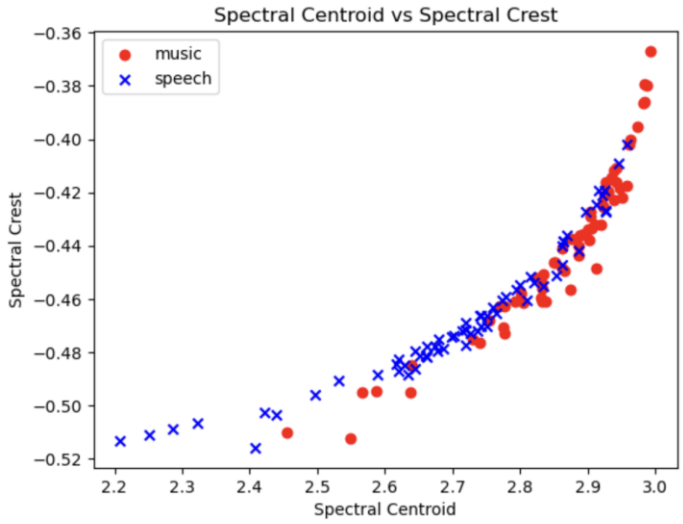
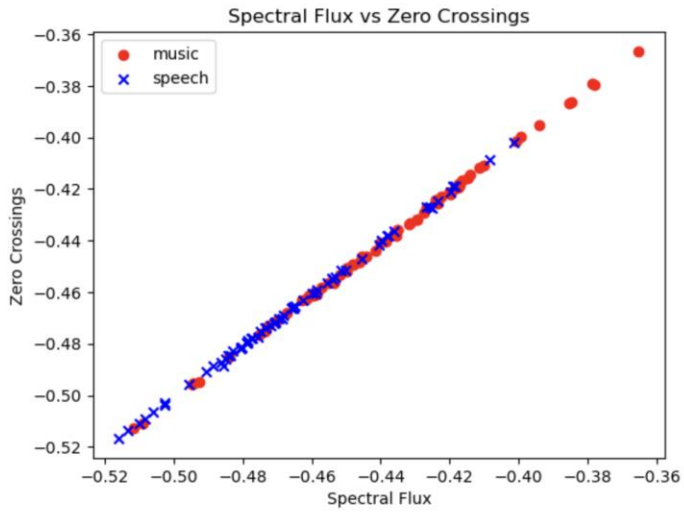
C. Feature Visualization Plots and Inferences

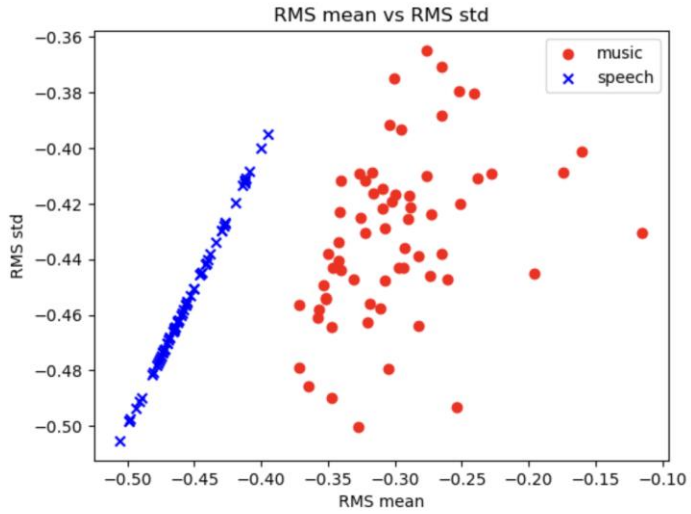
1. [SC mean, SCR mean]
    a. Spectral Centroid mean provides us with a spectrum representation of brightness where higher values would correspond with higher frequency information being detected like cymbals in a band recording. The plot allows us to infer that the bandwidth of the speech dataset frequency range had less high frequency influences on its mean value than the music dataset.
    b. Spectral Crest mean informs us of the signal to noise ratio in a spectrum where noisiness is indicated by a lower value and tonality is seen with a higher value. We would expect that speech would have lower values than music since most music is constructed around a tonal center. The plot follows that expectation. There are a few outliers that we suppose are due to the music dataset including songs not centered around a tonal center and the speech dataset including both stress and tonal languages.



2. [SF mean, ZCR mean]
    a. We can expect that the spectral flux mean will be sensitive to a spectrum that has significant frequency change over time. It is also expected that beats signifying tempo would result in music having a higher spectral flux mean than speech. The scatter plot supports that inference.
    b. Zero crossing rate, if higher, indicates a noisier signal. A higher mean would represent a signal with noisier frequencies and a more tonal signal would be represented by a lower value. This trend is present in the scatter plot with the music dataset having higher zero crossing rate mean values than the speech dataset.



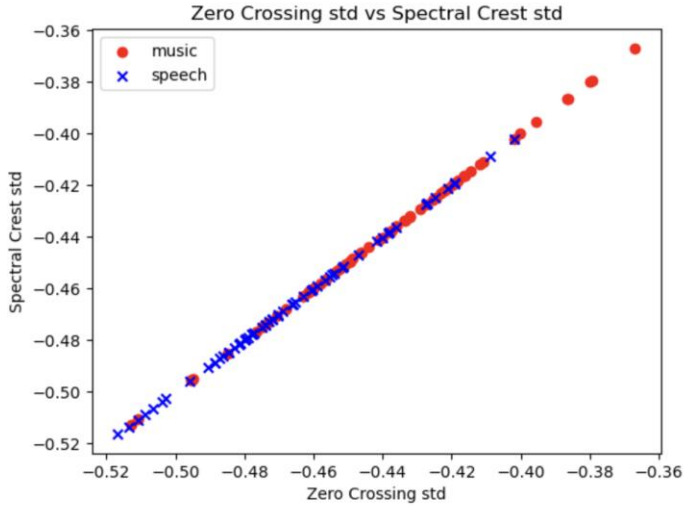
3. [RMS mean, RMS std]
    a. Root means square standard deviation is a value that shows variance in a signal's sound intensity where lower values indicate more constant or silent signals. Lower root mean square mean values indicate the same trend where higher values are representative of signals with more peaks of loudness in the signals dynamic range. The scatter plot allows us to infer that the speech dataset has more constant intensity levels than the music dataset. The narrowness of the speech values within a 0.12 range for both standard deviation and mean makes the music values appear more dynamic in the range of mean values of about 0.30 in the mean and 0.14 in the standard deviation. The scatter plot supports our inference that there is a larger dynamic range in music than speech.



4. [ZCR std, SCR std]
    a. Zero Crossing rate standard deviation indicates variation among the signal with higher values representing a signal with more changes in the distribution of noisy to tonal

sections. The scatter plot provides us with the inference that the music dataset has higher zero crossing rate standard deviations than the speech dataset.

b. Spectral Crest standard deviation provides a representation of the rate of brightness variation in a signal's spectrum. The scatter plot follows a similar distribution to the spectral crest mean where music has a higher value due to its frequency content having more variation than speech.



5. [SC std, SF std]
    a. Spectral Centroid standard deviation represents the variation in brightness value with lower values indicating less variation. The plot indicates a higher variability in the speech dataset as compared to the music dataset.
    b. Spectral Flux standard deviation is expected to be higher in spectrums that have pitch changes and frequent onsets detected. The scatter plot shows that the music dataset held a higher spectral centroid standard deviation than the speech dataset.