1.1

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

I considered country India by using country code 1
df = df[df['Country Code'] == 1]
And then I reset the index by using
df.reset_index(drop = True, inplace = True)
After that I replaced the NCR cities name to Delhi-NCR
And rest the cities of India as 'Rest of India'.
Then I used value count function to calculate no of restaurants in Delhi and Rest of India
And then plotted them on a bar graph using plt.bar
According to the pie chart we can see that the dataset contains 91.85% restaurants in Delhi-NCR .

1.2 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.

For this I created two set with name NCR_cuisines and Rest_cuisines
Then I used the split function to split the contents of df.Cuisines by ',' .

Then I created 2 datasets NCR and Rest with data of Delhi-NCR and Rest of India resp. And then added contents of NCR in NCR_cuisines and Rest in Rest_cuisines .
Then I used (Rest_cuisines - NCR_cuisines) to get unique items.

1.3

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

For this part of ques I created 2 dict with name ncr and rest resp. And then added cuisines of NCR in ncr and Rest in rest

And then I used a sort function to sort them and give top 10 cuisines only.

1.4 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.

In this ques I plotted 2 pie chart one with Delhi-NCR cuisines using Delhi_cuisines_names for name of cuisines and Delhi_cuisines_counts = [ ]
For the count of cuisines and the same for Rest of India.

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

1. **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

   For this ques I used plt.plot to plot aggregate rating and votes

   2.Restaurant serving more number of cuisines

   For this ques I splitted the cuisine column by ',' and and then created a new column which contains len of cuisines
   And then sorted the values by Aggregate rating and the using plt.scatter to plot them

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

   For this ques I used plt.hist to plot Average Cost for two on x axis and range[0,6000] on y axis.

   4.Restaurant serving some specific cuisines.

   For this ques I created 2 variables with Aggregate rating and Number of cuisines in resp variables and then I created a booleans variable where cuisine<=2
And then selected cuisine and rating where cuisine<=2

And plotted them suing plt.scatter.

2.2.1

Find the weighted restaurant rating of each locality and find out the top 10 localities with more weighted restaurant rating?

Weighted Restaurant Rating=Σ (number of votes * rating) / Σ (number of votes)

In this ques I used group by function to group the data by Locality and used apply function to find weighted Restaurant rating using

```
x = df.groupby(by = 'Locality').apply(lambda x:
((x['Votes'] * x['Aggregate rating']).sum()))
y = df.groupby(by = 'Locality').apply(lambda x:
(x['Votes'].sum()))
weighted_rating = x/y
round(weighted_rating, 3)
```

After that I sorted the data in variable sorted_weighted_rating where top 10 localities are printed using print function.

3.1

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

In order to plot top 15 restaurants I first created a list res_name

And added items of df[‘Restaurant Name’]

Then I created a dictionary y to count values of restaurant names

Then a=y.keys and b=y.values

Then I created a data frame of a and b and sorted them by No. of outlets and then printed them and plotted them using plt.bar

Result

```
Cafe Coffee Day 83
Domino's Pizza 79
Subway 63
```

```
Green Chick Chop 51
McDonald's 48
Keventers 34
Pizza Hut 30
Giani 29
Baskin Robbins 28
Barbeque Nation 26
```

3.2

Plot the histogram of aggregate rating of restaurant(drop the unrated restaurant.

In this Ques I dropped rows where Rating text === Not rated
And plotted df[‘plt.hist(df['Aggregate rating'] using plt.hist

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

In order to get restaurant with highest number of Votes I used groupby function and used sum function to get restaurant with highest number of Votes and plotted them using plt.bar

3.4

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

To do this question I used data where the country code was 216 in order to get data from the USA.
Then I created a dict and added cuisines with their count and then sorted them on the basis of items of dict and used[:10] to get the top ten cuisines and used plt.pie to plot it.

Result:

American 112
Seafood 59
Burger 49
Sandwich 49
Pizza 49
Steak 42
Italian 38
Breakfast 37
Mexican 36
Sushi 34

3.5

Plot the bubble graph of a number of Restaurants present in the city of India and keep the weighted restaurant rating of the city in a bubble.

In this ques I used groupby function to group data by City and then used apply function to find weighted_rating.
And then plotted restaurant_count on x axis and Restaurant names on y axis and weighted_rating in bubble using plt.scatter