VooDoo Doughnut Expansion

Applied Data Science Capstone Project

Week 2

Problem: Shall VooDoo Dougnut Open New Locations in Manhattan, NY and Downtown Toronto, ON?

Who is VooDoo Doughnut?

- Research revealed that downtown Portland lacked a single doughnut shop, so in 2003 they rented a hole-in-the-wall storefront scrunched between two Old Town nightclubs, joked to friends and family about being bent on "world doughnut domination," and opened Voodoo Doughnut. Their initial pastry offerings were a mix of the classic and the unconventional.

- Within a month of opening, Voodoo Doughnut received its first mention in the national press.

- To date, Voodoo Doughnut has been featured on The Travel Channel, The Cooking Channel, MTV, Good Morning America, The Today Show, The Tonight Show, Wheel of Fortune, Portlandia, Grimm, Leverage, The Amazing Race and Access Hollywood, and in print publications such as People, Conde Nast, Spin, GQ, Jane, Maxim, Food & Wine, and more.

- Voodoo Doughnut opened its second location in 2008 and as of 2019, has eight locations in five states, most recently opening a second location in Denver, Colorado. More often than not, the stores – open 24 hours a day, 7 days a week – have a line not just out the door, but around the block.

The information above was obtained from the VoodDoo Doughnut website www.voodoodoughnut.com

Stakeholders

- Voodoo Doughnut senior management and board of directors
- Senior management and board of directors will ultimately make the decision to expand to Manhattan and Toronto

Why Might Voodoo Doughnut Be Interested in Expanding to Manhattan and/or Downtown Toronto?

- Two years ago Voodoo Doughnut hired a new CEO. His background includes growing Mod Pizza from 1 location to more than 350 locations.

- 2 of their most recent locations are high profile: Universal theme park Orlando and Universal them park Hollywood

- Voodoo Doughnut has been expanding and is poised for future expansion

Data Sets to be used

- Foursquare location data shall be used for each location: Portland, Manhattan, Downtown Toronto

- The Foursquare location data will provide venue information for each neighbourhood in the cities being examined (E.g. coffee shops, bakeries, parks, …). It will show the most common venue, up through the 10th most common venue for that neighborhood.

Sample Manhattan Foursquare Location Data

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
6	Central Harlem	African Restaurant	Bar	American Restaurant	French Restaurant	Chinese Restaurant	Seafood Restaurant	Cosmetics Shop	Café	Library	Market
8	Upper East Side	Italian Restaurant	Exhibit	Art Gallery	Bakery	Coffee Shop	Boutique	French Restaurant	Hotel	Cocktail Bar	Gym / Fitness Center
10	Lenox Hill	Italian Restaurant	Coffee Shop	Sushi Restaurant	Pizza Place	Gym / Fitness Center	Café	Burger Joint	Gym	Cocktail Bar	Sporting Goods Shop
12	Upper West Side	Italian Restaurant	Coffee Shop	Wine Bar	Bar	Café	Cosmetics Shop	Vegetarian / Vegan Restaurant	Mediterranean Restaurant	Bakery	Indian Restaurant
13	Lincoln Square	Theater	Italian Restaurant	Café	Gym / Fitness Center	Plaza	Concert Hall	Performing Arts Venue	French Restaurant	Park	Indie Movie Theater

Sample Downtown Toronto Foursquare Location Data

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adelaide, King, Richmond	Coffee Shop	Café	Bar	Steakhouse	Breakfast Spot	Asian Restaurant	Cosmetics Shop	Restaurant	Hotel	Bakery
1	Berczy Park	Coffee Shop	Bakery	Cocktail Bar	Farmers Market	Steakhouse	Seafood Restaurant	Café	Cheese Shop	Beer Bar	Clothing Store
2	Brockton, Exhibition Place, Parkdale Village	Breakfast Spot	Café	Coffee Shop	Grocery Store	Caribbean Restaurant	Restaurant	Bar	Intersection	Bakery	Stadium
3	Business Reply Mail Processing Centre 969 Eas...	Light Rail Station	Pizza Place	Auto Workshop	Comic Shop	Recording Studio	Restaurant	Butcher	Burrito Place	Brewery	Skate Park
4	CN Tower, Bathurst Quay, Island airport, H...	Airport Service	Airport Lounge	Airport Terminal	Harbor / Marina	Airport	Airport Food Court	Airport Gate	Boutique	Bar	Boat or Ferry

Sample Portland Foursquare Location Data

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	97201	Breakfast Spot	Massage Studio	Farmers Market	Café	Park	Event Space	Fish Market	Fast Food Restaurant	Farm	Falafel Restaurant
2	97202	Food Truck	Mexican Restaurant	Garden	Brewery	Convenience Store	Farmers Market	Performing Arts Venue	Storage Facility	Auto Garage	Gym / Fitness Center
3	97203	Food Truck	Bar	Yoga Studio	Pizza Place	Park	Convenience Store	Coffee Shop	Rental Car Location	Falafel Restaurant	Mexican Restaurant
4	97204	Food Truck	Hotel	Coffee Shop	Restaurant	American Restaurant	Vietnamese Restaurant	Burger Joint	Pizza Place	Electronics Store	Thai Restaurant
5	97205	Cocktail Bar	Coffee Shop	Italian Restaurant	Bakery	Café	Furniture / Home Store	Thai Restaurant	Pub	Japanese Restaurant	History Museum

Other Data

- Demographic information is obtained from the USA Census Bureau and census data from the Canadian government. The demographic data is used to compare the similarities of the cities. Demographic information includes but is not limited to: population size, density, age, education, income

- The next slide contains example sample census data for Portland, Manhattan, and Central Toronto. More extensive data is available within the census data and that more extensive data will be utilized

- The number of existing donut shops in Manhattan is obtained from the New York City Health and Mental Hygience department. The number of donut shops in downtown Toronto is obtained from https://www.menupix.com/toronto/nc/1565/26/Downtown-Toronto-Donut-Shops/page_0

Sample Census Data

All Topics ▼	Q Portland city, Oregon ☒	Q New York County (Manhattan Borough), New York ☒
ℹ Population estimates, July 1, 2018, (V2018)	653,115	1,628,701
👤 PEOPLE		
Population		
ℹ Population estimates, July 1, 2018, (V2018)	653,115	1,628,701
ℹ Population estimates base, April 1, 2010, (V2018)	583,792	1,586,360
ℹ Population, percent change - April 1, 2010 (estimates base) to July 1, 2018, (V2018)	11.9%	2.7%
ℹ Population, Census, April 1, 2010	583,776	1,585,873

	Toronto Central, Ontario [Health region, December 2017]		
	Data quality Map Change geography		
	Total	Male	Female
Characteristic	Counts (unless ot		
Population and dwellings			
Population, 2016 [1]	1,232,258	...	...
Population, 2011 [1]	1,149,993	...	...
Population percentage change, 2011 to 2016	7.2	...	...
Total private dwellings [2]	611,053	...	...
Private dwellings occupied by usual residents [3]	568,161	...	...
Population density per square kilometre	6,412.6	...	...

How Data Will Be Used

- The target cities (Manhattan and Downtown Toronto) will be compared to Portland (existing location of VooDoo Doughnut)

- Foursquare location data can be used to determine type of businesses currently in the target cities as well as to compare them to the current location in Portland to look for opportunities for new locations

- Demographic data can be compared between Portland and the target cities using census data to identify similarities

Methodology

- Get the Toronto neighborhoods from Wikipedia. Screen scrape using BeautifulSoap. Put the data into a dataframe. Do cleanup on the data such that all rows are of the same format.

- Using the Foursquare data, get the top 100 venues are that within 500 meters of each Toronto neighborhood, plus the VooDoo Doughnut location in Portland, OR.

- Cluster the neighborhoods using k-means clustering machine learning. K-means clustering is used, since it is able to discover inherent groupings (clusters) of neighborhoods to determine which neighborhoods are most appropriate for a new VooDoo Doughnut location.

- Try to determine the optimal number of clusters using the Elbow Method and the Silhouette Method.

- Examine the clusters and determine which cluster the VooDoo Doughnut location resides.

Methodology continued

- For Manhattan, read in the New York City data from the json file at New York University into a data frame.

- Using the Foursquare data, get the top 100 venues are that within 500 meters of each Manhattan neighborhood, plus the VooDoo Doughnut location in Portland, OR.

- Cluster the neighborhoods using k-means clustering machine learning. K-means clustering is used, since it is able to discover inherent groupings (clusters) of neighborhoods to determine which neighborhoods are most appropriate for a new VooDoo Doughnut location.

- Try to determine the optimal number of clusters using the Elbow Method and the Silhouette Method.

- Examine the clusters and determine which cluster the VooDoo Doughnut location resides.

Results - Demographics

- There are currently 143 donut shops in Manhattan, with 129 of these being Dunkin' Donuts. Downtown Toronto has 72 donut shops, with 65 being Tim Horton's.

- Despite spanning two countries and being geographically dispersed, the three cities are more similar than one might suspect. The median age is similar. Each city has a high education level, with Manhattan having the highest. Manhattan also has the highest median household income. The one major difference is the population per square mile. Manhattan is much more densely populated than Portland or Toronto.
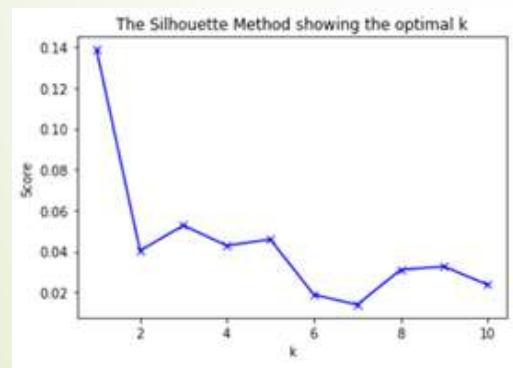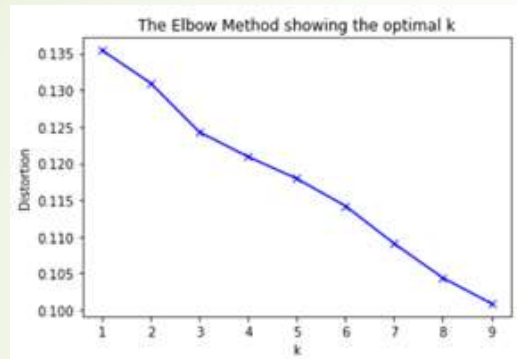
Results – Number of Clusters

- Five clusters were used with k-means clustering for both Manhattan and Toronto

- Two methods were tried to determine the optimal number of clusters for k-means clustering: Elbow Method and Silhouette Method. Neither method proved very useful.

- Manhattan - The Elbow Method essentially produced a straight line, indicating no optimal number of clusters. The Silhouette Method only indicated more than one cluster is needed, but was inconclusive beyond that. Therefore, the average number of clusters, five, were chosen.

- Toronto - Toronto has similar results to Manhattan, with the Elbow Method being virtually a straight line. The Silhouette Method gave a hint at two clusters, but it was not definitive. Therefore, five clusters were used.
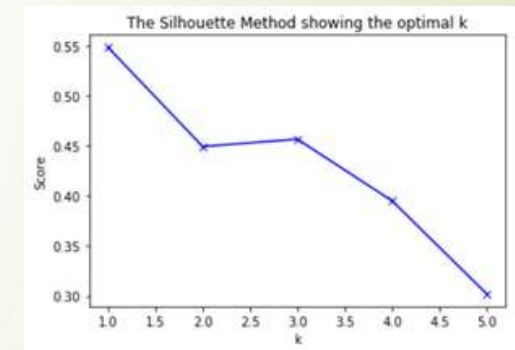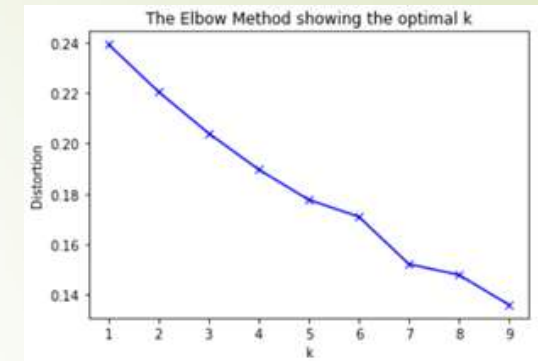
Results – Optimal Number of Clusters

■ Manhattan

The Elbow Method showing the optimal k

The Silhouette Method showing the optimal k

■ Toronto

The Elbow Method showing the optimal k

The Silhouette Method showing the optimal k

Results – Manhattan Cluster sample

- This is a sample of the most appropriate neighborhoods for a new location

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Chinatown	2	Chinese Restaurant	American Restaurant	Cocktail Bar	Vietnamese Restaurant	Optical Shop	Bakery	Spa	Salon / Barbershop	Asian Restaurant	Sandwich Place
6	Central Harlem	2	African Restaurant	Seafood Restaurant	Bar	American Restaurant	Cosmetics Shop	Fried Chicken Joint	French Restaurant	Chinese Restaurant	Café	Library
8	Upper East Side	2	Italian Restaurant	Exhibit	Bakery	Art Gallery	Coffee Shop	Hotel	Cocktail Bar	Gym / Fitness Center	French Restaurant	Mexican Restaurant
13	Lincoln Square	2	Theater	Plaza	Café	Italian Restaurant	Performing Arts Venue	Gym / Fitness Center	Concert Hall	Park	Indie Movie Theater	French Restaurant
14	Clinton	2	Theater	Coffee Shop	Gym / Fitness Center	American Restaurant	Italian Restaurant	Wine Shop	Gym	Hotel	Spa	French Restaurant
15	Midtown	2	Hotel	Theater	Sporting Goods Shop	Coffee Shop	Clothing Store	Spa	Bakery	Bookstore	Steakhouse	Food Truck
16	Murray Hill	2	Coffee Shop	Hotel	Sandwich Place	Japanese Restaurant	American Restaurant	Italian Restaurant	Gym / Fitness Center	Gym	Mediterranean Restaurant	Bakery
17	Chelsea	2	Coffee Shop	Italian Restaurant	Bakery	Nightclub	Ice Cream Shop	Hotel	Theater	American Restaurant	Seafood Restaurant	Market

Results – Toronto Cluster sample

- This is a sample of the most appropriate neighborhoods for a new location

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	East Toronto	0	Health Food Store	Trail	Pub	Department Store	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop
1	East Toronto	0	Greek Restaurant	Coffee Shop	Ice Cream Shop	Italian Restaurant	Furniture / Home Store	Bookstore	Restaurant	Brewery	Bubble Tea Shop	Pub
2	East Toronto	0	Park	Sushi Restaurant	Ice Cream Shop	Board Shop	Sandwich Place	Brewery	Burger Joint	Fast Food Restaurant	Burrito Place	Fish & Chips Shop
3	East Toronto	0	Café	Coffee Shop	Bakery	Italian Restaurant	American Restaurant	Comfort Food Restaurant	Bookstore	Seafood Restaurant	Sandwich Place	Brewery
5	Central Toronto	0	Gym	Hotel	Pizza Place	Clothing Store	Convenience Store	Sandwich Place	Breakfast Spot	Food & Drink Shop	Park	General Entertainment
6	Central Toronto	0	Sporting Goods Shop	Coffee Shop	Clothing Store	Yoga Studio	Rental Car Location	Diner	Spa	Dessert Shop	Salon / Barbershop	Burger Joint
7	Central Toronto	0	Pizza Place	Dessert Shop	Sandwich Place	Coffee Shop	Gym	Italian Restaurant	Café	Sushi Restaurant	Park	Seafood Restaurant
9	Central Toronto	0	Pub	Coffee Shop	American Restaurant	Liquor Store	Sports Bar	Restaurant	Bagel Shop	Supermarket	Sushi Restaurant	Fried Chicken Joint
11	Downtown Toronto	0	Coffee Shop	Restaurant	Chinese Restaurant	Pizza Place	Park	Pub	Market	Café	Italian Restaurant	Bakery
12	Downtown Toronto	0	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Gay Bar	Restaurant	Hotel	Gastropub	Men's Store	Café	Mediterranean Restaurant

Discussion

- VooDoo Doughnut should pursue opening new locations in Manhattan and Toronto. While barriers to entry are high, due to the both Manhattan and Toronto being dominated by a single donut shop chain, there still is opportunity for Voodoo Doughnut.

- Krispy Kreme, which has opened donut shops in both Manhattan and Toronto, despite the competition from the large, entrenched donut chains. Krispy Kreme started major expansion from a much smaller number of locations in the 1990s, to a nationwide (USA) and international presence today. Krispy Kreme can be seen as a model of successful donut shop expansion and a model VooDoo Doughnut may seek to emulate.

Conclusion

- Both Manhattan and Toronto are appropriate locations for VooDoo Doughnut expansion.

- Donuts are already a strong part of the culture in Manhattan and Toronto.

- The neighborhoods in Manhattan and Toronto most closely resembling the existing Portland VooDoo Doughnut location have been determined.

- While both Manhattan and Toronto are each dominated by a single large donut chain, Krispy Kreme has been able to open donut shops in both cities, clearly indicating expansion to Manhattan and Toronto is possible.