```
---
title: "Cyclistic SQL Analysis"
author: "Suzana Alípio"
date: "2024-12-20"
output: pdf
---
```

```
--Average Duration by User Type and Month
SELECT
    CASE
        WHEN EXTRACT(MONTH FROM started_at) = 1 THEN 'January'
        WHEN EXTRACT(MONTH FROM started_at) = 2 THEN 'February'
        WHEN EXTRACT(MONTH FROM started_at) = 3 THEN 'March'
        WHEN EXTRACT(MONTH FROM started_at) = 4 THEN 'April'
        WHEN EXTRACT(MONTH FROM started_at) = 5 THEN 'May'
        WHEN EXTRACT(MONTH FROM started_at) = 6 THEN 'June'
        WHEN EXTRACT(MONTH FROM started_at) = 7 THEN 'July'
        WHEN EXTRACT(MONTH FROM started_at) = 8 THEN 'August'
        WHEN EXTRACT(MONTH FROM started_at) = 9 THEN 'September'
        WHEN EXTRACT(MONTH FROM started_at) = 10 THEN 'October'
        WHEN EXTRACT(MONTH FROM started_at) = 11 THEN 'November'
        WHEN EXTRACT(MONTH FROM started_at) = 12 THEN 'December'
    END AS month_name,
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length_minutes), 2) AS avg_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    month_name, member_casual
ORDER BY
    CASE
        WHEN month_name = 'January' THEN 1
        WHEN month_name = 'February' THEN 2
        WHEN month_name = 'March' THEN 3
        WHEN month_name = 'April' THEN 4
        WHEN month_name = 'May' THEN 5
        WHEN month_name = 'June' THEN 6
        WHEN month_name = 'July' THEN 7
        WHEN month_name = 'August' THEN 8
        WHEN month_name = 'September' THEN 9
        WHEN month_name = 'October' THEN 10
        WHEN month_name = 'November' THEN 11
        WHEN month_name = 'December' THEN 12
    END,
    member_casual;
```

```sql
--Average Duration by User Type and Day of the Week
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 1 THEN 'Sunday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 2 THEN 'Monday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 3 THEN 'Tuesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 4 THEN 'Wednesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 5 THEN 'Thursday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 6 THEN 'Friday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 7 THEN 'Saturday'
    END AS day_of_week_name,
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length_minutes), 2) AS avg_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    day_of_week_name, member_casual
ORDER BY
    CASE
        WHEN day_of_week_name = 'Sunday' THEN 1
        WHEN day_of_week_name = 'Monday' THEN 2
        WHEN day_of_week_name = 'Tuesday' THEN 3
        WHEN day_of_week_name = 'Wednesday' THEN 4
        WHEN day_of_week_name = 'Thursday' THEN 5
        WHEN day_of_week_name = 'Friday' THEN 6
        WHEN day_of_week_name = 'Saturday' THEN 7
    END,
    member_casual;


--Total Average Duration by User Type
SELECT
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length_minutes), 2) AS avg_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    member_casual;

--Calculate the General Average
SELECT
    ROUND(AVG(ride_length_minutes), 2) AS avg_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`;
```

```sql
--Average by User Type
SELECT
    member_casual,
    ROUND(AVG(ride_length_minutes), 2) AS avg_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    member_casual;



--MAX value of ride_length_minutes
SELECT
    ROUND(MAX(ride_length_minutes), 2) AS max_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`;



--MAX value of ride_length_minutes per user type
SELECT
    member_casual,
    ROUND(MAX(ride_length_minutes), 2) AS max_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    member_casual;
```

```sql
--Maximum Amount per Month and User Type
SELECT
    CASE
        WHEN EXTRACT(MONTH FROM started_at) = 1 THEN 'January'
        WHEN EXTRACT(MONTH FROM started_at) = 2 THEN 'February'
        WHEN EXTRACT(MONTH FROM started_at) = 3 THEN 'March'
        WHEN EXTRACT(MONTH FROM started_at) = 4 THEN 'April'
        WHEN EXTRACT(MONTH FROM started_at) = 5 THEN 'May'
        WHEN EXTRACT(MONTH FROM started_at) = 6 THEN 'June'
        WHEN EXTRACT(MONTH FROM started_at) = 7 THEN 'July'
        WHEN EXTRACT(MONTH FROM started_at) = 8 THEN 'August'
        WHEN EXTRACT(MONTH FROM started_at) = 9 THEN 'September'
        WHEN EXTRACT(MONTH FROM started_at) = 10 THEN 'October'
        WHEN EXTRACT(MONTH FROM started_at) = 11 THEN 'November'
        WHEN EXTRACT(MONTH FROM started_at) = 12 THEN 'December'
    END AS month_name,
    member_casual,
    ROUND(MAX(ride_length_minutes), 2) AS max_ride_length_minutes
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    month_name, member_casual
ORDER BY
    CASE
        WHEN month_name = 'January' THEN 1
        WHEN month_name = 'February' THEN 2
        WHEN month_name = 'March' THEN 3
        WHEN month_name = 'April' THEN 4
        WHEN month_name = 'May' THEN 5
        WHEN month_name = 'June' THEN 6
        WHEN month_name = 'July' THEN 7
        WHEN month_name = 'August' THEN 8
        WHEN month_name = 'September' THEN 9
        WHEN month_name = 'October' THEN 10
        WHEN month_name = 'November' THEN 11
        WHEN month_name = 'December' THEN 12
    END,
    member_casual;
```

```sql
--Most Popular Stations to Start by user type
SELECT
    start_station_name,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
GROUP BY
    start_station_name, member_casual
ORDER BY
    total_rides DESC
LIMIT (10);
```

```sql
--Most Used Stations for End
SELECT
    end_station_name,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    end_station_name IS NOT NULL
GROUP BY
    end_station_name, member_casual
ORDER BY
    total_rides DESC
LIMIT (10);
```

```sql
--Average Distance between Start and End by User Type
SELECT
    member_casual,
    ROUND(AVG(ST_DISTANCE(ST_GEOGPOINT(start_lng, start_lat),
ST_GEOGPOINT(end_lng, end_lat)) / 1000), 2) AS avg_distance_km
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_lat IS NOT NULL AND start_lng IS NOT NULL AND end_lat IS NOT NULL
AND end_lng IS NOT NULL
GROUP BY
    member_casual;
```

```sql
--Most Popular Times:total rides taken based on the hour of the day based on
the type of rider

SELECT
    EXTRACT(HOUR FROM started_at) AS start_hour,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    start_hour, member_casual
ORDER BY
    start_hour, member_casual;
```

```sql
--Most Popular Days of the Week by user
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 1 THEN 'Sunday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 2 THEN 'Monday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 3 THEN 'Tuesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 4 THEN 'Wednesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 5 THEN 'Thursday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 6 THEN 'Friday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 7 THEN 'Saturday'
    END AS day_of_week_name,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    day_of_week_name, member_casual
ORDER BY
    CASE
        WHEN day_of_week_name = 'Sunday' THEN 1
        WHEN day_of_week_name = 'Monday' THEN 2
        WHEN day_of_week_name = 'Tuesday' THEN 3
        WHEN day_of_week_name = 'Wednesday' THEN 4
        WHEN day_of_week_name = 'Thursday' THEN 5
        WHEN day_of_week_name = 'Friday' THEN 6
        WHEN day_of_week_name = 'Saturday' THEN 7
    END,
    member_casual;
```

```sql
--Calculates the numerical value and percentage of total rides taken by type
of bike user
SELECT
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (), 2) AS percentage_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    member_casual
ORDER BY
    total_rides DESC;




--Total Trips by Station and User Type
SELECT
    CASE
        WHEN EXTRACT(MONTH FROM started_at) IN (12, 1, 2) THEN 'Winter'
        WHEN EXTRACT(MONTH FROM started_at) IN (3, 4, 5) THEN 'Spring'
        WHEN EXTRACT(MONTH FROM started_at) IN (6, 7, 8) THEN 'Summer'
        WHEN EXTRACT(MONTH FROM started_at) IN (9, 10, 11) THEN 'Autumn'
    END AS season,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    season, member_casual
ORDER BY
    season, total_rides DESC;



--Total rides taken and percentage based on weekday vs weekend based on the
type of rider
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 'Weekend'  --
Sunday (1) and Saturday (7)
        ELSE 'Weekday'                                                    --
Monday (2) to Friday (6)
    END AS day_type,
    COUNT(*) AS total_rides,
    ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (), 2) AS percentage_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    day_type
ORDER BY
    total_rides DESC;
```

```sql
--Total and Percentage by Day of the Week and Weekend by User Type
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 'Weekend'
        ELSE 'Weekday'
    END AS day_type,
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY member_casual),
2) AS percentage_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    day_type, member_casual
ORDER BY
    member_casual, day_type;
```

```sql
--Most Popular Stations by Week and Weekend and by user type
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 'Weekend'
        ELSE 'Weekday'
    END AS day_type,
    start_station_name,
    member_casual,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
GROUP BY
    day_type, start_station_name, member_casual
ORDER BY
    day_type, member_casual, total_rides DESC
LIMIT 30;
```

```sql
--Most Popular months by user type
SELECT
  CASE
      WHEN EXTRACT(MONTH FROM started_at) = 1 THEN 'January'
      WHEN EXTRACT(MONTH FROM started_at) = 2 THEN 'February'
      WHEN EXTRACT(MONTH FROM started_at) = 3 THEN 'March'
      WHEN EXTRACT(MONTH FROM started_at) = 4 THEN 'April'
      WHEN EXTRACT(MONTH FROM started_at) = 5 THEN 'May'
      WHEN EXTRACT(MONTH FROM started_at) = 6 THEN 'June'
      WHEN EXTRACT(MONTH FROM started_at) = 7 THEN 'July'
      WHEN EXTRACT(MONTH FROM started_at) = 8 THEN 'August'
      WHEN EXTRACT(MONTH FROM started_at) = 9 THEN 'September'
      WHEN EXTRACT(MONTH FROM started_at) = 10 THEN 'October'
      WHEN EXTRACT(MONTH FROM started_at) = 11 THEN 'November'
      WHEN EXTRACT(MONTH FROM started_at) = 12 THEN 'December'
  END AS month_name,
  member_casual AS user_type,
  COUNT(*) AS total_rides
FROM
  `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
  month_name, user_type, EXTRACT(MONTH FROM started_at)
ORDER BY
  total_rides DESC;


--calculate the count and percentage of trips taken during weekdays and
weekend by user type
SELECT
  member_casual AS user_type,
  SUM(CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (2, 3, 4, 5, 6) THEN 1
        ELSE 0
      END) AS total_weekday_rides,
  ROUND(SUM(CASE
              WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (2, 3, 4, 5, 6) THEN
1
              ELSE 0
            END) * 100.0 / COUNT(*), 2) AS weekday_percentage,
  SUM(CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 1
        ELSE 0
      END) AS total_weekend_rides,
  ROUND(SUM(CASE
              WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 1
              ELSE 0
            END) * 100.0 / COUNT(*), 2) AS weekend_percentage,
  COUNT(*) AS total_rides
FROM
  `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
  user_type
ORDER BY
  user_type;
```

```sql
--total number of trips taken at each start station per day of the week,
focusing only on casual users
WITH day_data AS (
  SELECT
    start_station_name,
    EXTRACT(DAYOFWEEK FROM started_at) AS day_of_week
  FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
  WHERE
    member_casual = 'casual'  -- Focus on casual users only
    AND start_station_name IS NOT NULL
)
SELECT
  start_station_name,
  CASE
      WHEN day_of_week = 1 THEN 'Sunday'
      WHEN day_of_week = 2 THEN 'Monday'
      WHEN day_of_week = 3 THEN 'Tuesday'
      WHEN day_of_week = 4 THEN 'Wednesday'
      WHEN day_of_week = 5 THEN 'Thursday'
      WHEN day_of_week = 6 THEN 'Friday'
      WHEN day_of_week = 7 THEN 'Saturday'
  END AS day_of_week,
  COUNT(*) AS total_rides
FROM
  day_data
GROUP BY
  start_station_name,
  day_of_week
ORDER BY
  total_rides DESC;  -- Sort by total number of trips (descending)
```

```sql
--total number of trips taken at each starting station per month, focusing
only on casual users
WITH month_data AS (
  SELECT
    start_station_name,
    EXTRACT(MONTH FROM started_at) AS month_number
  FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
  WHERE
    member_casual = 'casual'  -- Focus on casual users only
    AND start_station_name IS NOT NULL
)
SELECT
  start_station_name,
  CASE
      WHEN month_number = 1 THEN 'January'
      WHEN month_number = 2 THEN 'February'
      WHEN month_number = 3 THEN 'March'
      WHEN month_number = 4 THEN 'April'
      WHEN month_number = 5 THEN 'May'
      WHEN month_number = 6 THEN 'June'
      WHEN month_number = 7 THEN 'July'
      WHEN month_number = 8 THEN 'August'
      WHEN month_number = 9 THEN 'September'
      WHEN month_number = 10 THEN 'October'
      WHEN month_number = 11 THEN 'November'
      WHEN month_number = 12 THEN 'December'
  END AS month_name,
  COUNT(*) AS total_rides
FROM
  month_data
GROUP BY
  start_station_name,
  month_number
ORDER BY
  total_rides DESC;  -- Sort by total number of trips (descending)



----Total number of rides taken by hour by user type
SELECT
  EXTRACT(HOUR FROM started_at) AS ride_hour,
  member_casual AS user_type,
  COUNT(*) AS total_rides
FROM
  `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
  ride_hour, user_type
ORDER BY
  total_rides DESC;
```

```sql
--Total rides taken and percentage based on weekday vs weekend based on the
type of rider
SELECT
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) IN (1, 7) THEN 'Weekend'  --
Sunday (1) and Saturday (7)
        ELSE 'Weekday'                                                    --
Monday (2) to Friday (6)
    END AS day_type,
    member_casual AS user_type,
    COUNT(*) AS total_rides,
    ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY member_casual),
2) AS percentage_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    day_type, user_type
ORDER BY
    total_rides DESC;


--total trips per hour, day of the week and type of user
SELECT
    EXTRACT(HOUR FROM started_at) AS ride_hour,
    CASE
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 1 THEN 'Sunday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 2 THEN 'Monday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 3 THEN 'Tuesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 4 THEN 'Wednesday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 5 THEN 'Thursday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 6 THEN 'Friday'
        WHEN EXTRACT(DAYOFWEEK FROM started_at) = 7 THEN 'Saturday'
    END AS day_of_week,
    member_casual AS user_type,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
    ride_hour, day_of_week, user_type, EXTRACT(DAYOFWEEK FROM started_at)
ORDER BY
    total_rides DESC;



--Calculates total rides taken by preferred bike
SELECT
  rideable_type AS bike_type,
  member_casual AS user_type,
  COUNT(*) AS total_rides
FROM
  `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
  bike_type, user_type
ORDER BY
  total_rides DESC;
```

```sql
--Compare the preferred type of bike between the two types of users (members
and casuals),
SELECT
  rideable_type AS bike_type,
  member_casual AS user_type,
  COUNT(*) AS total_rides,
  ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY member_casual), 2)
AS percentage_rides
FROM
  `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
GROUP BY
  bike_type, user_type
ORDER BY
  user_type, total_rides DESC;


--Most popular starting stations for annual members (member)
SELECT
    start_station_name,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
    AND member_casual = 'member'  -- Focus on Annual Members Only
GROUP BY
    start_station_name
ORDER BY
    total_rides DESC
LIMIT 10;



--Most popular home stations for casual users (casual):
SELECT
    start_station_name,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
    AND member_casual = 'casual'  -- Focus on casual users only
GROUP BY
    start_station_name
ORDER BY
    total_rides DESC
LIMIT 10;
```

```sql
--Most popular routes when combining (start_station_name) and
(end_station_name), by user type
SELECT
    CONCAT(start_station_name, ' -> ', end_station_name) AS route,
    member_casual AS user_type,
    COUNT(*) AS total_rides
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
    AND end_station_name IS NOT NULL
GROUP BY
    route, user_type
ORDER BY
    total_rides DESC
LIMIT 20;

--Most popular routes when combining (start_station_name) and
(end_station_name), by user type, including location data
SELECT
    CONCAT(start_station_name, ' -> ', end_station_name) AS route,
    member_casual AS user_type,
    COUNT(*) AS total_rides,
    start_station_name,
    end_station_name,
    start_lat AS start_latitude,
    start_lng AS start_longitude,
    end_lat AS end_latitude,
    end_lng AS end_longitude
FROM
    `oval-airship-437710-
u0.my_cyclistic_csv_data.my_cyclistic_csv_cleaned_with_length_minutes`
WHERE
    start_station_name IS NOT NULL
    AND end_station_name IS NOT NULL
    AND start_lat IS NOT NULL
    AND start_lng IS NOT NULL
    AND end_lat IS NOT NULL
    AND end_lng IS NOT NULL
GROUP BY
    route, user_type, start_station_name, end_station_name, start_lat,
start_lng, end_lat, end_lng
ORDER BY
    total_rides DESC
LIMIT 10;
```