Impact of COVID-19 on Professor Reviews:

Study of reviews from the Rate My Professor website

Yu (Eason) Tsao, Hee Jae Yoon, Yuewen Su, Aditya Dutt, Aakanksha Sangwan

University of Southern California

DSO 560: Text Analytics and Natural Language Processing

Dr. Lorena Martin

April 29, 2024

1. Abstract

This paper goes over the effects of the COVID-19 pandemic on student learning by analyzing text student review data on the Rate My Professors website. Specifically, this analysis examines whether review sentiment was more negative during the pandemic, the correlation between review sentiment and grade/quality/difficulty, and if pandemic sentiment decreased differently by department (engineering/computer science < business/economics < theater). Utilizing, word clouds, sentiment analysis, T-test, semantic analysis, and LDA topic modeling we saw COVID-19 did not cause a statistically significant change in review sentiment, there was a positive correlation between sentiment and grade/quality and a negative correlation with difficulty, and department was a confounding variable in pandemic sentiment. Specifically, the business department had the greatest drop in sentiment, followed by engineering and computer science, with a slight drop in economics, and an increase in theater professor review sentiment all because of the pandemic. From this analysis, it is recommended that professors focus on balancing coursework with interactiveness, cater learning styles to the specific department, and regularly monitor student sentiment for continuous improvement.

Keywords: Rate My Professors, Text Analysis, Sentiment, Semantic, NLP, COVID-19

2. Introduction

The COVID-19 pandemic impacted many young adults, especially students and professors who had to change their learning and teaching styles. This shift led to ancillary effects on how students learn today. While studies indicate that 51% of college students can't concentrate as well as before COVID-19, there is more to understand concerning post-pandemic student learning (National Survey of Student Engagement, 2021).

Looking at student testimonies on professors' performance during the pandemic compared to outside the pandemic can give far greater insight into how student learning has evolved. This can be evaluated by looking at the website ratemyprofessor.com (RMP). Rate My Professor is a professor review site where any student can anonymously leave a review for a professor. By analyzing professor reviews from students during and outside the pandemic, there can be more honest and quantitative insights into the pandemic's effects on student learning.

3. Objective

This analysis aims to take away insights into what specific changes in learning styles occurred for students because of COVID-19, but also to understand what students now desire from their professors because of the pandemic. With this goal in mind, it would be appropriate to examine if there were more negative sentiments in professor reviews during COVID-19. Additionally, looking at those who earned higher grades and scored the quality of the professor highly could have a greater positive sentiment in their review. Lastly, there could be a possible confounding variable in the class subject since many majors may not be as affected as others with the shift to remote learning. Given that some classes already use computers like engineering or computer science it is important to examine if there is a smaller change in sentiment in these departments during COVID-19 than in majors like business/economics or especially in-person heavy classes like dramatic arts. By evaluating the aforementioned hypotheses, coupled with word choice and topic findings, this analysis can capture a holistic view of the effects of COVID-19 on student learning and insights for professors to adapt.

4. Literature Review

This analysis looked into two areas for previous research, the role of RMP among US universities and the impact of COVID-19 on student learning and engagement. Online review websites and Rate My Professor especially, play a crucial part in decision-making by students. This is shown by a study where the lack of access to official evaluations makes RMP a practical alternative for students (Brown et al., 2010). They compare RMP reviews with official student evaluation teachings and present that students believe RMP ratings are honest and representative of instructors' abilities, verifying the legitimacy of the RMP platform

. Another study looked into the biases of RMP reviews and found that prior online ratings, both positive and negative, do affect subsequent online ratings indicating a systemic bias in the dataset that is used in the study (Ackerman & Chung, 2017). To circumvent this bias, the analysis selected two professors from three ranking categories of low, middle, and high of star_rating.

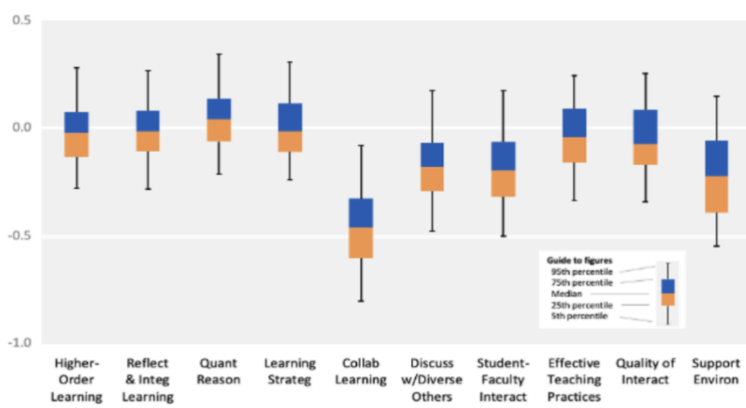
Furthermore, literature from the national survey of student engagement from 2021 showcased the pandemic impact on students' learning. The survey identified an overall decrease in student engagement, especially collaborative learning, decreasing by 50% (National Survey of Student Engagement, 2021). Further supporting this, a Malaysian study emphasized undergraduate students' level of attention and motivation is essential in online learning. It also highlights the importance of self-regulation to increase students' concentration in online lectures (Yusof & Fee, 2022). With these studies there is clear evidence of COVID-19's effect on student learning to use as a basis for our hypothesis and analysis.

4.1 Study Gap

This review pointed towards a gap in research around COVID-19 effects on online professor reviews. Another study explored similar dimensions by analyzing RMP reviews for the Computer Science department professors with special emphasis on how COVID-19 affected students' perception of their classes and professors (Parkin, 2023). The promising insights of such studies further shaped this study to look into COVID-19 impact on student learning and across departments within the same university.

Figure 1

First-year student engagement difference with COVID-19



5. Methodology

5.1 Data Sourcing

The study employed building and using a PyQuery-based web crawler that scraped data from the RMP website, focusing on USC departments with large class sizes, such as Engineering, Computer Science, Business, and Economics. The Theater department was also selected to analyze the impact of a pandemic on departments that typically relied heavily on offline instruction. For each department, data was sourced for six professors, including ten reviews per professor from the COVID-19 period (March 1st, 2020 - August 1st, 2022) and ten reviews from a non-COVID-19 period for comparison. For each professor data collected was the overall star rating, percentage of students who would take the class again, and difficulty. For each review, the contents of the review, quality, and difficulty ratings associated, the date of the review posted and other attributes such as if the class was online, textbook, would take again, taken for credit, attendance, and grade were collected.

5.2 Data Pre-Processing and Analysis

To prepare the collected data for analysis, several pre-processing and cleaning steps were undertaken. The main text cleaning was required on the reviews column. Regular expressions and natural language processing libraries were utilized to remove special characters, convert text to lowercase, and perform tokenization. Additionally, stopwords relevant to this project such as "professor," "class," "lecture," and "student" were removed for some analysis to focus on more meaningful content. Techniques such as lemmatization and joining tokens were applied to normalize the text data. The data was further cleaned by converting date information to a consistent format and data type. The processed data was then utilized as input for different analyses, which aimed to uncover patterns, sentiments, and trends within the student reviews, using methods such as word cloud visualization, bigram, trigram, semantic and sentiment analysis, LDA topic modeling, and Topic Coherence Evaluation.

6. Results

Subsequent sections elaborate on the methodologies employed and the insights gained, with an emphasis on the nuanced sentiment variations across the selected academic departments. The analytic investigation aims to extract actionable recommendations that could enhance the educational experience during disruptive periods.

6.1 Overall Sentiment Analysis

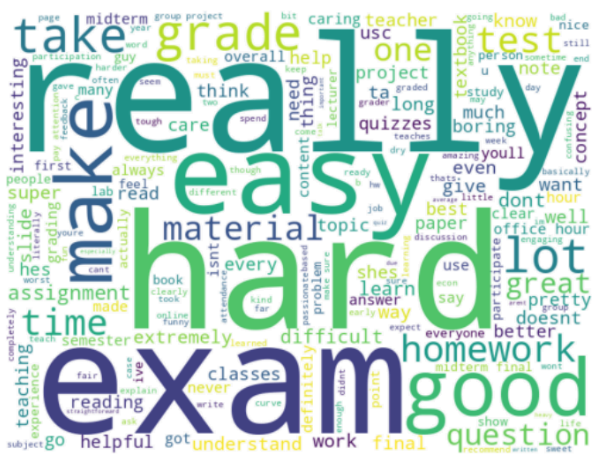
The sentiment analysis commenced with the generation of word clouds to synthesize the prevailing sentiments within the professor reviews. Distinctive approaches for unigrams, bigrams, and trigrams were employed to capture the different levels of expression.

6.1.1 Unigram Analysis

Some common stop words such as "professor" and "class" to emphasize more descriptive terms were removed, revealing appreciation and course rigor as key themes in student feedback.

Figure 2

Word Cloud for Unigram



6.1.2 Bigram and Trigram Word Clouds

Nest, for bigrams and trigrams stopwords, were retained to preserve the integrity of common phrases, illuminating the supportive nature of professors and the learning outcomes as seen by students. These word clouds provide a nuanced view of the student's educational experiences and the diverse teaching styles of professors.

Figure 3

Word Cloud for Bi-Gram

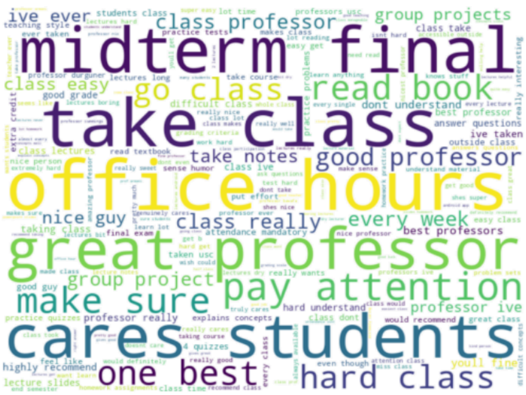


Figure 4

Word Cloud for Tri-Gram

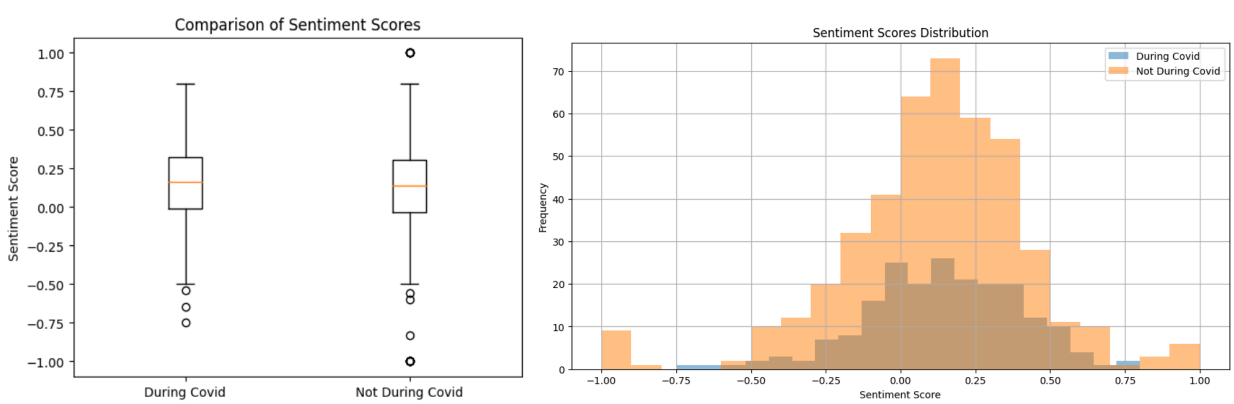
6.2 Further Sentiment Analysis by T-test/Boxplot

6.2.1 Overall Sentiment Analysis

The sentiment analysis conducted via T-tests and boxplots investigated whether the COVID-19 pandemic constituted a vital factor in review sentiment. The boxplot comparison of sentiment scores during and not during periods revealed that the central tendency of sentiments remained relatively stable, primarily neutral over time. This was supported by a T-test yielding a p-value of 0.768, which suggests no statistically significant difference in sentiment scores between the two.

Figure 5

Box Plot and Histogram for Overall Sentiment Scores

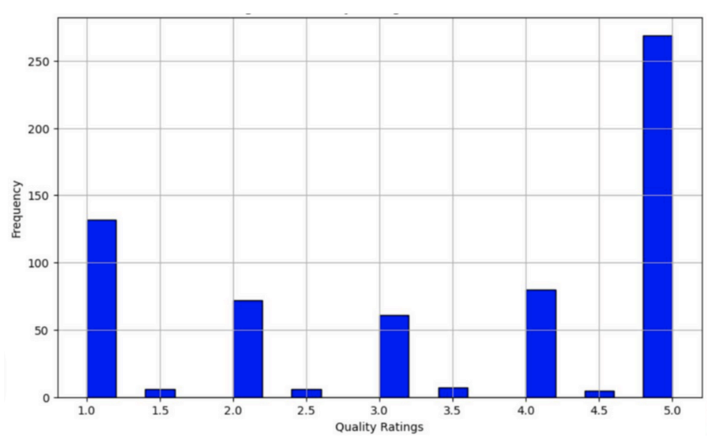


6.2.2 Overall Rating Distribution

On the other hand, the overall ratings distribution portrayed a different story. Histograms of quality ratings indicated a high polarization, with students inclined to give extreme scores—either very high or very low. This polarization in the rating distribution signifies that despite the neutrality in sentiment analysis, the intensity of student opinions is strongly reflected in the ratings they assign.

Figure 6

Bar Chart for Rating Frequency Across All Comments



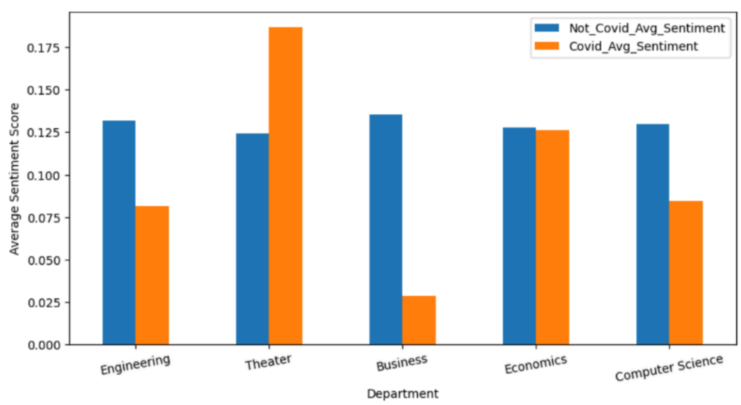
The juxtaposition of these findings—neutral sentiment scores but polarized quality ratings—emphasizes the complexity of review evaluations. It indicates that both the qualitative sentiments expressed in textual reviews and the quantitative ratings must be considered to form a comprehensive understanding of student feedback.

6.2.3 Sentiment Analysis by Different Schools

During COVID-19, Engineering, Business, and Computer Science schools at USC experienced a decline in sentiment scores, particularly in Business, highlighting the challenges of remote learning and the importance of in-person interactions. In contrast, the School of Theater at USC saw an improvement in sentiment scores, likely due to the department's efforts to enhance creativity and connectivity among students. The sentiment scores in the School of Economics at USC remained stable, suggesting that the curriculum and teaching methods were well-adapted to online learning, or that economics students were more pragmatic, focusing solely on professors' performances and not the remote environment. Overall, COVID-19's impact on student sentiments varied across schools, possibly influenced by the unique characteristics of the student body and the effectiveness of the faculty in each discipline.

Figure 7

Comparison of Average Sentiment Scores by Schools

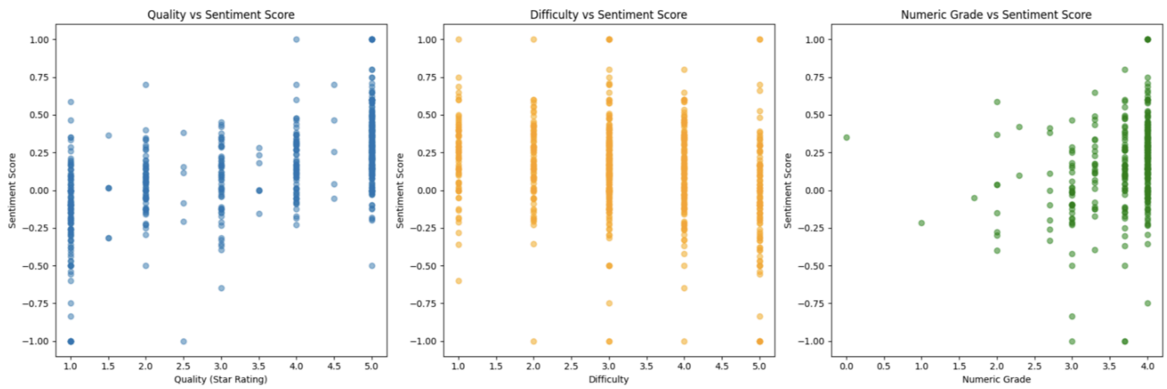


6.3 Sentiment Correlations

In the context of a generally neutral sentiment trend among student reviews, our analysis explored the nuances of how positive sentiments correlate with academic performance and course difficulty.

Figure 8

Correlation between Sentiment Score and (Quality, Difficulty, and Letter Grade)



- Quality and Sentiment Correlation: The data presented a moderate positive correlation (corr=0.593) between the quality of teaching and sentiment scores. This correlation suggests that even as the pandemic prompted a shift in instructional strategies, the essence of high-quality teaching remained pivotal to student satisfaction.
- Difficulty and Sentiment Correlation: A slight negative correlation (corr=-0.248) between the perceived difficulty of courses and sentiment scores indicated that more challenging classes may negatively influence student sentiment, albeit to a lesser extent.
- Grades and Sentiment Correlation: Positive sentiment also showed a correlation (corr=0.269) with the grades students attained, reinforcing the notion that academic success is often accompanied by positive feedback in reviews.

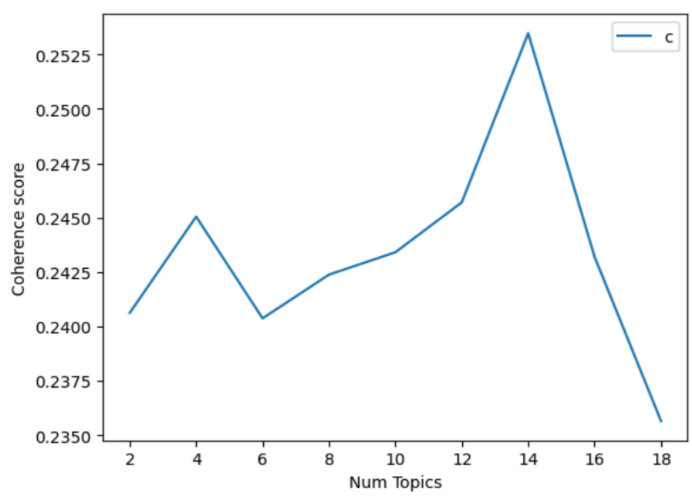
Despite these correlations, it is essential to note that the overall sentiment across reviews remained predominantly neutral. This neutrality points towards a balanced student body voice, where extreme sentiments are less common in textual feedback, even if they are pronounced in the quantitative ratings. Understanding this landscape of student sentiment, with its underlying neutral tone, is crucial for interpreting the more subtle shifts and patterns that emerge in relation to quality, difficulty, and grading.

6.4 Coherence Analysis

From the coherence plot, the peaks are at 4 to 5 topics and 14 to 15 topics, which helped in selecting the number of topics for building the LDA model. Considering comments from the Rate My Professor website, 14 to 15 topics appear to be too many for focused analysis. However, 4 to 5 topics provide a manageable range for in-depth analysis and categorizing the comments into specific groups. Consequently, five topics were selected to construct the LDA model.

Figure 9

Coherence Plot for Number of Topics Selection



6.5 Semantic Analysis

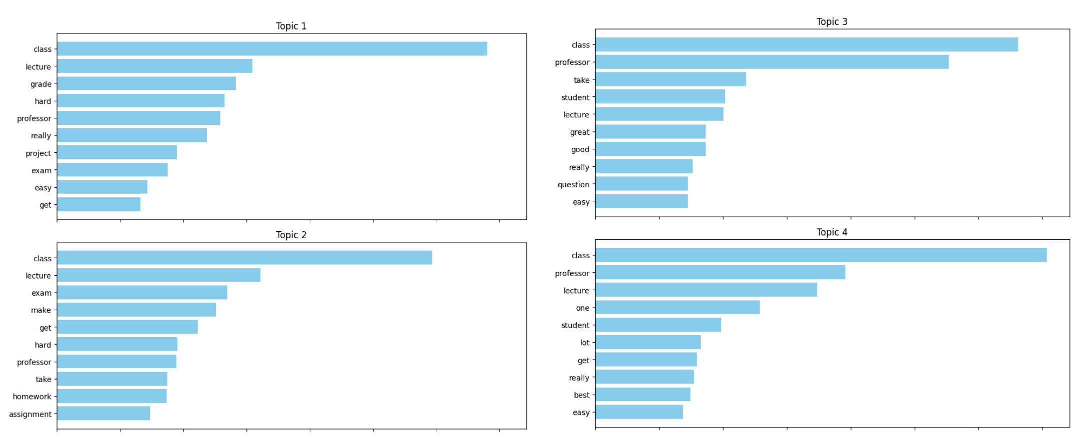
These analytical threads, woven together, offer a tapestry of insights into student sentiments during an unprecedented epoch in education. The interlacing of quantitative sentiment analysis with qualitative textual examination forms a comprehensive portrayal of the academic experience as refracted through the lens of student reviews during the COVID-19 pandemic.

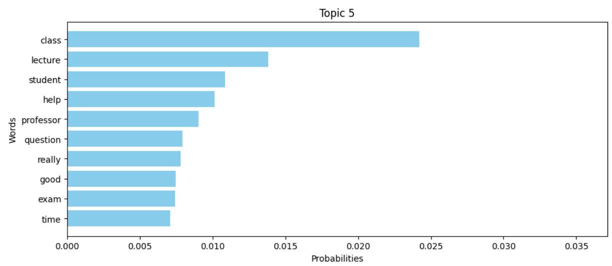
In the LDA model, the first topic includes words like "grade," "hard," and "project," indicating that students discuss the challenging aspects of their coursework. The second topic contains "exam" and "assignment," reflecting dialogues on the stress and preparations associated with assessments and completing assignments. The third topic features "great," "good," and "really," pointing to evaluative comments about the quality of teaching and classroom experiences. The fourth topic includes "best," "really," and "lot," encompassing students' perspectives on their overall educational journey, sharing personal experiences and significant learning aspects. Lastly, the fifth topic with the words "help," "question," and "time," represents discussions about study strategies, time management, and seeking academic assistance, likely reflecting the support structure students depend on.

Collectively, these five topics distinctively categorize the content of comments and provide a comprehensive view of students' perspectives on professors and their focal concerns, guiding professors to adapt and enhance their teaching methods based on student feedback.

Figure 10

Top 10 Most Frequent Words in 5 Topics





7. Recommendations

In effort·s to combat the demands of students as mentioned above, the following procedures are recommended for professors to improve sentiment, and thus feedback:

1. Balanced Coursework: Departments should constantly make efforts to balance course rigor with interactive elements, including group discussions, hands-on exercises, and 1 projects. This can help address students' desire for engagement and practical learning experiences.
2. Department-Specific Strategies: Based on the varying impacts of COVID-19 across departments, customized strategies should be developed. For instance, Engineering and Computer Science may benefit from additional support mechanisms while Teater could incorporate more student-driven performances or practical exercises.
3. Regular Sentiment Analysis: Academic institutions should conduct regular sentiment analysis on student reviews to monitor how course experiences evolve over time. This will provide insights into changing student needs and preferences, allowing for timely adjustments to teaching methods and curriculum design.

8. Limitations

The project faced a couple of limitations during the analysis. The reviews may reflect strong opinions, skewing results due to biased perspectives. RMP also lacks questions specifically addressing COVID-19 experiences, making it difficult to directly measure the pandemic's influence on student sentiment. Furthermore, limited demographic data hinders understanding how COVID-19 impacted different student segments, complicating the ability to propose extensive insights.

9. Discussions

The COVID-19 pandemic had a complex impact on the sentiments expressed in RMP reviews, reflecting varied experiences across departments. Reviews in the Theater department saw a positive increase in sentiment, while Engineering, Business, and Computer Science saw declines, likely due to the challenges of transitioning to online teaching. Economics reviews remained relatively stable, showing only a minor increase in sentiment.

Overall, star ratings correlated positively with review sentiment, indicating that students' satisfaction often translates into positive feedback. Conversely, course difficulty showed a slightly negative correlation, reflecting how challenges might impact students' perceptions.

The pandemic has reshaped the educational landscape, and insights from these reviews offer valuable guidance for professors to build a better learning environment. By incorporating sentiment analysis, balanced coursework, and department-specific strategies, they can adapt and thrive in this new era of learning, leading to more positive feedback from students.

References

1. Ackerman, D. S., & Chung, C. (2017). Is RateMyProfessors.com Unbiased? A Look at the Impact of Social Modelling on Student Online Reviews of Marketing Classes. Journal Name, 40(3). https://doi.org/10.1177/027347531773565

2. Brown, M. J., Baillie, M., & Fraser, S. (2010). Rating Ratemyprofessors.com: A Comparison of Online and Official Student Evaluations of Teaching. Communication Teacher, 57(2), 89-92. https://doi.org/10.3200/CTCH.57.2.89-92

3. Pandemic & Student Engagement: Annual Results 2021: Research: Evidence-Based Improvement in Higher Education. (2021). National Survey of Student Engagement. Retrieved April 29, 2024, from nsse.indiana.edu/research/annual-results/2021/story1.html

4. Parkin, D. (2023). Student sentiments: A sentiment analysis study of Rate My Professor reviews at CSULB (thesis), University Honors Program, California State University, Long Beach

5. Yusof, F. H., & Fee, L. Y. (2022). Mediation Effect of Self-Regulation in Relationship between Attention Span and Motivation in Online Learning during COVID-19 Pandemic among Undergraduate Students in The Faculty of Human Ecology, Universiti Putra Malaysia International Journal of Academic Research in Business and Social Sciences, 12(12), 2114–2124. DOI: 10.6007/IJARBSS/v12-i12/16009