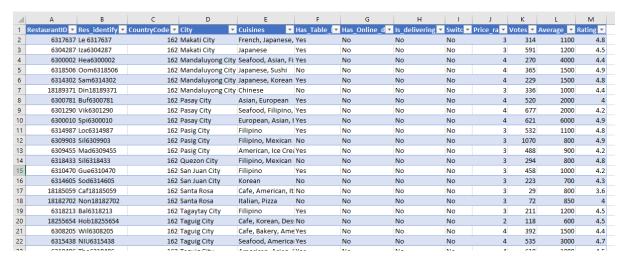
Zomato Restaurant Data Analysis using SQL (MySQL)

The Zomato Restaurant Data Analysis initiative is dedicated to conducting an in-depth exploration and examination of restaurant data sourced from Zomato, a prominent online platform for food delivery and restaurant discovery. By harnessing the power of SQL for data manipulation and querying, this endeavour strives to unveil valuable insights that can guide strategic business decisions and foster a deeper comprehension of the restaurant landscape within a specific geographic area. Through meticulous analysis of this comprehensive dataset, the project aims to uncover patterns, trends, and opportunities that can potentially shape the future course of the industry and enhance the overall dining experience for consumers.

Datasets: This data analysis project utilizes two datasets.

✓ **Zomato** – Containing comprehensive details of restaurants affiliated with Zomato. Containing 9551 rows and 13 columns: RestaurantID, Res_identify, CountryCode, City, Cuisines, Has_Table_booking, Has_Online_delivery, Is_delivering_now, Switch_to_order_menu, Price_range, Votes, Average_Cost_for_two, Rating



✓ **CountryTable** – containing 15 rows and 2 columns: CountryCode, Country

4	Α		В		
1	CountryCode	¥	Country	¥	
2		1	India		
3	21	4	UAE		
4	21	5	UK		
5	21	6	USA		
6	1	4	South Africa		
7	3	0	Malayasia		
8	3	7	Indonesia		
9	9	14	Singapore		
10	14	8	Hongkong		
11	16	2	Nigeria		
12	16	6	France		
13	18	4	Switzerland		
14	18	9	Australia		
15	19	1	New Zealand		
16	20	8	Canada		

Through this comprehensive analysis, the project aims to offer a deeper understanding of the Zomato restaurant ecosystem, enhance user experiences, and optimize business strategies.

While analyzing Zomato restaurant data, the focus was on the following aspects:

- ✓ City Ratings Analysis: Identified cities with poor restaurant ratings to help Zomato target areas for improvement.
- ✓ **Affordable and Highly Rated Restaurants:** Enlisted the most affordable and highly rated restaurants city-wise to assist users in finding value-for-money dining options.
- ✓ **Offline Services Evaluation:** Identified restaurants with poor offline services to enhance customer satisfaction.
- ✓ **Restaurant Categorization:** Grouped restaurants based on average cost for two into categories such as Luxurious Expensive, Very Expensive, Expensive, High, Medium High, and Average, then quantified the number of restaurants in each category.
- ✓ **Top Rated Restaurants:** Listed the top 5 restaurants with the highest rating and maximum votes to highlight premier dining options.
- ✓ High-Rated Cities Identification: Identified cities with at least 3 restaurants having ratings >=
 4.9, sorted alphabetically in case of ties.
- ✓ **Top Countries by Restaurant Count:** Identified the top 5 countries with the most restaurants linked with Zomato.

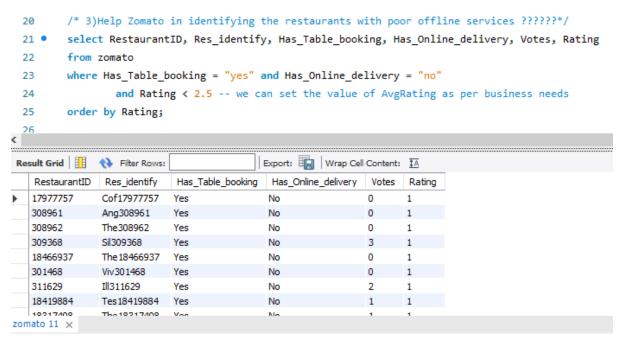
Analysis:

```
5
  6
         -- Business Questions: --
  7
         /* 1) Help Zomato in identifying the cities with poor Restaurant ratings */
  8
         select city, round(avg(Rating),1) as AvgRating
  9
         from zomato
 10
         group by city
 11
         having AvgRating < 2.5; -- we can set the value of AvgRating as per business needs
 12
<
                                        Export: Wrap Cell Content: IA
AvgRating
   city
   Mc Millan
            2.4
   Montville
   Faridabad
   Noida
           2.4
```

** Here, cities with poor restaurant ratings were identified. Initially, the data was grouped by cities and the average restaurant rating for each city was calculated. Subsequently, a condition was applied on the average rating. The 'HAVING' clause was utilized instead of the 'WHERE' clause because the condition was applied to aggregate values derived from the 'GROUP BY' clause.

```
13
         /* 2)Enlist most affordable and highly rated restaurants city wise. */
         select city, RestaurantID, Res_identify, Rating, Average_Cost_for_two
  14 •
      15
         rank() over (partition by city order by Rating desc, Average_Cost_for_two) as rnk
  16
       from zomato) as temp
  17
         where rnk = 1;
  18
<
 Result Grid
             Filter Rows:
                                        Export: Wrap Cell Content: TA
                                            Average_Cost_for_two
    city
              RestaurantID Res_identify
                                      Rating
   Abu Dhabi
              5703500
                         Pun5703500
                                     4.9
                                            330
                                            0
             3400346
                         She3400346
                                     4.9
   Agra
   Ahmedabad
             18385201
                         Cry 18385201
                                     4.6
                                            350
              17284409
   Albany
                         Gua 1728 4409
                                     3.9
                                            10
   Albany
              17284158
                         Jim 17284158
                                     3.9
                                            10
   Allahabad
             2400052
                         Eat2400052
                                     3.7
                                            200
             2200175
                         Gur2200175
                                            100
   Amritsar
                                     4.1
                         Gag6004011
             6004011
                                     4.9
                                            80
   Ankara
                         Whi16611114
   Armidale
              16611114
                                     3.5
                                            20
   Athens
             17293409
                         Sr. 17293409
                                     4.6
                                            10
Result 2 x
```

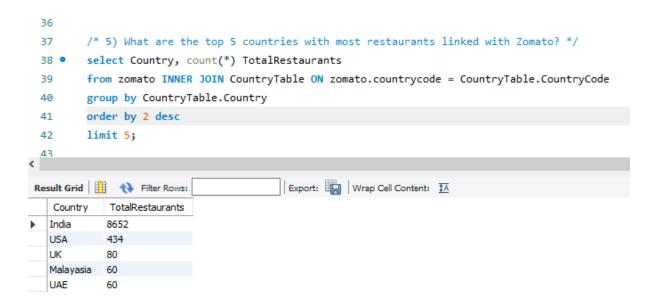
** Here, highly rated and most affordable restaurants for each city were identified. Windows functions were employed for the calculation. The RANK window function grouped the restaurants for each city and ranked them based on high ratings and low costs. This query was utilized as a Common Table Expression (CTE), and based on that another query was written where the condition 'rank=1' was applied to obtain the highly rated and most affordable restaurant for each city.



** For identifying restaurants with poor offline services, three conditions were applied. The restaurant should offer offline table booking services, while online delivery services should be unavailable. Additionally, restaurants with ratings below a specified threshold value were considered to provide poor service. The threshold value of the rating can be adjusted based on business needs.

```
26
    27
       In case there are two cities with the same result, sort them in alphabetical order.*/
28
29 •
      SELECT city,
30
           count(restaurantid) as noOfRestaurants
31
      FROM zomato
      WHERE rating >= 4.9
33
      GROUP BY city
34
      having count(restaurantid) >=3
35
      ORDER BY NoOfRestaurants DESC, city ASC;
Export: Wrap Cell Content: IA
         noOfRestaurants
  city
 Dubai
 Jakarta
 London
 New Delhi
        3
```

** To identify cities with at least 3 restaurants having ratings of 4.9 or higher, the cities and restaurants with ratings greater than or equal to 4.9 were selected. The WHERE clause was utilized for this purpose. Subsequently, the data was grouped by cities using the GROUP BY clause, and the count of restaurants for each city was determined. The HAVING clause was then applied to the aggregate value count to filter cities with at least 3 restaurants with ratings of 4.9 or higher.



** To determine the top 5 countries with the most restaurants linked with Zomato, two tables, Zomato (containing restaurant details) and CountryCode (containing country names), were joined using an INNER JOIN to obtain restaurant data along with country names. Next, the data was grouped by country using the GROUP BY clause, and the count of restaurants associated with Zomato in each country was calculated. The ORDER BY clause was then utilized to reorder the countries in descending order based on the number of restaurants. Finally, the LIMIT clause was applied to retrieve only the top 5 countries with the highest number of restaurants linked with Zomato.

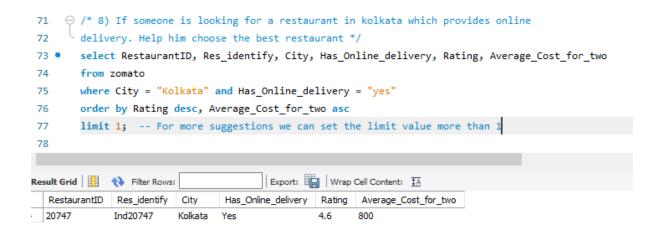
```
\ominus /* 6) Group the restaurants basis the average cost for two into:
45
       Luxurious Expensive, Very Expensive, Expensive, High, Medium High, Average.
       Then, find the number of restaurants in each category.
46
47
48 • ⊖ with tempdf as(
49
                        select RestaurantID,
50
51
                            when Average_Cost_for_two between 0 and 1000 then "Average"
52
                            when Average_Cost_for_two between 1001 and 5000 then "Medium High"
53
                            when Average_Cost_for_two between 5001 and 10000 then "High"
                            when Average_Cost_for_two between 10001 and 50000 then "Expensive"
54
                            when Average_Cost_for_two between 50001 and 100000 then "Very Expensive"
55
                            else "Luxurious Expensive"
56
57
                        end as Status
                        from zomato)
59
60
       select Status, count(*) as `Number of Resturant`
       from tempdf
62
       group by Status
       order by 2 desc;
63
Result Grid Filter Rows:
                                              Export: Wrap Cell (
                         Number of
    Status
                         Resturant
                        8376
    Average
    Medium High
                        1143
                        18
    Luxurious Expensive
                        11
    Very Expensive
                        3
```

** Here, the CASE WHEN clause was initially employed to group restaurants based on the average cost for two persons. The cost range can be adjusted according to business needs to effectively group the restaurants.

Subsequently, this query was utilized as a common table expression to construct another query aimed at determining the total number of restaurants belonging to each category. The GROUP BY clause was utilized to group restaurants by each category, and the count of restaurants within each category was calculated.

```
04
         /* 7) List the two top 5 restaurants with highest rating with maximum votes. */
 65
         select RestaurantID, Res identify, Rating, Votes
 66 •
 67
         from zomato
         order by Rating desc, Votes desc
 68
 69
         limit 5;
 70
                                            Export: Wrap Cell Content: IA
Result Grid
               Filter Rows:
   RestaurantID
               Res_identify
                                     Votes
                             Rating
  20842
               Bar 20842
                             4.9
                                     5966
  94286
               AB'94286
                             4.9
                                     5434
  17806994
               Mir 17806994
                             4.9
                                     3244
               Nat310143
                             4.9
  310143
                                     2620
  17580142
               McG17580142
                             4.9
                                     2238
```

** To identify the top 5 restaurants with the highest ratings and maximum votes, the dataset was reordered using the ORDER BY clause, sorting by rating and votes in descending order. Subsequently, the top 5 rows were extracted using the LIMIT clause.



** Here, an attempt was made to find the best restaurant in Kolkata that provides online delivery. To achieve this, the WHERE clause was utilized to apply conditions such as the city being Kolkata and the restaurant offering online delivery. For the best restaurant, one should consider establishments with high ratings and low costs. Consequently, the ORDER BY clause was employed to reorder the dataset, with ratings sorted in descending order and costs in ascending order. Finally, the LIMIT clause with a value of 1 was applied to identify the top restaurant in Kolkata providing online delivery. The limit value can be adjusted to explore more restaurant.

```
79
        /* 9) Help someone in finding the best rated Restaurant for Pizza in New Delhi. */
        select RestaurantID, Res_identify, City, Cuisines, Rating
 80 •
 81
        from zomato
        where Cuisines like "%pizza%" And city = "New delhi"
 82
        order by Rating desc
 83
        limit 1;
84
Export: Wrap Cell Content: IA
  RestaurantID Res_identify
                                    Cuisines
                                                                    Rating
                          City
 18400736
              Owl18400736
                          New Delhi
                                   Burger, American, Fast Food, Italian, Pizza
                                                                    4.5
```

** In this scenario, an attempt was made to identify the highest-rated restaurant for pizza in New Delhi. The WHERE clause was utilized to apply conditions such as the city being New Delhi and pizza being available in their cuisines. To determine the best-rated restaurant, the ORDER BY clause was used to reorder the dataset, with ratings sorted in descending order. Finally, the LIMIT clause with a value of 1 was applied to pinpoint the top-rated restaurant for pizza in New Delhi. The limit value can be adjusted to explore additional restaurant options.

Results:

- ✓ In the above analysis, cities with poor restaurant ratings and restaurants with poor offline services were identified. This analysis can assist Zomato in taking necessary actions to enhance their services in those areas.
- ✓ Additionally, the most affordable and highly rated restaurants were identified city-wise, and restaurants were categorized into various cost brackets, such as Luxurious Expensive, Very Expensive, Expensive, High, Medium High, and Average. This categorization aids customers in selecting restaurants according to their preferences.
- ✓ Furthermore, the top 5 countries with the highest number of restaurants linked with Zomato were identified. This information enables Zomato to concentrate on countries with a low number of restaurants linked with their platform to expand their business.
- ✓ Moreover, Zomato can assist its customers by providing information about the best restaurants offering online delivery in any city or allowing customers to order specific cuisines from highly rated restaurants in any city.