The Effect of Fare Evasion Crime on MTA IN NY

EXPLORATORY DATA ANALYSIS

Ahad Saad Almutairi

01. FARE EVASION CRIME

02. ARISING ISSUES

03. DATASETS + TOOLS

04. DATA CLEANING

05. EDA GRAPHS + INSIGHTS

06. CONCLUSION + RECCOMENDATION

FARE EVASION

is the act of travelling on public transport without paying

WHATS THE ISSUE?

FINANCIAL LOSS

HIGH FARE EVASION CRIME RATE

DATASET

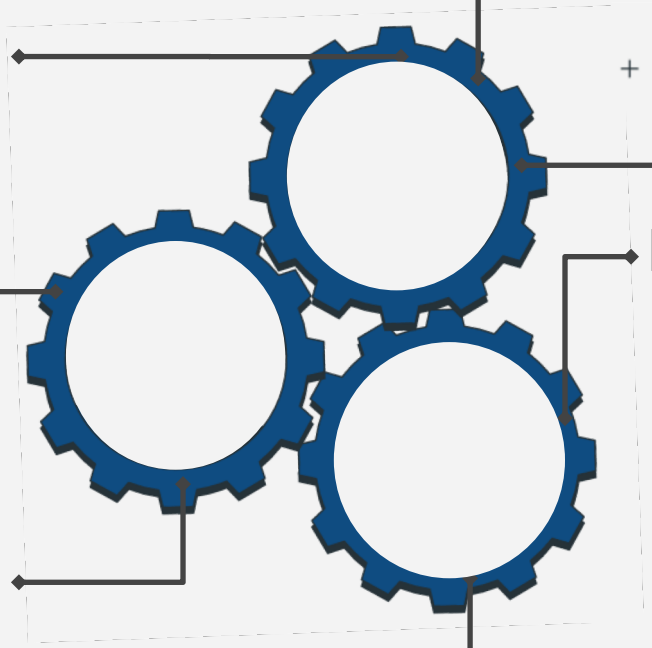
MTA TURN STILE DATASET

SAMPLE:2018

ARRESTS FOR FARE EVASION DATASET

SAMPLE:2018

TOOLS


NUMPY
JUPYTER NOTEBOOK
PANDAS
PYTHON
MATPLOTLIB
SQLALCHEMY
SEABORN

DATA CLEANING

01

Stripping whitespaces from columns

02

Duplicated values

03

Getting rid of irregular readings

04

Focusing on entries only

05

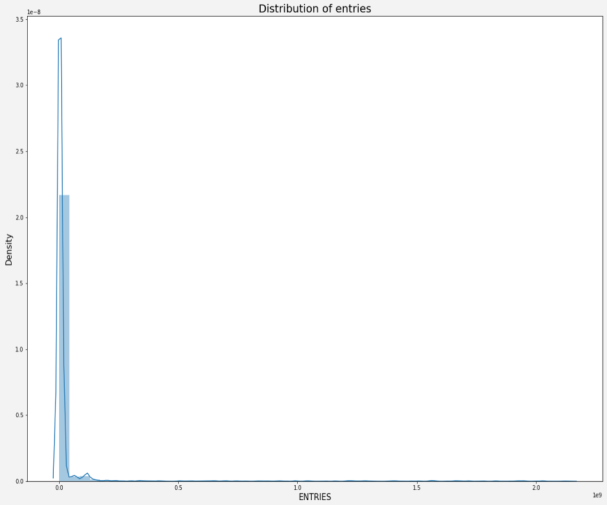
Eliminating outliers

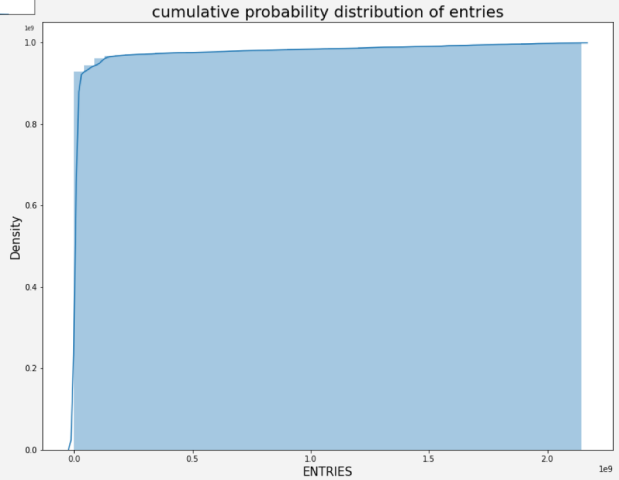
06

Eliminating false readings

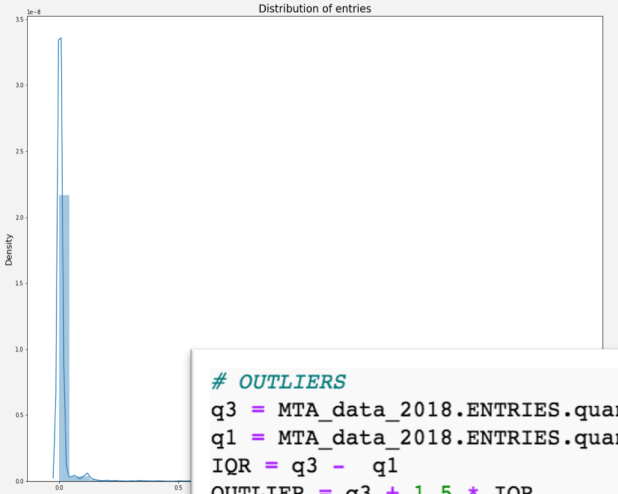
(negative values, huge values)

Data Cleaning

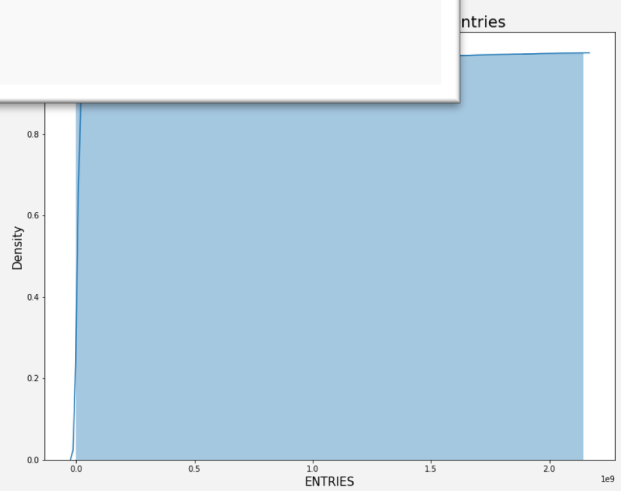




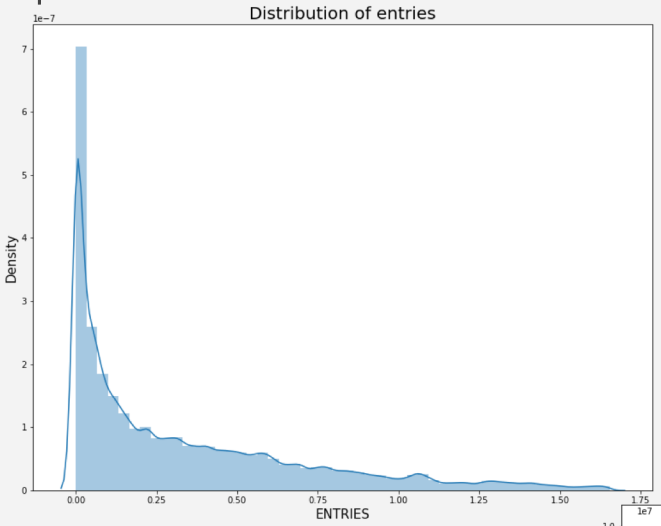
Data Cleaning

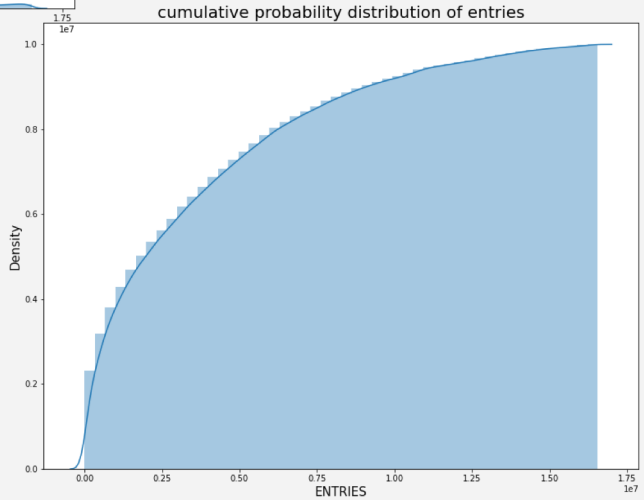



```python
# OUTLIERS
q3 = MTA_data_2018.ENTRIES.quantile(q=0.75)
q1 = MTA_data_2018.ENTRIES.quantile(q=0.25)
IQR = q3 -  q1
OUTLIER = q3 + 1.5 * IQR
OUTLIER
```

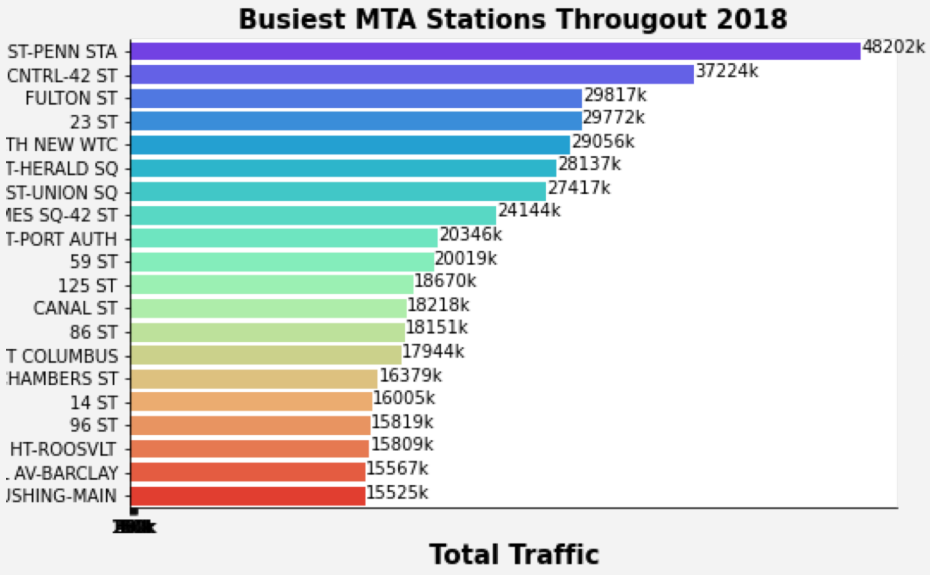


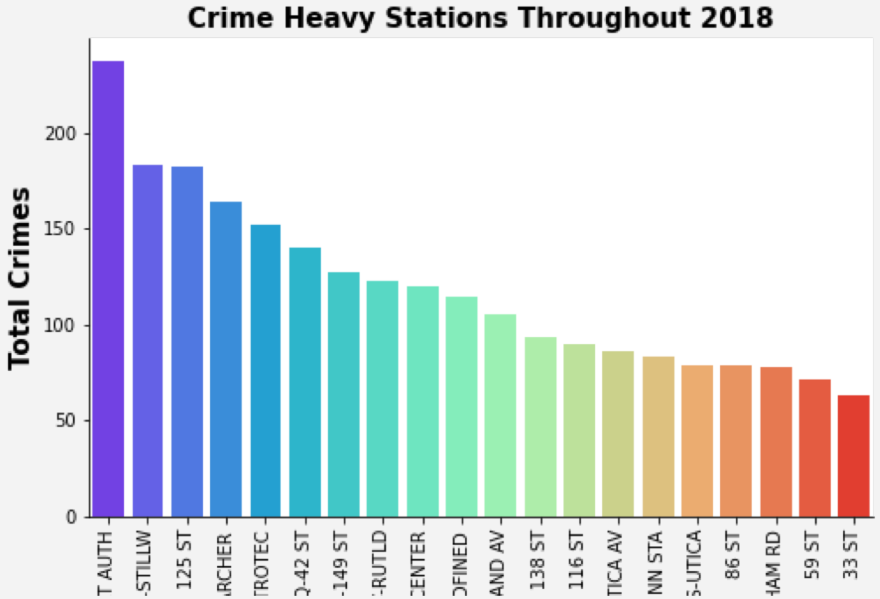
Data Cleaning



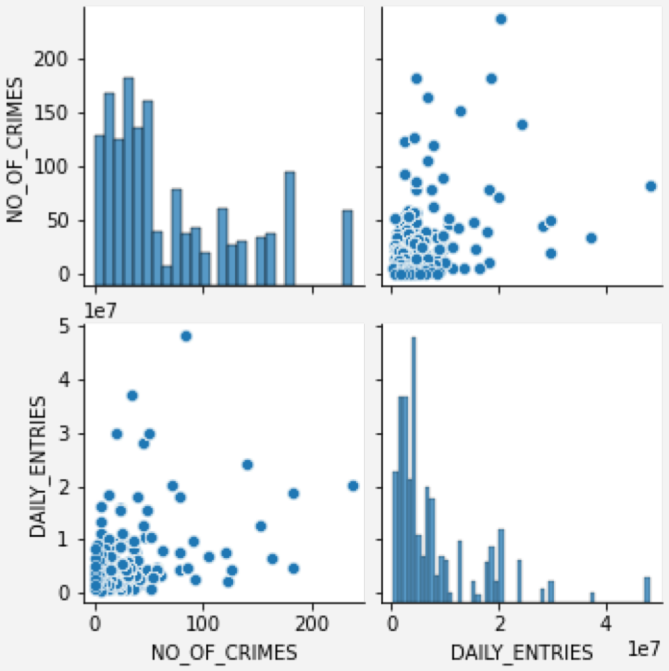


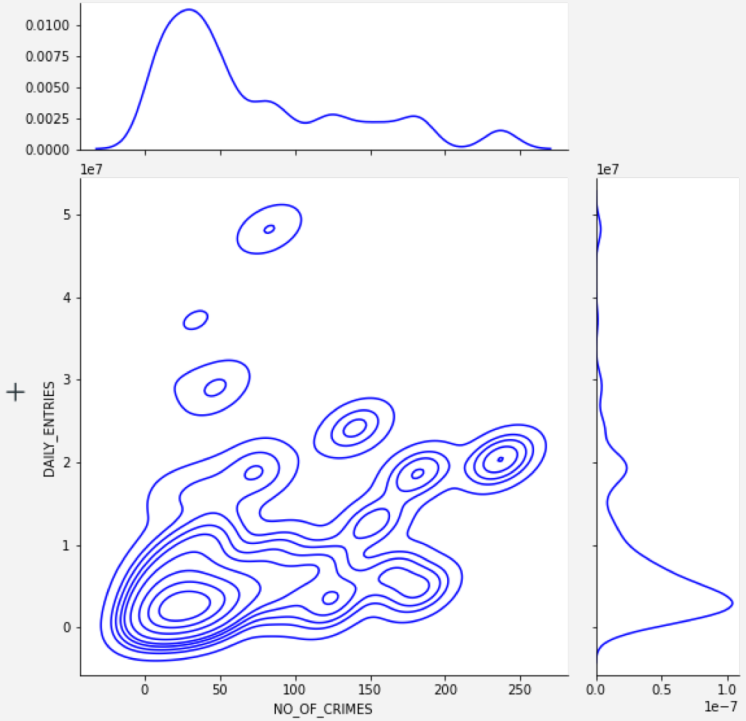
ANALYSIS


Busiest MTA Stations Througout 2018
ST-PENN STA
CNTRL-42 ST
FULTON ST
23 ST
TH NEW WTC
T-HERALD SQ
ST-UNION SQ
IES SQ-42 ST
T-PORT AUTH
59 ST
125 ST
CANAL ST
86 ST
T COLUMBUS
HAMBERS ST
14 ST
96 ST
HT-ROOSVLT
. AV-BARCLAY
JSHING-MAIN
48202k
37224k
29817k
29772k
29056k
28137k
27417k
24144k
20346k
20019k
18670k
18218k
18151k
17944k
16379k
16005k
15819k
15809k
15567k
15525k
Total Traffic

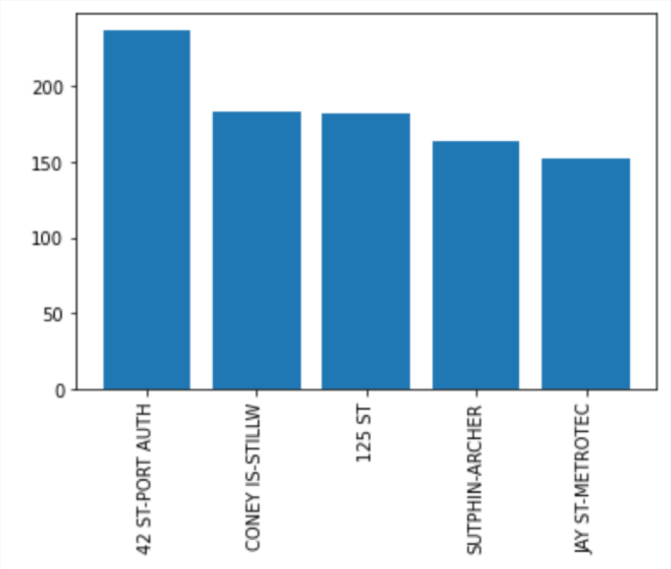

Crime Heavy Stations Throughout 2018
Total Crimes
0
50
100
150
200
T AUTH
-STILLW
125 ST
RCHER
TROTEC
Q-42 ST
-149 ST
-RUTLD
CENTER
DFINED
AND AV
138 ST
116 ST
TICA AV
NN STA
S-UTICA
86 ST
AM RD
59 ST
33 ST

CORRELATION

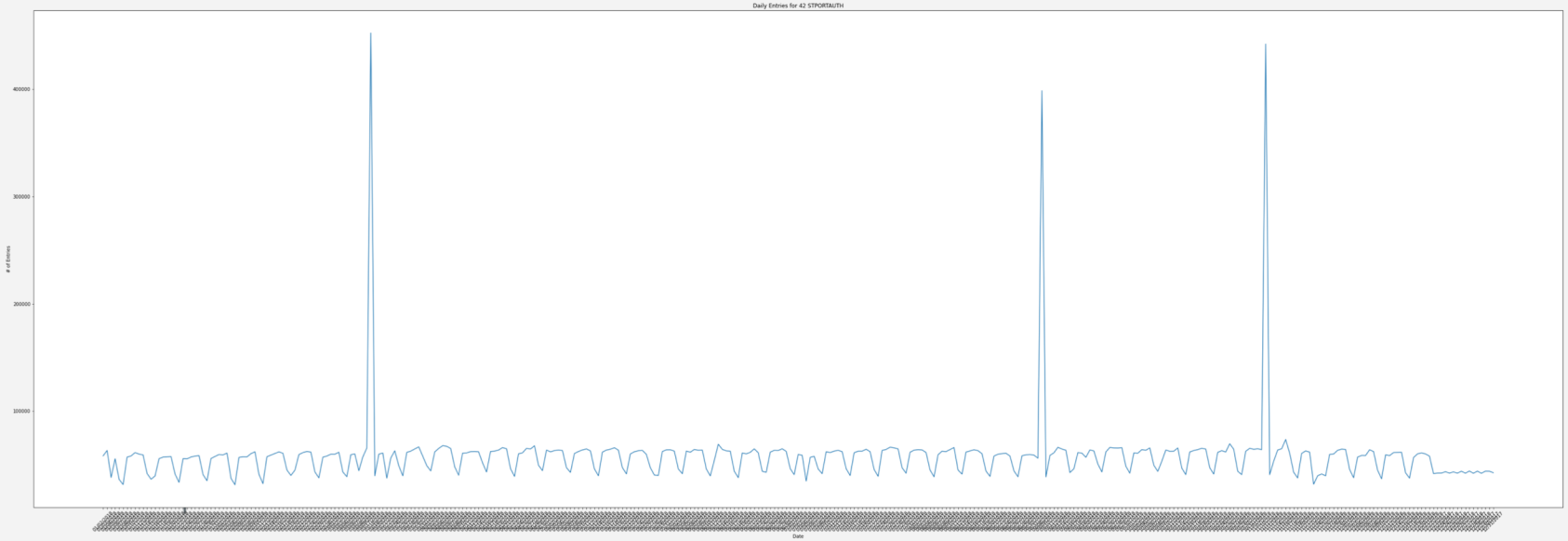




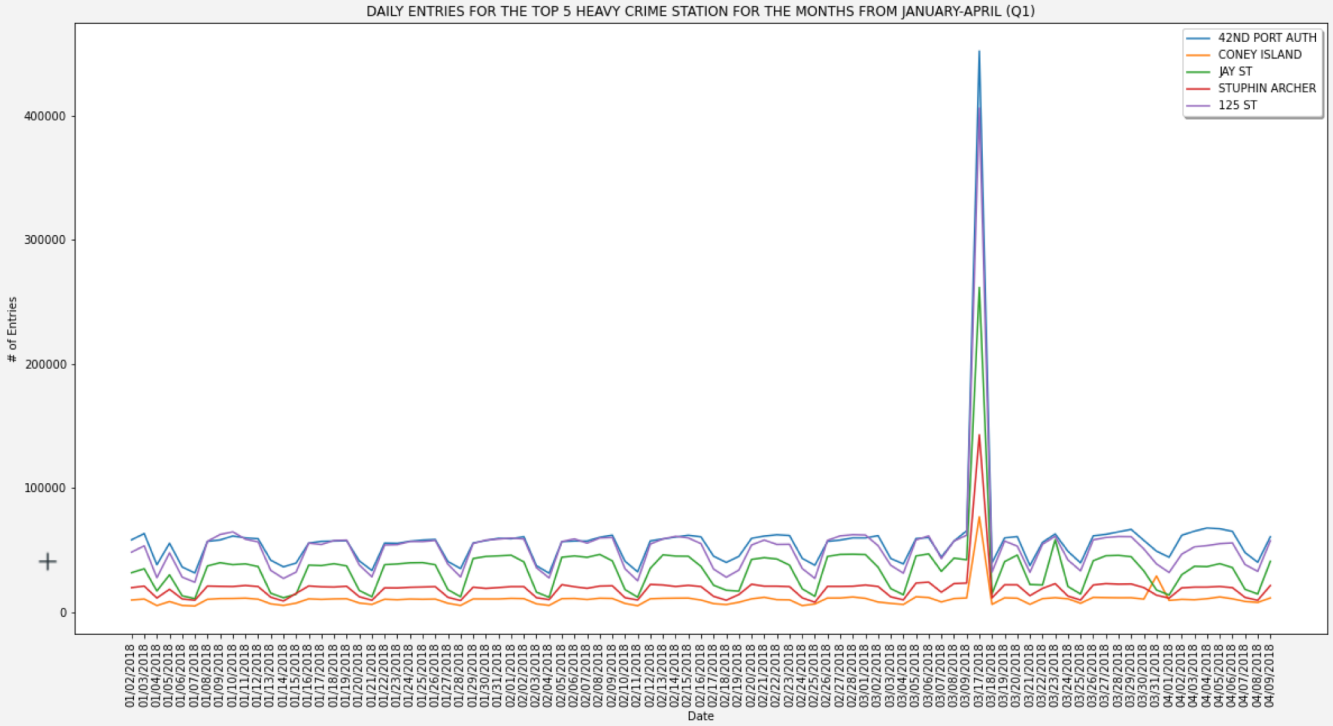
ANALYSIS OF THE TOP HEAVY CRIME STATIONS



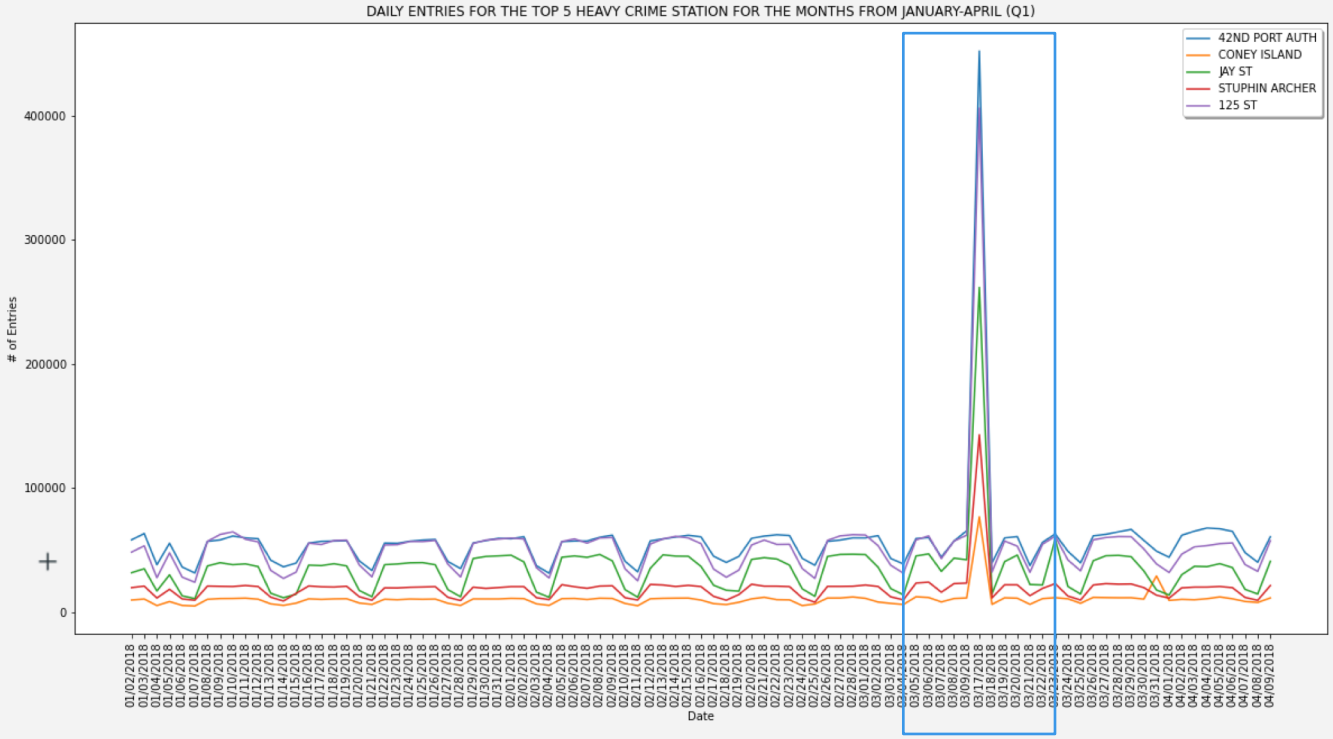
ANALYSIS OF THE TOP HEAVY CRIME STATIONS



ANALYSIS OF THE TOP HEAVY CRIME STATIONS



ANALYSIS OF THE TOP HEAVY CRIME STATIONS



ANALYSIS OF THE TOP HEAVY CRIME STATIONS

MERGING TWO DATASETS

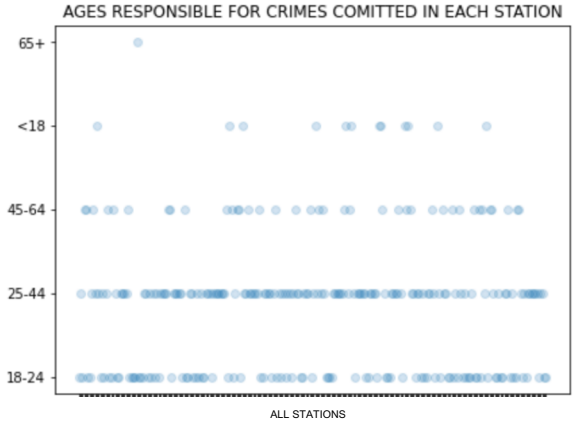
	ARREST_DATE	ARREST_BORO	AGE_GROUP	PERP_SEX	PERP_RACE	STATION	notes	NO_OF_CRIMES	DAILY_ENTRIES
0	2018-12-28	M	25-44	M	WHITE HISPANIC	UNION ST	L-all times	15	2082203.0
1	2018-12-27	M	25-44	M	WHITE	UNION ST	L-all times	15	2082203.0
2	2018-10-27	M	25-44	M	WHITE HISPANIC	UNION ST	L-all times	15	2082203.0
3	2018-10-18	M	45-64	M	WHITE	UNION ST	N,Q-all times, R-all times exc nights	15	2082203.0
4	2018-08-31	M	18-24	F	WHITE HISPANIC	UNION ST	L-all times	15	2082203.0
5	2018-08-31	M	45-64	M	WHITE HISPANIC	UNION ST	L-all times	15	2082203.0
6	2018-08-29	M	25-44	M	WHITE	UNION ST	L-all times	15	2082203.0
7	2018-08-05	M	18-24	M	WHITE HISPANIC	UNION ST	L-all times	15	2082203.0
8	2018-07-29	M	18-24	M	ASIAN / PACIFIC ISLANDER	UNION ST	L-all times	15	2082203.0
9	2018-07-22	M	25-44	M	BLACK	UNION ST	L-all times	15	2082203.0

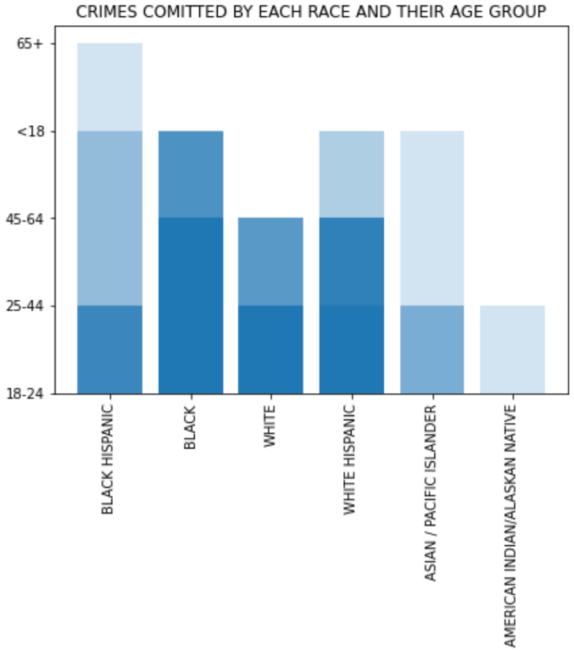
OBSTECLES

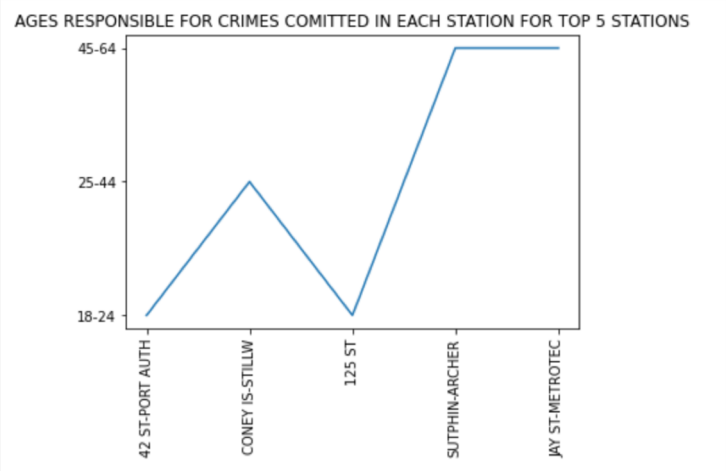
- MORE VALUES FOR MEANINGFUL CORRELATION

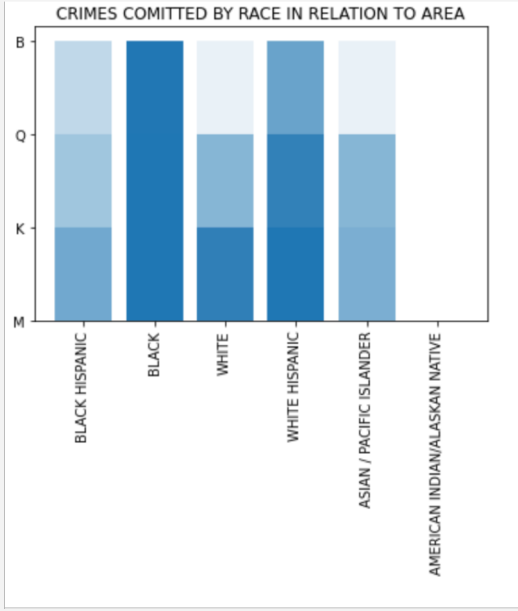
- DIFFERENT VALUES WHILE MERGING WILL RESULT IN EMPTY DATAFRAME

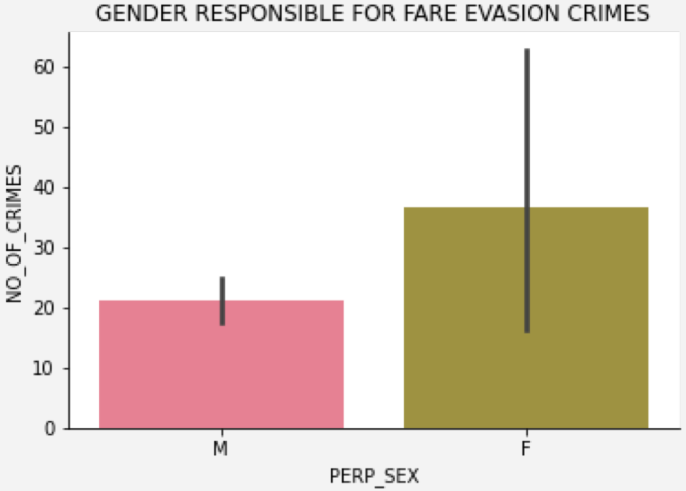
INSIGHTS


AGES RESPONSIBLE FOR CRIMES COMITTED IN EACH STATION
65+
<18
45-64
25-44
18-24
ALL STATIONS


CRIMES COMITTED BY EACH RACE AND THEIR AGE GROUP
65+
<18
45-64
25-44
18-24
BLACK HISPANIC
BLACK
WHITE
WHITE HISPANIC
ASIAN / PACIFIC ISLANDER
AMERICAN INDIAN/ALASKAN NATIVE


AGES RESPONSIBLE FOR CRIMES COMITTED IN EACH STATION FOR TOP 5 STATIONS
45-64
25-44
18-24
42 ST-PORT AUTH
CONEY IS-STILLW
125 ST
SUTPHIN-ARCHER
JAY ST-METROTEC


CRIMES COMITTED BY RACE IN RELATION TO AREA
B
Q
K
M
BLACK HISPANIC
BLACK
WHITE
WHITE HISPANIC
ASIAN / PACIFIC ISLANDER
AMERICAN INDIAN/ALASKAN NATIVE


GENDER RESPONSIBLE FOR FARE EVASION CRIMES
NO_OF_CRIMES
0
10
20
30
40
50
60
M
F
PERP_SEX

CONCLUSION

- There is a correlation between traffic and fare evasion crimes
- Holidays are a big cause of fare evasion crimes
- Almost everyone commits fare evasion crime
- Heavy crime stations will continue to grow

RECOMMENDATIONS

01

DISTIRBUTING MANPOWER (PLOICE OFFICERS) DURING THE **HOLIDAYS**

02

SURVAILING THE ALREADY **KNOWN STATIONS FOR CRIME**

03

DISTIRBUTING MANPOWER (PLOICE OFFICERS) DURING THE **WEEKENDS**

THANKS!

ANY QUESTIONS?