Ruoheng Du

Regression and Multivariate Data Analysis

Prof. Jeff Simonoff

Apr. 25, 2023

Examining the Relationship Between Health Expenditure, GDP per Capita, and Continent

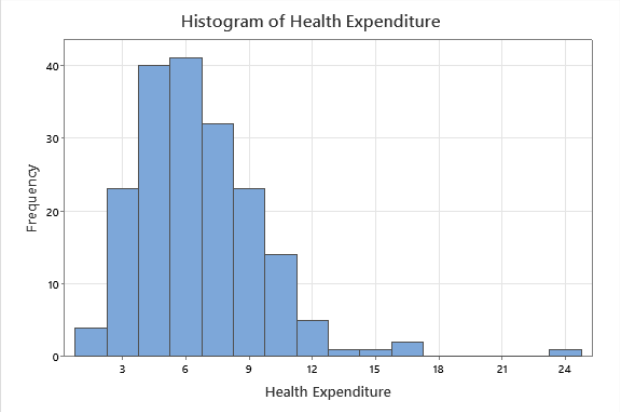
Research on health expenditure and its relationship with GDP per capita and continent is critical for evaluating a country's healthcare system and dedication to its population's well-being. The percentage of health expenditure in a country's GDP is an essential indicator that shows how much a country values its healthcare system. Considering that countries differ in terms of economic growth and geographical location, the continent-specific factor may have a role in setting health expenditure levels. Examining the relationship between health expenditure, GDP per capita, and continent can thus provide insights into the factors that influence government spending on healthcare, particularly during times of crisis, such as the COVID-19 pandemic, when the need for effective and sufficient healthcare policies has become more pressing than ever.

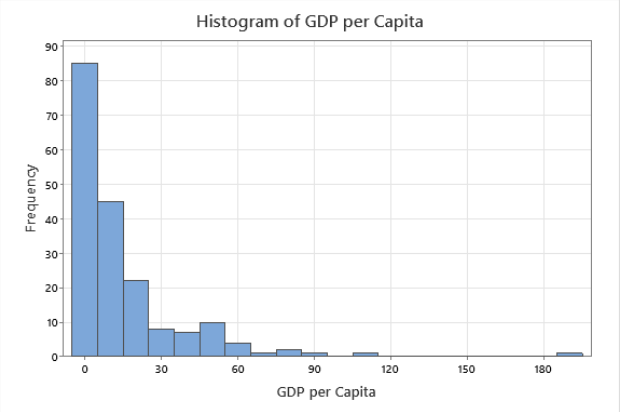
The following analysis in the report is based on data from 187 countries gathered from the UNdata (http://data.un.org/) health expenditure statistical table. This information is gathered by the United Nations Statistics Division and is included in the GlobalEdge Statistics Data Sources database (https://globaledge.msu.edu/global-resources/statistical-data-sources). The data used in this report has four important characteristics. To start with, GDP per capita (in US dollars) will be used as a measure for a country's economy in this report. GDP per capita is computed by dividing a country's total economic output by its population. The approach has significance because population size influences the level of resources and services required to

operate a functional healthcare system. Using GDP alone as a numerical predicting variable may result in skewed findings. For the sake of analysis, all GDP per capita numbers included in this report have been manually updated and expressed in thousands. Furthermore, health expenditure as a percentage of GDP is a more precise indicator than absolute expenditure on health. Relying just on the latter can be misleading, as countries with larger economies and greater GDP levels are tend to have higher levels of healthcare spending. This does not necessarily imply that these countries are prioritizing health expenditure and allocating a greater proportion of their resources to healthcare. The third characteristic of this report is the use of health expenditure and GDP per capita data from 2019, which aims to mitigate the impact of the COVID-19 pandemic. Lastly, this report specifies six major regions based on the suggested geographical regional classification provided by the United Nations (https://population.un.org/wpp/DefinitionOfRegions/): Africa, Asia, Europe, Latin America and the Caribbean, Northern America, and Oceania. The specific classification of countries within every region can be found on its website (https://unstats.un.org/unsd/methodology/m49/). Therefore, this report will use health expenditure (% of GDP) as the target variable, GDP per capita (in thousands) as the numerical predicting variable, and continent as the categorical predicting variable in the following ANCOVA model.

Here is the statistics and histograms of health expenditure and GDP per capita.

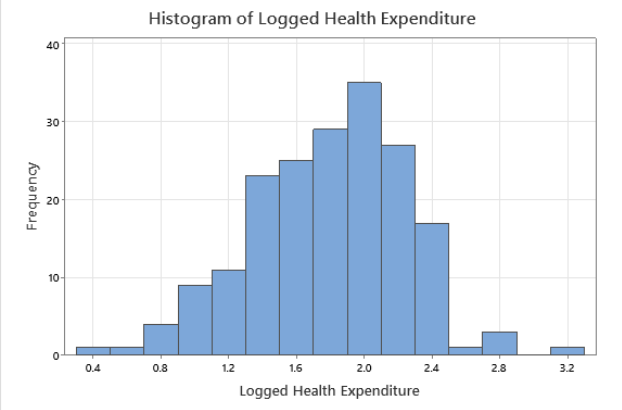
Variable	N	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Skewness
Health Expenditure	187	6.583	0.222	3.029	1.500	4.400	6.200	8.300	24.000	1.57
GDP per Capita	187	15.43	1.71	23.35	0.29	2.12	6.18	17.99	189.51	3.43

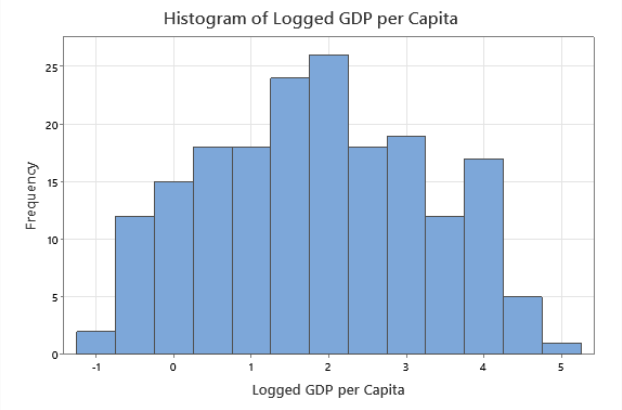




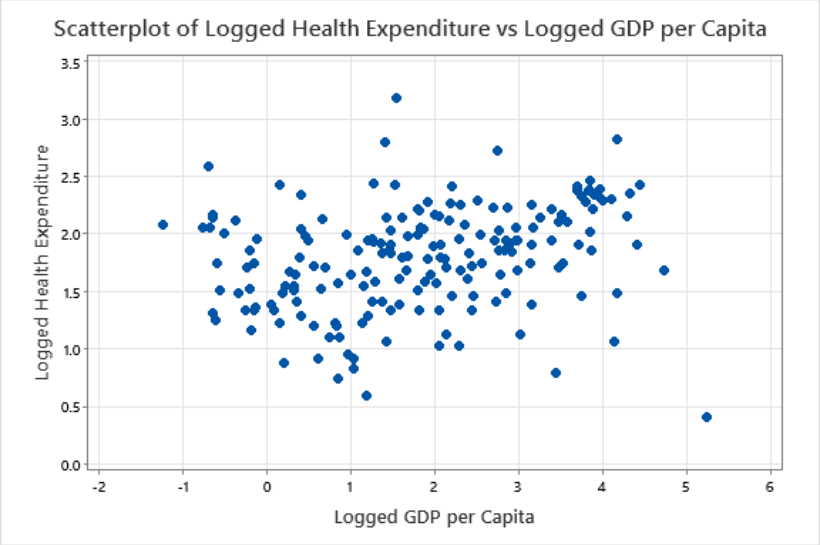
As we can see, health expenditure has a mean of 6.583 and a standard deviation of 3.029, implying that countries spend 6.583% of GDP on health on average, while GDP per capita has a mean of 15.43 and a standard deviation of 23.35, implying that the average GDP per capita for the countries listed is $15,430. In addition, the histograms and the values of skewness suggest a long right tail in these two variables, indicating the need of taking the natural logarithm.

Variable	N	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Skewness
Logged Health Expenditure	187	1.7853	0.0332	0.4535	0.4055	1.4816	1.8245	2.1163	3.1781	-0.23
Logged GDP per Capita	187	1.831	0.104	1.417	-1.231	0.749	1.821	2.890	5.244	0.03

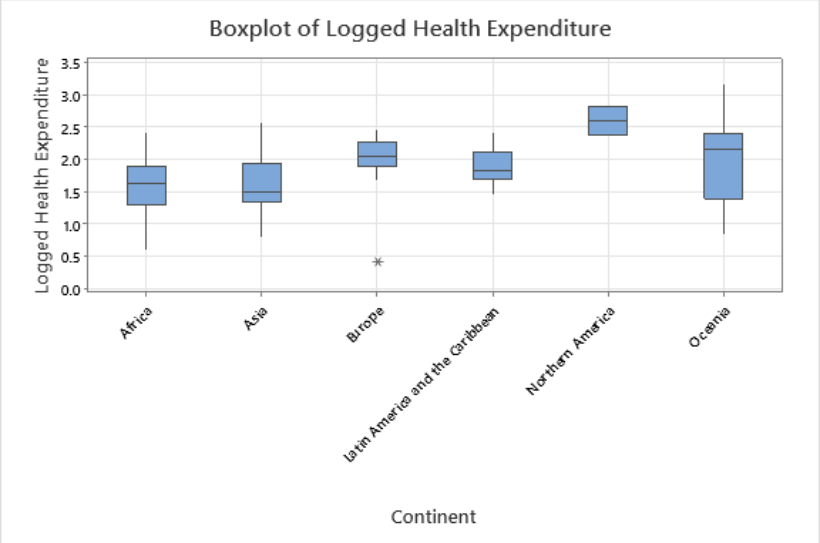




Next, in order to have a clear overview of the relationship of logged health expenditure versus logged GDP per capita and continent, let's take a look at some statistics and plots.



Based to the scatterplot, there appears to be a positive correlation between logged health expenditure and logged GDP per capita. However, the distribution of the data points on the scatterplot is relatively dispersed, suggesting that additional variables other than GDP per capita may also have an impact on health expenditure.



The boxplot emphasizes the geographical region effect in health expenditure. Oceania, for example, has larger variety in health expenditure than any other region according to its longer length of the box, while Northern America has much higher average health expenditure as a

percentage of GDP than other regions. These findings can be further supported by the descriptive statistics table presented below.

Descriptive Statistics: Logged Health Expenditure

Statistics

Variable	Continent	N	N*	Mean	SE Mean	StDev	Minimum
Logged Health Expenditure	Africa	52	0	1.5814	0.0571	0.4116	0.5878
	Asia	44	0	1.6055	0.0635	0.4215	0.7885
	Europe	41	0	2.0390	0.0530	0.3394	0.4055
	Latin America and the Caribbean	32	0	1.8841	0.0463	0.2622	1.4586
	Northern America	2	0	2.600	0.221	0.312	2.380
	Oceania	16	0	1.993	0.166	0.663	0.833

Variable	Continent	Q1	Median	Q3	Maximum	Skewness
Logged Health Expenditure	Africa	1.2878	1.6189	1.9018	2.4248	-0.22
	Asia	1.3350	1.5041	1.9459	2.5802	0.24
	Europe	1.8946	2.0541	2.2769	2.4596	-2.76
	Latin America and the Caribbean	1.6864	1.8325	2.1097	2.4069	0.24
	Northern America	*	2.600	*	2.821	*
	Oceania	1.393	2.150	2.408	3.178	-0.08

It can be inferred from the table that on average, the average and standard deviation of logged health expenditures are relatively similar in Africa and Asia, with both regions having the lowest mean values among the six regions. Northern America, on the one hand, has the highest mean value of 2.600, while on the other hand, Oceania, with the standard deviation value of 0.663, has the greatest variation in logged health expenditure.

Here's the ANCOVA output for the constant shift model.

General Linear Model: Logged Health Expenditure versus Logged GDP per Capita, Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Logged GDP per Capita	1	0.0283	0.02827	0.17	0.679
Continent	5	5.9395	1.18791	7.20	0.000
Error	180	29.6784	0.16488		
Lack-of-Fit	179	29.3325	0.16387	0.47	0.852
Pure Error	1	0.3459	0.34589		
Total	186	38.2608			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.406054	22.43%	19.85%	15.16%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.9787	0.0874	22.64	0.000	
Logged GDP per Capita	-0.0125	0.0301	-0.41	0.679	2.05
Continent					
Africa	-0.3917	0.0900	-4.35	0.000	3.00
Asia	-0.3501	0.0753	-4.65	0.000	1.92
Europe	0.1006	0.0810	1.24	0.216	2.14
Latin America and the Caribbean	-0.0692	0.0806	-0.86	0.392	1.84
Northern America	0.672	0.246	2.73	0.007	6.25

Regression Equation

Continent			
Africa	Logged Health Expenditure	=	1.5870- 0.0125 Logged GDP per Capita
Asia	Logged Health Expenditure	=	1.6287- 0.0125 Logged GDP per Capita
Europe	Logged Health Expenditure	=	2.079- 0.0125 Logged GDP per Capita
Latin America and the Caribbean	Logged Health Expenditure	=	1.9096- 0.0125 Logged GDP per Capita
Northern America	Logged Health Expenditure	=	2.650- 0.0125 Logged GDP per Capita
Oceania	Logged Health Expenditure	=	2.017- 0.0125 Logged GDP per Capita

Means

Term	Fitted Mean	SE Mean
Continent		
Africa	1.5642	0.0700
Asia	1.6058	0.0612
Europe	2.0565	0.0762
Latin America and the Caribbean	1.8867	0.0721
Northern America	2.628	0.294
Oceania	1.995	0.102

Means for Covariates

Covariate	Data Mean	StDev
Logged GDP per Capita	1.83	1.42

Here, the regression is quite weak as the R-squared value is 22.43% and adjusted R-squared value is 19.85%. The standard error of the estimate of 0.406054 says that this model could be used to predict logged health expenditure within ±0.812108, roughly 95% of the time. However, only the continent coefficient is statistically significant with a p-value of 0.000. As for logged GDP per capita coefficient, the p-value of 0.679 is greater than any reasonable significance level. This coefficient indicates that given the continent is held fixed, a 1% change in GDP per capita is associated with a -0.0125% change in health expenditure. This is beyond my expectation because I was expecting a positive relationship based on the scatterplot. As for the continent effect, as we can see from the regression equation, Africa's constant value is the lowest, and it slightly increases when it comes to Asia. Latin America and the Caribbean follows with the constant value of 1.9096. The next two regions are Oceania with the value of 2.017 and Europe with the value of 2.079. On top of these is Northern America. This trend can also be found under the means. The fitted means indicate that for a typical logged GDP per capita value of 1.83, as shown in the means for covariates, we estimate that the logged health expenditure for Asia is 1.6058. Besides, given the logged GDP per capita, the difference in fitted means between Europe and Oceania is 2.0565 – 1.995 = 0.0615, which implies that given the same logged GDP per capita, the logged health expenditure of Europe is 0.0615 higher than that of Oceania.

Given that this is a constant shift model without interaction, we can further discover the fitted means by Tukey comparisons.

Comparisons for Logged Health Expenditure

Tukey Pairwise Comparisons: Continent

Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping	
Northern America	2	2.62756	A	
Europe	41	2.05647	A	
Oceania	16	1.99465	A	
Latin America and the Caribbean	32	1.88674	A	
Asia	44	1.60585		B
Africa	52	1.56421		B

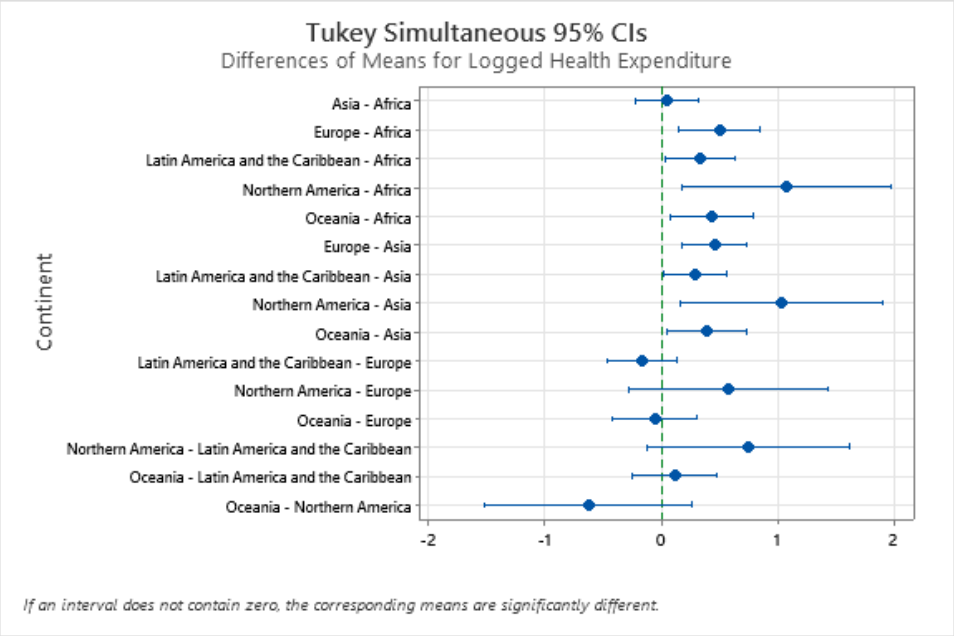
Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Continent Levels	Difference of Means	SE of Difference	Simultaneous 95% CI
Asia - Africa	0.0416	0.0933	(-0.2269, 0.3102)
Europe - Africa	0.492	0.119	(0.149, 0.835)
Latin America and the Caribbean - Africa	0.323	0.103	(0.026, 0.619)
Northern America - Africa	1.063	0.312	(0.167, 1.960)
Oceania - Africa	0.430	0.125	(0.072, 0.789)
Europe - Asia	0.4506	0.0974	(0.1703, 0.7309)
Latin America and the Caribbean - Asia	0.2809	0.0945	(0.0089, 0.5529)
Northern America - Asia	1.022	0.301	(0.157, 1.887)
Oceania - Asia	0.389	0.119	(0.048, 0.730)
Latin America and the Caribbean - Europe	-0.170	0.102	(-0.464, 0.125)
Northern America - Europe	0.571	0.295	(-0.278, 1.420)
Oceania - Europe	-0.062	0.126	(-0.424, 0.300)
Northern America - Latin America and the Caribbean	0.741	0.302	(-0.128, 1.609)
Oceania - Latin America and the Caribbean	0.108	0.124	(-0.250, 0.466)
Oceania - Northern America	-0.633	0.311	(-1.527, 0.261)

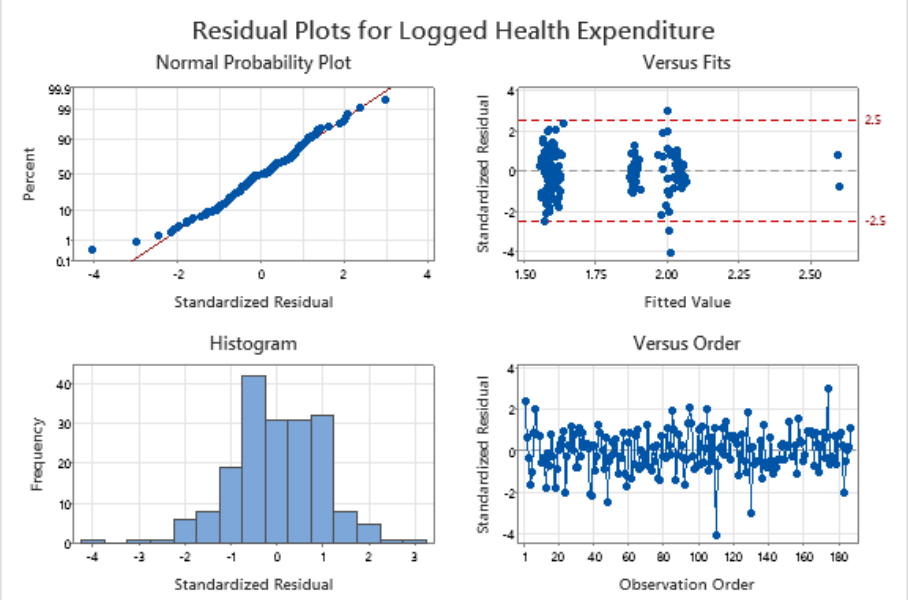
Difference of Continent Levels	T-Value	Adjusted P-Value
Asia - Africa	0.45	0.998
Europe - Africa	4.13	0.001
Latin America and the Caribbean - Africa	3.13	0.025
Northern America - Africa	3.41	0.010
Oceania - Africa	3.45	0.009
Europe - Asia	4.63	0.000
Latin America and the Caribbean - Asia	2.97	0.039
Northern America - Asia	3.40	0.011
Oceania - Asia	3.28	0.016
Latin America and the Caribbean - Europe	-1.66	0.560
Northern America - Europe	1.94	0.384
Oceania - Europe	-0.49	0.996
Northern America - Latin America and the Caribbean	2.45	0.144
Oceania - Latin America and the Caribbean	0.87	0.954
Oceania - Northern America	-2.04	0.326

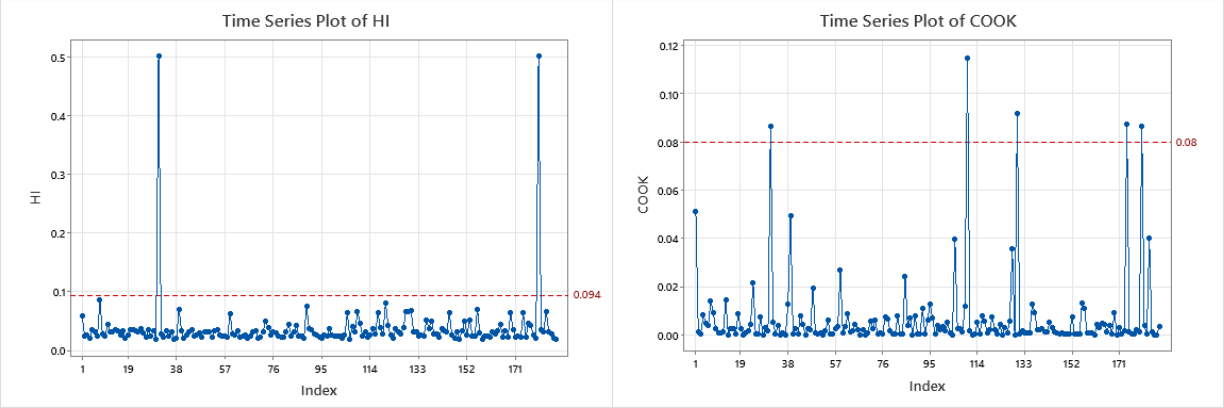
Individual confidence level = 99.55%



Based on the result above, we can find that there's no significant difference among Northern America, Europe, Latin America and the Caribbean, and Oceania, while there's no significant difference between Africa and Asia.

Now, let's take a look at the "four in one plot" and the Time Series Plot of HI and COOK.





From the normal probability plot, it is evident that some points on the lower-left side deviate significantly from the line, and in the histogram, although there is a concentration on zero, the distribution is skewed to the left, indicating non-normality. What's more, in the plot of standardized residual versus fitted values, it is obvious that there does not exhibit a lack of pattern with several points exceed the reference value of ±2.5, therefore indicating nonconstant variance and violation of this assumption.

As for the leverage values and Cook's distances, the reference line of this HI plot is $2.5 * \frac{p+1}{n} = 2.5 * \frac{6+1}{187} = 0.094$. Clearly, there are two points, the United States and Canada, significantly higher than this value. But we can't simply take these two points as the leverage points that need to be excluded. In fact, this is the result of a small sample size in the Northern America region. In the Time Series Plot of Cook's Distance, in addition to those two points, there are three points that are above my suggested value of 0.08: Monaco, Tuvalu, and Papua New Guinea.

Before continuing dealing with these unusual points, let's take a Levene's test of heteroscedasticity.

General Linear Model: absres versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Continent	5	10.24	2.0471	6.11	0.000
Error	181	60.67	0.3352		
Total	186	70.91			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.578976	14.43%	12.07%	8.98%

With the p-value of 0.000, there's sufficient evidence that there does exist nonconstant variance. Now, let's take a look at these unusual observations.

As aforementioned, one of the obviously unusual points is Monaco. It is a leverage point with an unusual GDP per capita value of $189,507, which is approximately 80.5 standard deviations higher than the average GDP per capita of all the 187 countries. This is also an outlier point because Monaco health expenditure as percentage of GDP is obviously lower than that of other countries. On the one hand, this can be linked to Monaco's small population of 0.037 million in 2019, according to the World Bank. On the other hand, its well-developed tertiary economy, particularly the gaming and finance industries, has contributed significantly to the country's economic prosperity. The second unusual point is Tuvalu. This country has a high health expenditure of 24% of GDP, which is 5.75 standard deviations higher than the global average. It is a small island country in the Pacific with limited resources and a high prevalence of noncommunicable diseases, which could be attributed to a poor diet and the effects of sea level rise (Global Nutrition Report). The country's unique environmental and sociological

characteristics, including high healthcare expenses and the need for continuing medical treatment and preventive measures, can be ascribed to the high proportion of health expenditure in GDP. Lastly, Papua New Guinea, which is also an island country in the Pacific, has a relatively low health expenditure. Compared with Tuvalu, it is a larger and more populous country with limited amount of funds available for healthcare, resulting in a lower GDP per capita. Another possible reason could be related to the political instability and corruption, which can play a role in hindering the proper allocation and utilization of funds towards healthcare. As a result, I try to remove these three countries from the data set and form a new ANCOVA model without any application to them. The following are the results.

General Linear Model: Logged Health Expenditure versus Logged GDP per Capita, Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Logged GDP per Capita	1	0.0016	0.00159	0.01	0.914
Continent	5	5.7372	1.14745	8.39	0.000
Error	177	24.2198	0.13683		
Lack-of-Fit	176	23.8739	0.13565	0.39	0.888
Pure Error	1	0.3459	0.34589		
Total	183	33.5052			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.369912	27.71%	25.26%	21.11%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.9503	0.0807	24.16	0.000	
Logged GDP per Capita	0.0030	0.0278	0.11	0.914	2.04
Continent					
Africa	-0.3702	0.0828	-4.47	0.000	2.91
Asia	-0.3503	0.0689	-5.08	0.000	1.84
Europe	0.1200	0.0740	1.62	0.107	2.01
Latin America and the Caribbean	-0.0723	0.0737	-0.98	0.328	1.76
Northern America	0.638	0.225	2.84	0.005	5.62

Regression Equation

Continent			
Africa	Logged Health Expenditure	=	1.5801 + 0.0030 Logged GDP per Capita
Asia	Logged Health Expenditure	=	1.6000 + 0.0030 Logged GDP per Capita
Europe	Logged Health Expenditure	=	2.070 + 0.0030 Logged GDP per Capita
Latin America and the Caribbean	Logged Health Expenditure	=	1.8780 + 0.0030 Logged GDP per Capita
Northern America	Logged Health Expenditure	=	2.588 + 0.0030 Logged GDP per Capita
Oceania	Logged Health Expenditure	=	1.985 + 0.0030 Logged GDP per Capita

Means

Term	Fitted Mean	SE Mean
Continent		
Africa	1.5855	0.0639
Asia	1.6054	0.0558
Europe	2.0757	0.0697
Latin America and the Caribbean	1.8834	0.0657
Northern America	2.594	0.269
Oceania	1.9906	0.0991

Means for Covariates

Covariate	Data Mean	StDev
Logged GDP per Capita	1.82	1.40

This time, the model improves a little bit compared with the previous one, but it is still quite weak as the R-squared value is 27.71% and adjusted R-squared value is 25.26%. The logged GDP per capita coefficient is still not statistically significant with the p-value of 0.914. The logged GDP per capita coefficient indicates that given the continent is held fixed, a 1% change in GDP per capita is associated with a 0.0030% change in health expenditure. This has changed from the previous negative one, which is consistent with my expectation now. Let's take a look at the Tukey comparisons for the fitted means.

Comparisons for Logged Health Expenditure

Tukey Pairwise Comparisons: Continent

Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping	
Northern America	2	2.59391	A	
Europe	40	2.07574	A	
Oceania	14	1.99057	A	
Latin America and the Caribbean	32	1.88343	A	
Asia	44	1.60542		B
Africa	52	1.58552		B

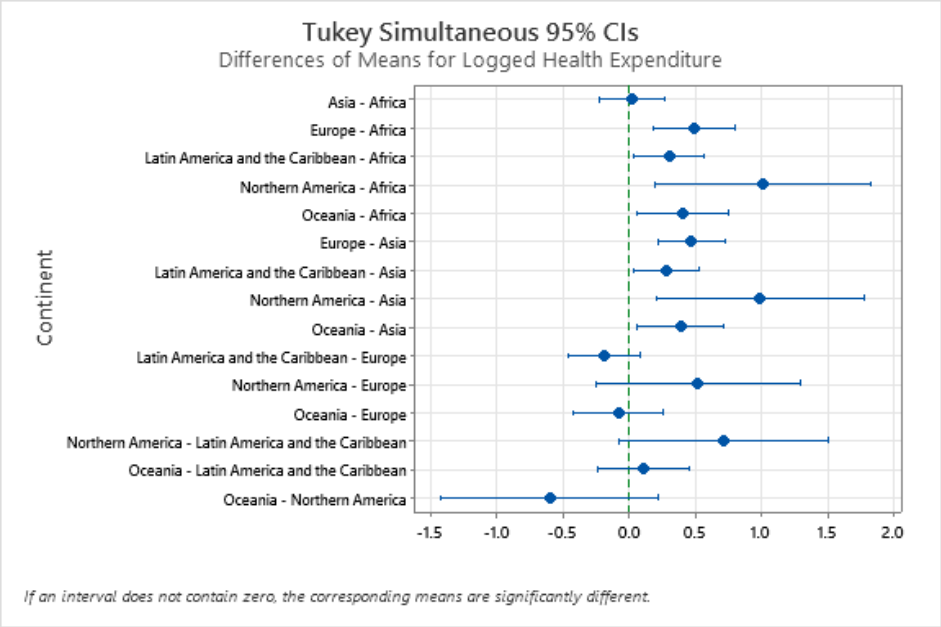
Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Continent Levels	Difference of Means	SE of Difference	Simultaneous 95% CI
Asia - Africa	0.0199	0.0853	(-0.2255, 0.2653)
Europe - Africa	0.490	0.109	(0.177, 0.803)
Latin America and the Caribbean - Africa	0.2979	0.0942	(0.0268, 0.5690)
Northern America - Africa	1.008	0.284	(0.190, 1.827)
Oceania - Africa	0.405	0.120	(0.060, 0.750)
Europe - Asia	0.4703	0.0888	(0.2146, 0.7260)
Latin America and the Caribbean - Asia	0.2780	0.0861	(0.0302, 0.5258)
Northern America - Asia	0.988	0.274	(0.200, 1.777)
Oceania - Asia	0.385	0.114	(0.058, 0.712)
Latin America and the Caribbean - Europe	-0.1923	0.0933	(-0.4608, 0.0762)
Northern America - Europe	0.518	0.269	(-0.256, 1.292)
Oceania - Europe	-0.085	0.119	(-0.428, 0.258)
Northern America - Latin America and the Caribbean	0.710	0.275	(-0.081, 1.502)
Oceania - Latin America and the Caribbean	0.107	0.119	(-0.234, 0.448)
Oceania - Northern America	-0.603	0.285	(-1.423, 0.216)

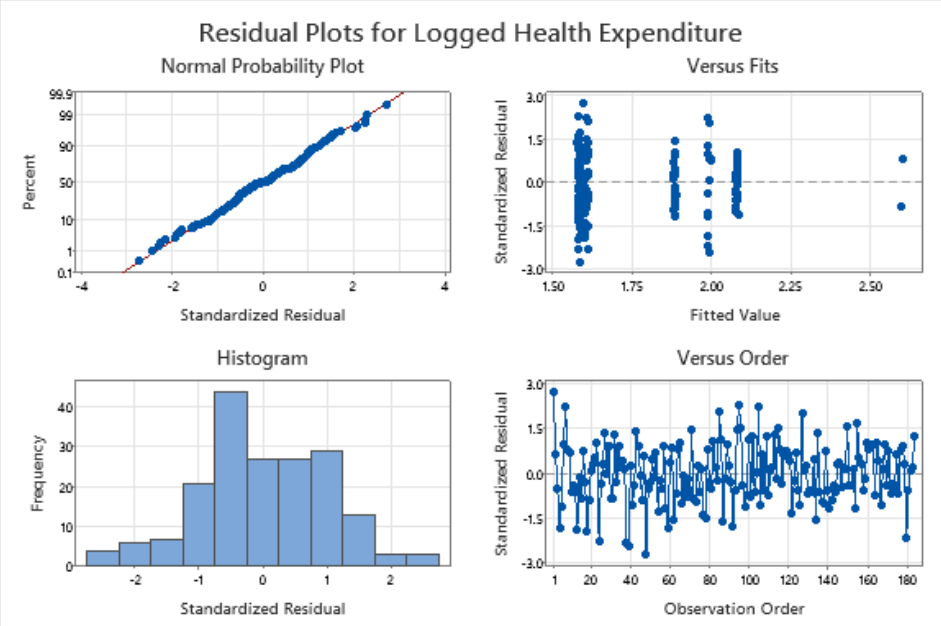
Difference of Continent Levels	T-Value	Adjusted P-Value
Asia - Africa	0.23	1.000
Europe - Africa	4.51	0.000
Latin America and the Caribbean - Africa	3.16	0.022
Northern America - Africa	3.55	0.007
Oceania - Africa	3.38	0.012
Europe - Asia	5.29	0.000
Latin America and the Caribbean - Asia	3.23	0.018
Northern America - Asia	3.61	0.005
Oceania - Asia	3.39	0.011
Latin America and the Caribbean - Europe	-2.06	0.313
Northern America - Europe	1.93	0.390
Oceania - Europe	-0.72	0.980
Northern America - Latin America and the Caribbean	2.58	0.107
Oceania - Latin America and the Caribbean	0.90	0.945
Oceania - Northern America	-2.12	0.283

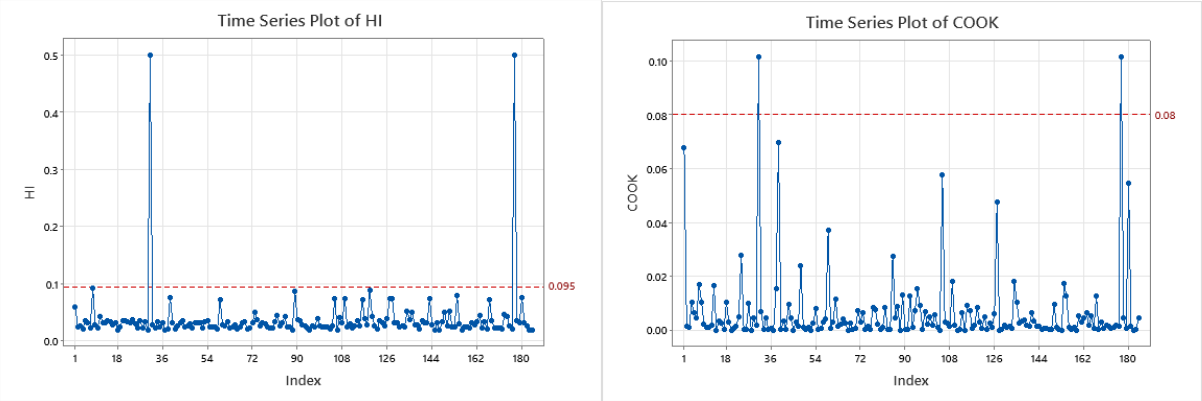
Individual confidence level = 99.55%



As we can see, the new grouping does not change compared with the previous model.

Then, let's check the assumptions.





Although the histogram is not perfect, it is apparent that the non-normality problem has been considerably addressed as almost all of the points are perfectly on the straight line. However, in the plot of standardized residual versus fitted values, nonconstant variance appears to be a problem. The new reference line of this HI plot is 0.095, and the only two points that lie above are the two Northern American countries. Meanwhile, in the Time Series Plot of Cook's Distance, we have no points other than those two countries that sit above if we use the same suggested value of 0.08 as before.

General Linear Model: absres versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Continent	5	9.490	1.8980	6.49	0.000
Error	178	52.056	0.2924		
Total	183	61.546			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.540785	15.42%	13.04%	10.27%

The result of the Levene's test is shown above. The p-value of 0.000 confirms the existence of nonconstant variance. Therefore, let's run the weighted least squares for the current ANCOVA model. The weights can be computed by using the standard deviation values shown below.

Descriptive Statistics: SRES

Statistics

Variable	Continent	N	N*	Mean	SE Mean	StDev	Minimum	Q1
SRES	Africa	52	0	0.001	0.156	1.127	-2.722	-0.800
	Asia	44	0	0.001	0.175	1.160	-2.264	-0.740
	Europe	40	0	-0.0003	0.0946	0.5984	-1.0975	-0.4968
	Latin America and the Caribbean	32	0	-0.000	0.127	0.721	-1.173	-0.550
	Northern America	2	0	0.000	0.843	1.192	-0.843	*
	Oceania	14	0	0.001	0.408	1.528	-2.427	-1.346

Variable	Continent	Median	Q3	Maximum
SRES	Africa	0.091	0.872	2.305
	Asia	-0.272	0.941	2.739
	Europe	-0.0026	0.5703	1.0356
	Latin America and the Caribbean	-0.143	0.632	1.435
	Northern America	0.000	*	0.843
	Oceania	0.439	1.046	2.252

The following are the results of WLS without Monaco, Tuvalu, and Papua New Guinea.

General Linear Model: Logged Health Expenditure versus Logged GDP per Capita, Continent

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Logged GDP per Capita	1	6.5183	18.70%	0.2007	0.2007	1.52	0.219
Continent	5	5.0254	14.41%	5.0254	1.0051	7.63	0.000
Error	177	23.3217	66.89%	23.3217	0.1318		
Lack-of-Fit	176	23.0493	66.11%	23.0493	0.1310	0.48	0.849
Pure Error	1	0.2723	0.78%	0.2723	0.2723		
Total	183	34.8653	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.362989	33.11%	30.84%	25.1573	27.84%	134.98	159.88

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	1.8875	0.0818	(1.7261, 2.0489)	23.08	0.000	
Logged GDP per Capita	0.0307	0.0249	(-0.0184, 0.0799)	1.23	0.219	2.14
Continent						
Africa	-0.3199	0.0877	(-0.4930, -0.1467)	-3.65	0.000	2.46
Asia	-0.3390	0.0793	(-0.4956, -0.1824)	-4.27	0.000	1.71
Europe	0.0945	0.0694	(-0.0424, 0.2314)	1.36	0.175	2.69
Latin America and the Caribbean	-0.0662	0.0704	(-0.2052, 0.0729)	-0.94	0.349	2.09
Northern America	0.590	0.260	(0.076, 1.104)	2.27	0.025	3.77

Regression Equation

Continent			
Africa	Logged Health Expenditure	=	1.5676 + 0.0307 Logged GDP per Capita
Asia	Logged Health Expenditure	=	1.5485 + 0.0307 Logged GDP per Capita
Europe	Logged Health Expenditure	=	1.9820 + 0.0307 Logged GDP per Capita
Latin America and the Caribbean	Logged Health Expenditure	=	1.8213 + 0.0307 Logged GDP per Capita
Northern America	Logged Health Expenditure	=	2.477 + 0.0307 Logged GDP per Capita
Oceania	Logged Health Expenditure	=	1.928 + 0.0307 Logged GDP per Capita

Means

Term	Fitted Mean	SE Mean
Continent		
Africa	1.6235	0.0662
Asia	1.6044	0.0635
Europe	2.0378	0.0483
Latin America and the Caribbean	1.8772	0.0466
Northern America	2.533	0.311
Oceania	1.984	0.148

Means for Covariates

Covariate	Data Mean	StDev
Logged GDP per Capita	1.82	1.40

Similar to the previous one, with the p-value of 0.219, the logged GDP per capita coefficient is not statistically significant. The logged GDP per capita coefficient now means that given the continent is held fixed, a 1% change in GDP per capita is associated with a 0.0307% change in health expenditure. This has increased a lot compared with the previous model without weights. The results of the Tukey comparisons are as follows.

Comparisons for Logged Health Expenditure

Tukey Pairwise Comparisons: Continent

Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping		
Northern America	2	2.53322	A	B	
Europe	40	2.03785	A		
Oceania	14	1.98401	A	B	C
Latin America and the Caribbean	32	1.87720	A		
Africa	52	1.62349		B	C
Asia	44	1.60436			C

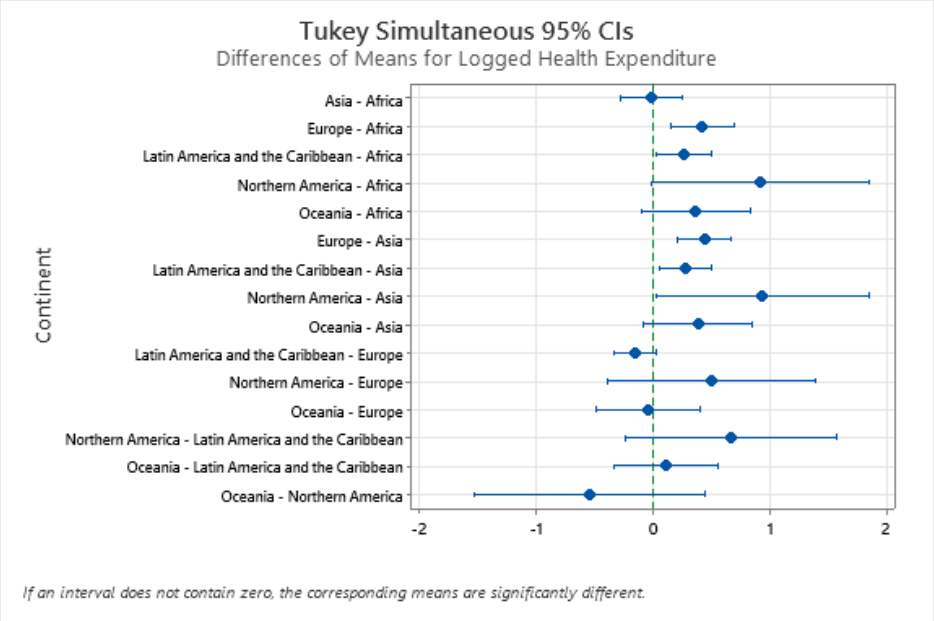
Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Continent Levels	Difference of Means	SE of Difference	Simultaneous 95% CI
Asia - Africa	-0.0191	0.0921	(-0.2841, 0.2458)
Europe - Africa	0.4144	0.0951	(0.1408, 0.6879)
Latin America and the Caribbean - Africa	0.2537	0.0833	(0.0141, 0.4933)
Northern America - Africa	0.910	0.324	(-0.021, 1.841)
Oceania - Africa	0.361	0.164	(-0.111, 0.832)
Europe - Asia	0.4335	0.0794	(0.2050, 0.6620)
Latin America and the Caribbean - Asia	0.2728	0.0787	(0.0464, 0.4993)
Northern America - Asia	0.929	0.317	(0.016, 1.841)
Oceania - Asia	0.380	0.161	(-0.085, 0.844)
Latin America and the Caribbean - Europe	-0.1606	0.0642	(-0.3455, 0.0242)
Northern America - Europe	0.495	0.309	(-0.393, 1.383)
Oceania - Europe	-0.054	0.155	(-0.499, 0.391)
Northern America - Latin America and the Caribbean	0.656	0.313	(-0.246, 1.558)
Oceania - Latin America and the Caribbean	0.107	0.155	(-0.340, 0.554)
Oceania - Northern America	-0.549	0.343	(-1.538, 0.439)

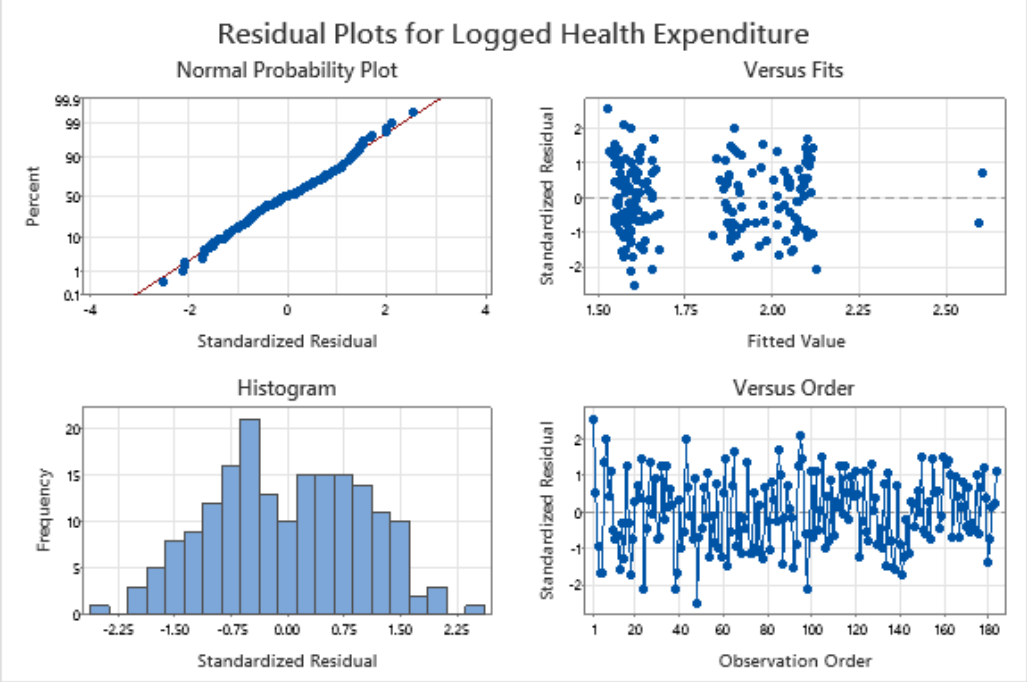
Difference of Continent Levels	T-Value	Adjusted P-Value
Asia - Africa	-0.21	1.000
Europe - Africa	4.36	0.000
Latin America and the Caribbean - Africa	3.05	0.031
Northern America - Africa	2.81	0.060
Oceania - Africa	2.20	0.242
Europe - Asia	5.46	0.000
Latin America and the Caribbean - Asia	3.47	0.009
Northern America - Asia	2.93	0.044
Oceania - Asia	2.35	0.179
Latin America and the Caribbean - Europe	-2.50	0.130
Northern America - Europe	1.61	0.596
Oceania - Europe	-0.35	0.999
Northern America - Latin America and the Caribbean	2.09	0.295
Oceania - Latin America and the Caribbean	0.69	0.983
Oceania - Northern America	-1.60	0.600

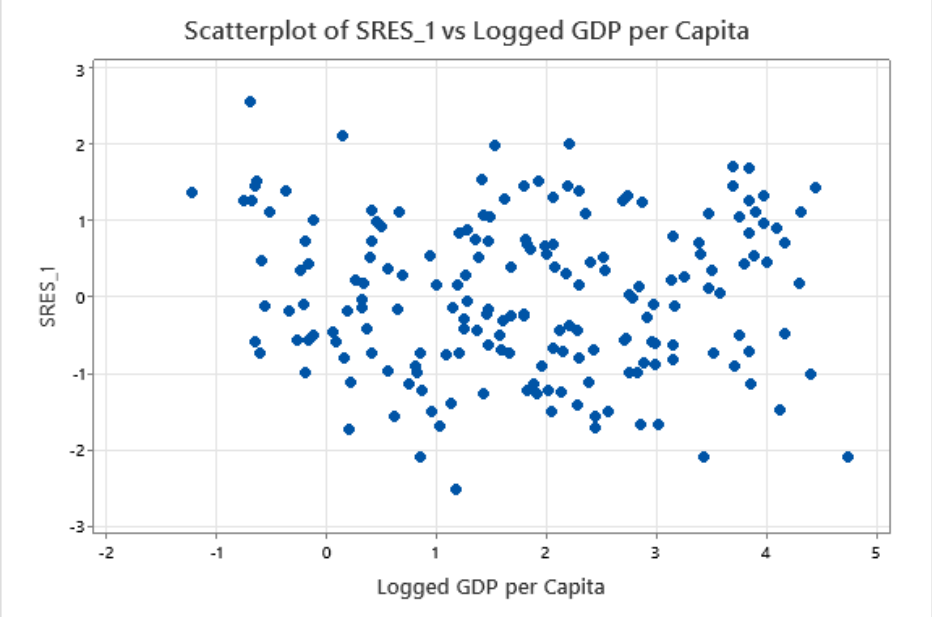
Individual confidence level = 99.55%

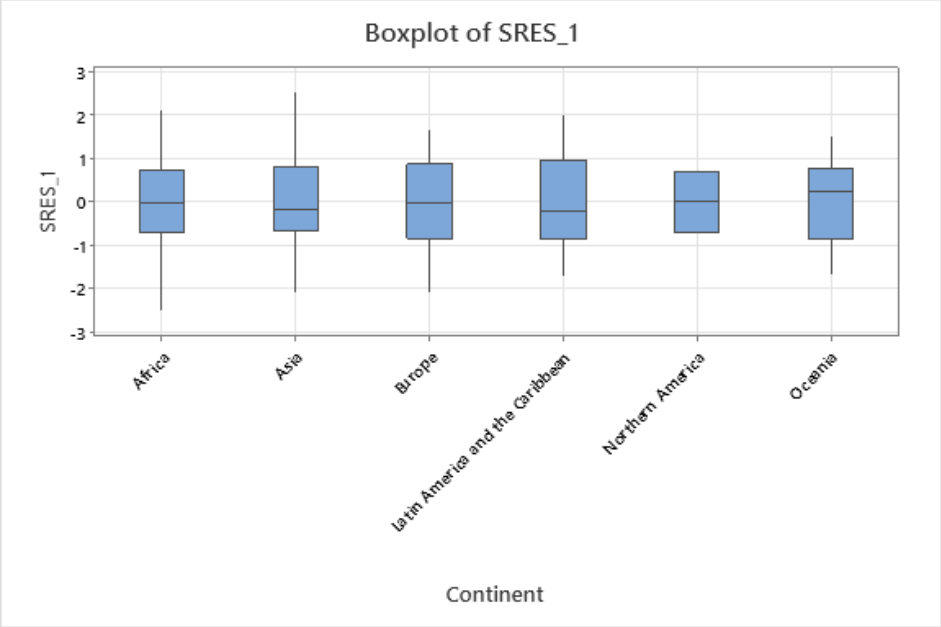


This time, with weights, the grouping has changed a lot. Northern America, Europe, Oceania, and Latin America and the Caribbean are considered to have similar fitted means. Besides, there's no significant difference among Northern America, Oceania, and Africa, and there's also no significant difference between Oceania, Africa, and Asia.

Let's take a look at the "four in one plot" and the plots of standardized residuals versus each predictor in this WLS regression. The Levene's test without weights of this model is also shown below.


Residual Plots for Logged Health Expenditure
Normal Probability Plot
Percent
Standardized Residual
Versus Fits
Standardized Residual
Fitted Value
Histogram
Frequency
Standardized Residual
Versus Order
Standardized Residual
Observation Order


Scatterplot of SRES_1 vs Logged GDP per Capita
SRES_1
Logged GDP per Capita



General Linear Model: absres_1 versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	0.0915	0.17%	0.0915	0.01831	0.06	0.997
Error	178	52.6408	99.83%	52.6408	0.29573		
Total	183	52.7324	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.543815	0.17%	0.00%	55.5500	0.00%	306.54	328.41

On the one hand, the non-normality problem is almost resolved with almost all the points sitting on the line. On the other hand, as proved by the p-value of 0.997 in the Levene's test and the lack of pattern of the points in the scatterplot, the nonconstant variance problem has been greatly addressed.

Up till now, we haven't looked at whether differing logged GDP per capita slopes for each continent would improve the model. The below are the results.

General Linear Model: Logged Health Expenditure versus Logged GDP per Capita, Continent

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Logged GDP per Capita	1	6.5183	18.70%	0.0769	0.07691	0.59	0.442
Continent	5	5.0254	14.41%	0.6228	0.12457	0.96	0.442
Logged GDP per Capita*Continent	5	1.0627	3.05%	1.0627	0.21255	1.64	0.151
Error	172	22.2589	63.84%	22.2589	0.12941		
Lack-of-Fit	171	21.9866	63.06%	21.9866	0.12858	0.47	0.853
Pure Error	1	0.2723	0.78%	0.2723	0.27233		
Total	183	34.8653	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.359739	36.16%	32.07%	*	*	137.72	177.37

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	1.00	1.22	(-1.40, 3.41)	0.82	0.411	
Logged GDP per Capita	0.235	0.305	(-0.367, 0.837)	0.77	0.442	326.80
Continent						
Africa	0.61	1.22	(-1.80, 3.02)	0.50	0.618	484.48
Asia	0.62	1.22	(-1.79, 3.03)	0.51	0.613	414.30
Europe	0.74	1.22	(-1.67, 3.16)	0.61	0.543	852.64
Latin America and the Caribbean	0.90	1.22	(-1.51, 3.32)	0.74	0.461	641.84
Northern America	-3.72	6.09	(-15.73, 8.29)	-0.61	0.542	2094.90
Logged GDP per Capita*Continent						
Africa	-0.307	0.309	(-0.916, 0.303)	-0.99	0.322	55.42
Asia	-0.244	0.308	(-0.851, 0.363)	-0.79	0.428	136.79
Europe	-0.131	0.306	(-0.736, 0.474)	-0.43	0.670	573.87
Latin America and the Caribbean	-0.247	0.310	(-0.858, 0.365)	-0.80	0.427	198.58
Northern America	1.09	1.52	(-1.90, 4.09)	0.72	0.473	969.18

Regression Equation

Continent			
Africa	Logged Health Expenditure	=	1.6137 - 0.0717 Logged GDP per Capita
Asia	Logged Health Expenditure	=	1.623 - 0.0092 Logged GDP per Capita
Europe	Logged Health Expenditure	=	1.748 + 0.1042 Logged GDP per Capita
Latin America and the Caribbean	Logged Health Expenditure	=	1.908 - 0.0116 Logged GDP per Capita
Northern America	Logged Health Expenditure	=	-2.72 + 1.33 Logged GDP per Capita
Oceania	Logged Health Expenditure	=	1.843 + 0.072 Logged GDP per Capita

The AIC_c value of this model with interaction is 137.72, which is higher than the AIC_c value of 134.98 in the previous constant shift model. Besides, the Logged GDP per capita*Continent coefficient has a p-value of 0.151, which is not statistically significant. As a result, this model with variable slopes does not outperform the preceding model with the same slope.

Finally, given that the p-values of the logged GDP per capita coefficient are all statistically insignificant in this report, I will try a one-way ANOVA model and compare it with my current constant shift ANCOVA model. Let's start from including all the 187 countries.

General Linear Model: Logged Health Expenditure versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	8.554	22.36%	8.554	1.7108	10.42	0.000
Error	181	29.707	77.64%	29.707	0.1641		
Total	186	38.261	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.405124	22.36%	20.21%	31.9823	16.41%	201.28	223.27

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	1.9506	0.0549	(1.8423, 2.0588)	35.56	0.000	
Continent						
Africa	-0.3692	0.0715	(-0.5103, -0.2281)	-5.16	0.000	1.90
Asia	-0.3451	0.0741	(-0.4913, -0.1988)	-4.65	0.000	1.87
Europe	0.0884	0.0754	(-0.0603, 0.2371)	1.17	0.242	1.86
Latin America and the Caribbean	-0.0665	0.0802	(-0.2247, 0.0917)	-0.83	0.408	1.83
Northern America	0.650	0.240	(0.176, 1.124)	2.71	0.007	5.96

Regression Equation

Logged Health Expenditure = 1.9506 - 0.3692 Continent_Africa - 0.3451 Continent_Asia
+ 0.0884 Continent_Europe - 0.0665 Continent_Latin America and the Caribbean
+ 0.650 Continent_Northern America + 0.0425 Continent_Oceania

Means

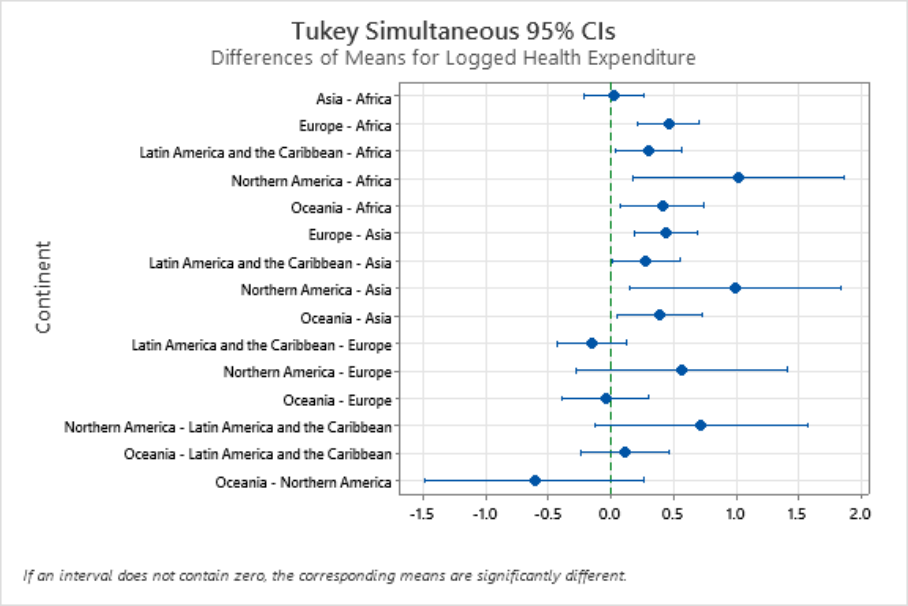
Term	Fitted Mean	SE Mean
Continent		
Africa	1.5814	0.0562
Asia	1.6055	0.0611
Europe	2.0390	0.0633
Latin America and the Caribbean	1.8841	0.0716
Northern America	2.600	0.286
Oceania	1.993	0.101

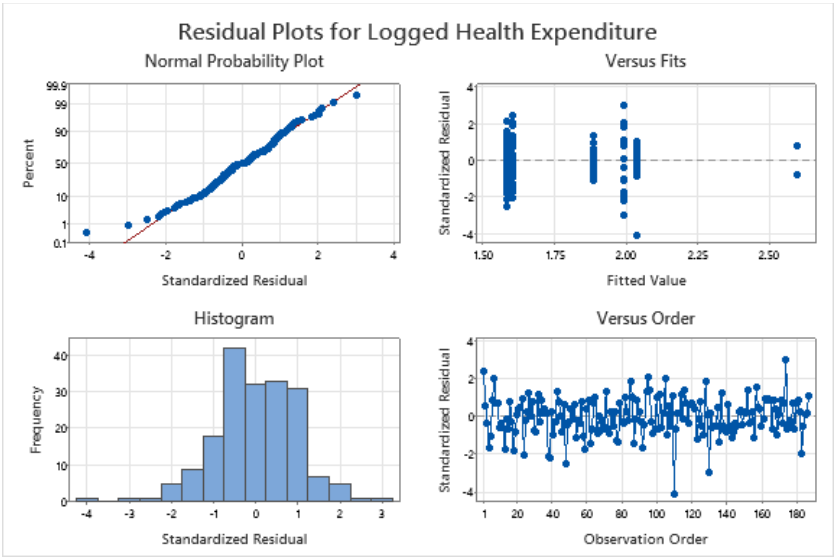
Comparisons for Logged Health Expenditure

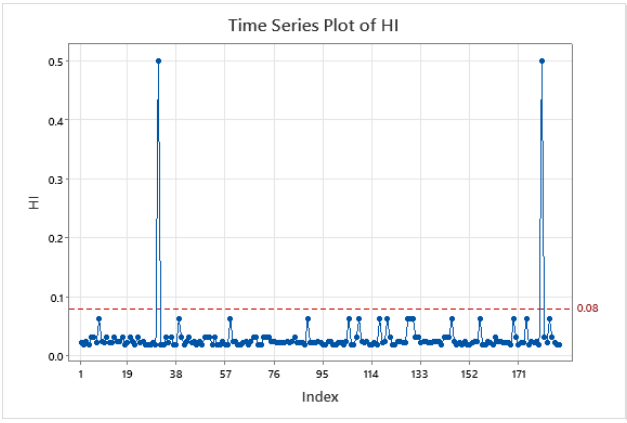
Tukey Pairwise Comparisons: Continent

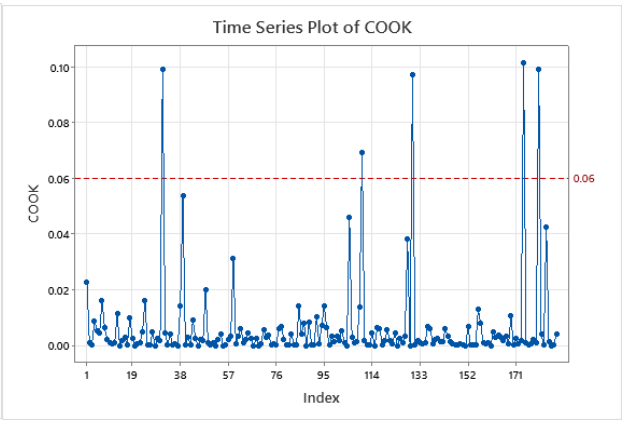
Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping	
Northern America	2	2.60046	A	
Europe	41	2.03899	A	
Oceania	16	1.99306	A	
Latin America and the Caribbean	32	1.88410	A	
Asia	44	1.60553		B
Africa	52	1.58142		B


Tukey Simultaneous 95% CIs
Differences of Means for Logged Health Expenditure
Continent
Asia - Africa
Europe - Africa
Latin America and the Caribbean - Africa
Northern America - Africa
Oceania - Africa
Europe - Asia
Latin America and the Caribbean - Asia
Northern America - Asia
Oceania - Asia
Latin America and the Caribbean - Europe
Northern America - Europe
Oceania - Europe
Northern America - Latin America and the Caribbean
Oceania - Latin America and the Caribbean
Oceania - Northern America
-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0
If an interval does not contain zero, the corresponding means are significantly different.


Residual Plots for Logged Health Expenditure
Normal Probability Plot
Percent
Standardized Residual
Versus Fits
Standardized Residual
Fitted Value
Histogram
Frequency
Standardized Residual
Versus Order
Standardized Residual
Observation Order


Time Series Plot of HI
HI
0.08
Index


Time Series Plot of COOK
COOK
0.06
Index

General Linear Model: absres versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	10.10	14.11%	10.10	2.0205	5.95	0.000
Error	181	61.51	85.89%	61.51	0.3398		
Total	186	71.61	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.582958	14.11%	11.73%	65.4386	8.62%	337.38	359.38

As we can see, there's no significant change in the R-squared value and adjusted R-squared value between this model and the constant shift ANCOVA model across 187 countries. But this time, the continent coefficient has a p-value of 0.000 and it is statistically significant. The Tukey comparisons have identical grouping, with no significant difference among Northern America, Europe, Oceania, and Latin America and the Caribbean, and no significant difference between Africa and Asia. The normal probability plot and the histogram are almost the same as before, indicating non-normality. Also, this model appears to have nonconstant variance, which can be supported by the plot of standardized residual versus fitted values and the Levene's test. In the Time Series Plot of HI, the only two points that are higher than the current reference value of 0.08 are the United States and Canada, which is the result of a small sample size. And in the Time Series Plot of COOK, except for those two countries, the other three points are Monaco, Tuvalu, and Papua New Guinea, which are the same unusual observations in the ANCOVA model. Now, let's move on to ruling out these three unusual observations.

General Linear Model: Logged Health Expenditure versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	9.284	27.71%	9.284	1.8568	13.65	0.000
Error	178	24.221	72.29%	24.221	0.1361		
Total	183	33.505	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.368883	27.71%	25.68%	26.0665	22.20%	163.71	185.58

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	1.9571	0.0503	(1.8578, 2.0564)	38.90	0.000	
Continent						
Africa	-0.3757	0.0654	(-0.5047, -0.2466)	-5.75	0.000	1.83
Asia	-0.3516	0.0678	(-0.4853, -0.2178)	-5.19	0.000	1.79
Europe	0.1227	0.0693	(-0.0140, 0.2594)	1.77	0.078	1.77
Latin America and the Caribbean	-0.0730	0.0733	(-0.2176, 0.0716)	-1.00	0.320	1.74
Northern America	0.643	0.219	(0.212, 1.075)	2.94	0.004	5.36

Regression Equation

Logged Health Expenditure = 1.9571 - 0.3757 Continent_Africa - 0.3516 Continent_Asia
+ 0.1227 Continent_Europe - 0.0730 Continent_Latin America and the Caribbean
+ 0.643 Continent_Northern America + 0.0342 Continent_Oceania

Means

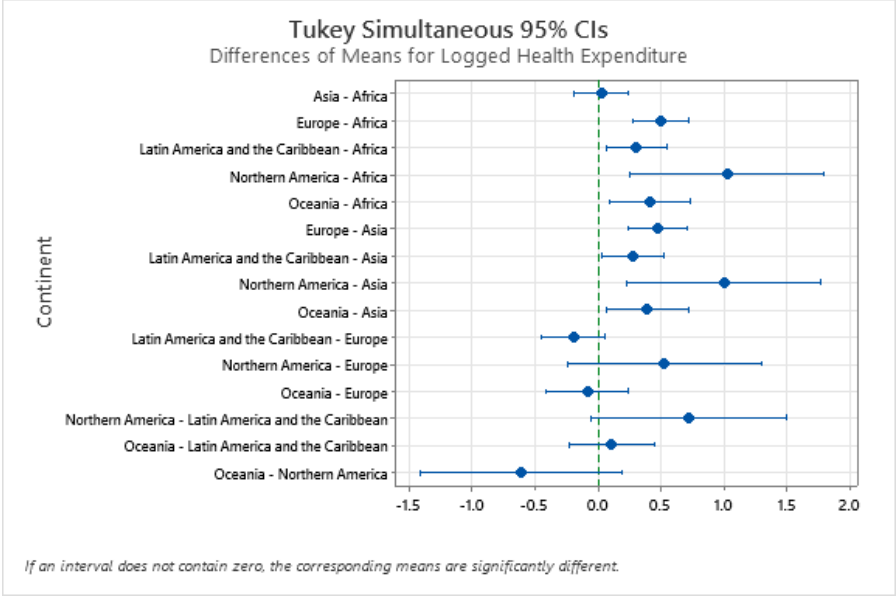
Term	Fitted Mean	SE Mean
Continent		
Africa	1.5814	0.0512
Asia	1.6055	0.0556
Europe	2.0798	0.0583
Latin America and the Caribbean	1.8841	0.0652
Northern America	2.600	0.261
Oceania	1.9913	0.0986

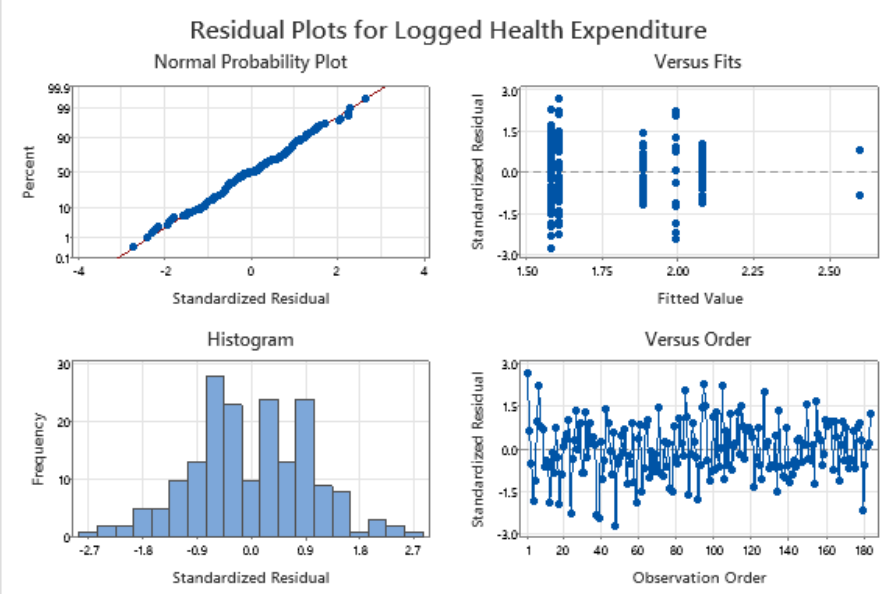
Comparisons for Logged Health Expenditure

Tukey Pairwise Comparisons: Continent

Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping
Northern America	2	2.60046	A
Europe	40	2.07983	A
Oceania	14	1.99128	A
Latin America and the Caribbean	32	1.88410	A
Asia	44	1.60553	B
Africa	52	1.58142	B


Tukey Simultaneous 95% CIs
Differences of Means for Logged Health Expenditure
Continent
Asia - Africa
Europe - Africa
Latin America and the Caribbean - Africa
Northern America - Africa
Oceania - Africa
Europe - Asia
Latin America and the Caribbean - Asia
Northern America - Asia
Oceania - Asia
Latin America and the Caribbean - Europe
Northern America - Europe
Oceania - Europe
Northern America - Latin America and the Caribbean
Oceania - Latin America and the Caribbean
Oceania - Northern America
-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0
If an interval does not contain zero, the corresponding means are significantly different.


Residual Plots for Logged Health Expenditure
Normal Probability Plot
Percent
99.9 99 90 50 10 1 0.1
-4 -2 0 2 4
Standardized Residual
Versus Fits
Standardized Residual
3.0 1.5 0.0 -1.5 -3.0
1.50 1.75 2.00 2.25 2.50
Fitted Value
Histogram
Frequency
30 20 10 0
-2.7 -1.8 -0.9 0.0 0.9 1.8 2.7
Standardized Residual
Versus Order
Standardized Residual
3.0 1.5 0.0 -1.5 -3.0
1 20 40 60 80 100 120 140 160 180
Observation Order

General Linear Model: absres versus Continent

Method

Factor coding	(-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	9.434	15.44%	9.434	1.8867	6.50	0.000
Error	178	51.658	84.56%	51.658	0.2902		
Total	183	61.092	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.538716	15.44%	13.07%	54.8076	10.29%	303.07	324.94

Removing the three unusual points does not change the one-way ANOVA model much. The p-value of the continent coefficient is still less than any reasonable significance level. The coefficient values, fitted means, and Tukey comparisons have not changed much. As for the assumptions, this time, the data of the model is more satisfied with the normal distribution. However, the problem of nonconstant variance has not been resolved. Therefore, I will use weighted least squares to try to solve this problem.

General Linear Model: Logged Health Expenditure versus Continent

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	11.31	32.50%	11.31	2.2622	17.14	0.000
Error	178	23.49	67.50%	23.49	0.1320		
Total	183	34.80	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.363274	32.50%	30.61%	25.0319	28.07%	134.39	156.25

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	1.9571	0.0595	(1.8398, 2.0745)	32.91	0.000	
Continent						
Africa	-0.3757	0.0754	(-0.5245, -0.2269)	-4.98	0.000	1.81
Asia	-0.3516	0.0788	(-0.5071, -0.1961)	-4.46	0.000	1.70
Europe	0.1227	0.0658	(-0.0072, 0.2526)	1.86	0.064	2.41
Latin America and the Caribbean	-0.0730	0.0705	(-0.2121, 0.0661)	-1.04	0.302	2.08
Northern America	0.643	0.258	(0.134, 1.153)	2.49	0.014	3.69

Regression Equation

Logged Health Expenditure = 1.9571 - 0.3757 Continent_Africa - 0.3516 Continent_Asia
+ 0.1227 Continent_Europe - 0.0730 Continent_Latin America and the Caribbean
+ 0.643 Continent_Northern America + 0.034 Continent_Oceania

Means

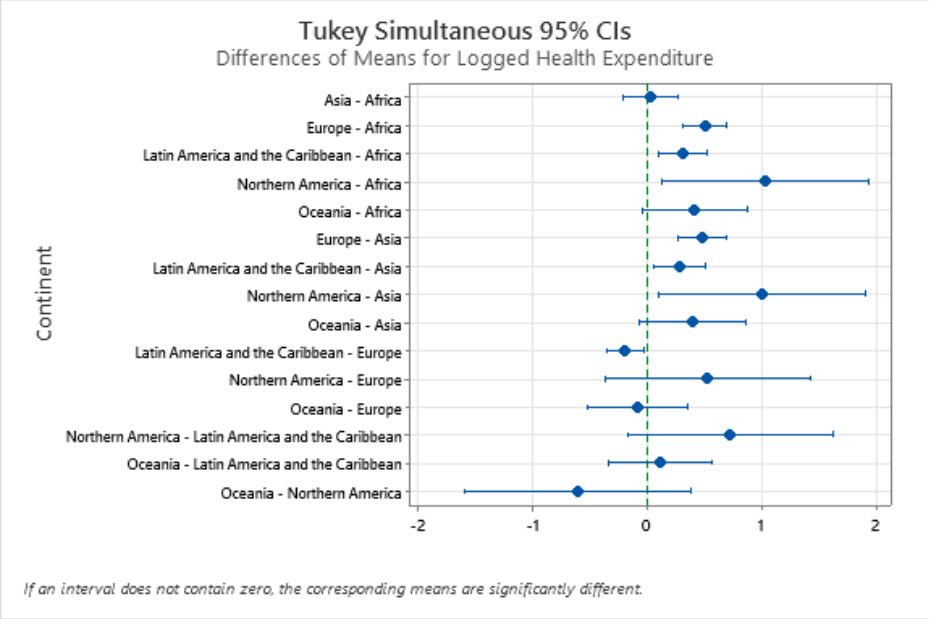
Term	Fitted Mean	SE Mean
Continent		
Africa	1.5814	0.0568
Asia	1.6055	0.0633
Europe	2.0798	0.0346
Latin America and the Caribbean	1.8841	0.0464
Northern America	2.600	0.308
Oceania	1.991	0.148

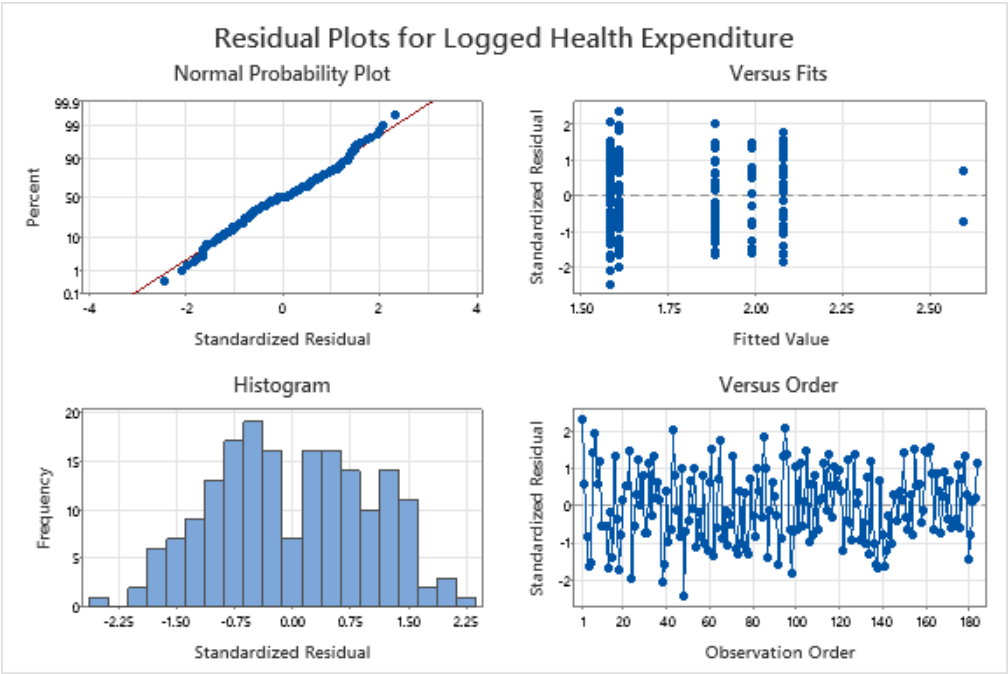
Comparisons for Logged Health Expenditure

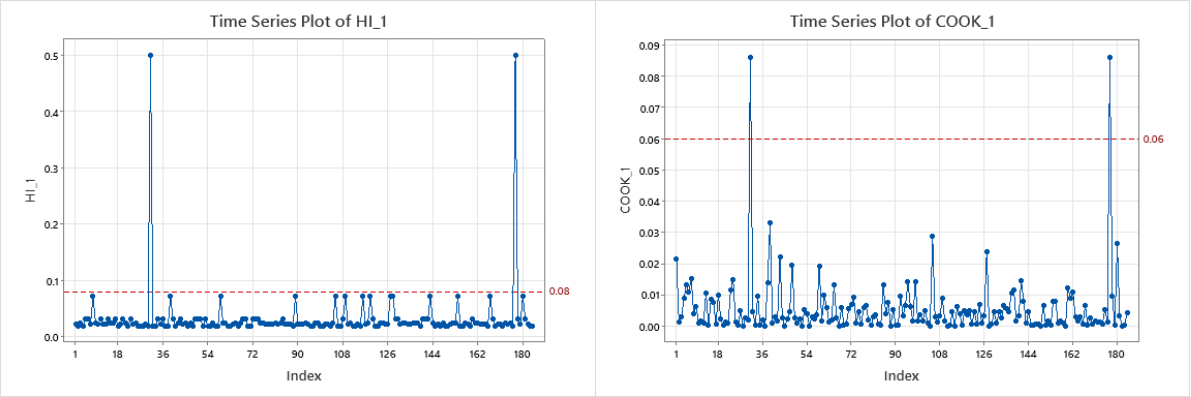
Tukey Pairwise Comparisons: Continent

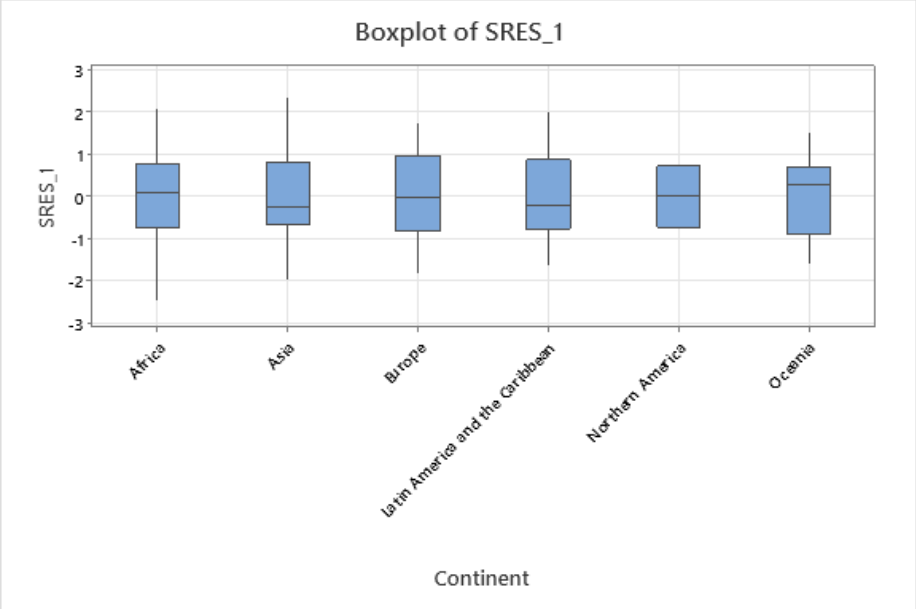
Grouping Information Using the Tukey Method and 95% Confidence

Continent	N	Mean	Grouping		
Northern America	2	2.60046	A	B	
Europe	40	2.07983	A		
Oceania	14	1.99128	A	B	C
Latin America and the Caribbean	32	1.88410		B	
Asia	44	1.60553			C
Africa	52	1.58142			C


Tukey Simultaneous 95% CIs
Differences of Means for Logged Health Expenditure
Continent
Asia - Africa
Europe - Africa
Latin America and the Caribbean - Africa
Northern America - Africa
Oceania - Africa
Europe - Asia
Latin America and the Caribbean - Asia
Northern America - Asia
Oceania - Asia
Latin America and the Caribbean - Europe
Northern America - Europe
Oceania - Europe
Northern America - Latin America and the Caribbean
Oceania - Latin America and the Caribbean
Oceania - Northern America
-2
-1
0
1
2
If an interval does not contain zero, the corresponding means are significantly different.


Residual Plots for Logged Health Expenditure
Normal Probability Plot
Percent
99.9
99
90
50
10
1
0.1
-4
-2
0
2
4
Standardized Residual
Versus Fits
Standardized Residual
2
1
0
-1
-2
1.50
1.75
2.00
2.25
2.50
Fitted Value
Histogram
Frequency
20
15
10
5
0
-2.25
-1.50
-0.75
0.00
0.75
1.50
2.25
Standardized Residual
Versus Order
Standardized Residual
2
1
0
-1
-2
1
20
40
60
80
100
120
140
160
180
Observation Order





General Linear Model: absres_1 versus Continent

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Continent	Fixed	6	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Continent	5	0.1169	0.18%	0.1169	0.02338	0.06	0.997
Error	178	64.8568	99.82%	64.8568	0.36436		
Total	183	64.9737	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.603626	0.18%	0.00%	68.3645	0.00%	321.26	343.12

The above are the results of the model from weighted least squares without Monaco, Tuvalu, and Papua New Guinea. The continent coefficient is still statistically significant and the problem of normality and constant variance have been largely resolved. Interestingly, with the AIC_c value of 134.39, across the same 184 countries, this model preforms slightly better than the constant shift ANCOVA model with weighted least squares. But the grouping results have changed a lot, as shown by the Tukey comparisons. This time, Northern America, Europe, and Oceania are considered to have similar fitted means. Also, there is no major difference between Northern America, Oceania, and Latin America and the Caribbean, as well as Oceania, Africa, and Asia.

In conclusion, the one-way ANOVA model with health expenditure as the response variable and continent as the categorical variable is the most effective when analyzing the relationship between health expenditure, GDP per capita, and continent, implying that the impact of GDP per capita appears to be minor. It can be seen that expenditure on health is an essential investment for countries to maintain a functional medical system and to guarantee the population with access to basic medical care. However, given that different geographical regions have varying levels of development and natural conditions, each country has its unique medical and health demands. As a result, the proportion of health expenditure in GDP is not only intimately tied to the geographical locations discussed in this report, but it is also influenced by a variety of potential factors that merit further investigation and analysis.

Works Cited

"Population, Total - Monaco." *The World Bank*, https://data.worldbank.org/indicator/SP.POP.TOTL?locations=MC. Accessed 26 Feb. 2023.

"Tuvalu Nutrition Profiles." *Global Nutrition Report*, https://globalnutritionreport.org/resources/nutrition-profiles/oceania/polynesia/tuvalu/. Accessed 26 Feb. 2023.