TLC licenses over 130,000 vehicles in New York City. Each vehicle receives comprehensive safety and emissions inspections by TLC and must be driven by TLC licensed drivers that have undergone background checks and passed TLC education requirements.

Green Taxis

* Green Taxis provide street hail service and prearranged service in northern Manhattan (above E 96th St and W 110th St) and in the outer boroughs.
* Green Taxis charge standard metered fares for all street hail trips. The price for prearranged trips are set by the base or smartphone app used to reserve the trip.
* Green Taxis are easily identified by their green color, taxi “T” markings, and license numbers on the roof and sides of the vehicle.

As TLC is license issuing authority, it needs to determine the number of new licenses to be released every year. For that purpose, its thinking of following a data driven analysis approach this time. It’s expecting your help to do a detailed analysis of the green taxi trip data for the year 2019

We are using three tables used for querying.

Table1: “green\_trip1”

1. The green taxis trip data VendorID - The record provider
2. lpep\_pickup\_datetime - The date and time when the meter was engaged.
3. lpep\_dropoff\_datetime - The date and time when the meter was disengaged.
4. store\_and\_fwd\_flag – If the record was present in vehicle memory
5. RatecodeID - The final rate code in effect at the end of the trip.
6. PULocationID- Zone in which the taximeter was engaged
7. DOLocationID – Zone in which the taximeter was diengaged
8. passenger\_count- No of passengers in vehicle
9. trip\_distance – trip distance reported by taximeter
10. fare\_amount – fare calculated by taximeter
11. extra – extra charges

Other taxes and fees :

1. mta\_tax
2. tip\_amount
3. tolls\_amount
4. ehail\_fee
5. improvement\_surcharge
6. total\_amount – Amount charged to passengers.

Other details:

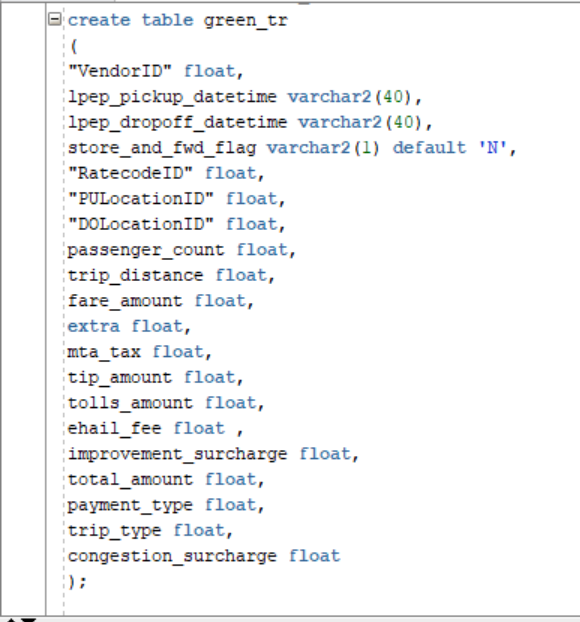
1. payment\_type
2. trip\_type
3. congestion\_surcharge

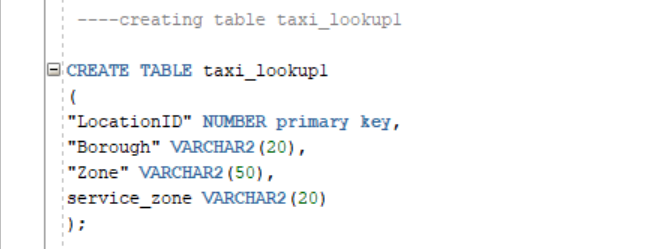
Table2: The Lookup table for Location Data Consisted of details of all the location IDs.

1. LocationID- This states the location ID for the service which has been engaged
2. Borough- District of location ID for the service which has been engaged
3. Zone - zone of location ID for the service which has been engaged
4. service\_zone- service zone of location ID for the service which has been engaged

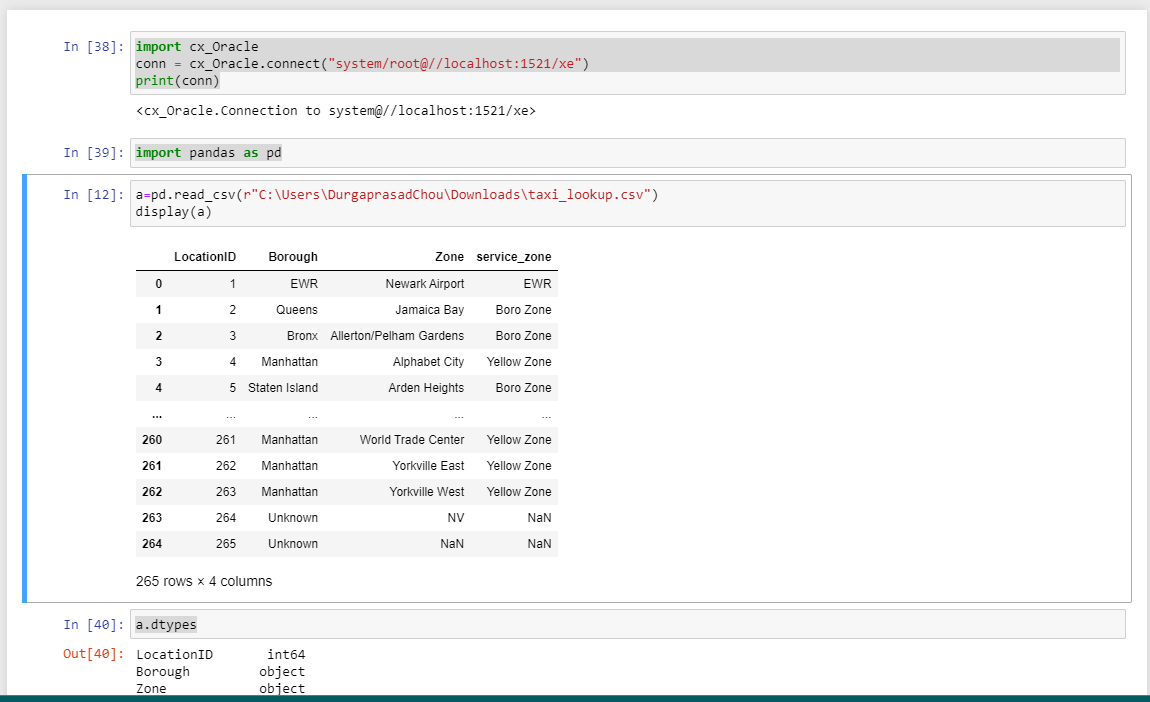
Table3: The payment table consists of details of all the payment types & Methods.

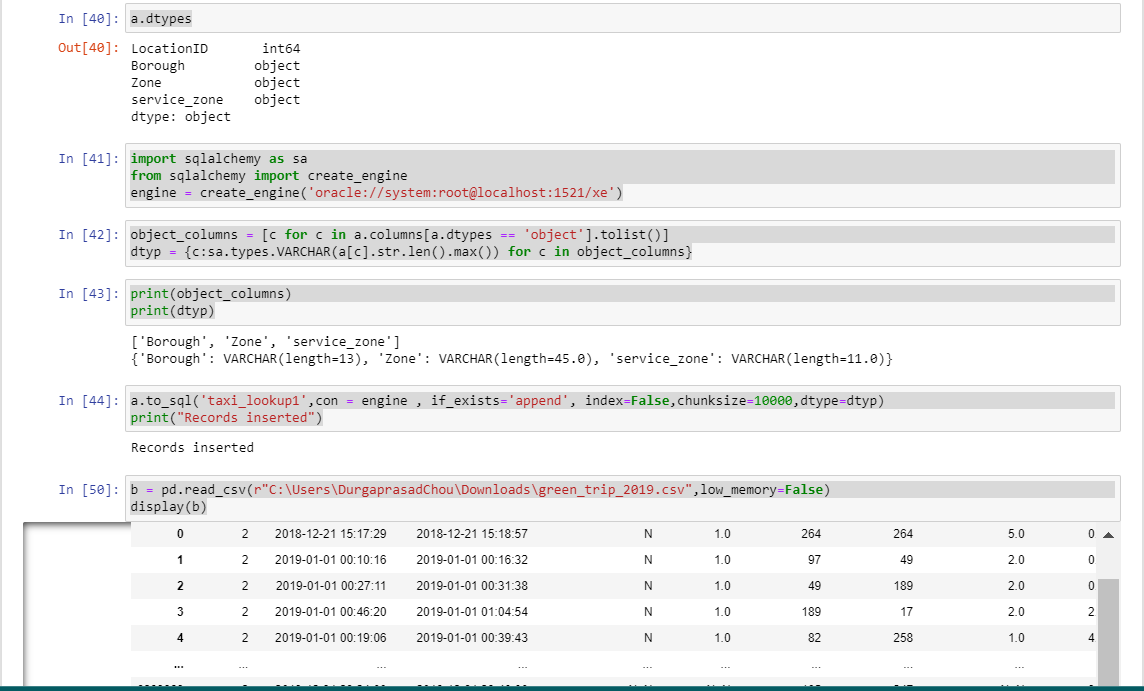
1. Payment\_name- method of payment used for the purpose of fare
2. Payment\_type- type of payment used for the purpose of fare
3. Create the table structure with appropriate data types before loading data





1. Use python to load the all the 2019 files to an Oracle Database table

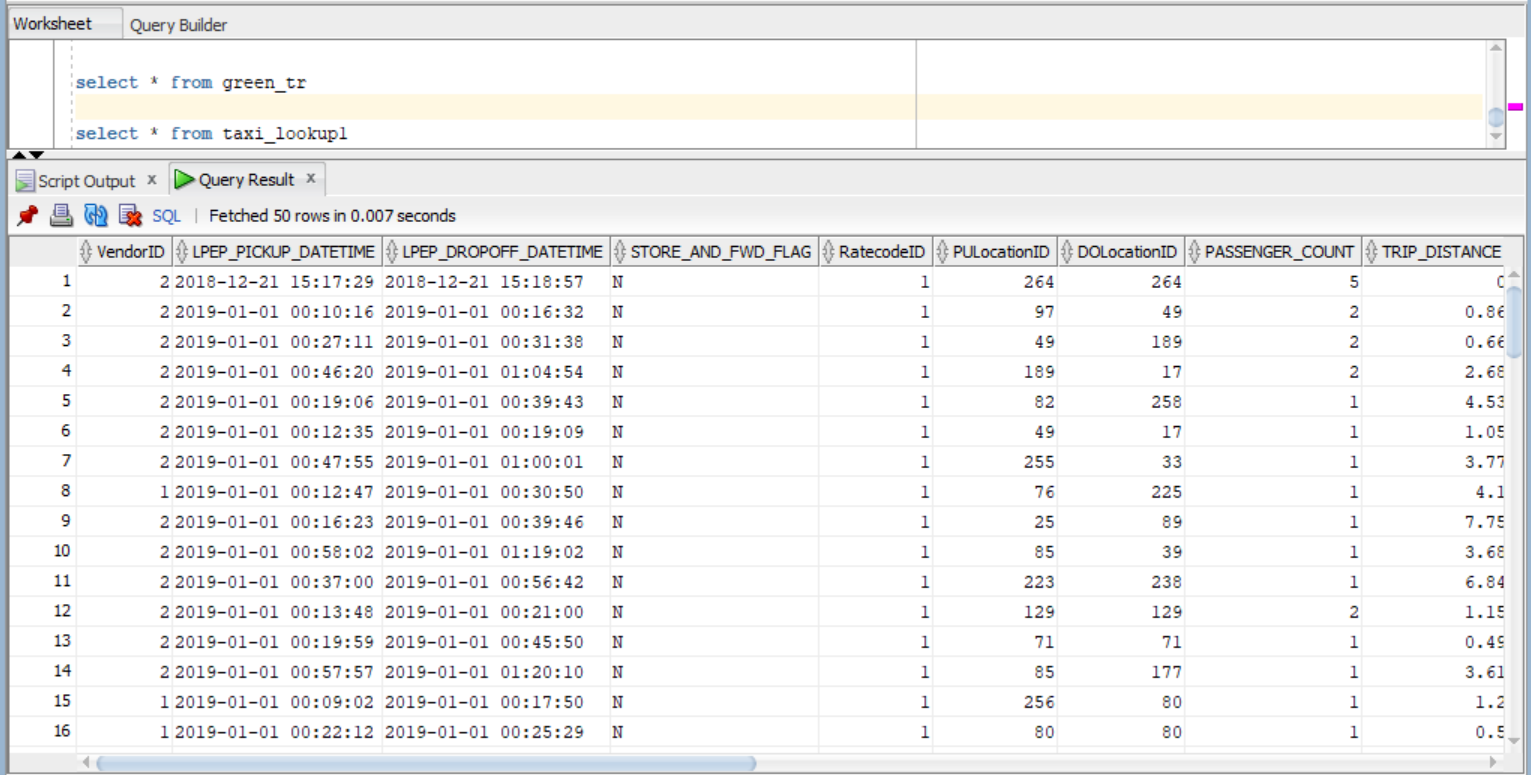


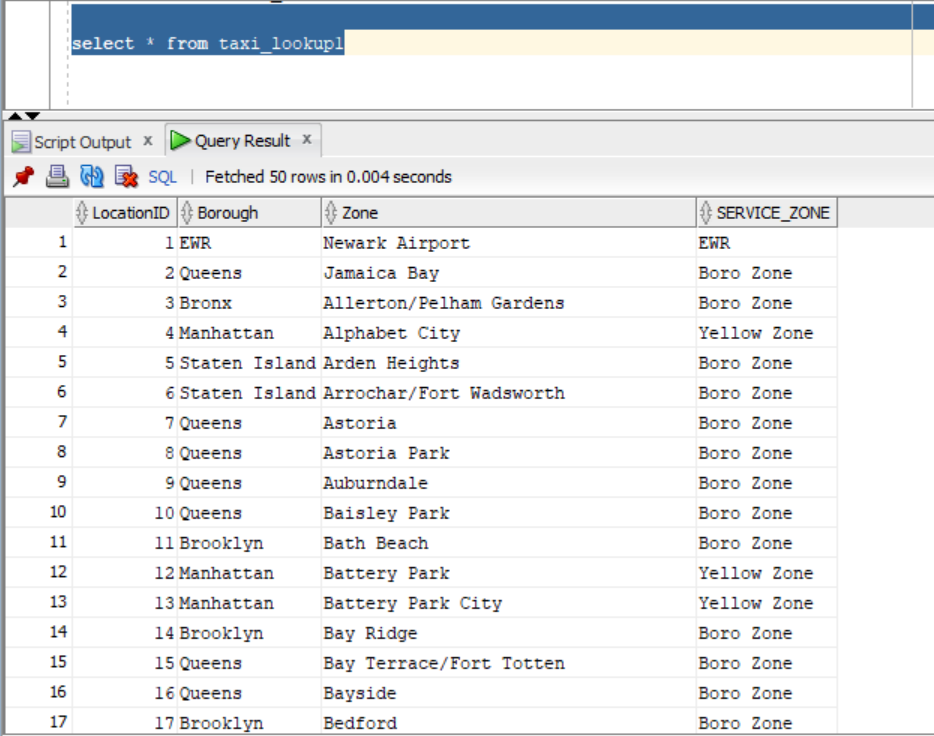




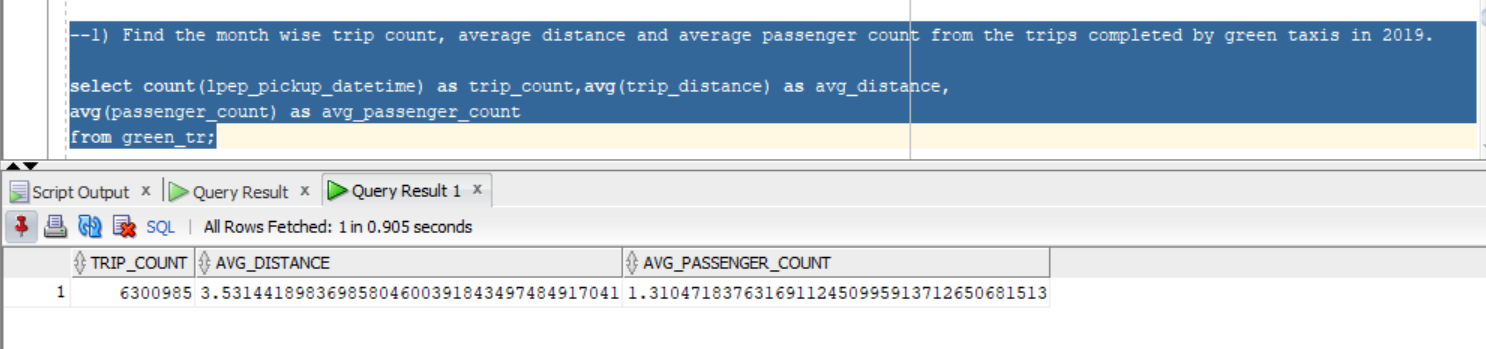
For Reference use this HTML file

Verifing data in oracle

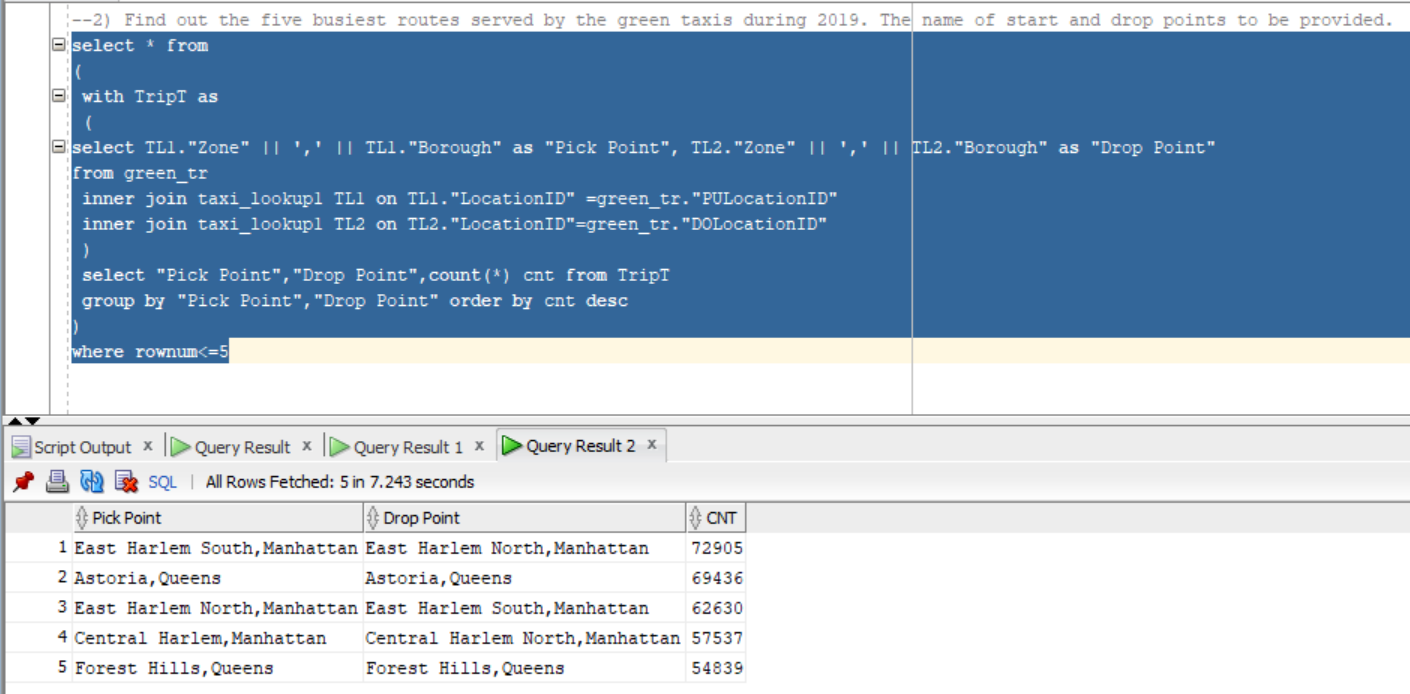




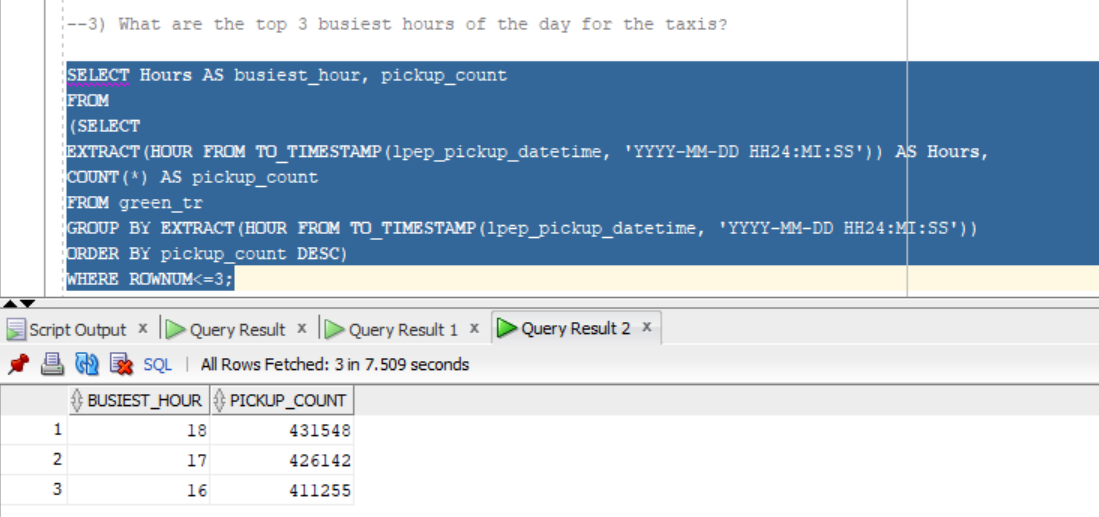
1. Answer the following questions
2. Find the month wise trip count, average distance and average passenger count from the trips completed by green taxis in 2019.



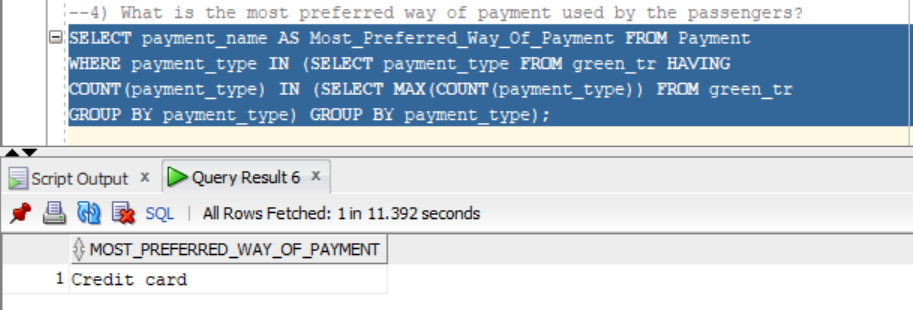
1. Find out the five busiest routes served by the green taxis during 2019. The name of start and drop points to be provided.

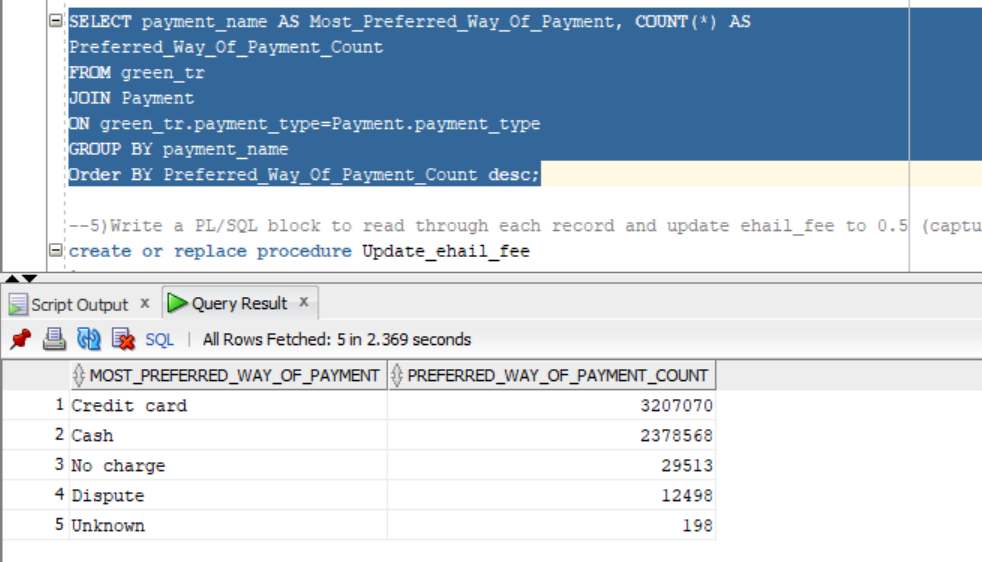


1. What are the top 3 busiest hours of the day for the taxis?

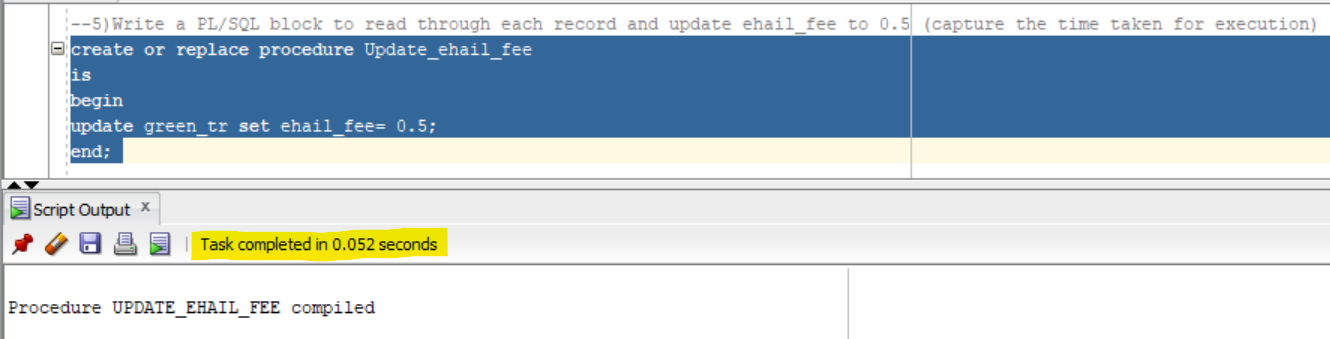


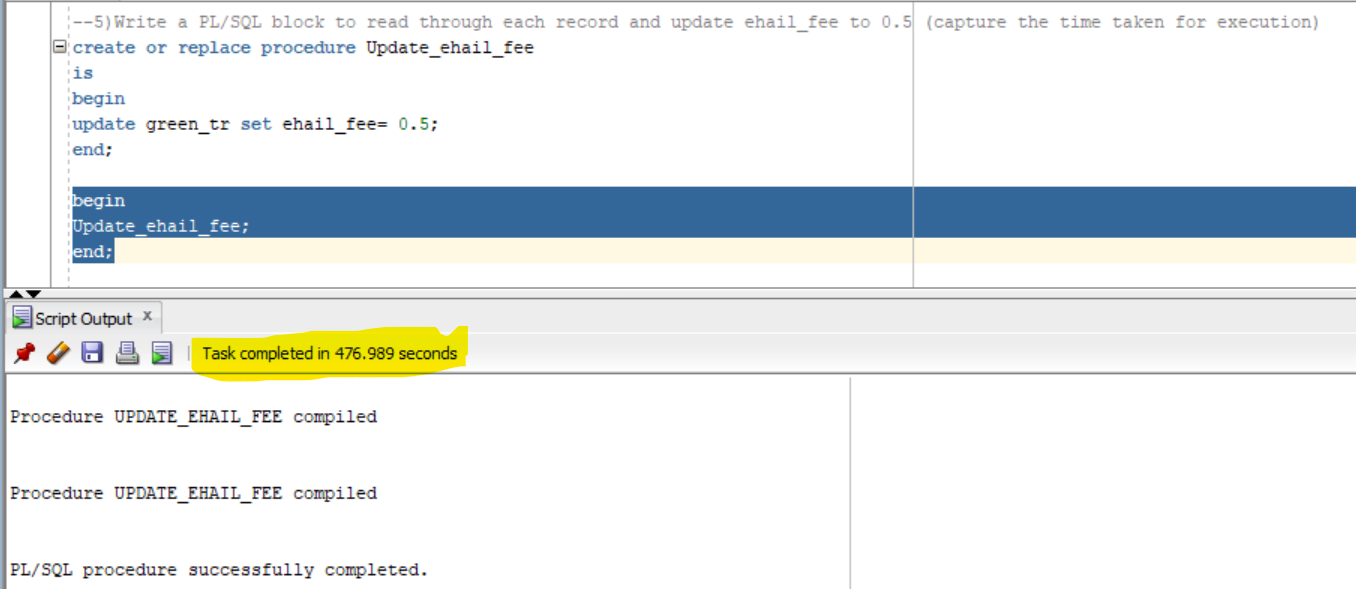
1. What is the most preferred way of payment used by the passengers?

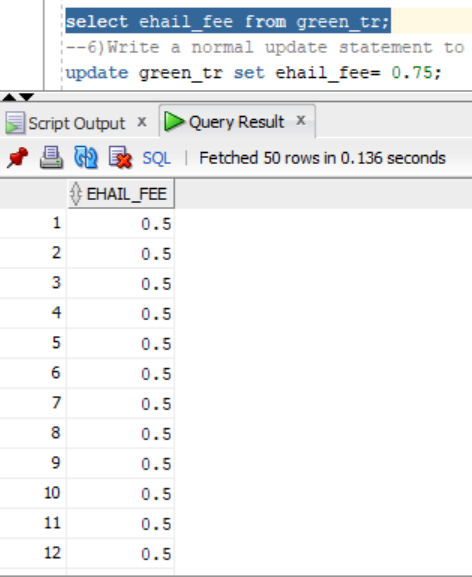




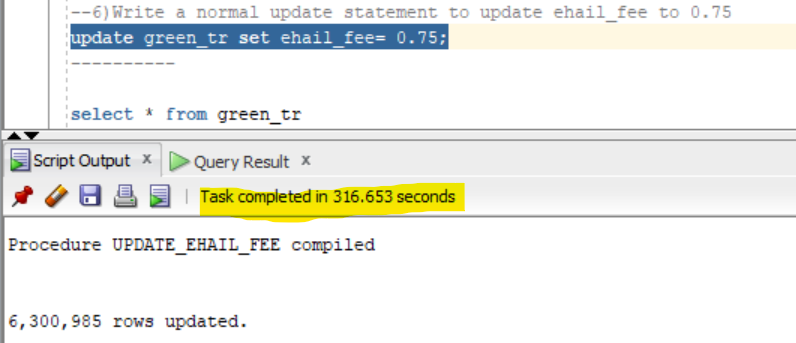
1. Write a PL/SQL block to read through each record and update ehail\_fee to 0.5 (capture the time taken for execution)

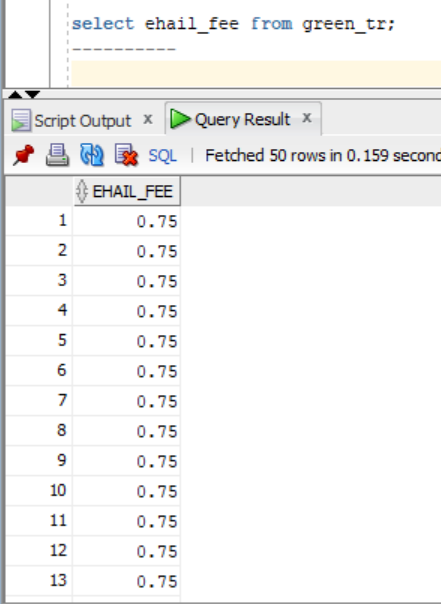






1. Write a normal update statement to update ehail\_fee to 0.75 (capture the time taken for execution)





1. Identify the time taken by e and f and provide your analysis on why each step took more/less time compared to other

Task E took 476.989Seconds to compile which it has 6,300,985 rows while Task F took 316.653seconds when we compare both the compilation data which task E took much time than task F because The procedure stored in SQL server while compiling for the first time this create the ever first ever plan this takes much time while task F has been done normally so it took less time when compared to task E.