0. Group Members & ID

Anson Ma, #20916612
Tahmid Bari, #20864394

1. Introduction

- Carbon dioxide (CO_2) is an important heat-trapping gas, which is released through human activities such as deforestation and burning fossil fuels, as well as natural processes such as respiration and volcanic eruption
- Over the past decade, human activities have raised atmospheric concentrations of CO_2 by 47% above pre-industrial levels and presented one of the world's most pressing challenges to combat climate change
- However, far from stabilizing concentrations, the global CO_2 emissions are in fact still rising and accumulating. The world has not yet peaked
- A changing climate has a range of potential ecological, physical, and health impacts, including extreme weathers, sea-level rise; altered crop growth; disrupted water systems. We have all been experiencing these negative impacts already
- If countries achieved their current pledges and targets, there would be an in the coming years. In this regard, the world is making some progress
- But if our aim is to limit global warming to "well below 2°C" as laid out in the Paris Agreement, we are clearly still far from the rates of progress we would need to achieve international targets. Countries will have to identify factors that affect their carbon dioxide emissions
- A number of countries, have managed to reduce CO_2 emissions whilst increasing gross domestic product (GDP)
- The objective of this project is to examine the relationship between CO_2 emissions, GDP and natural factors through statistical methods
- Natural factors include the differences in average climate temperatures, the proportion of urban land areas, and the availability of renewable and fossil fuel resources
- While many studies have shown the empirical relationship between CO_2 emissions and income, the question to what extent natural factor determines cross-country differences in CO_2 emissions has been somehow neglected
- The analysis would be mainly based on and compared against four countries of different income groups, namely Ethiopia (low-income), Vietnam (lower-middle-income), Thailand (middle-income), and Japan (high-income)

2. Literature Review

- We started off our project by researching for some literatures done in the past years
- The question that takes the central stage is the responsibility for climate change among nations across the globe. It is also central to public and policy discourse over actions to cut greenhouse gas emissions and limit adverse impacts
- The United Nations Framework Convention on Climate Change (UNFCCC) has established the principle of “common but differentiated responsibilities” among nations, signaling the recognition that nations that had produced the larger share of historical emissions will bear greater responsibilities. (UNFCCC, 1998)
- Reflecting this principle, the Paris Agreement establishes common commitments, for example to global net-zero greenhouse gas emissions in the second half of this century, while allowing flexibility in mitigation efforts to accommodate different national capacities and circumstances (United Nations, 2015)
- The relationship between economic growth and environmental quality does exist. This relationship could be positive before economies cross a certain level of income. Reaching a ‘sustainable development’ is possible, thus turning economic growth compatible with environmental quality, after that certain level of income. (Hu et al., 2011)
- The share of renewables and non-renewable energy sources are important to explain differences in emissions. They suggest a significant change in the trend of economic and environmental efficiency in European countries and put forward the high disparities existing among them. (Moutinho et al., 2017)
- The growth of renewable energy consumption has a significant positive and negative impact on economic output and CO2 emissions respectively. Institutions have a positive influence on economic growth and a reducing effect on CO2 emissions. (Bhattacharya et al., 2017)
- A comparison of the national results at the product level points out that country characteristics, like energy supply, population density and the availability of district heating, influence variation in household CO2 emissions between and within countries. (Kerkhof et al., 2009)

3. Methods

- RStudio and Microsoft Excel will be used to execute and present the methodologies
- Dependent Variable test of normality
    - Apply the Kolmogorov–Smirnov test (K–S test) and Shapiro–Wilk (S-W test) to check the normality of the Dependent Variable (DV)

- Univariate Analysis: Sample t-test
    - Apply Sample t-test to compare the mean of DV on 4 countries
    - Construct 95% confidence intervals around the largest and the smallest mean
- Multivariate analysis: OLS regression model
    - Apply Ordinary least squares (OLS) regression model
    - Residual analysis to check for normality
    - Breusch–Pagan test to check for heteroskedasticity

- Proposed model

$$Y_i = a_i + b_{i1}x_{i1} + b_{i2}{x_{i1}}^2 + b_{i3}x_{i3} + b_{i4}x_{i4} + b_{i5}x_{i5} + b_{i6}x_{i6} + b_{i7}x_{i7} + e_i$$

(Neumayer E, 2002)

Where

$Y = CO_2\ emission\ per\ capita$
$x_1 = GDP\ per\ capita$
${x_1}^2 = squared\ GDP\ per\ capita$
$x_3 = lowest\ monthly\ average\ temperature$
$x_4 = highest\ monthly\ average\ temperature$
$x_5 = percentage\ of\ urban\ land\ in\ total\ land\ area$
$x_6 = percentage\ of\ renewable\ energy\ use\ in\ total\ energy\ use$
$x_7 = percentage\ of\ fossil\ fuel\ consumption\ \ in\ total\ energy\ use$
$\varepsilon\ = error\ term$

$i = 1{:}\ Ethiopia;\ \ 2{:}\ Vietnam, 3{:}\ Thailand, 4{:}\ Japan$

4. Data

- The raw data is consisted of a panel covering 25 years from 1991-2015 on 4 countries collected form the World Bank to maintain consistency, expecting a better significant mode

Table.: Descriptive Statistics

Variable	Variable Name	Description	Type	Mean	SD	Minimum	Maximum
Y	co2_emission	carbon dioxide emission per capita	continuous	3.440	3.740	0.041	9.881
x1	gdp_capita	US$ gross domestic product per capita	continuous	10,764	12,139	347	40,396
x2	gdp_capita2	squared US$ gross domestic product per capita	continuous	261,744,033	430,488,940	120,120	1,631,856,465
x3	lowest_temp	lowest monthly average temperature	continuous	15.754	9.755	-1.908	24..544
x4	highest_temp	highest monthly average temperature	continuous	25.807	2.497	20.688	30.336
x5	urban_area	percentage of urban land in total land area	continuous	10.247	11.907	0.517	29.816
x6	renewable_energy	percentage of renewable energy use in total energy use	continuous	43.679	35.481	3.568	97.740
x7	fossil_fuel	percentage of fossil fuel consumption in total energy use	continuous	54.392	32.842	2.245	94.633

- CO2 emissions per capita, forming the dependent variable, is based on the data from Carbon Dioxide Information Analysis Center (CDIAC). CO2 emissions are those stemming from the burning of fossil fuels and the manufacture of cement. These emissions also include CO2 produced during consumption of solid, liquid, and gas fuels and gas flaring
- Income, as one of the independent variables, is based on the data from World Bank National Accounts and OECD National Accounts. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. GDP per capita is obtained by diving GDP by mid-year population. Higher income countries would have greater emissions to sustain economic development and activities, and vice versa
- The lowest and highest monthly average temperatures are based on the data from World Bank Climate Change Knowledge Portal (CCKP). The temperatures are the average of temperatures of a month, which are then compared against each other within the year to identify the lowest and the highest of the year. It is expected that cold countries would have greater heating demands while hot countries would have greater cooling demands. Cold countries and hot countries would therefore have higher CO2 emissions
- The percentage of urban areas is based on the data from CIESIN Urban-Rural Population and Land Area Estimates and the Food and Agriculture Organization. The urban area is computed on a combination of population counts, settlement points, and the presence of nighttime lights. The numbers are then divided by the country total land area to obtain the percentage. Countries with less urban areas are sparsely inhabited and have higher transportation demands to move goods and people over long distances. Higher transportation demands would have higher emissions, and vice versa
- The percentage of renewable energy in total energy use is based on the data from World Bank Sustainable Energy for All (SE4ALL). Renewable resources encompass hydroelectric, geothermal, solar and wind resources as well as “fuel and waste”,

which comprise biomass and animal products, gas/liquids from biomass, industrial waste, and municipal waste. It is expected that countries that have access to domestic renewable energy resources would have lower emissions than countries that lack such resources

- The percentage of fossil fuel consumed is based on the data from IEA Statistics. Fossil fuel comprises coal, oil, petroleum, and natural gas products. The higher the consumption, the fewer the reserve, vice versa. Countries that have fewer fossil fuel reserves should have lower CO2 emissions than countries that are rich in such reserves. This is for two reasons: First, because of the emissions generated in the extraction and possibly the transport and processing of such resources. Second, because of countries that lacked major domestic fossil fuel reserves have had strong incentives to develop in a less fossil fuel intensive way to cut down on energy import costs

5. Results

- Dependent Variable Test of Normality

    - Kolmogorov-Smirnov test

        - The Kolmogorov-Smirnov test (K-S test) compares the data with a known distribution and tells us if they have the same distribution. It is commonly used as a test for normality to see if our data is normally distributed.
        - The hypotheses:

          H0: The data is normally distributed
          H1: The data is not normally distributed

        - The K-S test p-value is less than α = 0.05, therefore did not fail to reject H0 concluding the data are not normally distributed. Again, this is expected given the dv_normality object was created via the rnorm() and co2 emission as runif() function

```
        Two-sample Kolmogorov-Smirnov test

data:  dv_normality and co2_emission
D = 0.48, p-value = 0.0002033
alternative hypothesis: two-sided
```

- Shapiro-Wilk test

  - The shapiro.test() function in R (program) employs the Shapiro-Wilk test (S-W test) on data to test whether the data are normally distributed. Use of the Shapiro-Wilk test is (less-sensitive) and contingent on univariate and continuous data.
  - The hypotheses:

    H0: The data is normally distributed
    H1: The data is not normally distributed

  - The S-W test p-value is greater than α = 0.05, therefore failing to reject H0 concluding the data are normally distributed. Again, this is expected given the dv_normality object was created via the rnorm() function.

```
        Shapiro-Wilk normality test

data:  dv_normality
W = 0.99876, p-value = 0.7254
```

  - The S-W test is also performed on the DV = CO2_emission, followed by another S-W test on the squared values to check the normality of the DV.

Using the CO2_emission (dependent variable), we have performed the S-W Test using R

Country	W	p-value	alpha	Results	Comments
Japan	0.9359	0.1072	0.05	p-value > alpha	Normally Distributed
Thailand	0.95715	0.3386	0.05	p-value > alpha	Normally Distributed
Vietnam	0.92054	0.04629	0.05	p-value < alpha	Not Normally Distributed
Ethiopia	0.86383	0.002671	0.05	p-value < alpha	Normally Distributed

By squaring the CO2_emission (dependent variable), we have performed the S-W Test using R

Country	W	p-value	alpha	Results	Comments
Japan	0.94169	0.1474	0.05	p-value > alpha	Normally Distributed
Thailand	0.96402	0.4768	0.05	p-value > alpha	Normally Distributed
Vietnam	0.87207	0.003933	0.05	p-value < alpha	Not Normally Distributed
Ethiopia	0.74232	2.13E-05	0.05	p-value > alpha	Normally Distributed

- Univariate Analysis

    - Sample t-test

	Japan	Thailand
Mean	9.391375457	3.103616614
Variance	0.108738244	0.568406843
Observations	26	26
Pearson Correlation	0.447874747	
Hypothesized Mean Difference	0	
df	25	
t Stat	47.55965533	
P(T<=t) one-tail	2.41779E-26	
t Critical one-tail	1.708140761	
P(T<=t) two-tail	4.83558E-26	
t Critical two-tail	2.059538553	

	Japan	Vietnam
Mean	9.39138	0.979067422
Variance	0.108738	0.299928883
Observations	26	26
Pearson Correlation	0.216848	
Hypothesized Mean Differenc	0	
df	25	
t Stat	74.6308	
P(T<=t) one-tail	3.35E-31	
t Critical one-tail	1.708141	
P(T<=t) two-tail	6.71E-31	
t Critical two-tail	2.059539	

	Japan	Ethiopia
Mean	9.391375457	0.070484373
Variance	0.108738244	0.000552898
Observations	26	26
Pearson Correlation	0.06671267	
Hypothesized Mean Difference	0	
df	25	
t Stat	144.4497966	
P(T<=t) one-tail	2.35866E-38	
t Critical one-tail	1.708140761	
P(T<=t) two-tail	4.71732E-38	
t Critical two-tail	2.059538553	

	Thailand	Vietnam
Mean	3.103616614	0.979067422
Variance	0.568406843	0.299928883
Observations	26	26
Pearson Correlation	0.937013138	
Hypothesized Mean Difference	0	
df	25	
t Stat	35.22874312	
P(T<=t) one-tail	3.93694E-23	
t Critical one-tail	1.708140761	
P(T<=t) two-tail	7.87387E-23	
t Critical two-tail	2.059538553	

- Japan: mean (H0 = 9.391375457) > Thailand: mean (Ha = 3.103617)
- Japan: mean (9.391375457) > Vietnam: mean (0.979067422)
- Japan: mean (Ho = 9.391375457) > Ethiopia: mean (Ha = 0.070484)
- Thailand: mean (Ho = 3.103616614) > Vietnam: mean (Ha = 0.979067422)

    - Confidence Interval

- Construct 95% confidence intervals around the largest mean and the smallest mean

(Largest) Mean - Japan	(Smallest) Mean - Ethiopia						
9.391375457	3.103616614	6.287759		(Largest) Mean	9.320891	Sample SD	4.206154
9.391375457	0.979067422	8.412308		(Smallest) Min	0.908583	Margin of Error	5.829324914
9.391375457	0.070484373	9.320891		(Largest) Mean LB (Mean - E)	3.491566	confidence coefficient	0.95
3.103616614	0.979067422	2.124549		(Largest) Mean UB (Mean + E)	15.15022	alpha	0.05
3.103616614	0.070484373	3.033132		(Smallest) Min LB (Mean - E)	-4.92074		
0.979067422	0.070484373	0.908583		(Smallest) Min UB (Mean + E)	6.737908		
	(Largest) Mean - Japan	9.32089		(Smallest) Mean - Ethiopia	0.90858		
	Sample SD	3.48848		Sample SD	3.48848		
	Margin of Error	3.660936		Margin of Error	3.660936		

- Multivariate Analysis

  - OLS Regression

    - Using RegressIt to generate the Model_1 based on the DV = CO_2_emission

Model:	Regression							
Dependent Variable:		co2_emission						
	R-Squared	**Adj.R-Sqr.**	**Std.Err.Reg.**	**Std.Dep.Var.**	**# Fitted**	**# Missing**	**Critical t**	**Confidence**
	0.998	0.998	0.181	3.740	100	4	1.986	95.0%
Variable	**Coefficient**	**Std.Err.**	**t-Statistic**	**P-value**	**Lower95%**	**Upper95%**	**VIF**	**Std. Coeff.**
Constant	0.009380	1.095	0.009	0.993	-2.166	2.184	0.000	0.000
fossil_fuel	0.015	0.009280	1.660	0.100	-0.003029	0.034	280.701	0.135
gdp_capita	0.000295	0.000015	19.538	0.000	0.000265	0.000325	101.212	0.956
gdp_capita2	-4.766E-09	2.981E-10	-15.986	0.000	-5.358E-09	-4.174E-09	49.776	-0.549
highest_temp	-0.034	0.026	-1.337	0.185	-0.085	0.017	12.302	-0.023
lowest_temp	-0.035	0.011	-3.037	0.003	-0.057	-0.012	37.162	-0.090
renewable_energy	0.013	0.010	1.289	0.201	-0.007050	0.033	389.579	0.124
urban_area	0.149	0.016	9.195	0.000	0.117	0.181	112.320	0.474
	Mean Error	**RMSE**	**MAE**	**Minimum**	**Maximum**	**MAPE**	**A-D* stat**	
Fitted (n=100)	0.000	0.174	0.130	-0.919	0.366	32.1%	1.00 (P=0.012)	

  - Residual Analysis

Figure.: Plot of the Residual -vs- Predicted

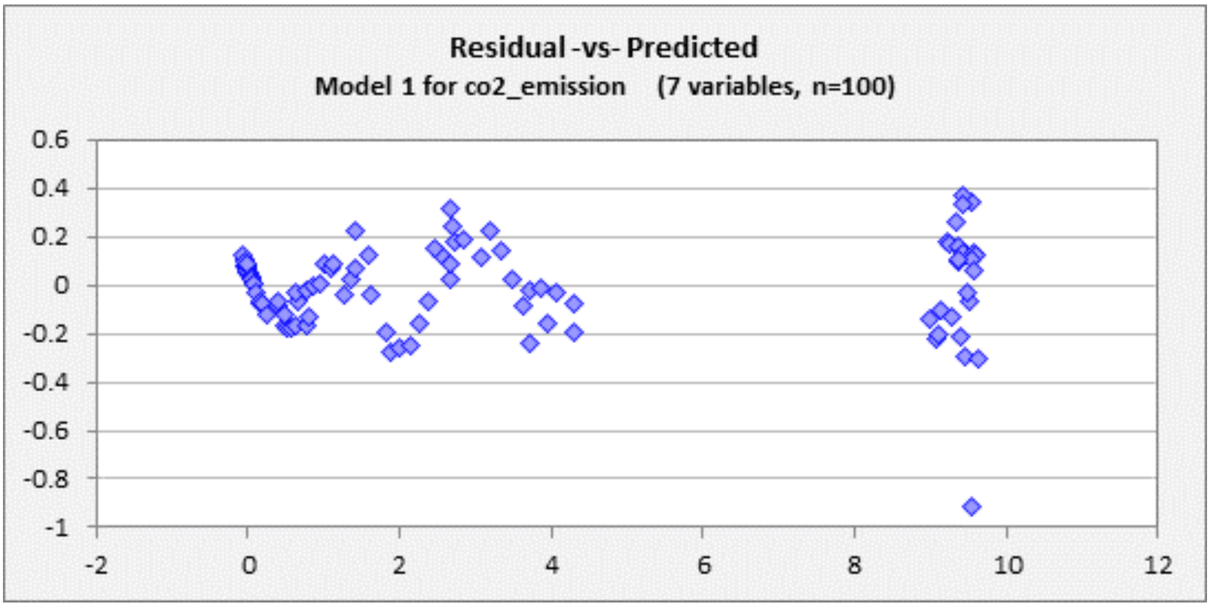
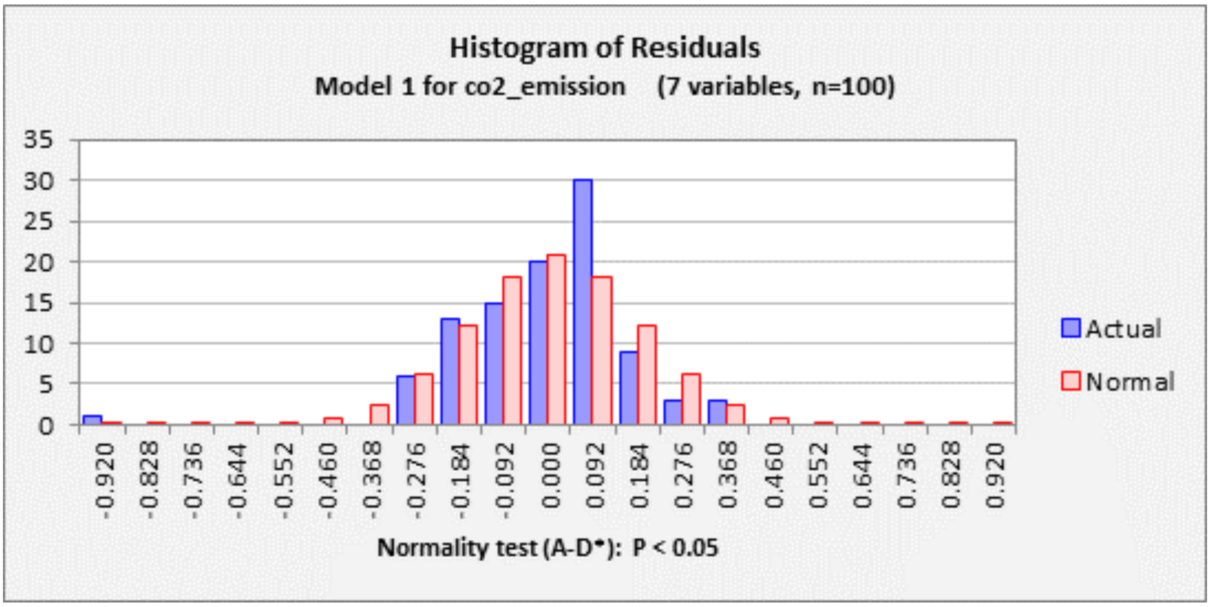


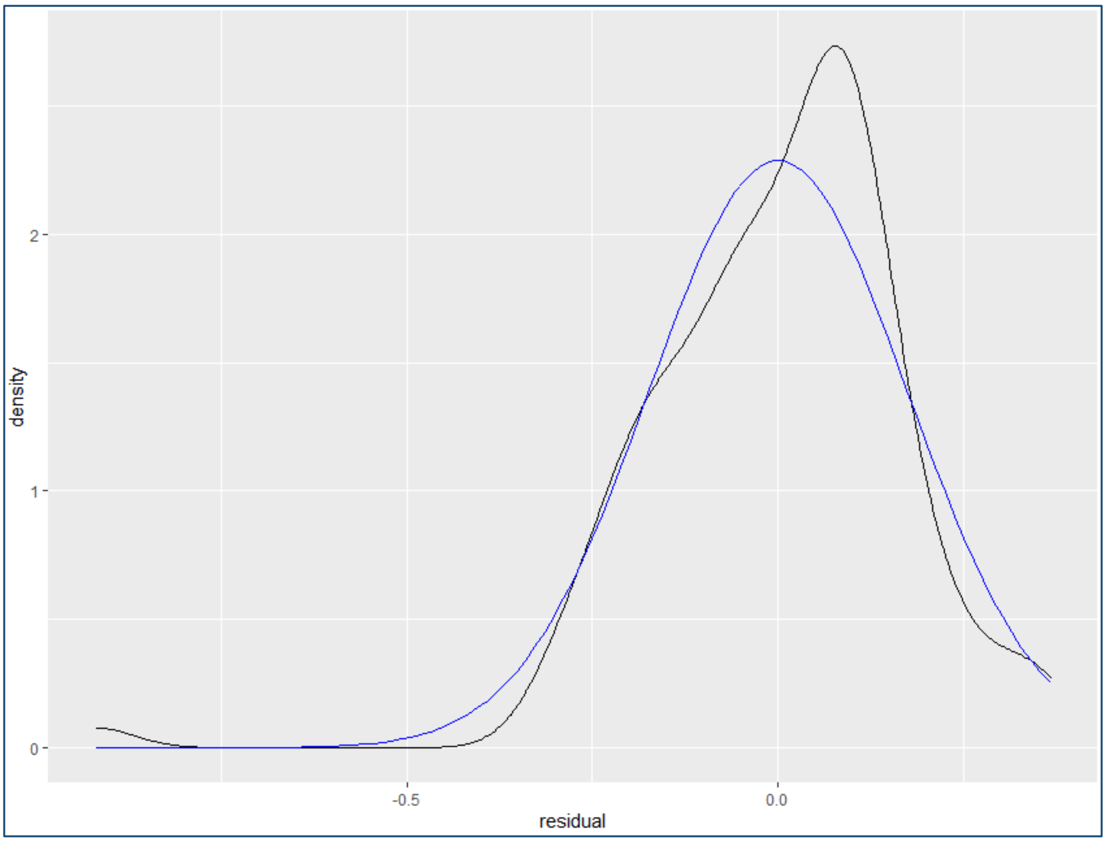
Figure.: Histogram of Residuals



- The density plot provides a visual judgment about whether the distribution is bell shaped. The black line represents the residual data, and the blue line represents the normal distribution given the μ and σ values of residual. Note: The residual data appears as bell shaped and does fit the normal distribution model given the parameters calculated via the residual data.

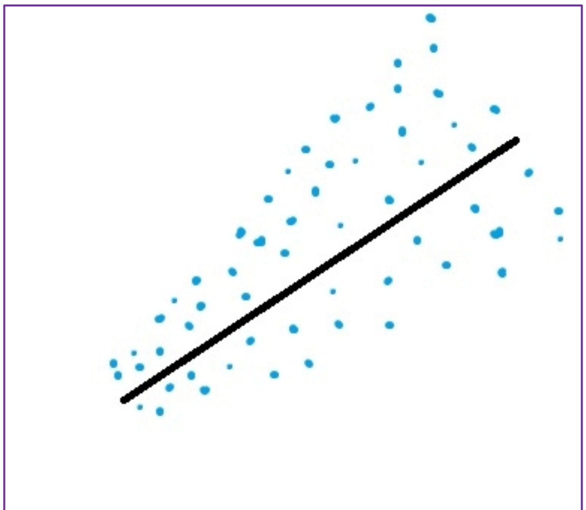
```
Shapiro-Wilk normality test
data:  d3_residual$residual
W = 0.91241, p-value = 5.696e-06

describe(d3_residual$residual)
   vars   n mean   sd median trimmed  mad   min  max range  skew kurtosis   se
X1    1 100    0 0.17   0.02    0.01 0.14 -0.92 0.37  1.29 -1.35     6.04 0.02
```

Figure.: Density plot of the residuals



- The Breusch-Pagan test is a test for heteroscedasticity of errors in regression. Heteroscedasticity means “differently scattered”; this is opposite to homoscedastic, which means “same scatter.” Homoscedasticity in regression is an important assumption; if the assumption is violated, you won’t be able to use regression analysis.

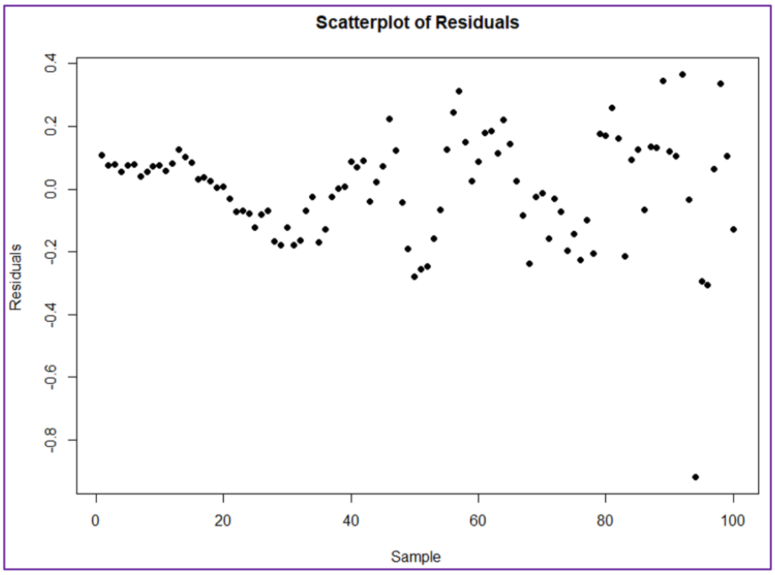
Figure.: Sample of heteroskedastic data



Heteroscedastic data.

- Check for heteroskedasticity: There seems to be no evident pattern. However, it does seem to look as if there's more variation in residuals in this sample data from the linear regression.

Figure.: Scatter plot of the residuals



6. Implications and Conclusion

- Natural factors can explain cross-country differences in CO2 emissions only to some limited extent only
- Countries having colder climate, or a lower availability of renewable resources have higher fossil fuel consumption than countries with warmer climates or a higher availability of renewable resources
- There is a relationship between CO2 emissions with GDP growth and fossil fuels. As the value of GDP, urban area, renewable energy increases; CO2 emissions decreases
- Countries with having low natural conditions (temperature) demanded higher carbon emissions than countries at roughly the same income levels
- Limitations of the model
    - There was a huge data gap in between the high-income and low-income country in terms of dependent and all independent variables
    - Used the square of GDP per-capita
    - Taken the percentage of urban land in total land area, renewable energy use in total energy use, fossil fuel consumption in total energy use
- Future work
    - During a pandemic situation, we would want to extend out the model.

- Further reading:
    - Impact of COVID-19 outbreak measures of lockdown on the Italian Carbon Footprint (Rugani, et al. 2020)
    - To study government initiatives that impact on the carbon-footprint
    - The dynamic impact of renewable energy and institutions on economic output and CO_2 emissions across region (Bhattacharya, et al., 2017)
    - To include the usage of plastic and bio-degradable items

7. References

Bhattacharya M, Churchill S A, Paramati S R (2017). 'The dynamic impact of renewable energy and institutions on economic output and CO_2 emissions across regions'. Renewable Energy 111, Pp 157-167

Hu Z G, Yuan J H, Hu Z (2011). 'Study on China's low carbon development in an Economy-Energy-Electricity-Environment framework.' Energy Policy 39 (5), Pp 2596–2605

Kerkhof C A, Benders M J R, Moll C H (2009), 'Determinants of variation in household CO_2 emissions between and within countries'. Energy Policy 37 (4), Pp 1509-1517

Moutinho V, Madaleno M, Margarita R (2017), 'The economic and environmental efficiency assessment in EU cross-country: Evidence from DEA and quantile regression approach' Ecological Indicators 78, Pp 85-97

Neumayer, E (2002). 'Can Natural Factors Explain Any Cross-Country Differences in Carbon Dioxide Emissions?' Energy policy, 30 (1). pp. 7-12.

Ritchie, H and Roser, M (2017). 'CO_2 and Greenhouse Gas Emissions' [Online Resource]. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions

UNFCCC (1998). 'Kyoto Proocol to the United Nations Framework Convention on Climate Change' [Online Resource]. http://unfccc.int/resource/docs/convkp/kpeng.pdf

United Nations (2015). 'Paris Agreement'. [Online Resource]. https://unfccc.int/process-and-meetings/the-paris-agreement/theparisagreement

World Bank, World Development Indicators. (2016). ‘CO2 emissions (metric tons per capita)’ [Data file]. https://data.worldbank.org/indicator/EN.ATM.CO2E.PC

World Bank, World Development Indicators. (2019). ‘GDP per capita (current US$)’ [Data file]. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

World Bank, World Development Indicators. (2016). ‘Climate Change Knowledge Portal’ [Data file]. https://climateknowledgeportal.worldbank.org/download-data

World Bank, World Development Indicators. (2013). ‘Urban land area (sq. km)’ [Data file]. https://data.worldbank.org/indicator/AG.LND.TOTL.UR.K2

World Bank, World Development Indicators. (2018). ‘Land area (sq. km)’ [Data file]. https://data.worldbank.org/indicator/AG.LND.TOTL.K2

World Bank, World Development Indicators. (2015). ‘Renewable energy consumption (% of total final energy consumption)’ [Data file]. https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS

World Bank, World Development Indicators. (2015). ‘Fossil fuel energy consumption (% of total final consumption)’ [Data file]. https://data.worldbank.org/indicator/EG.USE.COMM.FO.ZS

8. Appendix

Table.: Description of the Variables

Variable Name	Description	Type	Source	Website
country	countries under analysis	categorical	–	–
country_code	country code	categorical	–	–
year	calendar year	integer	–	–
co2_emission	carbon dioxide emission per capita	continuous	The World Bank	https://data.worldbank.org/indicator/EN.ATM.CO2E.PC
gdp_capita	US$ gross domestic product per capita	continuous	The World Bank	https://data.worldbank.org/indicator/NY.GDP.PCAP.CD
gdp_capita2	squared US$ gross domestic product per capita	continuous	The World Bank	https://data.worldbank.org/indicator/NY.GDP.PCAP.CD
lowest_temp	lowest monthly average temperature	continuous	The World Bank	https://climateknowledgeportal.worldbank.org/download-data
highest_temp	highest monthly average temperature	continuous	The World Bank	https://climateknowledgeportal.worldbank.org/download-data
urban_area	percentage of urban land in total land area	continuous	The World Bank	https://data.worldbank.org/indicator/AG.LND.TOTL.UR.K2
renewable_energy	percentage of renewable energy use in total energy use	continuous	The World Bank	https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS
fossil_fuel	percentage of fossil fuel consumption in total energy use	continuous	The World Bank	https://data.worldbank.org/indicator/EG.USE.COMM.FO.ZS