Music Atlas (@music_atlas)

Debjyoti Paul, Nishant Agarwal, Shweta Singhal,

November 10, 2016

1 Basic Info

Team Members:

Name	UID	Email
Debjyoti Paul	u0992708	deb@cs.utah.edu
Nishant Agarwal	u1010232	nishant@cs.utah.edu
Shweta Singhal	u1011542	shweta@cs.utah.edu

1.1 Code Repository:

https://github.com/MusicAtlas/musicatlas

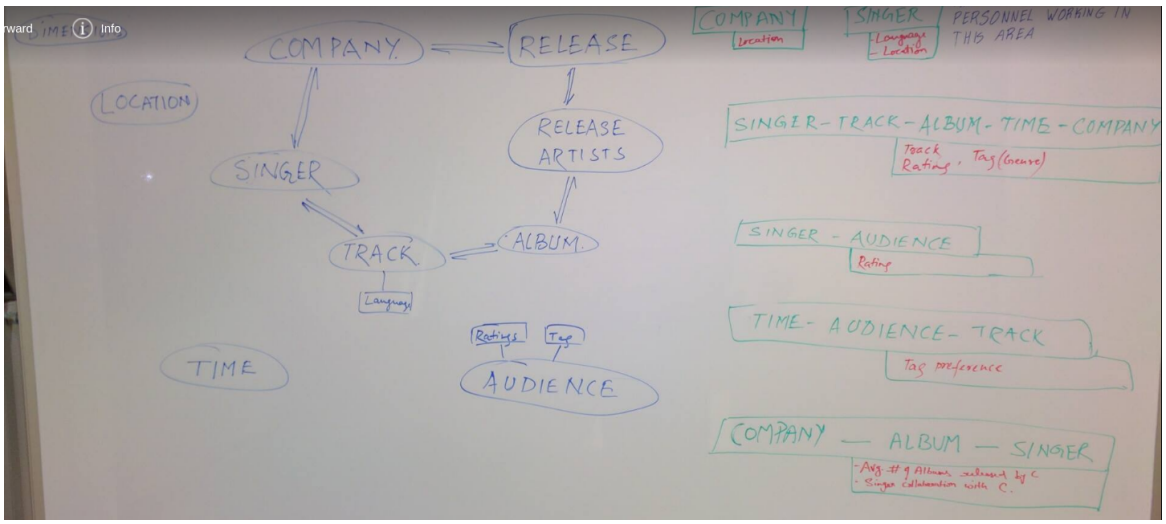
2 Background and Motivation

History shows music has always been a source of inspiration for the creative minds by the creative minds. Last couple of decades with the advent of information technology and social networking music has reached from one corner of the world to other. Exchange of such aesthetic medium has produced phenomenal creation. Recently some websites shows music related information, they are related to latest hit songs lacking any form of exploration feature. Our aim for this project is to build visualization for exploration and analysis of information related to music. Music enthusiasts want to explore songs of diverse taste. Our visualization will enable the end users (music enthusiast) to explore and learn facts about music of different geographical areas, cultures and genres. This will make an ultimate search experience for a music enthusiast. Currently this kind of information can only be accessed after an extensive google search. For example it is elusive to find any information about an artist or a genre based on locations. We would like to present MusicAtlas (@music_atlas) as a one-stop exploration and analysis tool for Music lovers.

3 Data

We are using the MusicBrainz open source database to answer interesting questions about the music world. We have collected the data by following the instructions provided by MusicBrainz for setting up the Postgres database of entire dataset.
Database: https://musicbrainz.org/doc/MusicBrainz_Database
Server Repository: https://bitbucket.org/lalinsky/mbslave

Some of the dimensions from MusicBrainz database we are using are artists, company/label, tracks, time, release, albums etc.



4 Project Objectives

The aim of this project is to provide user interface to educate, explore and analyze information related to music. The user should have the flexibility to explore the music, company(label/Recording Studio), artist or release data and see what the trend is with respect to location. We also want to show artist history, popularity etc. to see how the trend is changing by changing the parameter, the same could be observed for tracks or company. The musicbrainz dataset also has time dimension attached to various entities, we want to leverage that to see the timeline trend for given entity (like company, artist, track etc).

We would like to answer questions like which how many releases a particular artist made, which language he prefers for his song, what is the release details of a track, what genre was most favored by artists. how the genre favored changed over time based on the tracks released in a year. These are some questions which a music lover would like to know. These questions could also help them learn more about the music industry and trends.

5 Data Processing

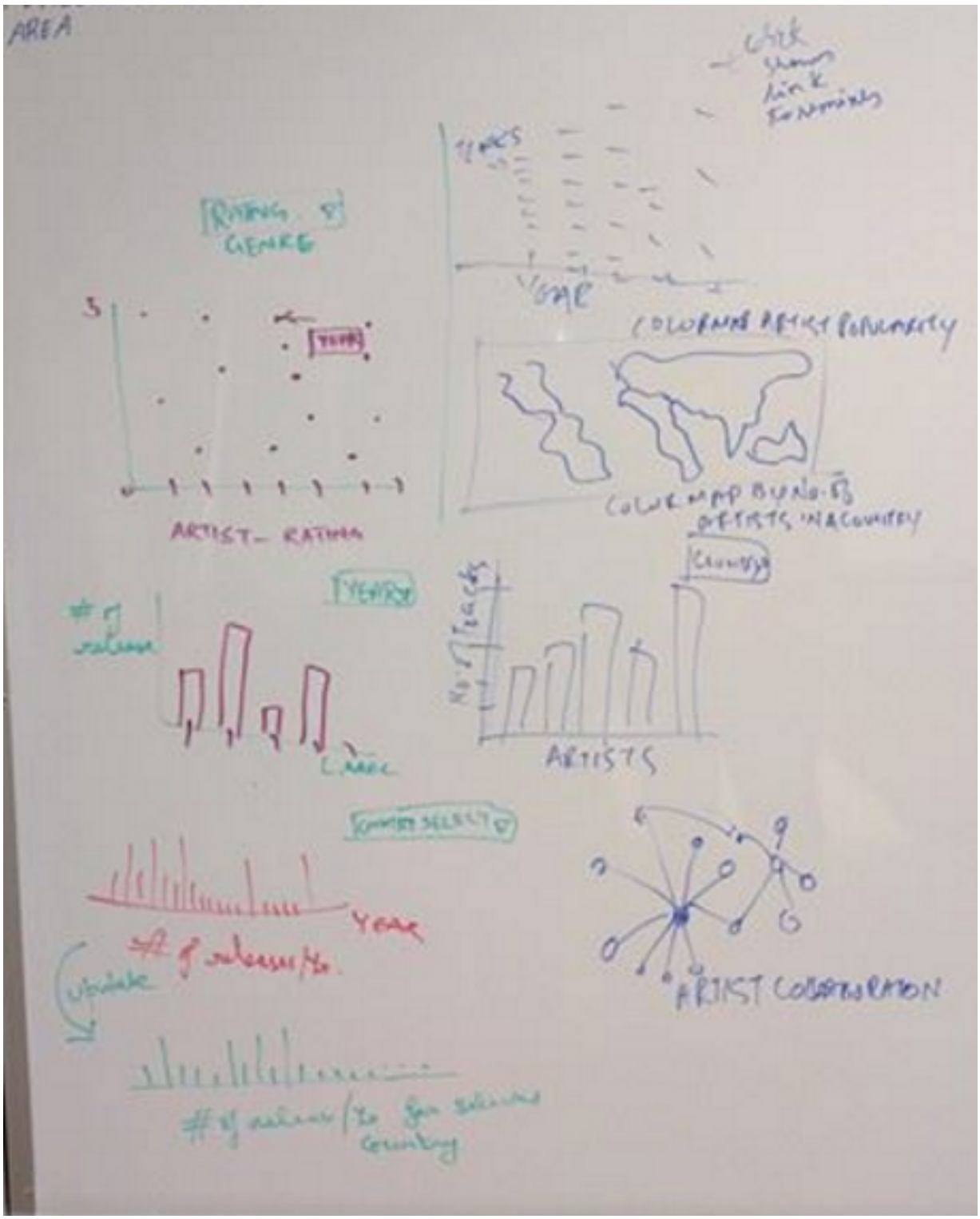
The dataset which we have chosen is huge(about 50-60 Gb) and in the form of relational tables which are linked with each other using keys. Also, the tables are huge with millions of entries in almost each entity tables. Also, the tables have various fields or meta-data information about the entities which we might not be using for our visualization. We will be using Artist, Company(Label), Release, Track, User(audience/listener) as main tables, from which we will be querying relevant and data and exporting them in CSV/JSON format.

We first figured out the dimensions which we will be using in the visualization process and then we have sorted out the relevant attributes/facts which will be interesting to analyze. From the database we have observed that year information or timeline is scarce among. Hence, we have filtered the dataset which has timeline infomation. With the exception that empty column with timeline information in Labels table could mean that company still exists and releasing new albums and launching new artists.

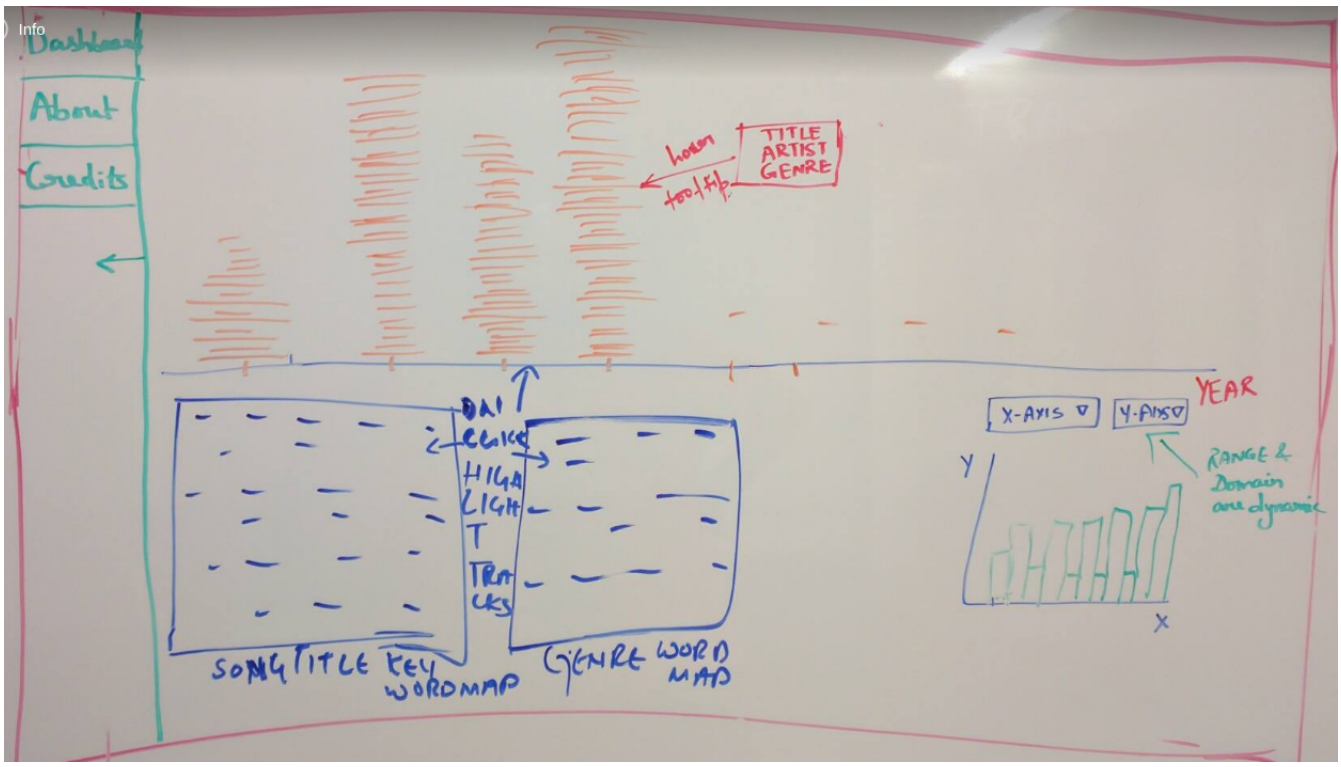
To produce our analysis we will be making queries to multiple tables and also join them to find the correlation among them (like we are planning to showcase the artist collaboration with company, for that purpose we will be joining the artist and company table with the release table which will result the releases that artist has done with the given label company and so on). For analysis purpose, we will be using the User taggings, annotations for the songs, artists and companies and also will analyze the tag preference according the edit history of the user. Artists origin, language, genre preference could be interesting in analysis, and the same applies to label companies based upon location.

Using the given dataset in hand, we expect to see the artists genre interest with his/her collaboration with the label along with that his popularity region charts too. Also, we are expecting various interesting results regarding the shift in users music taste.

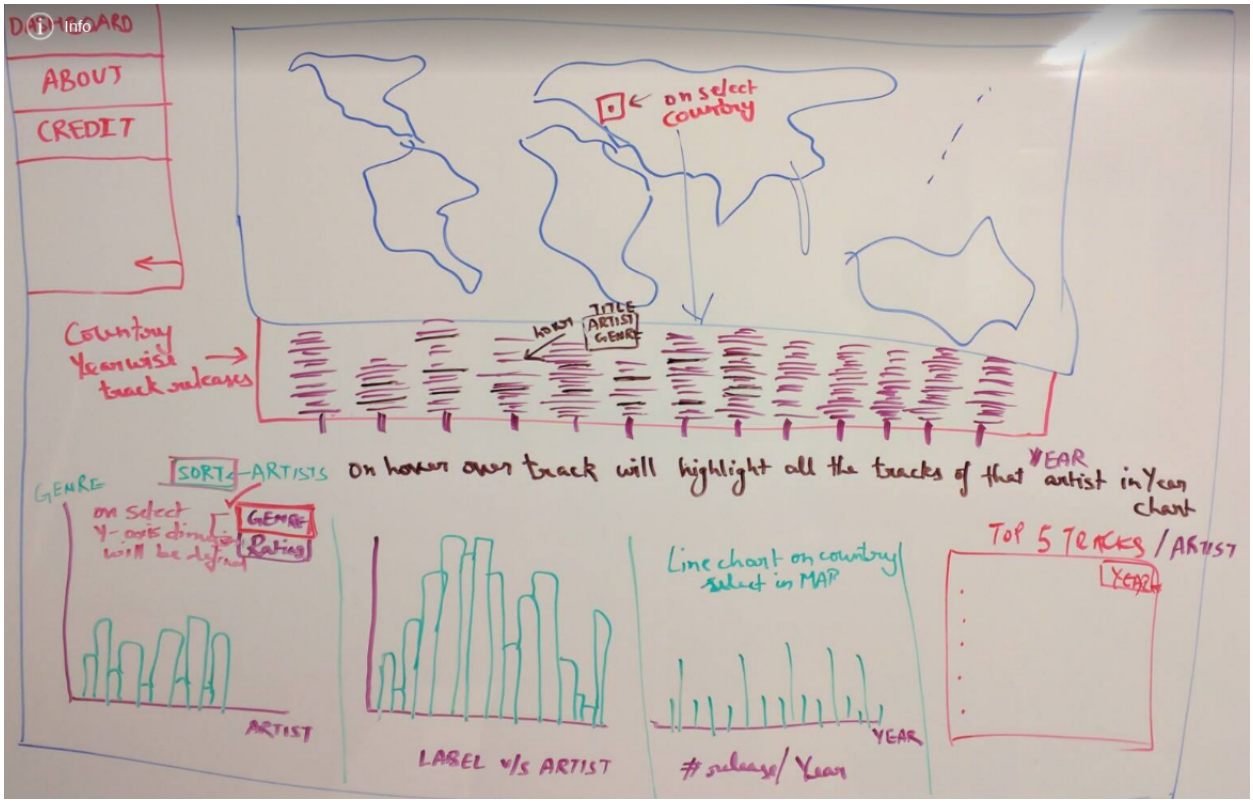
6 Visualization Designs



This chart is the initial layout which shows multiple SVG's fulfilling different purposes. First, the Artist rating is shown in scatter-plot which on hover display the year in which that rating is marked. Second, On selecting the year from the drop-down, the bar chart will display the number of releases by the label companies in that year, by default one year details will be displayed. Third, On country select from the dropdown on the top-right side of that SVG, the number of releases done by any label aggregate is displayed with respect to year. By default the aggregate releases are displayed per year in line chart. Fourth, Track year scatterplot will be displayed in the top right side, on click, will show the link to the remix(but was complicated to get). Fifth, colored map with number of aritsts in country in world map and colorscale is proportional to the number of artists. Sixth, number of tracks released by artist in bar chart and at last the artists force map in node chart.



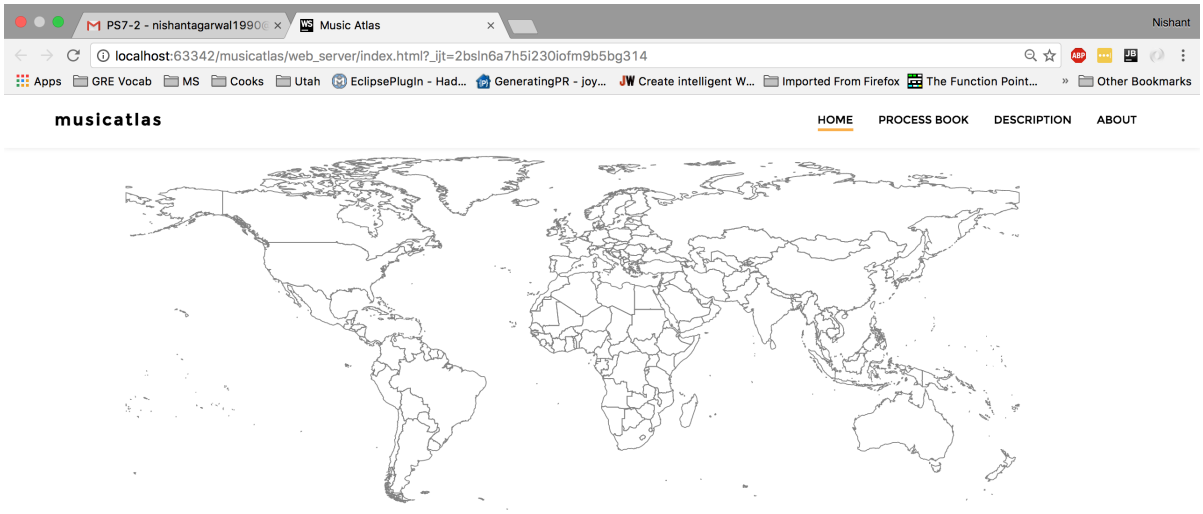
This design is more refined as compared to first design in managing the charts properly by providing the comparison axis selection in the right side in drop-down to render the below bar chart whose range and domain are dynamically loaded according to selection criteria in drop-down. Left-panel will have a Dashboard, About and Credits list which is collapsible. The main chart in this visualization is the year with track chart, where on hover on any track the tool-tip will be displayed with Full track title, artist and genre information and the track will be highlighted. Below, the song-title keyword mao and Genre word map which will interact with the year-track chart mentioned above.



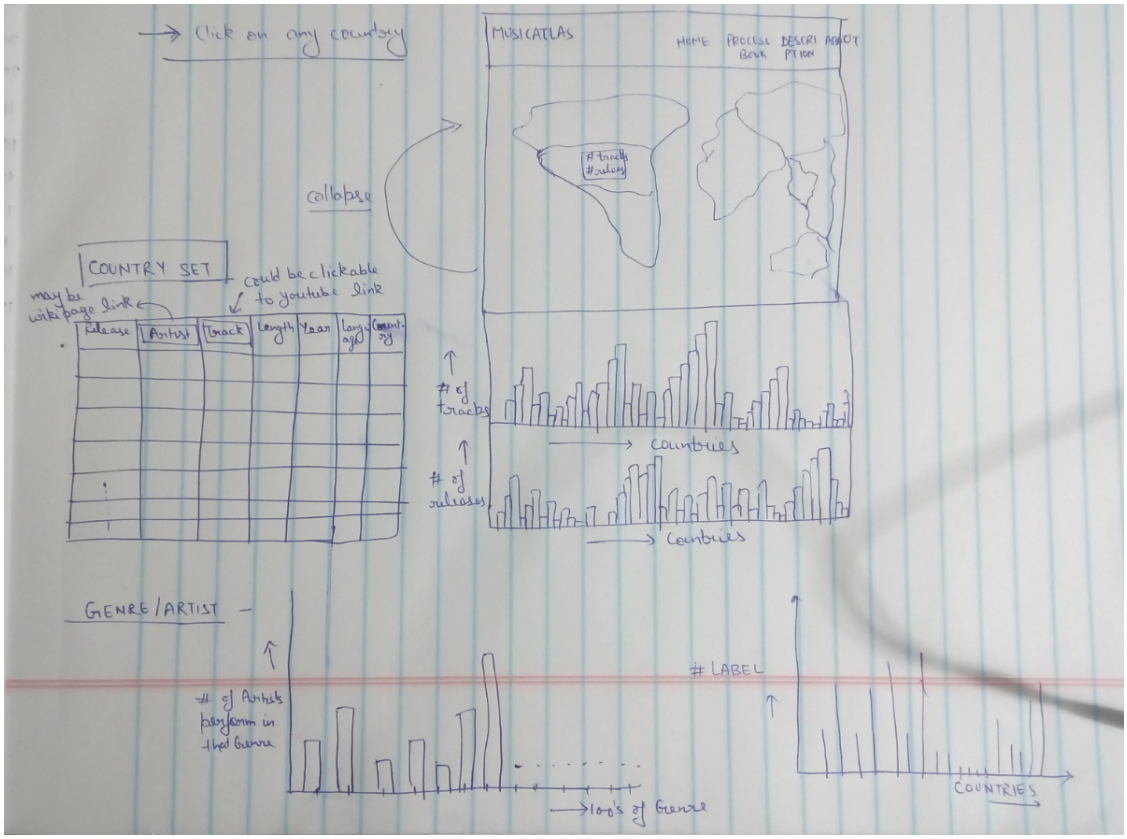
This design uses a map where countries are colored differently based on the number of tracks released. On selecting a country we display the tracks based on year below it. On clicking a track all tracks by that particular artist will get highlighted. Along with it we would show bar chart comparing artist based on genre and ratings, number of artists per label, number of releases per year for the selected country and the top 5 artists or tracks of that selected country. This is all aimed at providing the user with a burst of information which helps him/her decide on the kind of music they would possibly want to hear.We could also show a top 5 based on genre or other things. The ideas are quite infant at present.

This design incorporates the best of the previous design to the best of our understanding. We would use a choropleth map displaying based on number of tracks per country. The bars below it is the representation of the number of tracks per country. On clicking the bar we would display the data in the charts below. Our display would show labels vs artists. A multi-dimension or choice bar chart for artist vs genre/popularity. We are yet to decide based on data if it would be possible to show a timeline for an artist. This data provides a good representation of the information and would amuse a music enthusiast. They would be able to see most sought after labels by artists, the genres most popular by artists etc.

Current design :



Expected final design :



7 Must-Have Features

- User-artist and User-song chart with Genre
- Artist collaboration with label company
- Artist popularity chart
- Search on the basis of artists, company label, country
- Sorting on tracks title, artists, company labels (if data available)

8 Optional Features

- Node map to display the connections between company to multiple artists with their tracks
- Timeline trend of the artist popularity
- User Genre preference timeline trend
- Brush implementation and zooming animation on timeline chart of artists

9 Peer Review

Our team got review from one of the team who was really enthusiast about this visualization. Our earlier plan as proposed in the process book above was to show the artist and song history with respect to location and genre, but after review we get to know people are more interested in knowing more about artists(may be their history from wikipage whom they don't know but have listened to their songs) or making the tracks clickable which will redirect the user to the youtube page for that song.
Also, in review our reviewer suggested to show the tooltip when hover over any country which shows any information countable into the tooltip (like # of tracks relesed , # of artists born, # of label companies, # releases produced etc), which we are considering to go with. Apart from that the bar chart under the map which shows the # of tracks released per country involves the country name as labels on X-axis, which could be a issue, for which we will discuss with our mentor and work on, and the same applies for # of releases/country chart and # of label companies/country.

10 Project Schedule

Team Members:

Week	Date	Deadlines: Debjyoti Paul	Deadlines: Nishant Agarwal	Deadlines: Shweta Singhal
1	Oct 28	Data export		Data identify and clean
2	Nov 4	Base Structure	Explore Visualization	Analysis and design
3	Nov 11	World Map	Web site designing	Bootstrapping and Data filtering
4	Nov 18	Interaction world map and	Dynamic bar chart	Tabular data display and content
5	Nov 25	Making track searchable	Making artists search to wiki pages	Continue previous..
6	Dec 2	Project Due		