Data Journal

Instructions

The director of marketing believes the company's future success depends on maximizing the number of annual memberships. Therefore, Company wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, the Company will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.

This is where I come in the picture, My job is to gather the data from different sources, manipulate that data so that data can be used to analyze and then make the visualization of the data and find the insights so that I can answer that one question my stakeholder wants me to answer

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

To answer this Question we have to produce a report with the following deliverables:

1. **A clear statement of the business task**
2. **A description of all data sources used**
3. **Documentation of any cleaning or manipulation of data**
4. **A summary of your analysis**
5. **Supporting visualizations and key findings**
6. **Your top three recommendations based on your analysis**

Everything I do in this project is described briefly below

Date: March 1	**ASK PHASE- A clear statement of the business task**
Guiding Question:	● What is the problem we are trying to solve? ● How can our insights drive business decisions?
Answers:	● Main objective of this project is to identify the differrence between the use of cyclistic by casula user and member user ● After completing the analysis we can easily figure out that what are things which can help company to change their casual users into permanent member
Key task	1. Identify the business task 2. Consider key stakeholders

Date: March 1	**PREPARE PHASE-A description of all data sources used**
Guiding Questions:	● Where is your data located? ● How is the data organized? ● Are there issues with bias or credibility in this data? Does your data ROCCC? ● How are you addressing licensing, privacy, security, and accessibility? ● How did you verify the data's integrity? ● How does it help you answer your question? ● Are there any problems with the data?
Answers:	● Data is provided by the company itself ● Dat is organized in a several columns with many rows and every month has a separate data ● Data is provided by the company itself so there is very less chance of biasness in the data but we have checked data and no signs of biasness is found ● Security of data is ensured by only giving access of data to the team only ● Data is provided by the company itself ● After cleaning and filtering of data we will try to find some insights on two variable like member and casual ● If we talk about technical problem then there is no problem in that but if we talk about the values present in the data then there are lot of work to make data good for analysis
Key task:	1. Download data and store it appropriately. 2. Identify how it's organized. 3. Sort and filter the data. 4. Determine the credibility of the data
Deliverables:	1. The source of data is company itself so here we assume the data provides to us is not biased, it is credible, and Integrity of data is good 2. We also make sure that data is secured as the access is only provided to the team so there is no chance of leaking any type of information.

Date: March 2	**PROCESS PHASE-Documentation of any cleaning or manipulation of data**
Guiding Questions:	● 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:	• We are going to use **MS-Excel** for filtering/cleaning the data, **MSSQL server** for joining the 12 months data into one table and then manipulating that dat to find some interesting insights , **Tableau public** for the visualization and **Google slides** to showcase those visualizations as tableau public don't allow to download the visualizations. • Data is provided by the company itself • We have deleted all the blank rows and variables which are in not good format and make sure that every column contains variable of same data type of which column is • We have checked each and every row simultaneously for any error. • yes
Key Tasks:	1. Determine the best way to share your findings. 2. Create effective data visualizations. 3. Present your findings. 4. Ensure your work is accessible.
Deliverables:	1. All the documentation regarding any type of cleaning or manipulating the data is available in the folder

Date: March 4	**ANALYZE PHASE-A summary of your analysis**
Guiding Questions:	• 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?
Answers:	• We have make sure that we have every information we need to analyze like there are longitude and latitudes given in dat which we have converted into distance and then we have to make a single table for 12 months of data • Data is properly formatted as we delete all the columns which are of no use after using them and add some new column also • All the insights we found in this data is available separately in the folder • The main basic relationship is between only two variables member and casual all the other variables are depend on them • The insights have some really surprising outcomes which helps us to understand that how people think while using the services of the company
Key Tasks:	1. Aggregate your data so it's useful and accessible. 2. Organize and format your data. 3. Perform calculations. 4. Identify trends and relationships.

Deliverables:	This is glimpse of the insights which we found in this data full detailed information is available in other doc file in this folder total no of member riders = 29,03,846 total no of casual riders = 22,93,244 no of electric_bike riders = 19,25,467 no of docked_bike riders = 2,50,015 no of classic_bike riders = 30,21,608 no of riders on Monday = 7,87,636 no of riders on Tuesday = 6,51,946 no of riders on wednesday = 6,92,279 no of riders on thursday = 7,07,973 no of riders on friday = 6,91,107 no of riders on saturday = 7,51,748 no of riders on sunday = 9,14,401 total no of start_sattion_name = 1590 total no of end_sattion_name = 1573 least busiest start_station_name = Whipple St & Irving Park Rd with 1 rider in a year most busiest start_station_name = Streeter Dr & Grand Ave with 65,102 riders in a year total distance covered by cyclist = 3,25,11,529.79 miles

Date: March 5	**SHARE PHASE-Supporting visualizations and key findings**
Guiding Questions:	● Were you able to answer the question of how annual members and casual riders use Cyclistic bikes differently? ● What story does your data tell? ● How do your findings relate to your original question? ● Who is your audience? What is the best way to communicate with them? ● Can data visualization help you share your findings? ● Is your presentation accessible to your audience?
Answers:	● YES ● Data tells that there are more number of members than casual riders which means that people are more interested in taking a membership then being a casual riders more detailed insights is also available in another file ● According to data there is more traffic on weekend than weekdays not only this people prefer one specific type of ride than other ● My audience is the stakeholder of the company. The best way to communicate with them is through visualization ● Data visualization helps to make data look very much understandable

	which helps us to make future decisions ● Every thing we do in this project is accessible to the audience
Key Tasks:	1. Determine the best way to share your findings. 2. Create effective data visualizations. 3. Present your findings. 4. Ensure your work is accessible.
Deliverables:	1. All the visualization is present in other file

Date: March 5	**ACT PHASE-Your top three recommendations based on your analysis**
Guiding Questions:	● 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?
Answers:	● Conclusion is that Members are more in number than Casual riders but the use of Casual riders cover more distance than the members in fact casual members hire bike for more time than permanent members ● We can use the information to make some new plans which can grab attention of casual rider to become member ● Next step will be to make some change in the plans and make some new exciting offers for casual riders ● NO
Key Tasks:	1. Create your portfolio. 2. Add your case study. 3. Practice presenting your case study to a friend or family member.
Deliverables:	1. The company should introduce some new plans for New membership so that casual riders get attracted to it 2. People are using electric bike for very less time maybe because of the battery life so company can increase there electric bike battery life which can attract more people who are environment friendly