Assignment: Billboard Hits and Spotify Music Data

Questions I: Dual-Axis and Shared Axis Charts

1. Let's start by taking a look at how representation of two particular genres have changed over time on the Billboard charts: Pop and Rock. This question will only require the **BillboardHot100** data source.

   a. If you preview the Genre field for the Billboard data, you'll find that most songs have multiple genres together. To isolate individual genres, create two calculated fields *Is Pop* and *Is Rock* that returns whether or not each song's Genre contains the words 'Pop' or 'Rock'. (Aside: This means that variants of each genre will be counted as part of the main genre; 'Pop-Rock' songs will be counted as both.)

   b. Use the fields from part a to calculate additional fields *Prop.Pop* and *Prop.Rock* so that when we aggregate over songs, they return the proportion of the songs that fit the corresponding genre. To do this, ensure that your *Is Pop* and *Is Rock* values are 0s and 1s, and take the average of *Is Pop/Rock* in your calculated field.

   c. On Sheet 1, create a line chart showing the proportion of Pop and Rock songs in the Top 100, aggregated by month. There should be two lines: one for each genre. NOTE: The chart weeks are stored in the *Billboard Week* field.

   d. Interpret the chart in two to three sentences. Are there changes in listeners' preferences over the past two decades in these genres? Is one genre more consistently popular than the other?

Questions II: Interactive Visualizations

2. In order to look at how other genres have changed over time, we would probably rather not create an additional *Is <Genre>* and *Prop.<Genre>* field for each one. Instead, we should use a parameter that dictates which genre is plotted in a chart. This question will only require the **BillboardHot100** data source.

   a. Create a string parameter *p.Genre* that allows for selection between the following genres: Alternative, Country, Dance, Electro, Pop, R&B, Rap, and Rock.

   b. Use that parameter to create a calculated field *Is Genre* that returns whether or not a song's Genre includes the selected parameter value. Tip: This is identical to Question 1a but instead of searching for "Pop" or "Rock", it should look at the value of the parameter instead.

c. Use *Is Genre* to calculate an additional field *Prop.Genre* that returns the average number of Billboard songs with the selected *p.Genre* string. Tip: This is identical to Question 1b.

d. On Sheet 2, create a line chart that shows the proportion of songs matching the parameter Genre in the top 100, aggregated by month.

e. Let's make the plot more informative:

    i. Make sure the *p.Genre* card is visible for genre selection.

    ii. Revise the vertical axis so that units are in percentages.

    iii. Rename the axes titles and chart title. Make sure that the chart title changes its name to match the genre being plotted!

f. Explore the chart you've created. Of the eight genres that have been included in the parameter, which ones have seen a general increase in presence in the top 100 from the 2000s to the 2010s?

3. Finally, let's explore the relationship between pairs of Spotify features. Normally, we might use a scatter plot, but since there are so many data points, we'll use a density map instead so we can see the distribution of data points. This question will only require the **Spotify songAttributes** data source.

a. Create parameters *p.Audio Feature X* and *p.Audio Feature Y* that allow for selection between the following Spotify audio features: Acousticness, Danceability, Energy, Instrumentalness, Liveness, Speechiness, and Valence. Tip: Create the parameter once, then make a copy.

b. Create fields *Audio Feature X* and *Audio Feature Y* that take on values of the matching parameter. HINT: You'll need a logical expression (e.g. IF, CASE) to select the right values. If you are unsure on how to do this, rewatch the "Dynamically Displaying Dimensions" video.

c. On Sheet 3, create a scatter plot for all songs (not just Top 100 songs) using the fields you just created. Convert this into a density map by changing the Marks type to Density. Show the audio feature parameters.

d. Based on a visual inspection, are songs that are more positive (high *Valence*) more danceable? How would you describe the strength of this relationship?

Visualization Reference

You can check your work on the SkillBuilder assignment by referencing the images below. Each image is a general depiction of what you should see on each Sheet of the workbook, with various details hidden or omitted. **Your visualizations do not need to look exactly like these reference images!** These should just be a general guide to check if you're on the right track.

Sheet 1 (Question 1c)

Month of Billboard Week
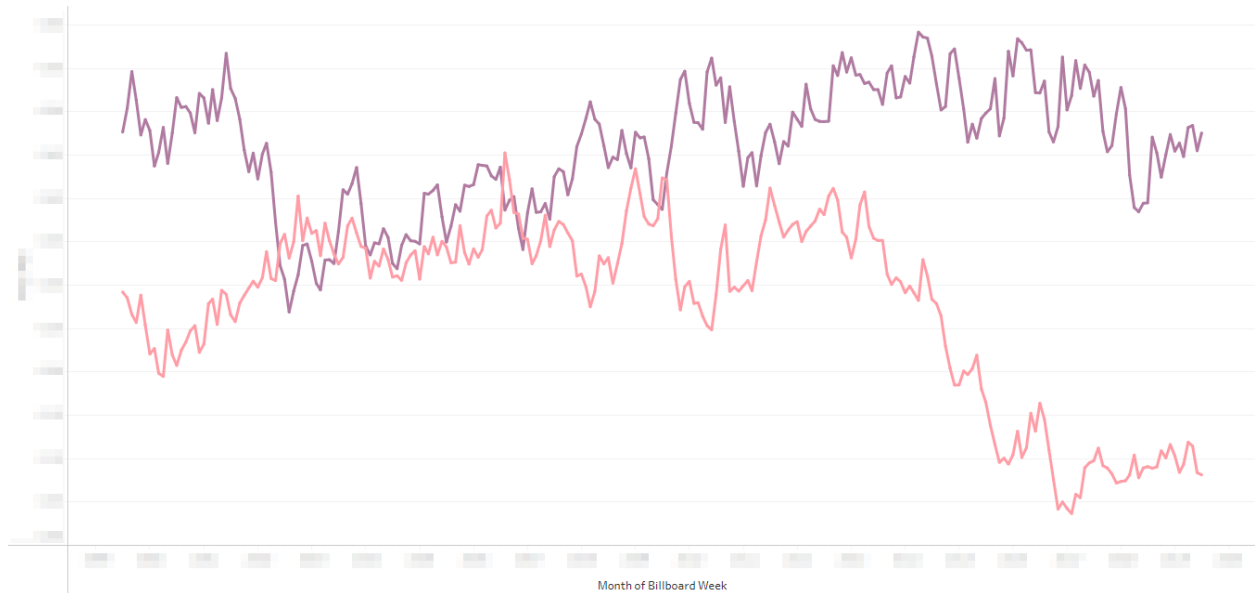
Sheet 2 (Question 2)

The chart below is plotted for the Alternative genre.
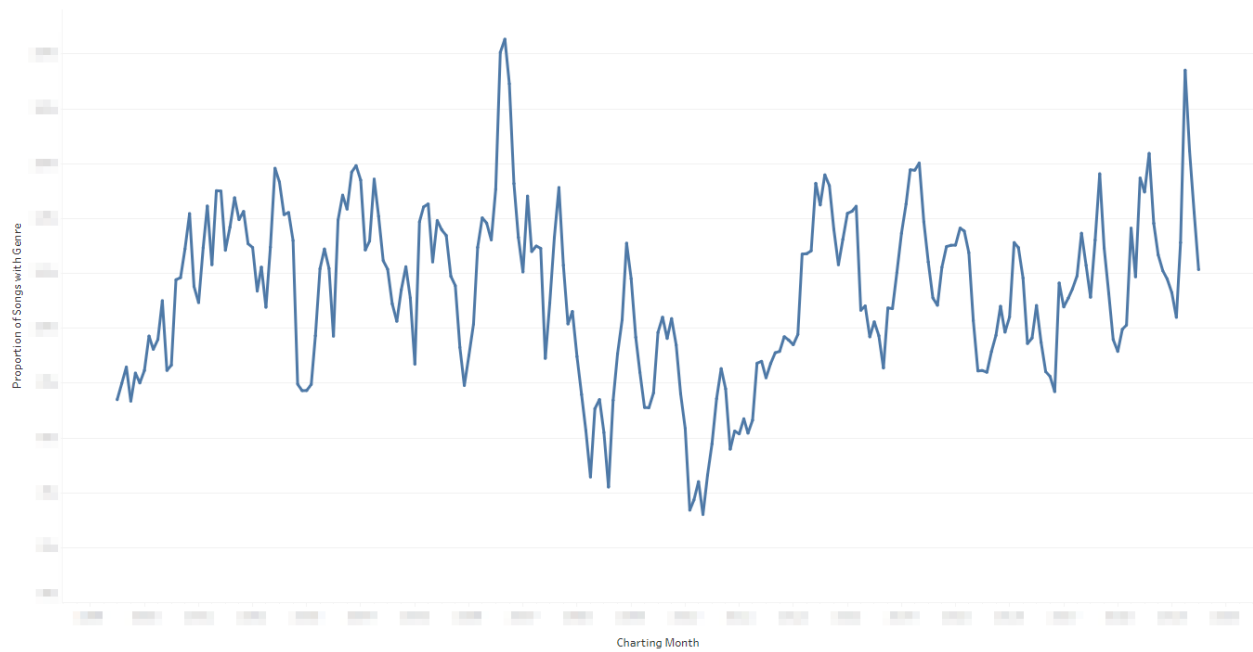
Sheet 3 (Question 3)

The chart below is plotted for the Liveness (on Y) and Energy (on X) features.