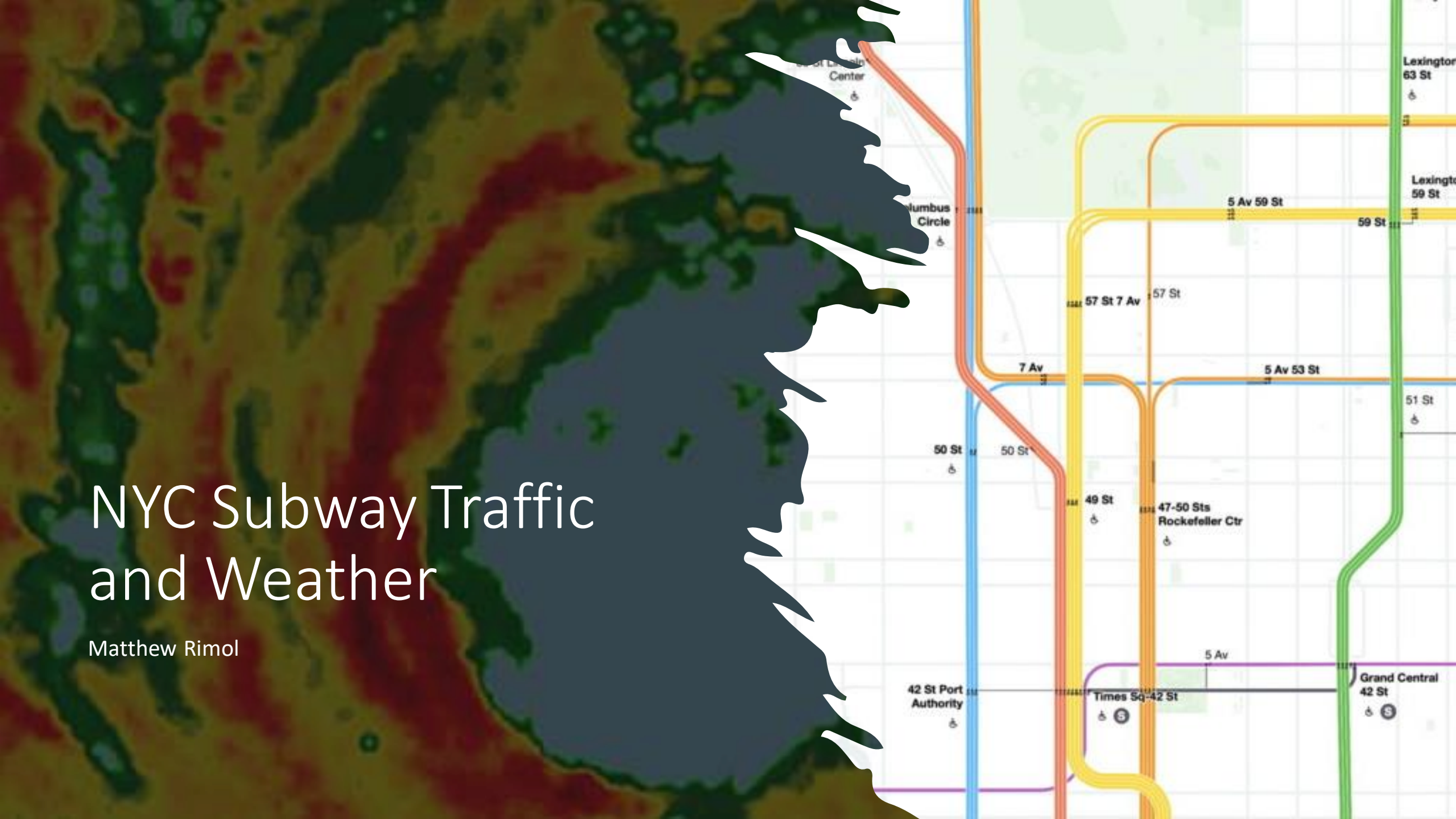
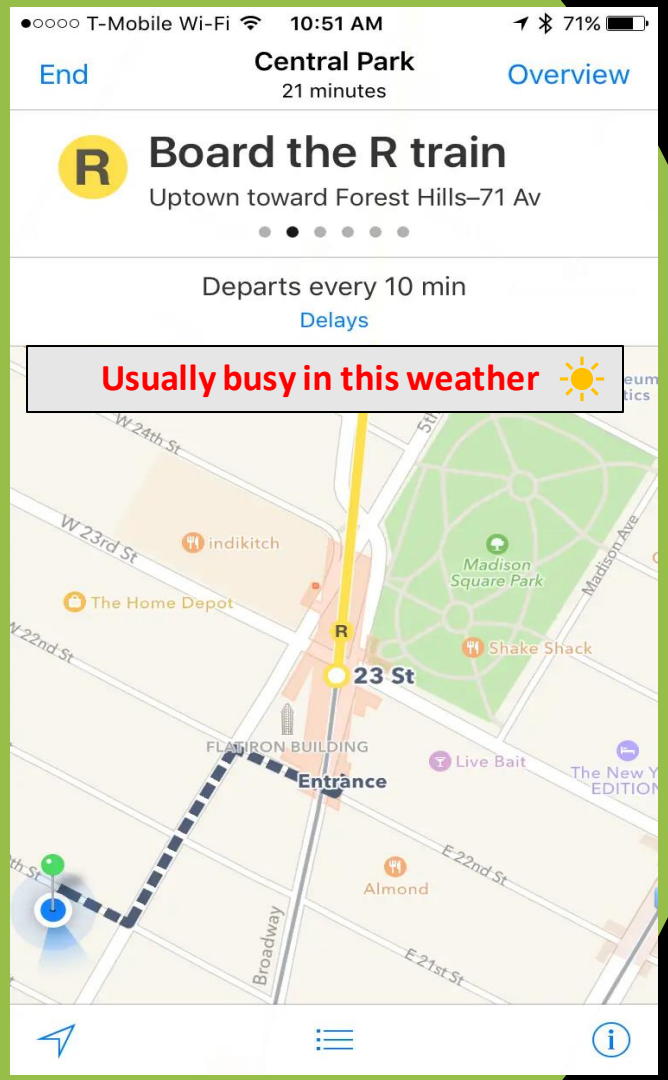
NYC Subway Traffic and Weather

Matthew Rimol





- Explore effects on subway traffic:
  - Temperature
  - Precipitation
  - Snowfall
- How do different stations respond?
- Characteristics of the relationship

MTA Data
(NYC MTA)

SQLite

SQLite3 DB

SQLite

NYC Weather Data
(Kaggle via National Weather Service)

pandas

pandas

jupyter

SQLAlchemy

pandas

	date	max_temp	min_temp	avg_temp	precipitation	snowfall	snow_depth
0	1-1-2016	42	34	38.0	0	0	0
1	2-1-2016	40	32	36.0	0	0	0
2	3-1-2016	45	35	40.0	0	0	0
3	4-1-2016	36	14	25.0	0	0	0
4	5-1-2016	29	11	20.0	0	0	0

pandas

	CA	UNIT	SCP	STATION	LINENAME	DIVISION	DATE	TIME	DESC	ENTRIES	EXITS
0	A002	R051	02-00-00	59 ST	NQR456	BMT	12/26/2015	03:00:00	REGULAR	5469540	1846630
1	A002	R051	02-00-00	59 ST	NQR456	BMT	12/26/2015	07:00:00	REGULAR	5469543	1846640
2	A002	R051	02-00-00	59 ST	NQR456	BMT	12/26/2015	11:00:00	REGULAR	5469617	1846739
3	A002	R051	02-00-00	59 ST	NQR456	BMT	12/26/2015	15:00:00	REGULAR	5469821	1846805
4	A002	R051	02-00-00	59 ST	NQR456	BMT	12/26/2015	19:00:00	REGULAR	5470180	1846874

pandas

	station	date	entries	daily_entries	max_temp	min_temp	avg_temp	precipitation	snowfall	snow_depth	day_of_week
2	1 AV	2016-01-02	4246496	3365.0	40	32	36.0	0.00	0.0	0.0	Saturday
3	1 AV	2016-01-03	4249815	3319.0	45	35	40.0	0.00	0.0	0.0	Sunday
9	1 AV	2016-01-09	4275815	4326.0	47	40	43.5	0.25	0.0	0.0	Saturday
10	1 AV	2016-01-10	4279327	3512.0	59	40	49.5	1.80	0.0	0.0	Sunday
16	1 AV	2016-01-16	4305378	4225.0	52	42	47.0	0.24	0.0	0.0	Saturday

matplotlib

seaborn

pandas

Data Ingestion

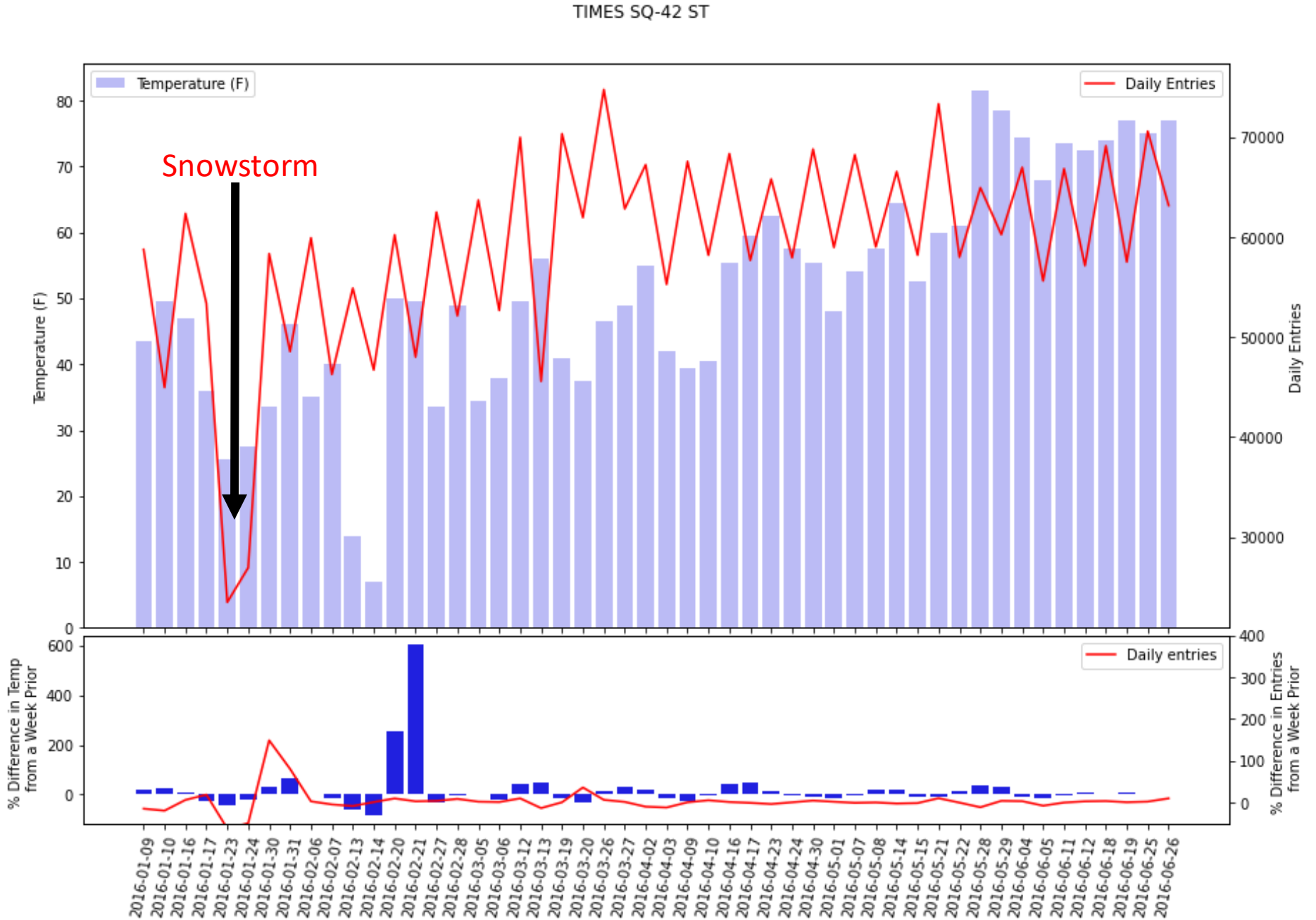
Initial Querying

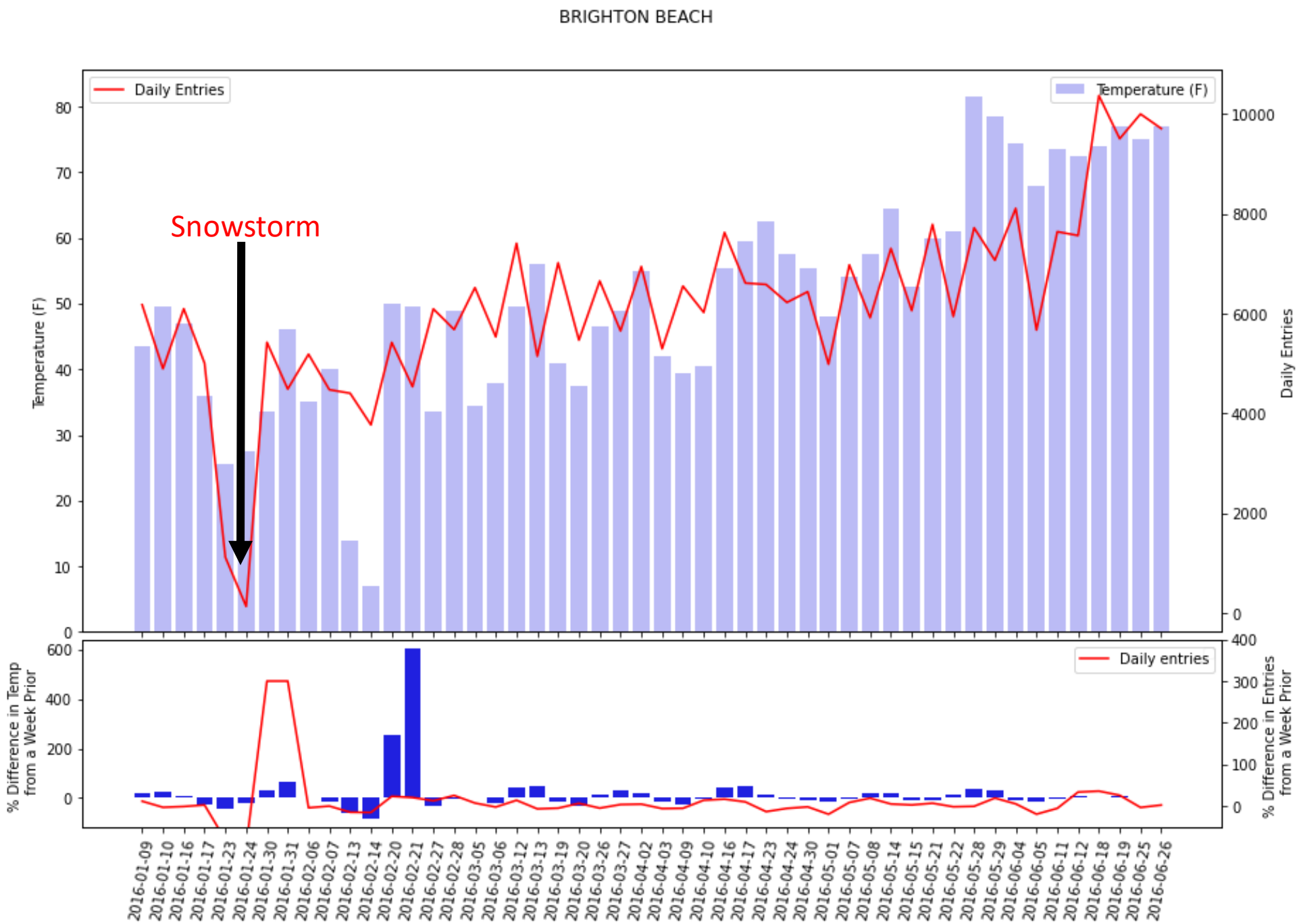
Data Cleaning

Analytics & Visuals

Analysis

- January 1 though June 30, 2016
- Weekends
- Temperature
- Stations more impacted by weather


TIMES SQ-42 ST
Temperature (F)
Daily Entries
Snowstorm
Temperature (F)
Daily Entries
Daily entries
% Difference in Temp from a Week Prior
% Difference in Entries from a Week Prior


BRIGHTON BEACH
Daily Entries
Temperature (F)
Snowstorm
Temperature (F)
Daily Entries
% Difference in Temp from a Week Prior
Daily entries
% Difference in Entries from a Week Prior
2016-01-09
2016-01-10
2016-01-16
2016-01-17
2016-01-23
2016-01-24
2016-01-30
2016-01-31
2016-02-06
2016-02-07
2016-02-13
2016-02-14
2016-02-20
2016-02-21
2016-02-27
2016-02-28
2016-03-05
2016-03-06
2016-03-12
2016-03-13
2016-03-19
2016-03-20
2016-03-26
2016-03-27
2016-04-02
2016-04-03
2016-04-09
2016-04-10
2016-04-16
2016-04-17
2016-04-23
2016-04-24
2016-04-30
2016-05-01
2016-05-07
2016-05-08
2016-05-14
2016-05-15
2016-05-21
2016-05-22
2016-05-28
2016-05-29
2016-06-04
2016-06-05
2016-06-11
2016-06-12
2016-06-18
2016-06-19
2016-06-25
2016-06-26

Findings

- Some stations (typically near beaches or parks) are much more sensitive to termperature changes
- Day of the week matters (even Saturday vs. Sunday)
- People do more when it is nice out but they still take the subway to get there

Future Work

- Focus on day of week
- Statistical analysis to quantify relationships at different stations
- Account for one time events (holidays, marathons, etc.)
- Regression Model