Guiding Questions

* Where is your data located?
  + The data has been hosted on a website
  + [Location of Data](https://divvy-tripdata.s3.amazonaws.com/index.html)
* How is the data organized?
  + Data is organized into 12 separate CSV files where each CSV file represents the month for the year of 2022
  + Within each CSV file, the data is organized based upon these parameters
    - ride\_id
    - rideable\_type
    - started\_at
    - ended\_at
    - start\_station\_name
    - start\_station\_id
    - end\_station\_name
    - end\_station\_id
    - start\_lat
    - start\_lng
    - end\_lat
    - end\_lng
    - member\_casual
* Are there issues with bias or credibility in this data? Does your data [ROCCC](https://www.coursera.org/learn/data-preparation/lecture/lHirM/what-is-bad-data)?
  + In terms of bias, I do not see any issues because the data is only capturing the specs of each trip which are very objective such as the start time and end time.
  + Although start\_station\_name and end\_station\_name weren’t captured for certain records in August, we can find this info through the latitude and longitude which were captured
  + ROCCC
    - Reliable
      * The data is reliable because there isn’t any sampling bias
    - Original
      * Yes the data is original because it is coming straight from the source which is coming from Lyft Bikes and Scooters who operate the City of Chicago’s Divvy Bicycle Sharing Service. The City of Chicago mandates that the data be made public.
    - Comprehensive
      * Yes it captures all of the necessary parameters for each trip such as start location/end location/start time/end time
    - Current
      * Yes its data for the year of 2022
    - Cited
      * Yes since its public data that has been provided by Motivate International Inc.
* How are you addressing licensing, privacy, security, and accessibility?
  + Data has been provided by Motivate International Inc
  + The data is public
  + Data can be used for ‘lawful’ purposes only
* How did you verify the data’s integrity?
  + I verified the data’s integrity by punching in the latitude and longitude for a couple of records to see if they match the station name
* How does it help you answer your question?
  + The data should help me answer the question because it provides us with the member\_type for each trip - through the duration of each trip and its respective locations, we can start to find any trends that causes a causal rider to be a causal rider and the steps we can take to entice them to sign up for an annual membership
* Are there any problems with the data?
  + The only problem with the data was for the months of June, August, and September where certain stations weren’t registered for certain trips