Data Visualization Project (Group 4)

Franz

Goal: We want to know which genre of music has more artists in each decade in the 1950s, 1960s, 1970s, 1980s, 1990s, 2000s and 2010s. And also give the option to the user to choose a particular country and see the graphic for that country.

List of attributes:

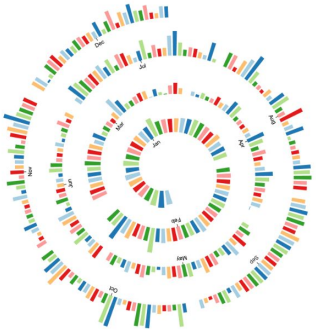
- genres. Ex. ["Alternative Rock","Hip Hop"]
- lifeSpan.begin. Ex. '1984-04-27'
- lifeSpan.end. Ex. '2010-04-18'
- locationInfo. Ex. ["United States","California","Oakland"]

Data Transformation:

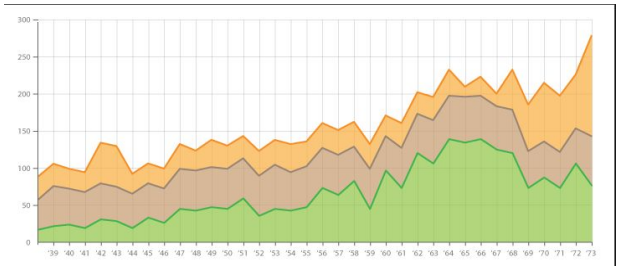
- genres, since we have a list of genres for each artist we are going to select the one with index 0. And we are going to plot a histogram of this variable so we can choose the 10 genres with higher frequencies and the eleven should be others.
- lifeSpan.begin, lifeSpan.end, in these 2 attributes we are going to split the String and get only the beginning year and end year of each artist. With that we can get the artist's lifetime.
- locationInfo, we are going to select only the country of each artist, that's the index 0 of the list locationInfo.

Visualization technique:

- Option 1. Spiral graph
  Each color represents one type of genre and the period of time is a decade.



- Option 2. Stacked area graph



Jasmine

Goal

We want to know which platform is highly used among various platforms and further divide them into genres.

List of attributes :

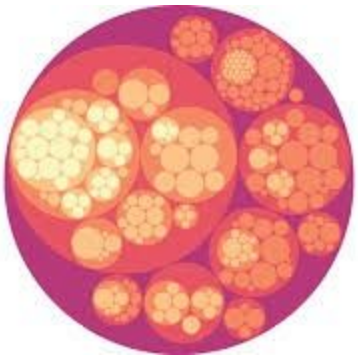
- Url amazon
- Url Spotify
- Url iTunes
- Url deezers
- Url AllMusic
- Url Discogs
- Genre

Data Transformation :

- We calculate the length of not null values for each platform and use them to visualise the frequency of each platform as circles.
- For each platform we consider the frequencies(value counts) for each genre in the dataset.
- We can show the frequencies of the division on hover.

Visualisation Technique :

Circle Packing : Representing the platforms inside the dataset and also the genre inside the platform circles.



Jean-Elisée

Goal

The goal of our visualisation is to present the most common language in the music across the countries and organised by release date.
Exemple: Showing that in France the most common language in June 2000 is French with more that 10000 songs released by different artists.

List of attributes required for the visualisation

The list of data that we need to perform the visualisation are

- The country
- The language
- The release date (To add an additional temporal dimension to the visualisation)

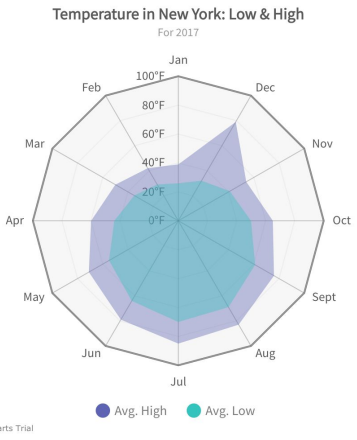
Data transformation

1. The wasabi contain multiple data which are relevant to different countries. Here the first goal will to group all the data according to the country
2. The result of the previous operation will then be grouped according to the year and month of release
3. For each subgroup create we will now calculate the frequency of usage of each language

Visualisation technique

The different options that can be used to represent the visualisation are:

- Graph char



t

There will be two informations that the user will have to select

- The country
- The year

Once both are selected the user will be displayed the graph showing the frequency of each language represented by different colours.

Limitation: The difficulty here will be to represent the temporal data.
Potential solution: Create a dropdown showing the list of countries and once the country selected the graph can be updated and displayed.

- Timeline

A representation country wise of the usage of languages in the music across the time.