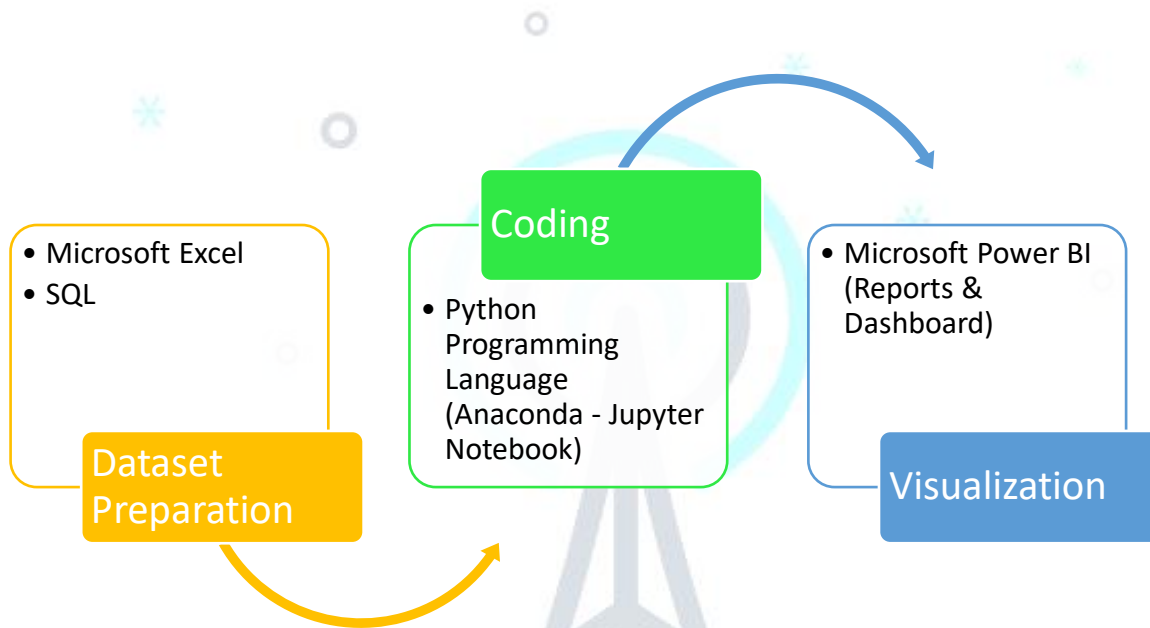


## METHODOLOGY

### SOFTWARE & TOOLS USED

- ✚ Dataset Preparation: Microsoft Excel & SQL
- ✚ Coding: Python Programming Language is used in Anaconda (Jupyter Notebook)
- ✚ Visualization: Microsoft Power BI is used for Final Reports & Dashboard



### OVERALL METHODOLOGY EXPLAINED IN BREIF

#### ➤ Collection of Data:

This project consists of two datasets:

- Spectrum dataset collected from Spectrum Management System Data which covered up cell towers information/data. (Government of Canada, n.d.)
- City Population collected from Statistics Canada. (Census Statistics Canada, n.d.)

#### ➤ Validating the datasets:

After getting both datasets initial analysis is done and information is validated so, research & analysis could be carried in right direction.

➤ Reverse-Geo Coding:

During my initial analysis I realised that Spectrum dataset was missing on proper location attributes. To deal with issue, I have gone with reverse-geo coding which helps to get accurate location using latitude and longitude. After the process completion, I was able to obtain attributes such as “City, Zip, Province” which help me build connection between two datasets and further analysis.

➤ MS SQL:

Once datasets are completed, I have added both datasets to MS SQL under one database named it as CMDB (Central Management Database). I used this technique to improve overall data consistency, better storage and sharing data across different platforms.

➤ Python Programming Language (Anaconda - Jupyter Notebook):

Moving forward I have started my analysis in Python for which Jupyter Notebook is used. The various activities performed using Python Programming Language as follows:

1. Connecting with MS SQL & Importing the Dataset from SQL:

To build connection with MS SQL SERVER - SQL ALCHEMY Library is used which helps to directly import datasets from MS SQL Server.

2. Data Preprocessing:

Then I start with data processing to understand my data in better manner. Under this python coding covers up:

- ✓ Checking shape of dataset
- ✓ Understanding the dataset
- ✓ Describe the dataset
- ✓ Checking duplicates and removing duplicates
- ✓ Checking NULL values

### 3. Data Manipulation:

Under data manipulation I covered up the following:

- ✓ Dropped the unnecessary attribute from datasets
- ✓ Changed data types for attributes in datasets
- ✓ Renaming Specific Column Values for better understanding & grouping
- ✓ Merged both datasets using Joins

### 4. Creating a Model in Python (Jupyter Notebook):

Finally, model is ready..! which is made using various calculation in python. Under model my coding covered up the following:

- ✓ Dropped the unnecessary attribute from datasets
- ✓ Performing Cell Tower Count & displaying under new column
- ✓ Using for loop
- ✓ Used lambda Function
- ✓ Performed calculations in Python
- ✓ Changed data type for attribute

### 5. Saving dataset after Data Manipulation:

Here, I have saved to the following:

- ✓ Saved dataset after Data Manipulation to MS SQL Server
- ✓ Saved dataset after Data Manipulation to CSV

➤ Pre-Modeling in Excel:

In MS Excel model I have done the count of towers in particular city and then done a lookup to get city's population. Then used excel if formula to get Green Zone and Red Zone based on criteria 1 tower/antenna could cover up 1000 people in the given area. The new attributes as expressed:

- **Green Zone** = The area has enough towers to cover up people in that region as shows in Figure 1- Green Zone.
- **Red Zone** = The area doesn't have enough towers to cover up people in that region as shows in Figure 2- Red Zone.

Prov	City	Cell Tower Cou	Populatio	Zone	Population Coverag	Recommen
Quebec	Varennnes	21	20994	Green	6	NA
Quebec	Pincourt	15	14558	Green	442	NA
British Columbia	West Vancouver	43	42473	Green	527	NA
Ontario	Welland	53	52293	Green	707	NA
New Brunswick	Douglas	7	6154	Green	846	NA
Ontario	Mapleton	14	13130	Green	870	NA
Ontario	Woodstock	42	40902	Green	1098	NA
Newfoundland and Labrador	Carbonear	6	4838	Green	1162	NA
British Columbia	Comox	16	14828	Green	1172	NA
Quebec	Amos	14	12823	Green	1177	NA
Quebec	Val-Shefford	8	6711	Green	1289	NA
Alberta	Ardrossan	9	7678	Green	1322	NA
Quebec	Lavaltrie	15	13657	Green	1343	NA

Figure 1 - Green Zone

Prov	City	Cell Tower Cou	Populatio	Zone	Population Coverag	Recommen
Quebec	Repentigny	40	84965	Red	-44965	45
British Columbia	Kincolith	6	49216	Red	-43216	43
Ontario	Barrie	111	147829	Red	-36829	37
British Columbia	North Cowichan	9	44451	Red	-35451	35
British Columbia	Port Coquitlam	105	139284	Red	-34284	34
Ontario	Cambridge	96	129920	Red	-33920	34
Quebec	Brossard	52	85721	Red	-33721	34
Ontario	Newmarket	51	84224	Red	-33224	33
British Columbia	Victoria	60	91,867	Red	-31867	32
Ontario	St. Thomas	30	61707	Red	-31707	32
Ontario	Peterborough	51	82094	Red	-31094	31
Quebec	Chateauguay	21	50815	Red	-29815	30
Quebec	Saint-Hyacinthe	31	59614	Red	-28614	29

Figure 2 - Red Zone

- Population Coverage = This attribute in my modeling shows how many people can cover up more which is basically the numbers shown in positive and numbers in negative shows how many people are losing the signal or struggling with signal because of not enough antenna's/ cell towers.
- Recommended = This attribute covers up the number of recommended towers to get losing signal area to green zone. Otherwise, NA shows as the practically is already in green zone with good signal.

➤ Python (Jupyter Notebook) Visualization:

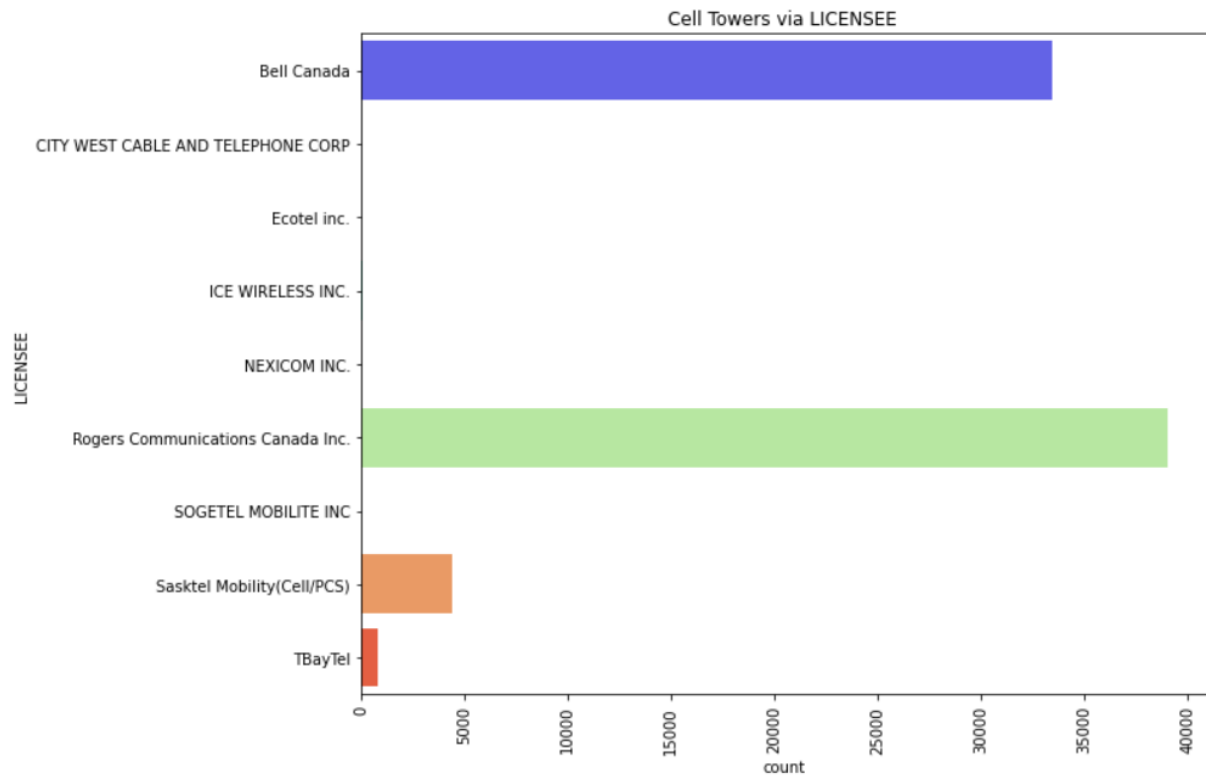


Figure 3 - Cell Towers via LICENSEE

From the above bar graph, we can see there are 8 major telecom services operators (Bell Mobility, TBayTel, Rogers Communication, Sogetel Mobilite, Nexcom Inc, Ice Wireless Inc, Ecotel Inc and Saskatel Mobility). From the above listed network providers Rogers Communication has highest number of tower and Bell Mobility stands at 2<sup>nd</sup> position. Hence these two (Rogers Communication and Bell Mobility) are the dominant network provider in Canada.

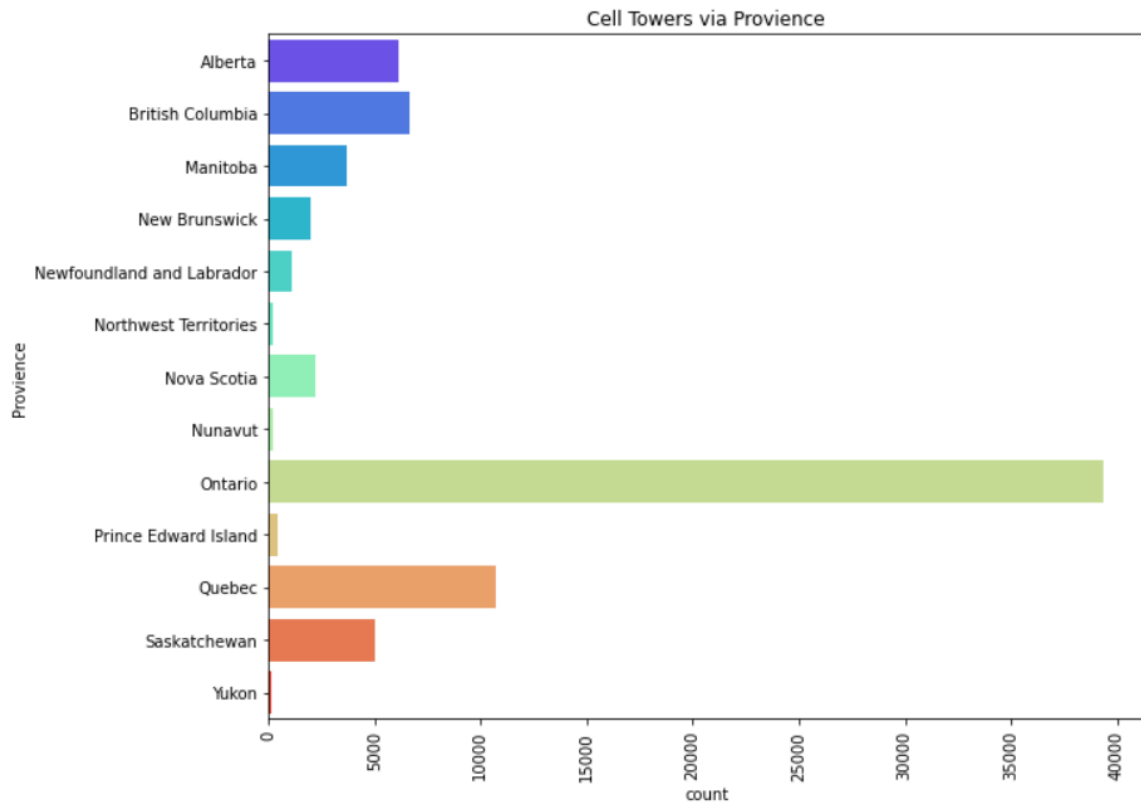


Figure 4 - Cell Towers via Province

In the above visualization we tried to figure out the number of towers each province has. So, Ontario ranked first in number of towers followed by Quebec. Saskatchewan holds third position whereas Nunavut, Yukon, Prince Edward Island and Northwest territories has similar number of towers.

➤ MS Power BI Visualization:

After Modeling, I have created dashboards in Power BI to make easy understanding on my Modeling. The Figure 5 & 6 dashboard shows that how many Cell Towers are in that area. The search can be City, Province, or Licensee. Other than dashboard covers up Max Distance the cell towers can cover and max area of reach

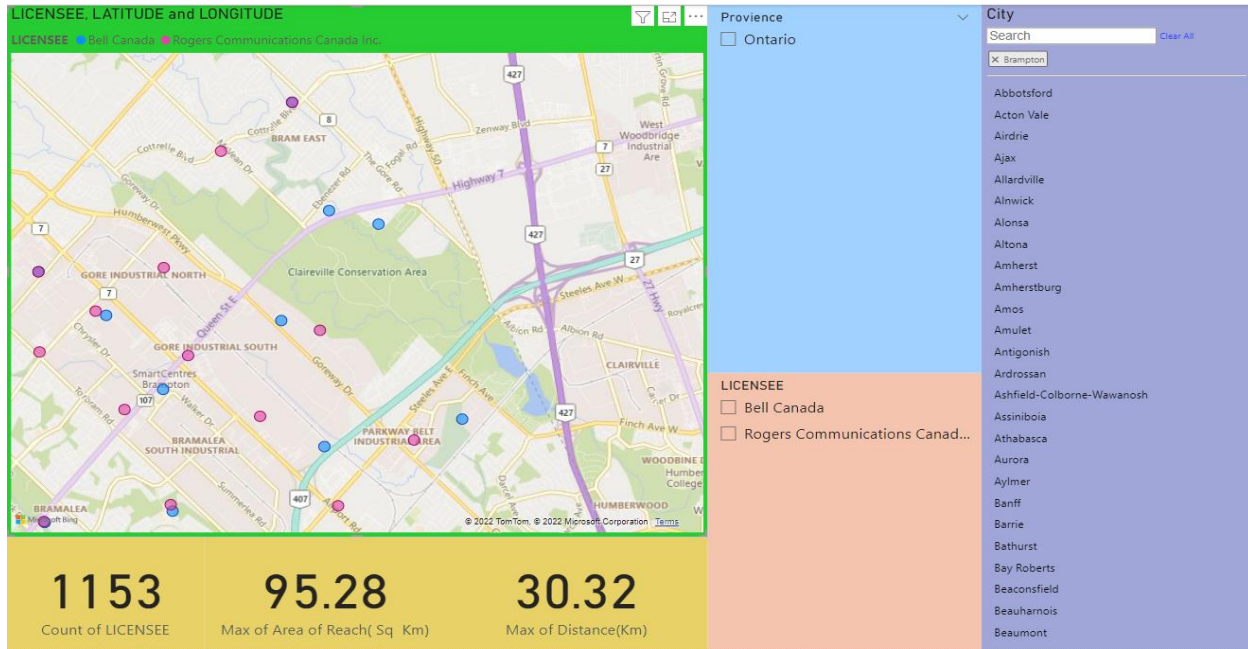


Figure 5

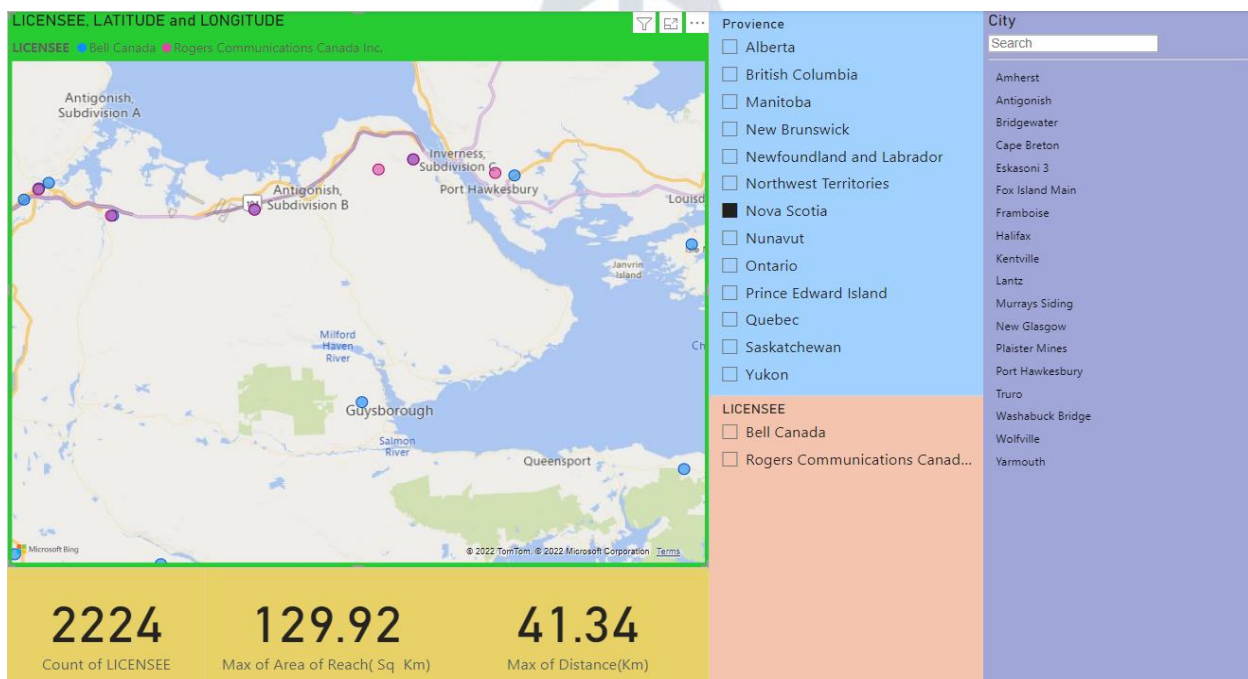


Figure 6



The second dashboard covers up Figure 7 & 8 dashboard illustrates how many Cell Towers are Recommended in that area. The search can be City, Province, Licensee and Zone. Other than dashboard covers up Max Distance the cell towers can cover along side population of that City and population coverage is covered or how much it is missed (which shows in negative).

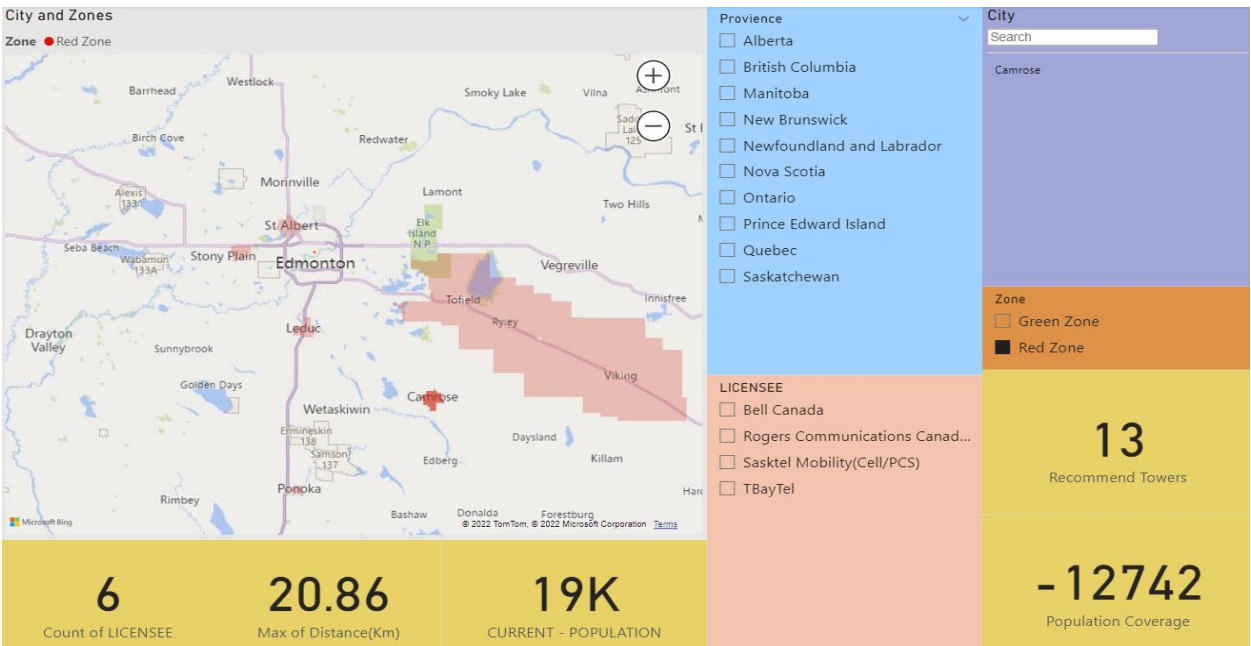


Figure 7

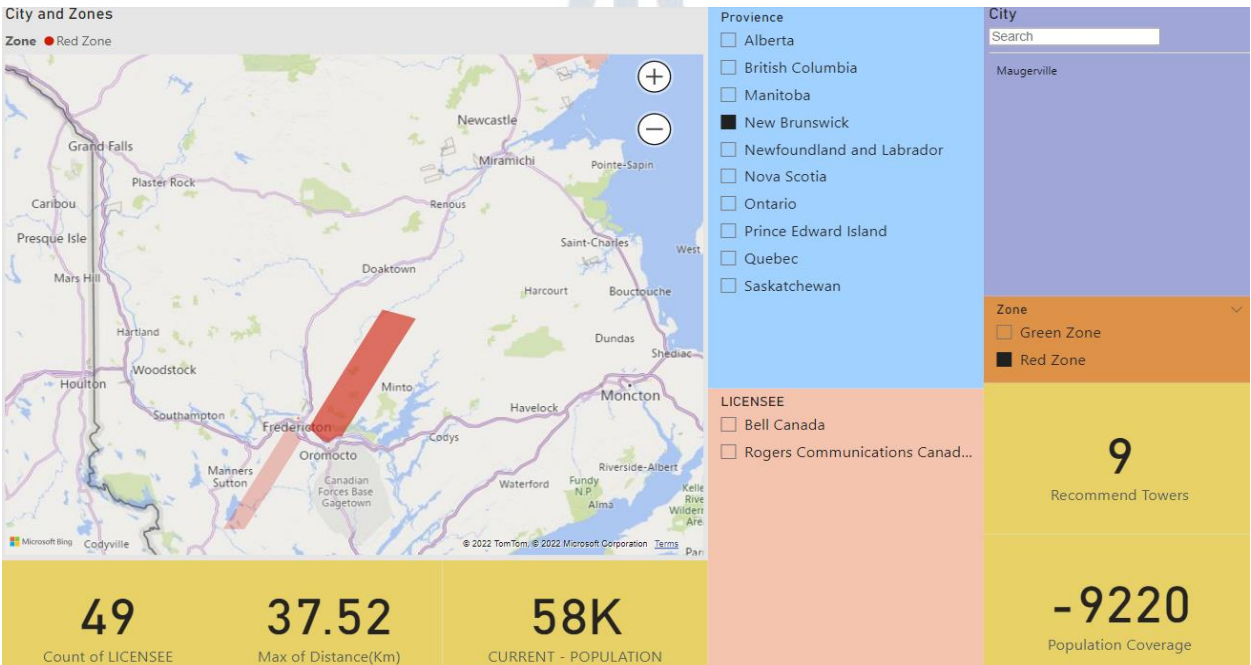


Figure 8

GitHub : <https://github.com/AkshayJTheAnalyst/Telecommunication-Networks-in-Canada.git>