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 do they use the same number of days. So, if we are going to create marketing campaigns for casual riders, the focus should be on Wednesdays and Sundays, as well as these top stations: Streeter Dr & Grand Ave, Lake Shore Dr & Monroe St & Shedd Aquarium. We could also explore synergy opportunities with Shedd Aquarium..

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

The majority of our users are millennials, so our marketing department needs to figure out how to connect with the younger Generation Z and prepare for Gen Alpha. This is in line with our mission to reduce greenhouse emissions and work with the planet, so we should leverage all available channels 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:

1st) To increase the number of women using our service, we propose the following strategies:

Launch a marketing campaign led by women and women influencers to promote our service and create a relatable message that resonates with women.

Invest in resources to identify the underlying reasons why men are more likely to use our service than women, and use the insights gained to create targeted solutions that address the barriers women face.

Test the hypothesis that comfort is an issue for women by examining factors such as clothing attire around office buildings, and explore ways to address these issues to make riding more comfortable and accessible for women.

2nd) The service is currently most popular among millennials, specifically those in their late 20s and mid-30s. In order to expand our user base, we should focus on reaching out to younger generations, such as Gen Z, by creating marketing campaigns centered around climate change and the benefits of using our services. Additionally, we should explore partnerships with companies that have sustainable practices to become a negative carbon producer and further promote our commitment to environmental responsibility.

Finally) We have noticed that usage of our service peaks at the start of each season and the beginning of the year. Specifically, we see a significant increase in usage at the start of Spring, Summer, Fall, and on January 1st. This could be attributed to the changing weather or new year's resolutions. We suggest taking advantage of these periods to promote annual memberships to our casual customers through targeted marketing campaigns.