

# World Birth Rate Analysis

Data 604

By: Saurabh Anand, Daniel Gu, Romi Punian & Charles  
Roberts

# Introduction

The role and status of women around the world bears a direct impact on birth rates and demographic shifts. Female education, female labour force participation and contraception are some very important topics we will consider in our study on birth rates. Our overall data revolves around worldwide data for birth rates along with various other data sets relating to socio-economic conditions affecting females around the world. We believe this data is not just interesting to analyze but also can help us better understand the world around us. This data would be useful for the governments of the world to understand how birth rates can be affected by different socio-economic conditions, and this data may be helpful in understanding how to solve or mitigate population issues. We understand that data on birthrates is an issue that predominantly affects females, and therefore the research questions we came up with, aligned with our datasets will be more focused on females.

## Role of Labour Force

When more women work, economies grow. Increasing females' education also contributes to women going into the workforce, and therefore higher economic growth. However, women still earn substantially less than men and thus there is a disproportionate participation of women in the labor force across countries worldwide. This disproportion also causes an opportunity cost for women in many cases, to choose between pursuing a career or raising a family. Some sociologists suggest that this opportunity cost can be mitigated by progressive social policy that distributes child rearing between both parents, and social welfare that allows women to balance their family lives with careers (Derose and Stone, 2021). Unemployment and reduced participation in the labor force can also affect birth rates across countries and we feel this would be interesting to investigate. We will try to assess these concerns and study how women's participation in the labor force will impact birth rates, as well as what may cause this correlation.

## Role of Contraception

We are progressing into a world where better and regular access to contraceptives is increasing consistently. While some countries have easy access to contraception, some don't and we wanted to study how the distribution of contraceptive prevalence across countries relates to birth rates.

## Role of Education

Female education becomes an interesting factor while considering birth rates. Investments in women's education have proven especially effective at lowering fertility rates, since better educated women tend to marry later and have fewer, healthier children. Studies by several countries show that as literacy rates rise, especially those of women, income levels, nutrition

levels, and child survival rates rise as well. Thus we feel assessing the role of education would be a suitable metric to study birth rates.

## Role of Mortality

Developing and developed countries have gone through the demographic transition from high birth rates and high death rates to low birth rates and low death rates this has led to population ageing. Population ageing is a serious problem throughout the world today, and we believe that exploring these trends deeper can give us a better understanding of this issue, perhaps allowing us to gather insights on how governments can work to solve this problem.

## Cleaning and Coding techniques

The majority of our datasets were sourced from the World Bank (other than the Continent dataset). As such there was a level of standardization in terms of the format of the datasets, all datasets included the same countries, and the same range of years (1960-2021). Because of this our team used a standardized approach to data cleaning which we applied to all of our individual datasets so that they would be organized and formatted the same way making it easier to join them once they were uploaded to our SQL database.

For our individual datasets sourced from the World Bank, we found that for many datasets there were empty values for a few countries that resulted in some inconclusive results. This was possibly due to the unavailability of data during those years but this is something which we wish to improvise on for our future analysis, perhaps by using new data wrangling techniques or research more intensively to gather organized cohesive datasets so that we obtain the correct conclusions. To deal with this we used SQL queries which ignored null values and stored these queries in data frames that we used for further exploration and visualizations.

We visualized our data with SQL queries we put into dataframes, from there utilizing matplotlib and plotly and were able to make some interactive animations to visualize the changes throughout the world.

```
cursor.execute('''create table world_birth_labor_contrc_liter as
    (select a.`Country Code`,a.`Country Name`,a.`Continent`,a.`Year`,a.`Birth Rate`,
        b.`Literacy rate`,c.`Contraceptive Prevalence`,
        d.`Ratio of female to male labor force participation rate (%)` Ratio_female
    from worldbirthrate a,
        worldliteracyrate b, worldcontraceptiveprevalence c,
        world_female_labour_force_participation_continents d
    where      a.`Country Code` = b.`Country Code`
        and    b.`Country Code` = c.`Country Code`
        and    c.`Country Code` = d.`Country Code`
        and    a.`Country Name` = b.`Country Name`
        and    b.`Country Name` = c.`Country Name`
        and    c.`Country Name` = d.`Country Name`
        and    a.Continent = b.Continent
        and    b.Continent = c.Continent
        and    c.Continent = d.Continent
        and    a.`Year` = b.`Year`
        and    b.`Year` = c.`Year`
        and    c.`Year` = d.`Year` ) ''')
print ("combined table world_birth_labor_contrc_liter is created \n")
```

After exploring our individual datasets we used the above SQL Join technique to join our datatables together into a large single datatable. We then did further SQL queries which were used to do a NumPy correlation analysis so we could explore how the different factors correlated with birth rates. We will discuss this further in the Data Exploration Section.

## Research Questions

We chose datasets that would likely be contributing factors to birth rates for a country. Our research questions will analyze if these contributing factors have an effect on birth rates for countries around the world. Our research questions are as follows:

1. What are the birth rate and mortality rate trends in the world?
2. What is the trend for female labour participation around the world? How does female labour participation correlate to birth rates?
3. Does the prevalence of contraceptive use correlate with low birth rates for a country?
4. Is there a correlation between female School Enrolment and birth rates?

## Data Exploration

We used many different queries throughout this project to learn more about our data and explore. With our individual and joined datasets we used SQL queries and Python, which helped us to explore our guiding questions.

From our Joined datasets we used queries to help us gather information to compare various factors with birth rates, and used python to find the correlation between factors and birth rates. For further exploration we looked deeper at these correlation coefficients for specific continents, or groups of countries we found interesting.

We will talk more in depth about what we learned from our exploration in the discussion section.

1. Individual tables, which were loaded into DB directly
  - Table: worldbirthrate
    - Main columns:Country Code, Country Name, Year, Birth rate
    - Content: All countries' birth rate over 60 years( 1960 to 2019)
    - Queries Used:
      - `select a.`Birth Rate`,a.`Country Code`,a.`Country Name`,a.`Year` from worldbirthrate a`
        - To show each country's birth rate among 60 years.
      - `select ROUND(AVG(`Birth Rate`),4),Continent,`Year` from worldbirthrate group by Continent,`Year`;`
        - This SQL lets us know birth rates' changing for each Continent among 60 years.
  - Table: worldmortalitydata

- Main columns: Country Code, Country Name, Year, Mortality rate
- Content: All countries' mortality rate over 60 years( 1960 to 2019)
- Queries Used:
  - `select ROUND(AVG(`Mortality Rate`),4),Continent,`Year` from worldmortalitydata group by Continent,`Year`;`
    - This SQL lets us know mortality rates' changing for each Continent among 60 years.
- Table: world\_female\_labour\_force\_participation\_continents
  - Main columns: Country Code, Continent, Country Name, Ratio of female to male labor force participation rate (%), Year
  - Content: all countries female labour force ratio's from 1960-2020
  - Queries Used:
    - `'SELECT * FROM `World_Female_Labour_Force_Participation_Continents` WHERE `Ratio of female to male labor force participation rate (%)` > 0 ORDER BY `Country Code` ASC ;'` takes every country in alphabetical order and displays the Ratio of female to male labor force participation rate (%) for each year that there is a recorded value (above 0) for that country.
    - `'SELECT * FROM `World_Female_Labour_Force_Participation_Continents` WHERE `Year` = 2010 ORDER BY `Ratio of female to male labor force participation rate (%)` DESC LIMIT 25;'` displays the top 25 countries with highest female labour force participation ratio for the year 2010.
    - `'SELECT * FROM `World_Female_Labour_Force_Participation_Continents` WHERE `Year` = 2010 AND `Ratio of female to male labor force participation rate (%)` > 0 ORDER BY `Ratio of female to male labor force participation rate (%)` ASC LIMIT 10;'` displays the bottom 25 countries with recorded values for the year 2010 (excluding countries with no recorded values/0 values as they would dominate the list).
    - `""select ROUND(AVG(NULLIF(`Ratio of female to male labor force participation rate (%)`,0)),4),Continent,`Year` from World_Female_Labour_Force_Participation_Continents group by Continent,`Year` ;""` gives us the average for every country for all years with recorded values for that country.
- Table: worldcontraceptiveprevalence
  - Main columns: Country Name, Country Code, Continent, Contraceptive prevalence, any methods (% of women ages 15-49), Years (1960-2020)
  - Content: all countries contraceptive prevalence % from 1960-2020
  - Queries:
    - `SELECT * FROM `worldcontraceptiveprevalence`.` Returns all world contraceptive prevalence for all countries and continents along with country code and year.

- **ROUND(AVG(NULLIF(`Contraceptive Prevalence`,0)),4),Continent,`Year`** from **worldcontraceptiveprevalence** group by **Continent,`Year`**. Returns mean contraceptive prevalence for all continents along with year and not including zero value cells for calculations.
- **SELECT \* FROM `worldcontraceptiveprevalence` WHERE `Contraceptive Prevalence` > 0 ORDER BY `Country Code` ASC.** Returns database with all zeros removed by country in ascending order along with country code, continent, contraceptive prevalence and year.
- **SELECT \* FROM `world\_contraceptive\_prevalence` WHERE `Year` BETWEEN 2010 AND 2019 ORDER BY `Contraceptive Prevalence` DESC LIMIT 10.** Returns the top 10 countries' contraceptive prevalence between 2010 AND 2019 along with country code and year.
- **SELECT ROUND(AVG(NULLIF(`Contraceptive Prevalence`,0)),4),`Country Name`** from **world\_contraceptive\_prevalence** group by **`Country Name`**. Returns the average contraceptive prevalence for each country (1960 to 2020).

- Table: worldBankLiteracyRate

- Main columns: Country Name, Country Code, Continent, Literacy Rates, any methods (% of women ages 15-49), Years (1960-2020)
- Content: all countries literacy rates % from 1960-2020
- Queries Used:
  - **"SELECT `Continent`,AVG(`Literacy rate`) as `Literacy rate` FROM `worldliteracyrate` WHERE `Year` = 2019 GROUP BY `Continent` ORDER BY AVG(`Literacy rate`) DESC.** Returns Best Literacy Rate wise Continent in the year 2019
  - **SELECT a.`Continent` as `Continent`,AVG(`Literacy Rate Year 2019`),AVG(`Literacy Rate Year 2001`)FROM ( (SELECT Continent,`Literacy rate` as `Literacy Rate Year 2019` FROM worldliteracyrate Where Year = 2019) INNER JOIN (SELECT Continent,`Literacy rate` as `Literacy Rate Year 2001` FROM worldliteracyrate Where Year = 2001) b on a.Continent = b.Continent) GROUP BY a.`Continent`.** Continent-wise Change in Literacy Rate of 2019 with Respect to the Literacy Rate of 2001
  - **SELECT CountryName,(100\*((Year\_2011-Year\_2001)/Year\_2001)) as GrowthDecade\_2001\_2011** FROM **`world\_bank\_literacy\_rates`** WHERE **Year\_2001 <> 0 ORDER BY GrowthDecade\_2001\_2011 DESC LIMIT 10.** Which Country had best Literacy Growth in the Decade (2001 - 2011)
  - **SELECT CountryName,(100\*((Year\_2011-Year\_2001)/Year\_2001)) as GrowthDecade\_2001\_2011** FROM **`world\_bank\_literacy\_rates`** WHERE **Year\_2001 <> 0 ORDER BY GrowthDecade\_2001\_2011 ASC LIMIT 10.** Which Country had lowest Literacy Growth in the Decade (2001 - 2011)- Which Country had lowest Literacy Growth in the Decade
  - **SELECT CountryName,(100\*((Year\_2019-Year\_1977)/Year\_1977)) as**

**GrowthRate\_1977\_2019 FROM `world\_bank\_literacy\_rates` WHERE Year\_1977 <> 0 ORDER BY GrowthRate\_1977\_2019 DESC.** Which Country showed the highest Growth in Literacy Rate from (1977-2019)

2. Joined tables (Individual Datasets combined for further exploration)

- Table: world\_birth\_labor\_contrc\_liter
  - Main columns:
    - Country Code, Continent, Country Name, Year, Birth Rate, Literacy rate, Contraceptive Prevalence, Ratio\_female (female labour force participation ratio)
  - Content: This table joined main columns from above individual tables related to our research topic, including 215 countries, 7 continents and 60 years, a total 12900 rows. We used it for correlation analysis mainly.

**SQL queries and code for Correlation between ratio of female labor and birth rate:**

```
sql = 'select * from world_birth_labor_contrc_liter a where a.`Ratio_female` != 0.0;'  
result = pd.read_sql(sql, conn)  
corrbr = np.corrcoef(result['Birth Rate'].astype(float), result['Ratio_female'].astype(float))
```

Above is an SQL query and the python code we used to find the overall correlation between birth rates and female labour participation from our joined datatable.

```
j = 0  
corrall = [0]*6  
for i in continents:  
    sql = 'select a.`Birth Rate` , a.`Ratio_female` from world_birth_labor_contrc_liter a where a.`Ratio_female` != 0.0 and a.`C`  
    result = pd.read_sql(sql, conn)  
    corrall[j] = np.corrcoef(result['Birth Rate'].astype(float), result['Ratio_female'].astype(float))  
    j += 1
```

Above is a loop that uses a query to get the correlation between birth rates and female labour participation for each continent.

```
countries = ['Sweden', 'Norway', 'Denmark', 'Finland', 'Iceland', 'Netherlands']  
j = 0  
corrall = [0]*6  
for i in countries:  
    sql = 'select a.`Birth Rate` , a.`Ratio_female` from world_birth_labor_contrc_liter a where a.`Ratio_female` != 0.0 and a.`C`  
    result = pd.read_sql(sql, conn)  
    corrall[j] = np.corrcoef(result['Birth Rate'].astype(float), result['Ratio_female'].astype(float))  
    j += 1
```

Above is a loop that uses a query to get the correlation between birth rates and female labour participation for each Nordic Country between 1985-1996 (did a similar query for all dates and a query for 1995-2006 to show how 1985-1996 is drastically different from the overall correlation).

**SQL queries and code for correlation between prevalence contraception and birth rate**



```
sql = 'select * from world_birth_labor_contrc_liter a where a.`Contraceptive Prevalence` != 0.0;'
result = pd.read_sql(sql,conn)
corrbc = np.corrcoef(result['Birth Rate'].astype(float), result['Contraceptive Prevalence'].astype(float))
```

Above is an SQL query and the python code we used to find the overall correlation between birth rates and contraceptive prevalence from our joined datatable.

```
continents = ['Asia', 'Africa', 'Northern Americas', 'Southern Americas', 'Europe', 'Oceania']
j = 0
corrall = [0]*6
for i in continents:
    sql = 'select a.`Birth Rate` , a.`Contraceptive Prevalence` from world_birth_labor_contrc_liter
    'a where a.`Contraceptive Prevalence` != 0.0 and a.Continent = ' +str(i)+' '
    result = pd.read_sql(sql,conn)
    corrall[j] = np.corrcoef(result['Birth Rate'].astype(float), result['Contraceptive Prevalence'].astype(float))
    j += 1
```

Above is a loop that uses a Query to get the correlation between contraceptive prevalence for each continent.

### SQL queries and code for correlation between literacy rate and birth rate

```
sql = 'select a.`Birth Rate` , a.`Literacy rate` from world_birth_labor_contrc_liter a where a.`Literacy rate` != 0.0
and a.Continent = ' +str(i)+' '
result = pd.read_sql(sql,conn)
corrall[j] = np.corrcoef(result['Birth Rate'].astype(float), result['Literacy rate'].astype(float))
```

Above is an SQL query and the python code we used to find the overall correlation between birth rates and literacy from our joined datatable.

## Discussion

### Datasets: World Birth and Mortality Rates (Daniel)

**Research Question:** What are the birth rate and mortality rate trends in the world?

#### Dataset: Birth Rates , Continent

##### Dataset Info:

The dataset is from The World Bank website in their data section and has a Creative Commons Attribution 4.0 International license. It contains the data from 1960 to 2020, including all the main countries in the world. The dataset displays the number of new babies in 1000 people each year. The dataset was last updated in 2019 and includes data for Country Name, Country Code, Birth rate, and Years from 1960 to 2020. The original number of rows is 15,596. After joining with the Continent dataset, the number of rows left 12,900 with 249 countries.

The Continent dataset is from the Kaggle, which was last updated in 2019, including Country code, Continent correspondence relationship.



## Dataset: Mortality Rates

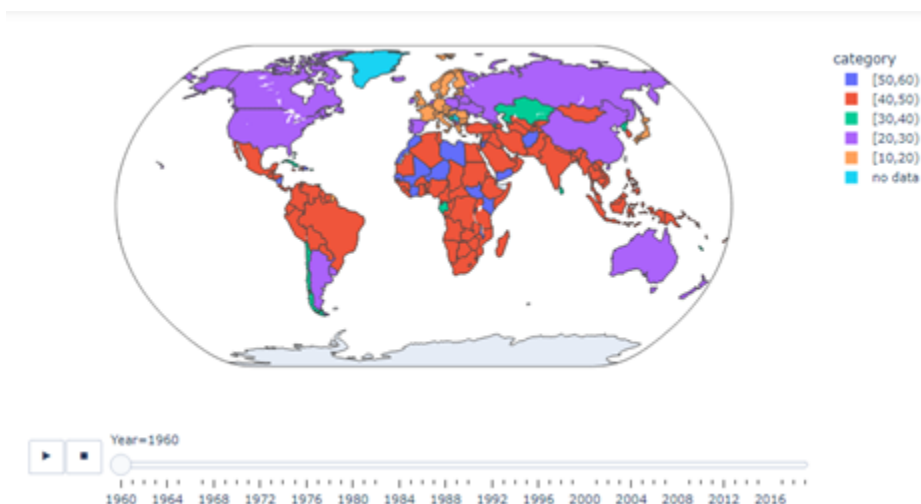
### Dataset Info:

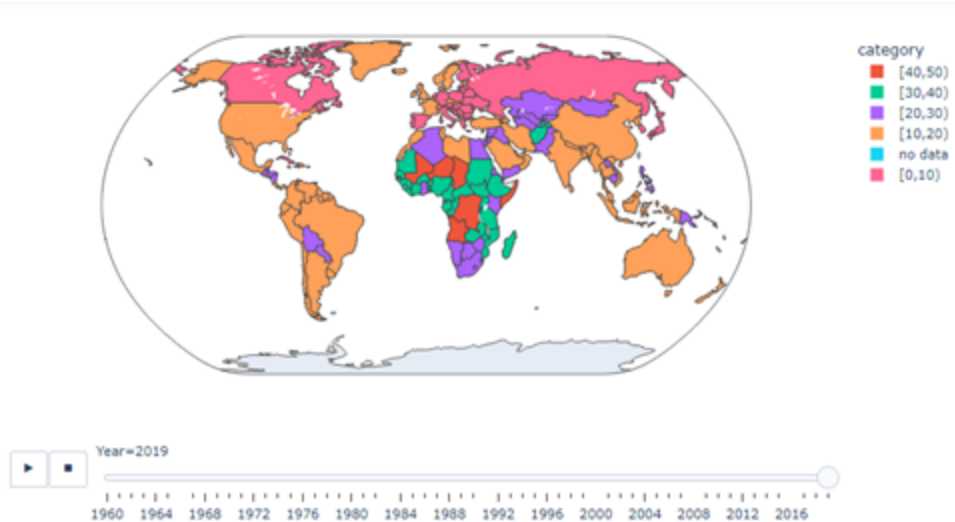
The dataset is from The World Bank website in their data section and has a Creative Commons Attribution 4.0 International license. It contains the data from 1960 to 2020, including all the main countries in the world. The dataset displays the number of deaths in 1000 people each year. The dataset was last updated in 2019 and includes data for Country Name, Country Code, Mortality rate, and Years from 1960 to 2020. The original number of rows is 15,596. After joining with the Continent dataset, the number of rows left 12,900, with 249 countries.

### Discussion/Analysis

As we know population ageing is a global phenomenon, and the two primary factors leading to it are mortality rate and birth rate. But what are the birth rate and mortality rate trends in the world?

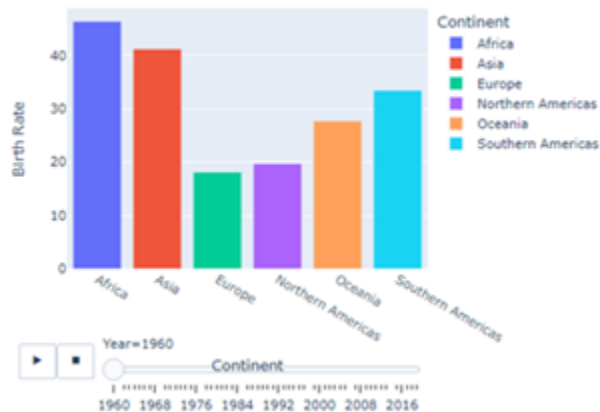
We explored all birth rates over 60 years in each country using animated charts, and found the birth rate is declining gradually. In 1960, there were a lot of countries' with birth rates over 40 new births per 1000 people (red or dark blue), but in 2019, most countries' birth rates lower 20 new births per 1000 people.



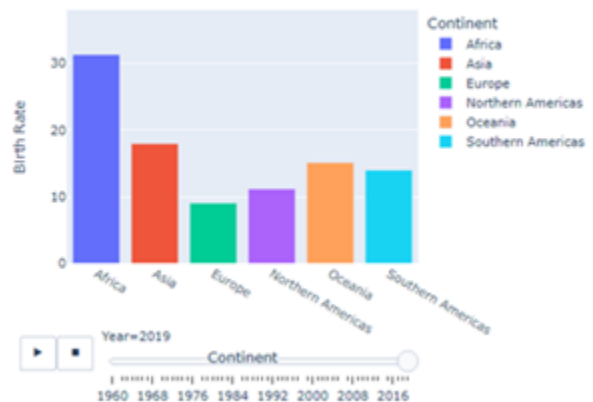


After focusing on continent mean birth rate data, it is more obvious that the birth rate is decreasing, from the highest over 40 in 1960 to below 30 in 2019. And we also found, the mean birth rate in Africa is always the highest among all continents for this period.

Trends for World Birth Rate by Continents from 1960 - 2019

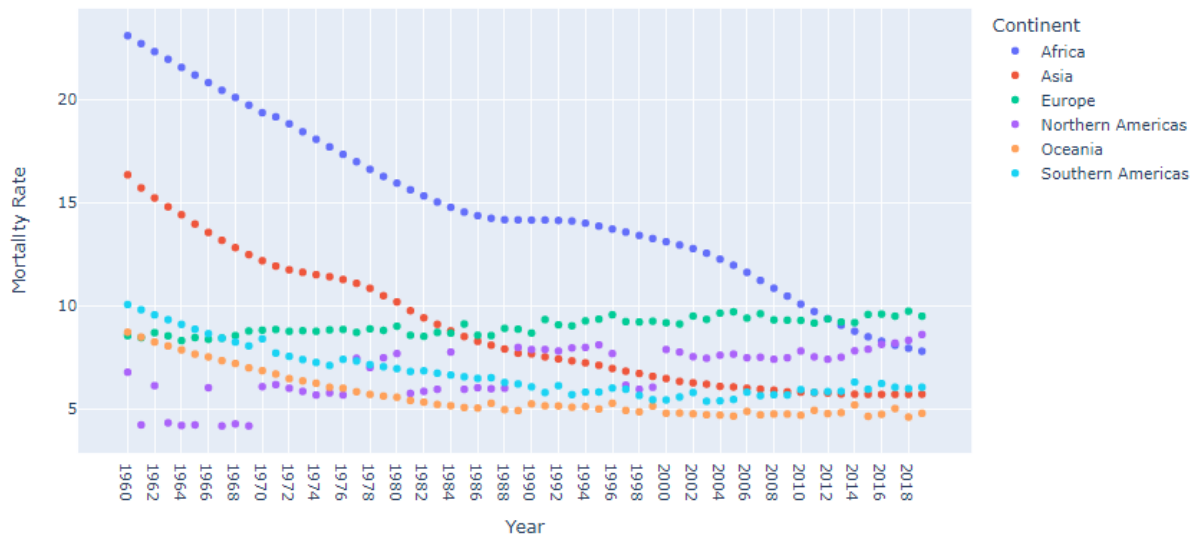


Trends for World Birth Rate by Continents from 1960 - 2019



Because mortality is one factor affecting population aging. From the chart, we can see the mortality rates in the world are continuing to decrease. As well we can observe that the mortality rate of Africa and Asia is reducing quickly after increasing the quantity and quality of modern health services. Although other continents' mortality rate decrease slightly, the general trends for most continents' mortality rate is stable and low.

Mortality rate per continent by 1960 - 2019



From the above analysis, we can draw one conclusion: the aging population issue in the world will become a problem in the future if action is not taken to solve it. Population aging is an extremely pressing matter for the world moving forward as an aging and increasing population will require more resources to sustain (United Nations, 2017). It is therefore imperative that a cooperative effort is made to address and mitigate the worst effects of this problem.

### Dataset: Female Labour Participation (Romi Punian)

**Research Question:** What is the trend for female labour participation around the world? How does female labour participation correlate to birth rates?

### Dataset Info:

The dataset for Labor force participation ratio displays the ratio of the population female in comparison to males. The dataset is from The World Bank website in their data section and has a Creative Commons Attribution 4.0 International license. It looks at individuals ages 15 and older that are economically active from the years 1960 to 2020 (although for some countries some cells are blank as no data was recorded for that year, this will have to be filtered out). This is calculated by dividing female labor force participation rate by male labor force participation rate and multiplying by 100. For the ratio in the dataset a ratio of 0 would mean no female participation, and 1 means equal female to male participation in the labor force. The dataset contains data for all people who supply labor for the production of goods and services during a specified period.

The dataset was last updated on 2020-09-29 and includes data for Country Name, Country Code, Indicator name, Indicator Code and Years from 1960 to 2020. Ratio of female to male labor force participation rate (%) (national estimate). Data on the labor force are compiled by the

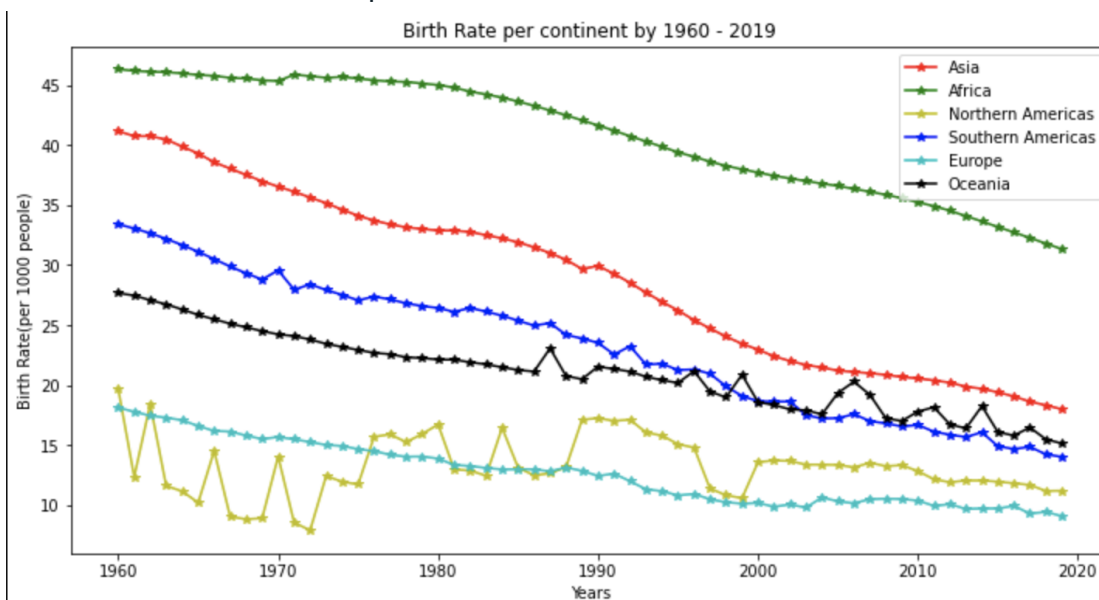
ILO from labor force surveys, censuses, and establishment censuses and surveys. For some countries a combination of these sources is used. Labor force surveys are the most comprehensive source for internationally comparable labor force data. The features we are interested in analyzing are the Indicator Code, Country Name and data for the Years. The dataset CSV file is presented in tabular format in 271 rows and 64 columns for a total of 17,615 files or cells.

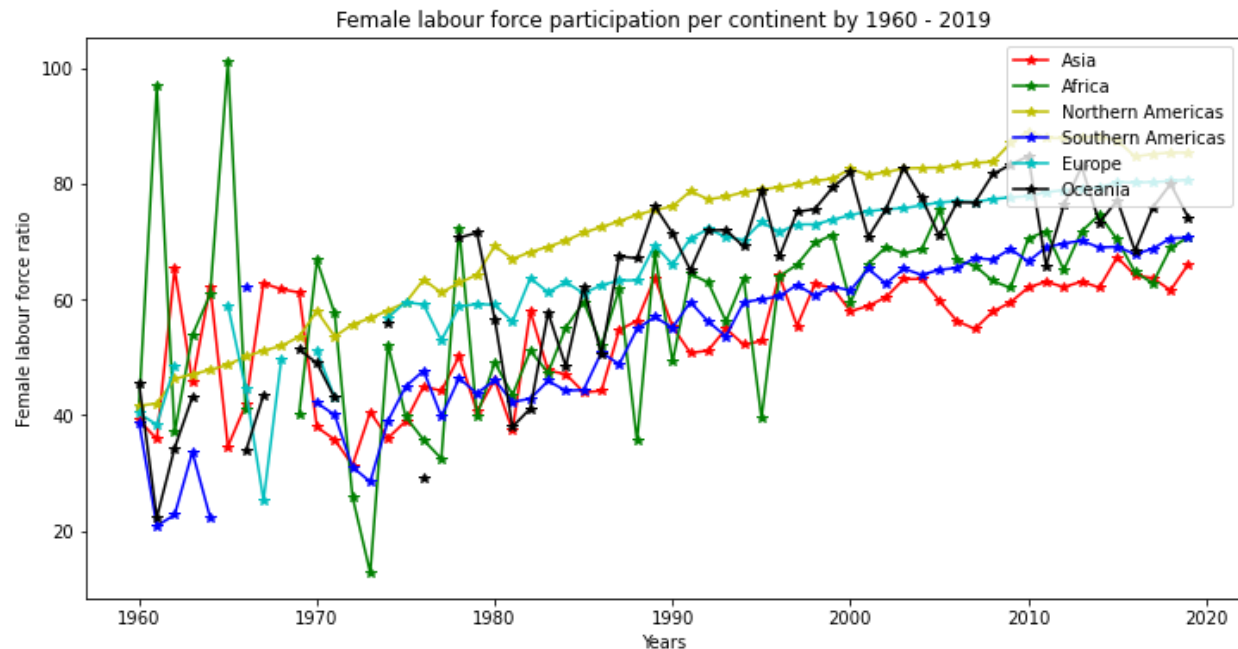
## Discussion/Analysis and Data Exploration

It is no secret that female labour force participation has increased significantly throughout the 20th and 21st centuries, the effects of this increase has been of great interest to economists and sociologists alike (Bloom et al., 2009). For the purposes of this paper we are particularly interested in looking at female labour force participation trends throughout the world, and how this correlates with birth rates.

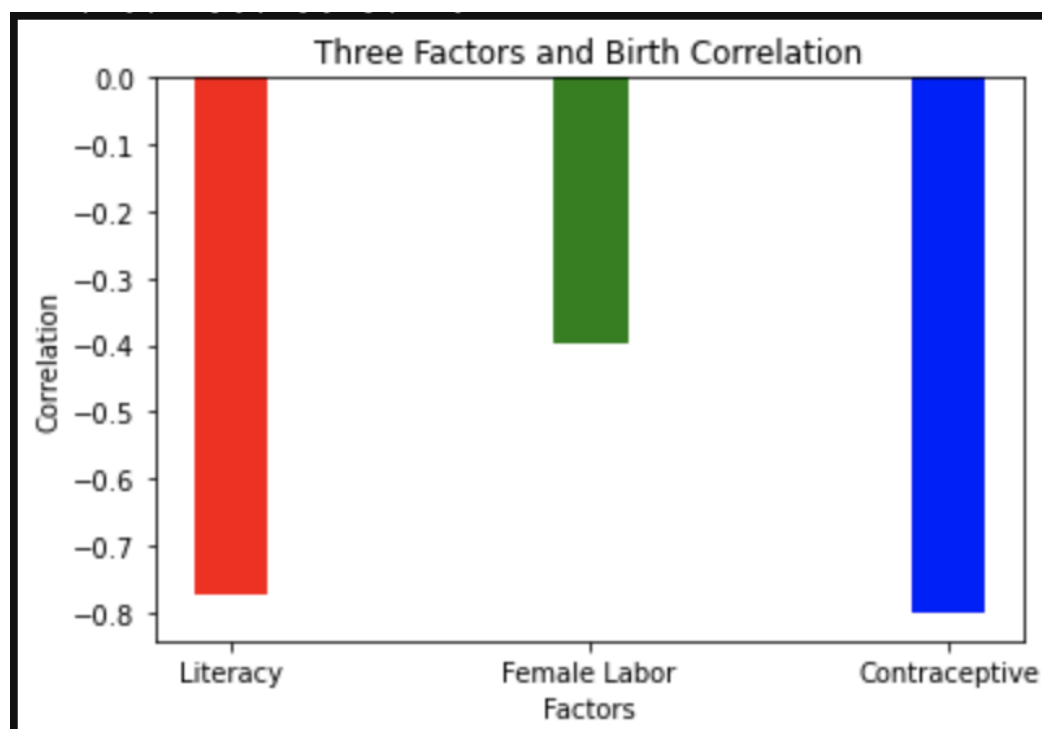
According to (Roser, 2014) following the world war period labor markets underwent historical changes which made the increasing labor force participation of women possible. In historically underdeveloped economies which tended to start off largely agrarian, a large majority of workers were involved in agriculture. Because of the physical demands of agrarian societies, men tended to have an advantage in the labour market over women.

However, as labour costs increased in developed countries, the developed economies of the west started outsourcing their manufacturing industries to the historically underdeveloped and exploited economies of the global south to capitalize on cheaper labour markets. This along with increasing importance of education due to technological development allowing for the erosion of the comparative advantage men had over women in these economies (Roser, 2014). That is to say Women's participation in the labour force increased as previously underdeveloped economies started to develop and industrialize.



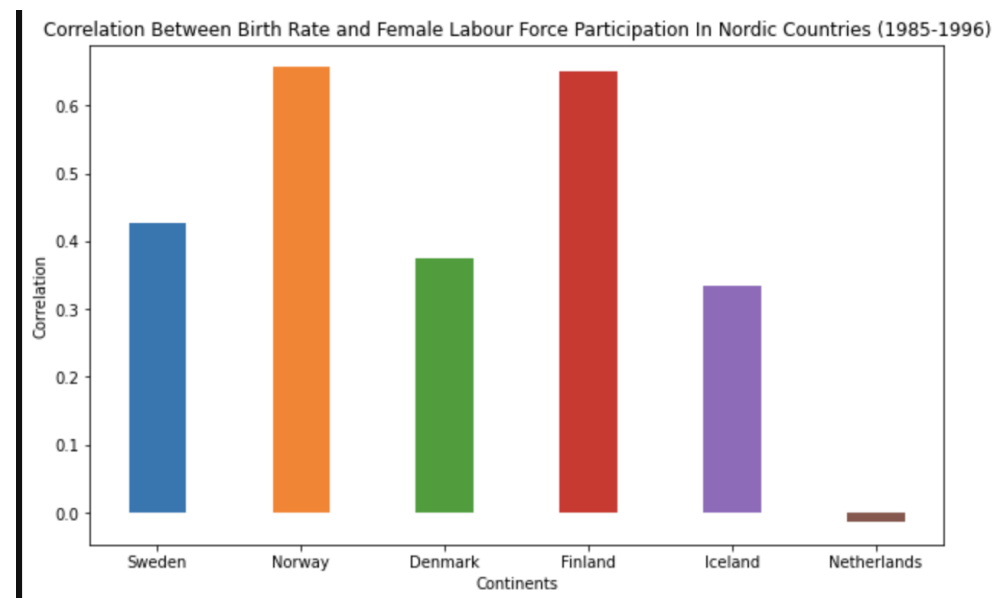


From the above charts we see that there is an overall birth rate decrease throughout the entire world as time goes on, and we see that inversely female labour participation increases as time goes on. This is expected, as many economists note that the overall trend is that as female labour rates increase there is a negative correlation with birth rates (Bloom et al., 2009). While this appears obvious enough we would like to explore this correlation in further detail.

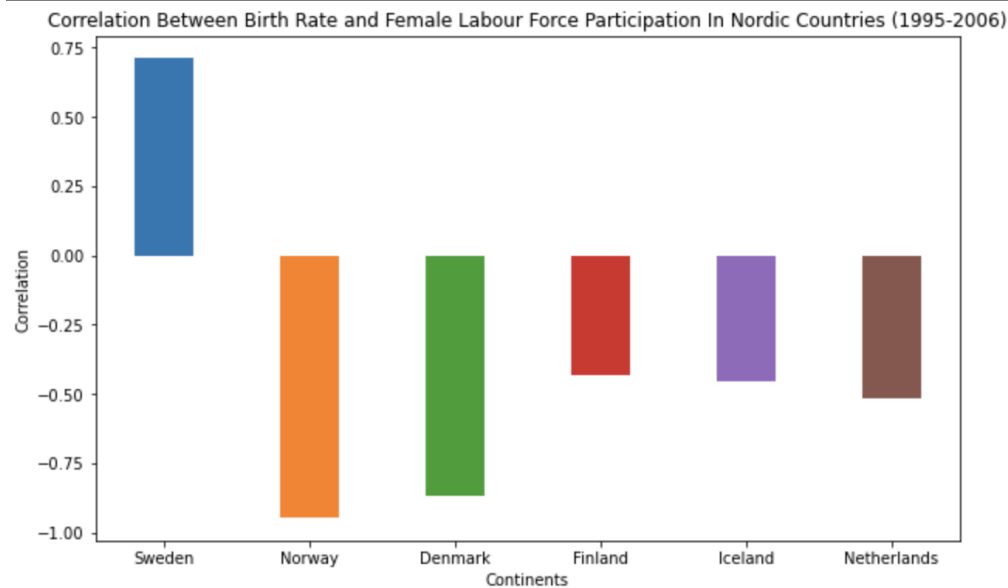


To explore our data further we looked at the overall correlation for female Labor and Birth rates and found that in comparison to Contraceptive Prevalence and Literacy which both were found to be highly negatively correlated with birth rates (correlation coefficients of -0.802 and -0.771 respectively), female labour participation had a lower negative correlation coefficient of -0.393. This means that female labour participation is loosely negatively correlated with birth rates.

We were interested in understanding further what could contribute to this loose correlation. We found that despite the overall trend for highly developed countries with high female labour participation to have low and decreasing birth rates, there were certain cases in which this wasn't true. Particularly within the “Nordic” Countries from the 80's to mid 90's (Derose and Stone, 2021).



As we see from the chart above all of the “Nordic” Countries have positive correlation with birth rates. This unique phenomenon was historically theorized to be the result of progressive social welfare policies that supported childbearing and promoted egalitarian views on gender common to these countries (Derose and Stone, 2021).



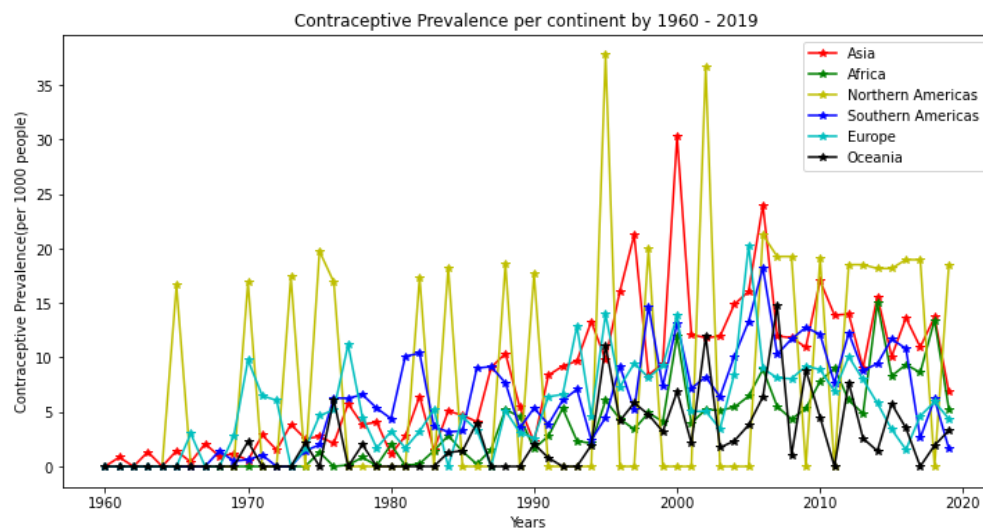
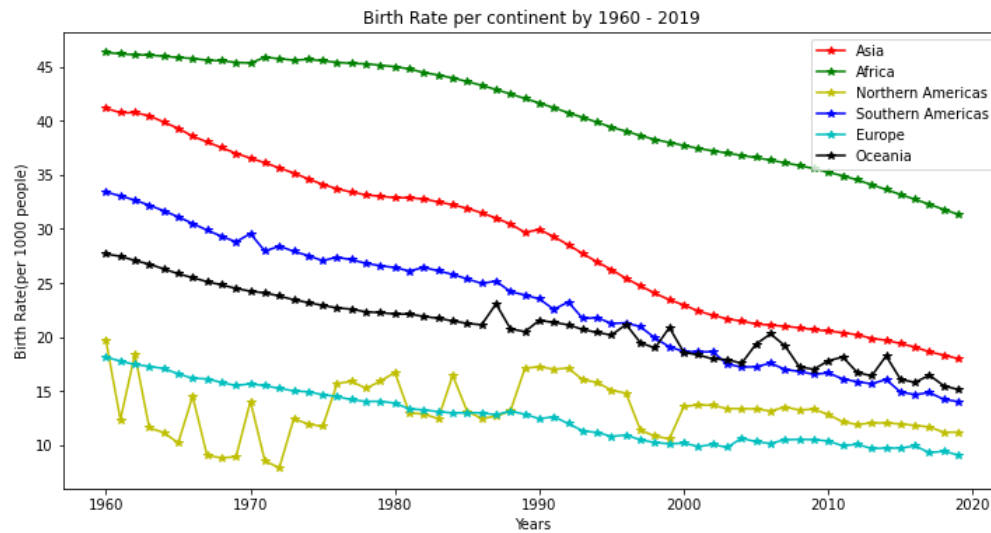
As we can see from the above charts, this decade-long correlation of high birth rates and high female labour force participation however came to an end in the following decade (1996-2006). According to Derose and Stone, 2021, this decrease in birth rates is related to an ideological shift in the population, a phenomenon common in many highly developed countries. This ideological shift marks a shift away from a focus on raising families, and instead a focus on career development dubbed “Workism”.

#### **Dataset: Contraceptive Prevalence (Charles Roberts)**

**Research Question:** Does the prevalence of contraceptive use correlate with low birth rates for a country?

We chose the Contraceptive Prevalence rate dataset because we hypothesised that the existence or acceptance of contraceptive use would influence birth rates for a country. The dataset is from The World Bank website in their data section and sourced from UNICEF's State of the World's Children and Childinfo, United Nations Population Division's World Contraceptive Use, household surveys including Demographic and Health Surveys and Multiple Indicator Cluster Surveys. The dataset includes data for Country Name, Country Code, Indicator name, Indicator Code and Years from 1960 to 2020.



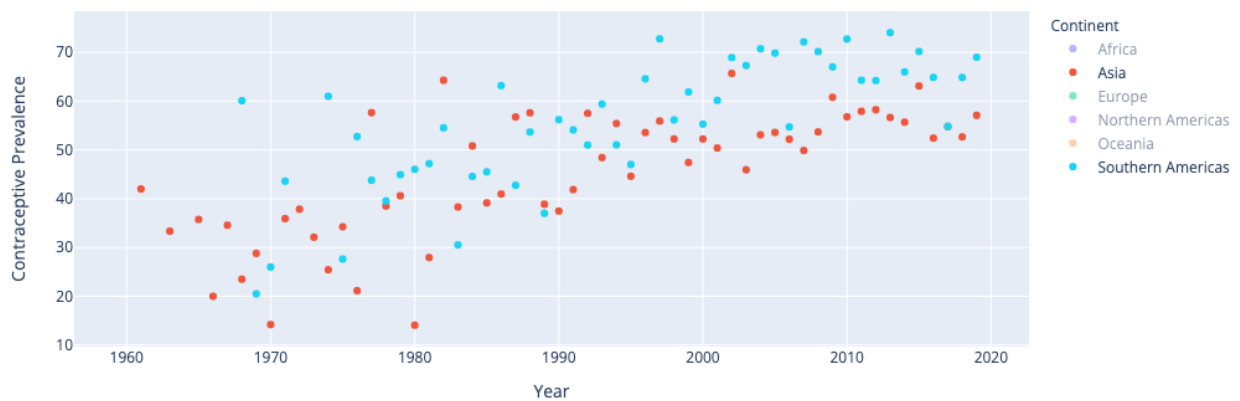


Based on our data analysis, the contraceptive prevalence mean rate among continents shows us that the world's contraceptive prevalence rate slightly increases for all continents from 1960 to 2020. Compared to the birth rates data set we can see a relationship between decreasing birth rates and rising contraceptive use. The United Nations has made a similar claim, *“There is an inverse relationship between contraceptive use and fertility. Countries with high proportions of women using contraception generally have lower levels of fertility”* (United Nations, 2020).

We found it helpful to plot all the contraceptive rates for each continent on a scatter plot and view the trends for one or two continents by turning them on and off. The continents that have the greatest increase in contraceptive prevalence rates from 1960 to 2020 are Southern Americas and Asia. The increase in rates in the Southern Americas may be explained by *“Permanent and long-acting methods, such as female and male sterilisation, IUD and implant are more common in Asia and in Latin America and the Caribbean than in other regions”* (United

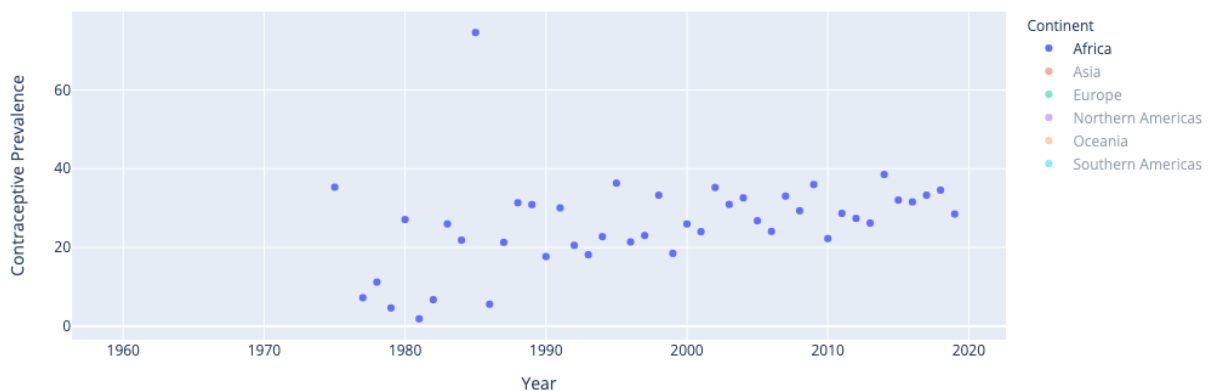
Nations, 2019).

Contraceptive Prevalence mean rate per continent by 1960 - 2019



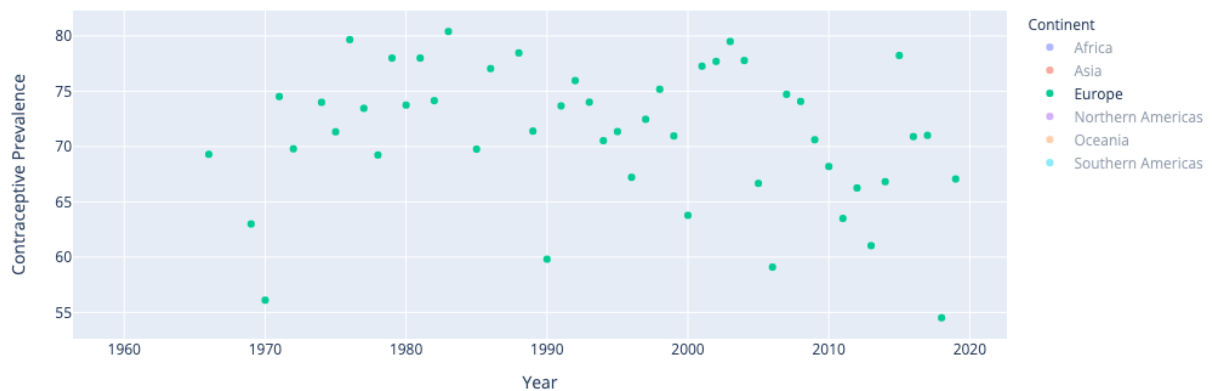
Africa had a slight rise in contraceptive prevalence in the 1980's but has been level since the 1990's. The rise in contraceptive use in Africa could be explained by the AID's epidemic that started in the 1980's; *“Contraception continues to play an integral role in global HIV prevention efforts in the era of increasing HIV treatment coverage, especially in sub-Saharan Africa”* (Sherwood, Lankiewicz, Roose-Snyder, 2021). The fact that this increase in rates only exists for Sub Saharan Africa and not Northern Africa could be explained by, *“More than a third of the countries in Northern Africa and Western Asia, Europe and Northern America and Eastern and South-Eastern Asia lack data for unmarried women”* (United Nations, 2019) and *“Sub-Saharan Africa (SSA) continues to have much higher fertility and lower contraceptive usage than any other region: the contraceptive prevalence rate of 22% is less than half that of South Asia (53%) and less than a third that of East Asia (77%)”* (Glick and Linnemayr, 2016). Considering the decreasing trends of world birth rates, contraceptive use remains prevalent over most major regions.

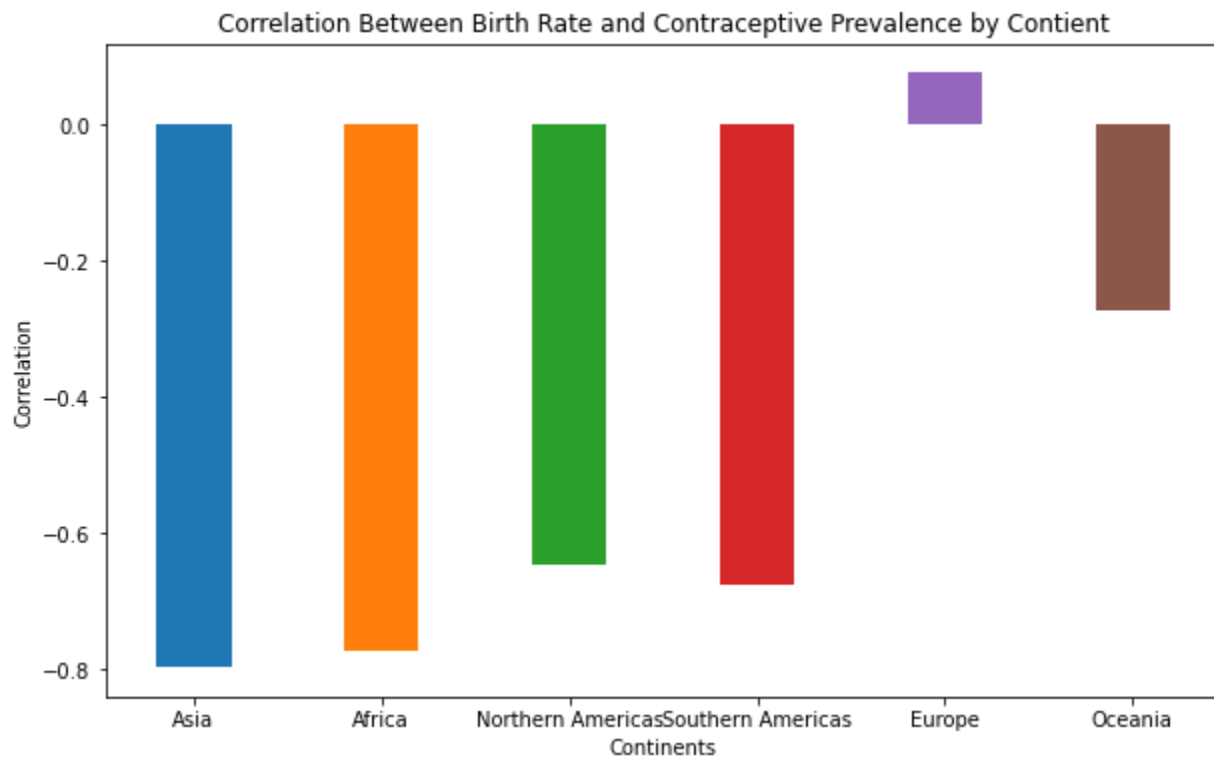
Contraceptive Prevalence mean rate per continent by 1960 - 2019



Contraceptive prevalence rates have maintained a constant rate throughout the years and we can see that most continents are gradually increasing. The only continent that doesn't seem to be following this trend is Europe where they have had consistently high contraceptive rates since the 1960's but seemed to have declined slightly in the last 10 years. The decline in contraceptive rates in Europe could be explained by *"At least one in ten married or in-union women in most regions of the world has an unmet need for family planning"* (United Nations, 2015).

Contraceptive Prevalence mean rate per continent by 1960 - 2019





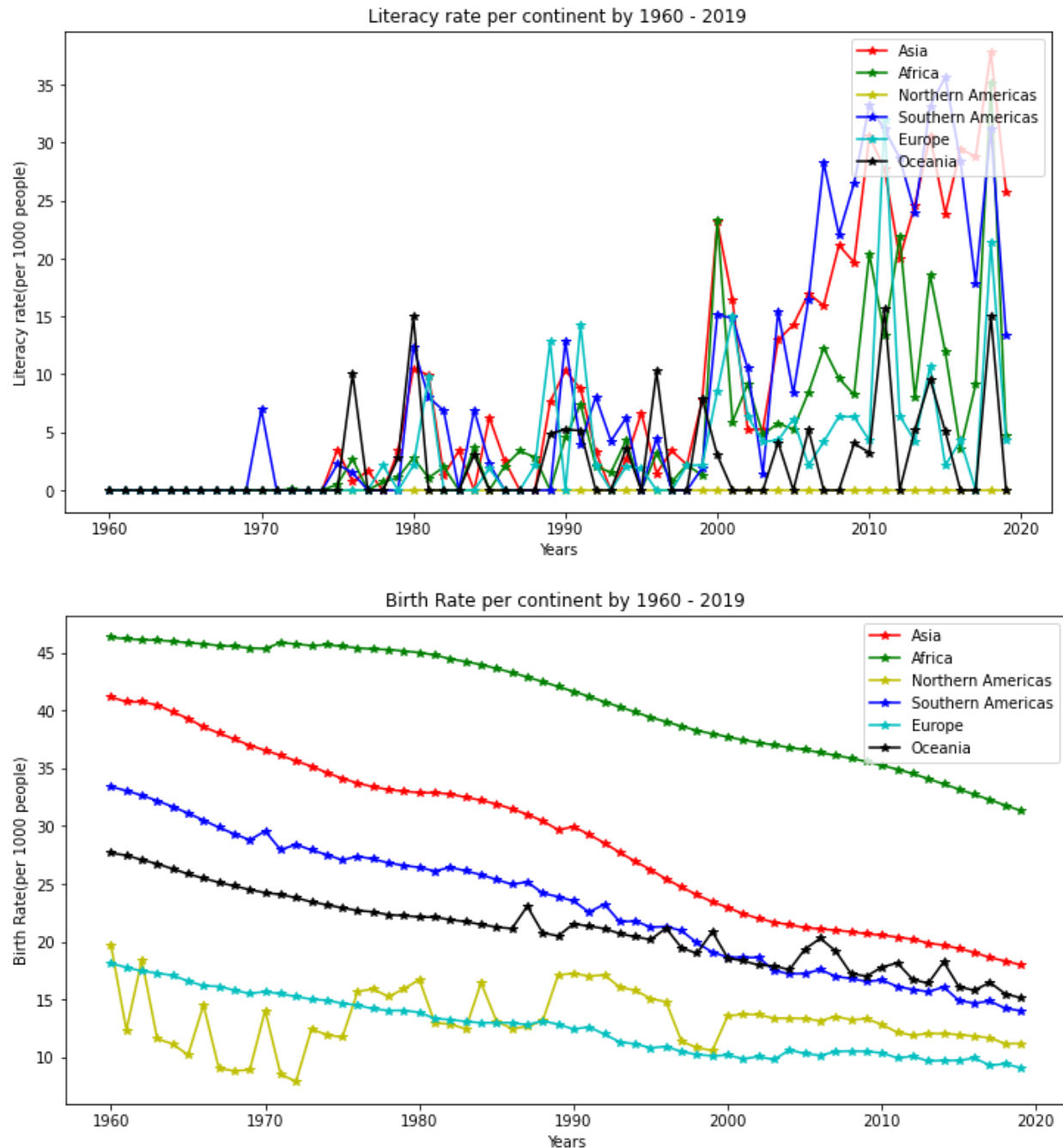
Based on our correlation analysis, birth rates have a negative correlation with contraceptive use except in Europe which has a slightly positive correlation in birth rates. Our correlation analysis shows us that contraceptive prevalence for the world has a 0.80 strong negative correlation against global birth rates. This strong negative correlation can be explained by a significant growth in the use of contraceptives, *“The greatest development in reproductive health in the past few decades has been a significant growth in the use of contraceptives (9% of married women in 1965-70 to 50% in 1985-90). This expanded use of contraceptives has resulted in a considerable fertility decline.”* (cited in National Library of Medicine, 2011). Therefore, as contraceptive use increases, birth rates decrease.

#### **Dataset: World Literacy Dataset (Saurabh Anand)**

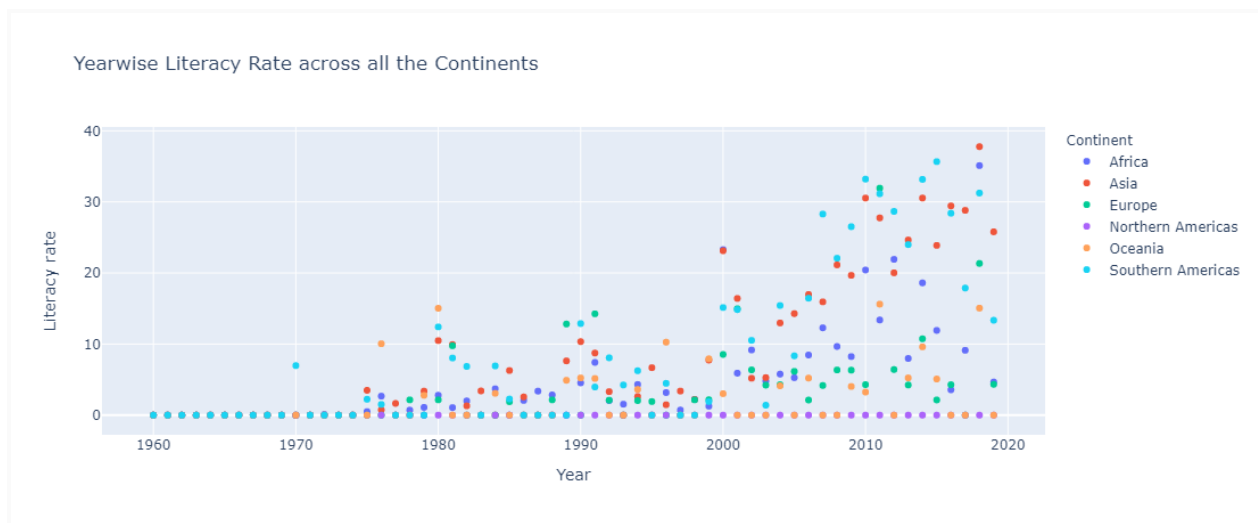
**Research Question: What is the overall literacy trend across countries? Is there a correlation between literacy and birth rates?**

The level of education in a society of women in particular is one of the most important predictors for the number of children families have. Women’s better education and employment opportunities both changed the role of women in society and increased the social status of women in society. Thus we chose to analyse literacy rates across countries and tried to investigate the relationship literacy rate would have with birth rates across world continents and countries.

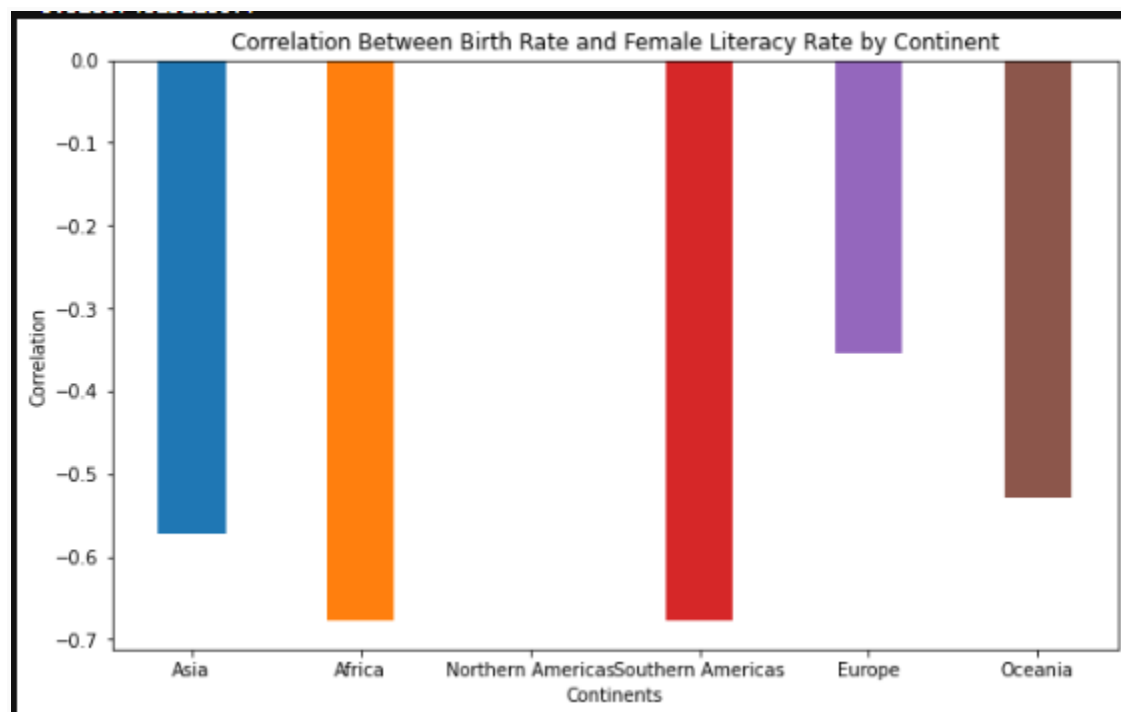
The dataset for literacy examines the literacy rates in secondary school for females across different countries from 1960 to 2020. The dataset includes data for Country Name, Country Code, Indicator name, Indicator Code and Years from 1960 to 2020. For our analysis we had some empty values for a few countries.



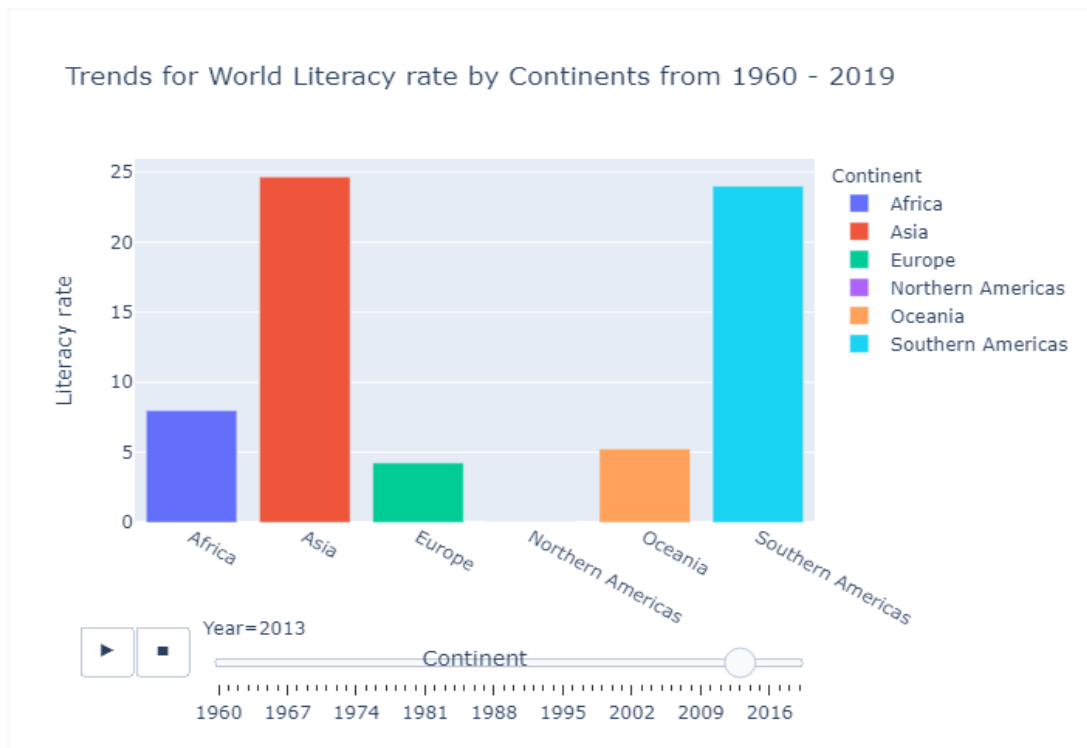
From our analysis we found that there was an overall increase in literacy rates across continents from 1960-2019, which can be testified in the paper by (Roser and Ortiz-ospina, 2018). According to this paper as of 2019, of the world population older than 15 years 86% are literate.



Asia had the highest rate of increase in literacy rates which was expected according to (www.macrotrends.net, n.d.) while continents like North America showed a declining trend which was surprising (Jameson, 2007). This can also be due to non-availability of data present in the dataset for specific countries and continents plus several socio-economic factors like government, economy, societal stability etc.



A great number of studies confirms that higher education of women is associated with lower birth rates as per this paper by (Saurabh, Pandey and Sarkar, 2013). We tried to explore the statistical relationship between birth rates and literacy rates across continents. The findings from the analysis suggested that continued trends in female educational attainment has hastened the decline in birth rates. The negative correlation coefficient from the analysis testifies for this claim. From the results, we see that for Africa though the birth rates have increased, the literacy rates have declined so we can conclude that birth rate bears an inverse relationship with literacy rates. While for Asia it has increased linearly as we can see from the bar while for North America, South America it shows a static and fluctuating trend.



Since the burden of child-birth and mostly also of the upbringing of children is borne by women, it is not surprising that birth rates tend to be high where women have a lower social status and few opportunities outside the household. It is only when greater importance is given to the interests of women that this changes. Women's better education and women's increased employment opportunities both changed the role of women in society and increased the social status of women in society. With more outside-options to having a large number of children, women were open to taking advantage of these options and consequently we observe the overall decline for continents and the inverse relationship of literacy to birth rates in our analysis.

### Technologies for the Future:

If we had more time we would be interested in refining our project further utilizing the following technologies and techniques:



- We would like to use more functions in NumPy and potentially utilize R for further statistical analysis. It would be interesting to create a regression analysis and predictive model that could predict future trends related to birth rates and the other factors. Further we could implement Shiny where we could create interactive web interfaces for our models and analysis.
- We would also be interested in learning Tableau and perhaps implementing that for more interactive visualizations.
- Cloud based Databases (such as Amazon AWS) could also be a potential technology to be utilized. This would potentially provide us with the benefits of: security, speed, scalability, and flexibility.
- A Github repository for this project would also be a good implementation that would have helped us to work on coding collaboratively, and ensuring version control for our code.

## Conclusion

Our study on birth rate analysis resulted in some insightful findings. We found out the overall birth rate trend is declining across the world from 1960 to 2019 where we saw Africa had the highest birth rate growth through the years. The overall mortality rate is also reducing from where we found that we are heading to an aging population. The demographics of not having enough people of working age is a very serious problem and one of the major concerns for countries worldwide.

Lower birth rates are good to prevent overpopulation but bad if you consider the increasingly aging population. Declining birth rates is a massive problem for a lot of countries because there are no good solutions being implemented. The aim should be to aim for a balanced demography. The demographics of not having enough people of working age is a very serious problem and one of the major concerns for governments worldwide.

We believe expanding employment opportunities, accelerating income convergence with developed countries, and strengthened ownership of national development processes by governments, offering mental health and social welfare support to women so that the opportunity costs between raising families versus pursuing a career are minimized. We also recommend promoting an equitable distribution of childrearing between men and women, social policies that promote economical and affordable options for childcare, better wages and working conditions for women, affordable education and affordable housing for families are some solutions that will help countries achieve a stable birth rate. As covered in our analysis, the Nordic Countries between the years of 1985 and 1996 managed to maintain a high birth rate with a high rate of female participation in the labour force, and the conditions that led to this balance should be held as a model to follow for the world.

Our analysis confirmed that there is a relationship between decreasing birth rates and rising contraceptive use. The inverse relationship between contraceptive use and birth rates explains why countries with high proportions of women using contraception generally have lower levels

of birth rates. We found that overall birth rates are decreasing as contraceptive prevalence in the world gradually increases.

We further delved into the key metrics of our analysis to compare factors that affect birth rates and tried to establish a relation between them. Since the burden of child-birth and mostly also of the upbringing of children is borne by women, it was not surprising to see that birth rates tend to be high where women have a lower literacy rate and few opportunities outside the household for countries, Africa in our analysis.

Thus we can conclude the overall effect of female labor force participation, literacy rates and contraceptive prevalence are some of the many complex factors affecting birth rates. Because birth rates are a complex topic, further investigation is recommended to gain a comprehensive understanding of all the factors involved.

## References

### Datasets

data.worldbank.org. (n.d.). Birth rate, crude (per 1,000 people) | Data. [online] Available at: <https://data.worldbank.org/indicator/SP.DYN.CBRT.IN?end=2019&start=2019&view=map&year=1960> [Accessed 3 Dec. 2021].

<https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf>

World Bank. (2020, September 29). Ratio of female to male labor force participation rate (%) (national estimate). Retrieved from <https://data.worldbank.org/indicator/SL.TLF.CACT.FM.NE.ZS>

The World Bank a, ,[Death rate, crude \(per 1,000 people\)](#). Available: Worldbank.org. (2019). Death rate, crude (per 1,000 people) | Data. [online] Available at: <https://data.worldbank.org/indicator/SP.DYN.CDRT.IN?end=2019&start=2019&view=map&year=1960> [Accessed 3 Dec. 2021].

World Bank. (2021, October 28). Contraceptive prevalence, any methods (% of women ages 15-49). Retrieved from <https://data.worldbank.org/indicator/SP.DYN.CONU.ZS>

World Bank. (2021, October 28). School enrollment, secondary, female (% gross). Retrieved from <https://data.worldbank.org/indicator/SE.SEC.ENRR.FE>

The World Bank. "Literacy Rate, Adult Total (% of People Ages 15 and Above) | Data." Worldbank.org, 2020, <https://data.worldbank.org/indicator/SE.ADT.LITR.ZS>

kaggle.com. (n.d.). Country Mapping - ISO, Continent, Region. [online] Available at: <https://www.kaggle.com/andradaolteanu/country-mapping-iso-continent-region> [Accessed 3 Dec. 2021].

## Birth Rates

<https://www.ncbi.nlm.nih.gov/books/NBK236385/>

National Research Council (US) Working Group on the Effects of Child Survival and General Health Programs on Mortality, Ewbank, D.C. and Gribble, J.N. (2018). Trends in Mortality and Causes of Death in Africa. [online] Nih.gov.

<https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf>

United Nations (2019) Department of Economic and Social Affairs, Population Division. World population ageing 2019: Highlights (ST/ESA/SER.A/440). pp. 4

## Mortality

Sciencedirect.com. (2017). *Ageing Population - an overview | ScienceDirect Topics*. [online] Available at: <https://www.sciencedirect.com/topics/social-sciences/ageing-population>.

United Nations (2017). *World Population Ageing*. [online] Available at: [https://ifa.ngo/wp-content/uploads/2018/12/WPA2017\\_Highlights.pdf](https://ifa.ngo/wp-content/uploads/2018/12/WPA2017_Highlights.pdf).

## Female Labour Force Participation

Adhikari, R. (2010). Demographic, socio-economic, and cultural factors affecting fertility differentials in Nepal. *BMC Pregnancy and Childbirth*, 10(1). doi:10.1186/1471-2393-10-19

Bloom, D.E., Canning, D., Fink, G. and Finlay, J.E. (2009). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(2), pp.79–101.

Bloom, D., Canning, D., Fink, G., & Finlay, J. (2007). Fertility, Female Labor Force Participation, and the Demographic Dividend. doi:10.3386/w13583

Da Rocha, J.M. and Fuster, L. (2006). Why Are Fertility Rates And Female Employment Ratios Positively Correlated Across O.E.C.D. Countries? *International Economic Review*, 47(4), pp.1187–1222.

Derosé, L. and Stone, L. (2021). *More Work, Fewer Babies: What Does Workism Have to Do with Falling Fertility? More Work, Fewer Babies: What Does Workism Have to Do with Falling Fertility?*.

Lim, L. (2001). *Female Labour Force Participation*. [online] Available at:  
[https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/unpd\\_egm\\_200203\\_backgroundpaper\\_female\\_labour-force\\_participation\\_lim.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/unpd_egm_200203_backgroundpaper_female_labour-force_participation_lim.pdf).

Roser, M. (2014). Fertility Rate. *Our World in Data*. [online] Available at:  
<https://ourworldindata.org/fertility-rate#what-explains-the-change-in-the-number-of-children-women-have> [Accessed 3 Dec. 2021].

## **Contraceptive Prevalence**

[https://www.un.org/en/development/desa/population/publications/pdf/family/World\\_Fertility\\_and\\_Family\\_Planning\\_2020\\_Highlights.pdf](https://www.un.org/en/development/desa/population/publications/pdf/family/World_Fertility_and_Family_Planning_2020_Highlights.pdf)

United Nations (2020) Department of Economic and Social Affairs, Population Division. *World Fertility and Family Planning 2020: Highlights* (ST/ESA/SER.A/440). pp. 2

<https://pubmed.ncbi.nlm.nih.gov/1462719/>

National Library of Medicine (1992) Steep decline in world fertility rates: contraceptive use up sharply. *Soz Praventivmed*, 37(5), pp. 254-5.

[https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesapd\\_report\\_2015\\_trends\\_contraceptive\\_use.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesapd_report_2015_trends_contraceptive_use.pdf)

United Nations (2015) Trends in Contraceptive Use Worldwide 2015. *Economics and Social Affairs*. pp. 7.

<https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10570-w>

Sherwood, J., Lankiewicz, E., Roose-Snyder, B. *et al.* (2021). The role of contraception in preventing HIV-positive births: global estimates and projections. *BMC Public Health*, 21(536)

[https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un\\_2019\\_contraceptiveusebymethod\\_databooklet.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2019_contraceptiveusebymethod_databooklet.pdf)

United Nations (2019) Department of Economic and Social Affairs, Population Division. *Contraceptive Use by Method 2019: Data Booklet* (ST/ESA/SER.A/435). pp. 10-14

<https://blogs.worldbank.org/health/economic-rationale-investing-family-planning-sub-saharan-africa>

Peter Glick and Sebastian Linnemayr (2016) The economic rationale for investing in family planning in Sub-Saharan Africa. *World Bank Blogs*

## **Literacy Rates**

[https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30677-2/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30677-2/fulltext)

Vollset, S.E., Goren, E., Yuan, C.-W., Cao, J., Smith, A.E., Hsiao, T., Bisignano, C., Azhar, G.S., Castro, E., Chalek, J., Dolgert, A.J., Frank, T., Fukutaki, K., Hay, S.I., Lozano, R., Mokdad, A.H., Nandakumar, V., Pierce, M., Pletcher, M. and Robalik, T. (2020). Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study, *The Lancet*

<https://ourworldindata.org/grapher/womens-educational-attainment-vs-fertility>.

Our World in Data. (2010). Women's educational attainment vs. number of children per woman. [online]

<https://www.jstor.org/stable/4411060>

Parikh, K.S. and Gupta, C. (2001). How Effective Is Female Literacy in Reducing Fertility? *Economic and Political Weekly*, [online] 36(35), pp.3391–3398.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4649870/>

Saurabh, S., Pandey, D. and Sarkar, S. (2013). Female literacy rate is a better predictor of birth rate and infant mortality rate in India. *Journal of Family Medicine and Primary Care*, 2(4), p.349

[https://www.theguardian.com/global-development-professionals-network/2017/mar/15/why-have-four-children-when-you-could-have-seven-contraception-niger?CMP=share\\_btn\\_tw](https://www.theguardian.com/global-development-professionals-network/2017/mar/15/why-have-four-children-when-you-could-have-seven-contraception-niger?CMP=share_btn_tw)

The Guardian. (2017). Why have four children when you could have seven? Family planning in Niger.

## **Birth Rate Research**

Adhikari, R. (2010). Demographic, socio-economic, and cultural factors affecting fertility differentials in Nepal. *BMC Pregnancy and Childbirth*, 10(1).

Bloom, D.E., Canning, D., Fink, G. and Finlay, J.E. (2009). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(2), pp.79–101.

Bloom, D., Canning, D., Fink, G. and Finlay, J. (2007). Fertility, Female Labor Force Participation, and the Demographic Dividend.

Kakarla, S. (2019). How to Clean Machine Learning Datasets Using Pandas. [online] ActiveState. Available at: <https://www.activestate.com/blog/how-to-clean-machine-learning-datasets-using-pandas/> [Accessed 3 Dec. 2021].

kaggle.com. (n.d.). Country Mapping - ISO, Continent, Region. [online] Available at: <https://www.kaggle.com/andradaolteanu/country-mapping-iso-continent-region>.

Olteanu, A. (2019, December 15). Country Mapping - ISO, Continent, Region. Retrieved from <https://www.kaggle.com/andradaolteanu/country-mapping-iso-continent-region>

Suhaib Peerzada (2018). 8 Ways to Clean Data Using Data Cleaning Techniques. [online] Digital Vidya. Available at: <https://www.digitalvidya.com/blog/data-cleaning-techniques/>.