Ask:

Guiding questions:
● **What is the problem you are trying to solve?** Converting casual riders into annual members.
● **How can your insights drive business decisions?** By making recommendations with compelling data insights and professional data visualizations.

Key tasks
1. Identify the business task:

To answer the following 3 questions:
A) How do annual members and casual riders use Cyclist bikes differently?
B) Why would casual riders buy Cyclist annual memberships?
And
C) How can Cyclistic use digital media to influence casual riders to become members?

2. Consider key stakeholders:
The marketing team,
Lily Moreno (director of marketing),
Cyclistic executive team

Deliverable: A clear statement of the business task.
Cyclistic, a bike-sharing company, is experiencing rapid success and wants to understand how to turn casual members into becoming annual members. Cyclistic operates in a competitive market in the Chicago metro area with other bike-sharing companies offering similar services. The objective of this business task is to identify trends and patterns in customer behavior, and to develop a marketing strategy that will increase annual memberships. The stakeholders for this business task include Cyclistic, Lily Moreno, Cyclistic marketing team, and executive team. Deliverables for this business task will include a description of all data sources used, documentation of any cleaning or manipulation of data, a summary of my analysis, supporting visualizations and key findings, and finally my top three recommendations based on my analysis.

Prapare

- Downloaded the data from the link provided.
- I took it upon good faith that the information provided was vetted properly and useful for the analyst. As this is part of the project.
- I downloaded 4 csv files with thousands of rows.
- I did a quick skim and noticed 1 file containing most of the info but missing key data along the gender lines. However, the missing data is not going to impact our analysis because our marketing will be inclusive and gender neutral; therefore, we do not need to know who is a female or male.

- Data is downloaded
- Identification is complete
- Determine how it needs to be sorted.

Deliverable: Downloaded the data from the provided link, sorted through the data and determined how to move forward. In this step I determined that I needed to rename the columns as well as reclassify the cells on the columns to match the rest.

PROCESS/ORGANIZING/CLEANING DATA

- After downloading the last year of data as per the instructions. I Chose SQL to clean and Organize The data was separated into quarters, downloaded the 4 previous quarters and noticed that one of them had a column that did not match the other 3, making sure that this 1 column matched was important to me because it was the usertype data which showed me who was a subscriber(member) and who was a customer(casual rider).
- After downloading the last year of data as per the instructions. I Chose SQL to clean and Organize The data was separated into quarters, downloaded the 4 previous quarters and noticed that one of them had a column that did not match the other 3, making sure that this 1 column matched was important to me because it was the usertype data which showed me who was a subscriber(member) and who was a customer(casual rider).
- Something that will help answer all 3 questions is which station is used by subscribers and customers the most and which ones are used the least. This will help us figure out which stations to offer promotional material through the app or signs encouraging customers to become subscribers.

- Created a new table with only the following columns: usertype, gender, birthyear, from_station_name, to_station_name, start_time, end_time, tripduration. This table combined data from Q2_2019 thru Q1_2020 during this propress I realize that it would be best to further clean the table. Documented the steps and saved a SQL query.
- I chose to work with R as well as the file was a way to work with a spreadsheet.
- Open clean data in a spreadsheet and add additional columns where I created one for the days of the week. And another one for the length of time between each station by subtracting the to_to time from end_time
- During my analysis phase i realize that i needed to further manipulate the data as some of the information i was using was not useful for my analysis. The birth year column needed to be converted to age. As using the year was not useful to answer the question i wanted to look into.

Three questions will guide the future marketing program:
1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?
3. How can Cyclistic use digital media to influence casual riders to become members?

The analysis: used spreadsheets and R.

- **I created a PIVOT table to see what days of the week were most frequent and least frequent for customers and subscribers. Monday thru Thursdays are the most used days of the week. With Tuesday and Wednesday being the highest. However, for customers it is Tuesday and Thursday. From this quick view the highest usage is Tuesday thru Thursdays for both. Which a quick assumption is that the usage is for commuting and work related usage.**

Findings:
Both users differ in that they do not use the same starting stations, nor they use the numbers of days. So if we are going to create marketing campaigns for casual riders the focus should be around wednesdays and sundays. As well as these top stations: Streeter Dr & Grand Ave, Lake Shore Dr & Monroe St & Shedd Aquarium. As well as exploring synergy opportunities with Shedd Aquarium.

Women are not using the service whether as casual or annual customers.

The majority of users are millennials, the marketing department has to find a way to reach out to Z'ers and prepare for Gen Alpha. As it is part of our mission to reduce greenhouse and work in conjunction with the earth. Utilize those avenues to promote our services.

While creating a viz, I discover high usage periods throughout the year. There were 4 peaks and they seem to correlate with the beginning of a new season as well as the beginning of the year. It is safe to make a prediction that taking actions during the windows of the beginning of Spring, Summer and Fall and the beginning of new year resolutions.

Recommendations:

Increase the number of usage by women by providing marketing made by women online influencers riding our bikes. Introducing a new product such as a standing bike for working women in skirts. Not as a woman but I could imagine being uncomfortable and degrading to ride a bike in skirts. And for business men in slacks or suits it is a lot more flexible to ride a bike.

The service is used by millennials as they are in their late 20s and mid 30s. Focus on reaching out to younger Gen by creating marketing around Climate change and the benefit for using our services. And explore becoming a Negative Carbon producer by partnering with companies with sustainable practices.

Turning our attention to creating the times when we see the most use, these dates coincide with the beginning of each season as well as the beginning of the year. We see at the start of Spring, Summer, Fall and January 1st and huge uptake in usage. This may be due to change of cooler or warmer weather as well as new year's resolutions. Take those times to market promotions for casual customers to become annual members.