



Ćwiczenie: Weryfikacja hipotez statystycznych dla jednej i dwóch średnich.

EXCEL

Do weryfikacji różnic między dwiema grupami jednostek doświadczalnych w Excelu wykorzystujemy funkcję o nazwie **T.TEST**. Zastosowana funkcja (test statystyczny) pozwala ocenić czy mamy podstawy do odrzucenia H_0 na korzyść hipotezy alternatywnej H_1 czy też podstaw tych nie ma. Informację tę otrzymujemy w formie prawdopodobieństwa popełnienia błędu I rodzaju (poziom istotności). Błąd ten informuje o tym, jakie jest prawdopodobieństwo odrzucenia H_0 , gdy jest ona prawdziwa. Tak też należy traktować wartości liczbowe będące skutkiem zadziałania funkcji **T.TEST**.

Składnia funkcji

- **T.TEST(tablica1;tablica2;ślady;typ)**
- **Tablica1** – pierwsza grupa obserwacji.
- **Tablica2** – druga grupa obserwacji.
- **Ślady** – określenie czy użyty test statystyczny ma być jednostronny (1) czy dwustronny (2).
- **Typ** - jest to rodzaj testu t, który ma być przeprowadzony:

Jeżeli typ równa się	Wariant testu
1	Sparowany
2	Wariancja równa dla dwóch prób (homoscedastyczna)
3	Wariancja nierówna dla dwóch prób (heteroscedastyczna)

Wartości krytyczne rozkładu t-Studenta można otrzymać w wyniku zastosowania funkcji:
=ROZKŁ.T.ODWR.DS (α; v)

Przykład 1 (stopy2009l.xls, arkusz SAS). Sprawdź czy kobiety i mężczyźni różnią się pod względem wzrostu, długości stopy oraz „indeksu” stopy. W tym celu wykorzystaj funkcję **T.TEST**, a ponadto moduł **Analiza danych**.

1. Obliczamy średnie arytmetyczne w odniesieniu do obu płci i wszystkich zmiennych.

Średnie	k	166.25	23.60	14.20
	m	180.63	26.59	14.72

2. Zanim przejdziemy do porównania średnich musimy sprawdzić czy zmienność w porównywanych grupach jest podobna czy też różna – czy wariancje są równe? Stosujemy test Fishera (**F.TEST**) =**F.TEST(E2:E88;E89:E115)**. Jeżeli wynik formuły jest mniejszy lub równy 0,05, to wnioskujemy że wariancje (zmienność) są nierówne. Informacja ta jest niezbędna do podjęcia decyzji o rodzaju testu t.

Hipoteza zerowa o równości wariancji w porównywanych populacjach posiada następującą postać:

$$H_0: \sigma_1^2 = \sigma_2^2$$

zaś alternatywna zakładająca różnice w zakresie zmienności:

$$H_1: \sigma_1^2 \neq \sigma_2^2$$



W MS EXCEL wykonujemy następujące obliczenia:

	Test F	=F.TEST(E2:E88;E89:E115)	0.745031382	0.664266133
--	--------	--------------------------	-------------	-------------

Przeprowadzona analiza statystyczna dowodzi, iż nie ma podstaw do odrzucenia H_0 o równości wariancji w odniesieniu do wszystkich badanych cech. Przyjmujemy zatem, że zmienność w populacji kobiet i mężczyzn jest statystycznie podobna.

3. Weryfikujemy hipotezę zerową o równości wartości oczekiwanych porównywanych populacji kobiet i mężczyzn:

$$H_0: \mu_1 = \mu_2;$$
$$H_1: \mu_1 \neq \mu_2$$

W MS EXCEL wykonujemy następujące obliczenia:

Argumenty funkcji

T.TEST

Tablica1 F2:F88 = {24;21;23;23;24;24;24;24;23;24;24;24}

Tablica2 F89:F115 = {25;27;25.5;28.5;27;25;27.5;27;28;...}

Ślady 2 = 2

Typ 2 = 2

= 1.53068E-20

Zwraca prawdopodobieństwo związane z testem t-Studenta.

Typ - określa rodzaj testu t: sparowany = 1, z dwiema próbkami o równej wariancji = 2, z dwiema próbkami o nierównej wariancji = 3.

Wynik formuły = 1.53068E-20

Pomoc dotycząca tej funkcji

OK Anuluj

	Test T	=T.TEST(E2:E88;E89:E115;2;2)	0.0000000	0.0000483
--	--------	------------------------------	-----------	-----------

4. **Interpretacja wyników.** Przeprowadzone badania pozwalają odrzucić H_0 traktującą o równości wartości oczekiwanych w odniesieniu do wzrostu, długości stopy i indeksy stopy. Wnioskujemy zatem, że płeć wpływa statystycznie bardzo wysoko istotnie na wyżej wymienione wymiary. Pomiędzy wymiarami kobiet i mężczyzn występują różnice statystycznie bardzo wysoko istotne.

Jaka różnica?????

Prawdopodobieństwo	Określenie istotności różnic
$p > 0,05$	nieistotna
$p \leq 0,05$	istotna
$p \leq 0,01$	wysoko istotna

$p \leq 0,001$ bardzo wysoko istotna

Moduł Analiza danych

	A	B	C	D	E	F	G	H	I	J	K
1	lp	ROK	studia	plec	wzrost	stopa	IndStopy				
2	1	2006	z	k	165	24	14.55		Test t: z dwiema próbami zakładający nierówne wariancje		
3	2	2006	z	k	160	21	13.13				
4	3	2006	z	k	162	23	14.20			Zmienna 1	Zmienna 2
5	4	2006	z	k	160	23	14.38		Średnia	166.25	180.1818182
6	6	2006	z	k	163	24	14.72		Wariancja	34.25869565	34.24524313
7	7	2006	z	k	171	24	14.04		Obserwacje	116	44
8	8	2006	z	k	166	24	14.46		Różnica średnich wg hipotezy	0	
9	9	2006	z	k	168	24	14.29		df	78	
10	10	2006	z	k	162	23	14.20		t Stat	-13.44558046	
11	11	2006	z	k	162	24	14.81		P(T<=t) jednostronny	2.60736E-22	
12	12	2006	z	k	165	24	14.55		Test T jednostronny	1.664624645	
13	13	2006	z	k	170	24.5	14.41		P(T<=t) dwustronny	5.21472E-22	
14	14	2006	z	k	155	24	15.48		Test t dwustronny	1.990847036	
15	15	2006	z	k	171	26	15.20				
16	16	2006	z	k	162	24	14.81				
17	17	2006	z	k	167	25	14.97				
18	21	2006	z	k	175	23	13.14				
19	22	2006	z	k	165	24	14.55				
20	23	2006	z	k	158	22	13.92				
21	24	2006	dz	k	182	26	14.29				
22	25	2006	dz	k	164	23	14.02				
23	26	2006	dz	k	150	20.5	13.67				
24	27	2006	dz	k	170	26.5	15.59				
25	28	2006	dz	k	167	23.5	14.07				
26	29	2006	dz	k	164	23	14.02				
27	30	2006	dz	k	166	24	14.46				
28	32	2006	dz	k	167	23.5	14.07				
29	33	2006	dz	k	160	23	14.38				
30	34	2006	dz	k	176	24	13.64				
31	35	2006	dz	k	164	23	14.02				

Test t: z dwiema próbami zakładający nierówne wariancje

Wejście

Zakres zmiennej 1: \$E\$2:\$E\$117

Zakres zmiennej 2: \$E\$118:\$E\$161

Różnica średnich wg hipotezy: |

Tyluły

Alfa: 0.05

Opcje wyjścia

Zakres wyjściowy: \$I\$2

Nowy arkusz:

Nowy skoroszyt

OK Anuluj Pomoc

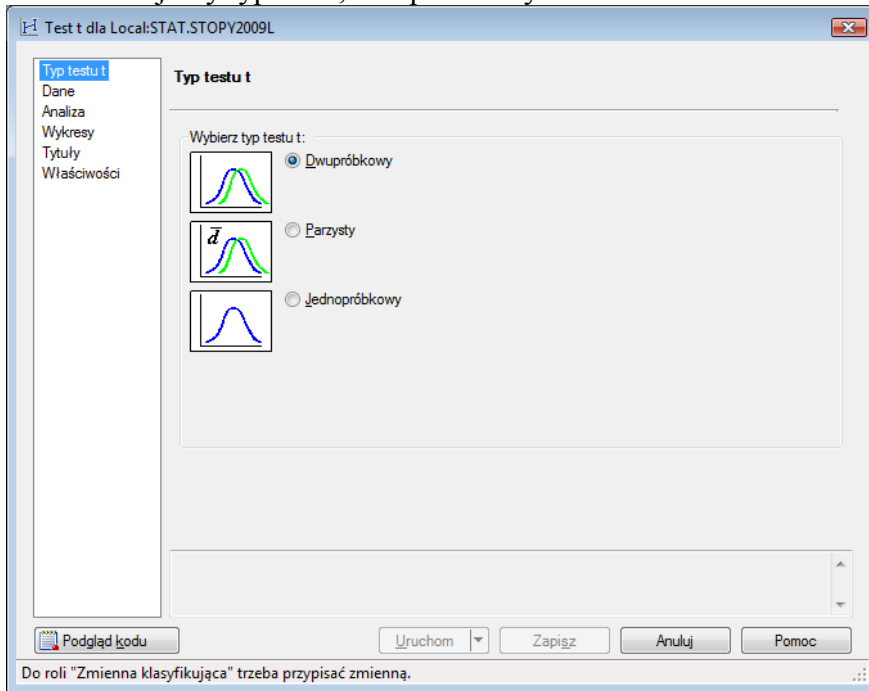
Wynik należy odczytać dla P(T<=t) dwustronny

Przykład 2. SAS Enterprise Guide PL, Ten sam problem w SAS

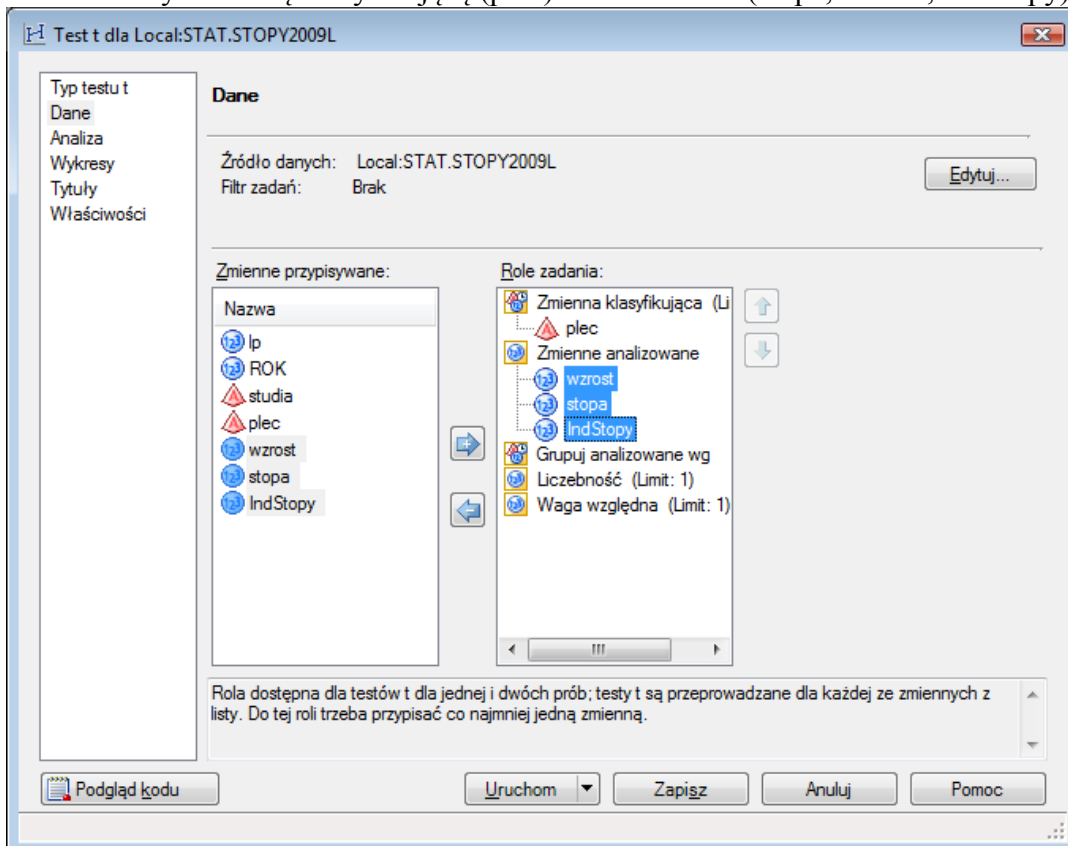
1. Uruchamiamy moduł dotyczący porównań dwóch grup.

The screenshot shows the SAS Enterprise Guide PL interface. The 'Zadania' (Tasks) menu is open, and the 'ANOVA' option is selected. The 'Test t...' dialog box is displayed, showing the same settings as the one in the previous image. The background shows a grid with data points and a 'Budowa zapytań' (Build queries) button.

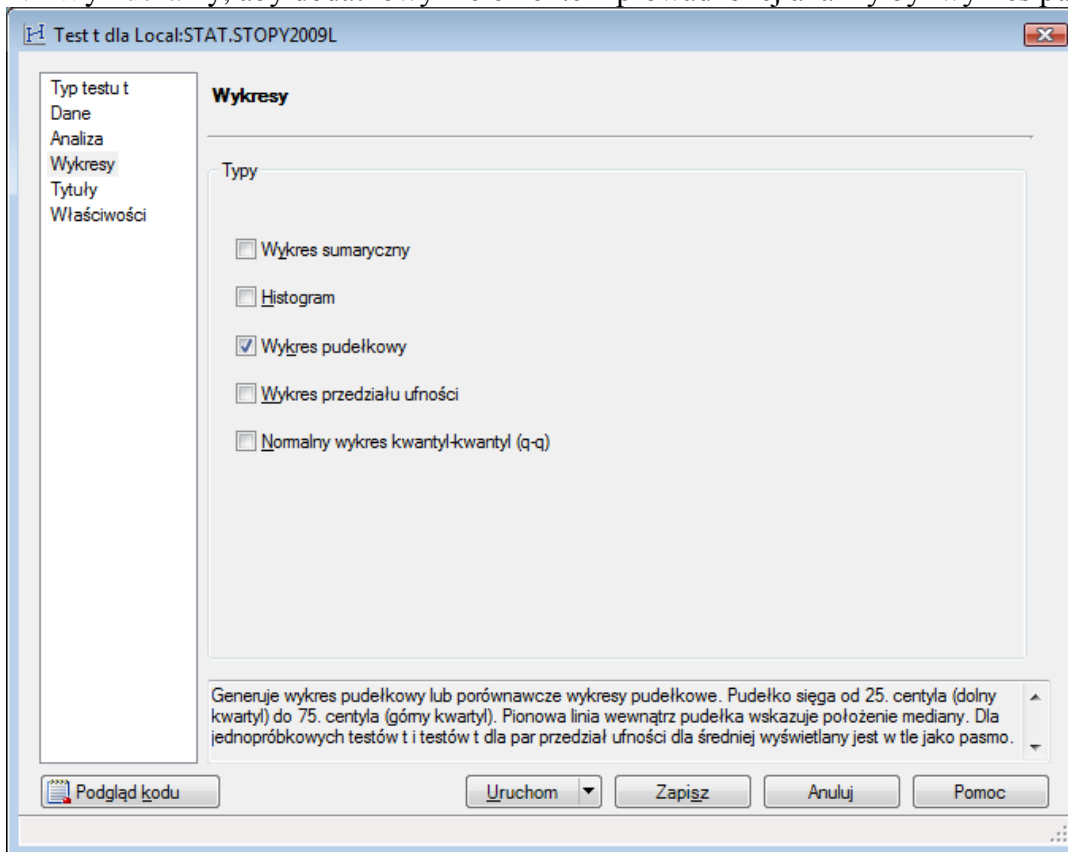
2. Wskazujemy typ testu, dwupróbkowy



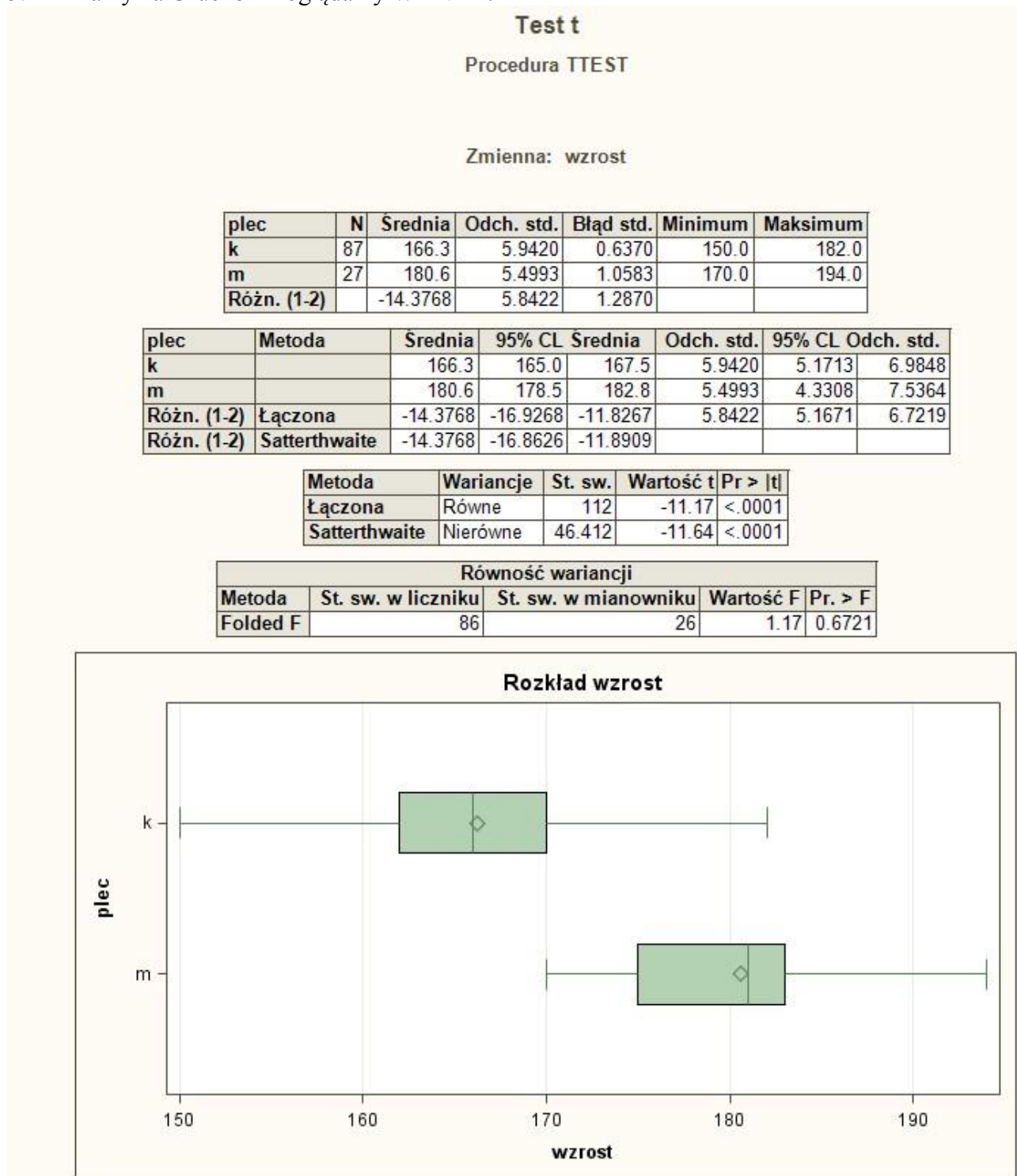
3. Ustalamy zmienną klasyfikującą (płeć) i analizowane (stopa, wzrost, IndStopy).



4. Wymuszamy, aby dodatkowym elementem prowadzonej analizy był wykres pudełkowy.

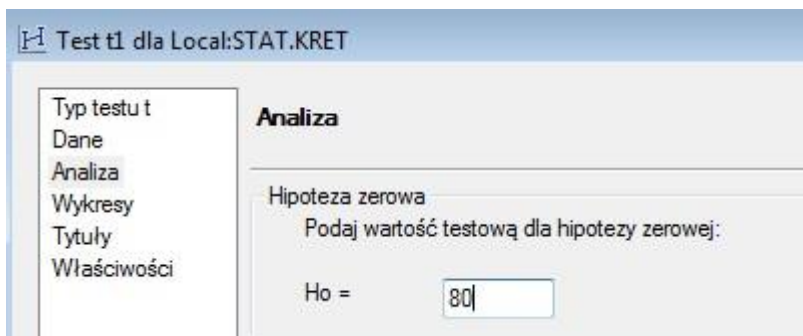
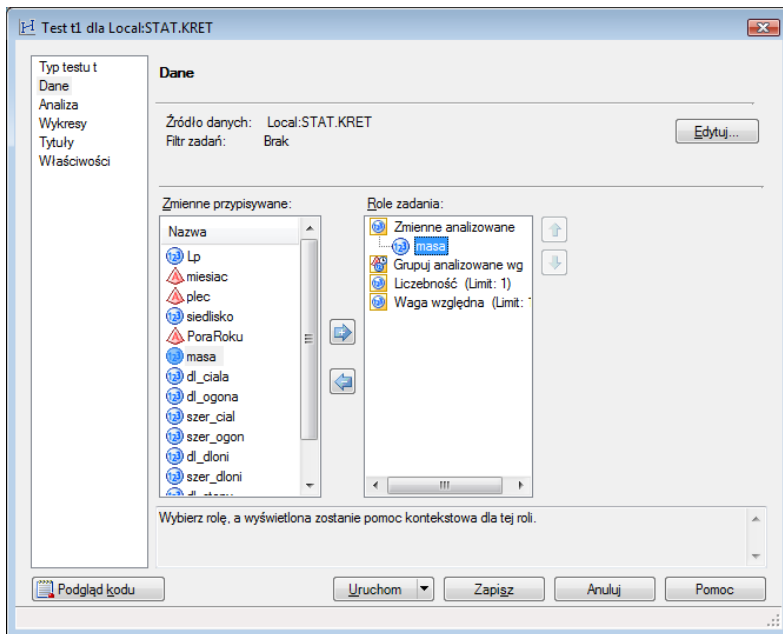
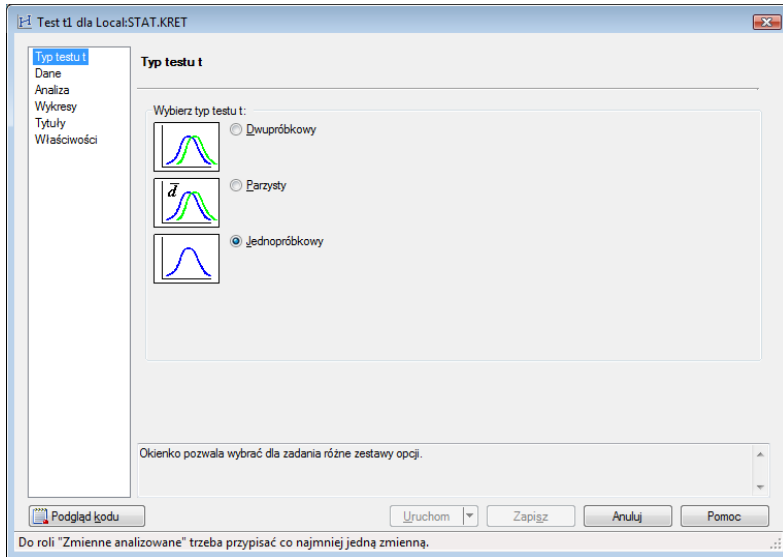


5. Klikamy na Uruchom i oglądamy WYNIKI.



Przykład 3. Średnia masa kretów w badanej populacji próbnej wyniosła 78,9 g. Zweryfikuj hipotezę zerową, zakładającą, iż średnia masa kretów w populacji generalnej wynosi 80 g.

$$H_0: \mu = 80 \text{ g}; H_1: \mu \neq 80 \text{ g}$$





WYNIKI

Test t

Procedura TTEST

Zmienna: masa

N	Średnia	Odch. std.	Błąd std.	Minimum	Maksimum
111	78.9024	14.4517	1.3717	58.0000	117.5

Średnia	95% CL Średnia	Odch. std.	95% CL Odch. std.
78.9024	76.1840 81.6207	14.4517	12.7683 16.6504

St. sw.	Wartość t	Pr > t
110	-0.80	0.4253

Prawdopodobieństw związane z testem T (0,425) dowodzi, iż nie mamy podstaw do odrzucenia hipotezy zerowej, zakładającej, iż średnia masa kretów wynosi 80 g. Nie oznacza to jednak, że omawiana masa właśnie tyle wynosi! Nie udało nam się wykazać, iż masa ciała jest różna aniżeli 80 kg.

Zadania

1. Zweryfikuj hipotezę zerową zakładającą, że średnia masa jaja stok wynosi 27 g (**jajaSROK.xls**).
2. Sprawdź czy istnieją różnice statystycznie istotne między płciami pod względem długości i szerokości pokryw. Analizy wykonaj dla każdego gatunku oddzielnie (**ChrzaszczMK.xls**).
3. Sprawdź czy Lipniki i Kamionki różnią się statystycznie pod względem emisji SO₂ i pyłu (**babulice100.xls**).
4. Zweryfikuj istotność różnic między grubością rogówki przed założeniem szkieł kontaktowych i w drugim tygodniu ich noszenia. Obliczenia wykonaj oddzielnie dla pomiarów wykonanych w centrum i na obwodzie rogówki (**oczy.xls**).
5. Porównaj maciorki i tryczki w zakresie wymiarów zoometrycznych (**wymiaryJag.xls**).