

Text Analysis Report: theory_of_relativity_deepseek

1. Basic Statistics

Metric	Value
Word Count	662
Sentence Count	28
Character Count	4361
Unique Word Count	305
Avg Word Length (chars)	5.44
Avg Sentence Length (words)	23.64
Avg Sentence Length (chars)	155.75
Sentence Length Std Dev (words)	10.10
Lexical Diversity (TTR)	0.4607
MATTR (Window=50)	0.7964

2. Entropy & Complexity

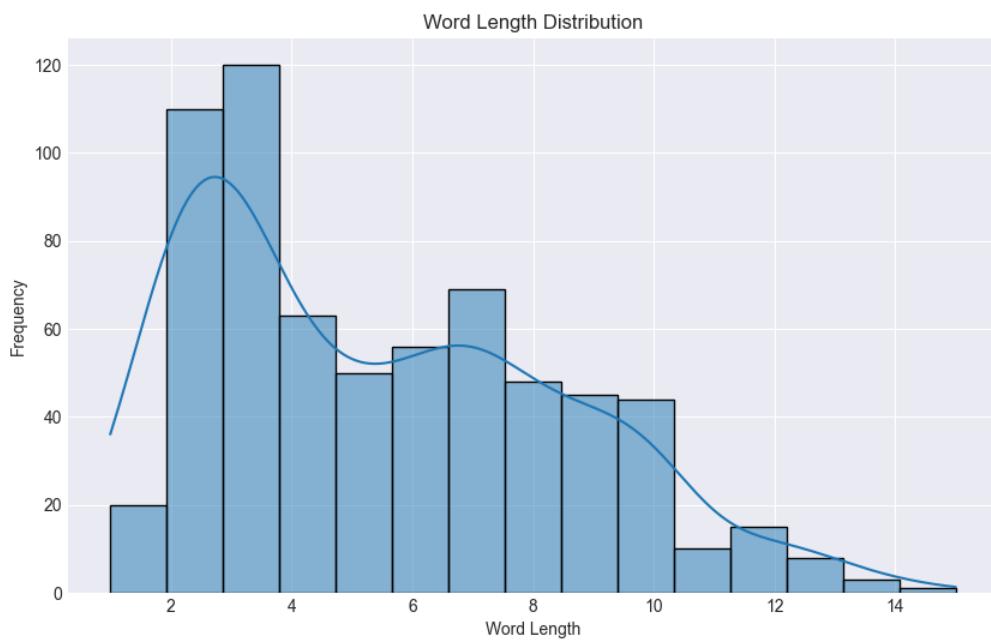
Metric	Value
Shannon Entropy (Words)	7.2859 bits
Shannon Entropy (Chars)	4.4526 bits
Function Word Ratio	0.3882

3. Readability Scores

Index	Score
Flesch Reading Ease	21.63
Flesch-Kincaid Grade Level	16.20
Gunning Fog Index	16.72
SMOG Index	17.50
Coleman-Liau Index	15.15
Automated Readability Index	17.40

4. Visualizations

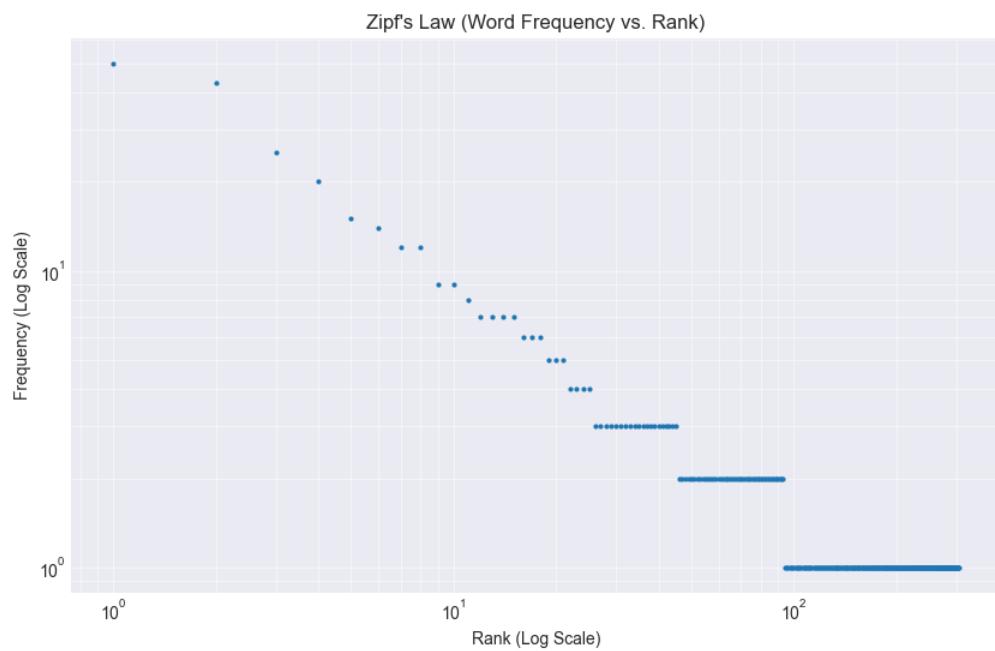
Word Length Distribution



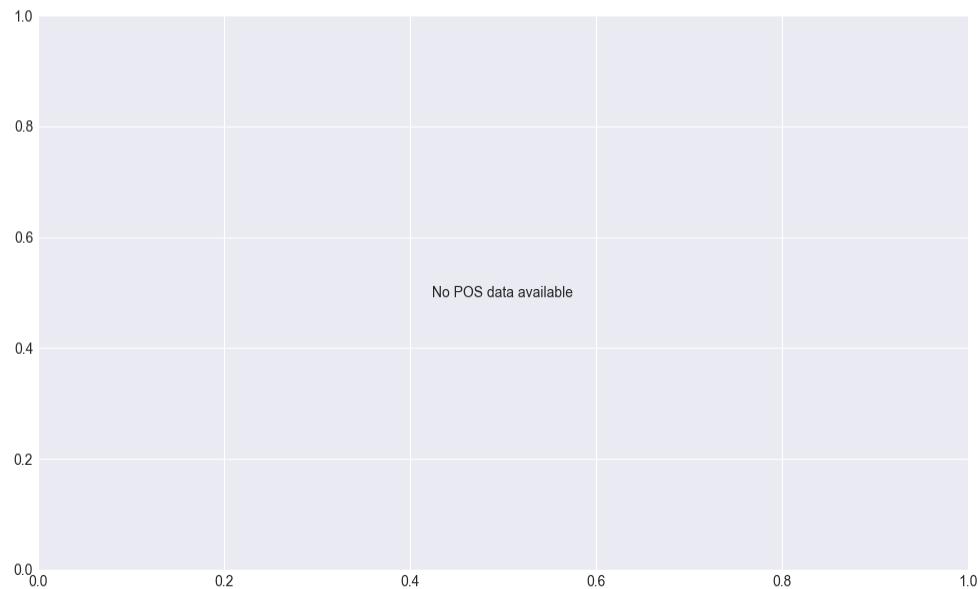
Sentence Length Distribution



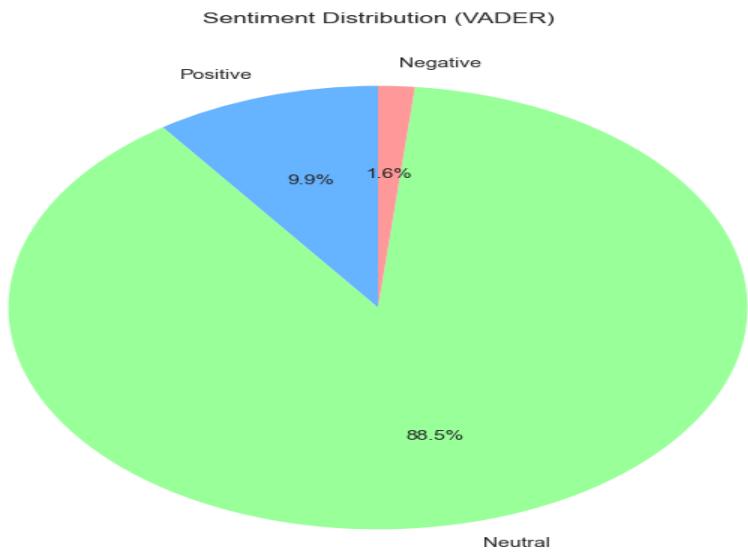
Zipf's Law Analysis (Word Frequency vs Rank)



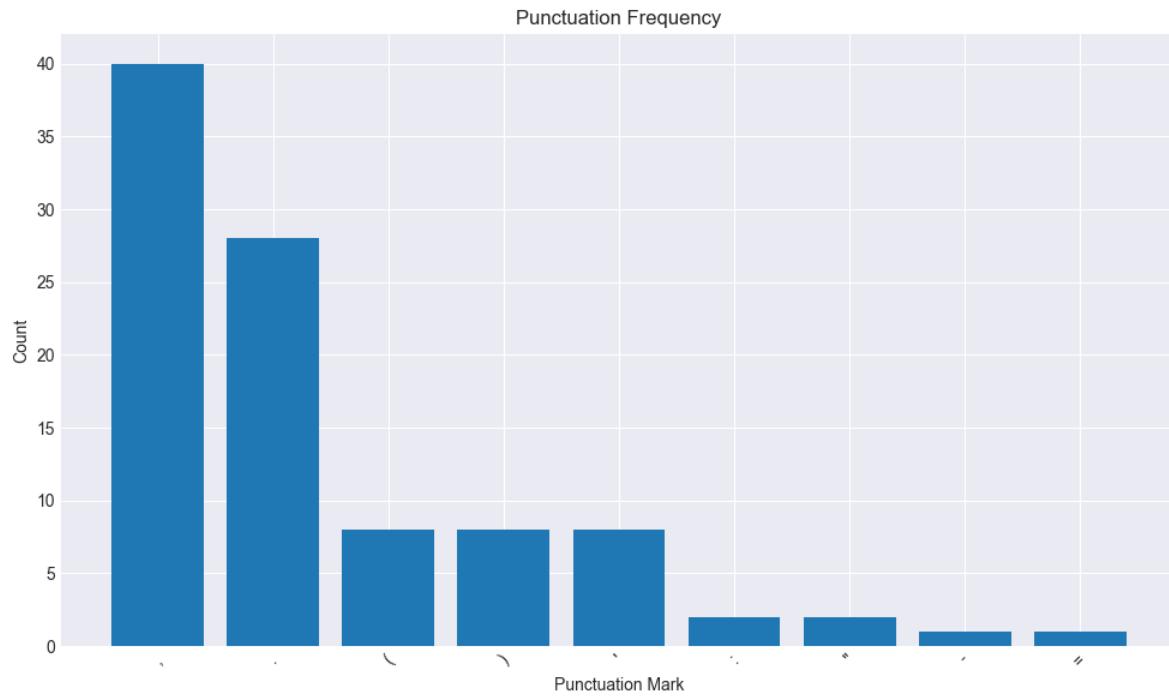
Part-of-Speech Distribution (Top 25)



Sentiment Distribution (VADER)



Punctuation Frequency



5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

the (50), of (43), and (25), relativity (20), in (15), as (14), theory (12), to (12), a (9), time (9), s (8), einstein (7), it (7), special (7), for (7)

Most Common Punctuation:

',' (40), '.' (28), '(' (8), ')' (8), '"' (8), ':' (2), """ (2), '=' (1), '-' (1)

6. N-gram Analysis

Common 2-grams (Top 10):

'theory of' (10), 'of relativity' (9), 'the theory' (7), 'and the' (7), 'of the' (7), 'special relativity' (6), 'such as' (6), 'general relativity' (5), 'einstein s' (5), 'as the' (4)

2-gram Repetition Rate: 0.1528

Common 3-grams (Top 10):

'theory of relativity' (9), 'the theory of' (6), 'such as the' (3), 'of black holes' (3), 'of relativity has' (2), 'relativity einstein s' (2), 'of relativity published' (2), 'relativity published in' (2), '1905 paper on' (2), 'the behavior of' (2)

3-gram Repetition Rate: 0.0500

Common 4-grams (Top 10):

'the theory of relativity' (6), 'theory of relativity has' (2), 'theory of relativity published' (2), 'of relativity published in' (2), 'the speed of light' (2), 'of mass and energy' (2), 'phenomena such as the' (2), 'theory of relativity was' (2), 'theory of relativity the' (1), 'of relativity the theory' (1)

4-gram Repetition Rate: 0.0182

Character N-gram Analysis:

Common Character 3-grams (Top 10):

'the' (69), '_th' (68), 'he_' (52), '_of' (43), 'of_' (43), 'ati' (38), 'lat' (31), 'ed_' (31), '_in' (30), 'nd_' (30)

Common Character 4-grams (Top 10):

'_the' (61), 'the_' (50), '_of_' (43), 'lati' (26), 'tion' (26), '_and' (25), 'and_' (25), 'rela' (24), 'elat' (24), 'vity' (24)

Common Character 5-grams (Top 10):

'_the_' (45), '_and_' (25), 'relat' (24), '_rela' (23), 'elati' (23), 'lativ' (23), 'ativi' (21), 'tivit' (20), 'ivity' (20), 'vity_' (20)