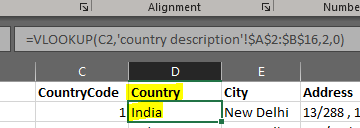
**Objectives**

**1.** Data cleaning

* I observed the Datekey\_Opening column had the date format: “23\_09\_2013” which is not an ideal format of the date. I used **text to columns** feature **=>** used Date as datatype **=>** Finish **=>** Changed the date format to Short date (dd-mm-yyyy)
* For Cuisines column **=>** Some of the cuisine’s names were blank **=>** I used Restaurant name and filled the blank cuisines with an appropriate value by googling it out.
* For Longitude and Latitude columns: Approx. 500 columns were blank **=>** I selected city column **=>** In data tab **=>** Changed data type to Geography **=>** Then in the blank cells there's a formula which I typed **=>** "*=cityname.Longitude*" **=>** Enter **=>** I got the coordinates.

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

* In the raw data, I added a column “Country” where I have applied VLOOKUP function to get the country name using country description sheet.



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

* Pivot table is the best representation to show number of restaurants opened alongside each country I have sorted the number of restaurants highest to lowest order to visualize in a better way (Please refer Solution of Objective problems sheet for the same).

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

* A column chart is the best fit for this problem to visualize. Where I have created a chart of year vs number of restaurants. To visualize quickly, I have highlighted highest and lowest by year where the restaurants are opened (Please refer Solution of Objective problems sheet for the same).

**5.** What is the total number of restaurants in India which are in the price range 4?

* A small pivot table can be drawn to see the number of restaurants in India which are in the price range 4 where in pivot we can take country in rows, price\_range in column and count of restaurant\_id in Values (Please refer Solution of Objective problems sheet for the same).

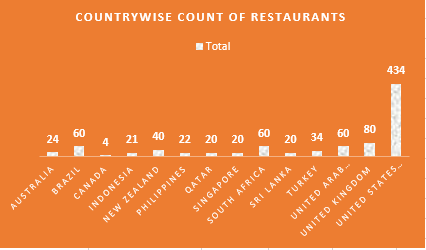
**6.** What is the average number of voters for the restaurants in each country according to the data?

* A column chart can be used here to visualize countries in x-axis vs average of votes taken in the Values (Please refer Solution of Objective problems sheet for the same).

**Subjective**

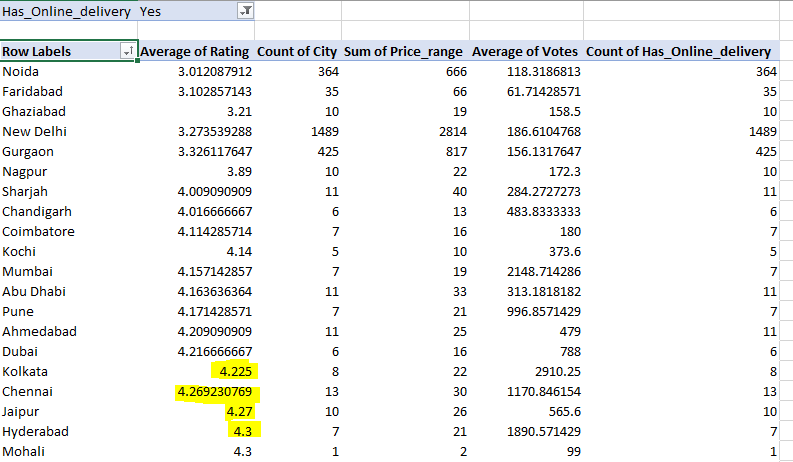
**Q.** Suggest few countries where the team can open newer restaurants with lesser competition.

* Considering the count of restaurants opened in each country as showed in the below chart. It is clearly showing Indonesia, Canada, Qatar, Singapore, Sri Lanka, and Philippines having least number of restaurants opened in their country therefore the lesser it makes competition to open a restaurant in these countries. Among which Canada has the lowest and hence, it could be considered the best choice to open few restaurants over there.



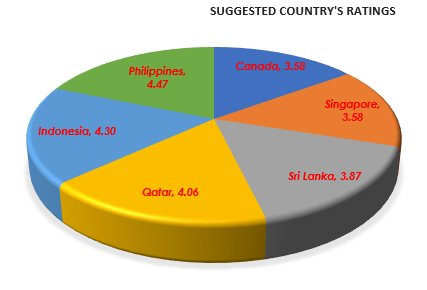
**Q.** Come up with the names of States and cities in the suggested countries suitable for opening restaurants.

* We can use a technique to get the cities from the data based on multiple factors:
  + The average Rating for each city to assess overall restaurant performance.
  + Restaurant count: Used COUNT function to find the total number of restaurants in each city.
  + Popularity Factors: Considering additional metrics like average votes, average price\_range, or percentage of restaurants offering Has\_Online\_delivery to gauge demand and market preferences.



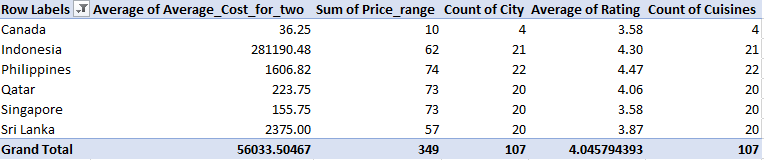
**Q.** According to the countries you suggested, what is current quality in terms of ratings for restaurants that are opened there?

* According to the six countries I have suggested above, we can clearly see in the below pie chart Philippines, Indonesia and Qatar have the ratings 4 or above which is Very good, out of 5 Please refer Solution of Subjective problems sheet for the same).



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

* We can narrow down the search to include each country, sum up the average of Average\_Cost\_for\_two for all restaurants within that country. This will give us the total expenditure on food in each suggested country. And I have added other factors like Price\_range, count of Cities and average of Rating as well to back my data so that we can analyse more in depth.



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

* There are multiple factors which can be included to set the criterion for restaurants name like:
* Similar offerings: I have curated the data based on similar cuisines and services (e.g., online delivery, table booking and Is\_delivering\_now)
* Price Range & Average Cost: price ranges or average costs for two people are in a similar bracket which for the below restaurants is 2 - 3.
* Rating Comparison: I have compared their ratings within the specified lower brackets which is 1-2.
* Here are the restaurants names:
  + Cake 24x7
  + Habibi Express
  + Yo! China
  + Moti Mahal Delux Tandoori Trail
  + Tmos Cafe Corner

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

* Frequency of Cuisines: I have created a pivot table to count the frequency of each cuisine offered across restaurants.
* Ratings by Cuisine: Calculated average ratings for each cuisine, this will show the average rating for each cuisine type. Below table belongs to the newer restaurants name.
* We can use correlation function (CORREL) to find if there's any correlation between specific cuisines and higher ratings. We can compare the frequency of a cuisine with the average ratings it receives.
* identifying cuisines that have higher average ratings and a significant presence in the market.

|  |  |  |
| --- | --- | --- |
| **Restaurant Names** | **Count of Cuisines** | **Average of Rating** |
| Talaga Sampireun | 3 | 4.9 |
| Silantro Fil-Mex | 2 | 4.85 |
| Fish Streat | 2 | 3.7 |
| The Manhattan FISH MARKET | 2 | 4 |
| Restaurant Andre | 1 | 3.8 |
| Malay Restaurant | 1 | 3.5 |
| Al'frank Cookies | 1 | 4.2 |
| Arabian Knights | 1 | 4.2 |
| OJJU | 1 | 3.9 |

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

* To analyze whether offering online delivery and table booking affects customer ratings, we can follow these steps in Excel:
  1. We can use AVERAGEIFS function to calculate the average ratings for restaurants based on whether they offer online delivery and table booking.
  2. Likewise, to compare the average ratings of restaurants that offer online delivery versus those that do not. Repeat the same for restaurants offering table booking.
  3. For calculating average ratings:

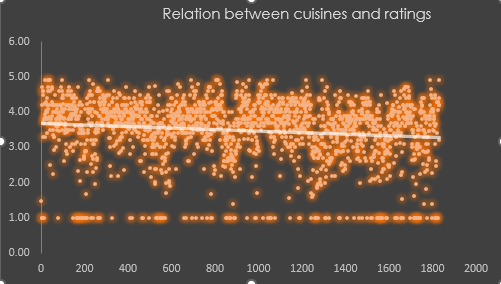
=AVERAGEIFS (Rating, Has\_Online\_delivery, "Yes")

=AVERAGEIFS (Rating, Has\_Online\_delivery, "No")

* We can use correlation function (CORREL) to analyze if there is any correlation between online delivery/table booking and higher ratings.
* Conclusion:
  + Based on the above analysis, we can conclude whether offering online delivery and table booking have an impact on customer ratings positively. If there is a significant correlation like higher average ratings for restaurants offering these services, then we can consider adding them into new restaurants.

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

* Ratings and Rate of Cuisines Analysis:
* I will first calculate the average ratings for each cuisine type (like the previous analysis).
* Secondly, I have drawn a scatter plot of the average cost for a cuisine vs average rating for that cuisine.
* I have used trendline feature to determine if there's a correlation between the rate of cuisines (average cost) and ratings.
* The trendline in my scatter plot is straight but slightly tilted downwards, it indicates a slight negative correlation between the variables being analyzed.
  + It suggests that –
    - The rate of cuisines will increase, the ratings tend to decrease slightly.
    - This negative correlation might be helpful to find a balance between offering quality cuisines and maintaining reasonable pricing to receive better feedback and ratings.
    - Please find attached below the scatter plot and trendline (also refer the same in the Excel sheet as well.)



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

* I have used stacked bar chart to show the distribution of number of restaurants of different prices along with the countries.
* With the help of pivot table countries vs number of restaurants along with Price\_range, I created a PivotChart.
  + For Price\_range – I created a calculated field in the pivot table that categorizes the price ranges into different groups like Price\_range: 1 will be categorized as Low, simultaneously for 2, 3 & 4 Price\_range will be categorized as Medium, High & Extreme.
* Additionally, I have used log scale for vertical axis to allow for better readability as there are small and large values.