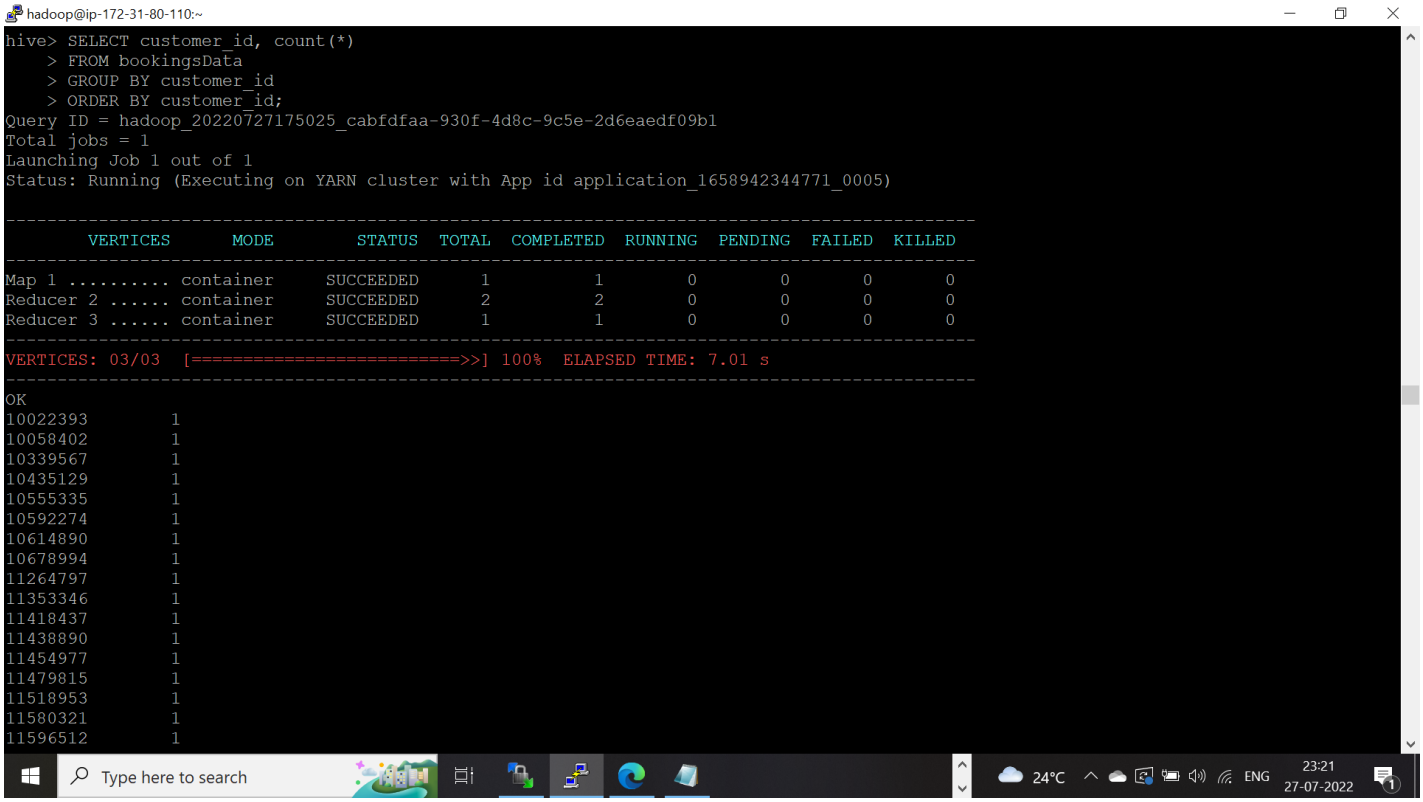
Queries

1. **Hive Query for Task 5:**
   Calculate the total number of different drivers for each customer.

```
SELECT customer_id, count(*)
FROM bookingsData
GROUP BY customer_id
ORDER BY customer_id;
```

Screenshot after executing Query:

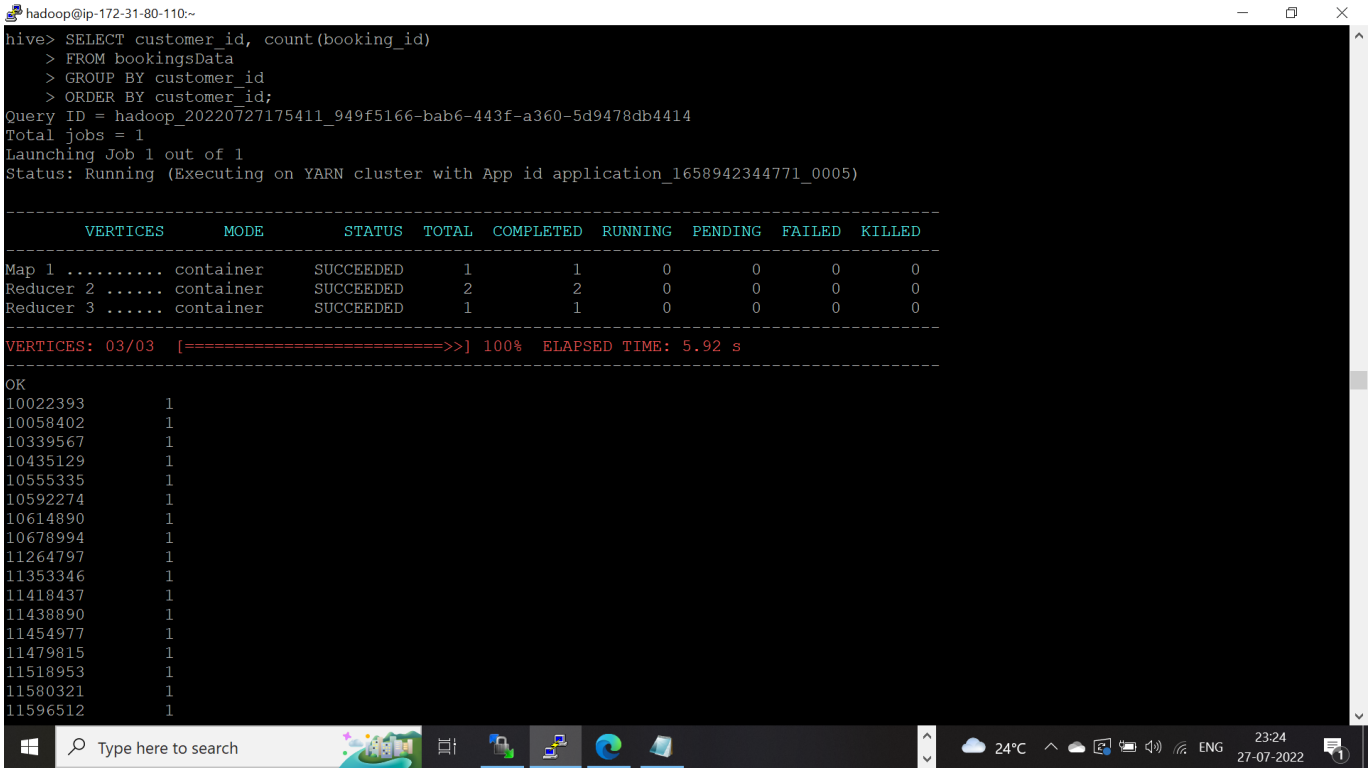


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

2. **Hive Query for Task 6:**
   Calculate the total rides taken by each customer.

```
SELECT customer_id,
count(booking_id)
FROM bookingsData
GROUP BY customer_id
ORDER BY customer_id;
```

Screenshot after executing Query:

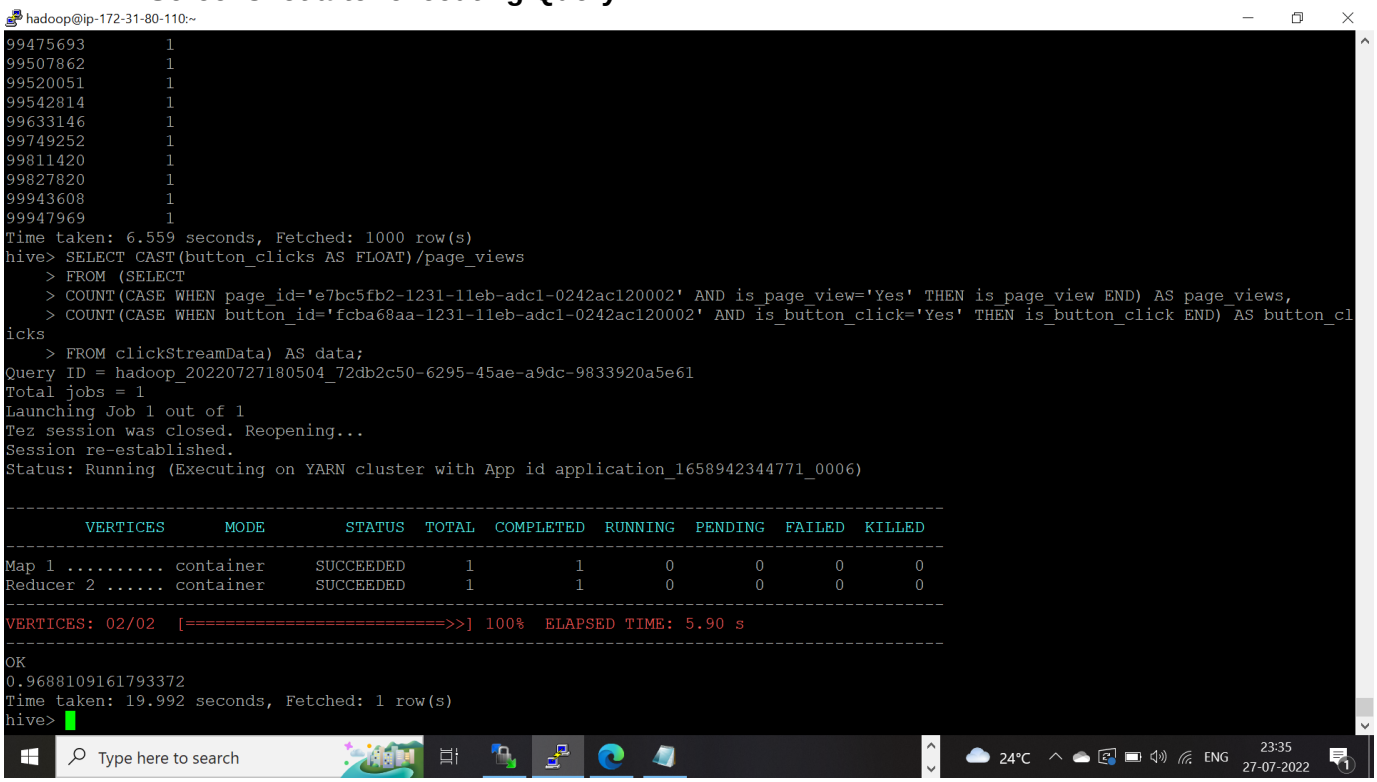


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

3. **Hive Query for Task 7:**
   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. The booking page id is 'e7bc5fb2-1231-11eb-adc1-0242ac120002'. The Book Now button id is 'fcba68aa-1231-11eb-adc1-0242ac120002'. You also need to calculate the conversion ratio as part of this task. Conversion ratio can be calculated as Total 'Book Now' Button Press/Total Visits made by customer on the booking page.

```
SELECT CAST(button_clicks AS FLOAT)/page_views
FROM (SELECT
COUNT(CASE WHEN page_id='e7bc5fb2-1231-11eb-
adc1-0242ac120002' AND is_page_view='Yes' THEN
is_page_view END) AS page_views,
COUNT(CASE WHEN button_id='fcba68aa-1231-11eb-
adc1-0242ac120002' AND is_button_click='Yes'
THEN is_button_click END) AS button_clicks
FROM clickStreamData) AS data;
```

Screenshot after executing Query:

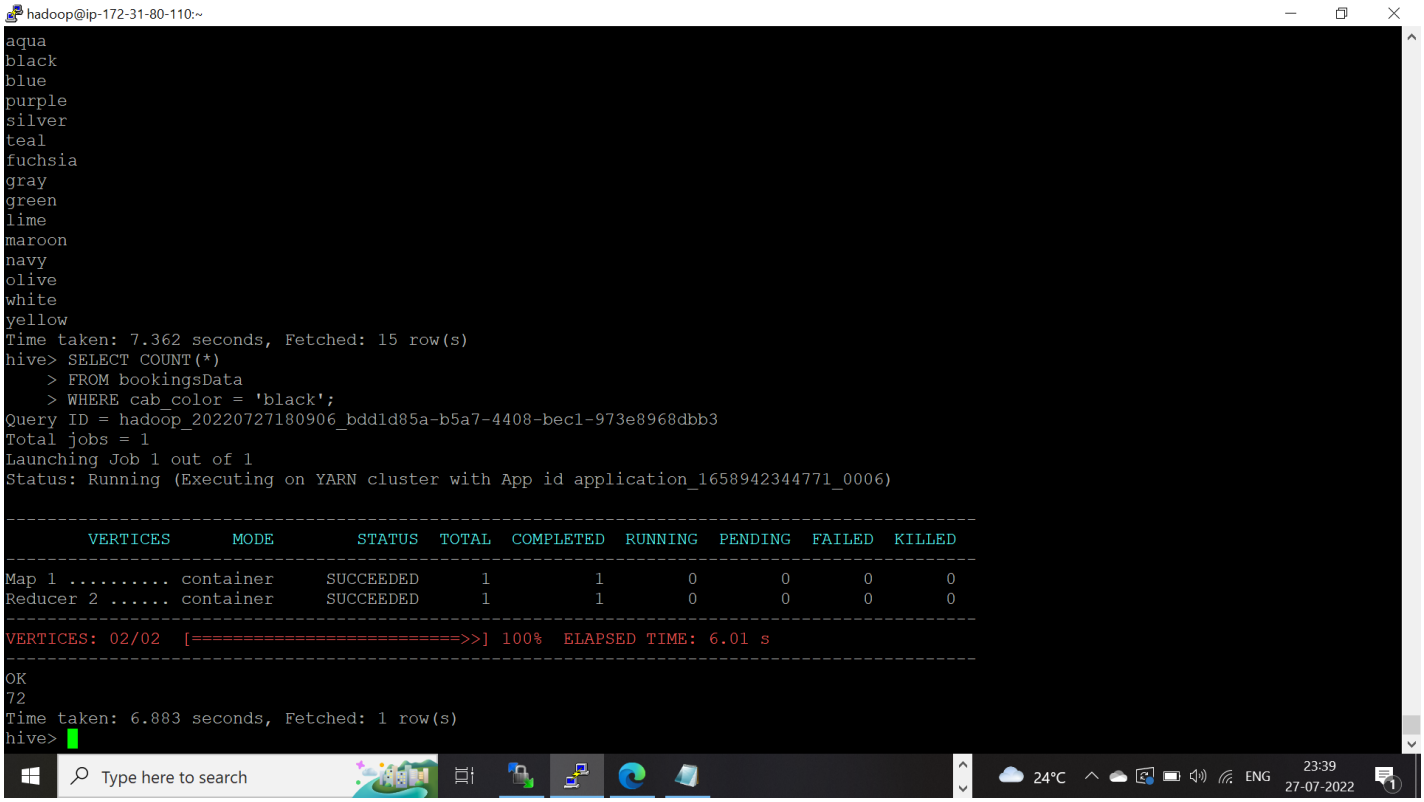


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

4. **Hive Query for Task 8:**
   Calculate the count of all trips done on black cabs.

```
SELECT COUNT(*)
FROM bookingsData
WHERE cab_color = 'black';
```

Screenshot after executing Query:

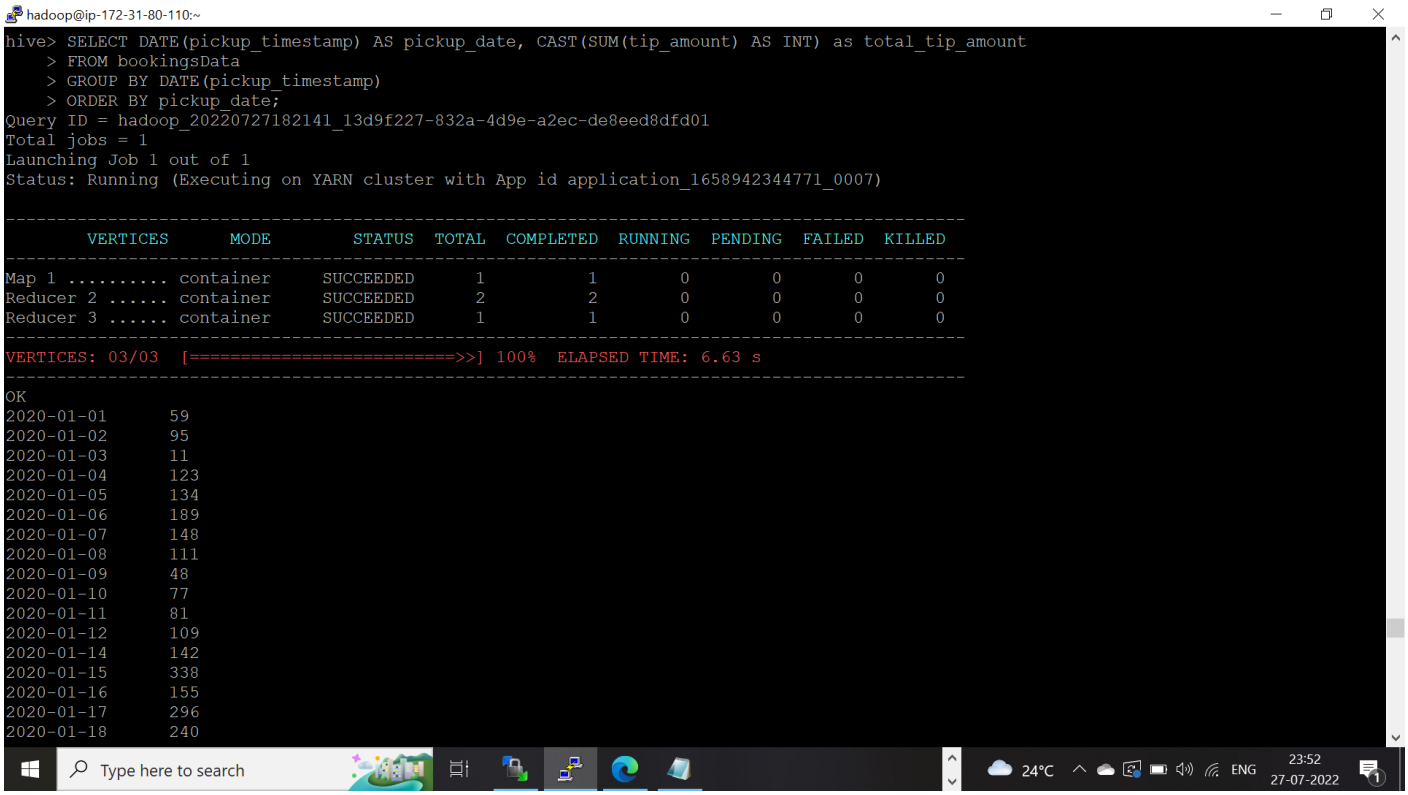


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

5. **Hive Query for Task 9:**
   Calculate the total amount of tips given date wise to all drivers by customers

```
SELECT DATE(pickup_timestamp) AS pickup_date,
CAST(SUM(tip_amount) AS INT) as
total_tip_amount
FROM bookingsData
GROUP BY DATE(pickup_timestamp)
ORDER BY pickup_date;
```

Screenshot after executing Query:

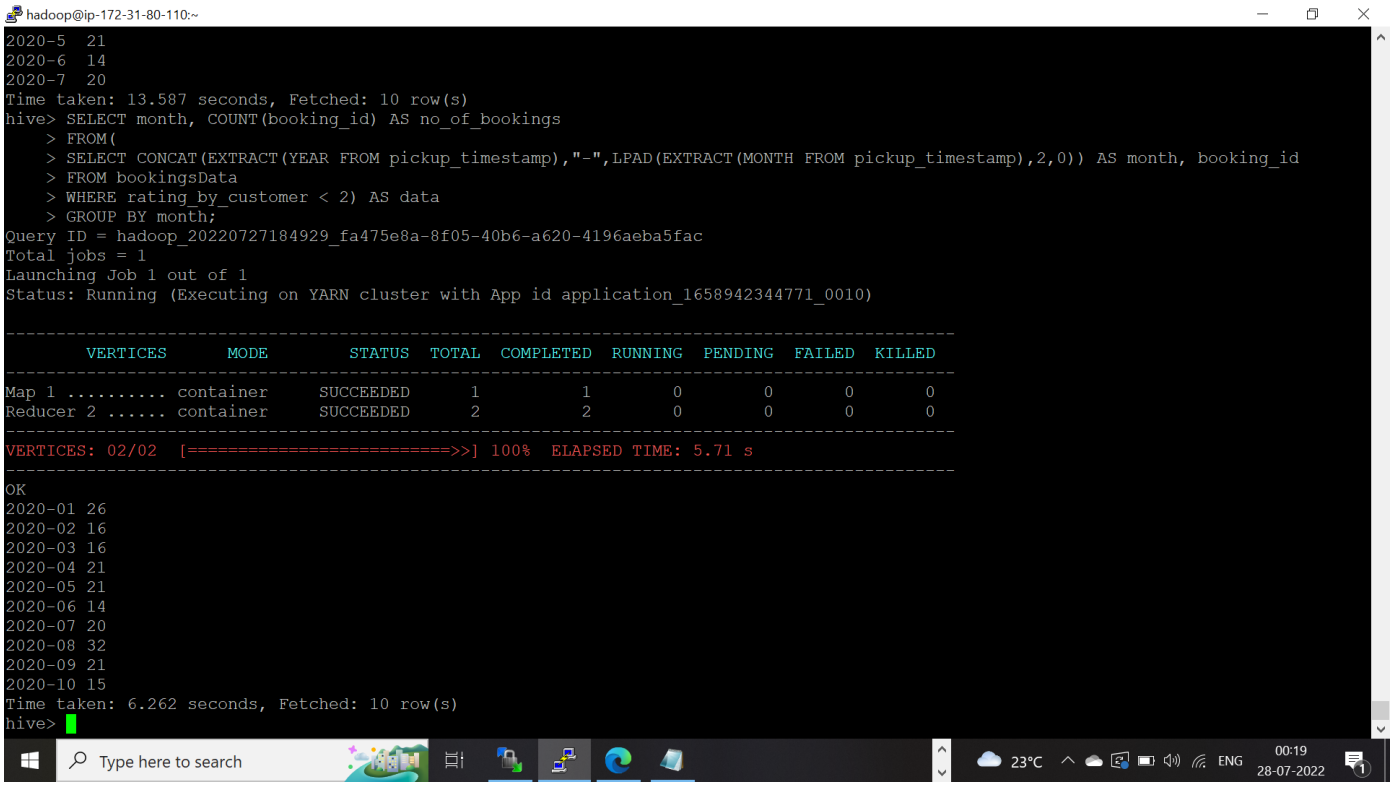


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

6. **Hive Query for Task 10:**
   Calculate the total count of all the bookings with ratings lower than 2 as given by customers in a particular month.

```
SELECT month, COUNT(booking_id) AS
no_of_bookings
FROM(
SELECT CONCAT(EXTRACT(YEAR FROM
pickup_timestamp),"-",LPAD(EXTRACT(MONTH
FROM pickup_timestamp),2,0)) AS month,
booking_id
FROM bookingsData
WHERE rating_by_customer < 2) AS data
GROUP BY month;
```

Screenshot after executing Query:

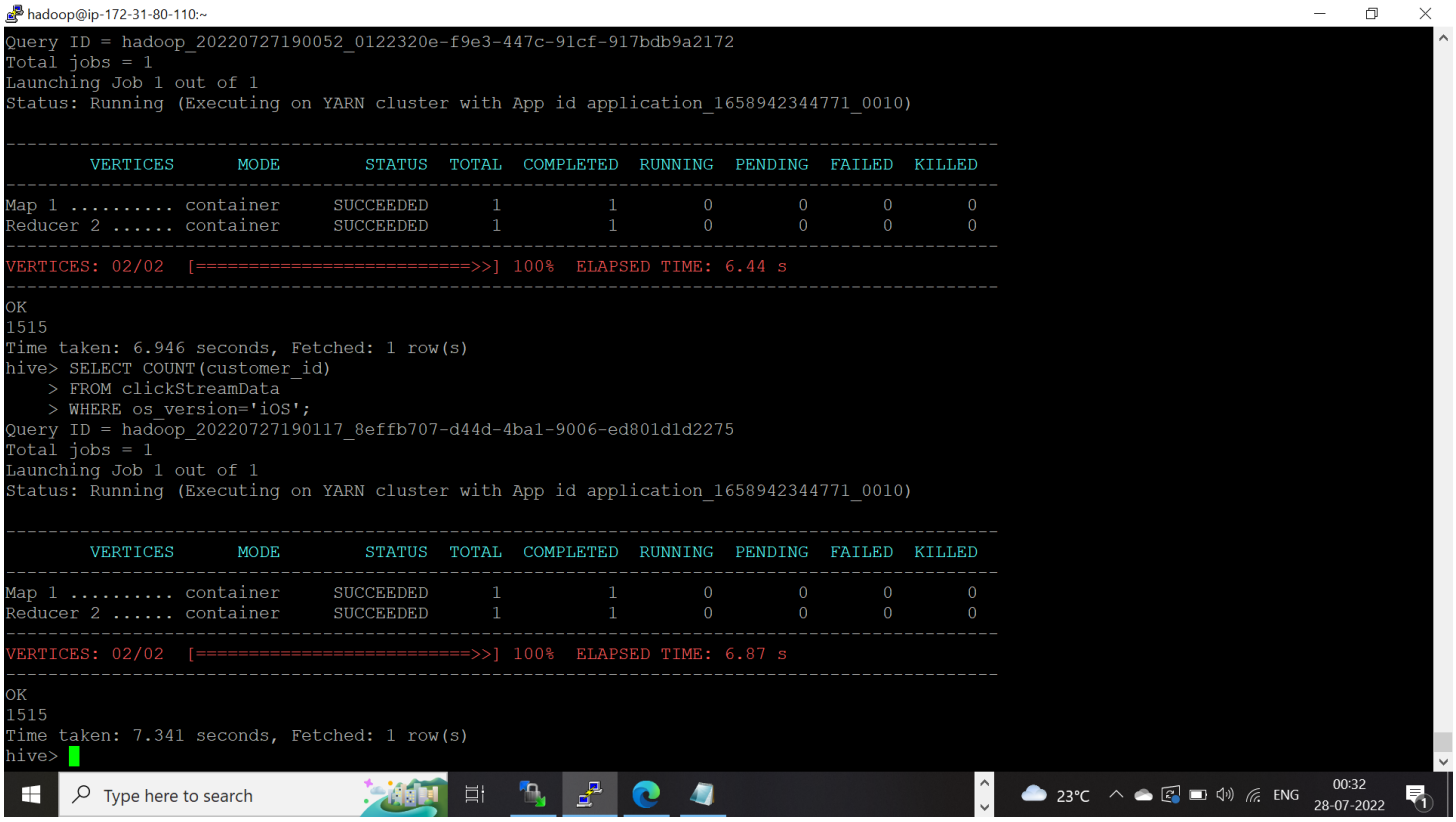


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

7. **Hive Query for Task 11:**
   Calculate the count of total iOS users.

```
SELECT COUNT(customer_id)
FROM clickStreamData
WHERE os_version='iOS';
```

Screenshot after executing Query:



© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved