

Badanie statystyczne wpływu różnych czynników na poziom samopoczucia

Wstęp

Celem naszego projektu jest analiza statystyczna wpływu różnych czynników na poziom samopoczucia.

Opis danych

(Import, obróbka, wstępna analiza, testy na normalność)

Hipotezy badawcze oraz ich weryfikacja

H_0 : nie istnieje zależność pomiędzy zmiennymi.

H_1 : istnieje zależność pomiędzy zmiennymi.

Hipotezy badane są na poziomie istotności $\alpha = 0.05$.

Hipoteza I

Badanie, czy istnieje zależność pomiędzy poczuciem upokorzenia a jakością snu.

P-value dla testu **rho-Spearmana** wynosi 0.0535368656, natomiast dla testu **tau-Kendalla** wynosi 0.0534696670. Nie ma podstaw do odrzucenia hipotezy H_0 , a więc nie istnieje istotna zależność pomiędzy poczuciem upokorzenia a jakością snu. Współczynniki korelacji dla testów **rho-Spearmana** oraz **tau-Kendalla** wynoszą odpowiednio -0.12478 oraz -0.10524, ale na poziomie istotności $\alpha = 0.05$ nie są istotne statystycznie.

Wykresy boxplot

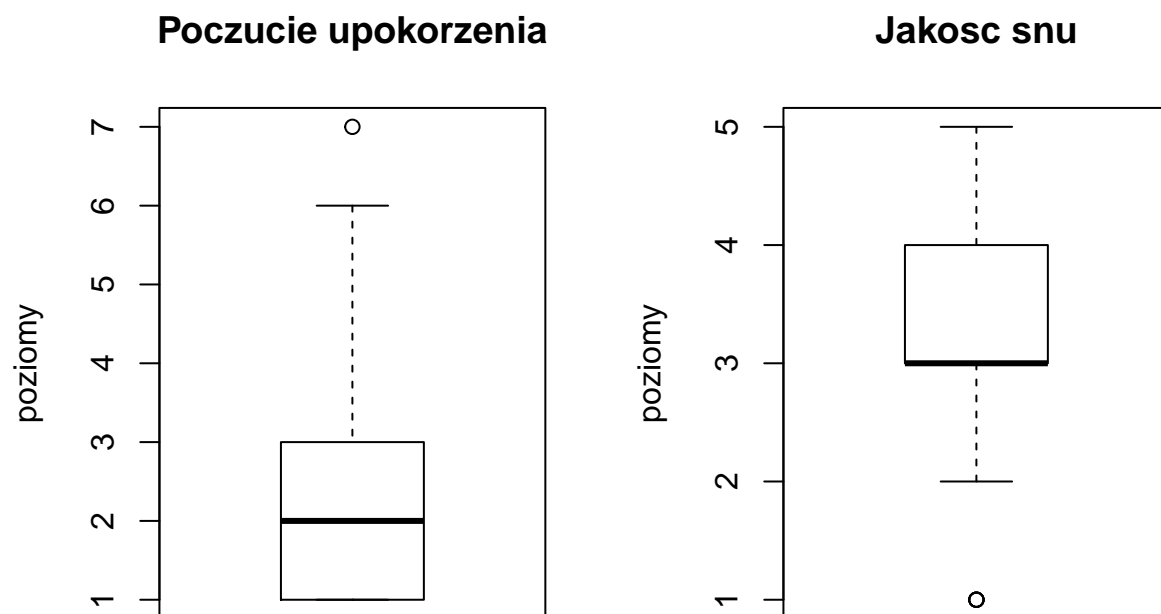
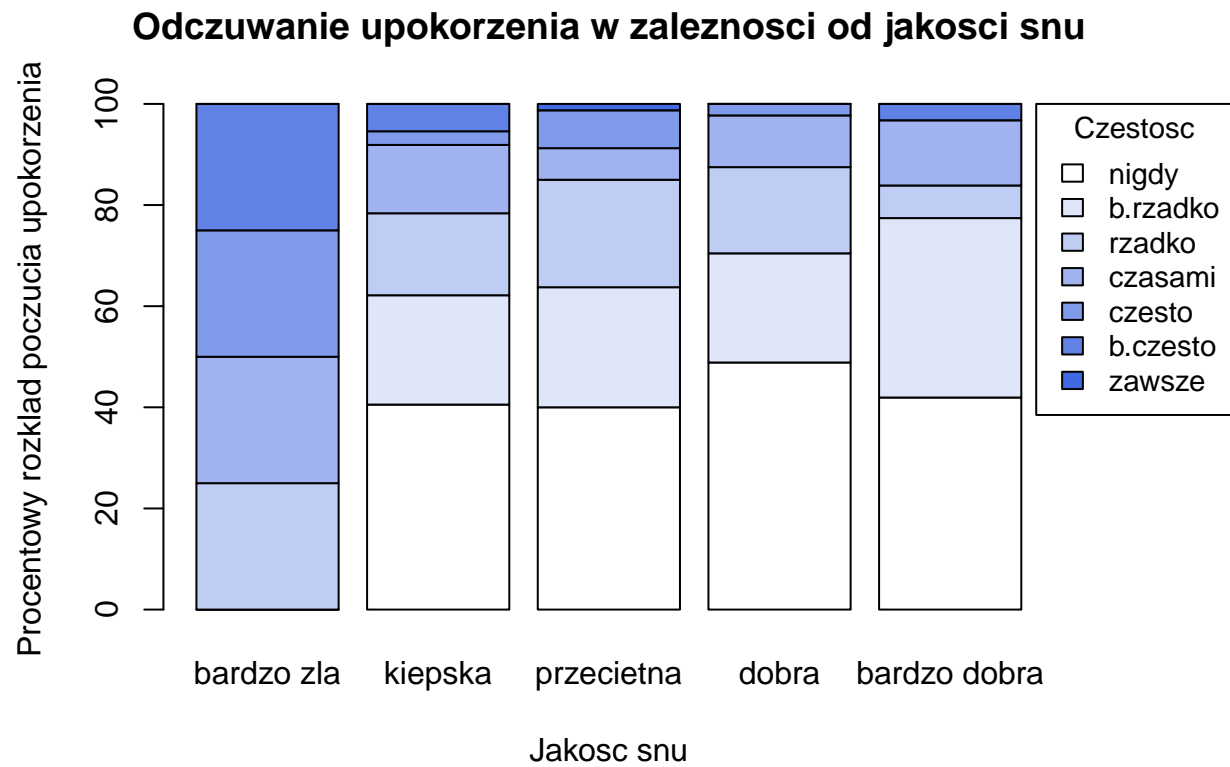


Tabela kontyngencji

Upokorzenie	Jakość snu				
	1	2	3	4	5
nigdy	0.0	40.5	40.0	48.9	41.9
bardzo rzadko	0.0	21.6	23.8	21.6	35.5
rzadko	25.0	16.2	21.2	17.0	6.5
czasami	25.0	13.5	6.2	10.2	12.9
często	25.0	2.7	7.5	2.3	0.0
bardzo często	25.0	5.4	0.0	0.0	3.2
zawsze	0.0	0.0	1.2	0.0	0.0



Hipoteza II

Badanie, czy istnieje zależność pomiędzy dietą a poczuciem radości.

P-value dla testu ρ -Spearmana wynosi 0.0839082132, natomiast dla testu τ -Kendalla wynosi 0.0839333604. Nie ma podstaw do odrzucenia H_0 . Nie istnieje istotna zależność pomiędzy dietą a poczuciem radości.

Wykresy boxplot

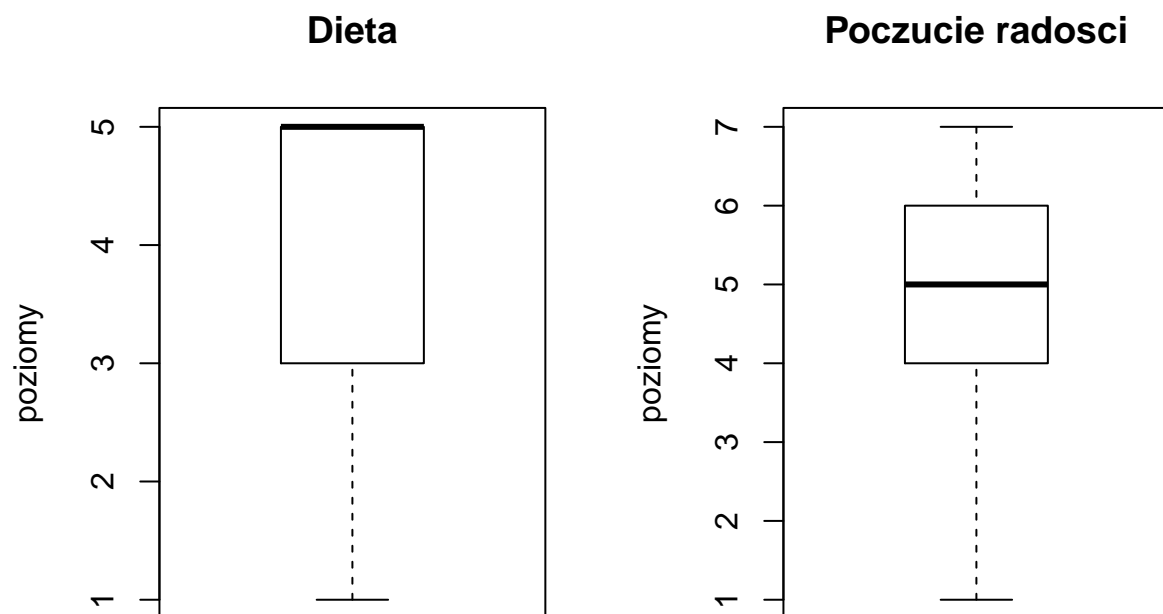
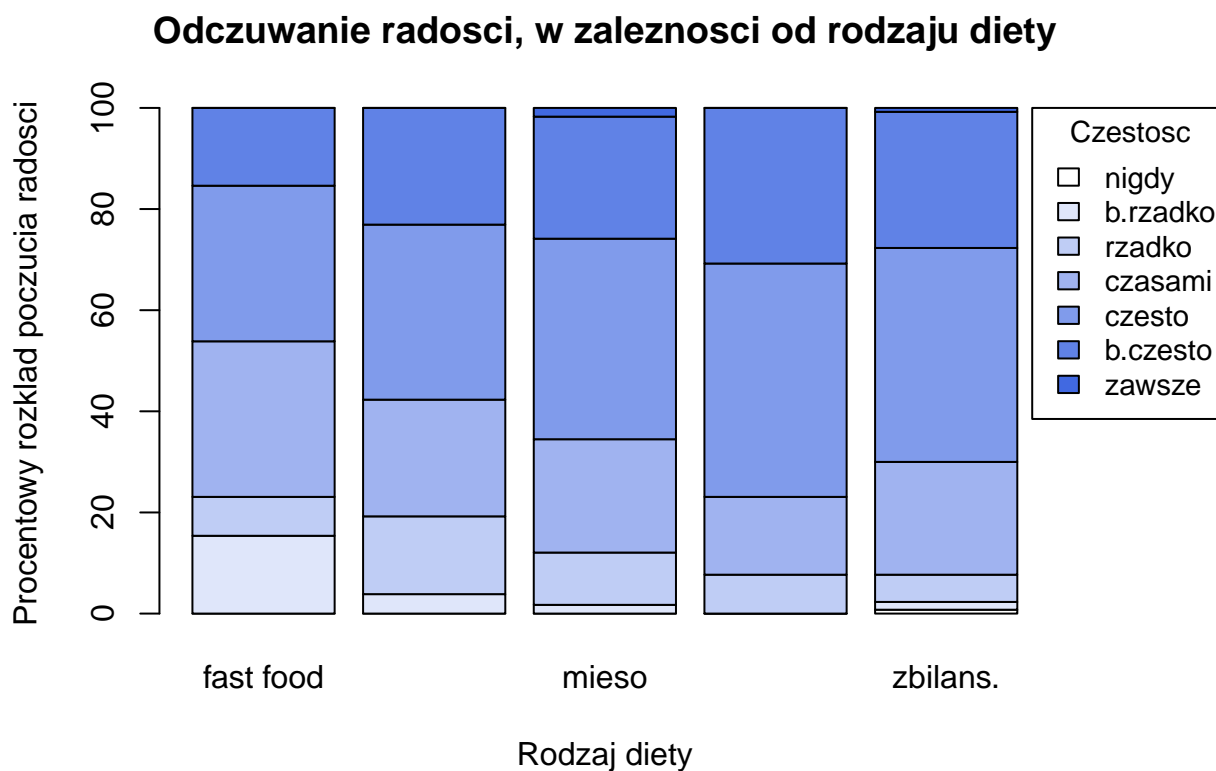


Tabela kontyngencji

Poczucie radości	Jakość diety				
	fast food	żywność mocno przetworzona (mrożonki, gotowe dania)	mięso i ciężkostrawne potrawy	warzywa i owoce	mam zbilansowaną dietę
nigdy	0.00	0.00	0.00	0.00	0.77
bardzo rzadko	15.38	3.85	1.72	0.00	1.54
rzadko	7.69	15.38	10.34	7.69	5.38
czasami	30.77	23.08	22.41	15.38	22.31
często	30.77	34.62	39.66	46.15	42.31
bardzo często	15.38	23.08	24.14	30.77	26.92
zawsze	0.00	0.00	1.72	0.00	0.77

Wykresy słupkowe



Hipoteza III

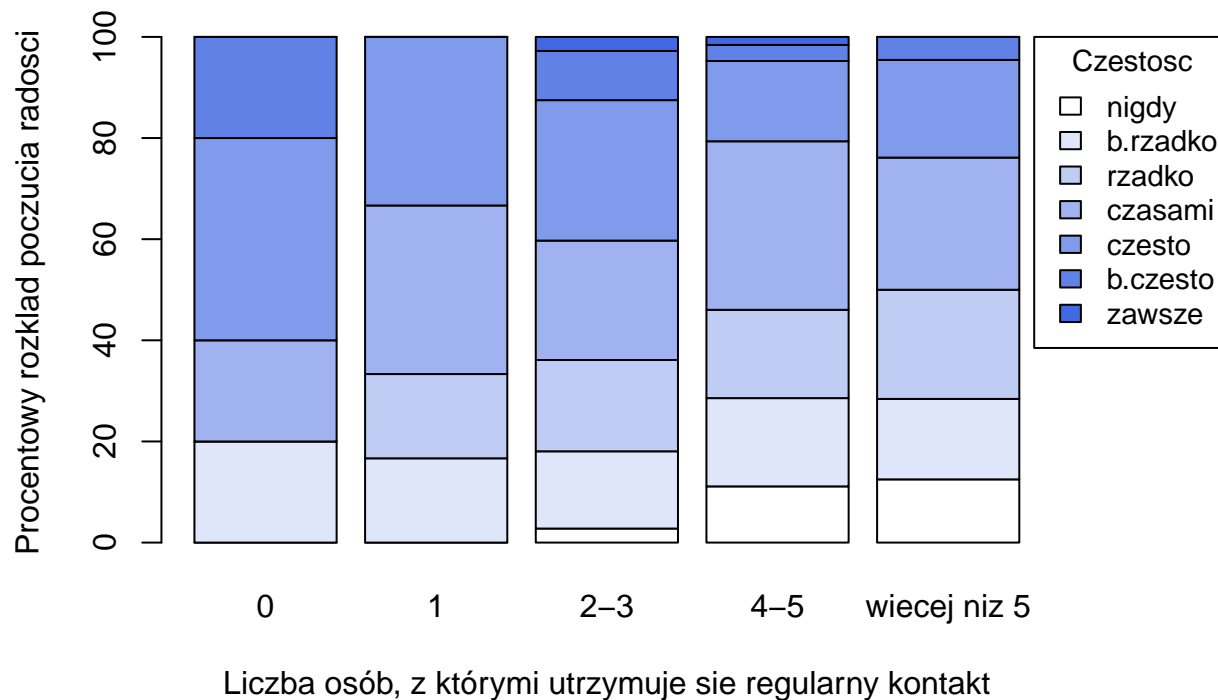
Badanie, czy istnieje zależność pomiędzy liczbą bliskich znajomych a odczuwaniem niepokoju.

P-value dla testu ρ -Spearmana wynosi 0.0055186575, natomiast dla testu τ -Kendalla wynosi 0.0057809380. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą osób, z którymi utrzymuje się regularny kontakt a odczuwaniem niepokoju.

Tabela kontyngencji

Niepokój	Liczba bliskich znajomych				
	0	1	2 - 3	4 - 5	5 i więcej
nigdy	0.0	0.0	2.8	11.1	12.5
bardzo rzadko	20.0	16.7	15.3	17.5	15.9
rzadko	0.0	16.7	18.1	17.5	21.6
czasami	20.0	33.3	23.6	33.3	26.1
często	40.0	33.3	27.8	15.9	19.3
bardzo często	20.0	0.0	9.7	3.2	4.5
zawsze	0.0	0.0	2.8	1.6	0.0

Odczuwanie niepokoju, w zależności od liczby bliskich znajomych



Współczynniki korelacji dla testów ρ -Spearmana i τ -Kendalla wynoszą odpowiednio -0.1786209 oraz -0.1473437. Interpretując te współczynniki oraz dane z wykresów, można stwierdzić, że wraz ze wzrostem liczby osób, z którymi utrzymuje się regularny kontakt, spada częstość odczuwania niepokoju.

Hipoteza IV

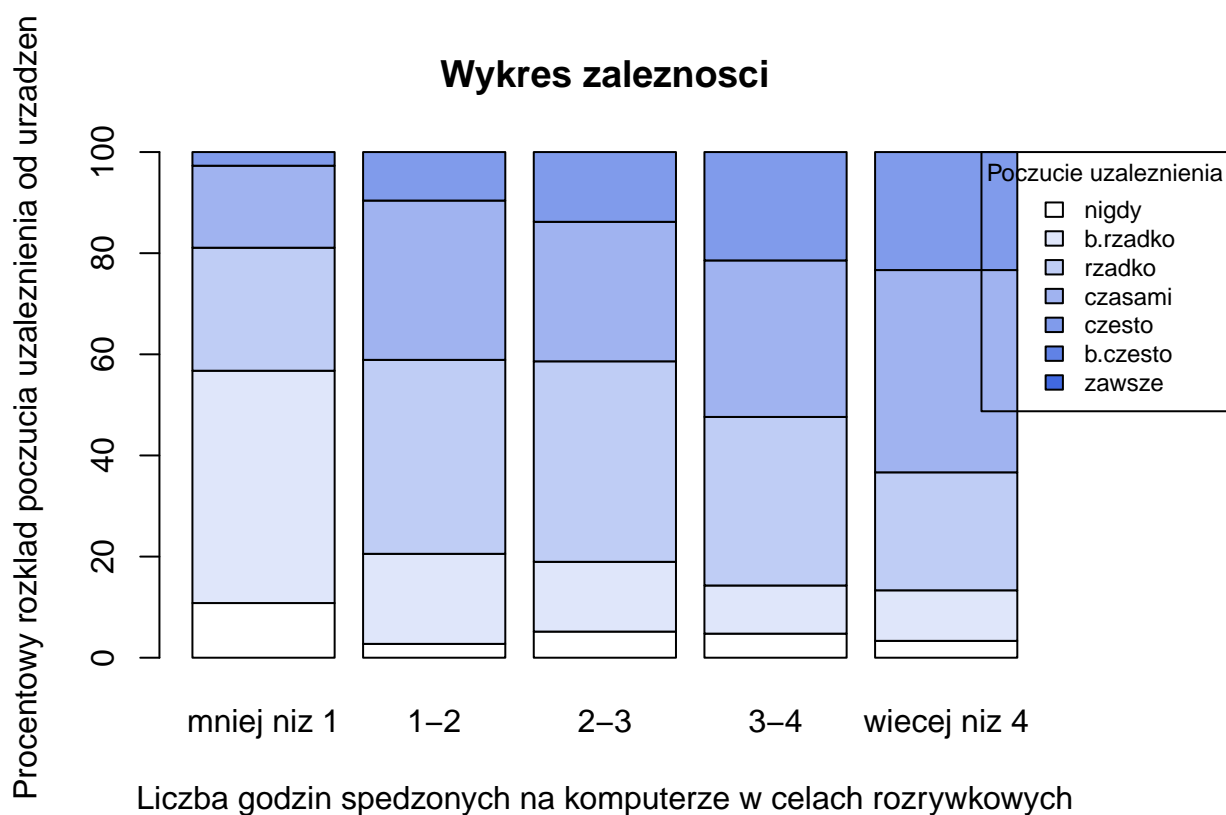
Badanie, czy istnieje zależność pomiędzy liczbą godzin przeznaczonych na korzystanie z urządzeń elektronicznych a poczuciem uzależnienia od urządzeń.

P-value dla testu ρ -Spearmana wynosi 0.0000046081, natomiast dla testu τ -Kendalla wynosi 0.0000047417. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą godzin spędzonych na używaniu urządzeń elektronicznych a poczuciem uzależnienia od nich.

Tabela kontyngencji

Poczucie uzależnienia	Godziny spędzone na używaniu urządzeń elektronicznych				
	mniej niż 1	1 - 2	2 - 3	3 - 4	więcej niż 4
nigdy	10.8	2.7	5.2	4.8	3.3
rzadko	45.9	17.8	13.8	9.5	10.0
czasami	24.3	38.4	39.7	33.3	23.3
często	16.2	31.5	27.6	31.0	40.0
bardzo często	2.7	9.6	13.8	21.4	23.3

Wykresy słupkowe



Współczynniki korelacji dla testów rho-Spearmana i tau-Kendalla wynoszą odpowiednio 0.2908585 oraz 0.2426912. Wniosek: wraz ze wzrostem liczby godzin, spędzonych na korzystaniu dla rozrywki z urządzeń elektronicznych, wzrasta znacznie poczucie uzależnienia od tych urządzeń.

Hipoteza V

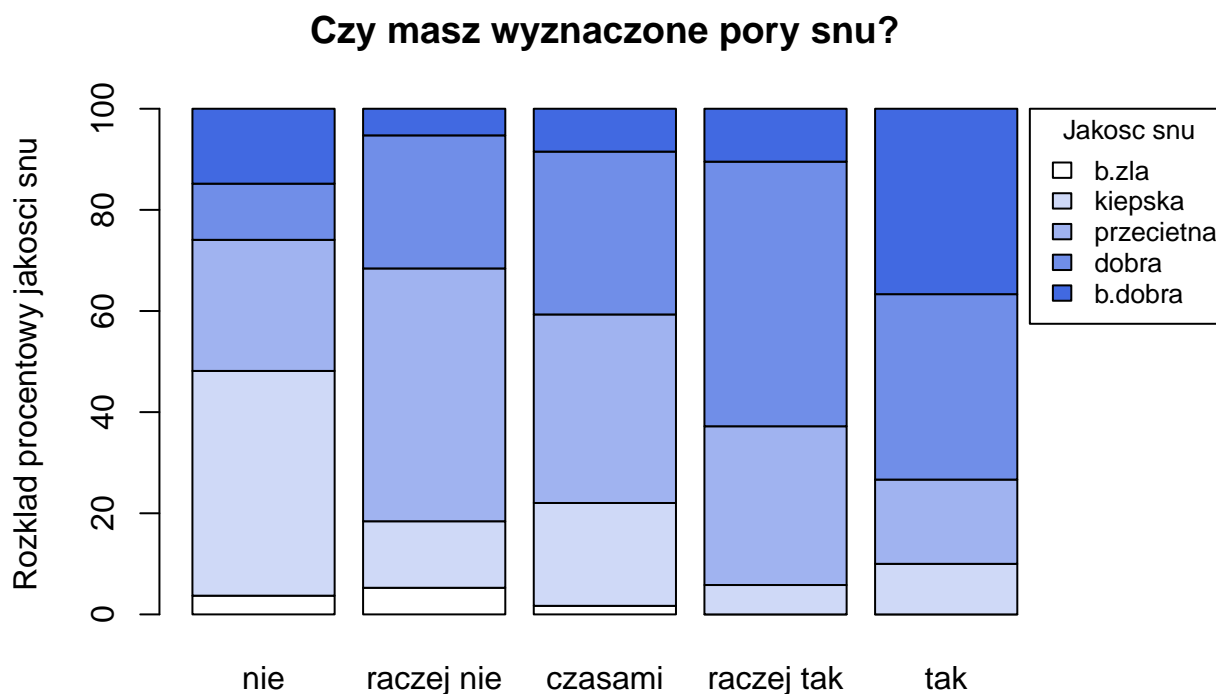
Badanie, czy istnieje zależność pomiędzy posiadaniem wyznaczonych pór snu a jego jakością.

P-value dla testu rho-Spearmana wynosi 0.0000000172, natomiast dla testu tau-Kendalla wynosi 0.0000000134. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy posiadaniem wyznaczonych pór snu a jego jakością.

Tabela kontyngencji

Jakość snu	Posiadanie wyznaczonych pór snu				
	1	2	3	4	5
1	3.7	5.3	1.7	0.0	0.0
2	44.4	13.2	20.3	5.8	10.0
3	25.9	50.0	37.3	31.4	16.7
4	11.1	26.3	32.2	52.3	36.7
5	14.8	5.3	8.5	10.5	36.7

Wykresy słupkowe



Współczynniki korelacji dla testów **rho-Spearmana** i **tau-Kendalla** są dość duże i wynoszą odpowiednio 0.3539432 oraz 0.3064142. Im bardziej regularne są pory snu, tym lepsza jest jego jakość.

Hipoteza VI

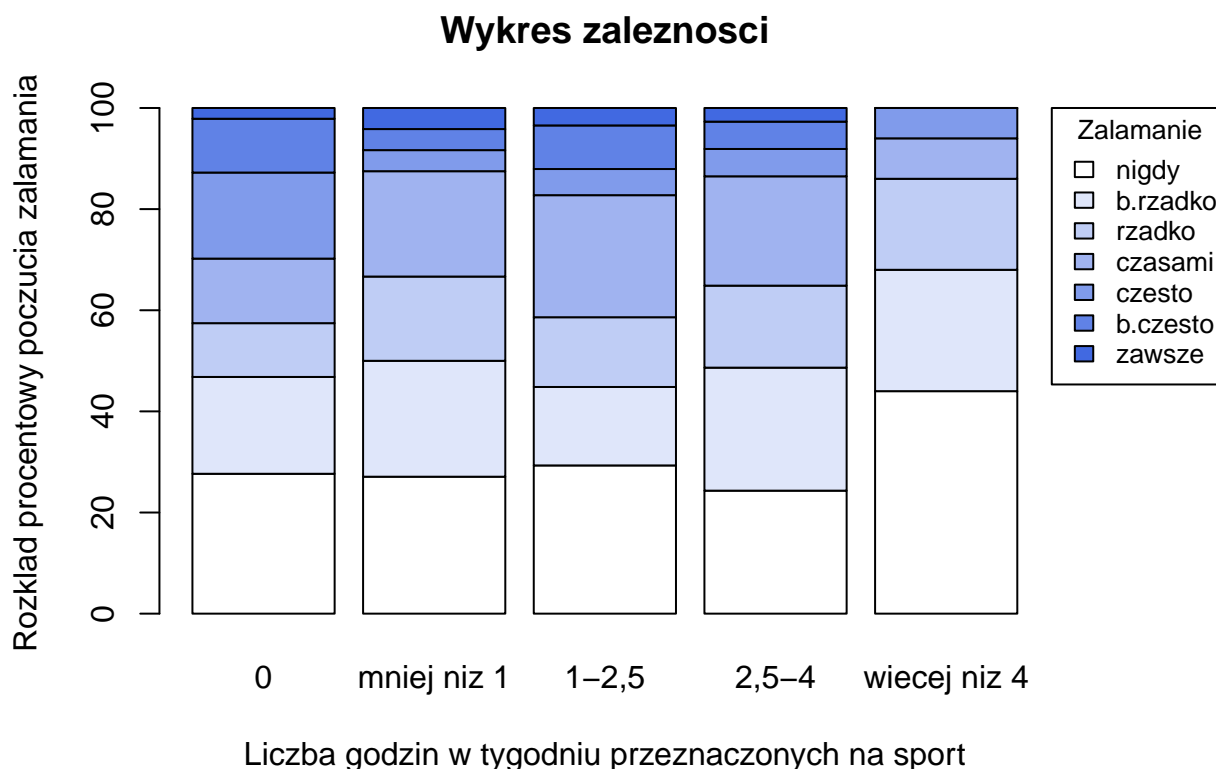
Badanie, czy istnieje zależność pomiędzy liczbą godzin przeznaczonych na uprawianie sportu a odczuwaniem załamania.

P-value dla testu **rho-Spearmana** wynosi 0.0114094634, natomiast dla testu **tau-Kendalla** wynosi 0.0107243185. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą godzin przeznaczonych na uprawianie sportu a odczuwaniem załamania.

Tabela kontyngencji

Załamanie	Liczba godzin przeznaczonych na sport				
	nie uprawiam żadnego sportu	mniej niż 1	1 - 2,5	2,5 - 4	więcej niż 4
nigdy	27.7	27.1	29.3	24.3	44.0
bardzo rzadko	19.1	22.9	15.5	24.3	24.0
rzadko	10.6	16.7	13.8	16.2	18.0
czasami	12.8	20.8	24.1	21.6	8.0
często	17.0	4.2	5.2	5.4	6.0
bardzo często	10.6	4.2	8.6	5.4	0.0
zawsze	2.1	4.2	3.4	2.7	0.0

Wykresy słupkowe



Współczynniki korelacji dla testów ρ -Spearmana i τ -Kendalla wynoszą odpowiednio -0.1630613 oraz -0.1318066. Im więcej godzin poświęca się tygodniowo na sport, tym rzadziej odczuwa się załamanie.

Hipoteza VII

Badanie, czy istnieje zależność pomiędzy jakością posiłków a dobrym humorem.

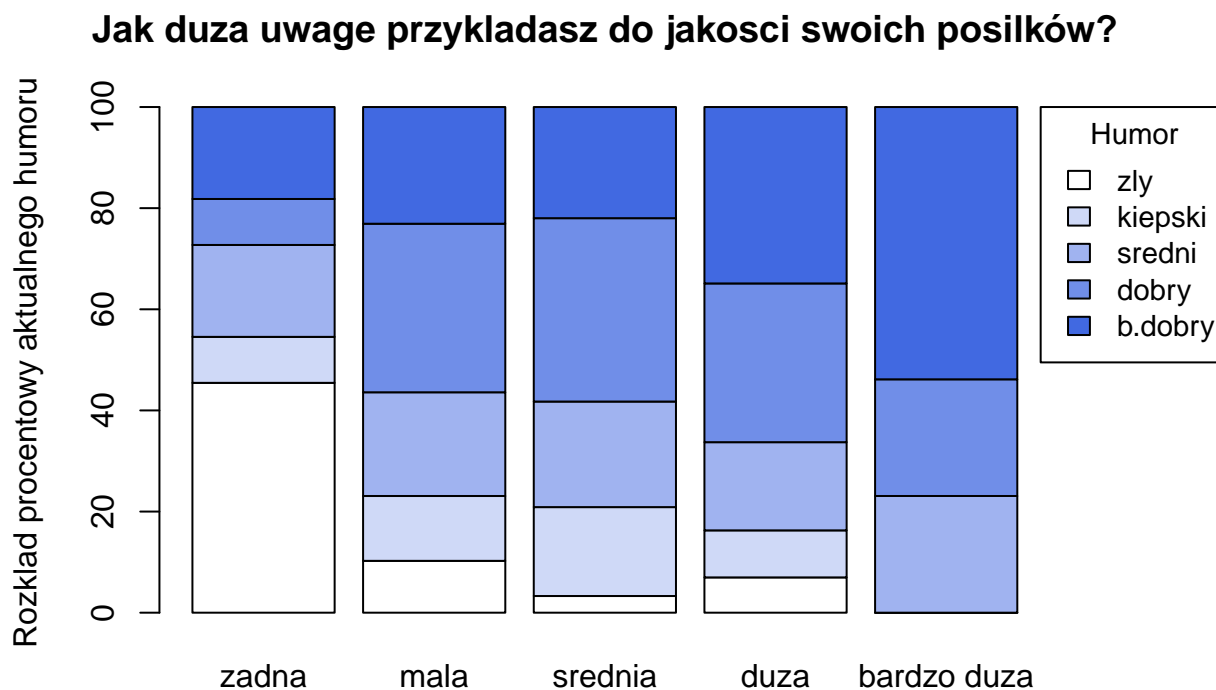
P-value dla testu ρ -Spearmana wynosi 0.0015012366, natomiast dla testu τ -Kendalla wynosi 0.0013687508. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy

jakością posiłków a dobrym humorem.

Tabela kontyngencji

Dobry humor	Jakość posiłków				
	wszystko mi jedno	nie jest to moim priorytetem	staram się jeść zdrowo, ale mi nie wychodzi	jest to dla mnie dość ważne	jestem tym, co jem!
nie zgadzam się	45.5	10.3	3.3	7.0	0.0
raczej się nie zgadzam	9.1	12.8	17.6	9.3	0.0
trochę tak, trochę nie	18.2	20.5	20.9	17.4	23.1
raczej się zgadzam	9.1	33.3	36.3	31.4	23.1
zgadzam się	18.2	23.1	22.0	34.9	53.8

Wykresy słupkowe



Zarówno współczynniki korelacji dla testów ρ -Spearmana i τ -Kendalla wynoszące odpowiednio 0.2038184 oraz 0.1738616, jak i wykresy słupkowe jasno wskazują na istotną zależność pomiędzy badanymi zmiennymi. Im większe znaczenie ma jakość spożywanych posiłków, tym lepszy humor się odczuwa.

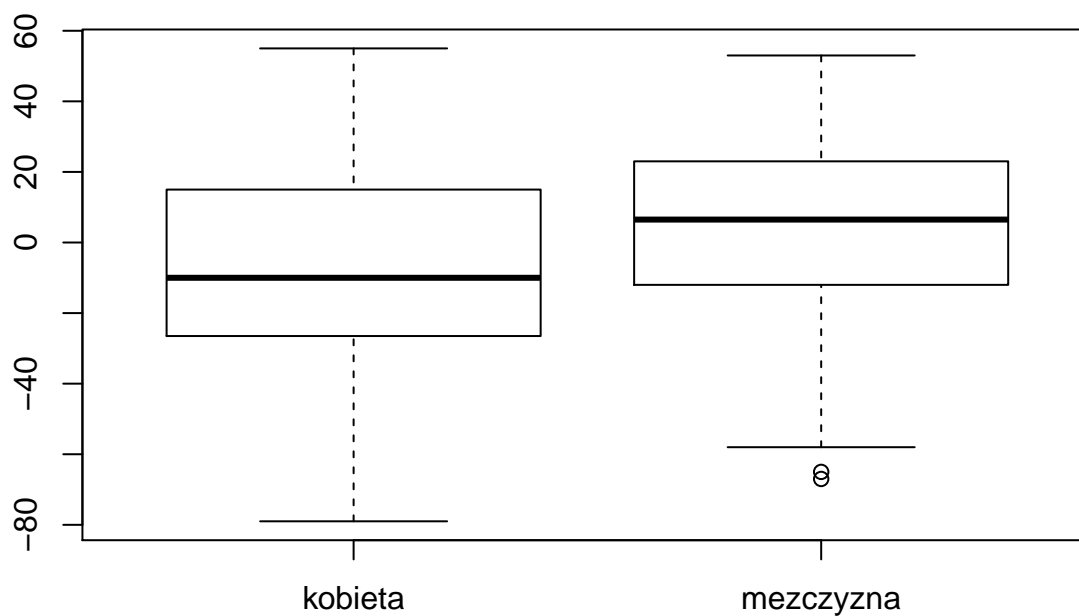
Hipoteza VIII

Badanie zależności między płcią a ogólną oceną samopoczucia.

Średnia ocena samopoczucia w grupie kobiet i mężczyzn

	Płeć	
	kobieta	mężczyzna
Ogólna ocena samopoczucia	-8.0	3.2

Wykres boxplot



Badanie zależności

P-value z testu Shapiro-Wilka dla cechy Ogólna ocena samopoczucia wynosi 0.00671, zatem ta cecha nie ma rozkładu normalnego. Stosujemy test Manna - Whitney'a - Wilcoxona.

P-value z testu Manna - Whitney'a - Wilcoxona wynosi 0.00150. Istnieje zależność pomiędzy płcią a ogólną oceną samopoczucia.

Badanie równości średnich

Test F wykazuje, że wariancje Ogólnej oceny samopoczucia są równe w grupach kobiet i mężczyzn (p-value 0.53174), zatem stosujemy test T dla średnich.

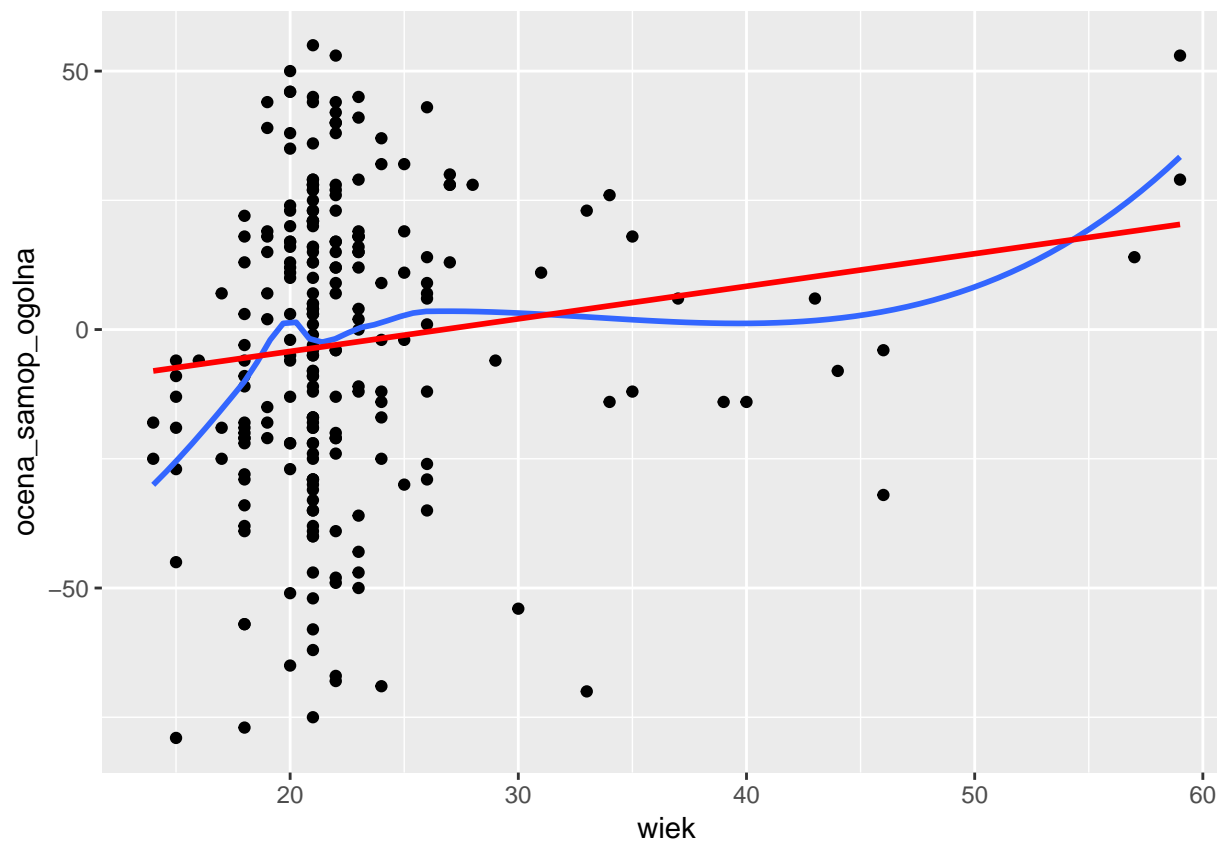
P-value z testu T wynosi 0.00236, zatem odrzucamy hipotezę o równości średnich. Średnia ogólna ocena samopoczucia różni się istotnie w grupach kobiet i mężczyzn.

Hipoteza IX

Badanie zależności między wiekiem a ogólną oceną samopoczucia.

Table 1:	
	<i>Dependent variable:</i>
	ocena_samop_ogolna
wiek	0.630** (0.292)
Constant	-16.832** (6.847)
Observations	240
R ²	0.019
Adjusted R ²	0.015
Residual Std. Error	28.413 (df = 238)
F Statistic	4.649** (df = 1; 238)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Istnieje istotna zależność pomiędzy ogólną oceną samopoczucia a wiekiem.



Na niebiesko regresja nieparametryczna, na czerwono liniowa.

Na podstawie wykresów diagnostycznych oraz odpowiednich testów mamy, że model spełnia założenia Gaussa - Markowa, a jedynym problemem jest brak normalności rozkładu reszt (choć na wykresach diagnostycznych wygląda trochę na heterogeniczność wariancji).

Wnioski

Bibliografia

Skale do Pomiaru Nastroju i Sześciu Emocji, Bogdan Wojciszke i Wiesław Baryła

Wspomaganie Statystycznej Analizy Wyników Badań Empirycznych w STATISTICA 9, Janusz Wątroba, StatSoft Polska Sp. z o. o.