The Consequences of Life Expectancy and Infant Mortality on the GDP per Capita of Countries

1 Introduction

Over the last 80 years, financial and health indices have been used to determine the standing of nations in the world today. Subsequently, these indices have come to the forefront in the myriad reports written about the relationship between health and wealth and whether there is a correlation and causation effect as regards both terms. Does health cause wealth? or is it the other way around. This research report aims to answer the former question by exploring whether attributes such as Infant mortality rate and life expectancy have effects on the Gross Domestic Product (GDP) per capita of countries around the world. Before the analysis can commence properly, key terms have to be expanded upon.

According to the Office for National Statistics (ONS), the GDP is the cumulative economic worth of a country (Office of National Statistics, 2016). It is a value used to know the standing of the economy of a country. GDP per capita on the other hand is the average contribution of each individual to the economy of a particular country. Introduced in the 1930s before the start of the second world war by Simon Kuznets, GDP has, today, become the most popular index for determining the economic standing of countries (Talberth et al., 2006).

The first health indicator that will be analysed is the ”infant mortality rate” indicator. The infant mortality rate shows the rate at which infants (children under the age of one (1) year) die, for every 1000 births, in a country (Nuffield Trust, 2021).

The final health indicator that will be analysed in this report is the life expectancy indicator. In line with the definition from the ONS, the life expectancy of nations is the average number of years a person is expected to live based on the economic, social, and geographic standing of their country (UK Government, 2017).

The data used for this analysis was obtained from the world bank (https://data.worldbank.org/), with the aid of an R studio package known as ”WDI”. The data is cross-sectional data of 178 countries from the year 2019. 2019 was chosen as that was the latest year with a minimal amount of missing data.

In subsequent sections, previous work on analyses with regards to all indicators will be touched upon. Afterward, the indicators will be explored in-depth and regression analysis will be performed to determine the relationships between GDP per capita and the health indicators.

2 Literature Review

Tons of research has been done in determining the relationship between health and wealth across nations of the world. First, in the May 2016 version of the Quarterly Journal of Economics, an article by Cesarini et al postulated, from data obtained from Swedish lottery players, that affluent countries have a better chance of being healthier than non-affluent ones (Cesarini et al., 2016).

Another research paper, which focuses on an analysis of GDP and life expectancy in Spain and Italy, by Felice et al, titled "GDP and life expectancy in Italy and Spain over the long run: A time-series approach" was published in the Demographic Research Journal in 2016. The article investigated the connection between GDP and life expectancy in Italy and Spain, with the aid of time series analysis, and concluded that a significant improvement in life expectancy leads to an equally significant improvement in the GDP of the countries (Felice et al., 2016). This and the previous article lead to a similar conclusion that health leads to wealth, at least concerning the countries mentioned above.

From the previous research explored so far, the consensus has been that countries with a high life expectancy tend to have a high value for GDP and in turn, GDP per capita. However, concerning infant mortality, the opposite seems to be the case. Erdogan et al concluded in their research article, that high infant mortality rates have a significant negative impact on the GDP of countries around the world based on their regression analysis performed (Erdogan et al., 2013).

Further research from Ensor et al buttressed the previous point in their research article on the effect of health on the recession, published in 2010. In the paper titled, "The impact of economic recession on maternal and infant mortality: lessons from history", after analysing the data ranging from 1936 to 2005, they concluded that there is a negative association between the infant mortality rate and GDP (Ensor et al., 2010). The time-series data used to obtain their findings were culled from an analysis of fourteen (14) upper-middle and high-income countries around the world (North American and European countries mostly).

Most of these papers mentioned above were analysed with the aid of linear regression or multiple linear regression; while the articles from Ensor et al and

Felice et al were analysed with the aid of time series techniques that will not be covered in this paper. These studies showed a positive correlation between life expectancy and GDP per capita and on the other hand, a negative correlation between the infant mortality rate and the GDP per capita.

There have also been numerous research on how these two health indicators interact with one another. A paper titled "Infant mortality and life expectancy in China" by Xu et al published in 2014 revealed the negative correlation between the infant mortality rate and life expectancy in China (Xu et al., 2014).

The next section will explore the data that will be used for the analysis and the techniques that will be employed in analysing it.

3 Data Exploration

As stated in section 1, the data used for this analysis was obtained from the World Bank online repository and was accessed with the aid of R studio. This analysis was carried out with the aid of the R programming language. Although, before the main analysis (linear regression) begins, insights can be gained from exploring the data .

The three(3) variables involved in the analysis, which are cross-sectional data from 178 countries for the year 2019, will be explored in detail. First of all GDP per capita, whose unit is in U.S. dollars ($), has a maximum value of $113,218.70 (Luxembourg) and a minimum value of $228.21 (Burundi). The first health indicator, infant mortality rate per 1000 births has a maximum value of 82.4 (Sierra Leone) and a minimum value of 1.6 (Iceland). The final variable, life expectancy, has a maximum value of 84.36 years (Japan) and a minimum value of 53.28 years (Democratic Republic of Congo).

Further insight can be viewed in figure 1 where the box plot shows that most countries in the world have GDP per capita values between $0 and $20000. Outliers are also present in the data for the GDP values which adds to the positive skew of the variable.

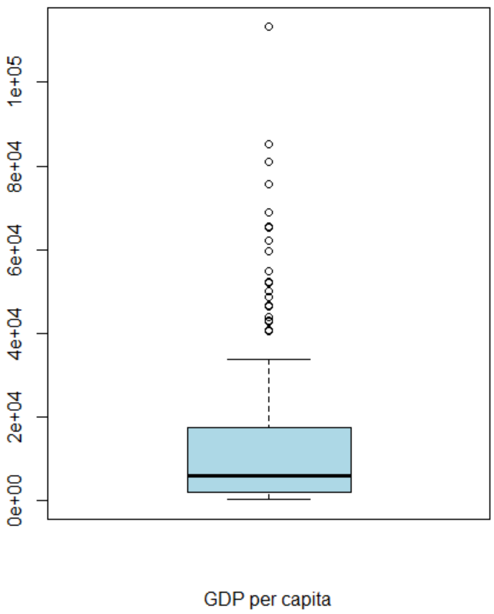

Figure 1: Box plot of the GDP per Capita.

Figure 2 shows that majority of the life expectancy values of countries in this analysis range from 67 to 77 with the values being slightly negatively skewed. Also, no outliers are present.

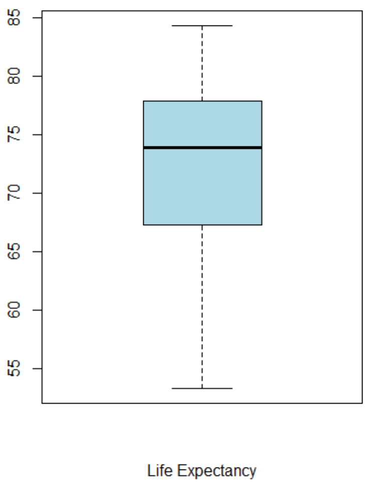

Figure 2: Box plot of the Life Expectancy of Countries.

In figure 3, the values of the infant mortality rate of countries lie between 10 and 35 with only 3 outliers. These outliers are Sierra Leone (82.4), Central African Republic (79.1), and Somalia (74.8). A common factor about these countries is that they are all African countries. To add to this, the skew of the infant mortality rate is positive.

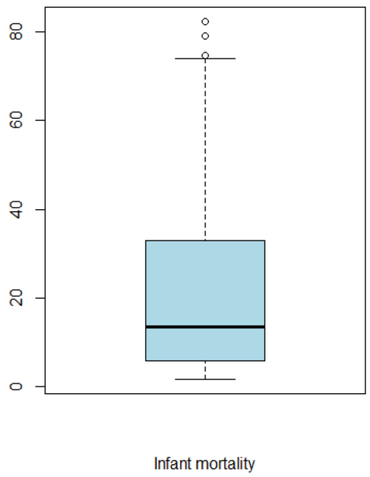

Figure 3: Box plot of the Infant Mortality Rate per 1000 births.

Figure 4 confirms the findings in figure 1 with most countries having GDPs per capita between $0 and $20000. Figure 4 also shows that the GDPs of countries are exponentially distributed.

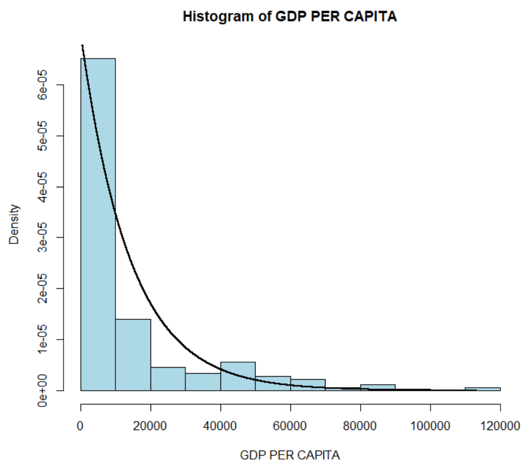

Figure 4: Histogram of GDP per Capita.

The histogram of the life expectancy of countries in figure 5 give an approximately normal distribution with a negative skew as confirmed from figure 2.

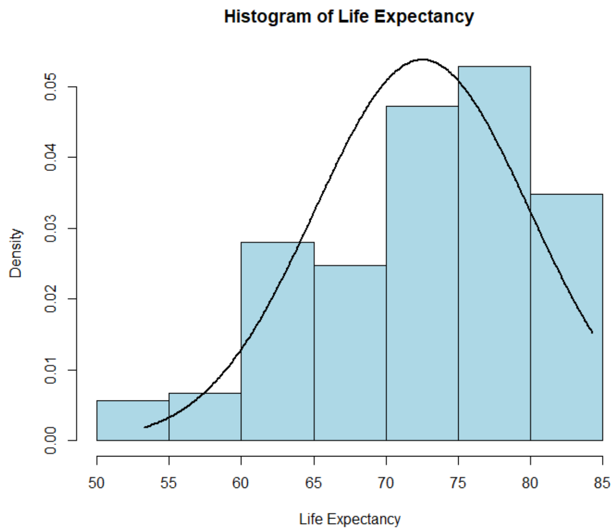

Figure 5: Histogram of Life Expectancy.

Similar to the GDP per capita in figure 4, the histogram of the infant mortality rate, in figure 6, shows the variable is also exponentially distributed.

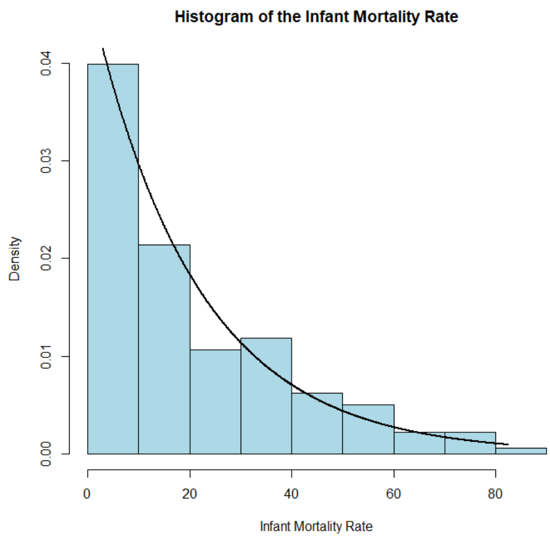

Figure 6: Histogram of Infant Mortality Rate.

Table 1 contains the summary statistics of all the variables for all 178 observations (n). The next section will go further into the relationship between the three variables and the regression analysis. Table 1 also shows that the values of the three variables are not commensurate; as GDP per capita has a higher scale than the other two.

Table 1: Descriptive Statistics of all Variables with n = 178

S/N	**Variable Name**	**Short Name**	**Unit**	**Mean (2 d.p.)**	**Median**	**Standard Deviation (2 d.p.)**
1	GDP per Capita	gdpPercap	U.S. Dollars ($)	14118.80	6025.50	19560.97
2	Life Expectancy	lifeExp	Years	72.51	73.93	7.42
3	Infant Mortality Rate	infRate	Deaths per 1000 births	20.91	13.45	19.46

The variable names and their short names in table 1 will be used interchangeably for the remainder of this paper.

4 Results and Analysis

In this section, the regression analysis will take place to decipher the relationship between GDP per capita, life expectancy, and infant mortality rate. For this analysis, the dependent variable (output) will be the GDP per capita while the independent variables (inputs) will be the infant mortality rate and the life expectancy. However, before the regression analysis commences, other methods can be used to view the relationships between the three variables.

Correlation is a statistical tool that can be used to examine the ways variables depend on one another. Correlation shows how variables change concerning one another; the change could be positive or negative (Schober et al., 2018). The common way of measuring correlation is with the Pearson correlation coefficient (r). From figure 7, GDP per capita and life expectancy are moderately correlated, while being slightly negatively correlated to the infant mortality rate. This means that countries with high values for GDP per capita tend to have slightly high values for life expectancy and moderately low values for the infant mortality rate. Life expectancy has a very strong negative correlation to the infant mortality rate. Therefore, countries with high life expectancy values tend to have low values for their infant mortality rate and vice versa.

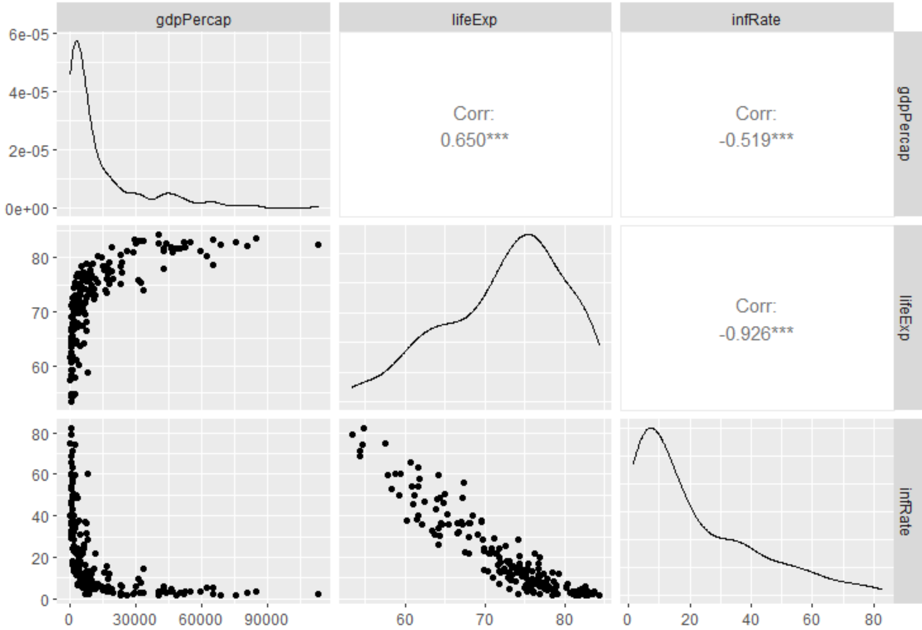

Figure 7: Correlation or Scatter Matrix of all Variables

4.1 Regression Analysis

The main aim of this report is to determine the relationship between GDP per capita, Life expectancy, and infant mortality rate. Regression analysis is a statistical technique that reveals the linear relationship between variables. GDP per capita will be the dependent variable for this analysis while Life expectancy and infant mortality rate will be the independent variables. Three regression models will be created; they are:

- Log-transformed GDP per Capita and Life Expectancy.
- Log-transformed GDP per Capita and Infant Mortality rate.
- Log-transformed GDP per Capita, Life Expectancy, and Infant Mortality Rate.

The reason for the log transformation is due to the positive skew of the GDP per capita variable as seen in figures 1 and 4. The log-transformed GDP per capita normalises the variable and makes it ready for modelling and interpretation. The linear regression models will be created with the aid of the generalised linear model (glm) function on R studio.

Model 1: After creating the model, the regression equation is:

$$log(gdpPercap) = -2.846899 + (0.158876 * lifeExp).$$

The model can be interpreted as a 1 unit change in the life expectancy of a country leads to a 0.158876 increase in the value of log-transformed GDP per capita; which in turn leads to a 1.1722($e^{0.158876}$) increase in the actual value of GDP per capita. This shows that an increase in the life expectancy of a country leads to an increase in the GDP per capita of a country. The r-squared value will be used to evaluate the model. The r-squared value shows if the independent variable(s) can explain the variation in the dependent variable (Figueiredo et al., 2011). The r-squared value of model 1 is about 70.09%. This means that the independent variable (lifeExp) in this model can tell us about 70% of the variation in the dependent variable (log(gdpPercap)). This is further confirmed in figure 8 with the plots of the independent variable against the dependent variable.

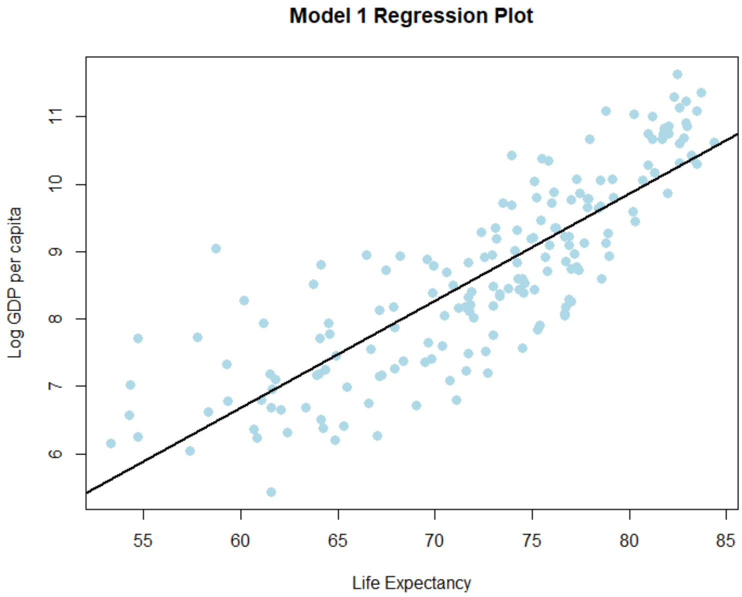

Figure 8: Regression plot of Model 1

Model 2: After creating the model, the linear regression equation is:

$$log(gdpPercap) = 9.851472 - (0.056348 * infRate).$$

The model can be interpreted as a 1 unit change in the infant mortality rate per 1000 births in a country leads to a 0.056348 decrease in the value of log-transformed GDP per capita; which also leads to a 1.0580 ($e^{0.056348}$)

decrease in the actual GDP per capita. This shows that an increase in the infant mortality rate of a country leads to a decrease in the GDP per capita of a country. The r-squared value of this model is about 60.73%. This is worse when compared to model 1 as the dependent variable (infRate) just tells us about 61% of the variation of the dependent variable. Figure 9 describes the negative relationship between the two variables.

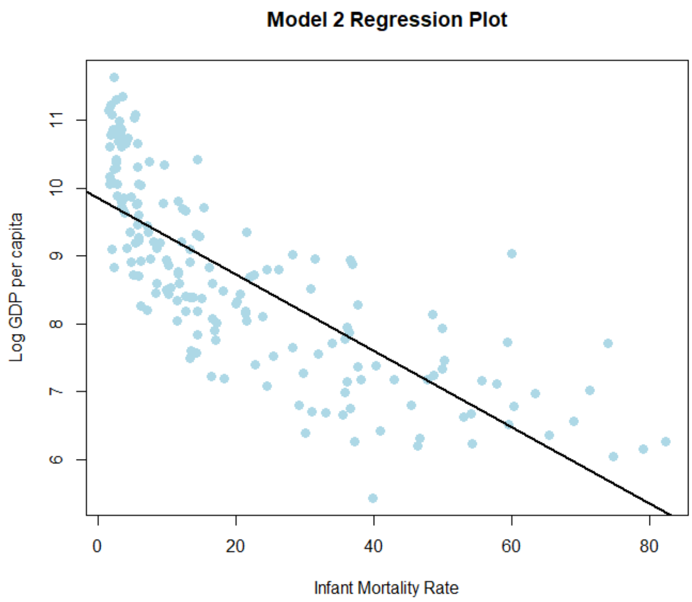

Figure 9: Regression plot of Model 2

Model 3: The final model is a multiple linear regression model with two (2) inputs and the equation is:

$$log(gdpPercap) = -2.419 + (0.1536 * lifeExp) - (0.002 * infRate).$$

The model can be interpreted as a 1 unit change in the life expectancy of countries, when the other input variable is kept at zero (0), leading to a 0.1536 increase in the value of log-transformed GDP per capita. Also, when the life expectancy is kept at zero (0), a 1 unit change in the infant mortality rate leads to a 0.002 decrease in the log-transformed GDP per capita. This confirms what was inferred from the previous models; an increase in life expectancy leads to a similar increase in GDP per capita; however, an increase in the infant mortality rate leads to a slight decrease in the GDP per capita. Also, from the model, life expectancy has more of an effect on GDP per capita than the infant mortality rate. The r-squared value of this model is about 70.10%. This tells us that both independent variables (lifeExp and infRate) show about 70% of the variation in the dependent variable. Compared to models 1 and 2, model 3 slightly provides more information than both models 1 and 2.

The next section will determine whether the coefficients of each model are statistically significant.

4.2 Statistical Significance of the Coefficients

Statistical significance is the probability of making a type I error (rejecting a null hypothesis when it is true). This gives validity to the statistical readings we obtained in the previous sections. The significance level of the test will be at the 5% level. Due to this, table 2 will show the 95% confidence intervals of the coefficients of all variables in the three models. Coefficients that have 0 between the lower and upper bounds of their confidence intervals will be deemed as statistically

Table 2: 95% Confidence Intervals and P-Values of all Model Coefficients.

MODEL 1					
S/N	**Independent Variable**	**Lower Bound**	**Upper Bound**	**P-value**	**Statistically Significant (5% Level)**
1	Life Expectancy	0.144	0.174	0.000	Yes
MODEL 2					
S/N	**Independent Variable**	**Lower Bound**	**Upper Bound**	**P-value**	**Statistically Significant (5% Level)**
1	Infant Mortality Rate	-0.063	-0.050	0.000	Yes
MODEL 3					
S/N	**Independent Variable**	**Lower Bound**	**Upper Bound**	**P-value**	**Statistically Significant (5% Level)**
1	Life Expectancy	0.113	0.194	0.000	Yes
2	Infant Mortality Rate	-0.018	0.013	0.784	No

Table 2 shows that all the coefficients in the first two models are statistically significant on the 5% level. This means that both life expectancy and infant mortality rate have a statistical impact on the output (log-transformed GDP per capita). However, in the third model (Model 3), life expectancy is statistically significant, while the infant mortality rate is not statistically significant at the 5%

level. Thus, in model 3, the infant mortality rate probably has no effect, with life expectancy included, on the log-transformed GDP per capita. This may be due to the strong negative correlation between life expectancy and the infant mortality rate as shown in figure 7.

5 Conclusion and Limitations

5.1 Overview of Findings

The outcomes from the regression models show that the two health variables (life expectancy and infant mortality rate) have different relationships with the log-transformed GDP per capita and the actual GDP per capita in turn. Based on the regression analysis, countries with a high life expectancy will most likely lead to a high value for the log-transformed GDP per capita; while countries with high infant mortality rates will have a low value for the log-transformed GDP per capita. However, the effect of life expectancy tends to be greater than that of the infant mortality rate.

5.2 Limitations

There are always limitations in every statistical analysis performed and this report is no different. A major limitation that may have affected the results of the regression analysis is the correlation between the input variables (model 3 in particular). Correlated inputs mask the effects of one of the coefficients leading to it being statistically insignificant.

Another limitation is the possibility of a Type I error occurring during the test for the significance of each coefficient in the three models. This gives leeway to the chance of false positives happening and could mean that coefficients deemed not statistically significant could be statistically significant.

5.3 Conclusion

This report aimed to answer the question "Does health cause wealth?" by finding the relationship between the GDP per capita of countries, life expectancy, and infant mortality per 1000 births. The data exploration and the regression analysis showed that GDP per capita has a negative relationship with the infant mortality rate and a positive relationship with life expectancy. The regression analysis also revealed that positive effects of life expectancy have a more sizeable impact on the GDP per capita of countries than the negative effects on the infant mortality rate. Therefore, based on the relationships established by the analysis, it is safe to come to the conclusion that health, based on the health indicators analysed, does indeed lead to more wealth in countries.

6 References

Cesarini, D., Lindqvist, E., Ostling, R., & Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on swedish lottery players. *Quartely Journal of Economics*, *131*(2), 687–738.

Ensor, T., Cooper, S., Davidson, L., Fitzmaurice, A., & Graham, W. J. (2010). The impact of economic recession on maternal and infant mortality: Lessons from history.

Erdogan, E., Ener, M., & Arica, F. (2013). The strategic role of infant mortality in the process of economic growth: An application for high income oecd countries. *Procedia Social and Behavioral Sciences*, *99*, 19–25.

Felice, E., Andreu, J. P., & D'Ippoliti, C. (2016). Gdp and life expectancy in italy and spain over the long run: A time-series approach. *Demographic Research*, *99*(28), 813–866.

Figueiredo, D., Júnior, S., & Rocha, E. (2011). What is r2 all about? *Leviathan-Cadernos de Pesquisa Polútica*, *3*, 60–68. https://doi.org/10.11606/issn.2237-4485.lev.2011.132282

Nuffield Trust. (2021). *Stillbirths and neonatal and infant mortality*. https://www.nuffieldtrust.org.uk/resource/infant-and-neonatal-mortality

Office of National Statistics. (2016). *What is gdp?* https://www.ons.gov.uk/economy/grossdomesticproductgdp/articles/whatisgdp/2016-11-21

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, *126*(5), 1763–1768.

Talberth, D. J., Cobb, C., & Slattery, N. (2006). *The genuine progress indicator 2006, a tool for sustainable development*. Redifining Progress.

UK Government. (2017). *Research and analysis, chapter 1: Life expectancy and healthy life expectancy*. https://www.gov.uk/government/publications/health-profile-for-england/chapter-1-life-expectancy-and-healthy-life-expectancy

Xu, Y., Zhang, W., Yang, R., Zou, C., & Zhao, Z. (2014). Infant mortality and life expectancy in china. *Medical Science Monitor*, *20*, 379–385.