

# **"A visit to Shimla City (Hill Queen) of India"**

Arun Kumar

**June 01, 2020**

## **1. Introduction**

To begin with, we will take a look at Shimla on the Map using the folium library. Shimla is the capital of the northern Indian state of Himachal Pradesh, in the Himalayan foothills. Once the summer capital of British India, it remains the terminus of the narrow-gauge Kalka-Shimla Railway, completed in 1903. It's also known for the handicraft shops that line The Mall, a pedestrian avenue, as well as the Lakkar Bazaar, a market specializing in wooden toys and crafts.

- ❖ Elevation: 2,276 m
- ❖ Area: 35.34 km<sup>2</sup>
- ❖ Weather: 11 °C, Wind NE at 8 km/h, 97% Humidity
- ❖ Population: 1.7 lakhs (2011)
- ❖ Local time: Sunday, 6:46 pm
- ❖ Metro population: 1,71,640

## **Problem Background:**

- The City of Shimla is the capital of State of Himachal Pradesh, in India. During British Raj, it was the summer capital of the Britishers. The City of Shimla is called the Summer Queen of India as lacs of the tourists visit this place during summers and camp around of the city.
- The temperature of India is hot and humid and therefore hill stations have a major attraction to the tourists. Therefore, Shimla is the major hill station that holds tourist's attention during summers as well as during winters. I have spent around two decades in Shimla and know the problem faced by them during their visit to the city.
- This work shall help the people in knowing about the places and eating joints in and around the Shimla city and can spend money as per their pocket.
- The information provided through the maps and charts shall prove its worth to the tourists who intend to visit the city.

- Every year more than **one million tourists** visit the Shimla City. The tourists find it difficult to search the best venues with budgetary constraints and they also want to visit the venues that are near to the tourist places of Shimla, as they have limited time. Therefore, I decided to help the tourists of this city by helping them by answering the following questions:
  - Get information for which of the eating/drinking **venues available** within the 4km periphery from main Shimla?
  - Get information for which of the eating/drinking venues available within the 4km periphery from main Shimla are **having low prices**?
  - Get information for which of the eating/drinking venues available within the 4km periphery from main Shimla **have good ratings**?
  - Get information for which of the eating/drinking venues available within the 4km periphery main Shimla are **cheaper and have good rating?**

## **2. Data Description**

### ***Data Description***

- The Map of Shimla city can be created using the **folium library**.
- The data about the different eating/drinking venues like name, longitude, latitude, etc. within the radius of 4km from the centre of Shimla shall be fetched from **Foursquare API**.
- The data about the different eating/drinking venues like name, longitude, latitude, address, rating, category, price etc. within the radius of 4km from the centre of Shimla shall be fetched from **Zomato API**.
- **Shimla\_LATITUDE = 31.104605**
- **Shimla\_LONGITUDE = 77.173424**

**Snapshot of the Map of Shimla:**

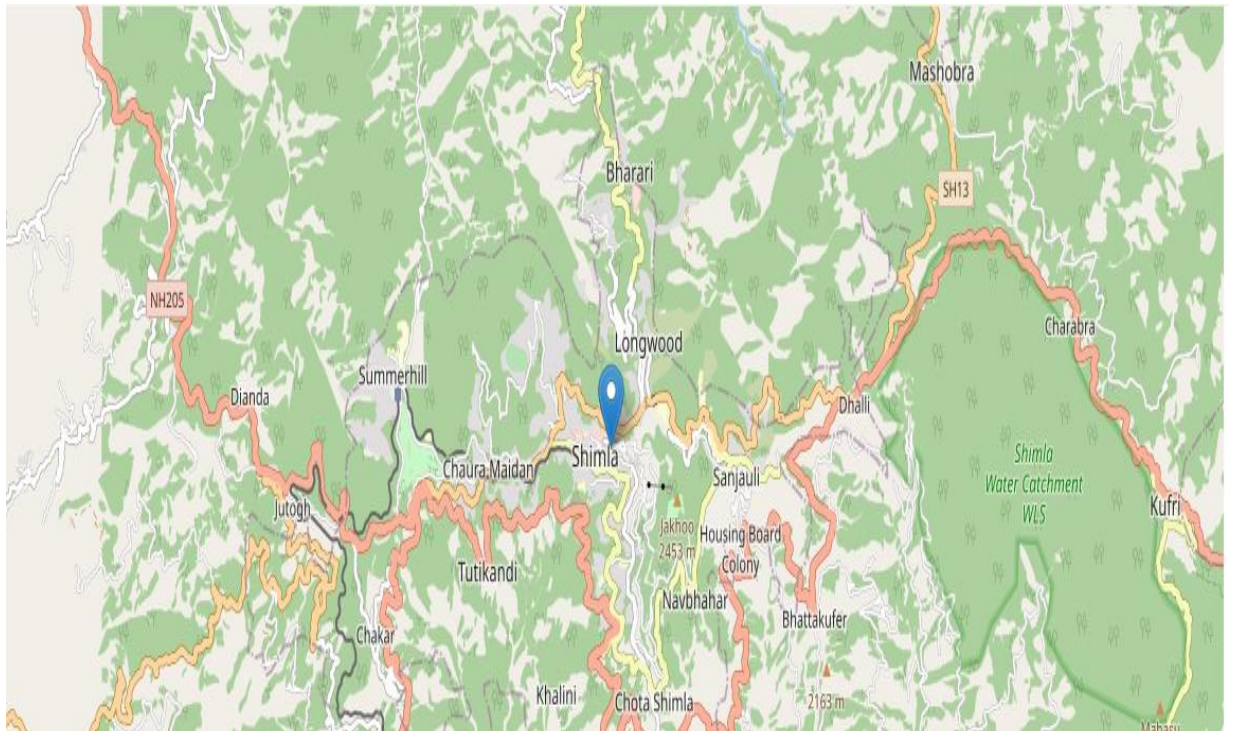


Figure1: **Snapshot Map of Shimla**

**Snapshot Map of Shimla refined data:**



**Figure: Snapshot Map of Shimla-Refined**

**Snapshot of the Data set captured:**

	name	categories	lat	lng	venue	latitude	longitude	price_for_two	price_range	rating	address	lat_diff	lng_diff
0	Shimla	City	31.1053	77.1717	Dobby's	31.1052	77.1716	350	2	0	Mall Road, Near Syndicate Bank, Summer Hill, S...	-0.0001	-0.0001
1	The Mall	Plaza	31.1046	77.1751	Honey Hut	31.1044	77.1750	450	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	-0.0002	-0.0001
2	Lakkar Bazaar	Furniture / Home Store	31.1065	77.1784	Fresh Juice Shop	31.1063	77.1782	150	1	0	Lakkar Bazar, Longwood, Shimla	-0.0002	-0.0002
3	The Ridge, Shimla	Monument / Landmark	31.1048	77.1758	Book Cafe	31.1050	77.1760	200	1	0	Takka Bench, Ridge, The Mall, Summer Hill, Shimla	0.0002	0.0002
4	Wake and Bake Cafe	Café	31.1047	77.1732	Wake & Bake	31.1047	77.1732	800	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	0.0000	0.0000
5	Indian coffee house	Coffee Shop	31.1054	77.1710	Indian Coffee House	31.1054	77.1713	350	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	0.0000	0.0003
6	The Oberoi Cecil	Hotel	31.1032	77.1549	The Restaurant - The Oberoi Cecil	31.1032	77.1550	2200	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	0.0000	0.0001

### Snapshot of the data captured (average price)

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	City	Dobby's	31.1052	77.1716	2	0	Mall Road, Near Syndicate Bank, Summer Hill, S...	175.0
1	Plaza	Honey Hut	31.1044	77.1750	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	225.0
2	Monument / Landmark	Book Cafe	31.1050	77.1760	1	0	Takka Bench, Ridge, The Mall, Summer Hill, Shimla	100.0
3	Café	Wake & Bake	31.1047	77.1732	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	400.0
4	Coffee Shop	Indian Coffee House	31.1054	77.1713	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	175.0
5	Hotel	The Restaurant - The Oberoi Cecil	31.1032	77.1550	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	1100.0
6	Indian Restaurant	Baljees Sweet Shop	31.1048	77.1729	1	4.1	Near Scandal Point, Summer Hill, Shimla	150.0
7	Indian Restaurant	Sher-e-Punjab	31.1033	77.1770	2	2.8	Mall Road, Summer Hill, Shimla	300.0
8	Pizza Place	Domino's Pizza	31.1049	77.1728	2	3.4	24, Mall Road, Summer Hill, Shimla	200.0
9	Coffee Shop	Cafe Coffee Day	31.1044	77.1747	2	2.8	Middle Bazaar, Mall Road, Summer Hill, Shimla	250.0

## 3. Methodology

In order to solve the problem at hand, the following methodology has been adopted for which the steps are as follows:

### 1: Data Collection

In our project work, we are dealing with the city of Shimla, therefore data about the category, name, longitude, latitude, rating, price etc. is collected from Foursquare API and Zomato API. Before fetching the data from these APIs, some credentials are required.

(a) The credentials required to fetch the data from Foursquare API are <CLIENT ID>, <CLIENT SECERT>

(b) The credential required to fetch the data from Zomato API is <ZOMATO API KEY>

These credentials are generated after creating the account on the following website: [www.foursquare.com](http://www.foursquare.com) and [www.Developers.zomato.com](http://www.Developers.zomato.com).

## 2: Data Cleaning

Data cleansing or data cleaning[1] is the process of detecting and correcting corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data.

As the dataset for this particular problem has been taken from two different APIs. First, the two datasets are to be merged and then duplicate values and mismatched values are to be removed to get the final dataset on which analysis is to be done.

## 3: Data Analysis

Data analysis is defined as a process of cleaning, transforming, and modelling data to discover useful information[1] for business decision-making. The purpose of Data Analysis is to extract useful information from data and taking the decision based upon the data analysis.

Different types of analysis can be done to answer the queries generated in the introduction section.

## 4: Results & Discussion

The focus of this step is on what **results** have been obtained and set out clearly what happened in the experiments and/or investigations without worrying about their implications.

In this step, different results can be shown with the help of different visualization tools like bar plots, maps etc. Through these different visualization tools, different results shall be discussed.

## 5: Conclusion

A conclusion summarizes the report as a whole, drawing inferences from the entire process about what has been found, or decided, and the impact of those findings or decisions.

## 4. Results

In this section, we shall discuss the results generated from the data collection, data cleaning and data analysis phases which are as follows:

### 4.1 Data Collection Results

The data obtained from the two APIs is merged and data about a total of 54 venues has been obtained. The sample of the merged dataset is shown in Fig. 4 below:

	name	categories	lat	lng	venue	latitude	longitude	price_for_two	price_range	rating	address	lat_diff	lng_diff
0	Shimla	City	31.1053	77.1717	Dobby's	31.1052	77.1716	350	2	0	Mall Road, Near Syndicate Bank, Summer Hill, S...	-0.0001	-0.0001
1	The Mall	Plaza	31.1046	77.1751	Honey Hut	31.1044	77.1750	450	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	-0.0002	-0.0001
2	Lakkar Bazaar	Furniture / Home Store	31.1065	77.1784	Fresh Juice Shop	31.1063	77.1782	150	1	0	Lakkar Bazar, Longwood, Shimla	-0.0002	-0.0002
3	The Ridge, Shimla	Monument / Landmark	31.1048	77.1758	Book Cafe	31.1050	77.1760	200	1	0	Takka Bench, Ridge, The Mall, Summer Hill, Shimla	0.0002	0.0002
4	Wake and Bake Cafe	Café	31.1047	77.1732	Wake & Bake	31.1047	77.1732	800	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	0.0000	0.0000
5	Indian coffee house	Coffee Shop	31.1054	77.1710	Indian Coffee House	31.1054	77.1713	350	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	0.0000	0.0003
6	The Oberoi Cecil	Hotel	31.1032	77.1549	The Restaurant - The Oberoi Cecil	31.1032	77.1550	2200	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	0.0000	0.0001

Figure.4: **Sample of merged Data set of venues fetched using APIs**

### 4.2 Data Cleaning Results

In this, the duplicate values and mismatched values from the data that is fetched through the APIs is dropped to make it ready for the analysis step. The Fig. 5 below shows the mismatched value where category is **"Furniture/Home Store"** so these types of records has been dropped.

2	Lakkar Bazaar	Furniture / Home Store	31.1065	77.1784	Fresh Juice Shop	31.1063	77.1782	150	1	0	Lakkar Bazar, Longwood, Shimla	-0.0002	-0.0002
---	---------------	------------------------	---------	---------	------------------	---------	---------	-----	---	---	--------------------------------	---------	---------

Fig.5: **Sample of mismatched Data set of venues fetched using APIs**

The average price per person has been added in the dataset for each eating/drinking venues. The Fig. 6 below shows the few data records.

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	City	Dobby's	31.1052	77.1716	2	0	Mall Road, Near Syndicate Bank, Summer Hill, S...	175.0
1	Plaza	Honey Hut	31.1044	77.1750	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	225.0
2	Monument / Landmark	Book Cafe	31.1050	77.1760	1	0	Takka Bench, Ridge, The Mall, Summer Hill, Shimla	100.0
3	Café	Wake & Bake	31.1047	77.1732	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	400.0
4	Coffee Shop	Indian Coffee House	31.1054	77.1713	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	175.0
5	Hotel	The Restaurant - The Oberoi Cecil	31.1032	77.1550	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	1100.0
6	Indian Restaurant	Baljees Sweet Shop	31.1048	77.1729	1	4.1	Near Scandal Point, Summer Hill, Shimla	150.0
7	Indian Restaurant	Sher-e-Punjab	31.1033	77.1770	2	2.8	Mall Road, Summer Hill, Shimla	300.0
8	Pizza Place	Domino's Pizza	31.1049	77.1728	2	3.4	24, Mall Road, Summer Hill, Shimla	200.0
9	Coffee Shop	Cafe Coffee Day	31.1044	77.1747	2	2.8	Middle Bazaar, Mall Road, Summer Hill, Shimla	250.0

Fig.6: Sample of Data set generated after data cleaning

After data cleaning, the data is ready for analysis.

### 4.3 Data Analysis Results

#### (a) Analysis on the basis of Category

The analysis on the basis of the category of the venue like hotel, café, restaurant etc. has been done. The bar graph below shows the number of venues in each category. It can be observed from the bar graph that we have maximum number of Hotels and Cafes as compare to other categories of eating/drinking venues.

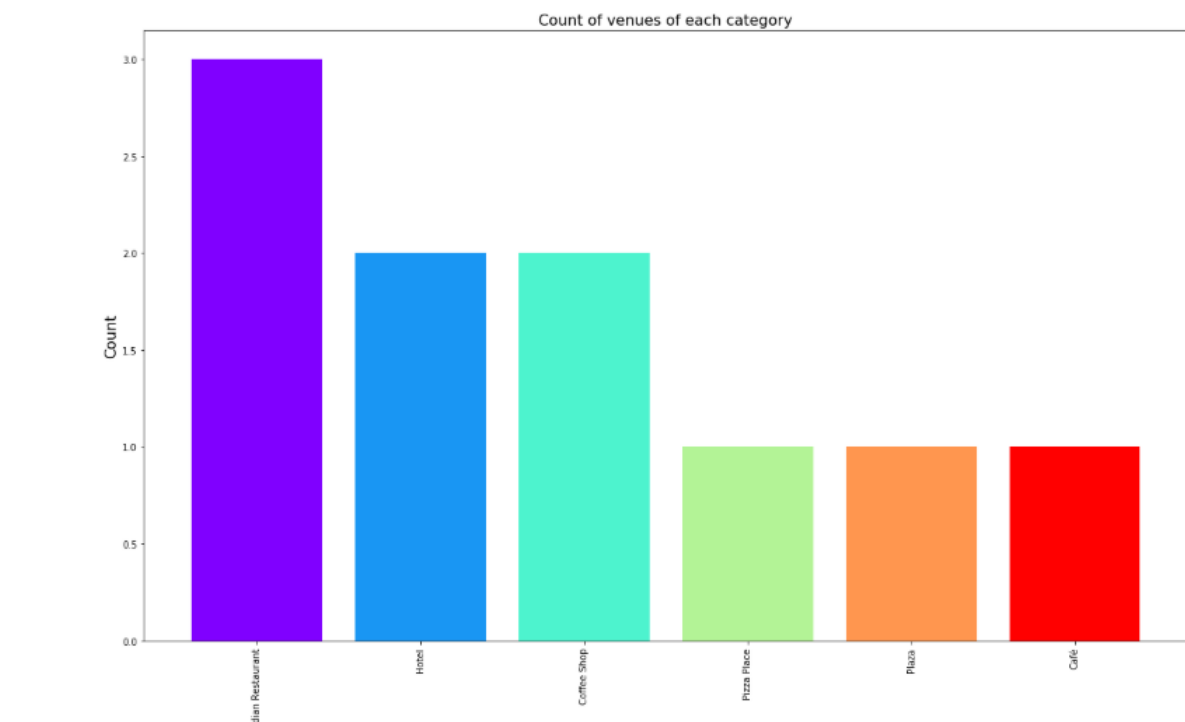


Figure.7: Bar Plot showing the number of venues for each category

**(b) Analysis on the basis of Rating of Venues**

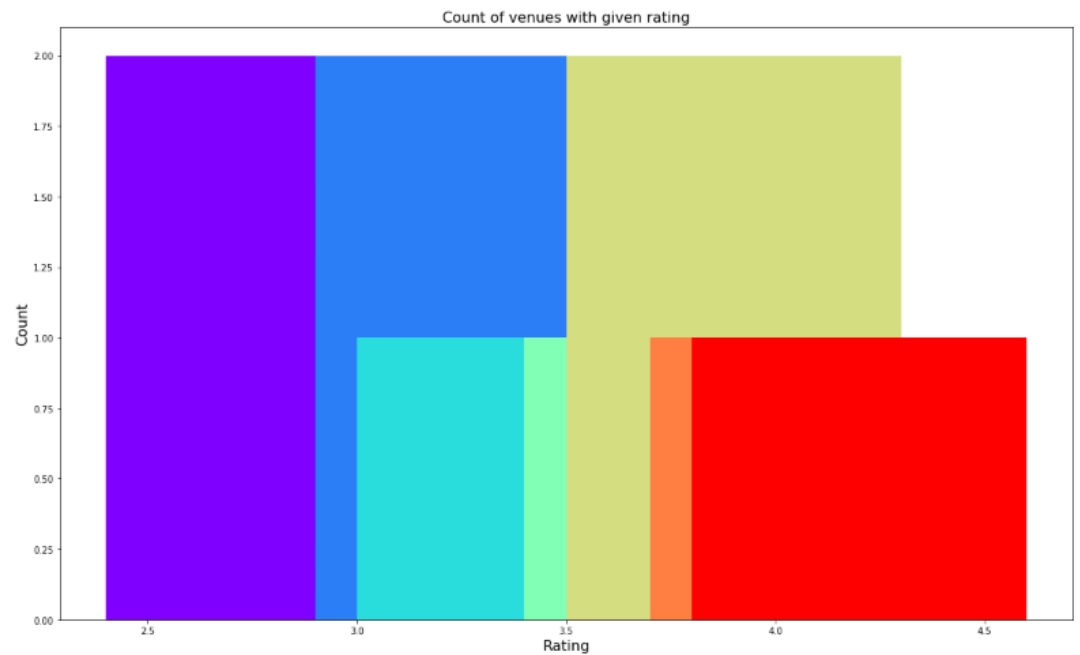


Figure.8: Bar Plot showing the number of venues for each rating

**(c) Analysis on the basis of dividing the venues on the basis of ratings into Low, Okay, Good and very good categories**

The rating of the venues has been divided in to four categories as follows and shown with different colors on the map shown in Fig. 9.

- 1 to 2: Low shown with RED
- 2 to 3: Okay shown with ORANGE
- 3 to 4 Good shown with GREEN
- 4 to 5 Very Good shown with DARKGREEN



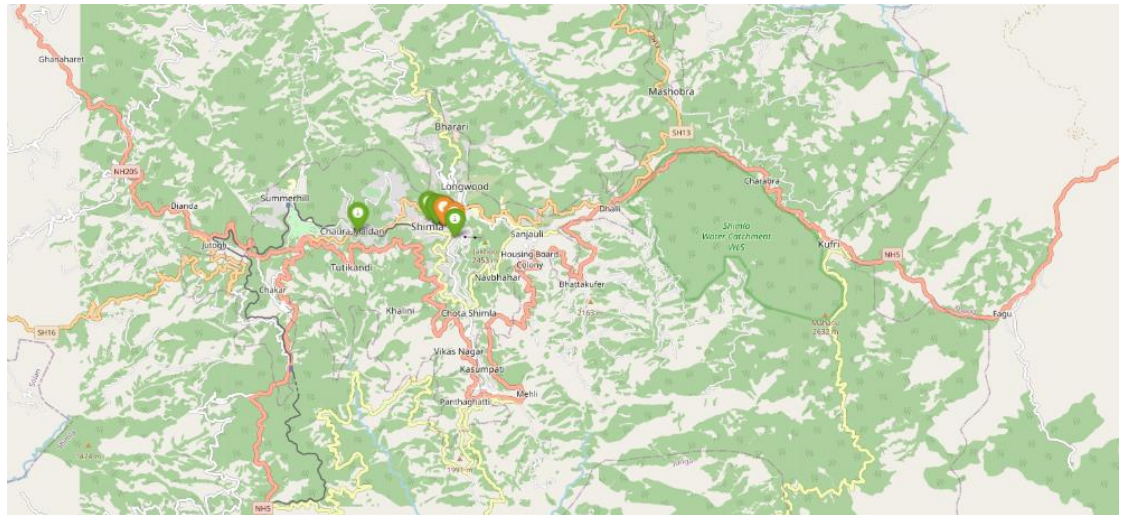


Figure.9: Division of the venues on the basis of ratings into Low, Okay, Good and very good categories

(d) Analysis on the basis of venues with Average Price per Person expenses

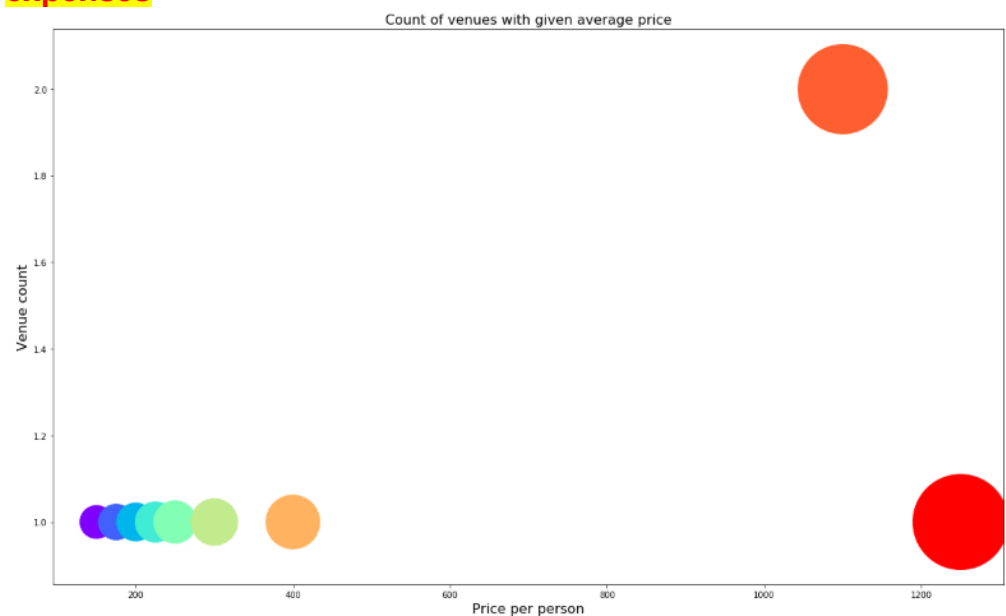


Figure.10: Division of the venues on the basis of Average Price per person expenses

(e) Use of KMeans Clustering

	cluster_labels	categories	venue	latitude	longitude	price_range	rating	address	average_price	rating_bin
1	0	Plaza	Honey Hut	31.1044	77.1750	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	225.0	Very good
3	0	Café	Wake & Bake	31.1047	77.1732	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	400.0	Good
4	0	Coffee Shop	Indian Coffee House	31.1054	77.1713	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	175.0	Good
5	1	Hotel	The Restaurant - The Oberoi Cecil	31.1032	77.1550	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	1100.0	Good
6	0	Indian Restaurant	Baljees Sweet Shop	31.1048	77.1729	1	4.1	Near Scandal Point, Summer Hill, Shimla	150.0	Very good

Figure.11: Sample dataset for dividing the venues into two clusters with label 1 and 0

	cluster_labels	categories	venue	latitude	longitude	price_range	rating	address	average_price	rating_bin
0	0	Plaza	Honey Hut	31.1044	77.1750	2	4.2	Khadi Bhawan, Middle Bazaar, Mall Road, Summer...	225.0	Very good
1	0	Café	Wake & Bake	31.1047	77.1732	3	3.9	34/2, The Mall Shimla, Teh. & Amp Distt, Shiml...	400.0	Good
2	0	Coffee Shop	Indian Coffee House	31.1054	77.1713	2	3.8	1, The, Mall Rd, Lower Bazar, Shimla, Himachal...	175.0	Good
3	0	Indian Restaurant	Baljees Sweet Shop	31.1048	77.1729	1	4.1	Near Scandal Point, Summer Hill, Shimla	150.0	Very good
4	0	Indian Restaurant	Sher-e-Punjab	31.1033	77.1770	2	2.8	Mall Road, Summer Hill, Shimla	300.0	Okay
5	0	Pizza Place	Domino's Pizza	31.1049	77.1728	2	3.4	24, Mall Road, Summer Hill, Shimla	200.0	Good
6	0	Coffee Shop	Cafe Coffee Day	31.1044	77.1747	2	2.8	Middle Bazaar, Mall Road, Summer Hill, Shimla	250.0	Okay

Fig.12: Sample dataset of Cluster 0

cluster_labels		categories	venue	latitude	longitude	price_range	rating	address	average_price	rating_bin
0	1	Hotel	The Restaurant - The Oberoi Cecil	31.1032	77.1550	4	3.9	The Oberoi Cecil, Nabha, State Museum Road, Lo...	1100.0	Good
1	1	Hotel	Seventh Heaven	31.1021	77.1773	4	3.3	Hotel Combermere, Mall Road, Summer Hill, Shimla	1250.0	Good
2	1	Indian Restaurant	HPTDC Goofa Ashiana Bar and Restaurant	31.1047	77.1742	4	3.3	Ridge, The Mall, Longwood	1100.0	Good

Fig.13: Sample dataset of Cluster 1

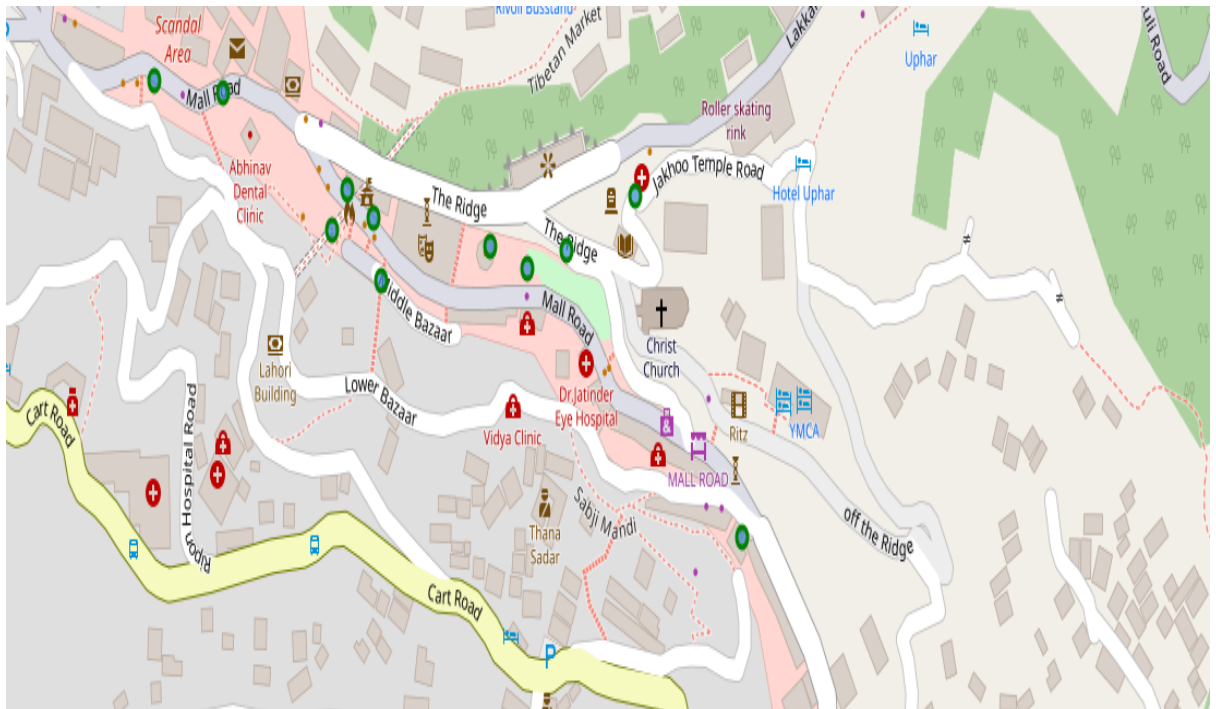


Fig.14: Map showing the two clusters 0 and 1 in green and red colours respectively

## 5. Discussion

For the analysis done in a number of conclusions can be drawn which can be useful to the tourists who want to visit the Hill Queen, Shimla.

### 1. Hotels/Restaurants with Price Range greater than Rupees 1000/-

- For cluster 1: The Restaurant-The Oberoi Cecil has a latitude value of 31.1032 and a longitude value of 77.1550 has rating 3.9
- For Cluster 1: The Restaurant-Seventh Heaven has a latitude value of 31.1021 and a longitude value of 77.1773 has rating 3.3
- For Cluster 1: The HPTDC Goofa Ashiana Bar and Restaurant has a latitude value of 31.1047 and a longitude value of 77.1742 has rating 3.3

### 2. Hotels/Restaurants with Price Range between Rupees 200/- to Rupees 400/-

- For cluster 0: The Restaurant- Honey Hutt has a latitude value of 31.1044 and a longitude value of 77.1550 has rating 4.2
- For Cluster 0: The Restaurant-Wake and Bake has a latitude value of 31.1047 and a longitude value of 77.1732 has rating 3.9
- For Cluster 0: The Balaji Sweet Shop has a latitude value of 31.1048 and a longitude value of 77.1729 has rating 3.3
- After applying KMeans clustering, it has been observed that venues with less average price have a mean rating of 3.88 and average price range of 1.79 whereas the venues with higher price have mean price range of 4.00 and rating spread around 3.60.
- It has been concluded from the analysis that if the tourist wants a cheaper venue for eating/drinking he/she should go to Lower Bazar, Mall Road, for Lakkar Bazar, if he/she is looking for best places

## **6. Conclusion**

The objective of this project is to help the visitors to find the eating/drinking venues who are visiting Shimla. The observations have been shown through maps and it is suggested to the tourists to visit Lower Bazar, Mall Road, for Lakkar Bazar for eating/drinking and can find the venues as per their budget as well as ratings of the venues.

**Future Scope: An Android App can be developed for this.**

## References

- [1] <https://en.wikipedia.org/wiki/>
- [2] <https://foursquare.com>
- [3] <https://developers.zomato.com>