



Analyzing Tweets: Cannes Film Festival



Amy Kim
June 25th 2022





TABLE OF CONTENTS

. . .

01

INTRODUCTION

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

. . .

02

METHODOLOGY

03

RESULTS

04

CONCLUSION & FURTHER WORK

.

.



NLP

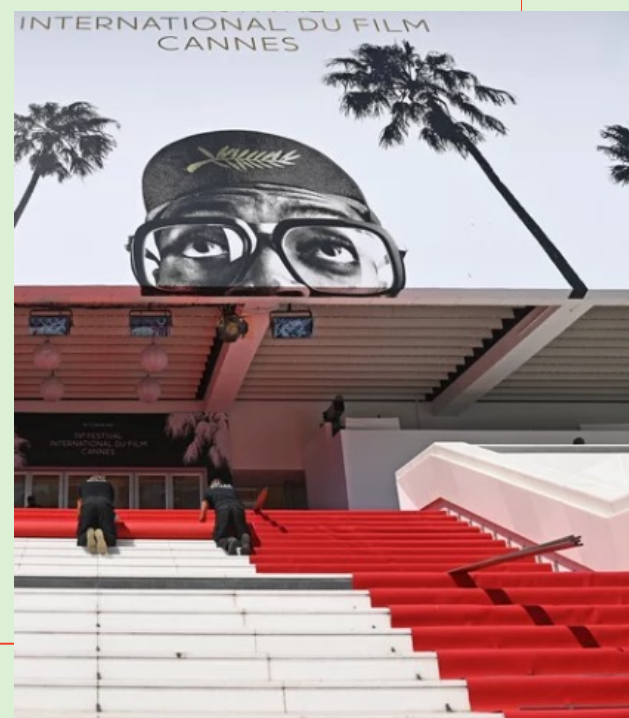


INTRODUCTION

Cannes Film Festival
May 17 – May 28 , 2022

GOAL:

Find which movies were discussed
the most, what was the sentiment &
conversation around these.





METHODOLOGY

1. Data Collection & cleaning

- Twint to scrape tweets ~29k
- Regex for text cleaning
- Stopword removal



3. Topic modelling

- NMF for identifying movies
- Sub-topics within movies were found

2. Tokenization and Vectorization

TF-IDF vectorization with
n_words =1

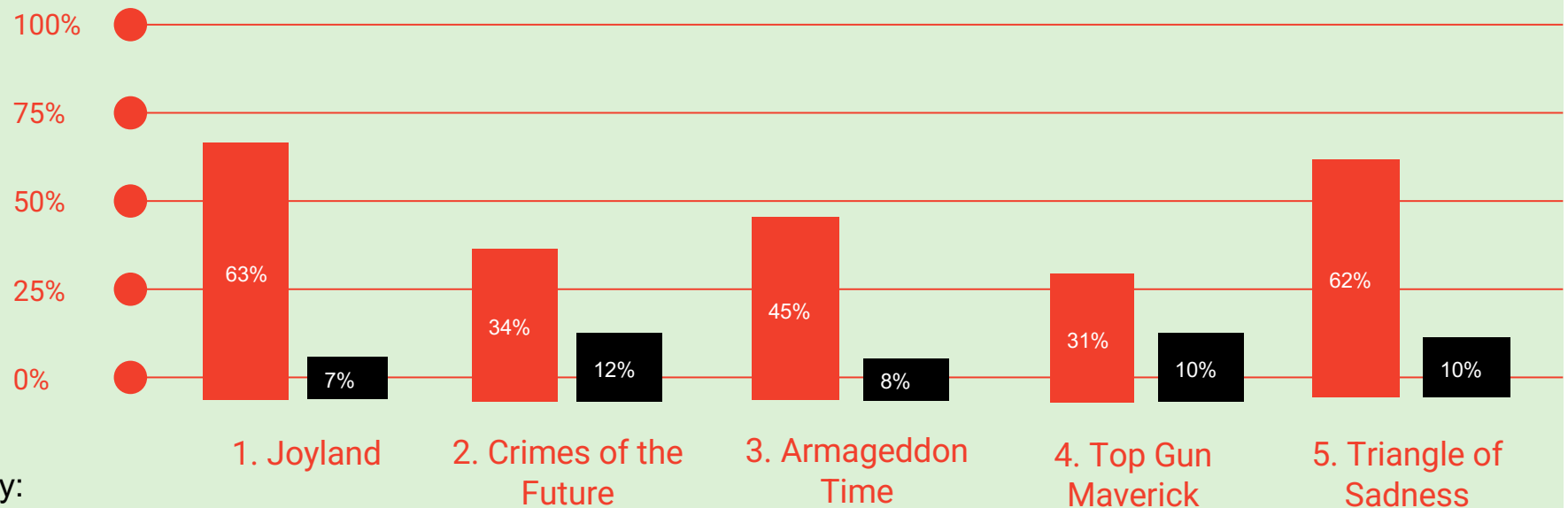
4. Sentiment Analysis

- NLTK VADER
- SpaCy for adjective tagging

NLTK

spaCy

TOP 5 TWEETED MOVIES



Key:

positive

negative



Most negative tweets from VADER Sentiment Analysis

tweet	topic_name	neg
broker team fighting	Broker	0.714
broker fighting broker brokeratcannes2022	Broker	0.714
broker team fighting	Broker	0.714
fuck shit am dead iu leejieun broker leejieuna...	Broker	0.695
mindfuck arthouse horror	Crimes of the Future	0.649
disgusting spectacle of meaningless human glut...	Crimes of the Future	0.612
poor claire denis	The Stars at Noon	0.608



Joyland Subtopics

Standing Ovation

ovation, standing, receives, minute, 12

Jury Prize

jury, prize, certain, regard, wins

Pakistan pride

proud, moment, Pakistan, representing,
Saim Sadiq

Pakistan debut

entry, Pakistan, debut, prolonged, screening

Congratulations

team, congratulations, winning, happy, entire



CONCLUSION



- Able to assess topics and sentiment through various methodology
- **Joyland** was the most talked about and positively tweeted movie

FURTHER WORK

- Try supervised training with a target e.g. can we identify official accounts?





THANKS!

....

Does anyone have any questions?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**.