CYCLISTIC BIKESHARE

Presented as a *Case Study* for *Google Data Analytics Capstone* in the *Google Data Analytics Professional Certificate Course* by **Afviya Nabila** *(2023, June)*

Visualization on Tableau
Presentation on Slides
Full Project on Github
Full Profile on Linkedin

CASE:

Cyclistic is a bike-share company in Chicago. Cyclistic has two types of customers, casual riders who purchase either single-ride passes or full-day passes and annual members who purchase annual membership. Cyclistic features more than 5,824 bicycles and 692 docking stations across Chicago. Customers more likely ride for leisure, but about 30% to commute to work each day. Finance analysts have concluded that annual members are much more profitable than casual riders. The director of marketing analyst team believe there is a better chance to convert casual riders into annual members instead of creating a marketing campaign that targets all-new customers, because casual riders are already aware of the Cyclistic program.

IDENTIFY THE PROBLEM:

BUSINESS PROBLEM:

Maximizing the number of annual memberships by designing marketing strategies aimed at converting casual riders into annual members

BUSINESS TASK:

Analyze Cyclistic historical trip data to identify trends of how casual riders and annual members use Cyclistic bikes differently

QUESTIONS:

1. How casual riders and annual members use Cyclistic bikes differently?
2. Why would casual riders purchase annual membership?

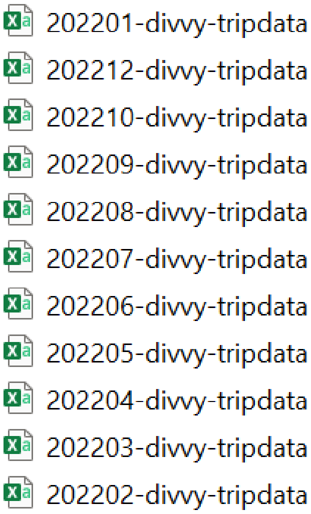
STAKEHOLDERS:

1. Cyclistic Executive Team
2. Director of Marketing Analyst Team
3. Marketing Analyst Team Members

DATA PREPARATION:

DATA SOURCES:

Dataset provided by Cyclistic (originally made available by Motivate International Inc. under license). The data used is Cyclistic historical trip data for 2022 saved per month as CSV files which are publicly available here.



202201-divvy-tripdata
202212-divvy-tripdata
202210-divvy-tripdata
202209-divvy-tripdata
202208-divvy-tripdata
202207-divvy-tripdata
202206-divvy-tripdata
202205-divvy-tripdata
202204-divvy-tripdata
202203-divvy-tripdata
202202-divvy-tripdata

The data is structured data where each CSV file consists of rows represent records per trip and columns represent fields of trip information. Each trip have unique id represent in the ride_id column.

Fields	Data Type	Description
ride_id	string	Unique ID
rideable_type	string	Bike types: classic, docked, electric
started_at	timestamp	Trip start day and time
ended_at	timestamp	Trip end day and time
start_station_name	string	Trip start station
start_station_id	string	Trip start station ID
end_station_name	string	Trip end station
end_station_id	string	Trip end station ID
start_lat	float	Trip start latitude
start_lng	float	Trip start longitude
end_lat	float	Trip end latitude
end_lng	float	Trip end longitude
member_casual	string	Customers type: casual or member

DATA CREDIBILITY:

Datasets come from reliable original sources cited by Cyclistic. Data consists of Cyclistic historical trip data per month in the past year which is current and comprehensive.

DATA PROCESSING:

DATA PRE-CLEANING:

The data prepared for analysis, to keep the original data, each CSV file is copied and saved as XLS for pre-cleaning. The following steps are performed and being saved as new CSV files:

- Removed duplicates
  - Data > Data Tools > Remove Duplicates

- Made sure no unneeded spaces and characters
  - Used TRIM and CLEAN function (example: =TRIM(CLEAN(A2)))
  - Used TRIM, CLEAN, LOWER function (example: =TRIM(CLEAN(LOWER(B2))))

- Made sure started_at and ended_at columns in the timestamp data type (DD/MM/YYYY hh:mm:ss)
  - Format > Cells > Custom > yyyy-mm-dd h:mm:ss

- Created a column called ride_date
  - Format > Cells > Date > yyyy-mm-dd

- Created a column called start_time
  - Format > Cells > Time > h:mm:ss

- Created a column called end_time
  - Format > Cells > Time > h:mm:ss

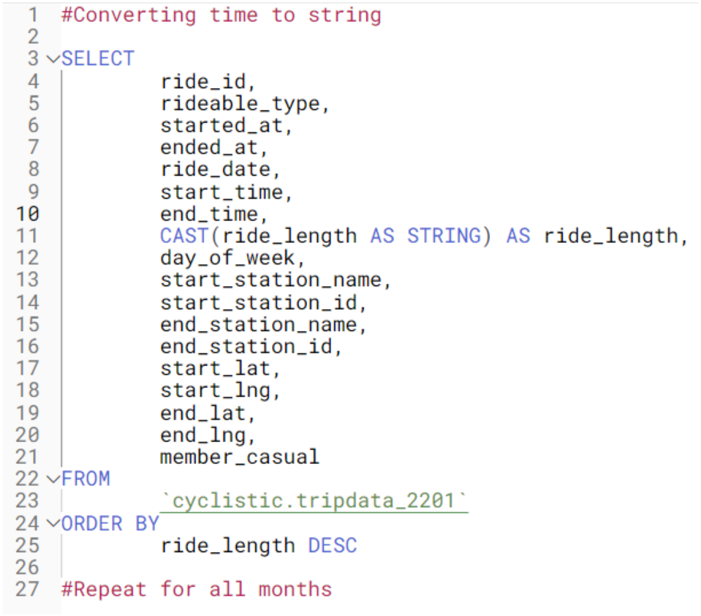
- Created a column called ride_length
  - Calculated each trip by subtracting end_time with start_time (example: =G2-F2)
  - Format > Cells > Time > h:mm:ss

- Created a column called day_of_week
  - Identified the day of the week of each trip using WEEKDAY function (example: =WEEKDAY(E2;2))
  - Return type used is 2 (where Number 1 = Monday through 7 = Sunday)
  - Format > Cells > Number

DATA CLEANING & MANIPULATION:

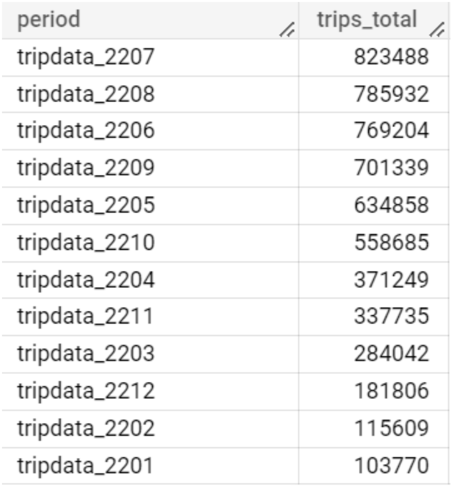
The process continued using SQL because each file has large size.

- Converted ride_length data type from date into string for easier further analysis



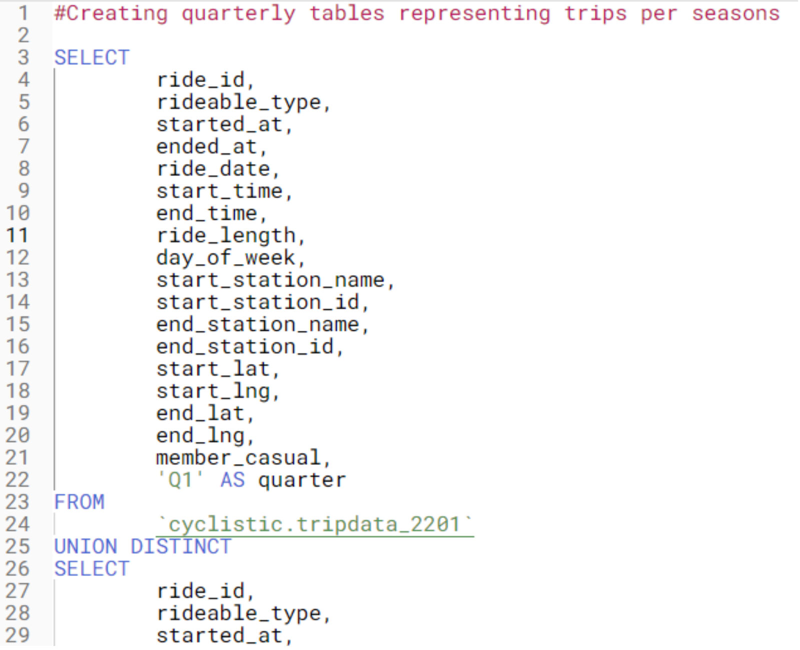
```
#Converting time to string

SELECT
        ride_id,
        rideable_type,
        started_at,
        ended_at,
        ride_date,
        start_time,
        end_time,
        CAST(ride_length AS STRING) AS ride_length,
        day_of_week,
        start_station_name,
        start_station_id,
        end_station_name,
        end_station_id,
        start_lat,
        start_lng,
        end_lat,
        end_lng,
        member_casual
FROM
        `cyclistic.tripdata_2201`
ORDER BY
        ride_length DESC

#Repeat for all months
```

- Counted how many trips each month has into one table using COUNT (DISTINCT). Generally we got to know that July, August, June, and September period (late Spring-Summer) have the most trips, followed by months in Spring and Fall season, then ended with months in Winter.

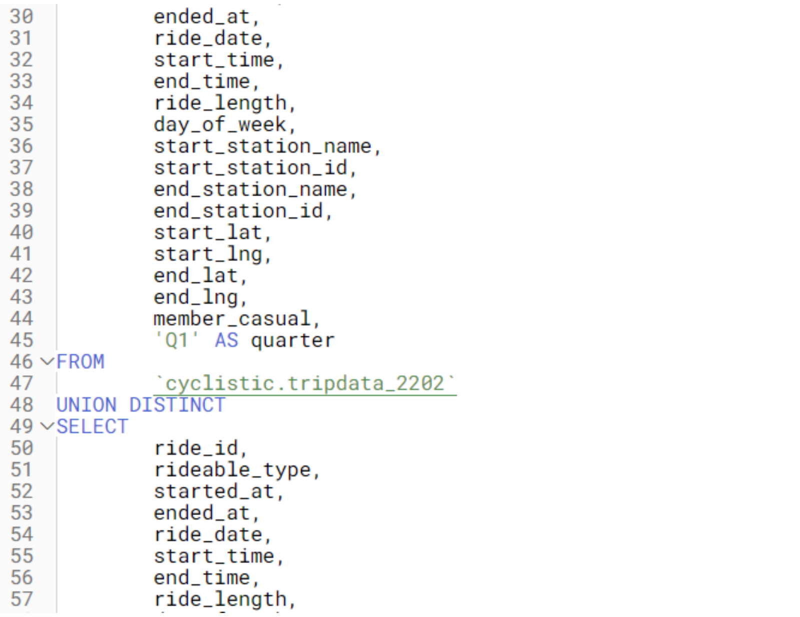


period	trips_total
tripdata_2207	823488
tripdata_2208	785932
tripdata_2206	769204
tripdata_2209	701339
tripdata_2205	634858
tripdata_2210	558685
tripdata_2204	371249
tripdata_2211	337735
tripdata_2203	284042
tripdata_2212	181806
tripdata_2202	115609
tripdata_2201	103770

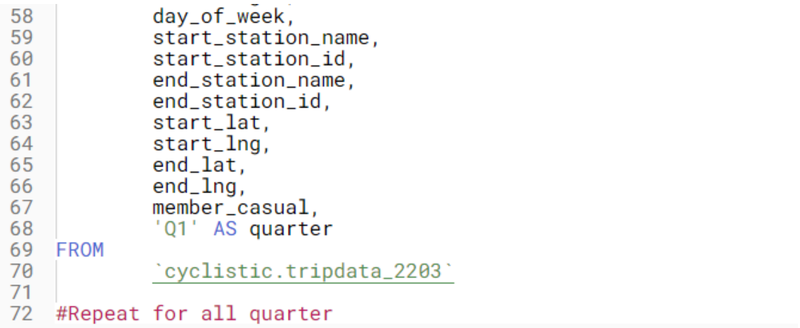
- Created quarterly tables representing trips per seasons (3-months period each) as Q1, Q2, Q3, and Q4
    - Q1 / Winter = January, February, March
    - Q2 / Spring = April, May, June
    - Q3 / Summer = July, August, September
    - Q4 / Fall = October, November, December



```
#Creating quarterly tables representing trips per seasons

SELECT
        ride_id,
        rideable_type,
        started_at,
        ended_at,
        ride_date,
        start_time,
        end_time,
        ride_length,
        day_of_week,
        start_station_name,
        start_station_id,
        end_station_name,
        end_station_id,
        start_lat,
        start_lng,
        end_lat,
        end_lng,
        member_casual,
        'Q1' AS quarter
FROM
        `cyclistic.tripdata_2201`
UNION DISTINCT
SELECT
        ride_id,
        rideable_type,
        started_at,
```

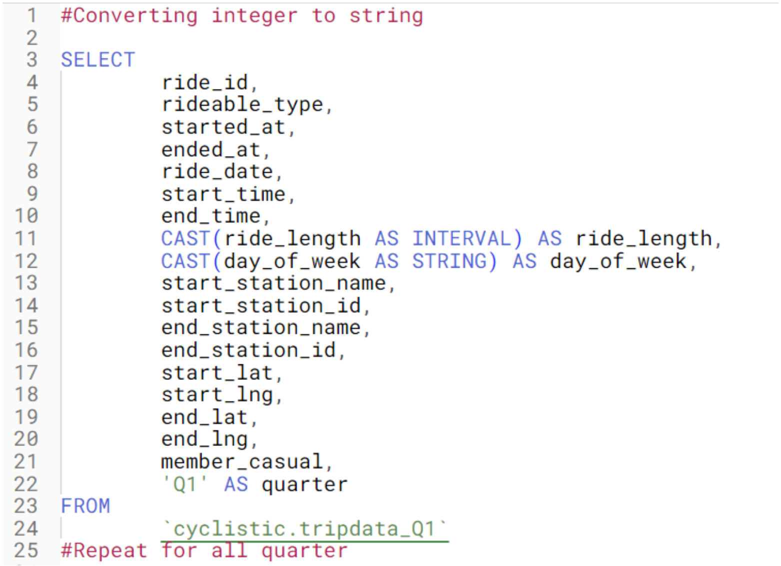


```
        ended_at,
        ride_date,
        start_time,
        end_time,
        ride_length,
        day_of_week,
        start_station_name,
        start_station_id,
        end_station_name,
        end_station_id,
        start_lat,
        start_lng,
        end_lat,
        end_lng,
        member_casual,
        'Q1' AS quarter
FROM
        `cyclistic.tripdata_2202`
UNION DISTINCT
SELECT
        ride_id,
        rideable_type,
        started_at,
        ended_at,
        ride_date,
        start_time,
        end_time,
        ride_length,
```



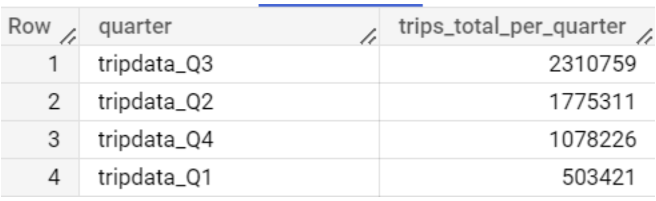
```
        day_of_week,
        start_station_name,
        start_station_id,
        end_station_name,
        end_station_id,
        start_lat,
        start_lng,
        end_lat,
        end_lng,
        member_casual,
        'Q1' AS quarter
FROM
        `cyclistic.tripdata_2203`

#Repeat for all quarter
```

- Converted ride_length data type to interval and day_of_week data type from integer into string for easier further analysis.



```
#Converting integer to string

SELECT
        ride_id,
        rideable_type,
        started_at,
        ended_at,
        ride_date,
        start_time,
        end_time,
        CAST(ride_length AS INTERVAL) AS ride_length,
        CAST(day_of_week AS STRING) AS day_of_week,
        start_station_name,
        start_station_id,
        end_station_name,
        end_station_id,
        start_lat,
        start_lng,
        end_lat,
        end_lng,
        member_casual,
        'Q1' AS quarter
FROM
        `cyclistic.tripdata_Q1`
#Repeat for all quarter
```

- Counted how many trips each quarter has using COUNT (DISTINCT). Generally we got to know Summer season (Q3) followed by Spring season (Q2) have the most trips.

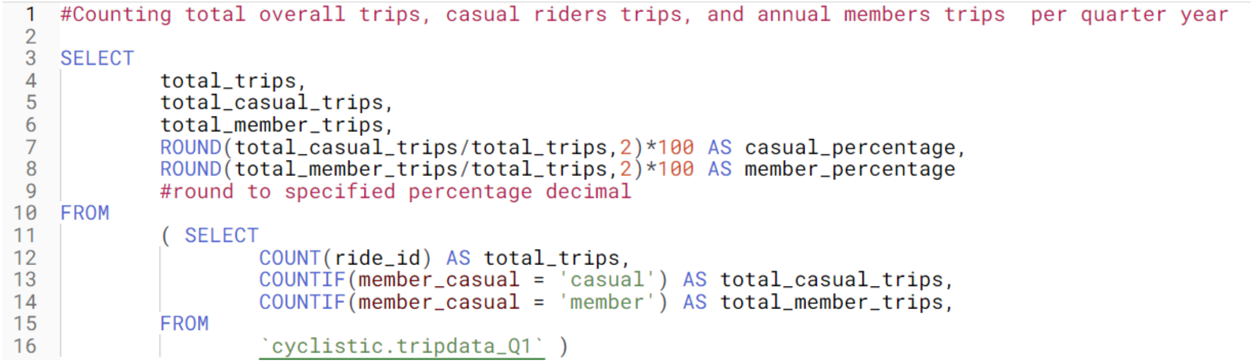


Row	quarter	trips_total_per_quarter
1	tripdata_Q3	2310759
2	tripdata_Q2	1775311
3	tripdata_Q4	1078226
4	tripdata_Q1	503421

EXPLORATORY DATA ANALYSIS:

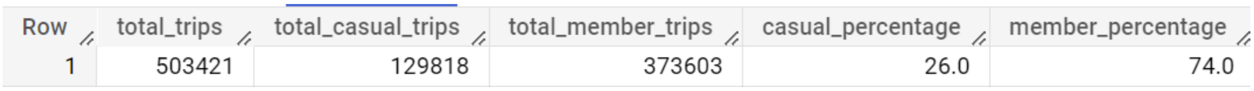
TOTAL TRIPS

We will count total overall trips, casual riders trips, and annual members trips per quarter year together with their percentages.



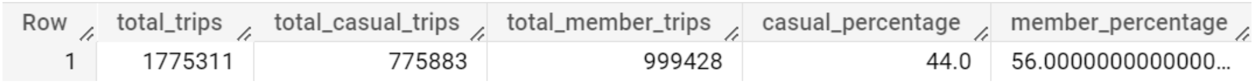
```
#Counting total overall trips, casual riders trips, and annual members trips  per quarter year

SELECT
        total_trips,
        total_casual_trips,
        total_member_trips,
        ROUND(total_casual_trips/total_trips,2)*100 AS casual_percentage,
        ROUND(total_member_trips/total_trips,2)*100 AS member_percentage
        #round to specified percentage decimal
FROM
        ( SELECT
                COUNT(ride_id) AS total_trips,
                COUNTIF(member_casual = 'casual') AS total_casual_trips,
                COUNTIF(member_casual = 'member') AS total_member_trips,
        FROM
                `cyclistic.tripdata_Q1` )
```

Q1



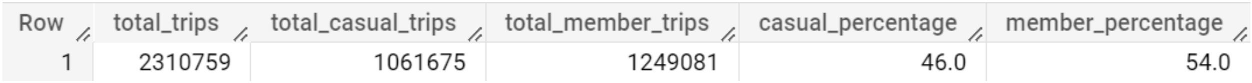
Row	total_trips	total_casual_trips	total_member_trips	casual_percentage	member_percentage
1	503421	129818	373603	26.0	74.0

Q2



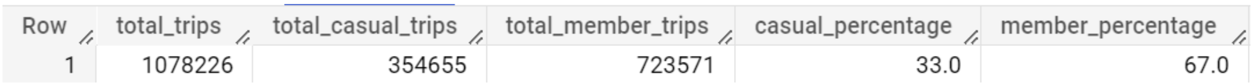
Row	total_trips	total_casual_trips	total_member_trips	casual_percentage	member_percentage
1	1775311	775883	999428	44.0	56.0000000000000...

Q3



Row	total_trips	total_casual_trips	total_member_trips	casual_percentage	member_percentage
1	2310759	1061675	1249081	46.0	54.0

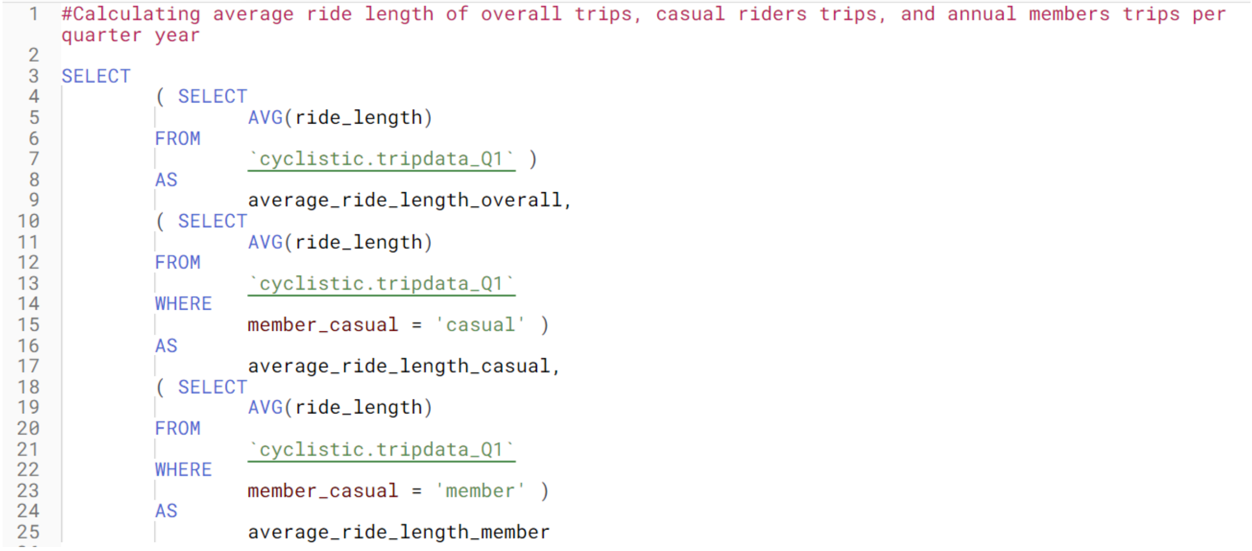
Q4



Row	total_trips	total_casual_trips	total_member_trips	casual_percentage	member_percentage
1	1078226	354655	723571	33.0	67.0

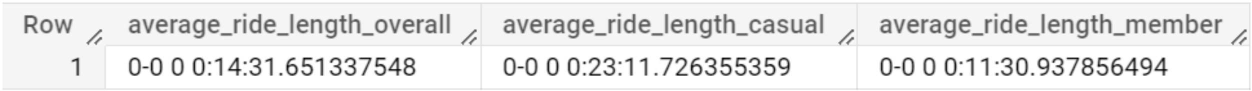
AVERAGE RIDE LENGTH

We calculate the average of ride_length for both casual riders and annual members trips per quarter year.



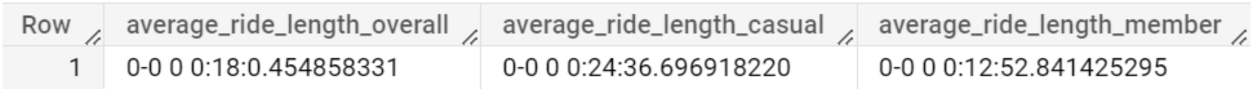
```
#Calculating average ride length of overall trips, casual riders trips, and annual members trips per
quarter year

SELECT
        ( SELECT
                AVG(ride_length)
        FROM
                `cyclistic.tripdata_Q1` )
        AS
                average_ride_length_overall,
        ( SELECT
                AVG(ride_length)
        FROM
                `cyclistic.tripdata_Q1`
        WHERE
                member_casual = 'casual' )
        AS
                average_ride_length_casual,
        ( SELECT
                AVG(ride_length)
        FROM
                `cyclistic.tripdata_Q1`
        WHERE
                member_casual = 'member' )
        AS
                average_ride_length_member
```

Q1



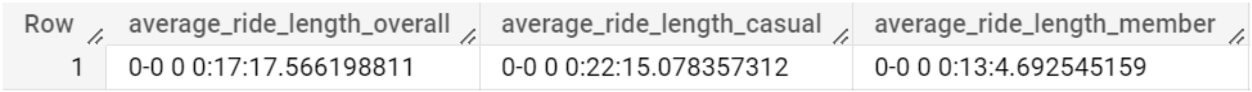
Row	average_ride_length_overall	average_ride_length_casual	average_ride_length_member
1	0-0 0 0:14:31.651337548	0-0 0 0:23:11.726355359	0-0 0 0:11:30.937856494

Q2



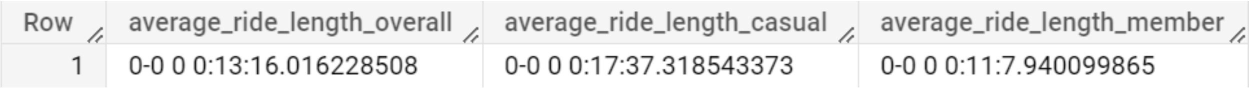
Row	average_ride_length_overall	average_ride_length_casual	average_ride_length_member
1	0-0 0 0:18:0.454858331	0-0 0 0:24:36.696918220	0-0 0 0:12:52.841425295

Q3



Row	average_ride_length_overall	average_ride_length_casual	average_ride_length_member
1	0-0 0 0:17:17.566198811	0-0 0 0:22:15.078357312	0-0 0 0:13:4.692545159

Q4

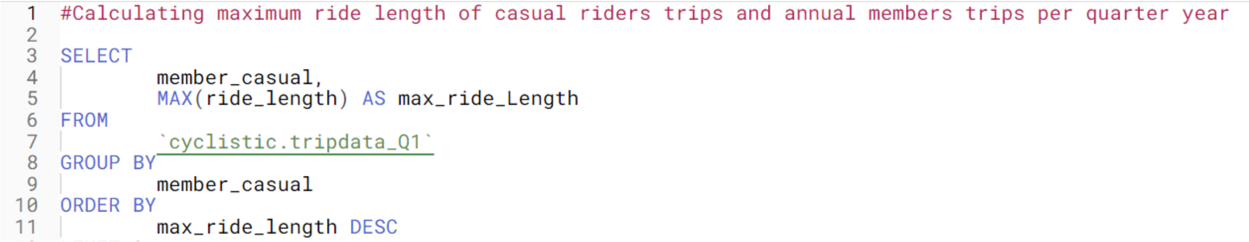


Row	average_ride_length_overall	average_ride_length_casual	average_ride_length_member
1	0-0 0 0:13:16.016228508	0-0 0 0:17:37.318543373	0-0 0 0:11:7.940099865

The table sowed that average ride_length of casual riders about 17 - 23 minutes while average ride_length of annual members about 11-14 minutes which show pretty much difference results.

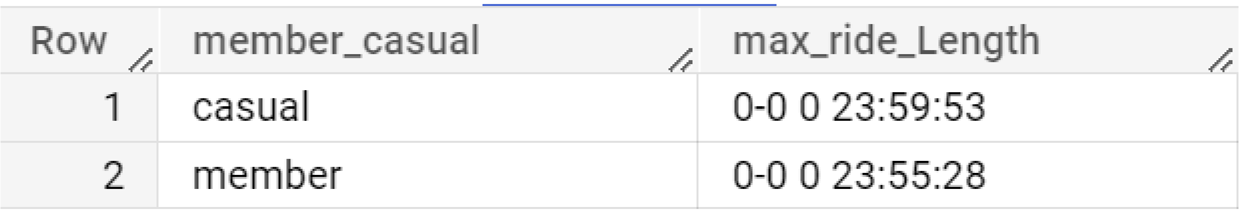
MAXIMUM RIDE LENGTH

We continue to calculate maximum ride_length of trips per quarter year.



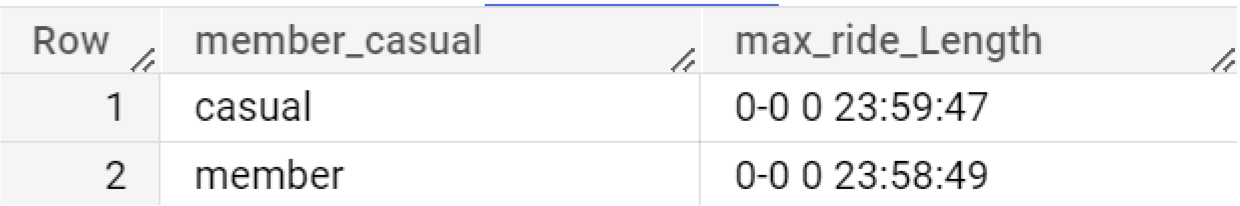
```
#Calculating maximum ride length of casual riders trips and annual members trips per quarter year

SELECT
        member_casual,
        MAX(ride_length) AS max_ride_Length
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        member_casual
ORDER BY
        max_ride_length DESC
```

Q1



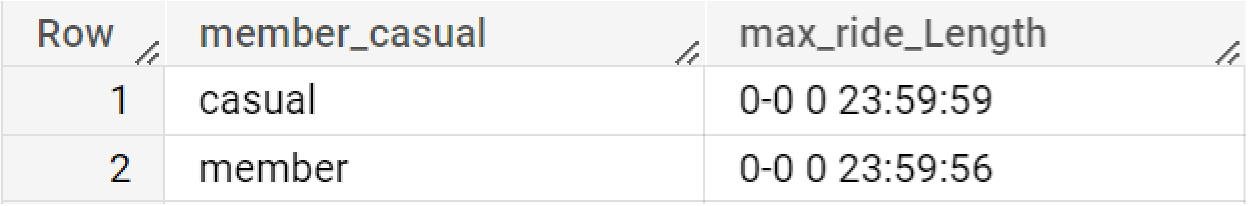
Row	member_casual	max_ride_Length
1	casual	0-0 0 23:59:53
2	member	0-0 0 23:55:28

Q2



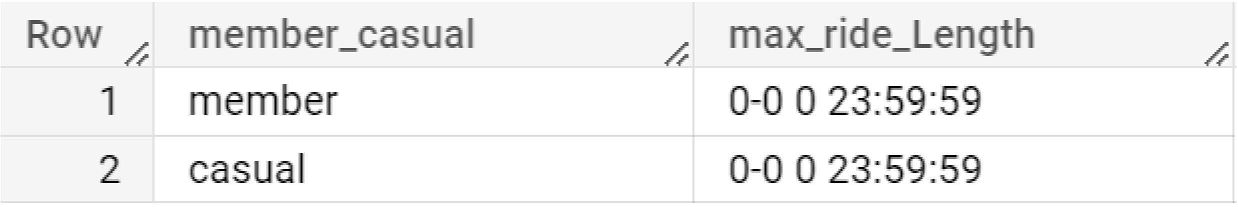
Row	member_casual	max_ride_Length
1	casual	0-0 0 23:59:47
2	member	0-0 0 23:58:49

Q3



Row	member_casual	max_ride_Length
1	casual	0-0 0 23:59:59
2	member	0-0 0 23:59:56

Q4

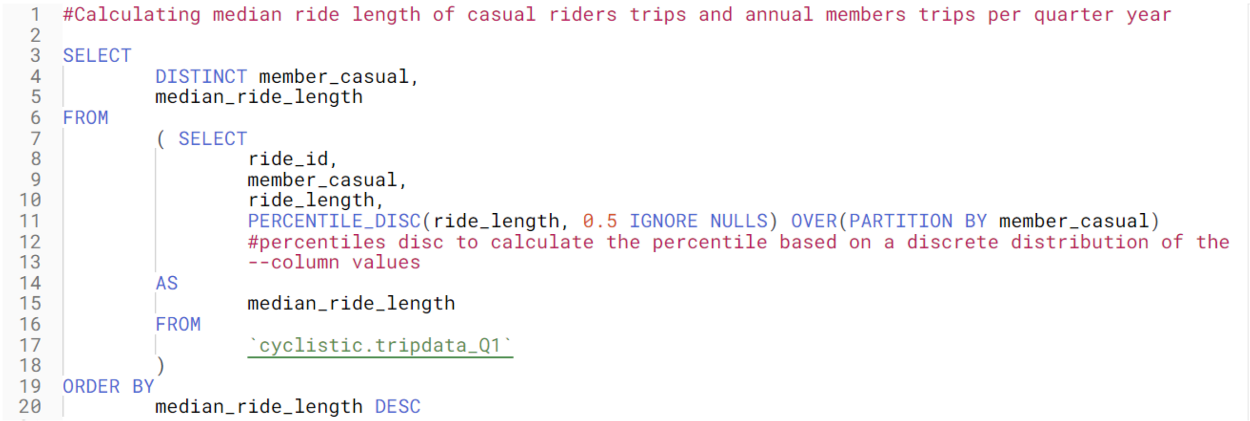


Row	member_casual	max_ride_Length
1	member	0-0 0 23:59:59
2	casual	0-0 0 23:59:59

The table showed that maximum ride_length for both type customers were around 1 day (24 hours).

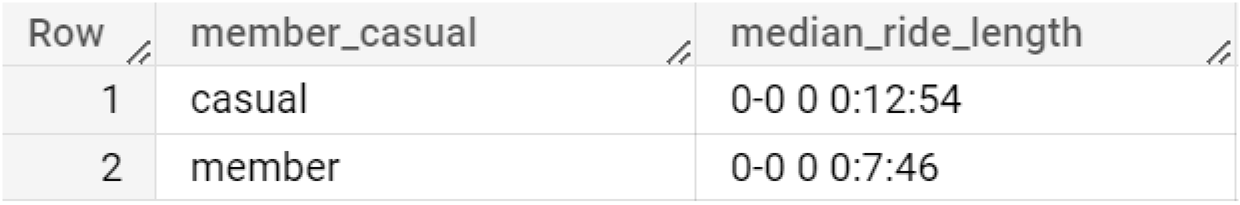
MEDIUM RIDE LENGTH

Maximum ride length might potentially influence the average of ride_length, instead we can use median ride_length to minimize the impact of that few outliers on further analysis.



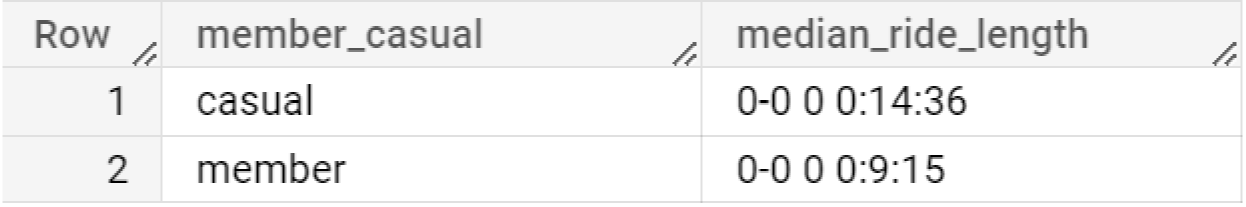
```sql
#Calculating median ride length of casual riders trips and annual members trips per quarter year

SELECT
        DISTINCT member_casual,
        median_ride_length
FROM
        ( SELECT
                ride_id,
                member_casual,
                ride_length,
                PERCENTILE_DISC(ride_length, 0.5 IGNORE NULLS) OVER(PARTITION BY member_casual)
                #percentiles disc to calculate the percentile based on a discrete distribution of the
                --column values
        AS
                median_ride_length
        FROM
                `cyclistic.tripdata_Q1`
        )
ORDER BY
        median_ride_length DESC
```

Q1



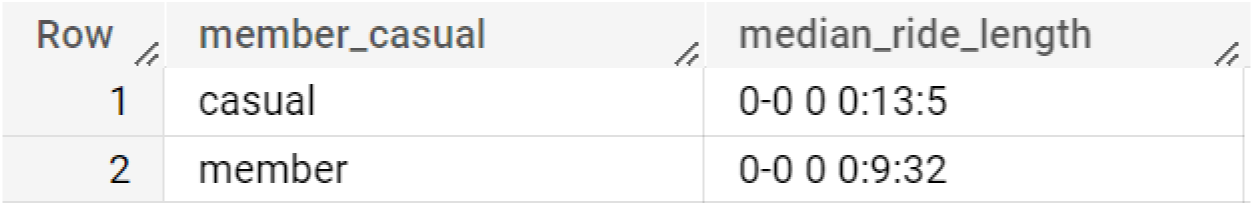
Row	member_casual	median_ride_length
1	casual	0-0 0 0:12:54
2	member	0-0 0 0:7:46

Q2



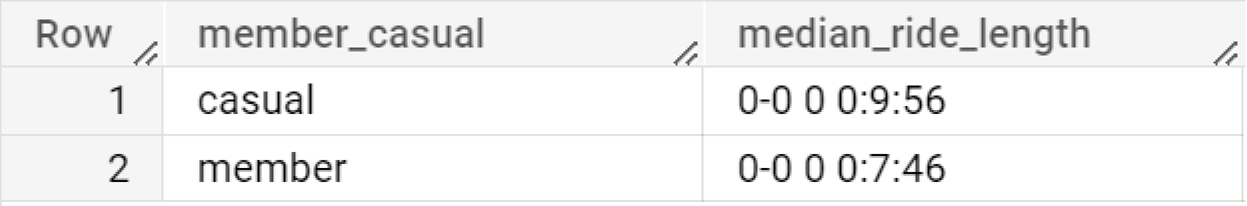
Row	member_casual	median_ride_length
1	casual	0-0 0 0:14:36
2	member	0-0 0 0:9:15

Q3



Row	member_casual	median_ride_length
1	casual	0-0 0 0:13:5
2	member	0-0 0 0:9:32

Q4

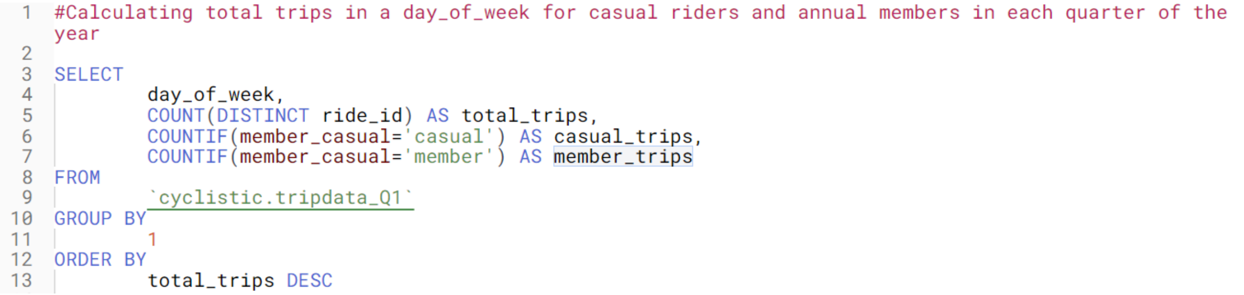


Row	member_casual	median_ride_length
1	casual	0-0 0 0:9:56
2	member	0-0 0 0:7:46

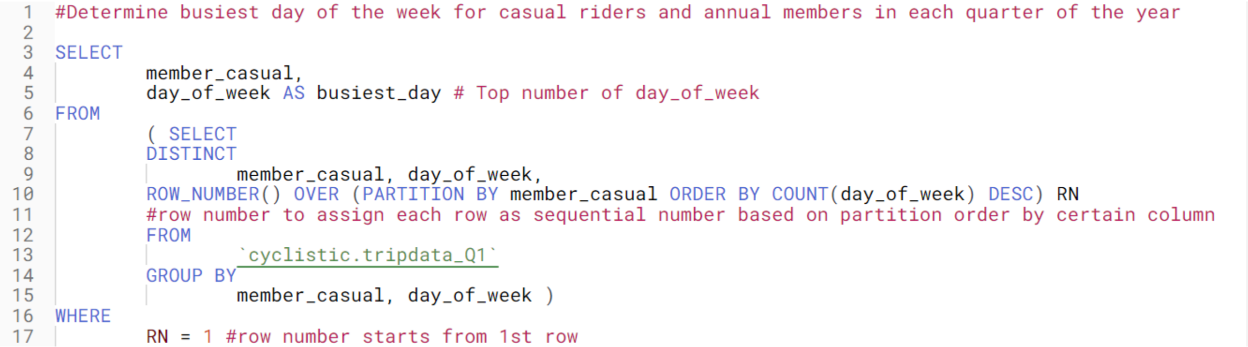
The results show that the median ride_length varies among quarters of the year, where casual riders between 9 minutes 56 seconds and 14 minutes 36 seconds, while annual members between 7 minutes 46 seconds and 9 minutes 32 seconds.

TOTAL TRIPS IN A DAY OF WEEK

We calculate total trips in a day_of_week to determine which day is the busiest_day of the week in each quarter of the year. In this case study, remember that day_of_week starts from Monday (1) to Sunday (7).

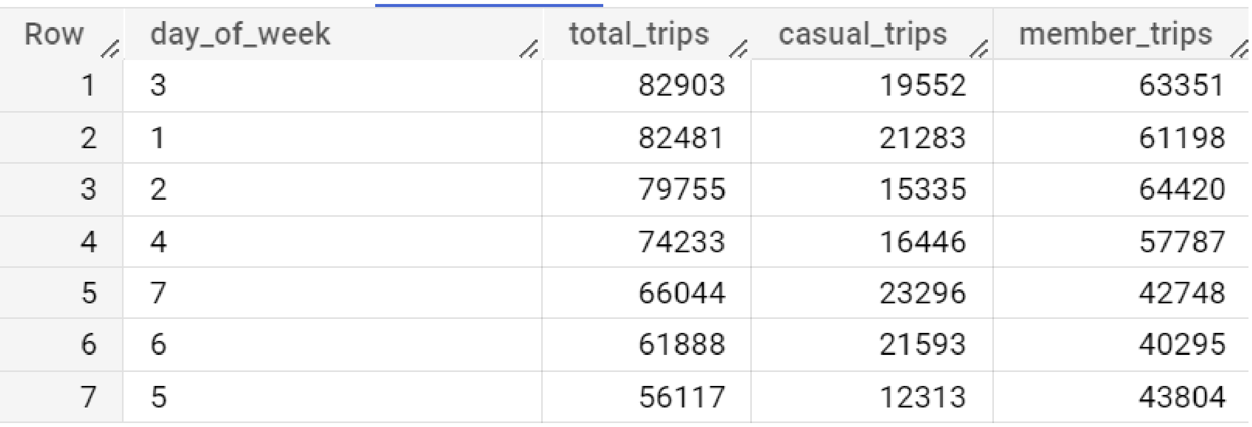


```
#Calculating total trips in a day_of_week for casual riders and annual members in each quarter of the year

SELECT
        day_of_week,
        COUNT(DISTINCT ride_id) AS total_trips,
        COUNTIF(member_casual='casual') AS casual_trips,
        COUNTIF(member_casual='member') AS member_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        1
ORDER BY
        total_trips DESC
```

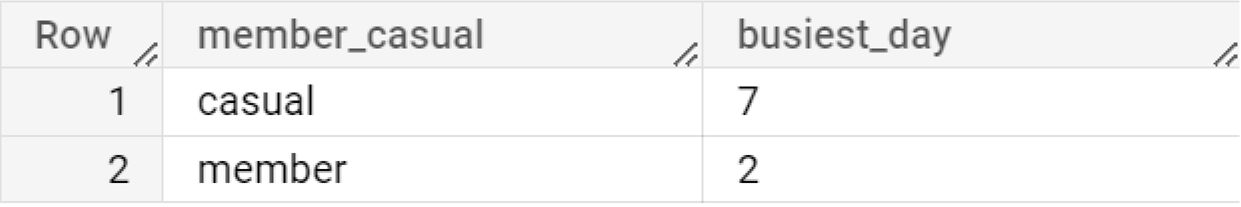


```
#Determine busiest day of the week for casual riders and annual members in each quarter of the year

SELECT
        member_casual,
        day_of_week AS busiest_day # Top number of day_of_week
FROM
        ( SELECT
        DISTINCT
                member_casual, day_of_week,
        ROW_NUMBER() OVER (PARTITION BY member_casual ORDER BY COUNT(day_of_week) DESC) RN
        #row number to assign each row as sequential number based on partition order by certain column
        FROM
                `cyclistic.tripdata_Q1`
        GROUP BY
                member_casual, day_of_week )
WHERE
        RN = 1 #row number starts from 1st row
```

Q1

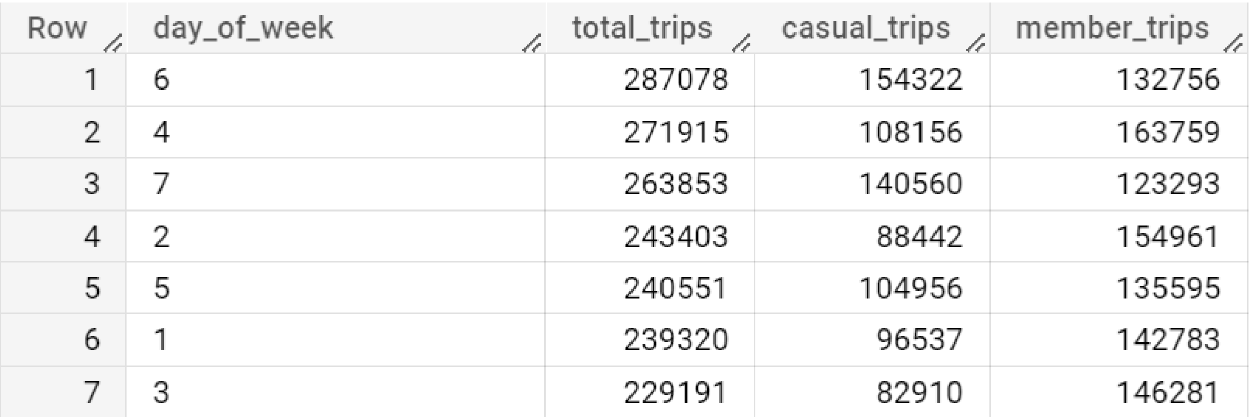


Row	day_of_week	total_trips	casual_trips	member_trips
1	3	82903	19552	63351
2	1	82481	21283	61198
3	2	79755	15335	64420
4	4	74233	16446	57787
5	7	66044	23296	42748
6	6	61888	21593	40295
7	5	56117	12313	43804

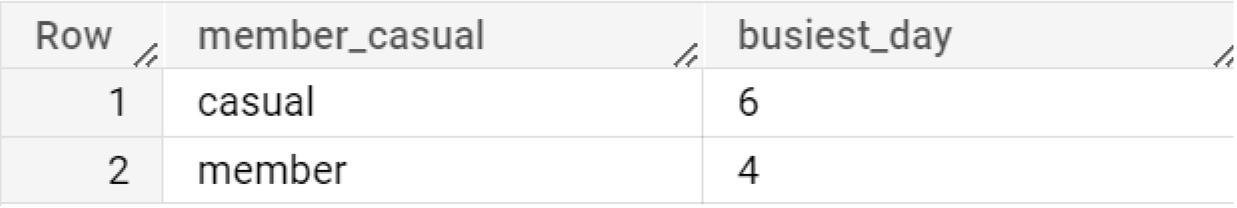


Row	member_casual	busiest_day
1	casual	7
2	member	2

Q2

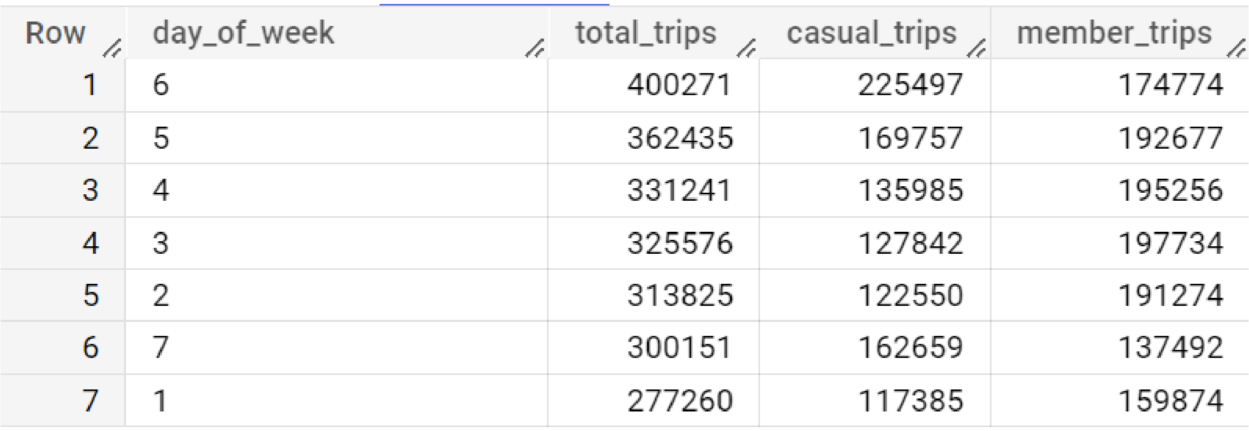


Row	day_of_week	total_trips	casual_trips	member_trips
1	6	287078	154322	132756
2	4	271915	108156	163759
3	7	263853	140560	123293
4	2	243403	88442	154961
5	5	240551	104956	135595
6	1	239320	96537	142783
7	3	229191	82910	146281

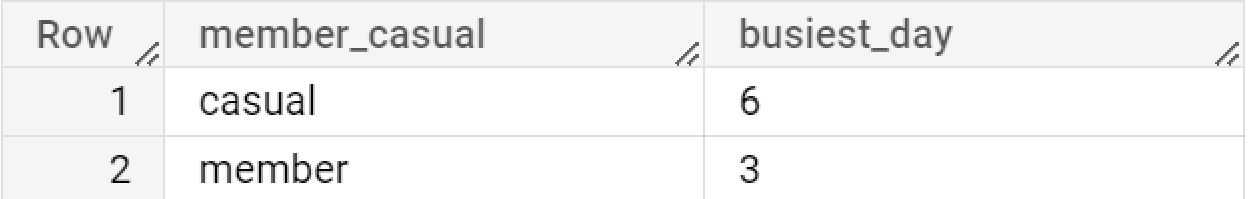


Row	member_casual	busiest_day
1	casual	6
2	member	4

Q3

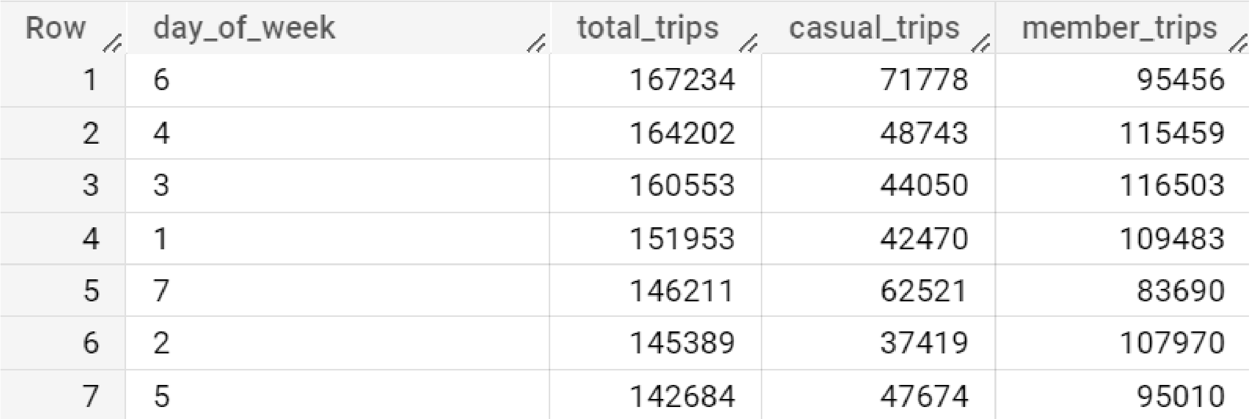


Row	day_of_week	total_trips	casual_trips	member_trips
1	6	400271	225497	174774
2	5	362435	169757	192677
3	4	331241	135985	195256
4	3	325576	127842	197734
5	2	313825	122550	191274
6	7	300151	162659	137492
7	1	277260	117385	159874

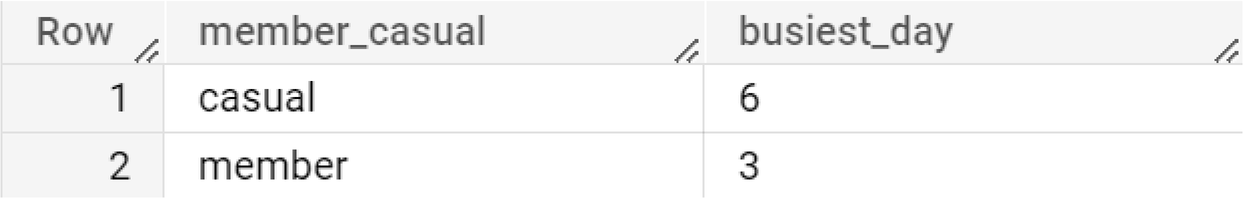


Row	member_casual	busiest_day
1	casual	6
2	member	3

Q4



Row	day_of_week	total_trips	casual_trips	member_trips
1	6	167234	71778	95456
2	4	164202	48743	115459
3	3	160553	44050	116503
4	1	151953	42470	109483
5	7	146211	62521	83690
6	2	145389	37419	107970
7	5	142684	47674	95010

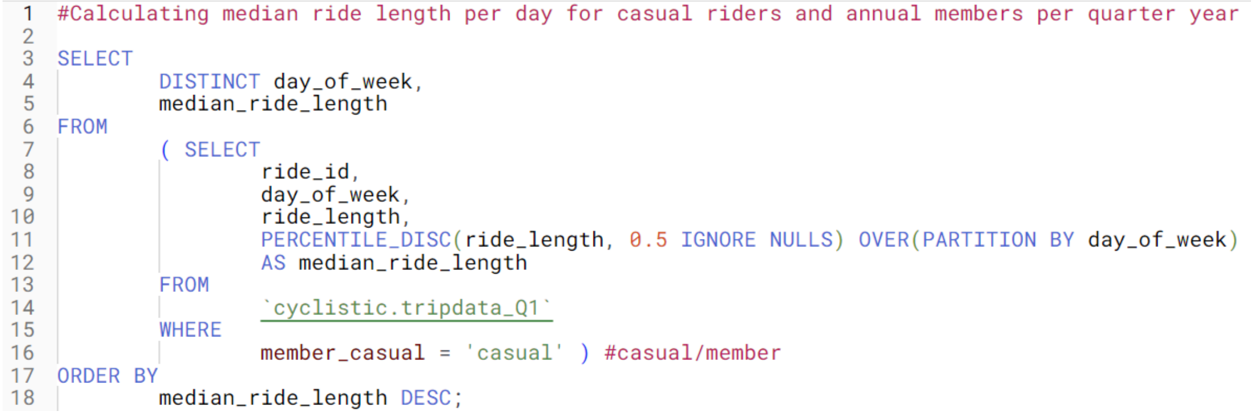


Row	member_casual	busiest_day
1	casual	6
2	member	3

Saturday followed by Sunday is the busiest_day for casual riders for most of the quarter year which is understandable because both days are Weekend, while annual members differ among Wednesday, Thursday, and Tuesday. From overall total_trips each day in the year, the busiest_day is Saturday.

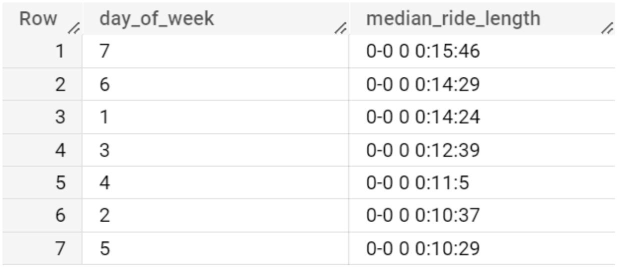
MEDIAN RIDE LENGTH PER DAY

Median ride_length each day determined for casual riders (*left*) and annual members (*right*).

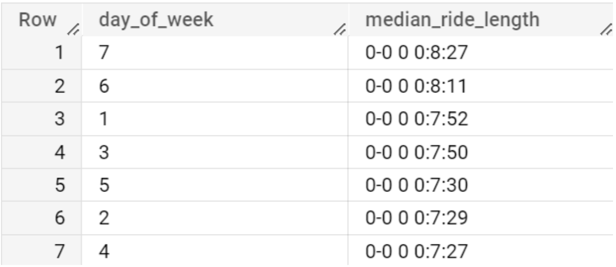


```
#Calculating median ride length per day for casual riders and annual members per quarter year

SELECT
        DISTINCT day_of_week,
        median_ride_length
FROM
        ( SELECT
                ride_id,
                day_of_week,
                ride_length,
                PERCENTILE_DISC(ride_length, 0.5 IGNORE NULLS) OVER(PARTITION BY day_of_week)
                AS median_ride_length
        FROM
                `cyclistic.tripdata_Q1`
        WHERE
                member_casual = 'casual' ) #casual/member
ORDER BY
        median_ride_length DESC;
```

Q1

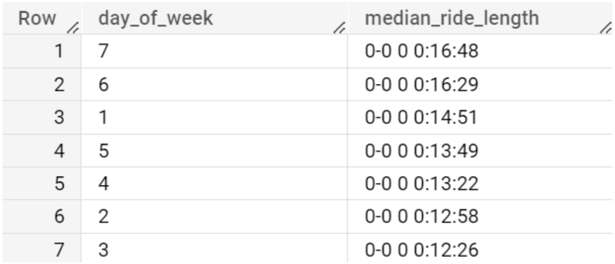


Row	day_of_week	median_ride_length
1	7	0-0 0 0:15:46
2	6	0-0 0 0:14:29
3	1	0-0 0 0:14:24
4	3	0-0 0 0:12:39
5	4	0-0 0 0:11:5
6	2	0-0 0 0:10:37
7	5	0-0 0 0:10:29

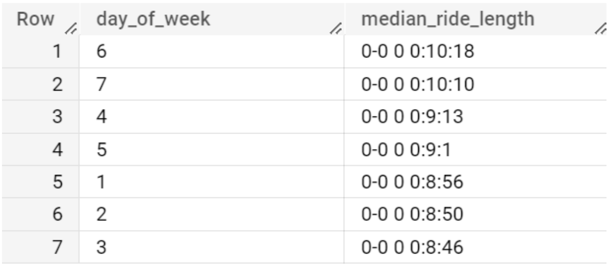


Row	day_of_week	median_ride_length
1	7	0-0 0 0:8:27
2	6	0-0 0 0:8:11
3	1	0-0 0 0:7:52
4	3	0-0 0 0:7:50
5	5	0-0 0 0:7:30
6	2	0-0 0 0:7:29
7	4	0-0 0 0:7:27

Q2

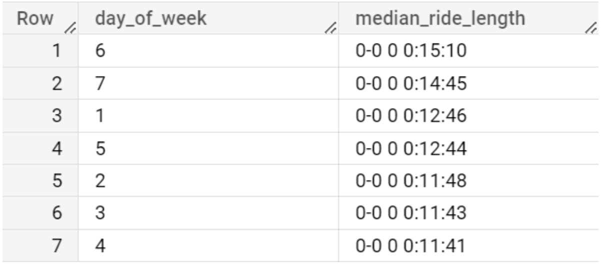


Row	day_of_week	median_ride_length
1	7	0-0 0 0:16:48
2	6	0-0 0 0:16:29
3	1	0-0 0 0:14:51
4	5	0-0 0 0:13:49
5	4	0-0 0 0:13:22
6	2	0-0 0 0:12:58
7	3	0-0 0 0:12:26

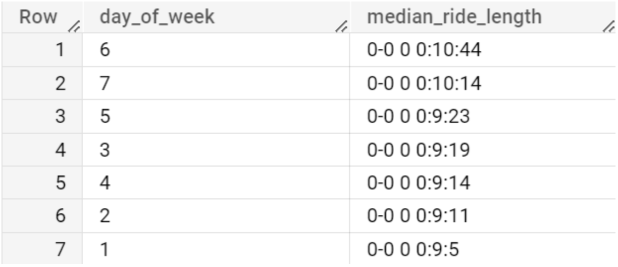


Row	day_of_week	median_ride_length
1	6	0-0 0 0:10:18
2	7	0-0 0 0:10:10
3	4	0-0 0 0:9:13
4	5	0-0 0 0:9:1
5	1	0-0 0 0:8:56
6	2	0-0 0 0:8:50
7	3	0-0 0 0:8:46

Q3

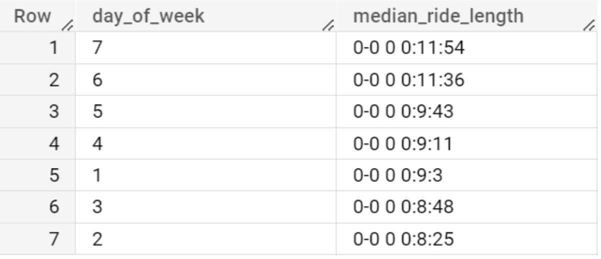


Row	day_of_week	median_ride_length
1	6	0-0 0 0:15:10
2	7	0-0 0 0:14:45
3	1	0-0 0 0:12:46
4	5	0-0 0 0:12:44
5	2	0-0 0 0:11:48
6	3	0-0 0 0:11:43
7	4	0-0 0 0:11:41

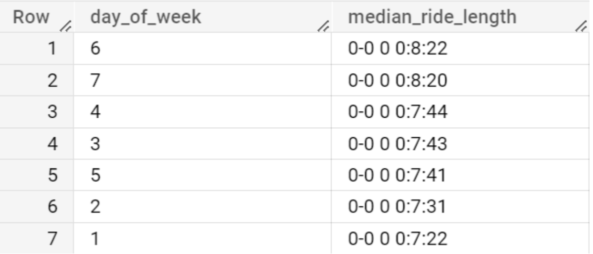


Row	day_of_week	median_ride_length
1	6	0-0 0 0:10:44
2	7	0-0 0 0:10:14
3	5	0-0 0 0:9:23
4	3	0-0 0 0:9:19
5	4	0-0 0 0:9:14
6	2	0-0 0 0:9:11
7	1	0-0 0 0:9:5

Q4



Row	day_of_week	median_ride_length
1	7	0-0 0 0:11:54
2	6	0-0 0 0:11:36
3	5	0-0 0 0:9:43
4	4	0-0 0 0:9:11
5	1	0-0 0 0:9:3
6	3	0-0 0 0:8:48
7	2	0-0 0 0:8:25

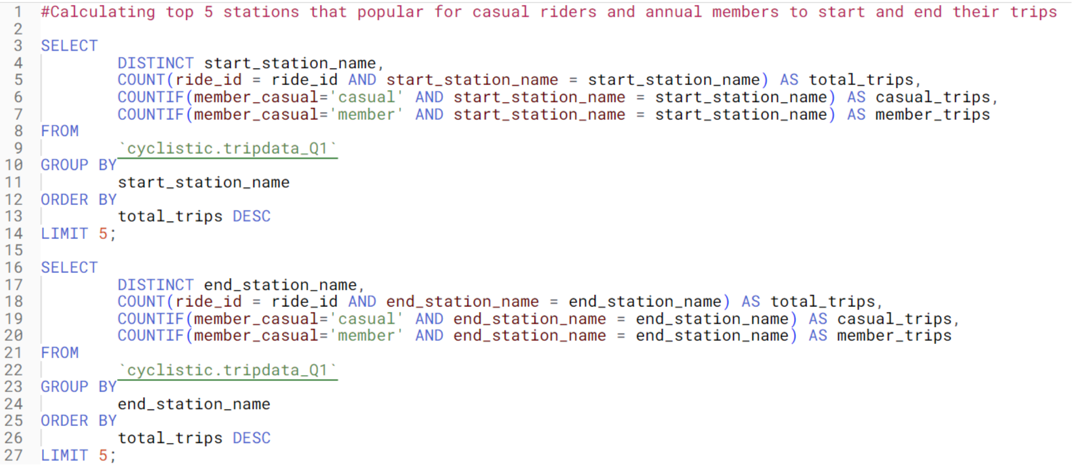


Row	day_of_week	median_ride_length
1	6	0-0 0 0:8:22
2	7	0-0 0 0:8:20
3	4	0-0 0 0:7:44
4	3	0-0 0 0:7:43
5	5	0-0 0 0:7:41
6	2	0-0 0 0:7:31
7	1	0-0 0 0:7:22

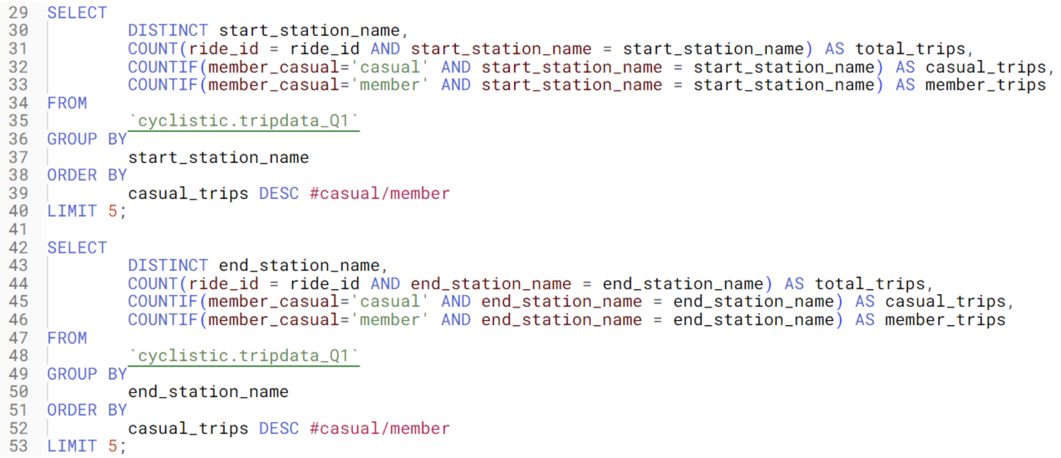
Days with the highest median ride_length for this year are Sunday and Saturday (Weekend) that showed quite same trends with busiest_day results. Median ride_length each day for casual riders is longer than annual members where most of them are more than 10 minutes, while annual members are less than it.

TOP START STATION & END STATION

We will look at the top 5 stations that popular for overall, casual riders (*left*) and annual members (*right*) to start and end their trips for each quarter of the year.

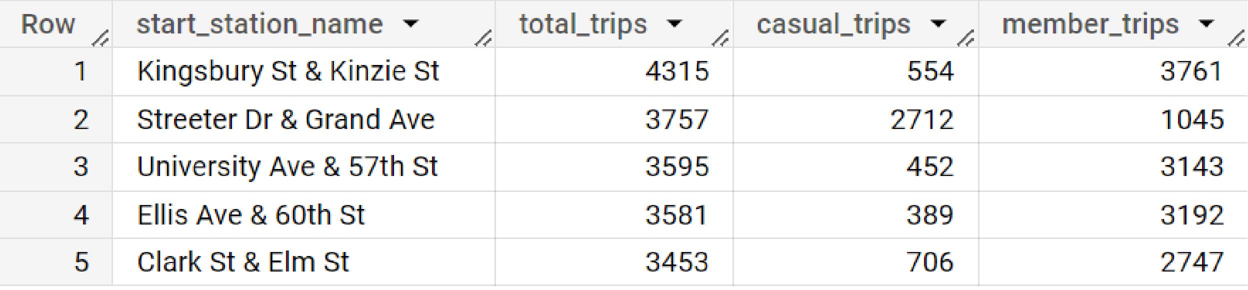


```
#Calculating top 5 stations that popular for casual riders and annual members to start and end their trips

SELECT
        DISTINCT start_station_name,
        COUNT(ride_id = ride_id AND start_station_name = start_station_name) AS total_trips,
        COUNTIF(member_casual='casual' AND start_station_name = start_station_name) AS casual_trips,
        COUNTIF(member_casual='member' AND start_station_name = start_station_name) AS member_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        start_station_name
ORDER BY
        total_trips DESC
LIMIT 5;

SELECT
        DISTINCT end_station_name,
        COUNT(ride_id = ride_id AND end_station_name = end_station_name) AS total_trips,
        COUNTIF(member_casual='casual' AND end_station_name = end_station_name) AS casual_trips,
        COUNTIF(member_casual='member' AND end_station_name = end_station_name) AS member_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        end_station_name
ORDER BY
        total_trips DESC
LIMIT 5;
```

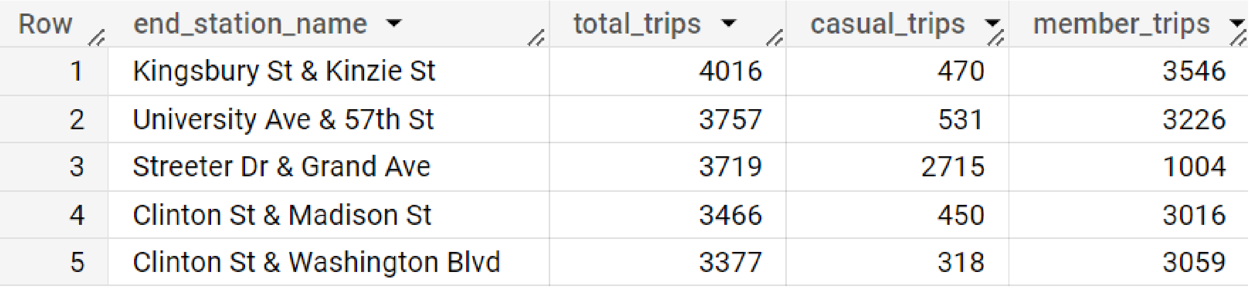


```
SELECT
        DISTINCT start_station_name,
        COUNT(ride_id = ride_id AND start_station_name = start_station_name) AS total_trips,
        COUNTIF(member_casual='casual' AND start_station_name = start_station_name) AS casual_trips,
        COUNTIF(member_casual='member' AND start_station_name = start_station_name) AS member_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        start_station_name
ORDER BY
        casual_trips DESC #casual/member
LIMIT 5;

SELECT
        DISTINCT end_station_name,
        COUNT(ride_id = ride_id AND end_station_name = end_station_name) AS total_trips,
        COUNTIF(member_casual='casual' AND end_station_name = end_station_name) AS casual_trips,
        COUNTIF(member_casual='member' AND end_station_name = end_station_name) AS member_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        end_station_name
ORDER BY
        casual_trips DESC #casual/member
LIMIT 5;
```

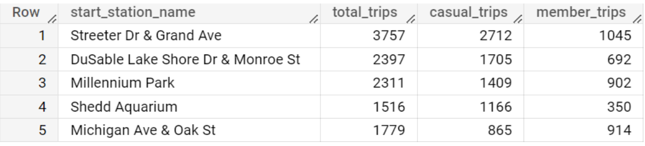
Q1



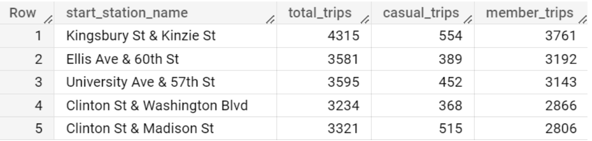
Row	start_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	4315	554	3761
2	Streeter Dr & Grand Ave	3757	2712	1045
3	University Ave & 57th St	3595	452	3143
4	Ellis Ave & 60th St	3581	389	3192
5	Clark St & Elm St	3453	706	2747



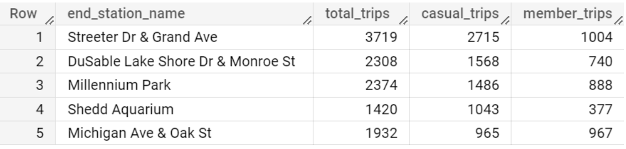
Row	end_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	4016	470	3546
2	University Ave & 57th St	3757	531	3226
3	Streeter Dr & Grand Ave	3719	2715	1004
4	Clinton St & Madison St	3466	450	3016
5	Clinton St & Washington Blvd	3377	318	3059



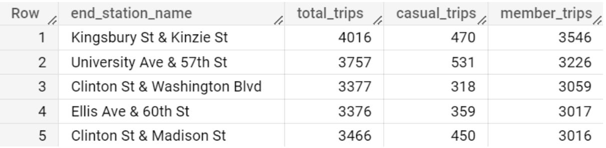
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	3757	2712	1045
2	DuSable Lake Shore Dr & Monroe St	2397	1705	692
3	Millennium Park	2311	1409	902
4	Shedd Aquarium	1516	1166	350
5	Michigan Ave & Oak St	1779	865	914



Row	start_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	4315	554	3761
2	Ellis Ave & 60th St	3581	389	3192
3	University Ave & 57th St	3595	452	3143
4	Clinton St & Washington Blvd	3234	368	2866
5	Clinton St & Madison St	3321	515	2806

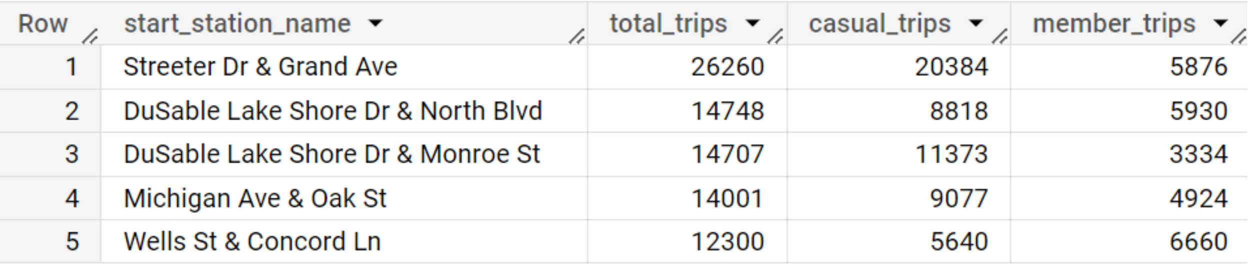


Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	3719	2715	1004
2	DuSable Lake Shore Dr & Monroe St	2308	1568	740
3	Millennium Park	2374	1486	888
4	Shedd Aquarium	1420	1043	377
5	Michigan Ave & Oak St	1932	965	967

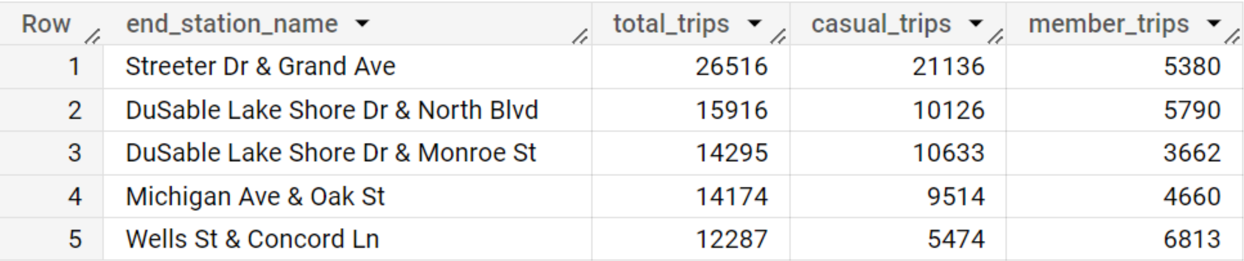


Row	end_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	4016	470	3546
2	University Ave & 57th St	3757	531	3226
3	Clinton St & Washington Blvd	3377	318	3059
4	Ellis Ave & 60th St	3376	359	3017
5	Clinton St & Madison St	3466	450	3016

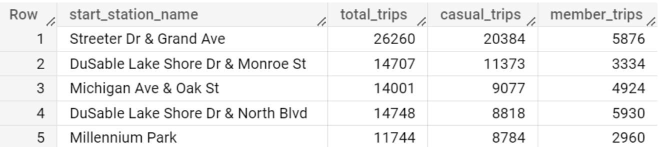
Q2



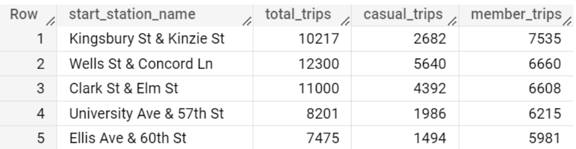
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	26260	20384	5876
2	DuSable Lake Shore Dr & North Blvd	14748	8818	5930
3	DuSable Lake Shore Dr & Monroe St	14707	11373	3334
4	Michigan Ave & Oak St	14001	9077	4924
5	Wells St & Concord Ln	12300	5640	6660



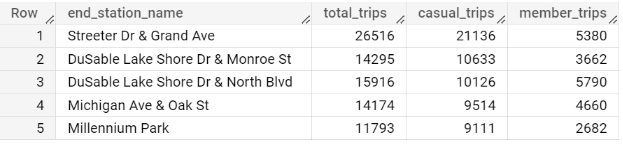
Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	26516	21136	5380
2	DuSable Lake Shore Dr & North Blvd	15916	10126	5790
3	DuSable Lake Shore Dr & Monroe St	14295	10633	3662
4	Michigan Ave & Oak St	14174	9514	4660
5	Wells St & Concord Ln	12287	5474	6813



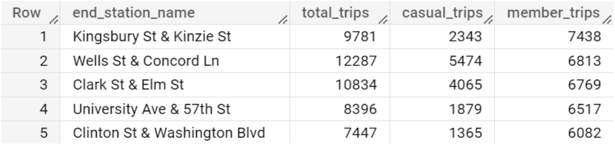
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	26260	20384	5876
2	DuSable Lake Shore Dr & Monroe St	14707	11373	3334
3	Michigan Ave & Oak St	14001	9077	4924
4	DuSable Lake Shore Dr & North Blvd	14748	8818	5930
5	Millennium Park	11744	8784	2960



Row	start_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	10217	2682	7535
2	Wells St & Concord Ln	12300	5640	6660
3	Clark St & Elm St	11000	4392	6608
4	University Ave & 57th St	8201	1986	6215
5	Ellis Ave & 60th St	7475	1494	5981

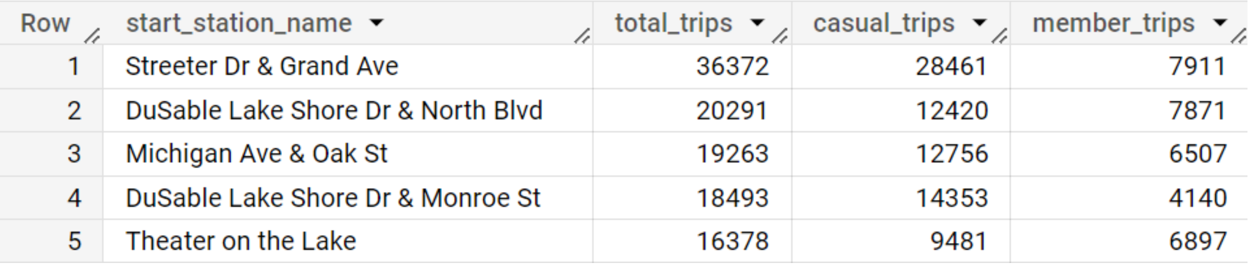


Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	26516	21136	5380
2	DuSable Lake Shore Dr & Monroe St	14295	10633	3662
3	DuSable Lake Shore Dr & North Blvd	15916	10126	5790
4	Michigan Ave & Oak St	14174	9514	4660
5	Millennium Park	11793	9111	2682

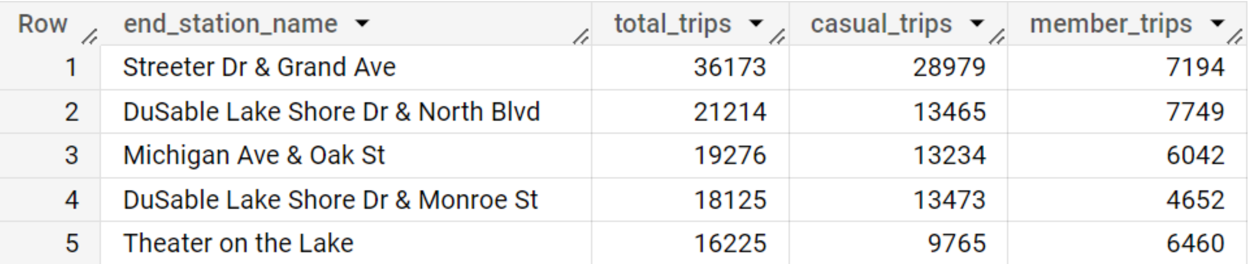


Row	end_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	9781	2343	7438
2	Wells St & Concord Ln	12287	5474	6813
3	Clark St & Elm St	10834	4065	6769
4	University Ave & 57th St	8396	1879	6517
5	Clinton St & Washington Blvd	7447	1365	6082

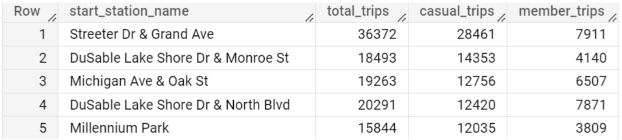
Q3



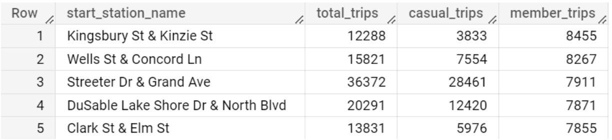
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	36372	28461	7911
2	DuSable Lake Shore Dr & North Blvd	20291	12420	7871
3	Michigan Ave & Oak St	19263	12756	6507
4	DuSable Lake Shore Dr & Monroe St	18493	14353	4140
5	Theater on the Lake	16378	9481	6897



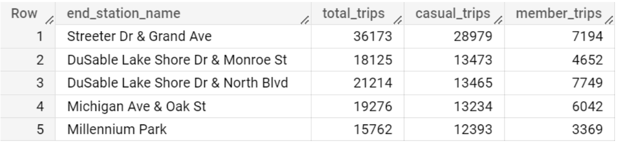
Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	36173	28979	7194
2	DuSable Lake Shore Dr & North Blvd	21214	13465	7749
3	Michigan Ave & Oak St	19276	13234	6042
4	DuSable Lake Shore Dr & Monroe St	18125	13473	4652
5	Theater on the Lake	16225	9765	6460



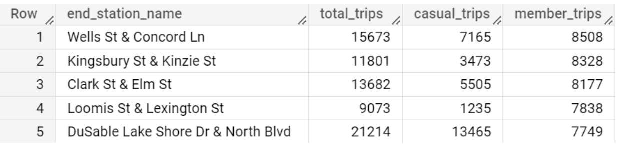
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	36372	28461	7911
2	DuSable Lake Shore Dr & Monroe St	18493	14353	4140
3	Michigan Ave & Oak St	19263	12756	6507
4	DuSable Lake Shore Dr & North Blvd	20291	12420	7871
5	Millennium Park	15844	12035	3809



Row	start_station_name	total_trips	casual_trips	member_trips
1	Kingsbury St & Kinzie St	12288	3833	8455
2	Wells St & Concord Ln	15821	7554	8267
3	Streeter Dr & Grand Ave	36372	28461	7911
4	DuSable Lake Shore Dr & North Blvd	20291	12420	7871
5	Clark St & Elm St	13831	5976	7855

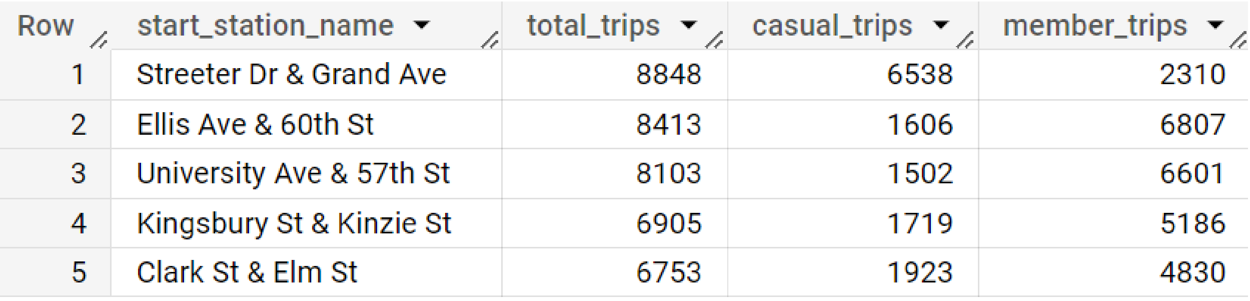


Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	36173	28979	7194
2	DuSable Lake Shore Dr & Monroe St	18125	13473	4652
3	DuSable Lake Shore Dr & North Blvd	21214	13465	7749
4	Michigan Ave & Oak St	19276	13234	6042
5	Millennium Park	15762	12393	3369

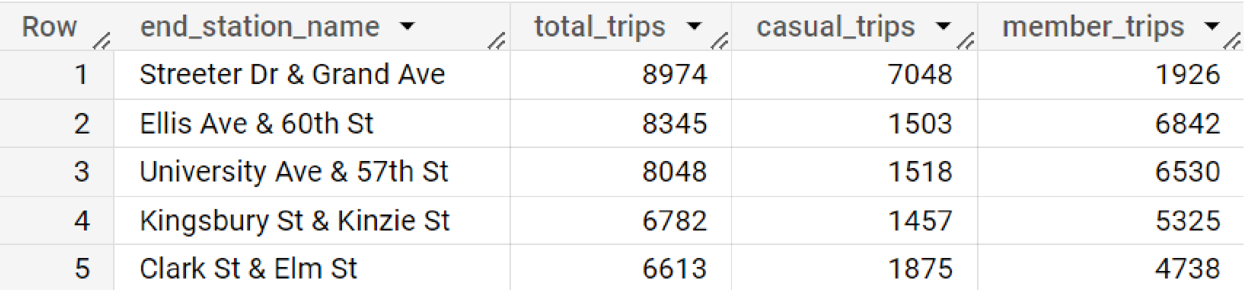


Row	end_station_name	total_trips	casual_trips	member_trips
1	Wells St & Concord Ln	15673	7165	8508
2	Kingsbury St & Kinzie St	11801	3473	8328
3	Clark St & Elm St	13682	5505	8177
4	Loomis St & Lexington St	9073	1235	7838
5	DuSable Lake Shore Dr & North Blvd	21214	13465	7749

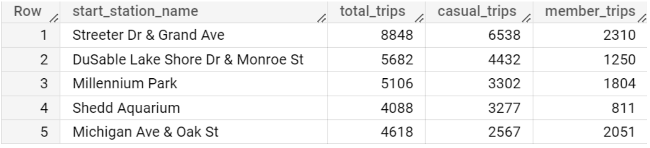
Q4



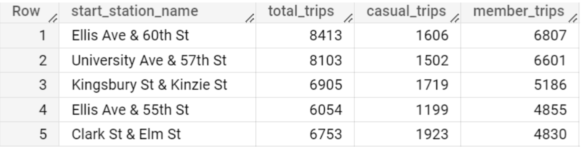
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	8848	6538	2310
2	Ellis Ave & 60th St	8413	1606	6807
3	University Ave & 57th St	8103	1502	6601
4	Kingsbury St & Kinzie St	6905	1719	5186
5	Clark St & Elm St	6753	1923	4830



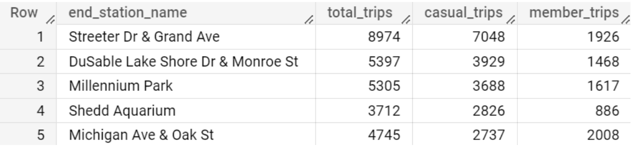
Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	8974	7048	1926
2	Ellis Ave & 60th St	8345	1503	6842
3	University Ave & 57th St	8048	1518	6530
4	Kingsbury St & Kinzie St	6782	1457	5325
5	Clark St & Elm St	6613	1875	4738



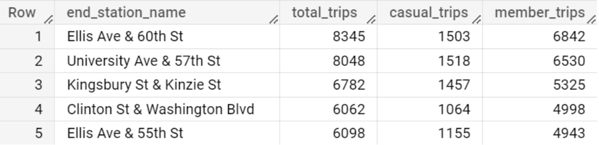
Row	start_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	8848	6538	2310
2	DuSable Lake Shore Dr & Monroe St	5682	4432	1250
3	Millennium Park	5106	3302	1804
4	Shedd Aquarium	4088	3277	811
5	Michigan Ave & Oak St	4618	2567	2051



Row	start_station_name	total_trips	casual_trips	member_trips
1	Ellis Ave & 60th St	8413	1606	6807
2	University Ave & 57th St	8103	1502	6601
3	Kingsbury St & Kinzie St	6905	1719	5186
4	Ellis Ave & 55th St	6054	1199	4855
5	Clark St & Elm St	6753	1923	4830



Row	end_station_name	total_trips	casual_trips	member_trips
1	Streeter Dr & Grand Ave	8974	7048	1926
2	DuSable Lake Shore Dr & Monroe St	5397	3929	1468
3	Millennium Park	5305	3688	1617
4	Shedd Aquarium	3712	2826	886
5	Michigan Ave & Oak St	4745	2737	2008

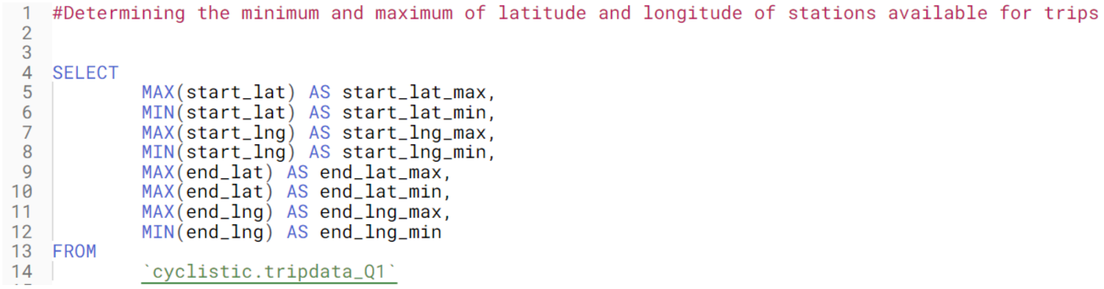


Row	end_station_name	total_trips	casual_trips	member_trips
1	Ellis Ave & 60th St	8345	1503	6842
2	University Ave & 57th St	8048	1518	6530
3	Kingsbury St & Kinzie St	6782	1457	5325
4	Clinton St & Washington Blvd	6062	1064	4998
5	Ellis Ave & 55th St	6098	1155	4943

The patterns we can find from the lists is the contrast of stations favored by casual riders and annual members. Most of the start stations are also popular as end stations of their trips throughout the year. Streeter Dr & Grand Ave is the most popular station for casual riders, while most of the time Kingsburry St & Kinzie St is popular for annual members.

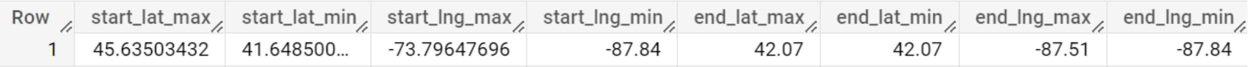
MINIMUM AND MAXIMUM LATITUDE AND LONGITUDE OF STATION

We also determine the minimum and maximum of latitude and longitude of stations available for trips.



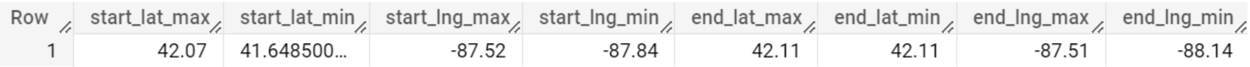
```
#Determining the minimum and maximum of latitude and longitude of stations available for trips


SELECT
        MAX(start_lat) AS start_lat_max,
        MIN(start_lat) AS start_lat_min,
        MAX(start_lng) AS start_lng_max,
        MIN(start_lng) AS start_lng_min,
        MAX(end_lat) AS end_lat_max,
        MAX(end_lat) AS end_lat_min,
        MAX(end_lng) AS end_lng_max,
        MIN(end_lng) AS end_lng_min
FROM
        `cyclistic.tripdata_Q1`
```

Q1



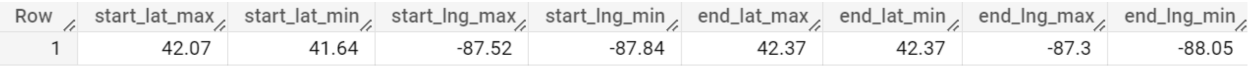
Row	start_lat_max	start_lat_min	start_lng_max	start_lng_min	end_lat_max	end_lat_min	end_lng_max	end_lng_min
1	45.63503432	41.648500...	-73.79647696	-87.84	42.07	42.07	-87.51	-87.84

Q2



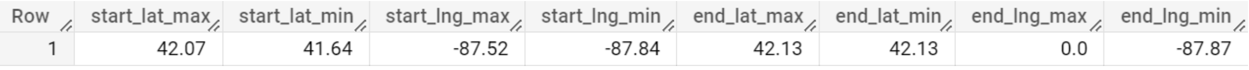
Row	start_lat_max	start_lat_min	start_lng_max	start_lng_min	end_lat_max	end_lat_min	end_lng_max	end_lng_min
1	42.07	41.648500...	-87.52	-87.84	42.11	42.11	-87.51	-88.14

Q3



Row	start_lat_max	start_lat_min	start_lng_max	start_lng_min	end_lat_max	end_lat_min	end_lng_max	end_lng_min
1	42.07	41.64	-87.52	-87.84	42.37	42.37	-87.3	-88.05

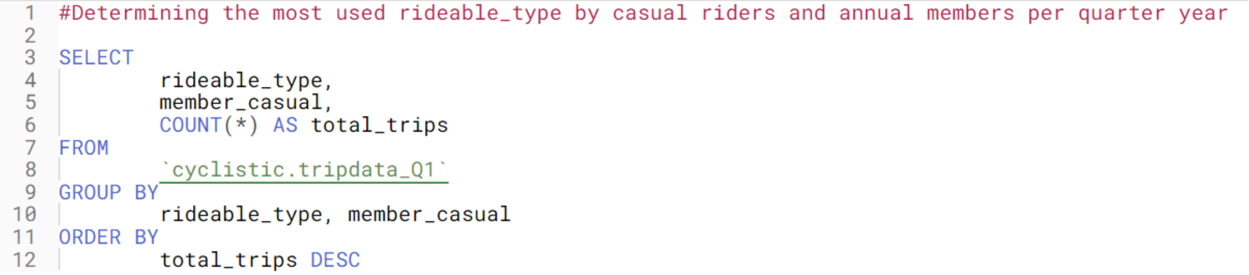
Q4



Row	start_lat_max	start_lat_min	start_lng_max	start_lng_min	end_lat_max	end_lat_min	end_lng_max	end_lng_min
1	42.07	41.64	-87.52	-87.84	42.13	42.13	0.0	-87.87

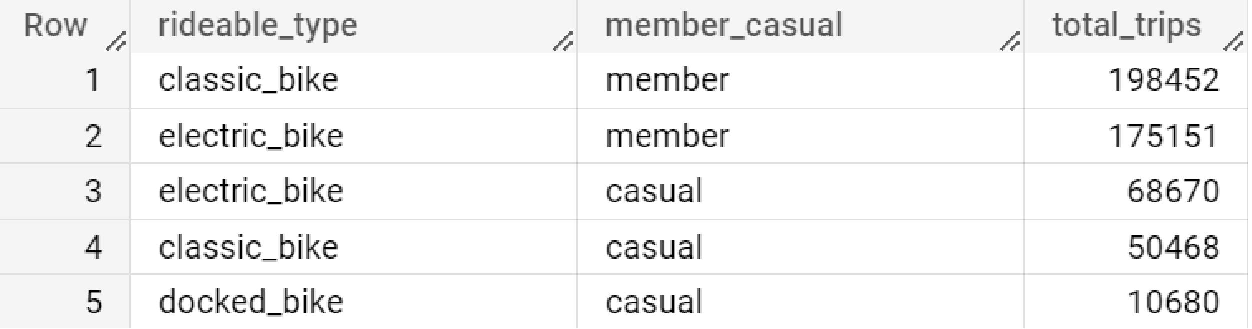
MOST RIDEABLE TYPE

We get to know the most used rideable_type by casual riders and annual members per quarter year.



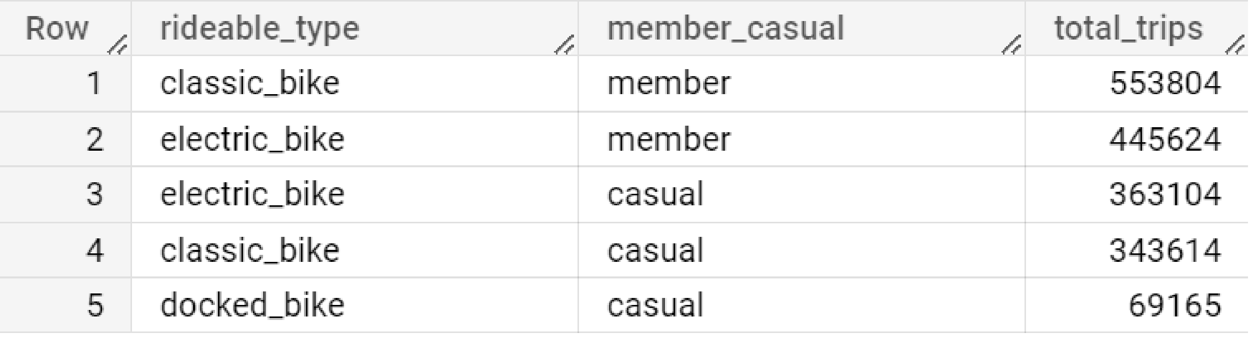
```
#Determining the most used rideable_type by casual riders and annual members per quarter year

SELECT
        rideable_type,
        member_casual,
        COUNT(*) AS total_trips
FROM
        `cyclistic.tripdata_Q1`
GROUP BY
        rideable_type, member_casual
ORDER BY
        total_trips DESC
```

Q1



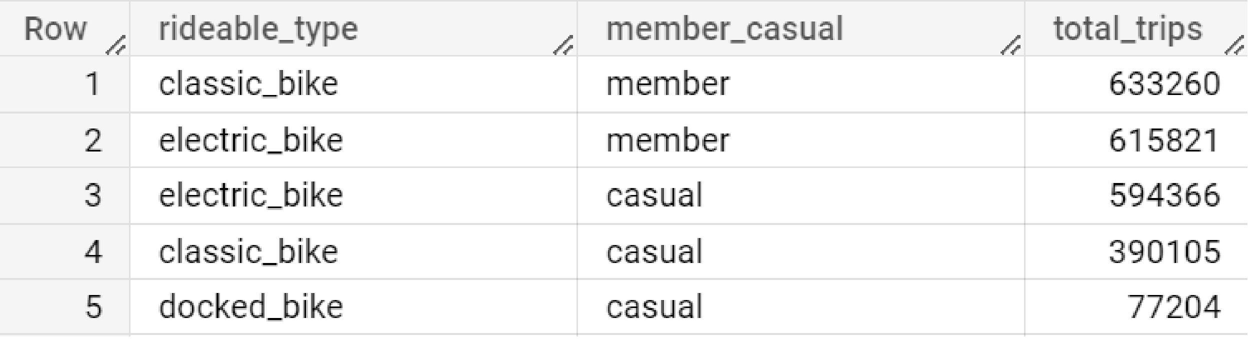
Row	rideable_type	member_casual	total_trips
1	classic_bike	member	198452
2	electric_bike	member	175151
3	electric_bike	casual	68670
4	classic_bike	casual	50468
5	docked_bike	casual	10680

Q2



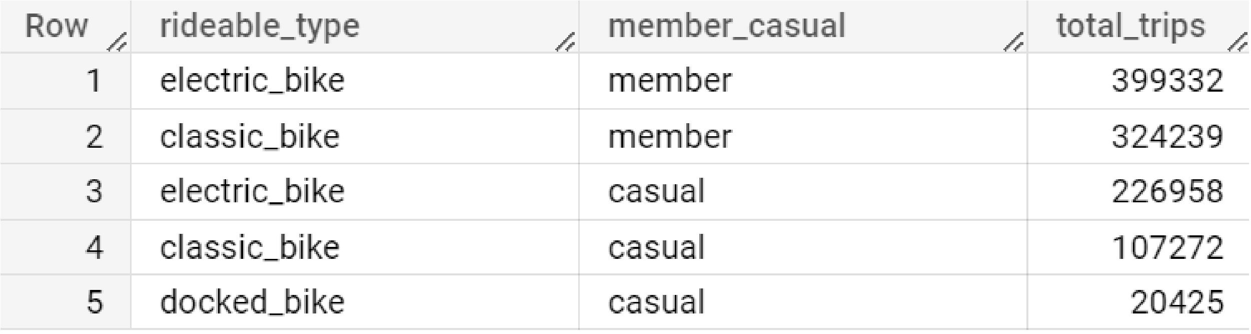
Row	rideable_type	member_casual	total_trips
1	classic_bike	member	553804
2	electric_bike	member	445624
3	electric_bike	casual	363104
4	classic_bike	casual	343614
5	docked_bike	casual	69165

Q3



Row	rideable_type	member_casual	total_trips
1	classic_bike	member	633260
2	electric_bike	member	615821
3	electric_bike	casual	594366
4	classic_bike	casual	390105
5	docked_bike	casual	77204

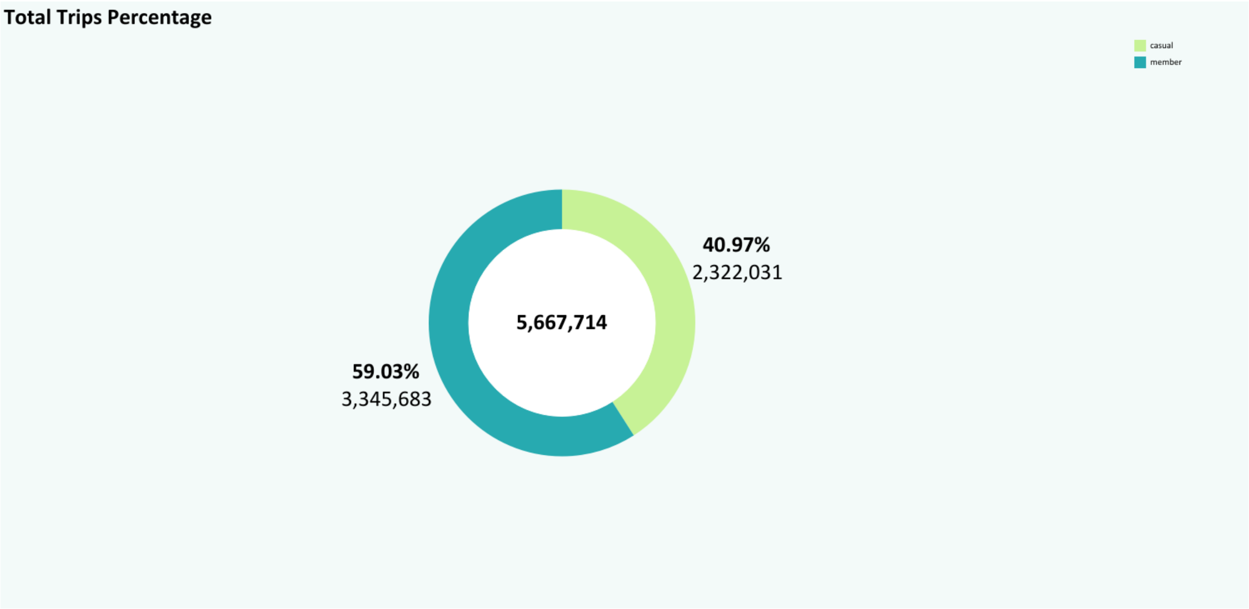
Q4



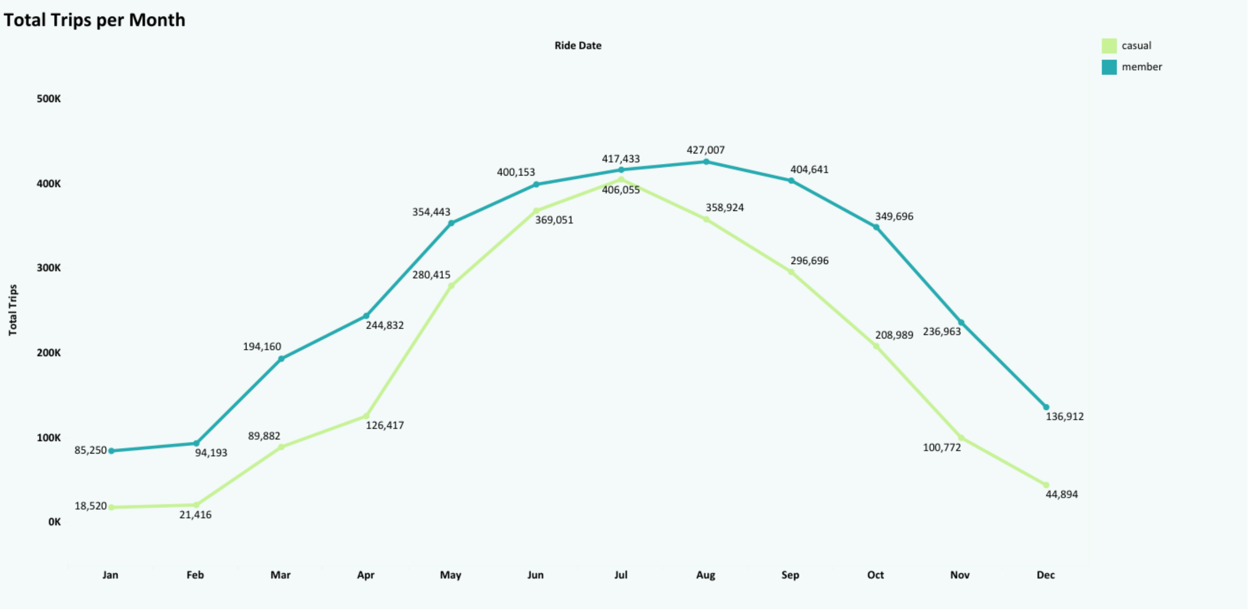
Row	rideable_type	member_casual	total_trips
1	electric_bike	member	399332
2	classic_bike	member	324239
3	electric_bike	casual	226958
4	classic_bike	casual	107272
5	docked_bike	casual	20425

The tables showed that the most used rideable_type for casual members is electric_bike, while for annual members is classic_bike, except for Q4 which is also electric_bike, and docked_bike being the less used through the year.

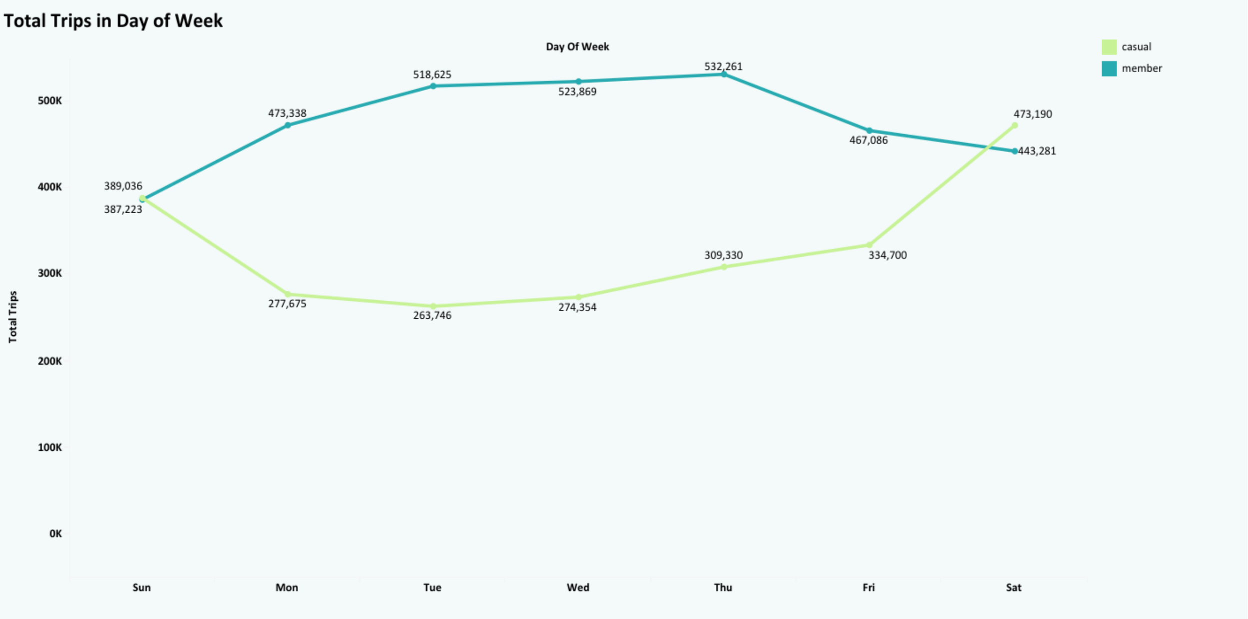
DATA ANALYSIS AND VISUALIZATION:



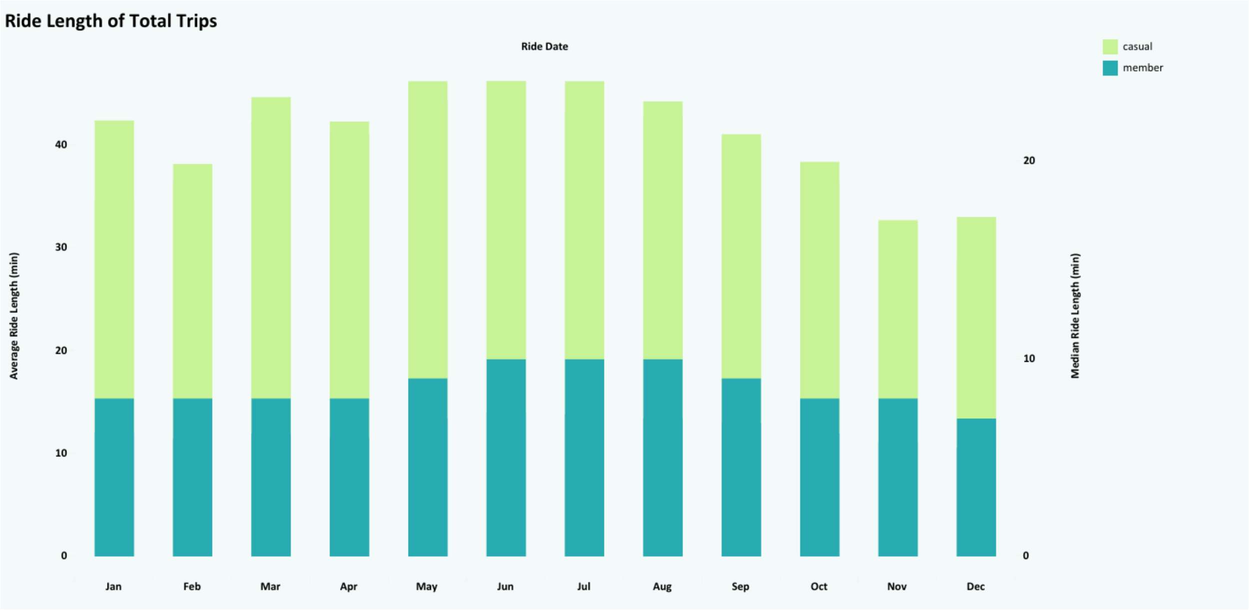
The total trip percentage of Cyclistic customers throughout the year showed that 59.03% customers are annual members, while the other 40.97% are casual riders.



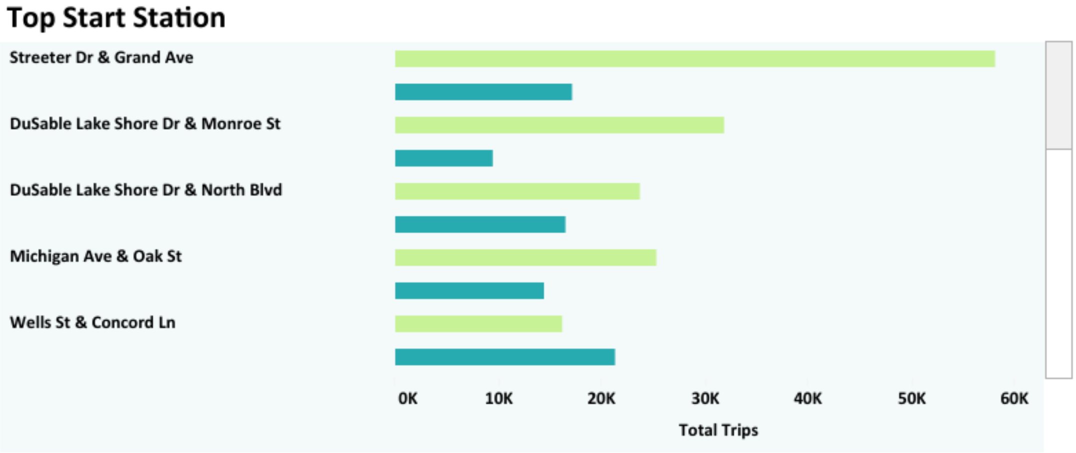
Total trips of annual members during the past 12 months exceed casual riders, yet casual riders' trips almost catch up with annual members' in June. Based on the data, we found a pattern of total trips influenced by seasons. Further analysis on a quarterly basis pointed out that the peak season for both customer types falls in Summer (Q3) and Spring (Q2) believed because at this range of time Chicago starts to feel warmer and the summer holiday also coming. The demand begins to increase in May (late spring) until September (early fall). The highest amount of total trips is 427,007 trips in August. However, the lowest total trips in early Winter did not even reach 100,000 trips.

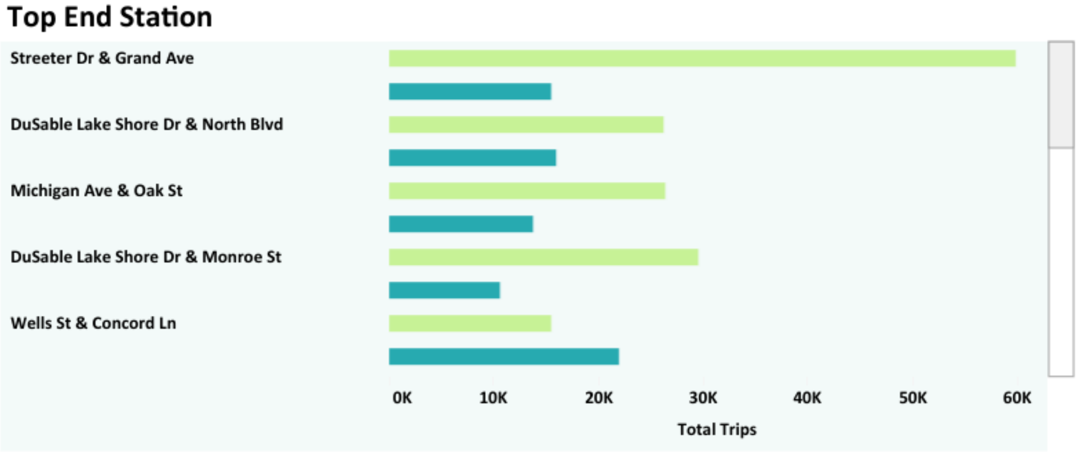


The popular days for annual members to do trips are between Tuesday to Thursday (mid-week), while the weekend consistently becomes the preference of casual riders. This trend indicated that annual members and casual riders have different purpose in using Cyclistic bikeshare. Annual members potentially take the trips for daily activities, like going to and from the office or school, so they need to use Cyclistic bikeshare routinely, whilst casual riders use Cyclistic bikeshare for recreation purpose on their off-day.

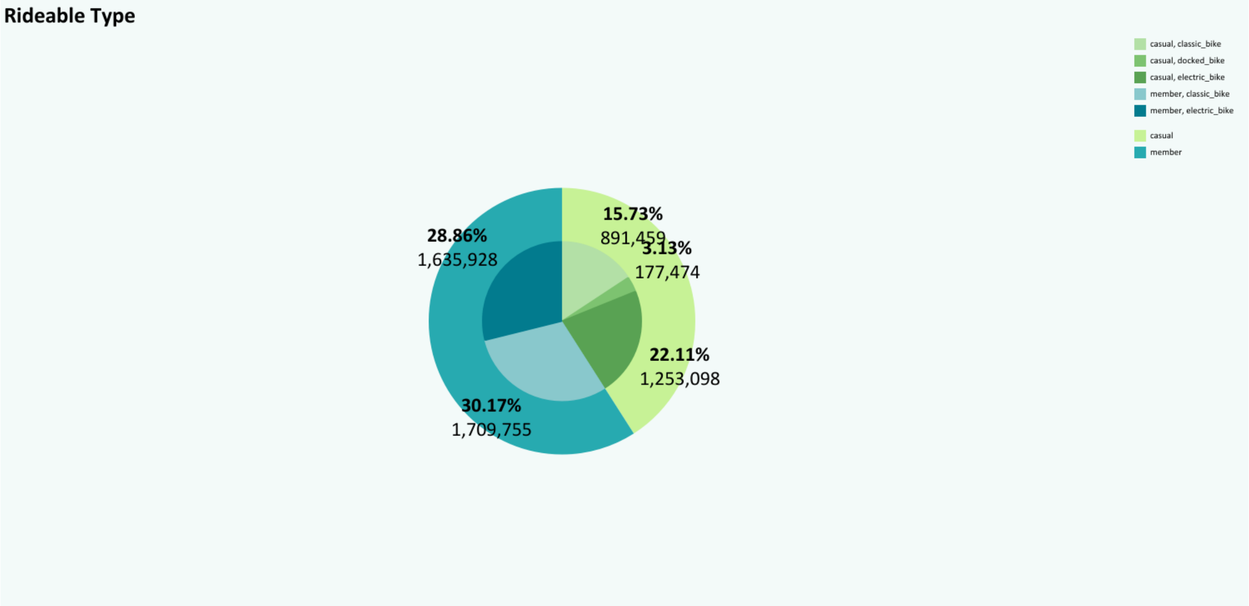


Ride length of total trips presented in minutes with both average and median. Focusing on the median results, as average results highly impacted with the outliers, annual members spent time around 7-10 minutes per trip, whereas casual riders mainly spent more than 10 minutes per trip. The results convinced that casual members spent their time for trips longer and more leisurely than the annual members as it is known casual members take the trips on their off-day on the weekend.





The busiest station locations for annual members and casual riders are quite different. Streeter Dr & Grand Ave station become the most crowded station for casual members over the year. The start stations and end stations for each customer most of the time are the same indicated that customers come from and go to the same routes.



Both classic bike and electric bike have almost the same percentage being annual members' favorite bikes with 30.17% and 28.86% respectively, while the electric bike becomes casual riders' favorite bike leading with 22.11%.

CONCLUSION:

Based on the insights obtained, Cyclistic should design marketing strategies by considering seasonal trends, usage day and time, station locations, and bike type preferences that affect how casual riders and annual members use Cyclistic bikes differently

RECOMMENDATIONS

1. Summer and Spring can be an opportunity for Cyclistic to create promotional campaigns that will attract casual riders to convert into annual members.
2. Prioritizing the weekend for marketing strategies implementation and focusing on giving more benefits for longer ride length casual riders who are willing to purchase an annual membership.
3. Top stations for casual riders, like Streeter Dr & Grand Ave are the best place to held promotional campaign locations.
4. Providing benefits for new joined annual members who prefer the electric bike as their rideable type.