Falling Behind: A Global Look at Declining Birth Rates and its Consequences

Jonathan Carona

Data Visualization

06.06.2023

Motivation

Just a few months ago, my family joyfully welcomed a new baby boy, making my eldest sister a mother. This is the first time any of my siblings bore a child and it felt for me like starting a new chapter with my family. At the same time, I have came by on my YouTube recommendation feed multiple videos of declining birth rates as if youtube was spying on my whole family. In addition as of late, I had to endure the past life stories of my parents on how many siblings they had.

Because of these recent events, my interest in that topic has greatly increased and I ultimately decided to write a my report on declining birth rates.

Story idea and intention

My intention in writing a data story about global declining birth rates is to inform readers about the ongoing crisis and encourage them to explore the topic further. By presenting survey data and analysis, the story aims to provide readers with a comprehensive understanding of the factors contributing to declining birth rates worldwide. It is designed to engage readers by allowing them to draw their own conclusions about the possible reasons based on data for the decline. To make the story also relevant for the reader, he will be informed on possible future consequenes and what we have already tried in the past to mitigate declining birth rates.

Target audience

General Public and Media. The data story aims to raise awareness among the general public about the global declining birth rate crisis. It provides accessible information that can be easily understood by a wide audience.

Exploratory work

Datasets

[Fertility rate]

This dataset contained all the mean birth rate per year by country.

Country_Name	Country_Code	1960	1961	1962	1963
Aruba	ABW	4.820000	4.655000	4.471000	4.271000
tern and Southern	AFE	6.724125	6.742752	6.762930	6.778712
Afghanistan	AFG	7.282000	7.284000	7.292000	7.302000
estern and Central	AFW	6.458448	6.471518	6.491826	6.506088
Angola	AGO	6.708000	6.790000	6.872000	6.954000

[Survey on Maternity]

	Gender	Age	Employment Status	Race/Ethnicity	Relationship Status	Currently have children	Q1	Q2	Q3	Q4
0	Female	19	Student	African American	Single, but not in a relationship	No	5	2	5.0	5
1	Female	27	Currently unemployed	African American	Single, but not in a relationship	No	4	4	5.0	5
2	Female	27	Employee	Haitian-American	Single, but not in a relationship	No	1	1	4.0	1
3	Female	45	Employee	Haitian-American	Married	Yes	2	2	4.0	4
4	Female	31	Employee	African American	Married	Yes	3	3	5.0	3

A group of Graduate researchers from Rider University in Lawrenceville, NJ were tasked with formulating a survey with 30 response questions to track social attitudes toward not having Children. In each data row the scores of questions and attributes of women is given.

[Population pyramid]

The website which provided this data had data on many countries age pyramids. I have only chosen the swiss data, since displaying multiple age pyramid plots would be just redundant and would be overwhelming for the reader to look at.

	Age	M	F
0	0-4	223985	212509
1	5-9	227226	214300
2	10-14	221517	209884
3	15-19	216231	204080
4	20-24	244226	228306

Implementation choices

Layout and colors

The data story was done with streamlit, since I wanted to make the story accessible from a browser. I was going for a similar look like the website [medium]. I wanted it to look like a typical publishing article. The reason being, if you search on any search engine and come across the title of story, you would expect it to look like a news article.

As for the colors, I have decided to optimize the data story on dark mode. I have no reasons why I set the setting to dark mode other than it makes reading for most people more comfortable. In the future, I can optimize the website also for white mode.

However since it is now on dark mode, I will optimize the plot colors and accents to dark background colors. To make everything consistent, all plots will use only the following colors:

| Color	In HEX	Description
	#00FFFF	Cyan
	#9370db	Medium purple
	#BA55D3	Medium Orchid
	#483D8B	Dark slate blue
	#00008B	Dark blue
	#FAFAFA	White (Text)
	#0E1117	Black (Background)

Charts

Birth Rate by Country (1960-2020)

This choropleth chart is placed in the beginning of the story. The whole point of the plot is to make the reader realize that there was and is still a global declining birth rate throught the years and beyond. I have chosen it to be an interactive year based choropleth since it is very simple and eye catching to look at. It also makes it relevant for the reader, since one's country is also listed on it. If you inspect the colors, the brighter colors have lower birth rates. This is because humans react on brighter colors and I wanted to highlight the change into lower birthrates.
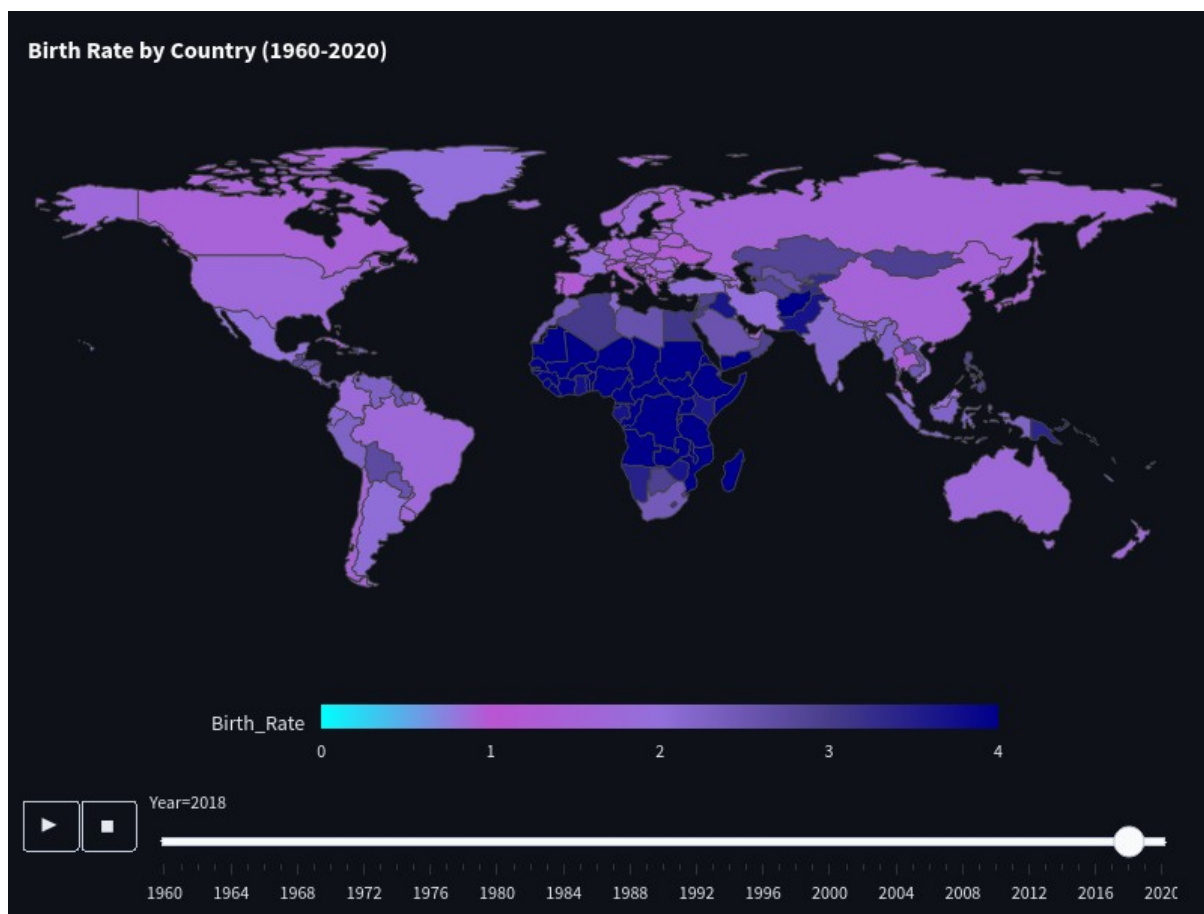
Survey: Women's backgrounds

In the survey, there were 124 women with different backgrounds. I wouldn't want for the reader to think the survey is biased, so I displayed the most important attributes for the story's context. The bright color cyan was used for the category that was prevalent. The positions of the plots: top left, top right and bottom were done with purpose. Humans tend to look at the top left and bottom position instead of the top right position.
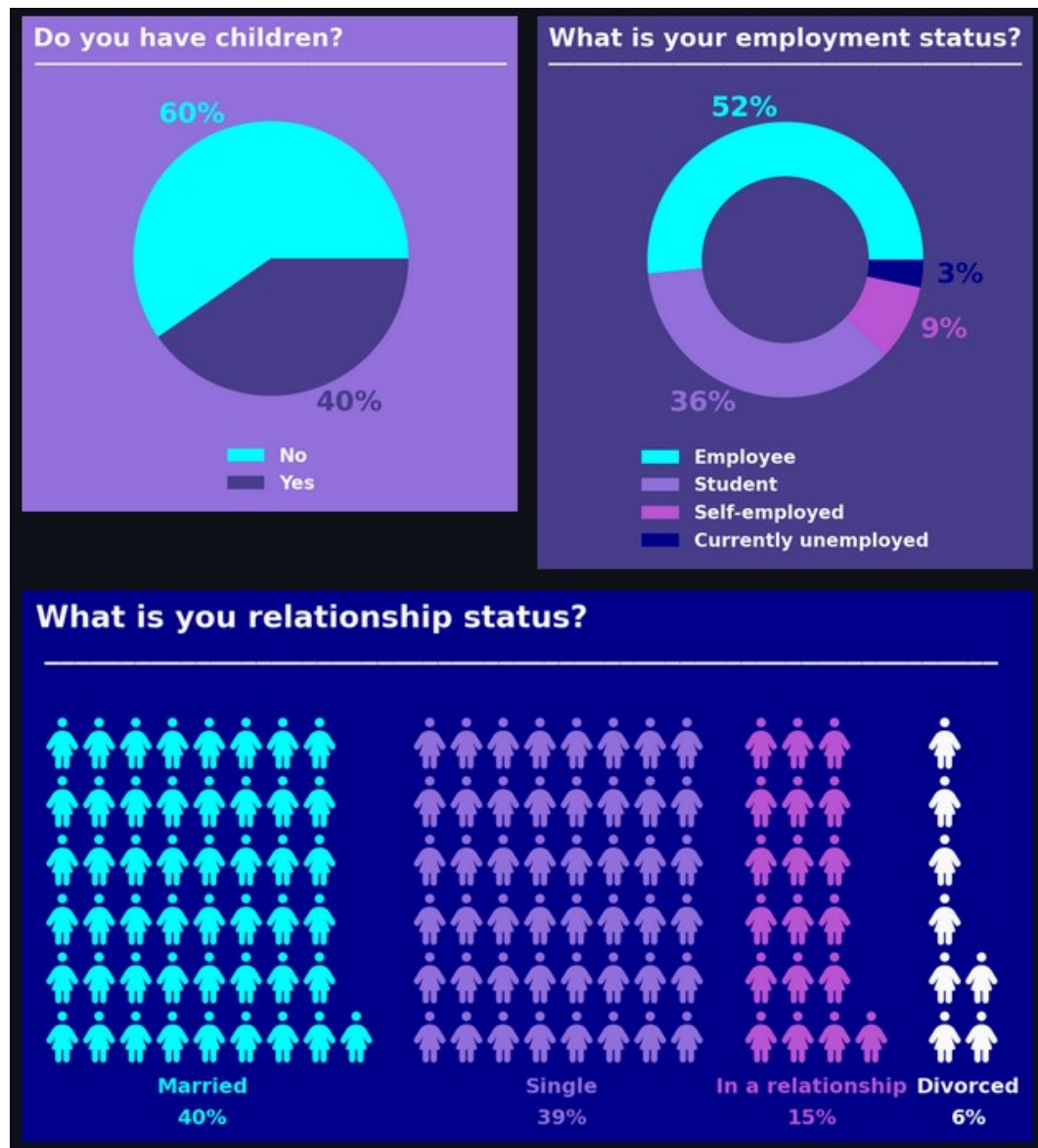
Do you have children?

This pie chart is to show that there is somewhat an equal distribution of women with children and without. A pie chart was chosen to show the data split as a whole. The percentages were rounded to no decimals. This is to make it simple and decimals would not add any value to the plot. The plot was placed top left, since I thought it was the most important.

What is your employment status?

Instead of a pie chart, a donut chart was chosen. This is because there were 4 categories in this feature. Again here, decimals were removed and it was ensured that it would total to 100%. This plot was placed top right. I thought it wasn't as important as the left and bottom.

What is your relationship status?

This pictogram chart was used to make all three plots more eye catching. If I would use another pie or any round plot, it would make it redundant and more boring. Woman icons were used to visualize the plot for obvious reasons. This plot was put at the bottom position, since it was more important than the top right plot, but still less important than top left.

Survey opinion results

I have spent many hours trying to visualize the data of the survey. At first I wanted to visualize all 30 questions and took the average of it. But it brought a problem that it was tedious for the reader to check every single question.
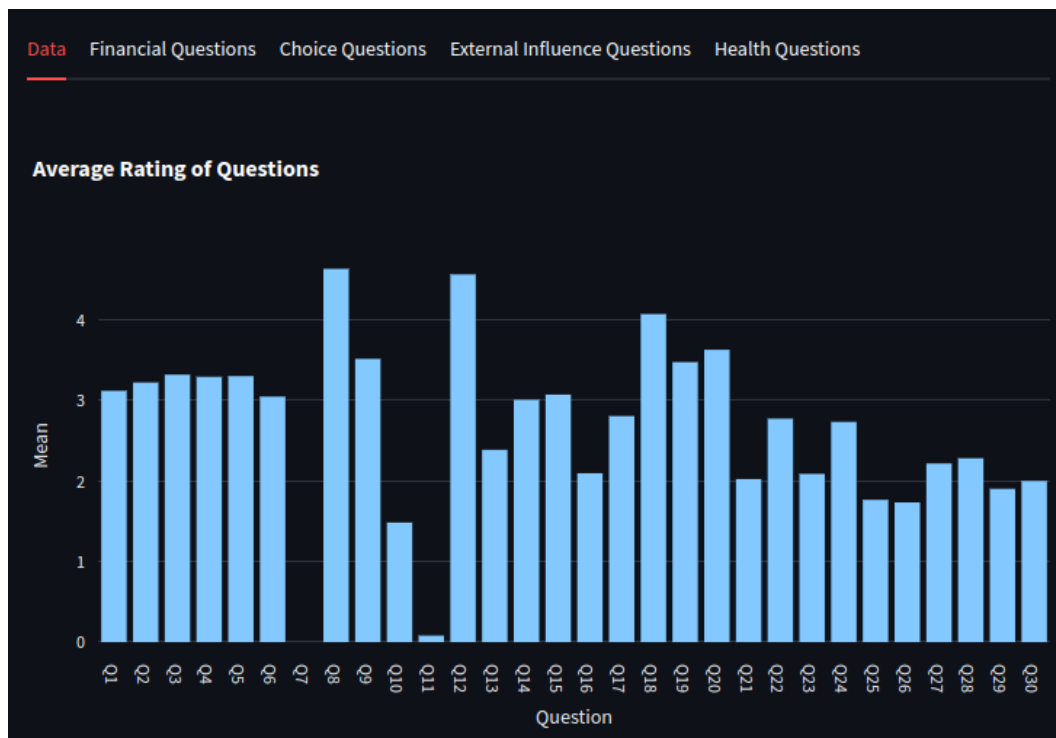
Then I tried to reduce the plots into the four categories and tried different plots. However I was still not satisfied with the following plots:

In then end I have opted into choosing two question of each category. I checked which questions were the most relevant to my story. I have chosen to use a grouped bar plot to distinguish the results between women with children and without children and wrote my analysis according to the differences of the results.
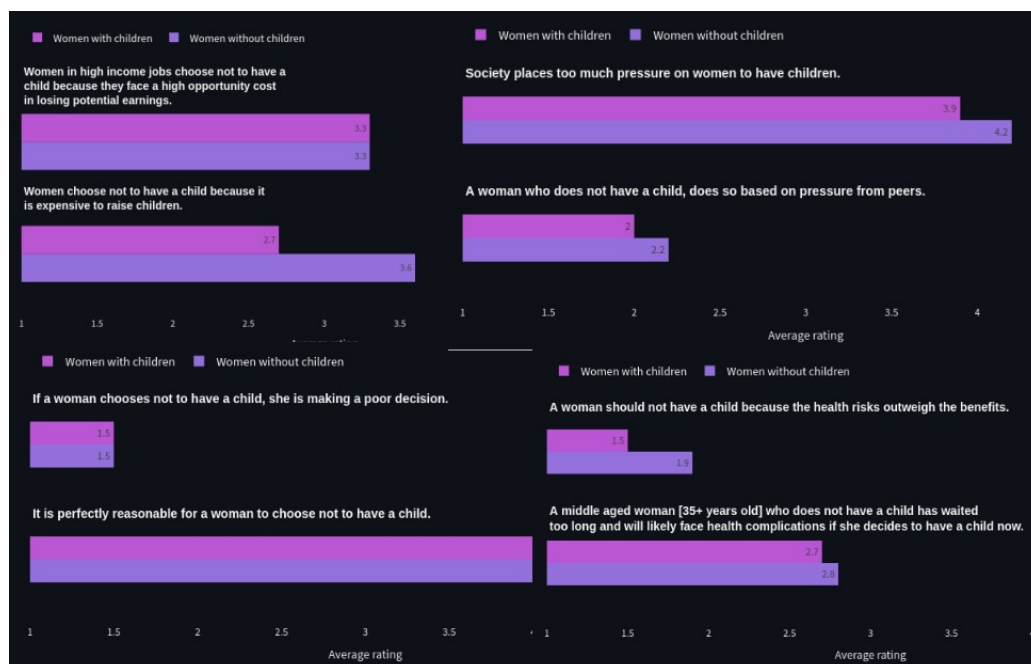
Swiss age pyramid

For the last plot, I have decided to implement an age pyramid chart. I have highlighted the age group that will leave the workforce in 10 years and the other that will enter the workforce. Later in my analysis I have mentioned that the discrepancy between these two age groups will be a problem in the future. In the corners, I have added the icons of a man and a woman to make it more clear which side is which. The data is based on swiss data so I added the swiss flag to the bottom left. To make the plot more self explanatory and more independent from the text analysis, I have included a small annotation two both groups. Indicating these groups are leaving and entering the work force
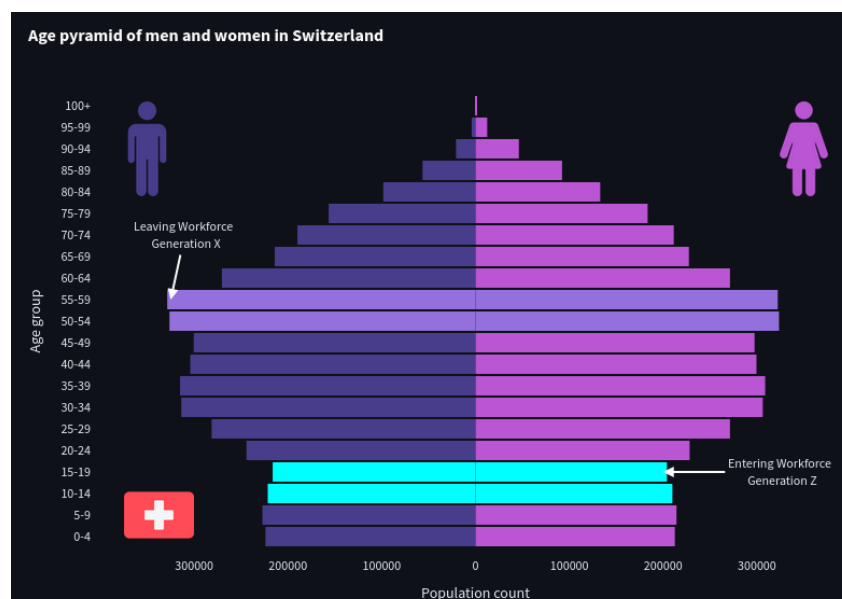
Packages

Pandas

Pandas was used to read and process the datasets. I have some experience with Pandas because of prior data science and ML courses.

Plotly

I found plotly much easier to use and more pleasing to look at than matplotlib.

Interactivity is to an extend already implemented, which made working with the package faster.

Streamlit

I have never used streamlit before, but it left a good first impression from just looking at the documentation. It is also a new oppurtunity to learn a new library.

Streamlit was used to host the datastory on a browser. It was simple to use and most of the default templates already followed my vision on how the data story should look like.

Matplotlib and PyWaffle

I initially planned to only use plotly, however i came across a problem where I wanted to create a pictogram chart.

There was a way to implement that chart type in plotly, but it was too tedious and in the end I opted to use PyWaffle which is based on matplotlib. Pywaffle is a library that made developing pictogram charts simple.

INFO: 8333 characters with whitespaces

Bibliography

Fertility rate: Worldbank, 2022, https://data.worldbank.org/indicator/SP.DYN.TFRT.IN

Survey on Maternity: RKKAGGLE, 2021, https://www.kaggle.com/datasets/rkkaggle2/social-attitudes-regarding-childlessness-nj-survey

Population pyramid: populationpyramid, 2022, https://www.populationpyramid.net/

medium: medium, 2023, https://medium.com/