DS 785 - Capstone Course Project Proposal

Project Description

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

- 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 of 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. An example of how these individuals could be identified is show in the paper *A Data-Based Approach to Social Influence Maximization* (Goyal, A., et. al.). 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.

Rationale

BA has realized a need for this project once the announcement from Facebook outlining changes to the algorithm used within their news feed. The quick overview of this announcement states that the company will introduce less paid content from media companies into users feed. BA like many other companies have used sponsored content on most of their social media platforms to create a wider reach with their messages, because of this change and the possibility of more like them on other social media platforms they likely could see a loss in the organic reach of their content. The overarching reason for this project is to counter act this problem by supplementing it by optimizing the content published and targeting specific users who are very likely to share/retweet BA's message.

Proposed Project Title

Social Media Content Optimization and Data Driven Brand Advocate Identification

Proposed Project Purpose

Client-based project: include the client organization description and client contact (name and position) and purpose

The purpose of this project is to server BA (a non-profit organization based in Colorado), and in turn the micro-breweries in which it serves, with the ability to better communicate with its social community.

To be able to do this it is important to understand the client of this project. BA mission state that it is focused on promoting and protecting American craft brewers, their beers and the community of brewing enthusiasts.

The primary contacts for this project includes Bart Watson - Chief Economist, and Lara Betthauser – Social Media Manager from BA.

Project Objectives

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.

Project Timeline and Activities.

Description	Week
• Use Twitter API and Tweepy (python package) to gather BA data related to each of BA's Brands	2/5 – 2/11
• Convert the JSON files from the API into a useable relational format	2/12 – 2/18
Activity Update 1 – (Spencer and Bart/Lara)	Due 2/19
• Begin analysis of data and interpreting answers for the above questions ○ Begin development of Tableau dashboard/s	2/19 – 2/25
• Continued analysis and development	2/26 – 3/4
Activity Update 2 – (Spencer and Bart/Lara)	Due 3/5
• Complete analysis and development of dashboard/s *Expert Interview – Bart Watson, BA (due 3/2)*	3/5 – 3/11
• Create write-up related to analysis and development • Research academic articles related to brand advocates within social media	3/12 – 3/18
Activity Update 3 – (Spencer and Bart/Lara)	Due 3/19
• Identify "advocates" model • Gather additional data from API if needed and convert	3/19 – 3/25
• Begin analysis and fitting of "advocates" model *Peer Review (due 3/23 – 3/27)*	3/26 – 4/1
• Continue analysis and fitting *Expert Interview – Yuchi Huang, ACT (due 4/6)*	4/2 – 4/8
Activity Update 4 – (Spencer and Bart/Lara)	Due 4/9
• Complete analysis and fitting of the "advocates" model	4/9 – 4/15
• Create write-up related to analysis and fitting	4/16 – 4/22
Activity Update 5 – (Spencer and Bart/Lara)	Due 4/23
• Allow time for delays that may have occurred • Work on final documentation for entirety of project	4/23 – 4/29

• Work on final documentation for entirety of project	4/29 – 5/6
Capstone Due (4/11)	4/5 – 5/11

Interviewees.

The two interviews that will be conducted in part of this project will include Bart Watson of BA, and Yuchi Huang of ACT.

Bart is the Chief Economist for the Brewers Association and is quite familiar with the data needs for BA and its members. Most if his work is around the economics of micro brewing, he reports his findings in data for on his public blog that can be found on BA's website. He will be a great asset for questions related to the project and to better understand what Data Science is like in the brewing industry and as a part of a non-profit organization.

Yuchi is a Senior Research Scientist in ACTNext, the business innovation department with ACT, his work consists of spearheading research in the use of machine learning/deep learning models in education research. His current areas of interest include automatic generation of multimodal educational resources, multimodal analytics for measurement of complex skills and competencies (such as communication ability and collaborative problem solving), and automatic photorealistic avatars generation for human-agent interaction. He will also be a great asset in understanding Data Science in the non-profit sector.

Application of Data Science Concepts

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

Description of Final Document

The final document will be a detailed report on the entirety of the project. Specifically, it will include:

- the steps taken to gather the data from the Twitter API,
- the manipulation and transformation of the JSON files returned,
- the work done in creating visualizations intended to answer the above questions,
- figures of the visualization, and the corresponding insight identified,
- a summary of the research conducted on identifying brand advocates,
- The analysis and application of the applied method identified within the research and the list of advocates identified.

Additionally, the final documents will also include a reference section and other supporting information related to the overall project.