

BACKGROUND AND MOTIVATION

We plan to develop an interactive visualization of sales data for key music formats (CD, cassette, vinyl, digital download, streaming, etc.) in the United States over the years. We think it would be interesting to explore how music consumption has changed from generation to generation and how evolving music format technologies have influenced that. Lastly, with many business articles reporting that the music industry is in decline, it would also be interesting to verify if this is the case or not based on trends and patterns in the past.

PROJECT OBJECTIVES

The primary questions we are trying to answer with our visualization are:

- How has the popularity of various music formats changed throughout the years?
- Have overall sales in music been decreasing gradually over time?
- Have new and evolving music format technologies hurt or helped music sales?

What we would like to learn is if any trends or patterns exist in the music sales dataset. In particular, the things we would like to accomplish include:

- ranking the various music formats to find out which ones dominated over different periods of time
- finding the peaks and valleys in music sales over time
- discovering when formats increased and decreased in popularity
- finding the greatest changes in sales throughout each year
- discovering any resurgences of older music formats (e.g. vinyl, cassette)

The benefits include:

- being able to compare sales between various music formats over time
- being able to compare music sales of different time periods
- being able to compare music sale dollar values accounting and not accounting for inflation

DATA

The United States music sales dataset comes from the Recording Industry Association of America (RIAA) Shipment Database. The database is accessible through the following login page: https://www.riaa.com/keystatistics.php?content_selector=riaa-shipment-database-log-in. The RIAA has provided us access to their database for academic use. The database provides an option to export the dataset as an Excel spreadsheet, which is how we are collecting the data.

DATA PROCESSING

We do not expect to do substantial data cleanup. We plan to save the Excel spreadsheet exported from the RIAA database as CSV files and process the data using the D3 CSV parsing API.

The dataset includes year-end sales statistics from 1973 to 2014 for the recorded music industry in the United States. The statistics comprise of both the number of units and dollar values sold for key music formats (CD, cassette, vinyl, digital download, streaming, etc.).

The quantities we plan to derive from the number of units and dollar values data include:

- the aggregate amounts over a specified time period
- the differences in amounts between two specified time periods
- the changes in amounts over each year

VISUALIZATION

The visualization will comprise of multiple coordinated charts that link together so that when a user interacts with one of the views, the others dynamically update through animated transitions. We will display the music sales data using line, stacked area, and bar charts.

MUST-HAVE FEATURES

The must-have features of the visualization include:

- toggling the displayed sale metric data (units, dollar values, dollar values adjusted for inflation) through selection
- ranking the sales of each music format through sorting
- focusing on the sales of a subset of music formats through filtering
- focusing on the sales of a specified time period through brushing
- viewing the changes in sales for each music format over time through sliding of the brushed selection

OPTIONAL FEATURES

The optional features of the visualization include:

- toggling between a line chart and stacked area chart
- toggling between displaying raw sales data and derived data, which is the change in sales over each year
- comparing sales between different time periods through multiple brush selections

PROJECT SCHEDULE

The following project schedule lists the objectives for each week leading up to the final project deadline.

Weekly Objectives

By 4/10/2015

-

By 4/17/2015

-

By 4/24/2015

-

By 5/4/2015

-