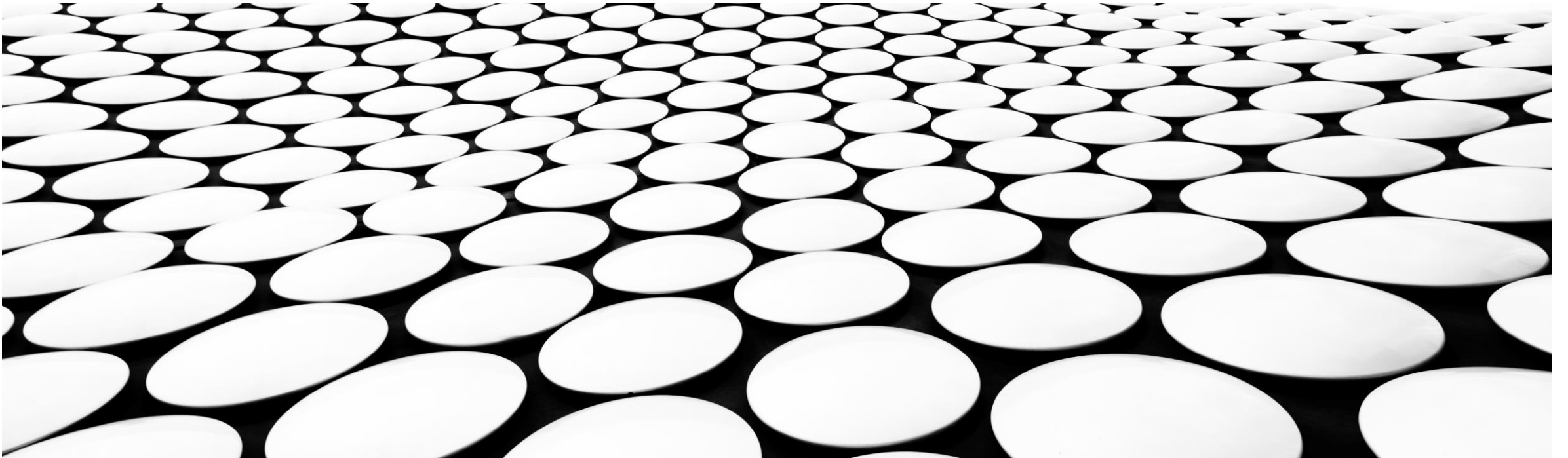# BATTLE OF THE NEIGHBOURHOODS – WEEK ONE

BRITTANY EASTON FOR IBM – APPLIED DATA SCIENCE CAPSTONE

AGENDA

- Business Problem

- Data

BUSINESS PROBLEM

- The client is a black belt in Brazilian Jiujitsu, and is looking to open up her own martial arts studio in the Greater Toronto Area (GTA). She has asked for a recommendation on where to start her business. She would like a shortlist of 3 locations provided to her so that she can explore them as her primary options.

The client asks that the criteria for the recommendation be based on the following priorities:

- First and foremost, she is focused on quickly creating a profitable business. Therefore the region she selects should have lower competition to allow for faster growth. The area she selects to start her business needs to have a smaller ratio of studios relative to the number of people so that she can easily compete as a newcomer.

- She is interested in starting her business in an area where she can make a big impact and have the greatest opportunity to grow a large membership base. An area with a higher population and a higher density will mean the business is closer and accessible to more people.

- Since doing martial arts can be an expensive hobby, she would like to make sure that the region where she sets up her business is an affluent one. A region where more households are making more than $100,000 mean that the people who live there can more easily afford Brazilian Jiujitsu for the whole family.

DATA

Data Source:	Foursquare	Wikipedia	AggData
Data Contents:	**Location & Venue data**	**GTA Regions**	**Coordinates by FSA**
What insight does this data provide?	The level of competition in a particular region.	The definition of all Greater Toronto Area regions.	The geographic coordinates of each FSA within the Greater Toronto Area
What data will be used?	The count of, and the location of martial arts studios in the GTA.	The regions included within the Greater Toronto Area	FSA, Latitude, Longitude, Province, Place Name
How will the data be used?	In determining population density per studio, and the number of high income households per studio in each region.	Used as a metric for inclusion in the analysis of where to open the business.	Map the Greater Toronto Area regions
Example Data	Action & Reaction MMA 485 McNicoll Ave, North York M2H 3H9	Durham region is in the GTA	L3R: Markham, ON 43.8479 -79.3288

DATA

	Canada Revenue Agency	Statistics Canada	Postal Codes in Canada
Data Source:	Canada Revenue Agency	Statistics Canada	Postal Codes in Canada
Data Contents:	Household Income by FSA	Population by FSA	FSAs by County
What insight does this data provide?	The ability of the households of the GTA to afford a recreational hobby in the martial arts.	How much room for growth exists for the business.	Which FSAs lay within which geographical region.
What data will be used?	The number of households per FSA which make a particular range of income.	Geographic Name, Population (2016), Total Private Dwellings (2016)	FSA, Type (indicates region, city or county name)
How will the data be used?	Proportion of households in each FSA making over $100,000 in income will inform a conversion rate.	Population size and density will inform metrics for growth	To define the geographical region of each FSA.
Example Data	FSA $100,000 to $149,999 A0A 1900	FSA POP (2016) DWELL (2016) A0A 46,587 26,155	FSA TYPE K1C OTTAWA

DATA

Data Source	Data Contents	Data Access Through
Foursquare	Location & venue data	Foursquare API
Wikipedia	GTA Regions	Greater Toronto Area – scrape the table "Municipalities in Greater Toronto Area and Related CMAs"
AggData	Coordinates by FSA	Complete List of Canada Postal Codes (FSAs)
Canada Revenue Agency	Household Income by FSA	Individual Tax Statistics by Forward Sortation Area (FSA) – 2019 Edition (2017 tax year)
Statistics Canada	Population by FSA	Population and Dwelling Count Highlight Tables, 2016 Census
Postal Codes in Canada	FSAs by County	Ontario Postal Codes List