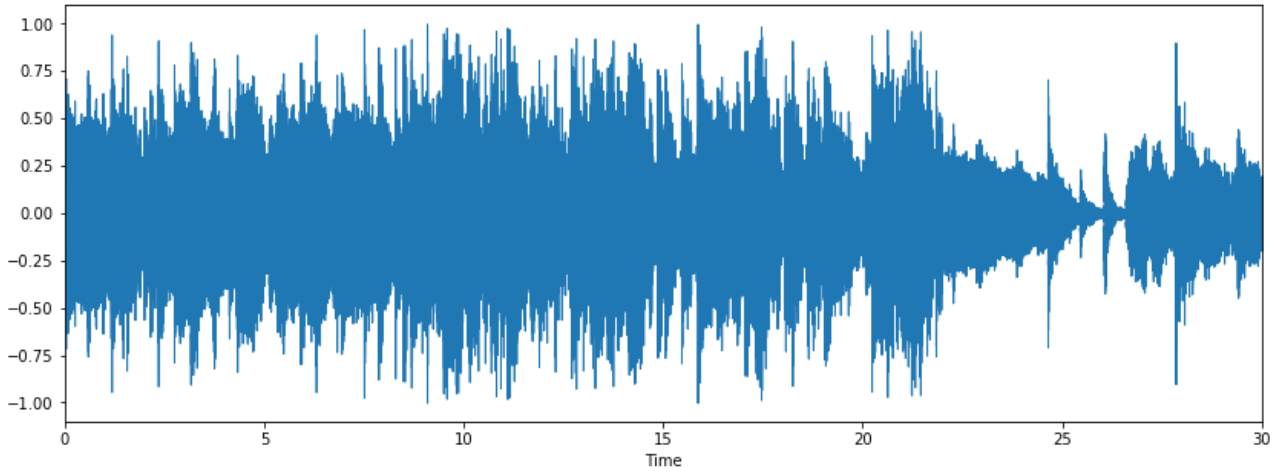
Analysis of audio signal(bonus task)

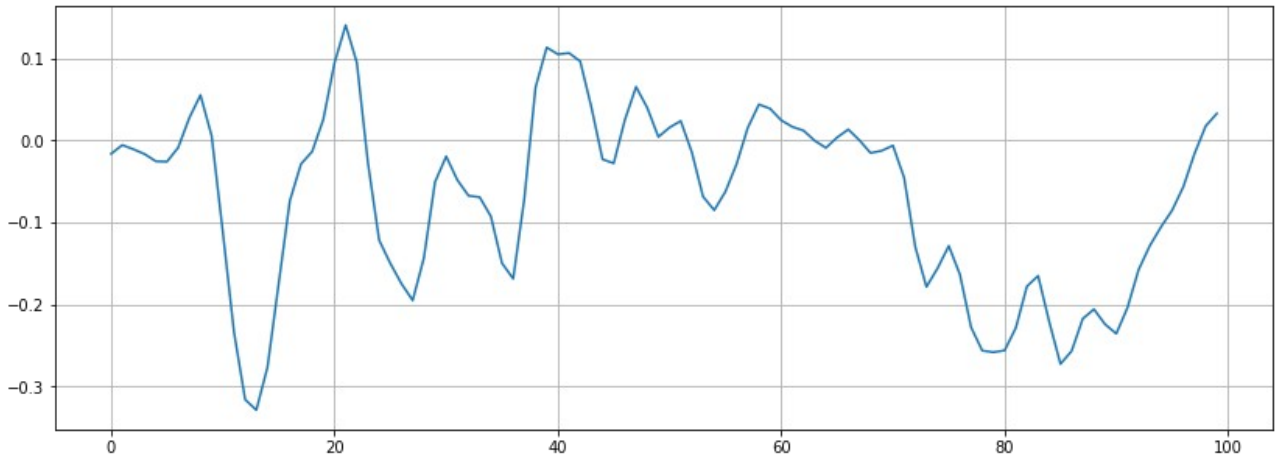
Note:Code is present in accompanying notebook
I have used one song of each genre to observe the feature values as representative of the genre.

1)Zero Crossing Rate

Rock

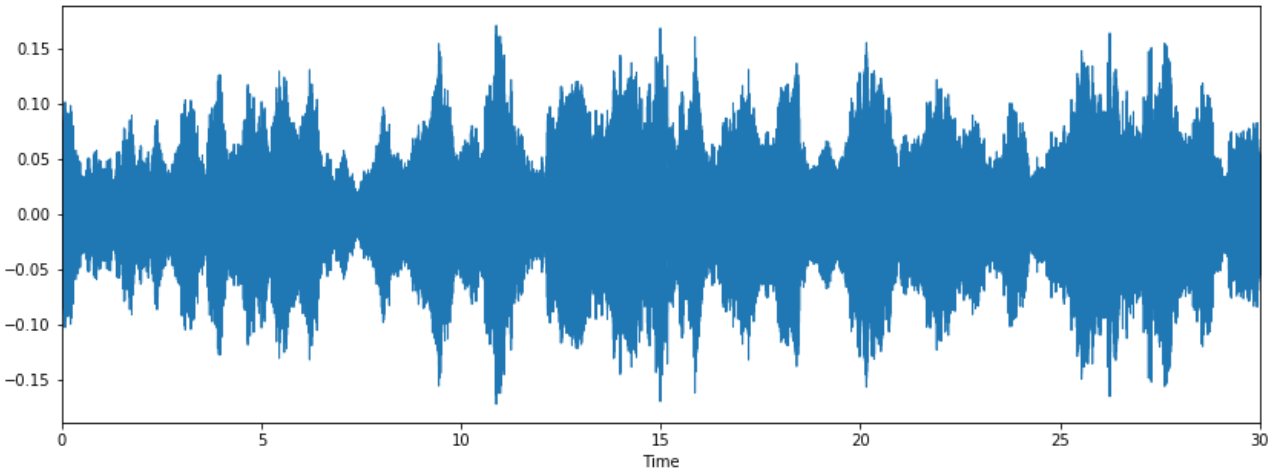


On zooming in,

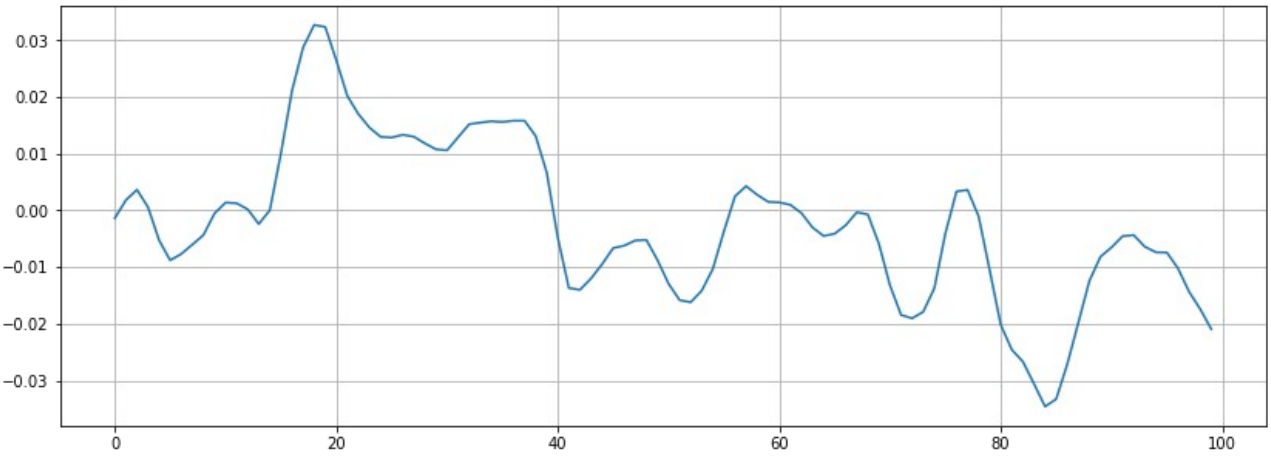


Total no. of zero crossings=36426

Classical



On zooming in,



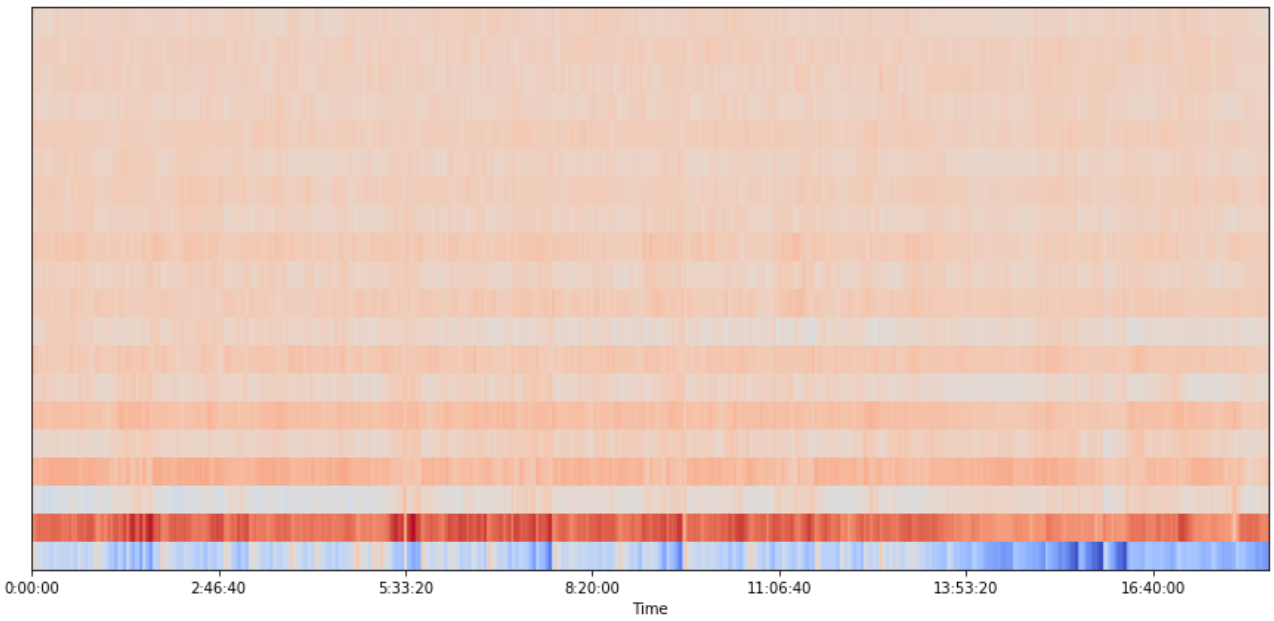
Total no. Of zero crossings=58180

On picturing the two waveforms side-by side, we see that the signal for rock music seems more packed toward a single amplitude(except towards the end) while the classical wave form shows variations in ampltitude.
Number of zero-crossings are more for classical music as can be seen by the numerically obtained result
Rock Music:36426
Classical Music:58180

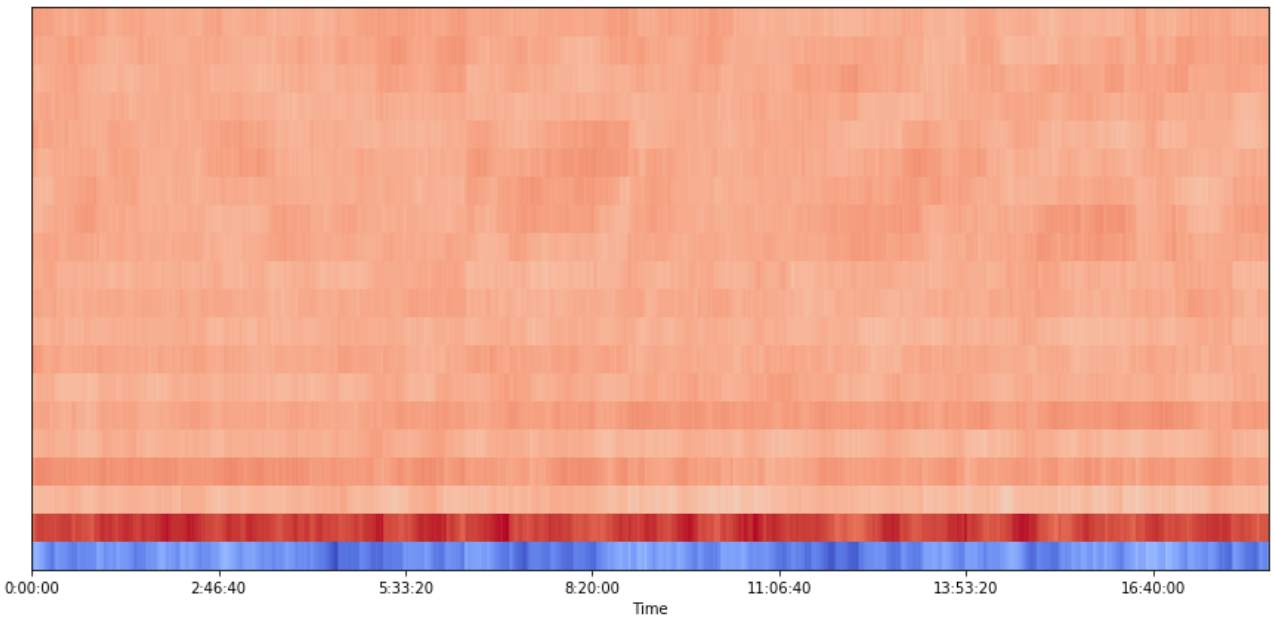
2)MFCC

MFCC takes into account human perception for sensitivity at appropriate frequencies by converting the conventional frequency to Mel Scale, and is thus useful in categorizing sound. It concisely describes the overall shape of a spectral envelope.
However, calculating the mfccs involves an involved procedure which includes linear cosine transform of a log power spectrum.
Therefore, interpretability directly from the diagram is difficult.
Nevertheless, we do see a difference in the two figures.

Rock



Classical

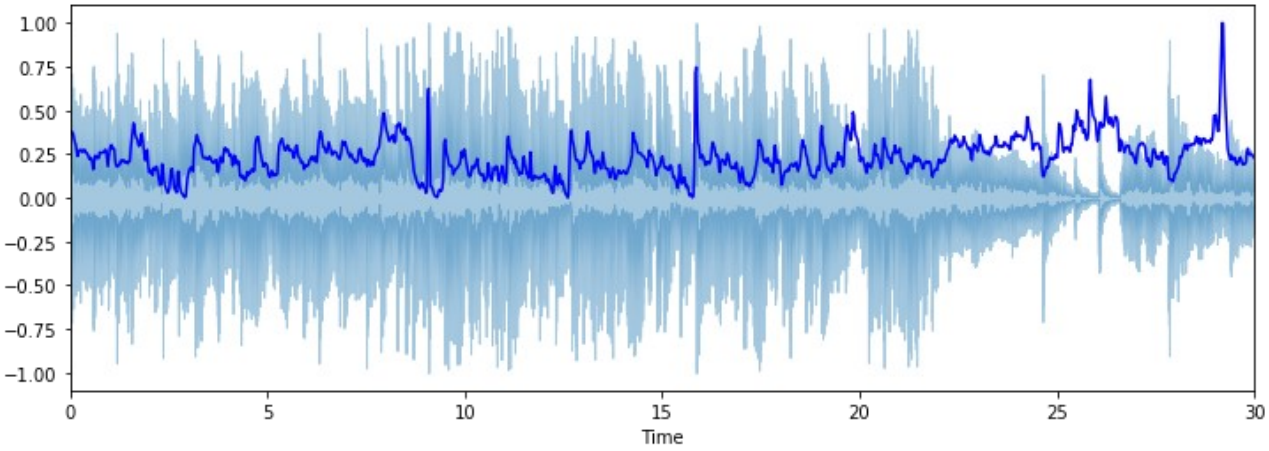


The plot is darker in shade for classical music than for rock music.

3)Spectral Centroid

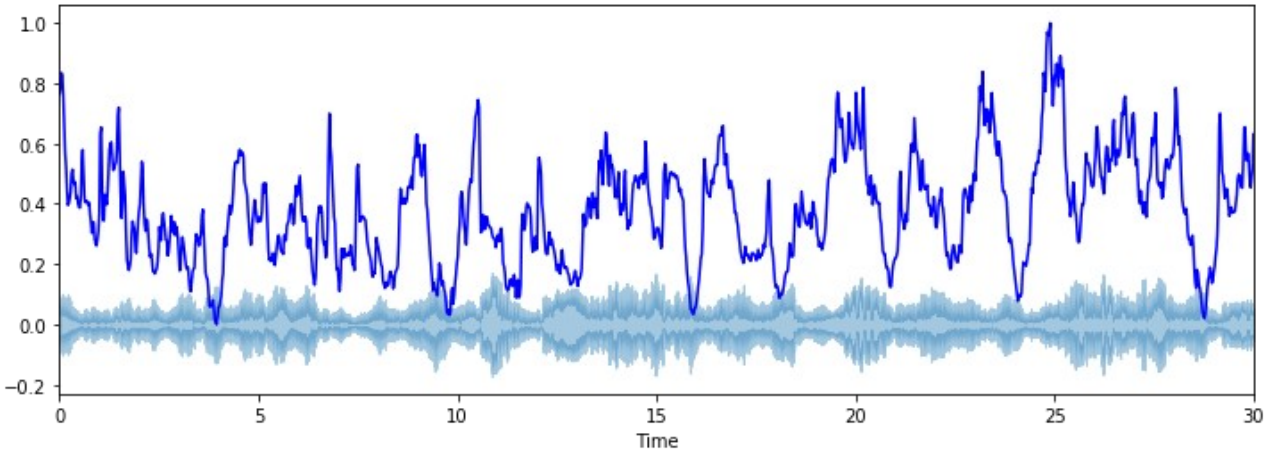
The spectral centroid indicates at which frequency the energy of a spectrum is centered upon or in other words It indicates where the "center of mass" for a sound is located. This is like a weighted mean.

Rock



The graph shows the variation of spectral centroid on an absolute scale as a continuous function of time. The wave form at that time is also plotted in a lighter shade alongside.
The spectral centroid appears fairly centred at the start and shoots up only occassionally throughout the playtime.

Classical



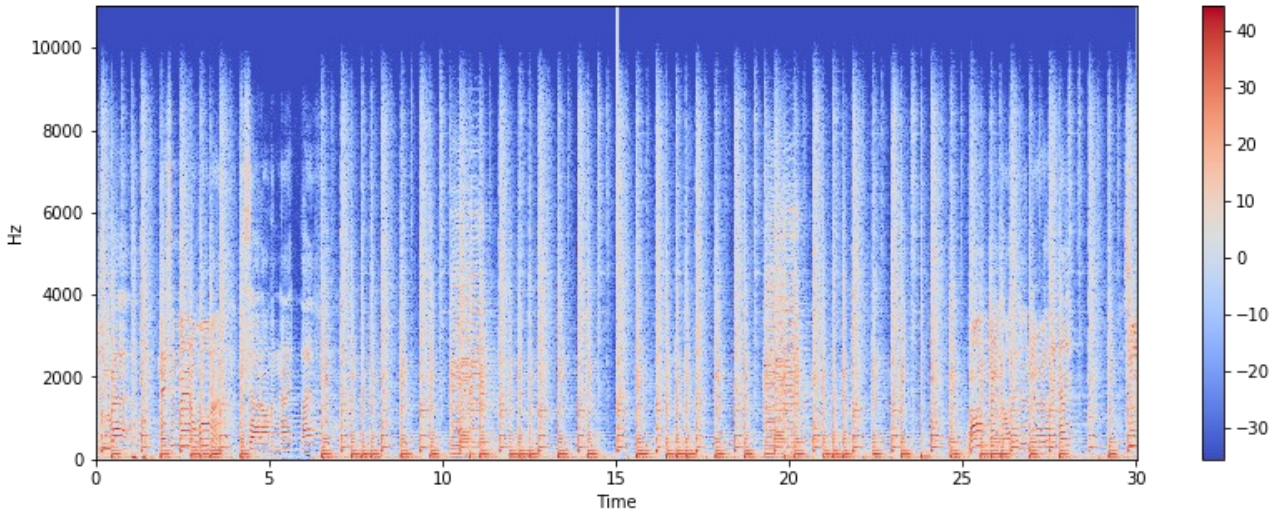
In case of our classical music sample, a greater variation in spectral centroid is seen with Max-Min almost 1.0 on a relative scale which is greater than the maximum variation in classical music. Furthermore, the change is more frequent.

4)Power Spectrogram

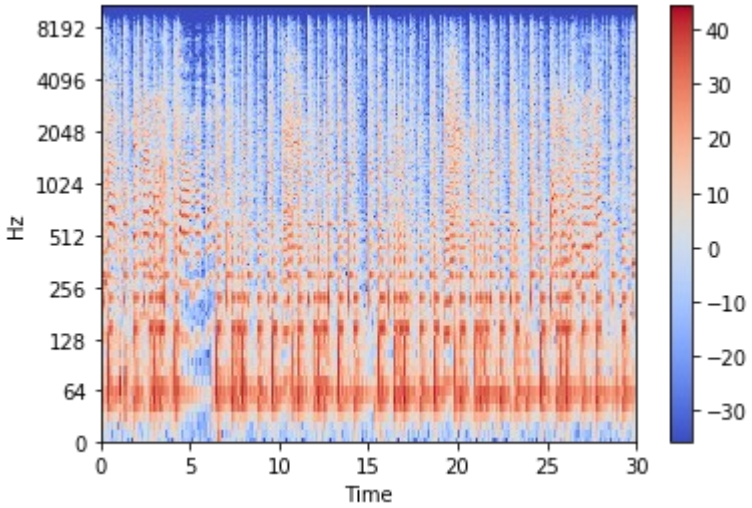
A spectrogram is a visual way of representing the signal strength, or “loudness”, of a signal over time at various frequencies present in a particular waveform. Not only can one see whether there is more or less energy at, for example, 2 kHz vs 10 kHz, but one can also see how energy levels vary over time.

A spectrogram is usually depicted as a heat map, i.e., as an image with the intensity shown by varying the color or brightness.

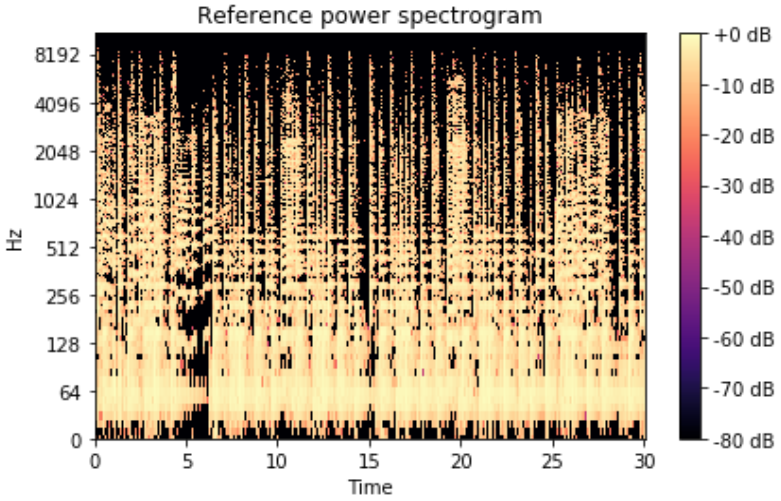
Rock



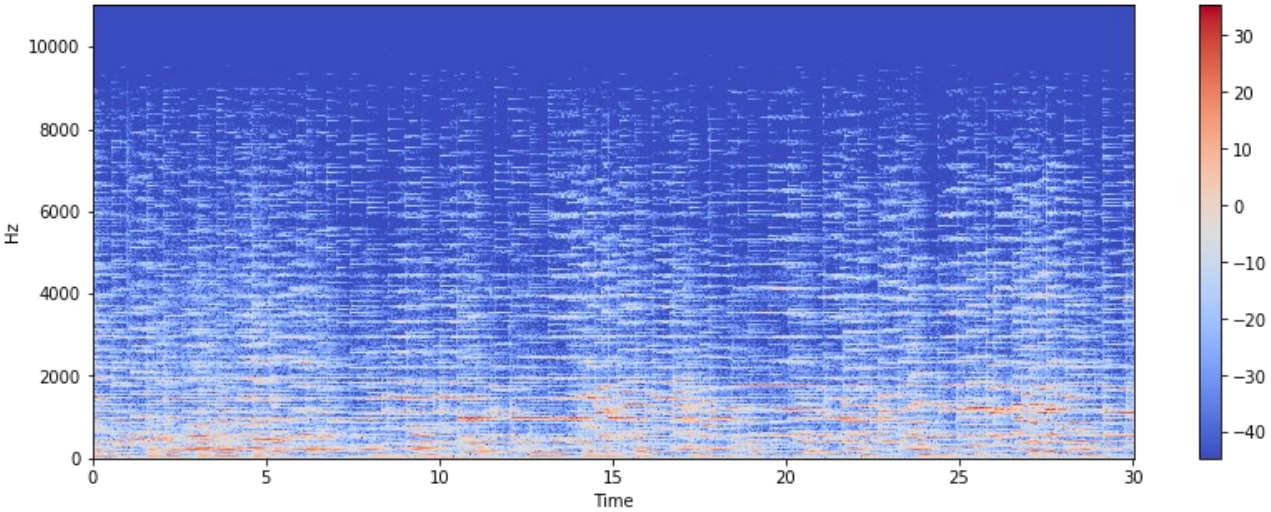
The below plot is a logarithmic plot(along y) of the above plot so as to make it clearer.



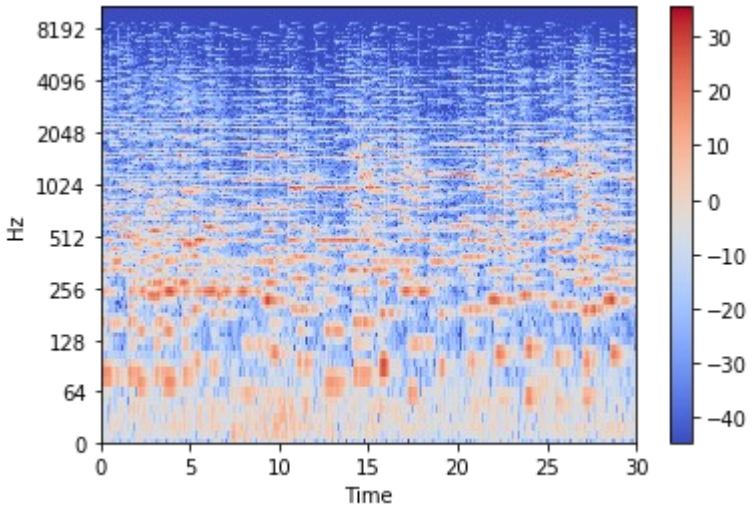
Decibel Plot



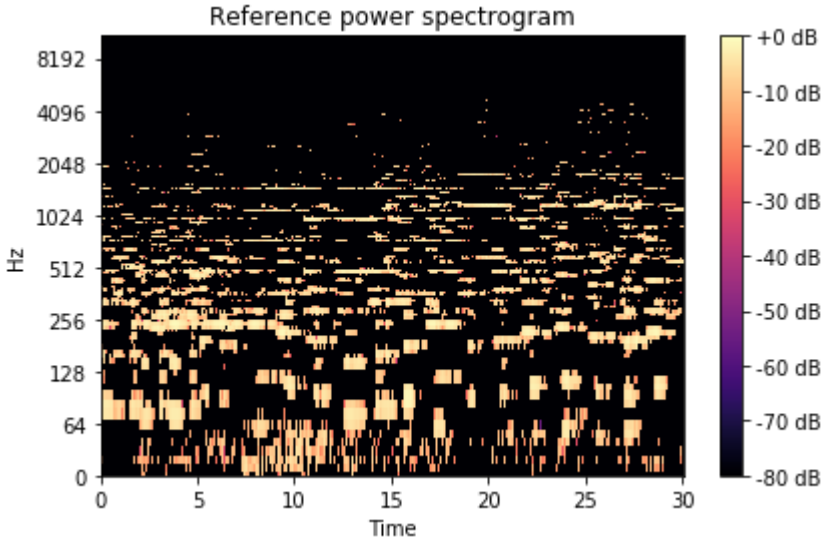
Classical Music



The below plot is a logarithmic plot(along y) of the above plot so as to make it clearer.



Decibel Plot



The logarithmic plots for both the songs show differences.
Greater decibel values are seen for a larger range of frequencies throughout the song length for rock music in contrast to classical music.
Therefore, one can expect rock music to have greater rms energy.

VIBHHU SHARMA
EE19B128
Phone-9881129406