Julianne Delia

COMM 420 – Taming Big Data

Professor Shahin

May 1, 2019

*The Bachelor*: Happily Never After?

**Abstract**

The purpose of this study was observe the viewer reaction and general sentiments of the ABC show, *The Bachelor*. This study was conducted over the span of three weeks during season 23 of *The Bachelor*. The hypothesis is that the viewer reaction is an overall negative sentiment. Tweets were collected in order to conduct a sentiment analysis of the audience members watching the show while it aired. This paper will discuss the methods involved in order to conduct the analysis. The results will also show that the original hypothesis was wrong in regards to the overall viewer sentiment.

**Literature Review**

The art of second screening has become wildly popular after the evolution of social media. Watching television with a second screen, the other screen usually being a smart phone or tablet, is the viewer’s attempt of interacting with the outside world from the comfort of their own home. The “complementary app may allow the viewer to interact with a television program in a different way…the second screen phenomenon represents an attempt to make TV more interactive for viewers, and help promote social buzz around specific programs” (“What”). Twitter is one of the most popular social media platforms for audience members to interact with one another. When it comes to reality TV shows, the stars are often social media influencers (Ilchi). This allows the audience members to interact with the stars themselves. Since devoted TV watchers love to voice their opinions on social media, I want to look at of those audience members, who is more likely to voice their opinions on social media; are the positive sentiments more often voiced than the negative sentiments? This report will look at different literature that explores second screening and Twitter sentiment analysis.

In their article “Who is on Your Sofa? TV Audience Communities and Second Screening Social Networks,” Mark Doughty, Duncan Rowland, and Shaun Lawson explore the different types of audience members in the second screening world. As mentioned earlier, Twitter is one of the most popular social networks for audience members to gather and interact. The development of Twitter hashtags that are show-specific “allow the audience to ‘tune’ their data streams to one focused upon the particular broad cast being viewed” (Doughty 80). These hashtags are either created by the devoted fans or are created by the show’s producers and pop up on the screen. Sometimes, the hashtags that become more popular than the rest have emojis attached to the words. This reiterates that the emoji hashtag is the unofficially official hashtag that fans of the show should use while watching said show. The show I’m following, *The Bachelor*, usually has a rose or a man with a suit emoji next to the hashtag #TheBachelor.

All of this action and interaction allows users to connect with one another from opposite ends of the globe. Doughty, Rowland, and Lawson point out that “the connections between audience members afforded by the retweet (RT) and mention structures within twitter allows us to view the information flows and links within a distributed TV audience. Opinions, observations and personal status updates related to the TV show are shared and connections made” (Doughty 81). The act of second screening has evolved and will continue to evolve as social media evolves. Fans interact with other fans, actors, producers, and writers. This makes the fans feel more connected to the actors, the characters, and the show. Since social allows fans to connect with actors and writers, the fans are more likely to voice their feelings and opinions. They think that if they show how upset or happy they are about something that happened on the show, then the actors/writers/producers will listen and change something. It lets the fans feel like they are a part of the creative process. “Connections with celebrities and stars of the shows were made through mentions and retweeted messages” (Doughty et al., 85). This observation made by Doughty, Rowland, and Lawson further proves that the likeliness of a connection to the show’s celebrities increases the likeliness of fans voicing their thoughts.

In her 2016 Masters Media Studies thesis, Alessia Borgomastro explores how online participation has affected the reality television watching experience. Borgomastro points out that “mainstream TV programs operate official Twitter accounts, and viewers are encouraged to follow and interact with them…these reality programs take the concept of audience feedback making it central to the organization of the program articulating it across social media” (Borgomastro 30). In this project’s example, *The Bachelor* has its own account for the show and the title character himself also has an account. On top of that, almost every contestant has an account, and they will often live-tweet what’s happening during the airing of each episode. When the show account, the bachelor himself, and the contestants all live-tweet during each episode, this invites the audience to feel connected with the faces they see on their television. This also allows them to feel more empathetic towards the stars because these people can be contacted online rather than the audience just seeing them as characters in a show. It allows people to express their emotions on a public platform.

Siqi Zhao, Lin Zhong, Jehan Wickramasuriya, and Venu Vasudevan conducted a sentiment analysis of an NFL game. They collected tweets in real-time and observed the occurrence of words indicating both positive and negative sentiments (Zhao 4). They found that extracting tweets real-time was the most effective way to conduct a sentiment analysis of a show or event that airs at a scheduled time on television. They were also able to distinctly acknowledge which words held positive sentiments and which words held negative sentiments. Their project was done on a sporting event that was happening live, which differs from the show that I am analyzing. The show I am doing my report on was recorded a few months before the air date. However, both a live sporting event and *The Bachelor* have a set time they are going to air on national television.

In conclusion, the art of second screening has allowed data analysts to conduct sentiment analysis of television segments. It is a natural human reaction to have a strong emotional connection to the television shows in which we get invested. Nowadays, social media makes it easy for people to connect with each other from across the globe without leaving their couch.

**History**

*The Bachelor* was created by Mike Fleiss and the first season aired in 2002 (“About”). The show revolves around a single bachelor and the bachelor is, on average, between the ages of 25 and 35. There are 25 female contestants who vie for the title character’s heart. On average, the women are between the ages of 24 and 32 (Yuan). Throughout the season, contestants are eliminated at the Rose Ceremony. The Rose Ceremony is when the bachelor hands out roses to the women he either already has a connection with, or would like to pursue further. Women who do not receive a rose are sent home and their time on the show ends. Eventually, the bachelor picks one woman to become his future bride.

Each season roughly follows the same format: the bachelor goes on dates every week. Some are one-on-one dates, which is when it’s just the bachelor and one woman. Another date is what’s called the “group date.” This is where the bachelor takes a larger group of women; the group dates have a minimum of three women. The last kind of date is the two-on-one date. That is when it is the bachelor and only two women. With this kind of date, the bachelor must eliminate at least one of the women. After a week of about three dates, there is a rose ceremony where the bachelor eliminates women he no longer feels a connection with.

During the course of the season, the cast travels to various exotic locations around the world. The entire season is usually nine weeks and during the last week, the bachelor chooses one woman. The intent is for the bachelor to propose to the last woman. However, that does not always happen. There are several unpredictable variables that can cause a twist during the season that deviates from the intended format. Those variables include when a candidate who was eliminated returns to the show to plead her case, when the bachelor distributes more roses or fewer roses than planned, when the bachelor eliminates a woman outside of the normal elimination process, and when the bachelor does not propose. There have been a total of 23 seasons, 15 proposals, and three marriages. Only one bachelor married the woman he chose in the season finale. The other two broke up with the woman they chose and then married the runner up.

There have been a total of four spin off shows: *The Bachelorette, Bachelor Pad, Bachelor in Paradise,* and *The Bachelor Winter Games*. On average, couples from *The Bachelorette* are more successful than couples from *The Bachelor* (Yuan).

**Method**

The stigma of *The Bachelor* franchise is usually one of two things: the system works and everyone should try it or the system is totally useless and people only do it to get their 15 minutes of fame. Either way, those people are sure to voice their opinions on social media and I wanted to analyze their thoughts and feelings. I conducted a sentiment analysis of tweets using the hashtag #TheBachelor. All of the tweets were collected using Netlytic. I only used the #TheBachelor when searching for datasets and each dataset was named based on the time and date at which the data was collected.

Fans of *The Bachelor* franchise (also known as “Bachelor Nation”) are encouraged to tweet along using the aforementioned hashtag. Along with devoted fans, the title character and some of his contestants live-tweeted the event. Also, previous bachelors and bachelorettes (also known as “Bachelor Alum”) tweeted along with the viewers at home (Beck). Therefore, there are an abundance of people using #TheBachelor, especially during the times that the show aired.

I collected 1,000 tweets on March 6, 2019, which was exactly one week before the two-night season finale. The season finale was a two-night special, so I collected tweets every three to five minutes during both nights of the season finale. On March 11, 2019 (which was night one), I collected 42,000 tweets. On March 12, 2019 (night two), I collected 45,000 tweets. All of the tweets on night one and night two were collected between 7pm and 11pm. I then collected 1,000 tweets on March 20, 2019, which was exactly one week after the two-night season finale. This entire process was completed over the course of three weeks in the month of March and a total of 89,000 tweets were collected.

I conducted a sentiment analysis using R to see if the majority of people live-tweeting the show have negative sentiments or positive sentiments. This analysis used the R package, “syuzhet.” I created bar charts in order to visualize the data. A problem that was considered before the analysis was conducted is if a tweet was misinterpreted by the program. For example, if someone tweeted “She does not look happy,” the program might see the word ‘happy’ and assume it’s a positive sentiment.

There are different groups of people that should care about this data. First of all, fans should want to know what other fans are saying about the show. As someone who follows the show, I like to see what other people say and if it lines up with my feelings. Therefore, it would be beneficial for fans to connect with one another if their opinions line up. Contestants of the show should also care about this data because tweeters are commenting on their actions on screen. Lastly, producers of the show should care the most about this data. It would be beneficial for them to what aspects of the show are starting conversations.

My hypothesis was that there would be an overwhelming amount of negative sentiments towards this season finale due to the events that led up to it. As someone who followed the show for the entire season, the Twitter users in my social media bubble appeared upset about the events that took place during the season (Brodeur).

**Results**

Figure

First, a bar chart was created for the first set of 1,000 tweets that was collected on March 6, 2019. This was done in order to gage the viewer’s sentiment exactly one week before the two-night season finale. I wanted to see how people were feeling before the finale began.

Figure 1 will show that generally, “positive” was the most overwhelming sentiment. After that, “anticipation” was really high. This makes sense because people were most likely anxious to see how everything would end in the following week. “Joy” was also relatively high. This makes sense because after sifting through the tweets of that dataset, people were “excited” to see what happened next.

 During that week of the show, a highly anticipated event finally happened. Throughout the season, the show teased the bachelor, Colton, jumping an 8-foot-high fence. Viewers referred to it as “the fence jump.” This fence jump was the most talked about thing throughout the duration of Colton’s season (Ivie). Since this event finally took place during the episode that aired on March 5, 2019, people were still talking about it on Twitter when the March 6 dataset was collected. The word “jump” was in several tweets, and that word was dubbed “positive” sentiments. However, the fence jump relates to something sad and negative that happened during the season: Colton jumped the fence in order to get away from the producers because he was so upset. Therefore, any tweet with the word “jump” in it was counted as a positive tweet but it actually relates to something negative.

Figure

Anticipation spiked after night one. Figure 2 shows that after “positive,” “anticipation” is the next highest sentiment. After that, “negative” and “trust” are high as well, which is incredibly contradictory.

It makes complete sense that “anticipation” is very high after night one because there was still one more episode to watch before the conclusion of this season. Therefore, people were very anxious to see how the season ended. The most common words that were used during night one were “jump”, “love”, “excited”, and “engaged.” Since the algorithm reads all four of those words as positive sentiments, it makes sense that positive is still the most overwhelming sentiment.

People were still talking about the fence jump during night one, hence the large number of tweets mentioning the word “jump.” Though the words “love”, “excited”, and “engaged” were used to express positive sentiments, they too were often used in a negative light. Cassie, the girl Colton confessed his love to, broke up with Colton and said she wasn’t ready for marriage. Colton did not want to take “no” for an answer, so he continued to chase her. Therefore, some Twitter users used those aforementioned “positive” words in a negative light. For example, @RealKDPFan4Eva tweeted “Never excited for someone to look like a fool but Colton earned it this season. Sorry, but it was clear as day that Cassie wasn’t feeling it and going after her after she has said she’s not even sure she loves you... is just.... silly. So meh, whatever. #TheBachelor.” This user used the word “excited” and the algorithm counted it as a positive sentiment, but this user is clearly unhappy with how the night unfolded.

Figure

Night two ended the season with a generally high positive sentiment. Figure 3 will show that “positive” continued to remain the highest sentiment according to the sentiment analysis. “Anticipation” was also still very high after night two. At first, that didn’t make sense because the season was over and there was nothing more to watch. However, the next title character of *The Bachelorette* was announced at the end of night two, so it was then clear why “anticipation” was so high. People were interested to see how the next bachelorette’s season would unfold.

As for the conclusion of Colton’s season, he left the season with the girl he wanted all along. They did not get engaged, but they decided to just continue dating. Twitter users had mixed feelings about that. One user, @realityspice1, did not like Colton and Cassie as a couple. This user wrote, “I’m good with never hearing the names Colton and Cassie again #ThankUNext #thebachelor #thebachelorfinale.” On the other hand, @emztho wrote “cassie and colton look so cute together! #TheBachelor #TheBachelorFinale.” Both sides of the spectrum voiced their opinions about either liking or disliking Colton and Cassie as a couple.

Figure

The tweets from one week after were generally positive. Figure 4 shows that “positive” remained the highest sentiment throughout this study. Most of the negative sentiments decreased. “Anger”, “disgust”, and “fear” were all at the bottom. “Anticipation” was not as high as it was during the first three datasets, but it still remained up in the higher numbers, most likely because people are anticipating what is going to happen during the next season of *The Bachelorette.*

For the most part, it seemed like people began to accept Colton’s current dating life, which is why there was a spike in positive emotions. Twitter user @grapeisnotdead wrote, “they’re actually the cutest happiest humans oh my goodness” which a picture of Colton and Cassie attached to the tweet. Another user, @nicolemdak retweeted a tweet from the official @BachelorABC account that said, “Love is worth the wait and we love these two” with a picture of Colton and Cassie. However, there are still people who are upset with how things turned out. User @amandalo926 wrote “I’m a week late on #TheBachelor. I feel like Colton decided he loved Cassie because she rejected him. It’s that wanting what you can’t have thing... This is dumb.”

Something that I noticed while sifting through the tweets for March 20 was that people were tweeting as if they were watching the finale live. This leads me to believe that those people were not able to watch the season finale live and were making it up by watching it one week later.

Figure

 I wanted to compare the sentiments from March 6 and March 20 to see how people’s sentiments towards the show changed before and after the season finale aired. Figure 5 shows that positive emotions increased and negative emotions decreased. People seemed angrier and more fearful one week before the season finale. Something that I found very interesting is the “sadness” sentiment. “Sadness” was the only negative emotion that was higher one week after the season finale. All of the other negative emotions dropped. This is most likely because people were still watching the season finale a week after it aired.

Figure

I also compared the sentiments from night one and night two of the season finale in order to see how people’s feelings changed over the course of the two-night whirlwind. Figure 6 shows that the sentiments remained relatively the same throughout the two-night special. Something that did surprise me was the fact that “anticipation” was higher after the season finale was over. I was also surprised that “trust” went down just a bit from night one to night two. Nevertheless, “positive” still remained the highest sentiment over the course of all four datasets.

**Discussion**

After carefully looking at the sentiment analysis, bar chart comparisons, and sifting through the tweets, I have concluded that my hypothesis was wrong. The six bar graphs proved that “positive” remained the overwhelming emotion throughout the course of this study. This proves that the users I’ve included in my social media bubble have negative feelings towards the show, but that does not mean all Twitter users feel the same. This study showed that humans who use social media are prone to fall into the so-called “social media bubble” that prevents us from seeing and experiencing thoughts and opinions outside of that bubble.

Since I was able to collected over 89,000 tweets from three weeks, this also shows that second screening is very popular. The fact that I was able to collect 45,000 tweets using #TheBachelor in just two hours shows that fans of *The Bachelor* enjoy the second screening experience. A lot of those tweets were also re-tweets. Therefore, people are re-tweeting one another, thus adding to the second screening experience.

I do want to explore a few errors to take in consideration while viewing this data. First of all, as mentioned earlier, the algorithm misread several tweets due to the sarcastic tone and/or misuse of positive words. The fence jump was one of the most talked about scenes from season 23 and it is something that correlates with something negative. However, the sentiment analysis read tweets with the word “jump” as a positive sentiment. More often than note, twitter users using the word “jump” were referring to something negative. They were either joking about the jump itself, or they were sad that Colton reached the point of sadness that caused such a dramatic event.

I also pointed out a tweet that used the word “excited” negatively. The user was excited for the show to be over, which means they had a negative sentiment about the show itself. Other words that were misinterpreted were things like “love”, “engaged/engagement”, and “marriage/marry/married.” The algorithm read all of these words as something positive, but users often used them to refer to something negative. This was a common error that occurred, and therefore could have skewed the results.

If the algorithm was able to understand the misuse of the word “jump” or any of the other aforementioned words, I think the “positive” bar on all the bar graphs would look a lot different. Granted, there were times when all of those words were actually referring to something positive, but not nearly as many as the algorithm interpreted.

There are a few changes I would make if I were to conduct this experiment again. First of all, I would make sure to collect an equal amount of data from both nights one and two. I had 3,000 extra tweets from night two, which could have potentially skewed the data. Secondly, while collecting the data from nights one and two, I would make sure to have a set timer to collect the tweets at the same exact time each night. I was collecting tweets while watching the show and scrolling through Twitter, so sometimes there would be seven minutes between datasets and sometime the would be two minutes between datasets. This is probably why I did not have an equal amount of tweets from nights one and two. If I were to do this again, I would make sure that every dataset was collected at the same time for both nights. Lastly, I wish I started this experiment when the season started. It would have been interesting to see how the sentiments changed as dramatic events unfolded during the season.

In conclusion, people genuinely enjoyed this season of *The Bachelor*. Clearly, this show creates a lot of conversation on Twitter, which enhances the show’s growing popularity. If I were to do this experiment again, I would make a few revisions, but overall follow the same format of collecting tweets and comparing them to each other. Lastly, these algorithms are definitely helpful when conducting an experiment like such, but we, as humans, cannot solely rely on it to do all of the work. It is hard enough for humans to read tone and sarcasm online. Therefore, we cannot expect machines to be any better at it.

Works Cited

“About The Bachelor TV Show Series.” *ABC*, abc.go.com/shows/the-bachelor/about-the-show.

Beck, Lia. “Follow These 'Bachelor' Nation Alums For The Best Insider Takes On Colton's Season.” *Bustle*, Bustle, 4 Feb. 2019, www.bustle.com/p/12-bachelor-alums-you-should-follow-on-twitter-for-their-witty-episode-commentary-15867684.

Borgomastro, Alessia. “Social TV: Second Screen and Audience Participation.” *Erasmus School*

*of History, Culture and Communication*, 2016.

Brodeur, Michael Andor. “Are You Living in a Social-Media Bubble?” *BostonGlobe.com*, The Boston Globe, 5 Dec. 2016, www2.bostonglobe.com/arts/2016/12/04/social-media-where-you-find-information-diversity/u7isCbOsiHSvBhxIyHGFdO/story.html.

Doughty, Mark, et al. “Who Is on Your Sofa? TV Audience Communities and Second Screening

Social Networks.” *EuroITV '12 Proceedings of the 10th European Conference on Interactive TV and Video*, 4 July 2012, pp. 79–86., doi:10.1145/2325616.

Ilchi, Layla. “From Reality TV Stars to Influencers: How 'The Bachelor' Contestants Sell Products.” *WWD*, WWD, 22 Jan. 2019, wwd.com/beauty-industry-news/beauty-features/bachelor-contestants-influencers-1202969257/.

Ivie, Devon. “Why The Bachelor's Colton Underwood Jumped That Fence.” *Vulture*, Vulture, 5 Mar. 2019, www.vulture.com/2019/03/the-bachelor-colton-cassie-fence-jump.html.

“What Is a Second Screen? - Definition from Techopedia.” *Techopedia.com*,

[www.techopedia.com/definition/29212/second-screen](http://www.techopedia.com/definition/29212/second-screen).

Yuan, Jada. “A History of The Bachelor, by the People Who Lived It.” *The Cut*, 4 Jan. 2016, www.thecut.com/2015/12/history-of-the-bachelor-c-v-r.html.

Zhao, Siqi, et al. “Analyzing Twitter for Social TV: Sentiment Extraction for Sports.” *CiteSeerX*,

Solr, 2011, citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.414.6401.