Utilizing Data Science to Satisfy Food Cravings

Peer-graded Assignment: Capstone Project The Battle of Neighborhoods

Choosing the best foodie city

- Overall, cities with diverse demographics have large food selections due to food “immigration”
- Report to give data for North America’s 5 largest cities’ demographics
  - May be interesting for food delivery companies
  - May be interesting for food travelers
- Data will help those compare city demographics

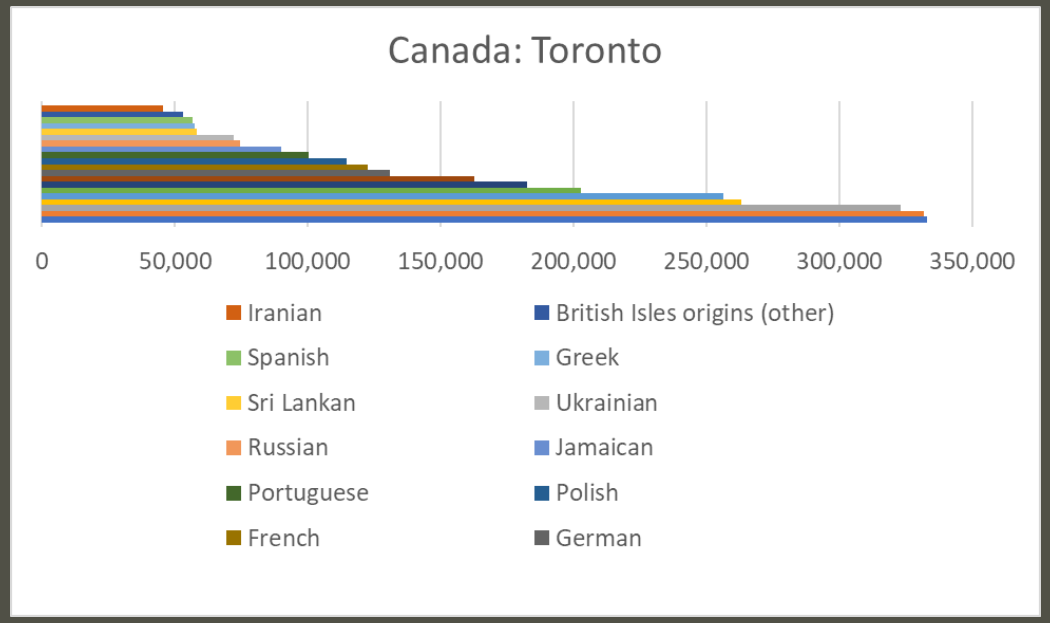
- Food has been the fueling backbone of society since the ancient days; some individuals travel the earth or pay exorbitant prices just to taste the flavors of the orient or exotic.
- Luckily in this modern day and age, these foodie explorers no longer even need to take a step out the door, thanks to delivery and restaurant selection.
- Naturally, cities with a higher population density bring forth people from all sorts of backgrounds, each with their own style of food (think China towns, little Italys etc!).
- New York alone has at least 8 different food districts with 8 million people, that’s a lot of demand… and data. So the target audience would be those looking to open a restaurant and those looking to eat good food, this is because they'd want to know where the dense populations are located in relation to what those populations are eating.

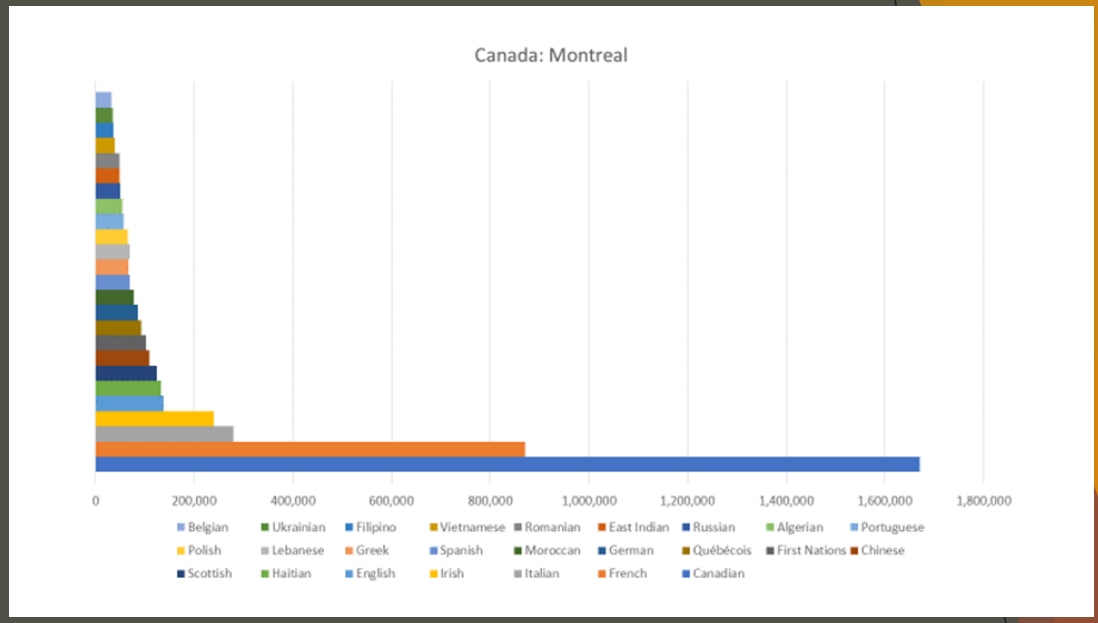
Data Sources

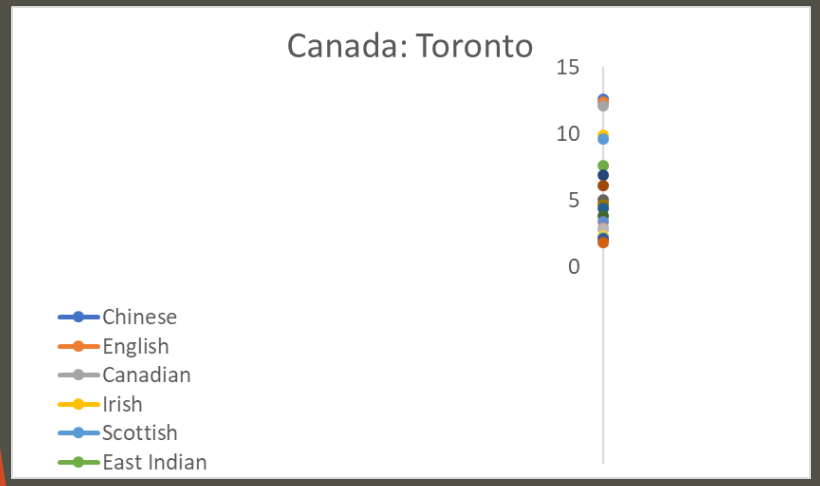
- Tables to be scraped from Wikipedia
  - Sources include Statistics Canada - Census Subdivision for Canada
  - Sources include U.S. Census Bureau for America
- Some data will be omitted
  - Age
  - Gender

With this in mind, the data report will analyze the population densities of well known urban cities, and compare to one another the different demographics allow future potential food or delivery companies to choose the optimal city for their hungry consumer food selection. Foursquare location data will inspect and explore the diversity of the data in correlation to the population density and location. For example data, we can see what food is commonly eaten (western, Chinese, etc), where they are commonly eaten (USA etc) and how many people typically live in that area (how many people live in that area, possibly by square km).

Data: Canada

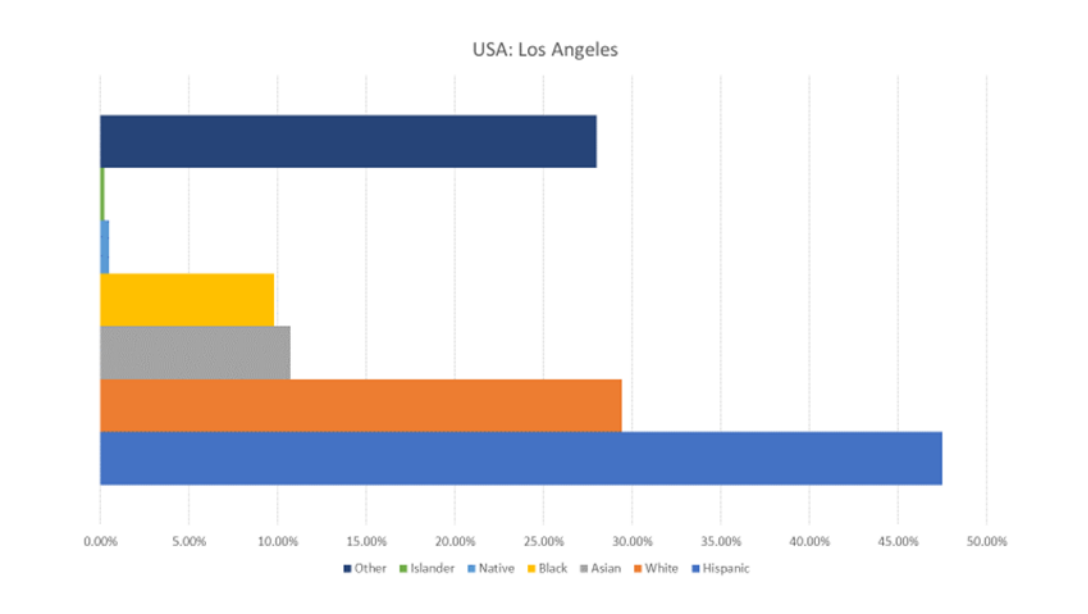


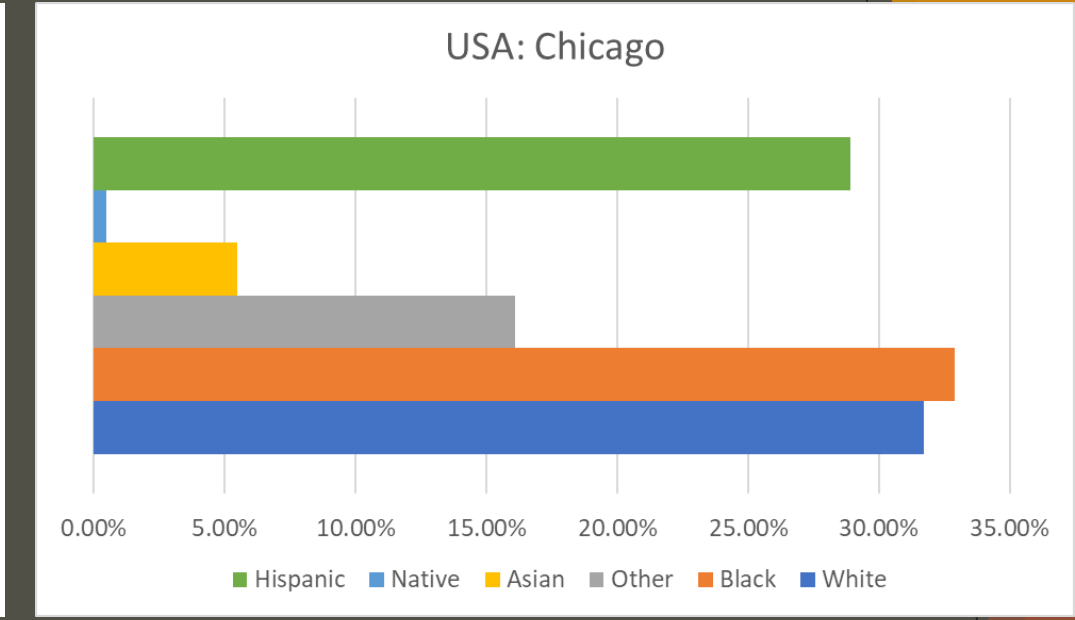


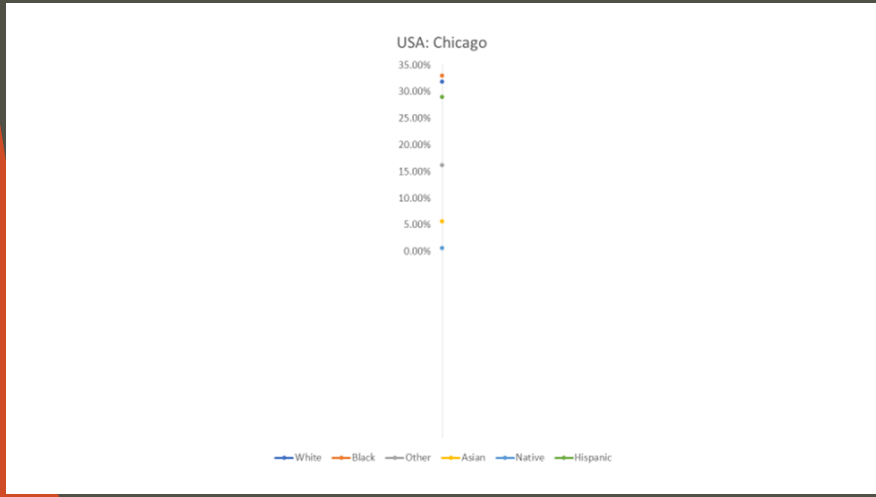


- 12 different demographic data, the median demographic being Spanish with +/- 200, 000 individuals
- Focused relationship on the lower spectrum ranging from +/- 2 - +/- 8
- Although the French is considered the median at approximately +/- 850,000 the mean is more towards the smaller demographic.

Data: USA







- Under 30% can be seen as the majority-minor outlier.
- USA has a larger spread of almost 12% per data set with larger demographics nearing 35%.

Conclusion: Canada & USA

- Clusters of data in +/- 4
- Restaurants with “exotic” food may wish not to open in Canada due to the clustering of minor demographics, and Caucasian restaurants may not wish to open in USA due to the overabundance of existing competition