Information Visualisation Project Report

HUANG Yidong
Grenoble, France
yidong.huang@etu.univ-grenoble-alpes.fr

January 2021

1 Introduction

In this report, we will briefly discuss the visualisation problem and the visualisation designer before jumping into the visualisation itself. We will discuss how and how well the visualisation answers the question afterwords.

2 Question to answer and the design

The question to answer in this visualisation is the following: "How are the CO2 emission across users." Histogram is the default pick for showing distribution. But due to the particularity of the continues data, showing number of user per emission value will make the graph too fragmented thus unreadable.

The total emission of each user is then transformed into emission categories thus regroups users that has similar emission values together and solve the fragmentation issue. To show an extra dimension of the data, a color coding representing the mode of transportation is used. This allows us to see what mode is responsible for the CO2 emission and thus gives us more insight into the emission composition. To offer both a global view and a more focused local view, 2 sliders were also added in the design allow user to select the range of emission value and how big each emission category is. With the sliders, user can go for a global view with big emission categories, but can also zoom in a specific region for better local insight.

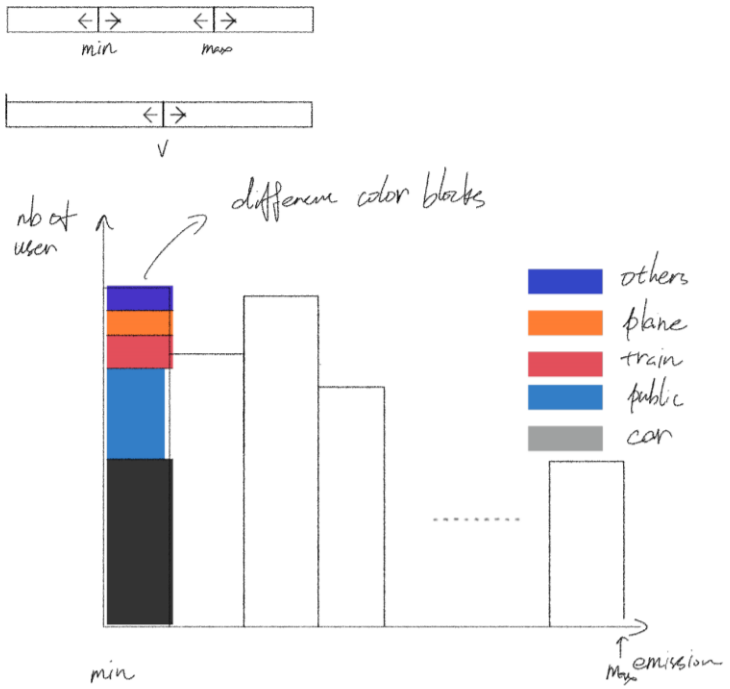

Figure 1: The sketch of the design for the visualisation answering the first question "How are the CO2 emission across users"

3 Implementation of visualisation and interaction

The implementation of this visualisation is done using $C3.js$, a JavaScript library based on $D3.js$. Once the data processing is done and we have the necessary data for each category, we can plot the graph. Another advantage of using $C3.js$ is part of the interaction is already integrated into the graph. We can select two or three desired mode of transportation to further compare their impact. Then the slides for adjusting the range and the category size(named *step* in the implementation) is added so the user can choose the desired value for those parameters.

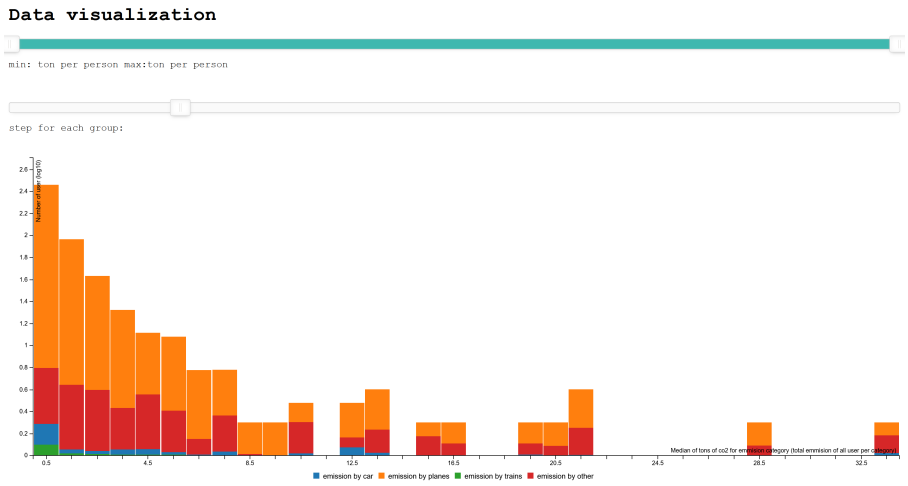

Figure 2: The actual implementation for the visualisation answering the question "How are the CO2 emission across users"

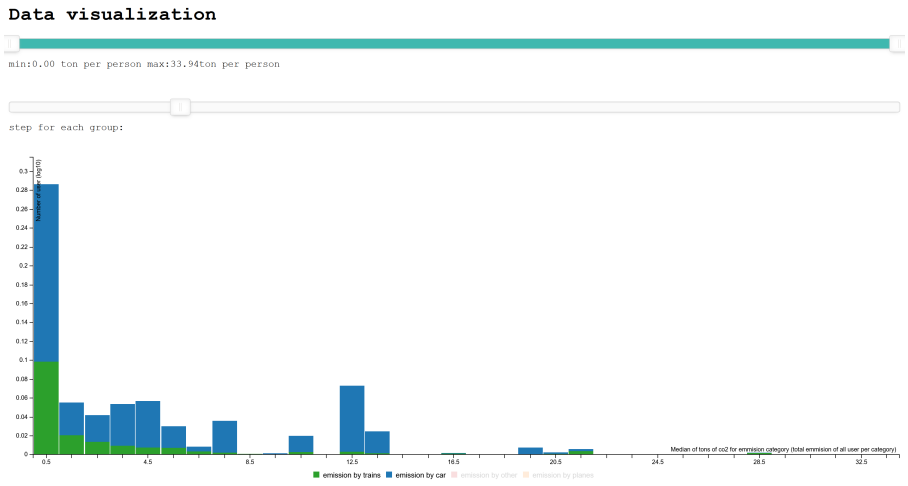

Figure 3: A demonstration of how we can select only 2 mode of transportation to better compare them per category. Here the transportation mode train and car is selected.

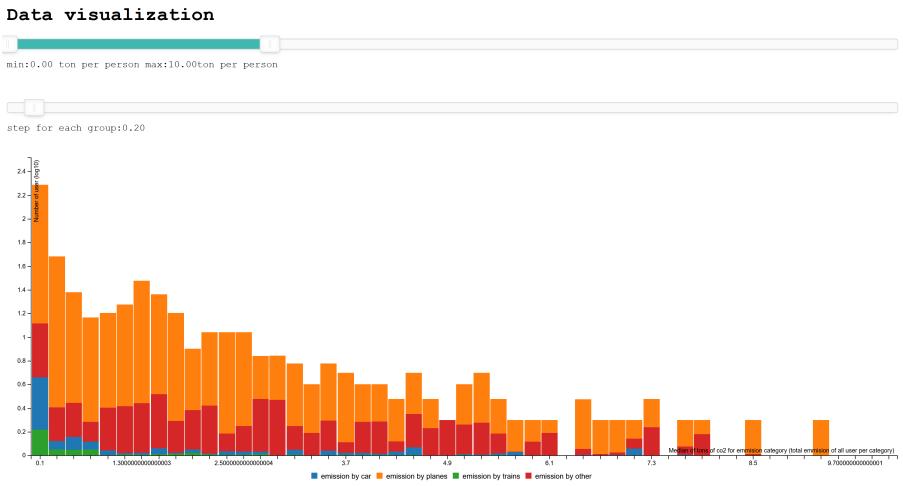

Figure 4: A demonstration of how we can reduce the displayed range for the emission and have a better local view with smaller emission category size.

Noted that the scale for the y axis is not linear. The linear scale was used in the first version of the implementation, but we quickly discovered that due to the distribution of the emission, the number of user per category decreases rapidly. This renders the color coding useless after the first few categories of emission. We then changed the graph to logarithmic scale for better visibility for the color coding.

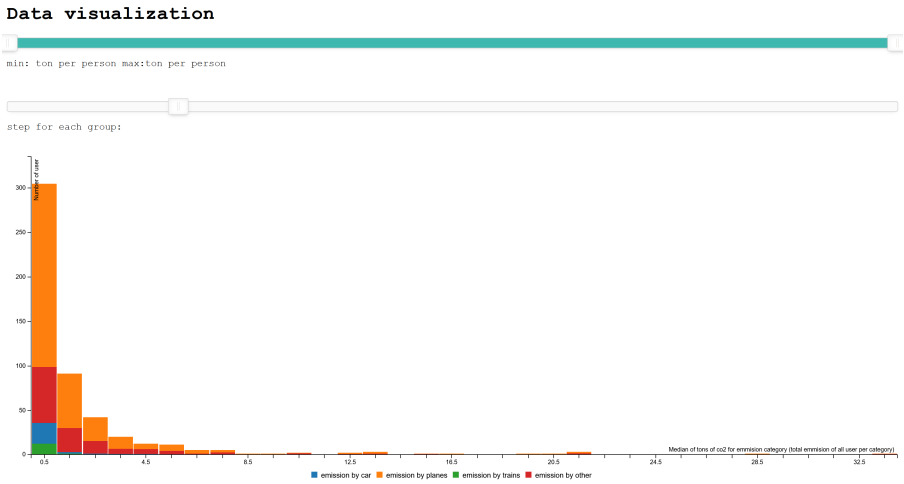

Figure 5: First implementation of the visualisation without using logarithmic scale for y axis. The rapid dropping of number of users per emission category renders the graph unreadable after the first few emission category.

4 Comments on the performance and potential improvements

This visualisation offers us a clear view on the distribution of the emission across users, also a decomposition depending the mode of transportation responsible for the CO2 emission. And thanks to the slides for adjusting the range and category size, user can have both a global and a local view of the distribution. For instance if we set the category size to the minimum and zoom on the lowest emission categories, we can observe that the lowest emission category have the lowest portion of plane emission. And the portion for plane grows gradually as we move up to higher emission category. We may conclude that plane can be a big contributor of CO2 emission for people with low emission level. Other methods of interaction can be easily added to the visualisation if desire. A drop down menu that let user choose what color codes represents(such as institution, rank or destination continent) can be the next step for example.

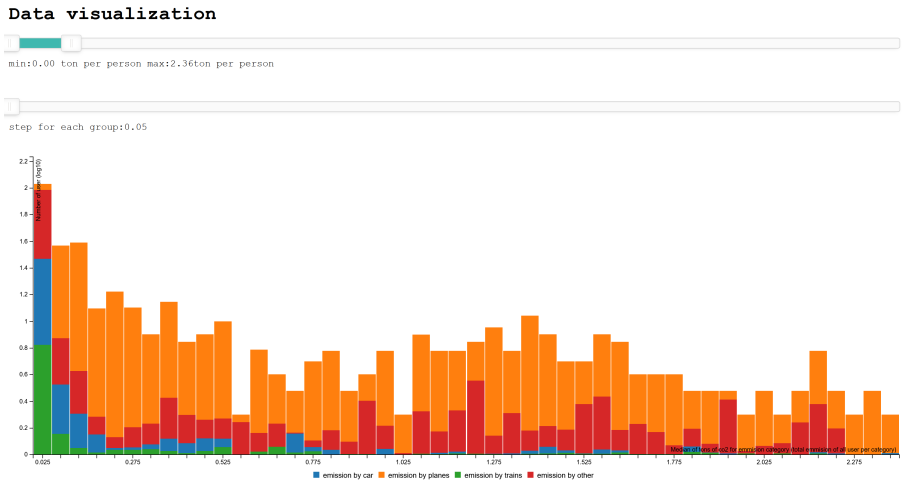

Figure 6: The range displayed here is limited to the lower emission categories and the size of the emission category is set to the minimum. We can see the portion representing plane in orange is piratically none existent.

However there still are some flaws for this implementation. First the color coding comparison can only show us the difference among transportation modes in one category, across category comparison does not offer any value since here the height of each bar represents the number of users and not the emission. This can cause confusion and might misled the user to draw incorrect conclusion if the user didn't notice the previous limitation.

5 Conclusion

As a prove of concept implementation for the visualization, this version led us to a deeper understanding of the adventage and disadvantage of the design. This visualisation design might be more suited for the other question I choose: "Does the rank/title/institution has an impact on the emmision" since in this case the emission level will be mapped to the y axis thus offer cross category compression for the color coding. In both cases the design is very similar, but we can see how the design might be more advantageous to answer the second question.

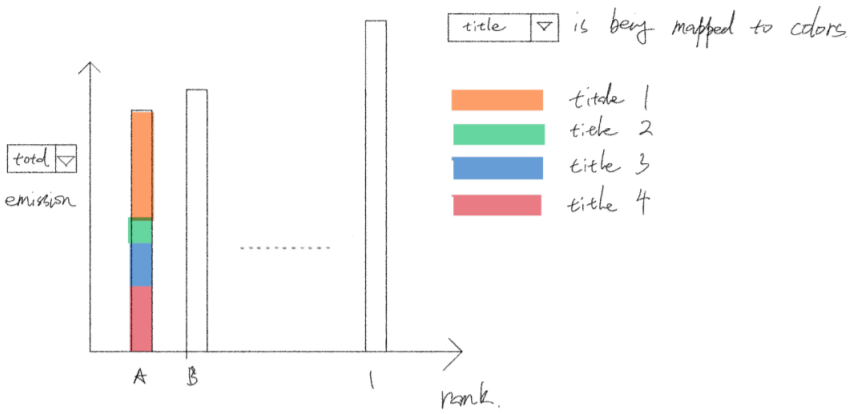

Figure 7: The sketch for the design to answer the second question "Does the rank/title/institution has an impact on the emmision". In this case the emission level is mapped to the y axis, the color coding will help the comparison across categories.