5/18/2025

**SUHAS J**

Zomato Restaurant Expansion

**Objective Observations**

1. The given data of Zomato Restaurants Expansion has **Two** tables. Those are:

* **Raw Data** – This consists of all the important data such as restaurant name, restaurant id, city, country code etc
* **Country description** – This has the country name with country code

1. The data has a total of **22** attributes. Here the attributes refer to the columns in the sheets given in the dataset. Each attribute refers to the specific characteristics or the property the data stored.

* **21** Attributes in the **Raw Data sheet**. Those are:

RestaurantID, RestaurantName, CountryCode, City, Address, Locality, LocalityVerbose, Longitude, Lattitude, Cuisines, Currency, Has\_Table\_Booking, Has\_Online\_Delivery, Is\_delivering\_now, Switich\_to\_order\_menu, Price\_range, Votes, Averege\_cost\_for\_two, Rating, Datekey\_opening

* **1** Attribute in **country description**

CountryName

**Note:**

* Among all the attributes CountryCode plays a key role in helping in connecting two tables. It acts as a Primary Key in CountryDescription sheet and Foreign Key in RawData sheet.
* Count of attributes that are mentioned only include that are given in the dataset, new attributes that are created for the ease of analysis are not considered.

1. Categorical and Numerical / Continuous Data are the part of any data set.

* Categorical Data represents distinct groups or categories in the dataset.

Numerical / Continuous Data refers to the numerical values that can take any value within a range.

* So, this Data set has **17 Categorical** Columns. Those are:

RestaurantID, RestaurantName, CountryCode, City, Address, Locality, LocalityVerbose, Cuisines, Currency, Has\_Table\_Booking, Has\_Online\_delivery, Is\_delivering\_now, Switch\_to\_order\_menu, Price\_range, DateKey\_Opening and CountryName.

* There are **5 Numerical / Continuous** data column in the given dataset

Longitude, Lattitude, Votes, Averege\_cost\_for\_two and Rating.

1. **Data Inconsistency** may happen due to human errors, system migration, synchronization issue between dataset etc…. It results in poor insights, reduced data integrity and compromised or biased decision making. To prevent these data cleaning is done.

* Given Data is checked for duplicate entries. **No duplicate** entries were found in the data.
* For ease of analysis **CountryName** is added as a new column in the RawData sheet with the help of **VLOOKUP** function of excel with exact match criteria.

After the extraction of the CountryName, it is pasted as value in the Country column in the RawData sheet.

* There are blanks cells in the given data. There are a total of **9 blanks** in the column Cuisines. These are **removed** from the dataset as it does not affect the data vehemently.
* Datekey\_opening column is modified and new column is created with name DateKey\_Opening\_Modified
* Additionally created Year, month using text functions

**Year** - LEFT(U2,FIND("\_",U2,1)-1)

**Month** - MID(U2,(FIND("\_",U2,1)+1),((FIND("\_",U2,(FIND("\_",U2,1)+1)))- (FIND("\_",U2,1)+1))).

With the help of the above two columns DateKey\_Opening\_Modified is created with the using

**DATE(year,month,day)** and is pasted as value.

Quarter column is created using if else statement from newly created DateKey\_Opening\_Modified column.

* **Tbl\_currency\_exchange** is created in the country description column for the ease of calculation of transaction amount in a single currency mode rather than in a multiple currency mode.
* With the help of the above table new column **average\_cost\_for\_two\_USD and**  **average\_cost\_for\_two\_INR** is created using VLOOKUP function.

=S2\*VLOOKUP(D2,'country description'!$E$1:$G$16,3,0).

1. Used **VLOOKUP** function of excel to create a new column **“CountryName”** in the RawData sheet extracting from CountryDescription sheet.

Here the CountryCode column in used as connecting column between two sheets.

We can say that CountryCode column acts as Primary Key in Country Description and as Foreign Key in RawData sheet.

VLOOKUP(C2,Country\_desc,2,0)

Here,

* + 1. C2 refers to the cell where the country code is present in the RawData sheet. It is used as **lookup value**.
    2. Country\_desc refers to the table in the Country Description sheet where country name exists. It is used as **lookup array**.
    3. 2 refers to the second column i.e. **CountryName** in the table country\_desc.
    4. 0 refers to the **exact match** criteria.

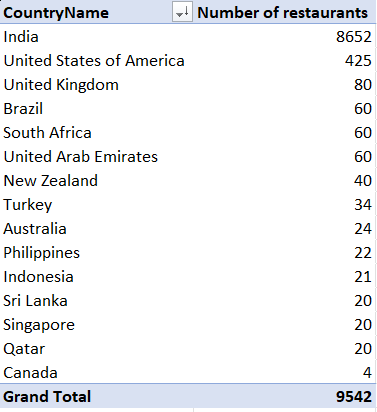
1. To get an overview on the total number of restaurants and number of restaurants in each country, a table is created with the help of Pivots which helps in aggregation of Total count of restaurants country wise.

Here we can observe that,

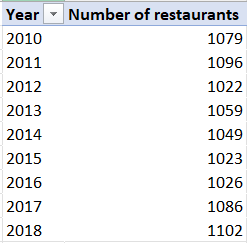
There are a total of **9542** restaurants across different parts of the world that are in associate with Zomato.

Over **91%** of the restaurants are in **India** with **8652** restaurants.

**Canada** has **least** number of restaurants associated with Zomato with only **4** restaurants.



1. Further continuing the analysis over year on year regarding the total number of restaurants opened.



The **average** number of restaurants opened per year for the last 8 years is **1061**.

However, **2018** has the most restaurants opened i.e. **1102** and **2012** has the least opening of restaurants with **1022.**

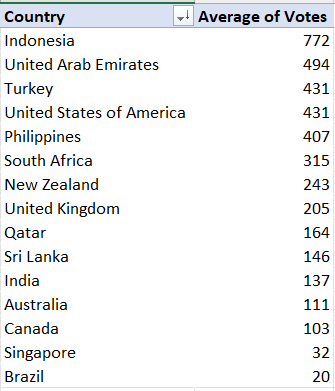
There is an increase of **2.13%** on restaurants opened from the year 2010 to 2018.

1. With over **91%** of the restaurants opened in **India**, we shall analyse the count of restaurants with price range 4.

COUNTIFS('Raw Data'!D2:D9543,"India",'Raw Data'!Q2:Q9543,"4")

From the above formula it is observed that **388** restaurants are present in India with price range 4. It suggests that out of 8652 restaurants only **4.5%** of the restaurants have **high price range**.

1. To get to know the success of the restaurant opened we have to consider attributes like votes and ratings received to the restaurants from the customers. We shall consider the voting in this as it shows the customer engagement in that country. We shall consider the average votes with regarding to each country



The average number of votes across the data is observed to be **157**.

**Indonesia** has the **highest** average in terms of votes compared to any other country in the dataset with **772** votes, whereas **Brazil** has **least** with only **20** votes.

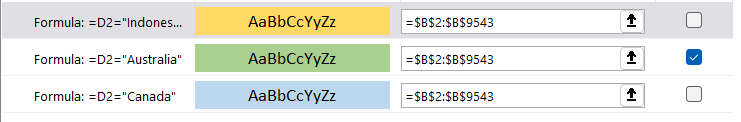
1. Now coming to the rating, we shall analyse the average rating of the restaurant which has **price range less than 4** and provides online delivery

AVERAGE(IF(('Raw Data'!Q:Q<4)\*('Raw Data'!N:N="Yes"),'Raw Data'!U:U))

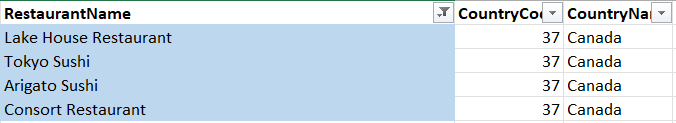
Above formula is used to calculate the average rating.

It is observed that **3.27** is the average rating of the restaurants for price range less than 4 and has online deliveries.

1. To highlight the restaurant name in the suggested country, **conditional formatting** is used to colour the specific restaurants in the country. On analysing the data, it is found that countries like **Canada, Australia and Indonesia** are suitable for opening restaurants.

In conditional formatting, “**use a formula to determine which cell to highlight**” is used in the **new rule** section and the below formula is used to highlight the cells.

According to this only the restaurants in the countries like Australia, Canada and Indonesia are highlighted with respective colours.



1. Averege\_cost\_for\_two in the data set is not properly formatted. So, a new column is created as Average\_cost\_for\_two\_mod where the money is concatenated with its respective currency symbol.

CONCATENATE(MID(L2,FIND("(",L2,1)+1,FIND(")",L2,1)-(FIND("(",L2,1)+1))," ",S2)

Above text formula is used to concatenate the Averege\_cost\_for\_two money with its respective currency symbol.

1. An analysis is made on the data with the count of restaurant that has average cost for two is less than or equal to 250 Rs, do not have online delivery and on the lowest price range.

To make such analysis on multiple condition with filtering the data and then summing up together to get the count, **SUMPRODUCT** of Arrayformula is used.

SUMPRODUCT(('Raw Data'!AB2:AB9543<=250)\*('Raw Data'!N2:N9543="No")\*('Raw Data'!Q2:Q9543=1))

Total of **1691** restaurants have satisfied the above assumed conditions.

**Subjective Question:**

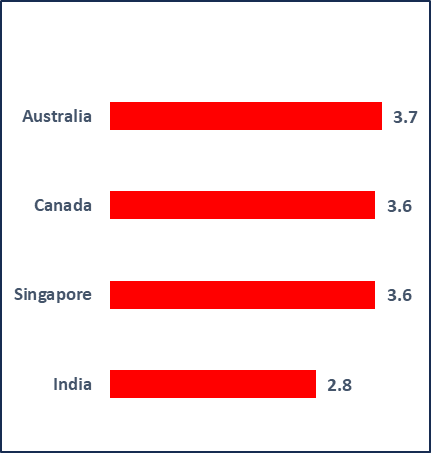
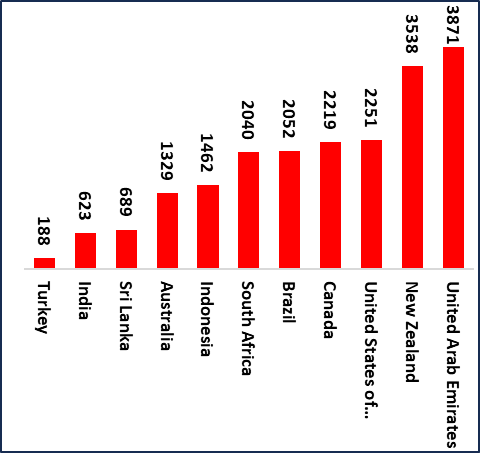
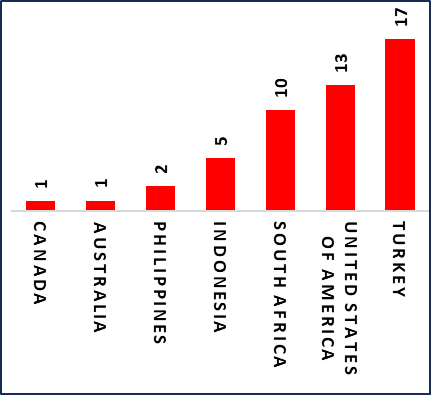
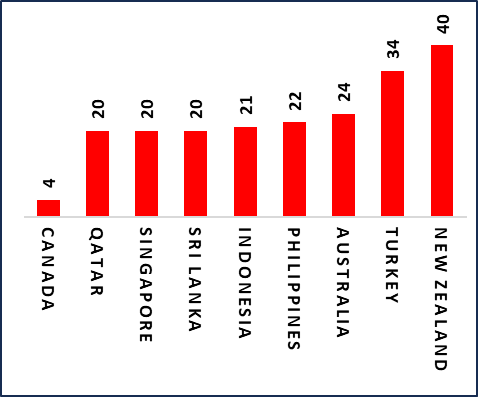
1. With over **9500** records in the data set it is important to take right decision regarding open of a new restaurant in a particular country or city where there are less competitors, has good returns in terms of money and easy to expand the restaurant chain in future.

On analysing various attributes given in the dataset few countries are suggest considering the following criteria,

* + Total **restaurant count** in a particular country is **less than 60**
  + Restaurant count to city **ratio** of a particular country is less than **18**
  + Average **rating** is less than **3.7**
  + Average **cost for two** is less than **Rs.4000**

**Restaurant to City Ratio**

**Country vs Restaurant Count**

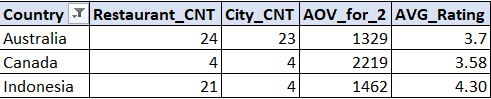


**Average rating**

**Restaurant vs AOV for 2**

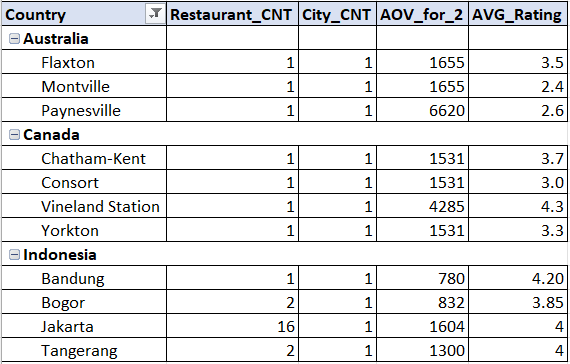
From the analysis made above, it is recommended to open the restaurants at **Canada, Australia and Indonesia**. Based on the addressable market, any of the mentioned country can be chosen since these countries have less competition and has huge potential in generating good returns on the money invested.

1. Analysing with suitable data points regarding the state and the city, few suggestions are made regarding the opening of the new restaurant. A pivot table is created to justify the suggestion given above.



Most suitable locations are:

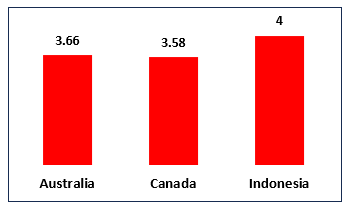
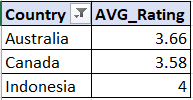
* + **Australia** – Flaxton, Paynesville, Montville
  + **Canada** - Chatham-Kent, Consort, Vineland Station, Yorkton
  + **Indonesia** – Tangerang

****

Reason for suggestion:

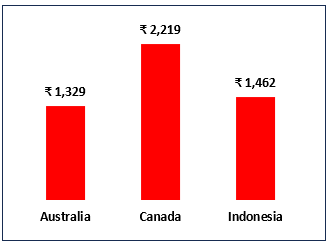
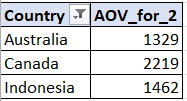
* + **Australia** 
    - Its City to Restaurant count ratio is 1, which means there is only one restaurant for each city.
    - Its average order value for 2 is around 1350, but in some places like Flaxton and Montville its over 2000 and in Paynesville its around 6500
    - Though the country average rating is 3.7, in high order value city its below 2.5 which means there is a huge potential for good quality and tasty food
  + **Canada**
    - Its City to Restaurant count ratio is 1, which means there is only one restaurant for each city.
    - Its average order value for 2 is over 2200, but in places like Vineyard station its over 4500 with only one restaurant
    - Though the country average rating is 3.5, in some cities around 3 which indicates there is a huge demand for good quality and tasty food
  + **Indonesia**
    - Though its City to Restaurant count ratio is 5, which means there are 5 restaurants for each city, its majority of the restaurants i.e. with over 65% of the restaurants are in Jakarta rest other cities have around 2 restaurants per city.
    - Its average order value for 2 is around 1450, but in some place like Tangerang its around 1300 with satisfactory rating of 4, which means with good quality and better food there can be increase in AOV.

1. To get into the **ratings** for the restaurants suggested, a pivot table is created where the analysis can be made regarding quality of the restaurant.



**Australia and Canda**’s rating are below **3.5** which means there is a need of improvement in the service and food at these restaurants, by proving good customer satisfaction there is a huge potential in these countries.

1. For maintaining the **expenditures** of the restaurants, analysis is made on average cost for 2 to forecast the financial expenditure.



The average cost for two in,

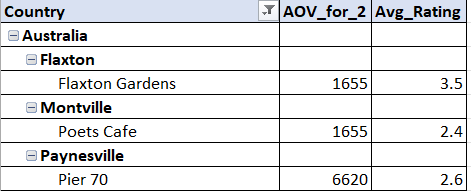
* + - **Australia** is around **Rs.1350** but in places like Flaxton and Montville its over Rs.2000 and in Paynesville its around Rs.6500.
    - **Canada** is over **Rs.2200**, but in places like Vineyard station its over Rs.4500 with only one restaurant
    - **Indonesia** is over **Rs.1450.**

According to the average cost for two in these countries, a forecast of expenditure has to be made so that the financial expenses are maintained efficiently.

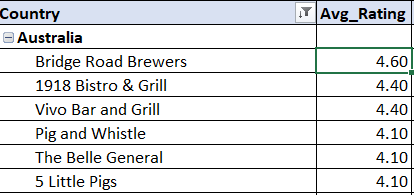
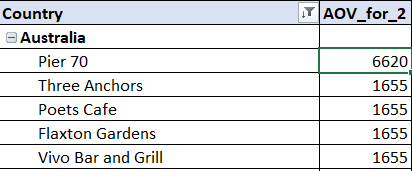
1. Having Idea about the **competitors** and making analysis on them is a must when it comes to entering into the market. Competitor analysis helps us to set a benchmark in the performance, discover opportunities and Improve Strategy.

When it Comes to the competitors in the suggested Countries and cities with price range 1-2 and 2-3 are,

* **Australia** 
  + Since there are only one restaurant in each city there is no much of competition when compared city wise, however in the suggested cities like Flaxton, Montville, Paynesville there are competitors like Flaxton Gardens, Poets Café and Pier 70 respectively.
  + Even though they have high AOV for 2 but their rating is very low. To stay ahead with the quality food and ambiance is sufficient.



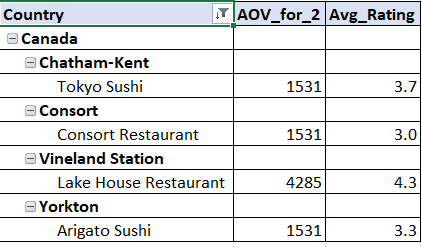
* However, when we compare the restaurants country wise,



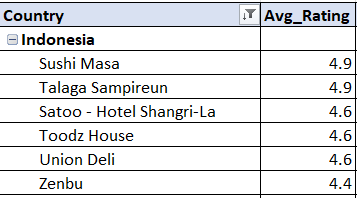
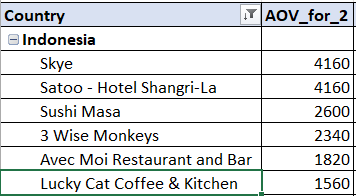
Top 5 competitors w.r.t AOV\_for\_2

Top 5 competitors w.r.t Avg\_rating

* **Canada**
  + Since there is only one restaurant in each city there is no much of competition when compared city wise, however in the suggested cities like Chatham Kent, Consort, Vineland Station, Yorkton there are competitors like Tokyo Sushi, Consort Restaurant, Lake House Restaurant, Arigato sushi respectively.
  + A very good competitor is Lake View Restaurant from Vineland Station which has a good AOV for 2 and rating, rest others have moderate AOV for 2 and below par rating



* **Indonesia** 
  + Most of the restaurants are situated in Jakarta, Indonesia. So, at the remaining location there is no much of restaurants. In the suggested city Tangerang “Talanga Sampireun” has moderate AOV for 2 and high rating, where as Onokabe has healthy AOV for 2 and satisfactory rating.



Top 5 competitors w.r.t AOV\_for\_2

Top 5 competitors w.r.t Avg\_rating

1. A restaurants success majorly lies upon the food that is served to the customer. The data clearly reflects that **cuisine popularity** varies across different countries, which in turn is essential to identify the best choice of cuisine for the recommended countries.

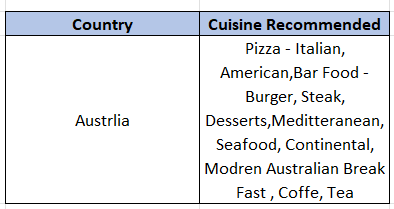
This analysis utilizes a Pivot table, taking into account the popularity of cuisines in the recommended region while also factoring in globally favoured cuisine types for a more comprehensive perspective.

Suggestion for the recommended countries,

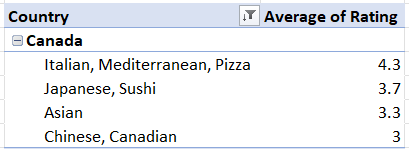
* **Australia**



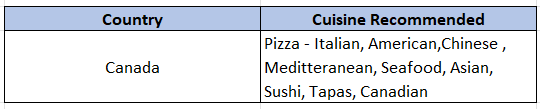
By analysing the average ratings, the top five most popular cuisines in Austria - Pizza, Bar Food, Mediterranean, Seafood, and Modern Australian are identified as key culinary preferences. To introduce more variety within similar cuisine types, we also explore globally popular cuisines that align with these, providing a broader perspective on food trends

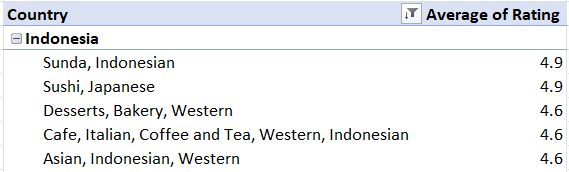


* **Canada**



Top Cuisines that are popular in Canada are Pizza, Japanese, Sushi, Asian, Chinese. To introduce more variety within similar cuisine types, we also explore globally popular cuisines that align with these, providing a broader perspective on food trends.



* **Indonesia**

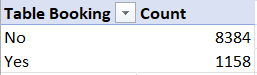
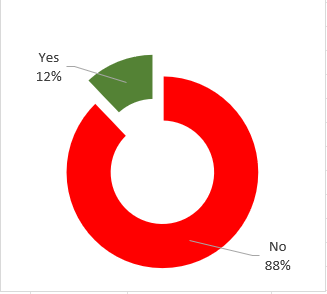
By analysing the average ratings, the top five most popular cuisines in Indonesia Sunda, Indonesian, Sushi, Japanese, Desserts, Bakery, Coffee, Italian, Tea, Asian are identified as key culinary preferences. To introduce more variety within similar cuisine types, we also explore globally popular cuisines that align with these, providing a broader perspective on food trends



Incorporating the recommended cuisines tailored to each country will greatly enhance customer feedback and restaurant ratings, while also driving significant revenue growth.

1. A pivot table was used to analyse how ratings are influenced by **online delivery** and **table bookings**, assessing whether these factors have a significant impact.

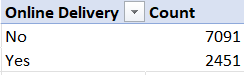
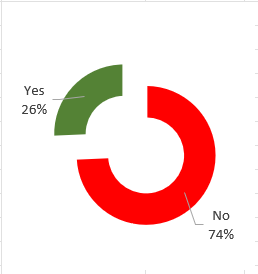
The following pivot table presents a breakdown of restaurants based on whether they offer table booking or not



This Shows over **88%** of the restaurants do not offer table bookings at the restaurant.

**Restaurants with table booking facility**

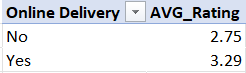
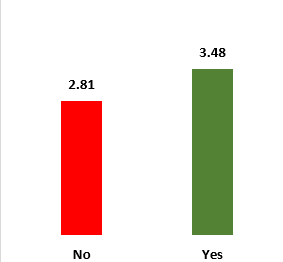
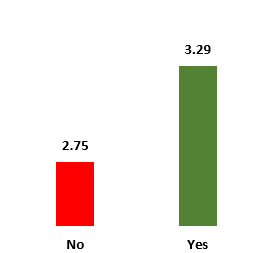
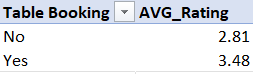
The following pivot table presents a breakdown of restaurants based on whether they offer online delivery or not



This Shows over **74%** of the restaurants do not offer online delivery

The following data illustrates the influence of online delivery and table bookings on restaurant ratings.

**Online Delivery vs AVG\_rating**



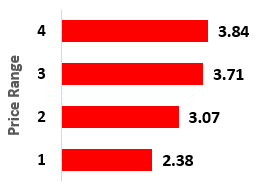
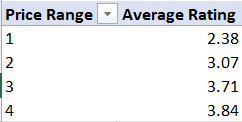
**Table \_booking vs AVG\_rating**

The graphs above indicate that, although only a small percentage of restaurants offer online delivery (26%) and table booking (12%), these establishments tend to have significantly **higher ratings** compared to those that do not. Therefore, incorporating online delivery and table booking options is recommended and has impact on customer ratings and overall restaurant performance.

**Recommendation:**

Since recommended countries like **Australia, Canada, and Indonesia** currently do not offer online delivery or table booking feature integrating this service could be a key strategy to enhance **restaurant ratings and customer satisfaction**. Providing a **seamless delivery and reservation system** would likely improve the convenience and establish a sense of reliability among customers, ultimately leading to higher engagement and better reviews.

1. Determining the appropriate **pricing** for a cuisine plays a crucial role in customer satisfaction and feedback. To facilitate analysis, a pivot table has been created to examine the provided data, specifically focusing on the price range of the cuisine and the average rating it has received.



The data clearly indicates a positive correlation between price range and average rating, with higher price ranges receiving significantly higher ratings. This is even confirmed with the formula to check the correlation between two ranges.

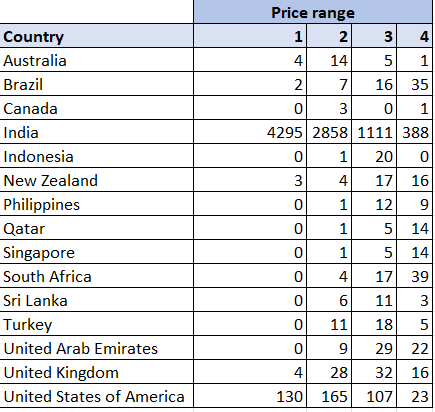
=CORREL(A210:A213,B210:B213) = **0.96**

Here, A column refers to Price Range and B column refers to AVG\_rating.

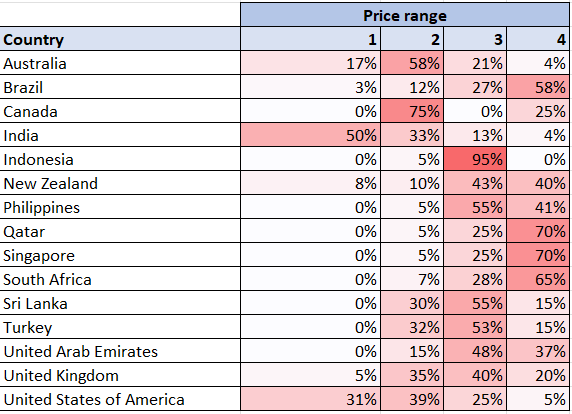
The graph and calculated score clearly demonstrate the influence of cuisine pricing on ratings, highlighting a measurable impact. A price range between 3 and 4 is advisable, as it tends to yield comparatively better ratings when assessed alongside other pricing brackets. This selection can help optimize customer satisfaction and feedback.

While the ratings for price ranges 3 and 4 fluctuate between 3.71 and 3.84, they do not stand out as exceptionally high. This suggests that pricing alone is unlikely to significantly influence restaurant ratings. To enhance overall customer satisfaction, it is crucial to consider multiple factors, including food quality, pricing strategy, and ambiance.

1. A pivot table is utilized to analyse the distribution of restaurants across different price ranges, providing precise insights into this aspect. This approach ensures a clear and structured understanding of how restaurant numbers vary based on pricing categories.



Conditional formatting is applied to enhance the visualization of percentage values in the table, making the data more interpretable and impactful.



Excluding India and the United States, most countries have a concentration of restaurants within the price range of 3 and 4. This distribution aligns with the observed correlation between pricing and average ratings in these categories, further reinforcing the impact of pricing on customer perceptions.

1. **Approach**

* **Understanding Business Requirements:** Define key performance indicators (KPIs) based on the business objectives and expectations.
* **Exploring Data Structure:** Examine the provided datasets, including table relationships and attribute significance.
* **Data Cleaning & Preparation:** Identify and resolve inconsistencies such as duplicates and null values to ensure accurate analysis.
* **Data Enhancement:** Create additional columns if needed to streamline analysis and improve insights.
* **Dynamic Data Summarization:** Utilize pivot tables to generate country-wise summaries and facilitate interactive analysis.
* **Visualizing Insights:** Develop charts for better comprehension of key observations and trends.
* **Restaurant Affordability Classification:** Establish a five-band price range to categorize restaurants based on affordability.
* **Quarterly Performance Tracking:** Analyse financial metrics to maintain cost efficiency and revenue optimization.
* **Regional Cuisine Analysis:** Evaluate the popularity and performance of various cuisines in different regions to inform menu decisions.
* **Comprehensive Reporting:** Design dashboards to present the analysis visually and effectively for decision-making.