Andrew Ferruzza 111616974

CSE 332 - Lab 5 Report

Prof. Klaus Mueller

**Lab 5**

In the fifth and final lab for this semester, I have pulled together four graphs that accurately sum up my dataset. This is the final visualization for the *Top Spotify Songs of the 2010s* that I will be doing using Python and handy libraries such as Dash for the webpage and Plotly for the graphing portion. Two of these graphs, Graph #3 and Graph #4, use advanced visualization techniques--filtering and highlighting, respectively. Let’s take a look.

**Graph #1 - Log Scaled Histogram**

The first graph displays a log scaled y-axis histogram which maps the duration of all of the songs in this dataset. I log scaled this one to bring out the shape of the graph more, and it turns out to look almost normal or a combination of normal and skewed right. I did predict early on in my first lab report that most songs would fall within the 3-4 minute range (180-240 seconds), as this tends to be the duration for most of the radio-friendly, popular songs. Turns out this is true, with a few outliers floating in the 5-7 minute range.

**Graph #2 - Pie Chart for Top Genres**

Taking a look at all of the genres in this dataset, this pie chart conveniently displays the hierarchy we have going on. The genre “dance-pop” took up more than half of the entire dataset, which didn’t surprise me that much, considering in my first lab report I predicted this would be the highest charting genre. If you listened to any radio music in the 2010s (especially the first leg of the decade), it wouldn’t be too hard to agree with me there…

**Graph #3 - Filtering With Years and Popularity**

On this next graph, a scatter plot, I decided to use a slider as a filtering method to select a year from the decade. The graph displays all the songs ordered by index, so selecting a year will just limit you to the songs from that respective. This is then mapped against a color-coordinated “popularity” y-axis, where the brighter colors represent the more popular songs of that year. I found this was interesting to note that there was a decent range in popularity (how well people know the songs) for the Spotify songs. Sometimes songs might chart in Spotify, but not everyone knows about them. This might be in part due to more of the younger generations using Spotify.

**Graph #4 - Highlighting with Years and Energy**

The final graph is a line chart that shows the energy of all of these songs as the decade went on. Energy is defined by the database owner as “how likely it is to make people dance.” Think songs that bring good feelings and vibrations. Ultimately, it looks like the energy of these songs steadily decreased as time went on, which I actually find to be quite interesting. Maybe as the decade progressed, people became more open to more sad or introspective music? Dance-pop, for example, still dominated the second half of the decade, but there are some other genres that get thrown into the mix at that point as well. Interesting find here.

**Comparison with Lab 1 Hypotheses**

In the end, all of my Lab 1 Hypotheses came out to be true. But along the way I was able to pick up some interesting finds, like Graph #4 here for example, and by using the advanced techniques we started touching on in the class around the time of Lab 3. So, as I also predicted in Lab 1, there were trends that I knew I wasn’t going to be able to see at first glance until I did a visual analysis. And that’s why visualization is very important!