UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

LEANDRO SOARES GUEDES

MusicVis: Interactive Visualization Tool For Exploring Music Rankings

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Advisor: Carla Maria Dal Sasso Freitas

Porto Alegre
January 2017

CIP — CATALOGING-IN-PUBLICATION

Guedes, Leandro Soares

MusicVis: Interactive Visualization Tool For Exploring Music Rankings / Leandro Soares Guedes. – Porto Alegre: PPGC da UFRGS, 2017.

68 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2017. Advisor: Carla Maria Dal Sasso Freitas.

1. Music data visualization. 2. Music rankings. 3. Music charts. 4. Interactive visualization. I. Freitas, Carla Maria Dal Sasso. II. Título.

"When we are not sure, we are alive."

— GRAHAM GREENE

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Prof. Carla Freitas. Her help and support during this research were essential. Her door was always open and I felt truly comfortable with her all the time. I had the feeling that everything was going to be all right.

A special recognition for my family. During these last years I learned how to value and appreciate them. They are everything for me and I try my best to cheer them up. Nara and Gilmar Guedes, my parents, I love you!

I found incredible friends and classmates at UFRGS. Thank you so much for your support and for our long conversations. This include our out-of-lab celebrations. Special thanks to Bernardo Henz, Bruno Pagno, Cleóbulo Neto, Daniele Fernandes, Gerson Groth, Guilherme Oliveira, Hernandi Krammes, Rodrigo Moni and Victor Oliveira.

It is inevitable to recognize how my experience as a teacher at IFSUL made me a better person and professional. My colleagues and my students made my days constantly better. Special thanks to Calebe Micael, Gustavo Igansi, Louize Leitzke, Maurício Escobar, Rodolfo Favaretto and Vinicius Borba. Also want to thank the 2016 INF-4AM class for choosing me as paranymph, INF-3AT for being such a great students, my four research students and Fátima Andrade for being my "foster mother" in Charqueadas.

Jean Medeiros, Lucas Maraschin, Roberta Brizola, Rodrigo Dal Forno and Vanessa Bugs, thank you so much for your long-term friendship. You are essential for my life.

I could not forget to thank my friends Alan Carvalho, Analu de Souza, Eduardo Pegoraro, Eduardo Peixoto, Guilherme Moimaz, Heitor Almeida, John Santos, Joyce Bastos, Leandro Villela, Luan Lotti, Lucyana Vulcão and Marcio Luza. Thank you for your friendship, you are important to me.

I also want to thank all the volunteers involved in this project, they had a valuable contribution to this work.

ABSTRACT

Music rankings are mainly aimed at marketing purposes but also help users in discovering new music as well as comparing songs, artists, albums, etc. This work presents an interactive way to visualize, find and compare music rankings using different techniques, including the display of music attributes. The technique was conceived after a remote survey we conducted to collect data about how people choose music. Our visualization makes easier to obtain information about artists and tracks, and also to compare the data gathered from the two major music rankings, namely Billboard and Spotify. The tool also provides interaction with personal data. The results obtained from experiments with potential users showed that the tool was considered interesting, with an attractive layout. Compared to traditional music ranking tools users preferred ours, but with not such a large difference from using Billboard or Spotify. However, when evaluating the usability of our tool, results are positive, mainly concerning to data filtering and comparison features. MusicVis was also considered easy to learn.

Keywords: Music data visualization. Music rankings. Music charts. Interactive visualization.

MusicVis: Ferramenta de Visualização Interativa para Explorar Rankings Musicais

RESUMO

Os rankings musicais destinam-se principalmente a fins de marketing, mas também ajudam os usuários a descobrir novas músicas, bem como a comparar artistas, álbuns, etc. Este trabalho apresenta uma ferramenta interativa para visualizar, encontrar e comparar rankings musicais usando diferentes técnicas além de exibir atributos das músicas. A técnica foi concebida após uma pesquisa remota que coletou dados sobre como as pessoas escolhem música. As técnicas de visualização tornam mais fácil obter informações sobre artistas e faixas, e também comparar os dados obtidos a partir dos dois principais rankings de música, Billboard e Spotify. A ferrament também permite a interação com dados pessoais. Resultados de experimentos conduzidos com usuários potenciais mostraram que a ferramenta foi considerada interessante, com um layout atrativo. Comparando com as formas tradicionais de visualizar rankings de músicas, usuários preferiram a ferramenta aqui desenvolvida, mas a diferença para Billboard e Spotify não foi grande. Entretanto, quando avaliada a usabilidade da ferramenta, os resultados foram melhores, principalmente no que se refere à filtragem e às técnicas de comparação. MusicVis foi também considerado fácil de aprender.

Palavras-chave: Visualização de Dados Musicais, Rankings de Músicas, Visualização Interativa.

LIST OF ABBREVIATIONS AND ACRONYMS

CS	Computer Science
GUI	Graphical User Interface
HCI	Human-Computer Interaction
MIR	Music Information Retrieval
NLT	Node-Link Tree
UX	User Experience

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

People listen to music everyday, some of them even all day long. Music became a huge industry, with several artists and groups competing for popularity and recognition, which is likely to result in earnings. The more fans they conquer, the more influence they have.

Due to the worldwide internet access, the way people listen to music is changing. Some years ago, the success of a certain artist was mainly calculated by how many LPs or CDs were sold, which we call physical music sales. Nowadays, the main way of listening to music is using online streaming, like websites/players such as Youtube (YOUTUBE, 2017) and Spotify (SPOTIFY, 2017a). In fact, Liikkanen and Åman (LIIKKANEN; ÅMAN, 2015) found out that among on-demand music services, Spotify and YouTube, are the most popular ones.

In general, rankings of many different things were always available and have been used to influence users' choices; music rankings (also called charts) would not be different. TV music channels, such as MTV and VH1, have most of their schedule based on music rankings: they show what most people want to see.

The Billboard (BILLBOARD, 2017) magazine produces the most famous music ranking, the "Hot 100" list, which shows the most played tracks (usually called singles, music that is being released on the media) based actually by streaming activity, radio airplay and sales data (respectively audience impressions measured and sales data compiled by Nielsen Music (COMPANY, 2017)). Spotify also produces rankings, which are based on users' streams, and can be filtered by location, daily or weekly. The data is available at Spotify Charts (SPOTIFY, 2017b).

These popular rankings reflect the marketing strategy of record labels. When data are easier to observe and compare, new strategies can be planned and put into practice, contributing to improve marketing and music quality.

Music rankings visualizations can help the analysis of data about artists and record labels, and also work as recommendation systems: users can use visualizations to compare and classify artists' information. For example, if the user prefers pop music, it is likely to be easier finding another pop music only looking at an interactive visualization. Regarding recommendation, a tool can analyze what the user is listening to and recommend other similar artists.

There are several works dealing with music visualization and analysis, but only a few are about music rankings. So, our motivation was the possibility of providing new ways of exploring music and artists data through visualization techniques. We also aimed at exploring ways of displaying personal data about music preferences.

We established as our research goal to develop a music ranking visualization tool that allows users to accomplish exploratory tasks over music rankings data sets.

1.1 Requirements and Overview

Music Rankings and Personal data are influenced by several factors, from huge companies to viral social network medias. To help analyzing these data, we built a tool called MusicVis.

Finding appropriate approaches to explore music data though visualizations, lead us to establish some requirements, which were confirmed by a remote survey we did prior to the tool development:

- R1: build visualizations that fit music data without the traditional tabular form;
- R2: allow users to have full control of what data, ranking and visualization they are interacting with;
- R3: filter data with specific keywords;
- R4: the ability of listen to any music straight through the tool;
- R5: compare the performance and position of any artist, track and genre;
- R6: compare personal data with the traditional rankings.

Using these requirements, we designed a tool to support music users, companies and artists in the analysis of music rankings, from traditional to personal ones. The result was a web-based tool that provides:

- Sunburst, Node-Link Tree, Treemap and Bubble visualizations;
- A Filter Section so the user can search for specific tracks, artists and music genres;
- A Comparison Section, for allowing the comparison of artists, tracks and genres in a specific week or month;
- Personal data loaded from Last.fm to compare with the current rankings from Billboard and Spotify.

1.2 Document Organization

This thesis is organized as follows. Chapter 2 reviews related works that helped us building ours. The Chapter also reviews some works on aesthetics, perception and accessibility, rankings and similarity visualizations, music rankings and personal data as well as mood and genre classification.

In Chapter 3, we firstly introduce the results of a remote user survey that we developed for requirement analysis. Then, we explain how we have chosen the current design and which data sets we decided to use. We proceed describing the MusicVis, mainly the implemented visualization techniques. The interactive features and the user interface provided by our web application is described in Chapter 4. In Chapter 5, we describe and discuss the experiment conducted with potential users, while in Chapter 6 we evaluate our work and draw some comments about possibilities of future research.

2 RELATED WORK

In this chapter, we describe works related to our project divided into sections according to each theme. We briefly survey the exploration of music data sets, visualization of user's personal music listening history and automatic genre classification. All these works helped us to choose solutions in different stages of our project.

2.1 Aesthetics, Perception and Accessibility

How easy information can be collected by users, how fast it can be perceived and comprehended,and how long it remains in one's mind are an important issues for decision making.

Visual representations that are meant to convey more information through a high speed and large communication channel must increase the amount of retained information. Borkin et al. (BORKIN et al., 2013) analyzed the memorability of a visualization, and concluded that a visualization is more memorable if (i) it is distinct; (ii) it is a distinct visualization type; (iii) it is colorful; (iv) it is visually dense and (v) it has a low data-to-ink ratio. Studying also memorability, Bateman et al. (BATEMAN et al., 2010) reported that people's accuracy in describing embellished charts was no worse than for plain charts, and that their recall after a two-to-three-weeks gap was significantly better.

Another issue regarding music is accessibility. Since music is not accessible to all people because they can be deaf, deafened, and hard of hearing, more information is needed to allow better understanding of the emotions conveyed by music. The work by Fourney and Fels (FOURNEY; FELS, 2009) shows a possible approach with music visualization. They explored several techniques for visualizing music that focus entirely on music notes and timing. Music consumers who are interested in being entertained are not necessarily interested in large amounts of information.

Another group developed a music visualization system prototype (HIRAGA; WATANABE; FUJISHIRO, 2002) that enables users to better understand a musical piece and its performance, especially for a cooperative performance. Users can better understand the performance expression by visualizing the performance with expressive cues of the qualitative music terms. Based on this understanding, users can exchange with other users their comments, recorded on visualized figures, through the Internet.

While improving aesthetics is important for music discovery, Libeks and Turnbull work (LIBEKS; TURNBULL, 2011) focuses on using techniques from computer vision to make additional use of music-related images. First, they propose a new measure of music similarity based on visual appearance. Second, images of artists also represent an unexplored source of music information that is useful for the automatic annotation of music: associating semantic tags with artists. They describe an image-annotation system that can both compute artists' similarity and annotate artists with a set of genre tags based on album cover artwork or promotional photographs.

Perception-based visualizations are also explored. Zhu et al. (ZHU; LU, 2005) developed an algorithm to automatically estimate human perceptions on rhythm and timbre of a music clip. Then, based on these two values, each music clip is mapped into a 2D (timbre-rhythm) space. Thus a 2D perception-based visualization is built. Experimental evaluation indicates that this kind of visualization is efficiently helpful in many cases of music management manipulations, such as music navigation, similar music search and music play list generation.

2.2 Rankings and Similarity Visualization

An interesting solution for the analysis of different rankings is LineUp (GRATZL et al., 2013). The authors presented a ranking scalable multi-attribute visualization technique based on bar charts. The technique allows tabular data sets to be sorted for creating different rankings, where the attributes values are represented by bars. Attributes can be grouped for sorting purposes, and different rankings for the same data set can be lined-up and compared. Figure 2.1 presents LineUp showing a ranking of the top Universities according to the QS World University Ranking 2012 dataset with custom attributes and weights, compared to the official ranking.

One of the simplest tasks when dealing with music collections is navigation and/or exploration. Ono et al. use a similarity graph (ONO, 2015) to enable the exploration of data sets in terms of hierarchical similarities. They built a methodology for users to visually explore music collections considering that the similarity can take place only in small parts of the song. It uses music information retrieval (MIR) techniques to find similar segments between pairs of audio files, and a graph metaphor to display the detected similarities.

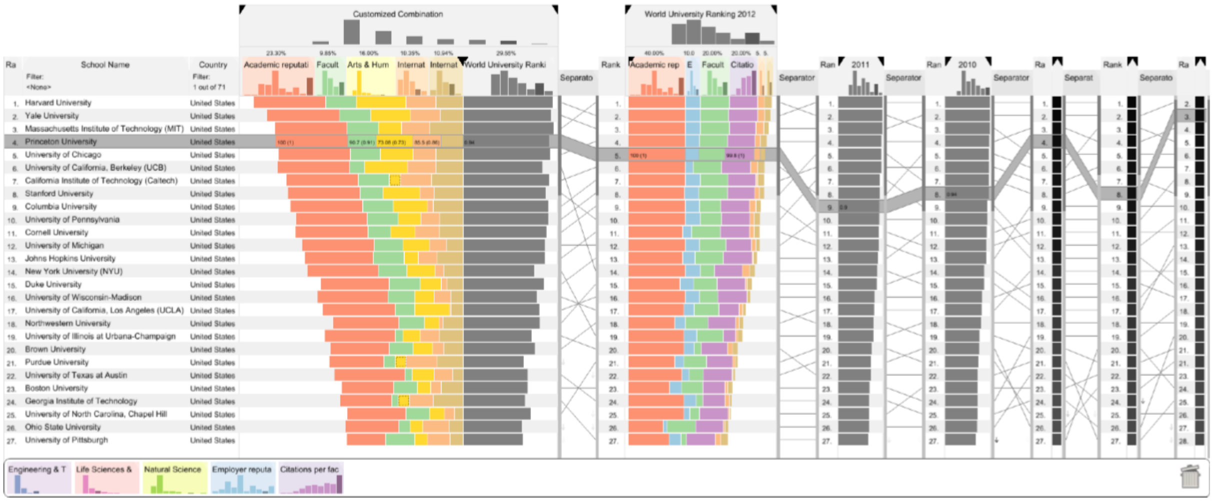


Figure 2.1: LineUp showing a ranking of the top Universities Ranking from 2012. Source: (GRATZL et al., 2013).

2.3 Music Rankings and Personal Data

Often, when dealing with rankings, time is an important attribute because rankings usually vary in time. This is especially true for music rankings. Thus, it is rather common finding works that are based on timeline visualizations. For example, Dias et al. (DIAS; FONSECA; GONcALVES, 2012) combine a timeline-based visualization with a set of synchronized views and an interactive filtering mechanism.

Also, it is also interesting to observe how music taste evolved along time: this has also drawn attention for showing Billboard data (POLYGRAPH, 2017). The data in the work is a time series starting in 1958: the top 5 artists for each week are shown in an interactive timeline, and the tool plays automatically the number-one track of each week. It is also possible to search a precise week, artist or track. Figure 2.2 shows a screen-captured image of Polygraph.

As for user's personal music listening history and lifelogging, there are some interesting works. LastHistory (BAUR et al., 2010) is an interactive visualization application for displaying Last.fm (LAST.FM, 2017) data, the music listening stories, along with contextual information from personal photos and calendar entries. The enthusiastic feedback that they received from average users shows a need for making personal data accessible.

Last Chart! (FORST, 2017) also uses personal data from Last.fm, and displays Bubble, Cloud and other visualization charts on the web. Another example is Peter Gilks' site (GILKS, 2017) that shows data from the tracking of his own music consumption on Spotify using Last.fm. He uses a handy script to download last.fm data into a CSV (FOXALL, 2017) for building the visualization.

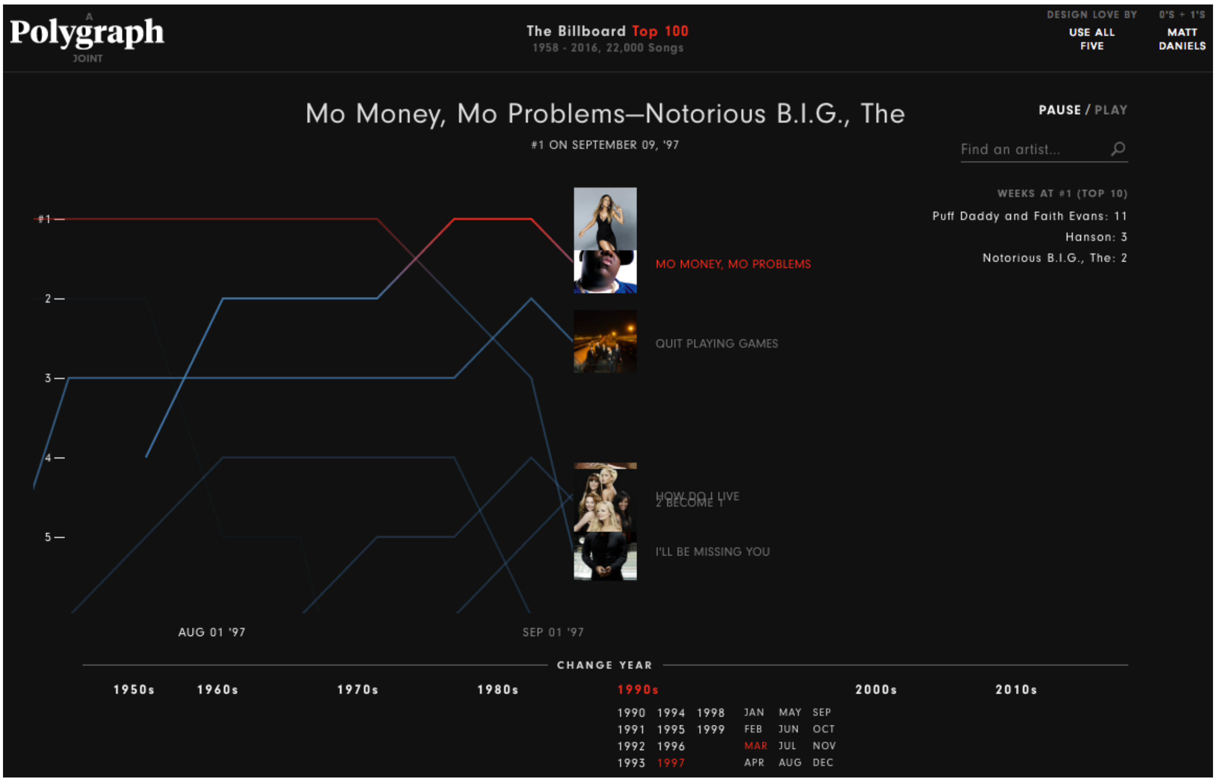


Figure 2.2: Polygraph playing Billboard Number 1 track of Sep. 1st 1997. Source: (POLYGRAPH, 2017).

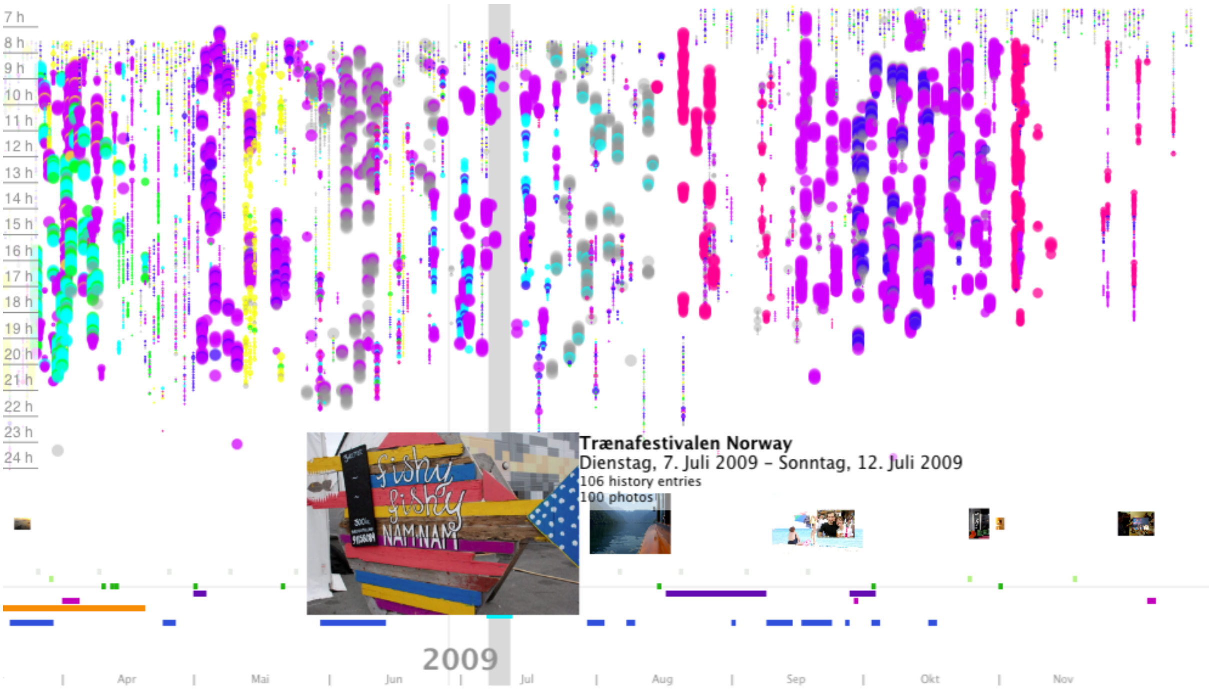


Figure 2.3: Personal Last.fm data from 2009. Source: (BAUR et al., 2010).

Listening factors (BAUR; BüTTGEN; BUTZ, 2012) present an empirical analysis of long-term music listening histories from Last.fm. Their sample contains 310 histories with up to six years duration and 48 associated variables describing various user and music characteristics.

They aggregated these variables into 13 components and found several correlations between them. The analysis showed the impact of seasons and a listener's interest in novelty on music choice. Using this information, a sample of a user's listening history or even just demographical data could be used to create personalized interfaces and recommendation strategies.

A group developed a visualization technique, called Hyper Word Clouds (NGUYEN; LE, 2016), for the examination of complex and multi-relationship Last.fm dataset. Through the text-based representation, tracks, albums, artists, and other Last.fm data items are visualized as words linked in parallel and anchor-based word clouds. The users can then interactively select to filter, highlight, and compare data and relationships of interest and to discover further insights.

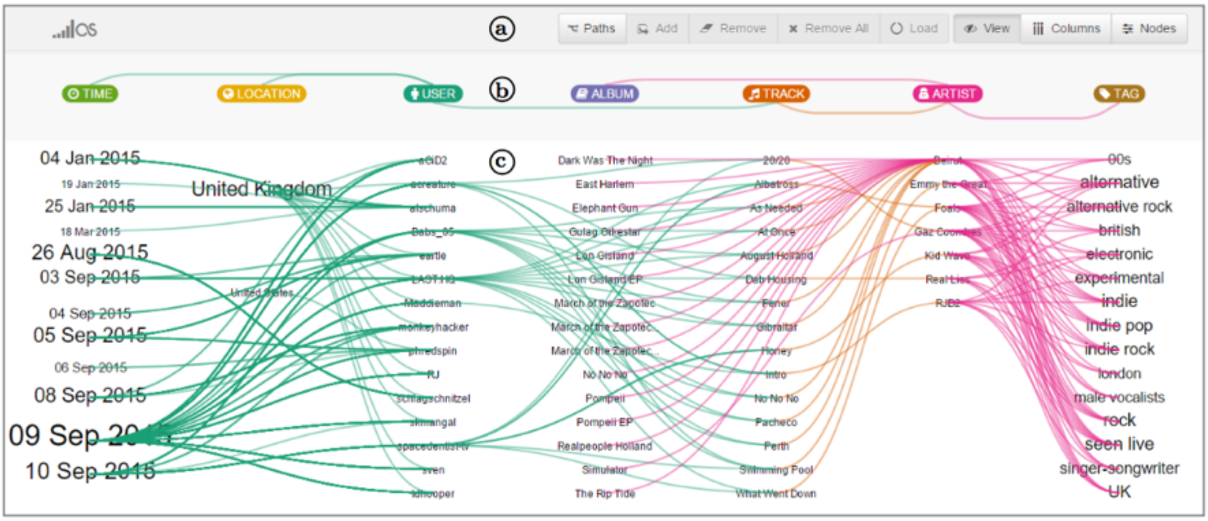


Figure 2.4: Hyper Word Clouds visualization design. Source: (NGUYEN; LE, 2016).

There are works related to other social networks. There is an interesting work based on Plurk (PLURK, 2017) social network. The work (LEE; DENG; LIU, 2013) provides several capabilities: (A) visualization for friends who share the same interest in music, (B) to group people who share the same interest in music into categories, and (C) to recommend songs function for an increase in the common interest in music. The research fellows need to handle the text information gathered from Plurk to carry out regularization. They used data mining method to analyze the information on the subject of music interest and they classify various types of songs. They also substitute these keywords called different degree of preference into the iSpreadRank algorithm to give different degree of preference.

Dealing with Twitter data, Streamwatchr (WEERKAMP; TSAGKIAS; RIJKE, 2013) is a real-time system for analyzing the music listening behavior of people around the world. It collects music-related tweets, extracts artists and songs, and visualizes the results in three ways: (i) currently trending songs and artists, (ii) newly discovered songs, and (iii) popularity statistics per country and world-wide for both songs and artists.

Music Tweet Map (HAUGER; SCHEDL, 2016) uses Twitter (TWITTER, 2017) data to build an interface for browsing music listening events on a global scale. These events were extracted automatically from a large set of microblogs harvested from Twitter. The major features are browsing music by time, set specific locations, topic clusters learned from tag information and music charts. Furthermore, music can be explored via artist similarity. They also present a music similarity measure, based on co-occurrence analysis of items in users' listening histories. A ranking of artists played in Brazil during the period of collected data from Music Tweet Map can be seen in Figure 2.5

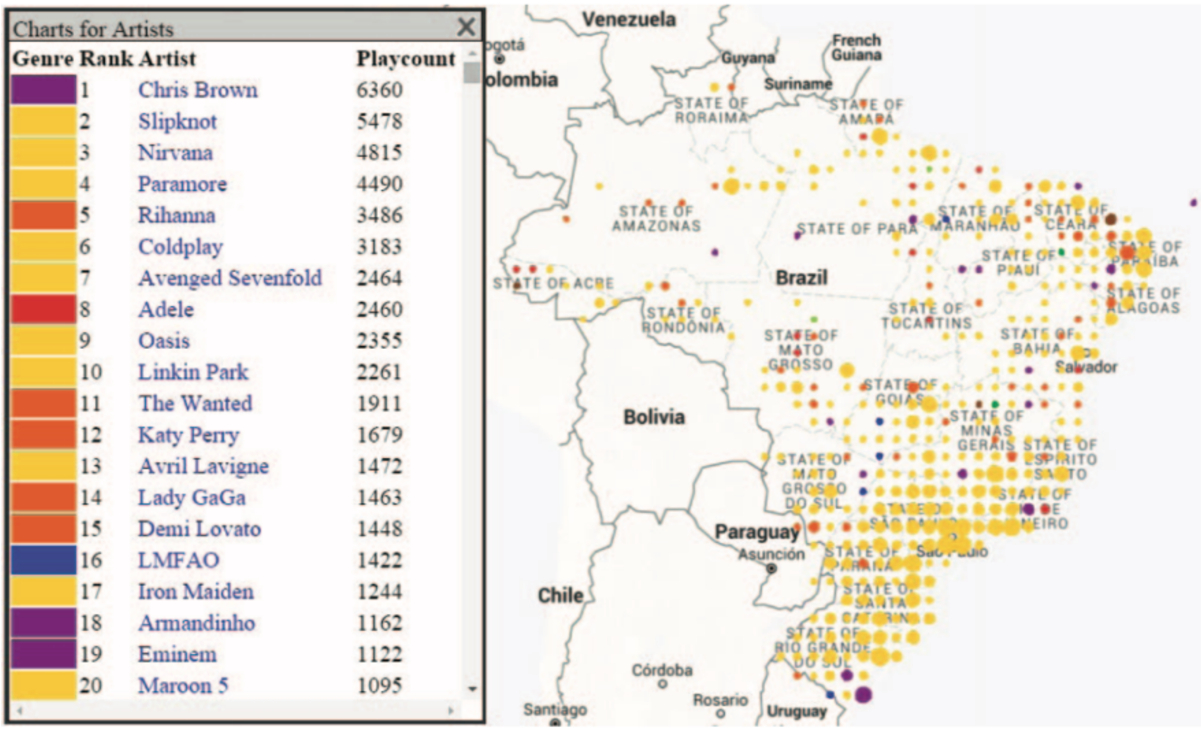


Figure 2.5: Rankings of artists played in Brazil based on Twitter data, considering the total collection period of the dataset of Music Tweet Map. Source: (HAUGER; SCHEDL, 2016).

2.4 Users, Mood and Genre Classification

Automatic genre classification is crucial for the organization, search, retrieval and recommendation of music. Valverde-Rebaza et al. (VALVERDE-REBAZA et al., 2014) investigate two components of the music genre classification process using traditional and relational approaches: a novel feature vector obtained directly from a description of the musical structure described in MIDI files (named as structural features), and the performance of relational classifiers compared to the traditional ones.

An analysis and visualization approach is reported by Zhang and Liu (ZHANG; LIU, 2014). They aimed at analyzing users' interests, and the work revealed the underlying relevance of music tracks based on metadata and also on users' votes, as a collaborative relevance.

Accessing personal and online music libraries with thousands of songs has become an everyday activity. Instead of textual lists, the libraries can also be accessed using graphical visualizations such as adaptive avatars. Holm et al. (HOLM; SIIRTOLA; LAAKSONEN, 2010) designed 17 stereotypical avatars representing different musical genres to study how well the avatars were recognized. The work discusses the design of the avatars, explains which musical genres were selected for the study and presents the results of the survey.

Another work from Holm et al. (HOLM; SIIRTOLA, 2012) presents and compares four different methods for visualizing musical genres: colors, icons, fonts and avatars. The findings from online surveys were utilized for designing novel graphical user interfaces (GUIs) for an existing music recommendation system. Based on the surveys' results, it was found that the best performance would be achieved by combining different visualization methods together. The best performing method (avatars) was partially based on the findings from color and icon questionnaires. The easiest genre to visualize was heavy metal; it performed better during their evaluations of online questionnaires and prototype design. Figure 2.6 shows part of this work with association percentages for genres and font collections.

Genre	Font collection	%	Genre	Font collection	%	Genre	Font collection	%
Alternative & Indie	BAND NAME Band Name Band Name	38	Blues	BAND NAME BAND NAME Band Name	5	Classical	Band Name Band Name Band Name	93
Country	BAND NAME BAND NAME BAND NAME	90	Electronica & Dance	BAND NAME BAND NAME Band Na	90	Folk	Band Name Band Name Band Name	33
Gospel	Band Name Band Name Band Name	17	Hip-Hop & Rap	BAND NAME BAND NAME BAND NAME	48	Jazz	BAND NAME BAND NAME BAND NAME	38
Latin	Band Name BAND NAME BAND NAME	10	Metal	BAND NAME BAND NAME BAND NAME	90	New Age	band na Band Name BAND NAME	7
Pop	BAND NA Band Name BAND NAME	21	Reggae	BAND NAME BAND NAME BAND NAME	12	Rock	BAND NA Band Name BAND NAME	43
Soul, RnB & Funk	Band Name BAND NAME Band Name	29	World Music	Band Name Band Name BAND NAME	62			

Figure 2.6: Association percentages for genres and font collections. Source: (HOLM; SIIRTOLA, 2012).

It is well known that music can convey emotion and modulate mood. This work (FENG; ZHUANG; PAN, 2003) concentrates on MIR by detecting mood, it is implemented by analyzing two music dimensions, tempo and articulation. They derive four categories of mood, happiness, anger, sadness and fear. They report the experimental result on a test corpus of 353 pieces of popular music with various genres.

Collaborative music discovery was studied in a work (LEHTINIEMI; OJALA; VÄÄNÄNEN, 2016) by creating playlists and associating them with mood pictures. The concept was evaluated in two field trials by a total of 45 individual users, with both trials containing 30 users and 15 of the users attending both of the trials. The results from the two field trials are presented under three main themes: socially augmented music discovery, user-generated content enhancing music discovery and social usage patterns emerging from the usage of such a system. Users formed ways to facilitate social interaction and music discovery through the playlist content they shared. The findings can be used as design implications for mood-based music service designers.

As can be noticed, none of the mentioned works deal directly with visualization of music rankings. Our tool intends to fill this gap by providing visualizations of music rankings and personal music data. We aim at supporting the comparison of rankings and, most important, showing attributes of the music tracks in the rankings, which is likely to make easier for a user to decide between listening a different music, exploring new alternatives based on genre, artist and position in the ranking, for example, or following the known path of listening the same music tracks.

3 MUSIC RANKINGS VISUALIZATION

People use rankings in general to compare data and get recommendations. Thus, the idea of using music datasets for building rankings is natural in the current streaming era. In 2015, at least two new big streaming services appeared, Apple Music and TIDAL, involving famous artists and record labels. This has impacted the traditional rankings. Playlists from these streaming services work as imperceptible merchandising. The users are attracted by titles such as "Top 100", "Hottest tracks", "The most played tracks", and start listening to brand new tracks, resulting in a recommendation cycle.

In this work, we decided to use two datasets acquired from Billboard and Spotify. They were chosen because they are the most used as observed from our Requirements Analysis, which is shown in the next section. Billboard data were acquired with a Java web crawler and stored in a MySQL database containing the track position, track name, artist, URL to listen on Spotify, last week position, weeks on chart and peak position. Spotify data were acquired from Spotify Charts as CSV files and also stored in the database, containing the track position, track name, artist, streams and URL. Music genre was an extra information mapped with iTunes (APPLE, 2017) to handle data redundancy, and added to each artist data. We only considered the major music styles, so we would have a small amount of data to represent, making easier to identify genres by color. Data gathered from the music rankings were all saved in a MySQL database, so, the application just loads the specific week content that the user is requiring.

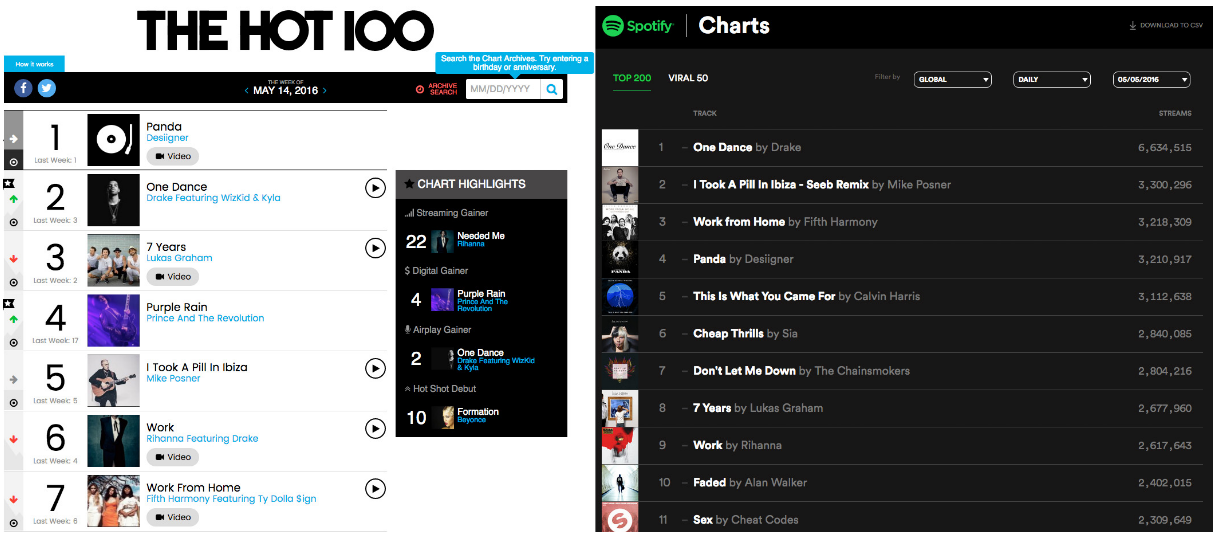


Figure 3.1: Billboard and Spotify Charts websites displaying the ranking of musics collected during the May 7th 2016 week. Billboard displays data one week further than the current week.

In order to propose a visualization for such music rankings, we started the project with a requirements analysis phase.

3.1 Requirement Analysis

A remote user survey was set up for 13 days to obtain data about users' preferences and habits in listening music or selecting new music tracks to listen. The survey was disseminated in mail lists, forums and social networks, so the users were totally random, and included who listen to music daily and who rarely do it. It was done in June 2015.

A total of 377 people from 11 countries answered our questionnaire. They were 23 years old in average, ranging from 13 to 60 years old, 130 females (34.5%) and 247 males (65.5%). Concerning to education, 37 (9.8%) have elementary school degree, 160 (42.5%) have high school degree, 131 (34.7%) have undergraduate education, 43 (11.4%) have Master degree and 6 (1.8%) have Ph.D. degree or Post-doctorate.

Similarly to Liikkanen and Åman (LIIKKANEN; ÅMAN, 2015), we found out that Youtube (84.8%), download (66.9%) and Spotify (44.7%) are the most used services to listen to music. They are followed by Radio (35.1%), CD/DVD (21.1%) and iTunes (18.3%). The preference for Youtube might be explained because it is easy to access as well as it is free.

We found out that people discover new music through the same services they use to listen to music, such as Youtube, Spotify or Radio (72.8%); through friends' recommendation (61.5%); music rankings (25.2%); clubs/concerts (12.1%); and forums (11.8%). So, the influence that music services have on users was confirmed, as the importance of friends' recommendation and music rankings.

Music genre influences most users choices for new music, with an influence rate of 93%, followed by artist (81%), music rankings (32%) and release date (22%). This influence affects how the music is chosen, as we expected. Music rankings are supposed to guide and rank general preferences, but not to influence so much the users' choices.

When users are interested in music rankings, they mainly look at Billboard (27%), followed by Spotify Charts (21.6%), and 37% of the people look at one of them at least. The interest criteria preferred by users when looking at music rankings visualization are: recommended tracks by each artists (51.7%), followed by total number of executions per track (51.4%), recommended tracks (51.1%) and total number of executions per artist (31.5%).

We asked what the potential users would expect from a website with music ranking visualization and answers vary from "an interface that's quick and easy to read, so you can get a decent amount of data without reading a wall of text" to "explain what counts as popular, and to maybe have a breakdown by region and by platform (Spotify plays vs. Radio plays, etc...)". Comments include demographics of listeners, play count of tracks and artists, playlists, popularity of tracks and artists by country.

Considering the results of our survey, we decided to investigate visualization techniques for music rankings, as well as integrating different attributes of the music tracks. Moreover, although the rankings visualization should include some of the attributes of the tracks, we also propose an alternative visualization to show the distribution of artists (and tracks) per genre. The following sections describe our design choices and the visualization techniques.

3.2 Design Choices

As mentioned before, we have chosen Billboard and Spotify as main data sources, Youtube was not considered because personal data services such as Last.fm do not consider scrobbles (when Last.fm automatically sends the name of each song played by an user) from this service. However, any service that provides the data we employ in our visualizations can be used as source.

With the data from rankings acquired and treated, we analysed carefully the results from the remote users survey to check what would be the best choice in visualizations.

Nowadays, the rankings on Billboard and Spotify are displayed as ordered lists of tracks based on the position in the ranking of the most listened music tracks/artists. As for Billboard, the interaction is basically, for some tracks, the possibility to get a link to the music video on Youtube and a link to the streaming on Spotify. The Chart Highlights section brings us some important events in the ranking. Regarding Spotify, the list is even simpler, exhibiting the position, track and artist, and the number of streams. The interaction is just the possibility of click on one of the tracks and listen to it. In the beginning of 2015, they stopped sharing gender and age information from users.

Since we found out 93% of the users are interested on the music genre, it became really important to add this information to the data. We chose to use colors for representing genres. All of the colors have similar tones and try to express the feeling or the major album colors of the music genres. The artist is also important to users, thus the visualizations explicitly represent them.

Finally, we have implemented the following data visualizations techniques: Sunburst (STASKO; ZHANG, 2000), Node-Link Tree, Bubble Chart and Treemap (JOHNSON; SHNEIDERMAN, 1991), all being able to represent music data content and music genre. They were chosen because they mix different types of visualizations, representing what related works taught us. They were implemented using D3.js (BOSTOCK; OGIEVETSKY; HEER, 2011), a

Javascript library for visualizing data, available at (BOSTOCK, 2017).

3.3 MusicVis overview

MusicVis is a web-base application, which shows a menu and a main data visualization area. The main option is "Music Data Visualization", which is explained below. Filter, Compare and About Us links, and a flag icon to change the language between English and Portuguese, are also included in the menu. An image of the final interface can be seen on Figure 3.2.

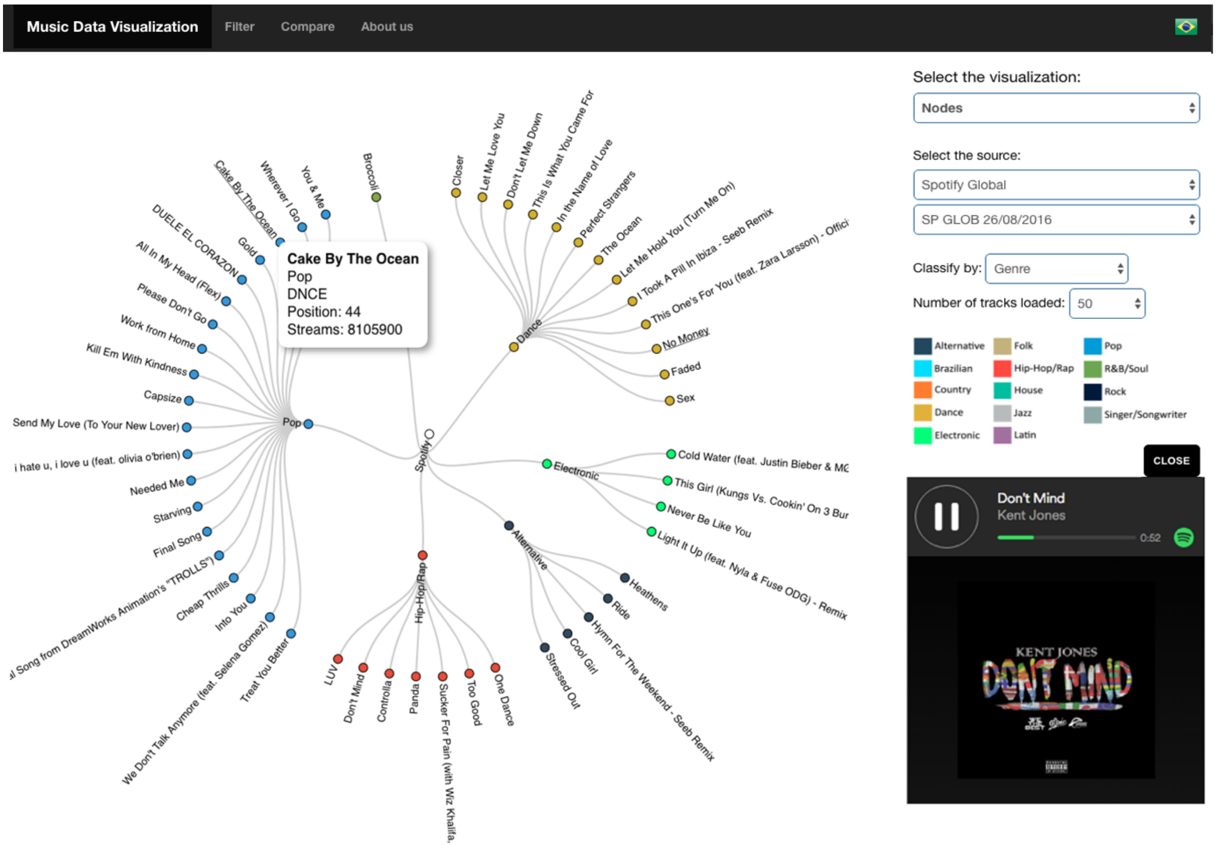


Figure 3.2: MusicVis Tool Interface: Data from Top 50 Spotify Global on August 26th 2016 week is exhibited by Genre as a Node-Link Tree visualization. The tooltip allows obtaining more information about each track, and when clicked, a Spotify player appears to listen to it.

In the main page, "Music Data Visualization", the user can visualize Music Rankings and Personal data. The user is able to:

- Select the visualization technique between Sunburst, Node-link Tree, Treemap and Bubble;
- Select the source of music data between Billboard, Spotify or personal data from Last.fm;
- Select the ranking week;

- Classify by Genre, Artist or Track;
- Select the number of loaded tracks.

Each visualization has a tooltip to present information about each artist, track or music genre. Also, it is possible to click and listen to the music on an Spotify embedded player.

When the page opens for the first time, some configurations are automatically loaded: a Sunburst visualization, the Spotify Global data from the current week, the classification by Artist and 50 tracks loaded. This decision was made to facilitate the user interaction with the tool and because our User Study (Chapter 5) showed they are the ones preferred by potential users.

In the next sections, we give details about how the visualizations work with the music ranking data.

3.4 Sunburst Music Visualization

A Sunburst chart (STASKO; ZHANG, 2000) displays a hierarchy of items layered in a circular arrangement. We created a Sunburst interactive visualization, similar to a previous work (GUEDES; FREITAS, 2016), to display and allow comparison of Billboard and Spotify rankings. Figure 3.3 presents a visual representation built with this technique.

The outer layer represents the music tracks, while the inner layer depends on the criteria used to order the data: when ordered by artist or position in the ranking, the inner sections represent the artists; when ordered by music genre, they represent the music genre. Color is used to represent music genre in both cases (there is a legend at the right side, not shown here).

The section size means the position (Billboard) or the streams (Spotify). When ordered by artist, the tracks from the same artists are clustered no matter what are their position in the ranking.

The usual Sunburst behavior in response to the selection of a section is implemented: when clicking on a section, the visualization changes for showing that specific music genre or artist occupying all the inner circle along with the related tracks in the outer sections. Figure 3.4 presents a visual representation when we click on the visualization. The central white circle is used to back to the main form.

A tooltip shows details about items. If the mouse is on a section of the inner circle, the tooltip will display the artist name or the music genre; if it is on the external circle, it will show the track name, music genre, position, and streams (Spotify) or last week position (Billboard).

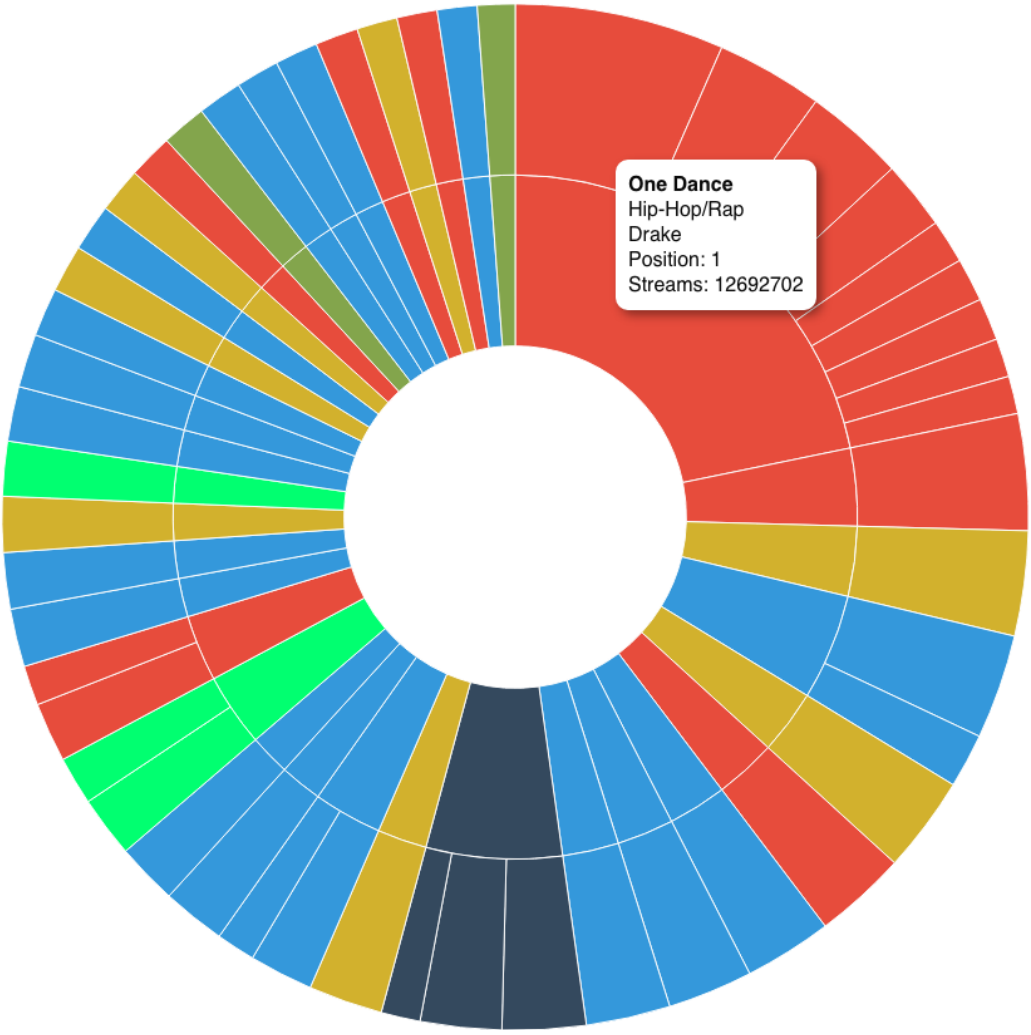


Figure 3.3: Sunburst visualization applied to Spotify's USA Top 50 during the June 17th 2016 week and ordered by artist.

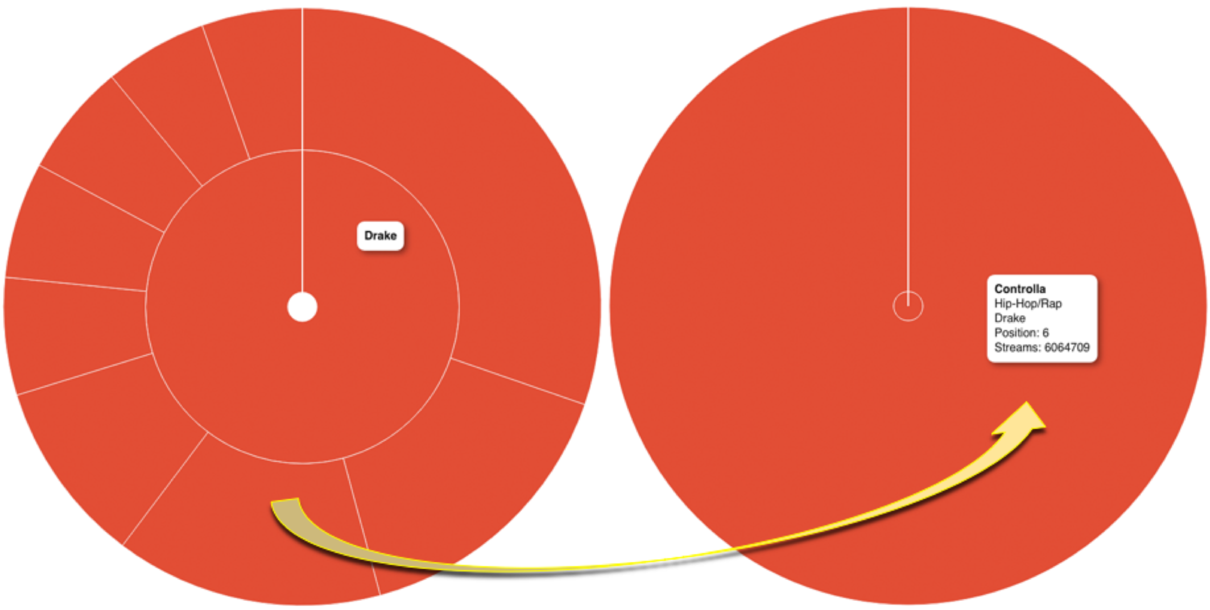


Figure 3.4: Sunburst clicking transformation applied to the rapper Drake on Spotify's USA Top 50 of June 17th 2016 (left) and zoom applied to an specific track (right).

3.5 Node-Link Tree Music Visualization

This visualization is a radial Reingold-Tilford tree (REINGOLD; TILFORD, 1981), with tidy arrangement of layered nodes. The central node represents the music ranking source. The depth of the nodes is computed by the distance from the root and the number of layers. Our technique can be seen on Figure 3.5.

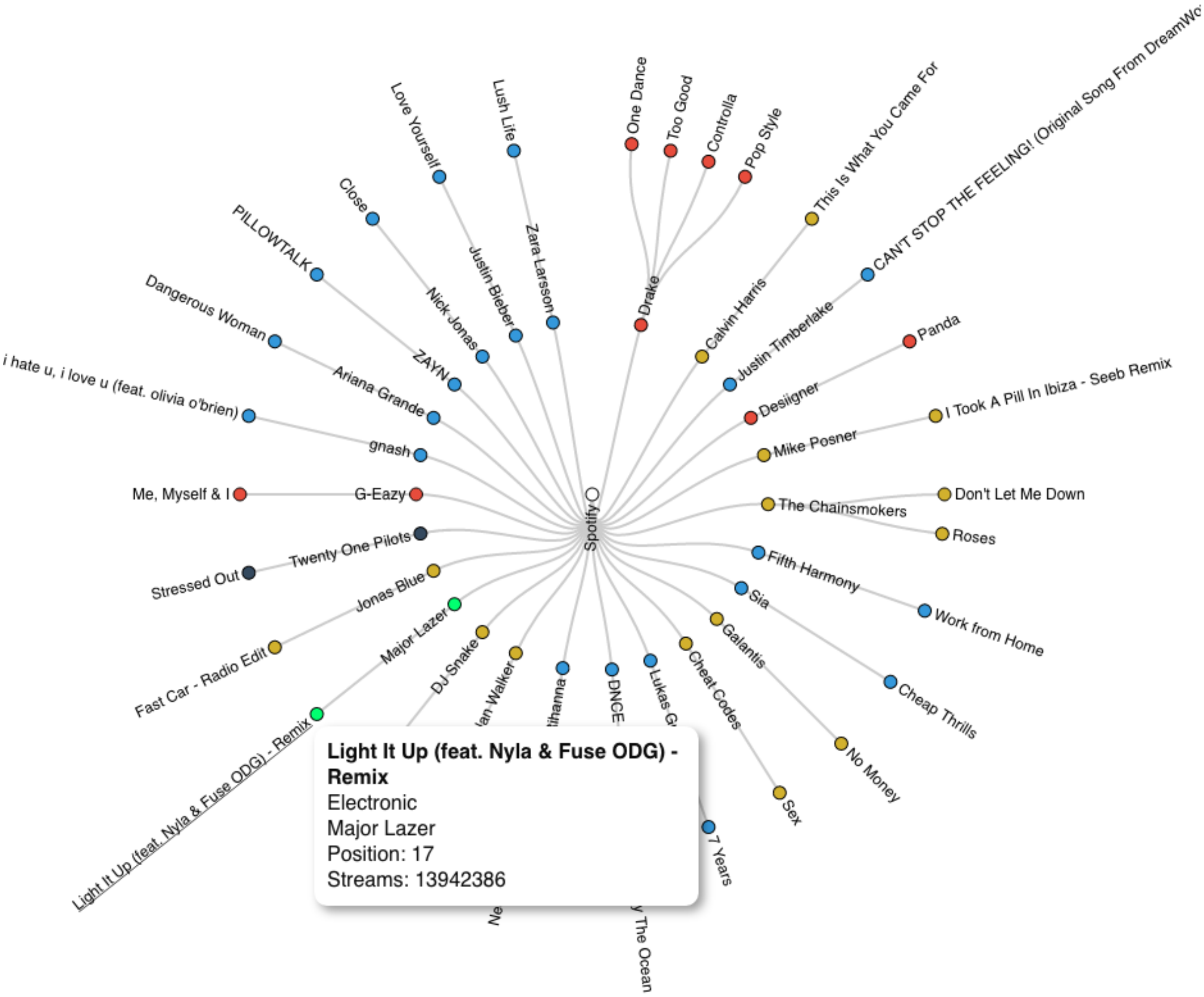


Figure 3.5: Node-Link Tree visualization.

The Node-Link Tree (NLT) visualization is displayed ordered by music genre. Each genre is represented by a node, and has one or more music tracks, which are represented by leaves connected to the music genre node.

The tooltip is available in all visualization techniques, and in NLT it is displayed when the user hovers the mouse over each node or text. It shows the music track name, music genre, artist, position and streams (Spotify) or last week position (Billboard). Also, all nodes and text are clickable: clicking on an leaf node (music track) it will open a Spotify online music player with the album cover. Clicking on an inner node (Music genre, Artist or Position), it will search this term on the web to give more information. This is useful when users are curious about where is the artist from, who are other famous artists of that music genre, and so on.

When ordered by artists, the inner nodes become the artists and the leaves are the music tracks. The same method is used when it is ordered by position, adding the position number beside the artist name.

3.6 Treemap

Treemaps (JOHNSON; SHNEIDERMAN, 1991) is a method for displaying hierarchical data by using nested rectangles. Figure 3.6 shows an example of this visualization applied to our dataset.

Figure 3.6: Treemap applied to Billboard Top 60 by genre, during the May 18th 2016 week.

As in Sunburst, when selecting Spotify data, the size of each section represents the amount of streams. The dataset shown in Figure 3.6 is classified by genre, which make easier to notice the proportion of each music genre on that specific week. Each rectangle can be clicked on to open a player.

3.7 Bubble Chart

Bubble charts represent data by circles of different sizes and colors. It is a widely know graphical representation.

In Figure 3.7 we can see such visualization as implemented in our work. It starts ordering the circles based on some criteria, then displays them from the center, spiralling data around. Data is classified by position or clustered by artist and music genre. The bubbles are clickable, so they allow listening each track. Tooltip and colors representing music genres are also available.

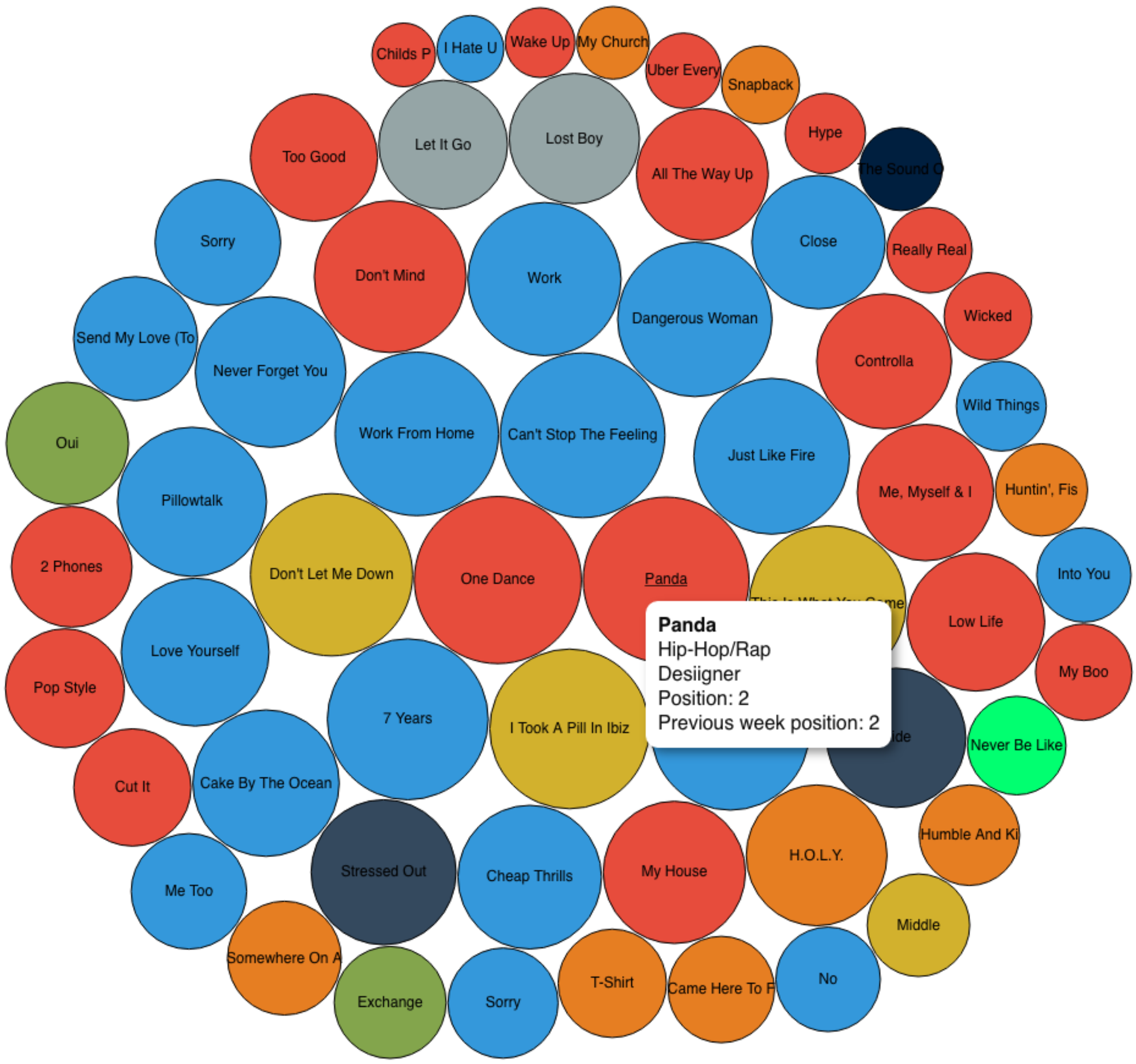


Figure 3.7: Bubble Chart applied to Billboard Top 60 by position, during the May 18th 2016 week.

3.8 Discussion

This section presents a brief summary of the visualization techniques found during the state-of-the-art review we performed. We added MusicVis to Table 3.1 to compare our tool to the ones available on the web.

Table 3.1 shows that Last Chart! (FORST, 2017) has the most number of visualization techniques implemented. This tool loads Last.fm data, and presents them with different kinds of visualizations.

Billboard (BILLBOARD, 2017) and Spotify Charts (SPOTIFY, 2017b) present their rankings in a list (or tabular) form. Peter Gilks' work (GILKS, 2017) presents his static personal Last.fm data, with Bubble and List visualization. Finally, Polygraph (POLYGRAPH, 2017) presents an interactive timeline-based visualization.

Table 3.1: Visualization techniques provided by web-based tools.

	Visualization						
Tool	Sunburst	Node-Link Tree	Bubble	Treemap	Force	Timeline	List
MusicVis	✓	✓	✓	✓			
Billboard							✓
Spotify Charts							✓
Peter Gilks			✓				✓
Polygraph						✓	
Last Chart!	✓		✓	✓	✓		✓

Table 3.2 summarizes the related works found in the literature. MusicVis was not included in this comparison table because it does not implement the same visualization techniques.

Table 3.2: Visualization techniques described in papers from the literature.

	Visualization			
Paper	Parallel C.	Timeline	Bars	Location
(BAUR et al., 2010)		✓		
(DIAS; FONSECA; GONcALVES, 2012)		✓		
(LEE; DENG; LIU, 2013)			✓	
(NGUYEN; LE, 2016)	✓			
(HAUGER; SCHEDL, 2016)			✓	✓

The most complete related work is Music Tweet Map (HAUGER; SCHEDL, 2016). It mixes location-based visualization with dots, with bar charts technique for comparison. Hyper Word Clouds (NGUYEN; LE, 2016) uses a kind of Parallel Coordinates visualization technique. Lee et al. work (LEE; DENG; LIU, 2013) only presents Bar Charts.

Both "The Streams of Our Lives" (BAUR et al., 2010) and "Music Listening History Explorer" (DIAS; FONSECA; GONcALVES, 2012) use a timeline-based visualization.

A more complete comparison between these works and MusicVis is presented in the next chapter.

4 INTERACTIVE FEATURES

This Chapter will discuss the interactive features provided by MusicVis: Filtering, Comparison and Personal Data Interaction. They can be employed with data from Billboard, Spotify Charts and also Last.fm personal data, allowing to interact with them.

4.1 Filtering Feature

Filtering and searching (Fig.4.1) are used to obtain data from our database. The query can be based on Artist, Track, Genre or Position. Once the type is selected and the name is typed, the user can filter specific ranking, country and week, or check the full results.

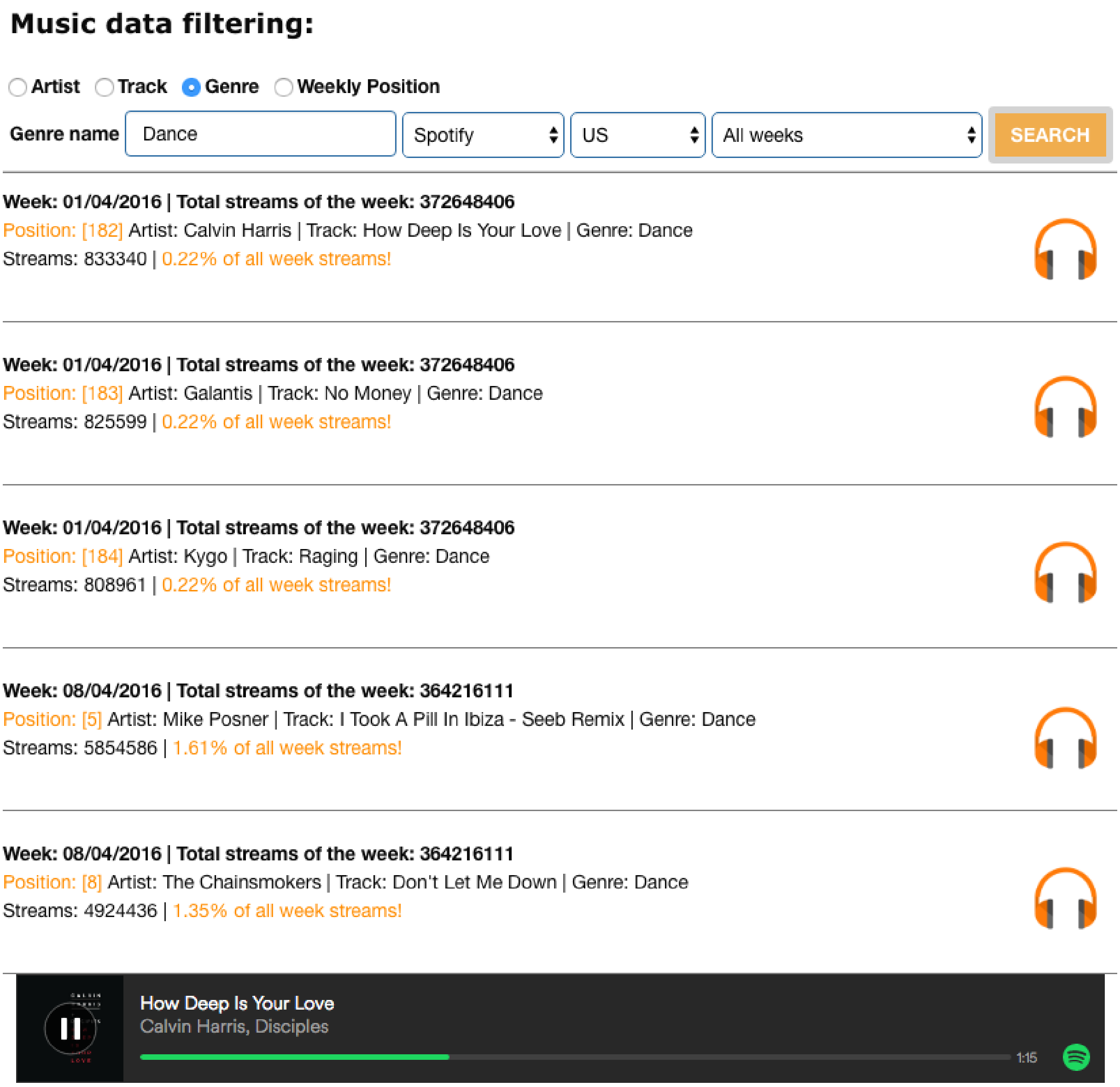


Figure 4.1: Filtering Dance music data from all Spotify USA ranking.

The results are shown as a list, with the following information: Week, Position, Artist, Track, Genre, Streams and Total number of streams on that week (Spotify) and the amount of listeners (in %) on that specific week.

The user can also order the result by new/old or relevance (representing the most listened tracks). When clicking on the headphone icon, a Spotify player is shown, and the user can listen to the selected track.

4.2 Comparison Feature

The user can compare artists and music tracks in a Multi-Series Line Chart, as can be seen in Fig. 4.2.

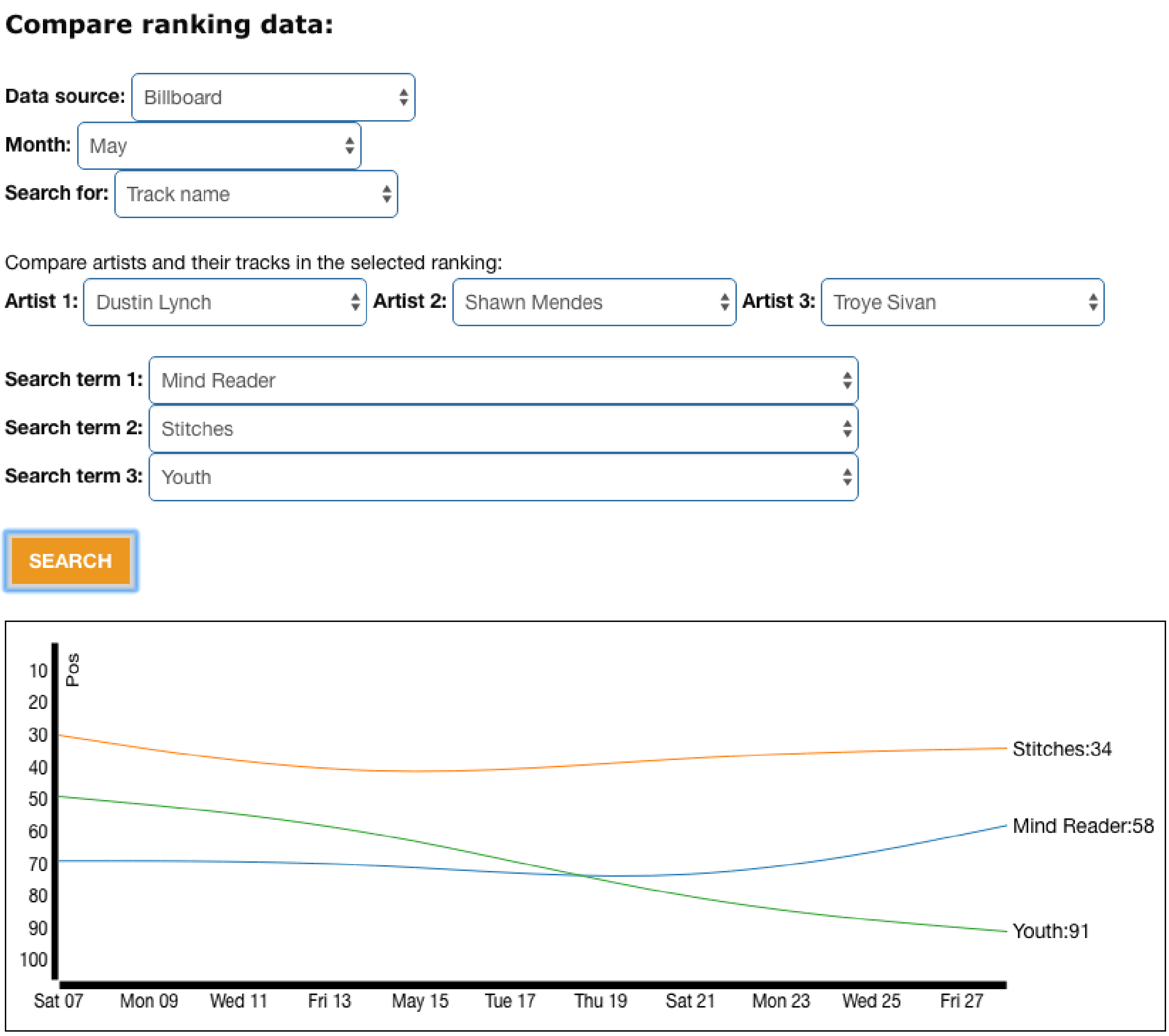


Figure 4.2: Comparison between 3 artists that appear in Billboard rankings of May 2016. They are all compared in a Multi-Series Line Chart.

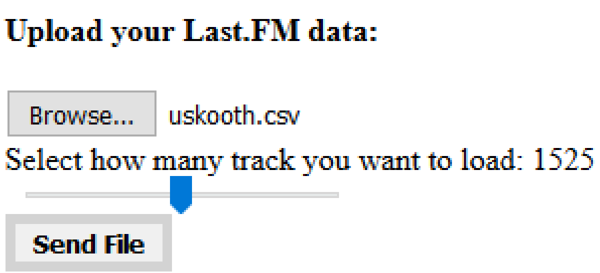
Firstly, one selects the data source (Billboard or Spotify) and Month, and then Search for Track Name or Artist. A drop-down list is shown with Artists to select. If we are searching for tracks, Search terms are available along the track names of each artist. After clicking on the search button, a Multi-Series Line Chart displays how the track oscillated during the month. In this graph, colors do not represent music genres, but are used to differentiate each result.

4.3 Personal Data Interaction Feature

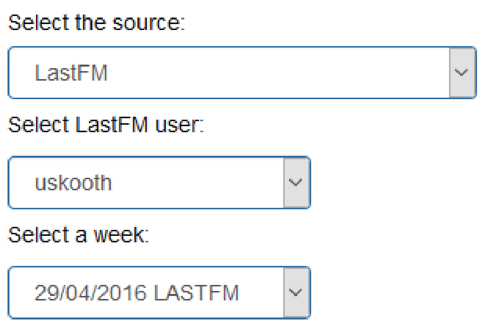
This lifelogging feature uses Last.fm data to extract users' personal music data. All of the Scrobbles from the user available at Last.fm are saved into a CSV file using Benjamin's script (FOXALL, 2017).

Once the user select its profile name, the next step is to select how many tracks from the CSV file our tool will synchronize. The upload supports reading up to 3000 lines. Effectively, this means that it can load data from up to 3000 user songs, which should normally represent data from a long period of months. The bigger the number of tracks loaded, the slower the system will perform for synchronizing the data. The interaction is represented in Figure 4.3a.

After this upload, the tool dynamically processes and transforms the data, separating them according to the weeks which they belong, categorizing them by genre, and even automatically counting the number of views each song had each week. All of this transform the data to the same pattern and organization of Billboard and Spotify, maximizing the compatibility with what was already developed. After synchronize the data, the user can select their account in the user selection option, and then select one of their personal music data weeks, as can be seen on Figure 4.3b.



(a) Selecting Last.FM data and amount of tracks.



(b) Selecting LastFM user and week.

Figure 4.3: MusicVis interaction with Last.FM data.

It is possible to exhibit the personal data into the four available visualizations: Sunburst, Node-Link Tree, Treemap and Bubble Chart. This allows users to visualize their personal information in interactive ways. In addition, the users receive information about the music data, such as Total number of streams and amount of Streams from that specific week. A screenshot of the full feature with Bubble visualization can be seen in Figure 4.4.

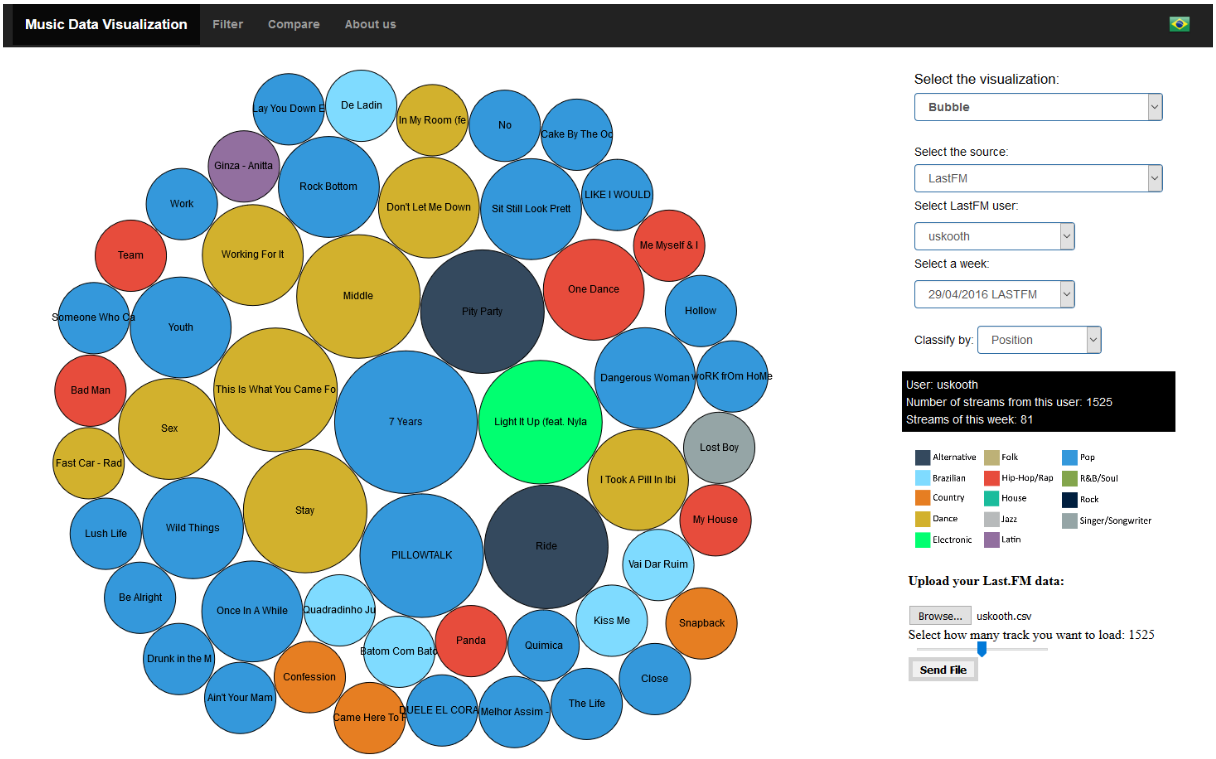


Figure 4.4: Page selecting Last.FM data from an user.

All of the interaction is available from the same main page, becoming easier to compare with Billboard and Spotify Data.

4.4 Discussion

This section summarizes the interactive features and personal data visualization techniques described in the related work found in the literature. Table 4.1 displays the summary of all music visualization related works.

MusicVis, Polygraph (POLYGRAPH, 2017), The Streams of Our Lives (BAUR et al., 2010), Hyper Word Clouds (NGUYEN; LE, 2016) and Music Tweet Map (HAUGER; SCHEDL, 2016) have both filter and comparison techniques.

Spotify Charts (SPOTIFY, 2017b), Last Chart! (FORST, 2017) and Music Listening History Explorer (DIAS; FONSECA; GONcALVES, 2012) have only the Filter feature while the work "Visualization for interest in music based on plurk social network" (LEE; DENG; LIU, 2013) has only the Comparison feature. Billboard (BILLBOARD, 2017) and Peter Gilks (GILKS, 2017) have no interactive features.

Table 4.1: Interactive features and datasets provided by Related Works.

	Features		Data Source				
Works	Filter	Compare	Billboard	Spotify	Last.fm	Twitter	Plurk
MusicVis	✓	✓	✓	✓	✓		
Billboard			✓				
Spotify Charts	✓			✓			
Peter Gilks					✓		
Polygraph	✓	✓	✓				
Last Chart!	✓				✓		
Music Listening...	✓				✓		
The Streams...	✓	✓			✓		
Hyper Word Clouds	✓	✓			✓		
Music Interest...		✓					✓
Music Tweet Map	✓	✓				✓	

About data source, MusicVis is the only work representing more than one data source: Billboard, Spotify and Last.fm. Billboard and Polygraph represent Billboard data. Spotify Charts exhibits Spotify data. Peter Gilks, Last Chart!, Music Listening History Explorer, The Streams of Our Lives and Hyper Word Clouds display only Last.fm data. The work "Visualization for interest in music based on plurk social network" exhibits Plurk data and Music Tweet Map displays Twitter data.

5 USER STUDY

In this chapter we describe the user study we performed for evaluating the visualization techniques as well as the filtering and comparison features. Since it was conducted in September 2016, when the Personal Data Interaction was not totally implemented yet, this feature was not included in the study. We aimed at assessing the usability and learnability of our tool.

5.1 Participants

After invitation on the mailing list of our University and Technical High School, an amount of 94 Brazilian people volunteered for the experiment: 44 males (46.8%) and 50 (53.2%) females, ranging from 15 to 33 years old, with mean and mode equal to 18. Concerning to education, 86 (91.5%) are students in a computing-oriented high school course, 3 (3.2%) had already graduated in Computer Science (CS), 3 (3.2%) had a M.Sc. degree in CS and 2 (2.1%) have concluded Ph.D. studies in CS. Among all, 85.1% consider important or very important to listen to music; 53.2% have already followed music rankings; 79.8% know Billboard and Spotify Charts or both rankings; 89.4% have already used Last.fm. And, finally, 53.2% are acquainted to data visualization.

5.2 Procedure

The experiment was performed on a local network, taking around 15 minutes each. Only few information were given in person concerning to the procedure itself.

As mentioned, subjects were invited by e-mail and upon arriving at the laboratory at a specific date/time, they were told to sign an agreement statement, and fill in a profile questionnaire (Appendix B). Then, they were invited to surf on Billboard, Spotify Charts, and our tool, freely, and answer our questionnaire (Appendix C). They were able to answer our questionnaire at the same time they were using our tool. The participants had to give their level of agreement to positive sentences using a 5-point Likert scale ranging for 1 (strongly disagree) to 5 (strongly agree). The last step was to answer a System Usability Scale (SUS) (BROOKE et al., 1996) questionnaire (Appendix D).

5.3 Results and Discussion

Figure 5.1 summarizes the overall feeling of subjects while using our tool. Aggregating answers "4" (Agree) and "5" (Strongly Agree) as positive feelings, we found out that:

- 67% found the layout attractive;
- 85.1% found that the tool is very interesting;
- 76.7% found it highly interactive, and
- 62.8% found the tool relevant.

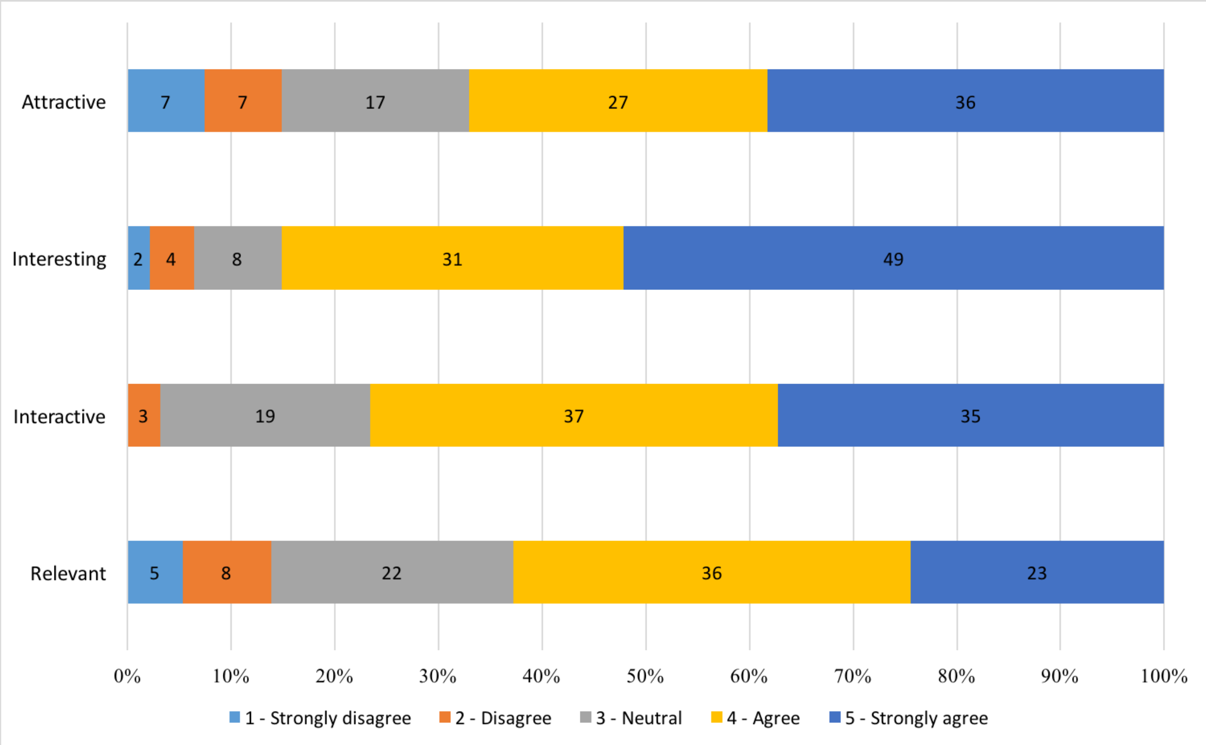


Figure 5.1: Summary of positive feelings about MusicVis. The figures inside the bars are number of users that answered the question with the corresponding score.

Regarding the evaluation of visualization techniques, the results are shown in Figure 5.2, and we can observe that:

- 66% liked Sunburst;
- 61.7% liked Node-Link Tree;
- 48.9% liked Bubble chart, and
- 44.7% liked Treemap.

The fact that users preferred Sunburst was already used in our design. As mentioned before, Sunburst is the visualization displayed when our tool is loaded.

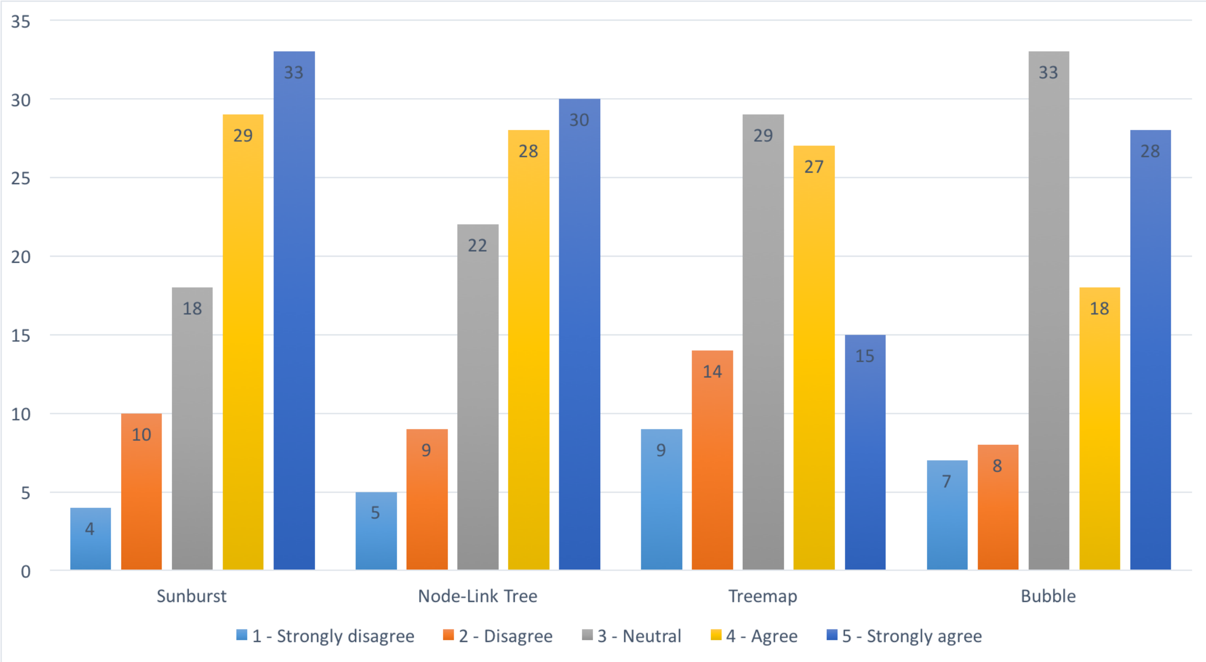


Figure 5.2: Number of users that liked each visualization technique.

Table 5.1: Results from preferences and features evaluation. Column numbers (1) to (9) correspond to sentences from the questionnaire. Line numbers correspond to Likert scale scores: 1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	5.3%	7.4%	1.1%	6.4%	4.3%	2.1%	3.2%	0%	2.1%
2	11.7%	12.8%	11.7%	8.5%	6.4%	7.4%	6.4%	3.2%	3.2%
3	26.6%	25.5%	21.3%	14.9%	20.2%	8.5%	14.9%	20.2%	17%
4	37.2%	30.9%	35.1%	19.1%	35.1%	41.5%	36.2%	41.5%	35.1%
5	19.1%	23.4%	30.9%	51.1%	34%	40.4%	39.4%	35.1%	42.6%

Other results about the visualization techniques and preferences are summarized in Table 5.1. Column numbers (1) to (9) represent the following sentences included in the evaluation questionnaire:

1. "I found that the proposed visualizations allow a better understanding of the ranking than the ones used by Billboard."
2. "I found that the proposed visualizations allow a better understanding of the ranking than the ones used by Spotify Charts."
3. "I liked the Sunburst visualization.", "I like the Node-link Tree visualization.", "I liked the Bubble visualization", "I liked the Treemap visualization."
4. "I found that the colors of musical genres were pleasant."
5. "I found the main page selection menu suitable for selecting the visualizations."
6. "I liked the data search on the filtering tab."
7. "I liked the data comparison in the comparison tab."

8. "I think the options for visualizing, filtering and comparing data are suitable."
9. "I think the visualizations could be applied to other areas."

The analysis of these results allowed us to trace some general conclusions, which we summarized below.

- 56.3% surely preferred using our tool instead of Billboard;
- 54.3% preferred our tool instead of Spotify Charts;
- 66% liked the chosen visualizations;
- 70.2% totally liked the music genre colors;
- 69.1% found the main selection menu suitable for selecting visualizations;
- 81.9% liked the data filtering feature;
- 75.5% liked the comparison feature;
- 76.6% found our design choices (for visualizing, searching and comparing data) suitable; and
- 77.7% think that our tool can be applied in other areas.

An interesting result is that these findings were confirmed by a SUS overall score of 79.2. Considering learnability, the score is 90.4. If we analyze only the group of those subjects acquainted to data visualization, we obtain a SUS overall score of 82 and a learnability score of 90.5.

The analysis of additional comments left by 18 participants allowed us to better understand what participants found about this work. Seven of these comments were about a good experience with the tool, such as: "user-friendly and interesting", "In certain environments where few people know other rankings, the application becomes very interesting to our knowledge about music", "I loved it, I discovered some interesting songs" and "I really liked it, great idea!". Calculating the scores for these users that gave us voluntarily a positive feedback the resulting SUS overall score grows to 85.7 and learnability increases to 94.7.

Noteworthy comments are reproduced below:

- "I really liked the Node-Link Tree because it was much more informative and its interaction was much more useful, with links to terms and the possibility of listening to the selected songs. In fact, the Node-Link Tree seems to replace Sunburst without any problems. Of all the options, Treemap was the least intuitive for me. I loved the comparison tab, but I think it could be on the main page somehow."
- "Some colors repeat themselves, leaving the visualization a bit confusing. In Sunburst, it

is not clear what is the difference between the "Size" and "count" options."

- "I liked the way it automatically changes the visualizations without submitting the form. You can improve the layout of the list in the comparison and allow the users to compare with as many artists as they want."

Subjects also reported a bug, criticized some aspects and made suggestions. The bug was found when changing the classification from artist to position in the Bubble chart visualization and the complain was regarding the colors we have chosen (1 user). Suggestions were made for adding the artist picture in the tooltip (1 user) and searching for non popular songs (1 user).

In summary, our tool was evaluated as interesting and relevant, in general, and our design choices for interaction and layout also yielded good results in terms of users' opinion. Next chapter concludes our work with a detailed analysis of our results and some comments on future work.

6 CONCLUSIONS AND FUTURE WORK

In this chapter we analyze our work based on the results obtained from the user study and discuss limitations and future work.

6.1 Analysis

From the analysis of the questionnaires we were able to obtain interesting insights about our project. Sunburst and Node-Link Tree were successful visualizations for our purpose. They are able to exhibit a large amount of data, and at the same time are pleasant. The traditional visualizations, Bubble Chart and Treemap, were not efficient in users' opinion: it was not easy to follow the ranking path in Bubble Chart, and Treemap was confusing, both non-intuitive.

Another important issue is that some artists are not available in Spotify, which makes the rankings different from one data source to another. Billboard considers radio, physical sales and other stream services, which make an artists like Beyoncé appear in this ranking and not in Spotify. Her new album was only available for streaming at TIDAL, for example.

This difficulty in mapping different artists is also present for music genres. Last.fm streams come from different sources, such as iTunes player, Windows Media Player, Spotify, etc. This can easily cause tag errors depending on the user because of special characters or wrong tag names. For example, the Pop group called *NSYNC is usually tagged as 'NSYNC, 'N-SYNC, *N-SYNC, N SYNC, and so on. This result in a hard work to match each artist to their real music genre.

A drawback of our work is due to the limited data sources. For example, we do not consider loading data from Youtube, the most used video streaming website. It has official and unofficial videos, causing major problems regarding tags and views count, and this would introduce errors in our music rankings.

Concerning the requirements we established in the beginning of the work (Section 1.1) we can observe that:

- R1: "build visualizations that fit music data without the traditional tabular form": we built visualizations that fit music data without the traditional tabular form: Sunburst, Node-Link Tree, Treemap and Bubble Chart are different from what one finds in the main music rankings websites (as seen in Figure 3.1).
- R2: "allow the users to have full control of what data, ranking and visualization they are

interacting with": our user study showed that 76.7% of the participants considered the tool very interactive and 69.1% found it suitable for controlling and interacting with.

- R3: "filter data with specific keywords": our tool allows filtering data from artists, track, genres and position; the user can also select a specific data source. This feature was approved by 81.9% of the users.
- R4: "the ability of listening to any music straight through the tool": this was implemented using Spotify plugin, as can be seen in Figure 3.2.
- R5: "compare the performance and position of any artist, track and genre": our Comparison feature allows comparing artists, tracks and genres from different data sources and dates. It can compare up to 3 items at the same time. This feature was approved by 75.5% of the users.
- R6: "compare personal data with the traditional rankings": our tool is partially compliant with this requirement since it was implemented but not tested properly. Our personal data system uses Last.fm, and can load personal data and easily compare it with the same week ranking from other data sources.

Each visualization had different advantages and disadvantages in user's opinion. Analysing results from the user study (Section 5.3), we can infer that Sunburst was the most liked visualization by 33 users (35.1%). We think that Sunburst caught people's attention because it is colorful and beautifully designed, changing its form. In second place, not far away from Sunburst, is Node-Link Tree which was liked by 30 users (31.9%). We had interesting feedback about why users liked it. In third place is Bubble Chart, liked by 28 users (29.8%). We think that this visualization got user's attention but was confusing due to its spiral form of organizing the ranking positions. The last one was Treemap, liked by 15 users only (16%), almost half of NLT and Bubble likings, and less then half of Sunburst. This traditional visualization was classified as not intuitive.

Comparing our tool with related works, we were in disadvantage compared to works like Polygraph, where the programmer had full access to all Billboard data, since 1958. This makes album cover and tags much more similar to the official results. Music Tweet Map and Listening factors works are considered complete in their segments, respectively, Twitter and Last.fm data. However, they do not have the intention of comparing music rankings, what MusicVis focused.

Differently from systems like Spotify (free account) and Youtube, where it is mandatory to hear merchandising of other tracks and products, causing interference in users' choices, we did not want to influence how users perceive and listen to music in our tool. It was satisfying to check users comparing data and discovering new music based on the our visualizations.

6.2 Future Work

Our motivation was the possibility of providing new ways of exploring music and artists data through visualization techniques. After evaluating MusicVis, our results show that our work is a new interactive way to present, find and compare music rankings, making easier for users to infer and become interested in music based on music genres and artists. Last.fm personal data was also implemented to have the ability to increase comparison and recommendation based on popular music rankings.

For future work, it is planned a side by side visualization comparison between music rankings and personal Last.fm data, being able to pick something in one ranking and highlight the same artist, track or genre in the other. Another future work is to compare artists and genres for more than a month: this temporal analysis would be valuable. Analysing the UX when dealing with these rankings is also planned for future work. We also want to provide new visualizations and obtain data from other music rankings.

REFERENCES

APPLE. **iTunes - Genre IDs Appendix**. 2017. Available at: <http://apple.co/2da4d4M>. Accessed: January 2017.

BATEMAN, S. et al. Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2010. (CHI '10), p. 2573–2582. Disponível em: <http://doi.acm.org/10.1145/1753326.1753716>.

BAUR, D.; BüTTGEN, J.; BUTZ, A. Listening factors: A large-scale principal components analysis of long-term music listening histories. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2012. (CHI '12), p. 1273–1276. Disponível em: <http://doi.acm.org/10.1145/2207676.2208581>.

BAUR, D. et al. The streams of our lives: Visualizing listening histories in context. **IEEE Transactions on Visualization and Computer Graphics**, v. 16, n. 6, p. 1119–1128, Nov 2010.

BILLBOARD. **Music Charts**. 2017. Available at: <http://www.billboard.com>. Accessed: January 2017.

BORKIN, M. A. et al. What makes a visualization memorable? **IEEE Transactions on Visualization and Computer Graphics**, v. 19, n. 12, p. 2306–2315, Dec 2013.

BOSTOCK, M. **Data-Driven Documents**. 2017. Available at: <https://d3js.org/>. Accessed: January 2017.

BOSTOCK, M.; OGIEVETSKY, V.; HEER, J. D3: Data-driven documents. **IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)**, 2011. Disponível em: <http://vis.stanford.edu/papers/d3>.

BROOKE, J. et al. Sus-a quick and dirty usability scale. **Usability evaluation in industry**, London, v. 189, n. 194, p. 4–7, 1996.

COMPANY, T. N. **Music Sales Measurement**. 2017. Available at: <http://www.nielsen.com>. Accessed: January 2017.

DIAS, R.; FONSECA, M. J.; GONcALVES, D. Music listening history explorer: An alternative approach for browsing music listening history habits. In: **Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces**. New York, NY, USA: ACM, 2012. (IUI '12), p. 261–264. Disponível em: <http://doi.acm.org/10.1145/2166966.2167013>.

FENG, Y.; ZHUANG, Y.; PAN, Y. Music information retrieval by detecting mood via computational media aesthetics. In: **Proceedings IEEE/WIC International Conference on Web Intelligence (WI 2003)**. [S.l.: s.n.], 2003. p. 235–241.

FORST, J. **Last Chart! - See the Data**. 2017. Available at: <http://www.lastchart.com>. Accessed: January 2017.

FOURNEY, D. W.; FELS, D. I. Creating access to music through visualization. In: **Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference**. [S.l.: s.n.], 2009. p. 939–944.

FOXALL, B. **Last.fm to CSV**. 2017. Available at: <http://benjaminbenben.com/lastfm-to-csv>. Accessed: January 2017.

GILKS, P. **What's Peter been listening to?** 2017. Available at: <http://public.tableau.com/s/gallery/whats-peter-been-listening>. Accessed: January 2017.

GRATZL, S. et al. Lineup: Visual analysis of multi-attribute rankings. **IEEE Transactions on Visualization and Computer Graphics**, v. 19, n. 12, p. 2277–2286, Dec 2013.

GUEDES, L. S.; FREITAS, C. M. D. S. Interactive music rankings visualization. In: **18th Symposium on Virtual and Augmented Reality - SVR 2016**. [S.l.: s.n.], 2016. p. 79–80.

HAUGER, D.; SCHEDL, M. Music tweet map: A browsing interface to explore the microblogosphere of music. In: **2016 14th International Workshop on Content-Based Multimedia Indexing (CBMI)**. [S.l.: s.n.], 2016. p. 1–4.

HIRAGA, R.; WATANABE, F.; FUJISHIRO, I. Music learning through visualization. In: **Web Delivering of Music, 2002. WEDELMUSIC 2002. Proceedings. Second International Conference on**. [S.l.: s.n.], 2002. p. 101–108.

HOLM, J.; SIIRTOLA, H. A comparison of methods for visualizing musical genres. In: IEEE. **2012 16th International Conference on Information Visualisation**. [S.l.], 2012. p. 636–645.

HOLM, J.; SIIRTOLA, H.; LAAKSONEN, L. Associating avatars with musical genres. In: IEEE. **2010 14th International Conference Information Visualisation**. [S.l.], 2010. p. 186–193.

JOHNSON, B.; SHNEIDERMAN, B. Treemaps: a space-filling approach to the visualization of hierarchical information structures. In: **Proc. 2nd International Visualization Conference 1991. IEEE**. [S.l.: s.n.], 1991. p. 284–291.

LAST.FM. **Listen to free music and watch videos**. 2017. Available at: <http://www.last.fm/>. Accessed: January 2017.

LEE, D. L.; DENG, L. Y.; LIU, Y.-J. Visualization for interest in music based on plurk social network. In: IEEE. **IT Convergence and Security (ICITCS), 2013 International Conference on**. [S.l.], 2013. p. 1–4.

LEHTINIEMI, A.; OJALA, J.; VÄÄNÄNEN, K. Socially augmented music discovery with collaborative playlists and mood pictures. **Interacting with Computers**, Oxford University Press, 2016.

LIBEKS, J.; TURNBULL, D. You can judge an artist by an album cover: using images for music annotation. **IEEE Multimedia**, Citeseer, v. 18, n. 4, p. 30–37, 2011.

LIIKKANEN, L. A.; ÅMAN, P. Shuffling services: Current trends in interacting with digital music. **Interacting with Computers**, Oxford University Press, p. iwv004, 2015.

NGUYEN, D. Q.; LE, D. D. Hyper word clouds: A visualization technique for last.fm data and relationships examination. In: **Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication**. New York, NY, USA: ACM, 2016. (IMCOM '16), p. 66:1–66:7. Disponível em: <http://doi.acm.org/10.1145/2857546.2857613>.

ONO, J. H. P. **Visualização de similaridades em bases de dados de música**. Tese (Doutorado) — Universidade de São Paulo, 2015.

PLURK. **Plurk**. 2017. Available at: <http://www.plurk.com/>. Accessed: January 2017.

POLYGRAPH. **How music taste evolved**. 2017. Available at: <http://polygraph.cool/history/>. Accessed: January 2017.

REINGOLD, E. M.; TILFORD, J. S. Tidier drawings of trees. **IEEE Transactions on Software Engineering**, IEEE, n. 2, p. 223–228, 1981.

SPOTIFY. **Music for everyone.** 2017. Available at: <http://www.spotify.com>. Accessed: January 2017.

SPOTIFY. **Spotify Charts**. 2017. Available at: <htp://www.spotifycharts.com>. Accessed: January 2017.

STASKO, J.; ZHANG, E. Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In: **Proceedings of the IEEE Symposium on Information Visualization 2000**. Washington, DC, USA: IEEE Computer Society, 2000. (INFOVIS ’00), p. 57–. Disponível em: <http://dl.acm.org/citation.cfm?id=857190.857683>.

TWITTER. **What’s happening?** 2017. Available at: <https://twitter.com/>. Accessed: January 2017.

VALVERDE-REBAZA, J. et al. Music genre classification using traditional and relational approaches. In: **Intelligent Systems (BRACIS), 2014 Brazilian Conference on**. [S.l.: s.n.], 2014. p. 259–264.

WEERKAMP, W.; TSAGKIAS, M.; RIJKE, M. de. Inside the world’s playlist. In: **Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management**. New York, NY, USA: ACM, 2013. (CIKM ’13), p. 2501–2504. Disponível em: <http://doi.acm.org/10.1145/2505515.2508216>.

YOUTUBE. **Youtube**. 2017. Available at: <http://www.youtube.com>. Accessed: January 2017.

ZHANG, J.; LIU, D. Visualization of user interests in online music services. In: **Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on**. [S.l.: s.n.], 2014. p. 1–6.

ZHU, J.; LU, L. Perceptual visualization of a music collection. In: IEEE. **2005 IEEE International Conference on Multimedia and Expo**. [S.l.], 2005. p. 1058–1061.

APPENDIX A — REQUIREMENT ANALYSIS SURVEY

Musical Charts Visualization Research

Hello!

This is an academic research about Music, Perception and data visualization of music charts. Your sincerity in the responses is extremely important for our work and all data will be used exclusively for purposes of this master's thesis work. The poll does not require data that may lead to the identification of the respondent.

The research is being conducted by Master Student Leandro Guedes, under the orientation of PhD. Carla Freitas, of the Graduate Program in Computer Science - UFRGS (Federal University of Rio Grande do Sul) - Brazil.

Thank you for your participation!
E-mail: lsguedes@inf.ufrgs.br

* Required

Characterization:

Gender *

- ◯ Female
- ◯ Male

Age (only numbers) *

e.g.: 27

Your answer

Occupation *

Your answer

Nationality *

Choose

Academic Degree *

Choose the higher level you completed

- ◯ Post-Doc
- ◯ PhD
- ◯ MsC
- ◯ Undergrad
- ◯ High School
- ◯ Elementary School

What are your preferred services to listen to music? *

- ☐ Apple Music
- ☐ CD/DVD/Blu-Ray
- ☐ Deezer
- ☐ Download
- ☐ Google Music
- ☐ Grooveshark
- ☐ Last.fm
- ☐ Musicmatch
- ☐ Pandora

- [ ] Radio
- [ ] SoundCloud
- [ ] Spotify
- [ ] TIDAL
- [ ] Youtube

How do you usually discover new music? *

- [ ] Same as last answer
- [ ] Radio
- [ ] Friends' recommendation
- [ ] Shows
- [ ] Music Charts
- [ ] Forums

What are your favorite music genres? *

- [ ] Alternative
- [ ] Blues
- [ ] Children's
- [ ] Classic
- [ ] Country
- [ ] Dance
- [ ] Electronic

- [ ] Gospel
- [ ] Heavy Metal
- [ ] Hip Hop/Rap
- [ ] Instrumental
- [ ] Jazz
- [ ] Pop
- [ ] Punk
- [ ] R&B
- [ ] Reggae
- [ ] Rock
- [ ] Other: ____________

Research

What is the level of influence that the ARTIST has in your choice for new music? *

	1	2	3	4	5	
Does not influence	○	○	○	○	○	Influence a lot

What is the level of influence that a CHART has in your choice for new music? *

	1	2	3	4	5	
Does not influence	○	○	○	○	○	Influence a lot

What is the level of influence that the MUSIC STYLE has in your choice for new music? *

	1	2	3	4	5	
Does not influence	○	○	○	○	○	Influence a lot

What is the level of influence that the RELEASE DATE has in your choice for new music? *

	1	2	3	4	5	
Does not influence	○	○	○	○	○	Influence a lot

When you are interested in music charts, where do you look for? *

- ☐ BBC Charts
- ☐ Billboard Charts
- ☐ iTunes Store
- ☐ Last.fm
- ☐ Official Charts
- ☐ Shazam Charts
- ☐ Spotify Charts
- ☐ Top40 Charts
- ☐ I do not search for music charts
- ☐ Other: ____________

Among the items below, which do you think are the most interesting for a music charts visualization website? *

- [ ] Total number of executions per track
- [ ] Total number of executions per artist
- [ ] Music recommendation by tracks
- [ ] Music recommendation by artists
- [ ] Cumulative ranking for a given time range
- [ ] Geographic information of listeners
- [ ] Age information from listeners
- [ ] Gender information from listeners

What do you expect from a website with music information and charts visualization?

Your answer

If you have any further comments or other suggestions, please use the space below:

Your answer

APPENDIX B — USER STUDY QUESTIONNAIRE - CHARACTERIZATION

Music data visualization survey

FREE AND EXCLUDED CONSENT TERM

You are being invited to participate in a survey on data visualization from musical rankings. Please read this document carefully and clarify any doubts before consenting to your participation.

Objective: This research aims to evaluate an interaction of users with different visualizations and musical data.

Procedures: The participant must answer all the next 5 sections, analyzing current visualization solutions and the proposal of this work of integrating techniques of data visualization coming from music rankings.

The total experiment time will be about 15 minutes.

Participants can, without any detriment and at any time, interrupt the test, if they want to.

Thank you,
Leandro Guedes and Carla Freitas.

* Required

I agree to take the test: *

- ○ Yes
- ○ No

NEXT

Page 1 of 6

Music data visualization survey

* Required

User Characterization

Age *
(only numbers)

Your answer

Gender *

- [ ] Female
- [ ] Male

Academic Degree *
(choose the higher level you completed)

- [ ] Post-Doc
- [ ] PhD
- [ ] MsC
- [ ] Undergrad
- [ ] High School
- [ ] Elementary School

Main activity area *

- [ ] Computer Science / Computer Engineering / Information Technology
- [ ] Engineering / Mathematics / Physics / Chemistry
- [ ] Humanities or Social Sciences
- [ ] Biological and Health sciences
- [ ] Other: ____________

Main Occupation *

- [ ] High School Student
- [ ] Undergrad Student
- [ ] Master Student
- [ ] PhD student
- [ ] Teacher
- [ ] Professional in company of its area of formation
- [ ] Professional in different area training company
- [ ] Other: ____________

Experience as a user of interactive systems *

(check all that apply)

- [ ] Yes, with the usual web systems (buying sites, social networks, banking websites, etc.)
- [ ] Yes, with systems required for my professional activities
- [ ] Yes, with computer games

☐ No, I just browse and query the web

What is the importance of listening to music for you? *

	1	2	3	4	5	
Not important	○	○	○	○	○	Very important

Have you ever looked for the popularity of a song or artist? *

○ Yes

○ No

Have you followed any music ranking? *

○ Yes

○ No

Do you know Billboard or Spotify Charts? *

○ Yes, both

○ Yes, just Billboard

○ Yes, just Spotify Charts

○ No

Have you used Last.fm? *

○ Yes

○ No

Are you familiar with data visualization? *

- Yes
- No

BACK NEXT

Page 2 of 6

APPENDIX C — USER STUDY QUESTIONNAIRE - TOOL EVALUATION

Music data visualization survey

Analysis of current systems

Review the following websites, from which we use data for our tool:

[1] http://www.billboard.com/charts/hot-100

[2] http://www.spotifycharts.com

After reviewing [1] and [2], proceed to the next section.

BACK NEXT Page 3 of 6

Music data visualization survey

Using the tool

Access and use all functionalities (Visualization, Filtering and Comparison):

Http://192.168.204.143:8888/MusicVis/

After you use it, go to the next section.

BACK NEXT

Page 4 of 6

Music data visualization survey

* Required

Questionnaire about the tool

This questionnaire aims at a specific evaluation of the MusicVis technique.

Statement must be answered using a 5-point Likert scale: "Strongly Disagree," "Disagree," "Neutral," "Agree," and "Strongly Agree."

I think it is an interactive tool. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the chosen visualizations good. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found that the proposed visualizations allow a better understanding of the ranking than the ones used by Billboard. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found that the proposed visualizations allow a better understanding of the ranking than the ones used by Spotify Charts. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the main page selection menu suitable for selecting the visualizations. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found that the colors of musical genres were pleasant. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I liked the Sunburst visualization. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I liked the Node-Link Tree visualization. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I liked the Bubble visualization. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Disagree

I liked the Treemap visualization. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I liked the data search on the filtering tab. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I liked the data comparison in the comparison tab. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think the options for visualizing, filtering and comparing data are suitable. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the layout of the tool attractive. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the tool interesting. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the tool relevant. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think the visualizations could be applied to other areas. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

If you have any suggestions or comments, please write below.

Your answer

Please, go to the last tab :)

BACK NEXT

Page 5 of 6

APPENDIX D — USER STUDY QUESTIONNAIRE - SUS

Music data visualization survey

* Required

Usability questionnaire

This questionnaire aims at an overall usability assessment of the MusicVis technique.

Statement must be answered using a 5-point Likert scale: "Strongly Disagree," "Disagree," "Neutral," "Agree," and "Strongly Agree."

I think that I would like to use this system frequently. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system unnecessarily complex. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought the system was easy to use. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think that I would need the support of a technical person to be able to use this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the various functions in this system were well integrated. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought there was too much inconsistency in this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I would imagine that most people would learn to use this system very quickly. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system very cumbersome to use. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I felt very confident using the system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I needed to learn a lot of things before I could get going with this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

BACK SUBMIT

Page 6 of 6