**THE BATTLE OF NEIGHBORHOODS**

|  |
| --- |
| **Barath Srinivasan** |

**INTRODUCTION/BUSINESS PROBLEM:**

This project is an attempt to leverage Four Square Location Service Provider API to address a unique problem applying learnings from earlier practical sessions. Commonly experienced problem of moving to a new apartment to a different location among considered areas was constructed.

Mr. John, 65 years old, has decided to move to a new apartment in Toronto with his wife, to one among the shortlisted Boroughs:

* Downtown Toronto
* East Toronto
* West Toronto
* Central Toronto

As with old age, they have some pre-existing medical conditions and they need a location in the close vicinity of Hospitals for their treatment to be continued. Mr. John is a **Heart Patient** and his requirement is that the new location should be nearer to **heart speciality clinics**. Mrs. John has been visiting **Toronto General Hospitals** for past few years and she needs atleast one among these hospitals accessible from the new location. Apart from this condition, they are flexible enough to handle other requirements.

They need our help to **suggest them an apt location** as near as possible according to above conditions, so that they can move without any hesitation and also continue to take their treatments in this old age without any hassle.

**DATA USED:**

The Toronto neighborhood data is extracted from their 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)

This has all the information we need to explore the neighborhoods in Toronto.

It is critical to convert this data into a pandas dataframe, in a structured format, clean it and make it inferable to carry out the analysis to explore and decide a prompt location according to Mr. and Mrs. Johns conditions.

*Structure of Data used:*

* The above data has three columns – Postal Code, Borough and Neighborhood of Toronto. We can see many NA values in them and they need to be handled. We processed only the cells with an assigned Borough and removed the rows if Borough Column had NA values. More than one neighbourhood can exist in one postal code area. For example, in the table on the Wikipedia page, we notice that **M5A** is listed twice and has two neighbourhoods: **Harbourfront**and **Regent Park**. These two rows were combined into one row with the neighbourhoods separated with a comma.

Table

Description automatically generated

**Figure 1. Sample of cleaned dataset – From Wiki**