# Analysis Report

Ashwin Goyal, University of Maryland College Park

## Introduction

As a finalist of The Data Incubator Challenge, I have prepared a brief report showcasing my project with a few analyses. The aim of the project, proposed at the Data Incubator, is to analyze the traffic stops data available for various states of the US to find a pattern which might help in decreasing the number of traffic stops. The motivation for this project arose from the first crude analyses of the traffic stops data set. It was observed that the number of traffic stops have not decreased over the past 2 decades. On the contrary, they have increased on an average over the past 2 decades. This observation held true both state-wise (Figure 1), where the maximum increase in the number of traffic stops occurred in Michigan, and violation-wise (Figure 2), where the number of speeding violations have increased the most.

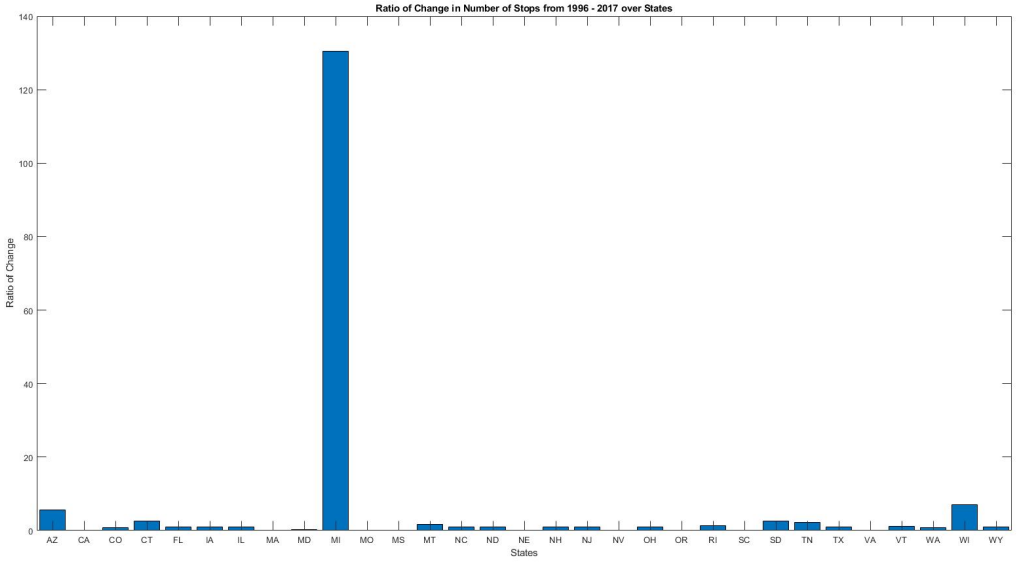


**Figure 1:** Change in Number of Stops over past 2 decades (State-wise)

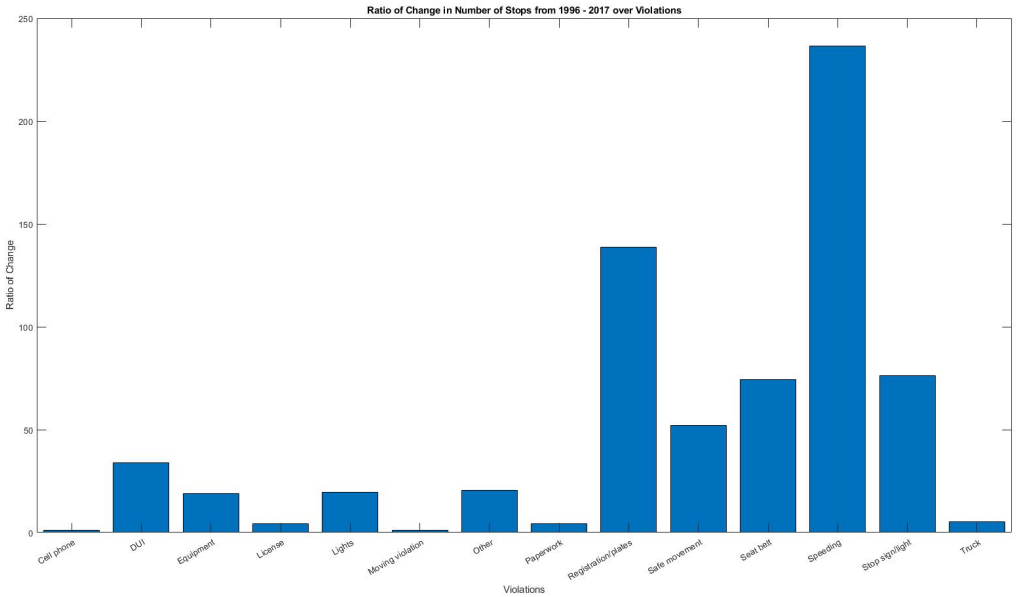


**Figure 2:** Change in Number of Stops over past 2 decades (Violation-wise)

## Methodology

This project is not a trivial one. It requires data other than the number of traffic stops in order to make sound analyses. For example, the increase in the number of cars on the road over the years is the missing factor which will justify the ratio of change in the number of traffic stops over the past two decades.

A few more analyses were done on the data. One of them was to observe the trend of the traffic stops throughout the day for different days of the week (Figure 3). It can be clearly seen that the trend remains almost similar for all days with a few deviations. On Sunday and Monday, the trend deviates from 0 to 5 hours, and on Friday and Saturday, the trend deviates from 18 to 24 hours. Now the latter observation makes sense as these are the weekend nights. But the former observation does not. It was also observed that minima occurs around 5 hours, and maxima occurs around 10 hours and 15 hours. This observation is trivial as the more number of cars on the road implies more number of traffic violations and the peaks occur around the work hours of the day, that is, the peak hours for the traffic on the road.

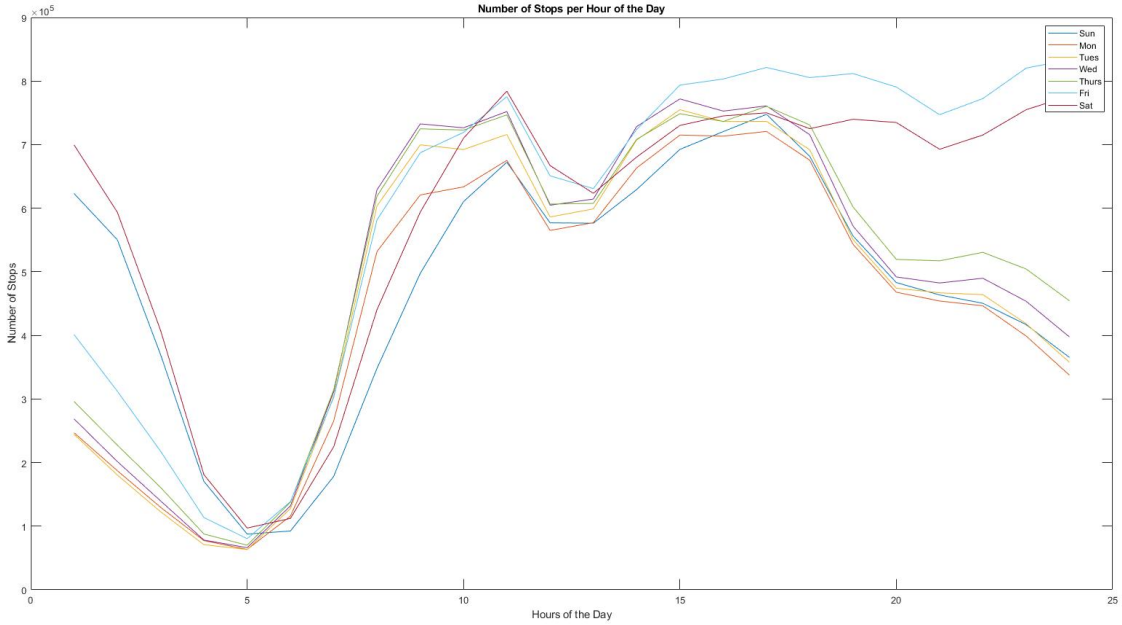


**Figure 3:** Trend of the violations throughout the Day (Day-wise)

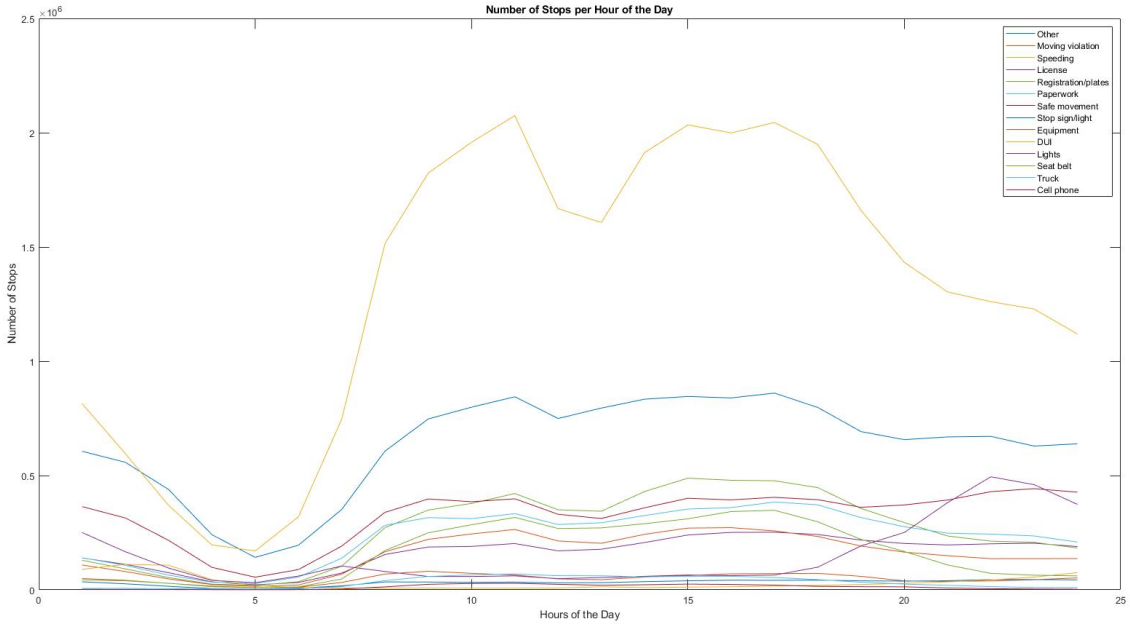


**Figure 4:** Trend of the violations throughout the Day (Violation-wise)

Another analyses was to observe the trend of the traffic stops throughout the day for different violations (Figure 4). It was observed that the trend remains almost similar for any violations with the only major difference being observed in the magnitude. This is not a trivial observation. It is generally thought of that the trend for, say, speeding and DUI, with be different from the trend for registration/plates. But the data is still inconclusive.

These analyses were further divided into sub-categories, that is, the day-wise analyses was done for individual violations and the violation-wise analyses was done for different days. The major differences have been identified below.

1. Trend of the violations for DUI (Figure 13) was different than the general trend. As one can predict, most DUI violations occur around 3 AM that too on Saturday and Sunday. Also, the minimum DUI violation occur during the day (office hours).

2. Similarly, the trend for Lights violation (Figure 14) differ from the general trend. The reason for the same is that lights are only used from sunset to sunrise in normal weather conditions. So, the traffic stops for the same can mostly happen during that time only. It should be noted that most violations occur around 9 PM on Friday and Saturday.

3. The trend of violation on Friday (Figure 24) and Saturday (Figure 25) differ from the general trend as well, especially for Speeding violation. In general trend, the number of speeding violation reduce after 8 PM, but that is not the case here. In fact, on Friday it becomes a constant.

The analyses plots are shown below and are also available on GitHub (https://github.com/Ghost1995/Data-Incubator.git).

These analyses show that there definitely is a trend in the number of traffic stops. If these patterns can be extracted from the analyses, then it might help to reduce the number of traffic violations per year. Next, I am planning to conduct analyses for number of traffic stops (age-wise). This way I am exploring individual aspects of the data (one-by-one) and later an N-dimensional predictor can be developed which takes the relevant data as input and predicts the likelihood of occurrence of various traffic violation on any particular day of the year in a particular location.

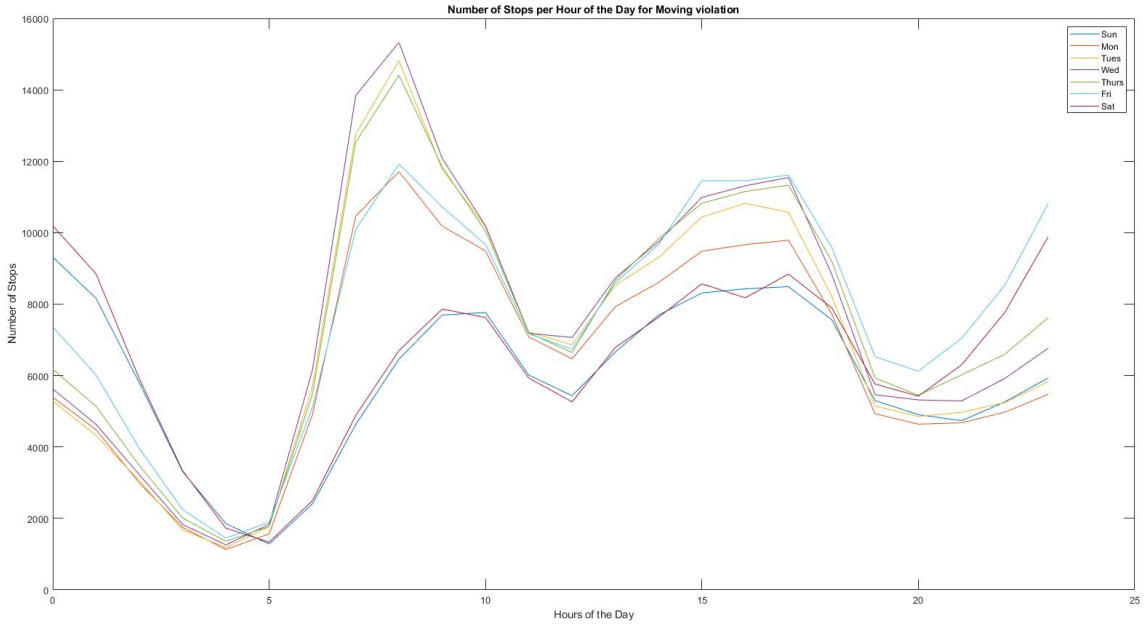


**Figure 5:** Trend of the violations throughout the Day for Moving Violation

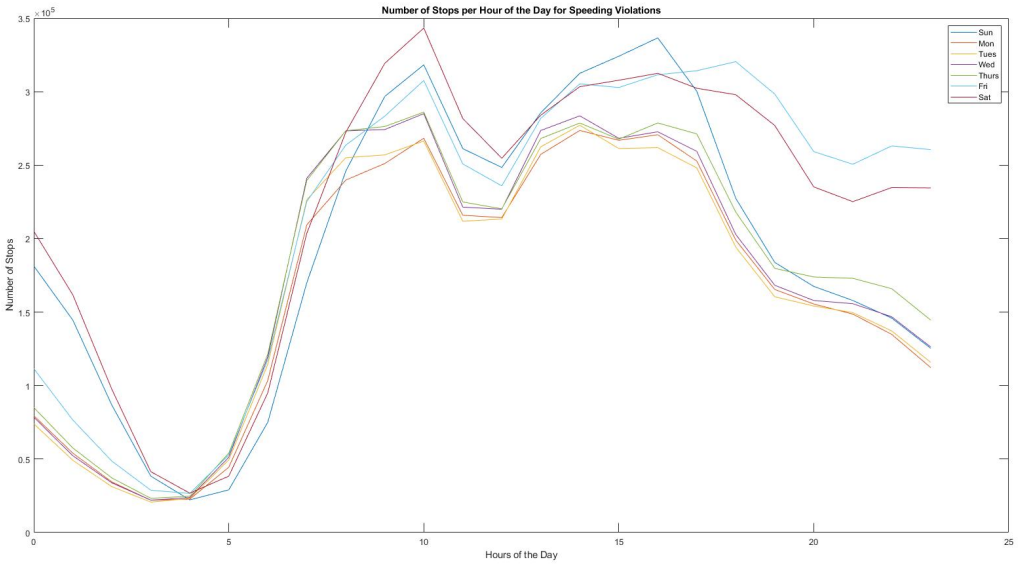


**Figure 6:** Trend of the violations throughout the Day for Speeding Violation

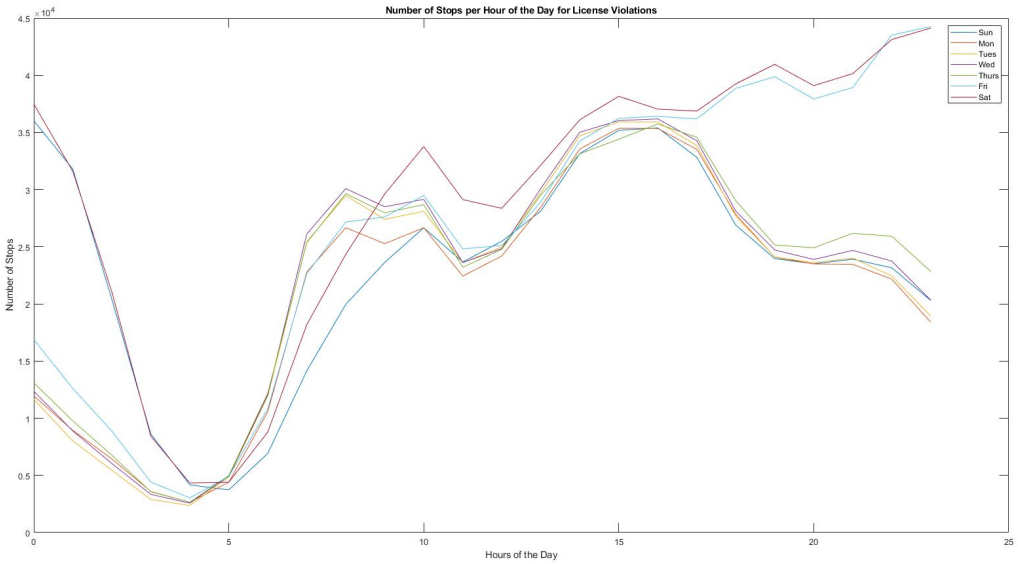


**Figure 7:** Trend of the violations throughout the Day for License Violation

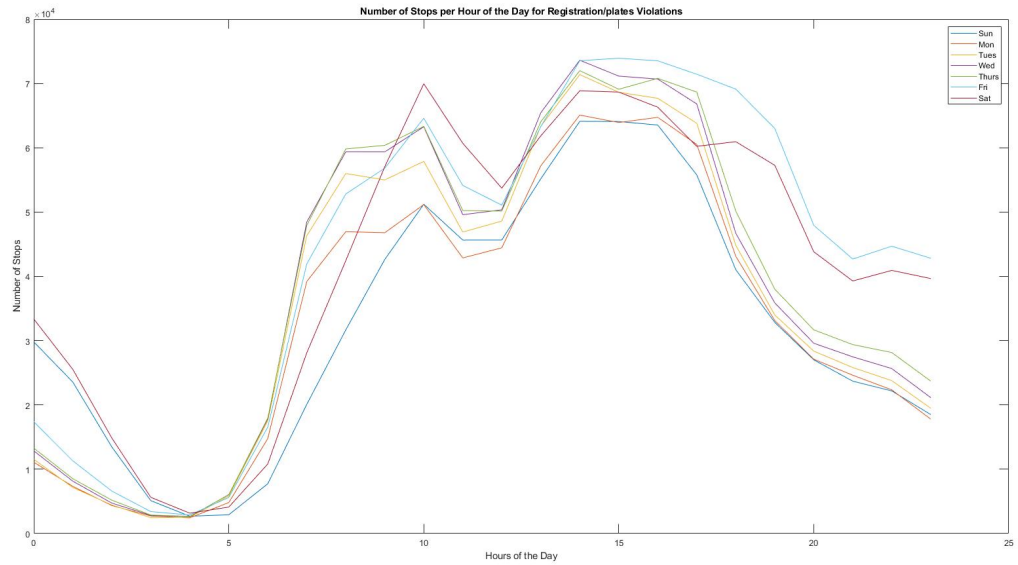


**Figure 8:** Trend of the violations throughout the Day for Registration/Plates Violation

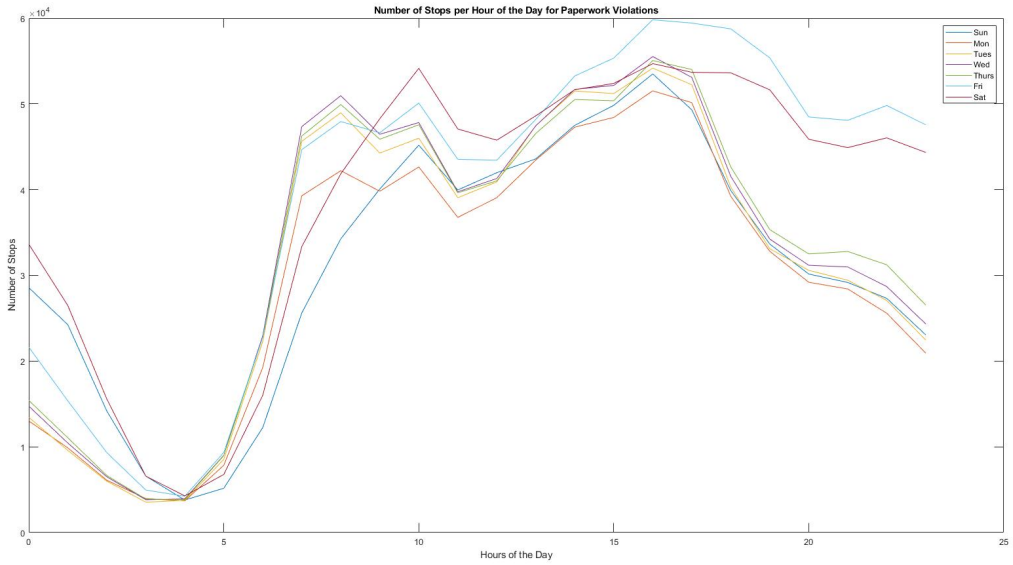


**Figure 9:** Trend of the violations throughout the Day for Paperwork Violation

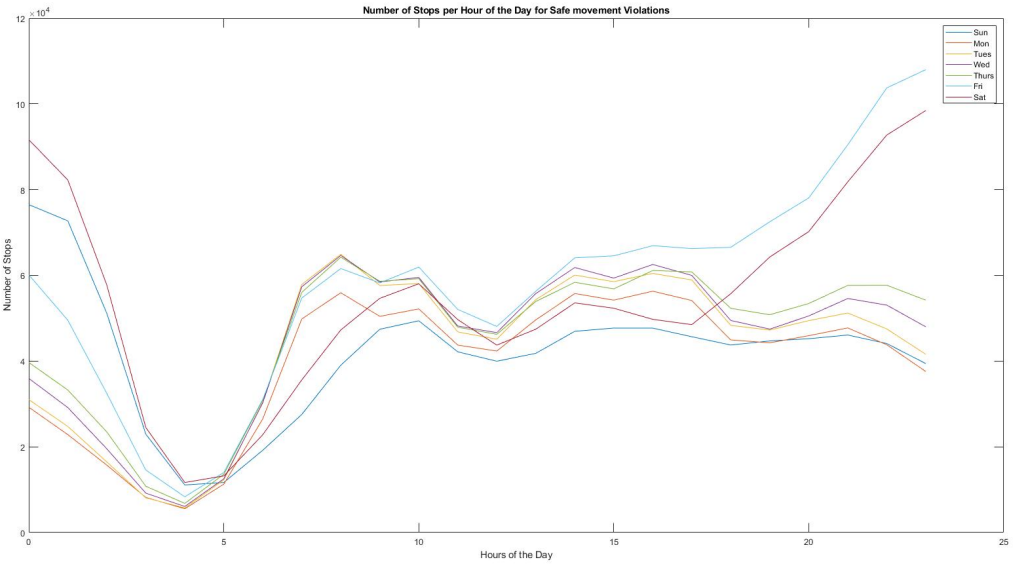


**Figure 10:** Trend of the violations throughout the Day for Safe Movement Violation

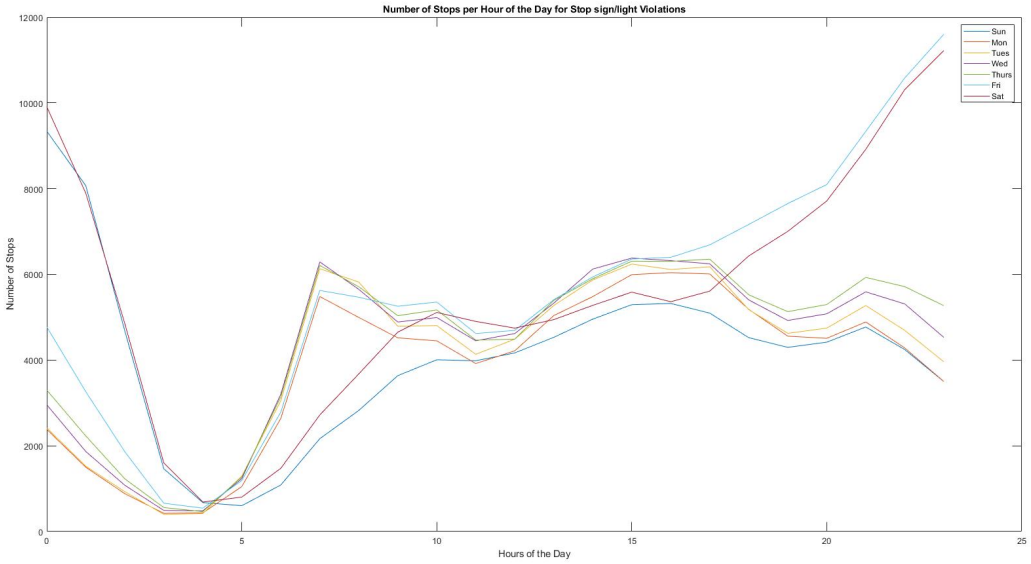


**Figure 11:** Trend of the violations throughout the Day for Stop Sign/Light Violation

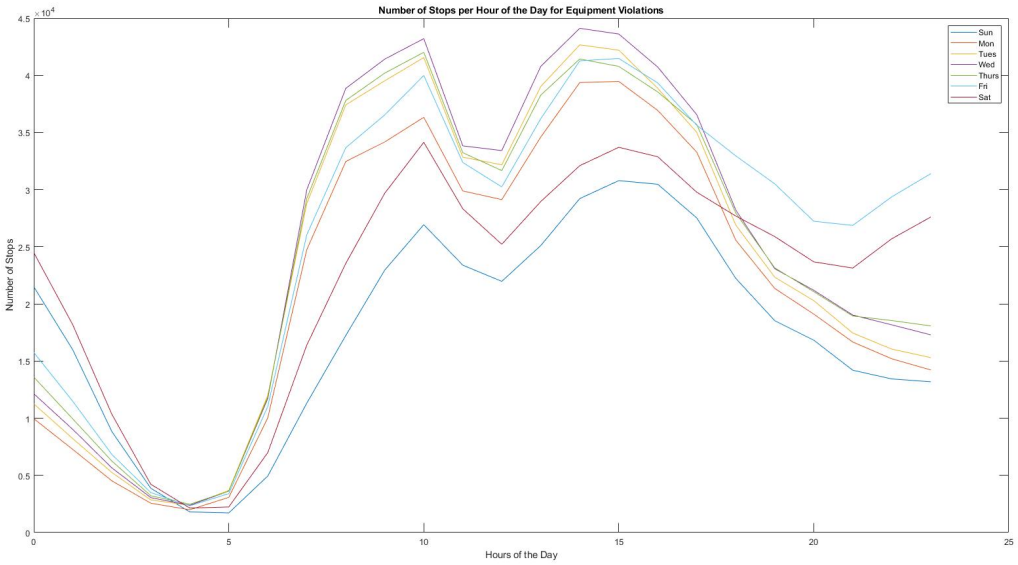


**Figure 12:** Trend of the violations throughout the Day for Equipment Violation

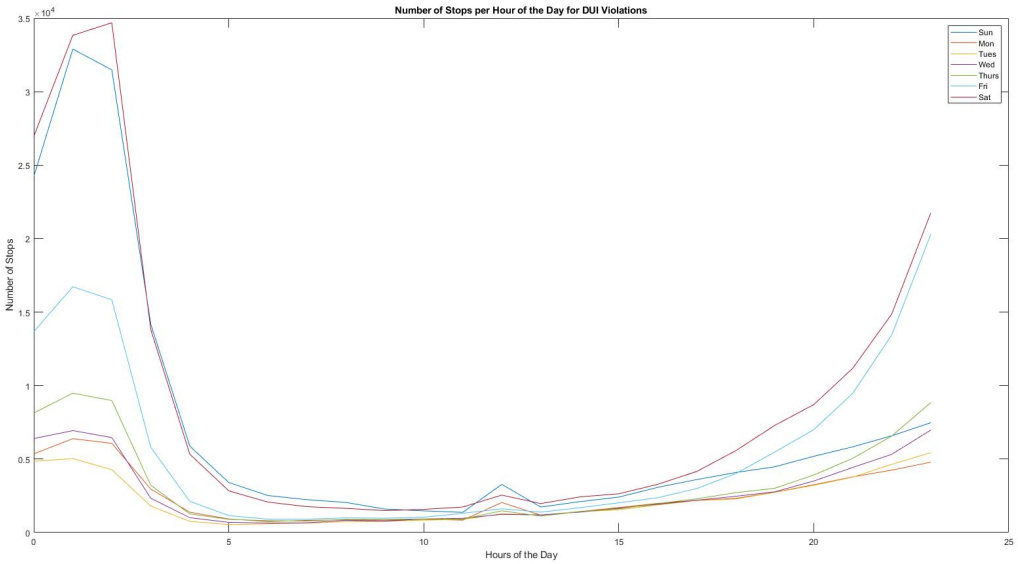


**Figure 13:** Trend of the violations throughout the Day for DUI Violation

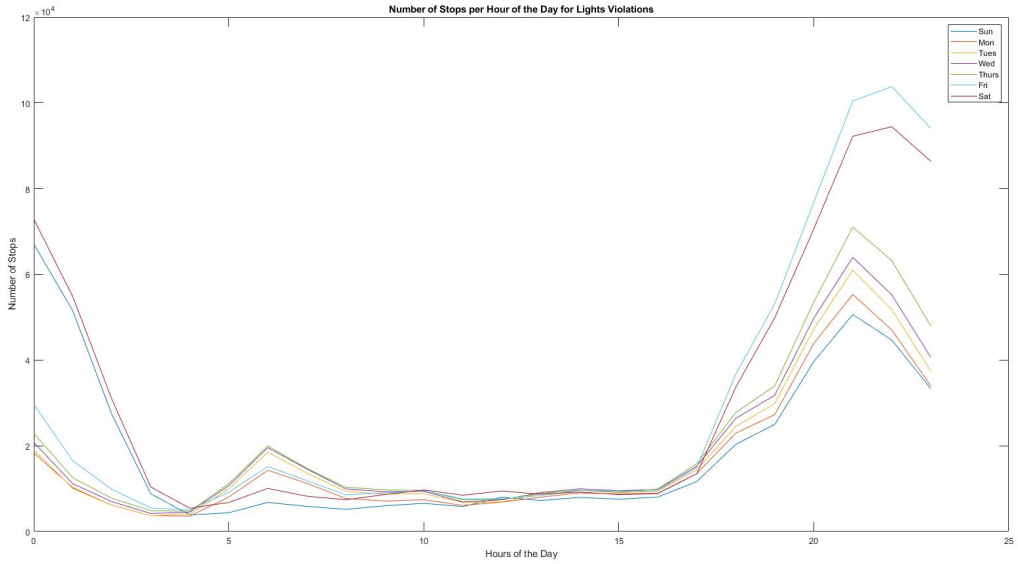


**Figure 14:** Trend of the violations throughout the Day for Lights Violation

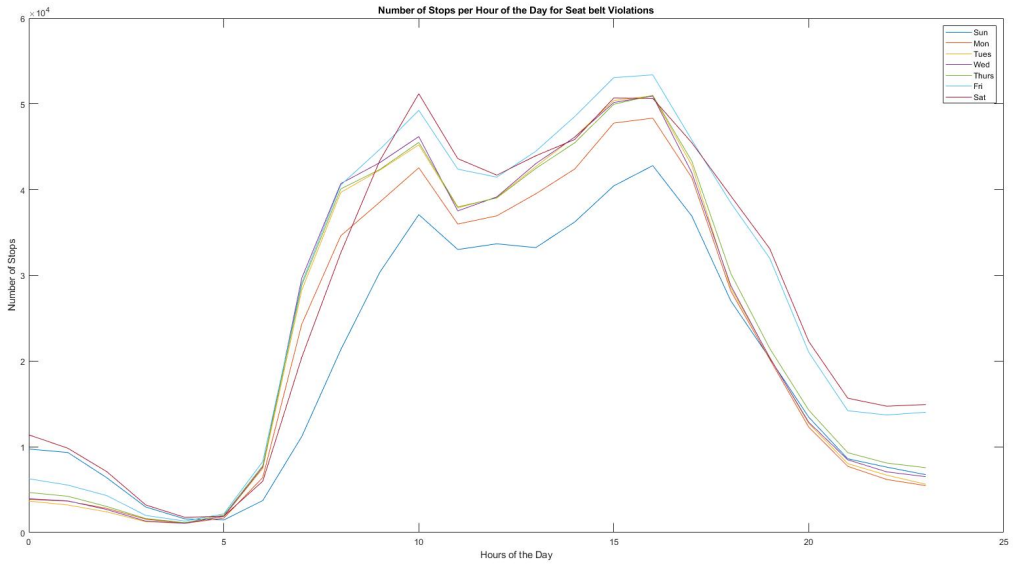


**Figure 15:** Trend of the violations throughout the Day for Seat Belt Violation

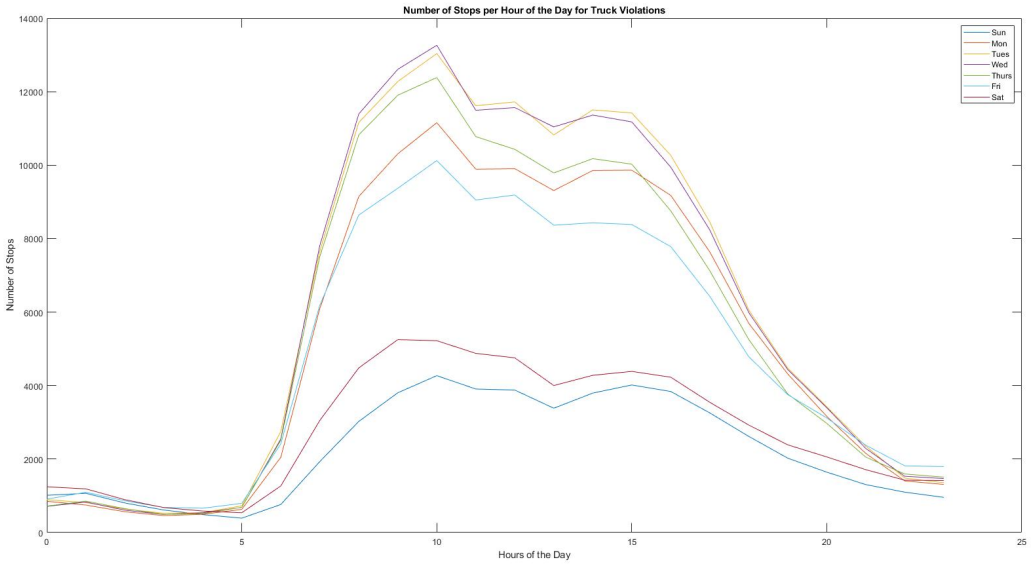


**Figure 16:** Trend of the violations throughout the Day for Truck Violation

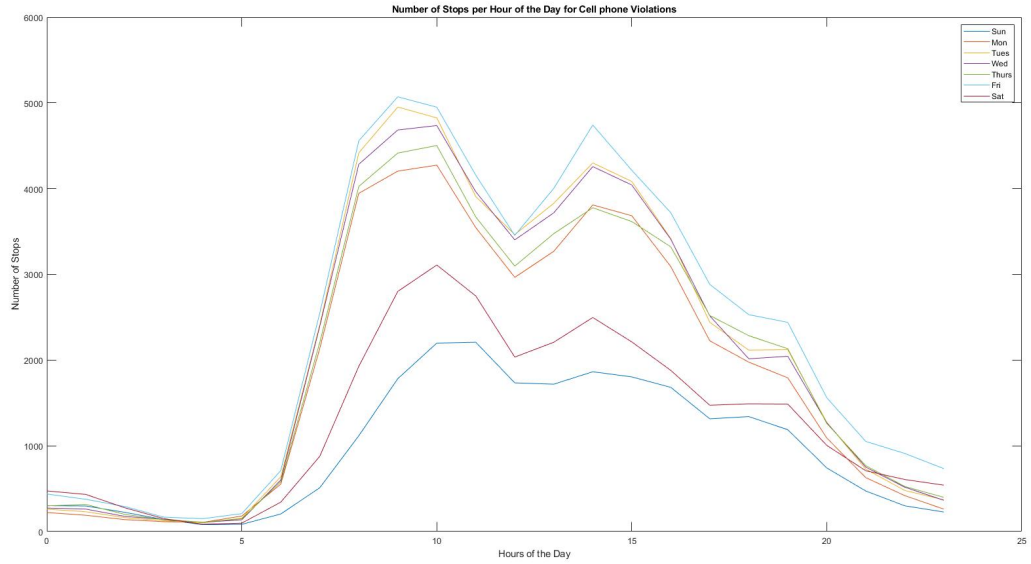


**Figure 17:** Trend of the violations throughout the Day for Cell Phone Violation

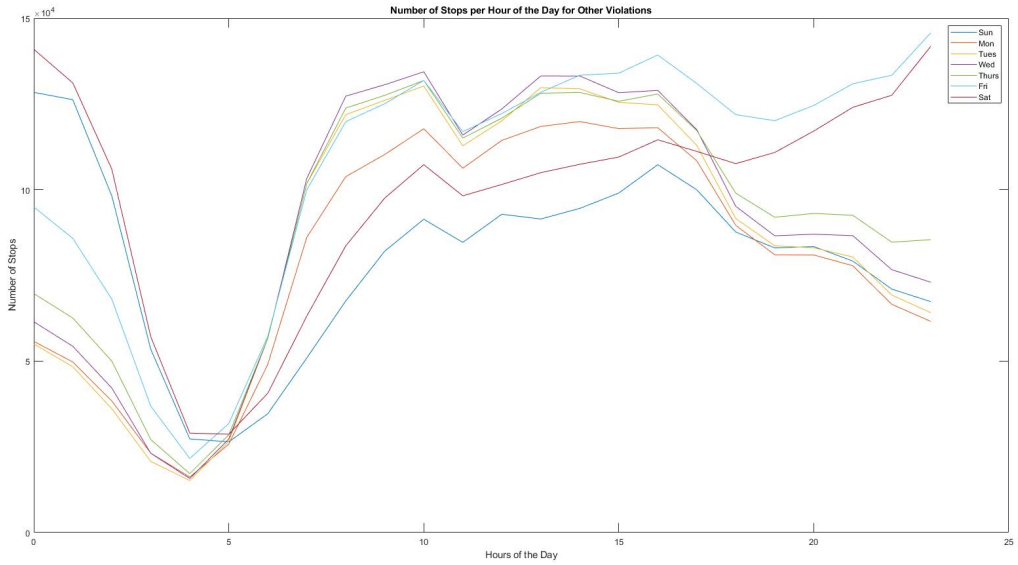


**Figure 18:** Trend of the violations throughout the Day for Other Violations

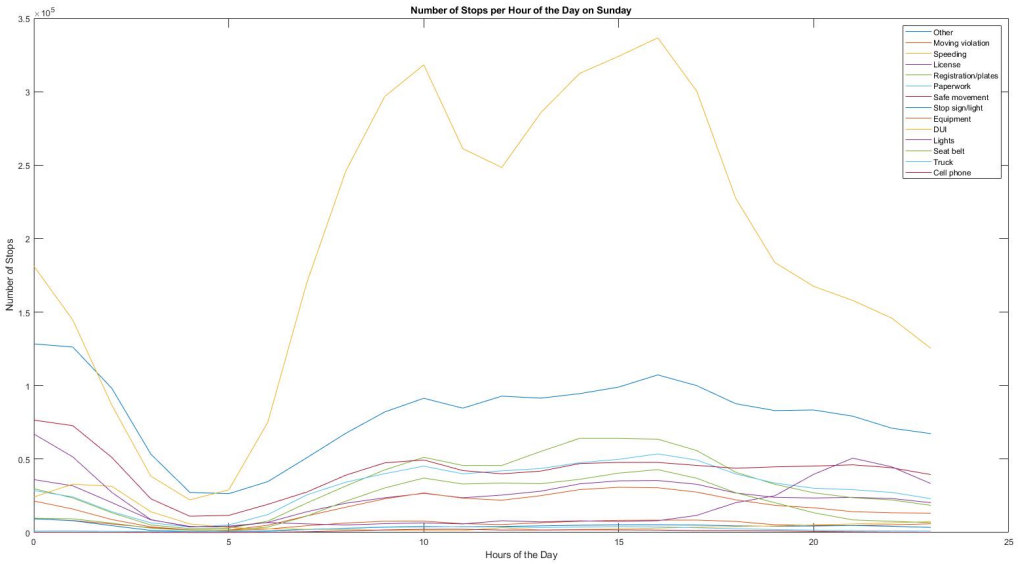


**Figure 19:** Trend of the violations throughout the Day on Sunday

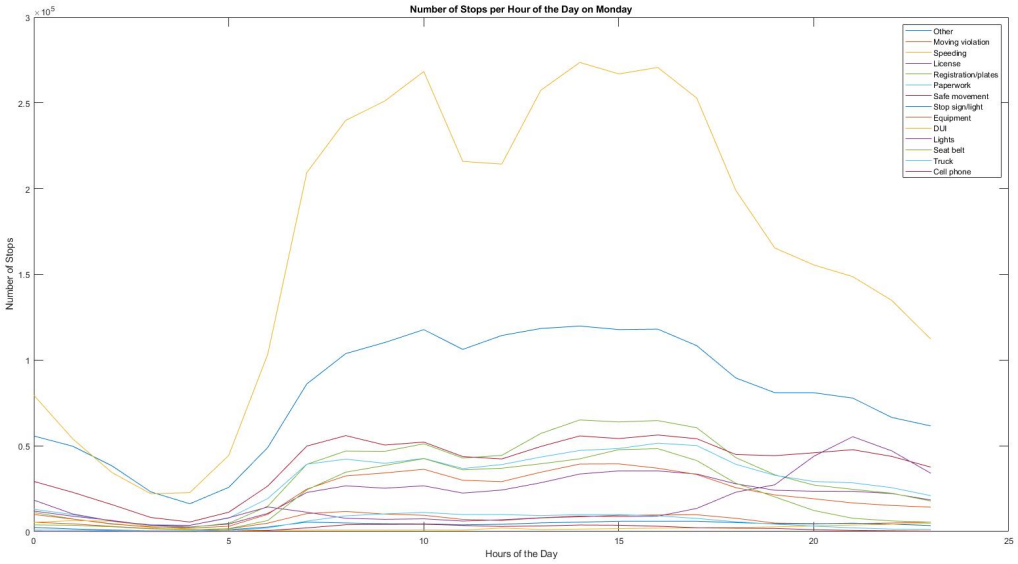


**Figure 20:** Trend of the violations throughout the Day on Monday

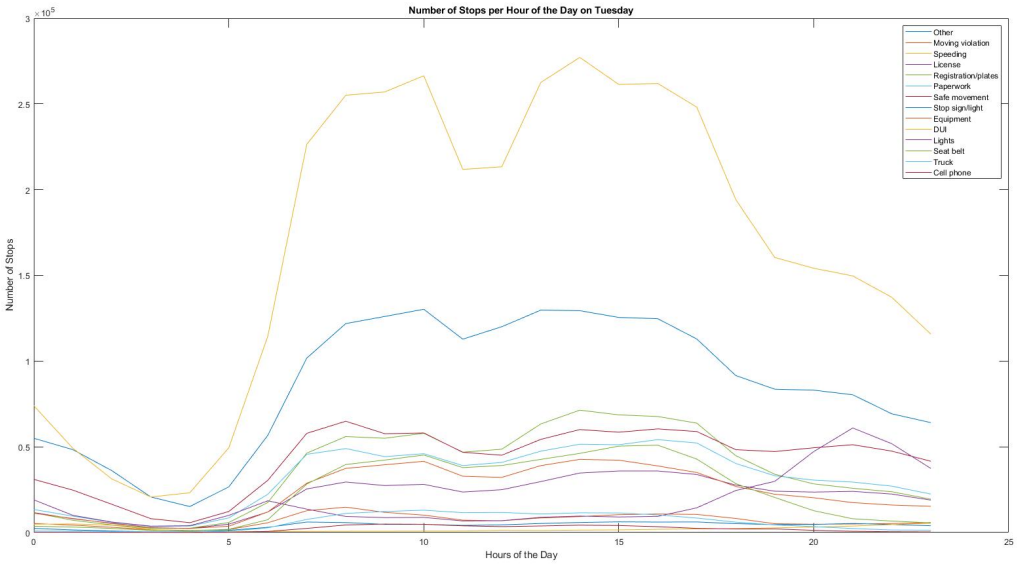


**Figure 21:** Trend of the violations throughout the Day on Tuesday

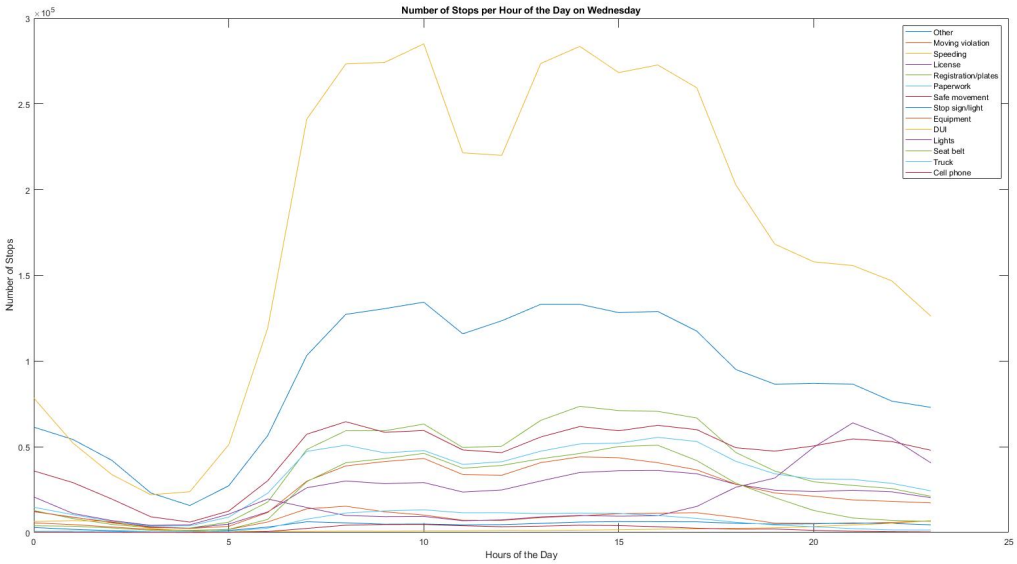


**Figure 22:** Trend of the violations throughout the Day on Wednesday

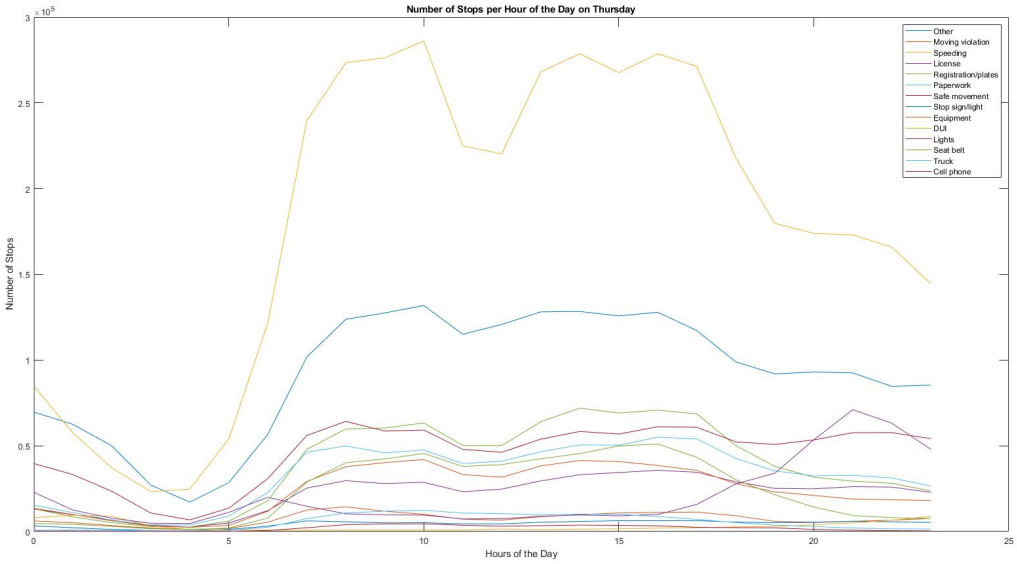


**Figure 23:** Trend of the violations throughout the Day on Thursday

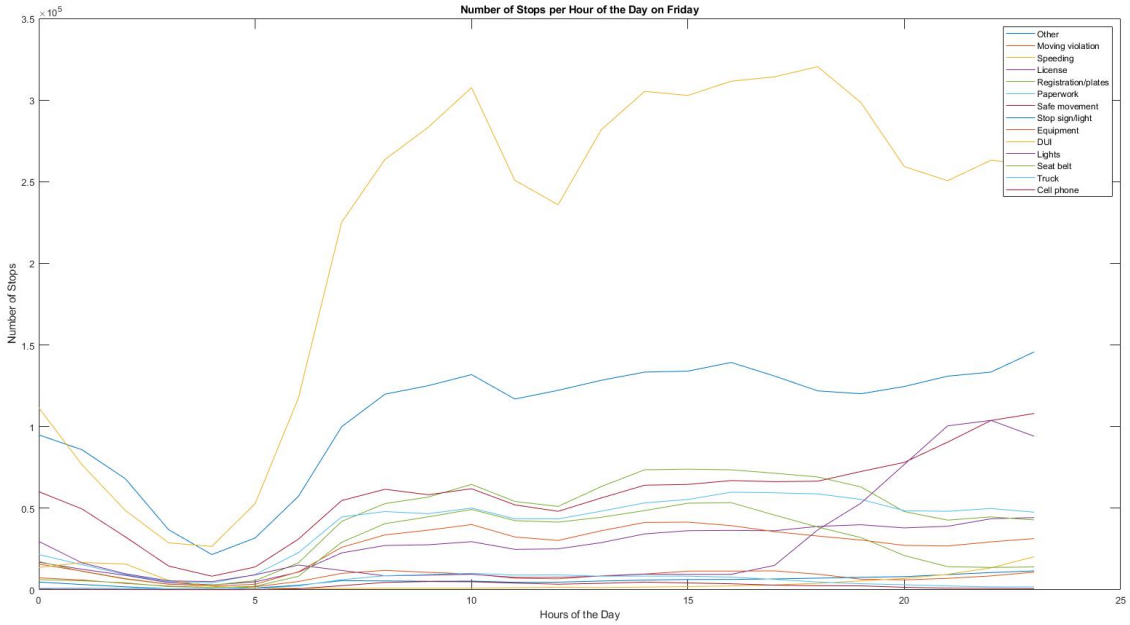


**Figure 24:** Trend of the violations throughout the Day on Friday

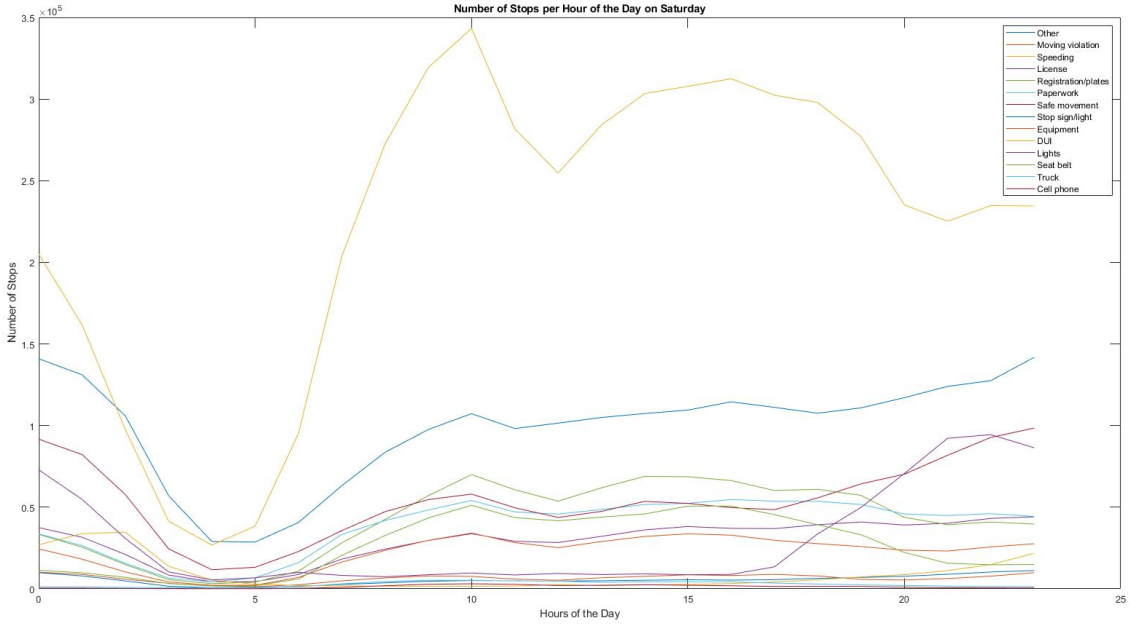


**Figure 25:** Trend of the violations throughout the Day on Saturday