# Provide a brief explanation/overview of your capstone project.

My project uses data gathered from the Twitter developer API to analyze the effectiveness of The Brewers Association (BA) social media efforts. It includes **two parts**, the first includes an analysis of historically published content. This analysis will be used to answer multiple questions similar to the following:

* What type(s) of content engages BA audiences the most (links, photos, videos, GIFs, etc.)?
* Are there words that are more appealing, certain images (color palette themes, lifestyle, technical, etc.)?
* Do specific days of the week, times, frequency, etc. matter?
* Do we see overlap from the various BA brands (Brewers Association, American Homebrewers Association and CraftBeer.com)?
* What are these brands similarities and differences?

The second part will focus on identifying brand advocates for BA. Brand advocates will be defined as users that help to spread the company’s message to the most unique followers. This portion of the project will lean on algorithms developed from academic research. The final algorithm developed for use in this second part of the project will be depended on what seems to fit best to BA’s data.

This project will be a used as a use case for BA to determine if an expanded project including all of their social media platforms could benefit from a similar analysis.

# List some goals, objectives or outcomes associated with the project.

The primary objectives for this project include the following:

1. Gather BA relevant and related data from the Twitter developer API
2. Use this data to analysis and create a dashboard/s answering the above questions as part 1 of the project description
3. Research and Implement an algorithm to identify individuals that are accredited with a high level of influence with BA related content.
4. The overall outcome of this project is to maintain or improve the effectiveness of BA’s social media campaigns regardless of the monetization changes of a given social media platform.

# Briefly describe how you came to choose this particular project.

BA has been a company that I have been interested in for the past four or five years. I have been interested in finding a way to interact with the company in some level and I figured this course could be an interesting start. Over the years of following BA I came to admire one of their employee’s [blog](https://www.brewersassociation.org/category/insights/), Bart Watson. He has put out numerus data related posts and in the past and I figured he would be a great contact to reach out to in the organization. After working through a few project ideas with Bart, we decided that this project seemed to fit the best with the requirements of this course, my own interests and most importantly it is something of use for BA.

# Briefly discuss some concepts, methods or skills involved that were obtained in courses you took in the MSDS program.

This project builds on concepts outlined in DS 745 (Visualization and Unstructured Data Analysis). Specifically it addresses creating effective visuals to maximize readability, comprehension, and understanding of complex datasets. Additional, it utilize techniques and methods used for analyzing semi-structured data, in the form of JSON. Finally, the project also addresses social network analysis.

# Provide insight into any challenges or difficulties you have encountered in undertaking your project.

One of the biggest difficulties that I have encountered so far is in the process of collecting the data. But this is not because of a lack of understand of the API or programming. It stems from the rate limits that Twitter imposes on developers (at least at the free use level). Depending on the type of call to the API different rate limits are imposed. Two calls in specific have been somewhat time consuming, the first being the user request, and the follower request (in essence they are the same call, looking up user profiles), these requests are limited to 15 calls per 15 mins. Because of this I have had to alter the second part of this project to not be holistic to all retweeters of BA and only collect a subset.

# Based on your experiences to this point in time, what, if anything, would you do differently?

To be honest I do wish that I would have looked into the API documentation a little bit earlier that the start of the project. If I would have done this it would have been more apparent that it would take quite a bit of time to collect the entirety of data that would be needed to be able to identify fully correct answers and advocates. But given that this project is truly a proof of concept one of the outcomes likely will be to recommend paying for a provider of this data or going directly to the paid API at the start.