

Cleaning industry(EDA)

❖ By: Yasir albahlal



CONTENT

- ❖ BACKSTORY
- ❖ HYPATHOSIS
- ❖ METHODOLOGY
- ❖ DATA PREPARATION
- ❖ ANALYSIS
- ❖ SUGGESTION



Backstory

- ❖ A cleaning company desires to discover the perfect time for cleaning in multiple New York busiest stations.
- ❖ Cleaning must be done when traffic is less.



Hypothesis

- ❖ I assume that that the station will be less busy between 00:00:00 and 08:00:00



Methodology

- ❖ Collecting data from Metropolitan Transportation Authority(MTA) New York subways
- ❖ Taking the past 6 months of data for study analysis, ' April-May-Jun-July-Aug-Sep' 2021

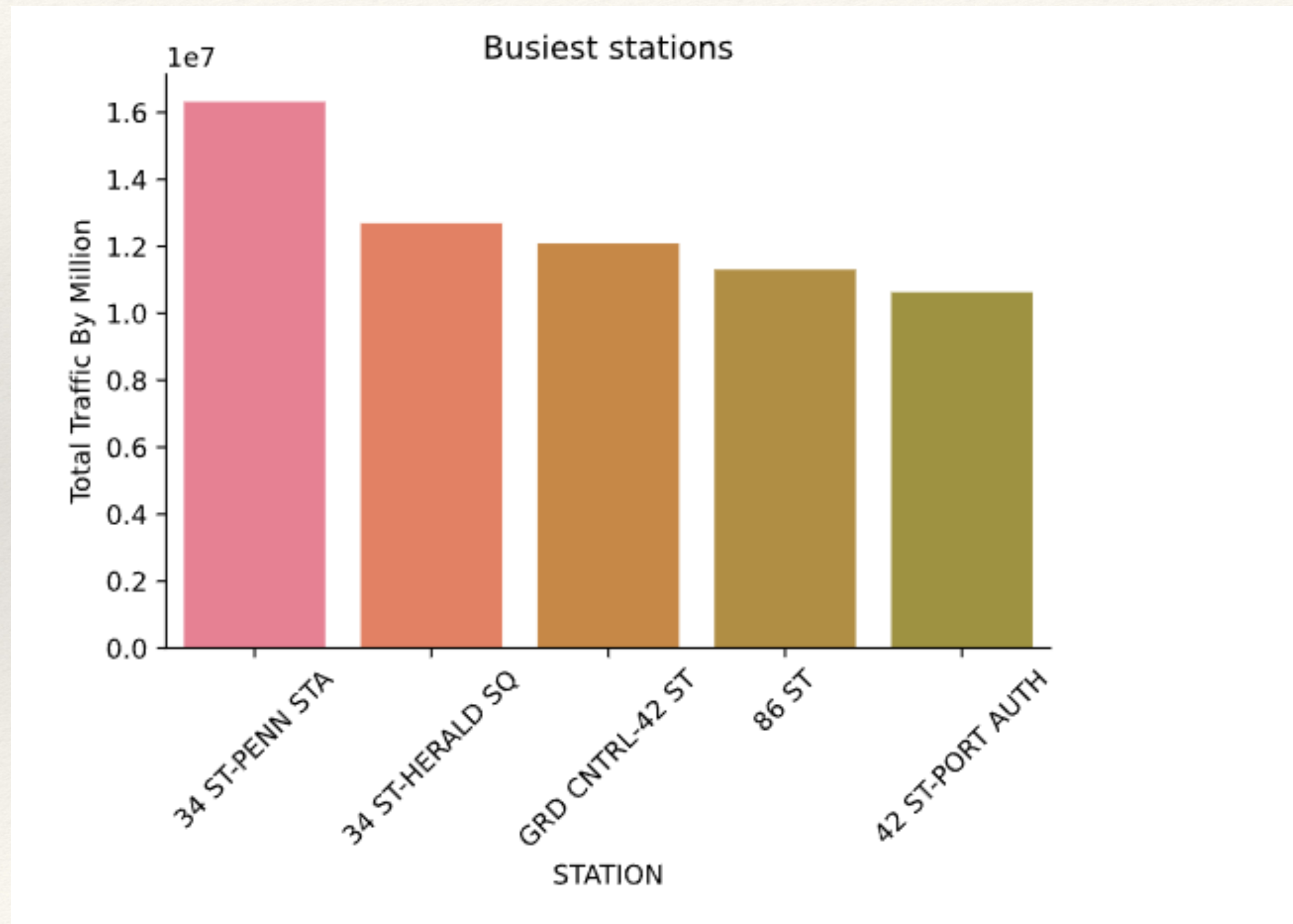


Data preparation

- ❖ Removing duplicates rows & null values
- ❖ Sorting Entries and Exits and add them together to get total traffic
- ❖ Sorting time reading for every 4 hours
- ❖ Adding weekday and weekend column

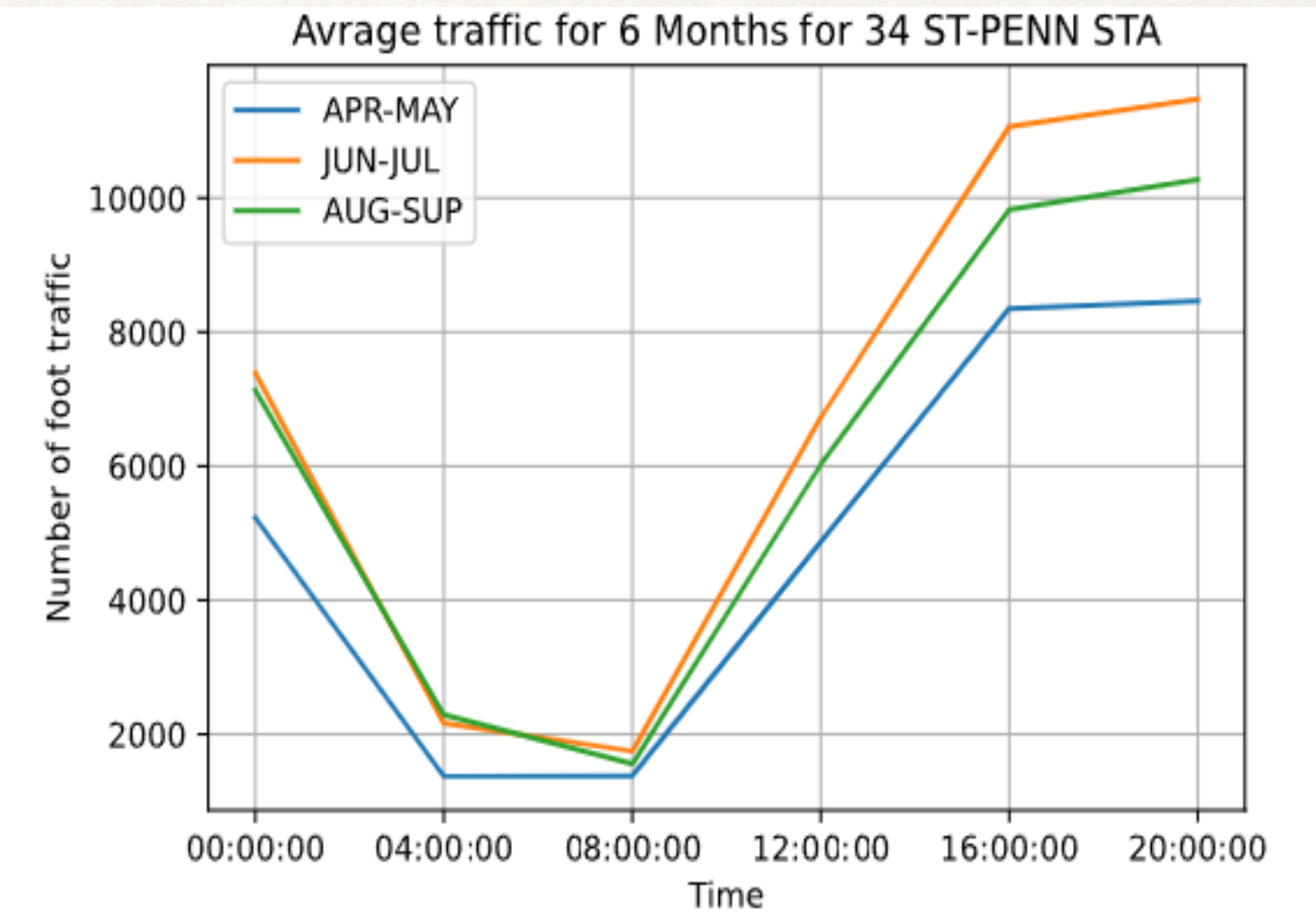
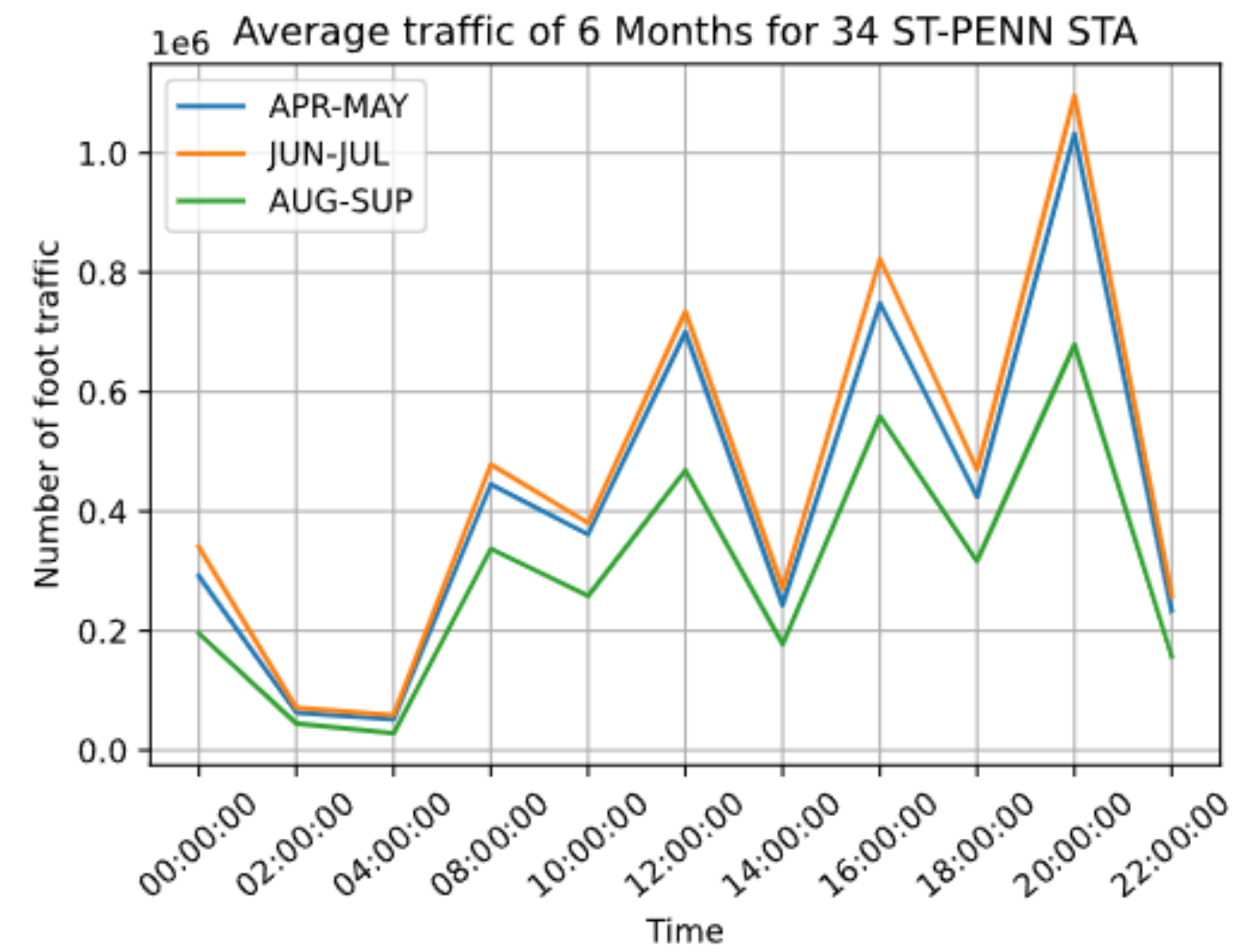
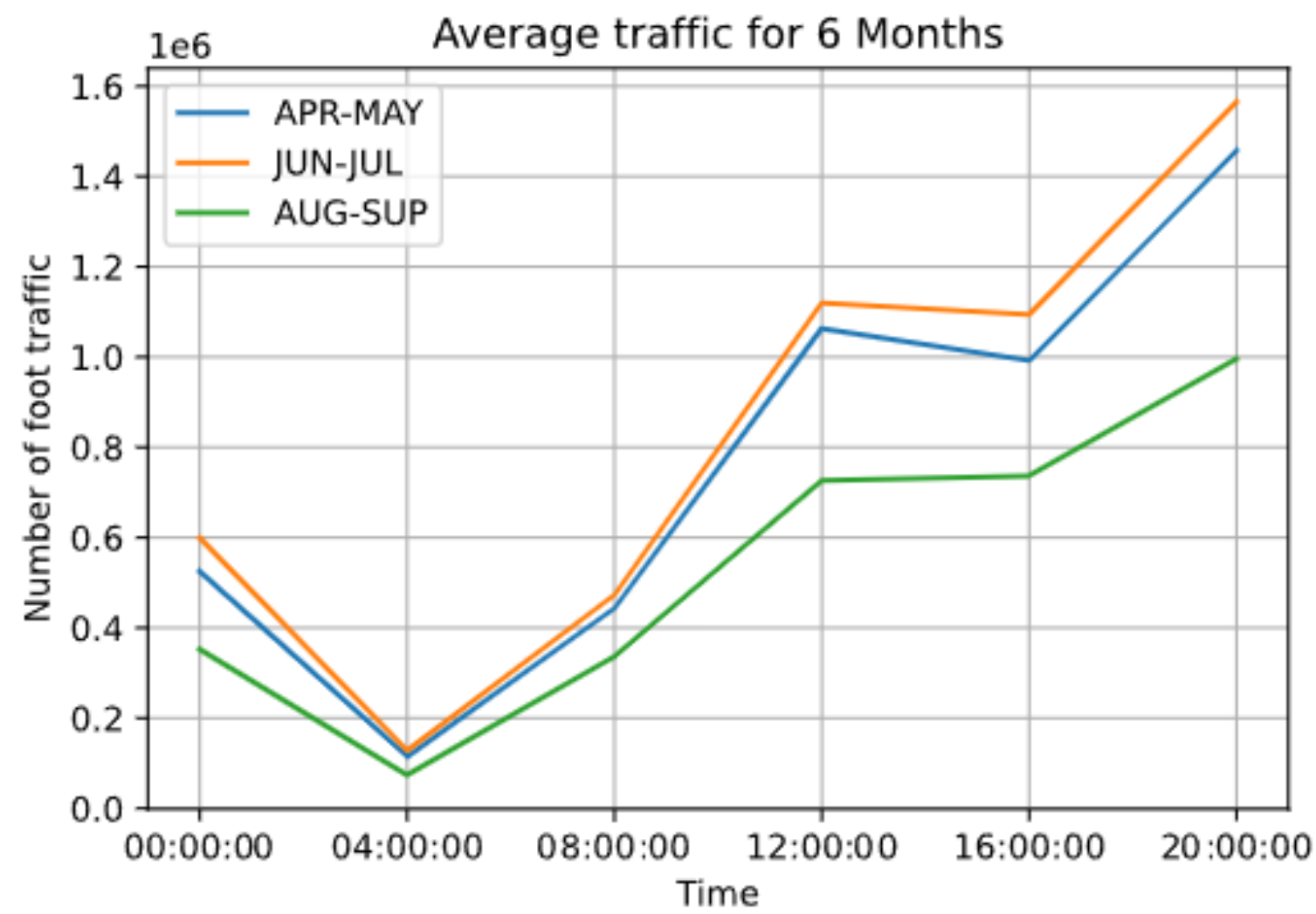
Analysis

TOP 5 STATIONS IN TERMS OF TRAFFIC



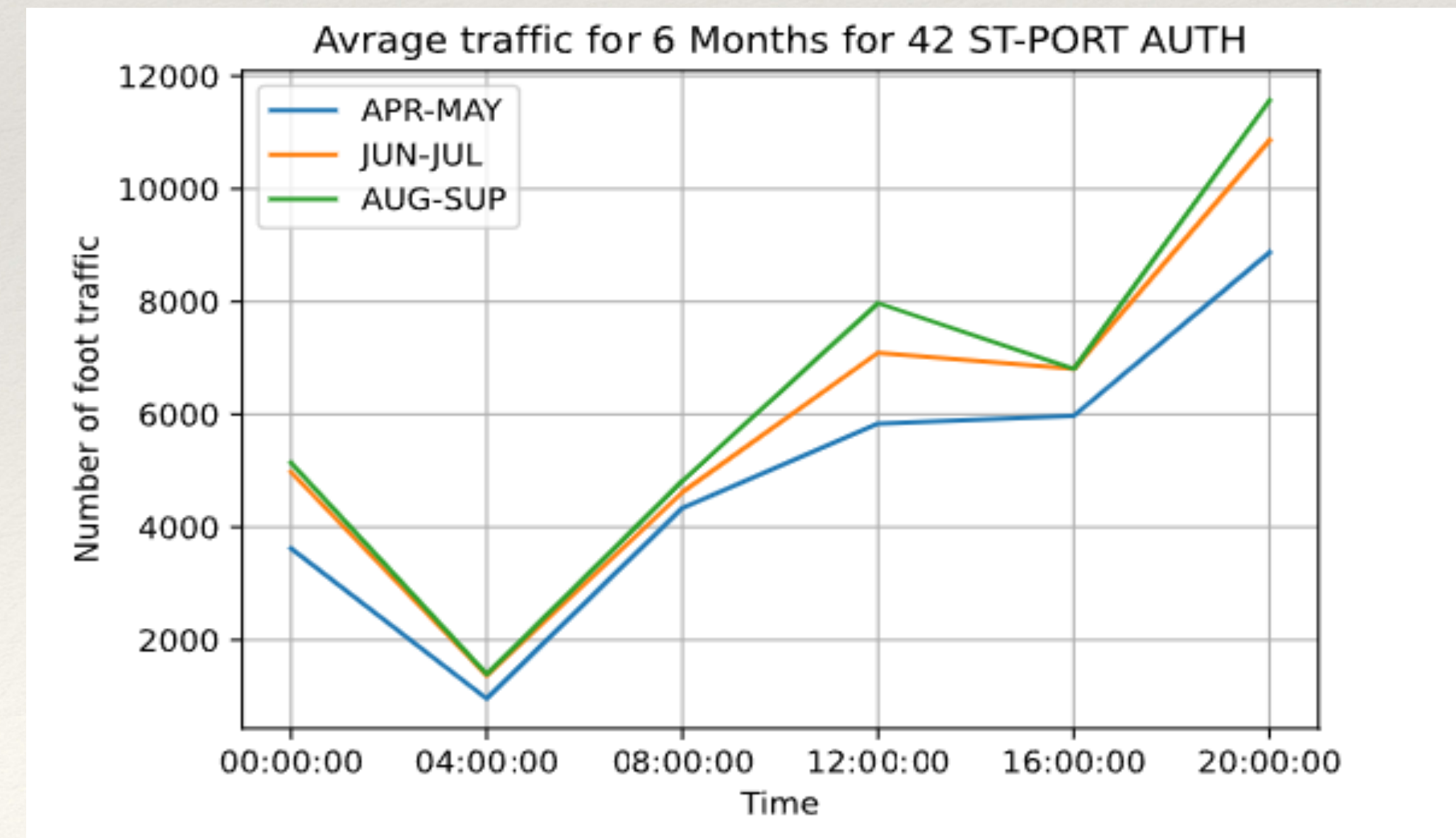
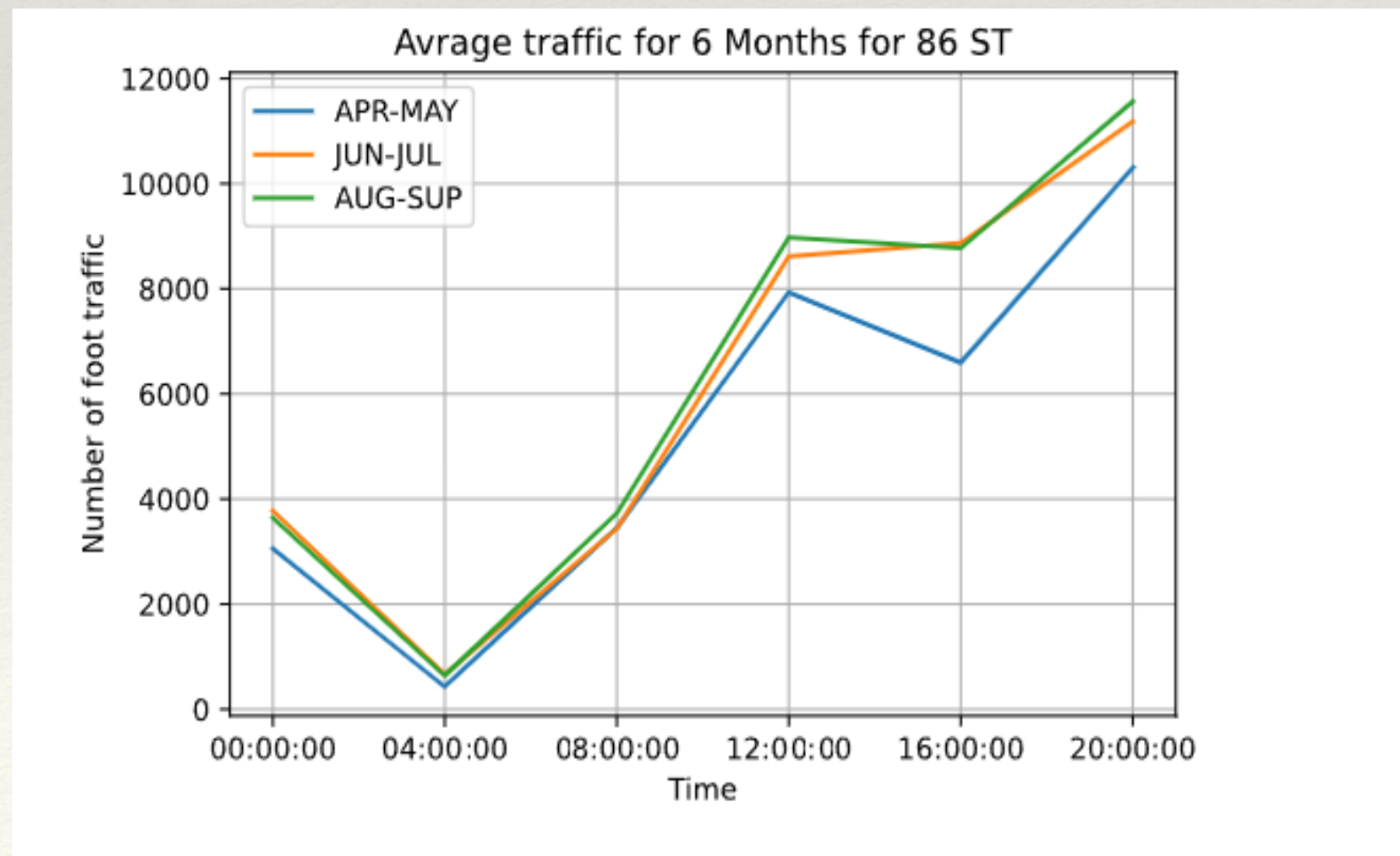
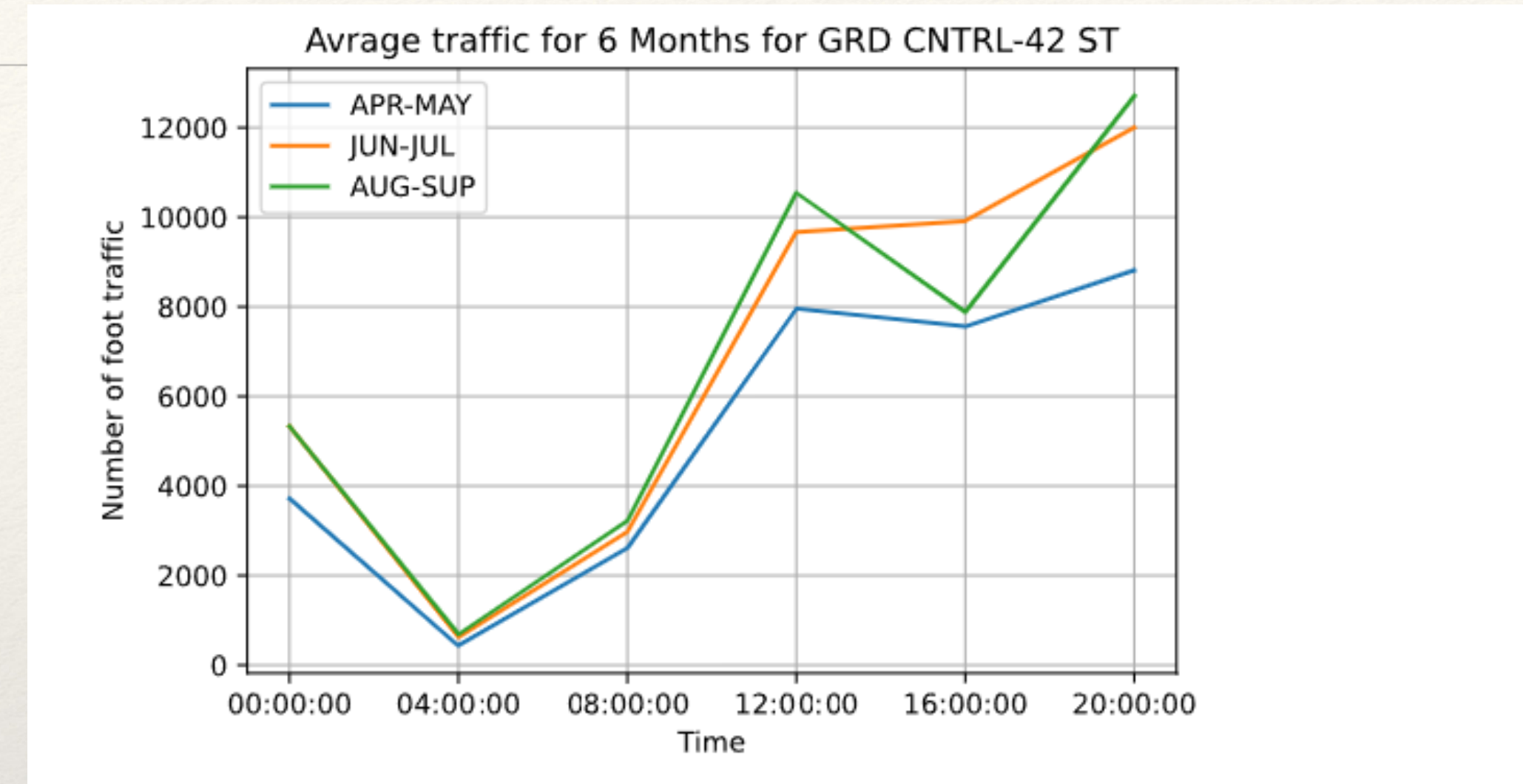
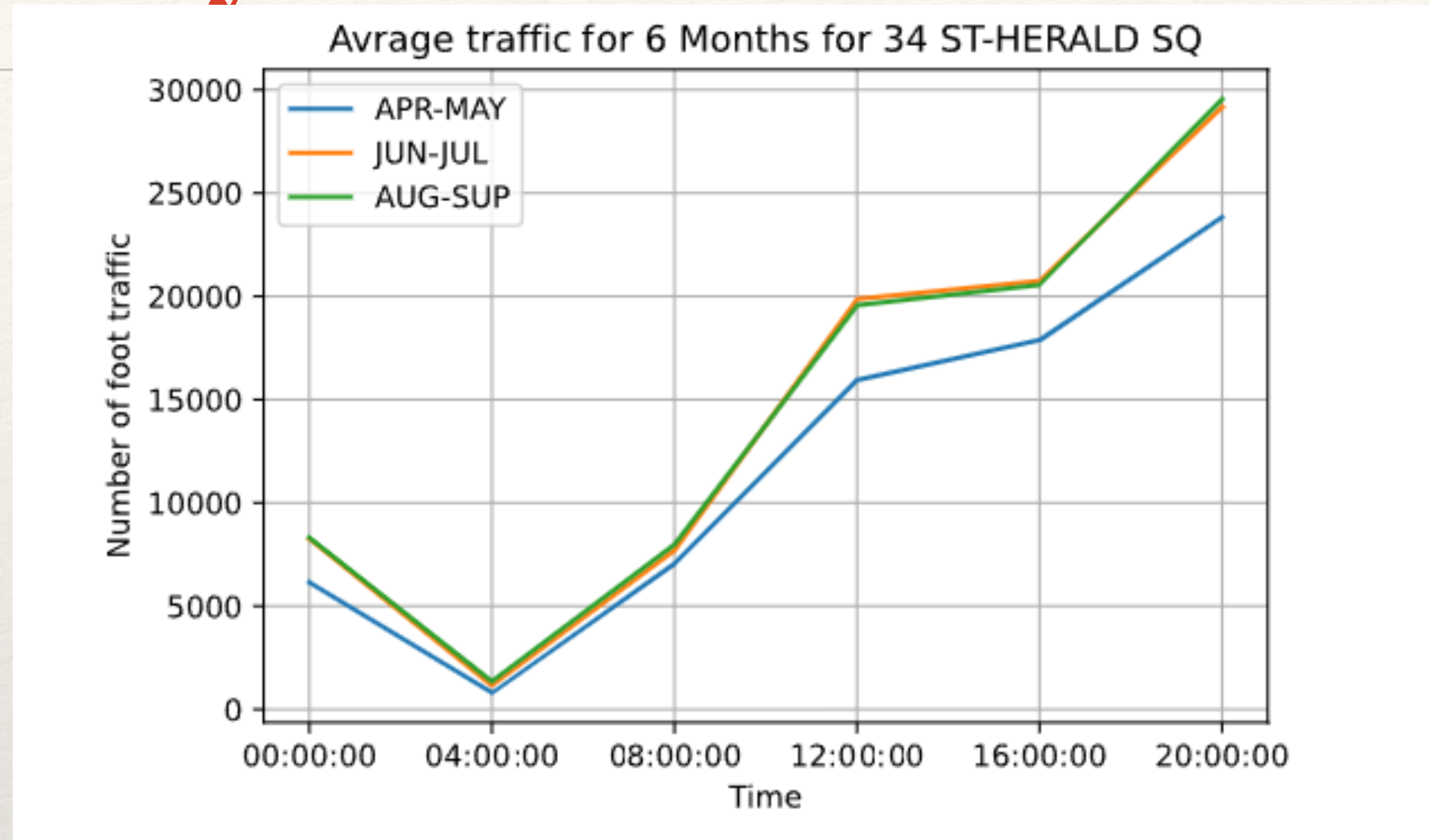
Analysis

❖ Average of 6 months the busiest station '34 ST-PENN STA' D



Analysis

Average of 6 months of the rest of
busiest stations



Suggestion

- ❖ Perfect time for cleaning stations is between 02:00:00 and 06:00:00 during weekday and 04:00:00 and 06:00:00 during weekend.

Any
questions ?

