Toronto Restaurants

By the numbers

Author: Rafik Matta

What's the problem?

How are Toronto's restaurants distributed by neighborhood?

1. What's the distribution of Toronto's restaurants?
2. What are the income and ethnic distributions of Toronto?
3. What type of restaurants does Toronto have?
4. Is it cheaper to eat in some neighborhoods and not others?
5. Do cheaper neighborhoods have worse quality restaurants?
6. How expensive/good are Toronto's restaurants anyway? (Pie Charts)

Data Sources

Statistics Canada Statistique Canada

Google Maps APIs

yelp

- https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#8c732154-5012-9afe-d0cd-ba3ffc813d5a
- https://www.yelp.com/dataset
- https://developers.google.com/maps/documentation/geocoding/
- http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CMACA&Code1=535&Geo2=PR&Code2=01&Data=Count&SearchText=toronto&SearchType=Begins&SearchPR=01&B1=Ethnic%20origin&TABID=1
- https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#a45bd45a-ede8-730e-1abc-93105b2c439f

Pre-processing

- Extract Toronto ethnic and income data from the Census 2016 dataset to prepare Toronto level visualizations (this was done using Excel).
- Convert ESRI Shape files to GeoJSON for use with Folium library
- Convert Toronto Neighborhoods names into decimal coordinates using Google Maps API Geolocation feature
- Append Yelp Restaurant's Toronto Neighborhood based on proximity
- Extract top neighborhoods with over 100 restaurants for comparative analysis
- Apply filtering on the Neighborhood Profiles dataset to find ethnic and income information per neighborhood.
- Apply filtering on the Yelp dataset with data from step 3 to create price and rating comparison charts.
- Apply pivot methods and aggregations to count and find means.

Tools Used

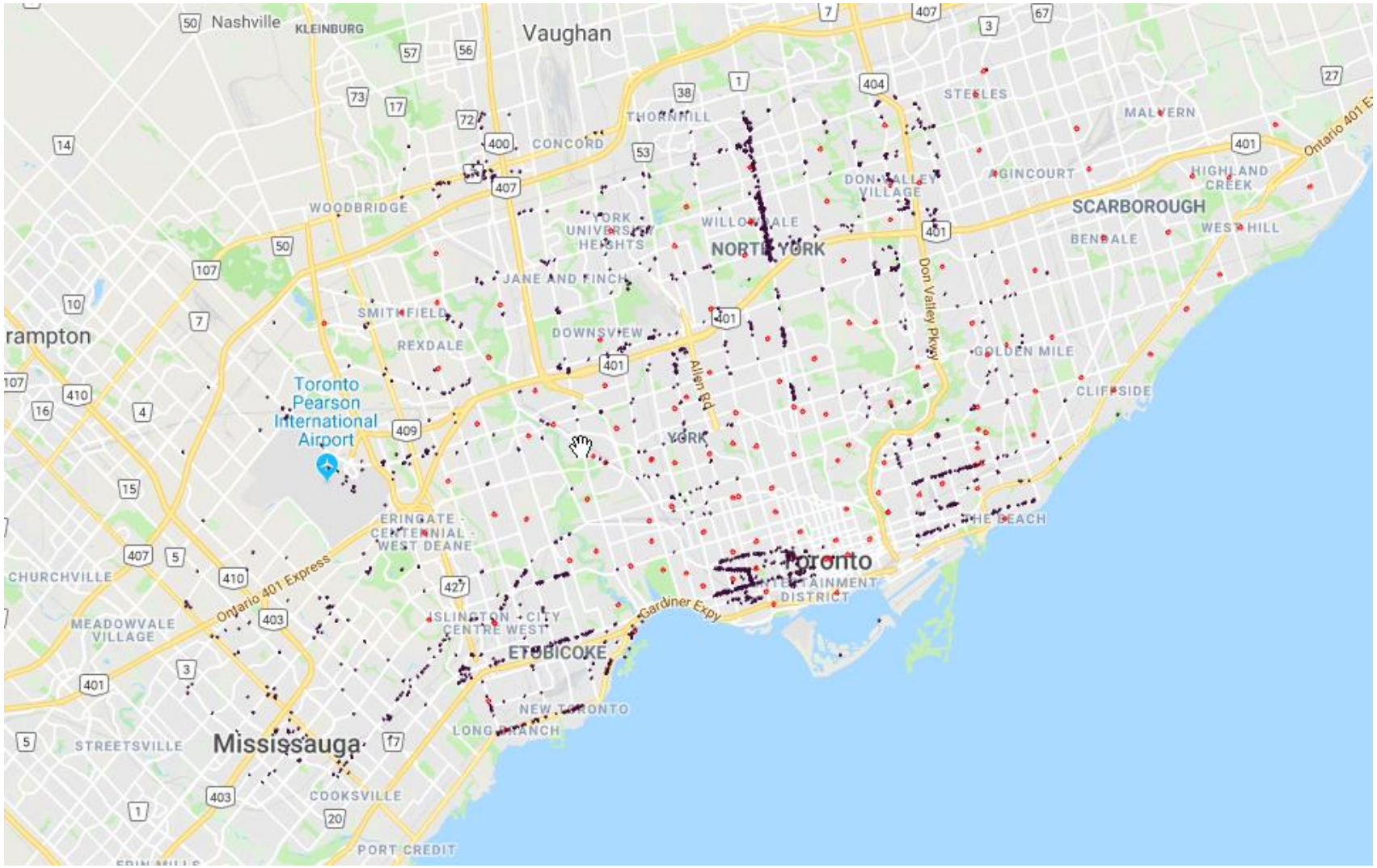
pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$

Google Maps APIs

Restaurant Distribution

where are all the restaurants?

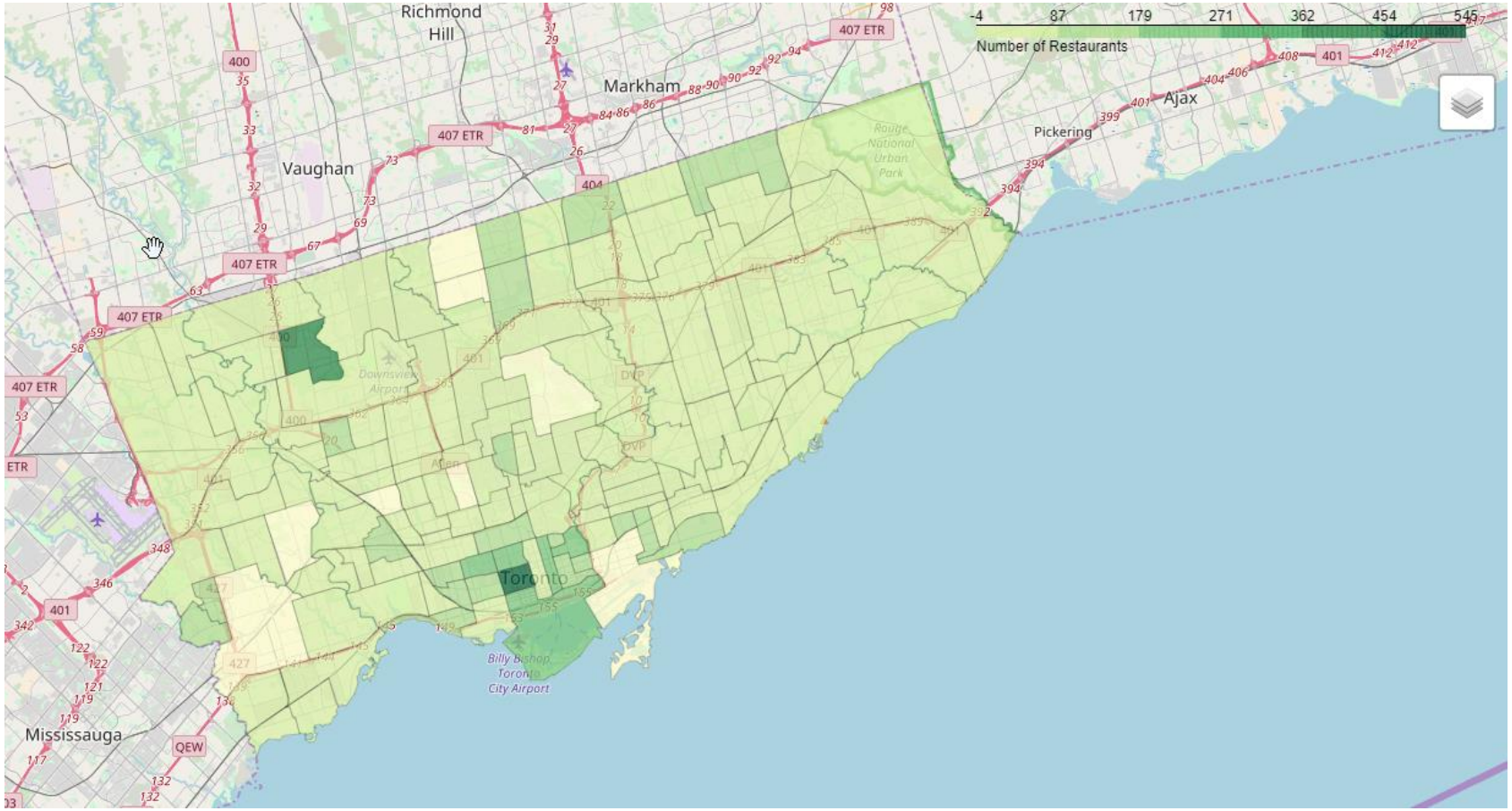


--- Neighborhood centroid

--- Restaurant

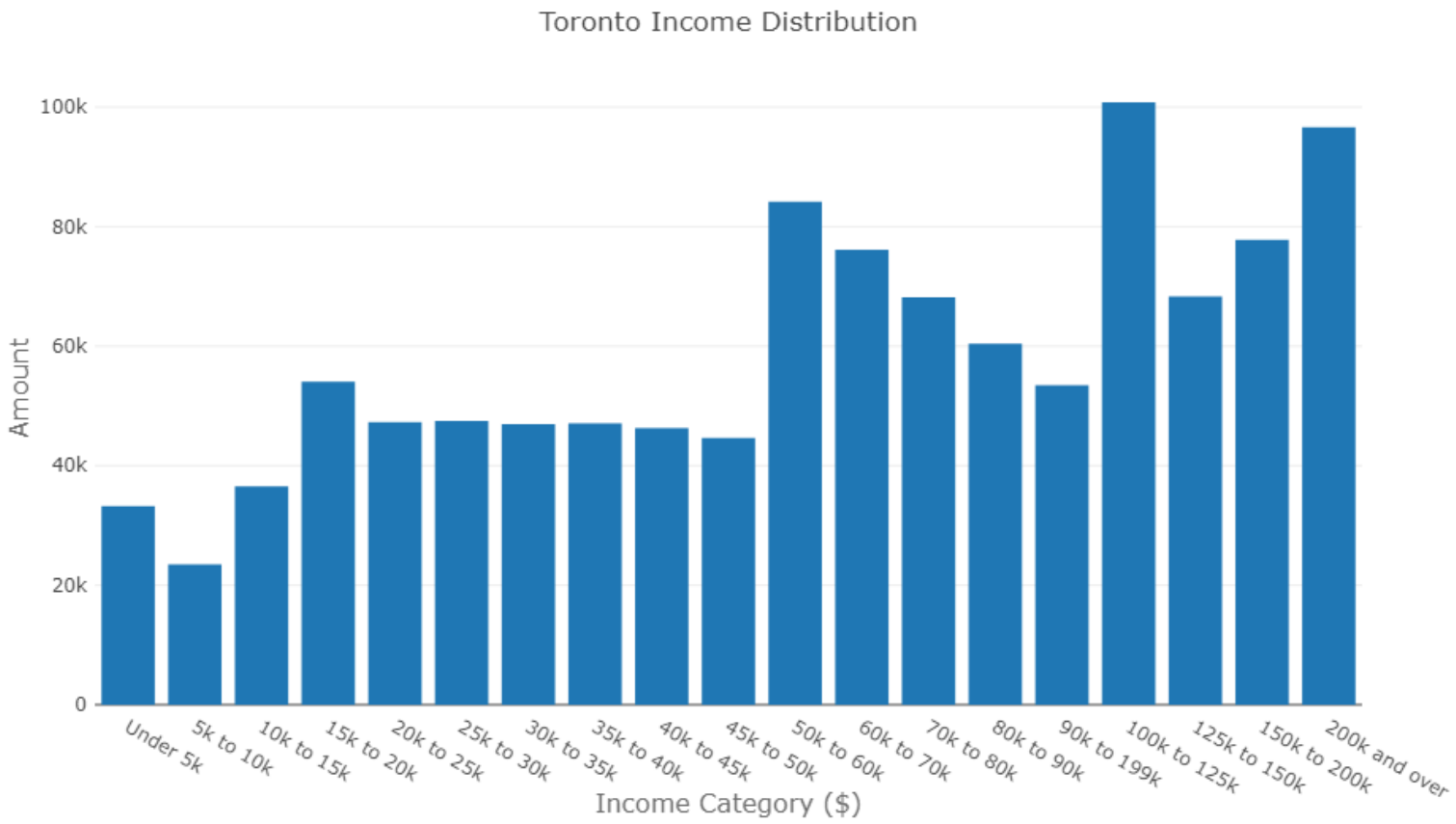
Where are all the restaurants?

Note: Open interactive Choropleth

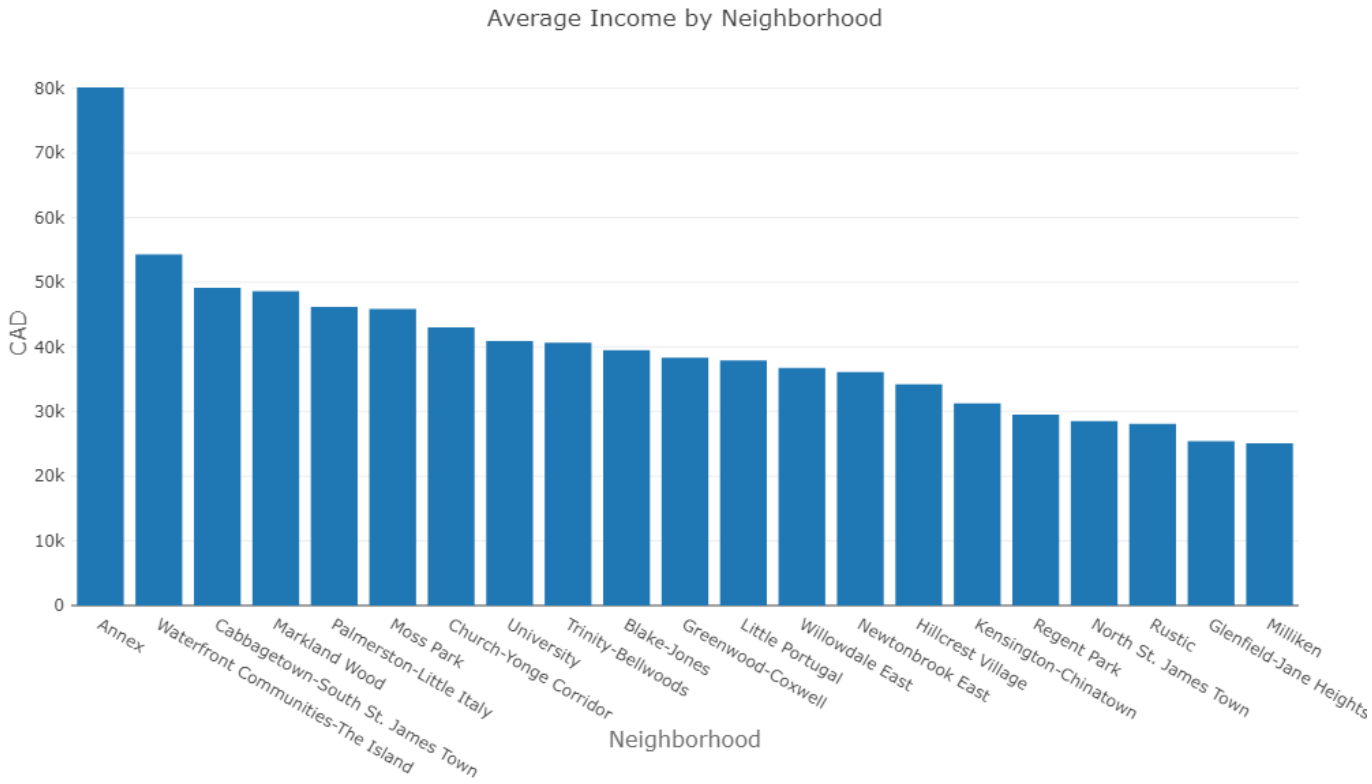


Income

Toronto Income Dist.

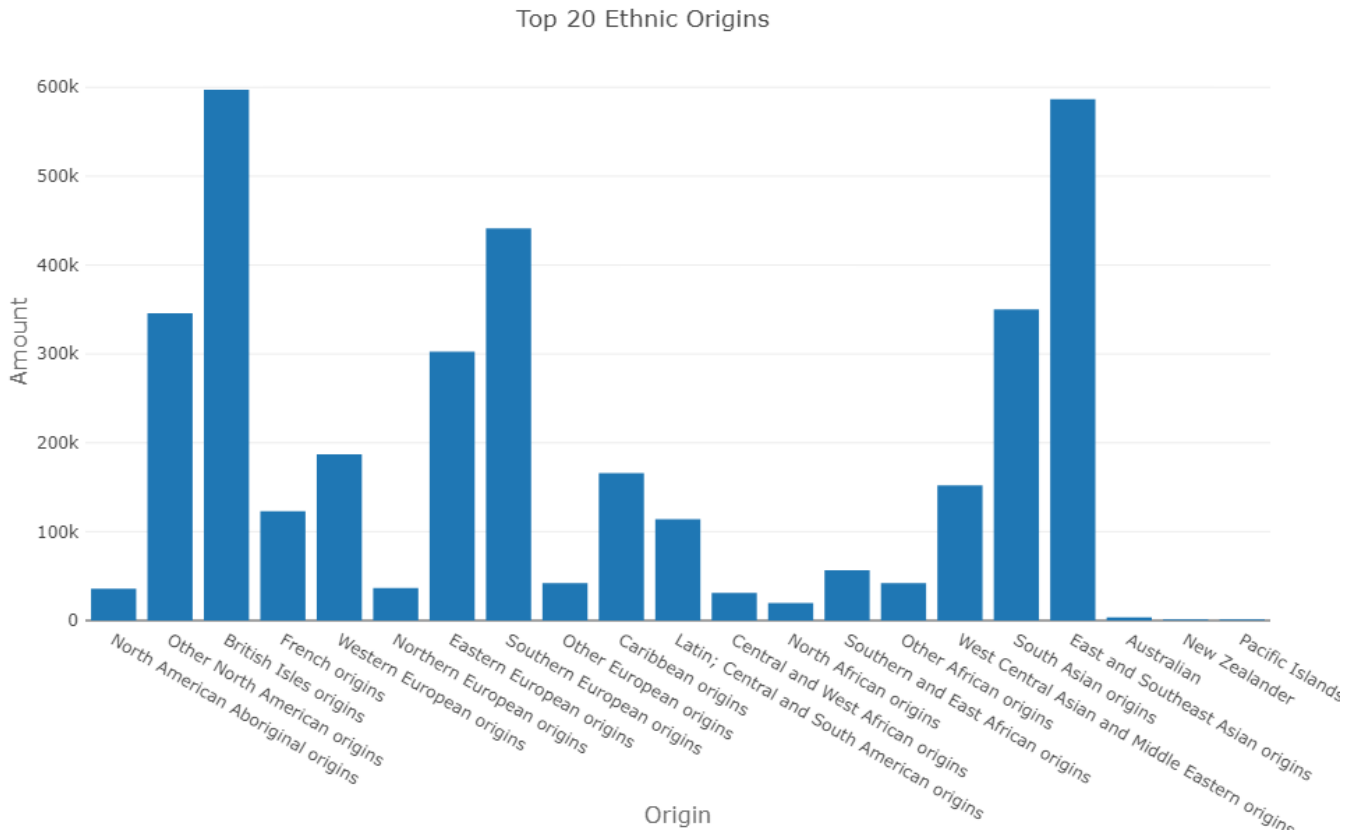


Income By Neigborhood

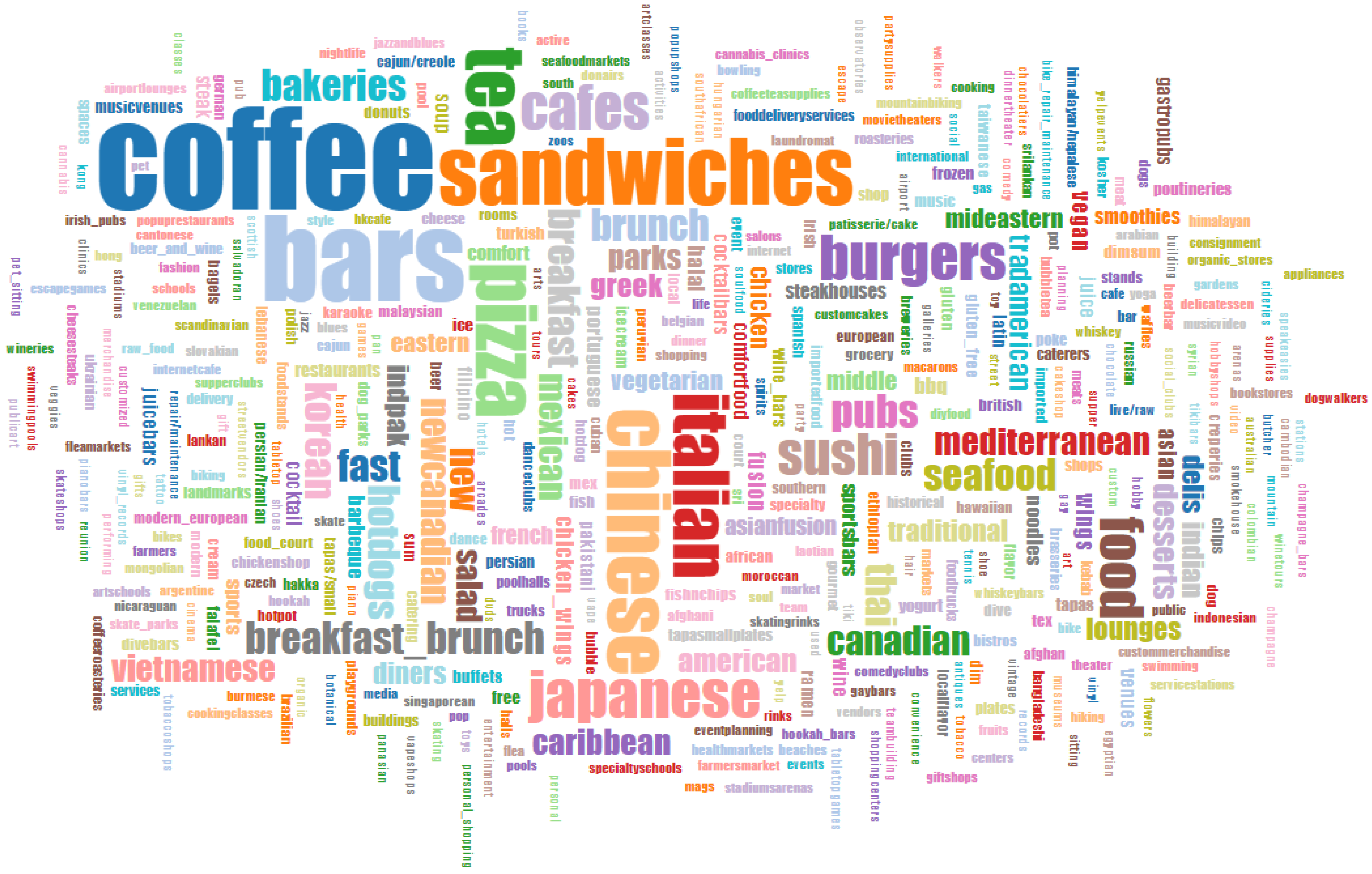


Ethnicities

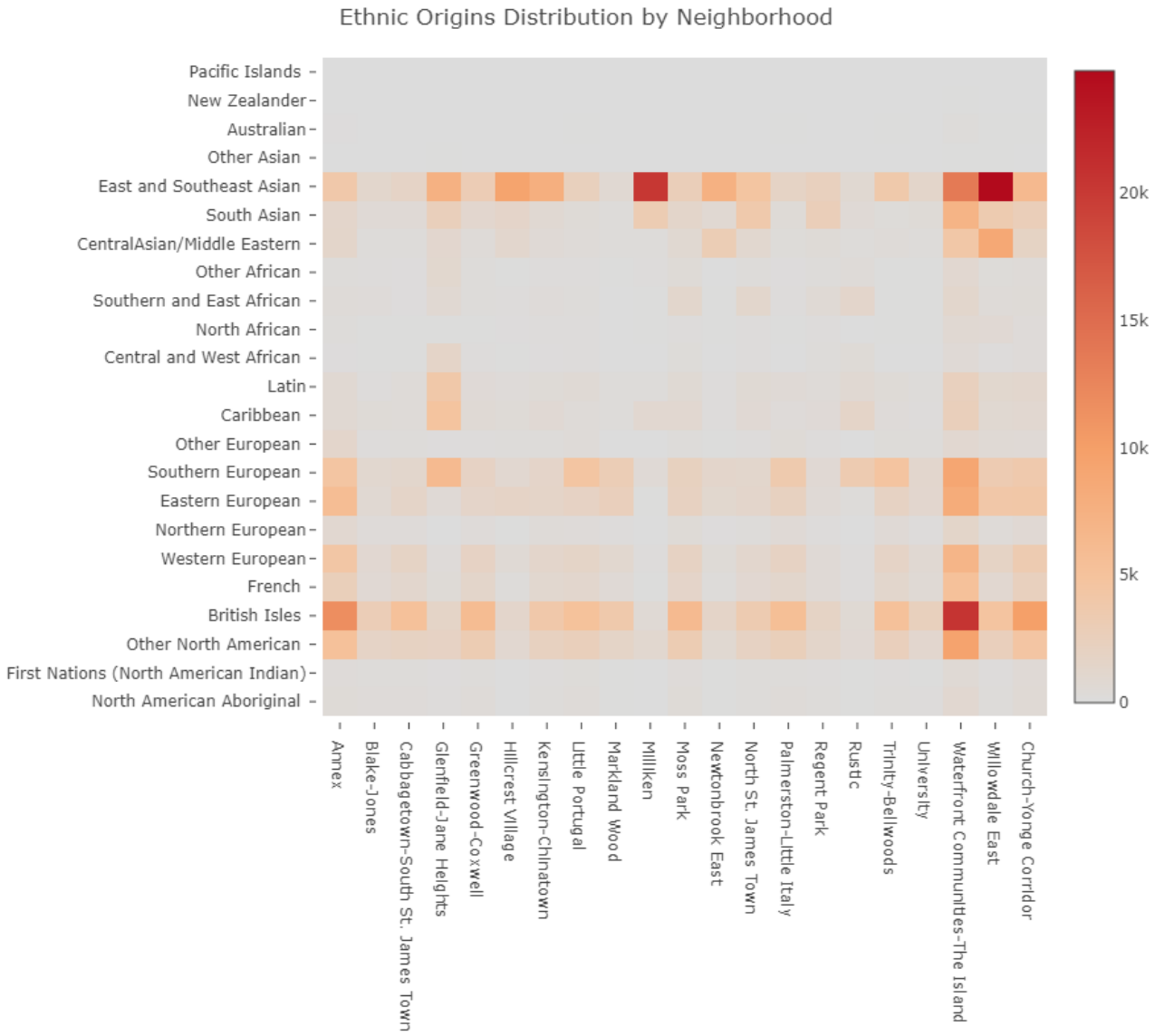
Top Ethnicities



Top Restaurant Categories



Ethnicities by Neighborhood



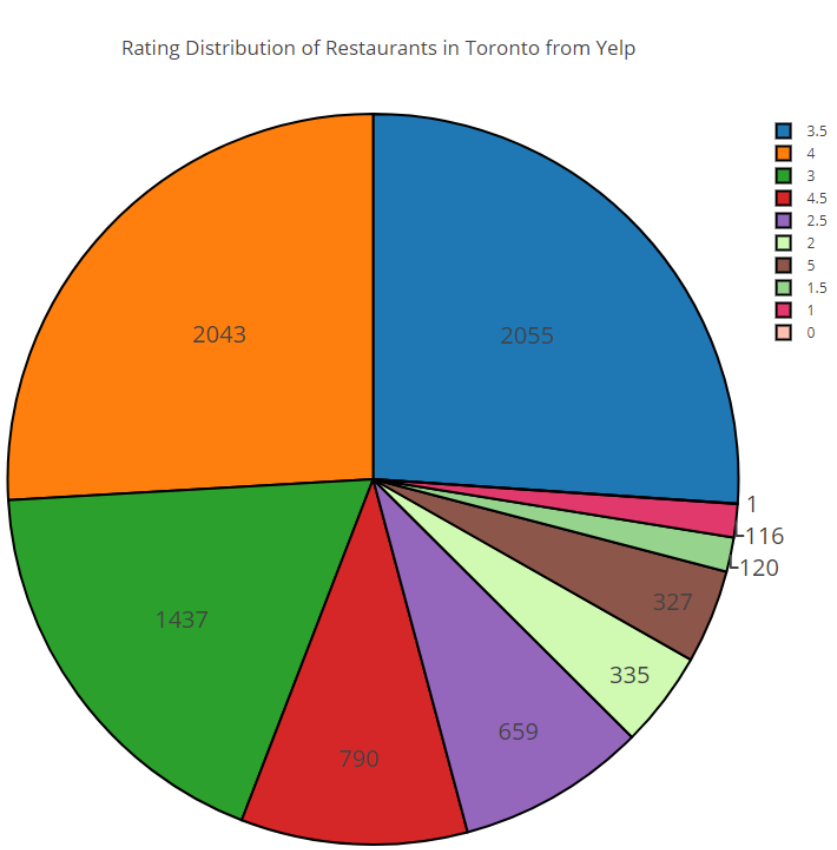
Top categories

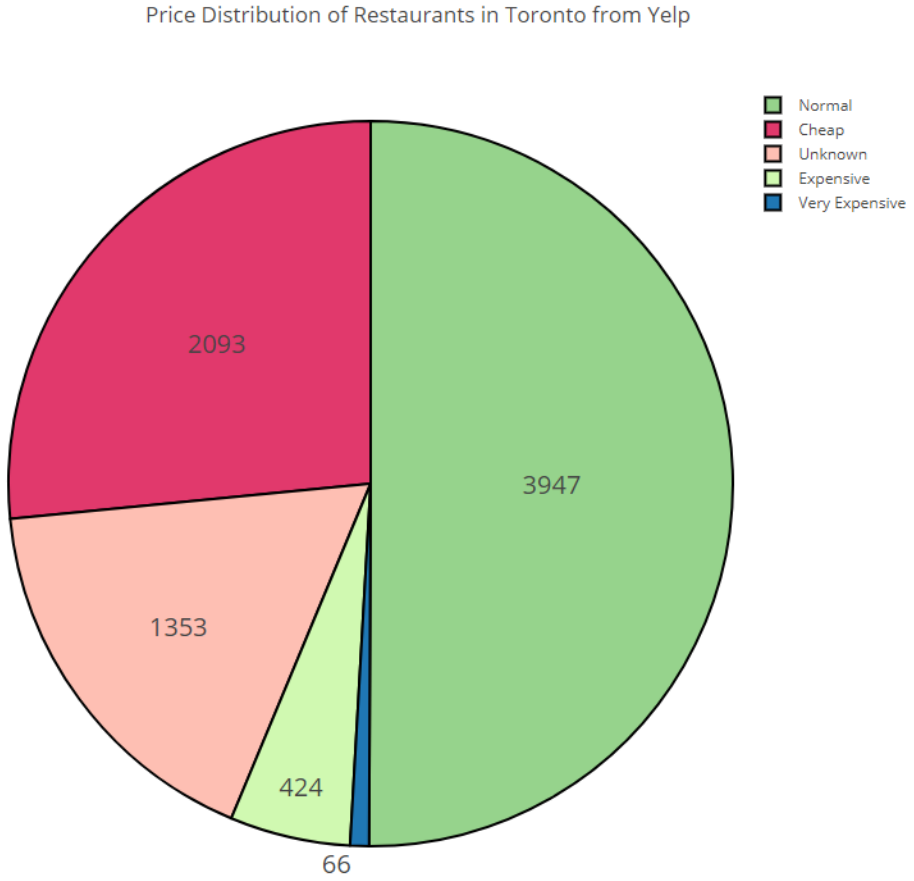
Kensington-Chinatown

Glenfield-Jane Heights

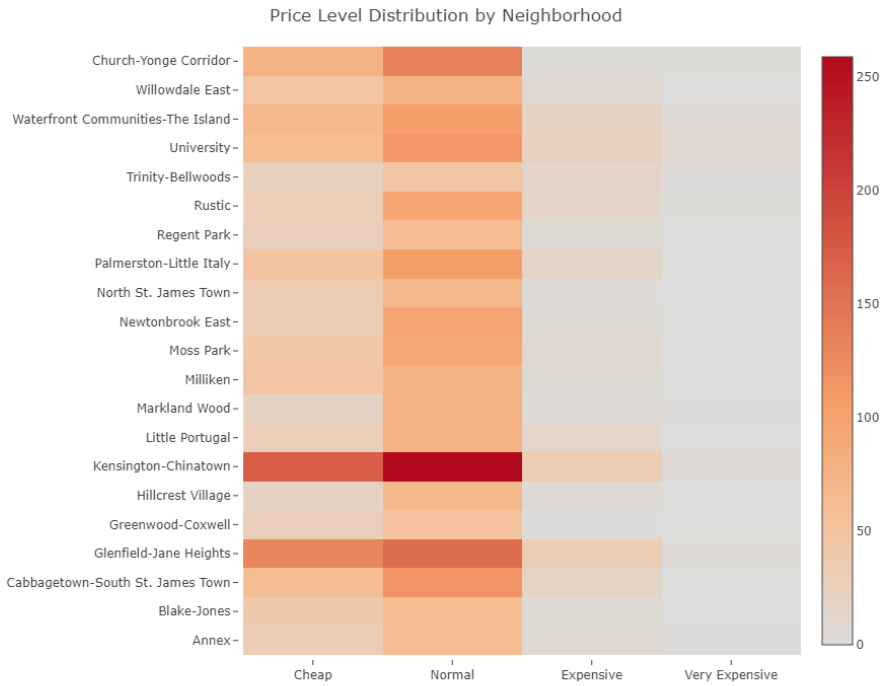
Price vs. Rating

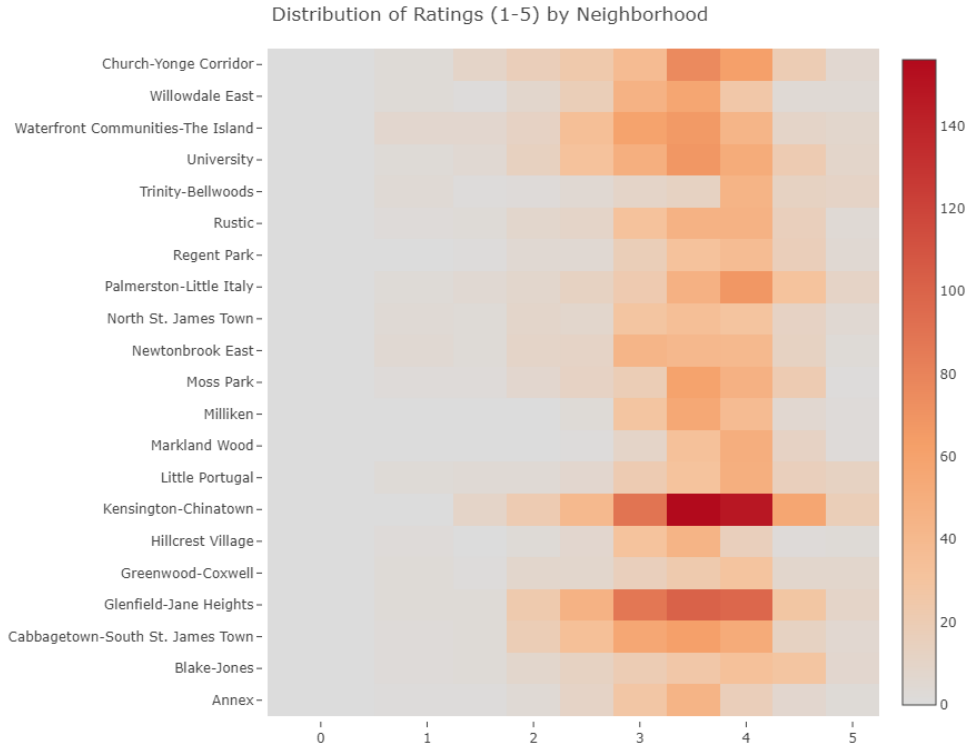
Toronto Restaurant Ratings Breakdown



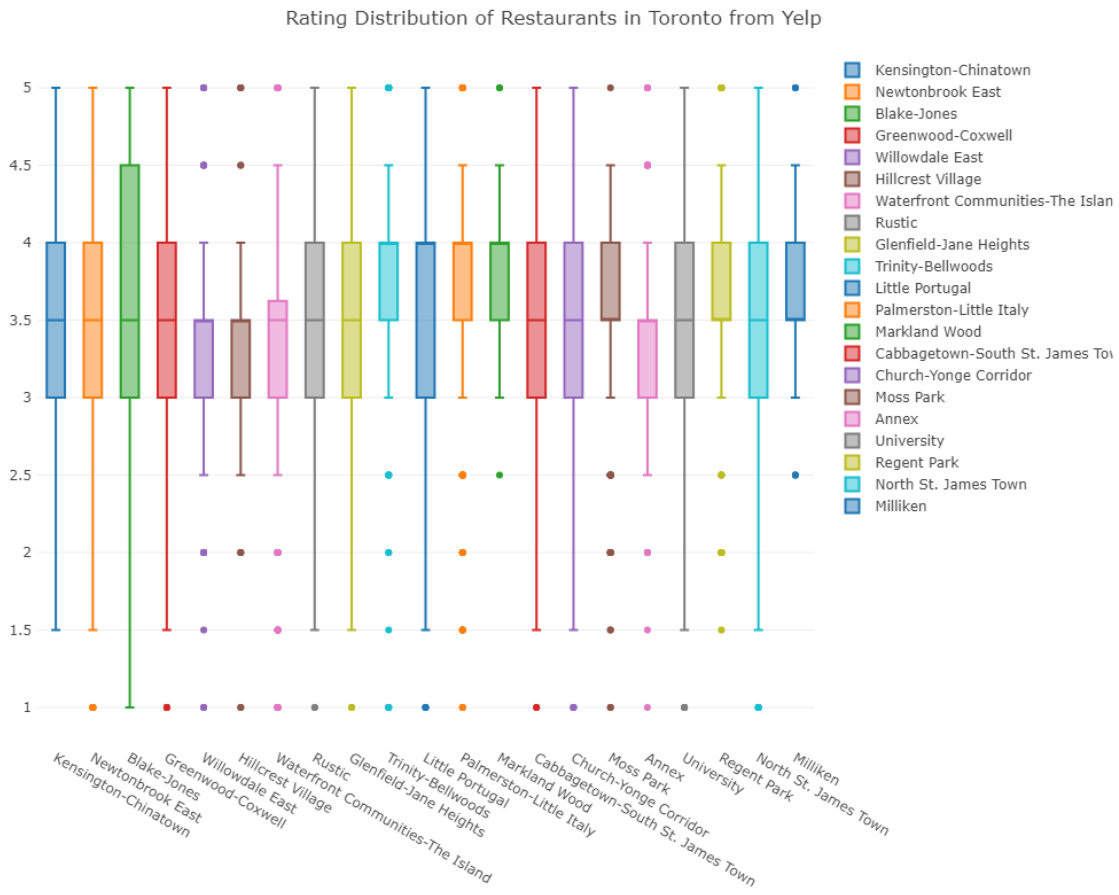


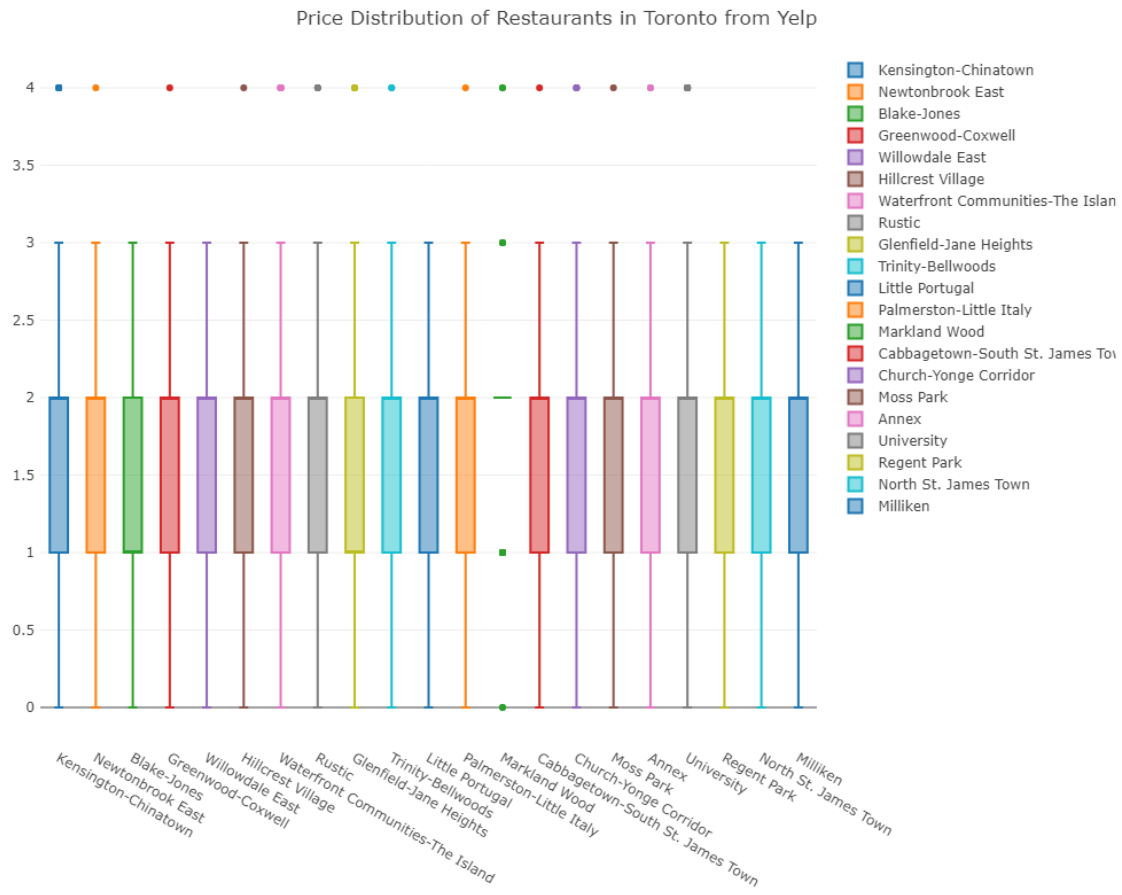
By Neighborhood





Boxplots of Ratings vs Price





Conclusions

Conclusions

- Toronto has a big income distribution
- Some neighborhoods are wealthier by a wide gap
- Toronto is quite multicultural, but Asian origins dominate
- Despite this, most popular restaurant type/category is…Coffee & Sandwiches!
- Kensington-Chinatown and Glenfield-Jane Heights have most places to eat
- Overall, price and ratings are somewhere in the middle
- By neighborhood, price level is consistent, but ratings are not (more expensive isn’t better)

Lessons Learned

Lessons Learned

- Finding good quality data is important in solving any problem
- Excel can be easier/more efficient for data cleansing and extraction than most other tools
- Watch out for Encoding!!!
- Be flexible-I allowed the data to tell me it’s story and decided to share it with the others.
- Data can serve more than one purpose
- Visualizing a large number of attributes/variables is decidedly quite hard to do effectively
- Pre-processing is paramount to preparing the data to be visualized, and almost always the hardest part