The Battle of Neighborhoods

**Introduction/Business Problem**

In this 21st Century there is a surge in Indians travelling abroad for education or employment. Studies reveal that students prefer to settle down in the US, the UK, Australia and Canada the most. It comes as no surprise that 1 in every 5 international students is from India. Canada is one the most popular destination for Indian migrants worldwide. Usually people tend to explore the places before moving to a new country, state, city or place for their work or to start a new life that involves so many factors including neighborhood analysis.

Now, Data Science plays a key role in analyzing each neighborhood in a city and gives insights about each and every neighborhood. Analyzing a neighborhood would be helpful for those who are travelling to a new country and get familiarized about their neighborhood even before landing.

The Project takes this requirement as its main idea to help the stakeholders to achieve the desired results, so as to help them spend their time and money in some other productive means rather keep them trapped in an infinite loop of extensive search engine.

My Objective is to get insights about the neighborhoods which has Indian Community living in Toronto, Ontario, Canada. Along with the other amenities like Indian Restaurants, Indian shops, Indian shopping mall, Indian Grocery Market, etc.. This helps in predicting the densely populated Indians living in those neighborhood. So, Naturally the Indian Migrants chose to live in the densely populated community as they feel safe and secure like their hometown.

This Project would help the stakeholders take a better decision on choosing the neighborhood of Indian Community out of many neighborhoods to live in Toronto city based on the distribution of various facilities in and around the neighborhood.

**Analytic Approach:**

**K-means clustering** unsupervised machine learning model is applied to identify the Indian Community in the Neighborhoods of Toronto, Canada with the help of **Foursquare location data and intelligence.**

*My Approach to Identify the Indian Community in Toronto:*

*In a city like Toronto, Canada, the details of all the neighborhoods and their Postal code are collected and they are matched with geographical coordinates of the center of each neighborhood. Then using a combination of location data and machine learning, the neighborhoods are explored and segmented into clusters to get deeper insights of each neighborhoods reputation*.

**Data Section:**

***Data Requirements and Data Collection***:

For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto.

[https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of­\_Canada:\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

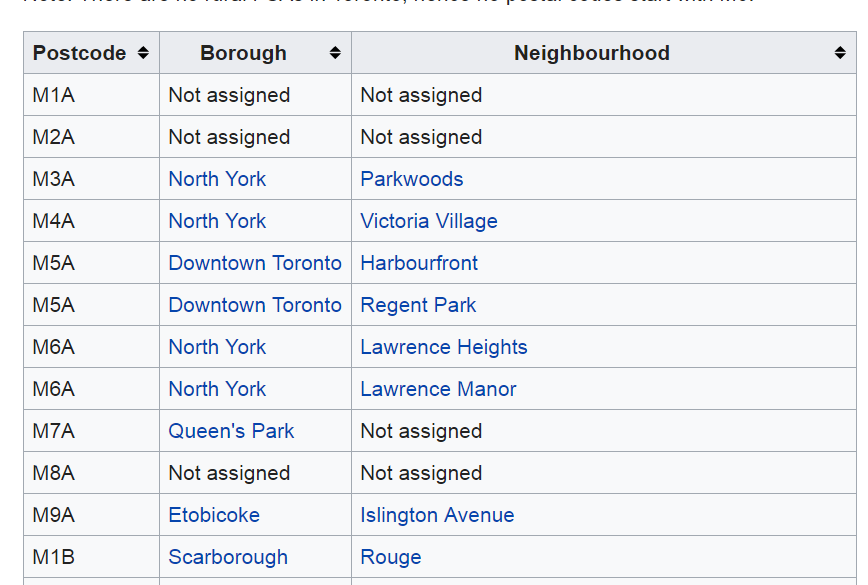
In order to utilize the Foursquare location data, we need to get the latitude and longitude coordinates of each neighborhood which is available as csv file: <https://cocl.us/Geospatial_data>

***Data Understanding***

This section allows us to check whether the data collected are relevant to solve the problem.

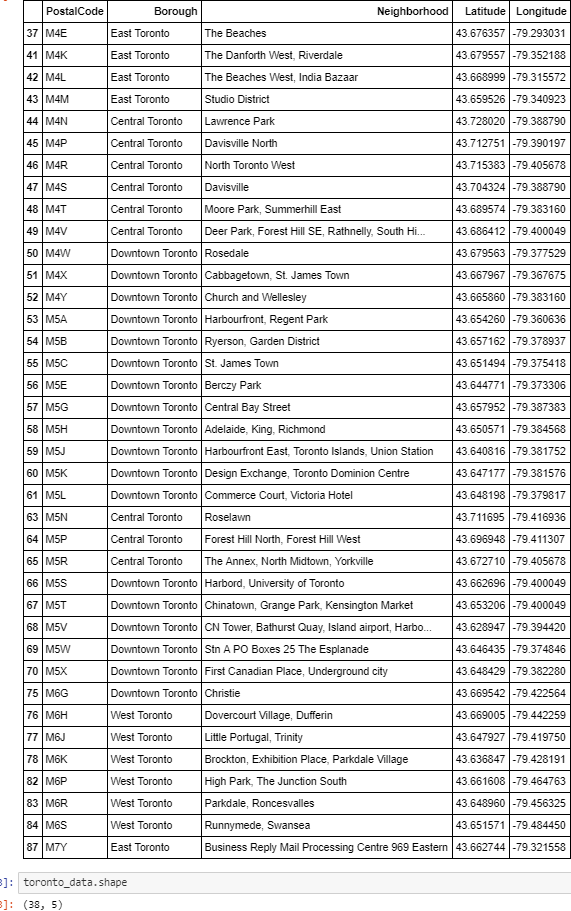
***Data Preparation:***

The data obtained from the Wikipedia page has to be transformed into a strucuted format with valid entries.



The Following changes are made to construct a Structured Dataset to work with:

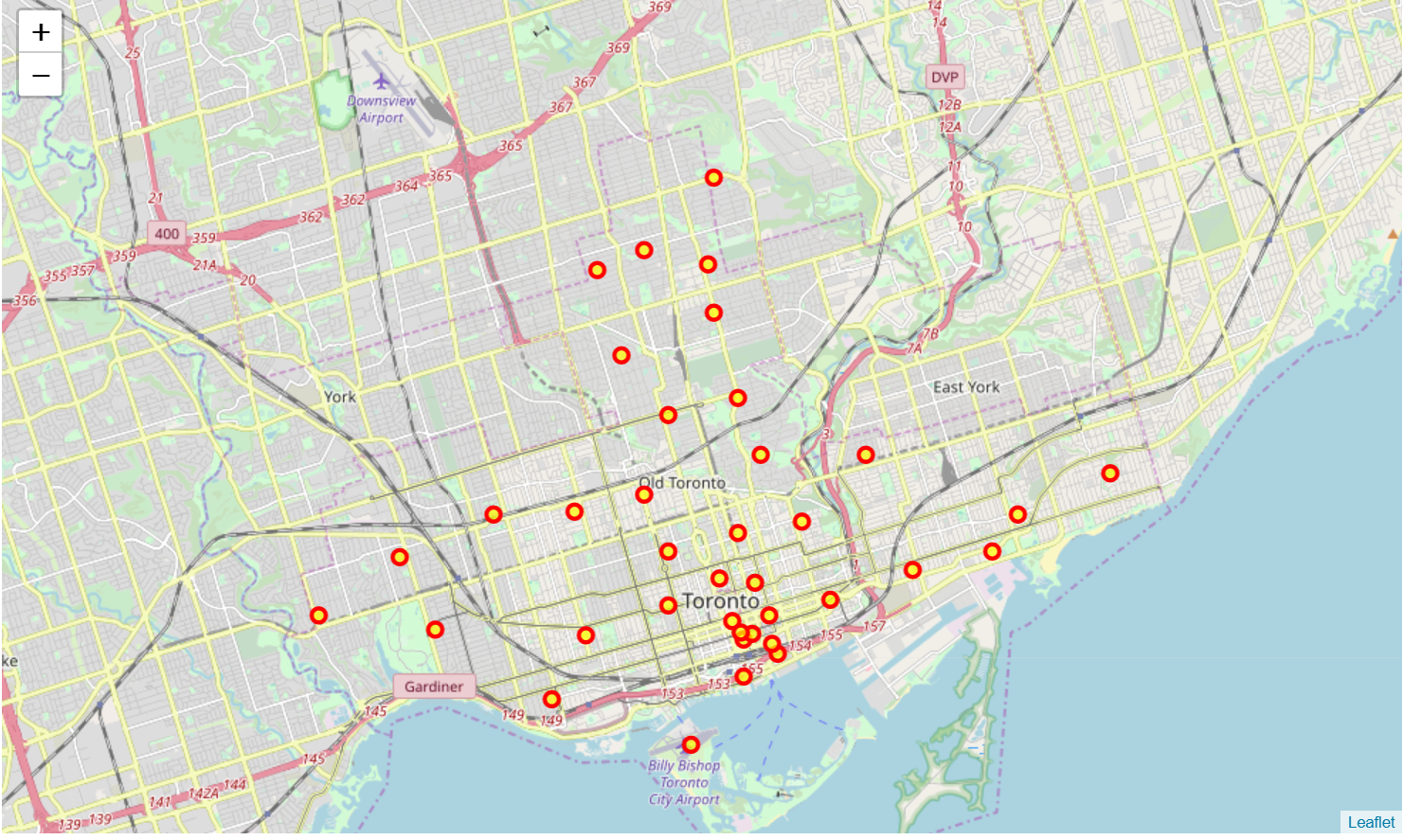
* Only process the cells that have an assigned *Borough*. Ignore cells with a *Borough* that is **Not assigned.**
* If a cell has a *Borough* but a **Not assigned** *Neighborhood,* then the *Neighborhood will be the same as the* Borough.
* More than one Neighborhood can exist in one Postal Code area. For eg. In the above table of Wikipedia page, you will notice that M5A is listed twice and has two *Neighborhoods: Harbourfront* and *Regent Park*. These two rows will be combined into one row with Neighborhoods separated with a comma.
* Then the Neighborhood of Canada dataset has to be merged with their latitudes and longitudes.
* Then get the dataset that has Toronto and their Neighborhoods and their coordinates to start the analysis.



Methodology Section :

**Explore Neighborhoods in Toronto :**

The Structured format – toronto\_data is now used for exploratory data analyses. Lets Visualize the map of Toronto with their neighborhoods.



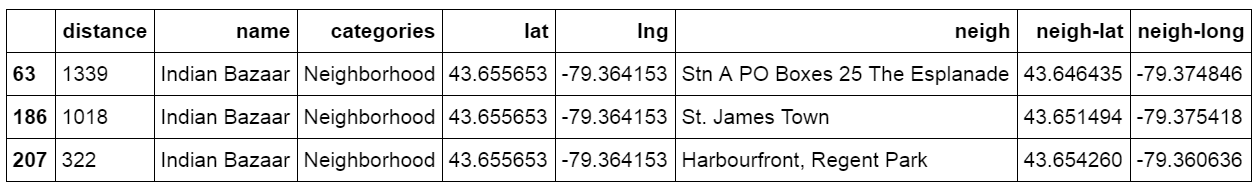
Foursquare location data and Intelligence plays a vital role in searching all the Indian Neighborhoods, Indian Restaurants, Indian shops, Indian markets etc., in all the neighborhoods of Toronto. The following Search query for “Indian” in 1000km radius was applied with the help of Foursquare API calls to all the neighborhoods.

The resulting dataset revealed the Names, Categories and their corresponding coordinates of Indian neighborhoods, Indian Restaurants, Park, Indian tattoo Parlor, Indian Astrologer, Indian Consulate/Embassy, Indian Affairs etc., in all the given Neighborhoods of Toronto.





It is observed that only East Toronto and Downtown Toronto has Indian Neighborhoods.

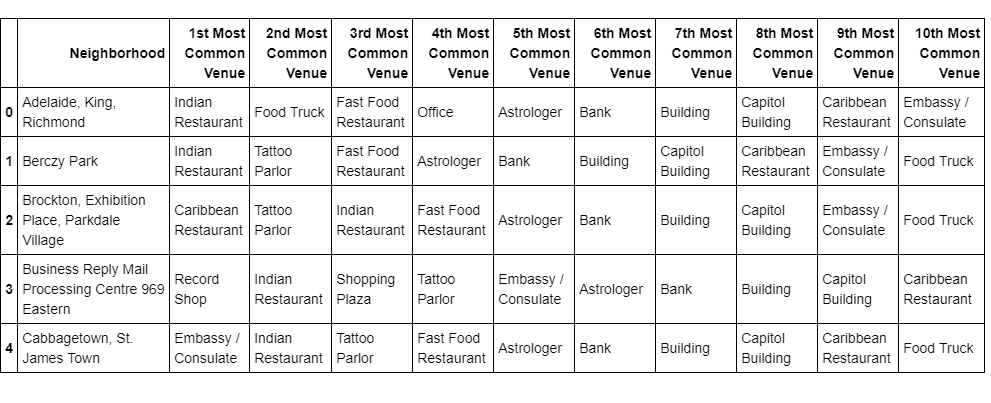




**Analyze Each Neighborhood:**

Analysis is done on the given exploratory data and thus revealed the Top 10 Venues for each Neighborhood in Toronto.

The below table shows top 10 venues ….

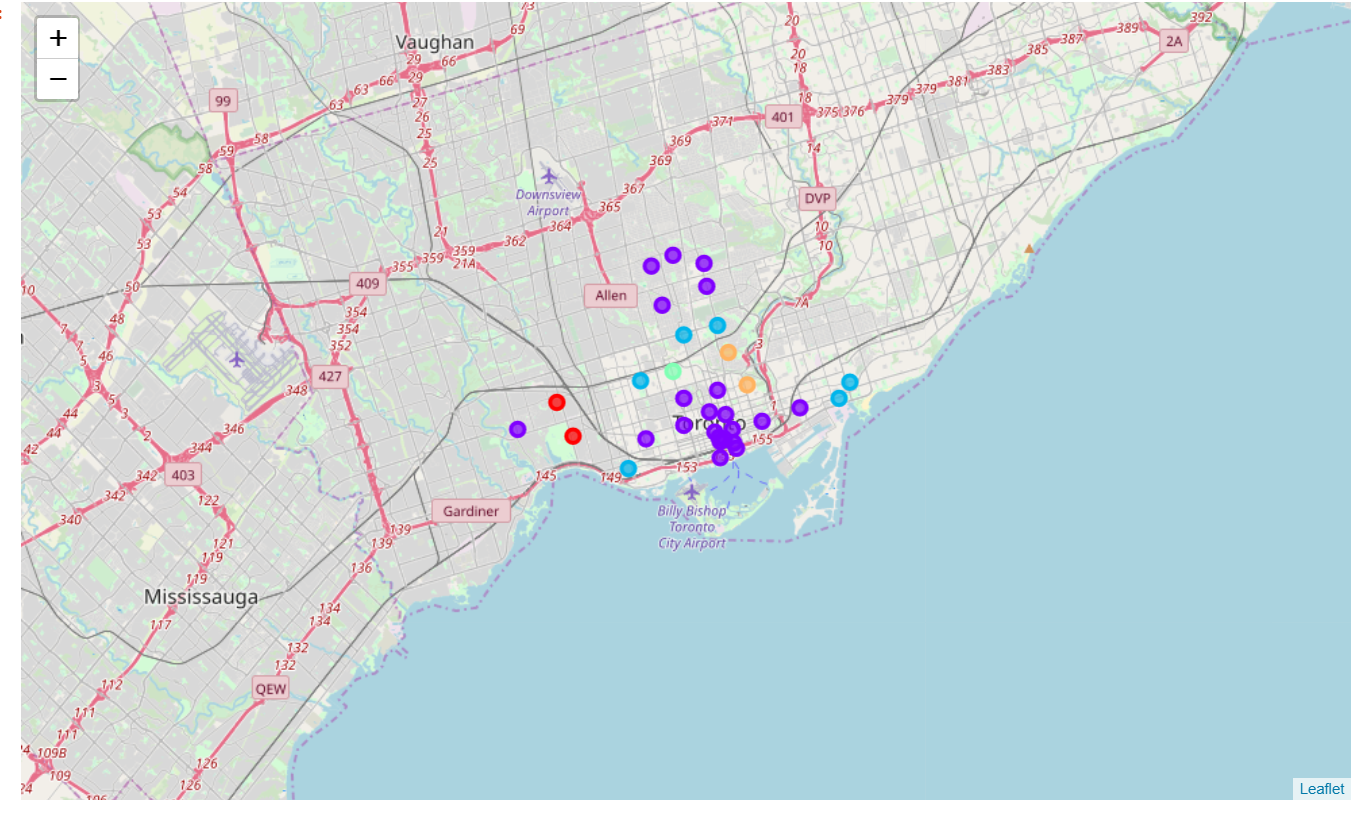


**Cluster Neighborhoods:**

A clustering task consists of creating groups of objects that have high intraclass similarity and low interclass similarity. In other words, objects are similar to other objects in the same cluster, but dissimilar to objects in other clusters,

The Unsupervised Machine Learning Model for clustering used for this project is Kmeans Clustering,

Using K-means the neighborhoods are clustered into 5 Clusters. Lets visualize the clusters.



Cluster\_0 - Red Cluster\_2 - cyan

Cluster\_1 - purple Cluster\_3 – light green Cluster\_4 – Orange

**Results Section:**

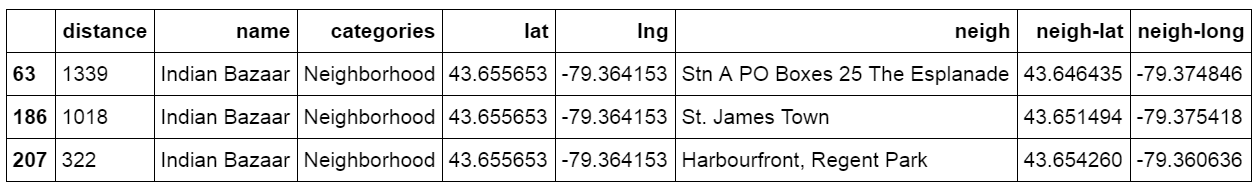
1. There are 2 neighborhoods in Cluster\_0. Park, Tattoo parlor, Astrologer, Consulate etc… are the most common venue.



#### There are 22 neighborhoods in Cluster\_1

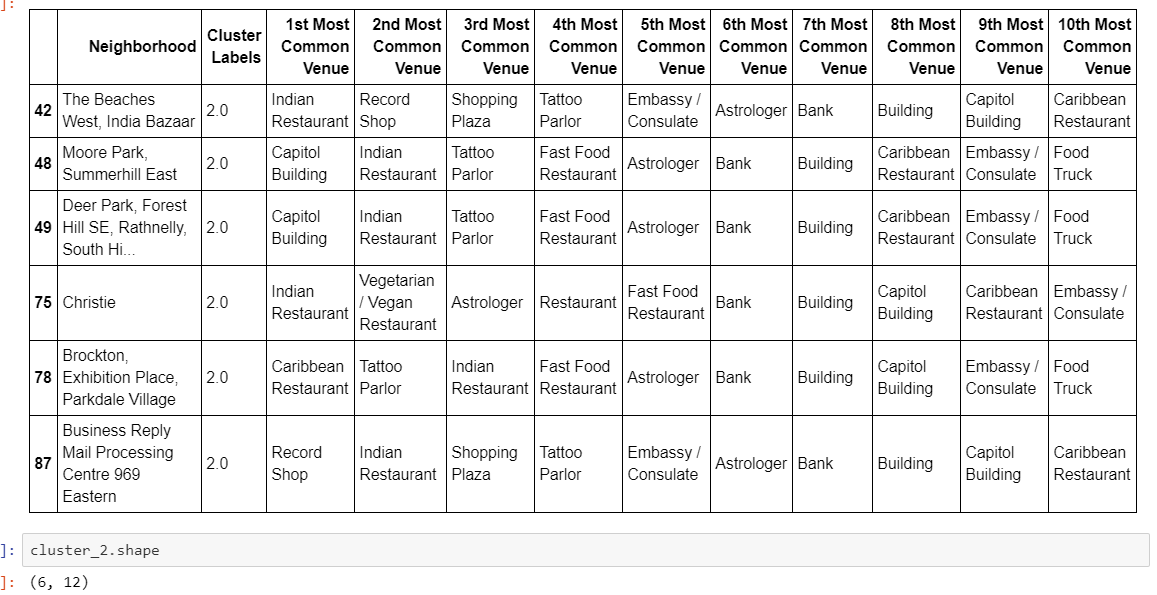
#### 

It is observed that the Indian Neighborhoods : **Indian Bazaar** in Esplanade, St. James Town, Harbourfront, Regent Park



are found in Cluster\_1.

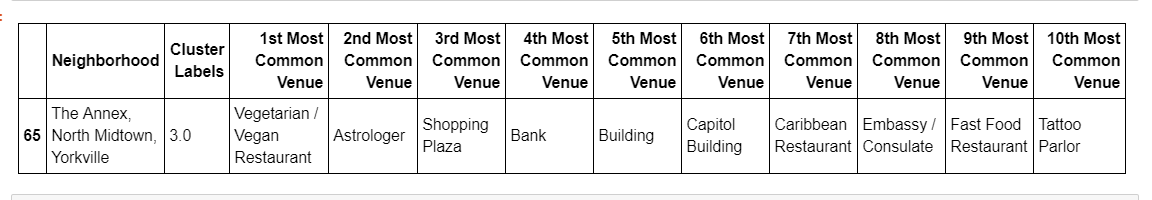
1. There are 6 Clusters in Cluster\_2. Bank, Astrologer, Shopping Plaza, Capitol Building, Indian Restaurant



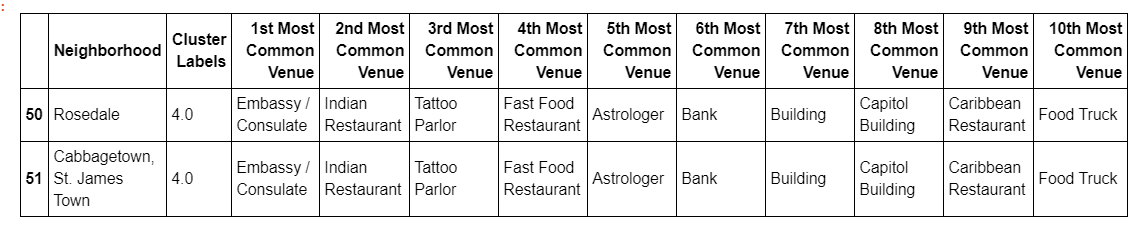
Etc., are the most common venues. Also observed that the Indian Neighborhood - **Little India Neighborhood** in the Beaches West, India Bazaar, Business Reply Mail Processing Centre 969 Eastern are in Cluster\_2.



1. There is 1 cluster in Cluster\_3 with the most common venues tabulated below.



1. There are 2 clusters in Cluster\_4 with the most common venues tabulated below.



**Discussion Section**

**Which is the best locality to settle down?**

* Evaluate Criteria for Selecting the community to re-locate to
* Compare localities based on priorities and preferences
* Review results to take “informed decisions”

**Is this the right location to invest for a new restaurant or community center?**

* Retail & Business- Understand potential growth opportunities
* Review trend of community expansion & settlement diversity
* Prioritize locations to set up new commercial establishments

**Which Neighborhood’s are popular by ethinicity, demand for new development initiatives?**

* Government & Municipal authorities – Town & Community planning
* Predict growth and improve services
* Proactively plan for newer zones development

**Conclusion :**

## **St.James Town, Harbourfront, Regent Park, The Esplanade are** the **top rated** neighborhood to Live In Toronto where most Indians live. With help of Kmeans analysis, It is observed that they are grouped in Cluster\_1 where other amenities like Indian Restaurants, Fast food Restaurant, Food Truck, Tattoo Parlor, and Corporate offices like Bank, Embassy, Capitol Buildings, Astrologer are well established for a comfortable stay in Toronto.

### Less densly populated Indian communities live in the neighborhoods are **Business Reply Mail Processing Centre 969 Eastern, The Beaches West, Indian Bazaar.** They are grouped under Cluster\_2 where other amenities like Indian Restaurant, Record Shop, Shopping Plaza, Tattoo Parlor, Embassy / Consulate Office, Astrologer, Bank and Capitol Buildings are available.

**Thank You!**