

Economical and political aspects of gender equality

Šimon Ružička, simon.ruzicka@aalto.fi

The latest version of the visualizations can be found on ruzicka02.github.io/gender-equality2025. Alternatively, the source code used to generate it, together with a fixed version is provided in a ZIP archive.

Due to how the Export to HTML feature works in the Python library Plotly, it is best to view this website on a wide (desktop/laptop) screen. I made my effort to make the styling phone-friendly, but the result is still lacking.

In my project, I focus on various aspects of gender inequality, as it is a very broad subject, affecting our lives in different ways. The most discussed issue of inequality is the so-called wage gap, that is how much less do women on average earn compared to men. This gap has two different components; different earnings on the same positions, and lower representation in high-earning jobs, or in other words, inequality in the workspace.

An important theme of this paper will be a deeper analysis of the effects of current gender inequality. While the indicators themselves may be relatively visible, and their social awareness may be high, some of the effects may stay hidden. Income inequality leads to an uneven distribution of wealth and lower financial independence, shaping our society.

Reaching more equal high-earning job representation is also important from a different point of view - having less women in decision-making positions, be it politics or corporate management, generally slows down any meaningful attempts to balance these inequalities. There is commonly a high correlation between high-paying and decision-making jobs, although they are not always equal.

Motivation

I believe that these issues are a topic for the entire society, as the entire society would benefit from addressing these issues, and they should therefore be treated as such. In addition, this topic is still very visible even in our developed “first world”, where hunger or lack of clean water may be hard to relate for us.

Key theses

My approach to this information visualization was to first ask the most elementary and crucial questions:

1. Are these gender equality issues decreasing over time?
2. Is there any significant difference between different countries or regions, and does it correlate for example with GDP per capita or the Human Development Index?
3. Is there a difference between different age groups – are older people more often discriminated based on their gender, or do they discriminate more themselves?

My expected results are that these described issues are getting better over time, although I don't know whether the change is fast enough for meeting any goals set by the United Nations. I would hopefully like to see a meaningful change in this matter in the following decade, so I hope that my final thesis and visualizations will have the potential to convince people to act better in this regard.

Target audience

I believe high/middle school students will learn the scale and effect of the wage gap issue when examining the time and space components of this discrimination, as well as the scale on which this is happening right now, in the world around them. We would know if we asked them about their opinions, how well they understand the issue and how it affects for example the lives of their families. However, this feedback generation is likely out of scope for this course.

In my subjective opinion, providing relevant information to motivated school children is very important. Their minds are commonly more open to new ideas, and they can bring new ideas to their parents and families as well, if they believe something is right. An example of this can be seen in Curitiba, Brazil, where a Recycling for Life program focused strongly on schools, providing incentives to children for recycling. This has created long-lasting habits in their entire families, earning Curitiba the nickname of the Greenest city on Earth.

Approach

I will first search for relevant data sources, trying to answer the posed questions to myself before attempting to create any presentable visualizations. After reaching

satisfying answers, I will sketch my own ideas of visual motives relevant for the topics, while also looking for existing visualizations of these topics for further inspiration.

Data sources

Many Europe-specific statistics were collected by Eurostat, many of those are linked on the page Statistics Explained - Gender statistics. In addition, the European Union has a dedicated European Institute for Gender Equality, providing its own independent statistics such as the Gender Equality Index.

Specifically, the Eurostat dataset earn_gr_gpgr2 contains gender wage gap data for all EU countries excluding Greece, for every year between 2014 and 2023. It shows how the EU27 average has decreased from 15.7 % to 12 % in these 9 years. The dataset sdg_05_50, also from Eurostat, then reports on share of women in national parliaments. Containing data from 2003 to 2024, it reports an increase from 21.1 % to 33.4 % of women.

There are also other data sources relevant for this analysis, which do not necessarily correspond with gender equality issues. I would like to analyze the correlation of the previously mentioned indicators with metrics that indicate the general degree of development of a country, such as GDP per capita, average salary etc.

Answers on theses

After analyzing the mentioned datasets, I found a positive trend of change in the past decade or so, in essentially all indicators I observed. This message is also carried in the attached visualizations. It is very subjective to say, whether the change is sufficient or not, and there are also many factors that will determine the future rate of change, which can both increase or decrease.

For some indicators, such as the representation of women in parliaments, there seemed to be a significant correlation between rich and equal countries. It would be very beneficial to do a proper correlation analysis with another metric, as my assumption was rather based on visual understanding. On the other hand, the wage gap did not seem to correlate with the standard separation of rich and poor countries, at least within the context of the EU.

Unfortunately, I did not find a good dataset or even a metric that would help me analyze the third thesis. It can only be assumed that the positive changes over time can roughly correspond to the changing decision-making generations, but it is a trend that is certainly hard to analyze.

Visualization tools

I did all my visualizations in Python, as I am already familiar with that language. I originally expected to do a big part of the heavy lifting with the `matplotlib` visualization backend, as many other plotting Python libraries (for example `seaborn`) use it as well. However, I eventually used the opportunity to learn working with a new library Plotly.

I originally found Plotly useful for special types of visualizations, such as for so-called choropleth visualizations, which are thematic maps that use different shades to represent indicator values across geographic areas. However, I eventually realized that it is beneficial to do any dynamic visualizations using Plotly to produce a coherent presentation, due to the way that the results can be rendered in a webpage using Javascript. With `matplotlib` or anything similar, I would probably have to embed a video, or something significantly less interactive.

Temporal data aspects

Analyzing the gender equality indicator datasets, we notice that both of them report a positive development in the following decades. Still, it is up to discussion, whether the development is fast enough.

Also, it is important to mention that we only have data for the past 10 to 20 years or so. This is a general trend for data monitoring more “modern” issues, such as gender equality, which were not perceived as relevant problems in our society longer time ago. Although the gap was likely even higher than now, nobody has seen the need to collect such data back then, and these datasets are usually collected/guessed retrospectively. In my opinion, doing such analysis and extrapolation would be out of the scope of this project, although the results would undoubtably be interesting.

User interaction

All interactivity of my visualizations will stem from the key theses stated earlier. In order to present, whether the gender equality issues are getting smaller over time, it is crucial to grasp the concept of time well, as described in the following section.

Another key thesis is to determine the difference in gender equality indicators between regions. This can be combined in a single visualization with the time frame representation by an interactive choropleth. That is, a single indicator will

be visualized on a world map for multiple countries, and using data from multiple points in time, the interactive time slider can be added.

Another set of visualizations should also be added to put the scale of gender inequality into perspective. For that purpose, I have produced a dynamic pair of gauges, that reduces the data from all observed countries into a single average, developing in time. For every year, two numbers are displayed – the current value, and the year to year change, leading to a quick understanding on how the world is developing.

Representation of time

Eventually, I have decided to make my main visualizations as dynamic and interactive choropleths, displaying the data for EU countries on a map. They serve the right balance of readability and data density for my target audience. However, throughout my exploration, I considered other forms of visualization as well.

One option was to create a visualization of multiple indicators in a static point in time. For example, correlation between the GDP per capita on x-axis and representation of women in national parliaments on y-axis, with different countries serving as the data points. This visualization will then feature an interactive slider for a point in time, with a “Play” option. This visualization style was popularized by Hans Rosling in his book *Factfulness*, and is often used by his Gapminder initiative, such as on the World Health Chart. For my purpose, I believe that the information density of such a visualization would be too high, so I didn’t do it.

For simpler data with lower dimensionality, it is also feasible to plot the time component on the x-axis, as is common in scientific visualizations. This lets the more experienced viewers to perceive all available data at once, making faster and more rigorous observations about attributes like rate of growth. However, for my purposes, the evolution over time will serve a better purpose for intuitive understanding in the masses.

Representation of space

As was already mentioned, when plotting data for many countries at once, I have decided to use choropleths. In my opinion, this is the most intuitive way of plotting this kind of data, enabling very quick and intuitive comprehension for anyone familiar with the described region. In this kind of visualization, the space on paper/screen corresponds to geographical space (as does in any map). This is intuitive for us, similarly to the dynamic year-to-year data mentioned before.

However, I felt that this kind of a visualization didn't really describe the problematic as a whole. I would say that the choropleths are good to answer the key thesis of whether some regions are better off than others, but not so much for the general improvement over time. It is quite hard to draw some generalized conclusion from it, as the yearly changes in shades are understandably quite small. One can notice a general trend, but how significant is it really? I wanted a clean and simple visualization for the general trend, a single average for every metric, evolving over time.

I have originally intended for it to be a bar chart, but I found a better alternative available in the Plotly library. So-called "gauge chart" is essentially a "curved bar chart", reminiscent of a tachometer found in a car or an airplane. With a bit of creativity, one can imagine the airplane pilot, carefully observing the dozens of gauges, ensuring that the plane is stable. For the intended non-scientific audience, this chart type seemed slightly more playful than plain bar charts, while still being serious enough to do the job.

I have also chosen the colors to represent the stereotypical gender colors; the gauge scale is a bipolar color map from red (women) to blue (men). This colormap is already implemented in Plotly as `RdBu`, or `RdBu_r` for the reversed variant.

Appendix - Visualization screenshots

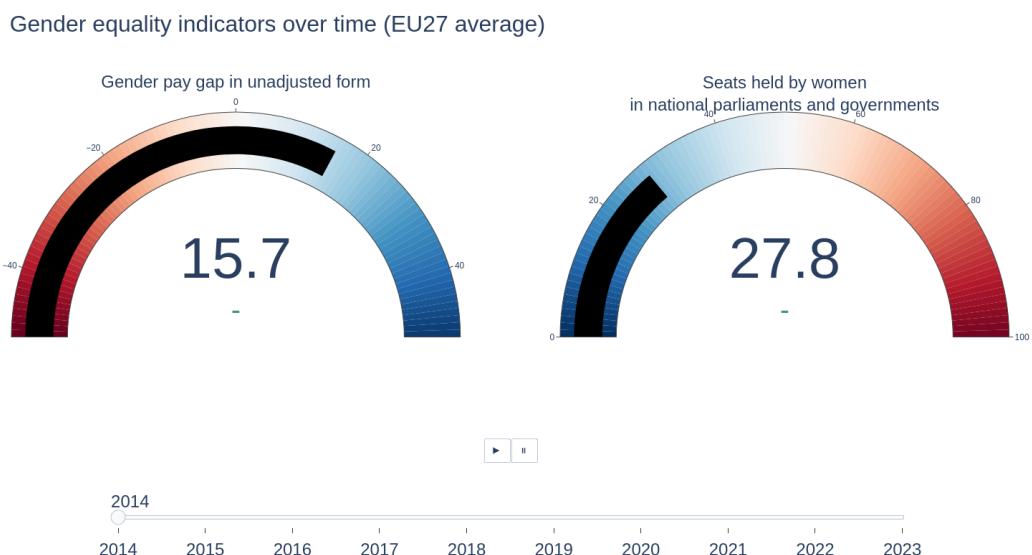


Fig 1: Summarizing gauge chart.

Gender pay gap in unadjusted form (earn_gr_gppr2)

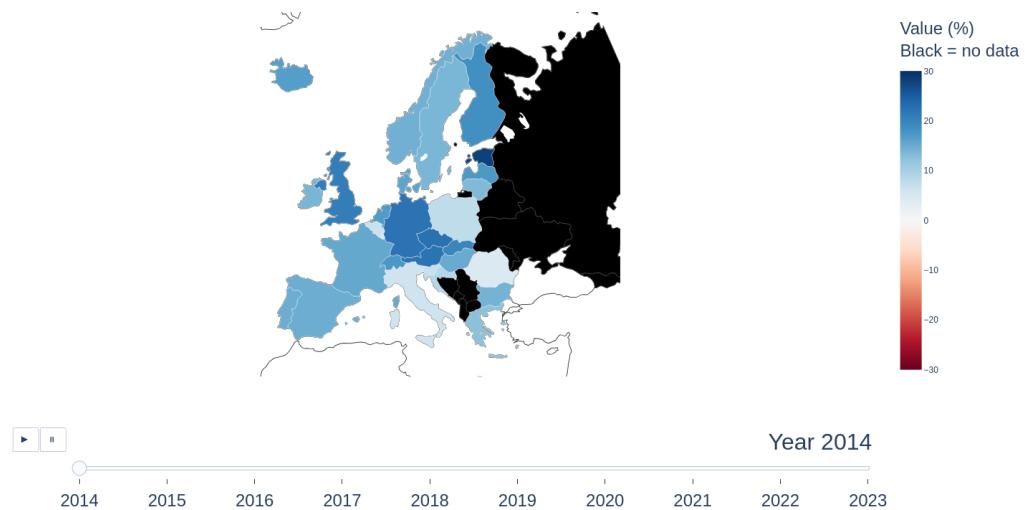


Fig 2: Example of the choropleth.