

REPORT ON WORLD BANK HUMAN CAPITAL PROJECT

The Human Capital Project was started in several countries around 2017. For the purpose of this report, I have referred to the World Bank website for data. (<https://data.worldbank.org/country/GH>)

I looked at datasets from various countries in the developing world. Most had similar patterns in available features. So eventually chose Ghana since it is a developing economy in Africa.

The policy was implemented in 2017. There is little data from 2018 to 2021 to determine the impact of policy on development indicators. The results of such large-scale policy implications can be seen only after a few years. However, some conclusions can still be drawn using techniques from Data Science to determine policy impact on indicators.

The dataset/technical challenges:

The raw dataset obtained from the world bank website for the country Ghana had to go through pre-processing, data cleaning, and ETL (extract, transform, and loading) process, as it had many missing values and redundant and insignificant features.

The dataset consists of records from the year 1960 -2021 and according to the world bank case study, Ghana's human capital project was established in the year 2018. There are almost 1444 columns or features also specifically called development indicators such as social, economic, and environmental which decide the overall growth.

All features are not very relevant hence some features could also be clubbed together and be considered an indicator or predictors.

There were missing values and characters like double dots (..) which had to be replaced also while plotting some values on the line plot there were issues due to data types

I transposed the dataset to get it in the desired row-column format.

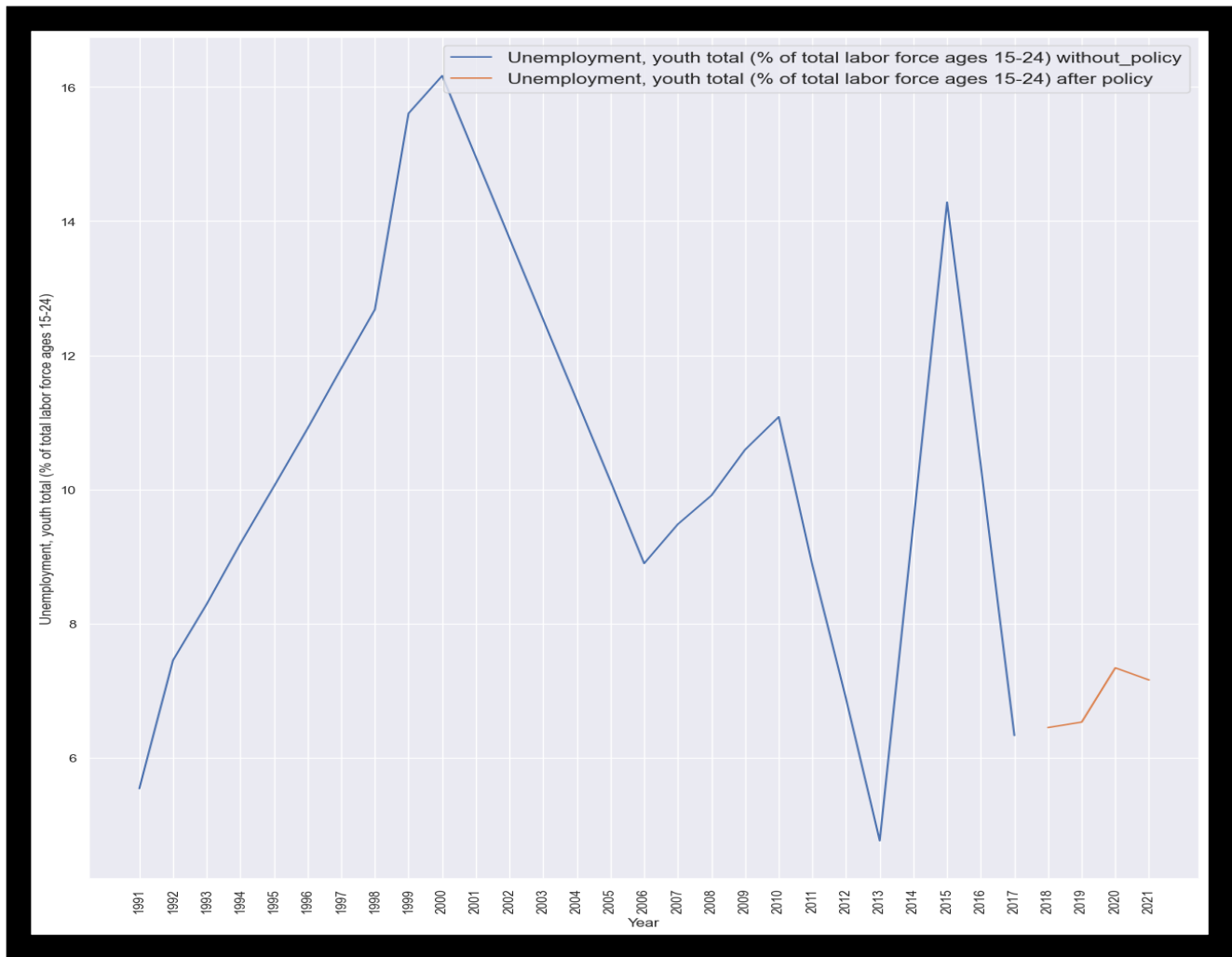
The dataset has 1444 indicators, and out of these several features had only sparse data of a few years. Hence, a few indicators were chosen depending on data available before and after the policy implementation date of 2018 and their general importance for a country's development.

The following part of this report consists of mainly plots of these key selected indicators plotted against the year. With these plots, we can visualize the evolution of these indicators from the starting year, 1960 to 2021, which was the last year for which data was available. After the plots, some insights were derived for a few key indicators using data science techniques like linear regression and correlation analysis. Some comments have been made on the efficacy of the Human Capital project on these indicators with this analysis. In the end, the report is concluded with some policy recommendations.

***** Graphical Visualisation of the impact of various world development indicators before and after implementation of human capital policy *****

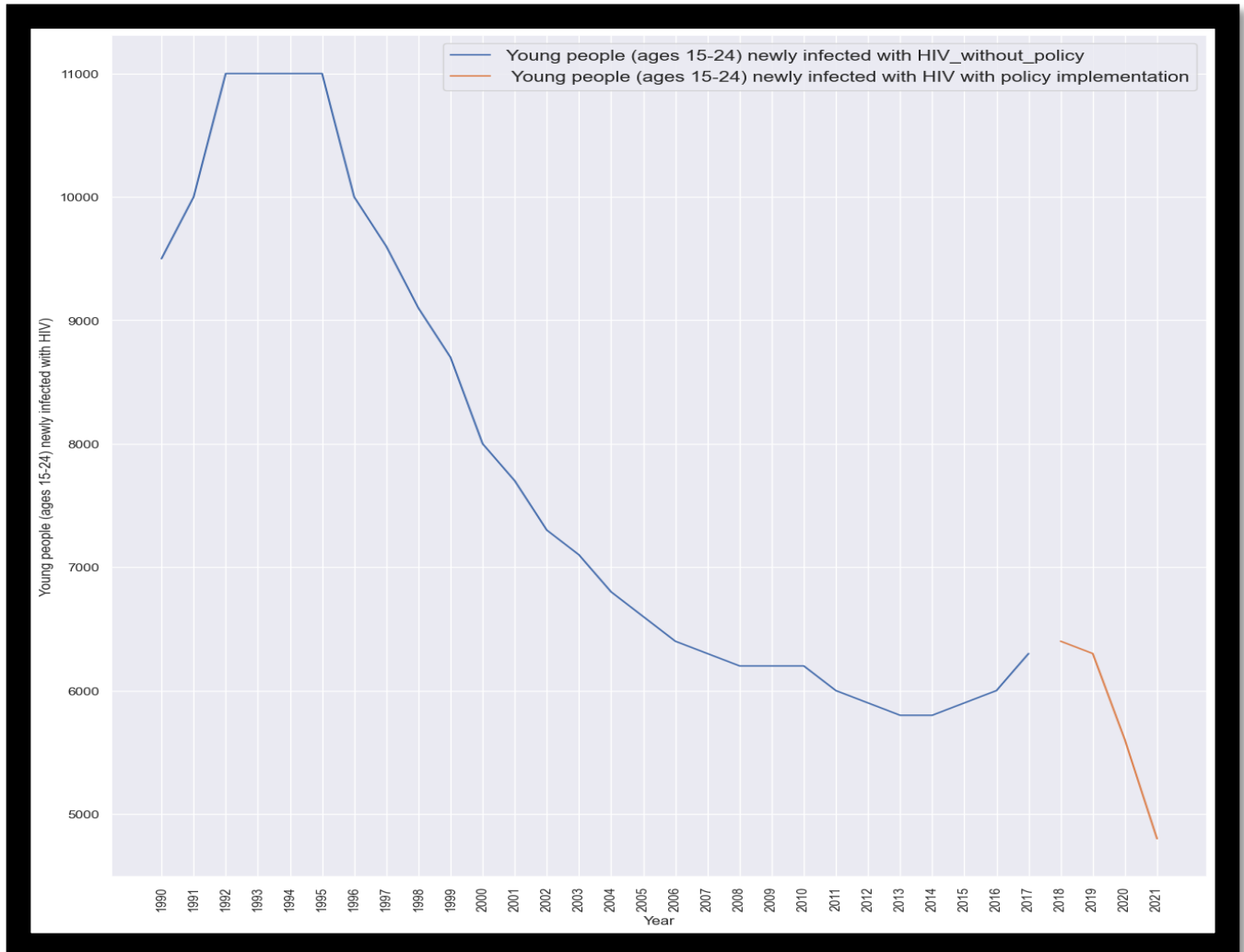
1. Unemployment, youth total (% of the total labor force)

There is significant variability in unemployment rates, as can be seen from the graph below. The reasons for this variation are not accounted for in this report. From the data from 2018 onwards, though, it looks like unemployment is a bit steady and on average much lower than pre-2018. Hence, we can say that the Human Capital project has contributed positively to this indicator.



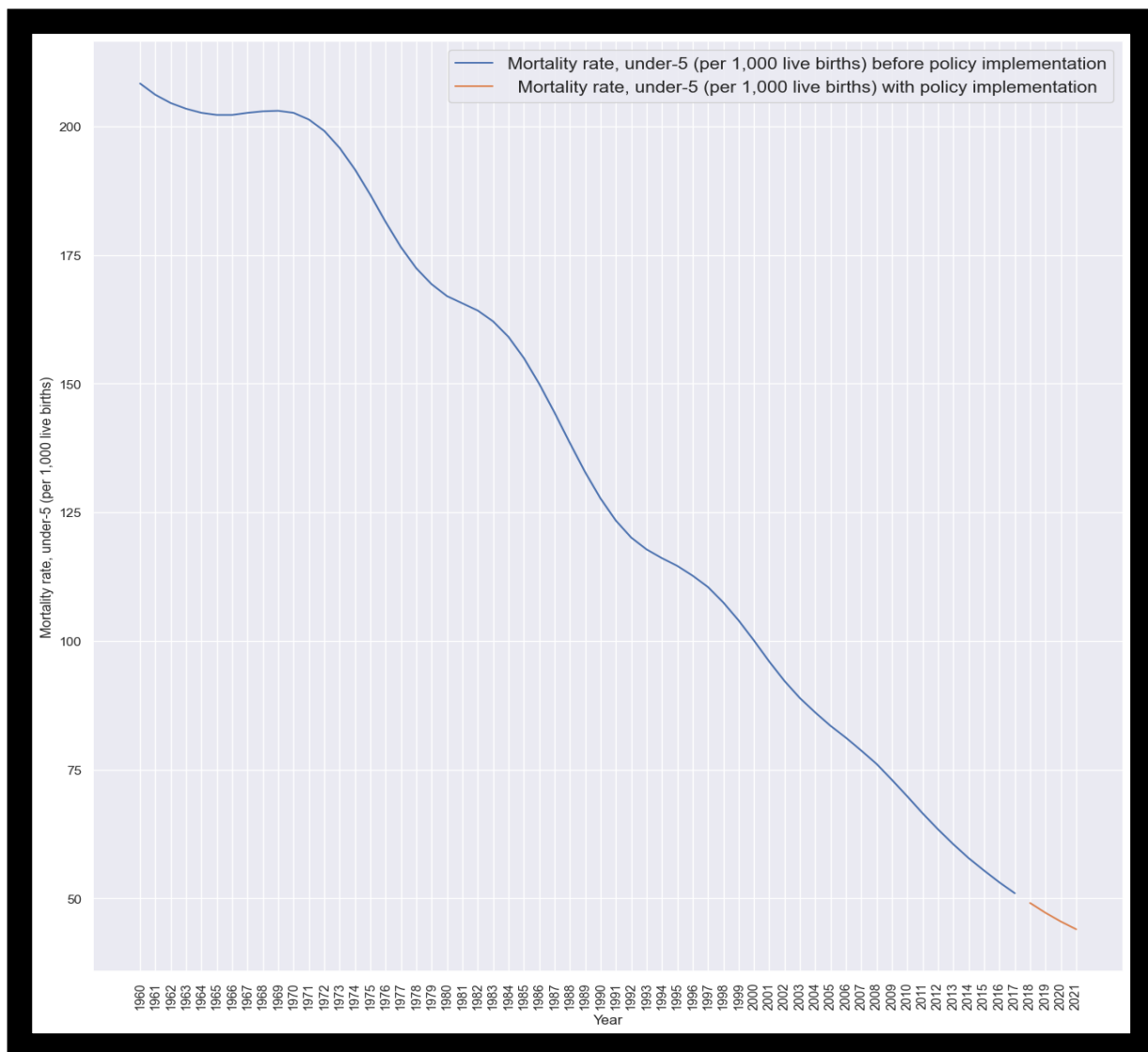
2. Young people (ages 15-24) newly infected with HIV

The below graph demonstrates that young people getting affected with HIV have drastically decreased which indicates that proper preventive measures have been adopted for spreading awareness and providing medications after policy implementation.



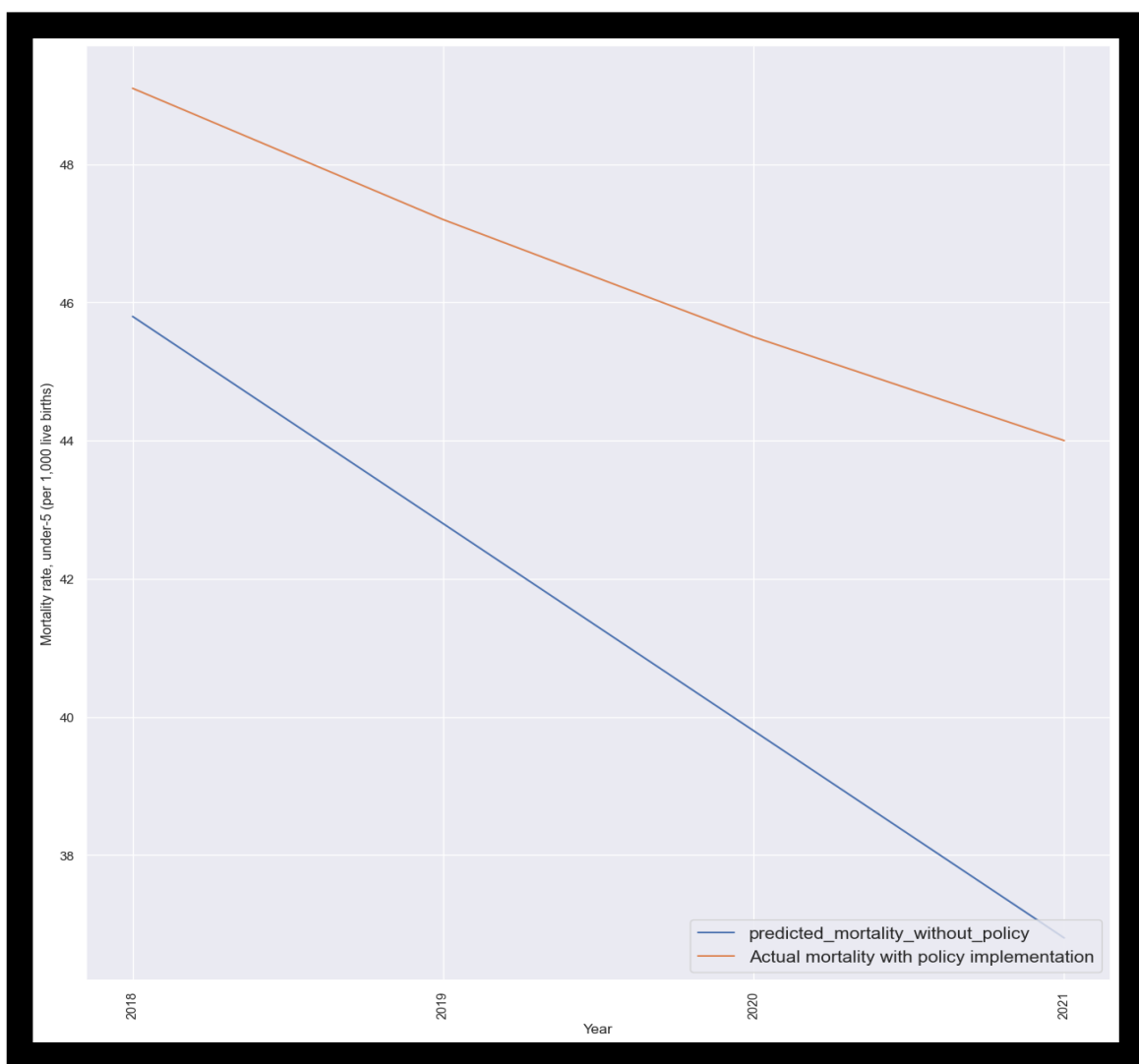
3. The trend of Mortality rate, under-5 (per 1,000 live births) before and after implementation of world human capital implementation with linear regression model implementation.

The mortality rate under 5 is an important indicator of child health, which ultimately impacts several other indicators. From the graph of the mortality rate, we see that it has been continuously decreasing since the early years, although at different rates at different times.



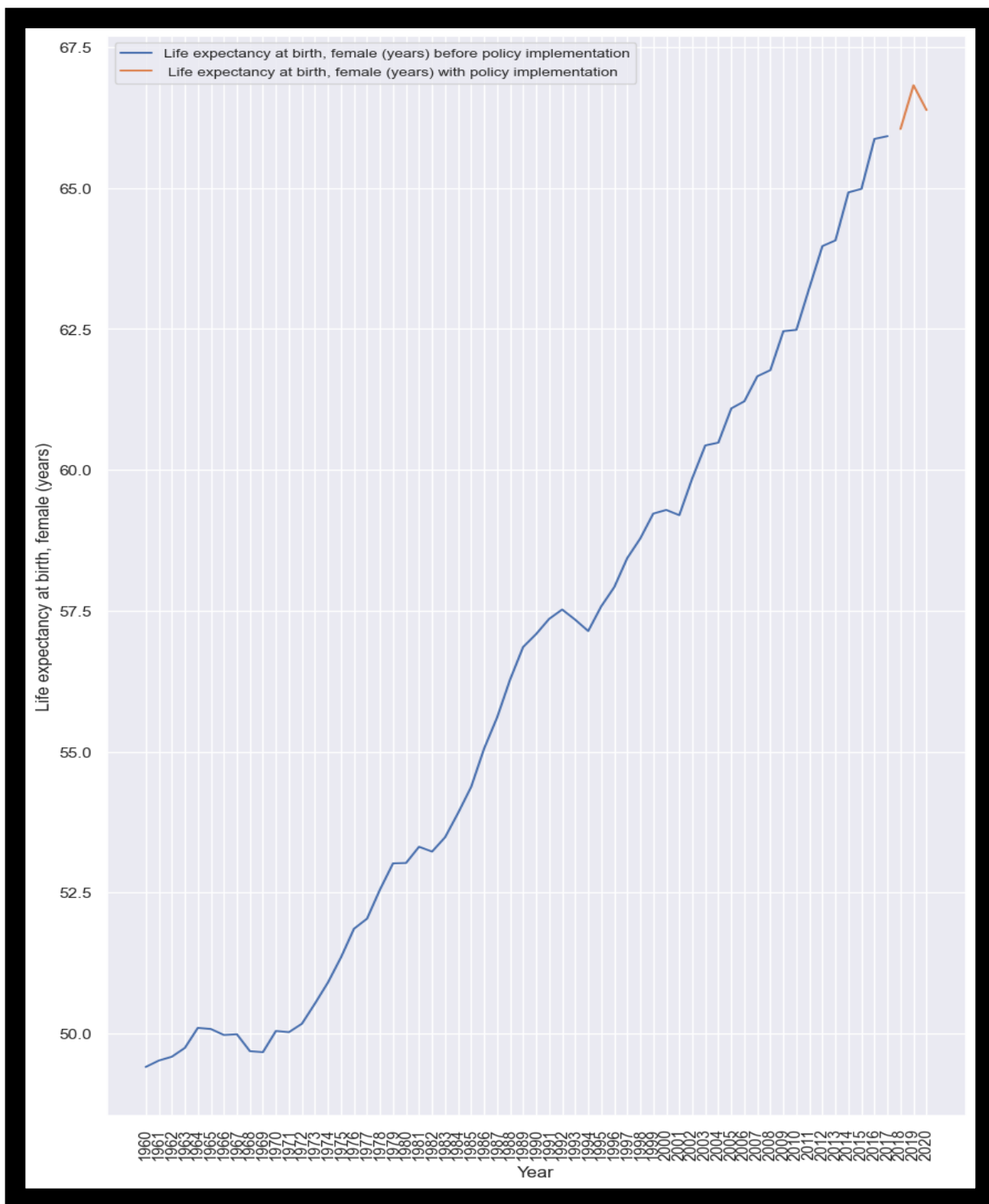
Hence, to predict the impact of human capital policy implementations on the mortality rate post-2017, I have implemented a linear regression (machine learning) algorithm to determine the trend of a mortality rate decrease, if there was no human capital policy implementation in 2017.

As you can observe from the graph of linear regression, the mortality rates after policy implementation (Red line) are slightly higher than what the linear regression trend predicts (blue line). This indicates that policy implementation might have made the mortality rate worse, as compared to the trend that would exist if there was no policy implementation. Hence, more efforts are needed in this area to reduce the mortality rates under 5.



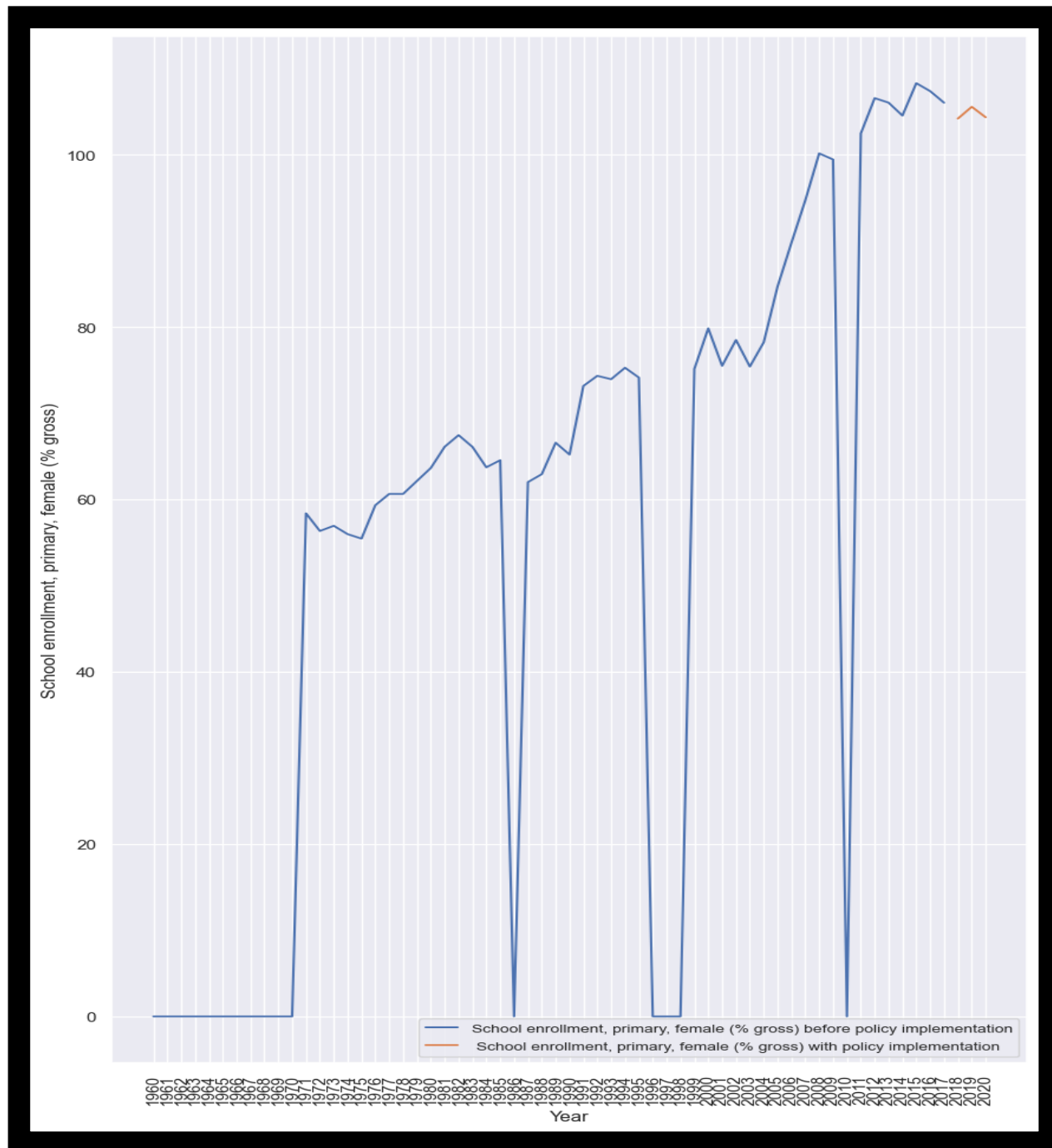
4. Life expectancy at birth, female (years)

Female life expectancy is following a positive trend and the policy seems to influence the trend positively.



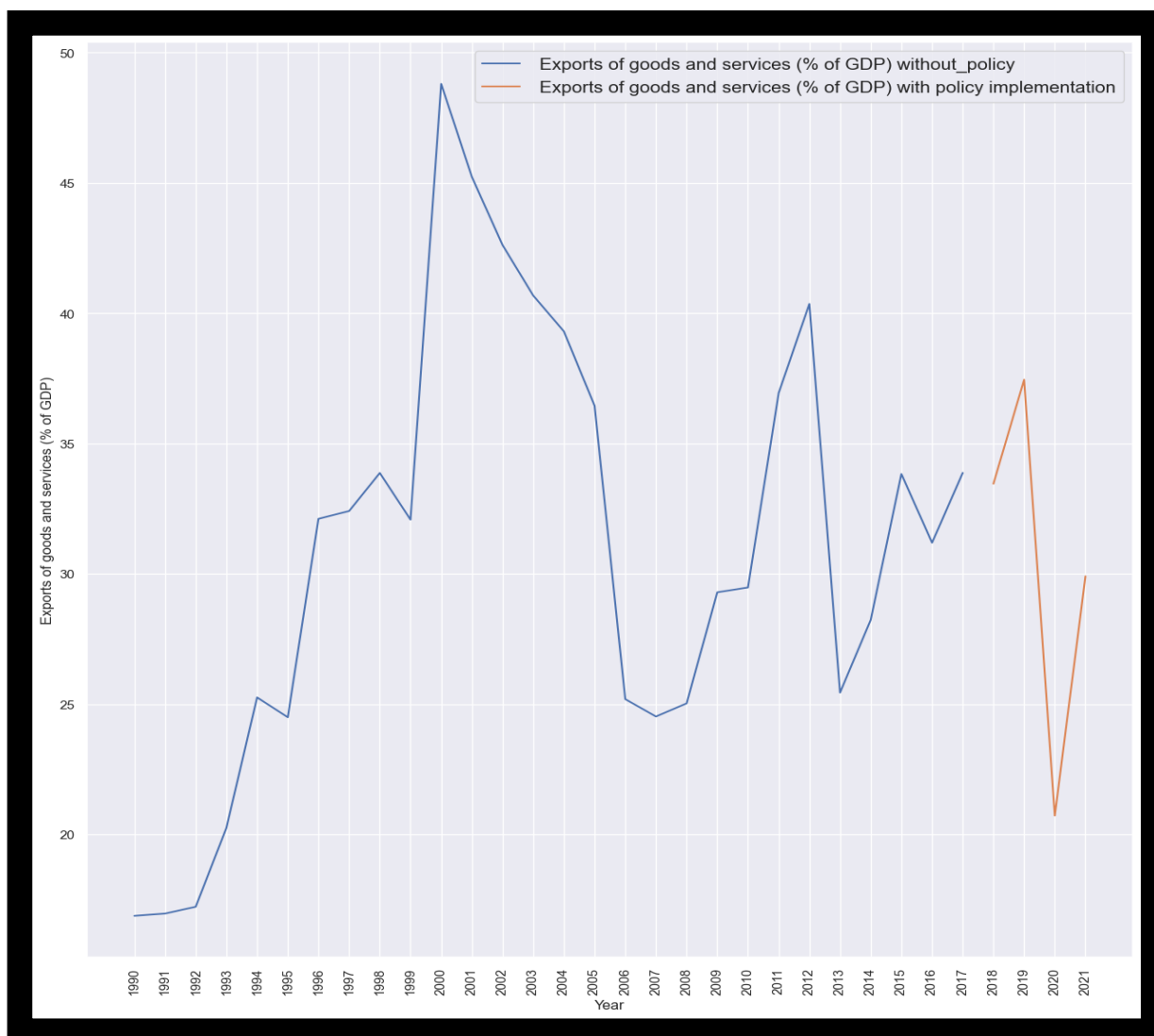
5. School enrollment, primary, female (% gross)

The school enrolment rate seems to be the same as before the policy implementation (some gaps in data exist, which correspond to the dips in the graph)



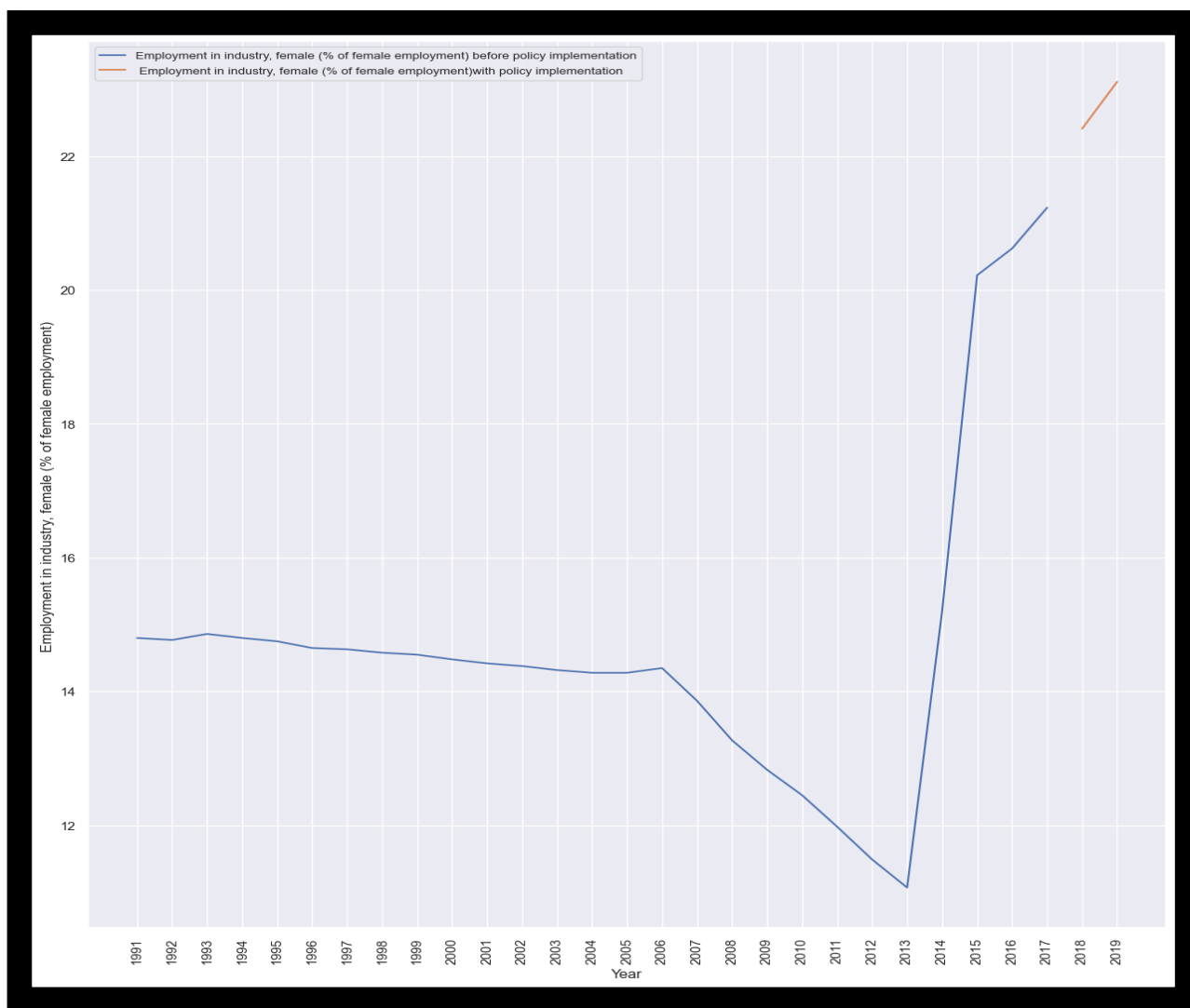
6. Exports of goods and services (% of GDP)

The export of goods and services as a percentage of GDP is an important indicator of the overall economy of a developing country. This indicator shows fluctuations. After policy implementation, there was the Covid pandemic as well in 2019-2020. Hence policy impact is not very clear.



7. Employment in industry, female (% of female employment)

This indicator shows a very positive trend post-policy implementation, although data for the final two years is lacking. Still, there is about a 2% growth in female employment, which is very important for economic and social growth.



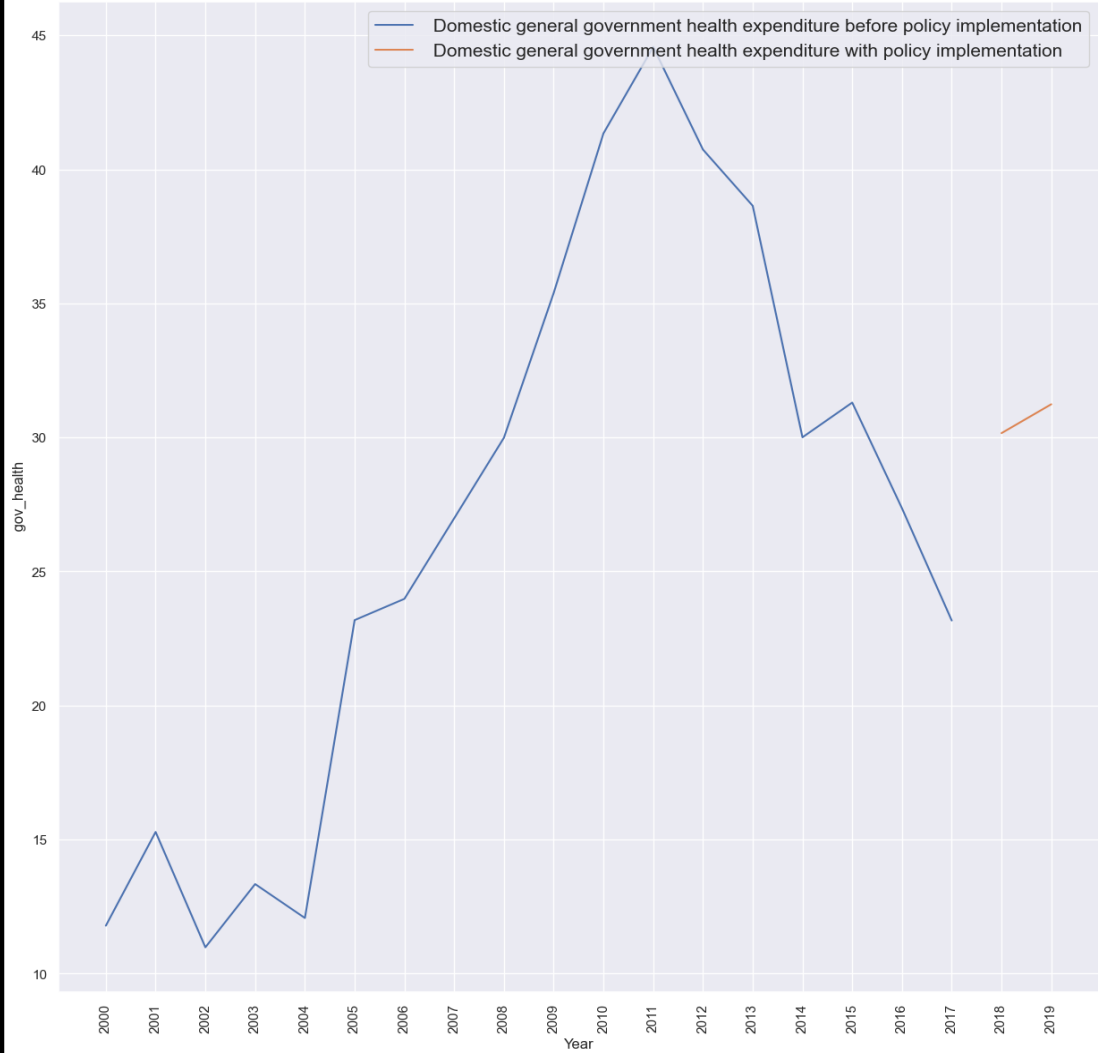
8. Domestic general government health expenditure before and after policy implementation

Health is a very important measure of development. There were many indicators related to health, but they lacked data. Hence, four indicators were chosen and their mean/average over all the years was computed and clubbed as a composite new indicator, which is called **Domestic general government health expenditure (% of current health expenditure)**. The following indicators from the dataset were used for averaging:

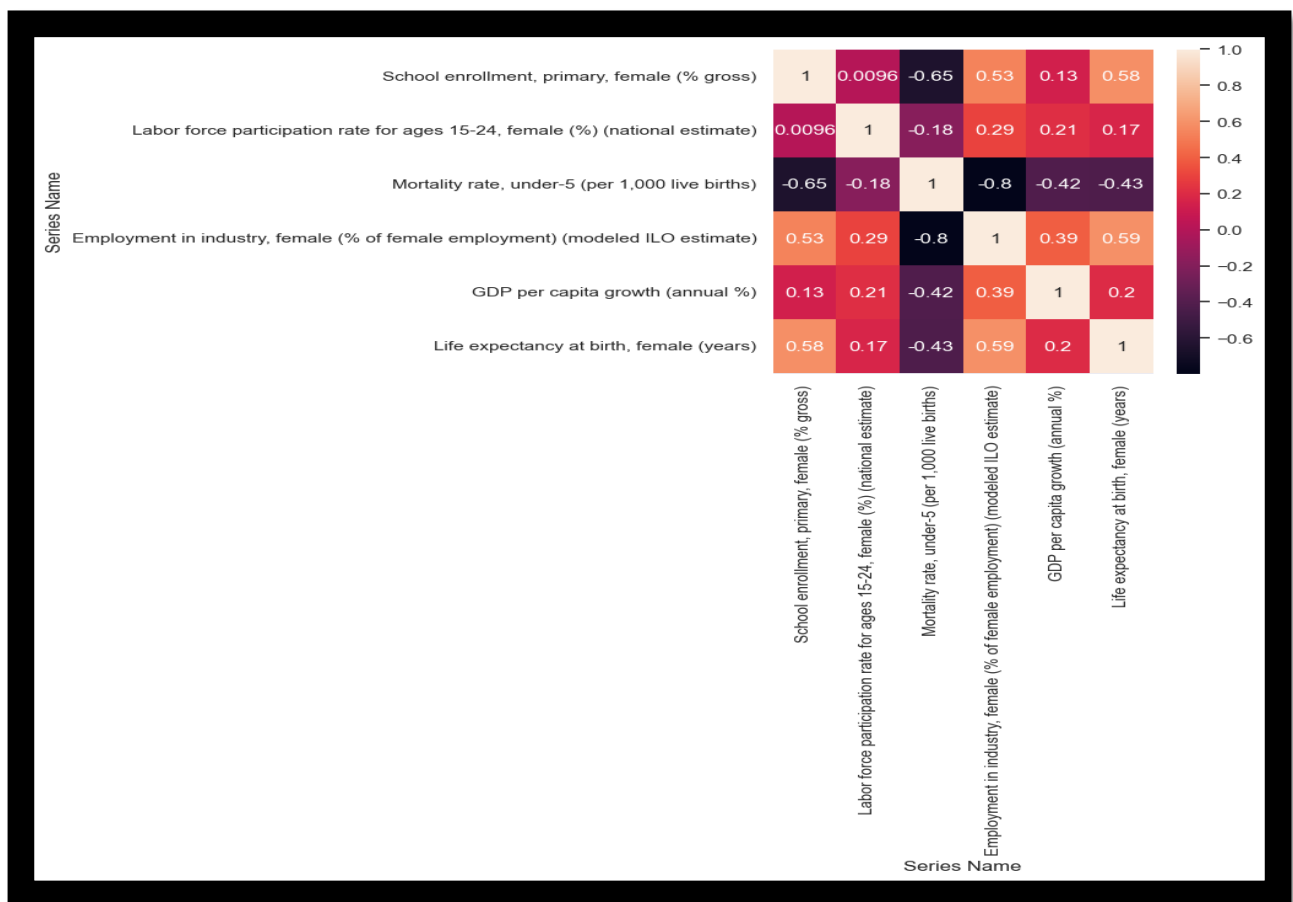
- i. Domestic general government health expenditure (% of GDP)
- ii. Domestic general government health expenditure (% of general government expenditure)
- iii. Domestic general government health expenditure per capita (current US\$)
- iv. Domestic general government health expenditure per capita, PPP (current international \$)

The below graph of this composite indicator tells an interesting story. Domestic health expenditure increases till 2011 and then there is a steep decrease, which is detrimental to the health of the population. However, after 2017, the declining trend is arrested, and expenditure has increased by more than 5% compared to 2017. This shows that the Human Capital policy has led to a significant increase in healthcare spending. This is an encouraging result.

This analysis shows that creating composite features from related similar indicators can help provide significant insights even when data is missing.



CORRELATION ANALYSIS OF CHOSEN WORLD DEVELOPMENT INDICATORS:



A correlation matrix shows the relationship among indicators. While it is well known that correlation does not imply causation, still it is quite useful to visualize the pairwise relationship between statistical variables, in this case, indicators. Please see the above matrix for numbers. Some aspects are commented on below.

The most important numbers from the matrix are the largest and smallest values. The largest value implies a positive correlation, and the most negative value implies a negative correlation. Consider the indicator of primary school enrolment, employment in the industry for females, and life expectancy at birth for females. The largest value of 0.59 is found in row 4, column 6. This shows that there is a very high correlation between life expectancy at birth for females and subsequent employment in industries. This is easy to guess, but a strong positive number here helps the government and other agencies to improve conditions for girl children in general. Similar positive trends are seen for school enrolment for girls and employment in industries/life expectancy.

Similarly, the mortality rate is negatively correlated with all other indicators, especially against employment in the industry.

RECOMMENDATIONS AND SUGGESTIONS

There does not exist much data after 2018 to conclusively determine if the Human Capital Project has positively impacted indicators. Hence the conclusions too must evolve with more data collection. However, the following are some recommendations to improve development indicators.

- i. Formal training and upskilling youth especially from rural areas relevant to industry needs, all such facilities should be cost-effective and affordable.
- ii. There should be better coordination among government, private agencies, and educational institutions to bridge the gap between the on-field reality and classroom teaching.
- iii. Some sectors still are lacking funding and cooperation, and investment from the government, the professionals working in such industries should be well-paid so that there is the motivation among youngsters to take up those jobs.

The correlation analysis clearly indicates that reducing mortality rate has a very positive effect on important indicators like employment, economy, GDP growth etc. The linear regression result however, has shown that mortality rate is slightly worse after human capital policy implementation. Some suggestions to reduce the mortality rate further, for which government should invest in the following areas:

- i. Child health care, and women empowerment which includes self-employment.
- ii. Campaigning and spreading awareness about hygienic practices, which also includes providing basic health care amenities.
- iii. Expenditure on the provision of education to rural women.
- iv. Employ strategies to engage women in self-employment so that they can also contribute equally to family income.
- v. Economically backward and underprivileged people should get access to all basic facilities at an affordable price which includes, health, education, employment etc.
- vi. Control measures on population growth and explosion as it directly impacts country's overall GDP.