



UNIWERSYTET
ZIELONOGÓRSKI

Analiza statystyczna

JAK?



Prace zaliczeniowe studentek i studentów
2 roku psychologii z przedmiotu Metodologia i statystyka - laboratorium
zebrał Paweł Kleka

Spis zawartości:

1. Algorytm wyboru testu w zależności od sytuacji badawczej.....	2
2. Obliczanie nowych zmiennych.....	3
3. Transformacja zmiennych na przykładzie rangowania.....	4
4. Statystyki opisowe zmiennych ilościowych.....	5
5. Statystyki opisowe zmiennych kategorialnych.....	6
6. Testy chi-kwadrat.....	7
7. Test t-Studenta dla 1 próby.....	8
8. Test t-Studenta dla prób niezależnych.....	10
9. Test t-Studenta dla prób zależnych.....	11
10. Test Manna-Whitneya.....	12
11. Analiza wariancji ANOVA.....	13
12. Test nieparametrycznej analizy wariancji Kruskala-Wallisa.....	16
13. Anova dla powtarzanych pomiarów.....	18
14. Korelacja i korelacja cząstkowe.....	20
15. Analiza regresji - podstawy.....	22
16. Test ekwiwalencji (braku różnicy między próbami).....	24
17. Glosariusz pojęć ze statystyki i metodologii.....	26

Jakiego rodzaju są dane?

1. niezależne (np. niezależne grupy)
2. zależne (np. dwukrotny pomiar)

Na jakiej skali jest wyrażona zmienna zależna

1. Ilościowa
2. Porządkowa
3. Nominalna

Ile jest grup porównawczych (kategorii w zmiennej grupującej)?

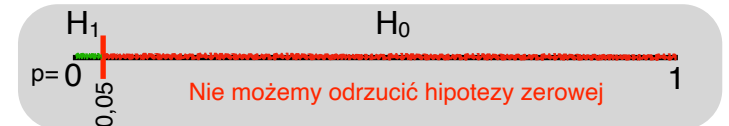
1. dokładnie 2 grupy
2. więcej niż 2 grupy

Czy rozkład zmiennej zależnej ilościowej jest normalny (test normalności (SW) przy Weryfikacji założeń)

$p < 0,05$

$p > 0,05$

Pamiętaj o wielkość efektu!



Porządkowa lub Ilościowa z rozkładem nienormalnym		Ilościowa oraz rozkład normalny				Porządkowa lub Ilościowa z rozkładem nienormalnym		Nominalna	
2				>2				2 lub >2	
niezależne	zależne	niezależne	zależne	niezależne	zależne	niezależne	zależne	niezależne	zależne
 T-Tests				 ANOVA				 Frequencies	
Test t dla danych niezależnych	Test t dla danych zależnych	Test t dla danych niezależnych	Test t dla danych zależnych	One-way lub ANOVA	Powtarzane pomiary ANOVA	Nieparametryczny		Tabele krzyżowe	
U Manna-Whitneya	Wilcoxon	Brak homogeniczności (p < 0,05) wpływa na wybór korekty dla t (Welcha)		Brak homogeniczności wpływa na wybór testu post hoc dla porównań parami i korekty dla F (Welcha)		Jednoczynnikowa ANOVA (test Kruskala-Wallisa)	ANOVA dla powtarzanych (test Friedmana)	Test zgodności χ²	Test McNemara

Dla związków między zmiennymi skorzystaj z odpowiednich współczynników korelacji: (Analizy $\rightarrow$ Test zgodności $\chi^2 \rightarrow$ Phi lub V: nominalna ~ nominalna) (Analizy $\rightarrow$ Regresja $\rightarrow$ **Macierz korelacji** oraz: ilościowa ~ ilościowa $\Rightarrow$ r Pearson'a I ilościowa lub porządkowa ~ porządkowa $\Rightarrow$ rho Spearmana) lub więcej >1 predyktory ilościowe lub nominalne ~ zmienna objaśniana ilościowa (regresja liniowa) lub nominalna (regresja logistyczna).

Obliczanie nowych zmiennych w Jamovi

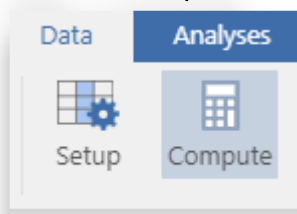
(jak obliczamy zmienne będące sumą lub średnią z innych wartości)

Przypuśćmy, że mamy już jakieś dane, np. :

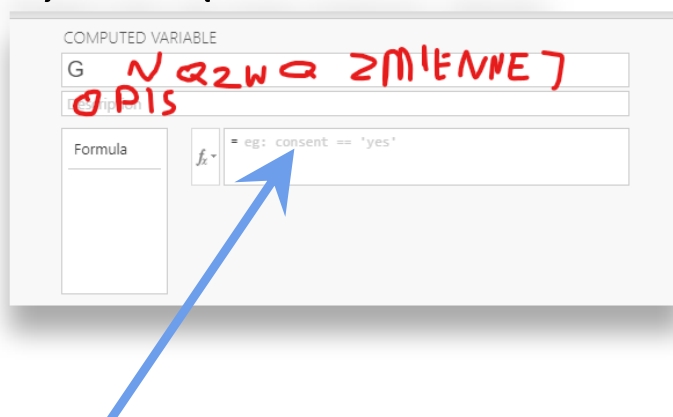
A		satysfakcja	afekt.poz	nadzieja.n...	pleć
1	1	1.720	-0.255	0.493	1
2	2	-1.246	0.141	-0.596	1
3	3	-0.590	-0.078	-0.170	1
4	4	1.095	0.510	0.600	1
5	5	0.205	1.094	0.296	1
6	6	-0.611	0.332	-1.160	1
7	7	-0.239	0.039	-0.264	1
8	8	-1.625	-0.257	-0.829	1
9	9	0.999	-0.585	0.516	1
10	10	1.948	-0.545	1.671	1
11	11	-0.277	-0.855	-1.188	1
12	12	1.237	-0.366	0.076	1
13	13	-0.390	0.142	0.105	1
14	14	-1.653	-1.257	-0.631	1
15	15	0.363	-0.116	0.179	1
16	16	-0.904	-1.203	-1.214	1
17	17	0.749	-0.193	1.542	1
18	18	-1.038	0.412	-1.009	1
19	19	-0.500	-1.094	0.934	1
20	20	1.463	-0.322	0.039	1
21	21	-0.372	1.285	2.111	1
22	22	-0.180	0.452	-0.297	1
23	23	-0.194	-0.018	-0.286	1
24	24	-0.817	-0.130	0.499	1
25	25	-0.299	1.232	-1.719	1
26	26	-0.766	1.306	0.858	1
27	27	0.986	-0.473	0.257	1
28	28	-1.153	0.600	0.417	1
29	29	-0.841	1.576	0.906	1
30	30	-1.460	0.078	-0.858	1
31	31	-0.042	-0.328	-0.058	1
32	32	0.072	1.816	-0.216	1
33	33	1.198	0.443	1.991	1
34	34	0.636	1.962	-0.943	1
35	35	-2.302	-0.382	-1.218	1
36	36	0.382	0.948	-0.024	1
37	37	0.923	1.287	-0.361	1
38	38	0.354	0.019	-1.798	1
39	39	-0.080	-0.576	-0.376	1
40	40	0.355	-0.816	-1.356	1
41	41	0.475	0.252	-0.184	1
42	42	-1.011	1.237	-2.094	1
43	43	-0.373	2.344	1.066	1
44	44	-0.663	0.493	0.509	1
45	45	0.363	1.166	0.127	1
46	46	-0.183	0.055	-1.243	1

Chcielibyśmy wyliczyć jaka jest suma satysfakcji, afektu oraz nadziei. Jak to zatem zrobić?

1. Data -> Compute



2. Pojawi nam się nowa zmienna i okienko



3. Na dole za znakiem równości musimy wpisać formułę, która ma być obliczeniem nowej zmiennej. W przypadku, gdy chcemy obliczyć sumę satysfakcji, afektu oraz nadziei, formuła będzie wyglądała następująco:

`SUM(satysfakcja, afekt.poz, nadzieja.na.sukces)`

(PS. Między zmiennymi musi występować przecinek (,), w przeciwnym wypadku funkcja nie zadziała)

4. Klikamy enter i mamy nową zmienną, która jest sumą trzech innych zmiennych.

satysfakcja	afekt.poz	nadzieja.n...	pleć	Satysfakcj...
1.720	-0.255	0.493	1	1.959
-1.246	0.141	-0.596	1	-1.702
-0.590	-0.078	-0.170	1	-0.838
1.095	0.510	0.600	1	2.205
0.205	1.094	0.296	1	1.595
-0.611	0.332	-1.160	1	-1.440
-0.239	0.039	-0.264	1	-0.464
-1.625	-0.257	-0.829	1	-2.711
0.999	-0.585	0.516	1	0.930
1.948	-0.545	1.671	1	3.073
-0.277	-0.855	-1.188	1	-2.319
1.237	-0.366	0.076	1	0.947

(Czarna kropka przy nazwie zmiennej oznacza, że jest to wynik innych zmiennych i dlatego nie można jej edytować)

Jak zatem wyliczyć średnią?

Otóż robimy to dokładnie tak samo, jednak w kroku 3. zmieniamy początek formuły z SUM na MEAN, a zatem:

`MEAN(satysfakcja, afekt.poz, nadzieja.na.sukces)`

pleć	Satysfakcj...	Średnia
1	1.959	0.653
1	-1.702	-0.567
1	-0.838	-0.279
1	2.205	0.735
1	1.595	0.532

I takim sposobem stworzyliśmy dwie nowe zmienne, które są średnią oraz sumą trzech innych zmiennych.

	A		C
1	1		
2	2		
3	3		
4	2		
5	2		
6	1		
7	2		
8	1		
9			
10			
11			
12			
13			
14			
15			

Wybierz dodanie nowej zmiennej Transformowanej

TRANSFORMED VARIABLE

A (2)

Description

Source variable A

using transform None Edit...

None

Create New Transform...

Zdefiniuj transformację

Wzór na transformację polegającą na rangowaniu to

$RANK(\$source)$

Nazwa transformacji jest dowolna, a poziom pomiaru najlepiej Ciągły

● TRANSFORM used by 1

rangowanie

Description

Variable suffix

+ Add recode condition

$= RANK(\$source)$

Measure type Continuous

	A	A - rang...
1	1	2.0
2	2	5.5
3	3	8.0
4	2	5.5
5	2	5.5
6	1	2.0
7	2	5.5
8	1	2.0

Jeśli wszystko poszło dobrze, to pojawia się wartość w kolumnie nowej zmiennej.

Wartości te, to rangi.

Opis podstawowych parametrów statystycznych (średnia, mediana, odchylenie standardowe)

Przypuśćmy, że mamy już jakieś dane np:

	plec	status ekonomiczny	typ szkoły	program	umiejętność matematyczna	umiejętność fizyka	historia
1	chłopiec	niski	publiczna	ogólny	57	52	41
2	dziewczyna	przeciętny	publiczna	zawodowy	68	59	53
3	chłopiec	wysoki	publiczna	ogólny	44	33	54
4	chłopiec	wysoki	publiczna	zawodowy	63	44	47
5	chłopiec	przeciętny	publiczna	profil.naukowy	47	52	57
6	chłopiec	przeciętny	publiczna	profil.naukowy	44	52	51
7	chłopiec	przeciętny	publiczna	ogólny	50	59	42
8	chłopiec	przeciętny	publiczna	profil.naukowy	34	46	45
9	chłopiec	przeciętny	publiczna	ogólny	63	57	54
10	chłopiec	przeciętny	publiczna	profil.naukowy	57	55	52
11	chłopiec	przeciętny	publiczna	zawodowy	60	46	51
12	chłopiec	przeciętny	publiczna	profil.naukowy	57	65	51
13	chłopiec	wysoki	publiczna	profil.naukowy	73	60	71
14	chłopiec	wysoki	publiczna	profil.naukowy	54	63	57
15	chłopiec	niski	publiczna	profil.naukowy	45	57	50
16	chłopiec	niski	publiczna	ogólny	42	49	43
17	chłopiec	wysoki	publiczna	profil.naukowy	47	52	51
18	chłopiec	przeciętny	prywatna	ogólny	57	57	60
19	chłopiec	wysoki	publiczna	profil.naukowy	68	65	62
20	chłopiec	przeciętny	publiczna	ogólny	55	39	57
21	chłopiec	przeciętny	publiczna	ogólny	63	49	35
22	chłopiec	przeciętny	publiczna	zawodowy	63	63	75
23	chłopiec	przeciętny	publiczna	profil.naukowy	50	40	45
24	chłopiec	wysoki	publiczna	profil.naukowy	60	52	57
25	chłopiec	przeciętny	publiczna	zawodowy	37	44	45
26	chłopiec	przeciętny	publiczna	zawodowy	34	37	46
27	chłopiec	wysoki	publiczna	profil.naukowy	65	65	66
28	chłopiec	przeciętny	prywatna	zawodowy	47	57	57
29	chłopiec	wysoki	prywatna	profil.naukowy	44	38	49
30	chłopiec	niski	publiczna	ogólny	52	44	49

Jak wyznaczyć średnią i medianę¹?

Krok 1. Wchodzimy w Exploration a następnie Descriptives

Krok 2. Zaznaczamy opcję „mean” i „median” i otrzymujemy wynik.

Statistics

Sample Size
☐ N ☐ Missing

Percentile Values
☐ Cut points for 4 equal groups
☐ Percentiles 25,50,75

Dispersion
☐ Std. deviation ☐ Minimum
☐ Variance ☐ Maximum
☐ Range ☐ IQR

Mean Dispersion
☐ Std. error of Mean
☐ Confidence interval for Mean 95 %

Central Tendency
☒ Mean
☒ Median
☐ Mode
☐ Sum

Distribution
☐ Skewness
☐ Kurtosis

Normality
☐ Shapiro-Wilk

Outliers
☐ Most extreme 5 values

W tym miejscu możemy włączyć wyliczanie różnych parametrów dotyczących opisu zmiennych w zależności od ich poziomu pomiaru i naszych potrzeb.

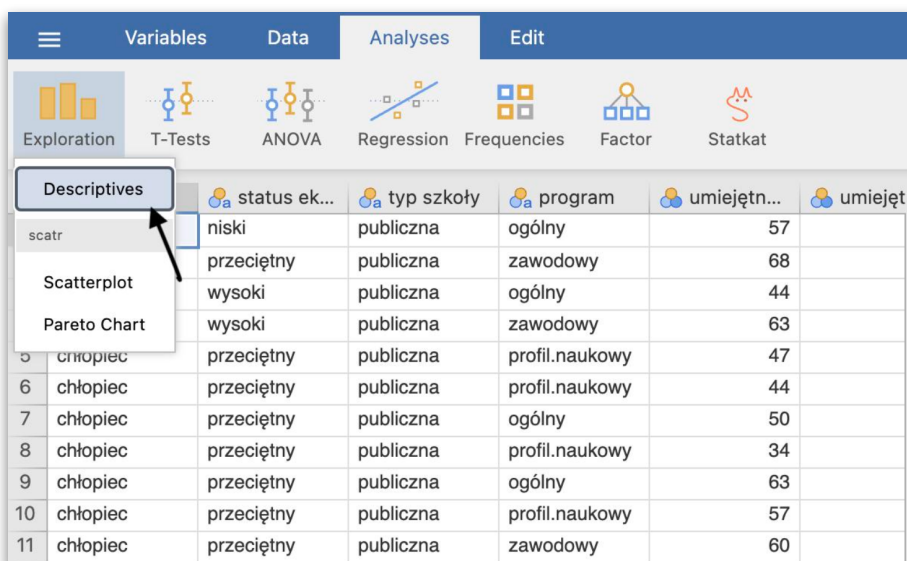
Poniżej znajduje się zakładka PLOTS, która umożliwia wizualizację rozkładów zmiennych. Zachęcam do eksperymentów i oglądania w [oknie raportu](#) efektów.

¹ Mediana jest wartością środkową zbioru, dzieli ona wszystkie nasze obserwacje na dwie równe części.

Opis podstawowych parametrów zmiennej kategorialnej

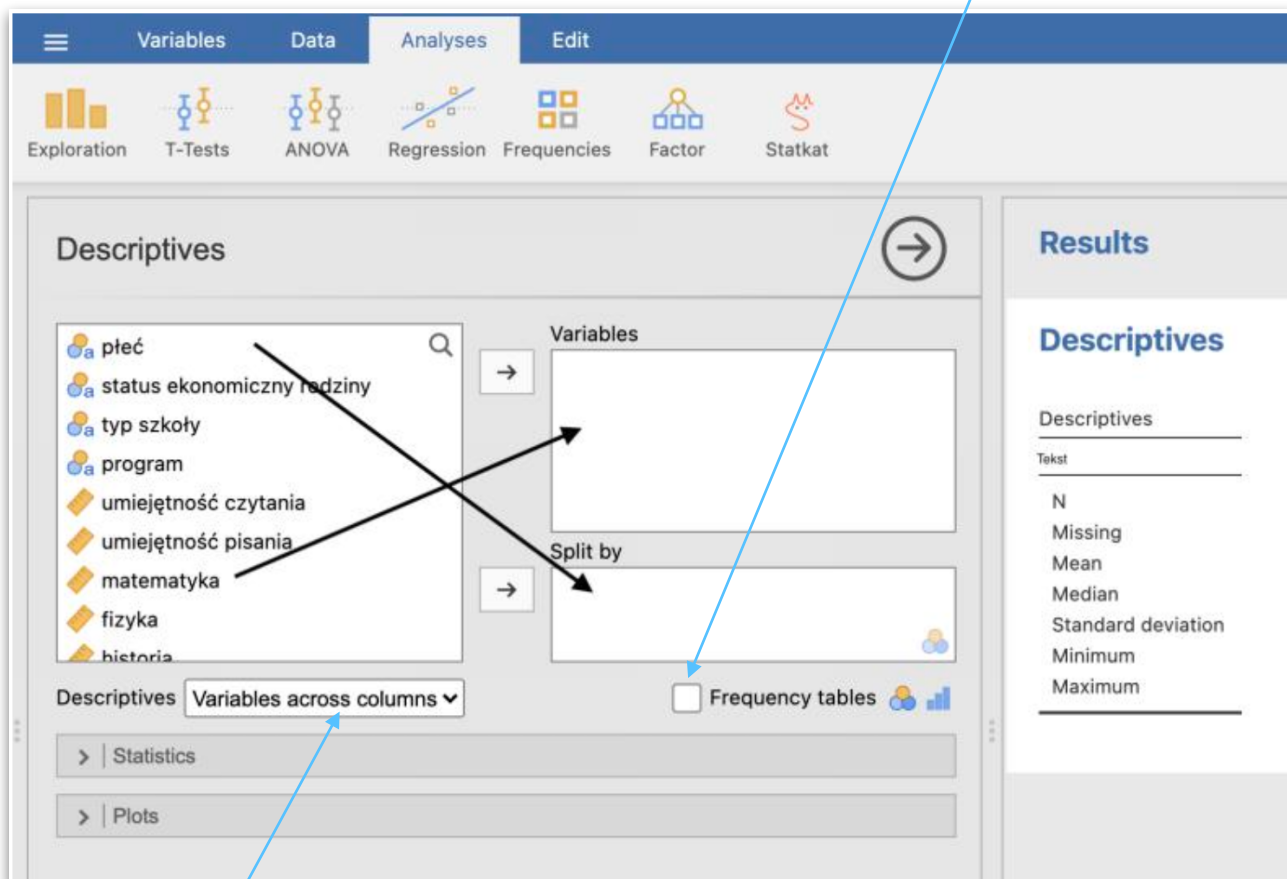
1. Przypuśćmy, że mamy już dane zawierające zmienną kategorialną (jeśli nie to musimy je zaimportować lub wprowadzić ręcznie).

2. Wchodzimy w zakładkę "Exploration" i wybieramy polecenie "Descriptives"



	status ek...	typ szkoły	program	umiejętn...	umiejęt
scatr	niski	publiczna	ogólny	57	
	przeciętny	publiczna	zawodowy	68	
	wysoki	publiczna	ogólny	44	
	wysoki	publiczna	zawodowy	63	
5 chłopiec	przeciętny	publiczna	profil.naukowy	47	
6 chłopiec	przeciętny	publiczna	profil.naukowy	44	
7 chłopiec	przeciętny	publiczna	ogólny	50	
8 chłopiec	przeciętny	publiczna	profil.naukowy	34	
9 chłopiec	przeciętny	publiczna	ogólny	63	
10 chłopiec	przeciętny	publiczna	profil.naukowy	57	
11 chłopiec	przeciętny	publiczna	zawodowy	60	

3. Z listy zmiennych przenosimy zmienne, które chcemy opisać do pola **Variables**. Zmienną kategorialną możemy opisać **tylko** przez częstości i w tym celu włączamy **tabele częstości**. Pole **Split by** służy do podzielenia naszego zbioru na podgrupy i przedstawieniu wyników osobno np. dla kobiet i dla mężczyzn.



Descriptives

Variables: płeć, status ekonomiczny rodziny, typ szkoły, program, umiejętność czytania, umiejętność pisania, matematyka, fizyka, historia

Split by:

Frequency tables: ☒

Descriptives: Variables across columns

Statistics: > | Statistics

Plots: > | Plots

Results

Descriptives

Descriptives

Tekst

N

Missing

Mean

Median

Standard deviation

Minimum

Maximum

Jeśli opisujemy kilka zmiennych na raz to możemy zmienić układ tabeli z kolumnowego na wierszowy

Testy χ^2

Wykorzystanie testu χ^2 , kiedy mamy zmienne nominalne.

Stosuje się go w celu sprawdzenia, czy między dwiema zmiennymi jakościowymi występuje istotna statystycznie zależność. Bazuje on na porównywaniu ze sobą liczebności obserwowanych, tj. takich, które uzyskaliśmy w badaniu, z liczebnościami oczekiwanymi, tj. takimi, które zakłada test, gdyby nie było żadnego związku między zmiennymi. Jeżeli różnica pomiędzy liczebnościami obserwowanymi, a oczekiwanymi jest duża (istotna statystycznie) to można uznać, że zachodzi zależność między jedną zmienną, a drugą (w Jamovi: test proporcji).

1. Od czego zacząć?

Wybieramy test χ^2 , gdy:

- mamy 2 zmienne
- są to zmienne nominalne
- są to zmienne niezależne

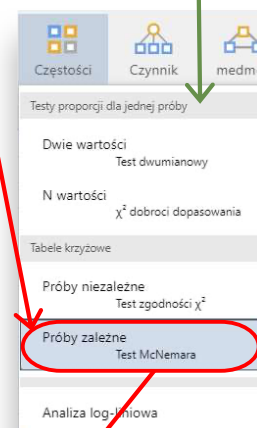
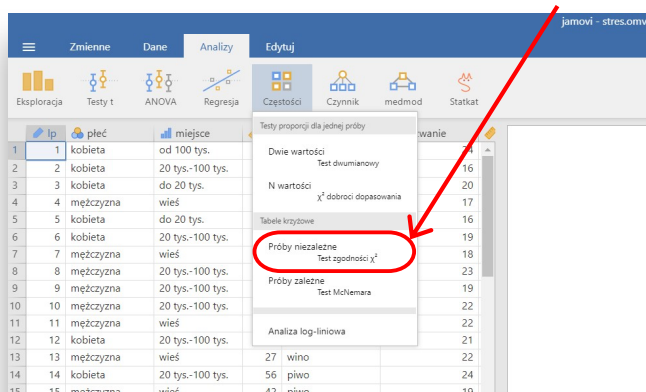


Wybieram test McNemara, gdy:

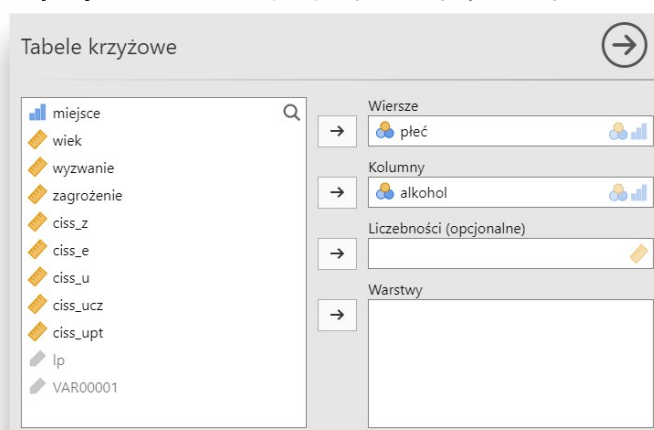
- mamy tylko 2 pomiary
- mierzymy zmienną nominalną z 2 kategoriami
- pomiary są zależne

2. Gdzie znaleźć ten test?

analizy → częstości → tabele krzyżowe → próby niezależne



3. Gdzie wpisujemy dane? Dane wpisujemy w rubrykę wierszy i kolumn



Tabele krzyżowe dla prób zależnych

Tabele krzyżowe

stres pomiar 1	stres pomiar 2		Całość
	brak stresu	odczuwanie stresu	
brak stresu	15	5	20
odczuwanie stresu	4	38	42
Całość	19	43	62

Test McNemara

	Wartość	df	p
χ^2	0.111	1	0.739
N	62		

Testy χ^2

	Wartość	df	p
χ^2	0.0261	1	0.872
N	67		

Miary porównawcze

95% przedziały ufności

	Wartość	Dolna granica	Górna granica
Iloraz szans	1.08	0.413	2.84

4. Jak odczytujemy dane?

Uzyskany wynik zapisujemy w następujący sposób:
 $\chi^2(df, N) = \text{wartość_}\chi^2, \text{wartość_}p, \text{wartość_}OR$

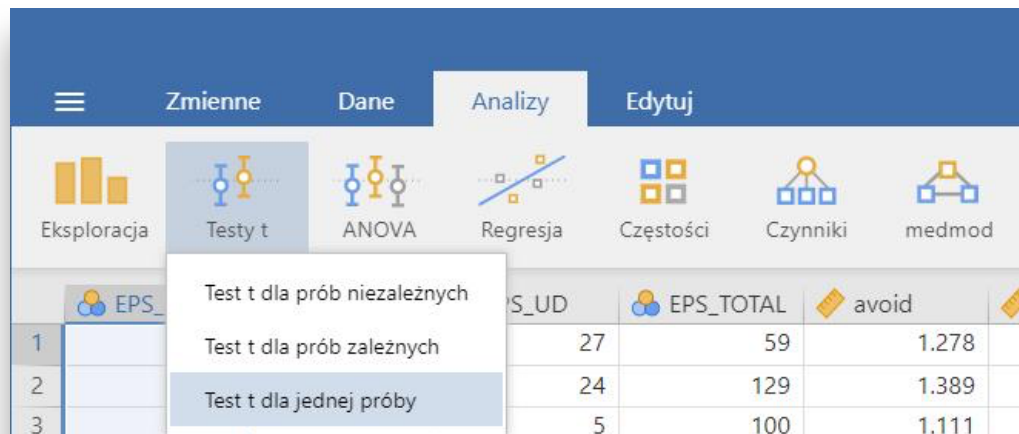
(gdzie df – stopień swobody; N – wielkość próby;
 χ^2 – wynik Naszego testu;
 p – wynik prawdopodobieństwa Naszej hipotezy; OR – iloraz szans).

W tym przypadku zapisalibyśmy to w następujący sposób:
 $\chi^2(1, N=67) = 0,026, p = 0,872, OR = 1,08; 95\% CI [0,41; 2,84]$

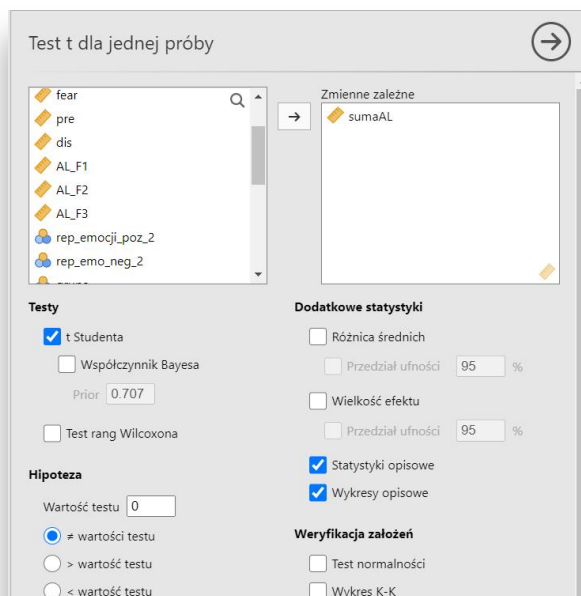
Test t- Studenta dla 1- próby

Użycie testu t-Studenta dla 1- próby umożliwia sprawdzenie, czy średnia jednej zmiennej różni się od określonej stałej średniej lub czy średnia wartość, którą uzyskaliśmy w badaniu różni się od zakładanej wartości.

1. Wprowadzamy lub importujemy dane.
2. Z analiz wybieramy **Testy t**, a następnie **Test t dla jednej próby**:



3. Przeciągamy zmienne na odpowiednie miejsca.
4. W polu **Hipoteza** wpisujemy stosowną wartość lub zostawiamy 0 (w zależności czemu ma służyć przeprowadzony test)



Moje dane przedstawiały sumę punktów jakie osoby uzyskały podczas testu na aleksytymię.

Do wykonania testu t-Studenta dla jednej próby rozkład zbliżonej zmiennej musi być zbliżony do rozkładu normalnego.

Możemy to sprawdzić używając **statystyk opisowych** i wybierając **test normalności** w **weryfikacji założeń**:

Test t dla jednej próby

☒ fear
☒ pre
☒ dis
☒ AL_F1
☒ AL_F2
☒ AL_F3
☒ rep_emocji_poz_2
☒ rep_emo_neg_2

Zmienne zależne
☒ sumaAL

Testy
☒ t Studenta
☐ Współczynnik Bayesa
 Prior: 0.707
☐ Test rang Wilcoxona

Hipoteza
 Wartość testu: 66
☒ ≠ wartości testu
☐ > wartości testu
☐ < wartości testu

Dodatkowe statystyki
☐ Różnica średnich
☐ Przedział ufności: 95 %
☐ Wielkość efektu
☐ Przedział ufności: 95 %
☒ Statystyki opisowe
☒ Wykresy opisowe

Weryfikacja założeń
☒ Test normalności
☐ Wykres K-K

Test normalności (Shapiro-Wilk)

	W	p
sumaAL	0.9563	0.031

Uwaga. Niska wartość p sugeruje naruszenie założenia o normalności rozkładu

Jak widać, rozkład mojej zmiennej **NIE** jest zbliżony do rozkładu normalnego, co sugeruje wartość $p < 0.05$.

5. Wyniki

Średnia mojej próby wyniosła 66,7, w hipotezie zazaczyłam, że wartość testu powinna być mniejsza od zera. Nie ma żadnego wyniku, który byłby mniejszy od zera, przez co wartość $p = 1$.

Hipoteza

Wartość testu: 0

☐ ≠ wartości testu
☐ > wartości testu
☒ < wartości testu

Test t dla jednej próby

Test t dla jednej próby

	Statystyka	df	p
sumaAL t Studenta	26.52	59.00	1.000

Uwaga. $H_0: \mu < 0$

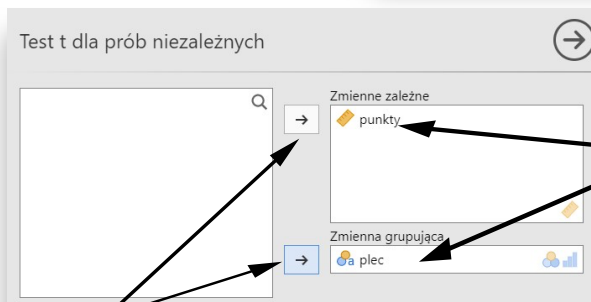
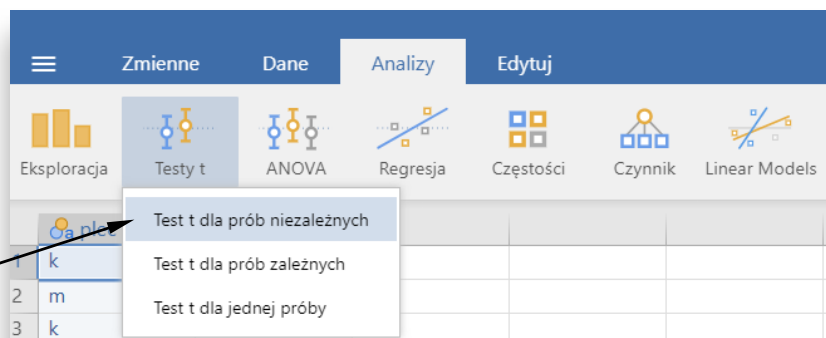
Otrzymany wynik nie jest istotny statystycznie, ponieważ ($p > 0.005$).

Oto przykładowy opis wyniku testu t-Studenta potwierdzającego różnicę między wynikami aleksytymii w próbie a normą:

Analiza t-Studenta dla jednej próby wykazała statystycznie istotną różnicę między wynikami aleksytymii w próbie ($M = 66,7$; $SD = 3,49$) a normą, $t(59) = 4,38$; $p < 0,001$. Średni wynik aleksytymii w próbie był istotnie wyższy niż średnia norma ($M = 50$), co sugeruje większe występowanie trudności w rozpoznawaniu i wyrażaniu emocji u badanych osób.

Test t- Studenta dla prób niezależnych

1. Jeśli chcesz porównać dwie różne grupy pod względem jakiejś zmiennej, to w zakładce „Analizy” wybierz „Testy t”, a następnie „Test t dla prób niezależnych”.



2. Przeciągnij i upuść odpowiednio zmienną zależną i zmienną grupującą.

Możesz zrobić to również za pomocą strzałek- kliknij na zmienną, a następnie użyj odpowiedniej strzałki.

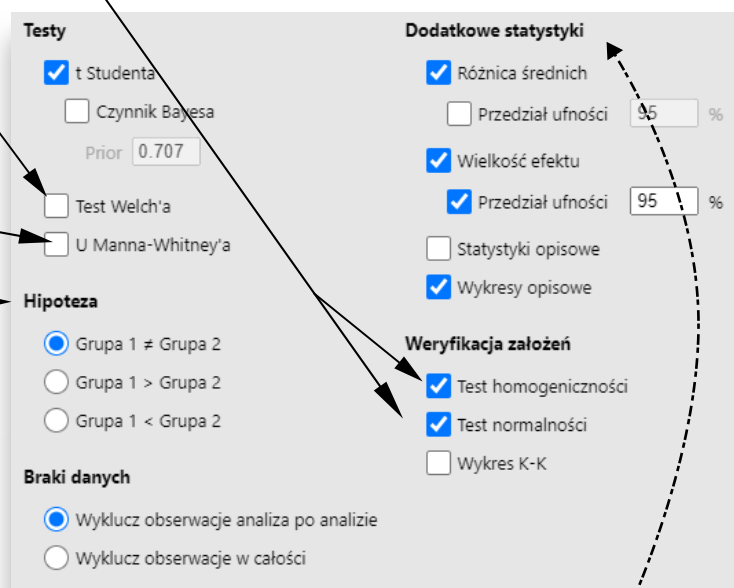
3. Pod spodem:

a) Wybierz test homogeniczności oraz test normalności.

b) Jeśli grupy nie są homogeniczne ($p < 0,05$), rozważ wybór Testu Welch'a.

c) Jeśli zmienna zależna nie ma rozkładu normalnego ($p < 0,05$), rozważ wybór testu U Manna- Whitney'a.

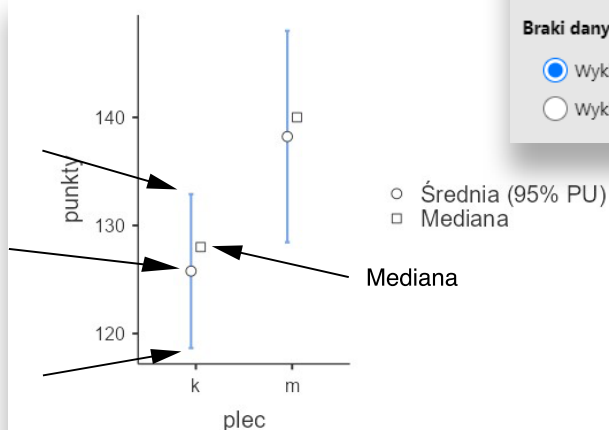
Zaznacz odpowiedni rodzaj hipotezy. Jeśli postawiona hipoteza jest kierunkowa, wybierz opcję drugą lub trzecią w zależności od treści hipotezy. Jeżeli hipoteza jest bezkierunkowa, wybierz opcję pierwszą.



Górna granica przedziału ufności średniej (95%)

Średnia

Dolna granica przedziału ufności średniej (95%)



Wybierz opcje „różnica średnich”, „wielkość efektu” z przedziałem ufności 95% oraz „wykresy opisowe”, żeby otrzymać szczegółowe informacje.

4. Wykres i tabela z wynikiem wyświetli się w prawym panelu strony.

Test t dla prób niezależnych

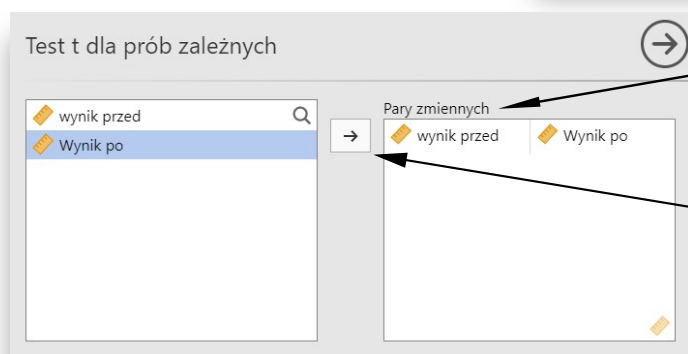
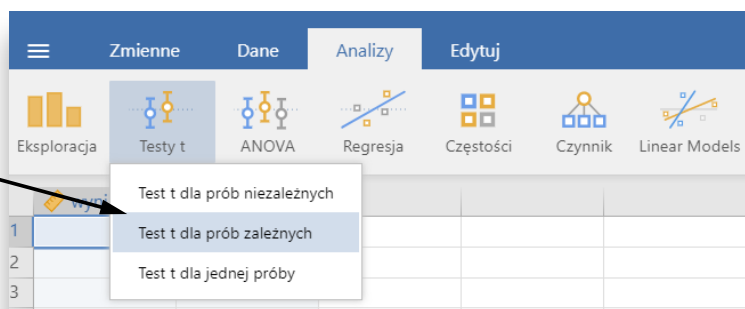
								95% przedział ufności		
		Statystyka	df	p	Różnica średnich	Różnica błędów standardowych	Wielkość efektu	Dolna granica	Górna granica	
punkty	t Studenta	-2.02	16.0	0.061	-12.4	6.17	d Cohena	-0.950	-1.96	0.104

Dla danych wykorzystanych w przykładzie poprawny zapis wyniku w tekście ma formę:
 $t(16) = 2,02; p = 0,061; d = -0,95, 95\% CI [-1,96; 0,10]$.

*Wynik ten nie jest istotny statystycznie.

Test t- Studenta dla prób zależnych

1. Jeśli chcesz sprawdzić, czy średnie pochodzące z dwóch powiązanych ze sobą prób różnią się między sobą, to zakładce „Analizy” wybierz „Testy t”, a następnie „Test t dla prób zależnych”.



2. Przecignij i upuść odpowiednio pary zmiennych.

Możesz zrobić to również za pomocą strzałki - kliknij na zmienną, a następnie użyj strzałki.

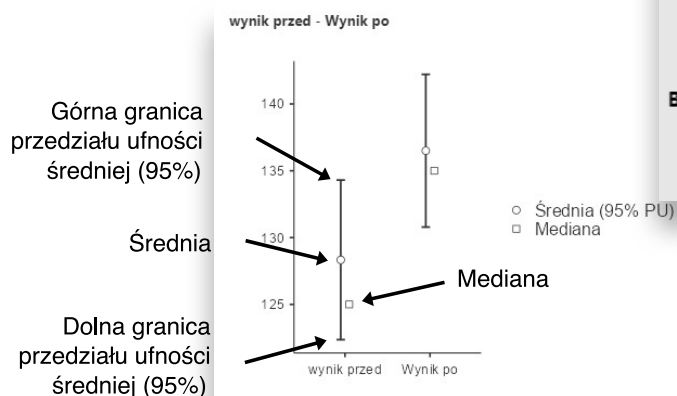
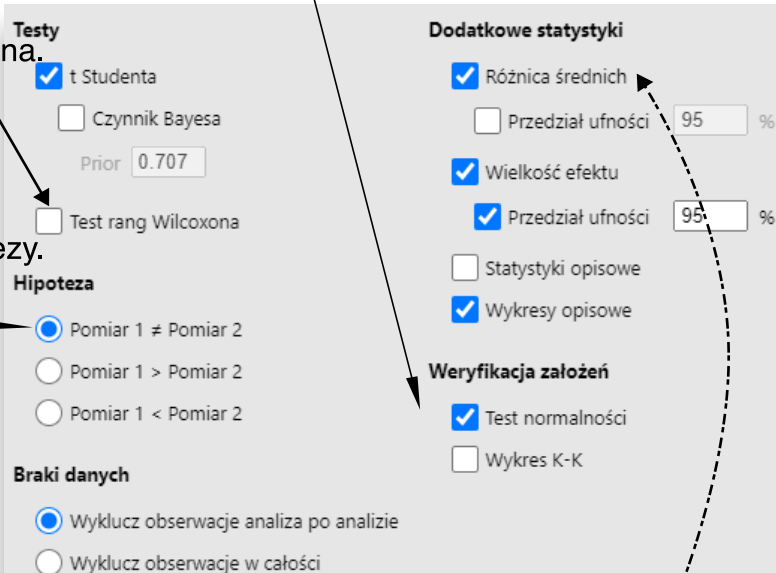
3. Pod spodem:

a) Wybierz test normalności, aby sprawdzić czy rozkład jest normalny (jeśli p jest większe niż 0,05 rozkład jest normalny).

b) Jeśli zmienna zależna wyrażona jest na skali porządkowej lub ilościowej, a rozkład nie jest normalny, rozważ zastosowanie Testu rang Wilcoxona.

c) Zaznacz odpowiedni rodzaj hipotezy.

Jeśli postawiona hipoteza jest kierunkowa, wybierz opcję drugą lub trzecią w zależności od treści hipotezy. Jeżeli hipoteza jest bezkierunkowa, wybierz opcję pierwszą.



Dla szczegółów wybierz opcje „różnica średnich”, „wielkość efektu” z przedziałem ufności 95% oraz „wykresy opisowe”.

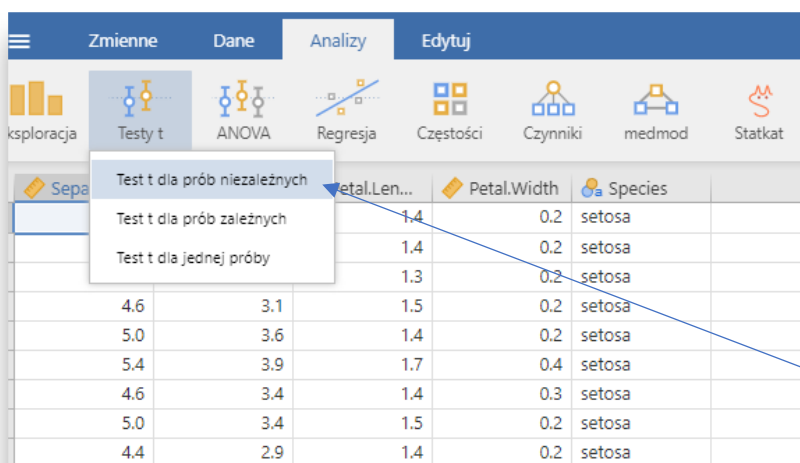
4. Wykres i tabela z wynikiem wyświetli się w prawym panelu strony.

Test t dla prób zależnych

								95% przedział ufności	
	statystyka	df	p	Różnica średnich	Różnica błędów standardowych	Wielkość efektu		Dolna granica	Górna granica
wynik przed	Wynik po	t Studenta	-4.65	17.0	< .001	-8.17	1.76	d Cohena	-1.10
								-1.68	-0.498

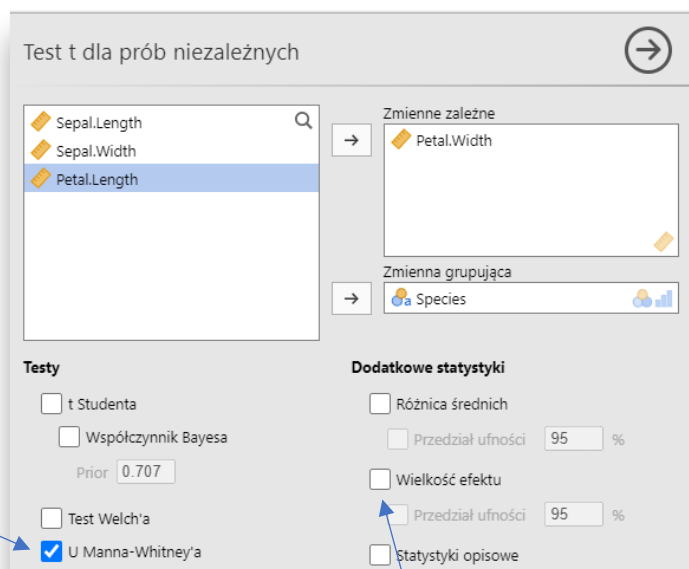
Dla danych wykorzystanych w przykładzie poprawny zapis wyniku w tekście ma formę:
 $t(17) = -4,65; p < 0,001; d = -1,10; 95\% CI [-1,68; -0,50]$.

Jak zrobić test U Manna – Whitneya w Jamovi - Poradnik



Aby móc wykonać w Jamovi test U Manna-Whitneya, pierwszym krokiem, który należy podjąć jest najechanie kursorem na ikonkę „Testy t”, po czym z opcji, które się pojawią pod spodem, kliknąć na tą o nazwie „Test t dla prób niezależnych”

Kiedy poprzedni krok został już wykonany na ekranie powiniensię pojawić taki oto widok. Teraz wystarczy już tylko zaznaczyć „ptaszka” w miejscugdzie znajdują się nazwa pożądanego testu! Brawo Ty :)



Niech cię tylko nie zmyli to zdanie :o

Wyniki

Test t dla prób niezależnych

		Statystyka	p
len	U Manna-Whitney'a	325	0.064

Uwaga. $H_0: \mu_{OJ} = \mu_{VC}$

A tak oto wygląda nasz test w obecności zmiennych ^^

A ten wynik to wielkość efektu ☺

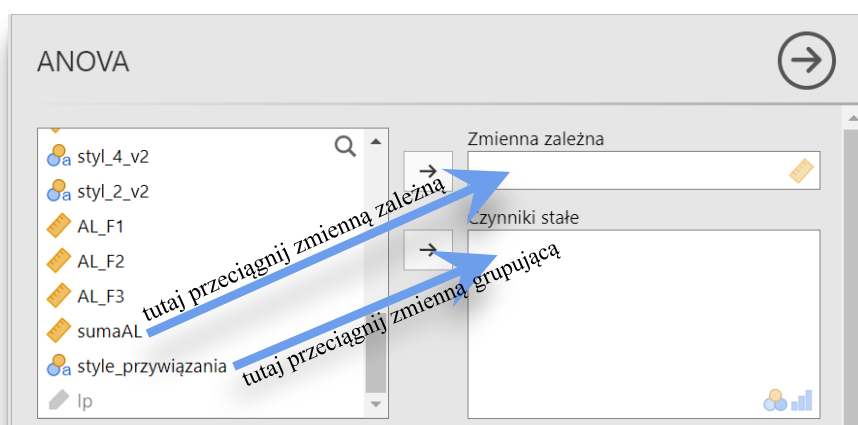
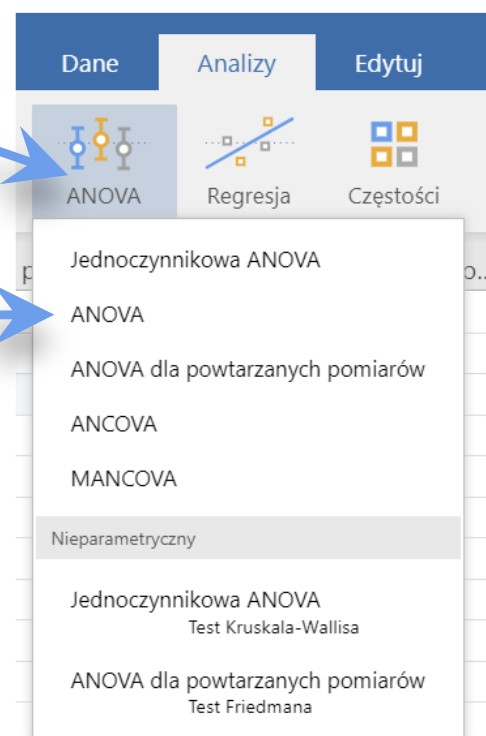
Jeśli chodzi o sam zapis wyników dla testu U, to należy zrobić go tak: $U = 325; p = 0,064; r_{pb} = 0,279$

ANOVA – jest to test porównujący średnie trzech lub więcej grup na podstawie prób niezależnych

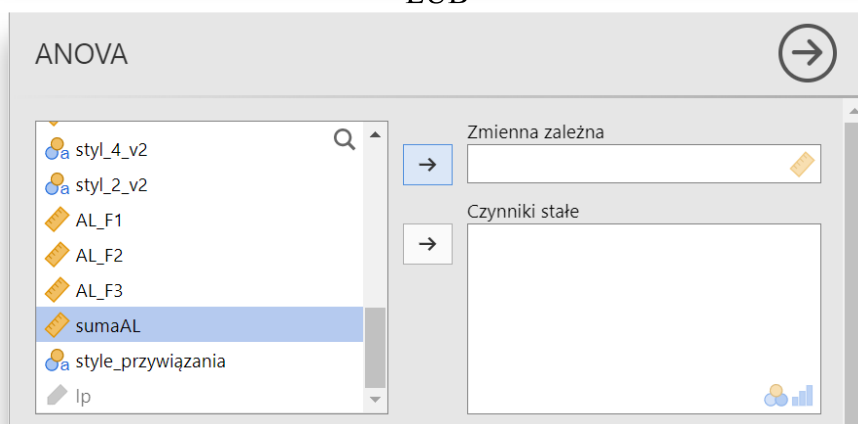
Aby go użyć:

wybierz zakładkę “ANOVA”

następnie kliknij opcję “ANOVA”



LUB



aby dodać wybraną zmienną, kliknij na nią (wtedy się podświetli) a następnie na strzałkę przy okienku na zmienną (wtedy „wskoczy” do tego okienka)

Po wprowadzeniu zmiennych w raporcie obok pojawi się tabela z wynikiem testu

ANOVA

ANOVA - sumaAL

	Suma kwadratów	df	Średni kwadrat	F	p	η^2
style_przywiązania	4502	3	1501	4.70	0.005	0.201
Reszty	17874	56	319			

Wielkość efektu

☒ η^2 ☐ częściowe η^2 ☐ ω^2

Aby obliczyć wielkość efektu, należy kliknąć interesującą nas opcję

Dla danych w tym przykładzie wynik analizy wariancji zapisujemy: $F(3, 56) = 4,70$; $p = 0,005$; $\eta^2 = 0,201$

Aby zweryfikować założenia o normalności i homogeniczności rozkładu zmiennej zależnej kliknij „Weryfikacja założeń” i wybierz obie opcje

> Model

▼ Weryfikacja założeń

☐ Test homogeniczności

☐ Test normalności

☐ Wykres K-K

> Kontrasty

jeżeli chcesz zrobić wykres K-K, kliknij tę opcję

W polu obok pojawiają się tabele z wynikami ($F(3, 56) = 5,02; p = 0,004$), ($p = 0,705$) i wykres:

Weryfikacja założeń

Test homogeniczności wariancji (Levene'a)

F	df1	df2	p
5.02	3	56	0.004

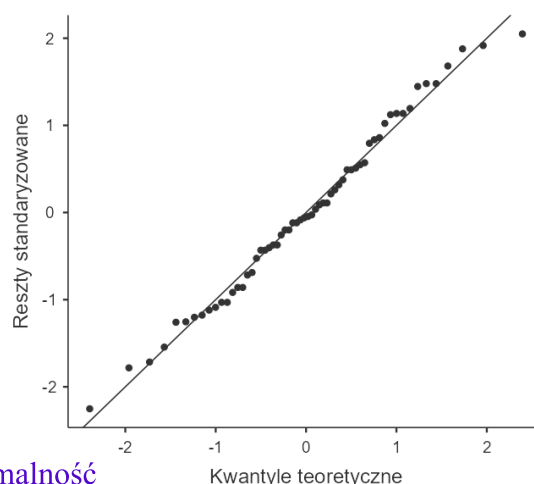
< 0,05 – wariancje nie są równe w grupach

Test normalności (Shapiro-Wilk)

Statystyka	p
0.986	0.705

> 0,05 – jest normalność

Wykres K-K



Nie spełnienie założeń ma wpływ na wynik ANOVA i może zafałszować jego interpretację

> Weryfikacja założeń

▼ Kontrasty

style_przywiazania

Brak

Brak

Odchylenie

Proste

Różnica

Helmert

Powtórzony

Wielomian

> Testy post-hoc

Żeby sprawdzić kontrasty (czyli zaplanowane różnice między wybranymi grupami) wybierz interesującą cię opcję – obok wyświetli się tabela prezentująca wyniki

Aby wykonać testy post-hoc (czyli porównać wszystkie grupy parami) wybierz tę opcję

▼ Testy post-hoc

style_przywiazania

→

przecignij wybraną zmienną do pola obok

Poprawka

☐ Bez korekty

☒ Test Tukey'a

☐ Test Scheffégo

☐ Test Bonferroniego

☐ Metoda Holma

Wielkość efektu

☐ d Cohena

☐ Przedział ufności

95 %

tutaj możesz wybrać korektę, której chcesz użyć (zaznacz wybraną opcję)

Aby obliczyć wielkość efektu, wybierz tę opcję

po wybraniu tej opcji będziesz mógł sprawdzić przedział ufności wokół efektu

Jeśli wybrałeś *post-hoc* w raporcie wyświetlą się porównania parami z wybraną korektą na poziom istotności ze względu na wielokrotne porównania.

Testy post-hoc

Porównania post-hoc – style_przywiazania

Porównanie		Różnica średnich	SE	df	t	p Tukey
style_przywiazania	style_przywiazania					
bezpieczny	- zaabsorbowany	-10.60	5.35	56.0	-1.981	0.207
	- lękowy	-24.89	6.77	56.0	-3.677	0.003
	- odrzucający	-17.19	11.10	56.0	-1.549	0.416
zaabsorbowany	- lękowy	-14.29	6.39	56.0	-2.236	0.126
	- odrzucający	-6.59	10.87	56.0	-0.606	0.930
lękowy	- odrzucający	7.70	11.64	56.0	0.661	0.911

Uwaga. Porównania są oparte na oszacowanych średnich brzegowych

Aby wyliczyć średnie brzegowe, wybierz tę opcję

następnie przeciągnij wybraną zmienną w odpowiednie miejsce

Jeżeli zaznaczysz opcję „Wykresy średnich brzegowych”, wykres pojawi się w polu obok (zaznacz tabelę, a również ona się pojawi).

Oszacowane średnie brzegowe

style_przywiazania

←

Średnie brzegowe

Składnik 1

Przeciągnij zmienną tutaj

Dodaj nowy składnik

Wyjście

☒ Wykresy średnich brzegowych
☐ Tabele średnich brzegowych

Wykres

Słupki błędów

Przedział ufności

☐ Wartości obserwowane

Opcje ogólne

☒ Równe wagi komórek

Przedział ufności 95 %

Oszacowane średnie brzegowe

style_przywiazania	Średnia	SE	95% przedział ufności
			Dolna granica Górna granica
bezpieczny	56.5	4.10	48.3 64.7
zaabsorbowany	67.1	3.44	60.2 74.0
lękowy	81.4	5.39	70.6 92.2
odrzucający	73.7	10.31	53.0 94.3

Oszacowane średnie brzegowe - style_przywiazania

style_przywiazania	Średnia	SE	95% przedział ufności	
			Dolna granica	Górna granica
bezpieczny	56.5	4.10	48.3	64.7
zaabsorbowany	67.1	3.44	60.2	74.0
lękowy	81.4	5.39	70.6	92.2
odrzucający	73.7	10.31	53.0	94.3

style_przywiazania

Warto wybierać „Wartości obserwowane” w opcjach wykresu, wtedy pojawiają się na wykresie ułatwiając interpretację wyniku.

Test Kruskala-Wallisa

Jak go przeprowadzić?

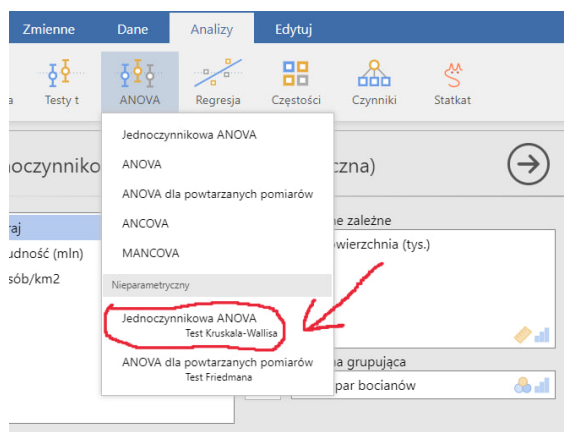
Test Kruskala-Wallisa stosowany jest do porównania mediany trzech lub więcej grup na podstawie prób niezależnych.

O to są kolejne kroki do przeprowadzenia tego testu:

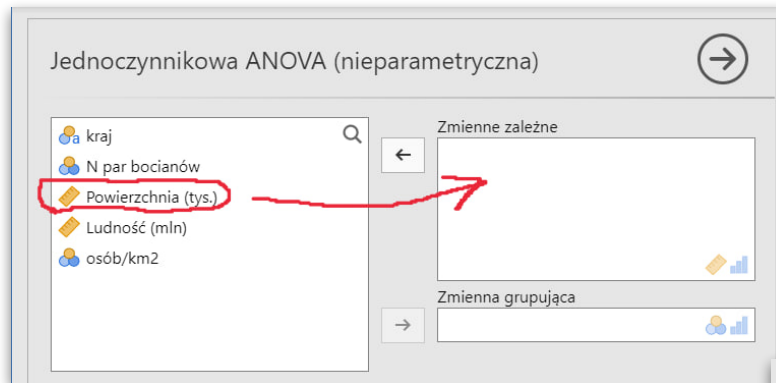
1. Zaimportować dane lub wpisać je ręcznie (po prostu je mieć)

	🌐 kraj	👤 N par boc...	📏 Powierzch...	👤 Ludność (...)	👤 osób/km2
1	Albania	20	28.7	3.1	108
2	Algeria	5147	2381.7	32.6	14
3	Austria	392	83.9	8.2	98
4	Belgia	50	30.5	10.4	341
5	Białoruś	20342	207.6	9.8	47
6	Bośnia i Herc...	40	51.1	4.5	88
7	Bułgaria	4956	110.9	7.5	68
8	Chorwacja	1700	56.6	4.5	80
9	Czechy	814	78.9	10.2	129
10	Dania	3	43.1	5.4	125
11	Estonia	2650	45.2	1.3	29
12	Finlandia	1	338.1	5.2	15
13	Francja	941	551.5	60.6	110
14	Grecja	2139	132.0	11.2	85
15	Hiszpania	33217	506.0	43.4	86
16	Holandia	528	41.5	16.3	393
17	Litwa	13000	65.3	3.4	52
18	Łotwa	10700	64.6	2.3	36
19	Maroko	5000	458.7	29.8	65
20	Mołdawia	491	33.8	3.9	115
21	Niemcy	4482	357.0	82.7	232
22	Polska	52500	312.7	38.1	122
23	Portugalia	7684	92.2	10.5	114
24	Rumunia	5500	237.5	21.4	90
25	Serbia i Czarn...	1250	88.4	10.0	113
26	Słowacja	1330	49.0	5.4	110
27	Słowenia	236	20.3	2.0	99
28	Szwajcaria	198	41.3	7.5	182
29	Szwecja	29	450.0	9.0	20

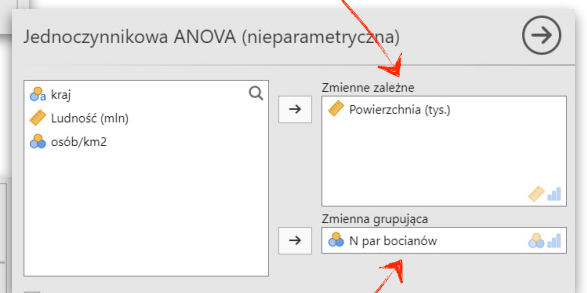
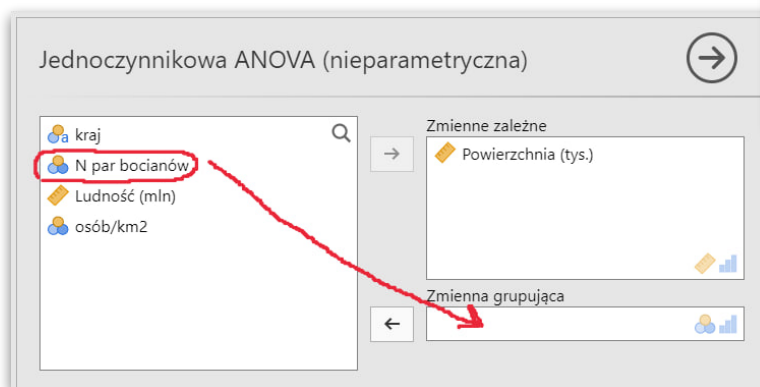
2. W rozdziale „Analizy” wybieramy „ANOVA” -> Jednoczynnikowa ANOVA Test Kruskala – Wallisa w podrozdziale „Nieparametryczny”



3. Do pola „Zmienne zależne” (Dependent Variable) przeciągnij i upuść zmienną zależną



4. Do pola „Zmienna grupująca” (Grouping Variable) przeciągnij i upuść zmienną niezależną



5. Używając testu Kruskala-Wallisa w danym przypadku chcieliśmy sprawdzić zależności ilości bocianów (N par bocianów) od powierzchni kraju (Powierzchnia (tys.)), wyrażanej w km². Raport przeprowadzonego testu wygląda w ten sposób:

Wyniki

Jednoczynnikowa ANOVA (nieparametryczna)

Test Kruskala-Wallisa

	χ^2	df	p
Powierzchnia (tys.)	33.0	33	0.467

Jak widać, test nie wykazał istotności statystycznej, czyli nie istnieje zależności między ilością bocianów a powierzchnią kraju, gdyż: $\chi^2(33) = 33,0$; $p = 0,467^1$.

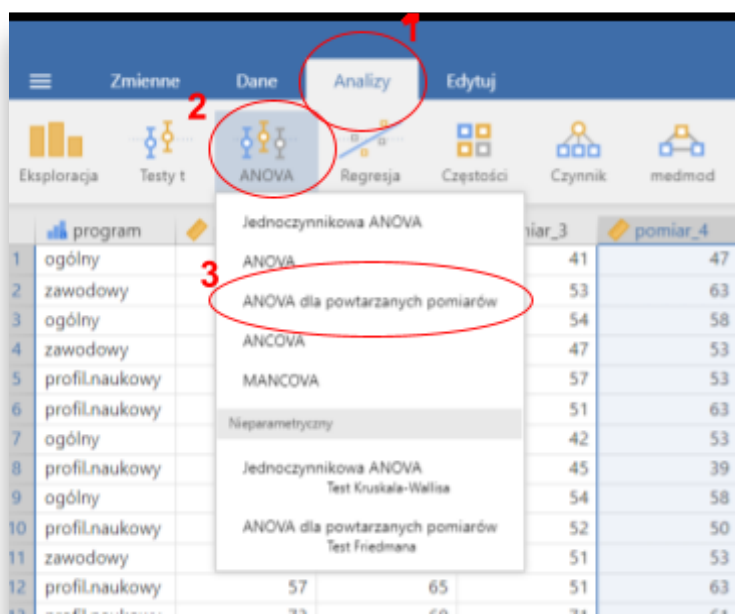
¹ Gdyby był istotny statystycznie powinniśmy podać także wielkość efektu *epsilon-kwadrat*: ϵ^2 oraz obliczyć różnice parami za pomocą testu *post hoc* DSCF, który wskaże, które grupy konkretnie się różnią między sobą. Obie opcje są dostępne w Jamovi pod okienkiem z listą zmiennych.

Analiza wariancji z powtarzanymi pomiarami, czyli **ANOVA dla powtarzanych pomiarów** przyda Ci się, kiedy masz więcej niż 3 grupy zmiennych oraz zmienne są od siebie zależne. Innymi słowy porównujesz w tym teście wyniki osób, które zostały poddane badaniu co najmniej 2-krotnie w różnych odstępach czasu, np. wyniki badań osób przed i po zastosowaniu danej terapii. Analiza wariancji dla powtarzanych pomiarów sprawdza, czy różnice między grupami są większe od różnic wewnątrz tych grup.

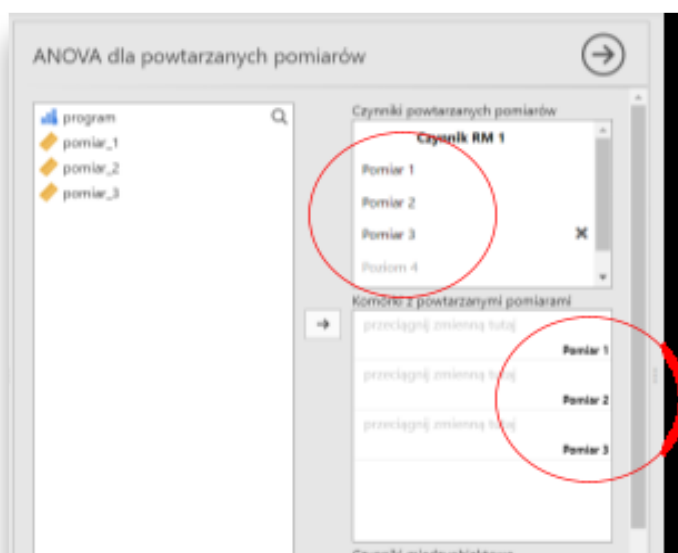
1. Upewnij się, że twoje zmienne zależne są zmiennymi ilościowymi (ciągłymi), czyli oznaczone **żółtą linijką**.

	program	pomiar_1	pomiar_2	pomiar_3
1	ogólny	57	52	
2	zawodowy	68	59	
3	ogólny	44	33	
4	zawodowy	63	44	
5	profil.naukowy	47	52	

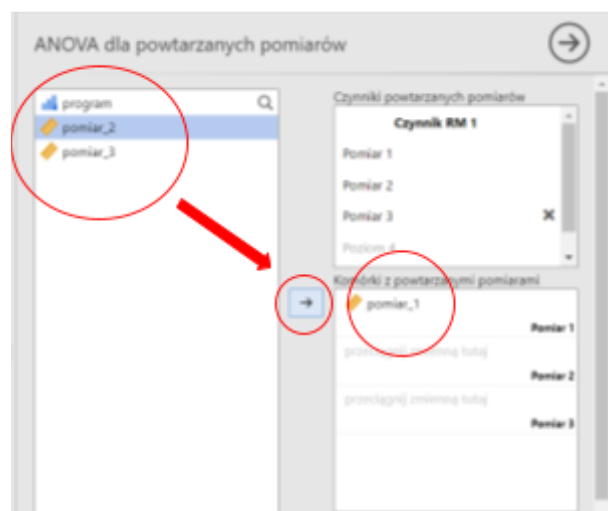
2. Wybierz opcję analizy -> ANOVA -> ANOVA dla powtarzanych pomiarów



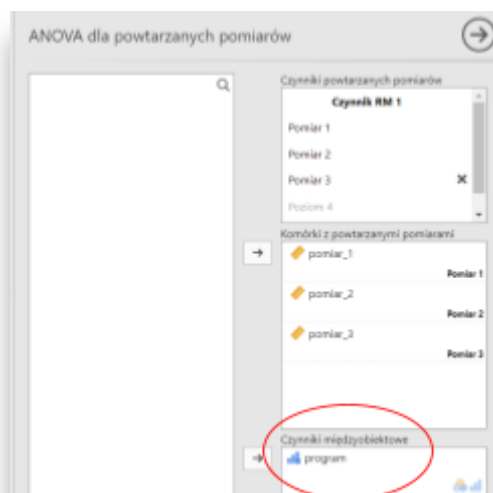
3. W oknie czynniki powtarzalnych pomiarów możesz zmienić nazwy swoich poziomów, ułatwi ci to później interpretację wyników.



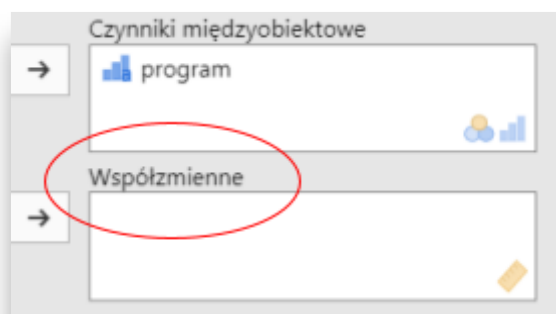
4. Przenieś wybrane zmienne do okna komórki z powtarzanymi pomiarami. Możesz to zrobić przeciągając je myszką lub klikając najpierw na zmienną, a następnie na strzałkę.



5. Niżej znajduje się okno czynniki międzyobiektywne, gdzie możesz dodać czynnik, który może wpływać na różnice między grupami, ale są niezależne od powtarzalnych pomiarów wewnątrz tych grup, jest to np. płeć, status zawodowy, wykształcenie. Musi być to zmienna porządkowa lub nominalna.



6. W oknie współzmiennie możesz wprowadzić zmienną, którą chcesz kontrolować podczas badania. Poprzez uwzględnienie współzmiennych jako dodatkowych czynników w analizie, możemy ocenić, czy różnice między grupami na pewno są wynikiem głównych zmiennych.



7. W ostatnim kroku pozostaje jedynie odczytanie i zinterpretowanie wyników testu, które po poprawnym ustawieniu zmiennych wyświetlą się po prawej stronie w formie tabeli.

Efekt powtarzanych pomiarów

Efekt interakcji między pomiarami a przynależnością do grupy: „program”

Efekt samej przynależności do grupy: „program” bez uwzględnienia powtarzanych pomiarów

ANOVA dla powtarzanych pomiarów

Efekty wewnątrzobiektywne

	Suma kwadratów	df	Średni kwadrat	F	p
Czynnik RM 1	48.2	2	24.1	0.677	0.509
Czynnik RM 1 * program	59.4	4	14.9	0.417	0.797
Reszta	14042.2	394	35.6		

Uwaga. Suma kwadratów typu 3

[3]

Efekty międzyobiektywne

	Suma kwadratów	df	Średni kwadrat	F	p
program	10835	2	5418	34.1	< .001
Reszta	31327	197	159		

Uwaga. Suma kwadratów typu 3

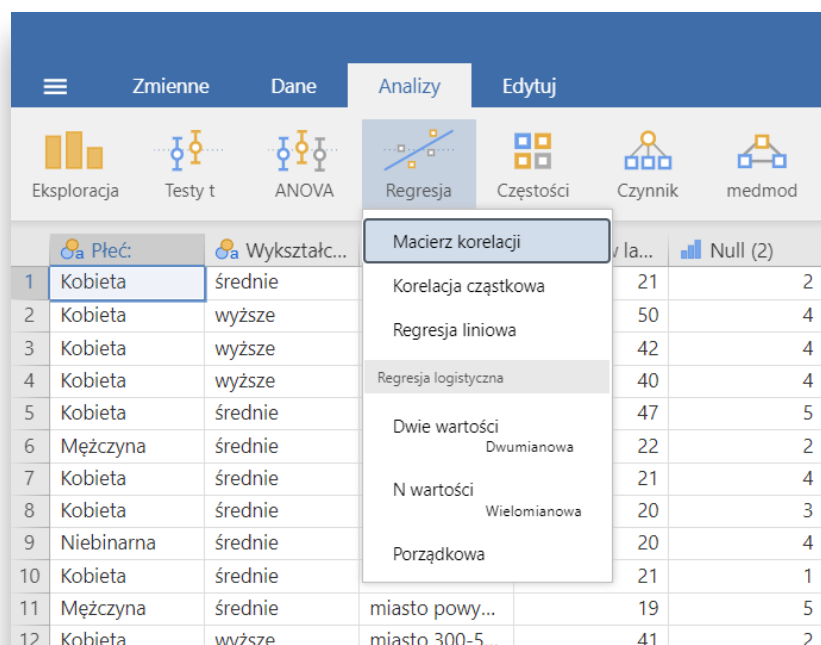
Różnice między pomiarami można opisać tak:

Wpływ czynnika „RM Factor 1” (inaczej pomiarów) na poziom zmiennej zależnej był nieistotny statystycznie: $F(2, 394) = 0,677$, $p = 0,509$.

Analiza korelacji i korelacji cząstkowej

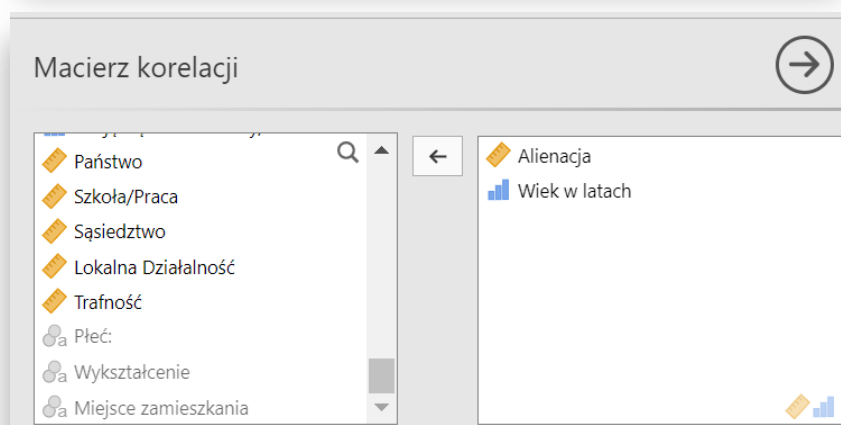
Analiza korelacji pozwala na analizowanie zależności między dwoma zmiennymi

1. Żeby zrobić analizę korelacji wchodzimy w zakładkę "Analizy" i spod ikonki "regresja" wybieramy "macierz korelacji"



2. Wybieramy zmienne, między którymi chcemy zbadać zależność

W moim przypadku badamy zależność między poczuciem alienacji a wiekiem



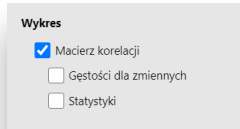
3. Wybieramy współczynnik korelacji. Pearsona wtedy, gdy obie badane zmienne są ujęte w formie liczbowej (są zmienną ciągłą) i gdy nie ma obserwacji odstających. Współczynnik Spearmana w innych przypadkach.

Współczynniki korelacji

- ☒ r Pearsona
☐ rho Spearmana

W naszym przypadku zaistniała umiarkowana korelacja ($r = 0.39$), która, jak wskazuje wartość p , mniejsza od 0.001, jest istotna statystycznie.

4. Klikając na "Macierz korelacji" pod sekcją "Wykres" możemy włączyć wykres korelacji.

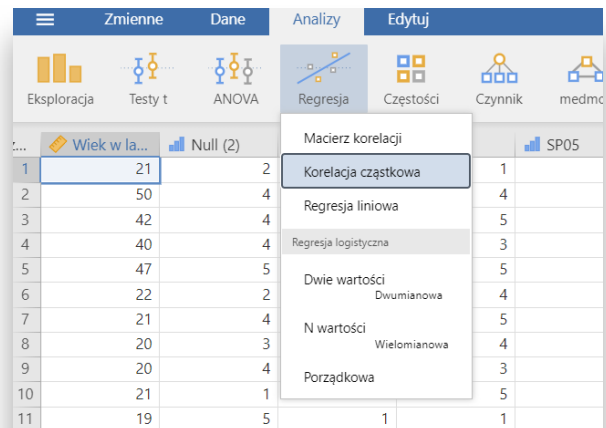


Wyniki			
Macierz korelacji			
		Alienacja	Wiek w latach
Alienacja	r Pearsona	—	
	p	—	
Wiek w latach	r Pearsona	0.39011 ***	—
	p	< .001	—
Uwaga. * $p < .05$, ** $p < .01$, *** $p < .001$			

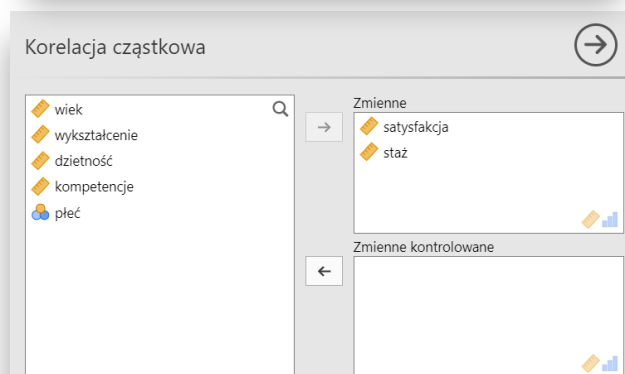
Korelacja cząstkowa

Korelację cząstkową nazywamy korelację pomiędzy parą zmiennych z uwzględnieniem związku tych zmiennych z inną, trzecią zmienną. Innymi słowy czy nadal zachodzi związek pomiędzy zmienną A i B, jeżeli wyeliminujemy z niej związek pomiędzy zmienną B i C i związek części wspólnej A i C ze zmienną B.

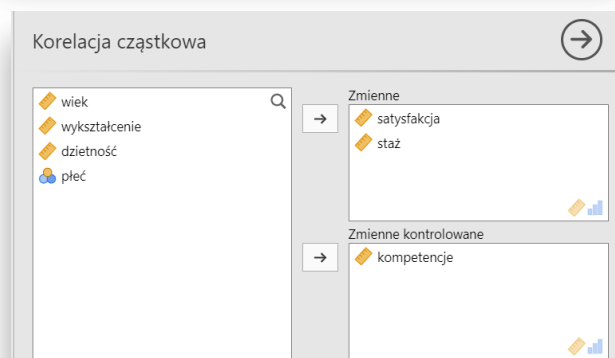
1. Żeby zrobić analizę korelacji wchodzimy w zakładkę "Analizy" i spod ikonki "regresja" wybieramy "korelacja cząstkowa"



2. Wybieramy zmienne, których korelację chcemy zobaczyć, umieszczamy je w pierwszej tabeli. W moim przypadku badamy związek satysfakcji z życia i stażu.



3. Wybieramy, którą zmienną chcemy kontrolować, czyli której zmiennej wpływ wyeliminujemy.



Uzyskujemy kontrolowaną dla trzeciej zmiennej korelację między dwiema pierwszymi zmiennymi.

W moim przypadku bardzo małą ujemną korelację między satysfakcją a stażem ($r = -0.04$), która jest nieistotna statystycznie ($p = 0.762 > 0.005$)

Wyniki				
Korelacja cząstkowa				
Korelacja cząstkowa				
			satysfakcja	staż
satysfakcja	r Pearsona		—	
	p		—	
staż	r Pearsona	-0.04034		—
	p	0.762		—
Uwaga. kontrolowane dla 'kompetencje'				
[3]				

Analiza regresji w Jamovi

Analiza regresji służy przewidywaniu danych dla konkretnej zmiennej, na podstawie innych zmiennych. Pozwala to na określenie wartości, którą przyjmie dana zmienna, gdy znamy wartość innej zmiennej.

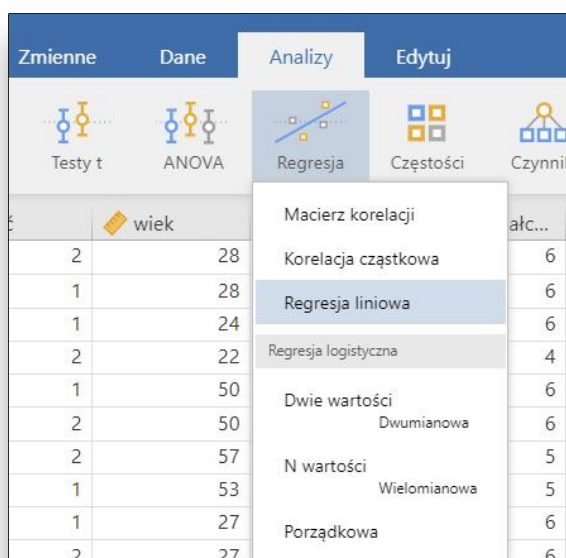
Analiza regresji może być wykorzystana np. w celu określenia, które zmienne opisujące są powiązane ze zmienną zależną.

Kroki do przeprowadzenia analizy regresji prostej/ liniowej:

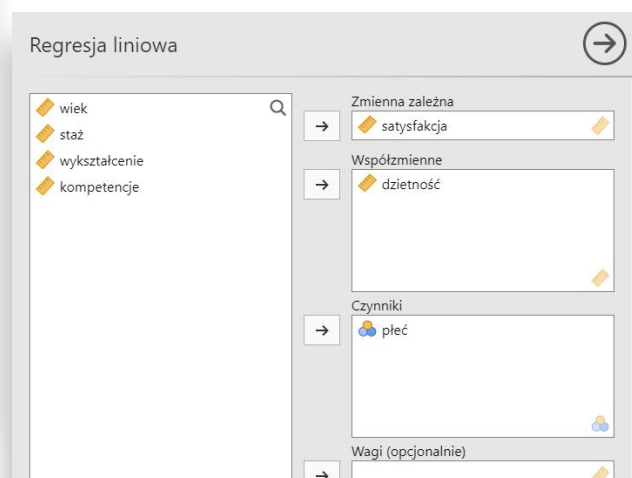
1. Wprowadzamy lub importujemy dane:

	płeć	wiek	staż	wykształc...	dzieciństwo	satysfakcja	kompeten...		
1	2	28	6	6	2	127	124		
2	1	28	6	6	2	130	126		
3	1	24	5	6	2	107	98		
4	2	22	5	4	2	98	73		
5	1	50	27	6	1	138	99		
6	2	50	27	6	1	118	97		
7	2	57	28	5	1	111	92		
8	1	53	28	5	1	102	87		

2. Z analiz wybieramy **Regresja**, a następnie **Regresja liniowa**



3. Przeciągamy wszystkie zmienne na odpowiednie miejsca



W naszym przypadku badamy zależność satysfakcji z życia od liczby posiadanych dzieci. Oznacza to, że satysfakcja jest naszą **zmienną zależną** (badaną), a dzieciństwo **zmienną niezależną** (współzmienną). Dodawanie danych w ramce **czynniki** nie jest konieczne, ale możemy to zrobić np. gdy chcemy sprawdzić, czy płeć też jest czynnikiem i różnicuje wyniki.

4. Chcieliśmy zbadać zależność satysfakcji z życia (satysfakcja) od liczby posiadanych dzieci (dzietność) i płci osób badanych.

Wyniki		
Regresja liniowa		
Miary dopasowania modelu		
Model	R	R ²
1	0.2222	0.04936

Model Coefficients - satysfakcja				
Predyktor	Oszacowanie	SE	t	p
Wyraz wolny *	110.100	6.116	18.0018	< .001
dzietność	5.833	3.915	1.4898	0.142
płeć:				
2 - 1	3.300	3.836	0.8602	0.393

* Reprezentuje poziom referencyjny

Przeprowadzona przez nas analiza regresji wykazała istnienie niewielkiej korelacji ($R = 0,2222$). Model regresji jednak nie jest informatywny, ponieważ tylko wyraz wolny jest istotny statycznie, co oznacza, że dzietność i płeć nie wyjaśniają zmienności w poziomie satysfakcji u osób badanych.

Gdyby wynik był istotny moglibyśmy go opisać w taki sposób:

Analiza regresji wykazała statystycznie istotny model przewidujący satysfakcję z życia ($R^2 = 0,049$; $F(2, 100) = 18,72$; $p < 0,001$). Czynniki dzietność i płeć miały istotny wkład w wyjaśnienie zmienności satysfakcji z życia.

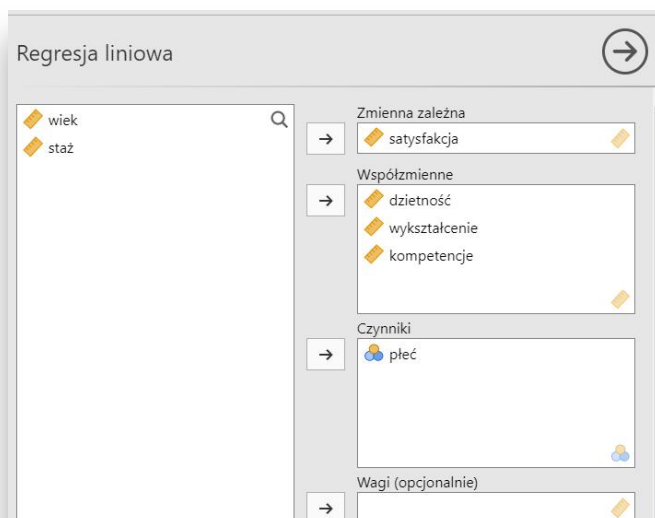
Wskaźniki regresji wykazały, że dzietność była istotnym predyktorem satysfakcji z życia ($B = 5,83$; $SE = 3,92$; $t = 4,13$; $p < 0,001$). Wyższa dzietność była związana ze wzrostem satysfakcji z życia. Płeć osoby badanej również miała istotny wpływ na satysfakcję z życia ($B = 3,30$; $SE = 3,84$; $t = 2,35$; $p = 0,021$). Osoby płci żeńskiej (oznaczone 1) wykazywały wyższy poziom satysfakcji z życia w porównaniu do osób płci męskiej.

Wyższa dzietność jest związana z większą satysfakcją z życia, podczas gdy bycie płci żeńskiej wiąże się z nieco wyższym poziomem satysfakcji z życia. Jednakże, warto zauważyć, że model regresji wyjaśniał tylko 4,9% zmienności satysfakcji z życia, co sugeruje, że inne czynniki mogą również odgrywać rolę w tej kwestii.

Kroki do przeprowadzenia analizy regresji wielorakiej:

Z regresji wielorakiej możemy skorzystać, gdy chcemy porównać związek zmiennej zależnej z wieloma zmiennymi niezależnymi (współzmiennymi). Ponownie będziemy sprawdzać korelację między satysfakcją z życia, a ilością posiadanych dzieci, jednak dodamy do zmiennych niezależnych (współzmiennych) także wykształcenie i kompetencje, sprawdzimy więc korelację satysfakcji z życia z innymi wybranymi zmiennymi.

Przeciągamy wszystkie zmienne na odpowiednie miejsca



W wynikach powinniśmy podawać wartość R_{adj}^2 , czyli skorygowaną wartość R^2 . Natomiast ocenę istotności statystycznej modelu regresji otrzymamy jeśli w Jamovi włączymy **Test ogólny modelu** [Overall Model Test]

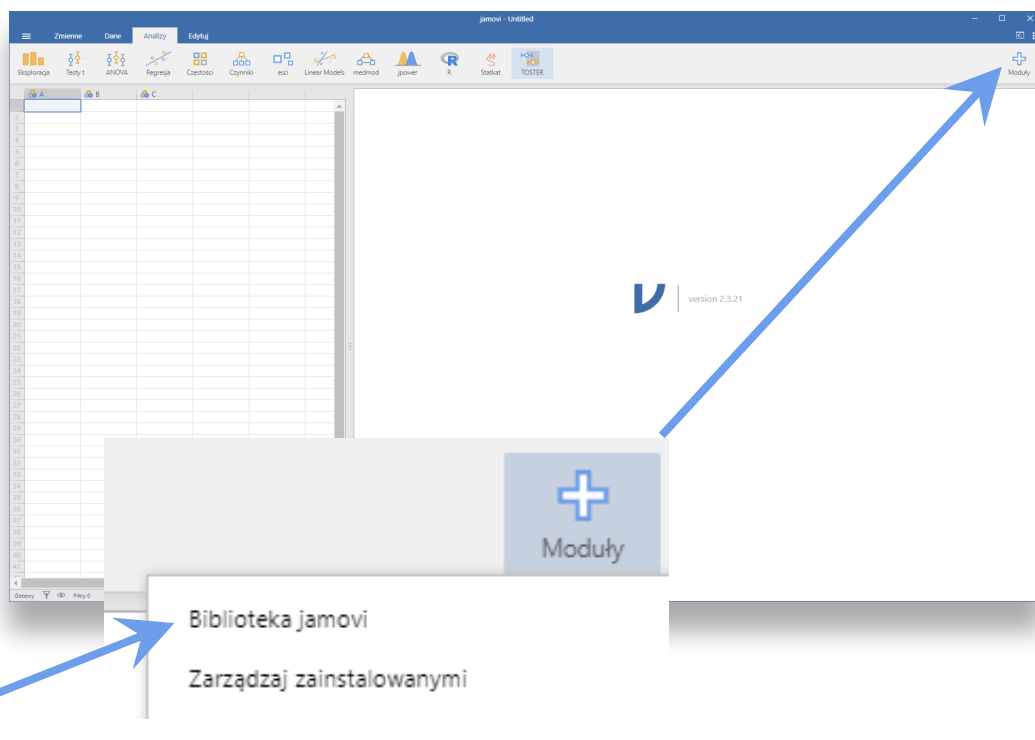
Test ekwiwalencji dla $H_0 \neq 0$

Test badający hipotezę o braku różnicy (TOSTER)

Jeśli chcemy udowodnić, że grupy się nie różnią – czyli hipoteza o braku efektu jest prawdziwa, musimy założyć, że się różnią (to nasza hipoteza zerowa) i zebrać dowody na to, że tak nie jest. Zazwyczaj (w NHST) robimy na odwrót – testujemy hipotezę zerową o braku różnic. Aby **udowodnić braku różnic** w Jamovi musimy najpierw pobrać odpowiedni moduł (Jamovi domyślnie nie ma go zainstalowanego).

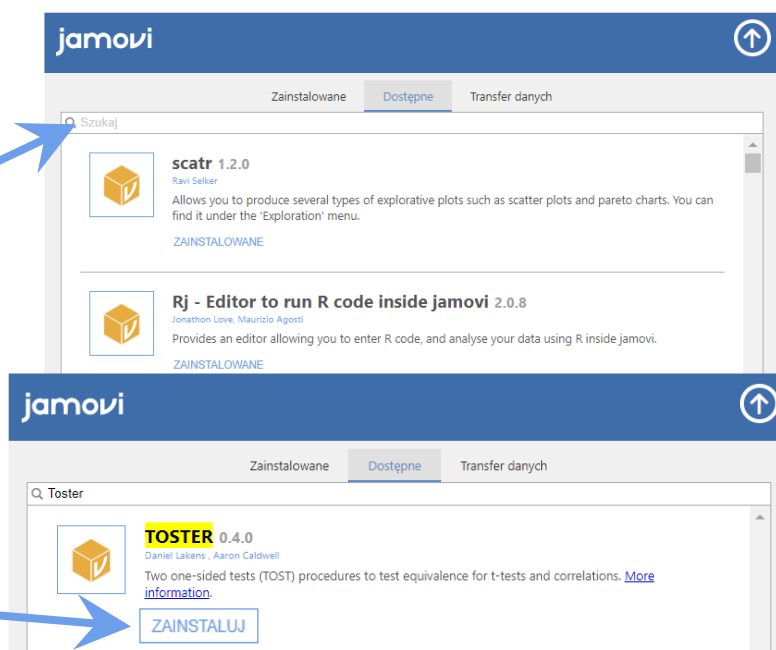
Instalowanie modułu

Klikamy w prawym, górnym rogu w biały plus z niebieskimi konturami, podpisany “Moduły”



Następnie klikamy w “Biblioteka jamovi” możemy w niej pobrać też różne inne moduły.

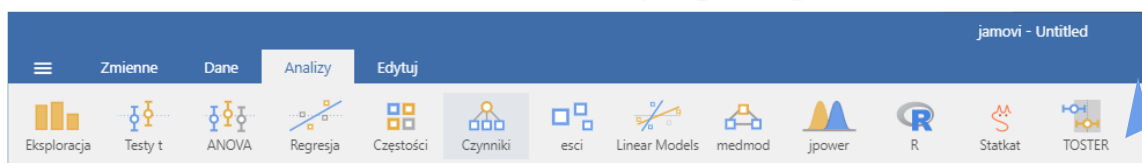
Gdy już to zrobimy pokaże nam się okno z dostępnymi modułami do pobrania. W pasku wyszukiwania wpisujemy “TOSTER”



Gdy już to zrobimy klikamy w przycisk “ZAINSTALUJ”

Teraz musimy chwilę poczekać, aż moduł się zainstaluje.

Ikona zainstalowanego modułu wyświetli nam się na pasku pozostałych testów/modułów.



Przykład

Chcemy wykazać, że płeć nie różnicuje wyników ekstrawersji.

Musimy więc zbadać hipotezę - Płeć różnicuje ekstrawersję.

Włączamy test ekwiwalencji w Jamovi, w tym przypadku dla zmiennych niezależnych (Independent Samples T-Test). Jako zmienną (Variable) wstawiamy "ekstrawersja", a zmienna grupująca (Grouping Variable) "płeć".

TOST Independent Samples T-Test

losób
miej.zam
droga
czas.dr
ojciec
matka
l.rodze
relacje

Variable
→ ekstrawersja

Grouping Variable
→ płeć

☐ Assume equal variances

☒ Descriptive statistics

☒ Plot Effect Sizes

☐ Plot Data

Hypothesis

☒ Equivalence Test

☐ Minimal Effects Test

Bound type

☒ Raw

☐ Standardized

Lower bound -0.5

Upper bound 0.5

Alpha level 0.05

Standardized Effect Size

☐ Hedges' g

☒ Cohen's d

TOST Independent Samples T-Test

Hypothesis Tested: Equivalence

Null Hypothesis: $-0.5 \geq (\text{Mean}_1 - \text{Mean}_2) \text{ or } (\text{Mean}_1 - \text{Mean}_2) \geq 0.5$

Alternative: $-0.5 < (\text{Mean}_1 - \text{Mean}_2) < 0.5$

✗ NHST: don't reject null significance hypothesis that the effect is equal to zero

✓ TOST: reject null equivalence hypothesis

Note: SMD confidence intervals are an approximation. See our [documentation](#).

TOST Results

		t	df	p
ekstrawersja	t-test	-1.08	602	0.281
	TOST Upper	5.57	602	< .001
	TOST Lower	-7.72	602	< .001

Uwaga. Welch's t-test

Equivalence Bounds

		Low	High
ekstrawersja	Cohen's d(av)	-0.527	0.527
	Raw	-0.500	0.500

Effect Sizes

		Estimate	90% Confidence Interval	
			Lower	Upper
ekstrawersja	Cohen's d(av)	-0.0855	-0.216	0.0448
	Raw	-0.0811	-0.205	0.0428

Uwaga. Denominator set to the average SD

Descriptives

	N	Mean	Median	SD	SE
chłopiec	283	2.95	2.96	0.956	0.0568
dziewczynka	365	3.03	3.04	0.942	0.0493

Uwaga! - Użyłam domyślnych ustawień Jamovi dla rodzaju granic. Granice można podawać albo w punktach (RAW → Δ), albo w wielkości efektu (STANDARDIZED → d). W tym przypadku 0.5 dotyczy różnicy w punktach ekstrawersji.

Gdy włączymy "Plot Effect Sizes" otrzymamy wykres.

ekstrawersja

Confidence Interval 0.68 0.9 0.95 0.999

-0.53 0.53

Cohen's d(av)

Test równoważności wykazał, że różnica w ekstrawersji między grupami jest statystycznie nieistotna, $t(602) = 1.08$, $p = .281$, a dla minimalnego efektu $d = 0.527$ wyniki obu grup są równoważne: $t(602) = 5.57$, $p < .001$, $d = 0.09$, 90% CI [-0.22, 0.04].

Interpretacja wyniku:

Wynik testu równoważności wskazuje, że grupy nie różnią i są statystycznie równoważne pod względem ekstrawersji.

Pojęcia z zakresu metodologii i statystyki

Pojęcie	Symbol	Znaczenie / definicja
Mediana	Md	wartość środkowa, dzieli zbiór wartości na dwa równoliczne zbiory (np. w ciągu liczb: 1, 1, 3, 4, 5, 5, 5, 8, 9 mediana wynosi 5 i po obydwu jej stronach jest tyle samo cyfr (wartości))
Odchylenie standardowe	SD	miara rozproszenia; mówi o tym, jak bardzo wartości danej wielkości są rozproszone wokół jej średniej; o ile średnio odchylają się wartości badanej cechy od jej średniej arytmetycznej; pomaga nam stwierdzić, czy w danej populacji jednostki są podobne ze względu na badaną cechę, czy znacznie się od siebie różnią; im wartość odchylenia mniejsza, tym mniej się różnią
Rozkład normalny	$\sim N(\mu, \sigma)$	Rozkład normalny, nazywany także rozkładem Gaussa lub krzywą dzwonową, jest jednym z najważniejszych rozkładów prawdopodobieństwa w statystyce. Charakteryzuje się on tym, że ma kształt symetrycznej krzywej dzwonowatej, którą określa dwie parametry: średnią (oznaczaną jako μ) i odchylenie standardowe (oznaczane jako σ). W rozkładzie normalnym większość obserwacji skupia się wokół średniej wartości, a im bardziej oddalone wartości od średniej, tym mniej prawdopodobne ich wystąpienie. Rozkład normalny jest bardzo ważny w statystyce, ponieważ wiele zjawisk w przyrodzie i społeczeństwie ma rozkład zbliżony do normalnego. Dlatego rozkład normalny jest często stosowany w analizie danych, aby opisać i prognozować zachowanie zmiennych losowych. Ponadto wiele testów statystycznych opiera się na założeniu, że dane mają rozkład normalny, co umożliwia dokładne przewidywanie wyników analiz.
Średnia	M	Miara tendencji centralnej; średnia arytmetyczna to suma zbioru liczb podzielna przez ich ilość. Średniej nie można wyznaczyć w zmiennej porządkowej i nominalnej.
Wariancja	V	Wariancja to miara ogólnej zmienności danych, możliwa do obliczenia tylko dla zmiennych o ilościowym poziomie pomiaru. Przyjmuje ona wartości od 0 do + nieskończoności, gdzie 0 to brak zróżnicowania wyników.
Istotność statystyczna	p	Ocena, czy wynik wystąpił przypadkowo. Wynik istotny statystycznie = małe prawdopodobieństwo przypadku. Przyjętym poziomem istotności (granica) jest $p < 0.05$

Pojęcie	Symbol	Znaczenie / definicja
p value	p	Wartość prawdopodobieństwa wystąpienia danego zdarzenia przypadkowo. Może osiągać wartości od 0 do 1, przy czym $p < 0.05$ jest istotne statystycznie, bo wtedy mamy prawdopodobieństwo przypadkowości mniejsze od 5%
Parametr	–	
Estymacja	–	Estymacja to szacowanie pewnych wartości (np. średniej, odchylenia standardowego) w całej populacji na podstawie części tej grupy (próby). Estymacja pozwala nam na uogólnienie wyników, stosowane np. w sondażach
Przedział ufności	CI	Przedział ufności dla danej miary statystycznej informuje nas “na ile możemy ufać danej wartości”, pokazuje nam, że poszukiwana przez nas, rzeczywista wartość mieści się w pewnym przedziale z określonym prawdopodobieństwem. Przykład: CI 95% = [-0,131; 0,209] interpretujemy jako: w 95 przypadkach na 100 w przedziale <-0,131; 0,209> mieści się rzeczywista (występująca w populacji) wartość.
Błąd pomiarowy	SEM	Błąd standardowy średniej (SEM, ang. Standard Error of the Mean) jest miarą rozrzutu próby średniej, która określa, jak bardzo różne wartości próby mogą różnić się od prawdziwej wartości populacji. SEM jest zdefiniowany jako odchylenie standardowe (SD) próby podzielone przez pierwiastek z liczby obserwacji (n) w próbie. Innymi słowy, jest to estymator odchylenia standardowego populacji i jest używany do obliczania przedziałów ufności dla średniej próby oraz wykorzystywany w kilku testach statystycznych (np. t-Studenta).
Wielkość efektu	d, eta, omega, ...	Jest to ilościowa miara siły zjawiska, która jest obliczana na podstawie danych. Stosuje się go do mierzenia siły związku między zmienną zależną i zmienną niezależną, czyli wpływu danego czynnika na wynik ogólny grupy.
Moc testu	1-beta	Moc testu ($1 - \beta$) to prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona fałszywa. Jest to prawdopodobieństwo niepopelnienia błędu drugiego rodzaju. Im większa jest moc testu, tym mniejsze jest ryzyko przyjęcia fałszywej hipotezy zerowej. W praktyce, moc testu oznacza prawdopodobieństwo podjęcia na podstawie testu decyzji o istnieniu efektu, gdy on rzeczywiście istnieje.

Pojęcie	Symbol	Znaczenie / definicja
Skośność	SKE	Skośność wyników egzaminu wynosi -1,39, co wskazuje, że rozkład jest lewoskośny. Skośność równa 0 oznacza symetryczny rozkład; większa niż 0 oznacza, że rozkład jest skośny w prawo (dłuższy ogon po prawej stronie); mniejsza niż 0 oznacza, że rozkład jest skośny w lewo (dłuższy ogon po lewej stronie).
Kurtoza	K	Kurtoza wyników egzaminu wyniosła 4,17, co wskazuje, że rozkład był bardziej skoncentrowany wokół średniej w porównaniu z rozkładem normalnym (ponieważ jest większy niż 3). Kurtoza równa 0 oznacza, że rozkład ma standardową kurtozę; większa niż 0 oznacza, że rozkład ma wyższą koncentrację wokół środka; mniejsza niż 0 oznacza, że rozkład ma niższą koncentrację wokół środka. Rozkłady, których kurtoza jest większa od 3 nazywa się rozkładami "szpiczastymi" (leptokurtycznymi), a rozkłady, których kurtoza jest mniejsza od 3 nazywa się rozkładami "płaskimi" (platykurtycznymi).
Rozstęp	–	Miara rozproszenia, rozrzutu. Różnica między wartością maksymalną, a minimalną danych. Często używany do wieku i dochodu. Określa wówczas jak duża różnica była między osobą najmłodszą a najstarszą lub o najniższych dochodach i najwyższych.
Dominanta	Mo	Zwana też modą. Miara tendencji centralnej. Oznacza wartość danego zbioru pojawiającą się najczęściej. Może być wykorzystywana przy każdych zmiennych. (nominalnych, porządkowych, ilościowych)
Zmienna porządkowa	–	Zmienne, które można ułożyć w określonym porządku, ale nie można określić o ile jednostek się od siebie różnią (używamy określeń: większe, mniejsze) np. wielkość miejsca zamieszkania, wykształcenie
Zmienna ilościowa	–	Zmienne, które można ułożyć na skali oraz dokładnie określić ich wartość. Możemy określić ich minimum, maksimum, medianę, dominantę, odchylenie standardowe, wariancję, skośność i kurtozę oraz standardowy błąd pomiaru. np. wiek, czas przebiegnięcia dystansu, natężenie cechy psychologicznej
Zmienna nominalna	–	Zmienne, które można przyporządkować do określonej kategorii na zasadzie - należy, nie należy; Nie można wykonywać na nich działań (wyliczać średniej, mediany) np. płeć, style przywiązania, marka samochodu, nazwy miast
Zmienna wyjaśniana	–	Inaczej zmienna zależna - zmienna, która jest przedmiotem badania i której związek z innymi zmiennymi badacz stara się określić (np. Poddawanie się sugestiom autorytetu w eksperymencie Milgrama z rażeniem prądem)

Pojęcie	Symbol	Znaczenie / definicja
Regresja	R ²	Analiza, która pozwala przetestować związek między zmiennymi ilościowymi. Można potraktować ją jako rozszerzenie analizy korelacji - korelacja dotyczy jedynie związku między parą zmiennych, a regresja liniowa pozwala na wprowadzenie wielu zmiennych wyjaśniających, innymi słowy, daje nam możliwość na przewidywanie wartości zmiennej zależnej na podstawie większej ilości zmiennych. Jeszcze inaczej, w regresji liniowej, interpretacja współczynników pozwala zrozumieć to, w jaki sposób szczególny układ zmiennych wyjaśnia zmienność wprowadzonej do modelu zmiennej niezależnej.
Błąd standardowy	SE	Jest to oszacowanie mówiące o tym, jak bardzo wartość statystyki testowej różni się w zależności od próbki - czyli jest to miara niepewności testu statystycznego.
Rangowanie	–	Polega na zastąpieniu konkretnej zmiennej przez wyliczoną dla niej rangę. Najprostszym sposobem rangowania jest numerowanie ich rosnąco, czyli od 1 wzwyż. Rangowanie jest szczególnie przydatne, gdy pracujemy z kilkoma wynikami o takiej samej wartości.
Kurtoza	–	Określa rozmieszczenie i koncentrację wartości. Im wyższa jest kurtoza, tym większe jest skupienie wartości wokół średniej. Dla rozkładu normalnego przyjmuje się wartość kurtozy = 3
Homoskedastyczność	–	Warunkowa jednorodność rozkładu wariancji – dla poszczególnych wartości zmiennej niezależnej obserwujemy zbliżoną wariancję zmiennej zależnej (dla wszystkich wartości X mamy takie same “choinki” w Y).