**Index**

1. Introduction

* Objective of Analysis
* Problem Statement

1. Data Overview
2. Methodology
3. Analysis of Objective Questions
4. Insights from Subjective Questions
5. Strategic Recommendations
6. Dashboard and Visualizations
7. Conclusion

**ZOMATO DATA ANALYSIS**

**Introduction**

Zomato is one of the world’s leading restaurant discovery and food delivery platforms, founded in 2008 and headquartered in India. With operations in over 20 countries, Zomato connects millions of users with restaurants, enabling them to browse menus, read reviews, place orders, and make reservations. Over the years, Zomato has become a vital part of the urban dining experience and continues to expand its services to new markets globally.

As the food and hospitality industry becomes increasingly data-driven, Zomato leverages insights from its massive dataset to make informed business decisions. One of the key areas of focus is strategic expansion—identifying new cities or countries with high growth potential and relatively low competition. This requires a deep understanding of customer preferences, operational metrics, and market dynamics.

The goal of this project is to support Zomato’s expansion strategy by analysing its restaurant data using Microsoft Excel. The dataset contains detailed information on restaurant locations, pricing, cuisines, delivery options, user ratings, and more. Through data cleaning, transformation, and visualization, this report aims to uncover actionable insights.

An interactive Excel dashboard will be created to help management explore various dimensions such as:

* **Market saturation by location**
* **Cuisine performance**
* **User rating patterns**
* **Pricing effectiveness**
* **Delivery and table booking trends**

This dashboard, along with data-backed recommendations, will serve as a decision-making tool for Zomato’s leadership team as they plan to enter new markets or strengthen existing ones.

**Objective of This Project**

The primary objective of this project is to analyse Zomato’s restaurant dataset to provide strategic recommendations for business expansion. By leveraging Excel’s data handling and visualization capabilities, the aim is to develop a comprehensive and interactive dashboard that enables management to:

* **Identify potential countries, cities, or localities** for opening new restaurants based on market gaps, customer demand, and competition levels.
* **Evaluate current restaurant performance** through key indicators such as ratings, price range, average cost for two, and customer feedback (votes).
* **Understand customer preferences** in terms of cuisines, pricing, and service options (online delivery, table booking) across different regions.
* **Compare expenditure trends** to optimize financial planning and maintain profitability in new markets.
* **Assess competition** by analysing top-performing and low-rated restaurants in targeted locations.
* **Provide insights into delivery and booking trends**, helping management decide whether to invest in these services in new outlets.
* **Create dynamic reports and dashboards** using slicers, pivot tables, charts, and lookup formulas to support scenario-based discussions and long-term planning.

Ultimately, this project aims to turn raw data into actionable insights, enabling Zomato to make informed, data-driven decisions for sustainable business growth.

**Problem Statement**

Zomato has appointed you as a data analyst to assist in its expansion strategy. Your task is to analyse the current dataset and suggest new locations (countries, states, cities) where opening restaurants would be strategically beneficial. The management is particularly interested in understanding market gaps, customer preferences, pricing strategies, and the competitive landscape in various regions. Your insights, supported by Excel dashboards, visualizations, and metrics, will form the basis for Zomato's decision-making process in restaurant expansion.

**Data Overview**

The dataset provided for this project contains detailed information about restaurants listed on Zomato across various countries and cities. Each row in the dataset represents a unique restaurant and includes multiple attributes related to its identity, location, service offerings, pricing, and customer feedback. Below is a brief description of each field in the dataset:

* **Restaurant ID**: A unique identifier assigned to each restaurant.
* **Restaurant Name**: The name or brand under which the restaurant operates.
* **CountryCode**: Numerical code representing the country in which the restaurant is located.
* **City**: The specific city where the restaurant is situated.
* **Address**: The detailed physical address of the restaurant.
* **Locality**: The neighbourhood or area within the city where the restaurant is found.
* **Locality Verbose**: A more detailed description of the locality, often including nearby landmarks or zones.
* **Longitude**: The geographic longitude coordinate of the restaurant’s location.
* **Latitude**: The geographic latitude coordinate of the restaurant’s location.
* **Cuisines**: A list of cuisines served by the restaurant (e.g., Indian, Chinese, Italian).
* **Currency**: The type of currency accepted at the restaurant for transactions.
* **Has\_Table\_booking**: Indicates whether the restaurant provides a table reservation option (Yes/No).
* **Has\_Online\_delivery**: Specifies if the restaurant offers online delivery services (Yes/No).
* **Is\_delivering\_now**: Shows whether the restaurant is currently accepting delivery orders (Yes/No).
* **Switch\_to\_order\_menu**: Indicates if users can switch directly to the online order menu (Yes/No).
* **Price\_range**: A numeric category (typically from 1 to 4) indicating the pricing tier of the restaurant.
* **Votes**: The total number of customer votes or ratings the restaurant has received.
* **Average\_Cost\_for\_two**: The average expenditure for two people dining together at the restaurant.
* **Rating**: The average customer rating based on reviews and feedback (usually on a scale of 1 to 5).
* **Datekey\_opening**: The date on which the restaurant was officially opened.

This structured dataset serves as the foundation for the analysis, helping uncover trends and insights critical for Zomato’s expansion decisions.

**Methodology**

The project methodology involves a step-by-step approach to transform raw data into actionable business insights using Microsoft Excel. The process is outlined as follows:

**1. Data Cleaning and Preparation**

* Removed duplicate records and handled missing values to ensure data consistency.
* Standardized inconsistent entries (e.g., “Yes/No” formatting, cuisine types).
* Used VLOOKUP to map CountryCode to actual country names.
* Verified geographic coordinates and ensured city names matched country entries.

**2. Data Transformation**

* Converted appropriate columns to categorical (e.g., Cuisines, Price\_range) or numerical formats (e.g., Votes, Rating).
* Created derived columns such as:
  + **Year of Opening** from Datekey\_opening
  + **Custom Price Tag** combining currency symbols with average cost.
* Used string functions to analyze cuisines and identify patterns.

**3. Exploratory Data Analysis (EDA)**

* Used pivot tables, slicers, and filters to:
  + Calculate distribution of restaurants across countries and cities.
  + Identify average cost, rating, and votes per country or city.
  + Compare features like table booking and delivery options.

**4. Dashboard Development**

* Designed a dynamic dashboard using charts, pivot tables, and slicers for:
  + Country-wise and year-wise trends.
  + Rating vs. pricing analysis.
  + Popular cuisines and delivery insights.
  + Competition mapping by rating and cost.
* Incorporated filters for year and country to make the dashboard interactive.

**5. Strategic Evaluation**

* Applied logical and statistical Excel functions (e.g., IF, AVERAGEIF, COUNTIFS, TEXT, LEFT, RIGHT) to answer both objective and subjective business questions.
* Highlighted recommended areas using conditional formatting and visual cues.

**Analysis of Objective Questions & Answers**

**1.** The total no. of tables present in the given data is 2 in two different sheets,

One is “Raw Data”

Another one is “country description”,

After analyse the data, it has 11 sheets with various tables which I used for getting answers of given subjective and objective questions.

Rest of 9 sheets name below—

|  |  |
| --- | --- |
| **Sheets Names** | **Consist Of** |
| Working Data | 26 Columns needed for Analysis |
| Dashboard | Created Dashboard |
| Objective\_Answers | Needed Pivot, Calculation and chart to answer given Objective Question |
| Sub\_Q1 | Needed Pivot, Calculation and chart to answer Subjective Q-1 |
| Sub\_Q2 | Needed Pivot, Calculation and chart to answer Subjective Q-2 |
| Sub\_Q3&4 | Needed Pivot, Calculation and chart to answer Subjective Q-3 & Q-4 |
| Sub\_Q5&6 | Needed Pivot, Calculation and chart to answer Subjective Q-5 & Q-6 |
| Sub\_Q7,8&9 | Needed Pivot, Calculation and chart to answer Subjective Q-7, Q-8 & Q-9 |
| Pivot Table and Chart | Needed Pivot, Calculation and chart to Visualize charts in dashboard |

**2.** The total no. of attributes present in the given data is 20.

**3. Categorical Variables:**

* Restaurant ID (identifier)
* Restaurant Name
* CountryCode / Country Name
* City
* Address
* Locality
* Locality Verbose
* Cuisines
* Currency
* Has\_Table\_booking (Yes/No)
* Has\_Online\_delivery (Yes/No)
* Is\_delivering\_now (Yes/No)
* Switch\_to\_order\_menu (Yes/No)
* Price\_range (ordinal category: 1 to 4)
* Datekey\_opening (can be grouped by Year/Month for analysis)

**Continuous Variables:**

* Longitude
* Latitude
* Votes
* Average\_Cost\_for\_two
* Rating

**4.** Step-1: 1st copied the ‘Raw Data’ into sheet named ‘Working Data’ for further analysis.

Step-2: Find and Replace Missing values:

* Select the entire dataset (e.g., click the top-left cell and press Ctrl + Shift + → then Ctrl + Shift + ↓).
* Press Ctrl + G (opens "Go To" dialog box).
* Click on Special button.
* Choose Blanks and click OK
* This will highlight all blank cells in the selected range.
* Without clicking anywhere else, press = then the VALUE (“Unknown”) to reference the value above.
* Press Ctrl + Enter

➤ This fills all blank cells with the value above.

Step-3: Checked duplicates through the whole dataset. **“No Duplicates Found”.**

Step-4: “Datekey\_Opening” Column in data contains “2013\_9\_21” format date instead of Proper Date Format.

Using **text to columns from data** and **delimiter [\_] separated** created a new column as “Datekey\_Opening” in “21-09-2013” Date format using [ **=DATE ()**] function and also created different **Year, Month & Day** column for date.

**5.** Filled up the countries in the original data using the country code from country description data using **VLOOKUP ().** This is the formula used**— [=VLOOKUP (C2,'country description'!$A$1:$B$16,2,FALSE)]**

**6.**

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

|  |  |
| --- | --- |
| **Row Labels** | **Count of RestaurantID** |
| 2010 | 1080 |
| 2011 | 1098 |
| 2012 | 1022 |
| 2013 | 1061 |
| 2014 | 1051 |
| 2015 | 1024 |
| 2016 | 1027 |
| 2017 | 1086 |
| 2018 | 1102 |

**7.**

**8.**The total number of restaurants in India in the price range of 4 is **388.**

|  |  |
| --- | --- |
| **Conditons** | **Number of Restaurant** |
| Price\_range = 4 | 388 |
| Country = India |

**Used Formula -- =COUNTIFS('Working Data'!D:D,"India",'Working Data'!Q:Q,4)**

**9.**

|  |  |
| --- | --- |
| **Row Labels** | **Average of Votes** |
| Indonesia | 772.0952381 |
| United Arab Emirates | 493.5166667 |
| Turkey | 431.4705882 |
| United States of America | 428.2211982 |
| Philippines | 407.4090909 |
| South Africa | 315.1666667 |
| New Zealand | 243.025 |
| United Kingdom | 205.4875 |
| Qatar | 163.8 |
| Sri Lanka | 146.45 |
| India | 137.212552 |
| Australia | 111.4166667 |
| Canada | 103 |
| Singapore | 31.9 |
| Brazil | 19.61666667 |

**10.** The average rating for all the restaurants that have price\_range < 4 and provide online delivery is 3.27381151.

|  |  |
| --- | --- |
| **Conditions** | **Average Rating** |
| Price\_range<4 | 3.27381151 |
| Has\_Online\_delivery = "Yes" |

Used only the “IF” function, Logical Operators (\*), and Aggregation functions (Average) to solve this problem. [Note: Didn’t use Conditional aggregation (AVERAGEIF) in this question.]

And **array formula**, I solved this problem.

**Used Formula -- {=AVERAGE (IF (('Working Data’! $Q$2:$Q$9552<4)\*('Working Data'!$N$2:$N$9552="Yes"),'Working Data'!$T$2:$T$9552))}**

**11.** The **countries or cities** I've already recommended for expansion:

Countries: *Qatar, Indonesia, Philippines, Turkey, New Zealand.*

Cities: *Tangerang, Bogor, ÛÁstanbul, Jakarta, Pasig City, Pasay City, Makati City, Taguig City, San Juan City, Auckland, Mandaluyong City, Quezon City.*

Go to:

Home > Conditional Formatting > New Rule > Use a formula to determine which cells to format

Using **Conditional formatting** with a custom formula to highlight rows that match these recommended locations with fill YELLOW colour in “Working Data” Sheet.

**The formula is:**

=OR ($D2="Indonesia", $D2="New Zealand", $D2="Philippines", $D2="Qatar", $D2="Turkey", $E2="Tangerang", $E2="Bogor", $E2="ÛÁstanbul", $E2="Jakarta", $E2="Pasig City”, $E2="Pasay City", $E2="Makati City", $E2="Taguig City”, $E2="San Juan City", $E2="Auckland", $E2="Mandaluyong City", $E2="Quezon City")

**12.** Created a new customized price column that consists of the abbreviation/symbol of the currency along with the Average\_cost\_for\_two values.

**Formula—=CONCATENATE (VLOOKUP (L2,'country description'!$C$2:$E$13,3,FALSE),S2)**

Also create another column called **“Average\_Cost\_for\_two\_INR”** for getting all the currency values in **INR rate.**

**13.**

|  |  |
| --- | --- |
| **Conditions** | **Number of Restaurant** |
| Has\_Online\_delivery = "No" | 1694 |
| Price\_range = Min(Price\_range) |
| Average\_Cost\_for\_two<= Rs. 250 |

**Formula- [** {**=COUNTIFS ('Working Data'!$N$2:$N$9552,"No",'Working Data'!$Q$2:$Q$9552,MIN('Working Data'!$Q$2:$Q$9552),'Working Data'!$Z$2:$Z$9552,"<=250")}]**

**Insights from Subjective Questions**

**1.** To identify countries with lower restaurant competition:

To identify countries with lesser competition but active customer engagement, I considered the following columns:

* Country – to group data.
* Restaurant ID – to count restaurants (measure of competition).
* Votes – to measure user engagement/popularity.
* Aggregate Rating – to assess customer satisfaction.
* **Criteria used**:

At first, Selected countries where the number of restaurants is **less than 50**.

Then, consider those countries with average votes should be **greater than 150 with decent rating (greater than 3.5).**

These thresholds were chosen manually to define low competition zones **but high customer engagement**, suggesting **untapped markets** with active users.

|  |  |  |  |
| --- | --- | --- | --- |
| **Row Labels** | **Count of RestaurantID** | **Average of Votes** | **Average of Rating** |
| Indonesia | 21 | 772.0952381 | 4.30 |
| Turkey | 34 | 431.4705882 | 4.30 |
| Philippines | 22 | 407.4090909 | 4.47 |
| New Zealand | 40 | 243.025 | 4.26 |
| Qatar | 20 | 163.8 | 4.06 |
| Sri Lanka | 20 | 146.45 | 3.87 |
| Australia | 24 | 111.4166667 | 3.66 |
| Canada | 4 | 103 | 3.58 |
| Singapore | 20 | 31.9 | 3.58 |

**Visualization Used**

* **Bar Chart:** Number of restaurants per country.
* **Scatter Plot:** X-axis = Number of restaurants, Y-axis = Average votes.

From the chart:

* Countries like **Indonesia, New Zealand**, **Qatar**, **Philippines**, and **Turkey** show low restaurant density.
* These countries still have growing urban populations and rising food delivery demand.

Based on the above analysis, the team should consider expanding into:

These above mention countries currently have fewer restaurants but **moderately high user votes and ratings**, indicating **low competition and high customer interest**. These countries could be strong opportunities for restaurant expansion.

**2.** To identify suitable **states and cities**, I followed these steps:

* Filtered the dataset by each of the 5 suggested countries.
* Created a **Pivot Table** in Excel showing the number of restaurants also the total votes by city.

Criteria used: Selected cities where Average of Votes is greater than 290. This threshold was chosen manually to define supports demand zones.

* Used a Bar Chart-Line Chart combo to highlight cities with:
  + A moderate number of restaurants (not zero, indicating some demand).
  + A population size that supports demand but is not overly saturated.

|  |  |  |
| --- | --- | --- |
| Country | (Multiple Items) |  |
|  |  |  |
| **Row Labels** | **Average of Votes** | **Count of RestaurantID** |
| Tangerang | 1183.5 | 2 |
| Bogor | 971 | 2 |
| ÛÁstanbul | 860.2142857 | 14 |
| Jakarta | 742.6875 | 16 |
| Pasig City | 696.6666667 | 3 |
| Pasay City | 606 | 3 |
| Makati City | 452.5 | 2 |
| Taguig City | 415.75 | 4 |
| San Juan City | 340.5 | 2 |
| Auckland | 339.75 | 20 |
| Mandaluyong City | 300 | 4 |
| Quezon City | 294 | 1 |

Based on this analysis:

* Zomato can expand into **mid-to-large cities** in each country that are **less saturated** yet have:
  + Strong demand drivers (population, tourism, economy)
  + Urban infrastructure
  + Growth potential in food services

Suggested cities in the table to prioritize. These are strategically positioned cities where competition is not very high and economic indicators support future restaurant growth.

**3.** To evaluate the **quality** of restaurants in the suggested countries (Indonesia, New Zealand, Qatar, Philippines, Turkey), I did the following:

* Created a Pivot Table in Excel showing the number of restaurants by country.
* **In rows-Country, in columns- Rating, in values- Count of RestaurantID.**
* **Filtered** the dataset for only these 5 countries.
* **Grouped** data by Average Rating → Calculated the **Count of Restaurants** in those rating group. [1-3, 3-5]

Created a **Bar Chart** to visually compare average ratings.

* **Criteria Used**:
  + Consider only restaurants that have a non-zero, non-null rating.
  + Focused on countries with an average rating of **3 or above**, indicating overall good quality.

|  |  |  |
| --- | --- | --- |
| **Count of RestaurantID** | **Column Labels** |  |
| **Row Labels** | **1-3** | **3-5** |
| Turkey |  | 34 |
| Qatar |  | 20 |
| Philippines |  | 22 |
| New Zealand | 1 | 39 |
| Indonesia |  | 21 |

**4.** Determine the total financial expenditure required to expand a restaurant chain across multiple countries based on their average cost for two.

**Criteria Used**:

* **Divider**: Since there is no total financial expenditure for countries present in the given data, so I’m set a **divider 100,** where if I’ve average cost for two for suggested countries I can get an approximate financial expenditure based on that formula.
* **Formula:** =
* **What-If Analysis – Goal Seek**: Applied in Excel to find how much financial expenditure would be needed so that the average cost for two in a given country scales to match the respective benchmark.

**Goal Seek Table:** This table shows the **"Average Cost for Two (INR)"** by **Country** and the **Financial Expenditure** derived using goal seek.

|  |  |
| --- | --- |
| Set\_Divider | 100 |
| Financial\_Expenditure | ₹ 2,00,000.00 |
| Average\_Cost\_for\_two\_INR | ₹ 2,000.00 |

|  |  |
| --- | --- |
| **Row Labels** | **Average of Average\_Cost\_for\_two\_INR** |
| Indonesia | ₹ 1,630.90 |
| New Zealand | ₹ 3,487.50 |
| Philippines | ₹ 9,801.59 |
| Qatar | ₹ 5,314.06 |
| Turkey | ₹ 381.84 |

**Highest Costs**: Philippines (₹9,801.59) require the highest investment to scale down to the target cost level, indicating expensive dining experiences in these countries.

**Lowest Costs**: Turkey (₹381.84) show the lowest average costs, implying more cost-effective expansion opportunities.

**5.** Identify **top competitors** (high-performing restaurants) and **low-rated** restaurants (ratings between 1–3) from the **recommended states**.

**➤ Criteria:**

1. **Filter for recommended cities or states only** (based on prior analysis).
2. **Segment restaurants based on rating brackets:**
   * **Top competitors** → Top 10 High Sum of votes.
   * **Low-rated** → Ratings between 1.0–3.0.
3. Use **Pivot Table** to get name of those restaurants.

|  |  |
| --- | --- |
| City | (Multiple Items) |
|  |  |
| **Row Labels** | **Average of Votes** |
| Aôôk Kahve | 901 |
| Ceviz AÛôacÛ± | 1034 |
| J'adore Chocolatier | 1311 |
| Karakí\_y Gí\_llí\_oÛôlu | 1305 |
| Lemongrass | 1159 |
| Skye | 1498 |
| Starbucks | 1042 |
| Talaga Sampireun | 1838 |
| Toodz House | 1476 |
| Union Deli | 903 |

|  |  |
| --- | --- |
| City | (Multiple Items) |
|  |  |
| **Row Labels** | **Average of Rating** |
| De Fontein Belgian Beer Cafe | 2.3 |

**6.** Identify cuisines that receive better feedback (in terms of ratings and votes) so they can be prioritized in future restaurant expansions.

**Criteria Used:**

* Considered only cuisines with **high number of votes** (Top 10 by Sum of Votes) as a proxy for popularity.
* Evaluated **average restaurant rating** to assess customer satisfaction.
* Focused on cuisines from **suggested cities/countries** (as per mentor's instruction).
* Excluded any cuisines with too few votes or inconsistent rating data.
* **Method:**
* Created a pivot table grouping by cuisine (or cuisine combinations).
* Used:
  + **Sum of Votes** = Measure of popularity
  + **Average of Rating** = Measure of feedback quality
* Sorted results by Sum of Votes and reviewed Top 10.

|  |  |  |
| --- | --- | --- |
| Country | (Multiple Items) |  |
|  |  |  |
| **Row Labels** | **Sum of Votes** | **Average of Rating** |
| Sunda, Indonesian | 5514 | 4.9 |
| Cafe | 3405 | 4.281818182 |
| Desserts | 2346 | 4.833333333 |
| Italian, Continental | 1498 | 4.1 |
| Cafe, Italian, Coffee and Tea, Western, Indonesian | 1476 | 4.6 |
| Restaurant Cafe | 1407 | 3.85 |
| Filipino, Mexican | 1364 | 4.85 |
| Desserts, Bí\_rek | 1305 | 4.7 |
| Filipino | 1201 | 4.5 |
| Peranakan, Indonesian | 1159 | 4 |

* **Sunda, Indonesian** cuisine had the **highest votes (5514)** and an **outstanding rating of 4.9**, making it a prime candidate for expansion.
* **Desserts** and **Filipino, Mexican and Desserts** cuisines also performed very well, with average ratings above **4.8**, showing strong customer preference.
* Combinations like **Cafe, Italian, Coffee and Tea, Western, Indonesian** also scored highly, suggesting **multi-cuisine hybrid menus** may appeal more to customers.
* Cuisines such as **Restaurant Cafe** showed high votes but relatively lower ratings (**3.85**), indicating that **popularity doesn't always mean satisfaction**.
* Overall, cuisines with both **high votes and high ratings** indicate areas where demand and satisfaction overlap — these should be prioritized.

Yes, the **choice of cuisine moderately affects restaurant ratings**.  
For **new restaurant launches**, we should focus on those mention in the pivot table cuisines have **proven appeal and high satisfaction** based on the data.

**7.** Built a **PivotTable** where:

* + **Rows**: Has\_Online\_delivery
  + **Columns**: Has\_Table\_booking
  + **Values**: Average of Rating

This helps us compare all four combinations:

* Neither service
* Only table booking
* Only online delivery
* Both services

This structured approach allows us to determine the impact of these two features on customer satisfaction.

**Clustered Column Chart** with:

* X-axis: Table Booking (No, Yes)
* Series: Online Delivery (No, Yes)

|  |  |  |
| --- | --- | --- |
| **Average of Rating** | **Column Labels** |  |
| **Row Labels** | **No** | **Yes** |
| No | 2.679786734 | 3.220585317 |
| Yes | 3.411618257 | 3.60045977 |

* Y-axis: Average Rating

From the PivotTable and the Clustered Column Chart, we observe:

* Both services individually have a **positive impact** on customer ratings.
* The **combination** of online delivery and table booking results in the **highest customer satisfaction**.
* Therefore, we **should implement both** services to enhance customer experience and drive better feedback.

**8.** Evaluate whether **increasing cuisine prices** (Average\_Cost\_For\_Two) leads to **better feedback** (restaurant ratings), and if there is a **statistical correlation** between pricing and feedback.

* Used **data** for both **Rating** and **Average\_Cost\_For\_Two**.
* Calculated the **Pearson correlation coefficient** between these two numeric variables.
* Created a **scatter plot** to visually validate the correlation.
* The **correlation coefficient is 0.31**, indicating a **moderate positive correlation** between price and feedback. As prices increase, ratings tend to increase as well — though not strongly.
* This suggests that **higher-priced cuisines tend to receive better ratings**, possibly due to better quality, service, or experience.
* However, a large portion of restaurants are still in the lower price tiers, so **customer sensitivity to pricing** should be considered in each region.

**Conclusion:**

* **Yes**, the team should consider **keeping the cuisine rates slightly higher**, especially if the experience justifies it.
* A **moderate positive correlation (0.31)** between price and rating suggests that **well-priced, premium offerings may lead to better feedback**.
* That said, pricing must still align with customer expectations — not too high to reduce footfall, but **high enough to signal quality and drive better feedback**.

**9.** To analyze how restaurants are distributed across different **price ranges (1 to 4)** across **all countries**.

1. Inserted a Pivot Table based on the dataset.
2. Used **Price\_range** as the Legend (category) field.
3. Used **Count of RestaurantID** as the value to aggregate.
4. Inserted a **Pie Chart** to clearly show the share of each price range across all countries.

|  |  |
| --- | --- |
| **Row Labels** | **Count of RestaurantID** |
| 1 | 46.53% |
| 2 | 32.59% |
| 3 | 14.74% |
| 4 | 6.14% |
| **Grand Total** | **100.00%** |

* A significant majority of restaurants (**46.53%**) fall into **Price Range 1**, indicating a **global leaning towards affordable dining options**.
* **Price Range 2** also holds a substantial portion (**32.59%**), suggesting mid-range pricing is common.
* The number of **high-end restaurants (Price Ranges 3 and 4)** is relatively small — only **14.74% and 6.14%**, respectively.
* This suggests that **most restaurants in the dataset cater to low- and mid-income customers**, with luxury dining forming a small niche.
* For businesses and stakeholders, this indicates **high demand and competition in the budget and mid-level segments**, whereas premium markets might offer **opportunities for niche positioning**.

**Strategic Expansion Insights (Zomato Data Analysis)**

To drive smart expansion, we identified 9 countries—including Canada, Singapore, and Turkey—with fewer than 50 restaurants, signalling low competition and high growth potential. Within these countries, cities like Jakarta, Auckland, and Istanbul showed strong demand (avg. votes > 290) but limited restaurant presence, making them ideal for pilot launches.

Quality analysis reveals countries like Australia, Indonesia, and Philippines have restaurants rated 3+, ensuring service standards can be maintained. Financially, countries like Turkey and Sri Lanka offer affordable expansion opportunities, while Singapore and Philippines may require higher initial investment due to higher dining costs.

We identified top competitors like Skye and Starbucks for benchmarking and flagged low-rated players to avoid poor formats. Top cuisines like Sunda (Indonesian), Desserts, and Filipino received high ratings and votes—indicating strong customer preference.

Service-wise, restaurants offering both online delivery and table booking earned the highest ratings, suggesting both should be offered in new markets. A moderate positive correlation (0.31) between pricing and ratings suggests premium pricing is effective if backed by quality. Lastly, with 79% of restaurants in price ranges 1 and 2, mid-budget offerings should be the core focus, with select upscale options for niche appeal.

**10.**  **Approach to Suggest Countries/Cities for New Restaurants**

* **Understand the Dataset Structure:**
  + Analyze the columns to identify key metrics: location (country, city), price range, online delivery, ratings, number of votes, cuisine, etc.
* **Data Cleaning & Preparation:**
  + Remove duplicates and handle missing or inconsistent values.
  + Standardize categorical values (like city names, price levels).
  + Use lookup to fill in missing data (e.g., country names via country codes).
* **Explore Restaurant Distribution:**
  + Create pivot tables to show number of restaurants per country and city.
  + Identify underrepresented countries/cities with relatively fewer restaurants.
* **Analyze Competition Level:**
  + Define a "competition score" based on the number of restaurants in an area.
  + Prefer locations with low competition but with moderate customer engagement (votes).
* **Customer Feedback & Ratings:**
  + Analyze average ratings per city and country.
  + Prefer locations where existing restaurants have high ratings (indicating good customer base and food culture).
* **Time-Based Performance Analysis:**
  + Calculate how many years it typically takes for restaurants in different cities to achieve good ratings or high vote counts.
  + Use year-wise data to track rating growth or customer engagement over time.
  + Prefer cities where restaurants show quicker performance improvements—indicating faster market acceptance and better potential for new entrants.
* **Economic Feasibility:**
  + Analyze average cost for two across countries.
  + Select locations with lower food expenditure for budget-friendly strategy or higher expenditure areas for premium strategy.
* **Cuisine Analysis:**
  + Check most popular cuisines by location.
  + Select countries where popular cuisines align with what we plan to offer.
* **Delivery & Booking Trends:**
  + Analyze if online delivery and table booking features are popular in certain countries.
  + Use these features as strategic advantages in those regions.
* **Create Visual Insights:**
  + Use charts and dashboards (heatmaps, bar charts) to visualize:
    - Restaurant count
    - Average ratings
    - Price range distribution
    - Cuisine popularity
* **Filter Using Slicers:**
  + Use slicers to interactively filter by Year and Country to observe trends.
* **Shortlist Ideal Locations:**
  + Based on low competition, decent rating culture, affordable average cost, and popular cuisines.