

Question 3

Peter Meihuizen^a

^a*Stellenbosch University*

1. Introduction

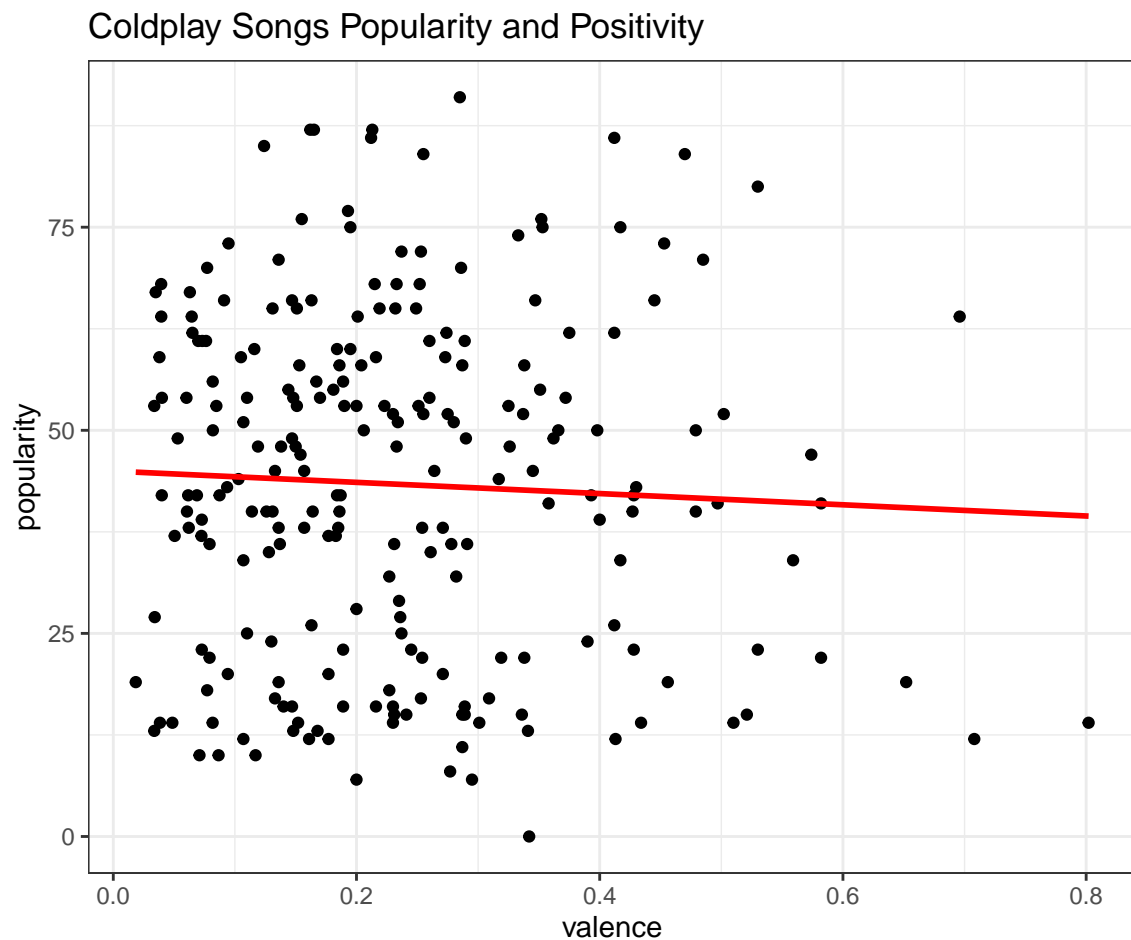
This question shows the popularity of different songs for Metallica and Coldplay. I will try to analyse these different songs to see what makes a song popular.

2. Analysis

I first need to import the data for both Coldplay and Metallica's songs.

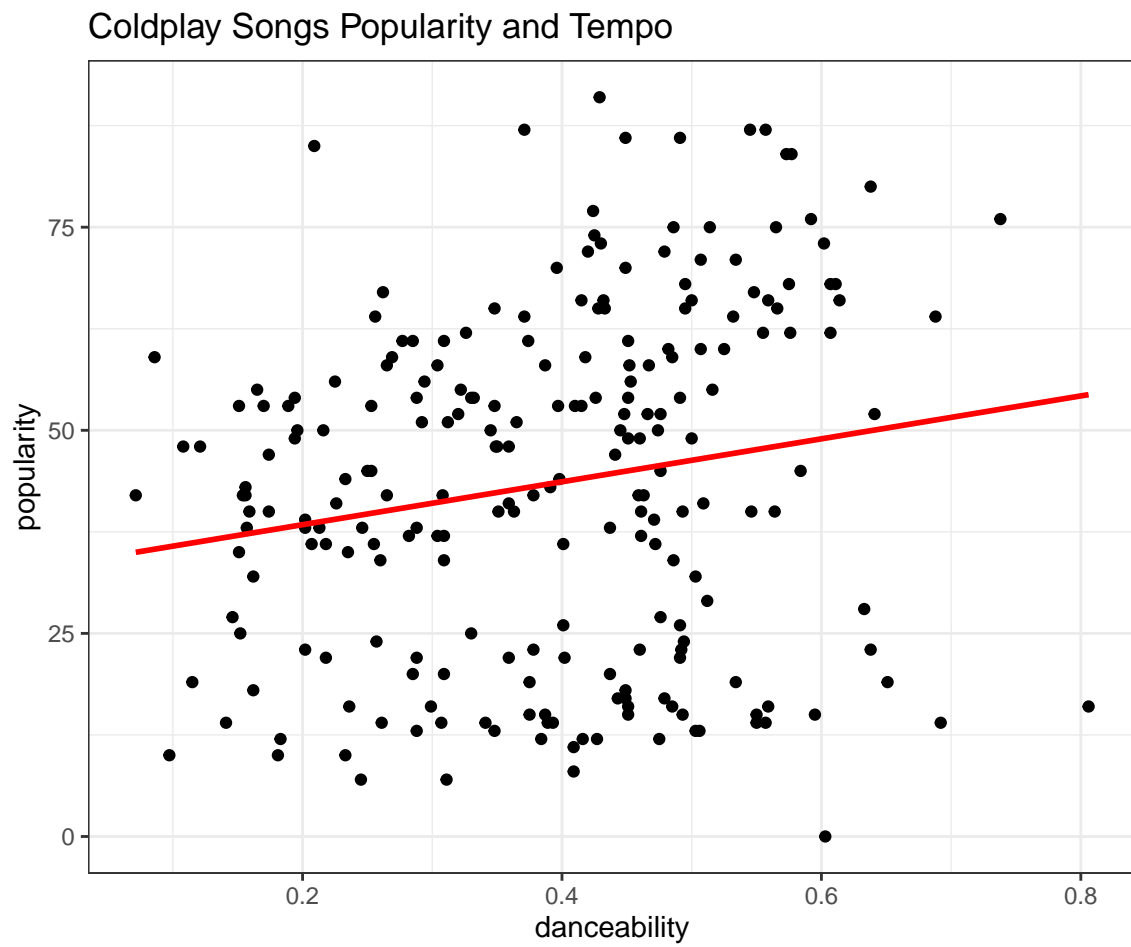
So I want to look at what makes songs popular, so I'm going to do some scatter plots to plot the songs various aspects with their popularity to see how they compare. To do this I'm going to build a function which makes a scatterplot.

Plotting the scatterplot shows as follows:



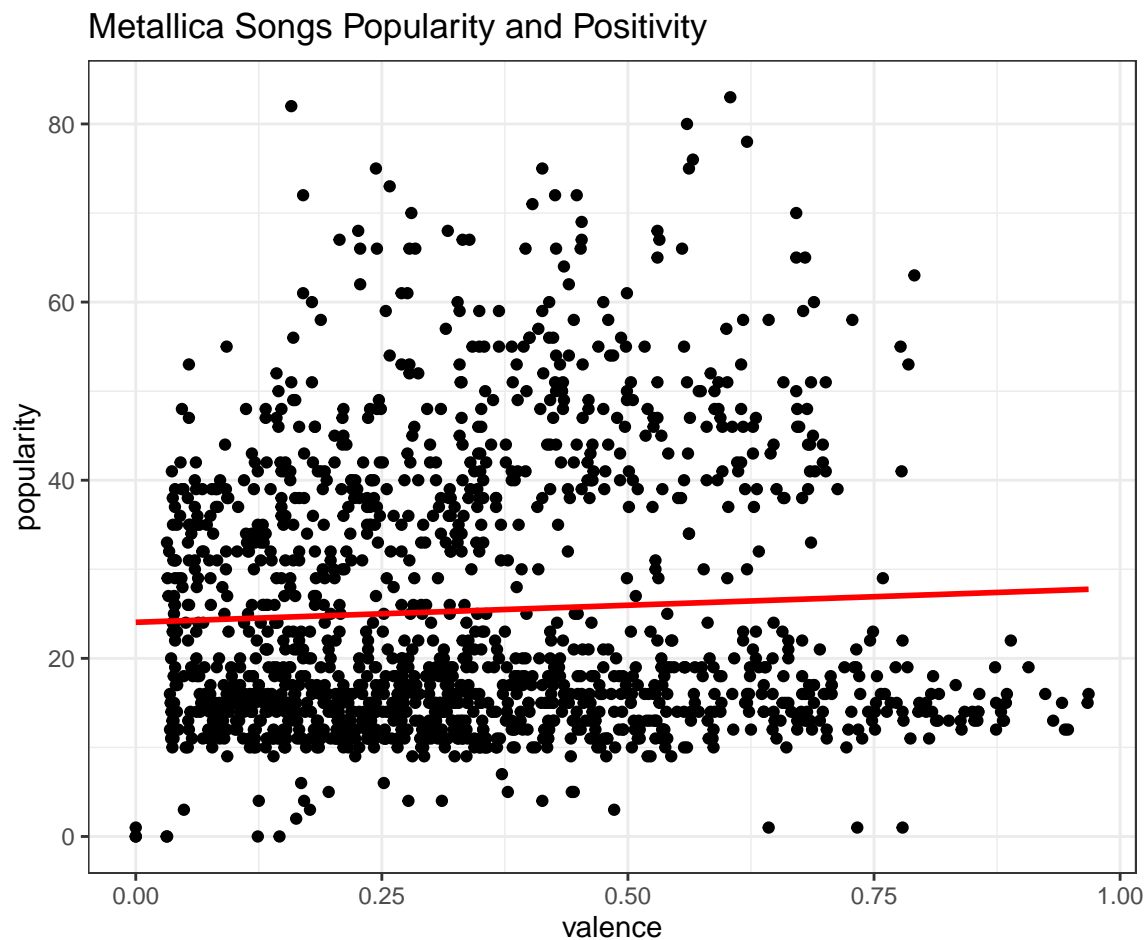
It can be seen that Coldplay's music mostly has a lower valence, meaning it resembles mostly negative emotions, and it appears that there is a slight negative correlation between valence and popularity, suggesting more positive music does slightly worse than more negative music.

Everyone likes to dance right? So what effect does the danceability have on the popularity of their songs?

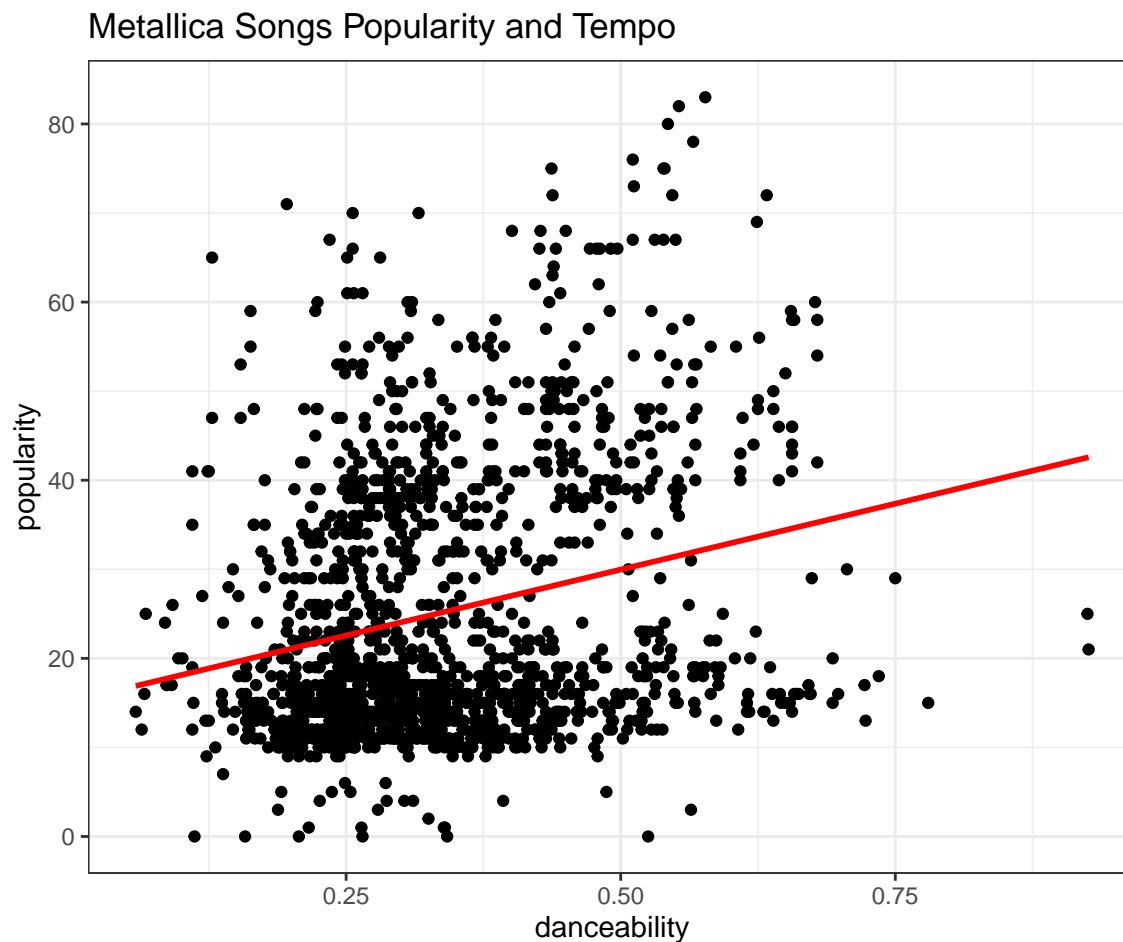


Again it appears that danceability has a positive effect on the popularity of Coldplay's songs.

Ok what about Metallica?



There seems to be minimal correlation between valence and popularity for Metallica's songs. It seems as though their songs are more concentrated around a low valence, suggesting they express negative emotions. However their more positive songs are slightly more popular as is suggested by the very gradual slope of the correlation line.

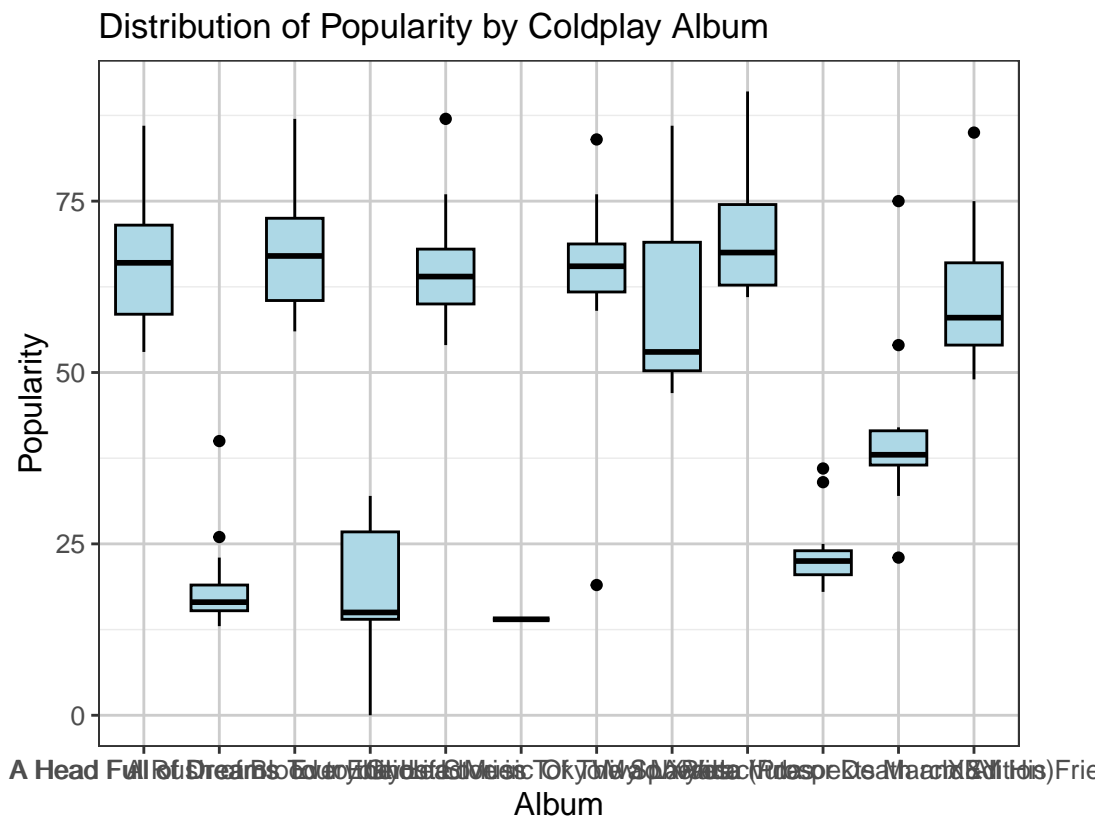


Again, the same as was seen with Coldplay, danceability increases the popularity of Metallica's songs, despite the majority of their songs being lower on the danceability scale.

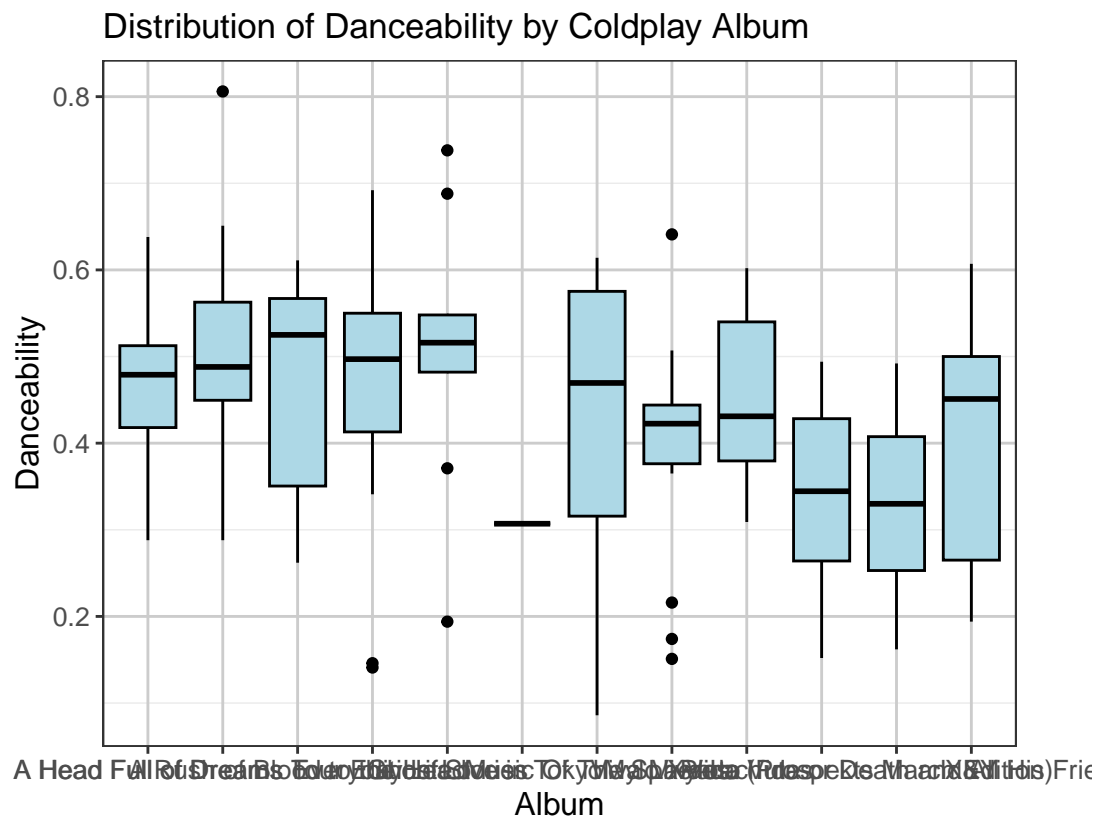
Ok so it seems as though danceability makes a particular bands songs more popular. So let's see how the popularity and danceability was distributed across their various albums and see whether these are linked. In order to do this lets make a function that looks at the distribution of the danceability of the different albums of Coldplay and Metallica.

I'm just going to rename the album_name column in the Coldplay data frame so that it is the same as in the Metallica data set. I am also going to get rid of the live albums and songs.

Now plotting the boxplots, first by popularity:



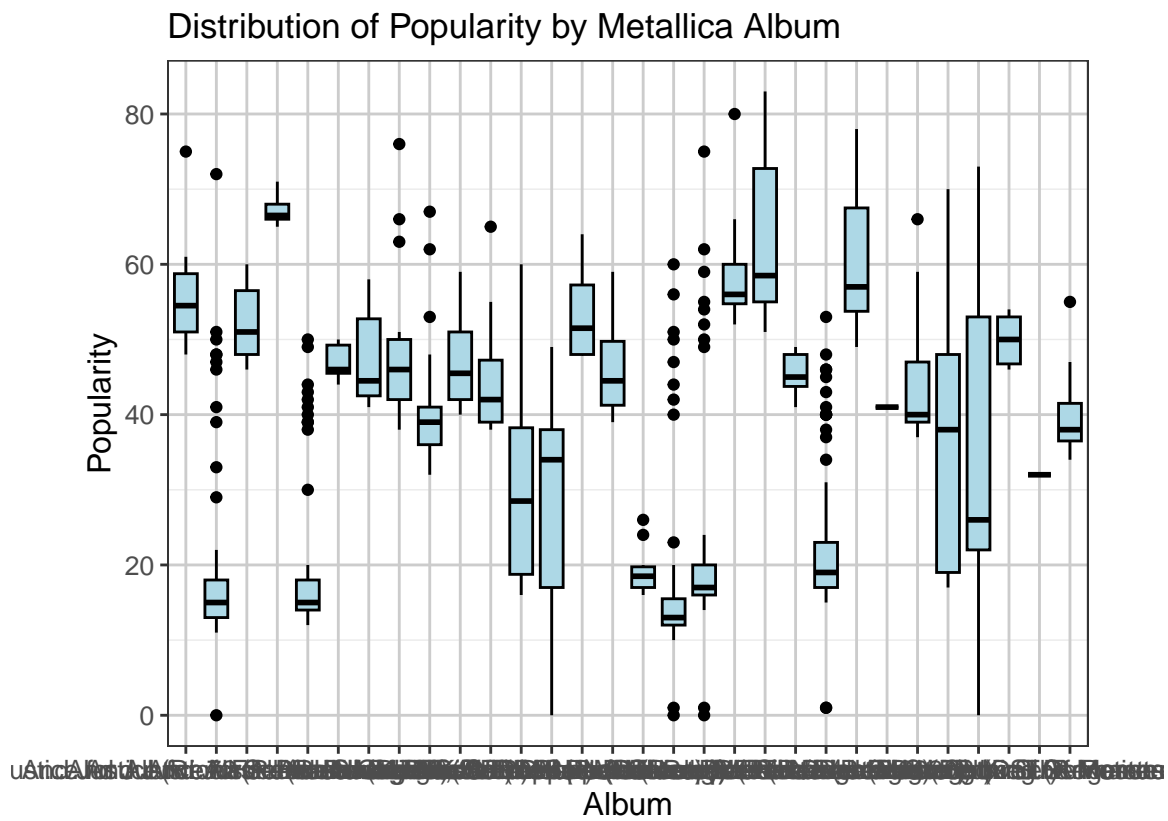
So as can be seen Coldplay's albums vary quite a lot in terms of there popularity, some were significantly more popular than others. What about the daceability of their albums?



It appears as though the distribution of danceable songs on Coldplay's albums have been pretty similar across albums.

How about Metallica now. First I need to remove the live songs and albums.

Now to look at the popularity and danceability of the albums. First popularity.



Ok similar to Coldplay there are some differences in popularity between Matallica's albums. Now let's look at their danceability.

