggle.com, and Bureau of Transportation Statistics by the Maryland Transportation Institute and Center for Advanced Transportation Technology Laboratory at the University of Maryland.

**METRICS**

The dataset used multiple combination differencing method (time shift, log, sqrt, cbrt). The model and forecasting method used are ARIMA and SARIMAX model, train & testing method, and the accuracy is measured by RMSE.

**STATISTICAL ANALYSIS**

Number of Trips was chosen as the main dataset because I want to capture the whole population movements, regardless of the distance. The data is not stationary, hence 8 differencing methods was performed & found time shift period of 1 gives the best p-value in ADF test. For loop was used to find the best combination of p,d and q parameters.   
After performing ARIMA model, the error of prediction on test data is extremely high.

I continued with SARIMAX model, which also shows extremely high error.

I reevaluated the data by combining log and time shift, that gives the second lowest p-value and all the RMSEs in modelling and forecasting shows around 0.06.

**CONCERNS**

There are few concerns: 1) I observed the data has 7-day cycle, but unable to use 7 in seasonal order in SARIMAX. 2) Does the data need to be exponentiated back to original value? 3) How do we come out with forecasted numbers?

**PROPOSED SOLUTION**

Necessary development such as increasing public transport coverage and frequencies, adding more highways or smart traffic lights shall be evaluated for future needs.