

EDA ON SHANG CHI TWEETS

Castillo, Moreno, Santos



01

Context

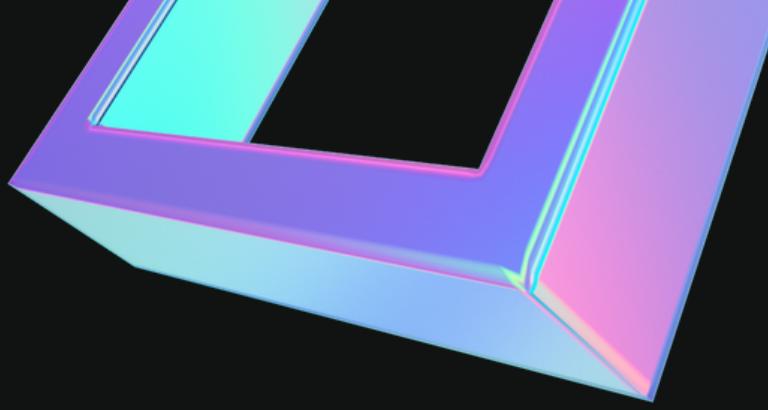
SHANG-CHI



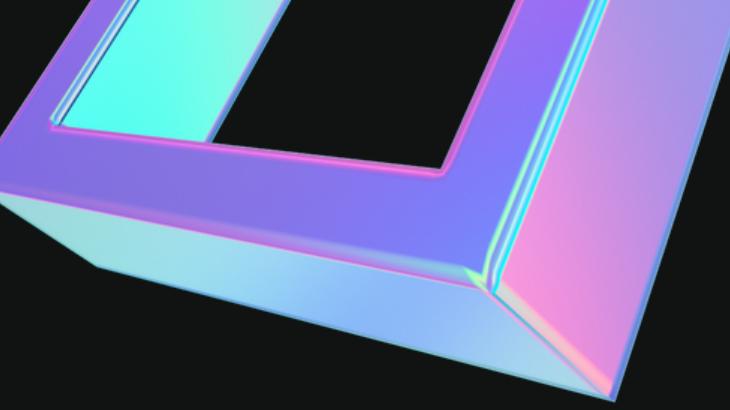
- Released on September 3, 2021.
- First Asian marvel super hero movie! (cool)
- More than just a successful movie, it presents representation.
- The film earned \$94 million domestically over its opening weekend, breaking the box-office record for a Labor Day release.
- It has a Rotten Tomatoes rating of 92%.
- Critics have said that the movie is a perfect mix of culture and Marvel action sequences.

Sources:

https://www.rottentomatoes.com/m/shang_chi_and_the_legend_of_the_ten_rings/reviews?intcmp=rt-what-to-know_read-critics-reviews



**How was the movie
Shang Chi received
by the people of
twitter?**

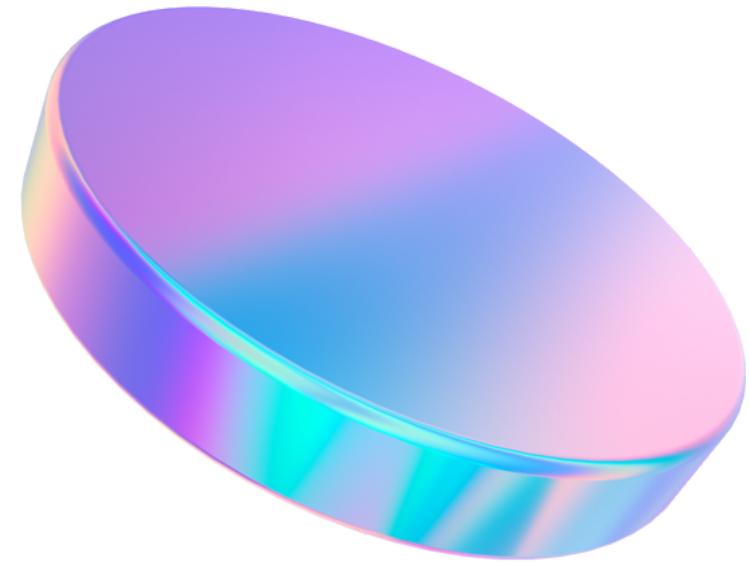


Hypothesis:
The movie was
received very
positively

Dataset

	date	time	username	tweet	language
0	10/15/2021	7:57:56	wolforbewolfed	Finally saw Shang Chi!!! It was dope. Dragons ...	en
1	10/15/2021	7:55:28	ging_hive	one of these days i will watch shang chi	en
2	10/15/2021	7:51:30	spideysbeloved	@BENSHAPIROFAN69 oh no... i could choose my favo...	en
3	10/15/2021	7:50:32	sudthequeen	@eupenobre Shang chi	en
4	10/15/2021	7:40:52	kadoochie	I still need to watch Shang Chi	en

- Social media platform: Twitter
- Tweet dates from October 6 until October 15
- keyword: Shang Chi
- limited to the english language
- total of ~10,000 tweets
- No specific demographic



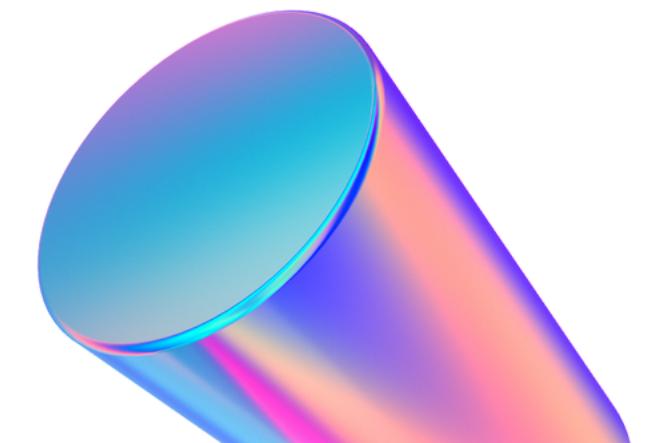
Preprocessing

tweet	language	cleaned_text
@MarvelStudios you fooled me once with space K...	en	bad
Quick movie reviews: No Time To Die: Brilliant... ...	en	much stereotypical good
@ParadoxThaGod 1.)Avengers Infinity War 2.)Ave...	en	civil

preview of adjectives only cleaned_text

- stop word removal
 - NLTK english stopwords
 - iterations of preprocessing:
 - removing other movie titles
 - keeping only adjectives
 - made use of NLTK POS tagging

- lemmatization
- removing trailing spaces, converting to lowercase
- removing other characters



Methods used for analysis

3-4 NGRAM

To see what type of phrases will arise from discussing this movie

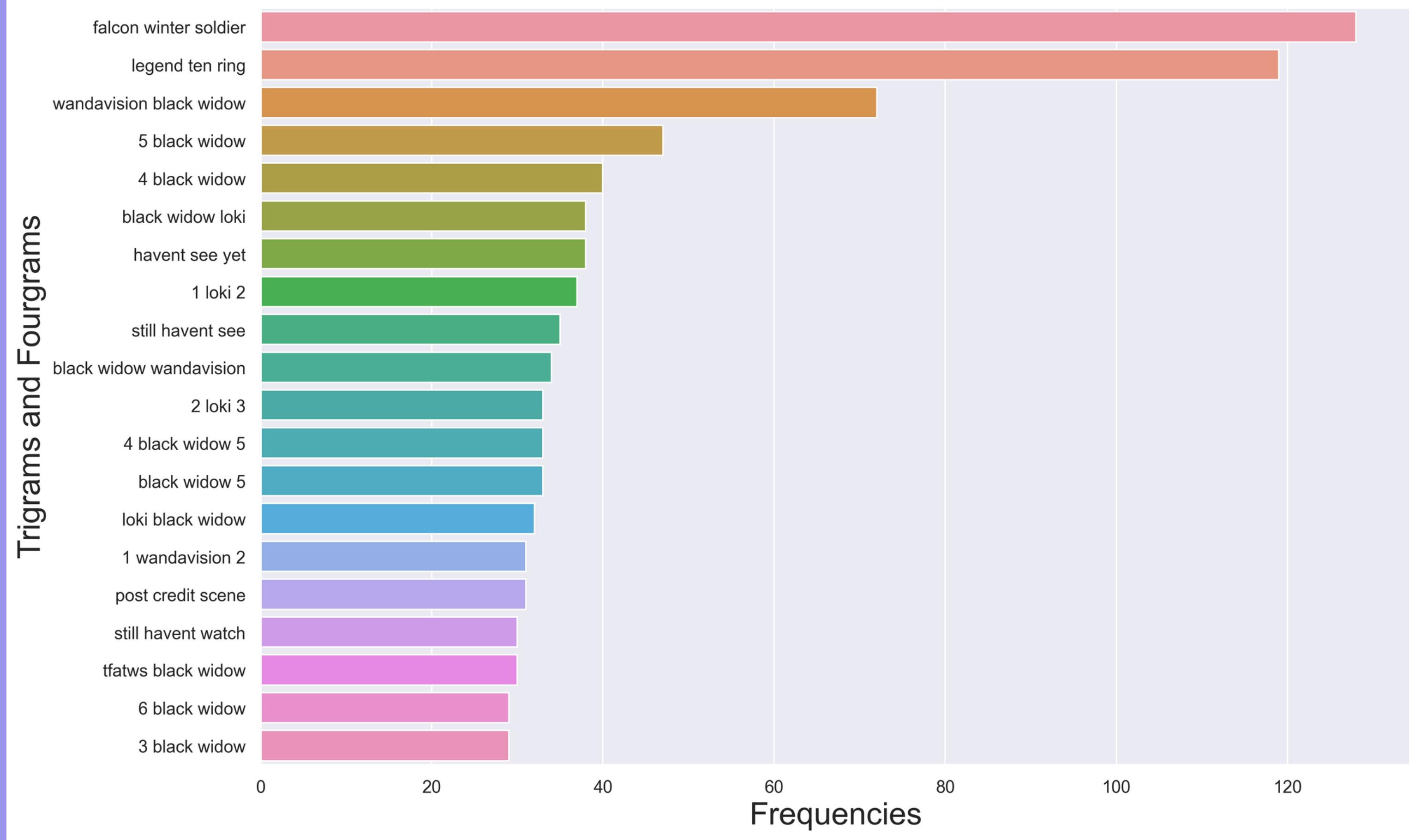
TFIDF

To check the most important/relevant terms across all the tweets/documents that we have collected

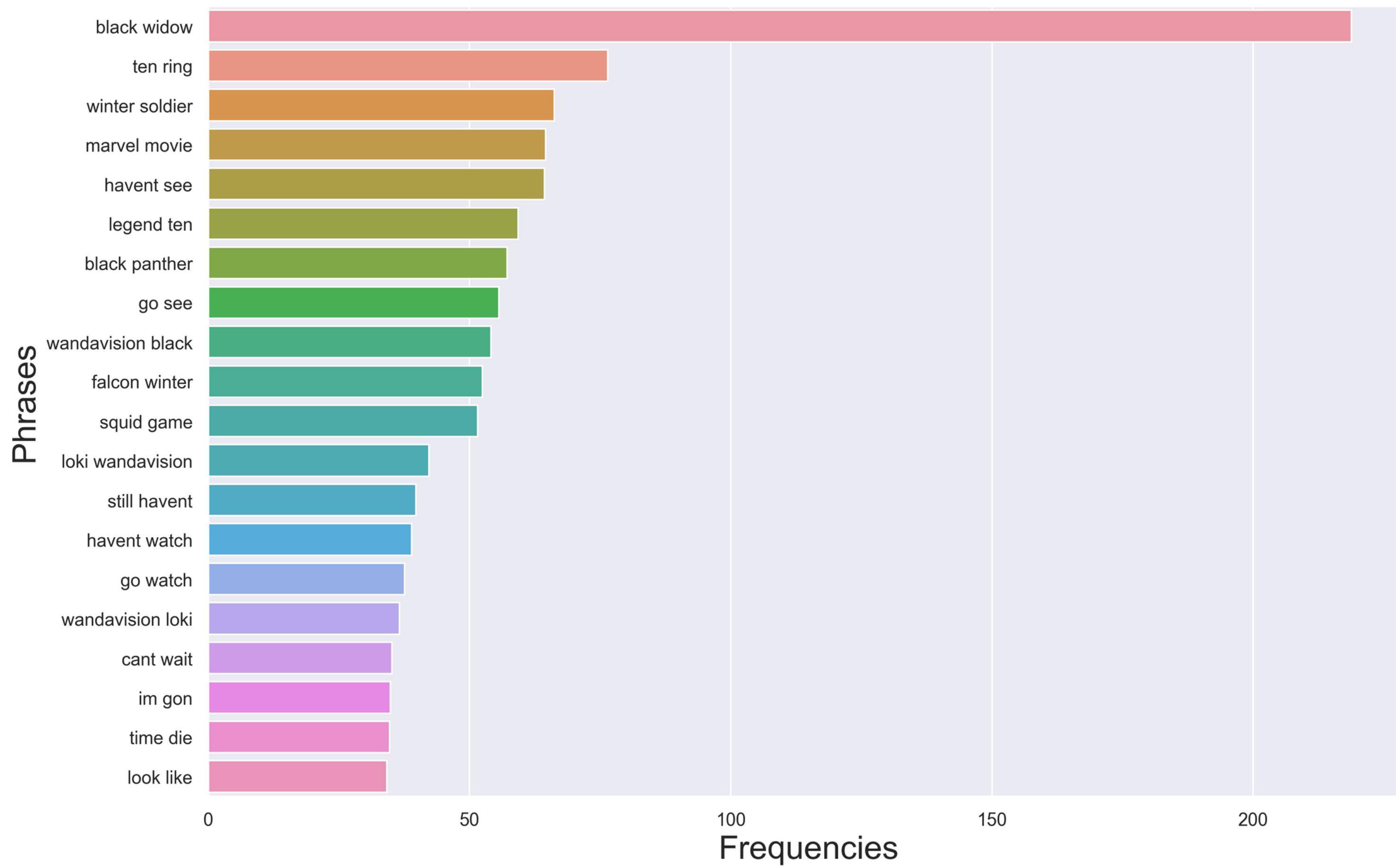
WORD CLOUD

To simply see the most repetitive words from all the tweets.

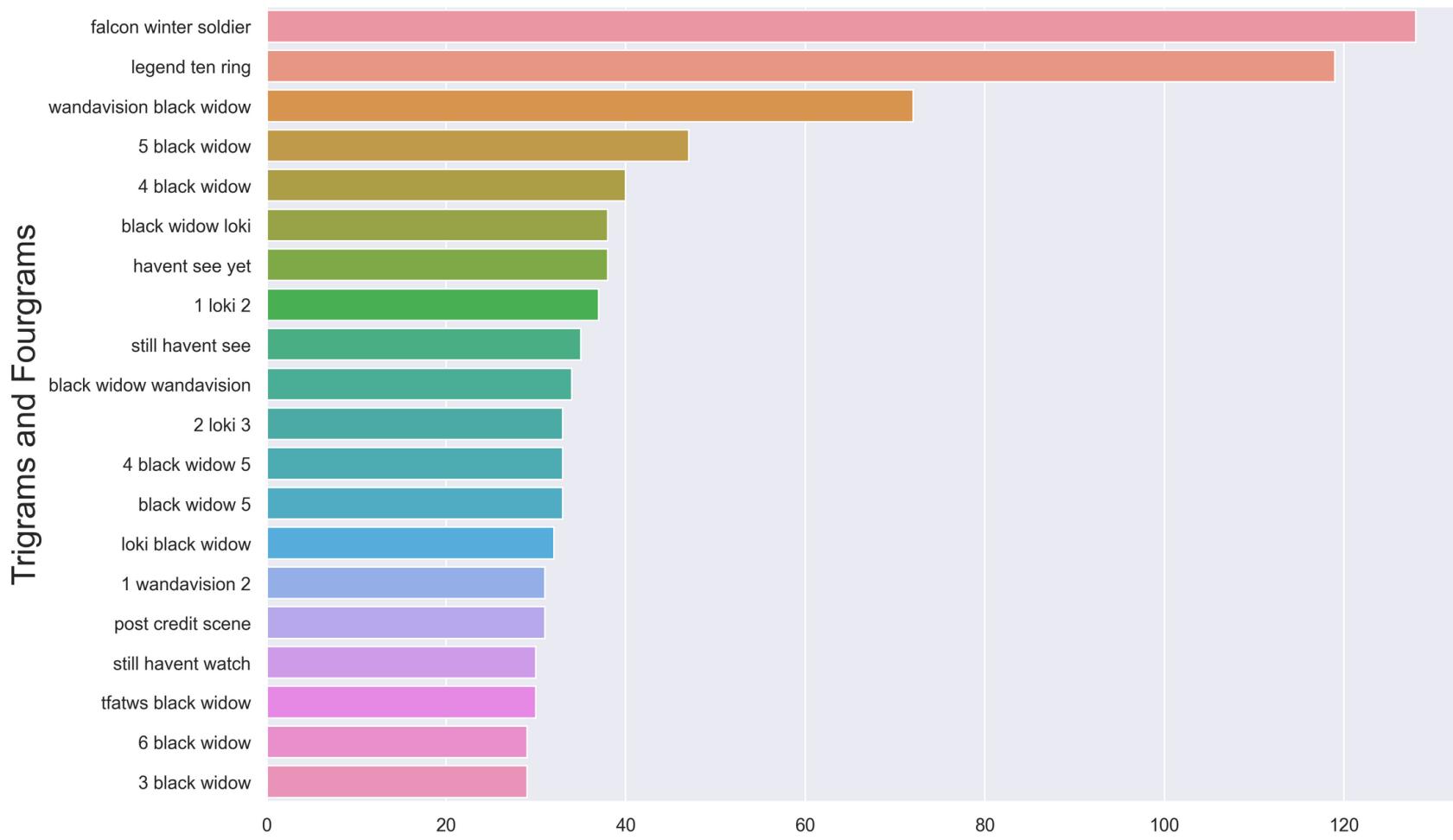
3-4Gram - Shang Chi



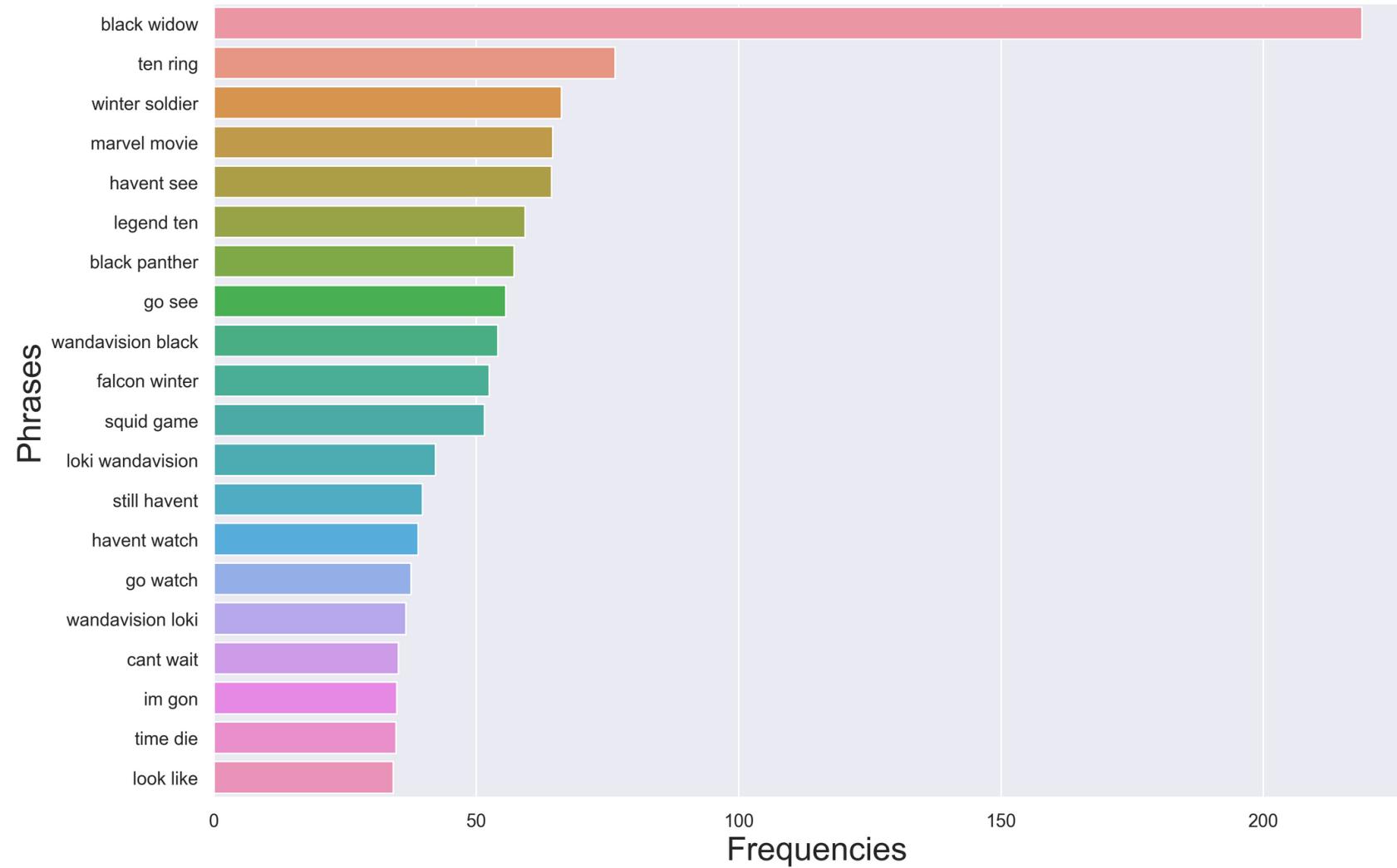
TFIDF - Shang Chi



3-4Gram - Shang Chi

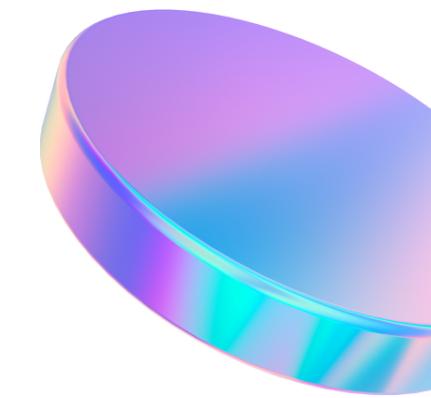


TFIDF - Shang Chi

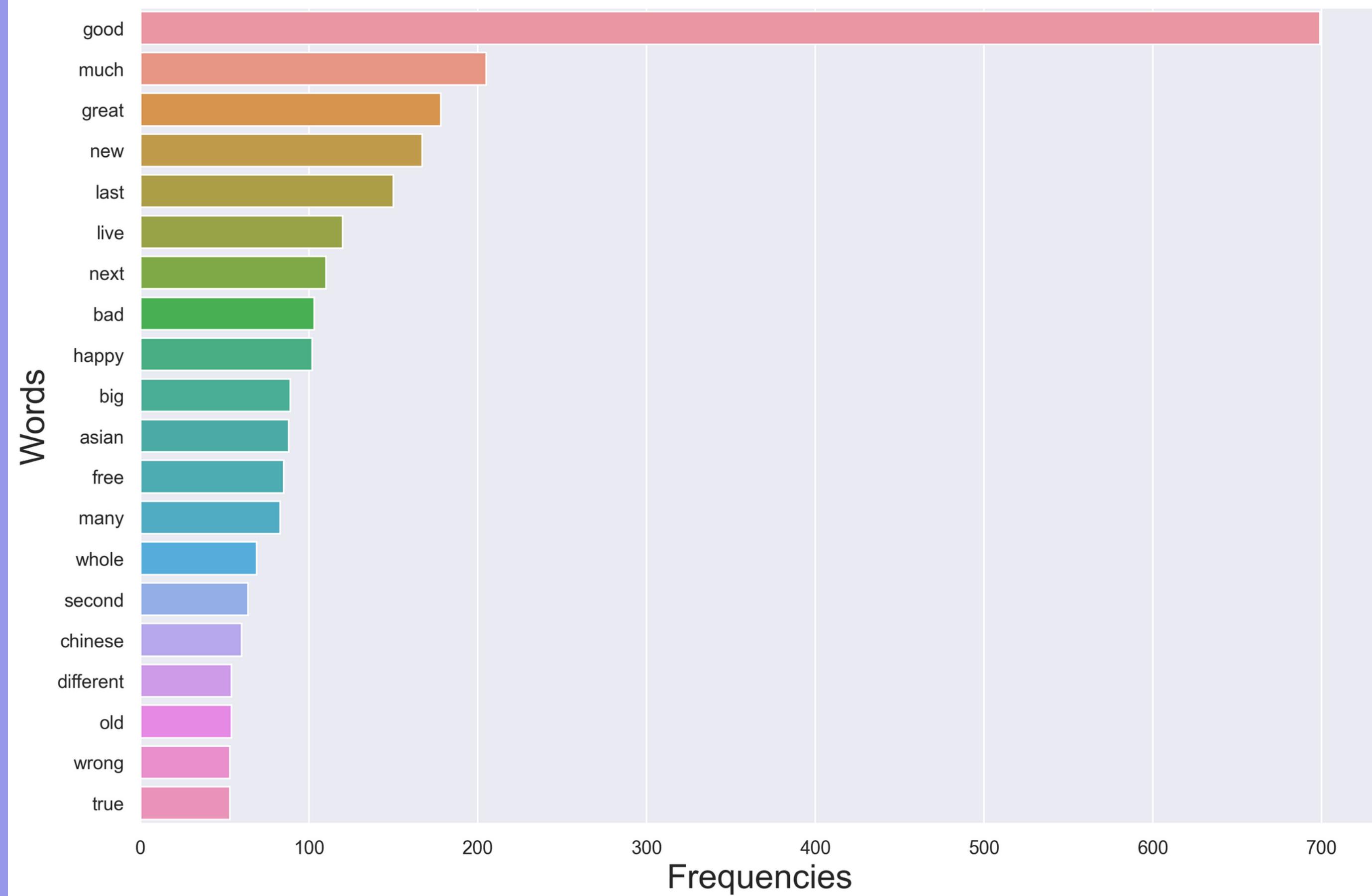


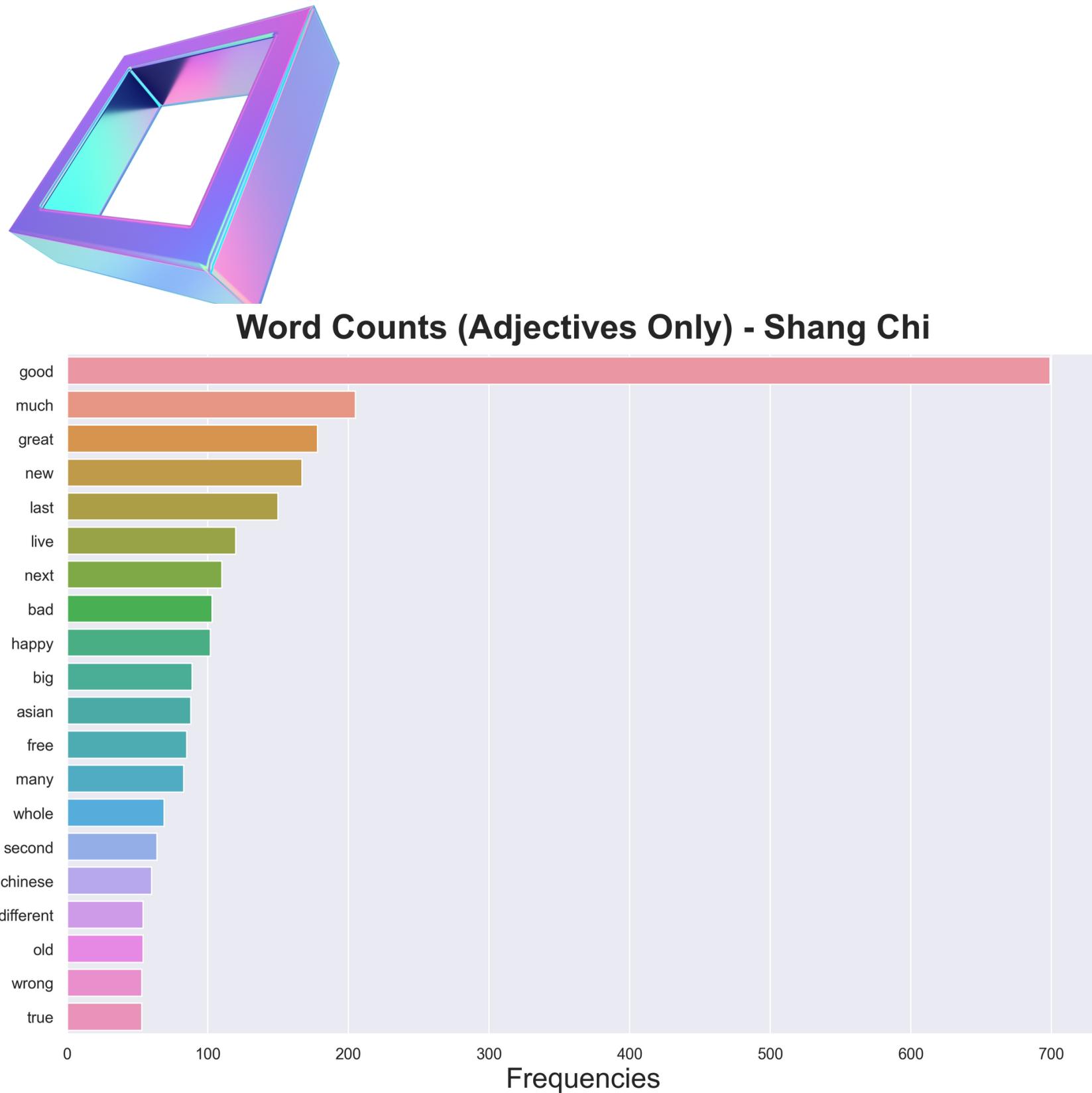
Analysis

- Other marvel movie titles emerged:
 - black widow, wandavision, winter soldier
 - some twitter movie fanpages have prompts asking people to rank films
- Sales comparison for film openings (during the pandemic)
 - black widow: \$80M
 - Shang chi: \$71.4M
- A lot of "go watch" / "go see" terms
 - sign of good perception
 - led us to create another iteration of analysis for adjective only words



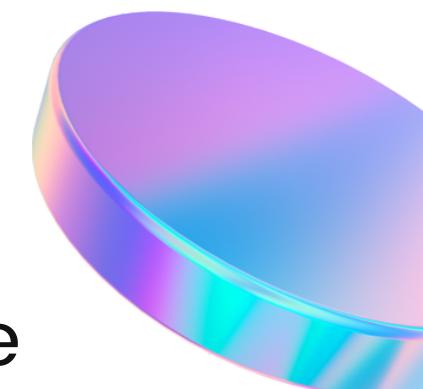
Word Counts (Adjectives Only) - Shang Chi



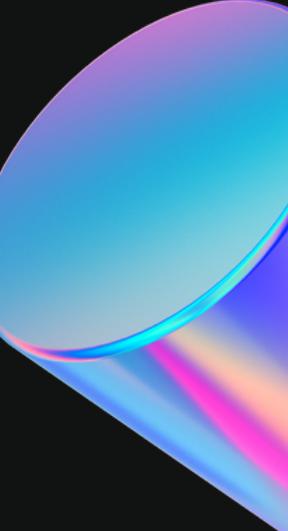
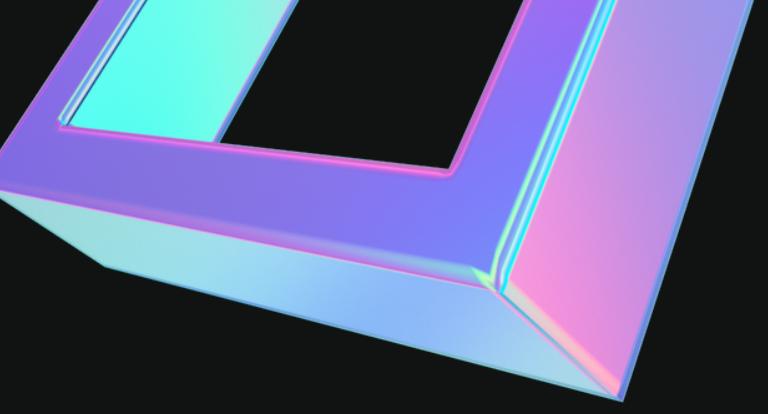


Analysis

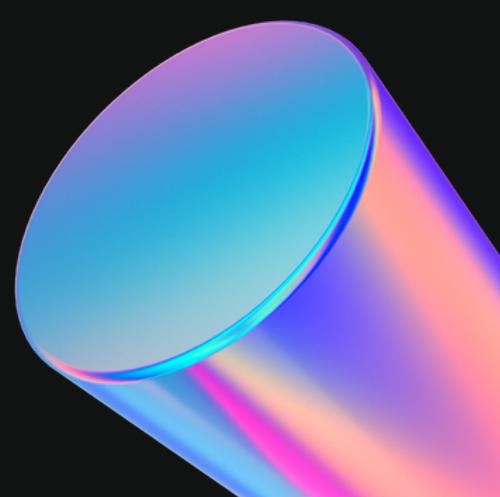
- Most frequently used words are "good", "much" and "great"
 - shows that many people say it's an enjoyable and "must watch" movie.
- "second"
 - second highest box office hit!
- "last"
 - potentially watched "last time" due to the short move run
 - comparison for the "last movie",
 - amazing "last fight"
- "asian" and "chinese"
 - for the positive and accurate representation!







**In conclusion:
Shang Chi did
receive mostly
positive feedback!**



Thank you!