



Data Journal

Date: Feb 1	Course/topic: Cyclistic
Prompt:	What does the word “data” mean to you?
Journal Entry:	When I think of data, I think of records and information. Data is a digital record. It could be a record of a sale, GPS records of where a car has been, or posts of images and pictures on social media. Data in itself isn’t useful, because there’s so much of it. However, enough data usually contains information and patterns. This information is valuable, and data analysts are the ones who extract it.
Other thoughts or questions:	The amount of data that is created each year gets bigger and bigger! Data is important! Nowadays, companies that don’t use their data are at a disadvantage to those that do.

Date: 2_8_2024	Course/topic: ASK
Prompt:	What do you hope to take away from this capstone project? What is one important skill you think you’ll learn? Which skills do you most look forward to demonstrating? What are some issues you might encounter?
Journal Entry:	I hope it will show my skill with data analytics and teach me something new. I hope to demonstrate a skill for discernment that will positively affect the outcome measures that will help solve our questions. I might encounter an issue with software and understanding how to manipulate the data the way I need.
Other thoughts or questions:	I hope I can do this capstone project and maybe a few more to highlight skills so I can get a job. I can’t wait for the new journey.

Date: 2_8_2024	Course/topic: PREPARE
-----------------------	------------------------------



Prompt:	How do annual members and casual riders use Cyclistic bikes differently? Why would casual riders buy Cyclistic annual memberships? How can Cyclistic use digital media to influence casual riders to become members?
Journal Entry:	Knowing the answers to these questions can lead to a clear picture of how to increase casual riders to annual members. Identify the business task Increase memberships Consider key stakeholders Cyclistic members Moreno Cyclistic marketing analytics team Cyclistic executive team
Other thoughts or questions:	Assumptions 1. Casual riders don't have a need for bike transportation as often as annual members. 2 It is too expensive to use the bike unless absolutely needed. 3 It is not convenient enough to become a member. 4 They don't have a strong interest in riding often.

Date: 2_9_2024	Course/topic: PROCESS
Prompt:	Download data and store it appropriately. Identify how it's organized. Sort and filter the data. Determine the credibility of the data.
Journal Entry:	Where is your data located? https://divvy-tripdata.s3.amazonaws.com/index.html How is the data organized? Zip files. Are there issues with bias or credibility in this data? Does your data ROCCC? Source evaluation: Credible source. Data collection methodology: Random selection, ranges seems appropriate. Potential biases: no biases detected. Missing data: Missing data will be addressed/fixed. Timeliness: Data is collected monthly and seems to fit needed timeline. Consistency: collection is consistent from all data collected. Relevance: extremely relevant. How are you addressing licensing, privacy, security, and accessibility? The data will be locked behind passwords and accessibility locks.
Other thoughts or questions:	Data is stored online in zip files containing the current month and also several years back worth of bike cycling data.



Date: 2_10_2024	Course/topic: PROCESS
Prompt:	What tools are you choosing and why? Have you ensured your data's integrity? What steps have you taken to ensure that your data is clean? How can you verify that your data is clean and ready to analyze? Have you documented your cleaning process so you can review and share those results?
Journal Entry:	<p>I will be using Big Query because it can handle the larger amounts of data I will be feeding it.</p> <p>I have ensured the data's integrity by</p> <ol style="list-style-type: none">1. Implement data validation2. Maintain clear documentation3. Establish data entry protocols4. Perform regular data quality checks5. Backup data6. Use access controls7. Implement change tracking8. Perform regular maintenance9. Use error-detection and correction methods10. Train staff <p>Steps I have taken to ensure that my data is clean.</p> <ol style="list-style-type: none">1. Remove duplicates2. Handle missing values3. Check for outliers4. Ensure consistency5. Correct errors6. Standardize data7. Validate data types8. Remove irrelevant data9. Perform feature engineering10. Documented my cleaning process <p>How can I verify that my data is clean and ready to analyze?</p> <ol style="list-style-type: none">1. Visual inspection2. Summary statistics3. Data profiling4. Cross-validation5. Validation rules6. Test your cleaning process7. Check for missing values8. Data consistency9. Model readiness10. Peer review



	Have you documented your cleaning process so you can review and share those results? YES!
Other thoughts or questions:	I have Checked the data for errors. I have Chosen my tools. I have transformed the data so I can work with it effectively. I have Documented the cleaning process. I have saved all my queries.

Date: 2_10_2024	Course/topic: Analyze
Prompt:	Aggregate your data so it's useful and accessible. Organize and format your data. Perform calculations. Identify trends and relationships.
Journal Entry:	<p>How should you organize your data to perform analysis on it? Has your data been properly formatted? What surprises did you discover in the data? What trends or relationships did you find in the data? How will these insights help answer your business questions?</p> <p>Question 1: Is the number of rides increasing? Answer: Yes, the number of rides appears to be increasing for both 'casual' and 'member' users from January 2023 to February 2024. The total number of rides is higher in February 2024 compared to January 2023 for both user types.</p> <p>Question 2: Is the proportion of casual users vs member users changing? Answer: To determine if the proportion of 'casual' users vs 'member' users is changing, you would need to calculate the proportion of rides for each user type in each month. Based on the provided results, the proportion of rides made by 'casual' users compared to 'member' users seems to be decreasing over time. In January 2023, the proportion of 'casual' rides to 'member' rides is approximately 1:3.7. In February 2024, the proportion is around 1:4.2, indicating that the proportion of 'member' rides is increasing compared to 'casual' rides.</p> <p>Question 3: Is the average distance travelled per ride changing? Answer: The average distance traveled per ride fluctuates for both 'casual' and 'member' users throughout the given period. For 'casual' users, the average distance per ride decreases from January 2023 to February 2023, then increases significantly in January 2024, and decreases again in February 2024. For 'member' users, the average distance per ride fluctuates less, but there is a general decrease from January 2023 to February 2024.</p> <p>Question 4: Is there a difference in rides made by casual users vs member users?</p>



	Answer: Yes, there is a difference in the rides made by 'casual' users vs 'member' users. 'Member' users consistently make more rides than 'casual' users, and the average distance per ride is generally lower for 'member' users compared to 'casual' users.
Other thoughts or questions:	Members use the services more often for shorter trips. Casual riders take longer trips less often. Targeted adds appealing to individual needs of casual riders.

Date: 2_11_2024	Course/topic: Share
Prompt:	Determine the best way to share your findings. Create effective data visualizations. Present your findings. Ensure your work is accessible.
Journal Entry:	Were you able to answer the question of how annual members and casual riders use Cyclistic bikes differently? yes. What story does your data tell? That members use bikes more often for shorter trips. Casual riders ride bikes farther for a longer period. How do your findings relate to your original question? It tells me how these two groups use this service differently and how these insights can help sell the product. Who is your audience? Stake holders. What is the best way to communicate with them? Showing the data in a visual way and explaining the patterns it has revealed. Can data visualization help you share your findings? Yes. Is your presentation accessible to your audience? Yes.
Other thoughts or questions:	Supporting visualizations and key findings

Date: 2_11_2024	Course/topic: Act
Prompt:	Create your portfolio. Add your case study. Practice presenting your case study to a friend or family member.



Journal Entry:	<p>What is your final conclusion based on your analysis? How could your team and business apply your insights? What next steps would you or your stakeholders take based on your findings? Is there additional data you could use to expand on your findings?</p> <p>Based on the findings from the data, annual members and casual riders use Cyclistic bikes differently in terms of ride frequency and distance traveled. Annual members generally take more rides compared to casual riders, suggesting that they are more frequent users of the service. On the other hand, casual riders tend to have a higher average distance per ride, indicating that they might use the bikes for longer trips or less frequently.</p> <p>Casual riders might be motivated to purchase Cyclistic annual memberships for several reasons:</p> <ol style="list-style-type: none"> 1. Cost savings: If casual riders use the bikes frequently enough, it might be more cost-effective for them to purchase an annual membership rather than paying for individual rides. 2. Convenience: Having an annual membership might offer more convenience, as they won't need to worry about paying for each ride individually. 3. Perks: Cyclistic could offer additional perks or benefits to annual members, such as priority access to bikes, discounted rates for friends and family, or special events for members. <p>To influence casual riders to become members using digital media, Cyclistic could consider the following strategies:</p> <ol style="list-style-type: none"> 1. Targeted advertising: Use online advertising platforms, such as social media ads or search engine ads, to target casual riders with messaging about the benefits of an annual membership, like cost savings, convenience, and exclusive perks. 2. Email marketing: Send personalized emails to casual riders highlighting the advantages of becoming an annual member based on their usage patterns. For example, if a casual rider frequently takes long rides, emphasize the potential cost savings of an annual membership. 3. Social media: Share testimonials or success stories from annual members on social media, showcasing the positive experiences and benefits they enjoy. This could encourage casual riders to consider upgrading their membership. 4. Content marketing: Create blog posts, infographics, or videos that explain the benefits of an annual membership and how it compares to casual ride pricing. Share these materials on social media and email newsletters to reach casual riders. 5. Retargeting campaigns: Set up retargeting campaigns to reach casual riders who have visited the Cyclistic website or used the service but have not yet purchased an annual membership. Offer limited-time promotions or discounts to incentivize them to upgrade.
Other thoughts or questions:	<p>Set up retargeting campaigns to reach casual riders who have visited the Cyclistic website or used the service but have not yet purchased an annual membership. Offer limited-time promotions or discounts to incentivize them to upgrade.</p>