**Business understanding**

Introduction

Toronto is not only the financial hub and largest populated city of Canada, It is also one of cities which is vouched for real estate investments in the world. Toronto real-estate market rise around 20% every year.

Many entreprenaurs and even mid sized investors prefer to set their rental business in neighbourhoods of Toronto. The neighbourhood of the business would impact ROI and Hence the investors would prefer to set up their properties near Universities, Tourist Places, Job hub and better school districts.

Context

Mr A is a mid size real estate investor and is looking to invest in residential properties to rent in Toronto.He is mainly looking into Condo properties as he thinks temporary tenants prefer economical condos in main city more than others as they provide sufficient space and less maintainence.

He wants us Data scientists team ,to explore various neighbourhoods of Toronto and understand which neighbourhood would give him best ROI.

The main objective of Mr A is that , the properties should not be vacant for more than 3 months and also he preferes tenants who dont live more than 2 years in a stretch in same condo for maintainence purposes. With this given criteria , we would target neighbourhoods with universities , school districts . He is also open to rent the property to temporary visitors . Hence a center location with easy public accessiblity to other venues is pereferable.

Summary :

- Which neighbourhood has universities ,schools and other colleges that would attract floating population to rent houses in that area?

- What is the public travel accessiblity rating of the neighbourhood? How far is it from bus terminal, train stop or others?

**Analytic Approach**

To answer the above business query we would ,

- We would get the list of neighbourhoods of Toronto city.

- We would consider only those neighbourhoods which have university/colleges/schools within a specific radius of distance.

- We would check the public transit stops/stations in the area.

- From the result , we will choose the best list of neighbourhoods based on their aggregated score .

- Further we will cluster the above neighbourhoods list and recommend the cluster that hat Mr A can

invest on

In the process , we plan to use python and pandas for retrieving and preparing the dataset.

For clustering we will be using popular k-means algorithm. For generating visualization maps we will use Folium. Also , we are using Four square API ( public location service) to get the venues.