

Edyta Karasek 157062

Gr. 9 AG

Kolokwium BAN gr. 9 zestaw B

1. Sprawdzić za pomocą analizy czynnikowej, czy pytania v45, v46, v48, v53 mogą stanowić jeden czynnik. Wkleić tabelkę z ładunkami czynnikowymi i uzasadnić odpowiedź.
Odp.

Zmienna	Ładunki czynnik.(Brak) (KolokwiumBAN) Wyodrębn. : Składowe główne (Oznaczone ładunki są >,500000)
	Czynnik 1
v45	-0,614258
v46	-0,642042
v48	-0,820902
v53	-0,723256
War.wyj.	1,986511
Udział	0,496628

Wszystkie zmienne mają ładunki o wartości większej niż 0,5 , z tego wynika ,że mogą stanowić jeden czynnik.

2. Sprawdzić za pomocą współczynnika alfa Cronbacha, czy powyższe pytania tworzą rzetelną skalę sumacyjną. Wkleić tabelkę i uzasadnić odpowiedź.

zmienna	Korelacje (KolokwiumBAN)			
	v45	v46	v48	v53
v45	1,000	0,323	0,308	0,201
v46	0,323	1,000	0,343	0,212
v48	0,308	0,343	1,000	0,553
v53	0,201	0,212	0,553	1,000

Korelacje w dwóch przypadkach nie przekraczają zwyczajowego progu na poziomie 0,3 ,reszta z nich charakteryzuje się dość niskimi współczynnikami korelacji.

zmie nna	Podsumow. skali:Średnia=8,63322 Odch.st=2,88432 N ważn.:578 (KolokwiumBAN) Alfa Cronbacha: ,645334 Alfa standaryzowana:,656617 Średnia kor. między poz.: ,329448					
	Śred.gdy usunięte	War. gdy usunięte	OdSt.gdy usunięte	Poz-Cał. Korel.	Alfa gdy usunięte	
	v45	6,922145	5,677330	2,382715	0,357241	0,621358
	v46	6,252595	5,431005	2,330452	0,376944	0,609945

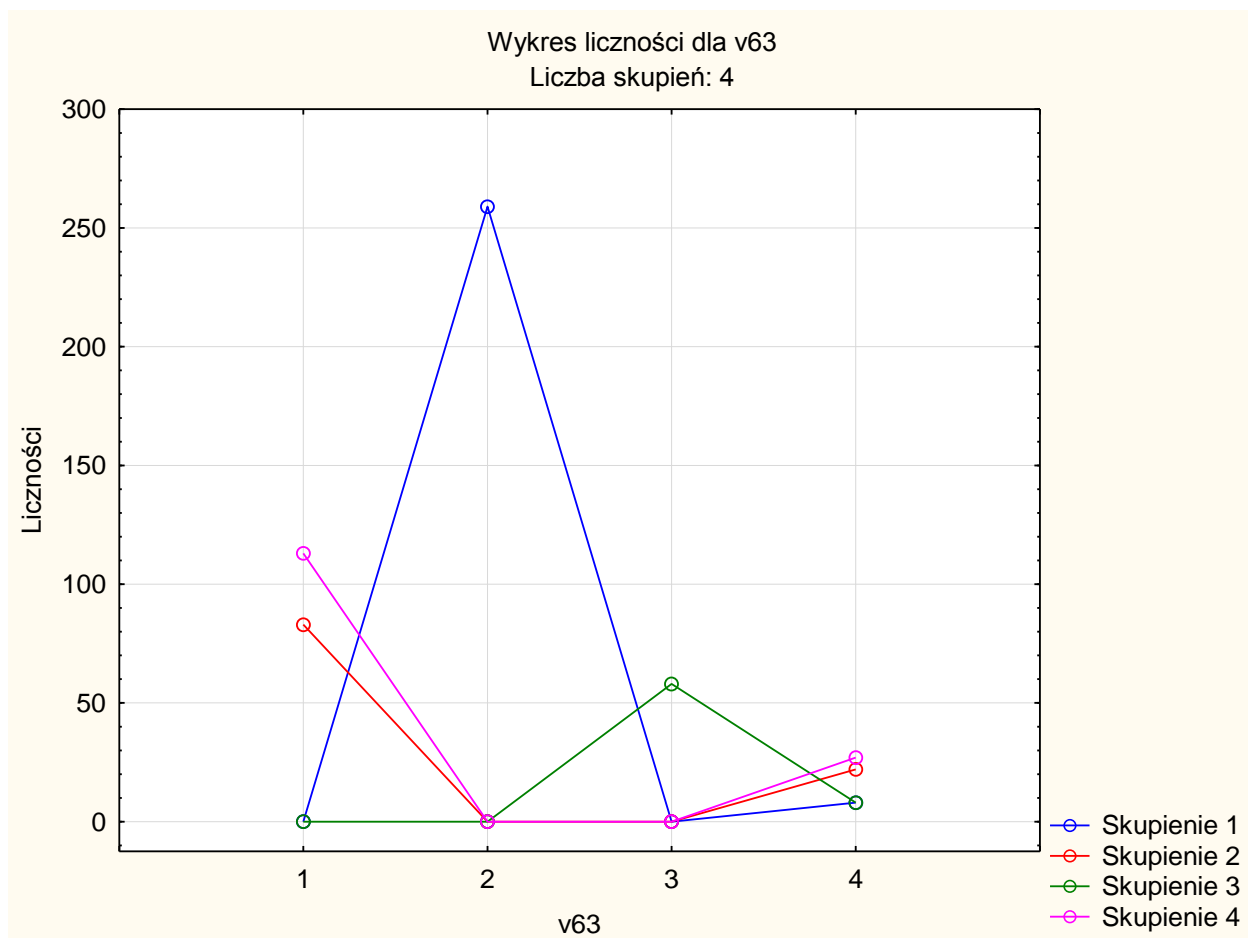
v48	6,519031	5,156213	2,270730	0,584506	0,483089
v53	6,205883	4,630622	2,151888	0,421400	0,589029

Alfa Cronbacha ma wartość mniejszą od przyjętego progu 0,7 . Z tego wynika, że pytania nie tworzą rzetelnej skali sumacyjnej.

W celu zwiększenia rzetelności należy usunąć te zmienne dla których wartość skorelowania jest mniejsza od 0,5. W tym przypadku jest to zmienna v45,v46 oraz v53.

3. Sprawdzić za pomocą klasycznej i nieparametrycznej analizy czynnikowej, czy średnia z wyżej wymienionych pytań istotnie różnicuje się względem wydziału (zmienna v63). Wkleić tabelki i udzielić odpowiedzi.

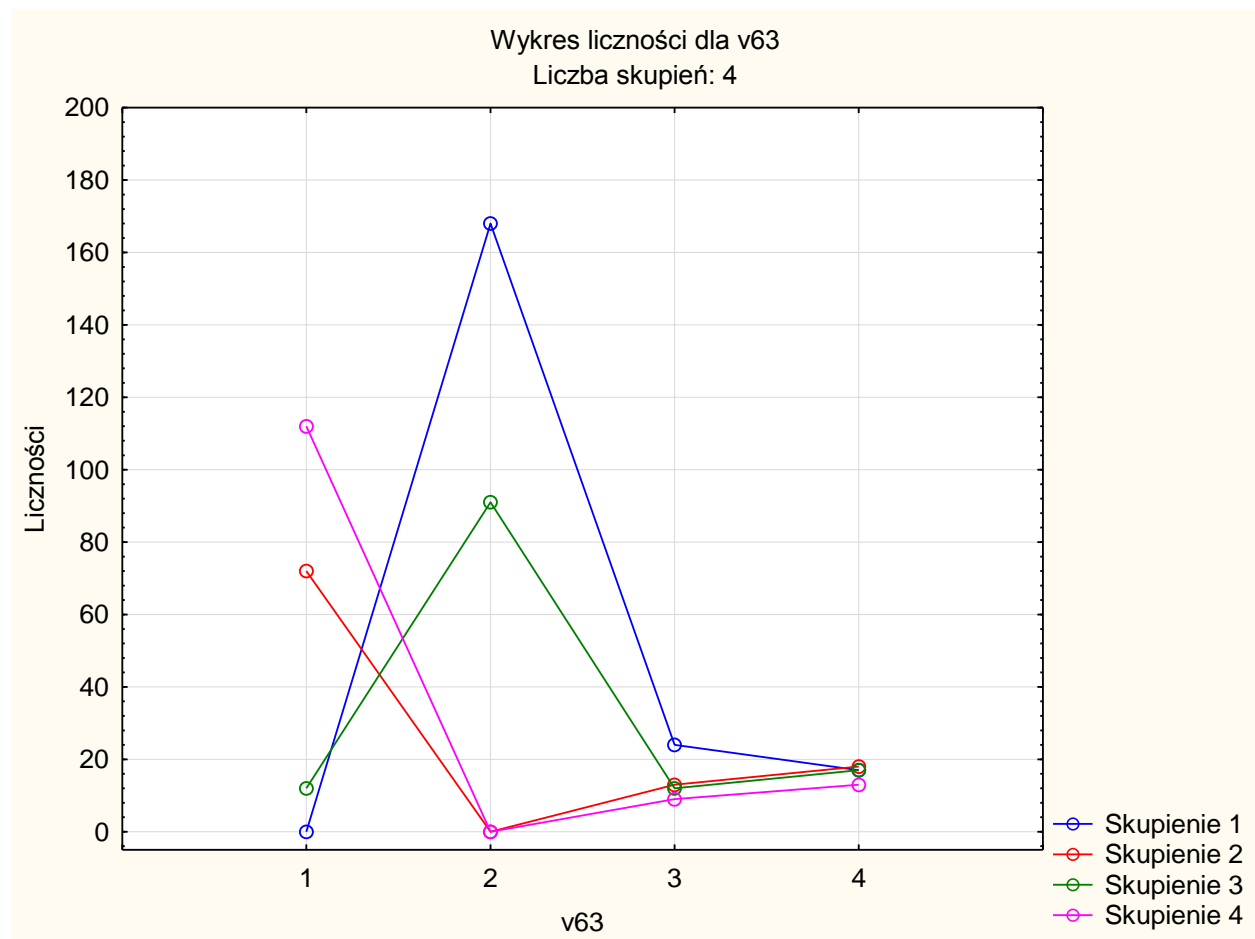
Skupienie	Średnie skupień (metoda k-średnich) (KolokwiumBAN) Liczba skupień: 4 Całkowita liczba przypadków uczących: 578			
	v63	Średnia	Liczba przypadków	Procent (%)
1	2	2,205056	267	46,19377
2	1	2,785714	105	18,16609
3	3	2,075758	66	11,41869
4	1	1,637500	140	24,22145



Analizując tabelę przedstawiającą średnią skupień można stwierdzić ,że 1 skupieniu sprzyjają respondenci z wydziału 2- ZIFu, 2 i 4 skupieniu studenci z NE , a trzeciemu uczniowie z IE. Żadne ze skupień nie wyróżniło wydziału GRiT jako zdominowanego.

Na podstawie wykresu można zauważyć ,że obliczona średnia nie do końca różnicuje wszystkie wydziały. Pierwsze skupienie charakteryzuje głównie studentów z wydziału 2 tj. ZIF-u , ale oprócz tej grupy dodatkowo można w nim wyróżnić osoby należące do GRiT. Drugie skupienie określa w zdecydowanej większości wydział 1 tj. NE oraz wydział 4 tj. GRiT. Trzecie skupienie składające się z najmniejszej ilości przypadków -66 określa głównie wydział 3 tj. IE oraz w nielicznych przypadkach GRiT. Skupienie 4 podobnie jak 2 zalicza do swojej grupy respondentów z wydziału NE oraz GRiT. Ich wykresy są bardzo podobne. Na podstawie tych obserwacji należy zastanowić się nad połączeniem skupienia 2 i 4 w jedno, ze względu na ich duże podobieństwo.

Nieparametryczna analiza skupień (brak standaryzacji)



Wykres licznosci jest zbliżony do otrzymanego w standardowej analizie skupien

ANOVA dla zmiennych ilościowych (KolokwiumBAN) Liczba skupień: 4 Całkowita liczba przypadków uczących: 578						
	Międzygru p. SS	df	Wewnątrzgr up. SS	df	F	Wartość p
Średnia	205,0663	3	94,94886	574	413,23 32	0,00

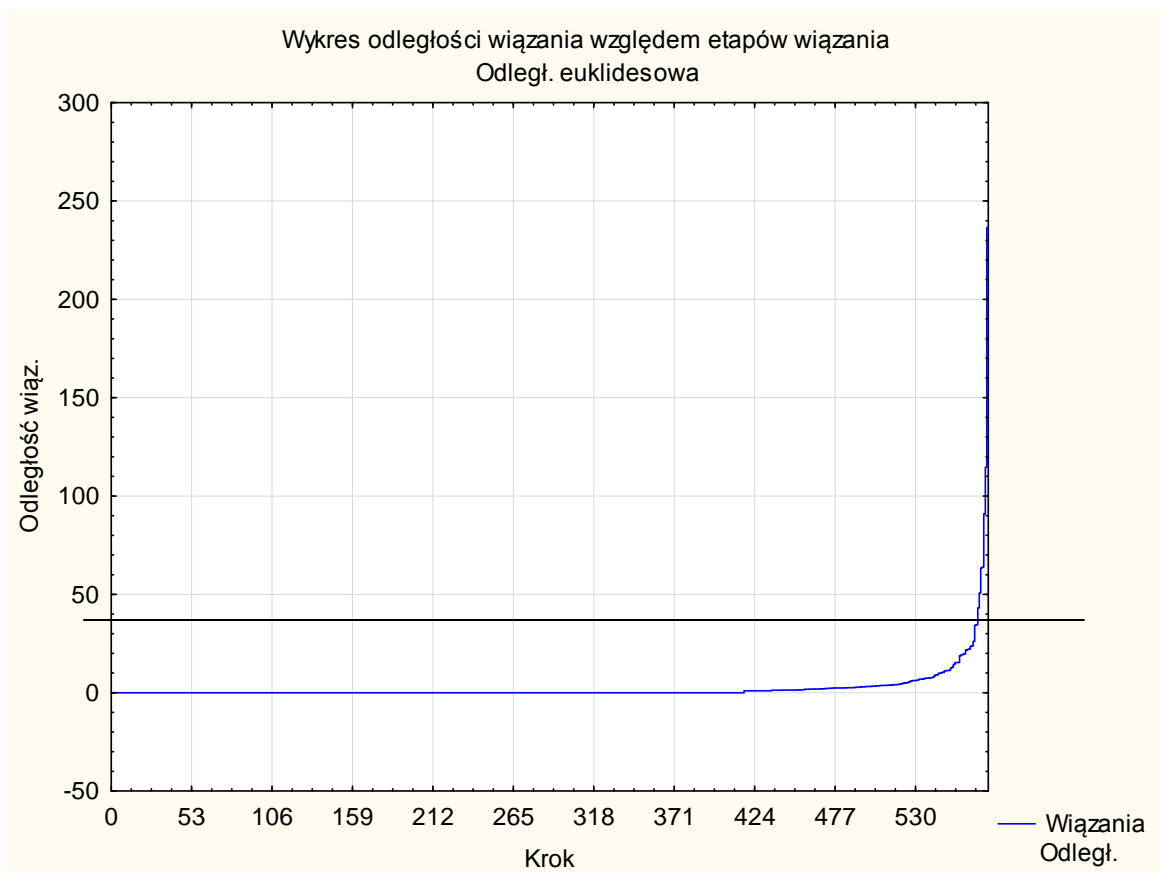
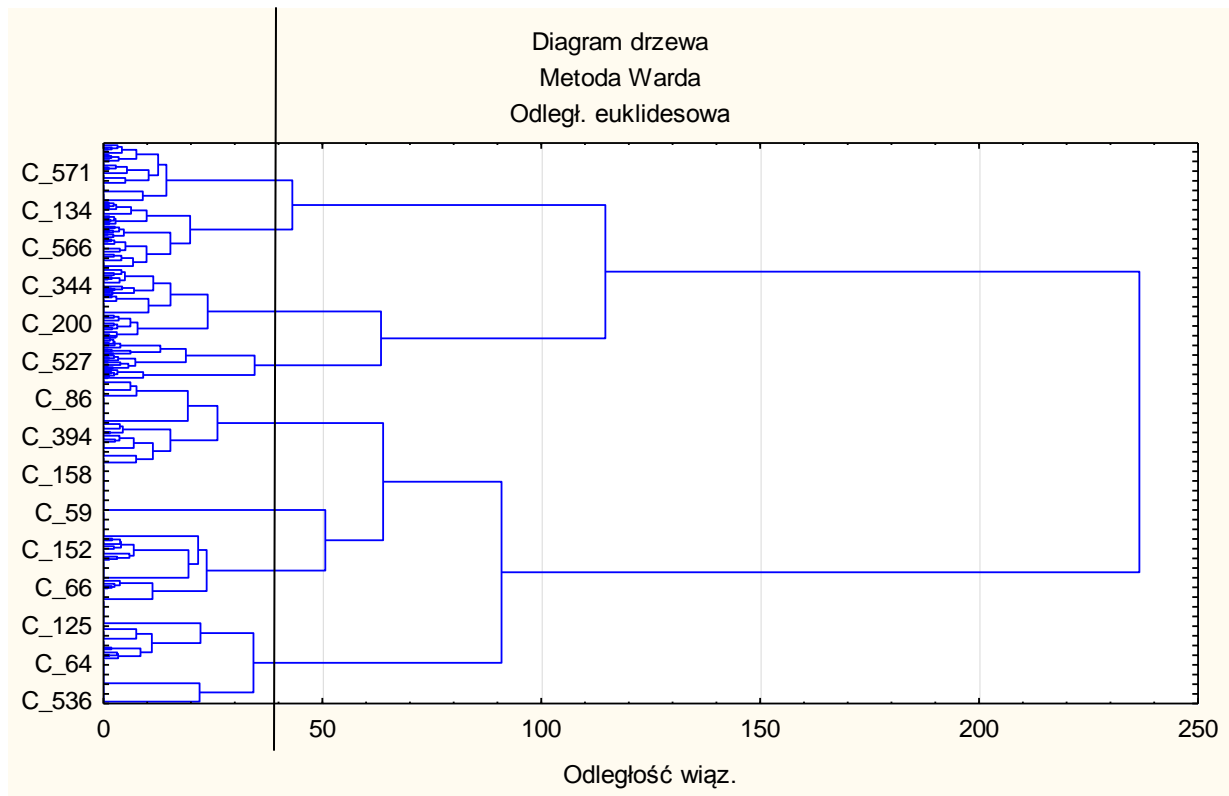
Średnie skupień (metoda k-średnich) (KolokwiumBAN) Liczba skupień: 4 Całkowita liczba przypadków uczących: 578				
Skupienie	v63	Średnia	Liczba przypadków	Procent (%)
1	2	1,777512	209	36,15917
2	1	1,453883	103	17,82007
3	2	3,104167	132	22,83737
4	1	2,361940	134	23,18339

Analizując powyższą tabelę można stwierdzić, że 1 oraz 3 skupieniu sprzyjają respondenci z wydziału 2-ZIFu, 2 i 4 skupieniu studenci z NE. Żadne ze skupień nie wyróżniło wydziałów IE oraz GRiT jako zdominowanego.

Testy niezależności dla zmiennych jakościowych (KolokwiumBAN) Liczba skupień: 4 Całkowita liczba przypadków uczących: 578					
	df	Chi-kwadrat	Wartość p	G-kwadrat	Wartość p
v6 3	9	422,7353	0,00	557,8705	0,00

Pomimo, że skupienia nie wyróżniły 3 i 4 wydziału to na podstawie test niezależności chi-kwadrat można twierdzić, że istnieje istotna zależność między wydziałem a skupieniem (p-value < 0,05).

4. Sprawdź za pomocą hierarchicznej analizy skupień (metoda Warda, odległość euklidesowa bez standaryzacji zmiennych), w ile skupień łączą się respondenci. Wklej wykres(y) i uzasadnij odpowiedź.



Na diagramie drzewa oraz na wykresie przebiegu aglomeracji zostały naniesione linie, które wskazują na miejsca przecięcia wiązań. W obydwu przypadkach przecinają one około 30 wiązanie. Na drugim wykresie, można też zauważyć, że dokładnie w tym miejscu następuje wyraźny uskok. W wyniku tych operacji powstanie 10 grup obiektów.

5.Sprawdzić w metodzie k-średnich, czy dla liczby skupień wyznaczonej w pytaniu 4 wszystkie zmienne istotnie się różnicują (domyślna metoda wstępnych skupień, bez standaryzacji zmiennych). Wkleić tabelkę analizy wariancji i uzasadnić odpowiedź.

Zmienna	Analiza wariancji (KolokwiumBAN)					
	Między SS	df	Wewn. SS	df	F	Istotn. p
v45	334,4828	9	220,2663	568	95,8366	0,00
v46	387,3509	9	226,9121	568	107,7340	0,00
v48	292,6027	9	163,8609	568	112,6961	0,00
v53	732,3942	9	119,0539	568	388,2462	0,00

Na podstawie wartości statystyki testowej F można twierdzić, że wszystkie zmienne charakteryzują się względnie dużymi wartościami, z tego wynika, że uwzględniony podział jest odpowiedni. Biorąc pod uwagę kolumnę z wartościami p-value, każda zmienna ma istotny wpływ na podział studentów na 10 skupień.

