

Text Analysis Report: theory_of_relativity_chatgpt

1. Basic Statistics

Metric	Value
Word Count	337
Sentence Count	14
Character Count	2349
Unique Word Count	191
Avg Word Length (chars)	5.84
Avg Sentence Length (words)	24.07
Avg Sentence Length (chars)	167.79
Sentence Length Std Dev (words)	6.71
Lexical Diversity (TTR)	0.5668
MATTR (Window=50)	0.8344

2. Entropy & Complexity

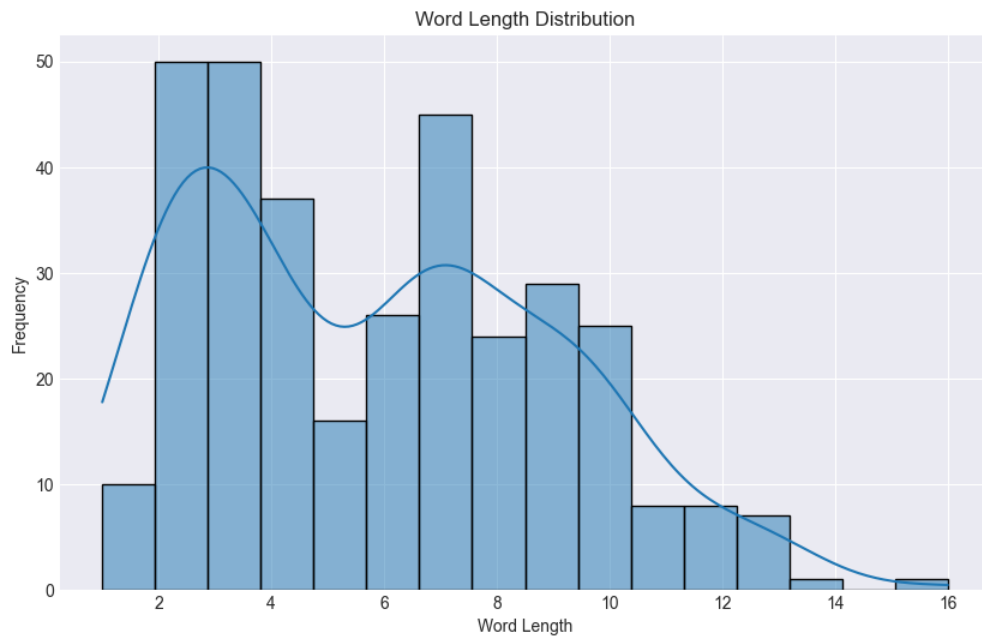
Metric	Value
Shannon Entropy (Words)	6.9733 bits
Shannon Entropy (Chars)	4.4639 bits
Function Word Ratio	0.3591

3. Readability Scores

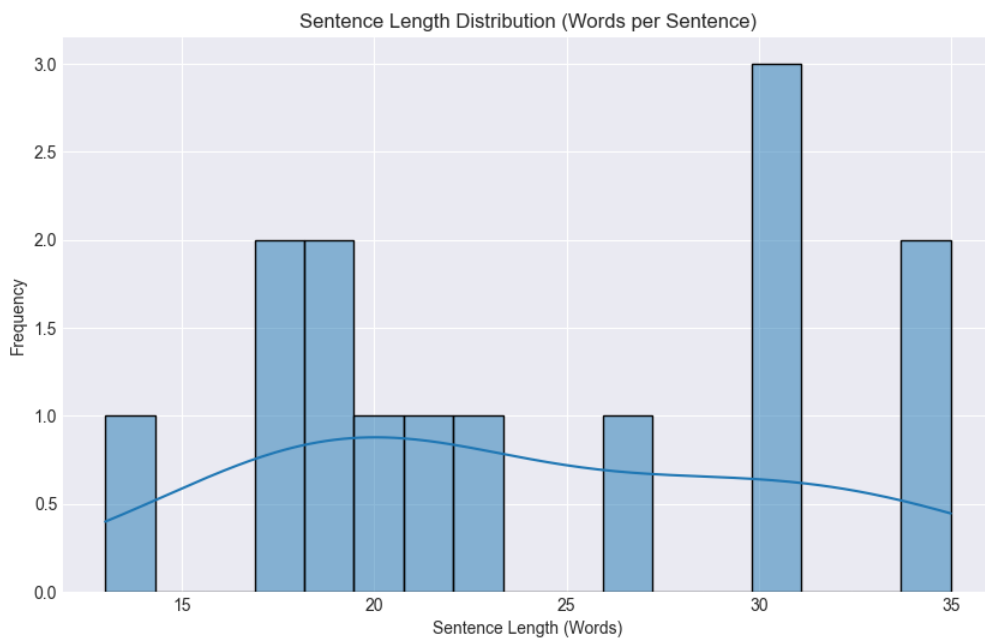
Index	Score
Flesch Reading Ease	13.38
Flesch-Kincaid Grade Level	17.30
Gunning Fog Index	17.68
SMOG Index	18.40
Coleman-Liau Index	17.06
Automated Readability Index	18.80

4. Visualizations

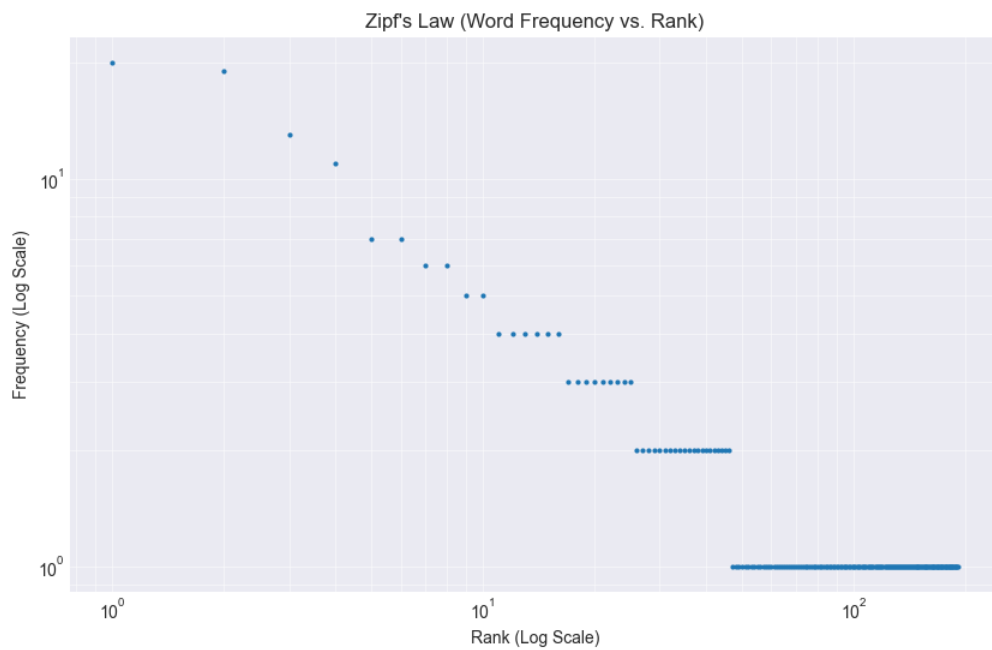
Word Length Distribution

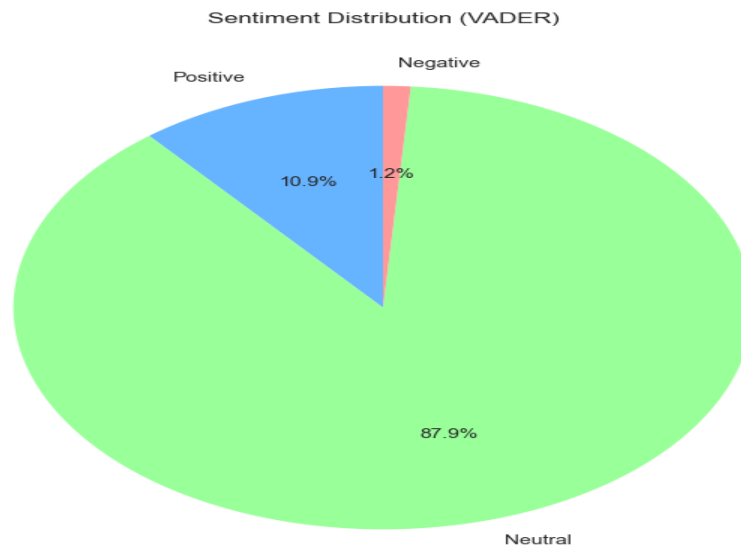


Sentence Length Distribution

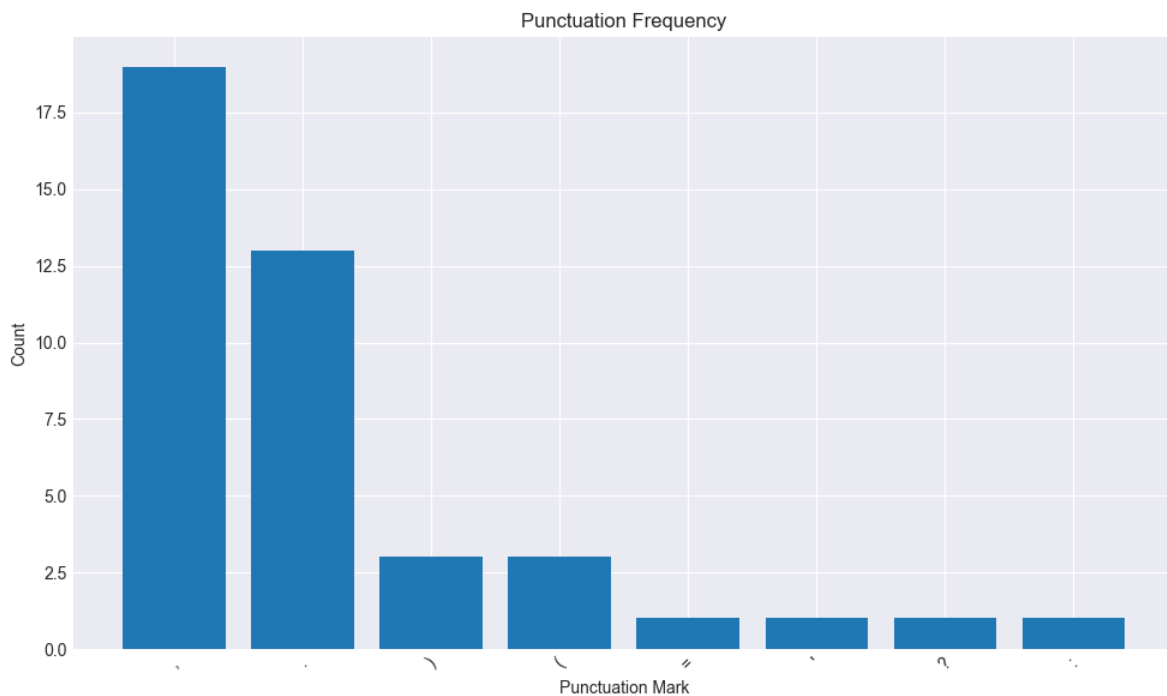


Zipf's Law Analysis (Word Frequency vs Rank)





Punctuation Frequency



5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

the (20), of (19), relativity (13), and (11), a (7), in (7), general (6), as (6), theory (5), that (5), physics (4), by (4), special (4), time (4), gravity (4)

Most Common Punctuation:

',' (19), ':' (13), '(' (3), ')' (3), ':' (1), '=' (1), '"' (1), '?' (1)

6. N-gram Analysis

Common 2-grams (Top 10):

'general relativity' (6), 'special relativity' (4), 'theory of' (3), 'of relativity' (3), 'such as' (3), 'time dilation' (3), 'the theory' (2), 'and time' (2), 'and the' (2), 'mass and' (2)
2-gram Repetition Rate: 0.0774

Common 3-grams (Top 10):

'theory of relativity' (3), 'the theory of' (2), 'mass and energy' (2), 'of relativity the' (1), 'relativity the theory' (1), 'of relativity is' (1), 'relativity is a' (1), 'is a fundamental' (1), 'a fundamental framework' (1), 'fundamental framework in' (1)
3-gram Repetition Rate: 0.0119

Common 4-grams (Top 10):

'the theory of relativity' (2), 'theory of relativity the' (1), 'of relativity the theory' (1), 'relativity the theory of' (1), 'theory of relativity is' (1), 'of relativity is a' (1), 'relativity is a fundamental' (1), 'is a fundamental framework' (1), 'a fundamental framework in' (1), 'fundamental framework in modern' (1)
4-gram Repetition Rate: 0.0030

Character N-gram Analysis:

Common Character 3-grams (Top 10):

'_th' (33), 'the' (30), 'ati' (26), '_re' (21), 'vit' (21), 'al_' (21), 'he_' (20), '_of' (19), 'of_' (19), 'lat' (19)

Common Character 4-grams (Top 10):

'_the' (25), 'the_' (20), '_of_' (19), 'tion' (18), 'lati' (17), 'vity' (17), 'rela' (15), 'elat' (15), '_rel' (14), 'ing_' (14)

Common Character 5-grams (Top 10):

'_the_' (18), 'relat' (15), '_rela' (14), 'elati' (14), 'lativ' (13), 'ativi' (13), 'tivit' (13), 'ivity' (13), 'vity_' (13), 'ation' (12)