

# How does childbirth affect women's life?

<b>1. Introduction</b>	<b>1</b>
Problem Description	1
Question	1
Motivation	1
<b>2. Data Wrangling</b>	<b>1</b>
Data sources	1
Tools	2
Steps for data wrangling	2
<b>3. Data Checking</b>	<b>3</b>
Overview	3
Tools	3
Steps for data checking	3
<b>4. Data Exploration</b>	<b>4</b>
Question 1: How is the childbirth situation across the world?	4
Question 2: What are potential factors affecting childbirth in China?	6
<b>5. Conclusion</b>	<b>8</b>
<b>6. Reflection</b>	<b>9</b>
<b>7. Bibliography</b>	<b>9</b>
<b>8. Appendix</b>	<b>11</b>
Appendix A	11
Appendix B	12
Appendix C	12
Appendix D	13
Appendix E	13
Appendix F	14
Appendix G	14
Appendix H	15
Appendix I	15
Appendix J	16
Appendix K	16

# 1. Introduction

## Problem Description

With the concerns about ageing population and economic growth bottleneck, China has shifted from one-child policy<sup>[1]</sup> since 1980, to two-child policy in 2016, and again three-child policy in 2021. As with these policy changes, when women are bearing the primary burden of childbirth, it is important to understand how childbirth affects a woman's life on an individual level rather than merely consider this question from a national perspective, so that women can make an informed decision.

## Question

In this data exploration project, I am interested to investigate the following questions:

1. How is the childbirth situation across the world?
2. What are potential factors affecting childbirth in China?

## Motivation

I want to take this opportunity to understand the childbirth situation across the world, and specifically how childbirth affects a woman's life in China, by exploring the population structure, gender equality situation, and career development. By performing this data exploration, I wish to produce insights for women considering childbirth, and potentially the policy maker who intends to encourage childbirth.

# 2. Data Wrangling

## Data sources

Three datasets are included in the proposal for this data exploration project, considering the complexity of the dataset and change of scope of topic, only two datasets are used.

1. World Development Indicators

<https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators>

This dataset has 395276 rows and 68 columns, it is time series data from 1960 to 2022, covers 217 countries and regions, as well as summary statistics across geographic location. There are 1486 features in the indicator list.

Following steps are taken to process this dataset:

Given the large magnitude of this dataset, the primary focus of data wrangling in this dataset is to find relevant features regarding the data exploration topic.

2. Social Expenditure

<https://data.oecd.org/socialexp/family-benefits-public-spending.htm>

This dataset is removed due to a change of topic scope. Given the large magnitude of the world development indicator dataset, we had to limit our exploration with focus in China. This social expenditure dataset does not contain information about China, therefore is no longer relevant.

3. Days of leaves (including paid maternity, paternity, and parental leave)  
<https://ourworldindata.org/grapher/days-of-paid-maternity-paternity-and-parental-leave?tab=table&time=2022>

## Tools

1. Google colab: to perform data processing in jupyter notebook for readability
2. Google sheet: to preview dataset, and convert csv to xlsx file for tableau visualisation
3. Tableau: to adjust data type
4. Pandas package: to perform dataframe manipulation
5. Openpyxl package: to read xlsx file into jupyter notebook for further processing
6. Google search: to confirm feature relevance

## Steps for data wrangling

Given the difference across datasets, the data wrangling steps taken for each dataset are different as well. The jupyter notebook containing steps for data wrangling of these datasets can be accessed via the following link:

<https://colab.research.google.com/drive/1Jaa0-lBp6ks3hUXx-HkEplnBnrMDMNih?usp=sharing> (please refer to “dataset and wrangling process” section)

### 1. World Development Indicators

- Inspect the shape of dataset
- Extract number of features and number of countries
- Remove summary statistics for region
  - We only perform the exploration and analysis on country level, instead of region or continent level, but keep the world data as a benchmark for comparison.
  - We remove countries first to reduce computational expenses.
- Manually select relevant feature & Perform preliminary visualisation
  - This is a recursive process.
  - We first extracted all features to a csv file, and used google sheet to group relevant features.
  - We extract the feature names to notebook, filter the dataframe based on selected features, and extract them to csv file.
  - Convert the csv file to xlsx file, open with tableau to perform preliminary visualisation.
- Confirm relevant features
- Reshape the dataframe
  - The outcome of this dataset are two dataframes, df\_world\_birth and df\_china.
  - The df\_world\_birth has the country name, indicator name, year and value as column name, where each row is one record.

- The df\_china have relevant indicator names as columns, and year as row.  
This dataframe is further filtered to keep the year from 1980 to 2022, which is the relevant period for China's childbirth policy change.
- The overview of the wrangled df\_world\_birth and df\_china are listed as appendix A.

## 2. Days of leaves

This is a smaller dataset, where less wrangling steps are needed

- Inspect shape and content of the dataset
- Drop unnecessary column
- Reshape the dataframe
  - The outcome of this dataset is df\_leave, which has entity, year, leave type and days as columns, where each row is one record.
- The overview of the wrangled df\_leave is listed as appendix B.

# 3. Data Checking

## Overview

There are three dataframes used in the exploration project:

1. df\_world\_birth
2. df\_china
3. df\_leave

The steps we took for data checking are similar, and given the credible source of those dataset, the dataframes are expected to be clean. Therefore, the primary purpose of this data checking process is to validate that data is ready to use.

The jupyter notebook containing steps for data checking of these datasets can be accessed via the following link:

<https://colab.research.google.com/drive/1Jaa0-lBp6ks3hUXx-HkEplnBnrMDMNih?usp=sharing> (please refer to “data checking” section)

## Tools

1. Google colab: to perform data checking in jupyter notebook for readability
2. Google sheet: to preview dataset, and perform range check with built in functionality
3. Pandas package: to perform dataframe checking
4. Tableau: to visualise data for understanding of missing pattern

## Steps for data checking

Given the similarity of data checking process, the data checking steps of three dataframes are summarised as follows:

1. Validate data types

The validation of data types are performed using the pandas package. The validated data types are shown in appendix C.

## 2. Range Check and looking for outliers

The range checks are performed in google sheet by manually confirming with built in summary statistics, and is performed on perview.csv and leave.csv. One example screenshot of the range checking is shown as appendix D.

## 3. Missing value check

The missing value check is also performed using the pandas package, the missing value checking results are listed as appendix E.

Both df\_world\_birth and df\_china have large amounts of missing value. However, after inspecting the raw data, I realised that those missing values are not missing at random, they are missing that depend on unobserved predictors, such as policy influence, social awareness ect. This would be further analysed in the data exploration section of this report. Therefore, to keep the relevant information, I made the decision to keep those missing values and study their pattern in the exploration project.

As for df\_leave, there are no missing values in this dataset, and it is ready for use.

## 4. Pattern checking for inaccurate data removal

Some outlier and inaccurate data are spotted through data visualisation. For example, the sex ratio of Liechtenstein is showing a sine wave pattern and is not consistent with other online data sources, therefore removed, see appendix F.

# 4. Data Exploration

The tools used for data exploration are tableau. Google search is also used to provide context for visualisation and analysis. The twb file and dataset used can be find via this link: <https://drive.google.com/drive/folders/18x7GDABW-X3ot78c1YL-jj4fLQhELr-8?usp=sharing>

## Question 1: How is the childbirth situation across the world?

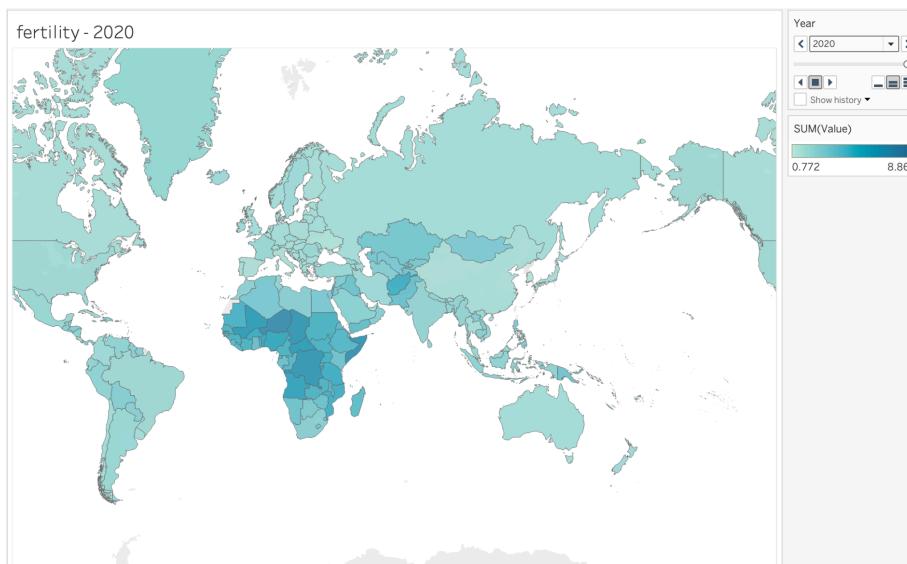
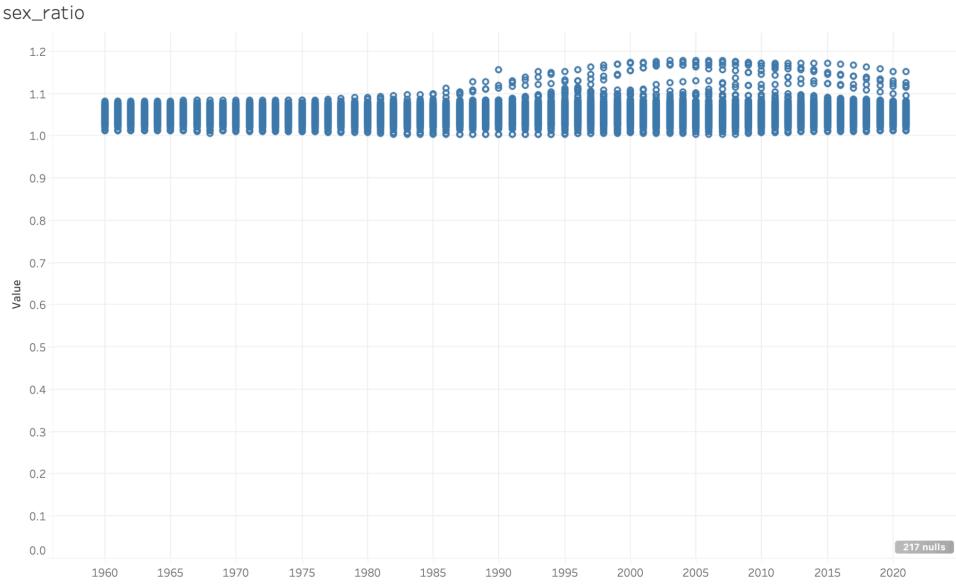


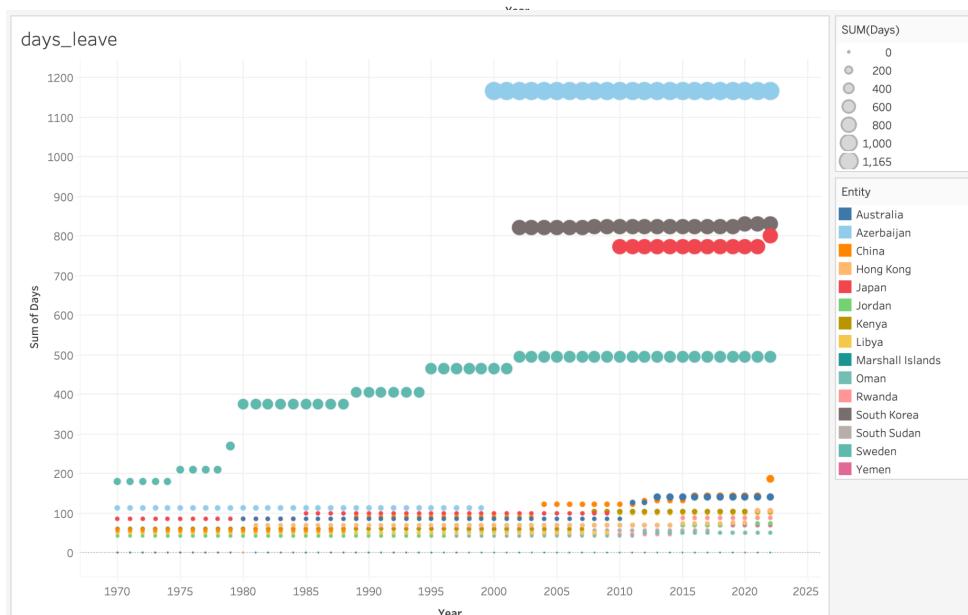
Figure 1: This is a visualisation of fertility rate (birth per women) across the world, with a year slider, containing data from 1960 to 2022. The minimum value 0.772 belongs to Hong Kong in 2021, see Appendix G.



**Figure 2:** This is a visualisation of maternal death, maternal death (1 in the number of women giving birth)



**Figure 3:** This is visualisation for sex ratio across countries (male births per female births), more information is available in this visualisation with value filter and year slider.



**Figure 4:** This is a visualisation about total days of leave related to childbirth, with a country filter which allows us to select a subset of countries of interest.

## Key Findings:

- Africa and central Asia have higher fertility ratios than the rest of the world.  
Women in Jordan, Kenya, Libya, Marshall Islands, Oman, Rwanda, South Sudan and Yemen Republic have had fertility rates greater than 8.  
Only until 1993, the world's fertility rate finally decreased under 8.  
See appendix J for more details.
- According to figure 2, maternal death did not make it to the world developer indicator until 2000.
  - In 2000, most countries with available data had the lifetime risk of maternal death less than 1 in 10000 women.
  - In 2020, the lifetime risk of maternal death is 1 in 15 women, there are numerous countries who have a lifetime risk of maternal death less than 1 in 100 women.
  - See appendix H for more details.
  - In 2019 to 2020, some countries show a sudden increase of maternal death risk, this could be relevant with covid-19<sup>[3]</sup>.
- Based on Figure 3, the majority of countries have sex ratio from 1 to 1.1. If we filter the country with higher sex ratio, we can get a subset of countries with “son preference”<sup>[2]</sup>.
  - When setting the year slider to 2000, the countries with son preference are: Armenia, China, Azerbaijan, Northern Mariana Islands, Georgia, Albania, and South Korea, see appendix I.
- According to Figure 4, the number of days for childbirth related leaves have some correlation with fertility rate.
  - The drastic increase in days of leave in Azerbaijan have slowed down the speed of decrease of its fertility rate, but the decrease is inevitable.
  - The generous childbirth related leaves in 1985 and 2010 Japan have slowed down or even slightly increased the fertility rate, but the recent adjustment in 2022 has not worked so far.
  - See appendix K for more information.
  - The longer childbirth related leaves may not necessarily improve fertility rate, it is affected by gender neutrality of the leaves, and this could affect both mother's and father's career<sup>[4]</sup>.

## Question 2: What are potential factors affecting childbirth in China?

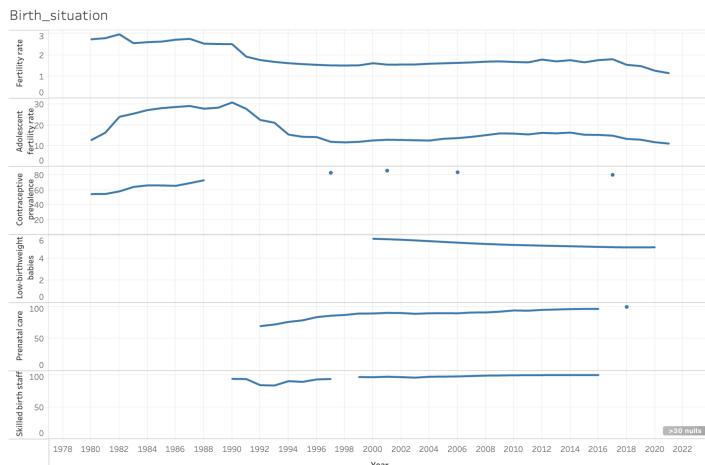
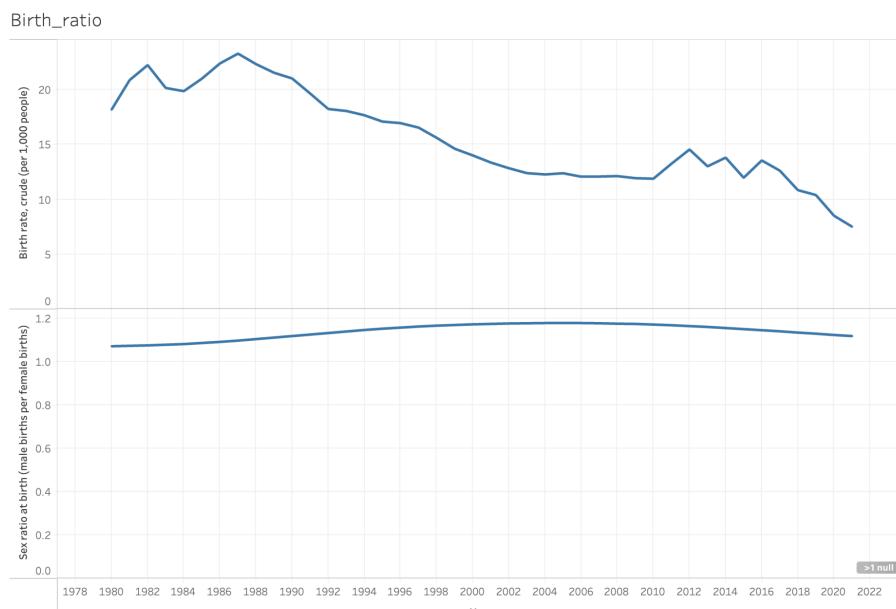
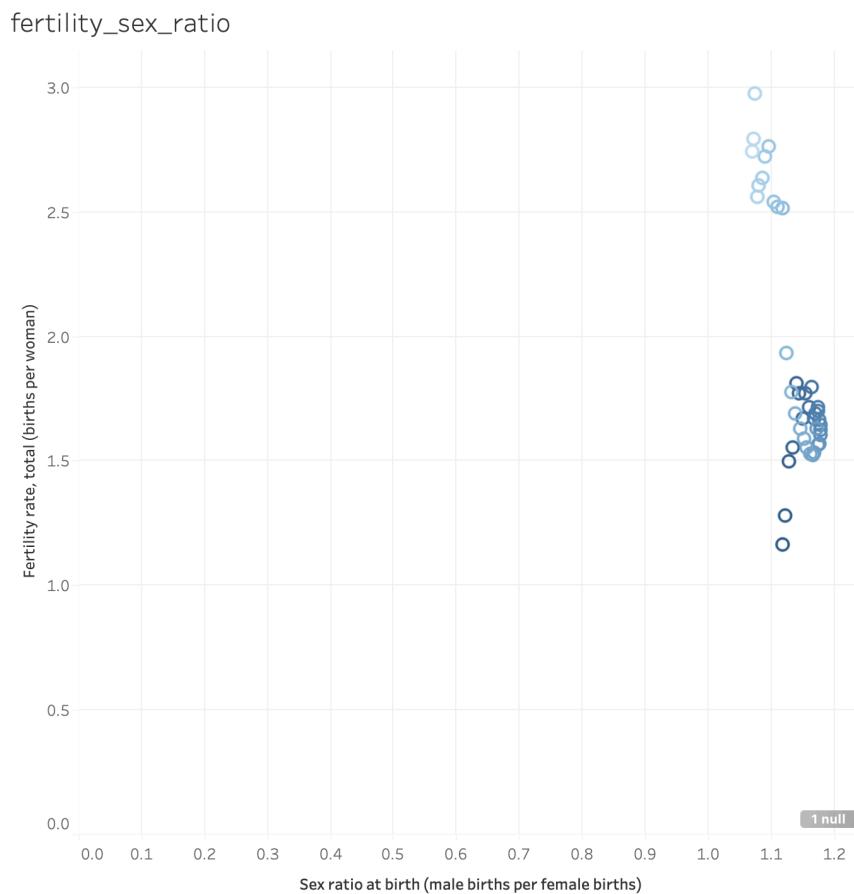


Figure 5: This is a visualisation about the fertility rate (births per woman), adolescent fertility rate (births per 1,000 women ages 15-19), contraceptive prevalence (% of married women ages 15-49), low-birth weight babies (% of births), parental care (pregnant women receiving prenatal care %) and skilled

birth staff attendance (births attended by skilled health staff % of total).



**Figure 6:** This is a visualisation of birth date (per 1000 people) and sex ratio (male births per female births).



**Figure 7:** This is a visualisation of the trajectory between the correlation of fertility rate and sex ratio at birth, where light blue indicates earlier years and dark blue later, covers 1980 to 2022.

#### Key findings:

- According to Figure 5, it is noticeable that visualisation about the birth situation in China has a lot of missing data, and as mentioned earlier, those missing values were not removed because they are missing due to unobserved features.

- The lack of data for contraceptive prevalence since 1988 could indicate the relaxation of one-child policy and the fact that the awareness of one-child has been widely accepted, which is reflected in the reduced fertility rate.
  - The percentage of low-birth weight babies are negatively correlated with parental care and skilled birth staff during childbirth.
  - It is also worth noting that the percentage of parental care and skilled birth staff during childbirth have only reached the highest 99.6 and 96.6 respectively. The number of females in China is 691 million in 2022, which means 2.8 million and 30 million did not receive those support during their childbirth.
- According to Figure 6, from 1985 to 2010, the birth ratio is showing a decreasing trend, while the sex ratio is increasing correspondingly. The trend of sex ratio reversed after 2010, while the birth ratio increased briefly, and fluctuated decreased, especially after covid.
- According to Figure 7, the fertility ratio was closer to world average until 1990, was going back and forth from average level to unusually high level especially after 2010, and is showing a reduced trend after 2015.

## 5. Conclusion

### Question 1

- More children are born in Africa and central Asia, as well as conventionally developing countries.
- In extreme cases, the average woman gives birth to 8 children in their lifetime, even when the risk of maternal death is high.
- The common sex ratio across the world is 1 to 1.1, with certain countries showing increased sex ratio as the development of prenatal sex discernment technology.
- Many countries attempt to increase childbirth related leaves to increase fertility, but those policies don't always work.

### Question 2

- Childbirth policies are the most important factor to affect childbirth in China, relevant time patterns are consistent with policy publish time.
- The support for childbirth, such as skilled birth staff and parental care for pregnant women has a slight positive effect on fertility rate.
- The relaxation of child numbers has improved the “son preference” in China, but the historical effect of it remains<sup>[5]</sup>.

### Question 3

(How does childbirth affect women's education experience and career development?)

I was unable to answer this question due to the limited information from the dataset, the difficult to extract time trend from incomplete dataset, as well as the time constraint.

Therefore, I replaced this question with question 1 which was not originally in the proposal.

## 6. Reflection

As is elegantly put in Invisible Woman by Caroline Criado-Perez, there is such a thing as being sexist without the realisation, and this has been reflected on lack of data regarding women. I wish I have had the dataset to explore women's suicide due to denied caesarean section<sup>[6]</sup>, the "husband stich"<sup>[7]</sup> without consent after natural labour, side effects<sup>[8]</sup> like urinary incontinence or depression after childbirth. But such dataset are usually qualitative or smaller in size, and are not sufficient to be visualised. In hindsight, I wish I had considered more methods to try and visualise text data, such as using web scraping to extract youtube comments related to such discussions.

Additionally, I could have chosen a less personal topic for the data exploration project. I was the one not aborted during the one-child policy, and a feminist, and that background may have affected my objectivity and added a sense of responsibility when completing this project. Had I chosen a less personal topic, I could probably put more focus and the exploration and visualisation part of the report.

However, I am still glad that I attempted this topic in my data exploration project, and I hope you enjoy reading this report, and it can provide you with some insights regarding how childbirth affects a woman's life, and how the childbirth situation is in countries around the world.

## 7. Bibliography

- [1] Kane, P., & Choi, C. Y. (1999). China's one child family policy. *BMJ (Clinical research ed.)*, 319(7215), 992–994. <https://doi.org/10.1136/bmj.319.7215.992>
- [2] Wikipedia contributors. (2024, March 8). Son preference. In Wikipedia, The Free Encyclopedia. Retrieved May 2, 2024, from [https://en.wikipedia.org/w/index.php?title=Son\\_preference&oldid=1212647036](https://en.wikipedia.org/w/index.php?title=Son_preference&oldid=1212647036)
- [3] Caryn Rabin, R. (2023, March 16). Covid Worsened a Health Crisis Among Pregnant Women. *The New York Times*. (This article may require subscription, but the link should give you full access to the article, please let me know if it doesn't work) [https://www.nytimes.com/2023/03/16/health/covid-pregnancy-death.html?unlocked\\_article\\_code=1.000.fZZa.QvxnoGMRjonj&smid=url-share](https://www.nytimes.com/2023/03/16/health/covid-pregnancy-death.html?unlocked_article_code=1.000.fZZa.QvxnoGMRjonj&smid=url-share)
- [4] Hideg, I., Krstic, A., Trau, R., & Zarina, T. (2018, September 14). Do Longer Maternity Leaves Hurt Women's Careers? *Harvard Business Review*. <https://hbr.org/2018/09/do-longer-maternity-leaves-hurt-womens-careers>
- [5] [Female Political Strategy]. (2022, April 18). 35 Million Surplus Chinese & Indian Men Struggle To Find Wives & A Purpose [Podcast]. Spotify. <https://open.spotify.com/episode/778Nf2StGeYxGsDV0WQ2XO?si=aKZ1cEO1Qf-LL2GERgvcg>

[6] Allen, K. (2017, September 7). Chinese debate medical rights after pregnant woman's suicide. BBC. <https://www.bbc.com/news/world-asia-china-41172675>

[7] Reilly, K. (2023, February 22). Is 'the husband stitch' a medical myth? Women speak out about their experience. Yahoo Life.

[https://www.yahoo.com/lifestyle/what-is-the-husband-stitch-episiotomies-130009654.html?guccounter=1&guce\\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xIImNvbS8&guce\\_referrer\\_sig=AQAAALS0cSBDqERaumITk4dqgGfqlxI6\\_gxbI4LyuYTsGODL4BEtrxR9r5V54WA4qY-XgmzRpUNby8DAX1kedPCiPqp7FArBZg2zL\\_Ex5amUUYEnC9giTyc3VWdoX7EVHg6XeqsDN3Wm94nmUfBri4FzLJMeZbyrHDdfHKgaopzNxoR](https://www.yahoo.com/lifestyle/what-is-the-husband-stitch-episiotomies-130009654.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xIImNvbS8&guce_referrer_sig=AQAAALS0cSBDqERaumITk4dqgGfqlxI6_gxbI4LyuYTsGODL4BEtrxR9r5V54WA4qY-XgmzRpUNby8DAX1kedPCiPqp7FArBZg2zL_Ex5amUUYEnC9giTyc3VWdoX7EVHg6XeqsDN3Wm94nmUfBri4FzLJMeZbyrHDdfHKgaopzNxoR)

[8] (2023, December 7). More than a third of women experience lasting health problems after childbirth, new research shows. WHO.

<https://www.who.int/news/item/07-12-2023-more-than-a-third-of-women-experience-lasting-health-problems-after-childbirth>

## 8. Appendix

Appendices are added to the table of content for convenience and quick reference.

### Appendix A

Screenshot of df\_world\_brith after wrangling:

	Country Name	Indicator Name	Year	Value
0	World	Adolescent fertility rate (births per 1,000 wo...	1960	91.748048
1	World	Births attended by skilled health staff (% of ...	1960	NaN
2	World	Contraceptive prevalence, any modern method (%...)	1960	NaN
3	World	Fertility rate, total (births per woman)	1960	4.695854
4	World	Lifetime risk of maternal death (1 in: rate va...	1960	NaN
...	...	...	...	...
97897	Zimbabwe	Fertility rate, total (births per woman)	2022	NaN
97898	Zimbabwe	Lifetime risk of maternal death (1 in: rate va...	2022	NaN
97899	Zimbabwe	Low-birthweight babies (% of births)	2022	NaN
97900	Zimbabwe	Pregnant women receiving prenatal care (%)	2022	NaN
97901	Zimbabwe	Sex ratio at birth (male births per female bir...	2022	NaN

97902 rows × 4 columns

Screenshot of df\_china after wrangling:

Given the large number of features contained in this dataset, the screenshot of features and records are incomplete. The dataframe has 43 rows × 73 columns, ranging from 1980 to 2022, containing the features previously deemed relevant.

Indicator Name	Adjusted net national income per capita (annual % growth)	Adjusted net national income per capita (constant 2015 US\$)	Adolescent fertility rate (births per 1,000 women ages 15-19)	Age dependency ratio (% of working-age population)	Age dependency ratio, old (% of working-age population)	Age dependency ratio, young (% of working-age population)	Birth rate, crude (per 1,000 people)	Births attended by skilled health staff (% of total)	Contraceptive prevalence, any method (% of married women ages 15-49)
1980	NaN	NaN	12.749	68.109636	7.379242	60.730393	18.21	NaN	NaN
1981	NaN	NaN	16.332	65.525607	7.416471	58.109137	20.91	NaN	NaN
1982	NaN	NaN	24.029	63.580339	7.463516	56.116823	22.28	NaN	NaN
1983	NaN	NaN	25.563	61.376407	7.488467	53.887940	20.19	NaN	NaN
1984	NaN	NaN	27.302	58.807946	7.501029	51.306917	19.90	NaN	NaN
1985	NaN	NaN	28.228	56.578883	7.540722	49.038161	21.04	NaN	NaN
1986	NaN	NaN	28.796	54.793170	7.597524	47.195646	22.43	NaN	NaN
1987	NaN	NaN	29.270	53.585668	7.664629	45.921040	23.33	NaN	NaN

## Appendix B

Screenshot of df\_leave after wrangling:

	Entity	Year	Leave Type	Days
0	Afghanistan	1970	Length of paid maternity leave (calendar days)	0
1	Afghanistan	1971	Length of paid maternity leave (calendar days)	0
2	Afghanistan	1972	Length of paid maternity leave (calendar days)	0
3	Afghanistan	1973	Length of paid maternity leave (calendar days)	0
4	Afghanistan	1974	Length of paid maternity leave (calendar days)	0
...	...	...	...	...
50080	Zimbabwe	2018	Length of paid shared parental leave (calendar...	0
50081	Zimbabwe	2019	Length of paid shared parental leave (calendar...	0
50082	Zimbabwe	2020	Length of paid shared parental leave (calendar...	0
50083	Zimbabwe	2021	Length of paid shared parental leave (calendar...	0
50084	Zimbabwe	2022	Length of paid shared parental leave (calendar...	0

50085 rows × 4 columns

## Appendix C

Screenshot of data type validation:

```
df_world_birth.dtypes
```

```
Country Name      object
Indicator Name    object
Year             object
Value            float64
dtype: object
```

```
df_china.dtypes
```

```
Indicator Name
Adjusted net national income per capita (annual % growth)           float64
Adjusted net national income per capita (constant 2015 US$)          float64
Adolescent fertility rate (births per 1,000 women ages 15-19)        float64
Age dependency ratio (% of working-age population)                   float64
Age dependency ratio, old (% of working-age population)                float64
Vulnerable employment, male (% of male employment) (modeled ILO estimate) float64
Vulnerable employment, total (% of total employment) (modeled ILO estimate) float64
Wage and salaried workers, female (% of female employment) (modeled ILO estimate) float64
Wage and salaried workers, male (% of male employment) (modeled ILO estimate) float64
Wage and salaried workers, total (% of total employment) (modeled ILO estimate) float64
Length: 73, dtype: object
```

```
df_leave.dtypes
```

```
Entity      object
Year       int64
Leave Type  object
Days       int64
dtype: object
```

## Appendix D

Screenshot example of performing range check in google sheet:

The screenshot shows a Google Sheets interface with a table of data. The table has columns labeled BC through BO and rows labeled 1 through 27. A formula is applied to the range D2:BN2, which highlights the entire table. The formula is =if(D2>=2011 and D2<=2022, 1, 0). This formula checks if the value in column D (the year) is between 2011 and 2022. If it is, the cell contains a 1; otherwise, it contains a 0. The formula is also shown in the formula bar at the top.

	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM	BN	BO
1	0	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
2	7	1.538782432	1.859800484	1.180956899	2.123053825	1.97093608	1.644139995	2.556436994	1.487428152	1.172793511	-4.961173445	4.557003073	
3	13	7756.114621	7900.362879	7993.662759	8163.372522	8324.267376	8461.129986	8677.433443	8806.504031	8909.786138	8467.756195	8853.632105	
4	4	51.65949147	51.43462742	51.2004455	50.80968892	47.19389179	45.66405213	44.92286463	44.07506449	43.37562576	42.74621887	42.47963455	
5	2	54.71124299	54.6648953	54.70387123	54.78451621	54.87728854	54.97918832	55.11579481	55.23714124	55.32325866	55.35538964	55.29001466	55.18912141
6	4	11.8392989	12.02291556	12.24573336	12.50318239	12.79247047	13.10564261	13.44631202	13.80241946	14.16863802	14.52818901	14.82147643	15.10430117
7	7	41.23041382	41.0349191	40.88762032	40.75360917	40.6034498	40.44310931	40.29946659	40.12250222	39.98316416	39.61323001	39.29130317	38.93966935
8	18	20.11913462	20.23409555	19.72808815	19.64298229	19.11284268	19.17389417	18.70009921	18.17880719	17.81769659	17.22661128	16.93865591	
9	14									82.55781655			
10	15									62.81778325			
11										55.74761455			
12	6	1.749834083	1.770413192	1.777990493	1.759808779	1.779897656	1.789047451	1.809217355	1.821607237	1.824351909	1.797448145	1.822937907	1.839893563
13	7	4.078784644	4.10181235	4.065538543	4.067228378	4.083686682	4.077674447	4.147035054	4.191479354	4.139155825	4.009074338	3.899437648	4.060175205
14	5	3.16059247	3.182984966	3.163837845	3.156823355	3.173849449	3.172509826	3.22104402	3.252155476	3.220237051	3.133462861	3.073801471	3.180267065
15	2	31.84301549	30.99829804	30.14515299	29.23905885	28.54090037	28.01193345	27.48015679	26.84599503	26.36584378	27.04690179	26.88078368	26.36668744
16	4	30.79609711	29.82397097	28.90425535	28.03928738	27.44496273	26.94921934	26.45841647	25.77559017	25.34014158	25.80196376	25.77625212	25.61260147
17	6	32.52439899	31.76213241	30.95252536	30.02082153	29.25621844	28.70723469	28.15024785	27.54884606	27.04102727	27.86271845	27.27775921	26.86175341
18	18	22.9230627	23.3406839	23.31998793	23.36576823	23.23341812	23.2847037	23.39283175	23.52590868	23.54688411	23.4162545	23.73513713	23.89216662
19	17	17.85394511	18.17660668	17.97295033	17.80478205	17.45553947	17.38426348	17.43773635	17.47470042	17.48158105	17.44303379	17.41977646	17.20674256
20	16	26.2222964	26.6997619	26.79909811	26.9892741	27.00468921	27.14523907	27.29866611	27.49927363	27.53973008	27.33106071	27.90360008	26.00000000
21	4	45.23392178	45.66101843	46.53485938	47.39917416	48.22568433	48.7033629	49.12701039	49.62809673	50.08727397	49.53684322	49.5840813	4 ✓ Sum: 61.51768113
22	2	51.34995847	51.99942203	53.122792	54.15593196	55.09950032	55.66651924	56.10384734	56.74970847	57.17827708	56.75500104	56.80397212	5 Avg: 1.537942028
23	9	41.25330405	41.53180553	42.24837652	42.98990502	43.7309467	44.14752613	44.55108651	44.95188001	45.41924417	44.80622151	44.8186413	4 Min: -4.961173445
24													Max: 4.557003073
25													Count: 40
26													Count Numbers: 40
27	9	2.534060066	2.549643099	2.514248426	2.506578481	2.477909572	2.4811556	2.459272238	2.397407425	2.360608718	2.299666176	2.273163352	

## Appendix E

Screenshot of missing value check:

The screenshot shows three Jupyter Notebook cells. The first cell displays the result of df\_world\_birth.isnull().sum(), showing the count of missing values for each column. The second cell displays the result of df\_china.isnull().sum(), showing the count of missing values for each indicator. The third cell displays the result of df\_leave.isnull().sum(), showing the count of missing values for each entity.

```
df_world_birth.isnull().sum()
Country Name      0
Indicator Name    0
Year              0
Value             45610
dtype: int64
```

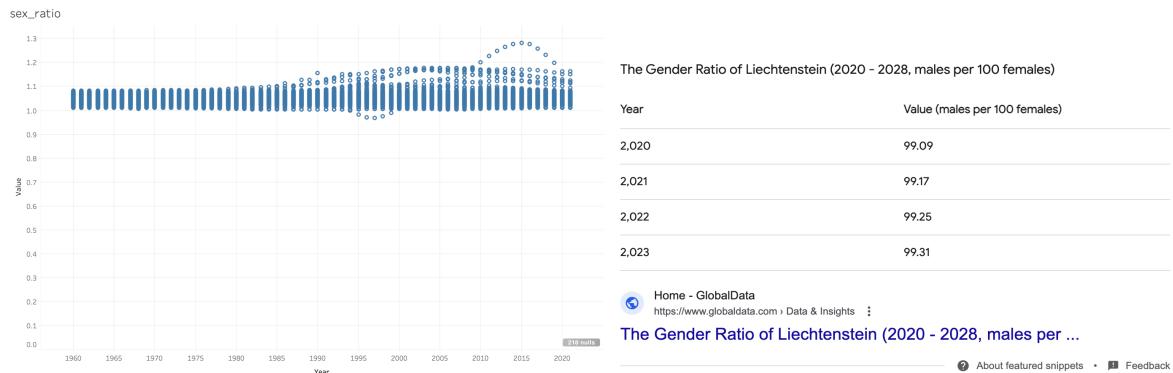
```
df_china.isnull().sum()
Indicator Name
Adjusted net national income per capita (annual % growth)          17
Adjusted net national income per capita (constant 2015 US$)        16
Adolescent fertility rate (births per 1,000 women ages 15-19)       1
Age dependency ratio (% of working-age population)                  0
Age dependency ratio, old (% of working-age population)                0
Vulnerable employment, male (% of male employment) (modeled ILO estimate) 11
Vulnerable employment, total (% of total employment) (modeled ILO estimate) 11
Wage and salaried workers, female (% of female employment) (modeled ILO estimate) 11
Wage and salaried workers, male (% of male employment) (modeled ILO estimate) 11
Wage and salaried workers, total (% of total employment) (modeled ILO estimate) 11
Length: 73, dtype: int64
```

```
df_leave.isnull().sum()
Entity      0
Year       0
Leave Type 0
Days       0
dtype: int64
```

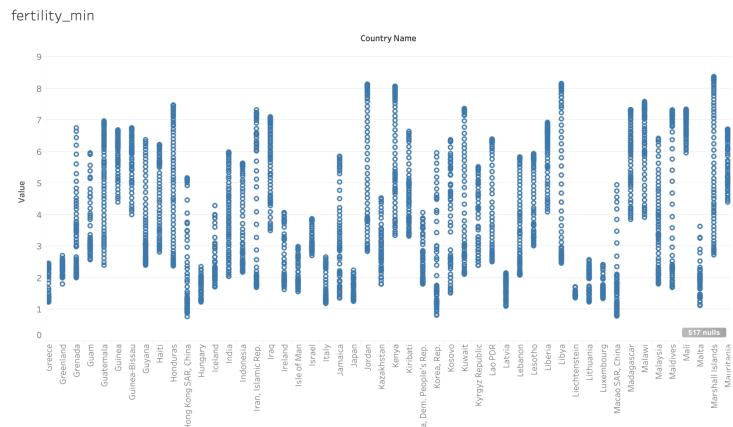
## Appendix F

Screenshot of visualisation of Liechtenstein, which is showing an unusual pattern and is inconsistent with google search results of other available dataset. (The sine wave is the pattern for Liechtenstein sex ratio)

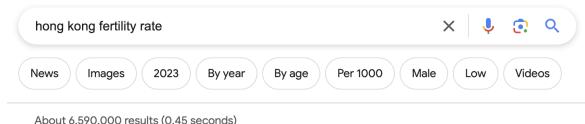


## Appendix G

Screenshot for visualisation used to find minimum value of fertility rate:

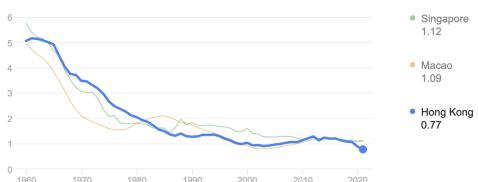


Screenshot for google search of Hong Kong fertility rate to validate the authenticity of this minimum value:



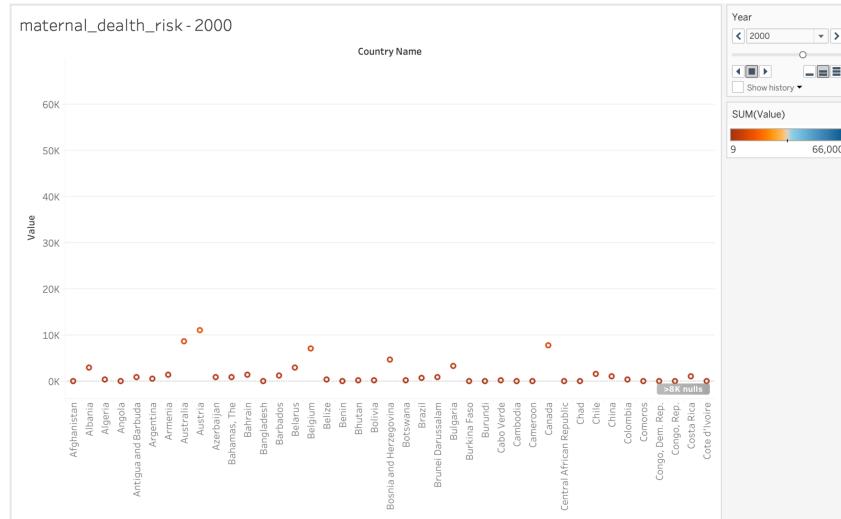
Hong Kong / Fertility rate

0.77 births per woman (2021)

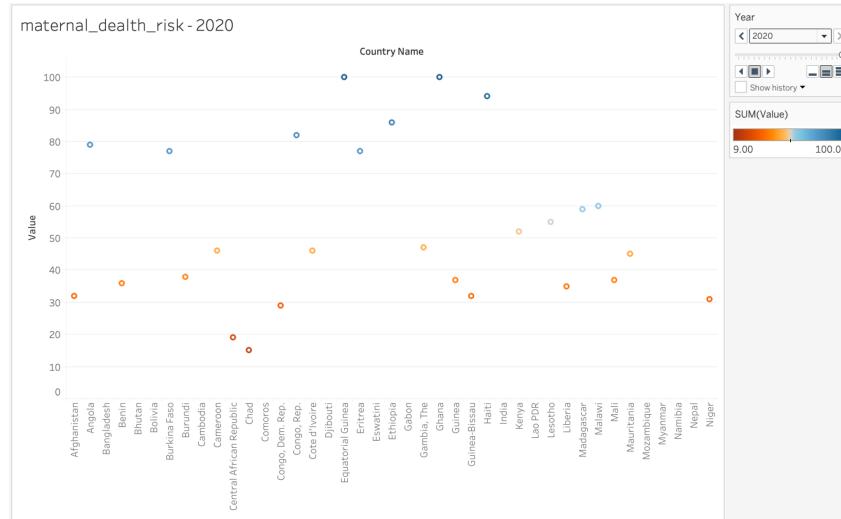


## Appendix H

Screenshot of maternal death risk level across countries in 2000:

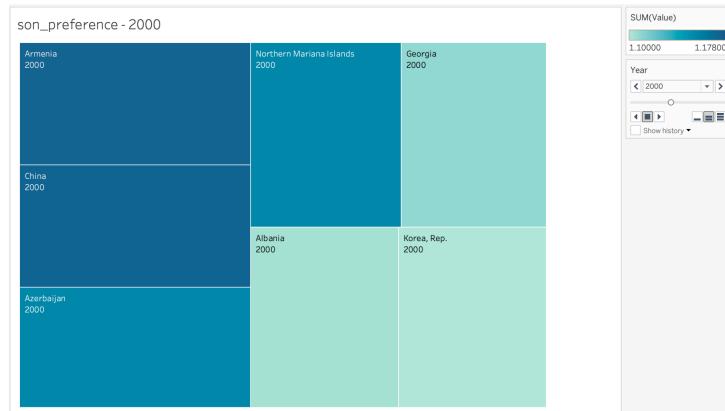


Screenshot of countries with high maternal death risk less than 1 in 100 women in 2020:



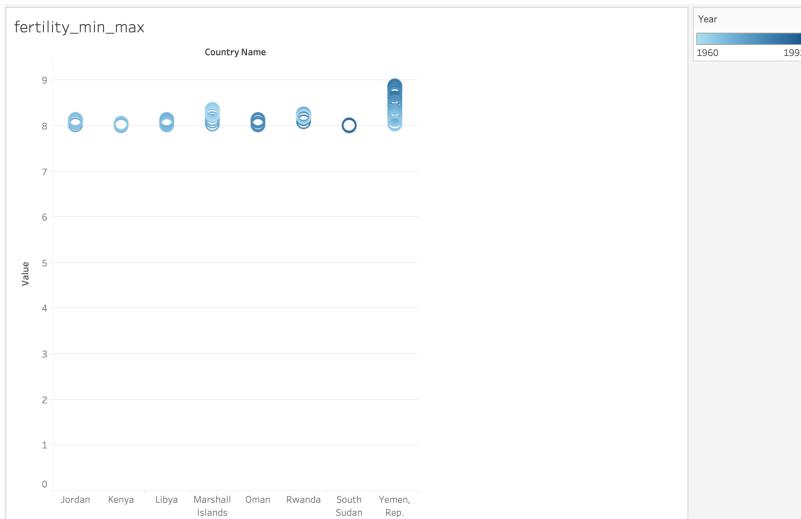
## Appendix I

Screenshot of countries with high sex ratio in 2000:



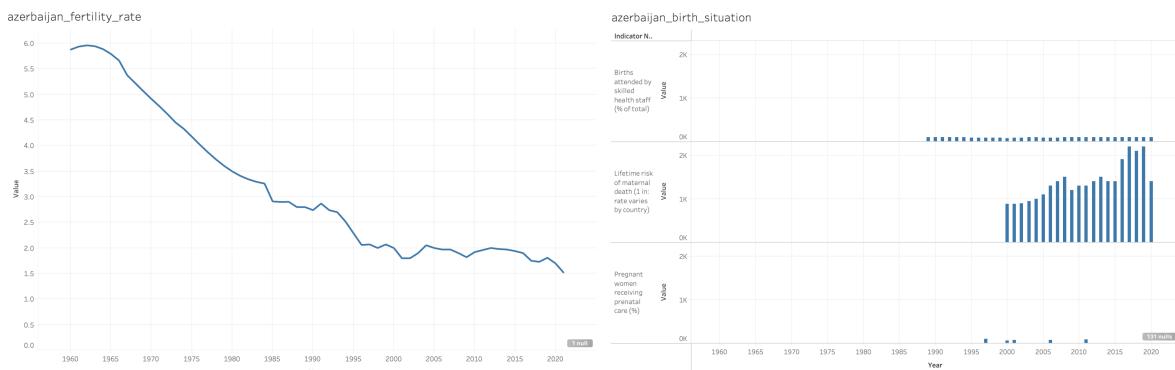
## Appendix J

Screenshot of fertility min max exploration, adjusted with both year slider and colour coded year during exploration process:



## Appendix K

Screenshot of fertility rate and childbirth related situation in Azerbaijan:



Screenshots of days of leave and fertility rate in Japan:

