Elijah Borque - Cyclistic Case Study Deliverable Run-Down

Ask - Business Task:
I will use data to determine how the behaviors and habits of casual bike riders using the Cyclistic bike service compare and/or contrast to those of bike riders who purchase and possess an annual membership of the Cyclistic bike share service.

Prepare - Description of Data Sources:
Where is your data located?

- Index of bucket "divvy-tripdata"
- Months from 2024
- Exported to Spreadsheets
- Cyclistic internal data

How is the data organized?

- CSV
- Contains columns: ride_id, rideable_type, started_at, ended_at, start_station_name, start_station_id, end_station_name, end_station_id, start_lat, start_lng, end_lat, end_lng, member_casual

Process - Documentation of any cleaning or manipulation of data:
Some entries are missing information, including: the ride's start station name, start station ID, end station name, end station ID. However, these entries do have starting latitudes and longitudes. These could possibly be attributed to customers using the bikes for their transportation, but deciding to not end their ride at a bike station. But there are duplicates of these longitudes and latitudes, which needs to be further investigated. For the sake of the business problem at hand, this information is not necessary at the moment.

Two columns were created for analysis: ride_length and day_of_week. The ride_length column is calculated by subtracting the column started_at from the column ended_at to get the duration of each ride. The day_of_week column is determined by using the WEEKDAY() function to attribute the date that each ride was started with a number to determine which day of the week the ride was done. For example, a 1 would mean the ride was done on a Sunday, 2 for Monday, and so on.

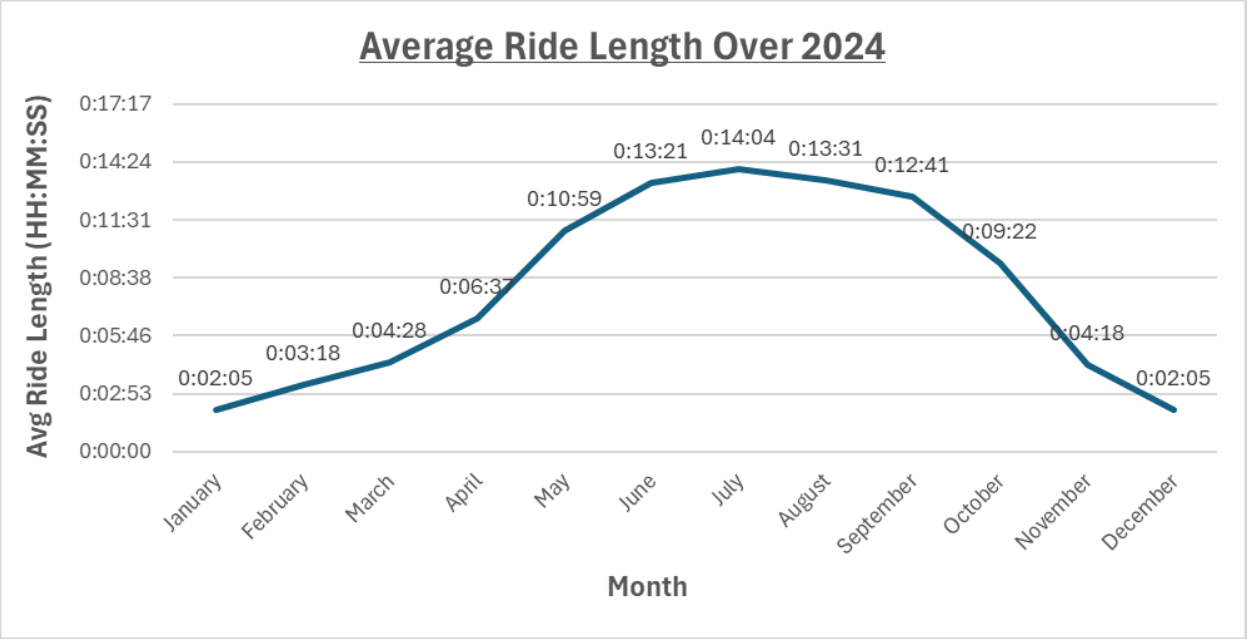
From these two columns, the following columns were created: avg_ride_length, max_ride_length, mode_day_of_week. The avg_ride_length column takes the mean of the ride_length column. The max_ride_length_column takes the maximum ride duration from the ride_length column. The mode_day_of_week takes the most frequent day value from the day_of_week column.

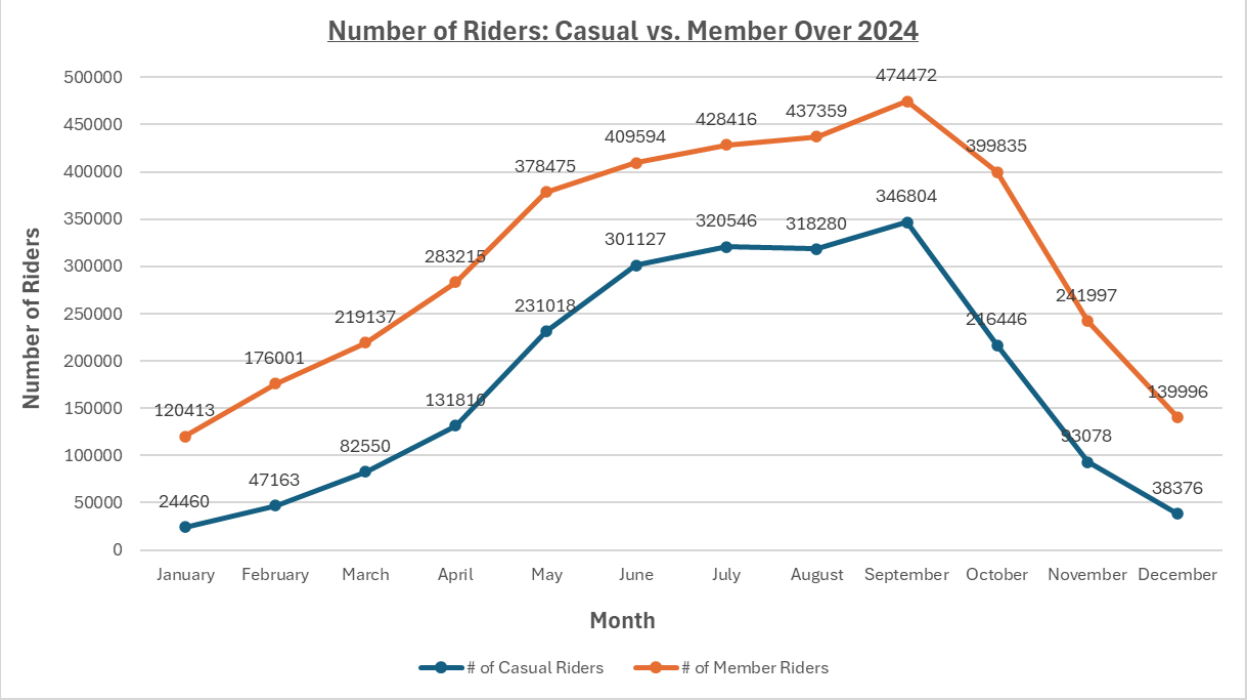
Some entries when calculating their ride_length, they came up as ######. This is because it was calculating a negative result due to the subtraction of a greater started_at value from a lower ended_at value. It is likely that the two values meant to be reversed, so to solve this issue, the equation for the ride_length column was modified to: =(D2-C2)*-1. This achieved the desired result for correctly calculating ride_length.

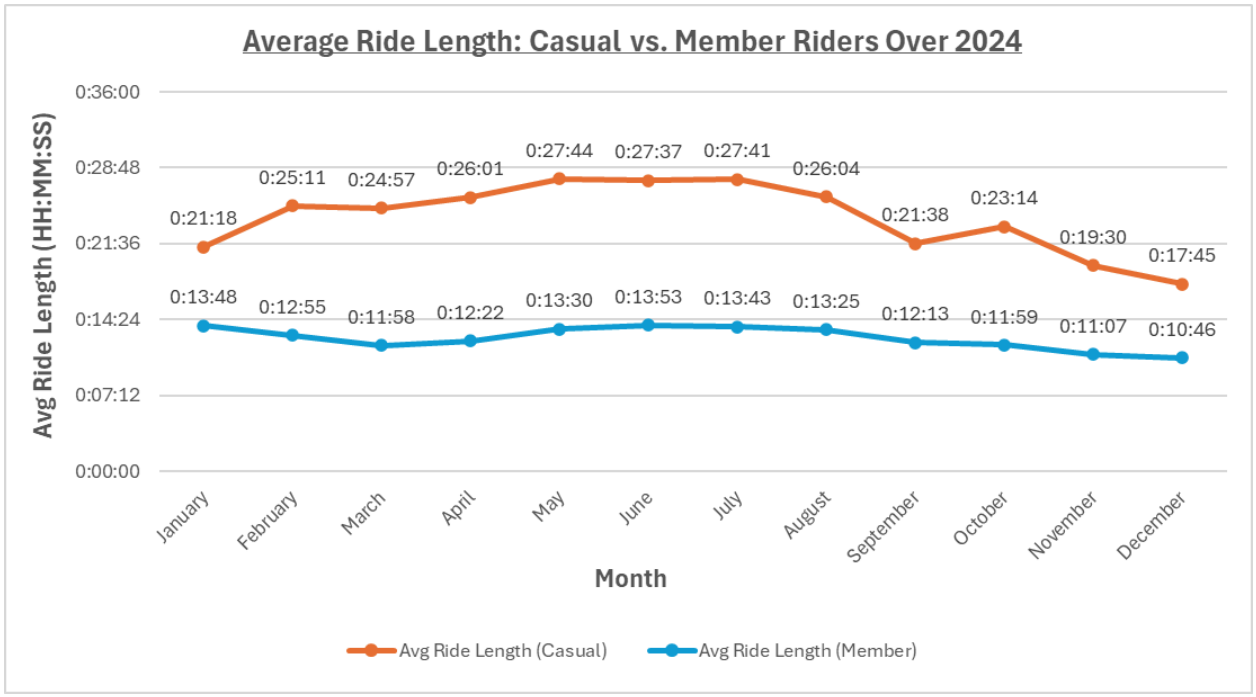
Analyze - A summary of your analysis:

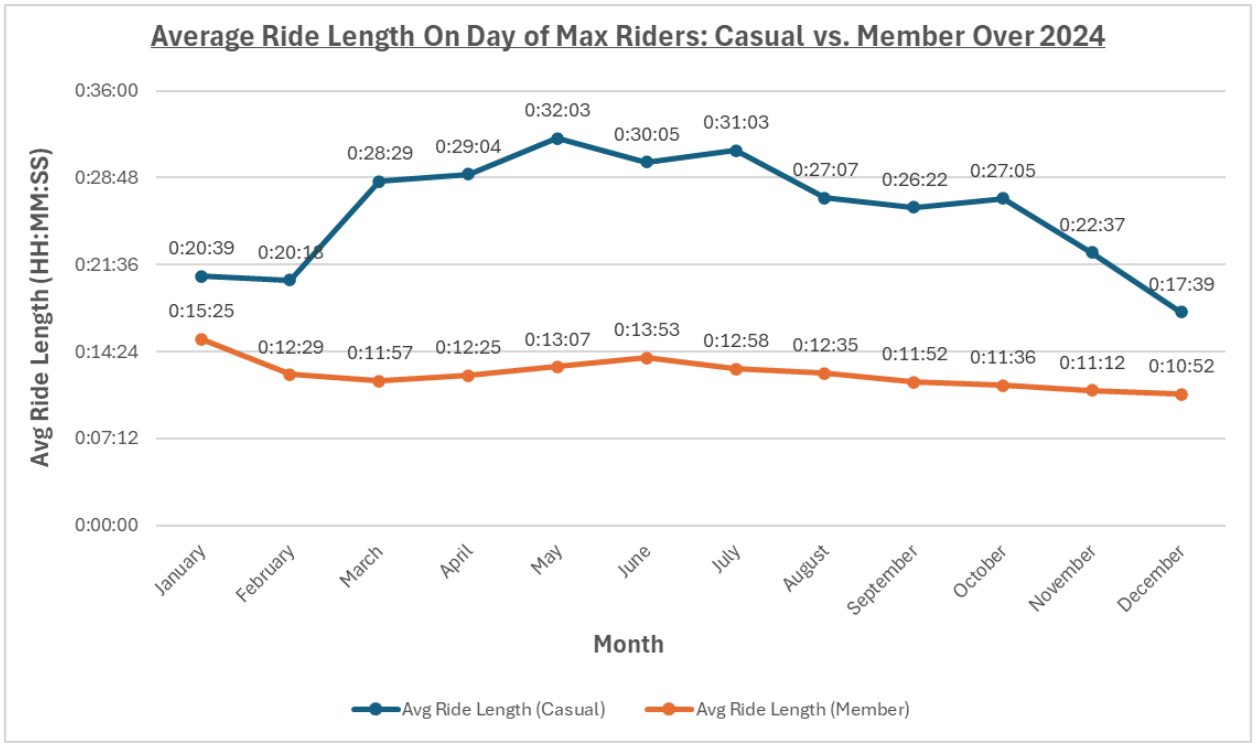
- Most frequent day for rides done for every month of 2024 was Saturday
- Average ride length line graph showed low average ride lengths in the beginning of the year, increasing until its peak in July, then decreasing for the rest of the year after July.
- Peak average ride length: 00:14:04 (hh:mm:ss) in July
- Lowest average ride length: 00:02:05 in January and December
- Possible correlation of weather and temperature impacting ride activity (higher ride activity during hotter seasons, lower ride activity during colder seasons)
- Casual riders had overall higher average ride lengths compared to member riders throughout the year
- Member riders had overall higher number of riders compared to casual riders throughout the year
- The number of riders for both casual riders and member riders follow trend of increasing until its peak in September, and then decreasing for the rest of the year
- Casual # of riders peak: 346804 in September; Member # of riders peak: 474472 in September
- Casual # of riders lowest: 24460 in January; Member # of riders lowest: 120413 in January
- Overall, throughout 2024, member riders had higher instances of max riders compared to casual riders.
- Overall, throughout 2024, casual riders had higher average ride length during the instances of max riders.
- Casual highest Max # of riders: 77329 in August; Member highest Max # of riders: 67845 in July
- For casual riders, for the days of the month where the max amount of riders occurred, Saturday was the most frequent (7 times) (0 times for members)
- For member riders, Wednesday (5 times) (1 time for casual)

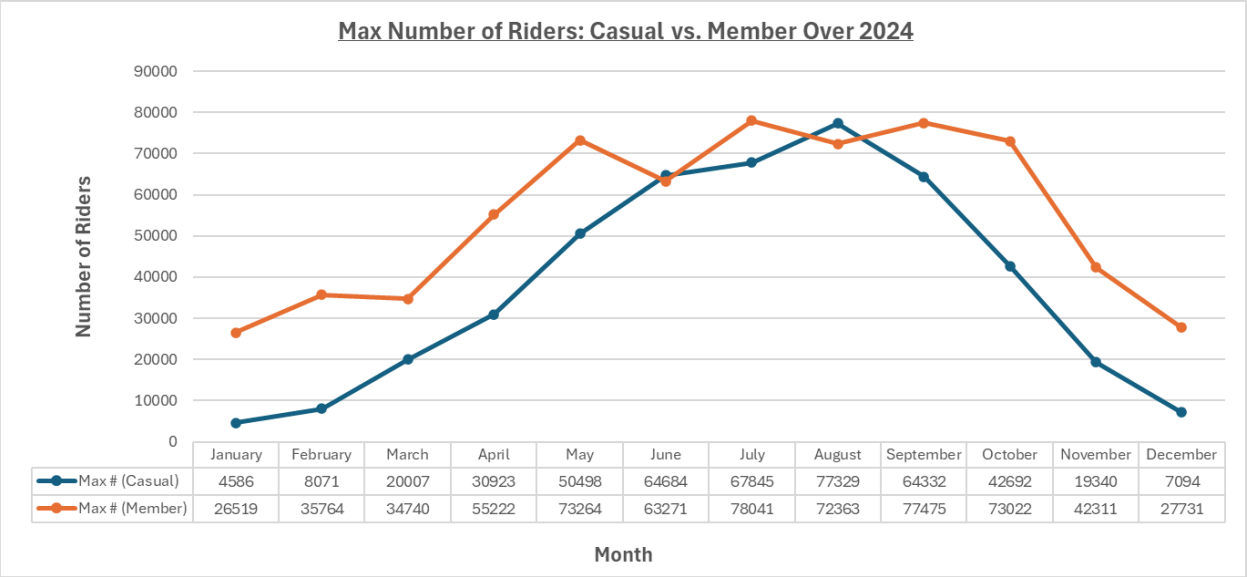
Share - Supporting visualizations and key findings



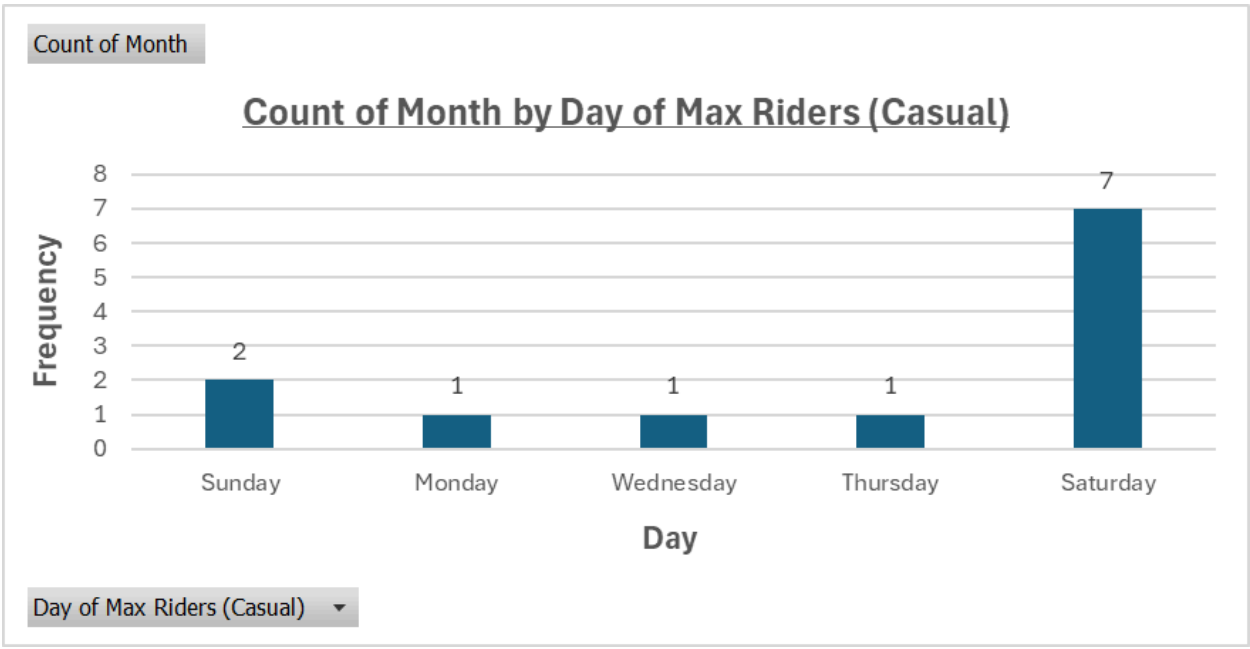


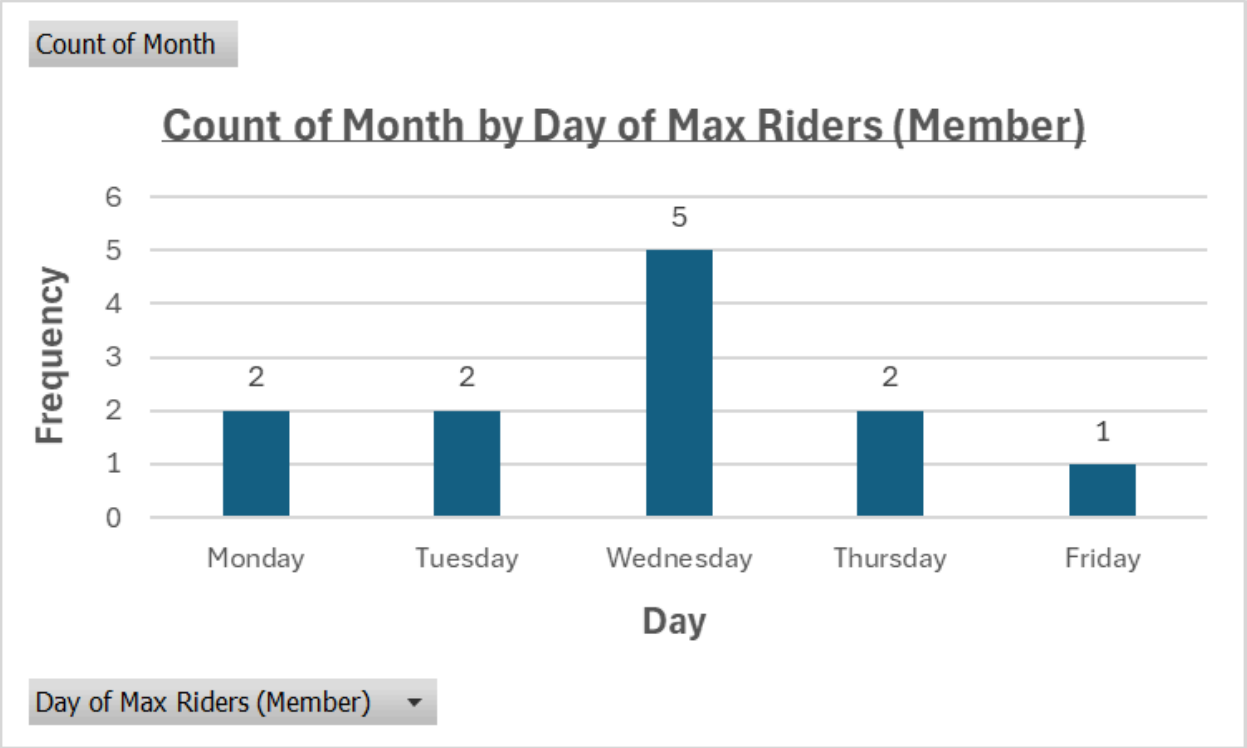

Average Ride Length: Casual vs. Member Riders Over 2024
Avg Ride Length (HH:MM:SS)
0:36:00
0:28:48
0:21:36
0:14:24
0:07:12
0:00:00
0:21:18
0:25:11
0:24:57
0:26:01
0:27:44
0:27:37
0:27:41
0:26:04
0:21:38
0:23:14
0:19:30
0:17:45
0:13:48
0:12:55
0:11:58
0:12:22
0:13:30
0:13:53
0:13:43
0:13:25
0:12:13
0:11:59
0:11:07
0:10:46
January
February
March
April
May
June
July
August
September
October
November
December
Month
Avg Ride Length (Casual)
Avg Ride Length (Member)


Average Ride Length On Day of Max Riders: Casual vs. Member Over 2024
Avg Ride Length (HH:MM:SS)
0:36:00
0:28:48
0:21:36
0:14:24
0:07:12
0:00:00
0:20:39
0:20:13
0:28:29
0:29:04
0:32:03
0:30:05
0:31:03
0:27:07
0:26:22
0:27:05
0:22:37
0:17:39
0:15:25
0:12:29
0:11:57
0:12:25
0:13:07
0:13:53
0:12:58
0:12:35
0:11:52
0:11:36
0:11:12
0:10:52
January
February
March
April
May
June
July
August
September
October
November
December
Month
Avg Ride Length (Casual)
Avg Ride Length (Member)



	January	February	March	April	May	June	July	August	September	October	November	December
Max # (Casual)	4586	8071	20007	30923	50498	64684	67845	77329	64332	42692	19340	7094
Max # (Member)	26519	35764	34740	55222	73264	63271	78041	72363	77475	73022	42311	27731





Act - Your top three recommendations based on your analysis

Spring/summer membership deals/promotions

Winter membership deals

Host summer marathon event

Friend Referral program

Seasonal merch accompanied with membership

Weekend deals for memberships (Saturday)