**Part 1: Ask the right questions Professionals with an analytics mindset need to develop a strong understanding of the economic and institutional factors that influence the data they are analyzing. They allow that understanding to guide their analysis and interpretation of the data. In this section of the tutorial, you will consider what insights you are seeking to find from Twitter data.**

**1. Think about Twitter and what appears on the platform. What interests you about Twitter?** When I look on Twitter, I see advertisements, pictures, and other interesting pictures that family, friends, and other followers have posted. These images and short messages are meant to share important or interesting information with viewers all around the word. What interests me most about Twitter is the news I can get at the touch of a button. The app is simple to use and great for all ages.

**2. What questions could you ask and “answer” using Twitter data? List at least 4.**

- (Pick a Random Celebrity) What percent of their followers are females? Males?

- What state has the most twitter accounts?

- What topic draws the most attention?

- What topic is the most talked about in California?

**3. Choose one of the questions from #2 and form a hypothesis. How specific is your question? Will you be able to “answer” it with Tweet data?** I picked question one. I believe this question is pretty straight forward. My hypothesis would be that male celebrities, especially younger celebrities like Justin Bieber, will have a higher percentage of female followers. I believe I will be able to answer this question just by seeing the celebrity’s profile picture.

**4. After identifying your question and forming your hypothesis identify a few hashtags you will use to mine tweets.** If I stayed with Justin Bieber, these would be the hashtags I would use. #pop #JustinBieber #Belieber #JustinBieberfanclub. I also believe that most of the fan clubs that follow Justin Bieber will most likely be females that follow him on Twitter.

**5. Modify the example Jupyter Notebook provided in Module 3 with your new file names and hashtag search terms. Run the file and open the resulting CSV file.**

**6. Examine the tweets and sentiment/polarity numbers. Did the sentiment analyzer accurately capture the sentiment of each tweet?**

**Part 2: NLP with tweets**

**1. Make a copy of your tweet CSV file and rename it (example: “covid\_tweets\_text.csv” if you were searching for tweets related to covid).**

**2. Download and save the tutorial Jupyter Notebook template file in the same folder you have been using for your other Jupyter Notebooks.**

**3. Follow the steps in the video and in the template.**

**4. What have you learned from this exercise? Were you able to answer your question from part 1? If not, what information would you need?**