

Project Overview: This project's goal is primarily to compare and contrast a corpus of texts by the authors Jane Austen and Charles Dickens. The data analysis will help recognize data among the selected group of texts. By doing so, we can gain knowledge of the surrounding theme, narratives, and use of language within the texts. To better understand what we will be analyzing, we will take a deeper look into the emotions and elements of the author's literary work.

Hypothesis/Introduction: We believe that Charles Dickens locations will correlate with more urban areas in England like London and will be more world traveled. We believe this is because he depicts the lower class with more male tones. While Jane Austen will correlate with rural areas in England and will be more condenced. We think this will show opposingly of Dickens because Jane Austen depicts the upper classes with more feminine themes in her novels.

After completing the code, it is clear that the majority of Charles Dickens' novels are located in Urban centers. The following five most popular cities are taken from the code: Birmingham, New York, Liverpool, Maypoole, and Canterbury. All of these areas that are used in Dickens' text are rated highly urban areas of various countries including the UK and US. The amount these cities were used in these texts are used 55-97 times throughout all of the novels which is a large amount. The cities most commonly use have high poverty rates especially during the times that Dickens wrote about them, which range from 17-52% poverty rates. These areas are densely populated and urban which proves how Dickens most commonly writes about urban centers to portray his take of the lower class and negative represented areas.

Jane Austen's locations seem to align with our hypotheses. It looks like they correlate with more rural areas. This makes sense because her characters in her novels often go to country estates and the English countryside. This reflects her focus on England's high society and middle-class families, rather than urban or cosmopolitan centers. The rural backdrop allows Austen to delve into the private lives and social circles of her characters in a more isolated setting, where reputations are emphasized in the context of limited geographical movement and societal structures. There are some urban cities that challenge our theory like Bath and London. This might show that Austen's works are not entirely insulated from urban areas, but rather selectively incorporate them to provide contrast. Overall her locations seemed to be more rural and less widespread than Dickens.

These divergences between the top 5 and bottom 5 locations by sentiment follow from the different ways these places are represented and spoken about in her text, reflecting both the emotions tied to specific geographic settings-such as London and other places in Europe-and the events or themes associated with those settings. Overall, places with higher sentiment scores, such as Dorset, Park, and Croydon, we believe may relate to more positive stories, experiences, or associations. Such places might be portrayed as tranquil, thriving, or culturally active, evoking

positive feelings from characters or the narrator. Positive scores may also indicate a location's appeal through natural beauty, a quiet atmosphere, and victories or positive events taking place there.

By contrast, the bottom 5, including Northamptonshire, Lambton, and Somerset, could indicate a number of negative events, such as conflict, hardship, or personal struggle. These locations may be represented in the novel as being inhospitable or may come under themes regarding, for instance, economic struggle or social and personal disasters. This postulates that the negative scores denote less favorable tone and context when talking about these places; thus, giving a negative perception of these places.

The overall difference in top and bottom locations is symbolic of how geography within the text may influence and mirror emotional dynamics. Places are not simply physical settings; they also can often be representative of the emotional states, challenges, or successes of a character. If a place does evoke a positive or negative feeling, it can be based upon the plot, the social or historical context of the location, or the specific experiences the location calls to mind. This juxtaposition reveals how the text differentiates between emotionally meaningful places, therefore building an emotional geography which influences the reader's perception of each single place.

Reflection:

Though this project was extremely difficult for me, I found the process of understanding and further exploring Jane Austen's literature and corpus to be very interesting. Further exploring the emotions that were depicted in Jane Austen's literature was very hard for me personally to understand the emotions of Jane Austen's. Something that I think we could have done differently for this project is that I might have experimented with multiple models or parameters to see if the results changed across regions with high and low scores. Additionally, I could have also had a deeper breakdown of positive, neutral, and negative emotions. This might have provided a richer context for understanding why certain regions were scored as they were, particularly if I examined specific passages or themes associated with those locations. Furthermore, filtering for more frequent mentions or applying weighted averages could have helped clarify the significance of some locations' scores, potentially leading to a more meaningful comparative analysis.

