

Statistika 2 z računalniško analizo podatkov

Univariatne analize in
oblikovanje izpisov



V Statistične analize v SPSS-ju

V.1 Splošno

V.2 Univariatne analize

Analyze – Descriptive statistics - Frequences

Analyze – Descriptive statistics - Descriptives

Analyze – Descriptive statistics - Explore

Graphs – Chart Builder

V.3 Oblikovanje izpisa



V.1 Splošno

Statistične analize (ali druge zahteve) lahko izvajamo na 2 načina:

- s pomočjo menijev (*Analyze in Graphs*), ki so dostopni v okviru SPSS-ja v vseh oknih;
- s pomočjo programskih stavkov, ki jih zapisujemo v okno Syntax Editor.



V.1 Splošno

Delo z meniji:

- **Glavna pogovorna okna** za posamezne analize (npr. *Analyze – Descriptive Statistics – Frequencies*) odpiramo z izbiro možnosti znotraj menija.
- V glavnem pogovornem oknu navadno določimo, katero(e) spremenljivko(e) obravnavamo (prenesemo jih iz seznama na levi strani v seznam na desni strani).
- **Dodatna pogovorna okna** za podrobnejše določanje zahtev (npr. *Statistics* ali *Charts*) odpiramo z gumbi v glavnih pogovornih oknih.
- V glavnih in dodatnih pogovornih oknih s klikanjem ali vpisovanjem določimo vse podrobnosti izvajanja analiz.



V.1 Splošno

Pomoč pri delu z menuji:

- Klik na gumb *Help*.

Izvedba ukaza:

- Z gumbom OK v glavnem pogovornem oknu sprožimo izvajanje analiz, ki smo jih zahtevali v tem in njemu dodatnih oknih.

V.2 Univariatne analize = analize ene spremenljivke

Primeri univariatne statistične analize	Procedura v SPSS-u
Frekvenčna porazdelitev za pregled vrednosti spremenljivke in njihovih frekvenc	<ul style="list-style-type: none">• Analyze – Descriptive statistics - Frequencies
Grafična predstavitev, npr. histogram, poligon, ogiva, strukturni stolpci, krogi – odvisno od merske lestvice	<ul style="list-style-type: none">• Analyze – Descriptive statistics – Frequencies – Charts (histogram, strukt. stolpci, krogi)• Analyze – Descriptive statistics – Explore (histogram)• Graphs – Chart builder
Opisne statistike: Kvantili, srednje vrednosti, mere varibilnosti (razpršenosti), koeficienti asimetrije, sploščenosti	<ul style="list-style-type: none">• Analyze – Descriptive statistics - Frequencies• Analyze – Descriptive statistics – Descriptives• Analyze – Descriptive statistics – Explore
Inferenčna statistika: interval zaupanja za aritmetično sredino, za delež	<ul style="list-style-type: none">• Analyze – Descriptive statistics – Explore



Procedura FREQUENCIES

Kako?

- V meniju izberemo: *Analyze – Descriptive statistics – Frequencies*.

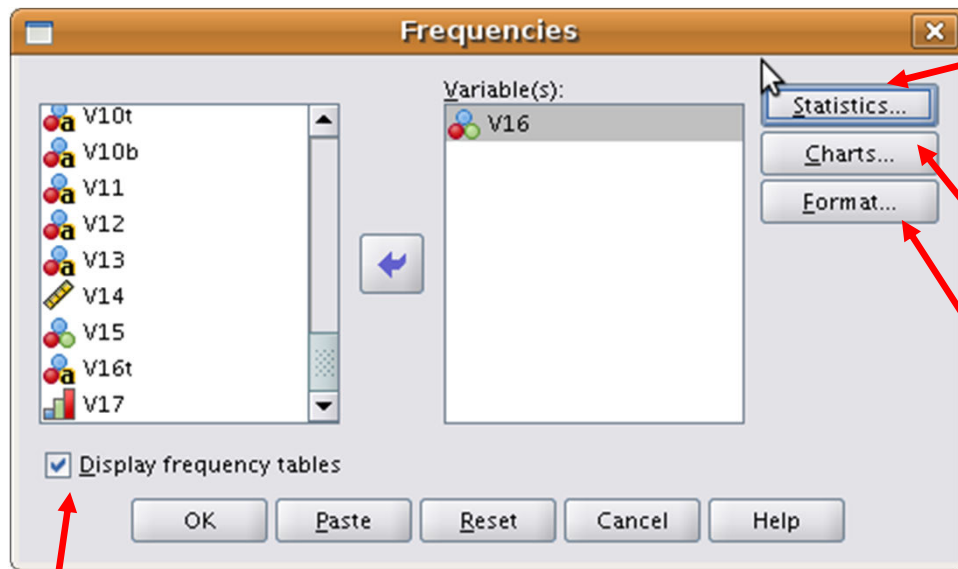
Kaj omogoča?

- Izpiše frekvenčno(e) tabelo(e), absolutne in relativne frekvence (strukturne odstotke), absolutne in relativne kumulative (izbrana opcija *Display frequency tables*).
- Izračuna kvantile, srednje vrednosti, mere variabilnosti, koeficiente asimetrije, sploščenosti (gumb *Statistics*).
- Nariše strukturne stolpce, kroge, histogram (gumb *Charts*).
- Določamo lahko vrstni red izpisa vrednosti, obliko tabele pri istočasnem ukazu za več spremenljivk, neizpis prevelikih tabel (gumb *Format*).

Kdaj je to uporabno?

- To naj bo vedno prvi korak v vsaki analizi, ker na ta način pregledamo vrednosti in ugotovimo morebitne neveljavne vrednosti, ki jih je potrebno določiti kot uporabniške manjkajoče vrednosti.
- Izpis frekvenčne tabele je manj smiseln, če gre za spremenljivko z velikim številom različnih vrednosti (npr. dohodek).
- Kadar želimo izračunati opisne statistike za neintervalne spremenljivke (mediana, modus, kvantili, ...).

Procedura FREQUENCIES



Izračuna kvantile, srednje vrednosti, mere variabilnosti, koeficiente asimetrije (*Skewness*), sploščenosti (*Kurtosis*)

Nariše strukturne stolpce, kroge, histogram

Določamo lahko vrstni red izpisa vrednosti, obliko tabele

Izpiše frekvenčno(e) tabelo(e),

Procedura FREQUENCIES

The image shows the 'Frequencies: Statistics' dialog box in SPSS. It is annotated with Slovenian text and red arrows pointing to various options:

- Percentile Values:**
 - Quartiles**: Annotated as 'Kvartili' (Quartiles).
 - Cut points for:** 10 equal groups: Annotated as 'Kvantili, ki razdeijo populacijo na x delov' (Quantiles that divide the population into x parts).
 - Percentile(s):**: Annotated as 'Poljubni centili – z njimi lahko dobimo poljuben kvantil' (Arbitrary percentiles – with them we can get an arbitrary quantile).
- Central Tendency:**
 - Mean**: Annotated as 'Srednje vrednosti (vsota (Sum) seveda ni srednja vrednost, čeprav je v tem okvirčku)' (Average values (the sum (Sum) is of course not the average value, even though it is in this box)).
 - Median**
 - Mode**
 - Sum**
- Dispersion:**
 - Std. deviation**: Annotated as 'Standardni odklon, varianca, ranžirni razmik' (Standard deviation, variance, range).
 - Variance**
 - Range**
 - Minimum**
 - Maximum**
 - S.E. mean**: Annotated as 'Standardna napaka aritmetične sredine' (Standard error of the mean).
- Distribution:**
 - Skewness**: Annotated as 'Mera asimetrije' (Measure of asymmetry).
 - Kurtosis**: Annotated as 'Mera sploščenosti oz. koničavosti' (Measure of flatness or peakedness).

Buttons: Continue, Cancel, Help.



Procedura DESCRIPTIVES

Kako?

- V meniju izberemo: *Analyze – Descriptive statistics – Descriptives*.

Kaj omogoča?

- Izračuna nekatere opisne statistike: aritmetično sredino, vsoto vrednosti, mere variabilnosti, koeficiente asimetrije, sploščenosti (gumb *Options*).
- Če opravljamo analize za več spremenljivk hkrati, lahko določamo vrstni red izpisa izračunov za več spremenljivk (gumb *Options*).
- Izračuna standardizirane vrednosti (opcija *Save standardized values as variables*).

Kdaj je to uporabno?

- Če želimo izračunati opisne statistike samo za vsaj intervalne spremenljivke, pa pri tem ne potrebujemo frekvenčne tabele.
- Če želimo v izpisu urediti spremenljivke po abecedi ali po povprečjih
- Če potrebujemo standardizirane spremenljivke (npr. za nadaljnje analize).



Procedura CASE SUMMARIES

Kdaj je to uporabno?

- Kadar želimo izpisati vrednosti izbranih spremenljivk za izbrane enote.
- Kadar želimo izračunati neke statistične karakteristike po skupinah enot.



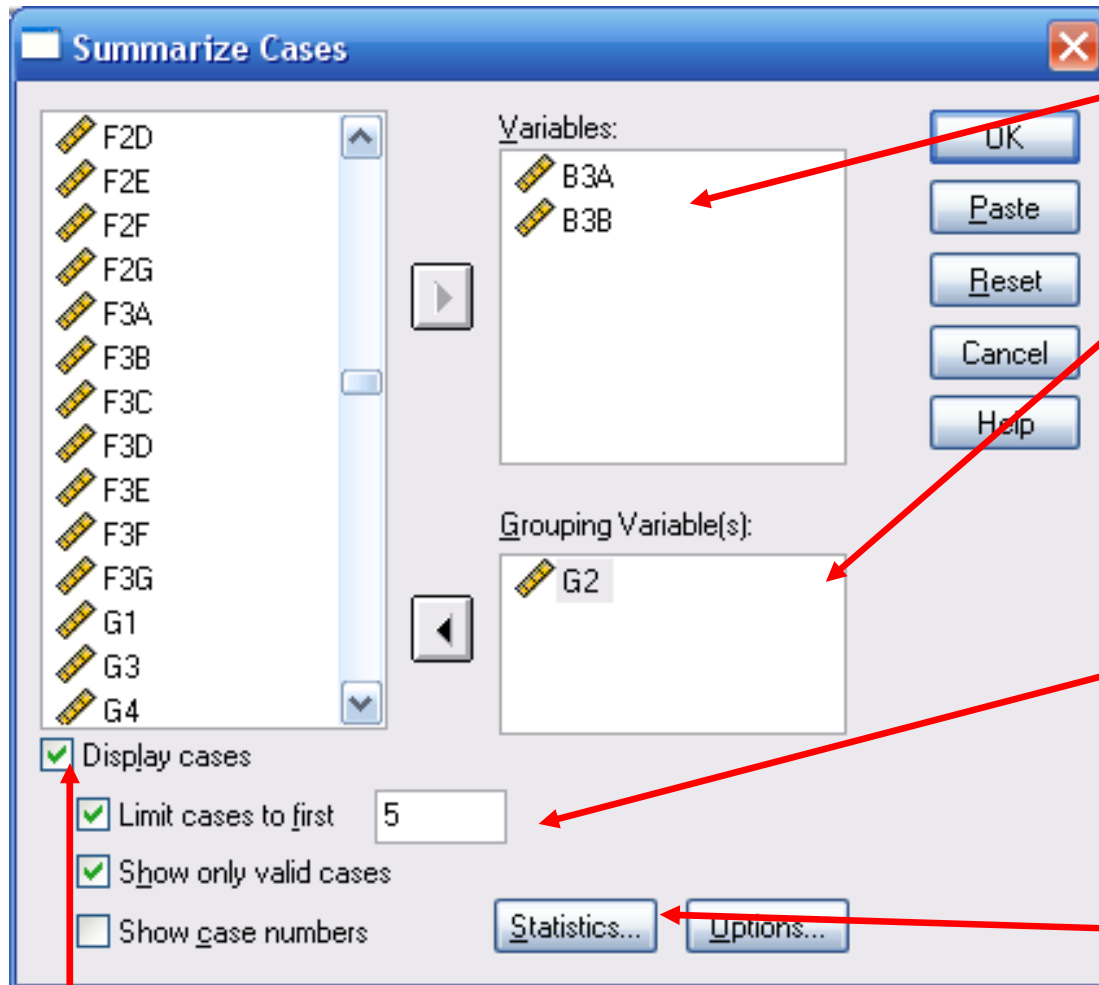
Procedura CASE SUMMARIES: Izpis vrednosti izbranih spremenljivk za izbrane enote

Analyze – Reports – Case Summaries

- Spremenljivko oz. spremenljivke, za katere zahtevamo izpis, prenesemo v polje *Variables*. Pri tem določimo vrstni red spremenljivk pri izpisu.
- Če želimo izpis po nekih podskupinah, v polje *Grouping Variable(s)* prenesemo spremenljivko oz. spremenljivke, katerih vrednosti določajo podskupine.
- Določimo lahko izpis za vse enote v datoteki ali samo za izbrano število enot iz začetka datoteke (opcija *Limit cases to first ...*).
- Določimo izpis samo za enote z veljavnimi vrednostmi ali za vse enote (opcija *Show only valid cases*).
- Določimo zapis zaporedne številke enote (*Show case numbers*).
- S specifikacijami v podoknu, ki ga odpremo z gumbom *Statistics*, lahko določamo, katere statistike naj se izračunajo za izpisane enote. Običajno izpišemo samo število enot (*Number of cases*).
- S specifikacijami v podoknu, ki ga odpremo z gumbom *Options*, določamo obliko izpisa (naslove, način izpisa manjkajočih vrednosti itd.).
- Pozor: V tabelo se izpisujejo lahko vrednosti spremenljivk, vrednosti in labele teh vrednosti ali samo labele. To je odvisno od nastavitve z ukazom *Edit – Options*.

Izpis vrednosti izbranih spremenljivk za izbrane enote

Ukaz: *Analyze – Reports – Case Summaries*



Spremenljivke, katerih vrednosti želimo izpisati

Spremenljivke, ki naj definirajo skupine, po katerih naj bo izpis narejen (tu spol, posebej za moške in posebej za ženske)

Ali naj izpišemo podatke samo za prvih nekaj enot (in koliko)

Katere statistike se naj izpišejo za vse izpisane enote oz. tudi po skupinah, če so določene

Ali naj izpišemo podatke za posamezne enote (ali samo statistike za celoto in po skupinah) – izberemo za izpis enot.

Izpis vrednosti izbranih spremenljivk za izbrane enote

Primer izpisa za 2 spremenljivki za prvih 5 enot v datoteki; izpisane samo enote z veljavnimi vrednostmi:

Izpis^a

	Zaporedna številka enote	B3A osebnega računalnika?	B3B prenosnega računalnika?
1	3	1 da, tudi jaz	3 ne, ne dostopamo
2	4	1 da, tudi jaz	3 ne, ne dostopamo
Skupaj	N	2	2

a. Omejeno na prvih 5 enot.

Primer izpisa za 2 spremenljivki za prvih 5 enot v datoteki po spolu; izpisane vse enote:

Izpis^a

				Zaporedna številka enote	B3A osebnega računalnika?	B3B prenosnega računalnika?
G2 Spol	1 moški	(manjka) 1		1	-2	-2
		1		4	1 da, tudi jaz	3 ne, ne dostopamo
		Skupaj	N		1	1
	2 ženski	(manjka) 2		2	-2	-2
		1		3	1 da, tudi jaz	3 ne, ne dostopamo
		(manjka) 5		5	-2	-2
		Skupaj	N		1	1
		Skupaj	N		2	2

a. Omejeno na prvih 5 enot.



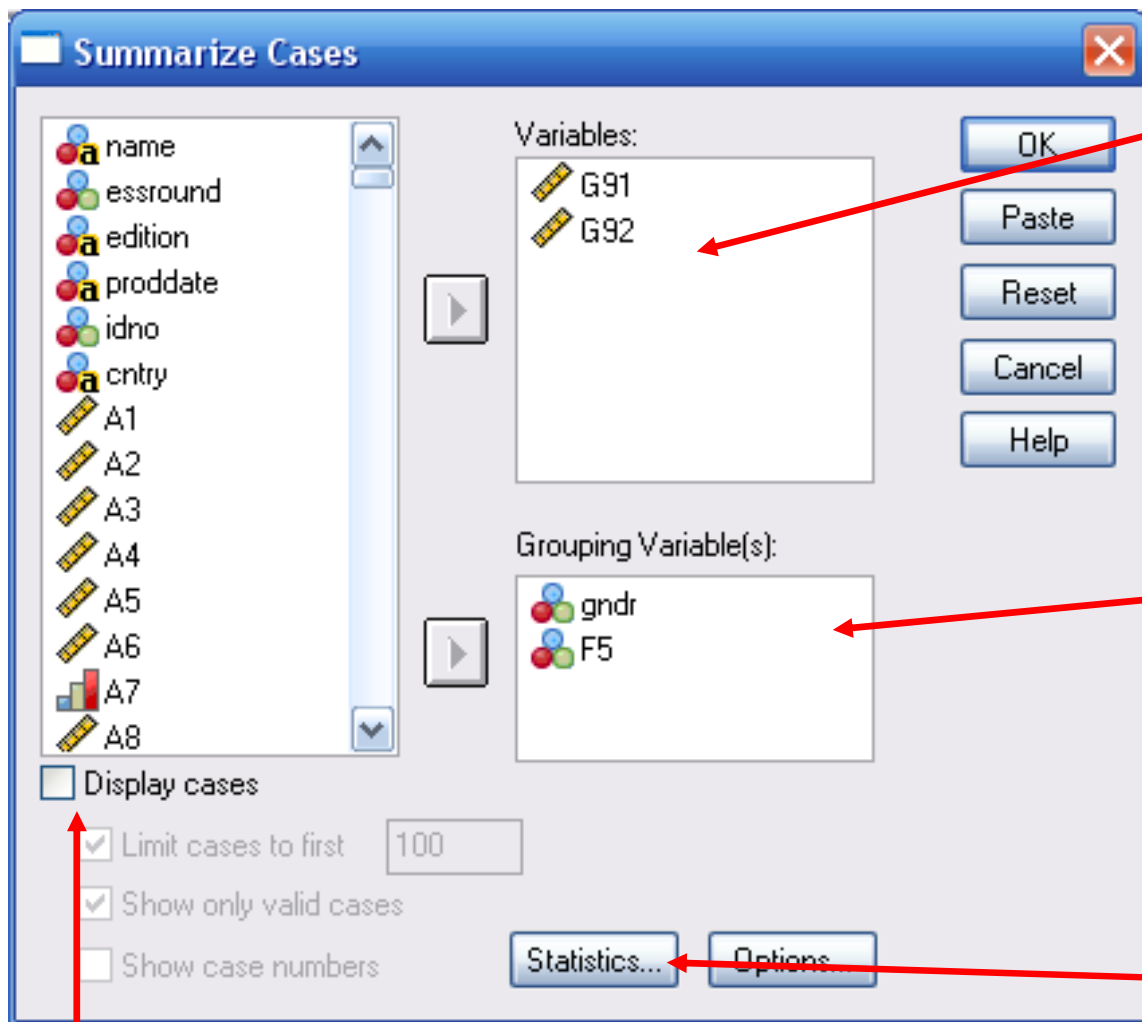
Procedura CASE SUMMARIES: Izračun statistik po skupinah

Analyze – Reports – Case Summaries

- Spremenljivko oz. spremenljivke, za katere zahtevamo izpis, prenesemo v polje *Variables*. Pri tem določimo vrstni red spremenljivk pri izpisu.
- V polje *Grouping Variable(s)* prenesemo spremenljivko oz. spremenljivke, katerih vrednosti določajo podskupine. Če določimo več spremenljivk, naredi skupine glede na vse možne kombinacije vrednosti izbranih spremenljivk.
- Ne izpišemo vrednosti enot - ne izberemo opcije *Display cases*.
- S specifikacijami v podoknu, ki ga odpremo z gumbom *Statistics*, lahko določamo, katere statistike naj se izračunajo za vsako skupino in celoto.
- S specifikacijami v podoknu, ki ga odpremo z gumbom *Options*, določamo obliko izpisa (naslove, način izpisa manjkajočih vrednosti itd.).

Izračun statistik po skupinah

Ukaz: *Analyze – Reports – Case Summaries*



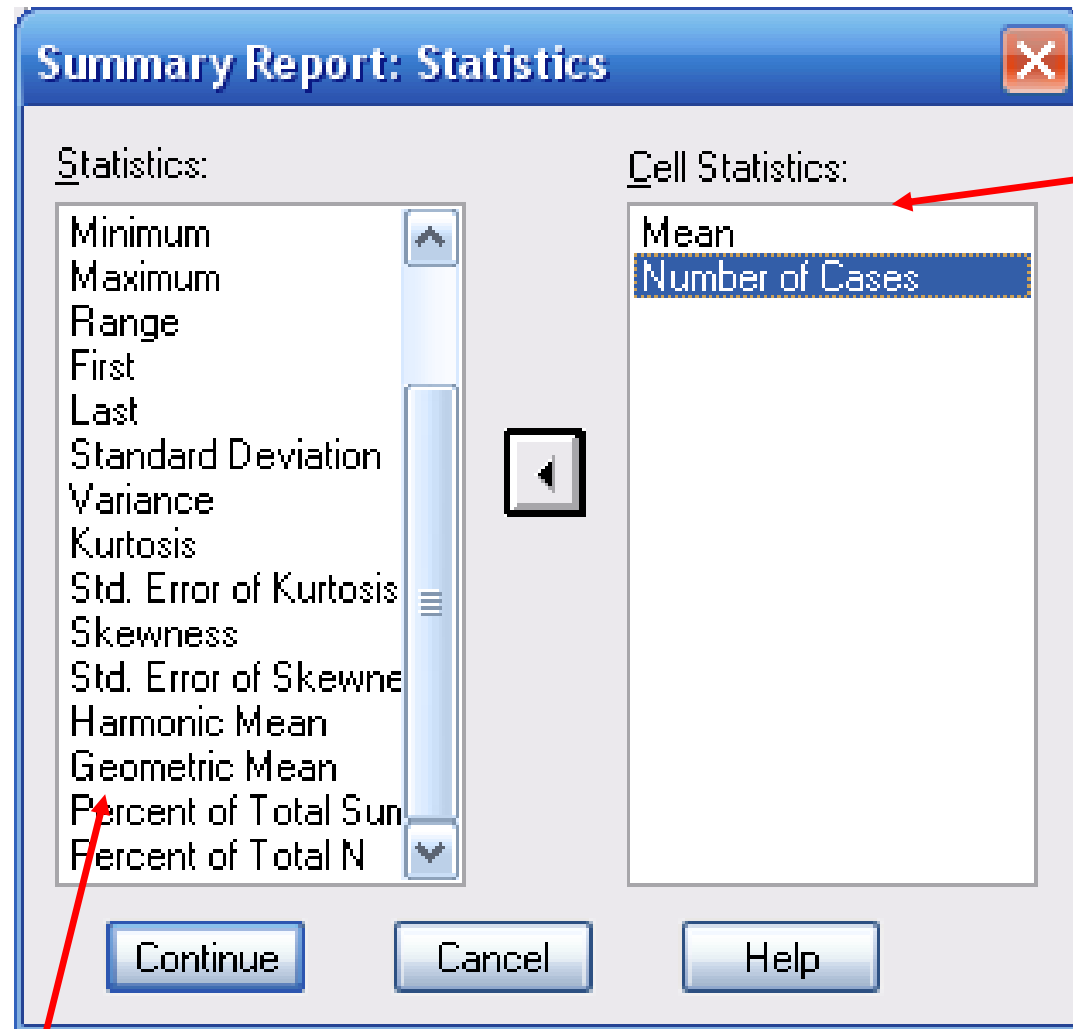
Spremenljivke, za katere želimo izračunati statistike

Spremenljivke, ki naj definirajo skupine, po katerih naj bo izpis narejen (tu spol in kraj bivanja)

Katere statistike se naj izpišejo za v se izpisane enote oz. tudi po skupinah, če so določene

Ali naj izpišemo podatke za posamezne enote (ali samo statistike za celoto in po skupinah).

Izračun statistik po skupinah



Statistike, ki smo jih izbrali.
Izračunale se bodo za vsako
podskupino in za vse enote.

Statistike, ki so na voljo, a niso izbrane.

Izračun statistik po skupinah

Case Summaries

gndr spol	F5 Kraj bivanja		G91 Bruto plača v 1000 sit	G92 Neto plača v 1000 sit
1 moški	1 veliko mesto	Mean	252,00	141,39
		N	14	23
	2 predmestje ali obrobje velikega mesta	Mean	231,13	117,38
		N	24	29
	3 manjše mesto	Mean	288,69	167,10
		N	39	49
	4 vas	Mean	227,09	118,80
N		69	70	
5 kmetija ali hiša na deželi	Mean	188,89	138,74	
	N	27	19	
Total	Mean	237,59	135,77	
	N	173	190	
2 ženski	1 veliko mesto	Mean	252,33	126,46
		N	15	28
	2 predmestje ali obrobje velikega mesta	Mean	209,24	136,85
		N	21	27
	3 manjše mesto	Mean	210,11	121,02
		N	35	46
	4 vas	Mean	202,81	101,53
N		67	83	
5 kmetija ali hiša na deželi	Mean	188,00	74,40	
	N	7	10	
Total	Mean	209,91	113,27	
	N	145	194	
Total	1 veliko mesto	Mean	252,17	133,20
		N	29	51
	2 predmestje ali obrobje velikega mesta	Mean	220,91	126,77
		N	45	56
	3 manjše mesto	Mean	251,53	144,79
		N	74	95
	4 vas	Mean	215,12	109,43
N		136	153	
5 kmetija ali hiša na deželi	Mean	188,71	116,55	
	N	34	29	
Total	Mean	224,97	124,40	
	N	318	384	

Povprečna bruto in neto plača ter število enot z veljavnimi vrednostmi po spolu in kraju bivanja.

Tabela bo bolj pregledna, če jo preoblikujemo.

Izračun statistik po skupinah

Case Summaries

		gndr spol					
		1 moški		2 ženski		Total	
		Mean	N	Mean	N	Mean	N
G91 Bruto plača v 1000 sit	F5 Kraj bivanja						
	1 veliko mesto	252,00	14	252,33	15	252,17	29
	2 predmestje ali obrobje velikega mesta	231,13	24	209,24	21	220,91	45
	3 manjše mesto	288,69	39	210,11	35	251,53	74
	4 vas	227,09	69	202,81	67	215,12	136
	5 kmetija ali hiša na deželi	188,89	27	188,00	7	188,71	34
	Total	237,59	173	209,91	145	224,97	318
G92 Neto plača v 1000 sit	1 veliko mesto	141,39	23	126,46	28	133,20	51
	2 predmestje ali obrobje velikega mesta	117,38	29	136,85	27	126,77	56
	3 manjše mesto	167,10	49	121,02	46	144,79	95
	4 vas	118,80	70	101,53	83	109,43	153
	5 kmetija ali hiša na deželi	138,74	19	74,40	10	116,55	29
	Total	135,77	190	113,27	194	124,40	384



Kako? Procedura EXPLORE

- V meniju izberemo: *Analyze – Descriptive statistics – Explore*.

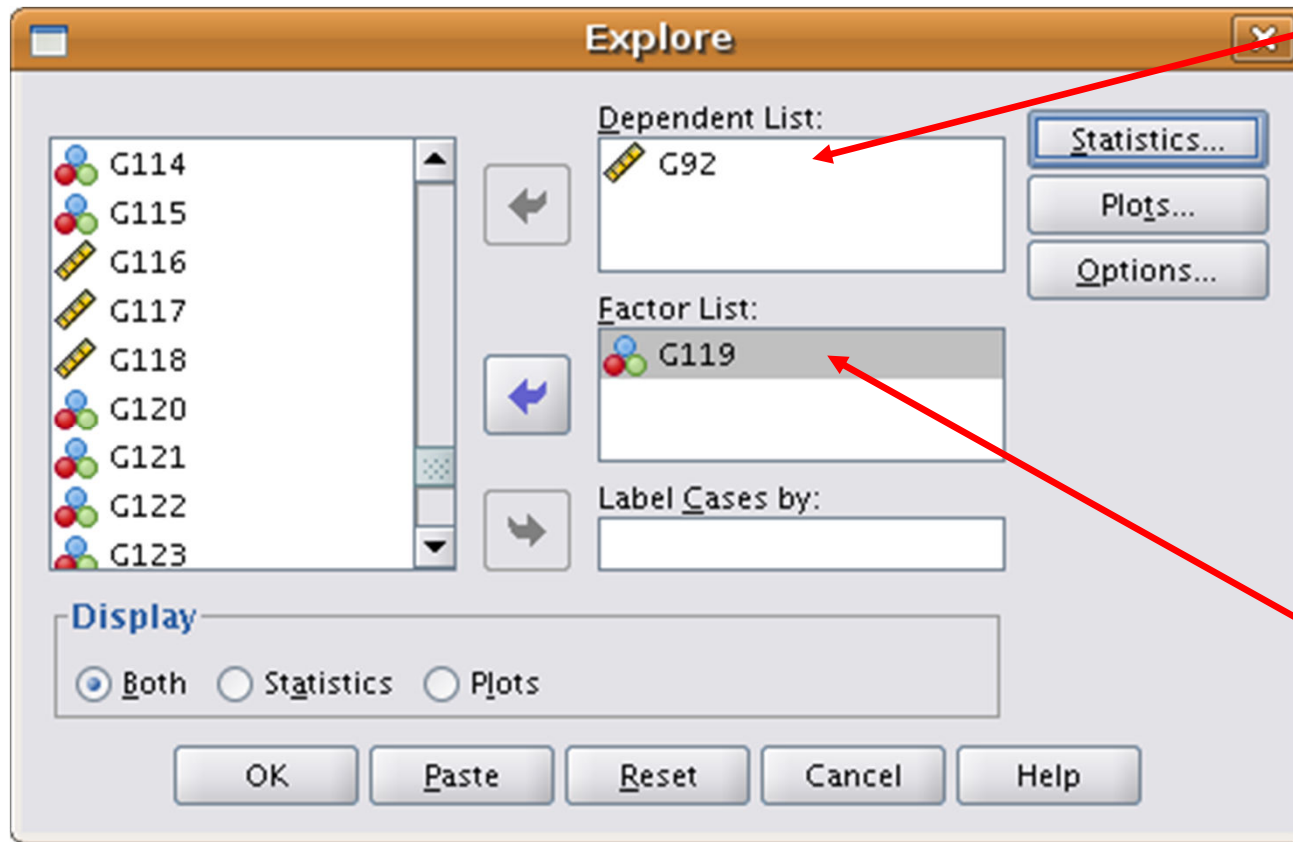
Kaj omogoča?

- Izračuna nekatere opisne statistike + poljuben interval zaupanja za aritmetično sredino ali delež (gumb *Statistics*).
- Izriše lahko histogram ali škatlasti grafikon (gumb *Plots*).
- Analiziramo lahko neko spremenljivko (prenešeno v okno *Dependent list*) za vse enote ali za posamezne skupine enot (določene s spremenljivko, prenešeno v okno *Factor list*).
- Omogoča še dodatne analize in grafe, ki jih zaenkrat ne bomo obravnavali.

Kdaj je to uporabno?

- Če želimo narediti analize za podskupine enot.
- Če želimo izračunati interval zaupanja za aritmetično sredino.

Procedura EXPLORE



Glavna oz.
odvisna
spremenljivka

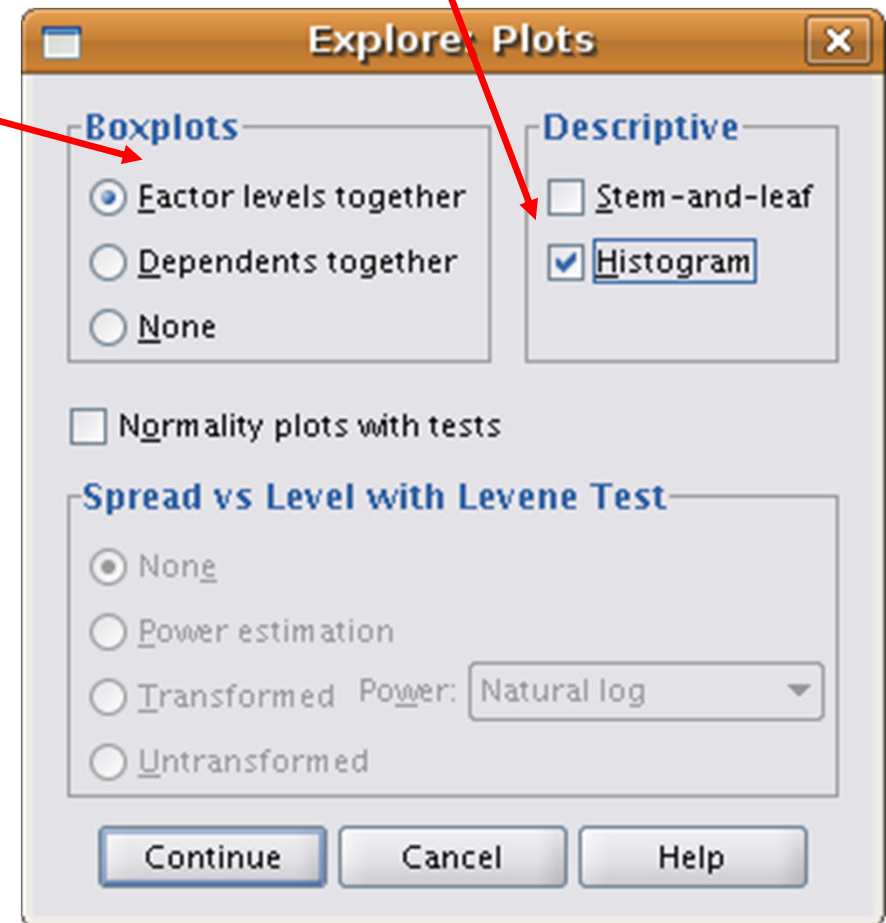
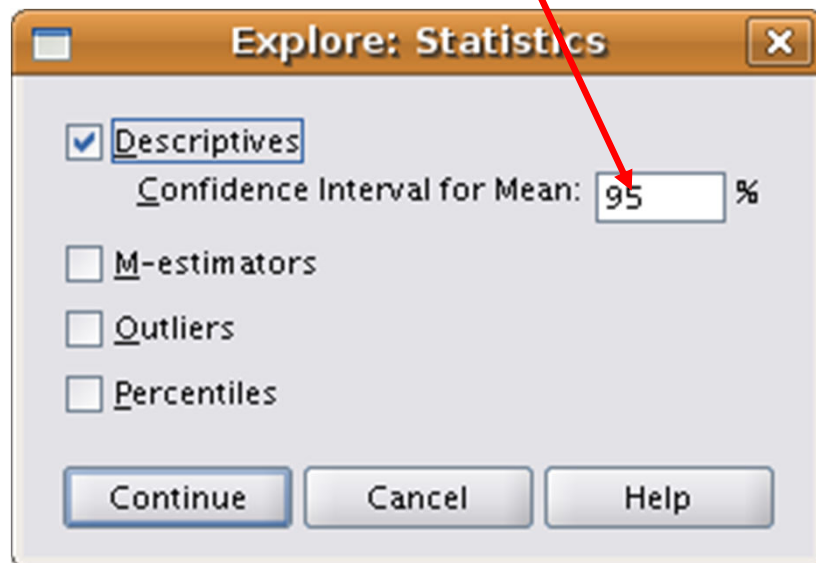
Spremenljivka, ki
določa skupine

Procedura EXPLORE

Izris histograma.

Izris škatlastega grafikona.

Nastavimo stopnjo gotovosti pri intervalu zaupanja za aritmetično sredino oz. delež.





V.3 Oblikovanje izpisa

- Večino izpisov v SPSS-u je mogoče poljubno preoblikovati.
- Glavna obliki izpis sta:
 - tabela (nekaj je tudi tekstovnega izpisa),
 - slika (graf).



Oblikovanje tabel

Pri tabelah lahko:

- Spremenimo besedilo.
- Spremenimo format številčk (npr. št. dec. mest,...) in besedila (pisava, velikost,...).
- Spremenimo, kje se različni elementi pojavljajo (v vrsticah (rows), stolpcih (columns) ali plasteh (layers)).

Oblikovanje začnemo tako, da dvakrat kliknemo na objekt (tabelo).



Oblikovanje tabel

Spreminjanje besedila:

- Kliknemo na celico, v kateri želimo spremeniti besedilo (postane črna).
- Dvakrat kliknemo na besedilo, ki ga želimo spremeniti.
- Besedilo spremenimo.
- Po koncu spreminjanja kliknemo ven iz celice (če kliknemo ven iz izbrane tabele, končamo z urejanjem celotne tabele).

Oblikovanje tabel

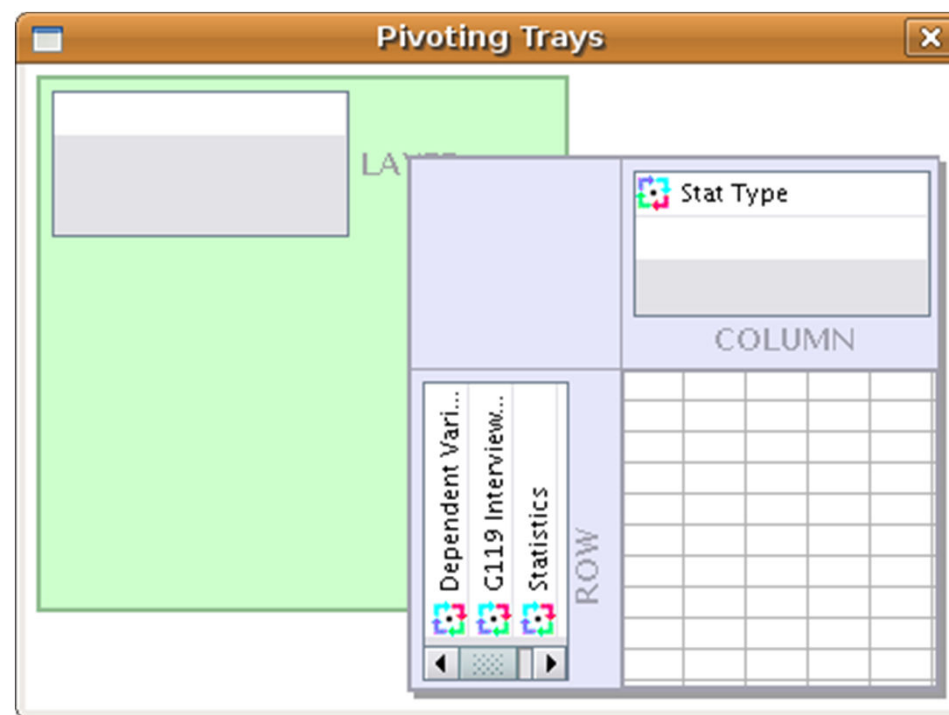
Spreminjanje formata številčk ali oblike besedila:

- Kliknemo na celico, v kateri želimo spremeniti besedilo (postane črna) – lahko naekrat izberemo tudi več celic (z vlečenjem ali držanjem tipke “Shift”).
- Izberemo iz menija “*Format*” (med ostalimi meniji okna) ali iz menija, ki se pojavi ob desnem kliku “*Cell Properties*”.
- Izberemo ustezen zavihek glede na to, kaj želimo oblikovati:
 - Font and Background* za oblikovanje besedila in ozadja,
 - Format Value* za oblikovanje formata številčk (decimalna mesta, zapis datuma, %, ...),
 - Alignment and Margins* za določanje poravnave in prostora ob robu celic.

Oblikovanje tabel

Spreminjanje postavitev določenih elementov:

- “Odpremo” tabelo.
 - V meniju izberemo “*Pivot/Pivoting trays*”.
 - Različne elemente lahko potem poljubno premikamo (sistem „primi“ in „prenesi“) med vrsticami, stolpci in plastmi.
- Določamo lahko tudi vrstni red (hierarhijo elementov) znotraj istega področja (vrstice, stolpci, plasti).



Oblikovanje tabel

Tako lahko na primer iz te tabele ...

Descriptives

G119 Interviewer		Statistic	Std. Error		
G92 Usual net pay in Euros, after deduction for tax/insurance	1 Woman	Mean	115,22	5,972	
		95% Confidence Interval for Mean	Lower Bound		103,44
			Upper Bound		127,00
		5% Trimmed Mean	111,25		
		Median	105,00		
		Variance	6739,532		
		Std. Deviation	82,095		
	Minimum	0			
	Maximum	380			
	Range	380			
	Interquartile Range	100			
	Skewness	,431	,177		
	Kurtosis	-,121	,352		
	2 Man	2 Man	Mean	135,09	7,346
95% Confidence Interval for Mean			Lower Bound	120,59	
			Upper Bound	149,58	
5% Trimmed Mean			128,18		
Median			130,00		
Variance			9821,937		
Std. Deviation			99,106		
Minimum		7			
Maximum		600			
Range		593			
Interquartile Range		101			
Skewness		1,032	,180		
Kurtosis		2,578	,358		

Oblikovanje tabel

naredimo tako tabelo ...

Descriptives

			G119 Interviewer code, gender			
			1 Woman		2 Man	
			Statistic	Std. Error	Statistic	Std. Error
G92 Usual net pay in Euros, after deduction for tax/insurance	Mean		115,22	5,972	135,09	7,346
	95% Confidence Interval for Mean	Lower Bound	103,44		120,59	
		Upper Bound	127,00		149,58	
	5% Trimmed Mean		111,25		128,18	
	Median		105,00		130,00	
	Variance		6739,532		9821,937	
	Std. Deviation		82,095		99,106	
	Minimum		0		7	
	Maximum		380		600	
	Range		380		593	
	Interquartile Range		100		101	
	Skewness		,431	,177	1,032	,180
	Kurtosis		-,121	,352	2,578	,358

ki je veliko bolj pregledna.



Oblikovanje grafov

Pri grafih lahko:

- Spreminjamo besedilo.
- Dodajamo besedilo.
- Brišemo, premikamo in spreminjamo velikost elementom (besedilo, legendo, ...).
- Spreminjamo graf (tip, barve).
- Dodajamo vrednosti na graf (izpis %, frekvenc, ...).

Oblikovanje začnemo tako, da kliknemo na objekt (graf), ki se nato odpre v novem oknu (okno *Chart Editor*).

Oblikovanje grafov

Nekaj splošnih pravil:

- Element (besedilo, del slike) izberemo z enojnim klikom nanj.
- Če je element moč zbrisati/premikati/ /povečati/zmanjšati, se okoli njega pojavi štirioglat okvir, ki ima na ogliščih in srediinah črt kvadratke.
- Element preoblikujemo, kot je običajno pri ostalih Windows programih (prenesemo, spremenimo velikost, zberišemo z pritiskom tipke "Del").
- Če elementa ni mogoče premikati/zbrisati/..., se okoli njega pojavi zaobljen okvir.
- Če je element besedilo, ga (potem ko je že izbran - ima nek okvir) lahko spremenimo še z enim klikom nanj.
- Z dvojnimi klikom na element prikličemo okno, kjer lahko urejamo lastnosti elementa in tudi nekatere lastnosti celotne slike.