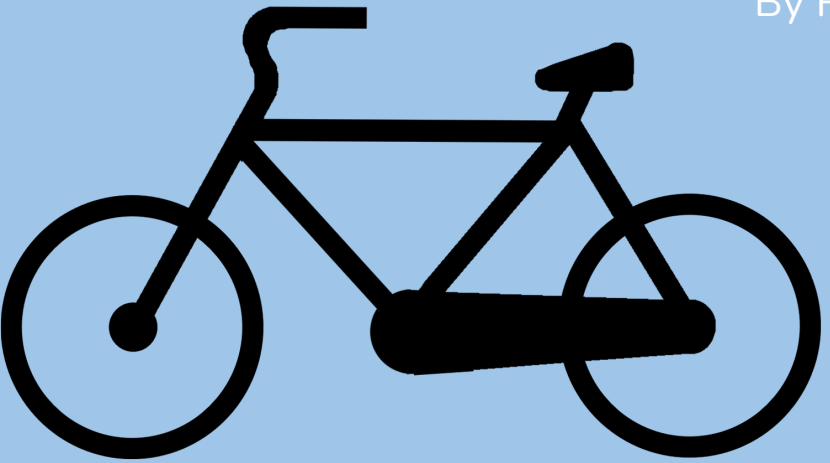
Cyclistic (Divvy) Bike-share Analysis Case Study

Date:1/6/2023
By Rita Lam



Background

Cyclistic (Divvy) provides bike-sharing service for riders who can unlock a bike from one station/location and return to another station/location at any time.

Casual riders can opt for single ride or day pass while members sign up for annual membership.

Annual memberships are more profitable but single rides or day pass provide the flexibility for riders which attract new customers.

Objective - Analyze how members and casual riders utilize bikes differently in order to determine marketing strategy for maximizing annual membership subscriptions

Consider the following factors:

- number of rides, ride duration vs time (time of day/month/year)
- Most popular bike stations and routes
- Usage of three bike types - electric, classic and docked

Data Source

Source: https://divvybikes.com/system-data

Public data provided by Motivate International Inc.

License: Data License Agreement | Divvy Bikes

Merged 12 monthly data sets from **Jan 2022 to Dec 2022** into one single csv data set for analysis

Data Integrity

Data Cleaning -Checked duplicate row and trimmed white space

-**2 separate analysis for data points with ride duration <3 hours and >3 hours** due to large data skew

- Removed rows with null values in bike return location

-Removed outliers

Data sets from single source

Primary data - original and credible

Data sets are current (Jan 2022 - Dec 2022)

Large data set ~ 5,500,000 data points for ride duration <3hrs (majority of data)
~15,000 data points for ride duration >3hrs

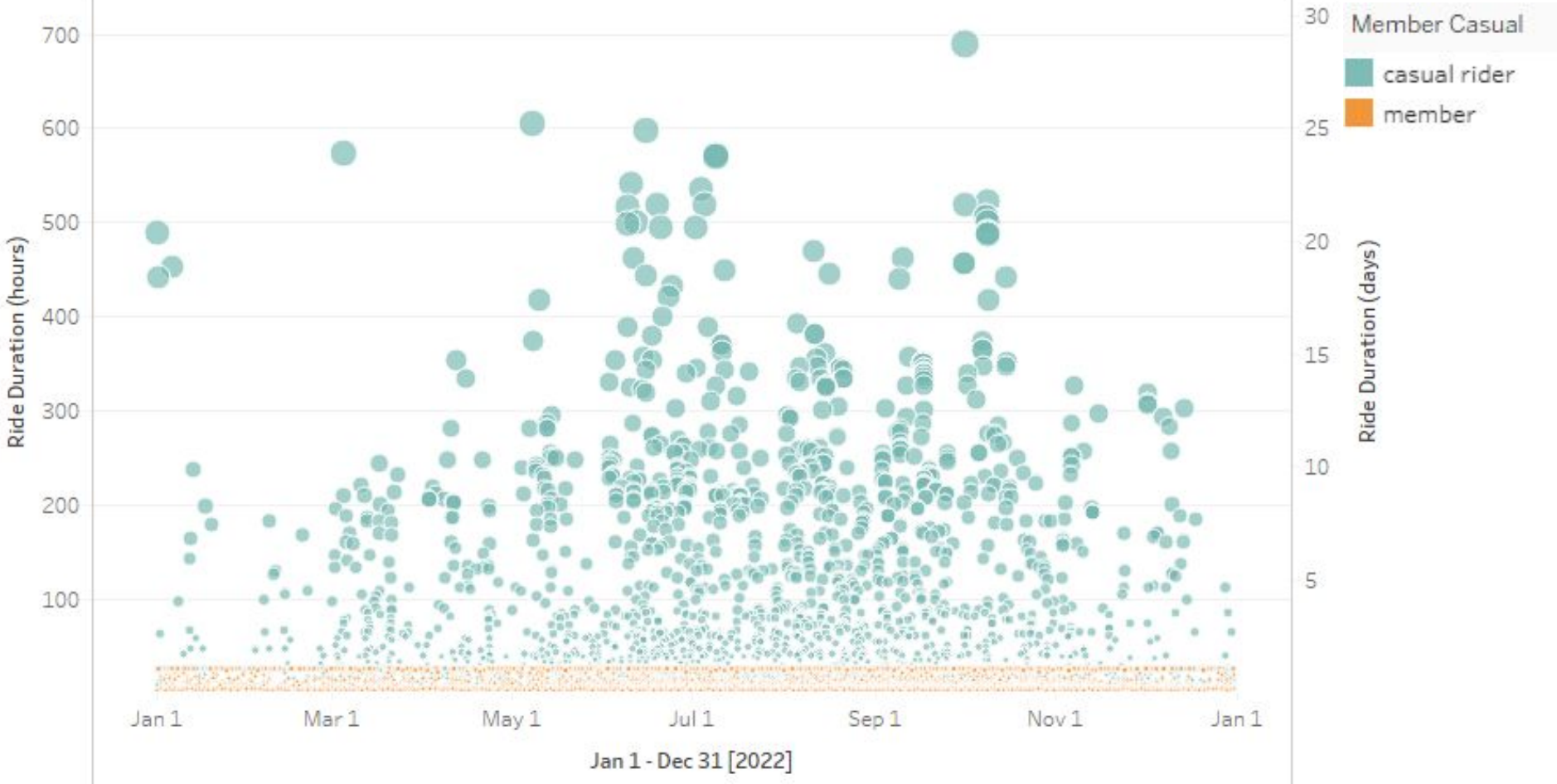
```
14    FROM "202212-divvy-tripdata"
15    WHERE ride_duration_hr>=25
16    AND member_casual IS NOT NULL
17    AND ride_duration_hr IS NOT NULL
18    AND rideable_type = 'docked_bike'
```

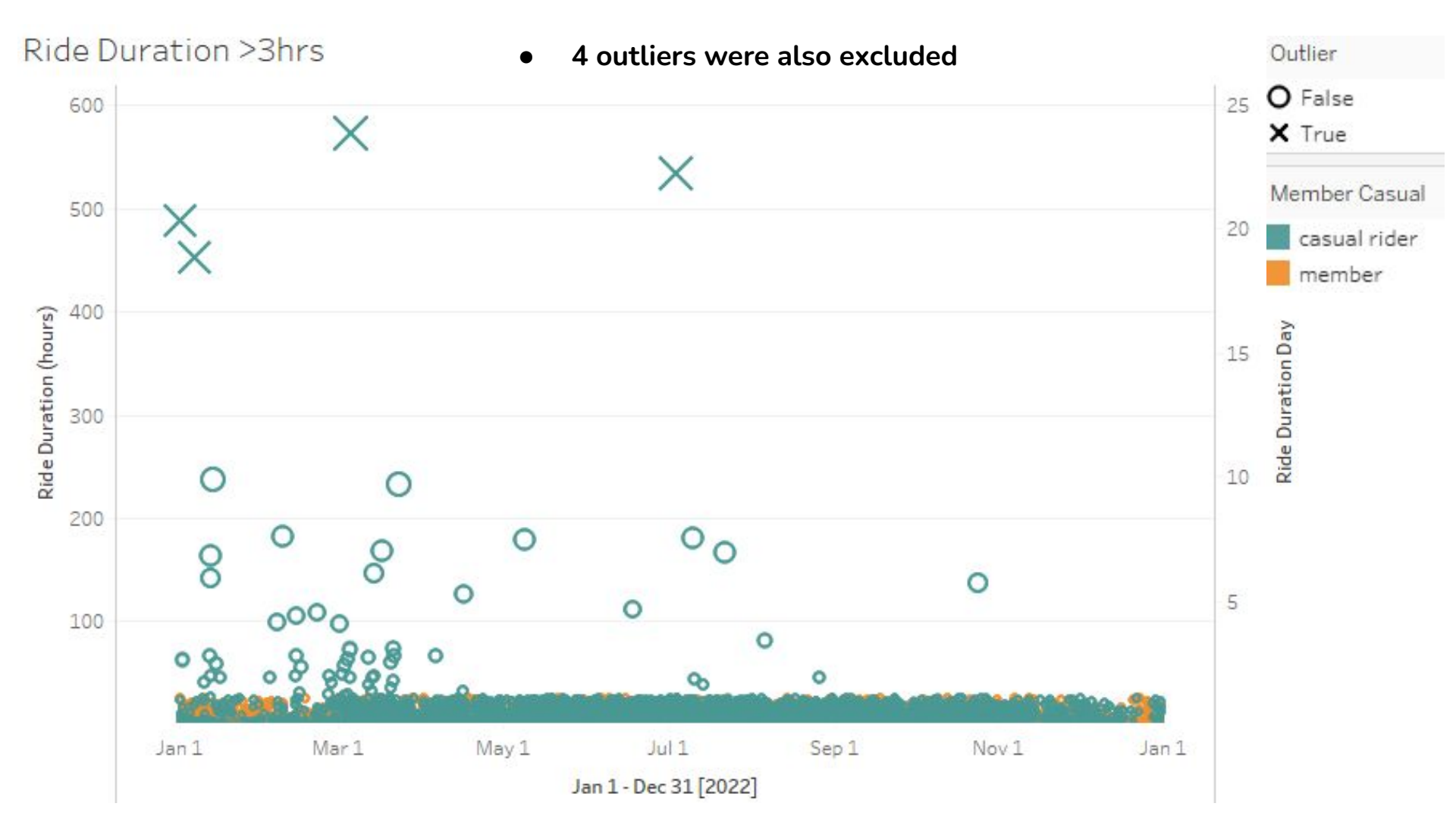
	member_casual	rideable_type	start_lat	start_lng	end_lat	end_lng	we
1	casual	docked_bike	41.860384	-87.625813	NULL	NULL	wee
2	casual	docked_bike	41.945529	-87.646439	NULL	NULL	wee
3	casual	docked_bike	41.889375	-87.627077	NULL	NULL	wee
4	casual	docked_bike	42.010587	-87.662412	NULL	NULL	wee
5	casual	docked_bike	41.886976	-87.612813	NULL	NULL	wee
6	casual	docked_bike	41.961108	-87.72897	NULL	NULL	wee
7	casual	docked_bike	41.883984	-87.624684	NULL	NULL	wee
8	casual	docked_bike	41.903486	-87.643353	NULL	NULL	wee
9	casual	docked_bike	41.857412	-87.613792	NULL	NULL	wee
10	casual	docked_bike	41.828792	-87.680604	NULL	NULL	wee
11	casual	docked_bike	41.93259	-87.665936	NULL	NULL	wee
12	casual	docked_bike	41.881032	-87.624084	NULL	NULL	wee
13	casual	docked_bike	41.880958	-87.616743	NULL	NULL	wee

- Bike return location becomes mostly null in data points with ride duration >=25 hours
- Need to verify if they are valid data points

Ride Duration >3hrs

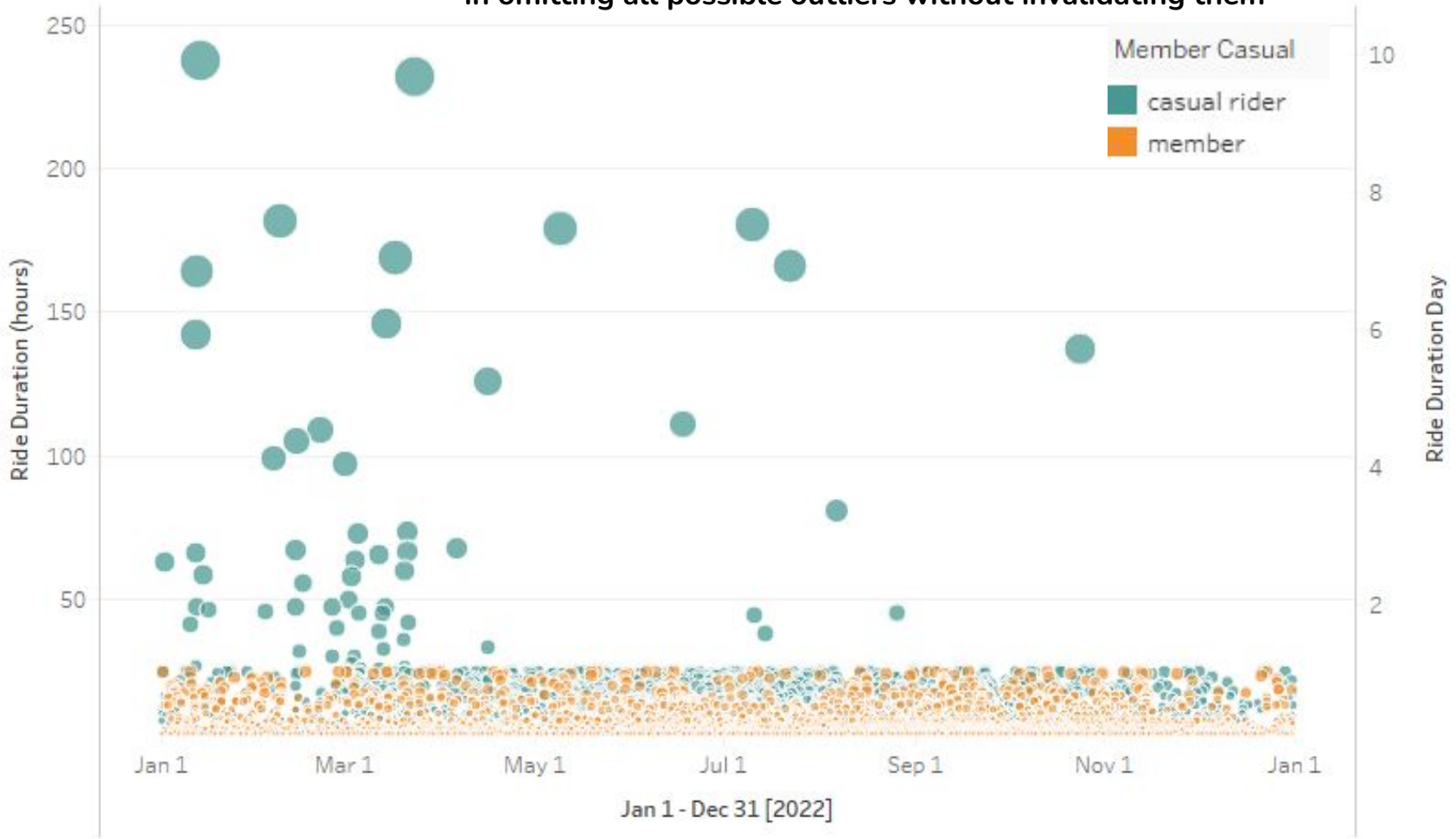
- **Scatter plot before eliminating NULL values**



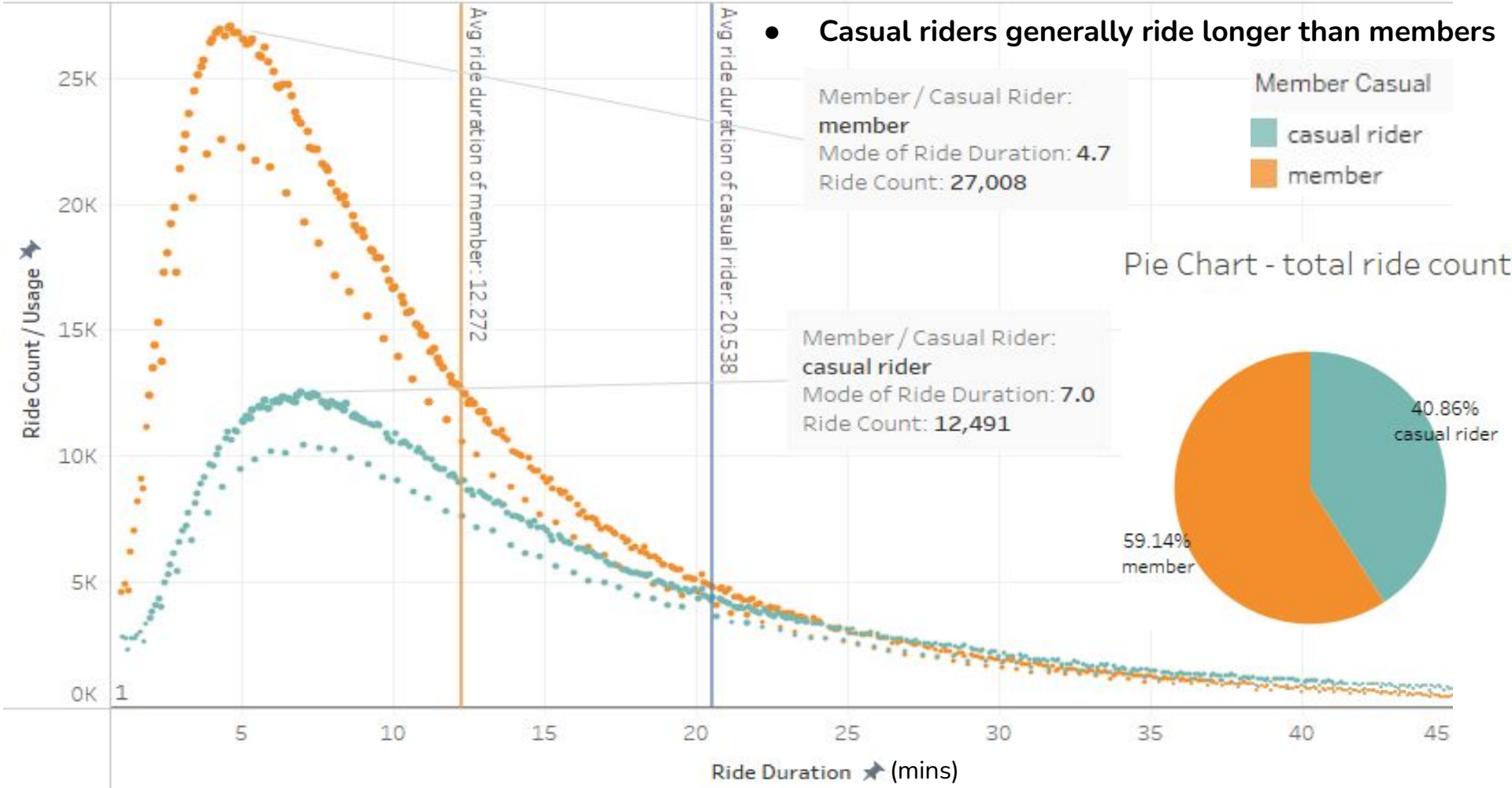

Ride Duration >3hrs
• 4 outliers were also excluded
Outlier
False
True
Member Casual
casual rider
member
Ride Duration (hours)
600
500
400
300
200
100
25
20
15
10
5
Ride Duration Day
Jan 1
Mar 1
May 1
Jul 1
Sep 1
Nov 1
Jan 1
Jan 1 - Dec 31 [2022]

Ride Duration >3hrs

- **Data points are still congested in the bottom range but not entirely confident in omitting all possible outliers without invalidating them**

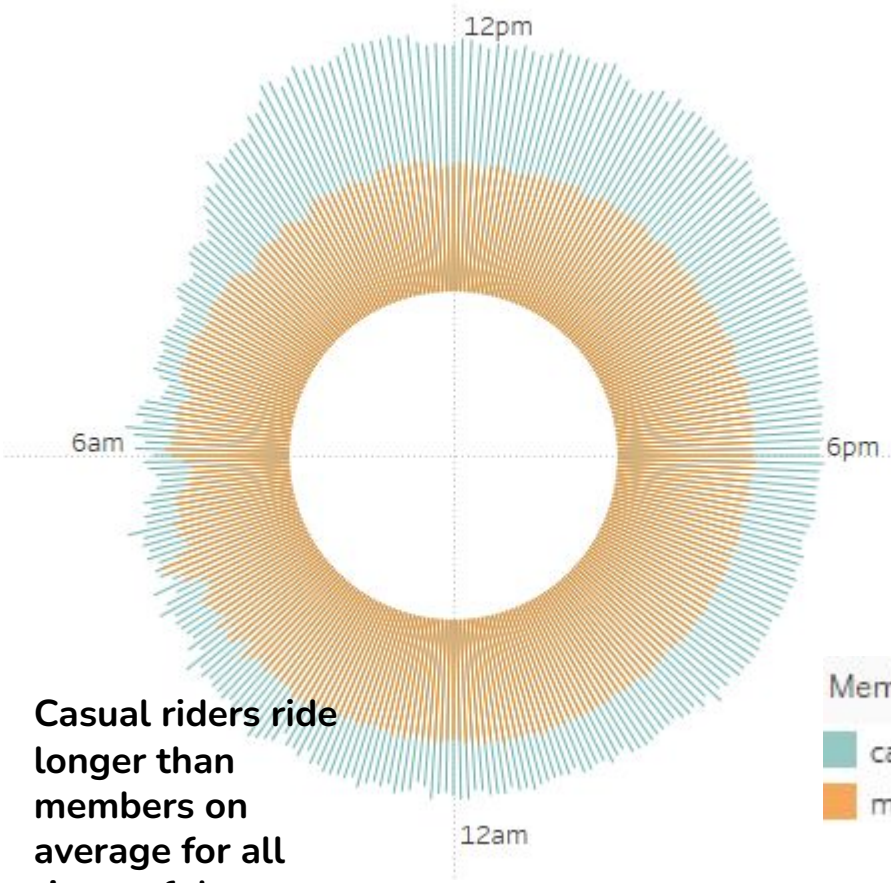


Analysis - Number of Rides vs Ride Duration < 3



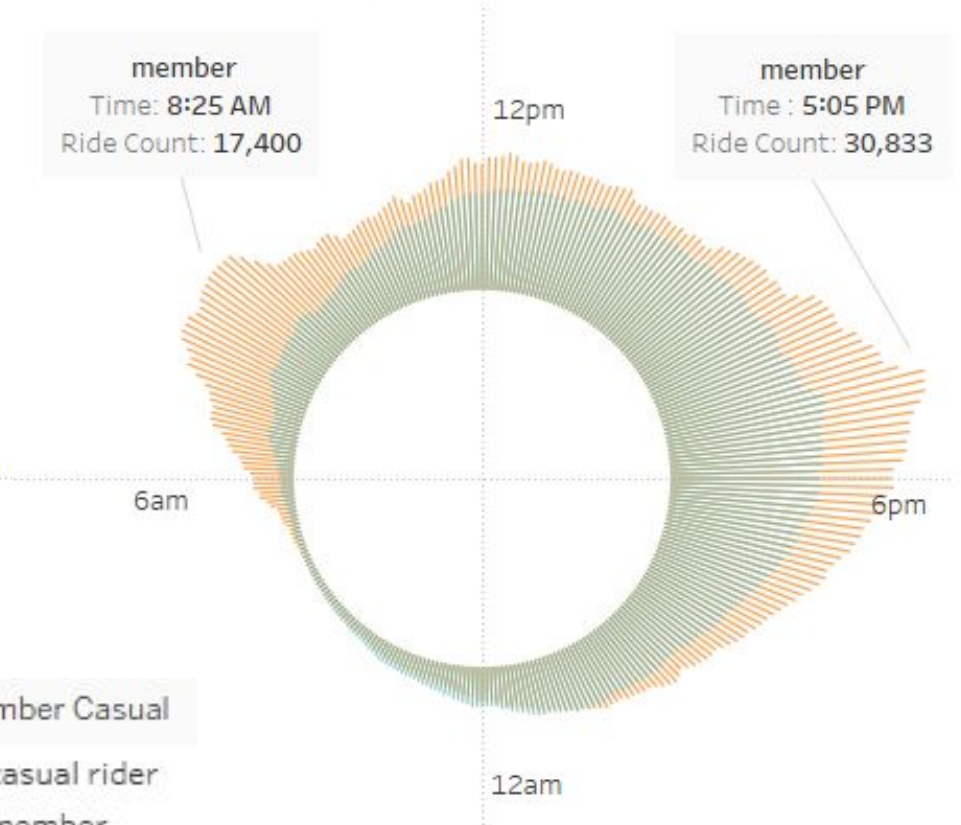
- **Casual riders generally ride longer than members**

Radial Avg Ride Duration by Time

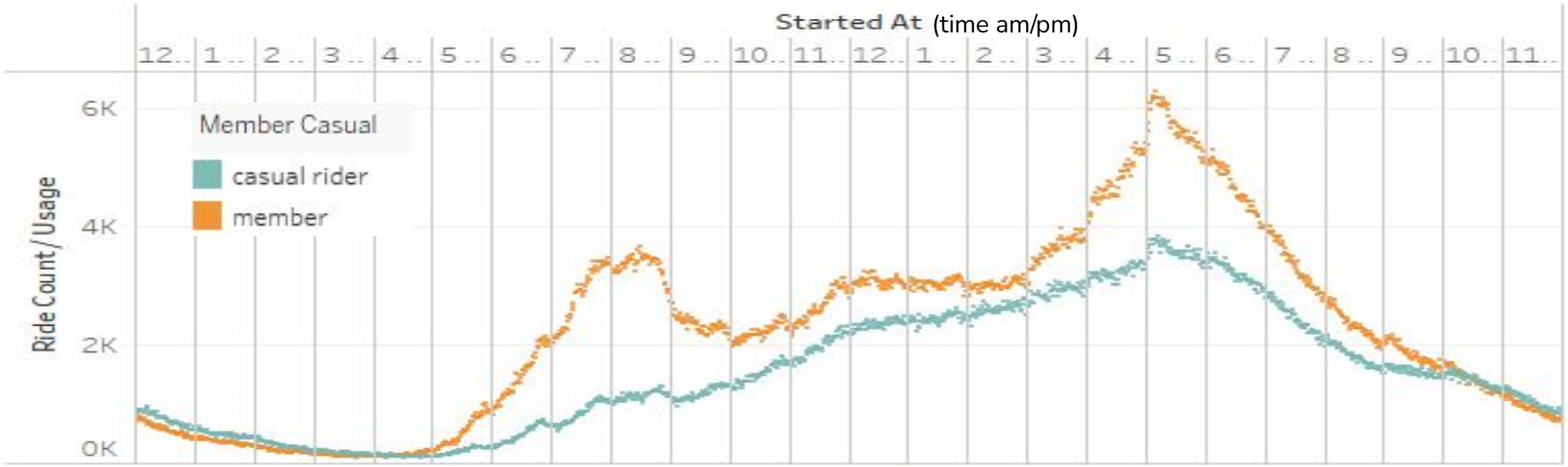


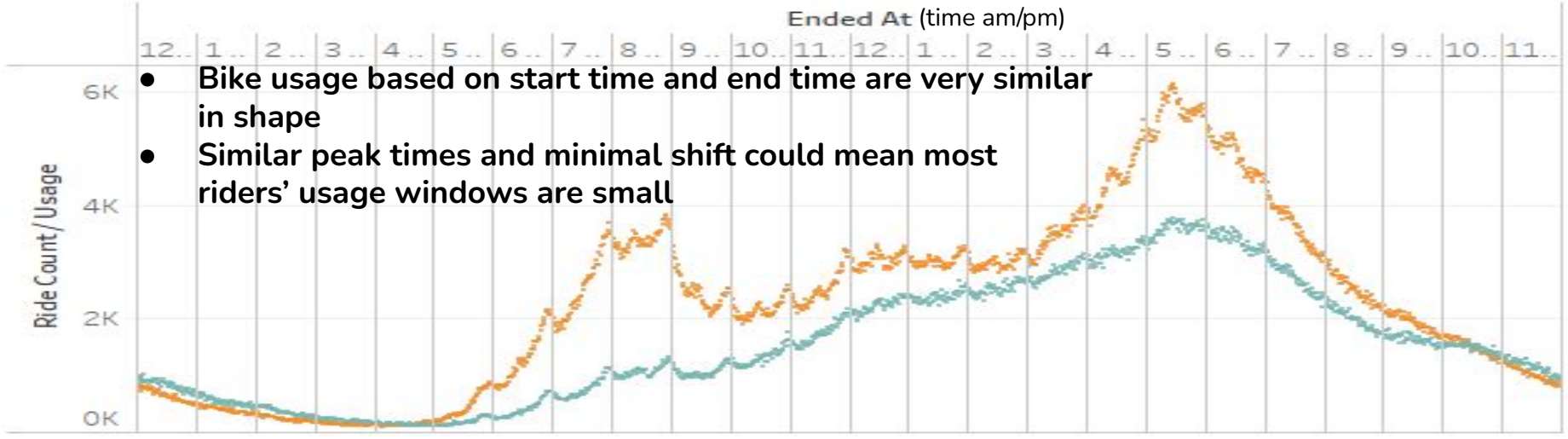
- **Casual riders ride longer than members on average for all times of day**

Radial Ride Count by Time

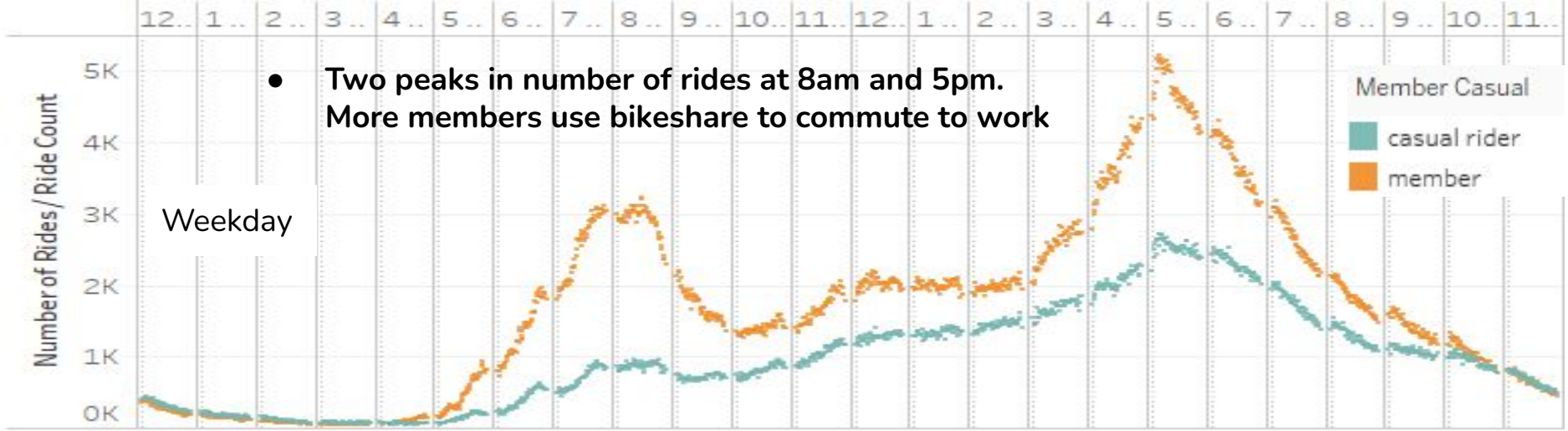


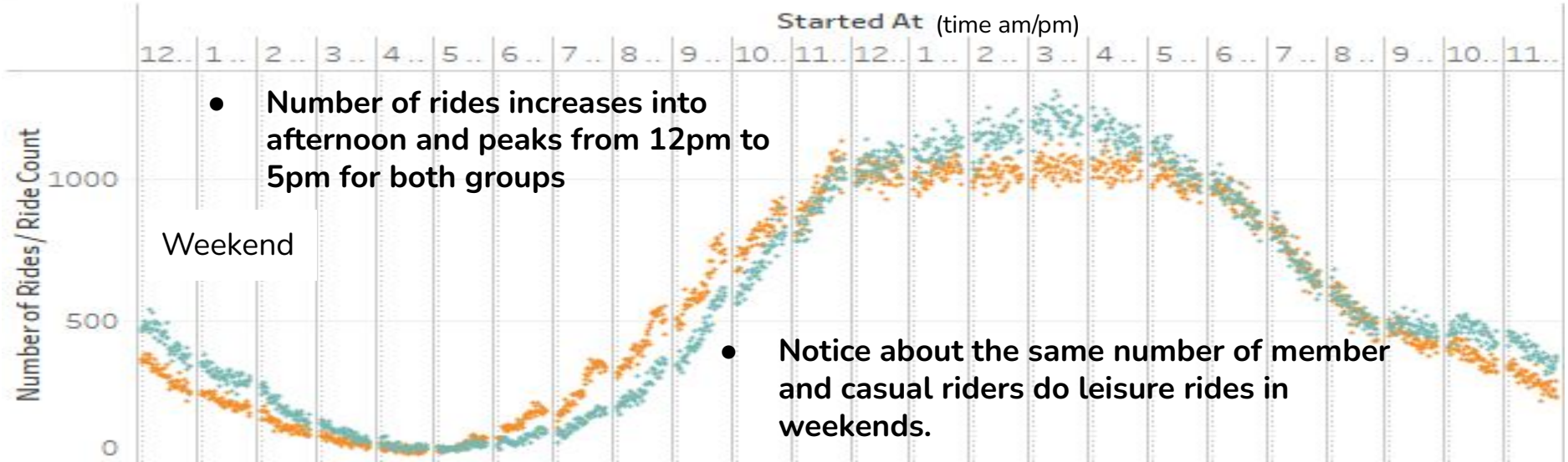
- **Number of rides peaks around 8am and 5pm for members and 5pm only for casual riders**





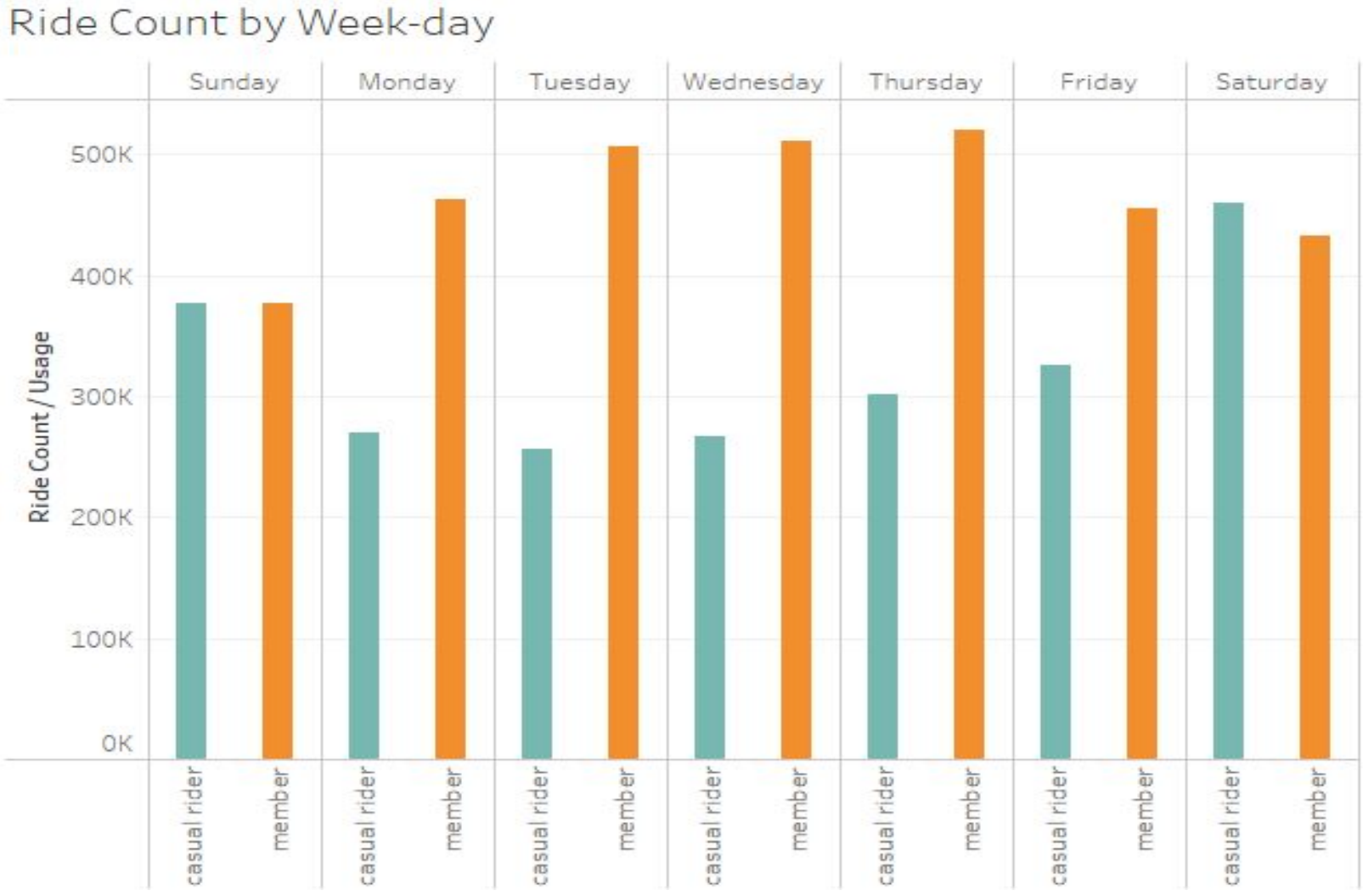
- **Bike usage based on start time and end time are very similar in shape**
- **Similar peak times and minimal shift could mean most riders' usage windows are small**


12.. 1.. 2.. 3.. 4.. 5.. 6.. 7.. 8.. 9.. 10.. 11.. 12.. 1.. 2.. 3.. 4.. 5.. 6.. 7.. 8.. 9.. 10.. 11..
Number of Rides / Ride Count
5K
4K
3K
2K
1K
0K
Member Casual
casual rider
member
Weekday
• Two peaks in number of rides at 8am and 5pm.
More members use bikeshare to commute to work


Started At (time am/pm)
12.. 1.. 2.. 3.. 4.. 5.. 6.. 7.. 8.. 9.. 10.. 11.. 12.. 1.. 2.. 3.. 4.. 5.. 6.. 7.. 8.. 9.. 10.. 11..
Number of Rides / Ride Count
1000
500
0
Weekend
• Number of rides increases into afternoon and peaks from 12pm to 5pm for both groups
• Notice about the same number of member and casual riders do leisure rides in weekends.

Avg Ride Duration by Week-day

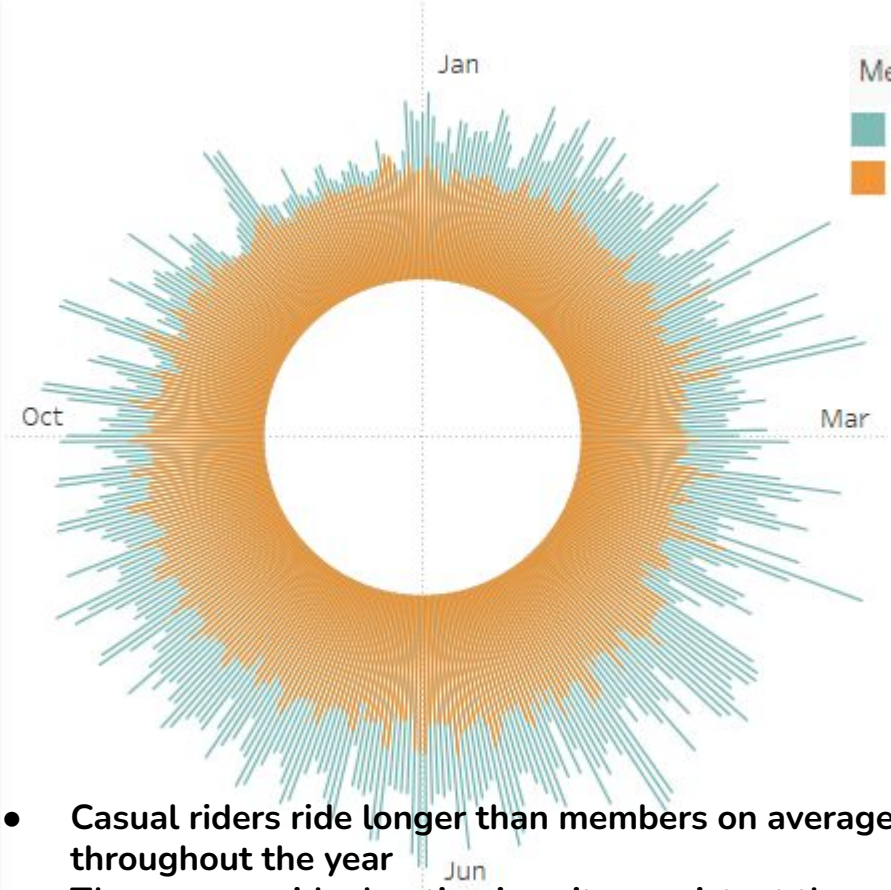
- **Both casual riders and members ride longer in weekend**
- **On average, casual riders ride 3-4 mins longer in weekend whereas members ride 1-2 mins longer**



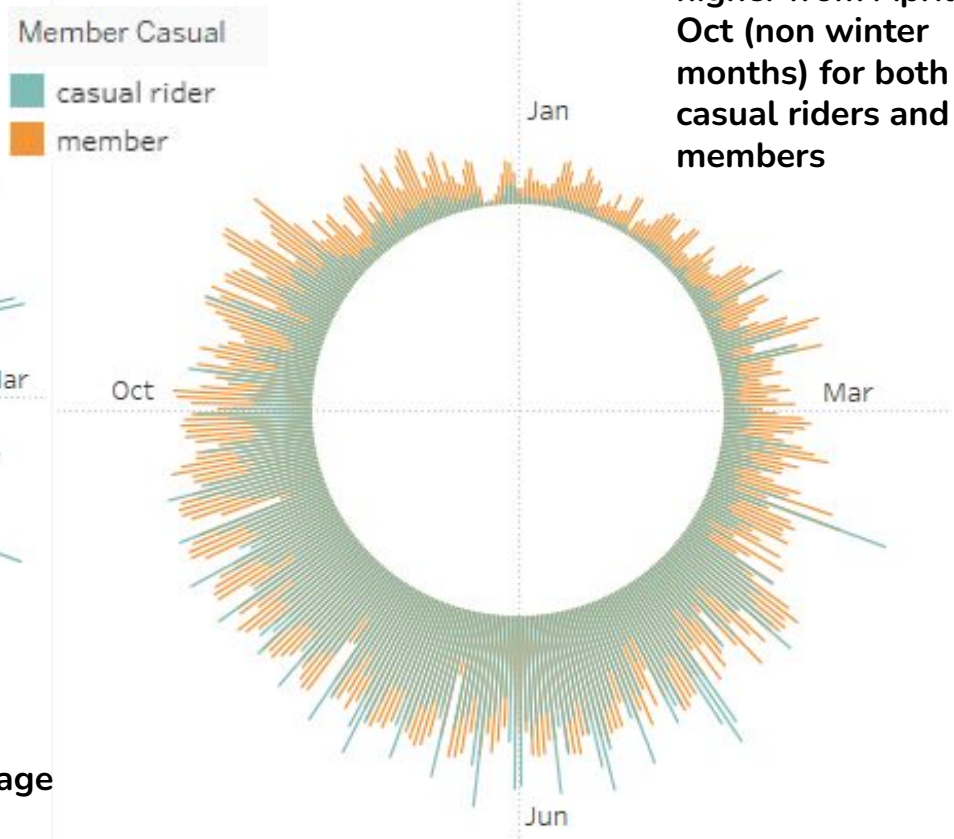
-Bike usage (number of rides) is higher in weekday for members but lower for casual riders.

Marketing suggestions:
-reach out to casual riders in weekends
-offer weekend discount in membership

Radial Daily Avg Ride Duration



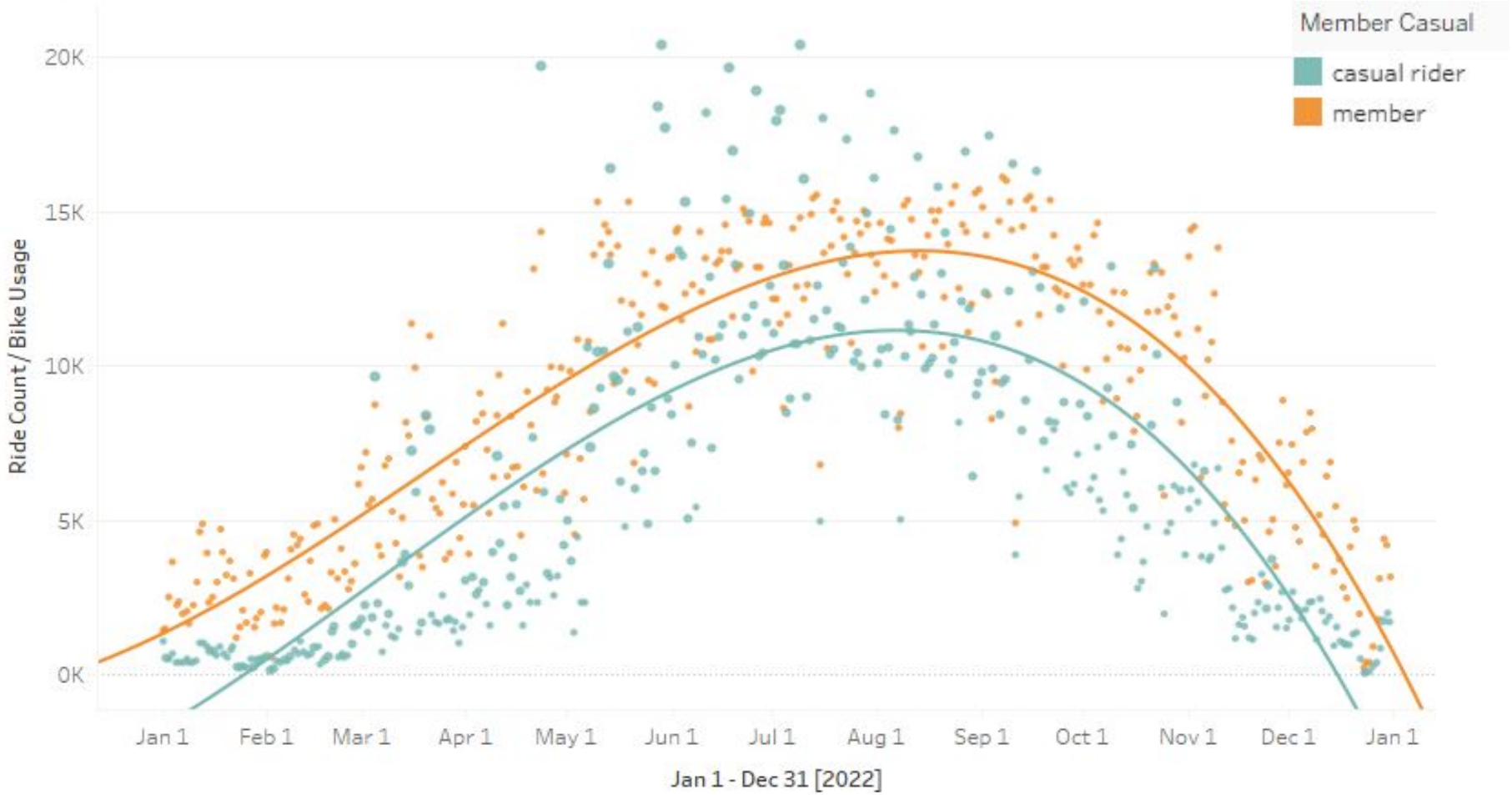
Radial Daily Ride Count



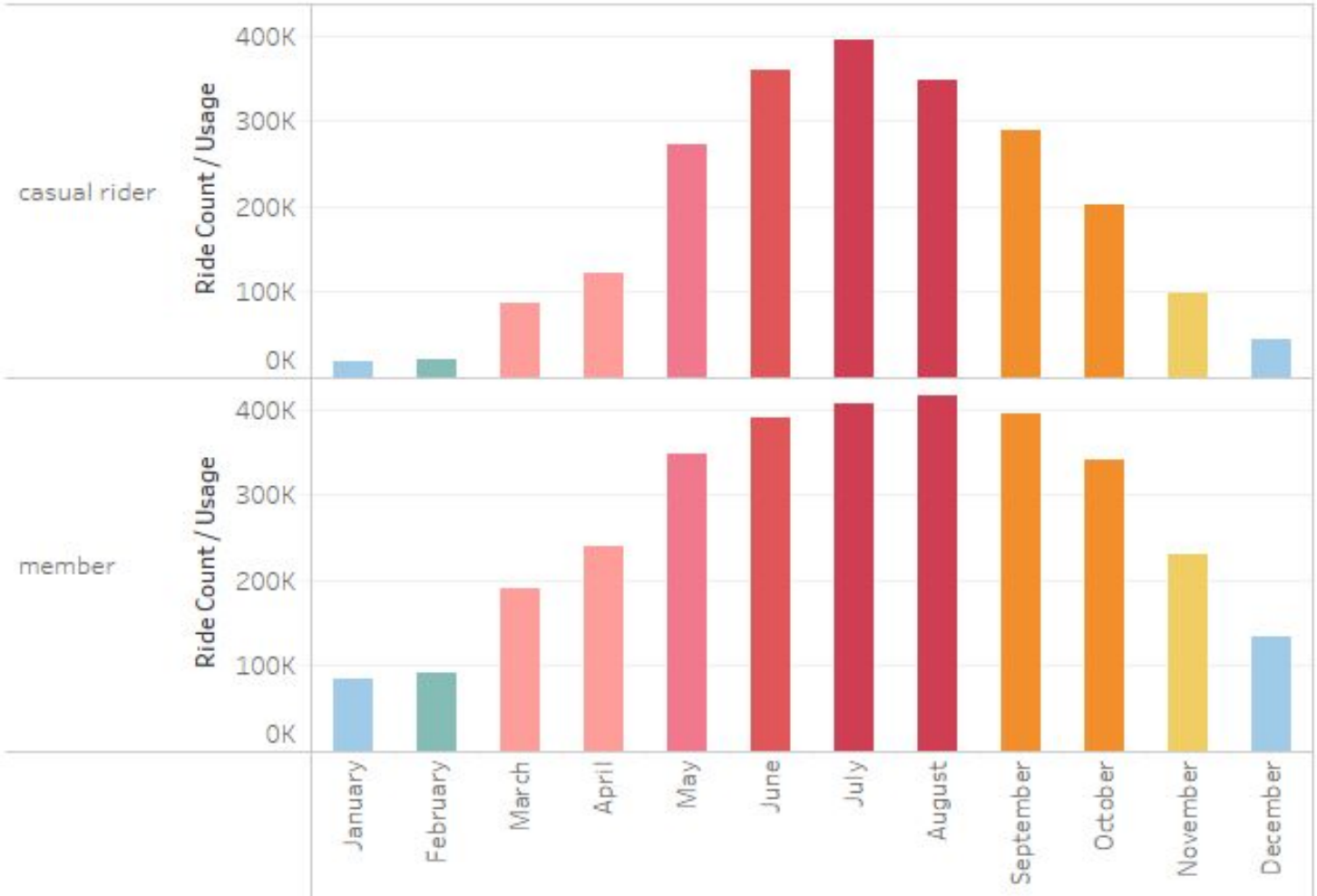
- **Number of rides is higher from April to Oct (non winter months) for both casual riders and members**

- **Casual riders ride longer than members on average throughout the year**
- **The average ride duration is quite consistent throughout the year for both members and casual riders**

Daily Ride Count



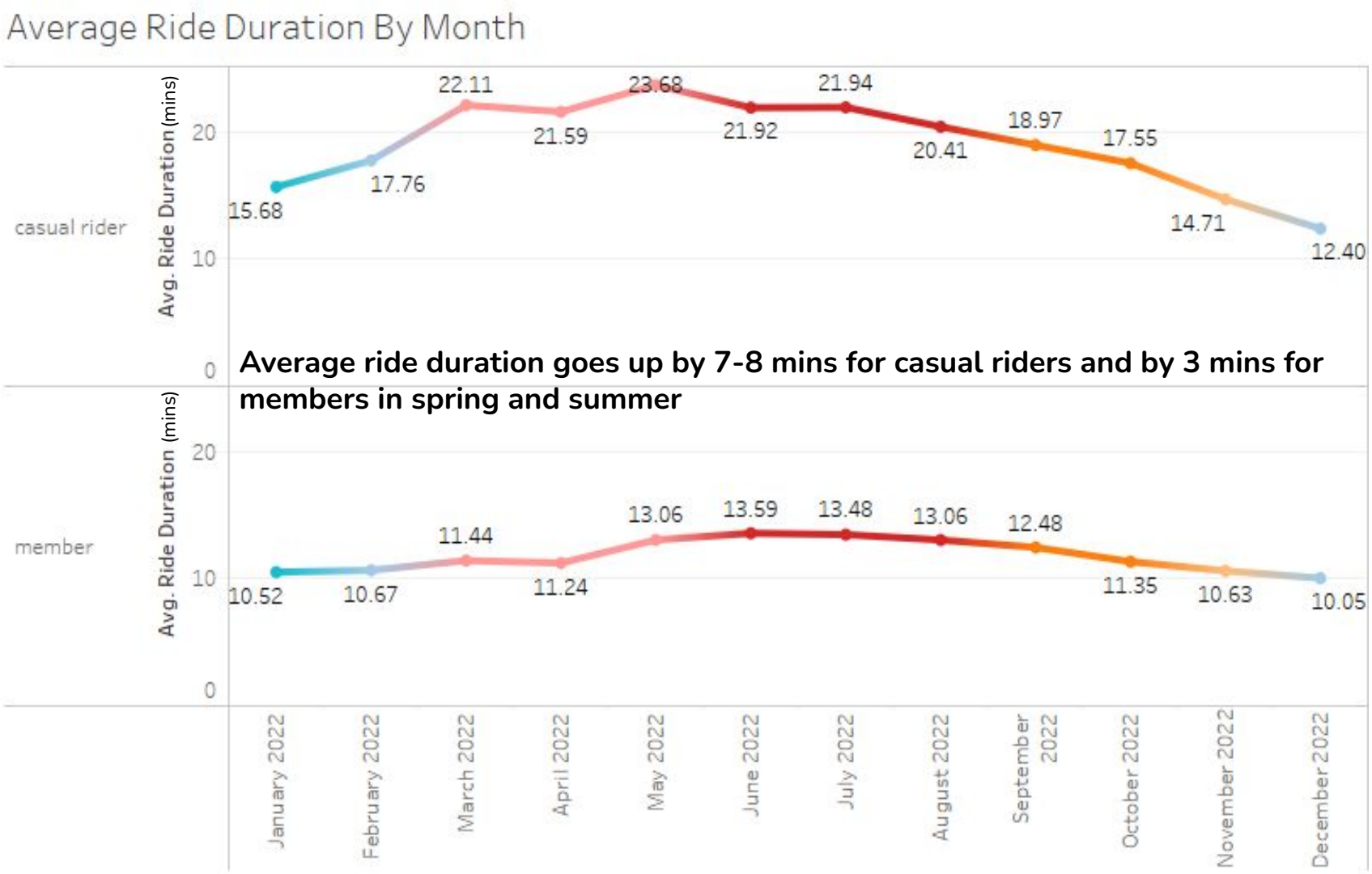
Monthly Ride Count



- **Bike usage peaks in summer months and slow down in winter months for both members and casual riders.**

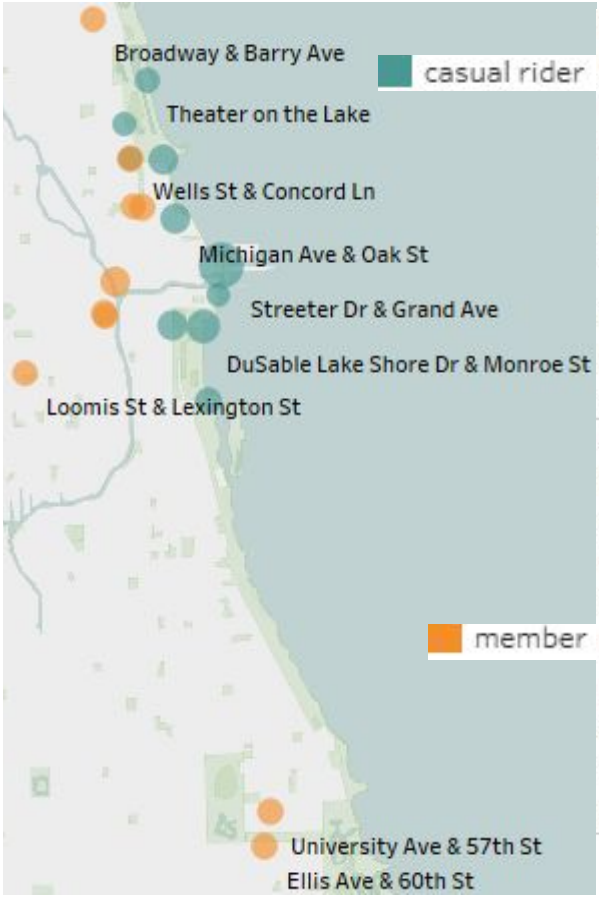
Marketing suggestions:

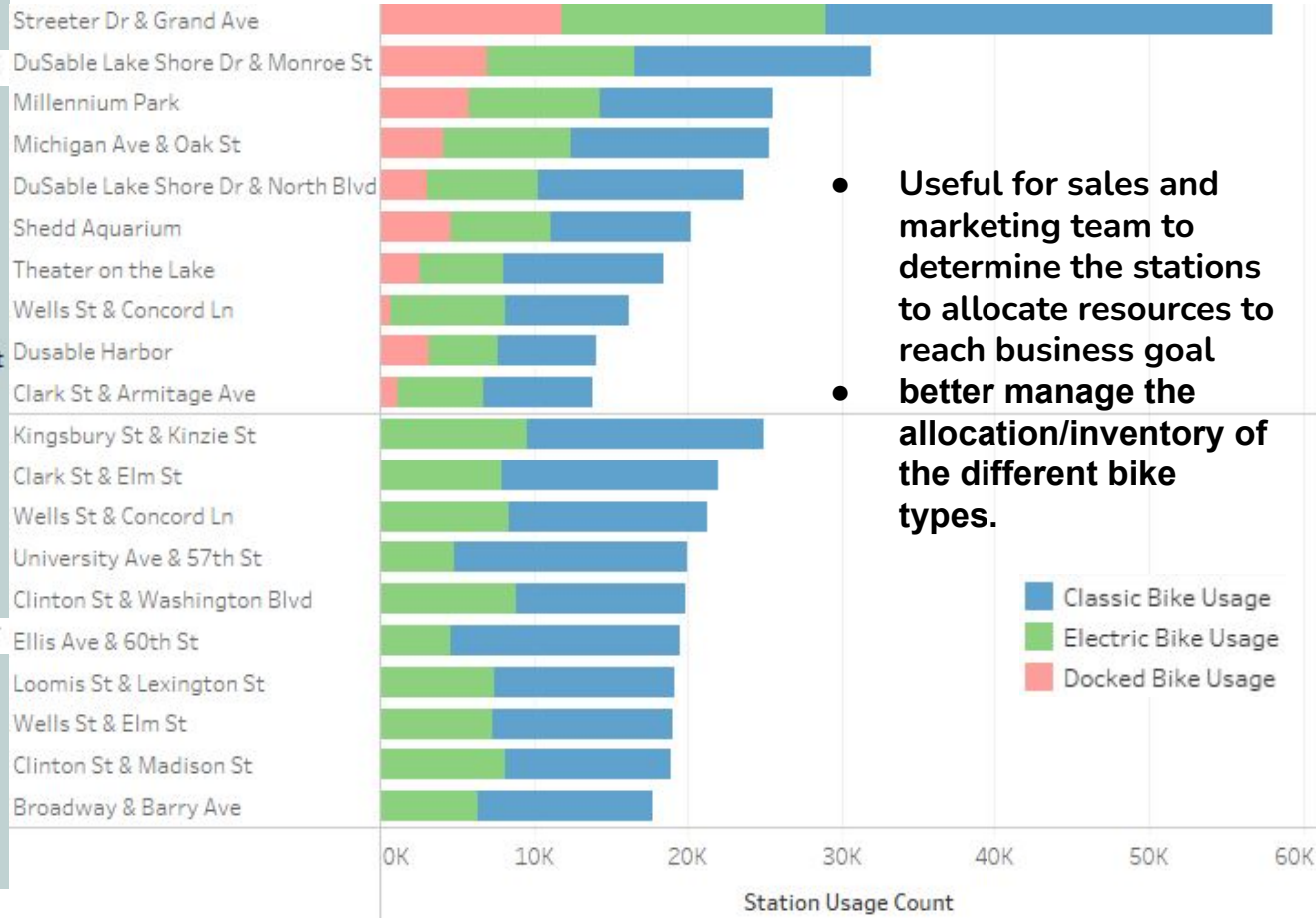
- **Reach larger customer pool in summer months**
- **Seasonal discount in slower months**



Average ride duration goes up by 7-8 mins for casual riders and by 3 mins for members in spring and summer

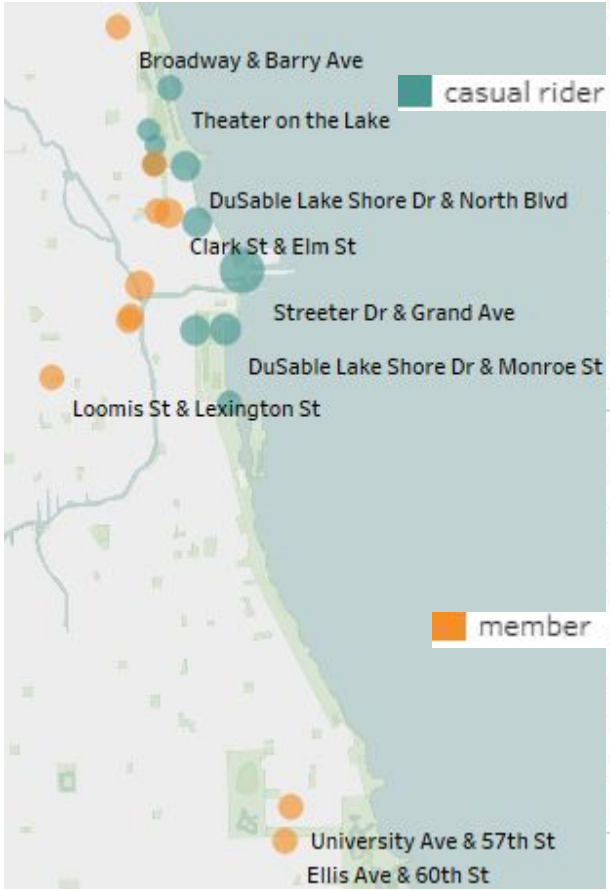
Top 10 Bike Share Start Stations

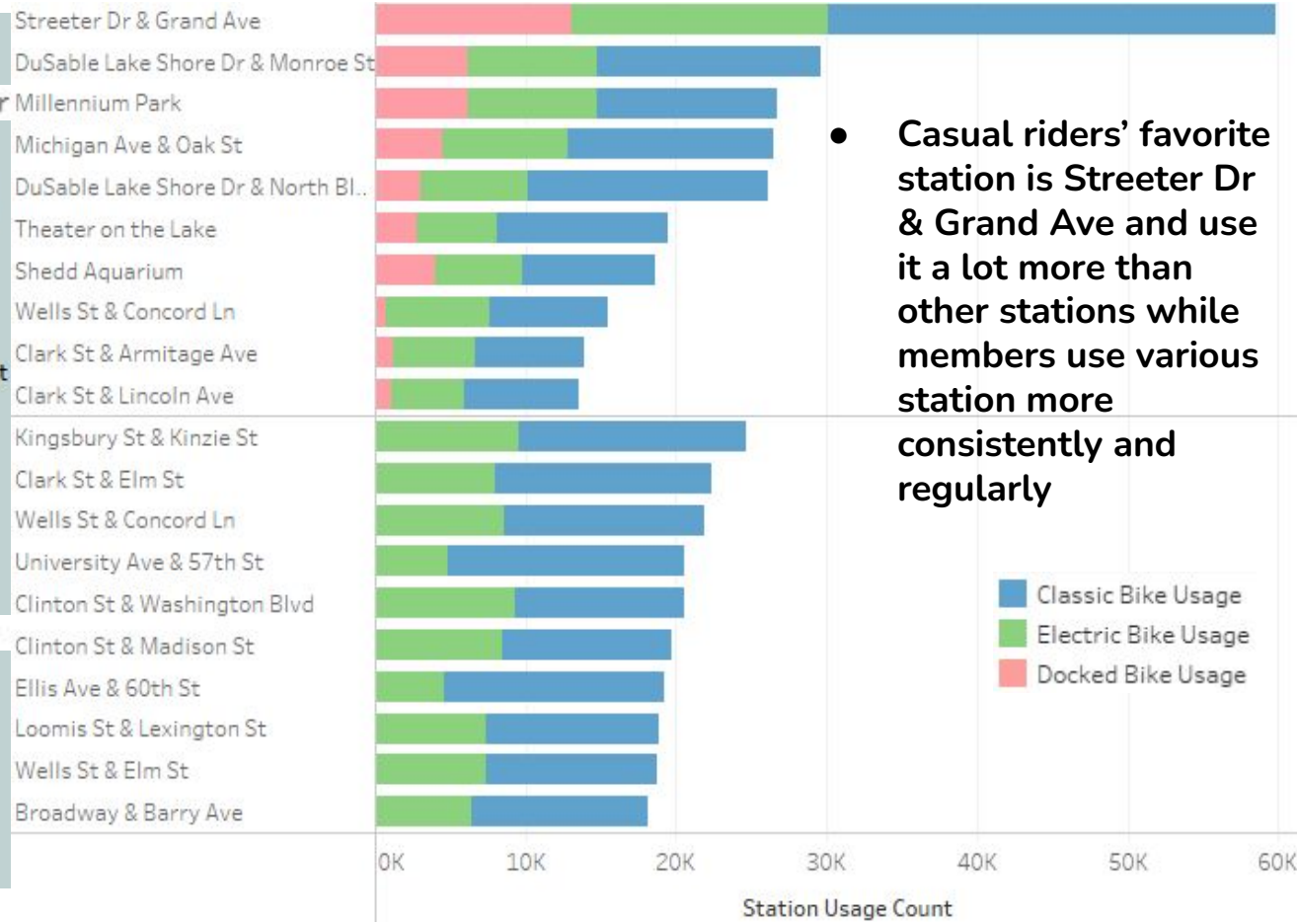




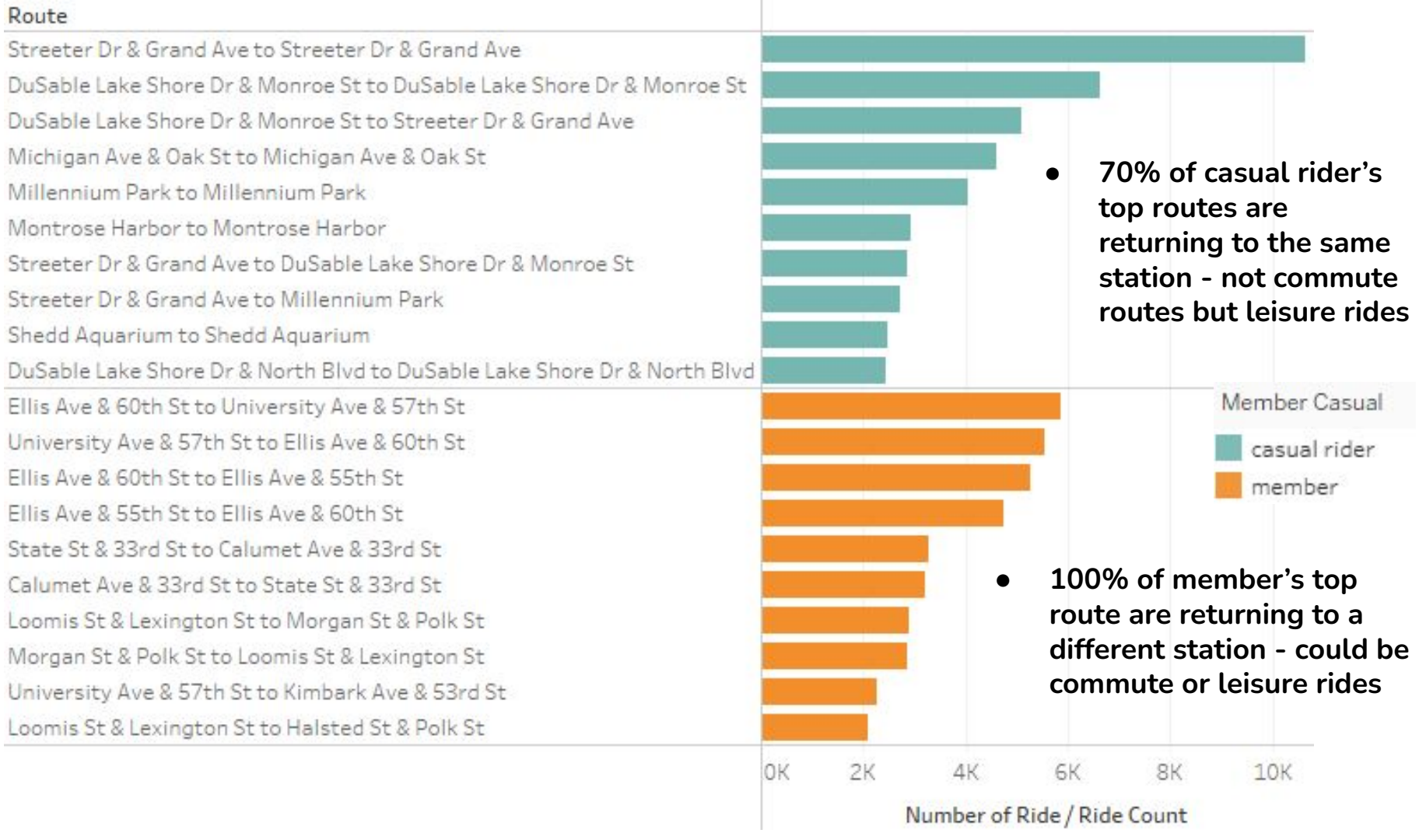
- **Useful for sales and marketing team to determine the stations to allocate resources to reach business goal**
- **better manage the allocation/inventory of the different bike types.**

Top 10 Bike Share Return Stations





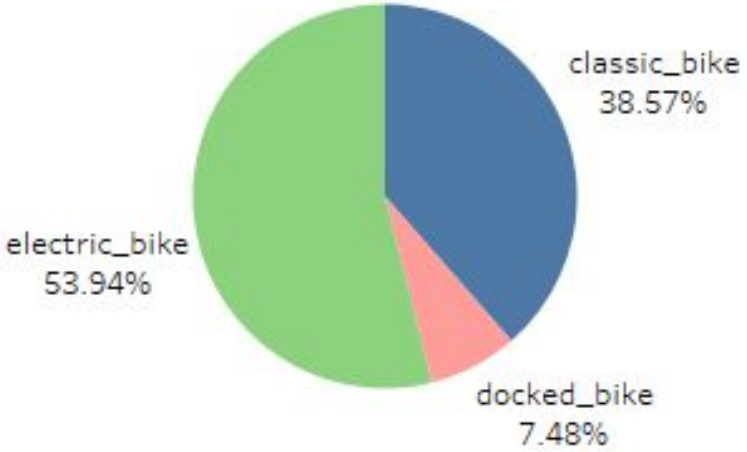
- **Casual riders' favorite station is Streeter Dr & Grand Ave and use it a lot more than other stations while members use various station more consistently and regularly**


Route
Streeter Dr & Grand Ave to Streeter Dr & Grand Ave
DuSable Lake Shore Dr & Monroe St to DuSable Lake Shore Dr & Monroe St
DuSable Lake Shore Dr & Monroe St to Streeter Dr & Grand Ave
Michigan Ave & Oak St to Michigan Ave & Oak St
Millennium Park to Millennium Park
Montrose Harbor to Montrose Harbor
Streeter Dr & Grand Ave to DuSable Lake Shore Dr & Monroe St
Streeter Dr & Grand Ave to Millennium Park
Shedd Aquarium to Shedd Aquarium
DuSable Lake Shore Dr & North Blvd to DuSable Lake Shore Dr & North Blvd
Ellis Ave & 60th St to University Ave & 57th St
University Ave & 57th St to Ellis Ave & 60th St
Ellis Ave & 60th St to Ellis Ave & 55th St
Ellis Ave & 55th St to Ellis Ave & 60th St
State St & 33rd St to Calumet Ave & 33rd St
Calumet Ave & 33rd St to State St & 33rd St
Loomis St & Lexington St to Morgan St & Polk St
Morgan St & Polk St to Loomis St & Lexington St
University Ave & 57th St to Kimbark Ave & 53rd St
Loomis St & Lexington St to Halsted St & Polk St
0K
2K
4K
6K
8K
10K
Number of Ride / Ride Count
Member Casual
casual rider
member
• 70% of casual rider's top routes are returning to the same station - not commute routes but leisure rides
• 100% of member's top route are returning to a different station - could be commute or leisure rides

Bike Type Usage <3 hr

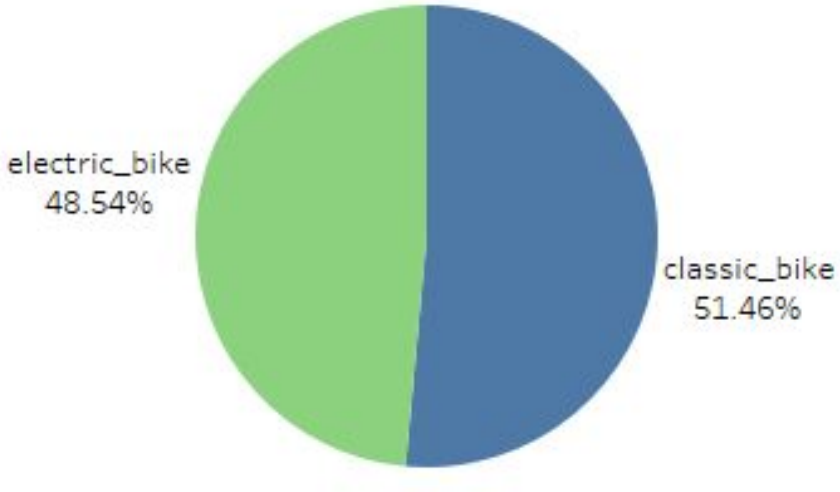
casual rider

- **Use electric bike slightly more than classic bikes but also use docked bike at a small percentage.**

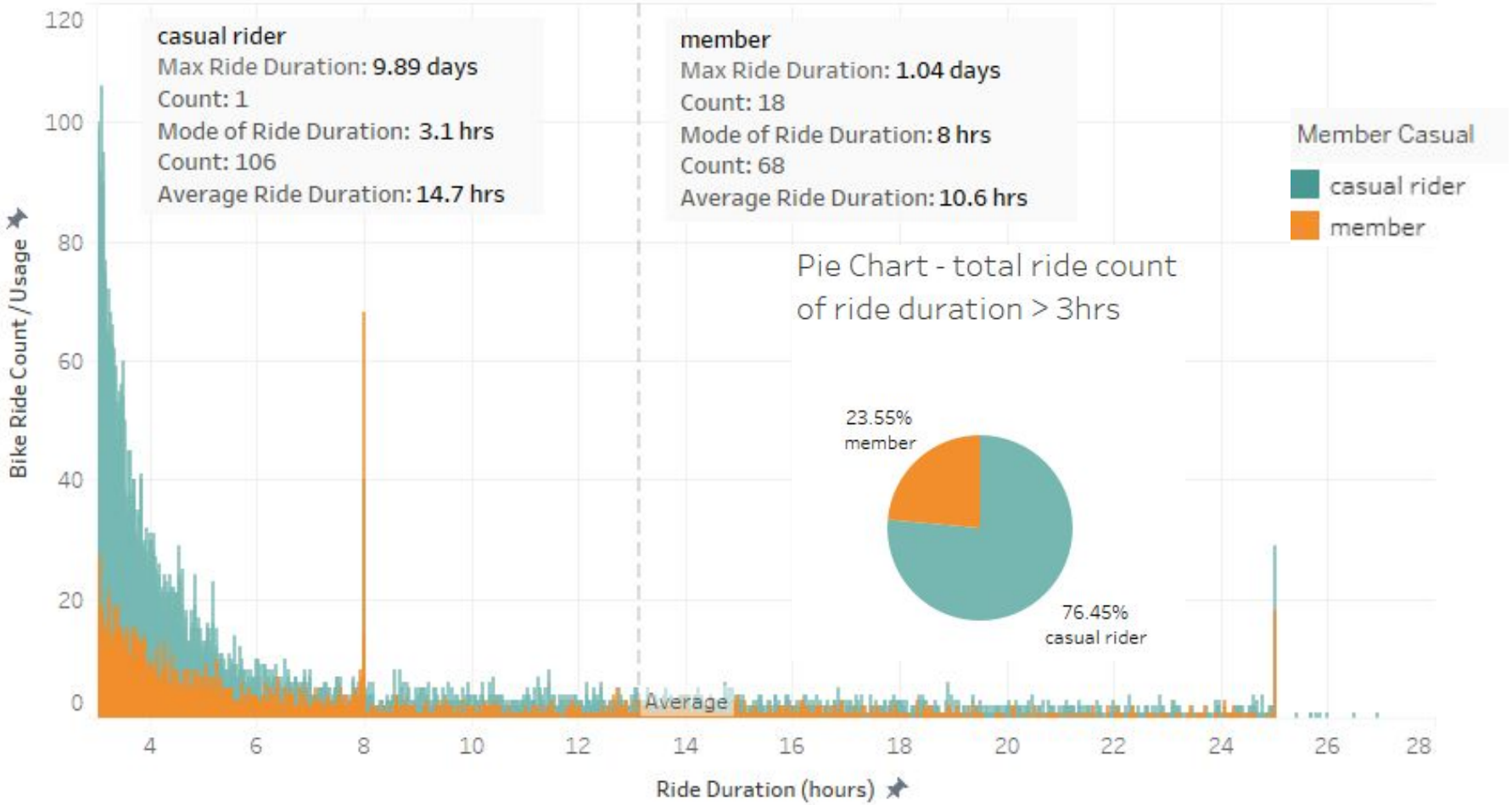


member

- **Use electric bikes as often as classic bikes**



Ride Duration vs Count

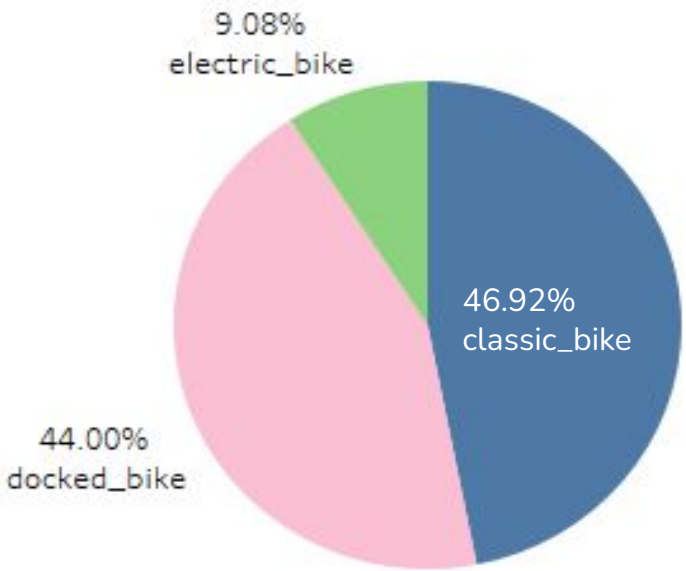


Bike Type Usage >3hr

Rideable Type
- classic_bike
- docked__bike
- electric_bike

casual rider

- **Large percentage of casual riders use docked bike for long rides (ride length > 3 hrs)**

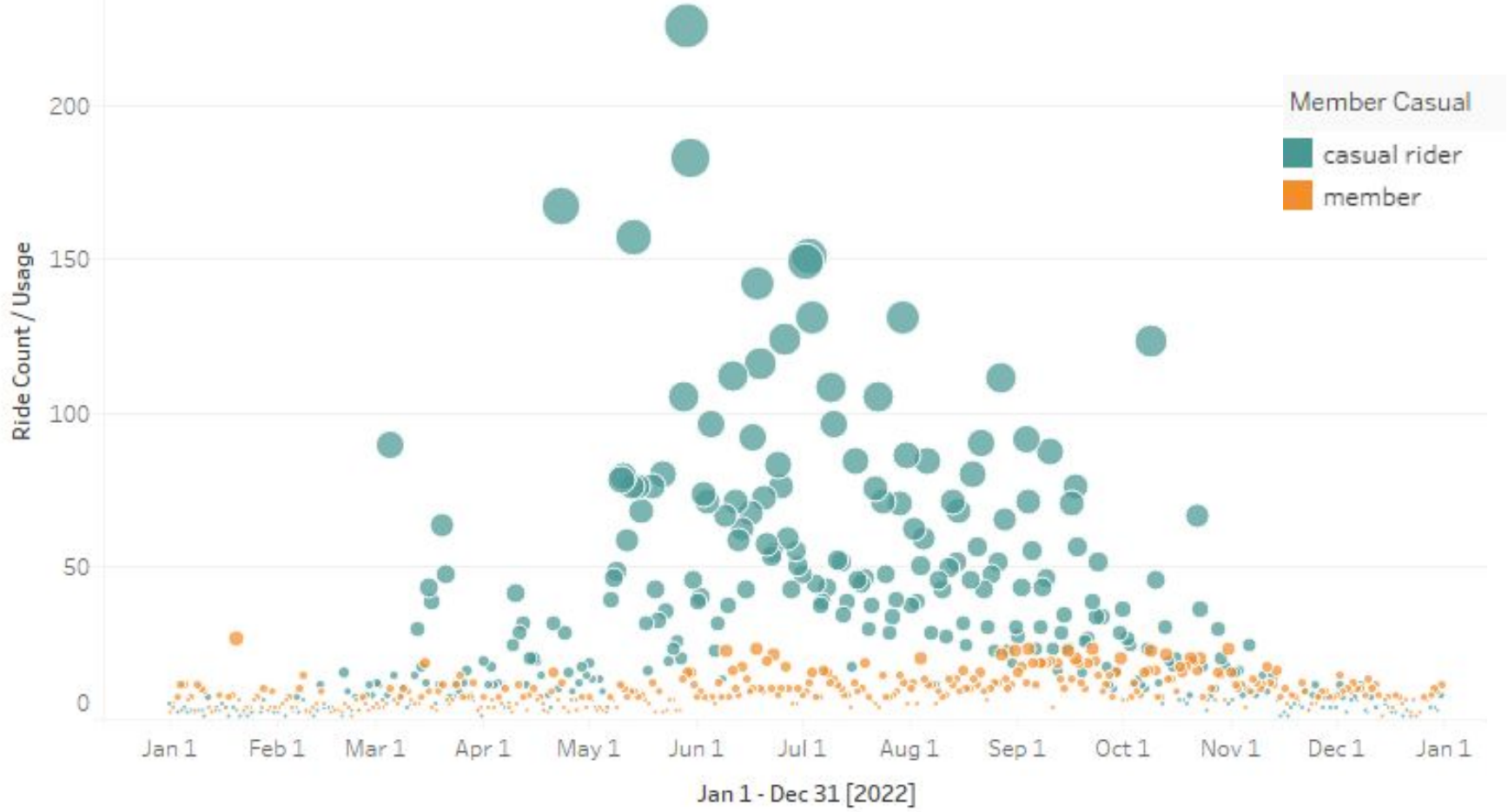


member

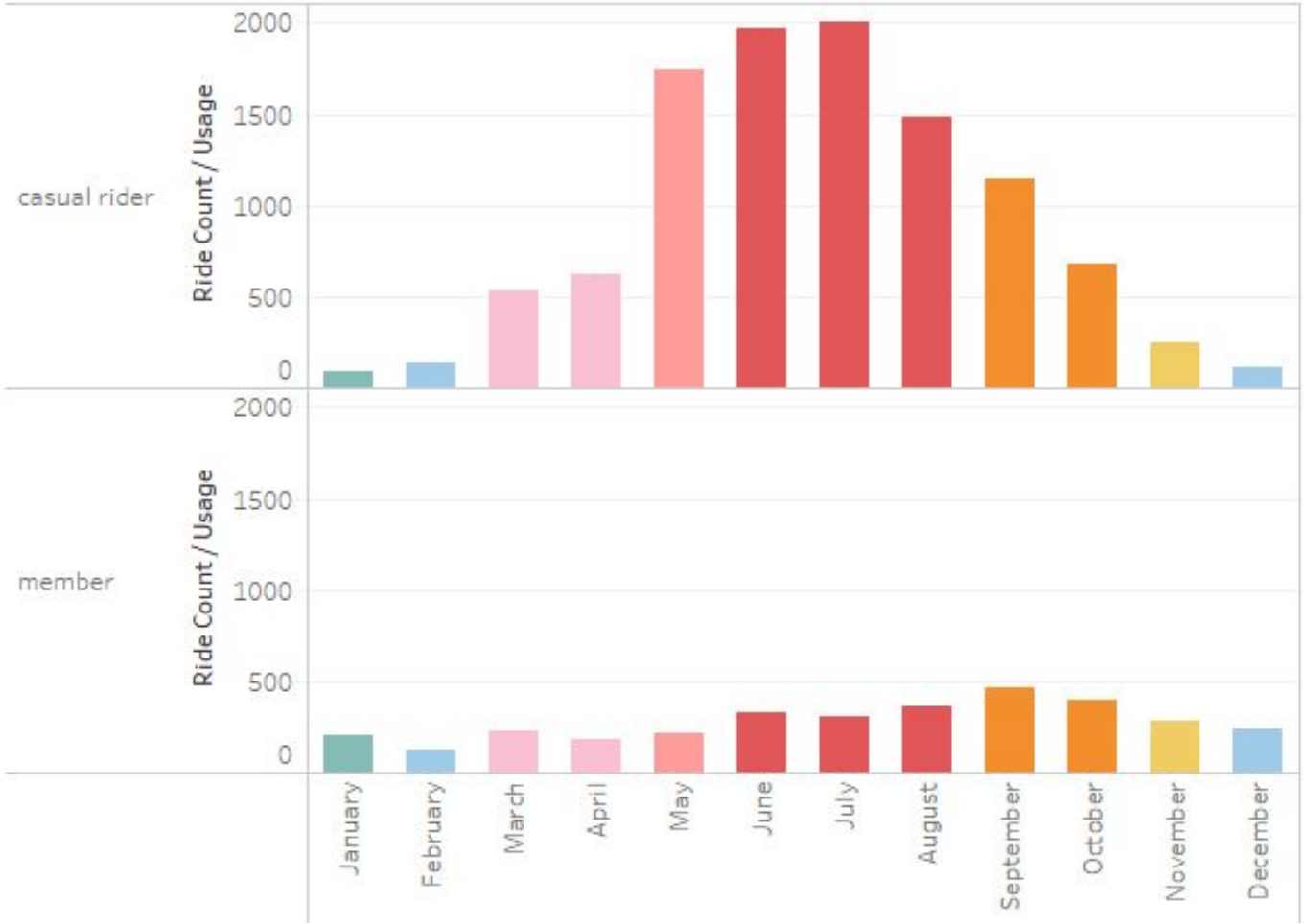
- **Docked bike probably is not an option offered in membership**
- **Adding the option of docked bike in membership may appeal to casual riders**

Ride Count >3hrs

- **Daily ride count scatter plot for ride duration > 3 hours**
- **Number of long rides increases in summer months for casual riders while it stay fairly low throughout for members**

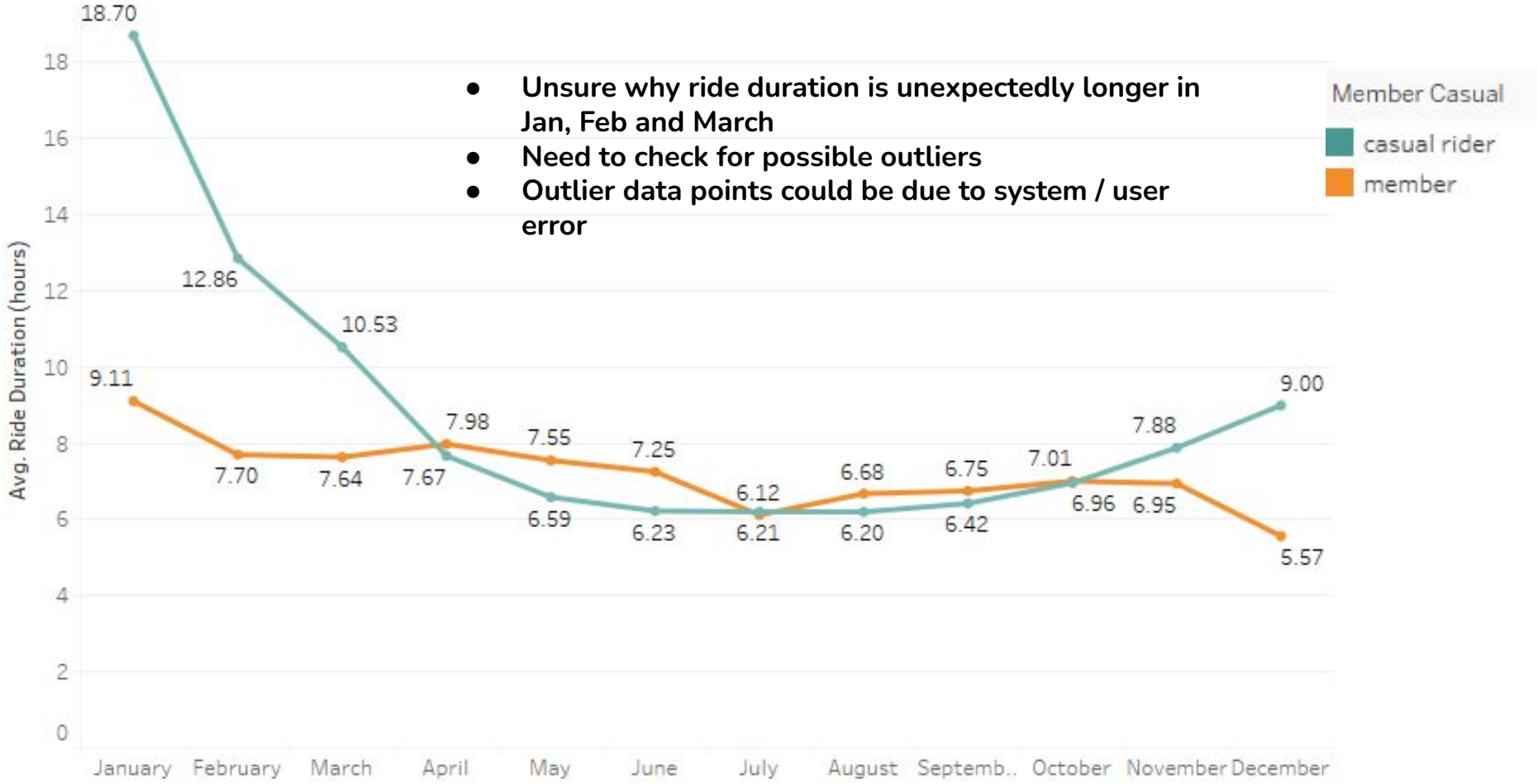


Ride Count by Month >3 hrs



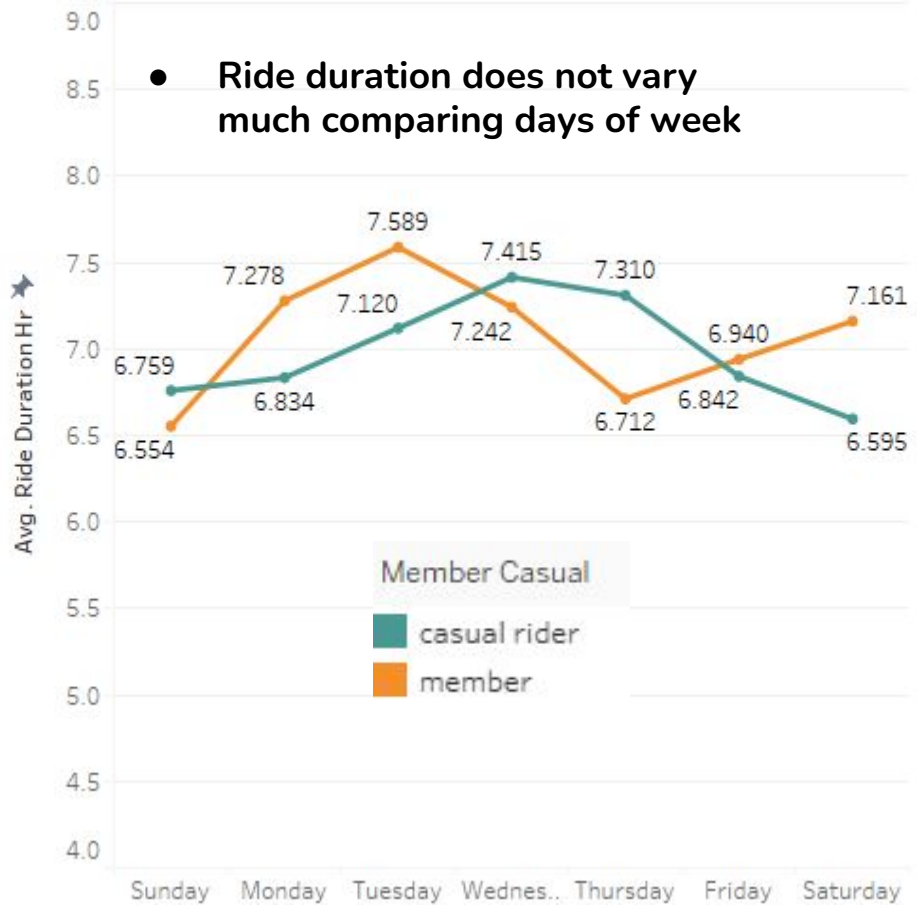
- Monthly ride count bar graph for ride duration > 3 hrs
- Casual riders do significantly more long rides than members
- Casual riders do most long rides in summer months
- Members do slightly more long rides in summer and fall months

Avg Ride Duration by Month >3hrs



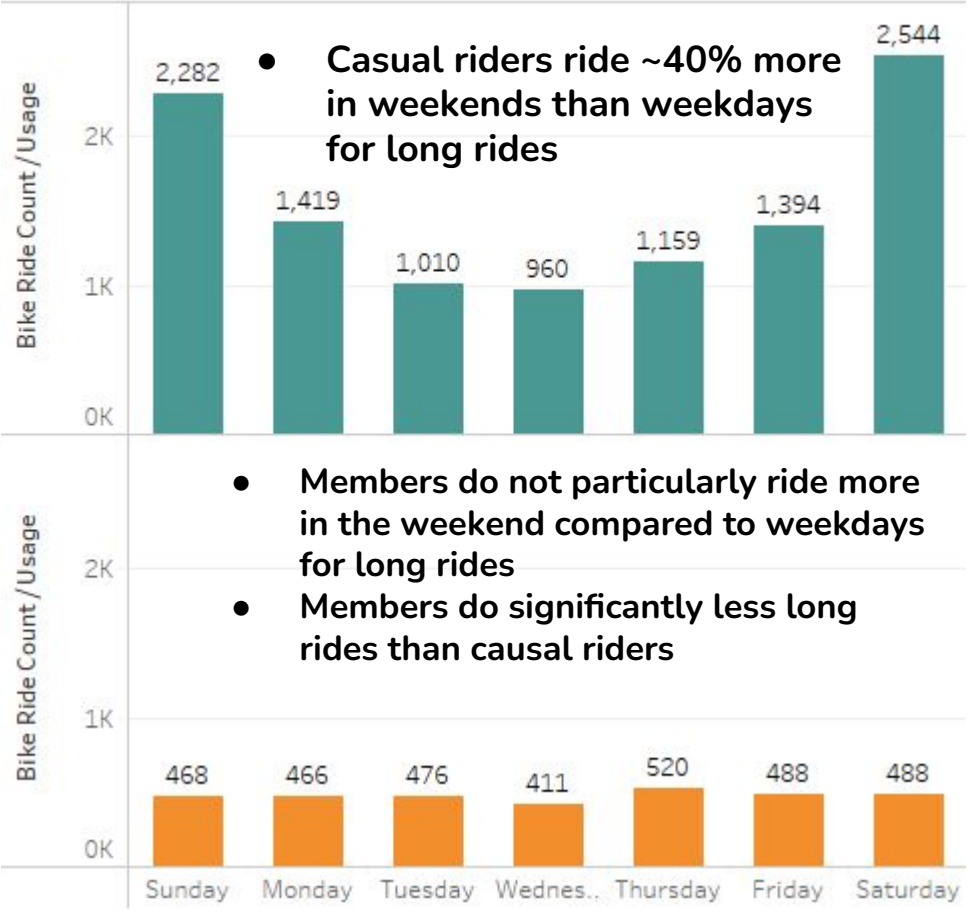
- **Unsure why ride duration is unexpectedly longer in Jan, Feb and March**
- **Need to check for possible outliers**
- **Outlier data points could be due to system / user error**

Ride Duration by Week-day >3 hrs



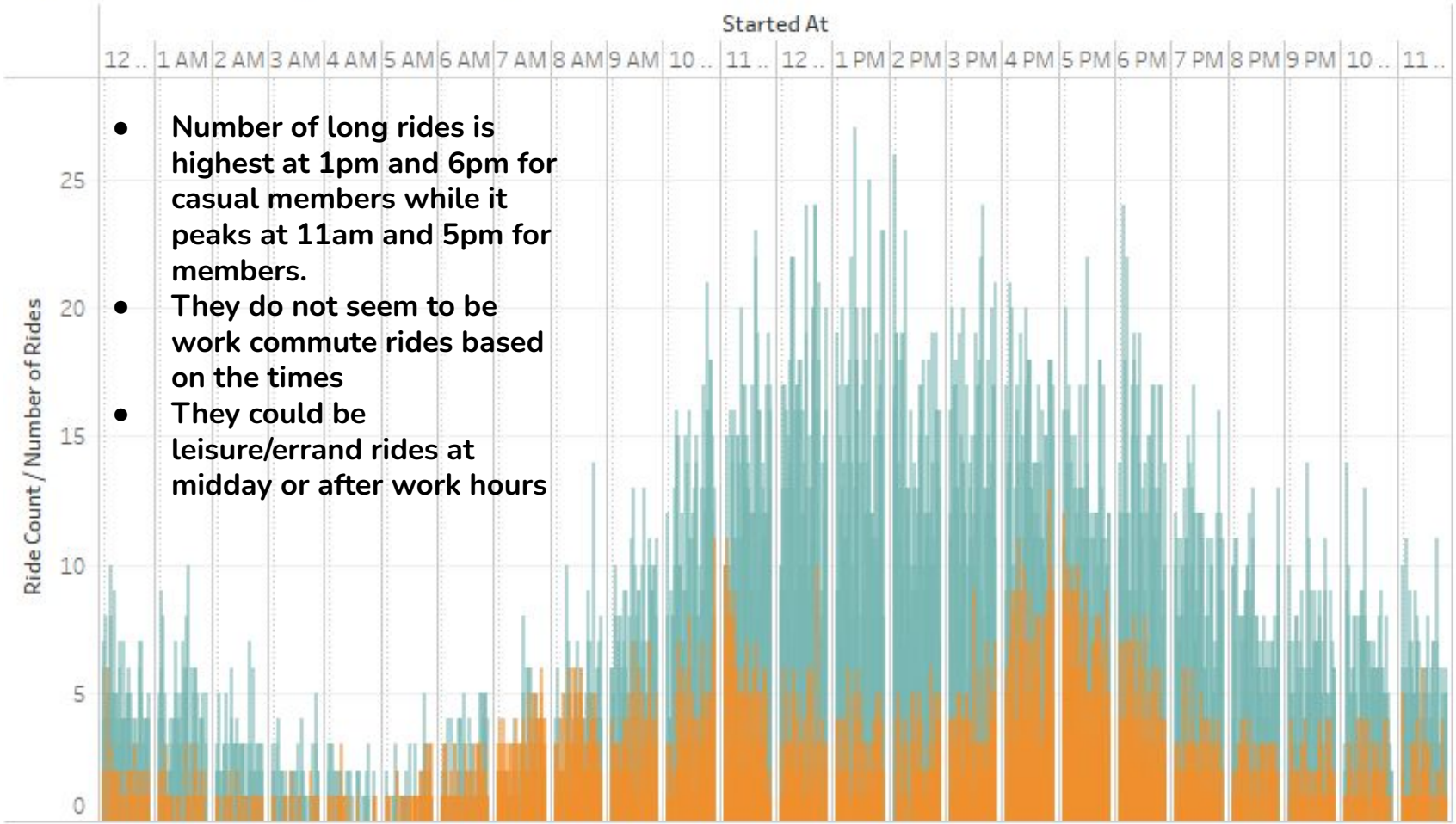
- **Ride duration does not vary much comparing days of week**

Ride Count by Week-day >3 hrs



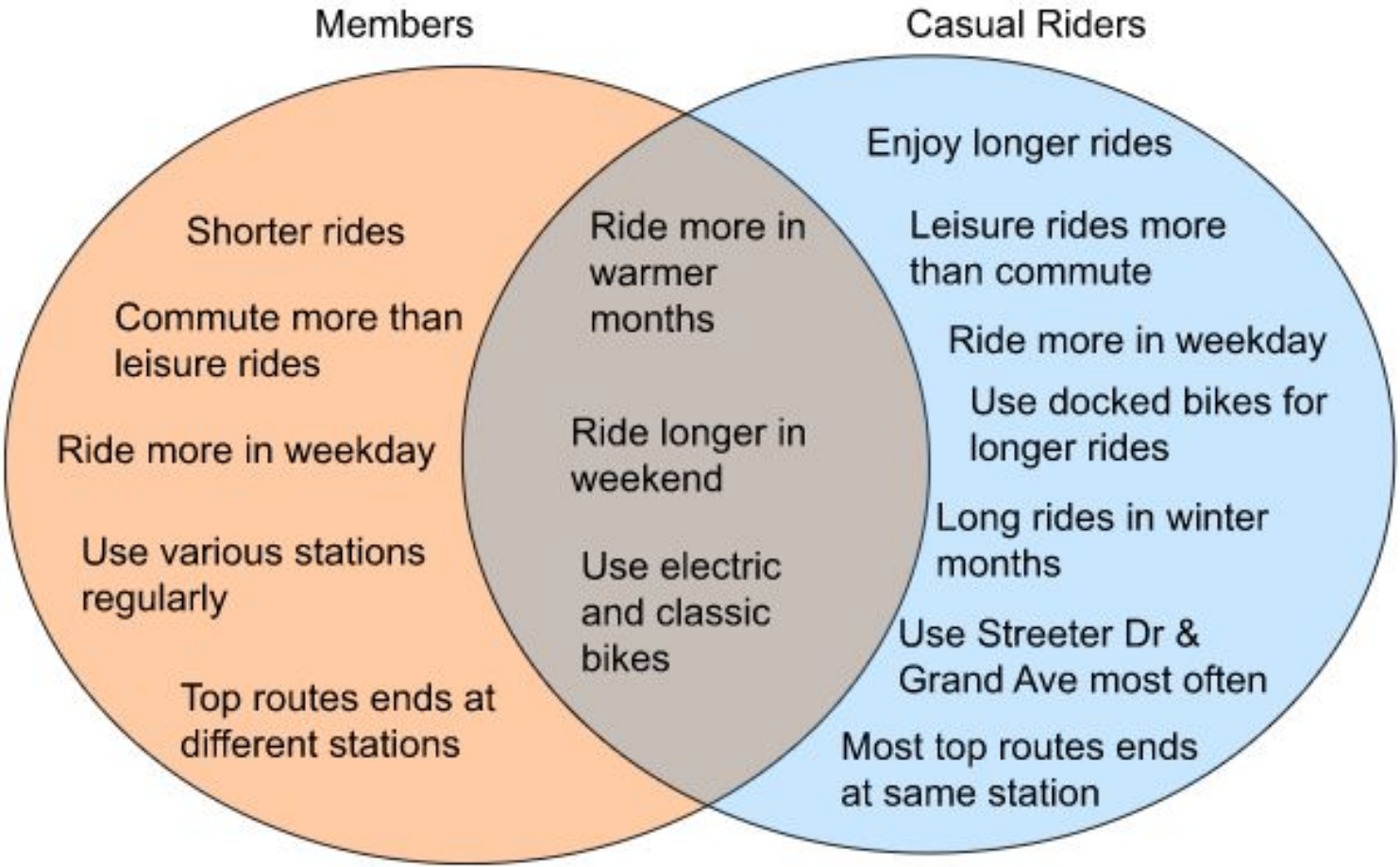
- **Casual riders ride ~40% more in weekends than weekdays for long rides**
- **Members do not particularly ride more in the weekend compared to weekdays for long rides**
- **Members do significantly less long rides than causal riders**

Ride Count by Start Time > 3hrs



- **Number of long rides is highest at 1pm and 6pm for casual members while it peaks at 11am and 5pm for members.**
- **They do not seem to be work commute rides based on the times**
- **They could be leisure/errand rides at midday or after work hours**

Summary


Members
Casual Riders
Shorter rides
Commute more than leisure rides
Ride more in weekday
Use various stations regularly
Top routes ends at different stations
Ride more in warmer months
Ride longer in weekend
Use electric and classic bikes
Enjoy longer rides
Leisure rides more than commute
Ride more in weekday
Use docked bikes for longer rides
Long rides in winter months
Use Streeter Dr & Grand Ave most often
Most top routes ends at same station

Recommendations

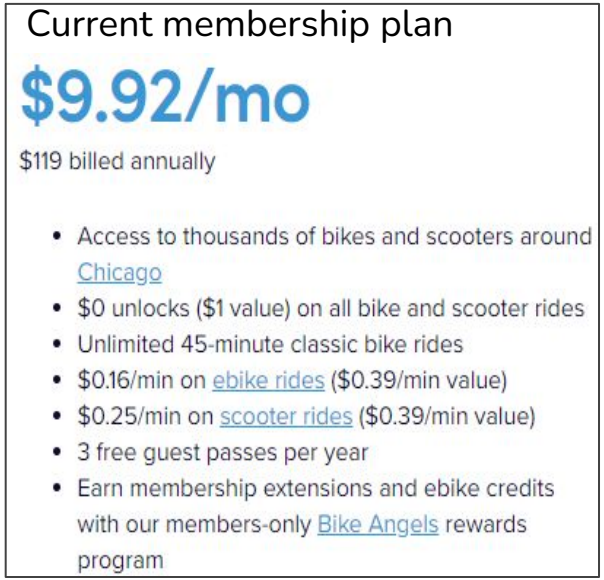
Sales and Marketing team's best date/time/location to promote to casual riders:

Best Date: weekends in summer, fall and spring

Best Time: 12pm - 6pm

Best Location: top 3 stations -Streeter Dr & Grand Ave
-DeSable Lake Shore Dr & Monroe St
-Millenium Park

Best Strategy: -Offer long ride discount in membership as incentive
-Offer docked bike as an option in membership
-Identify casual riders that use classic bike and advise on how membership can help them save
-Identify casual riders that regularly do short rides or commute and inform them about membership plan
-Tailor a membership plan for long ride users



Explore Further

-Validate/invalidate outliers

-Additional data column of member id or rider id

-Additional data column of price paid by users for each ride

-Compare to other yearly datasets

Questions?