

1. Choose one continuous variable in your data as the variable of interest (it can be the same one you used for assignment 1). Choose/create a dichotomous variable in the data that separates the cases into two groups.

Write a research hypothesis that examines the differences in the mean for the two groups. Explain why you think the variable is a good way to explain the differences you see in the dependent variable. Then, use the independent sample t-test to analyze whether the variable of interest exhibits a statistically significant difference between the two groups. Describe the results by using the output from Excel. Was there significance? Why and why not? If not, what variables would you try next?

Hypothesis: The research hypothesis is to investigate whether countries with higher CO₂ emissions have a significantly different Inequality-Adjusted Human Development Index (IHDI) compared to countries with lower emissions. To this end, the annual CO₂ emissions for all countries in 2021 were averaged, and the median was chosen as a measure of central tendency due to the positively skewed data. The countries were then divided into two groups based on whether their emissions were above or below the median. The research question is whether there is a significant difference in the IHDI between these two groups.

Explanation: While at this point I have not conducted a correlation analysis between CO₂ emissions and IHDI scores, dividing the data into two groups based on CO₂ levels to explain differences in IHDI between countries is a reasonable approach. This is because CO₂ emissions are a byproduct of economic activity, and the IHDI is a measure of human development that is closely linked to economic development as it measures GDP per capita, life expectancy, and access to education.

u_1 = mean IHDI for countries with annual CO₂ emissions above average

u_2 = mean for IHDI for countries with annual CO₂ emissions below average

$$H_0: u_1 = u_2$$

$$H_a: u_1 \neq u_2$$

Data results

Independent Sample T Tests: IHDI & Annual CO2 per capita

group_01_IHDI =	countries that emit above the 2021 annual_co2_per_capita median		
group_02_IHDI =	countries that emit below the 2021 annual_co2_per_capita median		
t-Test: Two-Sample Assuming Unequal Variances			
	group_01_IHDI	group_02_IHDI	
Mean	0.747068493	0.441521739	
Variance	0.012917842	0.015120812	
Observations	73	69	
Hypothesized Mean Difference	0		
df	137		
t Stat	15.35238416		
P(T<=t) one-tail	7.19523E-32		
t Critical one-tail	1.65605208		
P(T<=t) two-tail	1.43905E-31		
t Critical two-tail	1.977431212		

The t-test for independent means resulted in a t-statistic of 15.352, a critical two-tail t-value of 1.977, and a very low p-value of 1.439E-31, at a significance level of 0.05. These results suggest a significant difference in IHDI scores between the two groups of countries based on their CO2 emission levels. While the two-tail t-test does not suggest direction, it could be that countries with higher CO2 emissions have achieved better human development outcomes, possibly because they have emphasized economic growth, sustained by increased CO2 emissions. This could be because these countries have invested more in industrial development and infrastructure, leading to greater economic growth and better human development outcomes. While this approach has resulted in higher CO2 emissions, it has also produced greater economic and social benefits. These countries may have prioritized economic development over environmental sustainability, which has led to this trade-off between environmental impact and human development outcomes. While this independent variable (CO2 emissions) provides a good explanation for the differences between the groups, it would be interesting to explore other independent variables, such as the Democracy Index and the Corruption Perception Index (CPI), to determine if the quality of governance affects the quality of life of a country, as measured by the IHDI.

2. Using the same continuous variable, choose/create a nominal variable in the data that separates the cases into three/four groups (feel free to use one of the three variables you were thinking about in assignment 1). Write a research hypothesis that examines the differences in the mean for the three groups. Then, use ANOVA to analyze whether the variable of interest exhibits a statistically significant difference between the groups. Why and why not? If not, what variables would you try next? Do you need to perform a posthoc test? Describe the results by using the output from Excel.

Hypothesis: The research hypothesis is to investigate whether the type/quality of governance affects a country's human development. The Democracy Index measures the state of democracy in a country and scores it from 0-10, dividing it into four categories: Full Democracy, Flawed Democracy, Hybrid Regime, and Authoritarian Regime. An ANOVA test was performed on the data divided into these four groups to determine if the means among the groups were all equal or not.

Explanation: Using the Democracy Index as an independent variable is a good option to explain the differences in IHDI because it measures the level of democracy in a country, which is often linked to a higher quality of life. Countries with higher levels of democracy tend to have more stable and effective governance, which can lead to better economic development, social welfare, and human rights. These improvements can, in turn, enhance factors such as GDP per capita, life expectancy, and access to education. Therefore, exploring the relationship between the Democracy Index and IHDI can provide insights into how the quality of governance affects a country's human development.

Data Results

Anova Analysis: IHDI & Democracy Index

group_01_ihdi =	countries with Full Democracies					
group_02_ihdi =	countries with Flawed Democracies					
group_03_ihdi =	countries with Hybrid Regimes					
group_04_ihdi =	countries with Authoritarian Regimes					
Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
group_01_ihdi	18	15.338	0.85211	0.00519		
group_02_ihdi	50	34.208	0.68416	0.01879		
group_03_ihdi	32	15.399	0.48122	0.01762		
group_04_ihdi	42	20.056	0.47752	0.02769		
ANOVA						
Source of Variator	SS	df	MS	F	P-value	F crit
Between Groups	2.57945	3	0.85982	44.1021	4.7E-20	2.6702
Within Groups	2.69046	138	0.0195			
Total	5.26991	141				

At a significance level of 0.05, the ANOVA test resulted in an F-value of 44.1021 with 3 and 138 degrees of freedom and a p-value of 4.74E-20. The F-critical value was 2.6702, indicating that the means among the groups were not equal. Therefore, it can be concluded that the quality of governance in countries has an impact on their IHDI. While a correlation analysis would be necessary to determine the direction of this impact, it is plausible that countries with higher Democracy Index scores tend to have better human development outcomes.

One possible explanation for the observed difference is that countries with higher Democracy Indexes tend to have more transparent and accountable governance, leading to better allocation of resources towards human development. Additionally, democratic countries often have greater civil liberties and political stability, which may foster social and economic development.

Lastly, performing a posthoc test could provide valuable insights, as it would help identify which specific countries drive the observed significant difference in means, as determined by the F-statistic in the ANOVA test.

- Using the same continuous variable, choose/create another continuous variable in the dataset. Write a research hypothesis that examines the relationship between the two variables. Then, run a correlation analysis to evaluate the relationship between the two variables. Describe the results by using the output from Excel

Hypothesis: Finally, I will explore the correlation between the Corruption Perception Index (CPI) and the Inequality-Adjusted Human Development Index (IHDI). The CPI measures perceived levels of public sector corruption on a scale from 0 (highly corrupt) to 100 (very clean). My hypothesis is that there will be a positive correlation between these variables, meaning that countries with higher CPI, indicating less corruption, will have higher IHDI.

Explanation: Corruption can impact the allocation of resources and access to public services, which can affect a country's human development outcomes, making the Corruption Perception Index a useful independent variable to predict IHDI scores. Although I am not predicting a causal relationship as that would require more complex multivariable analysis, I am giving the relationship a direction and will consider CPI as the independent variable.

Data Results

Correlation Analysis Results: IHDI & Corruption Perception Index

	<i>ihdi</i>	<i>corruption_index</i>
<i>ihdi</i>	1	
<i>corruption_index</i>	0.50026	1

The moderately strong correlation coefficient of 0.5026 indicates that there is a tendency for the values of the two variables to increase together, a positive correlation. While this does not imply a causal relationship, the correlation between the Inequality-adjusted index and Corruption Perception Index could be attributed to the fact that lower levels of

corruption can lead to a more equal distribution of resources and opportunities, which in turn can result in higher levels of human development. This highlights the importance of good governance and transparency in creating a society that is more equitable and prosperous.

Regression Analysis Results

Regression Statistics									
Multiple R	0.50026								
R Square	0.25026								
Adjusted R Square	0.231517								
Standard Error	0.145886								
Observations	42								
ANOVA									
	df	SS	MS	F	Significance F				
Regression	1	0.284162884	0.284162884	13.35183128	0.000741791				
Residual	40	0.851307593	0.02128269						
Total	41	1.135470476							
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
Intercept	0.262448	0.063017926	4.164649162	0.00016141	0.135083573	0.389812	0.135083573	0.389812	
corruption_index	0.006997	0.001914896	3.654015774	0.000741791	0.00312691	0.010867	0.00312691	0.010867	

The regression analysis showed that 25% ($R^2 = 0.25026$) of the data fits the linear model predicted by the analysis, with a trend of 0.006997 IHDI/CPI. This indicates a statistically significant positive linear relationship ($P\text{-value} = 0.000741791$) between the two variables, reinforcing the results of the correlation analysis. However, as only two variables were included in this regression, more predictor variables are required to fully explain differences in IHDI. Nonetheless, this confirms the previous results of a positive correlation between CPI and IHDI.

- Find a peer reviewed article that discusses and statistically analyzes the same dependent variable in your dataset. It doesn't have to be exactly the same, but it should be close (conceptually). From the article, what variable(s) was used to explain the differences observed in the dependent variable? What is the relationship? Are you surprised by the findings? If you were to add one new explanatory variable to

your data, do you think you will find statistical significance? Why and why not?
(10)

With regard to previous studies that have looked into measures of human development, I found the following two. An Ethiopia-based analysis using 2015 data on the top 20 CO₂ emitters and 5 African Nations, found a positive correlation between a country's Human Development Index (HDI) and per capita CO₂ emissions. Teklu (2018) recommended that Ethiopia prioritize industrial development and increase its CO₂ emissions until it reaches a fair emission budget and achieves an HDI of $\geq .8$. Moreover, Costa et al. (2011) found a positive correlation between per capita CO₂ emissions and Human Development Index (HDI) in 147 countries using 2000 data. They also suggested that developing countries aim for HDI scores ≥ 0.80 . However, to achieve this and remain within the UN's 2°C warming limit by 2100, they estimated that developed nations (HDI $\geq .80$) needed to reduce their per capita emissions by 17% and 33% every five years until 2050.

These studies differ slightly from my investigation as they utilized the Human Development Index (HDI), whereas I used the Inequality-Adjusted Human Development Index (IHDI), which I believe to be a more accurate measure of human development. Although the data from these studies is a bit older, from 2015 and 2000, their predictions of a positive correlation between CO₂ emissions and human development align with my findings. Moreover, to find additional statistically significant explanatory variables, they must be good predictors of human development. Some potential variables that could be added to predict IHDI include Gini coefficient, which is a measure of income inequality within a population. Additionally, another option could be access to modern energy sources (energy consumption levels), as lack of access can hinder economic and social development. In conclusion, a country's economic and social situation must be influenced by the explanatory variables in order to have statistical significance in predicting differences in IHDI.

References

- Costa, L., Rybski, D., & Kropp, J. P. (2011). A human development framework for CO₂ reductions. *PLoS ONE*, 6(12). <https://doi.org/10.1371/journal.pone.0029262>
- Teklu, T. W. (2018). Should Ethiopia and least developed countries exit from the Paris Climate Accord? – geopolitical, development, and Energy Policy Perspectives. *Energy Policy*, 120, 402–417. <https://doi.org/10.1016/j.enpol.2018.04.075>

