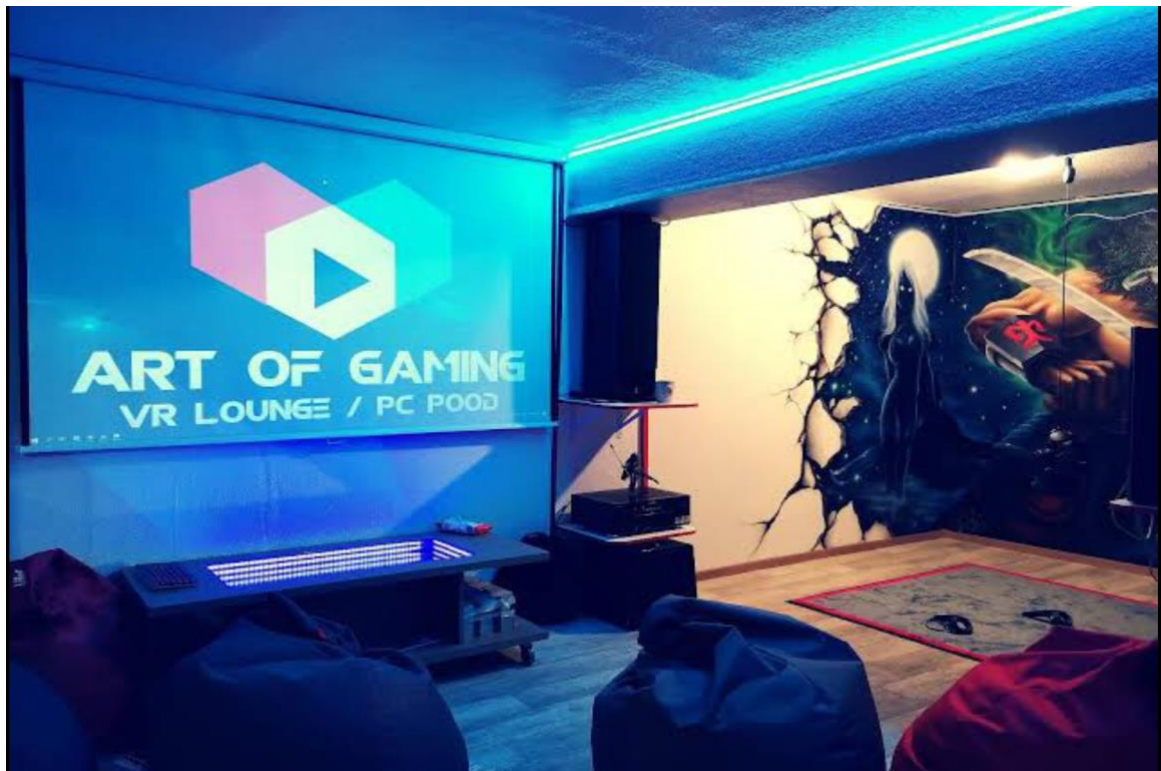


Opening a New Gaming Centre in Toronto, Canada.

Coursera Capstone Project



Vamshi Venkat Mallepalli

September 2019

TABLE OF CONTENTS

1. Introduction
2. Background
3. Business Problem
4. Description of the data
5. Solution outline
6. Target Audience
7. Methodology
8. Results
9. Recommendation and Conclusion

Introduction

The particular project is based upon a business problem where an entrepreneur is looking for a location to open a Gaming centre in Toronto, Canada. The main objective of this project is to find the favourable location to launch a gaming centre which can be a decent valuable one and by utilising particular unsupervised machine learning algorithms and foursquare API, I recommend the entrepreneur to open the gaming centre at a regional location.

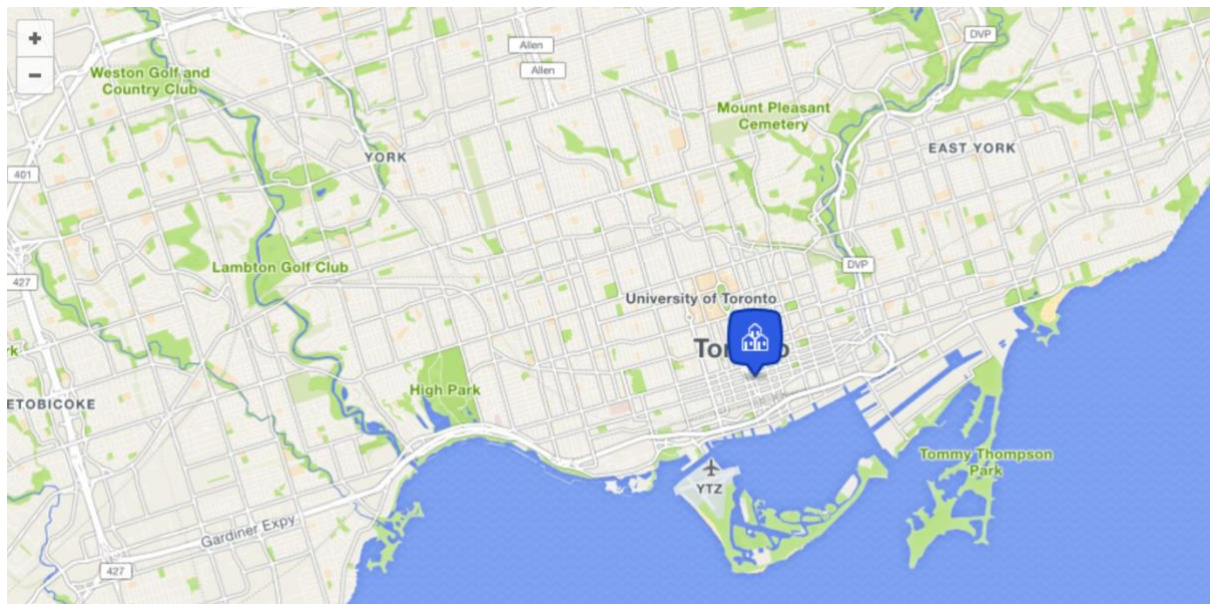


Fig:1. Map of Toronto, Canada from a foursquare map from the web.

Background

An arcade game or coin-op game is a coin-operated entertainment machine typically installed in open organisations, for example, restaurants, bars and amusement arcades. Most arcade games are computer games, pinball machines, electro-mechanical games, salvation games or trading games. the Sony PlayStation and Microsoft Xbox enhanced in their graphics and game- play capability and reduced in expense. So the advancement of gaming centres and arcade designers and entrepreneurs are stooped to a high rate to date.

Business problem

A speculative strategy for an entrepreneur who is gaming fancier wishes to open a gaming centre in Toronto, Canada. The main objective of this project is to recommend a specific place where the gaming centre can be a growing business in the Toronto area. However there are some arcade centres in Toronto, the entrepreneur is inquisitive to open the biggest arcade and gaming centre in a specified location. With this main objective keeping in my mind, uncovering a location to open the arcade centre/gaming centre in Toronto area is a speck challenging and to build the project to the entrepreneur to help his business to be a booming one in a specific location.

To get started right away with this project, we need some data.

- Data from websites
- Location data

Description Of Data

- Scraping of Toronto neighborhoods via Wikipedia.
- Getting latitudes and longitudes data of these neighborhoods via Geo-coder packages / with a .CSV file
- Using Foursquare API to get venue data related to neighbors.

NOTE: Here for finding the neighbourhoods that are most favourable for opening a gaming centre, we determine the venue data related to gaming venues. the purpose behind the selection of gaming venue is because maximum of the teenagers and children are more feasible to be playful and more delighted to get into various kinds of games, moreover, it can be relieving tension time for students and also develop their curiosity in computer knowledge and applications.

Discussing the solution to the problem.

Starting with

- Getting the list of neighbourhoods in Toronto extracting the list from wikipedia page.
- Web scraping by pandas HTML table scraping, pull tabular data directly from the web page into the data frame.
- Getting the coordinates to utilize foursquare to pull list of venues near the neighbourhood with the CSV file provided by IBM.
- Visualize Toronto map using folium package to verify the coordinates.
- Use Foursquare to pull the venues, names, categories, latitudes, longitudes with the help of client ID and secret ID.
- Check the unique categories from the venues.
- Analyze each neighborhood by grouping the rows by neighborhood taking the mean of frequency of occurrence of each venue category.
- After clustering venues, perform clustering methods by k-means clustering, an unsupervised machine learning algorithm.
- Based on the result, recommend the ideal location to open a gaming centre.

Target Audience

Schools, teenagers

Methodology

A. Business Understanding

- Comprehending the business problem and obtaining the appropriate solution to it is a challenging phase, identifying clearly the features and the factors that are acceptable to the solution that we present itself with. Evaluate all the targets and make a decision according to the desired problem.

B. Analytical Approach

- For the given business problem, we need to present itself with a solution which algorithm to be employed in such business problem circumstances and how can we solve, what are the possible suggestions and recommendations can be made. for this project problem, I use the unsupervised machine learning algorithm which is k-means clustering where it is used to group the neighbourhoods into clusters.

C. Data Requirements

- The given project is based upon the location data set so, I scraped the Toronto neighbourhoods data set and the location coordinates data set.

D. Data Collection

- Obtaining the list of neighbourhoods in Toronto, Canada by extracting the list of neighbourhoods from Wikipedia page.
("https://en.wikipedia.org/wiki/list_of_postal_codes_of_Canada:_M")
- Following step was the web scraping with a pandas HTML table, scraping method is easier to pull the tabular data promptly from the webpage into the data frame.
- The second data set is the location data of Toronto, Canada neighbourhoods, collected the data set as.CSV file from the IBM team.

E. Data Understanding

- By viewing the data frame, checking all the columns and rows, any missing values in the data frame and then we are ready for data cleaning.

- For the location data sets, we use the foursquare API to pull a list of venues near these neighbourhoods.

F. Data Preparation

- By collecting all the above data sets, cleaned the data, merged and grouped the two data frames.
- Subsequently gathering all these coordinates, I visualised the map of Toronto to check the coordinates of Toronto data.
- Following, I used the foursquare API to pull the list of top 200 venues within 500 metres radius, in order to do this, I have created a foursquare developer account ID and API key to pull the venues data.
- With this data, I can furthermore check how many unique categories that I can obtain from these venues.

G. Data Analysis

- Analysed each neighbourhood by grouping the rows by neighbourhood and taking the mean of frequency of occurrence of each category.
- Then, getting the number of venues in the neighbourhood.
- Here, I have opted for one category, as I have very less number of gaming categories in the neighbourhood, so I opted for all the related venues to gaming so that it is more feasible to get the desired location.
- Now, finding the clusters of all related venues in the neighbourhood.
- Then, grouped all the venues jointly incorporating neighbourhoods into 3 clusters by running k- means an unsupervised machine learning algorithm.

H. Data Visualisation

- After running the clustering method, I visualised clusters on the map with the help of folium software with a radius of 5.
- Examined the number of clusters and concluded the resulted location might be a good location to open the gaming centre in Toronto, Canada.

Results

Clustering

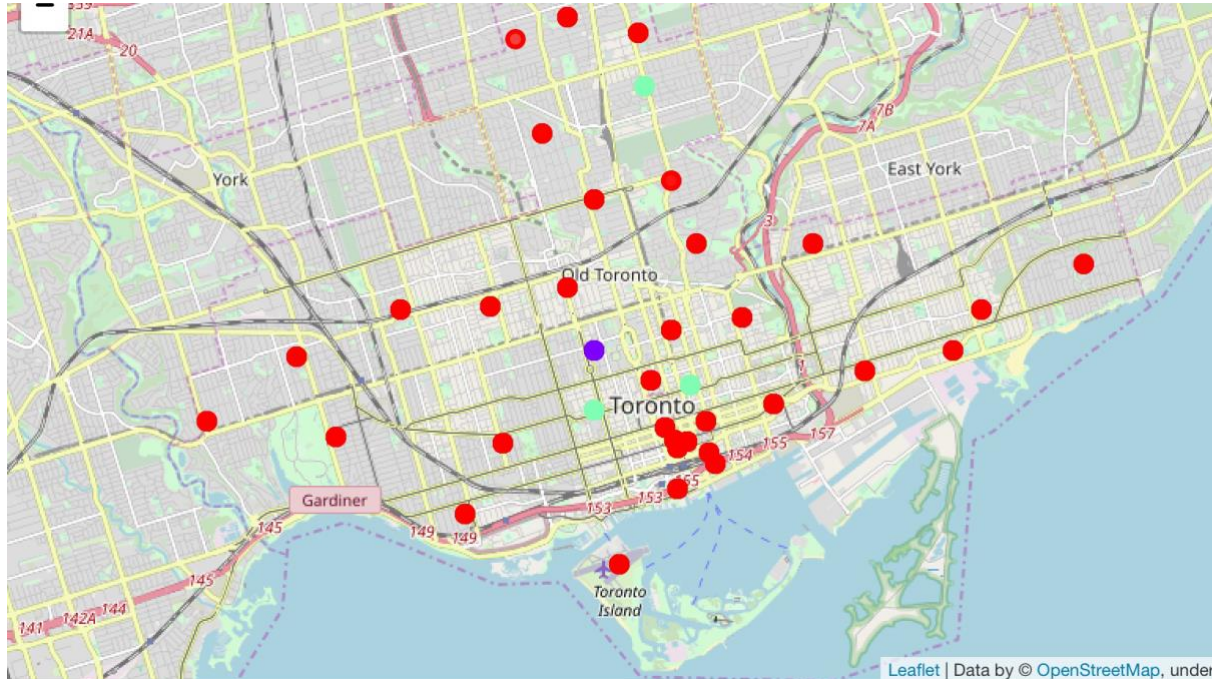


Fig 2. Map clusters of gaming venues by k-means clustering.

The results from k-means clustering show that we can classify Toronto neighbourhoods into 3 clusters based on how many gaming-related venues are in each neighbourhood.

- ❖ Cluster 0: Neighbourhood with little /no gaming venues.
- ❖ Cluster 1: Neighbourhood with no gaming venues.
- ❖ Cluster 2: Neighbourhood with high gaming venues.

The results which are envisioned in the above map with

- Cluster 0: Red
- Cluster 1: Purple
- Cluster 2: Green

Most of the Gaming related shops and cafes are in **Cluster 2** which is around Ryerson, Garden District and Chinatown, Grange Park, Kensington Market. areas in lowest (close to zero) fall in **Cluster 1** areas which are Harbord, University of Toronto nearby areas. Also, there are good opportunities to open near Central Bay street as the competition seems to be low. Looking at nearby venues, it appears **Cluster 1** might be a promising location as there are not a lot of Gaming venues in these areas. Therefore, this project recommends the entrepreneur to open a gaming centre or an arcade in these locations with no competition.



Fig 3. Highlighting the **cluster 1** with the label.

Recommendations

As there are numerous gaming stores there will be a lot more advantages and business outcomes with collaborating with other game stores and cafes nearby, you can make your business a booming one.

Conclusion

In this project report, we have gone through the process of identifying the business problem, specifying the data required, extracted and preparing the data, performing the machine learning by k- means clustering and provided a recommendation to the stakeholder.