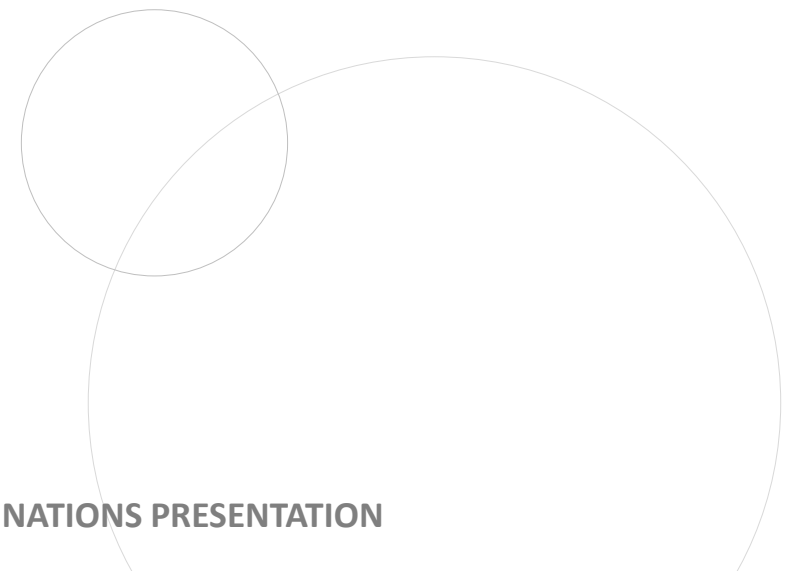# TORONTO DESTINATIONS

AN APPLIED DATA SCIENCE CAPSTONE PROJECT BY MORALES, J. 2020

WHY THIS PROJECT?

Planning a trip to a foreign city can be daunting

Often equipped with limited budget, time, and experience of the foreign city, visitors want to make the most of their stay.
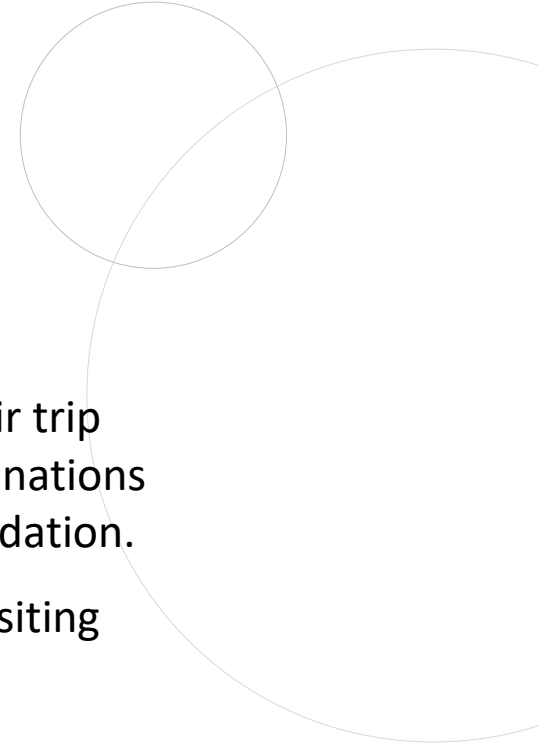
But how?

THE QUESTIONS

Which neighbourhoods are teeming with Toronto destinations for visitors?

Which neighbourhoods offer the most diverse experiences for Toronto visitors?

WHO IS THIS FOR?

- Toronto visitors who are planning their trip would be interested to see what destinations are closest to their place of accommodation.

- Torontonians playing tour guide for visiting family, friends, or acquaintances.

- Tour guides giving city-round tours.

- Travel agencies or consultants giving suggestions to clients.

- Anyone who is curious about Toronto.

ABOUT THE DATA

- Toronto dataset with 103 cases and 3 features (Postal Code, Borough, Neighbourhood). Scraped from Wikipedia.

- Geospatial dataset from the Applied Data Science Capstone Course week 3 learning materials.

- Venues dataset pulled from Foursquare.

- Cleansed, amalgamated, sliced.

METHODOLOGY

Data preparation and exploration

Data about neighbourhoods with postal codes starting with 'M' or Toronto neighbourhoods was mined from Wikipedia. Missing values were imputed. The cleansed Toronto dataset was combined with geo codes dataset containing latitudes and longitudes.

Data was visually explored by creating a map of Toronto and superimposing the neighbourhoods on top of the map. Data was further explored by slicing the dataset and obtaining only boroughts with the string 'Toronto', leading us to focus on the downtown-ish areas. Simple clustering was also applied to divide these neighbourhoods into areas.

More data preparation, machine learning techniques

The venues information was obtained from Foursquare and combined with the sliced dataset of Toronto neighbourhoods. Now we are working with Toronto neighbourhoods in the downtown area with knowledge of the venues in those neighbourhoods. Initial exploration showed a total of 1729 cases.

Clustering was applied throughout the process. Clustering helps us see similarities of elements within a cluster and dissimilarities of elements across cluster. It is also intuitive when used in conjunction with Folium / mapping.

RESULTS

Top 10 neighbourhoods with greatest number of venues

Neighbourhood	
Adelaide, King, Richmond	100
Design Exchange, Toronto Dominion Centre	100
St. James Town	100
Ryerson, Garden District	100
Harbourfront East, Toronto Islands, Union Station	100
Commerce Court, Victoria Hotel	100
First Canadian Place, Underground city	100
Stn A PO Boxes 25 The Esplanade	96
Chinatown, Grange Park, Kensington Market	90
Church and Wellesley	86

Top 10 neighbourhoods with most diverse venues
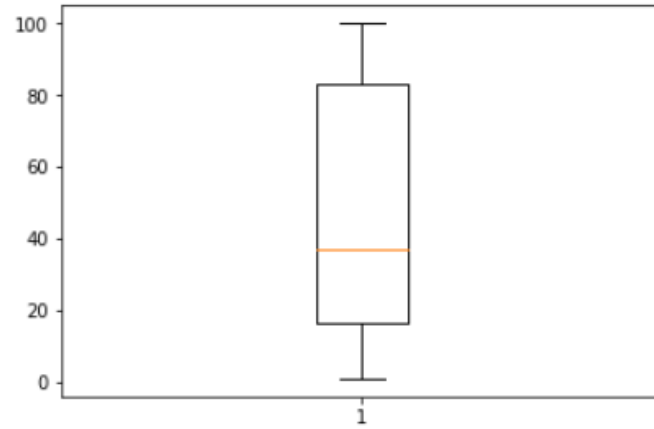
| Neighbourhood	
Ryerson, Garden District	65
Adelaide, King, Richmond	63
Stn A PO Boxes 25 The Esplanade	59
Harbourfront East, Toronto Islands, Union Station	59
St. James Town	58
Church and Wellesley	58
First Canadian Place, Underground city	54
Chinatown, Grange Park, Kensington Market	53
Commerce Court, Victoria Hotel	52
Design Exchange, Toronto Dominion Centre	49

RESULTS

Neighbourhood distribution with greatest number of venues

count	39.000000
mean	44.333333
std	35.742746
min	1.000000
25%	16.500000
50%	37.000000
75%	83.000000
max	100.000000

Neighbourhood distribution with most diverse venues

count	39.000000
mean	29.051282
std	19.999933
min	1.000000
25%	14.000000
50%	26.000000
75%	48.000000
max	65.000000
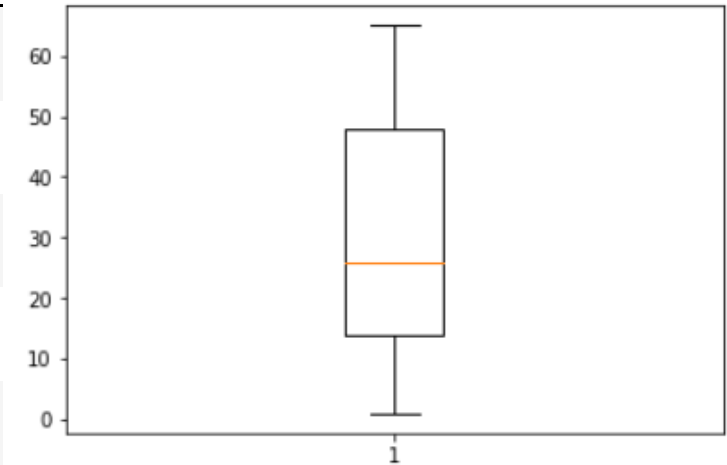
Appendices

Clustering Toronto Destinations - Project Report

https://github.com/jamiemorales/project-clustering-toronto-destinations/blob/master/clustering-toronto-destinations-report.ipynb

Clustering Toronto Destinations - Project Code

https://github.com/jamiemorales/project-clustering-toronto-destinations/blob/master/clustering-toronto-destinations-code.ipynb

Clustering Toronto Destinations - Project Abstract

https://github.com/jamiemorales/project-clustering-toronto-destinations/blob/master/clustering-toronto-destinations-abstract.ipynb

Clustering Toronto Destinations - Project Presentation

https://github.com/jamiemorales/project-clustering-toronto-destinations/blob/master/clustering-toronto-destinations-presentation.pdf

Clustering Toronto Destinations - Project Repo

https://github.com/jamiemorales/project-clustering-toronto-destinations

THANK YOU