

Praktikum 1

Salvestage kursuse kodulehelt omale arvutisse andmestik *kalad.xls*.
(http://www.eau.ee/~ktanel/DK_0007/kalad.xls)

Andmestik, mis on osa Mariann Nõlvaku poolt aastail 2004-2006 kogutud Eesti kalade andmebaasist, sisaldab järgmisi andmeid:

- kala number (lihtsalt identifitseerimiseks);
- liik (6 liiki: haug, särg, latikas, luts, ahven ja koha);
- rühm: röövkala või lepiskala;
- 5 püügikohta (Võrtsjärv, Kärevere, Kastre, Praaga ja Peipsi järv);
- püügisesoon (kevad-suvi või sügis-talv);
- kaal ja pikkus;
- sugu;
- lõpuse-, silma- ja sooleparasiitide ning laiussi (*Diphyllobothrium latum*) leidude arv kalal (arv_lõpusepar, arv_soolepar, arv_silmapar ja arv_Diphyllob).
- lõpuse-, silma- ja sooleparasiitidega ning laiussiga nakatumine (kas jah või ei kodeerituna 1 ja 0);
- parasiitide arv kala kohta (par_per_kala).

OSA 1

--- Arvkarakteristikud ---

❖ Funktsioonide kasutamine

- 1) Leidke andmetabeli alla püütud ahvenate arv, keskmine pikkus, pikkuse standardhälve ja standardviga, mediaan, alumine ja ülemine kvartiil, minimaalne ja maksimaalne väärtus (seejuures uurige välja, mida tähendab teine argument funktsioonil *QUARTILE*).
- 2) Püüdke sõnastada leitud suuruste alusel vähemalt 2 lauset (kasutades vähemalt ühes mediaani ja/või kvartiile).

1	A	B	C	D	E	F	G	H	1	A	B	C	D	E	F	G	H	I	J	P	Q				
kala nr				koht	sesoon	sugu	kaal	pikkus	ary	kala nr	liik	ryhm	pyygihoht	sesoon	sugu	kaal	pikkus	ary	lopuse	ary	silms	Diphylob	par	per	ka
1010	7			är	kevad-suvi	e	144	24	2	374	Ahven	rookkala	Võrtsjärv	kevad-suvi	i	131	22,5	10	0	0	0	0	0	2	
1011	7			i	kevad-suvi	e	108	22	3	372	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	202	25	0	0	0	0	0	0	2	
1012	6			a	kevad-suvi	e	135	23	4	560	Ahven	rookkala	Peipsi	sygis-talv	e	134	21,5	6	0	0	0	0	0	2	
1013	6			a	kevad-suvi	e	150	23,5	5	377	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	84	18,5	6	0	0	0	0	0	2	
1014	6			a	kevad-suvi	i	123	22,5	6	365	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	145	23	6	0	0	0	0	0	2	
1015	6			e	kevad-suvi	i	92	20	7	335	Ahven	rookkala	Praaga	kevad-suvi	i	95	19,5	20	0	0	0	0	0	2	
1016	6			e	kevad-suvi	i	103	20,5	8	64	Ahven	rookkala	Kastre	kevad-suvi	e	166	23	24	1	0	0	0	0	2	
1017	6			e	kevad-suvi	i	109	21	9	527	Ahven	rookkala	Võrtsjärv	sygis-talv	i	152	22	2	0	0	0	0	0	2	
1018	6			e	kevad-suvi	i	117	22	10	525	Ahven	rookkala	Võrtsjärv	sygis-talv	e	218	24,5	2	0	0	0	0	0	2	
1019	6			e	kevad-suvi	i	124	22,5	11	376	Ahven	rookkala	Võrtsjärv	kevad-suvi	i	111	22	4	0	0	0	0	0	1	
1020	6			e	kevad-suvi	i	130	22,5	12	375	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	132	23	0	0	0	0	0	0	2	
1021	6			e	kevad-suvi	i	109	21	13	360	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	351	31	22	0	0	0	0	0	2	
1022	6			e	kevad-suvi	i	117	22	14	347	Ahven	rookkala	Praaga	kevad-suvi	i	70	19	14	0	0	0	0	0	2	
1023	5			e	kevad-suvi	e	124	22,5	15	344	Ahven	rookkala	Praaga	kevad-suvi	e	61	18	30	1	0	0	0	0	2	
1024	5			e	kevad-suvi	e	130	22,5	16	343	Ahven	rookkala	Praaga	kevad-suvi	i	93	20	14	0	0	0	0	0	2	
1025	5			ere	kevad-suvi	e	308	27,5	17	310	Ahven	rookkala	Peipsi	kevad-suvi	e	87	19	18	1	0	0	0	0	1	
1026	5			ere	kevad-suvi	e	156	22,5	18	755	Ahven	rookkala	Peipsi	kevad-suvi	i	133	22	24	0	0	0	0	0	2	
1027	5			ere	kevad-suvi	e	133	21,5	19	754	Ahven	rookkala	Peipsi	kevad-suvi	i	194	23,5	0	0	0	0	0	0	2	
1028				i	kevad-suvi	e	267	28	20	753	Ahven	rookkala	Peipsi	kevad-suvi	i	192	23,5	2	0	0	0	0	0	2	
1029				i	kevad-suvi	e	116	22	21	521	Ahven	rookkala	Võrtsjärv	sygis-talv	i	229	24	0	0	0	0	0	0	2	
1030	10			ere	sygis-talv	e	121	21	22	487	Ahven	rookkala	Kastre	sygis-talv	e	45	15	2	1	0	0	0	0	2	
1031	10			e	sygis-talv	e	76	18	23	373	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	141	21	0	0	0	0	0	0	2	
1032	10			e	sygis-talv	e	89	19,5	24	371	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	248	27	4	11	0	0	0	0	2	
1033	10			e	sygis-talv	i	138	22	25	361	Ahven	rookkala	Võrtsjärv	kevad-suvi	e	250	27,5	10	1	0	0	0	0	2	
1034	10			e	sygis-talv	e	147	22	26	342	Ahven	rookkala	Praaga	kevad-suvi	i	43	15,5	18	6	0	0	0	0	2	
1035	10			e	sygis-talv	e	137	22,5	27	341	Ahven	rookkala	Praaga	kevad-suvi	i	58	18	26	0	0	0	0	0	2	
1036	10			e	sygis-talv	e	126	22	28	339	Ahven	rookkala	Praaga	kevad-suvi	i	70	18,5	28	0	0	0	0	0	2	
1037	9			är	sygis-talv	e	33	13	29	334	Ahven	rookkala	Praaga	kevad-suvi	e	123	22	28	0	0	0	0	0	2	
1038	9			är	sygis-talv	e	32	13	30	329	Ahven	rookkala	Kastre	kevad-suvi	e	76	18	50	0	0	0	0	0	1	
1039	904	Särg	lepiskala	Peipsi	sygis-talv	e	81	18	31	319	Ahven	rookkala	Peipsi	kevad-suvi	e	80	18,5	72	0	0	0	0	0	1	
1040	953	Särg	lepiskala	Peipsi	sygis-talv	i	119	21	32	318	Ahven	rookkala	Peipsi	kevad-suvi	e	77	18,5	4	0	0	0	0	0	1	
1041	933	Särg	lepiskala	Praaga	sygis-talv	e	134	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
1042	651	Särg	lepiskala	Kastre	kevad-suvi	i	101	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
1043	636	Särg	lepiskala	Kastre	kevad-suvi	e	119	21,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
1044																									
1045																									
1046		Ahven	kalade arv																						179
1047			keskmine																						20,45
1048			standardhälve																						3,55
1049			standardviga																						0,27
1050			mediaan																						20
1051			alumine kvartiil																						18
1052			ülemine kvartiil																						23
1053			min																						14
1054			max																						32

$se = \frac{s}{\sqrt{n}}$

=COUNT(H2:H180)
 =AVERAGE(H2:H180)
 =STDEV(H2:H180)
 =H1048/SQRT(H1046)
 =MEDIAN(H2:H180)
 =QUARTILE(H2:H180;1)
 =QUARTILE(H2:H180;3)
 =MIN(H2:H180)
 =MAX(H2:H180)

3) Leidke samad suurused ka ahvenate kaalule (massile).

Ahven	kalade arv									179
	keskmine									20,45
	standardhälve									3,55
	standardviga									0,27
	mediaan									20
	alumine kvartiil									18
	ülemine kvartiil									23
	min									14
	max									32

4) Arvutage laiussi noorvormide arv keskmiselt kõigi ahvenate kohta ja vaid nakatunud ahvenate kohta. Kui mitu % ahvenatest üleüldse olid laiussiga nakatunud?

`=SUM(L2:L180)/COUNTIF(L2:L180;">0")`

Keskmine laiussi noorvormide arv vaid nakatunud ahvenatel

	K	L	M	N	O	P	Q	R		
1	av	soolep	av	Diphyllob	lõpusepar	silmapar	soolepar	Diphyllob	par	kala
1039		0	0	0	1	0	0	0	1	
1040		0	0	0	1	0	0	0	1	
1041		0	0	0	1	0	0	0	1	
1042		0	0	0	1	0	0	0	1	
1043		0	0	0	1	0	0	0	1	
1044										
1045										
1046										
1047										
1048										
1049										
1050										
1051										

Keskmine laiussi noorvormide arv ahvenatel

`=SUM(L2:L180)/COUNTIF(L2:L180;">0")` `=AVERAGE(P2:P180)`

`=AVERAGE(L2:L180)`

`=COUNTIF(L2:L180;">0")/COUNT(L2:L180)`

Laiussi noorvormidega nakatunud ahvenate osakaal

	K	L	M	N	O	P	Q		
1	av	soolep	av	Diphyllob	lõpusepar	silmapar	soolepar	Diphyllob	par
1042		0	0	0	1	0	0		
1043		0	0	0	1	0	0		
1044									
1045									
1046									
1047									
1048									
1049									
1050									
1051									

Format Cells dialog box showing Percentage format with 2 decimal places.

Keskmine laiussi noorvormide arv nakatunutel:	1,3333	Laiussi noorvormidega nakatunud kalade osakaal:	1,68%
Keskmine laiussi noorvormide arv:	0,0223		
Laiussi noorvormidega nakatunud kalade osakaal:	1,68%		

- 5) Ja nüüd arvutage samad suurused ka haugide tarvis (kasutades maksimaalselt ära juba rakendatud funktsioone ja vorminguid – see tähendab, et tehke koopia kõigest ahvenate kohta kirjutatust ja arvutatust, edasi jääb üle muuta vaid funktsioonide argumente).

Ahven						Keskmine laiussi noorvormide arv nakatunutel:	Laiussi noorvormidega nakatunud kalade osakaal:
kalade arv		179	179			1,3333	1,68%
keskmine		118,60	20,45			Keskmine laiussi noorvormide arv:	
standardhälve		73,00	3,55			0,0223	
standardviga		5,46	0,27			Laiussi noorvormidega nakatunud kalade osakaal:	
mediaan		95	20			1,68%	
alumine kvartiil		70	18				
ülemine kvartiil		151,5	23				
min		34	14				
max		517	32				
Haug						Keskmine laiussi noorvormide arv nakatunutel:	Laiussi noorvormidega nakatunud kalade osakaal:
kalade arv		214	214			1,3214	13,08%
keskmine		785,63	48,76			Keskmine laiussi noorvormide arv:	
standardhälve		339,53	5,84			0,1729	
standardviga		23,21	0,40			Laiussi noorvormidega nakatunud kalade osakaal:	
mediaan		738	48,25			13,08%	
alumine kvartiil		578,75	45				
ülemine kvartiil		881	51,5				
min		238	26				
max		3013	74				

❖ Olulisemate arvarakteristikute leidmise teisi variante

Kuigi *MS Exceli* funktsioonidel on hulk positiivseid omadusi

- argumentide ette andmine on enamasti intuitiivselt mõistetav;
- tänu võimalusele funktsioone kopeerida on kord juba sisestatud käsud lihtsalt rakendatavad uutele argumentidele (väärtustele, tunnustele);
- funktsioonide omavaheline kombineerimine võimaldab väljastada keeruliste avaldiste tulemusi;

on teatud olukordades siiski sobivam kasutada mõnd teist *MS Exceli* vahendit. Näiteks

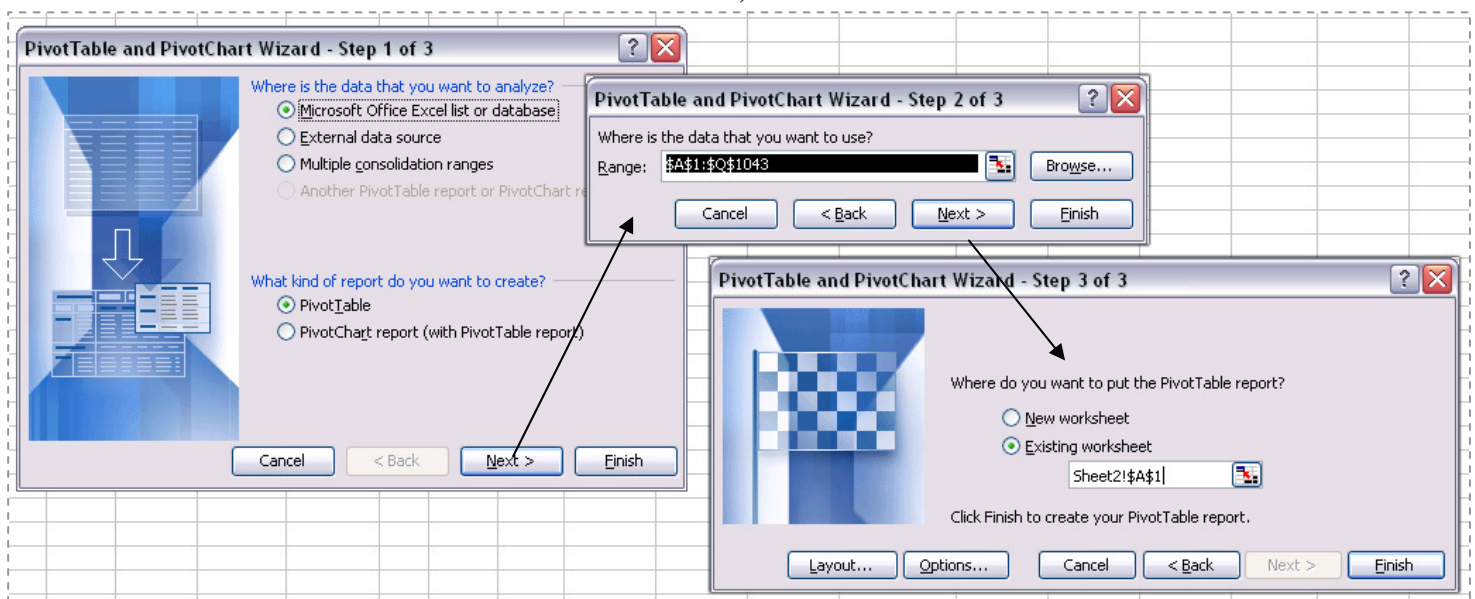
1. arvutamaks keskmisi väärtusi gruppides (näiteks keskmisi pikkusi erinevat liiki kaladel) on mugav kasutada *Pivot Table*'t (menüüst *Data / Andmed ->*);
2. leidmaks ühe korraga suurt hulka arvarakteristikuid, samuti usalduspiire keskmisele, võib kasutada statistikaprotseduuri *Descriptive Statistics (Tools / Tööriistad -> Data Analysis ->)*;
3. saamaks kiirelt teada mõne arvarakteristiku väärtust, ilma seda kuhugi töölehe lahtrisse arvutamata, võib kasutada töölehe allservas kuvatavaid selekteeritud lahtrite sisu kirjeldavaid väärtusi.

▪ *Pivot Table*

- 1) Arvutage kõigi kalaliikide kohta püütud kalade arv, keskmine, minimaalne ja maksimaalne pikkus ning laiussi noorvormiga nakatunud kalade osakaal vahendi *Pivot Table* abil.

--- Tööjuhend ---

- Paigutage kursor andmetabeli suvalisse lahtrisse.
- *Data / Andmed* → *PivotTable and PivotChart Report... / PivotTable ja PivotChart Aruanne...*
- Loodav tabel paigutage töölehe 'Sheet2' ülemisse vasakusse nurka (hiljem nimetage tööleht 'Sheet2' ümber leheks 'Pivot Table').



The image illustrates the steps to create a PivotTable in Microsoft Excel. It shows a PivotTable Field List, a PivotTable, a PivotTable Field dialog box, and a PivotTable with two data series.

PivotTable Field List (Initial):

- Fields: kala_nr, liik, ryhm, pyygikoht, sesoon, sugu, kaal, pikkus, arv_16pusepar, arv_silmapar, arv_soolepar

PivotTable (Initial):

liik	Total
Ahven	61256
Haug	69687
Koha	43300
Latikas	139832
Luts	62075
Särg	183458
Grand Total	559608

PivotTable Field Dialog (Initial):

- Source field: kala_nr
- Name: Count of kala_nr
- Summarize by: Count

PivotTable (Second):

liik	Total
Ahven	179
Haug	214
Koha	111
Latikas	173
Luts	137
Särg	228
Grand Total	1042

PivotTable Field List (Second):

- Fields: kala_nr, liik, ryhm, pyygikoht, sesoon, sugu, kaal, pikkus, arv_16pusepar, arv_silmapar, arv_soolepar

PivotTable (Final):

liik	Data	Total
Ahven	Count of kala_nr	179
	Sum of pikkus	3660,1
Haug	Count of kala_nr	214
	Sum of pikkus	10434
Koha	Count of kala_nr	111
	Sum of pikkus	4425
Latikas	Count of kala_nr	173
	Sum of pikkus	6833,5
Luts	Count of kala_nr	137
	Sum of pikkus	6272,5
Särg	Count of kala_nr	228
	Sum of pikkus	5005
Total Count of kala_nr		1042
Total Sum of pikkus		36630,1

- 2) Ümardage keskmised 1 kümnendkohani ja esitage laiussi noorvormide esinemissagedused %-des (2 kümnendkohaga); vaadake, kas oskate tabeli nõ pikali pöörata.

liik	Data	Total				
Ahven	Count of kala_nr	179				
	Average of pikkus	20,4				
	Min of pikkus2	14				
	Max of pikkus3	32				
	Average of Diphylob	1,68%				
Haug	Count of kala_nr	214				
	Average of pikkus	48,8				
	Min of pikkus2	26				
	Max of pikkus3	74				
	Average of Diphylob	13,08%				
Koha	Count of kala_nr	111				
	Average of pikkus	39,9				
	Min of pikkus2	22				
	Max of pikkus3	54				
	Average of Diphylob	0,00%				
Latikas	Count of kala_nr	173				
	Average of pikkus	39,5				
	Min of pikkus2	27				
	Max of pikkus3	48				
	Average of Diphylob	0,00%				
Luts	Count of kala_nr	137				

	A	B	C	D	E	F
1	Data					
2	liik	Count of kala_nr	Average of pikkus	Min of pikkus2	Max of pikkus3	Average of Diphylob
3	Ahven	179	20,4	14	32	1,68%
4	Haug	214	48,8	26	74	13,08%
5	Koha	111	39,9	22	54	0,00%
6	Latikas	173	39,5	27	48	0,00%
7	Luts	137	45,8	27	70	43,07%
8	Särg	228	22,0	11	45	0,00%
Total Count	Grand Total	1042	35,2	11	74	8,64%

Total Average of pikkus	35,2
Total Min of pikkus2	11
Total Max of pikkus3	74
Total Average of Diphylob	8,64%

- 3) Oskate seda tabelit viia ka järgmisele kujule?

	liik						
Data	Ahven	Haug	Koha	Latikas	Luts	Särg	Grand Total
Count of kala_nr	179	214	111	173	137	228	1042
Average of pikkus	20,4	48,8	39,9	39,5	45,8	22,0	35,2
Min of pikkus2	14	26	22	27	27	11	11
Max of pikkus3	32	74	54	48	70	45	74
Average of Diphylob	1,68%	13,08%	0,00%	0,00%	43,07%	0,00%	8,64%

▪ **Protseduur *Descriptive Statistics***

1) Arvutage nii palju erinevaid arvkarakteristikuid ahvenate pikkuse ja kaalu kohta, kui protseduur *Descriptive Statistics* võimaldab.

Descriptive Statistics

Input

Input Range:

Grouped By: Columns Rows

Labels in first row ← Kas on vaja?

Output options

Output Range:

New Worksheet Ply:

New Workbook

Summary statistics

Confidence Level for Mean: %

Kth Largest:

Kth Smallest:

<i>kaal</i>		<i>pikkus</i>	
Mean	118,6	Mean	20,44749
Standard Error	5,4561	Standard Error	0,26514
Median	95	Median	20
Mode	75	Mode	18
Standard Deviation	72,997	Standard Deviation	3,547326
Sample Variance	5328,6	Sample Variance	12,58352
Kurtosis	5,9845	Kurtosis	0,328945
Skewness	2,034	Skewness	0,771968
Range	483	Range	18
Minimum	34	Minimum	14
Maximum	517	Maximum	32
Sum	21230	Sum	3660,1
Count	179	Count	179
Largest(2)	393	Largest(2)	31
Smallest(2)	41	Smallest(2)	14
Confidence Level(95,0%)	10,767	Confidence Level(95,0%)	0,523221

- 2) *** Selles alapunktis räägitavaist usalduspiiridest ja hüpoteeside testimisest tuleb juttu alles järgmises loengus, aga eelnevate statistikursuste olemasolul ei tohiks antud leheküljel presenteeritu üle mõistuse käia. ***

Kas olulisuse nivool 0,05 võib väita, et uuritud ahvenate keskmine kaal erineb 100 grammist?

--- Lahenduse idee ja tööjuhend ---

Kui võrreldav konstant jääb keskmise 95%-usalduspiiride vahele, ei saa väita, et keskmine erineks konstandist, ja vastupidi:

$$\begin{array}{l}
 H_0: \mu = 100 \\
 H_1: \mu \neq 100
 \end{array}
 \quad \text{kui} \quad \begin{array}{c} | \quad | \quad | \\ \mu \quad 100 \quad \bar{\mu} \end{array} \quad \text{siis } H_0: \mu = 100,$$

$$\text{kui aga} \quad \begin{array}{c} | \quad | \quad | \\ \mu \quad \bar{\mu} \quad 100 \end{array} \quad \text{siis } H_1: \mu \neq 100.$$

	kaal		pikkus	
1070				
1071				
1072	Mean	118,6	Mean	20,44749
1073	Standard Error	5,4561	Standard Error	0,26514
1074	Median	95	Median	20
1075	Mode	75	Mode	18
1076	Standard Deviation	72,997	Standard Deviation	3,547326
1077	Sample Variance	5328,6	Sample Variance	12,58352
1078	Kurtosis	5,9845	Kurtosis	0,328945
1079	Skewness	2,034	Skewness	0,771968
1080	Range	483	Range	18
1081	Minimum	34	Minimum	14
1082	Maximum	517	Maximum	32
1083	Sum	21230	Sum	3660,1
1084	Count	179	Count	179
1085	Largest(2)	393	Largest(2)	31
1086	Smallest(2)	41	Smallest(2)	14
1087	Confidence Level(95,0%)	10,767	Confidence Level(95,0%)	0,523221
1088	Alumine usalduspiir	=B1072-B1087		
1089	Ülemine usalduspiir	=B1072+B1087		

Alumine usalduspiir	107,84
Ülemine usalduspiir	129,37

Kumb väidetest on korrektne?

- Ei ole alust väita, et uuritud ahvenate keskmine kaal erineks 100 grammist ($p > 0,05$).
- Uuritud ahvenate keskmine kaal erineb / on suurem kui 100 gr ($p < 0,05$).

- Kõige kiirem variant mõne olulisema arvarakteristiku väärtuse teada saamiseks *Excelis* (näiteks minimaalse ja maksimaalse väärtuse välja selgitamiseks andmete kontrollimise või sagedustabeli klasside moodustamise huvides) on järgmine:
 - võtate blokki huvipakkuvad lahtrid,
 - klikite hiire parempoolse klahviga *Exceli* tööakna alumisel ribal (vt joonist) ja valite avanenud rippmenüüst teid huvitava funktsiooni (nt. *Min*),
 - valitud funktsiooni väärtus kuvatakse samas akna alumisel serval, aga seda ei trükitu kuhugi tabeli lahtrisse,
 - mõne muu funktsiooni (näiteks *Max*) tarvis tuleb eelnevat korrata.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	kala_nr	liik	ryhm	pyygikoht	sesoon	sugu	kaal	pikkus	arv_lõpusep	arv_silmapar	arv_soolepa	arv_Diphylob	lõpusepar	silmapar	soolepar	Diphy
2	374	Ahven	roovkala	Võrtsjärv	kevad-suvi	i	131	22,5	10	22	25	0	1	1	1	1
3	372	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	202	25	0	49	36	0	0	1	1	1
4	560	Ahven	roovkala	Peipsi	sygis-talv	e	134	21,5	6	47	12	0	1	1	1	1
5	377	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	84	18,5	6	60	2	0	1	1	1	1
6	365	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	145	23	6	32	30	0	1	1	1	1
7	335	Ahven	roovkala	Praaga	kevad-suvi	i	85	19,5	20	0	24	0	1	0	1	1
8	64	Ahven	roovkala	Kastre	kevad-suvi	e	166	23	24	13	6	2	1	1	1	1
9	527	Ahven	roovkala	Võrtsjärv	sygis-talv	i	152	22	2	22	6	0	1	1	1	1
10	525	Ahven	roovkala	Võrtsjärv	sygis-talv	e	218	24,5	2	43	6	0	1	1	1	1
11	376	Ahven	roovkala	Võrtsjärv	kevad-suvi	i	111	22	4	33	11	0	1	1	1	1
12	375	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	132	23	0	0	0	0	0	0	0	0
13	360	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	351	31	22	2	33	0	1	1	1	1
14	347	Ahven	roovkala	Praaga	kevad-suvi	i	70	19	14	3	49	0	1	1	1	1
15	344	Ahven	roovkala	Praaga	kevad-suvi	e	61	18	30	19	38	0	1	1	1	1
16	343	Ahven	roovkala	Praaga	kevad-suvi	i	93	20	14	1	24	0	1	1	1	1
17	310	Ahven	roovkala	Peipsi	kevad-suvi	e	87	19	18	11	16	0	1	1	1	1
18	755	Ahven	roovkala	Peipsi	kevad-suvi	i	133	22	24	36	22	0	1	1	1	1
19	754	Ahven	roovkala	Peipsi	kevad-suvi	i	194	23,5	0	36	11	0	0	1	1	1
20	753	Ahven	roovkala	Peipsi	kevad-suvi	i	192	23,5	2	9	6	0	1	1	1	1
21	521	Ahven	roovkala	Võrtsjärv	sygis-talv	i	229	24	0	31	8	0	0	1	1	1
22	487	Ahven	roovkala	Kastre	sygis-talv	e	45	15	2	10	2	0	1	1	1	1
23	373	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	141	21	0	30	13	0	0	1	1	1
24	371	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	248	27	4	116	13	0	1	1	1	1
25	361	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	250	27,5	10	18	47	0	1	1	1	1
26	342	Ahven	roovkala	Praaga	kevad-suvi	i	43	15,5	18	84	17	0	1	1	1	1
27	341	Ahven	roovkala	Praaga	kevad-suvi	i	58	18	26	1	5	0	1	1	1	1
28	339	Ahven	roovkala	Praaga	kevad-suvi	i	70	18,5	28	38	17	0	1	1	1	1
29	334	Ahven	roovkala	Praaga	kevad-suvi	e	123	22	28	0	28	0	1	0	1	1
30	329	Ahven	roovkala	Kastre	kevad-suvi	e	76	18	50	1	12	0	1	1	1	1
31	319	Ahven	roovkala	Peipsi	kevad-suvi	e	80	18,5	72	0	4	0	1	0	1	1
32	318	Ahven	roovkala	Peipsi	kevad-suvi	e	77	18,5	4	0	19	0	1	0	1	1
33	316	Ahven	roovkala	Peipsi	kevad-suvi	e	67	18	4	26	6	0	1	1	1	1
34	313	Ahven	roovkala	Peipsi	kevad-suvi	e	119	22,5	2	26	42	0	1	1	1	1
35	309	Ahven	roovkala	Peipsi	kevad-suvi	e	123	21	2	0	16	0	1	0	1	1
36	306	Ahven	roovkala	Peipsi	kevad-suvi	e	115	21	14	0	15	0	1	0	1	1
37	244	Ahven	roovkala	Praaga	sygis-talv	e	34	14	38	4	5	0	1	1	1	1
38	67	Ahven	roovkala	Praaga	kevad-suvi	e	167	25	28	3	1	0	1	1	1	1
39	1035	Ahven	roovkala	Kastre	sygis-talv	i	74	18	0	12	3	0	0	1	1	1
40	834	Ahven	roovkala	Kastre	kevad-suvi	i	306	29	14	8	6	0	1	1	1	1
41	532	Ahven	roovkala	Võrtsjärv	sygis-talv	e	62	17	0	19	4	0	0	1	1	1
42	524	Ahven	roovkala	Võrtsjärv	sygis-talv	e	211	25,5	0	60	7	0	0	1	1	1
43	523	Ahven	roovkala	Võrtsjärv	sygis-talv	e	238	26	0	75	17	0	0	1	1	1
44	457	Ahven	roovkala	Kastre	sygis-talv	i	60	16	18	9	0	0	1	1	0	0
45	424	Ahven	roovkala	Võrtsjärv	sygis-talv	i	56	17	2	11	14	0	1	1	1	1
46	423	Ahven	roovkala	Võrtsjärv	sygis-talv	e	55	17	0	27	2	0	1	1	1	1
47	414	Ahven	roovkala	Võrtsjärv	sygis-talv	i	76	18	6	22	9	0	1	1	1	1
48	411	Ahven	roovkala	Võrtsjärv	sygis-talv	i	73	18	0	65	4	0	1	1	1	1
49	367	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	119	21,5	4	13	14	0	1	1	1	1
50	363	Ahven	roovkala	Võrtsjärv	kevad-suvi	e	159	23,5	0	64	9	0	1	1	1	1

OSA 2

--- Sagedustabelid ---

❖ Sagedustabel mittearvulisele või diskreetsele arvtunnusele

1) Leidke hariliku laiussi noorvormiga nakatunud ja mitte nakatunud haugide arvud ja osakaalud.

--- Tööjuhend ---

The image illustrates the process of creating a PivotTable and formatting it to show percentages. The steps are as follows:

- Initial Data:** A table with columns 'liik' (0, 1) and 'Grand Total'. The '0' column contains values 3, 28, 111, 59, 90. The '1' column contains values 179, 214, 173, 137, 228. The 'Grand Total' row contains values 1042, 214, 173, 137, 1042.
- PivotTable Wizard:** The 'Count of Diphylob' PivotTable is created with 'liik' as the row labels and 'Diphylob' as the column labels.
- PivotTable Field List:** The 'Diphylob' field is added to the report.
- Field Settings:** The 'Field Settings' task pane is opened for the 'Diphylob' field. The 'Summarize by' is set to 'Count', and 'Show data as' is set to '% of row'.
- Final PivotTable:** The resulting PivotTable shows the data as percentages of the row totals.

liik	Data	0	1	Grand Total
Haug	Count of Diphylob	186	28	214
	Count of Diphylob2	86,92%	13,08%	100,00%
Total Count of Diphylob		186	28	214
Total Count of Diphylob2		86,92%	13,08%	100,00%

❖ **Sagedustabel arvtunnusele**

1) Konstrueerige sagedustabel haugide pikkusele.

--- Tööjuhend ---

- Pidevale arvtunnusele sagedustabeli tegemiseks on *MS Excelis* kaks moodust – funktsioon *FREQUENCY* ja statistikaprotseduur *Histogram*. Mõlemad need variandid eeldavad, et kasutaja on eelnevalt välja mõelnud loodavad klassid ja sisestanud klasside ülemised piirid *Exceli* töölehele.
- Otsustamiseks klasside arvu ja suuruse üle, peab esmalt omama ülevaadet vaatluste arvust ja uuritava tunnuse väärtuste ulatusest.

Seega – **leidke** mistahes viisil **haugide arv** ning nende **vähim** ja **suurim pikkus** (soovi korral võib haugide kohta käivaist andmeist teha koopia eraldi töölehele).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	kala_nr	liik	ryhm	pyygikoht	sesoon	sugu	kaal	pikkus	arv_lõpuse	arv_silmapar	arv_soolepar	arv_Diphyllo	lõpusepar	silmapar	soolepar	Diphyllob	par_per_kala
2	291	Haug	roovkala	Peipsi	kevad-suvi	e	1071	52	22	39	2	4	1	1	1	1	6
3	380	Haug	roovkala	Võrtsjärv	kevad-suvi	i	602	45	50	4	5	0	1	1	1	0	6
4	230	Haug	roovkala	Praaga	sygis-talv	i	1108	55	52	18	3	0	1	1	1	0	5
5	179	Haug	roovkala	Praaga	sygis-talv	i	775	50	16	4	36	2	1	1	1	1	5
6	555	Haug	roovkala	Peipsi	sygis-talv	i	555	45,5	30	4	2	0	1	1	1	0	5
7	477	Haug	roovkala	Kärevere	sygis-talv	e	883	51,5	20	1	13	0	1	1	1	0	5
8	351	Haug	roovkala	Võrtsjärv	kevad-suvi	i	541	46	32	0	1	0	1	0	1	0	3
9	281	Haug	roovkala	Peipsi	kevad-suvi	e	1163	55	28	3	6	0	1	1	1	0	5
10	280	Haug	roovkala	Peipsi	kevad-suvi	i	717	49	46	4	3	0	1	1	1	0	6
11	198	Haug	roovkala	Kärevere	sygis-talv	i	819	50	26	0	5	0	1	0	1	0	4

- Vaatluste arv 214 annab vihje, et haugid võiks pikkuse järgi jagada $\sqrt{214} \approx 15$ klassi (see ei ole absoluutne tõde, kui klasside piiridega ja/või sisuliste järeldustega paremini sobib, võib väärtused jagada ka väiksemasse arvu klassidesse).
- Teadmise, et 214 haugi pikkused varieeruvad 26-st sentimeetrist 74 sentimeetrini, alusel võiks moodustada 13 4-sentimeetrist klassi: 25-28; 29-32; 33-36; ...; 69-72; 73-76.

NB! Toodud tähistus sobib üksnes täisarvuliste väärtuste korral (miks?), matemaatiliselt korrektne on esitada klassid poollõikudena: (24,28], (28,32], (32,36], ..., (68,72] ja (72,76].

- **Moodustatud klasside *Excelile* ette andmiseks tuleb sisestada klasside ülemisi piire sisaldav abitabel:**
- Märkusi:
 - mistahes abitabeli ja algse andmetabeli vahele on soovitatav jätta vähemalt üks tühi veerg (või rida) – **miks?** (kui ei tea, küsi!);
 - *Excel* tõlgendab ette antud väärtusi järgmiselt:
 - '28' ↔ ' ≤ 28 ';
 - '32' ↔ ' > 28 ja ≤ 32 '
 - jne (ehk igasse klassi kuuluvaks loetakse need väärtused, mis on väiksemad või võrdsed ette antud piirist ja mis ei kuulu eelnevatesse klassidesse);
 - viimast klassi ' 76 '=(72,76] ette ei anta, sest sagedustabeli moodustamisel teeb *Excel* ise täiendava klassi, kuhu loeb kokku kõik eelnevatesse klassidesse mitte kuuluvad väärtused.

T
pikkus_klassid
28
32
36
40
44
48
52
56
60
64
68
72

▪ **Funktsioon FREQUENCY**

Kõige kiirem variant lasta Excelil kokku lugeda, kui palju vaatlusi mingisse ette antud klassi kuulub, on kasutada funktsiooni FREQUENCY (kui statistikaprotseduuride pakett Data Analysis ei ole teie arvutis installeeritud, aitab just see funktsioon hädast välja).

Erinevalt enamusest MS Exceli funktsioonidest on funktsioon FREQUENCY **massiivifunktsioon**, st et selle funktsiooni tulemuseks ei puugi olla üks väärtus eelnevalt valitud lahtris, vaid hulk väärtusi eelnevalt valitud lahtriteblokkis.

- Esimese asjana peale klassipiiride Exceli töölehele sisestamist tuleb võtta blokki lahtrid töölehel kohas, kuhu soovitakse sagedusi arvutada; arvutatavate sageduste ja seeläbi blokki võetavate lahtrite arv on määratud konstrueeritava sagedustabeli klasside arvuga (üks täiendav blokki võetud lahter vastab Exceli poolt täiendavalt moodustatavale klassile).
- Trükite selekteeritud lahtriblokki (koheselt sellesse lahtrisse, millest blokki võtmist alustasite, uuesti klikkida esimesel lahtril ei tohi!!)

=FREQUENCY(H2:H215;T2:T13)

misjärel vajutate alla klahvid 'Shift' ja 'Ctrl' ning seejärel 'Enter' (st. 3 klahvi korruga).

- Alternatiiv taolisele funktsiooni klaviatuurilt sisestamisele on lisada funktsioon menüüsid ja abiaknaid kasutades:

The screenshot shows an Excel spreadsheet with the following data:

pikkus_klassi	Sagedused
28	
32	
36	
40	
44	
48	
52	
56	
60	
64	
68	
72	

The 'Function Arguments' dialog box shows:

- Function: FREQUENCY
- Data_array: H2:H215
- Bins_array: T2:T13

The 'Insert Function' dialog box shows 'FREQUENCY' selected under the 'Statistical' category.

▪ **Protseduur Histogram**

Tools / Tööriistad → Data Analysis... → Histogram

Ütleb Excelile, et ette antud lahtri-
blokkide esimeses reas on nimed

Väljundtabeli (vasaku ülemise nurga)
asukoht

Lisaks tavalistele sagedustele
arvutatakse ka kumulatiivsed
suhtelised sagedused e. jaotus

pikkus_klassid	Frequency	Cumulative %
28	1	0,47%
32	0	0,47%
36	2	1,40%
40	5	3,74%
44	37	21,03%
48	62	50,00%
52	68	81,78%
56	25	93,46%
60	7	96,73%
64	3	98,13%
68	1	98,60%
72	2	99,53%
More	1	100,00%

NB! Lahtri 'Bin Range' võib jätta ka tühjaks
– siis moodustab Excel klassid ise.

Proovige!

2) Arvutage lisaks absoluutsetele sagedustele ka suhtelised sagedused (%-des) ja illustreerige sagedustabelit histogrammiga (ükskõik, kas funktsiooni *FREQUENCY* või protseduuri *Histogram* abil konstrueeritud sagedustabeli baasil).

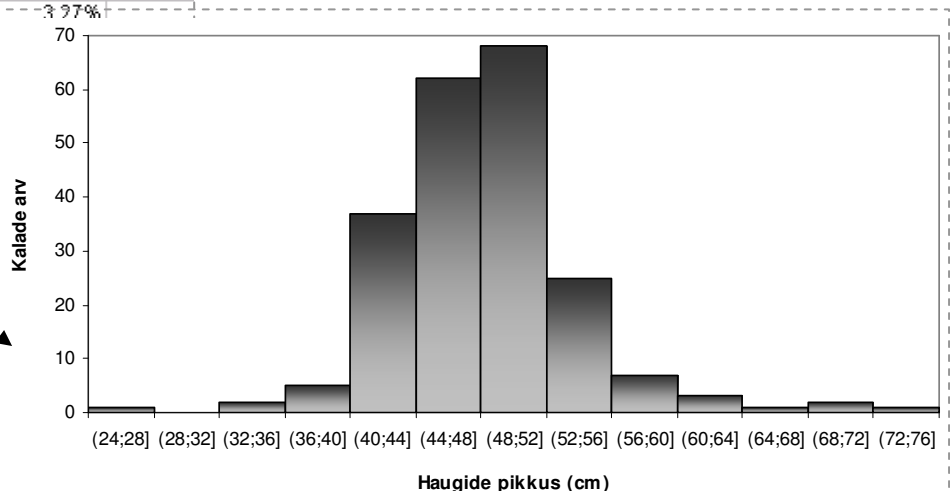
T	U	V	W
pikkus_klassid	Sagedused	Suhtelised sagedused	
28	1	=U2/SUM(U\$2:U\$14)	
32	0		
36	2		
40	5		
44	37		
48	62		
52	68		
56	25		
60	7		
64	3		
68	1		
72	2		
	1		

Suhtelised sagedused
0,0046729
0
0,00934579
0,02336449
0,1728972
0,28971963
0,31775701
0,11682243
0,03271028
0,01401869
0,0046729
0,00934579
0,0046729

Suhtelised sagedused
0,47%
0,00%
0,93%
2,34%
17,29%
28,97%
31,78%
11,68%
3,27%
1,40%
0,47%
0,93%
0,47%

Pikkused	pikkus_klassid	Sagedused	Suhtelised sagedused
(24;28]	28	1	0,47%
(28;32]	32	0	0,00%
(32;36]	36	2	0,93%
(36;40]	40	5	2,34%
(40;44]	44	37	17,29%
(44;48]	48	62	28,97%
(48;52]	52	68	31,78%
(52;56]	56	25	11,68%
(56;60]	60	7	3,27%
(60;64]	64	3	1,40%
(64;68]	68	1	0,47%
(68;72]	72	2	0,93%
(72;76]		1	0,47%

NB! Eraldi paiknevate lahtrite selekteerimiseks hoidke all 'Crlt'-klahvi ...

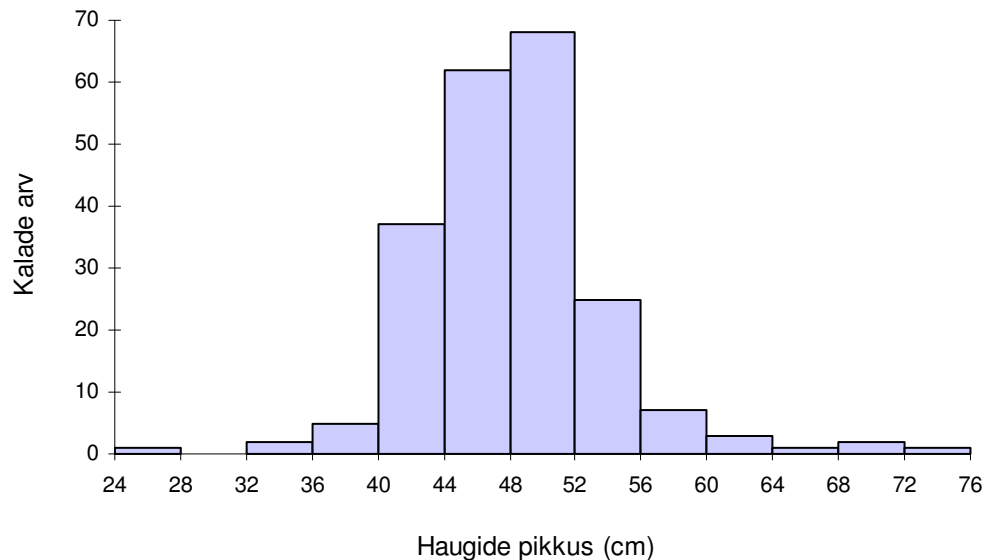


OSA 3

--- Pisut joonistest ---

❖ Histogramm

Eelmisel lehel esitatud histogramm matemaatiliselt korrektselt kujul (pideva arvteljena esitatud x-teljega) võiks olla järgmine:



Juhendi taolise pideva arvteljena esitatud x-teljega histogrammi tarvis leiate Internetist aadressilt http://www.treeplan.com/BetterHistogram_20041117_1555.pdf, samalt lehelt (<http://www.treeplan.com/better.htm>) on alla laaditav ka tasuta lisamoodul *Excelile* taolise joonise tegemiseks.

❖ Karp-vurrud diagramm

Karp-vurrud diagrammi (vt tulemust järgmise lehekülje lõpus) joonistamiseks *Excelis* tuleb esmalt teha abitabel ja arvutada sinna välja diagrammi aluseks olevad suurused – mediaan, alumine ja ülemine kvartiil ning minimaalne ja maksimaalne väärtus.

		Mediaan	25%-punkt	75%-punkt	Min	Max
Röövkalad	Ahven	20	18	23	14	32
	Koha	41	38	43,5	22	54
	Luts	46	41	51	27	70
	Haug	48,25	45	51,5	26	74
Lepsiskalad	Särg	22	20,5	24	11	45
	Latikas	40	37	42	27	48

Karp-vurrud diagrammile on aluseks tavaline tulpdiaagramm, mis on joonistatud eelnevas tabelis toodud karakteristikute baasil arvatud väärtuste põhjal:

		25%-punkt	Med - 25%-p	75%-p - med	25%-p - min	Max - 75%-p
	Series1	Series2	Series3	Error1	Error1	
Röövkalad	Ahven	18	2	3	4	9
	Koha	38	3	2,5	16	10,5
	Luts	41	5	5	14	19
	Haug	45	3,25	3,25	19	22,5
Lepsiskalad	Särg	20,5	1,5	2	9,5	21
	Latikas	37	3	2	10	6

