

Data Profile

This section gives a quick overview of the datasets I'm using for the project. I'm highlighting what kinds of data each one includes, how I cleaned and prepped them, and what patterns started to show up right away. It also includes some early checks for consistency and structure to make sure everything's solid before diving into deeper analysis. The goal here is to get familiar with the data, surface anything interesting upfront, and spot any limitations I'll need to keep in mind later.

Fertility Rate Dataset (Our World in Data)

This dataset tracks fertility rates by country over time. It shows how many children are born on average to women in each country, and how that number has changed across decades. I focused on the most recent data (from 2020) to capture the current state of global fertility.

I did a few basic cleaning steps removing regions like "World" or "High-income countries" and standardizing country names. The data is already formatted well, with clear numeric values and consistent country codes.

Early Patterns:

- Many high-income countries are well below the replacement level of 2.1 children per woman.
- Some of the sharpest drops are in places like South Korea, Japan, Italy, and Spain.
- Several African nations still have relatively high fertility rates, though they are also slowly declining.

OECD Family Database (PF3.4 – Childcare Support)

This part of the data zooms in on how much families are actually paying for childcare and whether the support systems in place make it affordable. I used three sheets from the PF3.4 file: gross childcare fees, net cost for two-earner families, and net cost for single-parent families. After cleaning each one, I looked at missing values, data types, and summary stats to get a feel for how the costs vary between countries and family structures.

- The **Gross Fees** sheet gave me a raw look at childcare costs before any financial help is applied. Some countries are clearly unaffordable at this level, making it harder for parents—especially mothers—to stay in the workforce.
- The **Net Cost for Two-Earner Families** sheet was more balanced, reflecting what couples actually pay after subsidies or tax breaks. But even here, affordability still depends a lot on where you live.
- The **Net Cost for Single-Parent Families** sheet was especially revealing. In many places, these families face high costs even after support is factored in, which says a lot about systemic inequality in family policy.

Together, these sheets show that not all government support is created equal. Some countries lighten the load for parents, while others leave them to figure it out alone. Comparing these costs side-by-side helps explain part of the fertility puzzle raising kids is simply too expensive in too many places.

Global Abortion Policies Database – Abortion Policy Data

This dataset adds a whole other layer to the story—one about rights. It tracks whether countries have any laws recognizing abortion across different types of legal frameworks, like constitutions, health codes, or civil laws.

To make the data easier to work with, I cleaned and converted the policy indicators into binary values (1 = legal provision exists, 0 = it doesn't). Then I created a composite flag to show whether any abortion law exists at all in each country. That way, I could see at a glance which countries provide some level of legal recognition—and which don't.

This dataset doesn't measure how accessible abortion is in practice, but it does provide a powerful reference point: **Does a government even acknowledge the right to choose at all?** It helps shift the focus from just economic support to also consider reproductive freedom and bodily autonomy—especially in countries that expect women to have more children without giving them any legal control over their own bodies.

UNDP Human Development Reports - Gender Inequality Index (GII)

This dataset gives context to the deeper gender dynamics that shape motherhood and opportunity. It measures inequality across labor force participation, education, and maternal health basically, all the key areas where gender roles get reinforced.

I cleaned the file to keep only what I needed: male labor force participation, education levels, and rankings that show where countries stand in terms of overall gender equality. I also removed regional headers, footnotes, and blank rows so it would merge smoothly with the rest of the data.

This isn't just about raw numbers, it's about contrast. When countries have high male participation but low female support, it reveals a lot about whose time and labor is truly valued. It helps show how traditional roles still dominate in places that say they're modern and developed.

Georgetown University's WPS Index

This dataset takes things beyond economics and into the lived experience of women how safe they are, whether they're included politically, and whether they have access to justice. It was one of the more complex files to clean, but totally worth it.

I renamed the long column headers so they made more sense in the analysis and dropped extra rows like region averages and footnotes. The final version includes core indicators like legal discrimination, political violence targeting women, intimate partner violence, and representation in parliament.

Where the GII helps compare opportunity, the WPS Index helps compare rights and freedoms. Together, they show how secure or vulnerable women are in different places especially when they become mothers or caretakers. It added depth to the picture I'm building with this project.

[The Cost of Motherhood in a Declining Population](#)

[→ It's fresh, multi-layered, globally relevant, and tells a human story that isn't overdone in portfolios.](#)

The Cost of Motherhood in a Declining Population: What Governments Ask of Women, and What They Refuse to Give"

Working Title:

“The Cost of Motherhood in a Declining Population: What Governments Ask of Women, and What They Refuse to Give”

Possible Analysis Themes:

1. Population Decline by Country/Region

- Trendlines showing fertility rate drops
- Maps comparing population growth vs. decline

2. Government Pressure vs. Policy Support

- Countries pushing pro-natalism
- But lacking support (e.g., paid leave, affordable healthcare, child care access)

3. Economic & Health Cost of Giving Birth

- Cost of labor and delivery by country
- Maternity-related death rates
- Post-birth depression rates or social dropout

4. Opportunity Cost for Women

- Female workforce participation
- Promotion drop-off after motherhood (“motherhood penalty”)
- Gender gaps that widen after having children

Possible Questions You Could Explore:

- Which countries are most aggressively pushing women to have more children?
- How much does it cost—financially and socially—for women to give birth and raise a child in those same countries?
- Do countries with higher pro-natal pressure offer **less** support for mothers?
- Is there a correlation between the cost of giving birth and birth rates?
- What is the “**motherhood penalty**” by country?

Geographic & Time-Based

- Population growth/decline = **time**
- Fertility & maternity support = **geography**
- Healthcare cost and maternal health = **continuous**
- Policy types, gender stats = **categorical**

This fits **all** the data requirements: ✓ continuous + categorical variables, ✓ geographic, ✓ >1500 rows if we pull from UN, World Bank, OECD, etc.

Datasets You Could Use:

- [World Bank Fertility Rate Data](#)
- [OECD Family Database](#)
- [UN Gender Data Portal](#)
- [KFF Global Maternal Health](#)
- [Eurostat Birth Rate + Maternity Leave](#)
- [Our World in Data – Childcare & Work](#)