

# Research of varying patterns of CO<sub>2</sub> emissions in 182 countries based on K-means method

Wenchang Zhou<sup>1,4</sup>, Jiangmeng Zhou<sup>2</sup>, Gaoyuan Hu<sup>3</sup>

<sup>1</sup>Tandon School of Engineering, New York University, Brooklyn, New York, 11201, the United States

<sup>2</sup>School of Physics and Electronics, Central South University, Hunan, Changsha, 410006, China

<sup>3</sup>RDFZ King's college School, Zhejiang, Hangzhou, 310000, China

<sup>4</sup>wz1455@nyu.edu

**Abstract.** Global climate change caused by greenhouse gases is becoming more and more serious, and many countries have proposed their own carbon neutrality goals in response to this problem. Although there are discrepancies in carbon emissions in different countries due to factors such as population, economy, technology, etc., from the perspective of the status quo and development trend, there should be certain similarities in carbon emissions between different countries. Therefore, the primary purpose of the analysis is to explore similar patterns from aggregate and per-capita CO<sub>2</sub> emissions from oil production and other field data in 182 countries using the K-means cluster analysis technique. The results revealed the existence of similar CO<sub>2</sub> emission patterns and divided all countries into 4 clusters. Furthermore, we analyzed CO<sub>2</sub> emission patterns for each cluster separately and discussed the reasons and potential relationships behind each emission pattern based on the results. The research provides a novel way to analyze global CO<sub>2</sub> emissions by applying taxonomic methods and provides insightful information to the international community for future low-carbon development.

**Keywords:** CO<sub>2</sub> emissions, K-means, clustering analysis.

## 1. Introduction

Global warming is one of the most critical issues of the international community and is a significant threat to the global environment. Excessive Greenhouse Gas (GHG) emissions are the most critical reason causing global warming, and the most dominant element of GHG is carbon dioxide (CO<sub>2</sub>). Fossil fuels' production and consumption are major sources of carbon dioxide emissions. As one of the most used fossil fuels, oil accounted for 35% of the total global energy demand in 2004 [1]. Although, to a certain extent, the oil demand has declined due to the epidemic, this decline will not last too long, as economic activity across countries is gradually returning to a pre-pandemic era (Ram et al., 2021).

Many studies explore the association between CO<sub>2</sub> emissions and elements, such as real GDP, oil production, and economic growth [2], and some studies focus on countries or regions that contribute a significant amount of CO<sub>2</sub> emissions. However, different countries might have some discrepancies in CO<sub>2</sub> emissions, but some patterns should exist, reflecting the similar progress and trend toward the carbon neutrality target, and few studies explore similar CO<sub>2</sub> emissions patterns on a global scale using

cluster analysis. Therefore, this study leverages K-means cluster analysis techniques to investigate similar patterns of global carbon emissions by analyzing CO<sub>2</sub> emissions from oil production data from 182 countries and regions.

This study innovatively applies K-means algorithms to explore global GHG emissions data, which provide a novel method to comprehensively understand the status of global CO<sub>2</sub> emissions and a practical implication for carbon reduction strategies or policies for different countries.

The following sections are as follows: Section 2 will include literature reviews of previous studies. Section 3 will introduce data and detailed methods of the research. Then, section 4 will present the results and analysis. Eventually, the study's results and discussion will be presented in Section 5.

## 2. Literature reviews

Looking at the related literature provides essential information. GHG emissions could be attributed to five factors: energy, industry, buildings, transport, and AFOLU (agriculture, forestry, and other land use). With the human world's development, GHG emissions have risen these years. Lamb [3] estimated the GHG emissions trends by the factor from 1990 to 2018, describing the major sources of emissions growth, stability, and decline across ten global regions.

To come up with valuable suggestions about reducing GHG emissions, researching developed countries like European countries is a good idea since those countries have made great efforts in this aspect. For example, Kijewska & Bluszczyk [1] researched the GHG emissions in European Union member states with an agglomeration algorithm, finding homogeneous countries with similar performance in GHG emissions. In addition, Alicja et al. [4] used cluster analysis, k-means, to group members of the Organization for Economic Cooperation and Development (OECD) into homogeneous subsets for similarities of agricultural variables affecting GHG emissions.

More specifically, Magazzino et al. [5] used a machine learning approach to explore the relationship among energy production, coal consumption, GDP, and CO<sub>2</sub> emissions in China, the USA, and India, which are the world's biggest energy consumers and CO<sub>2</sub> emitters. As for China, Li et al. [6] discussed the driving factors of carbon dioxide emissions based on machine learning methods, revealing that economic growth measured by GDP contributes to higher CO<sub>2</sub> emissions.

In recent years, the K-means algorithm has been applied to research in various fields. Tan et al. [7] used K-means clustering to find transaction abnormalities based on e-commerce data. Ahuja et al. [8] combined the K-means clustering algorithm with the K-Nearest Neighbor algorithm to build a movie recommendation system. Kijewska & Bluszczyk [9] used the K-means algorithm to study the distribution of different GHG emissions in Europe. Cui et al. [10] use K-means to investigate common development patterns of cities based on the CO<sub>2</sub> emissions data of different provinces in China.

## 3. Method

### 3.1. K-means clustering model

The K-means clustering algorithm is widely regarded as one of the most powerful unsupervised learning techniques when exploring the underlying data distribution patterns [11]. The rationale of this algorithm is to partition a data set  $D = \{x_1, \dots, x_N\}$  into  $k$  disjoint clusters  $C_1, C_2, \dots, C_K$  by minimizing the sum of squared Euclidean distance between data points and cluster centroids  $\mu_1, \mu_2, \dots, \mu_K$  [12]. The objective of the algorithm is to minimize the following function:

$$E(\mu_1, \dots, \mu_k) = \sum_{i=1}^N \sum_{k=1}^K I(x_i \in C_k) \|x_i - \mu_k\|^2 \quad (1)$$

Where  $I(x_i \in C_k) = 1$ , unless it will change to 0 if  $x_i$  not in the cluster  $C_k$ .

### 3.2. Data collection

The data is coming from the website of Our World in Data. This research is based on the CO<sub>2</sub> emission data from Hannah et al. (2020), which contains 60 columns and 29,589 rows about the CO<sub>2</sub> emission in

the different countries and continental plates worldwide in different years and fields, such as different sources (oil, coal, flaring, cement) of carbon emissions and followed by corresponding data (aggregate CO<sub>2</sub> emissions, cumulative CO<sub>2</sub>, share global CO<sub>2</sub>).

### 3.3. Data pre-processing

**3.3.1. Data cleansing.** The data contains many missing values, which might cause many problems in further analysis. Therefore, we select the time scale from 2001 to 2020 as the target years to achieve two goals. One is that most countries have valid data between 2001 to 2020. On the other hand, countries' policy decisions are highly dependent on recent years. Since we only focus on individual countries when applying the K-means model, we removed all continental plate and island data. Moreover, we select 9 variables from 60 columns to narrow down the scale of our research. Table 1. shows the columns we finally selected.

**Table 1.** Selected columns in data set.

Column Names	Description
year	Target time scale, 2001 – 2020
iso_code	Three-digit country code, defined by ISO 3166-1
co2	Annual production-based CO <sub>2</sub> emissions (million tonnes).
population	Population of each country
gdp	GDP measured in international dollar.
co2_per_capita	Annual production-based CO <sub>2</sub> emissions (tonnes per perso.
co2_per_unit_energy	Annual production-based CO <sub>2</sub> emissions (kilograms per kilowatt-hour)
oil_co2	Annual production-based CO <sub>2</sub> emissions from oil (million tonnes).
oil_co2_per_capita	Annual production-based CO <sub>2</sub> emissions from oil (tonnes per person).

However, a few null values are in the Gross Domestic Products (GDP) column. We filled some of them manually by finding data from World Bank. However, some small countries were still missing the data in 2020. Therefore, we deployed the zoo library in R language to fill them with the GDP in 2019. This step might import some bias, but since GDP is an indicator representing a country's economy and cannot have a huge discrepancy in one year, we use this technique to complete our data. Finally, we got 3640 valid data, including 182 individual countries with 9 features without any missing value.

**3.3.2. Features selection for K-means.** After cleaning the data, we select and construct some features that use to partition countries. The data set contains CO<sub>2</sub> emissions from oil and total CO<sub>2</sub> emissions from the whole country. Since the total emissions data is affected by the factor that reflects the country's volume, such as population, we only use per capita data as part of the features to pass into the K-means algorithm. In addition, since we also want to take CO<sub>2</sub> emissions from sectors other than oil into consideration, so we calculate the per capita CO<sub>2</sub> emissions by the following formula:

$$Q_{ot.p} = \frac{Q_{co_2} - Q_{oil}}{P} \quad (2)$$

Where  $Q_{ot.p}$  represents the other CO<sub>2</sub> emissions per capita  $Q_{co_2}$  represents the amount of CO<sub>2</sub> emissions from the whole country;  $Q_{oil}$  represents each country's oil CO<sub>2</sub> emissions; and P represents the population of each country. Moreover, to reflect the usage efficiency of the energy, we select CO<sub>2</sub> emissions per unit of energy to achieve this goal. We also selected the GDP per capita data in 2020 to represent each country's economic development degree, but this variable will not pass into the K-means algorithm because it is highly correlated with the oil CO<sub>2</sub> emissions per capita, which is 0.8.

Furthermore, the linear regression model is applied to acquire the trends of CO<sub>2</sub> emissions as clustering features in the K-means model. We deploy this technique on the data of each country's oil CO<sub>2</sub> emissions per capita, and other CO<sub>2</sub> emissions per capita from 2001 to 2020, and the coefficient

parameter of the fitting result can be used to represent the trend of these parameters in the past ten years. Table 2 shows the variables we select to pass into the K-means algorithm.

**Table 2.** Selected features for K-means.

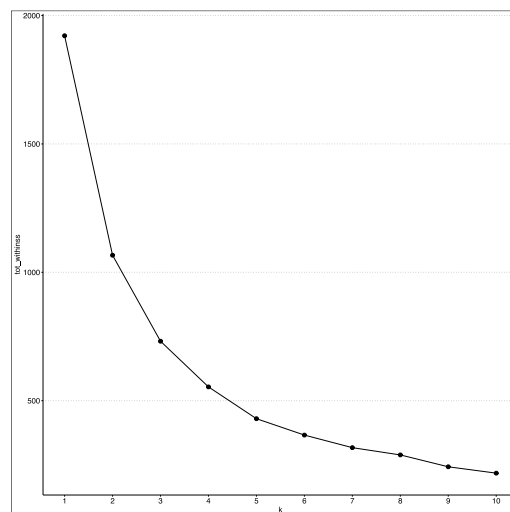
Variable	Description
oil.pc	CO <sub>2</sub> emissions per capita from oil production in 2020, normalized by logarithm.
other.pc	Other production-based CO <sub>2</sub> emissions per capita in 2020, normalized by logarithm.
co2.energy	CO <sub>2</sub> emissions per unit of energy, which represents the energy usage efficiency. Normalized by logarithm
oil.trend	Trend of total oil CO <sub>2</sub> emissions from 2001 to 2020 of each country.
oil.pc.trend	Trend of oil CO <sub>2</sub> emissions per capita from 2001 to 2020 of each country.
other.trend	Trend of CO <sub>2</sub> emissions from field other than oil of each country.
other.pc.trend	Trend of CO <sub>2</sub> emissions per capita from field other than oil of each country.
GDP.pc	GDP per capita of each country.

However, the results from the K-means algorithm are sensitive to outliers in the data since the outliers can increase the squared error drastically, which leads to the outliers tending to form a few small clusters, even if sometimes a cluster only contains a single value. The rest of the values will be classified as having a few distinct characteristics. Therefore, solving the outliers' problem is necessary before we pass the data into the algorithm. The outliers in our dataset cannot be removed since they are actual values rather than errors, and these outliers contain essential information about the discrepancies in the CO<sub>2</sub> emissions of different countries. To maintain the integrity of the information, we define threshold values for each parameter after exploring each parameter's distribution. Then, the outliers will be set equal to these threshold values.

## 4. Results and discussion

### 4.1. Selection of number of clusters

The K-means algorithm needs to initiate the k value to define the number of clusters. This study uses the elbow method to determine the optimal number of clusters for classification. Figure 1 shows the graph after using the elbow method on our data. According to the elbow method results [13], optimal number of clusters should be 4.



**Figure 1.** Elbow method results.

#### 4.2. Results

This study uses the built-in k-means function in R language to apply the K-means algorithm to the data. Table 3 shows the summary of the results from the algorithm. and the Figure 2 depict the distribution of GDP per capita in each cluster.

Table 3. Cluster Result Summary

Cluster	Amount	Mean (gdp.pc)	SD (gdp.pc)	Max (gdp.pc)	Min (gdp.pc)
1	38	10.48192	0.6890102	11.67148	8.447764
2	121	8.049994	1.110775	11.61284	5.455054
3	10	8.589461	0.8226463	9.913624	7.214956
4	13	9.270659	0.9102418	10.68602	7.866754

Note. gdp.pc stands for GDP per capita. SD stands for standard deviation. Table A1 in appendix shows the whole country list of each cluster.

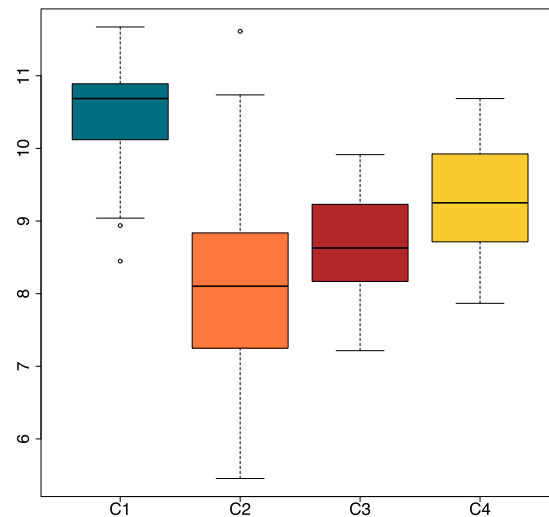
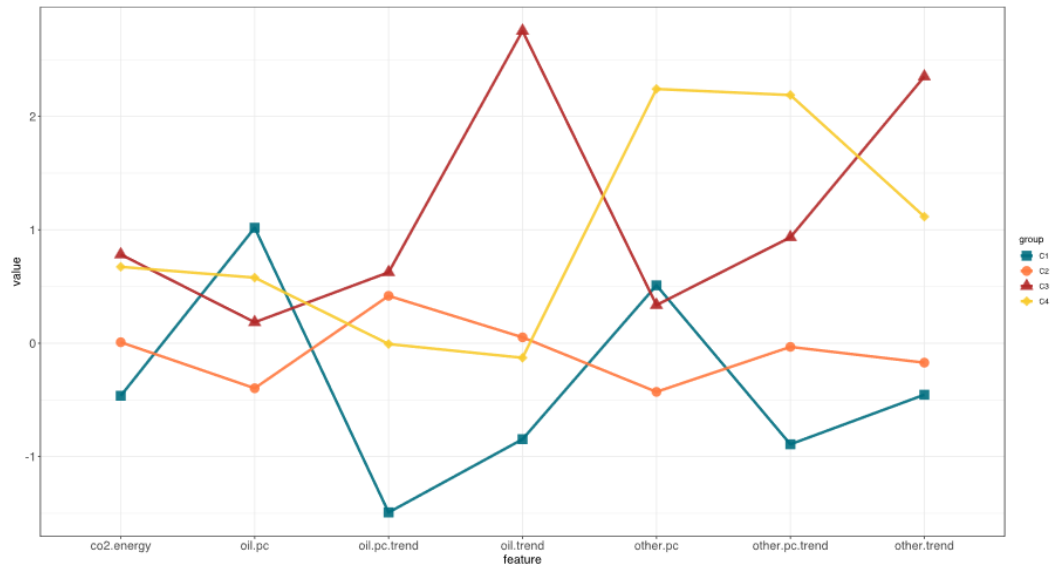


Figure 2. Distribution of GDP per capita in different clusters.

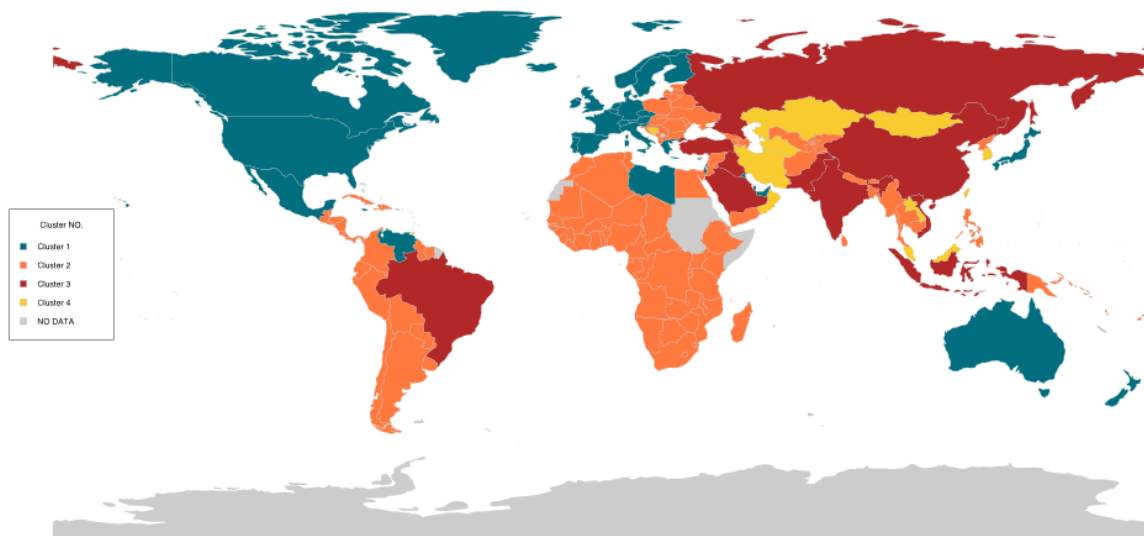
Table 4 and Figure 3 show the centers of each cluster, and Figure 4 shows a world map showing countries distributed in different clusters. These results depict the varying CO<sub>2</sub> emissions patterns of countries around the world.

Table 4. Centers of each cluster.

Feature	C1	C2	C3	C4
oil.pc	1.01817761	-0.3971615	0.18466744	0.57839351
other.pc	0.51053766	-0.4290412	0.33636033	2.24230407
co2.energy	-0.4635092	0.00851708	0.78249446	0.67368002
oil.trend	-0.846828	0.05208342	2.75490532	-0.128591
oil.pc.trend	-1.4924407	0.4177962	0.62530937	-0.0072067
other.trend	-0.4545577	-0.1715517	2.35179646	1.11638371
other.pc.trend	-0.8915221	-0.0323868	0.93434571	2.18870661



**Figure 3.** Centers of each cluster.



**Figure 4.** Distribution of countries in each cluster.

#### 4.3. Results analysis and discussion

The first cluster has the highest oil CO<sub>2</sub> emissions per capita among the four clusters. Based on the Table A1 and Figure 4, this cluster comprises oil-rich countries and many advanced economies with the world's highest CO<sub>2</sub> emissions per capita. However, these countries have negative value in oil and other CO<sub>2</sub> emissions per capita trends, indicating diminishing trends in recent years. In addition, CO<sub>2</sub> emissions per unit of energy in these countries are the lowest among all clusters, which means these countries have the highest energy efficiency.

The reason why these countries have such CO<sub>2</sub> emissions trends can be concluded into the following aspects:

Firstly, government policies have been essential in driving the carbon reduction process. For example, many EU countries have recently aimed to achieve sustainable development goals by building green economies. The green economy is a dynamic economic transformation process that aims to create new

occupations by adopting new technologies and innovations while increasing the well-being of people's livelihoods, changing the economy towards low-carbon development, and improving resource efficiency [14-15]. The adoption of such a policy also explains Europe's overall decline in carbon emissions in recent years.

Secondly, many countries invest significant amounts of money in pushing the adoption of renewable energy (i.e., wind, solar, nuclear) to modify the energy mix and reduce the dependencies on fossil fuels. For example, although many Gulf Cooperation Council (GCC) states such as Qatar, United Arab Emirates (UAE), and Kuwait have been led the CO<sub>2</sub> emissions in recent years, these countries have begun to diversify their energy structure and invest in renewable energy technologies, which contribute to the decreasing trend of CO<sub>2</sub> emissions in these countries. Qatar and UAE are two leading countries in GCC, which provide an exquisite environment for research on renewable technology development, such as Masdar Institute in UAE and Qatar Foundation's Education City. Furthermore, these countries also invest a vast amount of money in developing solar technologies for energy generation to reduce fossil fuel usage domestically [16].

Cluster 2 contains the most countries. However, this cluster has the lowest GDP per capita, based on Figure 3. For most countries in this cluster, the oil CO<sub>2</sub> emissions and CO<sub>2</sub> emissions from other fields are lower than the other clusters. Plus, the CO<sub>2</sub> emissions efficiency is about zero. Nevertheless, the oil CO<sub>2</sub> emissions per capita trends in these countries are greater than 0, which means these countries are developing, even though their level of economic development is still at the lowest level. The world map confirms the speculation from above. Most of the countries in cluster 2 are in Africa, and South America, which has many under-developed countries such as Afghanistan, Chile, and Zambia. Those countries should enhance their CO<sub>2</sub> emissions efficiency while developing economy. Specifically, they could promote green technological innovation and improve energy and carbon emission efficiency.

However, this cluster has some exceptions, such as Estonia, Latvia, and Poland, which are regarded as high-income economies. Most of these countries are from Europe, and the primary industries in these countries are concentrated in industries with low-carbon economies, such as services and tourism [17], compared to heavy industry.

Cluster 3 has the highest value in the oil CO<sub>2</sub> emissions and oil CO<sub>2</sub> emissions per capita trends. However, the massive discrepancy between these two parameters reveals that these countries have a vast population base. Furthermore, considering these countries are all giant developing economies, such as China, India, and Russia, the conspicuous increasing trends in CO<sub>2</sub> emissions indicate that slowing the growth rate of carbon dioxide emissions in these countries is one of the top priorities for slowing global warming. In addition, we notice that the CO<sub>2</sub> emissions efficiency in these countries is the worst. Compared to countries in cluster 1, they have the highest CO<sub>2</sub> emissions efficiency, which means improving energy consumption efficiency can be one of the most effective ways to reduce CO<sub>2</sub> emissions. Upgrading industry chain and exploring new energy sources to reduce the dependence to fossil fuel are good measures for these countries. For instance, they could promote electric vehicles to reduce carbon emissions [18].

The fourth cluster comprises countries with relatively high oil CO<sub>2</sub> emissions per capita but far less than non-oil CO<sub>2</sub> emissions per capita. Furthermore, it has the highest non-oil CO<sub>2</sub> emissions per capita trend among the four clusters, indicating that CO<sub>2</sub> emissions from non-oil sectors are still increasing dramatically in recent ten years. This cluster's primary sources of CO<sub>2</sub> emissions are from other fossil fuels such as coal or gas. For example, Kazakhstan, South Korea, and Mongolia heavily depend on coal as their primary energy source. On the other hand, Iran, Bahrain, and Brunei take gas as their primary energy source. In general, these countries have relatively high levels of CO<sub>2</sub> emissions in the world, and they face severe problems with the third cluster of countries in terms of progress in reducing carbon emissions, despite the growth in their total CO<sub>2</sub> emissions. The trend is no slower than that of the third group of countries, but after excluding population factors, the per capita emissions of such countries are higher than the per capita carbon emissions of the third group of countries in any sector. In addition, the growth rate of non-oil carbon dioxide emissions in these countries is much higher than that of the other three groups, which means that this trend may not be reversed in the next few years.

## 5. Conclusions

While there are different status quo of CO<sub>2</sub> emissions and different trends of CO<sub>2</sub> emissions in recent years in individual countries, this study proposes an innovative method to apply the K-means algorithm to explore similar CO<sub>2</sub> emissions patterns among countries and regions. We manage to separate all countries into four clusters, and countries in each cluster have some similar characteristics in their CO<sub>2</sub> emissions pattern. Moreover, the results reveal the problems in different regions and provide a novel path for future studies to analyze GHG emissions problems in homogeneous groups of countries in different regions.

However, there are still many deficiencies in our research. Environmental issues have always been very complex issues. Different observational scales and evaluation angles may lead to entirely different conclusions when judging a country's performance on environmental issues. This study only selects several elements that contribute to CO<sub>2</sub> emissions for analysis, which is highly one-sided. Therefore, future research will add more variables to the model, such as other components of GHG, or consider other industries, to build a more comprehensive clustering model. At the same time, we will analyze more fine-grained influencing factors in the future, such as household-level carbon emissions, to help us understand the carbon emission patterns of specific industries more accurately and propose more targeted emission reduction suggestions.

In addition, during our research, we found that during 2019-2020, global CO<sub>2</sub> emissions were lower than predicted by linear regression using data from 2001-2018, with more than 50% of countries experiencing a decline in CO<sub>2</sub> emissions, although this seems to be a good sign, another question is that many countries have also experienced a dramatic decline in GDP. Will this decline lead to a significant rebound in carbon emissions in the post-pandemic era to compensate for the loss of GDP? How we can avoid this backlash are questions we need to address in future research.

## Appendix

A1. Country List in Each Cluster

Cluster 1	Cluster 2			Cluster 3	Cluster 4
	Section 1	Section 2	Section 3		
Aruba	Afghanistan	Ethiopia	Nigeria	Brazil	Bahrain
Australia	Albania	Fiji	North Korea	China	Bosnia and Herzegovina
Austria	Algeria	Gabon	Palestine	India	Brunei
Belgium	Angola	Gambia	Panama	Indonesia	Iran
Canada	Antigua and Barbuda	Georgia	Papua New Guinea	Iraq	Kazakhstan
Cyprus	Argentina	Ghana	Paraguay	Pakistan	Laos
Czech Republic	Armenia	Grenada	Peru	Russia	Malaysia
Denmark	Azerbaijan	Guatemala	Philippines	Saudi Arabia	Mongolia
Finland	Bangladesh	Guinea	Poland	Turkey	Oman
France	Barbados	Guinea-Bissau	Republic of the Congo	Vietnam	South Korea
Germany	Belarus	Guyana	Romania		Taiwan
Greece	Belize	Haiti	Rwanda		Trinidad and Tobago
Greenland	Benin	Honduras	Saint Lucia		Turkmenistan
Iceland	Bermuda	Hong Kong	Samoa		
Ireland	Bhutan	Hungary	Sao Tome and Principe		
Israel	Bolivia	Ivory Coast	Senegal		
Italy	Botswana	Jordan	Serbia		



Jamaica	Bulgaria	Kenya	Seychelles
Japan	Burkina Faso	Kiribati	Sierra Leone
Kuwait	Burundi	Kyrgyzstan	Slovakia
Libya	Cambodia	Latvia	Solomon Islands
Luxembourg	Cameroon	Lebanon	South Africa
Malta	Cape Verde	Lesotho	Sri Lanka
Mexico	Central African Republic	Liberia	Suriname
Netherlands	Chad	Lithuania	Swaziland
New Zealand	Chile	Macedonia	Syria
Norway	Colombia	Madagascar	Tajikistan
Portugal	Comoros	Malawi	Tanzania
Qatar	Costa Rica	Maldives	Thailand
Singapore	Croatia	Mali	Togo
Slovenia	Cuba	Mauritania	Tonga
Spain	Democratic Republic of the Congo	Mauritius	Tunisia
Sweden	Djibouti	Moldova	Uganda
Switzerland	Dominica	Montenegro	Ukraine
United Arab Emirates	Dominican Republic	Morocco	Uruguay
United Kingdom	Ecuador	Mozambique	Uzbekistan
United States	Egypt	Myanmar	Vanuatu
Venezuela	El Salvador	Namibia	Yemen
	Equatorial Guinea	Nepal	Zambia
	Estonia	Nicaragua	Zimbabwe
	Ethiopia	Niger	

## References

- [1] A. Kijewska and A. Bluszcz, "Research of varying levels of greenhouse gas emissions in European countries using the k-means method," *Atmos Pollut Res*, vol. 7, no. 5, pp. 935–944, 2016.
- [2] K.-M. Wang, "Modelling the nonlinear relationship between CO2 emissions from oil and economic growth," *Econ Model*, vol. 29, no. 5, pp. 1537–1547, 2012.
- [3] W. F. Lamb *et al.*, "A review of trends and drivers of greenhouse gas emissions by sector from 1990 to 2018," *Environmental Research Letters*, vol. 16, no. 7. IOP Publishing Ltd, Jul. 01, 2021.
- [4] A. Kolasa-Więcek, "The use of cluster analysis in the classification of similarities in variables associated w/agricultural greenhouse gases emissions in OECD countries," *Village and Agriculture (Wieś i Rolnictwo)*, vol. 1, no. 1229-2016–98791, 2013.
- [5] C. Magazzino, M. Mele, and N. Schneider, "A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions," *Renew Energy*, vol. 167, pp. 99–115, Apr. 2021.
- [6] Y. Li, X. Yang, Q. Ran, H. Wu, M. Irfan, and M. Ahmad, "Energy structure, digital economy, and carbon emissions: evidence from China," *Environmental Science and Pollution Research*, vol. 28, no. 45, pp. 64606–64629, 2021.
- [7] X. S. Tan, Z. Yang, Y. Benlimane, and E. Liu, "Using Classification with K-means Clustering to Investigate Transaction Anomaly," in *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2020, pp. 171–174.

- [8] R. Ahuja, A. Solanki, and A. Nayyar, "Movie recommender system using K-Means clustering and K-Nearest Neighbor," in *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2019, pp. 263–268.
- [9] A. Kijewska and A. Bluszcz, "Analysis of greenhouse gas emissions in the European Union member states with the use of an agglomeration algorithm," *Journal of Sustainable Mining*, vol. 15, no. 4, pp. 133–142, Jan. 2016.
- [10] C. Cui, Z. Wang, B. Cai, S. Peng, Y. Wang, and C. Xu, "Evolution-based CO2 emission baseline scenarios of Chinese cities in 2025," *Appl Energy*, vol. 281, p. 116116, 2021.
- [11] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics (Basel)*, vol. 9, no. 8, p. 1295, 2020.
- [12] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognit*, vol. 36, no. 2, pp. 451–461, 2003.
- [13] C. Yuan and H. Yang, "Research on K-value selection method of K-means clustering algorithm," *J (Basel)*, vol. 2, no. 2, pp. 226–235, 2019.
- [14] I. Mikhno, V. Koval, G. Shvets, O. Garmatiuk, and R. Tamošiūnienė, "Green economy in sustainable development and improvement of resource efficiency," *Central European Business Review (CEBR)*, vol. 10, no. 1, pp. 99–113, 2021.
- [15] L. Dogaru, "Green economy and green growth—Opportunities for sustainable development," in *Proceedings*, 2021, vol. 63, no. 1, p. 70.
- [16] Y. Atalay, F. Biermann, and A. Kalfagianni, "Adoption of renewable energy technologies in oil-rich countries: Explaining policy variation in the Gulf Cooperation Council states," *Renew Energy*, vol. 85, pp. 206–214, 2016.
- [17] A. Panasiuk and E. Wszendybył-Skulska, "Social aspects of tourism policy in the European Union. The example of Poland and Slovakia," *Economies*, vol. 9, no. 1, p. 16, 2021.
- [18] J. Li and B. Yang, "Analysis of greenhouse gas emissions from electric vehicle considering electric energy structure, climate and power economy of ev: A China case," *Atmos Pollut Res*, vol. 11, no. 6, pp. 1–11, Jun. 2020.