

Biomeetria praks 4

Illustreeritud (mittetäielik) tööjuhend

Eeltöö

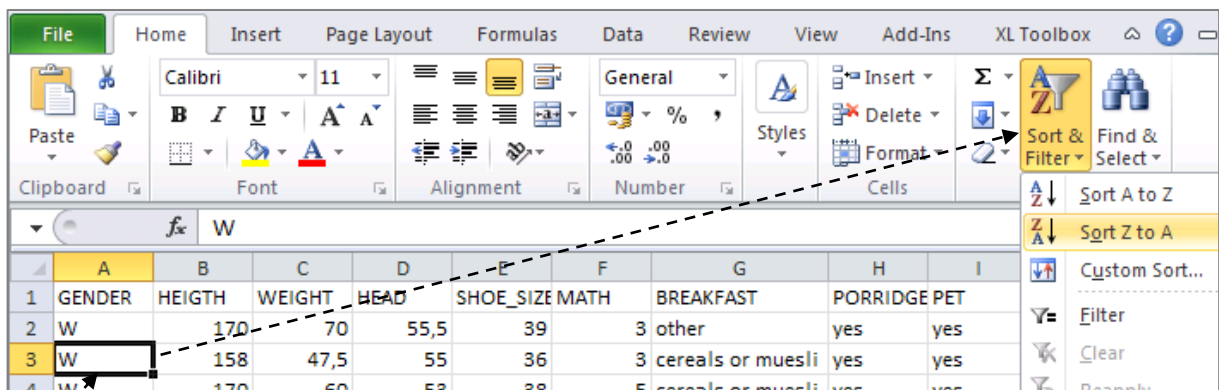
1. Avage MS Excel'is ankeedivastuseid sisaldav andmestik,
2. lisage uus tööleht, nimetage see ümber leheküljeks 'Praks4' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel uue lehekülje 'Praks4' ülemisse vasakusse nurka.

Ülesanne 1.

Oletame, et Veterinaarmeditsiini ja loomakasvatuse instituudi esimese kursuse tudengid ja nende ankeedivastused (see on see analüüsitava ankeedivastuste andmebaas) kujutavad enesest valimit kogu Maaülikooli esimese kursuse tudengitest. Uurige, kas Maaülikooli esimese kursuse neidude keskmine pikkus erineb Eesti standardist (Eesti naiste keskmine pikkus on 169 cm).

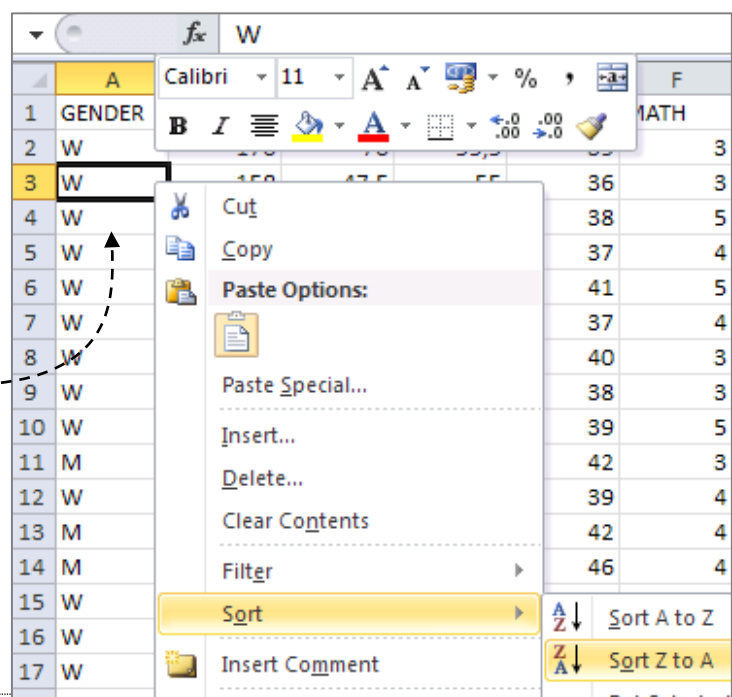
Tööjuhend

1. Sorteerige kogu andmetabel veeru 'SUGU' järgi.

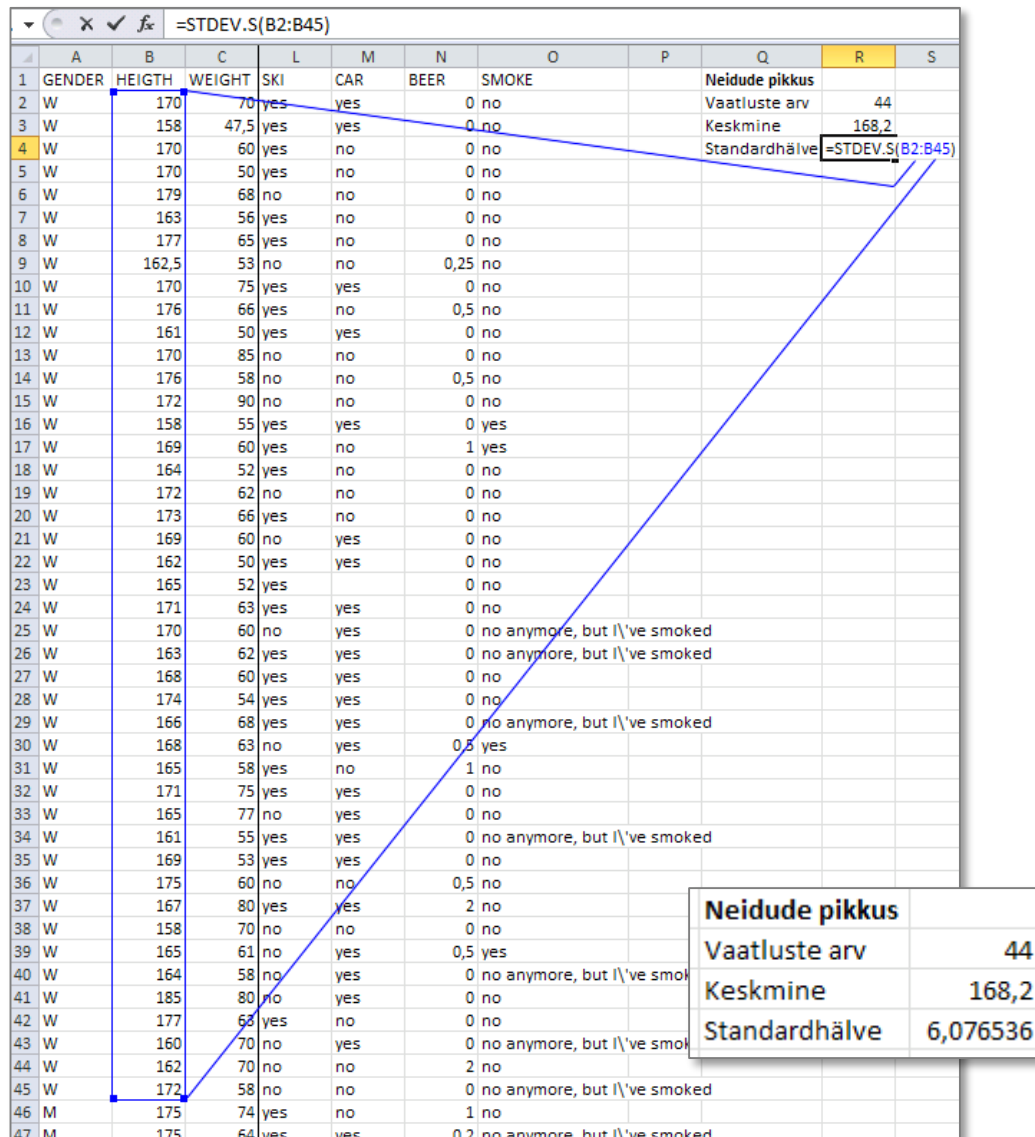


- 1) Näiteks pannes kursori veergu 'GENDER' ja valides *Home*-sakilt käsud
Sort & Filter ->
Sort Z to A
 (sorteerimaks neide noormeestest ettepoole).

- 2) Või klikkides suvalisel lahtril veerus 'GENDER' 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.



Seega on 44 teie andmestikus oleva neiu keskmine pikkus 168,2 cm standardhålbega 6,1 cm

- st, et keskmiselt erineb neidude tegelik pikkus 168,2 sentimeetrist 6,1 cm võrra (vastavalt standardhålbale olemusele);
- ehk, eeldades, et pikkus jaotub normaaljaotuse järgi, jääb vastavalt normaaljaotuse omadustele
 - *) ligikaudu 68,3% EMÜ esimese kursuse neidude pikkus vahemikku $168,2 \pm 6,1$ cm ($\bar{x} \pm s$) ja
 - *) ligikaudu 95,5% neidude pikkus vahemikku $168,2 \pm 12,2$ cm ($\bar{x} \pm 2s$).

Neidude pikkus	
Vaatluste arv	44
Keskmine	168,2
Standardhälve	6,076536
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.

Nåiteks:

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);

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!

Function Arguments

CONFIDENCE.NORM

Alpha: 0,05 = 0,05

Standard_dev: R4 = 6,076535729

Size: R2 = 44

Formula result = 1,795468576

Returns the confidence interval for a population mean, using a normal distribution.

Alpha is the significance level used to compute the confidence interval. It must be greater than 0 and less than 1.

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

Formula result = 1,795468576

Help on this function

OK

Funktsioon CONFIDENCE.NORM

=CONFIDENCE.NORM(0,05;R4;R2)

Funktsioon CONFIDENCE.NORM

1,795469

b) funktsiooni CONFIDENCE.T abil

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

Tulemus:

Funktsioon CONFIDENCE.NORM	
	1,795469
Funktsioon CONFIDENCE.T	
	1,847436

5. Arvutage alumine ja ülemine usalduspiir mõlema tulemuse alusel.

	P	Q	R	S	T	U	V	W
1		Neidude pikkus						
2		Vaatluste arv	44					
3		Keskmine	168,2					
4		Standardhälve	6,076536					
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			1,795469		Alumine usalduspiir	166,4	=R3-R22	
23					Ülemine usalduspiir	170,0		
24		Funktsioon CONFIDENCE.T						
25			1,847436		Alumine usalduspiir	166,4	=R3-R25	
26					Ülemine usalduspiir	170,1		
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 $t_{1-\alpha/2, n-1} = t_{0,975;43} = 2,017$ 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* 2013-s leitav näiteks funktsiooniga =T.INV(0,975;43) 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 lõppjärelus (koos põhjendusega) ja pange see kirja.

Spikker. Et Eesti naiste keskmine pikkus 169 cm jääb EMÜ esimese kursuse neidude keskmise pikkuse 95% usaldusintervalli sisse, $169 \in (166,4; 170,1)$, siis ei ole alust lugeda tõestatuks alternatiivset hüpoteesi keskmise 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 'WEIGHT' ja 'CAR' ning sorteerige abitabel veeru 'CAR' järgi.

	B	C	D	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	HEIGHT	WEIGHT	HEAD	CAR	BEER	SMOKE		Neidude pikkus								WEIGHT	CAR
2	170	70	55,5	yes	0	no		Vaatluste arv	44							60	no
3	158	47,5	55	yes	0	no		Keskmine	168,2							50	no
4	170	60	53	no	0	no		Standardhälve	6,0765							68	no
5	170	50	55	no	0	no										56	no
6	179	68	58	no	0	no										65	no
7	163	56		no	0	no		H_0 : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (163 cm)								53	no
8	177	65	55	no	0	no		H_1 : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (163 cm)								66	no
9	162,5	53	55	no	0,25	no										85	no
10	170	75	56	yes	0	no		või								58	no
11	176	66	57	no	0,5	no										90	no
12	161	50	55	yes	0	no		H_0 : EMÜ 1. kursuse neidude keskmine pikkus ei erine 163 sentimeetrist								60	no
13	170	85	57	no	0	no		H_1 : EMÜ 1. kursuse neidude keskmine pikkus erineb 163 sentimeetrist								52	no
14	176	58	52	no	0,5	no										62	no
15	172	90	58	no	0	no		või								66	no
16	158	55	57	yes	0	yes										58	no
17	169	60	55,5	no	1	yes		H_1 : $\mu_T = 169$	μ_T - EMÜ 1. kursuse neidude keskmine pikkus							60	no
18	164	52	56	no	0	no		H_1 : $\mu_T \neq 169$								70	no
19	172	62	56	no	0	no										63	no
20	173	66	56	no	0	no										70	no
21	169	60	55	yes	0	no		Funktsioon CONFIDENCE.NORM								58	no
22	162	50	50	yes	0	no		1,7955	Alumine usalduspiir	166,4 =R3-R22						74	no
23	165	52	50,5		0	no			Ülemine usalduspiir	170,0						73	no
24	171	63	57	yes	0	no		Funktsioon CONFIDENCE.T								70	yes
25	170	60	53	yes	0	no anymore, but li		1,8474	Alumine usalduspiir	166,4 =R3-R25						47,5	yes
26	163	62	55	yes	0	no anymore, but li			Ülemine usalduspiir	170,1						75	yes
27	168	60	55	yes	0	no										50	yes
28	174	54	55	yes	0	no										55	yes
29	166	68	56	yes	0	no anymore, but li										60	yes
30	168	63	53	yes	0,5	yes										50	yes
31	165	58	56	no	1	no										63	yes
32	171	75	55	yes	0	no										60	yes
33	165	77	58	yes	0	no										62	yes
34	161	55	57	yes	0	no anymore, but li										60	yes
35	169	53	55	yes	0	no										54	yes
36	175	60	57	no	0,5	no										68	yes
37	167	80	57,5	yes	2	no										63	yes
38	158	70	55	no	0	no										75	yes
39	165	61	57	yes	0,5	yes										77	yes
40	164	58	57	yes	0	no anymore, but li										55	yes
41	185	80	60	yes	0	no										53	yes
42	177	63	60	no	0	no										80	yes
43	160	70	57	yes	0	no anymore, but li										61	yes
44	162	70	55	no	2	no										58	yes
45	172	58	62	no	0	no anymore, but li										80	yes
46	175	74	57	no	1	no										70	yes
47	175	64	56	yes	0,2	no anymore, but li										64	yes
48	190	82	58	yes	3	no										82	yes
49	189	82		yes	2,5	yes										82	yes
50	170	80	56	yes	0	no										80	yes
51	176	74	56	yes	0,1	yes										74	yes
52	175	73	54	no	0,1	yes										74	yes
53	181	74	55	yes	1	no										75	yes
54	183	75		yes	3	no										87	yes
55	174	87	57	yes	0,5	no										52	

Copy -> Paste
+
Sort ...

2. Leidke nii autot omavate ja mitte omavate tudengite arvud, keskmised kehamassid ja kehamassi standardhälbed
(NB! Tudeng, kes ei tea, kas tal on auto või mitte, jätke analüüsist välja.)

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

WEIGHT	CAR		Auto		
			Ei	Jah	
60	no				
50	no	Tudengite arv	22	31	
68	no	Keskmine kehamass	64,4	66,6	
56	no	Standardhälve	9,95	11,11	
65	no				
53	no				
66	no				
85	no				
58	no				
90	no				
60	no				
52	no				
62	no				
66	no				
58	no				
60	no				
70	no				
63	no				
70	no				
58	no				
74	no				
73	no				
70	yes				
47,5	yes				
75	yes				
50	yes				
55	yes				
60	yes				
50	yes				
63	yes				
60	yes				
62	yes				
60	yes				
54	yes				
68	yes				
63	yes				
75	yes				
77	yes				
55	yes				
53	yes				
80	yes				
61	yes				
58	yes				
80	yes				
70	yes				
64	yes				
82	yes				
82	yes				
80	yes				
74	yes				
74	yes				
75	yes				
87	yes				
52					

	no	yes	Grand Total
Count of WEIGHT	22	31	53
Average of WEIGHT2	64,40909091	66,59677419	65,68867925
StdDev of WEIGHT3	9,945849055	11,10886985	10,53854236

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)

=F.TEST(Y2:Y23;Y24:Y54)

	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
1		WEIGHT	CAR				Auto				
2		60	no			Ei	Jah				
3		50	no		Tudengite arv		22	31			
4		68	no		Keskmine mass	64,4	66,6				
5		56	no		Standardhälve	9,95	11,11				
6		65	no								
7		53	no								
8		66	no								
9		85	no								
10		58	no		Values	Column Label:	no	yes	Grand Total		
11		90	no		Count of WEIGHT		22	31	53		
12		60	no		Average of WEIGHT2	64,40909091	66,59677419	65,68867925			
13		52	no		StdDev of WEIGHT3	9,945849055	11,10886985	10,53854236			
14		62	no								
15		66	no								
16		58	no		t-test						
17		60	no		H ₀ : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad						
18		70	no		H ₁ : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad						
19		63	no		või						
20		70	no		H ₀ : $\mu_{\text{Ei}} = \mu_{\text{Jah}}$	μ_{Ei} - autot mitte omavate tudengite keskmine kehamass					
21		58	no		H ₁ : $\mu_{\text{Ei}} \neq \mu_{\text{Jah}}$	μ_{Jah} - autot omavate tudengite keskmine kehamass					
22		74	no								
23		73	no		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada						
24		70	yes								
25		47,5	yes		F-test (võrdleme dispersioone)						
26		75	yes		H ₀ : $\sigma^2_{\text{Ei}} = \sigma^2_{\text{Jah}}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)					
27		50	yes		H ₁ : $\sigma^2_{\text{Ei}} \neq \sigma^2_{\text{Jah}}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)					
28		55	yes								
29		60	yes		funktsioon F.TEST	=F.TEST(Y2:Y23;Y24:Y54)					
30		50	yes								
31		63	yes								
32		60	yes								
33		62	yes								
34		60	yes								
35		54	yes								
36		68	yes								
37		63	yes								
38		75	yes								
39		77	yes								
40		55	yes								
41		53	yes								
42		80	yes								
43		61	yes								
44		58	yes								
45		80	yes								
46		70	yes								
47		64	yes								
48		82	yes								
49		82	yes								
50		80	yes								
51		74	yes								
52		74	yes								
53		75	yes								
54		87	yes								
55		52									

Function Arguments

F.TEST

Array1 Y2:Y23 = {60;50;68;56;65;53;66;85;58;90;60...}

Array2 Y24:Y54 = {70;47,5;75;50;55;60;50;63;60;62;...}

= 0,605560196

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 = 0,605560196

[Help on this function](#)

OK Cancel

6. Sõnastage järeldus F-testist, põhjendage.

See ongi põhjendus. Aru saite?

F-test (võrdleme dispersioone)		
H ₀ : $\sigma^2_{\text{Ei}} = \sigma^2_{\text{Jah}}$	(autot omavate	tudengite kehamasside varieeruvus ei ole erinev)
H ₁ : $\sigma^2_{\text{Ei}} \neq \sigma^2_{\text{Jah}}$	(autot omavate	tudengite kehamasside varieeruvus on erinev)
	Otsustusreegel ja otsus	Järeldus
funktsioon F.TEST	0,60556 = p > 0,05 =>	H ₀ : kehamasside varieeruvus võrreldavais gruppides ei ole erinev
Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.		

7. Viige läbi t-test võrdlemaks keskmisi kehamasse.

Tehke seda kahel viisil:

a) kasutades funktsiooni T.TEST:

=T.TEST(Y2:Y23;Y24:Y54;2;2)

X	Y	Z	AA	AB	AC	AD	AE	AF	AG
1	WEIGHT	CAR			Auto				
2	60	no			Ei	Jah			
3	50	no		Tudengite arv	22	31			
4	68	no		Keskmine mass	64,4	66,6			
5	56	no		Standardhälve	9,95	11,11			
6	65	no							
7	53	no							
8	66	no							
9	85	no							
10	58	no		Values	no	yes	Grand Total		
11	90	no		Count of WEIGHT	22	31	53		
12	60	no		Average of WEIGHT2	64,40909091	66,59677419	65,68867925		
13	52	no		StdDev of WEIGHT3	9,945849055	11,10886985	10,59854236		
14	62	no							
15	66	no							
16	58	no		t-test					
17	60	no		H ₀ : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad					
18	70	no		H ₁ : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad					
19	63	no		või					
20	70	no		H ₀ : $\mu_{\text{no}} = \mu_{\text{jah}}$	μ_{no} - autot mitte omavate tudengite keskmine kehamass				
21	58	no		H ₁ : $\mu_{\text{no}} \neq \mu_{\text{jah}}$	μ_{jah} - autot omavate tudengite keskmine kehamass				
22	74	no							
23	73	no		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.					
24	70	yes							
25	47,5	yes		F-test (võrdleme dispersioone)					
26	75	yes		H ₀ : $\sigma^2_{\text{no}} = \sigma^2_{\text{jah}}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)				
27	50	yes		H ₁ : $\sigma^2_{\text{no}} \neq \sigma^2_{\text{jah}}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)				
28	55	yes							
29	60	yes		funktsioon F.TEST	0,60556	= p > 0,05 => H ₀ : kehamasside varieeruvus võrreldavais gruppides.			
30	50	yes							
31	63	yes		Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.					
32	60	yes							
33	62	yes		Võrdleme keskmisi					
34	60	yes							
35	54	yes		funktsioon T.TEST	=T.TEST(Y2:Y23;Y24:Y54;2;2)				
36	68	yes							
37	63	yes							
38	75	yes							
39	77	yes							
40	55	yes							
41	53	yes							
42	80	yes							
43	61	yes							
44	58	yes							
45	80	yes							
46	70	yes							
47	64	yes							
48	82	yes							
49	82	yes							
50	80	yes							
51	74	yes							
52	74	yes							
53	75	yes							
54	87	yes							
55	52								
56									

Function Arguments

T.TEST

Array1 Y2:Y23 = {60;50;68;56;65;53;66;85;58;90;60;52}

Array2 Y24:Y54 = {70;47,5;75;50;55;60;50;63;60;62;...}

Tails 2

Type 2

Testime kahepoolset hüpoteesi!
Võrdse varieeruvuse eeldusel (miks võib seda eeldada?) sobiv t-testi tüüp

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 = 0,464386809

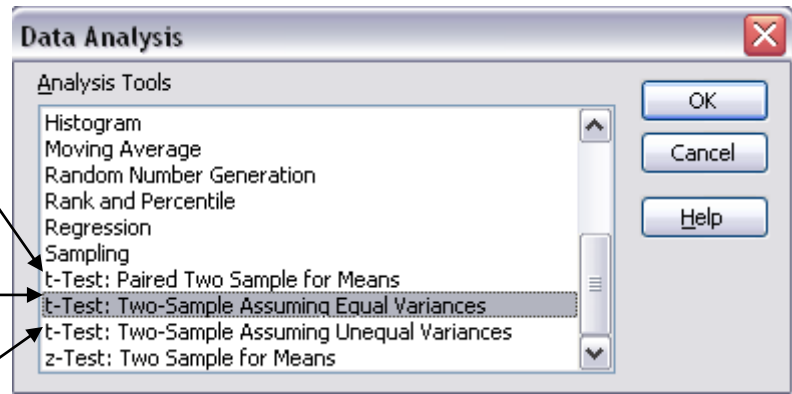
OK Cancel

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

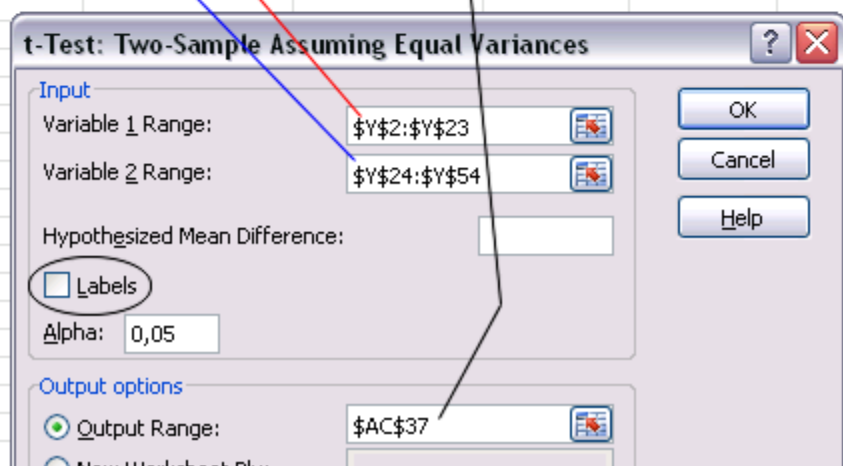
Sõltuvate gruppide (paariviisiline) võrdlus; funktsioonis T.TEST tüüp nr 1

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

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



	X	Y	Z	AA	AB	AC	AD	AE	AF
1		WEIGHT	CAR			Auto			
2		60	no			Ei	Jah		
3		50	no		Tudengite arv	22	31		
4		68	no		Keskmine mass	64,4	66,6		
5		56	no		Standardhälve	9,95	11,11		
6		65	no						
7		53	no						
8		66	no						
9		85	no						
10		58	no		Column Labels				
11		90	no		Values	no	yes	Grand Total	
12		60	no		Count of WEIGHT	22	31	53	
13		52	no		Average of WEIGHT2	64,40909091	66,59677419	65,68867925	
14		62	no		StdDev of WEIGHT3	9,945849055	11,10886985	10,59854236	
15		66	no						
16		58	no		t-test				
17		60	no		H ₀ : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad				
18		70	no		H ₁ : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad				
19		63	no		või				
20		70	no		H ₀ : $\mu_{\text{ei}} = \mu_{\text{jah}}$	μ_{ei} - autot mitte omavate tudengite keskmine kehamass			
21		58	no		H ₁ : $\mu_{\text{ei}} \neq \mu_{\text{jah}}$	μ_{jah} - autot omavate tudengite keskmine kehamass			
22		74	no						
23		73	no		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dis				
24		70	yes						
25		47,5	yes						
26		75	yes		F-test (võrdleme dispersioone)				
27		50	yes		H ₀ : $\sigma^2_{\text{ei}} = \sigma^2_{\text{jah}}$	(autot omavate ja mitte omavate tudengit			
28		55	yes		H ₁ : $\sigma^2_{\text{ei}} \neq \sigma^2_{\text{jah}}$	(autot omavate ja mitte omavate tudengit			
29		60	yes		funktsioon F.TEST	0,60556 = p > 0,05 => H ₀ : keham			
30		50	yes						
31		63	yes		Kasutada võib t-testi, mis eeldab kehamasside võrdset varieer				
32		60	yes						
33		62	yes		Võrdleme keskmisi				
34		60	yes						
35		54	yes		funktsioon T.TEST	0,464386809			
36		68	yes						
37		63	yes		protseduur t-test				
38		75	yes						
39		77	yes						
40		55	yes						
41		53	yes						
42		80	yes						
43		61	yes						
44		58	yes						
45		80	yes						
46		70	yes						
47		64	yes						
48		82	yes						
49		82	yes						
50		80	yes						
51		74	yes						
52		74	yes						
53		75	yes						
54		87	yes						



8. Sõnastage lõppjäreltus, põhjendage.

funktsioon T.TEST	0,464386809	$= p > 0,05 \Rightarrow H_0$: 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	64,40909091	66,59677419	
Variance	98,91991342	123,4069892	
Observations	22	31	
Pooled Variance	113,3240757		
Hypothesized Mean D	0		
df	51		
t Stat	-0,737186899		
P(T<=t) one-tail	0,232193405		
t Critical one-tail	1,67528495		
P(T<=t) two-tail	0,464386809	$= p > 0,05 \Rightarrow H_0$: autoga ja autota tudengite keskmised massid ei ole erinevad	
t Critical two-tail	2,00758377		

Veel võimalusi lõppjäreltuse sõnastamiseks:

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

Märkus. Olulisuse tõenäosuse väärtus $p = 0,464$ näitab, et

- väites, et autoomanikest tudengite keskmine kehamass erineb autot mitteomavate tudengite keskmisest kehamassist, eksiksime 46,4%-lise tõenäosusega;
- eeldades, et autot omavate ja mitteomavate tudengite keskmised kehamassid (arvatuna üle kõigi tudengite, st üldkogumis) on tegelikult võrdsed, siis tõenäosus saada nii suur kehamasside erinevus, nagu ilmnes meie andmetes, lihtsalt tänu juhusele, on 0,464. 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	64,40909091	66,59677419
Variance	98,91991342	123,4069892
Observations	22	31

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

Pooled Variance	113,3240757	← Võrdsete dispersioonide eeldusel arvutatud ühine dispersioon
Hypothesized Mean C	0	
df	51	← <i>t</i> -statistiku empiiriline (andmete alusel arvutatud) väärtus
t Stat	-0,737186899	← 1-poolsele hüpoteesile vastav <i>p</i> -väärtus
P(T<=t) one-tail	0,232193405	← 1-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
t Critical one-tail	1,67528495	← 2-poolsele hüpoteesile vastav <i>p</i> -väärtus
P(T<=t) two-tail	0,464386809	← 2-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
t Critical two-tail	2,00758377	

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 arvutatud) absoluutväärtuse $|t|$ ja teststatistiku kriitilise väärtuse $t_{critical}$ võrdlusest (vt hüpoteeside kontroll, 3. loeng).

Nimelt, kuna $|t| = 1,68 < 2,01 = t_{critical}$, 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?