

Billy Pilgrim was unstuck in time, and Kurt Vonnegut was ahead of his: How Combining Sentiment and Readerly Analysis is the Future of Literary Studies

By Madeline Vonk with help from Professors Elkins and Chun
Kenyon College, IPHS 200, Fall 2021

Introduction

In Kurt Vonnegut's lecture, "The Shape of Stories," he jokes, "I have tried to bring scientific thinking into literary criticism, and there's been very little gratitude for this."⁷ The same grudge between stem scholars and literary critics exists today, thus making the emerging field of digital humanities immensely controversial. Sentiment analysis is one of the main analytic tools used by digital humanities scholars, especially within the field of computational literary studies. This field can be defined as "the statistical representation of patterns discovered in a text mining fitted to currently existing knowledge about literature, literary history, and textual production."³ Sentiment analysis is a Natural Language Processing (NLP) tool that quantifies the emotions or affect of the words used in a text. Sentiment analysis can be used to perform several analytic tasks. It can be used to classify texts by genre or based on whether the ending is happy or sad. The tool can also be used to analyze characters and their interactions with one another. In this project, sentiment analysis will be used to identify the emotional peaks and valleys of *Slaughterhouse-Five's* plot to answer questions about how the novel's themes affect its structure.

In an early analysis of Kurt Vonnegut's tour de force, *Slaughterhouse-Five*, literary critic Wayne D. McGinnis analyzes how the themes of time, death, and renewal manifest in the novel's plot.⁴ McGinnis investigates the non-linear retelling of the story of the novel's protagonist, Billy Pilgrim. He argues, "The form created in the novel is essentially circular... *Slaughterhouse-Five* is a novel without climaxes, since its real subject matter and formal arrangement is renewal... [It is] a novel without beginning, middle, and end, without suspense and without a moral."⁴ However, I hypothesize that when graphed, *Slaughterhouse-Five* will have peaks and valleys just like any novel. Even so, the peaks and valleys may still reflect the cyclical theme of renewal as well as the other themes McGinnis identifies. Additionally, the use of sentiment analysis may help identify which aspects of the novel are real and which are manifestations of Billy Pilgrim's PTSD from fighting in World War II.

Why sentiment analysis?

In "The Shape of Stories," Vonnegut said, "The truth is, we know so little about life, we don't really know what the good news is and what the bad news is."⁵ By this, Vonnegut means that when graphed, the best stories are a plotless horizontal line because how could we, humans, identify the peaks and valleys, judge the good and the bad? But now, we have technology that can do just that. Today, this technology is used to identify emotional valence in novels that have been traditionally characterized as "plotless." The study "Can Sentiment Analysis Reveal Structure in a 'Plotless' Novel?" used various methods of sentiment analysis combined with readerly analysis to conduct a middle reading of Virginia Woolf's *To the Lighthouse*. The research showed that different models of sentiment analysis should not only be compared with each other to see if they agree but with readerly or human analysis as well. The paper found that comparisons between human and computational analysis usually agreed and "revealed thematic patterns that [were] significant to the novel."²

It is important to note that sentiment analysis can only reveal the emotional valence of a novel over time and not its exact plot; however, the two are closely correlated. In *Poetics*, Aristotle argues that a character must experience a discovery and a reversal that causes them to change their actions based on that discovery, and that reversal will cause the audience to experience catharsis or an emotional release. Sentiment analysis tracks the novel's emotional valence not the reader's, but as Aristotle states, the two go hand in hand. Because of that, I believe that analyzing *Slaughterhouse-Five* via sentiment analysis will enable me to see whether McGinnis is correct when he describes the novel's plot as cyclical. I expect my results to be similar to how the sentiment analysis of *To the Lighthouse* revealed that while many modernist writers considered their works to be "plotless," the text was still driven by an emotional arc.

Methodology

I started my project by finding a text file of Vonnegut's *Slaughterhouse-Five* online. I then uploaded that text file to a Google Colaboratory (Colab) notebook assembled by my professor, Jon Chun. Colab notebooks allow users to write and execute Python code from their web browser.

Before I could begin the process of sentiment analysis, I had to clean the text file. I began the cleaning process by stripping the text file with the help of the Python package Textthero. Textthero can be used to eliminate capitalization and remove digits alongside stop words and short sentences. This cuts down the number of sentences in the novel so that the sentiment analysis can focus on the more meaningful sentences. Here is an example of the shortest sentences before and after the cleaning process:

```
# View the shortest sentences before and after cleaning
```

```
novel_df.sort_values(by=['text_raw_len']).head(400)
```

	text_raw	text_clean	text_raw_len
765	In.		3
3374	No.		3
667	No.		3
3677	Oh.	oh	3
669	Oh.	oh	3
...
1908	The sun was high.	sun high	17
2724	'Many years ago.'	many years ago	17
2602	The sun was high.	sun high	17
2480	Barbara inquired.	barbara inquired	17
3670	'Where were you?'		17

400 rows x 3 columns

Once the text was clean, I could officially begin the sentiment analysis process. I executed this using the Python packages VADER and TextBlob. These tools score sentence inputs based on positivity or negativity with a numerical output. This is how VADER scores one of the most iconic lines from the novel:

```
vader_sa.polarity_scores('Everything was beautiful and nothing hurt.')
```

```
{'compound': 0.7701, 'neg': 0.0, 'neu': 0.375, 'pos': 0.625}
```

This is what the same sentence looks like when analyzed by TextBlob:

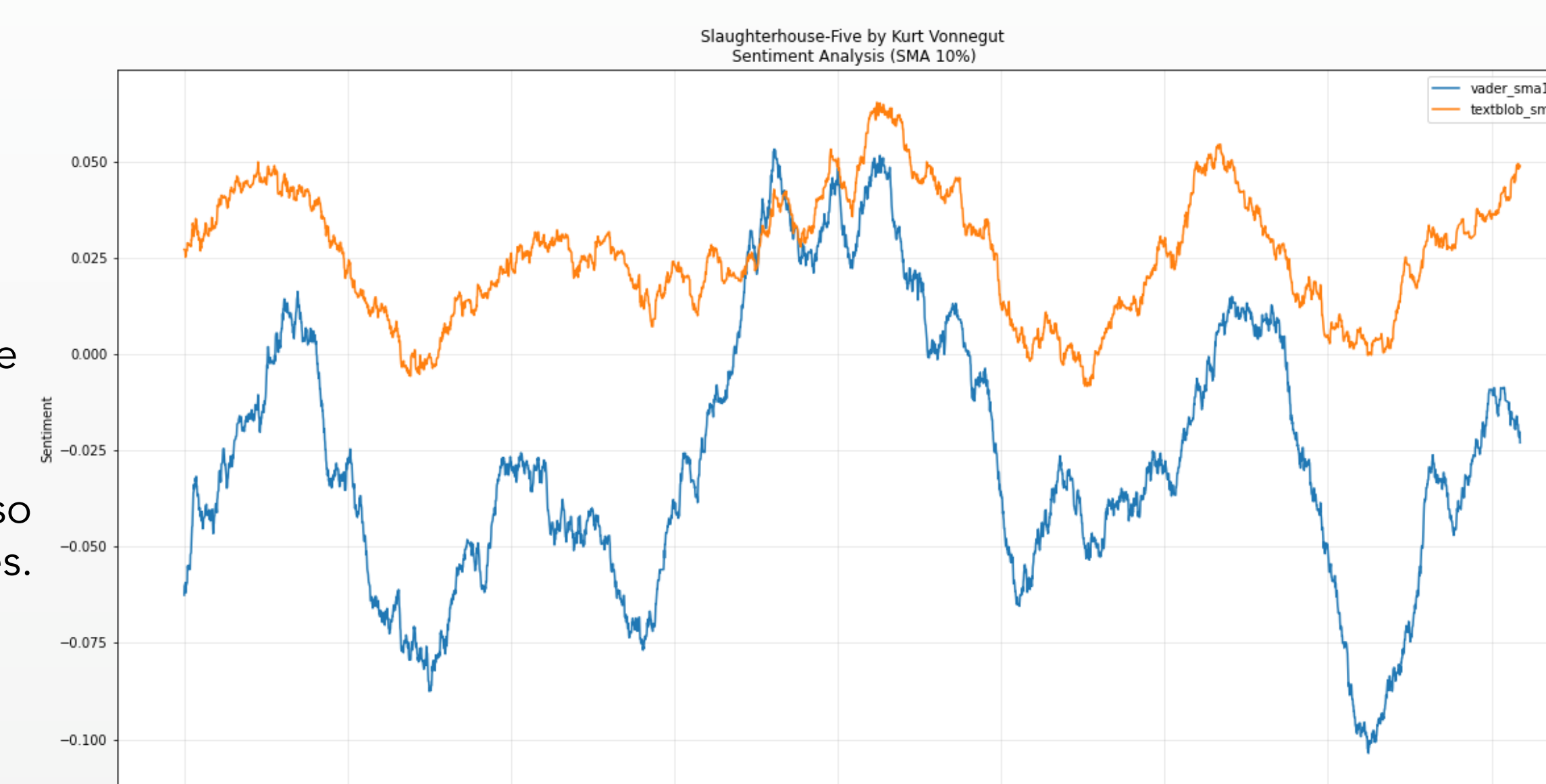
```
testimonial = TextBlob("Everything was beautiful and nothing hurt.")  
print(testimonial.sentiment.polarity)
```

```
0.85
```

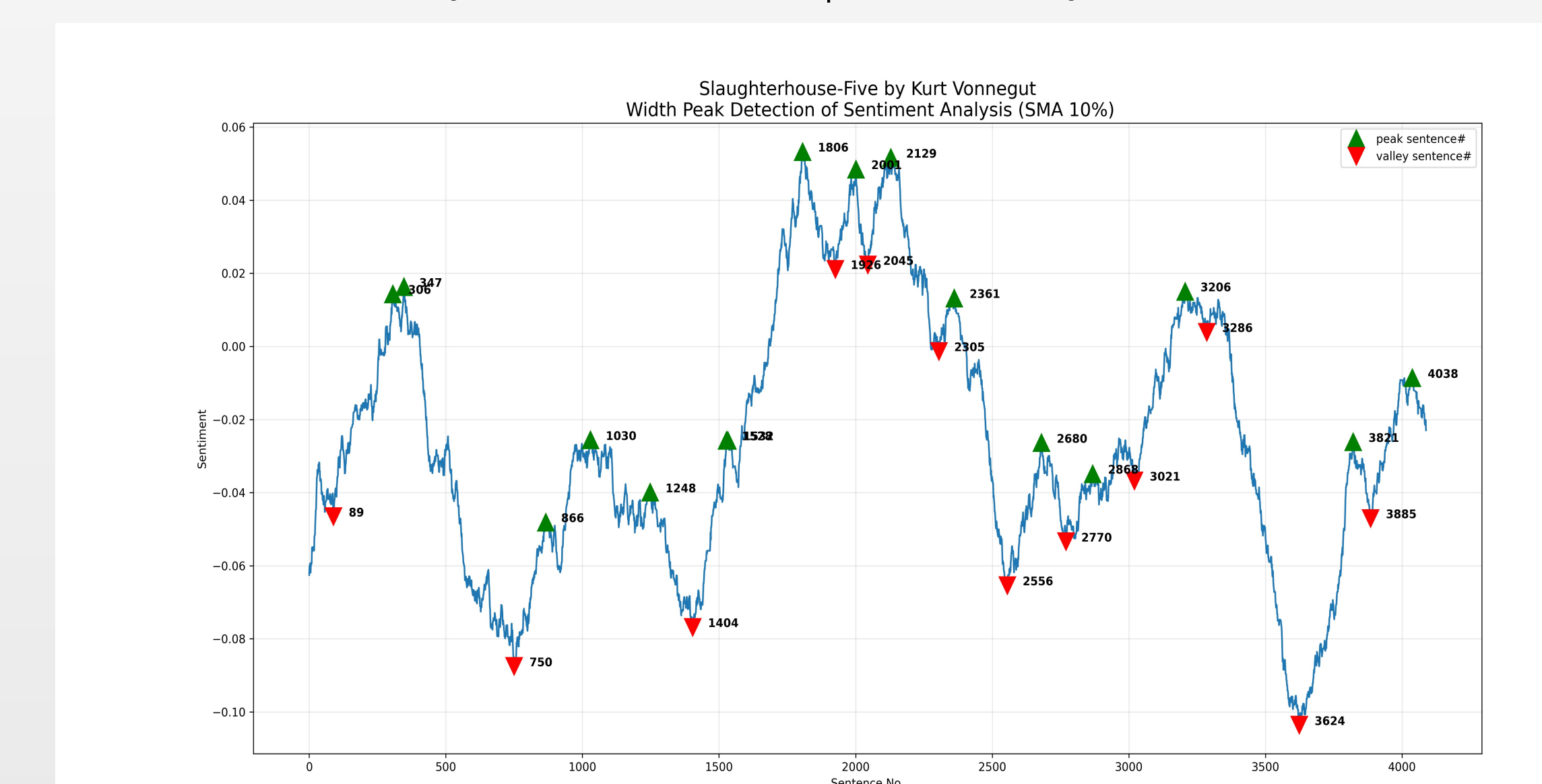
VADER and TextBlob both perform sentiment analysis, but they are programmed with different rules that weigh the input strings differently. Nevertheless, both outputs demonstrate that the sentence has an overall positive sentiment.

Finally, I ran the entire clean text through both programs and received a graph juxtaposing the two sentiment arcs. The graph shows that VADER is better at detecting more distinct peaks and valleys. VADER is also one of the most popular sentiment analysis tools and is highly rated on both github.com and the Natural Language Toolkit (NLTK).¹ For these reasons, I chose to continue my analysis with the sentiment analysis produced by the VADER graph. The VADER line plot was then analyzed with a peak finding algorithm from SciPy. I had to play with the parameters until I found a graph without an overwhelming range of peaks and valleys. From here, I could identify the sentences that the algorithm marked as the peaks and valleys, analyze them, and ascertain why they were chosen. The graphs enabled me to track the progress of the novel's plot as a whole to see if the plot truly is cyclical, as McGinnis argues.

Sentiment Analysis Graphs



Sentiment Analysis of the novel as performed by VADER and TextBlob



Cruxes identified on the VADER line graph

Results

Looking at the graphs above, it is fascinating to discover that when graphed, the sentiment driving *Slaughterhouse-Five's* plot follows a relatively cyclical pattern. The cyclical nature of the novel is especially highlighted by the TextBlob analysis. Even though the VADER graph has more disparities between its peaks and valleys, its structure is also relatively cyclical. The central peaks of both line graphs are book-ended by two smaller peaks, and they begin and end at similar points. The end of the novel is a little more positive than the beginning; however, this is to be expected since the problems identified at the start of a novel are more likely to be resolved by the end. This proves my hypothesis that peaks and valleys would manifest in the sentiment analysis graph. The cruxes demonstrate that McGinnis's claim that the novel's plot is cyclical is relatively correct; however, he also stated that the novel does not have climaxes. Looking at the VADER graph above, it is undeniable that the novel is driven by emotional climaxes.

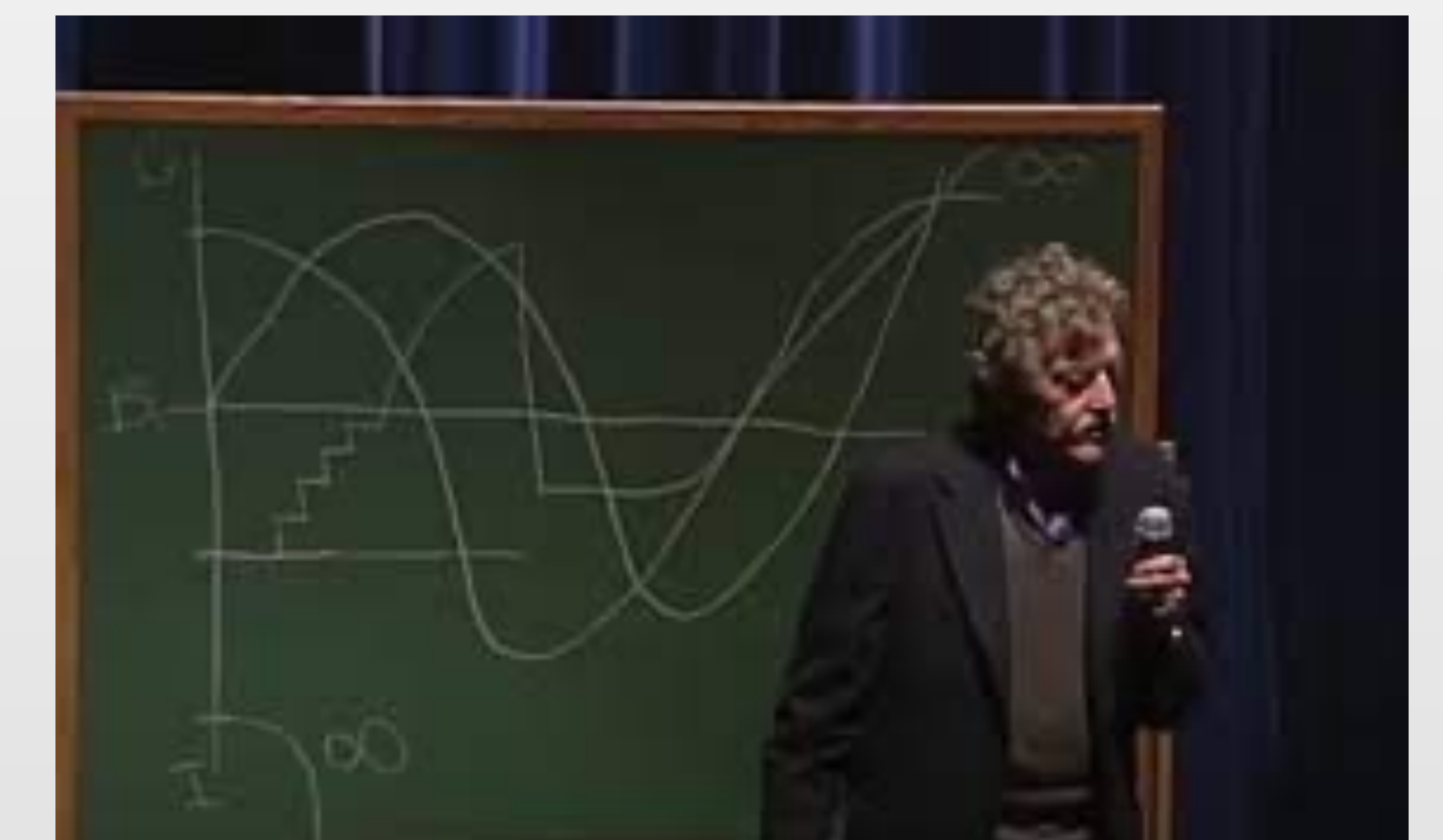
Additionally, each identified crux is from a scene before, during, and after Billy Pilgrim's time fighting in World War II. The novel begins with the line, "All of this happened, more or less," and analyzing the cruxes allows us to identify what actually happened.⁶ None of the cruxes take place on Tralfamador, the home planet of the aliens who abduct Billy. This supports the argument that Tralfamador is a product of Billy's PTSD and not a real aspect of the storyline.

Lastly, I feel it is important to note that because Vonnegut includes pictures in his novel, sentiment analysis cannot record the irony created by these images. For example, VADER and TextBlob gave "Everything was beautiful and nothing hurt" a positive value, but in the novel, this sentence is written on a gravestone and therefore, is deeply ironic and not as uplifting as one might believe.

Conclusion

I hope that this project highlights the pros of combining literary criticism with computational literary studies. The fact that sentiment analysis produced relatively cyclical line graphs reaffirms what the paper "Can Sentiment Analysis Reveal Structure in a 'Plotless' Novel" concludes about how human interpretations often agree with the computational analysis.² If McGinnis had access to sentiment analysis technology when he wrote his article, he could have included these graphs alongside his close readings to prove his thesis. Additionally, the cruxes answered my question of what aspects of the novel are real.

For further research, it would be helpful to produce more sentiment analysis models of *Slaughterhouse-Five*. These models could be averaged together to find a more accurate graph of the novel's sentiment arc. This would help ensure that none of the crux points identified in my graphs are outliers. In a study on sentiment arcs, the researchers averaged 34 models to find the best sentiment analysis plot, which is a much larger dataset than my analysis is based on.¹ Even though this project could be based on more data, it agrees with previous conclusions drawn by both literary critics as well as digital humanities scholars. It also shows that literary critics may be able to use sentiment analysis with novels like *Slaughterhouse-Five* where an element of magical realism obscures the reality of the text and the character's experiences. This project also supports what Vonnegut tried to do with "The Shapes of Stories." Scientific thinking and literary criticism should merge and are merging, and we should all give Kurt Vonnegut a little more gratitude for trying to accomplish that decades ago.



Vonnegut pictured with his plot graph from his lecture "The Shapes of Stories"

Sources and Acknowledgements

1. Chun, Jon. "SentimentArcs: A Novel Method for Self-Supervised Sentiment Analysis of Time Series Shows SOTA Transformers Can Struggle Finding Narrative Arcs." *ArXiv abs/2110.09454* (2021): n. pag.
2. Elkins, Kathrine and Jon Chun. "Can Sentiment Analysis Reveal Structure in a Plotless Novel?" *ArXiv abs/1910.01441* (2019): n. pag.
3. Kim, Evgeny and Roman Klinger. "A Survey on Sentiment and Emotion Analysis for Computational Literary Studies." *ArXiv abs/1808.03137* (2021): n. pag.
4. McGinnis, Wayne D. "The Arbitrary Cycle of *Slaughterhouse-Five*: A Relation of Form to Theme." *Critique* 17.1 (1975): 55. *ProQuest*. Web. 11 Dec. 2021.
5. Vonnegut, Kurt. "Here Is a Lesson in Creative Writing." *A Man without a Country*, edited by Daniel Simon, Seven Stories Press, New York, NY, 2005, pp. 23-37.
6. Vonnegut, Kurt. *Full Text of "Kurt Vonnegut Slaughterhouse Five Audiobook"*, https://archive.org/stream/KurtVonnegutSlaughterhouseFiveAudiobook/Slaughterhouse%20Five%20by%20Kurt%20Vonnegut_djvu.txt.
7. Vonnegut, Kurt. "Kurt Vonnegut, Shape of Stories." *Youtube.com*, 14 July 2018, https://www.youtube.com/watch?v=GOGru_4z1Vc.

A special thanks to Professor Chun and Professor Elkins for helping me with the code and concept for my project. An additional thanks to my uncle, Andrew Freed, who introduced me to Vonnegut and is always there to listen to me theorize about him.