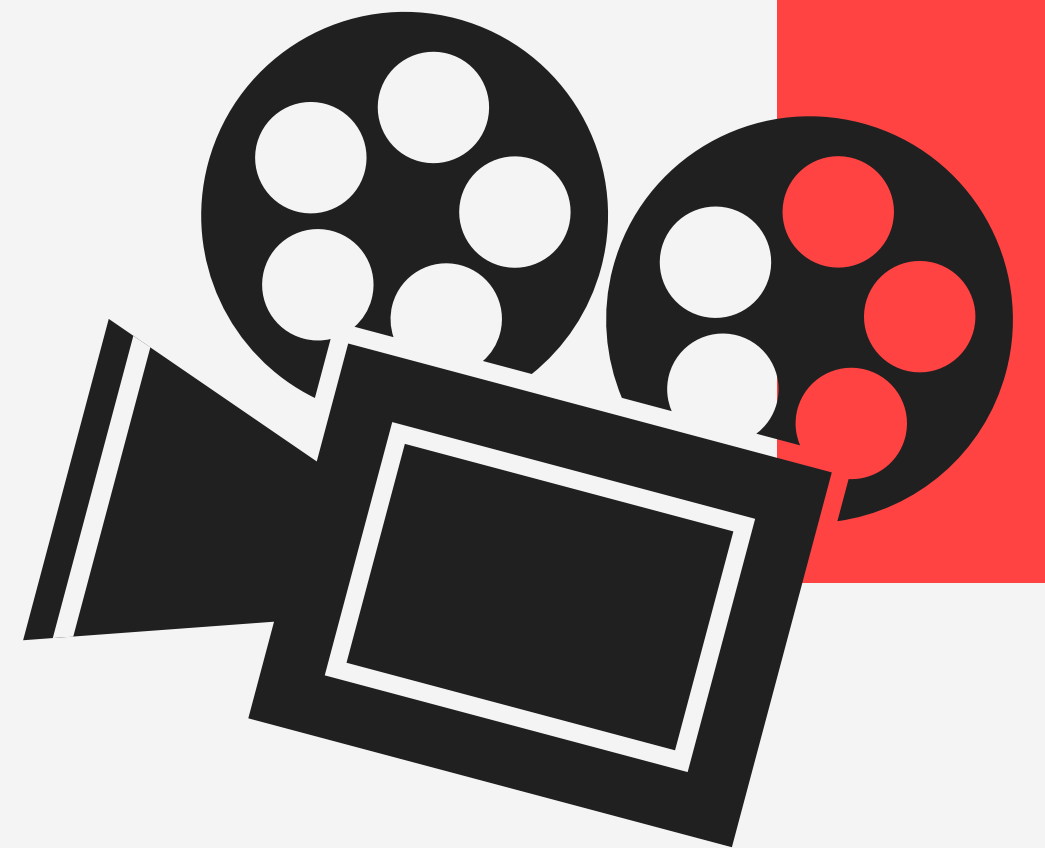# USING SUBREDDITS FOR HORROR FILM RECOMMENDATIONS

Maria Flores

Based on highly visited horror subreddits, a film studio wants to know:

- On average, what horror stories are more popular?
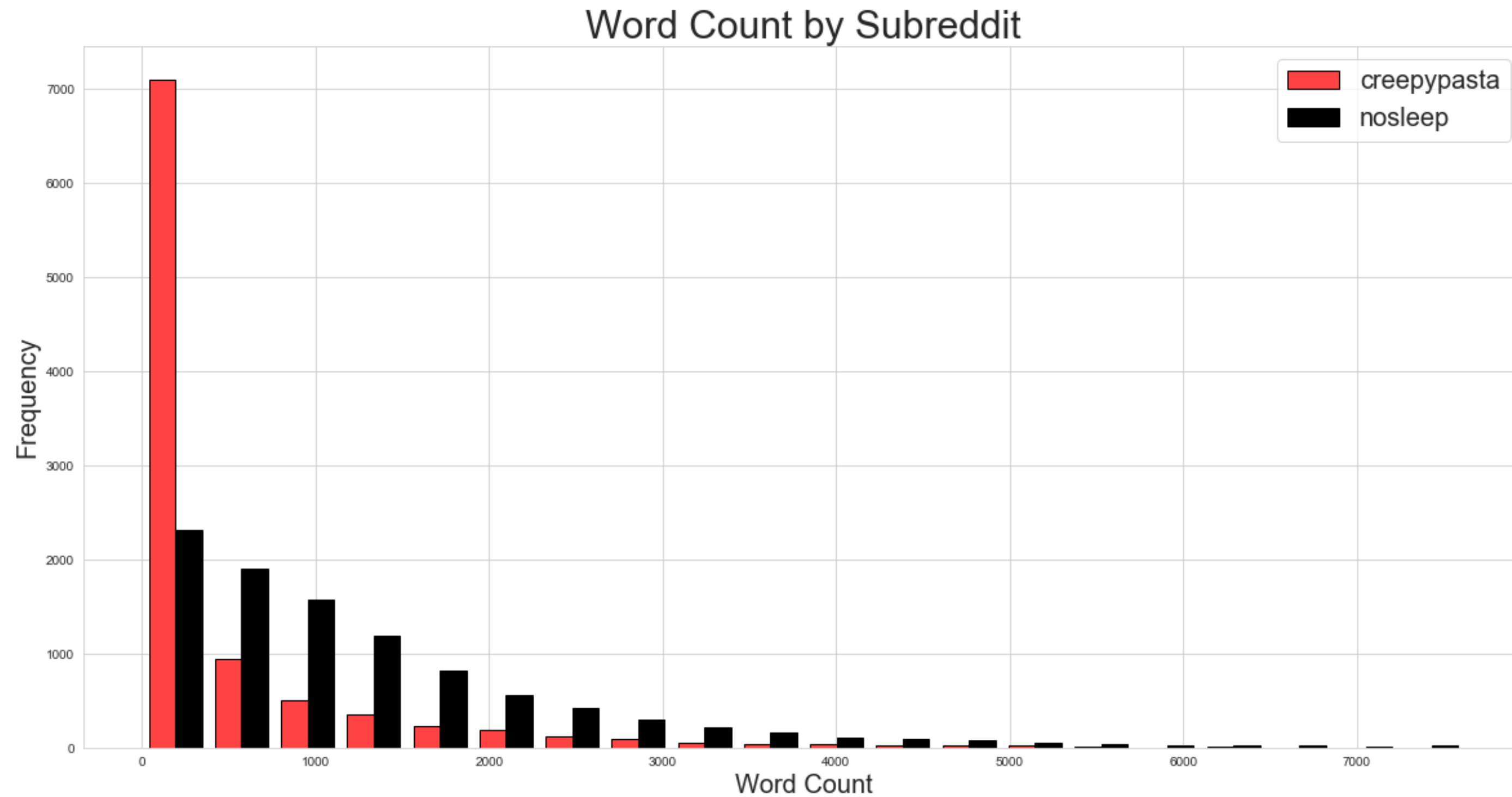- What subreddit did they come from?

What subreddits?

NOSLEEP

- User-generated horror stories with a believable plot written in the first-person
- Usually long stories
- Contains no images
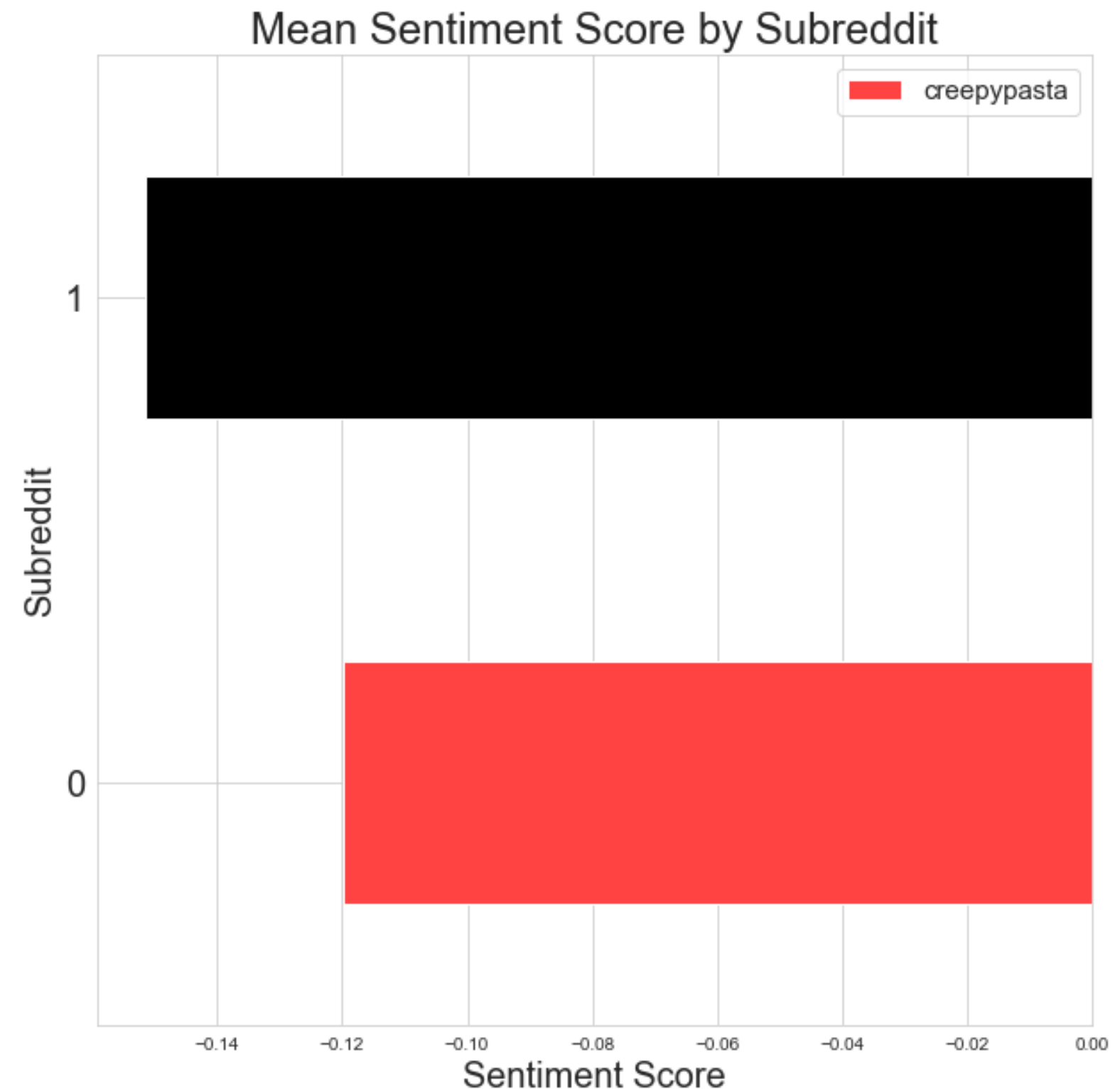- 13.7m members

CREEPYPASTA

- Short user-generated horror stories
- Usually contains an image or video
- Stories aimed to be virally shared across the internet
- 325k members

Which subreddit has a higher word count on average?
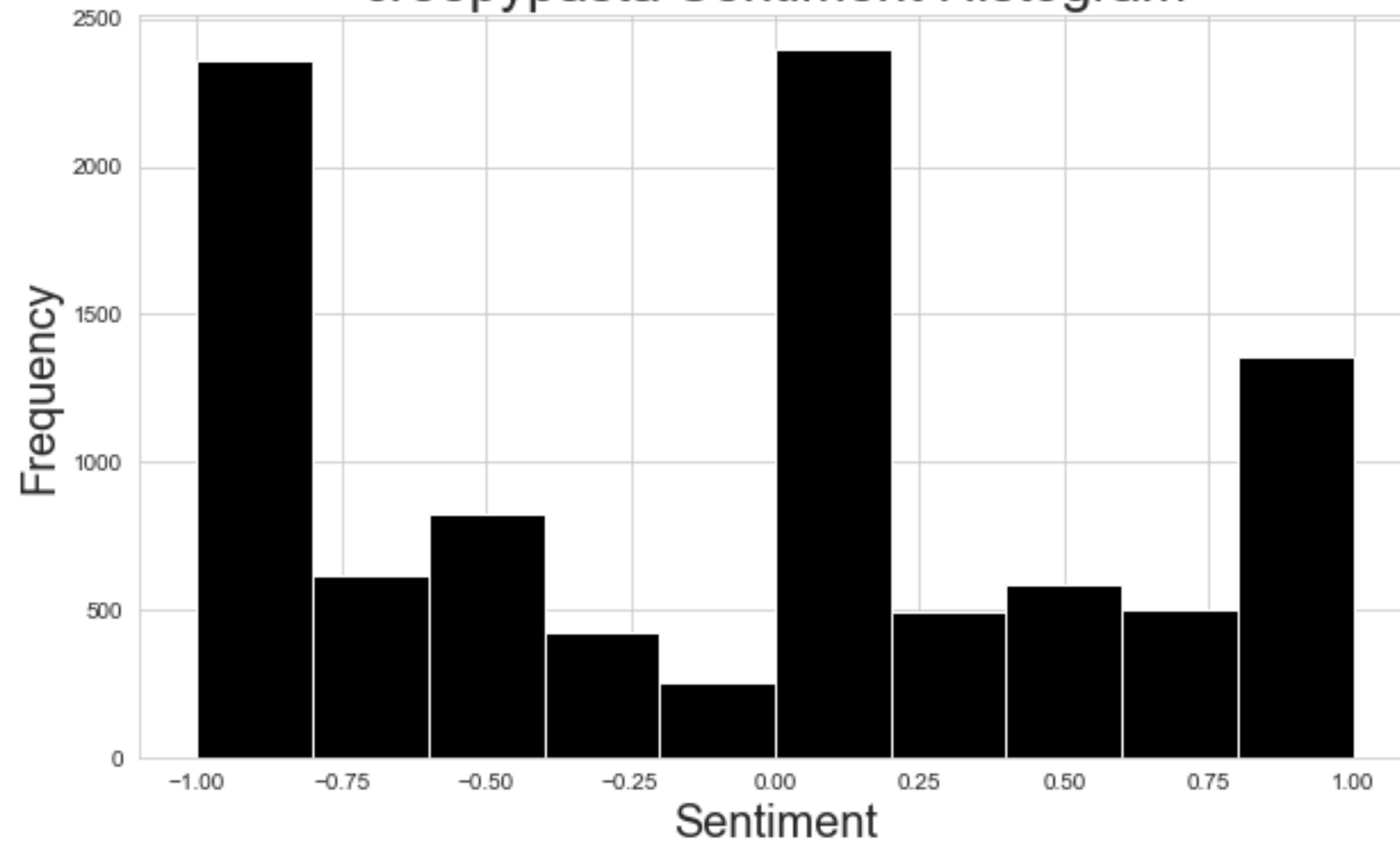

Word Count by Subreddit

Sentiment Analysis

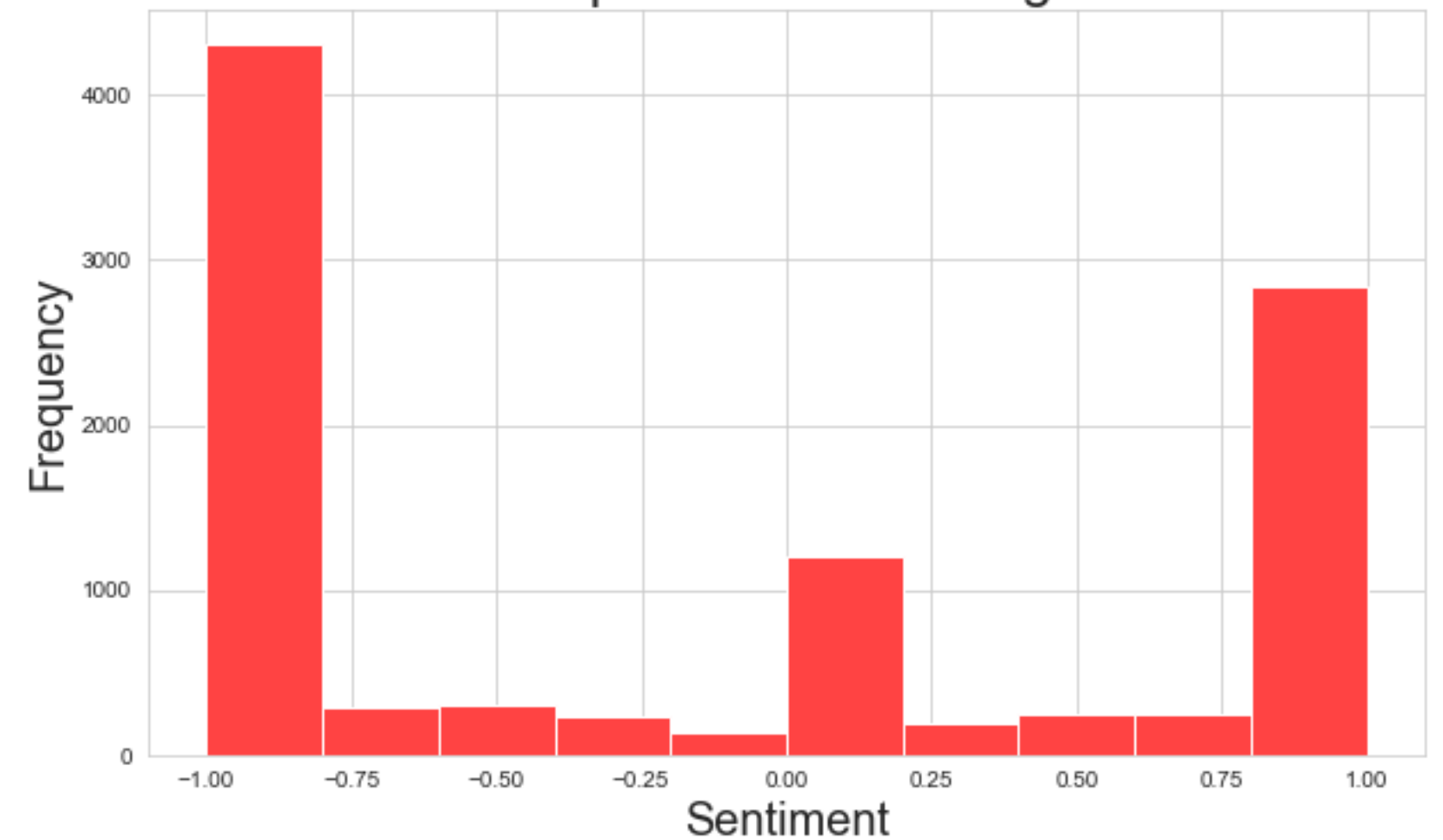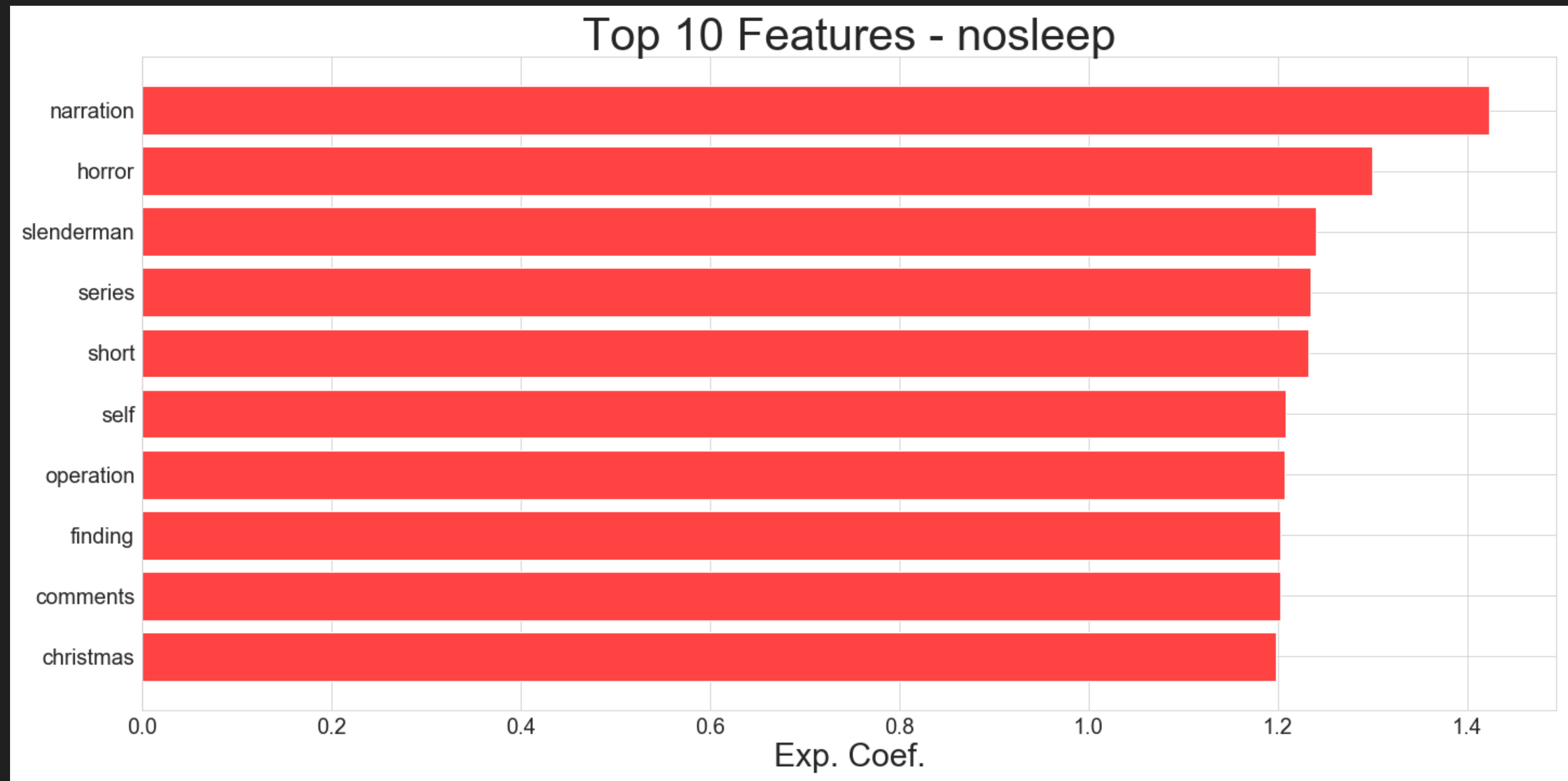- On average, **nosleep's** sentiment score, is slightly more negative than creepypasta's

Sentiment Score Distribution For Each Subreddit

creepypasta Sentiment Histogram

nosleep Sentiment Histogram

Top 10 Features - nosleep

	Exp. Coef.
narration	
horror	
slenderman	
series	
short	
self	
operation	
finding	
comments	
christmas	

As there is a one word increase in our word **'slenderman'**, the reddit post is 1.25 times as likely to be from the **NOSLEEP** subreddit.

Best Performing Model

Count Vectorizer with Extra Tree Classifier

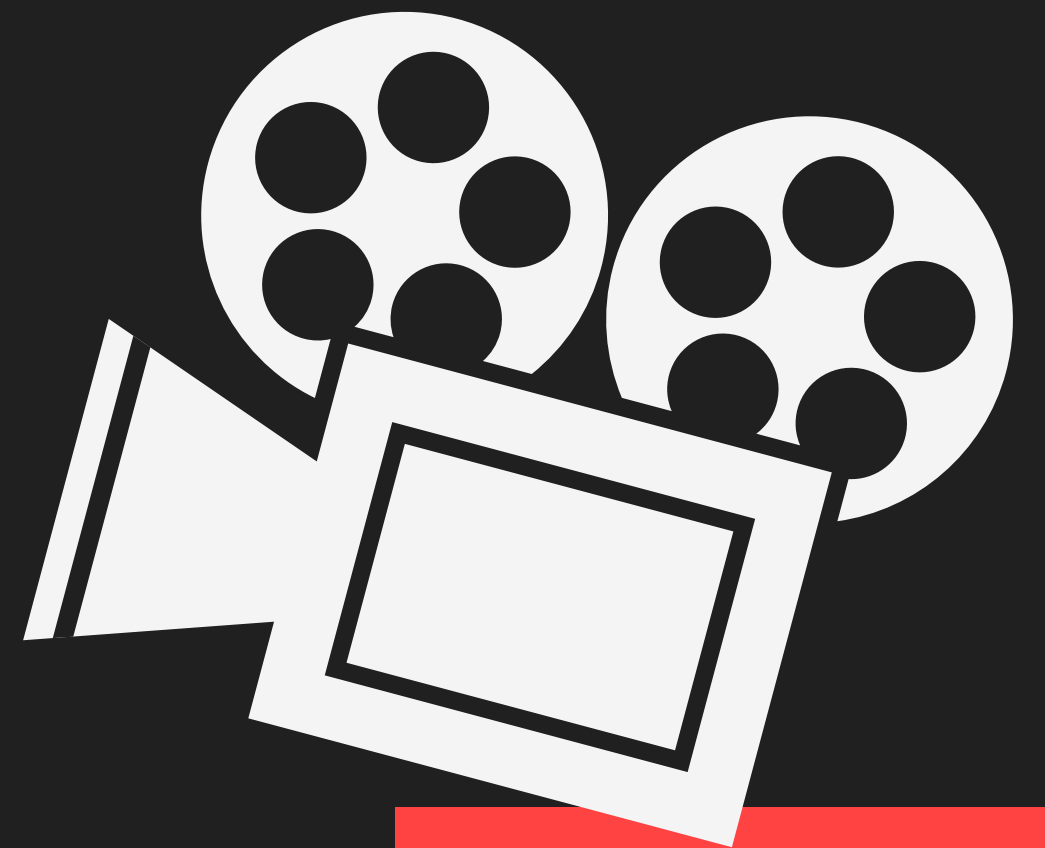- The model is able to predict the correct subreddit 78% of the time.
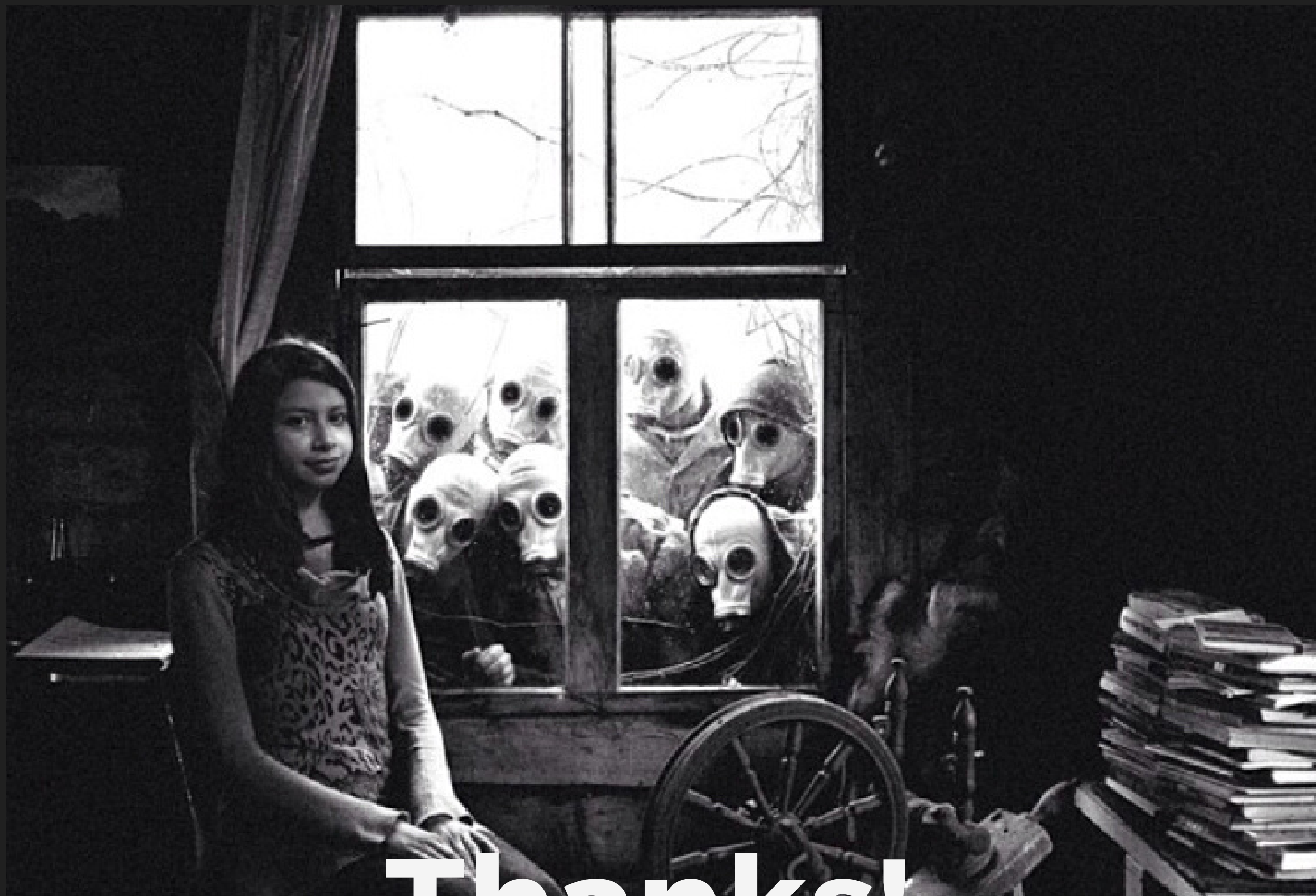
78%

Recommendations

Based on the most important features from the subreddit NOSLEEP, I would recommend making a horror film about Slender Man.

Overlapping Top Words
- 'like', 'just', 'said', 'time', 'know', 'eyes', 'looked', 'got', 'door', 'room', 'did', 'saw', 'night', 'went', 'way'

Thanks!