**Objective Questions**:

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

The total number of tables present in the data are 2 Tables (Raw data and country description**.**

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

Every single column is counted as attribute. So, there are **22 Attributes**.

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

Categorical columns are those that represent discrete values or categories. These columns typically contain data that can be divided into specific groups or categories, and they often include text values or codes that represent different categories

There are 15 categorical columns:

**SHEET - RAW DATA**

1. **RestaurantName**: Represents the name of each restaurant, which is a discrete category.
2. **Country code:** Each country code represents a different category
3. **City**: Represents the city where the restaurant is located, which is a discrete category.
4. **Address**: Represents the address of each restaurant, which is a discrete category.
5. **Locality**: Represents the locality within a city, which is a discrete category.
6. **LocalityVerbose**: A detailed description of the locality, which is a discrete category.
7. **Cuisines**: Represents the types of cuisines offered by the restaurant, which are discrete categories.
8. **Currency**: Represents the currency used, which is a discrete category.
9. **Has\_Table\_booking**: Indicates whether the restaurant offers table booking (yes/no), which is a discrete category.
10. **Has\_Online\_delivery**: Indicates whether the restaurant offers online delivery (yes/no), which is a discrete category.
11. **Is\_delivering\_now**: Indicates whether the restaurant is currently delivering (yes/no), which is a discrete category.
12. **Switch\_to\_order\_menu**: Indicates whether there is an option to switch to an order menu (yes/no), which is a discrete category.
13. **Datekey\_Opening**: Represents specific dates in a categorical manner.

**SHEET- COUNTRY DESCRIPTION:**

1. **Country code:** Each country code represents a different category
2. **Country Name**: Represents the name of each country, which is a discrete category
3. **The data consists of some inconsistent and missing values so ensure that the data used for further analysis is cleaned**

**MISSING VALUES**

1. **Cuisines:**

There were 9 missing values in Cuisines.

How I have figured it out:

* Find and select
* Go to special
* Blanks

All blank cells got highlighted with green colour. Applied ‘filter by colour’ on ‘Cuisines’ to check the blank values.

Reason: I observed all the 9 missing values belong to country code – 216. So, Created two pivot tables.

Table 1:

Rows- Country code

Values - distinct count of restaurant id & average rating

Output: Restaurant count- 434 (2nd highest) &

Avg. Rating-4.0

Table 2:

Rows- Cuisines

Columns - Country code

Values- Avg. Rating

Output – Blank restaurant – 3.9

1. Average of all restaurant is approximately equals to average of restaurants with blank cuisines. So, removing these restaurants with blank cuisines will not affect the overall rating of country.
2. 216 country code has large no. of restaurant, so removing 9 restaurants from this large no. will not affect the whole dataset

On the basis of above observation, **deleted the 9 rows**.

**INCONSISTENT VALUES**

1. **Restaurant Id**

On observation, it was clear that there is inconsistency in Restaurant Id

How I confirmed:

* I checked the maximum length of restaurant\_id. It comes out be 8.

**FUNCTION: =MAX (LEN (A2:A9552))**

* Then I applied conditional formatting to check how many restaurants are there which has length less than 8
* In conditional formatting- I applied the rule to highlight the restaurant ids with length <8.

**RULE: =LEN(A1)<8**

**ACTION:** To make all the restaurant id same in length, **I Formatted the cells in a custom format of “00000000”**

1. **Leading and trailing spaces**

How I figured it out:

* I selected the whole dataset and used conditional formatting with a new rule.

**RULE: =LEN(A1)<>TRIM(LEN(A1))**

The output: Adress, Locality and Locality Verbose had the extra spaces in values.

**ACTION:** So, **used Trim Function to make it consistent**.

**PRE-PROCESSING- ADDED COLUMN**

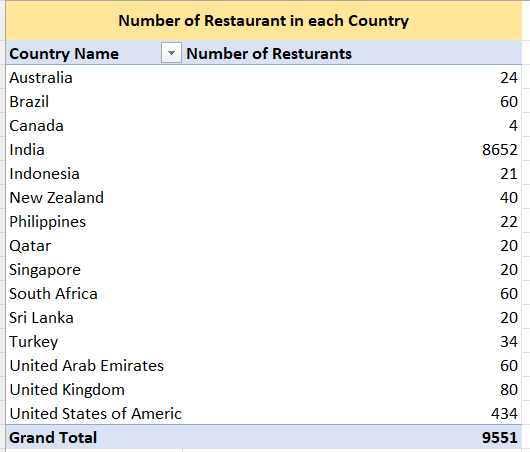
* **Country** - Added the country for respective country code using V-lookup Function. This column will make our analysis better as we don’t need to look for country code. Now we know the Country.
* **Average\_Cost\_in\_INR** - Changed all the ‘average price for two’ currency into INR. This will help us to know the average price in all country in the form of one currency.
* **Opening\_Month & Opening\_Year**- Extracted the opening year, month from ‘Date key opening’ using TEXT function. It will help us to analyse the data based on the year and month.
* **State-** Identified the state name from the cities name by using ChatGPT. This will help us in better understanding of density of restaurant in a particular country
* **RANGE\_OF Average\_Cost\_in\_INR-** Categorised ‘average cost in INR’ into price range. This will give us price bracket to compare

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

The XLOOKUP function, which can search both vertically and horizontally, can be used to accomplish this.

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

I used the pivot table for this. where "Country Name" is entered in the rows and "Restaurant Id" is entered in the value field. The field setting is then changed to the number of restaurants for each nation by entering the "Restaurant Id" of the pivot table.



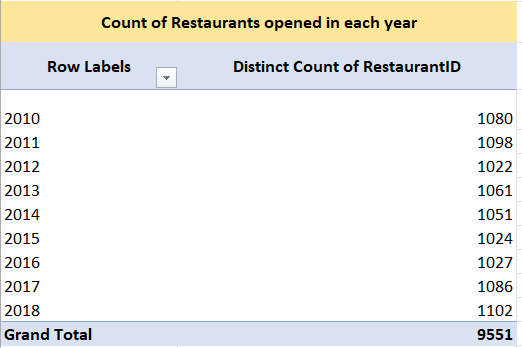
1. **Also, the management wants to look at the number of restaurants opened each year, so provide them with something here.**

Created PIVOT TABLE

ROWS: Opening year

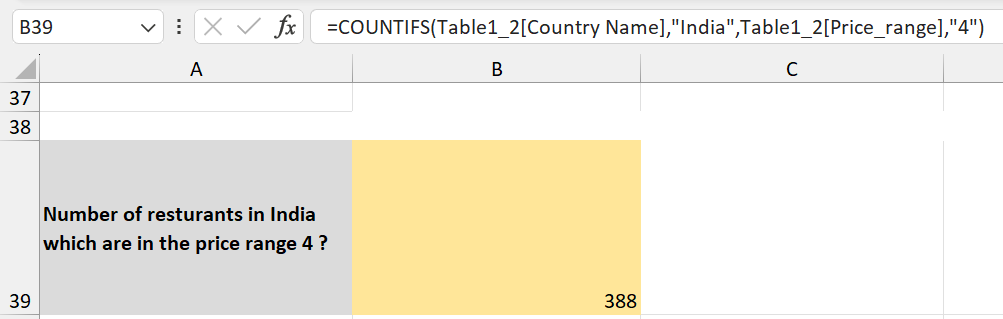
VALUES: Distinct Count of Restaurant

OBSERVATION:

 YEAR 2018 records the highest no. of restaurant opened.

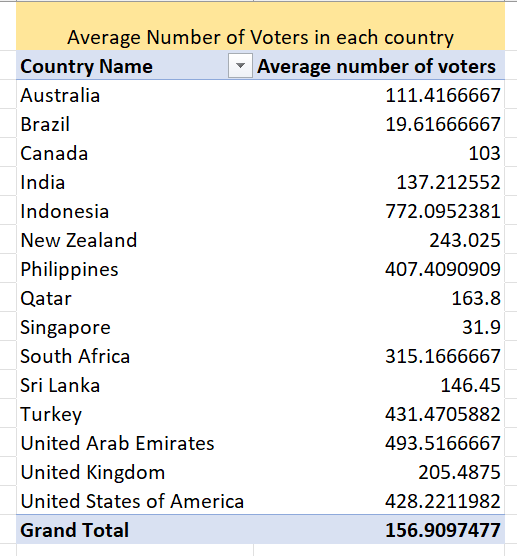
1. **What is the total number of restaurants in India in the price range of 4?**

We can accomplish this by using the conditional aggregation function COUNTIFS, since we need to search for two criteria: first, that the nation is "India," and second, that the "price range" equals four. To achieve the desired result, as indicated below, I entered the country name column range in criteria\_range\_1 (the first parameter of the COUNTIFS function), "India" in criteria1 (the second parameter of the COUNTIFS function), Price Range column range in criteria\_range\_2 (the third parameter of the COUNTIFS function), and "4" in criteria2 (the fourth parameter of the COUNTIFS function).



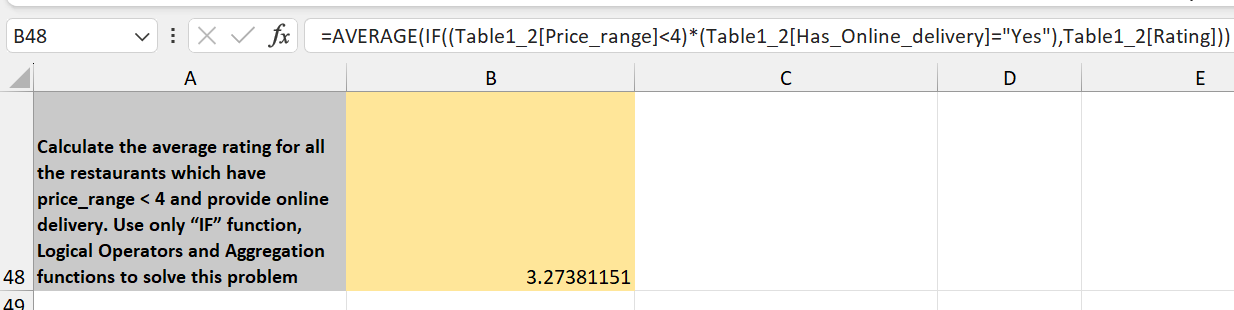
1. **What is the average number of voters for the restaurants in each country according to the data?**

For this I have used the pivot table. In which “Country Name” column is putted in the rows and “Voters” is putted in value field and changed the field setting to average of “Restaurant Id” of pivot table, so that we can get average number of voters for the restaurants in each country.



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.]**

I've done this by nesting the Average and If conditions rather than utilizing the conditional aggregation function "AVERAGEIF." First, the IF condition will search for cells with Price\_range less than 4 and Has\_Online\_delivery set to "Yes." The average of the Ratings column for the values obtained from the IF condition will then be determined by the AVERAGE function.

****

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.**

In the subjective questions 1 and 2, we can observe the name and method utilized to weed out the names of the recommended nations and cities, respectively. Next, I highlighted the names of restaurants within that proposed city using the city's name (found in column F). The conditional formatting formula is shown below, and the excel sheet titled "Restaurants in suggested Cities" displays the names of each restaurant.

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

**Step 1:** This is done first by fetching currency symbols from the “Currency” column using the MID, SEARCH function together, and putted it in “Currency (Imputed)” column.

**Step 2:** Then concatenating the Currency (Imputed)” column with “Average\_cost\_for\_two” and naming the new customized column as “Price (Imputed)”.

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?**

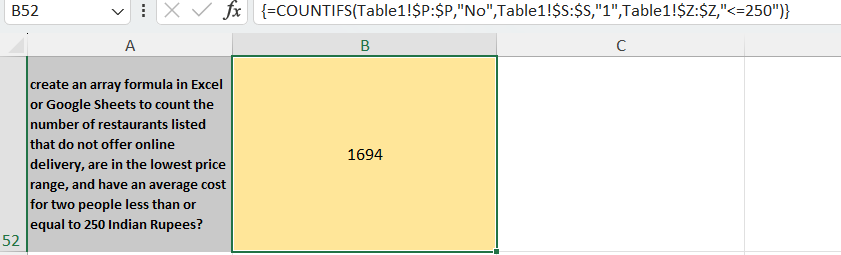
I used COUNTIFS Function to find the count of restaurants that do not offer online delivery, are in the lowest price range and have an average cost for two people less than or equal to 250 Indian Rupees?

{=COUNTIFS(Table1!$P:$P,"No",Table1!$S:$S,"1",Table1!$Z:$Z,"<=250")}

HOW TO CREATE ARRAY FUNCTION:

Firstly, write the required function according to the question,

and then press Ctrl+Shift+Enter

****

**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?**

**Suggested Countries: -**

|  |  |
| --- | --- |
| Australia | 1 |
| Canada | 2 |
| Indonesia | 3 |
| New Zealand | 4 |
| Philippines | 5 |
| Qatar | 6 |
| South Africa | 7 |
| Sri Lanka | 8 |
| Turkey | 9 |
| United Arab Emirates | 10 |

**Approach: -** We can analyse it by keeping following criteria in mind

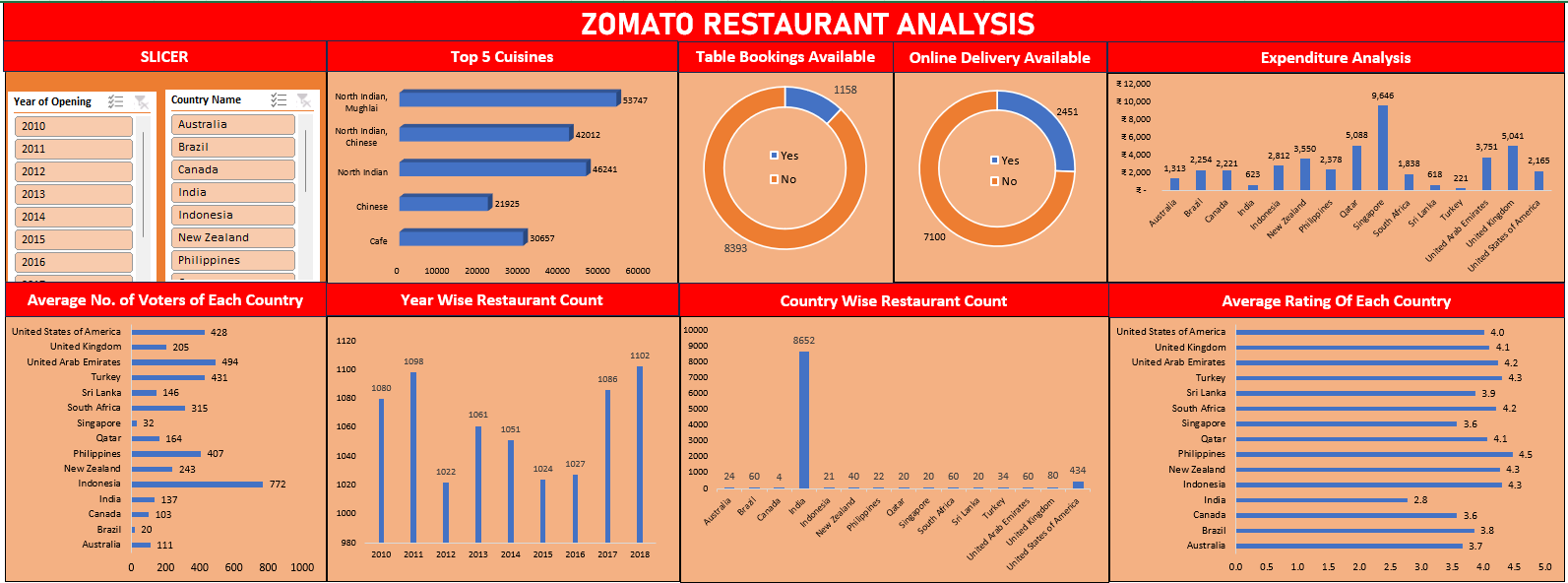
1. Country with less restaurants than 200.

2. Restaurants in that country which has rating>=3.

3. In that nation, restaurants with an average cost for two that is less than Rs 1600 rather than zero (after converting each average cost for two to any one currency). Since the typical cost of food for two people in Indian Rupees ranges from Rs 800-Rs 2000, a middle value of Rs 1600 is chosen for the analysis.

4. Votes (number of persons visited) >150 in that restaurant

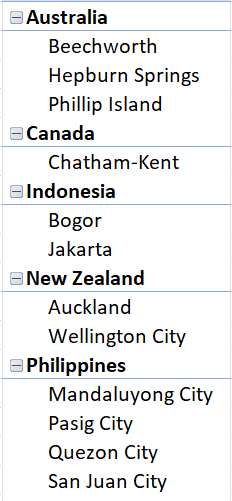
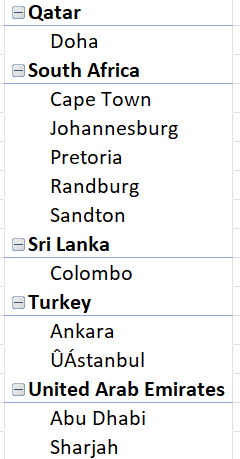
**Visualization: -** A dashboard can be used to display the analyses, and conditional formatting can be used to retrieve nations with less competition.



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

**Approach & Analysis**:

I looked at the number of eateries in each state and its corresponding city to generate the names of the cities and states. Therefore, I recommended that states and cities with fewer restaurants had a greater chance of growing.

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

**Approach: -** We can obtain average ratings for the restaurants in each nation, which are displayed below, by using a pivot table where the "Country Name" column is entered in the rows and "Ratings" is entered in the value field. The field setting is then changed to the average of the "Ratings" of the pivot table. The average rating is now shown using the star symbols in Excel's conditional formatting capability.

**OBSERVATIONS:**

**1**.**Highest and Lowest Average Ratings**

The Philippines has the highest average restaurant rating at 4.5

Canada has the lowest average restaurant rating at 3.6.

**2.General Quality Observation**

With remarkably high average scores, nations like Indonesia and the Philippines appear to have typically high-quality dining establishments.

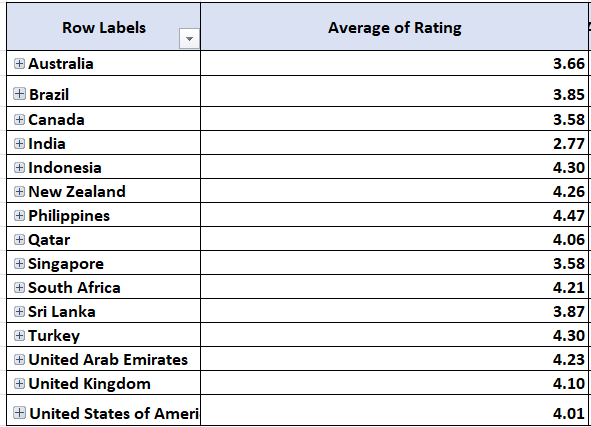
**ANALYSIS:**

Compared to the other countries on the list, Canada and Australia have somewhat lower average ratings, suggesting that the quality of the restaurants there may not be as good.

**Quality Tier Grouping**

The countries can be grouped into quality tiers based on average ratings:

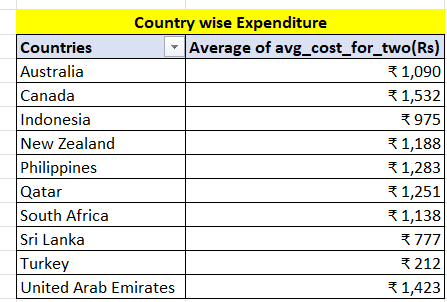
* + High Quality: Philippines (4.5), Indonesia (4.3)
  + Moderate to High Quality: South Africa (4.2)
  + Moderate Quality: Australia (3.7), Canada (3.6)



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

**Approach: -** We may determine the current food expenditure in the recommended nations by utilizing the "Average\_cost\_for\_two" column to calculate the average expenditure on food.

I used the pivot table for this. which sets the field setting to the average of the pivot table's "avg\_cost\_for\_two (Rs.)" and inserts the "Country Name" column in the rows and the "avg\_cost\_for\_two (Rs.)" column in the value field.



**OBSERVATION:**

1. **Highest and Lowest Expenditures**
   * Canada has the highest average food cost.
   * Turkey has the lowest average food cost.
2. **General Cost Observation**
   * Canada has notably high average food costs, suggesting that dining out in this country is relatively expensive.
   * Turkey has relatively lower average food costs, indicating more affordable dining options.
3. **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.**

**Identifying Major Rivals**

Understanding market leaders, their tactics, advantages, and areas of strength is facilitated by identifying top rivals.

**Finding Restaurants with Low Ratings**   
  
• Recognizing common problems and places for improvement is made easier by knowing about poorly rated eateries.   
  
• Offers information on what customers don't like, which is useful for preventing similar errors.

• Assists with market positioning through the identification of exploitable gaps and unmet client needs.

**Biggest Competitors criteria: -**

* + - 1. Restaurant rating>=4
      2. avg\_cost\_for\_two< Rs 1600
      3. Number of Voters > 150

**VISUALIZATION -** PIVOT TABLE

**Restaurant names with their Country and States (Biggest Competitors)-**

****

**Restaurant names with their Country and States those that are rated in the lower brackets, i.e. 1-2 or 2-3:**

****

**Insights and Recommendations**

* 1. **Competitive Landscape:**

Due to their high ratings and significant number of votes, Bridge Road Brewers in Beechworth and Fish Streat and Flip Burger in Jakarta are major rivals. For best practices, these restaurants' strategies should be carefully examined.   
 **B. Opportunities for Development:**Low-rated eateries, such as Vadakkan Pepper in Sharjah and Sangeetha Vegetarian Restaurant in Abu Dhabi, draw attention to potential shortcomings in consumer interaction, quality, or service. Avoiding similar problems can be facilitated by learning from their mistakes.

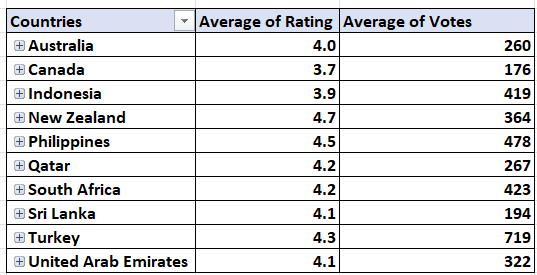
1. **Which cuisines should we focus on in the newer restaurants to get better feedback? Does the choice of cuisines affect the restaurant ratings?**

Indeed, the food has an impact on restaurant ratings. In order to improve customer feedback and ratings, I made the decision to identify which cuisines should be given priority in new restaurants.

New restaurants can make sure they satisfy patron expectations and preferences by concentrating on well-liked and highly-rated cuisines, which will increase customer satisfaction and improve ratings.

Relationship Between Ratings and Cuisines:   
aids in determining whether the kind of food has a direct impact on patron opinions and comments.

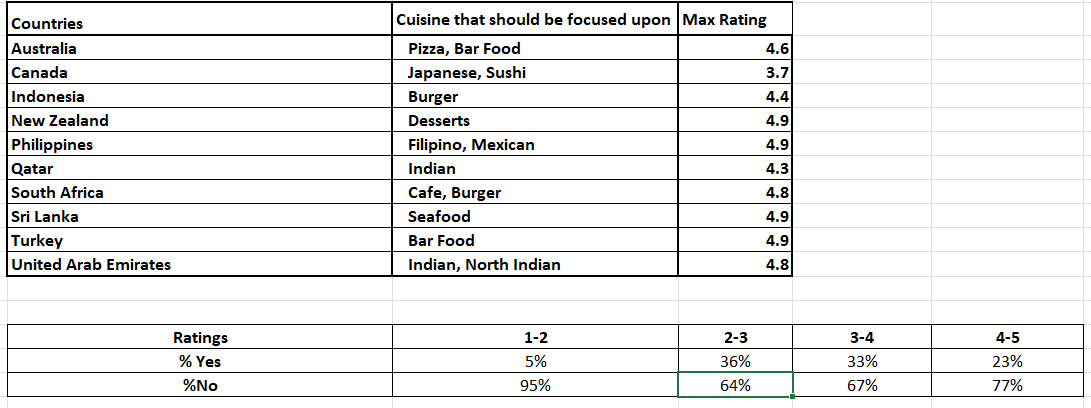
VISUALIZATION: PIVOT TABLES



OBSERVATION:

For countries- Turkey, South Africa, cuisines with no. of Rating and got highest no. of votes. So, focusing on the following Cuisines can enhance customer satisfaction and customer’s ratings.

For countries Philippines and New Zealand, some of the highest rating (4.5-4.7) cuisines have less no. of votes than some cuisines having rating 4.3- so, I tried to choose those cuisines which have rating>=3.5 and has good no. of votes.



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

**Approach: -** We must examine each country's online delivery and table booking trends as well as how ratings shift in order to provide recommendations for new eateries.   
  
In order to obtain ratings trends for each country about online delivery and table booking, I have utilized the PIVOT table and Pie Chart.

**Has Online Delivery Pie Chart**

**Has Table Booking Pie Chart**

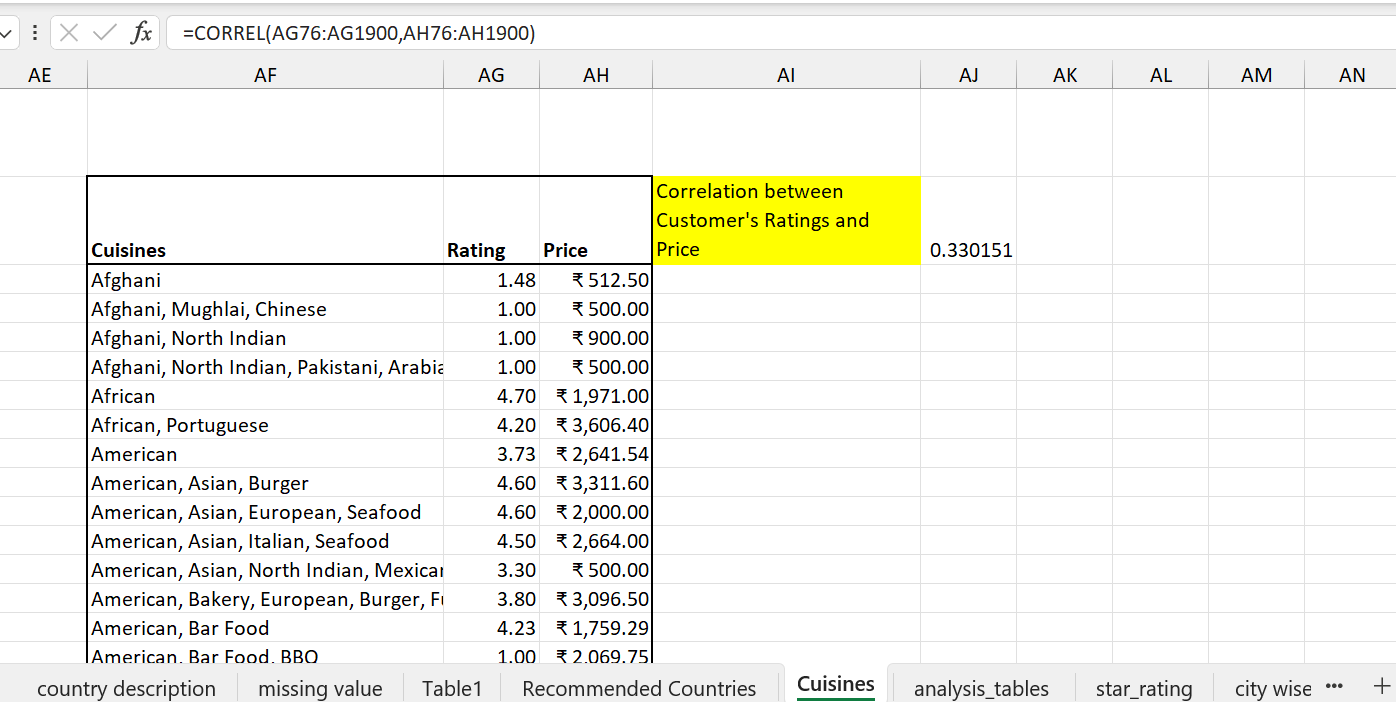
**Insights: -** Ratings for restaurants that offer both online ordering and table reservations are rising. Thus, we may conclude that online ordering and delivery do have an impact on consumer evaluations.

**Recommendations: -** For the convenience of the customer, incorporate online ordering and table reservations.

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?**

**Approach:** A scatter plot and a linear forecast trendline were used to illustrate the overall trend and direction of the data points, which made it simpler to see any linear relationship in the data. I used "Correlation Analysis" and a pivot table to visualize the distribution of cuisine ratings and the average cost for two for each country.

We may determine the degree and direction of the linear relationship between two variables by computing the correlation coefficient ("r") between cuisine ratings and avg\_cost\_for\_two.



[Where =CORREL (Rating, avg\_cost\_for\_two)]

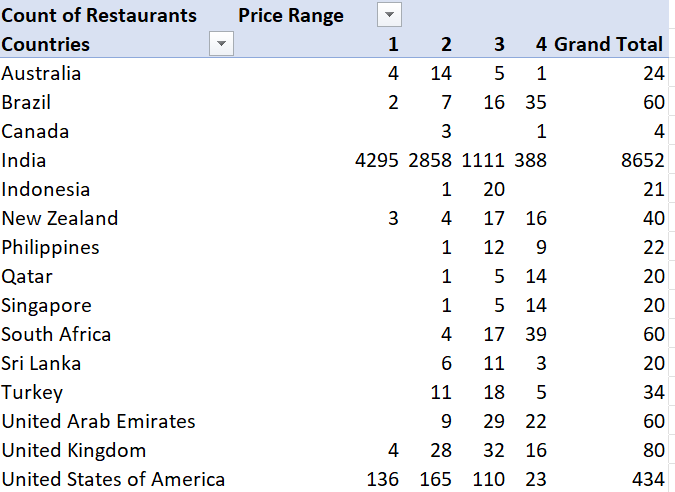
**Insights:** We may conclude that there is a positive correlation between the rates of cuisines and ratings because the linear forecast trendline is moving slightly upward and the correlation coefficient is trending towards +1. This means that as one variable rises, the other one tends to rise as well.

.

**Recommendation: -** Understanding the insights does not imply that we should raise the cost of every cuisine in every nation in order to improve ratings. Additionally, we should examine the rating and average price scatter plots by country and raise the rate when permitted by the government.

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

It can be found using a pivot table and line chart where the rows have the "Country Name" column, the price range is entered in the column, and the price range is entered in the values field. The field setting is changed to the count of the price ranges in various countries.



**Visualization: -** Doughnut chart since we need to show many parts (like price range) taking up space within a single unit (like country).

**OBSERVATIONS AND ANALYSIS**

1. **General Distribution:**
   * **Price Range 1**: This range has the highest number of restaurants (4444). It indicates a strong presence of budget-friendly or low-priced dining options across countries.
   * **Price Range 2**: This range has the second-highest number of restaurants (3113), suggesting a significant portion of mid-range dining options.
   * **Price Range 3**: There are (1408) restaurants in this range, showing a moderate presence of higher-priced dining options.
   * **Price Range 4**: The lowest number of restaurants (586) are in this range, indicating fewer high-end or premium dining options.
2. **Country-Specific Analysis:**
   * **Australia**: Shows a balanced distribution with some presence in all price ranges but with a higher number in the lower to mid-range categories (1 and 2).
   * **Brazil**: Exhibits a significant number of restaurants in the highest price range (35), indicating a robust market for premium dining options.
   * **Canada**: Data is sparse but shows a slight presence in the lower to mid-range categories.
   * **India**: Has a very high number of restaurants in the lowest price range (4295) and a substantial number in the mid-range (2858). This reflects a dominant trend towards budget-friendly and mid-range dining.
   * **Indonesia**: Limited data but shows some presence in lower and mid-range categories.
   * **New Zealand**: Has a relatively balanced distribution across price ranges, with notable numbers in the mid-range (17) and high-end (16) categories.
   * **Philippines**: Displays a range from lower to mid-range restaurants with a few in the high-end category.
   * **Qatar**: Shows a notable number of restaurants in the higher price ranges (5 and 14), suggesting a market for upscale dining.
   * **Singapore**: Similar to Qatar, there is a higher concentration of restaurants in the higher price ranges (5 and 14).
   * **South Africa**: The distribution is spread across all price ranges, with a notable number in the high-end range (39).
   * **Sri Lanka**: Has a more significant presence in the lower and mid-range categories.
   * **Turkey**: Shows a notable number of restaurants in the mid-range (18) and some in the lower (11).
   * **United Arab Emirates**: Displays a significant number of restaurants in the mid-range (29) and high-end (22) categories.
   * **United Kingdom**: Shows a high concentration in mid-range (32) and lower range (4) with fewer in the highest range (16).
   * **United States of America**: Has a high number of restaurants across all price ranges but is particularly notable in the lower (136) and mid-range (165) categories.
3. **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]**

Here is a methodical way to use the above data to recommend the best nations or locations for new restaurant openings:

**1.** **The number of eateries**   
• Using the Country and City columns, determine how many restaurants are located in each nation and city.   
  
• To identify places with less market saturation, highlight those with fewer eateries.   
 **2. Ratings and Votes:**   
  
• Analysing the rating and vote columns to learn about consumer preferences and feedback.   
  
• To guarantee client happiness and engagement, sites with higher average ratings and a respectable number of votes are chosen.   
  
  
**3. Popularity of Cuisine**   
  
• Examining the Cuisines column to determine which cuisines are underrepresented or popular in various geographical areas.   
  
• Identifying the cuisines that are most popular or underserved in particular regions so that you can customize your restaurant's menu.

**4. Evaluate Your Booking and Delivery Options**   
• To comprehend current services, look at the "Has\_Online\_delivery," "Has\_Table\_booking," and "Is\_delivering\_now" columns.   
• After determining the gaps in the current offerings, think about whether providing online ordering or table reservations could provide you a competitive edge.   
  
**5. Examine costs and affordability**   
• To comprehend price ranges and affordability, examine the RANGE\_OF Average\_Cost\_in\_INR.   
• Setting the restaurant's pricing in line with the competition and local spending patterns.   
  
**6. Take Geographical Distribution into Account**   
• Spatial analysis of restaurant locations using latitude and longitude.   
• Next, choose the best sites depending on how close they are to existing restaurant clusters and prospective patrons.

A synopsis of the main points

● Market Saturation: Pay attention to regions with fewer eateries already operating.   
● Customer Preferences: Focus on areas with a high vote count and rating.   
● Culinary Trends: Select restaurants according to popular or specialty cuisines.   
● Service Gaps: If online delivery or table reservations are not available, think about adding them.   
● Pricing Strategy: Comply with current pricing ranges and local affordability.   
● Local and Geographic Analysis: Select sites according to area features and ideal density.