Student Name

Professor

Course

Date

Correlation between HIV Mortality and GDP

Research question

Is there a correlation between HIV mortality and GDP per capita?

Introduction

HIV/AIDS was discovered in the United States in 1981, and since then, the condition has become one of the major global epidemics. It is estimated that in 2020 alone, at least 1 million people died from HIV-related complications, and 18% were young adults below 20. In 2020, it was estimated that 38 million people were living with HIV/AIDS (UNICEF).

I first learned about HIV/AIDS when I was 14 years old, and since that time, I have been very interested to know more about this disease. While I was looking at 2020 HIV deaths, I realized that in the top 10 nations with the highest HIV deaths, 8 of them are African nations. It si through this grasp, that I decided to use my mathematical prowess to investigate if there is a relatiomship between HIV deaths and GDP per capita in 29 nations with the highest HIV mortality.

Mathematical exploration

Correlation is an essential tool used in mathematics and statistics to indicate and describe the level of association/relationship between two variables. Correlation is based on strength

(strong or weak) and direction (negative or positive). A correlation is described as a strong correlation when the correlation value is >0.5; on the other, a relationship is described as a weak correlation when R-value is less than 0.5 (0.5) (Schober et al., 1768). In terms of direction, a correlation can be described as a negative correlation (-) when one variable increases while the other is decreasing. On the other hand, as both variables move in a single direction (both are increasing or decreasing), the correlation is described as a positive (+) correlation.

In this exploration, I will use both the Pearson correlation method and scatter plot method to calculate the correlation between HIV mortality and GDP per capita in 29 countries and territories.

Pearson correlation

The Pearson correlation method is also called Pearson product-moment (PPM). The formula to calculate the Pearson correlation is;

$$r = \frac{\sum (\mathrm{x} - \bar{\mathrm{x}})(\mathrm{y} - \bar{\mathrm{y}})}{\sqrt{\sum \left(x - x)^2\right)\left(y - y^2\right)}}$$

Where;

r = Pearson product-moment (PPM) (R-value)

x= dependent-variable in the dataset

$\bar{x}$= mean of x variable

y= independent-variable in the dataset

$\bar{y}$= Average of y variable

To calculate the average/mean, the following formula is used;

$$Mean = \frac{xi}{n}$$

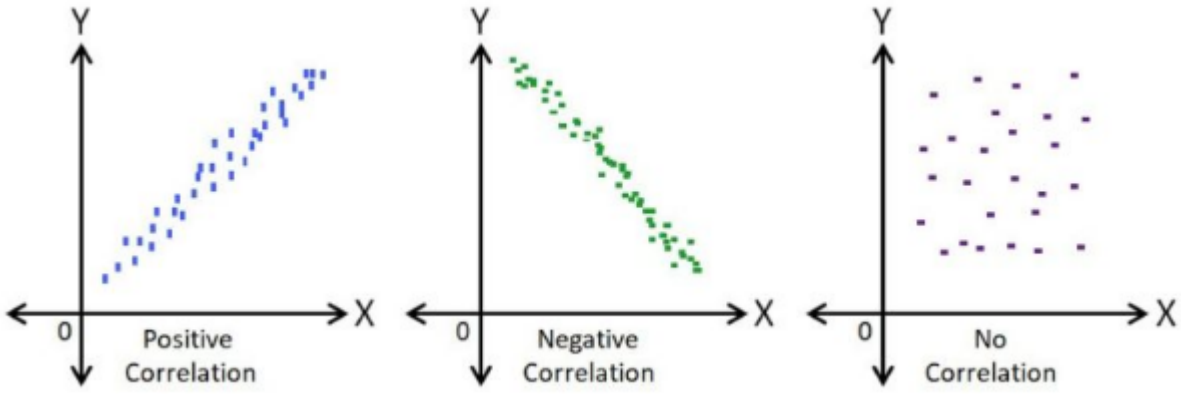
Where;

xi= sum of all the terms in the dataset

n= number of terms in the data set

Scatter plot

A Scatter plot is a graphical representation of data used to show the trend and type of the association between variables. The dependent variables appear on the (x-axis), and the independent variable appears on the (y-axis). The figure below indicates various types of scatter plots;



The line of best fit in the 1st image indicates an upward trajectory and thus meaning that there is a positive association between the two variables. The gradient of the graph will be

positive, further confirming that there is a positive correlation. The trend line in the 2nd graph indicates a decline movement and thus confirming a negative correlation between the two variables. The gradient in this graph will have a negative gradient, indicating that there is negative relationship between the two variables. The 3rd graph shows no trend line. This means there is no correlation or the correlation coefficient between two variables is (0).

Aim

This exploration aims to use mathematical tools (Pearson correlation and scatter plot) to find if there is a relationship between gross domestic product (GDP per capita) and HIV mortality/ deaths. In this exploration, I will provide GDP and HIV mortality data for 2020.

Hypothesis

In this exploration, I hypothesize that there is no correlation between gross domestic product GDP per capita and HIV mortality. The correlation coefficient (R-value) will be close to (0), indicating no correlation. The trend line in the scatter plot will indicate no direction, confirming no correlation between the two variables.

Data

The data for HIV deaths/mortality was obtained from: https://www.indexmundi.com/g/r.aspx?t=100&v=37&l=en

The data for GDP per capita for 29 countries was obtained from:

https://ourworldindata.org/grapher/gdp-per-capita-worldbank?tab=table&time=latest

Table 1: HIV mortality per country and GDP per capita

	Country	HIV deaths (000)	GDP
1	South Africa	72	12,666
2	India	69	6,166
3	Mozambique	51	1,230
4	Nigeria	45	4,917
5	Indonesia	38	11,445
6	Tanzania	27	2,625
7	Kenya	21	4,340
8	Uganda	21	2,175
9	Zimbabwe	20	3,353
10	Zambia	17	3,278
11	DRC	15	1,082
12	Ghana	14	5,446
13	Cameroon	14	3,666
14	Brazil	14	14,064
15	Thailand	14	17,285
16	Malawi	13	1,509
17	Angola	13	6,110
18	Ethiopia	12	2,297
19	South Sudan	9.1	3,114
20	Myanmar	7.7	4,875
21	Pakistan	6.8	4,563
22	Ukraine	5.9	12,376
23	Mali	5.8	2,226
24	Botswana	5.0	14,655
25	Vietnam	5.0	8,200
26	Lesotho	4.8	2,317
27	Congo	4.5	3,434
28	Colombia	4.1	13,449
29	Mexico	4.0	17,852

Pearson correlation

To compute the Pearson correlation of the two variables (HIV mortality and GDP per capita), I used the following method;

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum \left(x - x\right)^2 \left(y - y^2\right)}}$$

Where;

R= Pearson correlation coefficient

x= HIV mortality

$\bar{x}$= avergae of HIV mortality

y= GDP per capita

$\bar{y}$= mean value of GDP per capita

The first method is to compute the average for HIV deaths (yi) and GDP per capita (yi), as shown below;

$$Mean\ for\ HIV\ mortality\ (xi) = \frac{\Sigma x}{n}$$

$$Mean\ for\ GDP\ (yi) = \frac{\Sigma y}{n}$$

Table 2: Average

HIV deaths (000) (x)	GDP per capita (y)
72	12,666
69	6,166
51	1,230
45	4,917
38	11,445
27	2,625
21	4,340
21	2,175
20	3,353
17	3,278
15	1,082
14	5,446
14	3,666
14	14,064
14	17,285
13	1,509
13	6,110
12	2,297
9.1	3,114
7.7	4,875
6.8	4,563
5.9	12,376
5.8	2,226
5	14,655
5	8,200
4.8	2,317
4.5	3,434

4.1	13,449
4	17,852
Σx=552.7	Σy=190715

$$Mean\,for\,HIV\,mortality\,(xi) = \frac{552.7}{29}$$

$$Mean\,for\,HIV\,mortality\,(xi) = 19.05862$$

$$Mean\,for\,GDP\,(yi) = \frac{190715}{29}$$

$$Mean\,for\,GDP(yi) = 6576.379$$

I used the above formula to construct the Pearson correlation table 3 below;

Table 3: Pearson correlation table

HIV deaths (000) (x)	**GDP per capita (y)**	**Dx =(xi-x)**	**Dy=(yi-y)**	**dx*dy**	**dx*dx**	**dy*dy**
72	12,666	52.94138	6,089.621	322392.9	2802.79	37083483.9
69	6,166	49.94138	-410.379	-20494.9	2494.141	168410.924
51	1,230	31.94138	-5,346.379	-170771	1020.252	28583768.4
45	4,917	25.94138	-1,659.379	-43046.6	672.9552	2753538.67
38	11,445	18.94138	4,868.621	92218.4	358.7759	23703470.4
27	2,625	7.94138	-3,951.379	-31379.4	63.06552	15613396
21	4,340	1.94138	-2,236.379	-4341.66	3.768956	5001391.03

21	2,175	1.94138	-4,401.379	-8544.75	3.768956	19372137.1
20	3,353	0.94138	-3,223.379	-3034.42	0.886196	10390172.2
17	3,278	-2.05862	-3,298.379	6790.109	4.237916	10879304
15	1,082	-4.05862	-5,494.379	22299.6	16.4724	30188200.6
14	5,446	-5.05862	-1,130.379	5718.158	25.58964	1277756.68
14	3,666	-5.05862	-2,910.379	14722.5	25.58964	8470305.92
14	14,064	-5.05862	7,487.621	-37877	25.58964	56064468.2
14	17,285	-5.05862	10,708.621	-54170.8	25.58964	114674564
13	1,509	-6.05862	-5,067.379	30701.32	36.70688	25678329.9
13	6,110	-6.05862	-466.379	2825.613	36.70688	217509.372
12	2,297	-7.05862	-4,279.379	30206.51	49.82412	18313084.6
9.1	3,114	-9.95862	-3,462.379	34480.52	99.17411	11988068.3
7.7	4,875	-11.3586	-1,701.379	19325.32	129.0182	2894690.5
6.8	4,563	-12.2586	-2,013.379	24681.25	150.2738	4053695
5.9	12,376	-13.1586	5,799.621	-76315	173.1493	33635603.7
5.8	2,226	-13.2586	-4,350.379	57680.02	175.791	18925797.4
5	14,655	-14.0586	8,078.621	-113574	197.6448	65264117.3
5	8,200	-14.0586	1,623.621	-22825.9	197.6448	2636145.15
4.8	2,317	-14.2586	-4,259.379	60732.87	203.3082	18142309.5
4.5	3,434	-14.5586	-3,142.379	45748.7	211.9534	9874545.78
4.1	13,449	-14.9586	6,872.621	-102805	223.7603	47232919.4
4	17,852	-15.0586	11,275.621	-169795	226.762	127139629
19.05862	6576.379			-88451.8	9655.19	750220813

$$r=\frac{\sum\left((x-\bar{x})\right)((y-\bar{y})}{\sqrt{[\sum\left(x-x)^2\right)\left(y-y)2\right)]}}$$

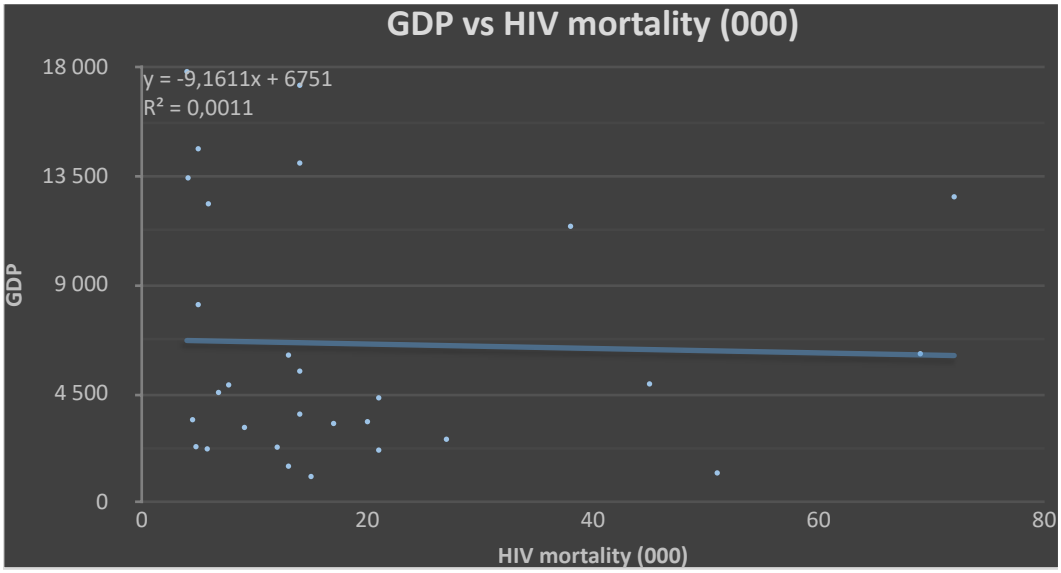
$$r=\frac{-88451.8}{\sqrt{(9655.19)(750220813)}}$$

r= -0.0329

The correlation coefficient based on the above calculation is -0.00329. This indicates that there is no relationship between HIV deaths and GDP per capita. This indicates that the GDP of the given nation does not affect HIV mortality.

Scatter plot method

Table 1 above can be represented by a graph (satter plot) below



Based on the graph above, the trend line indicates a slight negative movement. This indicates a less negative correlation between HIV mortality and GDP per capita. The GDP of a given nation/territory does not affect the mortality of HIV. The correlation coefficient of the above graph can be computed as follows;

$$R^2 = 0.0011$$

$$R = \sqrt{0.0011}$$

$R = -0.0329$

The correlation coefficient from thr image above is -0.0329. This clarifies that there is no correction between HIV mortality and GDP and thus confirms my hypothesis, which stated that "there is no association between HIV deaths and GDP per capita."

Conclusion

The primary aim of this exploration was to investigate if there is a relation between GDP per capita and HIV mortality in 29 countries. Before the investigation, it was hypothesized that "there is no association between HIV deaths and GDP per capita." While using the Pearson correlation method, the R-value was (-0.0329), indicating that there is no assocation between gross domestic product (GDP) and HIV deaths. When using the scatter plot, the R-value was (-0.0239), further indicating that there is assoctaion between the two variables. Based on this exploration, it can be concluded that the GDP of a given nation does not affect the mortality rate of the same nation.

Evaluation

The exploration was a big success as the aim and hypothesis were achieved. However, the data used in this exploration might be misleading and thus affect the final answer. The 2020 data (GDP and HIV mortality) was used in this exploration. This data might not be accurate due to COVID 19, declared a global pandemic in 2020. This might have impacted the GDP and HIV data and thus affected the final answer. In future investigations, it is imperative to use data for at least five years to increase data accuracy.

Works Cited

UNICEF. "HIV Statistics - Global and Regional Trends." *UNICEF DATA*, 2021, data.unicef.org/topic/hivaids/global-regional-trends/#:~:text=These%20hardships%20include%20prolonged%20illness. Accessed 11 July 2022.

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

Thank you for checking our sample!

Read our blog | Help Center

Start Your Order