

Analiza I

Potrzebujesz pomocy ? Wypełnij formularz

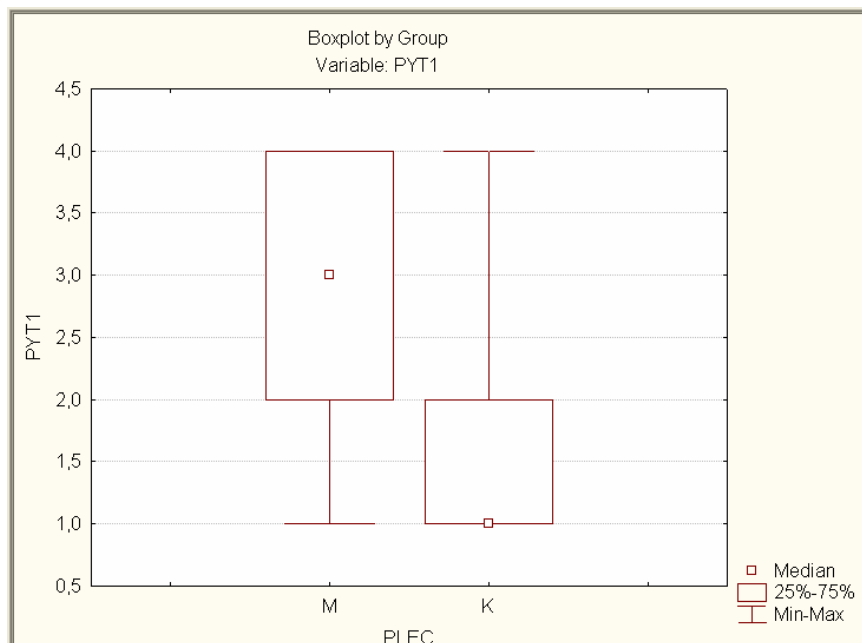
Dokonyjemy analizy mającej na celu pokazanie czy płeć jest istotnym czynnikiem różnicującym oglądalność w TV meczów piłkarskich. W tym celu zastosujemy test U, Manna – Whitneya do postawionego w ankiecie pytania1. Odpowiedzi wyrażone są w skali rangowanej, stąd też zastosowanie najmocniejszego testu nieparametrycznego dla dwóch populacji jest zasadne. W teście U, hipotezy H_0 i H_1 wyglądają następująco :

Test typu U Manna - Whitneya

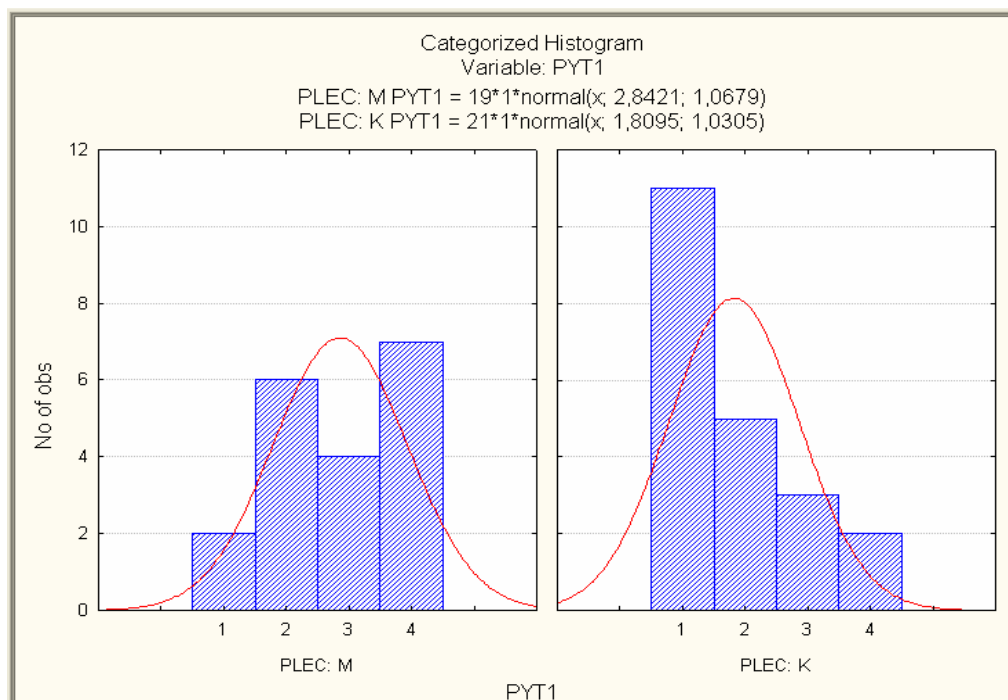
H_0 : Rozkłady populacji są takie same (płeć nie różnicuje)

H_1 : Rozkłady populacji nie są takie same (płeć różnicuje).

Zanim przeprowadzimy formalnie analizę testem U dokonamy wizualizacji danych stosując tzw. Wykres Pudełkowy (Box-Plot) oraz Histogram dla naszych danych empirycznych będących odpowiedziami na stawiane pytanie w przekroju płci .



Z powyższego widzimy dosyć istotne różnice i fakt że to Panowie znacznie częściej oglądają mecze piłkarskie niż Panie . Zachodzi więc przypuszczenie iż płeć będzie różnicować tę cechę . Poniżej zawarta jest analiza Histogramu .



Jak wyraźnie i istotnie płeć poróżniła nas w odpowiedzi na pytanie 1 , zauważamy po rozkładzie tych wyników przedstawionym za pomocą Histogramu . Zdecydowanie wśród Pań najwięcej jest jedynek , a najmniej czwórek , podczas gdy w przypadku Panów jest dokładnie na odwrót. Można więc śmiało powiedzieć że płeć różnicuje odpowiedzi na pytanie 1 , czyli powinniśmy hipotezę H_0 odrzucić . Poniżej znajduje się screen z Oprogramowania Statistica , będącym , analizą tego testu nieparametrycznego.

Mann-Whitney U Test (Slawek-ankieta)										
By variable PLEC										
Marked tests are significant at p < ,05000										
variable	Rank Sum M	Rank Sum K	U	Z	p-level	Z adjusted	p-level	Valid N M	Valid N K	2*sided exact p
PYT1	492,0000	328,0000	97,00000	2,776106	0,005502	2,880694	0,003968	19	21	0,004841

KOMENTARZ:

Z powyższego odczytujemy, poziom p-value = 0,005502 dla naszej statystyki testowej $U = 97$

Jak widać przy poziomie istotności $\alpha = 5\%$, mamy podstawy do odrzucenia hipotezy

zerowej. Jak widać z powyższego licznosc Panów i Pań wyniosła odpowiednio : 19 i 21 .

Obustronne p-value znacznie mniejsze od 5% jednoznacznie pozwala stwierdzić iż to Panie

zdecydowanie częściej oglądają mecze piłkarskie niż Panie , co zresztą już wcześniej

sugerowaliśmy .

Analiza II

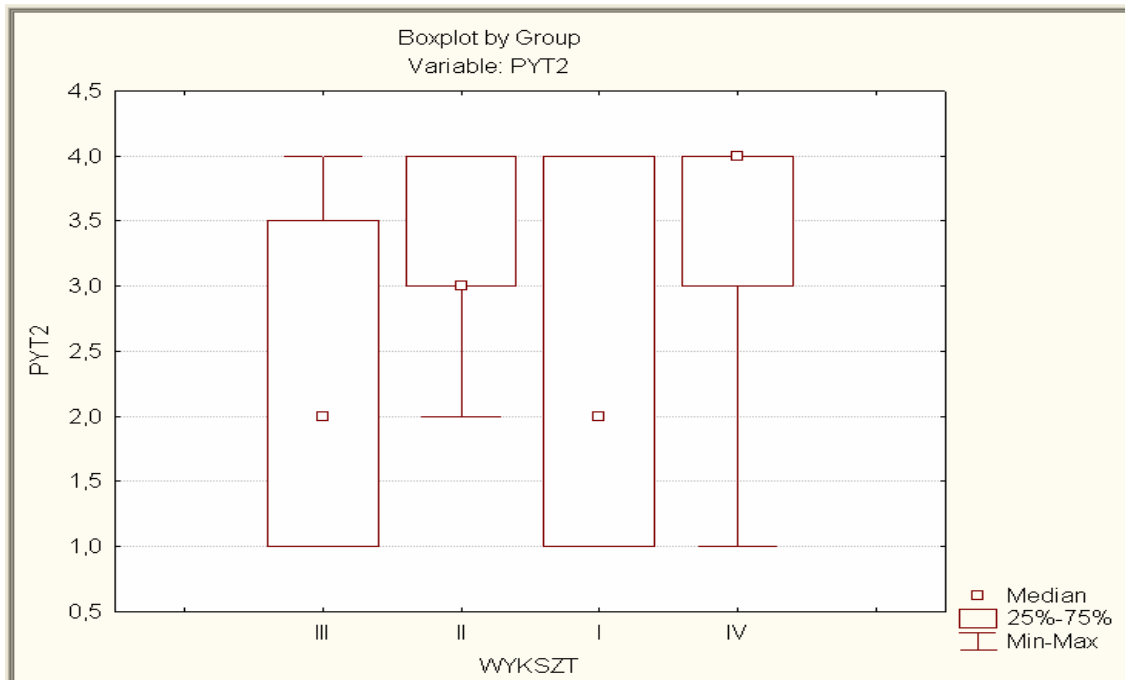
Analiza ta poświęcona będzie , tematowi ściśle związanemu z ostatnimi Mistrzostwami Świata 2006 . W pytaniu drugim pytaliśmy ile czasu Pan / Pani poświęcał (a) , na śledzenie bezpośrednich tematów poświęconych tematyce tego wydarzenia. Dokonałiśmy podziału na skale interwałową w skali dnia . Celem tej analizy będzie zweryfikowanie , czy ludzie o różnym wykształceniu poświęcali na śledzeni informacji z MŚ , mniej lub więcej czasu , czy też bez względu na stopień wykształcenia ten czas był mniej więcej taki sam . W tym celu zastosujemy mocny test nieparametryczny dla więcej niż dwóch prób tzw. Test H , gdyż dane tak wyrażone możemy traktować jako skalę rangowaną. W teście tym treści hipotez głównej i alternatywnej są jak poniżej :

Test typu H Kruskal – Wallis

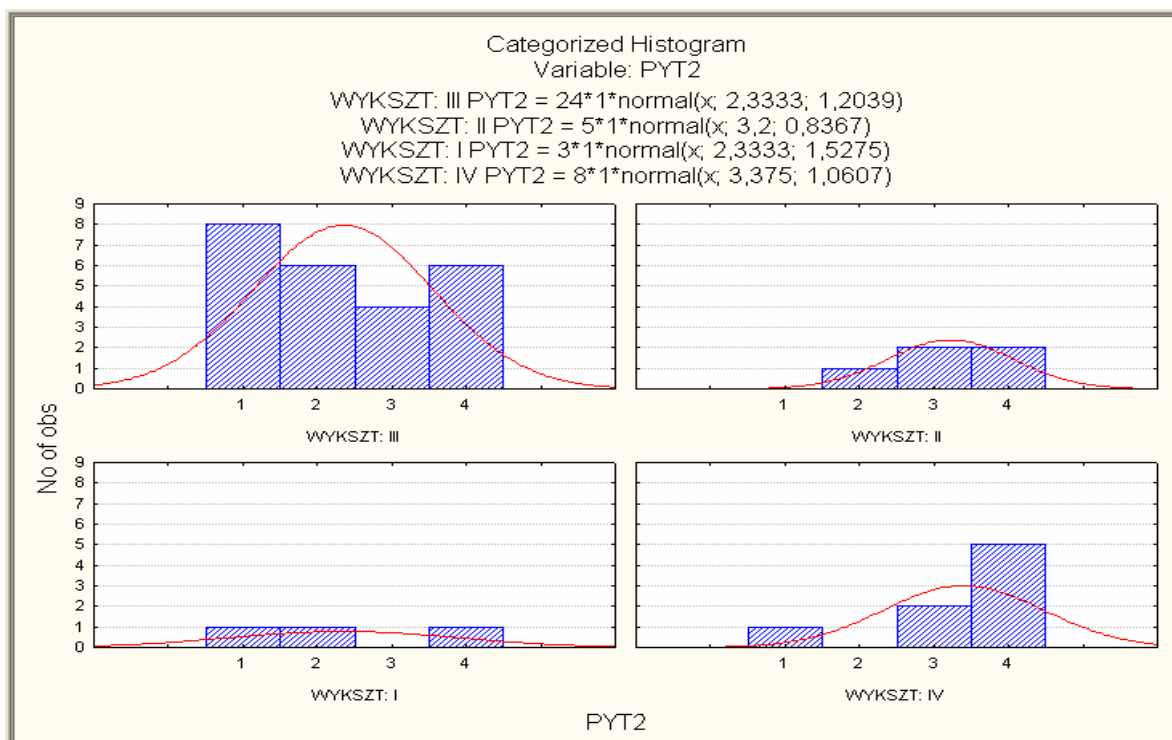
H0: Wszystkie cztery próby pochodzą z tej samej populacji (żadna się nie wyróżnia)

H1: Co najmniej jedna z czterech prób pochodzi z innej populacji niż pozostałe (co najmniej jedna różni się od innych).

Zanim przeprowadzimy formalnie analizę testem H dokonamy wizualizacji danych stosując Wykres Pudełkowy (Box-Plot) oraz Histogram dla naszych danych empirycznych będących odpowiedziami na stawiane pytanie 2 w przekroju wykształcenia .Z racji czterech grup I –IV , etykiet pokazujących różne wykształcenia będziemy analizować cztery Box-Ploty i cztery Histogramy jak poniżej . Należy zwrócić uwagę na to , że nie wszystkie grupy są tak samo liczne stąd też , znacznie lepiej , poprawnie było by zliczać w każdej z grup procent jaki stanowi poszczególna grupa ze względu na wykształcenie a nie samą licznosc !



Widzimy w oparciu o cztery pudełka z wąsami iż zjawisko jakim była oglądalność na co dzień programów poświęconych MŚ rozkłada się podobnie we wszystkich czterech grupach. Nie ma rażących różnic które by sugerowały iż hipotezę zerową należy odrzucić na korzyść alternatywnej H1. Poniżej podobna analiza za pomocą Histogramu .



Rozkłady odpowiedzi we wszystkich czterech grupach , nie są identyczne , ale też nie ma istotnych różnic , gdyż w każdej praktycznie grupie da się zauważyć odpowiedzi od wartości minimalnej do największej co zresztą było już wcześniej zauważone w Box-Plotach.

Poniżej znajduje się screen z Oprogramowania Statistica , będącym , analizą testu H .

Kruskal-Wallis ANOVA by Ranks; PYT2 (Sławek-ankieta)			
Independent (grouping) variable: WYKSZT			
Kruskal-Wallis test: H (3, N= 40) =5,571880 p =,1344			
Depend.: PYT2	Code	Valid N	Sum of Ranks
III	101	24	422,0000
II	102	5	126,5000
I	103	3	53,5000
IV	104	8	218,0000

KOMENTARZ:

Uzyskany poziom p-value = 0,1344 > 5% dla statystyki testowej $H = 5,571$, przy trzech stopniach swobody dla gęstości rozkładu chi-kwadrat , nie pozwala na odrzucenie hipotezy H_0 , stąd też ją przyjmujemy . Oznacza to , iż wykształcenie nie poróżniło nas , jeżeli chodzi o czas jaki poświęcaliśmy na obserwowanie zawodów sportowych . W zadaniu tym jak i w innych przyjęliśmy poziom istotności równy 5%. Reasumując każda z czterech grup „wykształceniowych “ poświęcała podobną ilość czasu na co dzień , oglądając sportowe wydarzenia.

Analiza I I I

Analiza zawarta w tym fragmencie dotyczy bardzo delikatnego aspektu kibicowania podczas oglądania meczów jakim jest zachowanie kibiców czy szeroko rozumiane bezpieczeństwo podczas trwania zawodów . Zapytaliśmy w pytaniu 4 , jak Pan / Pani ocenia poziom bezpieczeństwa w Europie podczas odbywanych zawodów piłkarskich . Będziemy chcieli zatem poznać opinie Pań i Panów odnośnie tego tematu . W tym celu zastosujemy słabszy test nieparametryczny od stosowanych wcześniej tzw. test serii Walda – Wolwofitza .

Test ten stosujemy do dwóch prób a formalnie treść hipotezy zerowej i alternatywnej jest jak poniżej :

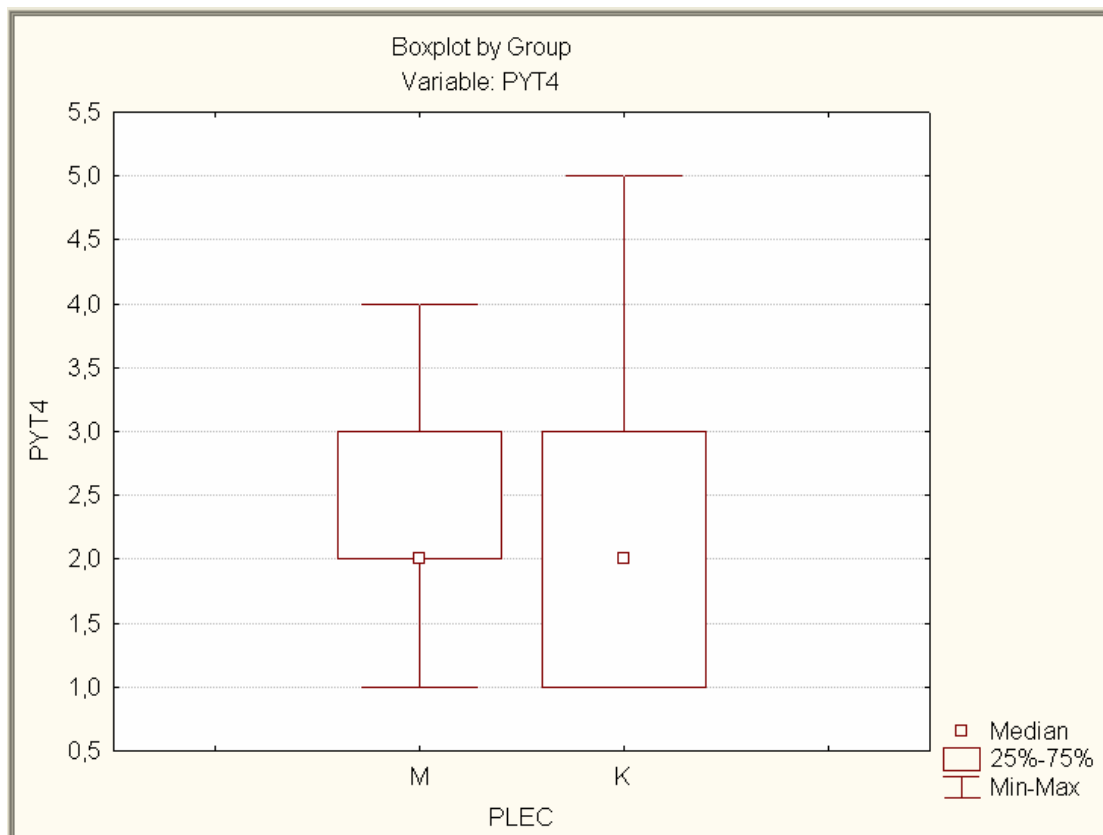
Test Walda - Wolfowitza (test dla dwóch prób)

Ho: Próby pochodzą z tej samej populacji (płęć nie różnicuje)

H1: Próby pochodzą z różnych populacji (płęć różnicuje).

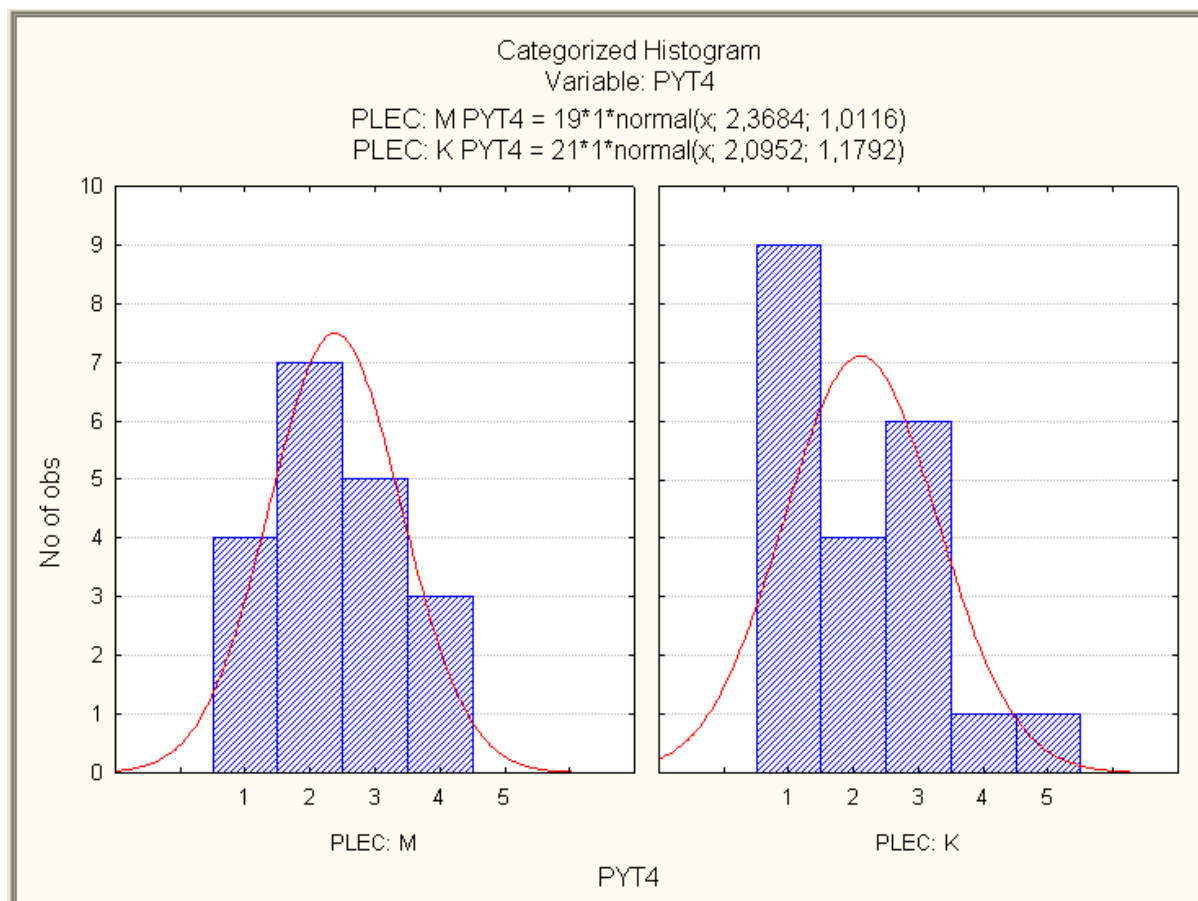
Analizę tego testu poprzedzimy graficzną prezentacją przy użyciu Box-Plotów i

Histogramów . Problem dotyczy więc tylko dwóch blisko-licznych prób .



Nie widać więc w oparciu o te dwa Box-Ploty , istotnego zróżnicowania , wydaje się iż problem wygląda tak samo w odbiorze obu płci . Można zaryzykować z małym błędem , że Panie i Panowie podobnie się wyrażają co do poziomu bezpieczeństwa , obawy może jedynie budzić ich niskie zaufanie co do tego faktu .

Poniżej dokonamy analizy Histogramu , a następnie testu W-W



W przekroju obu grup Panów i Pań, odpowiedzi stawiane na pytanie czwarte, przedstawiają się jak powyżej. Można śmiało powiedzieć, że nie występuje tutaj żadne silne zróżnicowanie tzn. płeć nie poróżniła nas w postrzeganiu bezpieczeństwa stadionowego w całej wspólnej Europie. Zarówno wśród Pań jak i Panów da się zauważyć odpowiedzi niskie jak i wysokie, a ich częstość występowania jest bardzo podobna, co skłania nas do przyjęcia hipotezy H0. Analiza p-value poniżej utwierdzi nas w tym, że tak przeprowadzone rozumowanie jest poprawne.

KOMENTARZ:

	Wald-Wolfowitz Runs Test (Slawek-ankieta)									
	By variable PLEC									
	Marked tests are significant at p < ,05000									
Variable	Valid N M	Valid N K	Mean M	Mean K	Z	p-level	Z adjstd	p-level	No. of Runs	No. of ties
PYT4	19	21	2,368421	2,095238	1,621988	0,104807	1,461395	0,143908	26	21

Widzimy z powyższego , że dla dwóch prób o licznosciach odpowiednio 19 – Panów i 21 Pań liczba serii jaką uzyskaliśmy wynosi 26 , jest duża , a odpowiadający tej wartości poziom $p\text{-value} = 0,1048$ jest większy niż teoretyczny równy 5% . Nie mamy więc podstaw do odrzucenia hipotezy H_0 , co już wcześniej zaznaczyliśmy w oparciu o analizę graficzną. W żaden sposób więc , nie można powiedzieć , że bezpieczeństwo na stadionach jest inaczej odbierane przez płeć – co swoją droga wydaje się być dosyć oczywiste . Poniżej już bez komentarza podajemy analizę tego problemu dokonaną przy użyciu najmocniejszego testu nieparametrycznego , testu U Manna – Whitneya.

Mann-Whitney U Test (Slawek-ankieta)										
By variable PLEC										
Marked tests are significant at $p < ,05000$										
variable	Rank Sum M	Rank Sum K	U	Z	p-level	Z adjusted	p-level	Valid N M	Valid N K	2*1sided exact p
PYT4	423,5000	396,5000	165,5000	0,920855	0,357127	0,958164	0,337981	19	21	0,361123

Można zauważyć , że decyzja odnośnie hipotezy H_0 , jest taka sama , a $p\text{-value} = 0,3571$.

Tym samym problem jest uwiarygodniony na dwa różne testy .

1. Jak często ogląda Pani/Pan mecze piłki nożnej ?

- rzadko
- czasami
- często
- bardzo często

2. Ile czasu Pani /Pan spędzaliście dziennie podczas oglądania Mundialu 2006 bądź programów ściśle związanych z tym wydarzeniem sportowym ?

- 0 - 1/2 godzin
- 1/2 - 1 godzin
- 1 - 2 godzin
- więcej niż 2 godziny

3. Czy opłaca Pan /Pani platforme cyfrową poświęconą w dużym stopniu wydarzeniom sportowym (Canal + , Polsat Sport i inne) :

- TAK NIE

Gdyby abonament miesięczny kilkakrotnie się zmniejszył , czy zdecydował (a) byś się na zakup takich stacji i kanałów sportowych ?

- TAK NIE

4. Jak Pani /Pan ocenia poziom bezpieczeństwa kibiców na stadionach europejskich



METRYCZKA

1.Płeć :

- kobieta mężczyzna

2.Wiek :

- do 20 lat 21 – 40 lat więcej niż 40 lat

3. Wykształcenie :

- podstawowe zawodowe średnie wyższe

4.Dochód miesięczny brutto:

- poniżej 1000 PLN 1000- 2000 PLN powyżej 2000 PLN

Dane empiryczne wykorzystane do analizy nieparametrycznej

PYT1	PYT2	PYT3-a	PYT3-b	PYT4	PLEC	WIEK	WYKSZ	DOCHOD
2	1	N	N	3	M	C	III	1
2	3	N	N	2	M	C	II	1
2	4	N	N	3	K	C	III	3
3	4	T	T	3	K	B	III	1
1	1	N	N	1	M	C	I	1
4	4	N	N	2	M	C	II	2
1	1	N	N	1	K	B	IV	1
4	4	N	T	2	M	B	IV	2
1	1	N	N	3	K	A	III	1
3	3	N	T	1	K	A	III	1
3	4	N	T	4	K	B	III	1
2	2	N	T	1	K	A	III	1
4	3	N	T	3	K	A	III	1
1	1	T	T	3	K	A	III	1
2	2	N	N	4	M	A	III	1
1	2	N	T	2	K	B	III	1
4	4	N	T	2	M	B	IV	2
1	2	N	N	2	K	A	III	3
2	2	N	N	3	M	A	II	1
3	4	N	T	4	M	B	III	2
2	3	N	T	1	K	B	III	1
1	2	N	T	1	K	A	III	2
4	4	N	T	2	K	C	I	1
4	4	N	T	3	M	C	IV	4
3	4	T	T	1	M	B	III	3
1	1	N	N	1	K	B	III	2
4	4	N	T	1	M	C	IV	2
1	1	N	T	1	K	C	III	2
4	3	N	T	1	M	A	IV	2
2	2	N	T	1	K	A	III	2
1	1	N	T	1	K	B	III	2
2	3	T	T	2	M	B	IV	3
3	4	N	T	3	M	B	III	2
2	3	N	T	5	K	C	III	1
3	3	N	T	2	M	A	II	1
2	4	N	T	4	M	A	II	1
1	2	N	N	2	M	A	I	1
1	1	N	N	3	K	C	III	1
1	1	N	N	2	K	C	III	2
4	4	T	T	3	M	A	IV	3

Potrzebujesz pomocy ?
Wypełnij formularz