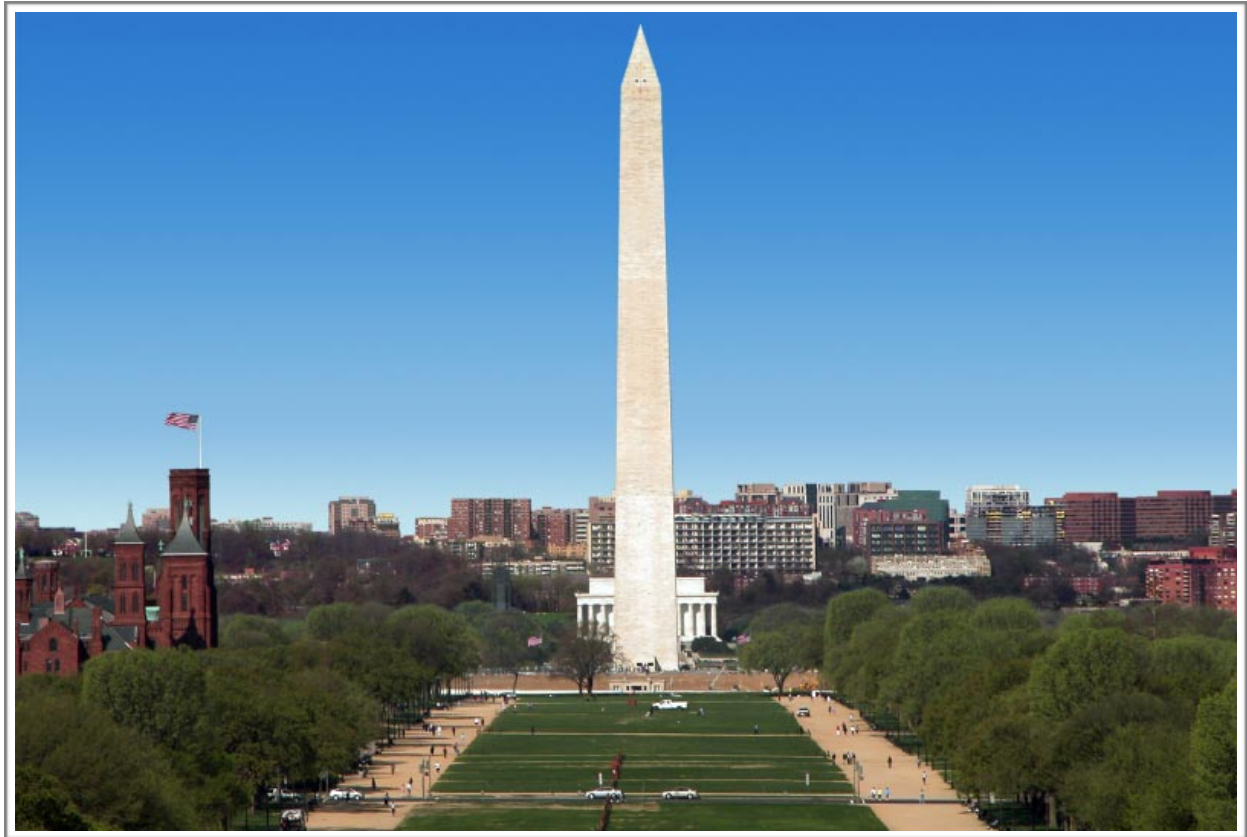# Coursera Capstone Project

From Washington DC to Toronto

Chris Petersen

Data Science Capstone

From Washington DC to Toronto

Introduction

Following the announcement of the results of the 2016 U.S. Presidential Elections, thousands of people made outrageous claims of moving to Canada. At the end of the day, many of those people did not move and the world continued on. With another election year right around the corner, I decided to examine what if I made those claims? Where in Canada would I move?

Data

My first thought was to move to Toronto, because I would be moving from one capital city (Washington DC) to another. But then I sheepishly realized Ottawa, not Toronto is the capital city of my neighbor to the north. So then I decided to move to a similarly-sized city. Washington DC's population is just under 700,000 with the surrounding metropolitan area housing approximately 5.7 million people. According to worldatlas.com, the largest cities in Canada are:

City	Population
Toronto	5.4 Million
Montreal	3.5 Million
Vancouver	2.3 Million
Calgary	1.2 Million
Edmonton	1.1 Million
Ottawa-Gatinau, Ontario/Quebec	.99 Million

Figure 1.1 Most Populated Canadian Cities

Toronto and the Washington DC area are both nearly the same size and have other similarities including the presence of multiple airports to travel from, top-ranked universities, and professional sports teams in the major US/Canada leagues. Being such a major city, lots of resources exist on the culture of the difference neighborhoods.

Because so many resources exist concerning the city of Toronto, I decided my study would be more interesting to compare what data could infer about a neighborhood (number of things, i.e. parks, museums, cafes, etc.) versus what professional sites say about the neighborhoods. Would raw data analysis make the same pick for me as the moving2canada.com website and other new resident sites? Do venues and their proximity make a neighborhood recommendable, or do people get too hung up on ambiance and culture?

Thus, my data comes from a comparison of foursquare data compared to open source information from mainstream sites used by tourists and new residents alike.

Methodology

Just as I was about to begin the study, I found a job in Downtown Toronto. This immediately made me focus my search in the Downtown Toronto area.

I sorted and cleaned the data and then used foursquare to discover the neighborhoods in the Downtown Toronto area. I identified the top five venues in each neighborhood and then used k-means clustering to cluster the 18 neighborhoods in Downtown Toronto to determine which area was most desirable.

Results

Ultimately, the top choice was Rosedale, Toronto.

Neighborhood	Most Common Venue	2nd Most Common	3rd Most Common	4th Most Common	5th Most Common
Rosedale	Park	Playground	Trail	Creperie	Dog Run
Cabbagetown, St. James Town	Coffee Shop	Pizza Place	Restaurant	Italian Restaurant	Pub
Church and Wellesley	Coffee Shop	Japanese Restaurant	Sushi Restaurant	Gay Bar	Burger Joint
Harbourfront, Regent Park	Coffee Shop	Bakery	Pub	Cafe	Park
Ryerson, Garden District	Coffee Shop	Clothing Store	Cafe	Cosmetics Shop	Middle Eastern Restaurant
Christie	Grocery Store	Cafe	Park	Baby Store	Coffee Shop

Figure 1.2 Top 5 venues in 6 neighborhood candidates

Rosedale's venues best matched what I was looking for in a neighborhood in Downtown Toronto. The top five venues were: Parks, Playgrounds, Trails, Creperies, and Dog Runs. Compared to the other locations, the venue locations portray a family-oriented neighborhood within the borough.

A close second was the neighborhood of Christie, containing several other top venues ideal for a family: grocery store, cafe, park, baby store, and coffee shop. Compared side-by-side this is how the top venues with frequency stack up:

Venue	Frequency
Grocery Store	0.19

Venue	Frequency
Cafe	0.19
Park	0.12
Diner	0.06
Italian Restaurant	0.06

Figure 1.3 Christie venues

Venue	Frequency
Park	0.5
Playground	0.25
Trail	0.25
Yoga Studio	0
Movie Theater	0

Figure 1.4 Rosedale venues

Discussion

Just as I was about to begin the study, I found a job in Downtown Toronto. This immediately made me focus my search in the Downtown Toronto area.

My venue searches using Foursquare set up only two "desirable" locations: Christie and Rosedale.

I compared this to the neighborhoods being advertised on moving2canada.com and they did not match. The site promoted: Yorkville, Downtown Core, Fashion District/Entertainment District, Queens Quay and Spadina. Of those, none of them made the cuts based off of venue locations. The study showed venues and facilities are but a small part of the location and while they can give you a good idea of what exists in an area, at most they should be used as a lunching pad for your investigation.

Conclusion

In conclusion, the venue locations point to moving to Rosedale while all of the sites pointed elsewhere. While I intended to do the study blind, I finally caved in and looked at what could possibly cause sites to not recommend Rosedale and this is what I found:

Figure 1.5 17,000 square-foot Georgian mansion in Rosedale, Toronto is listed for sale at C$22 million.

So maybe Rosedale is not the best choice after all.