

FRIENDS

THE ONE WITH THE NLP ANALYSIS

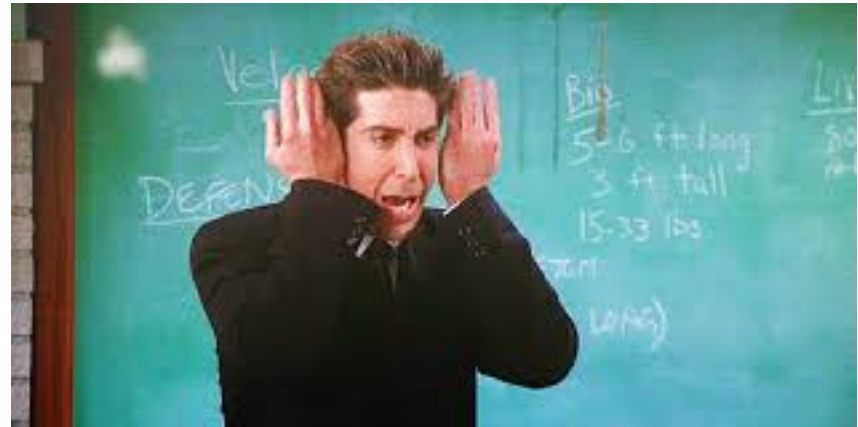
AGENDA

- MOTIVATION
- DATA COLLECTION
- FEATURES
- TEXT PREPROCESSING
- SENTIMENT ANALYSIS
- TOPIC MODELING
- FUTURE IMPROVEMENTS

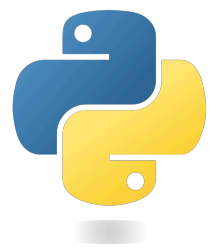


MOTIVATION

- HOW DEEP IS THE SHOW FRIENDS?
- ARE THE CHARACTERS GENERALLY HAPPY?
- WHAT ARE THE COMMON TOPICS?



DATA COLLECTION

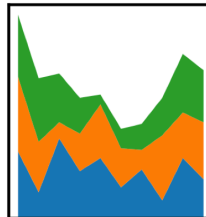
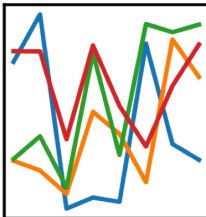
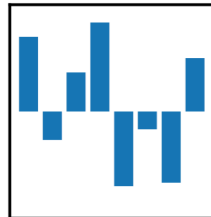


pythonTM

BeautifulSoup

pandas

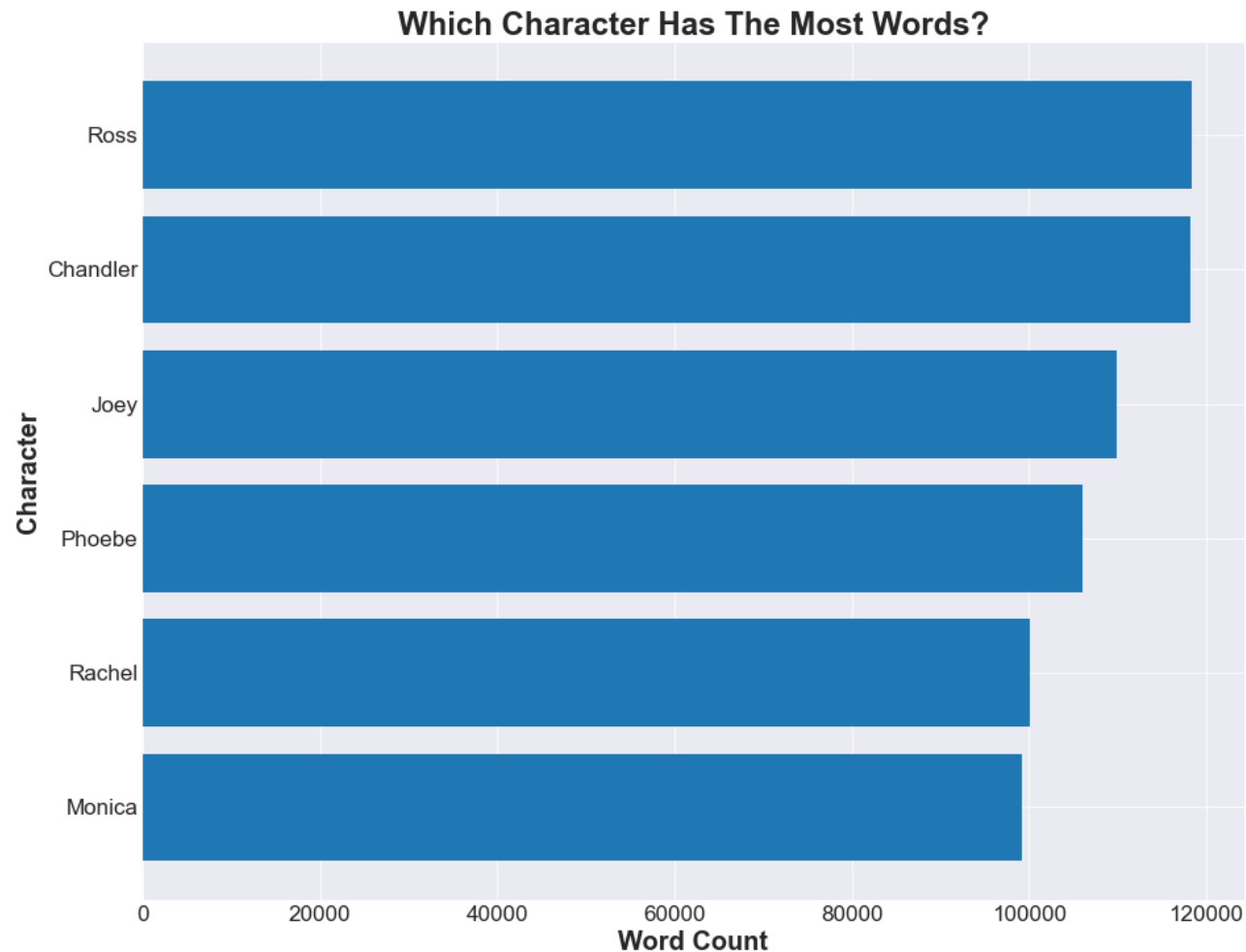
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



FEATURES

FEATURES	COUNT
SEASON	10
EPISODE	236
CHARACTER	800
LINE	61,827

EXPLORATORY DATA ANALYSIS



TEXT PREPROCESSING

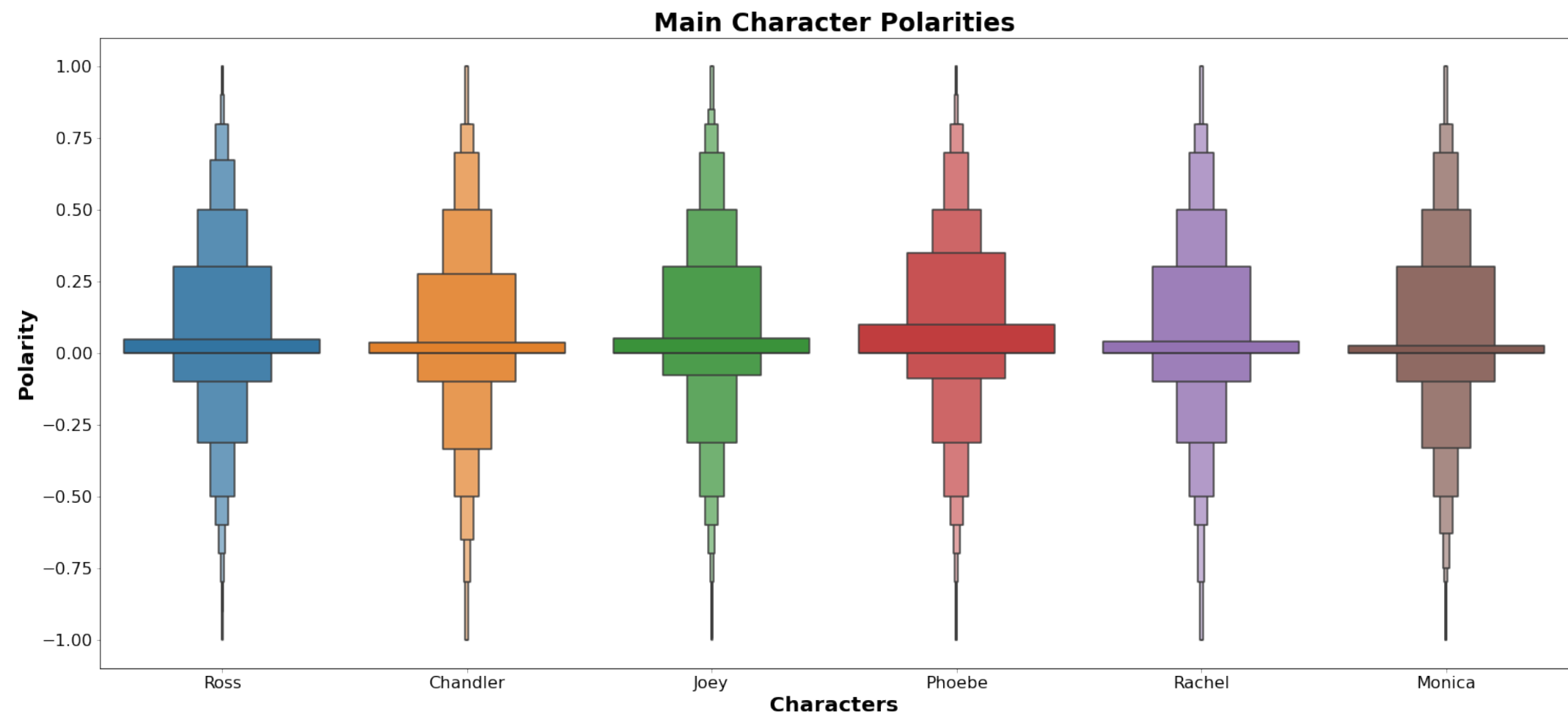
- PUNCTUATION
- CAPITAL LETTERS
- NUMBERS
- STOP WORDS



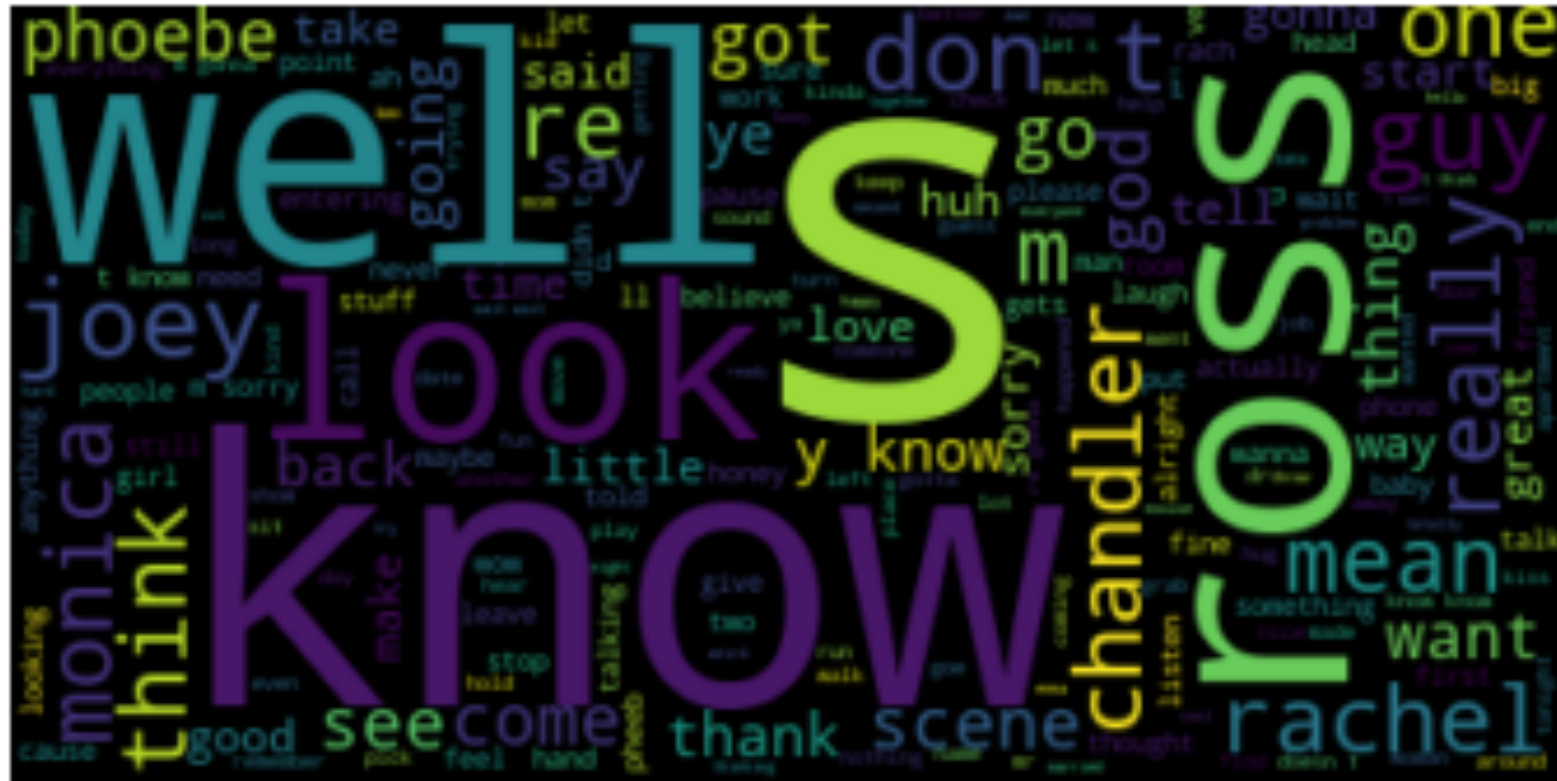
SENTIMENT ANALYSIS



SENTIMENT ANALYSIS



WORD CLOUDS



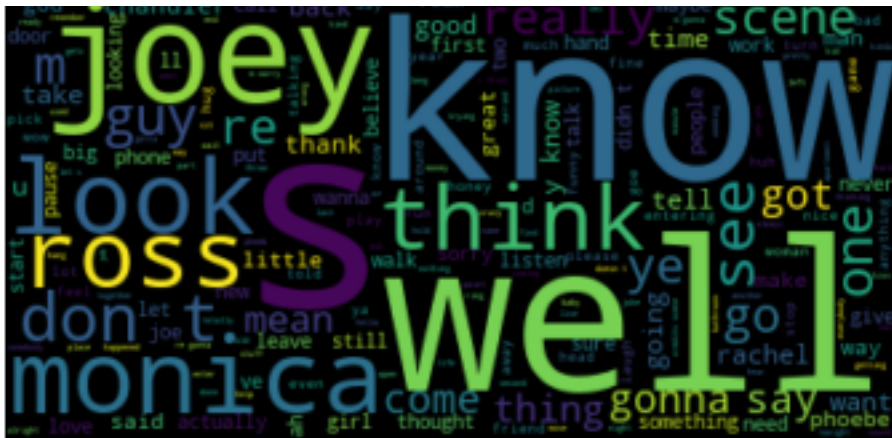
ROSS



RACHEL



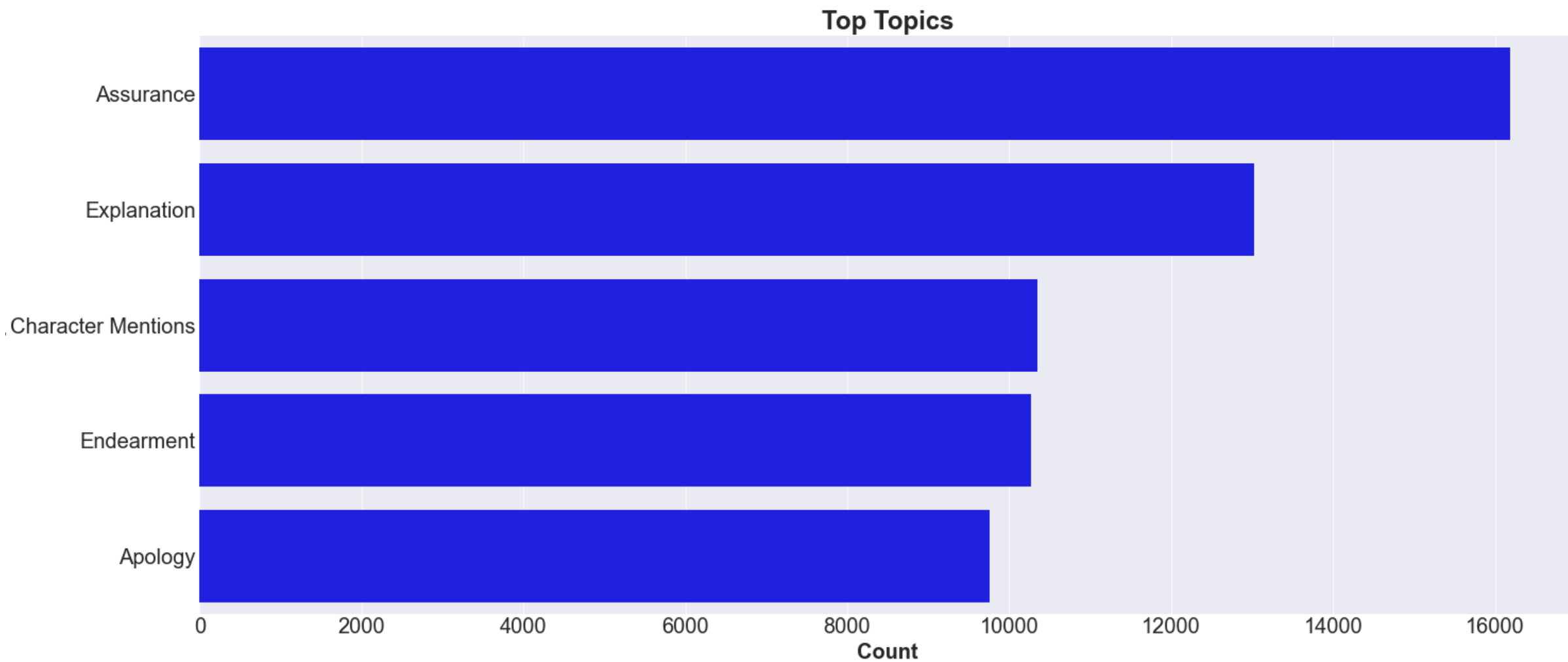
CHANDLER



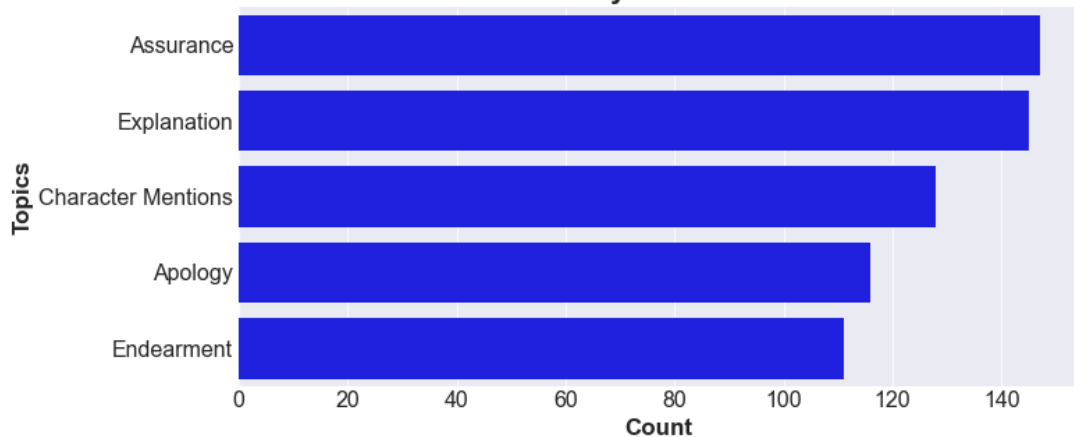
MONICA



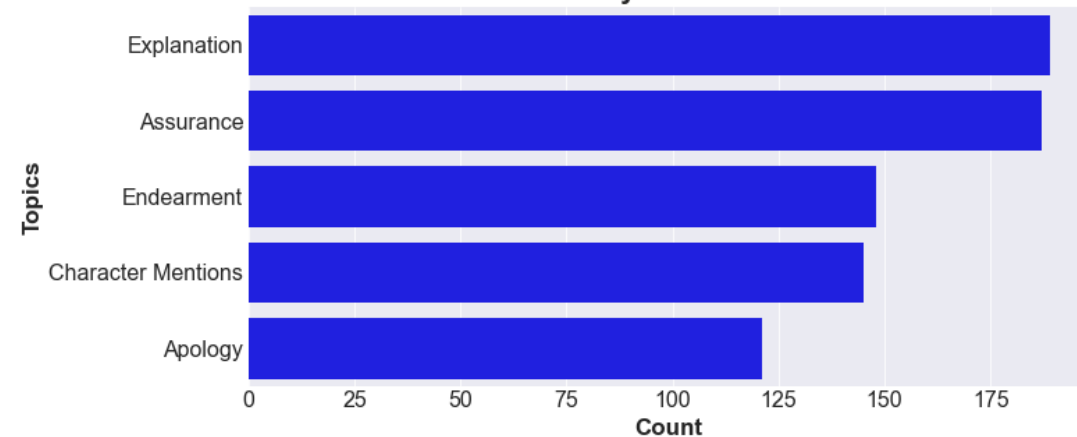
TOPIC MODELING



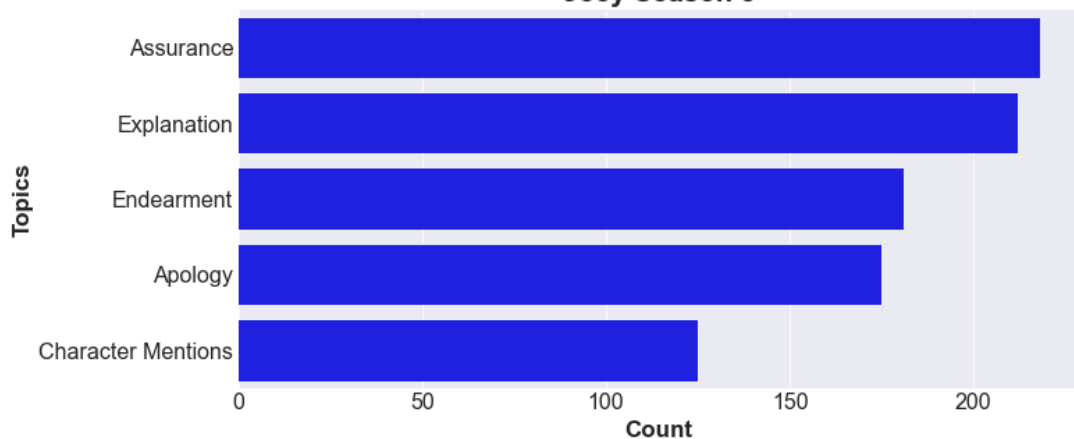
Joey Season 1



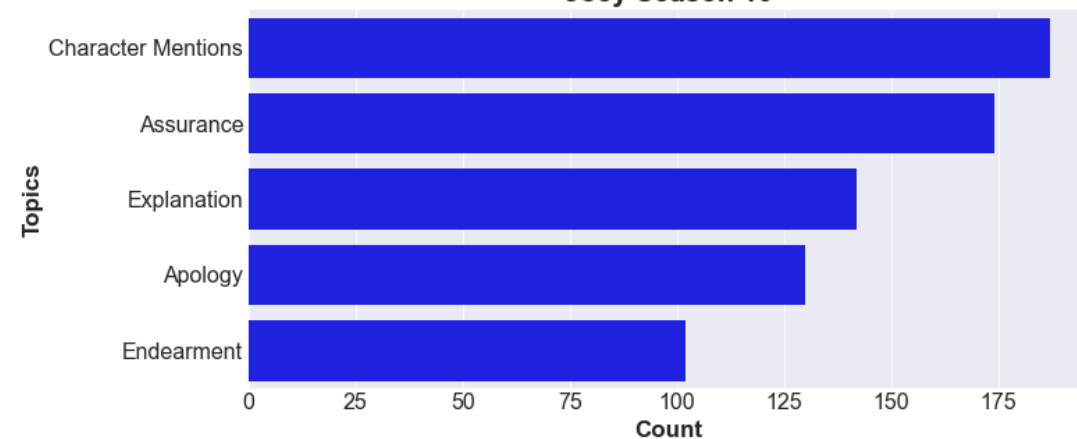
Joey Season 3



Joey Season 8



Joey Season 10



WHAT WE'VE LEARNED



FUTURE IMPROVEMENTS

- SCRIPT CLEANING
- STOP WORDS
- COMPARE WITH OTHER SITCOMS



QUESTIONS?

