

Exploratory Data Analysis

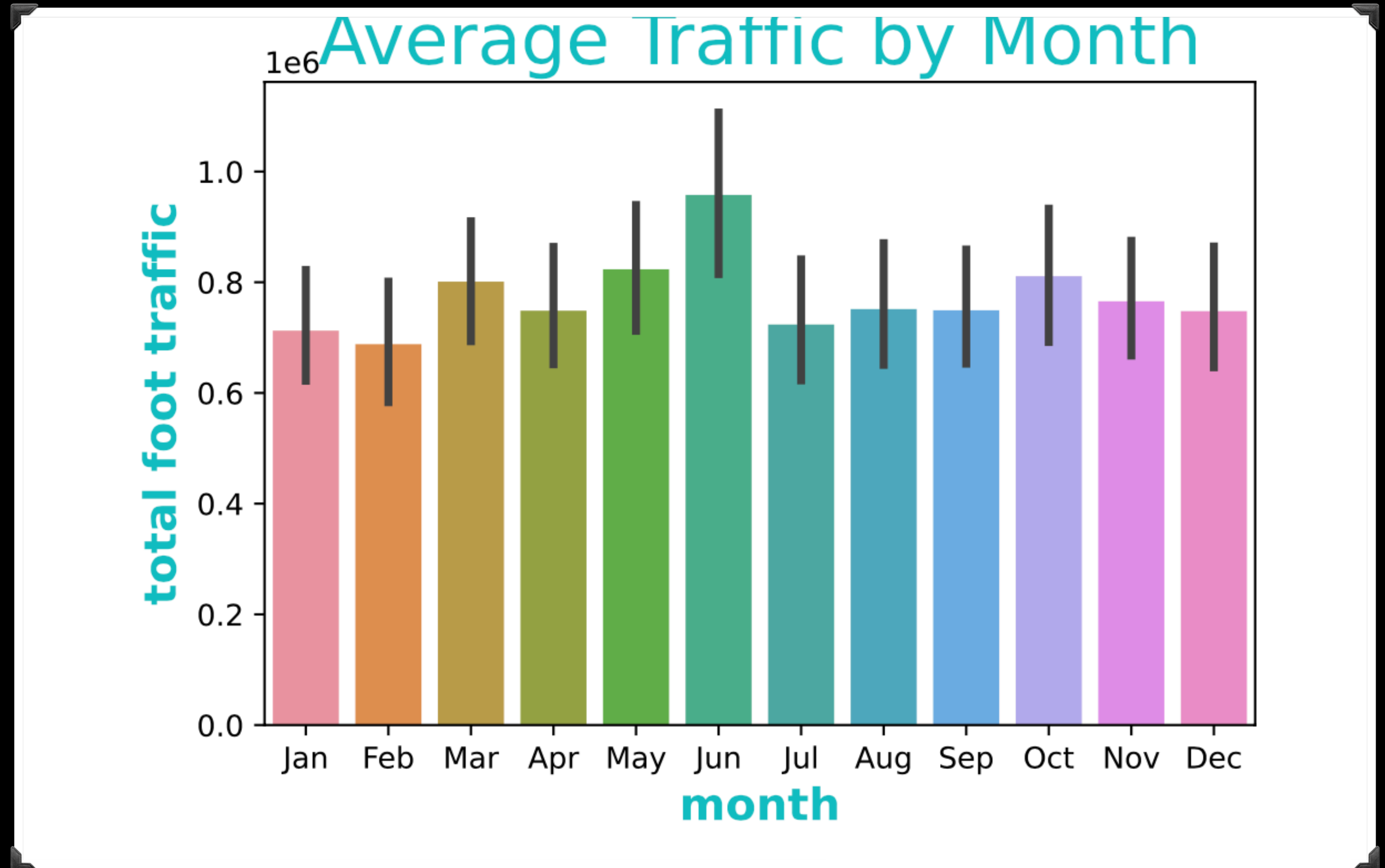
On MTA Transit Data

Proposal and Purpose

- Optimal windows for repairing and replacing turnstiles with least impact on commuters
- Identifying which months have lowest overall traffic
- Identifying stations with multiple turnstiles, allowing replacements to happen without significant impact



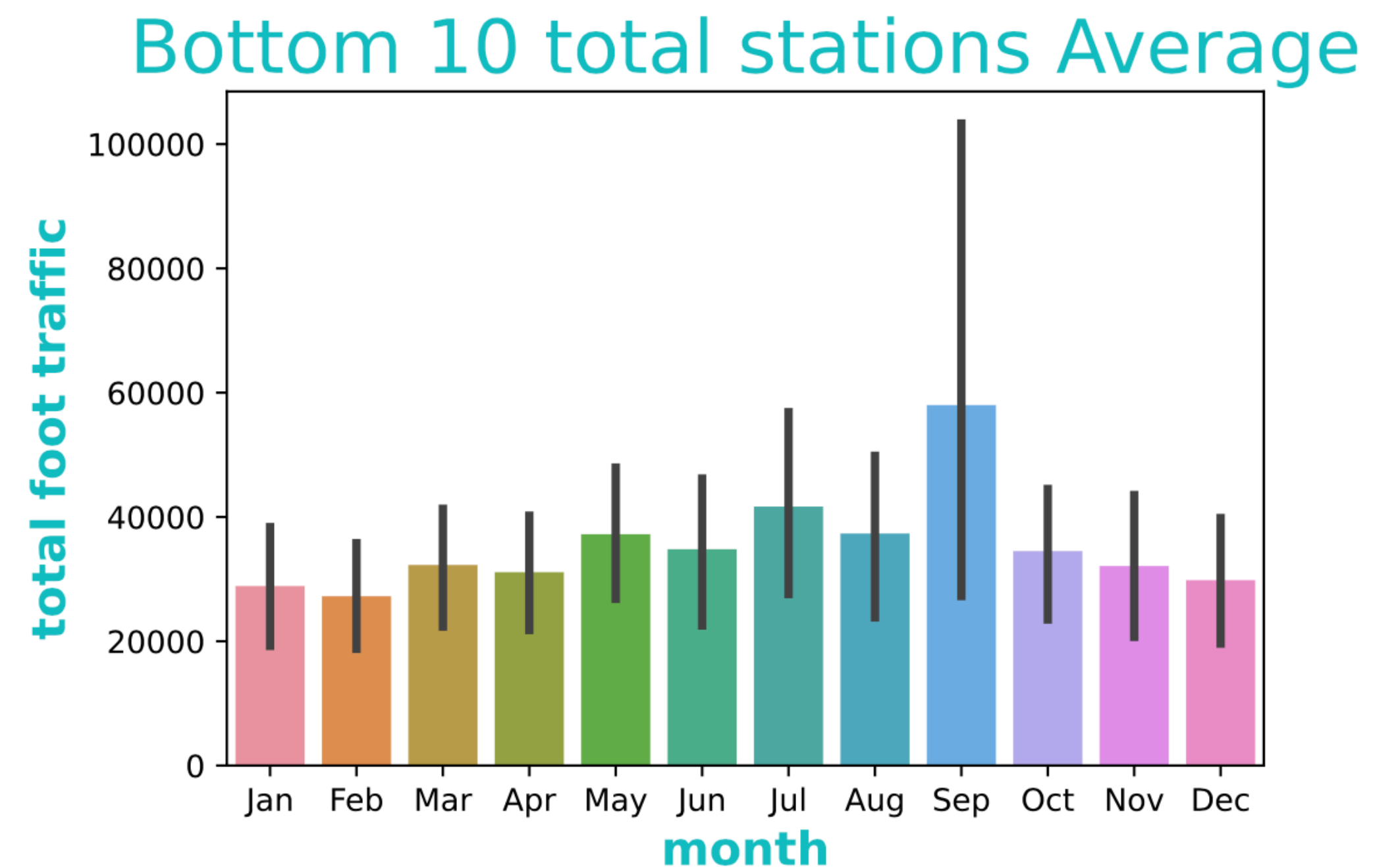
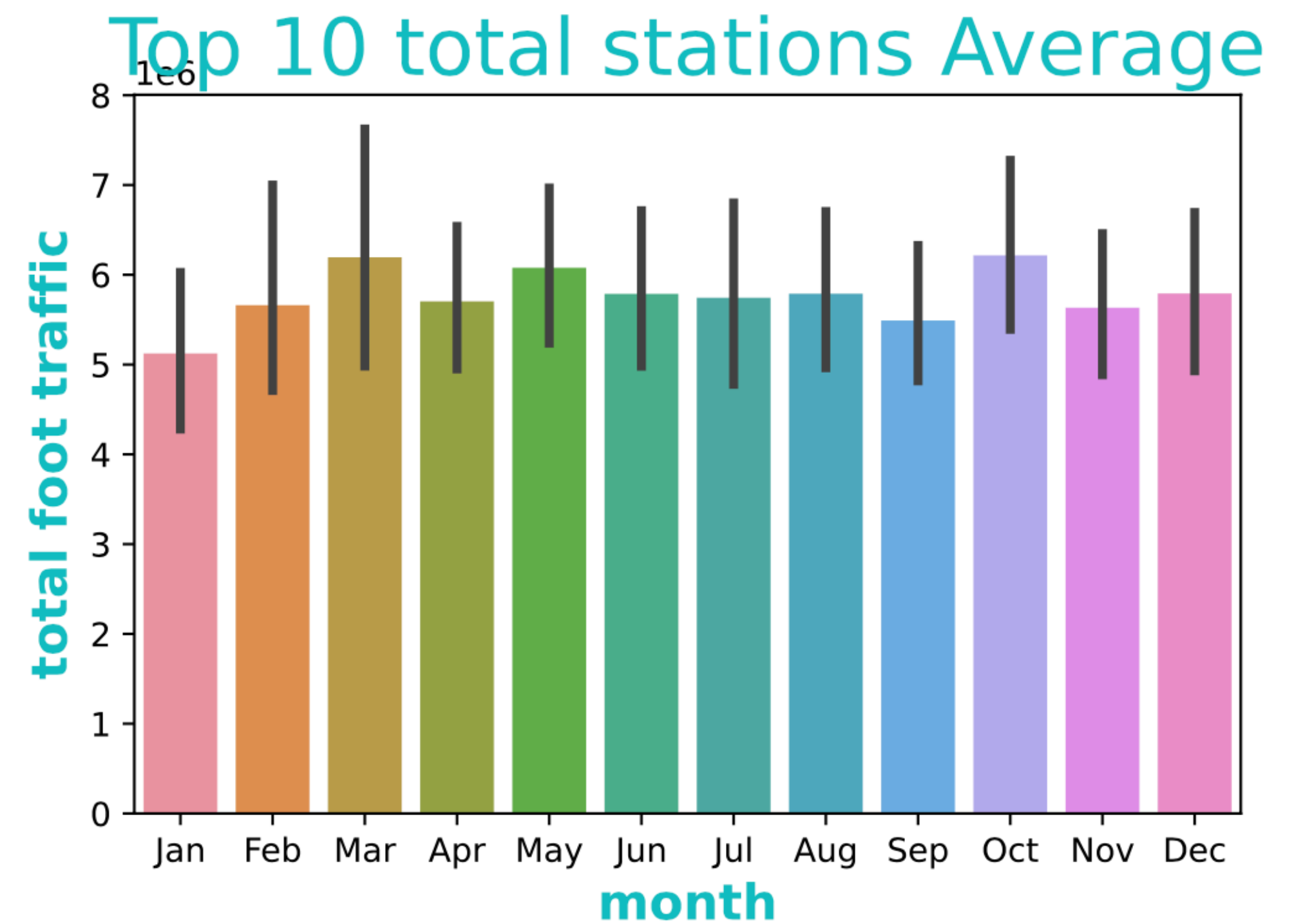
Preliminary Findings



- Graph shows the total average entries and exits by month for 2019

Comparing Trends

- Here we performed the same analysis for the top and bottom 10 stations, based on total yearly traffic.
- Best window seems to be Dec-Mar



Number of Turnstiles

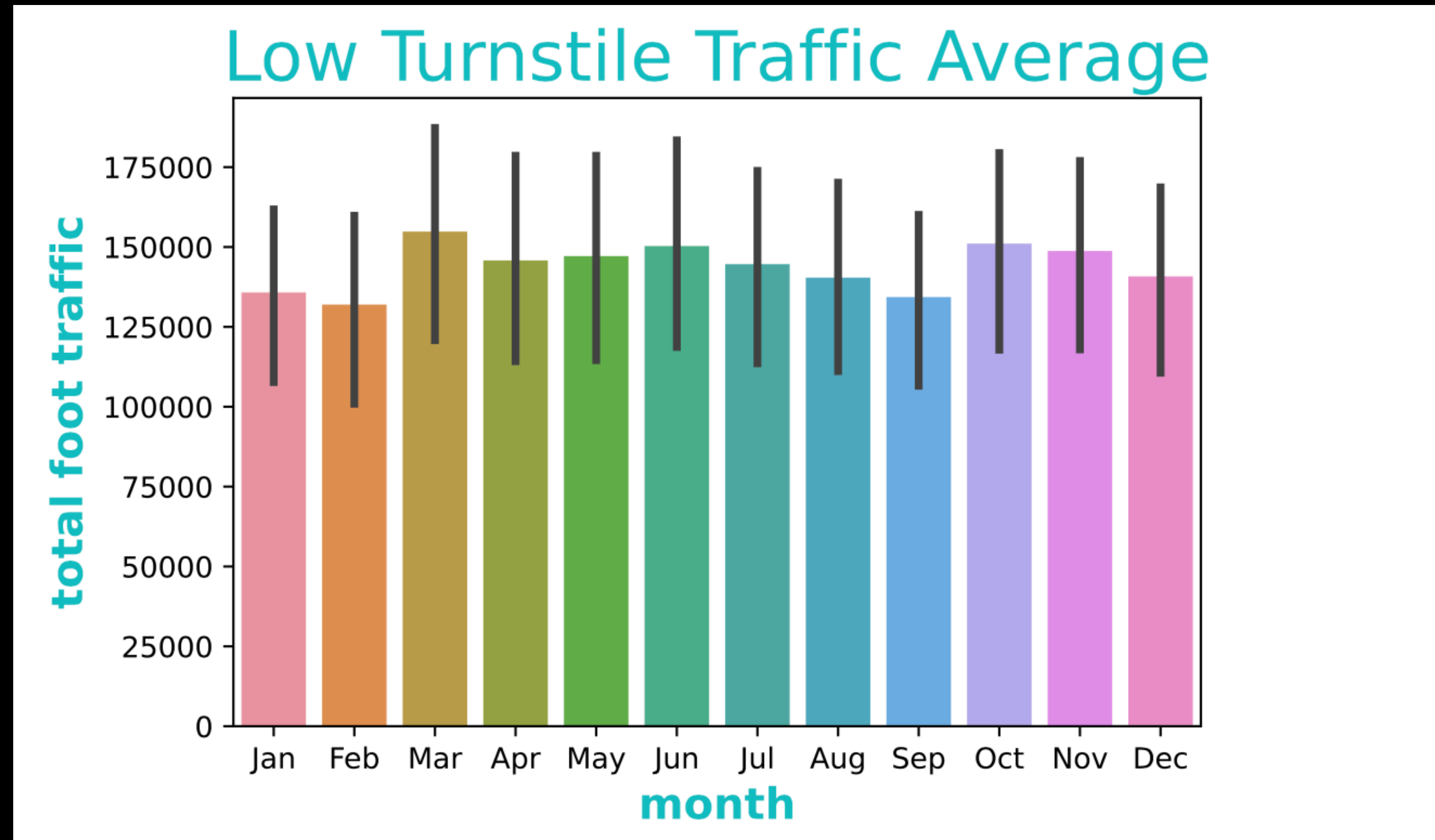
- Stark Difference in number of turnstiles for each station.

turnstile	
station	
34 ST-PENN STA	100
FULTON ST	96
GRD CNTRL-42 ST	74
23 ST	73
86 ST	63
CHAMBERS ST	59
CANAL ST	58
59 ST	57
34 ST-HERALD SQ	54
14 ST	53
42 ST-PORT AUTH	51
WALL ST	50
TIMES SQ-42 ST	50
ATL AV-BARCLAY	48
PATH NEW WTC	45
14 ST-UNION SQ	44
125 ST	43
59 ST COLUMBUS	41
28 ST	39
CORTLANDT ST	39

turnstile	
station	
CRESCENT ST	3
CYPRESS HILLS	3
BUSHWICK AV	3
NORWOOD AV	3
BURKE AV	3
NEW LOTS	3
225 ST	3
BROAD CHANNEL	3
NEREID AV	3
NEPTUNE AV	3
BOWERY	3
SHEPHERD AV	3
FOREST AVE	3
65 ST	3
BEVERLEY ROAD	3
MORRIS PARK	3
STERLING ST	3
219 ST	3
CLEVELAND ST	2
SUTTER AV	2

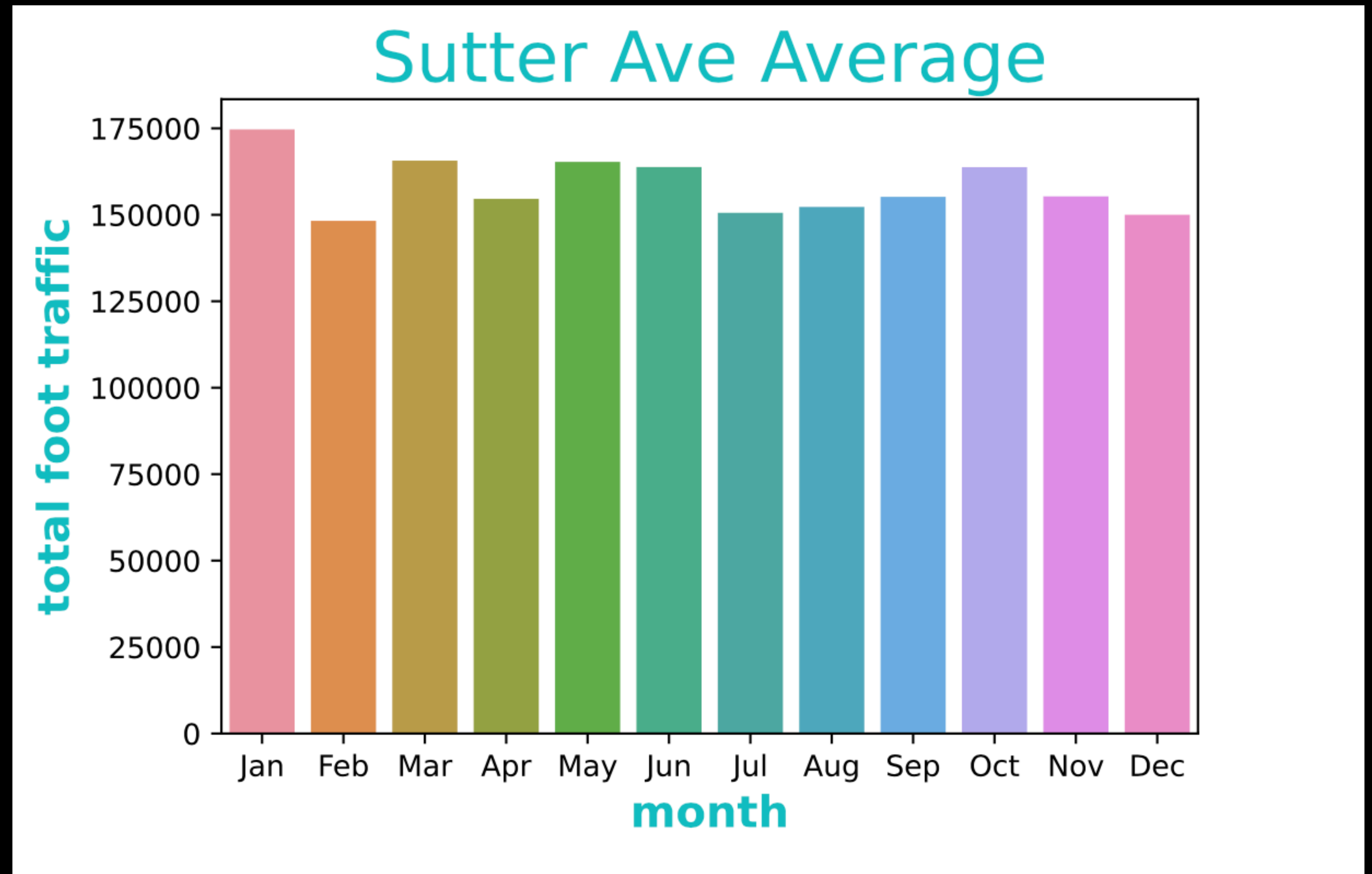
Low Turnstile Stations

- This graph has the same aggregation, but only using data from the lowest 20 station based on turnstile count.



Sutter Ave Totals

- Once again, same analysis done but on a station level.



Future Work

- Look at turnstiles by control area opposed to station.
- Complete same analysis on months found desirable, then week, then day.
- Isolate specific times for each station, opposed to average.
- Check with other years to ensure selected window is optimal.

Things that did not work

375]:

irnstile	datetime	station	starting_entries	starting_exits	jan_entries	jan_exits	march_entries	march_exits	feb_entries	feb_exits
A033 R170 02-00-00	2019-01-01	14 ST- UNION SQ	2058504	5184532	31568	77486	2117780	5334607	2086212	5257121
A033 R170 02-00-01	2019-01-01	14 ST- UNION SQ	1810289	2188156	56087	60203	1916183	2305572	1860096	2245369
A033 R170 02-00-02	2019-01-01	14 ST- UNION SQ	14202557	13711236	67574	73231	14331047	13850926	14263473	13777695
A033 R170 02-00-03	2019-01-01	14 ST- UNION SQ	6077136	7085828	95046	120544	6256372	7314371	6161326	7193827
A033 R170 02-00-04	2019-01-01	14 ST- UNION SQ	4939956	3757483	136620	105953	5196122	3959878	5059502	3853925
...
TRAM2 R469 00-00-01	2019-01-01	RIT- ROOSEVELT	2876975	33342	41667	517	2944684	34394	2903017	33877
TRAM2 R469 00-03-00	2019-01-01	RIT- ROOSEVELT	635516	24341	5719	293	647480	24914	641761	24621

[illegible]