# **Biometry practical 4**

# Illustrated (imperfect) practical guide

## **Preparatory work**

- 1. Open in MS Excel the questionary data (file analysed already in previous practical),
- 2. insert new worksheet, rename new worksheet to 'Praks4' (or 'Practical4') and
- 3. make a copy of the data table (from worksheet 'Andmed') and paste it into the upper left corner of the new worksheet.

# Exercise 1.

Suppose that the first year students of Institute of Veterinary Medicine and Animal Sciences (students in our dataset) are just a random sample from all first year students of Estonian University of Life Sciences. Knowing, that the average height of Estonian women is 169 cm, test the hypothesis: is the average height of first year female students of Estonian University of Life Sciences different from Estonian average 169 cm?

#### Guide

1. Sort the datatable by column 'GENDER'.



**2.** Find the number of girls and their average height and standard deviation of height using functions COUNT, AVERAGE and STDEV.S.

-	(= × ·	✓ f <sub>x</sub> =	STDEV.S	(B2:B45								
	Α	В	С	L	М	N	0 P	Р	Q	R	S	
1	GENDER	HEIGTH	WEIGHT	SKI	CAR	BEER	SMOKE		Neidude pikkus			
2	W	170	70	yes	yes	0	no		Vaatluste arv	44		
3	W	158	47,5	yes	yes	0	no		Keskmine	168,2		
4	W	170	60	yes	no	0	no		Standardhälve	=STDEV.S(	B2:B45)	
5	W	170	50	yes	no	0	no					
6	W	179	68	no	no	0	no					
7	W	163	56	yes	no	0	no					
8	W	177	65	yes	no	0	no					
9	W	162,5	53	no	no	0,25	no					
10	W	170	75	yes	yes	0	no					
11	W	176	66	yes	no	0,5	no		/			
12	W	161	50	yes	yes	0	no					
13	W	170	85	no	no	0	no					
14	W	176	58	no	no	0,5	no					
15	W	172	90	no	no	0	no		/			
16	W	158	55	yes	yes	0	yes		/			
17	W	169	60	yes	no	1	yes	_/	/			
18	W	164	52	yes	no	0	no					
19	W	172	62	no	no	0	no					
20	W	173	66	yes	no	0	no					
21	W	169	60	no	yes	0	no					
22	W	162	50	yes	yes	0	no					
23	W	165	52	yes		0	no					
24	W	171	63	yes	yes	0	no					
25	W	170	60	no	yes	0	no anymore, but I\'ve sn	nokei	1			
26	w	163	62	yes	yes	0	no anymore, but I\'ve sn	noke	3			
27	w	168	60	yes	yes	0	no				- 1	
28	w	1/4	54	yes	yes	0					- 1	
29	w	100	68	yes	yes	0	no anymore, but I ve sn	поке	3			
21	W	168	63	10	yes	2	yes no					
22	W	105	58	yes	10		10					
22	W	1/1	/5	yes	yes		10					
20	W	105		Noc	yes /		no anymore, but live co	noke	Girls' he	eight		
34	W	101	55	ves	ves	0	no anymore, but i ve sn	noke	1			
35	W	109	50	yes no	yes	0.5	00		Sample	size		4
30	W	1/5	90	Vec	10	0,5	00		Augran			160 000
38	W	107	70	no d	00	2	00		Average			108.23
30	W	100	- /U 		Vec	05	Vec		Standar	d devia	ation	6 0765
40	w	164	59	00	ves	0,5	no anymore, but l\'ve sn	noker	Standar	aucvie	actori	0.0700
41	w	194	90	~	ves	0	no anymore, but it ve sit	nokei				
42	w	103	50	ves	00	0	00					
44 43	w	160	70	100	ves	0	no anymore, but IV ve sr	noker	4			
44	w	162	70	00	00	2	no anymore, but i ve si	nokei	•			
45	w	172	58	00	00	0	no anymore, but IV ve sr	noke	4			
46	M	175	74	ves	no	1	no					
47	M	175	64	ves	ves	0.2	no anymore, but I\'ve sr	noker	4			

So, there are 44 girls with average height 168.2 cm and standard deviation 6.1 cm;

this says, that the average difference of girls' height from 168.2 cm is 6.1 cm; or, expecting that the height follows
normal distribution, then according to the properties of normal distribution

	*) approximately 68.3% of the heights of first year female students is in interval 168.2 $\pm$ 6.1 cm ( $\overline{x} \pm s$ ) and	Girls' height Sample size Average Standard deviation	44 168.2386 6.076536						
3.	*) approximately 95.5% of the heights of first year female students is in interval 168.2 $\pm$ 12.2 cm ( $\overline{x} + 2s$ ).	$H_0$ : the heights of fir $H_1$ : the heights of fir	st year fen st year fen	nale stude nale stude	nts corresp nts does n	oond to Est ot correspo	onian stan ond to Esto	dard (169 c nian stand	m) ard (169 cm)
	Formulate the hypothesis pair and write it down.	$H_0$ : the average heig $H_1$ : the average heig	ht of first y	year femal year femal	e students e students	does not o differs fro	differ from om 169 cm	169 cm	
	For example:	H <sub>0</sub> : μ <sub>F</sub> = 169 H <sub>1</sub> : μ <sub>F</sub> ≠ 170	μ <sub>F</sub> - avera	ge height o	f first year	female st	udents		

# Remainder from theory – relationships between hypothesis testing and confidence intervals

- If the task of hypothesis testing is to compare some estimated parameter with constant value, the decision is often made based on the confidence interval of the studied parameter:
  - if the constant is between confidence limits, then there is no reason to reject null hypothesis and it can be concluded that the studied parameter does not differ from given constant;
  - if constant is outside confidence interval, then the studied parameter is differ from given constant.
- For example, if you want to compare the average value with given constant (does the data correspond to standard), the hypothesis pair is:

--

$$H_0: \mu = c \text{ and } H_1: \mu \neq c.$$
  
if  $c \in [\underline{\mu}, \overline{\mu}]$ , then  $H_0: \mu = c$  is true; if  $c \notin [\underline{\mu}, \overline{\mu}]$ , then  $H_1: \mu \neq c$  is true.  
$$\underbrace{\mu}_{\underline{\mu}} \quad c \quad \overline{\mu}$$

**4.** Calculate the half of the 95% confidence interval using functions CONFIDENCE.NORM and CONFIDENCE.T:

#### a) function CONFIDENCE.NORM

(this function has 3 arguments:

significance level  $\alpha$ , standard deviation and number of female students in dataset);

Function CONFIDENCE.NOPM	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:
which function was used).	CHISQ.INV.RT
And after that put the cursor into result cell!	
	CORREL
	Returns the confidence interval for a population mean, using a normal distribution.
	Help on this function OK Cancel

СС	DUNT		$\checkmark f_x$	=CONFIDENCE.NORM(0.05;R4;R2)										
	Р	Q	R	S	т	U	V	w	x	Y	z			
1		Girls' height												
2		Sample size	44											
3		Average	168.2386					11 A						
4		Standard deviation	n 6.076536				FL	inction Arg	guments					
5				CONFR	DENCE.NORI	M								
6					Alpha 0.05									
7		H <sub>o</sub> : the heights of	first year fem		Stan	dard_dev R	4			= 6.0765357	29			
8		H <sub>1</sub> : the heights of	first year fem			Size R	2		55	= 44				
9							- \	<b>\</b>	1 Hour	4 705 4605				
10		H <sub>o</sub> : the average he	ight of first y	Returns	the confide	nce interval fo	or a populati	ion mean, usi	ing a normal	= 1.7954685 I distribution				
11		H₁: the average he	ight of first y		Alpha is the significance level used to compute the confid									
12		1 0				,	great	er than 0 and	less than 1	·	ie connu			
13		H <sub>o</sub> : μ <sub>ε</sub> = 169	u₌ - average				If we	want to cal	culate 95%	6 confidenc	e			
14		H.: II. ± 170		Formula	result = 1.	795468576	interval,	then the sig	gnificance	level $\alpha = 0$	.05.			
15		111. με <del>-</del> 170			i courte i i									
16				Help on	this functio	<u>n</u>				/	, ОК			
17		Function CONFIDE	NCF.NORM											
18			=CONFIDEN	ICE.NOR	/(0.05:R4:I	R2)								
_														
						Funct	ION CONF	IDENCE.N		7				
	1.795469													

## **b)** function CONFIDENCE.T

(arguments of this function are similar to the function CONFIDENCE.NORM).

Result:

Function CONFIDENCE.NORM							
	1.795469						
Function CONFIDENCE.T							
1.847436							

5. Calculate the lower and upper confidence limits based on both functions.

V2	21	▼ E 🗙	$\checkmark f_x$	=R3+R2	1						
	Ρ	Q	R	S	т	U	v	w			
1		Girls' height									
2		Sample size	44								
3		Average	168.2386								
4		Standard deviation	6.076536								
5											
6											
7		H <sub>o</sub> : the heights of f	irst year fer	nale stude	ents correspo	ond to Estoni	an standar	d (169 cm)			
8		H <sub>1</sub> : the heights of f	irst year fer	nale stude	ents does not	t correspond	to Estonia	n standard			
9											
10		H <sub>o</sub> : the average hei	ght of first	year fema	le students d	loes not diff	er from 169	9 cm			
11		H <sub>1</sub> : the average hei	ght of first	year fema	ale students differs from 169 cm						
12											
13		$H_0: \mu_F = 169$	μ <sub>F</sub> - averag	ge height o	of first year f	emale stude	nts				
14		H <sub>1</sub> : μ <sub>F</sub> ≠ 170									
15							$\langle \rangle$				
16											
17		Function CONFIDE	NCE.NORM		Lower confi	dence limit	166.4432	= R3-R18			
18			1.795469		Upper confi	dence limit	170.0341				
19								$\mathbf{h}$			
20		Function CONFIDER	NCE.T		Lower confidence limit 166.3912			= R3-R21			
21			1.84743 <u>6</u>		Upper confi	dence limit	170.0861				

#### Which of these 95% confidence intervals is wider? Why?

#### Answer.

The confidence interval got with function CONFIDENCE.T is slightly wider.

The reason is, that function CONFIDENCE.T calculates the confidence limits based on t-distribution following the formula  $\overline{x} \pm t_{1-\alpha/2,n-1} \frac{s}{\sqrt{n}}$ ,

but function CONFIDENCE.NORM calculates **asymptotic** (approximate) confidence interval (whereby the accuracy is increasing when the sample size is increasing) based on normal distribution:  $\overline{x} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}}$ , got range

is little under estimated in case of small data set.

Parameters  $t_{1-\alpha/2,n-1} = t_{0.975;43} = 2,017$  and  $z_{1-\alpha/2} = z_{0.975} = 1,96$  are 97.5%-points (values, from which the bigger values can occur only with probability 0.025) of t-distribution and standard normal distribution, respectively, the first of these quantities is calculable in Excel 2010 with function =T.INV(0,975;43) and the second with function =NORM.S.INV(0,975).

**NB!** In the older *Excel* versions there is no functions CONFIDENCE.T and CONFIDENCE.NORM. There is only function CONFIDENCE, which is equivalent to functions CONFIDENCE.NORM, to calculate the confidence interval based on t-distribution the corresponding option *Confidence Level for Mean* of procedure *Descriptive Statistics* can be used.

# 6. Make a final conclusion and write it down – does the average height of first year female students differ from Estonian average (169 cm)? The answer must contain also the argumentation, why you made this decision.

<u>Example.</u> As the average height of Estonian women (169 cm) is between 95% confidence limits of the first year female students average height:  $169 \in (166,4; 170,1)$ , then there is no reason to reject nullhypothesis H<sub>0</sub>: the average heights of first year female students does not differ from 169 cm.

#### 7. Supplementary task.

The average height of women over the world is 154 cm. Can you conclude that the average height of first year female students in Estonian University of Life Sciences differs from the world average?

NB! You don't need to make any additional calculations. The decision can be made just based on already calculated confidence interval.

# Exercise 2.

Are weights of students owning and not owning a car different?

#### Guide

1. Make an additional table containing only columns 'WEIGHT' and 'CAR' and sort it by column 'CAR'.

	в	C	D	E	F	G	н	1.1	1	K	L	M	N	0	P	Q	R	s	т U	v	w	x	Y Z
1	HEIGHT	WEIGHT	HEAD	SHOE_SIZ	MATH	BREAKFA:	PORRIDGE	PET	SICK	SPORT	SKI	CAR	BEER	SMOKE		Girls' height							WEIGHT CAR
2	170	70	35.5	39		3 other	yes	yes	no	yes	yes	yes		0 no		Sample size	44						70 yes
з	158	47.5	55	36		3 cereals or	yes	yes	no	yes	yes	yes		0 no		Average	168.239						47.5 yes
4	170	60	53	38		5 cereals or	yes	yes	no	yes	yes	no		0 no		Standard deviation	6.07654						60 no
5	170	50	55	37	4	4 sandwich	yes	yes	no	yes	yes	no		0 no									30 no
6	179	68	58	41		5 cereals or	yes	yes	no	yes	no	no		0 no									68 no
7	163	56		37	4	4 sandwich	ves	ves	no	no	ves	no		0 no		H <sub>a</sub> : the heights of f	rst year fem	ale students	correspond to Est	tonian stand	dand (169 on	n)	36 no
8	177	63	- 11	40		sandwich	sometime	ves	ves	ves	ves	-		0.00	_	H.: the beights of f	ist vear tem	ale students	does not correspo	and to Estar	vian standar	rd (169 m	63 00
	162.3			20		2 norridee	vec	vec	00	vec	00	<b>C</b>	0.7	3 00								- (	33.00
10	470			20		ather						<b>1</b>		0		U - the success he				dillar trans	100		75 110
	-//0	- 11				a conter	yes -	100			100	100		-		ne. the overage he	Burrow	car remarc s	codentis does not	uner nom.	205 011		75 765
11	1/6	66	2/	59		<ul> <li>sanowich</li> </ul>	sometime	no	no	yes	yes	<b>*</b>		5 no	_	H <sub>3</sub> : the average ne	gnt or first y	ear remaie s	tudents differs fro	m 169 cm			66 NO
12	161	50	22	57		+ notning	no	yes	yes	yes	yes	yes		u no	_								oo yes
13	170	85	57	41		4 cereals or	no	yes	no	no	no	no		0 no	_	H <sub>0</sub> : μ, = 169	μ, - average	height of firs	st year female stu	dents			85 no
14	176	58	52	39		5 cereals or	yes	yes	yes	no	no	no	0.	5 no		H <sub>2</sub> : µ, ≠ 170							38 no
15	172	90	36	41	-	4 porridge	yes	yes	no	no	no	no		0 no									90 no
16	158	- 33	57	38		4 cereals et	V. S.	yes	yes	no	yes	yes		0 yes									55 yes
17	169	60	35.5	41	4	4 cereals or	yes	yes	100	no	yes	no		1 yes		Function CONFIDEN	ICE.NORM	Lowe	er confidence limi	t 166.443	= R3-R18		60 no
18	164	52	56	37	4	4 other	sometime	no	no	yes	Yes	00		0 no			1.79547	Upp	er confidence limit	t 170.034			32 no
19	172	62	56	39		4 sandwich	yes	yes	no	no	no	no		0.00									62 no
20	173	66	56	40		5 cereals or	yes	yes	no	yes	yes	no		0 no		Function CONFIDEN	ICE.T	Lowe	er confidence limi	t 166.391	= R3-R21		66 no
21	169	60	55	39		3 other	yes	yes	no	yes	no	yes		0 no	_		1.84744	Upp	er confidence limit	t 170.086			60 yes
22	162	50	30	38		3 porridge	yes	yes	yes	no	yes	yes		0 no	_								30 yes
23	165	52	30.5	37		4 sandwich	sometime	yes	yes	yes	yes			0 no	_								32
24	171	63	57	39		o cereals or	yes	yes	yes	yes	yes	yes		0 no						_			63 yes
25	170	60	53	39		other	no	yes	yes	yes	no	yes		o ne enymore, e	at in ve	smoked							60 yes
26	165	64		58		o cereals or	yes	no	no	yes	yes	yes		u no anymore, o	ut IVve	smoked		·					62 yes
2/	100	50		33		<ul> <li>cereals or</li> <li>cereals or</li> </ul>	yes	yes	yes 	yes	yes	yes		0 no	_			Cot	ν -> F	Paste			eu yes
20	400	- 2				2 other	yes no	763			Yes	100				resolved			<i>y y y</i>	uste			24 yes
20	100	60	30	20		s ourier Crandwich	Ner I		Ner.	yes	yes 00	yes ver		o no anymore, o Siver	ut ive	shokeu			+				60 yes 67 yer
24	100		30	27		<ul> <li>senswich</li> </ul>	763	yes ver	yes m	yes	Ner.	100		1 m	_				1				59 pc
37	171	75		41		4 sandwich	vec	vec	00	Vec	vec	ver.		0.00					Sort				75 ves
33	165	77	38	29		sandwich	Ves	ves	00	00	00	Nes.		0 no					5011				77 yes
34	161	22	37	38		3 porridee	ves	ves	ves	ves	ves	ves		0 no anymore, b	ut I\'ve	smoked							33 yes
35	169	53	33	38		3 sandwich	sometime	ves	00	ves	ves	ves		0 no									53 yes
36	175	60	57	42		5 cereals or	ves	ves	no	no	no	no	0	5 no	_								60 no
37	167	80	57.5	41		5 other	yes	yes	no	yes	yes	yes		2 no									80 yes
38	158	70	55	38		5 cereals or	yes	yes	yes	yes	no	no		0 no									70 no
39	165	61	57	39		3 other	sometime	yes	yes	no	no	yes	0.	5 yes									61 yes
40	164	58	57	39		8 sandwich	yes	yes	yes	yes	no	yes		0 no anymore, b	ut I\'ve	smoked							58 yes
41	185	80	60	41	4	4 cereals or	sometime	yes	no	yes	no	yes		0 no									80 yes
42	177	63	60	40		2 sandwich	no	no	no	yes	yes	no		0 no									63 no
43	160	70	57	39	4	4 sandwich	sometime	yes	yes	yes	no	yes		0 no anymore, b	ut I\'ve	smoked							70 yes
44	162	70	55	40		5 sendwich	no	yes	no	no	no	no		2 no									70 no
45	172	58	62	39	4	4 other	sometime	yes	no	yes	no	no		0 no anymore, b	ut I\'ve	smoked							58 no
46	175	74	57	42		3 sandwich	yes	yes	no	yes	yes	no		1 no									74 no
47	175	64	56	42	4	4 other	yes	yes	no	yes	yes	yes	0.	2 no anymore, b	ut I\'ve	smoked							64 yes
48	190	82	58	46	4	4 other	yes	yes	no	yes	yes	yes		3 no									82 yes
49	189	82		43	4	4 cereals or	no	yes	no	yes	yes	yes	2.	5 yes	_								82 yes
50	170	80	56	41	4	4 cereals or	no	yes	no	yes	no	yes		0 no	_								80 yes
51	176	74	56	42		o porridge	yes	yes	yes	yes	yes	yes	0.	1 yes	_					_			74 yes
52	175	73	34	43		other	sometime	yes	yes	yes	no	no	0.	1 yes	_								73 no
53	181	74	35	44		+ senowich	yes	yes	nd	yes	yes	yes		1 10	-								74 yes
54	183	73		43		s pornioge	yes	no	nd	yes	yes	yes		s nd									75 yes
55	174	87	37	40	-	+ senowich	sometime	yes	yes	yes	no	yes	• •	5 nd	_								87 yes

2. Calculate the average and standard deviation of weights depending on the owning of car (NB! Omit the student, who does not know has she or he a car or not).

You can use corresponding <u>functions</u> or <u>PivotTable</u>. If you wish, you can try both variants.

			<b>`</b>					
WEIGHT	CAR				CAR			
60	no		$\mathbf{X}$	No		Yes		
50	no	No of students		1	22		31	$\mathbf{i}$
68		Average		64 4090	9091	66 59677	119	)
56		Standard deviation		9 94584	9055	11 108869	985	
65	00							
53						/		
	10		~ .					
00	no		Colu	imn Label	5 -			
85	no	Values	no			yes		Grand Total
58	no	Count of WEIGHT			22		31	53
90	no	Average of WEIGHT2		64.4090	9091	66.596774	419	65.68867925
60	no	StdDev of WEIGHT3		9.94584	9055	11.108869	985	10.59854236
52	no							
62	no							
66	00						_	
58	00							
60	00							
70	no							
63	no							
70	no							
58	no						_	
74	no							
73	no						_	
70	ves							
47.5	ves							
75	ves							
50	ves							
55	ves							
60	ves							
50	yes							
63	ves							
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	ves							

**3.** Formulate the hypothesis pair and write it down.

t-test										
H <sub>0</sub> : The <b>average weigh</b>	<b>its</b> of students owning a	and not owning a	car are not different							
H1: The average weights of students owning and not owning a car are different										
or										
$H_0: \mu_{No} = \mu_{Yes}$	$H_0: \mu_{No} = \mu_{Yes}$ $\mu_{No}$ - average weights of students not owning a car									
$H_1: \mu_{No} \neq \mu_{Yes}$ $\mu_{Yes}$ - average weights of students owning a car										

#### 4. Which t-test to use?

NB! There are three types of t-tests, look at page 12 (step 7b).

- As compared **groups are independent** (there are different students in groups), before the comparison of means the **variances must be compared** to decide, which t-test to use (this, which assumes equal variances, or this, which assumes unequal variances).
- To compare variances the F-test can be used.
- **5.** To decide is the weights' variability of students with and without car equal or not, formulate the **hypothesis pair for variances comparison** and **perform F-test** (**function F.TEST**).

NB! There is also statistical procedure *F*-test (*Data*-tab -> *Data Analysis...* -> *F*-Test Two-Sample for Variances), but this tests only one side hypothesis and can't be directly applied to decide about equality of variances.

t-test														
H <sub>0</sub> : The <b>average weigh</b>	<sub>0</sub> : The <b>average weights</b> of students owning and not owning a car <b>are not different</b>													
H <sub>1</sub> : The <b>average weig</b>	1: The average weights of students owning and not owning a car are different													
or	or													
H <sub>0</sub> : μ <sub>No</sub> =μ <sub>Yes</sub>	$\mu_{No}$ - average weights	s of students not o	owning a car											
H <sub>1</sub> : μ <sub>No</sub> ≠μ <sub>Yes</sub>	μ <sub>Yes</sub> - average weights of students owning a car													
The groups are indepe	ndent. Before compari	son of means the	variances must b	e compared to de	cide, which t-test to u	se.								
	F-test (comparison	of variances)												
	$H_0: \sigma^2_{No} = \sigma^2_{Yes}$	(the weights' var	riability of studen	ts owning and no	t owning a car <b>is not c</b>	lifferent)								
	$H_1: \sigma^2_{No} \neq \sigma^2_{Yes}$ (the weights' <b>variability</b> of students owning and not owning a car <b>is different</b> )													



#### 6. Write down justified conclusion based on F-test.

						This is the just	stification.	
F-test (compari	son of vari	ances)	-			Do you und	erstood?	
H <sub>0</sub> : σ <sup>2</sup> <sub>Ei</sub> =σ <sup>2</sup> <sub>Jah</sub> H <sub>1</sub> : σ <sup>2</sup> <sub>Ei</sub> ≠σ <sup>2</sup> <sub>Jah</sub>	(the weig (the weig	Decision 1 and decis	rule ion	x <sup>4</sup> -students with and withou	t car is n onclusi	on ent)		
function F.TEST	0.60556	= p > 0,05 =	> H <sub>0</sub>	10: the weights' variability in	compar	red groups is n	ot different	'
The t-test assur	ning equal	variances m	ust	t be used.				

# 7. Perform the t-test to compare average weights.

Using both

## **a)** function **T.TEST**:

-	(	X 🗸 🕽	🕯 =T.1	EST(Y2	Y23;Y24:Y54;2;	2)						
	X	Y	Z	AA	AB	AC		AD	AE	AF	AG	AH
1		WEIGHT	CAR			C	AR	-				
2		60	no			No		Yes				
3		50	no		No of students		22	31				
4		68	no		Average	64.409090	091	66.5967742				
5		56	no		Standard deviation	9.9458490	055	11.1088698				
6		65	no									
7		53	no									
8		66	no			Column Labels	<b>T</b> ,					
9		85	no		Values	no		yes	Grand Total			
10		58	no		Count of WEIGHT		22	31		53		
11		90	no		Average of WEIGHT	64.409090	091	66.5967742	65.688679	25		
12		60	no		StdDev of WEIGHT3	9.9458490	055	11 1088698	10.598542	36		
13		52	00									
14		62	00									
15		66	00		4 4 4							
16		58	00		t-test	able of students a		les and not		and not different		
17		03	00		no. The average weight	gitts of students of	owr	ning and not o	owning a ca	are not different		
19		70			H1: The average we	gnts of students	ow	ning and not	owning a ca	r are different		
19		63	no	1	or							
20		70	10		H <sub>0</sub> : μ <sub>No</sub> =μ <sub>Yes</sub>	μ <sub>No</sub> - average we	eigi	nts of studen	ts not ownin	ig a car		
20		10	no		H <sub>1</sub> : μ <sub>No</sub> #μ <sub>Yes</sub>	µ <sub>ves</sub> - average w	eig	hts of studen	ts owning a	car		
21		58	no									
22		74	no		The groups are inde	ependent. Before	e co	mparison of	means the v	ariances must be	compared to dec	ide, which t-test t
23		73	NO				_					
24		/0	yes			F-test (compari	ISO	n of variances	s) Lucalatilitas a	f an ideata anala	a and a statustan	a care la not differen
25		47,5	yes			Ho: 0"No=0"Yes		(the weights	variability o	f students ownin	g and not owning	a car is not different
26		15	yes			H <sub>1</sub> : 0 <sup>™</sup> N0 <sup>#</sup> 0 <sup>™</sup> Yes		(the weights	variability o	r students ownin	g and not owning	a car is different)
27		50	yes	$\mathbf{i}$		4		0.00000		11 other well-base		
28		55	yes			function F.TEST		0.60556	= p > 0,05 =	> H <sub>o</sub> : the weights	variability in com	npared groups is r
29		60	yes			The state of the st				- 4		
30		50	yes			The t-test assum	ing	equal variance	s must be us	ed.		
31		63	yes		Comparison of mo	20.0						
32		60	yes		comparison of me	ans						
24		62	yes		function TYPEST	-T TEST(V2-V23-V	124-	V54-2-21				
25		54	yes		Indication 1. Los	-1.1001(12.120,1		(34,2,2)				
36		68	yes	_			_					
37		63	yes	Eunet	an Armonte							2 🔽
38		75	ues	runct	ion Arguments		_					
39		77	ues	T TES		$\mathbf{i}$						
40		55	ves	1						_		
41		53	ves		Ar	ray1 Y2:Y23				(60;50)	68;56;65;53;66;8	5;58;90;60;52
42		80	yes			1004.UT	_		- C	- (70.47)	E.7E.E0.FE.40.FC	62,60,62
43		61	yes		Ar	1492 Y24:Y54				= {/0;4/,	5,75;50;55;60;50	,03;00;02;
44		58	yes			Tails 2			1	o test two-s	ide hypothe	sis
45		80	yes								<u>_</u>	
46		70	yes		0	Type 2	_		1	ype of t-tes	t assuming e	equal ¦
47		64	yes						L.	ariability in	compared g	roups
48		82	yes								population B	
49		82	yes	Returns	s the probability ass	ociated with a S	tud	ent's t-Test.				
50		80	yes			Type	is F	he kind of t-t	est: paired	= 1. two-sample	equal variance (h	omoscedastic) =
51		74	yes			1700	2.1	two-sample u	inegual vari	ance = 3.	equal variance (ii	
52		74	yes				-					
53		75	yes									
54		87	yes									
55		52		Formula	a result = 0,46438	6809						
56												
14 4	•	M Z Ar	dmed	I Help on	this function						OK	Cancel

**b**) corresponding statistical procedure (*Data*-tab -> *Data Analysis*... -> *t*-*Test*: ...):

	Data Analysis	
	<u>A</u> nalysis Tools	ОК
Comparison of <b>dependent groups</b> (pair-wise comparison); type 1 in function T.TEST Comparison of <b>independent groups</b> assuming <b>equal variances</b> ; type 2 in function T.TEST Comparison of <b>independent groups</b> assuming <b>unequal variances</b> ;	Histogram Moving Average Random Number Generation Rank and Percentile Regression Sampling t-Test: Paired Two Sample for Means t-Test: Two-Sample Assuming Equal Variances t-Test: Two-Sample Assuming Unequal Variances z-Test: Two Sample for Means	
type 5 in function 1.1ES1		

	X		7	۵۵	AB	AC		AD	ΔF	AF
- 1		VEICHT	CAP	~~~	- P	- AV	CAR	~D	AL	Ar
2		60 E0	CMN			Ne	CAR	Vee		
2		50	no 		No. of environments	NO	2.2	Tes		
3		50	no		No of students	64.4000	22	51		
4		50	no		Average	64.40905	9091	66.5967742		
5		56	no		Standard deviation	9.945849	9055	11.1088698		
Б		65	no							
7		53	no							
8		66	no			Column Labels	- <b>T</b>			
9		85	00		Values	00		Ves	Grand Total	
10		58	00		Course of MERCUIT	110		705	52	
11		90			Count of WEIGHT		22	51	55	
10			110		Average of WEIGHT2	64.40909	9091	66.5967742	65.68867925	
12		60	no		StdDev of WEIGHT3	9.945849	9055	11.1088698	10.59854236	
13		52	no							
14		62	no							
15		66	no							
16		58	 DO		t-test					
17		60			H <sub>0</sub> : The average weigh	hts of students ov	vning	g and not owni	ng a car are not d	lifferent
10			no		H <sub>1</sub> : The average weig	hts of students o	wnin	g and not own	ing a car are diffe	rent
18		/0	no		or					
19		63	no		H <sub>0</sub> : µ <sub>N0</sub> =µ <sub>Yes</sub>	μ <sub>No</sub> - average w	eight	s of students r	not owning a car	
20		70	no		Harman		aisht	s of students a	woine a car	
21		58	no		11- HNOFHYES	Pro - average W	eight	s or scudents o	wing a car	
22		74	no		-					
23		73	00		The groups are indep	endent. Before c	omp	arison of mear	ns the variances	must be compa
24		70								
24		47.5	yes			F-test (compar	ison	of variances)		
20		47,5	yes			H <sub>0</sub> : σ <sup>2</sup> <sub>No</sub> =σ <sup>2</sup> Yes		(the weights'	variability of stud	dents owning ar
26		75	yes			H <sub>1</sub> : σ <sup>2</sup> <sub>No</sub> ≠σ <sup>2</sup> <sub>Ves</sub>		(the weights'	variability of stud	dents owning ar
27		50	ves			1 10 10				
28		55	yes	$\mathbf{h}$		function E TEST		0.60556	- n > 0.05 -> H.:	the weights'va
29		60	yes			Turrettori T. Teor		0.00350	- p = 0,05 -= 110.	the weights va
30		50	ves							
31		63	ves			The t-test assum	ninge	equal variances	s must be used.	
32		00	105							
33		62	ues		Comparison of mean	ns				
24		60	yes							
34		50	yes		function T.TEST	0.4643868				
35		54	yes							
36		68	yes		procedure t-test			1		
37		63	yes				7			
38		75	yes			· · · · · · · · · · · · · · · · · · ·		1		
39		77	yes	t-Ter	st: Two-Sample	Assuming E	ma	Variance	s	2 🔽
40		55	yes		a ino sample	Associating L	qua	anance		
41		53	yes	Inpu	ut					
42		80	yes	Vari	iable 1 Papper	Autou	tyto	0		ОК
43		61	ues	Vari	abio <u>i</u> Kange:	\$7\$2:5	pr\$2	.s  [	× –	H
44		58	ues	Vari	iable 2 Rappe:	47404	.4v#	54 0		ancel
45		90	ves		inter Enteringen	\$1\$24	• • • •			H
40		70	yes							Help
40		70	yes	Hyp	oth <u>e</u> sized Mean Diff	erence:				[]
47		64	yes							
48		82	yes		Labels					
49		82	yes	Alet	0.05			/		
50		80	yes	Hibi	ia: 0,05			/		
51		74	yes					/		
52		74	yes	Out	put options		/	/		
53		75	ves		Output Range:	\$AC\$3	17	6	<b>S</b>	
54		87	ues		<u>s</u> aspacitionger	4.1.242		e		-
54			yes		Now Workshoot Dlu					

Comparison of m	eans								
function T.TEST	0,4643868	= p > 0,05	=> H <sub>0</sub> : the a	verage weights of students wi	th and with	out car <mark>are</mark>	not differe	ent	
procedure t-test	t-Test: Two-San	nple Assun	ning Equal V	ariances					
		Variable 1	Variable 2						
	Mean	64,40909	66,596774						
	Variance	98,91991	123,40699						
	Observations	22	31						
	Pooled Varianc	113,3241							
	Hypothesized N	0							
	df	51							
	t Stat	-0,73719							
	P(T<=t) one-tai	0,232193							
	t Critical one-ta	1,675285							
	P(T<=t) two-tai	0,464387	= p > 0,05 =	> H <sub>0</sub> : the average weights of st	tudents wit	h and with	out car <mark>are</mark>	not differ	ent
	t Critical two-ta	2,007584							

#### 8. Write down the final conclusion, justify it(!).

Other possibilities for the final conclusion:

"the average weights of students owning and not owning a car are not <u>statistically</u> <u>significantly</u> different (p > 0,05)",

this is a little more scientific and accurate conclusion;

"the students' weight does not depend on owning of a car" (this is little differently phrased conclusion but also correct).

<u>Remark</u>. Significance probability (p-value) p = 0.464 is showing that

- the probability to make a mistake concluding that the average weights are different is 46.4%;
- <u>assuming, that in population there is no difference</u> between compared groups, the <u>probability to get the</u> <u>sample with observed difference just by chance</u> is 0.464.

As from one hand the probability to make a wrong conclusion that the groups are different is too big (formally the traditional limit is 0.05) and from other hand the probability to get observed difference just by chance is also big (once again, the formal limit is 0.05), there is no reason to reject null hypothesis about equality of average weights.

- **9.** But what are showing the other quantities calculated by procedure *t-Test: Two-Sample Assuming Equal Variances*?
  - The first part of output contains basic descriptive statistics of compared groups (mean, variance and number of observations; NB! Be careful and don't mix up the compared groups;

you may want to write instead abstract names 'Variable 1' and 'Variable 2' the real group names):

	Car='No'	Car='Yes'
	Variable 1	Variable 2
Mean	64,40909091	66,59677419
Variance	98,91991342	123,4069892
Observations	22	31

• The other output parts are already related with hypothesis testing:

Pooled Variance	113,3240757	Common variance calculated assuming the equal	
Hypothesized Mean D	0	variances in compared groups	
df	51	Empirical (sample based) value of t-statistic	
t Stat	-0,737186899	- p-value corresponding to one-side hypothesis	
P(T<=t) one-tail	0,232193405		
t Critical one-tail	1,67528495	5 Critical value of t-statistic corresponding to one-side	
P(T<=t) two-tail	0,464386809	p-value corresponding to two-side hypothesis	
t Critical two-tail	2,00758377	Critical value of t-statistic corresponding to two-side h	

Two-side (or two-tailed) hypothesis means testing "non-equal against equal":

# $H_0: \mu_{No} = \mu_{Yes}, \\ H_1: \mu_{No} \neq \mu_{Yes}.$

**One-side (or one-tailed) hypothesis** means testing "less than *against* more or equal than" (or "more than *against* less or equal than"; in which direction to test depends on data, for example Excel performs one-side t-test testing always is bigger average also statistically significantly bigger):

 $\begin{array}{l} H_0: \mu_{No} \geq \mu_{Yes}, \\ H_1: \mu_{No} < \mu_{Yes}. \end{array}$ 

Instead of p-value the final conclusion can be made also based on **comparing the empirical value of teststatistic** |t| **with its critical value**  $t_{critical}$ . If the empirical value is smaller than critical value,  $|t|=1,68 < 2,01 = t_{critical}$ , then there is no reason to reject the null hypothesis (the empirical value of teststatistic is in interval, where it should be with 95% probability if the null hypothesis is true – so, our data does not allow to reject null hypothesis).

#### 10. Can you reject the null hypothesis of one-side t-test?

How about the conclusion of this one-side test (how to phrase it)?