Wydział Informatyki Przedmiot:	Data: 22.04.2024
Grupa PS2	Prowadzący:
Grupa 12	
Imię i nazwisko:	Mgr. Joanna Gościk
Michał Sienicki	
Mateusz Milewski	
Damian Patalan	

# 1. Analiza pierwszego zbioru danych:

- **1.1 Zmienne zależne:** Przeanalizowano dwie zmienne zależne, w celu ustalenia różnicy między nimi.
- **1.2 Utworzenie zmiennej 'RÓŻNICA':** Obliczono różnicę między zmiennymi pierwszą i drugą, której wartości zostały zapisane jako nowa zmienna 'RÓŻNICA'.

### 1.3 Statystyki opisowe:

N ważnych, średnia, mediana, odchylenie standardowe, przedział ufności dla odchylenia standardowego (95%), błąd standardowy średniej, przedział ufności dla średniej (95%), oraz ekstrema (minimum i maksimum) zostały obliczone dla trzech zmiennych, w tym zmiennej 'RÓŻNICA'.

	Statystyki opis	sowe (Procer	t_ATP.sta w	Statystyka 3	3.stw)						
	Nważnych	Średnia	Ufność	Ufność	Mediana Minimum		Maksimum Odch.std		P. ufności odch. std.	P. ufności odch. std.	Standard.
Zmienna			-95,000%	95,000%					-95,000%	+95,000%	Błąd
SZWAJCARIA	38	59,23684	57,47682	60,99687	59,50000	50,00000	69,00000	5,354635	4,365440	6,927545	0,868637
POLSKA	38	35,00000	33,99777	36,00223	35,50000	30,00000	40,00000	3,049147	2,485859	3,944826	0,494637
RÓŻNICA	38	24,23684	22,09924	26,37444	24,50000	11,00000	35,00000	6,503364	5,301957	8,413709	1,054985

#### 1.4 Normalność rozkładu:

Przeprowadzono test normalności rozkładu Shapiro-Wilka dla zmiennej 'RÓŻNICA', a wyniki, w tym histogramy, zostały umieszczone w sprawozdaniu.

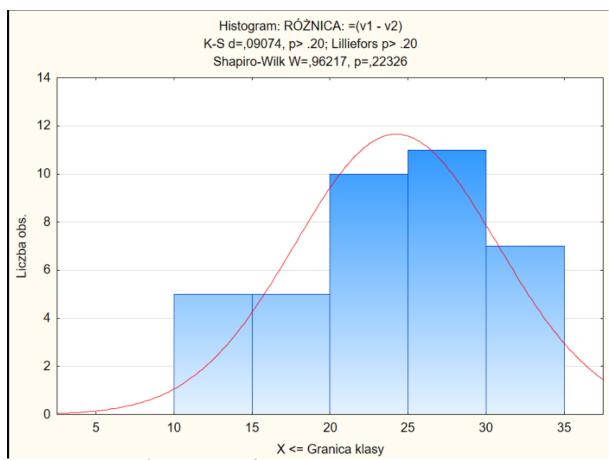


Tabela liczności: RÓŻNICA: =(v1 - v2) (Procent\_ATP.sta w Statystyka 3.stw) K-S d=.09074, p> .20; Lilliefors p> .20

1		,		.,			•
	Sha	piro-	-Wilk	W=	96217.	p=,22326	

	Liczba	Skumulow.	Procent	Skumul. %	% ogółu	Skumulow. %
Klasa		Liczba	Ważnych	Ważnych	Przypadki	Ogółu
5,000000 <x<=10,00000< td=""><td>0</td><td>0</td><td>0,00000</td><td>0,0000</td><td>0,00000</td><td>0,0000</td></x<=10,00000<>	0	0	0,00000	0,0000	0,00000	0,0000
10,00000 <x<=15,00000< td=""><td>5</td><td>5</td><td>13,15789</td><td>13,1579</td><td>13,15789</td><td>13,1579</td></x<=15,00000<>	5	5	13,15789	13,1579	13,15789	13,1579
15,00000 <x<=20,00000< td=""><td>5</td><td>10</td><td>13,15789</td><td>26,3158</td><td>13,15789</td><td>26,3158</td></x<=20,00000<>	5	10	13,15789	26,3158	13,15789	26,3158
20,00000 <x<=25,00000< td=""><td>10</td><td>20</td><td>26,31579</td><td>52,6316</td><td>26,31579</td><td>52,6316</td></x<=25,00000<>	10	20	26,31579	52,6316	26,31579	52,6316
25,00000 <x<=30,00000< td=""><td>11</td><td>31</td><td>28,94737</td><td>81,5789</td><td>28,94737</td><td>81,5789</td></x<=30,00000<>	11	31	28,94737	81,5789	28,94737	81,5789
30,00000 <x<=35,00000< td=""><td>7</td><td>38</td><td>18,42105</td><td>100,0000</td><td>18,42105</td><td>100,0000</td></x<=35,00000<>	7	38	18,42105	100,0000	18,42105	100,0000
Braki	0	38	0,00000		0,00000	100,0000

# 1.5 Testy statystyczne:

W zależności od wyniku testu normalności rozkładu, wybrano odpowiedni test statystyczny: Test t dla prób zależnych, jeśli dane miały rozkład normalny.

	Test średnic	h względem	stałej wartośc	ci odniesienia	a (Procent_A	TP.sta w Stat	ystyka 3.stw)			
	Średnia	Odch.st.	Ważnych	Bł. std.	Ufność	Ufność	Odniesienie	t	df	р
Zmienna			•		-95,000%	+95,000%	Stała			
RÓŻNICA	24,23684	6,503364	38	1,054985	22,09924	26,37444	0,00	22,97364	37	0,000000

Test kolejności par Wilcoxona, jeśli dane nie miały rozkładu normalnego.

	Test średnic	h względem	stałej wartośc	ci odniesienia	(Procent_A	TP.sta w Stat	ystyka 3.stw)								
	Średnia	irednia Odch.st. Ważnych Bł. std. Ufność Ufność Odniesienie t df p													
Zmienna					-95,000%	+95,000%	Stała								
RÓŻNICA	24,23684	6,503364	38	1,054985	22,09924	26,37444	0.00	22,97364	37	0.000000					

			Vilcoxona (Pr a istotne z p	ocent_ATP.s <,05000								
	N T Z p											
Para zmiennych	Ważny	ch										
SZWAJCARIA & RÓŻNIC	A	38 0,00	5,373093	0,000000								
	Test kolejnoś	ci par Wilco	xona (Procer	nt_ATP.sta w								
	Zaznaczone	wyniki są ist	otne z p <,05	5000								
N T Z p												
Para zmiennych	Ważnych											
POLSKA & RÓŻNICA	38	19 50000	5.090299	0.000000								

**1.6 Zapis skoroszytu:** Wszystkie wyniki i działania zostały zapisane.

### Statystyki opisowe:

Dla zmiennych SZWAJCARIA, POLSKA oraz RÓŻNICA obliczono średnią, medianę, odchylenie standardowe oraz przedział ufności. Wyniki wykazały, że średnie wartości dla SZWAJCARIA i POLSKA różnią się znacząco, co wskazuje na potencjalną różnicę w analizowanych danych.

# Normalność rozkładu zmiennej RÓŻNICA:

Test Shapiro-Wilka nie doprowadził do odrzucenia hipotezy o normalności rozkładu (p=0.22326), sugerując, że dane mogą pochodzić z populacji o rozkładzie normalnym.

# Test t dla zmiennej RÓŻNICA:

Przeprowadzony test t wykazał, że istnieje statystycznie istotna różnica względem wartości odniesienia (p < 0.00001). Wskazuje to na istotną różnicę między badanymi grupami.

# Test kolejności par Wilcoxona:

Ze względu na wynik testu Shapiro-Wilka dla zmiennej RÓŻNICA, przeprowadzono również test Wilcoxona, który potwierdził wcześniejsze wyniki testu t, wykazując istotność statystyczną różnicy (p < 0.00001).

### Wnioski dla pierwszego zbioru danych:

Średnia wartość zmiennej RÓŻNICA jest znacząco różna od zera, co sugeruje, że analizowane zmienne zależne różnią się między sobą w sposób statystycznie istotny.

Histogramy zmiennej RÓŻNICA wskazują na rozkład zbliżony do normalnego, co wspiera zastosowanie testu t.

# 2. Analiza drugiego zbioru danych:

**2.1 Zmienne niezależne:** Zanalizowano dwie zmienne niezależne.

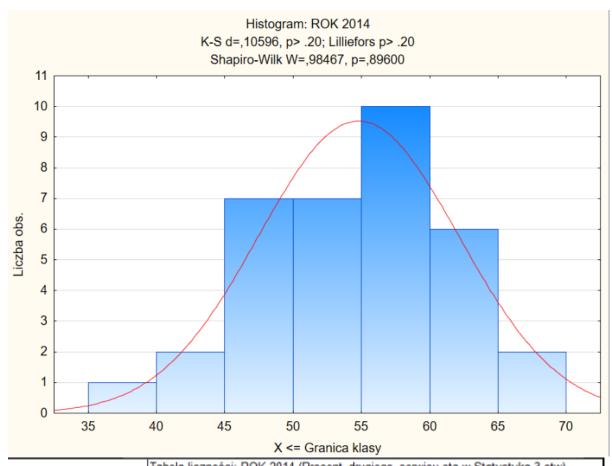
#### 2.2 Statystyki opisowe:

Podobnie jak w pierwszym zbiorze danych, obliczono statystyki opisowe dla obu zmiennych niezależnych.

	Statystyki opis	owe (Procer	nt_drugiego_s	serwisu.sta w	/ Statystyka :	3.stw)					
	Nważnych	Średnia	Ufność	Ufność	Mediana	Minimum	Maksimum	Odch.std	P. ufności odch. std.	P. ufności odch. std.	Standard.
Zmienna			-95,000%	95,000%					-95,000%	+95,000%	Błąd
ROK 2013	35	58,68571	56,20613	61,16530	58,00000	44,00000	74,00000	7,218324	5,838701	9,457463	1,220119
ROK 2014	35	54 80000	52 28287	57 31713	56 00000	40 00000	70.00000	7 327627	5 927113	9 600672	1 238595

# 2.3 Normalność rozkładów:

Dla każdej zmiennej niezależnej przeprowadzono test Shapiro-Wilka, a wyniki przedstawiono za pomocą histogramów oraz tabel.



	l abela licz	Tabela liczności: ROK 2014 (Procent_druglego_serwisu.sta w Statystyka 3.stw)													
	K-S d=,10	(-S d=,10596, p> .20; Lilliefors p> .20													
	Shapiro-W	hapiro-Wilk W=,98467, p=,89600													
	Liczba	Liczba Skumulow. Procent Skumul. % % ogółu Skumulow. %													
Klasa		Liczba Ważnych Ważnych Przypadki Ogółu													
35,00000 <x<=40,00000< td=""><td colspan="13">1 1 2,85714 2,8571 2,85714</td></x<=40,00000<>	1 1 2,85714 2,8571 2,85714														
40,00000 <x<=45,00000< td=""><td>2</td><td colspan="11">2 3 5,71429 8,5714 5,71429</td></x<=45,00000<>	2	2 3 5,71429 8,5714 5,71429													
45,00000 <x<=50,00000< td=""><td>7</td><td>10</td><td>20,00000</td><td>28,5714</td><td>20,00000</td><td>28,5714</td></x<=50,00000<>	7	10	20,00000	28,5714	20,00000	28,5714									
50,00000 <x<=55,00000< td=""><td>7</td><td>17</td><td>20,00000</td><td>48,5714</td><td>20,00000</td><td>48,5714</td></x<=55,00000<>	7	17	20,00000	48,5714	20,00000	48,5714									
55,00000 <x<=60,00000< td=""><td>10</td><td>27</td><td>28,57143</td><td>77,1429</td><td>28,57143</td><td>77,1429</td></x<=60,00000<>	10	27	28,57143	77,1429	28,57143	77,1429									
60,00000 <x<=65,00000< td=""><td>6</td><td>33</td><td>17,14286</td><td>94,2857</td><td>17,14286</td><td>94,2857</td></x<=65,00000<>	6	33	17,14286	94,2857	17,14286	94,2857									
65,00000 <x<=70,00000< td=""><td>2</td><td>35</td><td>5,71429</td><td>100,0000</td><td>5,71429</td><td>100,0000</td></x<=70,00000<>	2	35	5,71429	100,0000	5,71429	100,0000									
Braki	0	35	0,00000		0,00000	100,0000									

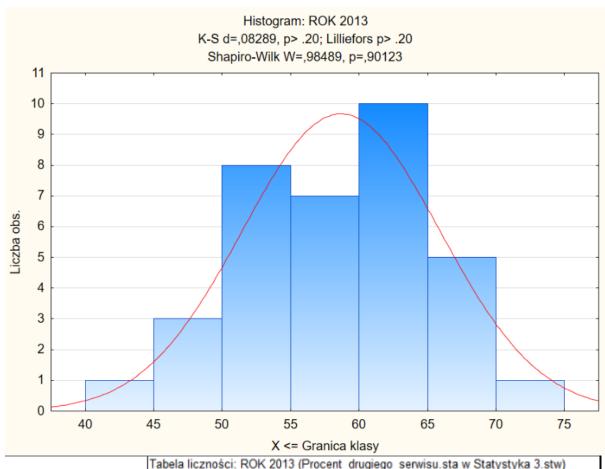


	Tabela liczności. Non 2013 (i locelit_druglego_serwisu.sta w otatystyka 5.5tw)														
	K-S d=,082	289, p> .20; Lilli	efors p> .20												
	Shapiro-W	ilk W=,98489, p	=,90123												
	Liczba														
Klasa		Liczba Ważnych Ważnych Przypadki Ogółu													
40,00000 <x<=45,00000< td=""><td>1</td><td>1</td><td>2,85714</td><td>2,8571</td><td>2,85714</td><td>2,8571</td></x<=45,00000<>	1	1	2,85714	2,8571	2,85714	2,8571									
45,00000 <x<=50,00000< td=""><td>3</td><td>4</td><td>8,57143</td><td>11,4286</td><td>8,57143</td><td>11,4286</td></x<=50,00000<>	3	4	8,57143	11,4286	8,57143	11,4286									
50,00000 <x<=55,00000< td=""><td>8</td><td>12</td><td>22,85714</td><td>34,2857</td><td>22,85714</td><td>34,2857</td></x<=55,00000<>	8	12	22,85714	34,2857	22,85714	34,2857									
55,00000 <x<=60,00000< td=""><td>7</td><td>19</td><td>20,00000</td><td>54,2857</td><td>20,00000</td><td>54,2857</td></x<=60,00000<>	7	19	20,00000	54,2857	20,00000	54,2857									
60,00000 <x<=65,00000< td=""><td>10</td><td>29</td><td>28,57143</td><td>82,8571</td><td>28,57143</td><td>82,8571</td></x<=65,00000<>	10	29	28,57143	82,8571	28,57143	82,8571									
65,00000 <x<=70,00000< td=""><td>5</td><td>34</td><td>14,28571</td><td>97,1429</td><td>14,28571</td><td>97,1429</td></x<=70,00000<>	5	34	14,28571	97,1429	14,28571	97,1429									
70,00000 <x<=75,00000< td=""><td>1</td><td>35</td><td>2,85714</td><td>100,0000</td><td>2,85714</td><td>100,0000</td></x<=75,00000<>	1	35	2,85714	100,0000	2,85714	100,0000									
Braki	0	35	0,00000		0,00000	100,0000									

# 2.4 Jednorodność wariancji:

Testowane za pomocą testu Levene'a w ramach przygotowania do testu t dla prób niezależnych.

# 2.5 Testy statystyczne:

W zależności od wyników testu Shapiro-Wilka i testu Levene'a, wybrano odpowiedni test statystyczny:

Test t dla prób niezależnych, jeśli dane spełniały założenia testu.

Test U Manna-Whitneya, jeśli dane nie spełniały założeń testu t.

	Testy dla prób niezależnych (Procent_drugiego_serwisu.sta w Statystyka 3.stw)																
	Uwaga: Zmienne traktowane są jako niezależne próby.																
	Średnia	Średnia	t	df	р	Nważnych	Nważnych	Odch.std	Odch.std	iloraz F	р	Levene'a	df	р	Średnia 1	Ufność	Ufność
Grupa 1 wz. Grupy 2	Grupa 1	Grupa 2				Grupa 1	Grupa 2	Grupa 1	Grupa 2	Wariancje	Wariancje	F(1,df)	Levene'a	Levene'a	- Średnia 2	-95,000%	+95,000%

**2.6 Zapis skoroszytu:** Wszystkie wyniki i działania zostały zapisane.

### Analiza drugiego zbioru danych:

#### Statystyki opisowe dla lat 2013 i 2014:

Obliczono średnie, mediany, odchylenia standardowe i przedziały ufności dla danych z obu lat. Wyniki wykazały niewielkie różnice między latami.

## Test t dla prób niezależnych (ROK 2013 vs. ROK 2014):

Przeprowadzono test t, który nie wykazał statystycznie istotnej różnicy w średnich między latami (p=0.930680).

### Test Levene'a na równość wariancji:

Test Levene'a wykazał, że wariancje między grupami są jednorodne (p=0.891406), co jest założeniem do przeprowadzenia testu t dla prób niezależnych.

### Wnioski dla drugiego zbioru danych:

Brak istotnych różnic w średnich między latami 2013 i 2014 może wskazywać na stabilność badanego zjawiska w czasie.

Jednorodność wariancji między grupami pozwala na stosowanie testów parametrycznych, takich jak test t.

#### **Podsumowanie:**

Analiza statystyczna dwóch zbiorów danych przeprowadzona za pomocą testu t, testu Shapiro-Wilka oraz testu Wilcoxona dostarczyła informacji o istotnych różnicach w obrębie pierwszego zbioru danych, co sugeruje, że analizowane zmienne zależne są od siebie statystycznie różne. W przypadku drugiego zbioru danych, brak wykazanych istotnych różnic w średnich między latami sugeruje stabilność analizowanego zjawiska w czasie.

#### **Dodatkowe uwagi:**

Dołączono histogramy i tabele liczności, które wizualizują rozkład danych i uzupełniają interpretację statystyczną.

Politechnika Białostocka Wydział Informatyki, 22.04.24r