Natural Language Processing

Peter Ganong and Maggie Shi

November 20, 2024

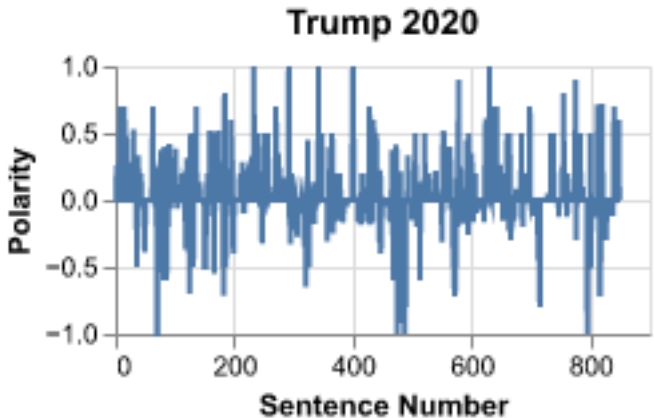
Table of contents I

Do-pair-share: solution

```
with open("files/trump_2020.txt", 'r') as file:
    trump_2020_text = file.read()

doc_trump_2020 = nlp(trump_2020_text)

trump_2020_sentence_polarities = []
for i, sentence in enumerate(doc_trump_2020.sents):
    polarity = sentence._.blob.polarity
    trump_2020_sentence_polarities.append({"n": i + 1, "trump_2020_polarity": polarity})

df_trump_2020 = pd.DataFrame(trump_2020_sentence_polarities)

print(f"Trump 2020 Polarity: {doc_trump_2020._.blob.polarity:.2f}")

chart3 = alt.Chart(df_trump_2020).mark_line().encode(
    x=alt.X('n', title='Sentence Number'),
    y=alt.Y('trump_2020_polarity', title='Polarity')
).properties(
    title = "Trump 2020",
    width=200,
    height=100
)
chart3
```

Do-pair-share: solution

```
Trump 2020 Polarity: 0.07
```



- Relative to 2016, Trump was more negative
- But also had less variance in his sentiment