Our investigation last quarter validated the hypothesis that people-of-color characters in young adult novels exhibit complex traumas by computationally identifying and visually representing the characters' emotional journeys. This quarter, we improved our analysis by addressing two technical challenges.

Primarily, we improved the accuracy of our name-mention extraction function to better track characters and their emotional journeys throughout novels, using dependency parsers that would extract only the sentence subjects. We strengthened our models with coreference resolution tasks that found all expressions that refer to the same person in a text, addressing the fragmentation in our time-series caused by multiple names characters have.

Character-level Visualizations - Styx from *The Season of Styx Malone*

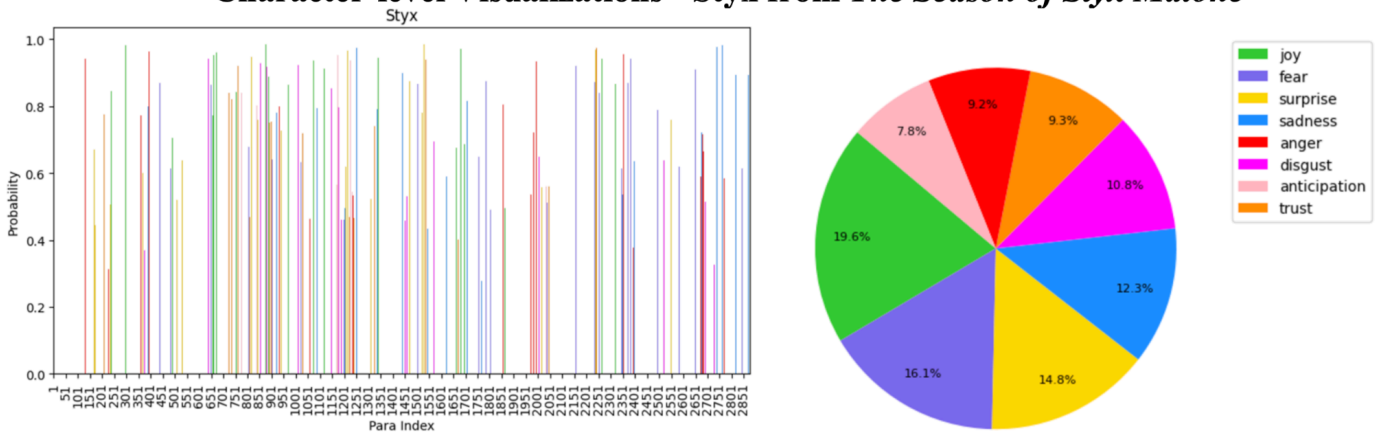

Figure 1. The diverging bar graph shows the change of emotion over time for the novel *The Season of Styx Malone* by Kekla Magoon, using our 8-class emotion classification model. It shows paragraph index along the X axis and the probability of an individual paragraph being an emotion, as determined by the legend. Here, we can see that the character Styx from the novel experiences a wide range of emotions, with particularly noticeable clusters of joy, fear, and surprise. The pie chart shows that these three emotions are in fact the ones with highest frequency compared to the other emotions.

You can see that the emotions in the bar graph are very sparse. This is since the dependency parsers tagged only the characters who are the protagonists in a paragraph, eliminating appearances of false positive name tagging. On average, only 38% of the characters from the overall character list appear in our final analysis.

Secondly, we strengthened the character-emotion mapping. To track characters' emotional journeys, we map paragraphs' emotion classifications to the related characters (Figure-1). We found that not all mentioned characters drive the expressed emotions; some are merely referenced in sentences. To reduce these false positive names, from our dependency parser, we filtered only nominal subjects and objects of 'by' preposition from active and passive voice sentences respectively.

As depicted in Figure-2, we also tried to expand our analyses to different gender groups, however due to constraints imposed by the books we examined, our external mentor opted to pause this exploration.

Novel-level Visualization -
Gender in *Home Is Not a Country* by Safia Elhillo

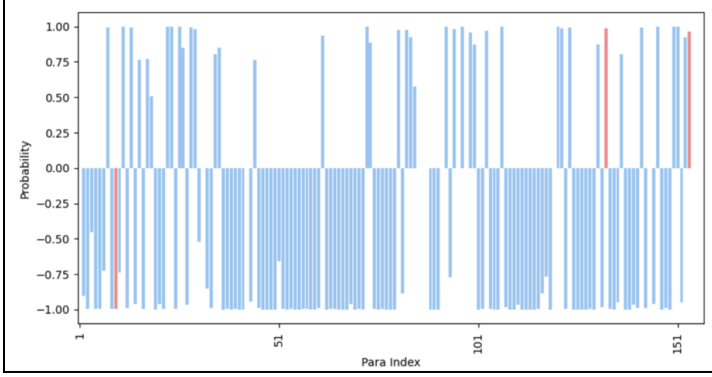

Figure 2. This bar chart considers our 3 class model predictions based on positive, negative, and neutral probabilities. Here, blue represents the emotions of male characters, and red represents the emotions of female characters. Unfortunately, there is too little information (i.e., of our 26 books, only a handful of them had a reasonable ratio of male to female characters deemed as "main" characters) to gauge whether the relationship between female and male characters' emotions is of high significance or not.