

OPENING A NEW HOTEL IN CANADA

NITIN VERMA
JUNE 2019

INTRODUCTION

Hotels play a great role in attracting the tourists and generating revenue. Hotels can also result in increase in employment rate. Various shops are also set near the hotels to attract the tourists. . However if a great number of hotels exist in an area, this results in competition, which can result in increasing the fare for a room in the hotels. Due to which the number of tourists visiting that place decreases.

BUSINESS PROBLEM

Studying the distribution and determining which places have high number of hotels already present in that area and determining which areas would result in greater revenue if a new hotel is setup in that area using machine learning and data science methodology.

DATA ACQUISITION

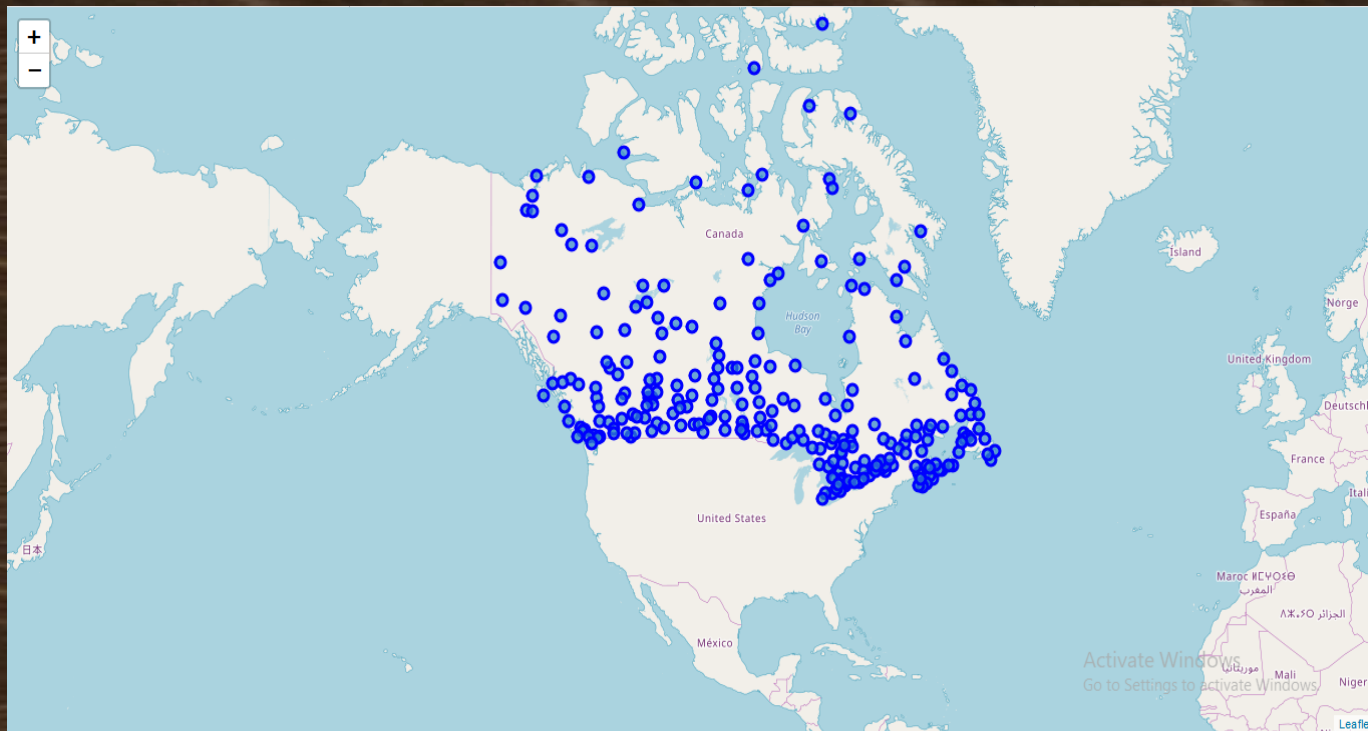
- The dataset can be acquired from <https://simplemaps.com/data/world-cities>.
- The attributes present are- name of the city, country, latitudes, longitudes, iso2, iso3, city_ascii, admin_name, capital, population, id.
- Required attributes- country, city, latitudes and longitudes.

METHODOLOGY

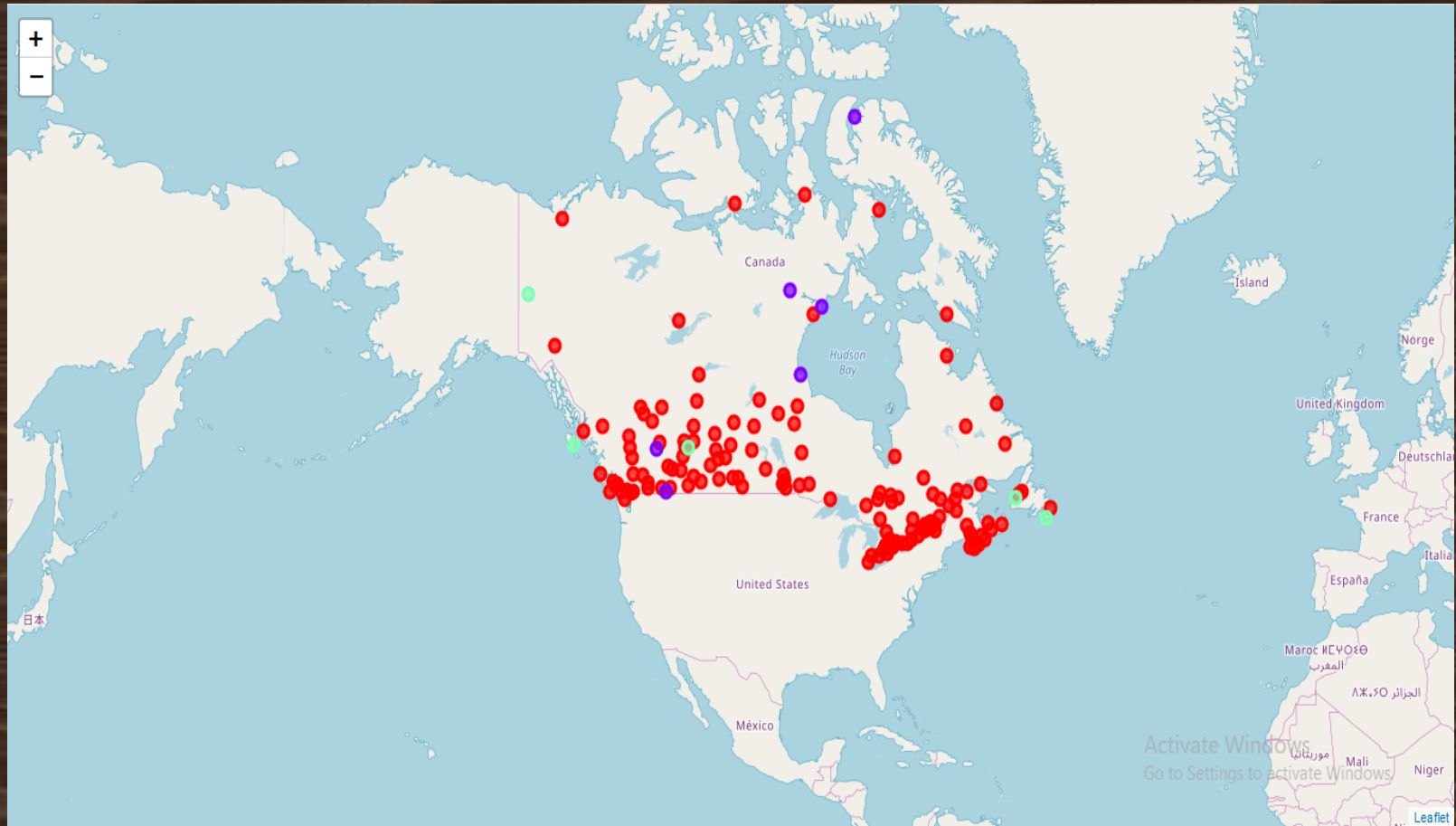
1. Acquiring dataset.
2. Cleaning the dataset.
3. Accessing the coordinates
4. Visualizing the dataset.
5. Using Foursquare API to access the VenueCategory.
6. Selecting hotels from the VenueCategory.
7. Use K-means cluster algorithms to cluster the result
8. Visualize the results
9. Analyze the results.

RESULTS

Visualizing the cities:



Visualizing the clusters:



OBSERVATIONS

On the basis of clusters the following observations were made:

- High number of hotels exists in cluster 2.
- Moderate number of hotels exists in cluster 1.
- Low number of hotels exists in cluster 0.

INFERENCE

Based upon the results that were obtained it is safe to say that setting up a new hotel at the location defined in cluster 2 will be very risky as very high concentration of hotels are already present. Setting up a new hotel there will result in competition. It is highly recommended that new hotel should be setup at either location mentioned in cluster 1 or cluster 0, but preferably cluster 1.