**ZOMATO RESTAURANT’S EXPANSION PROJECT !**

****

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

* **APPROACH** - For data cleaning, applied filter on each column and check for blanks. some BLANK values were found in the Cuisines Column, blanks values were replaced with FOOD.
* Also, need to extract the year from the Date-Key-Opening Column using split text to columns function.

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

* **APPROACH -** Used XLOOKUP Function in Raw data sheet for fetching the country name from the Country description sheet.
* Inserted a Country column next to the country code column in the raw data sheet for accurate analysis, Using XLOOKUP Function **“=XLOOKUP(C2,'country description'!$A:$A,'country description'!$B:$B)”.**

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

* **APPROACH -** Created a pivot table using Raw Data, in row added the country and in value added the Restaurant Name and summaries by count function.
* **FINDINGS -** The total number of restaurants in the given data set is **9551.** India has the highest number of restaurants among all countries around **8652** restaurants and Canada has the least number of restaurants, that is **4**.

| ***Country Name*** | **COUNT Of Restaurants** |
| --- | --- |
| Australia | 24 |
| Brazil | 60 |
| Canada | 4 |
| India | 8652 |
| Indonesia | 21 |
| New Zealand | 40 |
| Philippines | 22 |
| Qatar | 20 |
| Singapore | 20 |
| South Africa | 60 |
| Sri Lanka | 20 |
| Turkey | 34 |
| United Arab Emirates | 60 |
| United Kingdom | 80 |
| United States of America | 434 |

1. **Derive the number of Restaurants opened each year.**

* To answer this we need a Year column. For this, have Inserted columns next to the Datekey\_Opening column in the raw data sheet and then applied **‘split text to the column’** on the column, and then separated them using delimiters of ‘\_’. With this, I have extracted the Year, Month & Date from the Datekey\_Opening and **labeled it as Opening Year, Opening Month & Opening date.**
* **APPROACH -** Created a pivot table using Raw Data, add Year in row and Restaurant Name in value and summaries by count function. Along with this, I’ve also created a Bar chart for the better representation of data on how many restaurants have opened in each year.
* **FINDINGS -** This data confirms that more than 1000 restaurants are opened each year.
* As per the analysis, the maximum number of restaurants open in **2018 (i.e. 1102)** and the minimum number of restaurants open in **2012 (i.e. 1022)**.

| *Opening Year* | COUNT of Restaurants |
| --- | --- |
| 2010 | 1080 |
| 2011 | 1098 |
| 2012 | 1022 |
| 2013 | 1061 |
| 2014 | 1051 |
| 2015 | 1024 |
| 2016 | 1027 |
| 2017 | 1086 |
| 2018 | 1102 |

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

* **APPROACH -** Created a pivot table using Raw Data, add country in row and Restaurant Name in value, apply filter on price range then select 4 and filter on country name then select India.
* We can also derive The total number of restaurants in India using COUNTIFS Function **=COUNTIFS('Raw Data'!D2:D,"INDIA",'Raw Data'!Q2:Q,"4")**
* **FINDINGS -** The total number of restaurants in India which are in the price range of 4 are **388**.

| *Country Name* | COUNTA of Restaurant Name |
| --- | --- |
| India | 388 |
|  |  |

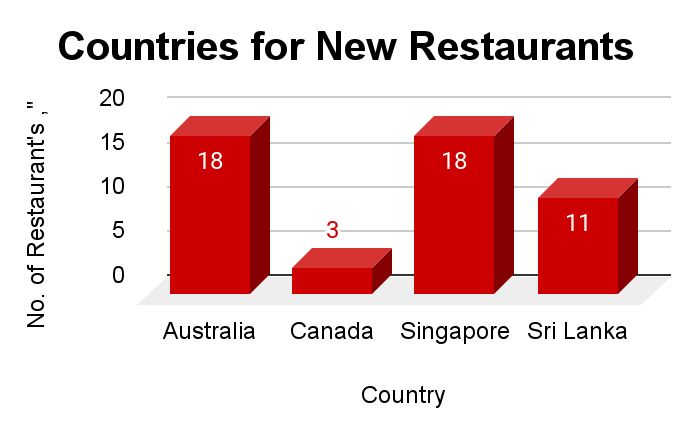
1. **According to the data, what is the average number of voters for the restaurants in each country?**

* **APPROACH -** Created the pivot table and add country in row and votes in value and summaries by Average function. Used stacked area chart for better visuals.
* **FINDINGS -** As per our data, Indonesia has the highest mean exactly **772** votes among all the countries and Brazil got minimum average of votes i.e. **19**.

| ***Country Name*** | **AVERAGE of Votes** |
| --- | --- |
| Australia | 111.42 |
| Brazil | 19.62 |
| Canada | 103.00 |
| India | 137.21 |
| Indonesia | 772.10 |
| New Zealand | 243.03 |
| Philippines | 407.41 |
| Qatar | 163.80 |
| Singapore | 31.90 |
| South Africa | 315.17 |
| Sri Lanka | 146.45 |
| Turkey | 431.47 |
| United Arab Emirates | 493.52 |
| United Kingdom | 205.49 |
| United States of America | 428.22 |

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

* **APPROACH -** In the Raw data, state is not available, Hence created a pivot table with country and city in row and Restaurants name in value & summaries by count function. Rating name in value & summaries by Average function. Filtered Countries where the number of restaurants is less than 50. Along with that Filtered number of restaurants whose rating is less than 4.
* Firstly, filtered out the countries where the number of restaurants was less than 50. This will help us to understand where the competition is less.
* **INSIGHTS** - But we also have to look at the current market scenario. There can be fewer popular restaurants which are highly in demand. These Restaurants can give us a tough competition and make it difficult to grab the market. To avoid this, have added the average rating in the Pivot Table to understand the market scenario. After this, those countries were automatically filtered having less than 4 average ratings.
* Having less rating clearly states that the people from those countries were not very satisfied with the quality of the food of the restaurants around them. Those were the right place for our expansion and will take care of their local food, market needs and the quality of food. We will offer more services like online delivery and table booking to increase the customer’s ratings because customers enjoy additional ease & convenience in their dining experience.
* **FINDINGS -** Here are a few countries & cities for opening new restaurants. As per the Survey, top 4 selected countries are given below, which are Canada, Singapore, Sri Lanka and Australia. Created a pivot chart describing the Count of Restaurants and average of ratings. Similarly created a Column Chart where plotted the Count of Restaurants for the countries which are selected for better visuals.

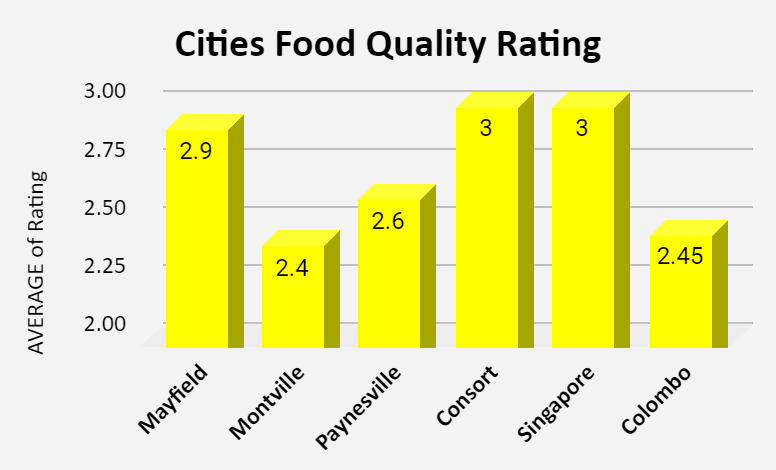


| *Country Name* | *COUNT of RestaurantID* | AVERAGE of Rating |
| --- | --- | --- |
| Australia | 18 | 3.45 |
| Canada | 3 | 3.33 |
| Singapore | 18 | 3.51 |
| Sri Lanka | 11 | 3.54 |

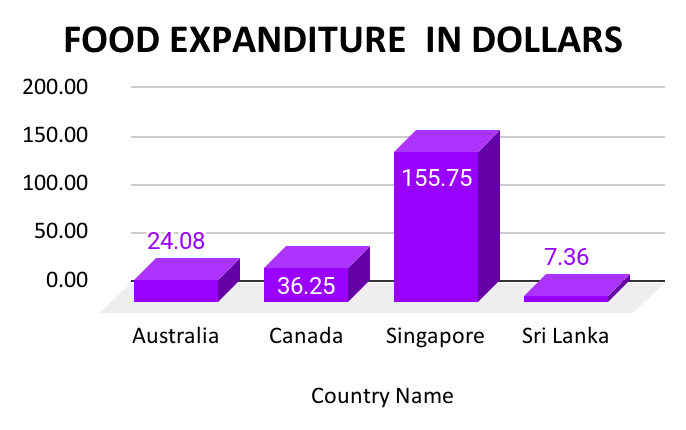
1. **Come up with the names of States and cities in the suggested countries suitable for opening restaurants and What is the current quality regarding ratings for restaurants that are open in suggested countries ?**

* **APPROACH -** Firstly, Listed down all the cities in the selected countries along with the number of Restaurants along with the average rating for the city. In that Pivot table, added the country column in row and restaurants name in value, summaries by count function.
* **FINDINGS -** Arranged all of the data in ascending order of Rating and then filter out only cities having less than 3 ratings for lesser competition. and as a result, we have **six** cities - “**Mayfield, Montville, Paynesville, Consort, Singapore and Colombo”** where we open new restaurants.

| ***Country Name*** | ***City*** | **AVERAGE of Rating** | **COUNT of RestaurantID** |
| --- | --- | --- | --- |
| **Australia** | **Mayfield** | 2.9 | 1 |
|  | Montville | 2.4 | 1 |
|  | Paynesville | 2.6 | 1 |
| Canada | Consort | 3 | 1 |
| Singapore | Singapore | 3 | 1 |
| Sri Lanka | Colombo | 2.45 | 2 |



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

* For this, created a separate sheet “**Currency Conversion**” and listed down the current conversion rate for all the different Currency types in the Raw Data. After that, inserted a new column next to the “Average\_Cost\_for\_two” column with the name of the column as **“Average\_Cost\_for\_two\_in\_Dollars”.** In this column, converted all the values in Dollars by using the VLOOKUP formula**“=VLOOKUP(L2,CURRENCY!$A:$B,2,0) \*S2”** from the Currency conversion.
* **APPROACH -** Created a Pivot Table added Country in Rows and Average of Average\_Cost\_for\_two\_in\_Dollars in Values and then filtered only the selected countries. Apart from this, Created the visualization with the column Chart where we can see what is the average cost for two in Dollars in the restaurants in the selected countries.
* **FINDINGS -** In Sri Lanka we spend an average amount like 7.36$ and in Singapore we need to spend 155.75 $ as the financial expenditure.
* We must keep our financial expenditure in control and have a balanced expenditure on food.

| Country | Cost\_for\_two\_in\_Dollars |
| --- | --- |
| Australia | 24.08 |
| Canada | 36.25 |
| Singapore | 155.75 |
| Sri Lanka | 7.36 |

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.**

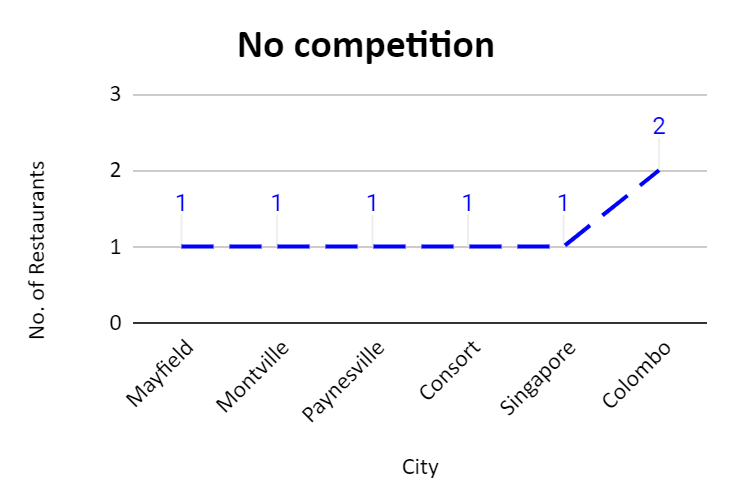
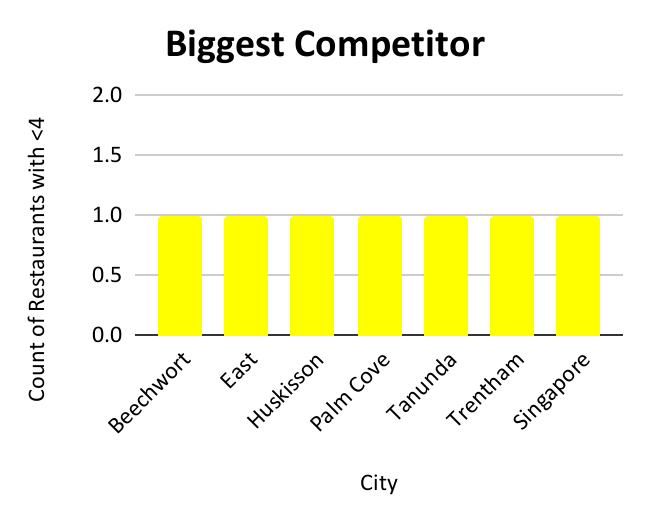
* **APPROACH -** Created the pivot table and used a rating filter in both cases, more than 4.5 for Biggest Competitors & less than 3 for no competition. Also used filter on cities. For better visualization, used the column chart to display this data.
* **FINDINGS -** As per the given data, Australia is the only country where we have 6 high competition cities **(Beechworth, East Ballina, Huskisson, Palm Cove, Tanunda, Trentham East)** as well as 3 low competition cities (**Mayfield, Montville, Paynesville).**

We can start from low competition cities and after that once we get a good grip on the Australian market then we can start with high competition cities.

* We have 5 cities suitable for opening our new restaurants with negligible competition. Those 5 cities are **Mayfield, Montville, Paynesville, Consort, Singapore.** Please refer to the column charts given below -

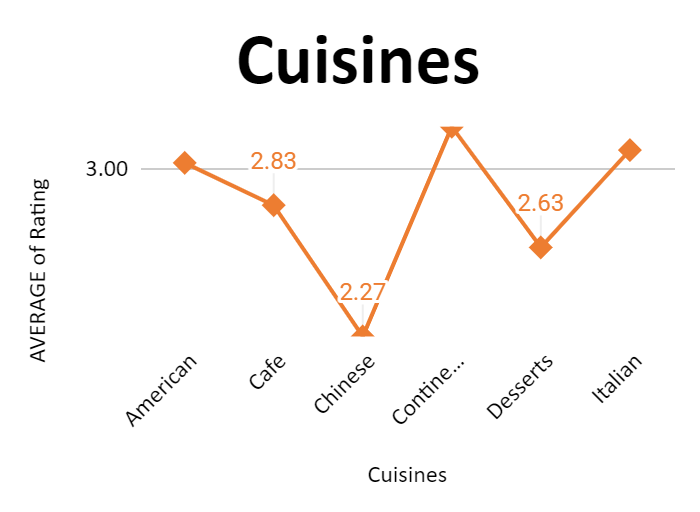
| *Country Name* | *City* | COUNTA of Restaurant NameChart | Cost\_for\_two\_in\_Dollars |
| --- | --- | --- | --- |
| Australia | Beechworth | 1 | 20.0 |
|  | East Ballina | 1 | 20.0 |
|  | Huskisson | 1 | 20.0 |
|  | Palm Cove | 1 | 30.0 |
|  | Tanunda | 1 | 30.0 |
|  | Trentham East | 1 | 20.0 |
| Singapore | Singapore | 1 | 20.0 |

| *Country Name* | *City* | COUNT of Restaurants | Cost\_for\_two\_in\_Dollars |
| --- | --- | --- | --- |
| Australia | Mayfield | 1 | 20.0 |
|  | Montville | 1 | 30.0 |
|  | Paynesville | 1 | 120.0 |
| Canada | Consort | 1 | 25.0 |
| Singapore | Singapore | 1 | 30.0 |
| Sri Lanka | Colombo | 2 | 5.9 |



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

* **APPROACH -** To determine Rating, we add cuisines in row, rate in value and summaries by Average function in the pivot table.
* Firstly, listed all the Cuisines which are getting served in the selected country along with the average of Rating and the total votes through which the rating has been decided and arranged the data into ascending order of average of rating. With this dataset, we will get to know the cuisines which are rated by a large number of people.
* **INSIGHTS -** we must know the cuisine is currently in demand. Now, we have taken the cuisines where the rating are below 3.8 rating. If the rating of any cuisine is above 3.8 rating then we should avoid the cuisines as there is a high chance of tough competition.
* the choices of cuisines do affect the rating of the restaurant because if we provide the customer with the type of food which is generally bad in the country with the good quality then there are high chances that they will like the food and give the restaurant higher ratings. But if we serve food which is already highly rated, then the smallest of the mistakes will lead to a bad rating.
* **FINDINGS -** In our newer restaurants, we should focus majorly on American food, Italian food and Continental food. Plotted a line chart for better visuals.



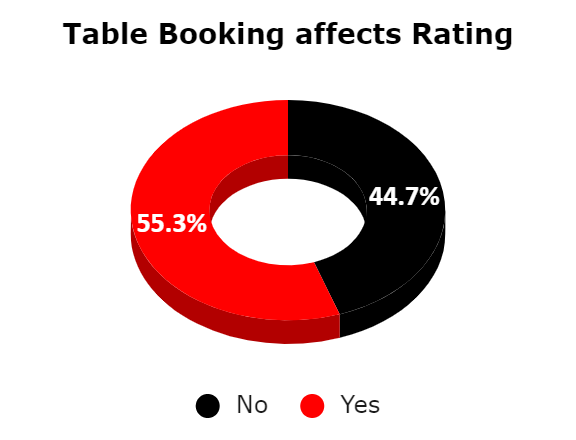
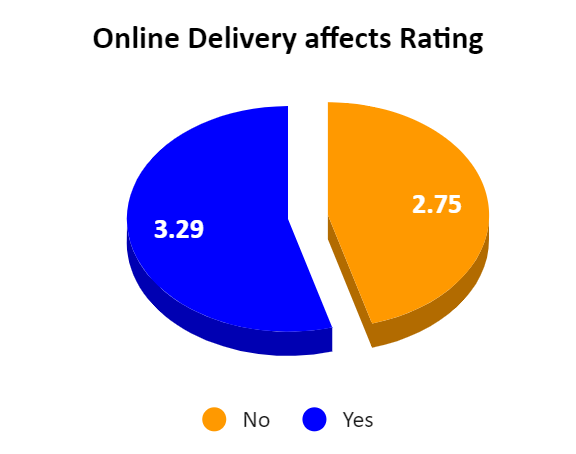
| *Cuisines* | AVERAGE of Rating | SUM of Votes |
| --- | --- | --- |
| American | 3.03 | 1053 |
| Cafe | 2.83 | 13799 |
| Chinese | 2.27 | 15222 |
| Continental | 3.22 | 1384 |
| Desserts | 2.63 | 1290 |
| Italian | 3.10 | 4174 |

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

* **APPROACH -** In Pivot chart, table booking in row and added average of rating in values. same with online delivery and created pie charts for better visualization.
* **INSIGHTS - Yes**. As per the survey, we must opt for online delivery and table booking, It affects the customer’s ratings.
* Restaurants offering more services (online delivery and table booking)tend to have higher ratings because customers enjoy additional ease & convenience in their dining experience. Refer to the pie charts for better clarity.
* We can offer the current high rating Restaurants to tie up with us. We can approach them.

| *Has\_Online\_delivery* | AVERAGE of Rating |
| --- | --- |
| No | 2.75 |
| Yes | 3.29 |

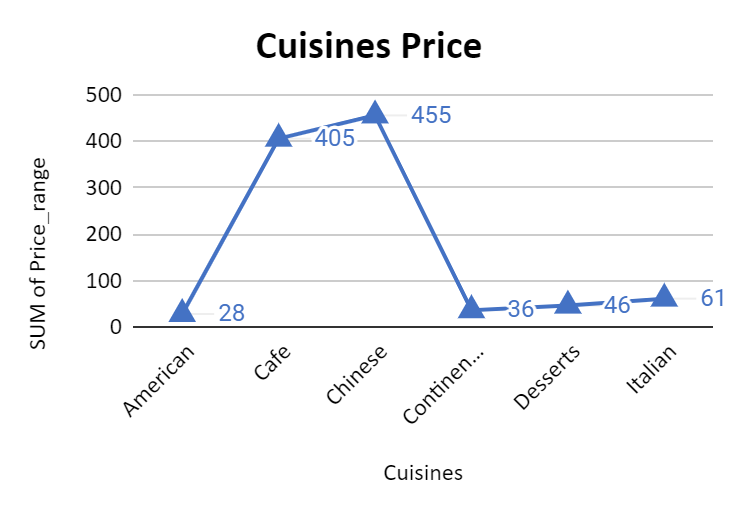
| *Has\_Table\_booking* | AVERAGE of Rating |
| --- | --- |
| No | 2.81 |
| Yes | 3.48 |



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?**

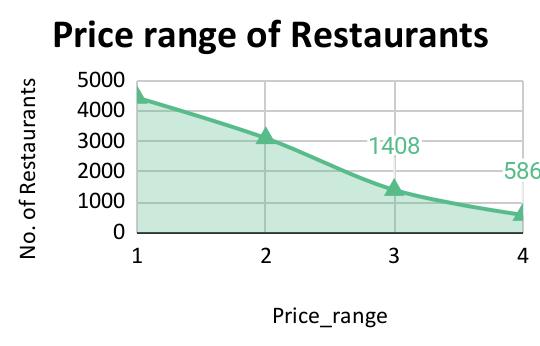
* **APPROACH -** Created Pivot chart, as cuisines in row, added average of rating and price range in value.
* **INSIGHTS -** According to our data, the rates of cuisines and ratings are not directly correlated. But the quality and ratings are directly related. To get a higher rating, We must keep the rate of cuisines in the middle range and keep high quality cuisines in our newer Restaurants.

| *Cuisines* | Rating | Price\_range |
| --- | --- | --- |
| American | 4 | 2.58 |
| Australian | 4 | 2.00 |
| Cafe | 3 | 1.73 |
| Chinese | 2 | 1.48 |
| Chinese, Canadian | 3 | 2.00 |
| Continental | 4 | 3.00 |
| Desserts | 3 | 1.15 |
| Italian | 4 | 2.98 |
| Seafood | 4 | 2.93 |
| Singaporean, Chinese, Seafood, Malay, Indian | 3 | 3.00 |
| Sri Lankan | 4 | 3.00 |



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

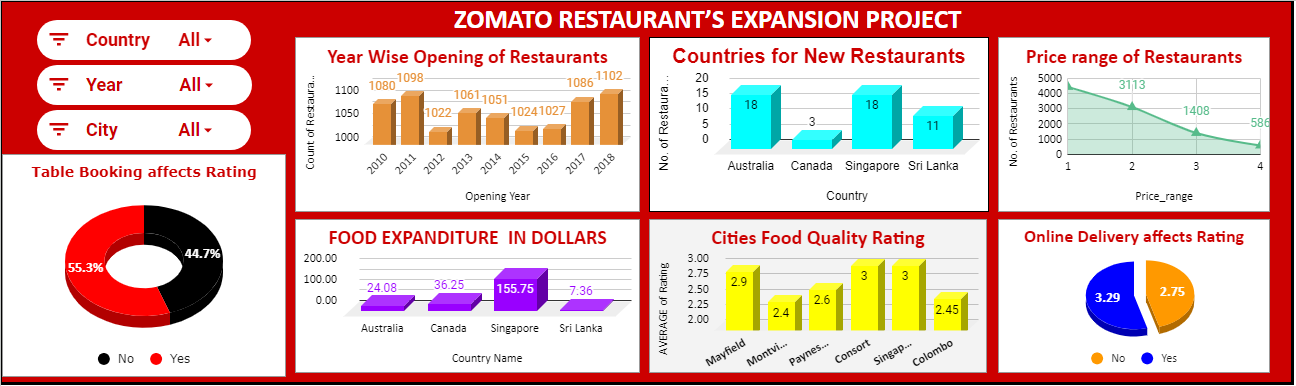
* **APPROACH -** Used a pivot table to fetch the different price range of restaurants of all the countries, add Price range in row and Restaurants Name in value and summaries by Count function.
* **FINDINGS -** As a result, we can see that there are less number of restaurants having higher price range.



| *Price\_range* | COUNTA of Restaurant Name |
| --- | --- |
| 1 | 4444 |
| 2 | 3113 |
| 3 | 1408 |
| 4 | 586 |

1. **The dashboard must consist of date and country slicers.**

* The Slicers added in the Dashboard for Country and Year. As we are dealing more with years not with exact dates. Also added Slicers to cities and cuisines, as these are our major focused areas.



**Main** **Findings**

* Three main findings emphasize on the following:

1. Higher the number of services/cuisines provided by a restaurant, more likely it is to receive higher ratings.
2. Variety & quality of cuisines attracts the customers.
3. Higher Cost-for-2 for a Restaurant, the less likely it is to be popular.

**Conclusion**

1. Top 4 selected countries for our expansion are Australia, Canada, Singapore, Sri Lanka.
2. Most suitable 6 cities for opening our new Restaurants are as follows - Mayfield, Montville, Paynesville, Consort, Singapore, Colombo. Our Major Focus will be on Australia, As this Country has 3 most suitable cities.
3. In our newer restaurants, we should focus majorly on American food, Italian food and Continental food. The Choice and variety of Cuisines affects the restaurantratings.
4. To get a higher rating, We must keep the rate of cuisines in the middle range and keep high quality cuisines in our newer Restaurants.
5. We must offer more services like online delivery and table booking to increase the customer’s ratings because customers enjoy additional ease & convenience in their dining experience.