

Opis próby

Niniejszą analizę wykonano w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023). Wizualizację wyników przeprowadzono z wykorzystaniem pakietu graficznego "ggplot2" (Wickham, 2016). Niniejsza analiza ma na celu opis zmiennych i wystąpień, a także wyliczenie jaki stanowiły one odsetek w analizowanej próbie. W celu weryfikacji czy te wystąpienia są tendencyjne lub przypadkowe wykonano test Chi Kwadrat dla proporcji (Pearson, 1900). W celu wyliczenia częstości i istotności występowania poszczególnych wartości zmiennych Płeć, Edukacja, Wiek wykonano serię 3 analiz częstości testem Chi Kwadrat.

Analiza zmiennej *Płeć* wykazała, że liczebność jej poszczególnych poziomów kształtowała się następująco: Kobieta: $n = 173$ (57.10%); Mężczyzna: $n = 130$ (42.90%), co łącznie dało $N = 303$ wystąpień (100%). Zaobserwowane różnice w liczebnościach kategorii były istotne statystycznie $\chi^2(1) = 6.10$; $p = 0.014$. Zróżnicowanie analizowanych wartości było słabe, miara siły efektu ϕ (Phi) wyniosła wartość $\phi = 0.14$. W rozkładzie zdecydowanie przeważała kategoria „Kobieta” (57.10%), najrzadziej zaś występowała kategoria „Mężczyzna” (42.90%).

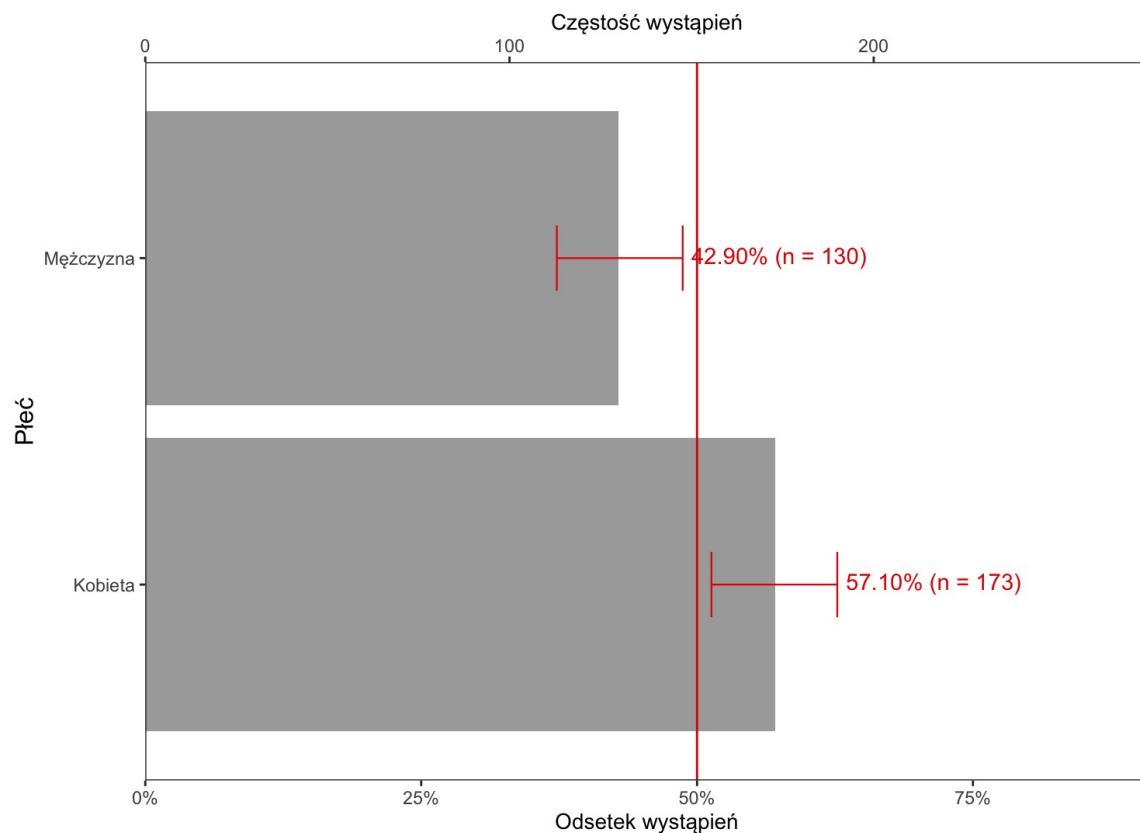
W przypadku zmiennej *Edukacja* stwierdzono, że liczebność jej poszczególnych kategorii kształtowała się następująco: Wyższe: $n = 171$ (56.44%); Średnie: $n = 130$ (42.90%); Podstawowe: $n = 2$ (0.66%), co łącznie dało $N = 303$ wystąpień (100%). Różnice te okazały się istotne statystycznie $\chi^2(2) = 153.88$; $p < 0.001$. Zróżnicowanie analizowanych wartości było silne, miara siły efektu V Craméra wyniosła wartość $V = 0.50$. W rozkładzie zdecydowanie przeważała kategoria „Wyższe” (56.44%), najrzadziej zaś występowała kategoria „Podstawowe” (0.66%). Powyżej wartości oczekiwanej (wynoszącej $n = 101.00$) znalazły się kategorie: Średnie, Wyższe, natomiast poniżej niej znalazła się kategoria: Podstawowe.

Rozkład zmiennej *Wiek* pokazał, że liczebność jej poszczególnych wariantów kształtowała się następująco: Od 20 do 29 lat: $n = 225$ (74.26%); Od 30 do 39 lat: $n = 59$ (19.47%); Od 40 do 51 lat: $n = 19$ (6.27%), co łącznie dało $N = 303$ wystąpień (100%). Analiza testem chi-kwadrat wykazała, że różnice między kategoriami były istotne statystycznie $\chi^2(2) = 236.28$; $p < 0.001$.

Zróźnicowanie analizowanych wartości było silne, miara siły efektu V Craméra wyniosła wartość $V = 0.62$. W rozkładzie zdecydowanie przeważała kategoria „Od 20 do 29 lat” (74.26%), najrzadziej zaś występowała kategoria „Od 40 do 51 lat” (6.27%). Powyżej wartości oczekiwanej (wynoszącej $n = 101.00$) znalazła się kategoria: Od 20 do 29 lat, natomiast poniżej niej znalazły się kategorie: Od 40 do 51 lat, Od 30 do 39 lat. Wyniki analiz przedstawia seria rysunków od nr 1 do nr 3.

Wykres nr 1

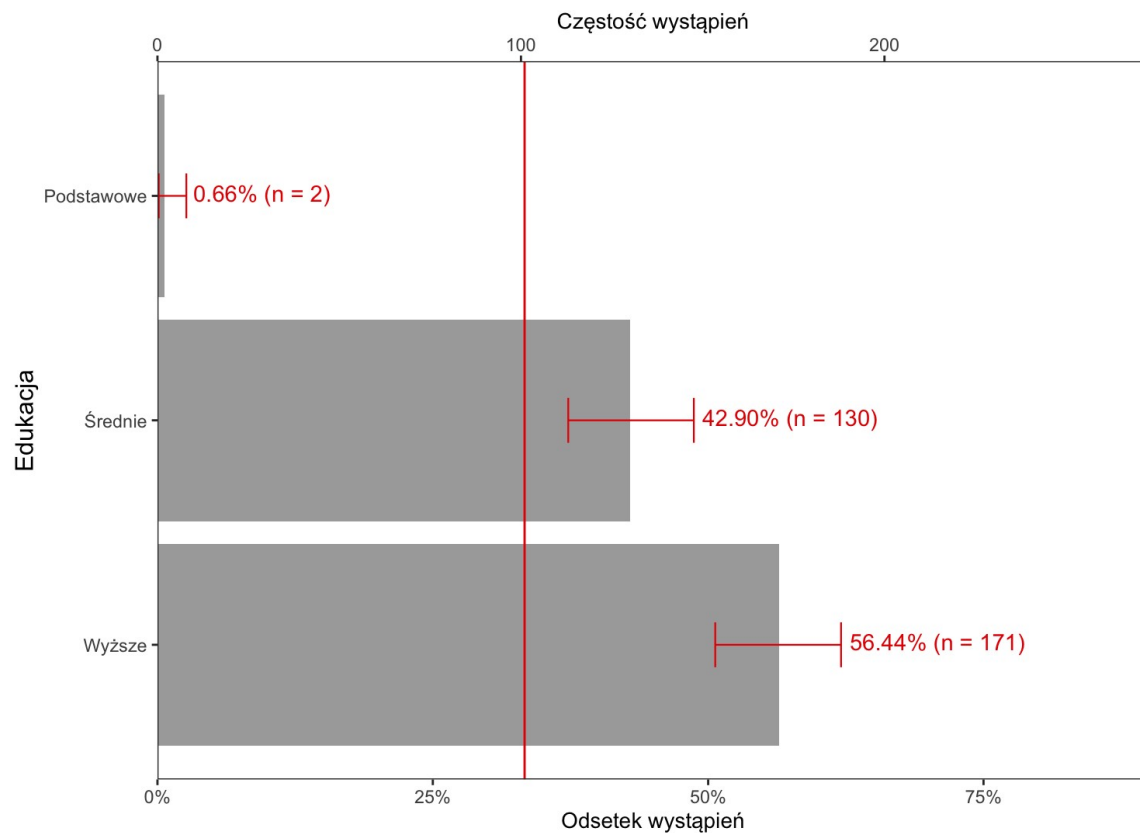
Częstość i odsetek występowania wartości zmiennej Płeć



Nota. Zróźnicowanie poziomów zmiennej Płeć było istotnie dalekie od wartości oczekiwanej wynoszącej 50.00% (czerwona linia). Wynik oszacowania testu Chi Kwadrat wyniósł: $\chi^2(1) = 6.10$; $p = 0.014$. Ciągła linia — oznacza istotne zróźnicowanie wartości zmiennej Płeć. Czerwone wąsy przy słupkach przedstawiają 95% przedział ufności dla odsetka wystąpień danej kategorii. Nakładanie się wąsów dwóch kategorii oznacza brak istotnych różnic między nimi, natomiast ich rozłączość świadczy o istotnych statystycznie różnicach.

Wykres nr 2

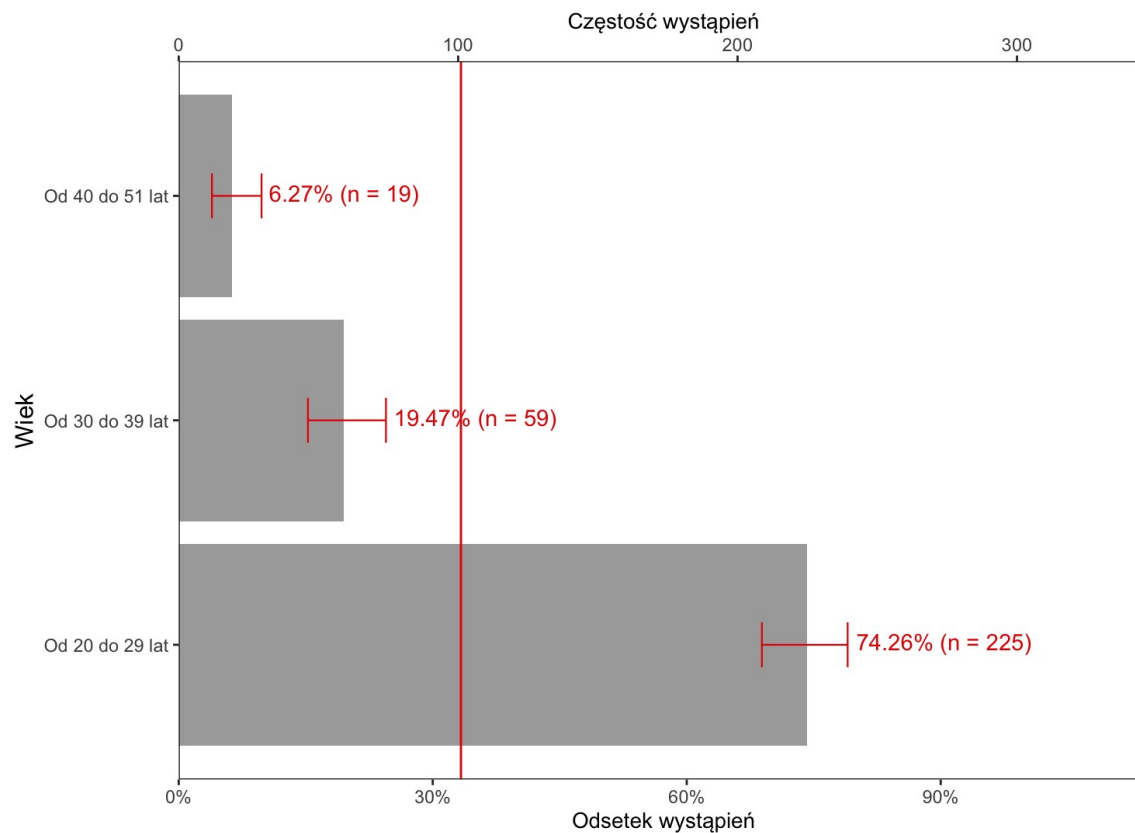
Częstość i odsetek występowania wartości zmiennej Edukacja



Nota. Zróżnicowanie poziomów zmiennej Edukacja było istotnie dalekie od wartości oczekiwanej wynoszącej 33.33% (czerwona linia). Wynik oszacowania testu Chi Kwadrat wyniósł: $\chi^2(2) = 153.88$; $p < 0.001$. Ciągła linia — oznacza istotne zróżnicowanie wartości zmiennej Edukacja. Czerwone wąsy przy słupkach przedstawiają 95% przedział ufności dla odsetka wystąpień danej kategorii. Nakładanie się wąsów dwóch kategorii oznacza brak istotnych różnic między nimi, natomiast ich rozłączość świadczy o istotnych statystycznie różnicach.

Wykres nr 3

Częstość i odsetek występowania wartości zmiennej Wiek



Nota. Zróżnicowanie poziomów zmiennej Wiek było istotnie dalekie od wartości oczekiwanej wynoszącej 33.33% (czerwona linia). Wynik oszacowania testu Chi Kwadrat wyniósł: $\chi^2(2) = 236.28$; $p < 0.001$. Ciągła linia — oznacza istotne zróżnicowanie wartości zmiennej Wiek. Czerwone wąsy przy słupkach przedstawiają 95% przedział ufności dla odsetka wystąpień danej kategorii. Nakładanie się wąsów dwóch kategorii oznacza brak istotnych różnic między nimi, natomiast ich rozłączość świadczy o istotnych statystycznie różnicach.

Podsumowanie wyników

Przeprowadzono serię trzech analiz częstości przy użyciu testu Chi Kwadrat dla proporcji, celem opisanego rozkładów zmiennych Płeć, Edukacja i Wiek oraz oceny, czy ich wystąpienia odbiegały od rozkładu losowego. Badana próba była nieco zdominowana przez kobiety, najczęściej wykształcenie uczestników było wyższe, a ich wiek w zdecydowanej większości był w przedziale od 20 do 29 lat

Analiza rzetelności pomiarów

Niniejszą analizę wykonano w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023). W celu weryfikacji dokładności pomiarowej testowanych zmiennych przeprowadzono analizę rzetelności z wykorzystaniem współczynnika Alfa Cronbacha (Cronbach, 1951) oraz współczynnika Omega McDonalda (McDonald, 1999). Uzupełniając, na podstawie jednoczynnikowej analizy składowych głównych (PCA), uwzględniono rzetelność kompozytową (CR; Jöreskog, 1971; Fornell & Larcker, 1981), rzetelność rhoA Dijkstry-Henselera (ρ_A ; Dijkstra & Henseler, 2015) oraz przeciętną wyodrębnioną wariancję (AVE; Fornell & Larcker, 1981). Rzetelność analizowanych skal oceniono łącznie, porównując poziom spójności wewnętrznej poszczególnych pomiarów. Wszystkie analizowane pomiary miały wysoką dokładność pomiarową. Wartości AVE dla skal Intencja zakupu, Użyteczność, Łatwość oraz Cele wskazują na wysoki poziom trafności zbieżnej. Akceptowalny poziom AVE uzyskano dla skali Obawy. Pełne wartości współczynników znajdują się w tabeli nr 1. W skrócie: Intencja zakupu ($\alpha = 0.87$, $\omega = 0.87$, CR = 0.91, $\rho_A = 0.87$, AVE = 0.71); Użyteczność ($\alpha = 0.87$, $\omega = 0.87$, CR = 0.91, $\rho_A = 0.87$, AVE = 0.73); Łatwość ($\alpha = 0.90$, $\omega = 0.90$, CR = 0.93, $\rho_A = 0.90$, AVE = 0.76); Cele ($\alpha = 0.91$, $\omega = 0.91$, CR = 0.93, $\rho_A = 0.91$, AVE = 0.73); Obawy ($\alpha = 0.86$, $\omega = 0.86$, CR = 0.90, $\rho_A = 0.86$, AVE = 0.64). Analizę rzetelności wykonano na $N = 303$ obserwacjach. Rezultaty tej analizy przedstawia tabela nr 1 oraz rysunki od 1 do 5.

Tabela nr 1

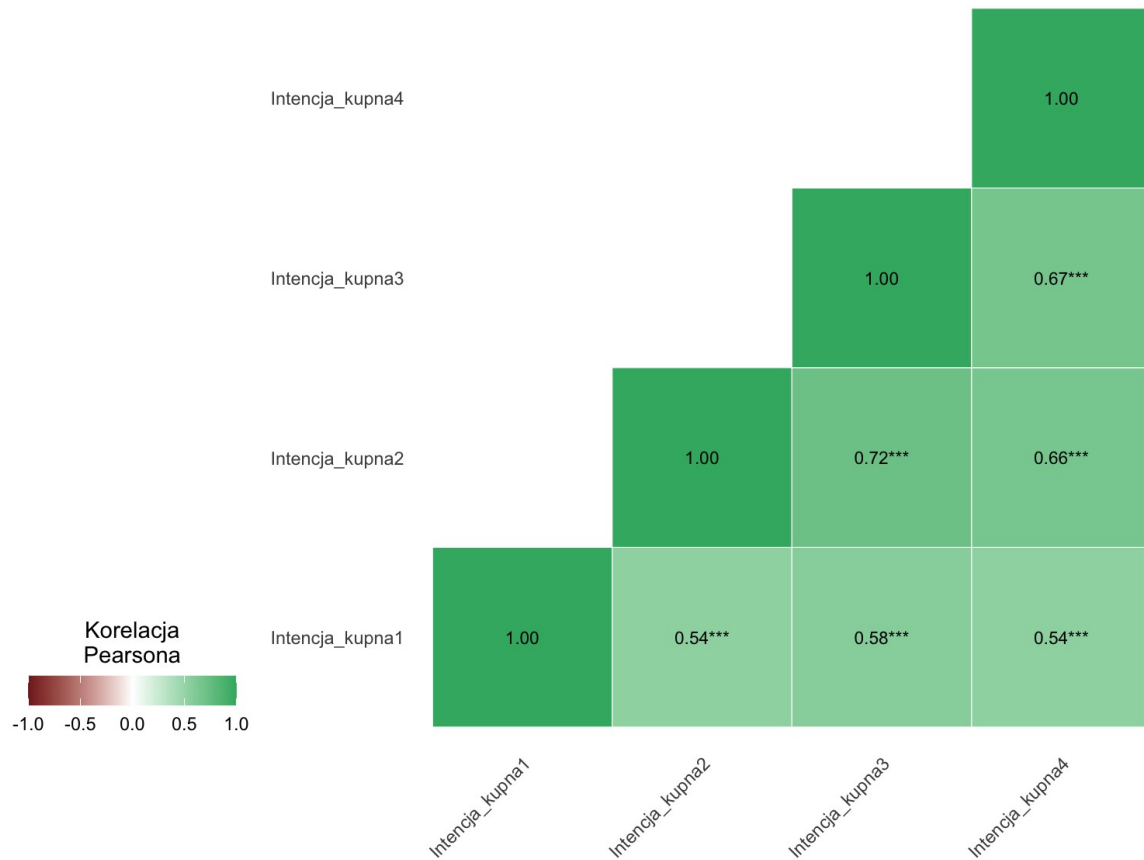
Wyniki analizy rzetelności dla skal Intencja zakupu, Użyteczność, Łatwość, Cele, Obawy

Skala	α	ω	CR	ρ_A	AVE	M	SD	N
Intencja zakupu	0.87	0.87	0.91	0.87	0.71	2.76	1.03	4
Użyteczność	0.87	0.87	0.91	0.87	0.73	3.24	0.89	4
Łatwość	0.90	0.90	0.93	0.90	0.76	3.61	0.79	4
Cele	0.91	0.91	0.93	0.91	0.73	2.17	0.74	5
Obawy	0.86	0.86	0.90	0.86	0.64	2.79	0.71	5

Nota. α = Alfa Cronbacha; ω = Omega McDonalda; CR = rzetelność kompozytowa; ρ_A = rzetelność rhoA Dijkstry-Henselera; AVE = przeciętna wyodrębniona wariancja. Współczynniki rzetelności α , ω , CR i ρ_A powyżej wartości 0.70 wskazują na akceptowalny poziom dokładności pomiarowej analizowanej skali (im bliżej 1, tym większa dokładność pomiarowa). AVE powyżej wartości 0.50 wskazuje na akceptowalny poziom trafności zbieżnej; N = Liczba pozycji testowych wchodzących w budowę skali; M = Średni poziom wyników pomiaru składowymi; SD = Odchylenie standardowe wyników pomiaru składowymi.

Wykres nr 4

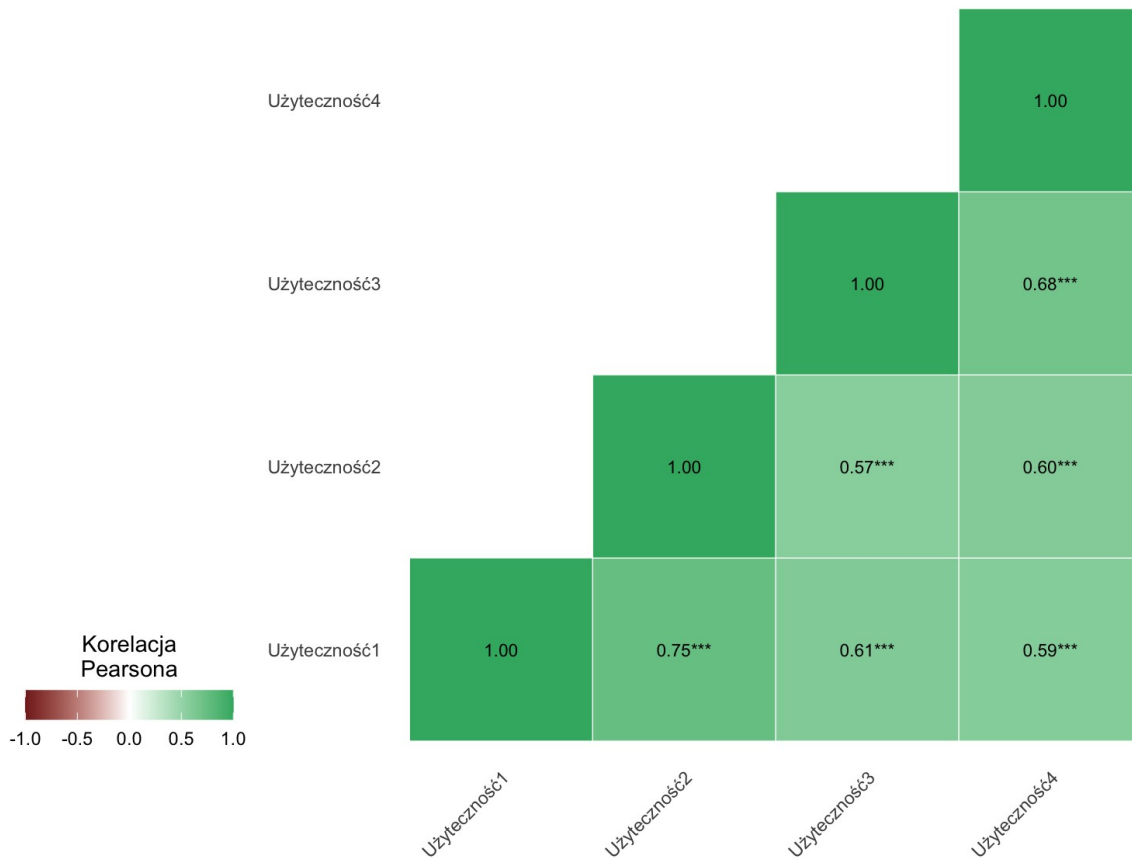
Korelacje pomiędzy składnikami budującymi pomiar Intencja zakupu



Nota. Rysunek przedstawia oszacowania współczynników korelacji Pearsona. Im bliżej wartości 1 (im większe nasycenie zielonego koloru), tym silniejsze powiązanie między składnikami testowymi. Diagonalne wartości 1 wskazują na korelację składnika z samym sobą. Gwiazdki oznaczają istotność statystyczną: *** $p < .001$.

Wykres nr 5

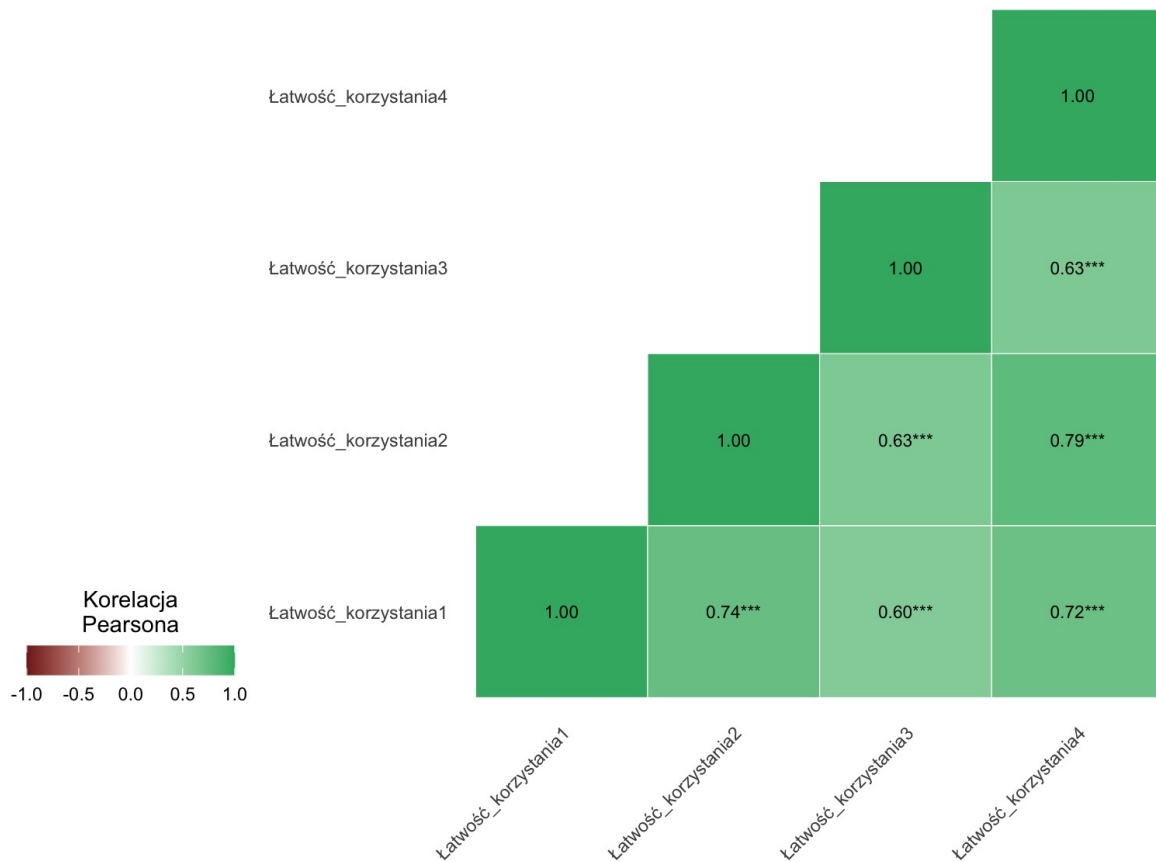
Korelacje pomiędzy składnikami budującymi pomiar Użyteczność



Nota. Rysunek przedstawia oszacowania współczynników korelacji Pearsona. Im bliżej wartości 1 (im większe nasycenie zielonego koloru), tym silniejsze powiązanie między składnikami testowymi. Diagonalne wartości 1 wskazują na korelację składnika z samym sobą. Gwiazdki oznaczają istotność statystyczną: *** $p < .001$.

Wykres nr 6

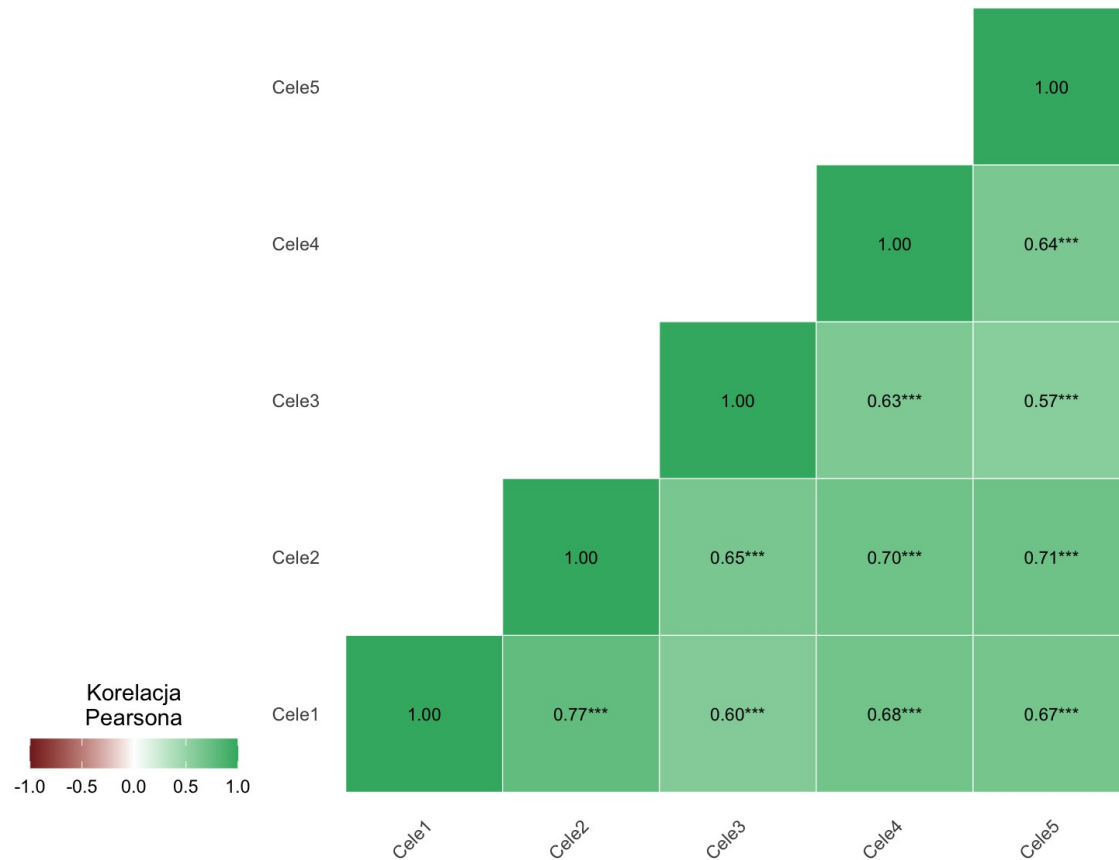
Korelacje pomiędzy składnikami budującymi pomiar Łatwość



Nota. Rysunek przedstawia oszacowania współczynników korelacji Pearsona. Im bliżej wartości 1 (im większe nasycenie zielonego koloru), tym silniejsze powiązanie między składnikami testowymi. Diagonalne wartości 1 wskazują na korelację składnika z samym sobą. Gwiazdki oznaczają istotność statystyczną: *** $p < .001$.

Wykres nr 7

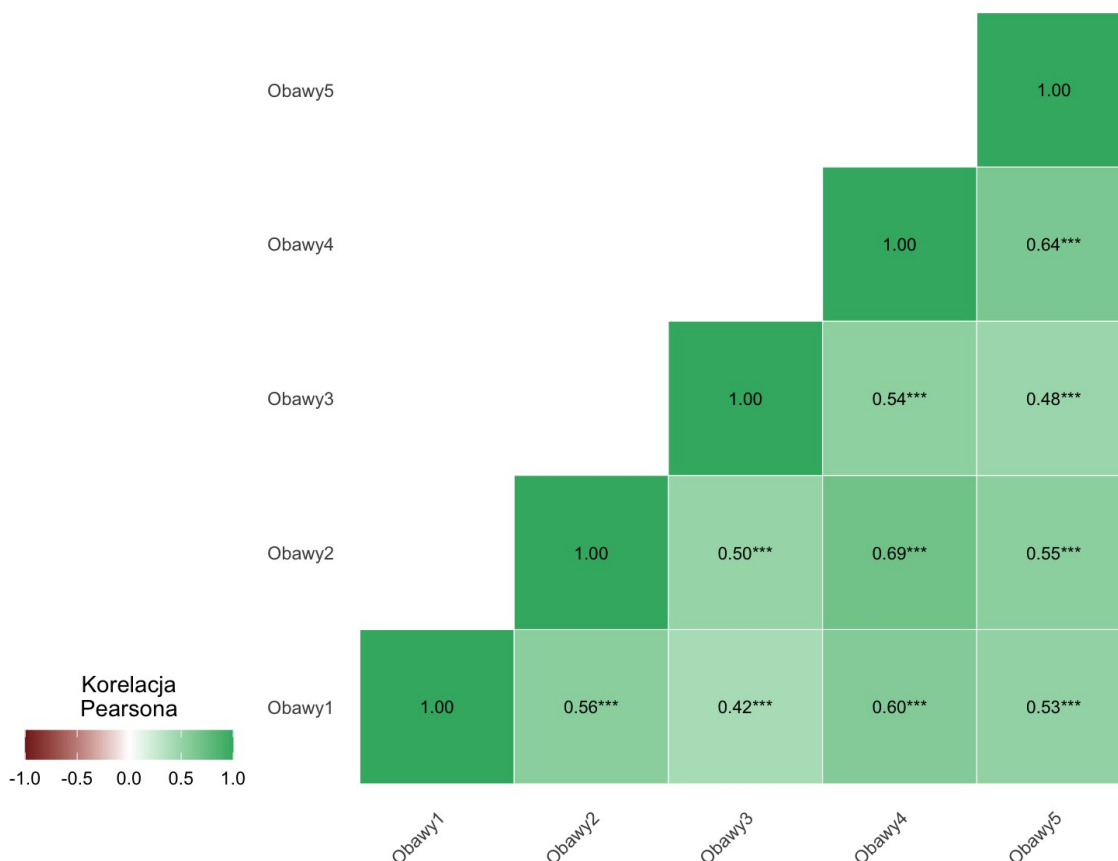
Korelacje pomiędzy składnikami budującymi pomiar Cele



Nota. Rysunek przedstawia oszacowania współczynników korelacji Pearsona. Im bliżej wartości 1 (im większe nasycenie zielonego koloru), tym silniejsze powiązanie między składnikami testowymi. Diagonalne wartości 1 wskazują na korelację składnika z samym sobą. Gwiazdki oznaczają istotność statystyczną: *** $p < .001$.

Wykres nr 8

Korelacje pomiędzy składnikami budującymi pomiar Obawy



Nota. Rysunek przedstawia oszacowania współczynników korelacji Pearsona. Im bliżej wartości 1 (im większe nasycenie zielonego koloru), tym silniejsze powiązanie między składnikami testowymi. Diagonalne wartości 1 wskazują na korelację składnika z samym sobą. Gwiazdki oznaczają istotność statystyczną: *** $p < .001$.

Podsumowanie wyników

Analizę rzetelności przeprowadzono na próbie 303 obserwacji, stosując współczynnik alfa Cronbacha i współczynnik omega McDonalda oraz, na podstawie jednoczynnikowej analizy składowych głównych, ocenę rzetelności kompozytowej, rhoA i AVE. Wszystkie badane skale wykazywały wysoką dokładność pomiarową i dobrą spójność wewnętrzną. Skale Intencja zakupu, Użyteczność, Łatwość i Cele wykazały wysoki poziom trafności zbieżnej oceniany przez AVE, natomiast skala Obawy osiągnęła poziom AVE uznany za akceptowalny; rzetelność kompozytowa i rhoA dla wszystkich pomiarów również były na wysokim poziomie. Ogólnie

rzecz biorąc, przeprowadzone analizy potwierdzały, że zastosowane pomiary miały odpowiednią rzetelność do dalszych analiz.

Wstępna weryfikacja hipotezy analizą korelacji Pearsona dla zmiennych Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu

Niniejszą analizę i raport opisowy wykonano analogicznie jak wcześniej w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023). Wizualizacje wyników przeprowadzono z wykorzystaniem pakietu graficznego “ggplot2” (Wickham, 2016). W celu weryfikacji zależności między zmiennymi Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu wykonano serię analiz korelacji metodą Pearsona (Pearson, 1895). Na analizę parametryczną zdecydowano się z powodu względnego spełnienia założenia o normalności rozkładów wyników w analizowanych zmiennych. Interpretację zakresu siły związku na podstawie 95% PU oparto na progach Cohena (1988). Analiza testem Pearsona wykazała, że:

Wzrost wyników zmiennej Cele wiązał się ze spadkiem wyników zmiennej Obawy, $r(301) = -0.39$; 95% PU $[-0.48, -0.29]$; $p < 0.001$ (95% PU wskazywał na zakres siły związku od słabej do umiarkowanej)

Wzrost wyników zmiennej Łatwość wiązał się ze spadkiem wyników zmiennej Obawy, $r(301) = -0.29$; 95% PU $[-0.39, -0.19]$; $p < 0.001$ (95% PU wskazywał na zakres siły związku od słabej do umiarkowanej)

Wzrost wyników zmiennej Użyteczność wiązał się ze spadkiem wyników zmiennej Obawy, $r(301) = -0.45$; 95% PU $[-0.53, -0.35]$; $p < 0.001$ (95% PU wskazywał na umiarkowaną siłę związku)

Wzrost wyników zmiennej Intencja zakupu wiązał się ze spadkiem wyników zmiennej Obawy, $r(301) = -0.41$; 95% PU $[-0.50, -0.31]$; $p < 0.001$ (95% PU wskazywał na umiarkowaną siłę związku)

Wzrost wyników zmiennej Łatwość wiązał się ze wzrostem wyników zmiennej Cele, $r(301) = 0.26$; 95% PU $[0.15, 0.36]$; $p < 0.001$ (95% PU wskazywał na zakres siły związku od słabej do umiarkowanej)

Wzrost wyników zmiennej Użyteczność wiązał się ze wzrostem wyników zmiennej Cele, $r(301) = 0.60$; 95% PU $[0.52, 0.67]$; $p < 0.001$ (95% PU wskazywał na umiarkowaną siłę związku)

Wzrost wyników zmiennej Intencja zakupu wiązał się ze wzrostem wyników zmiennej Cele, $r(301) = 0.63$; 95% PU [0.56, 0.70]; $p < 0.001$ (95% PU wskazywał na umiarkowaną siłę związku)

Wzrost wyników zmiennej Użyteczność wiązał się ze wzrostem wyników zmiennej Łatwość, $r(301) = 0.48$; 95% PU [0.39, 0.56]; $p < 0.001$ (95% PU wskazywał na umiarkowaną siłę związku)

Wzrost wyników zmiennej Intencja zakupu wiązał się ze wzrostem wyników zmiennej Łatwość, $r(301) = 0.35$; 95% PU [0.25, 0.45]; $p < 0.001$ (95% PU wskazywał na zakres siły związku od słabej do umiarkowanej)

Wzrost wyników zmiennej Intencja zakupu wiązał się ze wzrostem wyników zmiennej Użyteczność, $r(301) = 0.72$; 95% PU [0.66, 0.77]; $p < 0.001$ (95% PU wskazywał na zakres siły związku od umiarkowanej do mocnej). Rezultaty analizy przedstawia tabela nr 2 oraz rysunek nr 9.

Tabela nr 2

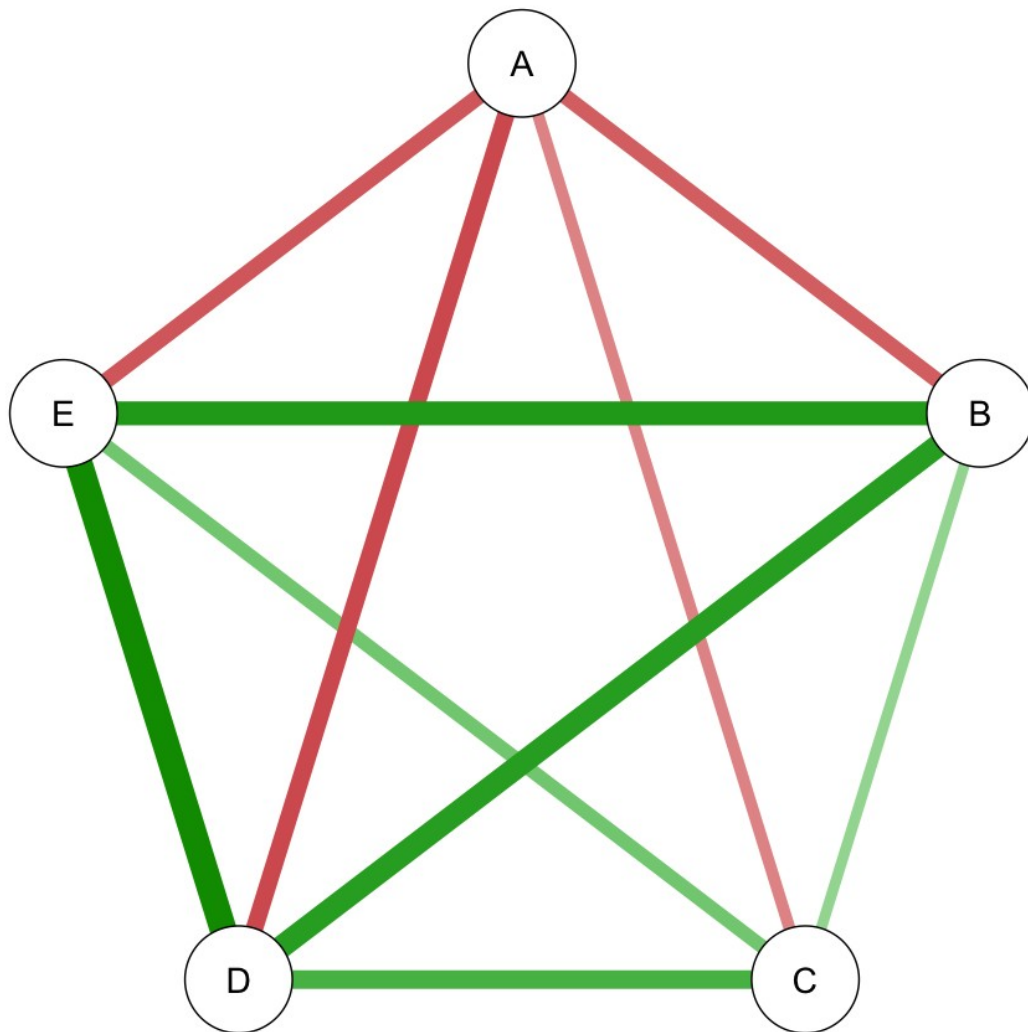
Wyniki analizy korelacji Pearsona między zmiennymi: Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu

Zmienna	Id.	1	2	3	4
Obawy	1				
Cele	2	-0.39 ***			
Łatwość	3	-0.29 ***	0.26 ***		
Użyteczność	4	-0.45 ***	0.60 ***	0.48 ***	
Intencja zakupu	5	-0.41 ***	0.63 ***	0.35 ***	0.72 ***

Nota. Liczba obserwacji w analizie wynosiła $N = 303$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Wykres nr 9

Wizualizacja relacji między zmiennymi Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu



Nota. A = Obawy, B = Cele, C = Łatwość, D = Użyteczność, E = Intencja zakupu

Im ciemniejszy kolor zielony = Bardziej dodatnia korelacja; Im ciemniejszy kolor czerwony = Bardziej ujemna korelacja. Ciągła linia - oznacza istotny związek między zmiennymi. Rysunek bazuje na uzyskanych oszacowaniach współczynników korelacji Pearsona

Porównanie korelacji między grupami

Dodatkowo porównano współczynniki korelacji między dwiema grupami wyznaczonymi przez zmienną Płeć: Kobieta oraz Mężczyzna. Różnice między współczynnikami korelacji oceniono z wykorzystaniem transformacji Fishera z (Fisher, 1921; Cohen et al., 2003). Interpretację zakresu siły związku na podstawie 95% PU oparto na progach Cohena (1988). Analiza wykazała, że:

Siła korelacji między zmiennymi Łatwość i Użyteczność różniła się istotnie między grupami. W grupie Mężczyzna zaobserwowano istotnie silniejszą dodatnią korelację ($r_{\text{Mężczyzna}} = 0.57$; 95% PU [0.44, 0.67]; $p < 0.001$; 95% PU wskazywał na umiarkowaną siłę związku) niż w grupie Kobieta ($r_{\text{Kobieta}} = 0.38$; 95% PU [0.25, 0.51]; $p < 0.001$; 95% PU wskazywał na zakres siły związku od słabej do umiarkowanej). Test różnicy: $z = -2.04$; $p = 0.041$. Bezwzględna różnica między korelacjami wynosiła $|\Delta r| = 0.18$, co wskazuje na małą różnicę praktyczną. Pozostałych 9 porównań współczynników korelacji nie okazało się istotnych statystycznie. Rezultaty analizy przedstawia tabela nr 3 oraz rysunek nr 10.

Tabela nr 3

Porównanie korelacji Pearsona między grupami zmiennej Płeć

Zmienna 1	Zmienna 2	r Kobieta	95% PU Kobieta	r Mężczyzna	95% PU Mężczyzna	z
Obawy	Cele	-0.36 ***	-0.48,-0.22	-0.47 ***	-0.59,-0.33	1.17
Obawy	Łatwość	-0.28 ***	-0.41,-0.14	-0.26 **	-0.41,-0.09	-0.21
Cele	Łatwość	0.17 *	0.03,0.31	0.36 ***	0.20,0.50	-1.74
Obawy	Użyteczność	-0.43 ***	-0.55,-0.30	-0.46 ***	-0.58,-0.31	0.26
Cele	Użyteczność	0.56 ***	0.45,0.66	0.64 ***	0.53,0.73	-1.11
Łatwość	Użyteczność	0.38 ***	0.25,0.51	0.57 ***	0.44,0.67	-2.04 *
Obawy	Intencja zakupu	-0.48 ***	-0.58,-0.35	-0.37 ***	-0.51,-0.21	-1.15
Cele	Intencja zakupu	0.58 ***	0.47,0.67	0.69 ***	0.59,0.77	-1.64
Łatwość	Intencja zakupu	0.32 ***	0.18,0.45	0.38 ***	0.22,0.52	-0.52
Użyteczność	Intencja zakupu	0.76 ***	0.68,0.81	0.69 ***	0.58,0.77	1.26

Nota. r = współczynnik korelacji w grupie; Liczebności w analizie wynosiły: n Kobieta = 173 i n Mężczyzna = 130; 95% PU w kolumnach grup oznacza 95% przedział ufności dla współczynnika korelacji w danej grupie, wyznaczony z wykorzystaniem transformacji Fishera; z = statystyka testu różnicy między niezależnymi współczynnikami korelacji po transformacji Fishera; gwiazdki przy wartościach r oraz z oznaczają poziom istotności: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Tabela nr 4

Łączona macierz korelacji Pearsona: dolny trójkąt = Kobieta, górny trójkąt = Mężczyzna

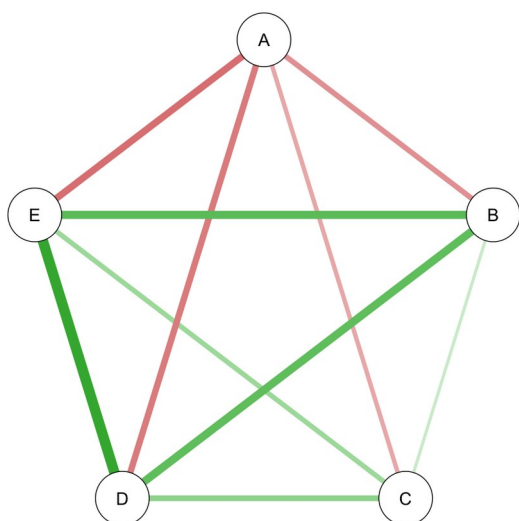
	Obawy	Cele	Łatwość	Użyteczność	Intencja zakupu
Obawy	-	-0.47***	-0.26**	-0.46***	-0.37***
Cele	-0.36***	-	0.36***	0.64***	0.69***
Łatwość	-0.28***	0.17*	-	0.57***	0.38***
Użyteczność	-0.43***	0.56***	0.38***	-	0.69***
Intencja zakupu	-0.48***	0.58***	0.32***	0.76***	-

Nota. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

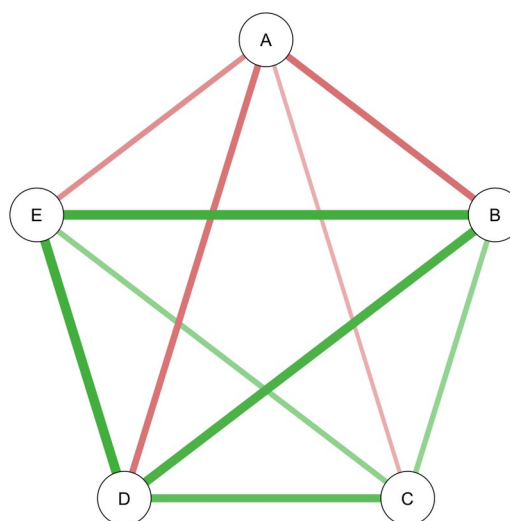
Wykres nr 10

Wizualizacja relacji między zmiennymi Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu w podziale na grupy zmiennej Płeć

Kobieta



Mężczyzna



Nota. A = Obawy, B = Cele, C = Łatwość, D = Użyteczność, E = Intencja zakupu

Im ciemniejszy kolor zielony = Bardziej dodatnia korelacja; Im ciemniejszy kolor czerwony = Bardziej ujemna korelacja. Ciągła linia - oznacza istotny związek między zmiennymi. Lewy panel przedstawia wyniki dla grupy Kobieta, a prawy panel przedstawia wyniki dla grupy Mężczyzna. Rysunek bazuje na uzyskanych oszacowaniach współczynników korelacji Pearsona.

Podsumowanie wyników analizy korelacji

Przeprowadzono analizę korelacji w całej próbie, w której stwierdzono szereg istotnych zależności między zmiennymi opisującymi postawy wobec pojazdów autonomicznych. Przede wszystkim zweryfikowano wstępnie postawione w pracy hipotezy. H1: Wzrost postrzeganych celów z wykorzystaniem pojazdów autonomicznych wiązał się ze wzrostem pomiarów akceptacji tej technologii została potwierdzona, ponieważ Cele korelowały dodatnio z miarami akceptacji. Hipoteza H2: Wzrost obaw przed pojazdami autonomicznymi wiązał się ze spadkiem pomiarów akceptacji tej technologii została potwierdzona, ponieważ Obawy korelowały ujemnie z miarami akceptacji. Hipoteza H3: Płeć moderowała związek pomiędzy łatwością korzystania z technologii pojazdów autonomicznych a ich użytecznością została wstępnie potwierdzona w analizie korelacji, gdyż związek między Łatwością a Użytecznością był silniejszy u mężczyzn niż u kobiet.

Model wpływu zmiennej Obawy, Cele na wyniki zmiennej Łatwość.

Niniejszą analizę i raport opisowy ponownie wykonano w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023). Wizualizacje wyników przeprowadzono z wykorzystaniem pakietu graficznego “*ggplot2*” (Wickham, 2016). Dla zmiennej Łatwość przeprowadzono wielozmienną analizę regresji liniowej (Fisher, 1922). W analizie wzięło udział $N = 303$ badanych obserwacji. Ze względu na odbieganie testowanego modelu dla wyników Łatwość od założeń modelu regresji, poddano tę zmienną procedurze normalizacji. W celu wyboru najlepszej metody normalizującej zmienną Łatwość wykonano serię przekształceń tej zmiennej szeregiem metod normalizujących. Wykorzystano w tym celu pakiet języka R *bestNormalize* (Peterson, 2021). Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika boxcox. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 31.37$. Zmienną tą wyrażono w postaci znormalizowanej (względem średniej wynoszącej 0 i odchylenia standardowego równego 1). Procedurze normalizacji poddano również zmienne wyjaśniające. Do testowanego modelu została przeliczona wizualna diagnostyka w pakiecie *performance* (Lüdtke et al., 2021), jest ona umieszczona w załączniku A do niniejszej analizy. Ocena statystyczna testowanego modelu wykazała istotne przewidywanie, $F(2, 300) = 18.68$; $p < 0.001$. Model wyjaśniał umiarkowaną część zmienności wyników zmiennej Łatwość, czyli około 11% (10% po skorygowaniu); $R^2 = 0.11$, $\text{adj.}R^2 = 0.10$. Oba predyktory uwzględnione w modelu okazały się istotne statystycznie. Analiza wykazała, że przewidywany przez model regresji średni poziom zmiennej Łatwość wynosił $M = 0.00$. Szczegółowa analiza poszczególnych predyktorów w modelu wykazała następujące rezultaty:

Najsilniejszym predyktorem zmiennej Łatwość okazała się zmienna Obawy, wykazująca umiarkowany ujemny związek z wynikami ($B = -0.25$; $t = -4.28$; $p < 0.001$; $\beta = -0.25$, 95%*PU* = $[-0.37; -0.14]$). Siła tego predyktora nie różniła się jednak istotnie od pozostałych istotnych zmiennych w modelu (nakładające się 95% przedziały ufności dla współczynników β). Istotny statystycznie związek ze zmienną Łatwość odnotowano również dla: Cele ($B = 0.14$; $t = 2.46$; $p = 0.015$; $\beta = 0.14$, 95%*PU* = $[0.03; 0.26]$).

Rezultaty oszacowań testowanego modelu przedstawia tabela nr 5. Wizualizacje wyników bazujących na oszacowaniach testowanego modelu przedstawia seria wykresów od nr 11 do nr 13

Tabela nr 5

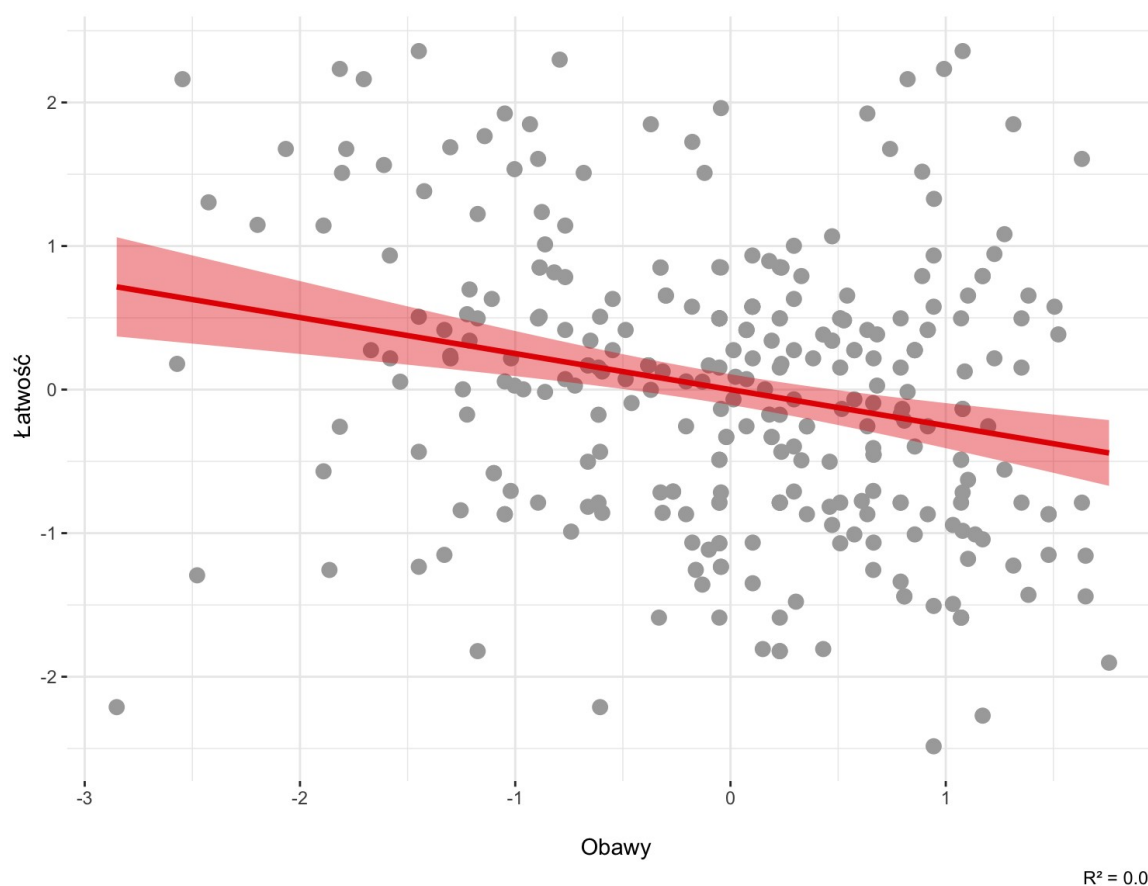
Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Łatwość

Zmienne w modelu	B	s.e.	t	DPU1	GPU1	p	β	DPU2	GPU2
Stała	0.00	0.05	0.00	-0.11	0.11	1.000	NA	NA	NA
Obawy	-0.25	0.06	-4.28	-0.37	-0.14	< 0.001	-0.25	-0.37	-0.14
Cele	0.14	0.06	2.46	0.03	0.26	0.015	0.14	0.03	0.26

Nota. B = Niestandaryzowany współczynnik regresji; s.e. = błąd standardowy dla B ; t = Statystyka t studenta; DPU = Dolny przedział ufności; GPU = Górny przedział ufności; DPU1 / GPU1 = 95% przedziały ufności dla B; p = Istotność statystyczna; β = Standaryzowany współczynnik regresji; DPU2 / GPU2 = 95% przedziały ufności dla β ; NA = Brak oszacowań standaryzowanych dla współczynnika stałej modelu regresji.

Wykres nr 11

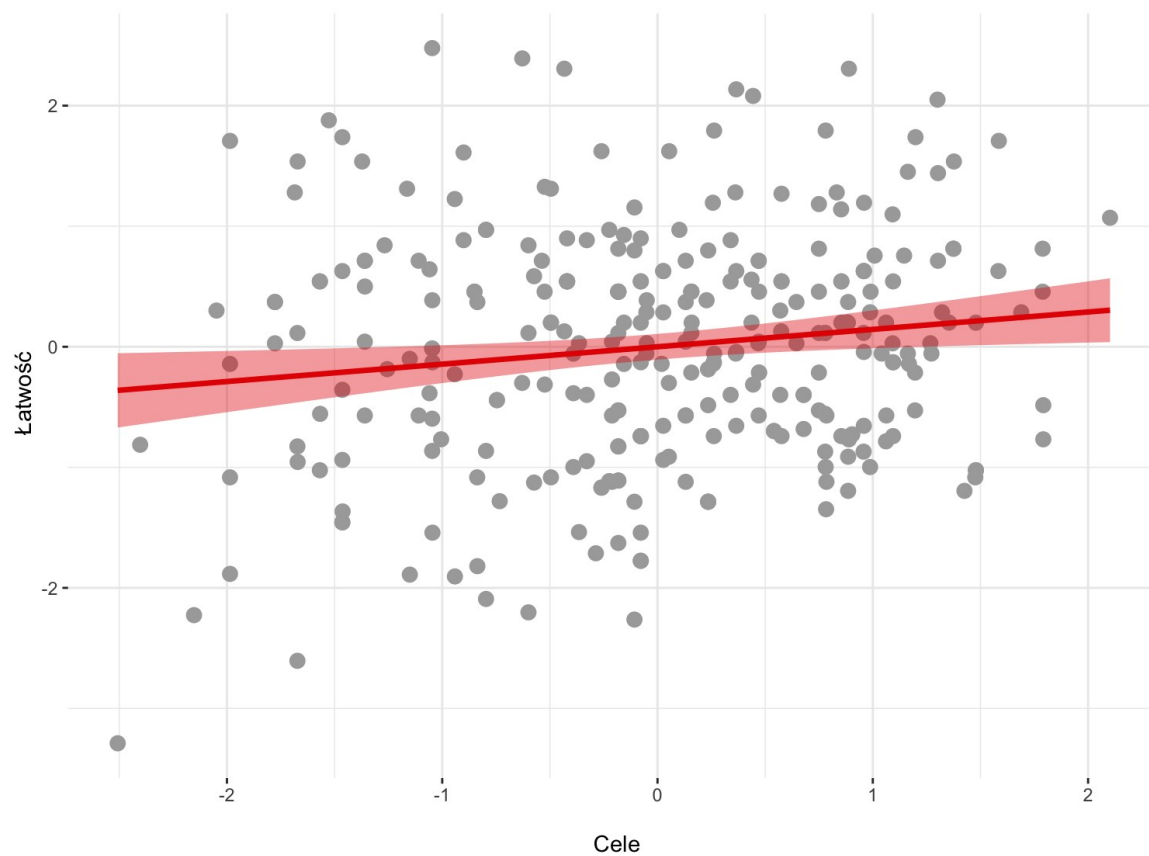
Wpływ zmiennej Obawy na zmienną Łatwość



Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 12

Wpływ zmiennej *Cele* na zmienną *Łatwość*

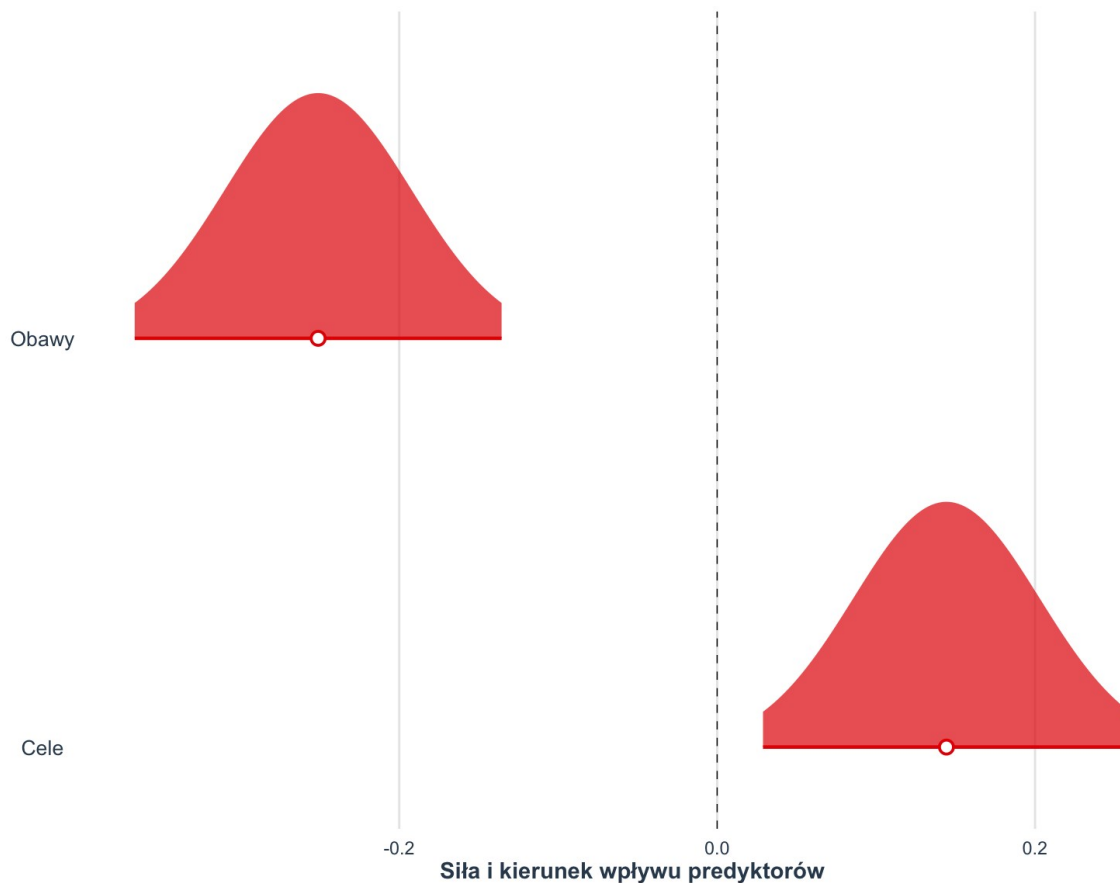


$R^2 = 0.02$

Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 13

Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Łatwość



Nota. Wąsy błędów przedstawiają 95% przedziały ufności dla oszacowania B. Linie zachodzące na siebie przedstawiają w przybliżeniu brak różnic między predyktorami we wpływie na poziom Łatwość. Natomiast, linie nie zachodzące na siebie przedstawiają w przybliżeniu istotne różnice we wpływie predyktorów na poziom zmiennej Łatwość.

Model wpływu zmiennej Obawy, Cele na wyniki zmiennej Użyteczność.

W celu wyjaśnienia zmiennej Użyteczność przeprowadzono wielozmienną analizę regresji liniowej (Fisher, 1922). W analizie wzięło udział $N = 303$ badanych obserwacji. Ze względu na odbieganie testowanego modelu dla wyników Użyteczność od założeń modelu regresji, poddano tę zmienną procedurze normalizacji. W celu wyboru najlepszej metody normalizującej zmienną Użyteczność wykonano serię przekształceń tej zmiennej szeregiem metod normalizujących. Wykorzystano w tym celu pakiet języka R *bestNormalize* (Peterson, 2021). Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika boxcox. W

porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 14.53$. Zmienną tą wyrażono w postaci znormalizowanej (względem średniej wynoszącej 0 i odchylenia standardowego równego 1). Procedurze normalizacji poddano również zmienne wyjaśniające. Do testowanego modelu została przeliczona wizualna diagnostyka w pakiecie *performance* (Lüdtke et al., 2021), jest ona umieszczona w załączniku A do niniejszej analizy. Analiza regresji wykazała istotne przewidywanie, $F(2, 300) = 107.57$; $p < 0.001$. W odróżnieniu od modelu dla zmiennej Łatwość ($R^2 = 0.11$), uwzględnione predyktory wyjaśniały znaczną część wariancji zmiennej Użyteczność, tj. około 42% (41% po skorygowaniu); $R^2 = 0.42$, $adj.R^2 = 0.41$. Oba predyktory uwzględnione w modelu okazały się istotne statystycznie. Analiza wykazała, że przewidywany przez model regresji średni poziom zmiennej Użyteczność wynosił $M = 0.00$. Analiza oszacowań poszczególnych predyktorów przyniosła następujące wyniki:

Największą siłę przewidywania zmiennej Użyteczność uzyskała zmienna Cele, dla której odnotowano bardzo silny dodatni związek ($B = 0.50$; $t = 10.44$; $p < 0.001$; $\beta = 0.50$, 95% $PU = [0.40; 0.59]$). Siła tego predyktora była istotnie większa niż pozostałych istotnych zmiennych w modelu (rozłączne 95% przedziały ufności dla współczynników β). Istotny wpływ na zmienną Użyteczność wykazano także w przypadku: Obawy ($B = -0.27$; $t = -5.68$; $p < 0.001$; $\beta = -0.27$, 95% $PU = [-0.36; -0.18]$). Rezultaty oszacowań testowanego modelu przedstawia tabela nr 6. Wizualizacje wyników bazujących na oszacowaniach testowanego modelu przedstawia seria wykresów od nr 14 do nr 16

Tabela nr 6

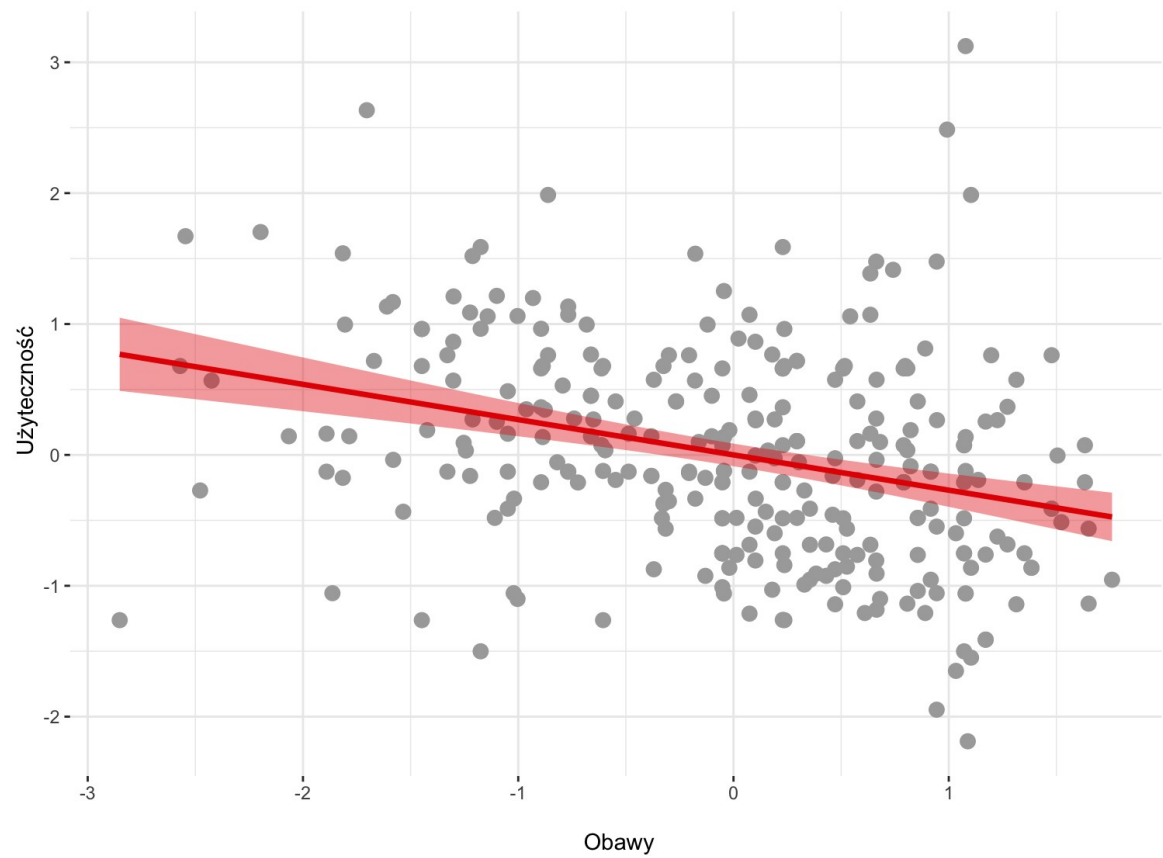
Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Użyteczność

Zmienne w modelu	B	$s.e.$	t	$DPU1$	$GPU1$	p	β	$DPU2$	$GPU2$
Stała	0.00	0.04	0.00	-0.09	0.09	1.000	NA	NA	NA
Obawy	-0.27	0.05	-5.68	-0.36	-0.18	< 0.001	-0.27	-0.36	-0.18
Cele	0.50	0.05	10.44	0.40	0.59	< 0.001	0.50	0.40	0.59

Nota. B = Niestandaryzowany współczynnik regresji; $s.e.$ = błąd standardowy dla B ; t = Statystyka t studenta; DPU = Dolny przedział ufności; GPU = Górny przedział ufności; $DPU1 / GPU1$ = 95% przedziały ufności dla B ; p = Istotność statystyczna; β = Standaryzowany współczynnik regresji; $DPU2 / GPU2$ = 95% przedziały ufności dla β ; NA = Brak oszacowań standaryzowanych dla współczynnika stałej modelu regresji.

Wykres nr 14

Wpływ zmiennej Obawy na zmienną Użyteczność

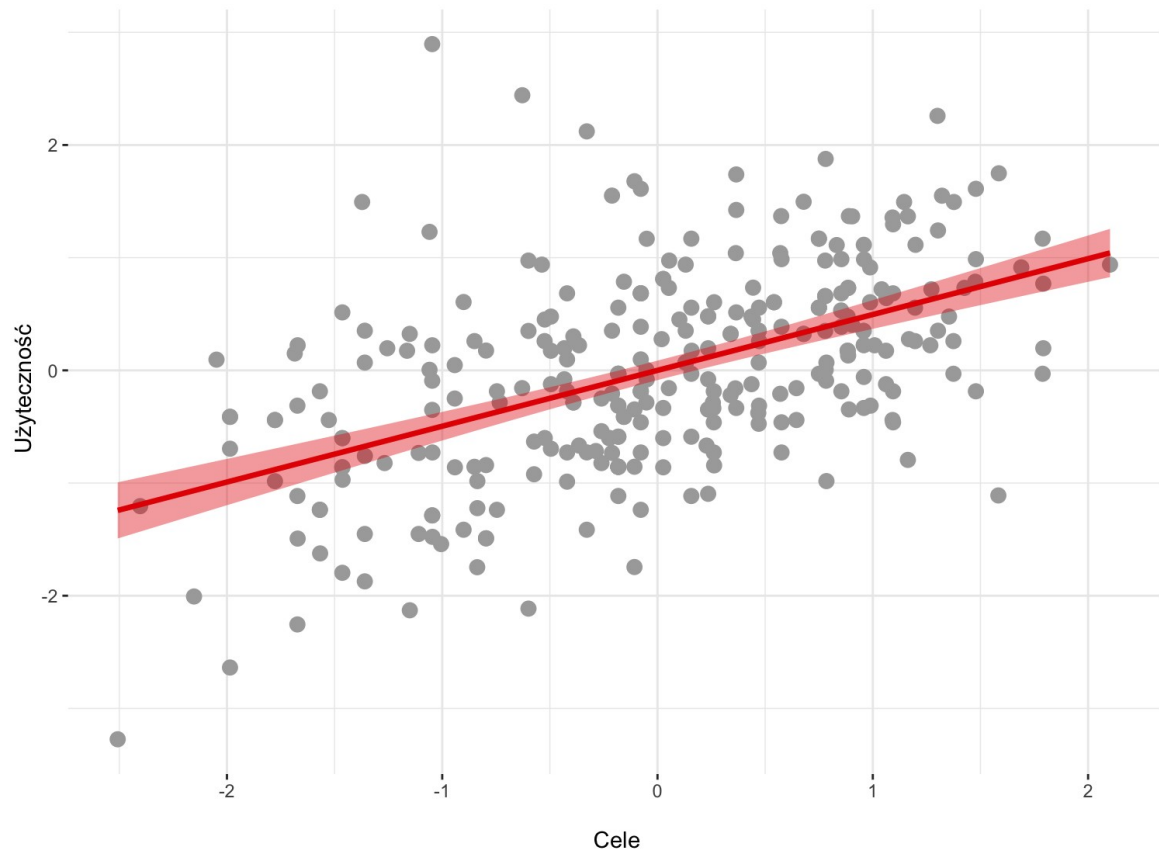


$R^2 = 0.10$

Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 15

Wpływ zmiennej *Cele* na zmienną *Użyteczność*

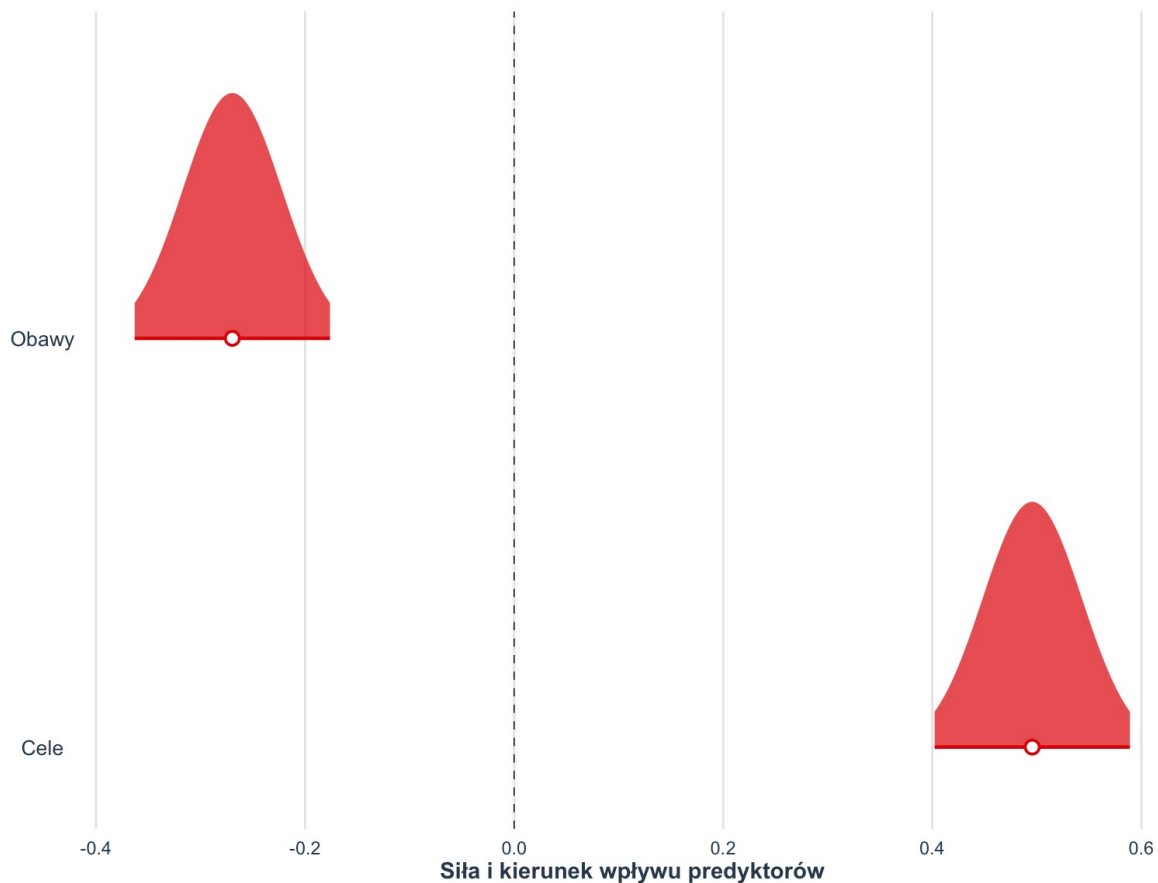


$R^2 = 0.27$

Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 16

Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Użyteczność



Nota. Wąsy błędów przedstawiają 95% przedziały ufności dla oszacowania B. Linie zachodzące na siebie przedstawiają w przybliżeniu brak różnic między predyktorami we wpływie na poziom Użyteczność. Natomiast, linie nie zachodzące na siebie przedstawiają w przybliżeniu istotne różnice we wpływie predyktorów na poziom zmiennej Użyteczność.

Model wpływu zmiennej Obawy, Cele na wyniki zmiennej Intencja zakupu.

W celu wyjaśnienia zmiennej Intencja zakupu przeprowadzono wielozmienną analizę regresji liniowej (Fisher, 1922). W analizie wzięło udział $N = 303$ badanych obserwacji. Ze względu na odbieganie testowanego modelu dla wyników Intencja zakupu od założeń modelu regresji, poddano tę zmienną procedurze normalizacji. W celu wyboru najlepszej metody normalizującej zmienną Intencja zakupu wykonano serię przekształceń tej zmiennej szeregiem metod normalizujących. Wykorzystano w tym celu pakiet języka R *bestNormalize* (Peterson, 2021). Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika boxcox. W

porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 8.31$. Zmienną tą wyrażono w postaci znormalizowanej (względem średniej wynoszącej 0 i odchylenia standardowego równego 1). Procedurze normalizacji poddano również zmienne wyjaśniające. Do testowanego modelu została przeliczona wizualna diagnostyka w pakiecie *performance* (Lüdtke et al., 2021), jest ona umieszczona w załączniku A do niniejszej analizy. Analiza regresji wykazała istotne przewidywanie, $F(2, 300) = 114.58$; $p < 0.001$. Podobnie jak w przypadku modelu dla zmiennej Użyteczność, uwzględnione predyktory wyjaśniały znaczną część wariancji zmiennej Intencja zakupu, tj. około 43% (43% po skorygowaniu); $R^2 = 0.43$, $\text{adj.}R^2 = 0.43$. Oba predyktory uwzględnione w modelu okazały się istotne statystycznie. Analiza wykazała, że przewidywany przez model regresji średni poziom zmiennej Intencja zakupu wynosił $M = 0.00$. Analiza oszacowań poszczególnych predyktorów przyniosła następujące wyniki:

Największą siłę przewidywania zmiennej Intencja zakupu uzyskała zmienna Cele, dla której odnotowano bardzo silny dodatni związek ($B = 0.56$; $t = 11.91$; $p < 0.001$; $\beta = 0.56$, $95\%PU = [0.47; 0.65]$). Siła tego predyktora była istotnie większa niż pozostałych istotnych zmiennych w modelu (rozłączne 95% przedziały ufności dla współczynników β). Istotny statystycznie związek ze zmienną Intencja zakupu odnotowano również dla: Obawy ($B = -0.20$; $t = -4.24$; $p < 0.001$; $\beta = -0.20$, $95\%PU = [-0.29; -0.11]$). Rezultaty oszacowań testowanego modelu przedstawia tabela nr 7. Wizualizacje wyników bazujących na oszacowaniach testowanego modelu przedstawia seria wykresów od nr 17 do nr 19

Tabela nr 7

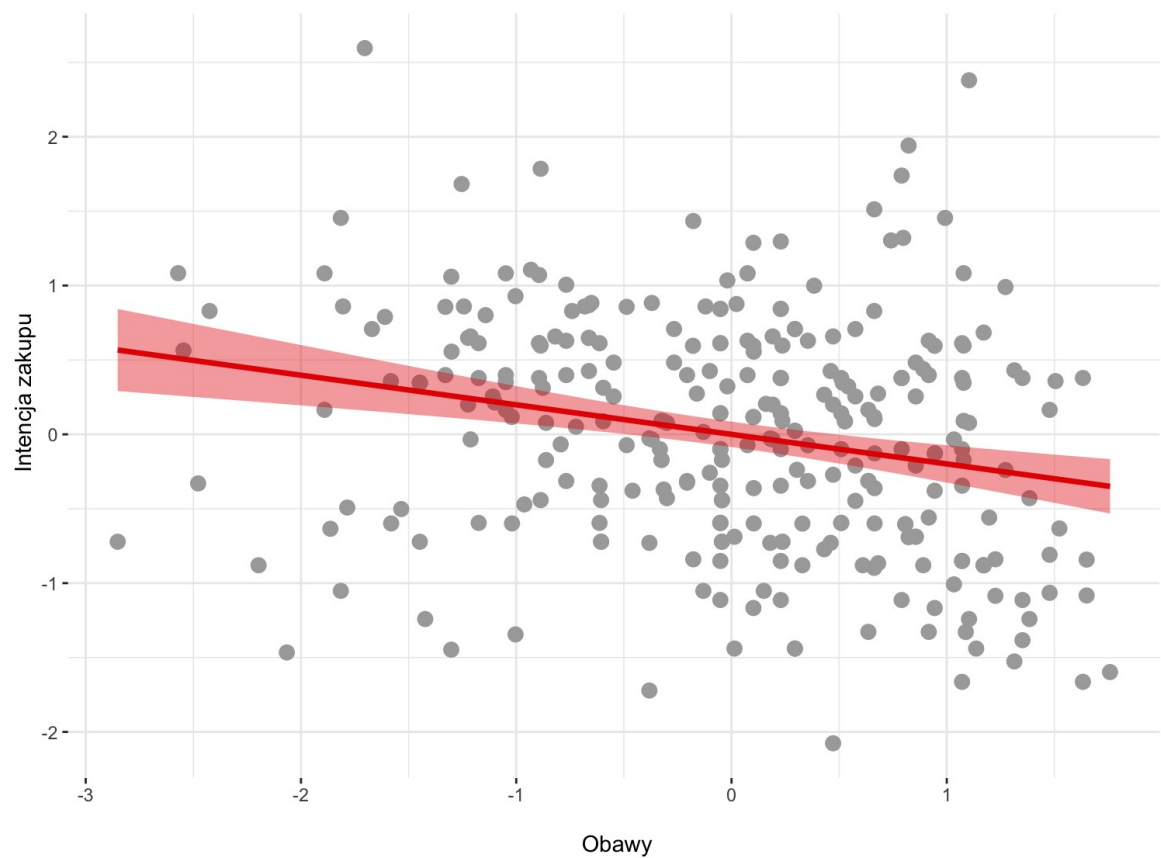
Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Intencja zakupu

Zmienne w modelu	<i>B</i>	<i>s.e.</i>	<i>t</i>	<i>DPU1</i>	<i>GPU1</i>	<i>p</i>	β	<i>DPU2</i>	<i>GPU2</i>
Stała	0.00	0.04	0.00	-0.09	0.09	1.000	NA	NA	NA
Obawy	-0.20	0.05	-4.24	-0.29	-0.11	< 0.001	-0.20	-0.29	-0.11
Cele	0.56	0.05	11.91	0.47	0.65	< 0.001	0.56	0.47	0.65

Nota. *B* = Niestandaryzowany współczynnik regresji; *s.e.* = błąd standardowy dla *B* ; *t* = Statystyka *t* studenta; *DPU* = Dolny przedział ufności; *GPU* = Górny przedział ufności; *DPU1* / *GPU1* = 95% przedziały ufności dla *B*; *p* = Istotność statystyczna; β = Standaryzowany współczynnik regresji; *DPU2* / *GPU2* = 95% przedziały ufności dla β ; NA = Brak oszacowań standaryzowanych dla współczynnika stałej modelu regresji.

Wykres nr 17

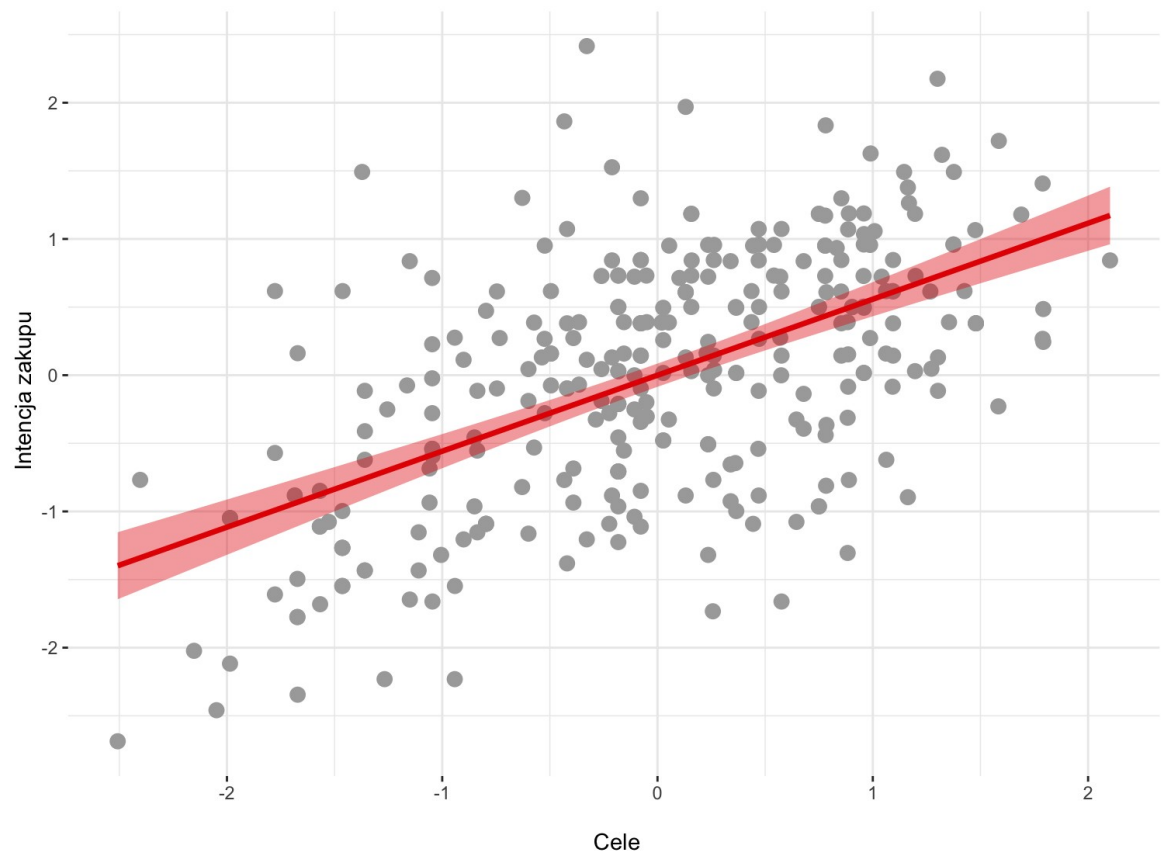
Wpływ zmiennej Obawy na zmienną Intencja zakupu



Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 18

Wpływ zmiennej Cele na zmienną Intencja zakupu

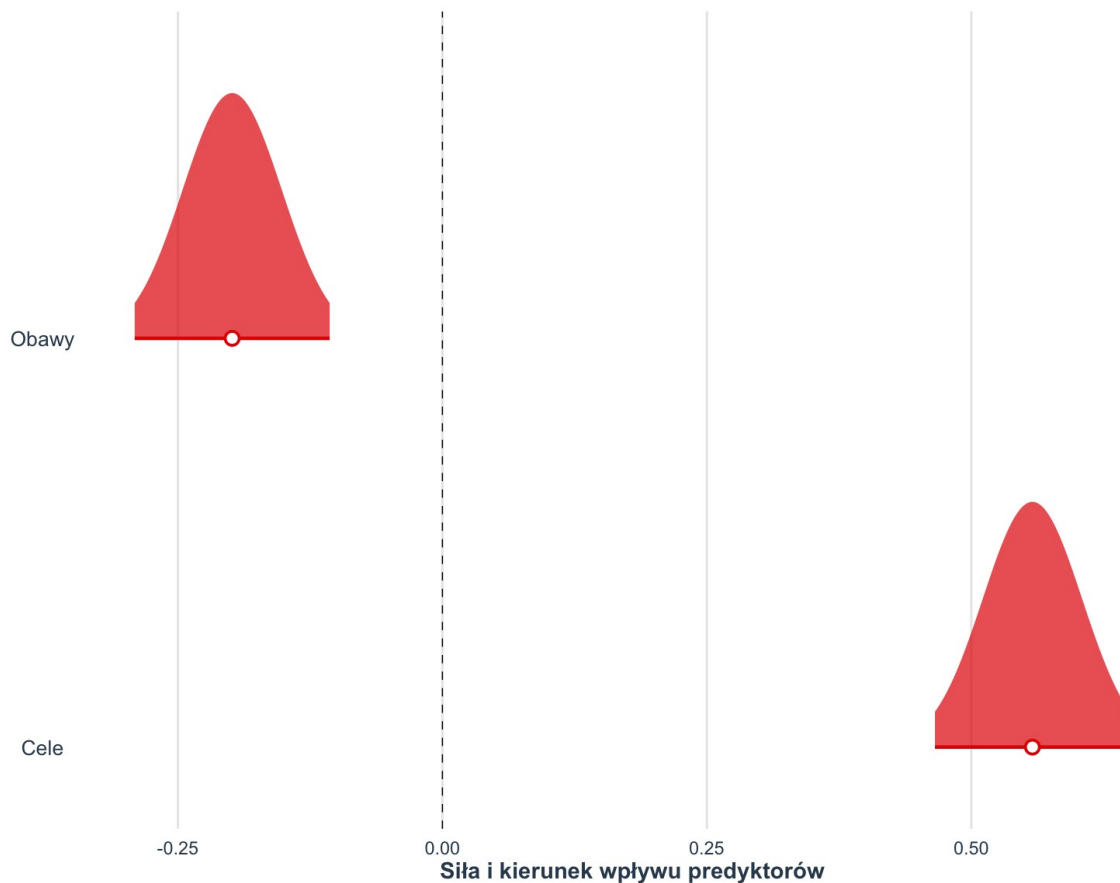


$R^2 = 0.32$

Nota. Ciągła linia - oznacza istotny wpływ predyktora.

Wykres nr 19

Wpływ zmiennych Obawy, Cele na poziom wyników zmiennej Intencja zakupu



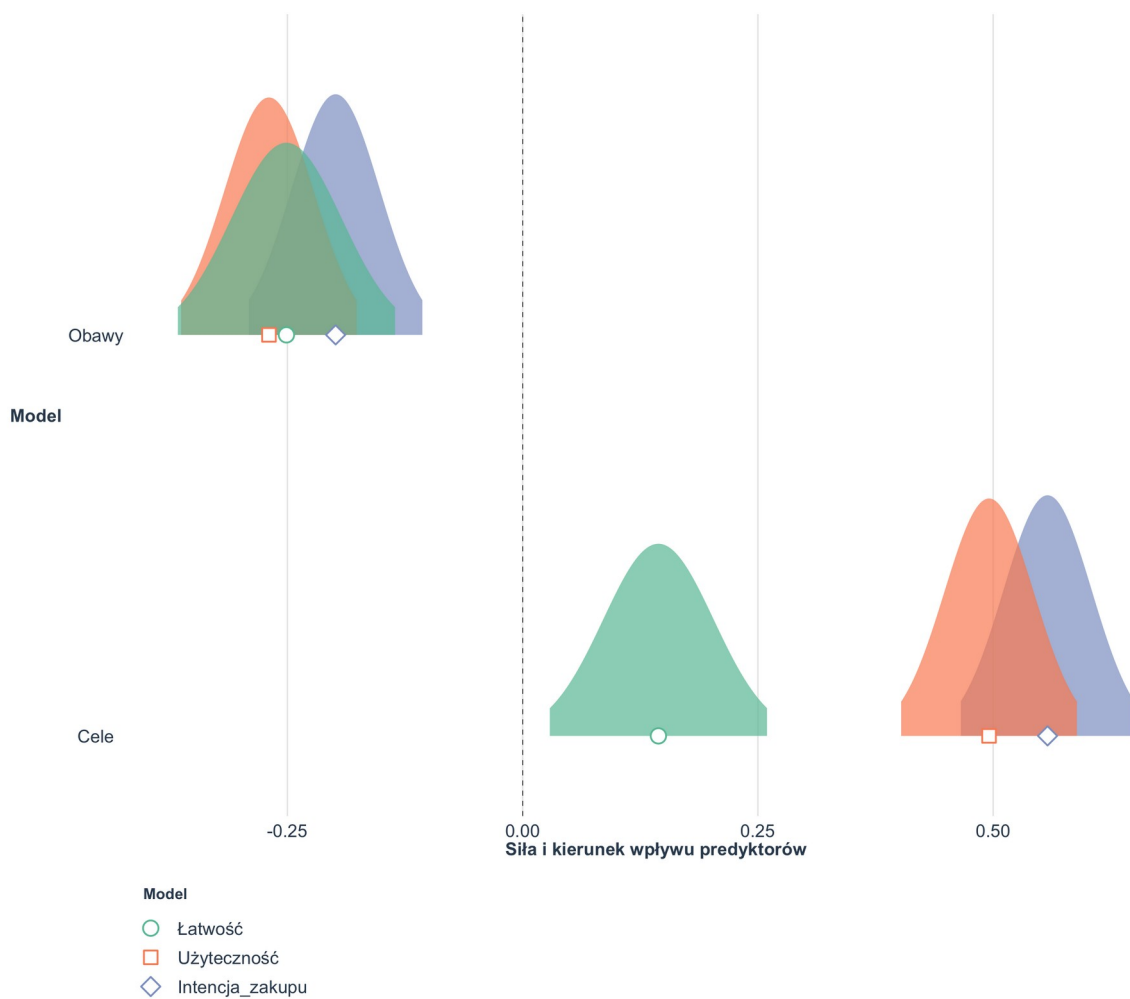
Nota. Wąsy błędów przedstawiają 95% przedziały ufności dla oszacowania B. Linie zachodzące na siebie przedstawiają w przybliżeniu brak różnic między predyktorami we wpływie na poziom Intencja zakupu. Natomiast, linie nie zachodzące na siebie przedstawiają w przybliżeniu istotne różnice we wpływie predyktorów na poziom zmiennej Intencja zakupu.

Porównanie testowanych modeli

Porównanie siły i kierunku wpływu predyktorów pomiędzy testowanymi modelami przedstawia Rysunek nr 20.

Wykres nr 20

Porównanie testowanych modeli



Nota. Wąsy błędów przedstawiają 95% przedziały ufności dla oszacowania B. Linie zachodzące na siebie przedstawiają w przybliżeniu brak różnic między predyktorami we wpływie na poziom zmiennych zależnych. Natomiast, linie nie zachodzące na siebie przedstawiają w przybliżeniu istotne różnice we wpływie predyktorów na poziom zmiennych zależnych.

Podsumowanie wyników

W trzech modelach regresji liniowej, przeprowadzonych na próbie 303 osób, analizowano wpływ postrzeganych celów i obaw związanych z pojazdami autonomicznymi na trzy miary akceptacji: Łatwość, Użyteczność i Intencję zakupu. Ze względu na naruszenia założeń modeli wszystkie zmienne zależne oraz predyktory zostały poddane normalizacji przy użyciu przekształcenia boxcox. Model wyjaśniał umiarkowaną część wariancji dla Łatwości,

natomiast znaczną część wariancji dla Użyteczności i Intencji zakupu. W modelu przewidującym Łatwość oba uwzględnione predyktory okazały się istotne; obawy wykazywały umiarkowany ujemny związek z oceną łatwości, natomiast cele wykazywały istotny, lecz słabszy dodatni związek; siła związku obaw nie różniła się istotnie od siły pozostałego predyktora. W modelu dla Użyteczności cele wykazywały bardzo silny dodatni związek i były najsilniejszym predyktorem tej zmiennej, natomiast obawy wykazywały istotny umiarkowany związek ujemny. W modelu dla Intencji zakupu cele również wykazywały bardzo silny dodatni związek i przewidywały zmienną silniej niż obawy, które miały istotny ujemny związek, lecz o mniejszej sile niż cele. Hipoteza H1, że wzrost postrzeganych celów wiązał się ze wzrostem pomiarów akceptacji technologii, potwierdziła się we wszystkich trzech modelach. Hipoteza H2, że wzrost obaw wiązał się ze spadkiem pomiarów akceptacji, także potwierdziła się we wszystkich trzech modelach.

Analiza moderacyjnej roli płci w relacji między Łatwością korzystania z technologii pojazdów autonomicznych z postrzeganą Użytecznością.

Niniejszą analizę i raport opisowy wykonano również w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023). Wizualizację wyników przeprowadzono z wykorzystaniem pakietu graficznego “ggplot2” (Wickham, 2016). W celu analizy wyników wykonano analizę regresji (Fisher, 1922), która wykazała, że na wyjaśnienie zmiennej Użyteczność wpływała istotnie zmienna Łatwość $b = 0.12$; $p < 0.001$, nie wpływała istotnie zmienna PłećMężczyzna $b = -0.87$; $p = 0.068$. Natomiast analiza efektów interakcyjnych w testowanym modelu wykazała, że w wyjaśnienie zmiennej Użyteczność zmienna Łatwość i zmienna PłećMężczyzna wchodziła ze sobą istotnie w interakcję $b = 0.06$; $p = 0.044$. Analiza współczynnika R^2 i jego skorygowanej wartości wykazała, że testowany model regresji wyjaśniał około 24.26% (23.50% po skorygowaniu) zmienności wyników zmiennej zależnej Użyteczność. Siłę wpływu testowanych zmiennych oceniano pod względem wartości statystyki d Cohena (Cohen, 1988). Wyniki oszacowań przedstawia tabela nr 8.

Tabela nr 8

Oszacowania analizy regresji dla zmiennej Użyteczność

Efekt	b	$s.e$	t	p	β	DPU	GPU
(Intercept)	-1.76	0.32	-5.49	0.000	NA	NA	NA
Łatwość	0.12	0.02	5.35	0.000	0.38	0.24	0.52
PłećMężczyzna	-0.87	0.48	-1.83	0.068	-0.87	-1.81	0.06
Łatwość*Płeć	0.06	0.03	2.02	0.044	0.06	0	0.13

Nota. b = niestandardyzowane oszacowanie analizy regresji; $s.e.$ = błąd standardowy oszacowania b , t = statystyka t Studenta; p = istotność statystyczna; β = Standaryzowany współczynnik regresji; DPU / GPU = 95% przedziały ufności dla β ; NA = Brak oszacowań standaryzowanych dla współczynnika stałej modelu regresji. R^2 i jego skorygowana wartość dla testowanego modelu wyniosła odpowiednio; 0.24 oraz 0.24.

Wpływ zmiennej Płeć na zmienną Użyteczność

Analiza wyników testowanego efektu głównego zmiennej Płeć w modelu regresji wykazała, że przyjmowanie kategorii Mężczyzna wiązało się ze spadkiem wyników zmiennej Użyteczność o wartość -0.87 w porównaniu do kategorii Kobieta. Uzyskany wynik nie był istotny statystycznie $b = -0.87$; $s.e. = 0.48$; $t = -1.83$; $p = 0.068$.

Wpływ zmiennej Łatwość na zmienną Użyteczność

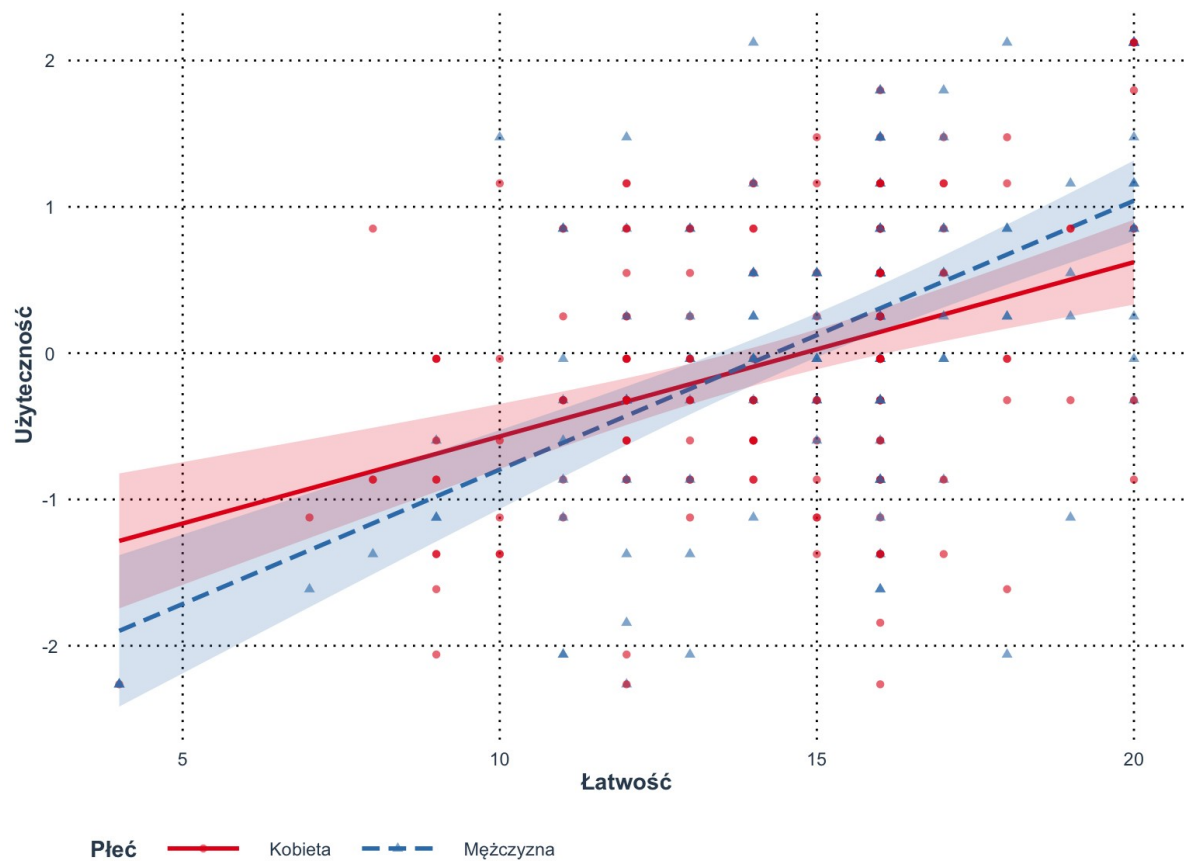
Analiza wyników testowanego efektu głównego w analizie regresji wykazała, że wzrost wyników zmiennej Łatwość o jeden punkt pomiarowy wiązał się ze wzrostem wyników zmiennej Użyteczność o wartość 0.12. Uzyskany wynik był istotny statystycznie $b = 0.12$; $s.e. = 0.02$; $t = 5.35$; $p < 0.001$.

Analiza wpływu zmiennej moderującej Płeć na wzorce wpływu zmiennej Łatwość na zmienność wyników Użyteczność

Analiza porównań nasilenia wpływu zmiennej Łatwość na zmienną Użyteczność wykazała, że w grupie Kobieta był statystycznie inny niż w grupie Mężczyzna $b = -0.06$; $t(299) = -2.02$; $p = 0.044$. Efekt siły różnic między tymi grupami był słaby, współczynnik d Cohena wyniósł $d = -0.07$; 95%PU[-0.15; 0]. W grupie Kobieta zaobserwowano istotnie słabszy wpływ zmiennej Łatwość na wzrost wyników zmiennej Użyteczność niż w grupie Mężczyzna, wyniki wynosiły odpowiednio $b_{\text{Kobieta}} = 0.12$; $t = 5.35$; $p < 0.001$ Vs $b_{\text{Mężczyzna}} = 0.18$; $t = 7.98$; $p < 0.001$. Wyniki przedstawia rysunek nr 21 oraz tabela nr 9 i 10.

Wykres nr 21

Analiza moderującej roli zmiennej Płeć na wzorcie wpływu zmiennej Łatwość na zmienność wyników Użyteczność



Nota. Kolorowe cieniowanie w okół linii regresji przedstawia 95% przedziały ufności dla oszacowań regresyjnych.

Tabela nr 9

Oszacowania istotności różnic pod względem wpływu zmiennej Łatwość na zmienną Użyteczność wynikającą z wpływu moderacyjnego efektu zmiennej Płeć

Porównanie	Różnica	t	p	d Cohena	DPU	GPU
Kobieta - Mężczyzna	-0.06	-2.02	0.044	-0.07	-0.15	0

Nota. Porównanie = Statystyczna ocena różnic we wpływie zmiennej Łatwość na zmienną Użyteczność w grupach zmiennej moderującej Płeć; Różnica = wielkość różnicy nasilenia wyników w poszczególnych grupach zmiennej Płeć; t = statystyka t Studenta; d Cohena; = współczynnik wielkości efektu d Cohena; DPU = Dolny przedział ufności dla efektu d Cohena; GPU = Górny przedział ufności dla efektu d Cohena. p = istotność statystyczna wyliczona bez korekty porównań wielokrotnych parami

Tabela nr 10

Oszacowania regresyjne wpływu zmiennej Łatwość na zmienną Użyteczność w podgrupach zmiennej moderującej Płeć

Grupa	<i>b</i>	<i>s.e.</i>	<i>DPU</i>	<i>GPU</i>	<i>t</i>	<i>p</i>
Kobieta	0.12	0.02	0.08	0.16	5.35	< 0.001
Mężczyzna	0.18	0.02	0.14	0.23	7.98	< 0.001

Nota. *b* = Oszacowanie wpływu zmiennej Łatwość na zmienną Użyteczność wynikającą z wpływu moderacyjnego efektu zmiennej Płeć; *s.e.* = błąd standardowy oszacowania *b* ; *t* = statystyka *t* Studenta; *DPU* = Dolny przedział ufności dla *b* ; *GPU* = Górny przedział ufności dla *b* . *p* = istotność statystyczna wyliczona bez korekty porównań wielokrotnych parami

Podsumowanie wyników

W pierwszym kroku poddawano zmienną Użyteczność normalizacji i wybierano najlepsze przekształcenie, przy czym metoda boxcox dawała najlepsze właściwości normalizacyjne. W analizie regresji wzrost zmiennej Łatwość wiązał się ze wzrostem zmiennej Użyteczność; jeden punkt wyższej Łatwości przekładał się na wzrost Użyteczności o 0.12 i ten związek był istotny. Sama Płeć nie wpływała istotnie na poziom Użyteczności, współczynnik dla kategorii Mężczyzna był ujemny, ale ten efekt nie był istotny. W modelu występowała istotna interakcja między Łatwością a Płcią, co oznaczało, że związek między Łatwością a Użytecznością różnił się w zależności od płci; wpływ Łatwości był silniejszy w grupie Mężczyzna niż w grupie Kobieta. Różnica ta była słaba pod względem wielkości efektu, współczynnik *d* Cohena wynosił około -0.07. Cały model regresji wyjaśniał około 24.26% wariancji zmiennej Użyteczność (około 23.50% po korekcie). Hipoteza H3, że Płeć moderowała związek pomiędzy łatwością korzystania z technologii pojazdów autonomicznych a ich użytecznością, potwierdziła się.

Podsumowanie

Przeprowadzono opisowe analizy częstości dla zmiennych demograficznych (Płeć, Edukacja, Wiek) na próbie $N = 303$. W próbie przeważały kobiety, dominowało wyższe wykształcenie, a największy udział miały osoby w wieku 20–29 lat. Siła efektu była słaba dla zmiennej Płeć, natomiast dla Edukacji i Wiek zróznicowanie kategorii osiągnęło wartości wskazujące na wyraźne przewagi występowania wymienionych grup.

Przeprowadzono kompleksową analizę rzetelności skal przy użyciu Alfa Cronbacha, Omega McDonalda, rzetelności kompozytowej (CR), rhoA oraz AVE na tej samej próbie. Wszystkie skale wykazywały dobrą spójność wewnętrzną i akceptowalną dokładność pomiarową; AVE wskazywała na wysoki poziom trafności zbieżnej dla skal Intencja zakupu, Użyteczność, Łatwość i Cele, natomiast skala Obawy osiągnęła poziom AVE uznany za akceptowalny. Na tej podstawie sklasyfikowano zastosowane pomiary jako wystarczająco rzetelne do dalszych analiz.

Analizy korelacji Pearsona ujawniły spójny wzorzec zależności między badanymi wymiarami: Obawy korelowały ujemnie z Cele, Łatwością, Użytecznością i Intencją zakupu, natomiast Cele, Łatwość, Użyteczność i Intencja zakupu wykazywały korelacje dodatnie między sobą. Najsilniejszy związek obserwowano między Użytecznością a Intencją zakupu, a związki między Obawami a pozostałymi zmiennymi miały siłę od słabej do umiarkowanej. Porównania korelacji wg płci wykazały jedynie jedną istotną różnicę: związek między Łatwością a Użytecznością był silniejszy u mężczyzn niż u kobiet, przy czym różnica ta miała niewielkie znaczenie praktyczne; pozostałe porównania korelacji nie były istotne statystycznie.

W trzech modelach regresji liniowej badano wpływ Celów i Obaw na Łatwość, Użyteczność i Intencję zakupu; ze względu na naruszenia założeń wszystkie zmienne zależne i predyktory zostały znormalizowane (metoda boxcox). Model dla Łatwości wyjaśniał umiarkowaną część wariancji, a oba predyktory były istotne: Obawy wykazywały umiarkowany ujemny wpływ, a Cele słabszy dodatni wpływ. Modele dla Użyteczności oraz Intencji zakupu wyjaśniały znaczną część wariancji; Cele były najsilniejszym, bardzo silnym predyktorem obu zmiennych, natomiast Obawy wykazywały istotny, umiarkowany efekt ujemny. Dodatkowo w analizie predykcji Użyteczności stwierdzono istotną interakcję Łatwość $\times$ Płeć: główny efekt

Płci był nieistotny, natomiast wpływ Łatwości na Użyteczność był silniejszy u mężczyzn niż u kobiet, z małym rozmiarem efektu. Hipotezy zakładające dodatni wpływ postrzeganych celów na akceptację technologii (H1) oraz ujemny wpływ obaw na akceptację (H2) potwierdziły się w analizach korelacji i we wszystkich modelach regresji, hipoteza o moderującej roli płci w związku między Łatwością a Użytecznością (H3) uzyskała wsparcie zarówno w analizie korelacji między grupami, jak i w modelu interakcyjnym, przy czym obserwowany efekt moderacji miał niewielką praktyczną wielkość.

Bibliografia

- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26 (2), 211–243.
- Cohen, J. (1988). The effect size. *Statistical Power Analysis for the Behavioral Sciences*. Abingdon: Routledge, 77–83.
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.
<https://doi.org/10.4324/9780203771587>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334, <https://doi.org/10.1007/BF02310555>
- Dallal, G.E. and Wilkinson, L. (1986) An analytic approximation to the distribution of Lilliefors' test for normality. *The American Statistician*, 40, 294–296.
- Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, 39(2), 297–316.
- Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3–32.
- Fisher, R. A. 1922. The goodness of fit of regression formulae, and the distribution of regression coefficients. *Journal of the Royal Statistical Society*. 85 (4), pp. 597–612.
<https://doi.org/10.2307/2341124>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.
<https://doi.org/10.1177/002224378101800104>

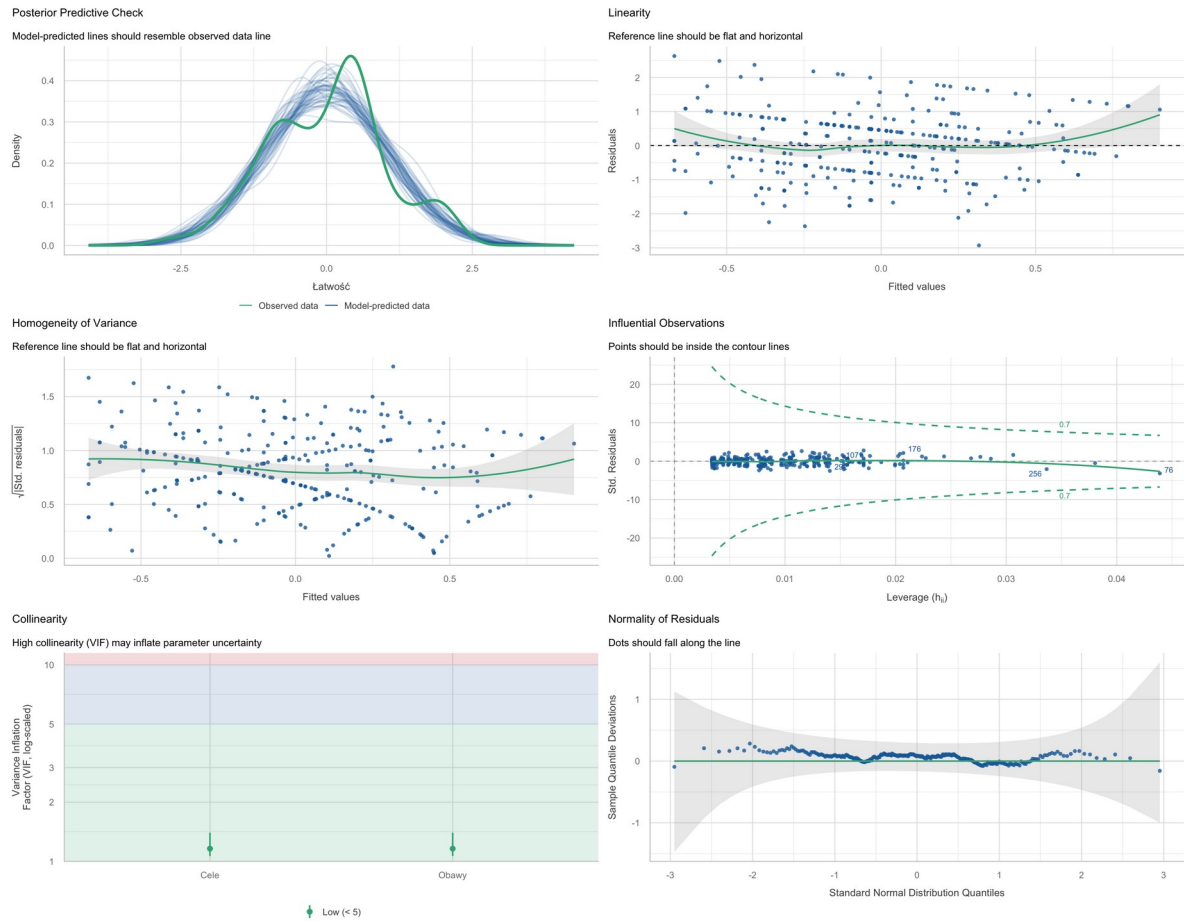
- Hryniewicz, K., Milewska, A. (2023). SZTOS: System Zautomatyzowanego Tworzenia Opisu Statystycznego (Wersja SZTOS) [Oprogramowanie]. <https://sztos-it.com/>
- Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47 (1), 183–189.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika* , 36, 109-133. <https://doi.org/10.1007/BF02291393>
- Karl Pearson F.R.S. (1900) X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling , *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50:302, 157-175, DOI: 10.1080/14786440009463897
- Lüdtke et al., (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- McDonald, R. P. (1999). *Test theory: A unified treatment* . Lawrence Erlbaum Associates.
- Pearson, K. (1895) Notes on Regression and Inheritance in the Case of Two Parents *Proceedings of the Royal Society of London*, 58, 240-242. <https://doi.org/10.1098/rspl.1895.0041>
- Peterson, R. A. (2021). Finding Optimal Normalizing Transformations via bestNormalize. *The R Journal*, 13:1, 310-329. DOI:10.32614/RJ-2021-041
- Royston, J. P. (1982). An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* , 31(2), 115–124. <https://doi.org/10.2307/2347973>
- Ryan A. Peterson & Joseph E. Cavanaugh (2020). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 47:13-15, 2312-2327, DOI: 10.1080/02664763.2019.1630372
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* . Springer-Verlag New York. ISBN 978-3-319-24277-4

Załącznik A

Wizualna diagnostyka testowanych modeli regresji

Wykres nr 1

Wizualna diagnoza modelu regresji dla zmiennej Łatwość

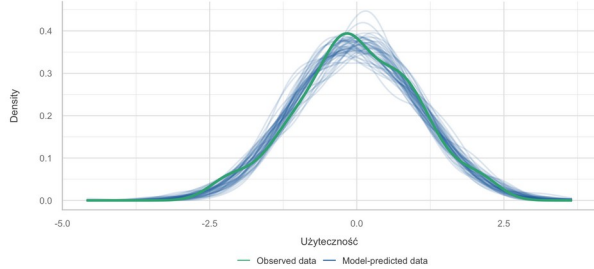


Wykres nr 2

Wizualna diagnoza modelu regresji dla zmiennej *Użyteczność*

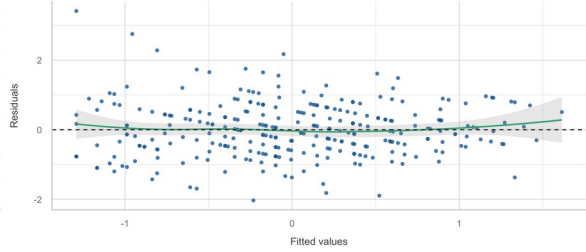
Posterior Predictive Check

Model-predicted lines should resemble observed data line



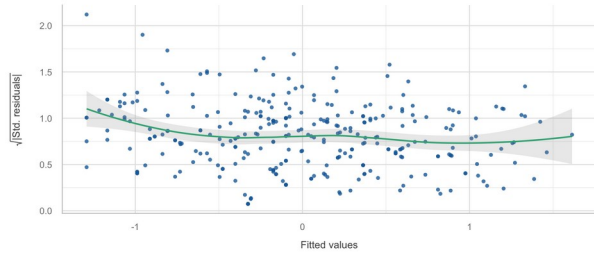
Linearity

Reference line should be flat and horizontal



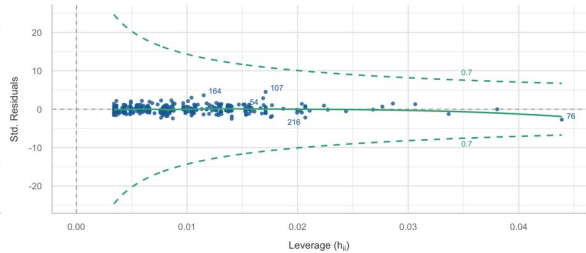
Homogeneity of Variance

Reference line should be flat and horizontal



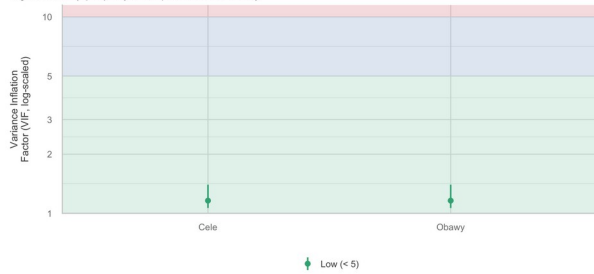
Influential Observations

Points should be inside the contour lines



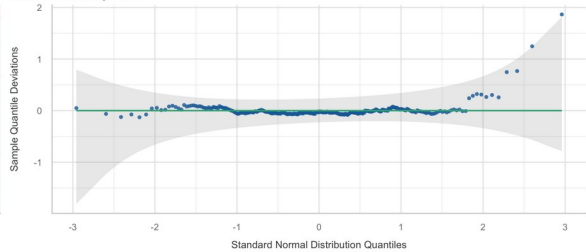
Collinearity

High collinearity (VIF) may inflate parameter uncertainty



Normality of Residuals

Dots should fall along the line



Wykres nr 3

Wizualna diagnoza modelu regresji dla zmiennej Intencja zakupu

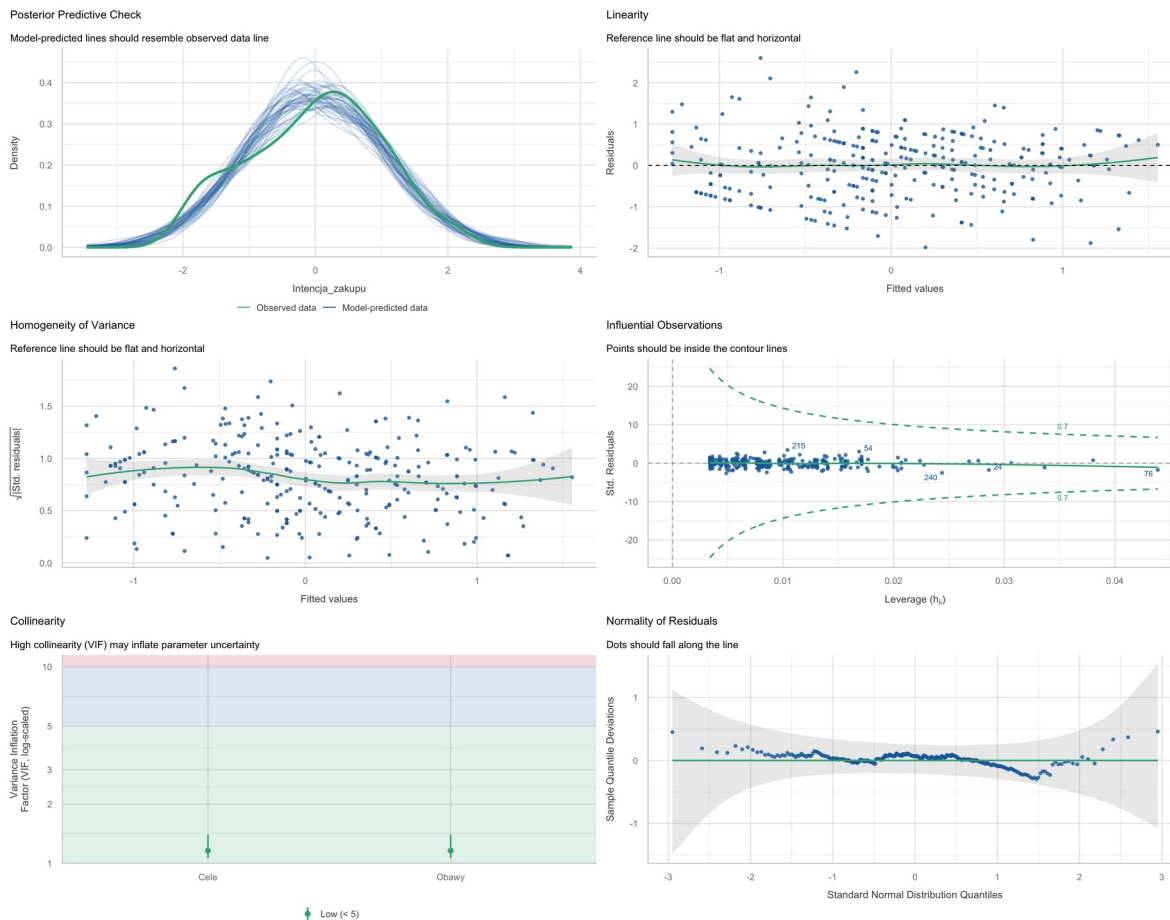


Tabela uogólniająca przeprowadzoną diagnozę modeli

Ocena modeli regresji dla zmiennych Łatwość, Użyteczność, Intencja zakupu

Tabela nr 1

Model	R^2	$adj.R^2$	RMSE	Sigma	AIC_{wt}	AIC_{wt}	BIC_{wt}	Performance Score
Intencja zakupu	0.43	0.43	0.75	0.76	0.98	0.98	0.98	1.00
Użyteczność	0.42	0.41	0.76	0.77	0.02	0.02	0.02	0.55
Łatwość	0.11	0.10	0.94	0.95	0.00	0.00	0.00	0.00

Nota. Uszeregowanie modeli w tabeli = Przedstawione modele w tabeli są uszeregowane (z góry na dół) słabnąco pod względem jakości. Objaśnienia wskaźników dopasowania modeli; Model = Nazwa zmiennej zależnej analizowanego modelu; R^2 = Współczynnik determinacji, czyli proporcja wariancji zmiennej zależnej wyjaśniona przez model. Wartości bliższe 1 oznaczają lepsze dopasowanie; $adj.R^2$ = Skorygowany współczynnik determinacji; uwzględnia liczbę predyktorów i wielkość próby. Chroni przed sztucznym zawyżaniem R^2 przy dodawaniu kolejnych zmiennych; RMSE (Root Mean Squared Error) = Średni błąd kwadratowy; informuje, jak bardzo średnio przewidywania modelu różnią się od wartości obserwowanych. Niższe wartości oznaczają lepsze dopasowanie; Sigma = Estymowana wartość odchylenia standardowego składnika resztowego (σ). Mniejsza sigma oznacza

mniejszy rozrzut reszt; AIC wt (Akaike Information Criterion weight) = waga modelu obliczona na podstawie kryterium informacyjnego AIC. Wyższa wartość oznacza lepszy model w porównaniu z innymi; AICc wt = Jak wyżej, ale skorygowane dla małych prób (AICc); BIC wt (Bayesian Information Criterion weight) = Waga modelu obliczona na podstawie kryterium informacyjnego BIC; Performance Score = Ogólny wskaźnik jakości modelu wyliczony na podstawie wielu miar; wartości bliższe 1 sugerują lepsze dopasowanie.

Załącznik B

Eksploracja zbioru danych

Niniejszą analizę eksploracyjną zbioru danych wykonano w Systemie Zautomatyzowanego Tworzenia Opisu Statystycznego - SZTOS (Hryniewicz, Milewska, 2023) w module eksploracja.

W pierwszej części raportu została umieszczona informacja o liczbie obserwacji oraz liczbie zmiennych we wprowadzonej bazie danych.

Aby usunąć z bazy danych a) obserwacje które nie zawierają żadnych danych, b) zmienne których liczba obserwacji jest mniejsza lub równa 1, zidentyfikowano takie przypadki. Jeśli takowe wystąpiły, zostały usunięte wraz z podaniem informacji o ich współrzędnych. Tej samej procedurze zostały poddane zmienne, które posiadały identyczne wartości.

Liczba obserwacji: 303

Liczba zmiennych: 8

W bazie danych **nie występują** OBSERWACJE, które zawierają wyłącznie braki danych. Obserwacje, które zawierały wyłącznie braki danych to: brak.

W bazie danych **nie wystąpiły** ZMIENNE, które zawierają wyłącznie braki danych lub 1 obserwację.

Zmienne, które zawierały wyłącznie braki danych lub 1 obserwację to: brak.

W bazie danych **nie wystąpiły** zmienne, które miały zerową zmienność.

Zmienne, które posiadały identyczne wartości to: brak.

Wyniki po wstępnym oczyszczeniu są zapisane w bazie danych

baza_wstepnie_oczyszczona.xlsx

Stan braków danych w analizowanej bazie danych

W tak przygotowanej bazie danych zidentyfikowano zmienne, które posiadają odsetek braków danych większy niż ten wyznaczony przez użytkownika, a informację podano do raportu.

W bazie danych **nie wystąpiły** ZMIENNE, które zawierają ponad lub równo: 5% braków danych.

Zmienne, które zawierają równo lub ponad 5% braków danych, to: brak.

Dokładne wyniki przedstawia Tabela nr 2.

Tabela nr 1

Następnie zidentyfikowano obserwacje, które zawierają ponad połowę braków danych i podano ich współrzędne. Taki stan rzeczy może wynikać np. z porzucenia badania w trakcie jego trwania. Może to powodować ograniczoną przydatność takich obserwacji w analizach

Zmienna	% braków danych	n/N braków danych
Płeć	0.00%	0/303
Edukacja	0.00%	0/303
Wiek	0.00%	0/303
Obawy	0.00%	0/303
Cele	0.00%	0/303
Łatwość	0.00%	0/303
Użyteczność	0.00%	0/303
Intencja zakupu	0.00%	0/303

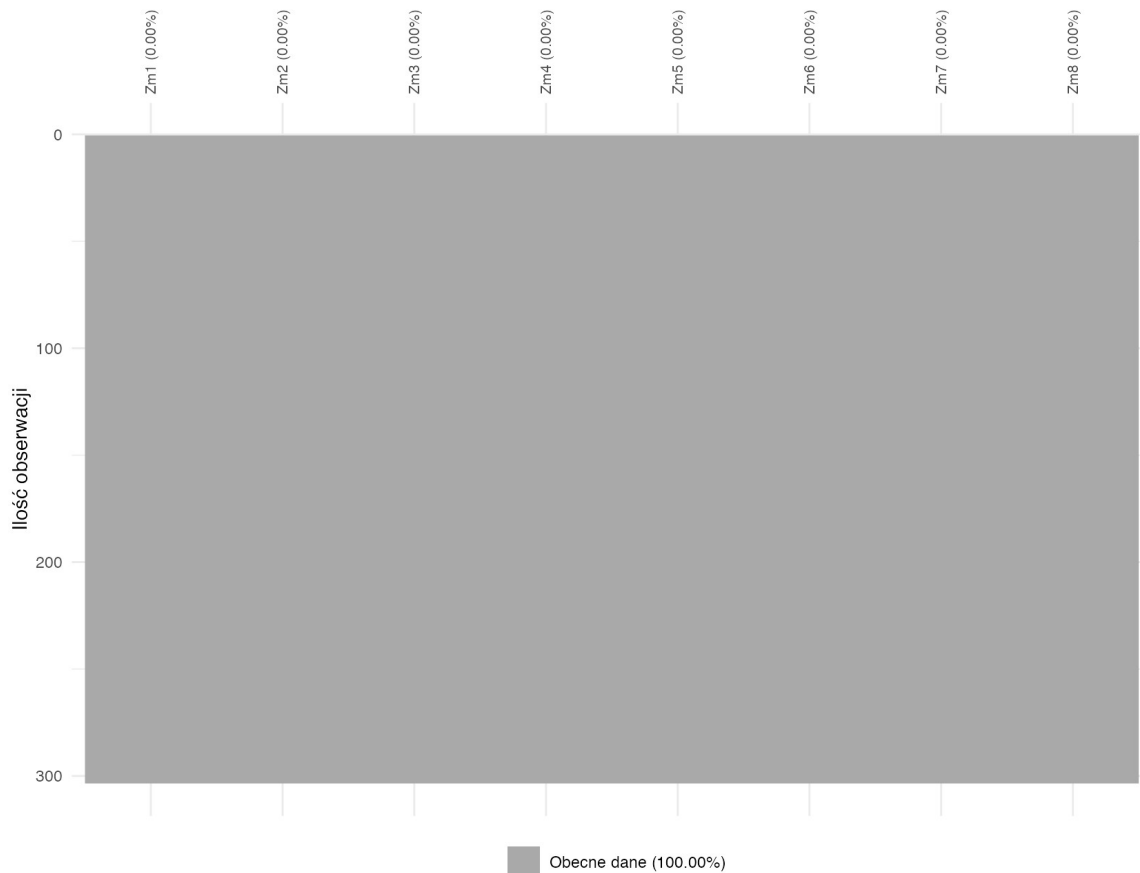
W bazie danych **nie wystąpiły** OBSERWACJE, które zawierają ponad połowę braków danych.

Numery wierszy, które zawierają ponad połowę braków danych: brak.

Wizualną prezentację braków danych stanowi Rysunek nr 1.

Wykres nr 1

Graficzna prezentacja braków danych występujących w zakresie zmiennych; Płeć, Edukacja, Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu .



Nota. Rysunek nr 1 Graficzna prezentacja braków danych występujących w zakresie zmiennych; Płeć, Edukacja, Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu . Nota : Wykres przedstawia następujące zmienne: Zm1 - Płeć, Zm2 - Edukacja, Zm3 - Wiek, Zm4 - Obawy, Zm5 - Cele, Zm6 - Łatwość, Zm7 - Użyteczność, Zm8 - Intencja zakupu.

Rozkład zmiennych jakościowych bez uwzględniania braków danych

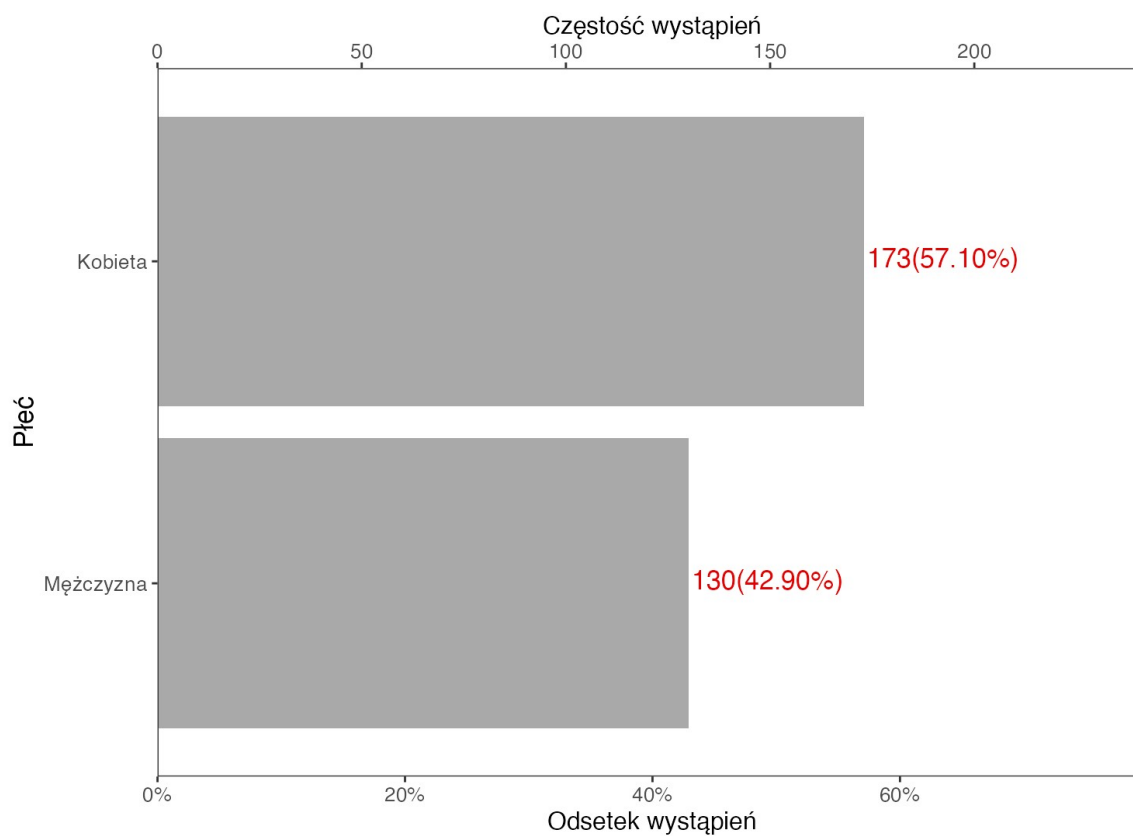
Analiza dla zmiennej Płeć wykazała, że liczebność jej poszczególnych poziomów Kobieta, Mężczyzna wynosiła odpowiednio 173, 130 (ich odsetek wynosił odpowiednio 57.10, 42.90), co stanowiło sumę 303 wystąpień (100%).

Analiza dla zmiennej Edukacja wykazała, że liczebność jej poszczególnych poziomów Wyższe, Średnie, Podstawowe wynosiła odpowiednio 171, 130, 2 (ich odsetek wynosił odpowiednio 56.44, 42.90, 0.66), co stanowiło sumę 303 wystąpień (100%).

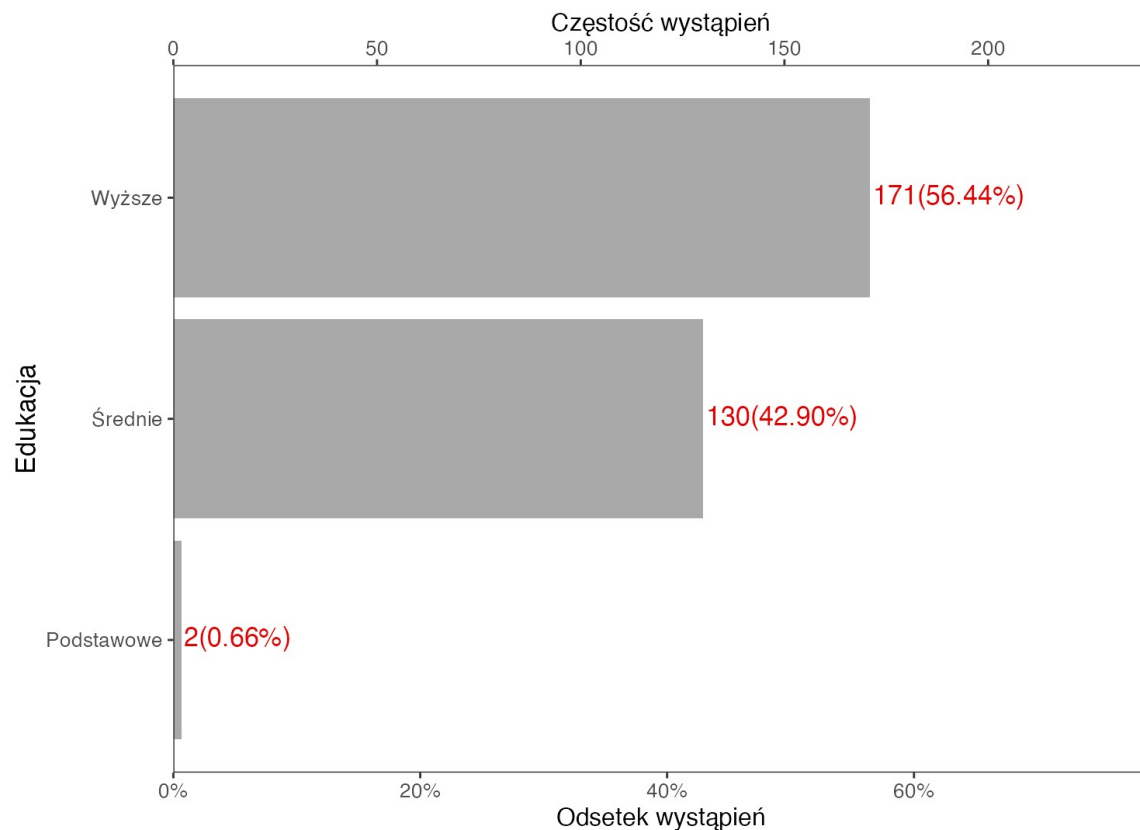
Otrzymane wyniki przedstawiają Rysunki od nr 2 do nr 5.

Wykres nr 2

Częstość występowania wartości zmiennej Płeć (bez uwzględniania braków danych)



Wykres nr 3



Nota. n (%); n = liczebność w podzbiorze. Rysunek nr 2 Częstość występowania wartości zmiennej Edukacja (bez uwzględniania braków danych)

Nota. n (%); n = liczebność w podzbiorze.

[1] “”

Statystyki opisowe i testy normalności analizowanych zmiennych

Analiza normalności rozkładu zmiennych: Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu.

W celu weryfikacji założeń dotyczących normalności rozkładów wyników analizowanych zmiennych przeprowadzono dwa testy statystyczne, był to test Shapiro - Wilka (stosowany dla małych próbek badawczych $N < 50$ lub $N < 100$) (Royston, 1982) oraz test Kolmogorova Smirnova z poprawką Liliefors'a (stosowany dla dużych próbek $N > 50$ lub $N > 100$) (Dallal i Wilkinson, 1986). Zdecydowano się na taką analizę ze względu na brak

jednoznacznych wytycznych dotyczących stosowania obu testów przy danej liczebności badanych obserwacji.

W celu weryfikacji normalności rozkładów zmiennych: Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu, przeprowadzono serię analiz weryfikujących podobieństwo rozkładu z próby do teoretycznego rozkładu normalnego testem Shapiro-Wilka (Royston, 1982). Analiza wykazała, że:

Rozkład wyników zmiennej Wiek był istotnie różny od rozkładu normalnego $SW = 0.85; p < 0.001$

Rozkład wyników zmiennej Obawy był istotnie różny od rozkładu normalnego $SW = 0.97; p < 0.001$

Rozkład wyników zmiennej Cele był istotnie różny od rozkładu normalnego $SW = 0.96; p < 0.001$

Rozkład wyników zmiennej Łatwość był istotnie różny od rozkładu normalnego $SW = 0.95; p < 0.001$

Rozkład wyników zmiennej Użyteczność był istotnie różny od rozkładu normalnego $SW = 0.98; p < 0.001$

Rozkład wyników zmiennej Intencja zakupu był istotnie różny od rozkładu normalnego $SW = 0.97; p < 0.001$

W celu dodatkowej weryfikacji normalności rozkładów zmiennych: Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu, przeprowadzono serię analiz weryfikujących podobieństwo rozkładu z próby do teoretycznego rozkładu normalnego testem Kolmogorova-Smirnova z poprawką Liliefors'a (Dallal i Wilkinson, 1986). Analiza wykazała, że:

Rozkład wyników zmiennej Wiek był istotnie różny od rozkładu normalnego $KS = 0.22; p < 0.001$

Rozkład wyników zmiennej Obawy był istotnie różny od rozkładu normalnego $KS = 0.12; p < 0.001$

Rozkład wyników zmiennej Cele był istotnie różny od rozkładu normalnego $KS = 0.12; p < 0.001$

Rozkład wyników zmiennej Łatwość był istotnie różny od rozkładu normalnego $KS = 0.17$; $p < 0.001$

Rozkład wyników zmiennej Użyteczność był istotnie różny od rozkładu normalnego $KS = 0.10$; $p < 0.001$

Rozkład wyników zmiennej Intencja zakupu był istotnie różny od rozkładu normalnego $KS = 0.08$; $p < 0.001$

Wyniki analiz przedstawia Tabela nr 2.

Tabela nr 2

Wyniki analizy rozkładu normalności i statystyk opisowych dla zmiennych: Wiek, Obawy, Cele, Łatwość, Użyteczność, Intencja zakupu

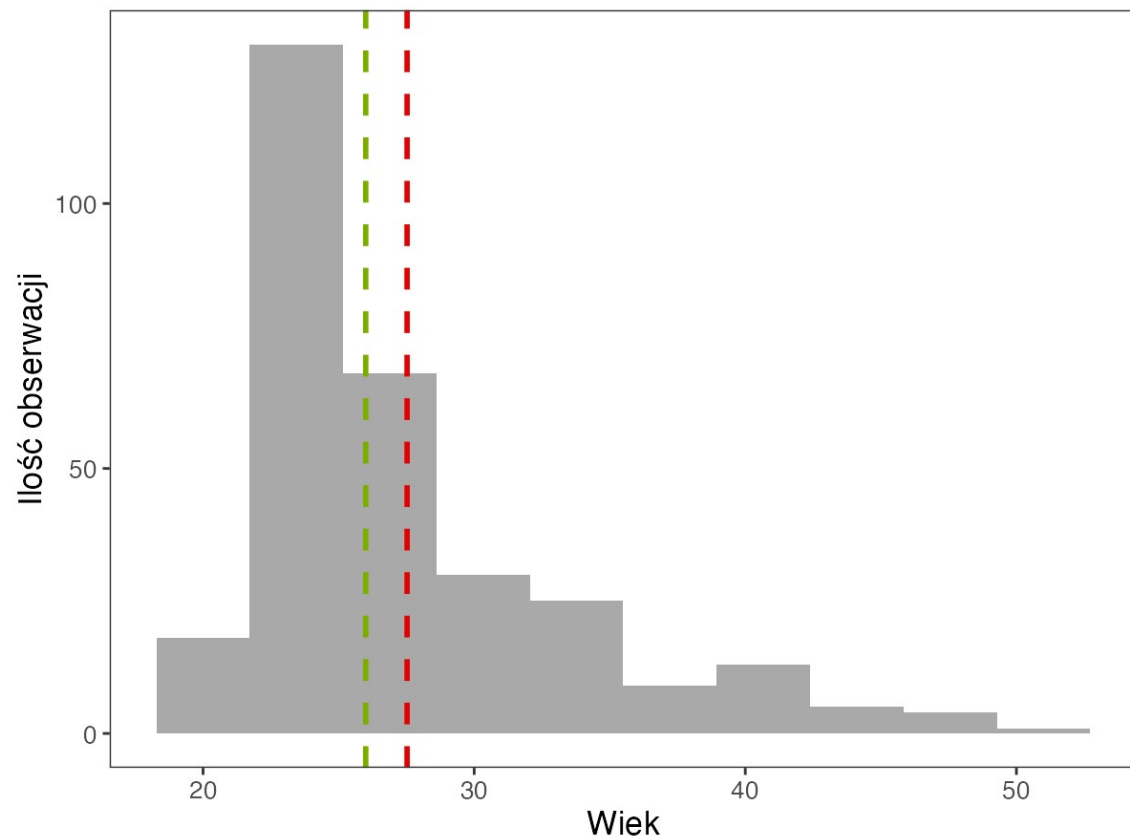
Zmienna	N	Min	Max	M	SD	SE	Me	Miary symetrii rozkładu		test KS		test SW	
								Skośność	Kurtoza	KS	p	SW	p
Wiek	303	20.00	51.00	27.52	5.87	0.34	26.00	1.47	1.86	0.22	< 0.001	0.85	< 0.001
Obawy	303	5.00	20.00	13.93	3.56	0.20	14.00	-0.31	-0.57	0.12	< 0.001	0.97	< 0.001
Cele	303	5.00	20.00	10.86	3.70	0.21	10.00	0.37	-0.23	0.12	< 0.001	0.96	< 0.001
Łatwość	303	4.00	20.00	14.45	3.17	0.18	15.00	-0.52	0.51	0.17	< 0.001	0.95	< 0.001
Użyteczność	303	4.00	20.00	12.97	3.57	0.21	13.00	-0.32	-0.05	0.10	< 0.001	0.98	< 0.001
Intencja zakupu	303	4.00	20.00	11.05	4.14	0.24	11.00	0.03	-0.68	0.08	< 0.001	0.97	< 0.001

Nota. N = Liczebność; Min = Wartość minimalna; Max = Wartość maksymalna; M = Średnia arytmetyczna; SD = Odchylenie standardowe; SE = Błąd standardowy średniej; Me = Mediana; p = Istotność statystyczna. Wyniki testu KS (Kolmogorov - Smirnov) i SW (Shapiro - Wilk), których $p < 0.05$ dla danej zmiennej, wskazują na istotną różnicę między rozkładem wyników w próbie a teoretycznym rozkładem normalnym.

Histogramy

Wykres nr 4

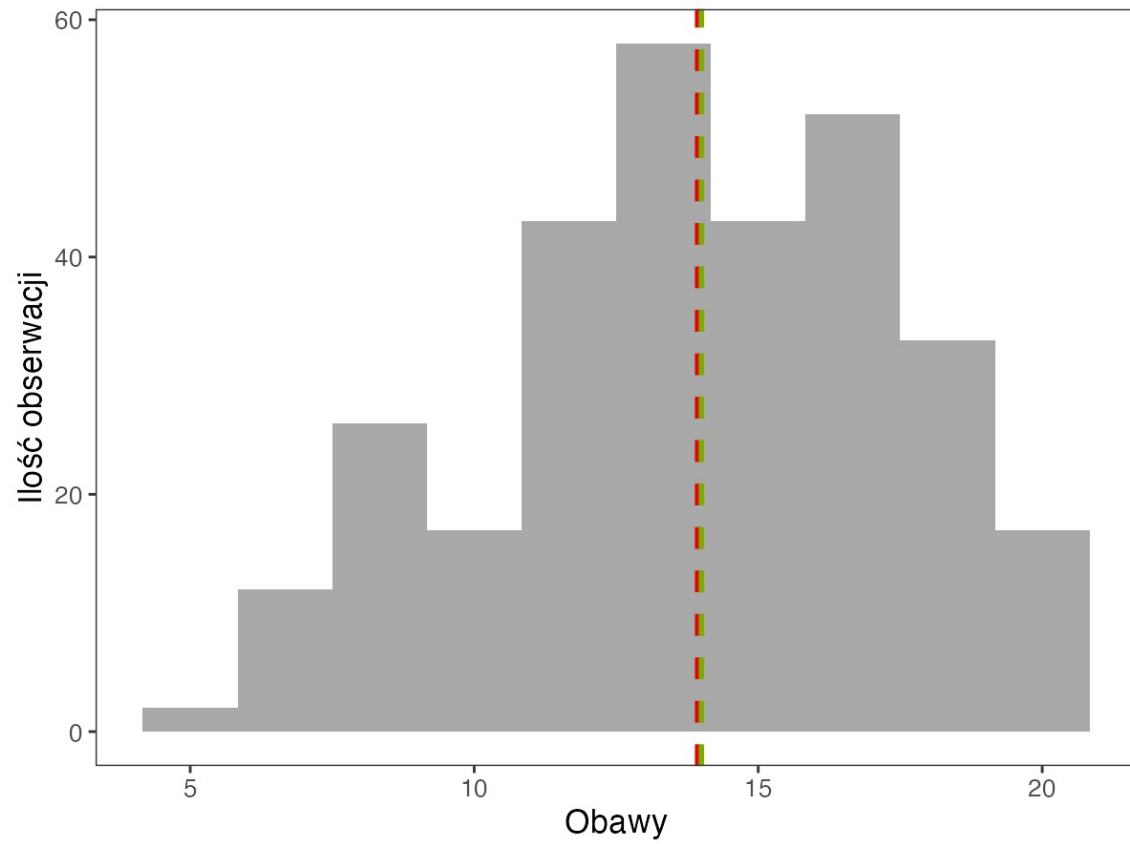
Rozkład wartości zmiennej Wiek



Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Wykres nr 5

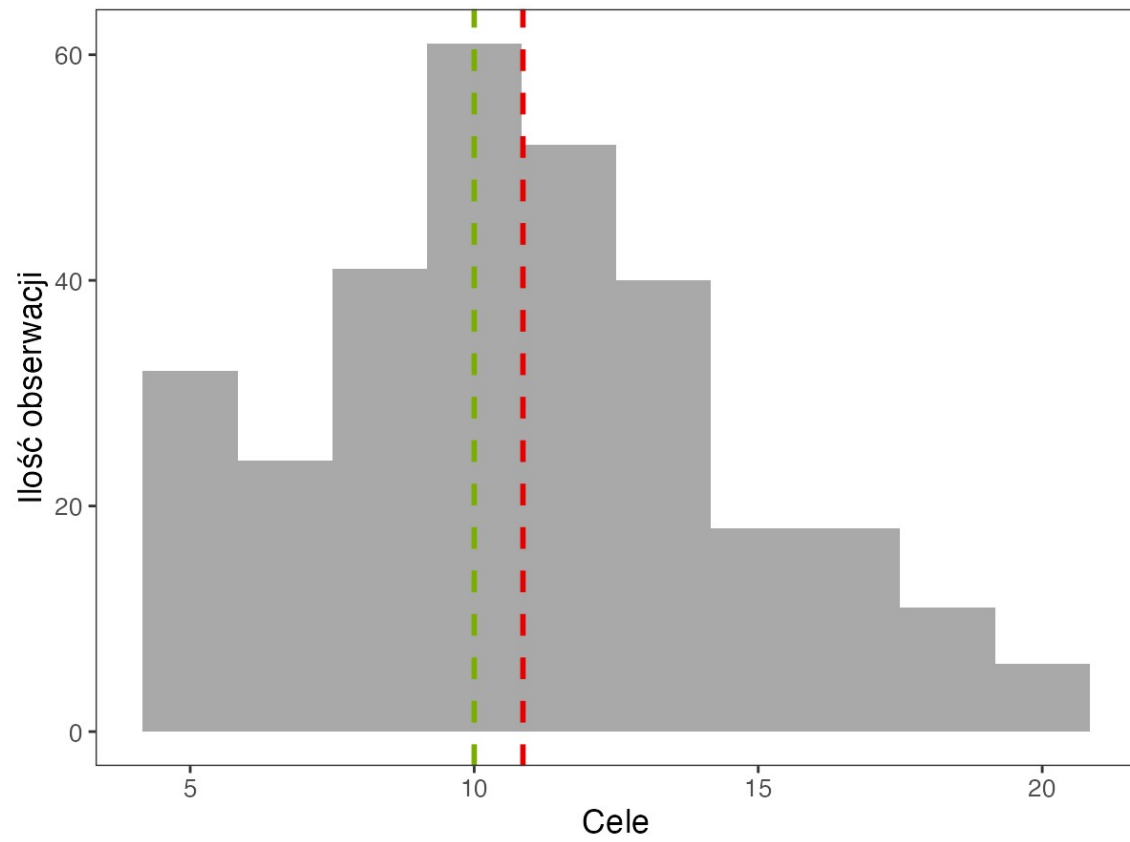
Rozkład wartości zmiennej Obawy



Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Wykres nr 6

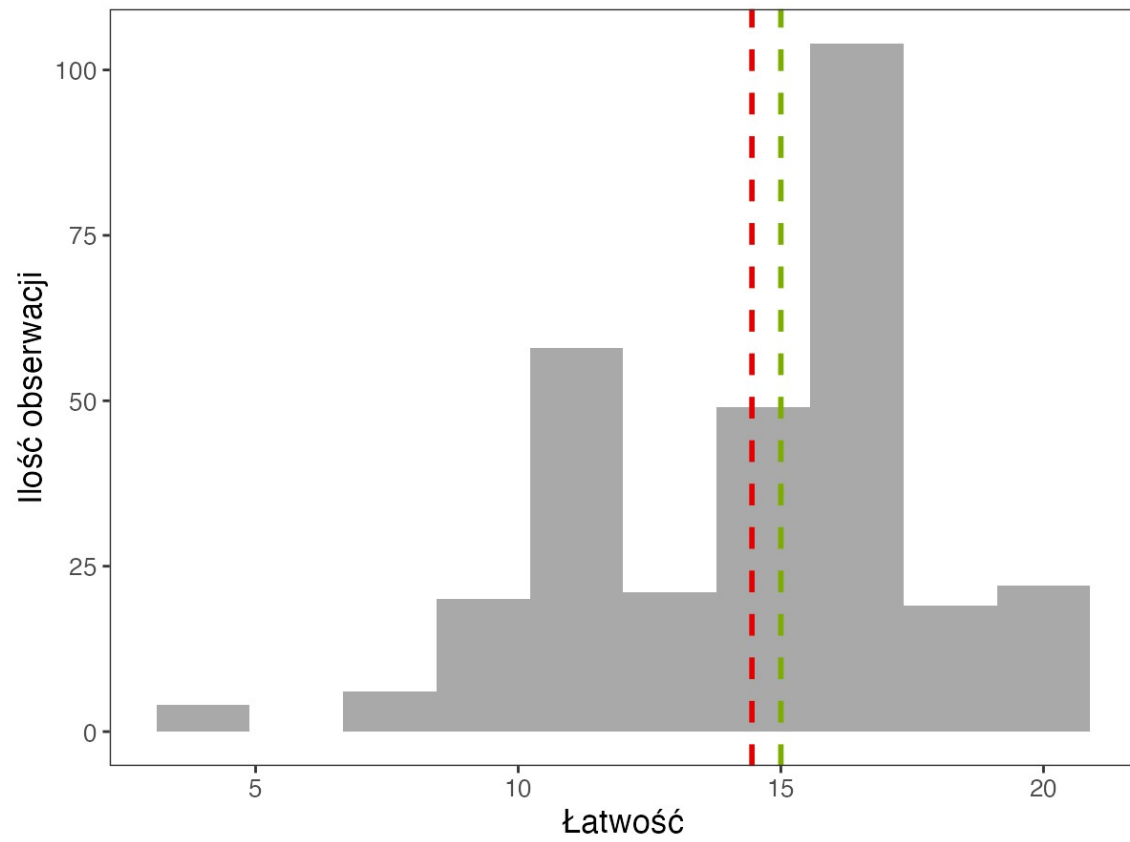
Rozkład wartości zmiennej Cele



Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Wykres nr 7

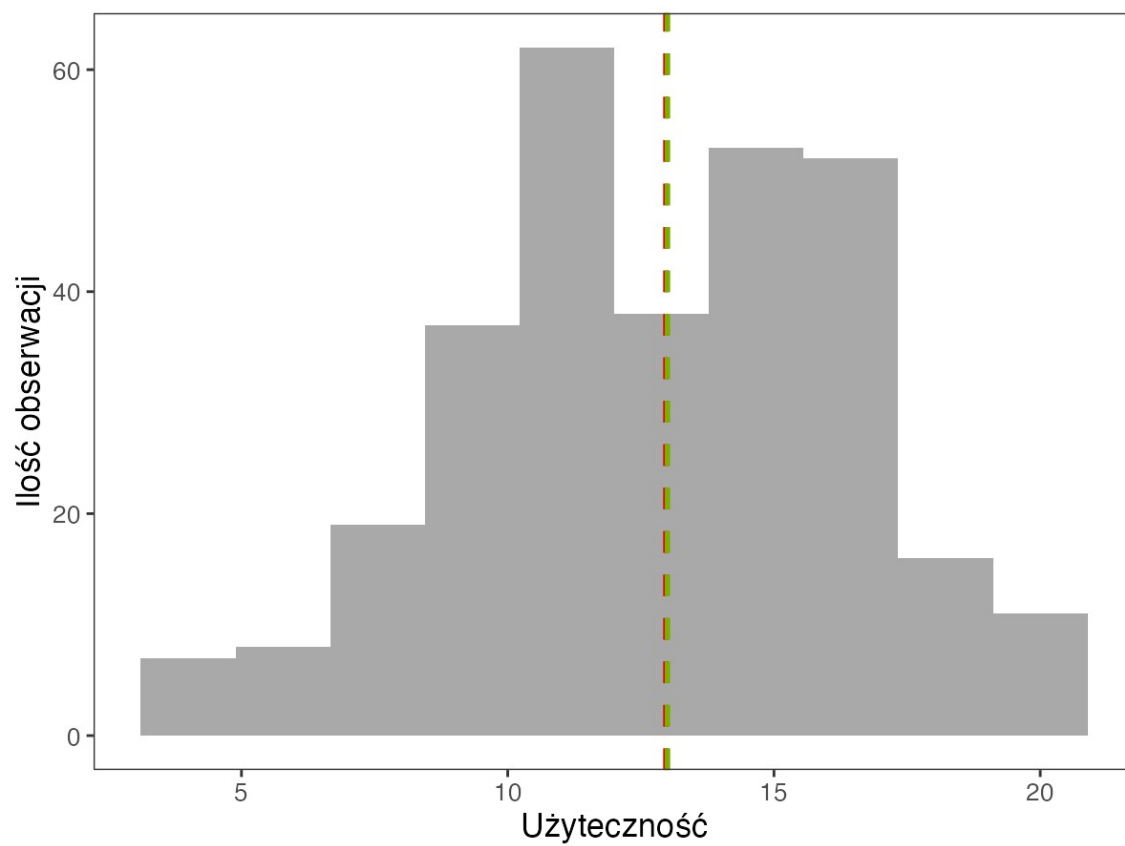
Rozkład wartości zmiennej Łatwość



Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Wykres nr 8

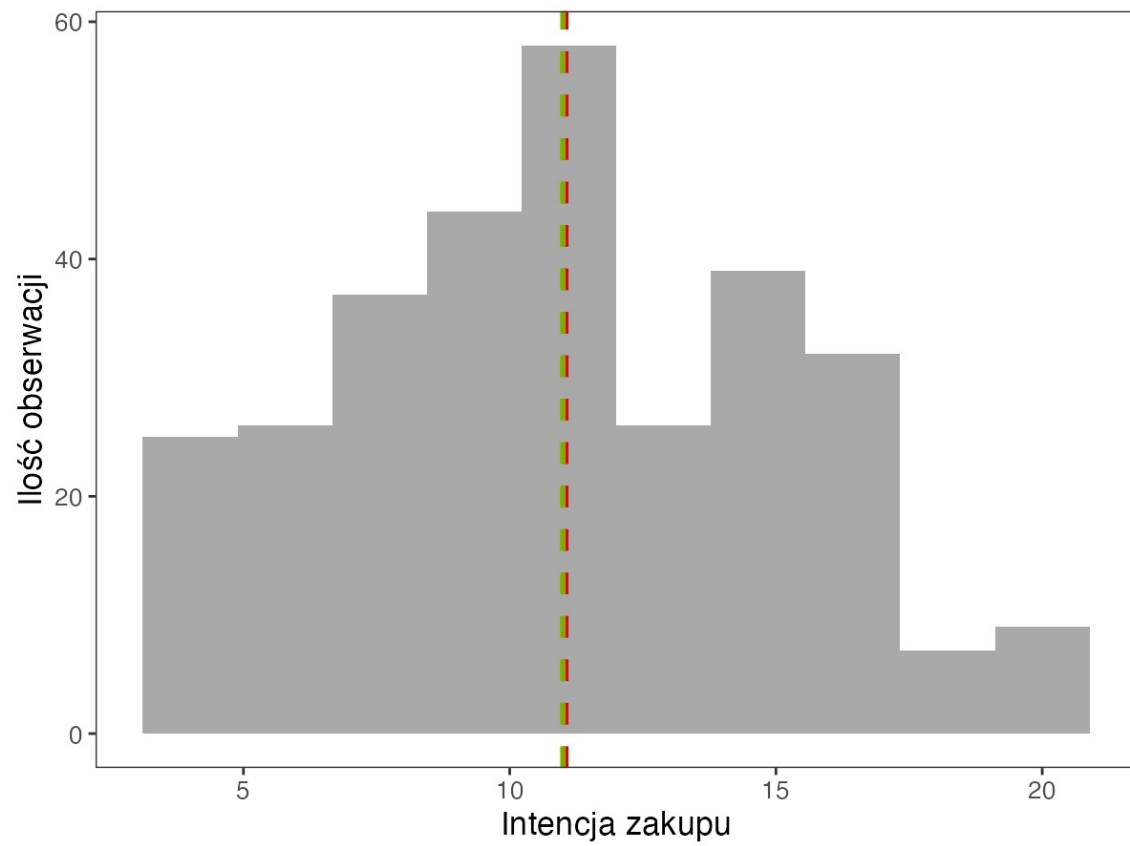
Rozkład wartości zmiennej Użyteczność



Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Wykres nr 9

Rozkład wartości zmiennej Intencja zakupu

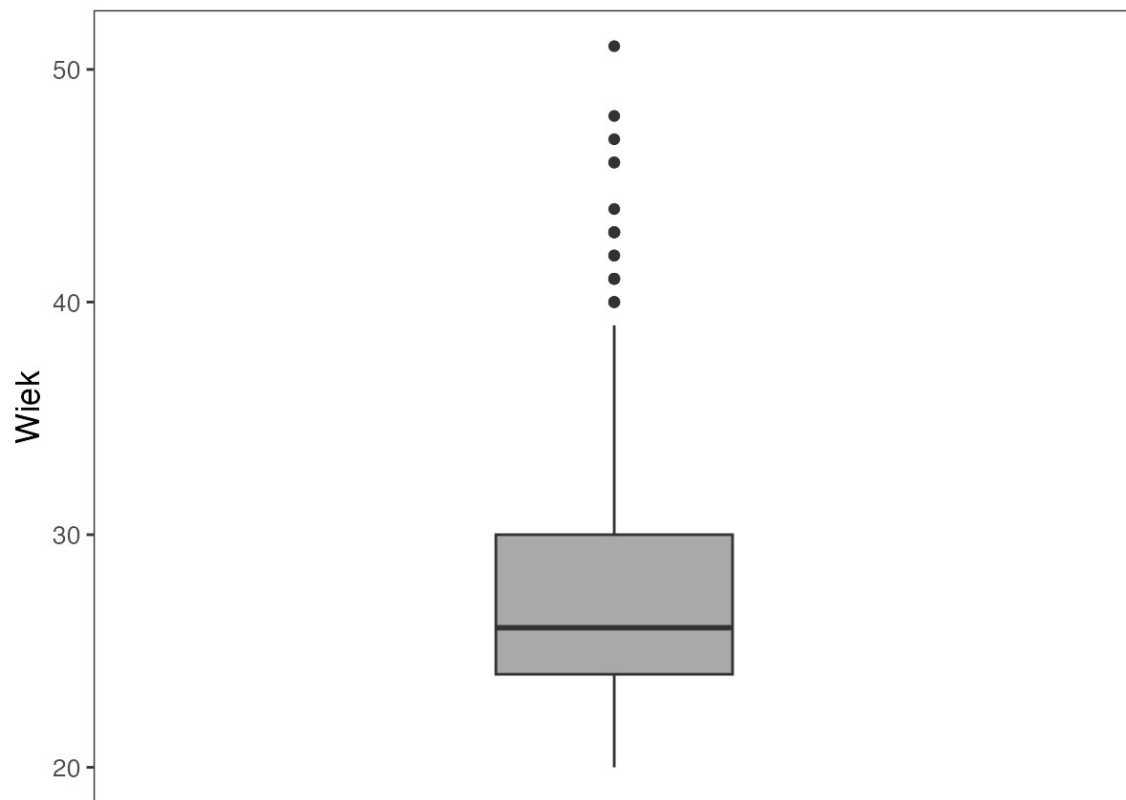


Nota. Kolor czerwony przedstawia Średnią ; Kolor zielony przedstawia Medianę .

Box-ploty

Wykres nr 10

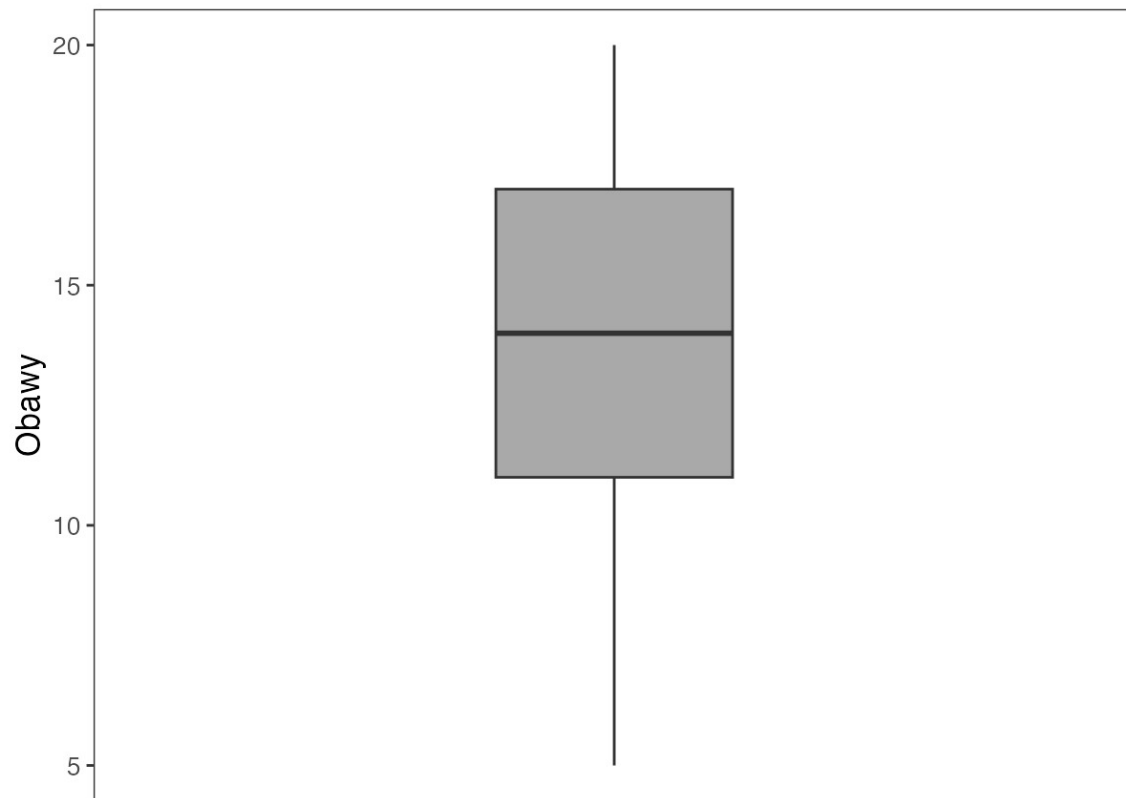
Rozkład wartości zmiennej Wiek



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Wykres nr 11

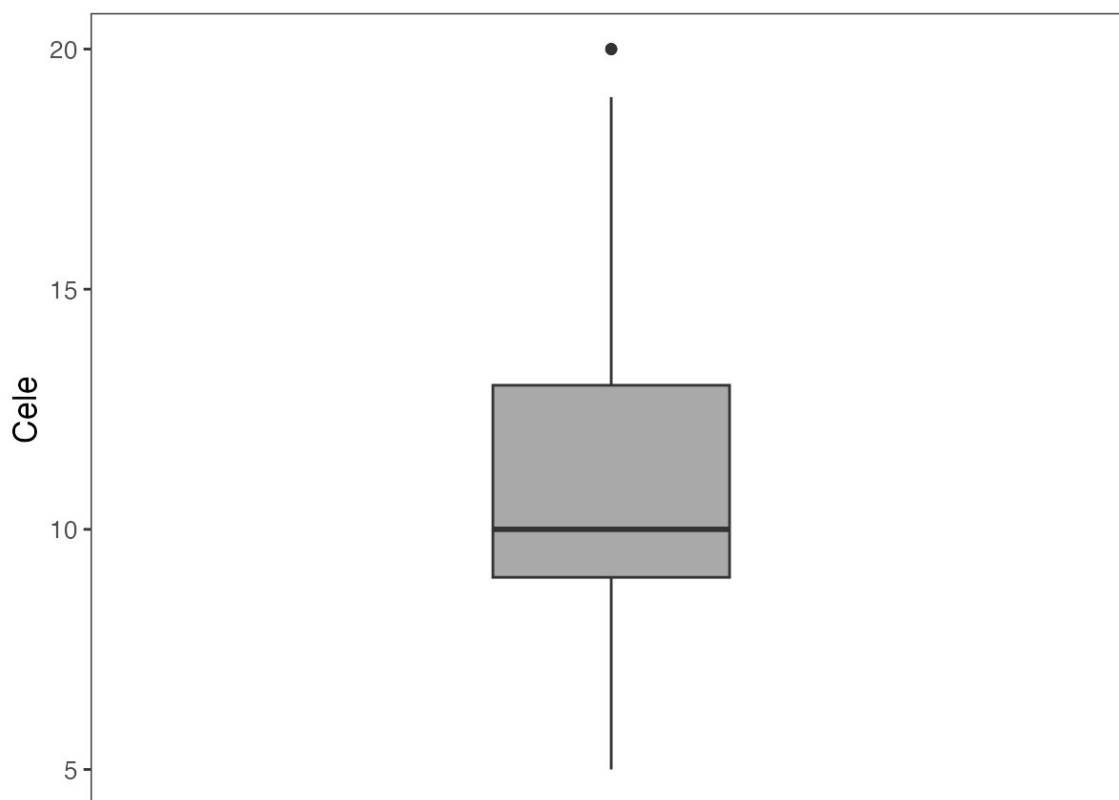
Rozkład wartości zmiennej Obawy



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Wykres nr 12

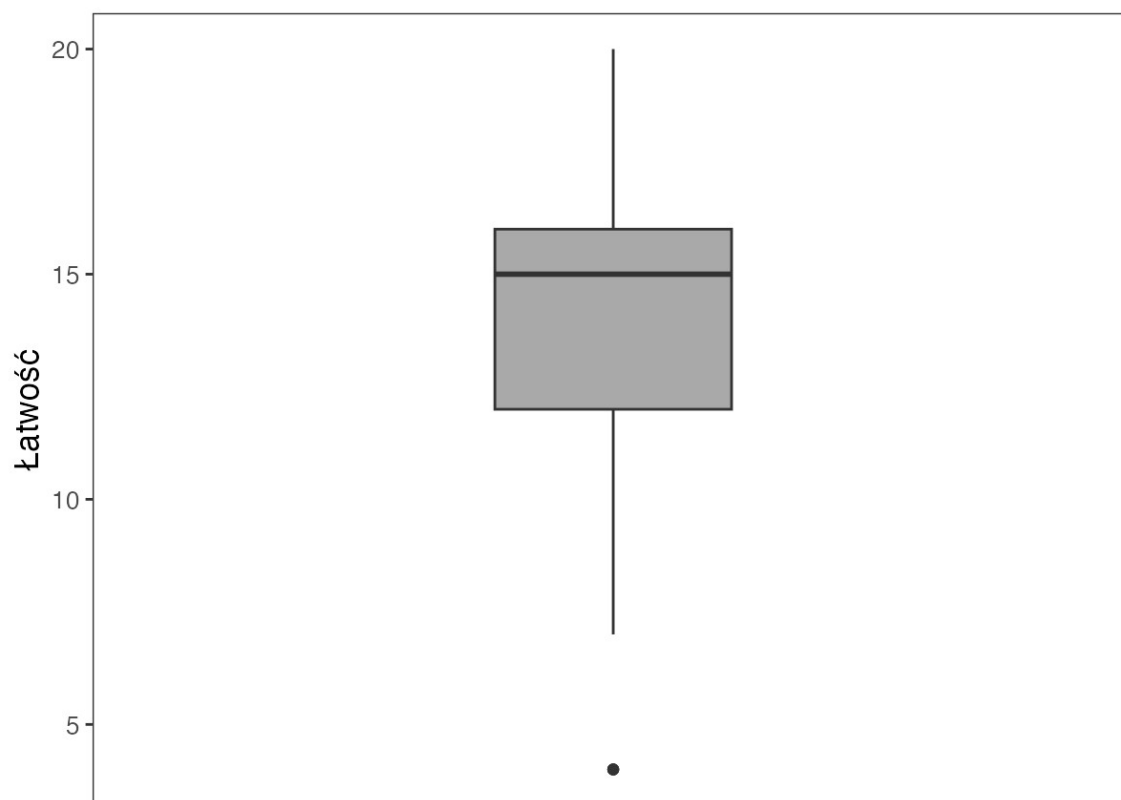
Rozkład wartości zmiennej Cele



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Wykres nr 13

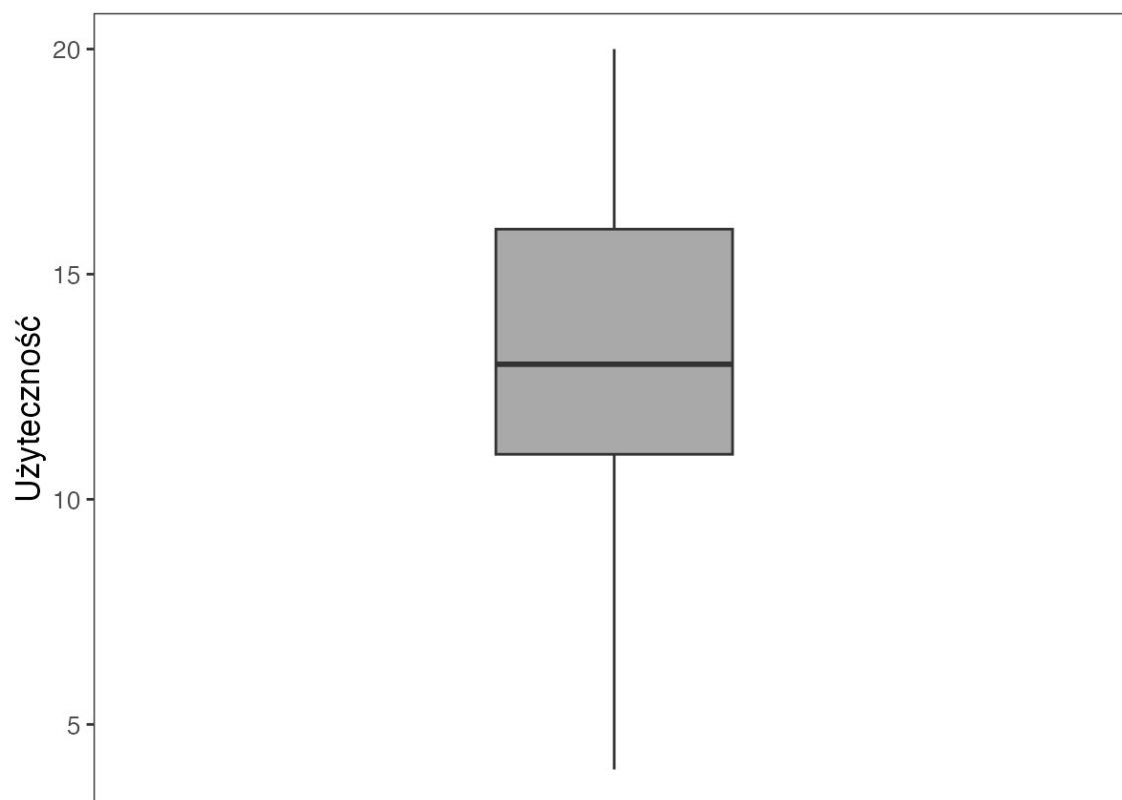
Rozkład wartości zmiennej Łatwość



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Wykres nr 14

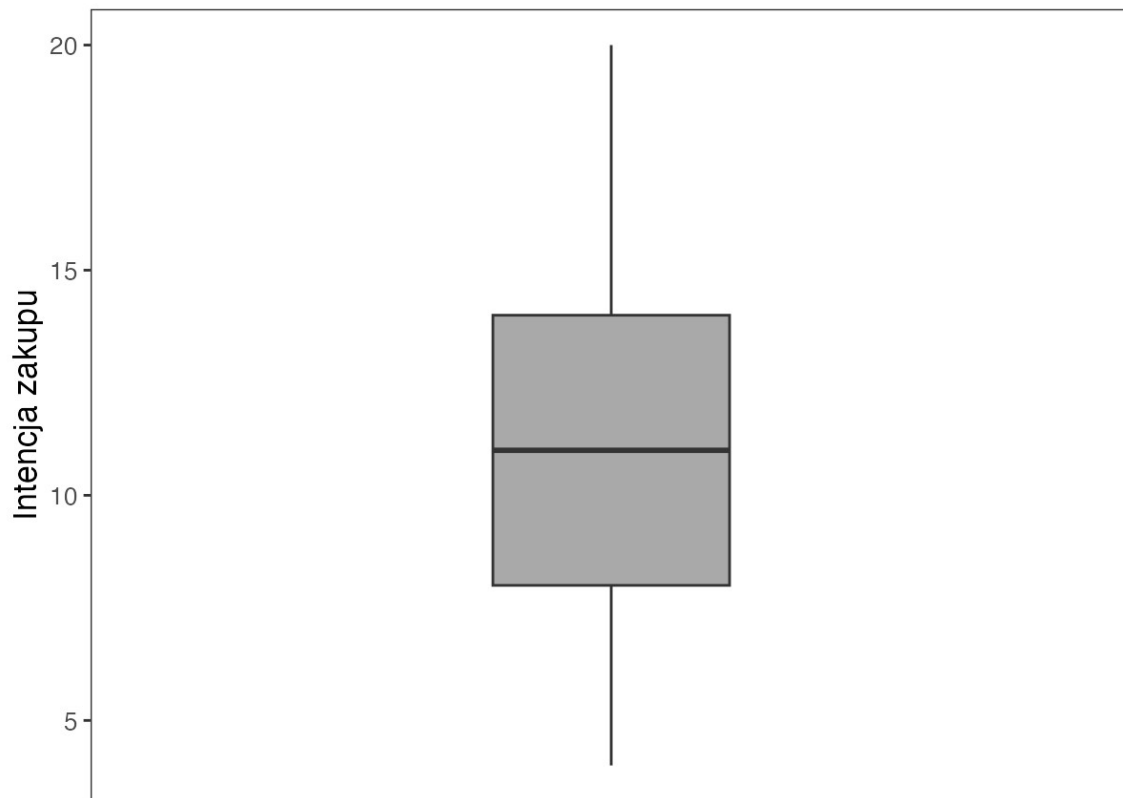
Rozkład wartości zmiennej Użyteczność



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Wykres nr 15

Rozkład wartości zmiennej Intencja zakupu



Nota. Pozioma linia przedstawia Medianę. Pudełko prezentuje Rozstęp międzykwartyłowy IQR. Wąsy sięgają do wartości powyżej/poniżej $1.5 \cdot \text{IQR}$. Kropki poza wąsami oznaczają wartości powyżej/poniżej $1.5 \cdot \text{IQR}$.

Opis przeprowadzonej normalizacji

W celu unormowania rozkładów w analizowanych danych, wykorzystano bibliotekę języka R o nazwie „bestNormalize” (Peterson, 2021; Peterson i inni, 2020), która pozwala na testowanie wielu przekształceń unormalniających rozkład. Logika tej funkcji polega na testowaniu 15 możliwych przekształceń, wraz z jednoczesnym ustaleniem optymalnych parametrów przekształceniowych np. hiperparametru Lambda w przekształceniu Boxa Coxa (Box & Cox, 1964).

Istnieje wiele przypadków, w których badacze mogą chcieć znormalizować zmienną. Oto jeden z najczęstszych:

Często problematyczne założenie o normalności zmiennej zależnej (uwarunkowanego współzmiennymi w modelu) w klasycznym problemie regresji liniowej. Przez lata stosowano wiele metod, aby złagodzić to założenie: uogólnione modele liniowe, regresja kwantylowa, modele przetrwania itp. Jedną z technik, która wciąż jest dość popularna w tym kontekście, jest transformacja danych, aby wyglądały normalnie. Może to być coś tak prostego jak transformacja logarytmiczna tzw. $\log()$ lub coś tak złożonego jak transformacja Yeo-Johnsona.

W celu wyboru metody normalizującej zmienną Wiek wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `orderNorm`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 1.81$. Po transformacji rozkład zmiennej stał się bardziej symetryczny (skośność: 0.06 vs 1.47 przed). W celu wyboru metody normalizującej zmienną Obawy wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `center_scale`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 1.61$. Po transformacji rozkład zmiennej stał się bardziej symetryczny (skośność: -0.31 vs -0.31 przed). W celu wyboru metody normalizującej zmienną Cele wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `sqrt_x`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 1.81$. Po transformacji rozkład zmiennej stał się bardziej symetryczny (skośność: -0.04 vs 0.37 przed). W celu wyboru metody normalizującej zmienną Łatwość wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `boxcox`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 3.21$. Po transformacji rozkład zmiennej stał się bardziej symetryczny (skośność: -0.04 vs -0.52 przed). W celu wyboru metody normalizującej zmienną Użyteczność wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `center_scale`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość statystyki dopasowania, $P/df = 1.53$. Po transformacji rozkład zmiennej stał się bardziej symetryczny (skośność: -0.32 vs -0.32 przed). W celu wyboru metody normalizującej zmienną Intencja zakupu wykonano serię przekształceń. Analiza wykazała, że najlepsze właściwości normalizacyjne zwracała technika `orderNorm`. W porównaniu do innych popularnych przekształceń, metoda ta zwróciła najniższą wartość

statystyki dopasowania, $P/df = 1.39$. Po transformacji rozkład zmiennej stał się mniej symetryczny (skośność: 0.08 vs 0.03 przed).

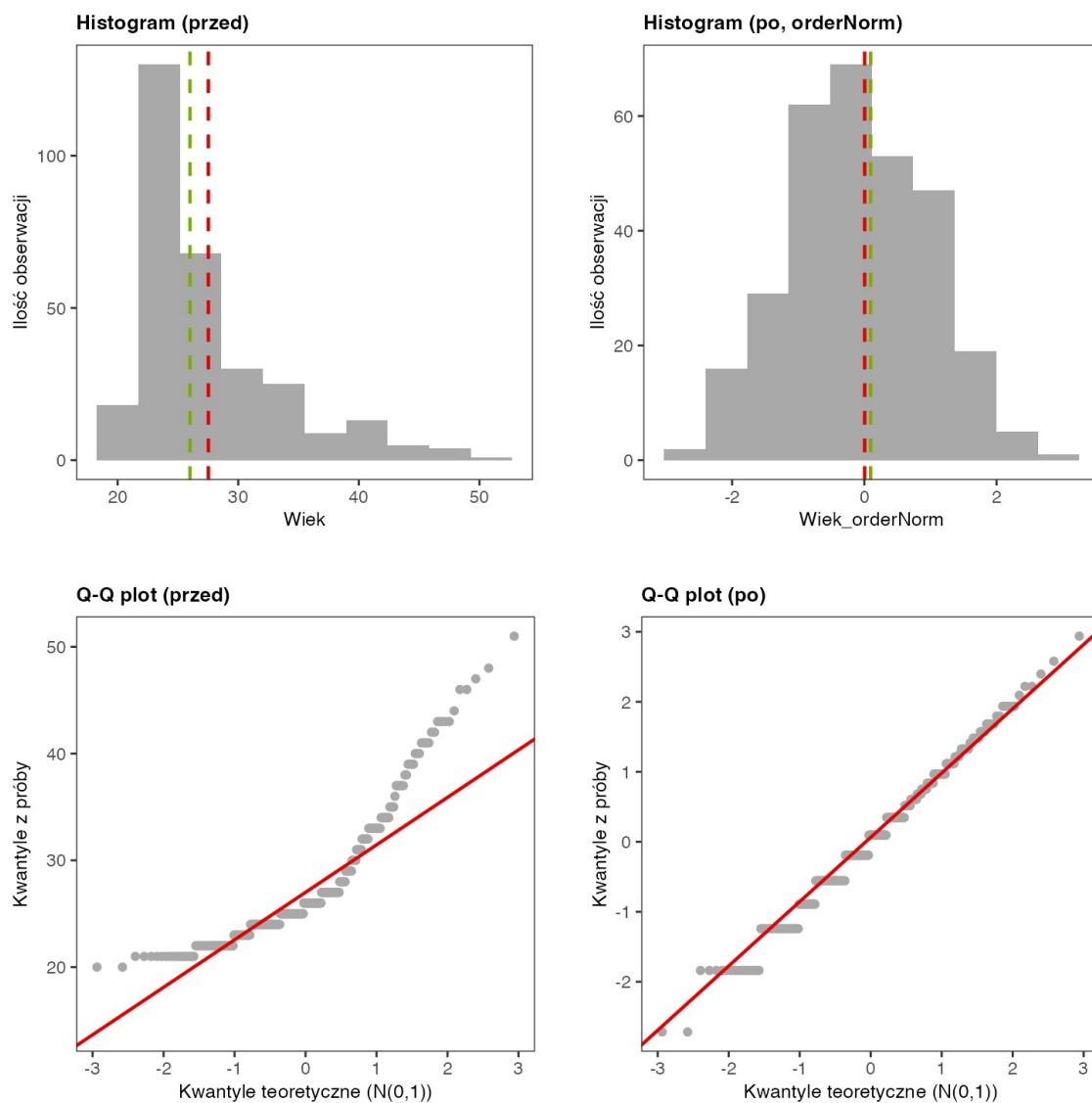
Wyniki przekształceń zostały zapisane w bazie danych

baza_trans.xlsx

Porównanie rozkładów przed i po normalizacji

Wykres nr 16

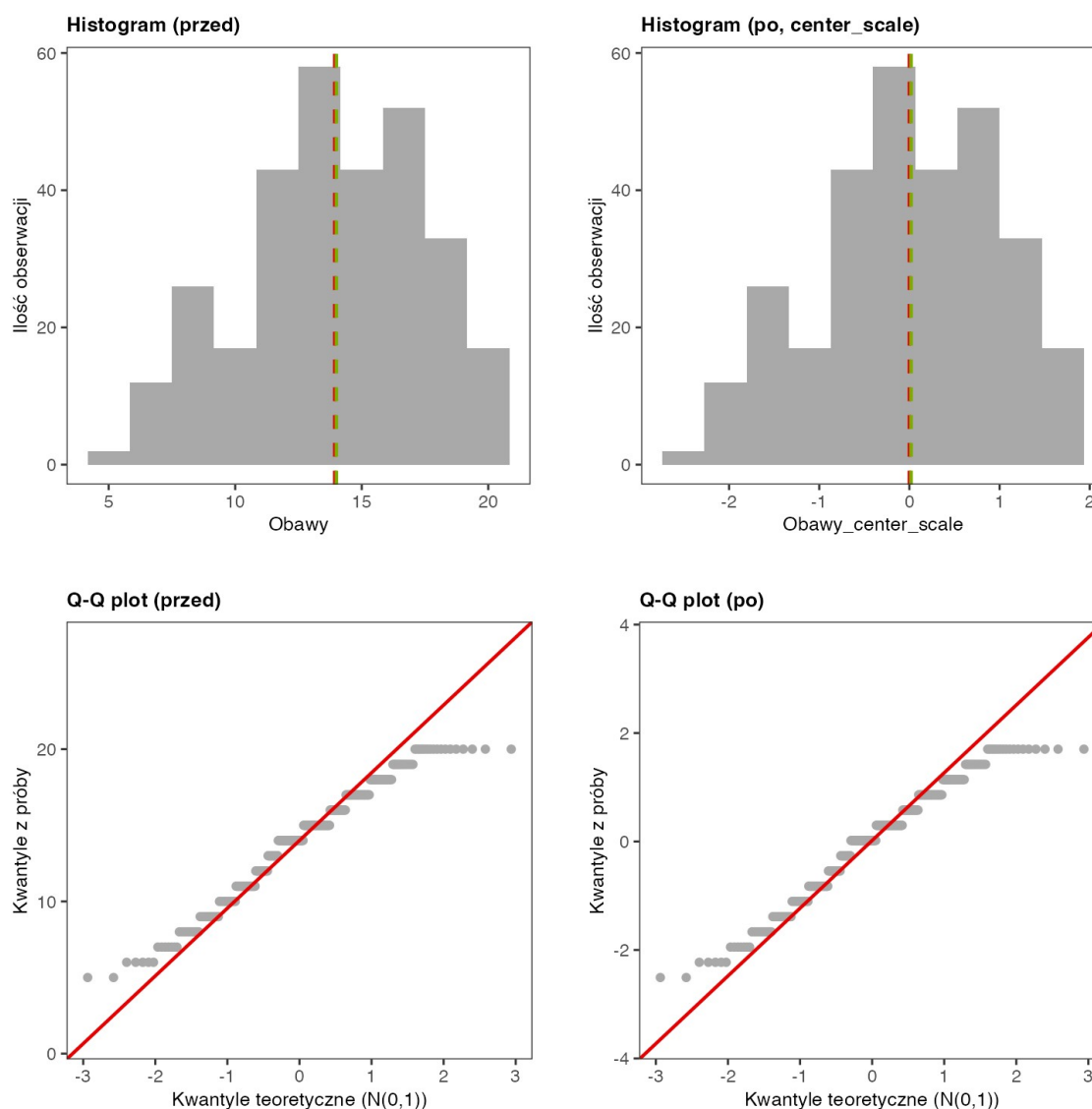
Aby umożliwić wizualną ocenę efektu zastosowanych przekształceń, dla każdej zmiennej numerycznej zestawiono histogram rozkładu oryginalnego (przed normalizacją) z histogramem rozkładu po normalizacji oraz wykresy kwantyl-kwantyl (Q-Q) względem teoretycznego rozkładu normalnego. Wizualizacje wyników przedstawia seria rysunków od nr 10 do nr 21. Rysunek nr 17 Porównanie rozkładu wartości zmiennej Wiek przed i po normalizacji (metoda: orderNorm).



Nota. Aby umożliwić wizualną ocenę efektu zastosowanych przekształceń, dla każdej zmiennej numerycznej zestawiono histogram rozkładu oryginalnego (przed normalizacją) z histogramem rozkładu po normalizacji oraz wykresy kwantyl-kwantyl (Q-Q) względem teoretycznego rozkładu normalnego. Wizualizacje wyników przedstawia seria rysunków od nr 1 do nr 6. Rysunek nr 1 Porównanie rozkładu wartości zmiennej Wiek przed i po normalizacji (metoda: orderNorm). *Nota :* Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Wykres nr 17

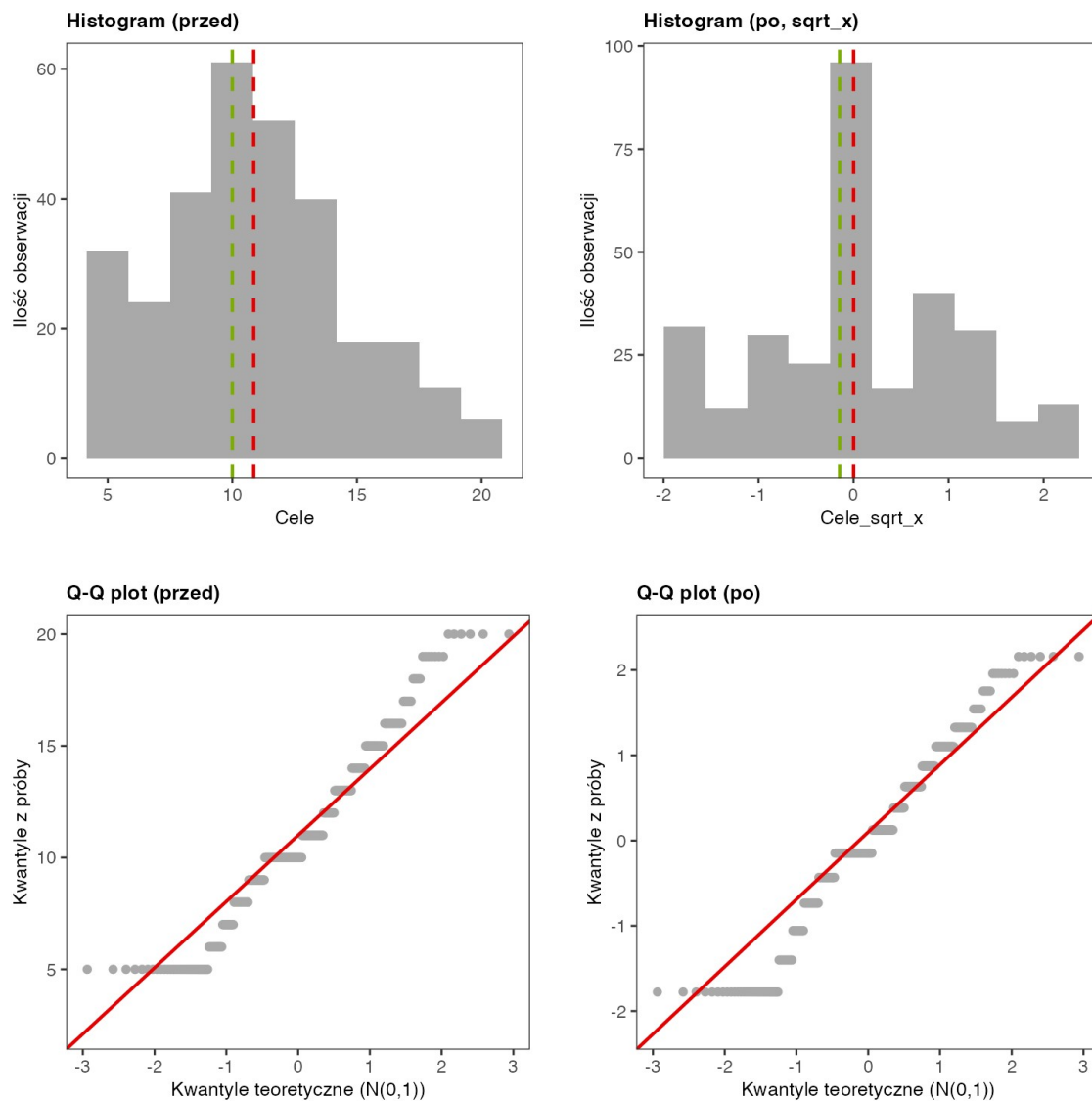
Porównanie rozkładu wartości zmiennej Obawy przed i po normalizacji (metoda: center_scale).



Nota. Rysunek nr 2 Porównanie rozkładu wartości zmiennej Obawy przed i po normalizacji (metoda: center_scale). *Nota :* Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Wykres nr 18

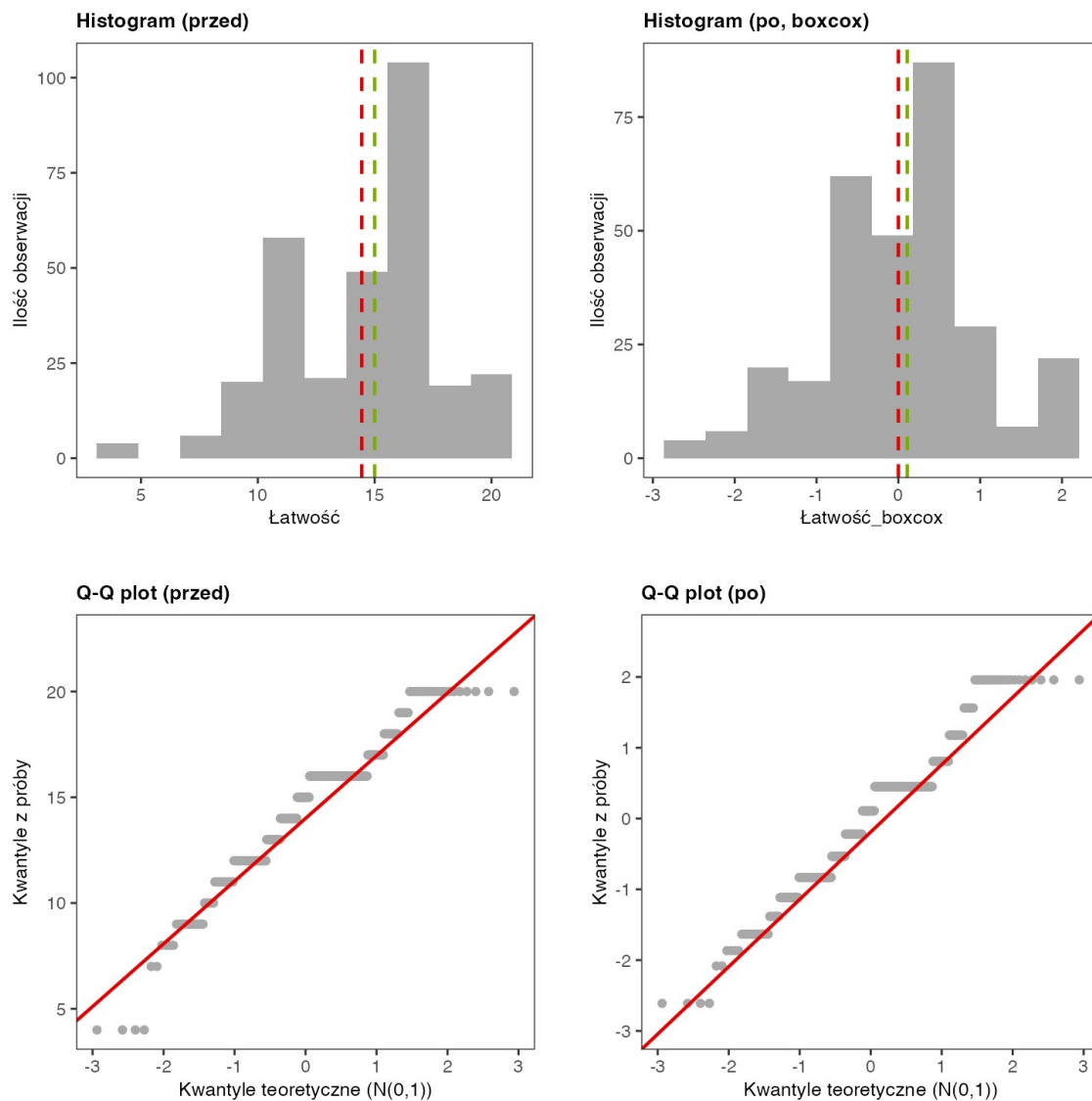
Porównanie rozkładu wartości zmiennej *Cele* przed i po normalizacji (metoda: *sqrt_x*).



Nota. Rysunek nr 3 Porównanie rozkładu wartości zmiennej *Cele* przed i po normalizacji (metoda: *sqrt_x*). *Nota :* Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Wykres nr 19

Porównanie rozkładu wartości zmiennej Łatwość przed i po normalizacji (metoda: boxcox).

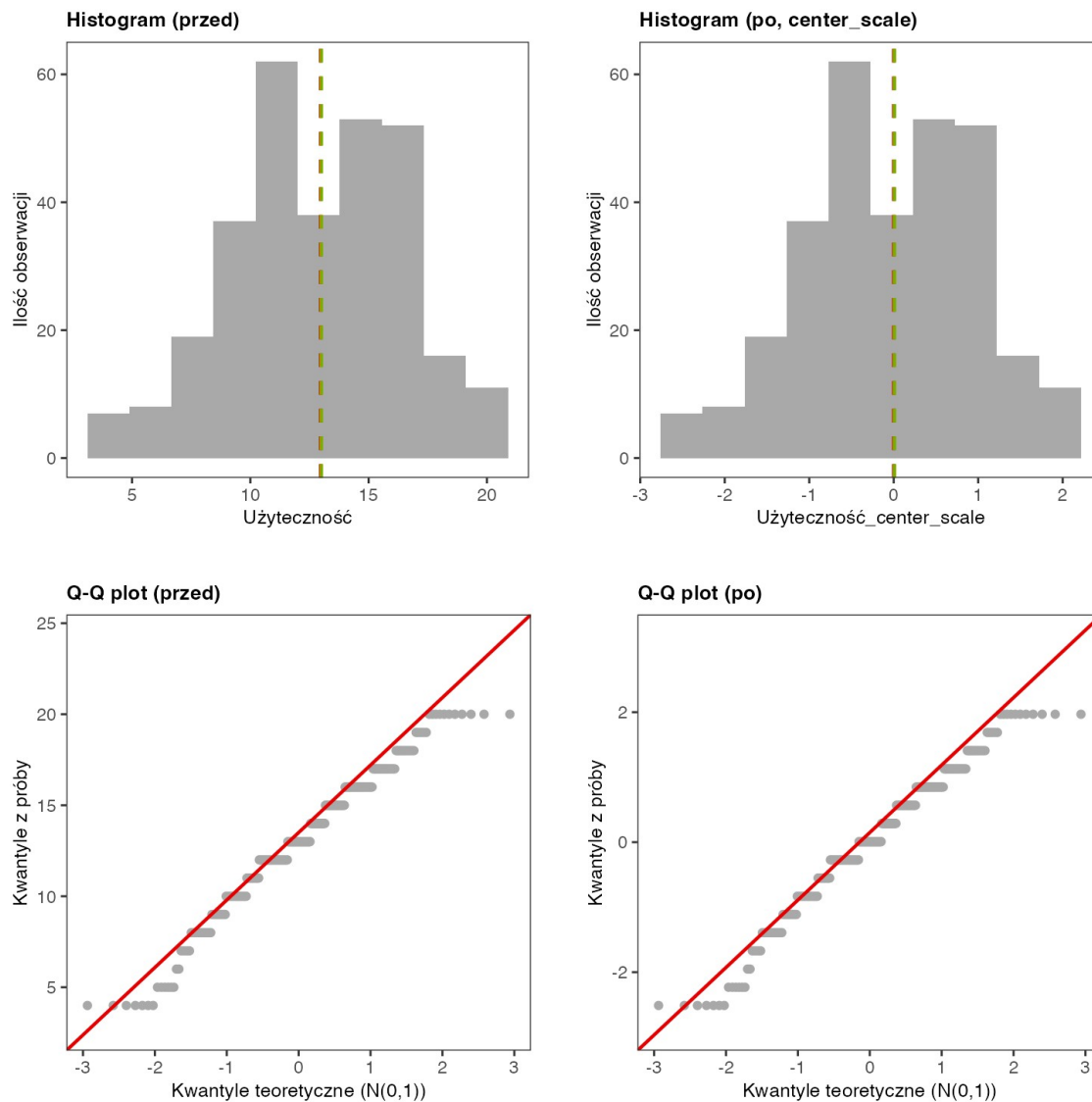


Nota. Rysunek nr 4 Porównanie rozkładu wartości zmiennej Łatwość przed i po normalizacji (metoda: boxcox).

Nota : Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Wykres nr 20

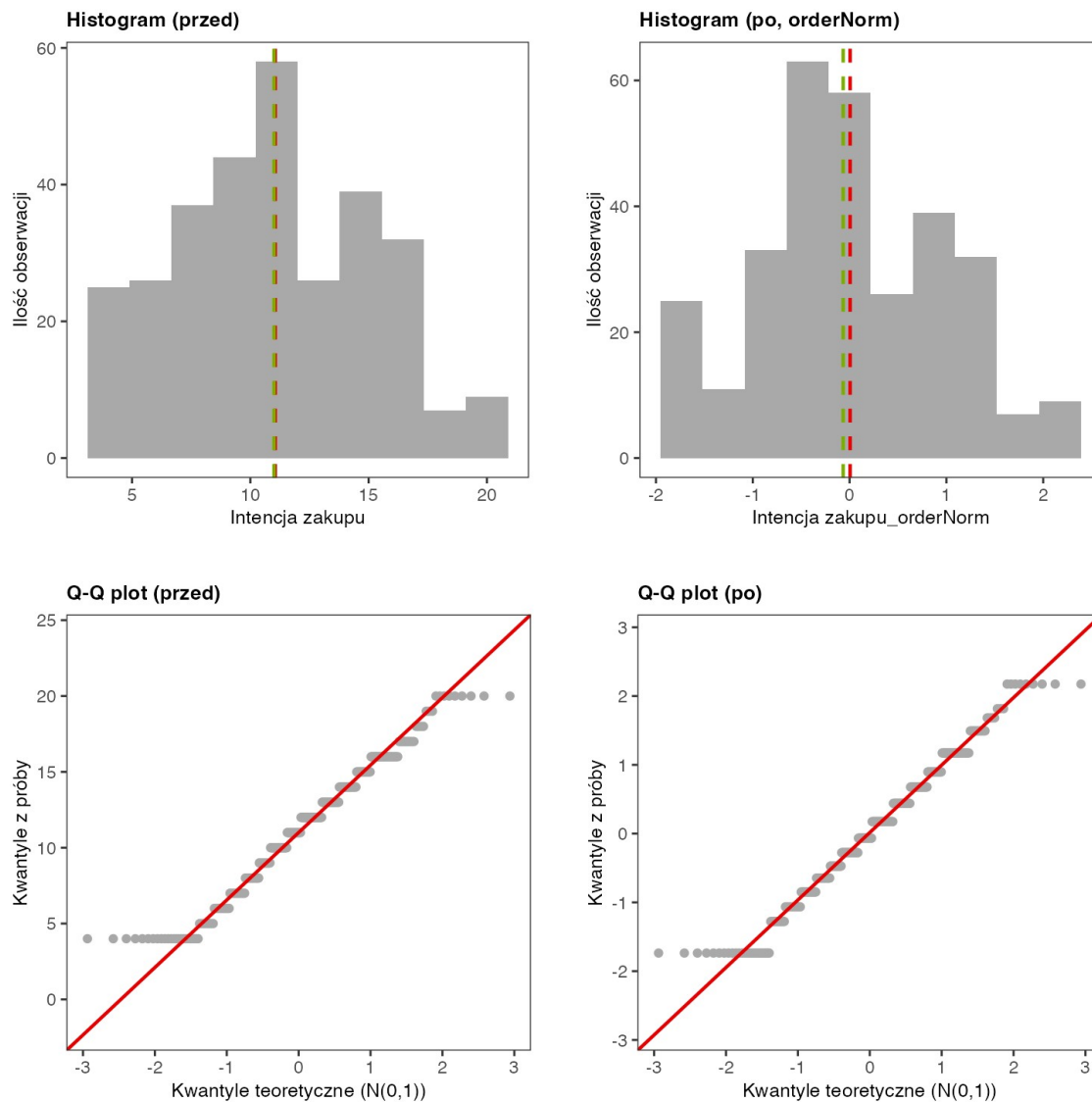
Porównanie rozkładu wartości zmiennej *Użyteczność* przed i po normalizacji (metoda: *center_scale*).



Nota. Rysunek nr 5 Porównanie rozkładu wartości zmiennej *Użyteczność* przed i po normalizacji (metoda: *center_scale*). *Nota :* Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Wykres nr 21

Porównanie rozkładu wartości zmiennej *Intencja zakupu* przed i po normalizacji (metoda: *orderNorm*).



Nota. Rysunek nr 6 Porównanie rozkładu wartości zmiennej *Intencja zakupu* przed i po normalizacji (metoda: *orderNorm*). *Nota :* Na histogramach kolor czerwony przedstawia Średnią ; kolor zielony przedstawia Medianę . Na Q-Q plot czerwona linia odpowiada idealnemu dopasowaniu do rozkładu normalnego — im bliżej punktów do tej linii, tym rozkład bardziej zbliżony do normalnego.

Tabela nr 3

Statystyki opisowe oraz wyniki testów normalności (Shapiro-Wilka i Kolmogorova-Smirnova z poprawką Lilieforce'a) przed i po normalizacji

Zmienna	Technika	N	Przed normalizacją						Po normalizacji					
			M	SD	Skoś.	Kurt.	p SW	p KS	M	SD	Skoś.	Kurt.	p SW	p KS
Wiek	orderNorm	303	27.52	5.87	1.47	1.86	< 0.001	< 0.001	0.00	0.99	0.06	-0.12	0.025	0.003
Obawy	center_scale	303	13.93	3.56	-0.31	-0.57	< 0.001	< 0.001	0.00	1.00	-0.31	-0.57	< 0.001	< 0.001
Cele	sqrt_x	303	10.86	3.70	0.37	-0.23	< 0.001	< 0.001	0.00	1.00	-0.04	-0.44	< 0.001	< 0.001
Łatwość	boxcox	303	14.45	3.17	-0.52	0.51	< 0.001	< 0.001	0.00	1.00	-0.04	-0.20	< 0.001	< 0.001
Użyteczność	center_scale	303	12.97	3.57	-0.32	-0.05	< 0.001	< 0.001	0.00	1.00	-0.32	-0.05	< 0.001	< 0.001
Intencja zakupu	orderNorm	303	11.05	4.14	0.03	-0.68	< 0.001	< 0.001	0.01	0.97	0.08	-0.48	< 0.001	0.021

Nota. Skoś. = skośność (typ 2); Kurt. = kurtóza (typ 2); p SW = istotność testu Shapiro-Wilka; p KS = istotność testu Kolmogorova-Smirnova z poprawką Lilieforce'a.