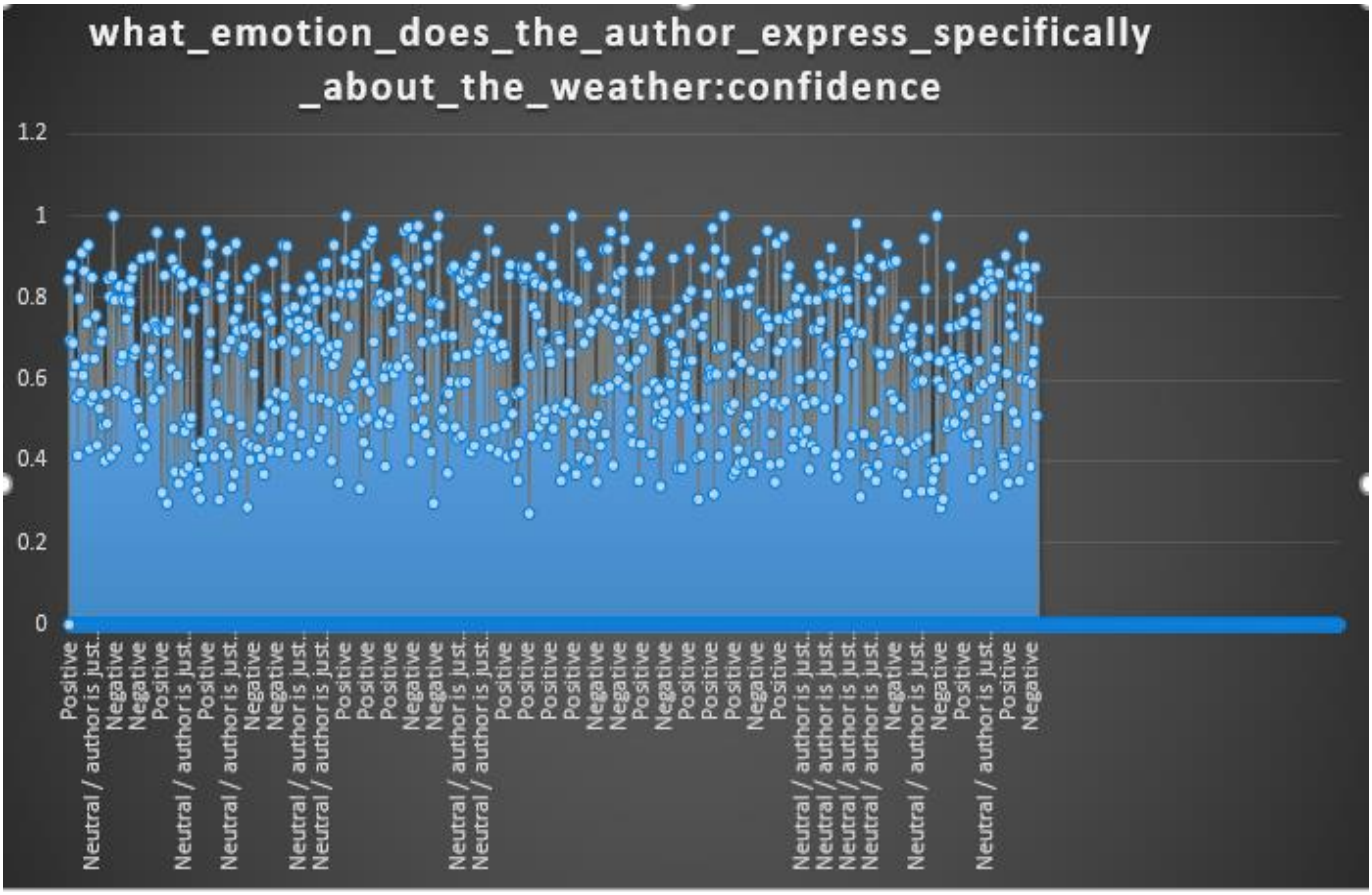
Mounika Reddy
A02236773



Data set and Cleaning: I have taken a data set on weather and cleaned it using ***Open Refine***. I cleaned all the rows where the response is not clear and where the tweet is not related to the weather.

Figure: Now that the data is cleaned, I chose two columns named "What_emotion_does_the_author_express_specifically_about_the_weather" and "what_emotion_does_the_author_express_specifically_about_the_weather: confidence" and drew a ***2D-Area Graph*** for these columns.

Representation: I took the author's emotion about the weather on X-Axis and the confidence value of it on Y-Axis which resulted in this figure.

Justification: I chose 2D Area figure because this is one of the techniques which is highly used for high dimensional data visualization. By looking at this figure, we can easily visualize authors responses and their corresponding confidence values. I have taken a black background on which blue is representing the area of authors responses so that it would be both appealing and easy to understand. The nodes in the figure represent the confidence value for each response. Range of the confidence value is [0 1].

Drawbacks: Although, 2D Area figure is good in many ways it has few drawbacks as well. It can only show two dimensions but when 3D Area figure is drawn, that would depict the data in a more precise and clear way when compared to 2D.