Data visualization: Personal report

Joris LIMONIER

January 4, 2022

Contents

1 Joris **2**
1.1 Introduction 2
1.2 Data pre-processing 2

2 Homework for 20/10/21 **3**

3 UX Protocol **4**
3.1 Presentation & training 4
3.2 User test 5
3.3 Debriefing 5

1 Joris

1.1 Introduction

My submission for the project of data visualization consists of a Sankey diagram. It is available here [1] and should be viewed in the browser (tested on Chrome), on a computer. The code for the project is posted on GitHub.
Sankey diagrams are defined as follows:"Sankey diagrams are a type of flow diagram in which the width of the arrows is proportional to the flow rate."

The diagram was created with the following structure:

- Python for

Let us first detail the data preprocessing part.

1.2 Data pre-processing

[1] If the link is dead, just go to https://jorislimonier.github.io/, navigate to "Projects", then look for "Collaborative Data Visualization"

2 Homework for 20/10/21

A paragraph describing the users.

Users could be people insterested in the music industry, who want to compare men and women artists for sociological interpretation.

The list of visual tasks supported by users and the visualization goals.

User task	Details
Overview	Flow between columns
Zoom	TBD
Filter	Filter by genres

The list of (raw) attributes you will need from the WASABI dataset you are going to use.

The following attributes will be needed:

- Album field
  - _id (album id)
  - id_artist
- Artist field
  - _id (artist id)
  - type (*e.g.* "Person", "Orchestra", "Group", "Choir", "Other" or "")
  - gender (male, female, unknown)
  - members (check this one, it may be the name of the members of a band)
- Song field
  - id_album
  - genre

The informal description of the processing of the raw data in order to make it to fit in the visualization technique. This might include calculated variables you must add in the process.

Clustering Artist - *type* variable to make *is_band* boolean variable

The name of visualization technique and the name of the member of the group who is going to implement it. Associate the visualization technique with the visual goal.

Sankey diagram with the following columns:

- Single artist or Band
- Male, Female, Unknown
- Number of Albums
- Number of Songs

A visual mapping of variables available in your data set (after data processing) and the visual variable available in the visualization technique you have chosen.

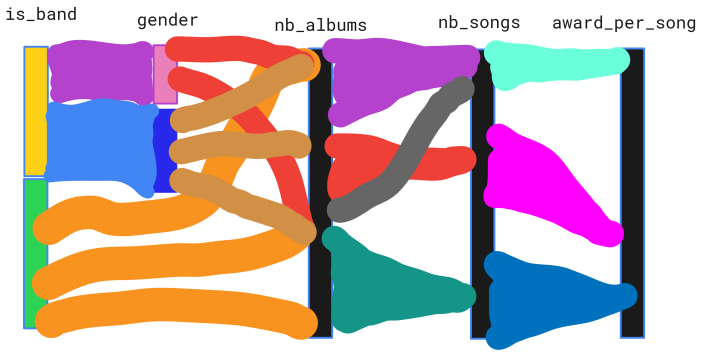

Figure 1: Representation of the Sankey diragram (JL)

3 UX Protocol

TODO:

1. Write text that will be read (to remove bias in the way I say things)
2. Test app on some people (minimum 3)

3.1 Presentation & training

1. Ask written consent for recording
2. Tell people why they are here and what we will do

3. Give the following information:

   Age

   Sex

   Highest level of education reached

4. Present Sankey diagram and explain what a Sankey diagram is.

5. Is anything unclear?

3.2 User test

1. Task 1

   (a) How many men made 5 albums.

   (b) On a scale from 1 to 5 (1 means very easy, 5 means very difficult), how hard was that task?

2. Task 2

   (a) How many rap solo artists are women.

   (b) On a scale from 1 to 5 (1 means very easy, 5 means very difficult), how hard was that task?

3. Task 3

   (a) Can you tell how many solo artists made 3 albums? Why/Why not?

   (b) On a scale from 1 to 5 (1 means very easy, 5 means very difficult), how hard was that task?

3.3 Debriefing

1. What are your three favorite feature?

2. What is your three least favorite feature?

3. Would you recommend this application to a friend?

4. What would you do differently?