Text Analysis Report: fall_of_roman_empire_chatgpt

1. Basic Statistics

Metric	Value
Word Count	448
Sentence Count	22
Character Count	2970
Unique Word Count	232
Avg Word Length (chars)	5.46
Avg Sentence Length (words)	20.36
Avg Sentence Length (chars)	135.00
Sentence Length Std Dev (words)	8.32
Lexical Diversity (TTR)	0.5179
MATTR (Window=50)	0.7881

2. Entropy & Complexity
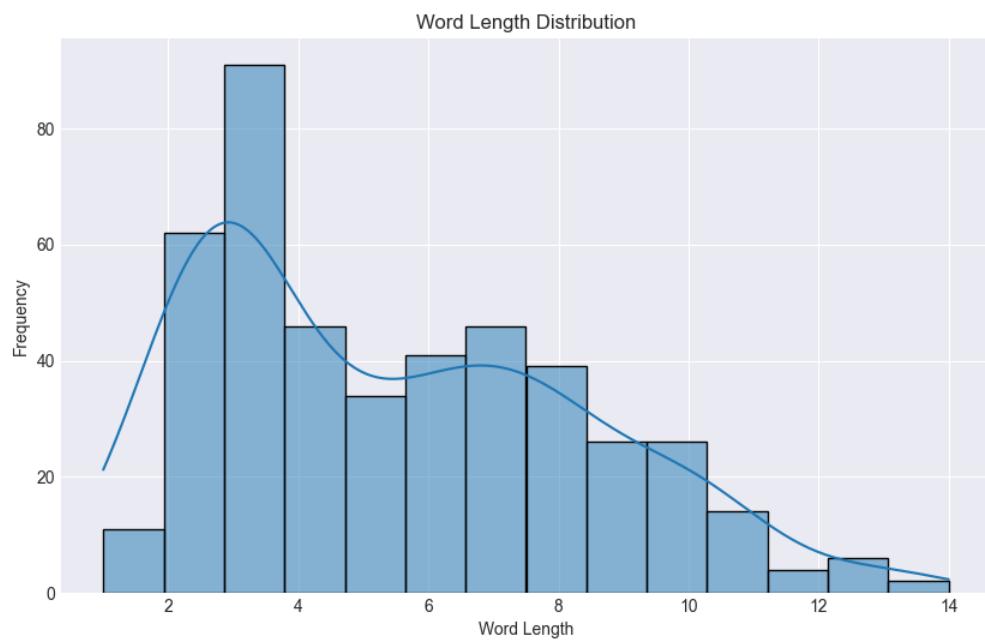
| Metric	Value
Shannon Entropy (Words)	6.9655 bits
Shannon Entropy (Chars)	4.5867 bits
Function Word Ratio	0.3549

3. Readability Scores
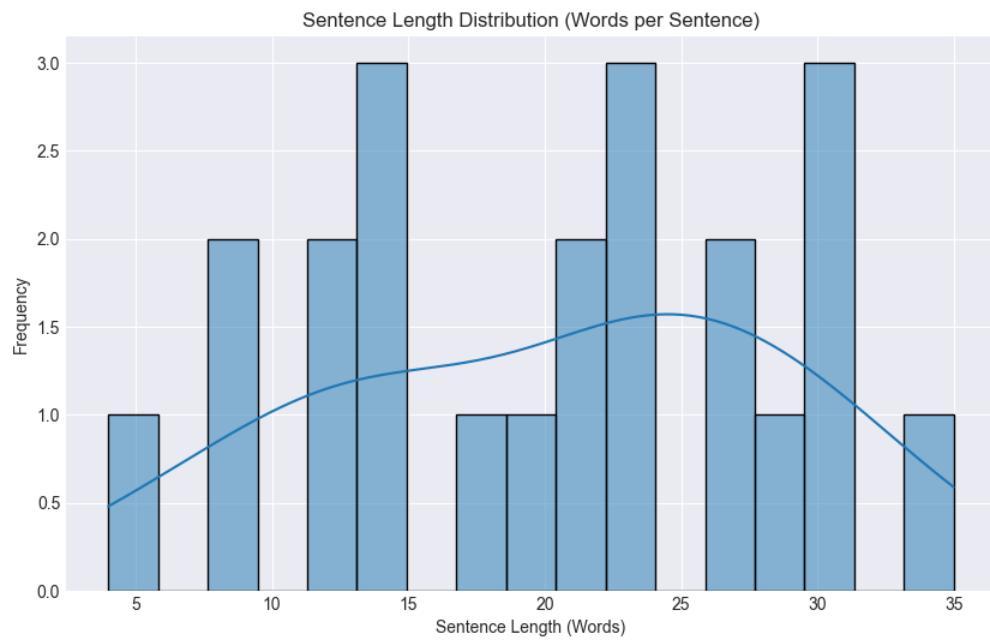
| Index	Score
Flesch Reading Ease	17.34
Flesch-Kincaid Grade Level	15.80
Gunning Fog Index	16.36
SMOG Index	17.00
Coleman-Liau Index	14.91
Automated Readability Index	15.60

4. Visualizations

Word Length Distribution

Word Length Distribution

Sentence Length Distribution
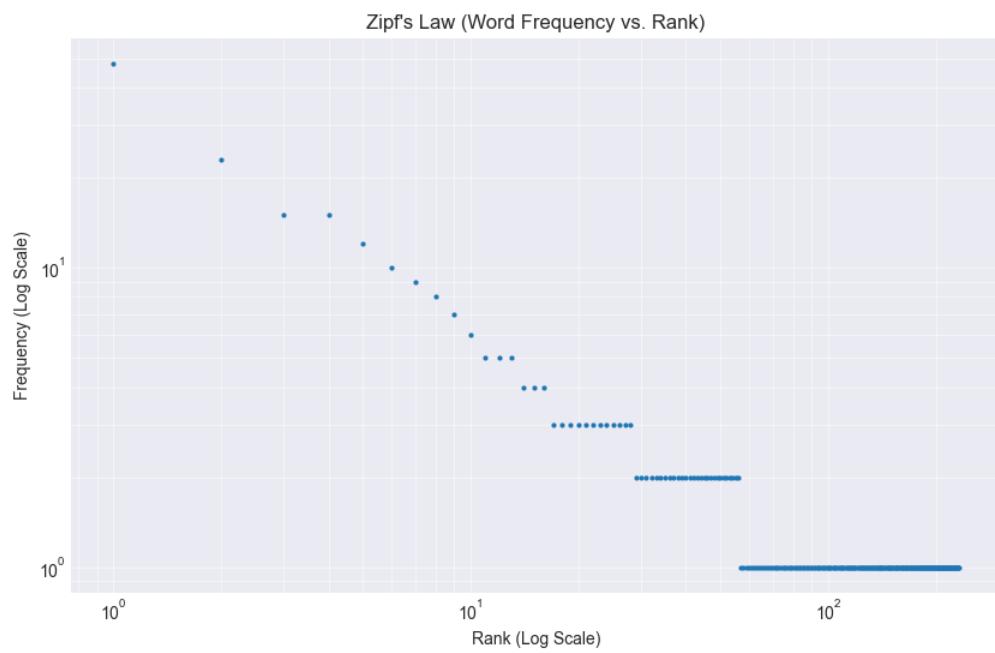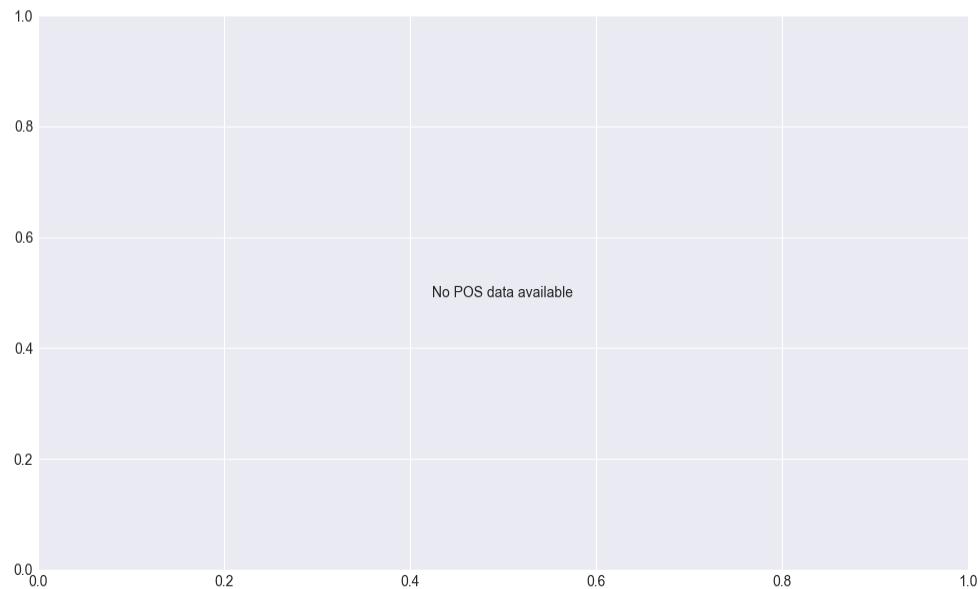
Sentence Length Distribution (Words per Sentence)
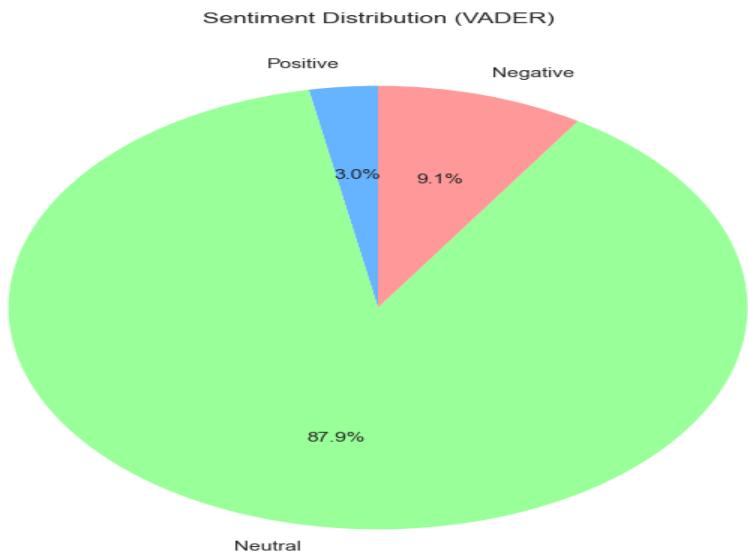
Zipf's Law Analysis (Word Frequency vs Rank)

Zipf's Law (Word Frequency vs. Rank)
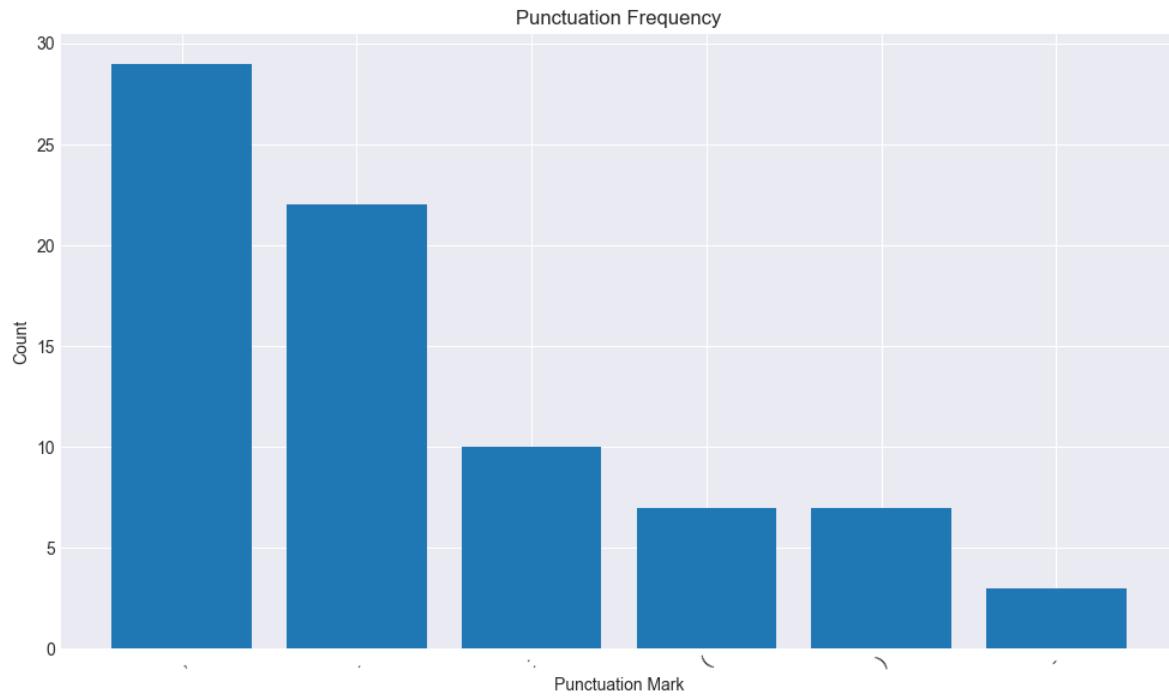
Part-of-Speech Distribution (Top 25)

No POS data available

Sentiment Distribution (VADER)

Sentiment Distribution (VADER)

Punctuation Frequency

5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

the (48), of (23), empire (15), and (15), roman (12), to (10), ce (9), western (8), in (7), fall (6), a (5), from (5), rome (5), decline (4), eastern (4)

Most Common Punctuation:

',' (29), '.' (22), ':' (10), '(' (7), ')' (7), '-' (3)

6. N-gram Analysis

Common 2-grams (Top 10):

'of the' (9), 'roman empire' (9), 'fall of' (6), 'the fall' (5), 'the western' (5), 'western roman' (5), 'to the' (4), 'the empire' (4), 'the roman' (3), '476 ce' (3)
2-gram Repetition Rate: 0.1566

Common 3-grams (Top 10):

'fall of the' (5), 'the fall of' (5), 'the western roman' (5), 'western roman empire' (5), 'of the western' (3), 'of the roman' (2), 'the roman empire' (2), 'the collapse of' (2), 'in 476 ce' (2), 'the medieval period' (2)
3-gram Repetition Rate: 0.0650

Common 4-grams (Top 10):

'the western roman empire' (5), 'the fall of the' (4), 'of the western roman' (3), 'fall of the roman' (2), 'of the roman empire' (2), 'the eastern roman empire' (2), 'many of whom had' (2), 'the roman empire the' (1), 'roman empire the fall' (1), 'empire the fall of' (1)
4-gram Repetition Rate: 0.0292

Character N-gram Analysis:

Common Character 3-grams (Top 10):

```
'_th' (54), 'the' (51), 'he_' (48), 'ed_' (29), '_of' (24), 'of_' (23), 'rom' (23), '_in' (23),
'ion' (21), 'emp' (19)
```

Common Character 4-grams (Top 10):

```
'the_' (48), '_the' (47), '_of_' (23), '_rom' (18), '_emp' (18), 'empi' (16), 'mpir' (16),
'pire' (16), 'tern' (16), 'and_' (16)
```

Common Character 5-grams (Top 10):

```
'_the_' (45), '_empi' (16), 'empir' (16), 'mpire' (16), '_and_' (15), '_roma' (12), 'roman'
(12), 'oman_' (12), 'stern' (12), 'tern_' (12)
```