Visualisation of highly uneven distributions: making sense of inequalities

Renaud Blanch

November 22, 2022

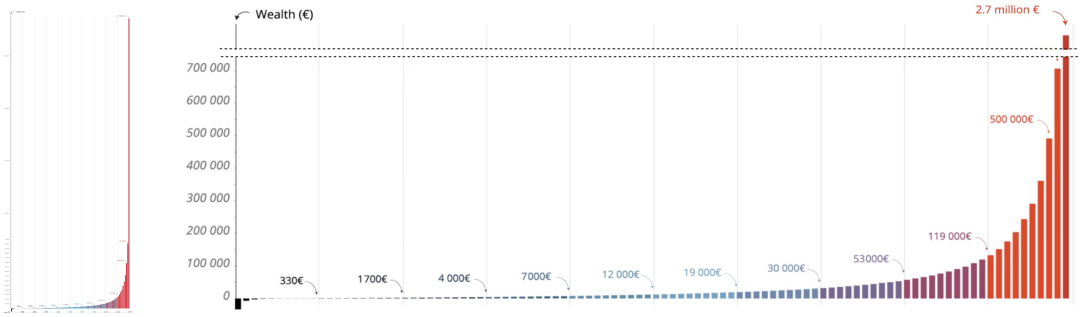

Figure 1: Wealth distribution among percentiles of world population, (left) regular, and (right) broken vertical axis (source World inequality report 2022 from the World inequality database).

1 Abstract

Highly uneven distributions are inherently hard to represent since they involve capturing different orders of magnitudes on the same visualisation. Common solutions to deal with that problem are: using log scales, using broken axis, etc. Their main drawback is that while they help make the visualisation readable, they tend to mask the extent of the inequalities. For exemple, on figure 1 (right), the low values of the distribution are much more readable than on figure 1 (left), but on that first visualisation, it is hard to tell how much higher the latest percentile's wealth is than even the second richest percentile of the population.

This kind of uneven distributions is very common when dealing with data sets related to social and environmental justice (e.g. income and wealth distribution, or greenhouse gas emissions distribution). In those domains, being able to show the data is a way to make people aware of the problems and help them develop their critical judgement. But for that, we need visualisations that do not minimize the phenomenon under scrutiny. We also need visualisation that are easy to understand in order to reach people that may lack the mathematical education needed to understand log axis for example.

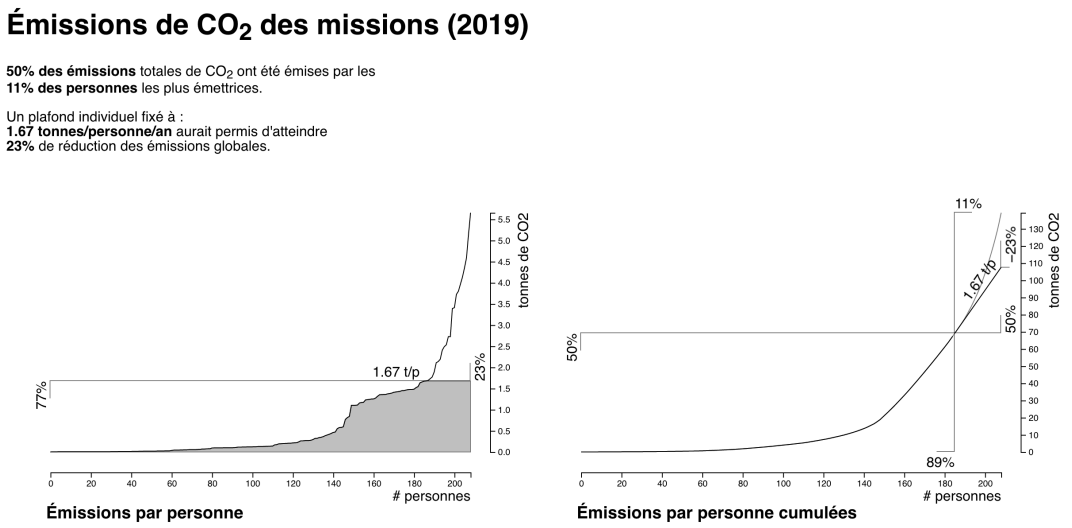

Figure 2: CO_2 emitted by the missions of the member of a research lab for year 2019, (left) by member, and (right) cumulated. <Online interactive version>

2 Work plan

The goal of the internship is to design visualisations for uneven distributions that suit those two requirements: first **make the inequalities visible**, and second do so **without altering their amplitude** so that people can grasp the extent of the inequalities.

The first step of the internship will be to list and study the existing visualisations used for inequal distributions. This work will allow to propose explicit criterions suitable to evaluate and compare the visualisations in terms of legibility and accuracy.

The second step will be to design visualisations that scores as much as possible given those criterions. Using interaction is probably a key to let the user get familiar with the data.

Interaction is also a way to allow users to explore different scenarios, e.g. when using the data to illustrate the effect of a specific policy. Figure 2 show a visualisation of the CO_2 emissions of a research lab induced by the researchers missions for the year 2019. The distribution of those emissions per member of the lab (left) show how they are unevenly distributed among people. Interaction allows to explore the effects of a policy that would enforce an individual ceiling on emissions: a ceiling of 1.67 ton per person would lead to a 23% reduction of the emissions. On the right, we see on the synchronised visualisation of the cumulated emissions that this ceiling impacts only the top 11% emitters (see the interactive version to explore other values of the ceiling).

The last step will be to evaluate the proposed visualisation techniques. In order to do so an experimental protocol will be designed to compare the visualisations and assert how useful they are for users.

3 Research group

The internship will take place in the Engineering Human-Computer Interaction research group of the Grenoble Informatics Laboratory (LIG) under the supervision of Renaud Blanch.
<mailto:blanch@imag.fr>
<http://iihm.imag.fr/en/member/blanch/>