

## Biomeetria praks 4

### Illustreeritud (mittetäielik) tööjuhend

#### Eeltöö

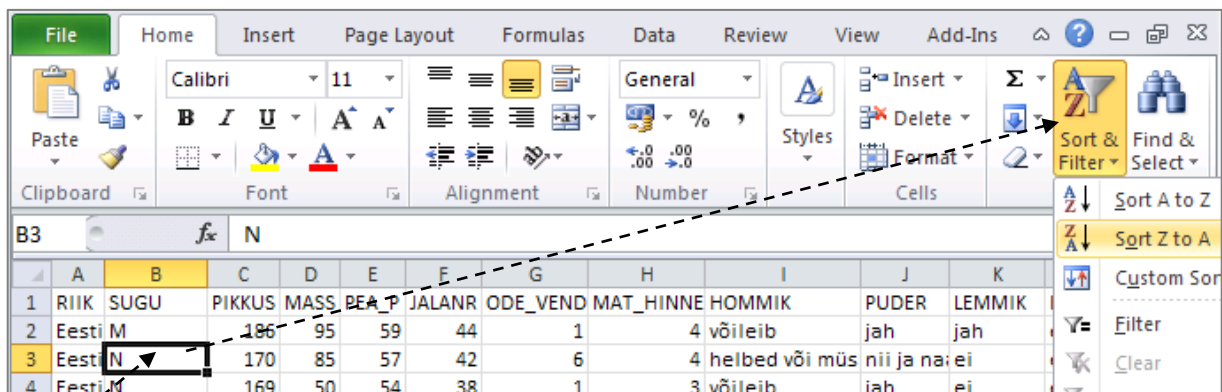
1. Avage MS Excel'is oma kursuse ankeedivastuseid sisaldav andmestik,
2. lisage uus tööleht, nimetage see ümber leheküljeks 'Praks4' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel lehekülje 'Praks4' ülemisse vasakusse nurka.

#### 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 168 cm)?

#### Tööjuhend

1. Sorteerige kogu andmetabel veeru 'SUGU' järgi.



1) Näiteks pannes kursori veergu 'SUGU' ja valides *Home*-sakilt käsud

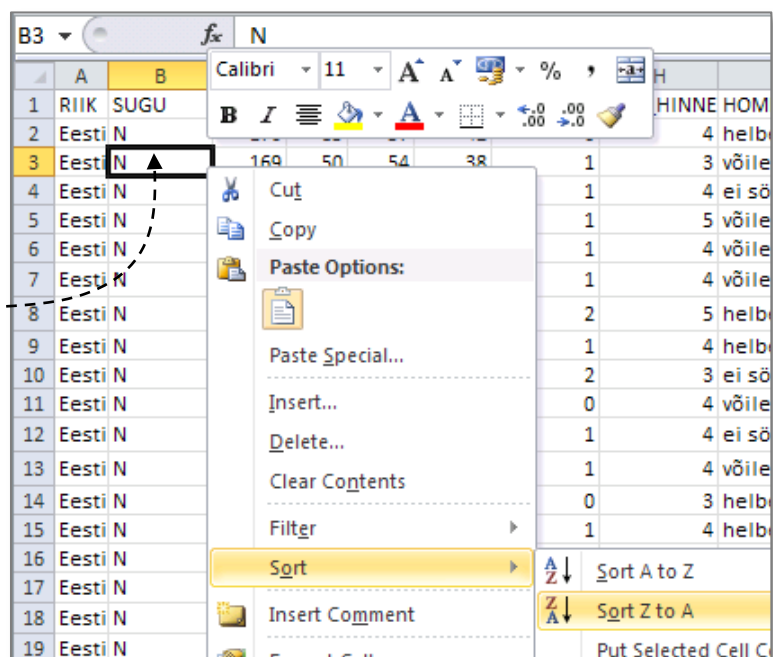
*Sort & Filter* ->

*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	O	P	Q	R	S	T	U	V	W
1	RIIK	SUGU	PIKKUS	MASS	AUTO	OLU	SUITS	TEATER	KINO		Neidude pikkus		
2	Eesti	N	170	85	ei	0	ei	rohkem kui aa	viimase kuu jooksul		Vaatluste arv	29	
3	Eesti	N	169	50	jah	0	ei	viimase aasta	viimase aasta jooksul		Keskmine	169,845	
4	Eesti	N	170	55	ei	0	enam ei	viimase aasta	viimase 10 päeva jooksul		Standardhälve	=STDEV.S(C2:C30)	
5	Eesti	N	160	58	ei	0	ei	rohkem kui aa	viimase kuu jooksul				
6	Eesti	N	161	57	jah	0	ei	rohkem kui aa	viimase 10 päeva jooksul				
7	Eesti	N	171,5	59	ei	0,5	ei	rohkem kui aa	viimase kuu jooksul				
8	Eesti	N	180	63	jah	0	ei	rohkem kui aa	viimase aasta jooksul				
9	Eesti	N	168	54	ei	0	ei	viimase aasta	viimase aasta jooksul				
10	Eesti	N	170	57	ei	0	ei	rohkem kui aa	viimase 10 päeva jooksul				
11	Eesti	N	163	61	ei	0	ei	rohkem kui aa	viimase aasta jooksul				
12	Eesti	N	157	63	ei	0	jah	rohkem kui aa	viimase kuu jooksul				
13	Eesti	N	175	60	ei	1	jah	viimase aasta	viimase kuu jooksul				
14	Eesti	N	176	75	ei	0	ei	rohkem kui aa	viimase aasta jooksul				
15	Eesti	N	170	100	ei	0	ei	rohkem kui aa	viimase aasta jooksul				
16	Eesti	N	169	60	ei	0	jah	rohkem kui aa	rohkem kui aasta tagasi				
17	Eesti	N	163	64	ei	0,5	jah	rohkem kui aa	rohkem kui aasta tagasi				
18	Eesti	N	181	74	ei	0	jah	viimase kuu jo	viimase aasta jooksul				
19	Eesti	N	172	65	ei	0	ei	viimase aasta	viimase kuu jooksul				
20	Eesti	N	173	80	ei	0	enam ei	rohkem kui aa	viimase kuu jooksul				
21	Eesti	N	167	61	ei	0	ei	rohkem kui aa	viimase kuu jook				
22	Eesti	N	171	64	ei	0,5	enam ei	viimase aasta	viimase kuu jook				
23	Eesti	N	173	55	ei	0	ei	viimase aasta	viimase aasta joo				
24	Eesti	N	174	75	jah	0	ei	viimase kuu jo	viimase kuu jook				
25	Eesti	N	162	55	ei	0	ei	viimase kuu jo	rohkem kui aasta				
26	Eesti	N	176	70	jah	0	ei	viimase aasta	viimase kuu jook				
27	Eesti	N	168	58	ei	0	ei	rohkem kui aa	viimase kuu jook				
28	Eesti	N	171	58	ei	1	ei	rohkem kui aa	viimase kuu jooksul				
29	Eesti	N	172	55	ei	0	ei	viimase kuu jo	viimase kuu jooksul				
30	Eesti	N	173	59	jah	0	ei	viimase aasta	viimase aasta jooksul				
31	Eesti	M	185	95	jah	2	enam ei	viimase aasta	viimase aasta jooksul				

Seega on 29 teie kursuse neiu keskmine pikkus 169,8 cm standardhällbega 5,7 cm

- st, et keskmiselt erineb neidude tegelik pikkus 169,8 sentimeetrist 5,7 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  $169,8 \pm 5,7$  cm ( $\bar{x} \pm s$ ) ja ligikaudu 95,5% neidude pikkus vahemikku  $169,8 \pm 11,4$  cm ( $\bar{x} \pm 2s$ ).

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

Näiteks:

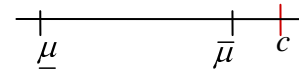
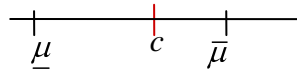
Neidude pikkus	
Vaatluste arv	29
Keskmine	169,8448
Standardhälve	5,680643
H <sub>0</sub> : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (168 cm)	
H <sub>1</sub> : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (168 cm)	
või	
H <sub>0</sub> : EMÜ 1. kursuse neidude keskmine pikkus ei erine 168 sentimeetrist	
H <sub>1</sub> : EMÜ 1. kursuse neidude keskmine pikkus erineb 168 sentimeetrist	
või	
H <sub>0</sub> : $\mu_T = 168$	$\mu_T$ - EMÜ 1. kursuse neidude keskmine pikkus
H <sub>1</sub> : $\mu_T \neq 168$	

### 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  $\alpha$ , 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!



## 5. Arvutage alumine ja ülemine usalduspiir mõlema tulemuse alusel.

	T	U	V	W	X	Y	Z	AA
1		<b>Neidude pikkus</b>						
2		Vaatluste arv	29					
3		Keskmine	169,8448					
4		Standardhälve	5,680643					
5								
6								
7		H <sub>0</sub> : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (168 cm)						
8		H <sub>1</sub> : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (168 cm)						
9								
10		või						
11								
12		H <sub>0</sub> : EMÜ 1. kursuse neidude keskmine pikkus ei erine 168 sentimeetrist						
13		H <sub>1</sub> : EMÜ 1. kursuse neidude keskmine pikkus erineb 168 sentimeetrist						
14								
15		või						
16								
17		H <sub>0</sub> : μ <sub>T</sub> = 168	μ <sub>T</sub> - EMÜ 1. kursuse neidude keskmine pikkus					
18		H <sub>1</sub> : μ <sub>T</sub> ≠ 168						
19								
20								
21		Funktsioon CONFIDENCE.NORM						
22			2,067505		Alumine usalduspiir	167,7773	=V3-V22	
23					Ülemine usalduspiir	171,9123		
24		Funktsioon CONFIDENCE.T						
25			2,160801		Alumine usalduspiir	167,684	=V3-V25	
26					Ülemine usalduspiir	172,0056		
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  $t_{1-\alpha/2, n-1} = t_{0,975;28} = 2,05$  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;28) 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 (168 cm)? **Sõnastage lõppjäreltus (koos põhjendusega) ja pange see kirja.**

Spikker. Et Eesti naiste keskmine pikkus 168 cm jääb EMÜ esimese kursuse neidude keskmise pikkuse 95% usaldusintervalli sisse,  $168 \in (167,7; 172,0)$ , siis ei ole alust lugeda tõestatuks alternatiivset hüpoteesi keskmise pikkuse erinevusest 168 sentimeetrist ja tuleb jääda nullhüpoteesi  $H_0$  juurde: EMÜ esimese kursuse neidude keskmine pikkus ei erine 168 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	N	O	P	Q	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	RIIK	SUGU	PIKKUS	MASS	PEA_P	SUUSK	AUTO	OLU	SUITS		Neidude pikkus								MASS	AUTO
2	Eesti	N	170	85	57	ei	ei	0	ei		Vaatluste arv	29							85	ei
3	Eesti	N	169	50	54	jah	jah	0	ei		Keskmine	169,845							50	jah
4	Eesti	N	170	55	55	ei	ei	0	enam ei, a		Standardhälve	5,68054							55	ei
5	Eesti	N	160	58	55	jah	ei	0	ei										58	ei
6	Eesti	N	161	57	55	ei	jah	0	ei										57	jah
7	Eesti	N	171,5	59	57	ei	ei	0,5	ei		H <sub>0</sub> : EMÜ 1. kursuse neidude pikkused vastavad Eesti standardile (168 cm)								59	ei
8	Eesti	N	180	63	58	jah	jah	0	ei		H <sub>1</sub> : EMÜ 1. kursuse neidude pikkused ei vasta Eesti standardile (168 cm)								63	jah
9	Eesti	N	168	54	57	jah	ei	0	ei										54	ei
10	Eesti	N	170	57	52	ei	ei	0	ei		või								57	ei
11	Eesti	N	163	61	57,5	ei	ei	0	ei										61	ei
12	Eesti	N	157	63	55,5	jah	ei	0	jah		H <sub>0</sub> : EMÜ 1. kursuse neidude keskmine pikkus ei erine 168 sentimeetrist								63	ei
13	Eesti	N	175	60	53	jah	ei	1	jah		H <sub>1</sub> : EMÜ 1. kursuse neidude keskmine pikkus erineb 168 sentimeetrist								60	ei
14	Eesti	N	176	75	56	jah	ei	0	ei										75	ei
15	Eesti	N	170	100	56	jah	ei	0	ei		või								100	ei
16	Eesti	N	169	60	56	ei	ei	0	jah										60	ei
17	Eesti	N	163	64	57	jah	ei	0,5	jah		H <sub>0</sub> : $\mu = 168$	$\mu =$ EMÜ 1. kursuse neidude keskmine pikkus							64	ei
18	Eesti	N	181	74	56	jah	ei	0	jah		H <sub>1</sub> : $\mu \neq 168$								74	ei
19	Eesti	N	172	65	53	jah	ei	0	ei										65	ei
20	Eesti	N	173	80	56	ei	ei	0	enam ei, a										80	ei
21	Eesti	N	167	61	57	ei	ei	0	ei		Funktsioon CONFIDENCE.NORM								61	ei
22	Eesti	N	171	64	54,5	ei	ei	0,5	enam ei, a		2,06751	Alumine usalduspiir	167,777	=V3-V22					64	ei
23	Eesti	N	173	55	57	ei	ei	0	ei			Ülemine usalduspiir	171,912						55	ei
24	Eesti	N	174	75	57	ei	jah	0	ei		Funktsioon CONFIDENCE.T								75	jah
25	Eesti	N	162	55	57	jah	ei	0	ei		2,1608	Alumine usalduspiir	167,684	=V3-V25					55	ei
26	Eesti	N	176	70	46	jah	jah	0	ei			Ülemine usalduspiir	172,006						70	jah
27	Eesti	N	168	58	55	ei	ei	0	ei										58	ei
28	Eesti	N	171	58	55	jah	ei	1	ei										58	ei
29	Eesti	N	172	55	53	ei	ei	0	ei										55	ei
30	Eesti	N	173	59	58	jah	jah	0	ei										59	jah
31	Eesti	M	186	95	59	jah	jah	2	enam ei, a										95	jah
32	Eesti	M	180	70	56	ei	jah	59	jah										70	jah
33	Eesti	M	172	66	54	ei	ei	2	ei										66	ei
34	Eesti	M	183	73	54,5	ei	jah	0	ei										73	jah
35	Eesti	M	185	72	56	jah	ei	1	jah										72	ei
36	Eesti	M	187	94	59	jah	jah	1	ei										94	jah
37	Eesti	M	183	83	56	ei	ei	0	ei										83	ei
38	Eesti	M	190	102	59	jah	ei	3	enam ei, a										102	ei
39	Eesti	M	173	58	55,5	ei	ei	0	enam ei, a										58	ei
40	Eesti	M	180	80	56	jah	jah	0	ei										80	jah
41	Eesti	M	180	84	60	ei	ei	0	ei										84	ei
42	Eesti	M	175	87	54	jah	ei	1	ei										87	ei
43	Eesti	M	181	81	55	ei	ei	0,3	ei										81	ei
44	Eesti	M	177	75	54	ei	jah	1	jah										75	jah
45	Eesti	M	185	100	67	ei	jah	2	jah										100	jah
46	Eesti	M	179	59	56	ei	jah	4	enam ei, a										59	jah
47	Eesti	M	193	75	55	jah	jah	0,5	ei										75	jah
48	Eesti	M	185	80	60	jah	jah	0,25	ei										80	jah
49	Eesti	M	191	70	59	jah	ei	0	ei										70	ei
50	Eesti	M	160	65	56	ei	ei	1	ei										65	ei
51	Eesti	M	173	67	55	ei	jah	1	ei										67	jah
52	Eesti	M	185	100	60	jah	jah	6	ei										100	jah
53	Eesti	M	182	100	ei	jah	20	ei											100	jah
54	Eesti	M	176	66	55	ei	ei	2	ei										66	ei
55	Eesti	M	188	95	59	ei	ei	5	enam ei, a										95	ei
56	Eesti	M	188	80	48	ei	ei	1	jah										80	ei
57	Eesti	M	180	74	56	ei	ei	4	jah										74	ei
58	Eesti	M	179	72	55	jah	jah	0	ei										72	jah

Copy -> Paste  
+  
Sort ...

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

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

MASS	AUTO	Auto	
		Ei	Jah
85	ei		
55	ei		
58	ei	Tudengite arv	36
58	ei	Keskmine mass	68,8
59	ei	Standardhälve	13,28
54	ei		
57	ei		
61	ei		
63	ei		
60	ei	Values	Column Labels, T
75	ei	ei	jah
75	ei	Count of MASS	36
100	ei	Average of MASS2	68,77777778
60	ei	StdDev of MASS3	13,28180519
64	ei		
74	ei		
65	ei		
80	ei		
61	ei		
64	ei		
55	ei		
55	ei		
58	ei		
58	ei		
55	ei		
66	ei		
72	ei		
102	ei		
58	ei		
84	ei		
87	ei		
81	ei		
70	ei		
65	ei		
66	ei		
95	ei		
80	ei		
74	ei		
50	jah		
57	jah		
63	jah		
75	jah		
70	jah		
59	jah		
95	jah		
70	jah		
73	jah		
94	jah		
80	jah		
75	jah		
100	jah		
59	jah		
75	jah		
80	jah		
67	jah		
100	jah		
100	jah		
72	jah		
83			



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

<b>t-test</b>	
H <sub>0</sub> : Autot omavate ja mitte omavate tudengite keskmised massid on võrdsed	
H <sub>1</sub> : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad	
või	
H <sub>0</sub> : $\mu_{EI} = \mu_{Jah}$	$\mu_{EI}$ - autot mitte omavate tudengite keskmine kehamass
H <sub>1</sub> : $\mu_{EI} \neq \mu_{Jah}$	$\mu_{Jah}$ - 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.

<b>t-test</b>	
H <sub>0</sub> : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad	
H <sub>1</sub> : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad	
või	
H <sub>0</sub> : $\mu_{EI} = \mu_{Jah}$	$\mu_{EI}$ - autot mitte omavate tudengite keskmine kehamass
H <sub>1</sub> : $\mu_{EI} \neq \mu_{Jah}$	$\mu_{Jah}$ - autot omavate tudengite keskmine kehamass
Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.	
<b>F-test (võrdleme dispersioone)</b>	
H <sub>0</sub> : $\sigma_{EI}^2 = \sigma_{Jah}^2$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)
H <sub>1</sub> : $\sigma_{EI}^2 \neq \sigma_{Jah}^2$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)



7. Viige läbi t-test võrdlemaks keskmisi kehamasse.

Tehke seda kahel viisil:

a) kasutades funktsiooni T.TEST:

The screenshot shows an Excel spreadsheet with the following data and analysis:

AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
	MASS	AUTO									
1						Auto					
2	85	ei			Ei	Jah					
3	55	ei		Tudengite arv	36	20					
4	58	ei		Keskmine mass	68,8	75,7					
5	59	ei		Standardhälve	13,28	15,23					
6	54	ei									
7	57	ei									
8	61	ei									
9	63	ei									
10	60	ei		Column Labels	.T						
11	75	ei		Values	ei	jah	Grand Total				
12	100	ei		Count of MASS	36	20	56				
13	60	ei		Average of MASS2	68,77777778	75,7	71,25				
14	64	ei		StdDev of MASS3	13,28180519	15,2318918	14,26916574				
15	74	ei									
16	65	ei		t-test							
17	80	ei		H <sub>0</sub> : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad							
18	61	ei		H <sub>1</sub> : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad							
19	64	ei		või							
20	55	ei		H <sub>0</sub> : $\mu_{ei} = \mu_{jah}$	$\mu_{ei}$ - autot mitte omavate tudengite keskmine kehamass						
21	55	ei		H <sub>1</sub> : $\mu_{ei} \neq \mu_{jah}$	$\mu_{jah}$ - autot omavate tudengite keskmine kehamass						
22	58	ei									
23	58	ei		tegu on sõitumate vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamiseks, millist t-testi kasutada.							
24	55	ei									
25	66	ei		F-test (võrdleme dispersioone)							
26	72	ei		H <sub>0</sub> : $\sigma^2_{ei} = \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus ei ole erinev)						
27	102	ei		H <sub>1</sub> : $\sigma^2_{ei} \neq \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamasside varieeruvus on erinev)						
28	58	ei									
29	84	ei		funktsioon F.TEST	0,40840484	= p > 0,05 => H <sub>0</sub> : kehamasside varieeruvus võrreldavais gruppides ei ole erinev					
30	87	ei									
31	81	ei		kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.							
32	70	ei									
33	65	ei		Võrdleme keskmisi							
34	66	ei									
35	95	ei		funktsioon T.TEST	=T.TEST(AC2:AC37;AC38:AC57;2;2)						
36	80	ei									
37	74	ei									
38	50	jah									
39	57	jah									
40	63	jah									
41	75	jah									
42	70	jah									
43	59	jah									
44	95	jah									
45	70	jah									
46	73	jah									
47	94	jah									
48	80	jah									
49	75	jah									
50	100	jah									
51	59	jah									
52	75	jah									
53	80	jah									
54	67	jah									
55	100	jah									
56	100	jah									
57	72	jah									
58	83										

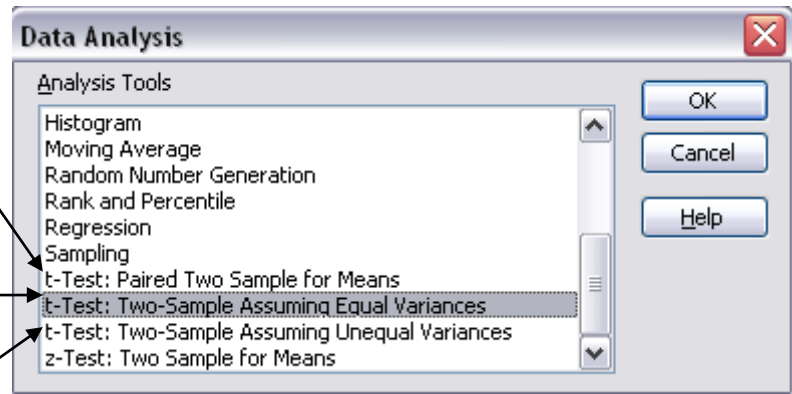
The 'Function Arguments' dialog box for T.TEST shows:

- Array1: AC2:AC37 = {85;55;58;59;54;57;61;63;60;75;100;6}
- Array2: AC38:AC57 = {50;57;63;75;70;59;95;70;73;94;80;...}
- Tails: 2
- Type: 2

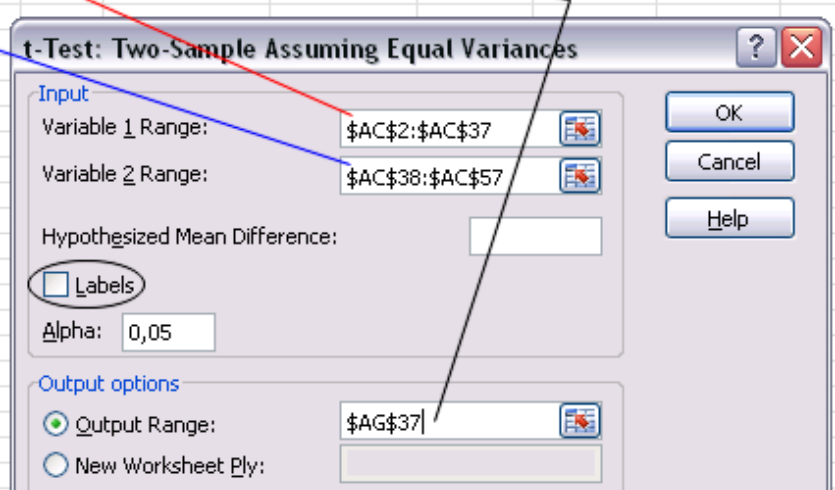
Annotations in the image explain that Type 2 is used for a two-sample unequal variance test, which is appropriate for testing the null hypothesis of equal variances.

**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
1		MASS	AUTO				Auto		
2		85	ei			Ei	Jah		
3		55	ei		Tudengite arv	36	20		
4		58	ei		Keskmine mass	68,8	75,7		
5		59	ei		Standardhälve	13,28	15,23		
6		54	ei						
7		57	ei						
8		61	ei						
9		63	ei						
10		60	ei		Values	Column Labels	.T		
11		75	ei		ei	jah	Grand Total		
12		100	ei		Count of MASS	36	20	56	
13		60	ei		Average of MASS2	68,77777778	75,7	71,25	
14		64	ei		StdDev of MASS3	13,28180519	15,2318918	14,26916574	
15		74	ei						
16		65	ei		t-test				
17		80	ei		H <sub>0</sub> : Autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad				
18		61	ei		H <sub>1</sub> : Autot omavate ja mitte omavate tudengite keskmised massid on erinevad				
19		64	ei		või				
20		55	ei		H <sub>0</sub> : $\mu_{ei} = \mu_{jah}$	$\mu_{ei}$ - autot mitte omavate tudengite keskmine kehamass			
21		55	ei		H <sub>1</sub> : $\mu_{ei} \neq \mu_{jah}$	$\mu_{jah}$ - autot omavate tudengite keskmine kehamass			
22		58	ei						
23		58	ei		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioon				
24		55	ei						
25		66	ei		F-test (võrdleme dispersioone)				
26		72	ei		H <sub>0</sub> : $\sigma^2_{ei} = \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamass)			
27		102	ei		H <sub>1</sub> : $\sigma^2_{ei} \neq \sigma^2_{jah}$	(autot omavate ja mitte omavate tudengite kehamass)			
28		58	ei						
29		84	ei		funktsioon F.TEST	0,40840484	= p > 0,05 => H <sub>0</sub> : kehamasside varieer		
30		87	ei						
31		81	ei		Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrre				
32		70	ei						
33		65	ei		Võrdleme keskmisi				
34		66	ei						
35		95	ei		funktsioon T.TEST	0,08185954			
36		80	ei						
37		74	ei		protseduur t-test				
38		50	jah						
39		57	jah						
40		63	jah						
41		75	jah						
42		70	jah						
43		59	jah						
44		95	jah						
45		70	jah						
46		73	jah						
47		94	jah						
48		80	jah						
49		75	jah						
50		100	jah						
51		59	jah						
52		75	jah						
53		80	jah						
54		67	jah						
55		100	jah						
56		100	jah						
57		72	jah						



### 8. Sõnastage lõppjärelus, põhjendage.

funktsioon T.TEST	0,08185954	<u>= p &gt; 0,05 =&gt; H<sub>0</sub>: autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad</u>	
protseduur t-test	t-Test: Two-Sample Assuming Equal Variances		
		<i>Variable 1</i>	<i>Variable 2</i>
Mean	68,7777778	75,7	
Variance	176,406349	232,0105263	
Observations	36	20	
Pooled Variance	195,970782		
Hypothesized Mea	0		
df	54		
t Stat	-1,7730546		
P(T<=t) one-tail	0,04092977		
t Critical one-tail	1,67356491		
P(T<=t) two-tail	0,08185954	<u>= p &gt; 0,05 =&gt; H<sub>0</sub>: autot omavate ja mitte omavate tudengite keskmised massid ei ole erinevad</u>	
t Critical two-tail	2,00487929		

Veel võimalusi lõppjärelduse 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  $p = 0,082$  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 8,2%-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 0,082. 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	68,77777778	75,7
Variance	176,4063492	232,0105263
Observations	36	20

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

Pooled Variance	195,9707819	← Võrdsete dispersioonide eeldusel arvutatud ühine dispersioon
Hypothesized Mean	0	
df	54	
t Stat	-1,773054598	← t-statistiku empiiriline (andmete alusel arvutatud) väärtus
P(T<=t) one-tail	0,04092977	← 1-poolsele hüpoteesile vastav p-väärtus
t Critical one-tail	1,673564906	← 1-poolsele hüpoteesile vastav t-statistiku kriitiline väärtus
P(T<=t) two-tail	0,08185954	← 2-poolsele hüpoteesile vastav p-väärtus
t Critical two-tail	2,004879288	← 2-poolsele hüpoteesile vastav t-statistiku kriitiline väärtus

**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

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

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

sest autoomanikest tudengite keskmine kehamass on suurem.

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

Nimelt, kuna  $|t| = 1,77 < 2,00 = t_{critical}$ , siis ei ole põhjust nullhüpoteesi ümber liikata

( $t$ -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?**