

**ANLAYSIS THE IMPACT OF TEMPERATURE, PRECIPITATION ON INCOME GROWTH**



Student Name

**Executive Summary**

In this project, we look at how temperature and precipitation affect income levels (as defined by per capita GDP) and income growth (measured by annual growth of per capita GDP). SLR (simple linear regression), LMR (multiple linear regression analysis), and correlation analysis were the methods we used. According to the conclusions of this study, warm countries are poor (having a lower per capita GDP), whereas dry nations are rich (having a higher per capita GDP). For a better analysis, we should include more variables such as education, unemployment, population, and so on.

**Introduction**

For a sample of 106 nations, historical data on temperature, precipitation, and economic outcomes is provided. We examine the influence of temperature and precipitation on income levels (as measured by per capita GDP in natural logarithm) and, more importantly, income growth in this assignment (measured by annual growth of per capita GDP). We used MS Excel to do our analysis (Data Analysis Tool Pack).

**Analysis and Interpretation**

**2.** Graphs:

The trend line in both plots is downward, indicating that the coefficient sign of the equation is negative, indicating that per capita GDP and temperature have an indirect relationship. In comparison to hot countries, we can argue that hot countries tend to be poor (with lower per capita GDP), whereas dry nations tend to be prosperous. The R square of both the model is very similar.

**3.** Correlation Matrix:

|  |  |  |
| --- | --- | --- |
| **Correlation Matrix** | | |
|  |  |  |
|  | *per capita GDP\_2000* | *tem1950\_1960* |
| *per capita GDP\_2000* | 1 |  |
| tem1950\_1960 | -0.643 | 1 |
|  |  |  |
|  |  |  |
|  | *tem1990\_2000* | *per capita GDP\_2000* |
| tem1990\_2000 | 1 |  |
| *per capita GDP\_2000* | -0.642 | 1 |

Covariance Matrix:

|  |  |  |
| --- | --- | --- |
| **Covariance Matrix** | | |
|  |  |  |
|  | *per capita GDP\_2000* | *tem1950\_1960* |
| *per capita GDP\_2000* | 2.768 |  |
| tem1950\_1960 | -7.427 | 48.225 |
|  |  |  |
|  |  |  |
|  | *tem1990\_2000* | *per capita GDP\_2000* |
| tem1990\_2000 | 48.3 |  |
| *per capita GDP\_2000* | -7.425 | 2.768 |

The correlation of per capita GDP and temperature (1950 to 1960) is -0.643, while per capita GDP and temperature (1990 to 2000) is -0.642. So, the negative 0.64 is moderate correlation for both. Additionally, we can say that the graphs depict that negative relation or downward trendline (answer2).

Similarly, the covariance of per capita GDP and temperature (1950 to 1960) is -7.427, while per capita GDP and temperature is 1990 to 2000 is -7.425. Thus, the negative sign of the covariance indicates that the downward direction of the linear relationship between variables.

**4. (i)** In this part, we are going to run the simple linear regression, where the response variable is annual growth of per capita GDP over 1990-2000 (Y) and the exposure variable is mean temperature over 1990-2000.

Now, run the simple linear regression.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Regression Statistics***  **(Dependent Variable: per\_cap\_GDP2000)** | | | | | | |
| Multiple R | 0.3374 |  |  |  |  |  |
| R Square | 0.1139 |  |  |  |  |  |
| Adjusted R Square | 0.1053 |  |  |  |  |  |
| Standard Error | 2.1651 |  |  |  |  |  |
| Observations | 106 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **ANOVA** | | | | | | |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 62.6388 | 62.6388 | 13.3626 | 0.0004 |  |
| Residual | 104 | 487.5116 | 4.6876 |  |  |  |
| Total | 105 | 550.1504 |  |  |  |  |
|  | | | | | | |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| **Intercept** | 3.6040 | 0.6379 | 5.6499 | 0.0000 | 2.3390 | 4.8689 |
| **tem1990\_2000** | -0.1106 | 0.0303 | -3.6555 | 0.0004 | -0.1706 | -0.0506 |

This is the summary output/result of simple linear regression. The regression equation is as follows.

We see that the sign of independent variable is negative indicating that per capita GDP and temperature have an indirect relationship.

Thus, the interpretation of this regression equation is that if temperature increases by 1 degree Celsius per year, then on an average per capita GDP decreases by 0.115\*100 (=11.5).

Hypothesis Testing:

Null Hypothesis (H­­­0):

Alternative Hypothesis (H­­­1):

Model Diagnostics: The p-value of the independent variable is close to zero and the critical value is at the 5% significance level. Thus, the general rule of p-value is if p-value is less than the critical value which means null go or reject the null hypothesis. Technically, p-value is less than the significance level i.e., 0.000 < 0.05. Therefore, we significantly reject the null hypothesis and support the alternative hypothesis.

**(ii)** Similarly, we are going to run the simple linear regression, where the dependent variable is annual growth of per capita GDP over 1990-2000 (Y) and the independent variable is mean precipitation over 1990-2000 (X),

Now, we perform the simple linear regression.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Regression Statistics***  **(Dependent Variable: per\_cap\_GDP2000)** | | | | | | |
| Multiple R | 0.1956 |  |  |  |  |  |
| R Square | 0.0382 |  |  |  |  |  |
| Adjusted R Square | 0.0290 |  |  |  |  |  |
| Standard Error | 1.6471 |  |  |  |  |  |
| Observations | 106 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **ANOVA** | | | | | | |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 11.2186 | 11.2186 | 4.1351 | 0.0445 |  |
| Residual | 104 | 282.1521 | 2.7130 |  |  |  |
| Total | 105 | 293.3707 |  |  |  |  |
|  | | | | | | |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| **Intercept** | 8.2791 | 0.3171 | 26.1115 | 0.0000 | 7.6504 | 8.9079 |
| **pre1990\_2000** | -0.0475 | 0.0233 | -2.0335 | 0.0445 | -0.0938 | -0.0012 |

This is the summary output/result of simple linear regression. Based on the excel regression, the estimated regression equation is as follows.

We see that the sign of independent variable is negative which indicating that per capita GDP and precipitation have an indirect relationship. Thus, the interpretation of this regression equation is that if precipitation increases by 1 mm per year, then on an average per capita GDP decreases by 0.0475\*100 (=4.75).

Hypothesis Testing:

Null Hypothesis (H­­­0):

Alternative Hypothesis (H­­­1):

Model Diagnostics: The p-value of the independent variable is close to zero and the critical value is at the 5% significance level. Thus, the general rule of p-value is if p-value is less than the critical value which means null go or reject the null hypothesis. Technically, p-value is less than the significance level i.e., 0.04 < 0.05. Therefore, we significantly reject the null hypothesis (H­­­0) and favour the alternative hypothesis.

Goodness of fit: The R\_square of the model (i) is 11% and model (ii) is 4% (approx.). Hence, a higher r-squared indicates a better fit for the model.

**5.** Here, we use the multiple regression analysis where the response variable is of annual growth of per capita GDP over 1990-2000 (Y) and the explanatory variable is mean temperature over 1990-2000 (X1), and mean precipitation over 1990-2000 (X2).

This is the excel generated regression output.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Regression Statistics***  **(Dependent Variable: per\_cap\_GDP2000)** | | | | | | |
| Multiple R | 0.6426 |  |  |  |  |  |
| R Square | 0.4130 |  |  |  |  |  |
| Adjusted R Square | 0.4016 |  |  |  |  |  |
| Standard Error | 1.2931 |  |  |  |  |  |
| Observations | 106 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 2 | 121.1533 | 60.5766 | 36.2298 | 0.0000 |  |
| Residual | 103 | 172.2174 | 1.6720 |  |  |  |
| Total | 105 | 293.3707 |  |  |  |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| **Intercept** | 10.7520 | 0.3937 | 27.3133 | 0.0000 | 9.9713 | 11.5327 |
| **tem1990\_2000** | -0.1557 | 0.0192 | -8.1086 | 0.0000 | -0.1938 | -0.1176 |
| **pre1990\_2000** | 0.0059 | 0.0195 | 0.3036 | 0.7621 | -0.0327 | 0.0445 |

This is the summary output/result of multiple linear regression, and the estimated regression equation is as follows.

In this regression equation, we see that the coefficient sign of X1 is negative which indicates that the negative relationship between X1 and Y, whereas the coefficient sign of X2 is positive is which indicates that the relationship between X2 and Y is positive. Apart from that the interpretation part is similar as we explained above except the beta coefficient.

Hypothesis Testing (Two-tailed test):

Null Hypothesis (H­­­0):

Alternative Hypothesis (H­­­1):

The p-value is less than the significance level i.e., 0.05 > 0.00. Therefore, we significantly reject the null hypothesis (H­­­0) and favour the alternative hypothesis.

On the other hand, the overall significance test.

Null Hypothesis (H­­­0):

Alternative Hypothesis (H­­­1):

The p-value of the model is less than the significance level i.e., 0.00< 0.05. Therefore, we significantly reject the null hypothesis (H­­­0) and favour the alternative hypothesis.

**6. a) and b).** In a regression analysis experiment, a hypothesis test is required/used to define the relationship between two variables. Hypothesis testing is used to determine if the null hypothesis (no difference, no effect) may be accepted/rejected. If H0 is rejected, then the research hypothesis do not reject. If H0 do not reject, then the research hypothesis is rejected. Because the test is for effects in both directions, 2-tailed hypothesis tests are also known as nondirectional and 2-sided tests. We split the significance threshold % between both tails of the distribution when performing a two-tailed test. When a (TS)test statistic is in the crucial zone, sample data are sufficiently incompatible with the null hypothesis that the null hypothesis can be rejected for the entire population.

**7. The Goodness of fit (or coefficient of determination): R\_squared:** The R\_square of the model is 41%, which indicate that the model explains 41% variability of the response data around its mean. In other words, the R2of this regression model indicates that 41% of the variation in Y is explained by X1 and X2. Hence, R-squared reveals that 41% of the data fit the regression model. After analysing the output of the regression model, we can say that the regression model is a normal fit.

On the other hand, the Adjusted R-squared (40%) corrects the measurement based on the number of independent variables in the model. If we add more and more useless variables to a model, adjusted r-squared will decrease. If we add more useful variables, adjusted r-squared will increase. Adjusted R2 will always be less than or equal to R2.

**8.** If we add two factors i.e., population and education that influence the annual GDP per capita. The population mass of the country affects the per capita GDP because it is the total GDP divided by the population of the country. Hence, the advanced the population is lesser the per capita GDP.

On the other hand, education might be effective on GDP per capita however is the country supplying equal education opportunities for its citizens or are there regional discrepancies. According to the Ilter, 2016, there is a positive and significant correlation between per capita GDP and compulsory education years.

# **References**

Ilter, C., 2016. WHAT ECONOMIC AND SOCIAL FACTORS AFFECT GDP PER CAPITA A STUDY ON 40 COUNTRIES. *ResearchGate,* 11(2)(2017), pp. 51-62. Retrived From: DOI:10.20460/JGSM.2018.252

Marshall Burke, Noah S. Diffenbaugh, 2019. Global warming has increased global economic inequality. *PNAS (Proceedings of the National Academy of Sciences of the USA),* 116 (20)(2019), pp. 9808-9813. Retrived From: https://www.pnas.org/content/116/20/9808