

Data Analysis

Tuesday, 30 May 2023

3:35 PM

```
-- find total number of members and casuals
SELECT member_casual, COUNT(member_casual) AS total FROM(
SELECT DISTINCT ride_id, member_casual
FROM `my-first-sandbox-project11112.project_cyclistic.cyclist_trip_data`
)
GROUP BY member_casual
-- find total trips by bike type and by member/casual
SELECT rideable_type, member_casual, COUNT(*) AS no_of_rides
FROM project_cyclistic.cyclist_trip_data
GROUP BY rideable_type, member_casual
ORDER BY rideable_type ASC, no_of_rides DESC
-- find total number of round trips
SELECT start_station_id AS station_id, start_station_name AS station_name,
COUNT(start_station_name) AS total_roundtrip
FROM `my-first-sandbox-project11112.project_cyclistic.cyclist_trip_data`
WHERE start_station_name = end_station_name
GROUP BY start_station_id, start_station_name
ORDER BY total_roundtrip DESC LIMIT 10
--find total number of trips from starting point
SELECT start_station_id, start_station_name, COUNT(start_station_name) AS total_start
FROM `my-first-sandbox-project11112.project_cyclistic.cyclist_trip_data`
WHERE start_station_id <> end_station_id
GROUP BY start_station_id, start_station_name
ORDER BY total_start DESC LIMIT 10
--find total number of trips from ending point
SELECT end_station_id, end_station_name, COUNT(end_station_name) AS total_end
FROM `my-first-sandbox-project11112.project_cyclistic.cyclist_trip_data`
WHERE start_station_id <> end_station_id
GROUP BY end_station_id, end_station_name
ORDER BY total_end DESC LIMIT 10
-- find which station has more casuals than members for starting point
WITH temp_table_member AS (
SELECT member_casual, start_station_id, start_station_name, COUNT(*) AS total
```

point

```
WITH temp_table_member AS (  
  SELECT member_casual, start_station_id, start_station_name,  
  COUNT(*) AS total  
  FROM `my-first-sandbox-  
project11112.project_cyclistic.cyclist_trip_data`  
  WHERE member_casual = 'member'  
  GROUP BY member_casual, start_station_id, start_station_name  
  ORDER BY start_station_id  
) ,  
temp_table_casual AS (  
  SELECT member_casual, start_station_id, start_station_name,  
  COUNT(*) AS total  
  FROM `my-first-sandbox-  
project11112.project_cyclistic.cyclist_trip_data`  
  WHERE member_casual = 'casual'  
  GROUP BY member_casual, start_station_id, start_station_name  
  ORDER BY start_station_id  
)  
SELECT cyclist_trip_data.start_station_id,  
cyclist_trip_data.start_station_name, temp_table_member.total AS  
total_member, temp_table_casual.total AS total_casual,  
temp_table_member.total + temp_table_casual.total AS total  
FROM `my-first-sandbox-  
project11112.project_cyclistic.cyclist_trip_data` AS  
cyclist_trip_data  
INNER JOIN temp_table_member ON  
cyclist_trip_data.start_station_id =  
temp_table_member.start_station_id  
INNER JOIN temp_table_casual ON  
cyclist_trip_data.start_station_id =  
temp_table_casual.start_station_id  
WHERE temp_table_casual.total > temp_table_member.total  
GROUP BY start_station_id, start_station_name, total_member,  
total_casual  
ORDER BY total DESC LIMIT 10  
  
--find number of rides by month  
SELECT FORMAT_DATE('%b-%Y', started_at) AS month_year, COUNT(*)  
AS total  
FROM `my-first-sandbox-  
project11112.project_cyclistic.cyclist_trip_data`  
GROUP BY month_year  
ORDER BY PARSE_DATE('%b-%Y', month_year)  
--find number of rides per member by month  
SELECT member_casual, FORMAT_DATE('%b-%Y', started_at) AS
```

```

SELECT member_casual, FORMAT_DATE('%D-%Y', started_at) AS
month_year, COUNT(*) AS total
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
WHERE member_casual = 'member'
GROUP BY member_casual, month_year
ORDER BY PARSE_DATE('%b-%Y', month_year)

```

--find number of rides per casual by month

```

SELECT member_casual, FORMAT_DATE('%b-%Y', started_at) AS
month_year, COUNT(*) AS total
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
WHERE member_casual = 'casual'
GROUP BY member_casual, month_year
ORDER BY PARSE_DATE('%b-%Y', month_year)

```

--Find which day do most trips start on

```

SELECT FORMAT_DATE('%A', started_at) AS day_of_week, COUNT(*) AS
total
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
GROUP BY day_of_week
ORDER BY total DESC

```

-- find out the day of the week where most members ride

```

SELECT member_casual, FORMAT_DATE('%A', started_at) AS
day_of_week, COUNT(*) AS total
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
WHERE member_casual = 'member'
GROUP BY member_casual, day_of_week
ORDER BY total DESC

```

-- find out the day of the week where most casuals ride

```

SELECT member_casual, FORMAT_DATE('%A', started_at) AS
day_of_week, COUNT(*) AS total
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
WHERE member_casual = 'casual'
GROUP BY member_casual, day_of_week
ORDER BY total DESC

```

-- find the longest trip and shortest trip by a member

```

SELECT member_casual, MAX(ride_length) AS longest_trip,
MIN(ride_length) AS shortest_trip,
ROUND(AVG(ride_length_minutes)) AS average_duration_minutes FROM
/

```

```

(
SELECT member_casual, ride_length, TIMESTAMP_DIFF(ended_at,
started_at, MINUTE) AS ride_length_minutes
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
)
WHERE member_casual = 'member'
GROUP BY member_casual

-- find the longest trip and shortest trip by a casual
SELECT member_casual, MAX(ride_length) AS longest_trip,
MIN(ride_length) AS shortest_trip,
ROUND(AVG(ride_length_minutes)) AS average_duration_minutes FROM
(
SELECT member_casual, ride_length, TIMESTAMP_DIFF(ended_at,
started_at, MINUTE) AS ride_length_minutes
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
)
WHERE member_casual = 'casual'
GROUP BY member_casual

-- find the longest trip and shortest trip by member/casual
SELECT member_casual, MAX(ride_length) AS longest_trip,
MIN(ride_length) AS shortest_trip,
ROUND(AVG(ride_length_minutes)) AS average_duration_minutes FROM
(
SELECT member_casual, ride_length, TIMESTAMP_DIFF(ended_at,
started_at, MINUTE) AS ride_length_minutes
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
)
GROUP BY member_casual

--average ride_length by member_type
SELECT member_casual, ROUND(AVG(ride_length_minutes)) AS
average_duration_minutes FROM (
SELECT member_casual, TIMESTAMP_DIFF(ended_at, started_at,
MINUTE) AS ride_length_minutes
FROM `my-first-sandbox-
project11112.project_cyclistic.cyclist_trip_data`
)
GROUP BY member_casual

-- find the longest,shortest,average duration by member type and
bike type
SELECT member_casual, rideable_type, MAX(ride_length) AS
longest_trip, MIN(ride_length) AS shortest_trip,
ROUND(AVG(ride length minutes)) AS average duration minutes FROM

```