

Hadi Kachmar

Data visualization project

Prof. Marco Winckler

The Project is based on the Wassabi dataset. There will be 2 types of visualization, The first one is a bubble map that will show the distribution of the artists according to their genre, while the second one is a funnel chart that will show the number of artists belong to a specific genre in order for a chosen country.

In order to implement the visualization, we need to go through few steps to deal with the data then implement the visualization. The data processing part is purely done using Rstudio.

In the Pre-processing, after fetching the dataset, only few columns will be considered to work with, The columns are [genres, location.country, id_artist_deezer]. The genres column is important for both visualizations, but this column contain some missing data, and some genres names are confusing, so to deal with, the best thing is to create a new column to create sub-classes for the genre, and fill the missing value with a string "Missing value" to be shown in the visualization later on. Next, "location.country" column is important for the bubble map visualization, this visualization will be based on the continent, not per country. To deal with that a new column for the continents can be added based on the "location.country" column. The column called "id_artist_deezer" will be used to identify each artist, then count the artist number for each genre. The last step in the processing part is to add 2 more columns "lat" and "lng" based on the "continent" column, these 2

columns will be helpful to identify each continent location to show the bubbles correctly on the map.

Once the processing part is done, now comes the implementation of the visualization.

The visualization dashboard will look like this:

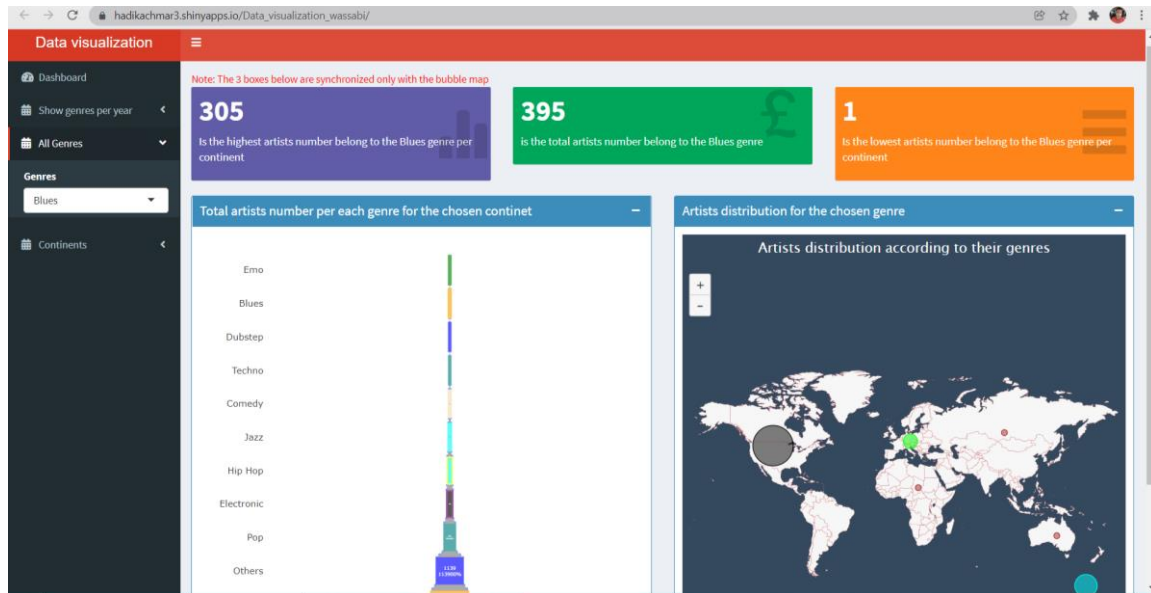


Figure 1: shows the visualization dashboard.

The 3 containers in the top are used to make it more user friendly, and make it easier to spot the changes.

1. The first container is to show the highest artists number belong to the chosen genre per continent
2. The second container is to show the total artists number belong to the chosen genre
3. The third container is to show the lowest artists number belong to the Blues genre per continent.

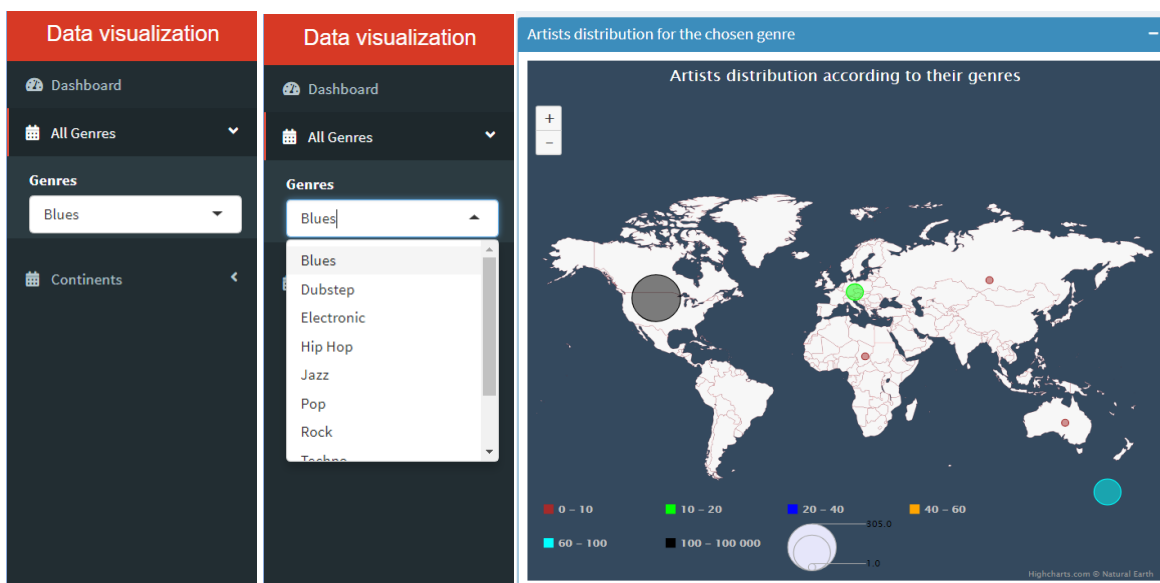


Figure 2-3-4: shows the bubble map visualization.

Talking about the main visualization, the user will be able to choose a specific sub-genre, and the changes will be shown on the bubble map, and in the 3 containers mentioned above, means that the Bubble map visualization is synchronized with the containers. Moreover, the colors palette shown down below in the bubble map, it's toggleable, the user is able to switch on and off, and the changes will be shown on the map. In other words, if the user press on the black color palette, the bubble will be removed from the map, and pressing it again will show it back.

For the funnel chart visualization,

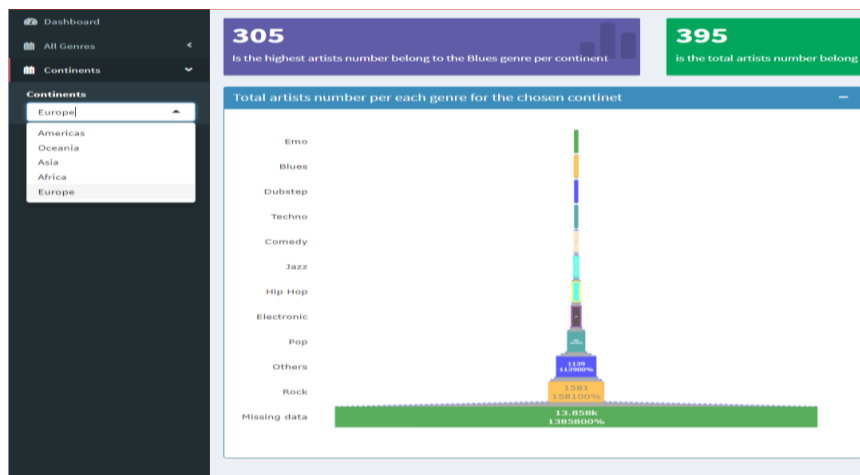


Figure 5: shows the funnel chart visualization.

The user will be able to choose any continent from the continent dropdown button, and the changes will appear on the funnel chart.

Rather not forget to mention about the tooltip as it is a crucial part in the visualization.

Once the user hover on the funnel chart parts, he will be able to see some statistical information, As shown in the figures below.

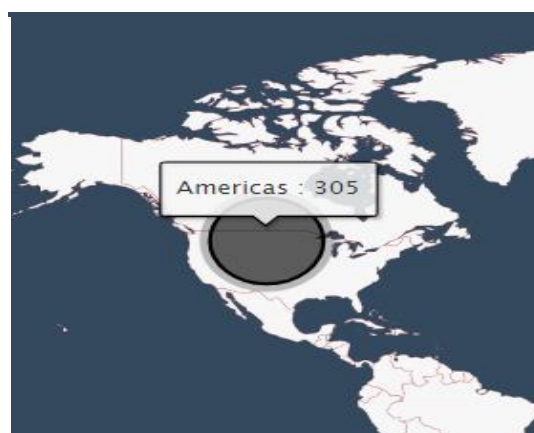
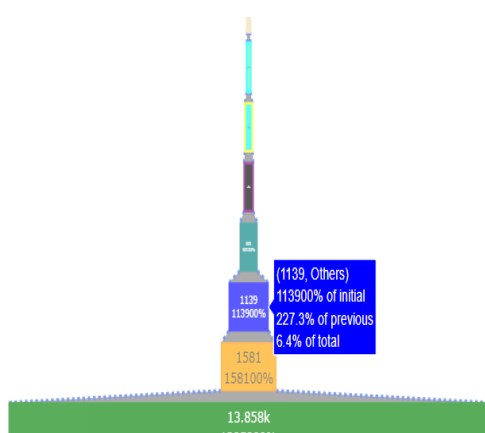


Figure 6-7: shows the tooltips.