

Guy de Maupassant Analysis

By: Oscar, Elijah, Jessica, Vikas, and Zexing

Discussion and Decision-Making

Guy de Maupassant is considered as the father of the modern short story form, whose work is characterized by economy of style. His work is well-known and is translated into different languages, influencing the literature across the world. Members of our groups all know and at least read one of his short stories. We are interested in why his work is favored by the readers and the common themes in his literature. Therefore, we all agree to take a closer look at short stories of Maupassant.

When deciding the text we want to analyze, we research about the work of Maupassant and choose eight of his representative works. The finest short story of Maupassant is probably *Boule De Suif*, while the *Diamond Necklace* is the most popular one. Besides, we choose *Found on a Drowned Man*, *Mademoiselle Fifi*, *Miss Harriet*, *My Uncle Jules*, *The Piece of String*, and *The Wreck*. We are interested in answering a few questions: What are the themes of the stories? Are there any trends in his story plots? How did his writing change overtime?

Summary and Contextualization

Below for each short story, is a quick summary of the tale, its interesting features along with the ending, all of which we used to complement and compare with our own analyses and generated story plots.

Boule De Suif - In this story, 10 travelers set forth from Rouen in a stagecoach to escape Prussian occupation, and the titular woman, “Boule de Suif”, is a prostitute in this carriage. This social setting proves to be a pessimistic microcosm of French society at the time, and the story concludes with Rousset (“Boule de Suif”) both starving and sobbing in the back of the carriage.

Found on a Drowned Man - A story without great deal of specifics about a man who has never been in love, but recounts a magical evening with a woman he had spent on a river. The story ends with the man seeing his companion as an embodiment of beauty, love, and life, and him trying to kiss her, to which she responds with the observation that he has a caterpillar in his hair. This is all within the frame story of a letter found on a man drowned in the Seine, investigated and shared for the purposes of identification

Mademoiselle Fifi - Another alike *Boule de Suif* in its exploration of themes of war (Franco-Prussian) and class interactions. Taking place at a manor house occupied by German officers, the titular character is an unpleasant man with an effeminate character (hence the epithet Mademoiselle). Fifi is stabbed by a French prostitute, who hides in a bell tower symbolic over her personal and national triumph over the Germans.

Miss Harriet - A less well-read story of Maupassant, in which a middle-aged woman in the countryside takes a fancy to a younger French adventurer who stays in her residence. Houses a religious and supernatural theme, and concludes with the woman's death.

The Diamond Necklace - An immensely famous story of Maupassant, in which Mathilde Loisel borrows and loses the titular piece of jewelry, keeping it a secret and having to commit to a life of poverty because of it.

The Piece of String - This story's setting takes place in the town of Goderville and concerns an animosity between Hauchecorne and Malandain. This underlies a controversy over Hauchecorne picking up a piece of string, which drives him to despair by the conclusion of the narrative.

The Wreck - Tale told within the frame story of Georges Garin, a mariner whose reception of a heavily stamped letter leads him to recount a tale of his youth as he was fixing a boat. Famous last line of Garin “How sad it all is!”.

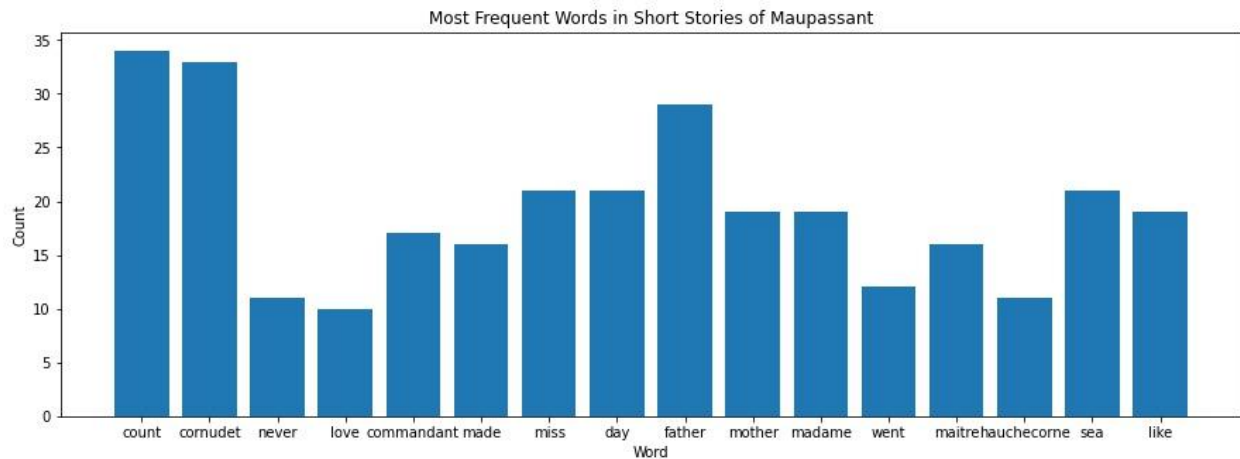
Rationale: We assembled this collection of short stories purposefully, choosing Maupassant most famous creations for two reasons. One, most of us were familiar with these works, and thus could judge with some degree whether our models were generating a representation of his work with identifiable features we could map between our personal understanding garnered from close readings, and the computational feature of the graphs we created. Secondly, we wanted to see any common themes or characteristic patterns of Maupassant which made these short stories his most well-promoted, outside of a contextual historical analysis on computational basis alone, a similar endeavor to Jockers associating the success of some stories with their rhythmic features.

Explanations and Visualizations

Most Frequent Words in Short Stories

Below is a bar graph of the top 2 frequent words in each of the eight short stories. Therefore, there are 16 graphs in total. We chose the top 2 frequent words rather than other numbers since one word might not be enough to reflect the theme. We noticed that words that appear the most times in text might overlap with the title of each story. Therefore, we filter out some particular names in the `special_name_list` during the text cleaning process and include more words in the bar graph to get a better sense of the content in text. We admit that including more words in this frequency distribution might present us with a more comprehensive view, but for the sake of clarity, we experiment with the top 2 frequent words for each text.

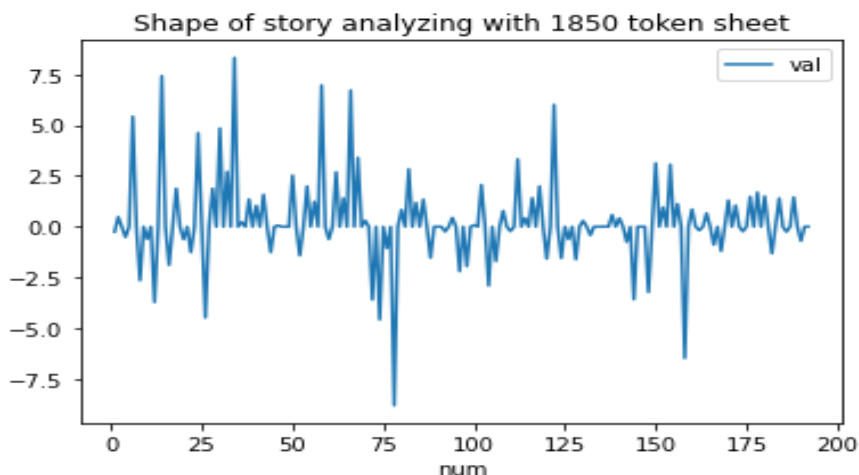
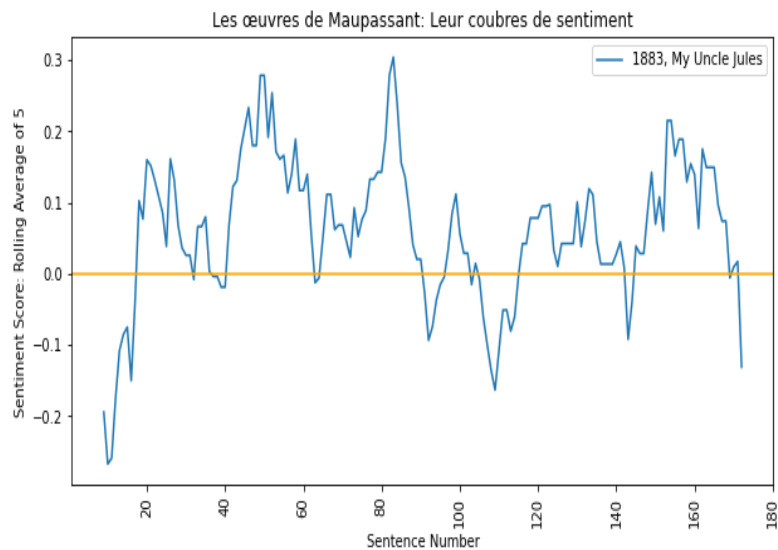
We can see that the words that repeat for a few times in text indeed indicate some overall themes. Except for the names of the characters, there are some words that indicate personal relationships or titles, such as father, mother, miss, madame, and maitre. This indicates the interpersonal dynamic within the text and reveals the potential social interaction between characters. Also, there are words related to emotions and values like love and like. Some terms that indicate strong emotional valence such as never. Time might also play a critical role as the recurrence of “day” emphasizes the fluidity of time.



Our group uses different methods when analyzing the 8 texts in terms of emotional valence (or, in other words, negativity or positivity). We use a machine learning model called 'VADAR' (Valence Aware Dictionary for Sentiment Reasoning) which is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. It was created based on modern-day Twitter posts. Similar to the 1850 valance sheet we did in assignments, Vadar sums up the sentiment score in each sentence. To be noted, it also understands basic sentence structure such as 'I did not like you' and classifies it into a negative score. As we could tell from vadar's results, it could not understand sarcastic languages in My Uncle Jules, and it might bring extreme results in each sentence and thus causing huge fluctuations in the sentiment-score graph. We thereby selecting rolling functions as a way to normalize our graphs and make it smoother. When it comes to the analysis based on 1850 token sheet, the result is a bit more extreme. Specifically, the score of adjacent sentences might indicate completely different emotions. There are two reasons for

it: first, the 1850 sheet analyzes the texts solely based on the score of each word within a sentence and it does not consider other situations such as negative sentences. Second, there are too many words not covered in the list, so each sentence's score might be solely calculated on a couple of words, which would bring bigger error in terms of total score. I also noticed that the French texts are translated into modern English, which means it has modern English words choices and sentence structure. Thereby using the 1850 sheet seems to be a bit inappropriate.

For example, the comparison between vader and token sheet on My Uncle Jules



Reflection

In the word frequency part, we filter out a few words in the title and words that carry no meaning to graph the top 2 frequent words. This analysis would be more comprehensive if we filter out the name of characters as well as include more words in our analysis. Through our research, we explore that the major themes of Maupassant's short stories are war, prostitution, and madness. Through counting the words, we can observe the social dynamic in his work and the element of femininity. Nevertheless, the theme of war and madness (possibly related to female characteristics in his work but is not clearly revealed in the word count) we could include more words in analysis that would possibly encompass more comprehensive elements.

I realized that there is a pretty interesting distinction between two sets of sentiment-score graphs(the vader and the 1850). The vader one has approximately three times more sentences than the ones in 1850. I am pretty sure that it is due to the differences in data cleaning and pre-processing, but I could not find it yet. In addition, I feel like our group's topic has some fundamental problems. The size of each text is small and conducting sentiment analysis might bring biased results when it comes to short-length texts, as it is more adept in analyzing lengthy texts. Plus, we are not able to use machine learning tools in this case. I feel like there are a couple of good models that are better at dealing with sentences and analyzing the texts in a multi-dimensional perspective. I would like to note in the end that as a group project, it is important for each of the group members to contribute at

least some effort. One of our group members neither contributed to the completion of our work nor came to the class for presentation, so I would not give credit to that member for this project and I feel like as a group we need to work cooperatively to get the work done.