# BUSINESS PROBLEM

Globally, the rates of obesity-related chronic diseases have become a public health concern. A modern diet is mainly processed food high in total fat, saturated fat, and simple sugars, which increases the risk of developing obesity [1]. In 2010, over 27 million (approximately 11%) adults were diagnosed with heart disease. Nearly 20 million Americans had cancer as of 2010, while 21 million Americans had diabetes mellitus. Most people are transforming into distinct lifestyles and dietary changes to reduce the high risk of chronic disease caused by modern nutritional patterns. Lately, doctors recommend that patients should live a healthy life and not depend on reactionary medicine. Vegetarianism is the best diet to enhance health. The study of [2] stipulates that vegetarians warranted a reputation of being overall healthier people. They are often associated with living more healthily and mindfully by participating in more regular exercise, avoiding fattening or sugary foods, and avoiding alcohol and smoking. Obesity is linked to lower rates of a vegetarian diet, heart disease, and diabetes mellitus (DiMarino, 2013). Figure 1 illustrates some of the food consumed by vegetarians.



Figure 1: Humankind Ventures Ltd, 2020

## Problem statement

There is a limited number of vegetarian restaurant, therefore, there is an opportunity to create employment through vegetarian businesses. Globally, unemployment is a huge challenge. According to [4] internationally, young people are three times unemployed as compared to adults, and over 350 million young people are not engaged in education, employment, or training. Therefore, there is a need to create employment. This study is significant since segmentation and clustering neighbourhoods’ information provides entrepreneurs with knowledge, for example (entrepreneurs will know their competitors locations and areas where there is a need for their products).

## Targeted audience

The skills gained on the segmentation and clustering neighbourhoods assignment can contribute immensely to the body of knowledge since the study findings can be used as a reference when conducting research. The target audience of the project is academics and non-academic researchers. Academics can use the information to educate learners and enhance the quality education. Tourists can study this kind of assignment to understand the area of interest before visiting Toronto. The entrepreneurs can use segmentation and clustering neighbourhoods’ information as a prototype when exploring new opportunities to open a thriving business in Toronto.

# DATA

In this assignment, segmenting and clustering the neighborhoods in the city of Toronto were explored. The initial stage followed to implement this assignment project was to use the Notebook to build the code to scrape the following Wikipedia page, [https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.) the head method was applied to display the columns and rows of the data. Columns which were not relevant were dropped from axis = 1 using the method drop (). The codes used to clean data are illustrated in the notebook provided.

The following instructions from Coursera course work were followed to create the dataframe:

* The dataframe will consist of three columns: PostalCode, Borough, and Neighborhood
* Only process the cells that have an assigned borough. Ignore cells with a borough that is Not assigned.
* The BeautifulSoup package was used for web scraping
* More than one neighborhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the neighborhoods separated with a comma as shown in row 11 in the above table.
* If a cell has a borough but a Not assigned neighborhood, then the neighborhood will be the same as the borough.
* Clean your Notebook and add Markdown cells to explain your work and any assumptions you are making.
* In the last cell of your notebook, use the .shape method to print the number of rows of your dataframe.
* Geocoder package was used to determine the geographical location of the neighbourhoods and venue data pertaining to vegetarian restaurants

# REFERENCES

1. DiMarino, A., 2013. *A Comparison of Vegetarian Diets and the Standardized Western Diet in Nutrient Adequacy and Weight Status* (Doctoral dissertation, The Ohio State University).
2. Berich, H.M., 2015. Knowledge and Perceptions of Vegetarian Diets Among College-Aged Students.
3. Humankind Ventures Ltd, 2020. How to Avoid Deficiencies on a Vegan Diet? [How to Avoid Deficiencies on a Vegan Diet - Forth Life (forthwithlife.co.uk)](https://www.forthwithlife.co.uk/blog/how-to-avoid-deficiencies-on-a-vegan-diet/)
4. Hanna, A.L., 2014. The Global Youth Unemployment Crisis: Exploring Successful Inifiafives and Partnering with Youths. *Durham, NC: Sanford School of Public Policy, Duke University. Retrieved August*, *14*, p.2018.