Psycho Music

Rahul Gomathi Sankarakrishnan
Paveethran Swaminathan
Akhil Yenisetty

Abstract:

The goal of Psychomusic is to examine and depict how music influences feelings and actions in people. There are moments when hearing a song makes you joyful and excited. Other times, it makes you nostalgic or melancholy because you remember certain things. These abrupt shifts in human feeling are linked to musical qualities. The listener's emotions are influenced by a song's frequency, volume, tones, tempo, genre, and instrumentation. This study will analyze and explore numerous musical and emotional trends through the use of various visualizations.

Introduction:

We saw a video where the emotion felt by us was changing with the change of music. We felt this was interesting and surprising how the thinking of the human brain changes with respect to music, then we started to see if there are any research papers or any data set we could find. Surprisingly we found that there are many medical treatments that are performed using music,and also many places like restaurants, clubs, theme parks, etc set the mood of the people with music. We thought if we could analyze and visualize the relation between music and human emotions, we could find some interesting facts and patterns between them. We believed that these results could help in various fields, where music plays an important role.

Music is a powerful stimulus and it does more wonders than we can imagine. People usually resolve to music as a form of enjoyment and happiness. The reason we feel mixed emotions is because in some way the musical waves interact with brain waves and different people feel different emotions.
We resonate more with the song when the two waves add up to create constructive interference and a negative or no emotion when it creates destructive interference.
In some recent research, music and single tone music are tested as a cure for Coma and Alzheimer's disease, where the researchers hope to use the nostalgic emotion to help the patients recall past events.

During the initial research we found plenty of research and existing visualizations regarding this subject matter. One research aimed to investigate frequency-dependent cues for human emotions. Experimentation was conducted on random people from different age groups, with their moods recorded before this observation.

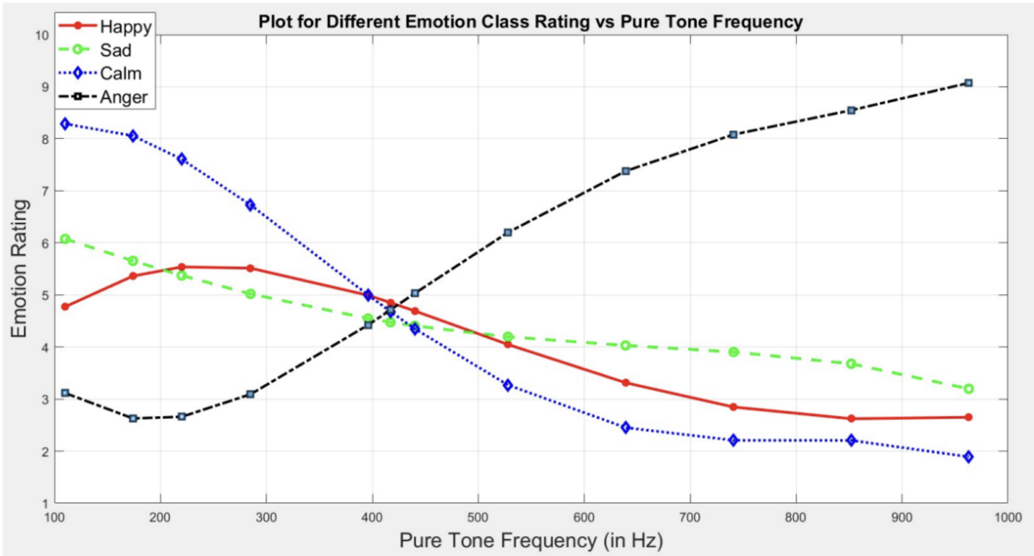


Figure 1. Plot for different emotion class rating vs Pure tone frequency. Here, the line graph for all four primary emotions—Happy, Sad, Anger, and Calm are plotted w.r.t pure tone frequency (in Hz).

Key findings from this line graph like the fact that crossover points for all four primary emotions lie in the frequency range of 417–440 Hz. This clearly depicts that the frequency range 417–440 Hz is neutral. However, there is no mention of accounting for the variables and confounding factors that the researchers documented, such as age, the subject's disposition before the experiment, the type of music they were listening to, etc. And this makes us believe that the plot is misleading in some way - we do not know if this emotion rating represents average values of a proper population with equal importance to each category of Age and Genre of music. Because there is no mention about how the data was sampled, whether they assumed equal distributions of different age groups or if they took into account the placebo effects by recording the mood prior to listening to the song, we can not conclude the information extracted from this plot.

The same researchers have tried to convince us by summarizing average emotion ratings for every emotion while listening to different frequencies, by using box-plots:

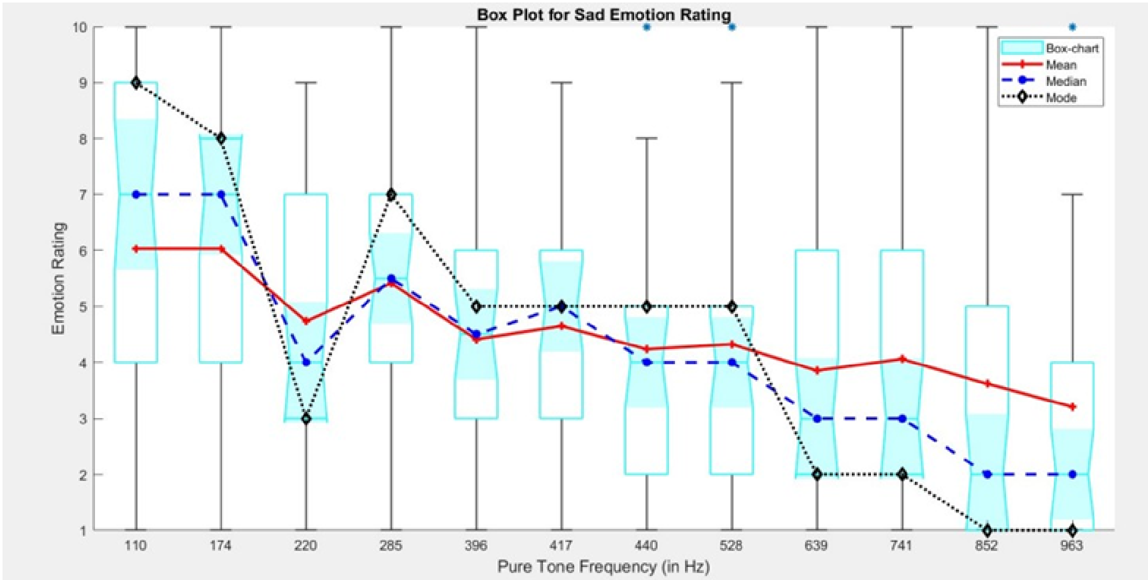


Figure 3. Plot for Sad emotion rating v/s Pure tone frequency. Here, the shaded area (cyan in color) of box chart represents distribution for annotated emotion corresponding to each pure tone frequency.

Acoustics **2022**, *4* 464

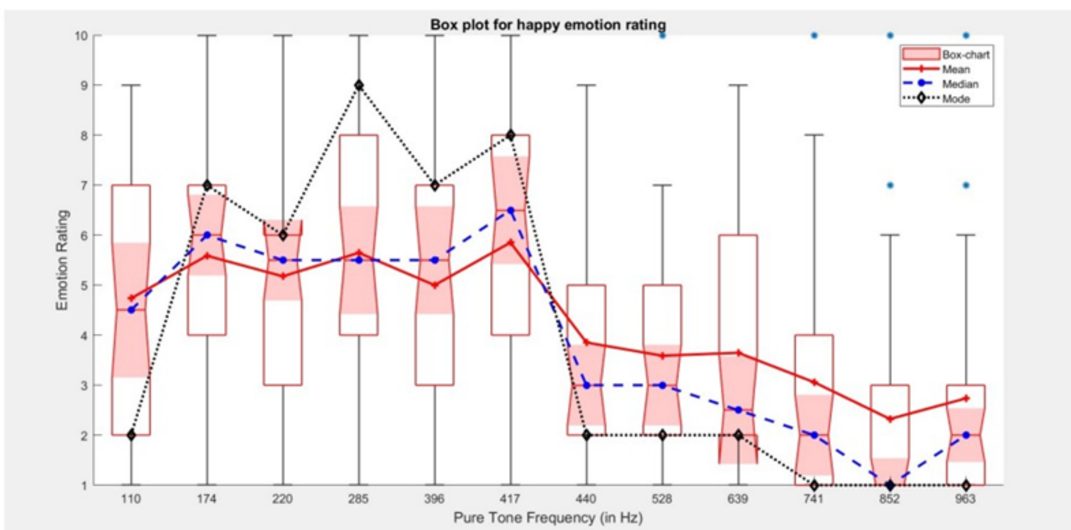
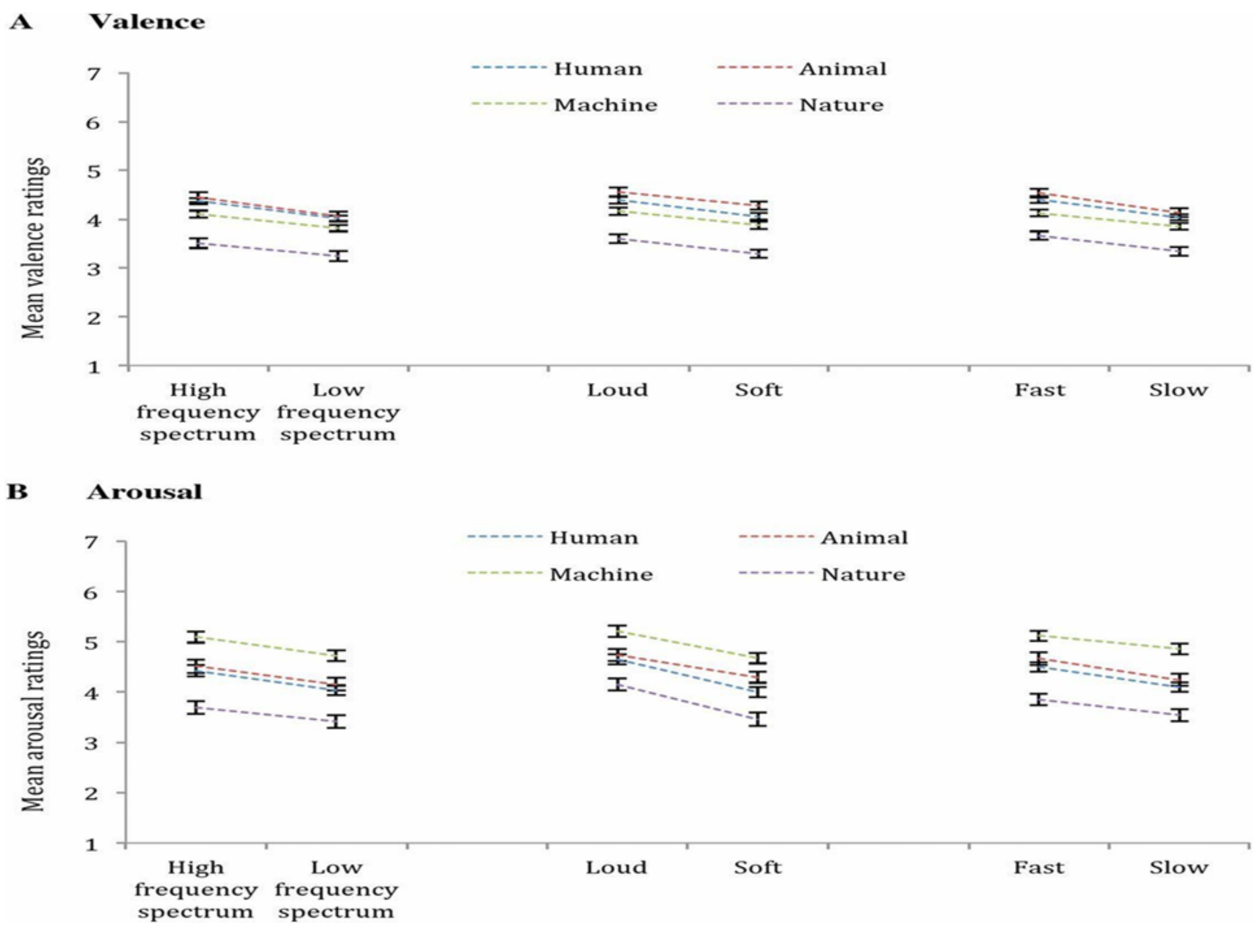


Figure 2. Plot for Happy emotion rating *v/s* Pure tone frequency. Here, the shaded area (red in color) of box chart represents distribution for annotated emotion corresponding to each pure tone frequency.

Similar to the previous critique, we can mention that the authors have not considered how this ratings by people vary for other characteristics like age of the rater, mood before he rated the song, if he likes or dislikes the song etc. These are some things that we plan to improve and visually present for rich information gathering.

A better way would be to use histograms or bar charts to differentiate average emotion ratings over different age groups over varying frequencies. We also plan to have an interactive plot (Altair based) to allow people to interact and change different attributes of x and y axis, to compare multiple subplots at the same time. This would make the visual more transparent and detect direct patterns. One more important point to notice in these box-plots are that there are shaded regions - we are not sure what those shaded regions represent or if they are supposed to depict some significant attribute. But having those shaded regions surely does not improve our understanding of the data and it only makes it more complicated.

Similarly, another research tracks human emotional responses to changes in the acoustic environment. And one of their plots is this:



The two features Arousal and Valence are compared against high-low frequency, load-soft sounds and the slow or fast beat-pace, even though it's intuitive to use these features, their poor choice of visualization plot has completely diminished our interest about this topic. With all the unnecessary overlaps and occlusions, it's very hard to get any information from this plot. Using continuous lines is better and helps people perceive the information much easily. Moreover, our choice of dataset is different from what these researchers have used, and we believe our dataset has the potential to uncover many different relationships and trends about human emotions and various acoustic attributes.

In our research, we sought to create visual representations of the variations in frequency in connection to the musical style. because a certain genre is frequently linked to a specific emotion. It must paint a vivid and precise picture of how a certain tone or beat can cause subtle to significant emotional changes in a person. Separate graphs that show the relationship between the listener's age and how they perceive a certain genre are also anticipated to be produced.Based on a particular range of frequencies and how they relate to related age groups, a pattern is anticipated to emerge. Additionally, a factor of preference is included to illustrate the correlation between how a certain age group perceives various genres in various ways. The initial idea is to develop a visual plot that is interactive and offers toggle options for switching between various age groups, genres, and likability variables.

Process:

The dataset to be used in the project is collectively present under the name of 'Emotifydata'. It is an openly available dataset which comprises 400 songs in the form of MP3 files, along with their corresponding instances of the times they were listened to by a listener. The CSV file is imported on to the Python file from its original path, while the audio files were made available to use through relevant modules.

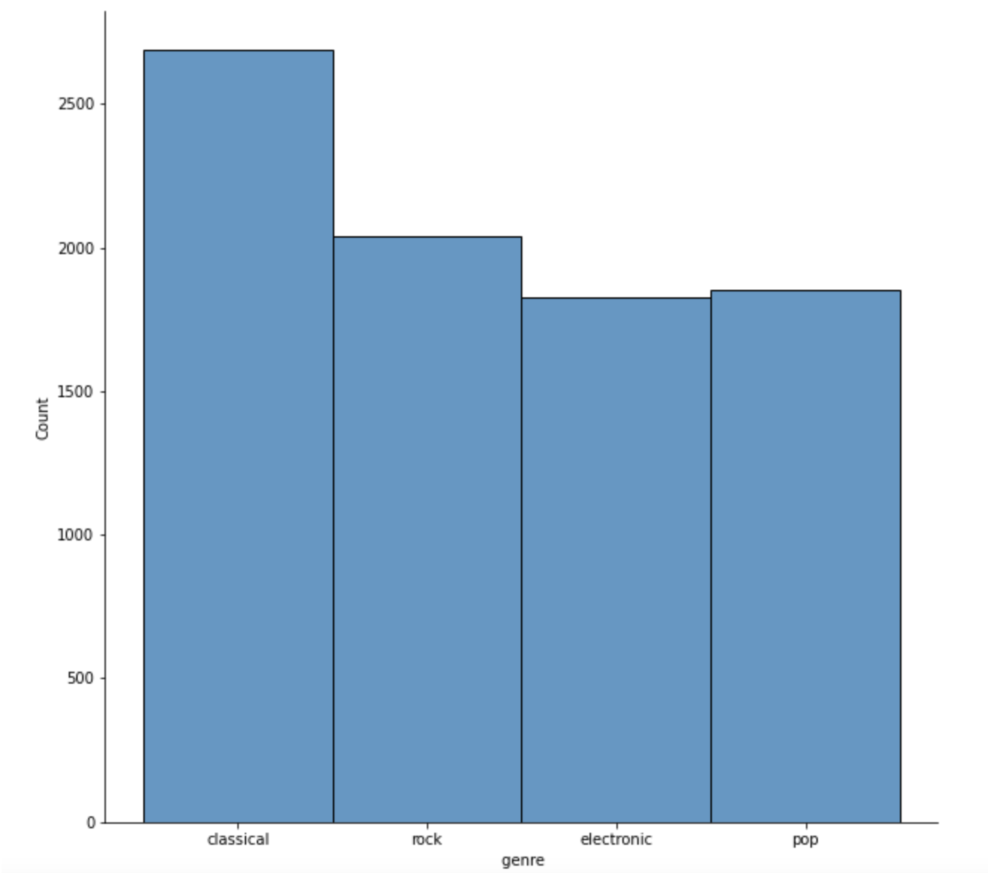
The primary task accomplished was to portray a depiction of the contents of the dataset through exploratory data analysis, consequently displaying numerous visualizations of the raw dataset. The visualizations show the distribution of the dataset in relation to its features. While most of the ideas and correlations came into fruition as coherent visualizations, one idea had to fizzle out.

Initially, a plan was made to create a cartogram to depict the various countries the listeners belonged to, and further depict its relevance to the liking pattern of each genre. However, the dataset only consisted of the mother tongue of the speaker rather than the country they are from, making it impossible to designate them to a specific country (especially for ubiquitous languages like English and Spanish). Moreover, the distribution of data was extremely skewed in favor of very few prominent languages, meaning there was no meaningful observation to be made.

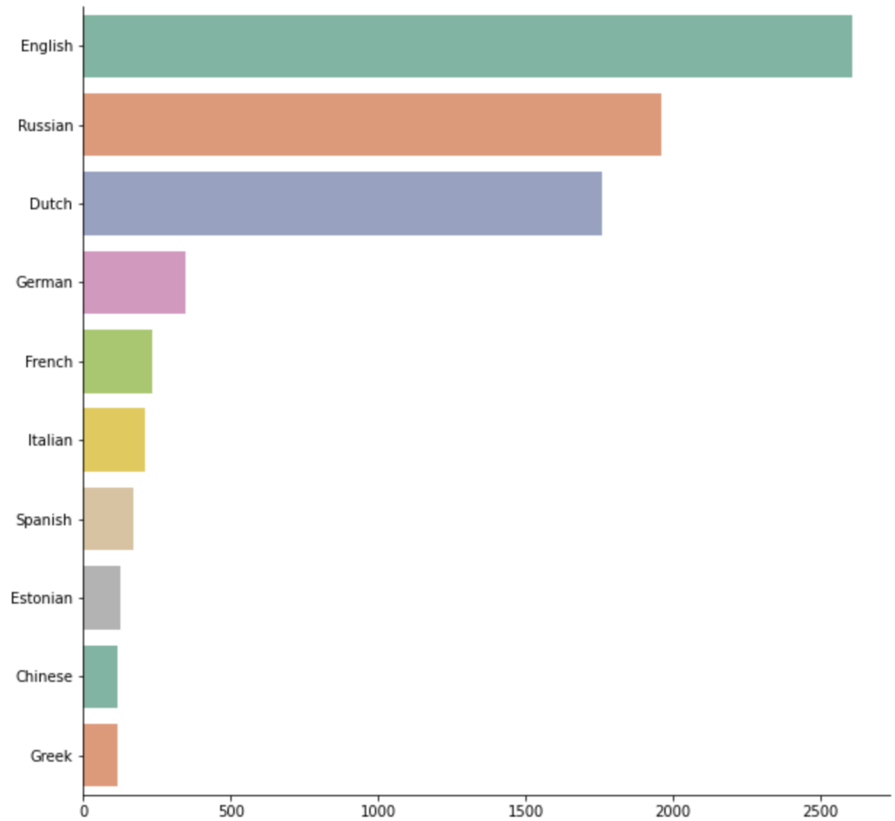
Subsequently, the audio files were analyzed to present visualizations which connected the frequency ranges of the songs with the likability of each genre. In general, the project makes use of numerous visualization techniques namely bar graphs, histograms, KDE plots, box plots, joint plots, radar charts, violin plots, heatmaps and line graphs. Other visualization techniques were also considered, but in the end failed to reciprocate the results obtained by the existing ones. Scatter plots and hex bins were prominent points of discussion regarding graphs to identify very specific portions of the dataset, but in the end failed to provide legible and understandable observations.

Results and insights:

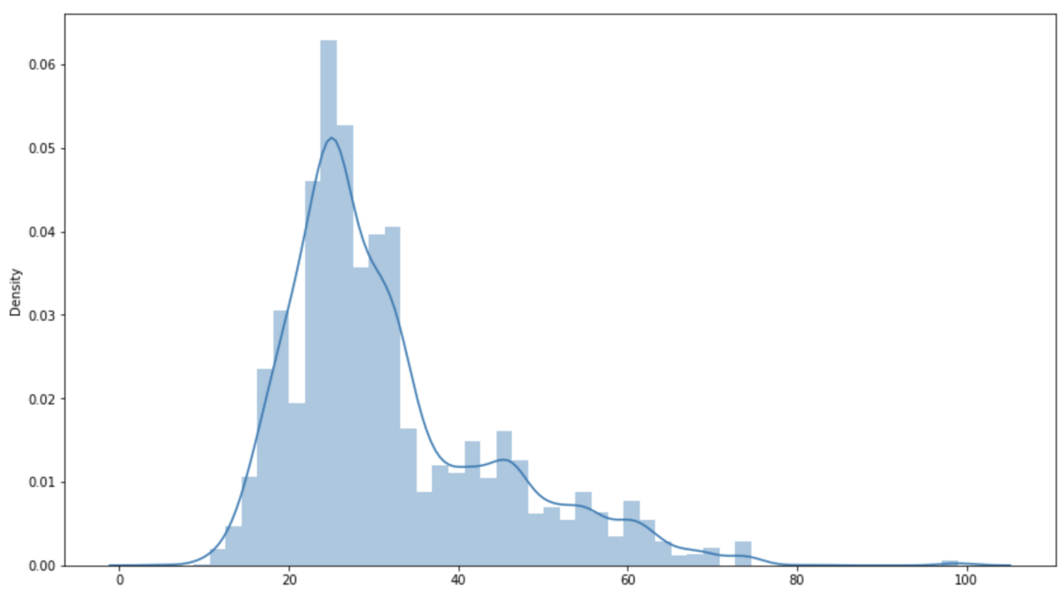
The initial plot of our study is how well the genres are distributed, when we have produced a graph between genre and their count we observed that the count for classical music is more. All the genres are unequally distributed because at the time of experiment, the volunteers were given a choice to skip the genre if they don't want to listen to it. By this we can see that many people like classical music compared with the other genres.



We wanted to see the native language of the volunteers and found that the majority of the people volunteering are either English, Russian and Dutch speaking. We were curious if language played any part in altering emotions.

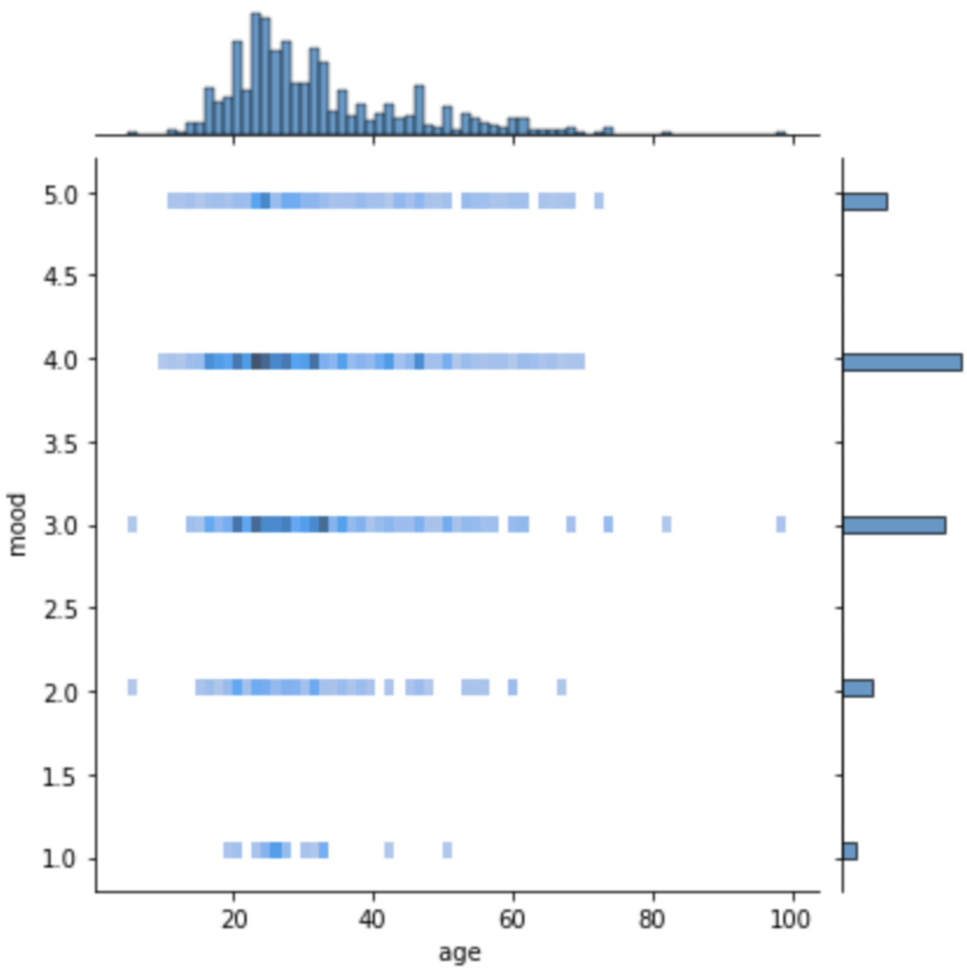


We have plotted a graph for age groups and see how the age of different volunteers are and we have seen this pattern.



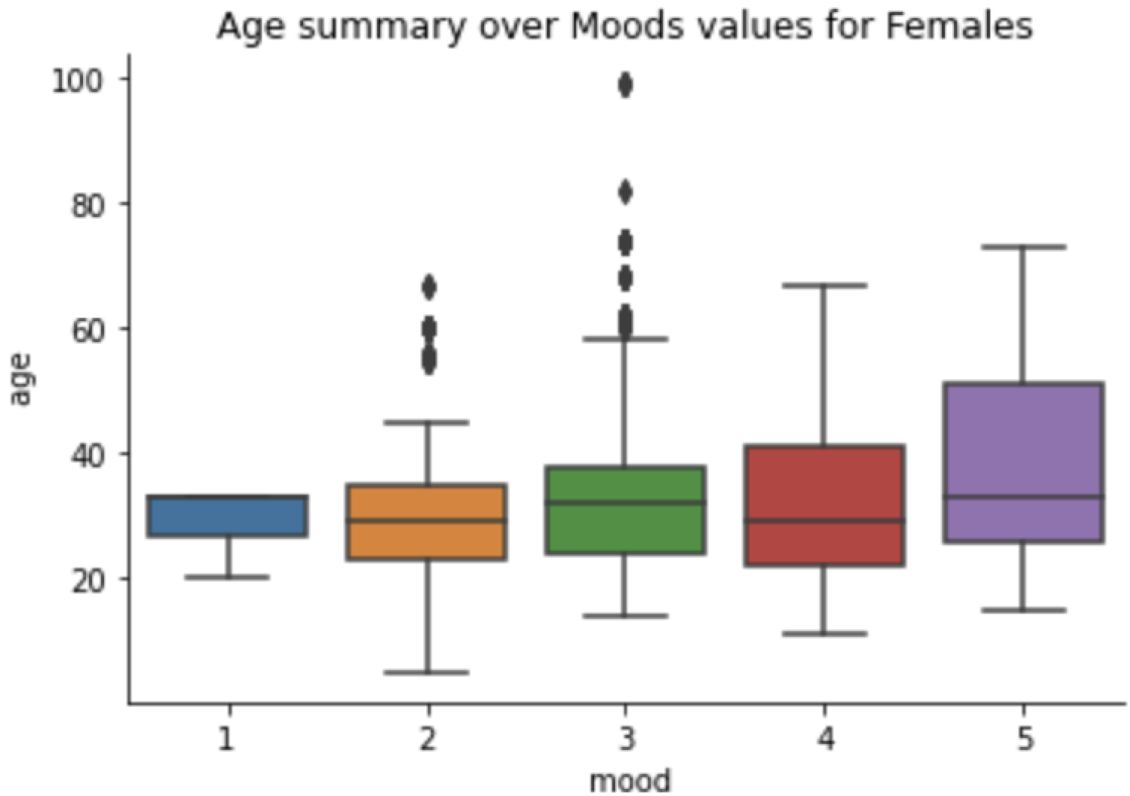
This shows us that the majority of the volunteers are from the age group 25 to 35. The age varies from 5 to 95 years. This distribution plot along with KDE made it easier to understand how the age is distributed.

We wanted to see a pattern or relation with age and mood. We used a joint plot for visualization.

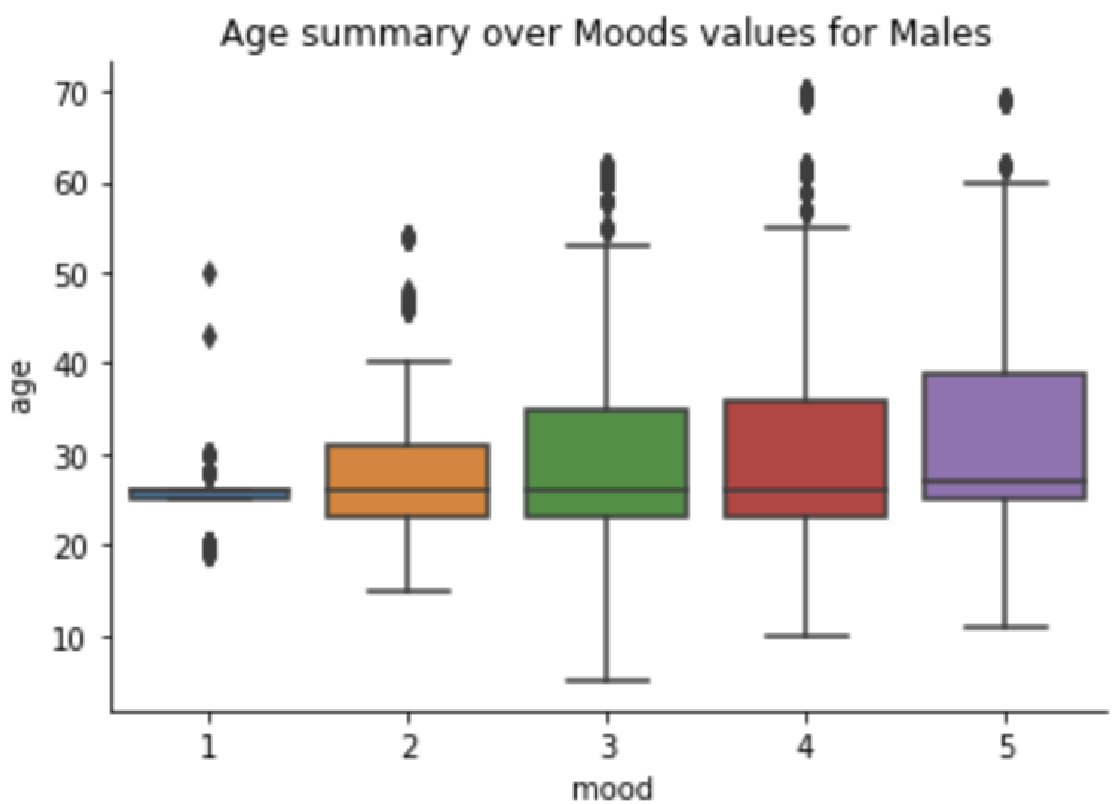


The histograms on top and right indicate that the majority of the people are in the age group 25 to 35 and are in mood 3 and 4. The dark spots in the graphs are in mood 3 and 4, apart from that there are also few spots where we can see how the mood is distributed among different age groups. This was a better representation for mood and age.

We have also tried a box plot to see if we can get extra information for males and females mood and age. We were able to see the mean of each age group.

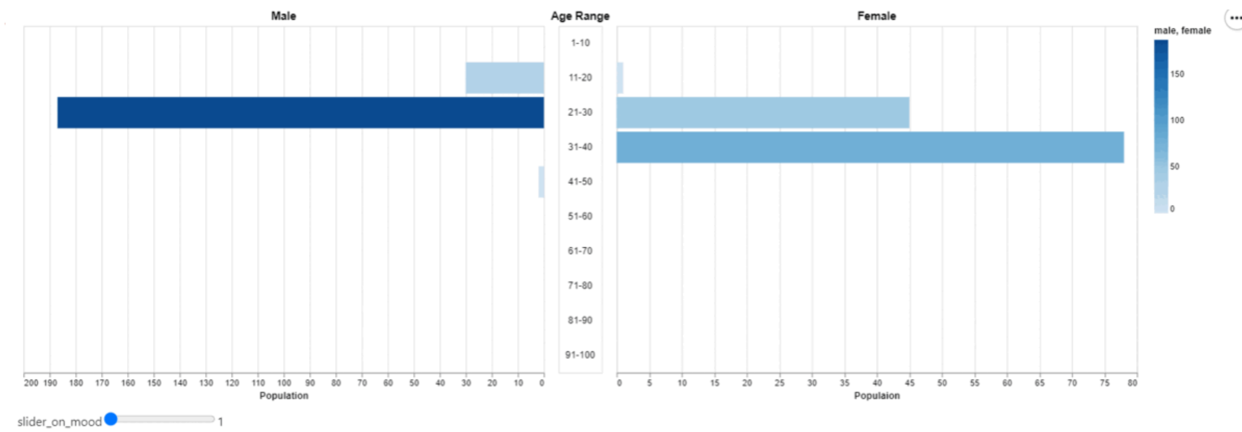


In Females the mean of all the moods are from age 25 to 35, we can see many outliers in mood 2 and 3 this indicate that these moods have a wide range of age groups.



Whereas in males box plot the outliers are seen in all the moods, this tells us that each mood has a wide range of age groups. Even in males box plot we can see that the mean for all the moods are in the range 25 to 35 years of age.

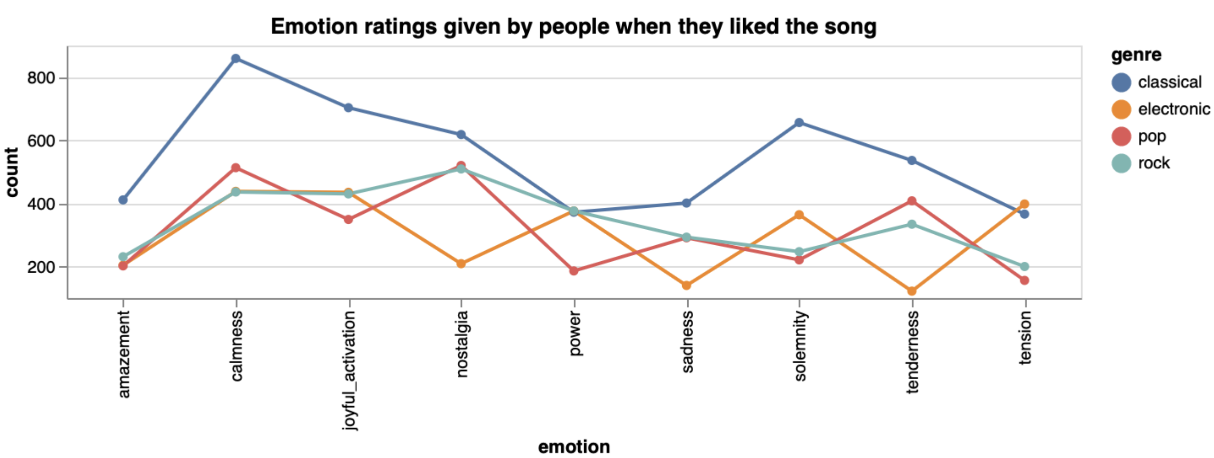
We have created an interactive plot to see how the variations are between males and females.



We have observed that males are dominant in mood 1 and mood 5, whereas females are more in mood 3. In both males and females only moods 3,4 and 5 have a wide range of age distribution. The interactive graph was more informative and better that the previous box plot.

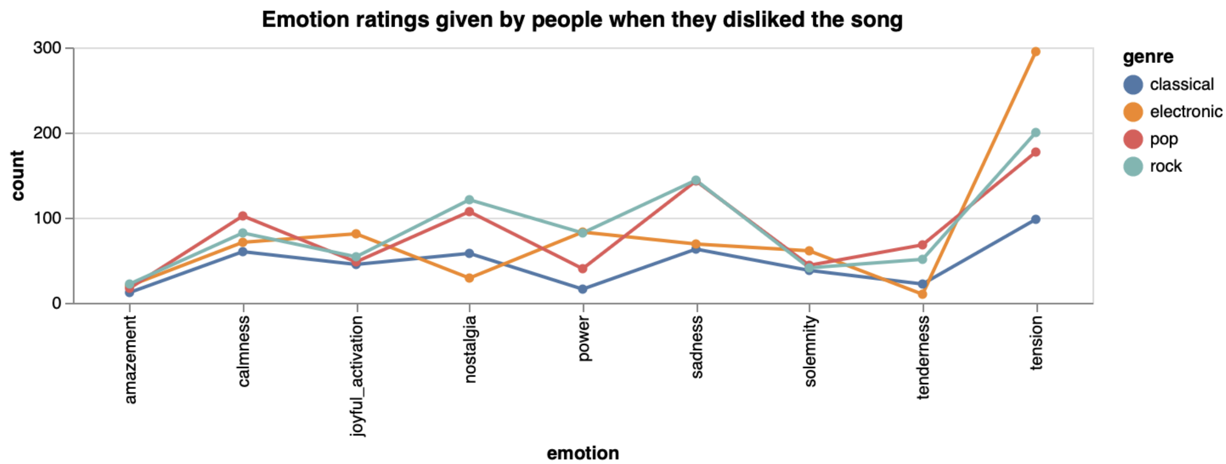
Emotions in liked/disliked songs:

We wanted to see if we can find and trend in emotions for liked and disliked songs. To visualize that trend we have used a line chart, where the emotions felt are in x-axis and the count for each emotion is in y-axis.



We can see that classical genre tops in all emotions, this is due to the skewness we have seen in the genre distribution. Interestingly amazement and power felt for all the genres is the same for liked songs. We can see that calmness is high for all genres, it tells us that we are calm when we are listening to the music we like. Sadness and tenderness for electronic music is very low. We can also tell that apart from the electronic genre, listening to all other genres makes you nostalgic for liked songs.

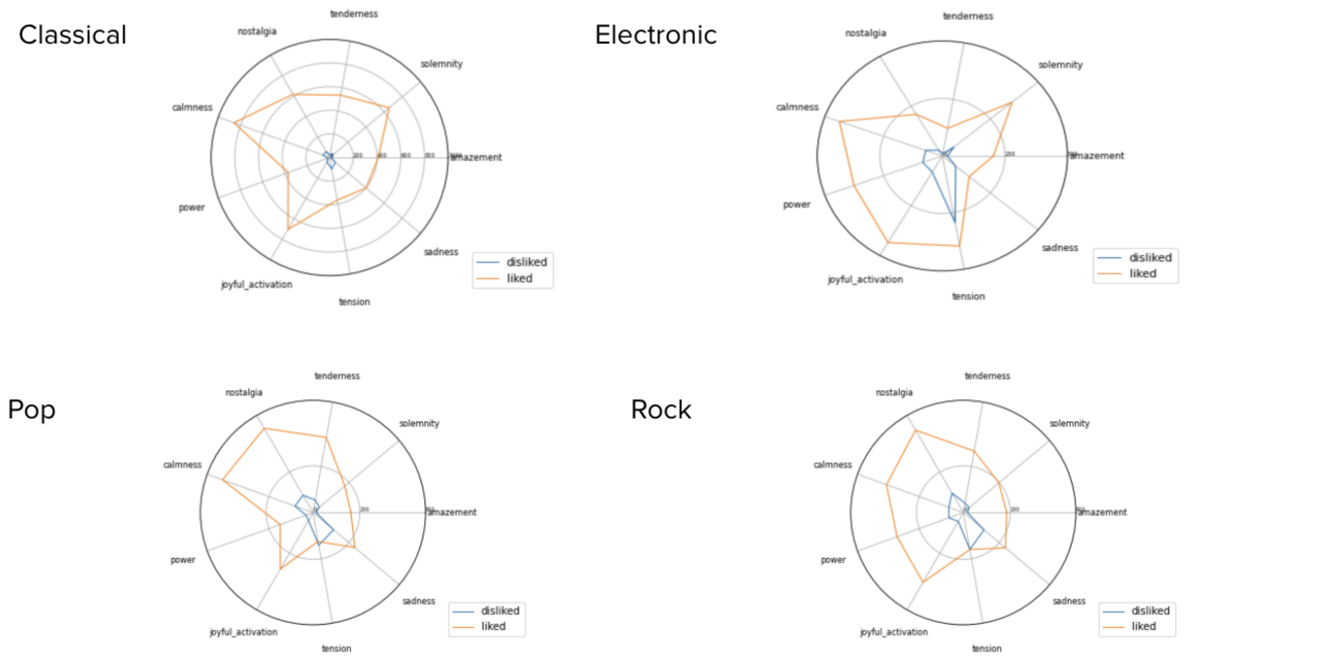
Lets see what we found for disliked songs:



We can clearly see that all the emotions felt irrespective of genre are the same except for sadness and tension. When you don't like pop or rock songs, listening to them can make you sad. Irrespective of genre, if you don't like a song, you tend to build a lot of tension listening to it.

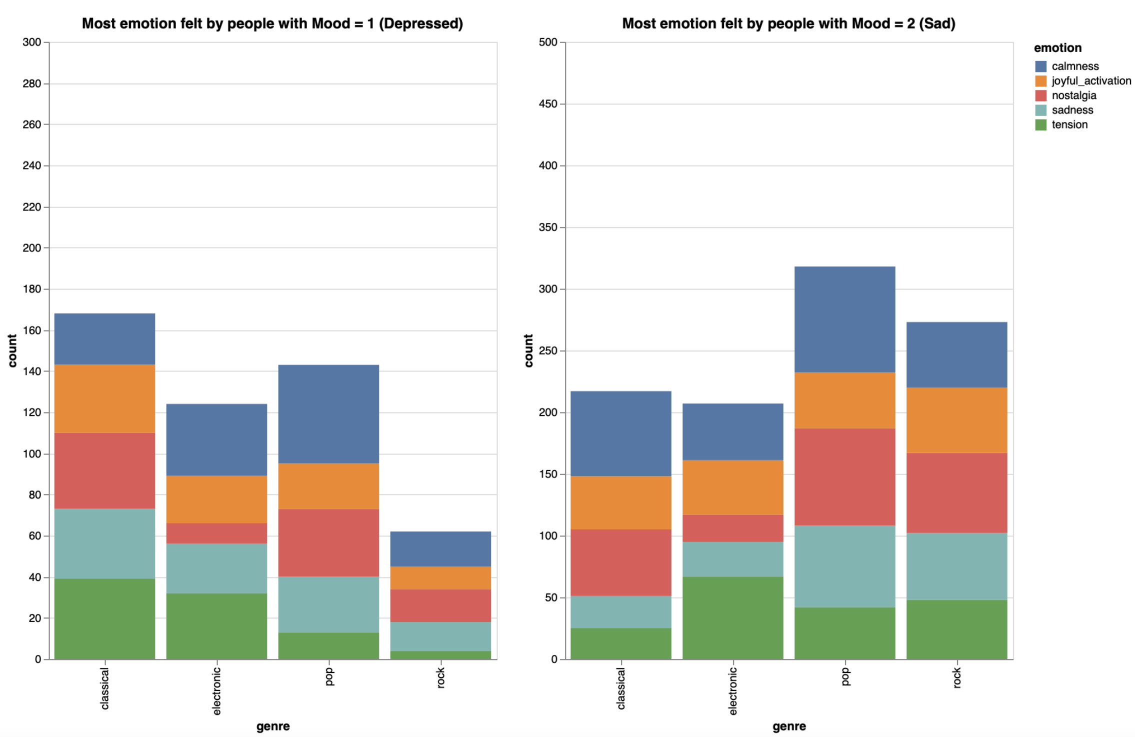
We can clearly see that different emotions are being triggered for liked and disliked songs.

We wanted to get more information about the emotions for liked and disliked songs. So, we have generated some radar charts for each genre to see the trend. The orange line represents the emotions for like songs, blue for disliked songs.

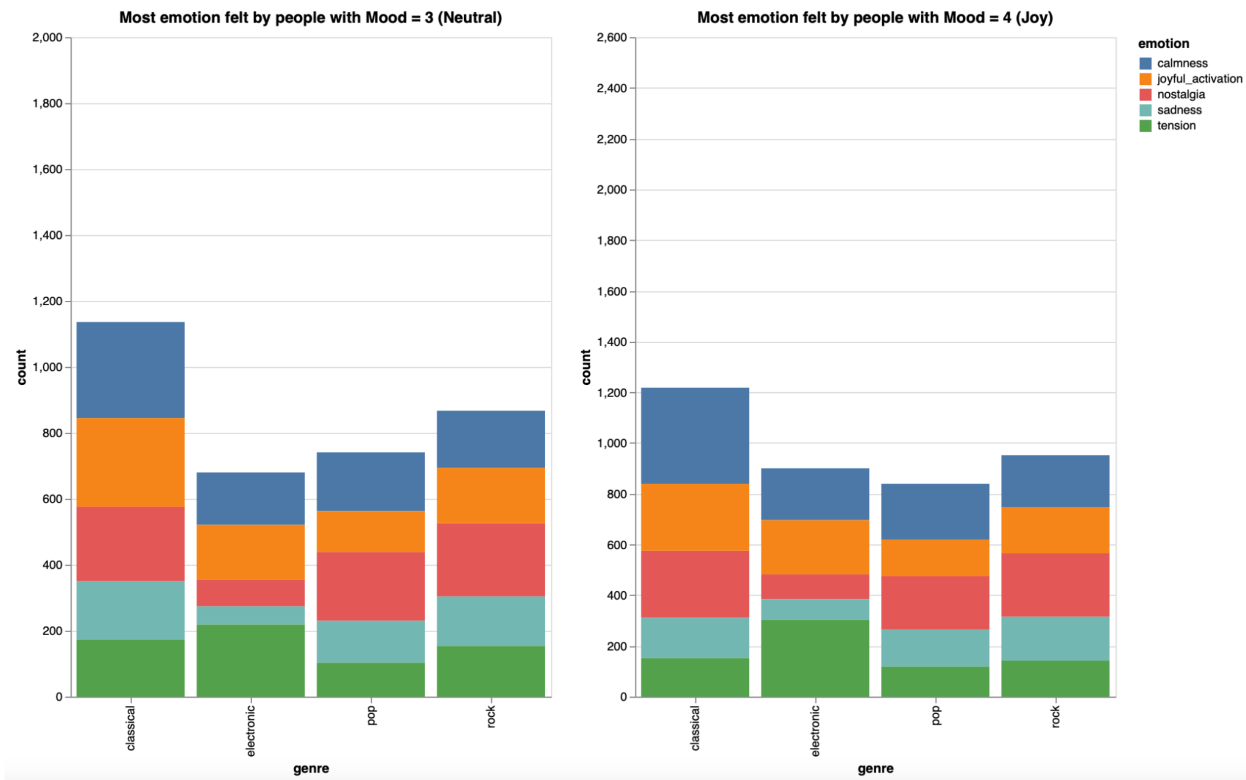


Generating these charts was very informative. We saw that there is a big variation in calmness and joyful activation for all the genres. If you dislike a song these emotions are very low but if you like the song these emotions are very high. Interesting observation here is that except for classical, the tension felt for liked/disliked songs is the same in other genres. In all genres if you dislike a song you tend to feel tension and sadness, whereas you feel nostalgic,calm and joyful if you like the song.

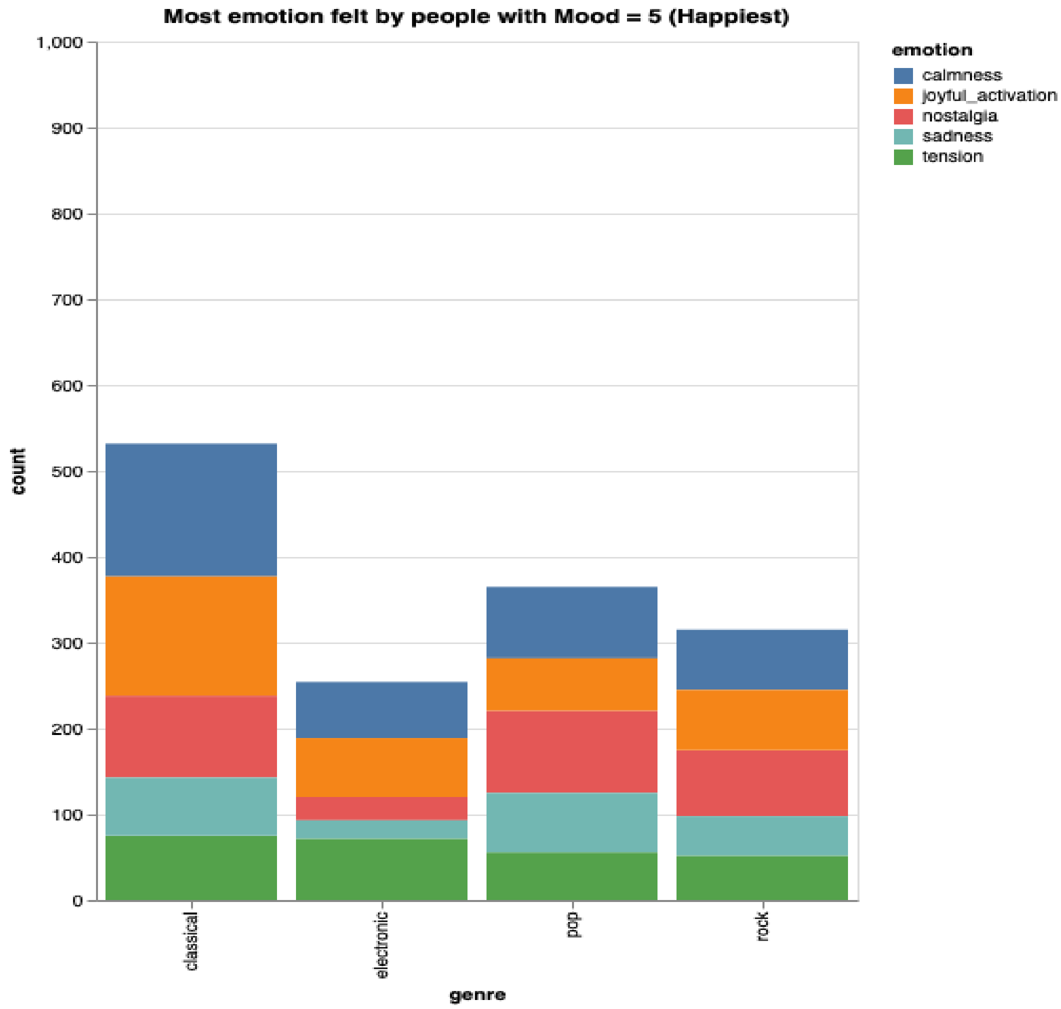
The next visualization we used was a stacked bar chart, we have generated each bar chart for every mood. We wanted to see what emotions are felt for each mood listening to different genres of music. The generated graphs are :



When you are depressed(mood 1), listening to classical music generates a lot of emotions. In electronic and pop music a lot of calmness and joyful activation is seen. When you are sad(mood 2), electronic music can give you a lot of tension, surprisingly the other genres can make you feel nostalgic. Calmness and joyful_activation for all the genres are similar.



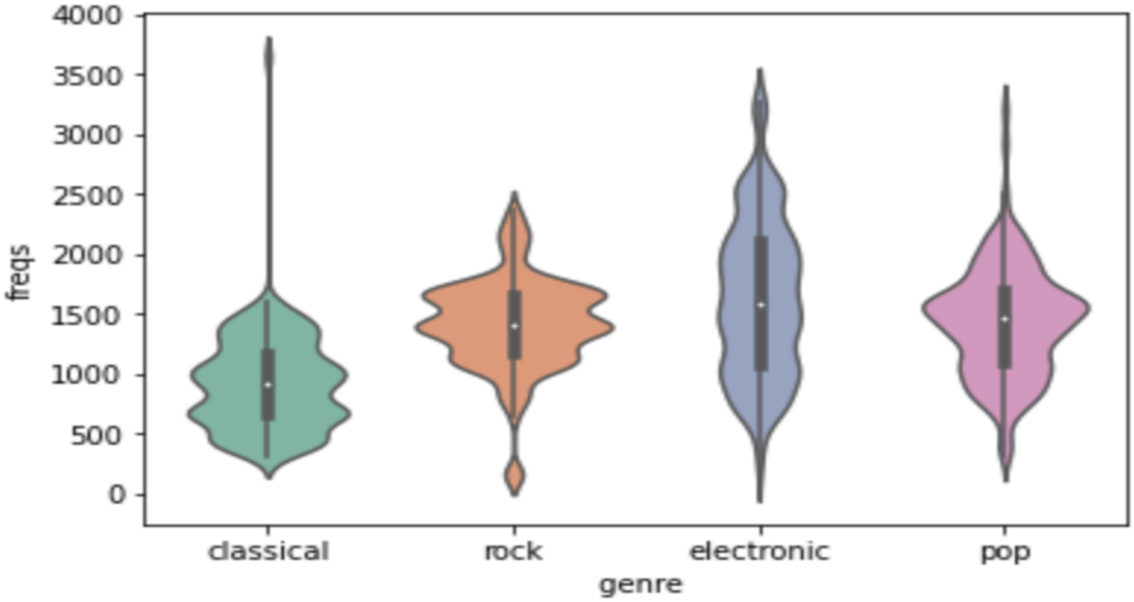
You feel more calm and joyful listening to classical music when you are in a neutral mood. Rock music is generating equal emotions meaning you may remain in the neutral mood. The sadness felt will almost be the same irrespective of the genre you are listening provided if you are in happy mood.(Mood 4). Interesting observation is that the tension felt is high when listening to electronic music.

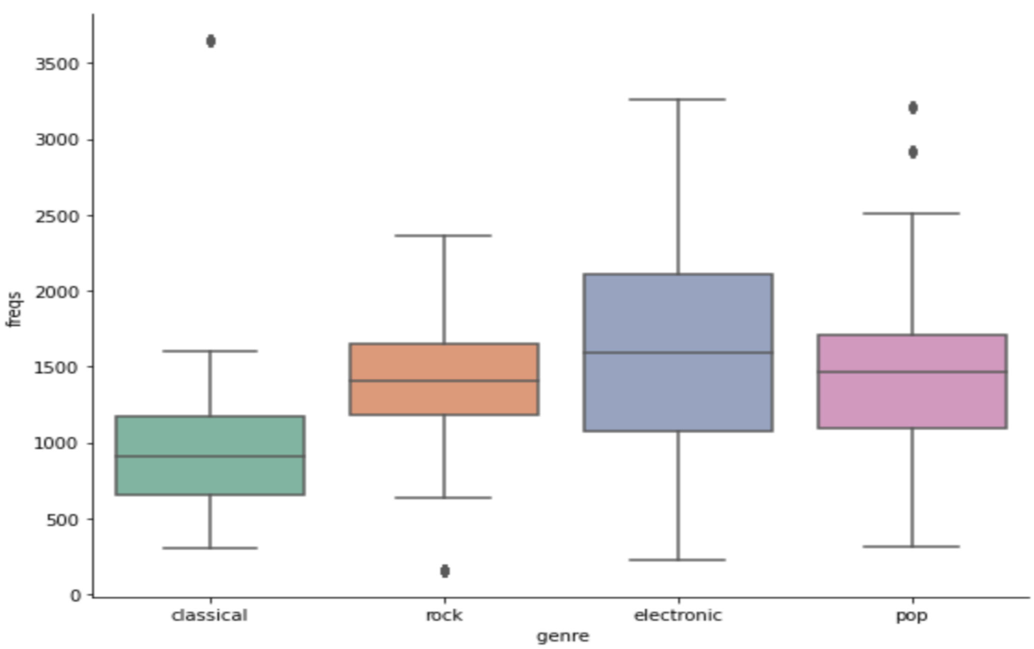


Person in mood 5 is also showing similar emotions to that of a person in Mood 4. The tension observed is lower than before. Sadness tends to be low in all the genres if the person is very happy.

Through these stacked bar charts we learned that calmness and joyful activation are the major emotions felt in all the moods. So, music does generate some sort of calmness in us regardless of your mood.

Using the audio files that we had, we calculated the average frequency values of all the songs used in this experiment. We used boxplot and violin plot as shown below:



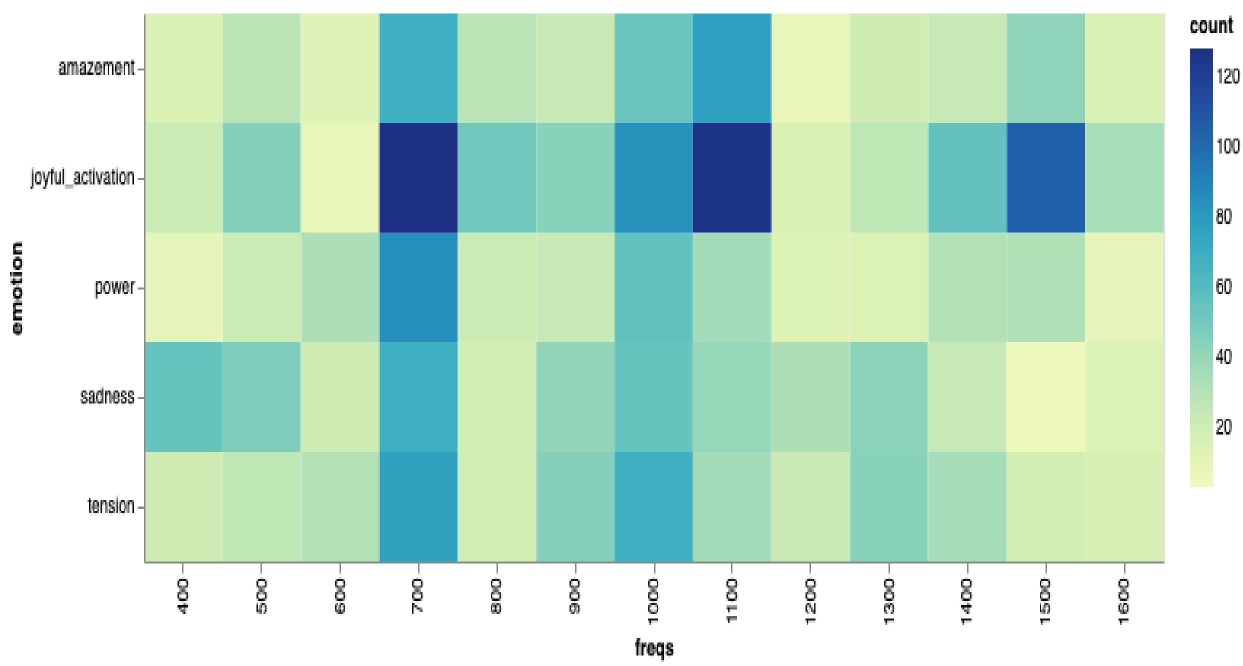


This shows a summary of the frequency values over the four genres. Humans can detect sounds between 20 Hz and 20000 Hz, and in this data, we found a few countable points that were outside this range were omitted. As we see from this graph, there are a few outliers in some genres - one noticeably in the rock genre, and one in classical music. These could be data entry error or it could also be that since we just took the Average frequency values, the song from which this point was calculated could have one outlier (or just one time where the frequency or spectral centroid was too high) and because of this one time, the average frequency for that song is seen as an outlier around 4000 Hz.

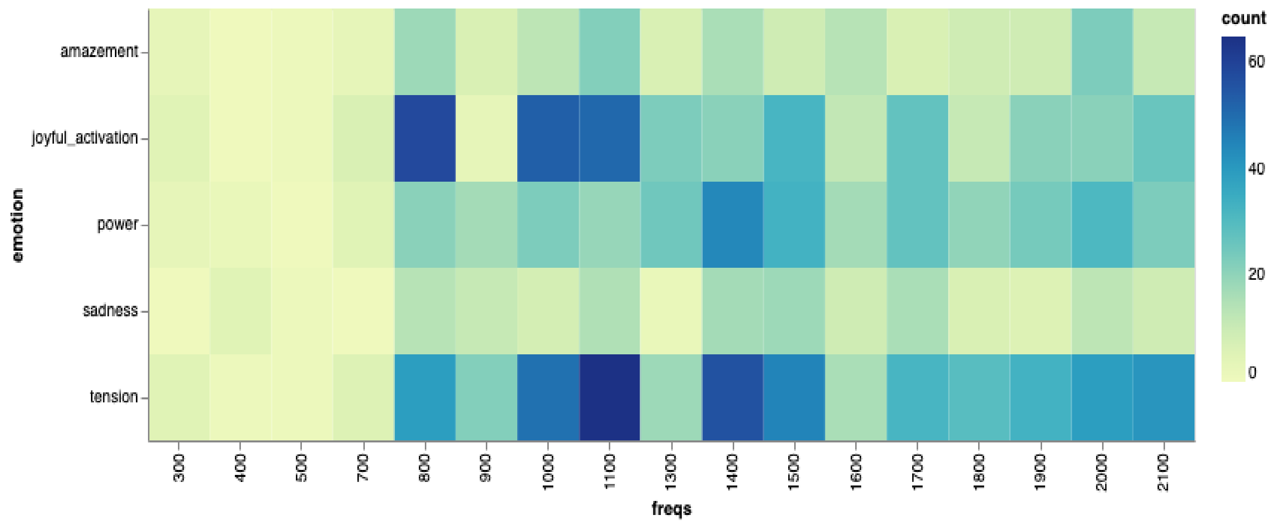
As you mentioned, a box plot can be used to identify outliers in a data set. Outliers are data points that are significantly different from the rest of the data and can indicate errors or anomalies in the data. In a box plot, outliers are typically represented by individual points outside the whiskers of the plot. However, without the actual data and the context in which it was collected, it's difficult to say for sure whether there are outliers in the data or what they might represent. It's important to carefully consider the context of the data and use appropriate methods to analyze and interpret it.

Next, we used heatmaps to explain the different frequencies and how emotion ratings vary by aggregating the number of people who felt a particular emotion at every frequency. The x-axis represents the different frequencies of each genre, and y-axis represents the number of people who felt the emotion.

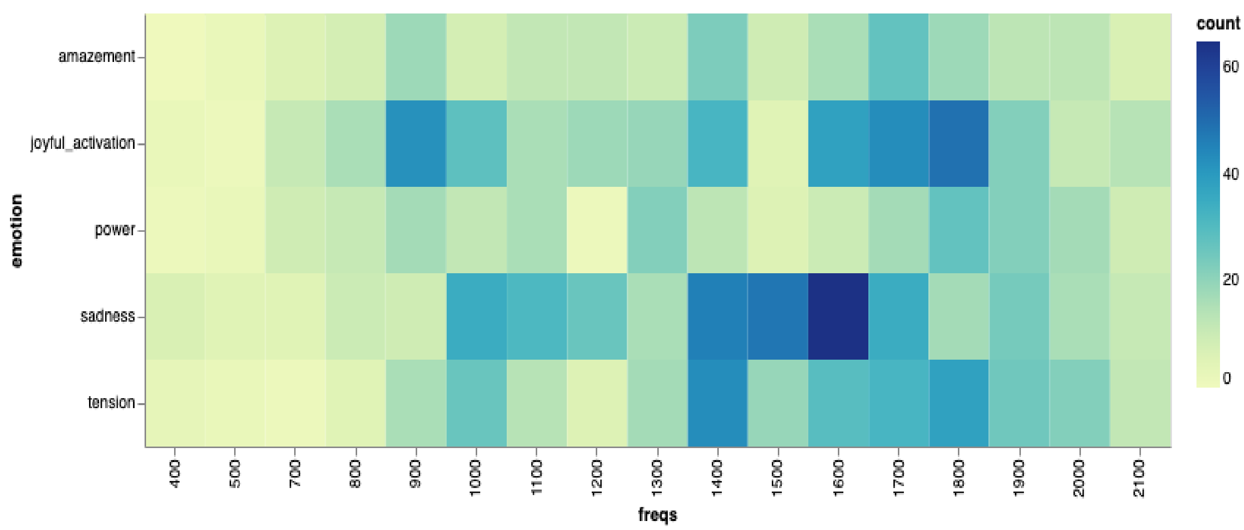
We should note that the scales of these different heatmaps are not the same - this is because we wanted to analyze each genre separately rather than have a common comparison.



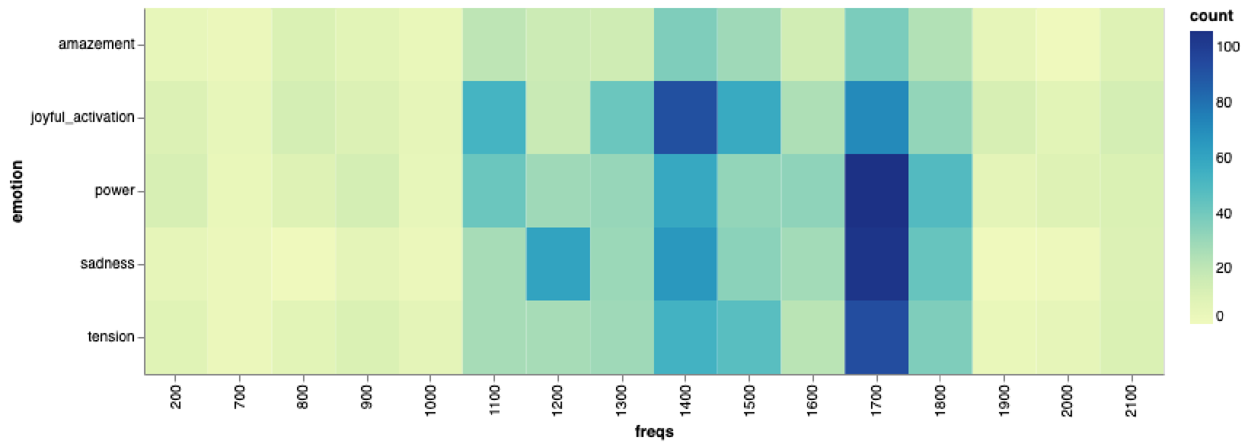
We can notice that at 700 hz, 1100 hz and 1500 hz, there has been significant activations in the joyful_activation emotions. We can also notice that at 700 hz, almost all the emotions are activated. This could mean that at this frequency, since all emotions are felt it is a neutral point or neutral frequency. Although we are not sure why this pattern occurs.



Next we have the heatmap for the electronic genre. It is quite visible that people have recorded values mostly for tension because that is the nature of electronic music, high intensity and high loudness tends to make people feel more tense. But it can also be said that if we like such kinds of music, then we feel happy listening to it, which can be seen from the density in dark boxes around 800 and 1000 frequencies.

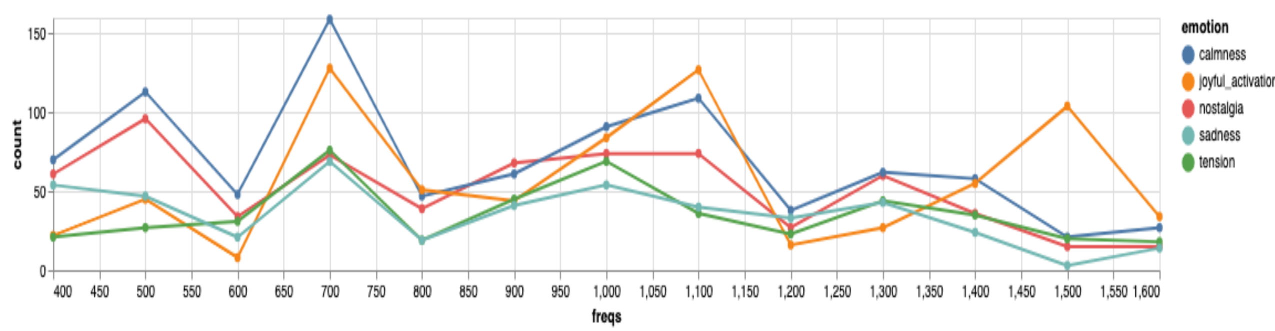


In the above heatmap, which shows a correlation between frequency and emotion ratings, surprisingly pop music makes most of the sampled people sad, especially at 1600 hz. This was unexpected as the most commonly heard genre is pop. And I personally feel happy listening to pop music.



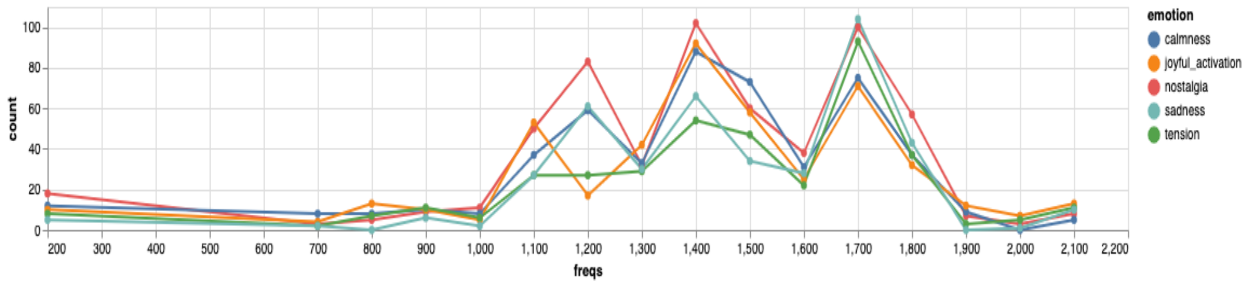
The last heatmap represents the rock genre and the 1700 hz is eye-catching. We see the dark boxes for power, sadness and tension, which resonate with rock music. Some rock songs from Nirvana make you a little sad whereas those from ACDC in particular makes you feel more powerful.

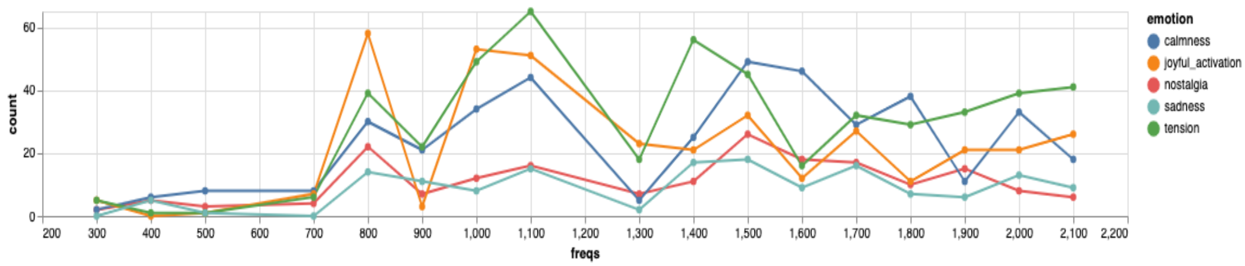
And even though the heatmaps gave a good amount of information, we still wanted a little more and the line charts gave us exactly what we needed.

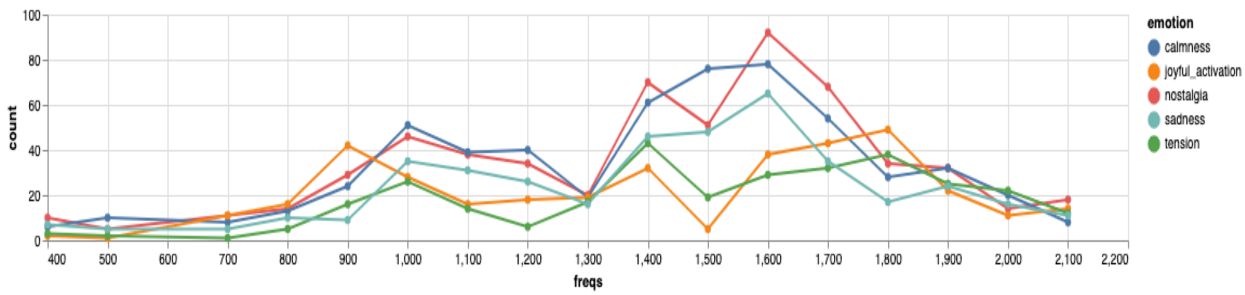


The first line chart is associated with the classical genre, and we see clearly that at 500 hz, there is something weird - people tend to feel both happy and sad at the same time. And at 1200 hz, all the emotion counts tend to converge indicating a neutral frequency at that point. This was one of the information we got only because we used a line chart.

For electronic and all the other genres, the emotion counts start to pick up at 700 hz, but before that everything seems to have no ratings at all. But somehow at 700 hz, things start to kick in and we see reactions from people.







But one most interesting thing to note here is that in pop and rock music, the red lines have higher peaks than the other emotions. This could be because most famous rock songs are from the 90s and from the age distribution histogram, we saw that most people are aged between 25 to 35, and so these people felt more nostalgic when their childhood songs were played to them during the experiment.

Comparison with previous study result:

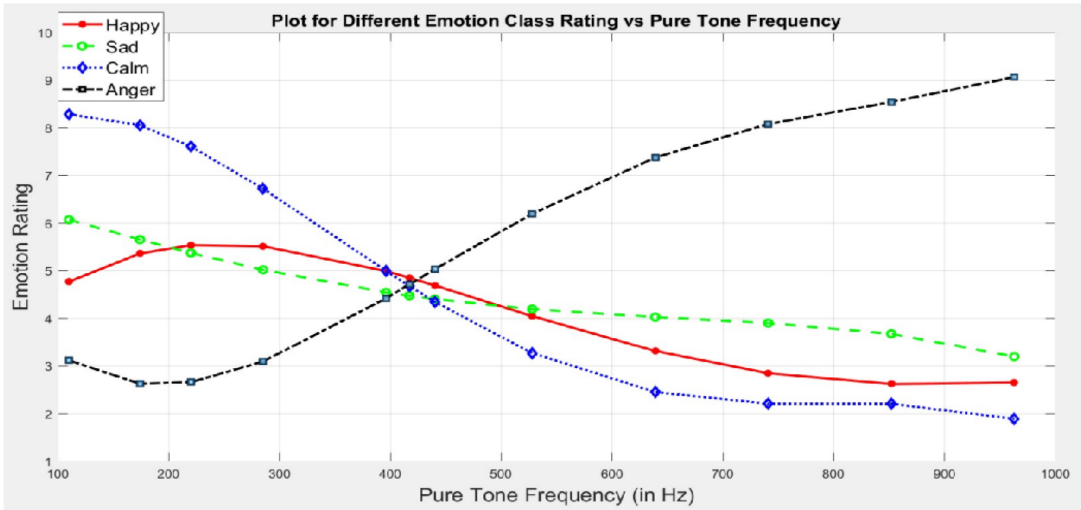
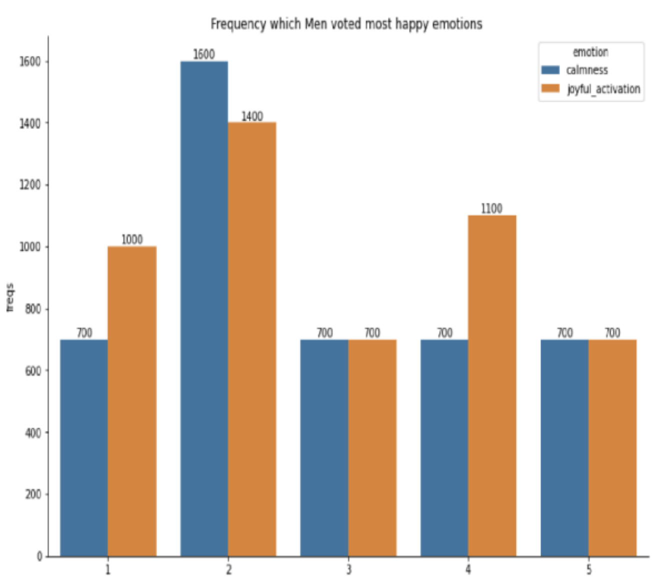
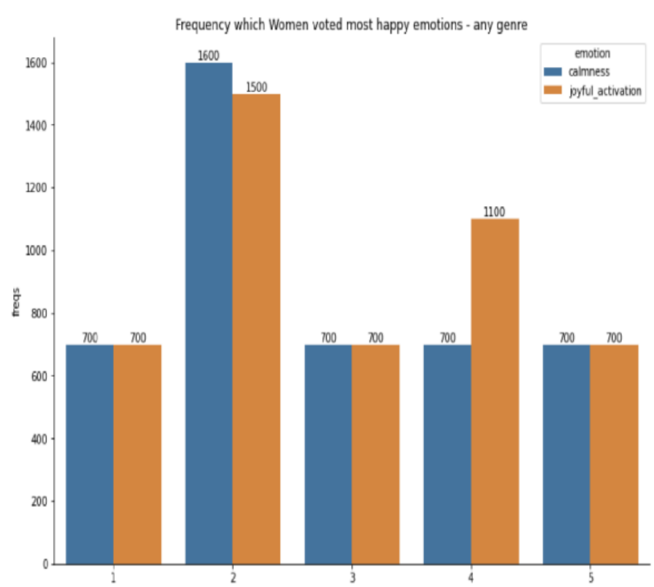
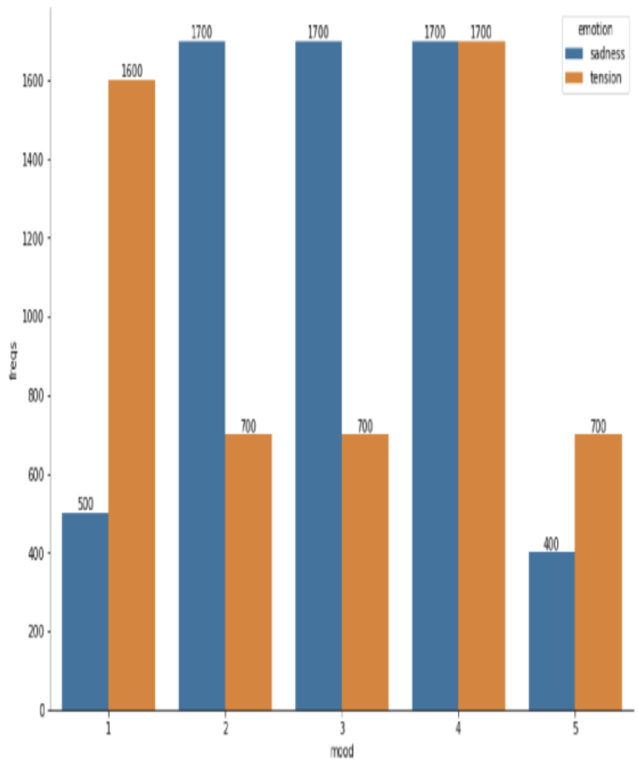


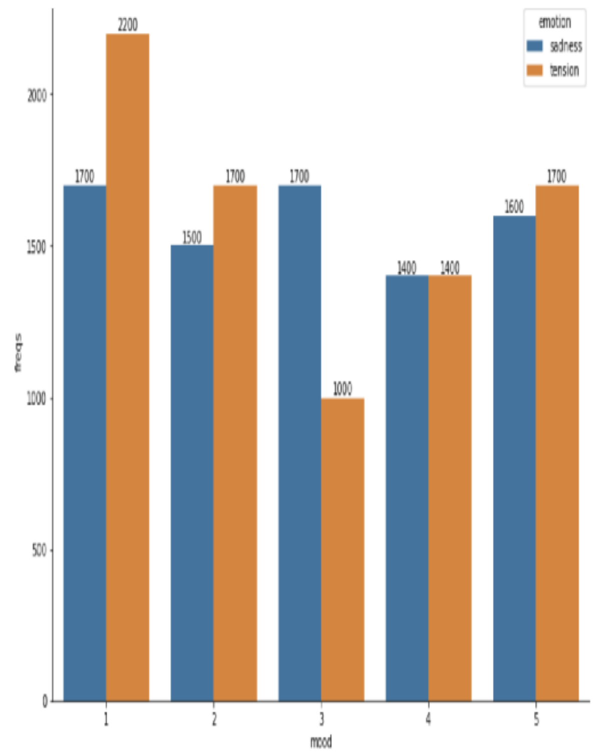
Figure 1. Plot for different emotion class rating vs Pure tone frequency. Here, the line graph for all four primary emotions—Happy, Sad, Anger, and Calm are plotted w.r.t pure tone frequency (in Hz).

From the previous study, the researchers have concluded that 400 to 417 hz is a neutral frequency because all emotions have been felt almost equally. But they have used single-tone music - which is listening to just one frequency at a time. In the real world, there are not a lot of people who listen to single tone music, usually a song consists of combination or mixed frequencies. And our dataset has samples of these songs, and so we have different genres with different behaviors for each.









Conclusion:

Our ultimate goal in this project is, given a person who is in mood X, what frequency Y should they listen to so that they feel the emotion Z.
In the first two bar charts, the x-axis shows each mood category and y-axis shows the frequency at which we can achieve the Z emotion, and it is evident that for both Men and Women with any mood from 1(being most sad) to 5(being most happy) songs with 700 hz average frequency makes them feel serene and happy.
And if for some reason, someone wants to feel more sad or remain sad, higher frequencies like 1600 to 1700 hz is the way to go - as we see from the second two bar charts.

One major disclaimer is that these results are not totally conclusive, as various other factors like the lyrics of the song, environment the person is staying in and many other external latent features should be considered before concluding any result.
Also, the most important thing is that a song can not be quantified using just one frequency value, so we can use some advanced methods like spectrograms of songs instead of just one frequency value, which is what we are planning to do as next steps for this project.

We have learned how to work on different datasets especially with audio format data, and more importantly we understood how to choose different visualization techniques.

Some visualization techniques will work in some cases, and in other cases we might have to think about other approaches. We tried a few charts, and most of them did not give us the specific information that we were looking for. And after trying many different charts, we were able to select the best visualization that conveys the information in the best way. We also learned how to tell the story in a good way - plotting different charts is one part of the project, the other and the most important is how to convey this information to other people, and here we had to learn story-telling techniques, and all these were very interesting to work and learn.

Future work:

The results generated gave us some good insights about relations among genre, frequency and human emotions. In our future work we would like to include other factors like lyrics, instruments used, song release time and artist to get a better understanding of the patterns between music and human emotions.

References:

A Preliminary Investigation on Frequency Dependant Cues for Human Emotions:
https://www.mdpi.com/2624-599X/4/2/28
Uncertain Emotion Discrimination Differences Between Musicians and Non-musicians Is Determined by Fine Structure Association: Hilbert Transform Psychophysics:
https://www.frontiersin.org/articles/10.3389/fnins.2019.00902/full
Human emotions track changes in the acoustic environment:
https://www.pnas.org/doi/10.1073/pnas.1515087112#abstract
https://towardsdatascience.com/get-to-know-audio-feature-extraction-in-python-a499fdaefe42
https://www.kdnuggets.com/2020/02/audio-data-analysis-deep-learning-python-part-1.html
https://www.fesliyanstudios.com/royalty-free-music/downloads-c/sleep-music/56
https://github.com/mdeff/fma
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3882084/
https://magenta.tensorflow.org/datasets/nsynth