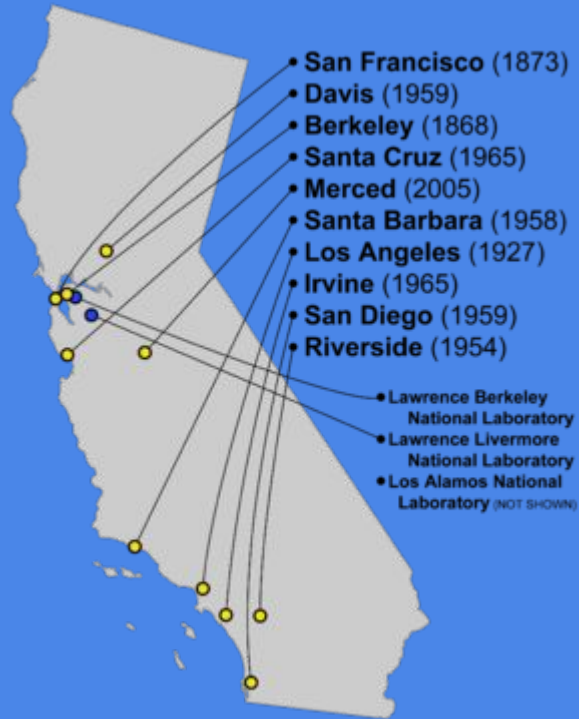




r/GradSchool r/College

Brianna Lytle
DSI - CC9 - LA

UC SYSTEM:



- UC Student Ambassador : goal to improve overall student experience
- 14 UC Schools
- 250,000 + Students

Goal:

- **College/Undergrad**
- **Grad/PHD**

- Understand the student experience at different levels.
 - What are the most common topics talked about?
- Create plan to improve student experience

Data:



- Scraped /r/college and /r/GradSchool **REDDIT** threads.
- Included all threads up to Oct 17, 2019
 - Categories: Hot, New, Controversial, Top, Rising

Methods:



- Cleaned and explored data
- Sentiment Analysis
- Automated 100 + different models to search to find most accurate model

Goal:

- **College/Undergrad**
- **Grad/PHD**

- Understand the student experience at different levels.

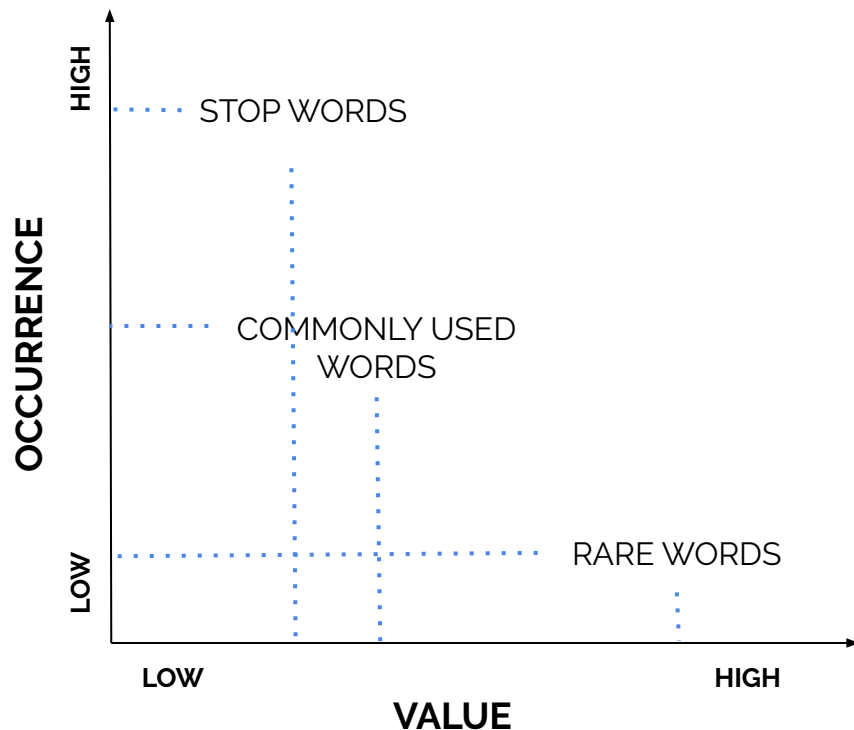
Model Transformer: Tf-IDF

- No custom stop words
- max_features = 1500
 - min_df = 0.85
 - max_df = 0.85

Model Estimator: Logistic Regression

Model Accuracy:
0.889

Tf-IDF

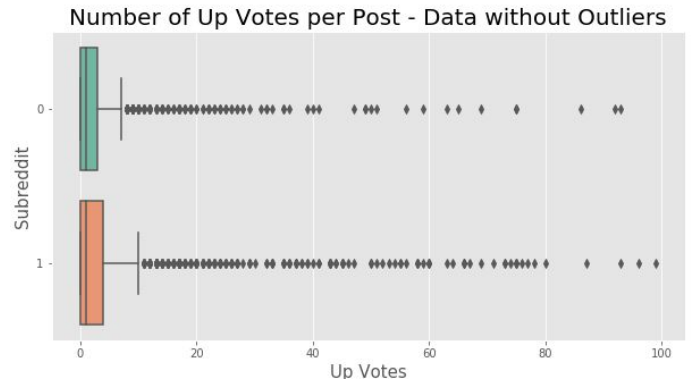
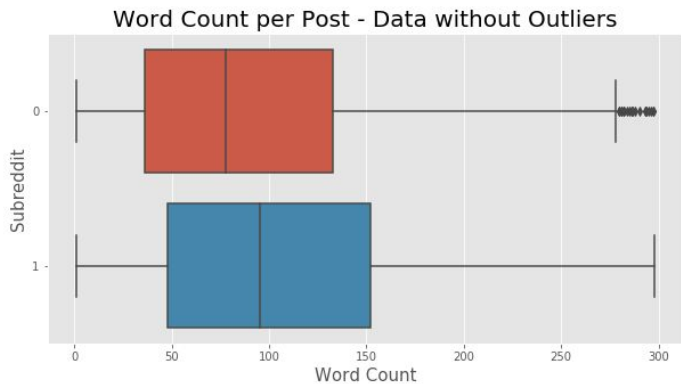
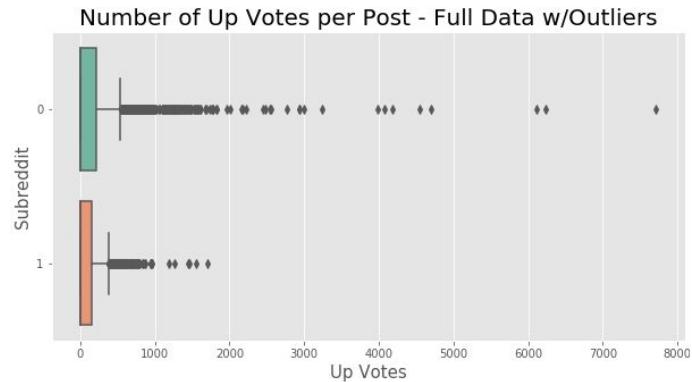
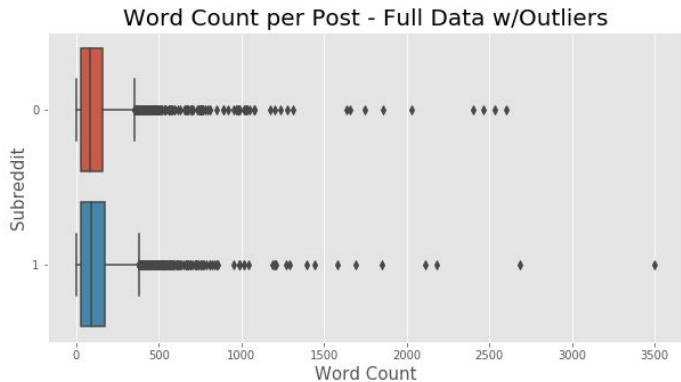


Data Exploration

/r/college = 2885 threads

/r/GradSchool = 3745 threads

GradSchool subreddit users are more active with posting & interacting with other threads.



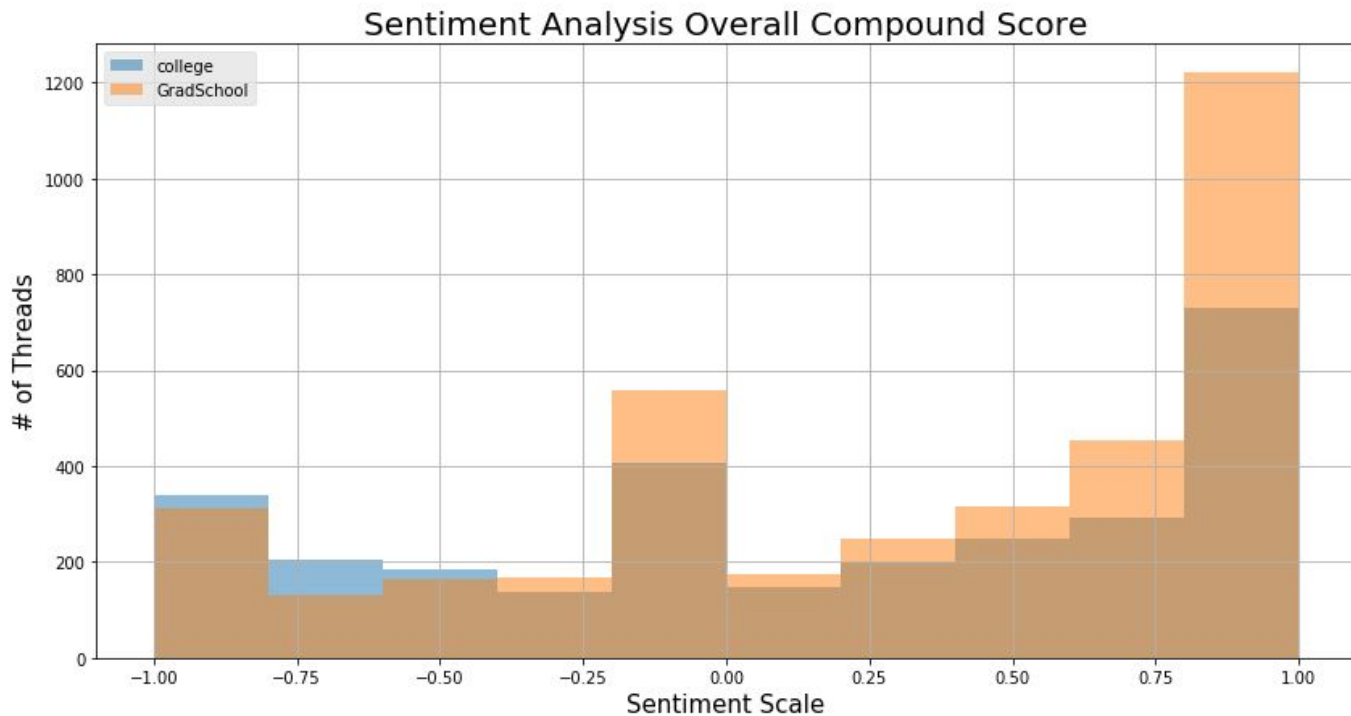
*Table adjusted for outliers

	Word Count	Up Votes
College = 0	25% = 36 50% = 78 75% = 133	25% = 0 50% = 1 75% = 3
Grad School = 1	25% = 48 50% = 96 75% = 152	25% = 0 50% = 1 75% = 4

Sentiment Analysis

College Thread:
More Negative, still
has posts in
positive range

GradSchool Thread: More
Positive/Neutral
Posts



College Thread: SIA Compound

25% = -0.4 | 50% = 0.22 | 75% = 0.8

GradSchool Thread: SIA Compound

25% = 0 | 50% = 0.46 | 75% = 0.88

Common Positive Terms

Compound Sentiment
Score > 0.6

/r/college = 1022
threads

/r/GradSchool = 3745
threads

College Thread: 2 & 3-ngram	
Financial aid	73 threads
Computer science	57 threads
Make sure	43 threads
Make friends	35 threads
Save lot time	8 threads
Computer science major	7 threads
Tax information irs	6 threads
Legal electronic signature	6 threads

Grad School Thread: 2 & 3-ngram	
Mental health	72 threads
Computer science	71 threads
Research experience	57 threads
Greatly appreciated	44 threads
Good happened week	20 threads
Advice greatly appreciated	14 threads
Years work experience	13 threads
Talk good happened	10 threads

Common Negative Terms

Compound Sentiment
Score < -0.6

/r/college = 1674
threads

/r/GradSchool = 442
threads

College Thread: 2 & 3-ngram	
Office hours	33 threads
Really want	24 threads
Computer science	22 threads
Feel bad	16 threads
Professor office hours	5 threads

Grad School Thread: 2 & 3-ngram	
Mental health	51 threads
Years ago	22 threads
Imposter syndrome	16 threads
Hard time	15 threads
High blood pressure	9 threads
Having hard time	8 threads
Mental health crisis	5 threads
Really really bad	5 threads

Goal:

- College/
Undergrad
- Grad/PHD

- Understand the student experience at different levels.
 - What are the most common topics talked about?
- Create plan to improve student experience

Recommendations & Next Steps:

- Grad Students
 - Very stressed, provide access mental health services directly to them/shaped for their experience
 - Looking for support - organize meetups, free events, social hour
- College Students
 - Financial Aid Services.
- Computer Science - more resources, professor help

- Run New models
- Further investigation on text from SIA analysis into full dataframe
- Read into comments

Sources

Reddit threads

- <https://www.reddit.com/r/GradSchool>
- <https://www.reddit.com/r/college>