

IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

The Battle of Neighborhoods – A new gym in Toronto

Prepared by Maciej Skoczny



WEEK 4 – INTRODUCTION

Background:

Our client, Jane Doe, is looking for a suitable location in Toronto to open her new gym. She has experience in the fitness industry and runs a handful of fitness studios in Canada's other regions. She would like to expand her Canadian business to Toronto; however, she doesn't know the city well.

Business problem:

What Toronto neighborhood/borough should Jane open her new gym in?

Data:

- List of Toronto neighborhoods and boroughs
- Geographical coordinates of Toronto, including neighborhoods and boroughs
- Gym venue data

Proposed solution:

We'll use the data science methods to determine which neighborhood of Toronto has the smallest number of gyms, and this could be the best place for Jane to open her newest venue. We'll use the Jupyter Notebook on the Anaconda platform and Python programming language. We'll use the following data science techniques:

- Data extraction
- Data scrapping
- Data segmentation
- Data clustering
- Data visualization

Data sources:

- List of Toronto neighborhoods and boroughs:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- Geographical coordinates of Toronto, including neighborhoods and boroughs: GeoCoder pack
- Gym venue data - <https://foursquare.com/>

WEEK 5 – FINAL REPORT

Methodology:

Results:

Discussion:

Conclusion: