

## Newyork Taxi Data – Exploratory Analysis:

### Yearly statistics for Green Taxi:

#### Total Trips:

```
SELECT count(*) trips FROM `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017`
```

**Trips: 11740667**

#### Trips with Tip:

with total as (

```
SELECT count(*) total from `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017` ),  
trips_with_tip as (
```

```
SELECT count(*) trips_with_tip
```

```
FROM `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017`
```

```
where tip_amount>0)
```

```
SELECT trips_with_tip.trips_with_tip, round((trips_with_tip.trips_with_tip*100/total.total),2) as  
Percentage_of_trips
```

```
FROM total,trips_with_tip
```

**Trips with Tip:4870683**

**Percentage of Trips:41.49 %**

#### Trips without tip:

with total as (

```
SELECT count(*) total from `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017` ),  
trips_without_tip as (
```

```
SELECT count(*) trips_without_tip
```

```
FROM `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017` where tip_amount>0)
```

```
SELECT trips_without_tip.trips_without_tip,  
round((trips_without_tip.trips_without_tip*100/total.total),2) as Percentage_of_trips
```

```
FROM total,trips_without_tip
```

**Trips without Tip:6869776**

**Percentage of trips:58.51 %**

**Number of trips during daytime(6 AM to 6 PM):**

```
SELECT count(*) as Number_of_day_trips
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
where HOUR(TIMESTAMP(pickup_datetime)) >=6 and HOUR(TIMESTAMP(pickup_datetime)) <=18
```

**Number of Trips: 6897228****Number of trips during the evening and night time:**

```
SELECT count(*) as Number_of_night_trips
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
where HOUR(TIMESTAMP(pickup_datetime)) >18 or HOUR(TIMESTAMP(pickup_datetime)) <6
```

**Number of Trips: 4843439****Total Earnings through 2017:**

```
SELECT ROUND(SUM(FLOAT(total_amount)),2) AS total_earnings
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
```

**Total Earnings: \$.161 million****Total distance covered in 2017:**

```
SELECT ROUND(SUM(FLOAT(trip_distance)),2) AS total_earnings
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
```

**Total Distance :3.15 million miles****Total Passenger served in 2017:**

```
SELECT SUM(passenger_count) total_passenger_served FROM [bigquery-public-
data.new_york_taxi_trips.tlc_green_trips_2017]
```

**Total Passenger served :16019537 (16 million)****Average of trip\_distance,passenger\_count,per\_ride\_earnings,per\_ride\_tip,average time per trip:**

```
SELECT round(AVG(trip_distance),2) avg_trip_distance,
round(AVG(passenger_count),2) avg_passenger_count,
round(AVG(tip_amount),2) avg_tip_per_ride,
round(AVG(total_amount),2) as average_fare_per_ride,
round(AVG(DATETIME_DIFF(dropoff_datetime,pickup_datetime,minute)),2) avg_time_per_trip
```

FROM `bigquery-public-data.new\_york\_taxi\_trips.tlc\_green\_trips\_2017`

avg\_trip\_distnace: 2.68 miles

avg\_passenger\_count : 1.36

avg\_tip\_per\_ride : \$1.15

average\_fare\_per\_ride : \$14.24

avg\_time\_per\_trip : 20.38 mins

### Average speed of a trip:

SELECT

ROUND(AVG(trip\_distance / DATETIME\_DIFF(dropoff\_datetime,  
pickup\_datetime,  
SECOND))\*3600, 1) Avg\_speed

FROM

`bigquery-public-data.new\_york\_taxi\_trips.tlc\_green\_trips\_2017`

WHERE

trip\_distance > 0

AND dropoff\_datetime > pickup\_datetime

ORDER BY

1 desc

**Avg\_Speed: 14.3**

### Number of trips by trip type : 1= Street-hail 2= Dispatch

SELECT trip\_type,count(\*) as trip\_type\_count

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by trip\_type

order by 2 desc

| Row | trip_type | trip_type_count |
|-----|-----------|-----------------|
| 1   | 1         | 11524542        |
| 2   | 2         | 216098          |
| 3   | null      | 2               |

### Number of trips by payment type: 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip

SELECT payment\_type ,count(\*) as trip\_type\_count

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 2 desc

| Row | payment_type | trip_type_count |
|-----|--------------|-----------------|
| 1   | 1            | 5914813         |
| 2   | 2            | 5740714         |
| 3   | 3            | 54447           |
| 4   | 4            | 30151           |
| 5   | 5            | 542             |

**Most important pickup location :**

SELECT pickup\_location\_id ,count(\*) as pickup\_location\_count

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 2 desc

limit 10

| Row | pickup_location_id | pickup_location_count |
|-----|--------------------|-----------------------|
| 1   | 74                 | 715713                |
| 2   | 41                 | 663025                |
| 3   | 75                 | 623706                |
| 4   | 7                  | 602222                |
| 5   | 82                 | 514707                |
| 6   | 166                | 464673                |
| 7   | 255                | 450229                |
| 8   | 42                 | 415940                |
| 9   | 181                | 395928                |
| 10  | 97                 | 373876                |

**Most important drop-off location:**

SELECT dropoff\_location\_id ,count(\*) as dropoff\_location\_count

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 2 desc

limit 10

| Row | dropoff_location_id | dropoff_location_count |
|-----|---------------------|------------------------|
| 1   | 74                  | 397510                 |
| 2   | 42                  | 391912                 |
| 3   | 7                   | 372001                 |
| 4   | 41                  | 350374                 |
| 5   | 129                 | 328481                 |
| 6   | 181                 | 291212                 |
| 7   | 75                  | 252898                 |
| 8   | 82                  | 233292                 |

| Row | dropoff_location_id | dropoff_location_count |
|-----|---------------------|------------------------|
| 9   | 166                 | 219011                 |
| 10  | 223                 | 212352                 |
| 11  | 97                  | 201692                 |
| 12  | 61                  | 201380                 |

### Longest trip in days:

SELECT \* from (

SELECT pickup\_datetime,dropoff\_datetime,DATEDIFF(dropoff\_datetime,pickup\_datetime ) as  
longest\_trip\_in\_days,trip\_distance,total\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1,2,3,4,5

order by 3 desc)

limit 1

| Row | pickup_datetime     | dropoff_datetime    | longest_trip_in_days | trip_distance | total_amount |
|-----|---------------------|---------------------|----------------------|---------------|--------------|
| 1   | 2017-04-08T23:35:14 | 2017-04-10T14:54:34 | 2                    | 6.9           | 28.8         |

### Tip amount analysis:

#### Average tip amount by passenger count:

SELECT passenger\_count,round(avg(tip\_amount),2) as avg\_tip\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 1

| Row | passenger_count | avg_tip_amount |
|-----|-----------------|----------------|
| 1   | 0               | 1.11           |
| 2   | 1               | 1.14           |
| 3   | 2               | 1.17           |
| 4   | 3               | 1.22           |
| 5   | 4               | 1.08           |
| 6   | 5               | 1.17           |

#### Average tip amount by trip type:

SELECT trip\_type,round(avg(tip\_amount),2) as avg\_tip\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 1

| Row | trip_type | avg_tip_amount |
|-----|-----------|----------------|
| 1   | null      | 0.13           |
| 2   | 1         | 1.15           |
| 3   | 2         | 0.92           |

**Average tip amount by pickup location:**

```
SELECT pickup_location_id ,round(avg(tip_amount),2) as avg_tip_amount  
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]  
group by 1  
order by 2 desc
```

| Row | pickup_location_id | avg_tip_amount |
|-----|--------------------|----------------|
| 1   | 23                 | 19.33          |
| 2   | 109                | 17.64          |
| 3   | 1                  | 11.67          |
| 4   | 214                | 10.33          |
| 5   | 172                | 8.81           |
| 6   | 115                | 6.78           |
| 7   | 132                | 5.60           |
| 8   | 176                | 4.57           |
| 9   | 138                | 4.47           |
| 10  | 84                 | 4.35           |

**Average tip amount by droopff location:**

```
SELECT dropoff_location_id ,round(avg(tip_amount),2) as avg_tip_amount  
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]  
group by 1  
order by 2 desc
```

| Row | dropoff_location_id | avg_tip_amount |
|-----|---------------------|----------------|
| 1   | 1                   | 12.49          |
| 2   | 23                  | 11.07          |
| 3   | 109                 | 8.87           |
| 4   | 84                  | 8.65           |
| 5   | 214                 | 6.68           |
| 6   | 5                   | 6.36           |
| 7   | 132                 | 5.49           |
| 8   | 176                 | 5.40           |
| 9   | 115                 | 5.38           |
| 10  | 251                 | 5.20           |

**Toll amount analysis:****Average toll amount by pickup location (Which location contributes to more toll):**

```
SELECT pickup_location_id ,round(avg(tolls_amount ),2) as avg_toll_amount  
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]  
group by 1  
order by 2 desc
```

| Row | pickup_location_id | avg_toll_amount |
|-----|--------------------|-----------------|
| 1   | 199                | 573.01          |
| 2   | 6                  | 4.27            |
| 3   | 194                | 3.23            |
| 4   | 115                | 3.14            |
| 5   | 5                  | 1.92            |
| 6   | 99                 | 1.75            |
| 7   | 221                | 1.54            |
| 8   | 253                | 1.26            |
| 9   | 93                 | 1.25            |
| 10  | 15                 | 1.19            |

**Average toll amount by dropoff location (Which location contributes to more toll):**

```
SELECT dropoff_location_id ,round(avg(tolls_amount ),2) as avg_toll_amount
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
group by 1
order by 2 desc
```

| Row | dropoff_location_id | avg_toll_amount |
|-----|---------------------|-----------------|
| 1   | 199                 | 616.23          |
| 2   | 1                   | 11.00           |
| 3   | 5                   | 7.85            |
| 4   | 84                  | 7.38            |
| 5   | 109                 | 6.71            |
| 6   | 204                 | 6.62            |
| 7   | 214                 | 6.40            |
| 8   | 172                 | 6.39            |
| 9   | 6                   | 6.26            |
| 10  | 176                 | 6.09            |

**Highest paid trip of the year:**

```
select vendor_id, pickup_datetime, dropoff_datetime, fare_amount , tolls_amount , total_amount
from [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
where tolls_amount in (SELECT max(tolls_amount )
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017] )
```

| Row | vendor_id | pickup_datetime     | dropoff_datetime    | fare_amount | tolls_amount | total_amount |
|-----|-----------|---------------------|---------------------|-------------|--------------|--------------|
| 1   | 1         | 2017-03-28T15:32:05 | 2017-03-28T15:34:40 | 999.99      | 7999.92      | 8999.91      |

**Quarterly Analysis :**

**Highest earning quarter:**

```
SELECT
    QUARTER(pickup_datetime) AS quarter,
    ROUND(SUM(total_amount ),2)
```

```

    AS Total_amt,
ROUND(AVG(total_amount ),2)

    AS Avg_Total_amt,

count(*) as no_of_trips
FROM

[bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY

quarter

ORDER BY

2 DESC

```

| Row | quarter | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|---------|---------------|---------------|-------------|
| 1   | 1       | 4.473036822E7 | 13.75         | 3252311     |
| 2   | 2       | 4.463086715E7 | 14.32         | 3117701     |
| 3   | 3       | 3.893431884E7 | 14.61         | 2664562     |
| 4   | 4       | 3.89150275E7  | 14.38         | 2706093     |

#### Quarter with most number of trips:

```

SELECT

QUARTER(pickup_datetime) AS quarter,

count(*) as no_of_trips
FROM

[bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY

quarter

ORDER BY

2 DESC

```

| Row | quarter | no_of_trips |
|-----|---------|-------------|
| 1   | 1       | 3252311     |
| 2   | 2       | 3117701     |
| 3   | 4       | 2706093     |
| 4   | 3       | 2664562     |

#### Quarter with highest passenger served:

```

SELECT

```



```

QUARTER(pickup_datetime) AS quarter,
round(sum(passenger_count ),2) as no_of_passenger,
round(avg(passenger_count ),2) as avg_passenger_count,
count(*) as no_of_trips
FROM
[bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
GROUP BY
quarter
ORDER BY
2 DESC

```

| Row | quarter | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|---------|-----------------|---------------------|-------------|
| 1   | 1       | 4438079.0       | 1.36                | 3252311     |
| 2   | 2       | 4242903.0       | 1.36                | 3117701     |
| 3   | 4       | 3702838.0       | 1.37                | 2706093     |
| 4   | 3       | 3635717.0       | 1.36                | 2664562     |

#### Tip amounts per quarter:

```

SELECT
QUARTER(pickup_datetime) AS quarter,
round(sum(tip_amount ),2) as tip_amount,
round(avg(tip_amount ),2) as avg_tip_amount,
count(*) as no_of_trips
FROM
[bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
GROUP BY
quarter
ORDER BY
2 DESC

```

| Row | quarter | tip_amount | avg_tip_amount | no_of_trips |
|-----|---------|------------|----------------|-------------|
| 1   | 2       | 3716939.57 | 1.19           | 3117701     |
| 2   | 1       | 3685718.11 | 1.13           | 3252311     |
| 3   | 3       | 3120907.49 | 1.17           | 2664562     |
| 4   | 4       | 2974552.04 | 1.1            | 2706093     |

#### Monthly statistics:

### Number of trips :

```
SELECT

    MONTH(pickup_datetime) AS month,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY

    month

ORDER BY

    2 DESC
```

| Row | month | no_of_trips |
|-----|-------|-------------|
| 1   | 3     | 1158641     |
| 2   | 4     | 1081771     |
| 3   | 1     | 1070399     |
| 4   | 5     | 1059463     |
| 5   | 2     | 1023271     |

### Monthly earnings:

```
SELECT

    MONTH(pickup_datetime) AS month,

    ROUND(SUM(total_amount ),2)

    AS Total_amt,

    ROUND(AVG(total_amount ),2)

    AS Avg_Total_amt,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY

    month

ORDER BY

    2 DESC
```

| Row | month | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|-------|---------------|---------------|-------------|
| 1   | 3     | 1.60176212E7  | 13.82         | 1158641     |
| 2   | 4     | 1.524857871E7 | 14.1          | 1081771     |
| 3   | 5     | 1.521545505E7 | 14.36         | 1059463     |
| 4   | 1     | 1.459105612E7 | 13.63         | 1070399     |

| Row | month | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|-------|---------------|---------------|-------------|
| 5   | 6     | 1.416683339E7 | 14.51         | 976467      |
| 6   | 2     | 1.41216909E7  | 13.8          | 1023271     |
| 7   | 10    | 1.350517949E7 | 14.59         | 925722      |
| 8   | 7     | 1.324072817E7 | 14.47         | 914783      |
| 9   | 9     | 1.311626901E7 | 14.86         | 882422      |
| 10  | 12    | 1.284243273E7 | 14.17         | 906154      |
| 11  | 8     | 1.257732166E7 | 14.5          | 867357      |
| 12  | 11    | 1.256741528E7 | 14.38         | 874217      |

#### Passengers served per month:

SELECT

MONTH(pickup\_datetime) AS month,

round(sum(passenger\_count ),2) as no\_of\_passenger,

round(avg(passenger\_count ),2) as avg\_passenger\_count,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

quarter

ORDER BY

2 DESC

| Row | month | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|-------|-----------------|---------------------|-------------|
| 1   | 3     | 1578619.0       | 1.36                | 1158641     |
| 2   | 4     | 1477670.0       | 1.37                | 1081771     |
| 3   | 1     | 1464835.0       | 1.37                | 1070399     |
| 4   | 5     | 1441878.0       | 1.36                | 1059463     |
| 5   | 2     | 1394625.0       | 1.36                | 1023271     |
| 6   | 6     | 1323355.0       | 1.36                | 976467      |
| 7   | 10    | 1263416.0       | 1.36                | 925722      |
| 8   | 7     | 1253084.0       | 1.37                | 914783      |
| 9   | 12    | 1239758.0       | 1.37                | 906154      |
| 10  | 11    | 1199664.0       | 1.37                | 874217      |
| 11  | 9     | 1197976.0       | 1.36                | 882422      |
| 12  | 8     | 1184657.0       | 1.37                | 867357      |

#### Tip data per month:

SELECT

MONTH(pickup\_datetime) AS month,

round(sum(tip\_amount ),2) as tip\_amount,

round(avg(tip\_amount ),2) as avg\_tip\_amount,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

month

ORDER BY

2 DESC

| Row | month | tip_amount | avg_tip_amount | no_of_trips |
|-----|-------|------------|----------------|-------------|
| 1   | 3     | 1299203.43 | 1.12           | 1158641     |
| 2   | 5     | 1274152.43 | 1.2            | 1059463     |
| 3   | 4     | 1262485.9  | 1.17           | 1081771     |
| 4   | 1     | 1207574.12 | 1.13           | 1070399     |
| 5   | 6     | 1180301.24 | 1.21           | 976467      |
| 6   | 2     | 1178940.56 | 1.15           | 1023271     |
| 7   | 7     | 1089993.15 | 1.19           | 914783      |
| 8   | 10    | 1060768.32 | 1.15           | 925722      |
| 9   | 9     | 1035108.3  | 1.17           | 882422      |
| 10  | 8     | 995806.04  | 1.15           | 867357      |
| 11  | 11    | 963153.56  | 1.1            | 874217      |
| 12  | 12    | 950630.16  | 1.05           | 906154      |

**Day wise statistics:**

**Number of trips:**

SELECT

dayofweek(pickup\_datetime) AS dayofweek,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | no_of_trips |
|-----|-----------|-------------|
| 1   | 7         | 2017033     |
| 2   | 6         | 1881905     |
| 3   | 1         | 1720831     |
| 4   | 5         | 1653340     |
| 5   | 4         | 1571155     |
| 6   | 3         | 1474205     |
| 7   | 2         | 1422198     |

### Day wise earnings:

```
SELECT

    dayofweek(pickup_datetime) AS dayofweek,

    ROUND(SUM(total_amount ),2)

        AS Total_amt,

    ROUND(AVG(total_amount ),2)

        AS Avg_Total_amt,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY

    dayofweek

ORDER BY

    2 DESC
```

| Row | dayofweek | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|-----------|---------------|---------------|-------------|
| 1   | 7         | 2.865920813E7 | 14.21         | 2017033     |
| 2   | 6         | 2.696677971E7 | 14.33         | 1881905     |
| 3   | 1         | 2.455372944E7 | 14.27         | 1720831     |
| 4   | 5         | 2.375835792E7 | 14.37         | 1653340     |
| 5   | 4         | 2.241299761E7 | 14.27         | 1571155     |
| 6   | 3         | 2.083678701E7 | 14.13         | 1474205     |
| 7   | 2         | 2.002272189E7 | 14.08         | 1422198     |

### Passenger spread across each day:

```
SELECT

    dayofweek(pickup_datetime) AS dayofweek,

    round(sum(passenger_count ),2) as no_of_passenger,

    round(avg(passenger_count ),2) as avg_passenger_count,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]

GROUP BY
```

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|-----------|-----------------|---------------------|-------------|
| 1   | 7         | 2779826.0       | 1.38                | 2017033     |
| 2   | 6         | 2549289.0       | 1.35                | 1881905     |
| 3   | 1         | 2383175.0       | 1.38                | 1720831     |
| 4   | 5         | 2246299.0       | 1.36                | 1653340     |
| 5   | 4         | 2135908.0       | 1.36                | 1571155     |
| 6   | 3         | 1998566.0       | 1.36                | 1474205     |
| 7   | 2         | 1926474.0       | 1.35                | 1422198     |

**Tip amount per day:**

SELECT

dayofweek(pickup\_datetime) AS dayofweek,  
round(sum(tip\_amount ),2) as tip\_amount,  
round(avg(tip\_amount ),2) as avg\_tip\_amount,  
count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | tip_amount | avg_tip_amount | no_of_trips |
|-----|-----------|------------|----------------|-------------|
| 1   | 7         | 2347336.88 | 1.16           | 2017033     |
| 2   | 6         | 2156131.71 | 1.15           | 1881905     |
| 3   | 1         | 2002705.64 | 1.16           | 1720831     |
| 4   | 5         | 1937998.89 | 1.17           | 1653340     |
| 5   | 4         | 1814424.37 | 1.15           | 1571155     |
| 6   | 3         | 1677582.57 | 1.14           | 1474205     |
| 7   | 2         | 1561937.15 | 1.1            | 1422198     |

**Hourly statistics:**

**Number of trips in each hour of the day:**

SELECT

hour(pickup\_datetime) AS hour,  
count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | no_of_trips |
|-----|------|-------------|
| 1   | 18   | 831658      |
| 2   | 19   | 794574      |
| 3   | 17   | 761054      |
| 4   | 20   | 696765      |
| 5   | 16   | 685998      |
| 6   | 21   | 650619      |
| 7   | 15   | 629425      |
| 8   | 22   | 605892      |
| 9   | 14   | 562230      |
| 10  | 23   | 552187      |
| 11  | 9    | 507294      |
| 12  | 13   | 495082      |
| 13  | 12   | 491282      |
| 14  | 11   | 484483      |
| 15  | 10   | 484299      |
| 16  | 8    | 481851      |
| 17  | 0    | 457892      |
| 18  | 1    | 351659      |
| 19  | 7    | 323088      |
| 20  | 2    | 256111      |
| 21  | 3    | 199872      |
| 22  | 4    | 165959      |
| 23  | 6    | 159484      |
| 24  | 5    | 111909      |

### Hourly earnings:

SELECT

hour(pickup\_datetime) AS hour,

ROUND(SUM(total\_amount ),2)

AS Total\_amt,

ROUND(AVG(total\_amount ),2)

AS Avg\_Total\_amt,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|------|---------------|---------------|-------------|
| 1   | 18   | 1.153654037E7 | 13.87         | 831658      |
| 2   | 17   | 1.102304354E7 | 14.48         | 761054      |
| 3   | 19   | 1.069822696E7 | 13.46         | 794574      |
| 4   | 16   | 1.017848547E7 | 14.84         | 685998      |
| 5   | 20   | 9256604.1     | 13.29         | 696765      |
| 6   | 15   | 9011950.87    | 14.32         | 629425      |
| 7   | 21   | 8911384.34    | 13.7          | 650619      |
| 8   | 22   | 8618224.43    | 14.22         | 605892      |
| 9   | 14   | 7981487.25    | 14.2          | 562230      |
| 10  | 23   | 7887496.04    | 14.28         | 552187      |
| 11  | 9    | 7346380.71    | 14.48         | 507294      |
| 12  | 8    | 7091232.02    | 14.72         | 481851      |
| 13  | 13   | 6972468.89    | 14.08         | 495082      |
| 14  | 12   | 6940096.17    | 14.13         | 491282      |
| 15  | 11   | 6856550.32    | 14.15         | 484483      |
| 16  | 10   | 6851400.82    | 14.15         | 484299      |
| 17  | 0    | 6433471.36    | 14.05         | 457892      |
| 18  | 1    | 4896788.48    | 13.92         | 351659      |
| 19  | 7    | 4885488.93    | 15.12         | 323088      |
| 20  | 2    | 3647799.4     | 14.24         | 256111      |
| 21  | 3    | 2959378.73    | 14.81         | 199872      |
| 22  | 6    | 2701014.94    | 16.94         | 159484      |
| 23  | 4    | 2576237.7     | 15.52         | 165959      |
| 24  | 5    | 1948829.87    | 17.41         | 111909      |

**Number of passengers every hour:**

SELECT

hour(pickup\_datetime) AS hour,

round(sum(passenger\_count ),2) as no\_of\_passenger,

round(avg(passenger\_count ),2) as avg\_passenger\_count,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC



| Row | hour | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|------|-----------------|---------------------|-------------|
| 1   | 18   | 1139293.0       | 1.37                | 831658      |
| 2   | 19   | 1091827.0       | 1.37                | 794574      |
| 3   | 17   | 1036238.0       | 1.36                | 761054      |
| 4   | 20   | 958457.0        | 1.38                | 696765      |
| 5   | 16   | 928211.0        | 1.35                | 685998      |
| 6   | 21   | 898678.0        | 1.38                | 650619      |
| 7   | 15   | 853184.0        | 1.36                | 629425      |
| 8   | 22   | 833748.0        | 1.38                | 605892      |
| 9   | 23   | 761368.0        | 1.38                | 552187      |
| 10  | 14   | 758611.0        | 1.35                | 562230      |
| 11  | 9    | 673210.0        | 1.33                | 507294      |
| 12  | 13   | 666786.0        | 1.35                | 495082      |
| 13  | 12   | 662163.0        | 1.35                | 491282      |
| 14  | 11   | 648465.0        | 1.34                | 484483      |
| 15  | 10   | 646824.0        | 1.34                | 484299      |
| 16  | 8    | 641748.0        | 1.33                | 481851      |
| 17  | 0    | 637470.0        | 1.39                | 457892      |
| 18  | 1    | 493092.0        | 1.4                 | 351659      |
| 19  | 7    | 431071.0        | 1.33                | 323088      |
| 20  | 2    | 363380.0        | 1.42                | 256111      |
| 21  | 3    | 286799.0        | 1.43                | 199872      |
| 22  | 4    | 238698.0        | 1.44                | 165959      |
| 23  | 6    | 212855.0        | 1.33                | 159484      |
| 24  | 5    | 157361.0        | 1.41                | 111909      |

### Tip amount across each hour of the day:

SELECT

hour(pickup\_datetime) AS hour,

round(sum(tip\_amount ),2) as tip\_amount,

round(avg(tip\_amount ),2) as avg\_tip\_amount,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | tip_amount | avg_tip_amount | no_of_trips |
|-----|------|------------|----------------|-------------|
| 1   | 18   | 953255.88  | 1.15           | 831658      |
| 2   | 19   | 889162.26  | 1.12           | 794574      |
| 3   | 17   | 855830.2   | 1.12           | 761054      |
| 4   | 20   | 768955.04  | 1.1            | 696765      |
| 5   | 21   | 760919.92  | 1.17           | 650619      |

| Row | hour | tip_amount | avg_tip_amount | no_of_trips |
|-----|------|------------|----------------|-------------|
| 6   | 22   | 760489.2   | 1.26           | 605892      |
| 7   | 16   | 752415.88  | 1.1            | 685998      |
| 8   | 23   | 692303.34  | 1.25           | 552187      |
| 9   | 15   | 657061.16  | 1.04           | 629425      |
| 10  | 9    | 636302.11  | 1.25           | 507294      |
| 11  | 8    | 633900.6   | 1.32           | 481851      |
| 12  | 14   | 576564.99  | 1.03           | 562230      |
| 13  | 0    | 548440.13  | 1.2            | 457892      |
| 14  | 10   | 543318.58  | 1.12           | 484299      |
| 15  | 11   | 519510.02  | 1.07           | 484483      |
| 16  | 12   | 518409.9   | 1.06           | 491282      |
| 17  | 13   | 508556.33  | 1.03           | 495082      |
| 18  | 1    | 409605.11  | 1.16           | 351659      |
| 19  | 7    | 407922.02  | 1.26           | 323088      |
| 20  | 2    | 303610.04  | 1.19           | 256111      |
| 21  | 3    | 238903.91  | 1.2            | 199872      |
| 22  | 6    | 221458.51  | 1.39           | 159484      |
| 23  | 4    | 187717.33  | 1.13           | 165959      |
| 24  | 5    | 153504.75  | 1.37           | 111909      |

### Hourly speed of the trips:

```

SELECT
  EXTRACT(HOUR
FROM
  pickup_datetime) hour,
  ROUND(AVG(trip_distance / DATETIME_DIFF(dropoff_datetime,
    pickup_datetime,
    SECOND))*3600, 1) speed
FROM
  `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017`
WHERE
  trip_distance > 0
  AND dropoff_datetime > pickup_datetime
GROUP BY
  1
ORDER BY
  2 desc

```

| Row | hour | speed |
|-----|------|-------|
| 1   | 6    | 24.0  |
| 2   | 5    | 23.2  |

| Row | hour | speed |
|-----|------|-------|
| 3   | 4    | 20.4  |
| 4   | 3    | 18.7  |
| 5   | 2    | 17.4  |
| 6   | 1    | 16.8  |
| 7   | 7    | 16.3  |
| 8   | 0    | 16.1  |
| 9   | 23   | 15.5  |
| 10  | 22   | 15.2  |
| 11  | 21   | 14.7  |
| 12  | 10   | 14.4  |
| 13  | 11   | 14.3  |
| 14  | 9    | 14.3  |
| 15  | 13   | 14.0  |
| 16  | 12   | 14.0  |
| 17  | 20   | 13.8  |
| 18  | 8    | 13.6  |
| 19  | 14   | 13.0  |
| 20  | 19   | 12.9  |
| 21  | 15   | 12.5  |
| 22  | 18   | 12.1  |
| 23  | 16   | 11.8  |
| 24  | 17   | 11.6  |

### Vendor data analysis:

#### Number of Trips:

```
SELECT vendor_id,count(*) trips FROM [bigquery-public-
data.new_york_taxi_trips.tlc_green_trips_2017]
group by 1
```

| Row | vendor_id | trips   |
|-----|-----------|---------|
| 1   | 1         | 2293365 |
| 2   | 2         | 9447302 |

#### Total Earnings:

```
SELECT vendor_id,ROUND(SUM(FLOAT(total_amount)),2) AS total_earnings
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
group by 1
```

| Row | vendor_id | total_earnings |
|-----|-----------|----------------|
| 1   | 2         | 1.3562386062E8 |
| 2   | 1         | 3.158672108E7  |

#### Total Passenger Count:

```
SELECT vendor_id, SUM(passenger_count) total_passenger_served
FROM [bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017]
group by 1
```

| Row | vendor_id | total_passenger_served |
|-----|-----------|------------------------|
| 1   | 1         | 2688424                |
| 2   | 2         | 13331113               |

#### Per trip details:

```
SELECT vendor_id, round(AVG(trip_distance),2) avg_trip_distance,
round(AVG(passenger_count),2) avg_passenger_count,
round(AVG(tip_amount),2) avg_tip_per_ride,
round(AVG(total_amount),2) as average_fare_per_ride,
round(AVG(DATETIME_DIFF(dropoff_datetime,pickup_datetime,minute)),2) avg_time_per_trip
FROM `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017`
group by 1
```

| Metrics               | Vendor 1 | Vendor 2 |
|-----------------------|----------|----------|
| avg_trip_distnace     | 2.61     | 2.7      |
| avg_passenger_count   | 1.17     | 1.41     |
| avg_tip_per_ride      | 1.14     | 1.15     |
| average_fare_per_ride | 13.77    | 14.36    |
| avg_time_per_trip     | 12.48    | 22.39    |

#### Average Speed of trip:

```
SELECT
  vendor_id,
  ROUND(AVG(trip_distance / DATETIME_DIFF(dropoff_datetime,
    pickup_datetime,
    SECOND))*3600, 1) Avg_speed
FROM
  `bigquery-public-data.new_york_taxi_trips.tlc_green_trips_2017`
WHERE
  trip_distance > 0
  AND dropoff_datetime > pickup_datetime
GROUP BY 1
ORDER BY
```

2 desc

| Row | vendor_id | Avg_speed |
|-----|-----------|-----------|
| 1   | 1         | 19.3      |
| 2   | 2         | 13.0      |

#### Trip type by vendor id:

```
SELECT vendor_id,trip_type,count(*) as trip_type_count
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
```

```
group by 1,2
```

```
order by 3 desc
```

| Row | vendor_id | trip_type | trip_type_count |
|-----|-----------|-----------|-----------------|
| 1   | 2         | 1         | 9272941         |
| 2   | 1         | 1         | 2251601         |
| 3   | 2         | 2         | 174361          |
| 4   | 1         | 2         | 41737           |
| 5   | 1         | null      | 27              |

#### Number of trips by payment type: 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip

```
SELECT vendor_id,payment_type ,count(*) as trip_type_count
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
```

```
group by 1,2
```

```
order by 1
```

| Row | vendor_id | payment_type | trip_type_count |
|-----|-----------|--------------|-----------------|
| 1   | 1         | 2            | 1140409         |
| 2   | 1         | 1            | 1093902         |
| 3   | 1         | 3            | 36332           |
| 4   | 1         | 4            | 22180           |
| 5   | 1         | 5            | 542             |
| 6   | 2         | 4            | 7971            |
| 7   | 2         | 2            | 4600305         |
| 8   | 2         | 1            | 4820911         |
| 9   | 2         | 3            | 18115           |

#### Most Popular Pickup location for Vendors:

1:

```
SELECT vendor_id,pickup_location_id ,count(*) as pickup_location_count
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_green_trips_2017]
```

```
group by 1,2
```

```
having vendor_id=string(1)
```

order by 3 desc

limit 5

| Row | vendor_id | pickup_location_id | pickup_location_count |
|-----|-----------|--------------------|-----------------------|
| 1   | 1         | 74                 | 156932                |
| 2   | 1         | 41                 | 132016                |
| 3   | 1         | 75                 | 122873                |
| 4   | 1         | 82                 | 105672                |
| 5   | 1         | 255                | 91521                 |

2:

SELECT vendor\_id,pickup\_location\_id ,count(\*) as pickup\_location\_count

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1,2

having vendor\_id=string(2)

order by 3 desc

limit 5

| Row | vendor_id | pickup_location_id | pickup_location_count |
|-----|-----------|--------------------|-----------------------|
| 1   | 2         | 74                 | 558781                |
| 2   | 2         | 41                 | 531009                |
| 3   | 2         | 7                  | 517761                |
| 4   | 2         | 75                 | 500833                |
| 5   | 2         | 82                 | 409035                |

**Average tip amount by vendors:**

SELECT vendor\_id,avg(tip\_amount) as avg\_tip\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_green\_trips\_2017]

group by 1

order by 1

| Row | vendor_id | avg_tip_amount     |
|-----|-----------|--------------------|
| 1   | 1         | 1.1446779557548348 |
| 2   | 2         | 1.1509056077616033 |

## Yellow Taxi Analysis:

### Total Trips:

```
SELECT count(*) trips FROM `bigquery-public-data.new_york_taxi_trips.tlc_yellowtrips_2017`
```

**Trips: 113496874**

### Trips with Tip:

with total as (

```
SELECT count(*) total from `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017` ),  
trips_with_tip as (
```

```
SELECT count(*) trips_with_tip
```

```
FROM `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017`
```

```
where tip_amount>0)
```

```
SELECT trips_with_tip.trips_with_tip, round((trips_with_tip.trips_with_tip*100/total.total),2) as  
Percentage_of_trips
```

```
FROM total,trips_with_tip
```

**Trips with Tip: 73001022**

**Percentage of Trips:64.32 %**

### Trips without tip:

with total as (

```
SELECT count(*) total from `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017` ),  
trips_without_tip as (
```

```
SELECT count(*) trips_without_tip
```

```
FROM `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017` where tip_amount=0)
```

```
SELECT trips_without_tip.trips_without_tip,  
round((trips_without_tip.trips_without_tip*100/total.total),2) as Percentage_of_trips
```

```
FROM total,trips_without_tip
```

**Trips without Tip: 40495008**

**Percentage of trips:35.68 %**

### Number of trips during daytime(6 AM to 6 PM):

```
SELECT count(*) as Number_of_day_trips
```

```
FROM [bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017]
```

where HOUR(TIMESTAMP(pickup\_datetime)) >=6 and HOUR(TIMESTAMP(pickup\_datetime)) <=18

**Number of Trips: 70155085**

**Number of trips during the evening and night time:**

SELECT count(\*) as Number\_of\_night\_trips

FROM [bigquery-public-data.new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

where HOUR(TIMESTAMP(pickup\_datetime)) >18 or HOUR(TIMESTAMP(pickup\_datetime)) <6

**Number of Trips: 43341789**

**Total Earnings through 2017:**

SELECT ROUND(SUM(FLOAT(total\_amount)),2) AS total\_earnings

FROM [bigquery-public-data.new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

**Total Earnings: 1.85484433372E9**

**Total distance covered in 2017:**

SELECT ROUND(SUM(FLOAT(trip\_distance)),2) AS total\_earnings

FROM [bigquery-public-data.new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

**Total Distance : 332 million miles**

**Total Passenger served in 2017:**

SELECT SUM(passenger\_count) total\_passenger\_served FROM [bigquery-public-data.new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

**Total Passenger served : 184464740 (184 million)**

**Average of trip\_distance,passenger\_count,per\_ride\_earnings,per\_ride\_tip,average time per trip:**

SELECT round(AVG(trip\_distance),2) avg\_trip\_distance,

round(AVG(passenger\_count),2) avg\_passenger\_count,

round(AVG(tip\_amount),2) avg\_tip\_per\_ride,

round(AVG(total\_amount),2) as average\_fare\_per\_ride,

round(AVG(DATETIME\_DIFF(dropoff\_datetime,pickup\_datetime,minute)),2) avg\_time\_per\_trip

FROM `bigquery-public-data.new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017`



avg\_trip\_distnace: 2.93 miles  
avg\_passenger\_count : 1.63  
avg\_tip\_per\_ride : \$1.84  
average\_fare\_per\_ride : \$16.34  
avg\_time\_per\_trip : 16.32 mins

### Average speed of a trip:

```
SELECT  
    ROUND(AVG(trip_distance / DATETIME_DIFF(dropoff_datetime,  
        pickup_datetime,  
        SECOND))*3600, 1) Avg_speed  
FROM  
    `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017`  
WHERE  
    trip_distance > 0  
    AND dropoff_datetime > pickup_datetime  
ORDER BY  
    1 desc
```

**Avg\_Speed: 13.3**

### Number of trips by payment type: 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip

```
SELECT payment_type ,count(*) as trip_type_count  
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]  
group by 1  
order by 2 desc
```

| Row | payment_type | trip_type_count |
|-----|--------------|-----------------|
| 1   | 1            | 76244327        |
| 2   | 2            | 36479462        |
| 3   | 3            | 599001          |
| 4   | 4            | 174081          |
| 5   | 5            | 3               |

### Most important pickup location :

```
SELECT pickup_location_id ,count(*) as pickup_location_count  
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

group by 1

order by 2 desc

limit 10

| Row | pickup_location_id | pickup_location_count |
|-----|--------------------|-----------------------|
| 1   | 237                | 4442294               |
| 2   | 161                | 4274032               |
| 3   | 236                | 4023665               |
| 4   | 186                | 3973288               |
| 5   | 162                | 3863529               |
| 6   | 230                | 3841712               |
| 7   | 234                | 3702711               |
| 8   | 170                | 3632477               |
| 9   | 48                 | 3566794               |
| 10  | 79                 | 3327326               |

### Most important drop-off location:

```
SELECT dropoff_location_id ,count(*) as dropoff_location_count
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

group by 1

order by 2 desc

limit 10

| Row | dropoff_location_id | dropoff_location_count |
|-----|---------------------|------------------------|
| 1   | 161                 | 4197865                |
| 2   | 236                 | 4150955                |
| 3   | 237                 | 3908234                |
| 4   | 170                 | 3618960                |
| 5   | 230                 | 3522229                |
| 6   | 162                 | 3450268                |
| 7   | 234                 | 3171724                |
| 8   | 186                 | 3090148                |
| 9   | 48                  | 3075486                |
| 10  | 142                 | 2930684                |

### Tip amount analysis:

#### Average tip amount by passenger count:

```
SELECT passenger_count,avg(tip_amount) as avg_tip_amount
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

group by 1

order by 1

| Row | passenger_count | avg_tip_amount |
|-----|-----------------|----------------|
| 1   | 0               | 1.77           |
| 2   | 1               | 1.84           |
| 3   | 2               | 1.85           |

| Row | passenger_count | avg_tip_amount |
|-----|-----------------|----------------|
| 4   | 3               | 1.75           |
| 5   | 4               | 1.63           |
| 6   | 5               | 1.86           |
| 7   | 6               | 1.84           |
| 8   | 7               | 4.87           |
| 9   | 8               | 5.19           |
| 10  | 9               | 6.93           |
| 11  | 192             | 1.24           |

#### Average tip amount by pickup location:

SELECT pickup\_location\_id ,round(avg(tip\_amount),2) as avg\_tip\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

group by 1

order by 2 desc

limit 10

| Row | pickup_location_id | avg_tip_amount |
|-----|--------------------|----------------|
| 1   | 104                | 12.02          |
| 2   | 117                | 11.58          |
| 3   | 1                  | 10.21          |
| 4   | 176                | 9.11           |
| 5   | 86                 | 8.87           |
| 6   | 265                | 8.85           |
| 7   | 5                  | 8.37           |
| 8   | 172                | 7.36           |
| 9   | 109                | 7.33           |
| 10  | 30                 | 7.02           |

#### Average tip amount by droopff location:

SELECT dropoff\_location\_id ,round(avg(tip\_amount),2) as avg\_tip\_amount

FROM [bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

group by 1

order by 2 desc

limit 10

| Row | dropoff_location_id | avg_tip_amount |
|-----|---------------------|----------------|
| 1   | 104                 | 14.16          |
| 2   | 204                 | 12.34          |
| 3   | 1                   | 11.32          |
| 4   | 44                  | 11.15          |
| 5   | 5                   | 10.54          |
| 6   | 84                  | 9.71           |
| 7   | 176                 | 9.68           |

| Row | dropoff_location_id | avg_tip_amount |
|-----|---------------------|----------------|
| 8   | 109                 | 9.54           |
| 9   | 265                 | 9.07           |
| 10  | 172                 | 8.28           |

#### Highest paid trip of the year:

```
select vendor_id, pickup_datetime, dropoff_datetime, fare_amount , tolls_amount , total_amount
from [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
where tolls_amount in (SELECT max(tolls_amount )
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017] )
```

| Row | vendor_id | pickup_datetime     | dropoff_datetime    | fare_amount | tolls_amount | total_amount |
|-----|-----------|---------------------|---------------------|-------------|--------------|--------------|
| 1   | 2         | 2017-07-03T17:15:15 | 2017-07-03T17:35:57 | 27.5        | 5743.51      | 5772.81      |

#### Quarterly Analysis :

##### Highest earning quarter:

```
SELECT
  QUARTER(pickup_datetime) AS quarter,
  ROUND(SUM(total_amount ),2)
  AS Total_amt,
  ROUND(AVG(total_amount ),2)
  AS Avg_Total_amt,
  count(*) as no_of_trips
FROM
  [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
GROUP BY
  quarter
ORDER BY
  2 DESC
```

| Row | quarter | Total_amt      | Avg_Total_amt | no_of_trips |
|-----|---------|----------------|---------------|-------------|
| 1   | 2       | 4.9444152519E8 | 16.59         | 29805318    |
| 2   | 4       | 4.7052474262E8 | 16.47         | 28561604    |
| 3   | 1       | 4.611169838E8  | 15.81         | 29173695    |
| 4   | 3       | 4.2876108223E8 | 16.52         | 25956257    |

### Quarter with most number of trips:

```
SELECT

    QUARTER(pickup_datetime) AS quarter,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]

GROUP BY

    quarter

ORDER BY

    2 DESC
```

| Row | quarter | Total amt      | Avg_Total amt | no_of_trips |
|-----|---------|----------------|---------------|-------------|
| 1   | 2       | 4.9444152519E8 | 16.59         | 29805318    |
| 2   | 4       | 4.7052474262E8 | 16.47         | 28561604    |
| 3   | 1       | 4.611169838E8  | 15.81         | 29173695    |
| 4   | 3       | 4.2876108223E8 | 16.52         | 25956257    |

### Quarter with highest passenger served:

```
SELECT

    QUARTER(pickup_datetime) AS quarter,

    round(sum(passenger_count ),2) as no_of_passenger,

    round(avg(passenger_count ),2) as avg_passenger_count,

    count(*) as no_of_trips

FROM

    [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]

GROUP BY

    quarter

ORDER BY

    2 DESC
```

| Row | quarter | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|---------|-----------------|---------------------|-------------|
| 1   | 2       | 4.8389585E7     | 1.62                | 29805318    |
| 2   | 1       | 4.7375952E7     | 1.62                | 29173695    |
| 3   | 4       | 4.6278513E7     | 1.62                | 28561604    |
| 4   | 3       | 4.242069E7      | 1.63                | 25956257    |

### Tip amounts per quarter:

SELECT

QUARTER(pickup\_datetime) AS quarter,  
round(sum(tip\_amount ),2) as tip\_amount,  
round(avg(tip\_amount ),2) as avg\_tip\_amount,  
count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

quarter

ORDER BY

2 DESC

| Row | quarter | tip_amount    | avg_tip_amount | no_of_trips |
|-----|---------|---------------|----------------|-------------|
| 1   | 2       | 5.550919902E7 | 1.86           | 29805318    |
| 2   | 4       | 5.317480428E7 | 1.86           | 28561604    |
| 3   | 1       | 5.254278706E7 | 1.8            | 29173695    |
| 4   | 3       | 4.730183975E7 | 1.82           | 25956257    |

### Monthly statistics:

#### Number of trips :

SELECT

MONTH(pickup\_datetime) AS month,  
count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

month

ORDER BY

2 DESC

| Row | month | no_of_trips |
|-----|-------|-------------|
| 1   | 3     | 10294633    |
| 2   | 5     | 10102128    |
| 3   | 4     | 10046197    |
| 4   | 10    | 9768740     |
| 5   | 1     | 9710231     |
| 6   | 6     | 9656993     |

| Row | month | no_of_trips |
|-----|-------|-------------|
| 7   | 12    | 9508086     |
| 8   | 11    | 9284778     |
| 9   | 2     | 9168831     |
| 10  | 9     | 8945574     |
| 11  | 7     | 8588486     |
| 12  | 8     | 8422197     |

### Monthly earnings:

SELECT

MONTH(pickup\_datetime) AS month,

ROUND(SUM(total\_amount ),2)

AS Total\_amt,

ROUND(AVG(total\_amount ),2)

AS Avg\_Total\_amt,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

month

ORDER BY

2 DESC

| Row | month | Total_amt      | Avg_Total_amt | no_of_trips |
|-----|-------|----------------|---------------|-------------|
| 1   | 5     | 1.6939136559E8 | 16.77         | 10102128    |
| 2   | 3     | 1.6701692953E8 | 16.22         | 10294633    |
| 3   | 4     | 1.6429674458E8 | 16.35         | 10046197    |
| 4   | 10    | 1.6354139304E8 | 16.74         | 9768740     |
| 5   | 6     | 1.6075341501E8 | 16.65         | 9656993     |
| 6   | 12    | 1.5389932022E8 | 16.19         | 9508086     |
| 7   | 11    | 1.5308402935E8 | 16.49         | 9284778     |
| 8   | 1     | 1.5076876101E8 | 15.53         | 9710231     |
| 9   | 9     | 1.4939402406E8 | 16.7          | 8945574     |
| 10  | 2     | 1.4333129325E8 | 15.63         | 9168831     |
| 11  | 7     | 1.4062107437E8 | 16.37         | 8588486     |
| 12  | 8     | 1.387459838E8  | 16.47         | 8422197     |

### Passengers served per month:

SELECT

MONTH(pickup\_datetime) AS month,

round(sum(passenger\_count ),2) as no\_of\_passenger,

round(avg(passenger\_count ),2) as avg\_passenger\_count,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

month

ORDER BY

2 DESC

| Row | month | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|-------|-----------------|---------------------|-------------|
| 1   | 3     | 1.6659032E7     | 1.62                | 10294633    |
| 2   | 5     | 1.6363377E7     | 1.62                | 10102128    |
| 3   | 4     | 1.6343798E7     | 1.63                | 10046197    |
| 4   | 1     | 1.5817806E7     | 1.63                | 9710231     |
| 5   | 10    | 1.5800531E7     | 1.62                | 9768740     |
| 6   | 6     | 1.568241E7      | 1.62                | 9656993     |
| 7   | 12    | 1.5520752E7     | 1.63                | 9508086     |
| 8   | 11    | 1.495723E7      | 1.61                | 9284778     |
| 9   | 2     | 1.4899114E7     | 1.62                | 9168831     |
| 10  | 9     | 1.4558792E7     | 1.63                | 8945574     |
| 11  | 7     | 1.4096141E7     | 1.64                | 8588486     |
| 12  | 8     | 1.3765757E7     | 1.63                | 8422197     |

**Tip data per month:**

SELECT

MONTH(pickup\_datetime) AS month,

round(sum(tip\_amount ),2) as tip\_amount,

round(avg(tip\_amount ),2) as avg\_tip\_amount,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

month

ORDER BY

2 DESC

| Row | month | tip_amount    | avg_tip_amount | no_of_trips |
|-----|-------|---------------|----------------|-------------|
| 1   | 3     | 1.911871493E7 | 1.86           | 10294633    |
| 2   | 5     | 1.905249014E7 | 1.89           | 10102128    |
| 3   | 10    | 1.839586948E7 | 1.88           | 9768740     |
| 4   | 4     | 1.831623642E7 | 1.82           | 10046197    |



| Row | month | tip_amount    | avg_tip_amount | no_of_trips |
|-----|-------|---------------|----------------|-------------|
| 5   | 6     | 1.814047246E7 | 1.88           | 9656993     |
| 6   | 11    | 1.75315074E7  | 1.89           | 9284778     |
| 7   | 12    | 1.72474274E7  | 1.81           | 9508086     |
| 8   | 1     | 1.700522893E7 | 1.75           | 9710231     |
| 9   | 9     | 1.668538263E7 | 1.87           | 8945574     |
| 10  | 2     | 1.64188432E7  | 1.79           | 9168831     |
| 11  | 7     | 1.545848338E7 | 1.8            | 8588486     |
| 12  | 8     | 1.515797374E7 | 1.8            | 8422197     |

#### Day wise statistics:

##### Number of trips:

SELECT

dayofweek(pickup\_datetime) AS dayofweek,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | no_of_trips |
|-----|-----------|-------------|
| 1   | 6         | 17387669    |
| 2   | 5         | 17069946    |
| 3   | 7         | 16878300    |
| 4   | 4         | 16743579    |
| 5   | 3         | 15824336    |
| 6   | 1         | 14996188    |
| 7   | 2         | 14596856    |

#### Day wise earnings:

SELECT

dayofweek(pickup\_datetime) AS dayofweek,

ROUND(SUM(total\_amount ),2)

AS Total\_amt,

ROUND(AVG(total\_amount ),2)

AS Avg\_Total\_amt,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | Total_amt      | Avg_Total_amt | no_of_trips |
|-----|-----------|----------------|---------------|-------------|
| 1   | 5         | 2.8788259204E8 | 16.86         | 17069946    |
| 2   | 6         | 2.8763827969E8 | 16.54         | 17387669    |
| 3   | 4         | 2.7881645746E8 | 16.65         | 16743579    |
| 4   | 3         | 2.5854841113E8 | 16.34         | 15824336    |
| 5   | 7         | 2.5755686686E8 | 15.26         | 1687830     |
| 6   | 1         | 2.438897822E8  | 16.26         | 14996188    |
| 7   | 2         | 2.4051194443E8 | 16.48         | 14596856    |

**Passenger spread across each day:**

SELECT

dayofweek(pickup\_datetime) AS dayofweek,

round(sum(passenger\_count ),2) as no\_of\_passenger,

round(avg(passenger\_count ),2) as avg\_passenger\_count,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

dayofweek

ORDER BY

2 DESC

| Row | dayofweek | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|-----------|-----------------|---------------------|-------------|
| 1   | 7         | 2.8391742E7     | 1.68                | 16878300    |
| 2   | 6         | 2.8141722E7     | 1.62                | 17387669    |
| 3   | 5         | 2.7314421E7     | 1.6                 | 17069946    |
| 4   | 4         | 2.6742097E7     | 1.6                 | 16743579    |
| 5   | 3         | 2.531922E7      | 1.6                 | 15824336    |
| 6   | 1         | 2.5095744E7     | 1.67                | 14996188    |
| 7   | 2         | 2.3459794E7     | 1.61                | 14596856    |

**Tip amount per day:**

SELECT

dayofweek(pickup\_datetime) AS dayofweek,

```

round(sum(tip_amount ),2) as tip_amount,
round(avg(tip_amount ),2) as avg_tip_amount,
count(*) as no_of_trips

```

FROM

```
[bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

GROUP BY

```
dayofweek
```

ORDER BY

```
2 DESC
```

| Row | dayofweek | tip_amount    | avg_tip_amount | no_of_trips |
|-----|-----------|---------------|----------------|-------------|
| 1   | 5         | 3.333347979E7 | 1.95           | 17069946    |
| 2   | 6         | 3.241416272E7 | 1.86           | 17387669    |
| 3   | 4         | 3.232978634E7 | 1.93           | 16743579    |
| 4   | 3         | 2.976053926E7 | 1.88           | 15824336    |
| 5   | 2         | 2.723601202E7 | 1.87           | 14596856    |
| 6   | 7         | 2.72296121E7  | 1.61           | 16878300    |
| 7   | 1         | 2.622503788E7 | 1.75           | 14996188    |

### Hourly statistics:

#### Number of trips in each hour of the day:

SELECT

```
hour(pickup_datetime) AS hour,
```

```
count(*) as no_of_trips
```

FROM

```
[bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

GROUP BY

```
hour
```

ORDER BY

```
2 DESC
```

| Row | hour | no_of_trips |
|-----|------|-------------|
| 1   | 18   | 7087838     |
| 2   | 19   | 6984603     |
| 3   | 20   | 6433718     |
| 4   | 21   | 6365977     |
| 5   | 17   | 6216293     |
| 6   | 22   | 6070820     |
| 7   | 14   | 6045924     |
| 8   | 15   | 5947460     |
| 9   | 13   | 5775167     |

|    |    |         |
|----|----|---------|
| 10 | 12 | 5733105 |
| 11 | 16 | 5459753 |
| 12 | 11 | 5435173 |
| 13 | 9  | 5265431 |
| 14 | 10 | 5204985 |
| 15 | 8  | 5189559 |
| 16 | 23 | 5084149 |
| 17 | 7  | 4231353 |
| 18 | 0  | 3836098 |
| 19 | 1  | 2789104 |
| 20 | 6  | 2563044 |
| 21 | 2  | 2021190 |
| 22 | 3  | 1461537 |
| 23 | 5  | 1162875 |
| 24 | 4  | 1131718 |

### Hourly earnings:

SELECT

hour(pickup\_datetime) AS hour,

ROUND(SUM(total\_amount ),2)

AS Total\_amt,

ROUND(AVG(total\_amount ),2)

AS Avg\_Total\_amt,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | Total_amt      | Avg_Total_amt | no_of_trips |
|-----|------|----------------|---------------|-------------|
| 1   | 18   | 1.1495686606E8 | 16.22         | 7087838     |
| 2   | 19   | 1.1077773581E8 | 15.86         | 6984603     |
| 3   | 17   | 1.073689509E8  | 17.27         | 6216293     |
| 4   | 21   | 1.0400913628E8 | 16.34         | 6365977     |
| 5   | 22   | 1.0271558487E8 | 16.92         | 6070820     |
| 6   | 20   | 1.0235562936E8 | 15.91         | 6433718     |
| 7   | 14   | 1.0175266413E8 | 16.83         | 6045924     |
| 8   | 15   | 9.964842665E7  | 16.75         | 5947460     |
| 9   | 16   | 9.834991651E7  | 18.01         | 5459753     |
| 10  | 13   | 9.550714506E7  | 16.54         | 5775167     |
| 11  | 12   | 9.087204337E7  | 15.85         | 5733105     |
| 12  | 23   | 8.683830582E7  | 17.08         | 5084149     |
| 13  | 11   | 8.537982795E7  | 15.71         | 5435173     |

| Row | hour | Total_amt     | Avg_Total_amt | no_of_trips |
|-----|------|---------------|---------------|-------------|
| 14  | 10   | 8.167212575E7 | 15.69         | 5204985     |
| 15  | 9    | 8.101259186E7 | 15.39         | 5265431     |
| 16  | 8    | 7.857432432E7 | 15.14         | 5189559     |
| 17  | 0    | 6.445948401E7 | 16.8          | 3836098     |
| 18  | 7    | 6.307429117E7 | 14.91         | 4231353     |
| 19  | 1    | 4.479261258E7 | 16.06         | 2789104     |
| 20  | 6    | 4.087017544E7 | 15.95         | 2563044     |
| 21  | 2    | 3.204914847E7 | 15.86         | 2021190     |
| 22  | 3    | 2.35512581E7  | 16.11         | 1461537     |
| 23  | 5    | 2.327283452E7 | 20.01         | 1162875     |
| 24  | 4    | 2.098325482E7 | 18.54         | 1131718     |

**Number of passengers every hour:**

SELECT

hour(pickup\_datetime) AS hour,

round(sum(passenger\_count ),2) as no\_of\_passenger,

round(avg(passenger\_count ),2) as avg\_passenger\_count,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | no_of_passenger | avg_passenger_count | no_of_trips |
|-----|------|-----------------|---------------------|-------------|
| 1   | 18   | 1.1544204E7     | 1.63                | 7087838     |
| 2   | 19   | 1.1422412E7     | 1.64                | 6984603     |
| 3   | 20   | 1.0535791E7     | 1.64                | 6433718     |
| 4   | 21   | 1.0515458E7     | 1.65                | 6365977     |
| 5   | 17   | 1.0120329E7     | 1.63                | 6216293     |
| 6   | 22   | 1.0073917E7     | 1.66                | 6070820     |
| 7   | 14   | 9858790.0       | 1.63                | 6045924     |
| 8   | 15   | 9728363.0       | 1.64                | 5947460     |
| 9   | 13   | 9405473.0       | 1.63                | 5775167     |
| 10  | 12   | 9316732.0       | 1.63                | 5733105     |
| 11  | 16   | 8899076.0       | 1.63                | 5459753     |
| 12  | 11   | 8822118.0       | 1.62                | 5435173     |
| 13  | 23   | 8438205.0       | 1.66                | 5084149     |
| 14  | 10   | 8377884.0       | 1.61                | 5204985     |
| 15  | 9    | 8343214.0       | 1.58                | 5265431     |
| 16  | 8    | 8156288.0       | 1.57                | 5189559     |
| 17  | 7    | 6576410.0       | 1.55                | 4231353     |
| 18  | 0    | 6365901.0       | 1.66                | 3836098     |
| 19  | 1    | 4631031.0       | 1.66                | 2789104     |

|    |   |           |      |         |
|----|---|-----------|------|---------|
| 20 | 6 | 3914787.0 | 1.53 | 2563044 |
| 21 | 2 | 3352144.0 | 1.66 | 2021190 |
| 22 | 3 | 2417582.0 | 1.65 | 1461537 |
| 23 | 4 | 1848676.0 | 1.63 | 1131718 |
| 24 | 5 | 1799955.0 | 1.55 | 1162875 |

#### Tip amount across each hour of the day:

SELECT

hour(pickup\_datetime) AS hour,

round(sum(tip\_amount ),2) as tip\_amount,

round(avg(tip\_amount ),2) as avg\_tip\_amount,

count(\*) as no\_of\_trips

FROM

[bigquery-public-data:new\_york\_taxi\_trips.tlc\_yellow\_trips\_2017]

GROUP BY

hour

ORDER BY

2 DESC

| Row | hour | tip_amount    | avg_tip_amount | no_of_trips |
|-----|------|---------------|----------------|-------------|
| 1   | 18   | 1.299922648E7 | 1.83           | 7087838     |
| 2   | 19   | 1.27104134E7  | 1.82           | 6984603     |
| 3   | 21   | 1.240392992E7 | 1.95           | 6365977     |
| 4   | 20   | 1.208436156E7 | 1.88           | 6433718     |
| 5   | 22   | 1.200854352E7 | 1.98           | 6070820     |
| 6   | 17   | 1.178739269E7 | 1.9            | 6216293     |
| 7   | 14   | 1.106308748E7 | 1.83           | 6045924     |
| 8   | 15   | 1.085885833E7 | 1.83           | 5947460     |
| 9   | 16   | 1.061994738E7 | 1.95           | 5459753     |
| 10  | 13   | 1.033803262E7 | 1.79           | 5775167     |
| 11  | 23   | 1.010537128E7 | 1.99           | 5084149     |
| 12  | 12   | 9968914.19    | 1.74           | 5733105     |
| 13  | 11   | 9377103.58    | 1.73           | 5435173     |
| 14  | 9    | 9375384.41    | 1.78           | 5265431     |
| 15  | 8    | 9343471.78    | 1.8            | 5189559     |
| 16  | 10   | 9030285.11    | 1.73           | 5204985     |
| 17  | 7    | 7443715.06    | 1.76           | 4231353     |
| 18  | 0    | 7299692.13    | 1.9            | 3836098     |
| 19  | 1    | 4915186.47    | 1.76           | 2789104     |
| 20  | 6    | 4581890.14    | 1.79           | 2563044     |
| 21  | 2    | 3334226.91    | 1.65           | 2021190     |
| 22  | 5    | 2511586.77    | 2.16           | 1162875     |
| 23  | 3    | 2362501.31    | 1.62           | 1461537     |
| 24  | 4    | 2005507.59    | 1.77           | 1131718     |

### Hourly speed of the trips:

```
SELECT  
  
    EXTRACT(HOUR  
  
FROM  
  
    pickup_datetime) hour,  
  
    ROUND(AVG(trip_distance / DATETIME_DIFF(dropoff_datetime,  
        pickup_datetime,  
        SECOND))*3600, 1) speed  
  
FROM  
  
    `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017`  
  
WHERE  
  
    trip_distance > 0  
  
    AND dropoff_datetime > pickup_datetime  
  
GROUP BY  
  
    1  
  
ORDER BY  
  
    2 desc
```

| Row | hour | speed |
|-----|------|-------|
| 1   | 5    | 23.3  |
| 2   | 4    | 21.9  |
| 3   | 3    | 19.3  |
| 4   | 2    | 18.0  |
| 5   | 6    | 17.9  |
| 6   | 1    | 17.5  |
| 7   | 0    | 16.9  |
| 8   | 23   | 15.9  |
| 9   | 22   | 14.7  |
| 10  | 21   | 14.1  |
| 11  | 7    | 13.8  |
| 12  | 20   | 13.5  |
| 13  | 16   | 12.5  |
| 14  | 19   | 12.3  |
| 15  | 13   | 11.8  |
| 16  | 15   | 11.8  |
| 17  | 9    | 11.7  |
| 18  | 17   | 11.7  |
| 19  | 10   | 11.7  |
| 20  | 8    | 11.6  |
| 21  | 14   | 11.6  |
| 22  | 12   | 11.5  |
| 23  | 18   | 11.5  |
| 24  | 11   | 11.3  |

### Vendor data analysis:

#### Number of Trips:

```
SELECT vendor_id,count(*) trips FROM [bigquery-public-  
data.new_york_taxi_trips.tlc_yellow_trips_2017]  
  
group by 1
```

| Row | vendor_id | trips    |
|-----|-----------|----------|
| 1   | 2         | 62336757 |
| 2   | 1         | 51160117 |

#### Total Earnings:

```
SELECT vendor_id,ROUND(SUM(FLOAT(total_amount)),2) AS total_earnings  
  
FROM [bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017]  
  
group by 1
```

| Row | vendor_id | total_earnings  |
|-----|-----------|-----------------|
| 1   | 2         | 1.02830515236E9 |
| 2   | 1         | 8.2653918146E8  |

#### Total Passenger Count:

```
SELECT vendor_id, SUM(passenger_count) total_passenger_served  
  
FROM [bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017]  
  
group by 1
```

| Row | vendor_id | total_passenger_served |
|-----|-----------|------------------------|
| 1   | 1         | 64157145               |
| 2   | 2         | 120307595              |

#### Per trip details:

```
SELECT vendor_id,round(AVG(trip_distance),2) avg_trip_distance,  
round(AVG(passenger_count),2) avg_passenger_count,  
round(AVG(tip_amount),2) avg_tip_per_ride,  
round(AVG(total_amount),2) as average_fare_per_ride,  
round(AVG(DATETIME_DIFF(dropoff_datetime,pickup_datetime,minute)),2) avg_time_per_trip  
FROM `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017`  
  
group by 1
```



| Metrics               | Vendor 1 | Vendor 2 |
|-----------------------|----------|----------|
| avg_trip_distnace     | 2.84     | 3.01     |
| avg_passenger_count   | 1.25     | 1.93     |
| avg_tip_per_ride      | 1.8      | 1.87     |
| average_fare_per_ride | 16.16    | 16.5     |
| avg_time_per_trip     | 13.57    | 18.58    |

### Average Speed of trip:

```

SELECT
  vendor_id,
  ROUND(AVG(trip_distance / DATETIME_DIFF(dropoff_datetime,
    pickup_datetime,
    SECOND))*3600, 1) Avg_speed
FROM
  `bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017`
WHERE
  trip_distance > 0
  AND dropoff_datetime > pickup_datetime
GROUP BY 1
ORDER BY
  2 desc

```

| Row | vendor_id | Avg_speed |
|-----|-----------|-----------|
| 1   | 1         | 15.2      |
| 2   | 2         | 11.7      |

### Number of trips by payment type: 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip

```

SELECT vendor_id,payment_type ,count(*) as trip_type_count
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
group by 1,2
order by 1

```

| Row | vendor_id | payment_type | trip_type_count |
|-----|-----------|--------------|-----------------|
| 1   | 1         | 4            | 157034          |
| 2   | 1         | 1            | 34201404        |
| 3   | 1         | 3            | 560339          |
| 4   | 1         | 2            | 16241337        |
| 5   | 1         | 5            | 3               |

| Row | vendor_id | payment_type | trip_type_count |
|-----|-----------|--------------|-----------------|
| 6   | 2         | 4            | 17047           |
| 7   | 2         | 3            | 38662           |
| 8   | 2         | 1            | 42042923        |
| 9   | 2         | 2            | 20238125        |

Most Popular Pickup location for Vendors:

1:

```
SELECT vendor_id,pickup_location_id ,count(*) as pickup_location_count
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
group by 1,2
having vendor_id=string(1)
order by 3 desc
limit 5
```

| Row | vendor_id | pickup_location_id | pickup_location_count |
|-----|-----------|--------------------|-----------------------|
| 1   | 1         | 237                | 2058182               |
| 2   | 1         | 161                | 1883145               |
| 3   | 1         | 236                | 1856896               |
| 4   | 1         | 186                | 1856054               |
| 5   | 1         | 230                | 1691971               |

2:

```
SELECT vendor_id,pickup_location_id ,count(*) as pickup_location_count
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
group by 1,2
having vendor_id=string(2)
order by 3 desc
limit 5
```

| Row | vendor_id | pickup_location_id | pickup_location_count |
|-----|-----------|--------------------|-----------------------|
| 1   | 2         | 161                | 2390887               |
| 2   | 2         | 237                | 2384112               |
| 3   | 2         | 162                | 2177867               |
| 4   | 2         | 236                | 2166769               |
| 5   | 2         | 230                | 2149741               |

**Average tip amount by vendors:**

```
SELECT vendor_id,avg(tip_amount) as avg_tip_amount
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
group by 1
```

order by 1

| Row | vendor_id | avg_tip_amount     |
|-----|-----------|--------------------|
| 1   | 1         | 1.7965909931372692 |
| 2   | 2         | 1.870723314978881  |

### Faulty records: (Yellow Taxi)

```
SELECT count(*)
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

```
where dropoff_datetime < pickup_datetime
```

Number of records: **1447**

### Faulty records: (Green Taxi)

```
SELECT count(*)
```

```
FROM [bigquery-public-data:new_york_taxi_trips.tlc_yellow_trips_2017]
```

```
where dropoff_datetime < pickup_datetime
```

Number of records: **51**