## Praks 2

## Eel- ja järeltöö

- 1. Salvestage arvutisse andmestik lammas.xls (http://ph.emu.ee/~ktanel/VL\_1112/lammas.xls).
- 2. Avage salvestatud fail MS Excel'is.

•••

...

3. Peale ülesannete lahendamist salvestage fail nimega '*perekonnanimi*\_lammas.xls' ja saatke e-meiliga aadressil tanel.kaart@emu.ee.

## Kommentaarid andmestiku kohta

Rümpade EUROP klassifitseerimine on EL riikides kehtiv lihakehade klassifitseerimissüsteem, kus hinnatakse iga lihakeha kommertsväärtust.

Rümbad jaotatakse järgmistesse kategooriatesse:

- alla 12 kuu vanuste lammaste e tallede rümbad (tähis "L") ning kõigi ülejäänud lammaste rümbad (tähis "S");
- visuaalselt hinnatatud lihakusklasside osas eristatakse: "E" (ekstra), "U" (väga hea), "R" (hea), "O" (rahuldav), "P" (lahja), "P-" (eriti lahja);
- visuaalselt hinnatud rasvasusklasside osas eristatakse: "1" (väherasvane), "2" (kergelt rasvane), "3" (keskmiselt rasvane), "4" (rasvane), "5" (väga rasvane).

Antud andmestik sisaldab 686 lambarümba andmeid (56-lt omanikult) 2002. aasta sügisest (tapetud ja hinnatud kõik samas tapamajas).

Iga lamba kohta on lisaks eelnevalt nimetatud kolmele rühmitavale tunnusele:

'rümba kategooria' $\in$  {L, S}, 'lihakusklass' $\in$  {E, U, R, O, P, P-} ja 'rasvasusklass' $\in$  {1, 2, 3, 4, 5}, fikseeritud ka see,

kas loom oli pärit jõudluskontrollialusest karjast või mitte (vastavalt 'jkk' = 1 või 0), kas realiseerimine leidis aset läbi ELaS-i turustusgrupi või mitte (vastavalt 'realis' = 1 või 0), samuti on teada rümba mass (kg) ja hind (EEK), mille alusel on arvutatud rümba 1 kg hind (EEK/kg).

# Ülesanded

Praktikumi tehniline pool hõlmab peamiselt *Pivot Table*'i ja diagrammide kasutamist *MS Excel*'is, lisaks ka veel  $\chi^2$ -testi ja regressioonanalüüsi.

- 1. Kirjeldage lammaste jagunemist EUROP klassifitseerimissüsteemi alusel, leides erinevatesse klassidesse kuuluvate rümpade arvud ja protsendid (seda siis 3 tunnuse tarvis rümba üldkategooria, lihakusklass ja rasvasusklass).
  - Kui mõnda lihakus- ja/või rasvasusklassi kategooriat esineb väga vähe, pange see kokku sarnase naaberkategooriaga.
  - Illustreerige saadud tabeleid sektordiagrammidega, kirjutades igale sektorile juurde sellele vastava väärtuse ja esinemise suhtelise sageduse protsentides.
- 2. Kas rümpade jagunemine rasvasusklassidesse sõltub rümba üldkategooriast?
  - Võimaliku seose kirjeldamiseks konstrueerige (uuele töölehele) vastav 2-mõõtmeline sagedustabel, viimasesse leidke nii rea- kui ka veeruprotsendid ja sõnastage lause(d), kasutades vähemalt kahte leitud suhtelistest sagedustest.
- **3.** Jätkuna punktile 2 testige rümpade üldkategooriatesse ja rasvasuklassidesse jagunemise vahelise seose statistilist olulisust.
  - Et oleks selge, mida te üldse testite, pange esmalt kirja kontrollitav hüpoteeside paar.
  - Järgnevalt konstrueerige uus 2-mõõtmeline sagedustabel, mis sisaldab üksnes absoluutseid sagedusi, selle alusel arvutage tunnuste sõltumatuse juhule (nullhüpoteesile) vastavad sagedused ja
  - o teostage funktsiooni CHISQ.TEST (*Excel* 2003-s CHITEST) abil  $\chi^2$ -test viimane võrdleb empiirilisi (andmetabelist arvutatud) sagedusi teoreetiliste (sõltumatuse juhule vastavate) sagedustega ja väljastab olulisuse tõenäosuse *p* väärtuse.
  - Sõnastage lõppjäreldus (viidates sõnastuses ka *p*-väärtusele, millel järeldus baseerub).
- **4.** Prognoosige tallerümpade 1 kg hinda lähtuvalt rümba massist. Kui palju võinuks 2002. aasta sügisel keskmiselt raha saada 20 kg kaaluva tallerümba eest.
  - Esmalt sorteerige/filtreerige algandmed vastavalt rümpade üldkategooriale ja tehke uuele töölehele koopia tallerümpade massidest ja 1 kg hindadest.
  - ✤ Teostage regressioonanalüüs graafiliselt.
    - Selleks laske *Excelil* joonistada hajuvusdiagramm (punktdiagramm), kus *x*-teljel paiknevad rümpade massid ja *y*-teljel hinnad.
    - \* Valmis diagrammile lisage regressioonisirge, regressioonivõrrand ja viimase baasil saadavate prognooside täpsust kirjeldav determinatsioonikordaja  $R^2$ .

- ★ Lisaks tavalisele lineaarsele regressioonanalüüsile sobitage punktiparvest läbi ka ruutfunktsiooni graafik ning tellige sellegi tarvis *Excel*'lt võrrand ja  $R^2$  (parema võrdlemise huvides värvige vastav joon ja parameetrid näiteks punaseks).
- Kumba seost lineaarset või ruutseost eelistada tallerümba 1 kg hinna prognoosimisel? Miks?
- Pange töölehele kirja regressioonivõrrand ja prognoosige 20 kg kaaluva tallerümba hinda.
- **5.** Lisaülesanne. Leidke *Pivot Table*'i abil uuele töölehele rümpade arv, keskmine, minimaalne ja maksimaalne mass ning massi standardhälve sõltuvalt lamba päritolust (jõudluskontrolli alusest karjast või mitte).
  - Illustreerige leitud keskmisi tulpdiagrammiga, kus rümba masside varieeruvust kirjeldavad standardhälbed on kujutatud nö veajoontena (joonise tegemiseks tehke vajalikest *Pivot Table*'i abil leitud väärtustest abitabel).
  - Sorteerige (või filtreerige) algandmed vastavalt jõudluskontrolli alla kuulumisele ning tehke leitud keskmistega samale lehele abitabel, mis sisaldab ühes veerus jõudluskontrollialusest karjast pärit rümpade masse ja teises veerus mitte jõudluskontrollialusest karjast pärit rümpade masse.
  - Testige keskmiste masside erinevuse statistilist olulisust (esmalt F-test ja selle tulemusest lähtuvalt õiget tüüpi t-test). Sõnastage lõppjäreldus.

Kui aru ei saa (näiteks, mida mingi funktsioon teeb või miks midagi just näidatud kujul tööjuhendis realiseeritud on), siis küsi!

#### Illustreeritud (ja mittetäielik) tööjuhend

- 1. Konstrueerime järgnevalt näitena sagedustabeli rümba lihakusklassi kohta, analoogselt käib sagedustabelite tegemine ka rümba üldkategooriale ja rasvasusklassile.
  - ♦ Paigutage kursor and metabeli suvalisse lahtrisse  $\rightarrow$  *Insert*-sakk  $\rightarrow$  *PivotTable*  $\rightarrow$

1	File H	Home	Inser	t Pa	ge Layout	Formulas	Data	Review	View	, A	Add-Ins						
	24			8	1	01-+	1	XX	0	=		$2^{1/2}$	C		$\sim$	ļ.	1101
Piv	otTable Ta	ble	Picture	Clip S	hapes SmartA	rt Screensh	ot Column	Line	Pie	Bar	Area	Scatter	Othe	er	Line	Column	Win/Loss
	*			Art	Wustrations	*	· ·	Ŧ	-		Ŧ	Ť	Chart	5 *		Coorklin	200
- 67	Pivo <u>t</u> Ta	ble 、			inustrations				C	iarts				121		Sparkin	les
1 0	Pivot <u>C</u> h	art	3														
	Α		В	R	D	E	F		G		Н			J		K	L
1	Lammas	; ,	Jkk	Realis	R_mass	R_kat	Lihakusklas	s Rasva	susklas	s R	_hind	R_kg_h	nind				
2	14265	6	1	1	23.8	L	R		3		928,2		39				
3	14268	7	1	1	15,2	<u> </u>	R		3	_	577,6		39				î
4	14236	6	1	1	15,6		R		2		624		40				
5	14233	5	1	1	24,4	L			 								
6	14246	5	1	1	21,1	L	Create	rvotla	ble						1		
7	14256	4	1	1	19,1	L	Choose th	ne data ti	hat you w	vant to	analyze						
8	14247	2	1	1	22,5	L	Sele	ct a table		_							
9	14226	7	1	1	21,7	L			s or range								
10	14229	8	1	1	23,8	L		Table/Ra	nge: ar	ndmed!:	\$A\$1:\$I\$	687					-
11	14257	1	1	1	23,2	L		an exter	nal data :	source							
12	14267	0	1	1	18,1	L		Choose	Connect	ion							
13	14231	1	1	1	14,4	L		Connecti	on name:					/			
14	14258	8	1	1	12,6	L		L	on namer				1				
15	14237	5	1	1	20,5	L	Choose w	inere you	i want the	e Pivoti	able rep	ort to be	placed	1			
10	16073	5	1	1	22,5			Workshi	eet								
1/	5287	0	1	1	18,9	L	<u>⊙ E</u> xis	ting Worl	sheet		/						
10	14020	9	1	1	13,4	L		Location:	andme	ed!\$L\$2	/						
19	14238	7	1	1	21,1	L		Looda	va tab	eli va	isaku			. –			
20	14239	0	1	1	19,9	L		ilemi	e nuro	19 961	ikoht	Ok	<		Cance		
21	9205	0	1	1	14,2	L		utenns	, nurg	a ast			0.7				



Tulemuseks saadud sagedustabel:

Sum of Lammas	Lihakusklass 💌					
	E	R	P	0	P-	Grand Total
Total	159845	47271377	6695601	36751218	621984	91500025

Vaikimisi arvutab *Excel* lammaste numbrite summa ...
Et selle asemel lihtsalt kokku lugeda, kui mitu lammast mingisse lihakusklassi kuulus, tuleb ära muuta *Pivot Table-s* kasutatav funktsioon (*Sum* asemel *Count*):

vicrosoft Exc	el	PivotTabl	e Tools							
View	Add-Ins	Options	Design							
A ↓ A Z A		3		B	T.		<b>)</b>	Σ	%	f <sub>x</sub>
Z↓ Sort	Insert Slicer *	Refresh Chan So	ige Data urce *	Clear	Select	Move PivotTable	Sum Valu	marize es By ₹	Show Values As ▼	Fields, It & Sets
Sort & F	Filter	Data			Action	15		<u>S</u> um		
						(		Cour	nt	
L		М	1	1		0		<u>A</u> vera	age	Q
								Max		
Sum of Lam	nmas <mark>L</mark> iha	kusklass 💌								
	ΪE		R		Р		(	Min		
Total		159845	4	7271377	'	6695601		Prod	uct	621

PivotTable Tools-sakk  $\rightarrow Options \rightarrow Summarize Values By \rightarrow$ 

Alternatiivina võib Pivot Table-s rakendatavat funktsiooni muuta ka

a) klikkides tabelil hiire parempoolse nupuga või

b) klikkides *Pivot Table* konstrueerimise aknas lahtris *Values* muuta soovitava funktsiooni järel paikneval kolmnurgal:

Sum of La	mmas	Lihakusklass 💌									
		E	R	P	0		P-	Grand Total			
Total		159845	47271377	6695601	367	51218	621984	91500025			
	<u> </u>				<b>_</b>			a)		Drag fields between an	eas below:
	Value	Field Settings		2	<u>×</u>	1	Move <u>U</u> p		· · · · ·	Report Filter	Column Labels
	Source	Name: Lammas					Move <u>D</u> own				Lihakusklass 🔻
	⊆ustor	n Name: Count of L	.ammas				Move to Be <u>g</u> inning				
	Sum	marize Values By	bow Values As				Move to <u>E</u> nd				
	Surr	marize value fiel	1 by			Y	Move to Report Filt	er			
	Choo	ose the type of calcu		to use to summarize	·		Move to Row Labels	5			
	data	from the selected fie	eld				Move to Column La	bels			
	Sum	) Int	~ ^			Σ	Move to Values		l	Row Labels	Σ Values
	Ave	rage				×	Remove Field				Sum of Lammas
	Min	` - =	~			•	Value Field Settings	···· 4			b)
	Pro					·					0)
		per Format	ОК	Cancel							

Lihakusklasside sisuliselt õiges järjekorras esitamiseks (Excel sorteerib tähestikulises, mitte sisulises järjekorras) on lihtsaim variant vales kohas olev klass lihtsalt ümber tõsta (klikkides selleks klassi nimel (näiteks lahtril 'P') ja tõstes lahtri servast kinni hoides õigesse kohta):

Count of Lammas	Lihakusklass 💌		_			
	E	R	P	0	P-	Grand Total
Total	1	347	53	277	8	686
Tulemus	3:					
Count of Lammas	Lihakusklass 💌					
	E	R	0	P	P-	Grand Total

277

8

686

53

347

✤ Lisaks absoluutsetele sagedustele võiks leida ka suhtelised sagedused.

R

1

Total

Excel	PivotTab	le Tools												- 0	23
Add-Ins	Options	Design											۵ (?)	- 6	23
t Insert Slicer •	Refresh Cha	nge Data	Clear	Select	Move PivotTable	Summarize Values By +	Show Values As	Fields, Items, & Sets *	PivotCharl		AP What-If	Field List	+/- Buttons He	Field eaders	
& Filter	Data	a		Actions	/	/	Mo №	Calculation		Too	DIS		Show		
						-	<b>%</b> (	of <u>G</u> rand Total							~
L		M	N		0	P	~\/ <sup>%</sup>	of <u>C</u> olumn Total			PivotTable Fie	eld List		-	×
	Lihakus	klass 🛫	/			-	<b>d)<sup>*</sup>%</b> ₀	of <u>R</u> ow Total			Choose fields t	o add to r	eport	- fill	7
Data Count of Lam	E mas	B		347	2	P 277	· % (	<u>)</u> f		-	Choose nelds t		eporti		3
Count of am	mas2	- 1		347	2	277	% (	of <u>P</u> arent Row To	otal						
<b>c</b> )							% (	of P <u>a</u> rent Columi	n Total	-	Realis		$\backslash$		
							% 0	of Par <u>e</u> nt Total			R_mass		<b>a</b> )		
							<u>D</u> if	ference From		-	R_kat				3
							% [	)i <u>f</u> ference From.			✓ Lihakuskl	ass	$\setminus$		
NT 1.	D'	JT 11	<b>D</b> . 1	1		7.	,	1		-	Rasvasusk	ass			
a) Lohista	age Pivo	otTable	e Field	l List	's tuni	nus Lan	nmas	ka teine			R_hind				
kord lahti	risse Val	lues;						ā	argest	-					-
<b>b</b> ) nõudko ridadeks, (lohistage	e, et <i>Exc</i> mitte ve e kastike	<i>cel</i> jaga eergudo ε Σ Va	iks tał eks <i>lues</i>	oeli e lahtri	rineva sse <i>Rc</i>	te funkts ww Label	sioonie ls);	le alusel	allest		Drag fields bet V Report Fil	ween area ter b)	as below: Colum Lihakuskla	n Labels Iss	
<b>c)</b> määrak ( <i>Sum</i> asei	ke vajad mel) nin	usel ka g	uue 1	ea ta	rvis fu	nktsioor	niks C	ount			Row Labe	lls	Σ Values	ammas V	
<b>d</b> ) nõudk	e väärtu	ste esit	amist	prote	sentide	ena							Count of I	ammk	5
(PivotTal	ble Tool	s-sakk	$\rightarrow O$	ntion	$s \rightarrow Sl$	how Vali	ues As								
$\rightarrow \% of$	Row To	tal).	51		. 51										

Tulemus:

	Lihakusklass 💌					
Data	E	R	0	Р	P-	Grand Total
Count of Lammas	1	347	277	53	8	686
Count of Lammas2	0,15%	50,58%	40,38%	7,73%	1,17%	100,00%

Et paremaks kui "hea" (kood "E") on hinnatud vaid üht rümpa, võiks selle ühendada grupiga "R" (moodustada uus grupp – vähemalt hindega "hea" rümbad).

Layout Formulas	Data Re	view	v	íew A	dd-Ins	Options	Design			
Expand Entire Field	Group Sele	ction	₽↓	AZA		3		<b>I</b>		Σ
Collapse Entire Field	Group Field		Z A •	Sort	Insert Slicer *	Refresh Chan So	ge Data urce *	Clear S	elect Move ▼ PivotTable	Summarize Sho Values By • Value
e Field	Group	1		Sort & Fil	ter	Data		ł	Actions	Calcu
L	M			Ν		0		Р	Q	R
	Lihakusklass	-	1	\						
Data	E		R	,	0		Ρ		P-	Grand Total
Count of Lammas		1		3	47	277		53		8 686
Count of Lammas2	0,1	15%		50,58	3%	40,38%		7,73%	1,17	% 100,00%

		Lihakusklass2 ~	Lihakusklass 🕆	-			
		80	Group1		Group2		Grand Total
I	Data	0	E	R	Р	P-	
I	Count of Lammas	277	1	347	53	8	686
1	Count of Lammas2	40,38%	0,15%	50,58%	7,73%	1,17%	100,00%

	Lihakusklass2 🔫	Lihakusklass 😁	_	
	80 💙	🗄 Group1 🔵	Group2	Grand Total
Data	0		P P-	
Count of Lammas	277	348	53 8	686
Count of Lammas2	40,38%	50,73%	7,73% 1,17%	100,00%

Grupeerige analoogsel viisil ka "lahjad" ja "eriti lahjad" rümbad (grupid "P" ja "P-", sest ega seal suurt vahet pole).

	Lihakusklass2 🗠	Lihakusklass 🗠		
	<b>0</b>	🗄 Group1 🤇	E Group2	Grand Total
Data	0	/		
Count of Lammas	277	348	61	686
Count of Lammas2	40,38%	50,73%	8,89%	100,00%

Pivot Table'i abil konstrueeritud tabeli põhjal kenade jooniste tegemiseks on sageli mõttekas teha vajalikest väärtustest abitabel ja joonistada diagramm abitabeli alusel.

Põhjuseks on see, et otse *Pivot Table* alusel joonise tegemise tulemuseks on nn *Pivot Chart*, mis on sarnaselt *Pivot Table*'ga lingitud andmetabeliga, seeläbi kergesti täiendatav ja ümberarvutatav, aga ei võimalda muuta kõike tavalisel *Excel*'i diagrammil muudetavat (või on see märksa keerulisem).

	Lihak	usklass2	-	Lihakusklass	-		
Conv	Ξ0			E Group1		∃ Group2	Grand Total
Data	0						
Count of Lammas		2	277	3	48	61	686
Count of Lammas2		40,3	8%	50,73	3%	8,89%	100,00%
	Ж	Cu <u>t</u>					
		<u>С</u> ору					
	1	Paste Op	otio	ns:			
		12	3	<i>f</i> x 🚹 %	2		
		Paste Sn Val	eci: lue:	s (V) on tells		▶	

Järgnevalt andke lihakusklassidele sisuliselt õiged nimed – trükkige need moodustatud abitabelisse, ja konstrueerige abitabeli vastavate lahtrite alusel sektordiagramm.

Kujundage saadud joonis (lisage sektoritele vastavate lihakusklasside nimed ja esinemissagedused, muutke soovi korral värve):



Järgnevalt konstrueerige sagedustabelid ja joonistage nende alusel sektordiagrammid ka rümba (üld)kategooriale ja rasvasusklassile. 2. Kas rümpade jagunemine rasvasusklassidesse sõltub rümba üldkategooriast?

Võimaliku seose kirjeldamiseks konstrueerige (uuele töölehele) vastav 2-mõõtmeline sagedustabel, viimasesse leidke nii rea- kui ka veeruprotsendid ja sõnastage lause(d), kasutades vähemalt kahte leitud suhtelistest sagedustest.



1

S

	F			Rasvasusklass -					
R_kat 🗠	Data	R_kat ~	Data	1	2	3	4	5	Grand Total
L	Sum of Lammas	L	Count of Lammas	69	301	132	14	10	526
1	Sum of Lammas2		Count of Lammas2	69	301	132	14	10	526
	Sum of Lammas3		Count of Lammas3	69	301	132	14	10	526
S	Sum of Lammas	S	Count of Lammas	2	35	58	34	31	160
١	Sum of Lammas2		Count of Lammas2	2	35	58	34	31	160
	Sum of Lammas3		Count of Lammas3	2	35	58	34	31	160
Total Sum of Lammas		Total Count of Lammas		71	336	190	48	41	686
Total Sum of Lammas2		Total Count of Lammas2		71	336	190	48	41	686
Total Sum of Lammas3		Total Co	unt of Lammas3	71	336	190	48	41	686
			1						

Muutke kõik kolm *Exceli* poolt leitud summat vaatluste arvudeks ( $Sum \rightarrow Count$ ):

#### Ning esitage 2. ja 3. vaatluste arv (Count of ... 2 ja 3) vastavalt rea- ja veeruprotsendina:

le H	lome Insert Pa	ige Layout Formula	s Data	Review View	Add-Ins	Options	Design				
tTable Nar tTable3 Options ▼ ivotTable	ne: Active Field: Count of Lammas: Pield Settings Ac fx Count of Lamma	●를 Expand Entire Field ■를 Collapse Entire Fiel tive Field mas2	Group S Ungrou d En Group F Grou	p Z↓ Z Field Z↓ S	iort Insert Slicer + ort & Filter	Refresh Char So Data	nge Data ource •	r Select	Move PivotTable	Σummarize Values By ▼	% Fields, Items, & Pivot & & Sets ▼ % Calculation % of Grand Total
Α	В	С	D	E	F	G	Н	1		J	% of <u>Column Total</u>
		Drop	Report Filter I	lelds Here							N <sub>% of Row Total</sub>
		Rasvasusklass 💌	_								% <u>O</u> f
R kat 🔻	Data	1	2	3	4	5	Grand Total				% of Parent Row Total
L	Count of Lammas	69	301	132	14	10	526				% of Parent Column Total
	Count of Lammas2 *	69	301	132	14	10	526				
	Count of Lammas3	69	301	132	14	10	526				% of Parent Total
S	Count of Lammas	2	35	58	34	31	160				Difference From
	Count of Lammas2	2	35	58	34	31	160				% Difference From
	Count of Lammas3	2	35	58	34	31	160				Running Total In
Total Co	unt of Lammas	71	336	190	48	41	686				of Description Table In
Total Co	unt of Lammas2	71	336	190	48	41	686				% K <u>u</u> nning Total In
Total Co	unt of Lammas3	71	336	190	48	41	686				Rank <u>S</u> mallest to Largest

#### Tulemus:

	-	Rasvasusklass 💌					
R_kat 💌	Data	1	2	3	4	5	Grand Total
L	Count of Lammas	69	301	132	14	10	526
	Count of Lammas2	13,12%	57,22%	25,10%	2,66%	1,90%	100,00%
	Count of Lammas3	97,18%	89,58%	69,47%	29,17%	24,39%	76,68%
S	Count of Lammas	2	35	58	34	31	160
	Count of Lammas2	1,25%	21,88%	36,25%	21,25%	19,38%	100,00%
	Count of Lammas3	2,82%	10,42%	30,53%	70,83%	75,61%	23,32%
Total Cou	unt of Lammas	71	336	190	48	41	686
Total Cou	unt of Lammas2	10,35%	48,98%	27,70%	7,00%	5,98%	100,00%
Total Cou	unt of Lammas3	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Kommentaarid. ... ... (sõnastage lause(d), kasutades vähemalt kahte leitud suhtelistest sagedustest)

- **3.** Rümpade üldkategooriatesse ja rasvasuklassidesse jagunemise vahelise seose statistiline olulisus.
  - ♦ Et oleks selge, mida te üldse testite, pange esmalt kirja kontrollitav hüpoteeside paar.
  - Järgnevalt konstrueerige uus 2-mõõtmeline sagedustabel, mis sisaldab üksnes absoluutseid sagedusi,

H <sub>0</sub> (nullhüpotees):								
H <sub>1</sub> (alternatiivne hü	potees):							
Empiirilised (and	<u>metest arvutatud)</u>	sagedused						
Count of Lammas	Rasvasusklass 💌							
R_kat 💌	1	2	3	4	5	Grand Total		
L	69	301	132	14	10	526		
S	2	35	58	34	31	160		
Grand Total	71	336	190	48	41	686		

tehke konstrueeritud tabeli <u>väärtustest</u> koopia

Empiirilised (ar	ndmetest arvuta	atud) sa	agedused				
Count of Lamma	s Rasvasusklas	s 🔻					
R_kat	<b>*</b>	1	2	3	4	5	Grand Total
L		69	301	132	14	10	526
S L		2	35	58	34	31	160
Grand Total		71	336	190	48	41	686
	1 64						
	s Cu <u>t</u>						
	<u>⊇</u> ору						
<b>[</b>	Raste Options						
	123 f		<b>%</b>				
	Paste Special Values (V Insert Copied	/) Cells	<b>&gt;</b>				

#### ✤ ja kustutage kopeeritud tabeli sisu (alles jätke rea- ja veerusummad!)

Count of Lammas	Rasvasusklass					
R_kat	1	2	3	4	5	Grand Total
L			Delete			526
S			Delele			160
Grand Total	71	336	190	48	41	686

 Arvutage uude tabelisse kustutatud sageduste asemele tunnuste sõltumatuse juhule (nullhüpoteesile) vastavad nö teoreetilised sagedused (kujul: reasumma × veerusumma / vaatluste arv). Teoreetiliste sageduste arvutamine:

	A	В	С	D	E	F	G	
25								
26	<u>Teoreetilised (nu</u>	<u>llhüpoteesile e sõ</u>	ltumatuse juhule	vastavad) sagedu	sed			
27	Count of Lammas	Rasvasusklass						
28	R_kat	1	2	- 3	4	5	Grand Total	
29	L	= <mark>\$G29</mark> *B\$31/\$G\$3					526	
30	S						160	
31	Grand Total	71	336	190	48	41	686	
20	1						T T	

٠	٠	٠	٠	٠	٠

	A	В	С	D	E	F	G
25							
26	<u>Teoreetilised (nu</u>	<u>llhüpoteesile e sõ</u>	ltumatuse juhule <sup>.</sup>	vastavad) sagedu	sed		
27	Count of Lammas	Rasvasusklass					
28	R_kat	1	2	3	4	5	Grand Total
29	L	54,44023324	257,6326531	145,6851312	36,80466472	31,43731778	526
30	S	16,55976676	78,36734694	44,3148688	11,19533528	<b>=\$G30*</b> F\$31/\$G\$31	( <u>160</u>
31	Grand Total	71	336	190	48	41	686

• Teostage funktsiooni CHISQ.TEST (*Excel* 2003-s CHITEST) abil  $\chi^2$ -test – viimane võrdleb empiirilisi (andmetabelist arvutatud) sagedusi teoreetiliste (sõltumatuse juhule vastavate) sagedustega ja väljastab olulisuse tõenäosuse *p* väärtuse

Et oleks lihtsam aru saada, milliste arvude võrdlemisel  $\chi^2$ -test baseerub (ehk siis millised tabelite osad tuleb *Excel*'i vastavale funktsioonile ette anda), võib vastavad lahtrid selguse mõttes näiteks ära värvida.

Järgnevalt, nagu funktsioonide puhul ikka, tuleb kursor panna lahtrisse, kuhu soovitakse tulemust saada (ja juurde võiks enne ka kirjutada, mida arvutama hakatakse).

•	(• ★ ✓ <i>f</i> <sub>x</sub>   =CHISQ.TEST(B21:F22;B29:F30)										
	Α	В	С	D	E	F	G				
17											
18	Empiirilised (and	metest arvutatud)	sagedused								
19	Count of Lammas	Rasvasusklass 💌									
20	R_kat 💌	1	2	3	4	5	Grand Total				
21	L	69	301	132	14	10	526				
22	S	2	35	58	34	31	160				
23	Grand Total	71	336	190	48	41	686				
24											
25											
26	Teoreetilised (nu	Ilhüpoteesile e sõ	ltumatuse juhule	vastavad) sagedus	sed						
27	Count of Lammas	Rasvasusklass									
28	R_kat	1	2	3	4	5	Grand Total				
29	L	54,44023324	257,6326531	145,6851312	36,80466472	31,43731778	526				
30	S	16,55976676	78,36734694	44,3148688	11,19533528	9,562682216	160				
31	Grand Total	71	336	190	48	41	686				
32											
33											
34	Hii-ruut-test	=CHISQ.TEST(B21	:F22;B29:F30)								
25											

Sõnastage lõppjäreldus (Kas seos on statistiliselt oluline? Miks te nii otsustasite?).

|--|

- **4.** Prognoosige tallerümpade 1 kg hinda lähtuvalt rümba massist. Kui palju võinuks 2002. aasta sügisel keskmiselt raha saada 20 kg kaaluva tallerümba eest.
  - Esmalt sorteerige/<u>filtreerige</u> algandmed vastavalt rümpade üldkategooriale ja tehke uuele töölehele koopia tallerümpade massidest ja 1 kg hindadest.



✤ ja tehke uuele töölehele koopia tallerümpade massidest ja 1 kg hindadest.

4	Α	В	С	D	E	F	G	Н	1
1	Lamm 😁	Jkk 🛫	Reali:	R_mass 🐣	R_ka 🐃	Lihakuskla 🕋	Rasvasuskia 🕋	R_hine ~	R kq hind *
2	142656	1	1	23,8	L	R	3	928,2	39
3	142687	1	1	15,2	L	R	3	577,6	39
4	142366	1	1	15,6	L	R	2	624	40
5	142335	1	1	24,4	L	R	3	951,6	39
6	142465	1	1	21,1	L	R	3	822,9	39
7	142564	1	1	19,1	L	R	4	706,7	37
8	142472	1	1	22,5	L	R	3	877,5	39
9	142267	1	1	21,7	L	R	3	846,3	39
10	142298	1	1	23,8	L	R	4	880,6	37
11	142571	1	1	23,2	L	R	5	788,8	34
12	142670	1	1	18,1	-	R	4	669,7	37
13	142311	1	1	14,4	L	Com	<b>~</b> 2	475,2	33
14	140588	1	1	12.6		Сору	2	415.8	22

- ✤ Teostage regressioonanalüüs graafiliselt.
  - Selleks laske *Excelil* joonistada hajuvusdiagramm (punktdiagramm), kus *x*-teljel paiknevad rümpade massid ja *y*-teljel hinnad.
  - Valmis diagrammile lisage regressioonisirge, regressioonivõrrand ja viimase baasil saadavate prognooside täpsust kirjeldav determinatsioonikordaja  $R^2$ .
  - Lisaks tavalisele lineaarsele regressioonanalüüsile sobitage punktiparvest läbi ka ruutfunktsiooni graafik ning tellige sellegi tarvis *Excel*'lt võrrand ja  $R^2$  (parema võrdlemise huvides värvige vastav joon ja parameetrid näiteks punaseks).
  - Kumba seost lineaarset või ruutseost eelistada tallerümba 1 kg hinna prognoosimisel? Miks?





Pange töölehele kirja regressioonivõrrand ja prognoosige 20 kg kaaluva tallerümba hinda.