

DS 785 Data Science Capstone

Capstone Activities Report

Name: Spencer Swartz

Project Title: Social Media Content Optimization and Data Driven Brand Advocate Identification

Activity Report #2

Agenda/goals:

At the completion of this activity period to remain on track for the project it is expected that I was able begin applying visual analytics on the data collected and shaped in the previous activity period.

Contacts Made/Method of Contact:

I have contacted Bart Watson (BA) via email with a similar update notifying him of the progress made in the data analysis period on 3/5/18.

Resources and Investigation Methods:

This portion of the project revolved around a few products from Tableau, first is there signature product Tableau Desktop (a drag and drop data visualization platform), and the second is Tableau Public (a hosting/publishing server for browser based interactive dashboards).

Tableau Desktop is used as the development platform for the investigative visual analytics that was performed in this activity period. It will also be the development platform for the final dashboard(s) that will be created in the next activity period.

Tableau Public is used as the portal for the Brewers Association to be able to interact with the investigative visuals created during this activity period, and will be where the team access the final dashboard(s).

Progress:

The project progress is on tract at the moment, I have been able to investigate all 5 of the initial questions outlined in the project proposal as well as below:

- What type(s) of content engages BA audiences the most?
- Are there words that are more appealing, certain images?
- 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?

Although I have investigated all of these questions I have not yet decided which content will be of most use for the customer (BA). For the most part I am using Retweets and Favorites as the definition of success for individual tweets. The following are screenshots of some of the views that have been created:

of combined hashtags

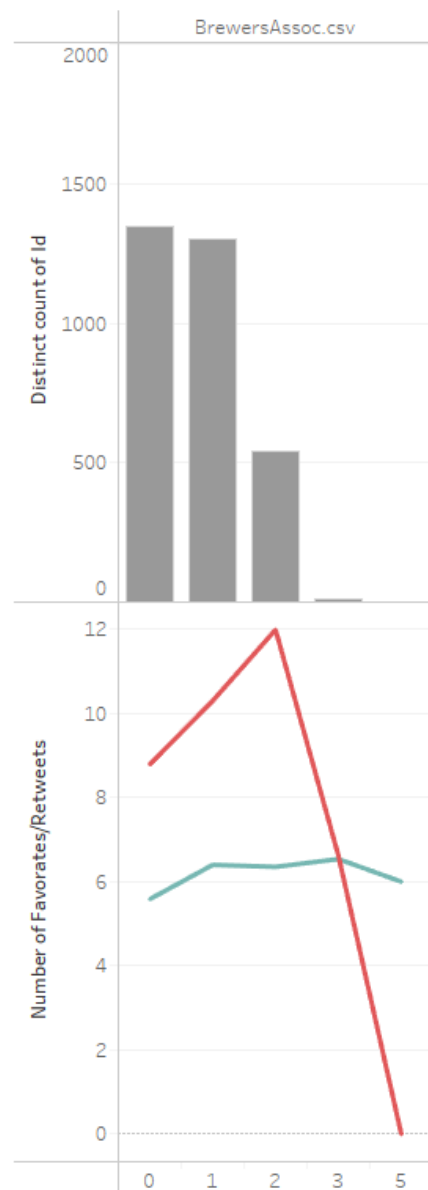


Figure 1 : Does creating mutiple hashtags help with tweet success?

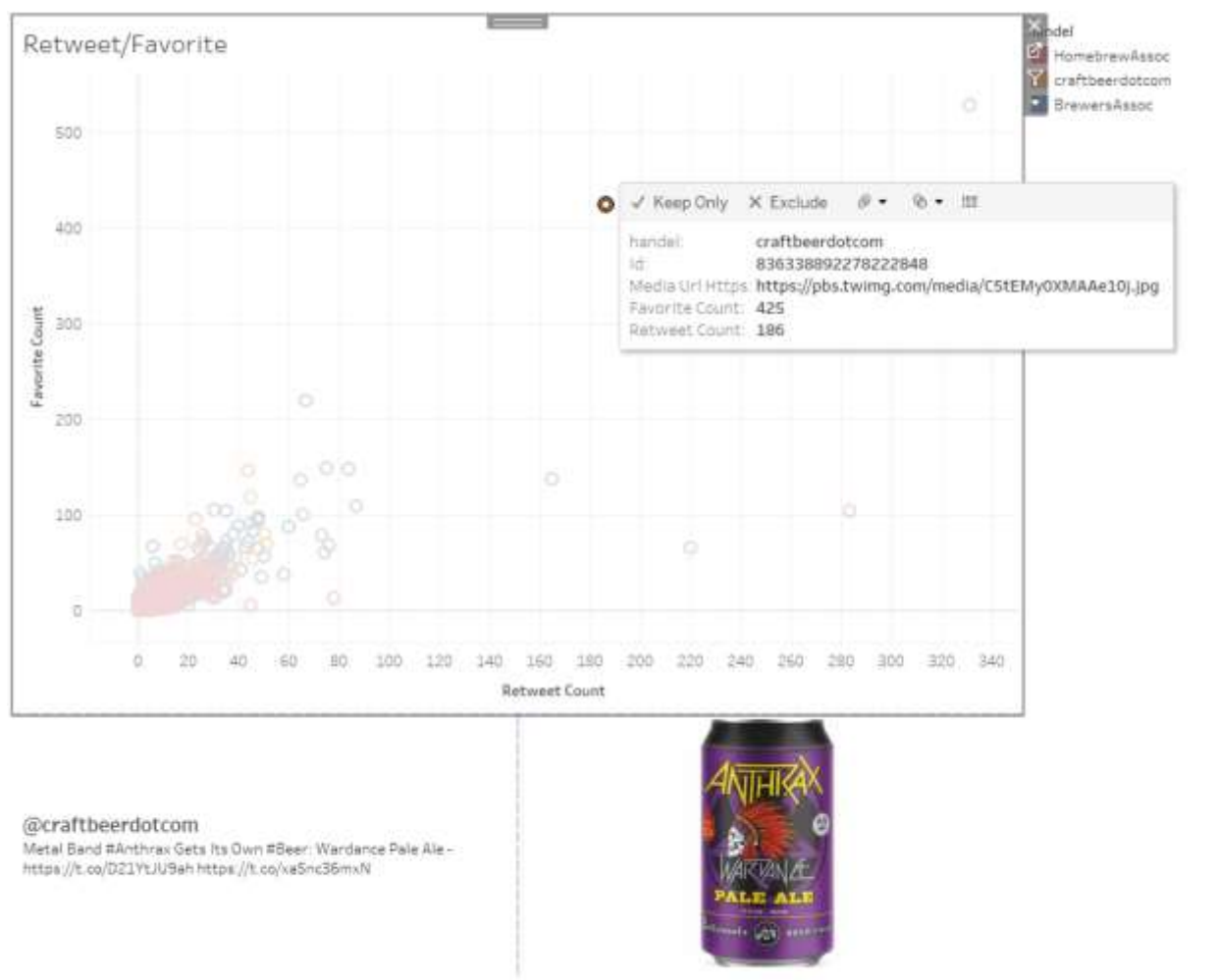


Figure 2: Recreating tweets and scatter plot based on Favorites and Retweets



Figure 3: Does Time of Day have an affect on success?

Achievements:

Once I had brought the data into Tableau it was quite obvious that some of the python functions created in the last activity period did not work as intended. I was able to identify the problem and fix it in a timely manner, I am now confident that the data collection process was a success. I have gathered some ideas into what I believe the final product should be for this part of the project and should be able to complete it on time.

Questions:

I currently have no questions for the instructor at this time, moving forward as stated in the last reporting period, I need to be aware of the time that may be required to research practices in identify brand advocates and possibly the time that will be needed to collect additional data. This is especially important as some request to the API take more time than others given the rate limits imposed.

Next step:

The next week will be focused finishing the visual modeling of the data within Tableau. A final product should be provided to BA within a week's time.

The second week of the next period will be used to identify research articles specific to identifying brand advocates as stated above.