

RateMyProfessors Snapshot

Rachel Harris, Shivam Patel, and Madeline Phillips

Abstract—This report analyzes a snapshot of student reviews from the review website RateMyProfessors.com for professors at The University of Tennessee. There are over 3700 professors from The University that have reviews. We will gather a section of reviews for each professor and analyze them using graphs and word clouds. We also separate reviews roughly based on each department within the University to determine trends in departmental student sentiment.

I. OBJECTIVE

Our objective is to extract students' opinions of The University of Tennessee's professors by extracting and analyzing their the reviews left for them on the on the website RateMyProfessors. We will break this information down by College to determine if any department has a higher satisfaction or dissatisfaction rate than the others.

II. MOTIVATION

RateMyProfessors can be a valuable tool when selecting classes. It's useful for determining which professors fit your personal learning style. Although we recognize that bias is present in the data, it can still provide a representative sample of students' sentiment regarding their college. The bias primarily results from the nature of a review site. Reviews can be polarized because students with neutral experiences rarely write reviews. Only the students who have strongly positive or negative experiences go on the site to review their professors, resulting in stronger opinions than those that are held by the general student body. Even so, the general results can provide an indication of student satisfaction in their departments.

For instance, student feedback is critical to the University's success in its journey to rank in the top 25 among the United States' public universities. The University garners this feedback in a variety of ways, but a primary obstacle is student unresponsiveness to surveys. Our RateMyProfessors sentiment analysis report can provide a partial solution to the University's lack of student feedback by making previously obscure information easily accessible. RateMyProfessors has thousands of students' reviews on their classes and instructors, making it a prime candidate for the University to receive candid and extensive student feedback. If our school wants to succeed in its journey to the top 25, acting on assessments from previously-ignored sources could be key.

On another front, our sentiment analysis report can be a practical and relatable tool for prospective students investigating our University. Hundreds of potential students in the incoming freshman class are undecided on their field of study. Being able to see fellow students' opinions on different departments and Colleges within the University

will help prospective students determine where they will be most successful and satisfied. This can significantly ease the difficult decision regarding field of study for numerous students.

Professors can also potentially benefit from seeing how well their department is doing. Professors will be more motivated to make a positive and impactful experience on students in their classes. Professors can try to adjust their course or teaching style based on any constructive criticism they may receive. This would greatly result in a more satisfactory classroom experience for both students and professors.

III. DATA WE OBTAINED

We will obtained review data from the site using a website scraper. We extracted reviews for over 3,700 professors. The site itself also allows us to filter by department and College, so we used this option to help us segment our data by department, thus analyzing and comparing the average numerical reviews for the professors in each college at The University of Tennessee. We can also compare reviews of professors within a department.

One issue with data processing that we encountered was RateMyProfessors' categorization of Colleges and departments; the site does not distinguish between them. It will categorize some reviews by College and others by department, but the data looks the same to the scraper. For example, there's a category for "Engineering", which encompasses the entire Tickle College of Engineering. However, there's no category for the College of Arts & Sciences, only departments within it such as "Social Sciences" and "Geology". Due to this categorization design error on the part of RateMyProfessors, we cannot specifically distinguish between Colleges and departments in this report, and we will treat them as the same entity, referring to them as "departments" due to the higher presence of departmental labels than College labels on the website.

Using the scraper, we gathered the numerical rating, review tags, professor, course, department, and the review body itself. We will gather a maximum of 20 reviews per professor to create an up-to-date "snapshot" of student sentiment.

IV. MODELS AND ALGORITHMS

We did not use specific models or algorithms to parse or interpret our data in this project. Rather, we used graphs to help the data speak clearly.

V. RESULTS

Our results displayed a variety of student opinions regarding their professors. All figures referenced can be found in the Appendix at full size. They were not included in this section because the small details were unreadable at this size. All scores are reported out of a 5-point scale.

Figure 1 presents the average difficulty rating per department. Some of the results were surprising: The College of Engineering only has a difficulty rating of 3.0, meaning that the students who wrote RateMyProfessors reviews consider their engineering classes to be an average difficulty. We expected this difficulty rating to be higher. The College of Law, another difficult field, also received a lower average. In fact, there were only five departments with a difficulty rating of 5.0/5.0. These were Anthropology, Architecture, Kinesiology, Social Sciences, and Theology. The only department out of this list that is widely considered to be challenging is Architecture, so the other results were unexpected. This discrepancy leads us to question whether there's a significant difference in the way students in different departments perceive difficulty. If we continued research on this, we would distribute a survey among the students of these departments to measure how they assess class difficulty.

Next, we look at the easiest and hardest professors in each department, as shown by Figure 2. All the departments except two had at least one professor considered to have an extremely easy class (a difficulty rating of 1.0). The first exception to this was the College of Law. The easiest class for students in the College of Law had a 4.0 rating, meaning that law students consider their classes to be more difficult than most other fields. Interestingly, the second department with no extremely easy classes was Hospitality. Students in Theology rank all their classes as "extremely difficult"; however, this does not hold much weight as there are only two Theology reviews for The University of Tennessee.

Now, we will move away from a discussion of difficulty and look at students' thoughts on the quality of their professors (see Figure 3). Most departments had professors who are dearly loved (with a score of 5.0) and professors who are less-enjoyed (with a score of 1.0). Hospitality students gave even their least favorite professor a higher score than 2.0, implying that Hospitality professors may be somewhat hospitable. Students of several departments gave their professors top quality ratings exclusively, including Exercise & Sport Studies, Educational Psychology & Counseling, and Criminal Justice. Conversely, two departments had exclusively lowest quality ratings, Design and Culinary Arts.

When a student writes her review, she is given the option to apply several tags to it from a selection provided by RateMyProfessors. These tags include informative names like "Gives Good Feedback" and "So Many Papers". Looking at the collection of the tags that the students chose gives us deeper insight into student opinion. Figures 4-6 concern these tags.

Figure 4 presents a word cloud created using the tags

for all the professors of the entire University. We can see that professors mainly defined by giving good feedback, that they're tough graders, and that class participation matters significantly. It is encouraging that as a whole, students find professors' feedback valuable. Let's move our focus to the College of Engineering specifically (see Figure 5). Students in our College have a slightly different perception of class than students overall. Similar to the University as a whole, Engineering students consider their professors to be tough graders. However, class participation is a much smaller part of an Engineering student's life. Instead, these tags put a strong emphasis on the "Amazing lectures" that their professors give. Furthermore, it seems that grades are based more on tests than other departments, as "Test heavy" is a dominant tag.

Last but not least, Figure 6 shows the ideal characteristics of one of the students' favorite Engineering professors, who received a score of 5.0. Qualities that are most important to students include the professor being "Inspirational", "Caring", and giving "Amazing lectures". As we saw in Figure 5, "Amazing lectures" are a defining trait of the College of Engineering. Thus, professors can set themselves apart to students by investing into being inspirational and caring.

VI. PRIMARY CHALLENGES

We encountered a variety of technical and logistical challenges during this project. As mentioned in Section III, we were not able to distinguish between Colleges and departments due to the categorization system of the website. We could not determine a way to counteract this issue.

Another aspect we struggled with was scraping all the reviews for each professor. Some professors had hundreds of reviews, while others had fewer than twenty. RateMyProfessors only stores 20 reviews per page. If the user wants to see (or send a request for) more than that, the user must click the button "Load More Ratings" at the bottom of the page to load 20 more reviews, and so on. We attempted to enable our scraper to click the button using Selenium Browser Automation and were successful. This led to the scraper sometimes being able to reach past the first page of professor reviews. However, our efforts were frustrated by the advertisements on the website. They would regularly pop-up, after which our scraper's clicking ability was rendered non-functional. We attempted to accommodate the advertisements in our scraper, but were unsuccessful. Ultimately, we only collected the first page of reviews for each professor, leading to our maximum of 20 reviews.

Finally, we had great difficulty parsing out the JSON data dump to isolate the specific elements we needed, such as professor names and ratings. Since we were parsing it using Python, much of the data was categorized into Python dictionaries. We could navigate these easily. However, some pieces of the data, such as the "Reviews" section for each professor, did not render correctly into dictionaries, instead remaining as lists. Converting the lists to dictionaries did not work, as the items were inconsistently formatted. We spent hours attempting to access each professor's review tags,

which were located at center of this nested monstrosity. In the end, we discovered Dictor for Python, a JSON handler. It allowed us to navigate the data as we had initially hoped to.

VII. FUTURE WORK

Our approach to future work would be threefold. Primarily, we would like to solve our limited review problem. Somehow, we need to evade the website's advertisements to collect all reviews for all professors. This improvement would give us the opportunity to perform a comprehensive review of student opinion on the departments. It is currently impossible to call our results comprehensive due to the missing data.

Secondarily, if we solve this issue, we would like to add a "time" component to our results. Each review is timestamped on the website. We would like to modify our scraper to extract the time as a part of the review, allowing us to delve into student opinions over time. Some questions we would like to investigate include "Have quality ratings increased or decreased over time per department?" and "Have "Have difficulty ratings increased or decreased over time per department?".

Thirdly, as mentioned in the Results section, we would investigate how students of different departments perceive class difficulty by distributing a survey with questions designed to measure how they assess the difficulty of their courses.

VIII. RESPONSIBILITIES

Overall, our team split the load fairly evenly. Shivam Patel was the primary writer of the scraper, although the design was developed by all three of us. Rachel Harris and Madeline Phillips wrote the program to parse the JSON data dump into a graph-able format, with Rachel specifically generating the graphs. Madeline analyzed the data and wrote the final report.

APPENDIX

Average Overall Difficulty Rating per Department

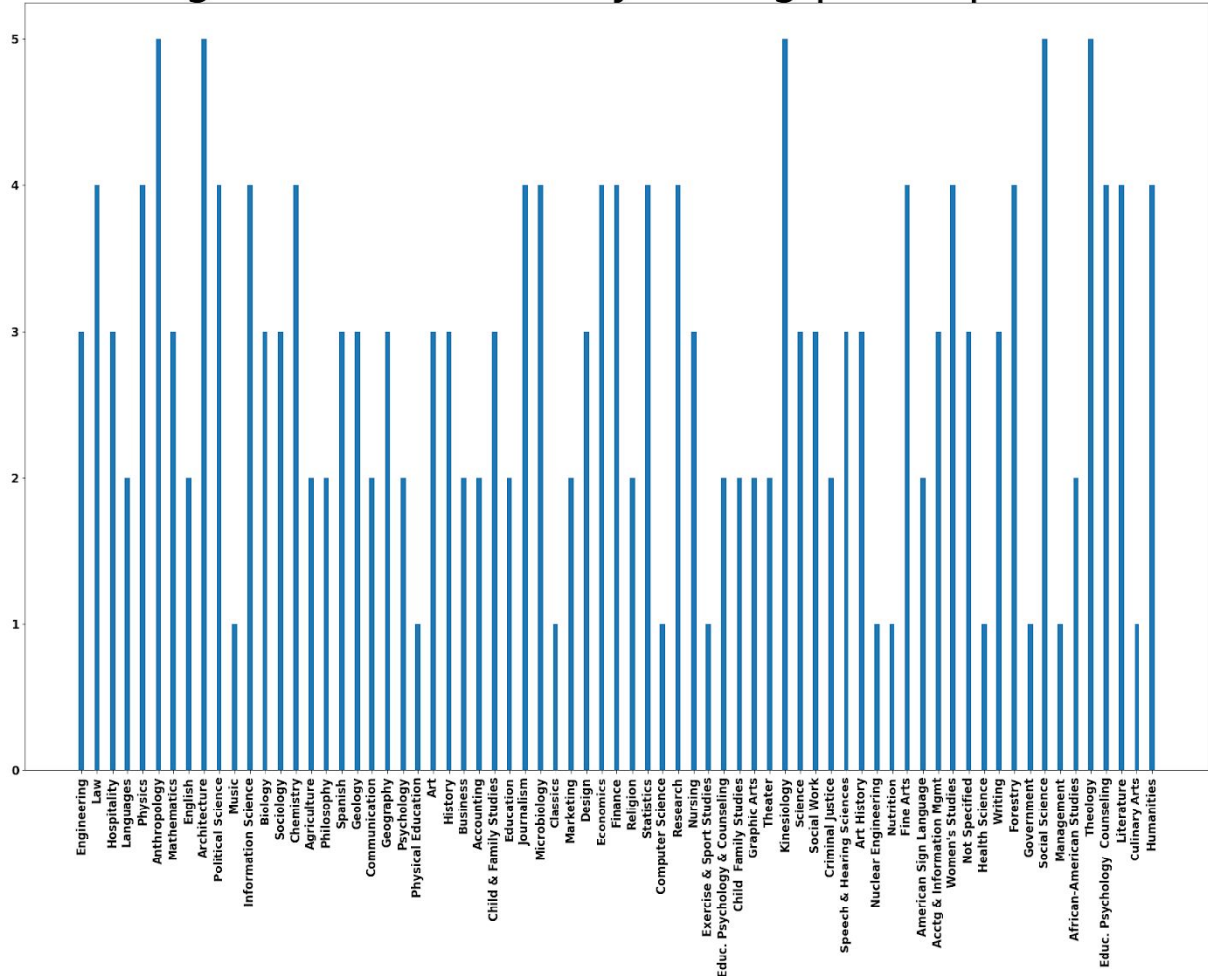


Figure 1: Average Difficulty Rating per Department

Best and Worst Professors by Difficulty Rating per Department

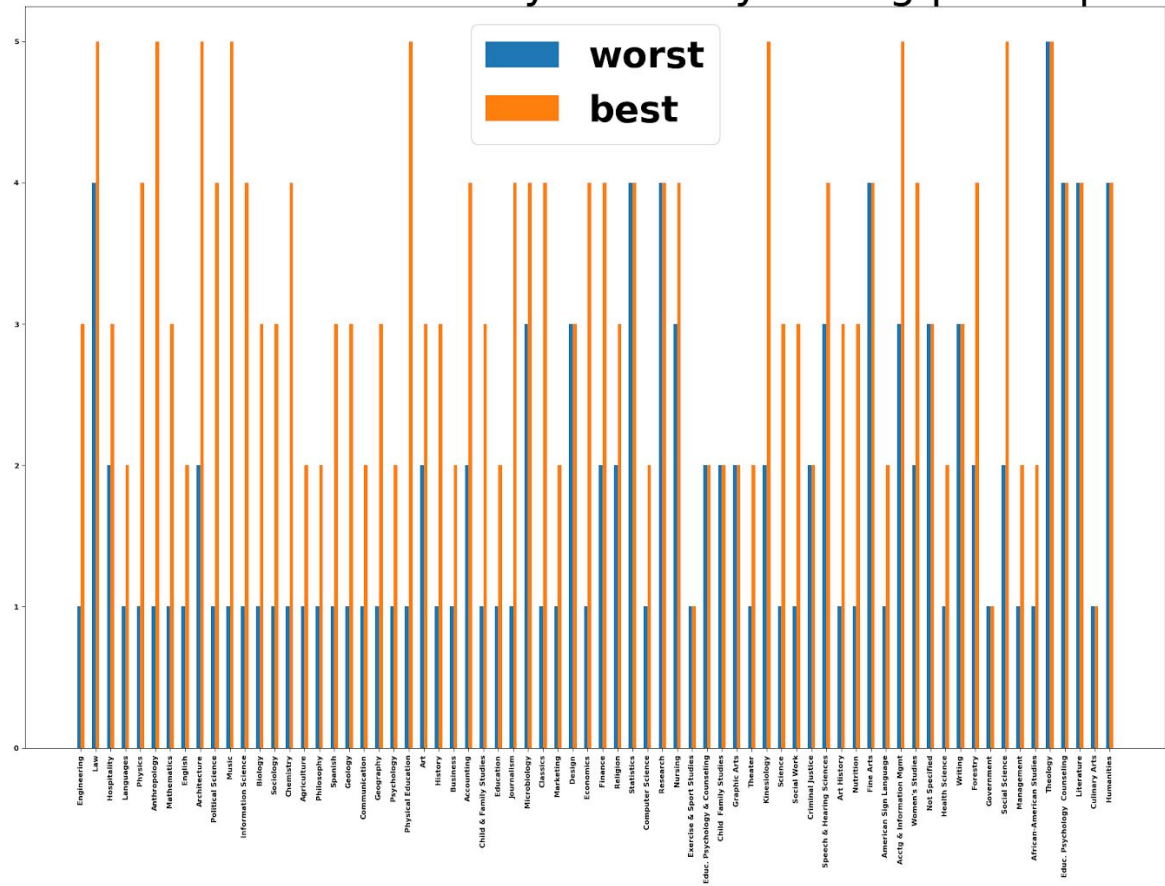


Figure 2: Easiest and Hardest Professor Rating per Department

Best and Worst Professors by Quality Rating per Department

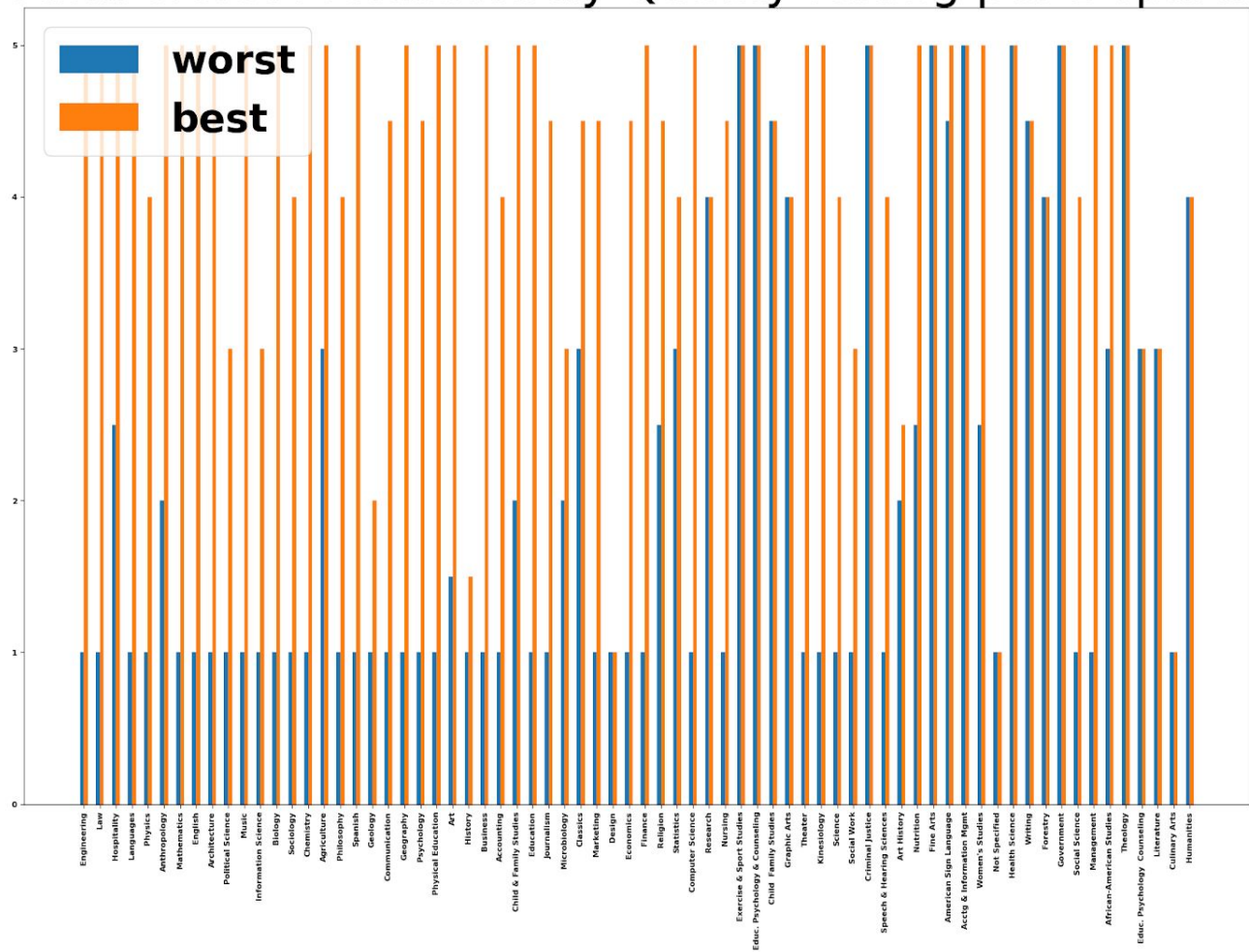


Figure 3: Best and Worst Professor Rating per Department

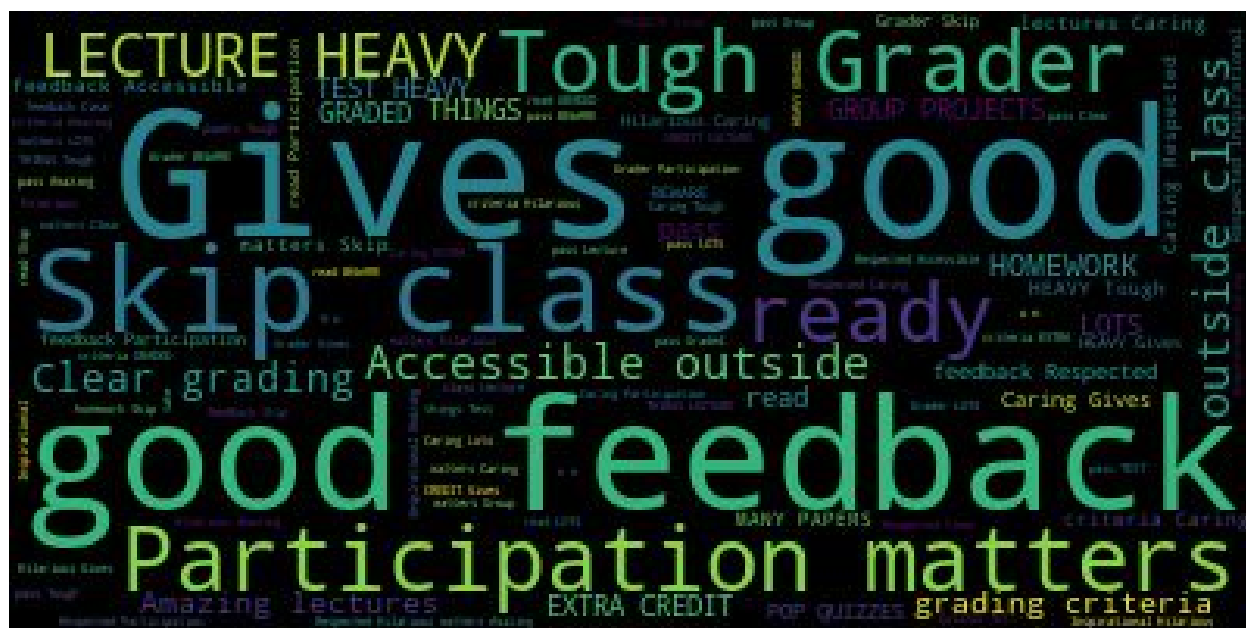


Figure 4: Word Cloud of Review Tags for all reviews

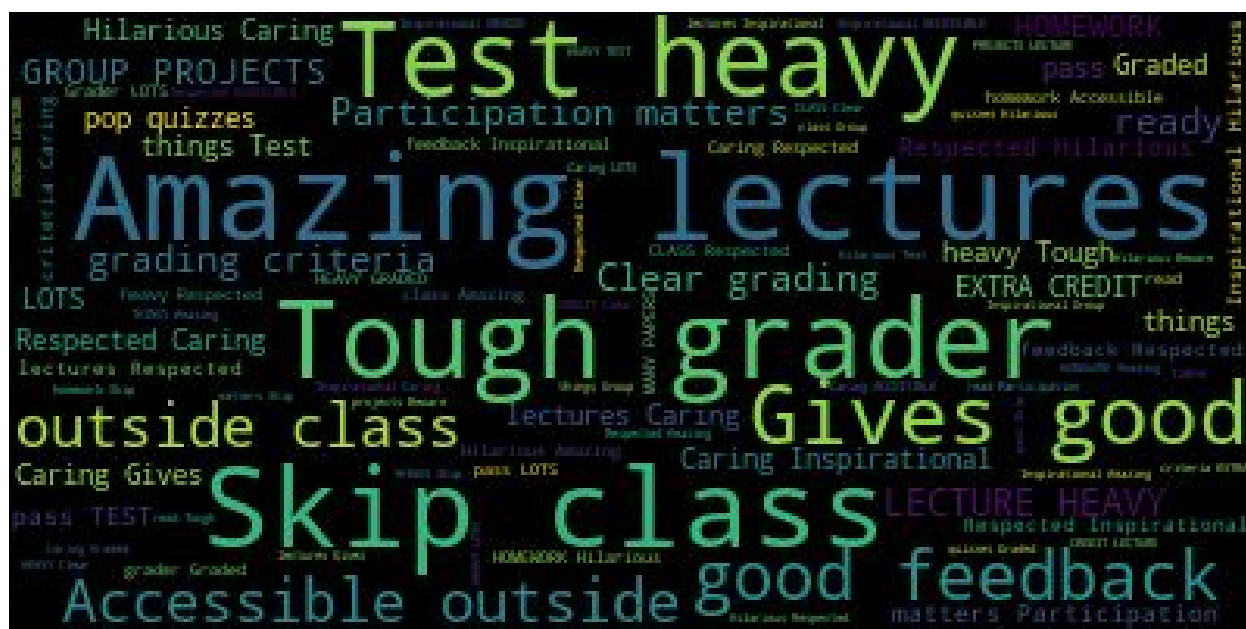


Figure 5: Word Cloud of Review Tags for the College of Engineering



Figure 6: Best Professor Review Tags Word Cloud