ZOMATO RESTAURANTS ANALYSIS
SQL QUERIES

Easy

---

1. Update the date format and modify the column

```
UPDATE sheet1
SET Datekey_Opening = REPLACE(Datekey_Opening, '_', '/')WHERE
Datekey_Opening LIKE '%_%';

ALTER TABLE sheet1
MODIFY COLUMN Datekey_Opening DATE;
```

View the updated table

```
SELECT * FROM sheet1;
```

---

2. Count of Restaurants by Cuisine Type

```
SELECT
   SUBSTRING_INDEX(cuisines, ',', 1) AS cuisine_type,
   COUNT(restaurantid) AS num_restaurants
FROM
   sheet1
GROUP BY
   cuisine_type;
```

---

3. Average Cost for Two People by Country

```
SELECT
    sheet2.country_name, AVG(Average_Cost_for_two) AS
    avg_cost_for_two
FROM
    sheet1
INNER JOIN
    sheet2 ON sheet1.country_code = sheet2.countryidGROUP BY
    sheet2.country_name;
```

4. Percentage of Restaurants based on Has_Online_delivery

```
SELECT
    has_online_delivery,
    CONCAT(ROUND(COUNT(Has_Online_delivery) / (SELECT COUNT(*) FROM sheet1) * 100, 1), '%') AS percentageFROM
    sheet1
GROUP BY
    has_online_delivery;
```

5. Percentage of Restaurants based on Has_Table_booking

```
SELECT
    has_table_booking,
    CONCAT(ROUND(COUNT(has_table_booking) / (SELECT COUNT(*) FROM sheet1) * 100, 1), '%') AS percentageFROM
    sheet1
GROUP BY
    has_table_booking;
```

6. Top Cities by Number of Restaurants

```
SELECT
    city,
    COUNT(restaurantid)  AS  num_restaurant
```

```
FROM
    sheet1
GROUP BY
    city ORDER
BY
    num_restaurants DESCLIMIT
10;
```

---

7. Split cuisines into separate columns

```
SELECT
    restaurantname,
    cuisines,
    SUBSTRING_INDEX(cuisines, ',', 1) AS cuisine1,
    SUBSTRING_INDEX(SUBSTRING_INDEX(cuisines, ',', 2), ',', -1) AS cuisine2,
    SUBSTRING_INDEX(SUBSTRING_INDEX(cuisines, ',', 3), ',', -1) AS cuisine3
FROM
    sheet1;
```

Intermediate

1. Count of Restaurants based on City and Country

```
SELECT
    sheet2.country_name,
    sheet1.city,
    COUNT(restaurantid) AS no_of_restaurants
FROM
    sheet1
INNER JOIN
    sheet2 ON sheet1.country_code = sheet2.countryidGROUP BY
    sheet2.country_name,
    sheet1.city;
```

2. Count of Restaurants based on Average Ratings

```
SELECT
   CASE
      WHEN rating <= 2 THEN '0-2'
      WHEN rating <= 3 THEN '2-3'
      WHEN rating <= 4 THEN '3-4'
      WHEN rating <= 5 THEN '4-5'
   END AS rating_range,
   COUNT(restaurantid)
FROM
   sheet1
GROUP BY
   rating_range
ORDER BY
   rating_range;
```

3. Create buckets based on Average Price and count restaurants in each bucket

```
SELECT
   CASE
      WHEN price_range = 1 THEN '0-500'
      WHEN price_range = 2 THEN '500-3000'
```

```
        WHEN price_range = 3 THEN '3000-10000'
        WHEN price_range = 4 THEN '>10000'
    END AS price_range, COUNT(restaurantid)
FROM
    sheet1
GROUP BY
    price_rangeORDER
BY
    price_range;
```

4. Highest-rating restaurants in each country

```
 SELECT
    country_name,
    restaurantname,
    MAX(rating) AS highest_ratingFROM
    sheet1
INNER JOIN
    sheet2 ON sheet1.country_code = sheet2.countryidGROUP BY
    sheet2.country_name;
```

5. Top restaurant with the highest rating and votes from each country

```
SELECT
    country_name,
    restaurantname,
    MAX(rating) AS highest_rating,MAX(votes)
    AS max_votes
FROM
    sheet1
INNER JOIN
    sheet2 ON sheet1.country_code = sheet2.countryidGROUP BY
    country_name
ORDER BY
    max_votes DESCLIMIT 5;
```

Advanced

1. Extract various date components

```
SELECT
   YEAR(Datekey_Opening) AS years,
   MONTH(Datekey_Opening) AS months, DAY(Datekey_opening)
   AS day, MONTHNAME(Datekey_Opening) AS monthname,
   QUARTER(Datekey_Opening) AS quarter,
   CONCAT(YEAR(Datekey_Opening), '-', MONTHNAME(Datekey_Opening)) AS yearmonth,
   WEEKDAY(Datekey_Opening) AS weekday,
   DAYNAME(Datekey_opening) AS dayname,
   CASE
      WHEN MONTHNAME(Datekey_Opening) IN ('January', 'February', 'March') THEN 'Q1'WHEN
      MONTHNAME(Datekey_Opening) IN ('April', 'May', 'June') THEN 'Q2'
      WHEN MONTHNAME(Datekey_Opening) IN ('July', 'August', 'September') THEN 'Q3'ELSE
      'Q4'
```

```
    END AS quarters,
    CASE
        WHEN MONTHNAME(Datekey_Opening) = 'January' THEN 'FM10'
        WHEN MONTHNAME(Datekey_Opening) = 'February' THEN 'FM11'
        WHEN MONTHNAME(Datekey_Opening) = 'March' THEN 'FM12' WHEN
        MONTHNAME(Datekey_Opening) = 'April' THEN 'FM1' WHEN
        MONTHNAME(Datekey_Opening) = 'May' THEN 'FM2' WHEN
        MONTHNAME(Datekey_Opening) = 'June' THEN 'FM3' WHEN
        MONTHNAME(Datekey_Opening) = 'July' THEN 'FM4' WHEN
        MONTHNAME(Datekey_Opening) = 'August' THEN 'FM5'
        WHEN MONTHNAME(Datekey_Opening) = 'September' THEN 'FM6'WHEN
        MONTHNAME(Datekey_Opening) = 'October' THEN 'FM7' WHEN
        MONTHNAME(Datekey_Opening) = 'November' THEN 'FM8' WHEN
        MONTHNAME(Datekey_Opening) = 'December' THEN 'FM9'
    END AS Financial_months,
    CASE
        WHEN MONTHNAME(Datekey_Opening) IN ('January', 'February', 'March') THEN 'Q4'WHEN
        MONTHNAME(Datekey_Opening) IN ('April', 'May', 'June') THEN 'Q1'
        WHEN MONTHNAME(Datekey_Opening) IN ('July', 'August', 'September') THEN 'Q2'ELSE
        'Q3'
    END AS financial_quarters
FROM
    sheet1;
```

---

2. Top 5 restaurants with the most number of votes

```
SELECT
    country_name,
    restaurantname, votes,
    Average_Cost_for_two
FROM
    sheet1
INNER JOIN
    sheet2 ON sheet1.country_code = sheet2.countryidGROUP BY
    sheet2.country_name,
    restaurantname, votes,
    Average_Cost_for_two
ORDER BY
    votes DESCLIMIT
5;
```

3. Numbers of Restaurants opening based on Year, Quarter, Month

```
SELECT
    YEAR(Datekey_Opening) AS year, QUARTER(Datekey_Opening)
    AS quarter, MONTHNAME(Datekey_Opening) AS monthname,
    COUNT(restaurantid) AS no_of_restaurants
FROM
    sheet1
GROUP BY
    YEAR(Datekey_Opening),
    QUARTER(Datekey_Opening),
    MONTHNAME(Datekey_Opening)
ORDER BY
    YEAR(Datekey_Opening),
    QUARTER(Datekey_Opening),
```

```
    MONTHNAME(Datekey_Opening);
```

4. Average Rating by Price Range

```
SELECT
    CASE
        WHEN price_range = 1 THEN '0-500'
        WHEN price_range = 2 THEN '500-3000'
        WHEN price_range = 3 THEN '3000-10000'
        WHEN price_range = 4 THEN '>10000'END
    AS price_range,
    AVG(rating) AS avg_rating
FROM
    sheet1
GROUP BY
    price_rangeORDER
BY
    price_range;
```