Text Analysis Report: fall_of_roman_empire_bard

1. Basic Statistics

Metric	Value
Word Count	267
Sentence Count	15
Character Count	1856
Unique Word Count	163
Avg Word Length (chars)	5.76
Avg Sentence Length (words)	17.80
Avg Sentence Length (chars)	123.73
Sentence Length Std Dev (words)	5.74
Lexical Diversity (TTR)	0.6105
MATTR (Window=50)	0.8161

2. Entropy & Complexity

| Metric	Value
Shannon Entropy (Words)	6.7526 bits
Shannon Entropy (Chars)	4.4640 bits
Function Word Ratio	0.3558

3. Readability Scores

Index	Score
Flesch Reading Ease	11.51
Flesch-Kincaid Grade Level	16.00
Gunning Fog Index	17.69
SMOG Index	17.20
Coleman-Liau Index	16.59
Automated Readability Index	16.00

4. Visualizations
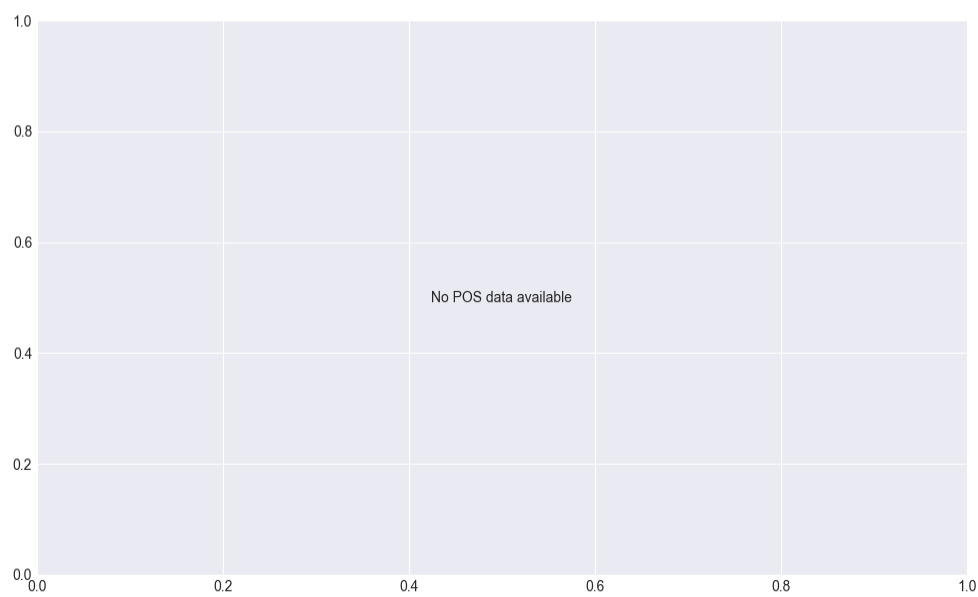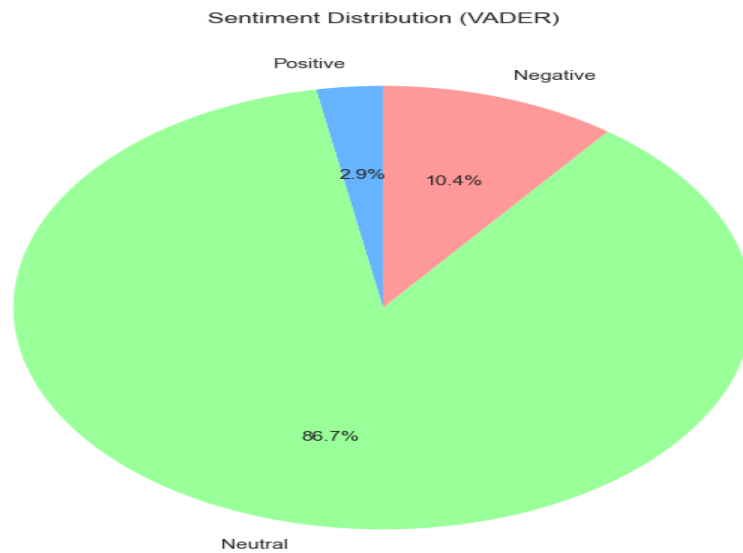
Word Length Distribution

Word Length Distribution

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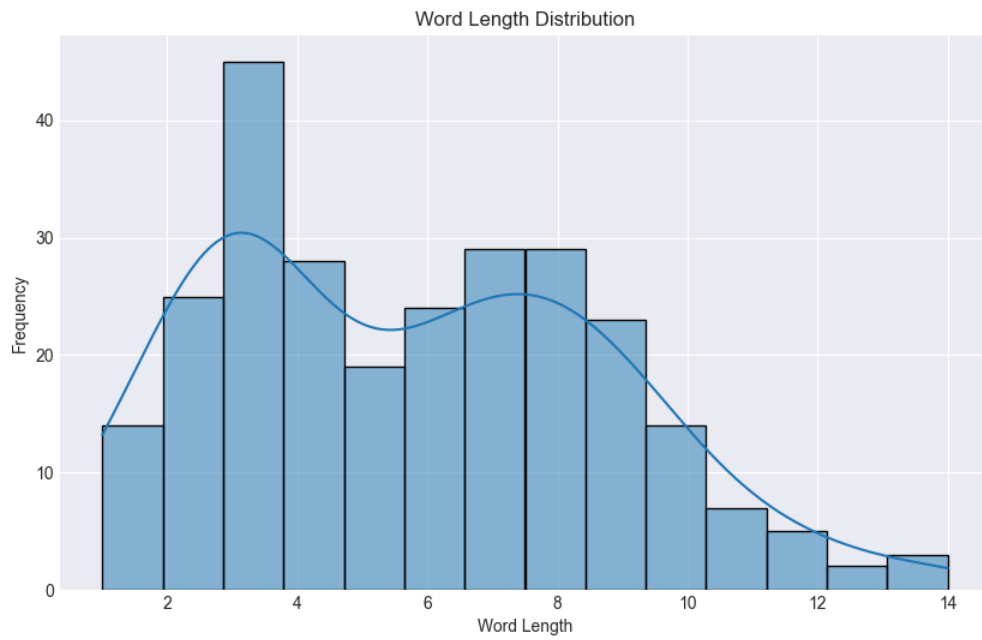
Sentence Length Distribution (Words per Sentence)
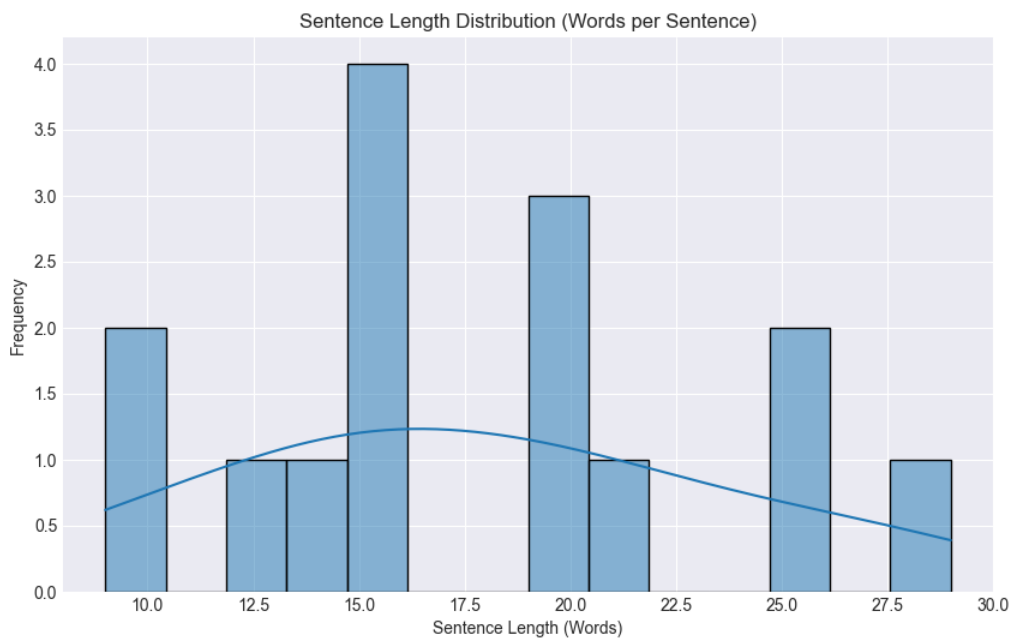
Zipf's Law Analysis (Word Frequency vs Rank)

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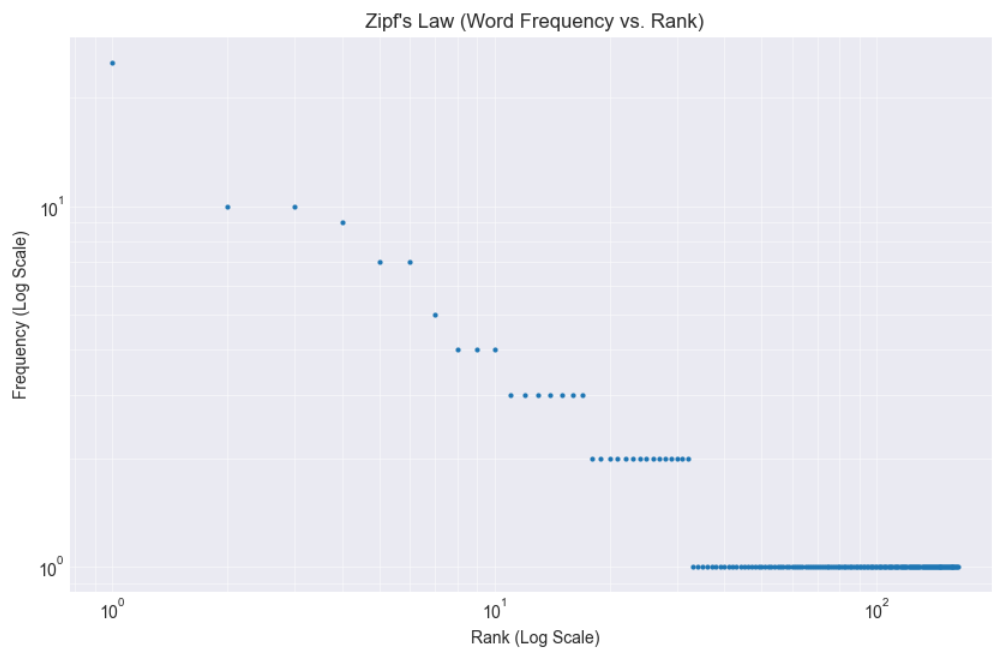
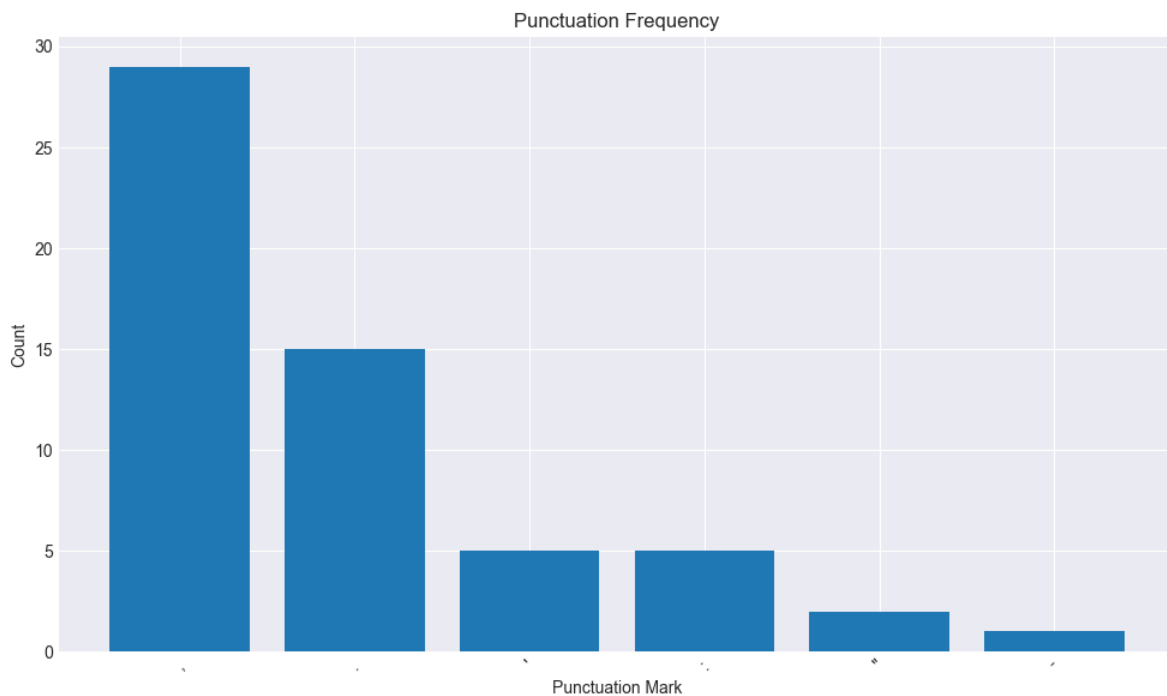## Part-of-Speech Distribution (Top 25)

No POS data available

Sentiment Distribution (VADER)

Sentiment Distribution (VADER)

Punctuation Frequency

5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

the (25), empire (10), and (10), a (9), roman (7), to (7), s (5), of (4), germanic (4), as (4), western (3), was (3), economic (3), it (3), decline (3)

Most Common Punctuation:

',' (29), '.' (15), ''' (5), ':' (5), '"' (2), '-' (1)

6. N-gram Analysis

Common 2-grams (Top 10):

'the empire' (6), 'empire s' (5), 'western roman' (3), 'roman empire' (3), 'as the' (3), 'contributed to' (3), 'the fall' (2), 'of the' (2), 'the western' (2), 'germanic tribes' (2)
2-gram Repetition Rate: 0.0789

Common 3-grams (Top 10):

'the empire s' (5), 'the western roman' (2), 'western roman empire' (2), 'the fall of' (1), 'fall of the' (1), 'of the western' (1), 'roman empire a' (1), 'empire a process' (1), 'a process spanning' (1), 'process spanning centuries' (1)
3-gram Repetition Rate: 0.0226

Common 4-grams (Top 10):

'the western roman empire' (2), 'the fall of the' (1), 'fall of the western' (1), 'of the western roman' (1), 'western roman empire a' (1), 'roman empire a process' (1), 'empire a process spanning' (1), 'a process spanning centuries' (1), 'process spanning centuries was' (1), 'spanning centuries was not' (1)
4-gram Repetition Rate: 0.0038

Character N-gram Analysis:

Common Character 3-grams (Top 10):

'the' (28), '_th' (27), 'he_' (25), 'ed_' (15), 'es_' (13), 'man' (12), '_em' (12), 'emp' (12), 'ion' (12), 'and' (12)

Common Character 4-grams (Top 10):

'the_' (25), '_the' (20), '_emp' (12), 'and_' (11), 'empi' (10), 'mpir' (10), 'pire' (10), 'tion' (10), '_and' (10), 'he_e' (9)

Common Character 5-grams (Top 10):

'_the_' (19), '_empi' (10), 'empir' (10), 'mpire' (10), '_and_' (10), 'the_e' (9), '_roma' (7), 'roman' (7), 'oman_' (7), 'ation' (7)