

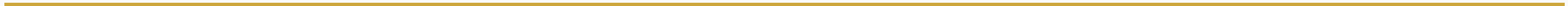
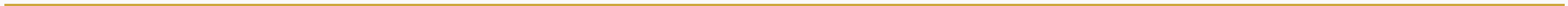


NYC MTA and maintenance company





Introduction



Data set and Tools

Data set

A series of data files containing numbers of cumulative entries and exits by stations, turnstile, along with their dates and time. MTA turnstile data

Sample

first three months of
2021 (January-Mar)
2719827 rows 11
columns

Tools

Numpy

Pandas

Matplotlib

Seaborn

Jupyter



Data

Step 1 Cleaning

Renaming columns

Step 2

Removing duplicates rows

Step 3

Calculating the daily exits and entries

Step 4

Removing outliers Using Boxplot

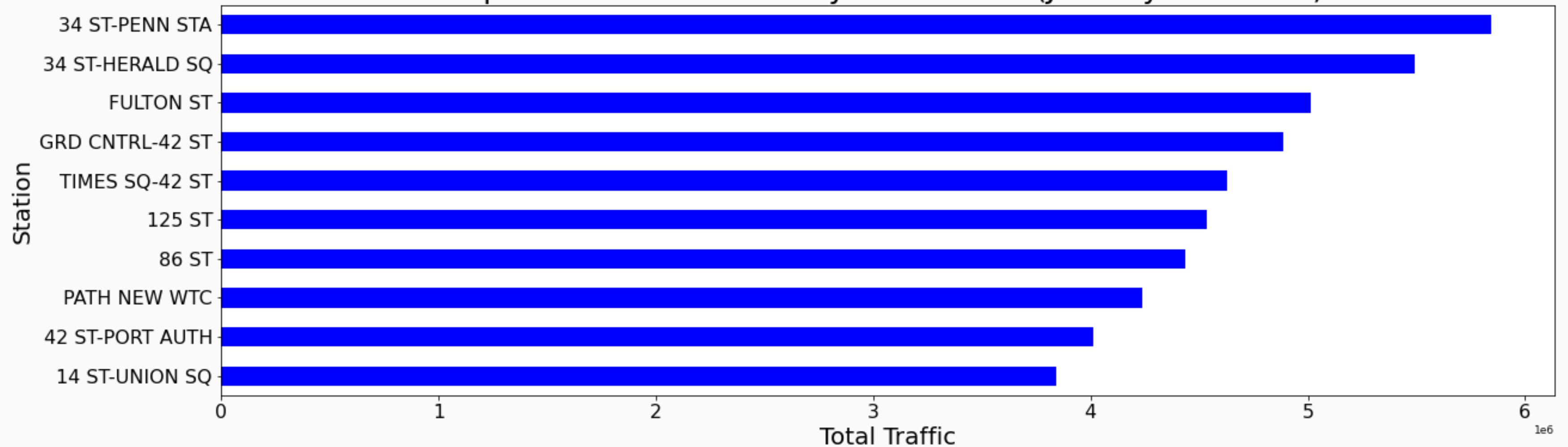




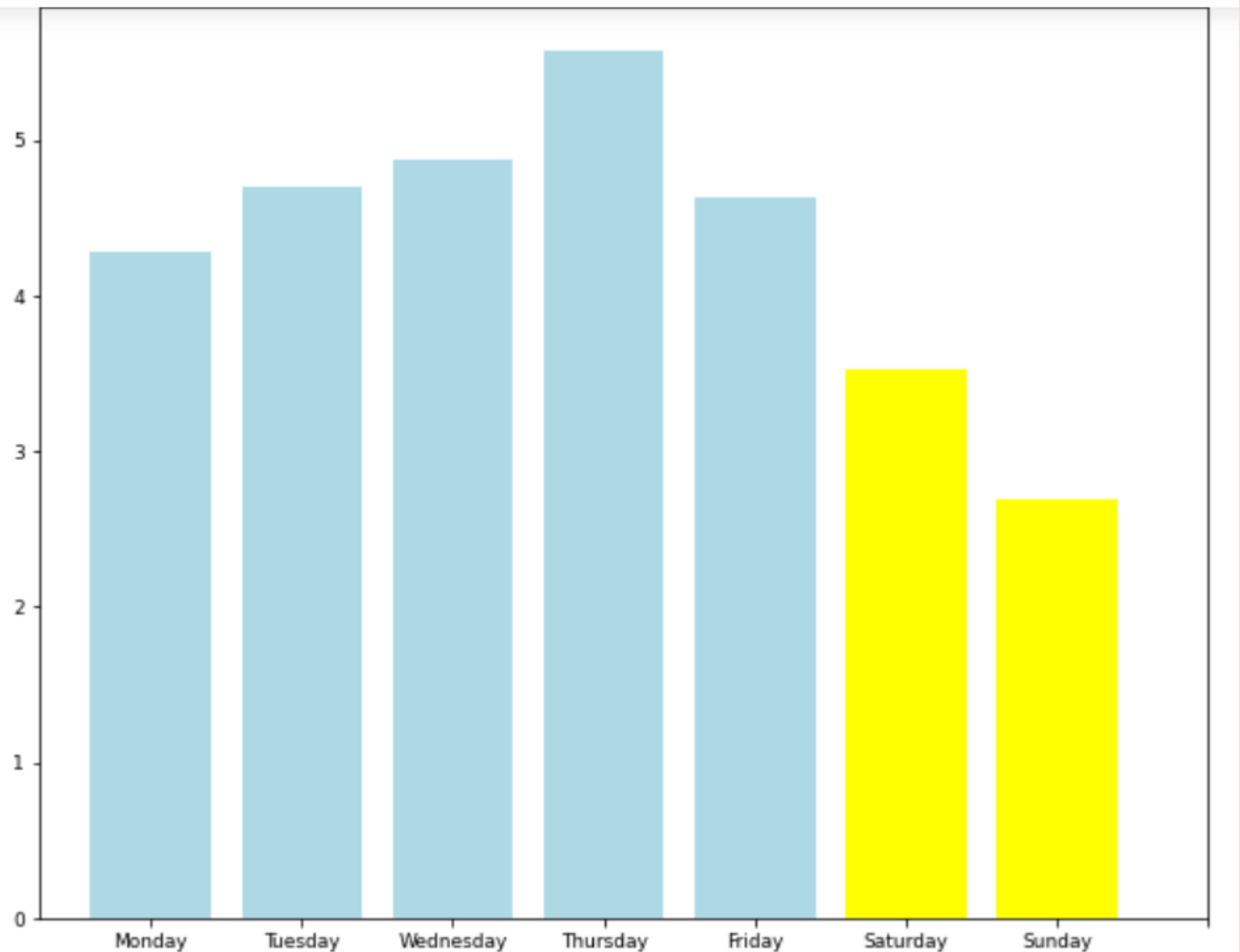
Data Analysis

What are the ten most active stations?

Top 10 Busiest Stations by Total Traffic (January-Mar 2021)



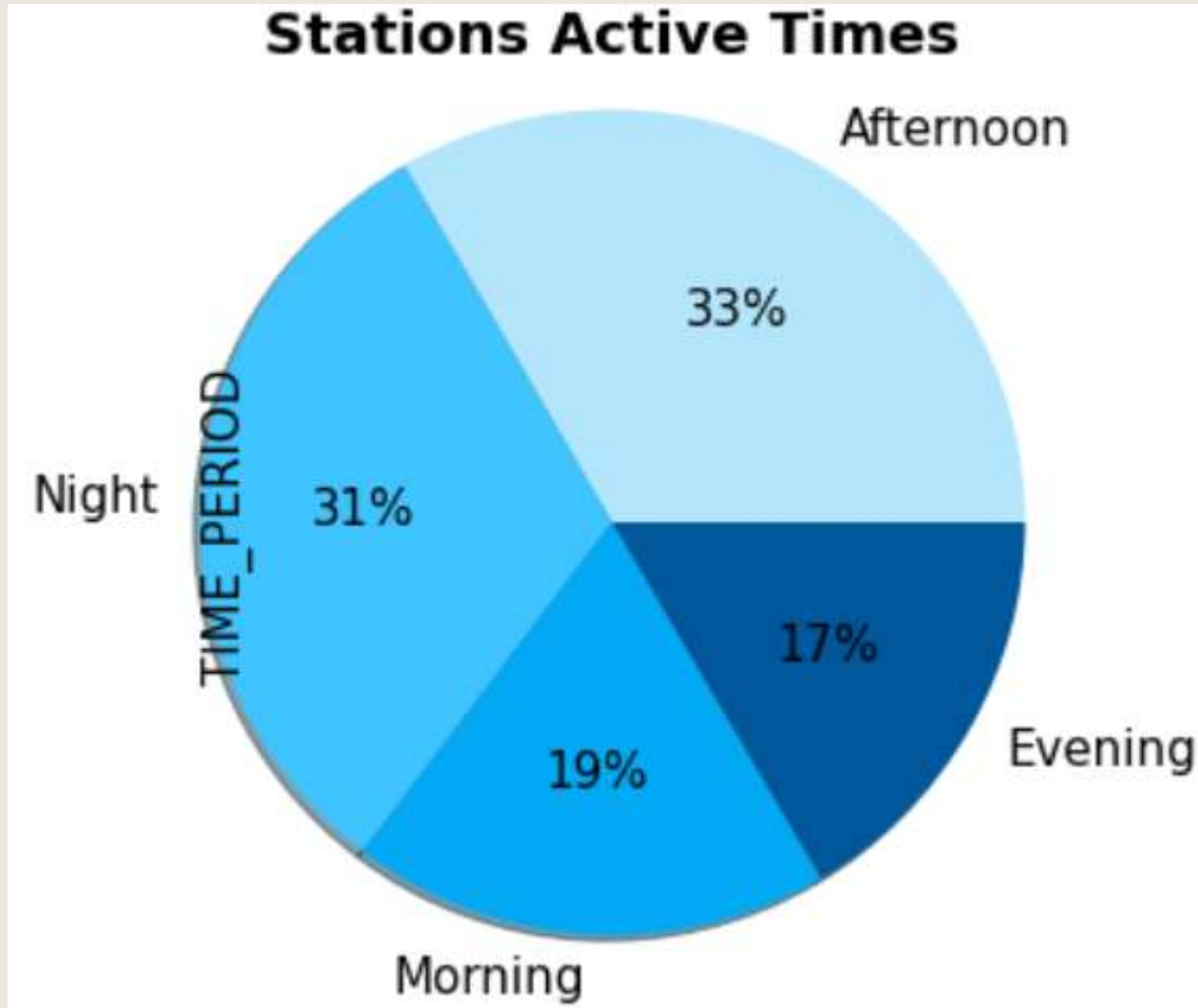
What are the most active days of the week?



TH



What are the most active times for each station?



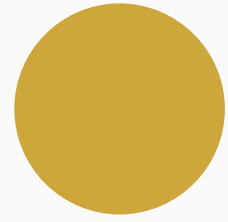
||

TH

●

○

○



Thank you

