***Zomato sales analysis and insights***

**1. What is the total no. of tables present in the data?**

A table is a way to organise data in rows as columns, making it easier to compare and analyse information. Tables is a means to display data into an orderly and structures way.

In the dataset provided the total numbers of tables present are two. One is the raw data and other one is the country description.

We have created one more table currency which will help in converting prices in US dollars.

**2. What is the total no. of attributes present in the data?**

Attributes are an identifying piece of information that describes an entity. Each attribute has a specific data type.

With respect to the data provided the attributes will be the different columns holding different information.

There are in total 20 attributes.

**3. How many categorical columns are there in the data?**

The categorical data consists of categories which represent the characteristic such as gender, hometown, date of birth. Categorical data is expressed in terms of natural descriptions.

In the dataset provided there are 14 categorical columns which are as below :

a. Country code

b. City

c. Address

d. Locality

e. Locality verbose

f. Longitude

g. Latitude

h. Date key opening

i. Currency

j. Cuisines

k. Has table bookings

l. Has online delivery

m. Price range

n. Rating

**4. The data consists of some inconsistent and missing values so ensure that the data used for further analysis is cleaned.**

The data had some inconsistent values in certain columns such as the Cuisines column, the average cost for two for certain restaurants had cost as zero and also the date key column was not formatted correctly.

For the cuisines column since the missing values were only for country united States of America, i calculated the most frequent value and filled the missing values.

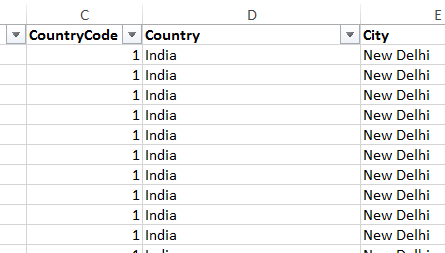
For the average cost for two column I calculated the mean and filled the values which were zero.

And for the date key opening column I used the find and replace functionality and replaced "\_" with "-" to make it a proper date.

**5. Using the LookUp functions, fill up the countries in the original data using the country code.**

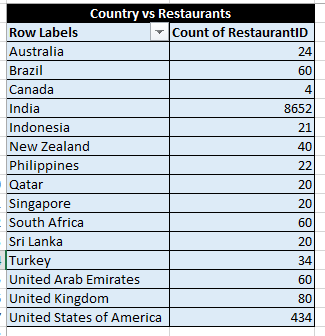
To fill up the countries in the original data using country country, I used a VLOOKUP function.

**=VLOOKUP($C2,'country description'!$A$1:$B$16,2,0)**



**6. Create a table to represent the number of restaurants opened in each country.**

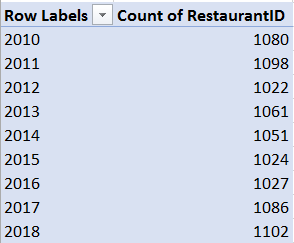
To achieve this I created a pivot table with the Country column is rows and the count of restaurant Ids in the values.



**7. Also, the management wants to look at the number of restaurants opened each year, so provide them with something here.**

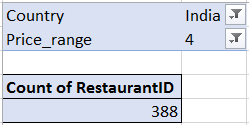
To achieve this I used pivot tables , Before that in the cleaned data I added new column which will have the opening date, opening month and opening year in separate columns using the TEXT function.

And then created a pivot table with the opening year in rows and the count of restaurant ID in the values.



**8. What is the total number of restaurants in India in the price range of 4?**

To find the total number of restaurants in India in price range 4, I created a pivot table taking country column in rows and count of restaurant id along with applying two filter one on country and other on price range.



**9. What is the average number of voters for the restaurants in each country according to the data?**

This can be done using Pivot tables. The country column will be in the rows and average of votes in the values.



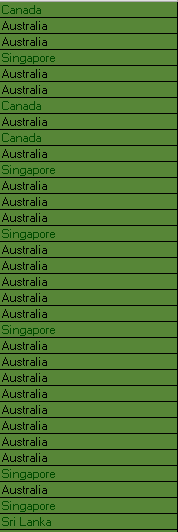
**10. Calculate the average rating for all the restaurants that have price\_range < 4 and provide online delivery. Use only the “IF” function, Logical Operators, and Aggregation functions to solve this problem. [Note: Don’t use Conditional aggregation in this question.]**

To calculate the average rating for all restaurants that have price\_range < 4 and provide online delivery the below formula is used.  
**=AVERAGEIFS('Cleaned data'!Y2:Y9552,'Cleaned data'!S2:S9552,"<4",'Cleaned data'!P2:P9552,"=Yes")**



**11. Using Conditional formatting highlight the rows of restaurants that are located in the countries or cities that you’ve suggested to the management for opening new restaurants.**

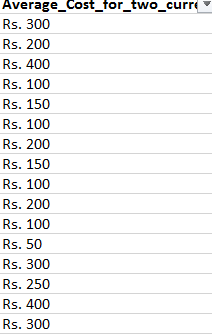
To apply the conditional formatting, select the data then click on conditional formatting>> Highlight cells rules > Text that contains



**12. Create a new customized price column that consists of the abbreviation/symbol of the currency along with the Average\_cost\_for\_two value. [Use string operations to do this task].**

To create a new column with customised price column that consists of abbreviation/symbol of the currency along with the average proce for two, I used the below formula.

**=CONCATENATE(MID(M:M,SEARCH("(",M:M)+1,SEARCH(")",M:M)-SEARCH("(",M:M)-1)," ",$V2)**



**13. How can you create an array formula in Excel or Google Sheets to count the number of restaurants listed that do not offer online delivery, are in the lowest price range, and have an average cost for two people less than or equal to 250 Indian Rupees?**

There are total 1293 restaurants having lowest price range, do not offer online delivery and have average cost less than 250 Rs.

**=COUNTIFS('Cleaned data'!Q:Q,"No",'Cleaned data'!T:T,"1",'Cleaned data'!Y:Y,"<2.99").**

2.99 because in the cleaned data, I have added a helper column where the price is converted to US dollars and 250 Rupees = 2.99 USD

**Subjective Questions :**

**1. Suggest a few countries where the team can open newer restaurants with lesser competition. Which visualization/technique will you use here to justify the suggestions?**

To suggest few countries where the team can open newer restaurants with lesser competition are :

**Canada, Australia, Singapore & Sri Lanka.**

To find out I created Pivot table with the country in the rows and the count of restaurant id along with the average rating in the values.

After analysing the data the countries with less number of restaurants and less rating were more suitable for opening of new restaurants as because less number of restaurants mean less competition.

**Reference – Subjective question sheet.**

**2. Come up with the names of States and cities in the suggested countries suitable for opening restaurants.**

To suggest the cities for opening restaurants same technique was considered i.e less the number of restaurants less the competition and lower the rating , greater chance of growth.

Pivot table was created with country and cities in rows and average rating and count of restaurant ID in values along with country in filter so that only suggested countries can be selected.

In **Canada** the restaurant can be opened in the **Consort** **city**.

In **Australia** the restaurants can be opened in **Montville, Paynesville & Mayfield.**

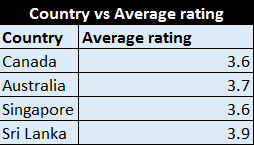
In **Singapore**, the restaurant can be opened in the **Singapore** city.

And in **Sri** **Lanka**, **Colombo** is suitable for opening new restaurants.

**Reference – Subjective question sheet.**

**3.According to the countries you suggested, what is the current quality regarding ratings for restaurants that are open there?**

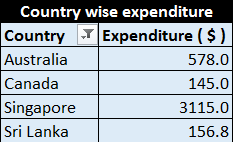
**According to the countries suggested, the current quality regarding ratings are as follows :**

 **Reference – Subjective question sheet. ( country vs average rating table )**

**Key insights –** The average rating in the suggested countries is between 3.6 to 3.9 which is on the lower side and suggests that the customers are not very satisfied .

**4.Also, what is the current expenditure on food in the suggested countries, so we can keep our financial expenditure in control?**

**The current expenditure of food in the suggested countries are as follows :**



**The above expenditure is calculated in US Dollars.**

**A new column was introduced in the cleaned Data, which converted the average price for two column in US Dollars for all countries.**

**Reference – Subjective question sheet. ( country wise expenditure table )**

**Key insights –** As compared to all the other countries the expenditure is on the lower side, so while opening new restaurants the expenditure can be slightly increased.

**5.Come up with the names of restaurants from the recommended states that are our biggest competitors and also those that are rated in the lower brackets, i.e. 1-2 or 2-3.**

According to the analysis the restaurants that are the biggest competitors in Australia are **Bridge Road Brewers, 1918 Bistro & Grill and Vivo Bar & Grill.**

In **Canada, Lake House Restaurant** is the biggest competition.

In **Singapore**, **Al’frank Cookies** and **Fratini La Trattoria** are biggest competiton.

And in **Sri Lanka, Ministry of Crab** and **Simply Strawberries by Jargo** are biggest competition.

**Reference – Country vs restaurant sheet (Biggest competitors and low rated restaurants table )**

**6.Which cuisines should we focus on in the newer restaurants to get better feedback? Does the choice of cuisines affect the restaurant ratings?**

Each country offers different cuisines, so in order to answer this question we need to first find out the most desired cuisines in every country.

A pivot table was created with country and cuisines in row, average of ratings in values and country in filter to filter out the suggested countries.

**After analysing the data the cuisines which should be focused are :**

**In Canada** since **Italian, Mediterranean, Pizza** are highly rated so these can be focused.

In **Australia Pizza, Seafood, Mediterranean & Bar Food** are rated highly so those can be offered in there.

In **Singapore Italian cuisines** should be focused.

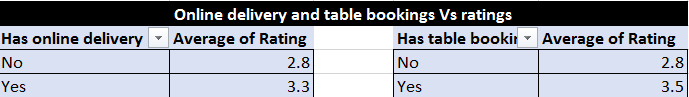
Similarly in **Sri Lanka** based on the average rating customers are most likely to eat **Seafood, Chinese and Juices & Desserts.**

**Reference – Cuisines sheet ( Cuisines vs average rating table )**

**7.According to our current data, should we go for online delivery and table booking? Does that affect the customer’s ratings?**

According to the current data, it is advisable to go for online delivery and table bookings since the data shows that the restaurants which have table bookings and provide online delivery are rated highly than the restaurants which do not offer.

This was analysed with the help of pivot tables.



**Reference – Ratings sheet ( Online delivery and table bookings vs ratings )**

**8.Should the team keep the rate of cuisines higher? Will that affect the feedback? According to our data are the rates of cuisines and ratings, correlated?**

To achieve this a Pivot table was created with the ratings in rows and the average cost of two in USD in the values.

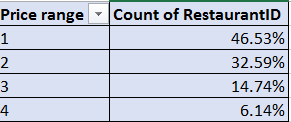
Then a correlation was calculated between the ratings and the average cost for two which came out to be 0.80784.

A positive sign means a string correlation which means when the restaurant ratings increases the prices tend to increase as well and vice versa.

**Reference – Ratings Sheet**

**9.What is the distribution of the number of restaurants of different price ranges in all the countries?**

For this I created a Pivot table with price range in rows and count of restaurant ID in the values.  
And it was clearly visible that almost 47 % of the total restaurants fall in the lower price range.



**10..Explain your approach in brief for suggesting countries/cities in order to open new restaurants, if the objective and subjective questions would have not been given to assist you. [you have to give bullet pointers in order to answer this question].**

* In order to suggest countries/cities for opening of new restaurants there were few data points which were considered.
* In the countries/cities which are already crowded and flooded with good number of restaurants, there is no point in opening new restaurants, so first of all figured out the countries which have less number of restaurants as lower the number of restaurants lower the competition for newer ones.
* The next key consideration was that the countries/cities which have the less average rating directly means that the customers are not satisfied with the quality of food. So new restaurants can be opened there with high chances of growth.
* Another key point would be countries/cities where the average cost of two would be lower since lower cost would attract more number of customers.
* The above approaches can be better analysed with the help of charts and Pivot tables.