# **Wealth and Wisdom: Unveiling the Contrast between Richness and Science**

## Introduction

*The Science of Getting Rich* authored by W. D. Wattles is a timeless exploration of wealth creation. The title itself hints at a unique perspective—the "science" not in a traditional sense but as a specific mindset and set of actions. Wattles proposes that adhering to a particular way of thinking, known as the "Certain Way of Thinking" can pave the path to prosperity. Rooted in the principles of the New Thought Movement and mental healing practices (Wikipedia contributors, 2023), this book challenges conventional wisdom and offers readers a paradigm shift in their pursuit of wealth and success.

Humanity Research Questions

The purpose of this project is to conduct a textual analysis of *The Science of Getting Rich* by W. D. Wattles, available on Project Gutenberg. The analysis aims to explore how the words related to “richness” were employed by the author to inspire and motivate readers. Additionally, we will investigate the author's use of terms like "science" and "law" to establish credibility and present a rational and logical system for wealth creation.

Research Question 1: How does the author employ specific words such as "rich", "wealth", "money" and "success" in *The Science of Getting Rich* to inspire and motivate the reader?

Research Question 2: In what manner does the author use words like "science" and "law" to establish credibility and present his wealth creation method as a rational and logical system?

Data and Methods

Data Retrieval

The primary source for this analysis is "The Science of Getting Rich" by Wallace D. Wattles which is available on Project Gutenberg, and the URL for retrieval is provided. Its cover and table of contents were removed in order to be focusing on the core content for analysis.

Text Preprocessing

To prepare the text for analysis, various preprocessing techniques were applied. Regular expressions were used to eliminate unnecessary line breaks and special characters. Additionally, the text was converted to lowercase for uniformity in word analysis. To enhance analysis accuracy, common English stop words and specific additional words, such as "project" and "Gutenberg," were deemed irrelevant to the study and were removed. Text processing involves tokenisation and other necessary adjustments, lays the foundation for a comprehensive and detailed exploration of the text's linguistic and thematic elements.

Word Frequency Analysis

The primary objective of the word frequency analysis is to explore the prominence of the terms about "richness". Using the Natural Language Toolkit (NLTK), the text is tokenised to individual words, and their frequency distribution is calculated. This focused analysis seeks to unveil the prevalence and significance of the term "rich" and its variants, providing insights into its dominance within the textual landscape of Wattles' work.

Sentiment Analysis with TF-IDF

To understand the sentiment conveyed in the text, sentiment analysis on sentences was performed using the TF-IDF (Term Frequency-Inverse Document Frequency) technique. The TfidfVectorizer from the scikit-learn library was used to convert the sentences into TF-IDF vectors. Simultaneously, the VADER sentiment intensity analyzer was utilized to assign sentiment scores to each sentence. This dual approach provided a nuanced understanding of how sentiments are distributed throughout the text.

Chi-squared Test for Independence

Additionally, a chi-squared test for independence examined the relationship between sentiment (negative or positive) and the presence of terms related to richness. Sentiment labels were mapped to binary values, and a contingency table was formed. The chi-squared test, conducted with appropriate statistical functions, provided results, including the chi-square statistic and p-value.Top of Form

Topic Modeling

Prior to constructing the topic modeling, sentences containing key terms such as 'science', 'law' and 'principle' were filtered to create a subset of relevant sentences. After tokenisation (removing common stop words and punctuation), a document-terminology matrix was constructed for Latent Dreichler Assignment (LDA) analysis. LDA is a probabilistic topic modeling technique used to discover 20 latent topics in the filtered sentences. Each topic has a set of popular keywords and the weight of each topic represents its significant contribution to the overall content. The resulting themes and their respective weights provide valuable insights into the thematic composition and focus of the text, revealing the author's treatment of key concepts in the content provided.

Findings

In the analyzed text, consisting of a total of 11,672 filtered words, terms associated with richness, including "rich," "wealth," "money," and others, were identified 223 times. This yields a frequency of approximately 0.019, suggesting that roughly 2 words related to wealth and affluence appear for every 100 words read. Hassan and Barber (2021), it is statistically trustworthy that the more times information is repeated, even when its veracity is unknown, the higher the level of trust in the information. This observation provides insight into the recurring thematic elements centered around prosperity and success within the content of the text.

In the text, the prominence of the term "rich" is notably pronounced, surpassing other significant words (Figure 1) and ranking at the top 1 frequent word as it recurrently appears a substantial 166 times, highlighting its central role in conveying key concepts and themes within the content. Figure 2 provides a captivating visual representation through a word cloud, showcasing the top 50 words where 'Rich' stands prominently as the largest, emphasizing its significance within the text.

A graph of blue bars

Description automatically generated

Figure 1: Word Frequency Comparison – Demonstrating the Dominance of 'Rich' in the Text Over Other Significant Terms

A close up of words

Description automatically generated

Figure 2: Wordcloud Visualization – Top 50 Words

Upon subjecting the 976 sentences from the article to a comprehensive sentiment analysis, TF-IDF vectorization processing and VADER sentiment intensity analyzer were employed to assess the sentiment of each sentence, considering both the vectorised content and inherent sentiment features (Figure 3). The outcomes of this analysis revealed that a majority of the sentences, precisely 807 out of 976, were classified as having a positive sentiment. Notably, 190 of these positive sentences contained words associated with richness, such as "rich," "wealth," and "success." This dual classification provides a nuanced understanding of the prevailing positive sentiment throughout the article, with a specific focus on the instances where themes of affluence and prosperity are explicitly mentioned.

A screenshot of a computer code

Description automatically generated

Figure 3: Sentiment Analysis Results and Classification Table Preview –Categorizing Each Sentence as Positive or Negative, with Additional Markings for Sentences Containing Words Related to Rich.

The Chi-squared test of independence was conducted to assess the relationship between two categorical variables: sentence sentiment (categorized as either negative or positive) and the presence of words related to richness (binary label for 1 and 0 as containing rich keywords or not). A contingency table (Figure 4) generated to form the aasis for the Chi-squared test of independence. The calculated Chi-square statistic, equal to 7.5117, signifies the overall test statistic, indicating the extent of deviation from independence. Simultaneously, the associated p-value, notably low at 0.0061, suggests that the observed association between sentiment and the presence of richness-related words is statistically significant. In practical terms, the results imply that there is a discernible relationship between the sentiment expressed in sentences and the inclusion of terms related to richness, reinforcing the notion that these variables are not independent in the analyzed text.

A close-up of a number

Description automatically generated

Figure 4: Contingency Table – Illustrating the Frequency Distribution of Sentiment and the Presence of Words Related to Richness

The author strategically employs a high frequency of terms associated with richness and maintains a predominantly positive tone throughout the text to effectively motivate the reader. The recurrent use of richness-related words underscores key motivational themes centered around affluence and prosperity. This intentional choice of language aims to inspire and uplift the reader by emphasizing the positive outcomes and rewards linked to the principles and practices outlined in the text. The cohesive blend of high-frequency richness and positivity serves as a motivational tool, encouraging the reader to adopt the prescribed mindset and actions for personal success and financial well-being.

The results of the topic modeling (Figure 5) shed light on how the author strategically employs words like "science," "law," and "principle" to establish credibility and present a rational and logical system for wealth creation. Topic 3, for instance, emphasizes the interplay between "science," "getting rich," and "law," suggesting an association between scientific principles and legal frameworks in the wealth creation method. Additionally, Topic 6 brings attention to the connection between "work," "things," and "force," reinforcing the systematic and principled approach advocated by the author. The prominence of these terms across multiple topics underscores their significance in conveying a methodical and rational foundation for the wealth creation process. The author's intentional use of language, as revealed through topic modeling, supports the notion that the proposed approach is rooted in a well-structured and logical system, reinforcing the credibility of the outlined wealth creation methodology. LDA introduces randomness in topic modeling, impacting result stability. Topic pruning, one of the post-processing techniques is crucial to refine and interpret LDA outcomes, mitigating the effects of randomness (Jelodar et al., 2018).

Each topic exhibits a unique combination of keywords, providing a snapshot of the prevalent themes within the content. However, the inherent randomness of LDA may introduce variations in topic distribution. To further enhance understanding, PyLDAVis, an interactive visualization tool, proves instrumental. The visualization presents the topics as circles, with size indicating prevalence, and their proximity reflecting similarity (Figure 6). As an interactive tool, the actual results cannot be shown in a screenshot, it can be accessed on the Implementation.ipynb by executing the code. Word clouds showcase key terms, aiding in topic interpretation. The interactivity allows users to explore relationships between topics, evaluate coherence, and identify relevant documents. By hovering over circles or words, users can access details such as top documents and coherence scores.

A close-up of a text

Description automatically generated

Figure 5: Results of Topic Modeling

A screenshot of a graph

Description automatically generated

Figure 6: PyLDAvis Visualization - an interactive exploration of the topics generated by the Latent Dirichlet Allocation (LDA) algorithm.

Conclusion

Wattles strategically employs language, sentiment, and thematic elements to create a motivational narrative focused on wealth creation. The deliberate use of positive language, the statistical validation of the relationship between sentiment and richness-related words, and the nuanced exploration of key concepts contribute to the overall effectiveness of the text in inspiring and motivating readers. The emphasis on the term "science" suggests a rational and logical foundation for the proposed wealth creation method, reinforcing the author's unique perspective on the science of getting rich.

Discussion

While the analysis provides insights, there are certain limitations and areas of ambiguity to consider. The use of words related to richness, such as "rich," "wealth," and "success," may have inherent ambiguity and subjectivity. The project's reliance on common synonyms might introduce bias, as the interpretation of these terms can vary among readers.

Moreover, the use of pre-trained models like TF-IDF and VADER, without detailed evaluation and parameterization, poses a risk of model overfitting. Fine-tuning these models to the specific context of Wattles' text could enhance the accuracy of sentiment analysis.

The stochastic nature of Latent Dirichlet Allocation (LDA) topic modeling and the data preprocessing steps may lead to under-representation of results. Sensitivity to the initial state and potential variations in tokenisation could affect the reliability of the topics extracted. A more robust approach might involve exploring different parameter settings and conducting multiple runs to ensure the stability of the results.

In future analyses, addressing these limitations and considering alternative methodologies could further enhance the depth and accuracy of the exploration. Despite these considerations, the present analysis provides a foundation for understanding how Wattles strategically combines language, sentiment, and thematic elements to inspire readers on the path to wealth creation.

Reference

Doll, T. (2019, August 20). *LDA Topic Modeling: An Explanation*. ThirdEye Data. https://thirdeyedata.ai/lda-topic-modeling-an-explanation/

Hassan, A., & Barber, S. J. (2021). The effects of repetition frequency on the illusory truth effect. *Cognitive Research: Principles and Implications*, *6*(1). <https://doi.org/10.1186/s41235-021-00301-5>

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2018). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Wikipedia contributors. (2023, December 17). *The Science of Getting Rich*. Wikipedia. https://en.wikipedia.org/wiki/The\_Science\_of\_Getting\_Rich