

# **FINDING THE BEST COMMUNITIES FOR NEWCOMERS TO CALGARY, ALBERTA, CANADA**

**ARPITA CHAKRAWARTI**

**NOVEMBER 29, 2020**

## **1.0 INTRODUCTION**

### **1.1 BACKGROUND**

Calgary, AB is a beautiful city in western Canada very well known as the centre of Canada's oil industry. Per the fact sheet provided [here](#),

- Calgary is currently the fourth largest city in Canada with the highest immigrant population next to Toronto, Vancouver and Montreal.
- By 2020, Calgary's total immigrant population is estimated to reach almost half a million.
- The Philippines, India, and China continue to be the lead source countries for immigrants to Calgary.

A very large percentage of the immigrant population belongs to working ages. It is often a struggle for new immigrants and newcomers to find the right neighbourhoods and communities, which can provide the right amenities, safety, affordable housing and good schools. It is therefore advantageous for newcomers to be aware of some general information about the various residential locations of this city.

### **1.2 PROBLEM**

Data that might help determine which neighborhoods or communities may be preferable might include house prices, crime rates, population, popular amenities or venues. This project aims to find groups of communities which are similar in nature, based on this data.

### **1.3 INTEREST**

The following may be interested in the analysis –

- New immigrants and newcomers to Canada
- The city of Calgary
- And others who may be interested in conducting a similar analysis of other cities

## **2.0 DATA ACQUISITION AND CLEANING**

### **2.1 DATA SOURCES**

All data for Calgary city, its crimes and property assessments can be found [here](#), [here](#) and [here](#) . Data for venues can be found using the foursquare.com Places API – described [here](#).

### **2.2 DATA CLEANSING**

Data downloaded from multiple sources have been combined into a single table with the communities as the key. Only residential communities have been considered in the analysis.

Some data cleansing has been required from each of the files, to retain the most recent data and to include only the communities of interest.

Features selected for the analysis include

- Total crimes by community
- Total population by community
- Median value of assessed property value by community
- Mean popularity for all categories of venues found by community