Paolo Malerba 22/05/2025

Generative AI for Work & Research Productivity Bootcamp 2025 Q2

Workshop 2: Automated Data Analysis & Reporting

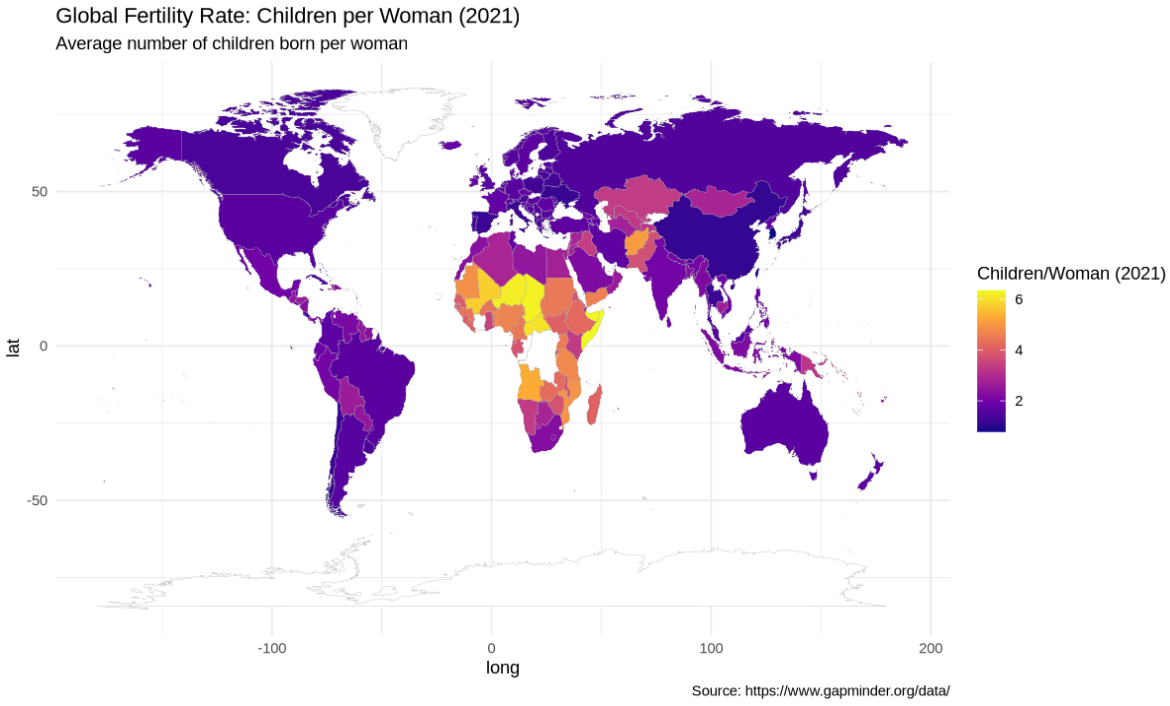
This analysis explores two key global development indicators for the year 2021:

- Fertility rate, measured as the average number of children born per woman, and
- Women's representation in national parliaments, measured as the percentage of seats held by women.

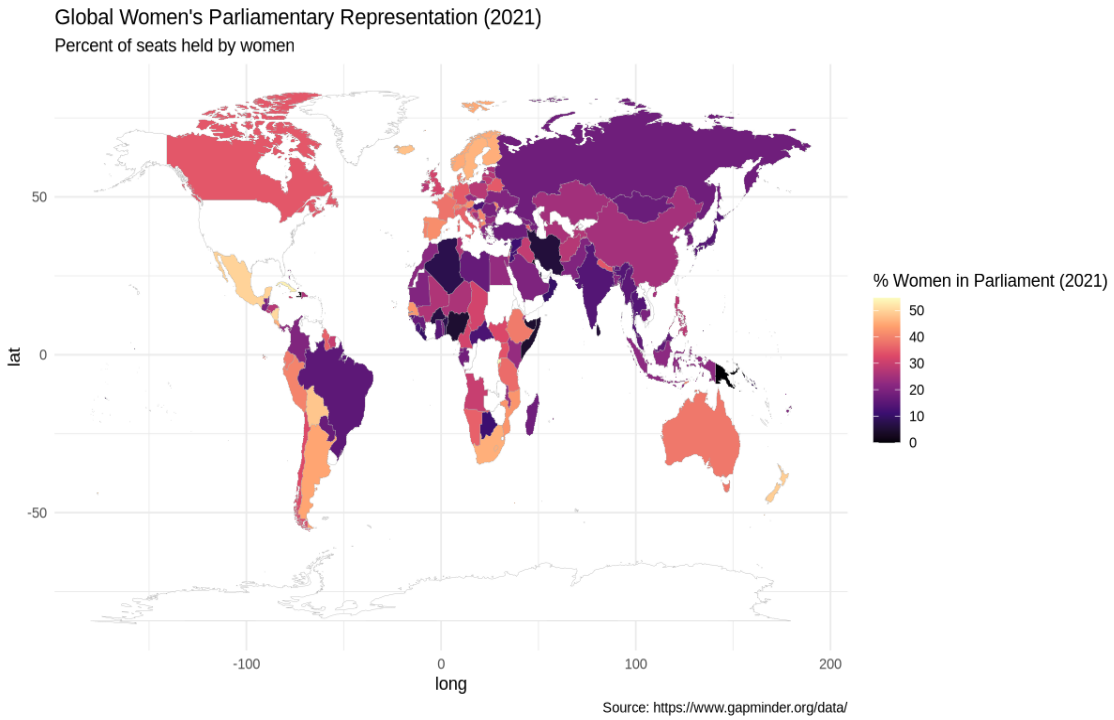
We began by loading and cleaning data from Gapminder for both indicators, then mapped each variable onto a world choropleth to reveal geographic patterns. Fertility rates use a plasma color scale—lighter hues where birth rates are higher, darker where lower. Women's parliamentary representation uses a magma scale, highlighting countries with stronger female legislative presence in deeper tones. White areas denote missing data.

A country-by-country merge allowed us to produce a scatterplot and calculate Pearson's correlation ($r \approx -0.24$), suggesting that higher fertility tends to coincide with lower female parliamentary representation. This introduction sets the stage for a deeper investigation into sociopolitical and economic factors underlying these patterns.

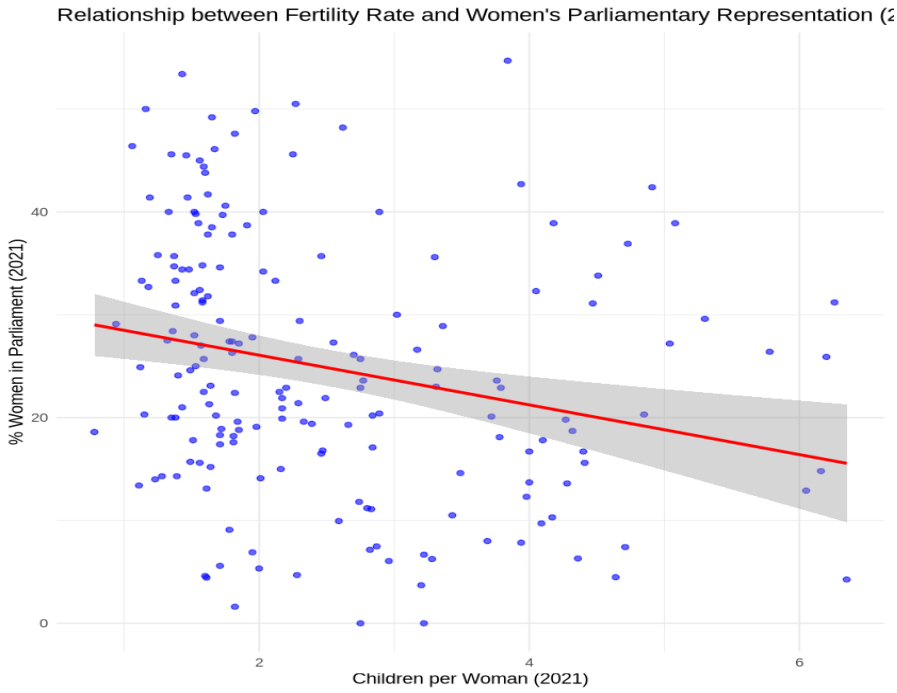
This first choropleth shows the global fertility rate ("Children per woman") in 2021:



This second choropleth shows the percentage of parliamentary seats held by women in 2021:



To examine their relationship, we merged the two indicators into the country and plotted a scatterplot with a fitted linear trend:



Finally, we computed Pearson's correlation coefficient between the two indicators. Interpretation: There is a modest negative correlation ($r \approx -0.243$), indicating that, across countries in 2021, higher fertility rates tend to be associated with lower shares of women in parliament.

Paolo Malerba 22/05/2025

Personal reflection:

This exercise offered me an opportunity to explore the application of AI for data analysis, a task that's pivotal in my daily routine as a researcher. While I'm thrilled about applying this newfound knowledge to my future research, I also experienced some frustration due to my lack of coding skills. Specifically, as I lack the knowledge to fully understand how the choropleths and scatterplot were generated through code, this prevents me from effectively modifying the generated code on my own. This feeling of grappling with tools that I don't fully comprehend evokes in me the impostor syndrome. However, it also motivates me to gain further knowledge and proficiency in using AI tools effectively.