**Segmenting and Clustering Neighbourhoods in Toronto**

Applied Data Science Capstone Project

Amira Kaskasoli

The Problem

A client is looking to move to Canada, specifically Toronto and would like advice on which neighbourhoods he should narrow his search to based on given criteria.

Toronto is a very large city spread across 630 square kilometres, and it is known for being a hub for finance, business, arts, culture and entertainment. With a population of over 6 million people and 27.5 million tourists visiting annually, it is considered to be in the top 5 liveable cities in the world, making it is a very attractive place to consider living.

Property costs in Toronto are among the highest in Canada. The client understands this and is willing to pay for the right location.

The client would like his new home to be near coffee shops, exercise facilities and his favourite type of cuisine: Italian food. House price is also a consideration for the client but at this point he is looking for information on which neighbourhoods to consider.

The client is looking to move ASAP so current data on Toronto is appropriate to review.

Using data analytics methodology and machine learning techniques such as clustering, this project will provide an answer to the client’s question regarding the most appropriate neighbourhood to live in in Toronto.

The Data

To solve the client’s problem, this report will be using data from various sources. The original datasource will be a Wikipedia page listing the various boroughs and their corresponding neighbourhoods and postcodes in Toronto. A second datasource containing geospatial data of Toronto will then be inputted and merged with the dataframe of neighbourhoods. The resulting dataframe will contain details of neighbourhoods and boroughs, alongside respective coordinates of each location. Folium will be used to visualise the data on a map of Toronto. Finally, Foursquare location data will be used to explore the neighbourhoods in Toronto and make informed statements about the data to provide to the client.