Mauricio Murillo, Johnny Hoang
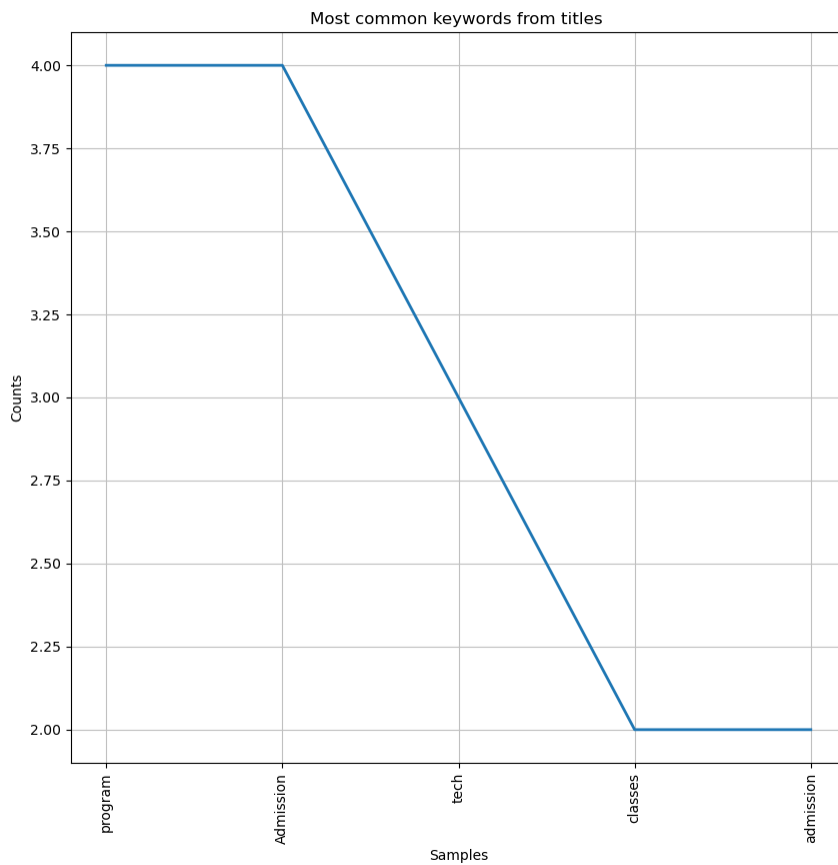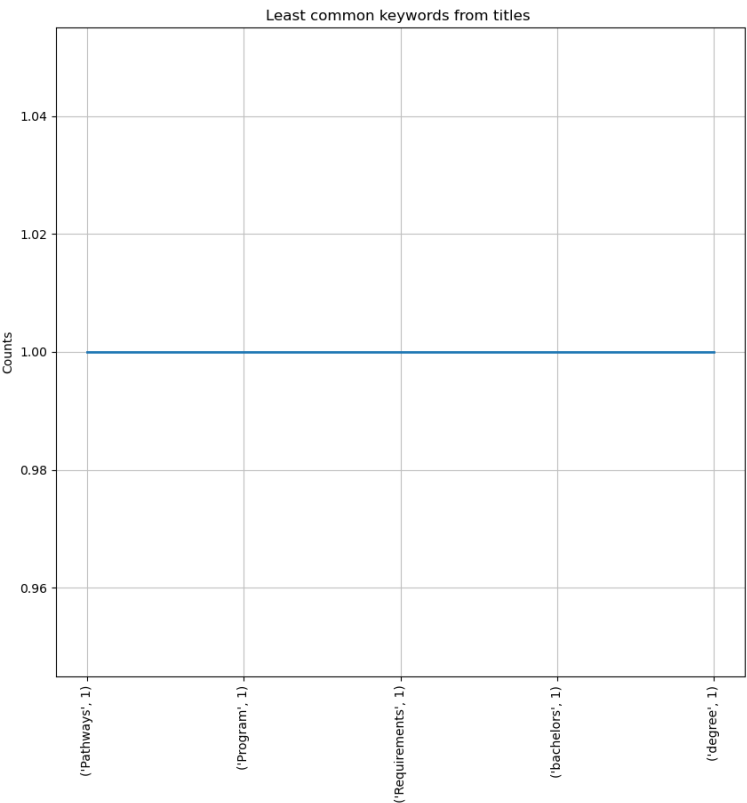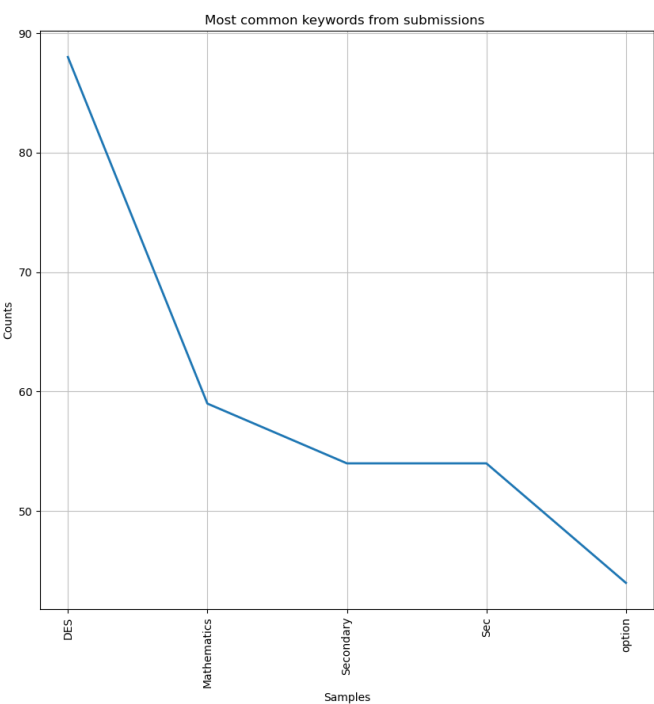
Dirk Dubois

Python Server-Side Programming

Analysis report of the Dawson college's subreddit

Accordingly to our research, the statistics that we collected prove that the main topic for computer science posts and comments are about admissions. We filtered out the computer science posts from the other posts in the Dawson subreddit by using keywords scrapped from the official dawson computer science website, which were stored in a database, and some additional manual entries like "computer science", "cs", etc. These posts were used as data to prove our hypothesis that the major demographic of the Dawson Subreddit are high-school students asking questions about Dawson College and the admission process.

We searched for the most common keywords from titles and removed the already used filtering keywords and stop words in order to obtain efficient data that can be analyzed. We found out that the word "admission" and "program" were the most relevant keywords, which shows that redditors are interested in knowing if they are eligible to join the computer science program at Dawson.
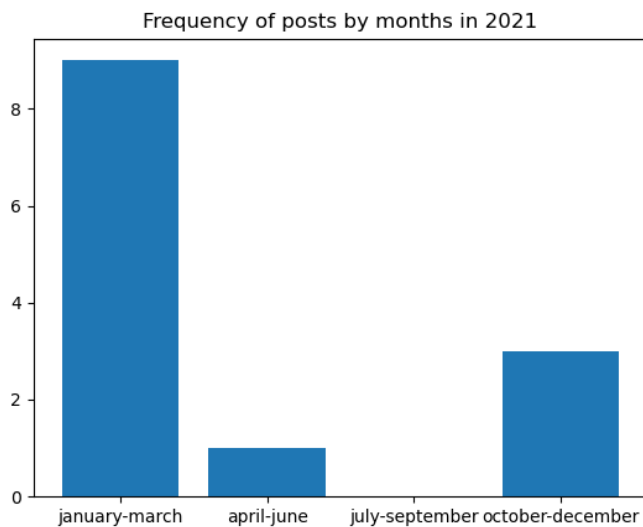
Most common keywords from titles

We then searched for the text of these submissions, and it further proves that our hypothesis is correct. People mention the keyword "DES" (Diploma of Secondary Studies) up to 88 times, "sec, and "Secondary" 54 times which demonstrate that they are looking to go to Dawson after high school. Mathematics appears 59 times meaning that people ask if their SEC IV and SEC V are enough to be accepted into the program.
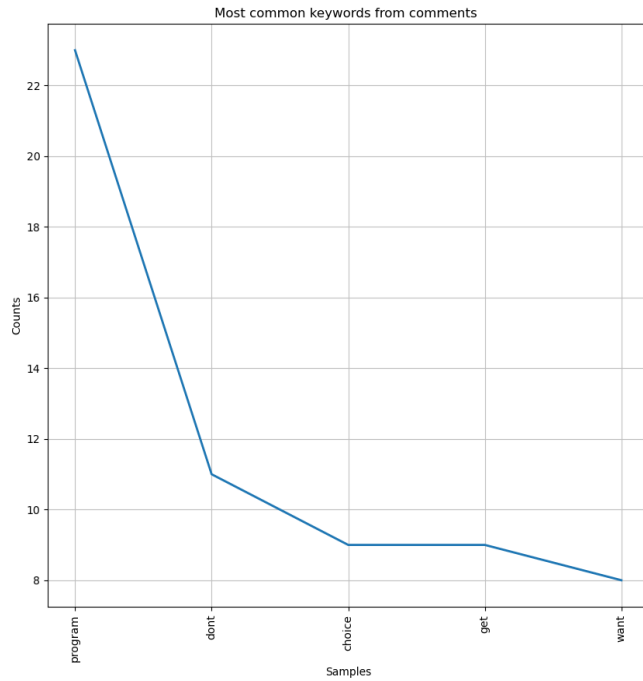
Most common keywords from submissions

While the questions about admissions are more prominent, we observed that the least common keywords were "pathways", "bachelors", "degree" which were only mentioned once. This shows that Redditors are not much interested about the career path that the program will lead to and the degrees that can be obtained after the program ends.

Least common keywords from titles

These findings coincide with our hypothesis because the new or not yet admitted students are not much concerned with the pathways after the degree because they are more concerned with the admission and applying process.

Frequency of posts by months in 2021

We analyzed the number of posts by months in 2021 to see which months get the greatest number of posts. It turns out that the most active time period of the subreddit is from January to March which is the time to apply for the Fall Semester. The only other period that is active with 3 posts is the time between October and December which is the time to apply for the Winter Semester.

Most common keywords from comments

Average number of sentences for posts and comments

According to the keywords that we collected from the comments, other redditors respond if people can get in the program with the words "don't" and "get" They also recommend which choice to do if they do not get into the program with the words like "program", "choice" and "want". The comments are usually shorter than submissions.