

Cyclistic Case Study

1. Ask

Question: How do annual members and casual riders use Cyclistic bikes differently?

- My insights can drive business decisions because understanding the difference can help develop a strategic business goal and transform those casual riders into annual members.

Key Tasks

- Identify the business task
- Consider key stakeholders

Deliverables

- A clear statement of the business task
 - To uncover how annual members and casual riders use Cyclistic bikes differently.

2. Prepare

Questions:

- Where is your data located? Using my Office 365 drive, I have stored the dataset under Desktop>Cyclistic Case Study>Excel Dataset
- How is the data organized? Zipped files are separated by year and month. They contain fields such as ride_id, started_at, start_station_name, start_long, and others.
- Are there issues with bias or credibility in this data? Does your data ROCCC? There seems to be little to no bias. My data does ROCCC because it is reliable, original, comprehensive, current, and cited.
- How are you addressing licensing, privacy, security, and accessibility? The data has been made available by Motivate International Inc. under the [Data License Agreement | Divvy Bikes](#).
- How did you verify the data's integrity? I verified the data's integrity because it comes

from a reliable source, includes a license, and is a public dataset.

- How does it help you answer your question? This dataset will allow me to uncover more about the ways different customer types are using the bikes at Cyclistic.
- Are there any problems with the data? Data appears to be reliable and comprehensive but I do have data for more than 12 months worth of data and am wondering which specific year I will be focusing on.

Key tasks

- Download data and store it appropriately.
- Identify how it's organized.
- Sort and filter the data.
- Determine the credibility of the data.

Deliverables

- A description of all data sources used.
 - I have decided to use Microsoft 365 as a central location for my case study
 - It is contained on my Desktop in a folder named Cyclistic Case Study in a hierarchical format

3. Process

Questions

- What tool are you choosing and why? I decided to use Excel because it is reliable and offers cloud services so I can work in different workstations.
- Have you ensured your data integrity? Yes, I ensured data integrity in the previous step.
- What steps have you taken to ensure that your data is clean? I have searched for duplicates, and filtered and sorted data.
- How can you verify that your data is clean and ready to analyze? Filtering and sorting is a great way to ensure that data is clean and ready to analyze. For example, ensuring that there are no duplicates and that there are no values outside a specified

range.

- Have you documented your cleaning process so you can review and share those results? I am documenting changes I have been making, particularly through this roadmap.

Key Tasks

- Check the data for errors
- Choose your tools
- Transform the data so you can work with it effectively
- Document the cleaning process

Deliverables

- Documentation of any cleaning or manipulation of data
 - Added ride_length and day_of_week columns
 - Had questions about abnormally small and large values but kept them in for the sake of the study. Also, there were blank values that I either wanted to fill in or delete
 - Sorted/ filtered the sheet based on ride_length values and deleted negative values because they showed that the started_at was greater than the ended_at which does not make sense
 - Continued to apply the same techniques to the other 11 tables
 - There was an abnormal amount of missing values in the 202012-divvy-tripdata

4. Analyze

Questions

- How should you organize your data to perform analysis on it? I reorganized my data to meet the needs of using R for data cleaning and manipulation. I ensured the columns were uniform among all datasets and combined the tables dating a year-long into one big dataset.
- Has your data been properly formatted? I my data has been properly formatted as I ensured that all of the tables were compatible and had appropriate column names so

that they could integrate together without data loss.

- What surprises did you discover in the data? I was surprised to find out that casual users had a greater average duration of rides compared to subscribed members.
- What trends or relationships did you find in the data? I found insightful trends in the data such as that subscribed members do consistently produce the highest number of rides in all days of the week. However, it is casual users had the highest average duration of bike rides in all weekdays. Thursdays were a peak day for the amount of time casual members preferred to ride bikes and had the lowest overall number of rides on Sundays and Saturdays. On the other hand, subscribed members utilized their bikes the most on the weekend.
- How will these insights help answer your business questions? These insights help answer business questions because they uncover that subscribed members do have higher number of rides than casual users but do not use the bikes longer than casual members. A great incentive for casual members would be to promote a program based on bike miles in exchange for some type of compensation such as chips/ merch/ points. This would encourage casual bike members to sign up for a membership and increase overall profits.

Key Tasks

- Aggregate your data so it's useful and accessible
- Organize and format your data
- Perform calculations
- Identify trends and relationships

Deliverables

- A summary of your analysis

I have combined all data from 2019 to 2020 into a single file using R Studio and conducted further cleaning and adding of data to prepare for analysis. Lastly, I conducted a descriptive analysis and found the following key data points:

- The minimum ride_length was 1
 - ◆ Casual: 2
 - ◆ Member: 1
- The maximum ride_length was 10628422
 - ◆ Casual: 10628422
 - ◆ Member: 6096428
- The mean ride_length was 1189
 - ◆ Casual: 5370.8
 - ◆ Member: 795.1
- The median ride_length was 539
 - ◆ Casual: 1393

◆ Member: 508

	member_casual	weekday	number_of_rides	average_duration
	<chr>	<ord>	<int>	<dbl>
1	casual	Sun	18652	5059.
2	casual	Mon	5591	4751.
3	casual	Tue	7311	4561.
4	casual	Wed	7690	4480.
5	casual	Thu	7147	8450.
6	casual	Fri	8013	6089.
7	casual	Sat	13473	4946.
8	member	Sun	60197	972.
9	member	Mon	110430	822.
10	member	Tue	127974	769.
11	member	Wed	121902	712.
12	member	Thu	125228	707.
13	member	Fri	115168	797.
14	member	Sat	59413	974.

- Number of rides/ average duration of casual v members based on days of the week.

5. Share

Questions

- Were you able to answer the question of how annual members and casual riders use Cyclistic differently? Yes, I was able to conduct an analysis and uncover some insights about how casual riders and annual members use Cyclistic differently.
- What story does your data tell? I found insightful trends in the data such as that subscribed members do consistently produce the highest number of rides in all days of the week. However, it is casual users had the highest average duration of bike rides in all weekdays. Thursdays were a peak day for the amount of time casual members preferred to ride bikes and had the lowest overall number of rides on Sundays and Saturdays. On the other hand, subscribed members utilized their bikes the most on the weekend.
- How do your findings relate to your original question? These insights help answer business questions because they uncover that subscribed members do have higher number of rides than casual users but do not use the bikes longer than casual members. A great incentive for casual members would be to promote a program based on bike miles in exchange for some type of compensation such as chips/ merch/ points. This would encourage casual bike members to sign up for a membership and increase overall profits.
- Who is your audience? What is the best way to communicate with them? My audience is the executive team and the best way to communicate with them would be either in a personal meeting or through a Zoom call in order to tell them all my findings.

- Can data visualizations help you share your findings? Data visualizations help share my findings as they help tell a story and allow data to be easily digested by the people in meetings.
- Is your presentation accessible to your audience? Yes, printouts/agendas would be appropriate for meetings as well as emailed presentations.

Key Tasks

- Determine the best way to share your findings
- Create effective data visualizations
- Present your findings
- Ensure your work is accessible

Deliverables

- Supporting visualizations and key findings
 - Casual riders ride longer but less frequently.
 - Subscribed members have a higher number of rides but shorter durations.
 - Potential for targeted marketing and incentives based on these patterns.
 - Casual riders have a higher average ride duration compared to annual members.
 - Thursdays are peak days for casual riders in terms of ride duration.
 - Casual riders have the lowest overall number of rides on Sundays and Saturdays, while subscribed members utilize their bikes the most on weekends.

6. Act

Questions

- What is your final conclusion based on your analysis? Annual members use bikes more frequently, while casual riders take longer rides, especially on Thursdays.
- How could your team and business apply your insights? By targeting casual riders with promotions to convert them to annual memberships, implementing a rewards program, and optimizing bike availability based on usage patterns.
- What next steps would you or your stakeholders take based on your findings?
 - Launch a rewards program for casual riders.

- Develop targeted marketing campaigns.
- Adjust bike distribution and maintenance schedules.
- Is there additional data you could use to expand on your findings? Yes, additional data such as rider demographics, seasonal trends, customer feedback, and competitive analysis can provide deeper insights and improve strategies.

Key Tasks

- Create your portfolio
- Add your case study
- Practice presenting your case study to a friend or a family member

Deliverables

- Your top three recommendations based on your analysis
 - Implement a rewards program for casual riders to convert them into annual members.
 - Promote weekend biking events or incentives for annual members.
 - Consider targeted marketing campaigns based on peak usage times.



