

Biomeetria praks 4

Illustreeritud (mittetäielik) tööjuhend

Eeltöö

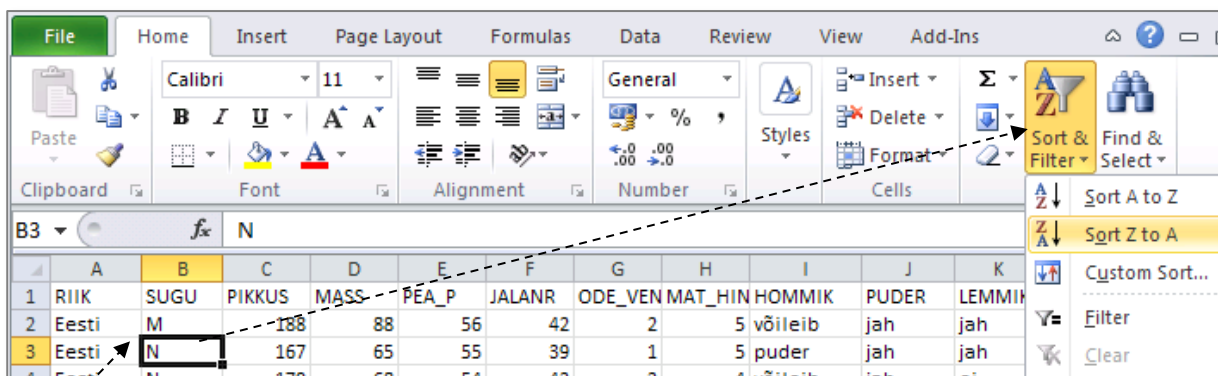
1. Avage MS Excel'is oma kursuse ankeedivastuseid sisaldav andmestik,
2. lisage uus tööleht, nimetage see ümber leheküljeks 'Praks4' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel lehekülje 'Praks4' ülemisse vasakusse nurka.

Ülesanne 1.

Tõlgendades teie kursuse neide kui valimit kogu Maaülikooli esimese kursuse neidudest, uurige, kas Maaülikooli esimese kursuse neidude keskmine pikkus erineb Eesti standardist (Eesti naiste keskmine pikkus on 169 cm)?

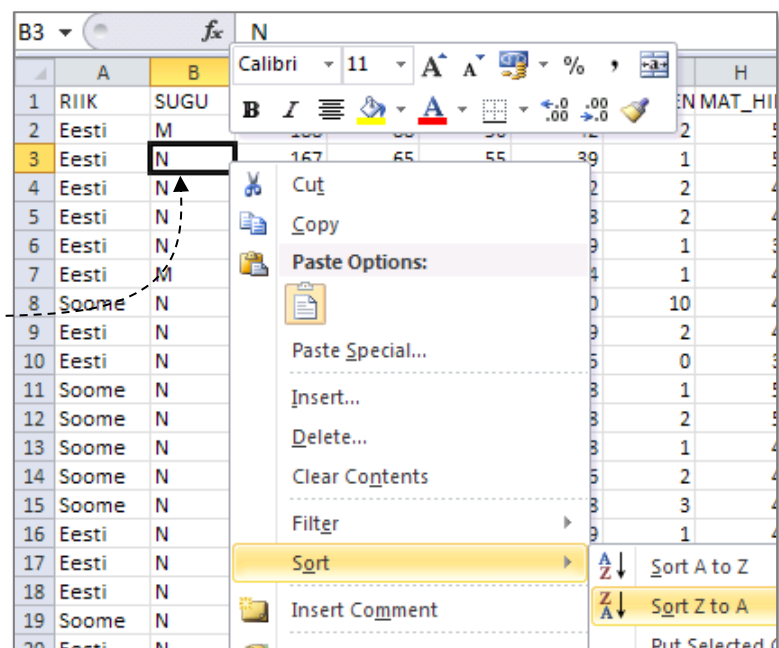
Tööjuhend

1. Sorteerige kogu andmetabel veeru 'SUGU' järgi.



1) Näiteks pannes kursori veergu 'SUGU' ja valides *Home*-sakilt käsud
Sort & Filter ->
Sort Z to A
 (sorteerimaks neide noormeestest ettepoole).

2) Või klikkides suvalisel lahtril veerus 'SUGU' hiire parempoolse klahviga ja valides avanenud rippmenüüst käsud
Sort -> *Sort Z to A*.



2. Leidke neidude arv, nende keskmine pikkus ja pikkuse standardhälve kasutades funktsioone COUNT, AVERAGE ja STDEV.S.

	A	B	C	D	P	Q	R	S	T	U	V	W
1	RIIK	SUGU	PIKKUS	MASS	OLU	SUITS	TEATER	KINO			Neidude pikkus	
2	Eesti	N	180	76	0	ei	viimase a	viimase kuu jooksul			Vaatluste arv	36
3	Eesti	N	178	65	0.25	ei	rohkem kui	viimase aasta jooksul			Keskmine	169.4
4	Eesti	N	165	62	0	ei	viimase a	viimase kuu jooksul			Standardhälve	=STDEV.S(C2:C37)
5	Eesti	N	170	67	0	ei	viimase a	viimase aasta jooksul				
6	Eesti	N	177	59	5	enam ei,	viimase k	viimase kuu jooksul				
7	Eesti	N	166	47	0	ei	viimase a	viimase kuu jooksul				
8	Eesti	N	165	55	0	ei	viimase a	viimase kuu jooksul				
9	Eesti	N	161	49	0	ei	rohkem kui	viimase 10 päeva jooksul				
10	Eesti	N	168	54	0	ei	viimase a	rohkem kui aasta tagasi				
11	Eesti	N	167	67	0	ei	rohkem kui	viimase 10 päeva jooksul				
12	Eesti	N	160	50	0	ei	viimase k	viimase 10 päeva jooksul				
13	Eesti	N	164	53	0	ei	viimase a	viimase aasta jooksul				
14	Eesti	N	171	57	0	enam ei,	rohkem kui	rohkem kui aasta tagasi				
15	Eesti	N	171	65	0	ei	viimase a	viimase 10 päeva jooksul				
16	Eesti	N	183	79	0	ei	viimase k	viimase aasta jooksul				
17	Eesti	N	177	75	2	enam ei,	viimase k	viimase 10 päeva jooksul				
18	Eesti	N	164	54	0	ei	rohkem kui	rohkem kui aasta tagasi				
19	Eesti	N	170	60	0	ei	viimase a	viimase 10 päeva jooksul				
20	Eesti	N	175	65	0	ei	viimase a	viimase aasta jooksul				
21	Eesti	N	168	69	0	ei	viimase a	rohkem kui aasta tagasi				
22	Eesti	N	160	55	0	ei	viimase a	viimase aasta jooksul				
23	Eesti	N	175	69	0	ei	viimase k	viimase kuu jooksul				
24	Eesti	N	151	53	0	ei	viimase a	viimase kuu jooksul				
25	Eesti	N	180	80	0	ei	viimase k	viimase kuu jooksul				
26	Eesti	N	171	72	0	ei	viimase a	viimase 10 päeva jooksul				
27	Eesti	N	169	80	0	enam ei,	viimase a	viimase kuu jooksul				
28	Eesti	N	160	64	0	jah	viimase a	rohkem kui aasta tagasi				
29	Eesti	N	170	65	2	jah	viimase a	viimase 10 päeva jooksul				
30	Eesti	N	176	58	0	ei	rohkem kui	viimase aasta jooksul				
31	Eesti	N	167	55	1	enam ei,	viimase a	viimase aasta jooksul				
32	Eesti	N	160	59	0	ei	viimase a	viimase aasta jooksul				
33	Eesti	N	172	80	0	ei	rohkem kui	rohkem kui aasta tagasi				
34	Eesti	N	170	60	0	jah	rohkem kui	viimase aasta jooksul				
35	Eesti	N	176	57	0	ei	viimase a	viimase 10 päeva jooksul				
36	Eesti	N	168	62	5	jah	rohkem kui	rohkem kui aasta tagasi				
37	Eesti	N	173	68	0	ei	rohkem kui	viimase aasta jooksul				
38	Eesti	M	177	70	5	jah	viimase a	viimase aasta jooksul				
39	Eesti	M	187	75	2	jah	viimase a	viimase aasta jooksul				

Neidude pikkus	
Vaatluste arv	36
Keskmine	169.4
Standardhälve	6.900426

Seega on 36 teie kursuse neiu keskmine pikkus 169,4 cm standardhällbega 6,9 cm

- st, et keskmiselt erineb neidude tegelik pikkus 169,4 sentimeetrist 6,9 cm võrra (vastavalt standardhällbe olemusele);
- ehk, eeldades, et pikkus jaotub normaaljaotuse järgi, jääb vastavalt normaaljaotuse omadustele ligikaudu 68,3% EMÜ esimese kursuse neidude pikkus vahemikku 169,4±6,9 cm ($\bar{x} \pm s$) ja

ligikaudu 95,5% neidude pikkus vahemikku 169,4±13,8 cm ($\bar{x} \pm 2s$).

Neidude pikkus	
Vaatluste arv	36
Keskmine	169,4
Standardhälve	6,900426
H ₀ : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (169 cm)	
H ₁ : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (169 cm)	
või	
H ₀ : EMÜ 1. kursuse neidude keskmine pikkus ei erine 169 sentimeetrist	
H ₁ : EMÜ 1. kursuse neidude keskmine pikkus erineb 169 sentimeetrist	
või	
H ₀ : $\mu_T = 169$	μ_T - EMÜ 1. kursuse neidude keskmine pikkus
H ₁ : $\mu_T \neq 169$	

3. Sõnastage kontrollitav hüpoteeside paar ja pange see leitud arvarakteristikute alla ka kirja.

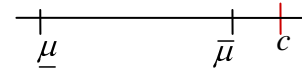
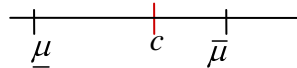
Näiteks:

Meeldetuletuseks teooriast - seos hüpoteeside kontrolli ja usalduspiiride vahel

- Juhul, kui kontrollitavaks hüpoteesiks on mingi andmete alusel hinnatud suuruse võrdlemine konstandiga (mingi arvuga), tehakse otsus sageli baseeruvana uuritava suuruse usaldusintervallil:
 - kui arv, millega andmeist arvatud suurust võrreldakse, jääb usalduspiiride vahele, siis **ei ole alust väita**, et **arvatud suurus erineb** ette antud **konstandist**;
 - kui aga ette antud arv jääb usaldusintervallist väljapoole, **on arvatud suurus konstandist erinev**.
- Näiteks kui soovitakse võrrelda andmeist arvatud keskmist mingi konstandiga (et kas andmed vastavad teatud standardile), on kontrollitav hüpoteeside paar kujul:

$$H_0: \mu = c \text{ ja } H_1: \mu \neq c.$$

Kui nüüd $c \in [\underline{\mu}, \bar{\mu}]$, siis kehtib $H_0: \mu = c$; kui aga $c \notin [\underline{\mu}, \bar{\mu}]$, siis kehtib $H_1: \mu \neq c$.



4. Arvutage liidetav neidude keskmise pikkuse 95% usaldusintervalli leidmiseks (so pool usaldusintervalli laiust) ja tehke seda **kahel viisil**:

a) funktsiooni CONFIDENCE.NORM abil

(funktsioonile tuleb ette anda 3 argumenti:

olulisuse nivoo α , neidude pikkuste standardhälve ja neidude arv);

The image shows two parts of an Excel interface. On the left, a cell contains the text 'Funktsioon CONFIDENCE.NORM'. A dashed box around it contains the text: 'Selle vahepealkirja võiks ise trükkida, et oleks selgem, mis funktsiooni on rakendatud. Ja kursor pange enne funktsiooni tellimist ikka sellesse lahtrisse, kuhu soovite tulemust saada!'. On the right, the 'Insert Function' dialog box is open. The 'Or select a category' dropdown is set to 'Statistical'. In the 'Select a function' list, 'CONFIDENCE.NORM' is selected. Below the list, the function signature is shown: 'CONFIDENCE.NORM(alpha;standard_dev;size)' with a description: 'Returns the confidence interval for a population mean, using a normal distribution.' Buttons for 'OK' and 'Cancel' are at the bottom.

The screenshot shows an Excel spreadsheet with the following data:

	T	U	V	W	X	Y	Z	AA	AB
1		Neidude pikkus							
2		Vaatluste arv	36						
3		Keskmine	169,4						
4		Standardhälve	6,900426						

The 'Function Arguments' dialog box for CONFIDENCE.NORM is open, showing:

- Alpha: 0,05
- Standard_dev: V4
- Size: V2

Formula result = 2,254097598

Note: Kuna tahame leida 95% usaldusintervalli, siis olulisuse nivoo $\alpha = 0,05$.

Formula in cell V2: =CONFIDENCE.NORM(0,05;V4;V2)

Result in cell V2: 2,254098

b) funktsiooni CONFIDENCE.T abil

(funktsiooni süntaks on analoogne funktsiooni CONFIDENCE.NORM süntaksiga).

Tulemus:

Funktsioon CONFIDENCE.NORM	
	2,254098
Funktsioon CONFIDENCE.T	
	2,334768

5. Arvutage alumine ja ülemine usalduspiir mõlema tulemuse alusel.

Z26		fx =V3+V25						
	T	U	V	W	X	Y	Z	AA
1		Neidude pikkus						
2		Vaatluste arv	36					
3		Keskmine	169,4					
4		Standardhälve	6,900426					
5								
6								
7		H ₀ : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (169 cm)						
8		H ₁ : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (169 cm)						
9								
10		või						
11								
12		H ₀ : EMÜ 1. kursuse neidude keskmine pikkus ei erine 169 sentimeetrist						
13		H ₁ : EMÜ 1. kursuse neidude keskmine pikkus erineb 169 sentimeetrist						
14								
15		või						
16								
17		H ₀ : μ _T = 169	μ _T - EMÜ 1. kursuse neidude keskmine pikkus					
18		H ₁ : μ _T ≠ 169						
19								
20								
21		Funktsioon CONFIDENCE.NORM						
22			2,254098	Alumine usalduspiir	167,1348	=V3-V22		
23				Ülemine usalduspiir	171,643			
24		Funktsioon CONFIDENCE.T						
25			2,334768	Alumine usalduspiir	167,0541	=V3-V25		
26				Ülemine usalduspiir	171,7237			
27								

Kumb 95%-usaldusintervallidest on laiem? Miks?

Vastus.

Funktsiooniga CONFIDENCE.T arvatatu on pisut laiem.

Põhjuseks see, et funktsioon CONFIDENCE.T arvutab usalduspiirid t -jaotuse baasil valemist $\bar{x} \pm t_{1-\alpha/2, n-1} \frac{s}{\sqrt{n}}$,

funktsioon CONFIDENCE.NORM arvutab aga **asümptootilised** (ligikaudsed, kusjuures täpsus on seda suurem, mida rohkem on andmeid) ja väikeste valimite korral pisut liiga kitsad usalduspiirid standardse normaaljaotuse baasil valemist $\bar{x} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}}$.

Suurused $t_{1-\alpha/2, n-1} = t_{0,975;35} = 2,030$ ja $z_{1-\alpha/2} = z_{0,975} = 1,96$ näitavad, milline on see väärtus, millest vastava t -jaotuse või siis standardse normaaljaotuse korral on väiksemad 97,5% väärtustest (ehk, millest suuremaid väärtusi võib esineda vaid tõenäosusega 0,025), neist suurustest esimene on *Excel* 2010-s leitav näiteks funktsiooniga =T.INV(0,975;35) ja teine funktsiooniga =NORM.S.INV(0,975).

Kui aru ei saanud, tutvu 3. loengu materjalidega.

NB! Vanemates *Exceli* versioonides ei leidu kumbagi funktsioonidest CONFIDENCE.T ja CONFIDENCE.NORM. Neist teisega, mis leiab usalduspiirid normaaljaotuse baasil, on analoogne funktsioon CONFIDENCE, t -jaotusel baseeruvate usalduspiiride leidmiseks aga vanemais *Exceli* versioonides funktsioon puudub ja kasutada tuleb protseduuri *Descriptive Statistics* valikut *Confidence Level for Mean*, mis arvutab usalduspiiride leidmiseks vajaliku liidetava t -jaotuse baasil (vt eelmise praktikumi viimast ülesannet).

6. Otsus püstitatud hüpoteesi osas – kas EMÜ esimese kursuse neidude keskmine pikkus erineb Eesti standardist (169 cm)?

Sõnastage lõppjäreltus (koos põhjendusega) ja pange see kirja.

Spikker. Et Eesti naiste keskmine pikkus 169 cm jääb EMÜ esimese kursuse neidude keskmine pikkuse 95% usaldusintervalli sisse, $169 \in (167,1; 171,7)$, siis ei ole alust lugeda tõestatuks alternatiivset hüpoteesi keskmine pikkuse erinevusest 169 sentimeetrist ja tuleb jääda nullhüpoteesi H_0 juurde: EMÜ esimese kursuse neidude keskmine pikkus ei erine 169 sentimeetrist.

7. Lisa. Naiste keskmine pikkus maailmas on 154 cm. Kas on alust väita, et EMÜ esimese kursuse tütarlaste pikkused erinevad maailma keskmisest?

Ülesanne 2.

Kas autot omavate ja mitte omavate tudengite keskmised kehamassid on erinevad?

Tööjuhend

1. Moodustage samale töölehele abitabel veergudest 'MASS' ja 'AUTO' ning sorteerige abitabel veeru 'AUTO' järgi.

	A	B	C	D	E	O	P	Q	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	RIIK	SUGU	PIKKUS	MASS	PEA_P	AUTO	OLU	SUITS		Neidude pikkus								MASS	AUTO
2	Eesti	N	180	76	56	jah		0	ei	Vaatluste arv	36								76 jah
3	Eesti	N	178	65	56	ei		0,25	ei	Keskmine	169,4								65 ei
4	Eesti	N	165	62	42	ei		0	ei	Standardhälve	6,900426								62 ei
5	Eesti	N	170	67	55,5	jah		0	ei										67 jah
6	Eesti	N	177	59	42	jah		5	enam ei, a										59 jah
7	Eesti	N	166	47	55	ei		0	ei	H ₀ : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (169 cm)									47 ei
8	Eesti	N	165	55	42	jah		0	ei	H ₁ : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (169 cm)									55 jah
9	Eesti	N	161	49	56	ei		0	ei										49 ei
10	Eesti	N	168	54	50	jah		0	ei	või									54 jah
11	Eesti	N	167	67	56	ei		0	ei										67 ei
12	Eesti	N	160	50	55	jah		0	ei	H ₀ : EMÜ 1. kursuse neidude keskmine pikkus ei erine 169 sentimeetrist									50 jah
13	Eesti	N	164	53	59	ei		0	ei	H ₁ : EMÜ 1. kursuse neidude keskmine pikkus erineb 169 sentimeetrist									53 ei
14	Eesti	N	171	57	50	ei		0	enam ei, a										57 ei
15	Eesti	N	171	65	55	jah		0	ei	või									65 jah
16	Eesti	N	183	79	54	ei		0	ei										79 ei
17	Eesti	N	177	75	55	jah		2	enam ei, a	H ₀ : $\mu_T = 169$	μ_T - EMÜ 1. kursuse neidude keskmine pikkus								75 jah
18	Eesti	N	164	54	52	jah		0	ei	H ₁ : $\mu_T \neq 169$									54 jah
19	Eesti	N	170	60	55	ei		0	ei										60 ei
20	Eesti	N	175	65	57	ei		0	ei										65 ei
21	Eesti	N	168	69	55	ei		0	ei	Funktsioon CONFIDENCE.NORM									69 ei
22	Eesti	N	160	55	55	ei		0	ei		2,254098		Alumine usalduspiir	167,1348	=V3-V22				55 ei
23	Eesti	N	175	69	53	jah		0	ei				Ülemine usalduspiir	171,643					69 jah
24	Eesti	N	151	53	52	jah		0	ei	Funktsioon CONFIDENCE.T									53 jah
25	Eesti	N	180	80	58	jah		0	ei		2,334768		Alumine usalduspiir	167,0541	=V3-V25				80 jah
26	Eesti	N	171	72		jah		0	ei				Ülemine usalduspiir	171,7237					72 jah
27	Eesti	N	169	80	57	jah		0	enam ei, a										80 jah
28	Eesti	N	160	64	54	jah		0	jah										64 jah
29	Eesti	N	170	65	55	jah		2	jah										65 jah
30	Eesti	N	176	58	54	ei		0	ei										58 ei
31	Eesti	N	167	55	54	ei		1	enam ei, a										55 ei
32	Eesti	N	160	59	56	ei		0	ei										59 ei
33	Eesti	N	172	80	54	ei		0	ei										80 ei
34	Eesti	N	170	60		jah		0	jah										60 jah
35	Eesti	N	176	57	56	jah		0	ei										57 jah
36	Eesti	N	168	62	56	ei		5	jah										62 ei
37	Eesti	N	173	68	45	ei		0	ei										68 ei
38	Eesti	M	177	70	57	ei		5	jah										70 ei
39	Eesti	M	187	75	45	jah		2	jah										75 jah
40	Eesti	M	186	74	43	jah		2	jah										74 jah
41	Eesti	M	180	68	51	ei		0	ei										68 ei
42	Eesti	M	194	105	57	jah		1	jah										105 jah
43	Eesti	M	178	65	53	ei		0	enam ei, a										65 ei
44	Eesti	M	177	90	57	jah		0	jah										90 jah
45	Eesti	M	187	99	58	jah		3	enam ei, a										99 jah
46	Eesti	M	189	81	54	ei		2	enam ei, a										81 ei
47	Eesti	M	186	98	56	ei		2	jah										98 ei
48	Eesti	M	183	110	57	jah		0	ei										110 jah
49	Eesti	M	193	100	58	jah		1	ei										100 jah
50	Eesti	M	186	94	58	jah		3	ei										94 jah
51	Eesti	M	198	110	60	ei		0,561	ei										110 ei
52	Eesti	M	174	120	56	jah		0	ei										120 jah
53	muu	M	189	78	57	jah		0	jah										78 jah

Copy -> Paste
+
Sort ...

2. Leidke nii autot omavate ja mitte omavate tudengite arvud, keskmised kehamassid ja kehamassi standardhälbed

Kasutada võite nii vastavaid funktsioone kui ka Pivot Table'i abi. Kui soovite, arvutage mõlemal viisil.

MASS	AUTO		Auto		
			Ei	Jah	
65	ei				
62	ei	Tudengite arv	24	28	
47	ei	Keskmine mass	66,8	75,0	
49	ei	Standardhälve	14,59	18,88	
67	ei				
53	ei				
57	ei				
79	ei				
60	ei	Values			
65	ei	Count of MASS	ei	jah	Grand Total
69	ei	Average of MASS2	24	28	52
55	ei	StdDev of MASS3	66,75	75	71,192308
58	ei		14,59079519	18,879736	17,377062
55	ei				
59	ei				
80	ei				
62	ei				
68	ei				
70	ei				
68	ei				
65	ei				
81	ei				
98	ei				
110	ei				
76	jah				
67	jah				
59	jah				
55	jah				
54	jah				
50	jah				
65	jah				
75	jah				
54	jah				
69	jah				
53	jah				
80	jah				
72	jah				
80	jah				
64	jah				
65	jah				
60	jah				
57	jah				
75	jah				
74	jah				
105	jah				
90	jah				
99	jah				
110	jah				
100	jah				
94	jah				
120	jah				
78	jah				

3. Sõnastage kontrollitav hüpoteeside paar ja pange see ka kirja.

t-test	
H_0 : Autot omavate ja mitte omavate tudengite keskmised massid on võrdsed	
H_1 : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad	
või	
$H_0: \mu_{EI} = \mu_{Jah}$	μ_{EI} - autot mitte omavate tudengite keskmine kehamass
$H_1: \mu_{EI} \neq \mu_{Jah}$	μ_{Jah} - autot omavate tudengite keskmine kehamass

4. Millist t-testi nende keskmiste võrdlemisel kasutada?

NB! t-testi on 3 tüüpi, vt lk 12 (punkt 7 b).

- Et tegu on sõltumatute vaatlustega (võrreldavad grupid koosnevad erinevatest tudengitest), tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada (kas seda, mis eeldab keskmiste võrdlemisel võrdset varieeruvust, või seda, mis arvutab mõlema grupi tarvis eraldi dispersioonid).
- Varieeruvuse (dispersioonide) võrdlemiseks kasutatakse **F-testi**.

5. Pange kirja kontrollitav hüpoteeside paar ja viige läbi **F-test**, otsustamaks autot omavate ja mitte omavate tudengite kehamasside varieeruvuse võrdumise või mittevõrdumise üle (**funktsioon F.TEST**).

NB! Protseduur *F-test (Data-sakk -> Data Analysis... -> F-Test Two-Sample for Variances)* testib vaid ühepoolset hüpoteesi ega ole seetõttu otseselt rakendatav, otsustamaks dispersioonide võrdumise või mittevõrdumise üle.

t-test	
H_0 : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad	
H_1 : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad	
või	
$H_0: \mu_{EI} = \mu_{Jah}$	μ_{EI} - autot mitte omavate tudengite keskmine kehamass
$H_1: \mu_{EI} \neq \mu_{Jah}$	μ_{Jah} - autot omavate tudengite keskmine kehamass
Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.	
F-test (võrdleme dispersioone)	
$H_0: \sigma_{EI}^2 = \sigma_{Jah}^2$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)
$H_1: \sigma_{EI}^2 \neq \sigma_{Jah}^2$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)

MID =F.TEST(AC2:AC25;AC26:AC53)

AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO
	MASS	AUTO				Auto							
1	65	ei			Ei	Jah							
2	62	ei		Tudengite arv	24	28							
3	47	ei		Keskmine mass	66,8	75,0							
4	49	ei		Standardhälve	14,59	18,88							
5	67	ei											
6	53	ei											
7	57	ei											
8	79	ei											
9	60	ei		Column Labels	ei	jah	Grand Total						
10	65	ei		Values	24	28	52						
11	69	ei		Count of MASS	66,75	75	71,192308						
12	55	ei		Average of MASS2	14,59079519	18,879736	17,377062						
13	58	ei		StdDev of MASS3									
14	55	ei											
15	59	ei											
16	80	ei		t-test									
17	62	ei		H ₀ : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad									
18	68	ei		H ₁ : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad									
19	68	ei		või									
20	70	ei		H ₀ : μ _{Ei} =μ _{Jah}	μ _{Ei} - autot mitte omavate tudengite keskmine kehamass								
21	68	ei		H ₁ : μ _{Ei} ≠μ _{Jah}	μ _{Jah} - autot omavate tudengite keskmine kehamass								
22	65	ei											
23	81	ei		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi									
24	98	ei											
25	110	ei		F-test (võrdleme dispersioone)									
26	76	jah		H ₀ : σ ² _{Ei} =σ ² _{Jah}	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)								
27	67	jah		H ₁ : σ ² _{Ei} ≠σ ² _{Jah}	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)								
28	59	jah											
29	55	jah		funktsioon F.TEST	=F.TEST(AC2:AC25;AC26:AC53)								
30	54	jah											
31	50	jah											
32	65	jah											
33	75	jah											
34	54	jah											
35	69	jah											
36	53	jah											
37	80	jah											
38	72	jah											
39	80	jah											
40	64	jah											
41	65	jah											
42	60	jah											
43	57	jah											
44	75	jah											
45	74	jah											
46	105	jah											
47	90	jah											
48	99	jah											
49	110	jah											
50	100	jah											
51	94	jah											
52	120	jah											
53	78	jah											

Function Arguments

F.TEST

Array1 AC2:AC25 = {65;62;47;49;67;53;57;79;60;65;69...

Array2 AC26:AC53 = {76;67;59;55;54;50;65;75;54;69;53...

= 0,212947189

Returns the result of an F-test, the two-tailed probability that the variances in Array1 and Array2 are not significantly different.

Array1 is the first array or range of data and can be numbers or names, arrays, or references that contain numbers (blanks are ignored).

Formula result = 0,212947189

[Help on this function](#)

OK Cancel

6. Sõnastage järeldus F-testist, põhjendage.

F-test (võrdleme dispersioone)	
H ₀ : σ ² _{Ei} =σ ² _{Jah}	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)
H ₁ : σ ² _{Ei} ≠σ ² _{Jah}	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)
funktsioon F.TEST	0,212947189 = p > 0,05 => H ₀ : kehamasside varieeruvus võrreldavais gruppides ei ole erinev

Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.

See ongi põhjendus. Aru saite?

7. Viige läbi t-test võrdlemaks keskmisi kehamasse.

Tehke seda kahel viisil:

a) kasutades funktsiooni T.TEST:

Summary Table:

Values	ei	jah	Grand Total
Count of MASS	24	28	52
Average of MASS2	66,75	75	71,19230769
StdDev of MASS3	14,59079519	18,87973634	17,37706225

t-test
 H_0 : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad
 H_1 : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad

F-test (võrdleme dispersioone)
 $H_0: \sigma^2_{ei} = \sigma^2_{jah}$ (autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)
 $H_1: \sigma^2_{ei} \neq \sigma^2_{jah}$ (autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)

funktsioon F.TEST 0,212947189 = $p > 0,05 \Rightarrow H_0$: kehamasside varieeruvus võrreldavais gruppides ei ole erinev

Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.

Võrdleme keskmisi
 funktsioon T.TEST =T.TEST(AC2:AC25;AC26:AC53;2;2)

Function Arguments

- Array1: AC2:AC25 = {65;62;47;49;67;53;57;79;60;65;69;55}
- Array2: AC26:AC53 = {76;67;59;55;54;50;65;75;54;69;53;80}
- Tails: 2
- Type: 2

Returns the probability associated with a Student's t-Test.

Type is the kind of t-test: paired = 1, two-sample equal variance (homoscedastic) = 2, two-sample unequal variance = 3.

Formula result = 0,08795331

b) vastava statistikaprotseduuri abil (Data-sakk -> Data Analysis... -> t-Test: ...):

Sõltuvate gruppide (paariviisiline) võrdlus; funktsioonis T.TEST tüüp nr 1

Sõltumatute gruppide võrdlus võrdsete dispersioonide eeldusel; funktsioonis T.TEST tüüp nr 2

Sõltumatute gruppide võrdlus mittevõrdsete dispersioonide eeldusel; funktsioonis T.TEST tüüp nr 3

	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK
1		MASS	AUTO				Auto			
2		65	ei			Ei	Jah			
3		62	ei		Tudengite arv	24	28			
4		47	ei		Keskmine mass	66,8	75,0			
5		49	ei		Standardhälve	14,59	18,88			
6		67	ei							
7		53	ei							
8		57	ei							
9		79	ei							
10		60	ei		Values	ei	jah	Grand Total		
11		65	ei		Count of MASS	24	28	52		
12		69	ei		Average of MASS2	66,75	75	71,19230769		
13		55	ei		StdDev of MASS3	14,59079519	18,87973634	17,37706225		
14		58	ei							
15		55	ei							
16		59	ei		t-test					
17		80	ei		H ₀ : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad					
18		62	ei		H ₁ : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad					
19		68	ei		või					
20		70	ei		H ₀ : $\mu_{ei} = \mu_{jah}$	μ_{ei} - autot mitte omavate tudengite keskmine kehamass				
21		68	ei		H ₁ : $\mu_{ei} \neq \mu_{jah}$	μ_{jah} - autot omavate tudengite keskmine kehamass				
22		65	ei							
23		81	ei		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone					
24		98	ei							
25		110	ei		F-test (võrdleme dispersioone)					
26		76	jah		H ₀ : $\sigma^2_{ei} = \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamass)				
27		67	jah		H ₁ : $\sigma^2_{ei} \neq \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamass)				
28		59	jah							
29		55	jah		funktsioon F.TEST	0,212947189	= p > 0,05 => H ₀ : kehamasside varieer			
30		54	jah							
31		50	jah		Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrre					
32		65	jah							
33		75	jah		Võrdleme keskmisi					
34		54	jah							
35		69	jah		funktsioon T.TEST	0,08795331				
36		53	jah							
37		80	jah		protseduur t-test					
38		72	jah							
39		80	jah							
40		64	jah							
41		65	jah							
42		60	jah							
43		57	jah							
44		75	jah							
45		74	jah							
46		105	jah							
47		90	jah							
48		99	jah							
49		110	jah							
50		100	jah							
51		94	jah							
52		120	jah							
53		78	jah							
54										

8. Sõnastage lõppjärelus, põhjendage.

funktsioon T.TEST	0,08795331	$= p > 0,05 \Rightarrow H_0$: autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad	
protseduur t-test	t-Test: Two-Sample Assuming Equal Variances		
		<i>Variable 1</i>	<i>Variable 2</i>
Mean		66,75	75
Variance		212,8913043	356,4444444
Observations		24	28
Pooled Variance		290,41	
Hypothesized Mean Difference		0	
df		50	
t Stat		-1,740328089	
P(T<=t) one-tail		0,043976655	
t Critical one-tail		1,675905025	
P(T<=t) two-tail	0,08795331	$= p > 0,05 \Rightarrow H_0$: autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad	
t Critical two-tail	2,008559112		

Veel võimalusi lõppjäreluse sõnastamiseks:

- „**autot omavate ja autot mitteomavate tudengite keskmised kehamassid ei ole statistiliselt oluliselt erinevad ($p > 0,05$)**“,
so selline pisut korrektsemalt ja teaduslikumalt sõnastatud järelus;
- „kehamass ei sõltu sellest, kas tudengil on auto või mitte“ (vähe teise nurga alt sõnastatud järelus, aga ka õige).

Märkus. Olulisuse tõenäosuse väärtus $p = 0,088$ näitab, et

- väites, et kõigi autoomanikest tudengite (ka nende, kelle kohta meil andmeid ei ole) keskmine kehamass erineb kõigi autot mitteomavate tudengite keskmisest kehamassist, eksiksime 8,8%-lise tõenäosusega;
- eeldades, et autot omavate ja mitteomavate tudengite keskmised kehamassid (arvutatuna üle kõigi tudengite, st üldkogumis) on tegelikult võrdsed, siis tõenäosus saada nii suur kehamasside erinevus, nagu ilmnes meie andmetes lihtsalt tänu juhusele, on 0,088. Et see tõenäosus ei ole väga väike (formaalselt on piiriks 0,05), siis ei ole alust lükata ümber nullhüpoteesi kehamasside võrdsusest.

9. Aga mida nendest protseduuri *t-Test: Two-Sample Assuming Equal Variances* tulemustest veel välja saab lugeda?

- Näiteks keskmise kehamassi, kehamasside dispersiooni ja vaatluste arvu võrreldavates gruppides (seejuures on mõistlik peale analüüsi teostamist kirjutada veergudele peale, mis grupi kohta seal olevad arvud käivad):

	Auto='Ei'	Auto='Jah'
	Variable 1	Variable 2
Mean	66,75	75
Variance	212,8913043	356,4444444
Observations	24	28

- Ülejäänud osa väljatrükkist on seotud juba kahe grupi keskmiste võrdlemisega:

Pooled Variance	290,41	← Vördsete dispersioonide eeldusel arvatud ühine dispersioon
Hypothesized Mean Difference	0	
df	50	← <i>t</i> -statistiku empiiriline (andmete alusel arvatud) väärtus
t Stat	-1,740328089	← 1-poolsele hüpoteesile vastav <i>p</i> -väärtus
P(T<=t) one-tail	0,043976655	← 1-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
t Critical one-tail	1,675905025	← 2-poolsele hüpoteesile vastav <i>p</i> -väärtus
P(T<=t) two-tail	0,08795331	← 2-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
t Critical two-tail	2,008559112	

Kahepoolne hüpotees tähendab statistikas võrdumise ja mittevõrdumise testimist:

$$H_0: \mu_{Ei} = \mu_{Jah},$$

$$H_1: \mu_{Ei} \neq \mu_{Jah}.$$

Ühepoolne hüpotees testib vaid ühepoolset erinevust – kas näiteks üks keskmine on teisest väiksem või mitte. Seejuures konstrueerib *Excel* hüpoteeside paari vastavalt keskmistele väärtustele, testides alati seda, kas suurem keskmine on ikka statistiliselt oluliselt suurem või mitte. Seega on antud juhul testitav ühepoolsete hüpoteeside paar kujul

$$H_0: \mu_{Ei} \geq \mu_{Jah},$$

$$H_1: \mu_{Ei} < \mu_{Jah},$$

sest autoomanikest tudengite keskmine kehamass on suurem.

Olulisuse tõenäosuse *p* asemel võib otsuse vastu võtmisel lähtuda ka teststatistiku empiirilise (andmete alusel arvatud) väärtuse absoluutväärtuse $|t|$ ja teststatistiku kriitilise väärtuse $t_{critical}$ võrdlusest (vt hüpoteeside kontroll, 3. loeng).

Nimelt, kuna $|t| = 1,74 < 2,01 = t_{critical}$, siis ei ole põhjust nullhüpoteesi ümber lükata (*t*-statistiku väärtus jääb sellesse piirkonda, kuhu ta nullhüpoteesi kehtides 95%-tõenäosusega peakski jääma).

10. Kas testides ühepoolset hüpoteesi võinuks lugeda alternatiivse hüpoteesi H_1 tõestatuks?

Kuidas kõlab lõppjärelus testitud ühepoolse hüpoteesi kohta?