

Zomato_Analysis_API

Consider only Indian restaurants in this analysis -

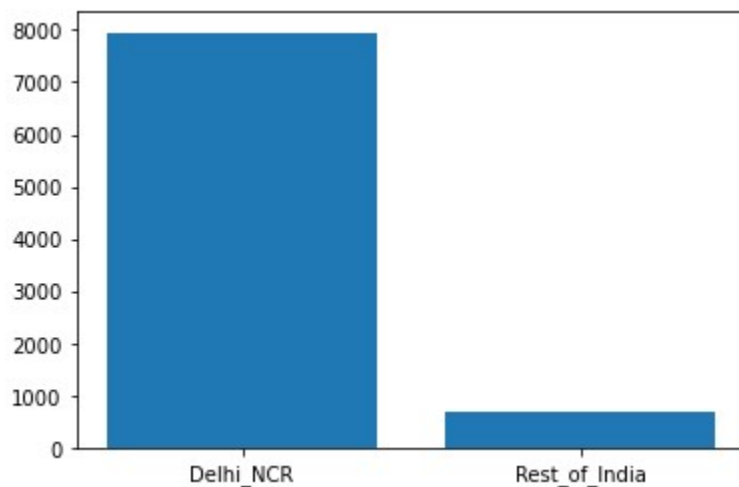
1. The dataset is highly skewed toward the cities included in Delhi-NCR. So, we will summarise all the other cities in Rest of India while those in New Delhi, Ghaziabad, Noida, Gurgaon, Faridabad to Delhi-NCR. Doing this would make our analysis turn toward Delhi-NCR v Rest of India.
 - Plot the bar graph of number of restaurants present in Delhi NCR vs Rest of India.
 - Find the cuisines which are not present in restaurant of Delhi NCR but present in rest of India. Check using Zomato API whether this cuisines are actually not served in restaurants of Delhi-NCR or just it due to incomplete dataset.
 - Find the top 10 cuisines served by maximum number of restaurants in Delhi NCR and rest of India.
 - Write a short detailed analysis of how cuisine served is different from Delhi NCR to Rest of India. Plot the suitable graph to explain your inference.
2. User Rating of a restaurant plays a crucial role in selecting a restaurant or ordering the food from the restaurant.
 - Write a short detail analysis of how the rating is affected by restaurant due following features: Plot a suitable graph to explain your inference.
 1. Number of Votes given Restaurant
 2. Restaurant serving more number of cuisines.
 3. Average Cost of Restaurant
 4. Restaurant serving some specific cuisines.
 - Find the weighted restaurant rating of each locality and find out the top 10 localities with more weighted restaurant rating?
 1. $\text{Weighted Restaurant Rating} = \frac{\sum (\text{number of votes} * \text{rating})}{\sum (\text{number of votes})}$.
3. Visualization
 - Plot the bar graph top 15 restaurants have a maximum number of outlets.
 - Plot the histogram of aggregate rating of restaurant(drop the unrated restaurant).
 - Plot the bar graph top 10 restaurants in the data with the highest number of votes.
 - Plot the pie graph of top 10 cuisines present in restaurants in the USA.
 - Plot the bubble graph of a number of Restaurants present in the city of India and keeping the weighted restaurant rating of the city in a bubble.

Your project will be evaluated on following parameters -

- ☐ Plots and graphs (Max Score 20)
- ☐ Justification (Max Score 20)
- ☐ Answer correctness (Max Score 20)

The dataset is highly skewed toward the cities included in Delhi-NCR. So, we will summarise all the other cities in the Rest of India while those in New Delhi, Ghaziabad, Noida, Gurgaon, Faridabad to Delhi-NCR. Doing this would make our analysis turn toward Delhi-NCR v Rest of India.

1. Plot the bar graph of the number of restaurants present in Delhi NCR vs the Rest of India.



From the output we can see that according to this dataset the total restaurants present in the Delhi-NCR is very large(almost Eight Times) as compare to total restaurants present in the rest of India

Delhi-NCR restaurants-7947

Rest of India restaurants-705

2. Find the cuisines which are not present in the restaurant of Delhi NCR but present in the rest of India.

Check using Zomato API whether these cuisines are not served in restaurants of Delhi-NCR or just due to an incomplete dataset.

Answer: I find four cuisine names with the provided data set where they are not present in Delhi-NCR restaurants and they are German Cajun Malwani BBQ, but I found that some of them are present in Delhi-NCR restaurants after checking with Zomato API. So the dataset has some missing information.

3. Find the top 10 cuisines served by a maximum number of restaurants in Delhi NCR and the rest of India.

Top 10 Cuisines in Delhi NCR

North Indian 3597

Chinese 2448

Fast Food 1866

Mughlai 933

Bakery 697

South Indian 569

Continental 547

Desserts 542

Street Food 538

Italian 535

Top 10 Cuisines in Rest of India

North Indian 349

Chinese 242

Continental 177

Italian 147

Cafe 136

Fast Food 97

South Indian 62

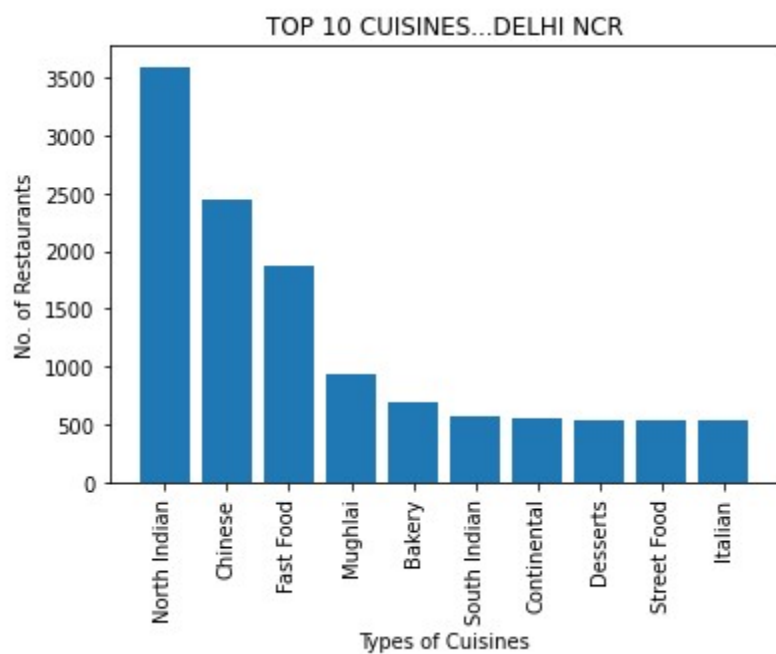
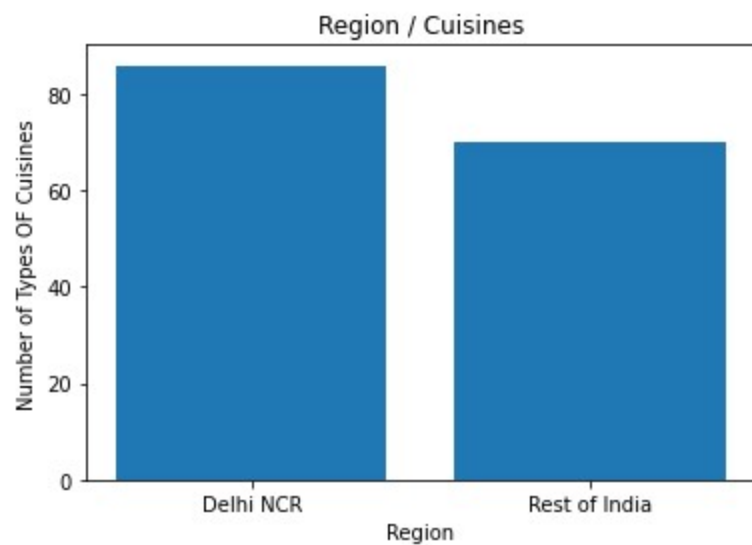
Mughlai 59

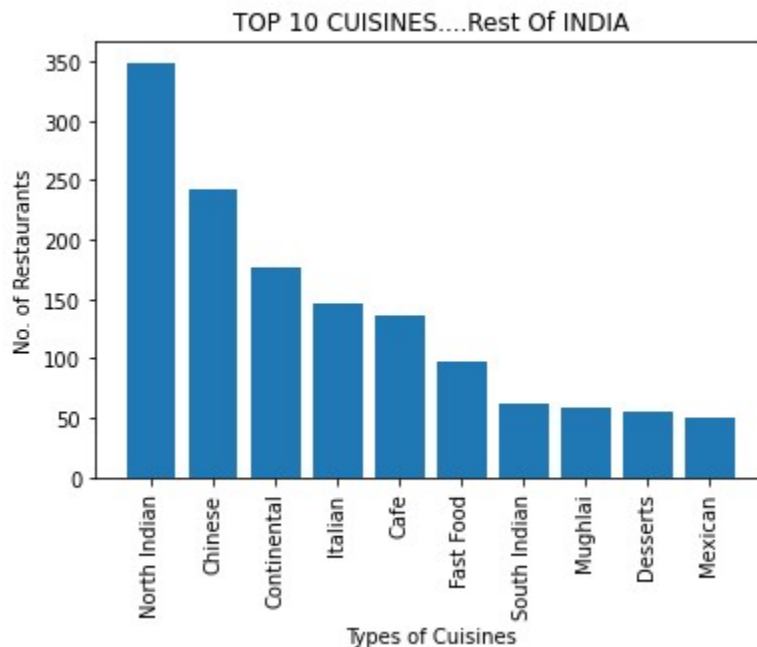
Desserts 55

Mexican 50

4. Write a short detailed analysis of how cuisine served is different from Delhi NCR to the Rest of India.

Plot a suitable graph to explain your inference.



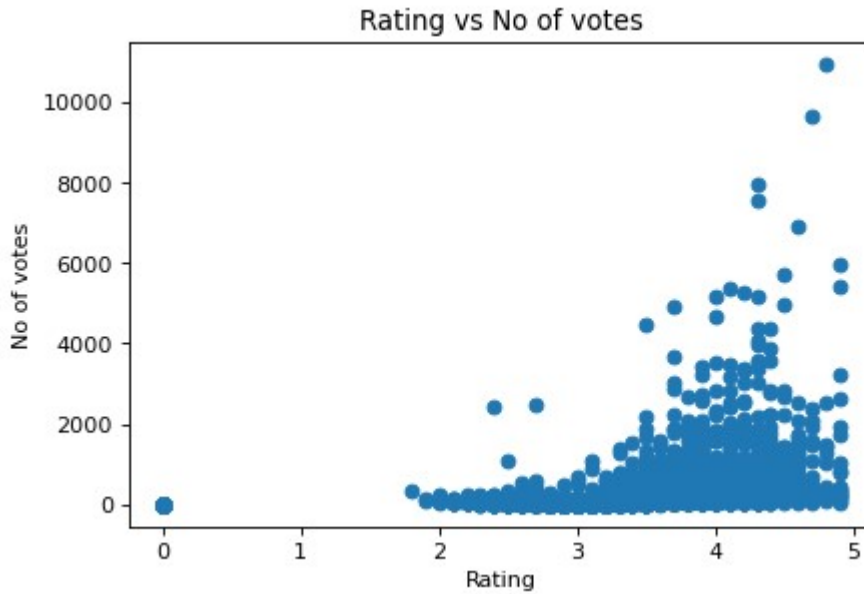


From the graph we can see that North Indian, Chinese Cuisine is served most in both places. But number of restaurants serves has a huge difference, North Indian cuisines is served by more than 3500 restaurants in Delhi NCR but this data changes drastically in Rest of India, and that is only 350 or so and same goes for Chinese cuisine. Third most served cuisine in Delhi NCR is Fast Food wherein Rest India is Continental and there is some order difference in terms of restaurant serves. There are total 86 different cuisines that are offered in Delhi-NCR wherein Rest India offers 70 different cuisines.

2. User Rating of a restaurant plays a crucial role in selecting a restaurant or ordering the food from the restaurant.

Write a short detailed analysis of how the rating is affected by the restaurant due following features: Plot a suitable graph to explain your inference.

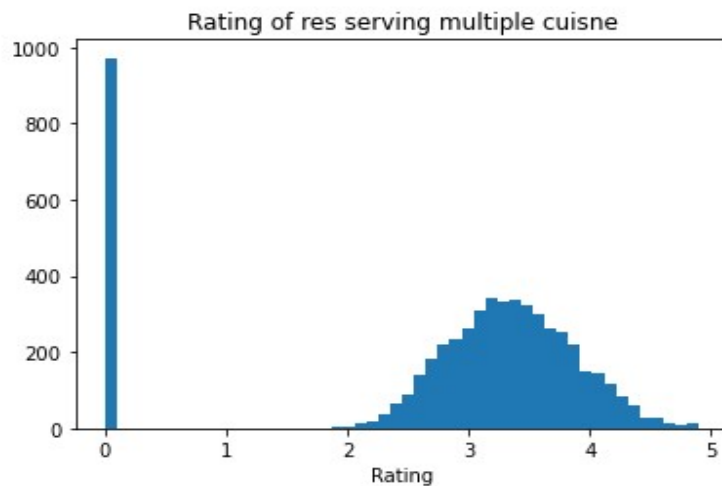
1. Number of Votes given Restaurant



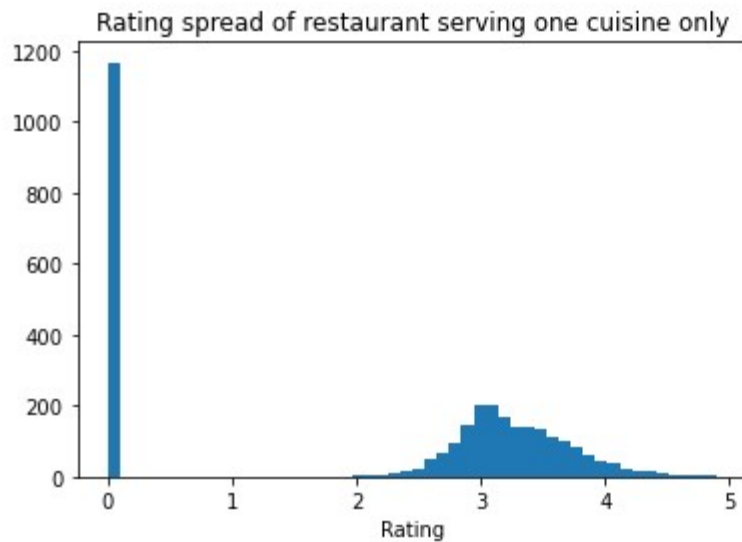
From the above graph, I can say that if we go with a large number of voting for a restaurant we will get good food usually.

Restaurant serving more number of cuisines.

Average rating of restaurants serving multiple cuisine : 2.784374437443748



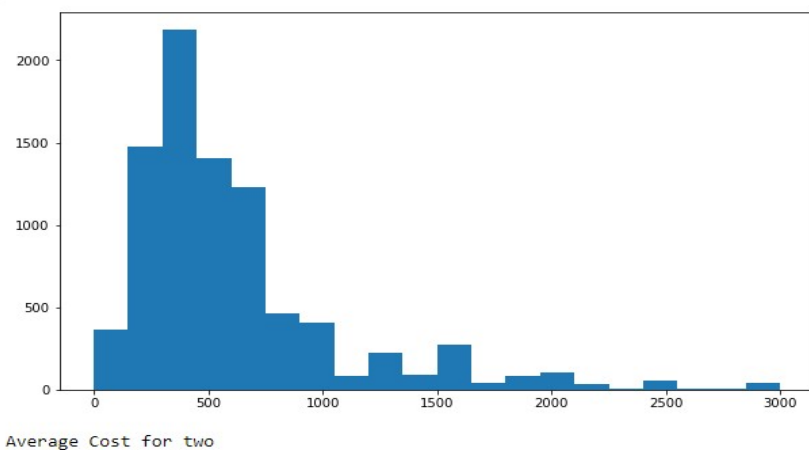
Average rating of restaurants serving one cuisine : 2.0550855666774277



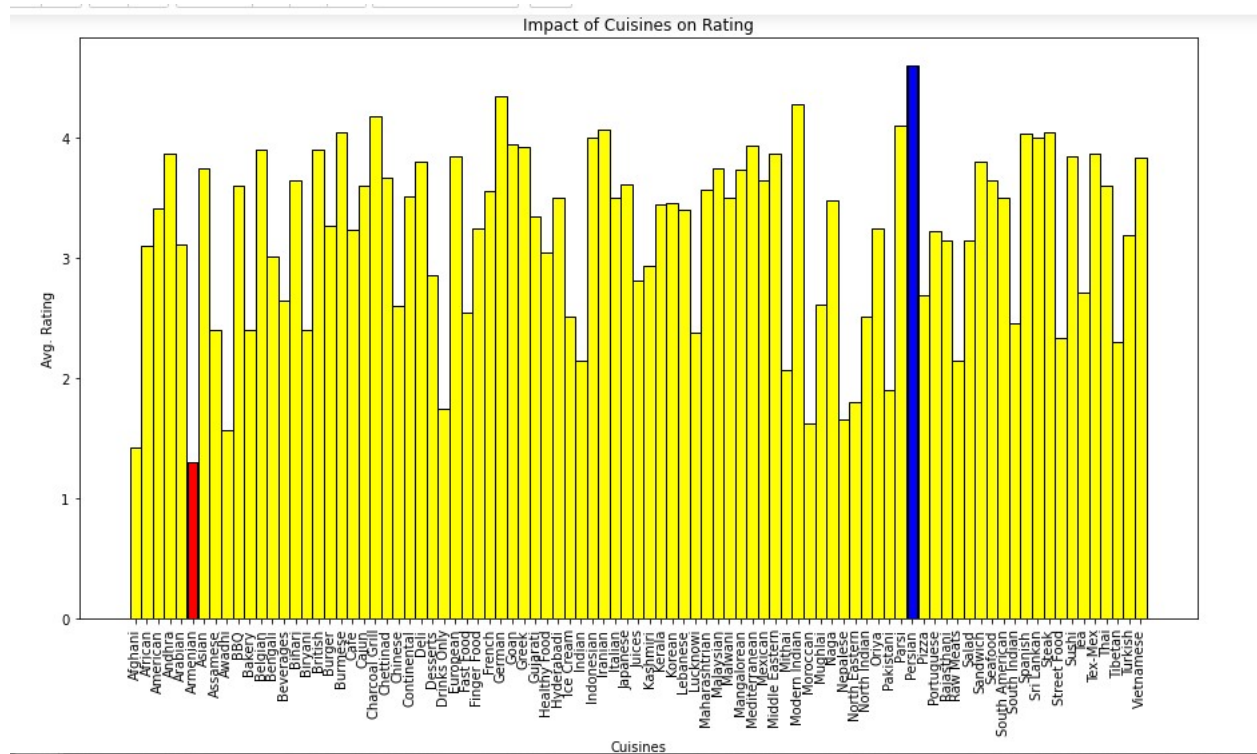
Hence, we can see that more number of restaurants serving multiple cuisine have a rating between 3 - 4 than compared to those serving only one Cuisine. Both multiple and single Cuisine restaurants are almost equally likely to have a rating below 1 Average rating of single Cuisine restaurant is 2.05 whereas for multiple cuisine restaurant is 2.7

Average Cost of Restaurant

NOTE: Data Compression is applied as numbers are huge and it might not fit in the graph properly.



Restaurant serving some specific cuisines.



1. Restaurants serving “Persian” cuisine have highest rating.
2. Restaurants serving “Armenian” cuisine have lowest rating.

Q. Find the weighted restaurant rating of each locality and find out the top 10 localities with more weighted restaurant ratings? $\text{Weighted Restaurant Rating} = \frac{\sum (\text{number of votes} * \text{rating})}{\sum (\text{number of votes})}$.

```
Aminabad 4.9
Friends Colony 4.886916367367881
Express Avenue Mall, Royapettah 4.8
Deccan Gymkhana 4.8
Sector 5, Salt Lake 4.707022552098202
Chittoor Road 4.6
Arambol 4.6
Marathahalli 4.57508238276299
Anjuna 4.5446681580909765
Ellis Bridge 4.438470451911935
```

Visualization

3.1 Plot the bar graph top 15 restaurants have a maximum number of outlets.

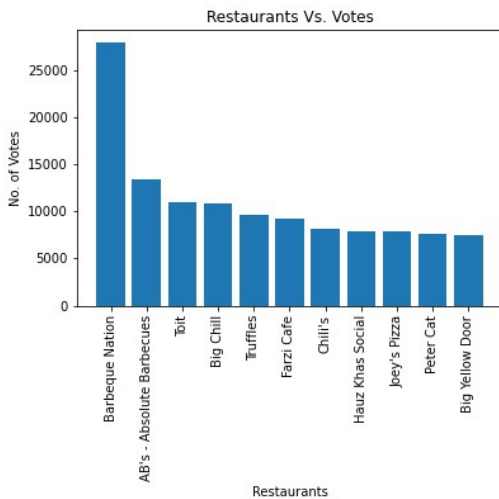


X

3.2 Plot the histogram of the aggregate rating of a restaurant(drop the unrated restaurant).



3.3 Plot the bar graph top 10 restaurants in the data with the highest number of votes

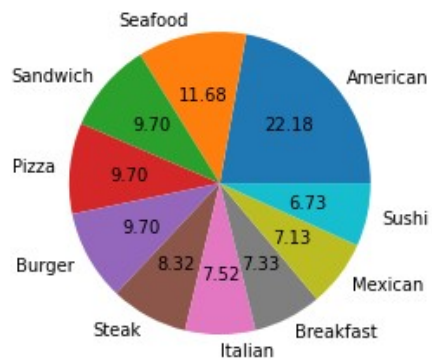


```
[[27835, 'Barbeque Nation'], [13400, "AB's - Absolute Barbecues"], [10934, 'Toit'], [10853, 'Big Chill'], [9682, 'Truffles'], [9189, 'Farzi Cafe'], [8156, "Chili's"], [7931, 'Hauz Khas Social'], [7807, "Joey's Pizza"], [7574, 'Peter Cat']]
```

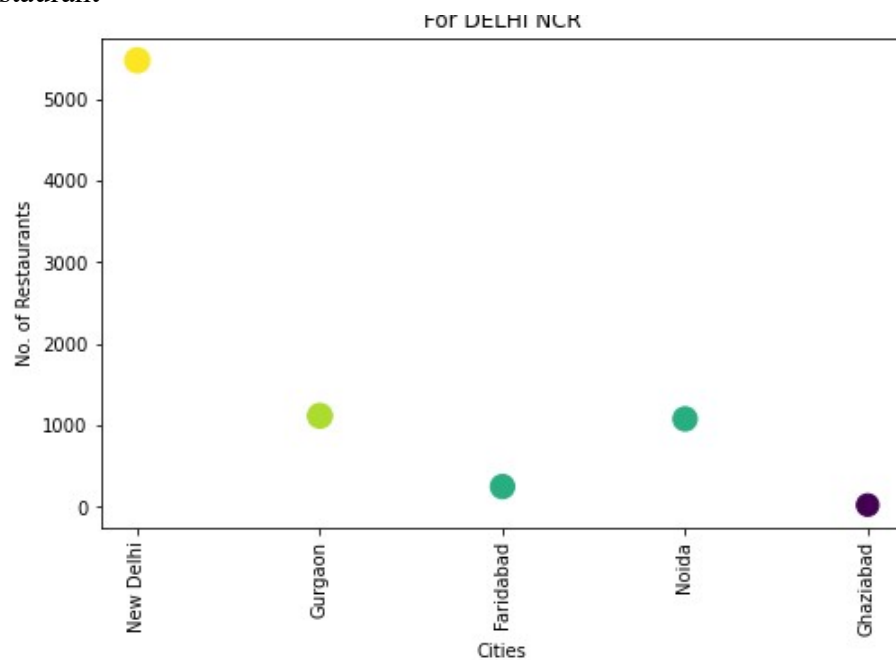
3.4 Plot the pie graph of the top 10 cuisines present in restaurants in the USA.

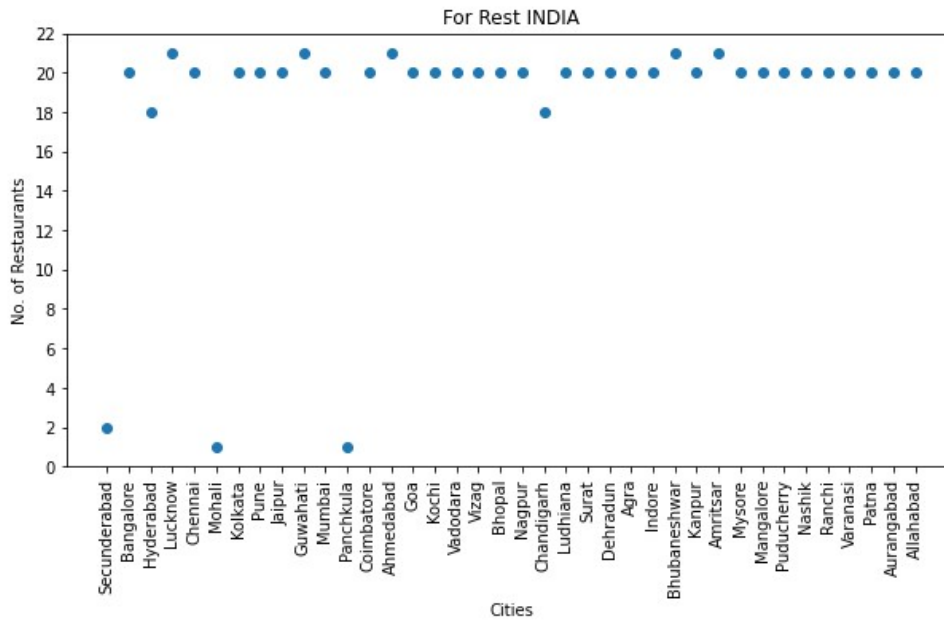
Rank	Cuisine	%
1	American	22.18
2	Seafood	11.68
3	Sandwich	9.7
4	Pizza	9.7
5	Burger	9.7
6	Steak	8.32
7	Italian	7.52
8	Breakfast	7.33
9	Mexican	7.13
10	Sushi	6.73

-
- 1 American 22.18
 - 2 Seafood 11.68
 - 3 Sandwich 9.7
 - 4 Pizza 9.7
 - 5 Burger 9.7
 - 6 Steak 8.32
 - 7 Italian 7.52
 - 8 Breakfast 7.33
 - 9 Mexican 7.13
 - 10 Sushi 6.73



3.5 Plot the bubble graph of several restaurants present in the city of India and keeping the weighted restaurant





Since given Dataset is skewed towards Delhi NCR, therefore I plotted graphs separately for Delhi NCR and Rest of India. Otherwise, representation would not be helpful in getting insights from data. Sizes of bubble may not reflect weighted rating of restaurant as these are very close. But slight change in Colors of bubbles definitely show difference in their size