Analysis for life expectancy

Group 8

Lachlan Thomas Moody

Joyce Lee

Cuiping Wei

Dang Thanh Nguyen

Thursday, June 11 2020

1 Introduction

Health and economic are considered as important indicators for a country's development. Pardi, Nawi, and Salleh (2016) states that many countries have adjusted their national objectives toward refining what can be done with the evolvement of national economic. This is closely linked to the establishment of Sustainable Development which embodies the vision of "Meeting the needs of the present without compromising the ability of future generations to meet their own needs" (Venkataraman 2009). Although health and economic seem to be two unrelated fields of studies, they are inseparable as they exist in a two-way relationship when it comes to achieving the concept sustainable development. The ideal behind sustainability embraces the nature of longevity as one of the parameters, mentioned by Patten and Costanza (1997). Therefore, life expectancy will be represented as longevity in this report to investigate in both health and economic perspectives.

For health aspects, our team has two main research questions, based on the difference in life expectancy between genders and the progression on life expectancy at birth and at the age of 60. For economic aspects, our team analyze health expenditure and GDP per capita in order to answer these questions: Whether high GDP per capita and health expenditure lead to higher life expectancy? What is the relations between health expenditure and life expectancy?

2 Data source

The data for Life expectancy and healthy life expectancy(LE and HALE) by country(Organization and others 2018) is collected by World Health Organization(WHO), which contains LE and HALE data from 2000 to 2016.In this report, year, sex, country, LE and HALE at birth or at age 60(years) variables were mainly used for analysis.

The data for Current health expenditure (CHE) per capita in US$ by country(Organization and others 2020) is collected by WHO, which contains CHE data from 2000 to 2017.In this report, year, country and CHE variables from 2000 to 2016 were mainly used for analysis.

The data for GDP per capita (current US$) by country(Bank 2018b) is from The World Bank(TWB), which contains CHE data from 1960 to 2018. In this report, year, country and GDP variables were mainly used for analysis.

The data for Population Total by country(Bank 2018a) is from TWB, which contains CHE data from 1960 to 2018. In this report, year, country and population variables were mainly used for analysis.

The data for continent in Section 4.3 is read by using URL(Dbouquin 2016) directly, which contains conntry and continent.

3 Methodology

3.1 Linear regression model

The linear regression model is mainly used for exploring relationships between data in this report, the formula can be represented by:

$$\hat{y} = \beta_0 + \beta_1 x$$

In this report, we used some linear regression models to study the differences between sex for LE and HALE, the relationship bettwen HALE and CHE.

4 Results

4.1 Differences between men and women

The disparity in the life expectancy between men and women was first recognised in the 1920's (Luy and Minagawa 2014). This section will examine this difference in detail.
The first step was to tidy the life expectancy data into a format R. The original variables were condensed into Country, Year, Type (expectancy measure), Gender and Expectancy. The first 10 resulting rows are displayed in Table 4.1.

Table 4.1: Life expectancy data

Country	Year	Type	Gender	Expectancy
Afghanistan	2016	Life expectancy at birth (years)	Both	62.7
Afghanistan	2016	Life expectancy at birth (years)	Male	61.0
Afghanistan	2016	Life expectancy at birth (years)	Female	64.5
Afghanistan	2016	Life expectancy at age 60 (years)	Both	16.3
Afghanistan	2016	Life expectancy at age 60 (years)	Male	15.5
Afghanistan	2016	Life expectancy at age 60 (years)	Female	17.1
Afghanistan	2016	Healthy life expectancy (HALE) at birth (years)	Both	53.0
Afghanistan	2016	Healthy life expectancy (HALE) at birth (years)	Male	52.1
Afghanistan	2016	Healthy life expectancy (HALE) at birth (years)	Female	54.1
Afghanistan	2016	Healthy life expectancy (HALE) at age 60 (years)	Both	11.3

An average was then calculated across the recorded gender types for life expectancy at birth which is shown in Table 4.2. It appears that the average life expectancy for both genders is slightly over 69 years with females having the higher value of 71.6 years and males almost 5 years lower at 66.9 years.

Table 4.2: Average life expectancy across genders

Gender	Average Life Expectancy

Gender	Average Life Expectancy
Both	69.2
Female	71.6
Male	66.9

The yearly averages were then visualised using the boxplot below in Figure 4.1 with gender on the x-axis and the average life expectancy on the y-axis. This further illustrates that females on average have a higher life expectancy across all years in the data set, having a higher median and interquartile range than men with only the male maximum and female minimum values overlapping.

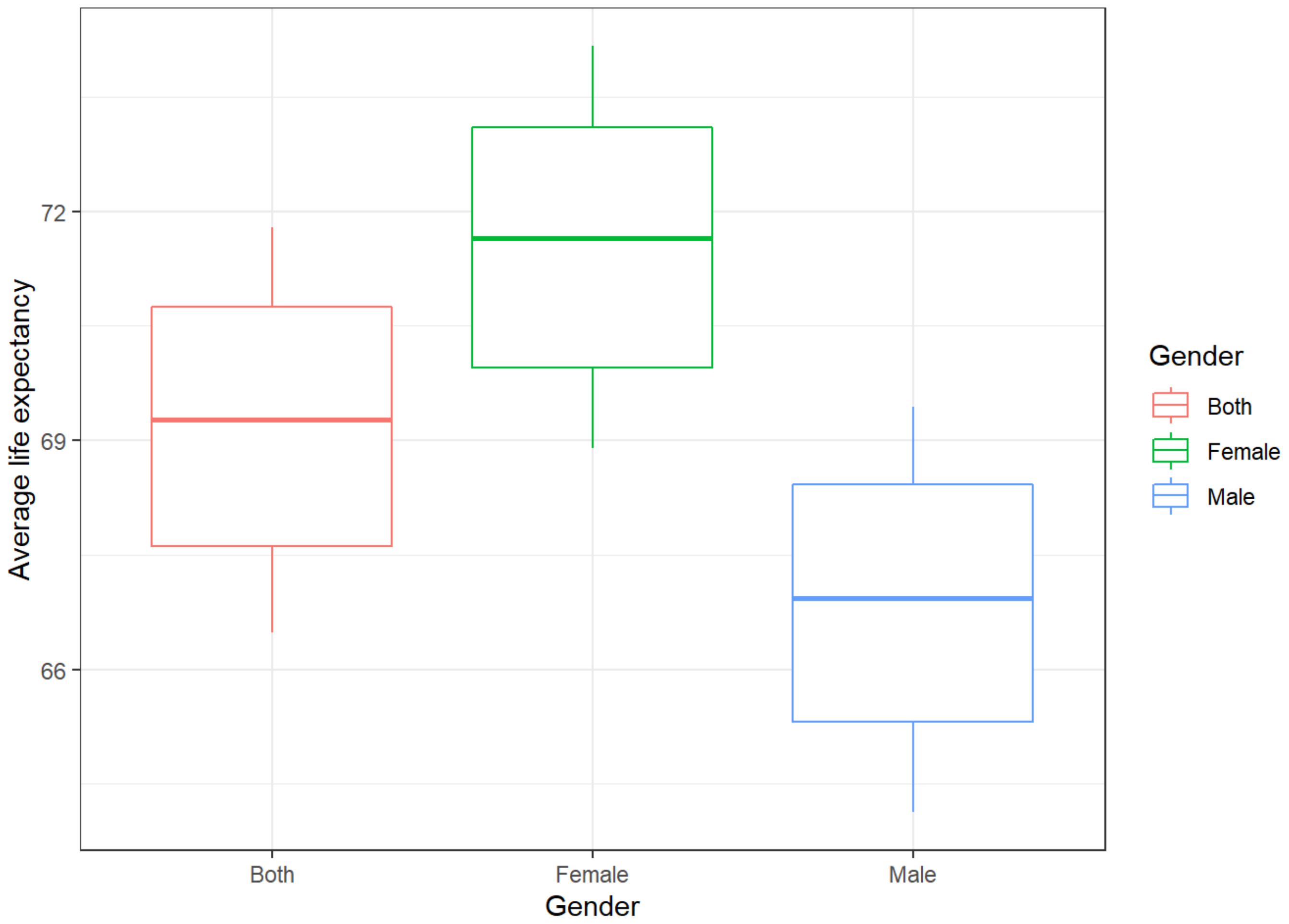

Figure 4.1: Comparison of life expectancy across genders per year

A line graph was then produced with year on the x-axis and the average life expectancy on the y-axis to illustrate this trend. As can be seen in Figure 4.2, across all years recorded in the data set, females have had a higher life expectancy than males by approximately the same amount.

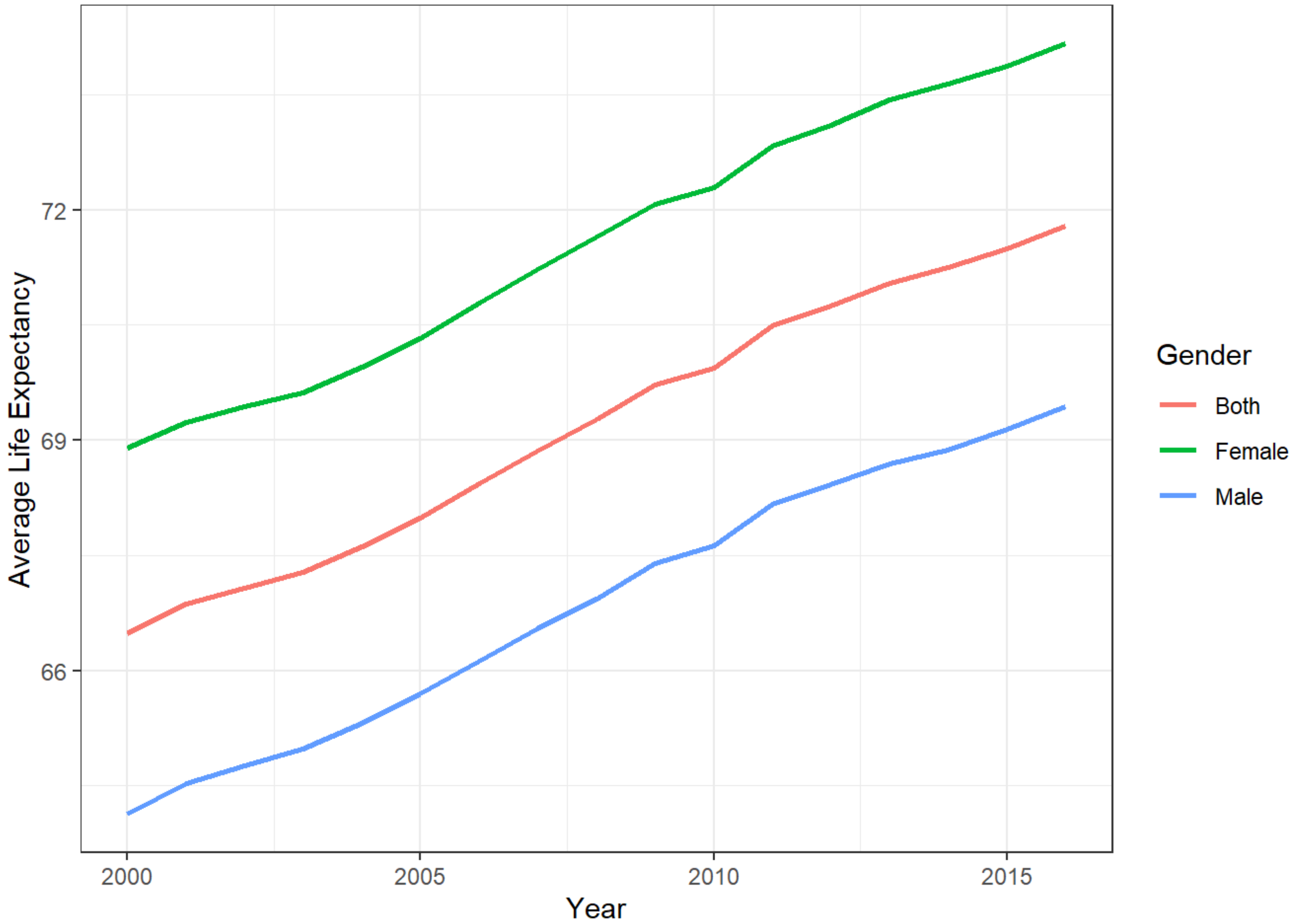

Figure 4.2: Change in life expectancy by year for each gender

A linear model was then produced to see if this trend was due to gender or year. The outputs are shown in Table 4.3.

Table 4.3: Model estimates for life expectancy based on gender and year

Model	term	estimate	r.squared	adj.r.squared
Gender	(Intercept)	71.562	0.652	0.641
Gender	GenderMale	-4.692	0.652	0.641
Year	(Intercept)	66.428	0.346	0.325
Year	Year	0.349	0.346	0.325
Both	(Intercept)	68.774	0.998	0.998
Both	GenderMale	-4.692	0.998	0.998
Both	Year	0.349	0.998	0.998

Several insights are provided from this table. Firstly, the GenderMale term for the Gender model suggests that males on average live 4.7 years less than the female average of 71.6. Secondly, the Year term for the Year model shows that for each after 2000, people on average are living 0.35 years longer. Finally the adjusted.r.squared values describes the explanatory power of these models. The results have been plotted in Figure 4.3 with the model name on the x-axis and the adjusted r squared on the y-axis.

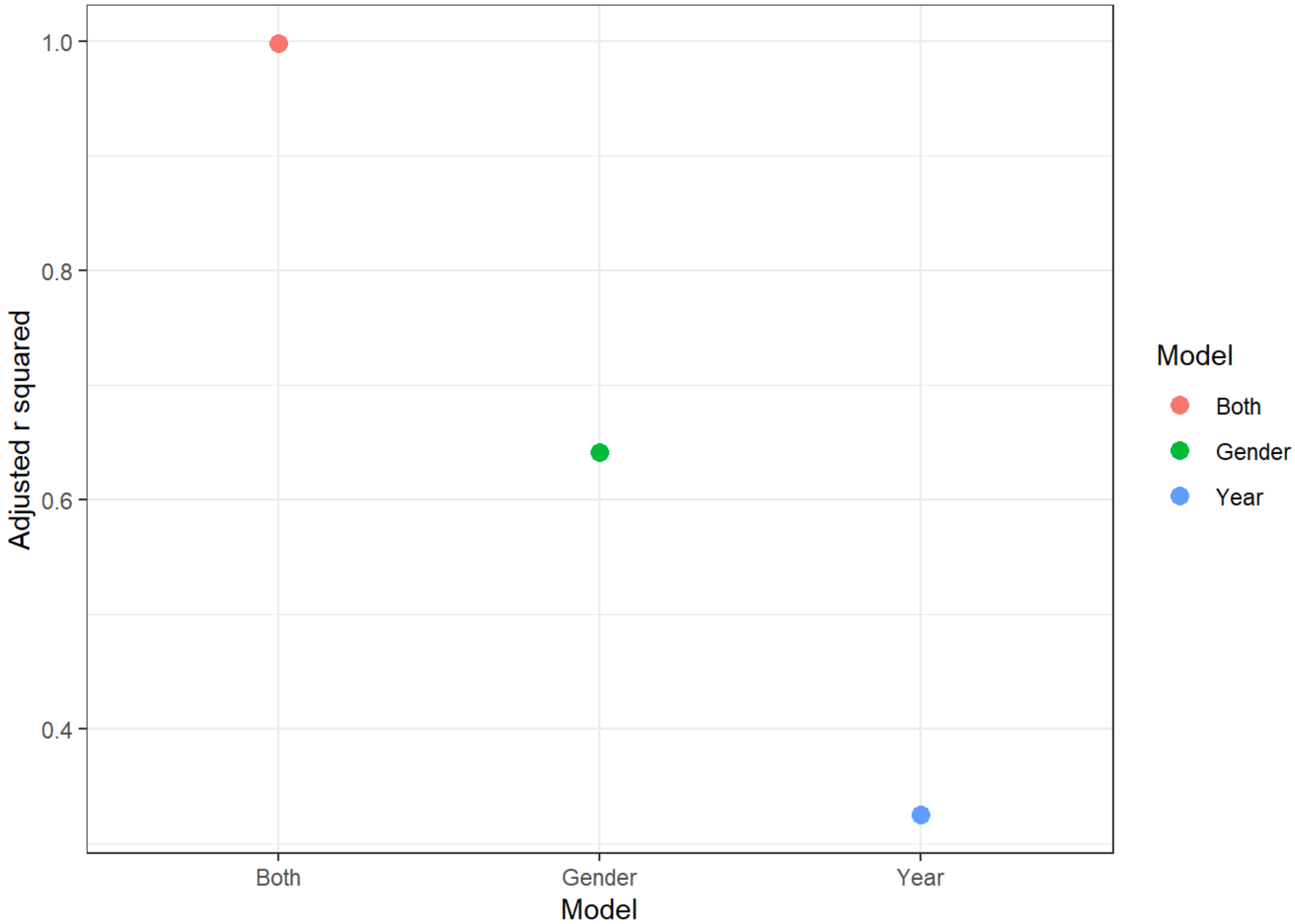

Figure 4.3: Comparison of model fits

This shows that the gender model is superior to the year mode, explaining over 64% of the variation in the life expectancy value compared to under 33% for the year variable. Interestingly, almost the entire variation in life expectancy can be explained by including both variables.

4.2 Life expectancy at birth vs at age 60 years

Life expectancy is a measure of population longevity which indicates how long a person is expected to live (Tosato et al. 2007). It can be measured with different levels, Rabbi (2013) raises the point that life expectancy at birth would be higher than at any particular age. In this section, life expectancy at birth and at the age of 60 are examined to depict the changes within these indicators across the world from 2000 to 2016.

Figure 4.4 presents the overall distribution for life expectancy at birth (years) and life expectancy at age 60 (years) from 2000 to 2016. On the x-axis, Types consist of life expectancy at birth (years), which is positioned as the left boxplot in each sub-graph, and life expectancy at age 60 (years), being the one on the right. The y-axis provides the scale of life expectancy, measured in years. From 2000 to 2016, there is a slight upward-lifting trend for both types of boxplots, implying that people at birth and at the age of 60 are expected to live longer with the progress in time.

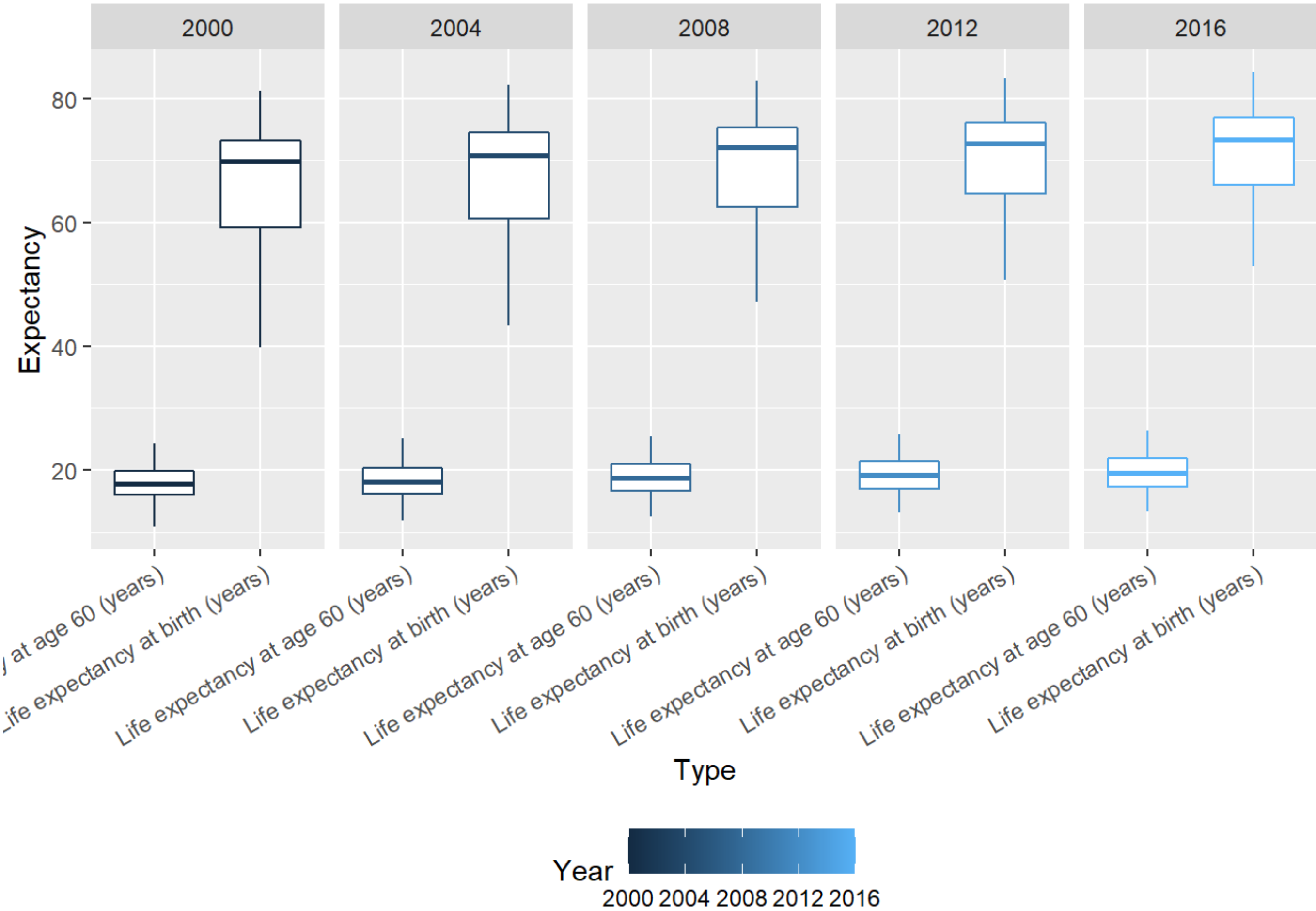

Figure 4.4: Life expectancy at birth (years) and at age 60 (years) boxplot in 2000, 2004,2008, 2012, 2016

Table 4.4 lists the average life expectancy at birth and at 60, taking into account of all countries where their life expectancy informations are available. Yearly increments at birth and at 60 represent the increasing proportion of the average each year comparing with the average of the former year. Similar to the results of boxplots, the averages had continued increasing since 2001. However, the extent of the rising pattern for average life expectancy at 60 is relatively insignificant to that at birth.

Table 4.4: Average life expectancy at birth, at age 60 (years) and their yearly increment

Year	average LE at birth	average LE at 60	Yearly increment at birth	Yearly increment at 60
2016	71.79	19.76	0.29	0.12
2015	71.50	19.64	0.25	0.04
2014	71.25	19.60	0.20	0.08
2013	71.05	19.52	0.30	0.12
2012	70.75	19.40	0.26	0.07
2011	70.49	19.33	0.54	0.21
2010	69.95	19.12	0.22	0.01
2009	69.72	19.11	0.45	0.13
2008	69.27	18.98	0.41	0.12
2007	68.86	18.85	0.42	0.13

Figure 4.5 exhibits the yearly improvements in average life expectancy both at birth and at 60 from 2000 to 2016. The increment for life expectancy at 60 was lower than that at birth for all times from 2000 to 2016. Nevertheless, the patterns for both of them are similar. For example, their increments in 2016 nearly drop to zero. On the next year, they both increase to the maximum increment from 2000 to 2016, as shown as the peaks.

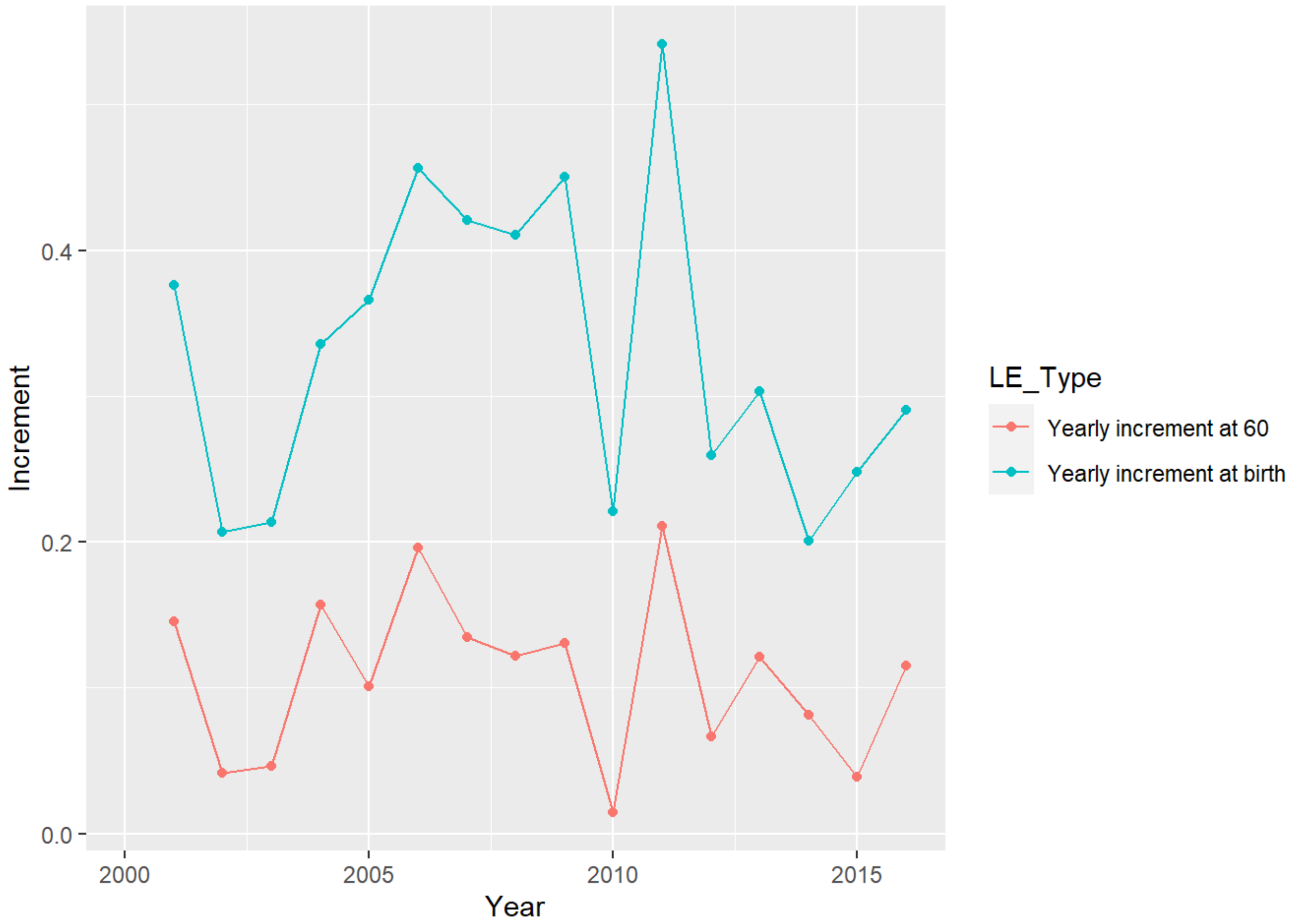

Figure 4.5: Yearly increment of average life expectancy at birth and at age 60 (years)

Table 4.5 contains the number of countries which had and had not met the average life expectancy from 2000 to 2016.

Table 4.5: Number of countries met and unmet average life expectancy at birth, at age 60 (years) each year

Year	meet_avg	number_countries_birth	number_countries_60
2016	met	121	97
2016	unmet	62	86
2015	met	116	96
2015	unmet	66	86
2014	met	114	95
2014	unmet	68	87
2013	met	116	92
2013	unmet	66	90

Year	meet_avg	number_countries_birth	number_countries_60
2012	met	115	90
2012	unmet	67	92

Figure 4.6 combines the number of countries that had met the average life expectancy at birth, at 60 and number of countries that had not met the averages from 2000 to 2016. As years pass on, more countries had been meeting and exceeding the average life expectancy at birth and at 60 each year. However, a pheonomenon is revealed here - although the number of countries meeting average life expectancy at 60 is lower than that at birth, the number of countries which did not meet the average is much higher. Furthermore, since 2001, the number of countries meeting average life expectancy at birth had already surpassed that of not meeting. Yet, for life expectancy at 60, it was not until 2013 that the number of countries meeting the average life expectancy at 60 began to exceed that of unmet number of countries.

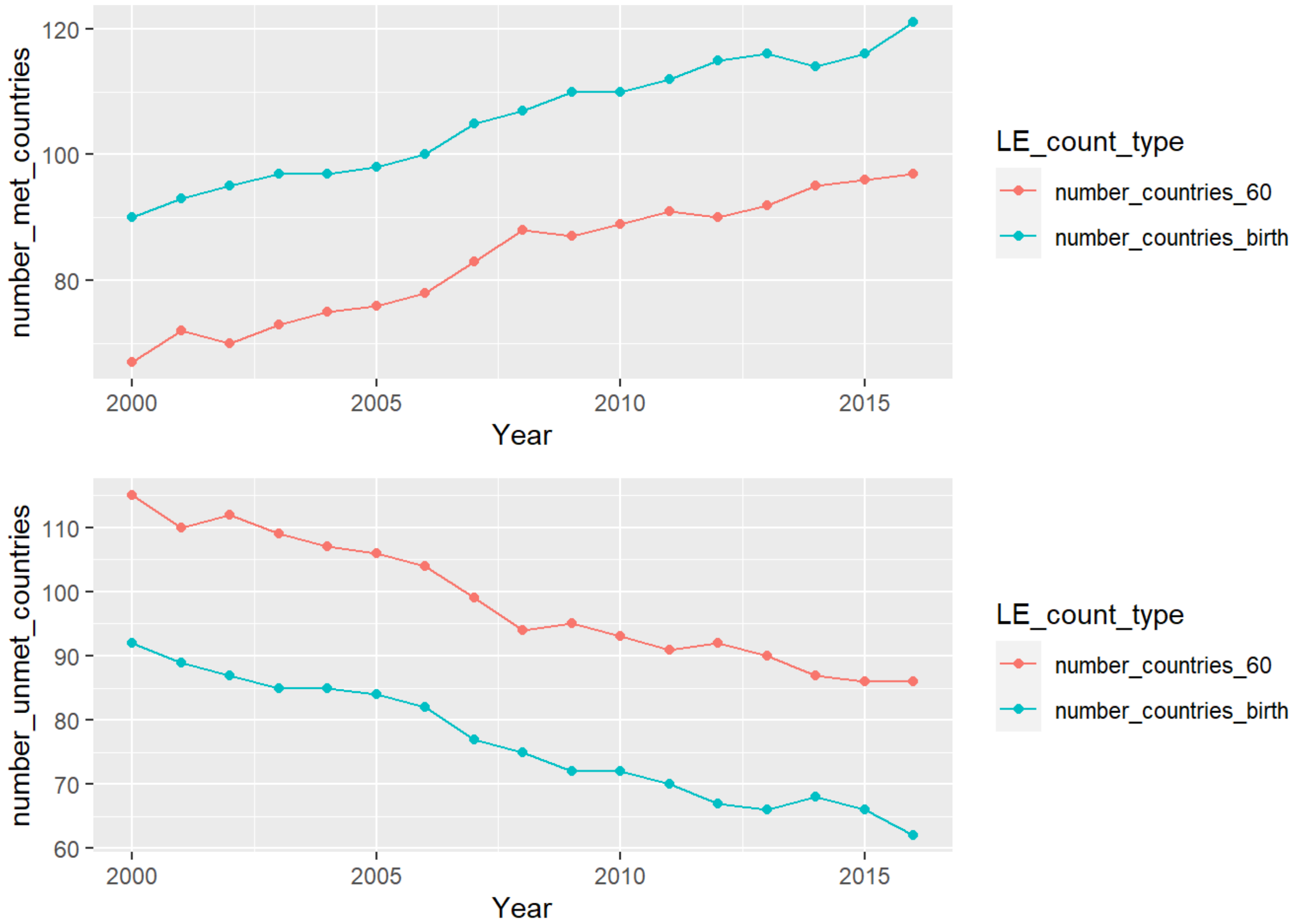

Figure 4.6: Trend of number of countries meeting and not meeting each year average LE at birth and 60

4.3 Life expectancy and Gross Domestic Product

Table 4.6: Average life expectancy and GDP per capita by continents, 2016

Continent	Average Life Expectancy	Average GDP per capita
Europe	78.83	35236.88
North America	75.08	12585.28

Continent	Average Life Expectancy	Average GDP per capita
South America	74.53	7937.75
Asia	73.76	12527.48
Oceania	73.18	11090.62
Africa	63.42	2478.58

Table 4.6 shows world's average life expectancy and GDP per capita by continents in 2016. People in Europpe had the highest income and lifespan. In contrast, Africa has lowest GDP per capita and life expectancy. This finding goes in line with that of Preston: That individuals born in richer countries, on average, can expect to live longer than those born in poor countries, and the curve between income and life expectancy will flat out (Preston 1975). However, others researchs suggest that there is causality that goes from income to health (Preston 2007). Because of limited scope of this study, we decide not go into detail of this relationship.

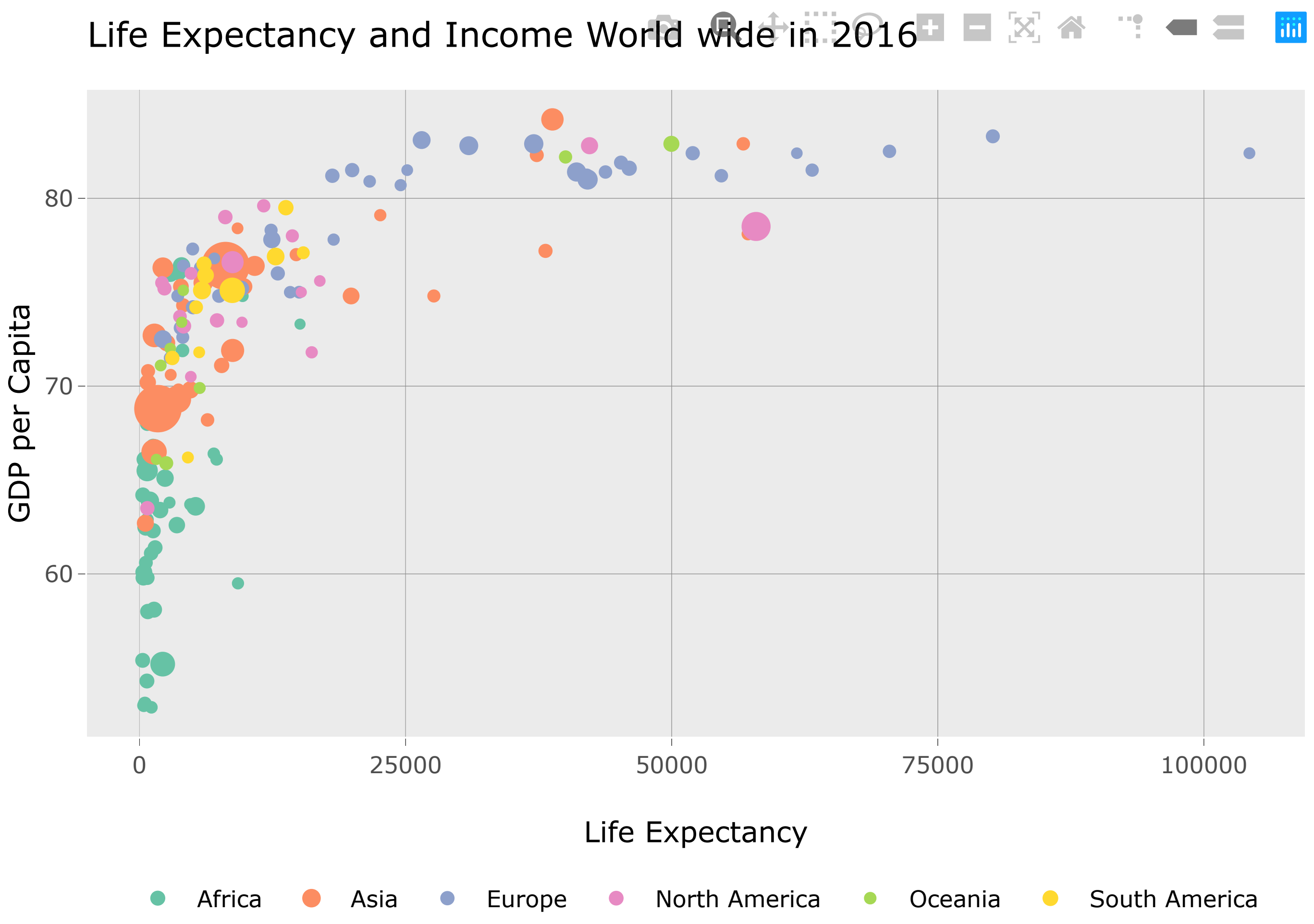

Figure 4.7: Life Expectancy and Income World wide in 2016

Figure 4.7 shows income and lifespan in the world in 2016. Each point is a country. Color shows region while size shows population. X axis shows GDP per capita and Y axis shows lifespan in years. The countries with lower income tend to have lower life expectancy. No high income countries have short life expectancy, and no low income countries have long life expectancy. Most people live in mid income countries. Therefore, for a more detailed view on the majority of population worldwide, we create figure 4.8. Figure 4.8 shows the life expectancy of countries with GDP per capita less then 20000 USD. There are huge different in lifespan on the same income level. This difference may be explained by how the money is used, or in other words, the country's health expenditure.

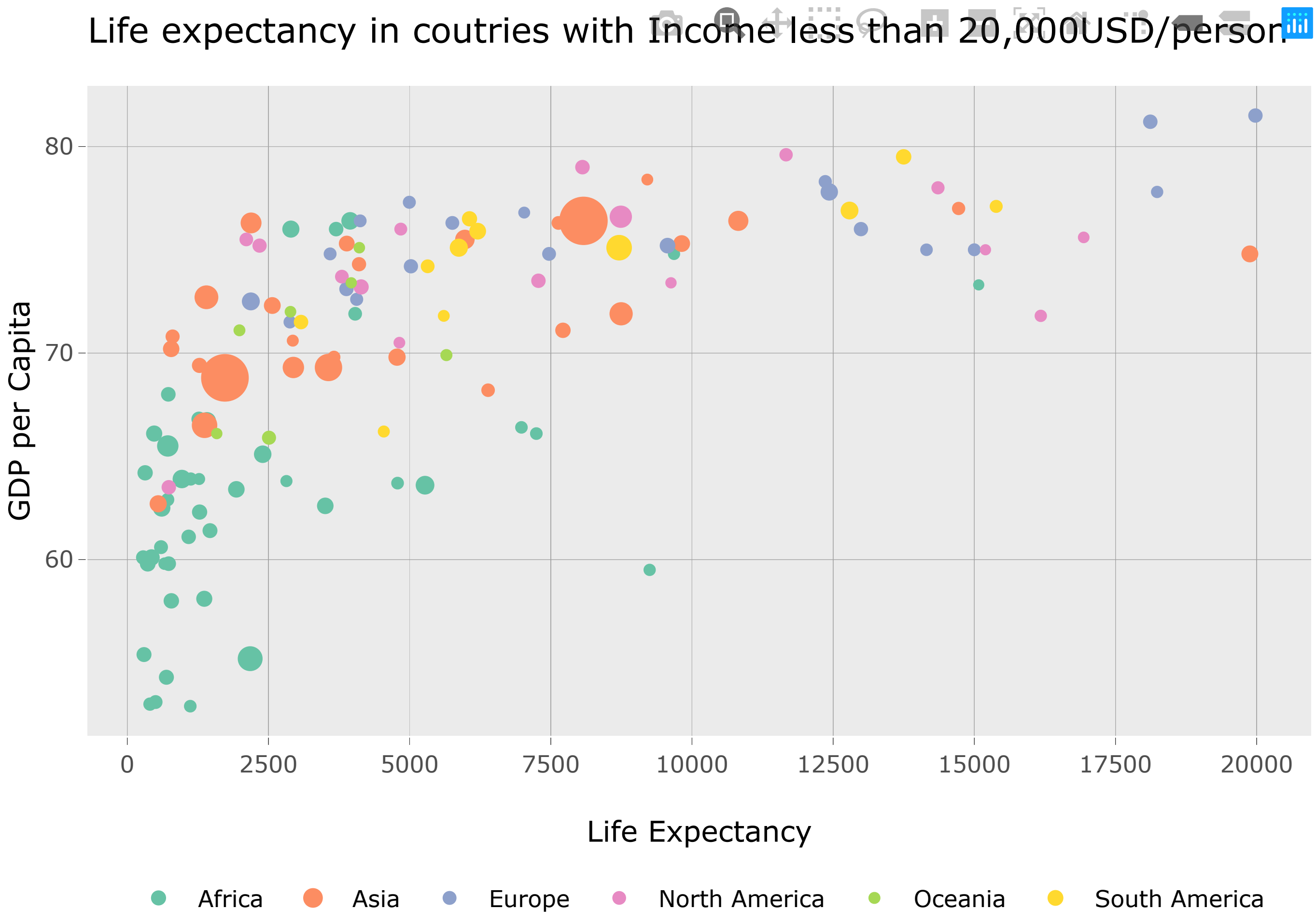

Figure 4.8: Life expectancy in coutries with Income less than 20,000USD/person

4.4 Life expectancy and Health expenditure

What is the relationship between life expectancy and health expenditure? Do countries that spend more on health care have higher life expectancy? We will explore this in this section.

4.4.1 Dynamic plots for Life expectancy and Health expenditure

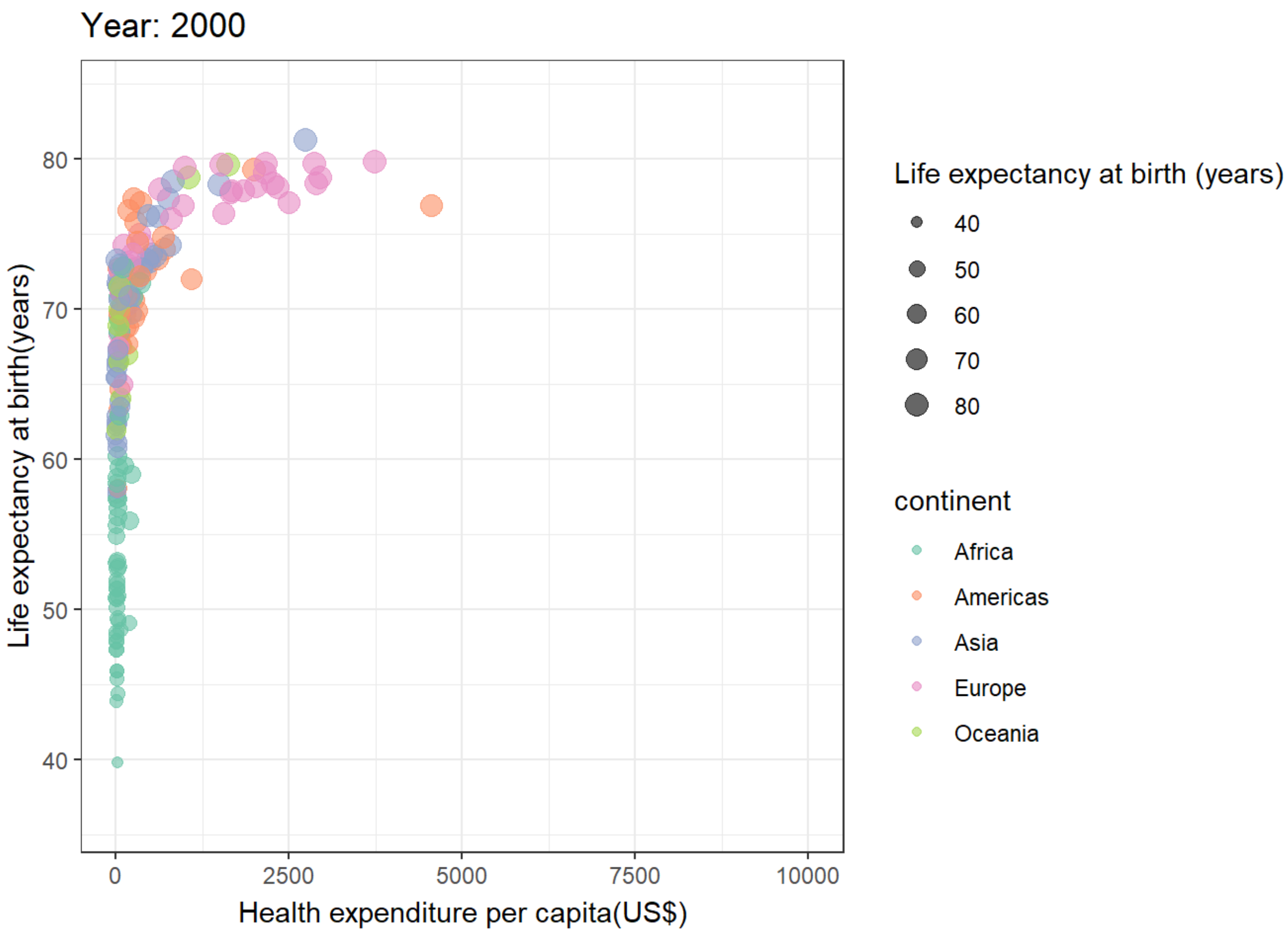

Figure 4.9: The plot for Life expectancy at birth and Health expenditure by year

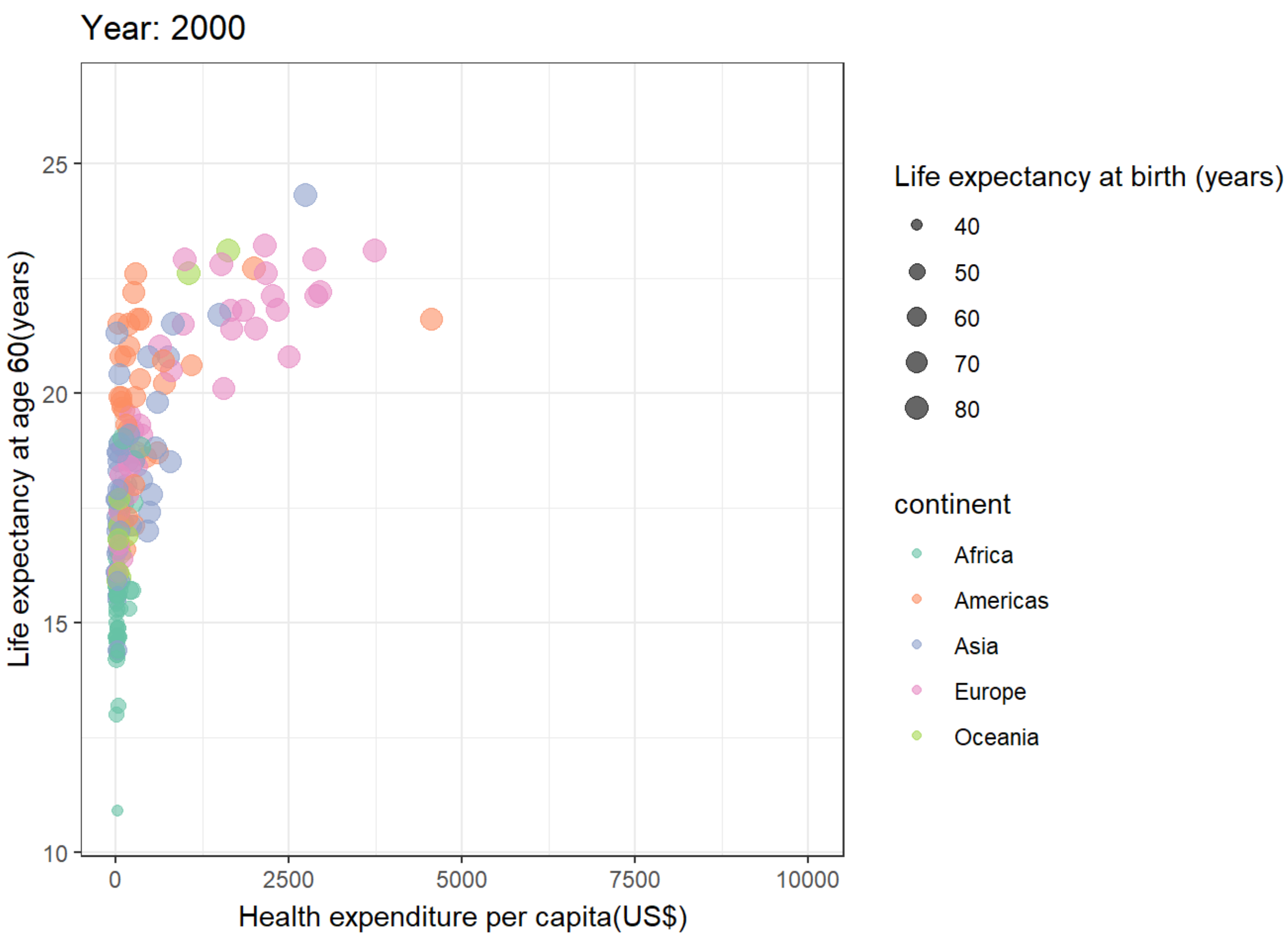

Figure 4.10: The plot for Life expectancy at age 60(years) and Health expenditure by year

Figure (4.9,4.10)show the dynamic relationship between life expectancy at birth(at age 60) and health expenditure from 2000 to 2016. We find that the countries with the lowest life expectancy and health expenditure are mainly in Africa, while Europe, Oceania and some countries in the Americas are far ahead. Life expectancy in most Asian and American countries is concentrated between 65 and 75 years. Also, almost all countries have increased their spending on health over time and improved life expectancy over time.

However, there was an outlier country in the Americas showed a significant drop in life expectancy in 2010. Combined with Table 4.7, we can find that this country is Haiti, which was hit by a large-scale earthquake in 2010 that killed over 230,000 people(Bilham 2010), which led to a significant drop in life expectancy in Haiti. Another interesting fact is that the country with the highest health expenditure is not the country with the highest life expectancy. Combine with Table 4.8, and we can find that the country with the highest life expectancy is in Asia, Japan. Japan has the highest life expectancy in the world mainly because of its government regulation and support for health care(Ikeda et al. 2011).

Table 4.7: The lowest life expectancy in the Americas

Country	Year	Life expectancy at birth (years)	Life expectancy at age 60 (years)	Health expenditure per capita(US$)
Haiti	2010	36.2	10.7	54.6

Table 4.8: The highest life expectancy in the Asia

Country	Year	Life expectancy at birth (years)	Life expectancy at age 60 (years)	Health expenditure per capita(US$)
Japan	2016	84.2	26.4	4174.9

4.4.2 Map for average health expenditure

The average health expenditure per by country

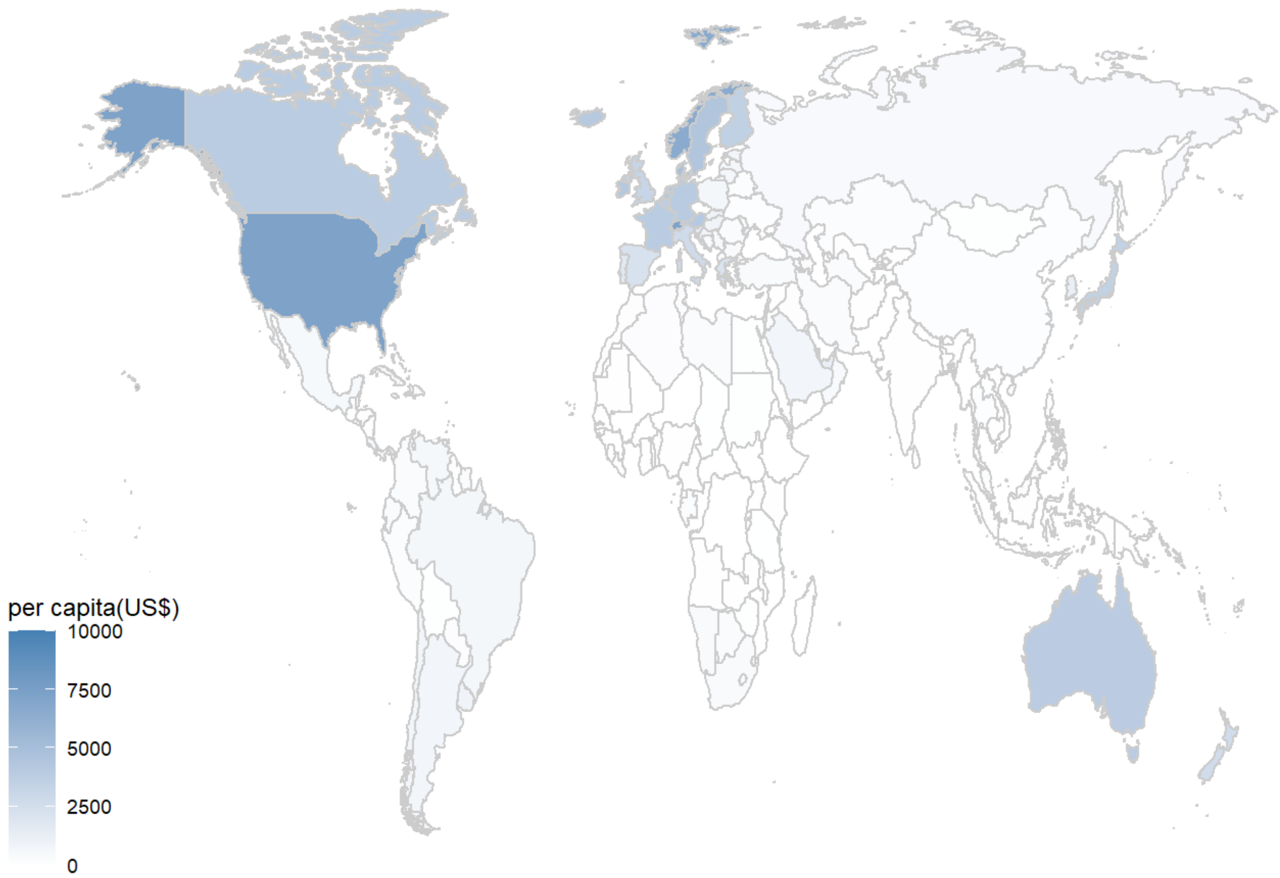

Figure 4.11: The average health expenditure per capita by country

By mapping Figure 4.11, it can be found that the highest per capita health expenditure is mainly in European countries, the United States, Canada, Australia, New Zealand and Japan. Some countries are 0 due to missing health expenditure data.

4.4.3 Model for Healthy life expectancy and Health expenditure by continent

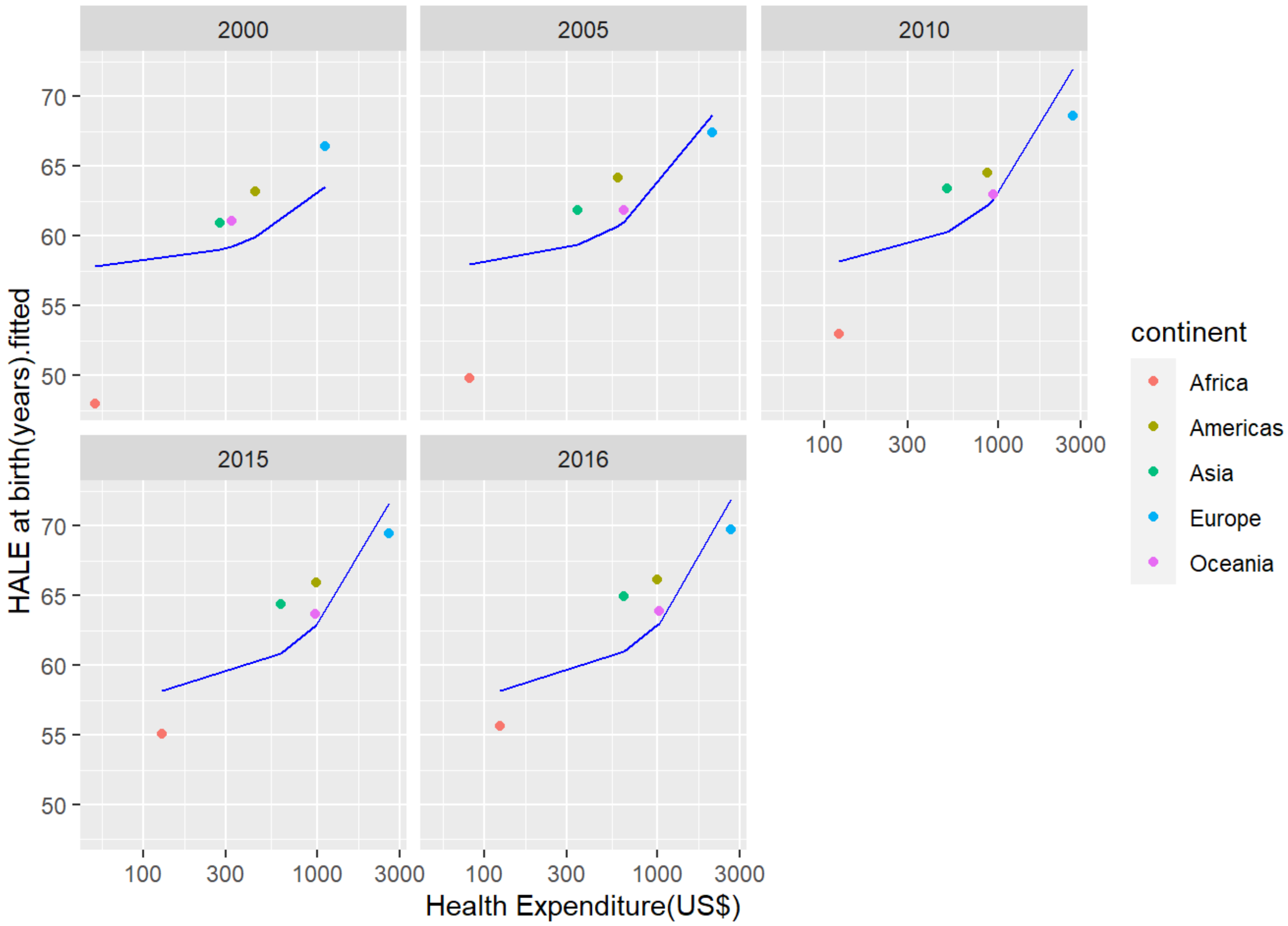

Figure 4.12: The plot for Healthy life expectancy at birth(years) and Health expenditure

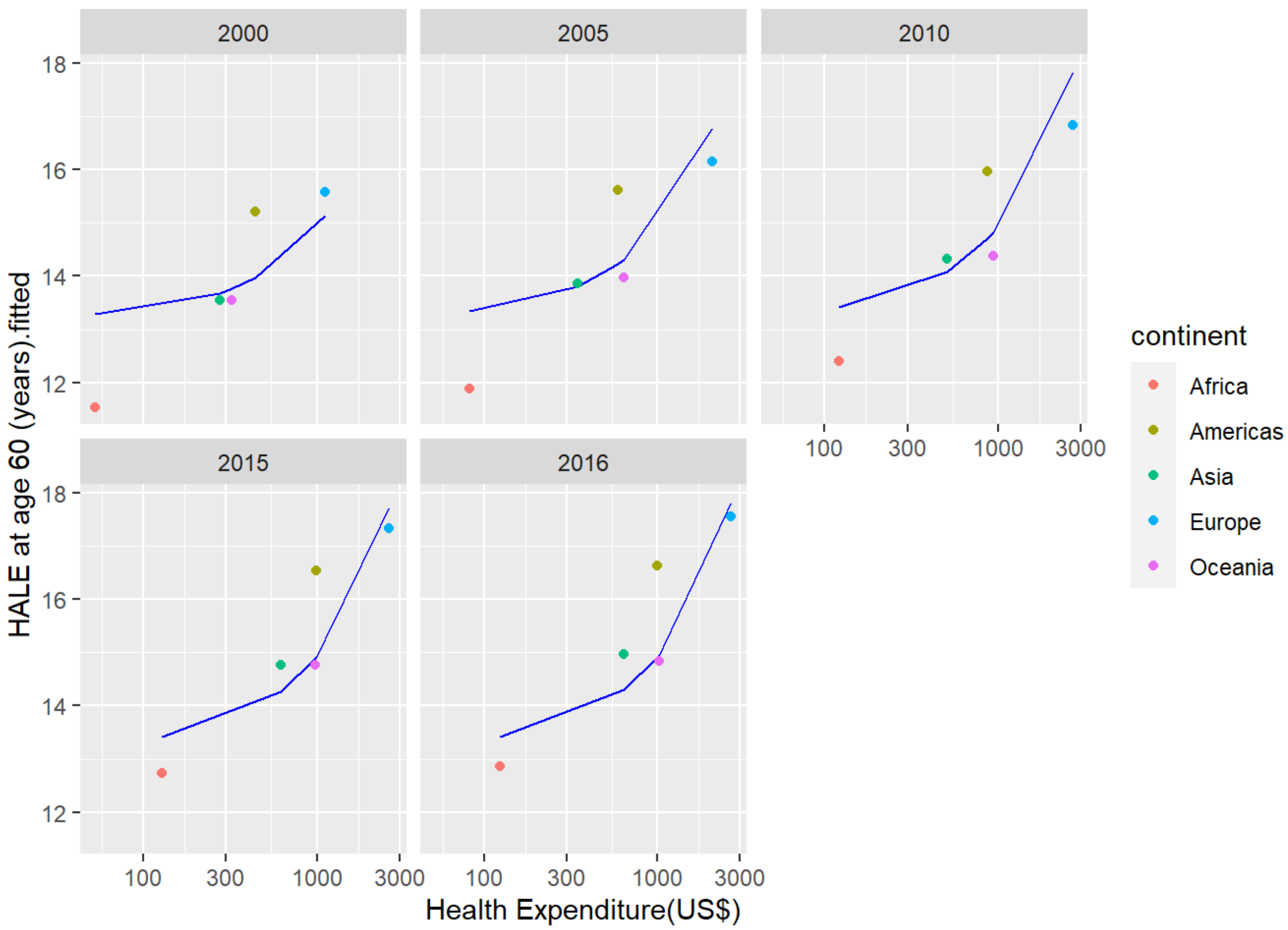

Figure 4.13: The plot for Healthy life expectancy at age 60(years) and Health expenditure

From Figure (4.12, 4.13), we can see that Africa is the worst fitting one both in Model for Birth and Model for Age 60, while Oceania has better fitting in both. As shown in Table 4.9, we can infer that the Model for Age 60 is better than Model for Birth, because it has the higher r.squared which explains the variance in the model. Model for Age 60 explains 69.11% of the dependent variable. Also, Model for Age 60 has the lower sigma, AIC and BIC, which gives it more confidence and makes it the best model fit(Kenny 2015).

Table 4.9: Model information

Model	r.squared	sigma	AIC	BIC
Birth	0.5721	3.85	142.22	145.87
Age 60	0.6911	0.95	72.26	75.92

Table 4.10: Confidence interval for model

Model	term	estimate	std.error	statistic
Birth	(Intercept)	57.521	1.147	50.15
Birth	`Health expenditure per capita(US$)`	0.005	0.001	5.55
Age 60	(Intercept)	13.204	0.283	46.64
Age 60	`Health expenditure per capita(US$)`	0.002	0.000	7.17

Through Table 4.10, we can get the formula of Model for Birth as follows:

$$HALE(\hat{at\ birth}) = 57.521 + 0.005 Health\ expenditure$$

And the formula of Model for Age 60 can be represented by:

$$HALE(\hat{at\ age\ 60}) = 13.204 + 0.002 Health\ expenditure$$

From the above analysis, we can learn that the amount of health expenditure per capita has a more significant impact on healthy life expectancy at age 60 than healthy life expectancy at birth. It’s possibly because people are more likely to get sick as they get older and need more health care to live longer.

5 Conclusion

The aim for this report was to respond to the four research questions outlined in the introduction. Based on the above analysis, the following conclusions can be drawn.

Firstly, the oft-reported gap in life expectancy for men and women is supported by the WHO data. Across the 16 years of recorded data, women were expected to outlive men by, on average, 4.7 years, possessing an average life expectancy of 71.6 years compared to 66.9 years for men. Furthermore, the gender variable appears to explain over 64% of this difference in the linear model produced.

Secondly, it was found that both the average life expectancy measured at birth and measured again at age 60 were increasing across the time period captured by the dataset. However, the rate of increase for measurements taken at birth was significantly higher than those taken at age 60. Furthermore, as time progressed an increased number of countries had met or exceeded both the average life expectancy at birth and at age 60, albeit at a slower rate for the measurements taken at age 60.

Thirdly, a positive relationship was found between a country's GDP per capita and its life expectancy with countries with a lower GDP having a lower expected lifespan. Furthermore no countries with a high GDP have a short life expectancy and no low GDP countries have a long life expectancy, perhaps indicating that this variable is an important indicator of an individuals expected lifespan.

Finally, similar to GDP, a dynamic relationship was found between life expectancy and health expenditure. The average life expectancy increased alongside the average health expenditure for almost all countries included in the WHO data although external forces can impact this trend. While generally, increased health expenditure was associated with increased life expectancy, the relationship was not perfectly linear. However, health expenditure was found to more accurately predict life expenditure at age 60, rather than at birth, possibly due to an increased need for medical care as one ages.

6 Acknowlegments

Our team would like to thank all the contributors to the following R package: Wickham et al. (2019), Zhu (2019), Wickham (2016), Xie (2020), Müller (2017), R Core Team (2020), Arnold (2019), Arel-Bundock, Enevoldsen, and Yetman (2018), Wickham, Hester, and Chang (2020), Pedersen (2018), Pedersen and Robinson (2020), Robinson and Hayes (2020), Richard A. Becker, Ray Brownrigg. Enhancements by Thomas P Minka, and Deckmyn. (2018), Sievert (2020), Xie (2016), Auguie (2017).

References

Arel-Bundock, Vincent, Nils Enevoldsen, and CJ Yetman. 2018. "Countrycode: An R Package to Convert Country Names and Country Codes." *Journal of Open Source Software* 3 (28): 848. https://doi.org/10.21105/joss.00848 (https://doi.org/10.21105/joss.00848).

Arnold, Jeffrey B. 2019. *Ggthemes: Extra Themes, Scales and Geoms for 'Ggplot2'*. https://CRAN.R-project.org/package=ggthemes (https://CRAN.R-project.org/package=ggthemes).

Auguie, Baptiste. 2017. *GridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra (https://CRAN.R-project.org/package=gridExtra).

Bank, TW. 2018a. "Population Total by Country." *The World Bank*.

———. 2018b. *The World Bank*.

Bilham, Roger. 2010. "Lessons from the Haiti Earthquake." *Nature* 463 (7283): 878–79.

Dbouquin. 2016. "Countries-Continents." https://raw.githubusercontent.com/dbouquin/IS_608/master/NanosatDB_munging/Countries-Continents.csv (https://raw.githubusercontent.com/dbouquin/IS_608/master/NanosatDB_munging/Countries-Continents.csv).

Ikeda, Nayu, Eiko Saito, Naoki Kondo, Manami Inoue, Shunya Ikeda, Toshihiko Satoh, Koji Wada, et al. 2011. "What Has Made the Population of Japan Healthy?" *The Lancet* 378 (9796): 1094–1105.

Kenny, David A. 2015. "Measuring Model Fit."

Luy, Marc, and Yuka Minagawa. 2014. "Gender Gaps-Life Expectancy and Proportion of Life in Poor Health." *Health Reports* 25 (12): 12.

Müller, Kirill. 2017. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here (https://CRAN.R-project.org/package=here).

Organization, World Health, and others. 2018. “Life Expectancy and Healthy Life Expectancy. Data by Country” 688. https://apps.who.int/gho/data/view.main.SDG2016LEXv?lang=en (https://apps.who.int/gho/data/view.main.SDG2016LEXv?lang=en).

———. 2020 688. https://apps.who.int/gho/data/view.main.GHEDCHEpcUSSHA2011REGv?lang=en (https://apps.who.int/gho/data/view.main.GHEDCHEpcUSSHA2011REGv?lang=en).

Pardi, Faridah, Abdol Samad Nawi, and Arifin Md Salleh. 2016. “Determining Factors Toward a Sustainable Development Path in Selected Asean Countries and Japan-Evidence from Panel Data Analysis.” *International Journal of Business and Information* 11 (2): 125.

Patten, Bernard C, and Robert Costanza. 1997. “Logical Interrelations Between Four Sustainability Parameters: Stability, Continuation, Longevity, and Health.” *Ecosystem Health* 3 (3): 136–42.

Pedersen, Thomas Lin. 2018. *Transformr: Polygon and Path Transformations*. https://CRAN.R-project.org/package=transformr (https://CRAN.R-project.org/package=transformr).

Pedersen, Thomas Lin, and David Robinson. 2020. *Gganimate: A Grammar of Animated Graphics*.

Preston, Samuel H. 1975. “The Changing Relation Between Mortality and Level of Economic Development.” *Population Studies* 29 (2): 231–48. http://www.jstor.org/stable/2173509 (http://www.jstor.org/stable/2173509).

Preston, Samuel H. 2007. “The changing relation between mortality and level of economic development.” *International Journal of Epidemiology* 36 (3): 484–90. https://doi.org/10.1093/ije/dym075 (https://doi.org/10.1093/ije/dym075).

Rabbi, AM Fazle. 2013. “Imbalance in Life Table: Effect of Infant Mortality on Lower Life Expectancy at Birth.” *Journal of Scientific Research* 5 (3): 479–88.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/ (https://www.R-project.org/).

Richard A. Becker, Original S code by, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka, and Alex Deckmyn. 2018. *Maps: Draw Geographical Maps*. https://CRAN.R-project.org/package=maps (https://CRAN.R-project.org/package=maps).

Robinson, David, and Alex Hayes. 2020. *Broom: Convert Statistical Analysis Objects into Tidy Tibbles*. https://CRAN.R-project.org/package=broom (https://CRAN.R-project.org/package=broom).

Sievert, Carson. 2020. *Interactive Web-Based Data Visualization with R, Plotly, and Shiny*. Chapman; Hall/CRC. https://plotly-r.com (https://plotly-r.com).

Tosato, Matteo, Valentina Zamboni, Alessandro Ferrini, and Matteo Cesari. 2007. “The Aging Process and Potential Interventions to Extend Life Expectancy.” *Clinical Interventions in Aging* 2 (3): 401.

Venkataraman, Bhawani. 2009. “Education for Sustainable Development.” *Environment: Science and Policy for Sustainable Development* 51 (2): 8–10.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org (https://ggplot2.tidyverse.org).

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686 (https://doi.org/10.21105/joss.01686).

Wickham, Hadley, Jim Hester, and Winston Chang. 2020. *Devtools: Tools to Make Developing R Packages Easier*. https://CRAN.R-project.org/package=devtools (https://CRAN.R-project.org/package=devtools).

Xie, Yihui. 2016. *Bookdown: Authoring Books and Technical Documents with R Markdown*. Boca Raton, Florida: Chapman; Hall/CRC. https://github.com/rstudio/bookdown (https://github.com/rstudio/bookdown).

———. 2020. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://yihui.org/knitr/ (https://yihui.org/knitr/).

Zhu, Hao. 2019. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra (https://CRAN.R-project.org/package=kableExtra).