

Koostatud juhend on mõeldud lisamaterjalina kasutamiseks „Andmeanalüüsi“ kursuse kuulajatele.

Näidiste ning õpetuste loomisel on kasutatud andmestiku „firma.sav“ andmeid.

Kõik näited põhinevad statistikapaketi SPSS versioonil 18.0. Juhendi koostas K.Osula.

Juhendi sisukord:

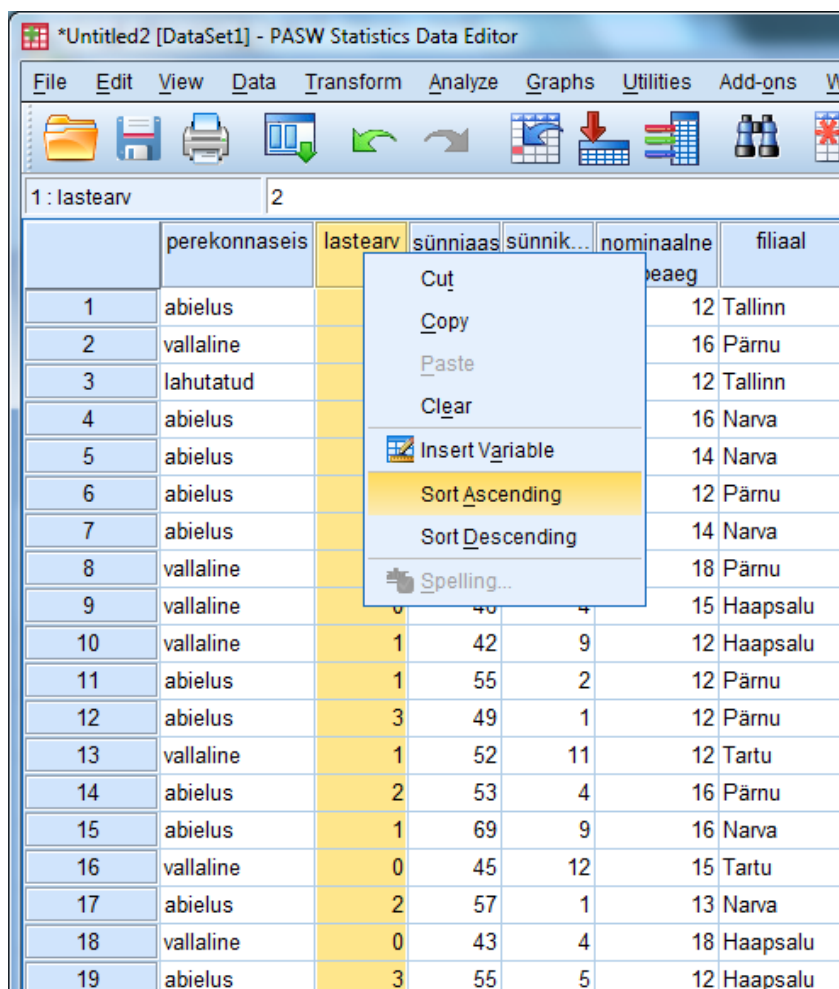
1. Tunnuse väärtuste järjestamine	1
2. Tekstitunnuse muutmine numbriliseks	2
3. Arvutamine.....	3
4. Tunnuse väärtuste jagamine gruppidesse	4
5. Väärtuste selekteerimine	5
6. Andmestiku jagamine osadeks.....	6

1. Tunnuse väärtuste järjestamine

Tee tunnuse **nime peal** hiire parem klõps ning vali:

Sort Ascending: kasvavas järjekorras järjestamiseks (1, 2, 3, 4, ... jne)

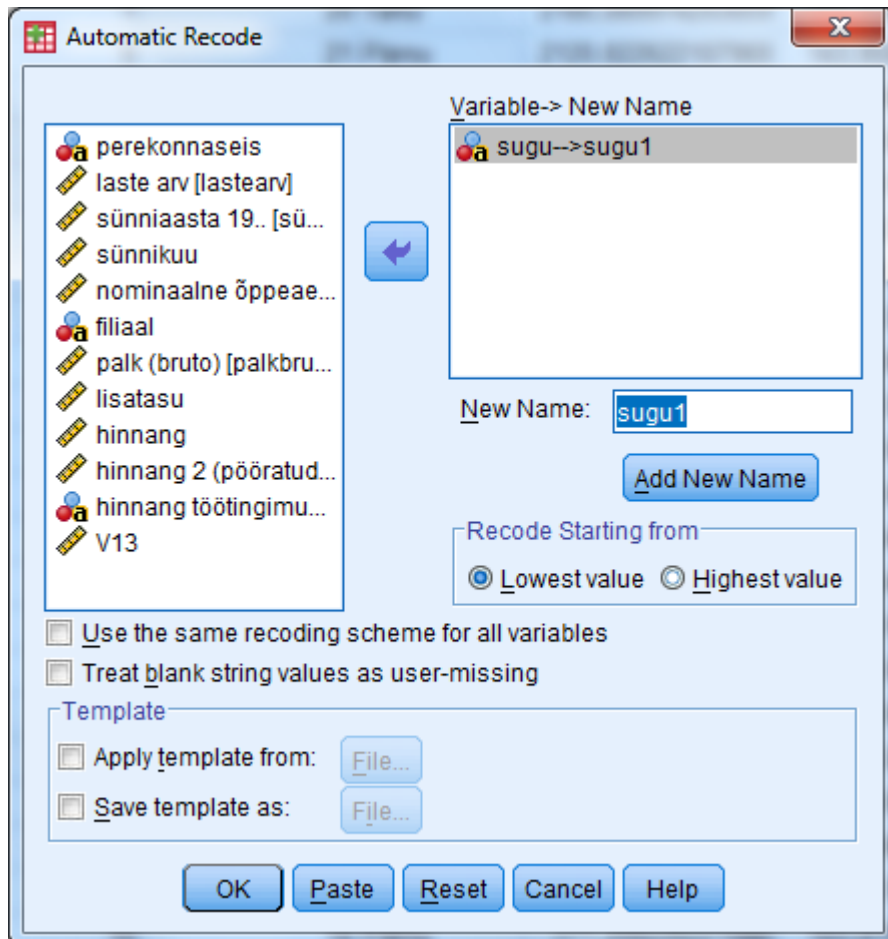
Sort Descending: kahanevas järjekorras järjestamiseks (9, 8, 7, 6, ... jne)



2. Tekstitunnuse muutmine numbriliseks

Tunnuse **Sugu** väärtused olid Exceli andmestikus esitatud sõnadena (mitte kodeeritult). Andmestiku SPSS-i ületoomisel loodi tekstitüüpi (*String*) tunnus. See saada takistuseks edasise andmeanalüüsi tegemisel, seega teisendame tekstitüüpi (*String*) tunnuse numbriliseks (*Numeric*) tunnuseks.

Vali *Transform/Automatic Recode*



Vali tunnus **sugu** ning vii see *Variable* lahtrisse.

Kuna teisenduse käigus luuakse uus tunnus, siis on vaja sellele panna nimi. Sisesta loodava tunnuse nimi lahtrisse *New Name* ning klõpsa *Add New Name* nupul.

Muuta on veel võimalik seda, kas kodeerimist alustatakse kasvavas järjestuses (*Lowest value*) või kahanevas järjestuses (*Highest value*). Koodide määramisel on aluseks vastusevariantide tähestikuline järjestus.

Kasvav järjestus (*Lowest value*) on näiteks selline:

Mees - 1

Naine - 2

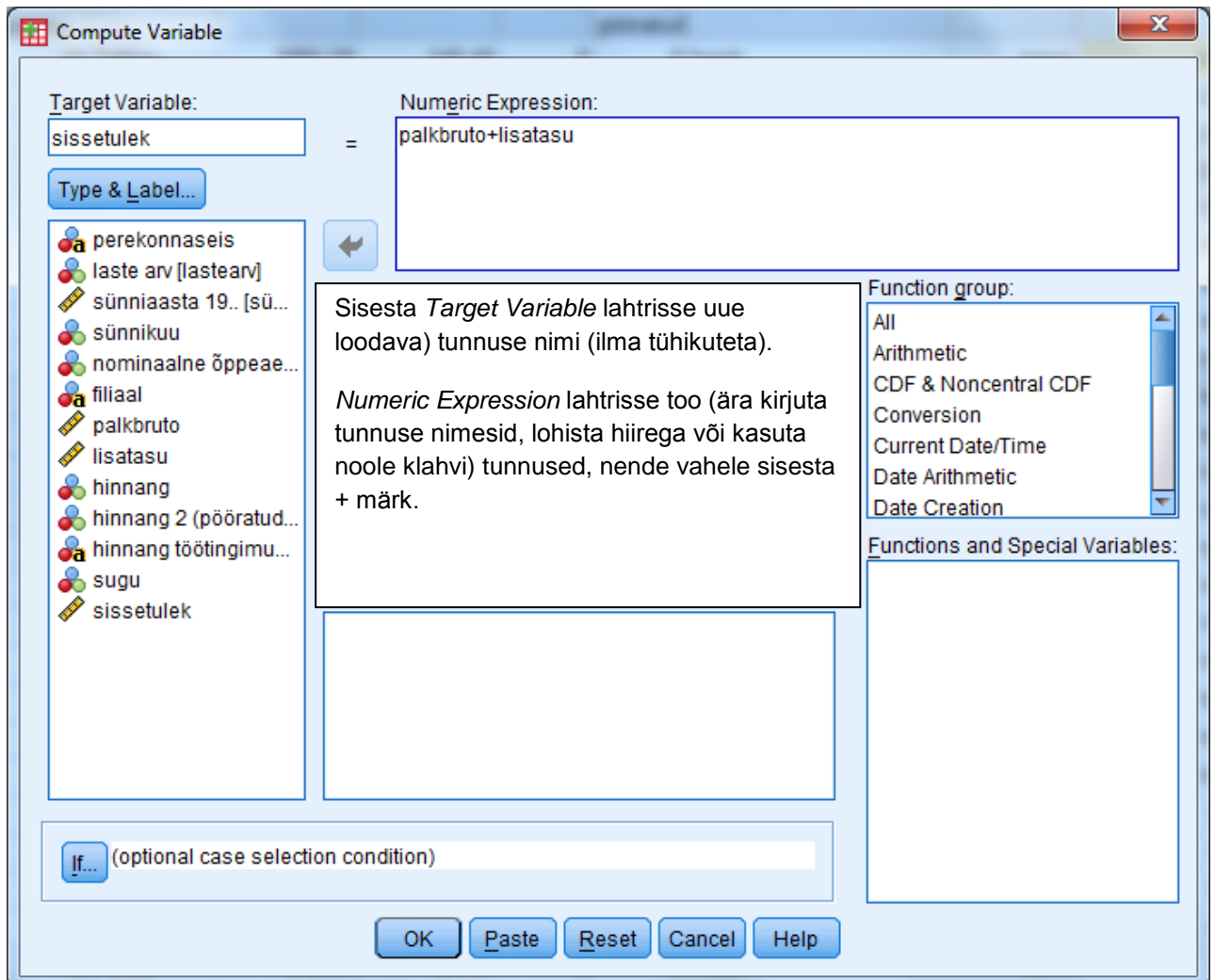
Teisenduse tulemusena loodi SPSS-i tabelisse (kõige lõppu) uus tunnus *sugu1*:

	sugu	perekonnaseis	la	s	s	n	fili	p	lis	hi	hi	hinnangtööt	tingimu	sugu1
			st	ü	ü	o	al	al	at	n	n	stele		
1	mees	vallaline	0	...	5	0	0	head			mees
2	mees	abielus	4	...	1	0	0	head			mees
3	mees	vallaline	1	...	9	7	7	head			mees
4	naine	vallaline	0	4	7	head			naine
5	mees	vallaline	1	...	8	1	7	head			mees

3. Arvutamine

Arvutame vastajate sissetuleku, liites kokku palga ja lisatasu

Vali *Transform/Compute*



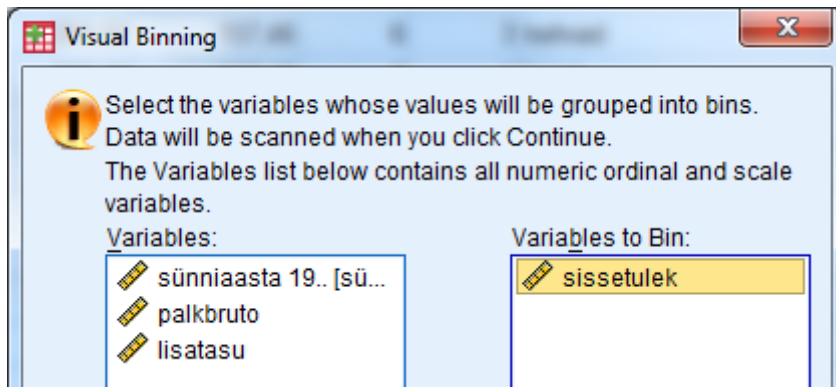
Arvutuse tulemusena loodi SPSS-i tabelisse (kõige lõpu) uus tunnus sissetulek:

palkbruto	lisatasu	hinnang	hinnang2 pööratud	hinnangtöötajimustele	sugu	sissetulek	var
3984,58	148,46	0	0	head	mees	4133,03	
2185,09	152,96	0	0	head	mees	2338,05	
2120,82	143,96	7	7	head	mees	2264,78	
1863,75	152,96	4	7	head	naine	2016,71	
1799,49	157,46	1	7	head	mees	1956,94	

4. Tunnuse väärtuste jagamine gruppidesse

Jagame loodud tunnuse sissetulek väärtuse vahemikesse.

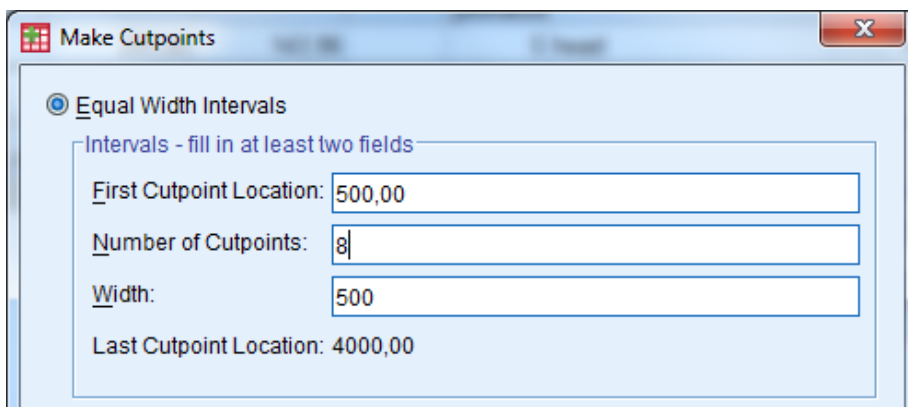
Vali *Transform/Visual Binning*



Vali tunnus sissetulek ja vii see *Variables to Bin* väljale.

Jätkamiseks klõpsa nupul *Continue*

Vahemike loomiseks klõpsa *Make Cutpoints* nupul.



Sisestame loodavate vahemike laiuise *Width* väljale ning esimese vahemiku lõpppunkti asukoha *First Cutpoint Location* lahtrisse. Klõpsa kursor kolmandasse, veel tühja lahtrisse (*Number of Cutpoints*), mille sisu täidab programm automaatselt.

Sisesta *Binned Variable* lahtrisse uue loodava) tunnuse nimi (ilma tühikuteta).

Andmesiltide loomiseks klõpsa *Make Labels* nupul. Andmesildid luuakse automaatselt vastavalt eelnevalt määratud intervalli laiusele.

	Value	Label
1	500,000	<= 500,00
2	1000,000	500,01 - 1000,00
3	1500,000	1000,01 - 1500,00
4	2000,000	1500,01 - 2000,00
5	2500,000	2000,01 - 2500,00
6	3000,000	2500,01 - 3000,00
7	3500,000	3000,01 - 3500,00
8	4000,000	3500,01 - 4000,00

Vahemikeks jagamise tulemusena loodi SPSS-i tabelisse (kõige lõppu) uus tunnus sissetulek1:

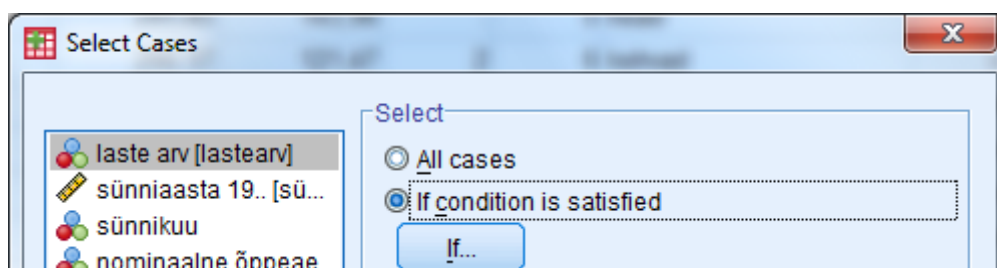
sugu	sissetulek	sissetulek1
mees	435,41	<= 500,00
naine	411,63	<= 500,00
mees	387,85	<= 500,00
naine	427,38	<= 500,00
naine	422,56	<= 500,00
mees	420,95	<= 500,00
naine	447,94	<= 500,00

5. Väärtuste selekteerimine

Oletame, et me soovime uurida mingis analüüsi osas ainult neid, kellel on 1 laps.

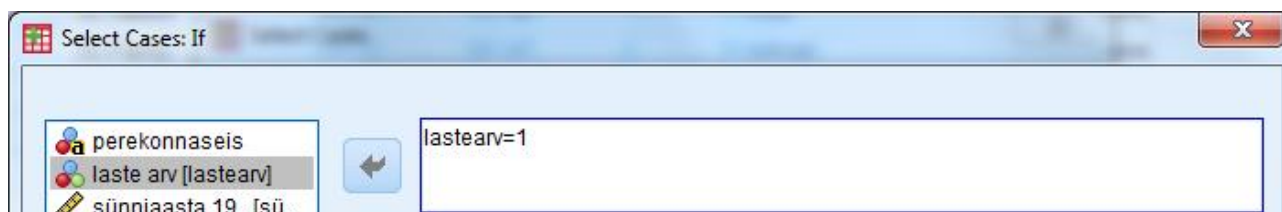
Seega me soovime, et meie andmetabelis oleks vaid need vastajad, kellel on 1 laps.

Vali *Data/Select Cases*



Kasutame selekteerimiseks *If* tingimust.

Need väärtused, mis peaksid andmetabelisse jääma, kirjeldame ära tingimuses: lastearv=1



Selekteerimise tulemusena kasutatakse edaspidi analüüsis vaid neid vastajaid, kellel on 1 laps. Teistele objektidele tõmmati „joon peale“ ning nende vastused analüüsis ei sisaldu.

	perekonnaseis	lastearv	sünniaas ta19	sünnik...	nominaalne õppeaeg	filiaal	palkbruto
1	abielus	2	52	11	12	Tallinn	291,45
2	vallaline	0	47	11	16	Pärnu	290,17
3	lahutatud	3	45	7	12	Tallinn	275,39
4	abielus	2	55	5	16	Narva	269,92
5	abielus	4	55	7	14	Narva	265,10
6	abielus	2	55	12	12	Pärnu	263,50
7	abielus	1	61	6	14	Narva	263,50
8	vallaline	1	47	10	18	Pärnu	263,50
9	vallaline	0	46	4	15	Haapsalu	257,07
10	vallaline	1	42	9	12	Haapsalu	257,07
11	abielus	1	55	2	12	Pärnu	257,07
12	abielus	3	49	1	12	Pärnu	257,07

Kõikide andmete analüüsimiseks e selekteeringu eemaldamiseks kustuta andmetabelist tunnus *filter_\$*

	aline eg	filiaal	palkbruto	lisatasu	hinnaang	hinnaang2 pööratud	hinnaangtöötajumustele	sugu	sissetulek	sissetulek1	filter_\$
1	12	Tallinn	291,45	143,96	.	5	head	mees	435,41	<= 500,00	Not Selected
2	16	Pärnu	290,17	121,47	2	6	kehvad	naine	411,63	<= 500,00	Not Selected
3	12	Tallinn	275,39	112,47	1	3	kehvad	mees	387,85	<= 500,00	Not Selected
4	16	Narva	269,92	157,46	6	3	kehvad	naine	427,38	<= 500,00	Not Selected
5	14	Narva	265,10	157,46	2	4	head	naine	422,56	<= 500,00	Not Selected
6	12	Pärnu	263,50	157,46	2	3	head	mees	420,95	<= 500,00	Not Selected
7	14	Narva	263,50	184,45	4	8	kehvad	naine	447,94	<= 500,00	Selected
8	18	Pärnu	263,50	121,47	2	2	kehvad	naine	384,96	<= 500,00	Selected
9	15	Haapsalu	257,07	116,97	6	7	kehvad	mees	374,04	<= 500,00	Not Selected

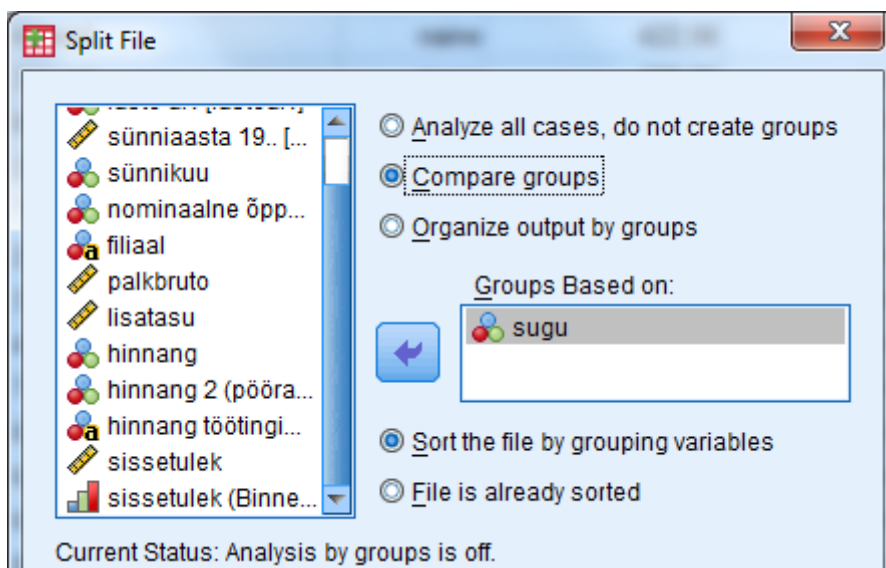
6. Andmestiku jagamine osadeks.

Oletame, et me soovime tellida analüüsi erinevate andmegrupidde kohta (nt. meeste ja naiste jaoks eraldi).

Selleks tuleb järgida skeemi:

* jagame andmestiku osadeks (Split File)

Vali *Data/Split File*



Vali *Compare groups* või *Organize output by groups*. Vt. erinevust lehe alumises osas paiknevas tabelis.

Vii tunnus, mille põhjal andmestik osadeks jagatakse, väljale *Groups Based on* väljale.

* tellime vajaliku analüüsi (sagedustabeli, diagrammi vms)

Compare groups

perekonaseis				
sugu			Frequency	Percent
mees	Valid	abielus	58	56,9
		lahutatud	5	4,9
		vallaline	39	38,2
		Total	102	100,0
naine	Valid	abielus	44	60,3
		lahutatud	2	2,7
		vallaline	27	37,0
		Total	73	100,0

Organize output by groups

perekonaseis ^a				
		Frequency	Percent	
Valid	abielus	58	56,9	
		lahutatud	5	4,9
		vallaline	39	38,2
		Total	102	100,0

a. sugu = mees

perekonaseis ^a				
		Frequency	Percent	
Valid	abielus	44	60,3	
		lahutatud	2	2,7
		vallaline	27	37,0
		Total	73	100,0

a. sugu = naine

* eemaldame gruppideks jagamise (Split File) *Analyze all cases, do not create groups*