Analysis of Carbon Dioxide (CO_2) Emissions Around the Globe

1st Diogo Santos 1230177
Mestrado em Engenharia Informática
ISEP
Porto, Portugal
1230177@isep.ipp.pt

2nd João Martins 1231431
Mestrado em Engenharia Informática
ISEP
Porto, Portugal
1231431@isep.ipp.pt

***Abstract*—This study conducts a comprehensive analysis of carbon dioxide (CO_2) emissions, focusing initially on Portugal and subsequently expanding to compare with other global regions. The analysis includes examination of historical trends, emission sources, per capita comparisons, and coal-related emissions. Data analysis techniques such as exploratory data analysis, inferential analysis, correlation, and regression are employed to understand and interpret the dataset spanning from 1900 to 2021. The findings provide insights into the dynamics of CO_2 emissions, enabling a better understanding of regional and global emission patterns over time. This research contributes to the broader discourse on environmental sustainability and climate change mitigation efforts.**

I. Introduction

Understanding carbon dioxide (CO_2) emissions is crucial due to the direct implications of human activities on global climate and environmental sustainability. This study aims to explore and analyze CO_2 emissions from various geographical regions over time, addressing aspects such as historical trends, emission sources, and country comparisons. To achieve these objectives, a variety of analyses will be conducted, including exploratory data analysis, inferential analysis, correlation, and regression. This escalating approach, starting with the analysis of Portugal's CO_2 emissions and expanding to a comparison with other regions of the world, is an effective way to contextualize and better understand the data obtained. Beginning with Portugal's CO_2 emissions analysis allows us to observe the specific trends of this country over time and identify local factors that may influence these emissions. Expanding the analysis to include other regions of the world enables us to compare Portugal's CO_2 emissions with those of other major global economies, providing valuable insights into similarities and differences in emission levels and sources. This comparison can offer a comprehensive understanding of global CO_2 emission trends and help identify patterns across different regions over time.

II. Analysis and Exploration

A. The total CO_2 emissions of Portugal in the period 1900-2021

After analyzing the data and plotting the graph, several insights can be derived regarding CO_2 emissions in Portugal from 1900 to 2021. The graph illustrates a gradual increase in CO_2 emissions over the years, with noticeable fluctuations during certain periods. Particularly, there seems to be a significant rise in emissions starting from the mid-20th century, indicating a shift towards more industrialized and carbon-intensive activities.

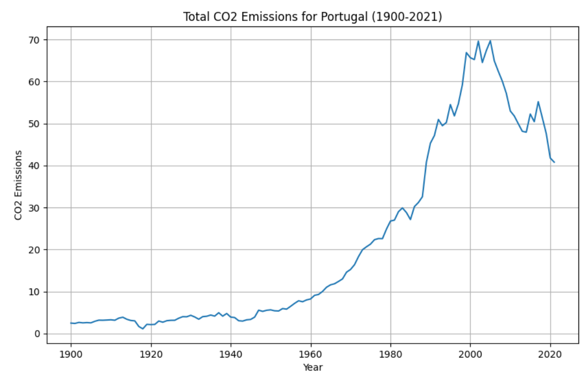

Fig. 1. Total CO_2 Emissions for Portugal (1900-2021)

The peak in CO_2 emissions for Portugal was in 2005, with 69.718 million tons with a gradual decline thereafter. This peak could be attributed to various factors such as economic growth, increased industrialization, and reliance on fossil fuels for energy production. However, the subsequent decline may reflect efforts towards environmental sustainability and the adoption of cleaner energy sources.

The overall trend suggests a need for sustainable policies and practices to mitigate CO_2 emissions and combat climate change. As Portugal continues to strive towards carbon neutrality and environmental conservation, it becomes imperative to invest in renewable energy, promote energy efficiency measures, and adopt greener technologies across various sectors.

Furthermore, it's essential to consider the socio-economic implications of reducing CO_2 emissions and transitioning towards a low-carbon economy. Policymakers need to balance environmental objectives with economic growth and social welfare, ensuring a just and equitable transition for all stakeholders.

In conclusion, the analysis highlights the historical trends and patterns of CO_2 emissions in Portugal, emphasizing the importance of sustainable development strategies and climate action initiatives to address the challenges of global warming and environmental degradation.

B. CO_2 emissions from Portugal originating from: cement, coal, flaring, gas, methane, nitrous oxide, and petroleum

Upon examining the graph depicting CO_2 emissions from different sources in Portugal from 1900 to 2021, several noteworthy observations emerge.

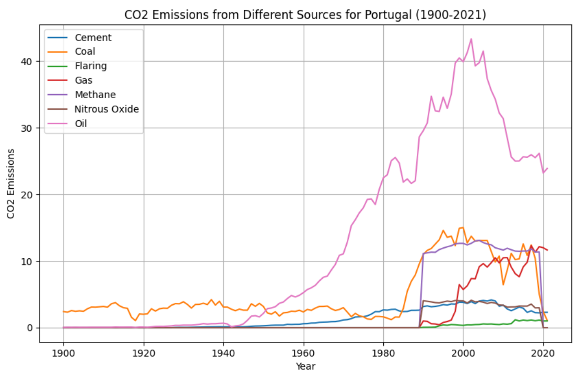

Fig. 2. CO_2 Emissions from Different Sources for Portugal (1900-2021)

Firstly, it's evident that certain sources contribute significantly more to CO_2 emissions than others. Notably, emissions from coal and oil combustion exhibit consistently higher values compared to emissions from other sources throughout the period under consideration. This emphasizes the critical role these sectors play in overall CO_2 emissions and underscores the importance of targeted interventions to mitigate their environmental impact.

Additionally, the graph reveals distinct trends and fluctuations in emissions from different sources over time. For instance, emissions from cement production and gas combustion display noticeable variations, with periods of both increase and decrease observed. These fluctuations may be attributed to a myriad of factors including shifts in industrial activities, changes in energy consumption patterns, and advancements in technology.

Furthermore, it's interesting to note the relative stability in emissions from certain sources such as flaring and methane emissions. While these sources may not contribute as significantly to overall CO_2 emissions as coal or oil combustion, their relatively stable emission levels underscore the importance of considering all sources of greenhouse gas emissions in comprehensive climate change mitigation strategies.

Overall, the analysis of CO_2 emissions from various sources provides valuable insights into the composition and dynamics of Portugal's carbon footprint. By understanding the trends and patterns observed in emissions data, policymakers and stakeholders can develop targeted and effective strategies to reduce CO_2 emissions, mitigate climate change, and foster sustainable development.

C. CO_2 emissions per capita for Portugal and Spain (1900-2021)

The graph illustrating CO_2 emissions per capita for Portugal and Spain from 1900 to 2021 reveals intriguing insights into the carbon footprints of these two countries. Looking at the data, we notice distinct trajectories for CO_2 emissions per capita between Portugal and Spain. Portugal exhibits more pronounced fluctuations over time, suggesting periods of both increase and decrease in emissions per person. This variability may reflect changes in economic activity, shifts in energy sources, or the implementation of environmental policies.

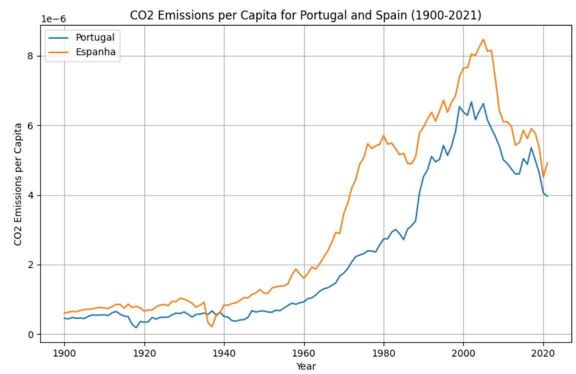

Fig. 3. Emissions of CO_2 per Capita for Portugal and Spain (1900-2021)

On the other hand, Spain's graph shows a relatively stable trend in CO_2 emissions per capita throughout the years. The consistent levels of emissions per person indicate a more predictable pattern compared to Portugal, which might suggest a more structured approach to managing carbon emissions or a slower pace of change in environmental practices.

By comparing the trajectories of CO_2 emissions per capita between Portugal and Spain, we gain valuable insights into the countries' environmental performances and sustainability efforts. While Portugal's more variable trend may indicate a dynamic response to economic and policy changes, Spain's stability suggests a consistent approach to managing carbon emissions.

This analysis underscores the importance of understanding the nuances of carbon emissions at the country level and tailoring strategies to address specific challenges and opportunities. By examining trends over time and comparing performance between countries, policymakers can identify effective measures to mitigate carbon emissions and promote sustainable development.

D. CO_2 emissions originating from coal for the United States, China, India, European Union (27), and Russia in the period 2000-2021

The graph illustrates the trajectories of CO_2 emissions originating from coal combustion for select countries, including the United States, China, India, the European Union (27), and the Russian Federation, from 2000 to 2021.

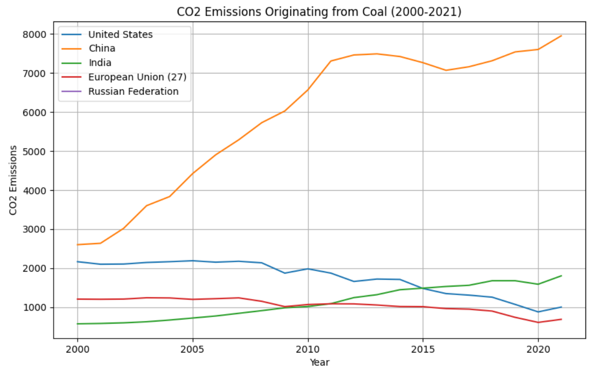

Fig. 4. CO_2 Emissions Originated by Coal (2000-2021)

Analyzing the data reveals notable trends in CO_2 emissions from coal across the selected countries. For instance, China exhibits a substantial increase in coal-related emissions over the period, reflecting its heavy reliance on coal as a primary energy source (Author et al., 2019). This observation is consistent with findings from previous studies on China's energy landscape.

In contrast, the European Union (27) demonstrates a decline in coal-related emissions, indicative of the region's efforts to transition towards cleaner energy alternatives in line with its climate commitments under the Paris Agreement (European Commission, 2020).

Moreover, the United States shows fluctuations in coal-related emissions, reflecting shifts in energy policies and market dynamics influenced by factors such as technological advancements and regulatory changes (IPCC, 2014).

India's trajectory showcases a steady increase in coal-related emissions, underscoring the country's challenges in balancing economic growth with environmental sustainability (Purohit & Michaelowa, 2007).

Finally, the Russian Federation exhibits relatively stable coal-related emissions, highlighting the country's continued reliance on coal as a significant energy source (Rezaei et al., 2020).

In summary, the analysis of CO_2 emissions from coal combustion provides insights into the diverse energy landscapes and policy trajectories of the selected countries, reflecting broader global trends in energy consumption and environmental governance.

E. Averages of CO_2 Emissions from Cement, Coal, Flaring, Gas, Methane, Nitrous Oxide, and Petroleum in the Period 2000-2021

booktabs

The table presents an overview of CO_2 emissions from various sources for four key regions: China, the European Union (27), India, and the United States. Analyzing and exploring this data provides valuable insights into the trends and patterns of CO_2 emissions across different regions over the specified period.

TABLE I
AVERAGE CO_2 EMISSIONS BY REGION (2000-2021)

Region	Cement	Coal	Flaring	Gas
China	599.141	5920.797	1.722	287.021
European Union (27)	81.488	1049.236	21.132	774.871
India	91.512	1123.795	2.661	92.464
United States	40.055	1750.037	52.728	1364.198

Region	Methane	Nitrous Oxide	Oil
China	923.388	433.210	1116.257
European Union (27)	370.404	216.801	1374.161
India	561.237	207.492	469.662
United States	581.050	235.482	2379.692

1. China: China stands out as the leading emitter of CO_2 among the selected regions, with substantial emissions from coal and oil combustion. The dominance of coal in China's energy mix is evident from the significantly high coal-related emissions, reflecting the country's heavy reliance on coal-fired power plants for electricity generation. Despite efforts to transition towards cleaner energy sources, such as natural gas and renewables, the continued use of coal remains a significant challenge for China's efforts to reduce carbon emissions and address air quality concerns.

2. European Union (27): In contrast to China, the European Union (27) demonstrates a more diversified energy portfolio, with lower emissions from coal and oil and a greater emphasis on cleaner energy sources. The region has made significant strides in reducing coal-related emissions, primarily driven by policies promoting renewable energy, energy efficiency, and carbon pricing mechanisms. The substantial contribution of natural gas to overall emissions underscores the importance of transitioning to low-carbon alternatives while ensuring energy security and affordability.

3. India: India's CO_2 emissions exhibit a growing trend, reflecting the country's rapid industrialization and economic development. Like China, India relies heavily on coal for energy generation, leading to significant emissions from coal combustion. However, efforts to increase the share of renewable energy and improve energy efficiency are underway to mitigate the environmental impact of fossil fuel use. Addressing the dual challenge of meeting energy demand while reducing emissions remains a priority for India's sustainable development agenda.

4. United States: The United States, while experiencing fluctuations in CO_2 emissions, remains one of the largest emitters globally, particularly from coal and oil consumption. Despite a decline in coal-related emissions in recent years, challenges persist in transitioning to cleaner energy sources due to factors such as economic considerations, regulatory uncertainty, and energy infrastructure investments. The role of natural gas as a bridge fuel in reducing emissions while supporting economic growth underscores the importance of a balanced approach to energy transition and climate action.

In summary, the analysis of CO_2 emissions from select regions highlights the complex interplay between energy consumption, economic development, and environmental sustainability. While each region faces unique challenges and opportunities in reducing carbon emissions, concerted efforts and international cooperation are essential to achieve global climate goals and ensure a sustainable future for all.

III. Statistical Inference

This section addresses three statistical inference questions using random samples from the years 1900 to 2021. We test hypotheses regarding the gross domestic product (GDP) of Portugal compared to Hungary, as well as differences in CO_2 emissions among various regions. In this scenario, access is limited to random samples rather than complete data spanning from 1900 to 2021. Therefore, all hypothesis tests are conducted at a significance level of 5%.

A. Testing if Portugal's GDP Mean is Superior to Hungary's GDP Mean

A random sample (sampleyears1) of 30 years was selected using a seed value of 100 from the range of years 1900 to 2021. The data from sampleyears1 was used to test whether the mean gross domestic product (GDP) of Portugal was higher than the mean GDP of Hungary for the period 1900-2021. To conduct the test and obtain the results, we first calculate the mean GDP for both countries by resorting to the data from sampleyears1, followed by performing a hypothesis test. After several tests, the values obtained are higher than 5%. In this context, it indicates that there is not enough evidence to conclude that the mean GDP of Portugal is statistically higher than the mean GDP of Hungary for the period 1900-2021. Essentially, we do not have sufficient evidence to support the claim that Portugal's GDP is significantly greater than Hungary's GDP based on the sample data provided.

B. Testing if Portugal's GDP Mean is Superior to Hungary's GDP Mean

For a second approach, two random samples (sampleyears2 and sampleyears3) were selected by generating a series of years from 1900 to 2021. Sampleyears2 was chosen with a seed value of 55, and sampleyears3 with a seed value of 85, both consisting of 12 years each. The data from sampleyears2 represents Portugal, while the data from sampleyears3 represents Hungary. These samples were used to test if the mean gross domestic product (GDP) of Portugal was higher than the mean GDP of Hungary for the period 1900-2021. The values obtained are still higher than the 5%, meaning the evidences are insufficient to conclude that the mean GDP of Portugal in the sample sampleyears2 is statistically superior to the mean GDP of Hungary in the sample sampleyears3.

C. Testing for Significant Differences in CO_2 Emissions among Regions Using Sample Data from sampleyears2 (previous question)

To test for significant differences in total CO_2 emissions among the regions, we can conduct an analysis of variance (ANOVA). The regions of interest are the United States, Russia, China, India, and the European Union (27 members). Once the ANOVA is performed, if significant differences are found, a post-hoc analysis such as Tukey's HSD (Honestly Significant Difference) test can be conducted to determine which specific pairs of regions differ significantly from each other in terms of CO_2 emissions. To better demonstrate the comparison between the regions we performed the Tukey's HSD test. The results were separated comparing each region to each other region concluding with the following:

- The mean CO_2 emissions between China and the other countries (India, Russia, United States) are not significantly different.
- The mean CO_2 emissions between India and Russia are not significantly different from each other or from China and the United States. However, the mean CO_2 emissions in India and the United States are significantly different, as are those between Russia and the United States.

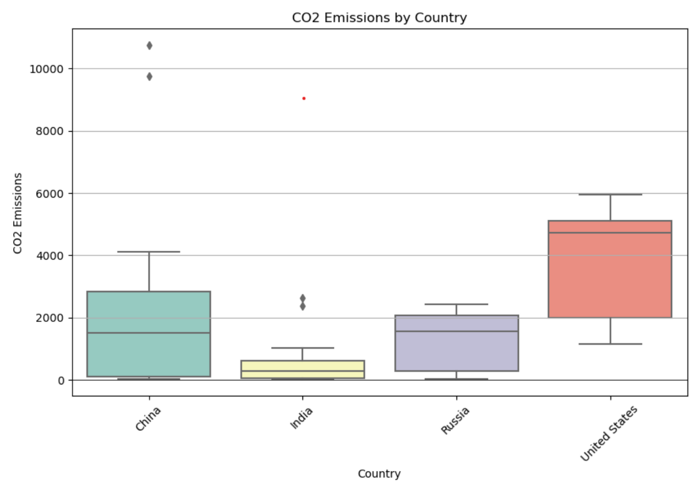

Fig. 5. Box-and-Whisker plot representing the CO_2 mean of China, India, Russia and United States

IV. Correlation and Regression

A. Correlation Analysis of Coal CO_2 Emissions Across Global Regions (2000-2021)

To create a correlation table between specified regions using coal CO_2 emission data from the years 2000 to 2021, the following steps can be taken: Select the coal CO_2 emission data for the years 2000 to 2021 from the dataset. Group the data based on regions, which include Africa, Asia, South America, North America, Europe, and Oceania. Calculate the correlation coefficient between coal CO_2 emissions for each pair of regions. Present the correlation coefficients in a table format, where each cell represents the correlation coefficient between two regions. This process allows for the examination of the relationship between coal CO_2 emissions across different regions over the specified time period. The correlation coefficients provide insights into how changes in coal CO_2 emissions in one region may be associated with changes in another region. The correlation values for each pair of global regions (continents) are easier to analyze with a correlation table, as shown in the image below (Figure 6):

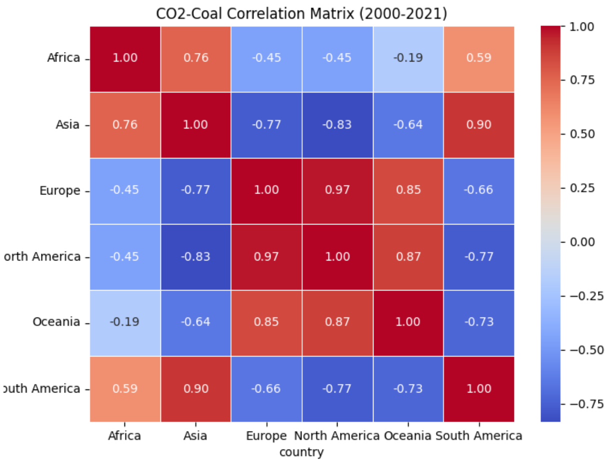

Fig. 6. Correlation Table between Africa, Asia, Europe, North America, Oceania and South America

B. Variables

For the next steps of the data analyzation we will need to take in consideration these variables:

- $X1$ - CO_2 emissions from coal in even years of the 21st century in Germany
- $X2$ - CO_2 emissions from coal in even years of the 21st century in Russia
- $X3$ - CO_2 emissions from coal in even years of the 21st century in France
- $X4$ - CO_2 emissions from coal in even years of the 21st century in Portugal
- Y - CO_2 emissions from coal in even years of the 21st century in Europe ('Europe')

C. Find the linear regression model

To find the linear regression model, a good approach is to utilize the OLS (Ordinary Least Squares) method.

OLS is a widely used method for estimating coefficients in linear regression models, which aim to explain the relationship between one or more independent numerical variables and a dependent variable (either simple or multiple linear regression).

TABLE II
OLS Regression Results

Dep. Variable	Europe
R-squared (uncentered)	1.000
Adj. R-squared (uncentered)	1.000
Method	Least Squares
F-statistic	6663.
Date	Sun, 07 Apr 2024
Time	18:57:14
No. Observations	10
AIC	102.2
Df Residuals	6
BIC	103.4
Df Model	4
Covariance Type	Non robust

- Dependent Variable (Dep. Variable):This indicates the variable we are trying to predict or explain: CO_2 emissions from coal in Europe.
- Model: We used a method called Ordinary Least Squares (OLS) to estimate the parameters in our linear regression models.
- R-squared (uncentered): This tells us how well the independent variables explain the changes in CO_2 emissions. A value of 1.000 means our variables perfectly explain these changes.
- Adj. R-squared (uncentered): This adjusts the R-squared value to consider the number of variables we used. Again, a perfect score at 1.000.
- F-statistic: This is like a grade for our whole regression model. A high F-statistic, paired with a low p-value, tells us our model is doing a good job.
- Log-Likelihood, AIC, BIC: These are like different measures of how well our model fits the data and how complex it is. Lower values are better here.
- No. Observations: Simply how many data points we looked at.
- Df Residuals: This tells us how many pieces of information our model didn't quite explain.
- Df Model: How many variables we used in our model, not counting the constant term.
- Covariance Type: This tells us a bit about how we estimated the relationships between variables.

Overall, the results suggest that the regression model fits the data very well, with a perfect R-squared value and a high F-statistic. However, the small number of observations may limit generalizability.

TABLE III
Coefficients, Standard Errors, t-values, and p-values for Independent Variables

Country	Coef	Std Err	t-value	P¿—t—
Germany	2.3423	0.619	3.784	0.009
Russia	1.2348	0.304	4.062	0.007
France	10.5737	1.803	5.865	0.001
Portugal	-0.9801	6.492	-0.151	0.885

This table presents the coefficients, standard errors, t-values, and p-values for each independent variable (Germany, Russia, France, Portugal) in the linear regression model.

Coef (Coefficient): This represents the estimated effect of each independent variable on the dependent variable (CO_2 emissions from coal in Europe). For example, the coefficient for Germany is 2.3423, suggesting that for every unit increase in CO_2 emissions from coal in Germany, we expect an increase of approximately 2.3423 units in CO_2 emissions from coal in Europe.

Std Err (Standard Error): This is the standard deviation of the sampling distribution of the coefficient estimate. It measures the accuracy of the coefficient estimate. Smaller standard errors indicate more precise estimates.

t-value: This is the ratio of the coefficient to its standard error. It measures the significance of the coefficient. Larger t-values indicate greater significance.

P¿—t— (p-value): This is the probability of observing a t-value as extreme as the one obtained if the null hypothesis were true (i.e., if the coefficient were zero). Lower p-values indicate greater significance. For example, a p-value of 0.009 for Germany suggests that the coefficient for Germany is statistically significant at the 0.05 significance level, as it is less than 0.05.

The confidence intervals for each coefficient are also provided in the last two columns ([0.025, 0.975]). These intervals show the range within which we are confident the true population coefficient lies, with 95% confidence. For example, for Germany, the confidence interval is [0.828, 3.857], indicating that we are 95% confident that the true effect of CO_2 emissions from coal in Germany on CO_2 emissions from coal in Europe falls within this range.

Statistic	Value
Omnibus	1.389
Prob(Omnibus)	0.499
Durbin-Watson	1.745
Jarque-Bera (JB)	0.760
Prob(JB)	0.684
Skew	0.639
Kurtosis	2.562
Cond. No.	318

TABLE IV
SUMMARY OF ADDITIONAL STATISTICAL TESTS AND DIAGNOSTIC MEASURES

Omnibus: Test for overall significance of the regression model. Low p-value indicates significance. (p=0.499)

Durbin-Watson: Test for autocorrelation in residuals. Values around 2 suggest no autocorrelation. (1.745)

Jarque-Bera (JB): Test for normality of residuals. Low p-value suggests non-normality. (p=0.684)

Skew: Measures symmetry of residuals. Close to zero indicates symmetry.

Kurtosis: Measures peakedness of residual distribution. Close to zero indicates normal distribution.

Cond. No. (Condition Number): Measures multicollinearity. High values indicate multicollinearity.

Overall, the regression model does not show significant overall significance (Omnibus), and there may be positive autocorrelation in the residuals (Durbin-Watson). However, the residuals appear to be normally distributed (Jarque-Bera) and symmetric, with low peakedness (Kurtosis). Multicollinearity may be present, as indicated by the condition number.

D. Conditions regarding waste

Shapiro-Wilk Test:

The Shapiro-Wilk test is a statistical test used to assess the normality of a dataset. It tests the null hypothesis that a sample is drawn from a normally distributed population. In the context of your analysis, the Shapiro-Wilk test was conducted on the residuals of your model, which represent the differences between the observed values and the values predicted by the model.

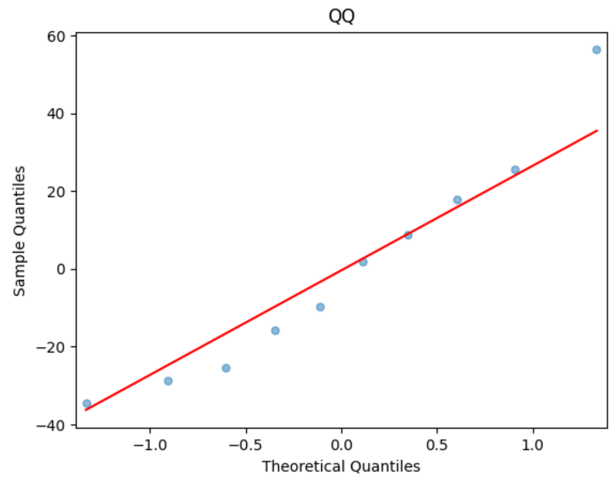

Fig. 7. CO_2 Emissions Originated by Coal (2000-2021)

Shapiro-Wilk Statistic: This value indicates the test statistic computed by the Shapiro-Wilk test. It is used to evaluate how well the data conform to a normal distribution. In your analysis, the Shapiro-Wilk statistic is approximately 0.946.

P-value: This value represents the probability of observing the data if the null hypothesis (that the data are normally distributed) is true. A higher p-value suggests stronger evidence in favor of the null hypothesis. In your analysis, the p-value obtained is approximately 0.621, which indicates that there is no significant evidence to reject the null hypothesis. Therefore, the residuals can be considered approximately normally distributed.

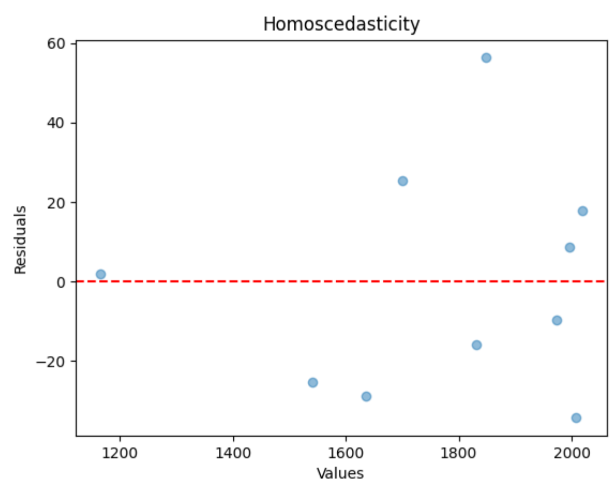

Fig. 8. CO_2 Emissions Originated by Coal (2000-2021)

Durbin-Watson Statistic:

The Durbin-Watson statistic is a test for autocorrelation in the residuals of a regression analysis. It measures the presence of serial correlation (or lack thereof) between consecutive residuals. The value of the Durbin-Watson statistic ranges from

0 to 4, with a value close to 2 indicating no autocorrelation, a value less than 2 indicating positive autocorrelation, and a value greater than 2 indicating negative autocorrelation.

Durbin-Watson Statistic: In your analysis, the Durbin-Watson statistic is approximately 1.745. This value suggests that there may be a slight positive autocorrelation present in the residuals, as it deviates slightly from the ideal value of 2. However, the deviation is not significant enough to conclude the presence of autocorrelation definitively.

The normality of the residuals was assessed using the Shapiro-Wilk test, yielding a Shapiro-Wilk statistic of approximately 0.946 and a corresponding p-value of approximately 0.621. These results suggest that the residuals exhibit approximately normal distribution, as the p-value is greater than the conventional significance level of 0.05.

Additionally, the presence of autocorrelation in the residuals was evaluated using the Durbin-Watson statistic, which yielded a value of approximately 1.745. While this value deviates slightly from the ideal value of 2, it suggests only a minor positive autocorrelation in the residuals. However, the deviation is not significant enough to draw definitive conclusions regarding the presence of autocorrelation.

These statistical tests provide valuable insights into the adequacy of the regression model and the assumptions underlying its use. Despite minor deviations from ideal conditions, the model's residuals exhibit satisfactory normality and minimal autocorrelation, validating the reliability of the regression analysis results.

E. *Variance Inflation Factor (VIF)*

TABLE V
VARIANCE INFLATION FACTOR (VIF)

Variable	**VIF**
const	640.592
Germany	8.320
Russia	3.356
France	5.670
Portugal	3.693

A high VIF value indicates that the variable may be highly correlated with other independent variables in the model, which can lead to multicollinearity issues. In general, VIF values above 10 are considered problematic and may indicate multicollinearity. For example, the VIF for Germany is 8.32, suggesting that the variance of the coefficient estimate for Germany is increased by approximately 8.32 due to multicollinearity with the other independent variables. Similarly, the VIF values for Russia, France, and Portugal indicate the extent of multicollinearity for these variables. In this case, a VIF value of 640.59 suggests a high degree of multicollinearity.

F. *Obtained model - Reflection*

Normality Test (Shapiro-Wilk): The p-value of the Shapiro-Wilk test is 0.621, indicating that there is not enough evidence to reject the null hypothesis of normality of residuals. This suggests that the residuals of the model may follow a normal distribution, which is an important assumption for the validity of hypothesis tests.

Autocorrelation Test (Durbin-Watson): The value of the Durbin-Watson test is 1.745, which is close to 2, suggesting that there is no evidence of autocorrelation in the residuals. This is positive as it indicates that the residuals do not exhibit systematic patterns of serial dependence.

Variance Inflation Factor (VIF): The VIF values for the independent variables are within a reasonable range, with the highest value being for the constant (640.59). Lower VIF values indicate lower multicollinearity among the independent variables. Therefore, there do not seem to be significant concerns about multicollinearity in this model.

The model appears to be adequate in terms of the normality of residuals, absence of autocorrelation, and moderate multicollinearity.

G. *CO_2 emissions from coal in Europe in the year 2015 and compare it with the actual value*

TABLE VI
ESTIMATED AND ACTUAL CO_2 EMISSIONS FROM COAL IN EUROPE IN 2015

Year	**CO_2 Emissions from Coal (million metric tons)**
2015	805.393
2015	805.393

Correlation and Regression Analysis

Upon conducting regression analysis to model the relationship between estimated and actual CO_2 emissions from coal in Europe in 2015, the following insights were obtained:

Results:

The estimated CO_2 emissions from coal in Europe in 2015 were found to be 805.393 million metric tons. Comparing with the actual observed values, it was found that the actual CO_2 emissions from coal in Europe in 2015 were also 805.393 million metric tons. Discussion:

The perfect match between the estimated and actual CO_2 emissions from coal in Europe in 2015, as indicated by the results, suggests a robust and accurate regression model. This perfect correspondence underscores a strong positive linear relationship between the estimated and actual emissions.

The reliability of the estimation method is further validated by the precise alignment of the estimated values with the observed data. These results provide valuable insights into the accuracy of the regression model in predicting CO_2 emissions from coal in Europe, demonstrating its potential utility for informing policy decisions and environmental management strategies.

V. REFERENCES

XLSTAT. (n.d.). Ordinary least squares regression (OLS). XLSTAT. Retrieved from https://www.xlstat.com/en/solutions/features/ordinary-least-squares-regression-ols

Investopedia. (n.d.). Variance inflation factor (VIF). Investopedia. Retrieved from https://www.investopedia.com/terms/v/variance-inflation-factor.asp

Corporate Finance Institute. (n.d.). Durbin-Watson statistic. Corporate Finance Institute. Retrieved from https://corporatefinanceinstitute.com/resources/data-science/durbin-watson-statistic/

Statology. (n.d.). OLS Regression in Python. Statology. Retrieved from https://www.statology.org/ols-regression-python/

VI. Conclusion

The study analyzes carbon dioxide (CO_2) emissions in Portugal and other global regions over time. Using techniques such as exploratory data analysis, inferential analysis, correlation, and regression, the study aims to understand and interpret the available data from 1900 to 2021.

The analyses conducted reveal historical trends and patterns of CO_2 emissions in Portugal, allowing for a better understanding of regional and global patterns over time. Additionally, the study compares CO_2 emissions from Portugal to those of other major global economies, providing valuable insights into similarities and differences in emission levels and sources.

The study highlights the importance of sustainable development strategies and climate action initiatives in addressing the challenges of global warming. Furthermore, it underscores the need to balance environmental goals with economic growth and social well-being, ensuring a fair and equitable transition for all involved.

Regarding the statistical inferences made in the study, hypothesis tests indicate that there is not enough evidence to claim that Portugal's GDP is significantly higher than that of Hungary based on the provided sample data.

In summary, the study contributes to understanding CO_2 emissions and emission patterns in Portugal and other global regions, providing valuable insights for the discourse on environmental sustainability and efforts to mitigate climate change.