

# Analiza statystyczna danych – emisja CO<sub>2</sub> na świecie.

Tomasz Wira

Wrocław 2023

# Wstęp

W dzisiejszych czasach temat globalnego ocieplenia jest wszystkim dobrze znany. Emisja gazów cieplarnianych w ostatnim czasie dość często pojawia się w debacie publicznej. Problem ten dotyczy całego świata ponieważ procesy przebiegające w przyrodzie oraz w poszczególnych gospodarkach są od siebie zależne. Aby zrozumieć jak silnie oddziałuje gospodarka na środowisko naturalne w niniejszej pracy zostanie przeprowadzona analiza emisji gazów cieplarnianych w odniesieniu do poszczególnych gospodarek.

Poniższa praca przedstawia analizę emisji gazów cieplarnianych w roku 2018 i została ona podzielona na trzy części w pierwszej zostaną przedstawione dane statystyczne metodą opisową oraz statystyczną. W drugiej części pracy zostanie przeprowadzona dokładna analiza danych statystycznych które zostały pozyskane z pomocą serwisu internetowego CORGIS Project<sup>1</sup>. Zebrane dane zostaną poddane analizie poprzez zastosowanie metod tj. test zależności, T-testy, analiza wariancji, analiza regresji wielorakiej oraz analiza regresji logistycznej. W trzeciej części niniejszej pracy zostaną podsumowane wyniki przeprowadzonych badań i pomiarów. W ostatniej części zostaną także wyciągnięte wnioski na podstawie wyników przeprowadzonego badania analizy statystycznej.

W pracy tej zostaną wykorzystane metody badawcze tj. metoda statystyczna, metoda opisowa, metoda analizy statystycznej oraz metoda heurystyczna. Aby jak najlepiej zobrazować wyniki prowadzonego badania będą dodawane rysunki przedstawiające wyniki analiz. Do badania w celach statystycznych zostanie wykorzystany program Microsoft Excel. W celach przeprowadzenia dogłębnej analizy statystycznej zostaną wykorzystane typowe dla tego celu programy jak Jamovi oraz GRETL. Za zaplecze teoretyczne posłużą informacje zdobyte na kursie oraz Elektronic Textbook StatSoft<sup>2</sup>.

A więc rozważania w poniższej pracy powinny pozwolić na uzyskanie odpowiedzi pytanie czy wielkość populacji oraz wielkość gospodarki ma wpływ na rozmiar produkcji gazów cieplarnianych? Dane zebrane oraz poddane analizie statystycznej powinny jednoznacznie odpowiedzieć czy takie zależności istnieją oraz jak silne one są.

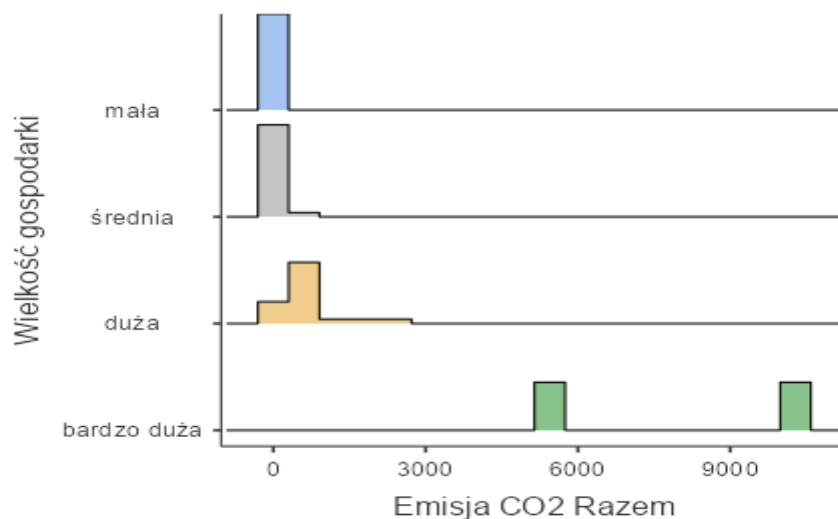
---

<sup>1</sup> CORGI Project, <https://think.cs.vt.edu/corgis/csv/> ([dostęp.25.01.2023r](#))

<sup>2</sup> StatSoft, [https://www.statsoft.pl/textbook/stathome\\_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fadwans4.html](https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fadwans4.html) ([dostęp.25.01.2023r](#))

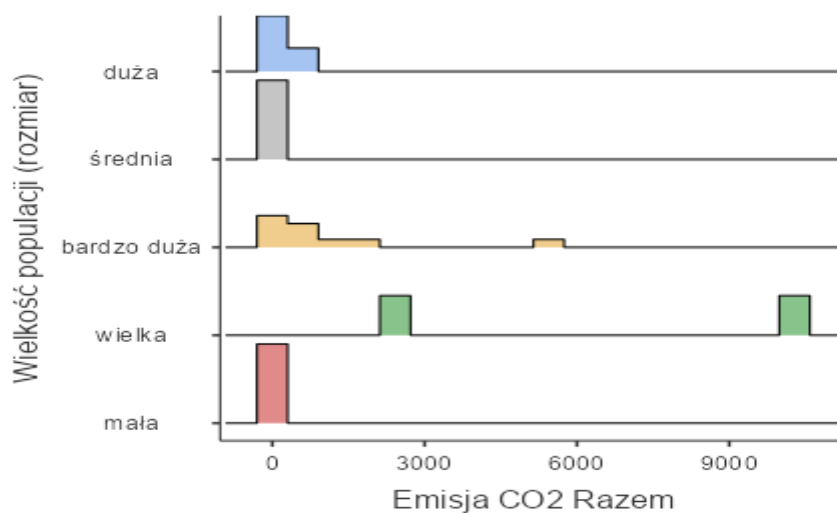
# Opis danych

Emisja CO2 Razem



Na podstawie wykresu można zaobserwować, że kraje posiadające bardzo duże gospodarki emitują większą ilość gazów cieplarnianych, natomiast w pozostałych grupach emisja CO2 utrzymuje się na niższym poziomie.

Emisja CO2 Razem



Na powyższym wykresie można zaobserwować, że w przypadku bardzo dużej oraz wielkiej populacji emisja CO2 jest na wysokim poziomie, natomiast pozostałe grupy krajów utrzymują ją na niższym poziomie.

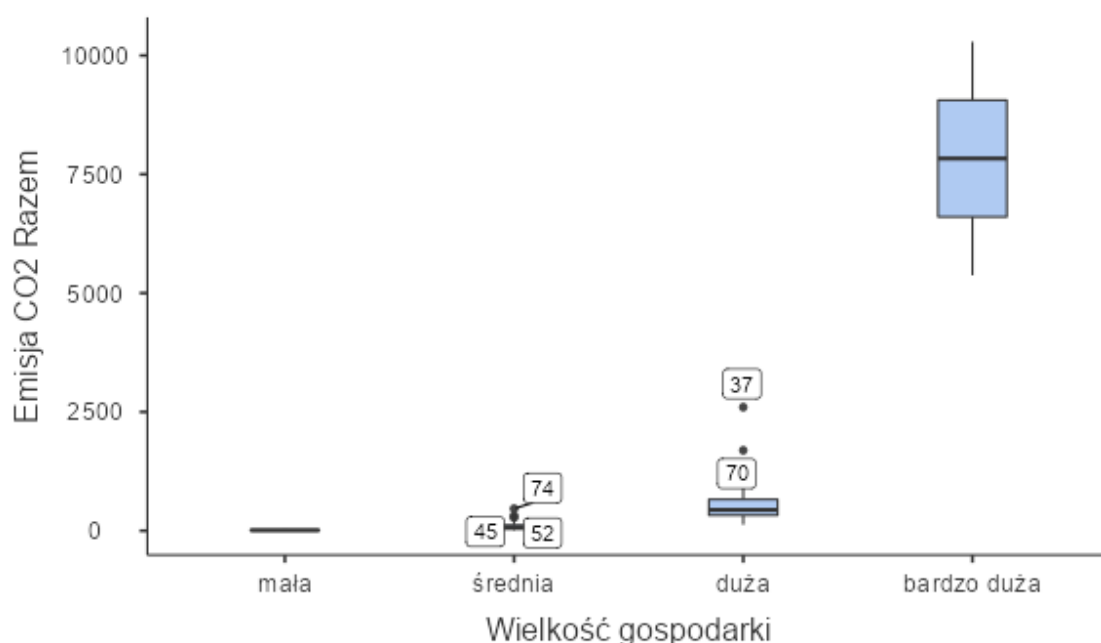
	Wielkość gospodarki	Emisja CO2 Razem
Mean	mała	12.0
	średnia	100
	duża	622
	bardzo duża	7833
Median	mała	9.34
	średnia	64.4
	duża	437
	bardzo duża	7833
Standard deviation	mała	7.40
	średnia	88.9
	duża	560
	bardzo duża	3475
IQR	mała	10.2
	średnia	76.3
	duża	333
	bardzo duża	2457
Minimum	mała	3.66
	średnia	11.8
	duża	128
	bardzo duża	5375
Maximum	mała	30.4
	średnia	464
	duża	2600
	bardzo duża	10290

Na podstawie powyższych danych można zaobserwować, że rozmiar gospodarki ma istotny wpływ na wielkość emisji CO<sub>2</sub>. Na przykładzie średniej można zaobserwować jak mała gospodarka emituje niewiele dwutlenku w porównaniu z gospodarką bardzo dużą. Natomiast w przypadku mediany można zaobserwować jak wraz z rozmiarem gospodarki rośnie ilość emitowanego CO<sub>2</sub>. Także w przypadku minimum i maksimum można zaobserwować że najwięcej CO<sub>2</sub> emitują duże oraz bardzo duże kraje.

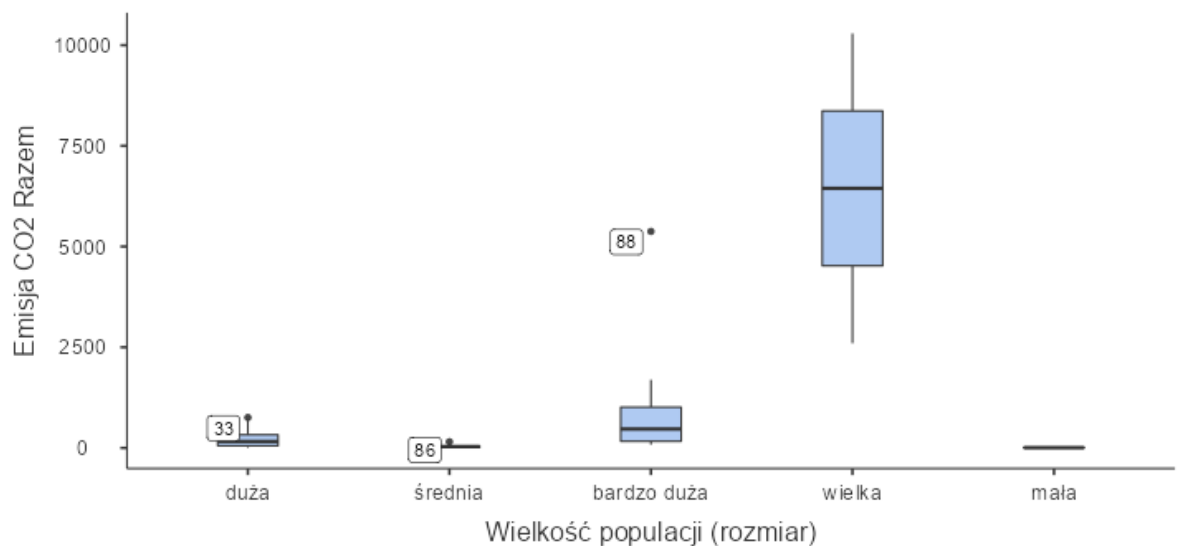
	Wielkość populacji (rozmiar)	Emisja CO2 Razem
Mean	duża	221
	średnia	36.3
	bardzo duża	1035
	wielka	6445
	mała	6.62
Median	duża	160
	średnia	35.2
	bardzo duża	469
	wielka	6445
	mała	6.62
Standard deviation	duża	207
	średnia	33.5
	bardzo duża	1608
	wielka	5438
	mała	4.18
IQR	duża	272
	średnia	39.2
	bardzo duża	846
	wielka	3845
	mała	2.95
Minimum	duża	8.35
	średnia	4.73
	bardzo duża	82.6
	wielka	2600
	mała	3.66
Maximum	duża	754
	średnia	154
	bardzo duża	5375
	wielka	10290
	mała	9.57

Także w przypadku wielkości populacji mniejsze produkują zdecydowanie mniej CO<sub>2</sub> niż kraje posiadające bardzo duże populacje. Na podstawie średniej oraz mediany można zaobserwować, że grupy krajów z małą oraz średnią populacją produkują mało dwutlenku węgla. Natomiast w przypadku grupy krajów posiadających dużą populację następuje duży wzrost emisji CO<sub>2</sub>. Kraje posiadające bardzo dużą populację generują zdecydowanie więcej CO<sub>2</sub> niż pozostałe grupy krajów. W przypadku grupy krajów posiadających małą populację zakres międzykwartylowy pokazuje, że nie jest to grupa jednorodna. W grupie krajów posiadających średnią populację zakres międzykwartylowy jest zbliżony do średniej co oznacza

że wartości w emisji są podobne między poszczególnymi krajami. W grupach krajów posiadających bardzo dużą oraz wielką populację odchylenie standardowe posiada dużą rozbieżność ze średnią co oznacza, że w grupach tych emisja CO<sub>2</sub> poszczególnych krajów posiada duży rozrzut.



Na powyższym wykresie można zaobserwować jak wraz ze wzrostem wielkości gospodarki rośnie rozrzut w emisji CO<sub>2</sub>. Grupa krajów posiadających małą gospodarkę emituje mało CO<sub>2</sub> oraz jest to grupa jednolita. Natomiast w przypadku średniej gospodarki emisja CO<sub>2</sub> w ujęciu ogólnym nie zmienia się, ale następuje rozproszenie w grupie oraz pojawiają się państwa odstające których objętość emisji CO<sub>2</sub> wykracza poza zakres maksymalny. W przypadku grupy krajów posiadających dużą gospodarkę rozrzut pomiarów jest jeszcze większy, a pomiary odstające posiadają jeszcze większe odstępstwa od mediany. Grupa krajów posiadających bardzo dużą gospodarkę posiada największe rozproszenie z pośród prezentowanych grup co oznacza, że emisja CO<sub>2</sub> w tej grupie krajów jest niejednorodna i zależy od innych czynników.



Także w tym przypadku można zauważyć wzrost rozrzutu w zależności od poszczególnych grup. W grupach małej i średniej populacja rozrzut względem mediany jest niewielki co sugeruje że są to grupy jednolite i emitują niewiele CO<sub>2</sub>. Także grupa o dużej populacji emituje niewiele CO<sub>2</sub> ale rozrzut w tej grupie się zwiększa. W przypadku grupy posiadającej bardzo dużą populację emisja dwutlenku węgla zwiększa się oraz następuje wzrost rozrzutu w grupie. Wartym odnotowania jest fakt outlierów który mimo, iż posiada populację niższą niż grupa krajów z wielką populacją to emituje podobną ilość CO<sub>2</sub>. Natomiast w przypadku krajów z wielką populacją widać tu wyraźny wzrost emisji dwutlenku węgla względem populacji oraz wielkość rozrzutu w grupie.

# Analiza Danych

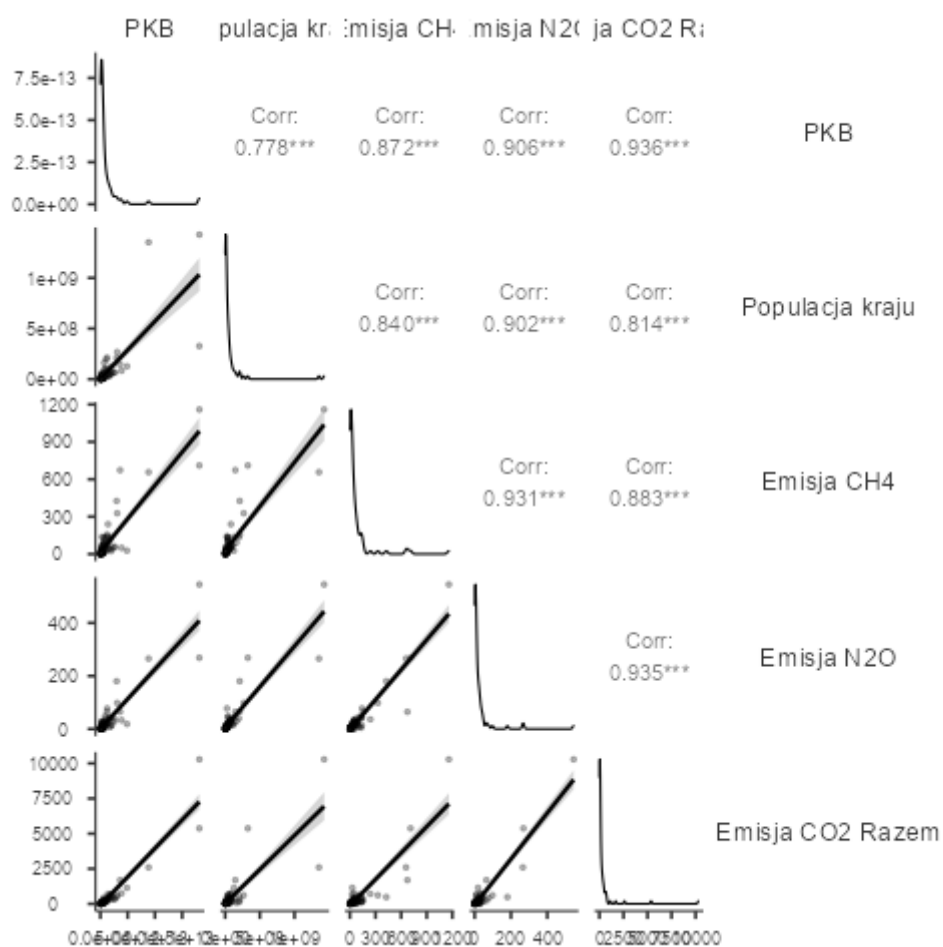
## Korelacja

Korelacja określa wzajemne powiązania pomiędzy wybranymi zmiennymi. Charakteryzując korelację podajemy dwa czynniki: kierunek oraz siłę. Wyrazem liczbowym korelacji jest współczynnik korelacji zawierający się w przedziale [-1; 1]. Podamy teraz próbie korelację pomiędzy emisją gazów cieplarnianych i PKB oraz korelację pomiędzy emisją gazów cieplarnianych a populacją kraju. Wyniki przedstawia poniższa tabela:

Correlation Matrix		PKB	Populacja kraju	Emisja CH4	Emisja N2O	Emisja CO2 Razem
PKB	Pearson's r	—				
	p-value	—				
Populacja kraju	Pearson's r	0.778	—			
	p-value	< .001	—			
Emisja CH4	Pearson's r	0.872	0.840	—		
	p-value	< .001	< .001	—		
Emisja N2O	Pearson's r	0.906	0.902	0.931	—	
	p-value	< .001	< .001	< .001	—	
Emisja CO2 Razem	Pearson's r	0.936	0.814	0.883	0.935	—
	p-value	< .001	< .001	< .001	< .001	—

W powyższej tabeli można zaobserwować silną dodatnią korelację wielkości PKB i emisji CH4 ponieważ wynik korelacji Pearsona jest bliski 1. Również w przypadku populacji ta korelacja jest silna. Także przypadek emisji N2O oraz wielkości PKB pokazuje bardzo silną korelację pomiędzy tymi zmiennymi. Podobna sytuacja występuje w przypadku wielkości populacji państw. Natomiast w przypadku emisji dwutlenku węgla także występuje korelacja, ale jest ona silniejsza przy wielkości PKB krajów.





Powyższe wykresy przedstawiają silną dodatnią korelację emisji gazów cieplarnianych oraz wielkości PKB i wielkości populacji. Co oznacza, że razem ze wzrostem PKB oraz wzrostem populacji zwiększa się liniowo emisja gazów cieplarnianych.

## Zależność

Zależnością korelacyjną nazywamy szczególny przypadek zależności stochastycznej, który charakteryzuje się tym, że określonym wartościom jednej zmiennej (zmiennej niezależnej) odpowiadają ściśle określone warunkowe średnie drugiej zmiennej (zmiennej zależnej). W poniższym przypadku będzie testowana zależność pomiędzy rozmiarem gospodarki a wielkością emisji dwutlenku węgla. Hipotezy dla tego przypadku prezentują się następująco:

$H_0$ : rozmiar gospodarki i wielkość emisji CO<sub>2</sub> są niezależne

$H_1$ : Rozmiar gospodarki i wielkość emisji CO<sub>2</sub> są zależne

Wyniki prezentują się następująco:

Contingency Tables

Wielkość emisji CO <sub>2</sub>		Wielkość gospodarki				Total
		mała	średnia	duża	bardzo duża	
mała	Observed	12	2	2	0	16
	% within column	54.5 %	4.3 %	9.1 %	0.0 %	17.4 %
średnia	Observed	10	30	0	0	40
	% within column	45.5 %	65.2 %	0.0 %	0.0 %	43.5 %
duża	Observed	0	14	17	0	31
	% within column	0.0 %	30.4 %	77.3 %	0.0 %	33.7 %
bardzo duża	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 %

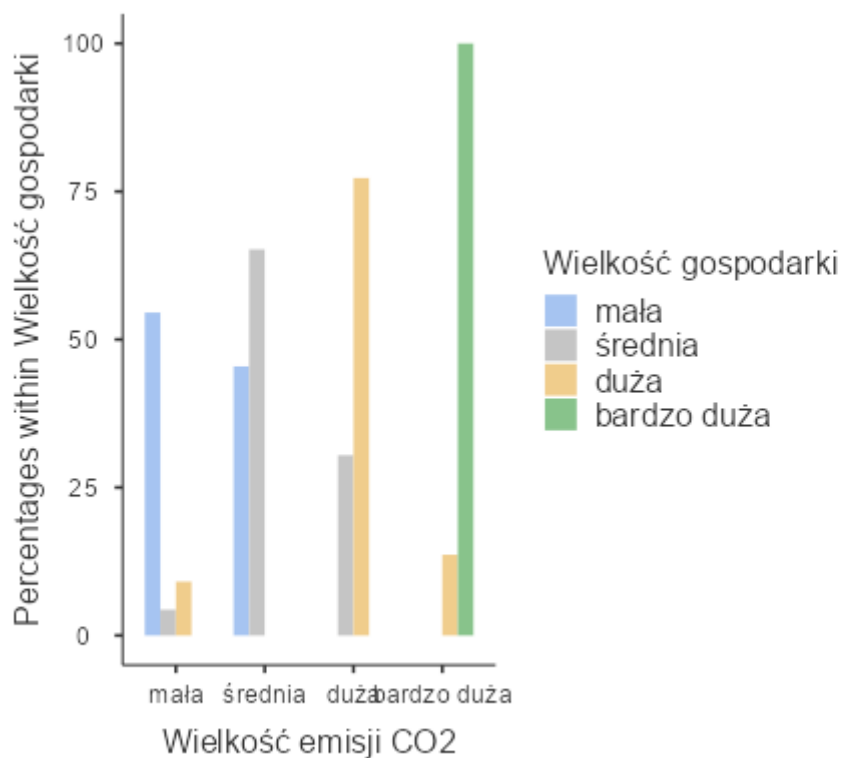
$\chi^2$  Tests

	Value	df	p
$\chi^2$	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

Na podstawie wyniku testu V-Cramera można zaobserwować umiarkowaną zależność pomiędzy wielkością gospodarki a wielkością emisji dwutlenku węgla. Ponieważ wynik jest większy niż 0,3 co oznacza umiarkowaną zależność. Dodatni współczynnik Tau-b Kendalla potwierdza tę zależność co oznacza, że gdy gospodarka rośnie zwiększa się emisja dwutlenku węgla.



Na powyższym wykresie można zaobserwować zależność wzrostu emisji CO<sub>2</sub> wraz ze wzrostem rozmiarów gospodarki. Nie są to dane jednoznaczne, ale wyniki w poprzednich tabelach to potwierdzają tę zależność.

W wyniku przeprowadzonej powyżej analizy statystycznej można odrzucić hipotezę  $H_0$  co oznacza, zmienne są zależne i należy przyjąć hipotezę  $H_1$ .

W następnym przypadku będzie testowana zależność pomiędzy emisją dwutlenku węgla a wielkością populacji. W tym przypadku hipotezy prezentują się następująco:

$H_0$ : wielkość populacji i emisja dwutlenku węgla są niezależne

$H_1$ : wielkość populacji i emisja dwutlenku węgla są zależne

Contingency Tables

Wielkość emisji CO2		Wielkość populacji (rozmiar)					Total
		mała	średnia	duża	bardzo duża	wielka	
mała	Observed	2	10	4	0	0	16
	% within column	100.0 %	29.4 %	9.1 %	0.0 %	0.0 %	17.4 %
średnia	Observed	0	23	16	1	0	40
	% within column	0.0 %	67.6 %	36.4 %	10.0 %	0.0 %	43.5 %
duża	Observed	0	1	24	6	0	31
	% within column	0.0 %	2.9 %	54.5 %	60.0 %	0.0 %	33.7 %
bardzo duża	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 %

$\chi^2$  Tests

	Value	df	p
$\chi^2$	90.9	12	< .001
N	92		

Nominal

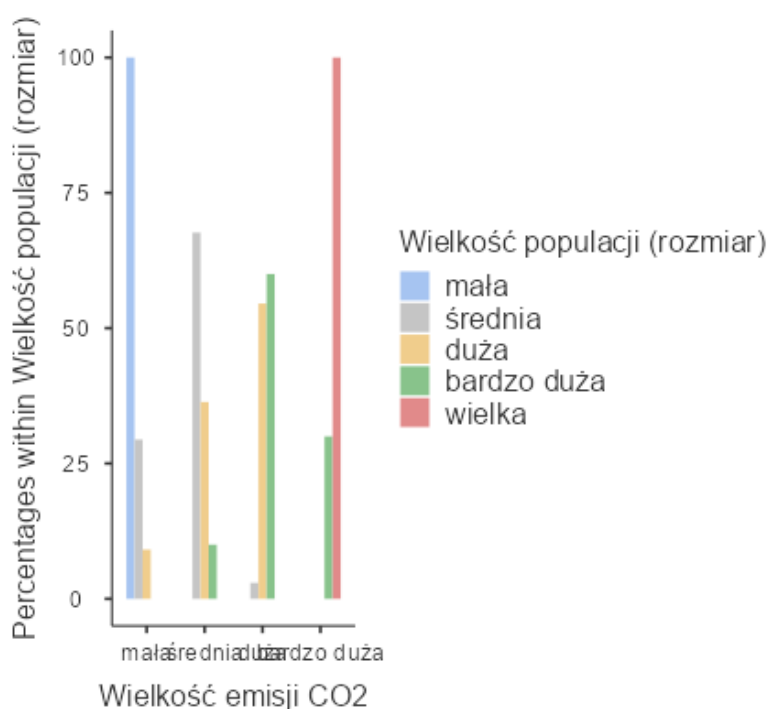
	Value
Phi-coefficient	NaN
Cramer's V	0.574

Kendall's Tau-b

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

Także w przypadku wielkości populacji wielkość wyniku testu V-Cramera świadczy o umiarkowanej zależności względem wzrostu emisji tlenku węgla. Także Tau-b Kendalla potwierdza tę zależność.

#### Plots



Powyższy wykres prezentuje procent danych populacji w wielkości emisji CO<sub>2</sub>, także w tym przypadku widoczna jest zależność. Ponieważ wraz ze wzrostem populacji rośnie zanieczyszczenie dwutlenkiem węgla.

Wyniki dowodzą że należy odrzucić hipotezę  $H_0$ , oznacza to prawdziwość hipotezy o zależności pomiędzy wielkością populacji a emisją dwutlenku węgla.

## T-testy

Testy t-Studenta służą do porównania ze sobą dwóch grup. Korzystamy z nich wtedy, gdy mamy wyniki dla dwóch grup i chcemy porównać je ze sobą - tzn. stwierdzić, czy wyniki w jednej grupie są większe bądź mniejsze niż w drugiej grupie. Nie można porównywać ze sobą kilku grup, wykonując kilkakrotnie test t-Studenta. W tym przypadku testowaniu zostanie poddana zmiana wielkości emisji dwutlenku węgla w roku 2018 względem 2017 roku. Dla tego testu hipotezy kształtują się następująco:

$H_0$ : wielkość emisji dwutlenku węgla w 2017 i 2018 roku były takie same

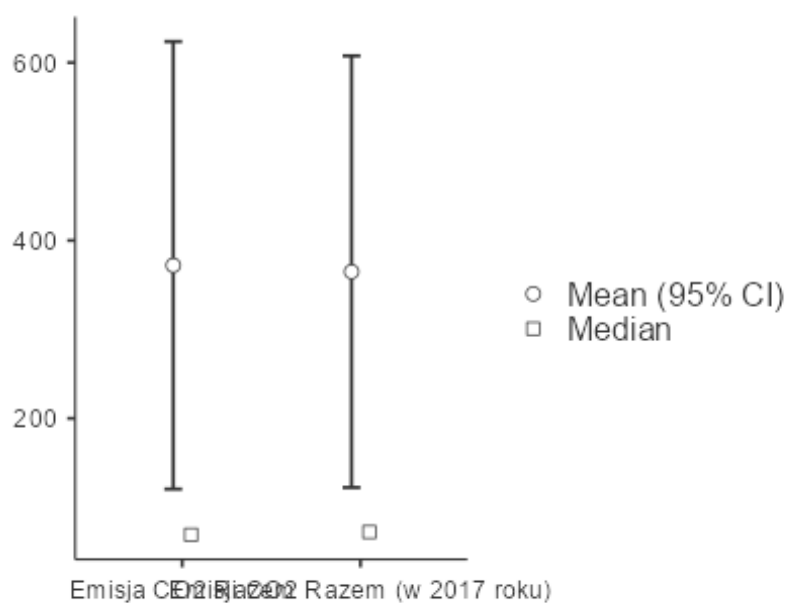
$H_1$ : wielkość emisji dwutlenku węgla w 2017 i 2018 roku były różne

Paired Samples T-Test

			statistic	df	p	Mean difference
Emisja CO2 Razem	Emisji CO2 Razem (w 2017 roku)	Student's t	1.47	91.0	0.145	7.19

Z testu T-Studenta wynika, że nastąpił wzrost emisji CO<sub>2</sub> w 2018 roku względem 2017 roku. Wysoka wartość p powyżej  $\alpha=0,05$  dowodzi dużej wiarygodności testu że wartości nie są równe. Natomiast przesunięcie średniej o 7,19 dowodzi o wzroście emisji dwutlenku węgla.

Emisja CO2 Razem - Emisji CO2 Razem (w 2017 roku)



Powyższy wykres również dowodzi wzrostu emisji CO<sub>2</sub> na przestrzeni dwóch lat. Wyraźnie widać przesunięcie średniej w górę oraz zwiększenie zakresu rozproszenia wśród wyższych wartości.

Wyniki dowodzą, że należy odrzucić hipotezę  $H_0$  oznacza to prawdziwość hipotezy o zmianie wielkości emisji dwutlenku węgla na przestrzeni 2017 i 2018 roku.



## Analiza wariancji

Jednoczynnikowa analiza wariancji jest to najprostsza forma analizy wariancji, gdzie zmienna odpowiedzi może zależeć tylko od jednego czynnika. Jednoczynnikowa ANOVA zawiera minimum dwa warunki eksperymentalne i pozwala na porównanie średnich z dwóch lub z większej ilości grup jednocześnie. W tym przypadku posłużymy się parametrycznym testem ANOVA w celu dokonania analizy zmian emisji dwutlenku węgla na przestrzeni pięciu lat w gospodarkach o różnej wielkości. Dla tego testu hipotezy prezentują się następująco:

$H_0$ : na przestrzeni lat wielkość emisji CO<sub>2</sub> dla różnych gospodarek nie zmieniała się

$$U_1=U_2=U_3=U_4=U_5$$

$H_1$ : na przestrzeni lat wielkość emisji CO<sub>2</sub> dla różnych gospodarek była różna

$$U_1 \neq U_2 \neq U_3 \neq U_4 \neq U_5$$

One-Way ANOVA (Welch's)

	F	df1	df2	p
Emisji CO2 Razem (w 2014 roku)	22.1	3	4.68	0.003
Emisji CO2 Razem (w 2015 roku)	23.0	3	4.68	0.003
Emisji CO2 Razem (w 2016 roku)	22.2	3	4.68	0.003
Emisji CO2 Razem (w 2017 roku)	22.1	3	4.68	0.003
Emisja CO2 Razem	21.0	3	4.68	0.004

Group Descriptives

	Wielkość gospodarki	N	Mean	SD	SE
Emisji CO2 Razem (w 2014 roku)	mała	22	11.6	7.69	1.64
	średnia	46	97.4	87.80	12.95
	duża	22	594.7	497.38	106.04
	bardzo duża	2	7754.2	3155.66	2231.39
Emisji CO2 Razem (w 2015 roku)	mała	22	11.5	7.50	1.60
	średnia	46	98.0	84.54	12.46
	duża	22	602.6	506.68	108.02
	bardzo duża	2	7610.1	3165.47	2238.32
Emisji CO2 Razem (w 2016 roku)	mała	22	11.8	7.82	1.67
	średnia	46	99.8	87.25	12.86
	duża	22	606.8	521.34	111.15
	bardzo duża	2	7484.2	3162.48	2236.21
Emisji CO2 Razem (w 2017 roku)	mała	22	12.0	7.94	1.69
	średnia	46	99.9	87.03	12.83
	duża	22	617.3	529.42	112.87
	bardzo duża	2	7564.1	3332.39	2356.35
Emisja CO2 Razem	mała	22	12.0	7.40	1.58
	średnia	46	100.3	88.91	13.11
	duża	22	622.2	559.80	119.35
	bardzo duża	2	7832.7	3475.07	2457.25

Wyniki testu Welcha wskazują na nieprawdziwość tezy która stwierdza, że wielkość emisji dwutlenku węgla na przestrzeni lat w gospodarkach o różnej wielkości były takie same. Ponieważ wartość p jest mniejsza od poziomu istotności  $\alpha=0,05$ . Potwierdzenie tego wyniku można zaobserwować na przykładzie średniej która rosła na przestrzeni lat. A więc założenie o równości emisji CO<sub>2</sub> na przestrzeni lat jest błędne. Można także założyć, że wraz ze wzrostem gospodarki rośnie ilość emitowanego dwutlenku węgla.

Powyższe wyniki dowodzą, że hipoteza  $H_0$  jest błędna co oznacza że wielkość emisji dwutlenku węgla na przestrzeni lat zmieniała się.

W tym przypadku posłużymy się parametrycznym testem ANOVA w celu dokonania analizy zmian emisji dwutlenku węgla na przestrzeni pięciu lat w populacjach o różnej wielkości. Dla tego testu hipotezy prezentują się następująco:

$H_0$ : na przestrzeni lat wielkość emisji CO<sub>2</sub> dla różnych populacji nie zmieniała się

$$U_1=U_2=U_3=U_4=U_5$$

$H_1$ : na przestrzeni lat wielkość emisji CO<sub>2</sub> dla różnych populacji była różna

$$U_1 \neq U_2 \neq U_3 \neq U_4 \neq U_5$$

One-Way ANOVA (Welch's)

	F	df1	df2	p
Emisji CO <sub>2</sub> Razem (w 2014 roku)	13.2	4	6.68	0.003
Emisji CO <sub>2</sub> Razem (w 2015 roku)	13.4	4	6.77	0.002
Emisji CO <sub>2</sub> Razem (w 2016 roku)	13.8	4	6.78	0.002
Emisji CO <sub>2</sub> Razem (w 2017 roku)	14.3	4	6.71	0.002
Emisja CO <sub>2</sub> Razem	14.2	4	6.65	0.002

	Wielkość populacji (liczebność)	N	Mean	SD	SE
Emisji CO2 Razem (w 2014 roku)	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
Emisji CO2 Razem (w 2015 roku)	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
Emisji CO2 Razem (w 2016 roku)	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
Emisji CO2 Razem (w 2017 roku)	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
Emisja CO2 Razem	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

Wyniki testu Welcha wskazują na nieprawdziwość tezy która stwierdza, że wielkość emisji dwutlenku węgla na przestrzeni lat w populacjach o różnej wielkości były takie same. Ponieważ wartość p jest mniejsza od poziomu istotności  $\alpha=0,05$ . Potwierdzenie tego wyniku można zaobserwować na przykładzie odchylenia standardowego, które rosło na przestrzeni lat (z wyłączeniem małej populacji). A więc założenie o równości emisji CO2 na przestrzeni lat jest błędne. Można także założyć, że wraz ze wzrostem populacji rośnie ilość emitowanego dwutlenku węgla.

Powyższe wyniki dowodzą, że hipoteza  $H_0$  jest błędna co oznacza że wielkość emisji dwutlenku węgla na przestrzeni lat zmieniała się.

Test Kruskala-Wallisa jest rozszerzeniem testu U Manna-Whitneya. Test ten to nieparametryczny odpowiednik jednoczynnikowej analizy wariancji. Wykrywa różnice w położeniu rozkładu. Ten test zakłada brak uporządkowania a priori k populacji, z których pobrano próby. W celu przeprowadzenia testu czy gospodarki produkują gazy cieplarniane w różnej wielkości posłużymy się nieparametryczną analizą wariancji. Dla testu należy zastosować poziom istotności  $\alpha=0,05$ . Hipotezy są następujące:

$H_0$ : wielkość emisji gazów cieplarnianych dla różnych gospodarek jest taka sama

$H_1$ : wielkość emisji gazów cieplarnianych dla różnych gospodarek jest różna

Kruskal-Wallis

	$\chi^2$	df	p
Emisja CH4	52.2	3	< .001
Emisja N2O	56.7	3	< .001
Emisja CO2 Razem	73.4	3	< .001

Niski wynik test Kruskala-Wallisa dowodzi, że gospodarki o różnej wielkości produkują gazy cieplarniane w różnej ilości. Ponieważ p-wartość jest zdecydowanie niższa od poziomu istotności  $\alpha=0,05$ , można jednoznacznie odrzucić hipotezę o równości.

Wyniki testu dowodzą, że należy odrzucić hipotezę  $H_0$  co oznacza, że różne gospodarki produkują gazy cieplarniane w różnej wielkości.

W celu przeprowadzenia testu czy populacje produkują gazy cieplarniane w różnej wielkości posłużymy się nieparametryczną analizą wariancji. Dla testu należy zastosować poziom istotności  $\alpha=0,05$ . Hipotezy są następujące:

$H_0$ : wielkość emisji gazów cieplarnianych dla różnych populacji jest taka sama

$H_1$ : wielkość emisji gazów cieplarnianych dla różnych populacji jest różna

Kruskal-Wallis			
	$\chi^2$	df	p
Emisja CH <sub>4</sub>	56.3	4	< .001
Emisja N <sub>2</sub> O	59.2	4	< .001
Emisja CO <sub>2</sub> Cement	47.4	4	< .001

Niski wynik test Kruskala-Wallisa dowodzi, że populacje o różnej wielkości produkują gazy cieplarniane w różnej ilości. Ponieważ p-wartość jest zdecydowanie niższa od poziomu istotności  $\alpha=0,05$ , można jednoznacznie odrzucić hipotezę o równości.

Wyniki testu dowodzą, że należy odrzucić hipotezę  $H_0$  co oznacza, że różne gospodarki produkują gazy cieplarniane w różnej wielkości.

## Analiza regresji wielorakiej

Ogólnym celem regresji wielorakiej (termin ten został po raz pierwszy użyty przez Pearsona w 1908 roku) jest ilościowe ujęcie związków pomiędzy wieloma zmiennymi niezależnymi (objaśniającymi) a zmienną zależną (kryterialną, objaśnianą). Na przykład pośrednik w handlu nieruchomościami (agent) zbiera dane dotyczące budynków - wielkość (w m<sup>2</sup>), liczba sypialni, średni dochód mieszkańców dzielnicy oraz subiektywna ocena atrakcyjności obiektu. Jeśli dysponuje już jakąś bazą danych tego typu, to może pokusić się o odpowiedź na następujące pytanie: Jak poszczególne wielkości wpływają na cenę budynku? Można w ten sposób na przykład dowiedzieć się, że liczba sypialni w lepszy sposób objaśnia cenę budynku niż na przykład to, jak ładny wydaje się on na podstawie oceny na oko (subiektywna atrakcyjność). Można też odkryć obiekty "odstające", to znaczy budynki, które mają większą wartość niż wynika to z danych zebranych przez agenta.

Przy pomocy regresji wielorakiej zostanie ustalone jaki wzrost PKB powoduje zwiększenie się emisji gazów cieplarnianych o 1 Mt oraz jak wpływa wzrost populacji na emisję CO<sub>2</sub>.

Model Fit Measures

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	AIC	BIC	RMSE	Overall Model Test			
							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
PKB	3.22e+7	1	3.22e+7	197.5	< .001
Populacja kraju	2.51e+6	1	2.51e+6	15.4	< .001
Residuals	1.45e+7	89	162816		

Note. Type 3 sum of squares

Model Coefficients - Emisja CO<sub>2</sub> Razem

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

### Assumption Checks

Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.0615	1.88	0.434

[3]

Collinearity Statistics

	VIF	Tolerance
PKB	2.54	0.394
Populacja kraju	2.54	0.394

[3]

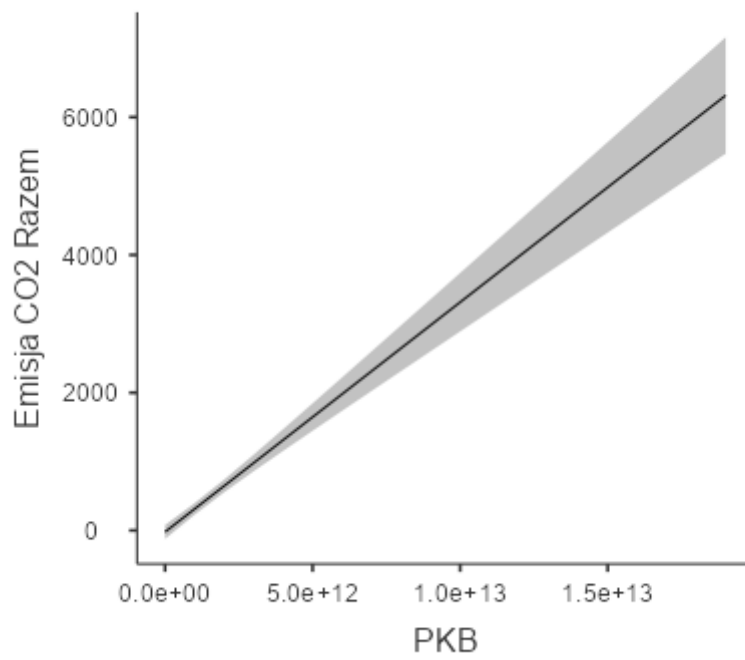
Normality Test (Shapiro-Wilk)

Statistic	p
0.564	< .001

Wskaźnik korelacji wykazuje niską korelację pomiędzy tymi trzema zmiennymi, ale poziom jego istotności jest bardzo wysoki. Niska korelacja oznacza brak naruszenia zasad modelu. Współczynnik VIF wskazuje na bardzo silną liniowość modelu. Natomiast test normalności nie spełnia założeń rozkładu normalnego. Współczynnik R-kwadrat wskazuje na wysokie dopasowanie modelu.

Wyniki uzyskane poprzez modelowanie wskazują, że wraz ze wzrostem Produktu Krajowego Brutto o 33,4 miliarda dolarów rośnie emisja dwutlenku węgla o 1 megatonę. Również w przypadku wzrostu populacji o 12,8 miliona rośnie produkcja CO<sub>2</sub> o 1 megatonę. Przyglądając się odchyleniu standardowemu możemy zauważyć, że wzrost PKB bardziej przyczynia się do zwiększenia emisji dwutlenku węgla niż wzrost populacji. Wnioskując można przyjąć założenie, że poprawa standardu życia przyczynia do wzrostu emisji gazów cieplarnianych.

## PKB



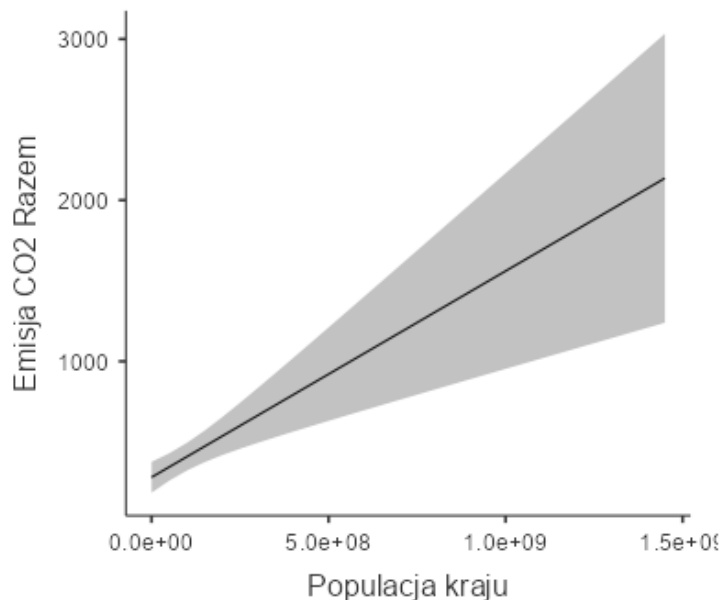
Estimated Marginal Means - PKB

PKB	Marginal Mean	SE	95% Confidence Interval	
			Lower	Upper
-1.66e-12 <sup>-</sup>	-575	79.4	-733	-417
1.18e+12 <sup>μ</sup>	372	42.1	289	456
4.02e+12 <sup>+</sup>	1319	79.4	1161	1477

Note. <sup>-</sup> mean - 1SD, <sup>μ</sup> mean, <sup>+</sup> mean + 1SD



#### Populacja kraju



Estimated Marginal Means - Populacja kraju

Populacja kraju	Marginal Mean	SE	95% Confidence Interval	
			Lower	Upper
-1.35e-8 <sup>~</sup>	107	79.4	-50.3	265
7.15e+7 <sup>μ</sup>	372	42.1	288.5	456
2.78e+8 <sup>+</sup>	637	79.4	478.9	795

Note. <sup>~</sup> mean - 1SD, <sup>μ</sup> mean, <sup>+</sup> mean + 1SD

Wnioski wyciągnięte z tabel również są zauważalne na wykresach. W przypadku PKB widać zależność wzrostu emisji dwutlenku węgla oraz wzrost odchylenia standardowego wraz ze wzrostem emisji. Natomiast emisja CO<sub>2</sub> do populacji kraju wykazuje niższy wzrost przy zwiększaniu się populacji. Porównanie obu wykresów potwierdza tezę, że wzrost PKB silniej wpływa na wzrost emisji dwutlenku węgla niż wzrost populacji.

Analiza regresji logistycznej (dane odnawialne źródła energii nie są w pełni wiarygodne)

Model regresji logistycznej jest szczególnym przypadkiem uogólnionego modelu liniowego. Znajduje zastosowanie, gdy zmienna zależna jest dychotomiczna, to znaczy przyjmuje tylko dwie wartości takie jak na przykład sukces lub porażka, wystąpienie lub brak pewnej jednostki chorobowej, kobieta lub mężczyzna. W zapisie matematycznym wartości te reprezentowane są jako 1 i 0. Regresja logistyczna nie ma wielu z kluczowych założeń regresji liniowej i ogólnych modeli liniowych, które bazują na algorytmach metody najmniejszych kwadratów – zwłaszcza jeżeli chodzi o liniowość, normalność, homoskedastyczność i poziom pomiaru. Regresja logistyczna wymaga aby zmienna zależna była podwójna, a porządkowa regresja logistyczna wymaga aby zmienna zależna była porządkowa. Zredukowanie porządkowej albo metrycznej zmiennej do dychotomicznego poziomu powoduje stratę dużej ilości informacji, które sprawiają, że taki test jest gorszy w porównaniu z porządkową regresją logistyczną w takim przypadku. W celu zbadania wpływu odnawialnych źródeł energii na emisje dwutlenku węgla zostanie zastosowana regresja logistyczna do trzech paliw kopalnych.

Model Fit Measures

Model	Deviance	AIC	R <sup>2</sup> <sub>McF</sub>
1	123	131	0.0360

Model Coefficients - Odnawialne źródła energii

Predictor	Estimate	SE	Z	p	Odds ratio	95% Confidence Interval	
						Lower	Upper
Intercept	0.16407	0.23108	0.710	0.478	1.178	0.749	1.85
Emisja CO2 Węgiel	-0.00213	0.00175	-1.215	0.224	0.998	0.994	1.00
Emisja CO2 Gaz	-0.00566	0.00458	-1.236	0.216	0.994	0.985	1.00
Emisja CO2 Ropa	0.00576	0.00403	1.429	0.153	1.006	0.998	1.01

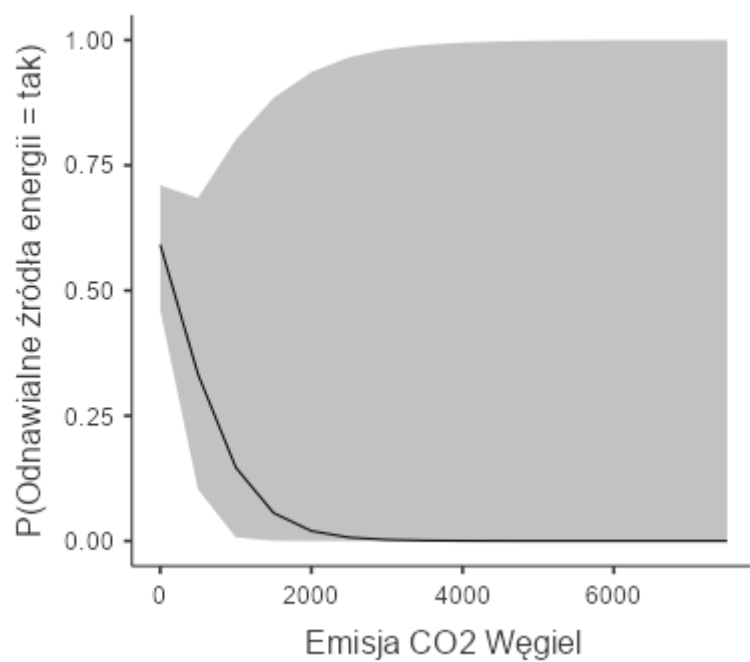
Note. Estimates represent the log odds of "Odnawialne źródła energii = tak" vs. "Odnawialne źródła energii = nie"

### Assumption Checks

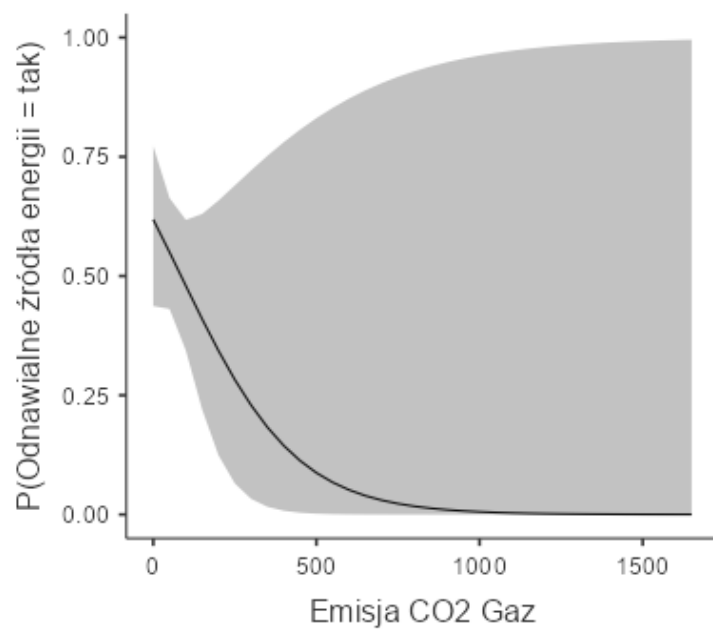
Collinearity Statistics

	VIF	Tolerance
Emisja CO2 Węgiel	3.18	0.3144
Emisja CO2 Gaz	10.46	0.0956
Emisja CO2 Ropa	16.12	0.0620

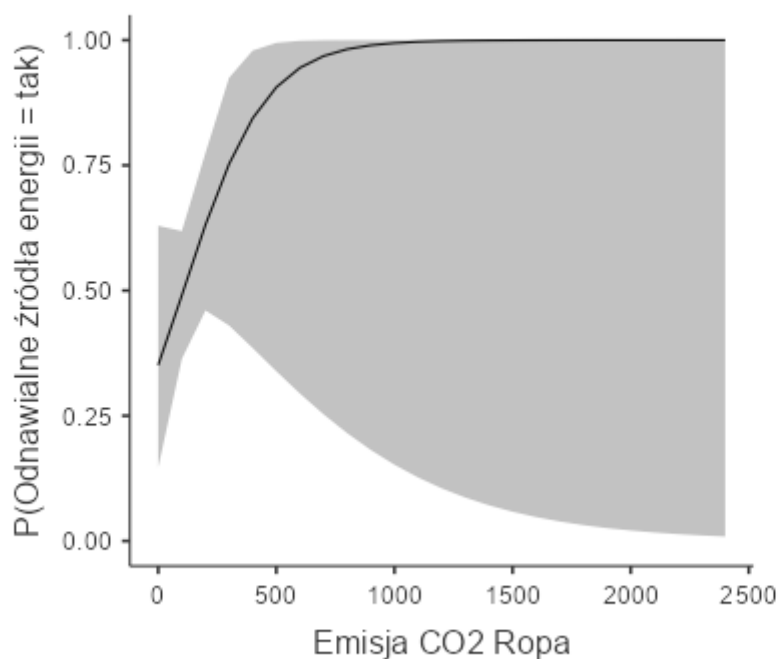
### Emisja CO2 Węgiel



### Emisja CO2 Gaz



### Emisja CO2 Ropa



Wskaźnik R-kwadrat wskazuje na niskie dopasowanie modelu. Wartość  $p$  we wszystkich przypadkach jest wyższa od poziomu istotności  $\alpha=0,05$ . Wartość estymacji wskazuje, że ilość emisji gazów cieplarnianych z węgla i gazu będzie spadać w momencie posiadania przez państwo OZE. W przypadku ropy oraz posiadania odnawialnych źródeł energii wartość emisji dwutlenku węgla będzie wzrastać. Wskaźnik VIF wskazuje na silną liniowość modelu.

Wykresy zamieszczone potwierdzają wyniki z tabeli. W przypadku korzystania przez dany kraj z odnawialnych źródeł energii prawdopodobieństwo emisji CO2 z węgla będzie niskie. Także gdy kraj korzysta z odnawialnych źródeł energii prawdopodobieństwo emisji dwutlenku węgla z gazu spada. Odwrotna sytuacja pojawia się w przypadku korzystania z odnawialnych źródeł energii oraz emisji dwutlenku węgla z ropy. Ponieważ zwiększa się prawdopodobieństwo emisji dwutlenku węgla gdy dany kraj korzysta z odnawialnych źródeł energii.

## Podsumowanie

Na podstawie powyższego badania można stwierdzić, że istnieje korelacja pomiędzy wzrostem Produktu Krajowego Brutto a emisją gazów cieplarnianych. Korelacja jest również widoczna w przypadku populacji oraz emisji gazów cieplarnianych. Obie te korelacje również posiadają zależność, co oznacza że gdy rośnie PKB i liczba ludności zwiększa to emisję gazów cieplarnianych. Także test T-Studenta dowiódł, że na przestrzeni lat 2017 i 2018 produkcja CO<sub>2</sub> wzrosła. Natomiast analiza parametryczna wariancji dowodzi, że wzrost emisji gazów cieplarnianych w roku 2018 nie był wyjątkiem wzrost ten jest tendencją stale rosnącą. Następnie wynik analizy parametrycznej wskazuje, że rozmiar gospodarki i populacji mają wpływ na wielkość emisji gazów cieplarnianych. Natomiast analiza regresji liniowej dostarczyła bardziej szczegółowych danych. Według których wraz ze wzrostem Produktu Krajowego Brutto o 33,4 miliarda dolarów rośnie emisja dwutlenku węgla o 1 megatonę. Również w przypadku wzrostu populacji o 12,8 miliona rośnie produkcja CO<sub>2</sub> o 1 megatonę. Regresja liniowa wykazała również, że do wzrostu emisji dwutlenku węgla bardziej przyczynia się wzrost PKB niż wzrost populacji. Można z tego wywnioskować, że do wzrostu emisji CO<sub>2</sub> przyczynia się wzrost poziomu życia ludzi. Natomiast niespodziewanego wyniku dostarczyła analiza logistyczna która potwierdziła, że przy odchodzeniu od gazu i węgla na rzecz OZE zmniejsza się ilość emisji CO<sub>2</sub>, ale w przypadku ropy odejście może spowodować wzrost emisji dwutlenku węgla.

Podsumowując można stwierdzić, że rozmiar gospodarki ma istotny wpływ na emisję dwutlenku węgla. Natomiast rozmiar populacji nie odgrywa tak kluczowej roli jak poziom na którym ona żyje. Większość emisji jest skoncentrowana w bardzo dużych populacjach i gospodarkach, nie należy także pomijać liczebności tych małych i średnich. Należało by także przyjrzeć się odchodzeniu od paliw kopalnych na rzecz odnawialnych źródeł energii aby nie wywołało to sprzężenia zwrotnego.