

Logic For Final Submission

<Explain the queries, list them and attach screenshots after successful execution of queries>

We're supposed to write seven queries for the tasks listed from 5 to 11. Every task, query and the logic behind it is given below-

1. Calculate the total number of different drivers for each customer.

```
SELECT customer_id, count(DISTINCT driver_id)
FROM booking_data
GROUP BY customer_id
ORDER BY customer_id ASC;
```

Since, we are asked to calculate number of different users for each customer, we have to select two columns customer-id and driver-id. The drivers are supposed to be **counted distinctly** for every customer; this is why we have grouped the results by customer-id.

2. Calculate the total rides taken by each customer.

```
SELECT customer_id, count(DISTINCT booking_id)
FROM booking_data
GROUP BY customer_id
ORDER BY customer_id ASC;
```

Here, we have calculated total number of rides by keeping booking-id as the counter field. Again since the rides are to be calculated as per distinct customers, we have grouped results by customer-id.

3. Find the total visits made by each customer on the booking page and the total 'Book Now' button presses. This can show the conversion ratio.

```
SELECT (SUM(CASE WHEN button_id = "fcb68aa-1231-11eb-adc1-0242ac120002" AND
is_button_click = 'Yes' THEN 1 END) /
SUM(CASE WHEN page_id = "e7bc5fb2-1231-11eb-adc1- 0242ac120002" AND
is_page_view = 'Yes' THEN 1 END))
AS conversion_ratio
FROM clickstream_data;
```

With the given button-id and page-id, we have picked relevant data from the clickstream-data table. We then divided the sum of total 'Book Now' button presses with the sum of total 'Booking' page visit.

4. Calculate the count of all trips done on black cabs.

```
SELECT COUNT (booking_id)
FROM booking_data
WHERE cab_color in ('black');
```

We simply picked the count of all the booking-ids and filtered the results by cab-color 'black'. It gave us the count 72.

5. Calculate the total amount of tips given date wise to all drivers by customers.

```
SELECT date_format(pickup_timestamp,'yyyy-MM-dd'),
SUM(tip_amount)
FROM booking_data
GROUP BY date_format(pickup_timestamp,'yyyy-MM-dd');
```

To calculate tips amount, we picked the tip-amount field from table booking-data. Then to show them cumulatively on the basis of date, we grouped the results by date filtered from the pickup-timestamp field.

6. Calculate the total count of all the bookings with ratings lower than 2 as given by customers in a particular month.

```
SELECT date_format(pickup_timestamp,'yyyy-MM'),
COUNT(rating_by_customer)
FROM booking_data
WHERE rating_by_customer < 2
GROUP BY date_format(pickup_timestamp,'yyyy-MM');
```

To find all the bookings with below 2 ratings we picked out all the rows with value less than 2 for field 'rating-by-customer'. Then we grouped the results by month filtered from the pickup-timestamp field.

7. Calculate the count of total iOS users.

```
SELECT os_version, COUNT(DISTINCT customer_id)
FROM clickstream_data
WHERE os_version in ('iOS')
GROUP BY os_version;
```

We simply used the value of field 'os-version' from the table 'clickstream-data'. Then all the values that had 'iOS' in its strings were counted for the final results.