A Data Mining Approach to Analyzing World Energy

Omar Ammar, Hevra Petekkaya, Maryam Imran,

Abstract—**In this paper, we introduce our topic for the data mining course, which is to work with a world energy dataset. The world energy dataset consists of various indicators related to energy consumption, production, and access, among other factors. The primary objective of this project is to mine this dataset to gain insights into energy patterns and trends across countries and regions. We plan to use various data mining techniques such as data preprocessing, clustering, and association rule mining to extract valuable information from the dataset. Our project will also involve visualization techniques to help us understand and communicate the results of our analysis. Through this project, we hope to contribute to the global discourse on energy and provide useful insights for policymakers, researchers, and the general public.**

Index Terms—**Data Mining, Clustering, Classification, Energy, CO2**

I. Introduction

This is our project proposal as part of our CNG 514 project for the spring semester 2022/2023

A. Problem Statement

The aim of this study is to address the problem of the effect of energy consumption and energy conservation on the economic growth of various geographic regions. Does conservation of energy affect production? Do CO2 emissions correlate with increased exports? Does energy export support or hamper energy conservation goals? These are the kind of questions that need to be answered in the process of analyzing energy usage and conservation of different countries to be able to gain any meaningful conclusions from the data.

The transportation sector is one of if not the, most important industries for any country with respect to economic growth as well as energy consumption. This means that there might be a direct correlation between the economic prosperity of a region and its energy consumption, mainly due to the increase in energy consumed by the transportation sector as it expands to support a growing economy. As has been stated time and again, for the world to stay within the safety threshold of average temperature increase, the transport sector needs to be de-carbonized [1]. This indicates two things. One, that the transport sector plays a major enough role in energy consumption that controlling just this industry would have a significant global impact. And secondly, the increase in CO2 emissions correlates with economic growth since that correlates with the expansion of the transportation network.

Given that we know that energy consumption and economic growth are correlated, as we explore the patterns energy patterns hidden in our dataset, we will get closer to an answer to how significant this correlation is and weather it is affected by energy conservation measures in any non-negligible manner. We will also explore how carbon emissions fit into the big picture and if measures to curb them have a positive or negative impact on economic prosperity. Questions such as this and more are the target of this research and we hope to find some interesting and maybe unexpected patterns.

B. Literature Review

A total of 6 research papers are planned to be read and reviewed that are related to energy consumption patterns of countries that we believe will guide us throughout this project.

The first paper to be studied is ”The Drivers of Chinese CO2 Emissions from 1980 to 2030” [5]. The paper discusses China’s economic growth and the resulting increase in energy consumption and greenhouse gas emissions. The authors argue that it is crucial to understand the key drivers of China’s growing energy consumption and greenhouse gas emissions to develop effective climate and energy policies. They use the IPAT model to assess the key drivers of China’s CO2 emissions and to develop three scenarios for potential levels of CO2 emissions in 2030. Examining this paper, which provides an in-depth analysis of the drivers of CO2 emissions in China, can serve as a valuable reference to guide the analysis of energy consumption patterns and CO2 emission contributors in other countries.

The second paper, ”Consumption-based accounting of CO2 emissions” [4], is about the consumption-based accounting of CO2 emissions that takes into account the emissions associated with producing goods and services that are consumed in one country but occur in another country due to the global nature of trade and supply chains. This approach considers the entire life cycle of the products and services consumed, highlighting the need for global cooperation and responsibility to address climate change. Studying the concept of consumption-based accounting of CO2 emissions highlighted in the paper can provide a more comprehensive understanding of the indirect CO2 emissions associated with a country’s consumption patterns, which is important in identifying the main CO2 contributors. Moreover, it will provide valuable insights into the techniques and approaches required for data mining projects to extract meaningful insights beyond mere numerical figures and comprehend the underlying context.

Moreover, the paper ”Global and regional drivers of accelerating CO2 emissions” [3] analyzes the increase in global CO2 emissions due to rising energy intensity, carbon intensity of energy, population growth, and rising per capita GDP since 2000.

This research paper holds significant relevance to our study as it systematically analyzes the impact of critical factors, such as the energy intensity of Gross Domestic Product (Energy/GDP) and carbon intensity of energy, on global CO2 emissions. Additionally, since these attributes/factors and other similar ones also appear in our dataset, this paper can provide valuable insights into their utilization and facilitate the extraction of meaningful insights from them.

Additionally, the paper "Modeling Economic Growth and Energy Consumption in Arab Countries: Cointegration and Causality Analysis" [6] uses an Auto Regressive Distributed Lag (ARDL) model to determine the relationship between energy consumption and real economic growth in 17 Arab countries from 1980-2011. The study finds that the neutrality hypothesis is supported in 16 out of 17 countries, meaning that energy conservation will not have a significant impact on economic growth and economic growth will have an insignificant effect on changes in energy consumption. The critiques raised about the paper's results suggest that further research is needed to validate its findings. Additionally, the fact that Arab countries export most of their energy does not necessarily imply that energy conservation cannot be implemented. It is possible to conserve energy while exporting and extracting it. Finally, the availability of relevant attributes in your dataset presents an opportunity to test the hypotheses regarding the relationship between energy consumption and economic growth in Arab countries.

As another paper, "Assessments of primary energy consumption and its environmental consequences in the United Arab Emirates" [7] examines the United Arab Emirates' (UAE) high energy consumption and resulting carbon emissions over the past two decades due to rapid economic growth and urbanization. The study suggests key remedial measures, such as public awareness campaigns, utilizing renewable energy, and implementing carbon sequestration technology, to mitigate the country's energy consumption and environmental impact. Given the hypotheses surrounding the relationship between CO2 emissions and economic growth, and the presence of relevant attributes in our dataset, analyzing this paper's findings could provide valuable insights for us while trying to test our own hypotheses.

Finally, the paper "The effect of renewable energy consumption on economic growth: Evidence from top 38 countries" [8] analyzes the impact of renewable energy consumption on the economic growth of 38 top consuming countries between 1991 and 2012. Panel estimation techniques reveal that renewable energy consumption has a significant positive impact on economic output for 57% of the countries. The study highlights the need for increased renewable energy investment to promote low-carbon growth. Given the relevance of our dataset's attributes on renewable energy and GDP, investigating this paper's findings on the effects of renewable energy consumption on economic growth in top renewable energy-consuming countries can offer valuable insights into the role of renewable energy in promoting sustainable economic development and reducing CO2 emissions.

C. Hypothesis and Methodology

In this project, we will be using a publicly available dataset on global energy consumption . The first step will be to preprocess the data and remove any missing or noisy data points. We will then explore different clustering algorithms to analyze the data and identify any patterns or trends. Specifically, we will be using unsupervised clustering algorithms such as k-means clustering and hierarchical clustering to group countries based on their energy consumption and economic development.

After clustering, we will conduct a statistical analysis to investigate the relationship between energy consumption and economic development in various countries and regions. We will use appropriate statistical techniques to identify any correlations between the two variables and determine the strength of the relationship. as well as visualize our findings using data visualization techniques where necessary. Finally, we will evaluate the performance of the clustering algorithms using appropriate metrics and assess the validity of our findings. we will be testing two hypotheses related to global energy consumption and environmental goals. The first hypothesis is whether countries that have promised to switch from fossil fuels and reduce carbon emissions are sticking to their environmental goals. We hypothesize that there is a discrepancy between the promises made by countries and their actual progress towards reducing carbon emissions. We expect to find evidence of countries falling short of their environmental goals, which could have significant implications for global efforts to combat climate change.

The second hypothesis is related to the relationship between energy consumption and economic development in various countries and regions. We hypothesize that third world countries will be the highest contributors to CO2 emissions due to their high levels of energy consumption and limited access to clean energy sources. We expect to find evidence of a positive correlation between energy consumption and economic development in developing countries, which could highlight the need for increased investment in clean energy and sustainable development in these regions.

D. Dataset Description

This dataset, obtained from the open-source platform Kaggle [2], provides a comprehensive view of global energy generation and consumption patterns. It covers a period spanning from 1990 to the end of 2020 and comprises annually collected data from countries across the world. EnerDATA, a prominent organization in the energy sector, is the main contributor to this dataset, ensuring the data is meticulously collected and curated.

The dataset includes a diverse range of columns that describe different aspects of energy generation and consumption, Specifically, the dataset contains the following columns: country, year, region, CO2 emissions from fuel combustion, average CO2 emission factor, CO2 intensity at constant purchasing power parities, total energy production, total energy consumption, the share of renewables in electricity production, the share of electricity in total final energy consumption, oil products domestic consumption, refined oil products production, natural gas production, natural gas domestic consumption, the energy intensity of GDP at constant purchasing power parities, electricity production, electricity domestic consumption, coal and lignite domestic consumption, the share of wind and solar in electricity production, crude oil production, and coal and lignite production.

E. Project Timeline

Looking at the dataset we have and the hypothesis that we are working on, the project would hopefully be done 2 weeks before the end of the semester. In the following 2 weeks, we will be working on data pre-processing. We will identify the outliers in the data and the noise. Next, clean the data and prepare the dataset for further analysis. We are going to split the work after that among us and try to use multiple data Mining techniques to achieve the results we are hoping for (or negate our Hypothesis). We will split the work among us equally where each of us will handle a different approach to the solving and then compare our approaches to come up with the best methodology.

F. Risks Associated

Conducting a data mining study on a world energy dataset carries several inherent risks. Firstly, the quality and reliability of the dataset itself pose a significant challenge. Energy data from different countries and sources may vary in terms of accuracy and completeness. Additionally, data mining studies can produce correlations or patterns that may not necessarily imply causation, leading to misleading interpretations or flawed policy recommendations. Therefore, careful statistical analysis and domain expertise are crucial to avoid drawing erroneous conclusions from the data. Finally, ethical considerations must be addressed, particularly when using data mining techniques for predictive modeling, as unintended consequences or discriminatory outcomes could arise from biased algorithms or unfair data sampling.

II. Data Understanding & Hypothesis Discussion

A. Data Pre-processing & Understanding

Discovering and pre-processing data are essential steps in data analysis and machine learning research. Before diving into detailed analysis, it is crucial to explore the dataset's structure, contents, and limitations. This helps us understand the variables, their relationships, and any hidden patterns. After discovery, we need to clean, transform, and format the data during pre-processing. This ensures accuracy, consistency, and suitability for further analysis through careful data discovery and pre-processing. First, let us discuss the dimensions and spread of our data. As it can be seen in the following Diagram (Fig. 1.), which shows the countries and the regions studied in this report, each region is color coded in a certain color.

Our dataset includes the energy consumption patterns for these countries between the dates 1990-2020. it includes details about general consumption, shares of different energy resources (renewable or not) as well as comparison with the sale price of energy unit in the country. In Fig 2. we can see how many countries are studied per region. Also, we can see that some regions are under-represented in our dataset; while, others include the full extent of countries in that region. It is crucial to be aware of such data distributions because it can possibly affect our interpretations of the results we get.

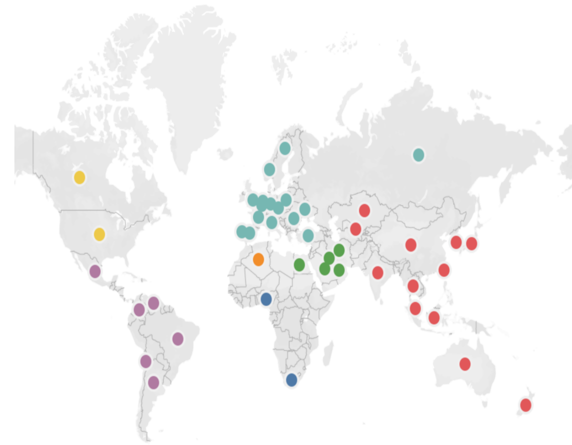

Fig. 1. Countries Studied

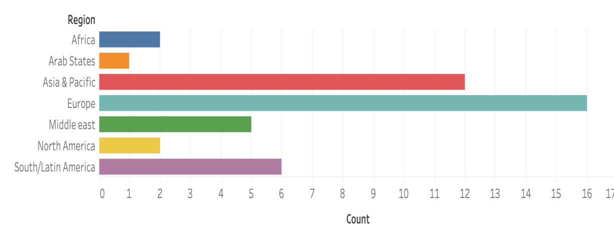

Fig. 2. Countries per Region

The data cleaning process for the dataset provided by the dataset provider involved several steps to ensure its quality and reliability. Initially, the dataset was carefully inspected to identify any missing or incomplete values. Any missing data points were either imputed or, if necessary, the corresponding entries were removed. Next, data outliers or anomalies were identified and addressed by applying appropriate techniques, such as statistical analysis or domain knowledge. Inconsistencies in the dataset, such as formatting issues or duplicate entries, were resolved by standardizing the data format and removing any duplicate records. Lastly, the dataset was thoroughly reviewed to ensure data integrity and accuracy. Through these cleaning steps, the dataset was refined and made ready for subsequent analysis and exploration. Due to the short space on this report we have decided not to include every column analysis in our data.

Our dataset excluding the names of the countries and the regions is numerical. In other words, there are no categorical or ordinal dimensions. Additionally, there are around 1364 entries in the dataset.

To detect the outliers in our dataset, we have plotted multiple box diagrams as well as implemented isolation forest algorithm. We can see from analyzing our dimensions that lots of the data is classified as outliers. This is not because this data is mistakenly taken or extreme. Rather, it is because countries vary quite drastically in their energy production, consumption, and energy usage patterns. Thus, we can not judge countries by their average or mean. This discards the false assumption that our data is full of outliers. Extreme outliers may have a different discussions, but from the numbers that we are having right now we can see that our data does not include such values. Fig. 3, which illustrates some of the block diagrams that we have obtained from our data study, can be examined.

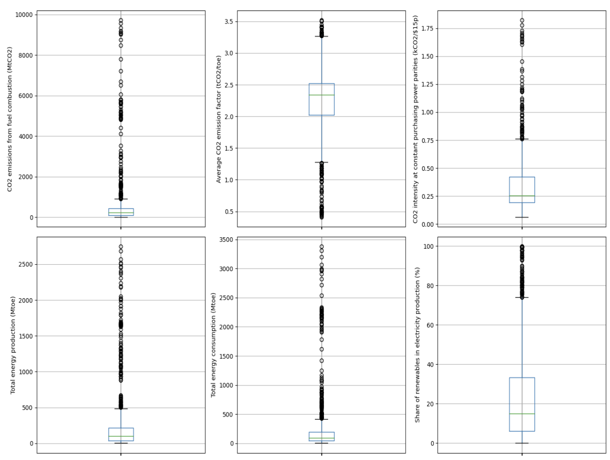

Fig. 3. Box Plots of some of our dimensions.

The final part of our Dataset study includes the section regarding checking whether the rest of our data makes sense. This included the manual and tedious analysis of our dataset tuples to ensure the numbers in it are fitting the use case we are discussing. We can now finally move onto discussing our hypothesis and approach towards proving them.

B. First Hypothesis: Countries and Promises

As we have mentioned in the Hypothesis and Methodology part, we will run data mining techniques on our hypothesis to test it. In order to investigate our hypothesis that countries in our dataset have not been adhering to their environmental commitments, we have visualized and studied the relationship between four main variables and time; Share of renewable energy in overall consumption, Share of fossil fuels in overall consumption, Average CO2 emission factor, and Average total energy consumption. These dimensions will help us understand better the patterns and behaviors of countries and would show us the outcome of the environmental promises.

We have decided to use regression techniques alongside correlation analysis and plotting the data. Below you can see the region wide analysis for each dimension vs time. Below you can see the initial visualizations of these relationships. Fig. 4. shows how are some regions reacting to renewables with time. but for our purpose it would be hectic to focus on every country in this dataset. Thus we will focus on 3 countries each from a different region. Our choice fell on these countries (United States - Latin America region, India - Asia and pacific Region , Russia - European Region) Due to the fact that they have promised to reduce their carbon footprint and become more Environmentally sustainable by 2030. and Due to them being some of the biggest countries that are major contributors to CO2 Emissions world wide according to an article by ClimateTrade.com in 2021[9].

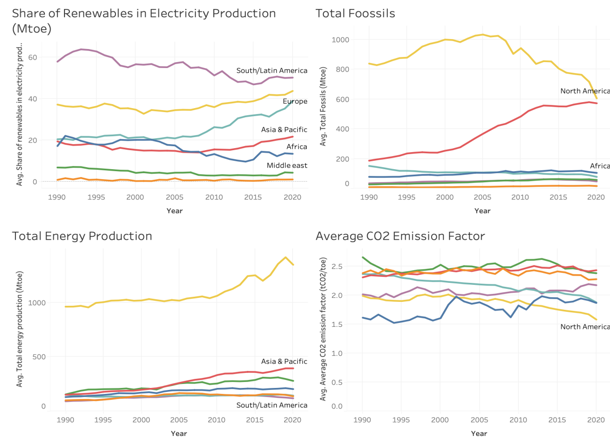

Fig. 4. Initial Visualizations of Energy Patterns/Region.

We will study these countries energy patterns comprehensively in the following paragraphs. However, before that, let us first discuss the most important factor, which is the CO2 Emissions factor. Fig. 5 can be examined to see how these countries acted upon their promises.

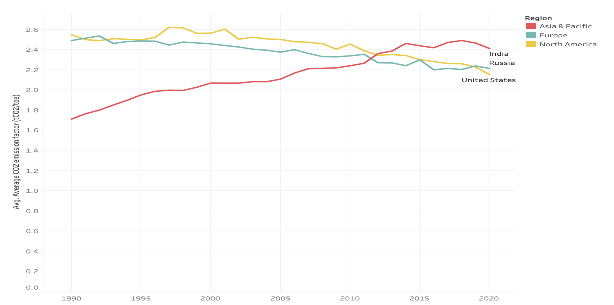

Fig. 5. side by side comparison of CO2 emissions from our studied countries.

We can see from Fig. 5. how USA and Russia have successfully reduced their CO2 emission factor. While India has drastically increased it. However, this data isn't sufficient by itself to explain why would a country like India increase its CO2 emissions which goes against its goals. It is worth mentioning that a country like India is a rapidly growing country with an exploding population. Between the years 1990 till 2020, India's population has increased from around 685 Million to around 1,395 Billion people, which is more than the US and Russia's Population Combined according to World-Data.info [10]. Thus we have to also study the implementation of renewable resources in India closely to justify this discrepancy. But first, we have to discuss how strong the relationship is between the Average Share of Renewable energy production and the Average CO2 emissions factor.

In Fig. 6, we observe a negative trend line, which indicates a negative relationship with the R-Squared value of 0.607641.

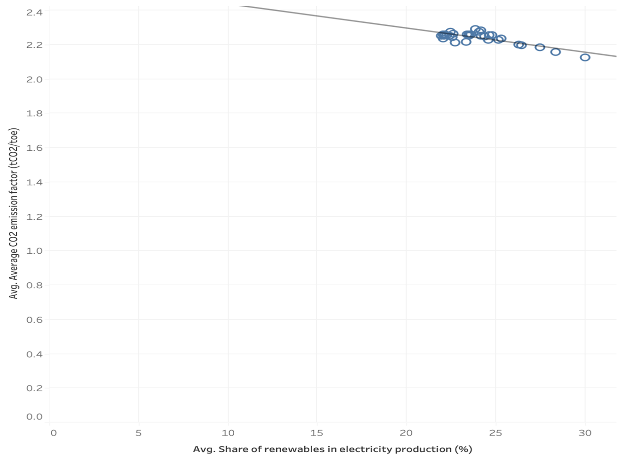

Fig. 6. Average Renewables Share in Total Energy Production vs Average CO2 Emissions Factor Worldwide

Furthermore, since the value of $R > 0.6$ and we have a negative slope that can suggest a very strong relationship. If electricity was the only form of energy to be considered, then increasing the share of renewable in electricity production would drastically decrease the CO2 coefficient factor even with an overall energy output increase due to population increase like in India's case. This would suggest that if India was implementing sufficient renewable policies it would help them reduce their carbon footprint. However, electricity is not the only form of energy present as can be seen in Fig. 7. In fact, at its maximum usage, it contributes by 17% to the total energy consumption. This makes it clear that using renewable energy for only electricity production is not enough for reducing CO2 emissions because there are other sources of CO2 emissions.

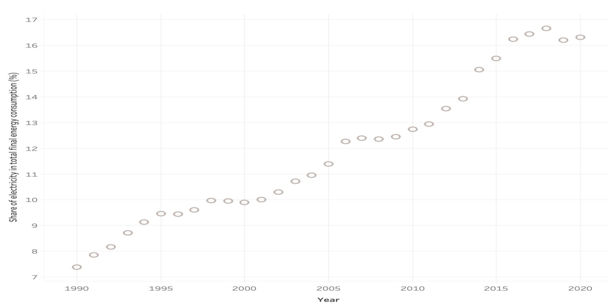

Fig. 7. Share of electricity in total final energy consumption in India

Let us see how did India change its share of Renewable in the past 30 years.

In Fig. 8, we can see the Share of renewable energy in India between the years 1990-2020. The renewable energy share of total energy production was declining between 1990 and 2003 then it started to increase again, However, it did not fully recover to the level it was at before in 1990. Although India has promised to reduce its carbon footprint, it has struggled to do so. There might be many reasons for this discrepancy which is outside the scope of this paper which just aims to prove the existence of such facts. However, it's worth mentioning that during the study of our data, we have noticed that India has

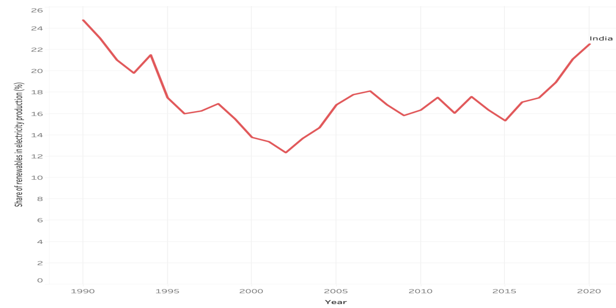

Fig. 8. Renewable Share in Total Energy Production in India (1990-2020)

increased its share of Solar/Wind energy between 1990 to 2020 from 0% up to 8%. which goes along with their promises to implement cleaner energy policies. Thus we can suggest that although India is trying to stick to their environmental promises the drastically increasing population did not give India the chance to implement these policies fast enough thus the discrepancy between the promises and the data.

C. Second Hypothesis: Economic factors and Renewable energy policy implementation

Energy systems and sustainability have become critical topics in the global discourse, as countries strive to transition towards cleaner and more renewable sources of energy while ensuring economic growth. Understanding the relationship between key energy indicators can provide valuable insights for policymakers and stakeholders. In this study, we investigate the relationship between the share of renewables in electricity production and the energy intensity of GDP at constant purchasing power parities. The analysis aims to uncover potential patterns and clusters among countries, shedding light on the interactions between renewable energy adoption and economic factors. By applying clustering techniques to a comprehensive dataset spanning the years 1990 to 2020, we explore distinct groups of countries sharing similar characteristics and identify the contributing factors that drive their positioning within these clusters.

We suggest that countries where Renewables have affected the purchasing power negatively i.e. electricity becomes cheaper by implementing new renewable energy resources. while others do the opposite or have no change in their economic policies. First, we have clustered our countries based on the similarity of countries in terms of their share of renewables in electricity production and the energy intensity of GDP at constant purchasing power parities. which helped us group the countries in 3 different clusters. You can check Table 1 in the appendix for the full list of countries in each cluster. Below we have the figures for each cluster.

We can see in Fig. 8. That all the countries in this cluster have a very faint relationship between Energy Purchasing power and Share of renewables. Although renewables offer cheaper and cleaner options. The countries in this cluster with increasing Share of renewables They generally still sell energy units at the same price of fossil fuels.

Similarly countries in Cluster 1 have the same traits between the two mentioned variables. However, The pearson correlation

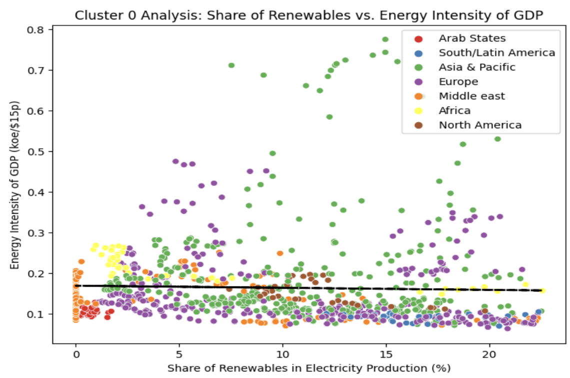

Fig. 9. Cluster 0 with $pearson\_corrleation = -0.03$

factor suggests a stronger relationship than the one in Cluster 0 Although still considered a weak negative relationship.

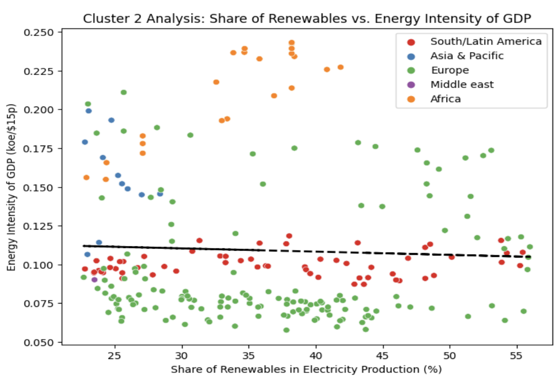

Fig. 10. Cluster 1 with $pearson\_corrleation = -0.04$

Finally in Fig. 10. We can See a moderate negative relationship between the two variables. This supports our hypothesis that countries in this cluster are increasing the Energy intensity with the increase of Renewable resources in the overall energy production in the country. This means that citizen of these countries would be able to buy more energy than before, This energy is also cleaner and more sustainable.

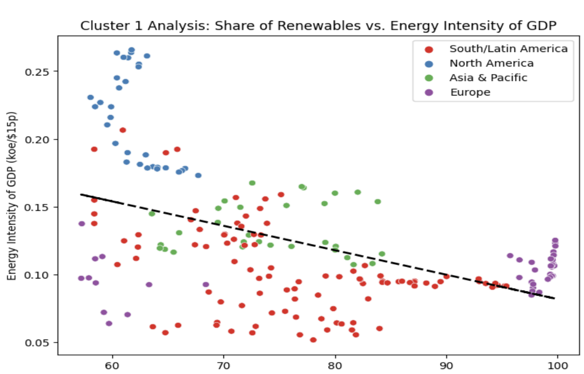

Fig. 11. Cluster 2 with $pearson\_corrleation = -0.48$

D. Third Hypothesis: Future of Renewables

The transition to renewable energy sources has become an urgent global priority in the face of increasing concerns about climate change and the need to reduce greenhouse gas emissions. However, the implementation of renewable energy resources varies across different regions, driven by unique priorities and circumstances. In this study, we investigate the relationship between the adoption of renewable energy resources and the overall CO2 emission factor in various regions over the next 10 years. We hypothesize that different regions will exhibit varying speeds of renewable energy implementation, reflecting their individual priorities and commitments towards sustainable energy solutions. By examining the interplay between regional factors and renewable energy adoption, we aim to gain insights into the potential impact on reducing CO2 emissions and fostering sustainable development strategies worldwide. Understanding these dynamics will contribute to the formulation of targeted policies and strategies to accelerate the global transition to a low-carbon future.

We will start our analysis by building exponential smoothing for our data in a time series model. Below you can see the first model that we have developed , the model aims to predict the change in Average CO2 emissions factor across each region. we have decided in our model to focus on 3 main regions that we have handled before in our first hypothesis; Europe, North America and Middle East. But rather than focusing on one country by itself we will do the study region wide. You can see in Fig. 11. The results we have obtained from each of our models.

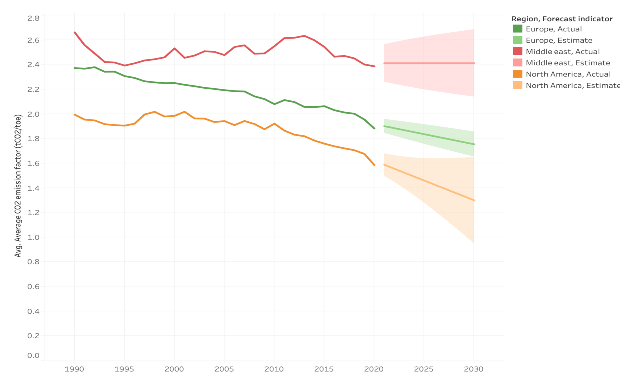

Fig. 12. Forecast of CO2 emission factor in the next 10 years with confidence bands

Avg. Average CO2 emission factor (tCO2/toe)

Color	Model			Quality Metrics					Smoothing Coefficients		
Region	Level	Trend	Season	RMSE	MAE	MASE	MAPE	AIC	Alpha	Beta	Gamma
North America	Additive	Additive	None	0.035	0.027	0.91	1.4%	-197	0.500	0.277	0.000
Middle east	Additive	None	None	0.059	0.050	1.33	2.0%	-170	0.500	0.000	0.000
Europe	Additive	Additive	None	0.022	0.017	0.85	0.8%	-227	0.500	0.000	0.000

Fig. 13. Forecast of CO2 emission factor in the next 10 years with confidence bands

You can see from our model result table that although our model has achieved impressive results regarding the closeness of prediction indicated by our low overall error metrics; it failed to discover a pattern or a seasonality in the Energy data.

That's mainly of how unpredictable and how many factors affect the behaviours and patterns of energy consumption in countries. but looking at the prediction of our model, we can that that European and North American Region are expected to bring their CO2 emissions lower in the next 10 years. While in the middle East region the CO2 emissions are expected to stay constant. This is mainly because in the Middle East Region the countries are very dependent on oil and fossil fuel production , as it is a big part of their economy.

We have implemented another model that would be able to forecast the Share of Renewable Energy in the total Energy production across the regions.

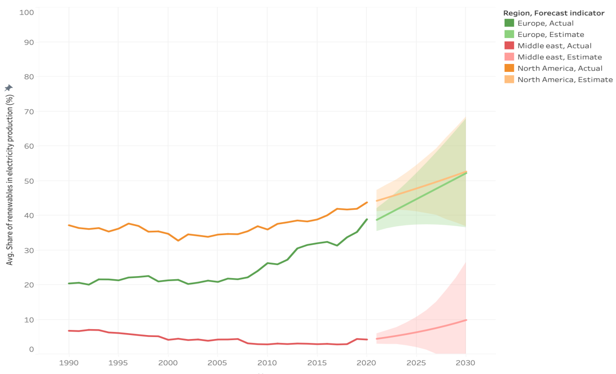

Fig. 14. Forecast of Implementation of Renewable energy resources in the next 10 years with confidence bands

We can see in the above figure that both of the European and North American regions are going to achieve part their environmental promises by having around 50% of their total electricity be generated form renewable resources. You can check the table below to see our parameters and model type. similar to the previous model it was very difficult to obtain a seasonality of our data. The error rate was quite high while predicting the European region pattern and North America, The confidence bands are also wider. but even with such error our lower bound still suggest an overall increase in implementation of cleaner energy.

Avg. Share of renewables in electricity production (%)

Color	Model			Quality Metrics					Smoothing Coefficients		
Region	Level	Trend	Season	RMSE	MAE	MASE	MAPE	AIC	Alpha	Beta	Gamma
North America	Multiplicative	Multiplicative	None	0.98	0.82	0.98	2.3%	9	0.500	0.321	0.000
Middle east	Multiplicative	Multiplicative	None	0.47	0.34	1.08	8.3%	-37	0.500	0.314	0.000
Europe	Additive	Additive	None	1.27	0.94	0.95	3.6%	25	0.500	0.389	0.000

Fig. 15. Forecast of Renewable energy share in overall production in the next 10 years with confidence bands

III. Results

A. First Hypothesis: Promises are true! or are they?

In this study, our hypothesis that countries have not been adhering to their environmental commitments was examined using data mining techniques. By analyzing the relationship between key variables - the share of renewable energy, the share of fossil fuels, the average CO2 emission factor, and the average total energy consumption - over time, we aimed to determine the extent to which countries have fulfilled their environmental goals.

The findings of our analysis provide insights into the behavior and patterns of countries with respect to their environmental promises. Among the selected countries, the United States and Russia demonstrated success in reducing their CO2 emission factors, aligning with their commitments. However, India experienced a significant increase in its CO2 emissions, suggesting a discrepancy between its goals and the actual outcomes.

By examining the relationship between the average share of renewable energy production and the average CO2 emission factor, we established a strong negative correlation. This indicates that increasing the share of renewable energy in electricity production has the potential to reduce CO2 emissions, even in the face of population growth. However, our analysis also revealed that electricity production alone is insufficient to address the overall energy consumption and emissions from other sources.

In conclusion, our study provides evidence that countries may not be fully adhering to their environmental goals. While some nations have made progress, others face challenges in achieving their desired outcomes. The findings underscore the importance of comprehensive strategies that consider various dimensions of energy consumption and population dynamics.

B. Second Hypothesis: Cheaper Energy in the Future!

The results of our analysis support the second hypothesis regarding the relationship between the share of renewables in electricity production and the energy intensity of GDP at constant purchasing power parities. Through clustering techniques applied to the dataset from 1990 to 2020, we have identified three distinct clusters of countries.

In Cluster 0, we observe a weak relationship between energy purchasing power and the share of renewables. This suggests that, despite the increase in renewable energy adoption, countries in this cluster have not significantly impacted energy prices. In Cluster 1, there is a slightly stronger negative relationship between the variables, indicating that countries have been able to strike a balance between renewable energy expansion and maintaining affordable energy prices.

The most notable finding comes from Cluster 2, where we identify a moderate negative relationship. This indicates that countries in this cluster have successfully increased their share of renewables in energy production while also experiencing an increase in energy intensity. In other words, as these countries adopt more renewable energy sources, their citizens are able to access greater amounts of energy. This outcome aligns with our hypothesis, demonstrating that these countries have effectively enhanced energy affordability and sustainability.

Overall, the results of our analysis provide valuable insights into the relationship between renewable energy adoption and economic factors. The clustering of countries based on their share of renewables and energy intensity highlights the diverse approaches taken by different nations. These findings contribute to the understanding of how countries navigate the challenges of energy transition while striving for economic growth and sustainability.

C. Third Hypothesis: Cleaner energy for everyone!

In line with our third hypothesis, we explore the relationship between the adoption of renewable energy resources and the overall CO2 emission factor in different regions. Our analysis focuses on Europe, North America, and the Middle East, aiming to understand the varying speeds of renewable energy implementation and their impact on reducing CO2 emissions.

Using exponential smoothing in a time series model, we forecasted the change in the average CO2 emission factor for each region over the next 10 years. The results from our models, as shown in Fig. 11, indicate that European and North American regions are expected to lower their CO2 emissions, while the Middle East region is projected to maintain relatively constant emissions. This discrepancy can be attributed to the Middle East's heavy reliance on oil and fossil fuel production, which plays a significant role in their economy.

Furthermore, we developed another model to forecast the share of renewable energy in total energy production across the regions. Fig. 12 illustrates that both Europe and North America are predicted to achieve a significant portion of their environmental commitments, with approximately 50

The unpredictability and numerous factors influencing energy consumption patterns contribute to the difficulty in identifying clear patterns or seasonality in the data. Nonetheless, our model predictions provide valuable insights into the future trajectory of renewable energy adoption and highlight the progress made by European and North American regions in their sustainable energy goals.

IV. Conclusion

In conclusion, this academic paper investigates three hypotheses related to energy systems, sustainability, and the transition to renewable energy sources. The findings shed light on important aspects of renewable energy adoption, environmental commitments, and the impact on CO2 emissions at both regional and global levels.

First, the study examines the adherence of countries to their environmental goals. The analysis reveals that while some countries have made significant progress in reducing their CO2 emissions, others struggle to meet their commitments. Factors such as population growth and the pace of renewable energy implementation contribute to the observed discrepancies. Overall, the hypothesis that countries are not uniformly adhering to their environmental goals is supported.

Secondly, the research explores the relationship between the share of renewables in electricity production and the energy intensity of GDP. The findings demonstrate that different clusters of countries exhibit distinct patterns in terms of the economic impact of renewable energy adoption. While some countries experience positive effects on purchasing power, others show no significant change or even negative effects. These results emphasize the complexity of the interplay between renewable energy and economic factors.

Lastly, the study examines the regional variations in renewable energy implementation and its impact on CO2 emissions. The forecasts suggest that European and North American regions are expected to make substantial progress in reducing CO2 emissions over the next decade. In contrast, the Middle East region, heavily reliant on fossil fuel production, is projected to maintain relatively constant emissions. The models also indicate an overall increase in the share of renewable energy, although challenges in capturing seasonality and higher error rates are observed.

Overall, this research contributes to the understanding of the complexities and dynamics involved in the transition to renewable energy sources. The findings underscore the need for targeted policies and strategies to accelerate the global adoption of cleaner energy and achieve sustainable development goals. By exploring regional variations, policy-makers and stakeholders can gain insights into the factors influencing renewable energy adoption and design effective measures to mitigate climate change and reduce greenhouse gas emissions on a global scale.

References

[1] Santos, G. (2017). Road transport and CO2 emissions: What are the challenges? Transport Policy, 59, 71-74. https://doi.org/10.1016/j.tranpol.2017.06.007

[2] https://www.kaggle.com/datasets/shub218/energy-data-1990-2020

[3] Raupach, M. R., Marland, G., Ciais, P., Le Quéré, C., Canadell, J. G., Klepper, G., & Field, C. B. (2007). Global and regional drivers of accelerating CO2 emissions. Proceedings of the National Academy of Sciences, 104(24), 10288-10293. https://doi.org/10.1073/pnas.0700609104

[4] Davis, S. J., & Caldeira, K. (2010). Consumption-based accounting of CO2 emissions. Proceedings of the National Academy of Sciences, 107(12), 5687-5692. https://doi.org/10.1073/pnas.0906974107

[5] Dabo Guan, Klaus Hubacek, Christopher L. Weber, Glen P. Peters, David M. Reiner, The drivers of Chinese CO2 emissions from 1980 to 2030, Global Environmental Change, Volume 18, Issue 4, 2008, Pages 626-634, ISSN 0959-3780, https://doi.org/10.1016/j.gloenvcha.2008.08.001. (https://www.sciencedirect.com/science/article/pii/S095937800800068X)

[6] Shahateet, M. (2014). Modeling Economic Growth and Energy Consumption in Arab Countries: Cointegration and Causality Analysis. International Journal of Energy Economics and Policy, 4, 349-359.

[7] Ayoub M. Kazim, Assessments of primary energy consumption and its environmental consequences in the United Arab Emirates, Renewable and Sustainable Energy Reviews, Volume 11, Issue 3, 2007, Pages 426-446, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2005.01.008. (https://www.sciencedirect.com/science/article/pii/S1364032105000390)

[8] Mita Bhattacharya, Sudharshan Reddy Paramati, Ilhan Ozturk, Sankar Bhattacharya, The effect of renewable energy consumption on economic growth: Evidence from top 38 countries, Applied Energy, Volume 162, 2016, Pages 733-741, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2015.10.104. (https://www.sciencedirect.com/science/article/pii/S0306261915013318)

[9] ClimateTrade. *Which Countries Are the World's Biggest Carbon Polluters?* https://climatetrade.com/which-countries-are-the-worlds-biggest-carbon-polluters/ [Accessed: 2023-06-09].

[10] World-Info. *Population growth in India* https://www.worlddata.info/asia/india/populationgrowth.php [Accessed: 2023-06-09].

V. Appendix: I

TABLE I
Countries in Clusters

Cluster 0	**Cluster 1**	**Cluster 2**
Algeria	Brazil	Argentina
Argentina	Canada	Australia
Australia	Chile	Belgium
Belgium	Colombia	Chile
China	New Zealand	China
Czechia	Norway	Egypt
Egypt	Portugal	France
France	Sweden	Germany
Germany	Venezuela	India
India		Indonesia
Indonesia		Italy
Iran		Mexico
Italy		Netherlands
Japan		Nigeria
Kazakhstan		Portugal
Kuwait		Romania
Malaysia		Spain
Mexico		Sweden
Netherlands		Turkey
Nigeria		United Kingdom
Poland		
Portugal		
Romania		
Russia		
Saudi Arabia		
South Africa		
South Korea		
Spain		
Taiwan		
Thailand		
Turkey		
Ukraine		
United Arab Emirates		
United Kingdom		
United States		
Uzbekistan		