

# Biomeetria praks 4

## Illustreeritud (mittetäielik) tööjuhend

### Eeltöö

1. Avage MS Excel'is oma kursuse ankeedivastuseid sisaldav (**elmites praktikumides puhastatud**) andmestik,
2. lisage uus tööleht (*Insert / Lisa -> Worksheet / Arvutustabel*), nimetage see ümber leheküljeks 'Praks4' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel lehekülje 'Praks4' ülemisse vasakusse nurka.

### Ülesanne 1.

Kas teie kursuse tütarlaste pikkused vastavad Eesti standardile (Eesti naiste keskmine pikkus on 165 cm)?

### Tööjuhend

1. Sorteerige andmetabel veeru 'SUGU' järgi (*Data / Andmed -> Sort / Sordi*).

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

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	SUGU	PIKKUS	MASS	PEA_P	MAT_HINNE	EBA_AINE	AINEKOOD	PUDER	ÕNNELIK				
2	M	186	75		4	estli keel	Hum	jah	jah				
3	M	182	73	56	4	keemia	Reaal	nii ja naa	jah				
4	M	174	93	58	3	matemaatika	Reaal	jah	jah				
5	M	191	88	60	3	vene keel	Hum	jah	nii ja naa				
6	M	175	70	56	5	vene keel	Hum	jah	nii ja naa				
7	M	187	70	56	3	antsuse ke	Hum	jah	jah				
8	M	197	56		3	inglise keel	Hum	jah	jah				
9	M	185	80	56	3	matemaatika	Reaal	ei	jah				
10	M	179	89	56	4	füüsika	Reaal	nii ja naa	nii ja naa				
11	M	191	80	60	4	kirjandus	Hum	ei	nii ja naa				
12	M	188	78		4	nuusika ajal	Hum	ei	nii ja naa				
13	M	186	80	66	3	keemia	Reaal	jah	jah				
14	M	190	85		3	keemia	Reaal	jah	nii ja naa				
15	M				5	shaline kasv	Muu	jah	er				
16	M	181	69	51	3	keemia	Reaal	jah	jah				
17	M	195	87	58	3	matemaatika	Reaal	jah	jah				
18	M	186	95	50	3	vene keel	Hum	jah	nii ja naa				
19	M	183	80		3	matemaatika	Reaal	jah	nii ja naa				
20	M	179,5	78	52	3	saksa keel	Hum	ei	ei				
21	M	183	85	55	3	keemia	Reaal	ei	nii ja naa				
22	M	177	71		3	keemia	Reaal	jah	jah				
23	M	177	75	56	4	estli keel	Hum	jah	nii ja naa				
24	M	176	80	52	4	saksa keel	Hum	nii ja naa	ei				
25	M	185	100	58	3	matemaatika	Reaal	ei	jah				
26	M	177	75		3	ajalugu	Hum	jah	jah				
27	M	186	76	55	3	keemia	Reaal	jah	jah				
28	M	178	68	57	4	vene keel	Hum	jah	jah				
29	M	173	72	58	4	vene keel	Hum	nii ja naa	nii ja naa				
30	M	186	95	57	3	matemaatika	Reaal	ei	jah				
31	M	197	85	50	3	matemaatika	Reaal	ei	jah				
32	N	165	59		4	füüsika	Reaal	jah	jah				
33	N	165	49		4	kirjandus	Hum	jah	jah				
34	N	171	70	56	3	ajalugu	Hum	ei	jah				
35	N	172,5	60		4	füüsika	Reaal	jah	jah				
36	N	167	75		4			ei	jah				
37	N	175	64	50	3	inglise keel	Hum	jah	jah				
38	N	169	72	55	3	estli keel	Hum	jah	jah				
39	N	157	57		5	geograafia	Hum	ei	jah				
40	N	164	50		4	matemaatika	Reaal	jah	jah				
41	N	177	79	53	3	füüsika	Reaal	ei	nii ja naa				
42	N	162	54		3	hisikonnoop	Hum	ei	nii ja naa				
43	N	168	55	50	3	füüsika	Reaal	jah	jah				
44	N	175	70	56	5	ajalugu	Hum	jah	jah				
45	N	169	55		3	matemaatika	Reaal	jah	ei				
46	N	171	54	53,2	3	keemia	Reaal	ei	jah				
47	N	178	80	56	5	keemia	Reaal	nii ja naa	jah				
48	N	157	48		4	füüsika	Reaal	jah	jah				
49	N	180	67	61	5	soome keel	Hum	ei	jah				
50	N	173	70		4	matemaatika	Reaal	ei	jah				
51	N	175	70		3	matemaatika	Reaal	ei	jah				
52	N	156	42		4	keemia	Reaal	ei	jah				
53	N	164	58		4	bioloogia	Hum	jah	jah				
54	N	165	46		3	estli keel	Hum	jah	nii ja naa				
55	N	177	62	56	3	shaline kasv	Muu	ei	jah				
56	N	165	59	56	3	vene keel	Hum	jah	nii ja naa				
57	N	167	66		4	matemaatika	Rea						
58	N	177	85		4	keemia	Re						
59	N	161	44	56	3	ajalugu	Hu						
60	N	159	50		4	füüsika	Re						
61	N	176	51		3	geograafia	Hu						
62	N	158	50	55	4	vene keel	Hum	jah	jah				
63	N	167	62,5	55	3	füüsika	Re						
64	N	170	60	55	3	füüsika	Re						
65	N	153	65	57	4	ajalugu	Hu						
66	N	170	60	58	3	keemia	Re						
67	N	170	73	57,5	4	saksa keel	Hu						

Tütarlaste pikkus		
Vaatluste arv		36
Keskmine		167,9583
Standardhälve		6,895521

Seega on 36 teie kursuse tütarlaste keskmine pikkus 168,0 cm standardhälbega 6,9 cm

- st, et keskmiselt erineb tütarlaste tegelik pikkus 168,0 sentimeetrist 6,9 cm võrra;
- ehk, eeldades, et pikkus jaotub normaaljaotuse järgi, jääb ligikaudu 68,3% tütarlaste pikkus vahemikku 168,0±6,9 cm ( $\bar{x} \pm s$ ) ja ligikaudu 95,5% tütarlaste pikkus vahemikku 168,0±13,8 cm ( $\bar{x} \pm 2s$ ).

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

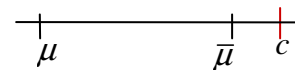
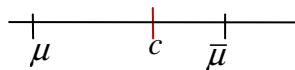
Tütarlaste pikkus	
Vaatluste arv	36
Keskmine	167,9583
Standardhälve	6,895521
H <sub>0</sub> : kursuse tütarlaste pikkused vastavad Eesti standardile (165 cm).	
H <sub>1</sub> : kursuse tütarlaste pikkused ei vasta Eesti standardile (165 cm).	
või	
H <sub>0</sub> : kursuse tütarlaste keskmine pikkus ei erine 165 sentimeetrist.	
H <sub>1</sub> : kursuse tütarlaste keskmine pikkus erineb 165 sentimeetrist.	
või	
H <sub>0</sub> : $\mu_T = 165.$	$\mu_T$ - kursuse tütarlaste keskmine pikkus
H <sub>1</sub> : $\mu_T \neq 165.$	

**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 välja poole, **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 tütarlaste keskmise pikkuse 95% usaldusintervalli leidmiseks (so pool usaldusintervalli laiust) ja tehke seda **kahel viisil**:

a) funktsiooni CONFIDENCE abil (sellele funktsioonile tuleb ette anda 3 argumenti: olulisuse nivoo  $\alpha$ , tütarlaste pikkuste standardhälve ja tütarlaste arv (kelle pikkus oli teada);

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!

**Insert Function**

Search for a function:  
Type a brief description of what you want to do and then click Go  
Or select a category: **Statistical**

Select a function:  
CHIDIST  
CHIINV  
CHITEST  
**CONFIDENCE**  
CORREL  
COUNT  
COUNTA

**CONFIDENCE(alpha;standard\_dev;size)**  
Returns the confidence interval for a population mean.

Help on this function OK Cancel

	J	K	L	M	N	O	P	Q	R	S	T
1											
2		<b>Tütarlaste pikkus</b>									
3		Vaatluste arv	36								
4		Keskmine	167,9583								
5		Standardhälve	6,895521								
6											
7											
8		HO: kursuse tütarlaste pikkused va									
9		H1: kursuse tütarlaste pikkused ei									
10											
11		või									
12											
13		HO: kursuse tütarlaste keskmine p									
14		H1: kursuse tütarlaste keskmine p									
15											
16		või									
17											
18		H <sub>0</sub> : $\mu_T = 165$ .									
19		H <sub>1</sub> : $\mu_T \neq 165$ .									
20											
21											
22		Funktsioon CONFIDENCE									
23		=CONFIDENCE(0,05;L5;L3)									

**Function Arguments**

CONFIDENCE

Alpha 0,05 = 0,05

Standard\_dev L5 = 6,895521321

Size L3 = 36

= 2,252495574

Returns the confidence interval for a population mean.

Size is the sample size.

Formula result = 2,252495574

Help on this function OK Cancel

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

Funktsioon CONFIDENCE  
2,252496

**b) statistikaprotseduuri *Descriptive Statistics* abil**  
 (märkides üksnes lisavaliku *Confidence Level for Mean*).

*Tools / Tööriistad -> Data Analysis... -> Descriptive Statistics*

Ka selle vahepealkirja võiks ise trükkida, eristamiseks erinevate meetoditega arvutatud usalduspiire.

	A	B	C	D	E	F	G	H	I	J	K	L
21	M	183	85	55	3	keemia	Reaal	ei	nii ja naa			
22	M	177	71		3	keemia	Reaal	jah	jah		Funktsioon CONFIDENCE	
23	M	177	75	56	4	eesti keel	Hum	jah	nii ja naa			2,2525
24	M	176	80	52	4	saksa keel	Hum	nii ja naa	ei			
25	M	185	100	58	3	matemaatik	Reaal	ei	jah			
26	M	177	75		3	ajalugu	Hum	jah	jah		Descriptive Statistics	
27	M	186	76	55	3	keemia	Reaal	jah	jah			
28	M	178	68	57	4	vene keel	Hum	jah	jah			
29	M	173	72	58	4	vene keel	Hum	nii ja naa	nii ja naa			
30	M	186	95	57	3	matemaatik	Reaal	ei	jah			
31	M	197	85	50	3	matemaatik	Reaal	ei	jah			
32	N	165	59		4	füüsika	Reaal	jah	jah			
33	N	165	49		4	kirjandus	Hum	jah	jah			
34	N	171	70	56								
35	N	172,5	60									
36	N	167	75									
37	N	175	64	50								
38	N	169	72	55								
39	N	157	57									
40	N	164	50									
41	N	177	79	53								
42	N	162		54								
43	N	168	55	50								
44	N	175	70	56								
45	N	169	55									
46	N	171	54	53,2								
47	N	178	80	56								
48	N	157	48									
49	N	180	67	61								
50	N	173	70									
51	N	175	70									
52	N	158	42									
53	N	164	58									
54	N	165	46									
55	N	177	62	56								
56	N	165	59	56								
57	N	167	66									
58	N	177	85									
59	N	161	44	56								
60	N	159	50									
61	N	175	51		3	geograafia	Hum	jah	nii ja naa			
62	N	158	50	55	4	vene keel	Hum	ei	jah			
63	N	167	52,5	55	3	füüsika	Reaal	jah	jah			
64	N	170	60	55	3	füüsika	Reaal	ei	jah			
65	N	153	65	57	4	ajalugu	Hum	jah	jah			
66	N	170	60	58	3	keemia	Reaal	jah	jah			
67	N	170	73	57,5	4	saksa keel	Hum	jah	ei			

**Descriptive Statistics**

Input  
 Input Range: \$B\$32:\$B\$67

Grouped By:  
 Columns  
 Rows

Labels in first row

Output options  
 Output Range: \$K\$27  
 New Worksheet Ply:  
 New Workbook

Summary statistics  
 Confidence Level for Mean: 95 %  
 Kth Largest: 1  
 Kth Smallest: 1

OK  
 Cancel  
 Help

Descriptive Statistics	
Column1	
Confidence Level(95	2,333109

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

	J	K	L	M	N	O	P	Q
1								
2		<b>Tütarlaste pikkus</b>						
3		Vaatluste arv	36					
4		Keskmine	167,9583					
5		Standardhälve	6,895521					
6								
7								
8		H <sub>0</sub> : kursuse tütarlaste pikkused vastavad Eesti standardile (165 cm).						
9		H <sub>1</sub> : kursuse tütarlaste pikkused ei vasta Eesti standardile (165 cm).						
10								
11		või						
12								
13		H <sub>0</sub> : kursuse tütarlaste keskmine pikkus ei erine 165 sentimeetrist.						
14		H <sub>1</sub> : kursuse tütarlaste keskmine pikkus erineb 165 sentimeetrist.						
15								
16		või						
17								
18		H <sub>0</sub> : $\mu_T = 165$ .	$\mu_T$ - kursuse tütarlaste keskmine pikkus					
19		H <sub>1</sub> : $\mu_T \neq 165$ .						
20								
21								
22		Funktsioon CONFIDENCE						
23			2,252496		Alumine usalduspiir	165,7058	=L4-L23	
24					Ülemine usalduspiir	170,2108		
25								
26		Descriptive Statistics						
27		Column1			Alumine usalduspiir	165,6252	=L4-L29	
28					Ülemine usalduspiir	170,2914		
29		Confidence Level(95	2,333109					
30								

## Kumb 95% -usaldusintervallidest on laiem? Miks?

Vastus. Protseduuriga *Descriptive Statistics* arvatatu on pisut laiem.

Põhjuseks see, et protseduur *Descriptive Statistics* arvutab usalduspiirid  $t$ -jaotuse baasil valemist  $\bar{x} \pm t_{1-\alpha/2, n-1} \frac{s}{\sqrt{n}}$ , funktsioon CONFIDENCE arvutab aga **asümptootilised** (ligikaudsed, kusjuures täpsus on seda suurem, mida rohkem on andmeid) usalduspiirid standardse normaaljaotuse baasil valemist  $\bar{x} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}}$ .

Suurused  $t_{1-\alpha/2, n-1} = t_{0,975;35} = 2,03$  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).

**Kui aru ei saanud, tutvu 3. loengu materjalidega.**

6. Otsus püstitatud hüpoteesi osas – kas teie kursuse tütarlaste pikkused vastavad Eesti standardile (165 cm)? **Sõnastage lõppjärelus (koos põhjendusega) ja pange see kirja.**

Spikker. Et Eesti naiste keskmine pikkus 165 cm ei jää teie kursuse tütarlaste keskmise pikkuse 95% usaldusintervalli sisse,  $165 \notin (165,6; 170,3)$ , siis võib olulisuse nivool 0,05 lugeda tõestatuks alternatiivse hüpoteesi  $H_1$ : kursuse tütarlaste keskmine pikkus erineb 165 sentimeetrist.

**Mida tähendab väljend „olulisuse nivool 0,05“?**

Agaseda, et kui väita, et kursuse tütarlaste pikkus ei vasta Eesti standardile (165 cm), on eksimise tõenäosus väiksem kui 5%.

## Ülesanne 2.

Kas mannaputru söövate ja mittesöövate tudengite keskmised kehamassid on erinevad?

### Tööjuhend

- Moodustage samale töölehele abitabel veergudest 'MASS' ja 'PUDER' ning sorteerige abitabel pudru järgi.

SUGU	PIKKUS	MASS	PEA_P	MAT	HINNE	EBA_AINE	AINEKOOD	PUDER	ÖNNELIK				MASS	PUDER
M	188	75		4		eesti keel	Hum	jah	jah	Tütarlaste pikkus			80	ei
M	182	73	56	4		keemia	Reaal	nii ja naa	jah	Vaatluste arv	36		80	ei
M	174	93	58	3		matemaatika	Reaal	jah	jah	Keskmine	167,9583		78	ei
M	191	88	60	3		vene keel	Hum	jah	nii ja naa	Standardhälve	6,895521		78	ei
M	175	70	56	5		vene keel	Hum	jah	nii ja naa				85	ei
M	187	70	56	3		antsuse keel	Hum	jah	jah				100	ei
M	187	56	3			inglise keel	Hum	jah	jah	H0: kursuse tütarlaste pikkused vastavad Eesti standardile (165 cm).			95	ei
M	185	80	56	3		matemaatika	Reaal	ei	jah	H1: kursuse tütarlaste pikkused ei vasta Eesti standardile (165 cm).			85	ei
M	179	89	56	4		füüsika	Reaal	nii ja naa	nii ja naa				70	ei
M	191	80	60	4		kijandus	Hum	ei	nii ja naa	või			75	ei
M	188	78		4		nuusika ajal	Hum	ei	nii ja naa				57	ei
M	186	80	66	3		keemia	Reaal	jah	jah	H0: kursuse tütarlaste keskmine pikkus ei erine 165 sentimeetrist.			79	ei
M	190	85		3		keemia	Reaal	jah	nii ja naa	H1: kursuse tütarlaste keskmine pikkus erineb 165 sentimeetrist.				ei
M				5		ahaline kasv	Muu	jah	ei				54	ei
M	181	69	51	3		keemia	Reaal	jah	jah	või			67	ei
M	195	87	58	3		matemaatika	Reaal	jah	jah				70	ei
M	186	95	50	3		vene keel	Hum	jah	nii ja naa	H.: $\mu_1 = 165$ .			70	ei
M	183	80		3		matemaatika	Reaal	jah	nii ja naa	H.: $\mu_1 \neq 165$ .			42	ei
M	179,5	78	52	3		saksa keel	Hum	ei	ei				62	ei
M	183	85	55	3		keemia	Reaal	ei	nii ja naa				44	ei
M	177	71		3		keemia	Reaal	jah	jah	Funktsioon CONFIDENCE			50	ei
M	177	75	56	4		eesti keel	Hum	jah	nii ja naa	2,252496	Alumine usalduspiir	165,7058	60	ei
M	176	80	52	4		saksa keel	Hum	nii ja naa	ei		Ülemine usalduspiir	170,2108	75	jah
M	185	100	58	3		matemaatika	Reaal	ei	jah				93	jah
M	177	75		3		ajalugu	Hum	jah	jah	Descriptive Statistics			88	jah
M	186	76	55	3		keemia	Reaal	jah	jah	Column?	Alumine usalduspiir	165,6252	70	jah
M	178	68	57	4		vene keel	Hum	jah	jah		Ülemine usalduspiir	170,2914	70	jah
M	173	72	58	4		vene keel	Hum	nii ja naa	nii ja naa	Confidence Level(2,333109)			56	jah
M	186	95	57	3		matemaatika	Reaal	ei	jah				80	jah
M	197	85	50	3		matemaatika	Reaal	ei	jah				85	jah
N	165	59		4		füüsika	Reaal	jah	jah				69	jah
N	165	49		4		kijandus	Hum	jah	jah				87	jah
N	171	70	56	3		ajalugu	Hum	ei	jah				95	jah
N	172,5	60		4		füüsika	Reaal	jah	jah				80	jah
N	167	75		4				ei	jah				71	jah
N	175	64	50	3		inglise keel	Hum	jah	jah				75	jah
N	169	72	55	3		eesti keel	Hum	jah	jah				75	jah
N	157	57		5		geograafia	Hum	ei	jah				75	jah
N	164	50		4		matemaatika	Reaal	jah	jah				76	jah
N	177	79	53	3		füüsika	Reaal	ei	nii ja naa				68	jah
N	162		54	3		ihiskonnaõp	Hum	ei	nii ja naa				59	jah
N	168	55	50	3		füüsika	Reaal	jah	jah				49	jah
N	175	70	56	5		ajalugu	Hum	jah	jah				60	jah
N	169	55		3		matemaatika	Reaal	jah	ei				64	jah
N	171	54	53,2	3		keemia	Reaal	ei	jah				72	jah
N	178	80	56	5		keemia	Reaal	nii ja naa	jah				50	jah
N	157	48		4		füüsika	Reaal	jah	jah				55	jah
N	180	67	61	5		soome keel	Hum	ei	jah				70	jah
N	173	70		4		matemaatika	Reaal	ei	jah				55	jah
N	175	70		3		matemaatika	Reaal	ei	jah				48	jah
N	158	42		4		keemia	Reaal	ei	jah				58	jah
N	164	58		4		bioloogia	Hum	jah	jah				46	jah
N	165	46		3		eesti keel	Hum	jah	nii ja naa				59	jah
N	177	62	56	3		ahaline kasv	Muu	ei	jah				66	jah
N	165	59	56	3		vene keel	Hum	jah	nii ja naa				85	jah
N	167	66		4		matemaatika	Reaal	jah	jah				50	jah
N	177	85		4		keemia	Reaal	jah	jah				51	jah
N	161	44	56	3		ajalugu	Hum	ei	jah				62,5	jah
N	159	50		4		füüsika	Reaal	jah	jah				65	jah
N	175	51		3		geograafia	Hum	jah	nii ja naa				60	jah
N	158	50	55	4		vene keel	Hum	ei	jah				73	jah
N	167	62,5	55	3		füüsika	Reaal	jah	jah				73	nii ja naa
N	170	60	55	3		füüsika	Reaal	ei	jah				89	nii ja naa
N	153	65	57	4		ajalugu	Hum	jah	jah				80	nii ja naa
N	170	60	58	3		keemia	Reaal	jah	jah				72	nii ja naa
N	170	73	57,5	4		saksa keel	Hum	jah	ei				80	nii ja naa

Kopeeri / Copy -> Kleebi / Paste  
+  
Andmed / Data -> Sordi ... / Sort ...

2. Leidke nii mannaputru söövate kui ka mittesöövate tudengite arv, keskmised kehamassid ja kehamasside standardhälbed ('nii ja naa' vastanud jätke kõrvale).

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

MASS	PUDER	Mannapuder		
80	ei			
80	ei	Tudengite arv	21	38
78	ei	Keskmine kehamass	70,6	67,6
78	ei	Standardhälve	15,6	13,1
85	ei			
100	ei			
95	ei			
85	ei			
70	ei			
75	ei			
57	ei			
79	ei			
	ei			
54	ei			
67	ei			
70	ei			
70	ei			
42	ei			
62	ei			
44	ei			
50	ei			
60	ei			
75	jah			
93	jah			
88	jah			
70	jah			
70	jah			
56	jah			
80	jah			
85	jah			
	jah			
69	jah			
87	jah			
95	jah			
80	jah			
71	jah			
75	jah			
75	jah			
76	jah			
68	jah			
59	jah			
49	jah			
60	jah			
64	jah			
72	jah			
50	jah			
55	jah			
70	jah			
55	jah			
48	jah			
58	jah			
46	jah			
59	jah			
66	jah			
85	jah			
50	jah			
51	jah			
62,5	jah			
65	jah			
60	jah			
73	jah			
73	nii ja naa			
89	nii ja naa			
80	nii ja naa			
72	nii ja naa			
80	nii ja naa			

		PUDER		
		ei	jah	Grand Total
Count of MASS		21	38	59
Average of MASS2		70,52380952	67,64473684	68,66949153
StdDev of MASS3		15,59044274	13,11431439	13,98077811

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

<b>t-test</b>	
$H_0$ : Mannaputru söövate ja mittedöövate tudengite keskmised kehamassid on võrdsed.	
$H_1$ : Mannaputru söövate ja mittedöövate tudengite keskmised kehamassid on erinevad.	
või	
$H_0$ : $\mu_{ei} = \mu_{jah}$	$\mu_{ei}$ - mannaputru mittedöövate tudengite keskmine kehamass
$H_1$ : $\mu_{ei} \neq \mu_{jah}$	$\mu_{jah}$ - mannaputru söövate tudengite keskmine kehamass

### 4. Millist t-testi nende keskmiste võrdlemisel kasutada?

- 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 erinevatesse gruppidesse kuuluvate tudengite kehamasside varieeruvuse võrdumise või mittevõrdumise üle (funktsioon FTEST).

NB! Protseduur *F-test* (Tools / Tööriistad -> 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>	
$H_0$ : Mannaputru söövate ja mittedöövate tudengite keskmised kehamassid on võrdsed.	
$H_1$ : Mannaputru söövate ja mittedöövate tudengite keskmised kehamassid on erinevad.	
või	
$H_0$ : $\mu_{ei} = \mu_{jah}$	$\mu_{ei}$ - mannaputru mittedöövate tudengite keskmine kehamass
$H_1$ : $\mu_{ei} \neq \mu_{jah}$	$\mu_{jah}$ - mannaputru söövate tudengite keskmine kehamass
<u>Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.</u>	
<b>F-test (võrdleme dispersioone)</b>	
$H_0$ : $\sigma_{ei}^2 = \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mittedöövate tudengite gruppides ei ole erinev)
$H_1$ : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mittedöövate tudengite gruppides on erinev)



**=FTEST(R2:R23;R24:R62)**

R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
MASS	PUDER			Mannapuder								
80	ei			Ei	Jah							
80	ei		Tudengite arv	21	38							
78	ei		Keskmine kehamass	70,5	67,6							
78	ei		Standardhälve	15,6	13,1							
85	ei											
100	ei											
95	ei											
85	ei											
70	ei											
75	ei											
57	ei											
79	ei											
54	ei											
67	ei		t-test									
70	ei		H <sub>0</sub> : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on võrdsed.									
70	ei		H <sub>1</sub> : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on erinevad.									
42	ei		või									
62	ei		H <sub>0</sub> : $\mu_{ei} = \mu_{jah}$	$\mu_{ei}$ - mannaputru mitesöövate tudengite keskmine kehamass								
44	ei		H <sub>1</sub> : $\mu_{ei} \neq \mu_{jah}$	$\mu_{jah}$ - mannaputru söövate tudengite keskmine kehamass								
50	ei											
60	ei		Tegu on sõitumate vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.									
75	jah											
93	jah											
88	jah		F-test (võrdleme dispersioone)									
70	jah		H <sub>0</sub> : $\sigma_{ei}^2 = \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mitesöövate tudengite gruppides ei ole erinev)								
70	jah		H <sub>1</sub> : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mitesöövate tudengite gruppides on erinev)								
56	jah											
80	jah		funktsioon FTEST	=FTEST(R2:R23;R24:R62)								

**Function Arguments**

FTEST

**Array1** R2:R23 = {80;80;78;78;85;101}

**Array2** R24:R62 = {75;93;88;70;70;56}

= 0,354632166

Returns the result of an F-test, the one-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 = 0,354632166

[Help on this function](#) OK Cancel

**6. Sõnastage järeldus F-testist, põhjendage.**

<b>F-test (võrdleme dispersioone)</b>	
H <sub>0</sub> : $\sigma_{ei}^2 = \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mitesöövate tudengite gruppides ei ole erinev)
H <sub>1</sub> : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(kehamasside varieeruvus mannaputru söövate ja mitesöövate tudengite gruppides on erinev)
funktsioon FTEST	0,354632166 = $p > 0,05 \Rightarrow H_0$ : 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 TTEST:

**=TTEST(R2:R23;R24:R62;2;2)**

R	S	T	U	V	W	X	Y	Z	AA	AB
	MASS	PUDER		Mannapuder						
				Ei	Jah					
	80	ei								
	80	ei	Tudengite arv	21	38					
	78	ei	Keskmine kehamass	70,5	67,6					
	78	ei	Standardhälve	15,6	13,1					
	85	ei								
	100	ei								
	95	ei								
	85	ei								
	70	ei								
	75	ei								
	57	ei								
	79	ei								
		ei								
	54	ei								
	67	ei	t-test							
	70	ei	H <sub>0</sub> : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on võrdsed.							
	70	ei	H <sub>1</sub> : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on erinevad.							
	42	ei	või							
	62	ei	H <sub>0</sub> : $\mu_{ei} = \mu_{jah}$							
	44	ei	H <sub>1</sub> : $\mu_{ei} \neq \mu_{jah}$							
	50	ei								
	60	ei	Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist							
	75	jah								
	93	jah	F-test (võrdleme dispersioone)							
	88	jah	H <sub>0</sub> : $\sigma_{ei}^2 = \sigma_{jah}^2$ (kehamassi varieeruvus võrreldavates gruppides ei ole erinev)							
	70	jah	H <sub>1</sub> : $\sigma_{ei}^2 \neq \sigma_{jah}^2$ (kehamassi varieeruvus võrreldavates gruppides on erinev)							
	70	jah								
	56	jah	funktsioon FTES	0,35463217	= p > 0,05 => H <sub>0</sub> : kehamasside varieeruvus ei ole erinev					
	80	jah	Kasutada võib t-testi, mis eeldab kehamasside võrdset varieeruvust võrreldavates gruppides.							
	85	jah								
		jah	Võrdleme keskmisi							
	69	jah	funktsioon TTEST	=TTEST(R2:R23;R24:R62;2;2)						
	87	jah								
	95	jah								
	80	jah								
	71	jah								
	75	jah								
	75	jah								
	76	jah								
	68	jah								
	59	jah								
	49	jah								
	60	jah								
	64	jah								
	72	jah								
	50	jah								
	55	jah								
	70	jah								
	55	jah								
	48	jah								
	58	jah								
	46	jah								
	59	jah								
	66	jah								
	85	jah								
	50	jah								
	51	jah								
	62,5	jah								
	65	jah								
	60	jah								
	73	jah								

**Function Arguments**

TTEST

**Array1** R2:R23 = {80;80;78;78;85;100}

**Array2** R24:R62 = {75;93;88;70;70;56;69;87;95;80;71;75;75;76;68;59;49;60;64;72;50;55;70;55;48;58;46;59;66;85;50;51;62,5;65;60;73}

**Tails** 2 Testime kahepoolset hüpoteesi

**Type** 2 Et võime eeldada võrdset varieeruvust (miks?), siis on sobiv kasutada 2. tüüpi t-testi.

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 = 0,453635361

[Help on this function](#) OK Cancel

**b) vastava statistikaprotseduuri abil (Tools / Tööriistad -> Data Analysis... -> t-Test: ...):**

Sõltuvate gruppide (paariviisiline) võrdlus; funktsioonis TTEST tüüp nr 1

Sõltumatute gruppide võrdlus võrdsete dispersioonide eeldusel; funktsioonis TTEST tüüp nr 2

Sõltumatute gruppide võrdlus mittevõrdsete dispersioonide eeldusel; funktsioonis TTEST tüüp nr 3

MASS	PUUDER
80	ei
80	ei
78	ei
78	ei
85	ei
100	ei
95	ei
85	ei
70	ei
75	ei
57	ei
79	ei
	ei
54	ei
67	ei
70	ei
70	ei
42	ei
62	ei
44	ei
50	ei
60	ei
75	jah
93	jah
88	jah
70	jah
70	jah
56	jah
80	jah
85	jah
69	jah
87	jah
95	jah
80	jah
71	jah
75	jah
75	jah
78	jah
68	jah
59	jah
49	jah
60	jah
64	jah
72	jah
50	jah
55	jah
70	jah
55	jah
48	jah
58	jah
46	jah
59	jah
66	jah
85	jah
50	jah
51	jah
62,5	jah
85	jah
80	jah
73	jah

**t-test**  
 $H_0$ : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on võrdsed.  
 $H_1$ : Mannaputru söövate ja mitesöövate tudengite keskmised kehamassid on erinevad.  
 või  
 $H_0: \mu_{ei} = \mu_{jah}$   
 $H_1: \mu_{ei} \neq \mu_{jah}$

Tegu on sõltumatute vaatluste võrdlusega

**E-test**  
 $H_0: \sigma_{ei} = \sigma_{jah}$   
 $H_1: \sigma_{ei} \neq \sigma_{jah}$

funkt

Kasu

**Võrdleme keskmisi**  
 funktsioon TTEST  
 protseduur t-test

## 8. Sõnastage lõppjärelus, põhjendage.

funktsioon TTEST	0,453635361 = $p > 0,05 \Rightarrow H_0$ : mannaputru söövate ja mannaputru mitesöövate tudengite		
			keskmised kehamassid ei ole erinevad
protseduur t-test	t-Test: Two-Sample Assuming Equal Variances		
		<i>Ei</i>	<i>Jah</i>
Mean	70,52380952	67,64473684	
Variance	243,0619048	171,9852418	
Observations	21	38	
Pooled Variance	196,9244218		
Hypothesized Mean	0		
df	57		
t Stat	0,75453307		
P(T<=t) one-tail	0,22681768		
t Critical one-tail	1,672028889		
P(T<=t) two-tail	0,453635361 = $p > 0,05 \Rightarrow H_0$ : mannaputru söövate ja mannaputru mitesöövate		
t Critical two-tail	2,002465444		tudengite keskmised kehamassid ei ole erinevad

Veel võimalusi lõppjäreluse sõnastamiseks:

- „mannaputru söövate ja mitesöövate tudengite keskmiste kehamasside vaheline erinevus ei ole statistiliselt oluline ( $p > 0,05$ )“;  
so sihuke teaduslikumalt sõnastatud järelus;
- „tudengite kehamass ei sõltu sellest, kas nad söövad või ei söö mannaputru“ (vähe teise nurga alt sõnastatud järelus, aga kah õige).

Märkus. Kui nüüd tulemusi veelkord vaadata, siis ilmneb, et mannaputru söövad tudengid kaaluvad peaaegu kolm kilogrammi vähem võrreldes mannaputru mitesöövate tudengitega. Tõsi ta on. Ainult et tulemuse üldistamiseks (et näiteks „mannaputru dieediga“ turule tulla) on andmeid liiga vähe (või on varieeruvus liiga suur, et erinevust selgelt tuvastada; või on kolme-kilogrammiline erinevus lihtsalt liiga väike).

Seda, et tuvastatud erinevus eksisteerib vaid analüüsitud valimis ega pruugi esineda üldkogumis, näitabki olulisuse tõenäosuse  $p$  väärtus 0,45. Seega väites, et kõik mannaputru söövad tudengid (ka need, kelle kohta meil andmeid ei ole) kaaluvad keskmiselt vähem kui mannaputru mitesöövad tudengid, eksiksime me 45%-lise tõenäosusega.

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

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

	<i>Ei</i>	<i>Jah</i>
Mean	70,52380952	67,64473684
Variance	243,0619048	171,9852418
Observations	21	38

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

Pooled Variance	196,9244218	← Võrdsete dispersioonide eeldusel arvatud ühine dispersioon
Hypothesized Mean	0	
df	57	← <i>t</i> -statistiku empiiriline väärtus
t Stat	0,75453307	← 1-poolsele hüpoteesile vastav <i>p</i> -väärtus
P(T<=t) one-tail	0,22681768	
t Critical one-tail	1,672028889	← 1-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
P(T<=t) two-tail	0,453635361	← 2-poolsele hüpoteesile vastav <i>p</i> -väärtus
t Critical two-tail	2,002465444	← 2-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus

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.

Nimelt, kuna  $|t| = 0,75 < 2,00 = t_{critical}$ , siis ei ole põhjust nullhüpoteesi ümber lükata

(*t*-statistiku väärtus jääb sellesse piirkonda, kuhu ta nullhüpoteesi kehtides 95%-tõenäosusega peabki jääma).

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

Mida see ühepoolne hüpotees õieti väidab (kuidas on sõnastatud)?