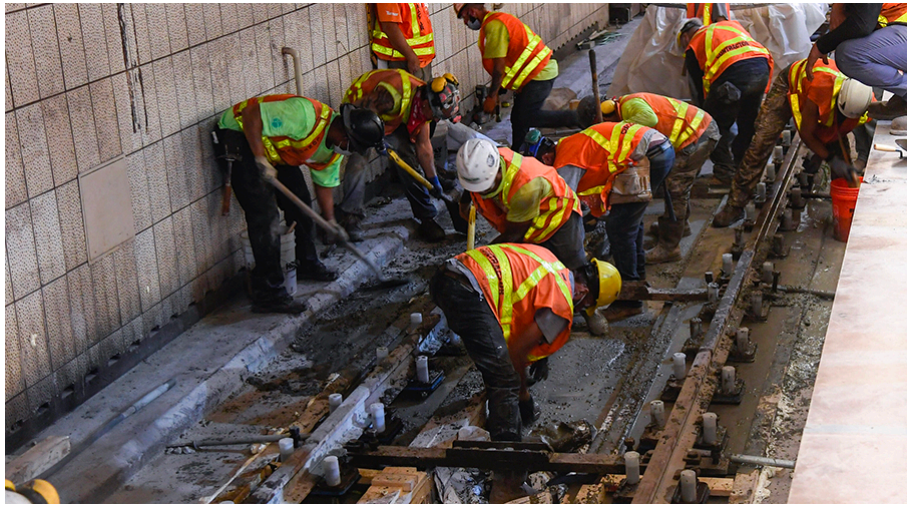
Module 1 Project: MTA Data

Kenneth Hua

MTA Maintenance Projects

Maintenance Project 1:

After every winter, the MTA needs to inspect/replace damaged tracks. This project takes 3 days to complete. What is the optimal 3 day span in spring to complete this project with minimal disruption?



Maintenance Project 2:

The MTA needs to lubricate the tracks on a weekly basis, to prevent friction from wearing down subway tracks. This process takes 3 hours to complete. What is the optimal time span in a week to regularly schedule this service with minimal disruption?

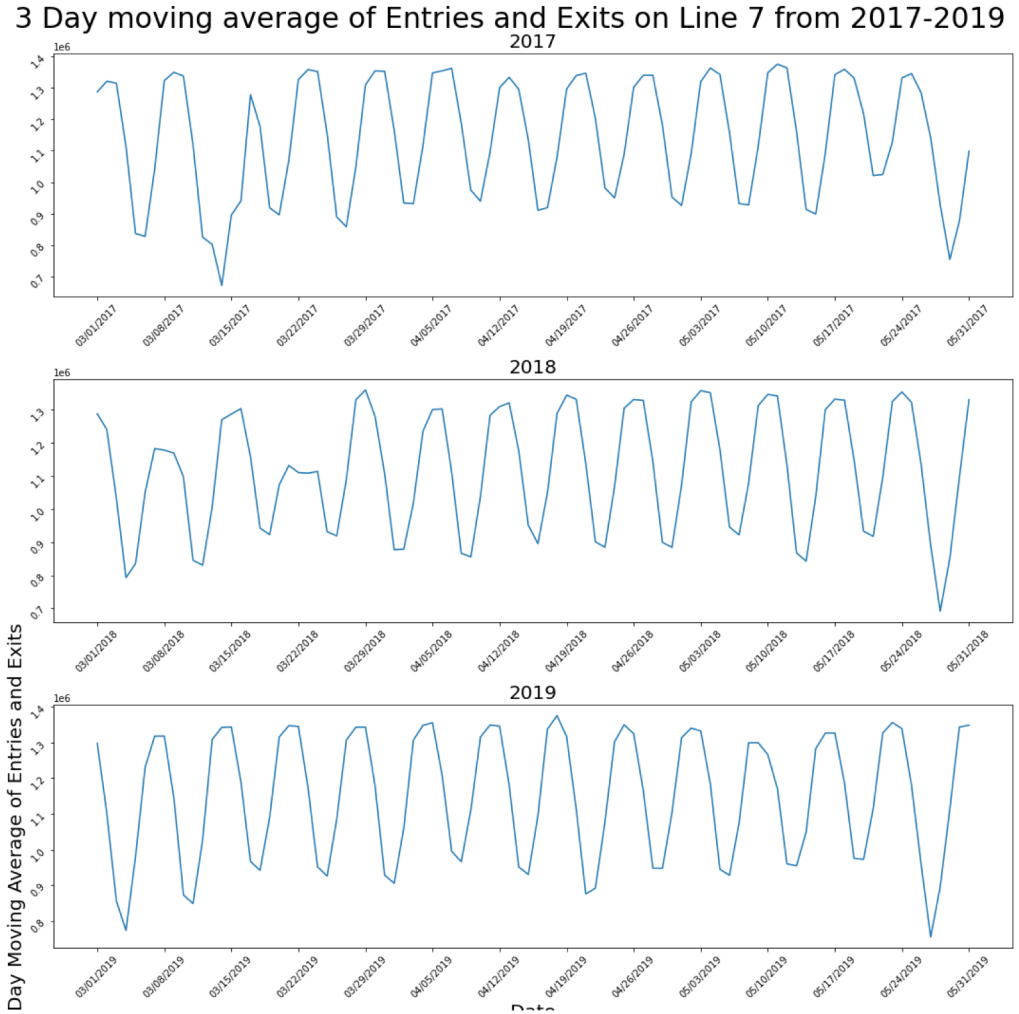
Data Source + Process

	C/A	UNIT	SCP	STATION	LINENAME	DIVISION	DATE	TIME	DESC	ENTRIES	EXITS
0	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	00:00:00	REGULAR	5799442	1966041
1	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	04:00:00	REGULAR	5799463	1966044
2	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	08:00:00	REGULAR	5799492	1966079
3	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	12:00:00	REGULAR	5799610	1966155
4	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	16:00:00	REGULAR	5799833	1966214
5	A002	R051	02-00-00	59 ST	NQR456	BMT	08/27/2016	20:00:00	REGULAR	5800121	1966271
6	A002	R051	02-00-00	59 ST	NQR456	BMT	08/28/2016	00:00:00	REGULAR	5800252	1966295
7	A002	R051	02-00-00	59 ST	NQR456	BMT	08/28/2016	04:00:00	REGULAR	5800281	1966303
8	A002	R051	02-00-00	59 ST	NQR456	BMT	08/28/2016	08:00:00	REGULAR	5800295	1966317
9	A002	R051	02-00-00	59 ST	NQR456	BMT	08/28/2016	12:00:00	REGULAR	5800377	1966387

Steps:

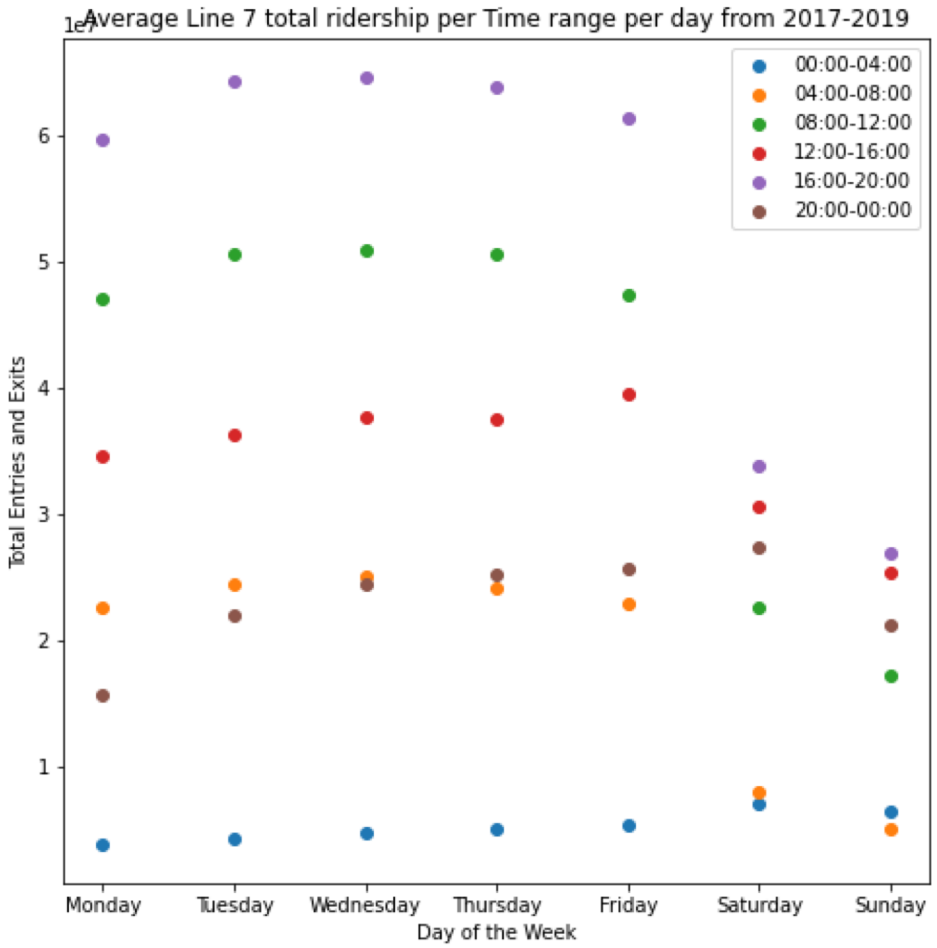
- Data Sourcing (Web Scraping, SQL)
- Data Cleaning(Pandas)
- Aggregation(Pandas groupby)
- Visualization (Matplotlib)
- Analysis + Conclusions

Maintenance Project 1: 7 Line Track Replacement



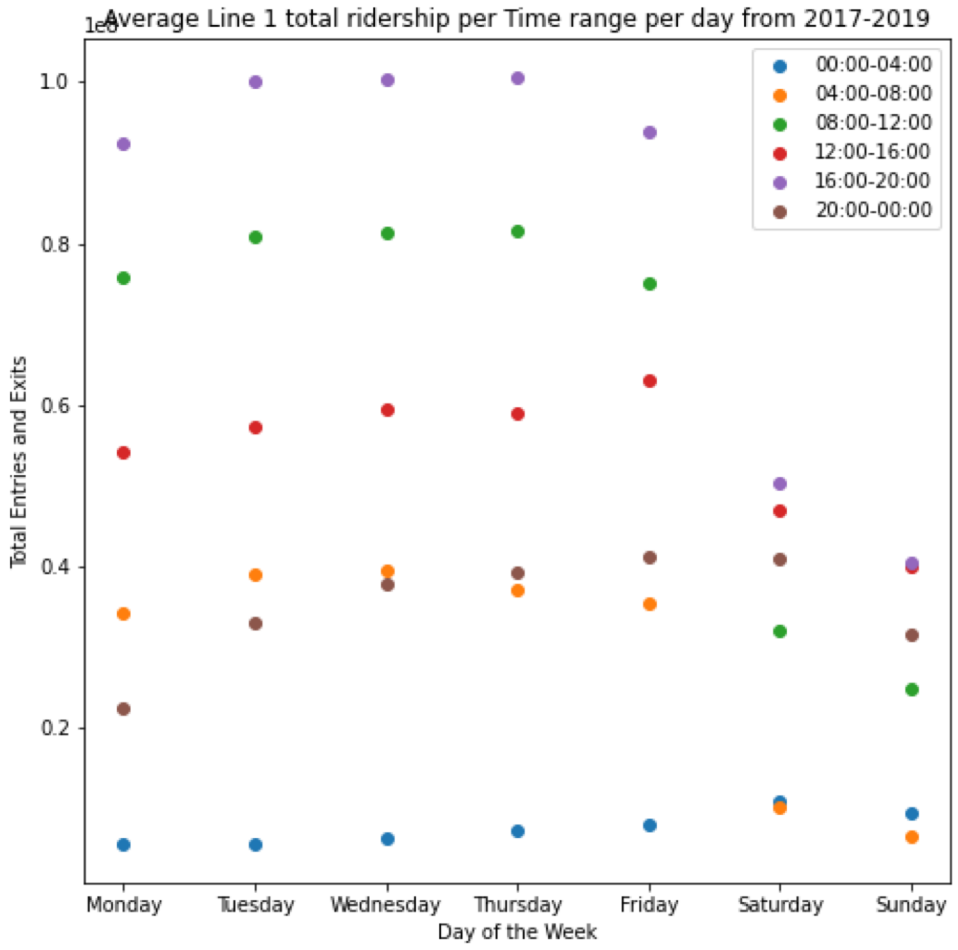
- Data is periodic, with peak ridership on the weekdays and lowest ridership on the weekends.
- The lowest 3-day average ridership is at the end of May, which is Memorial Day Weekend.
- The other low 3-day averages are in the first few weeks of March, before the Spring traffic has picked up.
- **On average, a weekend in the beginning of March is the most optimal choice for the planning of this line track replacement work.**

Maintenance Project 2: 7 Line Track Lubrication



- Line 7 is a commuter line running from Manhattan to Queens.
- Ridership is highest on weekday rush hours 08:00-12:00 and 16:00-20:00.
- Weekend ridership is lower and the behavior is different. (08:00-12:00 ridership is much lower on weekends).
- **For Line 7, the lowest ridership is on Monday from 0:00 – 4:00**

Maintenance Project 2: 1 Line Track Lubrication



- Line 1 is a central line running through Manhattan. It is used for various purposes, not just commuting
- The data is mostly similar to Line 7, but there are some slight differences (ridership on 00:00-4:00 Saturday & Sunday are relatively higher)
- **For Line 1, the lowest ridership is either on Monday from 0:00-4:00 or Sunday from 4:00 – 8:00.**
- **Because the maintenance team is scheduling work on Line 7 on Monday from 0:00-4:00, Sunday from 4:00-8:00 seems to be the best time.**

Questions?

- For more information/code: visit github page: https://github.com/kenhua15/MTA-Exploratory-Data-Project