



# Data Journal

## Instructions

Use this document as a template for the Journal Entry activities Project. Additionally, you can use the templates to take notes on what and How I've learned or just to jot down your thoughts about data analytics.

<b>Date:</b> May 30	<b>Course/topic:</b> Course 1: Data, data everywhere?
<b>Prompt:</b>	What does the word "data" mean to you? <b>Data mean to me in three ways?</b>
<b>Journal Entry:</b>	<p>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.</p> <p>1.</p>
<b>Other thoughts or questions:</b>	<p>The amount of data that is created each year gets bigger and bigger!</p> <p>Data is important! Nowadays, companies that don't use their data are at a disadvantage to those that do.</p>

<b>Date:</b> 30th May 2022	<b>Course/topic:</b> Capstone Project 1
<b>Prompt:</b>	<ol style="list-style-type: none"><li>1. What do you hope to take away from this capstone project? What is one important skill you think you'll learn?</li><li>2. Which skills do you most look forward to demonstrating?</li></ol>



	3. What are some issues you might encounter?
<b>Journal Entry:</b>	<ol style="list-style-type: none"> <li>1. My focus during this capstone project will be what are the key points and how to convert each step into a process.</li> <li>2. Reporting making, dashboard creative ideas and how to ask “what, how” related analytics questions while data analysis. And From to begin data cleaning.</li> <li>3. Sharing related processes, whom to ask, what points to remember while asking. How to network within the office for future connections.</li> </ol>
<b>Other thoughts or questions:</b>	Will come back when I remember while moving ahead.

<b>Date:</b>	<b>Course/topic:Ask i.e whole motive of the project</b>
<b>Prompt:</b>	<p>Three questions will guide the future marketing program:</p> <ol style="list-style-type: none"> <li>1. How do annual members and casual riders use Cyclistic bikes differently?</li> <li>2. Why would casual riders buy Cyclistic annual memberships?</li> <li>3. How can Cyclists use digital media to influence casual riders to become members?</li> </ol>
<b>Journal Entry:</b>	<ol style="list-style-type: none"> <li>1. Customers who purchase single-ride or full-day passes are referred to as casual riders. . Customers who purchase annual memberships are Cyclistic members.</li> <li>2.</li> </ol>
<b>Other thoughts or questions:</b>	

<b>Date:</b> 31-05-2022	<b>Course/topic:Capstone Track 1, Task : Company Overall Task.</b>
<b>Prompt:</b>	<p>Moreno has assigned you the first question to answer: How do annual members and casual riders use Cyclistic bikes differently?</p> <p>You will produce a report with the following deliverables:</p> <ol style="list-style-type: none"> <li>1. A clear statement of the business task</li> <li>2. A description of all data sources used</li> </ol>



	<ol style="list-style-type: none"> <li>Documentation of any cleaning or manipulation of data</li> <li>A summary of your analysis</li> <li>Supporting visualizations and key findings</li> <li>Your top three recommendations based on your analysis</li> </ol>
<b>Journal Entry:</b>	<ol style="list-style-type: none"> <li>The one line difference between Casual riders and annual members Cyclist is that Casual riders are new riders or one time riders. While of those Annual Members riders are those who prefer to take annual subscriptions for cycling.</li> <li>The data consist of 13 column which has distinct id of each rider, rider able type, start and end date of ride, Start &amp; End station name and id, latitude and longitude of start and end journey and rider status.</li> <li></li> </ol>
<b>Other thoughts or questions:</b>	

<b>Date:</b> 31st May 2022	<b>Course/topic:</b> Data Analysis First Step. ASK
<b>Prompt:</b>	<b>Guiding questions :</b> <ul style="list-style-type: none"> <li>What is the problem you are trying to solve?</li> <li>How can your insights drive business decisions?</li> </ul>
<b>Journal Entry:</b>	<b>Key tasks:</b> <ul style="list-style-type: none"> <li>Identify the business task? <ul style="list-style-type: none"> <li>To increase number of yearly annual subscription of Riders,</li> </ul> </li> <li>How can your insights drive business decisions? <ul style="list-style-type: none"> <li>If decision taker are convinced where have to put there effort eg in their marketing or sales or change plans Rates.</li> </ul> </li> </ul>
<b>Other thoughts or questions:</b>	<b>Key tasks:</b> To whom we need to create our analysis report keeping this people in mind. <ol style="list-style-type: none"> <li>Identify the task.</li> <li>Consider the Stakeholder Eg. VP or Team leader.</li> </ol> <b>Deliverable:</b> A clear statement of the business task.

<b>Date:</b> 31st May 2022	<b>Course/topic:</b> Data Analysis Second Step. PREPARE Prepare data for analysis using the following Case Study:
----------------------------	--



<b>Prompt:</b>	<b>Guiding Questions:</b> <ol style="list-style-type: none"> <li><b>Where is your data located?</b> <ol style="list-style-type: none"> <li>Data is being saved on a yearly basis. Starting from 2014 to 2022. Data before 2019 no data is of the whole year and located or stored at divy-tripdata their self database.</li> </ol> </li> <li><b>How is the data organized?</b> <ol style="list-style-type: none"> <li>Data is organized from 2014 to 2022. Since 2018 the data is collected monthly and save to the company dataset in CSV format.</li> </ol> </li> <li><b>Are there issues with bias or credibility in this data? Does your data ROCCC?</b> <ol style="list-style-type: none"> <li>Since the database is collected by Company self so the data is Primary data. So the issue of data being biased or credibility is not an issue.</li> </ol> </li> <li><b>How are you addressing licensing, privacy, security, and accessibility?</b> <ol style="list-style-type: none"> <li>This is public data that we used to explore how different customer types are using Cyclistic bikes. But with data-privacy usage of data prohibition we won't be able to use riders' personally identifiable information. This means we won't be able to connect pass purchases to credit card numbers to determine.</li> </ol> </li> <li><b>How did you verify the data's integrity?</b> <ol style="list-style-type: none"> <li>As we move ahead and do analysis we will get more insight like how many riders participate each year and whether each year the number of riders are increasing.</li> </ol> </li> <li><b>How does it help you answer your question?</b> <ol style="list-style-type: none"> <li>Analysis are there any trends or seasonal in increase or decrease of annual subscription..</li> </ol> </li> <li><b>Are there any problems with the data?</b> <ol style="list-style-type: none"> <li></li> </ol> </li> </ol>
<b>Journal Entry:</b>	<b>TASK :</b> <ol style="list-style-type: none"> <li>Download data and store it appropriately.             <ol style="list-style-type: none"> <li>Data is being store in CSV format and as the data is of 12 month from April 2020 to March 2021 so total size is 500 MB</li> </ol> </li> <li>Identify how it's organized.             <ol style="list-style-type: none"> <li>Data is well organized in tabular format and needs some cleaning procedure.</li> </ol> </li> <li>Sort and filter the data.             <ol style="list-style-type: none"> <li>Also need to see whether the data is sort date and time wise.</li> </ol> </li> <li>Determine the credibility of the data.             <ol style="list-style-type: none"> <li></li> </ol> </li> </ol>
<b>Other thoughts or questions:</b>	<ol style="list-style-type: none"> <li>A description of all data sources used             <ol style="list-style-type: none"> <li>Data has 14 columns where 8 are categorical and 4 are float data type</li> <li></li> </ol> </li> </ol>

<b>Date:</b> 1st June	<b>Course/topic:</b> Data Analysis Step Third. Process
-----------------------	--



2022	process your data for analysis using the following Case Study Roadmap as a guide:
<b>Prompt:</b>	<b>Guiding Questions:</b> <ol style="list-style-type: none"> <li><b>What tools are you choosing and why?</b> <ol style="list-style-type: none"> <li>As the data was stored monthly for each year, in order to analyze them we needed to save them in a single file for this purpose we use python coding.</li> </ol> </li> <li><b>Have you ensured your data's integrity?</b> <ol style="list-style-type: none"> <li>Yes data is genuine</li> </ol> </li> <li><b>What steps have you taken to ensure that your data is clean?</b> <ol style="list-style-type: none"> <li>Nope data is not clean, there is lots of missing value in Start and end station name columns.</li> </ol> </li> <li><b>How can you verify that your data is clean and ready to analyze?</b> <ol style="list-style-type: none"> <li>Python pandas and data cleaning libraries tools can be used to clean data.</li> </ol> </li> <li><b>Have you documented your cleaning process so you can review and share those results?</b> <ol style="list-style-type: none"> <li>Yes, written in short and simple words.</li> </ol> </li> </ol>
<b>Journal Entry:</b>	<b>Key tasks:</b> <ol style="list-style-type: none"> <li><b>Check the data for errors</b> <ol style="list-style-type: none"> <li></li> </ol> </li> <li><b>Choose your tools.</b> <ol style="list-style-type: none"> <li>Python pandas and datetime library.</li> </ol> </li> <li><b>Transform the data so you can work with it effectively.</b></li> <li><b>Document the cleaning process.</b> <ol style="list-style-type: none"> <li>Created two new columns naming it as "Weekday" and "ride_length".</li> </ol> </li> </ol>
<b>Other thoughts or questions:</b>	<b>Deliverable:</b> Documentation of any cleaning or manipulation of data

<b>Date:</b> 2 June	<b>Course/topic:</b> Analyze
<b>Prompt:</b>	<b>Guiding questions</b> <ul style="list-style-type: none"> <li>How should you organize your data to perform analysis on it?</li> <li>Has your data been properly formatted?</li> <li>What surprises did you discover in the data?</li> <li>What trends or relationships did you find in the data?</li> <li>How will these insights help answer your business questions?</li> </ul>
<b>Journal Entry:</b>	<b>Key tasks</b> <ol style="list-style-type: none"> <li>Aggregate your data so it's useful and accessible.</li> <li>Organize and format your data.</li> <li>Perform calculations.</li> <li>Identify trends and relationships.</li> </ol>



---

<b>Other thoughts or questions:</b>	Deliverable A summary of your analysis
-------------------------------------	--

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
--------------	----------------------



---

<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	



---

<b>Other thoughts or questions:</b>	
-------------------------------------	--

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
--------------	----------------------





---

<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	



---

<b>Other thoughts or questions:</b>	
-------------------------------------	--

---

<b>Date:</b>	<b>Course/topic:</b>
<b>Prompt:</b>	
<b>Journal Entry:</b>	
<b>Other thoughts or questions:</b>	

---

### Need another journal entry template?

If you have more journal entries to complete, please copy and paste one of the tables above and use it as a template for future entries.