FILM PRODUCTION IN THE SUBWAY

USING MTA TURNSTILE DATA

FAISAL MIRZA

OUTLINE

- INTRODUCTION
- DATASET
- DATA VALIDATION
- DATA VISUALIZATION

INTRODUCTION

- Scenario: Helping a film production company in finding the best place to film a movie in the subway, where it has the least amount of people.

- Goal: Finding subway stations with the least amount of traffic

DATASET

- Data used: NY MTA turnstile data

- Time period: from March 17 2018 - June 22 2018

- Number of rows: 2,756,607

- Number of columns: 11

DATA VALIDATION

ASSUMPTIONS:

- 0 daily entries means construction days
- Negative turnstile counter is a reverse counter, use absolute value

FINDINGS:

- Actual period was from 3/10/2018 to 6/14/2018
- There are no NA, NaN values

DATA VALIDATION

FINDINGS:

- There are 5 duplicate values

- There are a lot or rows where the daily entries was 0

DATA CLEANING:

- Removed 5 duplicate values

- Removed 7781 rows where the DESC column equals RECOVER AUD

- Removed rows where daily entries = 0

DATA VISUALIZATION

Finding stations with lowest entries:
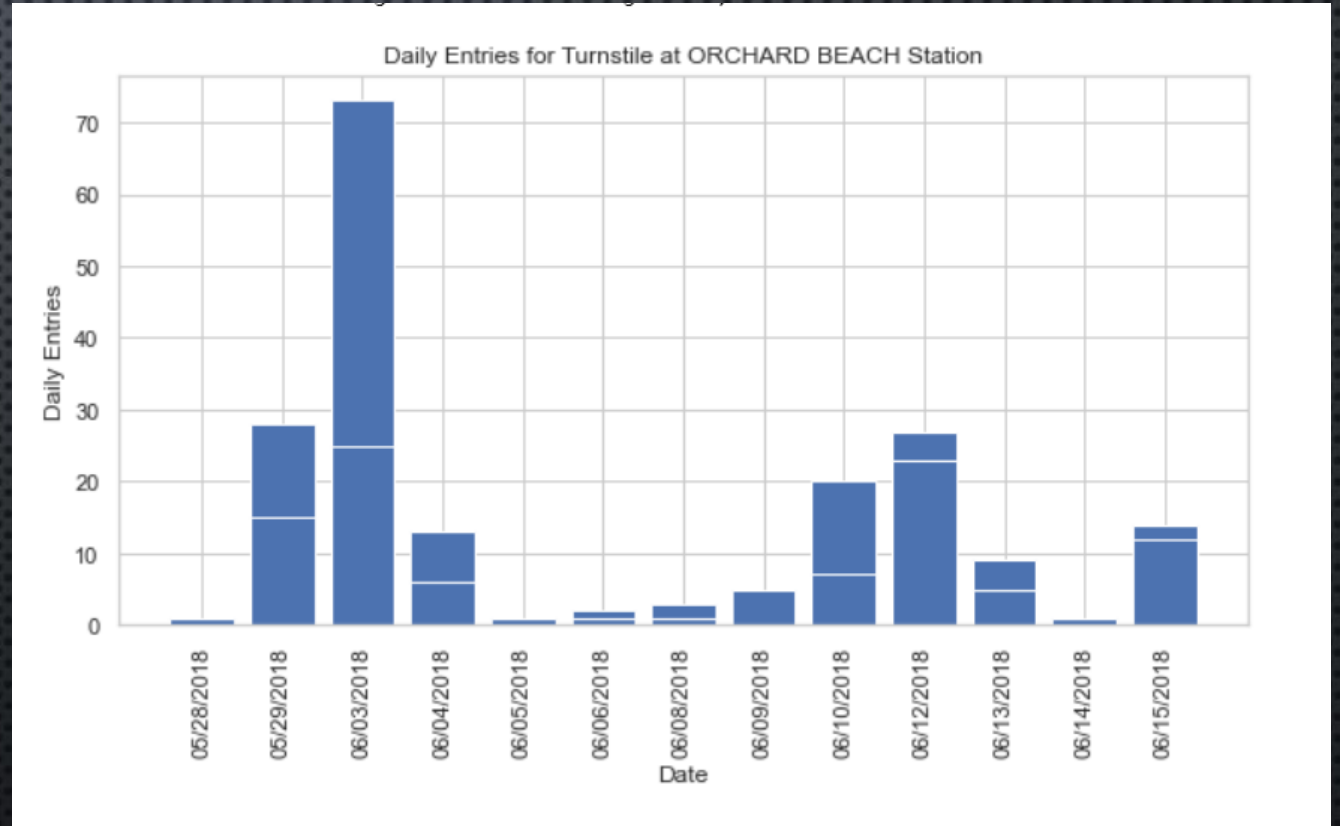
- Lowest daily entree stations by median:

1. ORCHARD BEACH

2. BEACH 105 ST

3. BROAD CHANNEL

STATION	DAILY_ENTRIES
ORCHARD BEACH	7.0
BEACH 105 ST	45.0
BROAD CHANNEL	75.0
75 AV	77.0
AVENUE I	80.5
E 143/ST MARY'S	90.5
14TH STREET	96.0
NEWARK HM HE	104.5
NEWKIRK AV	108.5
AQUEDUCT N.COND	112.0

DATA VISUALIZATION

ORCHARD BEACH
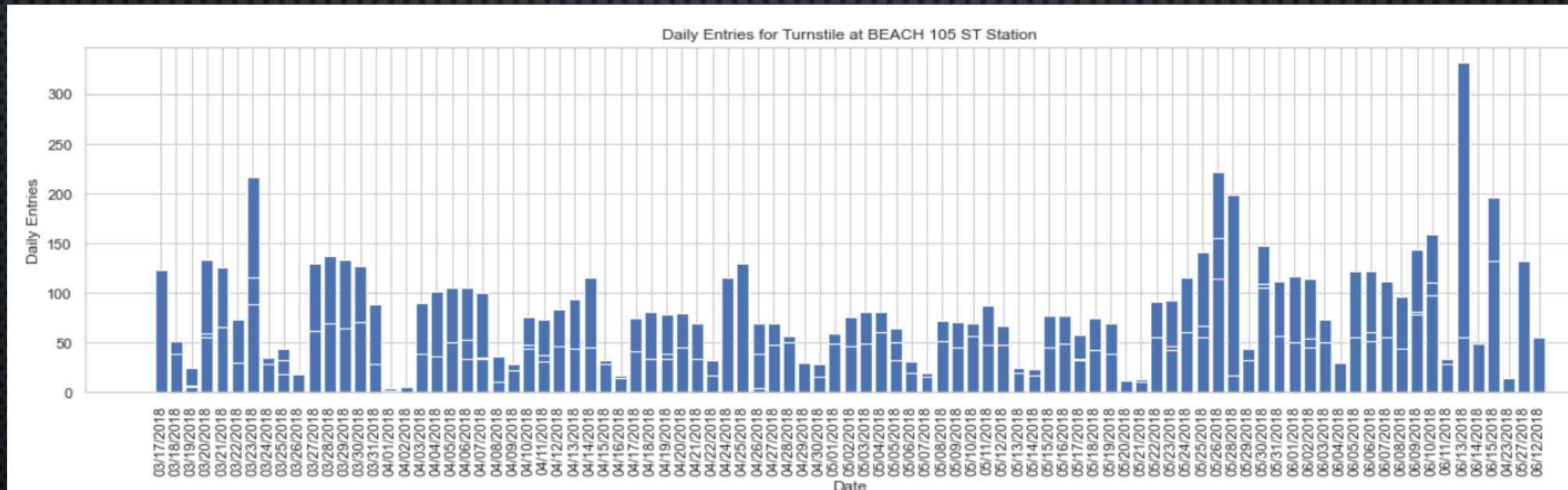
- Most daily entries are under 30

- June 3 2018 had 75 entries but still considered low
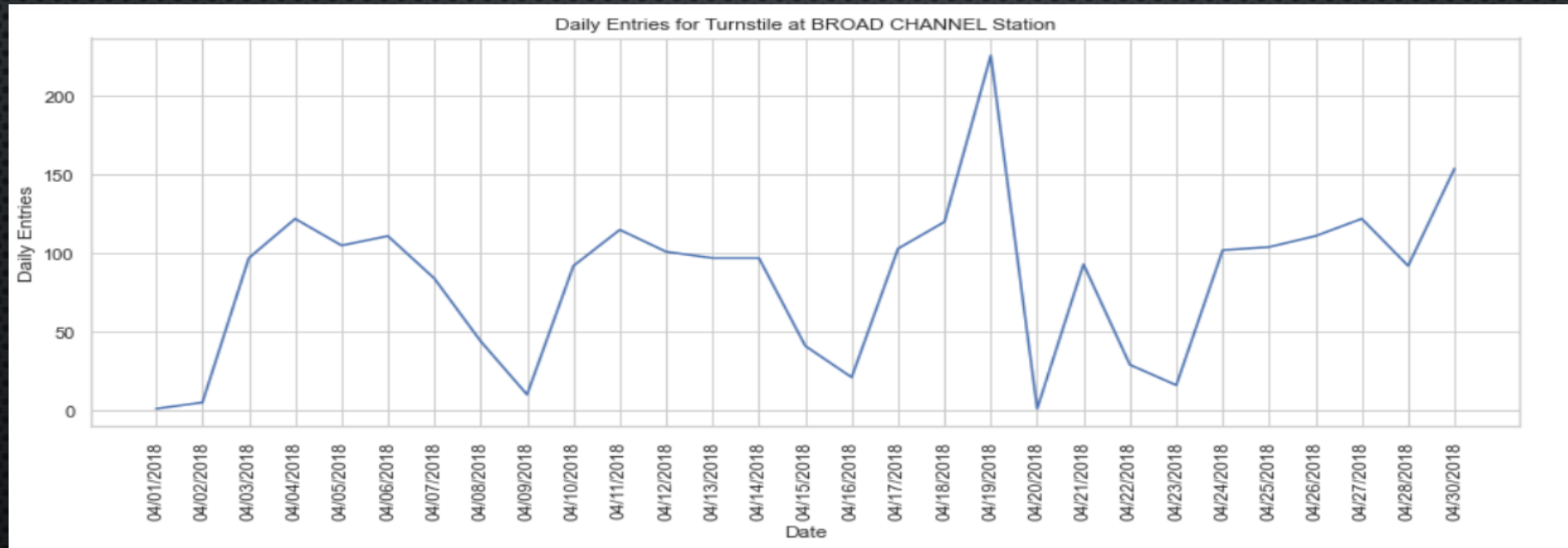
- Overall a low traffic station

Daily Entries for Turnstile at ORCHARD BEACH Station

DATA VISUALIZATION

BEACH 105 ST

- Most daily entries are around 100
- best days are from 3/31/2018 to 5/23/2018
- Overall a low traffic station

Daily Entries for Turnstile at BEACH 105 ST Station

DATA VISUALIZATION

BROAD CHANNEL

- Most daily entries are around 100
- The whole of April is a good time
- Overall a low traffic station

Daily Entries for Turnstile at BROAD CHANNEL Station

THANK YOU