When is the Witching Hour?
Sentiments of Fear and Social Media

James M. Clawson

23 Oct. 2019

Abstract

Using social media postings about Halloween, this sentiment analysis correlates fearfulness to time of day, testing the hypothesis that posts made at night show more fearful content. Ultimately, it is determined that posts made at night are *not* more fearful than posts made during the day. The hour with the greatest average fear value is 6:00 AM.

Introduction

On the Internet, nobody knows you're agog—but sentiment analysis can get close. As Halloween nears, social media posts polish a sweet tooth and turn toward ghosts and witches. Not everyone gets excited about costumes, though; Halloween's associated nighttime festivities make some people downright afraid.

This project analyzes emotions in social media posts about topics related to Halloween. Using a popular method of sentiment analysis, it correlates values of "fear" to a comment's posting time to determine when the most fearful posts are made: during the day or at night. The hypothesis is that nighttime posts will show more fear than posts made during the day.

Just as the night is dark, it should, I presume, also be full of terrors. By measuring fear values of posts made at different times of the day, analysis should make clear when commenters show their greatest trepidation.

Methods

Respective CSV files were downloaded from social-searcher.com for five different search terms: halloween, ghosts, candy, vampires, and witches.

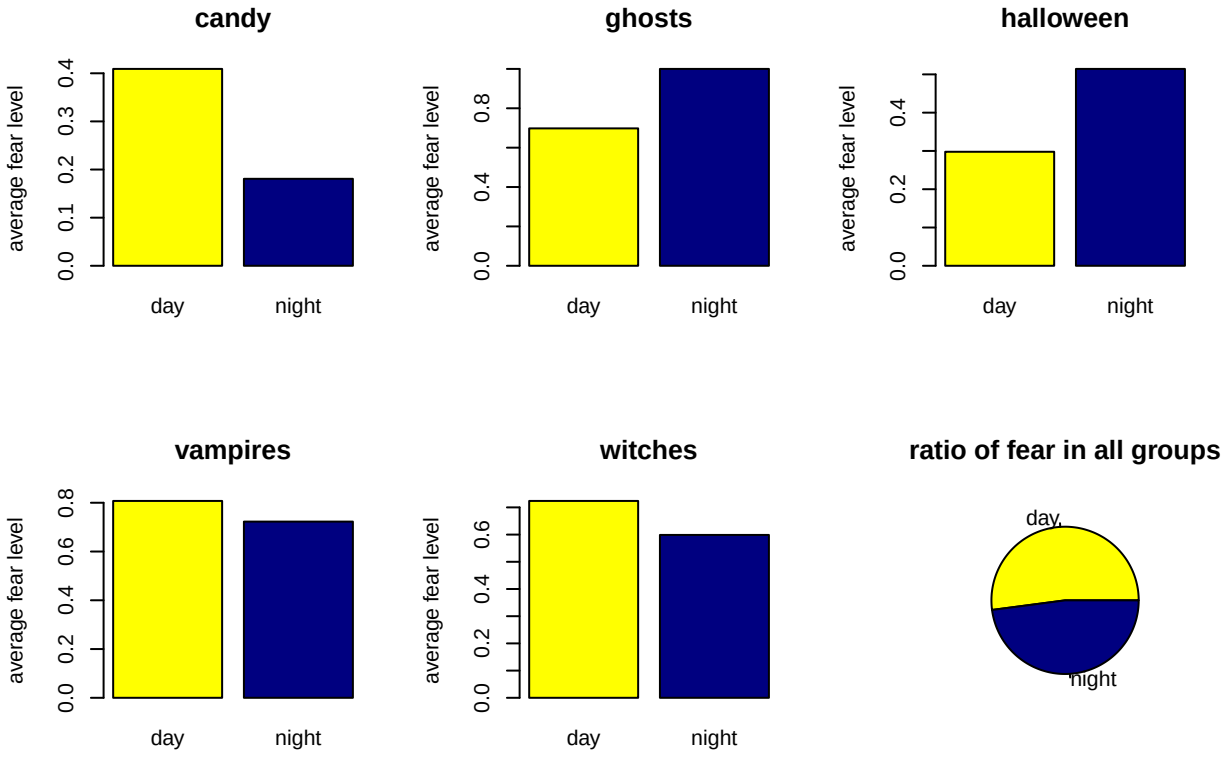
After text was cleaned and regularized, the `syuzhet` package was used to measure comments along eight axes of emotion: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

Finally, groups were considered at two different divisions: first, divided at each hour of the day; and second, divided into groups of daytime and nighttime, with daytime corresponding to posts made between 6 AM and 6 PM. After the `dplyr` package was used to group data by these categories, average values were recorded for each grouping. These divisions made it possible to test the hypothesis, comparing the average of fear values at night versus the average of fear values during the day; they also made it possible to determine which hour is the one in which commenters show the most fear.

Results

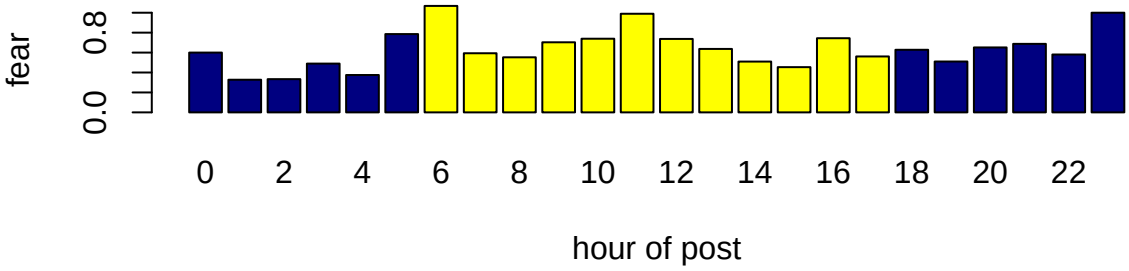
Night and day differences

Analysis found daytime posts to have higher fear values for three of the five search terms. Moreover, the combination of all groups found the most fearful comments to be posted during daytime hours, 6 AM to 6 PM. Visualized below, the bar charts show the night-and-day comparison of fear values for each search term, and the pie chart shows the average ratio of night-fear values to day-fear values for all combined groups.



Hourly differences

Rather than daytime or nighttime, more granular analysis highlights posts made between 6–7 AM to have the highest average levels of fear.



Discussion

Entries with the highest fear values show a flaw in the experiment: short texts are poorly suited for measuring emotional valence. One post with a high fear value is about choosing appropriate shoes for running; another advertises a book. It's unclear that these posts *should* be among those with the greatest fear in the whole set.

Additionally, posts' time zones may be inaccurate. Of 2,685 values, eleven are from Central Standard Time; the rest are Central Daylight Time. Some probably don't reflect the time zone of the person posting

Ignoring these potential flaws, results show the hypothesis mostly disproved. Rather than night time, social media posts about topics related to Halloween show the most fear at daybreak.

References

- social-searcher.com, with search terms candy, ghosts, halloween, vampires, and witches.
- `R` with the `dplyr` and `syuzhet` packages