2023

STEM and Pre-Health Students' Diminishing Online Conversation Volume Regarding Majors and Careers Within the Field: Why Higher Education Should Care

Susan Roh
South Dakota State University, nasusbackward@gmail.com

STEM AND PRE-HEALTH STUDENTS' DIMINISHING ONLINE CONVERSATION VOLUME REGARDING MAJORS AND CAREERS WITHIN THE FIELD: WHY HIGHER EDUCATION SHOULD CARE

by

Susan Roh

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Biological Sciences

Specialization in Microbiology

South Dakota State University

2023

THESIS ACCEPTANCE PAGE

Susan Roh

This thesis is approved as a creditable and independent investigation by a candidate for the master's degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Greg Heiberger
Advisor Date

Radhey Kaushik
Department Head Date

Nicole Lounsbery, PhD
Director, Graduate School Date

ACKNOWLEDGMENTS

Thank you to all the professors, faculty, staff, family, and friends that helped to make this achievement possible. Thank you to South Dakota State University for supporting this research and to our partners at Campus Sonar for their leadership, expertise, and collaboration.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

STEM AND PRE-HEALTH STUDENTS' DIMINISHING ONLINE CONVERSATION VOLUME REGARDING MAJORS AND CAREERS WITHIN THE FIELD: WHY HIGHER EDUCATION SHOULD CARE

SUSAN ROH

2023

The purpose of this study was to determine the STEM and pre-health online mention volume change regarding STEM and pre-health majors, careers, and universities, with the data collection period encompassing August of 2018 through June of 2021. We evaluated if the mention sentiment shifted between academic year periods. The data gathered for this study was procured in collaboration with Campus Sonar and the social-listening tool of Brandwatch, which allowed for specifications of the data collected. Kruskal-Wallis tests were used to determine the means of the majors, careers, and universities mentions and mean sentiment volumes. Changes in not only online conversation volume but sentiment associated with each STEM and pre-health mention signals the utility of social listening by university leaders to understand prospective student career interest more rapidly in the current and future highly dynamic world.

1. Introduction

Social listening is an emerging technique that broadens the ability for online conversation tracking, with uses postulated within organizational communication, including university admissions. Social listening encompasses social media and online communication sites, with the general intention of better understanding individuals via the information they post publicly. From the level of the individual, social listening can be used to form groups that encompass similar interests or attributes that the individuals freely stated. These groupings can then be utilized within data analyses to determine if there are trends within the populations, which may be helpful for organizations to better target the interests or better understand the issues common among the group, and thus address them in a more personalized manner.

One way to predict the number of future or incoming STEM and pre-health students to the university academic programs may be to use social listening. Social listening gives insight into trends of conversation volume on the topic of STEM and pre-health, sentiment about STEM and pre-health majors and careers, as well as how these conversations have changed in a pre- versus post-pandemic world. This insight can enable university recruiters and the healthcare field to gain a better understanding of not only the number of students or future career professionals to expect, but what these students are excited or concerned about as they prepare for their future careers.

In this study, I investigated how prospective and admitted STEM and pre-health students discussed majors of interest, STEM and health careers, and universities. Data used in this study was collected using Brandwatch Boolean from August 2018 to June 2021. STEM and pre-health mentions in online posts or forums were compared on an

academic year basis, with each academic year encompassing August until the next July, or June in the case of the third academic year. The sentiment associated with the mentions were used as a basis for comparison to determine not only how conversation volume changed but also how the sentiment associated with the posts potentially shifted across years. I also analyzed the difference in mention volume from the pre-COVID-19 pandemic, August 2018 through March 2020, to post-COVID-19 pandemic, April 2020 through June 2021. In addition, I mapped geographical data mentioned within the posts to view a breakdown of which areas of the United States had larger mention volume, as well as sentiment associated with geographical mentions.

2. Literature Review

2.1. Social media

Social media is a prevalent form of online communication in the modern world. Between the accessibility of online platforms via cellphones, tablets, computers, and more, most pre-higher education teens stay connected through social media platforms such as Facebook and Twitter. As of 2016, there are over two billion active users of social media globally, allowing for the connection of individuals and businesses on a mass scale (Kietzmann et al., 2016; Number, 2016). The term 'electronic word of mouth' relates heavily to social media, with reblogs, retweets, and hashtags allowing for information to be shared in an easily found form, with social media increasingly influencing consumer behaviors (Barnes & Jacobsen, 2014; Fotis et al., 2012). Social media and internet use are the second most prevalent forms of smartphone use by young adults, only surpassed by texting (Smith, 2015). Such frequent usage suggests a marketing opportunity with each social media platform engaged by a university or company.

2.2. Social listening

Social listening is a form of online data gathering regarding a specific area of interest, allowing for social intelligence within that field to be improved (Stewart & Arnold, 2018). The data gathered can be used to influence online relationships and alter communication to allow for a more targeted approach by companies or universities (Stewart & Arnold, 2018; Walters, 2013). From the data that has been gathered, companies are able to analyze the behavior of their customers or consumers, as well as the sentiment associated with their product and compare the information against data

associated with competitors (Gillin, 2008; TrackMaven, 2015). Besides simply tracking consumers, social listening allows companies to be able to actively respond and potentially increase consumer engagement with the company or product (Stewart & Arnold, 2018). Social listening is so impactful that leading research and advisory company Gartner Inc. considered it to be a top ten technology that impacted higher education in 2016 and 2017 (Gross & Horn, 2017; Lowendahl et al., 2016). Higher education is also considered to be one of the fastest growing service sectors, so utilization of a top ten technology within such a grow-heavy market provides boundless opportunities (Alwi & Kitchen, 2014; Bolat & O'Sullivan, 2017).

2.3. Quantitative versus qualitative measures

Due to the nature of social listening, the data that is retrieved is qualitative, with the sentiment of the words used being a key factor in the analysis by a company or university to determine whether a post or forum is discussing the terms of interest in a positive, neutral, or negative way. Engagement with a topic, a seemingly quantitative measurement, can even be broken down into qualitative measures of emotional, behavioral, and cognitive experiences that a user has when creating a post, helping to dictate the sentiment associated with it (Calder et al., 2016). After many online conversations have been collected and a sentiment has been associated with them, a quantitative analysis can be performed.

2.4. Ethics and biases

While social listening is an approach that can generate qualitative data, biases and ethics must be considered. Biases in social listening can include population bias, with a difference in the quantity of data versus the population said data represents within the

country or area of interest existing as a potential limitation to the usefulness of the data (Ruths & Pfeffer, 2014). Ethical considerations are also important when utilizing social listening, with studies having to consider confidentiality and the privacy of the authors of the online content being observed and analyzed. Ethical concerns on anonymity and data sensitivity also must be considered when utilizing data gathered through social listening (Fuchs, 2017). Social media tends to be considered within an ethical gray area, with blurred lines in what is considered informed consent when posting to an online platform. This blurring of boundaries is even more considerable when determining what is public or private, such as hobbies that may consume a business product (Fuchs, 2015). There are currently two main views within social media research, that of big data positivism and research ethics fundamentalism (Fuchs, 2017). Big data positivism suggests that information posted on social media is predominantly public and therefore essentially free to use without considering ethical concerns (Fuchs, 2017). In contrast, research ethics fundamentalism states that there is no informed consent when information is posted on a social media platform, therefore you must obtain consent to use the data (Fuchs, 2017). Studies utilizing social listening should thus be reviewed by an accredited board before data gathering takes place to ensure ethical gathering and use of data, specifically the level of privacy and consent an individual has upon posting on a public site or forum (Golder et al., 2017).

2.5. Social listening as a tool for universities

Social listening broadens the scope of higher education studies and enrollment predictions, with studies in equitable markets such as election predictions, product sales, and stock market fluctuations having already been studied (Schoen et al., 2013;

Greenhow et al., 2019). Just like these business ventures, universities can utilize social listening as a predictive technology for determining interest in the university, as well as offered majors and the career pipeline that feeds from higher education. With the appropriate use of social media, increased student engagement can be achieved, especially in the generations that have grown up with technology and the internet as a commonplace within their daily lives (Ratliff, 2011; Underwood et al., 2010). Deciding upon what is considered appropriate use of social media for student outreach is dynamic and constantly changing, just like social media itself. Realizing, understanding, and embracing trends at a specific period allows for the highest chance of engagement with the target audience (Olson & Martin, 2010).

2.6. Social media expectations of businesses and universities

It has now become almost an expectation that universities and businesses have a strong online presence, especially across social media platforms. The ability for a business to mimic the communication their consumers have with non-businesses is also expected by students, and a failure to do so could lead to a decrease in student engagement with the company or an increase in negative sentiment associated with them (Ratliff, 2011). Twitter is considered an especially useful platform to reach potential students, with the limited character allotment per tweet allowing for a concise message to be relayed that can keep attention due to its brevity (Ratliff, 2011).

2.7. Twitter and Facebook integration into university use

Twitter and Facebook have already become integrated into university recruiting across the United States, with the usage increasingly being utilized to communicate with and draw in potential students (Ganim Barnes & Mattson, 2009; Wankel & Wankel,

2011). Due to the successful integration of these technology habits within universities, social media sites such as Facebook and Twitter have been considered key communication platforms for the current generation of students (NACAC, 2009). While communication with students is essential, sharing of information is not the only thing that should be focused on. The ability of university recruiters to listen to what is being shared and the sentiment behind the words within a post are just as essential for engagement (Ouellette, 2018).

2.8. Utilization of student feedback

The utilization of student feedback is important to the students that have posted, as well as those who see the response of the university, ultimately affecting the satisfaction and potential future post sentiment that could come from an active university response (Douglas et al., 2006; Kuo et al., 2013). Social media has allowed for multiple platforms of communication between universities and students, with the ability for immediate response to feedback gained in student conversations (Peruta & Shields, 2016). Students who feel as though their online feedback has made an impact that they can see are more likely to participate in conversations in the future (Bennet & Nair, 2010; Coates, 2006; Shah & Nair, 2009). Students are also more likely to use social media to voice their opinions or concerns about a university as social media platforms are generally seen as informal environments in which their opinion can be seen without potential repercussions (Barlow & Thomas, 2011).

2.9. Earning a competitive edge using social media as a university

Even with the success seen in use of social media platforms at the university recruiting level, little research has been completed on the use of social media and social

listening in higher education recruitment. This is an issue that deserves consideration, as social media use by universities gives a competitive edge over those that do not interact with and respond to potential students (Finch et al., 2013; Koshkin et al., 2016).

3. Materials and Methods

3.1. Study Design

This study analyzed online conversation about incoming STEM and pre-health undergraduate students' college going experiences to gain a deeper understanding of their journeys to higher education – their motivations, obstacles, and opportunities. This research highlights first-person perspectives from more than 5,000 STEM and pre-health students over a three-year collection period to help campuses meet them where they are at and provide focused support to guide them on pathways to success, all while informing strategies to develop the STEM and pre-health pipeline. This study aims to understand STEM and pre-health students' evolving perspectives about, motivations for, and obstacles to enrolling in college. The study also explores how social listening captures authentic student conversations about perceptions of college and identifying strategies to best meet student needs (Gross, 2019). Prospective and admitted STEM and pre-health students' journey to higher education, view on majors of interest, potential careers, and specific universities are utilized to understand conversation trends around these topics. The study also explores how STEM and pre-health student conversation volume decreased from pre-pandemic to post-pandemic periods within the data collection window and how the sentiment of these conversations changed within that same period. Similarly, we explore how STEM and pre-health student conversation volume and associated sentiment varies across the major geographical area of the United States and the top 200 cities by population.

This study achieved these goals through a collaboration with *Campus Sonar*, a social listening company who utilizes Brandwatch Consumer Research, an enterprise-level

social listening software used to gather the data for this study. Brandwatch has access to over 100 million publicly-available online sources and over 1.3 trillion individual posts across a variety of social media sites, including Twitter, Reddit, Instagram, Tumblr, and YouTube, as well as sites such as blogs, news outlets, forums, and review sites. From these many publicly available sources, Brandwatch can be used to filter the data to fit the question at hand. Within this study, the research team identified unique key terms and phrases and built a Boolean query to collect online social data from the Brandwatch archive (Gross, 2021).

3.2. Experimental/Data Limits and Query Factors – Brandwatch

For this study, data collected within the Brandwatch Boolean had to be validated via the use of four conditions, of which each data piece had to fit within the boundaries (Jeppson, 2021). The first condition that data had to meet was that of the mention being in first-person and stated as being from a prospective or admitted undergraduate student who wants to attend a higher education institution within the United States. The second condition was that the mention had to be from a publicly accessible site such as the social media sites stated above, or a blog or forum site as designated by the Brandwatch Consumer Research software. The third condition was that the mention could not be a retweet or a post by another individual that has been shared, as this is not inherently indicative of the person who retweeted or shared the post also fitting within the criteria laid out within the first condition. The fourth and final condition for a mention to be included within the study's dataset was that the mention had to originate from a location within the United States or indicate within the mention text itself that the student was located within the United States. This fourth condition was achieved using Brandwatch

leveraging any location metadata included with each mention to determine whether it originated in the United States, filtering with the top 200 United States cities by population, and any references of nine set geographic regions (New England, Mid East, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, Far West, and Outlaying Areas).

Once each data point was determined to likely be from a student within the United States, additional steps were taken by the Campus Sonar team to further validate the data. A propriety Boolean was used to identify common irrelevancies within the entire dataset, such as international higher education terms, historical/reminiscent conversation, already graduated, post graduate parents and students, and undergraduate parents. Top phrases, top sites, and top authors were also reviewed for the dataset (Sandall, 2020). Each conversation also had to return for the research questions examined by site and audience, such as: prospective student + health careers + Reddit.

The dataset was segmented to allow for the analysis of research questions more easily, with Campus Sonar creating a set of Boolean rules that coded the data by enrollment, science and pre-health studies, health and science careers, socioeconomic status, and racial identities. Enrollment was defined as either prospective or admitted students, with prospective broken down further into subcategories of prospective, visit, and applicant. Science and pre-health studies was defined as the area of study for science and pre-health students, and segmented into science majors, with science majors identifying conversations within STEM and pre-health studies that contained major-related terms. Health and science careers were defined as containing careers that STEM and pre-health majors tend to enter post-graduation. Socioeconomic status was broken

into three categories: high, middle, and low. Racial identities were divided into given races and ethnicities, as well as Native American specifically due to the geographical prevalence of Native Americans within South Dakota and the surrounding regions.

Limitations within the dataset included the unpredictability and idiosyncrasies of language use within conversations, thus the inability to have definite sentiment intention on every datapoint (Sandall, 2021). To try to counteract this, the research team utilized a list of key terms when crafting the Boolean query and rules based on industry expertise, secondary research, client conversations, and a variety of tests to ensure that a robust and relevant sample of social data was collected. Another limitation was due to privacy terms changes within Instagram, Tumblr, and YouTube that limited how much of the site's social listening software can be accessed and for how long. A third limitation related to the collection of geographic location based on the social data, as not all data collected has associated geographic indicators. To counteract this, the Campus Sonar analysts applied linguistic and contextual tactics to identify mentions in which the author self-identifies as being from the United States (Miller & Sandall, 2020).

3.3. Statistical Analyses

The raw data within this study was analyzed using RStudio. The model utilized to analyze the data was a Kruskal-Wallis test due to the nature of the data's dependent variable being ordinal or ranked and there being one independent variable with 2+ categories or groups, which are independent of the other categories or groups present. The independent variables within this study were defined dependent on the research question being addressed. Positive sentiment was compared from academic year to

academic year, with the independent variables as each academic year period, resulting in three groups that could be used within the Kruskal-Wallis model. Similarly, the neutral and negative sentiment volumes were compared across the three years, once again broken into three groups, one representing each of the academic years that had been observed. The analyses of prospective and admitted STEM and pre-health student conversation about STEM and pre-health majors and careers were performed in a similar way, with a yearly breakdown to compare against the other two years/groups with the Kruskal-Wallis test. The pre-pandemic and post-pandemic time periods were also analyzed using this method, with pre- and post-pandemic giving the two required groups to use a Kruskal-Wallis model. The geographical data was also analyzed with independent variables of United States geographical regions, resulting in nine independent variables: New England, Mid East, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, Far West, and Outlaying Areas. The dependent variables of all the Kruskal-Wallis tests performed were the conversation sentiments of positive, neutral, and negative, as these values depended on the time periods or regions for their change. An alpha of 0.05 was utilized to determine whether there was a significant difference observed between the groups being compared. After creating the null and alternate hypotheses and stating an alpha value, RStudio was utilized to create a boxplot comparing the data in a visual form. Similarly, the Kruskal-Wallis test code was executed through the program to provide a Kruskal-Wallis chi-squared value, a degree of freedom for the data, and a p-value which then can be compared to the stated alpha to determine if there was a significant change between the two sets of data being compared. If the p-value is less than 0.05, the null hypothesis is rejected, meaning that the mean ranks of the groups were not the same and

there is a difference between the mean ranks of the two groups in the Kruskal-Wallis test. Alternatively, if the p-value is greater than 0.05, we fail to reject the null hypothesis, meaning that the mean ranks of the two groups were the same in the Kruskal-Wallis test.

4. Results & Discussion

4.1. How do prospective and admitted STEM and pre-health students talk about their journey to higher education in the United States?

4.1.1. Prospective STEM and pre-health student positive sentiment mentions

Our analysis indicated that prospective STEM and pre-health student online mentions with positive sentiment decreased over the thirty-five-month collection period. 3494 total mentions were made that indicated that a prospective student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 3494 mentions, 724 mentions had positive sentiment associated with the mention, accounting for approximately 20.7% of all mentions (Table 1, Figure 1). From the 2018/2019 academic year, there were 410 positively associated prospective STEM and pre-health student mentions. There was a decrease of 232 positive mentions from the first academic year to the second academic year, a decrease of approximately 56.6%. From the 2019/2020 academic year to the third academic year studied, a decrease of 42 positive mentions occurred, resulting in a loss of approximately 23.6% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 3.783×10^{-5}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 1. Prospective student STEM and pre-health mentions with positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Prospective Student Positive Sentiment	3494	410	178	136	20.365	3.783×10^{-5}

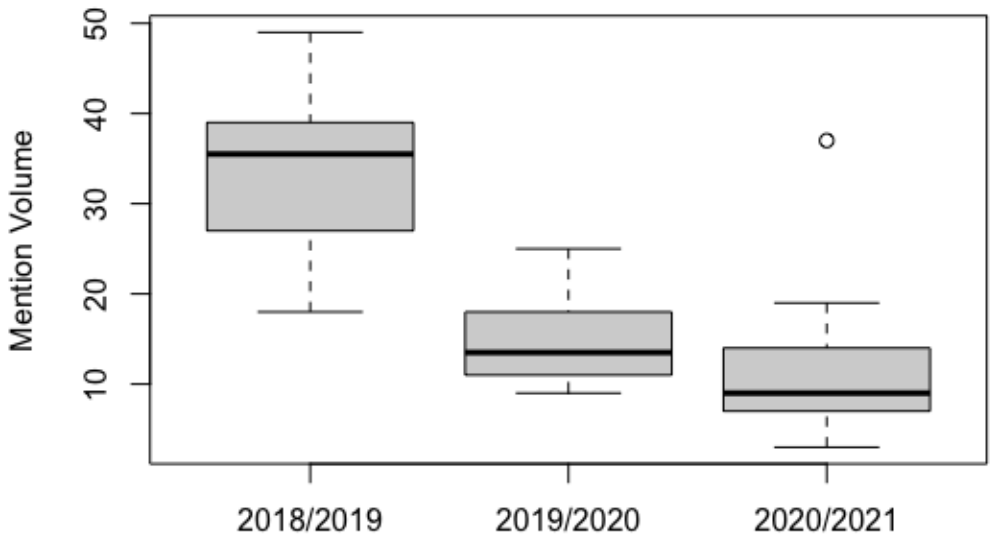

Figure 1. Prospective student STEM and pre-health mentions with positive sentiment

4.1.2. Prospective STEM and pre-health student neutral sentiment mentions

Our analysis indicated that prospective STEM and pre-health student online mentions with neutral sentiment decreased over the thirty-five-month collection period. 3494 total mentions were made that indicated that a prospective student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 3494 mentions, 1633 mentions had neutral sentiment associated with the mention, accounting for approximately 46.7% of all mentions (Table 2, Figure 2). From the 2018/2019 academic year, there were 842 neutrally associated prospective STEM and pre-health student mentions. There was a decrease of 327 neutral mentions from the first academic

year to the second academic year, a decrease of approximately 38.8%. From the 2019/2020 academic year to the third academic year studied, a decrease of 187 neutral mentions occurred, resulting in a loss of approximately 36.3% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 1.708×10^{-5}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 2. Prospective student STEM and pre-health mentions with neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Prospective Student Neutral Sentiment	3494	842	515	328	21.958	1.705×10^{-5}

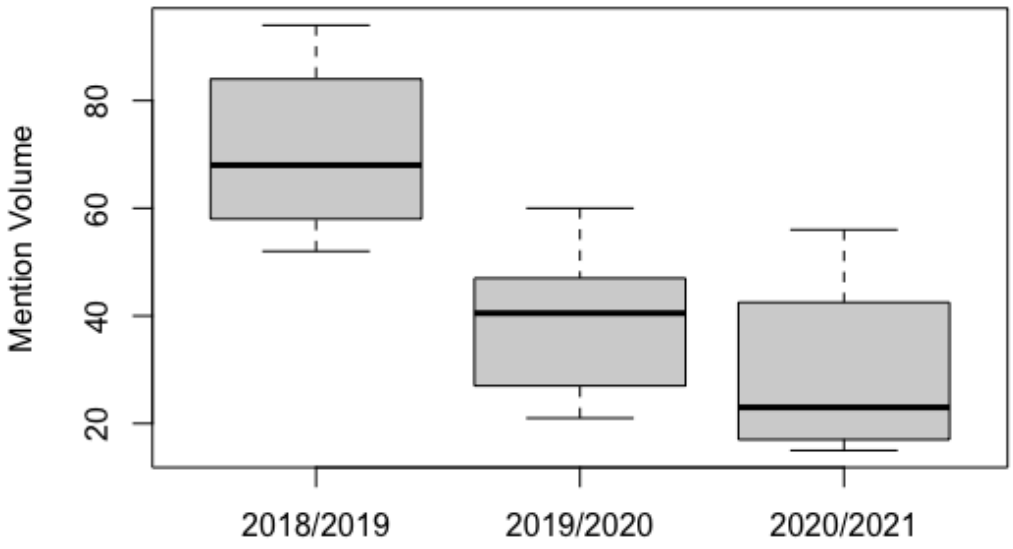

Figure 2. Prospective student STEM and pre-health mentions with neutral sentiment

4.1.3. Prospective STEM and pre-health student negative sentiment mentions

Our analysis indicated that prospective STEM and pre-health student online mentions with negative sentiment decreased over the thirty-five-month collection period. 3494 total mentions were made that indicated that a prospective student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 3494 mentions, 1137 mentions had negative sentiment associated with the mention, accounting for approximately 32.5% of all mentions (Table 3, Figure 3). From the 2018/2019 academic year, there were 726 negatively associated prospective STEM and pre-health student mentions. There was a decrease of 461 negative mentions from the first academic year to the second academic year, a decrease of approximately 63.5%. From the 2019/2020 academic year to the third academic year studied, a decrease of 119 negative mentions occurred, resulting in a loss of approximately 44.9% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 1.658×10^{-6}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 3. Prospective student STEM and pre-health mentions with negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value

Prospective Student Negative Sentiment	3494	726	265	146	26.62	1.658 x 10^{-6}

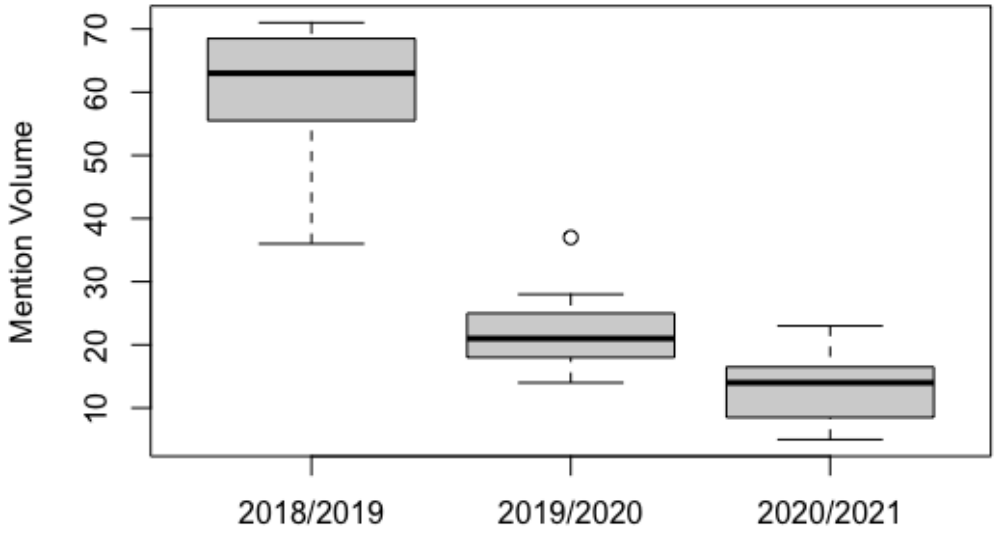

Figure 3. Prospective student STEM and pre-health mentions with negative sentiment

4.1.4. Admitted STEM and pre-health student positive sentiment mentions

Our analysis indicated that admitted STEM and pre-health student online mentions with positive sentiment decreased over the thirty-five-month collection period. 958 total mentions were made that indicated that an admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 958 mentions, 351 mentions had positive sentiment associated with the mention, accounting for approximately 36.6% of all mentions (Table 4, Figure 4). From the 2018/2019 academic year, there were 161 positively associated prospective STEM and pre-health student mentions. There was a decrease of 43 positive mentions from the first academic year to the second academic year, a decrease of approximately 26.7%. From the 2019/2020 academic year to the third academic year studied, a decrease of 46 positive mentions occurred, resulting in a loss of approximately 39.0% from the second academic year to

the third. The null hypothesis was not rejected, indicating that the mean ranks of the three compared groups (the three academic years) were the same; there was not a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.1574, a value greater than the set alpha value of 0.05, indicated no statistical significance within the difference between the means of the compared groups.

Table 4. Admitted student STEM and pre-health mentions with positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Admitted Student Positive Sentiment	958	161	118	72	3.698	0.1574

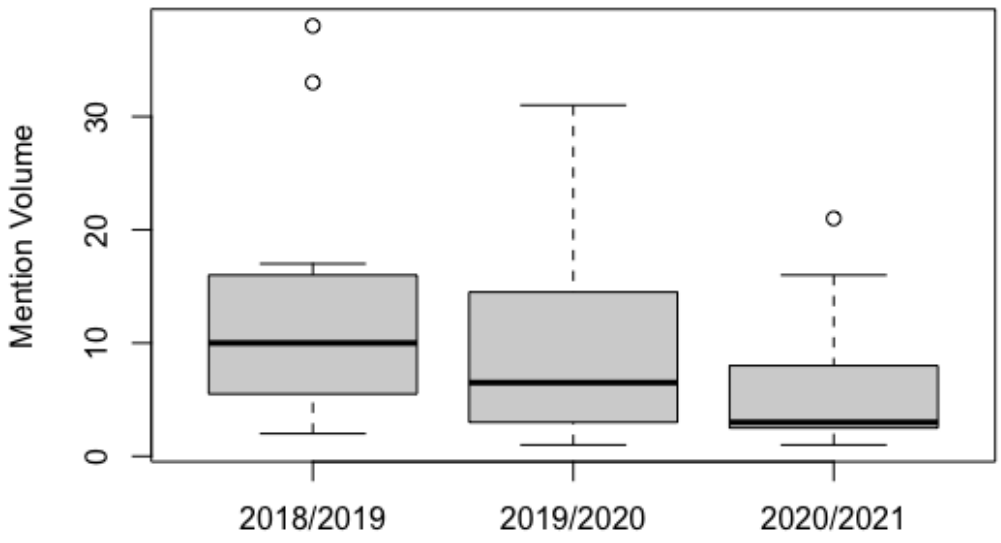

Figure 4. Admitted student STEM and pre-health mentions with positive sentiment

4.1.5. Admitted STEM and pre-health student neutral sentiment mentions

Our analysis indicated that admitted STEM and pre-health student online mentions with neutral sentiment decreased over the thirty-five-month collection period. 958 total mentions were made that indicated that an admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 958 mentions, 294 mentions had neutral sentiment associated with the mention, accounting for approximately 30.7 % of all mentions (Table 5, Figure 5). From the 2018/2019 academic year, there were 161 neutrally associated prospective STEM and pre-health student mentions. There was a decrease of 83 neutral mentions from the first academic year to the second academic year, a decrease of approximately 51.6%. From the 2019/2020 academic year to the third academic year studied, a decrease of 23 neutral mentions occurred, resulting in a loss of approximately 29.5% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.01992, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 5. Admitted student STEM and pre-health mentions with neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Admitted Student Neutral Sentiment	958	161	78	55	7.8313	0.01992

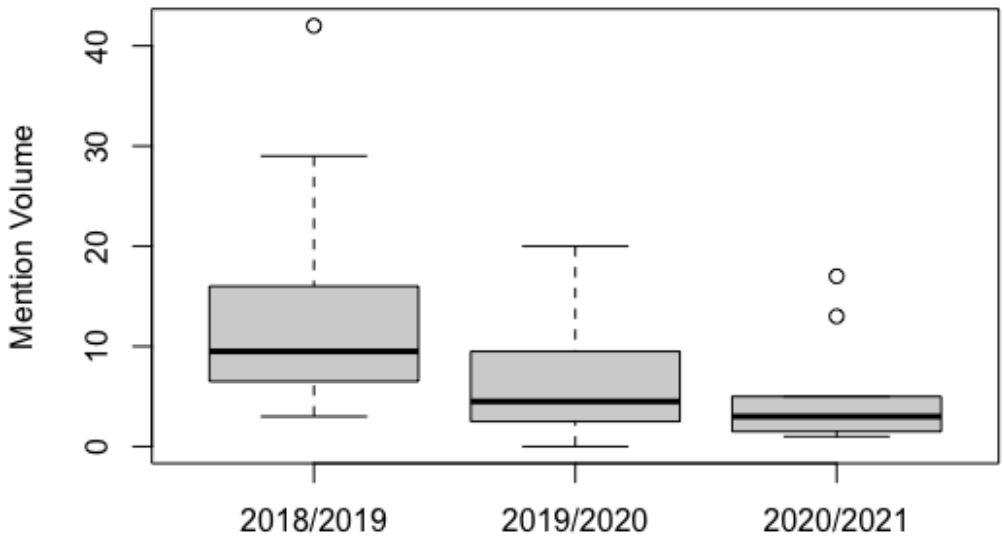

Figure 5. Admitted student STEM and pre-health mentions with neutral sentiment

4.1.6. Admitted STEM and pre-health student neutral sentiment mentions

Our analysis indicated that admitted STEM and pre-health student online mentions with negative sentiment decreased over the thirty-five-month collection period. 958 total mentions were made that indicated that an admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 958 mentions, 313 mentions had negative sentiment associated with the mention, accounting for approximately 32.7 % of all mentions (Table 6, Figure 6). From the 2018/2019 academic year, there were 204 negatively associated prospective STEM and pre-health student mentions. There was a decrease of 129 negative mentions from the first academic year to the second academic year, a decrease of approximately 63.2%. From the 2019/2020 academic year to the third academic year studied, a decrease of 41 negative mentions occurred, resulting in a loss of approximately 54.7% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the

Kruskal-Wallis test was 1.584×10^{-4}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 6. Admitted student STEM and pre-health mentions with negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Admitted Student Negative Sentiment	958	204	75	34	17.501	1.584×10^{-4}

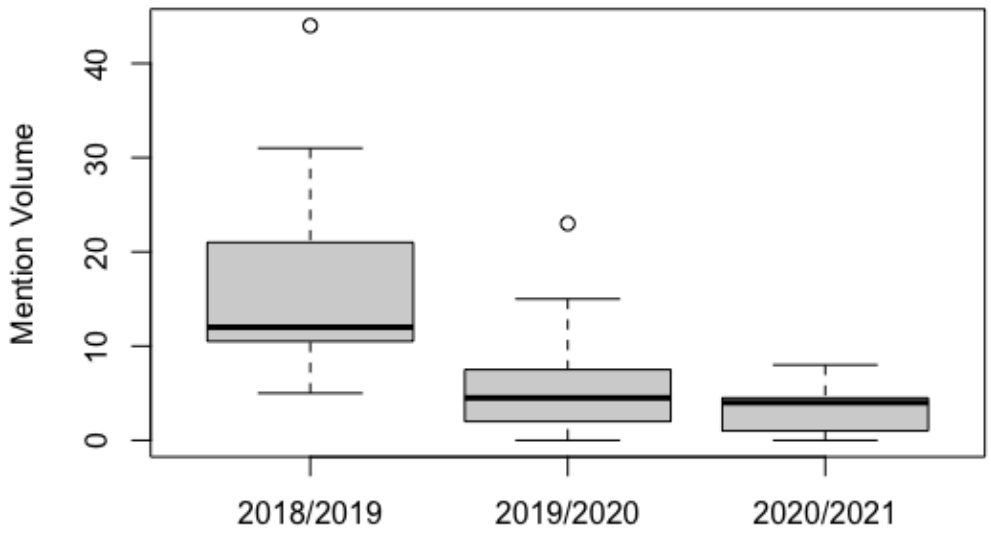

Figure 6. Admitted student STEM and pre-health mentions with negative sentiment

4.2. What do prospective and admitted STEM and pre-health students think about STEM and pre-health majors?

4.2.1. STEM and pre-health academic major positive sentiment mentions

Our analysis indicated that STEM and pre-health student academic major related online mentions with positive sentiment fluctuated over the thirty-five-month collection

period. 1466 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health academic major topic that was within the Boolean query terms. Out of the 1466 mentions, 336 mentions had positive sentiment associated with the mention, accounting for approximately 22.9% of all mentions (Table 7, Figure 7). From the 2018/2019 academic year, there were 161 positively associated prospective or admitted STEM and pre-health academic major student mentions. There was a decrease of 77 positive mentions from the first academic year to the second academic year, a decrease of approximately 47.8%. From the 2019/2020 academic year to the third academic year studied, an increase of 7 positive mentions occurred, resulting in a gain of approximately 8.3% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.01105, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 7. STEM and pre-health academic major related mentions with positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Academic Majors Positive Sentiment	1466	161	84	91	9.0105	0.01105

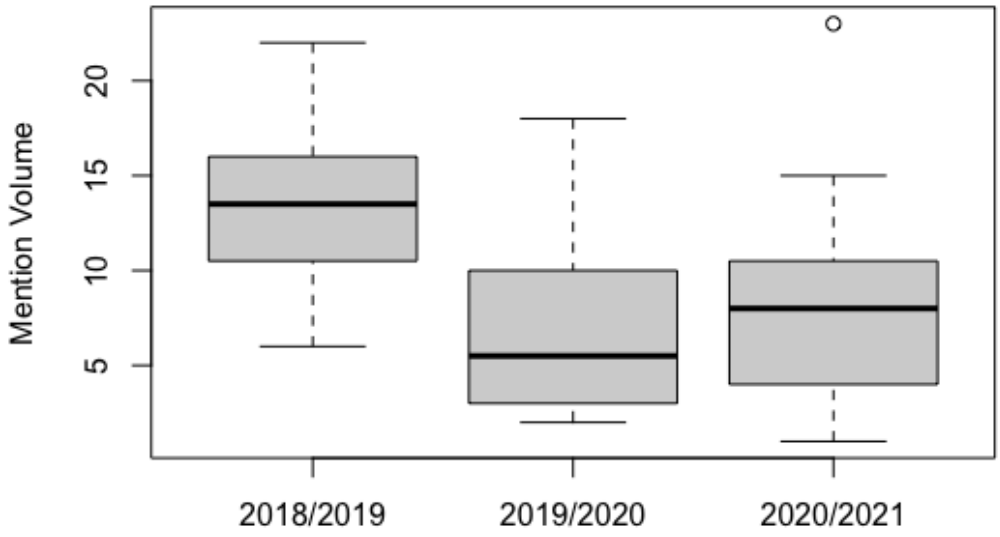

Figure 7. STEM and pre-health academic major related mentions with positive sentiment

4.2.2. STEM and pre-health academic major neutral sentiment mentions

Our analysis indicated that STEM and pre-health student academic major related online mentions with neutral sentiment decreased over the thirty-five-month collection period. 1466 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health academic major topic that was within the Boolean query terms. Out of the 1466 mentions, 734 mentions had neutral sentiment associated with the mention, accounting for approximately 50.1% of all mentions (Table 8, Figure 8). From the 2018/2019 academic year, there were 334 neutrally associated prospective or admitted STEM and pre-health academic major student mentions. There was a decrease of 114 neutral mentions from the first academic year to the second academic year, a decrease of approximately 34.1%. From the 2019/2020 academic year to the third academic year studied, a decrease of 40 neutral mentions occurred, resulting in a loss of approximately 18.2% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks

of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.004906, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 8. STEM and pre-health academic major related mentions with neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Academic Majors Neutral Sentiment	1466	334	220	180	10.635	0.004906

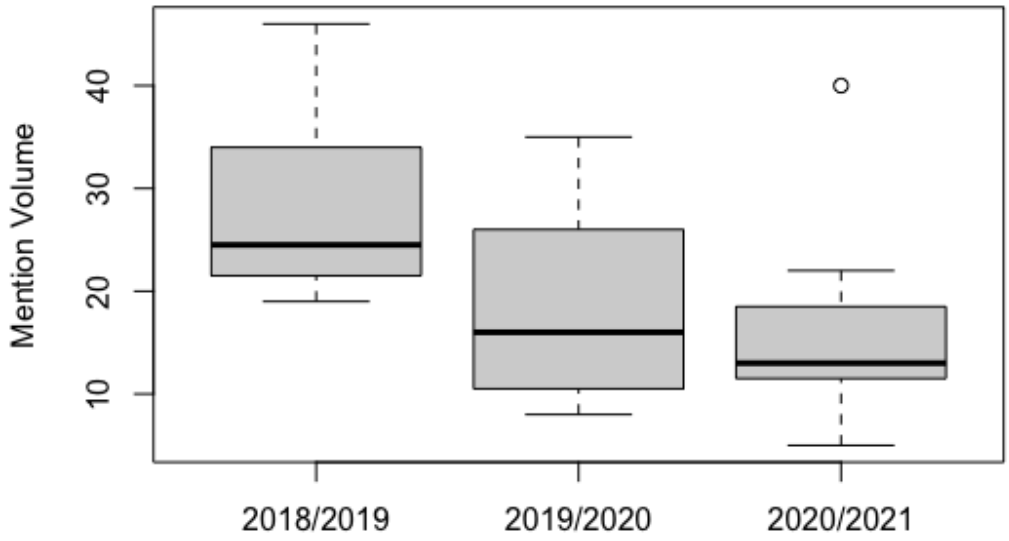

Figure 8. STEM and pre-health academic major related mentions with neutral sentiment

4.2.3. STEM and pre-health academic major negative sentiment mentions

Our analysis indicated that STEM and pre-health student academic major related online mentions with negative sentiment decreased over the thirty-five-month collection period. 1466 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health academic major topic that was within the

Boolean query terms. Out of the 1466 mentions, 396 mentions had negative sentiment associated with the mention, accounting for approximately 27.0% of all mentions (Table 9, Figure 9). From the 2018/2019 academic year, there were 239 negatively associated prospective or admitted STEM and pre-health academic major student mentions. There was a decrease of 136 negative mentions from the first academic year to the second academic year, a decrease of approximately 56.9%. From the 2019/2020 academic year to the third academic year studied, a decrease of 49 negative mentions occurred, resulting in a loss of approximately 47.6% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 4.698×10^{-5}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 9. STEM and pre-health academic major related mentions with negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Academic Majors Negative Sentiment	1466	239	103	54	19.931	4.698×10^{-5}

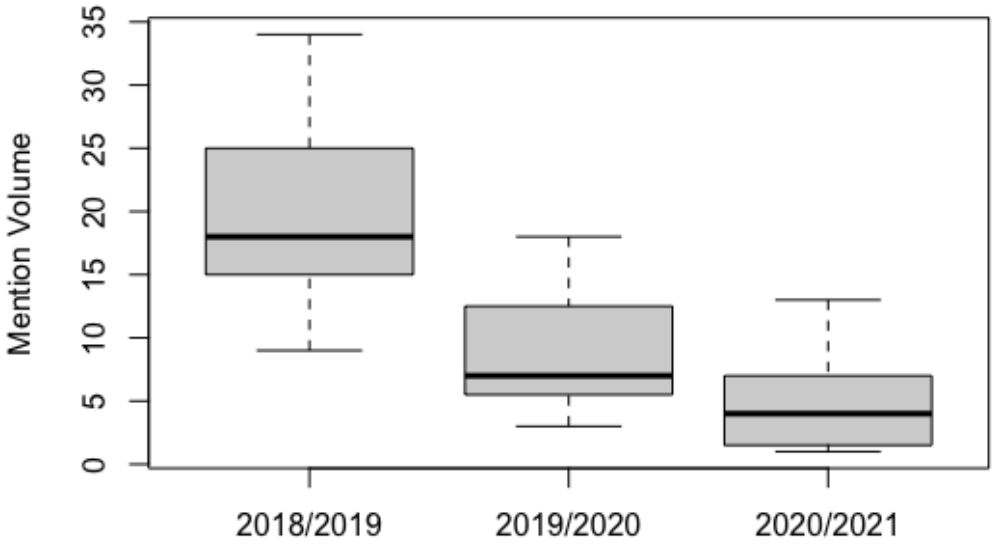

Figure 9. STEM and pre-health academic major related mentions with negative sentiment

4.3. What do prospective and admitted STEM and pre-health students think about STEM and pre-health careers?

4.3.1. STEM and pre-health career positive sentiment mentions

Our analysis indicated that STEM and pre-health student career related online mentions with positive sentiment decreased over the thirty-five-month collection period. 972 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health career topic that was within the Boolean query terms. Out of the 972 mentions, 120 mentions had positive sentiment associated with the mention, accounting for 27.7% of all mentions (Table 10, Figure 10). From the 2018/2019 academic year, there were 120 positively associated prospective or admitted STEM and pre-health career student mentions. There was a decrease of 34 positive mentions from the first academic year to the second academic year, a decrease of 28.3%. From the 2019/2020 academic year to the third academic year studied, a decrease of 22 positive mentions occurred, resulting in a loss of 25.6% from the second academic year to the third. The null hypothesis was not rejected, indicating that the mean ranks of the three compared groups (the three academic years) were the same; there was not a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.3013, a value greater than the set alpha value of 0.05, indicated no statistical significance within the difference between the means of the compared groups.

Table 10. STEM and pre-health career related mentions with positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-	p-value

					squared (H)	
Career Positive Sentiment	972	120	86	64	2.3995	0.3013

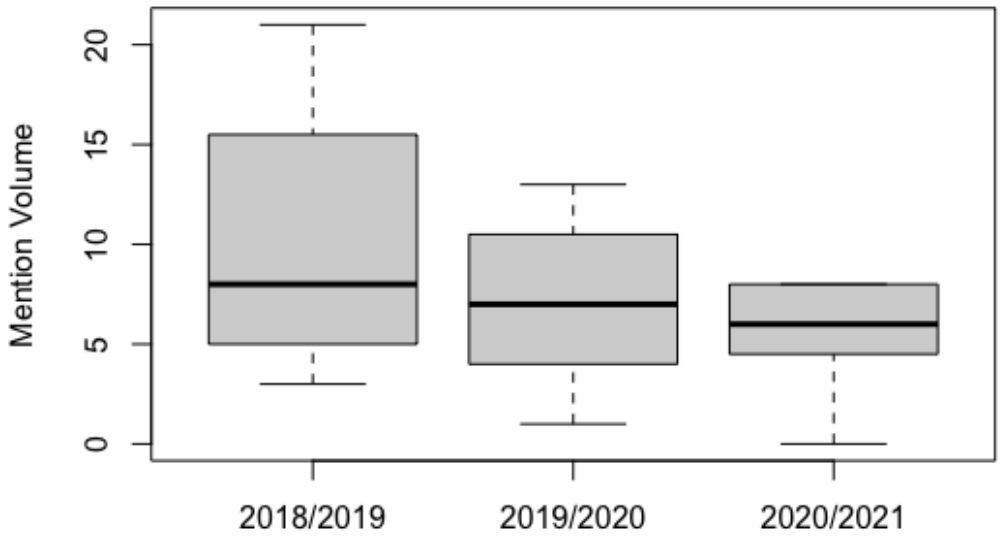

Figure 10. STEM and pre-health career related mentions with positive sentiment

4.3.2. STEM and pre-health career neutral sentiment mentions

Our analysis indicated that STEM and pre-health student career related online mentions with neutral sentiment fluctuated over the thirty-five-month collection period. 972 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health career topic that was within the Boolean query terms. Out of the 972 mentions, 458 mentions had neutral sentiment associated with the mention, accounting for 47.1% of all mentions (Table 11, Figure 11). From the 2018/2019 academic year, there were 149 neutrally associated prospective or admitted STEM and pre-health career student mentions. There was an increase of 25 neutral mentions from the first academic year to the second academic year, an increase of 16.8%. From the 2019/2020 academic year to the third academic year studied, a decrease of 39

neutral mentions occurred, resulting in a loss of 22.4% from the second academic year to the third. The null hypothesis was not rejected, indicating that the mean ranks of the three compared groups (the three academic years) were the same; there was not a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.4094, a value greater than the set alpha value of 0.05, indicated no statistical significance within the difference between the means of the compared groups.

Table 11. STEM and pre-health career related mentions with neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Career Neutral Sentiment	972	149	174	135	1.7859	0.4094

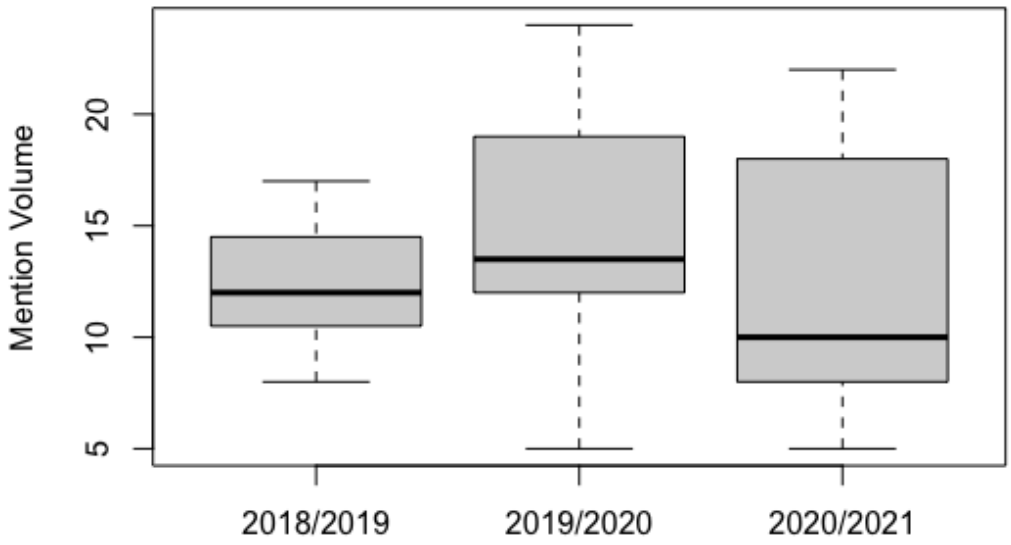

Figure 11. STEM and pre-health career related mentions with neutral sentiment

4.3.3. STEM and pre-health career negative sentiment mentions

Our analysis indicated that STEM and pre-health student career related online mentions with negative sentiment decreased over the thirty-five-month collection period. 972 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health career topic that was within the Boolean query terms. Out of the 972 mentions, 244 mentions had negative sentiment associated with the mention, accounting for 25.1% of all mentions (Table 12, Figure 12). From the 2018/2019 academic year, there were 126 negatively associated prospective or admitted STEM and pre-health career student mentions. There was a decrease of 56 negative mentions from the first academic year to the second academic year, a decrease of 44.4%. From the 2019/2020 academic year to the third academic year studied, a decrease of 22 negative mentions occurred, resulting in a loss of 31.4% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.00127, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 12. STEM and pre-health career related mentions with negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Career Negative Sentiment	972	126	70	48	13.576	0.001127

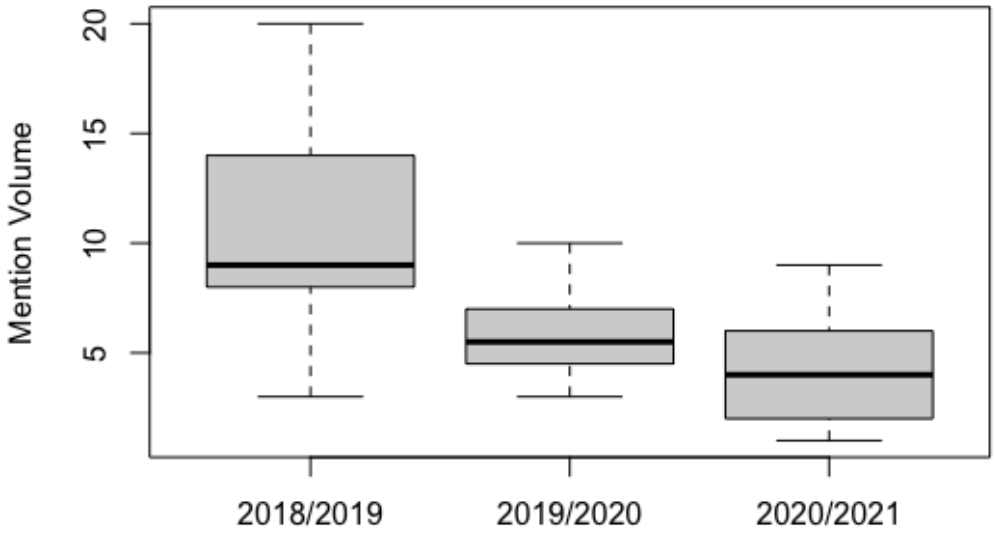

Figure 12. STEM and pre-health career related mentions with negative sentiment

4.4. How do prospective and admitted STEM and pre-health students talk about specific universities?

4.4.1. Top 5 and Non-Top 5 mentioned universities

Our analysis indicated that STEM and pre-health student online conversation volume regarding universities could be broken into the top 5 mentioned universities and the non-top 5 mentioned universities. 11516 total mentions were made that indicated that a prospective or admitted student discussed a university that was within the Boolean query terms. Out of the 11516 mentions, 3843 mentions contained the top 5 mentioned universities, accounting for 33.4% of all mentions (Table 13, Figure 13). The remaining 7673 non-top 5 university mentions account for 66.6% of all mentions. The null hypothesis was rejected, indicating that the mean ranks of the 2 compared groups (top 5 and non-top 5 mentioned universities) were not the same; there was a difference between the mean ranks of the two groups in this test. The p-value of the Kruskal-Wallis test was 1.889×10^{-7}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 13. Top 5 and Non-Top 5 mentioned universities conversation volume

	Total Mentions (n)	Top 5 Universities Mention Volume	Non-Top 5 Universities Mention Volume	Kruskal-Wallis chi-squared (H)	p-value
Top 5 and Non-Top 5 Universities Sentiment	11516	3843	7673	27.143	1.889×10^{-7}

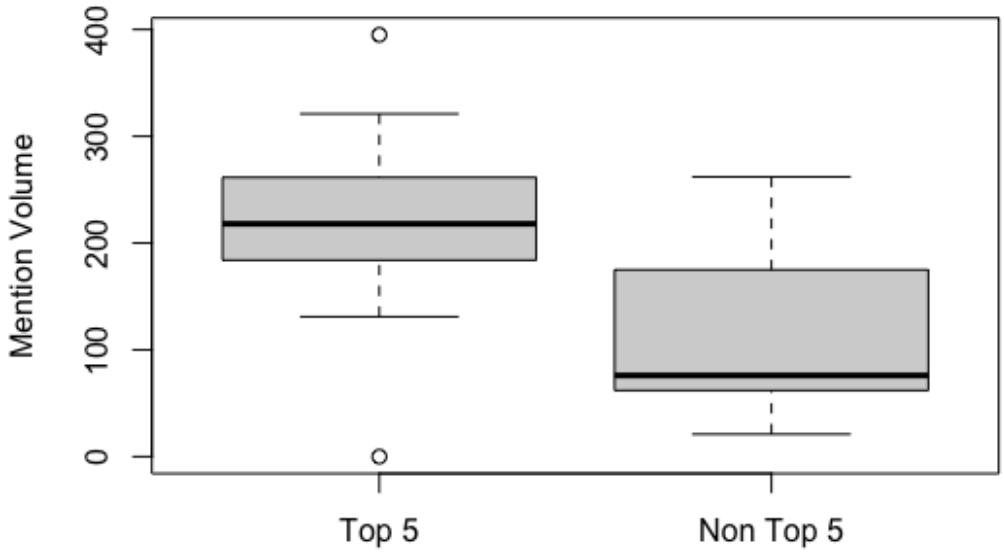

Figure 13. Top 5 and Non-Top 5 mentioned universities conversation volume

4.4.2. Top 5 mentioned universities positive sentiment

Our analysis indicated that STEM and pre-health student top 5 mentioned universities related online mentions with positive sentiment decreased over the thirty-five-month collection period. 3843 total mentions were made that indicated that a prospective or admitted student discussed a top 5 mentioned university that was within the Boolean query terms. Out of the 3843 mentions, 936 mentions had positive sentiment associated with the mention, accounting for 24.4% of all mentions (Table 14, Figure 14). From the 2018/2019 academic year, there were 528 positively associated mentions involving a top 5 mentioned university. There was a decrease of 299 positive mentions from the first academic year to the second academic year, a decrease of 56.6%. From the 2019/2020 academic year to the third academic year studied, a decrease of 50 positive mentions occurred, resulting in a loss of 21.8% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference

between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.0001569, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 14. Top 5 mentioned universities positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Top 5 Universities Positive Sentiment	3843	528	229	179	17.52	0.0001569

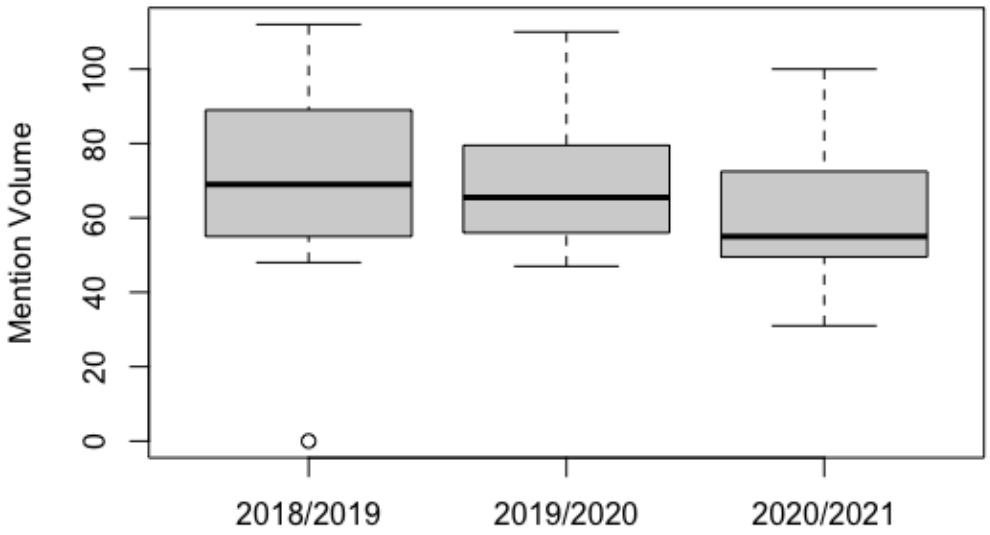

Figure 14. Top 5 mentioned universities positive sentiment

4.4.3. Top 5 mentioned universities neutral sentiment

Our analysis indicated that STEM and pre-health student top 5 mentioned universities related online mentions with neutral sentiment decreased over the thirty-five-month collection period. 3843 total mentions were made that indicated that a prospective

or admitted student discussed a top 5 mentioned university that was within the Boolean query terms. Out of the 3843 mentions, 1571 mentions had neutral sentiment associated with the mention, accounting for 40.9% of all mentions (Table 15, Figure 15). From the 2018/2019 academic year, there were 885 neutrally associated mentions involving a top 5 mentioned university. There was a decrease of 481 neutral mentions from the first academic year to the second academic year, a decrease of 54.5%. From the 2019/2020 academic year to the third academic year studied, a decrease of 122 neutral mentions occurred, resulting in a loss of 30.2% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 6.333×10^{-6}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 15. Top 5 mentioned universities neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Top 5 Universities Neutral Sentiment	3843	885	404	282	23.939	6.333×10^{-6}

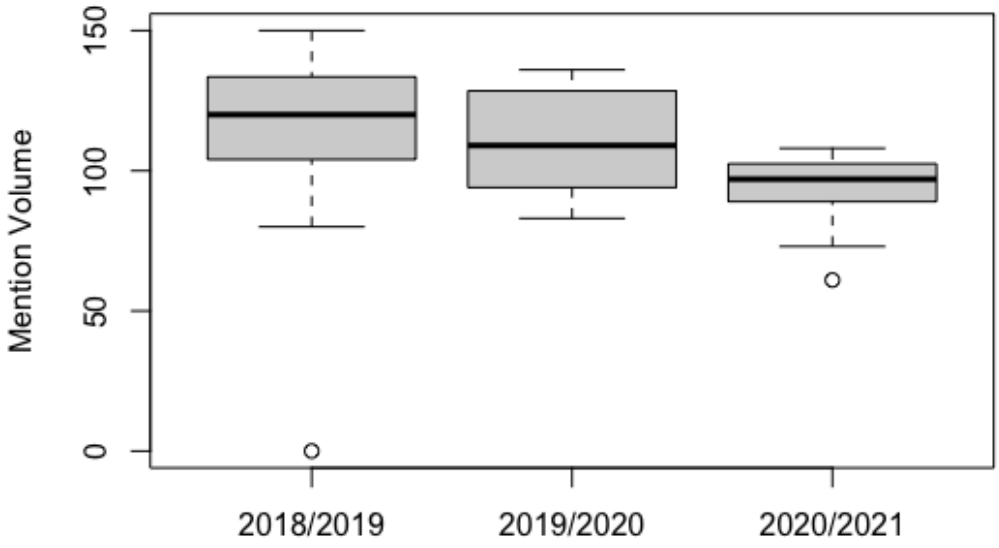

Figure 15. Top 5 mentioned universities neutral sentiment

4.4.4. Top 5 mentioned universities negative sentiment

Our analysis indicated that STEM and pre-health student top 5 mentioned universities related online mentions with negative sentiment decreased over the thirty-five-month collection period. 3843 total mentions were made that indicated that a prospective or admitted student discussed a top 5 mentioned university that was within the Boolean query terms. Out of the 3843 mentions, 1336 mentions had negative sentiment associated with the mention, accounting for 34.8% of all mentions (Table 16, Figure 16). From the 2018/2019 academic year, there were 888 negatively associated mentions involving a top 5 mentioned university. There was a decrease of 582 negative mentions from the first academic year to the second academic year, a decrease of 65.5%. From the 2019/2020 academic year to the third academic year studied, a decrease of 164 negative mentions occurred, resulting in a loss of 53.6% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference

between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 8.198×10^{-7}, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 16. Top 5 mentioned universities negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Top 5 Universities Negative Sentiment	3843	888	306	142	28.028	8.198×10^{-7}

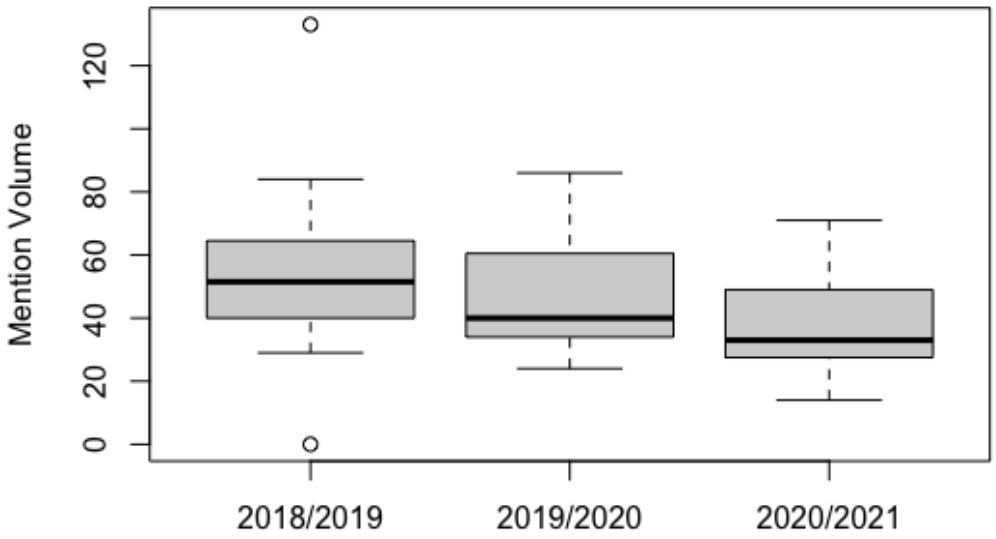

Figure 16. Top 5 mentioned universities negative sentiment

4.4.5. Non-Top 5 mentioned universities positive sentiment

Our analysis indicated that STEM and pre-health student non-top 5 mentioned universities related online mentions with positive sentiment fluctuated over the thirty-five-month collection period. 7673 total mentions were made that indicated that a

prospective or admitted student discussed a non-top 5 mentioned university that was within the Boolean query terms. Out of the 7673 mentions, 2345 mentions had positive sentiment associated with the mention, accounting for 30.6% of all mentions (Table 17, Figure 17). From the 2018/2019 academic year, there were 830 positively associated mentions involving a non-top 5 mentioned university. There was an increase of 12 positive mentions from the first academic year to the second academic year, an increase of 1.4%. From the 2019/2020 academic year to the third academic year studied, a decrease of 169 positive mentions occurred, resulting in a loss of 20.1% from the second academic year to the third. The null hypothesis was not rejected, indicating that the mean ranks of the three compared groups (the three academic years) were the same; there was not a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.4369, a value greater than the set alpha value of 0.05, indicated no statistical significance within the difference between the means of the compared groups.

Table 17. Non-Top 5 mentioned universities positive sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Non-Top 5 Universities Positive Sentiment	7673	830	842	673	1.656	0.4369

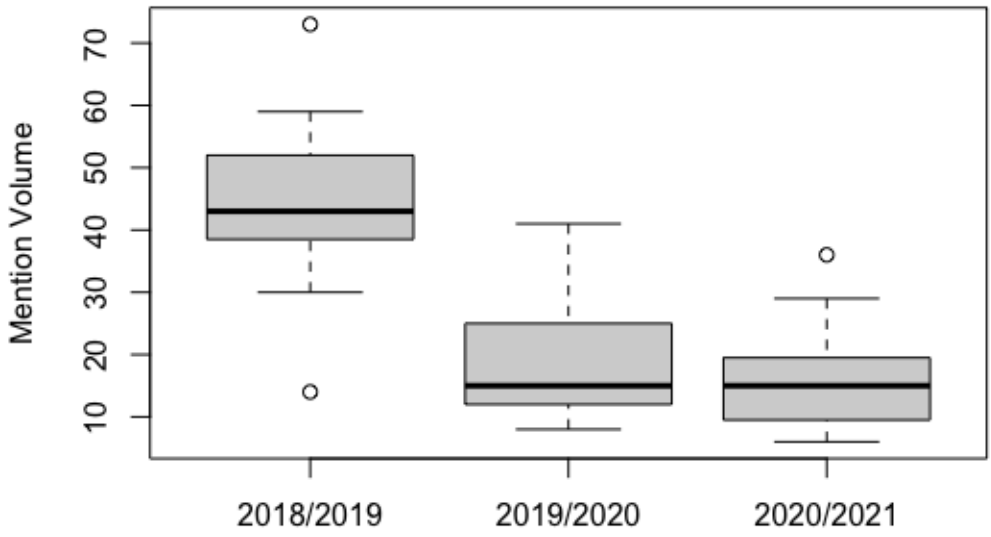

Figure 17. Non-Top 5 mentioned universities positive sentiment

4.4.6. Non-Top 5 mentioned universities neutral sentiment

Our analysis indicated that STEM and pre-health student non-top 5 mentioned universities related online mentions with neutral sentiment fluctuated over the thirty-five-month collection period. 7673 total mentions were made that indicated that a prospective or admitted student discussed a non-top 5 mentioned university that was within the Boolean query terms. Out of the 7673 mentions, 3672 mentions had neutral sentiment associated with the mention, accounting for 47.9% of all mentions (Table 18, Figure 18). From the 2018/2019 academic year, there were 1324 neutrally associated mentions involving a non-top 5 mentioned university. There was an increase of 4 neutral mentions from the first academic year to the second academic year, an increase of 0.3%. From the 2019/2020 academic year to the third academic year studied, a decrease of 308 neutral mentions occurred, resulting in a loss of 23.2% from the second academic year to the third. The null hypothesis was rejected, indicating that the mean ranks of the three compared groups (the three academic years) were not the same; there was a difference

between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.02317, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 18. Non-Top 5 mentioned universities neutral sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Non-Top 5 Universities Neutral Sentiment	7673	1324	1328	1020	7.5302	0.02317

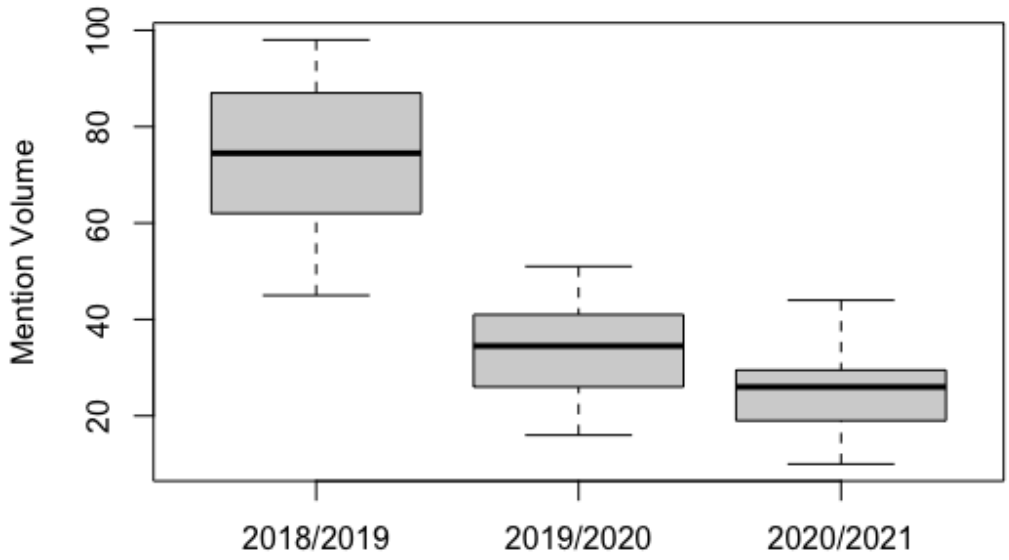

Figure 18. Non-Top 5 mentioned universities neutral sentiment

4.4.7. Non-Top 5 mentioned universities negative sentiment

Our analysis indicated that STEM and pre-health student non-top 5 mentioned universities related online mentions with negative sentiment decreased over the thirty-five-month collection period. 7673 total mentions were made that indicated that a prospective or admitted student discussed a non-top 5 mentioned university that was

within the Boolean query terms. Out of the 7673 mentions, 1656 mentions had negative sentiment associated with the mention, accounting for 21.6% of all mentions (Table 19, Figure 19). From the 2018/2019 academic year, there were 663 negatively associated mentions involving a non-top 5 mentioned university. There was a decrease of 93 negative mentions from the first academic year to the second academic year, a decrease of 14.0%. From the 2019/2020 academic year to the third academic year studied, a decrease of 147 negative mentions occurred, resulting in a loss of 25.8% from the second academic year to the third. The null hypothesis was not rejected, indicating that the mean ranks of the three compared groups (the three academic years) were the same; there was not a difference between the mean ranks of at least two of the three groups in this test. The p-value of the Kruskal-Wallis test was 0.1616, a value greater than the set alpha value of 0.05, indicated no statistical significance within the difference between the means of the compared groups.

Table 19. Non-Top 5 mentioned universities negative sentiment

	Total Mentions (n)	Mention Volume 2018/19	Mention Volume 2019/20	Mention Volume 2020/21	Kruskal-Wallis chi-squared (H)	p-value
Non-Top 5 Universities Negative Sentiment	7673	663	570	423	3.6448	0.1616

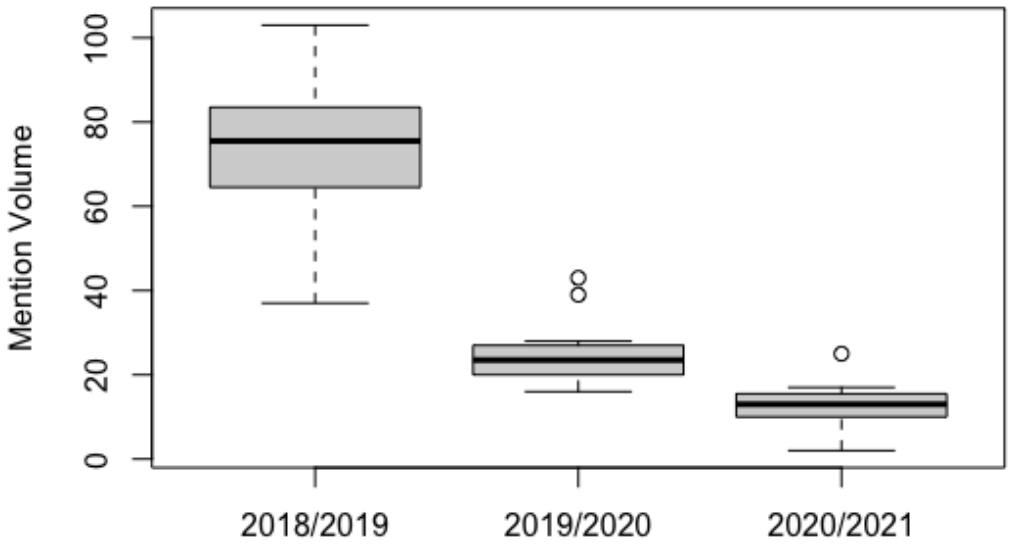

Figure 19. Non-Top 5 mentioned universities negative sentiment

4.5. How did prospective and admitted STEM and pre-health post-secondary students perception of STEM and pre-health change pre- versus post-pandemic?

4.5.1. Pre- and post-pandemic positive sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with positive sentiment decreased from pre- to post-pandemic announcement. 4457 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 4457 mentions, 1076 mentions had positive sentiment associated with the mention, accounting for 15.5% of all mentions (Table 20, Figure 20). There was a decrease of 302 positive mentions from the pre-pandemic to post-pandemic period, a decrease of 43.8%. The null hypothesis was rejected, indicating that the mean ranks of the 2 compared groups (pre- and post-pandemic) were not the same; there was a difference between the mean ranks of the two groups in this test. The p-value of the Kruskal-Wallis test was 0.04856, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 20. STEM and pre-health pre- and post-pandemic mentions with positive sentiment

	Total Mentions (n)	Pre-pandemic Mention Volume	Post-pandemic Mention Volume	Kruskal-Wallis chi-squared (H)	p-value
Pre- and Post-pandemic Positive Sentiment	4457	689	387	3.8906	0.04856

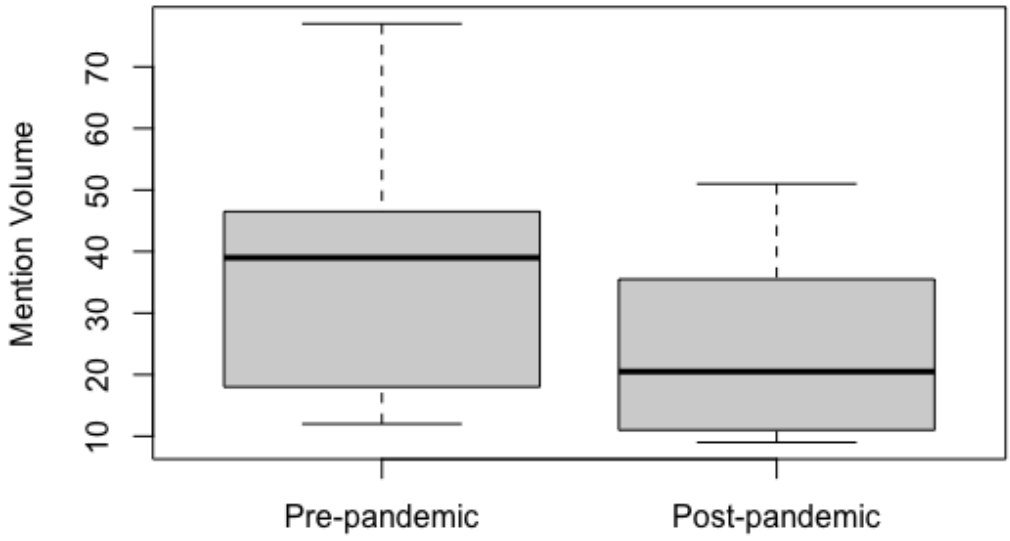

Figure 20. STEM and pre-health pre- and post-pandemic mentions with positive sentiment

4.5.2. Pre- and post-pandemic neutral sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with neutral sentiment decreased from pre- to post-pandemic announcement. 4457 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 4457 mentions, 1931 mentions had neutral sentiment associated with the mention, accounting for 43.4% of all mentions (Table 21, Figure 21). There was a decrease of 661

neutral mentions from the pre-pandemic to post-pandemic period, a decrease of 51.0%. The null hypothesis was rejected, indicating that the mean ranks of the 2 compared groups (pre- and post-pandemic) were not the same; there was a difference between the mean ranks of the two groups in this test. The p-value of the Kruskal-Wallis test was 0.000774, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 21. STEM and pre-health pre- and post-pandemic mentions with neutral sentiment

	Total Mentions (n)	Pre-pandemic Mention Volume	Post-pandemic Mention Volume	Kruskal-Wallis chi-squared (H)	p-value
Pre- and Post-pandemic Neutral Sentiment	4457	1296	635	11.303	0.000774

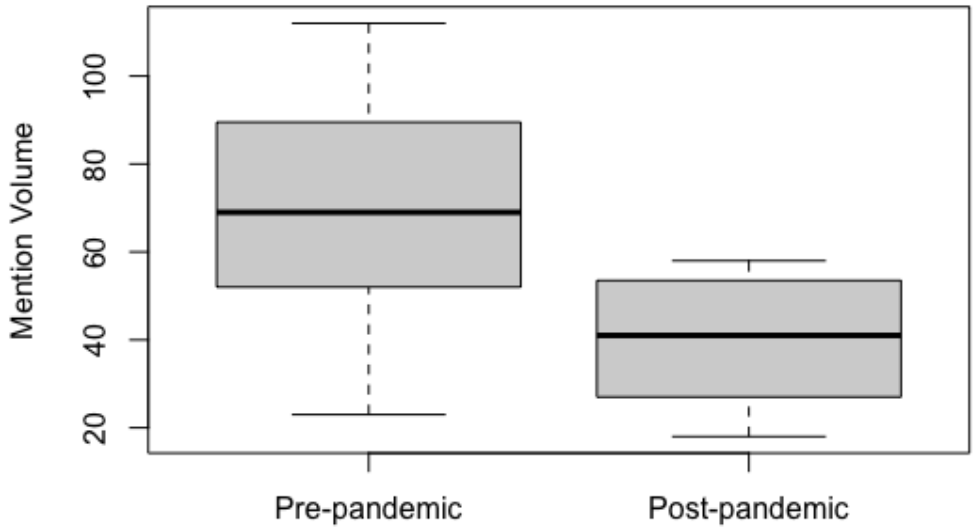

Figure 21. STEM and pre-health pre- and post-pandemic mentions with neutral sentiment

4.5.3. Pre- and post-pandemic negative sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with negative sentiment decreased from pre- to post-pandemic announcement. 4457 total mentions were made that indicated that a prospective or admitted student discussed a STEM or pre-health topic that was within the Boolean query terms. Out of the 4457 mentions, 1450 mentions had negative sentiment associated with the mention, accounting for 32.5% of all mentions (Table 22, Figure 22). There was a decrease of 744 negative mentions from the pre-pandemic to post-pandemic period, a decrease of 57.8%. The null hypothesis was rejected, indicating that the mean ranks of the 2 compared groups (pre- and post-pandemic) were not the same; there was a difference between the mean ranks of the two groups in this test. The p-value of the Kruskal-Wallis test was 0.0001925, a value less than the set alpha value of 0.05, indicated statistical significance within the difference between the means of the compared groups.

Table 22. STEM and pre-health pre- and post-pandemic mentions with negative sentiment

	Total Mentions (n)	Pre-pandemic Mention Volume	Post-pandemic Mention Volume	Kruskal-Wallis chi-squared (H)	p-value
Pre- and Post-pandemic Negative Sentiment	4457	1097	353	13.903	0.0001925

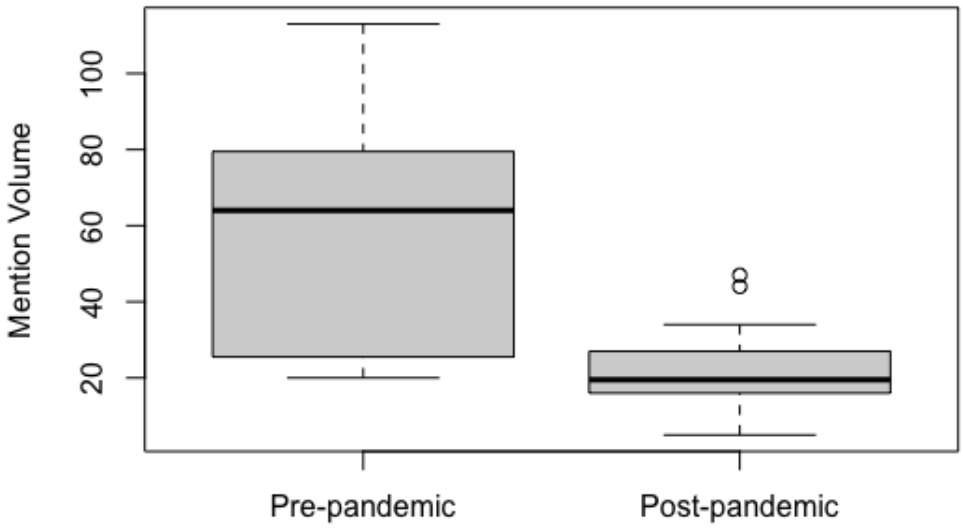

Figure 22. STEM and pre-health pre- and post-pandemic mentions with negative sentiment

4.6. How does geographical location/region within the United States affect STEM and pre-health student sentiment of STEM and pre-health?

4.6.1. Geographical location positive sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with positive sentiment varied across the 9 designated regions of the United States. The designated regions were as follows: FW indicated the Far West, GL indicated the Great Lakes area, ME indicated Mid East, NE signified the New England area, OA signified Outlying Areas, P signified the Plains, RM indicated the Rocky Mountains, SE indicated the Southeast, and SW signified the Southwest of the United States (Table 23, Figure 23). 2379 total mentions were made that indicated that a prospective or admitted student stated that they were in a state, city, or region that was within the Boolean query terms. Out of the 2379 mentions, 560 mentions had positive sentiment associated with the mention, accounting for 23.5% of all mentions (Table 23, Figure 23). Due to the discrepancies between population of each area that could not be accounted for in the analysis, no conclusion can be drawn from this data. Further investigation could allow for conclusions so be drawn regarding how geographical location within the United States could affect STEM and pre-health sentiment.

Table 23. Positive sentiment by geographic location in the United States

	Total Mentions (n)	FW	GL	ME	NE	OA	P	RM	SE	SW
Geographic Location Positive Sentiment	2379	99	63	154	51	0	10	19	93	71

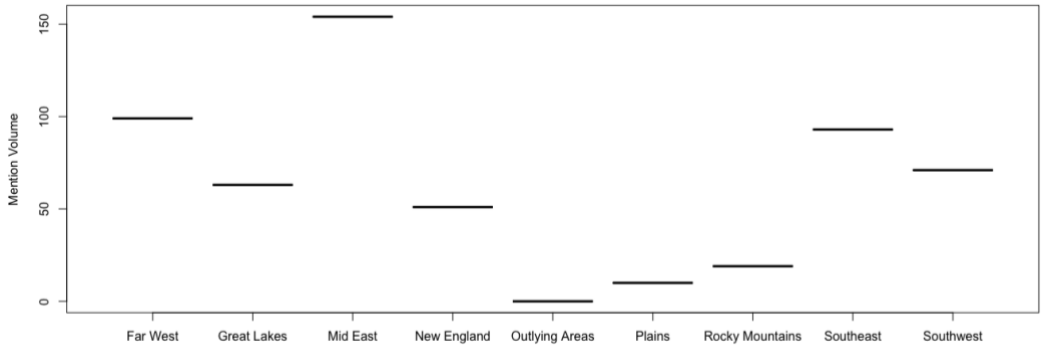

Figure 23. Positive sentiment by geographic location in the United States

4.6.2. Geographical location neutral sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with neutral sentiment varied across the 9 designated regions of the United States. The designated regions were as follows: FW indicated the Far West, GL indicated the Great Lakes area, ME indicated Mid East, NE signified the New England area, OA signified Outlying Areas, P signified the Plains, RM indicated the Rocky Mountains, SE indicated the Southeast, and SW signified the Southwest of the United States (Table 24, Figure 24). 2379 total mentions were made that indicated that a prospective or admitted student stated that they were in a state, city, or region that was within the Boolean query terms. Out of the 2379 mentions, 1196 mentions had neutral sentiment associated with the mention, accounting for 50.3% of all mentions (Table 24, Figure 24). Due to the discrepancies between population of each area that could not be accounted for in the

analysis, no conclusion can be drawn from this data. Further investigation could allow for conclusions so be drawn regarding how geographical location within the United States could affect STEM and pre-health sentiment.

Table 24. Neutral sentiment by geographic location in the United States

	Total Mentions (n)	FW	GL	ME	NE	OA	P	RM	SE	SW
Geographic Location Neutral Sentiment	2379	285	137	350	161	2	18	11	146	86

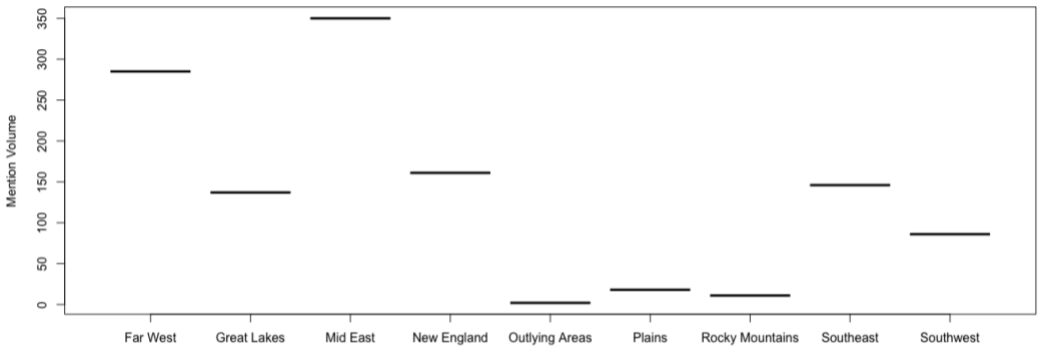

Figure 24. Neutral sentiment by geographic location in the United States

4.6.3. Geographical location negative sentiment mentions

Our analysis indicated that STEM and pre-health student online conversation volume with negative sentiment varied across the 9 designated regions of the United States. The designated regions were as follows: FW indicated the Far West, GL indicated the Great Lakes area, ME indicated Mid East, NE signified the New England area, OA signified Outlying Areas, P signified the Plains, RM indicated the Rocky Mountains, SE indicated the Southeast, and SW signified the Southwest of the United States (Table 25, Figure 25). 2379 total mentions were made that indicated that a prospective or admitted

student stated that they were in a state, city, or region that was within the Boolean query terms. Out of the 2379 mentions, 623 mentions had negative sentiment associated with the mention, accounting for 26.2% of all mentions (Table 25, Figure 25). Due to the discrepancies between population of each area that could not be accounted for in the analysis, no conclusion can be drawn from this data. Further investigation could allow for conclusions so be drawn regarding how geographical location within the United States could affect STEM and pre-health sentiment.

Table 25. Negative sentiment by geographic location in the United States

	Total Mentions (n)	FW	GL	ME	NE	OA	P	RM	SE	SW
Geographic Location Negative Sentiment	2379	147	42	140	48	0	14	5	135	92

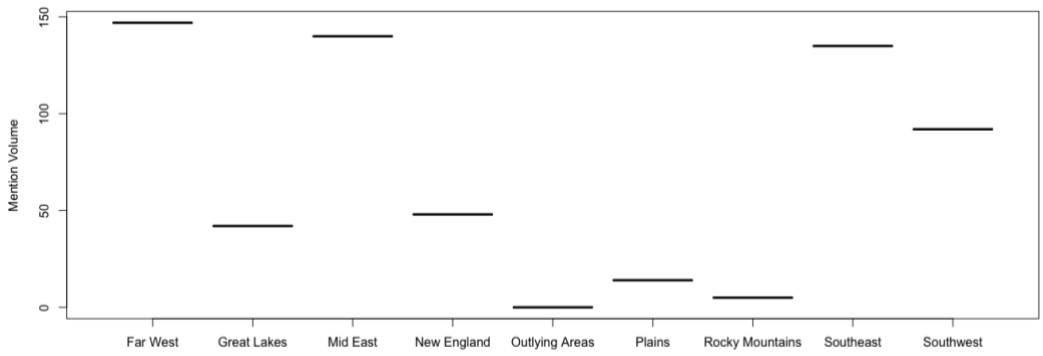

Figure 25. Negative sentiment by geographic location in the United States

5. Conclusion

Table 26. Sentiment changes and statistical significance

	Student Type	Sentiment	Significance Direction	Significance
STEM + Pre-Health Majors + Careers	Prospective	Positive	Negative	*
		Neutral	Negative	*
		Negative	Negative	*
	Admitted	Positive	Not Significant	ns
		Neutral	Negative	*
		Negative	Negative	*
	Prospective + Admitted	Positive	Negative	*
		Neutral	Negative	*
		Negative	Negative	*
Majors	Prospective + Admitted	Positive	Negative	*
		Neutral	Negative	*
		Negative	Negative	*
Careers	Prospective + Admitted	Positive	Not Significant	ns
		Neutral	Not Significant	ns
		Negative	Negative	*
Top 5 Universities	Prospective + Admitted	Positive	Negative	*
		Neutral	Negative	*
		Negative	Negative	*
Non-Top 5 Universites	Prospective + Admitted	Positive	Not Significant	ns
		Neutral	Negative	*
		Negative	Not Significant	ns
Pre- and post-COVID-19 pandemic	Prospective + Admitted	Positive	Negative	*
		Neutral	Negative	*
		Negative	Negative	*

In this study, 16 of the 22 Kruskal-Wallis tests performed on the data resulted in statistically significant data, indicating that the p-values of the 16 tests were equal to or below the defined alpha value of 0.05. All 16 statistically significant data sets had a decrease in the mean mention volume from the start of the data collection period to the

end of the study. The remaining 6 Kruskal-Wallis test results were neutral, with no statistical significance between the mean mention volumes between the compared values, indicated by a p-value greater than the alpha value of 0.05. Data that had the smallest p-values had the greatest difference between mean mention volumes within the Kruskal-Wallis test. The p-values that were less than or equal to 0.05 were signified with a single star, with p-values greater than 0.05 signified with ns, no significance (Table 26).

One research question that had negative statistical significance in positive, neutral, and negative sentiment was that of prospective student mentions, with positive sentiment at a p-value of 3.783×10^{-5}, neutral sentiment at a p-value of 1.705×10^{-5}, and negative sentiment at a p-value of 1.658×10^{-6}. This indicates that the mean mention volume significantly decreased across the three academic years that were compared. In comparison, admitted student mentions with neutral and negative sentiments were the data sets that resulted in a statistically significant decrease in mean mention volume across the study, with a p-value of 0.01992 for neutral sentiment mentions and a p-value of 1.584×10^{-4} for negative sentiment mentions. The p-value of admitted student positive mentions was 0.1574, indicating that over the three academic years of data collection, positive sentiment associated with admitted STEM and pre-health student mentions did not significantly decrease.

Academic major positive, neutral, and negative sentiment mean mention volume also significantly decreased across the study, with p-values of 0.01105, 0.004906, and 4.698×10^{-5} respectively. Due to the p-values of all three sentiments being less than or equal to the alpha of 0.05, it can be concluded that there was a significant decrease in mean mention volume across all sentiments regarding academic majors.

Negative sentiment regarding STEM and pre-health careers also significantly decreased in mean mention volume across the 35 months analyzed, with a p-value of 0.001127 indicating that a statistically significant decrease occurred. The p-value of positive sentiment career mentions was 0.3013, indicating that no significant decrease in online discussion of careers with positive sentiment occurred. Similarly, STEM and pre-health career conversation with neutral sentiment had no significant decrease in volume, with a p-value of 0.4094 indicating that neutral sentiment associated with careers stayed steady across the study duration. The significant decrease in only negative sentiment regarding STEM and pre-health careers indicated that there was less negative conversation about careers across the three years.

The top 5 most mentioned universities across the data collection period had a statistically significant decrease in the positive, neutral, and negative sentiment mean mention volumes. Respective p-values of 0.0001569, 6.333×10^{-6}, and 8.198×10^{-7} were obtained using the Kruskal-Wallis test, all three of which are below the alpha of 0.05, indicating that all three sentiments had a significant decline in the mean mention volume across the study.

The non-top 5 most mentioned universities across the period of the study had a statistically significant decrease in mean mention volume in only the neutral sentiment mentions, with a p-value of 0.02317 indicating such. Upon comparing the top-5 and non-top 5 university data, it was seen that the top 5 universities had a significant decrease in positive sentiment while non-top 5 universities did not have a significant decrease in positive sentiment.

Social listening is a technology that can be used by universities to predict quantities of future STEM and pre-health students, as well as the sentiment the incoming students have about specific majors, careers, and universities themselves. Insight gained through social listening can enable university recruiters to gain a better understanding of the number of students to expect in the upcoming year, as well as specifically target areas with high negative sentiment to attempt to appeal to more students.

In this study, many of the research questions' conversation volumes decreased between the three academic years observed. The observed general decrease in online STEM and pre-health conversation volume could be influenced by many factors, with pre- versus post-pandemic announcement volume differences among the most notable. The COVID-19 pandemic was an extraneous factor, unable to be predicted within the academic or university setting, but seemingly had an impact on prospective or admitted STEM or pre-health student conversation, which could potentially equate to less STEM and pre-health students in higher education. Due to the pandemic, stability and increase of interest in STEM and healthcare careers are more important now than ever before.

The changes in mention volume and sentiment of online conversations carried out by STEM and pre-health pre-college students suggests that there is potential in the use of social listening by university leaders to better understand prospective student career interests more rapidly and conclusively in the highly dynamic and heavily online-based world. Further research could assist in proving the suggested utility, something even more likely if other factors such as enrollment, science and pre-health studies, health and science careers, socioeconomic status, geographical location, and racial identities were investigated. Geographical location data might prove the most helpful for universities and

their admission and outreach efforts, with specific regions being studied assisting in associating sentiment to their university. In this study, too little geographical data was procured, with even less ability to work with said data to produce significant statements or interpretations of the data.

6. Conflicts of Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

APPENDIX

HOW PROSPECTIVE POST-SECONDARY STUDENT PERCEPTION OF STEM AND PRE-HEALTH CAREERS CHANGED PRE- AND POST-PANDEMIC

Susan Roh[1], Greg Heiberger[1], and Katlin Swisher[2]

[1] Department of Biology and Microbiology, South Dakota State University, Brookings, SD

[2] Campus Sonar, 38 Buttonwood Court, Madison, WI, 53718

Abstract. The purpose of this study is to determine the STEM and pre-health online mention volume change over a pre-pandemic period as compared to a post-pandemic period, with the data collection period encompassing August of 2018 through June of 2021. We also evaluated if the mention sentiment shifted between the periods. The data gathered for this study was procured in collaboration with Campus Sonar and the social-listening tool of Brandwatch, which allowed for specifications of the data collected. Kruskal-Wallis tests were used to determine the means of the pre- versus post-pandemic mentions and mean sentiments volumes. Over the span of the 35-month data gathering period, spanning from August of 2018 until June of 2021, there was a difference of 1,707 mentions, decreasing from 3,082 pre-pandemic to 1,375 post-pandemic. Positive mention sentiment changed from accounting for 22.06% of pre-pandemic mentions to 28.15% of post-pandemic mentions. Inversely, negative mention sentiment shifted from contributing to 35.59% of pre-pandemic mentions to 25.67% of post-pandemic mentions. This change in not only online conversation volume but the sentiment associated with each STEM and pre-health mention signals the utility of social listening by university leaders to

understand prospective student career interest more rapidly in the current and future highly dynamic world.

Keywords: STEM; pre-health; COVID; COVID-19; pre-pandemic; post-pandemic

1. Introduction

Social listening is an emerging technology that broadens the ability for online interest tracking, with uses postulated within organizational communication, including university admissions. Social listening encompasses social media and online communication sites, with the general intention of better understanding individuals via their self-admitted information they post publicly. From the level of the individual, social listening can be used to form groups that encompass similar interests or attributes that the individuals freely stated. These groupings can then be utilized within data analyses to determine if there are trends within the populations, of which may be helpful for organizations to better target the interests or better understand the issues the grouping has caused for the individuals within it and thus address them in a more personalized manner.

Due to extraneous factors that occurred between the data collection period of August 2018 through June 2021, stability and growth of interest in STEM and healthcare careers are more important now than ever before. COVID-19 started to widely affect the United States and its universities in March 2020, and its lasting impact is still felt today within the entire country. This pandemic possesses massive potential to have thus altered prospective college student opinions on STEM and pre-health related majors, and career perspectives. One way to predict the amount of future or incoming STEM and pre-health students to the university academic programs is using social listening. Social listening gives insight into trends of conversation volume on the topic of STEM and pre-health,

sentiment about STEM and pre-health majors, and careers, as well as how these conversations have changed in a pre- versus post-pandemic world. This insight can enable university recruiters and the healthcare field to gain a better understanding of not only the number of students or future career professionals to expect, but what these students are excited or concerned about as they prepare for their future careers

2. Materials and Methods

2.1 Study Design

This study analyzed online conversation about incoming STEM and pre-health undergraduate students' college going experiences to gain a deeper understanding of their journeys to higher education – their motivations, obstacles, and opportunities. This research highlights first-person perspectives from more than 5,000 STEM and pre-health students over a three-year collection period to help campuses meet them where they are at and provide focused support to guide them on pathways to success, all while informing strategies to develop the STEM and pre-health pipeline. This study aims to understand STEM and pre-health students' evolving perspectives about, motivations for, and obstacles to enrolling in college. The study is also exploring how social listening captures authentic student conversations about perceptions of college and identifying strategies to best meet student needs[1]. Similarly, the study explores how STEM and pre-health student conversation volume decreased from pre-pandemic to post-pandemic periods within the data collection window and how the sentiment of these conversations changed within that same period.

This study achieved the goals stated above by collaboration with Campus Sonar, who utilizes Brandwatch Consumer Research, an enterprise-level social listening software

Campus Sonar staff used to gather the data for analysis. Brandwatch has access to over 100 million publicly-available online sources and over 1.3 trillion individual posts across a variety of social media sites, including Twitter, Reddit, Instagram, Tumblr, and YouTube, as well as non-social sites such as blogs, news outlets, forums, and review sites. From these many publicly available sources and individual posts, Brandwatch can be used to filter the data to fit the study at hand. Within this study, the research team identified unique key terms and phrases and built a Boolean query[2] to collect online social data from the Brandwatch archive.

2.2 Experimental/Data Limits and Query Factors

For this study, data collected within the Brandwatch Boolean had to be validated via the use of four conditions of which each data piece had to fit within the boundaries[3]. The first condition that data had to meet was that of the mention being in first-person and stated as being from a prospective or admitted undergraduate student who wants to attend a higher education institution within the United States. The second condition was that the mention had to be from a publicly accessible site such as the social media sites stated above, or a blog or forum site as designated by the Brandwatch Consumer Research software. The third condition was that the mention could not be a retweet, as this is not inherently indicative of the person who retweeted it also fitting within the criteria laid out within the first condition. The fourth and final condition for a mention to be included within the study's dataset was that the mention had to originate from a location within the United States or indicate within the mention text itself that the student was located within the United States. This fourth condition was achieved using Brandwatch leveraging any location metadata included with each mention to determine whether it originated in the

United States, filtering with the top 200 United States cities by population, and any references of nine set geographic regions (New England, Mid East, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, Far West, and Outlaying Areas).

Once each data point was determined to likely be from a student within the United States, additional steps were taken by the Campus Sonar team to further validate the data. A propriety Boolean was used to identify common irrelevancies within the entire dataset, such as international higher education terms, historical/reminiscent conversation, already graduated, post graduate parents and students, and undergraduate parents. Top phrases, top sites, and top authors were also reviewed for the dataset[4]. Each conversation also had to return for the research questions examined by site and audience, such as: prospective student + health careers + Reddit.

The dataset was segmented to allow for the analysis of research questions more easily, with Campus Sonar creating a set of Boolean rules that coded the data by enrollment, science and pre-health studies, health and science careers, socioeconomic status, and racial identities. Enrollment was defined as either prospective or admitted students, with prospective broken down further into subcategories of prospective, visit, and applicant. Science and pre-health studies was defined as the area of study for science and pre-health students, with science majors identifying conversations within STEM and pre-health studies that contained major-related terms. Health and science careers were defined as containing careers that STEM and pre-health majors tend to enter post-graduation.

Limitations within the dataset included the unpredictability and idiosyncrasies of language use within conversations, thus the inability to have definite sentiment intention on every datapoint[5]. To try to counteract this, the research team utilized a list of key terms when crafting the Boolean query and rules based on industry expertise, secondary research, client conversations, and a variety of tests to ensure that a robust and relevant sample of social data was collected. Another limitation was due to privacy terms changes within Instagram, Tumblr, and YouTube that limited how much of the sites social listening software can be accessed and for how long. A third limitation related to the collection of geographic location based on the social data, as not all data collected has associated geographic indicators. To counteract this, the Campus Sonar analysts applied linguistic and contextual tactics to identify mentions in which the author self-identifies as bring from the United States[6].

2.3 Analyses

The raw data within this study was analyzed using RStudio. The model utilized to analyze the data was a Kruskal-Wallis test due to the nature of the data's dependent variable being ordinal or ranked and there being one independent variable with 2+ categories or groups, of which are independent of the other categories or groups present. The independent variables within this study were defined as pre-pandemic and post-pandemic time periods, giving the two required groups to use a Kruskal-Wallis model. The dependent variables were the conversation sentiments of positive, neutral, and negative, as these values depended on the time periods for their change. An alpha of 0.05 was utilized to determine whether there was a significant difference observed between the groups being compared. After creating the null and alternate hypotheses and stating an

alpha value, RStudio was utilized to create a boxplot comparing the data in a visual form. Similarly, the Kruskal-Wallis test code was executed through the program to provide a Kruskal-Wallis chi-squared value, a degree of freedom for the data, and a p-value which then can be compared to the stated alpha to determine if there was a significant change between the two sets of data being compared. If the p-value is less than 0.05, the null hypothesis can be rejected, meaning that the mean ranks of the groups were not the same and there is a difference between the mean ranks of the two groups in the Kruskal-Wallis test. Alternatively, if the p-value is greater than 0.05, we fail to reject the null hypothesis, meaning that the mean ranks of the two groups were the same in the Kruskal-Wallis test.

3. Results

3.1 Mention Volume and Sentiment Change Over Pre- to Post-pandemic Period

Table 1. Pre- v. Post-pandemic Mention Volume and Sentiment Change Over Collection Period

Comparisons	**Total Mentions (n)**	**Pre Pand-emic**	**Post Pand-emic**	**Kruskal-Wallis chi-squared (H)**	**Degrees of Freedom (df)**	**p-value**
Total Mentions	4457			17.316	2	0.0001737***
Pre v Post Pandemic Positive	1076	689	387	3.8906	1	0.04856*
Pre v Post Pandemic Neutral	1931	1296	635	11.303	1	0.000774***
Pre v Post Pandemic Negative	1450	1097	353	13.903	1	0.0001925***

3.2 Mention Volume Decrease from COVID-19's Widespread Effects

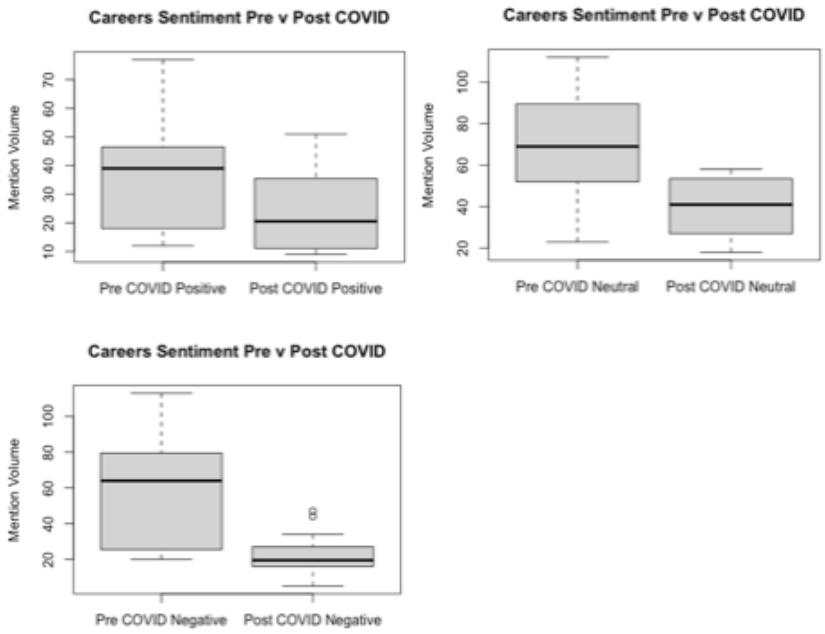

Figure 1. Pre- v. Post-pandemic Mention Volume for Positive, Neutral, and Negative Sentiments

4. Discussion

Over the span of the 35-month data gathering period, spanning from August of 2018 until June of 2021, there was a difference of 1,707 mentions, decreasing from 3,082 pre-pandemic to 1,375 post-pandemic, as seen in Table 1. Positive mention sentiment changed from accounting for 22.06% of pre-pandemic mentions to 28.15% of post-pandemic mentions. Inversely, negative mention sentiment shifted from contributing to 35.59% of pre-pandemic mentions to 25.67% of post-pandemic mentions, suggesting while there was less overall STEM and pre-health mentions online, there was a shift towards a more positive overall sentiment, approximately 6.09%, associated with the mentions that occurred in the post-pandemic collection period. Similarly, there was a

decrease in negative sentiment associated with the STEM and pre-health mentions, with a decrease of approximately 9.92% in negative sentiment from the pre- to post-pandemic collection period. Figure 1 supports this data, as the mean of the careers sentiment for the positive sentiment shows a decrease in average mention volume from approximately 40 positive mentions per month pre-pandemic to about 20 positive mentions per month post-pandemic. Similarly, the neutral sentiment mention mean decreased from around 70 mentions per month pre-pandemic to about 40 mentions per month post-pandemic. The negative sentiment mentions also decreased from the pre- to post-pandemic time periods, with a mean of approximately 65 negative mentions per month to a mean of about 20 negative mentions per month. All of the mention volume and sentiment mean decreases from pre- to post-pandemic were statistically significant, with all but the pre- versus post-pandemic positive sentiment volumes having three stars of significance, indicating an alpha value of less than 0.001.

5. Conclusions

Over the span of the 35-month data gathering period, spanning from August of 2018 until June of 2021, analysis of the data dictates that the appearance of COVID-19 on a widescale within the United States led to less STEM and pre-health pre-college student online conversation but an increase in the positive sentiment and decrease in the negative sentiment associated with the online conversations that occurred. The changes in mention volume and sentiment of online conversations carried out by STEM and pre-health pre-college students suggests that there is potential in the use of social listening by university leaders to better understand prospective student career interests more rapidly and conclusively in the highly dynamic and heavily online-based world. Further

research could assist in proving the suggested utility, something even more likely if other factors such as enrollment, science and pre-health studies, health and science careers, socioeconomic status, and racial identities were investigated.

Acknowledgements: Thank you to South Dakota State University for supporting this research and to our partners at Campus Sonar for their leadership, expertise, and collaboration.

Author Contributions: Katlin Swisher and the Campus Sonar team collected and refined data, Susan Roh and Greg Heiberger analyzed data, and Susan Roh wrote the manuscript.

Conflicts of Interest: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

1. Gross, L. (2019). *The higher ed social listening handbook.* Retrieved from: https://info.campussonar.com/handbook
2. Gross, L. (2021). *What the Boolean! Getting the most out of your social listening software*. Retrieved from https://blog.campussonar.com/blog/what-the-boolean-get-the-most-out-of-your-social-listening
3. Jeppson, A. (2021). *Doing it by the book: Data segmentation, validation, and analysis*. Retrieved from https://blog.campussonar.com/blog/doing-it-by-the-book-data-segmentation-validation-and-analysis
4. Sandall, A. (2020). *Puzzling through social data source*. Retrieved from https://blog.campussonar.com/blog/puzzling-through-social-data-sources

5. Sandall, A. (2021). *Seven reasons why human analysts are critical to social listening.* Retrieved from https://blog.campussonar.com/blog/six-reasons-why-human-analysts-are-critical-to-social-listening
6. Miller, B. & Sandall, A. (2020). *Digging into social data—Indicators and context clues*. Retrieved from https://blog.campussonar.com/blog/digging-into-social-data

REFERENCES

1. Alwi, S. F. S., & Kitchen, P. J. (2014). *Projecting corporate brand image and behavioral response in business schools: Cognitive or affective brand attributes?* Journal of Business Research, 67(11), 2324–2336. doi:10.1016/j.jbusres.2014.06.020
2. Barlow, M., & Thomas, D. B. (2011). *The executive's guide to enterprise social media strategy: how social networks are really transforming your business*. New Jersey: John Wiley.
3. Barnes, N. G., & Jacobsen, S. L. (2014). *Missed eWOM opportunities: A cross-sector analysis of online monitoring behavior*. Journal of Marketing Communications, 20(1–2), 147–158. doi:10.1080/13527266.2013.797788
4. Bennett, L., & Nair, C. S. (2010). *A recipe for effective participation rates for web-based surveys*. Assessment and Evaluation in Higher Education, 35, 357–365.
5. Bolat, E., & O'Sullivan, H. (2017). *Radicalising the marketing of Higher Education: Learning from Student-Generated Social Media Data.* Retrieved from http://eprints.bournemouth.ac.uk/24441/3/Radicalising%20the%20marketing%20of%20higher%20education.pdf
6. Calder, B. J., Malthouse, E. C., & Maslowska, E. (2016). *Brand marketing, big data and social innovation as future research directions for engagement*. Journal of Marketing Management, 32(5–6), 579–585. doi:10.1080/0267257X.2016.1144326

7. Coates, H. (2006). *Student engagement in campus-based and online education: university connections*. London: Taylor and Francis.
8. Douglas, J., Douglas, A., & Barnes, B. (2006). *Measuring student satisfaction at a UK university. Quality Assurance in Education*, 14, 251–267.
9. Finch, D., McDonald, S., & Staple, J. (2013). *Reputational interdependence: an examination of category reputation in higher education*. Journal of Marketing for Higher Education, 23(1), 34–61.
10. Fotis, J., Buhalis, D., Rossides, N., 2012. *Social media use and impact during the holiday travel planning process.* In: Fuchs, M., Ricci, F., Cantoni, L. (Eds.), Information and Communication Technologies in Tourism. Springer-Verlag, Vienna, Austria, pp. 13–24.
11. Fuchs, C. (2015) *Culture and Economy in the Age of Social Media*. New York: Routledge.
12. Fuchs, C. (2017). *From Digital positivism and administrative big data analytics towards critical digital and social media research!* Retrieved from https://journals.sagepub.com/doi/full/10.1177/0267323116682804
13. Ganim Barnes, N., & Mattson, E. (2009) *Social media and college admissions: The first longitudinal study.* Retreived from http://www.umassd.edu/cmr/studiesresearch/mediaandadmissions.pdf
14. Gillin, P. (2008). *New media, new influencers and implications for public relations.* Retreived from http://sncr.org/wp-content/uploads/2008/08/new-influencers-study.pdf

15. Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). *Attitudes toward the ethics of research using social media: A systematic review*. Retrieved from https://pubmed.ncbi.nlm.nih.gov/28588006/

16. Greenhow, C., Cho, V., Dennen, V. P., & Fishman, B. J. (2019). *Education and social media: Research Directions to guide a growing field*. Retrieved from https://www.researchgate.net/publication/356330867_Education_and_Social_Media_Research_Directions_to_Guide_a_Growing_Field

17. Gross, L., & Horn, M. (2017). *Opportunities for social listening in higher education*. HSTalks. Retrieved from https://hstalks.com/article/1621/opportunities-for-social-listening-in-higher-educa/

18. Gross, L. (2019). *The Higher Ed Social Listening Handbook*. Campus Sonar. Retrieved from https://info.campussonar.com/handbook

19. Gross, L. (2021). *What the Boolean! Getting the most out of your social listening software*. Retrieved from https://blog.campussonar.com/blog/what-the-boolean-get-the-most-out-of-your-social-listening

20. Jeppson, A. (2021). *Doing it by the book: Data segmentation, validation, and analysis*. Retrieved from https://blog.campussonar.com/blog/doing-it-by-the-book-data-segmentation-validation-and-analysis

21. Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). *Social Media? get serious! understanding the functional building blocks of social media*. Business Horizons. Retrieved from https://www.sciencedirect.com/science/article/pii/S0007681311000061

22. Koshkin, A. P., Rassolov, I. M., & Novikov, A. V. (2016). *Monitoring Social Media: Students Satisfaction With University administration activities - education and Information Technologies*. SpringerLink. Retrieved from https://link.springer.com/article/10.1007/s10639-016-9558-x
23. Kuo, Y. C., Walker, A. E., Belland, B. R., & Schroder, K. E. (2013). *A predictive study of student satisfaction in online education programs.* The International Review of Research in Open and Distributed Learning, 14(1), 16–39.
24. Lowendahl, J., Thayer, T. B. and Morgan, G. (2016). *Top 10 strategic technologies impacting higher education in 2017*. Gartner, 28th December, p. 31. Retrieved from www.gartner.com/newsroom/ id/3225717
25. Miller, B. & Sandall, A. (2020). *Digging into social data—Indicators and context clues*. Retrieved from https://blog.campussonar.com/blog/digging-into-social-data
26. National Association for College Admissions Counseling. (2009*). Reports finds use of social networking tools on the rise in college admissions offices*. [Press Release]. Retreived from http://www.nacacnet.org/AboutNACAC/PressRoom/2009/Pages/SocialNetworking.aspx
27. Ng, P., Lee, D., Wong, P., & Lam, R. (2020). *Making a higher education institution choice: Differences in the susceptibility to online information on students' advice-seeking behavior*. Online Information Review. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/OIR-07-2019-0218/full/html

28. Number of social media users worldwide from 2010 to 2020 (in billions), 2016. Retrieved from https://www.statista.com/statistics/278414/number-of- worldwide-social-network-users/.
29. Olson, D. & Martin, Q. (2010). *Engaging college students through online social networks.* The Journal of Technology in Student Affairs, 11(1). Retrieved from http://studentaffairs.com/ejournal/Winter_2010/EngagingCollegeStudents.html
30. Ouellette, M. M. (2018). *Higher education, social media and an era of emergencies: To*. Journal of Education Advancement & Marketing. Retrieved from https://ideas.repec.org/a/aza/jeam00/y2019v4i2p165-174.html
31. Peruta, A., & Shields, A. B. (2016*). Social media in higher education: understanding how colleges and universities use Facebook*. Journal of Marketing for Higher Education. doi:10.1080/08841241.2016.1212451.
32. Ratliff, A. F. (2011). *Are they listening? social media on campuses of Higher Education*. Retrieved from https://www.researchgate.net/publication/290901972_Are_they_listening_Social_media_on_campuses_of_higher_education
33. Ruths, D., & Pfeffer, J. (2014). *Social Media for large studies of behavior*. Retrieved from https://www.science.org/doi/10.1126/science.346.6213.1063
34. Sandall, A. (2020). *Puzzling through social data sources*. Retrieved from https://blog.campussonar.com/blog/puzzling-through-social-data-sources
35. Sandall, A. (2021). *Seven reasons why human analysts are critical to social listening*. Retrieved from https://blog.campussonar.com/blog/six-reasons-why-human-analysts-are-critical-to-social-listening

36. Schoen, H., Gayo-Avello, D., Takis Metaxas, P., Mustafaraj, E., Strohmaier, M., & Gloor, P. (2013). *The power of prediction with social media*. Retrieved from https://www.researchgate.net/publication/263231768_The_power_of_prediction_with_social_media

37. Shah, M., & Nair, C. S. (2009). *Using student voice to improve student satisfaction: two Australian universities the same agenda*. Journal of Institutional Research (South East Asia), 7, 43–55.

38. Smith, A. (2015). *Chapter three: A "Week in the Life" analysis of smartphone users*. Retrieved from http://www.pewinternet.org/2015/04/01/chapter-three-a-week-in-the-life-analysis-of-smartphone-users/

39. Stewart, M. C., & Arnold, C. L. (2018). *Defining social listening: Recognizing an emerging dimension of listening.* Retrieved from https://www.tandfonline.com/doi/full/10.1080/10904018.2017.1330656

40. TrackMaven. (2015). *Social listening*. Retrieved from http://trackmaven.com/marketing-dictionary/social-listening/

41. Underwood, S. J., & Austin, C., & Giroir, C. (2010). *Squeezing the virtual turnip: Introducing student affairs professionals to open source technologies*. The Journal of Technology in Student Affairs. 11(1). Retrieved from http://studentaffairs.com/ejournal/Winter_2010/SqueezingtheVirtualTurnip.html

42. Walters, S. (2013). *Beyond listening: Six steps for integrating and acting on social media*. Business Intelligence Journal, 18(1), 13–17.

43. Wankel, L. A., & Wankel, C. (2011). *Higher Education Administration with social media: Including applications in Student Affairs, Enrollment Management,*

alumni relations, and career centers. Google Books. Retrieved from https://books.google.com/books/about/Higher_Education_Administration_with_Soc.html?id=Z6wrQdcPdhgC

44. Wong, P., Ng, P.M., Mak, C.K. and Chan, J.K. (2016), *Students' choice of sub-degree programmes in self-financing higher education institutions in Hong Kong*. Higher Education, Vol. 71 No. 4, pp. 455-472.