**Zomato Restaurant Analysis**

****

**By: Kumar Prakash**

**Objective Questions with Answer**

**Q1. The data consists of some inconsistent and missing values so ensure that the data used for further analysis is cleaned.**

* Steps taken for data cleaning**:**
  + Used =TRIM() and =IF(ISBLANK(), "Unknown", value) for City/Restaurant Name/Locality/Address.
  + All numeric columns checked for errors, dates parsed to years.

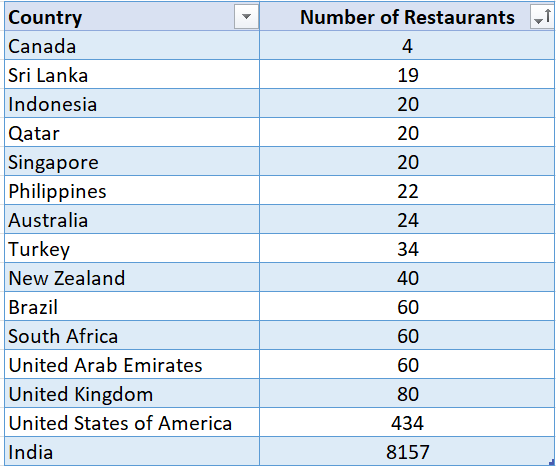
**Q2. Using the Lookup functions, fill up the countries in the original data using the country code.**

* Country filled via lookup: =VLOOKUP(Country Code, Lookup Table, 2, FALSE) in Country column.

For example: In D2 cell we used: = VLOOKUP(C2, 'Country Description'!$A$2:$B$16, 2, False) and by using fill handle we fill rest of the cells.

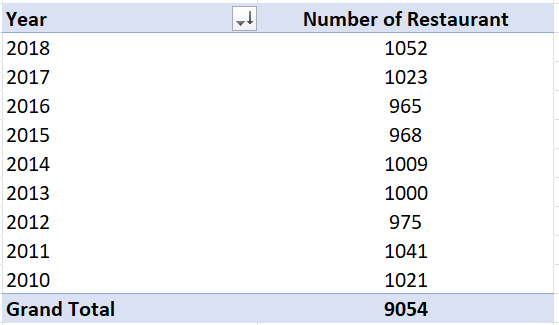
**Q3. The Create a table to represent the number of restaurants opened in each country.**

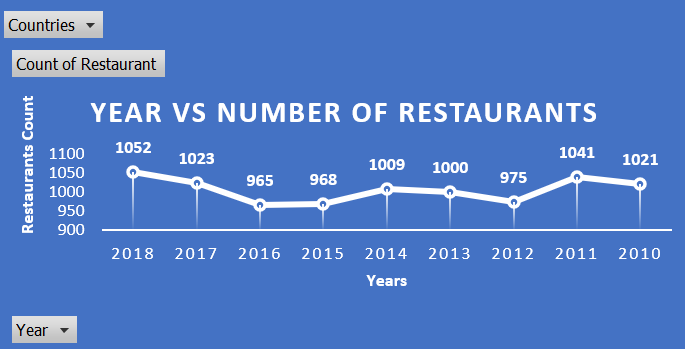
* Formula**:** I used pivot table then copy data and made convert into table to show the number of Restaurants in order.
* Now Country Vs Number of Restaurants in increasing order are shown below.

****

**Q4. Also, the management wants to look at the number of restaurants opened each year, so provide them with something here.**

* Formula: I used pivot table and short them based on year in decreasing order.
* Now Year Vs Number of Restaurants are below:

****

****

**Q5. What is the total number of restaurants in India in the price range of 4?**

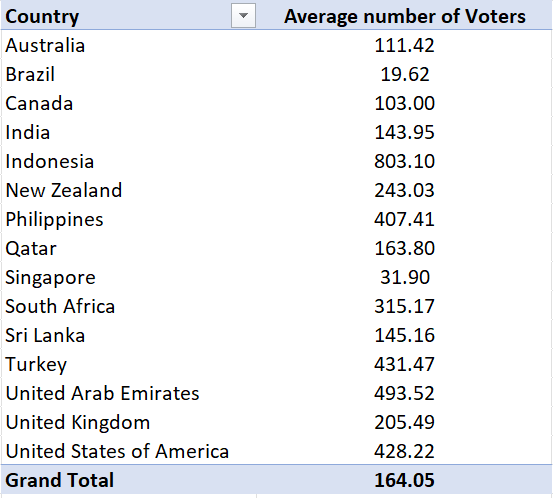
* Formula**:**= COUNTIFS(Table1[Country], "India", Table1[Price\_range], 4)

**(Note: Table1 refer to “Answer Data” Sheet.)**

* In our case value is: **372**

**Q6. According to the data, what is the average number of voters for the restaurants in each country?**

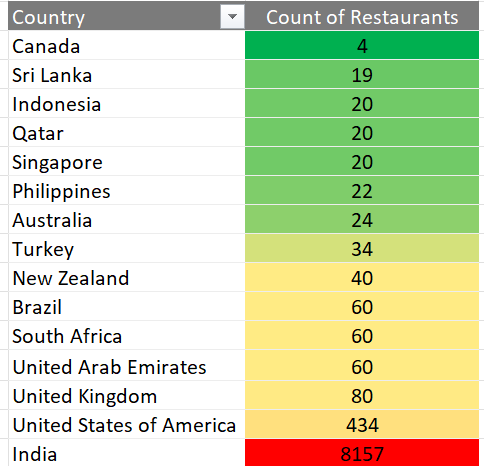
* Formula: For this I create Pivot table.
* Now Country Vs Average number of Voters are shown below.

****

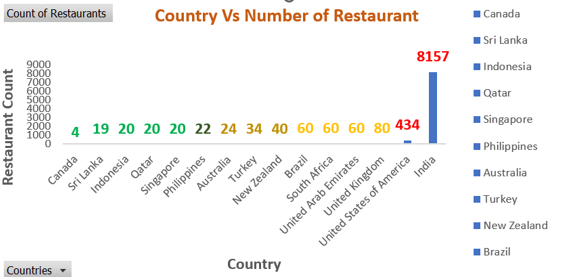
**Subjective Questions with Answer**

**Q1. 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?**

* Formula: I create a pivot table and then apply conditional formation (by selection Count).
* Note: Green (most priority) and Red (least priority).



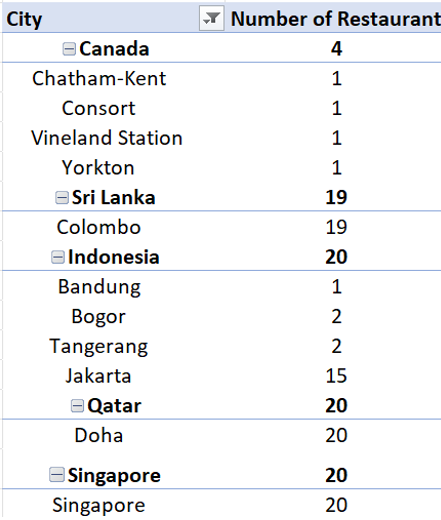
* Chart: Country Vs Number of Restaurant.

**

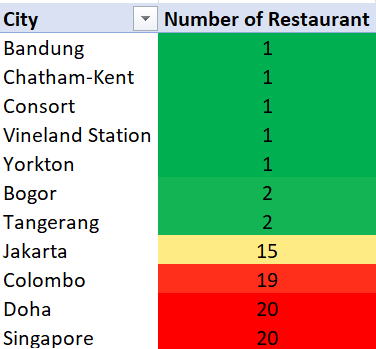
* Result: by analysing table and chart my choices are:
  1. Canada
  2. Sri Lanka
  3. Indonesia
  4. Qatar
  5. Singapore

**Q2. Come up with the names of States and cities in the suggested countries suitable for opening restaurants.**

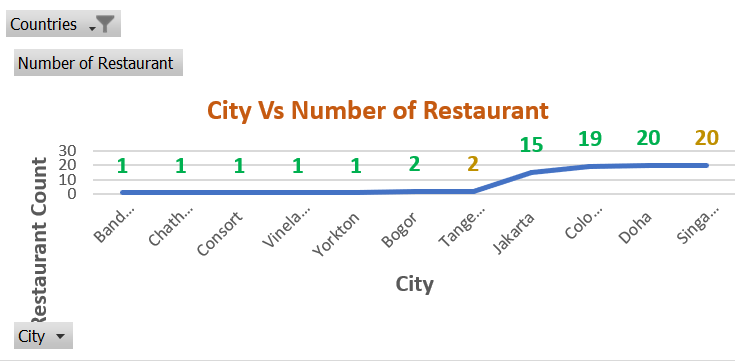
* Formula: I create a pivot table to show Country with City Vs Number of Restaurant.
* Table1: Country with City Vs Number of Restaurant is shown below.



* Table2: I used pivot table to show City Vs Number of Restaurant is shown below and country in filter section.
* At the end we apply Conditional Formatting by selecting Number of Restaurant Column.
* Set Green (Highest priority), Yellow ( medium priority) and Red (Lowest priority).

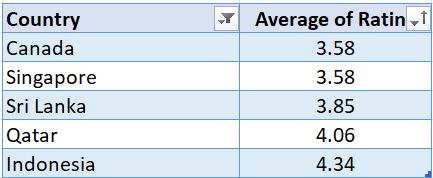


* Chart: City Vs Number of Restaurant.

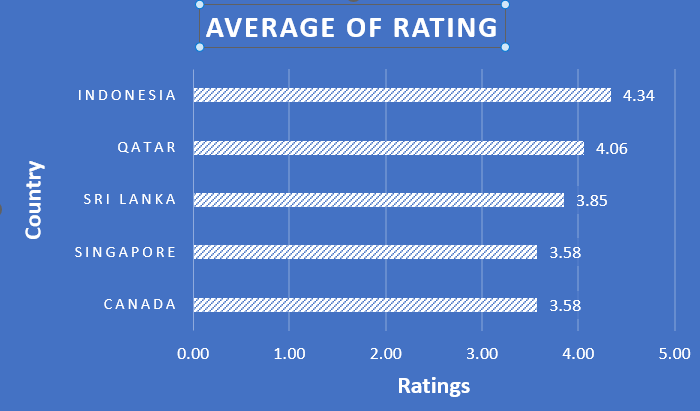


**Q3. According to the countries you suggested, what is the current quality regarding ratings for restaurants that are open there?**

* Formula: I create a pivot table and took rating in value and round it two decimal points.
* Table1: Country Vs Average of Rating is shown below.

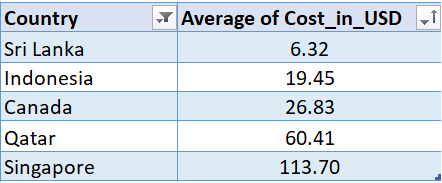


* Chart: Country Vs Average Rating

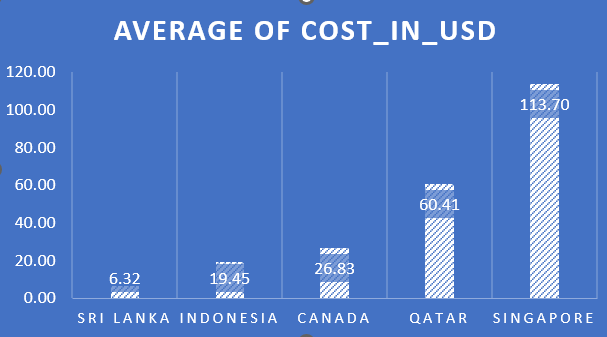


**Q4. Also, what is the current expenditure on food in the suggested countries, so we can keep our financial expenditure in control?**

* Formula: I create a pivot table and took rating in value and round it two decimal points.
* Table: Country Vs Average of Cost in USD.

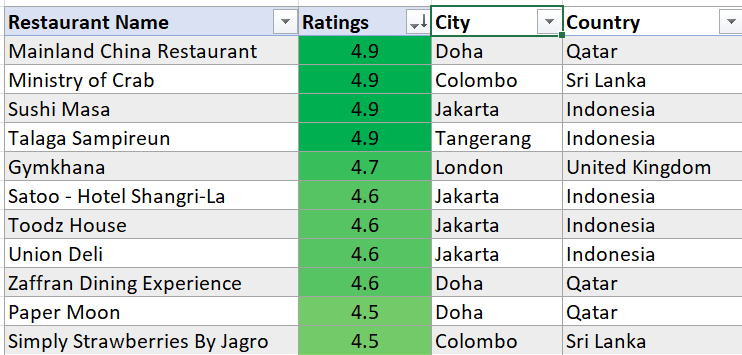


* Chart: Country Vs Average of Cost in USD.

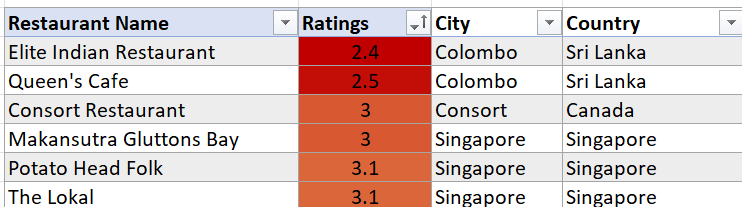


**Q5. Come up with the names of restaurants from the recommended states that are our biggest competitors and also those that are rated in the lower brackets, i.e. 1-2 or 2-3.**

* Formula: I create a pivot table and then apply conditional formation (by selection Count).
* Biggest Competitors (lies in upper bracket):

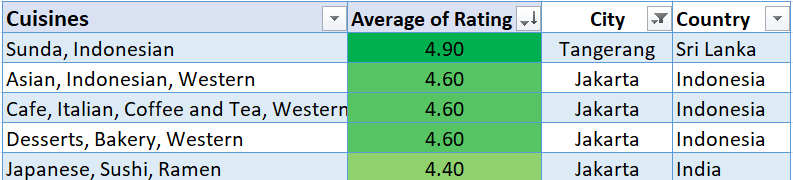
****

* Lies in lower bracket (i.e. 1-2 or 2-3):

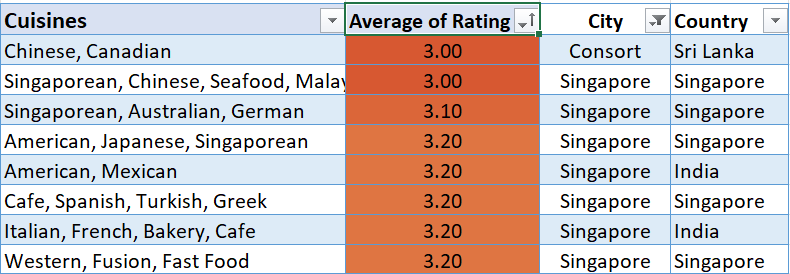
****

**Q6. Which cuisines should we focus on in the newer restaurants to get better feedback? Does the choice of cuisines affect the restaurant ratings?**

* Formula: I create a pivot table and then copy tables data and made a table and apply conditional formatting (by selecting Average Rating).
* Then add City and Country and apply XLOOKUP function (like: =VLOOKUP(G96,'Answer Data'!$B:$E,4,FALSE) ) to fill data inside these columns.
* Table1: below table show, Cuisines with High Average Rating (in decreasing order).

****

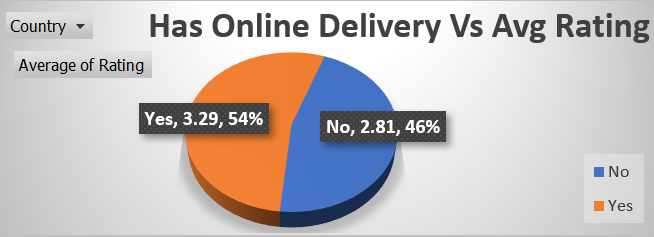
* Table2: below table show, Cuisines with Low Average Rating (in increasing order).

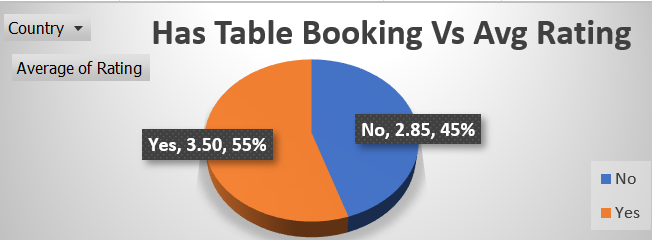
****

* Note**: Yes**, Choice of Cuisine is very important step for targeting customer and market

**Q7. According to our current data, should we go for online delivery and table booking? Does that affect the customer’s ratings?**

* Formula: I create a pivot table and then apply conditional formation (by selection Count).
* Below Two 3D pi chart given (i.e. for 1. Has online Delivery vs Average Rating and 2. Has Table Booking vs Average Rating.
* From both the Charts, Rating of YES is greater that NO also percentage.

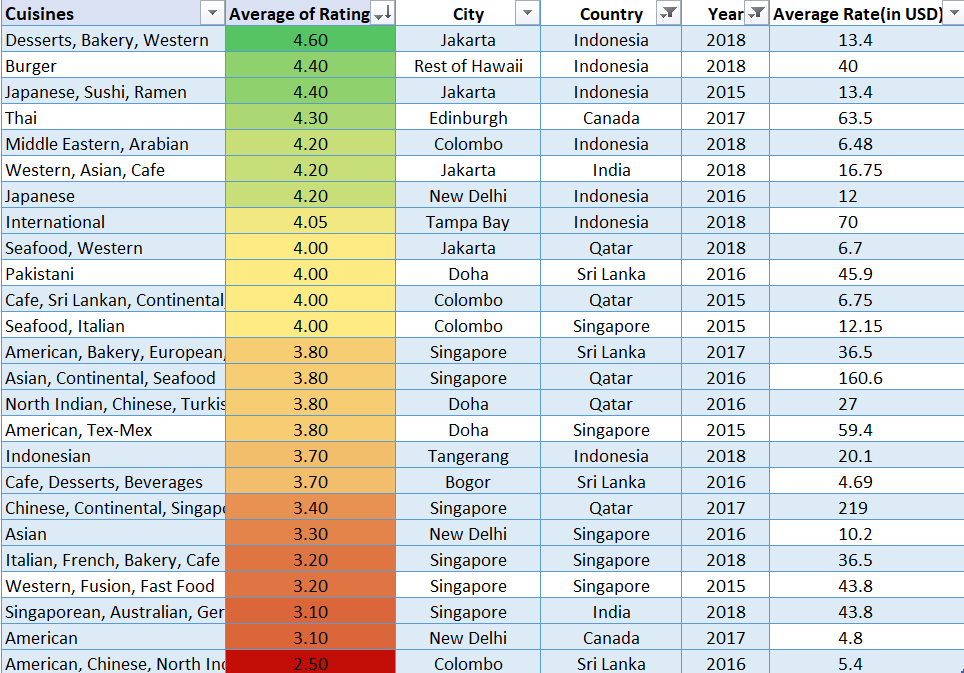
****

****

* From both the Charts, Rating and percentage of YES is greater that NO. It means we should go with Online delivery and Table booking.

**Q8. 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?**

* For this we need to calculate Correlation between Price and Rating.
* Formula: = CORREL([Price], [Rating])
* By using above formula, we got Correlation value = 0.06 (i.e. towards zero). It means both are not correlated.
* Also, I crosscheck with finding table between them with suggested country and City for last 4years. Got this.



* From above table we got same conclusion.
* So, Rate of Cuisine not affect the feedback.

**Q9. What is the distribution of a number of restaurants of different price ranges in all the countries?**

* For this we need to add helper column in answer data sheet, for Price Range (e.g. Low, Medium, High).
* Since price range vary between $0 to $365 dramatically. So, I assume less than $50 as Low, $50 - $100 as Medium and above $100 as High.
* Formula (In Z2 I used this): =IFS(T2<50, "Low", T2 < 100, "Medium", T2 >= 100, "High")



****