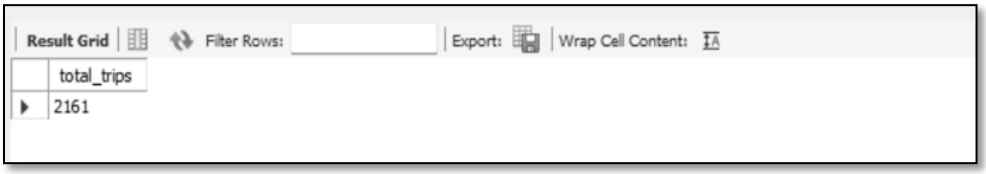
NAMMA YATRI'S ONE DAY DATA SQL ANALYSIS

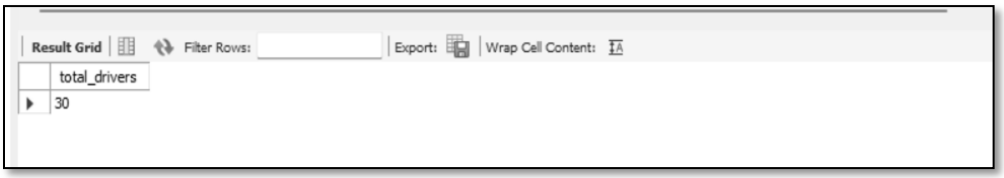
1. Total Trips

```
select count(distinct tripid) as total_trips from trips_details4;
```



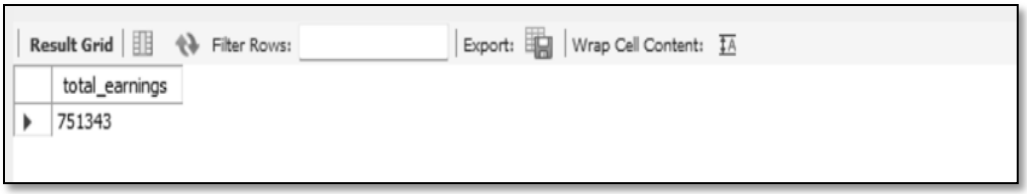
2. Total Drivers

```
select count(distinct driverid) as total_drivers from trips;
```



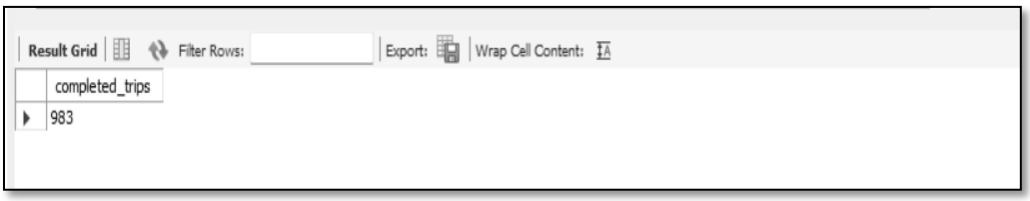
3. Total Earnings

```
select sum(fare) as total_earnings from trips;
```



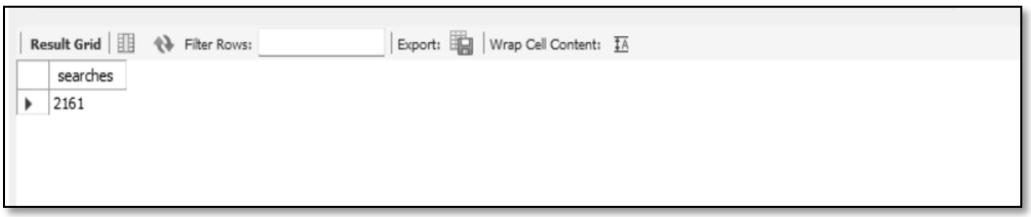
4. Total Completed Trips

```
select sum(end_ride) as completed_trips from trips_details4 ;
```



5. Total Searches

```
select sum(searches) searches from trips_details4;
```

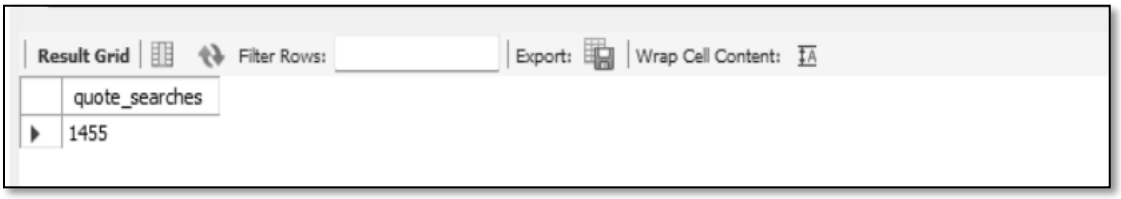


6. Total fare estimate searches

```
select sum(searches_got_estimate) fare_estimate_searches from trips_details4;
```

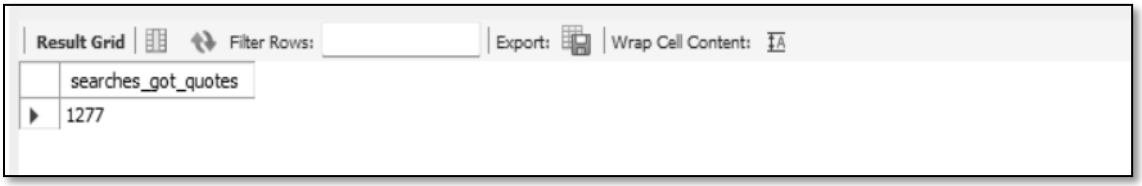
7. Total searches for quotes/drivers

```
select sum(searches_for_quotes) quote_searches from trips_details4;
```



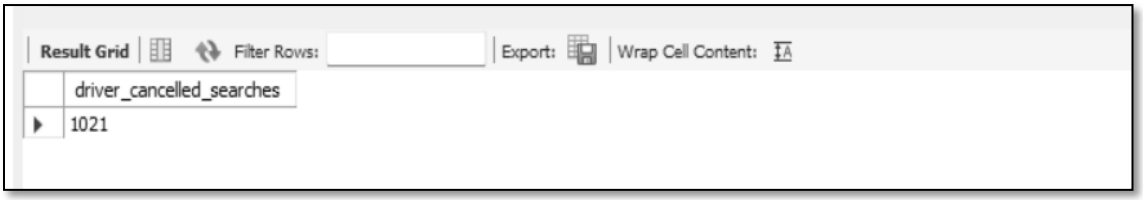
8. Total searches got quotes

```
select sum(searches_got_quotes) searches_got_quotes from trips_details4;
```



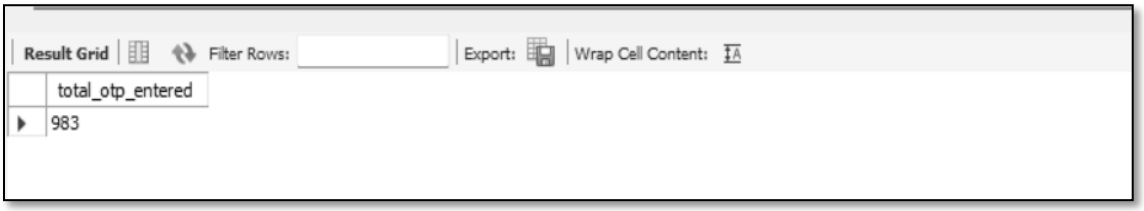
9. Trips cancelled by driver

```
select count(*) - sum(driver_not_cancelled) driver_cancelled_searches from trips_details4;
```



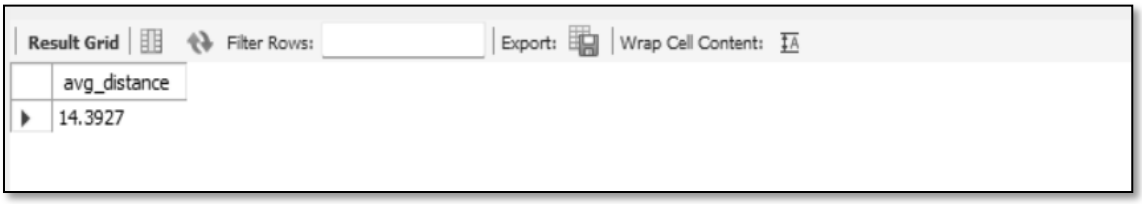
10. Total otp entered

```
select sum(otp_entered) total_otp_entered from trips_details4;
```



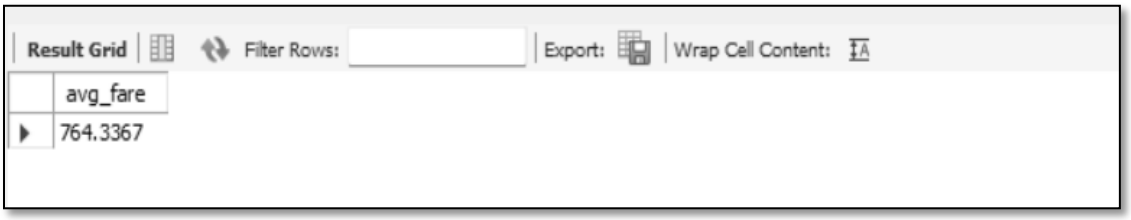
11. Average distance per trip

```
select avg(distance) as avg_distance from trips;
```



12. Average fare per trip

```
select avg(fare) as avg_fare from trips;
```

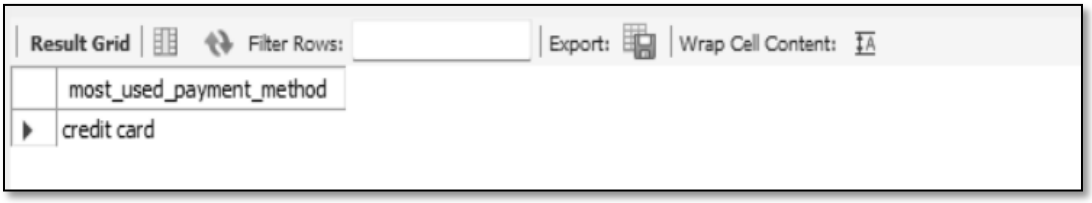


13. Total distance travelled

```
select sum(distance) as distance_travelled from trips;
```

14. Which is most used payment method ?

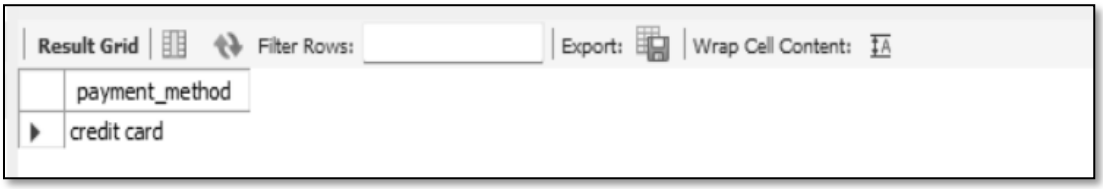
```
select a.method as most_used_payment_method
from payment a
join
(select faremethod, count(distinct tripid) cnt
from trips
group by faremethod order by cnt desc limit 1) b
on a.id = b.faremethod;
```



15. The highest payment was made through which payment method?

```
select a.method as payment_method from payment a
 join
```

```
(select  faremethod, sum(fare)  total_payment
 from trips group  by  faremethod  order by  total_payment desc  limit 1) b
on a.id = b.faremethod;
```



payment_method
credit card

16. Which two locations had the most trips?

(loc_from and loc_to are pair of locations where people are travelling most)

```
select  *  from
(select  * ,  dense_rank()  over(order by trip desc)  rnk
from
(select  loc_from, loc_to, count(distinct  tripid)  trip from trips
group by loc_from, loc_to  order by count(distinct tripid)  desc) a) b
where  rnk = 1;
```

loc_from	loc_to	trip	rnk
16	21	5	1
35	5	5	1

17. Top 5 earning drivers

```
select * from
(select *, dense_rank()  over(order by fare desc) rnk
from
(select  driverid,  sum(fare)  fare from trips
Group  by driverid  order by fare desc) b ) a
Where  rnk < 6;
```

driverid	fare	rnk
12	36787	1
8	30101	2
21	29787	3
24	28870	4
30	28853	5

18. Which duration had more trips?

```
select  *  from
(select *, dense_rank()  over(order by no_of_trips desc)  rnk from
(select duration, count(distinct tripid)  as no_of_trips
from trips
```

```
group  by duration order by no_of_trips desc) b) a
where rnk = 1;
```

duration	no_of_trips	rnk
1	53	1

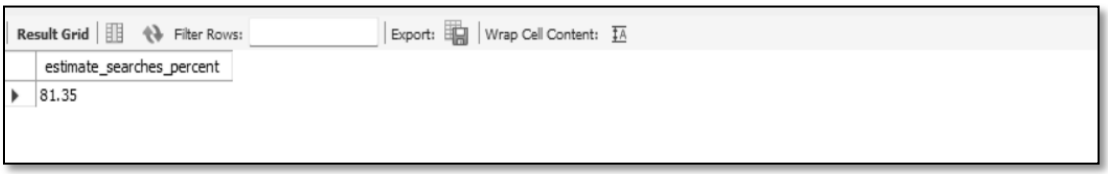
19. Which driver and customer pair had more orders?

```
select * from
(select *, dense_rank()  over(order by driver_cust_pair desc) as rnk
from
(select driverid,  custid , count(*)  as driver_cust_pair
from  trips group by driverid,  custid  order by  driver_cust_pair  desc) b) a
where rnk = 1;
```

driverid	custid	driver_cust_pair	rnk
17	96	4	1
28	15	4	1

20. Searches for estimate fare rate/percent

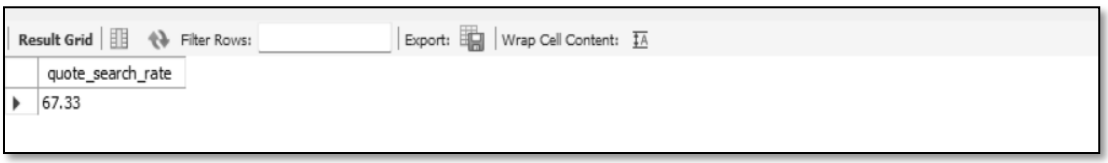
```
select round(sum(searches_got_estimate)/ sum(searches) * 100,2) as estimate_searches_percent
from trips_details4;
```



estimate_searches_percent
81.35

21. Searches for quotes rate

```
select round(sum(searches_for_quotes)/sum(searches) * 100,2)
as quote_search_rate
from trips_details4;
```

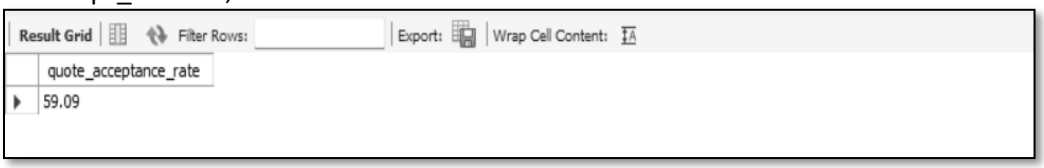


quote_search_rate
67.33

22. Quote acceptance rate

```
select  round(sum(searches_got_quotes)/sum(searches) * 100,2)
as quote_acceptance_rate
```

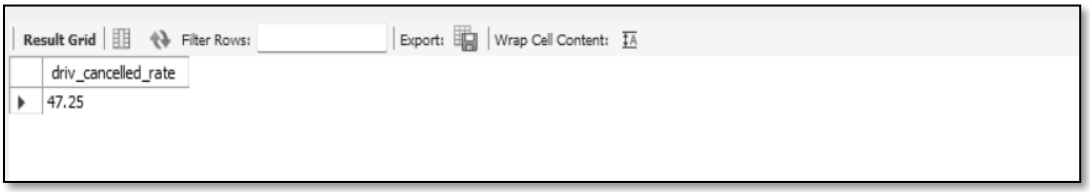
```
from  trips_details4;
```



quote_acceptance_rate
59.09

23. Booking cancellation rate by driver

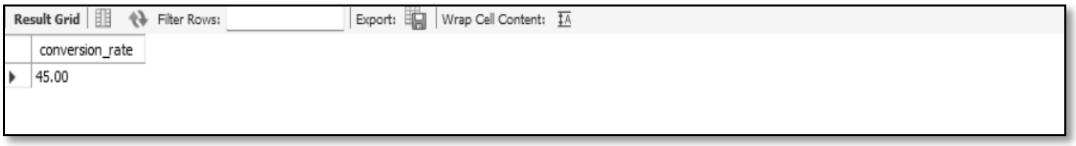
```
 select round((sum(searches)- sum(driver_not_cancelled))/sum(searches)*100,2)
 as driv_cancelled_rate
 from trips_details4;
```



driv_cancelled_rate
47.25

24. Calculate conversion rate?

```
 select  round(sum(end_ride)/sum(searches),2)*100  as  conversion_rate
from trips_details4;
```



conversion_rate
45.00

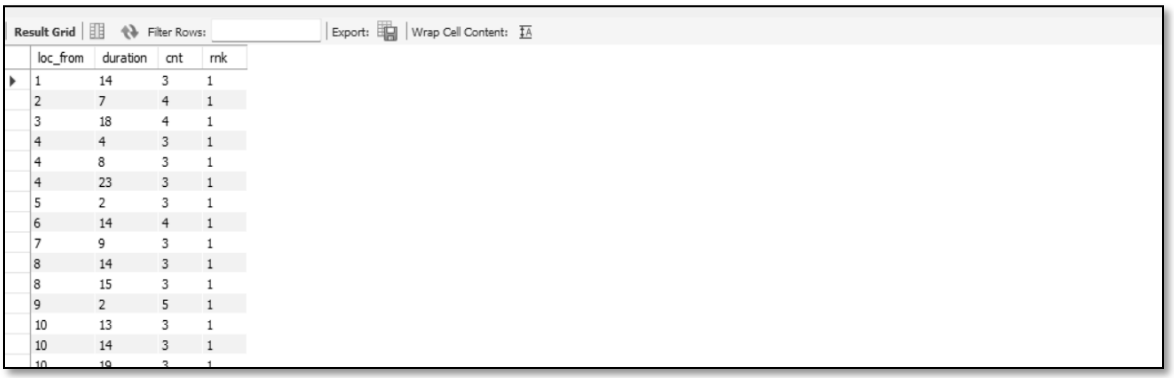
25. Which area got highest trips in which duration(based on duration)?

```
 select * from
 (select *, dense_rank()  over(partition by duration order by cnt desc)  rnk
 from
 (select duration, loc_from , count(distinct tripid) cnt
 from trips
 group by duration,loc_from) a)b
 where rnk = 1;
```

duration	loc_from	cnt	rnk
1	20	4	1
2	9	5	1
3	13	3	1
3	22	3	1
3	34	3	1
4	4	3	1
5	12	3	1
5	26	3	1
5	36	3	1
6	9	4	1
7	19	7	1
8	4	3	1
8	24	3	1
8	26	3	1
8	35	3	1

26. Which duration got highest trips in which area(based on location)?

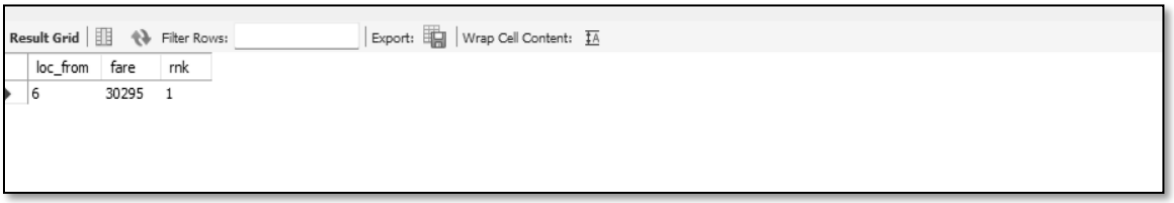
```
select  *  from
(select *, dense_rank() over(partition by loc_from order by cnt desc) rnk
from
(select loc_from , duration,  count(distinct tripid) cnt   from trips
group by loc_from,duration) a)b
where rnk = 1;
```



loc_from	duration	cnt	rnk
1	14	3	1
2	7	4	1
3	18	4	1
4	4	3	1
4	8	3	1
4	23	3	1
5	2	3	1
6	14	4	1
7	9	3	1
8	14	3	1
8	15	3	1
9	2	5	1
10	13	3	1
10	14	3	1

27. Which area got highest fares?

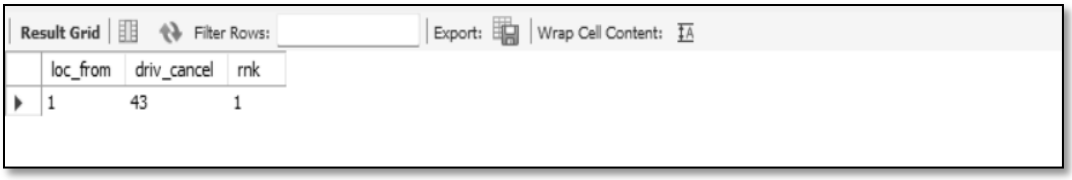
```
select  *  from
(select *, dense_rank() over(order by fare desc) rnk
from
(select loc_from ,  sum(fare) fare from trips
group by loc_from) a)b
where rnk = 1;
```



loc_from	fare	rnk
6	30295	1

28. Which area got highest driver cancellations ?

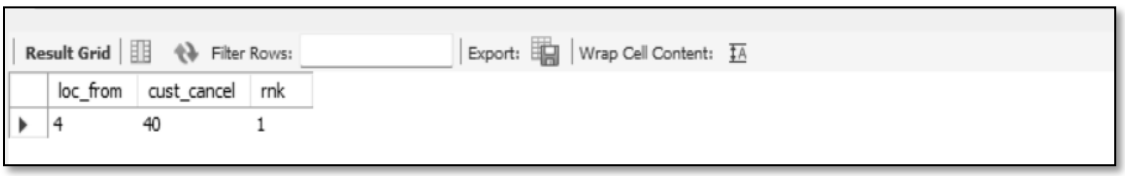
```
select  *  from
(select  *, dense_rank() over(order by driv_cancel desc) rnk
from
(select loc_from, count(*) - sum(driver_not_cancelled) driv_cancel  from trips_details4
group by loc_from) a)b
where rnk = 1;
```



loc_from	driv_cancel	rnk
1	43	1

29. Which area got highest customer cancellations?

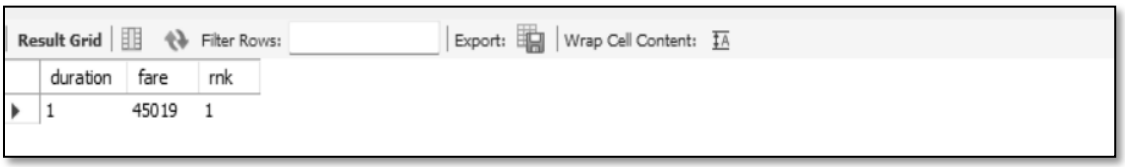
```
select * from
(select *, dense_rank() over(order by cust_cancel desc) rnk from
(select loc_from, count(*) - sum(customer_not_cancelled) cust_cancel  from trips_details4
 group by loc_from) a)b
 where rnk = 1;
```



loc_from	cust_cancel	rnk
4	40	1

30. Which duration got highest fares?

```
select  *  from
(select *, dense_rank() over(order by fare desc) rnk
from
(select duration ,  sum(fare) fare from trips
group by duration) a)b
where rnk = 1;
```



duration	fare	rnk
1	45019	1