University of Glasgow
Department of Statistics

Examining the Relationship of GDP on Life Expectancy: A Global Perspective

Introduction

The analysis is aimed at investigate the relationship between Life Expectancy at birth [1] and the Gross Domestic Product (GDP) of nations in the year 2021 [2], while also exploring the potential influence of continents.

Exploratory Data Analysis

Europe: Highest life expectancies and GDP.
Africa: Lowest life expectancies and GDP.
Americas, Asia, Oceania: Moderate life expectancies and GDP.

Table 1: Summary statistics on life expectancy by continent of 189 countries.

continent	Mean	Median	Std. Dev	Minimum	Maximum	IQR	Sample Size
Asia	74.07	73.47	5.77	61.98	85.47	9.13	45.00
Europe	78.24	80.11	4.42	68.85	83.99	7.66	39.00
Africa	62.86	61.65	5.62	52.53	76.38	6.09	52.00
Americas	73.21	72.83	4.18	63.19	82.66	4.95	39.00
Oceania	70.11	68.88	6.33	63.62	84.53	5.40	14.00

Table 2: Summary statistics on GDP by continent of 189 countries.

continent	Mean	Median	Std. Dev	Minimum	Maximum	IQR	Sample Size
Asia	24,687.91	14,193.12	24,807.56	1,516.31	106,032.23	33,049.55	45.00
Europe	41,075.95	38,717.69	21,844.46	12,943.61	115,683.49	23,705.53	39.00
Africa	5,737.47	3,279.17	6,000.30	705.03	28,760.52	4,494.70	52.00
Americas	21,873.94	18,512.46	17,080.51	2,870.14	80,271.13	12,546.79	39.00
Oceania	11,818.82	5,747.97	15,124.30	1,937.09	49,774.34	8,144.56	14.00

The "GDP" data exhibits skewness. To reduce this skewness and facilitate more robust analysis, a logarithmic transformation is used.

In Fig 2, life expectancy vs log(GDP) shows a strong positive linear relationship as suggested by the correlation coefficient (0.74)

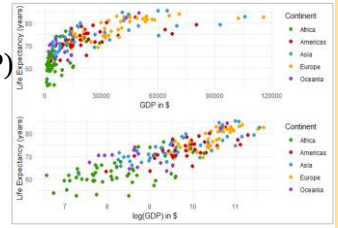

Figure 1: Relationship between Life Expectancy and GDP of countries

Formal Analysis

Model: Multiple regression with parallel slopes model

$$\text{life\_exp} = \alpha + \beta_{\text{log\_GDP}} \cdot \text{log\_GDP}_i + \beta_{\text{Americas}} \cdot \mathbb{I}_{\text{Americas}}(i) + \beta_{\text{Asia}} \cdot \mathbb{I}_{\text{Asia}}(i) + \\ + \beta_{\text{Europe}} \cdot \mathbb{I}_{\text{Europe}}(i) + \beta_{\text{Oceania}} \cdot \mathbb{I}_{\text{Oceania}}(i) + \epsilon$$

- Life expectancy as the ***outcome variable*** (life_exp)
- The log of GDP as the ***continuous explanatory variable*** (log_GDP)
- Continent as the ***categorical explanatory variable*** (continent)
- $\mathbb{I}$ is an indicator function such that

$$\mathbb{I}_{continent}(i) = \begin{cases} 1; \text{if the } i\text{th observation is in the continent} \\ 0; \text{otherwise} \end{cases}$$

- $\epsilon \sim N(0, \sigma^2)$

Table 3 : Regression Results

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	27.9228	2.835	9.851	0.000	22.330	33.515
continent[T.Americas]	3.8631	0.962	4.014	0.000	1.964	5.762
continent[T.Asia]	5.2004	0.915	5.682	0.000	3.395	7.006
continent[T.Europe]	5.7562	1.117	5.154	0.000	3.553	7.960
continent[T.Oceania]	4.6440	1.172	3.963	0.000	2.333	6.955
log_GDP	4.2461	0.338	12.549	0.000	3.579	4.914

From Table 3, we can observe that coefficients are statistically significant with p-values less than 5% and the 95% CI does not contain zero.

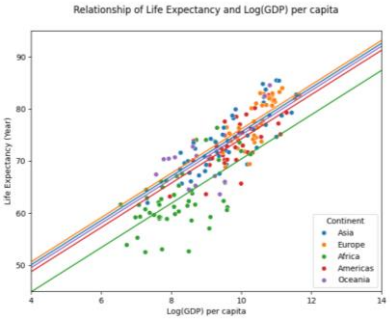

Figure 2: Life Expectancy vs log(GDP)

In this case, Africa is the baseline. Hence, the regression lines for each continent is:

Africa:
$$\widehat{\text{life\_exp}} = 27.9228 + 4.2461 \cdot \text{log\_GDP}$$

America:
$$\widehat{\text{life\_exp}} = 31.7858 + 4.2461 \cdot \text{log\_GDP}$$

Asia:
$$\widehat{\text{life\_exp}} = 33.1232 + 4.2461 \cdot \text{log\_GDP}$$

Europe:
$$\widehat{\text{life\_exp}} = 33.6790 + 4.2461 \cdot \text{log\_GDP}$$

Oceania:
$$\widehat{\text{life\_exp}} = 32.5687 + 4.2461 \cdot \text{log\_GDP}$$

Assessing Model Fit

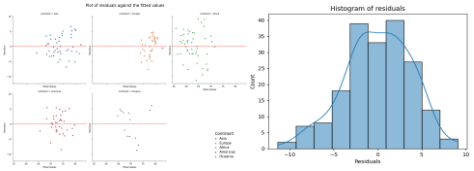

Figure 3: Plot of residuals against fitted values.

Figure 4: Residuals histogram of the whole

The residuals have zero mean and constant variance normally distributed.

Conclusions & Discussion

Table 4 : Model Comparison

Model	Log Transformation	AIC	BIC	Rsquared-adj
Parallel	yes	1050.1156	1069.5661	0.751
Simple	yes	1079.1234	1085.6069	0.7036
Interaction	yes	1051.2827	1083.7002	0.7544

In comparison to other two models, Parallel model has minimum AIC and BIC and considerably high R^2 adjusted value. **Multiple regression with parallel slopes model is the most appropriate model.**

In summary, it can be observed life expectancy has a statistically significant relationship with GDP. Furthermore, the influence of continents on life expectancy is evident.

References:
[1] Saloni Dattani, Lucas Rodés-Guirao, Hannah Ritchie, Esteban Ortiz-Ospina and Max Roser (2023) - "Life Expectancy" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/life-expectancy' [Online Resource]
[2] Max Roser, Pablo Arriagada, Joe Hasell, Hannah Ritchie and Esteban Ortiz-Ospina (2023) - "Economic Growth" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/economic-growth' [Online Resource]