1. Introduction

1.1 Background

Business recommendation System is very important for everyone that wants to start any business in any city or area. For this project we have selected PWD Islamabad as a central location but for the business review whole city data is important for us

In this project we will try to find the business opportunities in different area is **Islamabad Pakistan** .
Since there are many businesses already running in different areas, we will try to get the list of businesses that anyone can start in any new area.
For example There is any area where there is no Gym , so we can recommend about
Gym. We are getting Information about recommended businesses from other similar areas because new businesses should be similar to surrounding businesses. We can not start all businesses in all areas.

1.2 Problem

One of the problems will be data, it will be a challenge to get data about different businesses around any area.

1.3 Interest

This project is very useful for anyone that wants to start and wanna see business opportunities in any area. This project also gives good information about different businesses in any specific area that can help anyone that wants to move in that area. User ratings about different venues can help to plan visits to that venue.

2. Data acquisition and cleaning

According to pour problem, we have decided to get following data:
* All the current venues around the central location. It will includes all
* Most popular businesses around the central location.
* Rating and users interest/Like in all venues

2.1 Data sources

For this project we will get data from **FourSquare** API. This API will provide us all the surrounding information about any specific area , there are different API endpoints that we can

use to get data. For this project we will mainly use 3 API calls which are Search Venues, Recommendation Venus and Information about Venue.

2.2 Data cleaning

After getting data from API calls, first we need to convert that data into tabular form from json format. Data cleaning will be challenging and depends upon Client's requirement. We need to delete or manually add the category of Venus. We may need to remove some Venues information that are not useful for us like 'Police Station' or any other government building information.

Information about different residential areas will be important for us to know about the population of that area etc.

2.3 Feature selection

From all the data about different venues around any specific area, we need to select some important information about those venues. Distance from location, Rating from other users,

3. Exploratory Data Analysis

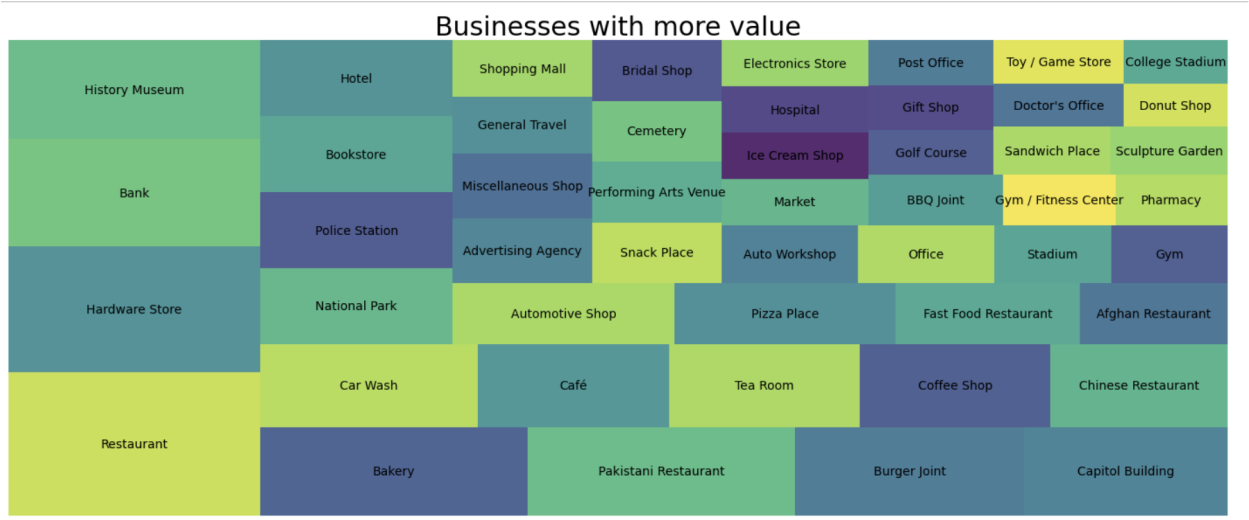
3.1 Calculation of target variable

This project will recommend some **Location** and Possible **Businesses** to start in that location.

4. Predictive Modeling :

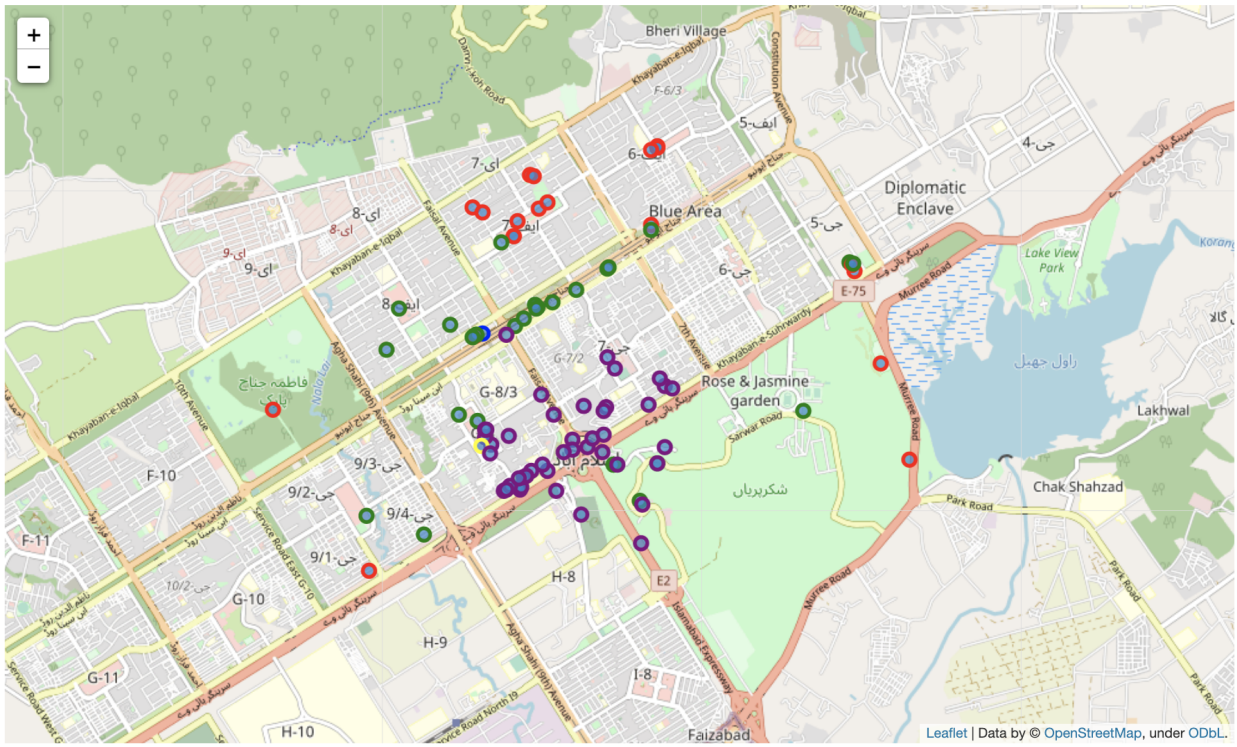
For this application, we have tried two predictive modeling.

Content-Based recommendation system:



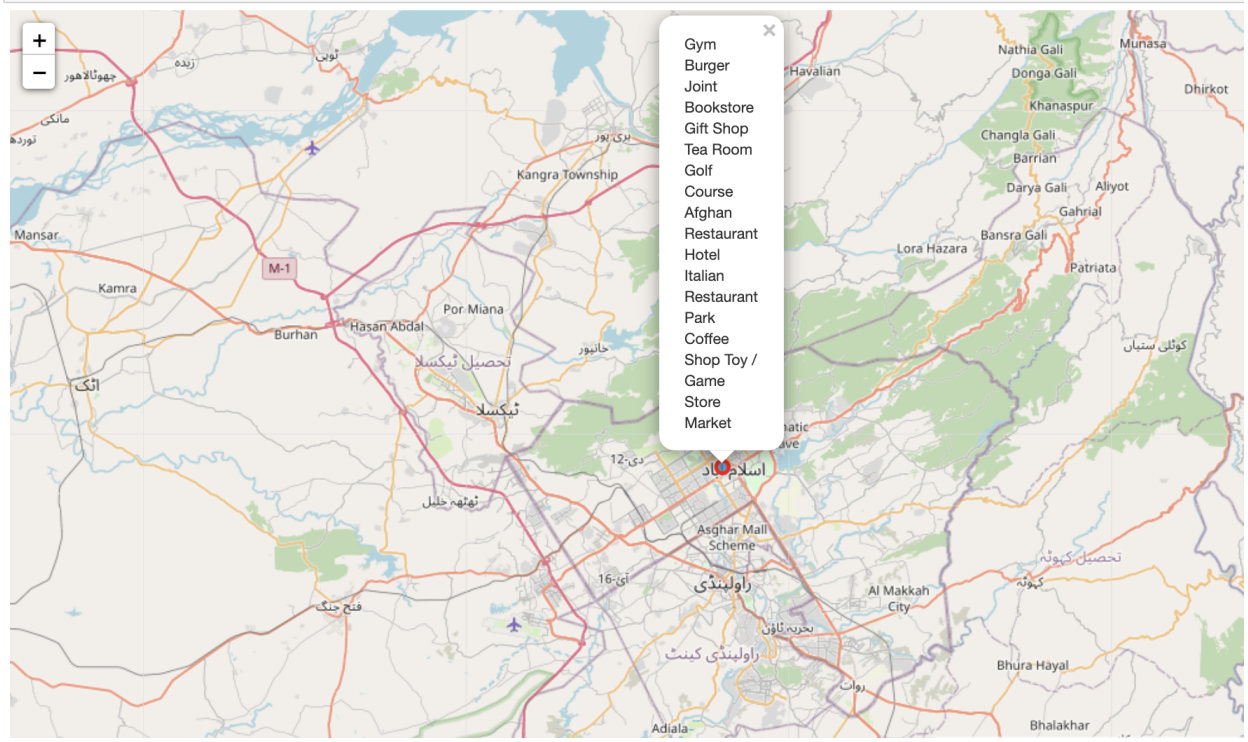
K-Mean Clustering:

label	venue_id	venue_name	venue_distance_from_location	venue_latitude	venue_longitude	category	venue_rate	venue_likes
0	1	1	1	1	1	1	1	1
1	21	21	21	21	21	21	20	20
2	30	30	30	30	30	30	30	30
3	2	2	2	2	2	2	0	0
4	46	46	46	46	46	46	0	0



5. Results are verification:

Our System recommended these businesses in G8 area:



We can see that these business have lack in this area, while these businesses are common in other areas.

```
getVenueSearch("G8",venue_latitude, venue_longitude)
```

	venue_id	venue_name	venue_distance_from_location	venue_latitude	venue_longitude	category
0	502a6e93e4b0c903102169f9	Rooh Afza Market	334	33.703993	73.057519	Rest Area
1	5dd3da22e100e00007888276	Medi Test Labs	468	33.709100	73.056740	Research Laboratory
2	579c628f498e7d496fd9e410	BeeWell Hospital	499	33.710237	73.058361	Hospital
3	4dbec7f04b222080d3c7ed14	Shaheen Chemist	537	33.708667	73.055329	Pharmacy
4	4ef1b2c48b81368cf76fd515	New Links International	473	33.709917	73.058142	Mobile Phone Shop
5	4e7759aeb0fb96803404e6a2	Daewoo Bus Terminal	583	33.704950	73.054014	Bus Station
6	50b99eb6e4b071a4ba221855	National Defence University	484	33.706483	73.054983	University
7	4c764c4ac219224bbec0a428	ASIA CARGO SERVICES	130	33.706442	73.058874	General Travel
8	58084b9938fa7ce987a128bf	NADRA Mega Service Centre	487	33.709023	73.056362	Government Building
9	4c60d09813791b8df9a150af	Hotel Crown Plaza- Oak Room	374	33.708026	73.056937	Not Availble

6. Conclusions

In this project I learned how we can review any city or are to start any business. Few things are important for us to start any new business.

People's interest or Trend (This is very important, because whatever we are going to start people should have interest in that. Like if we are going to start any Pizza shop , people should know about Pizza. In order to start any new Food, first we have to spend some time and investment to guide about new product in people)

Location: Location is also very important, Suppose if we are going to start any Food business, any existing food street will be a good option. If we are going to start any office, then any location where there are already many offices, can help us to get employees.

Similar Businesses: In order to start any business, a similar business with the same product is very important. For example if we need to start Pizza Shop and our shop/cafe is in a location where we already have a very popular Pizza Shop, we should consider that.

Price: Price is one of the main items to run any business. In order to start any business our price should be able to compete with similar businesses and we should know How many people in that area will be able to buy that product.

By using this Application, we can Plot these kind of information for someone that wanna start new business

7. Future directions

More data can improve the implementation of this application. I have selected some data related to Islamabad Pakistan. There is lots of information available on FourSquare API , but in reality some data is still missing. So if we can get more data we can improve our application more.