Catherine Taft

Udacity

Data Analyst Nanodegree – Project 2

Impact of Women in the Workforce

4/30/22

Questions for Analysis

First I'd like to point out that this is the report document only. Because I'm more writing-oriented I chose to keep my report separate from my code, and I kept the explanations for my programming thought process within the Jupyter notebook alongside the code itself. This is just the visualizations I made along with the interpretations.

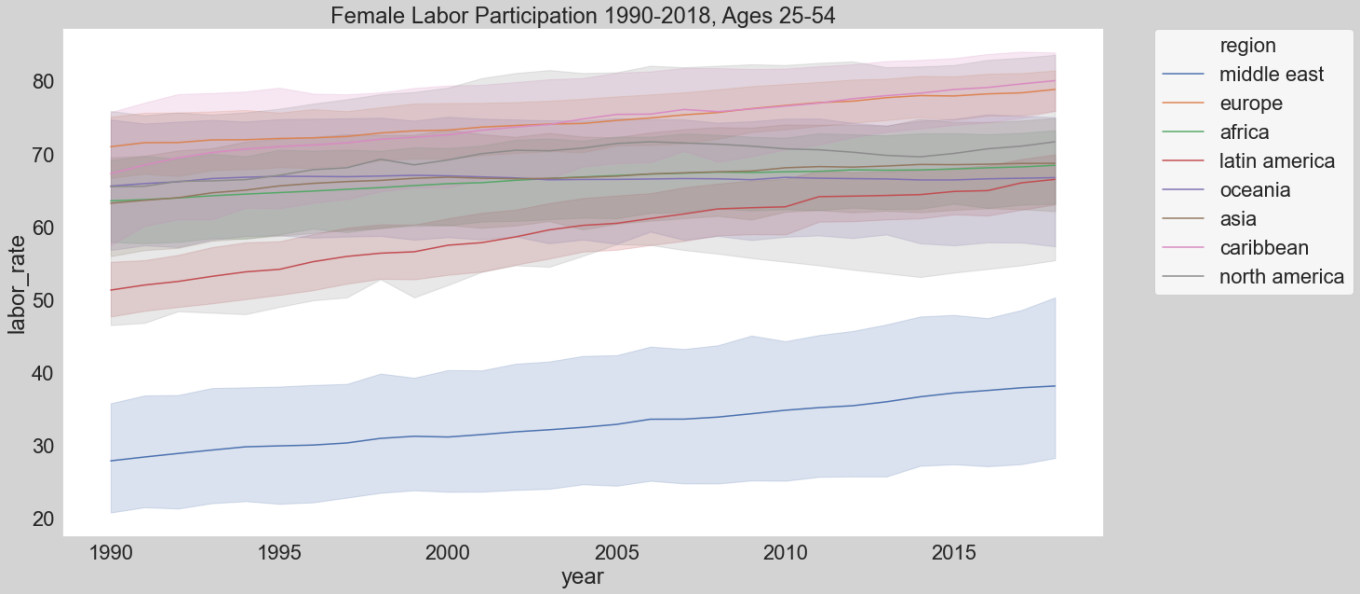
Gapminder has a pretty huge breadth of data but I chose a specific topic that has been on my mind lately for a number of reasons, which is the impact of women joining the workforce in increasing numbers. Obviously there have always been women doing some type of work, but in the last few decades women have joined the workforce alongside men all over the world in pretty much every industry. However, different countries industrialize at different times, some have been set back by war, and some have colonial histories that affect their development, culture, and even functionality in different ways, etc.

So I chose to look at two different questions, principally. The first is simply how the proportion of women in the workforce has changed over time. For this I used Gapminder's dataset that is exactly this – the proportion of women in the workforce. This dataset was out of the box 1989-2018. I shaved it to 1990-2018 in order to correspond to the other datasets. Having added a column to all datasets with geographic region, I wanted to see how the rates compared across different regions over time.

The second question I wanted to look at was what the relationship is, if any, between education rate, birthrates and women in the workforce. This is actually sort of two questions tied up into one. The reason I chose these three variables is because when people discuss women joining the workforce, it's often mentioned in parallel with declining birthrates (although there are surely other contributors), and looking at the relationship between work rates and education rates can potentially give insight into the types of work women are doing, because as women have easier access to education, they are likely to move into the work force in higher number and take better jobs. It can also be a signifier for a country shifting and modernizing, as the concept of traditional female gender roles (i.e. exclusively domestic labor/child rearing) changes. So these three things are linked.

Question 1 – How has the rate of women joining the workforce changed over time, and how has it changed within/between regions?

The range of years for this dataset was 1990-2018. I decided first to just look at the trajectories of each of the regions over this whole time span.

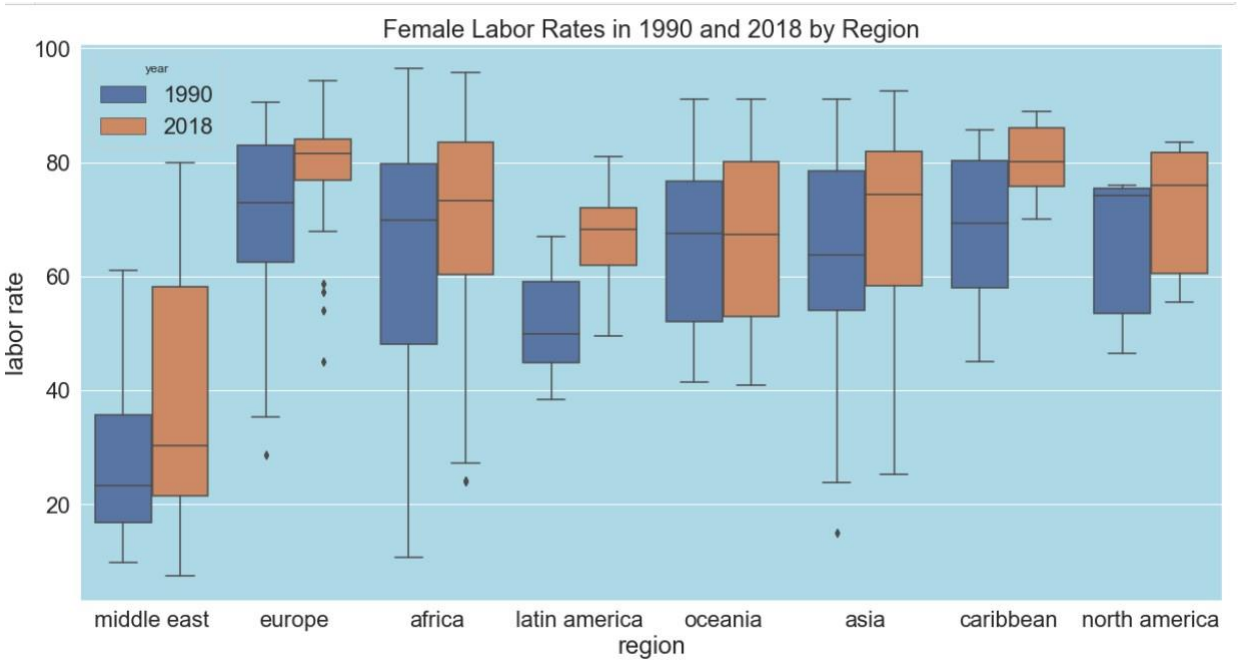


There were two things that struck me about this. One was that the Caribbean (which is part of North America but I separated out due to the fact that it has a more recent colonial history) ended up at the top of the list, just slightly popping out over Europe. The other was exactly how far below everyone else the Middle East is (also noteworthy how wide the confidence interval is for that).

I take the Caribbean result with a grain of salt, because there are simply fewer countries, and Europe has some very large countries in it comparatively, which could skew the data, as some large countries in Eastern Europe still adhere to traditional gender roles in many segments of the population. However, I still thought it was exciting as an isolated data point.

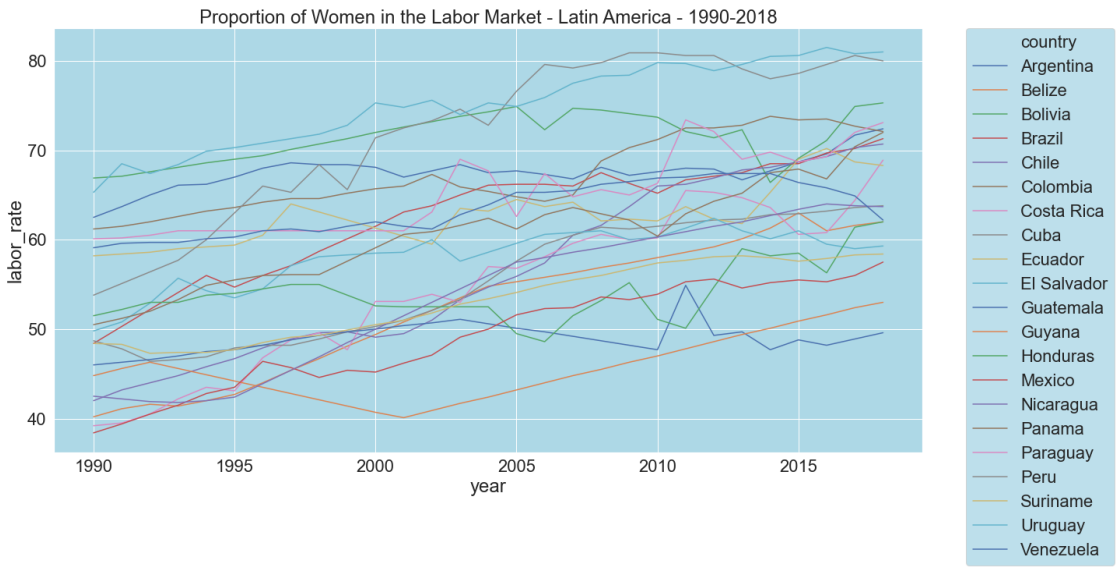
The Middle East is not exactly surprising as much as interesting. Those countries generally adhere to religious and traditional cultural norms. Anecdotally however, knowing some people from that region (mostly older gen Z and millennials), the women in this age range in 2022 are well educated and tend to be employed like anyone else so I was also not surprised to see that the rate of women in the workforce is steadily increasing.

At this point I decided to compare each region's data in 1990 (the beginning of the dataset) to 2018 (the end of the dataset).

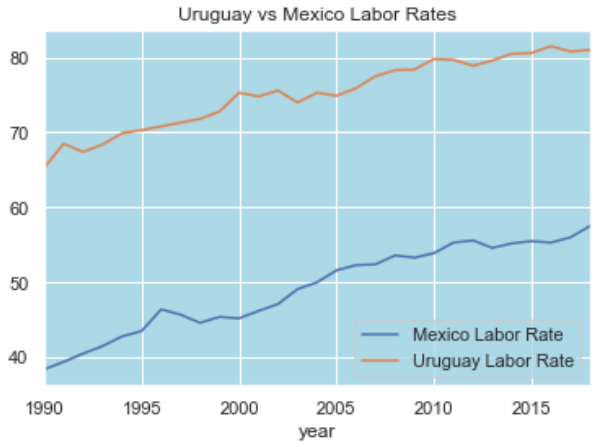


Here we see some pretty remarkable shifts. Latin America completely jumped up, with the mean going from around 50 to 70%. The Middle East's upper quartile has rocketed up more than 20% (which makes sense at least with my anecdotal experience), Africa's mean has stayed about the same but the upper quartile has increased on its upper bound and the lower quartile has compressed into a higher range. The Caribbean as I mentioned before has jumped dramatically. I do find the string of outliers for Europe to be very interesting, and it makes me think of countries like Russia, Bulgaria, some places in the Balkans, where due to any number of factors (poverty levels, access to higher education, traditional or religious cultures) women may not have transitioned into the work force in the same numbers.

At this point I decided to look at a couple of specific regions. I currently live in Panama, which is a modernized middle-income country approaching developed, and have spent a great deal of time in Brazil, so I'm interested in Latin America of course. The result from the box plots that really intrigued me was the Middle East, so I chose that as the other region to look at. I decided to look at the region as a whole and then compare the min and max from each region.

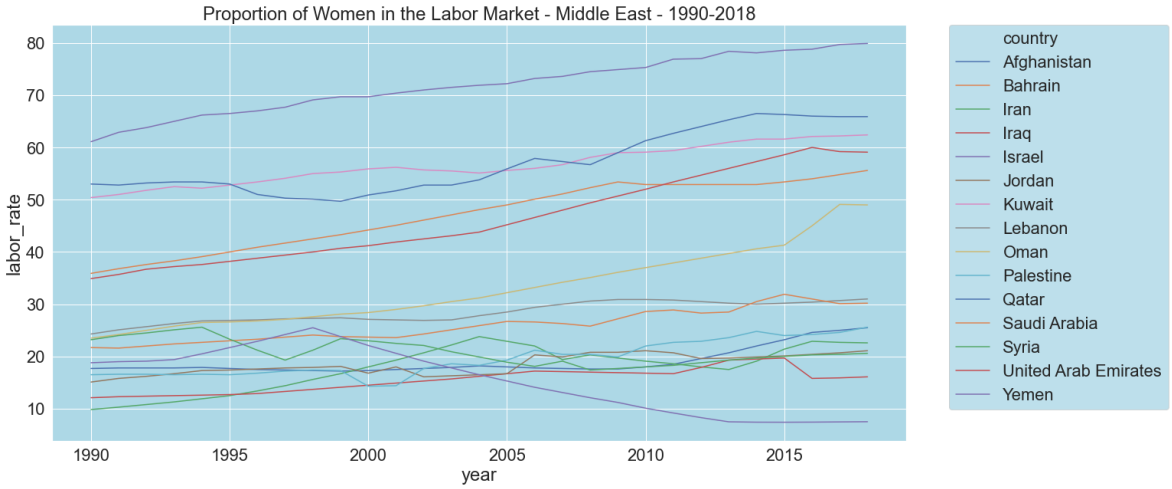


There was nothing overly surprising about this, as Latin America has developed significantly in the last 30 years. It was interesting to see Mexico go from the very bottom to #7, but again not surprising. So I looked at the countries that had the minimum and maximum rates of female employment over this time period:



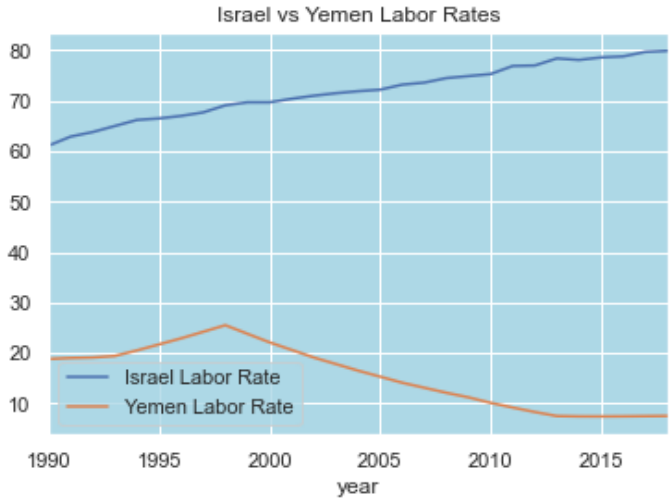
So despite Mexico's stark improvement, it is still far below Uruguay. Uruguay has long been a stable democracy with a high quality of life and a strong economy (Mourot, 2020), so its starting point at a fairly high level as well as its continued improvement makes sense to me.

Next I looked at the Middle East:



Here we can see Israel at the top of the pack, and Yemen way at the bottom. I was a bit surprised to see that there is a clump at the bottom with just a couple of countries above that. Afghanistan and the UAE are among those, which are noteworthy because of UAE's highly developed status, and Afghanistan's improvement in gender equality during its time as a democracy. Israel is unsurprising in its position given that it's a fairly diverse, developed country known to be a leader in industry.

Here's a look at the top and bottom two:



The timeline of Yemen here is pretty depressing. They had a civil war in 1994 (Montgomery, 2021), and after that we see a sharp increase in women entering the workforce, presumably because of

civil stability, but then it peaks in the late 90s and begins to decline again. They've been in another series of conflicts since 2014 post-Arab-Spring which is where the rate really hits bottom and stays there.

Question 2 – Relationship between childbirth rates, education, and women in the labor force

For my second question I wanted to look at the relationship between women's presence in the workforce, fertility rates, and education. Generally, it can be expected that as women have access to education, they will enter the workforce in greater numbers and birthrates may decrease as they decide to focus on their careers. However, it has to be mentioned that there are a lot of other contributing factors to the decision not to have children, including social inequality, financial and economic stress, (i.e. rising cost of living against flattening wages), climate change, etc.

The data sets for these are the one from the previous question that describes the proportion of women between ages 25-54 in the workforce except this time edited to end at 2008 to coincide with the others, one with the average number of children per woman, and one with the average years of education (primary to tertiary) for women ages 15-44.

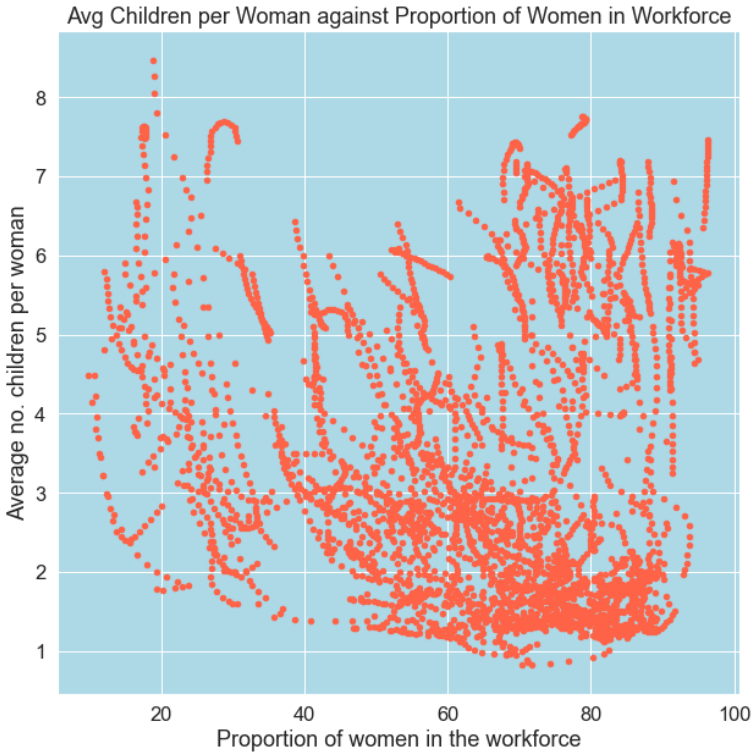
I started with some scatterplots to examine the relationships between these variables.

This was probably the most striking one, describing the relationship between birthrates and the average years of schooling that women receive. There is a stark negative correlation here, where we can see that once people achieve a high school diploma and beyond, the number of children is on average between 1-2. People having 6-8 children tend to have 6 or fewer years of education.

Next I looked at years of schooling against the proportion of women in the workforce:

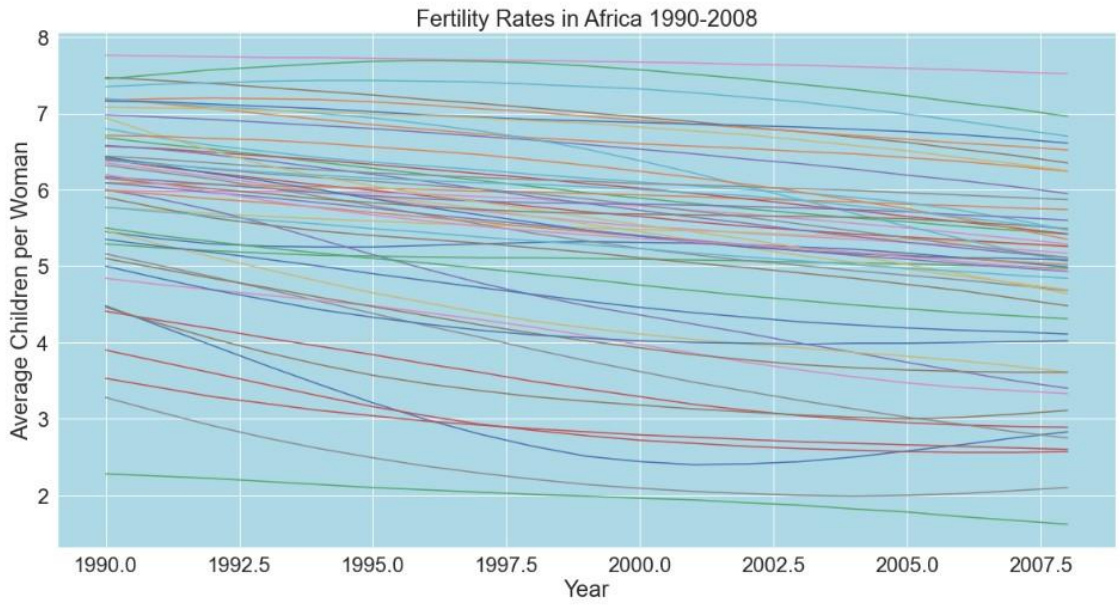
There is not as strong a relationship between these variables, which makes sense; not all work requires an education. Based on the extremely high proportions of women in the workforce in this data I'm assuming that this is not strictly salaried employees and includes manual or other unskilled labor. However, there is a noticeable clumping on the upper right of the plot, where societies with a high proportion women in the workforce tend to have a higher education rate.

The last pairing I looked at was fertility versus the proportion of women in the work force:

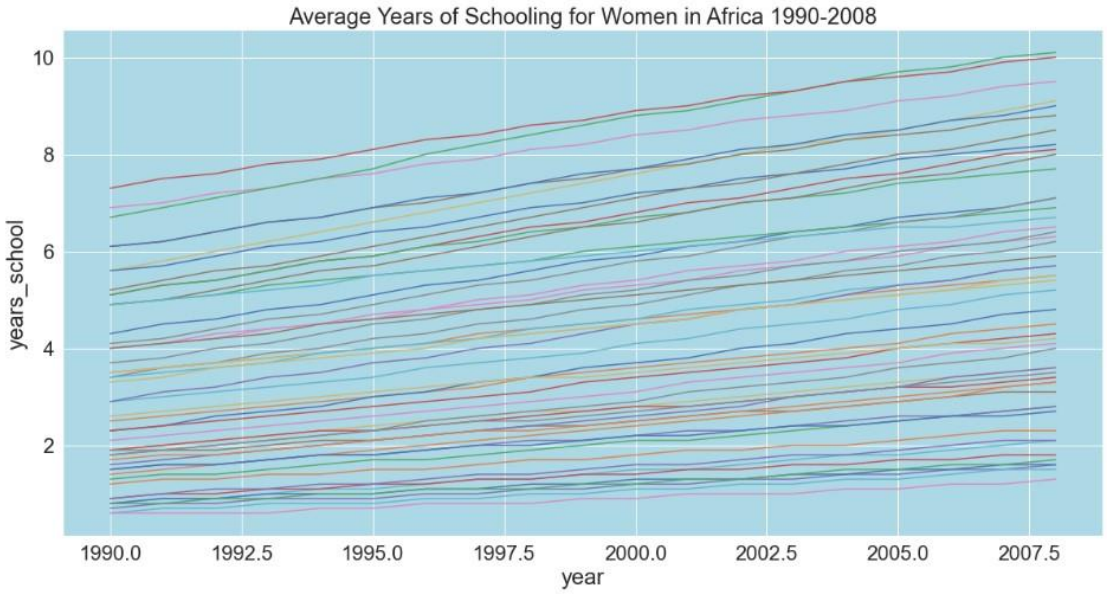


Again there isn't a strong linear relationship between these variables, but we see a noticeable clumping towards the bottom of the plot, where societies with a high percentage of employed women tend to see fewer children per women.

Next I decided again to look at some regional data examining these factors. Looking back at the box plot of changes in workforce participation, another one that jumps out as interesting is Africa, so I decided to examine that. First I looked at the fertility rates and then ran some comparisons on the means.



I didn't include the legend here because it's incredibly long and not immediately relevant. But we can see a general, if slight, decrease here. I ran the means for the dataframe and the mean fertility rate in 1990 was 5.95 children per woman, and in 2008 it was 4.81, which is a decrease of only 1.14. Next I looked at the average years of schooling and ran the means for those as well.



Here we can see a slightly more marked increase with years of schooling. The average years in school for girls in 1990 was 3.11, and in 2008 it was 5.16 which is an increase of 2.05. That's in a positive

direction, and given how many countries are in Africa making up that mean, that could indicate some really positive shifts in certain countries.

Conclusions

With both questions we can see the impact of development and modernizing culture shifts across the world, as women enter the workforce in increasing rates. Alongside this, we see education rates increasing and birthrates declining. These obviously don't have a causal relationship to one another, but they are part of the same shift that's allowing women to work in greater numbers. Since regions and countries develop at different rates and at different times for a variety of reasons, some shifts are more stark than others within the given time periods. I would be very interested to look at data from Europe and North America during the course of the Industrial Revolution, for example. But in general, the data here is in a positive direction and pretty optimistic.

Limitations

The most frustrating limitation with this data is that Gapminder didn't specify exactly what exactly constitutes "workforce." There was another dataset in their menu for salaried workers specifically, but when I tried to go to its page, the page was blank, so I was not able to download it. Because they specifically have a dataset for salaried workers and because the proportions are so high here, I'm assuming that it includes unskilled labor as well.

It was also somewhat unhelpful that it didn't include any grouping of the data by region or even continent, which is a somewhat obvious reference point for this type of data, so I had to add that manually.

Resources

Mourot, J. 2020. *Uruguay – Serving in Latin America's Beacon of Democracy.* State Magazine.

https://statemag.state.gov/2020/07/0720pom/

Montgomery, Marcus. 2021. *A Timeline of The Yemen Crisis From the 90s to the Present*.

Arab Center Washington DC.

https://arabcenterdc.org/resource/a-timeline-of-the-yemen-crisis-from-the-1990s-to-the-present/