Global suicide Rates and its correlation to country income, region, age, and gender in 2019*

An analysis of global suicide rates in 2019 using multiple linear regression

Hyuk Jang

April 16, 2024

Suicide has become a main cause of deaths globally over the decades. This creates curiosity in what immutable factors may be affecting suicide rates. This paper looks into the correlation between suicide rates and gender, sex, age, country income, and region. Then further analysis is done by looking into human resources available in each country.

Table of contents

*Code and data are available at: https://github.com/anggimude/Health.

1 Introduction

Suicide is defined as the act of intentionally causing one's own death. The main risk factors of suicide are mental disorders, physical disorders, and substance abuse. In particular there is a significant correlation between suicidal behaviour and mental health[write more about this based on research]. In 2019, 1.3% of all deaths resulted from suicide and is the fourth leading cause of deaths in 15~29 years old. This paper will dive deeper into suicide rates and its correlation between country income, age, gender, and region in 2019. The most recent year with relevant data is for 2019, from the WHO[citation] and since we are only looking into one year, we look at the global trends and world bank income group trends to get a more thorough understanding of the trends of suicide rates of 2019. Because one of the goals of the paper is to observe the correlation between suicide rates and country income and region, the countries are classified based on the world bank income group; with four income groups - low, lower-middle, upper-middle, and high. Low-income economies is defined as a country with a gross national income less than $1135, and in the manner lower-middle is between the range $1136 to $4465, upper-middle in the range of $4455 to $13845 and high is $13846 or more. Region is defined based on the continents; Asia, Africa, North America, South America, and Europe. The total number of countries is 183 which are all mapped to different regions and income groups.

This paper will produce a multiple linear regression to analyze the correlation coefficients of each factor. The results of the model will be checked by doing a 95% confident interval on the regression coefficients. In addition, we will be able to study the trends of the factors of suicide. We want to analyze the change in demographics for the different age groups, country income, region, and gender. The data was obtained from the WHO database[citaiton] under the mental health category. R Core Team (2023) was used to clean the raw data to what we use for modelling in other to write an analysis of the intended study. The estimand of this paper are the intercepts and coefficients of each predictors which correspond to region/income/sex and the different age categories.

This paper has 4 sections in total not including the introduction. In Section 2 we look at the data that used to carry out the reports including tables and graphs of cleaned data that will be used for the models and the summary statistics. In the next section, we discuss about the models that will be used to analyze our cleaned data, how it is set up and the justifications of it. Next we display and examine the results obtained from the models including tables of the model summaries which helps us make predictions. Lastly, we make final discussions of our results and research based on each cause and dive into some weaknesses that our paper has. In addition, we explore some next steps we or anyone else interested is willing to take after reading this paper. [make this paragraph more specific]

2 Data

2.1 Raw Data

The data used in this paper is derived from WHO((**who2021global?**)), and WHO data repository() and was downloaded from the data available in the links. WHO((**who2021global?**)) provides data of age-standardized suicide rates per 100,000 population for 183 countries for the years 2000-2019. WHO data repository() provides data of crude suicide rates per 100,000 population for the same 183 countries for the year 2019. Since the only data avaiable data for crude suicide rates depending on age is 2019, we only look into the year 2019 for WHO() as well. The data the two tables include are the 183 countries, sex (both, male, female), age groups of 85+, 75-84, 65-74, 55-64, 45-54, 35-44, 25-34, 15-24, and age standardized suicide rates. The age-standardized suicide rate is used as the dependent variable for our model as we are interested in the correlation between the independent variables which are gender, age group, country income, and region.

The cleaning, testing, and modelling of the data for this paper was done through R (R Core Team 2023) with the aid of the following packages: tidyverse (**citetidyverse?**), dplyr (**citedplyr?**), rstanarm (**citerstanarm?**), ggplot2 (**citeggplot2?**), modelsummary (**citemodelsummary?**), kableExtra (**citekableExtra?**), arrow(**citearrow?**), tidybayes(**citetidybayes?**).

2.2 Cleaned Data

The goal of the cleaning process of this paper is to create a table including income group(high, upper middle, lower middle, low), sex(both, male, female), region(Asia, Europe, North America, South America, Africa, Oceania) as the rows, and age-standardized suicide rates, and the 10 year age group suicide rates. This is so that we can create a model which will provide us the correlation of global suicide rates of 2019 against income group/sex/region and age groups. To do this we merge the two raw data tables we have downloaded. Before we merge the tables, WHO() includes the age-standardized suicide rates as a form of the actual value[lower limit, upper limit]. When we create our models or graphs, the interval is not necessary, so we first remove the interval so that each column only consists of the actual rate. After merging and doing some minor cleaning like removing repeated rows and changing names of columns and rows, the table will have country, sex, age-standardized suicide rates, and the 10 year age groups. We must also include the region and income group of each data so we create a mapping so that we country will have additional columns of continent and world bank income group. This was done based on the information provided from [citation] [citatiom]. Since we are looking into the global suicide rates and how the independent variables affect it, we must calculate the averages of the age-standardized suicide rates and age groups depending on sex, income group, region. The table with age-standardized suicide rates and suicide rates for the 10 year age groups depending on the 183 countries is available in Section A.

Below, Table 1 is a table of the averages of suicide rates of age standardized and 10 year age groups depending on region, country income, and sex. The rows include global, male, female, lower income, lower-middle income, upper-middle, high income, Asia, Europe, North America, South America, Africa, and Oceania. We observe the average suicide rates of each of these variables for age-standardized and the 10 year age groups. The most noticeable trend from Table 1 is the difference in suicide rates in male and female. While female had a suicide rate of 4.54, male's was 16.14 and the global average being 10.09 exhibiting significance of further research. Within the income groups, we see that the lower income groups tend to have a higher average suicide rate compared to the higher income countries. The lower income groups have an average of around 11.50 while the comparatively higher income is 9.00. The difference is not significant as the global average is 10.00, but it is worth looking into these trends in Section 5. Comparing the data of each region, we can notice that Asia, North America, and Europe are below the average, on the other hand, South America, Oceania, and Africa have suicide rates above the average. For the regions, the deviation from the global average is larger than income groups, but smaller than sex. Here we can notice the similarities with the income group data. For example, a large proportion of African and South American countries lie in the lower and lower-middle income bracket which has a comparably higher suicide rate, and it shows that South America and Africa have a suicide rate of 11.44 and 13.40 respectively.

Figure 1 and Figure 2 show the Table 1 as a scatter plot. The graph is divided into two to improve clarity and legibility. The split allows for easier viewing of labels, axis, and point, ensuring a better understanding of the data. The scatter plot is build for an analysis of the age group trends. We can observe that the older age groups, Age 85+ and Age 75-84 have the highest suicide rates. Especially the 85+ age group is an outlier in every region/income/sex group. We can also notice that the younger the age, the lower suicide rates tends to be. Some notable exceptions are Oceania's high suicide rate among the ages 15-44. Unlike the other region/income/sex groups, Oceania's younger generation and 85+ has the highest suicide rates. In addition, the older the age group, we notice more extreme suicide rates that goes above 100 per 100,000 population. This occurs three times all of which are in 85+ of age, and male, lower income, and Africa displays such results. The x-axis variables other than Africa, male, and lower income, the data points tend to be gathered up around the average of each variable, meaning that the standard deviation of each region/income group/sex is small.

Table 1: Global Suicide Rates by Age, Region, Sex, and Income Group

Region/Income/Sex	Age-standardized suicide rates (per 100 000 population)	Age 85+	Age 75-84	Age 65-74	Age 55-64	Age 45-54	Age 35-44	Age 25-34	Age 15-24
Global	10.09	59.32	27.47	20.90	17.56	15.46	12.90	10.80	7.79
Male	16.14	121.59	48.42	34.47	28.35	24.87	20.75	17.01	11.33
Female	4.54	27.26	13.56	10.12	7.91	6.40	5.12	4.42	4.06
Lower Income	11.08	131.75	50.75	37.67	25.74	17.13	9.64	6.68	4.54

Region/Income/Sex	Age-standardized suicide rates (per 100 000 population)	Age 85+	Age 75-84	Age 65-74	Age 55-64	Age 45-54	Age 35-44	Age 25-34	Age 15-24
Lower Middle	11.97	75.86	31.40	24.30	19.63	18.29	16.23	12.71	9.54
Upper Middle	8.25	40.10	20.62	15.29	13.09	11.97	11.01	9.87	6.95
High Income	9.39	25.35	18.55	14.50	15.60	15.00	12.88	11.71	8.38
Asia	6.92	28.96	15.51	11.13	9.54	9.28	9.42	9.10	7.10
Europe	9.73	32.25	22.08	17.31	17.89	16.80	13.20	10.55	7.10
North America	5.60	14.08	10.96	8.76	8.30	7.42	7.34	7.39	5.84
South America	11.44	34.30	21.25	17.77	16.79	14.37	14.70	16.32	13.03
Africa	13.40	131.11	50.86	37.99	27.77	22.81	15.69	9.88	5.74
Oceania	15.17	43.67	17.93	15.25	15.86	14.77	21.15	24.73	22.62

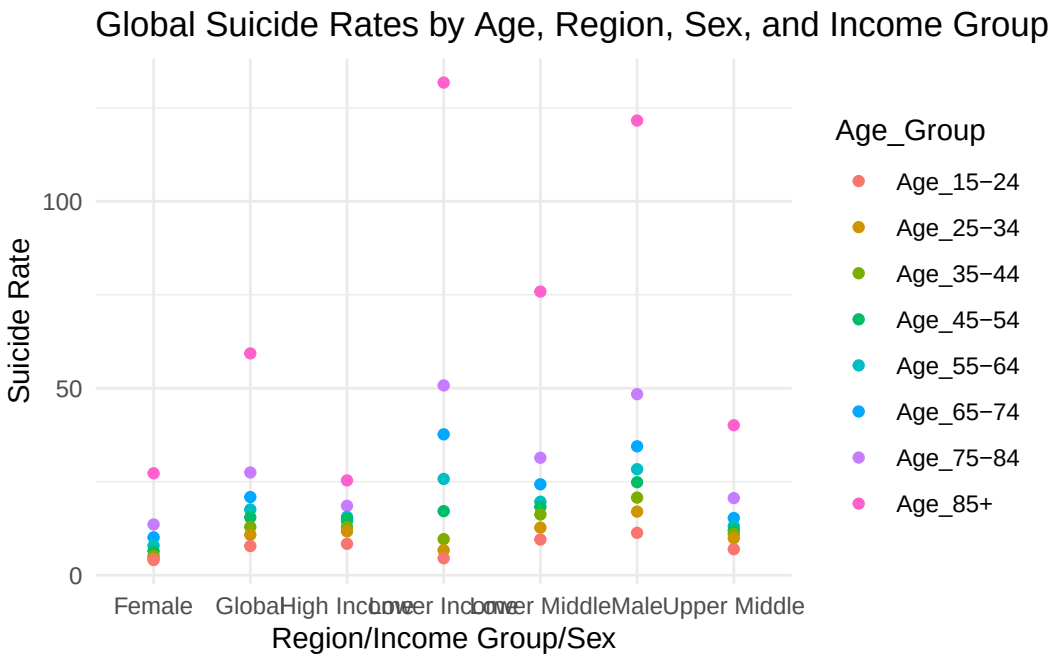


Figure 1: Global Suicide Rates by Age, Region, Sex, and Income Group - 1

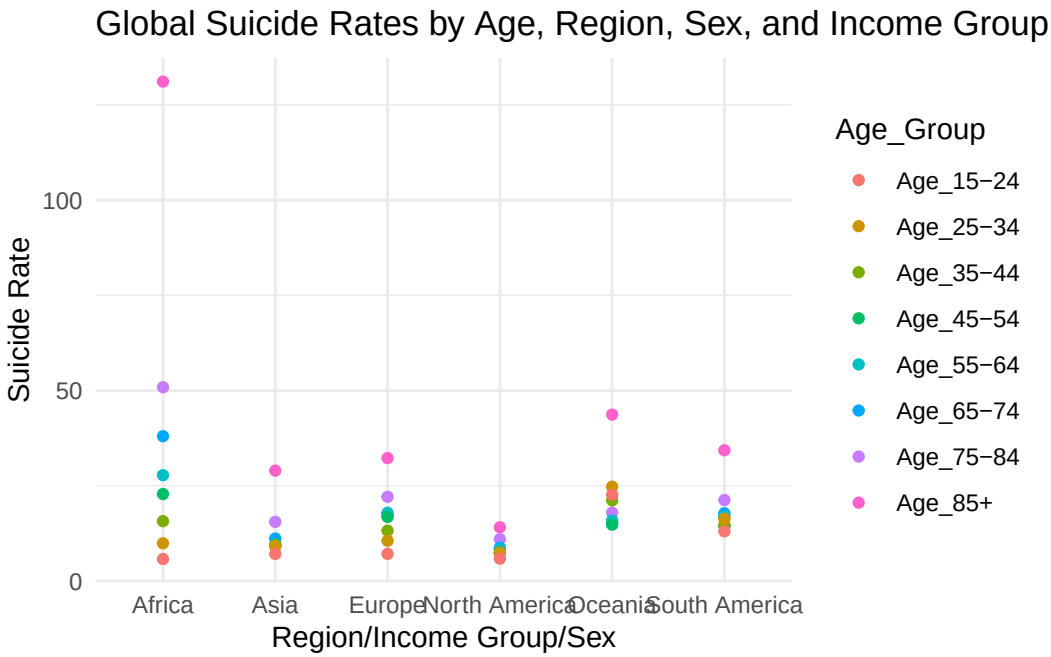


Figure 2: Global Suicide Rates by Age, Region, Sex, and Income Group - 2

2.3 Basic Summary Statistics

ea is a representation of the_data_one showing its minimum, mean, maximum, standard deviation, variance, and sample size. The summary statistics shows that the mean of the number of deaths are 18320, with a minimum of 6860, and a maximum of 44109 for a sample size of 24 as the data selected is from tersect_data_one. The standard deviation and variance may be abnormally high because the range of the data is large, in other words, because the mean is 18320, and the maximum is 44109, there is likely a few outliers in the data creating such a high standard deviation and variance.

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

Table 2: Summary statistics of suicide rates of different age groups

	Min	Mean	Max	SD	Var	N
Age-standardized suicide rates (per 100 000 population)	5	10	16	3	12	13
Age_85+	14	59	132	42	1806	13
Age_75-84	11	27	51	14	203	13
Age_65-74	9	20	38	10	104	13
Age_55-64	8	17	28	7	47	13
Age_45-54	6	15	25	5	29	13
Age_35-44	5	13	21	5	22	13
Age_25-34	4	12	25	5	28	13
Age_15-24	4	9	23	5	24	13

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular…

We can use maths by including latex between dollar signs, for instance θ.

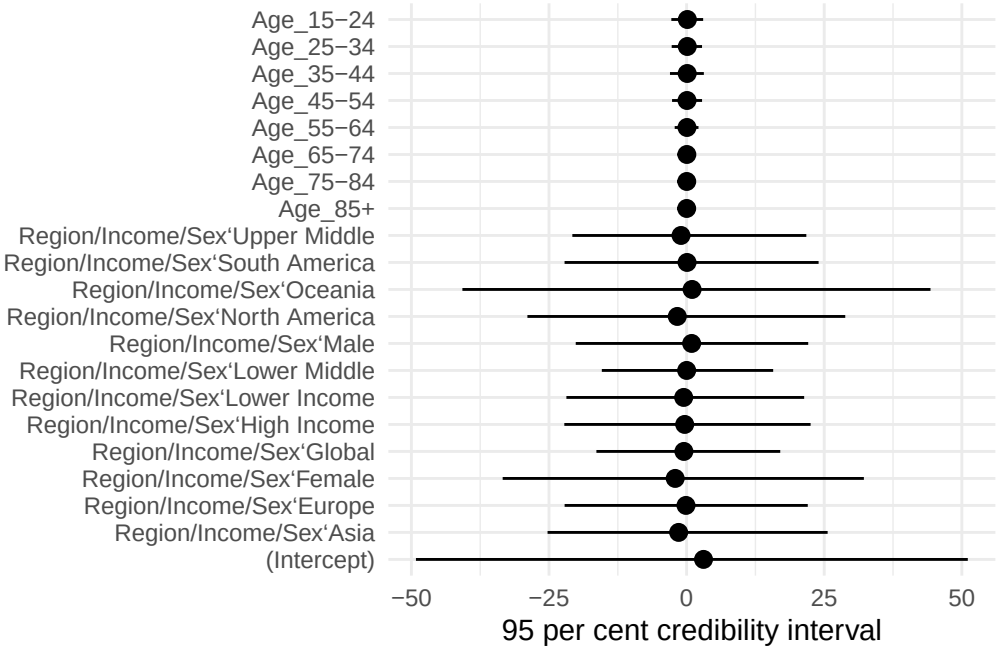
4 Results

Our results are summarized in Table 3.

Table 3: Explanatory models of flight time based on wing width and wing length

	Gaussian(Normal)
(Intercept)	3.07
	(26.07)
Region/Income/Sex‘Asia	−1.45
	(12.58)
Region/Income/Sex‘Europe	−0.11
	(11.24)
Region/Income/Sex‘Female	−2.07
	(16.91)
Region/Income/Sex‘Global	−0.49
	(7.91)
Region/Income/Sex‘High Income	−0.35
	(11.26)
Region/Income/Sex‘Lower Income	−0.53
	(10.81)
Region/Income/Sex‘Lower Middle	0.01
	(7.81)
Region/Income/Sex‘Male	0.91
	(10.74)
Region/Income/Sex‘North America	−1.70
	(14.63)
Region/Income/Sex‘Oceania	0.99
	(22.82)
Region/Income/Sex‘South America	0.08
	(11.91)
Region/Income/Sex‘Upper Middle	−1.02
	(10.47)
Age_85+	0.01
	(0.17)
Age_75-84	0.02
	(0.54)
Age_65-74	0.04
	(0.73)
Age_55-64	0.07
	(1.16)
Age_45-54	0.07
	(1.37)
Age_35-44	0.11
	(1.59)
Age_25-34	0.12
	(1.42)
Age_15-24	0.13
	(1.50)
Num.Obs.	13
R2	0.791
R2 Adj.	0.596
Log.Lik.	−26.819
ELPD	−39.0
ELPD s.e.	0.3
LOOIC	77.9
LOOIC s.e.	0.7
WAIC	73.7
RMSE	0.64

Table 4: **?(caption)**



5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

Table 5: Suicide rates for different countries for each age group

Country	Age-standardized suicide rates	Income Group	Region
Afghanistan	6.0	lower_income	Asia
Albania	3.7	upper_middle	Europe
Algeria	2.6	lower_middle	Africa
Angola	12.6	lower_middle	Africa
Antigua and Barbuda	0.3	high_income	North America
Argentina	8.1	upper_middle	South America
Armenia	2.7	upper_middle	Europe
Australia	11.3	high_income	Oceania
Austria	10.4	high_income	Europe
Azerbaijan	4.0	upper_middle	Europe
Bahamas	3.4	high_income	North America
Bahrain	7.2	high_income	Asia
Bangladesh	3.9	lower_middle	Asia
Barbados	0.3	high_income	North America
Belarus	16.5	upper_middle	Europe
Belgium	13.9	high_income	Europe
Belize	7.7	upper_middle	North America
Benin	12.7	lower_middle	Africa
Bhutan	5.1	lower_middle	Asia
Bolivia (Plurinational State of)	6.8	lower_middle	South America
Bosnia and Herzegovina	8.3	upper_middle	Europe
Botswana	20.2	upper_middle	Africa
Brazil	6.4	upper_middle	South America
Brunei Darussalam	2.5	high_income	Asia
Bulgaria	6.5	upper_middle	Europe
Burkina Faso	14.4	lower_income	Africa

Burundi	12.1	lower_income	Africa
Cote d'Ivoire	15.7	lower_middle	Africa
Cabo Verde	15.2	lower_middle	Africa
Cambodia	5.5	lower_middle	Asia
Cameroon	15.9	lower_middle	Africa
Canada	10.3	high_income	North America
Central African Republic	23.0	lower_income	Africa
Chad	13.2	lower_income	Africa
Chile	8.0	high_income	South America
China	6.7	upper_middle	Asia
Colombia	3.7	upper_middle	South America
Comoros	8.5	lower_middle	Africa
Congo	11.6	lower_middle	Africa
Costa Rica	7.6	upper_middle	North America
Croatia	11.0	high_income	Europe
Cuba	10.2	upper_middle	North America
Cyprus	3.2	high_income	Europe
Czechia	9.5	high_income	Europe
Democratic People's Republic of Korea	8.2	lower_income	Asia
Democratic Republic of the Congo	12.4	lower_income	Africa
Denmark	7.6	high_income	Europe
Djibouti	11.9	lower_middle	Africa
Dominican Republic	5.1	upper_middle	North America
Ecuador	7.7	upper_middle	South America
Egypt	3.4	lower_middle	Africa
El Salvador	6.1	upper_middle	North America
Equatorial Guinea	13.5	upper_middle	Africa
Eritrea	17.3	lower_income	Africa
Estonia	12.0	high_income	Europe
Eswatini	40.5	lower_middle	Africa
Ethiopia	9.5	lower_income	Africa
Fiji	9.5	upper_middle	Oceania
Finland	13.4	high_income	Europe

France	9.7	high_income	Europe
Gabon	13.1	upper_middle	Africa
Gambia	9.6	lower_income	Africa
Georgia	7.7	upper_middle	Europe
Germany	8.3	high_income	Europe
Ghana	10.5	lower_middle	Africa
Greece	3.6	high_income	Europe
Grenada	0.6	upper_middle	North America
Guatemala	6.2	upper_middle	North America
Guinea	12.3	lower_middle	Africa
Guinea-Bissau	12.4	lower_income	Africa
Guyana	40.9	high_income	South America
Haiti	11.2	lower_middle	North America
Honduras	2.6	lower_middle	North America
Hungary	11.8	high_income	Europe
Iceland	11.2	high_income	Europe
India	12.9	lower_middle	Asia
Indonesia	2.6	upper_middle	Asia
Iran (Islamic Republic of)	5.1	lower_middle	Asia
Iraq	4.7	upper_middle	Asia
Ireland	8.9	high_income	Europe
Israel	5.2	high_income	Asia
Italy	4.3	high_income	Europe
Jamaica	2.3	upper_middle	North America
Japan	12.2	high_income	Asia
Jordan	2.0	lower_middle	Asia
Kazakhstan	18.1	upper_middle	Europe
Kenya	11.0	lower_middle	Africa
Kiribati	30.6	lower_middle	Oceania
Kuwait	2.7	high_income	Asia
Kyrgyzstan	8.3	lower_middle	Asia
Lao People's Democratic Republic	6.0	lower_middle	Asia
Latvia	16.1	high_income	Europe
Lebanon	2.8	lower_middle	Asia
Lesotho	87.5	lower_middle	Africa
Liberia	7.4	lower_income	Africa

Libya	4.5	upper_middle	Africa
Lithuania	20.2	high_income	Europe
Luxembourg	8.6	high_income	Europe
Madagascar	9.2	lower_income	Africa
Malawi	10.6	lower_income	Africa
Malaysia	5.8	upper_middle	Asia
Maldives	2.8	upper_middle	Asia
Mali	8.0	lower_income	Africa
Malta	5.3	high_income	Europe
Mauritania	5.5	lower_middle	Africa
Mauritius	8.8	upper_middle	Africa
Mexico	5.3	upper_middle	North America
Micronesia (Federated States of)	29.0	lower_middle	Oceania
Mongolia	18.0	lower_middle	Asia
Montenegro	16.2	upper_middle	Europe
Morocco	7.3	lower_middle	Africa
Mozambique	23.2	lower_income	Africa
Myanmar	3.0	lower_middle	Asia
Namibia	13.5	upper_middle	Africa
Nepal	9.8	lower_middle	Asia
Netherlands (Kingdom of the)	9.3	high_income	Europe
New Zealand	10.3	high_income	Oceania
Nicaragua	4.7	lower_middle	North America
Niger	10.1	lower_income	Africa
Nigeria	6.9	lower_middle	Africa
North Macedonia	7.2	upper_middle	Europe
Norway	9.9	high_income	Europe
Oman	4.5	high_income	Asia
Pakistan	9.8	lower_middle	Asia
Panama	2.9	high_income	North America
Papua New Guinea	3.6	lower_middle	Oceania
Paraguay	6.2	upper_middle	South America
Peru	2.7	upper_middle	South America
Philippines	2.5	lower_middle	Asia
Poland	9.3	high_income	Europe
Portugal	7.2	high_income	Europe

Qatar	4.7	high_income	Asia
Republic of Korea	21.2	high_income	Asia
Republic of Moldova	12.2	upper_middle	Europe
Romania	7.3	high_income	Europe
Russian Federation	21.6	upper_middle	Asia
Rwanda	9.5	lower_income	Africa
Saint Lucia	6.9	upper_middle	North America
Saint Vincent and the Grenadines	1.0	upper_middle	North America
Samoa	14.6	lower_middle	Oceania
Sao Tome and Principe	2.2	lower_middle	Africa
Saudi Arabia	5.4	high_income	Asia
Senegal	11.0	lower_middle	Africa
Serbia	7.9	upper_middle	Europe
Seychelles	7.7	high_income	Africa
Sierra Leone	11.3	lower_income	Africa
Singapore	9.7	high_income	Asia
Slovakia	9.3	high_income	Europe
Slovenia	14.0	high_income	Europe
Solomon Islands	17.4	lower_middle	Oceania
Somalia	14.7	lower_income	Africa
South Africa	23.5	upper_middle	Africa
South Sudan	6.7	lower_income	Africa
Spain	5.3	high_income	Europe
Sri Lanka	12.9	lower_middle	Asia
Sudan	4.8	lower_income	Africa
Suriname	25.9	upper_middle	South America
Sweden	12.4	high_income	Europe
Switzerland	9.8	high_income	Europe
Syrian Arab Republic	2.1	lower_income	Asia
Turkiye	2.3	upper_middle	Asia
Tajikistan	5.3	lower_middle	Asia
Thailand	8.0	upper_middle	Asia
Timor-Leste	4.5	lower_middle	Asia

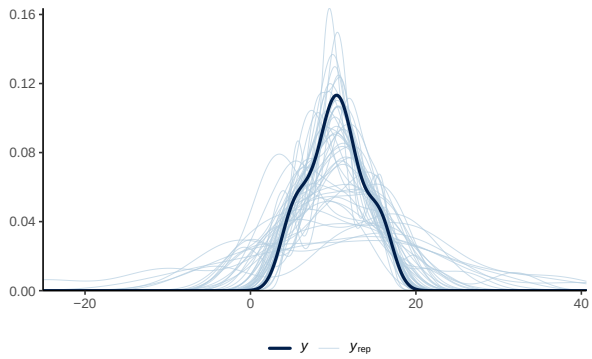
Togo	14.8	lower_income	Africa
Tonga	4.4	upper_middle	Oceania
Trinidad and Tobago	8.3	high_income	North America
Tunisia	3.2	lower_middle	Africa
Turkmenistan	6.1	upper_middle	Asia
Uganda	10.4	lower_income	Africa
Ukraine	17.7	lower_middle	Europe
United Arab Emirates	5.2	high_income	Asia
United Kingdom of Great Britain and Northern Ireland	6.9	high_income	Europe
United Republic of Tanzania	8.2	lower_middle	Africa
United States of America	14.5	high_income	North America
Uruguay	18.8	high_income	South America
Uzbekistan	8.3	lower_middle	Asia
Vanuatu	21.0	lower_middle	Oceania
Venezuela (Bolivarian Republic of)	2.1	upper_middle	South America
Viet Nam	7.2	lower_middle	Asia
Yemen	7.1	lower_income	Asia
Zambia	14.4	lower_middle	Africa
Zimbabwe	23.6	lower_middle	Africa

B Model details

B.1 Posterior predictive check

In we implement a posterior predictive check. This shows…

.



(a) Posterior prediction check

Figure 3: Examining how the model fits, and is affected by, the data

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.