Statistical analysis of data –
CO2 emissions in the world.

Tomasz Wira

Wrocław 2023

Introduction

Nowadays, the topic of global warming is well known to everyone. Greenhouse gas emissions have recently appeared quite often in public debate. This problem concerns the whole world because the processes taking place in nature and in individual economies are interdependent. In order to understand how strongly the economy affects the natural environment, this work will analyze greenhouse gas emissions in relation to individual economies.

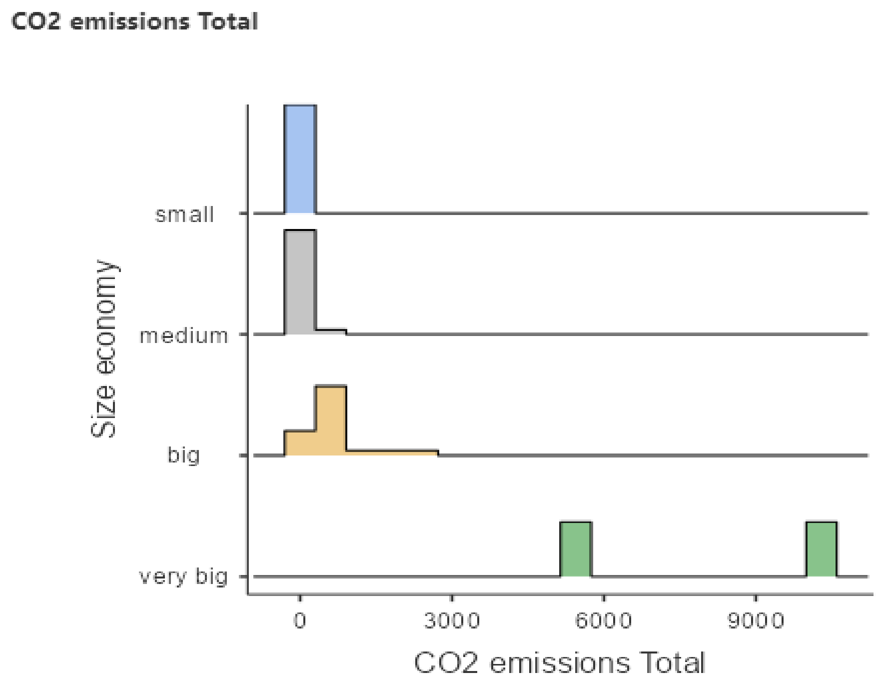
The following work presents an analysis of greenhouse gas emissions in 2018 and has been divided into three parts, in the first part statistical data will be presented using descriptive and statistical methods. In the second part of the work, a detailed analysis of statistical data will be carried out, which were obtained using the CORGIS Project website. The collected data will be analyzed using methods such as dependency test, T-tests, variance analysis, multiple regression analysis and logistic regression analysis. The third part of this work will summarize the results of the conducted research and measurements. The last part will also draw conclusions based on the results of the conducted statistical analysis study.

In this work, research methods will be used, i.e. statistical method, descriptive method, statistical analysis method and heuristic method. In order to best illustrate the results of the conducted research, drawings presenting the results of the analyses will be added. Microsoft Excel will be used for the study for statistical purposes. In order to conduct an in-depth statistical analysis, typical programs for this purpose will be used, such as Jamovi and GRETL. The theoretical background will be provided by information acquired during the course and the Electronic Textbook StatSoft[1].

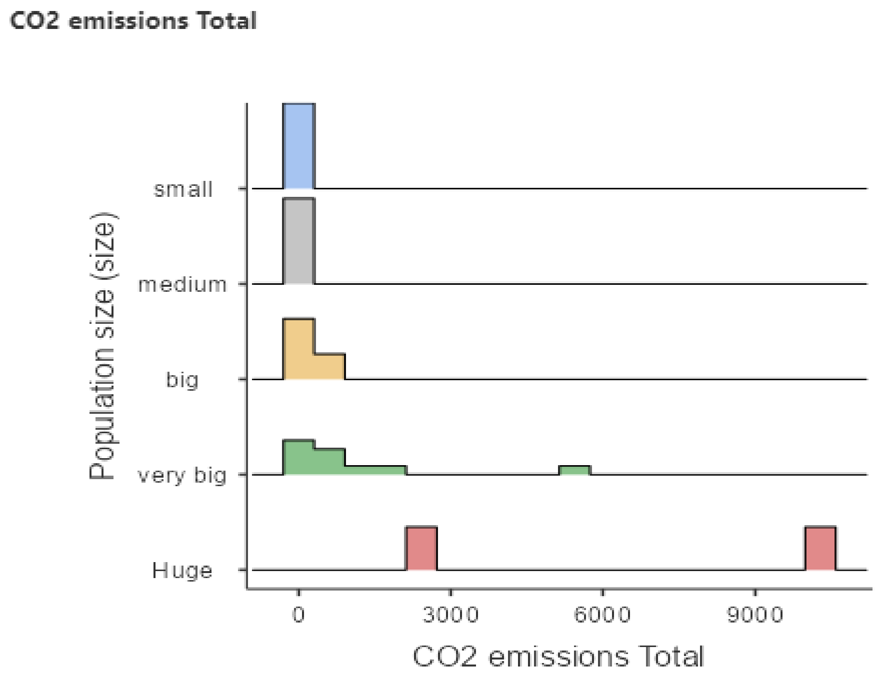
So the considerations in the following work should allow us to obtain an answer to the question whether the size of the population and the size of the economy have an impact on the size of greenhouse gas production? The data collected and statistically analyzed should clearly answer whether such relationships exist and how strong they are..

[1] StatSoft, https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fadvans4.html (dostęp.25.01.2023r)

Data description



From the graph it can be observed that countries with very large economies emit more greenhouse gases, while in the remaining groups CO2 emissions remain at a lower level.



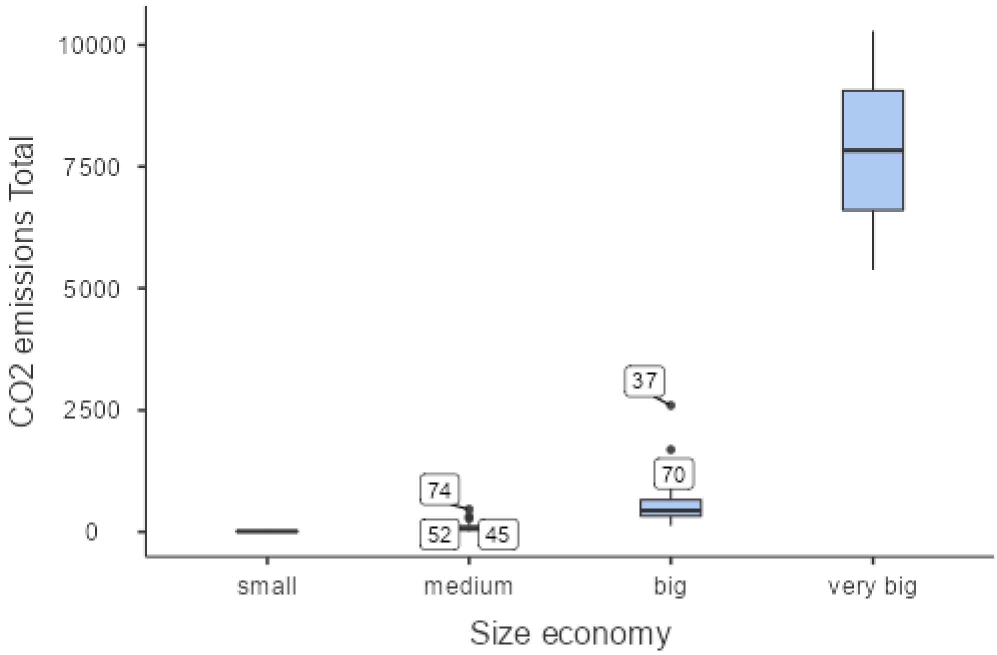
In the graph above, it can be observed that for very large and large populations, CO2 emissions are high, while the remaining country groups maintain them at a lower level.

	Size economy	CO2 emissions Total
Mean	small	12.0
	medium	100
	big	622
	very big	7833
Median	small	9.34
	medium	64.4
	big	437
	very big	7833
Standard deviation	small	7.40
	medium	88.9
	big	560
	very big	3475
IQR	small	10.2
	medium	76.3
	big	333
	very big	2457
Minimum	small	3.66
	medium	11.8
	big	128
	very big	5375
Maximum	small	30.4
	medium	464
	big	2600
	very big	10290

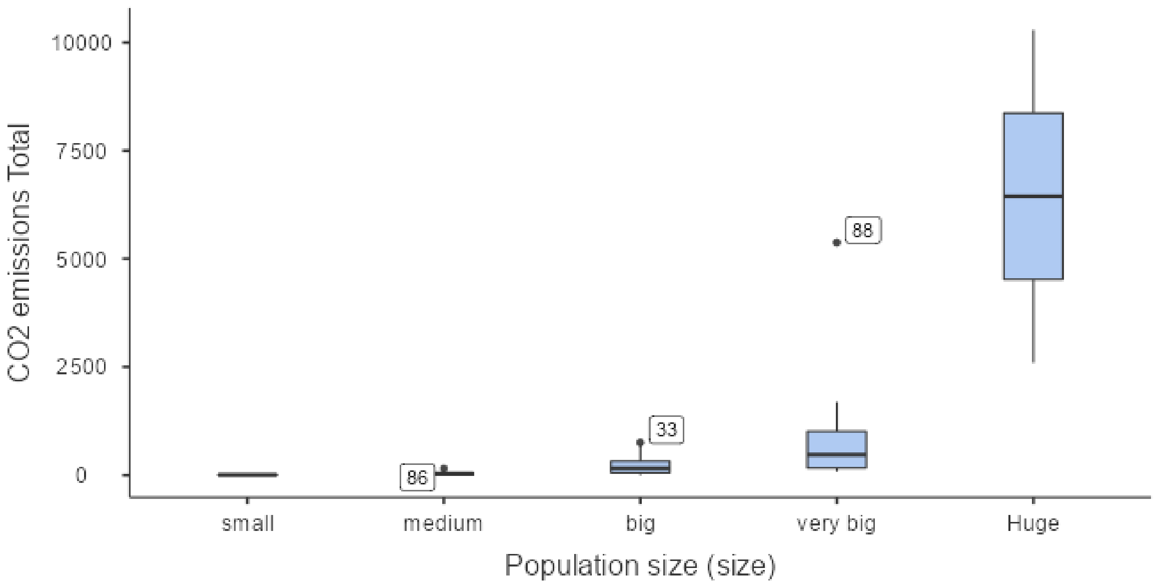
Based on the above data, it can be observed that the size of the economy has a significant impact on the amount of CO2 emissions. In the example of the average, it can be seen how a small economy emits little carbon dioxide compared to a very large economy. In the case of the median, however, it can be observed how the amount of CO2 emitted increases with the size of the economy. Also in the case of the minimum and maximum, it can be observed that the largest amounts of CO2 are emitted by large and very large countries.

	Population size (size)	CO2 emissions Total
Mean	small	6.62
	medium	36.3
	big	221
	very big	1035
	Huge	6445
Median	small	6.62
	medium	35.2
	big	160
	very big	469
	Huge	6445
Standard deviation	small	4.18
	medium	33.5
	big	207
	very big	1608
	Huge	5438
IQR	small	2.95
	medium	39.2
	big	272
	very big	846
	Huge	3845
Minimum	small	3.66
	medium	4.73
	big	8.35
	very big	82.6
Maximum	small	9.57
	medium	154
	big	754
	very big	5375
	Huge	10290

Also in the case of population size, smaller ones produce significantly less CO2 than countries with very large populations. Based on the average and median, it can be observed that groups of countries with small and medium populations produce little carbon dioxide. In the case of the group of countries with large populations, there is a large increase in CO2 emissions. Countries with very large populations generate significantly more CO2 than other groups of countries. In the case of the group of countries with small populations, the interquartile range shows that this is not a homogeneous group. In the group of countries with medium populations, the interquartile range is close to the average, which means that the values in emissions are similar between individual countries. In the groups of countries with very large and large populations, the standard deviation has a large discrepancy with the average, which means that in these groups the CO_2 emissions of individual countries have a large spread.



The above graph shows how the dispersion of CO_2 emissions increases with the growth of the size of the economy. The group of countries with a small economy emits little CO_2 and is a uniform group. In the case of a medium economy, CO_2 emissions do not change in general, but there is dispersion in the group and outlier countries appear whose volume of CO_2 emissions exceeds the maximum range. In the case of the group of countries with a large economy, the dispersion of measurements is even greater, and outlier measurements have even greater deviations from the median. The group of countries with a very large economy has the greatest dispersion among the presented groups, which means that CO_2 emissions in this group of countries are not uniform and depend on other factors.



In this case too, we can see an increase in the dispersion depending on the individual groups. In the groups of small and medium population, the dispersion relative to the median is small, which suggests that these are uniform groups and emit little CO2. The group with a large population also emits little CO2, but the dispersion in this group increases. In the case of the group with a very large population, carbon dioxide emissions increase and there is an increase in the dispersion in the group. It is worth noting that the outliner, although it has a lower population than the group of countries with a large population, emits a similar amount of CO2. In the case of countries with a large population, there is a clear increase in carbon dioxide emissions relative to the population and the size of the dispersion in the group.

Data Analysis

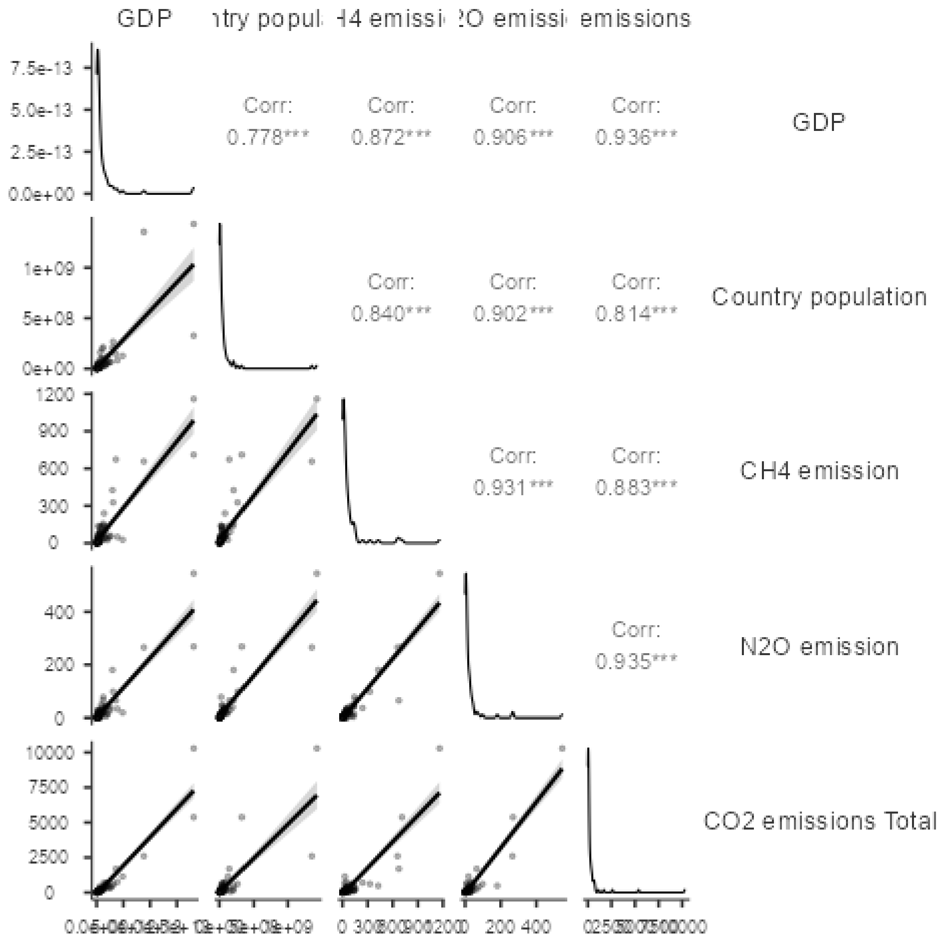
Correlation

Correlation defines the mutual relationships between selected variables. When characterizing correlation, we provide two factors: direction and strength. The numerical expression of correlation is the correlation coefficient contained in the interval [-1; 1]. We will now provide the sample with the correlation between greenhouse gas emissions and GDP and the correlation between greenhouse gas emissions and the country's population. The results are presented in the table below:

Correlation Matrix

		GDP	Country population	CH4 emission	N2O emission	CO2 emissions Total
GDP	Pearson's r	—				
	p-value	—				
Country population	Pearson's r	0.778	—			
	p-value	< .001	—			
CH4 emission	Pearson's r	0.872	0.840	—		
	p-value	< .001	< .001	—		
N2O emission	Pearson's r	0.906	0.902	0.931	—	
	p-value	< .001	< .001	< .001	—	
CO2 emissions Total	Pearson's r	0.936	0.814	0.883	0.935	—
	p-value	< .001	< .001	< .001	< .001	—

In the above table, a strong positive correlation can be observed between the size of GDP and CH4 emissions because the result of Pearson correlation is close to 1. Also in the case of population this correlation is strong. Also the case of N2O emissions and the size of GDP shows a very strong correlation between these variables. A similar situation occurs in the case of the size of the population of countries. However, in the case of carbon dioxide emissions there is also a correlation, but it is stronger with the size of the GDP of countries.



The above graphs show a strong positive correlation between greenhouse gas emissions and GDP and population size. This means that as GDP and population grow, greenhouse gas emissions increase linearly.

Relationship

A correlation relationship is a special case of stochastic relationship, which is characterized by the fact that specific values of one variable (independent variable) correspond to precisely defined conditional means of the second variable (dependent variable). In the following case, the relationship between the size of the economy and the amount of carbon dioxide emissions will be tested. The hypotheses for this case are as follows:

H_0: The size of the economy and the amount of CO2 emissions are independent

H_1: The size of the economy and the amount of CO2 emissions are dependent

The results are as follows:

Contingency Tables

		Size economy				
CO2 emission volume		small	medium	big	very big	Total
small	Observed	12	2	2	0	16
	% within column	54.5 %	4.3 %	9.1 %	0.0 %	17.4 %
medium	Observed	10	30	0	0	40
	% within column	45.5 %	65.2 %	0.0 %	0.0 %	43.5 %
big	Observed	0	14	17	0	31
	% within column	0.0 %	30.4 %	77.3 %	0.0 %	33.7 %
very big	Observed	0	0	3	2	5
	% within column	0.0 %	0.0 %	13.6 %	100.0 %	5.4 %
Total	Observed	22	46	22	2	92
	% within column	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %

χ^2 Tests

	Value	df	p
χ^2	98.6	9	< .001
N	92		

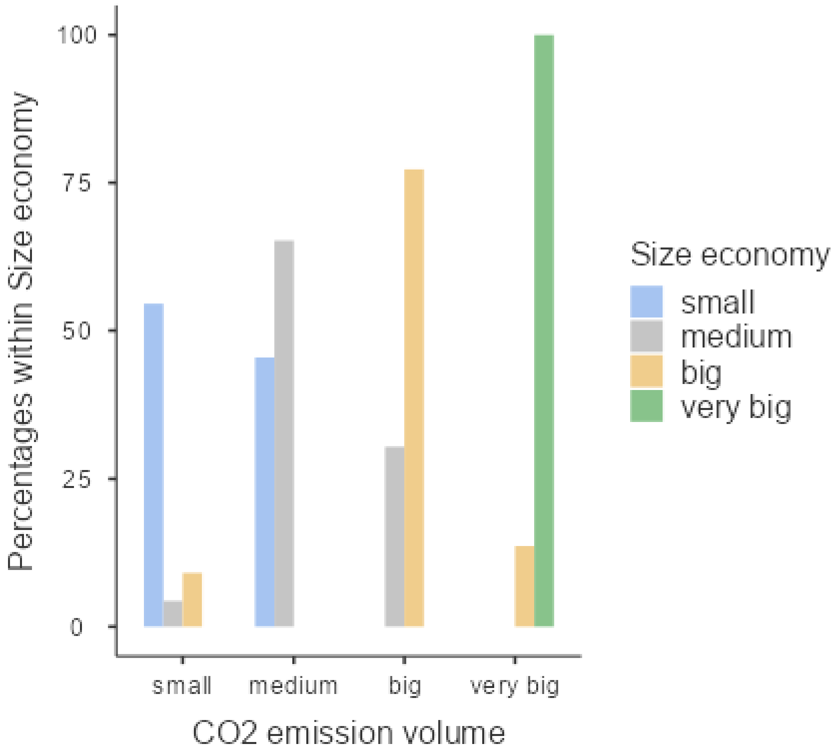
Nominal

	Value
Phi-coefficient	NaN
Cramer's V	0.598

Kendall's Tau-b

Kendall's Tau-B	t	p
0.668	7.21	< .001

Based on the result of the V-Cramer test, a moderate relationship can be observed between the size of the economy and the amount of carbon dioxide emissions. Since the result is greater than 0.3, which means a moderate relationship. The positive Kendall's Tau-b coefficient confirms this relationship, which means that when the economy grows, carbon dioxide emissions increase.



The graph above shows the relationship between the growth of CO2 emissions and the growth of the economy. These data are not clear, but the results in the previous tables confirm this relationship.

As a result of the statistical analysis conducted above, hypothesis H0 can be rejected, which means that the variables are dependent and hypothesis H1 should be accepted.

In the next case study, the relationship between carbon dioxide emissions and population size will be tested. In this case, the hypotheses are as follows:

H_0: population size and carbon dioxide emissions are independent

H_1: population size and carbon dioxide emissions are dependent

Contingency Tables

		Population size (size)					
CO2 emission volume		small	medium	big	very big	Huge	Total
small	Observed	2	10	4	0	0	16
	% within column	100.0 %	29.4 %	9.1 %	0.0 %	0.0 %	17.4 %
medium	Observed	0	23	16	1	0	40
	% within column	0.0 %	67.6 %	36.4 %	10.0 %	0.0 %	43.5 %
big	Observed	0	1	24	6	0	31
	% within column	0.0 %	2.9 %	54.5 %	60.0 %	0.0 %	33.7 %
very big	Observed	0	0	0	3	2	5
	% within column	0.0 %	0.0 %	0.0 %	30.0 %	100.0 %	5.4 %
Total	Observed	2	34	44	10	2	92
	% within column	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %

χ^2 Tests

	Value	df	p
χ^2	90.9	12	< .001
N	92		

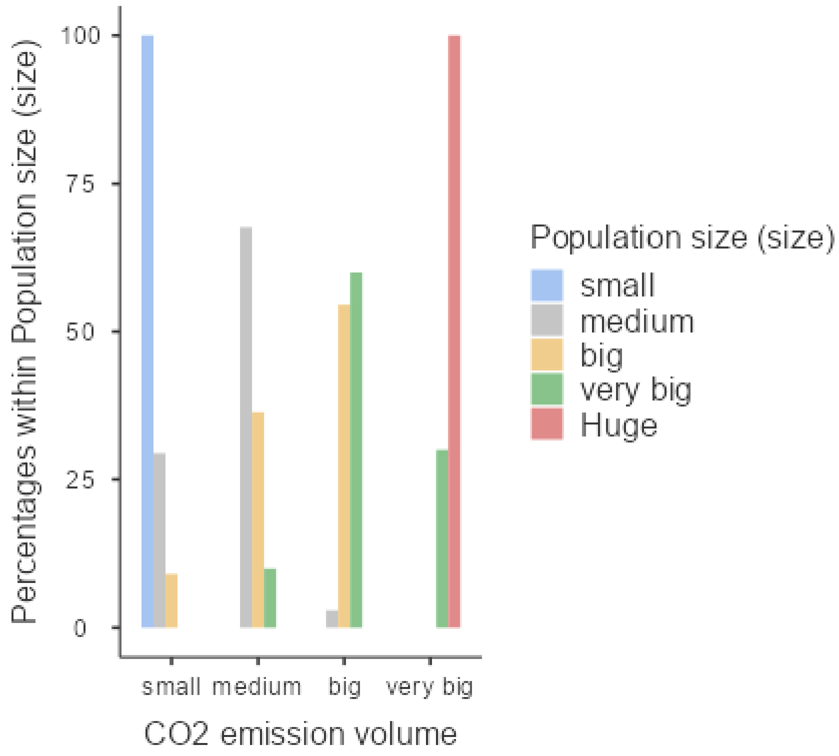
Nominal

	Value
Phi-coefficient	NaN
Cramer's V	0.574

Kendall's Tau-b

Kendall's Tau-B	t	p
0.626	6.72	< .001

Also in the case of population size, the value of the Cramer's V test result indicates a moderate relationship with the increase in carbon monoxide emissions. Kendall's Tau-b also confirms this relationship.



The above graph shows the percentage of a given population in the amount of CO2 emissions, also in this case the relationship is visible. Because with the increase in population, carbon dioxide pollution increases.

The results prove that hypothesis H0 should be rejected, which means that the hypothesis about the relationship between population size and carbon dioxide emissions is true.

T-tests

Student's t-tests are used to compare two groups. We use them when we have results for two groups and we want to compare them - that is, to determine whether the results in one group are greater or less than in the other group. You cannot compare several groups by performing the Student's t-test several times. In this case, the change in the amount of carbon dioxide emissions in 2018 compared to 2017 will be tested. For this test, the hypotheses are as follows:

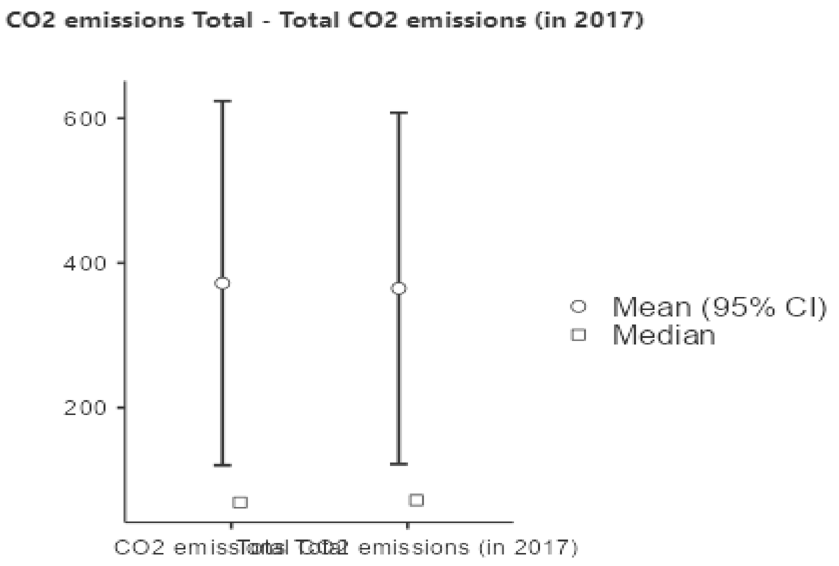
H_0: carbon dioxide emissions in 2017 and 2018 were the same

H_1: carbon dioxide emissions in 2017 and 2018 were different

Paired Samples T-Test

			statistic	df	p	Mean difference	SE difference
CO2 emissions Total	Total CO2 emissions (in 2017)	Student's t	1.47	91.0	0.145	7.19	4.89

The Student's T-test shows that there was an increase in CO2 emissions in 2018 compared to 2017. A high p-value above alpha=0.05 proves the high reliability of the test that the values are not equal. On the other hand, a shift in the mean by 7.19 proves an increase in carbon dioxide emissions.



The above graph also shows an increase in CO_2 emissions over the two years. There is a clear shift in the average upwards and an increase in the range of dispersion among the higher values.

The results show that hypothesis H0 should be rejected, which means that the hypothesis about the change in carbon dioxide emissions in 2017 and 2018 is true.

Variance analysis

Single-factor ANOVA is the simplest form of ANOVA, where the response variable can depend on only one factor. Single-factor ANOVA contains at least two experimental conditions and allows for the comparison of means from two or more groups simultaneously. In this case, we will use the parametric ANOVA test to analyze the changes in carbon dioxide emissions over five years in economies of different sizes. The hypotheses for this test are as follows:

H_0: Over the years, the amount of CO2 emissions for different economies has not changed $U1=U2=U3=U4=U5$

H_1: Over the years, the amount of CO2 emissions for different economies has varied $U1 \neq U2 \neq U3 \neq U4 \neq U5$

One-Way ANOVA (Welch's)

	F	df1	df2	p
CO2 emissions Total	21.0	3	4.68	0.004
Total CO2 emissions (in 2017)	22.1	3	4.68	0.003
Total CO2 emissions (in 2016)	22.2	3	4.68	0.003
Total CO2 emissions (in 2015)	23.0	3	4.68	0.003
Total CO2 emissions (in 2014)	22.1	3	4.68	0.003

Group Descriptives

	Size economy	N	Mean	SD	SE
CO2 emissions Total	small	22	12.0	7.40	1.58
	medium	46	100.3	88.91	13.11
	big	22	622.2	559.80	119.35
	very big	2	7832.7	3475.07	2457.25
Total CO2 emissions (in 2017)	small	22	12.0	7.94	1.69
	medium	46	99.9	87.03	12.83
	big	22	617.3	529.42	112.87
	very big	2	7564.1	3332.39	2356.35
Total CO2 emissions (in 2016)	small	22	11.8	7.82	1.67
	medium	46	99.8	87.25	12.86
	big	22	606.8	521.34	111.15
	very big	2	7484.2	3162.48	2236.21
Total CO2 emissions (in 2015)	small	22	11.5	7.50	1.60
	medium	46	98.0	84.54	12.46
	big	22	602.6	506.68	108.02
	very big	2	7610.1	3165.47	2238.32
Total CO2 emissions (in 2014)	small	22	11.6	7.69	1.64
	medium	46	97.4	87.80	12.95
	big	22	594.7	497.38	106.04
	very big	2	7754.2	3155.66	2231.39

The results of Welch's test indicate that the thesis that the amount of carbon dioxide emissions over the years in economies of different sizes was the same is false. Because the p value is less than the significance level of alpha = 0.05. Confirmation of this result can be observed in the example of the average that increased over the years. Therefore, the assumption of the equality of CO2 emissions over the years is incorrect. It can also be assumed that with the growth of the economy, the amount of carbon dioxide emitted increases.

The above results prove that hypothesis H0 is incorrect, which means that carbon dioxide emissions have changed over the years.

In this case, we will use a parametric ANOVA test to analyze the changes in carbon dioxide emissions over five years in populations of different sizes. The hypotheses for this test are as follows:

H_0: Over the years, the amount of CO2 emissions for different populations has not changed U1=U2=U3=U4=U5

H_1: Over the years, CO2 emissions for different populations have varied

U1 ≠U2 ≠ U3≠U4 ≠U5

One-Way ANOVA (Welch's)

	F	df1	df2	p
Total CO2 emissions (in 2014)	13.2	4	6.68	0.003
Total CO2 emissions (in 2015)	13.4	4	6.77	0.002
Total CO2 emissions (in 2016)	13.8	4	6.78	0.002
Total CO2 emissions (in 2017)	14.3	4	6.71	0.002
CO2 emissions Total	14.2	4	6.65	0.002

	Population size (number)	N	Mean	SD	SE
Total CO2 emissions (in 2014)	1	2	6.65	4.51	3.19
	2	35	36.78	37.52	6.34
	3	43	219.99	203.12	30.98
	4	10	1039.84	1658.84	524.57
	5	2	6085.72	5515.24	3899.86
Total CO2 emissions (in 2015)	1	2	6.44	4.11	2.91
	2	35	37.32	39.82	6.73
	3	43	222.11	205.20	31.29
	4	10	1024.92	1611.02	509.45
	5	2	6058.49	5359.77	3789.93
Total CO2 emissions (in 2016)	1	2	6.29	3.96	2.80
	2	35	37.61	39.74	6.72
	3	43	223.46	204.28	31.15
	4	10	1012.83	1571.26	496.88
	5	2	6051.33	5188.91	3669.11
Total CO2 emissions (in 2017)	1	2	6.43	4.00	2.83
	2	35	36.74	34.41	5.82
	3	43	227.36	208.90	31.86
	4	10	1014.20	1557.87	492.64
	5	2	6177.16	5293.83	3743.30
CO2 emissions Total	1	2	6.62	4.18	2.95
	2	35	36.29	32.98	5.58
	3	43	225.72	207.01	31.57
	4	10	1035.40	1607.75	508.41
	5	2	6444.90	5437.78	3845.09

The results of Welch's test indicate that the thesis that the amount of carbon dioxide emissions over the years in populations of different sizes was the same is false. Because the p value is less than the significance level of alpha = 0.05. Confirmation of this result can be observed in the example of the standard deviation, which increased over the years (excluding the small population). Therefore, the assumption of the equality of CO2 emissions over the years is incorrect. It can also be assumed that with the growth of the population, the amount of carbon dioxide emitted increases.

The above results prove that hypothesis H0 is incorrect, which means that carbon dioxide emissions have changed over the years.

The Kruskal-Wallis test is an extension of the Mann-Whitney U test. This test is a nonparametric equivalent of the one-way analysis of variance. It detects differences in the location of the distribution. This test assumes that there is no a priori ordering of the k populations from which the samples were taken. In order to test whether economies produce greenhouse gases in different amounts, we will use nonparametric analysis of variance. The test should be performed at a significance level of alpha=0.05. The hypotheses are as follows:

H_0: the amount of greenhouse gas emissions for different economies is the same

H_1: the amount of greenhouse gas emissions for different economies is different

Kruskal-Wallis

	χ^2	df	p
CH4 emission	52.2	3	< .001
N2O emission	56.7	3	< .001
CO2 emissions Total	73.4	3	< .001

The low result of the Kruskal-Wallis test proves that economies of different sizes produce greenhouse gases in different amounts. Since the p-value is significantly lower than the significance level of alpha=0.05, the hypothesis of equality can be unequivocally rejected.

The test results show that the hypothesis H0 should be rejected, which means that different economies produce greenhouse gases to different extents.

In order to test whether populations produce greenhouse gases to different extents, we will use nonparametric analysis of variance. The significance level for the test should be alpha=0.05. The hypotheses are as follows:

H_0: the amount of greenhouse gas emissions for different populations is the same

H_1: the amount of greenhouse gas emissions for different populations is different

Kruskal-Wallis

	χ^2	df	p
CH4 emission	56.3	4	< .001
N2O emission	59.2	4	< .001
CO2 Emissions Cement	47.4	4	< .001

The low result of the Kruskal-Wallis test indicates that populations of different sizes produce different amounts of greenhouse gases. Since the p-value is significantly lower than the significance level of alpha=0.05, the hypothesis of equality can be unequivocally rejected.

The test results show that the hypothesis H0 should be rejected, which means that different economies produce greenhouse gases to different extents.

Multiple regression analysis

The general goal of multiple regression (the term was first used by Pearson in 1908) is to quantify the relationships between a number of independent (explanatory) variables and a dependent (criterion) variable. For example, a real estate agent collects data on buildings—size (in square meters), number of bedrooms, average income of residents in a neighborhood, and subjective ratings of attractiveness. If they already have a database of this type, they might try to answer the following question: How do the individual variables affect the price of a building? In this way, they might find out that the number of bedrooms is a better explanation of the price of a building than, for example, how nice it looks to the naked eye (subjective attractiveness). They might also discover "outliers," that is, buildings that are worth more than the data collected by the agent.

Using multiple regression it will be determined what GDP growth causes an increase in greenhouse gas emissions by 1 Mt and what effect population growth has on CO2 emissions.

Model Fit Measures

							Overall Model Test			
Model	R	R^2	Adjusted R^2	AIC	BIC	RMSE	F	df1	df2	p
1	0.946	0.895	0.893	1370	1380	397	379	2	89	< .001

Omnibus ANOVA Test

	Sum of Squares	df	Mean Square	F	p
GDP	3.22e+7	1	3.22e+7	197.5	< .001
Country population	2.51e+6	1	2.51e+6	15.4	< .001
Residuals	1.45e+7	89	162816		

Note. Type 3 sum of squares

[3]

Model Coefficients - CO2 emissions Total

			95% Confidence Interval					95% Confidence Interval	
Predictor	Estimate	SE	Lower	Upper	t	p	Stand. Estimate	Lower	Upper
Intercept	−113	45.6	−204	−22.5	−2.48	0.015			
GDP	3.34e−10	2.37e−11	2.86e−10	3.81e−10	14.05	< .001	0.769	0.660	0.878
Country population	1.28e−6	3.26e −7	6.32e−7	1.93e−6	3.93	< .001	0.215	0.106	0.324

Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.0615	1.88	0.440

[3]

Collinearity Statistics

	VIF	Tolerance
GDP	2.54	0.394
Country population	2.54	0.394

[3]

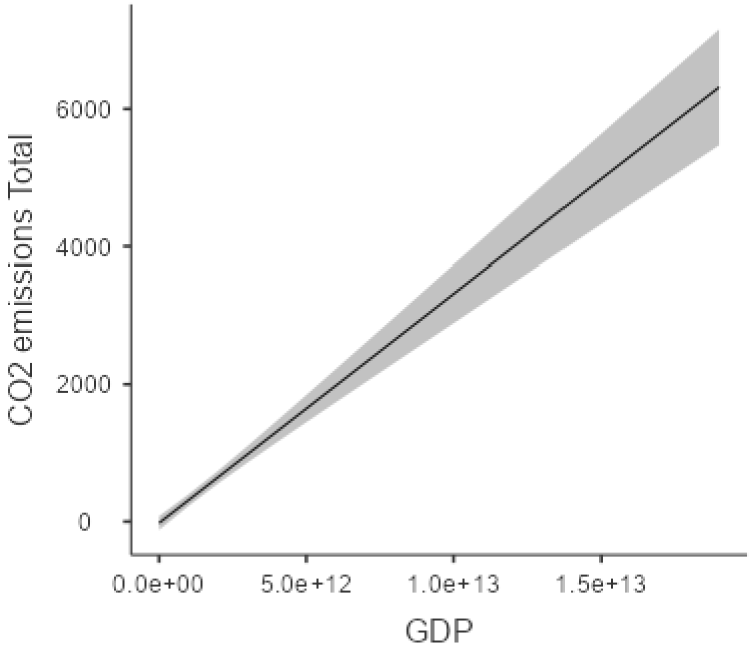
Normality Test (Shapiro-Wilk)

Statistic	p
0.564	< .001

The correlation coefficient shows a low correlation between these three variables, but its significance level is very high. A low correlation means that the model rules are not violated. The VIF coefficient indicates a very strong linearity of the model. However, the normality test does not meet the assumptions of the normal distribution. The R-squared coefficient indicates a high model fit.

The results obtained through modeling indicate that with the increase of Gross Domestic Product by 33.4 billion dollars, the emission of carbon dioxide increases by 1 megaton. Also, in the case of the increase of the population by 12.8 million, the production of CO_2 increases by 1 megaton. Looking at the standard deviation, we can see that the increase of GDP contributes more to the increase of carbon dioxide emissions than the increase of the population. In conclusion, we can assume that the improvement of the standard of living contributes to the increase of greenhouse gas emissions.

GDP

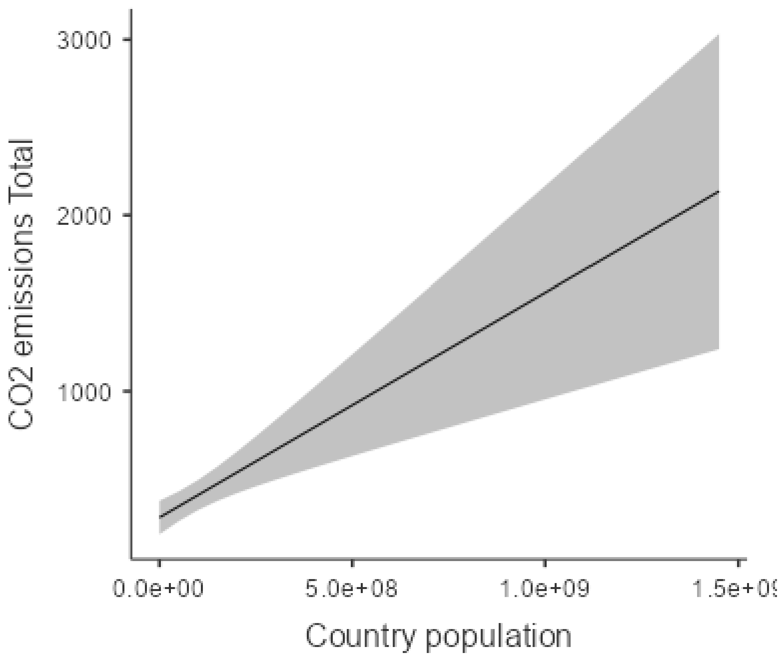


Estimated Marginal Means - GDP

			95% Confidence Interval	
GDP	Marginal Mean	SE	Lower	Upper
−1.66e−12 ⁻	−575	79.4	−733	−417
1.18e+12 μ	372	42.1	289	456
4.02e+12 ⁺	1319	79.4	1161	1477

Note. ⁻ mean - 1SD, μ mean, ⁺ mean + 1SD

Country population



Estimated Marginal Means - Country population

			95% Confidence Interval	
Country population	Marginal Mean	SE	Lower	Upper
−1.35e−8 ⁻	107	79.4	−50.3	265
7.15e+7 µ	372	42.1	288.5	456
2.78e+8 ⁺	637	79.4	478.9	795

Note. ⁻ mean - 1SD, µ mean, ⁺ mean + 1SD

The conclusions drawn from the tables are also visible in the graphs. In the case of GDP, there is a relationship between the growth of carbon dioxide emissions and the growth of the standard deviation with the growth of emissions. On the other hand, CO2 emissions to the country's population show a lower growth with the growth of the population. Comparison of both graphs confirms the thesis that GDP growth has a stronger effect on the growth of carbon dioxide emissions than population growth.

Logistic regression analysis (renewable energy data not fully reliable)

The logistic regression model is a special case of the generalized linear model. It is used when the dependent variable is dichotomous, that is, it takes only two values, such as success or failure, the presence or absence of a certain disease entity, female or male. In mathematical notation, these values are represented as 1 and 0. Logistic regression does not have many of the key assumptions of linear regression and general linear models, which are based on least squares algorithms - especially linearity, normality, homoscedasticity, and level of measurement. Logistic regression requires that the dependent variable be binary, and ordinal logistic regression requires that the dependent variable be ordinal. Reducing an ordinal or metric variable to a dichotomous level results in the loss of a large amount of information, which makes such a test inferior compared to ordinal logistic regression in such a case. In order to examine the effect of renewable energy sources on carbon dioxide emissions, logistic regression will be applied to three fossil fuels.

Model Fit Measures

Model	Deviance	AIC	R^2_{McF}
1	123	131	0.0360

Model Coefficients - Renewable energy sources

						95% Confidence Interval	
Predictor	Estimate	SE	Z	p	Odds ratio	Lower	Upper
Intercept	0.16407	0.23108	0.710	0.478	1.178	0.749	1.85
CO2 Emissions Coal	−0.00213	0.00175	−1.215	0.224	0.998	0.994	1.00
CO2 Emissions Gas	−0.00566	0.00458	−1.236	0.216	0.994	0.985	1.00
CO2 emissions Oil	0.00576	0.00403	1.429	0.153	1.006	0.998	1.01

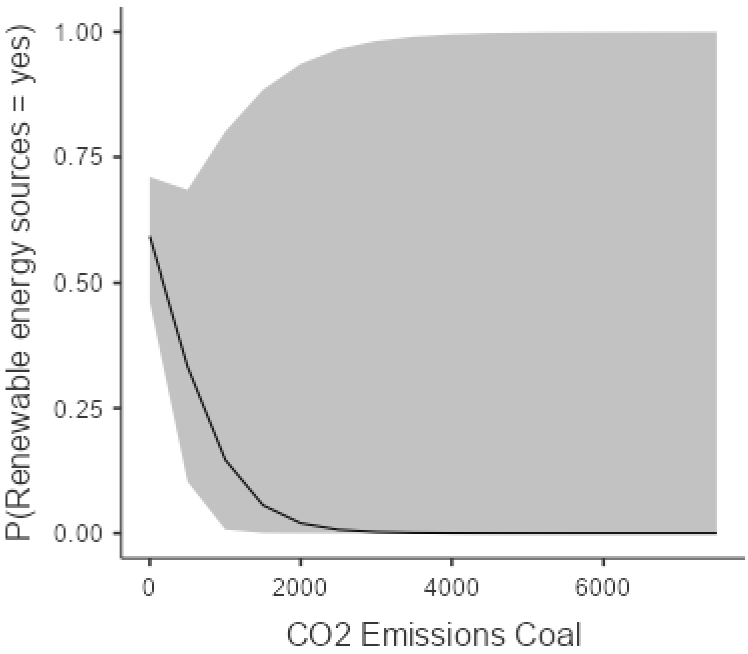
Note. Estimates represent the log odds of "Renewable energy sources = yes" vs. "Renewable energy sources = no"

Assumption Checks

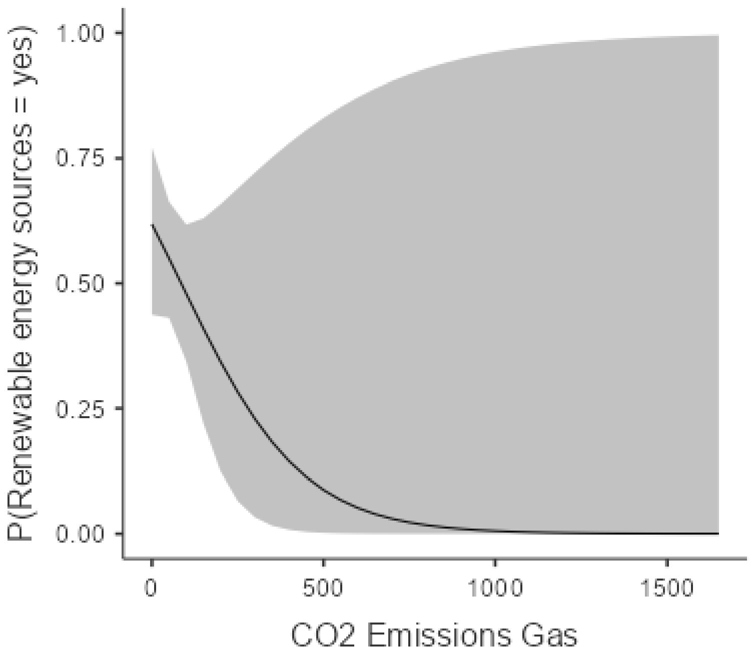
Collinearity Statistics

	VIF	Tolerance
CO2 Emissions Coal	3.18	0.3144
CO2 Emissions Gas	10.46	0.0956
CO2 emissions Oil	16.12	0.0620

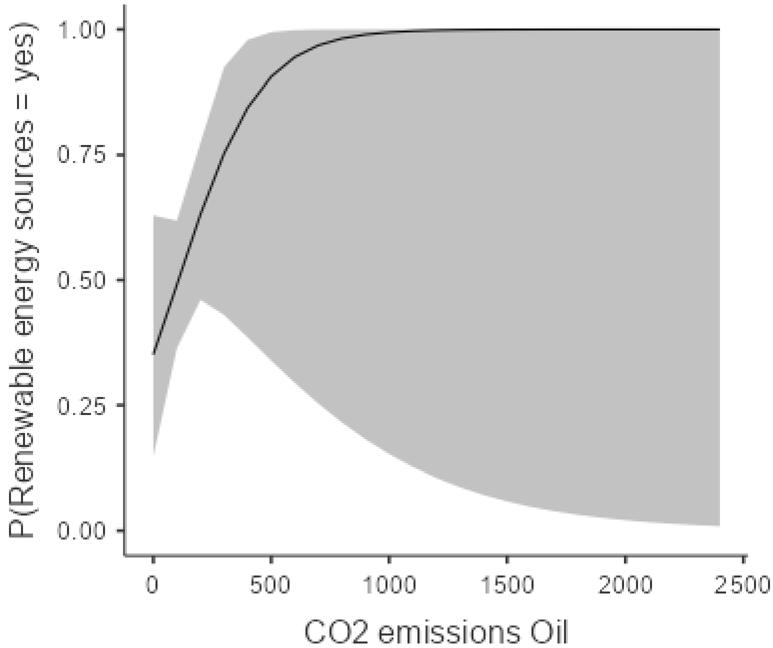
CO2 Emissions Coal



CO2 Emissions Gas



CO2 emissions Oil



The R-squared index indicates a low fit of the model. The p-value in all cases is higher than the significance level of alpha=0.05. The value of the estimate indicates that the amount of greenhouse gas emissions from coal and gas will decrease when the country has RES. In the case of oil and renewable energy sources, the value of carbon dioxide emissions will increase. The VIF index indicates a strong linearity of the model.

The graphs provided confirm the results from the table. In the case of using renewable energy sources, the probability of CO2 emissions from coal will be low. Also, when a country uses renewable energy sources, the probability of carbon dioxide emissions from gas decreases. The opposite situation occurs in the case of using renewable energy sources and carbon dioxide emissions from oil. Because the probability of carbon dioxide emissions increases when a country uses renewable energy sources.

Summary

Based on the above study, it can be concluded that there is a correlation between the growth of Gross Domestic Product and greenhouse gas emissions. The correlation is also visible in the case of population and greenhouse gas emissions. Both of these correlations also have a relationship, which means that when GDP and population increase, greenhouse gas emissions increase. Also, the Student's T-test proved that over the years 2017 and 2018, CO_2 production increased. On the other hand, the parametric analysis of variance proves that the increase in greenhouse gas emissions in 2018 was not an exception, this increase is a constantly increasing trend. Then the result of the parametric analysis jumps in that the size of the economy and population have an impact on the amount of greenhouse gas emissions. On the other hand, the linear regression analysis provided more detailed data. According to which, with the increase in Gross Domestic Product by 33.4 billion dollars, carbon dioxide emissions increase by 1 megaton. Also, in the case of an increase in population by 12.8 million, CO_2 production increases by 1 megaton. Linear regression also showed that the increase in GDP contributes more to the increase in carbon dioxide emissions than the increase in population. From this, it can be concluded that the increase in the standard of living of people contributes to the increase in CO_2 emissions. However, an unexpected result was provided by the logistic analysis, which confirmed that when moving away from gas and coal to renewable energy sources, the amount of CO_2 emissions decreases, but in the case of oil, the move may cause an increase in carbon dioxide emissions.

In summary, it can be said that the size of the economy has a significant impact on carbon dioxide emissions. However, the size of the population does not play as key a role as the level at which it lives. Most emissions are concentrated in very large populations and economies, and the numbers of small and medium-sized ones should not be ignored. It would also be necessary to look at the transition from fossil fuels to renewable energy sources so that this does not cause a feedback loop.