**Namma – Yatri Data analysis Report with SQL**

**Introduction about the Namma- yatri:-**

* Namma Yatri is an application for people in Mysore & Bangalore like ola & uber which provides the transportation services to different customers in Bangalore & Mysore through their mobile phone application.
* So the aim is to grow this application among the people of Bangalore in comparision to their competitors which are already very well established ola & uber.
* Thus we need to help them to analyze their data and help generate some insights to take some business decisions as per the analysis.

**Data Walkthrough:-**

For this project as we don’t have the real data we will be creating our own dummy data to solve some complex data analysis questions with the help of SQL through querying from database.

For doing the analysis part from database, which is querying the data will be done in sql server.

We have a database named namma\_yatri\_db and we have 5 main tables included in it.

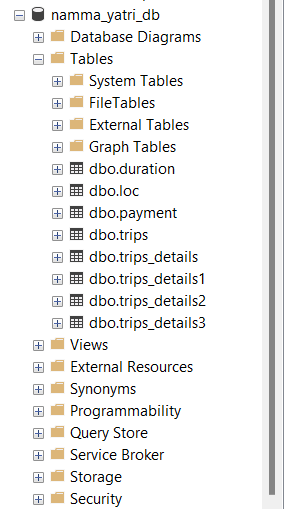
1) trips table

2) trips details table

3) duration table

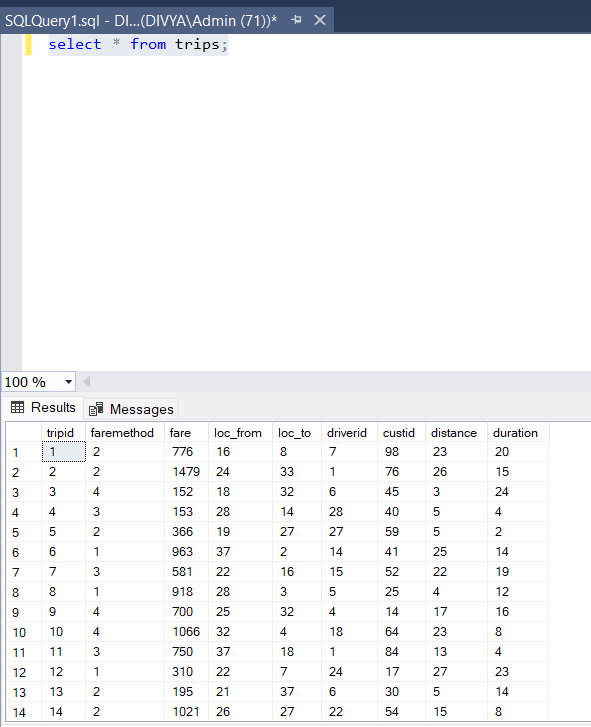
4) payment table

5) loc table

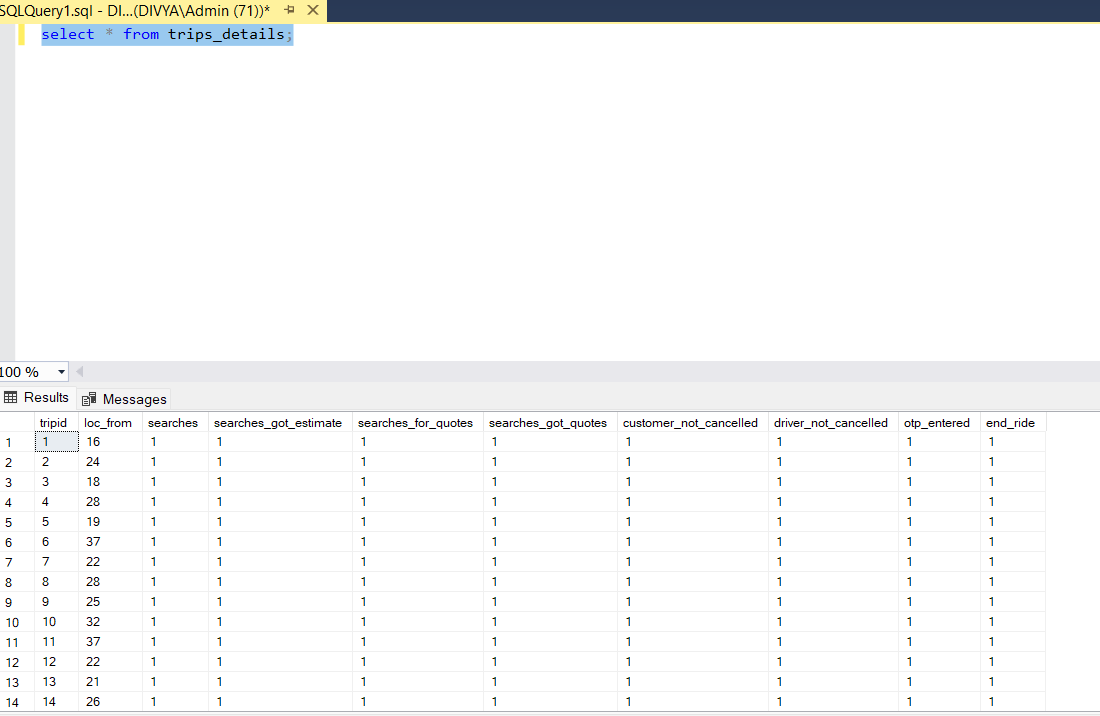


Now lets understand each table in some detail.

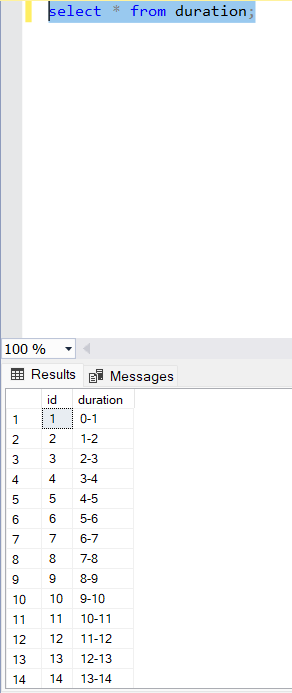
1. Trips table:-This table gives us the overall idea about the trip like the the pickup location to final destination, driver associated, duration taken for trip, how much it costed ,whats the payment mode for the fare.



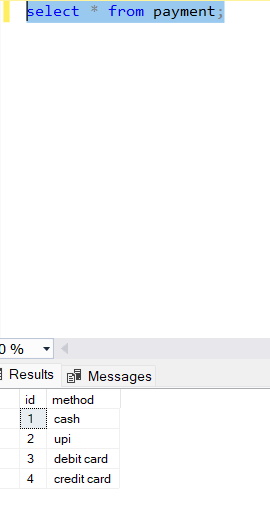
1. trip details table:- It gives the information about all the data points the customer has to go through during the process of booking a ride until the end of ride. It covers the whole journey of the customer. It includes all the successful rides as well as unsuccessful ones.



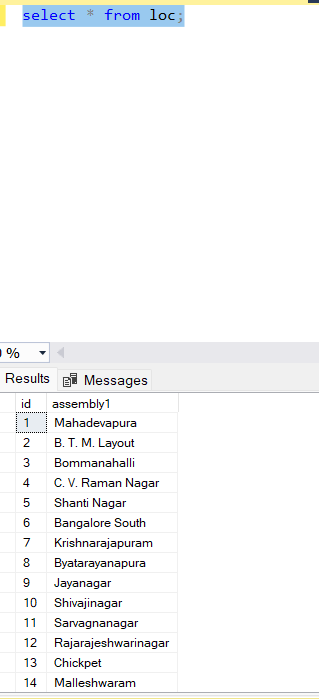
1. Duration table:- It gives the idea about the duration time of the ride.



1. Payment table:- It gives us the info about the payment method through credit card, upi etc.



1. loc table:- Gives us the information about the locations involved in the trips on namma yatri app.



* **Questions for Namma yatri analysis are as follows:-**

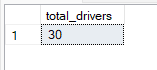
1. Count the total number of trips?

select count(distinct tripid) from trips\_details;

 2161 are the total trips.

1. What are total number of drivers?

select count(distinct driverid) as total\_drivers from trips;

 there are total 30 drivers

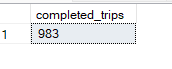
1. what is the total earnings from the app? (this value is on the day basis because we have data for one day)

select sum(fare) as earnings from trips;

 751343 rs. Is the total earnings of the day.

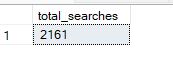
1. How many are total completed trips?

select count(distinct tripid) as completed\_trips from trips;

 983 are total completed trips.

1. How many searches have been there for location on the app?

select sum(searches) as total\_searches from trips\_details;

2161 were the total searches.

1. How many searches which got estimates on app? (means how many users actually entered the location from and location to and tried out to find the price of trip)

select sum(searches\_got\_estimate) as estimated\_searches from trips\_details;



1. How many searches which got quotes? ( here quotes means how many users got the options of riders available after getting the estimates and searching for the riders around to pick them up)

select sum(searches\_got\_quotes) as quoted\_searches from trips\_details;

1277 are the searches which got quotes.

1. Number of trips which are cancelled by the drivers?

select count(\*) - sum(driver\_not\_cancelled) as total\_drivers\_cancelled from trips\_details;

1021 trips are cancelled by drivers.

1. Number of trips which were successfully ended ?

select sum(end\_ride) as total\_end\_ride from trips\_details;

983 were the trips which were successful.

1. What is the average distance per trip?

select avg(distance) as avg\_distance from trips;

14 km is the average distance.

1. What is the average fare per trip?

select avg(fare) as avg\_fare from trips;

764 rs. Is the average fare.

1. Which is the most preferred payment method?

with trip\_payment\_method as (

select a.\*, b.method from trips a inner join payment b

on a.faremethod = b.id )

select top 1 method, count(distinct tripid) as total\_count\_payment\_used from trip\_payment\_method

group by method

order by total\_count\_payment\_used desc;

credit card is most used payment method with 262 trips being payed by this method.

1. The highest payment made through which instrument?

select a.method from payment a inner join

(select top 1 faremethod, sum(fare) as fare from trips

group by faremethod

order by sum(fare) desc) b

on a.id=b.faremethod

highest payment is done through credit card.

1. Which two locations has the most trips?

select b.\* from

(select a.\*,dense\_rank() over(order by total\_trips desc) as rnk

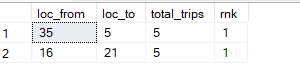
from

(select loc\_from, loc\_to, count(distinct tripid) as total\_trips

from trips

group by loc\_from,loc\_to) as a) as b

where rnk=1;



1. Which are the top 5 earning drivers?

select b.\* from

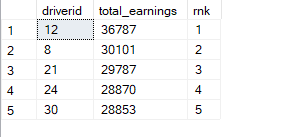
(select a.\*,dense\_rank() over(order by total\_earnings desc) as rnk

from

(select driverid, sum(fare) as total\_earnings from trips

group by driverid) as a) as b

where b.rnk <=5;



1. Which duration had most trips?

select \* from (select a.\*, dense\_rank() over(order by cnt desc) as rnk from

(select duration, count(distinct tripid) as cnt from trips

group by duration) as a) as b

where b.rnk=1;

duration 1 which is from 12 am to 1 am has most trips.

1. Which area producing highest fares?

select \* from

(select a.\*, rank() over(order by total\_fare desc) as rnk

from

(select loc\_from, sum(fare) as total\_fare from trips

group by loc\_from) as a) as c

where c.rnk=1;

Bangalore south produces highest fares.

1. What is the trip completion rate?

select sum(end\_ride) \* 100.0/sum(searches)

from trips\_details;

 45% of the customers successfully ends their ride.

1. Which area gives highest customer cancellations?

select \* from

(select a.\*, rank() over(order by cnt desc) as rnk

from

(select loc\_from, count(\*) - sum(customer\_not\_cancelled) as cnt

from trips\_details

group by loc\_from) as a) as c

where c.rnk=1;

 C.V Raman nagar has highest customer cancellations which is 40.

1. Which duration gives the highest fares?

select \* from

(select a.\*, rank() over(order by total\_fare desc) as rnk

from

(select duration, sum(fare) as total\_fare from trips

group by duration) as a) as c

where c.rnk=1;

the duration of 12am-1am gives highest fare of rs. 45019