Toronto Bike Share System

Introduction

- While bike share has similarities to traditional bike rental, it has many advantages as well.
- The sturdy-framed bikes can be taken from any station and returned to any station in the bike share system for up to 30 minutes.
- Bike share is perfect for short trips around town for both locals and visitors as an enjoyable and cost-effective option to walking, taxis, and public transportation.
- For this project I looked trip volume changes over time. I used an autoregression model to predict future trip volumes for three stations in the system, one with highest volume, one with lowest volume and one with most round trips.

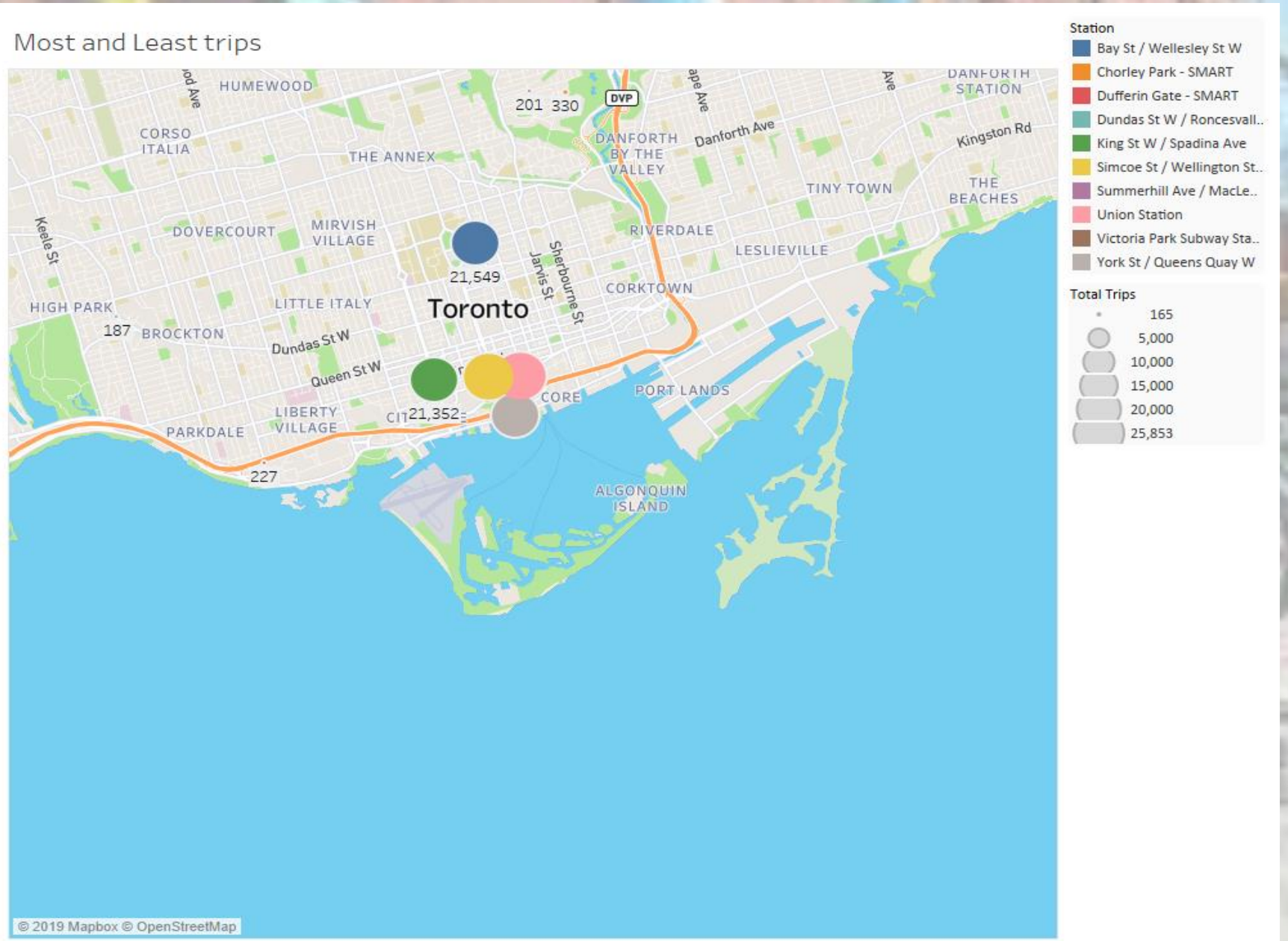
Methodology

- Ride trip data is from the City of Toronto's open data portal. The data is from Q4 of 2016 to Q4 of 2017 from a set of Excel and CSV files. The data contains columns for trip ID, trip start time, trip end time, trip duration, starting station, ending station and whether the user was a member.
- The data was preprocessed to get number of trips originating from each station and number of round trips for each station.
- Latitude and longitude data was obtained from an external source to plot location in a Tableau workbook.
- Time series data for three stations was used to create autoregression models for the station with most trips, the one with least trips, and the one with most round trips.

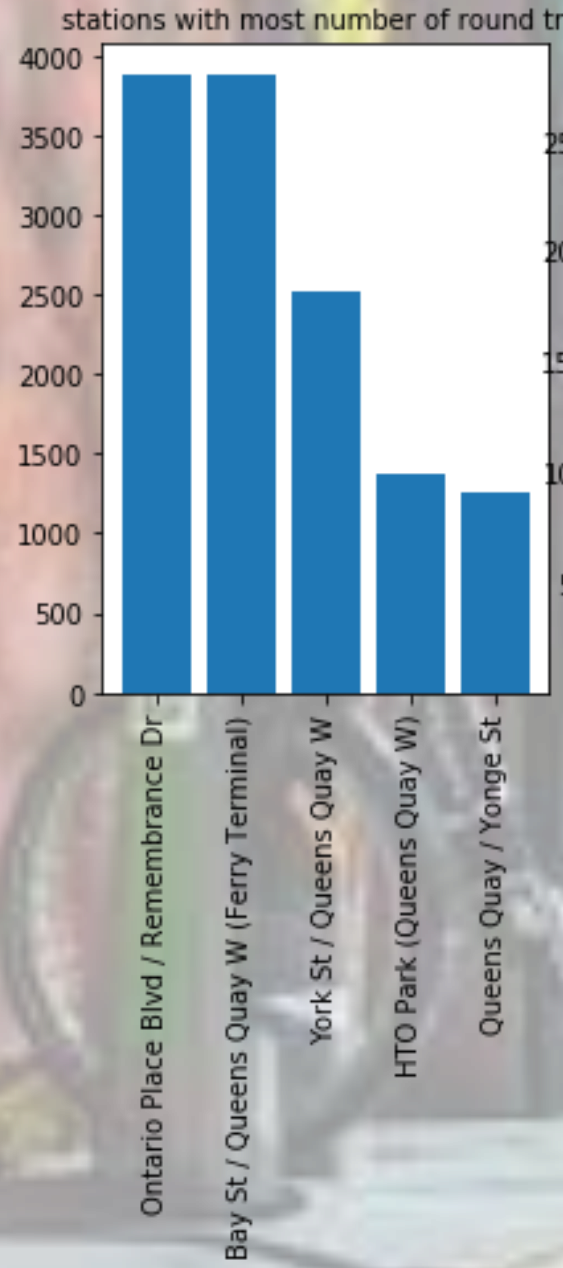
Bikeshare Data Source

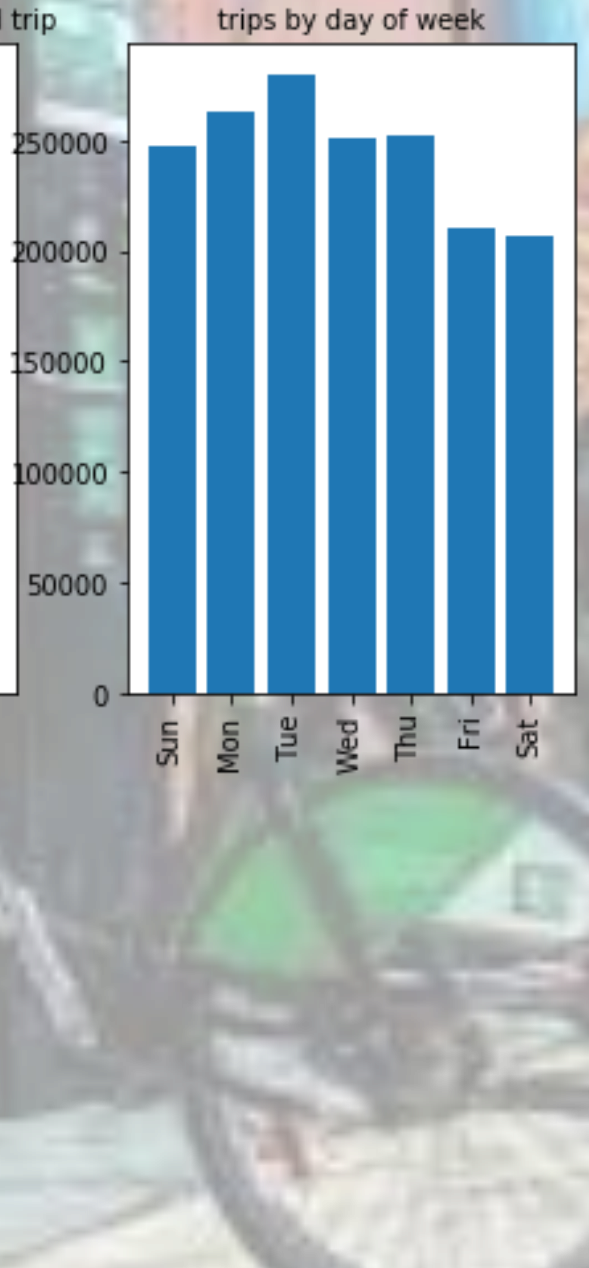
	trip_id	trip_start_time	trip_stop_time	trip_duration_seconds	from_station_name	to_station_name	user_type
0	462305	2016-01-10 00:00:00	2016-01-10 00:07:00	394	Queens Quay W / Dan Leckie Way	Fort York Blvd / Garrison Rd	Casual
1	462306	2016-01-10 00:00:00	2016-01-10 00:09:00	533	Sherbourne St / Wellesley St	Edward St / Yonge St	Member
2	462307	2016-01-10 00:00:00	2016-01-10 00:07:00	383	Queens Quay W / Dan Leckie Way	Fort York Blvd / Garrison Rd	Casual
3	462308	2016-01-10 00:01:00	2016-01-10 00:27:00	1557	Cherry St / Distillery Ln	Fort York Blvd / Capreol Crt	Casual
4	462309	2016-01-10 00:01:00	2016-01-10 00:27:00	1547	Cherry St / Distillery Ln	Fort York Blvd / Capreol Crt	Casual

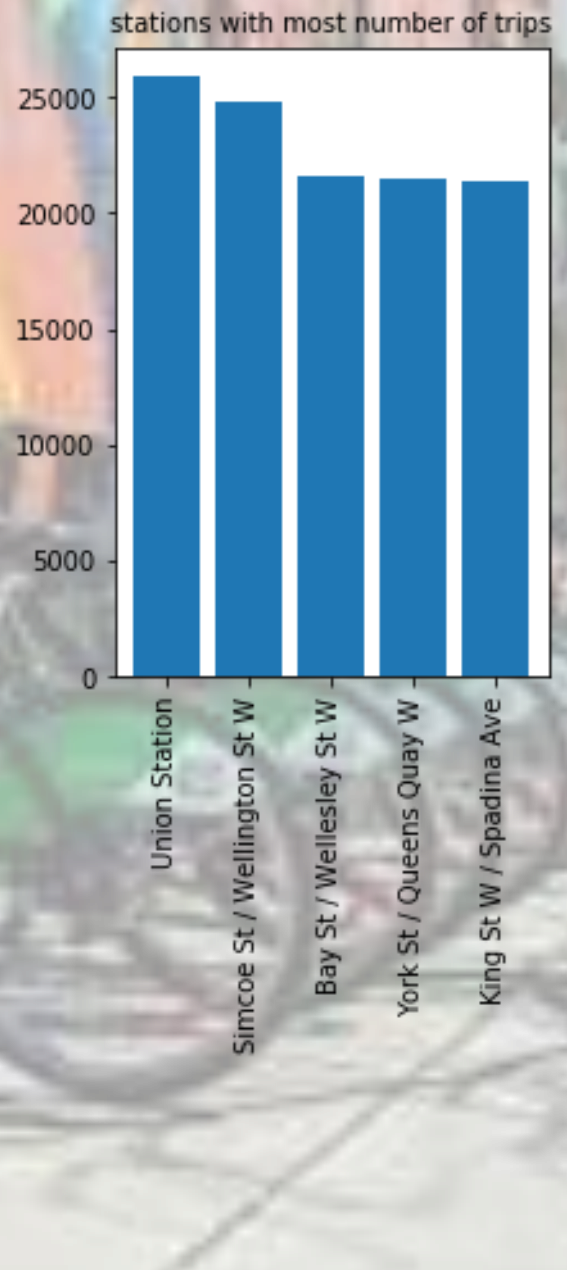
Stations with the most trips and with the least trips

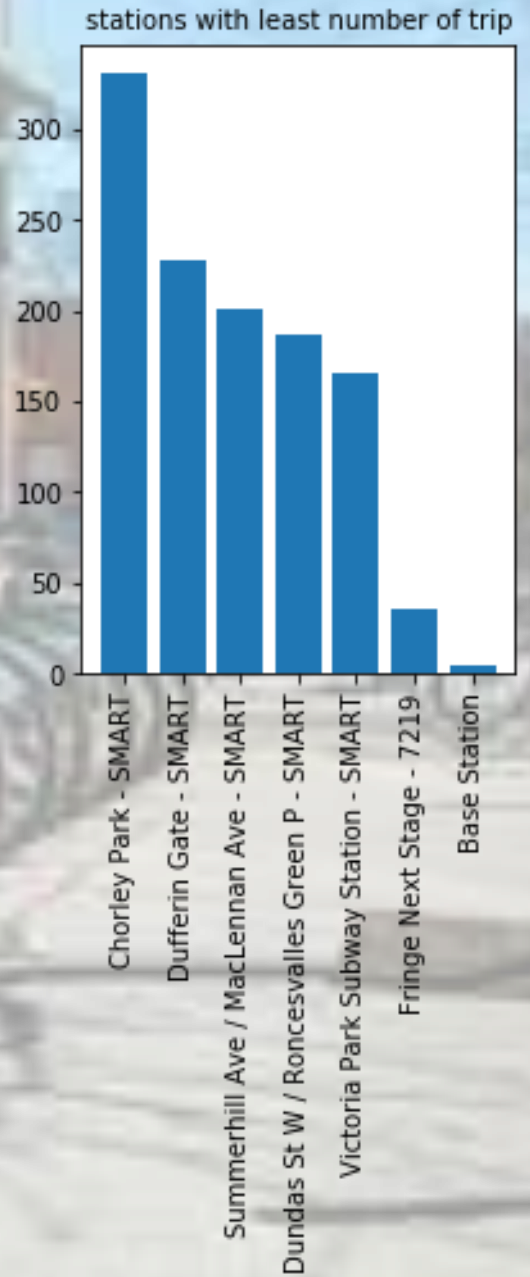


Trip Characteristics

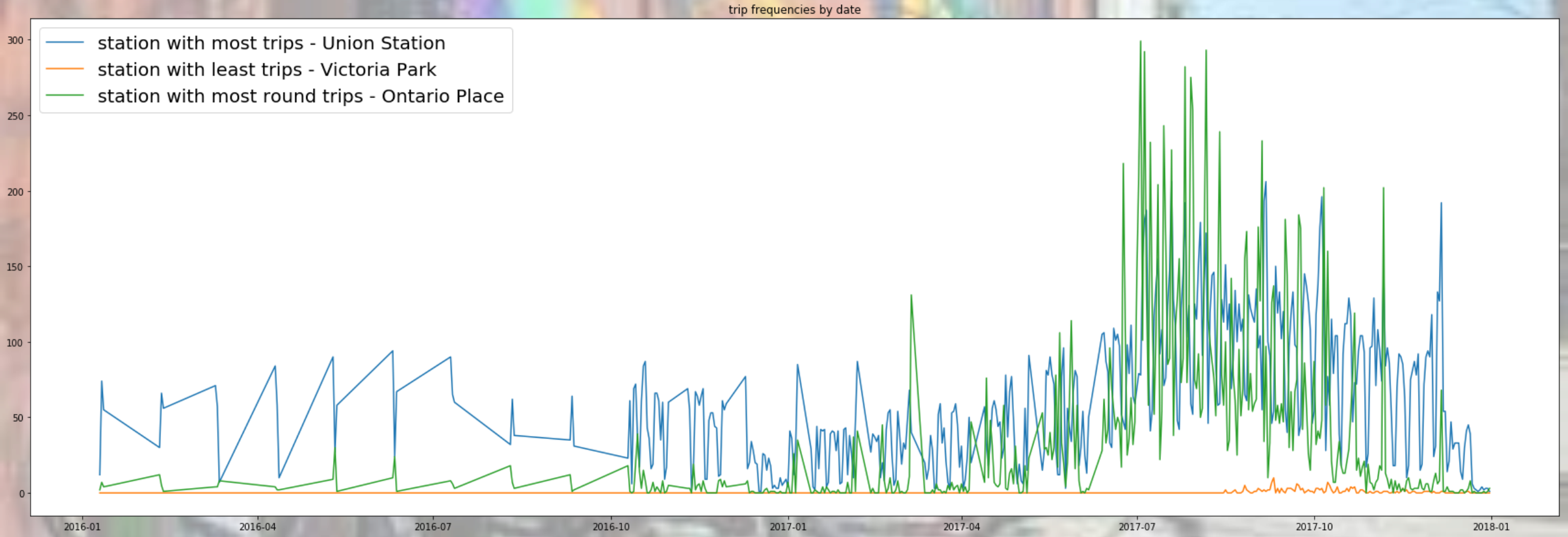




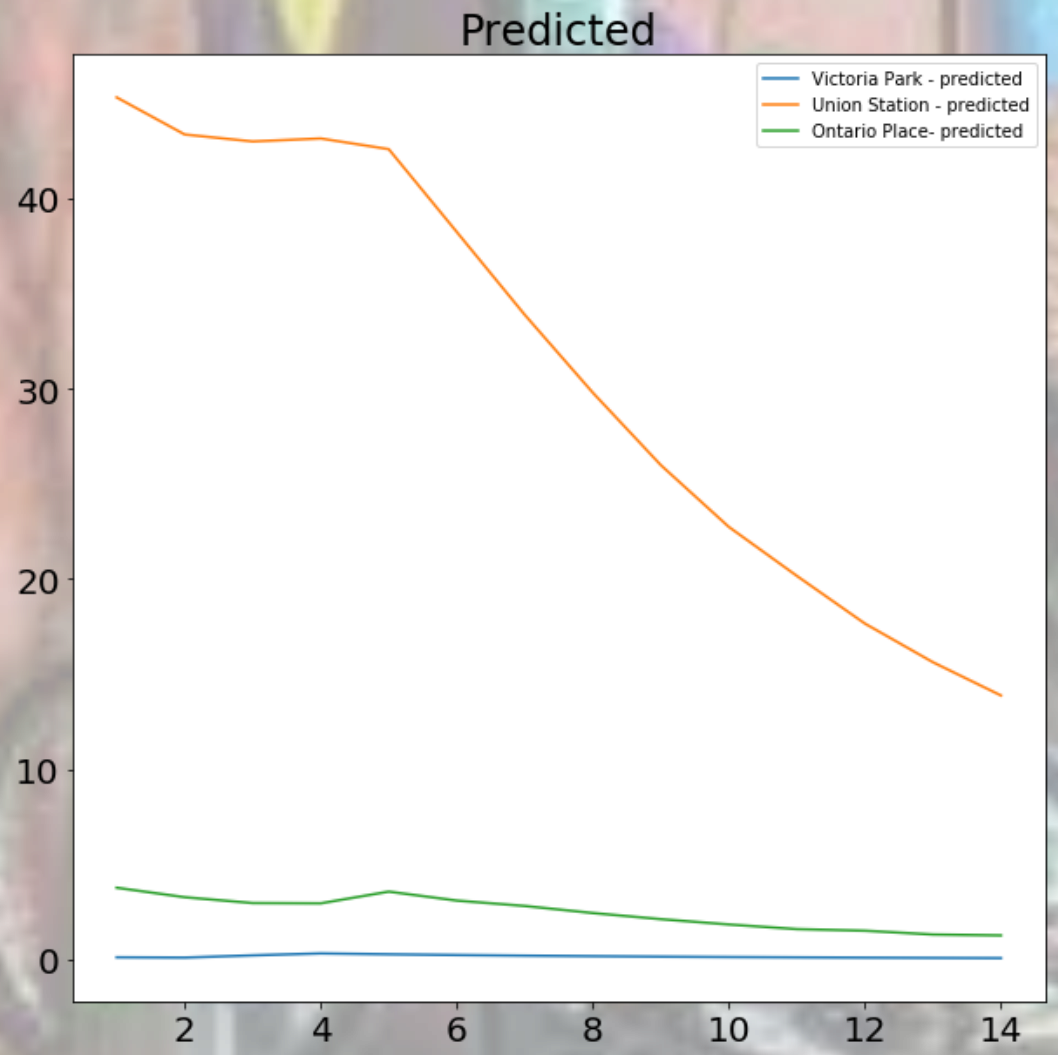


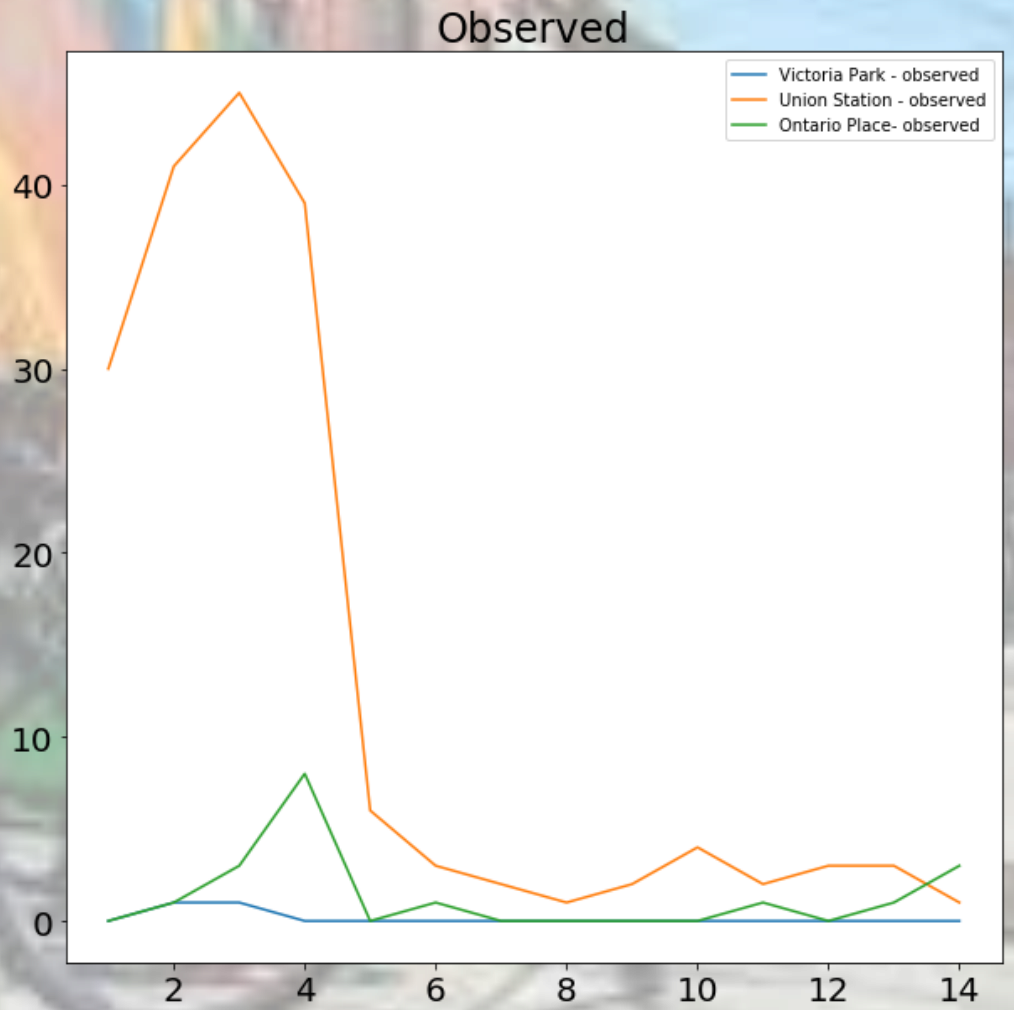


Trip Volume for Union Station (most), Victoria Park(least) and Ontario Place (most round trips)



Predicted and Observed Trip Volume of the Three Stations





Conclusions and Future Objectives

- Stations in downtown Toronto had the highest volume of rides over the time period while stations in more suburban areas had the fewest. This is likely that more short-distance trips within the initial 30 minute charge are possible and also that many trips are to and from workplaces or by visitors to the city.
- Ridership was fairly steady for the first part of the time period but shows a sharp increase in 2017Q3. This would coincide with the period when there are many large events in the city, such as the Canadian National Exhibition and Toronto International Film Festival.
- The autoregression model was somewhat poorer at predictions for daily volume for Union Station, with higher RSME. This was likely due to higher variance in the daily trip numbers which made it harder to predict future trip numbers.

Future objectives

- The data used is historical data. There is more current data that is available through an JSON API and I will look to updating the code to pull more recent data and test against the model.
- My initial objective had been to predict the number of bikes at each station for a given hour of the day, however the data available did not have sufficient information to do this. While the bikeshare website and app does show number of bikes at each station in real time, it does not guarantee that a bike will be available when a system user arrives at a station. I would like to build a predictive model to determine the number of bikes at each station and number of free docks.

Thank You