

Report on data selection

[Life](#), [Death](#)



Final Project: Statistics

Child mortality is dependent on many factors. For our analysis we will first have a look at the child mortality rates for nations across the world. We will then look at the two possible contributing factors to child mortality-1) GDP per capita of a nation and 2) health Expenditure per capita of a nation.

Hypothesis

Our assumption is that GDP per capita and Health Expenditure per capita of a nation is negatively co-related with Child Mortality rate. In simpler terms when the GDP of a country increases and the per capita health expenditure increases then child mortality rate of that country comes down.

Note: For our analysis we will use 2010 data for 165 countries. (Data source: GAPMINDER)

Child Death per ' 1000 ---- Mean= 1. 25; Std. Dev. = 1. 77

Outliers: We define outliers in this case as the observations which fall beyond 3 Std. Dev. or more away from the mean. Following are the outliers

Health Expenditure \$ per Capita ---- Mean= 1081; Std. Dev. = 1441

Outliers:

GDP Per Capita (\$) --- Mean = 7302; Std Dev. = 10, 468

Outlier

Frequency Distribution and Histogram

Child death per ' 1000

Health Expenditure \$ per Capita

GDP Per Capita

Individual data plots show that most of the countries have GDP per capita, health expenditure and child mortality rate all skewed towards the left hand side of the median.

Let's now see the relationship between variables

Scatter Plots

Health Expenditure vs. Child Mortality Rate

The scatter plots shows that there is some kind of relation between the two. One increases when the other decreases. From the trend line equation it is evident that X (health expenditure) is negatively correlated to Y (Child Mortality)

GDP per capita vs Child Mortality

Correlation Coefficients

Testing of Hypothesis

Finally let us do a testing of hypothesis using ANOVA technique to see how good the variables explain each other.

For our purpose we will test the below model

Child death per thousand = $a_1 + a_2 \cdot \text{GDP per capita} + a_3 \cdot \text{health Expenditure per capita}$

We can see from the above results the following

Regression

Only 17% of the variation of the dependent variable (Child Death) is explained by the dependent variables.(health Expenditure and GDP)

The t Statistic for GDP are 0. 8 and 1. 69 with p values 0. 419 and 0. 09 respectively so with 95% confidence level (or at p value 0. 05) we cannot

reject the null hypothesis that the variables are un correlated.

We can thus conclude that both Health Expenditure and GDP per capita are not correlated to child death rate and may be child death is influenced by some other economic or social factor.

However at 90% confidence level we can say that health expenditure is negatively correlated to child death.

Probably at lower level of GDP or health expenditure child death rate is more connected with the dependent variables selected but that needs further investigation and selective choice of data from the data set.

Works Cited :

Data for the analysis was taken from : GAPMINDER

<http://www.gapminder.org/data/>