

Explotary Analysis on Zomato Dataset



SQL stands for Structured Query Language. It is the standard language to interact with databases. **SQL** is the most important tool, a data analyst uses to manipulate and gain insights from the data. In this project, I will try to process, and analyze the data, for this purpose we will use this [Zomato Dataset](#)

And in this file there are 2 Excel files with name Country & Zomato

For this project, I will be using Microsoft SQL Server Management Studio 18 . You can download it from [here](#). Microsoft SQL Server Management Studio 18 is an open-source relational database management system. Feel free to use any other RDBMS software. So that now we have our software setup, then let's import the dataset into our database.

IMPORTING DATA

To import the data in Microsoft SQL Server Management Studio 18 follow the given steps, as the dataset is in Excel format we have to import the data using *the* SQL import and Export tool provided by Microsoft .

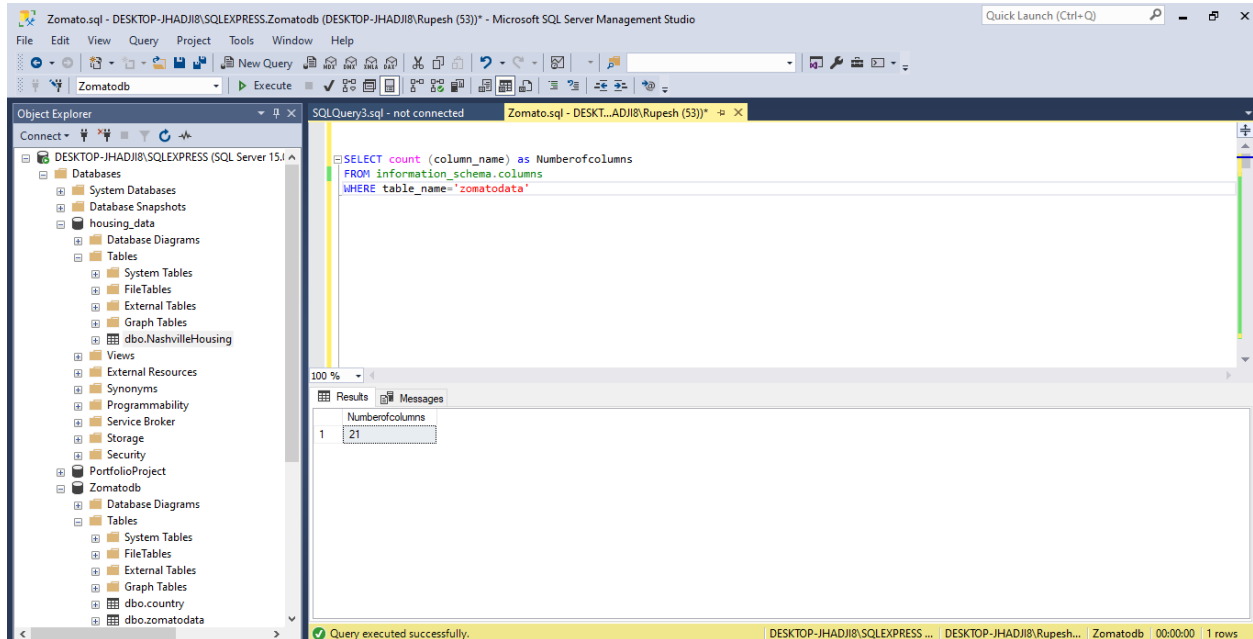
1. Create a database named **zomatodb**.
2. Then open SQL import and Export Server tool then load the Excel File.
3. Finally, click on the Next button to execute the task, and the dataset is successfully imported

Final Summary after EDA:

- 1) There are 15 Countries and 141 Cities where Zomato are Present .
- 2) India has the highest 8652 restaurant which are connected to Zomato.
- 3) There are 4049 restaurant has poor rating less than 3.
- 4) In India (Toit-Restaurant) Has the Highest number of Votes and Excellent Rating given by Customer.
- 5) Average price spent for 2 people on Zomato in india is 623 Rs.
- 6) 5 countries with most restaurant linked with zomato India US,UK,Brazil,South Africa.
- 7) North indian cuisines are most popular in zomato India.
- 8) In this 5 cities zomato must have to improve their food quality and services and those cities are Faridabad,Ghaziabad Gurgaon Newdelhi and Noida.

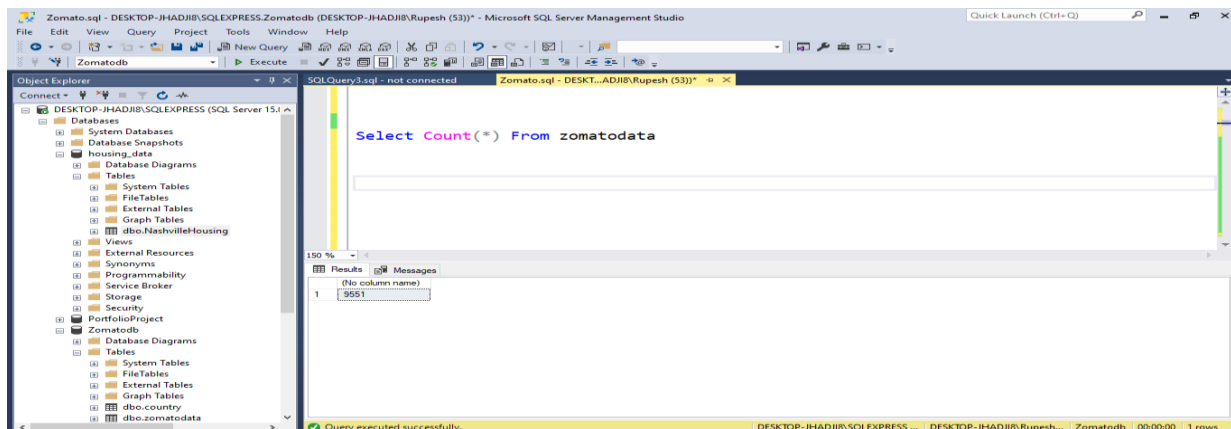
Number of Columns in Zomatodata Table

```
SELECT count (column_name) as Numberofcolumns  
FROM information_schema.columns  
WHERE table_name='zomatodata'
```



Number of rows in Data

```
Select Count(*) From zomatodata
```



Find out number of Distinct Country Code

Select Distinct [Country Code] From zomatodata

The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL code:

```
---Find out number of Distinct Country Code  
Select Distinct [Country Code] From zomatodata
```

The query has been executed successfully, and the results are displayed in the Results pane:

Country Code
14
30
184
166
208
189
216
191
148
215
1

The status bar at the bottom indicates "Query executed successfully." and "15 rows".

Find out List of country Names served by Zomato

select Distinct zomatodata.[Country Code],country.[Country] From
zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
order by [Country Code]

The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL code:

```
---Find out List of country Names served by Zomato  
select Distinct zomatodata.[Country Code],country.[Country] From zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]  
order by [Country Code]
```

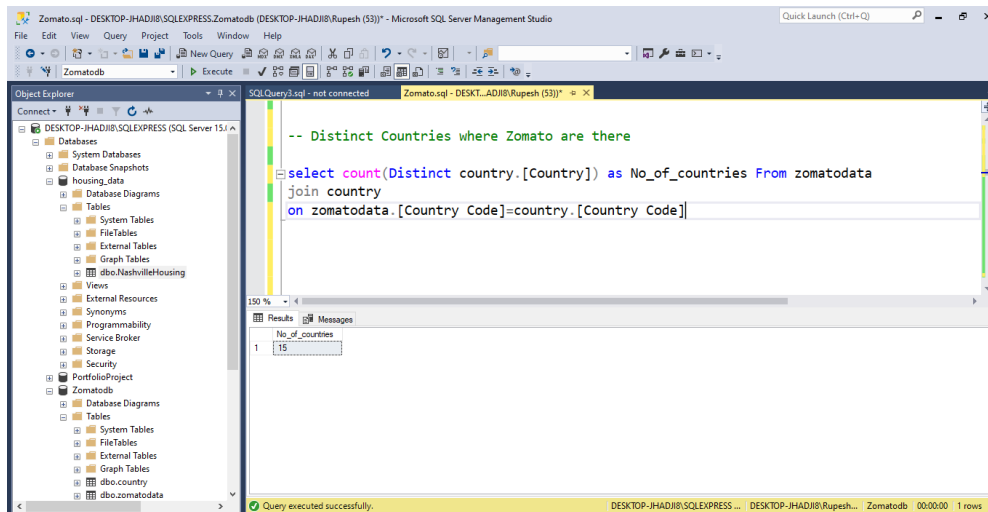
The query has been executed successfully, and the results are displayed in the Results pane:

Country Code	Country
14	India
30	Australia
30	Brazil
37	Canada
94	Indonesia
148	New Zealand
162	Philippines
166	Qatar
184	Singapore
189	South Africa
191	Sri Lanka

The status bar at the bottom indicates "Query executed successfully." and "15 rows".

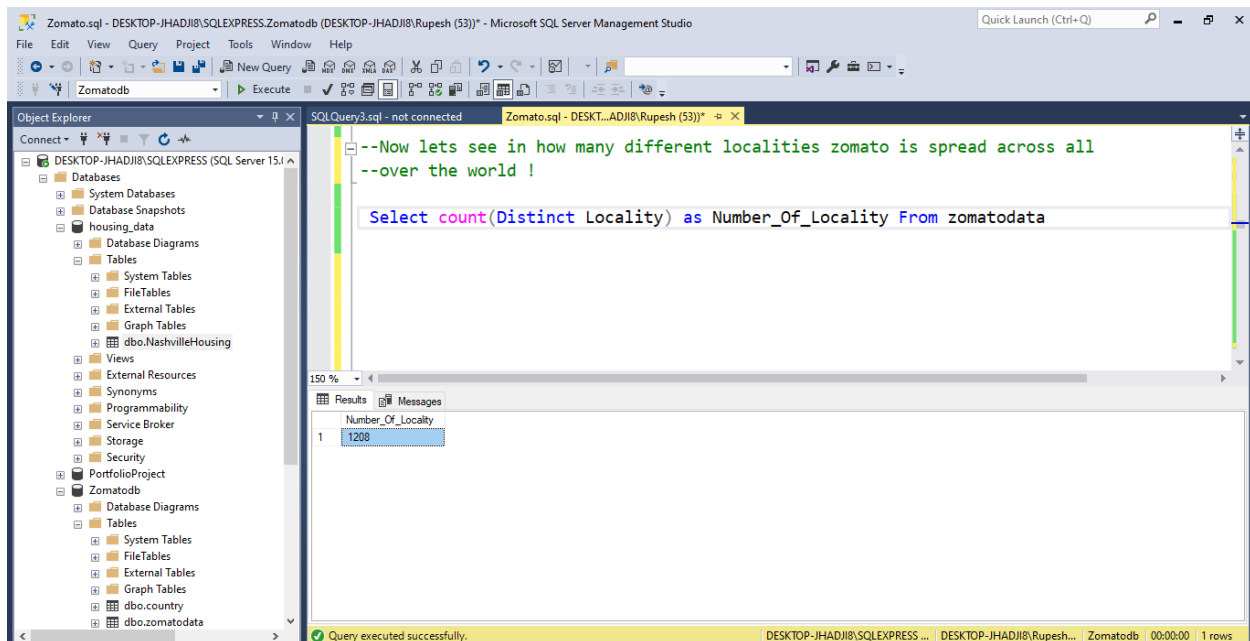
Distinct Countries where Zomato are there

```
select count(Distinct country.[Country]) as No_of_countries From  
zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]
```



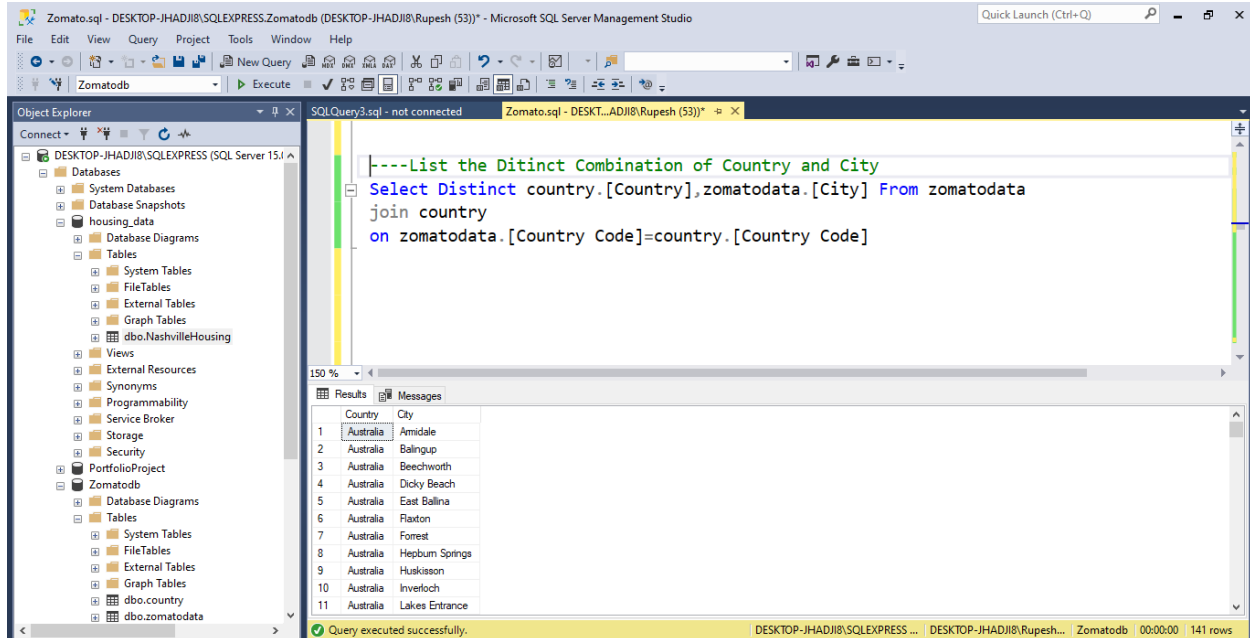
Now lets see in how many different localities zomato is spread across all over the world !

```
Select count(Distinct Locality) as Number_Of_Locality From zomatodata
```



List the Distinct Combination of Country and City

```
Select Distinct country.[Country],zomatodata.[City] From zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor displays the following SQL query:

```
----List the Distinct Combination of Country and City  
Select Distinct country.[Country],zomatodata.[City] From zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]
```

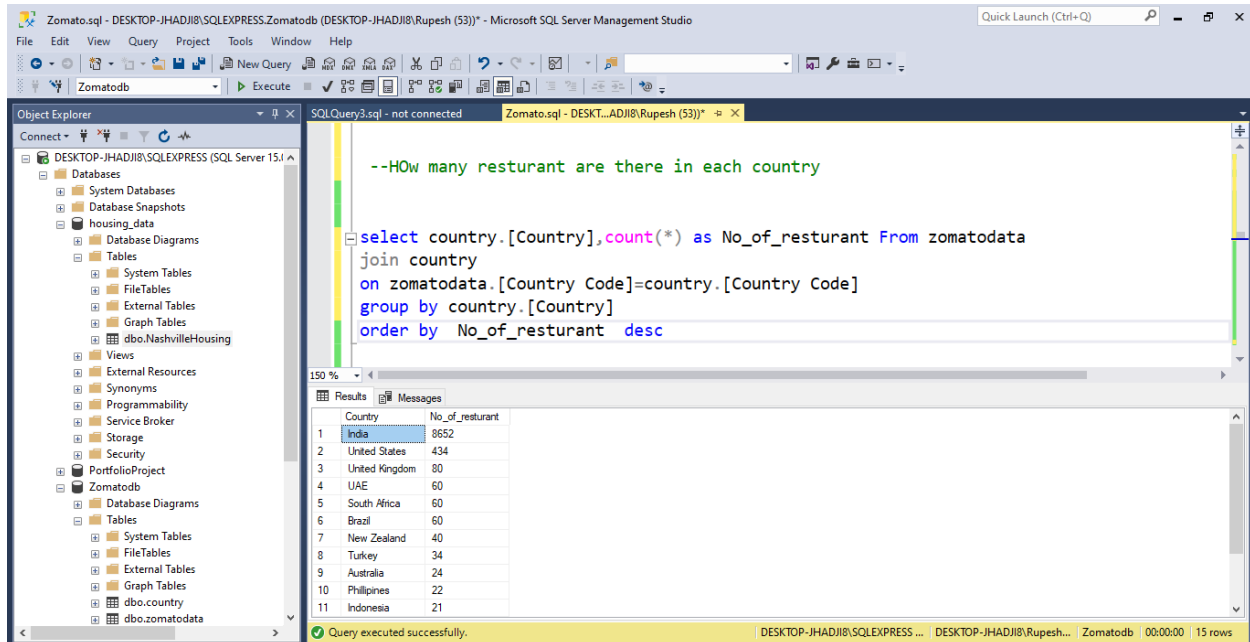
The query has been executed successfully, and the results are displayed in the Results pane. The results show 11 rows of data, all from Australia, with various cities listed.

Country	City
Australia	Amidale
Australia	Balingup
Australia	Beechworth
Australia	Dicky Beach
Australia	East Ballina
Australia	Flaxton
Australia	Forest
Australia	Hepburn Springs
Australia	Huskisson
Australia	Inverloch
Australia	Lakes Entrance

The status bar at the bottom indicates: Query executed successfully. DESKTOP-JHADJ18\SQLEXPRESS ... DESKTOP-JHADJ18\Rupesh ... Zomatodb 00:00:00 141 rows

How many restaurant are there in each country

```
select country.[Country],count(*) as No_of_resturant From zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country]
order by No_of_resturant desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL query:

```
--How many restaurant are there in each country

select country.[Country],count(*) as No_of_resturant From zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country]
order by No_of_resturant desc
```

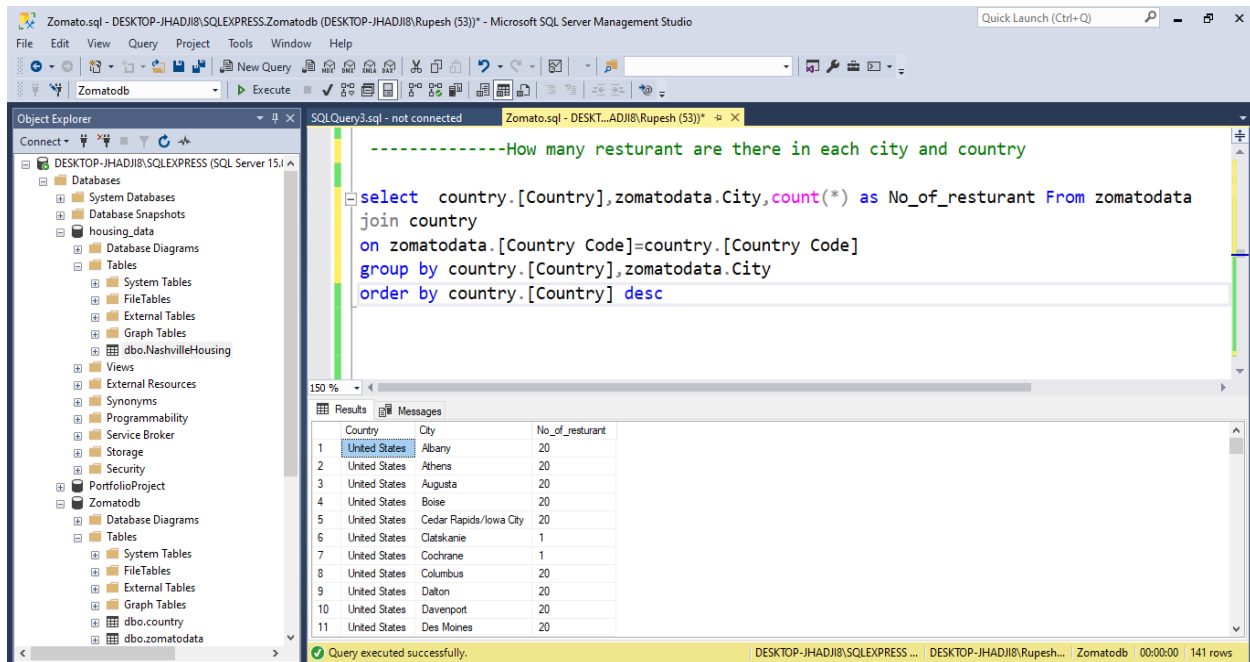
The query results are displayed in the Results pane, showing a table with two columns: Country and No_of_resturant. The results are ordered by the number of restaurants in descending order.

Country	No_of_resturant
India	8652
United States	434
United Kingdom	80
UAE	60
South Africa	60
Brazil	60
New Zealand	40
Turkey	34
Australia	24
Philippines	22
Indonesia	21

The status bar at the bottom indicates: Query executed successfully. 15 rows.

How many restaurant are there in each city and country

```
select country.[Country],zomatodata.City,count(*) as No_of_resturant
From zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country],zomatodata.City
order by country.[Country] desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL query:

```
-----How many restaurant are there in each city and country

select country.[Country],zomatodata.City,count(*) as No_of_resturant From zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country],zomatodata.City
order by country.[Country] desc
```

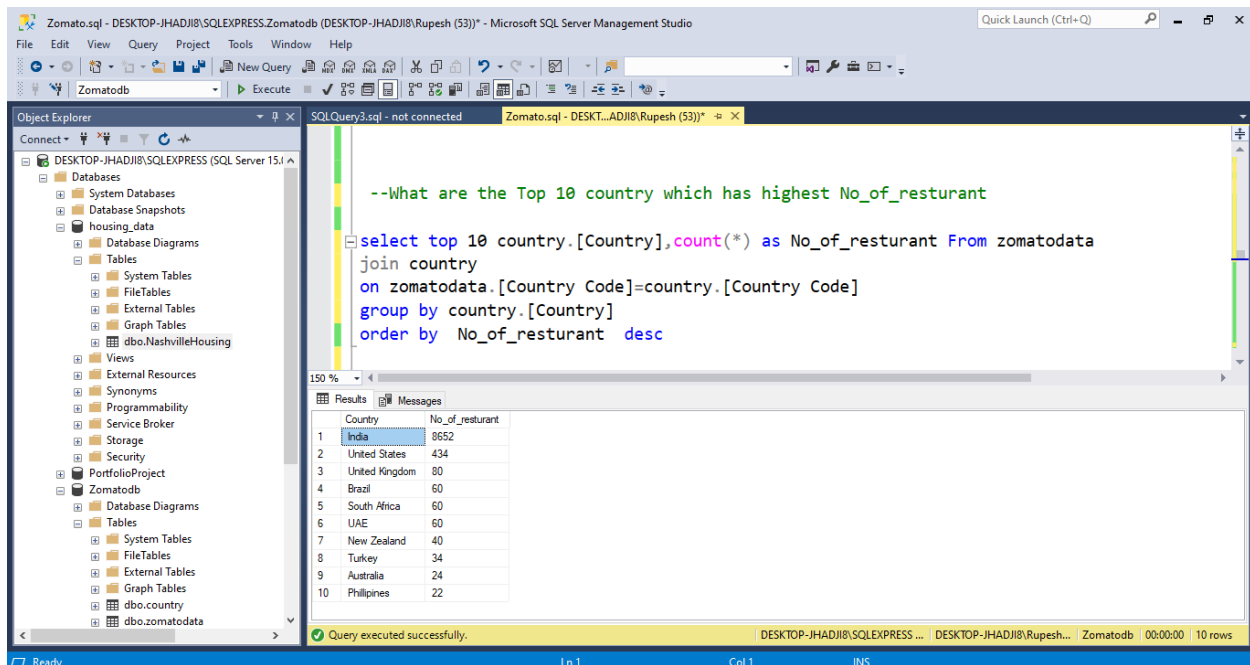
The query results are displayed in the Results pane, showing 11 rows of data:

	Country	City	No_of_resturant
1	United States	Albany	20
2	United States	Athens	20
3	United States	Augusta	20
4	United States	Boise	20
5	United States	Cedar Rapids/Iowa City	20
6	United States	Clatskanie	1
7	United States	Cochrane	1
8	United States	Columbus	20
9	United States	Dalton	20
10	United States	Davenport	20
11	United States	Des Moines	20

The status bar at the bottom indicates: Query executed successfully. DESKTOP-JHADJ18\SQLEXPRESS ... | DESKTOP-JHADJ18\Rupesh... | Zomatodb | 00:00:00 | 141 rows

What are the Top 10 country which has highest No of resturant

```
select top 10 country.[Country],count(*) as No_of_resturant From
zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country]
order by No_of_resturant desc
```



The screenshot displays the Microsoft SQL Server Management Studio interface. The 'Object Explorer' on the left shows the database structure, including 'Databases', 'Tables', and 'Views'. The 'Query Editor' in the center contains the following SQL query:

```
--What are the Top 10 country which has highest No_of_resturant

select top 10 country.[Country],count(*) as No_of_resturant From zomatodata
join country
on zomatodata.[Country Code]=country.[Country Code]
group by country.[Country]
order by No_of_resturant desc
```

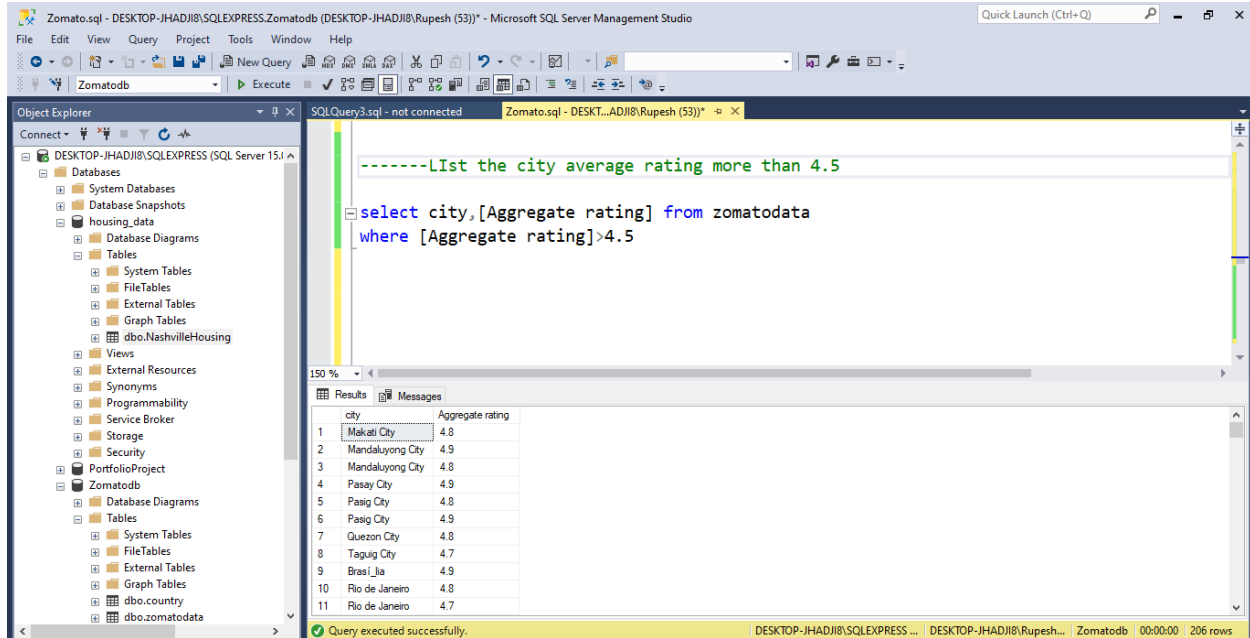
The 'Results' pane at the bottom shows the output of the query, which is a table with two columns: 'Country' and 'No_of_resturant'. The results are as follows:

Country	No_of_resturant
India	8652
United States	434
United Kingdom	80
Brazil	60
South Africa	60
UAE	60
New Zealand	40
Turkey	34
Australia	24
Philippines	22

The status bar at the bottom indicates 'Query executed successfully.' and '10 rows'.

List the city average rating more than 4.5

```
select city,[Aggregate rating] from zomatodata
where [Aggregate rating]>4.5
```



The screenshot displays the Microsoft SQL Server Management Studio interface. The 'Object Explorer' on the left shows the database structure, including 'Zomatodb' and its tables. The 'SQL Query Editor' in the center contains the following query:

```
-----List the city average rating more than 4.5

select city,[Aggregate rating] from zomatodata
where [Aggregate rating]>4.5
```

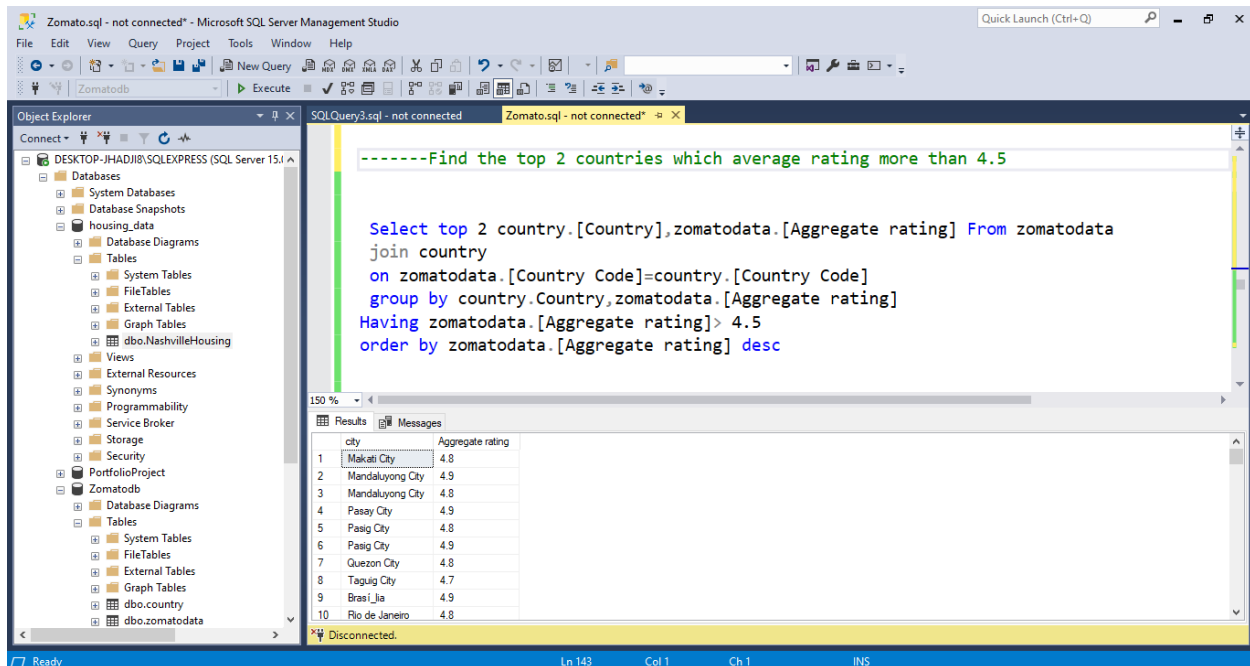
The 'Results' pane at the bottom shows the output of the query, which is a table with two columns: 'city' and 'Aggregate rating'. The table contains 11 rows of data, with the first row highlighted.

	city	Aggregate rating
1	Makati City	4.8
2	Mandaluyong City	4.9
3	Mandaluyong City	4.8
4	Pasay City	4.9
5	Pasig City	4.8
6	Pasig City	4.9
7	Quezon City	4.8
8	Taguig City	4.7
9	Brasília	4.9
10	Rio de Janeiro	4.8
11	Rio de Janeiro	4.7

The status bar at the bottom indicates that the query was executed successfully, returning 206 rows.

Find the top 2 countries which average rating more than 4.5

```
Select top 2 country.[Country],zomatodata.[Aggregate rating] From  
zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]  
group by country.Country,zomatodata.[Aggregate rating]  
Having zomatodata.[Aggregate rating]> 4.5  
order by zomatodata.[Aggregate rating] desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL query:

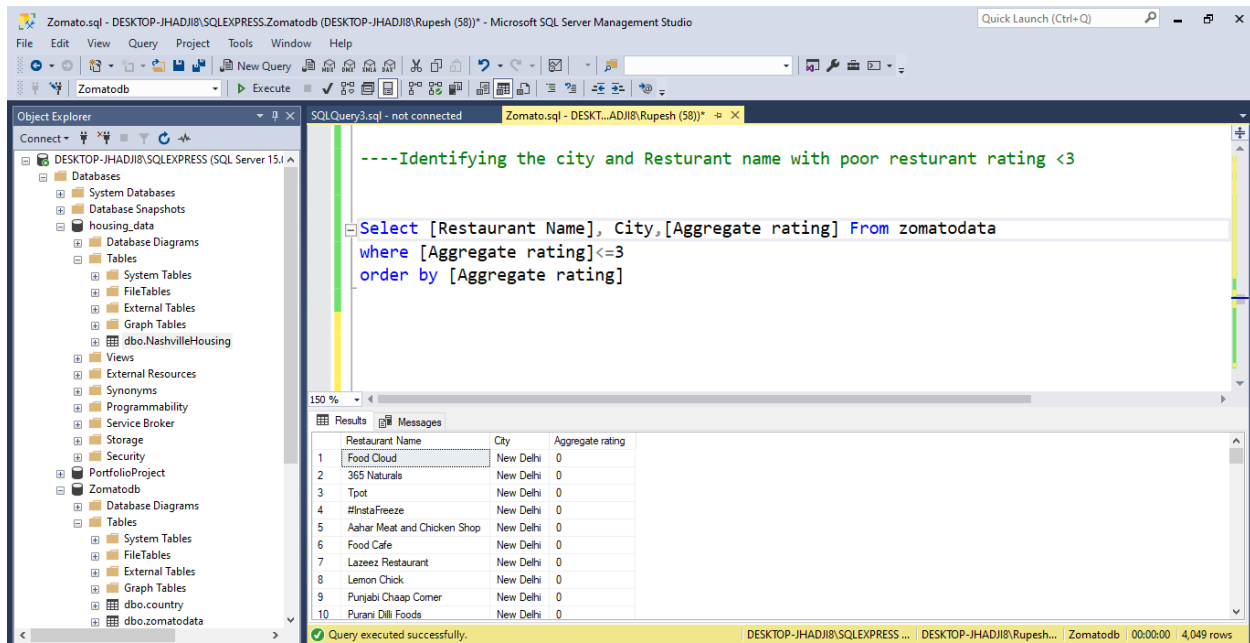
```
-----Find the top 2 countries which average rating more than 4.5  
  
Select top 2 country.[Country],zomatodata.[Aggregate rating] From zomatodata  
join country  
on zomatodata.[Country Code]=country.[Country Code]  
group by country.Country,zomatodata.[Aggregate rating]  
Having zomatodata.[Aggregate rating]> 4.5  
order by zomatodata.[Aggregate rating] desc
```

The query results are displayed in the Results pane, showing a table with two columns: city and Aggregate rating. The results are sorted by Aggregate rating in descending order.

	city	Aggregate rating
1	Makati City	4.8
2	Mandaluyong City	4.9
3	Mandaluyong City	4.8
4	Pasay City	4.9
5	Pasig City	4.8
6	Pasig City	4.9
7	Quezon City	4.8
8	Taguig City	4.7
9	Brasília	4.9
10	Rio de Janeiro	4.8

Identifying the city and Restaurant name with poor restaurant rating <3

```
Select [Restaurant Name], City,[Aggregate rating] From zomatodata
where [Aggregate rating]<=3
order by [Aggregate rating]
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL statement:

```
----Identifying the city and Restaurant name with poor restaurant rating <3

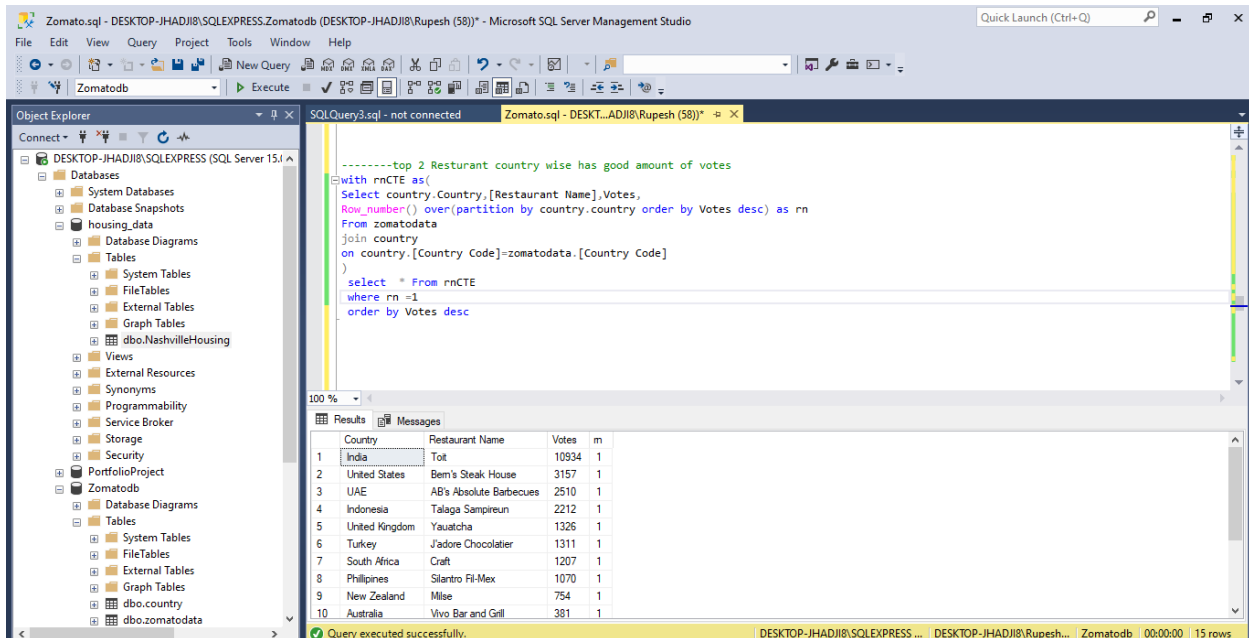
Select [Restaurant Name], City,[Aggregate rating] From zomatodata
where [Aggregate rating]<=3
order by [Aggregate rating]
```

The query has been executed successfully, and the results are displayed in the Results pane. The results show 10 rows of data, all from New Delhi, with an aggregate rating of 0. The status bar at the bottom indicates "Query executed successfully." and "4,049 rows".

	Restaurant Name	City	Aggregate rating
1	Food Cloud	New Delhi	0
2	365 Naturals	New Delhi	0
3	Tpot	New Delhi	0
4	#InstaFreeze	New Delhi	0
5	Aahar Meat and Chicken Shop	New Delhi	0
6	Food Cafe	New Delhi	0
7	Lazeez Restaurant	New Delhi	0
8	Lemon Chick	New Delhi	0
9	Punjabi Chaap Corner	New Delhi	0
10	Purani Dilli Foods	New Delhi	0

Top 2 Resturant country wise has good amount of votes

```
with rnCTE as(  
Select country.Country,[Restaurant Name],Votes,  
Row_number() over(partition by country.country order by Votes desc) as  
rn  
From zomatodata  
join country  
on country.[Country Code]=zomatodata.[Country Code]  
)  
select * From rnCTE  
where rn =1  
order by Votes desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor displays the following SQL query:

```
-----top 2 Resturant country wise has good amount of votes  
with rnCTE as(  
Select country.Country,[Restaurant Name],Votes,  
Row_number() over(partition by country.country order by Votes desc) as rn  
From zomatodata  
join country  
on country.[Country Code]=zomatodata.[Country Code]  
)  
select * From rnCTE  
where rn =1  
order by Votes desc
```

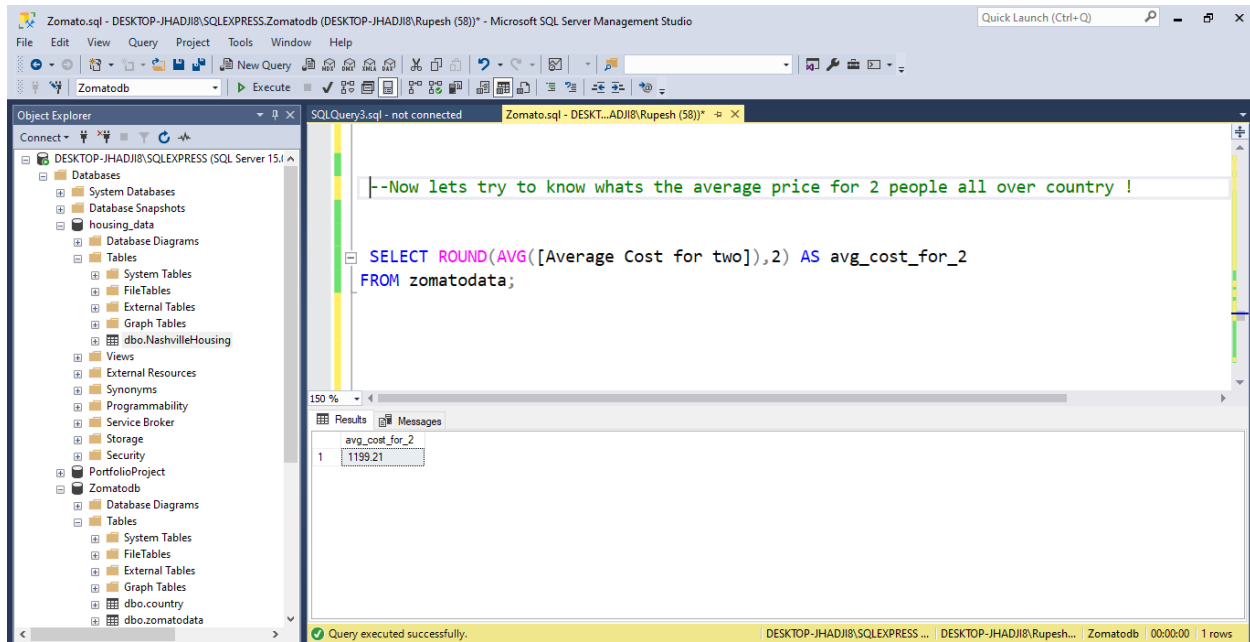
The query results are displayed in the Results pane, showing the top 2 restaurants for each country based on votes. The results are as follows:

Country	Restaurant Name	Votes	m
1 India	Toit	10934	1
2 United States	Bem's Steak House	3157	1
3 UAE	AB's Absolute Barbecues	2510	1
4 Indonesia	Talaga Sampireun	2212	1
5 United Kingdom	Yauatcha	1326	1
6 Turkey	Jadore Chocolatier	1311	1
7 South Africa	Craft	1207	1
8 Philippines	Silantro Fil-Mex	1070	1
9 New Zealand	Mise	754	1
10 Australia	Vivo Bar and Grill	381	1

The status bar at the bottom indicates "Query executed successfully." and "15 rows".

Now lets try to know whats the average price for 2 people all over country !

```
SELECT ROUND(AVG([Average Cost for two]),2) AS avg_cost_for_2  
FROM zomatodata;
```



The screenshot displays the Microsoft SQL Server Management Studio interface. The 'Object Explorer' on the left shows the database structure, including 'Zomatodb' and its tables. The 'SQL Query Editor' in the center contains the following SQL query:

```
--Now lets try to know whats the average price for 2 people all over country !  
  
SELECT ROUND(AVG([Average Cost for two]),2) AS avg_cost_for_2  
FROM zomatodata;
```

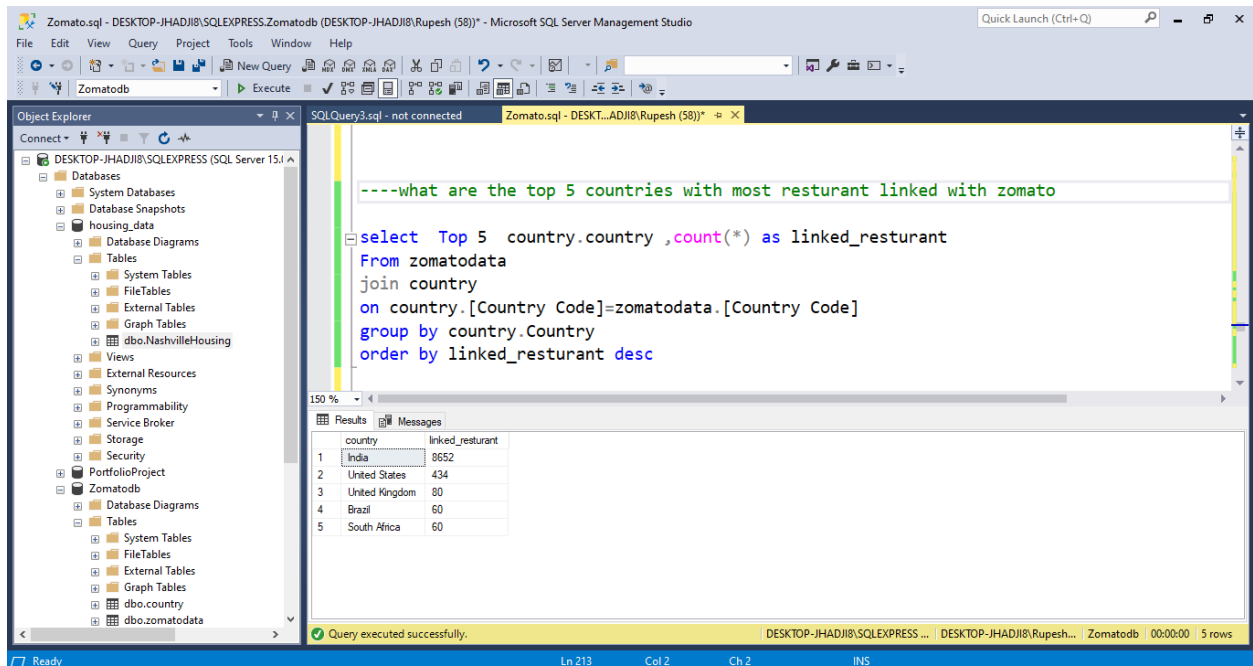
The 'Results' pane at the bottom shows the output of the query, which is a single row with the value 1199.21 for the column 'avg_cost_for_2'.

avg_cost_for_2
1199.21

The status bar at the bottom indicates 'Query executed successfully.' and '1 rows'.

what are the top 5 countries with most restaurant linked with zomato

```
select Top 5 country.country ,count(*) as linked_resturant
From zomatodata
join country
on country.[Country Code]=zomatodata.[Country Code]
group by country.Country
order by linked_resturant desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL query:

```
----what are the top 5 countries with most restaurant linked with zomato

select Top 5 country.country ,count(*) as linked_resturant
From zomatodata
join country
on country.[Country Code]=zomatodata.[Country Code]
group by country.Country
order by linked_resturant desc
```

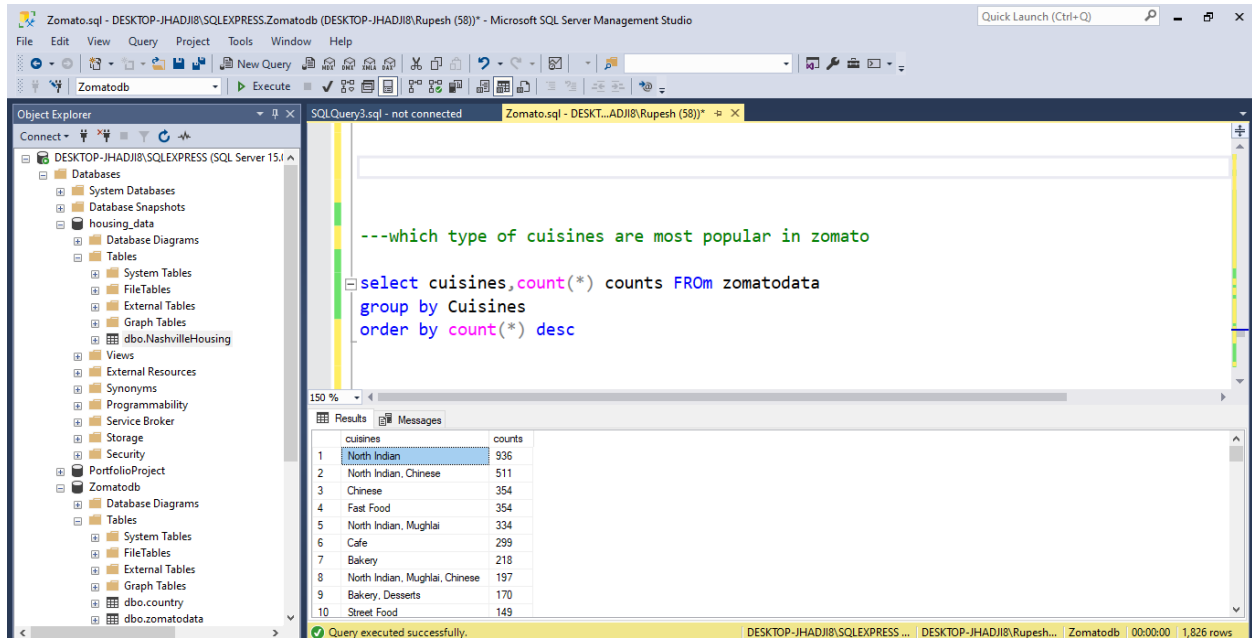
The query results are displayed in the Results pane, showing the top 5 countries with the most restaurants linked to Zomato:

country	linked_resturant
1 India	8652
2 United States	434
3 United Kingdom	80
4 Brazil	60
5 South Africa	60

The status bar at the bottom indicates "Query executed successfully." and "5 rows".

which type of cuisines are most popular in zomato

```
select cuisines,count(*) counts FROM zomatodata
group by Cuisines
order by count(*) desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor contains the following SQL query:

```
--which type of cuisines are most popular in zomato

select cuisines,count(*) counts FROM zomatodata
group by Cuisines
order by count(*) desc
```

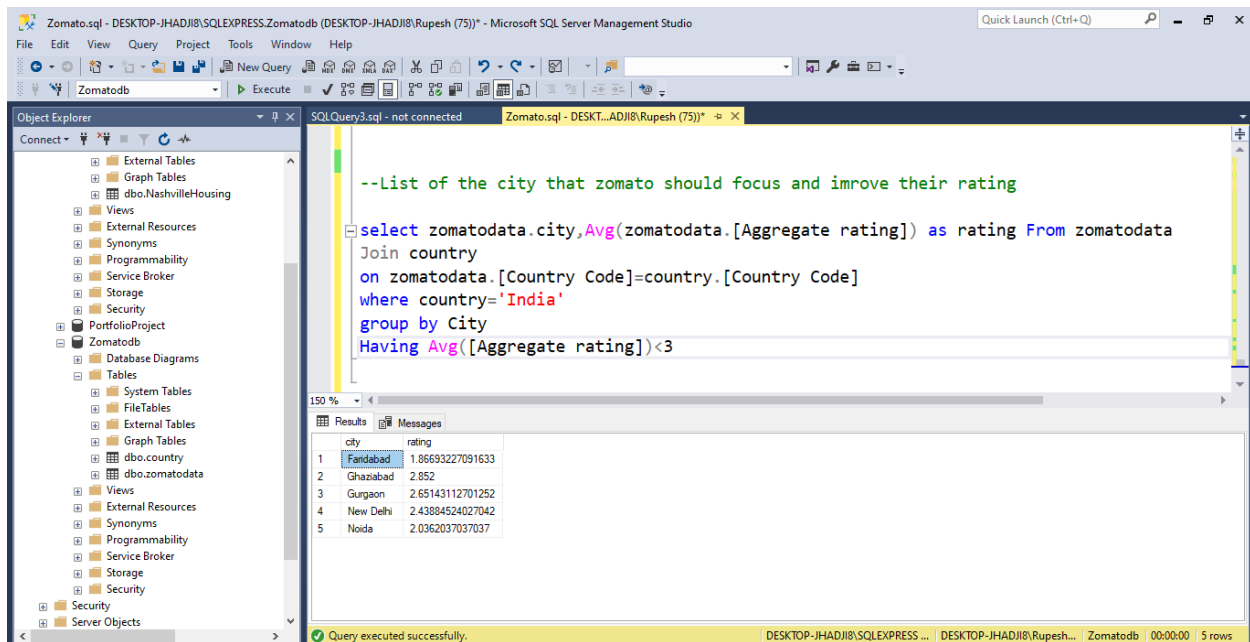
The query results are displayed in the Results pane, showing a table with two columns: cuisines and counts. The results are ordered by count in descending order.

	cuisines	counts
1	North Indian	936
2	North Indian, Chinese	511
3	Chinese	354
4	Fast Food	354
5	North Indian, Mughlai	334
6	Cafe	299
7	Bakery	218
8	North Indian, Mughlai, Chinese	197
9	Bakery, Desserts	170
10	Street Food	149

The status bar at the bottom indicates: Query executed successfully. | DESKTOP-JHADJI8\SQLEXPRESS ... | DESKTOP-JHADJI8\Rupesh ... | Zomatodb | 00:00:00 | 1,826 rows

List of the city that zomato should focus and improve their rating

```
select zomatodata.city,Avg(zomatodata.[Aggregate rating]) as rating
From zomatodata
Join country
on zomatodata.[Country Code]=country.[Country Code]
where country='India'
group by City
Having Avg([Aggregate rating])<3
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left displays the database structure, including the 'Zomatodb' database and its tables. The SQL Query Editor in the center contains the following query:

```
--List of the city that zomato should focus and improve their rating

select zomatodata.city,Avg(zomatodata.[Aggregate rating]) as rating From zomatodata
Join country
on zomatodata.[Country Code]=country.[Country Code]
where country='India'
group by City
Having Avg([Aggregate rating])<3
```

The Results pane at the bottom displays the output of the query, showing a table with 5 rows and 2 columns: 'city' and 'rating'.

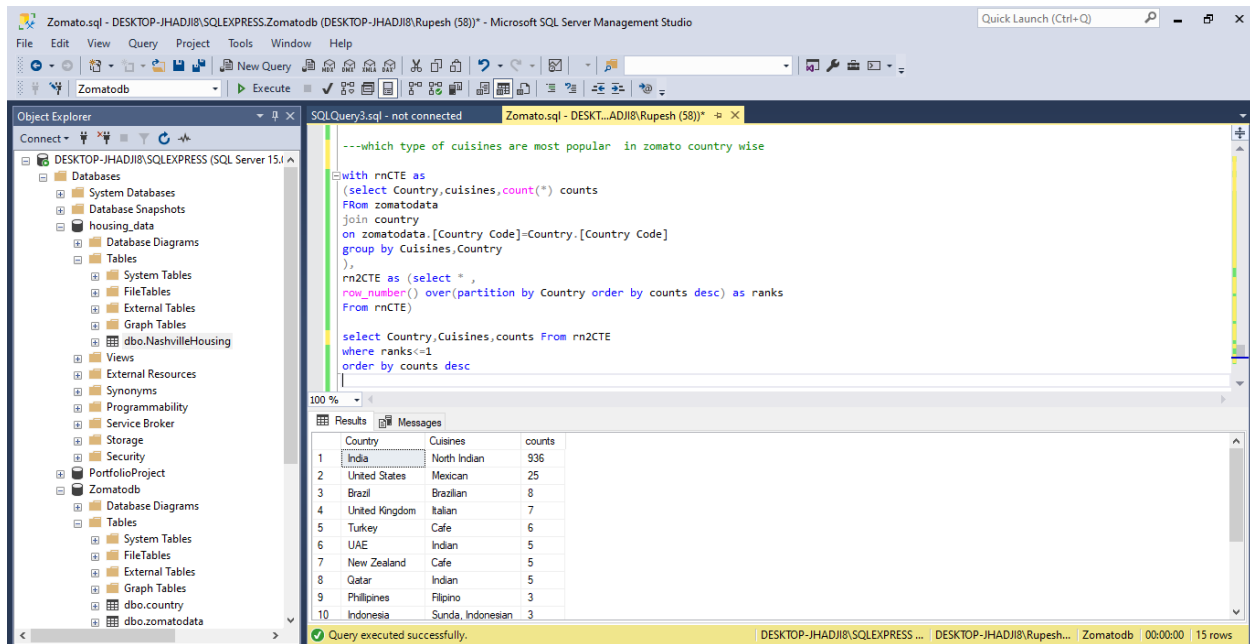
	city	rating
1	Fardabad	1.86693227091633
2	Ghaziabad	2.852
3	Gurgaon	2.65143112701252
4	New Delhi	2.43884524027042
5	Noida	2.0362037037037

The status bar at the bottom indicates that the query was executed successfully, returning 5 rows in 00:00:00 seconds.

which type of cuisines are most popular in zomato country wise

```
with rnCTE as
(select Country,cuisines,count(*) counts
From zomatodata
join country
on zomatodata.[Country Code]=Country.[Country Code]
group by Cuisines,Country
),
rn2CTE as (select * ,
row_number() over(partition by Country order by counts desc) as ranks
From rnCTE)

select * From rn2CTE
where ranks<=1
order by counts desc
```



The screenshot shows the Microsoft SQL Server Management Studio interface. The query editor displays the following SQL query:

```
---which type of cuisines are most popular in zomato country wise

with rnCTE as
(select Country,cuisines,count(*) counts
From zomatodata
join country
on zomatodata.[Country Code]=Country.[Country Code]
group by Cuisines,Country
),
rn2CTE as (select * ,
row_number() over(partition by Country order by counts desc) as ranks
From rnCTE)

select Country,Cuisines,counts From rn2CTE
where ranks<=1
order by counts desc
```

The query results are displayed in the Results pane, showing a table with 3 columns: Country, Cuisines, and counts. The results are ordered by counts in descending order.

Country	Cuisines	counts
1 India	North Indian	936
2 United States	Mexican	25
3 Brazil	Brazilian	8
4 United Kingdom	Italian	7
5 Turkey	Cafe	6
6 UAE	Indian	5
7 New Zealand	Cafe	5
8 Qatar	Indian	5
9 Philippines	Filipino	3
10 Indonesia	Sunda, Indonesian	3

The status bar at the bottom indicates "Query executed successfully." and "15 rows".