Scott Doyle

The Relationship between Geography and Development: A Regression Analysis of Countries' Physical Characteristics and their Non-Zero Contributions towards Economic Advancement

I. Introduction and Literature Review

In the modern world today, much effort has been made to combat the disparities of economic development between countries. International financial institutions such as the World Bank and the International Monetary Fund, along with an assortment of non-government agencies (NGO's), are presently working across the world to research and provide remedies to persistent economic inequalities that exist from the remote mountainous villages of Nepal's Himalayas, to the pacific islands of Vanuatu, and the sandy deserts of the Algerian Sahara. Although these agencies aim to countermeasure the financial, political, and economic issues that persist in developing countries, it is important to take into account the root causes of these problems in order to create sustainable policies that contextualize a remedy based on a country's specific characteristics.

During the cold war of the 20th century, labels such as 'first world,' 'second world,' and 'third world' sprung up to describe the U.S., Western Europe, and their allies as the first world, the Soviet Union, Eastern Bloc countries, China and their allies as the second world, while the remaining neutral nations were assigned to the third world. Commentators lament that "this 1-2-3 classification is now out of date, insulting and confusing. Who is to say which part of the world is first? And how can an affluent country like Saudi Arabia, neither Western nor communist, be part of the Third World? Plus, the Soviet Union doesn't even exist anymore" (Silver, 2015).

Nowadays, labels like ‘developed’ or ‘developing’ countries are used to showcase a framework that seeks improvements to certain indicators, such as increased health outcomes, improved schooling, and having more convenient ways to access water and electricity.

In order create a benchmark for these indicators, a ranking system was created to adequately categorize universal developmental objectives that countries can strive to attain. The United Nations established the Human Development Index to provide “a summary measure of achievements in three key dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living” (United Nations Development Program, 2022). Specifically, this measurement indexes data from a 191 countries’ life expectancy at birth, expected years of schooling, mean years of schooling, and gross national income per capita to rank from ‘very high’ human development to as close to 1 as possible, and ‘low’ human development that is closest to 0. Although there are critiques as to why these components were chosen and the methods to which the data is gathered and calculated, this framework is currently the most popular and widely used by policy-makers to gage a reference point in a country’s overall development profile.

From this, it is clear that there is a unit of measurement in a county’s ability to achieve greater improvements in their respective standards of living; therefore, a more pressing question comes into play as to what specific characteristics in a country would be able to predict better outcomes for sustainable development? Moreover, if policy-makers and governments knew how much influence a respective country’s characteristic would impact their HDI ranking the most, they could create targeted policies with these references in mind to better suite their outcomes. Of course, as previously mentioned, life expectancy at birth, years of schooling, and income per capita are obviously what make up the index, but there are certainly general characteristics that impact these goals altogether. This paper seeks to demonstrate that there is a non-zero relationship between a country's geographical features and its ranking in the Human Development Index.

Notably, there is already a large literature of knowledge written on the basis that geographical characteristics may demonstrate a correlation in a country's standard of living. In 1999, Gallup et al. put this idea into the fruition by concluding that "geographical considerations should be re-introduced into the econometric and theoretical studies of cross-country economic growth, which so far have almost completely neglected geographical themes." The impetus for geographical explanations derives form the idea that a country's proximity to coastal ports, deserts, mountains, and plains may drastically influence its ability to grow crops, supply a population with necessities, and exchange in global commerce. Landlocked countries are particularly highlighted in this research by their restriction on nautical commerce and greater dependence on friendly relations with neighbors.

Furthermore, there has also been a continual effort by researchers to discuss the impacts of starting geographical conditions and how it plays a role into a country's decision-making process down the road. One article discussed the notion that "initial conditions decisively affected the timing of transitions to agriculture, and through that route they affect contemporary levels of national prosperity" (Hibbs & Olsson, 2003). This research seeks to demonstrate that the prevalence of rainfall, biodiversity, and appropriate levels of conditions that support agricultural production have the ability to 'jump-start' a country's development path by ensuring that foundational elements to human survival, such reliable sources of food, water, and shelter, are already taken care of ahead of time. This then leaves room for governments to enact stronger policies to improve human capital, such as education, creativity, and technical skills.

Most recently, research has aimed to shed light on a country's precise geographical location and how the influence of a region can inadvertently spread its wealth among partner nations. An article about this idea discusses that "the development of the economy of a region depends on the firms that decide to locate there, the infrastructure and institutions in the region, as well as its natural resources and geography. Well-connected regions have an advantage because firms and individuals located there can more easily trade goods and services" (Rossi-Hansberg, 2019). The research that is used here puts modern issues at the forefront by encompassing the effects of globalization and the spread of knowledge gained from advances in information technology. This research highlights the notion that proximity to certain 'hubs' in technology or industry can matter a lot for the wealth of a country, despite a lack of necessary geographical characteristics that would normally slow down the normal trajectory toward higher development.

Lastly, recent research also highlights the impact of natural resource endowments that countries possess and how it leads to the creation of a resource 'curse' that stifles economic growth. An article describes the curse as referring "to the paradox that countries endowed with natural resources such as oil, natural gas, minerals etc. tend to have lower economic growth and worse development outcomes than countries with fewer natural resources" (Badeed et al., 2016). In principle, natural resource assets should create large benefits for poorer economies by developing reliable income streams and financing higher levels of consumption. Unfortunately, when governments aren't set up properly withhold revenues from foreign companies that extract resource wealth, or when much of the income stream in lost through corruption, the ownership of natural resources may actually incentivize greater exploitation and hinder the development of public institutions to provide an adequate standard of living to its population.

Based on the literature, this paper seeks to showcase an econometric model that will demonstrate a non-zero relationship between geographical features of a country the HDI, along with explaining to what degree its impact the ranking certain variables may have. The rest of the paper will be organized as follows: A description of data, variables, and sources in Section II, discussion and interpretation of regression results in Section III, concluding remarks in Section IV, references in Section V, tables and figures in Section VI, and the programming output code in the appendix Section VII.

II. Data Description

The data used this model came from reliable sources that are consistent with professional economic and statistical research. These sources consisted of the United Nations Development Program, the United States Census Bureau, and the World Bank among others listed in the references section of this paper. In order to explain the endogenous variable of the Human Development Index measured in 2021, seven (7) exogenous variables were initially chosen based on review from the literature and economic theory. These variables in the model are listed and described in Table 1. There are 5 variables based on geographic categories, and 2 separate variables, GDP per capita (2018) and rank in the Corruption Perceptions Index (2021), based on economic and political factors, respectively.

In Table 2, the summary statistics presents the 4 separate quartiles of each variable, in addition to the mean and standard deviation. Based on the summary, some important observations are as follows: (1) Three variables, RESRENT, AREA, and GDPTOTAL contain means that are nearly twice as high as the median, along with fairly large standard deviations—this demonstrates a high range along with some large clusters in these points that skew the data towards one area over another, (2) the mean and median of the dependent variable, HDISCORE, are very close to one-another, along with a predictable standard deviation range—this demonstrates that the data of the dependent variable likely acts in a normalized manner, and (3) LANDLOC is a dummy variable that only acts in accordance to data points of 0 not landlocked and 1 if landlocked, so the mean of .22 shows that most countries are not landlocked.

Lastly, it is worth mentioning that additional variables could have been chosen and included in this research based on increasingly specific geographical concepts. One could argue that there may be an infinite number of geographical characteristics, based on a country's profile, that could be modeled in order to examine their relationship with the Human Development Index. The 5 geographic variables expressed in this model were chosen based on their general relevance to the literature that has been thoroughly explored in the first section of this paper, along with the impacts they are likely to hold based on economic theory. Temperature, annual rainfall, area, landlocked status, and percent of natural resource rents all resemble impactful characteristics that are worth analyzing. Additionally, only 172 out of a potential 191 countries were included in this model. Many small-island nations in the Caribbean Sea and southern Pacific Ocean, along with countries with data that is hard to access due to unstable governments, such as North Korea and Myanmar (Burma), had to be excluded from this study, even though they may have potential to play an influential role in modeling the relationship.

III. Results

<u>Model Specification 1</u>

In Figure 1, this model specification contains 6 exogenous variables and a constant term to explain its non-zero relationship with HDISCORE. The economic interpretation of this model demonstrates that the average yearly temperature, the status of being landlocked, amount of rainfall, and score on the Corruption Perceptions Index all have a significant effect on a country's score in the HDI. Figure 1 illustrates a linear regression model, with AVETEMP and RAINFALL being multiplied by one-another. The economic interpretation of these two variables is based on the interaction of the terms themselves—this means that AVETEMP will have an effect of HDISCORE that is dependent on the value of RAINFALL, and vice-versa. So, the effect of AVETEMP on HDISCORE is -0.004 times a degree change in Fahrenheit, minus 0.0002 times a mm change in RAINFALL. The justification for the multiplication of these two variables stems from the idea that the addition of mm of rain is capable of affecting the outcome of the average annual temperature of a country. More rain tends to generally ensue a temperature change in a geographical area based on precipitation levels, humidity, and levels of drought. Therefore, it is logical to conclude that the interaction between these two variables can have an effect on the outcome of an HDI Score, which this model demonstrates. To test for any multicollinearity between these variables, Figure 6 graphs the strength of their relationship—which is not at all apparent.

Additionally, if a country's rank in the Corruption Perception Index goes up by 1 point (meaning that the country is getting less corrupt), this will lead to a 0.006 increase on a country's HDI ranking, which implies that countries with less corruption tend to have a higher HDI score. If a country is considered landlocked (meaning it does not border any coastlines), then it will cause a 0.07 decrease in a country's HDI ranking, implying that countries which are not landlocked will have a higher HDI score.

Figure 1 also showcases a R-squared of .627, meaning that 62.7% of the variation in the endogenous variable of HDI can be explained by the exogenous variables of average annual temperature, amount of rainfall, landlocked status, and rank in the Corruption Perceptions Index. Generally speaking, this R-squared indicates that this combination of variables is a decent fit for the model. However, it is important to recognize that 2 variables weren't significant at a minimum of 10% level of significance, 1 variable was omitted from this model specification, and the functional form of this model needs to be explored to see if this truly encapsulates the best version of the model.

In looking at the omitted variable, GDP per capita was not included. When a regression model was run with all variables independently included, neither RESRENT nor AREA were significant at a minimum of the 10% level of significance, meaning that these two variables on their own were not statistically significant predictors of HDI score; however, it is necessary to include these in the model since the literature and economic theory has demonstrated that they have the capability of explaining a relationship, even if the results show that they were not significant estimators. GDP per capita was actually a statistically significant variable but a problem with multicollinearity came into play. Essentially, since HDI is an index created from gross national income per capita, among other factors, GDP per capita is a variable that highly correlated with gross national income per capita, since they both contain similar elements of measuring the production generated by those who work in a single country. Figure 2 graphs the relationship between HDI rank and GDP per capita, which demonstrates a highly correlated view of the data. Due to this, GDP per capita was omitted from this specification model.

Lastly, to decide the functional form of the model, analysis needed to be conducted on the data from HDISCORE to see if it acted in a normalized manner to decide if other steps must be taken to adjust the model. Moreover, a test was conducted to demonstrate how normalized this data was by looking at it from the view of its overall distribution. Figure 3 illustrates a histogram of HDISCORE, while also graphing a curve that overlays the results. This shows a mean around .71 with a fairly equal amount of data on both sides of the distribution. There might be slightly more weight on the higher end of the distribution, but this figure generally demonstrates that there is not an apparent skewedness in the endogenous variable of the model; therefore, there is not the need to take the natural log of HDISCORE. Additionally, Figure 4 also shows a plot of HDISCORE, and it is clear that the data is plotted without being skewed toward one end of the spectrum over another.

Model Specification 2

In Figure 6, this illustrates the secondary model specification containing 6 exogenous variables and a constant term to explain the non-zero relationship between countries' geographical characteristics and their HDI score. This model specification is similar to the previous one, with the same justifications of functional form and interpretations of the exogenous variables; however, rather than multiplying AVETEMP with RAINFALL to see their interaction, AVETEMP was actually squared by itself in this case.

The justification for squaring this term stems from the notion that the relationship of AVETEMP with HDISCORE will wear off after a certain point. Since temperature is a causal factor in the growth and diversity of crops and natural resources, there will be a point will temperature will grow so high that it will not have as much of an effect on the development of a country. This squared term on Figure 6 shows that there will inevitably be a negative relationship between AVETEMP and HDISCORE as the temperature increases to a larger value. This undoubtably leads to a greater explanation of the variation in the endogenous variable, because the R-squared value actually goes up to 0.677, which is .05 more than the previous model. There were some slight adjustments in the numerical value of the remaining variables compared to the previous model specification, but the relationships managed to stay consistent overall.

IV. Conclusion

This linear regression model tested to see if there was a non-zero relationship between a country's geographical features and a score on the Human Development Index. The model also sought to explain to what degree (if any) a geographical variable could have on increasing or decreasing a country's score in the Human Development Index. Although literature pointed to a variety of variables that could be used, this model slimmed down an original 7 variables to 6 in two different models, due to prevalence multicollinearity. One of the most salient observations was that there wasn't a significant effect brought from the amount natural resource rents as a percent of GDP—there is clearly still a mixed message as to the benefit or cost of natural resource wealth, and if there truly is a 'curse' at all.

Lastly, the main takeaway from this model is that the geographical factors of a country's temperature, rainfall, and status of being landlocked or not can have a small, but significant effect on a country's overall HDI score. Policy-makers can conclude from this that the effect of geography can play a role in shaping the development of their country, and that implementing policies to counteract any negative effects of these geographical characteristics can have a significant effect. Of course, the level corruption has always been something that countries must be aware of, but combining these observations with the effects of geographical characteristics can certainly have a large impact on the overall development of a county and may lead to beneficial outcomes in the future.

V. References

Works Cited:

Badeed, R.A., Lean, H.H., & Clark, J. (2016). The evolution of the natural resource curse thesis: A critical literature survey. *Resources Policy, 51*. 123-134. http://dx.doi.org/10. 1016/j. resourpol. 2016.10.015

Gallup, J.L., Sachs, J.D., & Mellinger, A. (1999). Geography and Economic Development. *Working Papers: Center for International Development at Harvard University, 1*. 1-59. https://www.hks.harvard.edu/sites/default/files/centers/cid/files /publications/faculty-working-papers/001.pdf

Hibbs, D.A. & Olsson, O. (2003). Geography, biogeography, and why some countries are rich and others are poor. *Proceedings of the National Academy of Sciences, 101*. 3715-3720. https://www.pnas.org/doi/epdf/10.1073/pnas.0305531101

Silver, N. (2015). *If you shouldn't call it the Third World, what should you call it?* National Public Radio. https://www.npr.org/sections/goatsandsoda/2015/01/04/372684438/if-you-shouldnt-call-it-the-third-world-what-should-you-call-it#:~:text=The%20First%20World%20consisted%20of,assigned%20to%20the%20Third%20World.

Rossi-Hansberg, E. (2019). Geography of Growth and Development. *Oxford Research Encyclopedia, Economics and Finance.* 1-22. DOI: 10.1093/acrefore/9780190625979.013.273

United Nations Development Program. (2022). *Technical Notes: Calculating the Human Development Indices*. Human Development Report. https://hdr.undp.org/sites/default /files/2021-22_HDR/hdr2021-22_technical_notes.pdf

Data Sources:

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M., New, M. (2003). A Comprehensive Set of High-Resolution Grids of Monthly Climate for Europe and the Globe: The Observed Record (1901-2000) and 16 Scenarios (2001-2100).

Transparency International. (2022). *Corruption Perceptions Index.* https://www.transparency.org/en/cpi/2021

United States Census Bureau. (2022). *International Database: Demographic Overview.* https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

United Nations Development Program. (2021). *Human Development Index.* https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

World Bank. (2022). *Metadata Glossary*. https://data.worldbank.org/indicator/AG.LND.PRCP.MM

VI. Tables and Figures

Table 1: Variable Definition

HDISCORE	Country Score on Human Development Index in 2022 (on scale of 0-1, greater score amounts to higher development)
AVETEMP	Average yearly temperature in degrees Fahrenheit from 1961-1990
LANLOC	Status of being landlocked (Dummy)
AREA	Geographic area measured in square km

RESRENT	Total natural resource rents as a % of GDP in 2018
RAINFALL	Average rain precipitation measured in depth of mm per year from 1962-2017
GDPTOTAL	GDP per capita in 2018 measured in millions of $US
CORRUP	Level of corruption based on the Corruption Perceptions Index in 2021 (on scale of 1-100, higher score amounts to less corruption)

Table 2: Summary Statistics

Figure 1: Linear Regression Results of Model Specification 1

Variable	**Min**	**Quartile 1**	**Median**	**Average**	**Quartile 3**	**Max**	**Standard Deviation**
HDISCORE	0.394	0.593	0.735	0.719	0.841	0.962	0.152
AVETEMP	22.37	51.057	71.6	65.433	77.607	82.922	14.966
LANLOC (Dummy)	0.000	0	0	0.2209302	0	1	0.414
AREA	298	30,989	165,507.5	740,698.95	580,987.5	16,377,742	1,932,286.895
RESRENT	0.000	0.539	2.2439526	7.003274	8.070	72.215	11.001
RAINFALL	18.000	560.5	1027	1158.343	1701	3240	798.536
GDPTOTAL	238.034	1,993.177	6,127.28	14,709.36	17,844.492	117,254.74	20,280
CORRUP	13.000	30	39	43.680	55.75	88	18.476

Model Specification 1

	HDISCORE
AVETEMP	-0.004***
	(0.001)
RAINFALL	-0.0002**
	(0.0001)
CORRUP	0.006***
	(0.0005)
LANLOC	-0.070***
	(0.020)
AREA	-0.00004
	(0.000)
RESRENT	0.0005
	(0.001)
AVETEMP:RAINFALL	0.00000**
	(0.00000)
Constant	0.784***
	(0.078)
Observations	173
R2	0.627
Adjusted R2	0.612
Residual Std. Error	0.099 (df = 165)
F Statistic	39.701*** (df = 7; 165)

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 2: Relationship between HDI Score and natural log of GDP per Capita

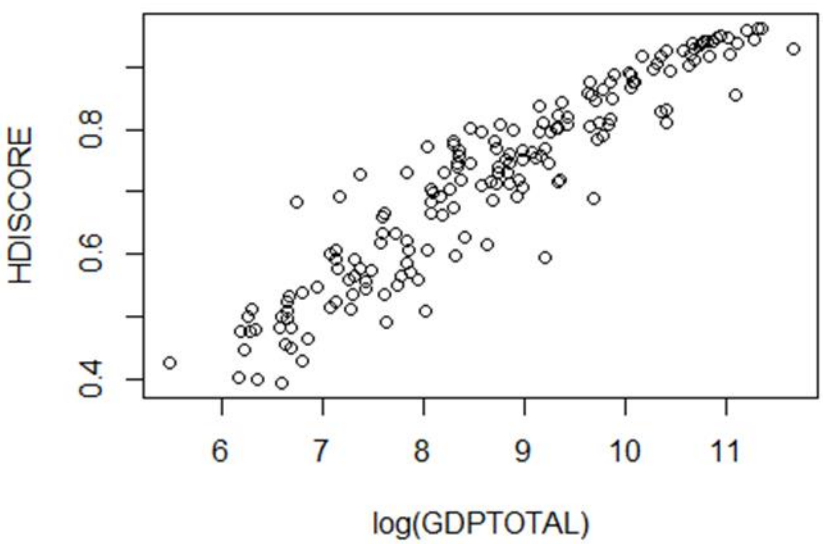

Figure 3: Frequency Distribution of HDI Score

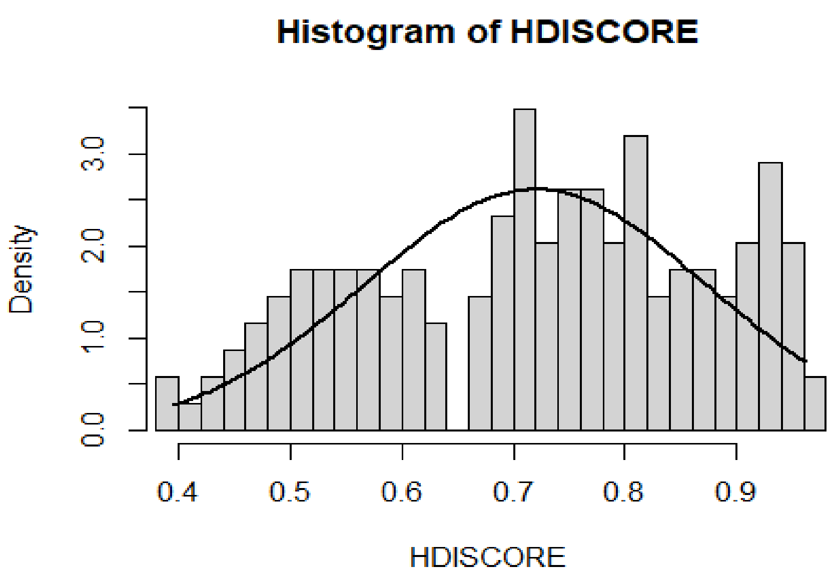

Figure 4: Plotted Data of HDI Score

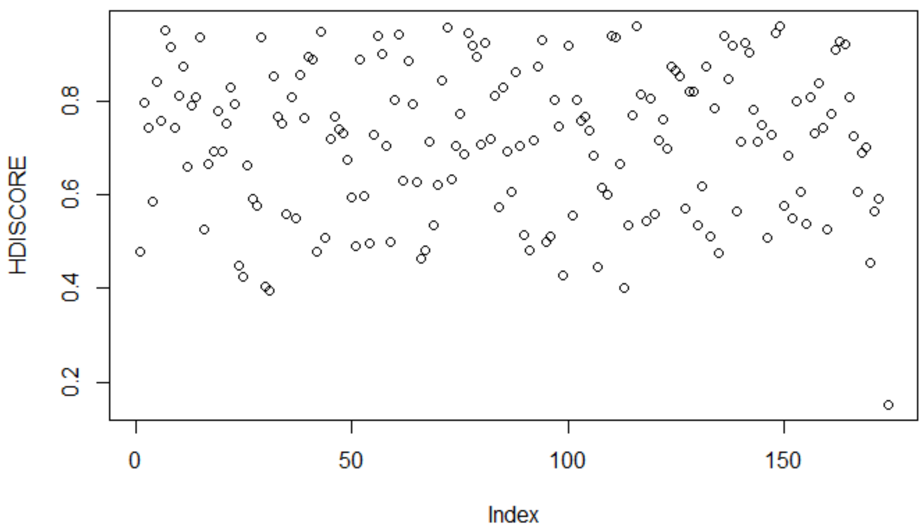

Figure 5: Relationship between Average Annual Temperature and Rainfall

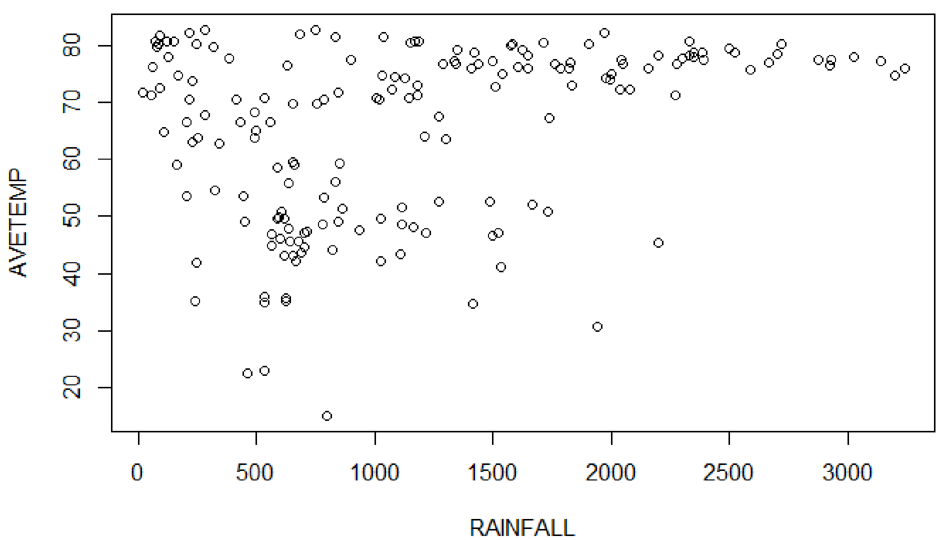

Figure 6: Linear Regression Results of Model Specification 2

	Model Specification 2

	HDISCORE

AVETEMP	0.016***
	(0.004)
I((AVETEMP)2)	-0.0002***
	(0.00003)
LANLOC	-0.088***
	(0.019)
AREA	-0.003
	(0.004)
RAINFALL	-0.00002
	(0.00001)
CORRUP	0.005***
	(0.0005)
RESRENT	0.001
	(0.001)
Constant	0.226*
	(0.121)

Observations	173
R2	0.677
Adjusted R2	0.664
Residual Std. Error	0.092 (df = 165)
F Statistic	49.467*** (df = 7; 165)
===	
Note:	*p<0.1; **p<0.05; ***p<0.01

VII. Programming Code

R Markdown

```
library(rio)
library(stargazer)

library(tidyverse)

library(MASS)

library(stargazer)

hdindic <- import("Research Paper Data1.xlsx")
view(hdindic)

HDISCORE <- hdindic$`HDI Score (2021)`
AVETEMP <- hdindic$`Average yearly temperature  (1961-1990)`
LANLOC <- hdindic$`Land Locked`
AREA <- hdindic$`Area (sq km)`
RESRENT <- hdindic$`Total Natural Resource Rents (% of GDP) in 2018`
RAINFALL <- hdindic$`Average rain precipitation in depth (mm per year) from 1
962-2017`
CORRUP <- hdindic$`Corruption Perceptions Index (2021)`
GDPTOTAL <- hdindic$`GDP per capita in 2018 (Millions of $US)`

summary(HDISCORE)

summary(AVETEMP)

summary(LANLOC)

summary(AREA)

summary(RESRENT)

summary(RAINFALL)

summary(CORRUP)

summary(GDPTOTAL)

m<-mean(HDISCORE)
std<-sqrt(var(HDISCORE))
hist(HDISCORE,prob=T, xlim = c(.394,.962), breaks = 40, freq = FALSE)
curve(dnorm(x, mean=m, sd=std), col="black", lwd=2, add=TRUE)

plot(log(GDPTOTAL),HDISCORE)
```

```
boxcox(lm(HDISCORE~1))
```

```
lAREA<-log(AREA)
lGDPTOTAL<-log(GDPTOTAL)

REG.ALL<-lm(HDISCORE~AVETEMP+LANLOC+lAREA+RESRENT+RAINFALL+CORRUP+lGDPTOTAL)
summary(REG.ALL)

##
## Call:
## lm(formula = HDISCORE ~ AVETEMP + LANLOC + lAREA + RESRENT +
##     RAINFALL + CORRUP + lGDPTOTAL)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.149751 -0.030383  0.001298  0.029325  0.121064
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.092e-01  5.440e-02   2.007   0.0463 *
## AVETEMP     -2.239e-03  3.255e-04  -6.878 1.22e-10 ***
## LANLOC      -1.635e-02  9.883e-03  -1.654   0.1000 .
## lAREA       -1.305e-03  1.777e-03  -0.734   0.4639
## RESRENT     -4.592e-04  3.527e-04  -1.302   0.1947
## RAINFALL     2.242e-06  4.986e-06   0.450   0.6536
## CORRUP      -2.662e-04  3.324e-04  -0.801   0.4244
## lGDPTOTAL    9.056e-02  4.321e-03  20.960  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04632 on 164 degrees of freedom
## Multiple R-squared:  0.9117, Adjusted R-squared:  0.9079
## F-statistic: 241.8 on 7 and 164 DF,  p-value: < 2.2e-16

REG.NOGDP<-lm(HDISCORE~AVETEMP+LANLOC+lAREA+RESRENT+RAINFALL+CORRUP)
summary(REG.NOGDP)

##
## Call:
## lm(formula = HDISCORE ~ AVETEMP + LANLOC + lAREA + RESRENT +
##     RAINFALL + CORRUP)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.175233 -0.072058  0.007871  0.069172  0.172358
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.623e-01  7.811e-02  11.039  < 2e-16 ***
## AVETEMP     -4.052e-03  6.000e-04  -6.754 2.35e-10 ***
## LANLOC      -8.605e-02  1.780e-02  -4.835 3.03e-06 ***
## lAREA       -4.514e-03  3.385e-03  -1.333    0.184
```

```
## RESRENT      7.864e-04  6.648e-04   1.183    0.239
## RAINFALL    -3.807e-06  9.518e-06  -0.400    0.690
## CORRUP       4.417e-03  4.706e-04   9.385  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08857 on 165 degrees of freedom
## Multiple R-squared:  0.6751, Adjusted R-squared:  0.6633
## F-statistic: 57.14 on 6 and 165 DF,  p-value: < 2.2e-16
```

```
REG.NOCORRUP<-lm(HDISCORE~AVETEMP+LANLOC+lAREA+RESRENT+RAINFALL+lGDPTOTAL)
summary(REG.NOCORRUP)
```

```
##
## Call:
## lm(formula = HDISCORE ~ AVETEMP + LANLOC + lAREA + RESRENT +
##     RAINFALL + lGDPTOTAL)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.142906 -0.031501  0.003134  0.029404  0.121920
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.112e-01  5.429e-02   2.048   0.0422 *
## AVETEMP     -2.188e-03  3.190e-04  -6.860 1.32e-10 ***
## LANLOC      -1.652e-02  9.870e-03  -1.673   0.0962 .
## lAREA       -1.030e-03  1.742e-03  -0.591   0.5551
## RESRENT     -3.934e-04  3.426e-04  -1.148   0.2525
## RAINFALL     1.972e-06  4.969e-06   0.397   0.6920
## lGDPTOTAL    8.824e-02  3.196e-03  27.611  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04627 on 165 degrees of freedom
## Multiple R-squared:  0.9113, Adjusted R-squared:  0.9081
## F-statistic: 282.6 on 6 and 165 DF,  p-value: < 2.2e-16
```

```
REG.GEOGONLY<-lm(HDISCORE~AVETEMP+LANLOC+lAREA+RESRENT+RAINFALL)
summary(REG.GEOGONLY)
```

```
##
## Call:
## lm(formula = HDISCORE ~ AVETEMP + LANLOC + lAREA + RESRENT +
##     RAINFALL)
##
## Residuals:
```

```
##       Min        1Q    Median        3Q       Max
## -0.206948 -0.080116  0.002448  0.079406  0.303373
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.388e+00  6.719e-02  20.661  < 2e-16 ***
## AVETEMP     -6.993e-03  6.318e-04 -11.067  < 2e-16 ***
## LANLOC      -1.352e-01  2.100e-02  -6.439 1.24e-09 ***
## lAREA       -1.532e-02  3.931e-03  -3.897 0.000141 ***
## RESRENT     -2.375e-04  8.097e-04  -0.293 0.769628
## RAINFALL    -3.334e-07  1.174e-05  -0.028 0.977385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1094 on 166 degrees of freedom
## Multiple R-squared:  0.5016, Adjusted R-squared:  0.4866
## F-statistic: 33.42 on 5 and 166 DF,  p-value: < 2.2e-16
```

```
REG.GEOGMULTI<-lm(HDISCORE~AVETEMP+((LANLOC)*(lAREA)))
summary(REG.GEOGMULTI)
```

```
REG.TRCL<-lm(HDISCORE~(AVETEMP*RAINFALL)+CORRUP+LANLOC)
summary(REG.TRCL)
```

```
##
## Call:
## lm(formula = HDISCORE ~ (AVETEMP * RAINFALL) + CORRUP + LANLOC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53789 -0.06630  0.00546  0.07364  0.19335
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        7.871e-01  7.156e-02  11.000  < 2e-16 ***
## AVETEMP           -4.224e-03  9.729e-04  -4.341 2.45e-05 ***
## RAINFALL          -1.663e-04  6.992e-05  -2.379 0.018505 *
## CORRUP             5.545e-03  4.797e-04  11.561  < 2e-16 ***
## LANLOC            -7.035e-02  1.957e-02  -3.595 0.000427 ***
## AVETEMP:RAINFALL   2.150e-06  9.418e-07   2.283 0.023712 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09811 on 167 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.6265, Adjusted R-squared:  0.6153
## F-statistic: 56.02 on 5 and 167 DF,  p-value: < 2.2e-16
```

```
stargazer(REG.TRCL, type = "text")
```

```
---------------------------------------------------------------------------

REG.TCL<-lm(HDISCORE~AVETEMP+CORRUP+LANLOC)
summary(REG.TCL)

##
## Call:
## lm(formula = HDISCORE ~ AVETEMP + CORRUP + LANLOC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55058 -0.06528  0.01126  0.07224  0.19731
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.6756169  0.0518866  13.021  < 2e-16 ***
## AVETEMP     -0.0026141  0.0005676  -4.605 8.07e-06 ***
## CORRUP       0.0052069  0.0004639  11.225  < 2e-16 ***
## LANLOC      -0.0721528  0.0195648  -3.688 0.000304 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09923 on 169 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.6133, Adjusted R-squared:  0.6064
## F-statistic: 89.33 on 3 and 169 DF,  p-value: < 2.2e-16

stargazer(REG.TCL, type = "text")

plot(RAINFALL,AVETEMP)
```