STEM Fellowship Big Data Challenge 2022

# Exploring Economic Factors in Clean Energy Adoption Worldwide

Hanz Nathan Qua Po

Francis Libermann Catholic High School

February 2, 2022

## Abstract

Climate change has been and is still a major problem throughout the world. Some of the top contributors to CO2 emissions are coal, oil, and gas [1], making the switch to clean energy an obvious and necessary one. This study investigated how economic factors influence the adoption of clean energy across the world, by looking at Gross Domestic Product (GDP) per capita, oil production per capita, and Human Development Index (HDI) values. This information was then compared to the proportion of energy from low-carbon sources in each country. Pearson Correlation Coefficients and $R^2$ values were examined in conjunction with Polynomial Regression analysis, primarily with Python.

Since GDP per capita is a reliable indicator of economic strength [2], it was the first factor considered in the analysis, in addition to HDI, which takes into account several economic and developmental factors, offering a different perspective of a country's success [3]. Finally, oil production per capita was also chosen, because it was hypothesized that countries which benefit economically from oil production may be less inclined to adopt clean energy.

Based on these data, results from the analysis indicated that countries with a higher GDP per capita and HDI were more likely to adopt clean energy. Countries that produce more oil are less likely to adopt clean energy, however, there are numerous exceptions. Therefore, economic factors play a significant role in clean energy adoption, but there may be other contributing factors. Future studies may look into exactly what other factors affect it.

**Keywords**
renewable, energy, economy, GDP, development

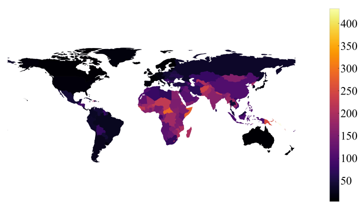


Figure 1: Number of deaths due to air pollution every year per 100,000 people.

## 1 Introduction

As climate change continues to become a universal threat to the world [4], fossil fuels and other non-renewable energy sources still primarily power countries around the world [5], acting as the lifeblood to many economies and industries. The production of energy from fossil fuels results in greenhouse gas emissions, which, in turn, causes a variety of disastrous effects, such as rising sea levels, inflating global temperatures, a weakened ozone layer, extreme weather events, and countless others [4]. Already, an estimated 3.61 million people die as a result of air pollution every year [6], six times more than the yearly number of war casualties, murders, and terrorist attacks combined [1].

With these in mind, it is of utmost importance for nations to take this issue seriously, reduce reliance on fossil fuels, and make the transition to cleaner sources of energy, such as solar, wind, hydropower, biomass, geothermal, and nuclear [7, 8].

This study aimed to investigate the relationship between the economies of various countries and how it relates to the share of their energy that comes from clean sources. Economic indicators such as GDP per Capita and the United Nations' Human Development Index were compared to the percentage of a country's energy production that comes from low-carbon sources.

Using a variety of data analysis and visualization libraries for Python, numerous datasets were processed into a format easily compatible for use in Python. Polynomial regression analysis was then utilised to observe possible trends or patterns in the data. Graphs and diagrams were generated based on the values calculated. The results demonstrated that there is a very weak correlation between a country's economic strength and its adoption of low-carbon energy.

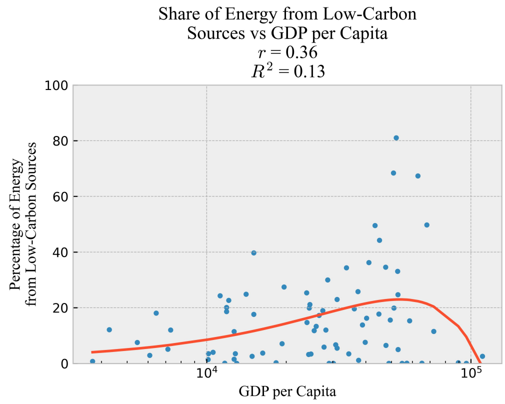


Figure 2: Relationship between proportion of energy from low-carbon sources and GDP per capita.

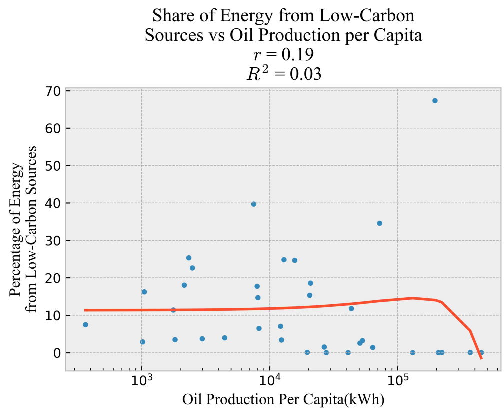


Figure 3: Relationship between proportion of energy from low-carbon sources and oil production per capita.

## 2 Materials & Methods

Data were gathered from the *Our World in Data* project compiled by the University of Oxford [1]. The data gathered from their project was compiled from a variety of sources, primarily the World Bank, the United Nations, and numerous government websites. The format of the datasets was in the comma separated values format, or .csv, so the pandas library was used to process them and create a dataframe. Some of the information gathered from these datasets contained information about previous years, which were filtered out using the same library.

From there, the scikit-learn library's LinearRegression and PolynomialFeatures functions were used to train polynomial regression models based on the data provided in each dataset. Coefficients of determination, also known as $R^2$, mean squared errors, and Pearson correlation coefficients were also calculated using this library.

When creating maps to visualise the comparison between certain values for countries around the world, GeoPandas was the library used, which itself relies on other libraries to generate maps such as shapely, fiona, GDAL. Vector and raster data for the world map were obtained from the *Natural Earth* project. Country codes from the International Organization for Standardization were used to align geographical data to data from other sources, in order to ensure that discrepancies between country names would be avoided. The final map was then generated by matplotlib.

The other diagrams containing the model generated by scikit-learn and data from the datasets were plotted using matplotlib's pyplot class. The platforms used were Microsoft Excel for previewing .csv files, and Visual Studio Code for utilizing Python with the previously mentioned libraries. Diagrams.net was used to create the flowchart.

## 3 Results

Results were generated into a plot through the aforementioned process. In all graphs, $r$ is used to denote the Pearson Correlation Coefficient of each model, and $R^2$ is used to denote the Coefficient of Determination. Based on the $r$ and $R^2$ values and through visual observation, a weak but present correlation exists in each of the graphs. In Figures 2 and 3, a logarithmic scale was used for the GDP per Capita Oil Production Per Capita axes to improve the clarity of the graph. Figures 5 and 6 do not present an analysis, however, they have been added for clarity on how these factors are related to geographical location.

## 4 Discussion

The correlation between all three factors and the availability of low-carbon energy sources is

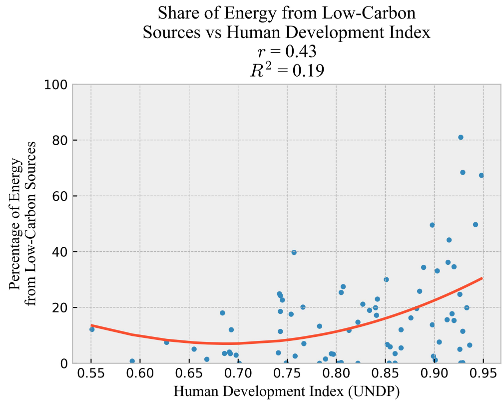


Figure 4: Relationship between proportion of energy from low-carbon sources and human development index.

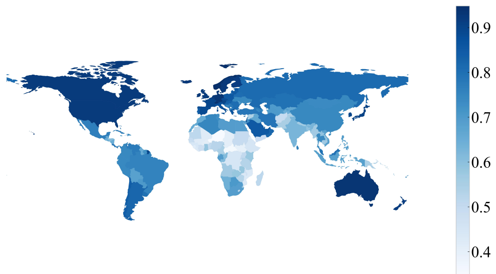


Figure 5: World map of Human Development Indexes.

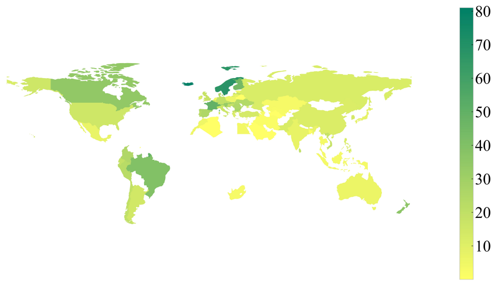


Figure 6: World map of proportion of energy derived from low-carbon sources. Note: Data are not available for some countries.

very apparent, especially when considering the $r$ and $R^2$ values. While they may not be particularly strong, they show that GDP, oil production, and Human Development Index values are linked to the availability of clean energy.

The metric chosen to reflect a country's adoption of clean energy was the proportion of their energy that came from low-carbon sources for a variety of reasons. Choosing to measure it in terms of total kilowatt-hour (kWh) would not have been as effective because different countries have differing energy needs, introducing factors that were irrelevant to the topic. Even measuring by total kWh per capita would only take into account the population factor of energy generation. Measuring only renewable energy was initially considered, however, this excludes nuclear energy, which, while not technically renewable, is widely considered to be clean energy, and is the second most popular source of low-carbon energy [8].

In Figure 2, the model compared the GDP per capita of a country to the availability of clean energy. GDP per capita was used because it is a reliable indicator of an economy's size [2]. The results show that generally, countries with a higher GDP per capita have a higher share of energy from low-carbon sources. However, the correlation is relatively weak, with a Pearson Correlation Coefficient value of just 0.36, and an $R^2$ value of 0.13, meaning that there are other factors that contribute. For example, Qatar, despite having a very high GDP per capita, does not use a lot of low-carbon energy sources.

In Figure 3, oil production per capita was compared to clean energy use. As depicted by the graph, a Pearson Correlation Coefficient value of 0.19, and an $R^2$ value of only 0.03, the correlation is very weak. As such, we can come to the conclusion that the amount of oil a country produces has no bearing on their adoption of clean energy. One example where this is clear is when comparing Canada to the United States. Canada produces approximately three times more oil than the United States by kWh per capita, yet 33.94% of its energy comes from low-carbon sources, as opposed to the United States' 16.74%. This is in stark contrast to what was hypothesised, as at first glance, one may be led to believe that oil-producing countries would be less likely to adopt clean energy.

Finally, Figure 4 shows a clear correlation between a country's Human Development Index (HDI) value and their low-carbon energy share. As depicted in Figure 7, the HDI value is calculated from a variety of factors, creating a final value that summarises the level of human development in a given country, including economic

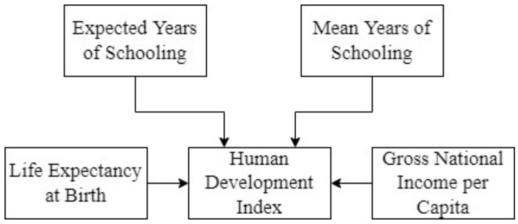


Figure 7: Factors that are taken into account in the United Nations Development Programme's Human Development Index.

growth, but also introducing other factors. The Pearson Correlation Coefficient and $R^2$ values are the highest of all of the values compared in this study, showing that there is a distinct link between a country's human development and its adoption of clean energy. Yet once again, with a mean squared error of 230.07, there is a lot of variation in the data, meaning that there are certainly other factors that contribute.

## Conclusions

Based on the analysis performed in this study, it is evident that economic factors influence clean energy adoption in a given country. GDP per capita, a measure of economic power, has a weak correlation. On the other hand, HDI, a measure of human development, has a much stronger correlation. A country's oil production also has a weak, but still non-negligible impact. Therefore, while the study has shown the clear role that economic factors play in clean energy adoption, other, non-economic factors also can affect it, which could be the topic of future studies.

## Acknowledgements

I would like to thank my supervisor, Mr. Keung, without whom this report would not have been possible. I would also like to thank everyone at STEM Fellowship involved in organising this challenge, and everyone who presented during one of the workshop events, through which they shared their valuable knowledge, experience, and expertise.

## References

[1] Hannah Ritchie and Max Roser. Energy. *Our World in Data*, 2020. https://ourworldindata.org/energy.

[2] Elvis Picardo. The importance of gdp. *Investopedia*, Sep 2021.

[3] United Nations Human Development Programme. Human development reports. *Human Development Index (HDI).*

[4] Marcia McNutt. Climate change impacts. *Science*, 341(6145):435–435, 2013.

[5] British Petroleum Company. Statistical review of world energy: Energy economics: Home. *bp global*, 2021.

[6] J. Lelieveld, K. Klingmüller, A. Pozzer, R. T. Burnett, A. Haines, and V. Ramanathan. Effects of fossil fuel and total anthropogenic emission removal on public health and climate. *Proceedings of the National Academy of Sciences*, 116(15):7192–7197, 2019.

[7] U.S. Energy Information Administration. Eia - independent statistics and analysis. *Renewable energy explained - U.S. Energy Information Administration (EIA)*, 2020.

[8] U.S. Department of Energy. 3 reasons why nuclear is clean and sustainable. *Energy.gov*, Mar 2021.