

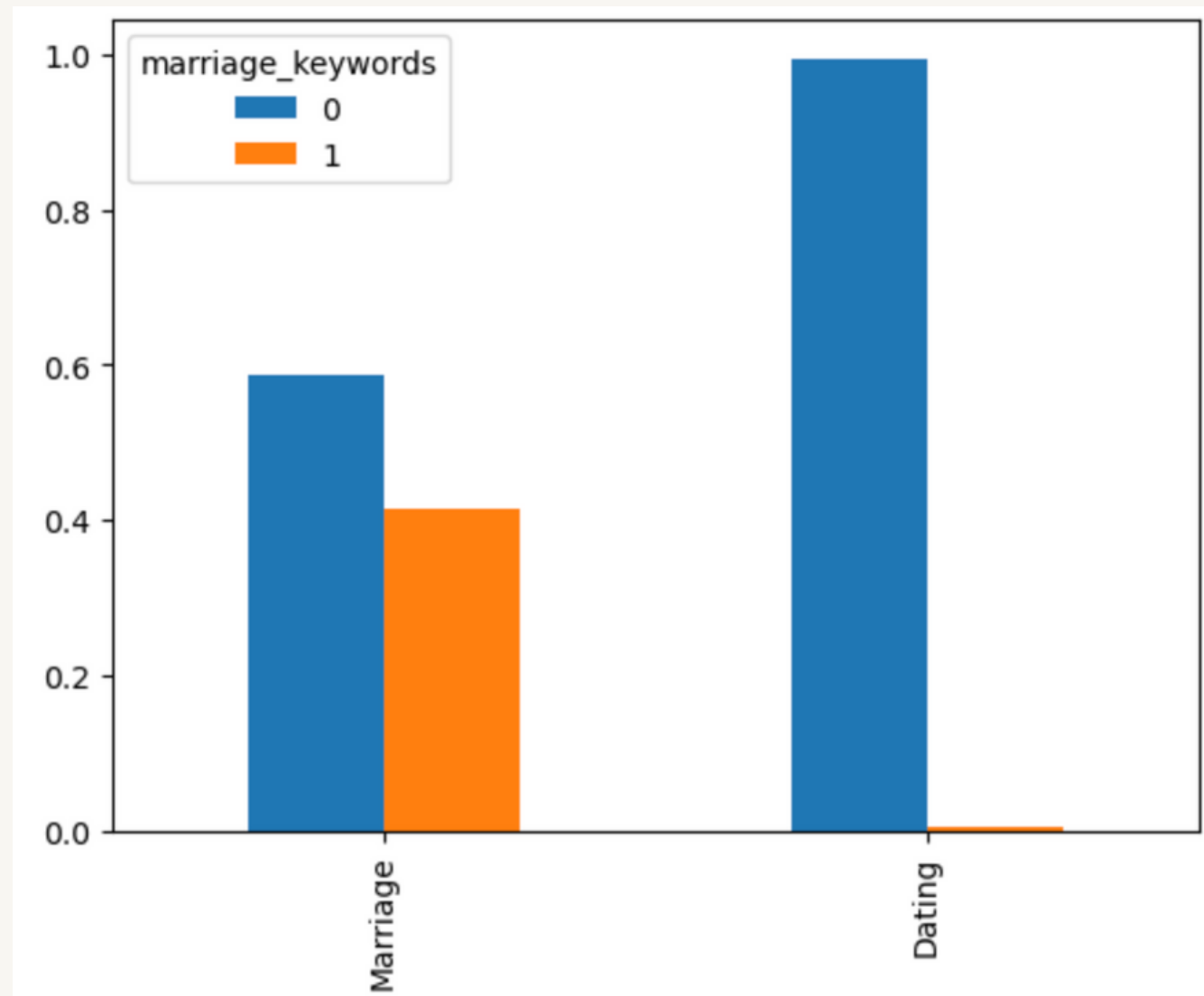
Dating x Marriage

Users Overlap: Approximately 6%

Dating and Marriage subreddits: Top 1% in size of Reddit

Title, Text and Subreddit:
40% Marriage 60% Dating

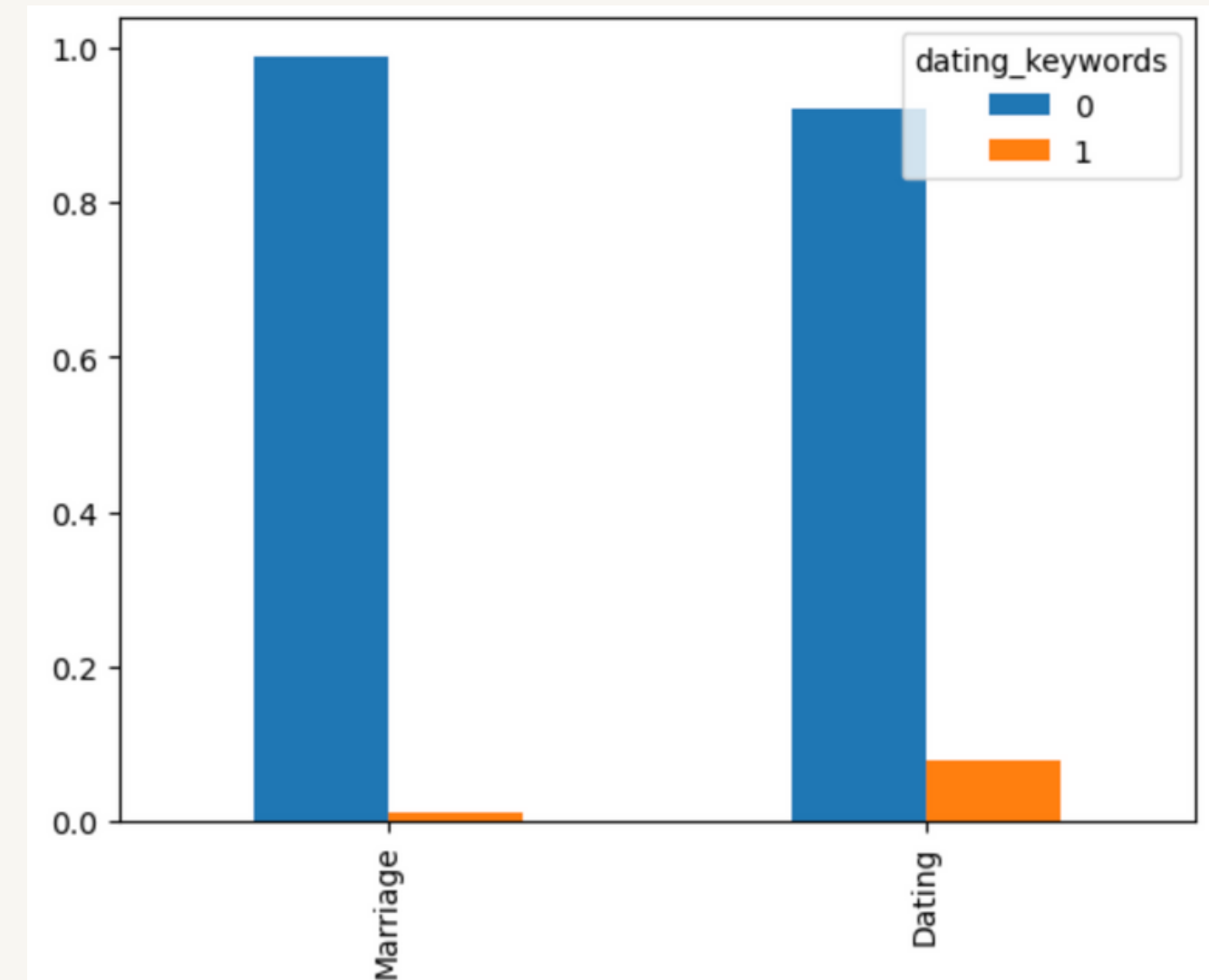
How specific?



Marriage keywords:

- Husband, Wife, Wedding, Marry

At least one is present in 40% of marriage's posts' texts

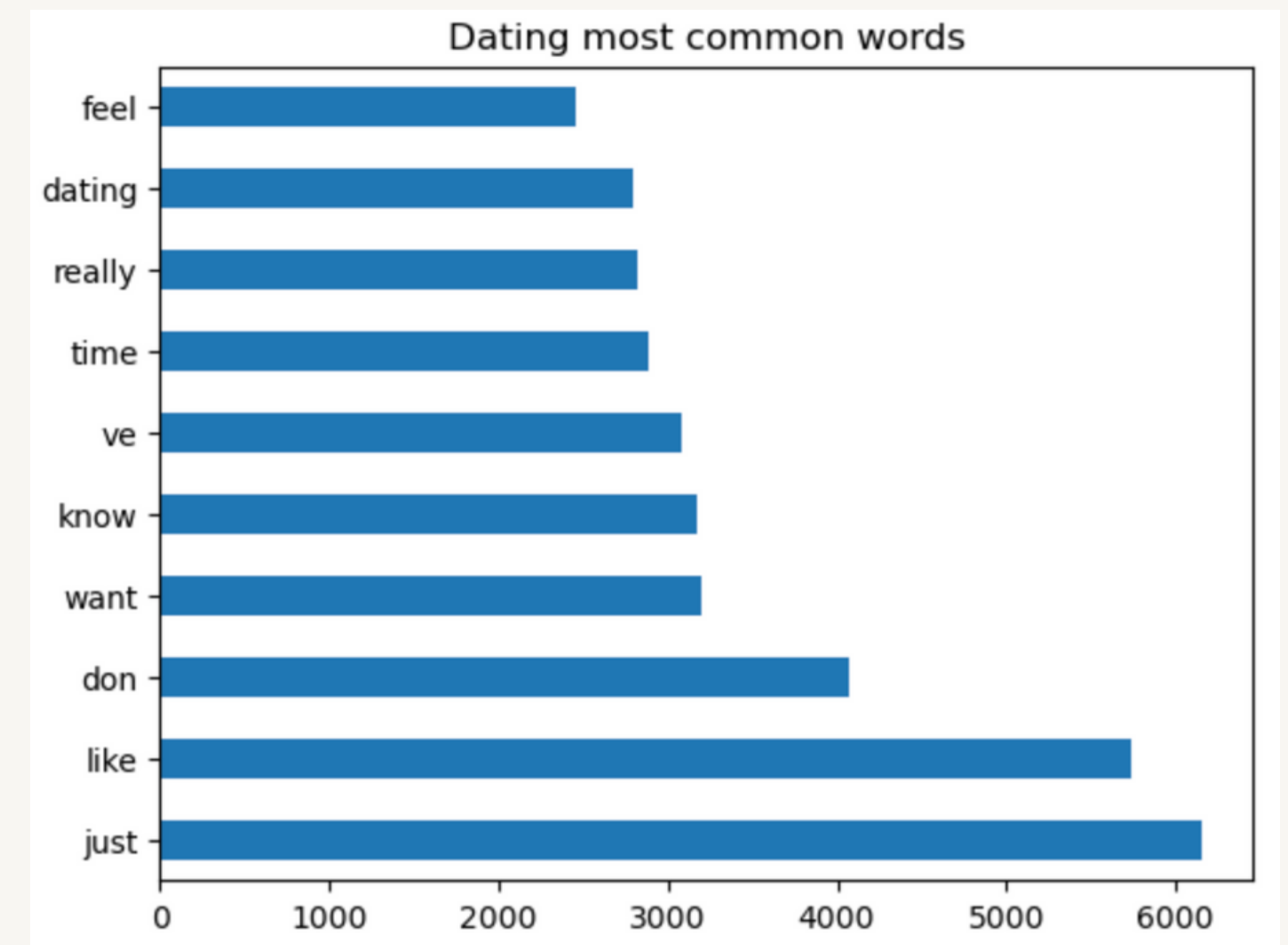
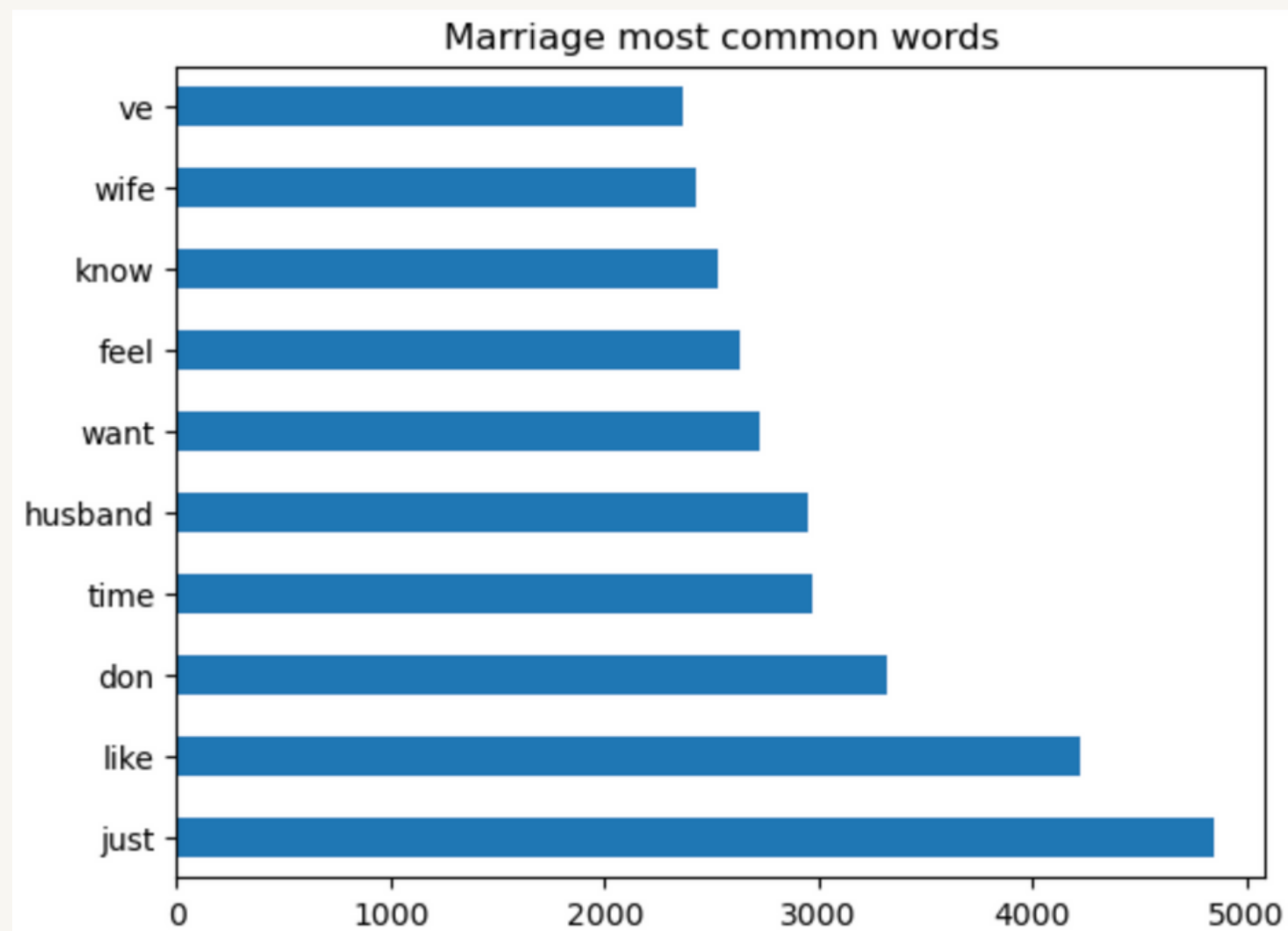


Dating keywords:

- Boyfriend, Girlfriend

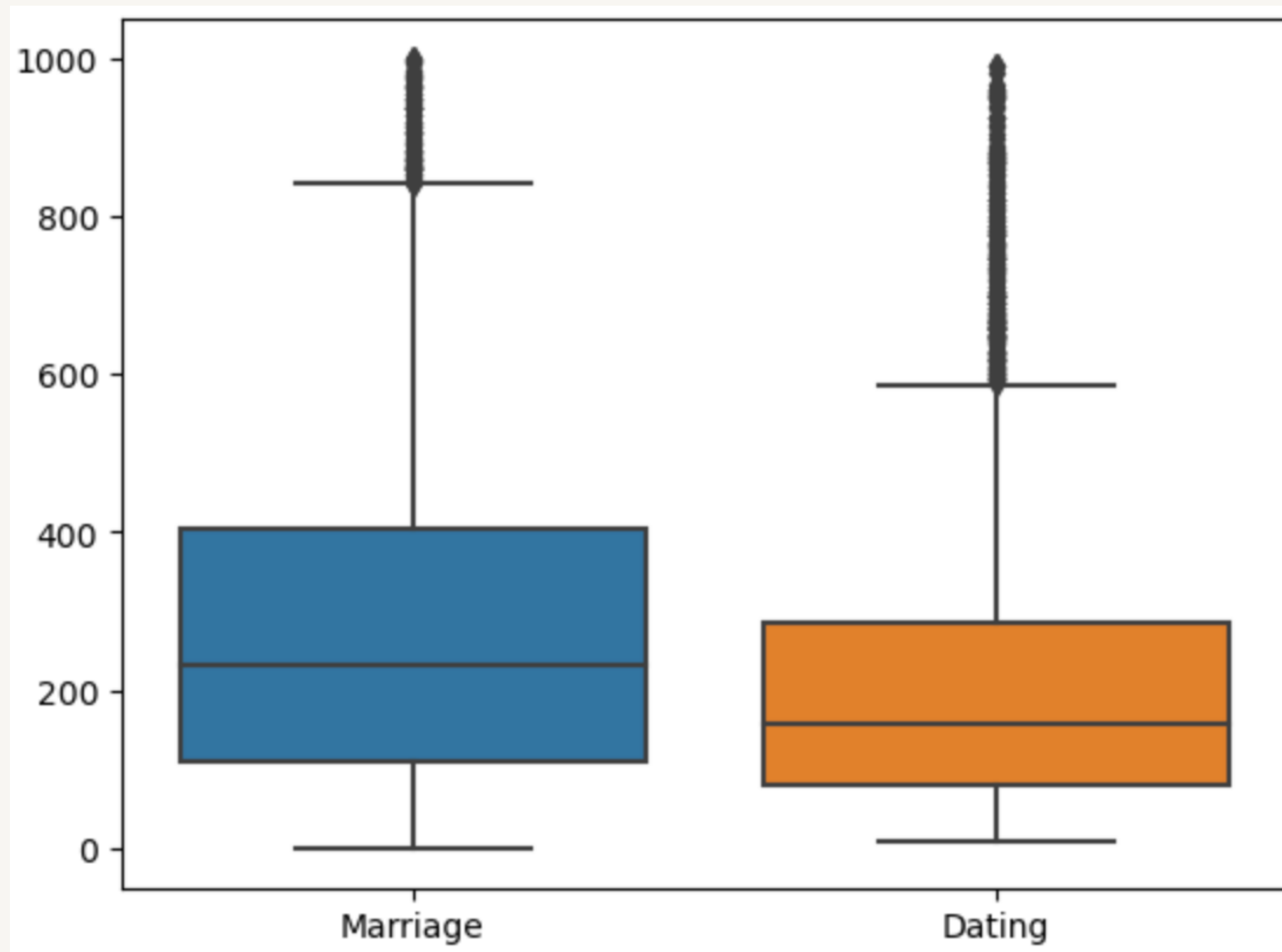
At least one is present in less than 20% of dating's posts' texts

How specific?



After removing stop words (words with no meaning without context)
Words most common are pretty similar, except for specific ones

Special Characters



Marriage subreddits usually use more special characters in the posts' texts

!@#\$%&*;?

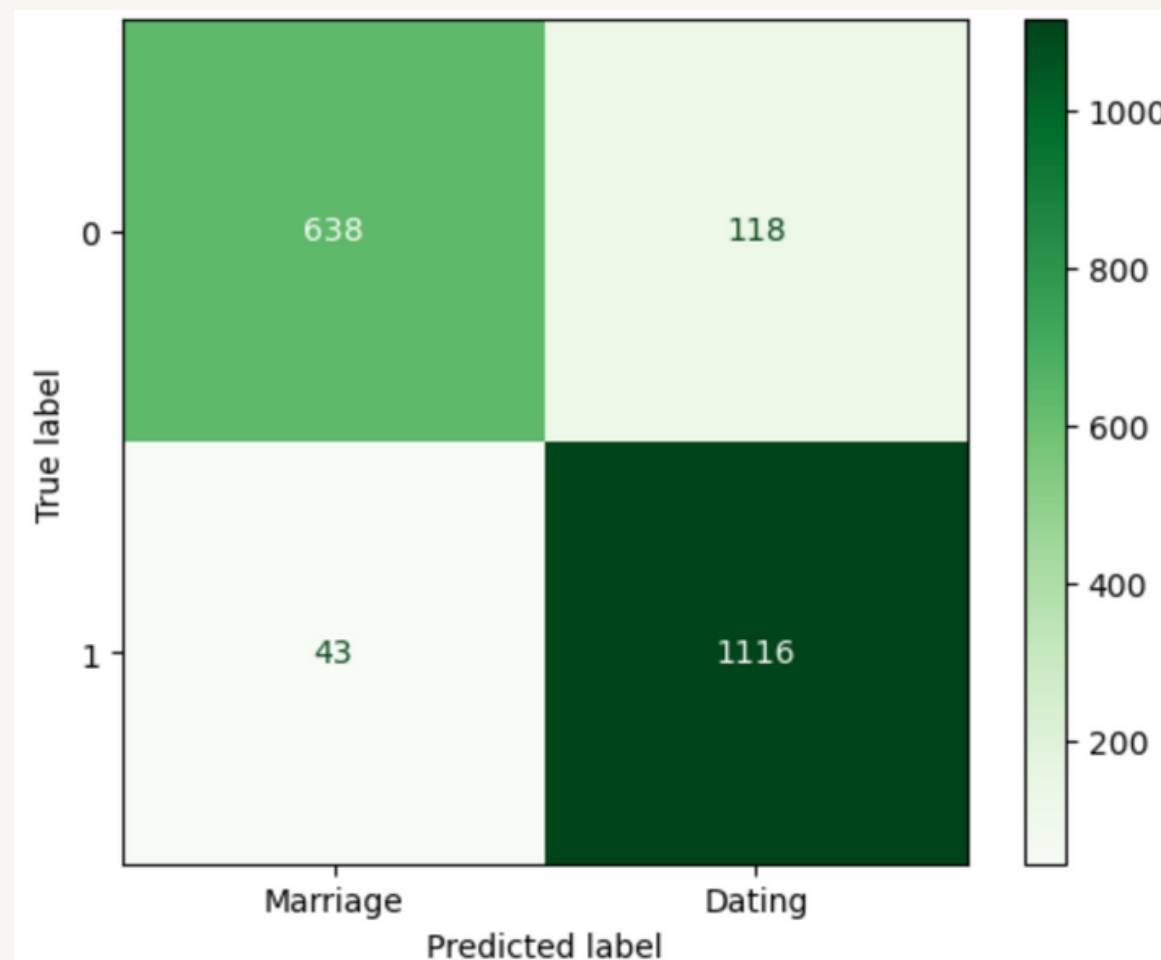
Sentiment Analysis

	compounds	neg	neu	pos
subreddit				
0	0.60305	0.087	0.783	0.117
1	0.71635	0.075	0.782	0.132

Sentiment Analysis scores'
median are pretty similar
between subreddits

Model containing only sentiment analysis: score equal to the baseline

Based on posts' texts



Considering and not considering keywords: accuracy decreased more than 5%

- Husband, Wife, Dating
- 91.4 to 86.3% accuracy

Conclusion and Recommendations

- Further investigation of the structure to better understand their concerns
- Further investigation if the posts' focus more on past or future
- There are differences even when considering only words and excluding keywords, showing their concerns are different

Any questions?