

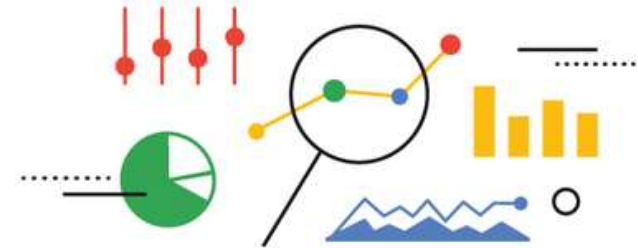
Cyclistic Bike Share Case Study Report

By: Kyle Heng



Project Introduction

- This is the capstone project component of my Google Data Analytics Career Certificate curriculum
- The project will follow the main steps of the data analysis process:
 - Ask, Prepare, Process, Analyze, Share, Act
- The data being analyzed is from “Cyclistic”, a fictional bike-share company based in Chicago
- The tools I use to conduct this analysis are Microsoft Excel, R Studio, and Tableau



Case Study Background

- Cyclistic is a bike-share company based in Chicago that features more than 5800 bicycles and 600 docking stations
- Director of marketing believes that company success is driven by maximizing the number of annual memberships
- Main Objective of Analysis:
 - Determine how casual riders and annual members use Cyclistic Bikes
 - Create visualizations that highlight trends in bike usage between the two groups
 - Provide insights backed by the analysis that would increase annual membership



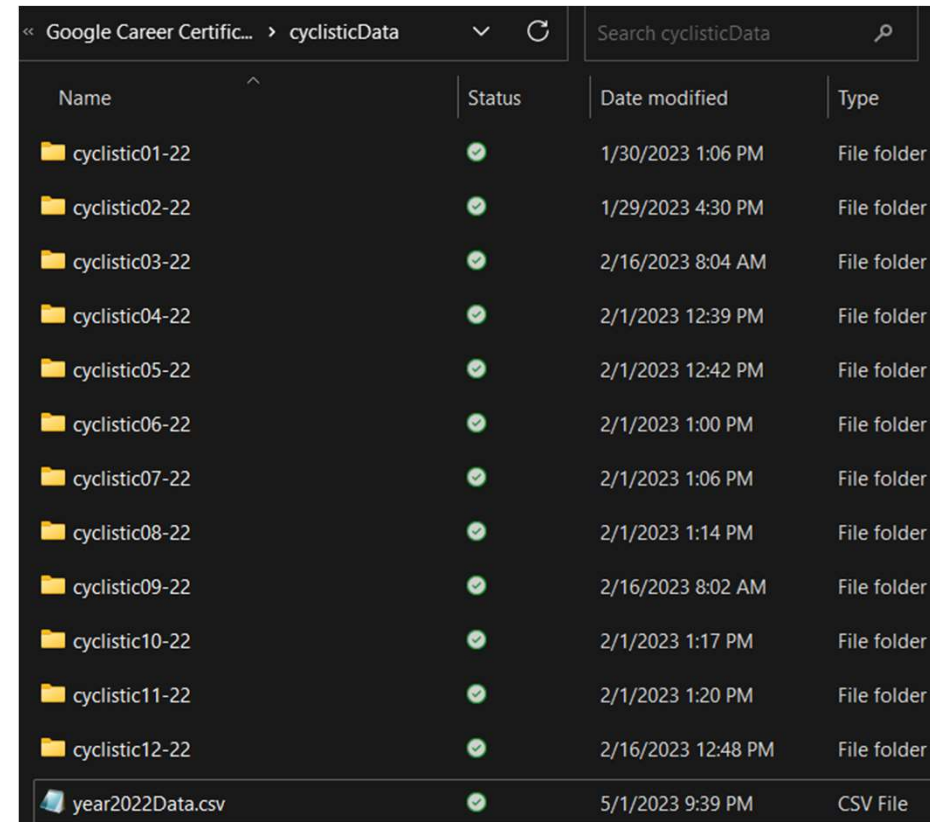
Ask:

- Questions to consider for Cyclistic Marketing:
 - How do annual members and casual riders use the bikes differently?
 - What would incentivize casual riders to buy annual memberships?



Prepare:

- I will be using public Cyclistic Trip Data provided by the course
 - For clarification, datasets have a different name (Divvy) since Cyclistic is a fictional company.
 - Data was made publicly available by Motivate International Inc and contains no personally identifiable information of riders.
- Downloading data and organizational structure:
 - Downloaded 12 datasets: one for each month of the year 2022
 - Created a folder named `cyclisticData` in project directory with folders for each month
 - Each folder contains the original `.csv` file and an `.xls` file that I will use for the processing step



Name	Status	Date modified	Type
cyclistic01-22	✓	1/30/2023 1:06 PM	File folder
cyclistic02-22	✓	1/29/2023 4:30 PM	File folder
cyclistic03-22	✓	2/16/2023 8:04 AM	File folder
cyclistic04-22	✓	2/1/2023 12:39 PM	File folder
cyclistic05-22	✓	2/1/2023 12:42 PM	File folder
cyclistic06-22	✓	2/1/2023 1:00 PM	File folder
cyclistic07-22	✓	2/1/2023 1:06 PM	File folder
cyclistic08-22	✓	2/1/2023 1:14 PM	File folder
cyclistic09-22	✓	2/16/2023 8:02 AM	File folder
cyclistic10-22	✓	2/1/2023 1:17 PM	File folder
cyclistic11-22	✓	2/1/2023 1:20 PM	File folder
cyclistic12-22	✓	2/16/2023 12:48 PM	File folder
year2022Data.csv	✓	5/1/2023 9:39 PM	CSV File

Process:

- Microsoft Excel Steps:

- For each month's data, two new columns were created:
 - ride_length: length of each ride
 - Calculated as difference between started_at and ended_at datetime columns
 - Formatted as HH:MM:SS using the 37:30:55 time format
 - day_of_week:
 - Day of the week each ride started
 - Used WEEKDAY() on started_at column
 - Values range from 1-7 with 1: Sunday and 7: Saturday



	C	D	E	F	G	H	I	J	K	L	M	N	O
1	started_at	ended_at	start_station_name	start_station_id	end_station	end_station_id	start_lat	start_lng	end_lat	end_lng	member_casual	ride_length	day_of_week
2	2/19/2022 18:08	2/19/2022 18:23	State St & Randolph St	TA1305000029	Clark St &	13179	41.88462	-87.6278	41.91569	-87.6346	member	0:15:15	7
3	2/20/2022 17:41	2/20/2022 17:45	Halsted St & Wrightwood Ave	TA1309000061	Southport	TA1307000	41.92914	-87.6491	41.92877	-87.6639	member	0:04:26	1
4	2/25/2022 18:55	2/25/2022 19:09	State St & Randolph St	TA1305000029	Canal St &	13011	41.88462	-87.6278	41.87926	-87.6399	member	0:13:38	6
5	2/14/2022 11:57	2/14/2022 12:04	Southport Ave & Waveland Ave		13235 Broadway	13323	41.94815	-87.6639	41.95283	-87.65	member	0:06:57	2

Analyze:



- [Link to R Markdown File](#)
 - Contains R code and analysis steps in further detail
- R Studio Primary Steps:
 - Imported each month's data from Excel as separate data frames
 - Merged the month data into a single data frame called year2022_data
 - Created day, month, and year columns and a numerical ride_length column in seconds.
 - Conducted Descriptive Analysis of ride length and member status
 - Used ggplot2() to create visualizations
 - Exported .csv file of the Year 2022 trip data for further analysis in Tableau

- Summary Statistics of Ride Length:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.0	322.0	559.0	913.8	989.0	1206314.0

- Summary Statistics for Members vs. Casual Riders:

Ride Length (Seconds)	Casual Riders	Members
Mean	1252.1912	708.6305
Median	675	505
Min	0	0
Max	1206314	89996

R Visualizations of Ridership and Ride Duration By Member Status



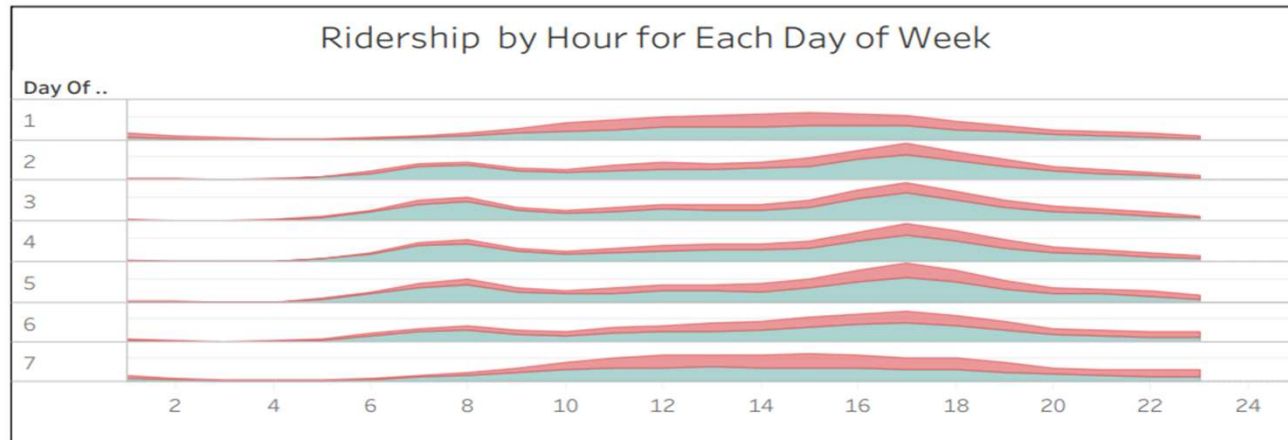
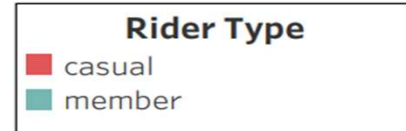
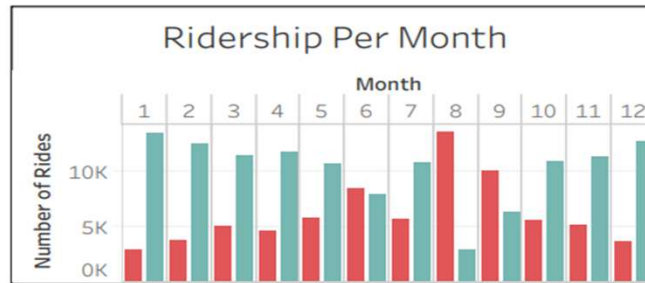
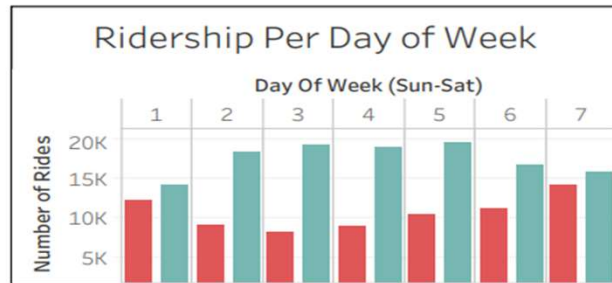
- Ridership for casual members increases during the weekends while members ride more on weekdays.
 - Many members are likely weekday commuters while casual members ride for leisure on the weekends
- Average Ride Duration decreases towards the middle of the week and increases during the weekends for casual members
- Average Ride Duration for members is more consistent with a slight increase on weekends
 - Members who commute likely go to the same point A and point B daily
- Members make up a majority of Cyclistic's ridership, but casual riders take longer rides

Share:

- After merging the monthly trip data with R into a single year of 2022 data frame, I exported the file to Tableau to create more visualizations for further analysis
- Building upon the visualizations I created in R, I wanted to further analyze other variables that could affect ridership and average ride duration between members and casual riders
- [Link to Tableau Workbook on Tableau Public](#)



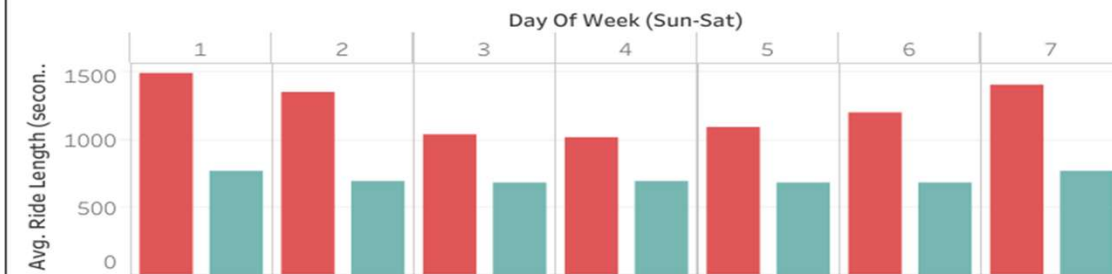
Cyclistic Bike Share Ridership Volume Trends



- For members, ridership slightly decreases going into summer and increases going into winter.
 - Since August and September already has missing location data, there might be missing data for members due to noticeably lower number of rides
- Casual ridership increases going into the summer months and decreases going into winter months
 - Colder temperatures would likely deter casual riders from going on more rides
- For weekdays, ridership peaks during commuting times around 8 am and 5 pm regardless of member status.
- Despite members taking more rides than casual riders all week, casual ridership out beats member ridership at peak hours
- Classic bikes are mostly used by members, docked bikes are only used by casual riders, and electric bikes are almost equally favored with a slight edge towards members

Ride Duration Trends

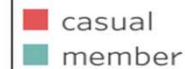
Average Ride Duration per Day of Week



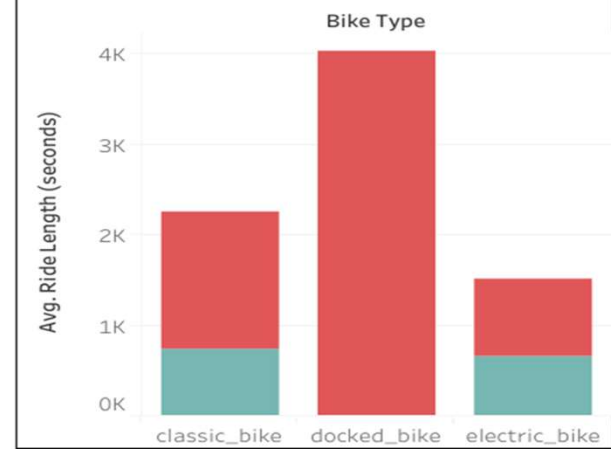
Average Ride Duration per Month



Rider Type

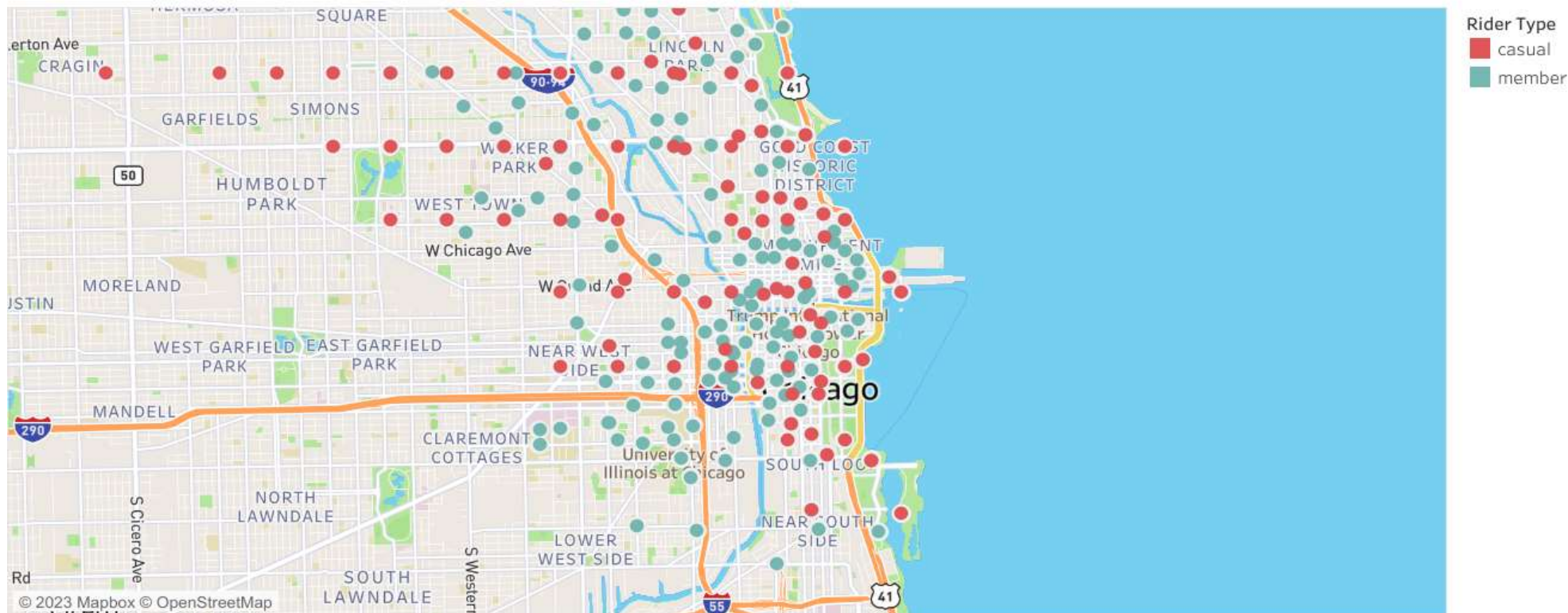


Average Ride Duration by Bike Type



- Ride duration for casual riders increases going from winter to spring, peaking in May. These numbers drop during the summer and slowly increase until November before decreasing again in December.
 - May is likely the most comfortable month to go on longer rides since winter and summer have extreme temperatures. As fall approaches, ride duration begins to increase again due to more moderate temperatures.
- Riders who use docked bikes take the longest rides while electric bike users take the shortest rides
 - Distance between docks could force riders to ride for more time while electric bikes are limited by their battery
- Casual riders take longer rides than members, regardless of day of week, month, or the type of bike.

Popular Starting Points for Riders by Rider Type



- Both members and casual riders start rides the most around downtown
 - There is a higher concentration of members towards the center of downtown and by the University of Illinois at Chicago, where there are likely more commuters
 - Casual riders typically start rides along the lake closer to the tourist attractions or in residential area further away from downtown.

Act:

- Revisiting Questions from the Ask Stage:

- **How do annual members and casual riders use the bikes differently?**

- Annual members likely use the bikes to commute to work or school, resulting in trips that are shorter but more consistent than casual riders. They typically start rides around centralized areas of downtown with more offices or around school campuses.
- Casual riders typically ride more on the weekends than during the week but will take longer rides than members. There is a higher concentration of them around areas of the city with more attractions such as the lake or in neighborhoods further from downtown.
- Members will opt for classic or electric bikes over the docked bikes while casual riders seem to favor electric bikes the most.

- **What would incentivize casual riders to buy annual memberships?**

- A membership option for weekends only that has a reduced cost
- Increasing the number of electric bikes available
- Advertising membership benefits at tourist hot spots
- Membership discounts for students who attend school in the city

