



# OPTIMAL LOCATION FOR NEW ASIAN RESTAURANT IN TORONTO

# RECOMMENDATION

- 4 optimal areas where demand might be unmet and have a possibility of being heavily trafficked due to other commercial companies being located there:
  1. **Burnfield Avenue, Christie Pits, Toronto**
    - Road connected to park in downtown Toronto. Multicultural neighborhood and close to subway station
  2. **Hart House Circle, Bloor Street Culture Corridor, Toronto**
    - Part of diverse cultural district with many distinct, culture and art organizations. Has very few Asian restaurants in direct vicinity. Multicultural space open for specialty type foods. In Downtown Toronto
  3. **Ashby Place, Corktown, Toronto**
    - Old residential area in the process of being gentrified, with many residential areas being torn down for commercial projects. There are a number of vacant industrial buildings there repurposed for commercial locations such as food and shops. Next to Downtown Toronto
  4. **Kingston Road, Cliffside, Toronto**
    - Upscale suburban housing area with many parks and waterfront areas. Population primarily of British Isles descent, but there might be desire for exotic/ethnic food. It has some distance from highly trafficked Downtown Toronto, but that also means less competition

# BACKGROUND

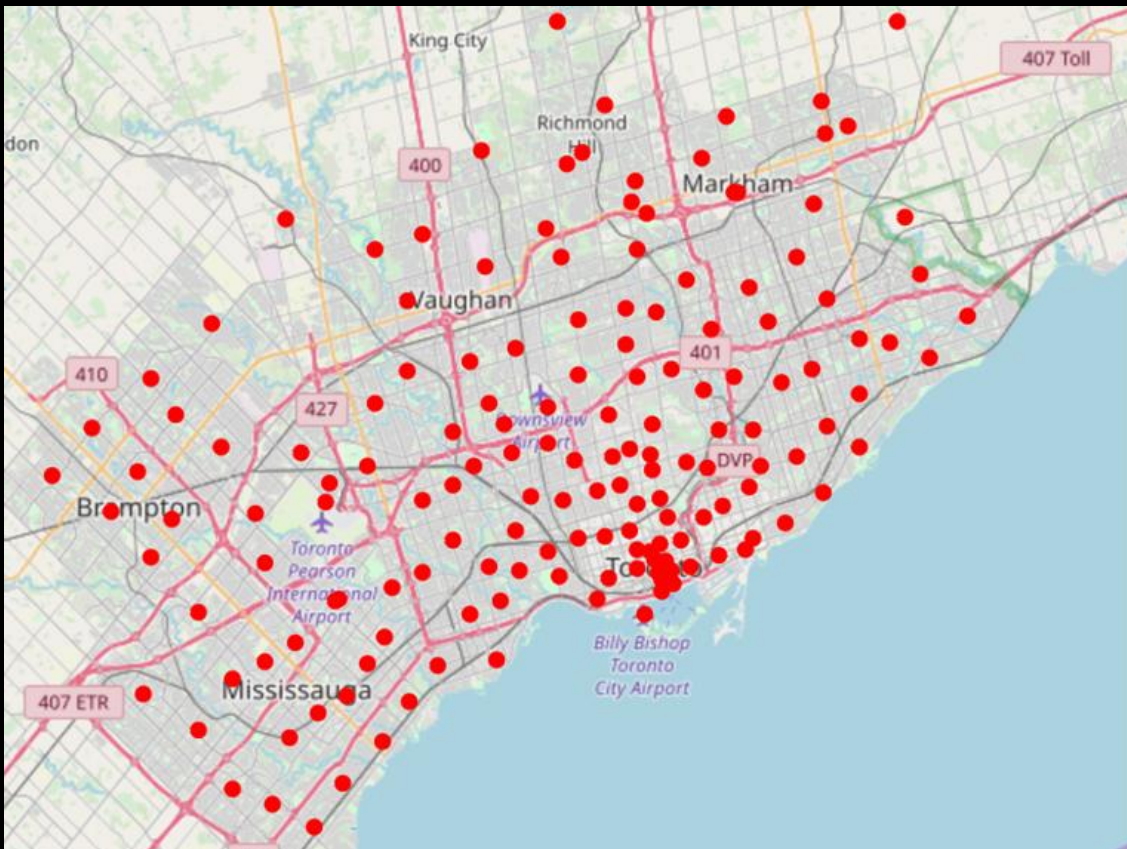
- Toronto's food and beverage sector produces annual sales over \$17.75 billion
- The industry has an average growth of 4% to 5%
  - Specialty (ethnic) food services are projected to grow at 12% over next five years
- Given this demand, what would the optimal locations to began a search for a new Asian restaurant?
  - Stakeholders: potential entrepreneurs, shareholders (and other backers), policy makers
- Assumptions for optimal location:
  - Contains potential for unmet demand
  - Has limited competition from similar themed products/restaurants
  - Competition is other mid and high tier restaurants

# DATA ACQUISITION AND CLEANING

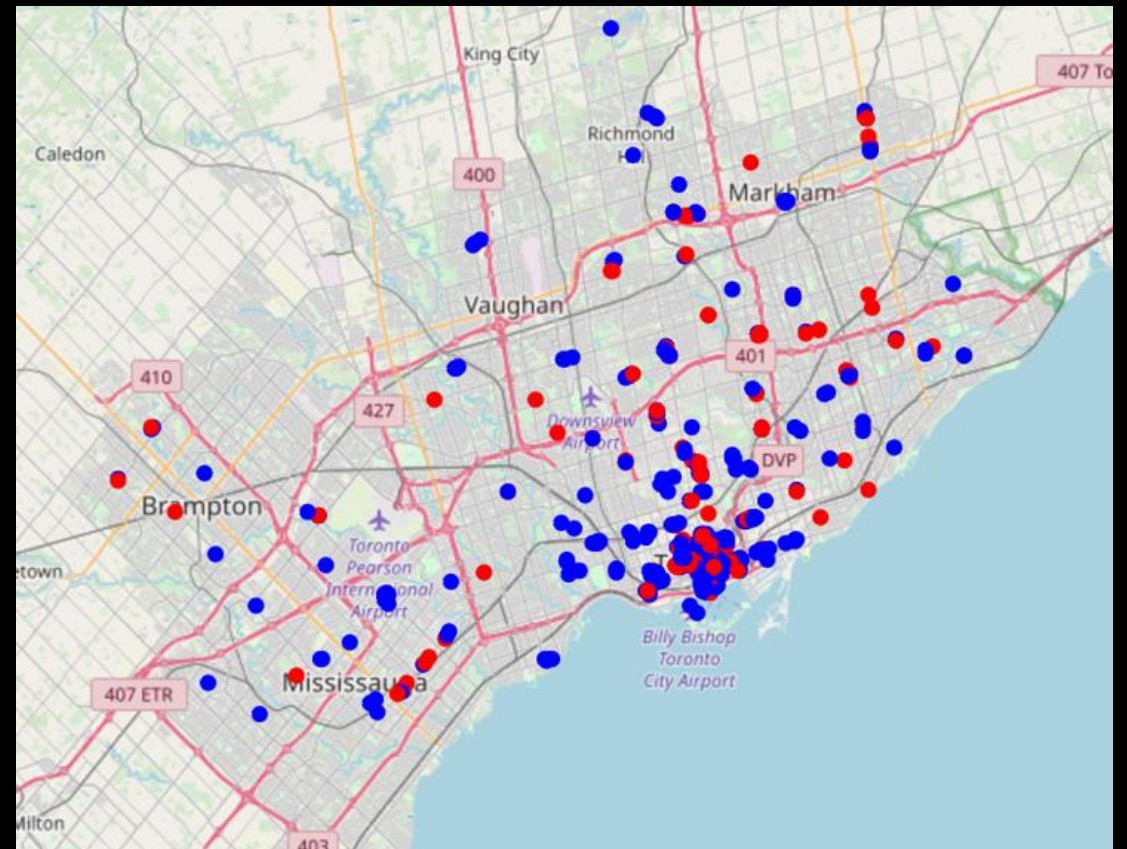
- Postcodes, Borough, Neighborhood, Latitude, Longitude from these sets:
  - [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
  - [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data)
  - [https://fusiontables.google.com/DataSource?docid=1H\\_cl-oyeG4FDwqJUTel\\_aGKmmkJdPDzRNccp96M&hl=en\\_US&pli=1](https://fusiontables.google.com/DataSource?docid=1H_cl-oyeG4FDwqJUTel_aGKmmkJdPDzRNccp96M&hl=en_US&pli=1)
- Population, housing, total occupied housing from Canada census for postcodes:
  - <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hltfst/index-eng.cfm>
- Restaurant with name, address, distance from postcode, category from Foursquare API: <https://foursquare.com/>
- Data cleaning
  - Empty rows were dropped, much of the data was joined together
  - Two features, count of restaurants and count of category restaurants added to main table
- 11 total features and lists containing restaurants



# VISUALIZATION OF GEOGRAPHIC LOCATIONS



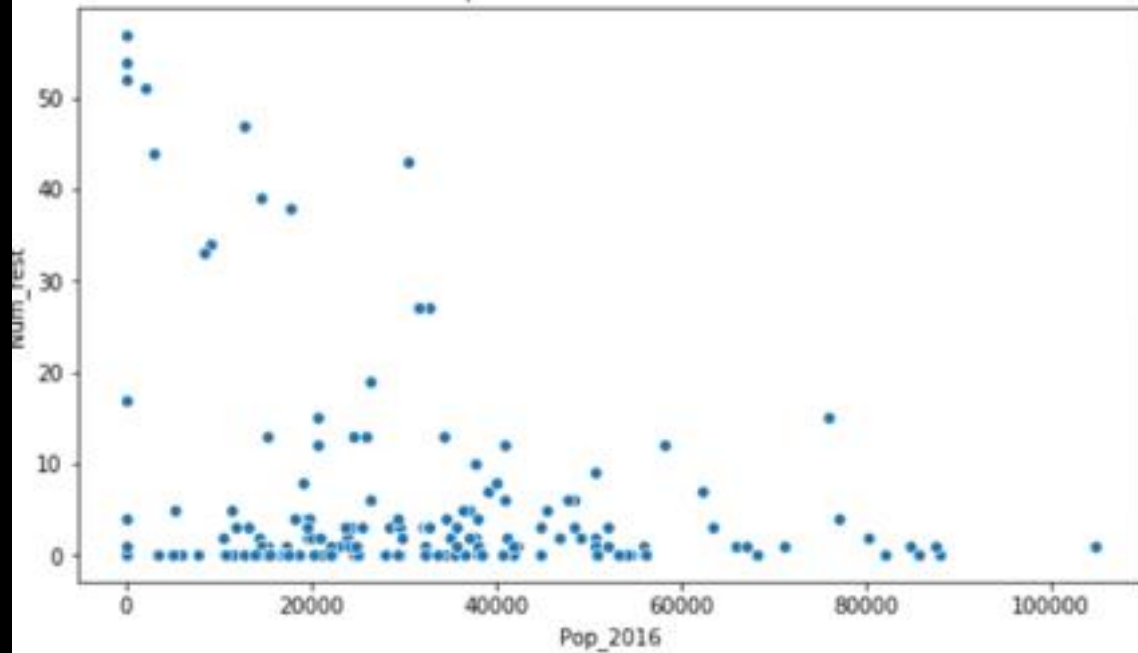
Postcode locations throughout Toronto



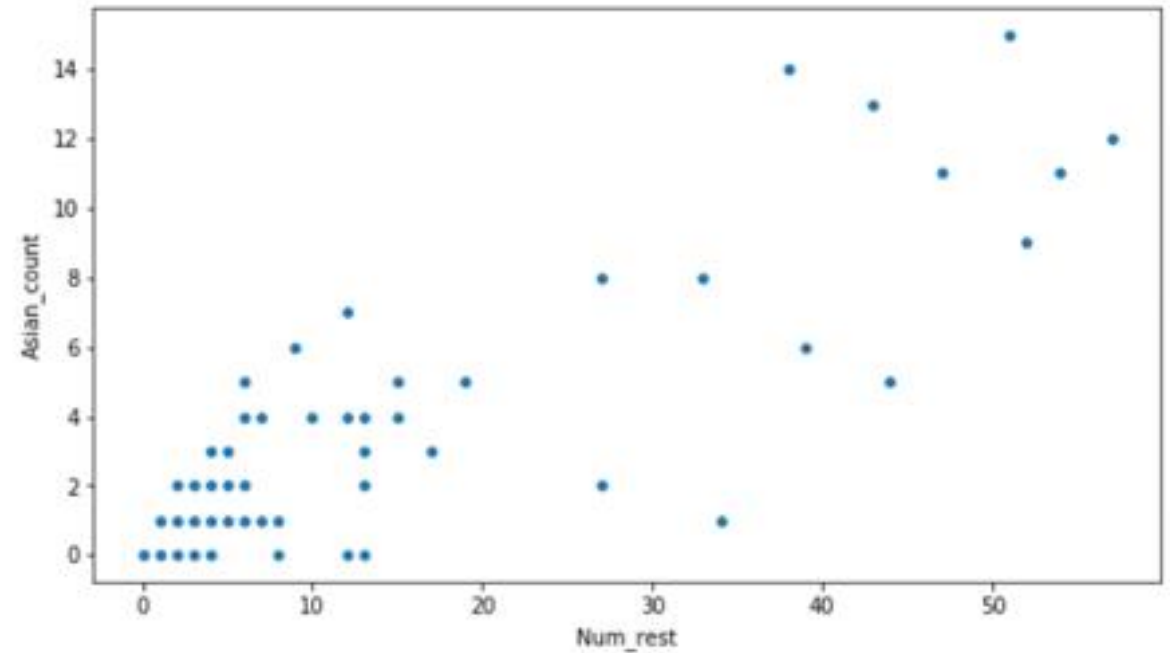
Restaurant locations throughout Toronto

# SCATTERPLOTS OF RESTAURANT COUNTS

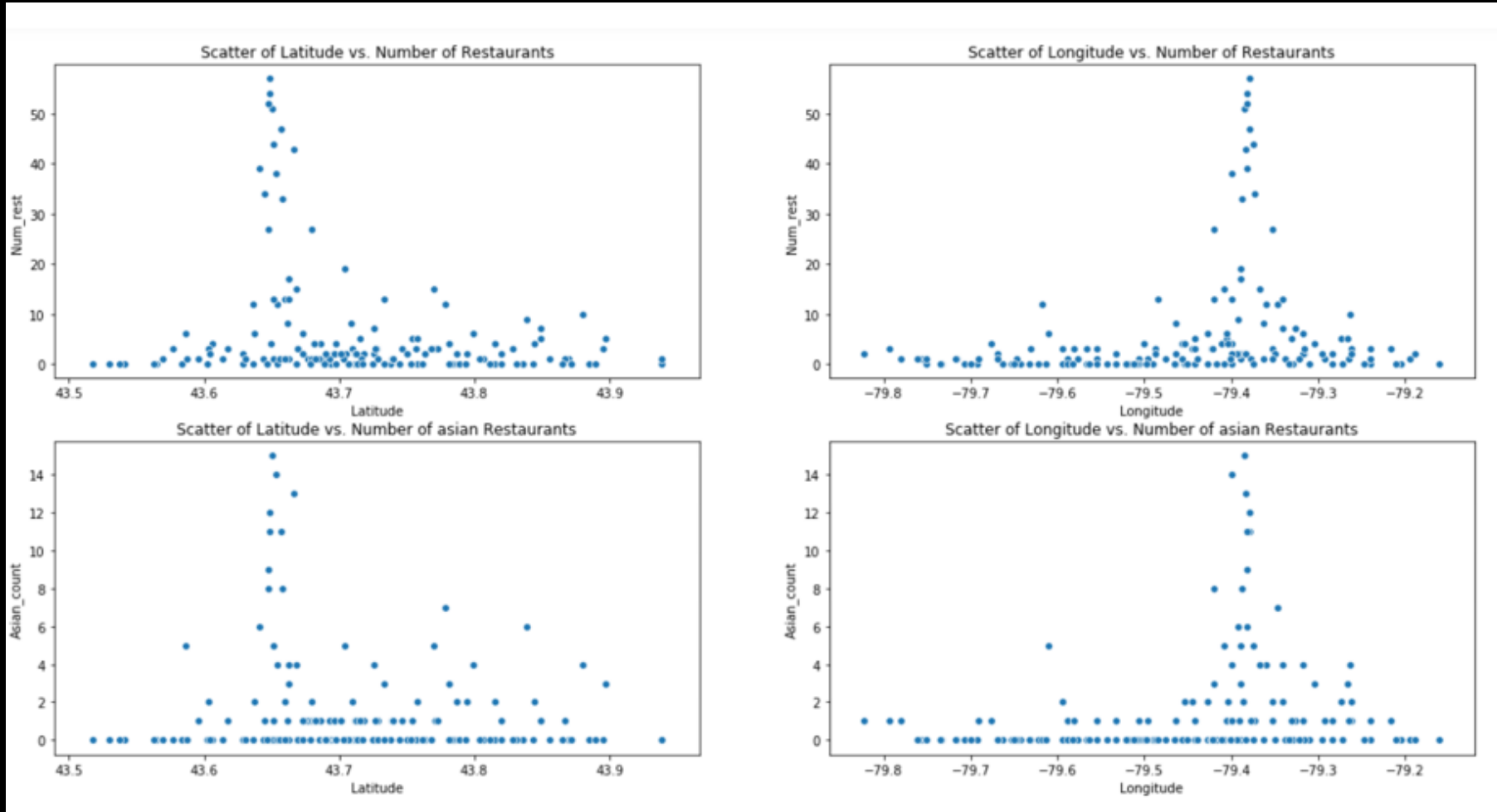
Scatter of Population vs. Number of Restaurants



Scatter of Number of Asian Restaurants vs. Number of Restaurants

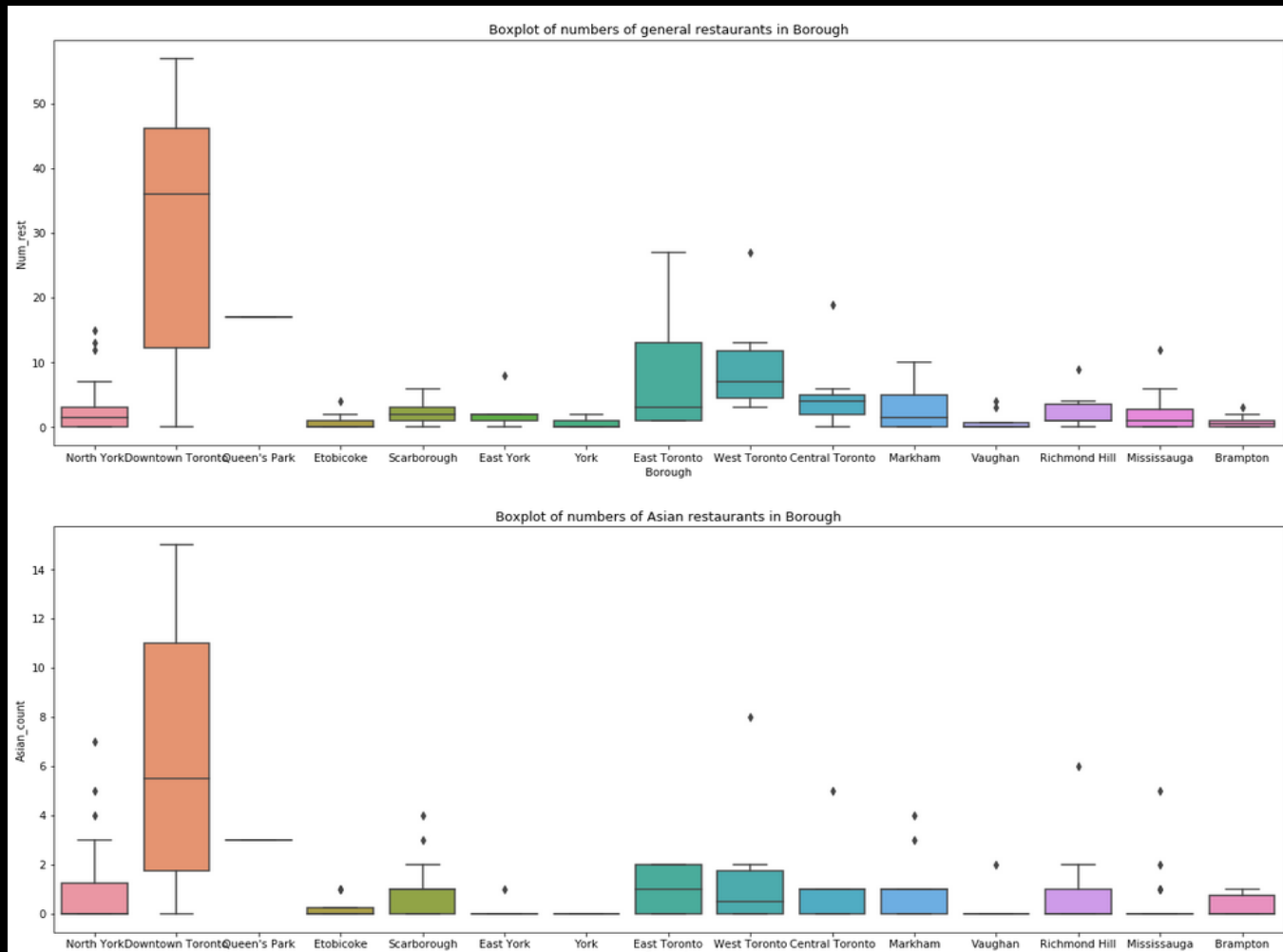


# LATITUDE/LONGITUDE VS. RESTAURANT NUMBER





# BOROUGH VS. RESTAURANT COUNT



- Downtown Borough have higher counts as percentiles than all other boroughs
- New binary feature, asking if postcode is in downtown borough or not

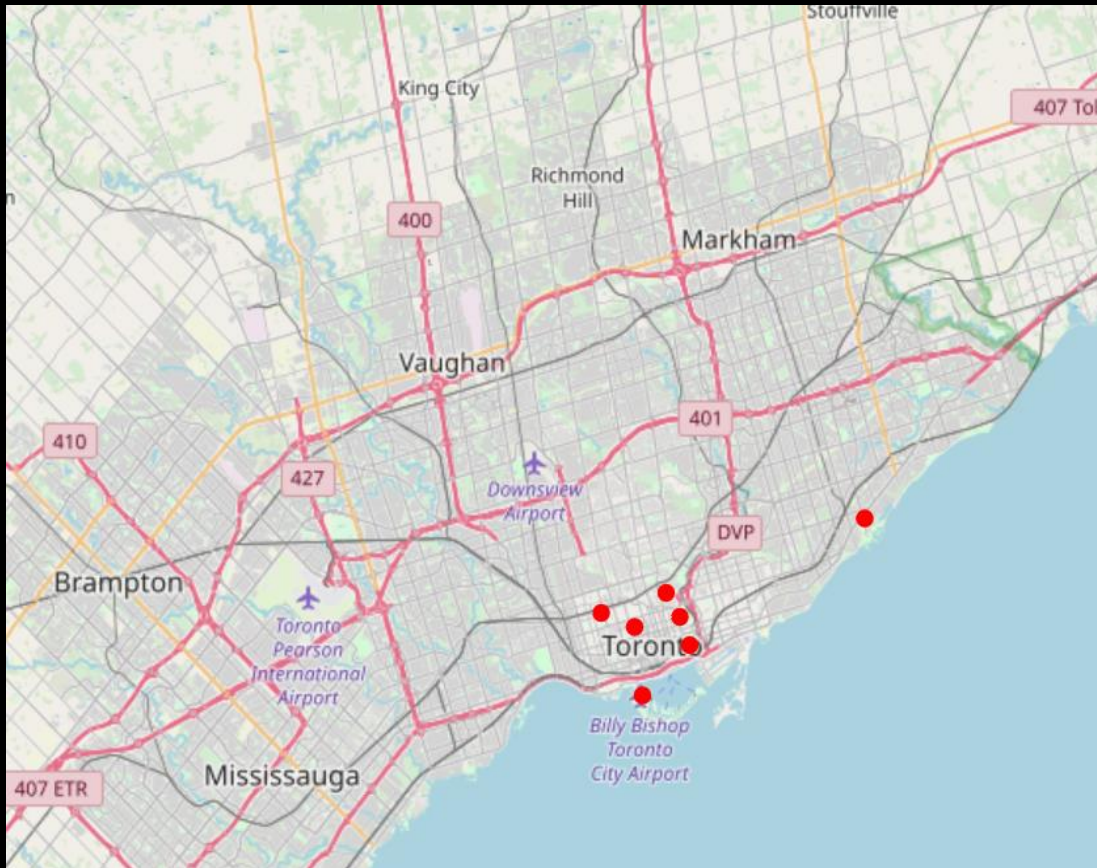


# REGRESSION METHOD

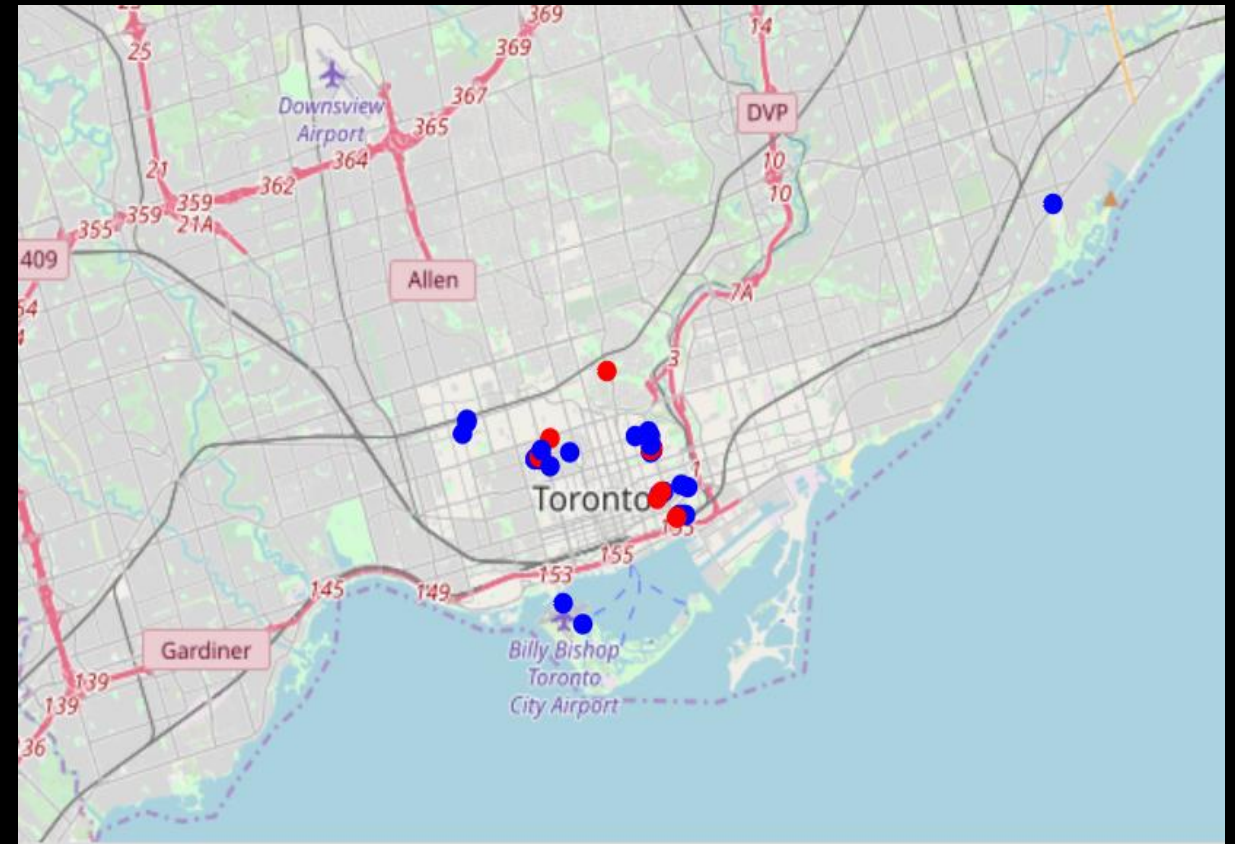
- Two linear models were built
  - Features were scaled by subtracting mean and dividing by standard deviation
  - No division of data into training, test, validation sets
  - Trained on whole data set then predicted on the same dataset
  - One predicts count of restaurants generally, the other specifically Asian restaurants
- Objective was to ask 2 questions: which postcodes were restaurant counts overpredicted by model? Which postcodes were Asian restaurant counts overpredicted by model?
- Only get postcodes where this is true

	Mean Squared Error	R <sup>2</sup> score
Model for Asian prediction	4.48	0.4448
Model for general restaurant prediction	59.4797	0.5667

# GEOGRAPHIC VISUALIZATION OF FINAL POSTCODES



Postcodes where both questions are true



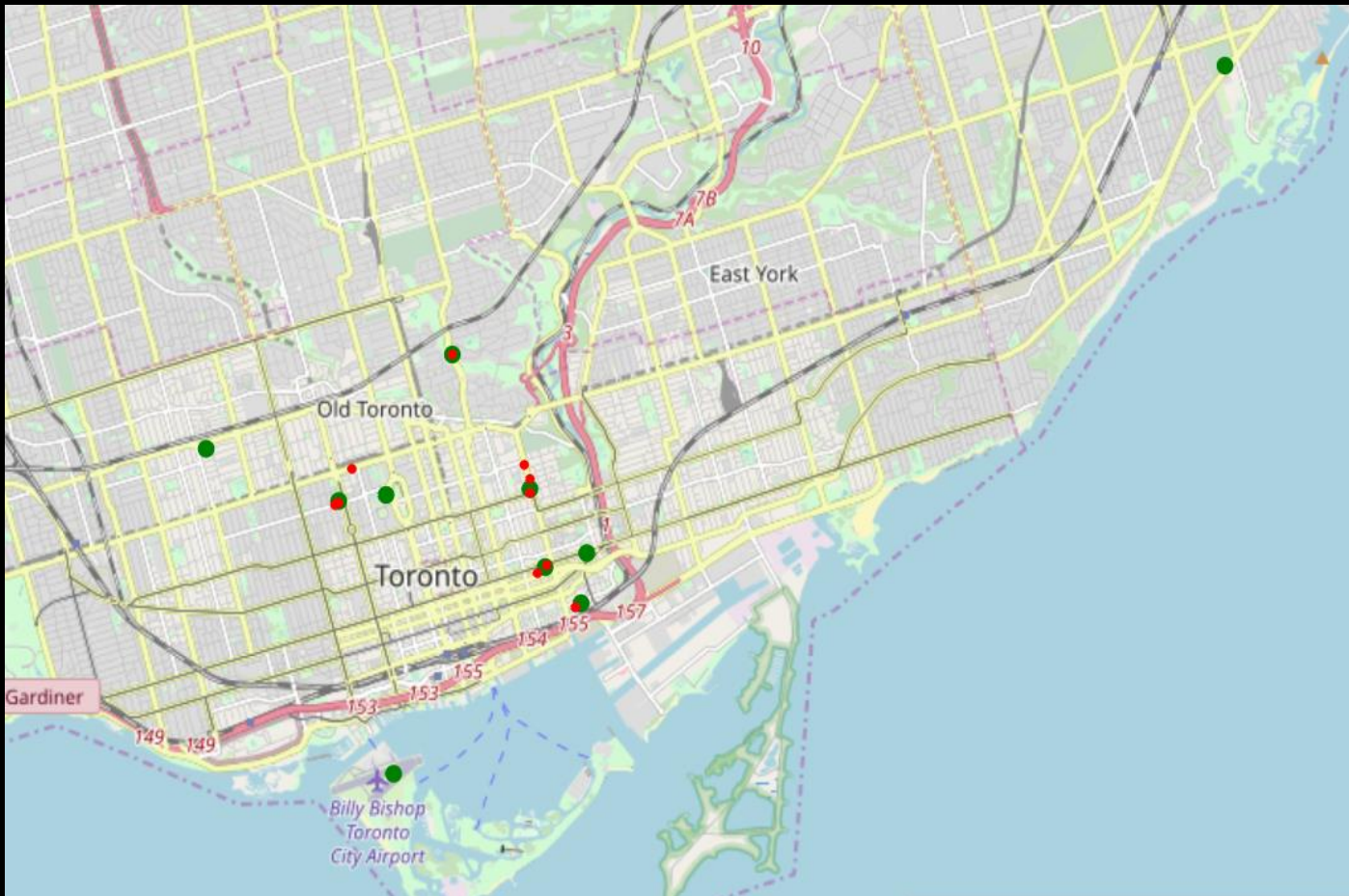
Restaurants in those postcodes

# CLUSTERING METHOD

- Using the final set of restaurants, creating 11 random centroids
- Nonhierarchical method: Kmeans
  - Objective is to find addresses around which restaurants are located
  - These addresses are assumed to be common commercial areas which will be highly trafficked
  - When people in area are searching for food, they will go for those addresses
- Out of 11 Clusters trained, 5 of them had no Asian restaurants in their direct vicinity
  - 1 at airport, this was ignored as actual potential locations may be artificially depressed
  - 3 near or in Downtown Toronto
  - 1 in Eastern Toronto



# VISUALIZATION OF THESE AREAS



1. Burnfield Avenue, Christie Pits, Toronto
2. Hart House Circle, Bloor Street Culture Corridor, Toronto
3. Ashby Place, Corktown, Toronto
4. Kingston Road, Cliffside, Toronto





# FUTURE DIRECTION

- Generalizing linear models, thus removing features that are specific to geographic location and adding features that exist in every city
  - Using different cities for training/testing
  - Using other methods of regression such as perceptrons and Tree regressors
- In general, in-depth understanding of factors that influence demand, and data collection of those factors in connection with geography
- Looking at the relationship between different types of restaurant and location