

MTA Exploratory Data Analysis

Presented by Ghlih maher

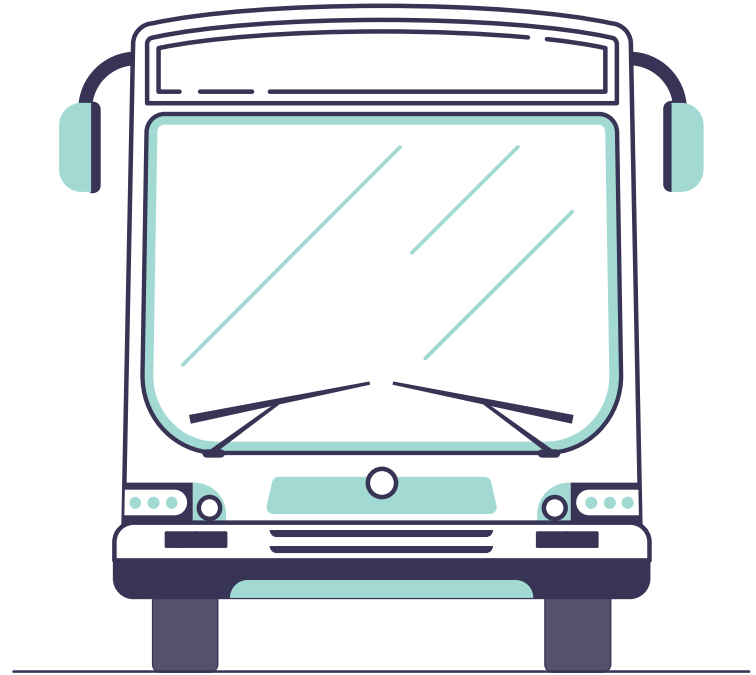
CONTENTS

- Introduction.
- Problem.
- Algorithm.
- Data.
- Conclusion.



Introduction

- MTA is the Metropolitan transportation authority in New York.
- The first case of covid-19 arrived in march.
- The total lockdown happened in March until April 2020.
- There was partial lockdown after April.





Problem Statment

Did the use of public transportation differ six months before
and after the total and partial lockdown?

Algorithm

Data cleaning:

- Striped the columns of all the white spaces and next line space.
- Dropped the C\A, unit, linenname, division, and the desc columns.
- Convert the date into datetime.
- Drop the negative values after calculations.
- Drop the outliers.
- Drop the NA values.

Data analysis:

- added three new columns named: TOTAL_ENTRIES, TOTAL_EXITS, TOTAL_TRAFFIC
- calculated the total traffic in one month, and append it to another data frame, in order to easily display the chart without the need to re-run the whole code.

Data before cleaning

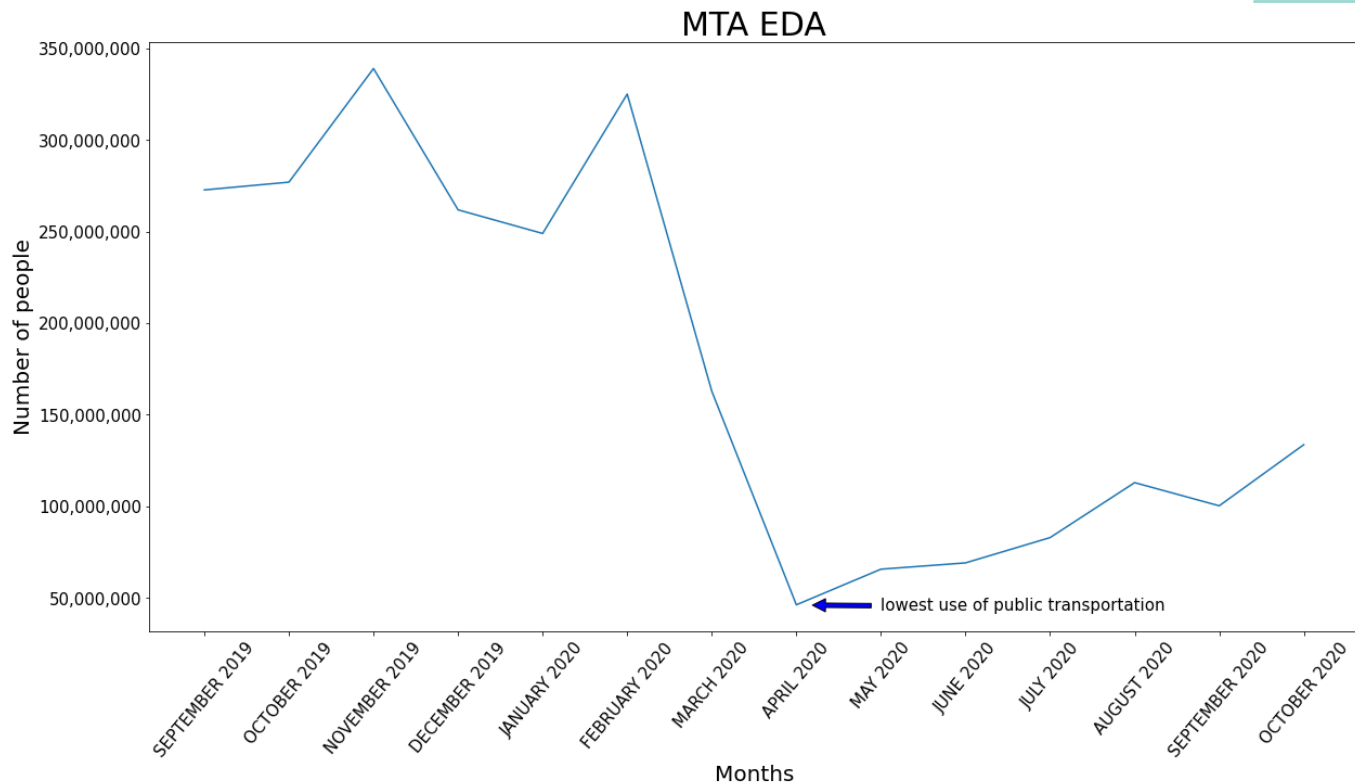
	C/A	UNIT	SCP	STATION	LINENAME	DIVISION	DATE	TIME	DESC	ENTRIES	EXITS
0	A002	R051	02-00-00	59 ST	NQR456W	BMT	10/24/2020	00:00:00	REGULAR	7477474	2544276
1	A002	R051	02-00-00	59 ST	NQR456W	BMT	10/24/2020	04:00:00	REGULAR	7477477	2544276
2	A002	R051	02-00-00	59 ST	NQR456W	BMT	10/24/2020	08:00:00	REGULAR	7477488	2544294
3	A002	R051	02-00-00	59 ST	NQR456W	BMT	10/24/2020	12:00:00	REGULAR	7477522	2544334
4	A002	R051	02-00-00	59 ST	NQR456W	BMT	10/24/2020	16:00:00	REGULAR	7477644	2544363

Data after cleaning

	STATION	DATE	TIME	ENTRIES	EXITS	TOTAL_ENTRIES	TOTAL_EXITS	TOTAL_TRAFFIC
1	59 ST	2020-10-24	04:00:00	7477477	2544276	3.0	0.0	3.0
2	59 ST	2020-10-24	08:00:00	7477488	2544294	11.0	18.0	29.0
3	59 ST	2020-10-24	12:00:00	7477522	2544334	34.0	40.0	74.0
4	59 ST	2020-10-24	16:00:00	7477644	2544363	122.0	29.0	151.0
5	59 ST	2020-10-24	20:00:00	7477785	2544386	141.0	23.0	164.0

Conclusion

	MONTH	TOTAL_TRAFFIC
0	OCTOBER 2020	133596429
1	SEPTEMBER 2020	100216541
2	AUGUST 2020	112864786
3	JULY 2020	82855166
4	JUNE 2020	69046048
5	MAY 2020	65611019
6	APRIL 2020	46167307
7	MARCH 2020	162791444
8	FEBRUARY 2020	325014415
9	JANUARY 2020	248903149
10	DECEMBER 2019	261815687
11	NOVEMBER 2019	338969289
12	OCTOBER 2019	276936304
13	SEPTEMBER 2019	272686638





**THANKE YOU FOR
LISTENING**

Any questions?