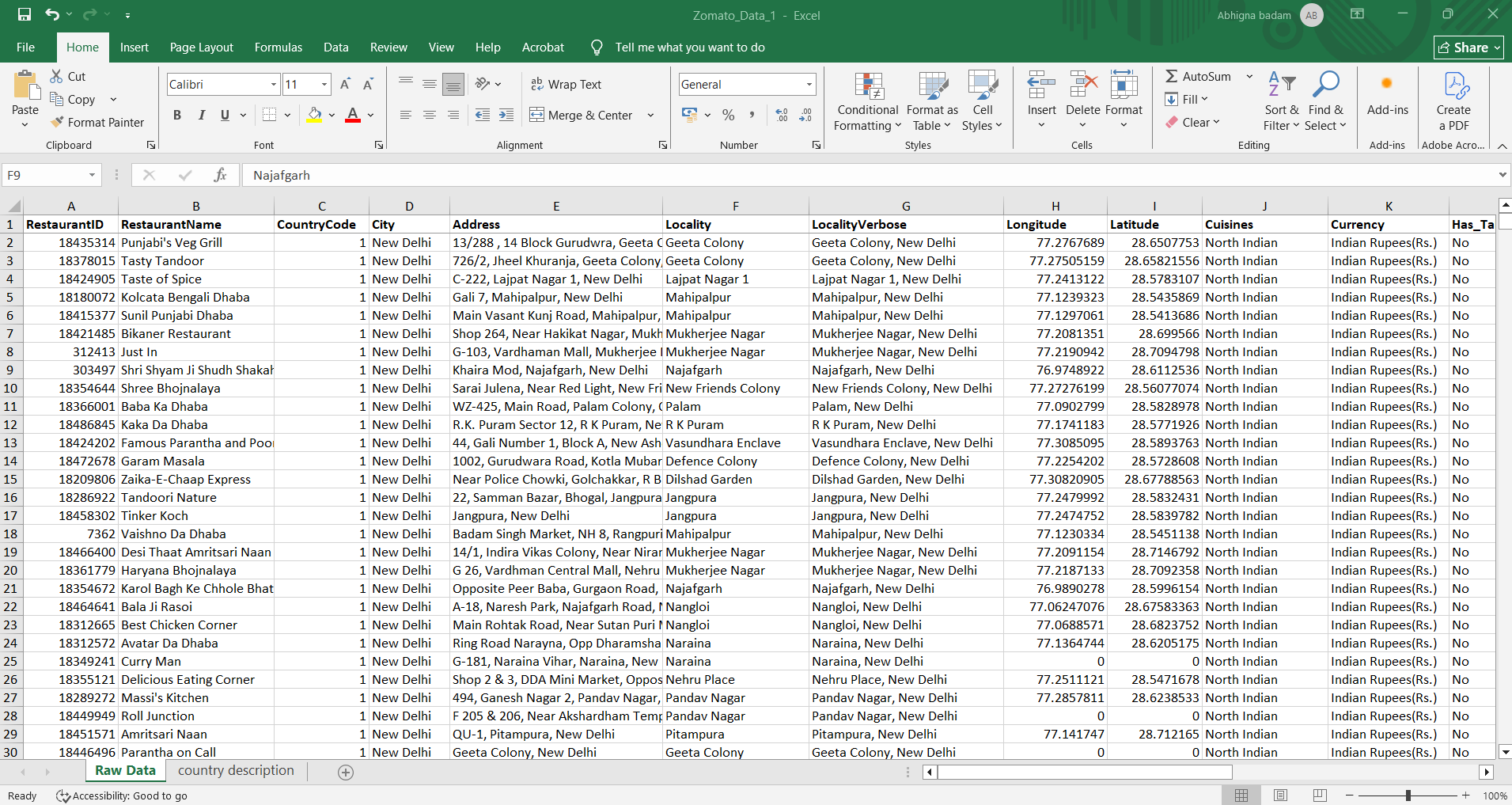
# **Zomato SS Project**

# **Tasks**

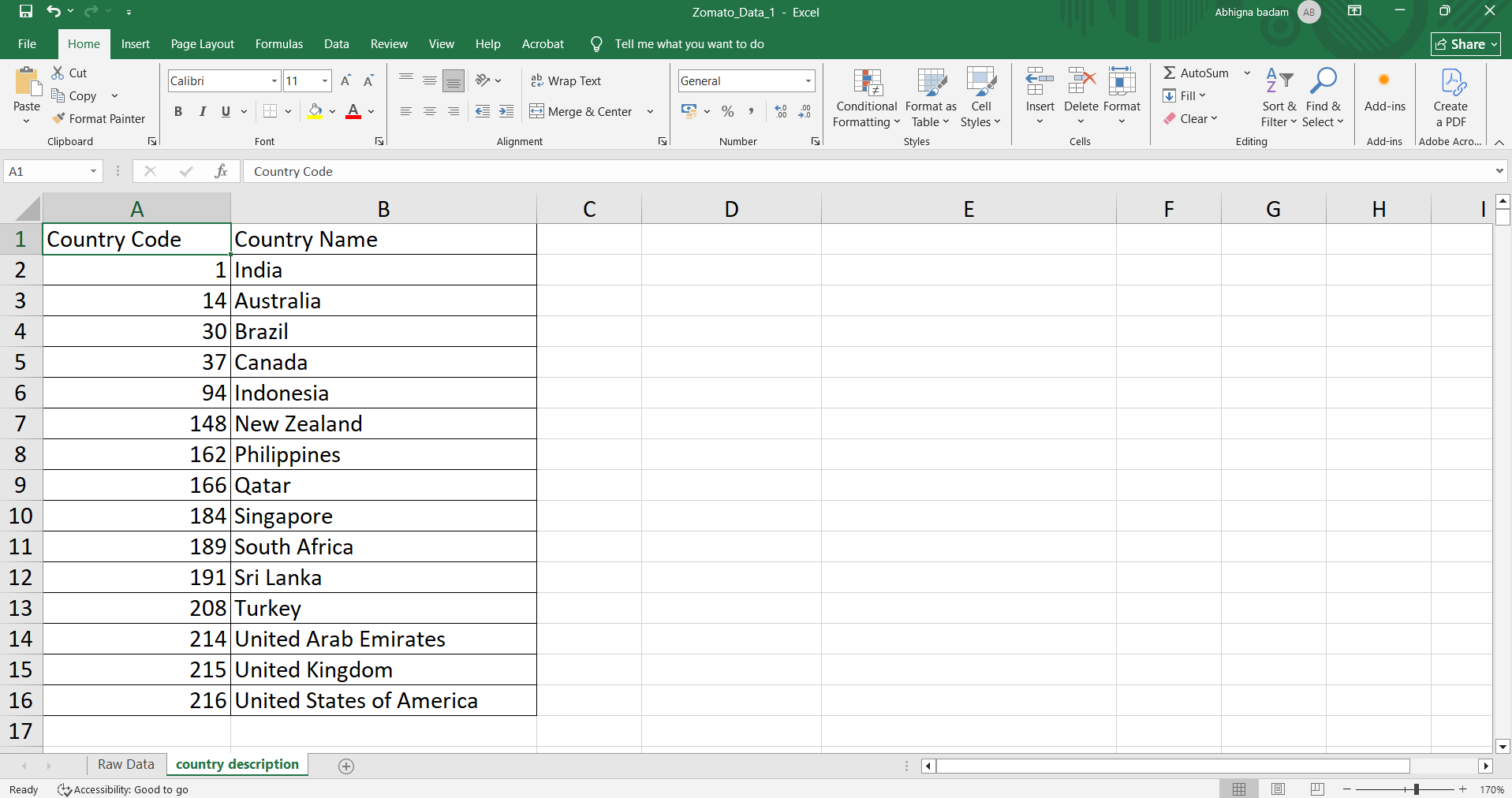
**Objective Questions:**

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

There are two sheets in the workbook that has project data and there is one table in each sheet. So, there are **2 tables** in the data   
Sheet named “**Raw Data**” has one table consisting of data of restaurants



And in that worksheet named “**country description”** has one table consisting of country code and corresponding country name



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

There are **20** attributes in worksheet “**Raw Data**” and there is one more extra attribute in worksheet “**country description**”. So, in total there are **21** unique attributes. The attributes are: -

1. **Restaurant ID** : **Primary key**, Unique identifier of each restaurant.
2. **Restaurant Name :** The name of the restaurant
3. **CountryCode :** Country Code of the location where the restaurant is situated.
4. **Country :** Country where the restaurant is situated.
5. **City :** The city in the country where the restaurant is located.
6. **Address:** The specific address of the restaurant.
7. **Locality:** The locality or neighbourhood where the restaurant is situated.
8. **Locality Verbose:** Detailed information about the locality.
9. **Longitude :** The geographical longitude coordinate of the restaurant.
10. **Latitude :** The geographical latitude coordinate of the restaurant.
11. **Cuisines :** The type of cuisine offered by the restaurant.
12. **Currency :** The currency used for the transaction in the restaurant
13. **Has\_table\_booking :** Indicates whether the restaurant has a table booking option.
14. **Has\_online\_delievery :** Indicates whether the restaurant offers online delivery.
15. **Is\_delievering\_now :**Indicates whether the restaurant is currently delivering.
16. **Switch\_to\_order\_menu :** Indicates whether the users can switch to the order menu.
17. **Price\_range :** The numeric value indicating the price range category of the restaurant.
18. **Votes :** No of votes or ratings received by restaurant.
19. **Average\_Cost\_for\_two :** The average cost for two people dining at the restaurant.
20. **Rating :** The overall rating of the restaurant is based on user reviews.
21. **Datekey\_opening :** The date when the restaurant was opened**.**
22. **How many categorical columns are there in the data? [Search about categorical and continuous data, and try to answer this question]**

Total Categorical columns = 15

1. **Restaurant ID**
2. **Restaurant Name**
3. **Country Code**
4. **Country Name**
5. **City**
6. **Address**
7. **Locality**
8. **Locality Verbose**
9. **Cuisines**
10. **Currency**
11. **Has Table Booking**
12. **Has Online Delivery**
13. **Is Delivering Now**
14. **Switch to Order Menu**
15. **Price Range**
16. **The data consists of some inconsistent and missing values so ensure that the data used for further analysis is cleaned.**
17. First checked for duplicate values. There are no duplicate values in the data.
18. Coming to missing values, applied filter for all columns to check the blank values, only **Cuisines** column has 9 missing values (0.1%) and all restaurants belongs to country USA. Inserted a pivot table to understand which cuisine is more famous but didn’t find anything with required condition. So, I removed the rows having blank values.

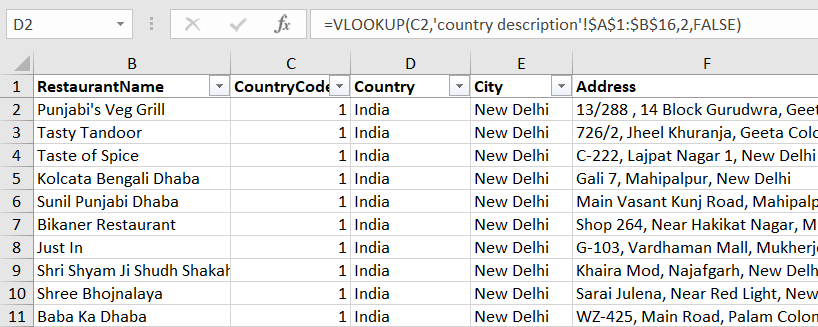
In original data there 9551 rows after removing this rows there are 9542 rows.

1. In case of inconsistent values, I formatted the values of **Datekey\_opening** column replaced the “\_” with “-“and fixed their datatype as “Date” as formatted as “dd-mm-yyyy”
2. Changed data type of all columns.
3. Didn’t found any column having extra space or special characters to remove or trim and all are in proper format.
4. **Using the LookUp functions, fill up the countries in the original data using the country code.**

Inserted new column on the right side of column “CountryCode” or column C  
I named the column as “Country”   
Used VLOOKUP to fill the values of “Country” column taking table array as table in “country description” sheet

The formula in D2 cell is :

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

(from before question there is no missing value in countrycode column so don’t need to use ”ifna” function)

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

Inserted a pivot table adding country in Rows and RestaurantID in values and used field setting to count

****

**India** has the most no of restaurants opened around 8652 restaurants which is 90% of total restaurants and next **USA** which has 425 which is only 4%.

Least no of restaurants in **Canada**

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

****

Observing no of restaurants opened in every year there is increment in year 2011 but there is big drop in year 2012 and next 2 years there is good no of restaurants opened but again there is a drop in year 2015 and 2016 and there is good increment in year 2017 and 2018.

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

****

No of restaurants with a price range of 4 in India is 388 which is around 4% of no of restaurants in India

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



The average number of voters is bigger in country Indonesia with having restaurants count 21 followed by UAE and Turkey and least is in country Brazil having no of restaurants 60 .

Even having high no of restaurants India has less no of average votes .

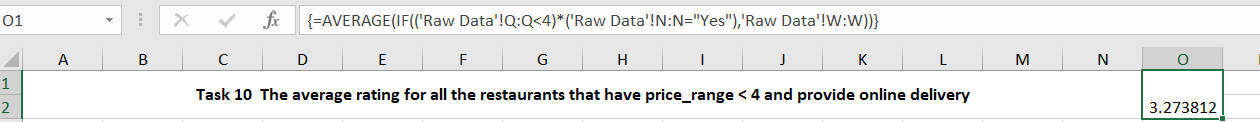
1. **Calculate the average rating for all the restaurants that have price\_range < 4 and provide online delivery. Use only the “IF” function, Logical Operators, and Aggregation functions to solve this problem. [Note: Don’t use Conditional aggregation in this question.]**

It can be done using two methods

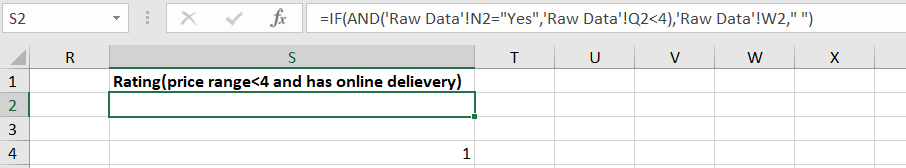
Method 1: - Using Array formula



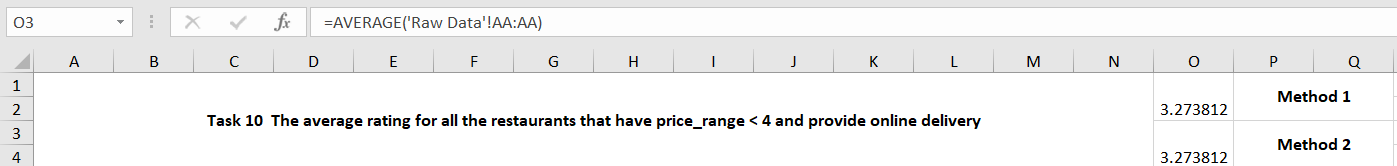
Where in Raw data sheet Column Q is price range as column N is has\_online\_delievery as column W consists of Rating.



Method 2 In this method created a customized column in a new sheet. In values of these column if our conditions are true then the value is our rating if not it is null. These is achieved by using IF and AND function and then applied average of the values in these columns



And did the average of the values of the column



Average rating using both methods is same and equals to **3.273812 approximately 3.3**

1. **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.**

The list of I’ve suggested to open new restaurants are

1. Indonesia: Bandung, Jakarta, Tangerang
2. New Zealand: Auckland, Wellington City
3. Philippines: Makati City, Mandaluyong City, Pasay City, Pasig City, Quezon City, San Juan City, Tagaytay City, Taguig City
4. Qatar : Doha
5. Sri Lanka: Colombo
6. Turkey : Ankara, Istanbul

Applied the following steps to do conditional Formatting:

1. Selected the entire data of Raw Data Sheet
2. Home → Conditional Formatting → New Rule.
3. Choose “Use a formula to determine which cells to format”.
4. Entered the formula

=OR($E2="Bandung",$E2="Jakarta",$E2="Tangerang",$E2="Auckland",$E2="Wellington City",$E2="Makati City",$E2="Mandaluyong City",$E2="Pasay City",$E2="Pasig City",$E2="Quezon City",$E2="San Juan City",$E2="Tagaytay City",$E2="Taguig City",$E2="Doha",$E2="Colombo",$E2="Ankara",$E2="ÛÁstanbul")

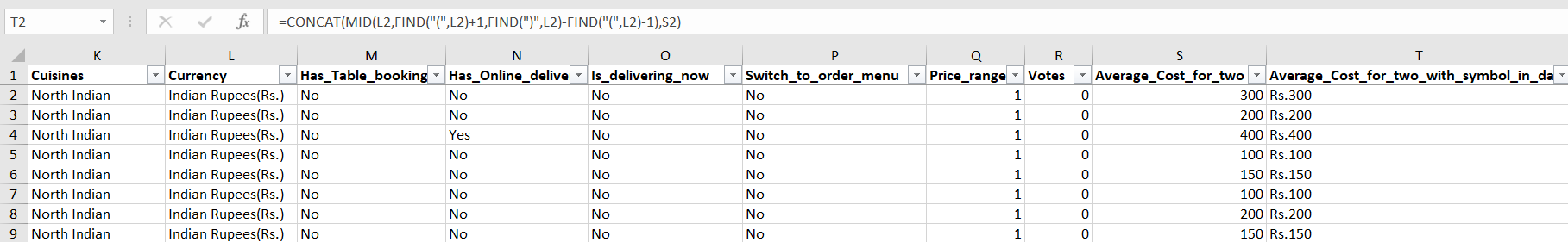
1. And next Format option I choose fill colour as Yellow
2. And next Ok

The restaurants of the cities the I’ve suggested got Highlighted.

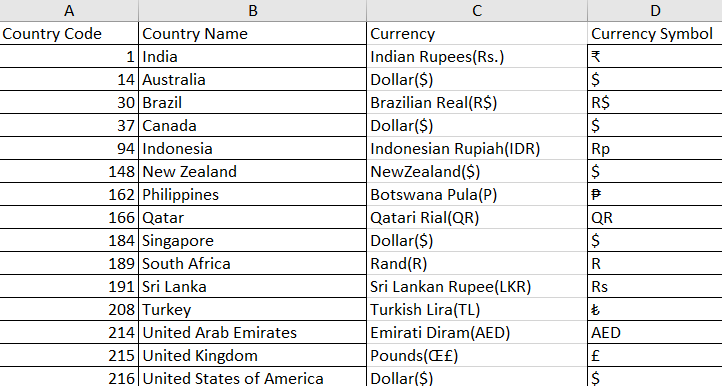
1. **Create a new customized price column that consists of the abbreviation/symbol of the currency along with the Average\_cost\_for\_two value. [Use string operations to do this task]**

For this one I have created two customized columns   
One with extracting the currenct abbreviation from currency column(column L) using RIGHT and FIND and used CONCAT function to merge both currency abbreviation and Average\_cost\_for\_two(column S)

=CONCAT(MID(L2,FIND("(",L2)+1,FIND(")",L2)-FIND("(",L2)-1),S2)

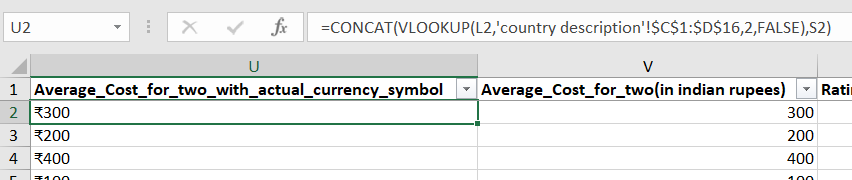


And about other column so I myself googled and entered actual currency symbol for all the countries in country description sheet



And using VLOOKUP and CONCAT merged this currency symbol and average\_cost\_for\_two

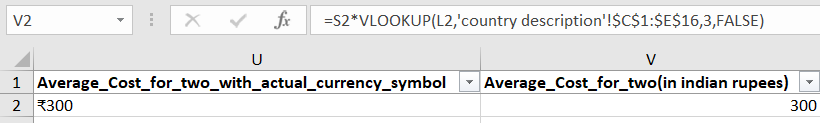
=CONCAT(VLOOKUP(L2,'country description'!$C$1:$D$16,2,FALSE),S2)



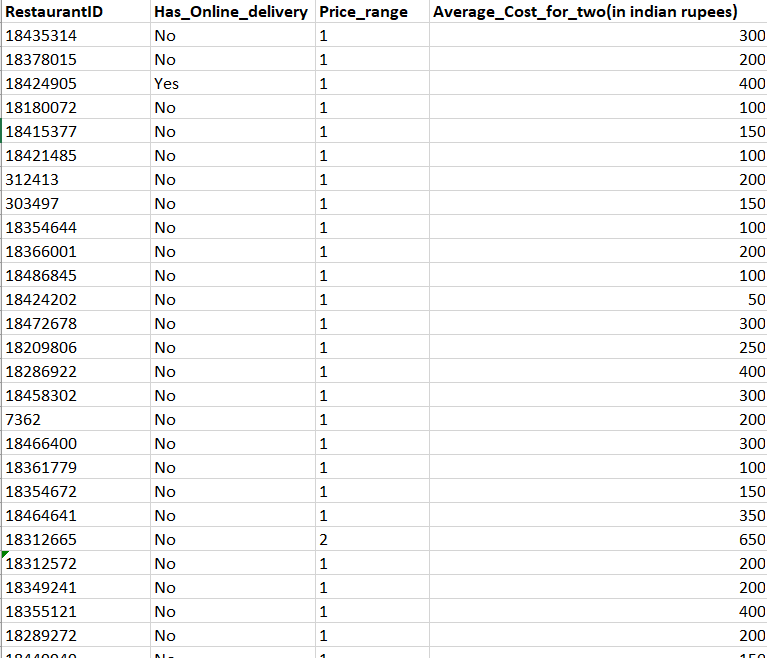
1. **How can you create an array formula in Excel or Google Sheets to count the number of restaurants listed that do not offer online delivery, are in the lowest price range, and have an average cost for two people less than or equal to 250 Indian Rupees?**

With the help of google created a new column in country description sheet the column consists of conversion rate that can help to convert the different currency into Indian currency

Created a new column Average\_cost\_for\_two(in Indian rupees) in raw Data sheet using VLOOKUP multiplied the average\_cost\_for\_two and the conversion rate

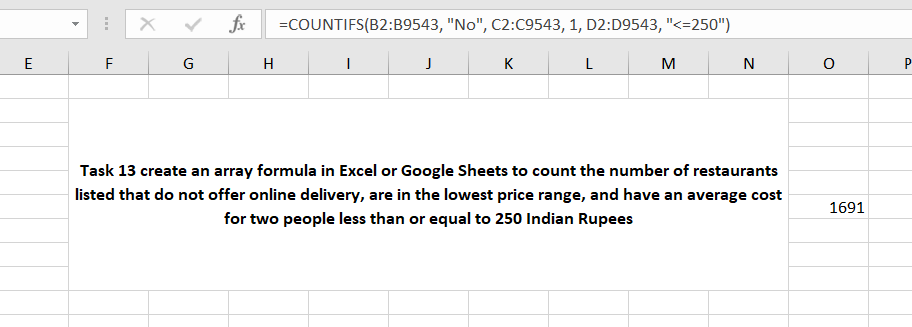
 

Copied the columns RestaurantID, Has\_online\_delivery, Price Range and Average\_cost\_for\_two into new sheet named “Objective task 13”



And next applied Array Formula to count the restaurants that do not offer online delivery,are in lowest price range and have an average cost for two people less than or equal to 250 Indian Rupees.

=COUNTIFS(B2:B9543, "No", C2:C9543, 1, D2:D9543, "<=250")

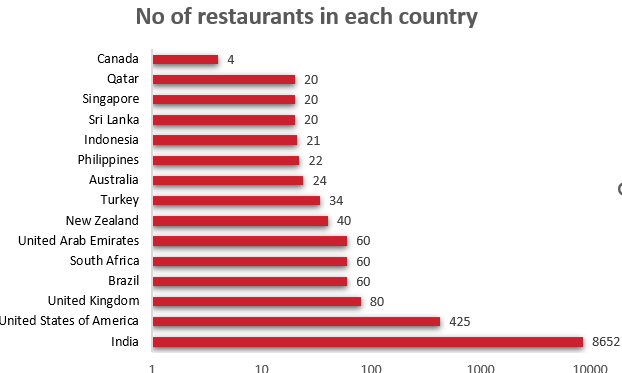


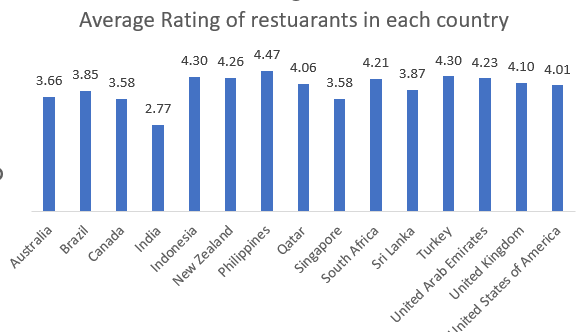
**Subjective Question:**

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?**

Created a pivot table added Country in Rows, RestaurantID (summarized values by Count) and Rating (summarized by average) in Values

****

****

****

**Insights**

The countries India and USA the restaurant market is highly saturated and average ratings of restaurants is very low in India  
For countries like Brazil, UAE, UK, South Africa the restaurant market is not that highly saturated or not that low and average ratings are pretty good.

In Australia, Singapore, Sri Lanka the saturation is low and rating are moderate

And countries Indonesia, New Zealand, Philippines, Qatar, Turkey the market is low and the ratings are high

And lastly Canada has very low number of restaurants and rating is moderate

**Suggestion**

**The countries that are best to open new restaurant are**

**1.Indonesia 4.Qatar**

**2.New Zealand 5.Turkey**

**3.Phillipines 6.Sri Lanka**

Because the number are very low and ratings of restaurants are high in these countries so that we can the competition is very low in these countries and people are willing to go restaurants and restaurants has good market in these countries. So, these countries are better option to open new restaurant.

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

Created a pivot table with Country and City in rows and RestaurantID(count) and rating(average) and applied filter for country and filtered the countries that are suggested in the previous question.

****

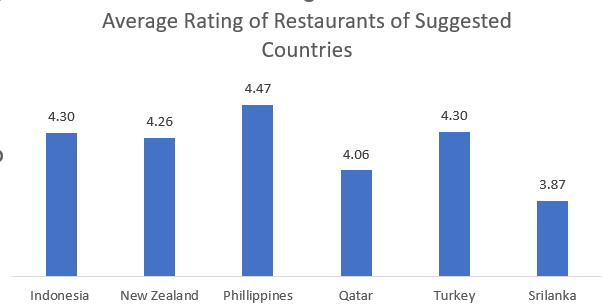
**Insights and suggestions**

* **Indonesia :** Good options to open new restaurants are **Bandung** and **Tangerang** for lower competition; consider select locations in **Jakarta** for scaling.
* **New Zealand:** Market is balanced but the count of restaurants is still low. The new restaurants can be opened in both **Auckland** and **Wellington City.**
* **Philippines:** Focus on cities with fewer restaurants but high ratings — **Quezon City, Tagaytay, Taguig, Makati, Pasig,Mandaluyong,Makati** we can also recommend other cities **San Juan** and **Pasay City.**
* **Qatar : Doha** has low count of restaurants and the avg rating is good. Explore areas outside Doha for potential expansion.
* **Sri Lanka: Colombo** has low count of restaurants and rating is some what good. Can open new restaurants.
* **Turkey :** Prioritize **Istanbul** (slightly less saturated) with high ratings; **Ankara** still viable**.**

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

Created a new table consisting of suggested countries in a column and the average rating of restaurants in another column

****



**Insights:**

* **Average Rating Range**: The average ratings in the data range from 3.87 to 4.47 on a scale of 5.0.
* **Philippines** leads with the highest average rating of **4.47 and Turkey, Indonesia, New Zealand has almost same average rating.**
* **Qatar has somewhat low but Sri Lanka has lowest compared to all.**

**Suggestions:**

* **Almost all the countries have good average rating implicating good customer experience but the restaurants in Sri Lanka have low keeping this in mind while opening new restaurant in that country we can overcome the drawbacks of current restaurants.**

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

Created two pivot tables one consisting of Country in Rows and Price range has column and Average\_cost\_for\_two (in Indian rupees) in Values(summarized as average) and other pivot table same as the first one but price range is not added in the column

****

****

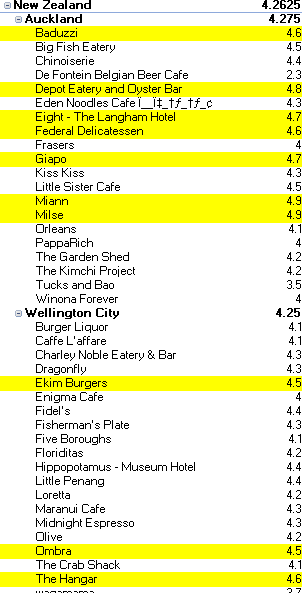
**Insights and suggestions :**

* As we can observe the food expenditure is high in countries New Zealand, Philippines, Qatar compared to Indonesia, Sri Lanka and Turkey.
* This variation suggests that pricing strategies should be tailored to local economic conditions and consumer expectations.
* The developed countries prices may need to reflect the higher cost of living and dining out, while in countries like **Sri Lanka**, **Turkey** more affordable pricing could help attract and satisfy customers.

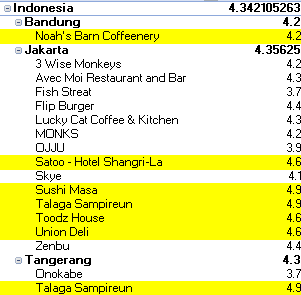
1. **Come up with the names of restaurants from the recommended states that are our biggest competitors and also those that are rated in the lower brackets, i.e. 1-2 or 2-3.**

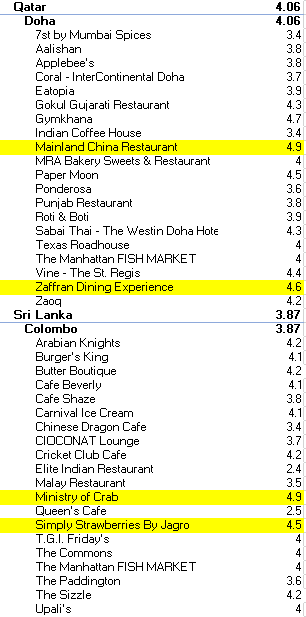
Created a pivot table and added country, city and restaurant name in Rows and added the rating in values (summarized as average).

The biggest competitors are the restaurants with highest ratings. All of the restaurants of my suggested countries have high ratings from 3.5-5. According to me the restaurants having rating >4.5 are the biggest competitors   
  
Highlighted the countries of all the restaurants that I believe is the biggest competitor that we have while opening new restaurant there







**** 

Coming the restaurants that belongs to lowest btracket that is 1-3 there are two restaurants one is



in AuckLand (New Zealand) and 

in Colombo (Sri Lanka)

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

Created a pivot table as country and Cuisines as Rows and Rating in Values (summ as average)

Identified the cuisines consisting of highest average for each country for better recommendation and highlighted them.

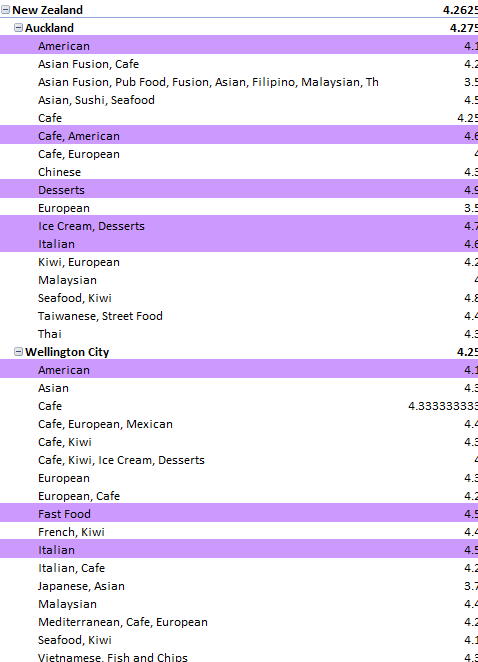
**Indonesia:**



Insight and suggestions:-

* As we can observe the cuisines like Indonesian, Japanese, Western and the cuisines that available in cafes like tea, coffee, desserts and all the things are highly rated.
* The best cuisines for new restaurant in Indonesia are **Indonesian, Japanese, if possible Western** and the basic cuisines that available in Cafes.

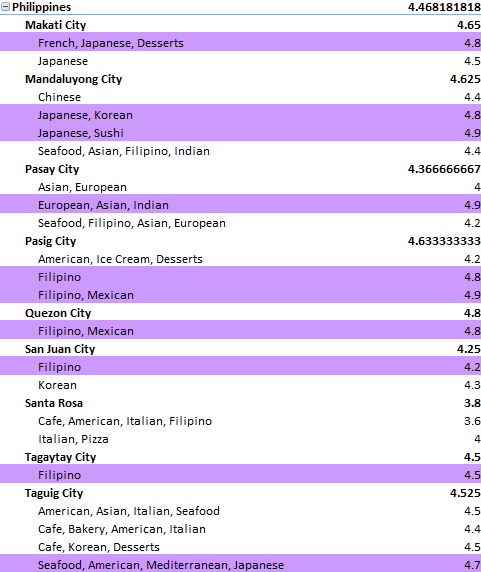
**New Zealand :-**

****

**Insights and suggestions**

* The restaurants consisting of American.Italian,Desserts, Ice creams and Café items.
* So, the cuisines suggestions for new restaurants are American,Italian,and desserts.

Philippines ;-



Insights and Suggestions;-

* As observing the cuisines and rating the suggestion for new restaurants will be Filipino, Japanese, French and seafood

**Qatar :**-



Insights and suggestions:-

* In Qatar there is no locally popularly cuisines the high rated cuisines are all international.
* To open a new restaurant in Qatar it is better to check which foreign country people are high in population and high tourists and making decisions. But according to data the Chinese cuisine is highly recommended.

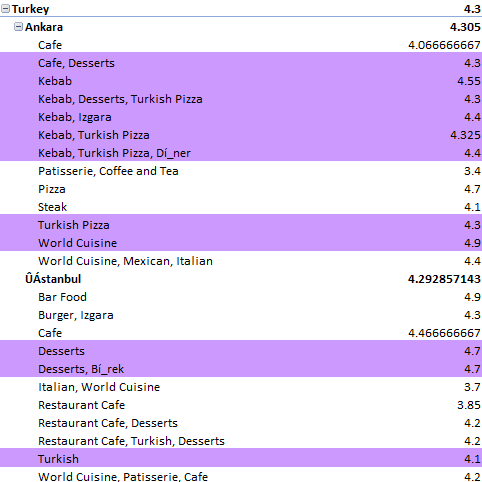
**Sri Lanka :-**

****

**Insights and Suggestions :-**

* Observing the above table we can see the cuisines like Seafood, American, Juices and Desserts are highly rated.
* The good recommendation for new restaurant will be Seafood, American and also Desserts.

**Turkey :-**

****

**Insights and Suggestions :-**

* The suggested cuisines for new restaurants will be Turkish, Turkish Pizza, and Desserts.

**Conclusion :-**

* The cuisine suggestion definitely depends upon the country and it is different for different countries.
* In every country the cuisine including desserts and basic cuisines like tea, coffee, ice creams are highly rated.
* Definitely the choice of cuisine effects the rating of the restaurants only popular cuisines in that region has good rating.

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

**Online delivery :-**

* The restaurants having online delivery option is around 25% of total restaurants.
* The average rating of restaurants having online delivery is high as compared to restaurants not having that option.

****

**Table Booking:-**

* The restaurants having table booking option is around 12% of total restaurants.
* The average rating of restaurants having table booking facility is high as compared to restaurants not having that option.

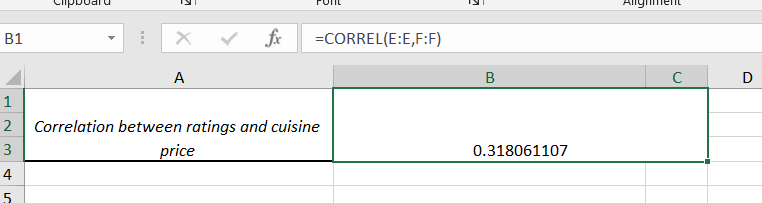
****

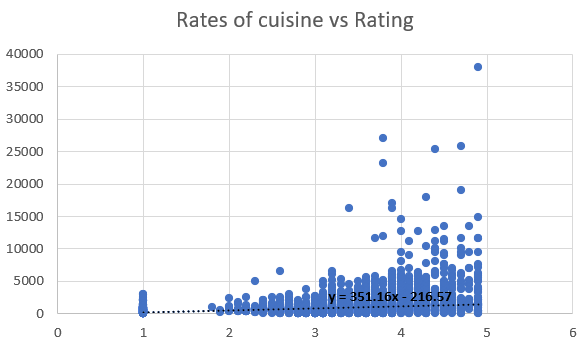
* The customer’s rating is really depended on the online delivery and table booking options.
* This strategic move could give us a competitive edge. However, before proceeding, conducting a survey in the respective countries to gauge consumer interest and willingness towards online delivery and table booking would be prudent**.**

1. **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?**

Copied and pasted Average\_Cost\_for\_two(in Indian rupees)(which will work as average cuisine cost) and Ratings Columns into new shee

And applied CORREL function for these two columns





* There is a positive correlation between rates of cuisines and ratings. It is low but not that close to zero.
* We can see that the restaurants that has lower prices has lower ratings and as the ratings increases the average cost also increases. This shows that there is a correlation between the two.
* In conclusion we can say that if the ratings of our restaurants are high then we can keep the rates of our cuisines high. But also, we should not keep the high rates of every cuisine available we must always have a low rates cuisine also.

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

Created a pivot table price range as rows and restaurantID as values(summarized as count)

****

**Insights and Suggestions**

* Almost 47% of restaurants belongs to lowest price range.
* Only about 20.86% of restaurants fall into the higher price brackets.
* In the lower price range, the number of restaurants is notably higher, this indicates that the more number of restaurants concentrate on keeping their price range low and attract more number of customers
* The market is **heavily skewed towards affordability**, likely indicating a **price-sensitive customer base** or a **larger lower-middle-income population** in the area.

1. **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]**

* If the objective and subjective questions are not given then I would have to almost same approach.
* My idea for suggesting countries/cities for opening the new restaurants would the country with low restaurant market density and I would also keep the mind the population of the city/country. Some cities won’t need a greater number of restaurants even though the restaurants count is very low.
* The average ratings also play crucial role while selecting place to open new restaurants.
* I would the most popular cuisine in that region by surveying or we can just find out by the data that is provided.
* And I would also keep in mind how much money the client can invest because the value of money is different in different countries. so I would have selected the place to open new restaurants keeping that in mind.
* Definitely I would have suggested having online delivery option is best and also table booking.
* And coming to the prices I would suggest keeping to low price range initially and after getting high ratings maybe price can be increased.