

Choropleth map exploration of the WASABI music dataset



Quentin Le Roux

January 18, 2021

Introduction

Why a choropleth map for music data

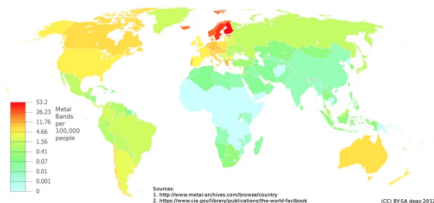
- A Choropleth map displays geographical areas that are coloured in relation to a variable. It allows to highlight patterns across the displayed location.
- Very common to see in news media, such as in music
- Example in The Atlantic:

THE ATLANTIC

Annotations

A World Map of Metal Bands Per Capita

DINO GRANDONI MARCH 29, 2012



Implementation

Presentation of the choropleth map implementation

- **Goal:** reproduce & expand the visualization found in news media
- We want to show the countries with the most country bands per capita, **per genres and decades since the 1960s.**

Let's launch the visualization!

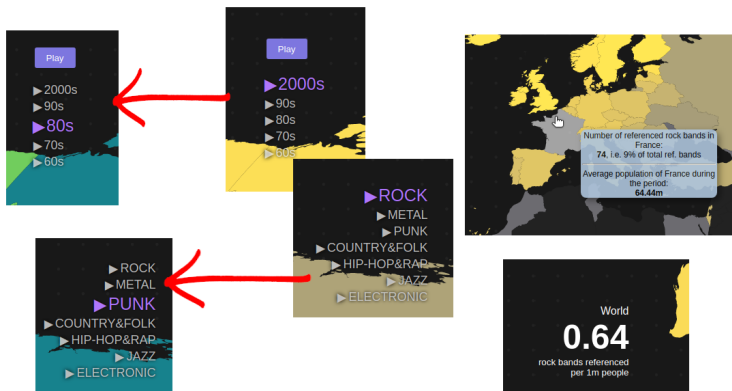
On a terminal:

```
$ git clone https://github.com/LMquentinLR/choropleth_wasabi_dataset.git
$ sudo apt install npm
$ npm i
$ sudo npm install -g parcel-bundler
$ parcel index.html
```

Possible interactions with the map

What can it do?

Tooltips per countries! Interactivity! Animation!



How does it work?

Folder Structure

```
.
├── css
│   └── base.css
├── data
│   ├── genre-summary
│   │   └── JSON with tree-like struct: {genre:{decade:{band number}}}
│   ├── music-data
│   │   └── JSON with tree-like struct: { country:{ genre:{ decade:{ band number, population}}}}
│   └── music-world-data
│       └── JSON with tree-like struct: {country:{ decade:{band number}}}
├── js
│   └── main.js
├── index.html
└── README.md
```

Superfluous folders and files omitted

How does it work?

Data Consumption by D3.js

Start the server

└─> D3.js function inside main.js

└─> fetch(music-data)

└─> fetch(genre-summary)

└─> fetch(music-world-data)

└─> render()←──────────┐

└─> user interaction (click, zoom, etc.)

WASABI dataset. Tables used:

- ALBUMS, used attributes: _ID, LOCATION.COUNTRY
- ARTISTS, used attributes: ID_ARTIST, GENRE, PUBLICATIONDATE

World Bank population dataset

- Used attributes: YEAR, COUNTRY.CODE, COUNTRY.NAME

Data processing

Processing Pipeline - 1

1. Simplify GENRES in the ARTISTS table by generalizing the values (e.g. 'punk rock', 'math rock', etc. → 'rock')
2. Remove rows not sorted into genre families: "rock", "metal", "punk", "country/folk", "hiphop/rap", "jazz", "electro"
3. Lower-case all string data
4. Transform all PUBLICATIONDATE (i.e. years) into DECADES (e.g. {1980, ..., 1989} → 1980)
5. ALBUMS and ARTISTS tables are joined via the keys _ID and ID_ARTIST
6. LOCATION.COUNTRY and COUNTRY.NAME are standardized and used as keys to join the WASABI with the WORLD BANK tables
7. The resulting table is grouped by COUNTRIES, DECADES, and GENRES

We dump the data in three different JSON files with the following tree-like structure:

- **genre-summary:**
`{GENRE:{DECADE:{# BANDS}}}`
- **music-data:**
`{COUNTRY:{GENRE:{DECADE:{POP.,# BANDS}}}}`
- **music-world-data:**
`{COUNTRY:{DECADE:{# BANDS}}}`

The goal of such file structures is to minimize the number of computations performed by the D3.js rendering function, which renders the choropleth map. The cost is the inclusion of 2 additional JSON weighting a total of 20Kb.

To summarize, a user can use the Choropleth map to perform the following:

- Explore band concentration per country, genre, and decade
- Access a tooltip by hovering on a country that provides demographic and market share data on a specific genre and decade
- Play a scrolling animation over the decade range (from the 1960s to the 2000s) for each genre

This can represent a **complex number of interactions**. As such, we can be **interested in evaluating how people would react to them**.

- 3 steps:
 - ▶ Welcoming, briefing, and interview of the participants (with a questionnaire)
 - ▶ The participants discover the application and must go through a series of monitored tasks
 - ▶ Debriefing and interview of the participants (questionnaire and end survey)
- 8 tasks to be performed. Example:
 - ▶ The participant was successful in the 8 tasks
 - ▶ Find the average population for the United States of America in the 2000s
- Result: 85 out of 100 on the System Usability Scale (a pretty good score)

- **The Atlantic:**
www.theatlantic.com/culture/archive/2012/03/world-map-metal-band-population-density/329913/
- **World Bank:** <https://data.worldbank.org/indicator/SP.POP.TOTL>