

Project Benson: MTA Data Challenge

Justin Kuan, Metis 2019 Winter cohort

Project Goals

- Cleanly organize MTA data
- Visualize data in several different ways
 - Day-to-day
 - Week-by-week
- Analyze data trends
- Suggest top stations based on ridership

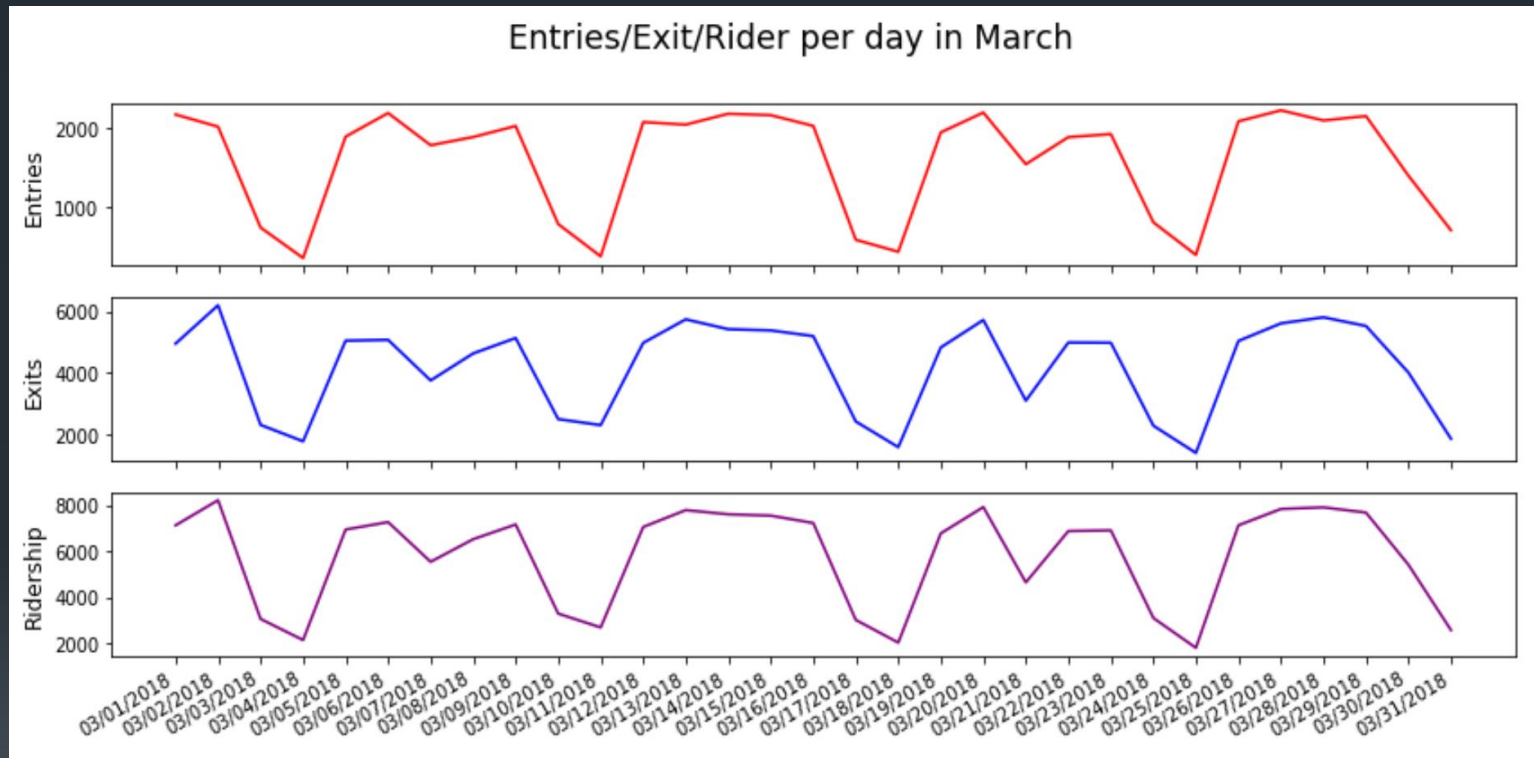
Data wrangling

- Choice of data (March 2018)
- Cleaning time stamps
- *date_time* for time-series
- Sorting the line name problem
- Too few stations
- Cumulative counts

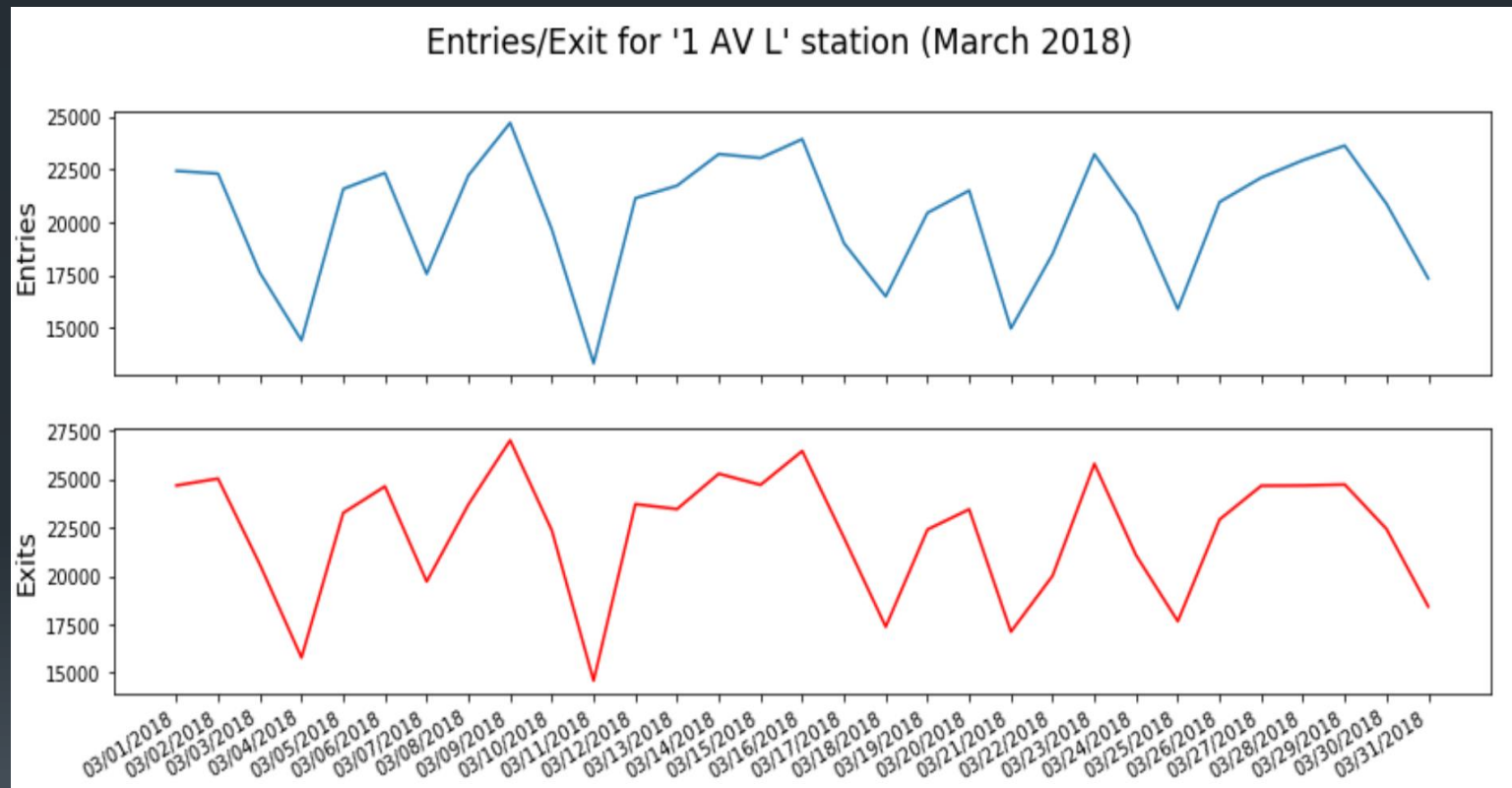


[This Photo](#) by Unknown author is licensed under [CC BY-SA](#).

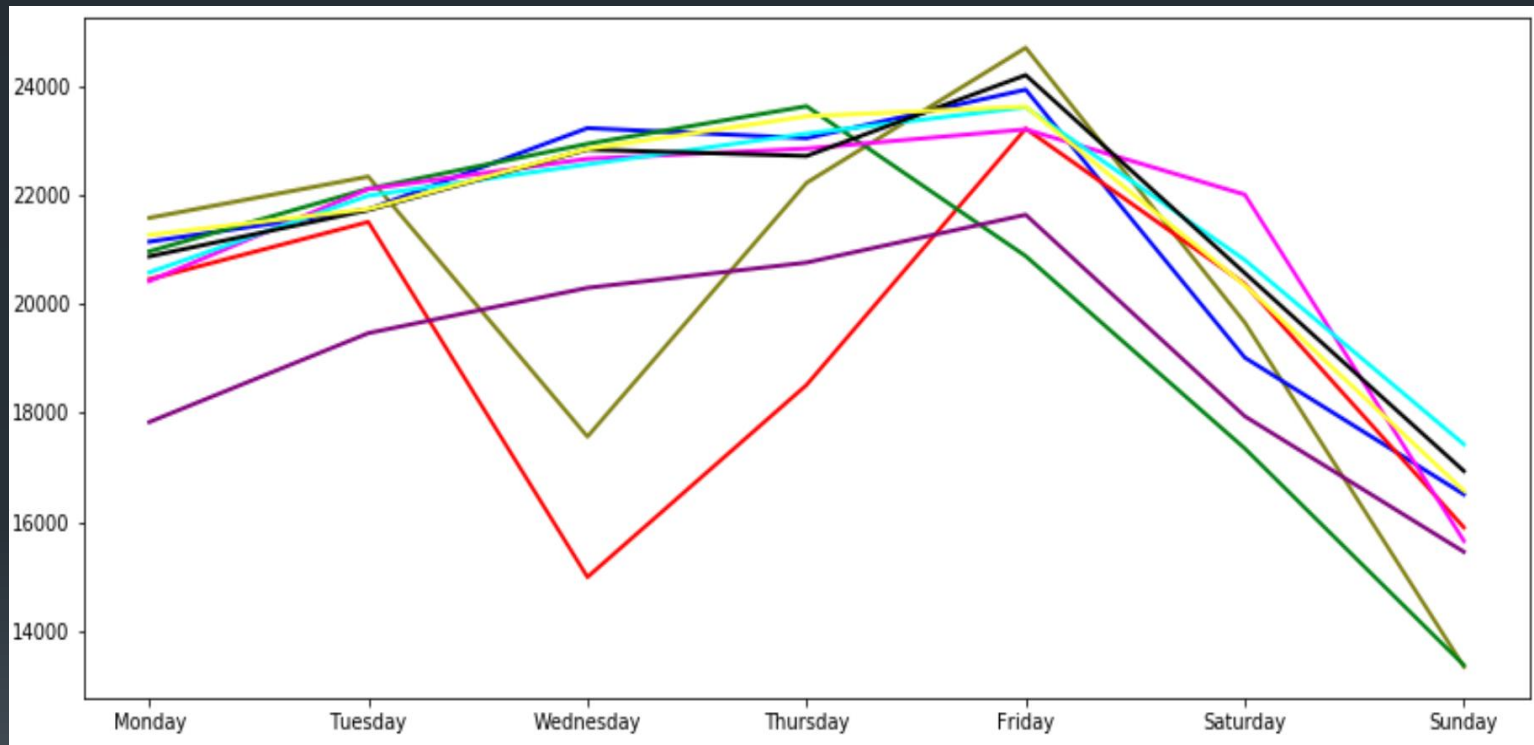
A sample turnstile



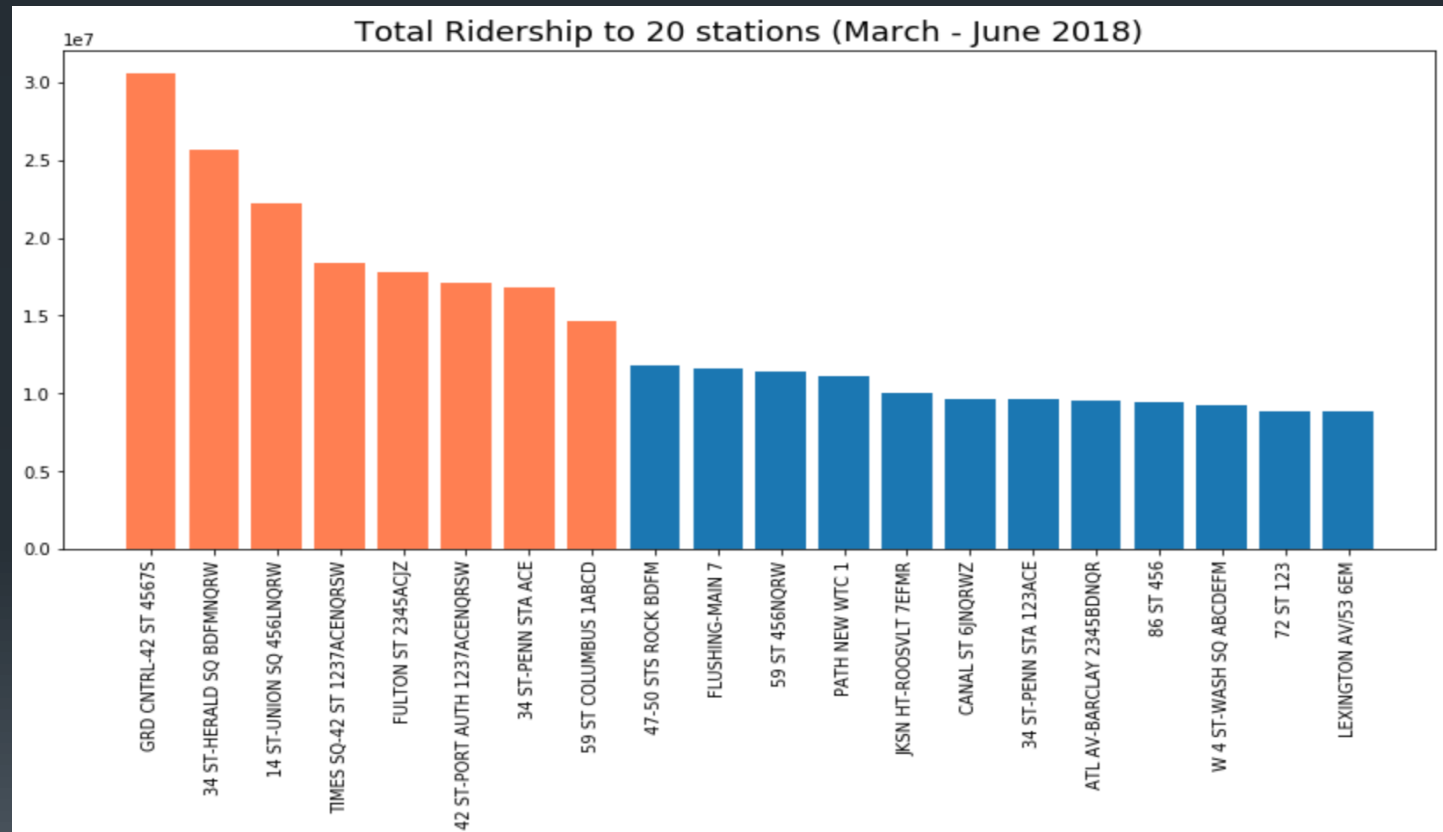
A sample station (1 AV L)



Week-by-week (Mar-Apr)



Total ridership



Conclusion

- GRD CNTRL-42 ST, 34 ST-HERALD SQ,
- 14 ST-UNION SQ, TIMES SQ-42 ST, FULTON ST, 42 ST-PORT AUTH, 34 ST-PENN STA, 59 ST COLUMNBUS

Further Work:

- Filter out weekend count
- Time of day optimization

Conclusion