

Maintenance Schedule for Turnstiles

Ali Albayat





Introduction

Backstory

- In October 2019, maintenance will be made on NYC subway turnstiles.
- Working when the station has minimum traffic.
- More convenient for the workers.
- Avoid people injuries.

BUT

What is the best time to do the maintenance ?



Data Sets and Tools



Data sets

Oct 2016

Oct 2017

Oct 2018

18 weeks

3,464,706 rows

11 columns



Tools

SQL

Pandas

Matplotlib

Seaborn

NumPy

Check

Check for NaN.

Drop

Drop duplicates.

Exclude

Exclude unwanted rows and columns.
(LINENAME, DIVISION, DESC).

Adjust

Adjust the time to be exactly every 4 hours.

Drop

Drop duplicates.

Find

Find the entries and the exits every 4 hours.

Check

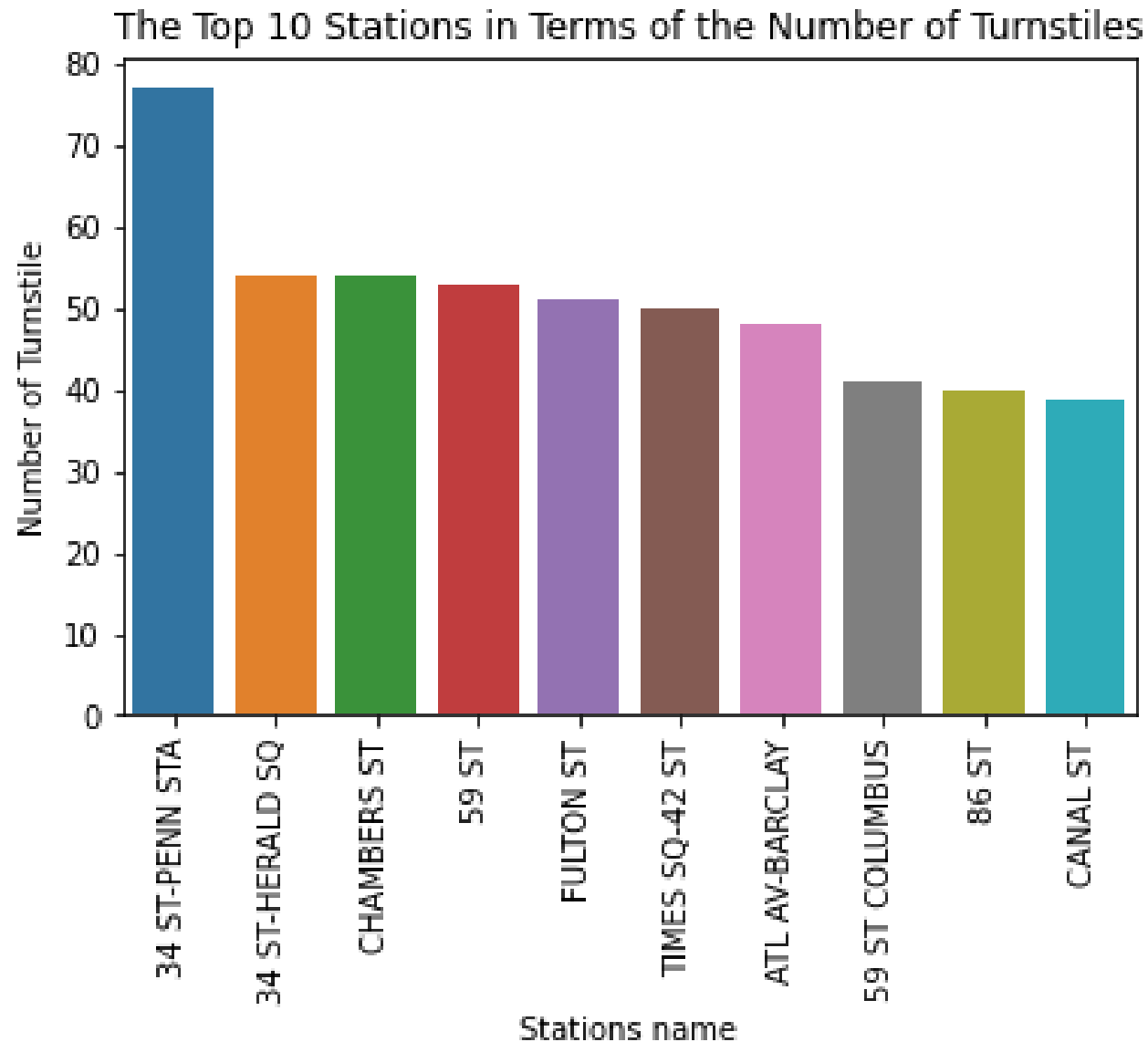
Check outliers.

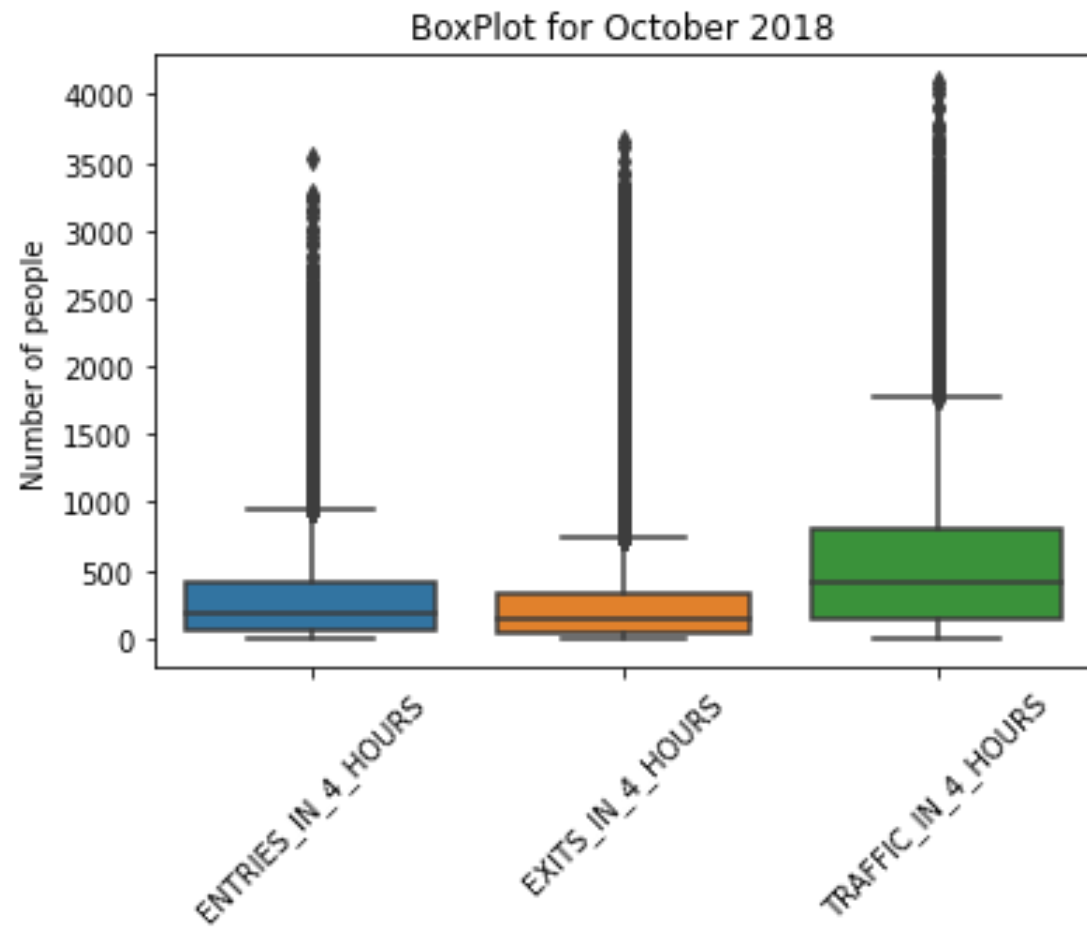
Data cleaning



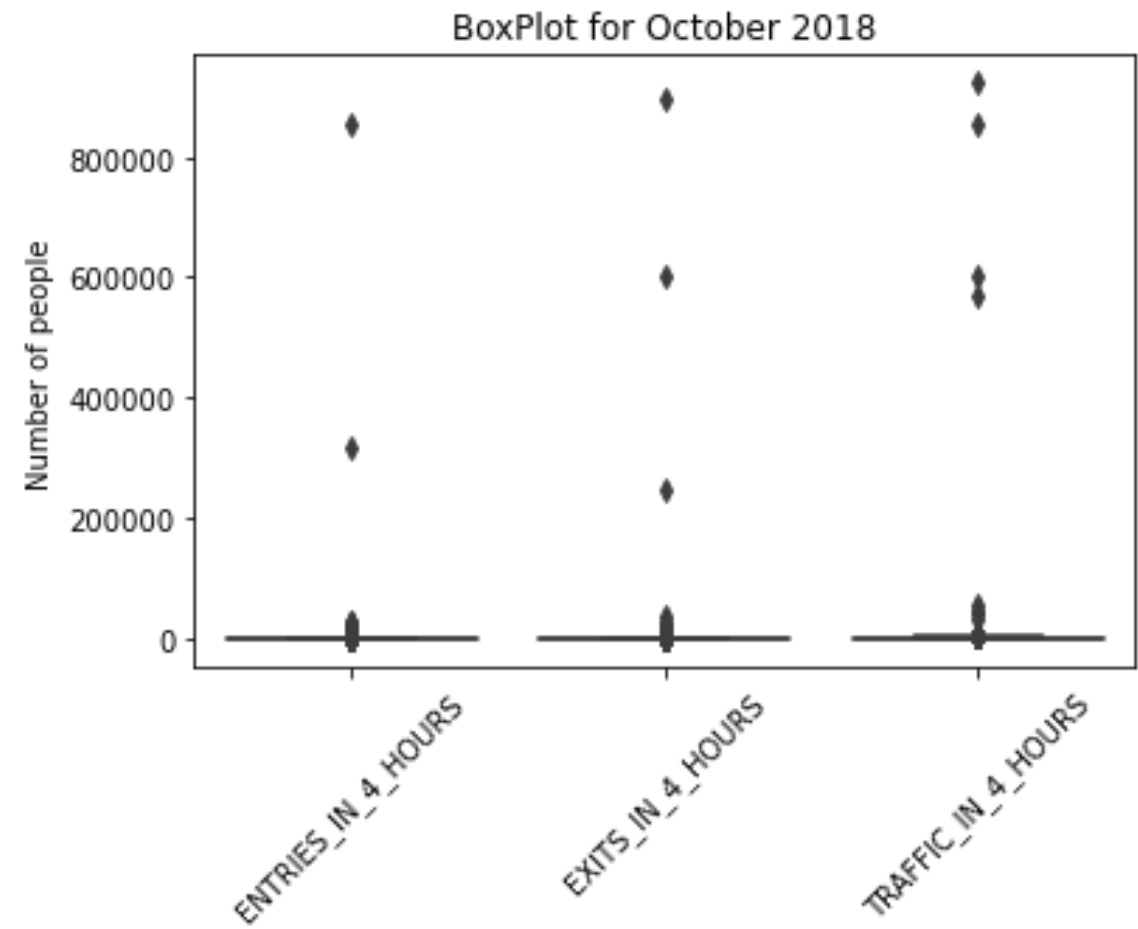
Station with highest number of turnstiles

Data Analysis



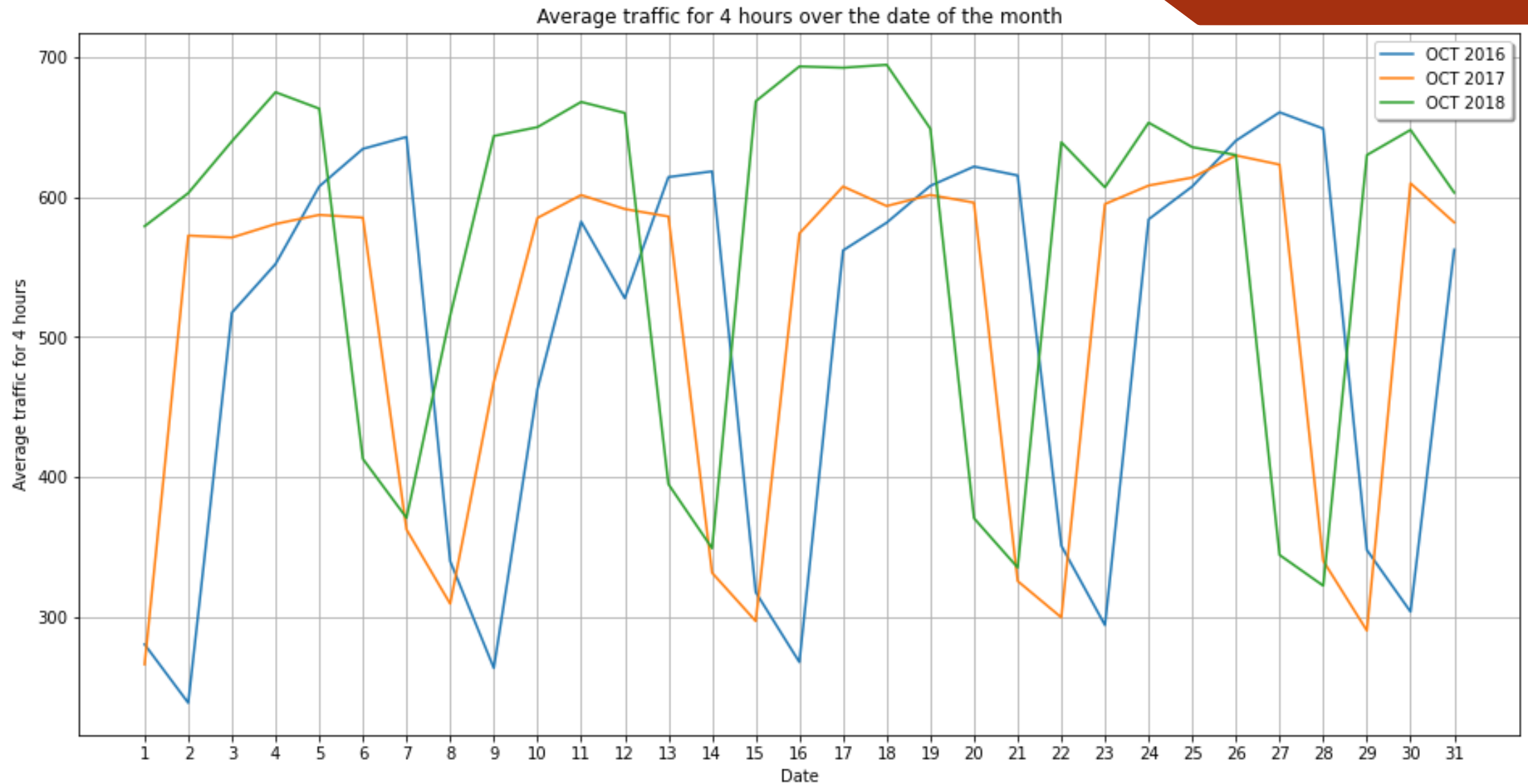


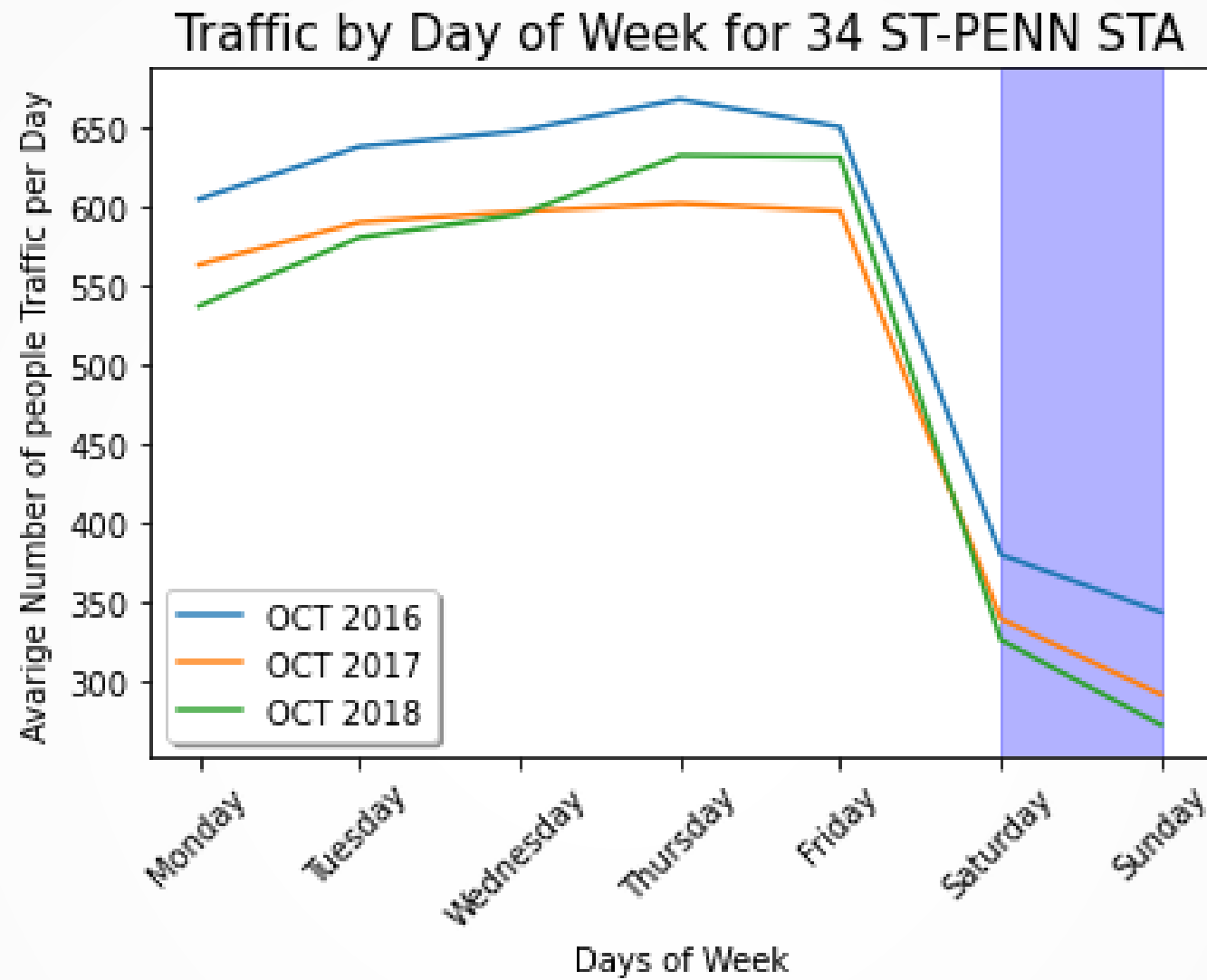
After
Data cleaning



Before
Data cleaning

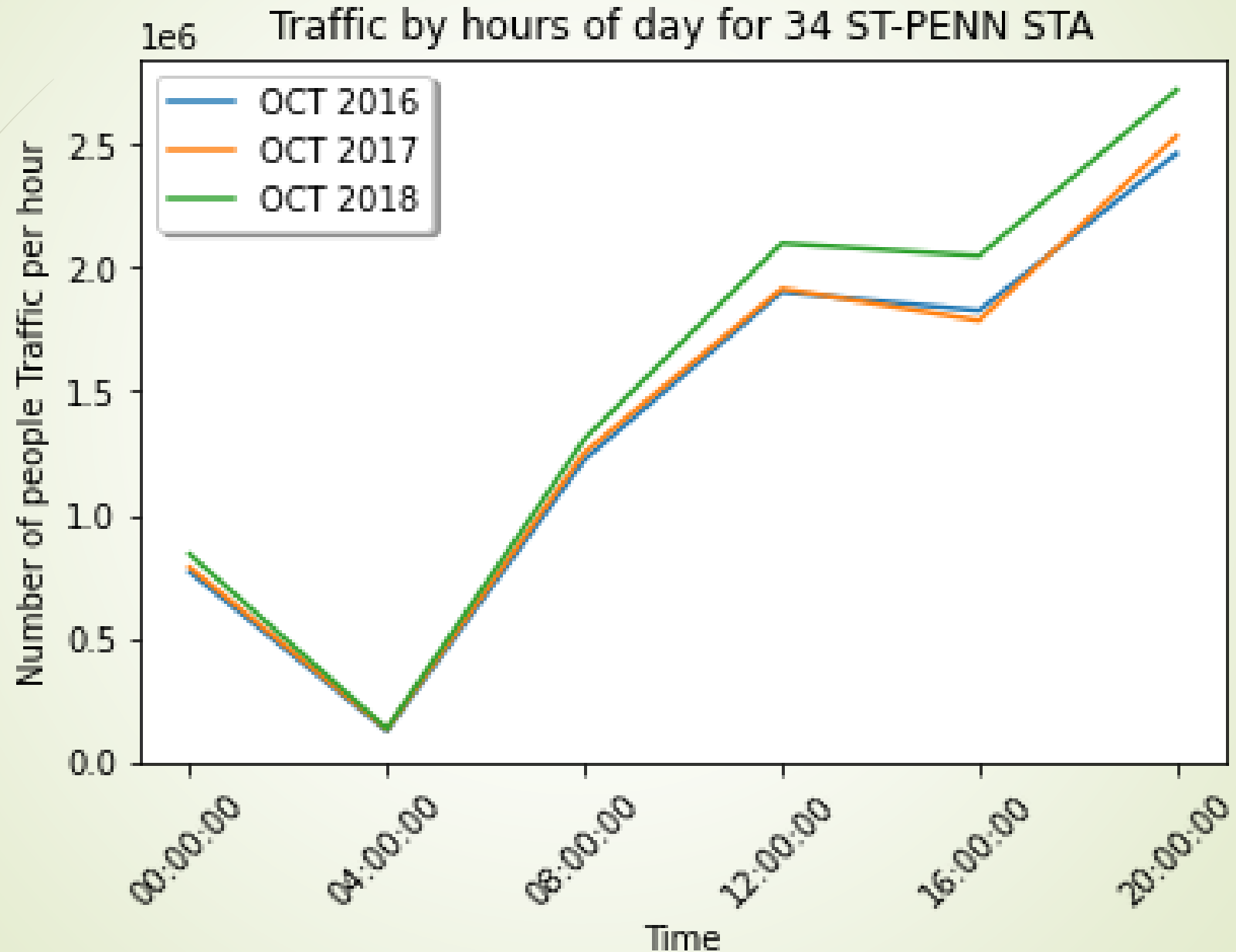
Data Analysis





Traffic by day of week

Traffic over hour of day





Future work

01

Do the analysis for the 13 control areas of 34 ST-PENN STA station to find the exact time where the traffic is at its minimum.

02

Plot the heatmap for each control area over the time of the day.

03

Try to link this data with another data

Conclusion

The background features several thin, light green lines. One line is a long, shallow curve starting from the top center and extending towards the right edge. Another line is a steeper curve starting from the top right and extending towards the bottom right. A third line is a straight vertical line located to the right of the word 'Conclusion'.



Thank you
for listening

