

3. Statistik



3.1 Ikke-grupperede observationer

1 Introduktion

I en løbeklub tog man en stikprøve, hvor 30 løberes puls blev målt lige efter en træningstur:

147, 156, 176, 157, 155, 167, 138, 167, 176, 159, 165, 181, 148, 169, 156, 167, 165, 147, 154, 153, 172, 132, 163, 170, 152, 160, 153, 174, 148, 175.

Det er svært at se det overordnede mønster i tallene. Statistik handler blandt andet om at beskrive sådanne tal på en overskuelig måde.

2 Definition

En **stikprøve** består af nogle elementer udvalgt fra en **population**. Populationen er den mængde, der er genstand for undersøgelsen. Samlingen af observationer i stikprøven kaldes et **datasæt** eller **observationssæt**. I et **ordnet observationssæt** er observationerne ordnet i rækkefølge.

3 Eksempel

- I en klasse lavede man et skema med elevernes skostørrelser. Hvis man opskriver skostørrelserne i ordnet rækkefølge 35, 35, 36, 37, 37, 38 ... osv., får man et ordnet observationssæt.
- Løbeklubbens målinger i stikprøven på 30 løbere er et datasæt med enkeltstående tal, der ikke er ordnet i rækkefølge.

4 Definition

I et **ikke-grupperet observationssæt** beskrives observationerne med enkeltstående tal, farver, rejsemål eller lignende. Ikke-grupperede observationer kan beskrives med forskellige tal, såkaldte **statistiske deskriptorer**. Den simpleste er **observationssættets størrelse, N** . Hvis alle observationerne er tal, er **variationsbredden** lig med forskellen på værdien af den største og mindste observation.

5 Eksempel

Blandt 25 elever undersøgte man elevernes antal søskende.

Svarene var: 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4. Forskellen på det største og det mindste tal er $4 - 0 = 4$, så variationsbredden er altså 4. Observationssættet har størrelsen $N = 25$.

6 Definition

Hypigheden $h(x)$ af en observation x er det antal gange, som observationen forekommer i observationssættet, og **frekvensen** $f(x)$ er den brøkdelt eller procentdel, hypigheden udgør af hele observationssættets størrelse. Den **kumulerede hypighed** $H(x)$ og den **kumulerede frekvens** $F(x)$ af en observation x er summen af hypighederne eller frekvenserne af alle de observationer, der er mindre eller lig med x .

7 Eksempel

Observationssættet med antal søskende fra eksempel 5 kan skrives op i en **hyppighedstabel**, hvori vi har beregnet de ovennævnte deskriptorer:



Observation	Hyppighed	Kumuleret hyppighed	Frekvens	Kumuleret frekvens
0	4	4	$4/25 = 0,16 = 16\%$	16 %
1	8	12	32 %	48 %
2	6	18	24 %	72 %
3	4	22	16 %	88 %
4	3	25	12 %	100 %

8 Definition

Et observationssæts **typetal** er den observation, der optræder flest gange. **Gennemsnittet** eller **middelværdien**, \bar{x} , er summen af observationerne divideret med observationssættets størrelse, N .

9 Eksempel

Typetallet for observationssættet: 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4 er lig med 1, fordi tallet 1 er det tal med den største hyppighed i datasættet.

Gennemsnittet er

$$\bar{x} = \frac{0+0+0+0+1+1+1+1+1+1+1+1+2+2+2+2+2+2+3+3+3+3+4+4+4}{25} = \frac{44}{25} = 1,76$$

10 Øvelse

- Opstil observationerne fra stikprøven i eksempel 1 i rækkefølge (omform til et ordnet observationssæt).
- Beregn variationsbredden og gennemsnittet.

11 Øvelse

- Slå op på 20 forskellige sider i denne bog og tæl antallet af billeder på siden.
- Stil observationerne op på en række i et ordnet observationssæt.
- Opstil en tabel med observationer, hyppigheder, frekvenser, de kumulerede hyppigheder og de kumulerede frekvenser.

12 Øvelse

- Opstil en tabel med observationer, hyppigheder, frekvenser, de kumulerede hyppigheder og de kumulerede frekvenser for observationssættet: 5, 5, 6, 7, 7, 7, 7, 8, 8, 9.

13 Øvelse

- Bestem typetallet og gennemsnittet for observationssættet: 5, 5, 6, 7, 7, 7, 7, 8, 8, 9.



3.2 Diagrammer og kvartilsæt

14 Introduktion

Et malet portræt viser noget væsentligt om en person, men er mindre detaljeret end et fotografi. På samme måde viser statistiske diagrammer væsentlige træk ved et observationssæt uden at give alle detaljer.

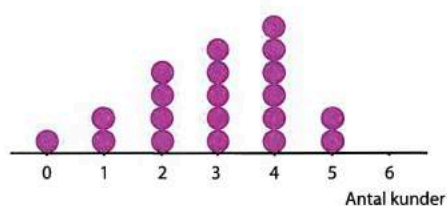
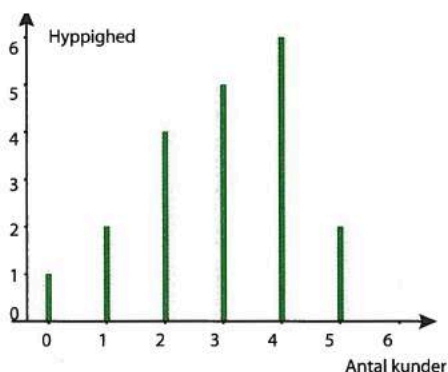
15 Definition

I et **pindediagram** vises observationernes hyppighed eller frekvens som længder af pinde. I et **prikdiagram** (eller punktdiagram) vises observationernes hyppigheder med prikker.

16 Eksempel

En lille tebutik ønskede at få et overblik over, hvor mange kunder der kom i den første time af åbningstiden. De valgte at registrere antallet af kunder i den første time i 20 dage. Resultaterne fremgår af tabellen. Nederst på siden er hyppigheden vist med et pindediagram og et prikdiagram.

Observation	Hyppighed $h(x)$	Kumuleret hyppighed $H(x)$	Frekvens $f(x)$	Kumuleret frekvens $F(x)$
0	1	1	0,05	0,05
1	2	3	0,10	0,15
2	4	7	0,20	0,35
3	5	12	0,25	0,60
4	6	18	0,30	0,90
5	2	20	0,10	1,00



17 Kvartilsæt for ikke-grupperede observationer

For **ikke-grupperede observationer** i et ordnet observationssæt defineres kvartiler således:

Medianen, M , eller **anden kvartil**, Q_2 , er den **midterste observation**. Ved et lige antal observationer vil der ikke være en midterste observation. Så tager man gennemsnittet af de to midterste.

Første kvartil, Q_1 , er medianen af de observationer, der står til venstre for M .

Tredje kvartil, Q_3 , er medianen af de observationer, der står til højre for M .

Tilsammen udgør de tre tal **kvartilsættet** (Q_1, Q_2, Q_3).

Kvartilafstanden er $Q_3 - Q_1$ (kaldes også kvartilbredden).

I regneark og programmer som Geogebra kan du også udregne kvartilsættet. Tallene kan afvige lidt fra dem, der fås manuelt, fordi beregningsmetoden kan variere en lille smule mellem programmerne. Endvidere kan et kvartilsæt aflæses på diagrammer.



18 Eksempel

Herunder ses højden i meter på et udvalg af 6 høje egetræer.

23,4 24,1 28,0 31,7 31,9 33,8

Medianen er gennemsnittet af de **to midterste** grå tal $M = \frac{28,0 + 31,7}{2} = 29,85$ og de andre kvartiler er lig med de blå tal $Q_1 = 24,1$ og $Q_3 = 31,9$

19 Øvelse

Herunder ses højden i meter for en gruppe på 9 høje træer.

23,4 24,1 28,0 31,7 31,9 32,9 33,1 33,8 37,2

a. Opskriv kvartilsættet for dette nye datasæt.



20 Øvelse

a. Tegn et pindediagram og et trappediagram for besøgende om formiddagen på en kaffebar. Observationerne for 100 dage fremgår af tabellen:

Antal kunder, x	0	1	2	3	4	5	6
Hyppighed, $h(x)$	5	5	20	30	20	10	10

21 Øvelse

5 gamle fjender møder hinanden hver onsdag i kiosken, hvor de kan skule til hinanden, mens de spiller lotto. En uge ser gevinsterne i kr. således ud: 0, 0, 0, 43, 1 000 000.

a. Bestem medianen.

b. Beregn gennemsnittet og forklar, hvad forskellen er på de to begreber median og gennemsnit.



3.3 Bokplot

22 Introduktion

En elev vil gerne skabe sig et overblik over de karakterer, hun fik ved sin hf-eksamen, det kan hun gøre med et bokplot.

23 Definition

De fem værdier: mindste værdi, første kvartil, median, tredje kvartil og største værdi kaldes det **udvidede kvartilsæt**.

Et **bokplot** er en grafisk fremstilling af de fem værdier.

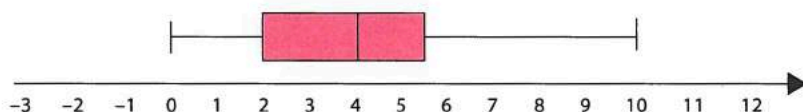
24 Eksempel

En student fik eksamenskaraktererne: 00, 02, 02, 02, 02, 02, 4, 4, 4, 4, 4, 4, 4, 7, 7, 10, 10. Det udvidede kvartilsæt er (0, 2, 4, 5,5, 10).

25 Eksempel

Sådan kan et bokplot tegnes: Tegn en tallinje, hvor værdierne af observationerne er afsat. Sæt 5 lodrette streger ovenover: ud for (1) mindste værdi, (2) nedre kvartil, (3) median, (4) øvre kvartil og (5) største værdi. Tegn en boks rundt om selve kvartilsættet og en streg ud fra enderne af boksen til mindste og største værdi. Boksens højde har ingen betydning. Mange matematikprogrammer kan også tegne bokplot.

Studenten fra eksempel 24 tegnede dette bokplot over karaktererne i studentehuen:



Sammenligning af to datasæt

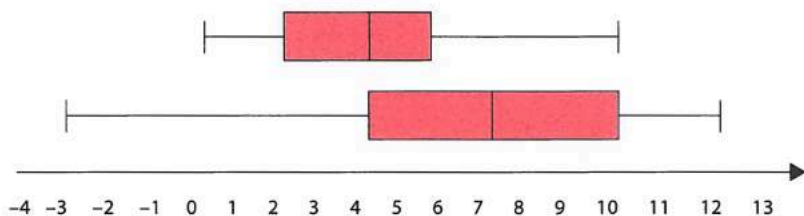
fx for to elevers karakterer, sker ved at tegne de to bokplot over den samme værdiakse, sådan som det er vist på næste side. Man kan derefter fx sammenligne **niveauforskelle** ved at sammenligne medianerne. Man kan sammenligne **spredningsforskelle** ved at sammenligne variationsbredderne eller sammenligne kvartilafstandene ($Q_3 - Q_1$) altså boksens bredde.

Kender man yderligere middelværdien af de enkelte datasæt, kan man også undersøge **symmetriforskelle**: højre-/venstreskævhed (middelværdiens placering i forhold til medianen). Ligger middelværdien fx til højre for medianen, kaldes datasættet **højreskævt**.

En **outlier** er en observation, der **ligger mere end 1,5 gange kvartilbredden** væk fra boksen.

26 Eksempel

To elever registrerede deres karakterer efter 1.g. Resultaterne blev illustreret med boksploot.



- Sammenligner man boksploottene herover, ser man først, at der er en niveauforskel, idet medianen på det øverste er 4, mens den er 7 i det nederste. Niveaueet er altså en del højere på det nederste.
- Kvartilafstanden er $5,5 - 2 = 3,5$ på det øverste, mens det er $10 - 4 = 6$ på det nederste, der dermed har en større variation end det øverste.
- Dette understøttes også ved at se, at variationsbredden er 10 på det øverste, men 15 på det nederste.
- Middelværdien i det øverste datasæt kan ud fra tallene i eksempel 24 udregnes til 4,24.

Da gennemsnittet er større end 4 og derved ligger til højre for medianen på 4, er dette datasæt højreskævt, det kan måske også anes ved at bemærke den store afstand fra medianen, 4, til størsteværdien, 10. Vi kender ikke middelværdien af det nederste datasæt og kan derfor ikke sammenligne de to sæt med hensyn til skævhed.

27 Øvelse

En dyrlæge udarbejder sin egen lille kvantitative undersøgelse af granddanoishundes levealder (målt i antal år). Registreringerne i notesbogen er:

5, 6, 4, 8, 7, 6, 9, 8, 9, 7, 7, 5, 9, 11, 8, 12, 6, 8

- Beregn det udvidede kvartilsæt.
- Tegn boksplottet.

28 Øvelse

Dyrlægen i øvelse 27 fik blod på tanden og besluttede sig for også at undersøge levealderen hos de gadekryds, der havde været tilknyttet lægeklinikken.

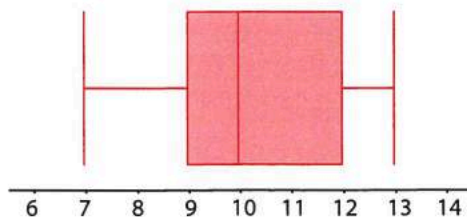
Her var det udvidede kvartilsæt (1, 9, 12, 14, 17).

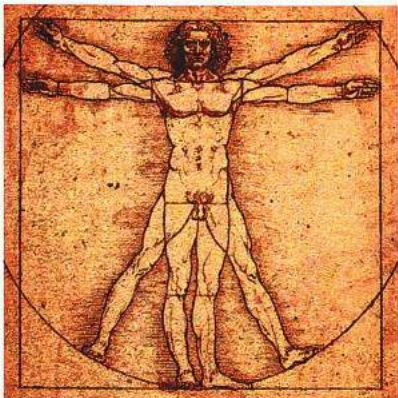
- Tegn boksplottet og sammenlign de to undersøgelser.



29 Øvelse

- Aflæs største- og mindsteværdi.
- Aflæs kvartilsættet.
- Hvad fortæller medianen om datasættet?





3.4 Grupperede observationer

30 Introduktion

I bogen "Da Vinci mysteriet" af Dan Brown påstår hovedpersonen, at højden af et menneske divideret med afstanden fra navlen til gulvet er lig tallet 1,618 (det gyldne snit). For at undersøge det kan man jo måle efter på nogle tilfældigt valgte mennesker. Hvis der er mange mennesker med i undersøgelsen, kan man med fordel samle resultaterne i grupper for at få et bedre overblik.

31 Definition

I et **grupperet observationssæt** samles observationerne i intervaller. Grupperede observationssæt bruges fortrinsvis til store datasæt med mange observationer, der er tal. Ligger en observation netop på grænsen mellem to intervaller, tælles observationen med i det nærmeste lavere interval. **Intervalhyppigheden** er antallet af observationer i intervallet, **intervalfrekvensen** er den brøkdelt eller procentdel, som intervalhyppigheden udgør af observationssættets størrelse. **Den kumulerede intervalhyppighed og den kumulerede intervalfrekvens** er summen af intervalhyppighederne eller frekvenserne i intervaller op til og med det aktuelle interval.

Højde
i cm

166
159
173
173
170
192
158
183
188
175
184
171
174
177
180
185
170
169
159
182
176
163

32 Eksempel

En klasse på en ungdomsuddannelse ville undersøge menneskekroppens proportioner og startede med at måle hinandens højder. Resultatet ses i tabellen til venstre. Derefter samlede de observationerne i intervaller (åbne til venstre og lukkede til højre) og beregnede værdien af deskriptorerne intervalhyppighed, kumuleret intervalhyppighed, intervalfrekvens og kumuleret intervalfrekvens for de enkelte intervaller som vist i tabellen herunder:

Højde-interval; cm	Intervalhyppighed	Kumuleret intervalhyppighed	Intervalfrekvens	Kumuleret intervalfrekvens
]155, 165]	4	4	$\frac{4}{22} = 0,18$	0,18
]165, 175]	9	13	$\frac{9}{22} = 0,41$	0,59
]175, 180]	7	20	$\frac{7}{22} = 0,32$	0,91
]185, 195]	2	22	$\frac{2}{22} = 0,09$	1,00
Antal i alt	22		1,00	

33 Definition

Et grupperet observationssæts **typeinterval** er det interval, der har flest observationer. **Gennemsnittet** eller **middelværdien**, \bar{x} , beregnes ved at gange midtpunktet af hvert interval med intervalhyppigheden, lægge alle resultaterne sammen og dividere det hele med observationssættets størrelse. Alternativt kan man beregne middelværdien ved at gange hvert intervalmidtpunkt med den tilsvarende intervalfrekvens og lægge resultaterne sammen.

34 Eksempel

I tabellen i eksempel 32 er typeintervallet]165, 175]. Den gennemsnitlige højde

$$\text{beregnes til } \bar{x} = \frac{160 \cdot 4 + 170 \cdot 9 + 180 \cdot 7 + 190 \cdot 2}{22} = 173,2$$

35 Øvelse

På et gymnasium fordelte den samlede gennemsnitlige månedlige indtægt blandt de 692 elever sig således:

Månedssindtægt i kr.]0;2000]]2000;4000]]4000;6000]]6000;8000]
Antal (hyppighed)	384	207	78	23

- Bestem typeintervallet.
- Beregn middelværdien.

36 Øvelse

22 elever målte afstanden mellem navle og gulv og fik resultaterne (målt i cm):

100, 95, 104, 104, 100, 116, 97, 114, 116, 105, 107, 103, 109, 108, 114, 116, 106, 102, 95, 109, 108, 94

- Grupper observationerne i passende intervaller og opstil en hyppighedstabel.
- Bestem typeintervallet og beregn middelværdien.

37 Beregning af gennemsnit, grupperet stikprøve

Gennemsnittet \bar{x} af tallene i en **grupperet stikprøve** med intervalmidtpunkter m_1, m_2, m_3, \dots og de tilsvarende intervalhyppigheder h_1, h_2, h_3 eller intervalfrekvenser f_1, f_2, f_3 kan beregnes med formlen:

$$\bar{x} = \frac{m_1 \cdot h_1 + m_2 \cdot h_2 + \dots}{N}, \text{ eller } \bar{x} = m_1 \cdot f_1 + m_2 \cdot f_2 + \dots, \text{ hvor } N \text{ er observationssættets samlede størrelse, og hvor alle intervalmidtpunkter og deres hyppigheder eller frekvenser regnes med.}$$



38 Beregning af gennemsnit, ikke-grupperet stikprøve

Gennemsnittet \bar{x} af tallene i en **ikke-grupperet stikprøve** med hyppigheder $h(x)$ og samlet størrelse N kan beregnes med formlen:

$$\bar{x} = \frac{x_1 \cdot h(x_1) + x_2 \cdot h(x_2) + \dots}{N}, \text{ hvor } h(x_i) \text{ er hyppigheden af observation } x_i \text{ osv.}$$

39 Øvelse

- Hyppighedstabellen viser hvilepuls hos 50 unge. Brug formlen i sætning 37 til at beregne gennemsnits-hvilepuls.

Hvilepuls-interval	Intervalhyppighed	Kumuleret intervalhyppighed	Intervalfrekvens	Kumuleret intervalfrekvens
]40, 60]	21	21	0,42	0,42
]60, 80]	18	39	0,36	0,78
]80, 100]	11	50	0,22	1,00
Antal i alt	50		1,00	



3.5 Diagrammer for grupperede observationer

40 Introduktion

I en hyggelig lille dansk by var der et år 100 fødsler.

Hvis mødrenes alder samles i intervaller på 5 år, vil man ikke kunne beregne median og kvartilsæt på samme måde som i et ugrupperet datasæt.

I stedet for kan vi bruge den såkaldte sumkurve til at aflæse kvartilerne.

Sumkurven er en af de grafer, vi ser nærmere på her.

41 Definition

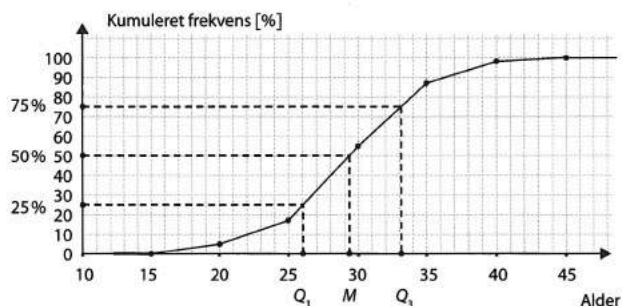
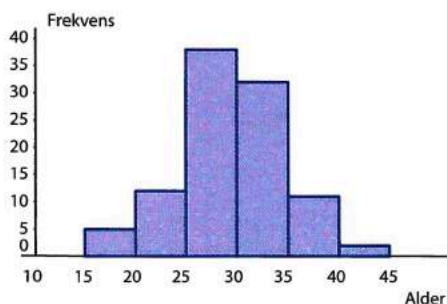
I et **histogram** vises observationernes hyppighed eller frekvens som arealet af søjler, der er tegnet med samme bredde som intervallerne. Hvis intervallerne har samme bredde, viser søjlernes højde observationernes hyppighed/frekvens i det pågældende interval.

De kumulerede hyppigheder eller frekvenser kan afbildes i en **sumkurve** med punkter, hvor **x-koordinaten er intervallerne højre endepunkt**, og **y-koordinaten er intervallerne kumulerede frekvens**. Den højeste værdi på y-aksen er altså 100%. Punkterne forbindes med rette linjestykker.

42 Eksempel

Herunder ses en grupperet opgørelse over 100 mødres alder med tilsvarende histogram og sumkurve.

Alder]15;20]]20;25]]25;30]]30;35]]35;40]]40;45]
Frekvens	5%	12%	38%	32%	11%	2%
Kumuleret frekvens	5%	17%	55%	87%	98%	100%



Kvartilsæt for grupperede observationer

43 Flere definitioner (deskriptorer for grupperede datasæt)

Kvartilerne aflæses på x -aksen i en sumkurve ud fra y -værdierne 25%, 50% og 75%

- **Første kvartil, Q_1** , er det tal, der skiller de mindste 25% af observationerne fra resten.
- **Medianen, M** , eller andet **kvartil, Q_2** , er det tal, der skiller de mindste 50% af observationerne fra de største 50%.
- **Tredje kvartil, Q_3** , er det tal, der skiller de mindste 75% af observationerne fra resten.
- **Kvartilsættet** er talsættet ($Q_1; M; Q_3$)
- **Kvartilafstanden** er $Q_3 - Q_1$ (kaldes også kvartilbredden).

44 Eksempel

I eksempel 42 på forrige side aflæses kvartilsættet på x -aksen i sumkurven ud fra y -værdierne 25%, 50% og 75%. Vi får: $Q_1 = 26$, $M = 29,3$ og $Q_3 = 33$.



45 Øvelse

I et boligkvarter i København fordelte de barslende kvinder sig således mht. alder:

Alder (år)]15;20]]20;25]]25;30]]30;35]]35;40]]40;45]
Frekvens	3%	11%	35%	28%	19%	4%
Kumuleret frekvens	3%	14%	49%	77%	96%	100%

a. Tegn et histogram over frekvenserne og en sumkurve over de kumulerede frekvenser.

46 Øvelse

- Aflæs kvartilsættet på sumkurven fra øvelse 45.
- Sammenlign kvartilerne med kvartilerne i eksempel 42. Hvad siger det om forskellen på storbyen og den lille by?

47 Øvelse

En pakkecentral registrerede vægten i kg af 50 pakker:

7	19	24	38	34	11	5	12	14	33
4	28	4	18	13	14	3	9	6	16
31	2	32	17	7	21	19	23	20	5
26	9	22	36	29	10	4	13	14	22
21	15	8	21	17	12	4	10	7	16

- Grupper observationerne i fire grupper og opstil en hyppighedstabel.
- Tegn sumkurven.
- Aflæs kvartilsættet på sumkurven.
- Bestem den vægt, der afgrænser de letteste 75% fra de tungeste 25%.



Opgave 301

Hans opdrætter høns. Han har en hane, ni høns, der ruger på æg, og en enkelt høne, som er pensioneret. Antallet af æg i rederne en dag i maj er 0, 3, 4, 5, 4, 4, 4, 4, 3, 4, 4



- a. Hvilke tal skal der stå i den nederste række i skemaet?

Antal æg	0	3	4	5
Hyppeghed				

- b. Hvad er antallet af observationer (observationssættets størrelse)?
 c. Redegør for, at det er et ikke-grupperet observationssæt.
 d. Beregn gennemsnittet af æg i rederne.
 e. Find største og mindste værdi, samt typetal.

Opgave 302

Foreningen af Gymnasiale Knallertkørere spurgte klasserne på et gymnasium, hvor mange der havde kørt på deres egen knallert til skole i løbet af året. Svarene fra de forskellige klasser var:

1, 6, 4, 4, 8, 7, 3, 2, 4, 6, 2, 3, 7, 8,
 2, 0, 5, 2, 4, 1, 3, 0, 6, 3, 3, 1, 0, 3, 1, 2

Observation																			
Hyppeghed																			

- a. Udfyld skemaet.
 b. Hvor mange klasser er der på gymnasiet?
 c. Redegør for, at det er et ikke-grupperet observationssæt.

- d. Bestem største og mindste værdi.
 e. Bestem variationsbredden.
 f. Bestem typetallet.
 g. Beregn gennemsnittet.
 h. Bestem medianen.

Opgave 303

En cyklist synes altid, der er rødt, når hun skal over et bestemt kryds, og beslutter sig for at finde ud af, om det er en rigtig formodning. Ud af 42 gange hun ankom til krydset, var der rødt 21 gange, gult 4 gange og grønt 17 gange.

- a. Hvad er observationerne?
 b. Opstil resultaterne i et skema, hvor hyppighederne fremgår.
 c. Hvorfor kan gennemsnittet ikke beregnes?

Opgave 304

I opgaven med den irriterede cyklist kunne gennemsnittet ikke beregnes. Hvis nu vi sætter 0 = rød, 1 = gul og 2 = grøn, så går det bedre.

- a. Hvad er observationerne nu?
 b. Udregn gennemsnittet.
 c. Hvilken farve er gennemsnitslyset?
 d. Bestem typetallet.

Opgave 305



En tidligere erhvervsøkonomielever inspireres af sin lærer og åbner en genbrugsbutik med modetøj. Eleven kan stadig høre lærerens belæring om, at en "lærende organisation" udfører systematiske kundeundersøgelser. Eleven spørger derfor en tilfældig

dag sine kunder om, hvor mange kilometer de har kørt for at besøge butikken. Svarene i km var:

11, 3, 5, 9, 8, 21, 14, 6, 1, 4, 2, 4, 2, 27, 7, 4, 4

- Bestem kvartilsættet.
- Forklar, hvad kvartilsættet fortæller om hendes kunders afstand til butikken.
- Budskabet i tallene er stadig lidt uklart for den friske iværksætter. Du bedes tegne et plot, så det ses tydeligt, hvordan kundernes afstande fordeler sig.

Opgave 306

I et samfund besad de rigeste 10 % af befolkningen 90% af den samlede kapital. Vil medianen eller gennemsnittet være det bedste tal at beskrive indkomsterne med?

Opgave 307

Halvdelen af de handlende kvinder på en mode-messe var 25 år eller derunder. Hvad er medianen?

Opgave 308

I statistikafdelingen på et lille analysebureau lavede man et forsøg med at opsætte en automat med chokoladeknapper. Direktøren, som holdt meget af den slags undersøgelser, foretog en rundspørge over forbruget på en tilfældig dag. 2 medarbejdere havde ikke fået chokoladeknapper den dag, 3 havde smagt mellem 1 og 5 knapper, og 3 havde spist mellem 25 og 30.

- Direktøren vil gerne have et enkelt tal, der kan måle interessen for chokoladeknapper. Er medianen eller gennemsnittet mon det bedste tal?

Opgave 309

30 ansatte samler LEGO-figurer på tid. Tabellen viser deres tider for at samle en bestemt LEGO-figur (i minutter).

11	15	15	14	18	10	16	12	14	12
17	13	15	12	14	13	11	18	17	13
16	19	10	15	15	12	17	13	15	16

- Lav en tabel for de forskellige tider med hyppighed, frekvenser og kumulerede frekvenser.
- Tegn et stolpediagram over stikprøvens frekvenser.
- Find typetal og gennemsnit.
- Hvor stor forskel er der på typetal og gennemsnittet fundet i spørgsmål c.?

Opgave 310

- Tæl antallet af bogstaver i hvert hele ord i denne opgave.
- Lav så en tabel over antal bogstaver med frekvenser og kumulerede frekvenser.
- Tegn et stolpediagram.
- Tegn et trappediagram.
- Find kvartilerne og kvartilafstanden.

Opgave 311

Tabellen viser alderen på 13 unge i alderen 18–21 år, der i et lille landdistrikt afsluttede en ungdomsuddannelse i 2015.

Alder i år	Antal
18	1
19	3
20	5
21	4

- Tegn et stolpediagram.
- Bestem kvartilerne.
- Beregn gennemsnitsalderen på de 13 unge.

Opgave 312

11 unge har målt deres armlængde fra skulder til fingerspidser i dm.

Resultaterne blev: 7, 7, 8, 8, 8, 9, 6, 8, 7, 7, 8, 8, 8, 7, 8, 8, 7, 7, 7, 7, 8, 7.

- Tegn et stolpediagram.
- Tegn et trappediagram.
- Bestem kvartilsættet.

Opgave 313

En skeptisk tilskuer har fået lov til at undersøge en tryllekunstners terning. Han slår 20 tilfældige kast med terningen og får følgende resultater (i ordnet rækkefølge).

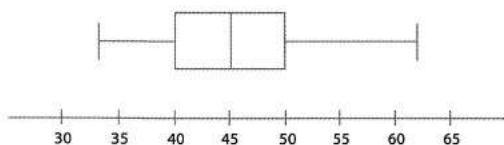
1, 1, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5, 6, 6, 6, 6, 6, 6

- Tegn et stolpediagram.
- Tegn et trappediagram.
- Bestem medianen.
- Tror du, terningen er fair, eller er den måske præpareret til at vise flere 6'ere? Begrund svaret.

Opgave 314

Her er et bokspot.

- Marker kvartilsættet med stiplede linjer, og opskriv det på formen (Q_1, M, Q_3) .



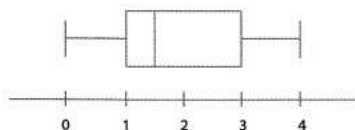
Opgave 315

Her er en række svar på spørgsmålet "Hvor mange søskende har du?"

1, 0, 2, 1, 4, 1, 0, 2, 1, 4, 0, 7, 0, 1, 2, 1

- Bestem kvartilsættet.
- Tegn bokspotet.

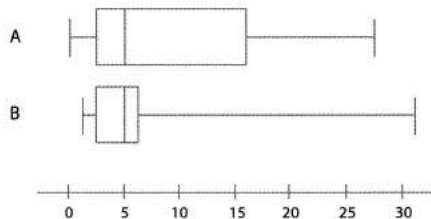
Resultaterne af en anden søskendeundersøgelse ses i dette bokspot.



- Bestem kvartilsættet for den anden undersøgelse.
- Sammenlign resultaterne af de to undersøgelser.

Opgave 316

To naboveje i et finere villakvarter i det østlige Jylland battlede om, hvem der havde den nyeste bilpark. Her er bilernes aldre på de to veje angivet:

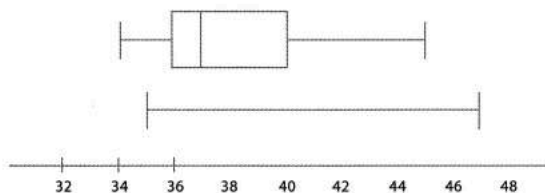


- Skriv kvartilsættet op for de to bokspot.
- Sammenlign de to bokspot.
- Skriv en kort overskrift på undersøgelsen set fra vej A's side, hvor du inddrager ordene median, største og mindste værdi.
- Gør det samme set fra vej B's side, hvor du inddrager ordene variationsbredde og øvre kvartil.

Opgave 317

I to klasser – en med sproglige studieretningsfag og en med naturvidenskabelige studieretningsfag – blev kosttørrelsen undersøgt. Resultaterne fremgår af de to bokspot.

- Opskriv det udvidede kvartilsæt for begge klasser.
- Sammenlign resultaterne, der er vist i de to bokspot.



Opgave 318

- Tegn et bokspot over de afsluttende karakterer til eksamen, som en elev angiver i dette udvidede kvartilsæt $(0, 2, 4, 7, 12)$.
- Tegn et bokspot mere. Denne gang over en elev, der fik "bedre" karakterer.

Opgave 319

I et land målttes følgende antal badedage i juni måned:

År	Badedage
1999	14
2000	20
2001	14
2002	11
2003	8
2004	21
2005	15
2006	14
2007	12
2008	16
2009	13
2010	11
2011	14

- Bestem det gennemsnitlige antal badedage i juni måned.
- Bestem kvartilsættet.
- Vis resultaterne i et boksplot.
- Er det korrekt, at der i over halvdelen af årene var mere end 15 badedage?

Opgave 320

På et biologihold registreredes elevernes skostørrelser i forbindelse med et forsøg. Observationerne var:

43, 37, 38, 37, 36, 39, 39, 43, 43, 38, 39, 44, 43, 38

- Udregn gennemsnittet.
- Udregn kvartilsættet.
- Angiv typetallet.
- Tegn en hyppighedstabel.
- Tegn et stolpediagram.
- Tegn et boksplot.

En elev kom for sent på trods af store, sunde fødder.



- Hvor meget ændres gennemsnittet, hvis vi indregner én elev mere, og denne elev bruger størrelse 46?
- Ændres typetallet?
- Hvor mange af de fem størrelser, der indgår i boksplottet, ændres (begrund svaret)?
- Hvad ændres i stolpediagrammet?

Opgave 321

Her er døgnlængden for otte af solsystemets planeter målt i jorddøgn (tallene er afrundede).

Merkur: 176 Jupiter: 0,4
Venus: 117 Saturn: 0,4
Jorden: 1 Uranus: 0,7
Mars: 1 Neptun: 0,7

- Hvad er et jorddøgn i timer?
- Beregn gennemsnittet af døgnlængden i timer.
- Lav et boksplot.

Opgave 322

- Forklar, hvad et boksplot viser, som et stolpediagram ikke viser.
- Forklar, hvad man hurtigt ser på stolpediagrammet, som man ikke lige ser på boksplottet.
- Kan man tegne et boksplot, hvis man kun har et stolpediagram?
- Kan man tegne et stolpediagram, hvis man kun har boksplottet?

Opgave 323



Sidsel opdrætter røde slanger. Hun har 110 ny-udklækkede unger, hvor længderne i cm fordeler sig på følgende intervaller:

Længde i cm]10;13]]13;16]]16;19]]19;22]]22;25]
Hyppighed	7	31	55	15	2

- Er det et grupperet eller et ikke-grupperet observationssæt?
- Hvad er observationssættets størrelse?
- Bestem typeintervallet.
- Beregn gennemsnittet.

Opgave 324

Gennem 42 år havde en metrolog registreret minimumstemperaturerne i juni måned. Her er resultaterne i grader celsius:

12,2	13,5	15,3	16,3	12,1	10,2	9,4
15,4	11,4	10,4	9,5	17,3	9,4	12,6
15,3	8,2	11,3	16,3	12,9	14,3	17,9
15,2	20,3	17,2	14,4	18,4	22,6	15,4
18,4	12,8	22,5	16,4	16,9	19,3	18,5
19,6	13,6	21,6	20,1	17,3	14,7	19,4

- Grupper data fra ovenstående skema som foreslået i efterfølgende tabel og udfyld resten af den.

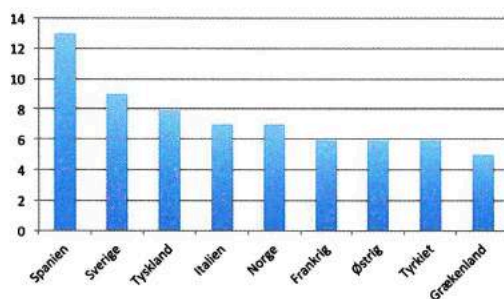
Observations-interval	Interval-hyppighed	Interval-frekvens	Kumuleret interval-frekvens
]8;11]			
]11;14]			
]14;17]			
]17;20]			
]20;23]			100
Sum		100	

- Bestem typeintervallet.

Opgave 325

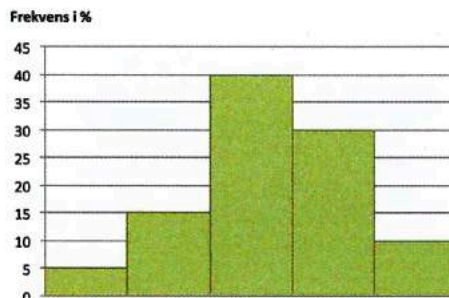
I NYT fra Danmarks Statistik nr. 320 fra 2011 kan man se en opgørelse over danskernes foretrukne feriesteder i udlandet (med mindst 4 overnatninger) i 2010.

Vi har indtastet i Excel, markeret cellerne og fået indsat et diagram.



- Er der tale om et søjlediagram eller om et histogram?
- Begrund dit svar i spørgsmål a.

Opgave 326



Histogrammet viser aldersfordelingen i en forening.

a. Udfyld tabellen.

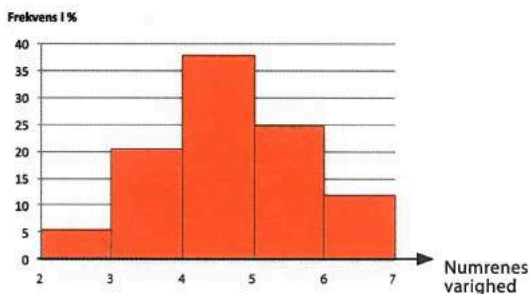
Alder i år	15-30	30-45	45-60	60-75	75-90
Intervalfrekvens					
Kumuleret intervalfrekvens					

- b. Tegn sumkurven.
 c. Hvor mange procent af foreningens medlemmer er under 40 år?
 d. Hvor mange procent af foreningens medlemmer er over 67 år?

Opgave 327

Et band har nu 11 egne numre, der varer mellem 3 og 7 min. Deres gennemsnitlige nummer varer 3:47 min. (3 min. og 47 sek.). De undersøger, hvor lang tid 93 tilfældigt udvalgte sange varer på YouTube i deres favoritgenre: "stonerrock".

Resultaterne ses i dette histogram:



- a. Udfyld på baggrund af histogrammet en tabel, hvor den kumulerede intervalfrekvens fremgår.
 b. Tegn sumkurven.
 c. Brug sumkurven til at aflæse, hvor mange procent af "stonerrock"-sangene, der varer mere end 4 min.
 d. Hvor mange procent af sangene varer mindre end 3 min.?

Opgave 328

En dyreværnsforening, der kæmper for flere fritgående svin, undersøgte aldersfordelingen blandt deres medlemmer:



Alder i år	15-30	30-45	45-60	60-75	75-90
Antal	370	1.549	2.031	963	462

- a. Bestem intervalfrekvenser og kumulerede intervalfrekvenser for observationerne.
 b. Tegn et histogram for denne fordeling.
 c. Tegn sumkurven for fordelingen.
 d. Hvor mange procent af medlemmerne er over 50 år?
 e. Bestem kvartilsættet for fordelingen.
 f. Forklar, hvad nedre kvartil fortæller om fordelingen af medlemmernes alder.

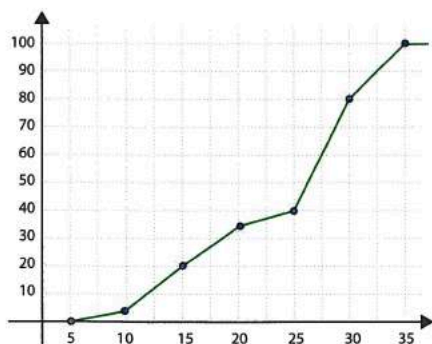
Opgave 329

- a. Tegn sumkurven for observationssættet her.
 b. Hvad fortæller kurvens udseende om observationerne?

Observation]0;10]]10;20]]20;30]]30;40]]40;50]
Hyppeghed	5	5	5	5	5

Opgave 330

Her ses en sumkurve over et grupperet observationsæt over nogle skruers længde.



- Den mindste fjerdedel af skruerne skal kasseres. Hvilke længder er det?
- Halvdelen af skruerne er under længden x . Bestem x .
- Hvilken skruelængde adskiller de korteste 75% af skruerne fra de længste 25%?

Opgave 331

En mekaniker registrerer antal klager fra sine kunder. Han mistænker, at det er dem med ældre biler, der klager, og at han dermed ikke har hele skylden i hvert fald.

- Opstil en hyppighedstabel med intervalfrekvens og kumuleret frekvens for observationsættet i tabellen (bilens alder er målt i år):

Bilens alder]2;5]]5;8]]8;11]]11;14]]14;17]]17;20]]20;23]
Antal klager	4	5	8	7	4	9	3

- Tegn sumkurven.
- Aflæs kvartilsættet for observationerne, idet du viser, hvordan du aflæser det, ved at sætte stiplede pile på sumkurven.
- I sit nyhedsbrev skriver mekanikeren: "Kundetilfredsheden afhænger af bilens alder! En undersøgelse viser, at 75% af klagerne kommer fra dem af jer, der har biler, der er mere end 10 år gamle." Er det korrekt?

Opgave 332

En ven til mekanikeren fra opgave 331 foreslår ham at se på de klagende kunders alder i stedet for bilernes alder:

Kundens alder i år]20;30]]30;40]]40;50]]50;60]]60;70]]70;80]
Antal klager	4	8	14	10	3	1

- Opstil en hyppighedstabel med intervalfrekvens og kumuleret frekvens for observationsættet i tabellen.
- Tegn sumkurven.
- Aflæs kvartilsættet for observationerne, idet du viser, hvordan du aflæser det, ved at sætte stiplede pile på sumkurven.
- Skriv en overskrift til mekanikerens næste nyhedsbrev, hvor du inddrager ordet "median".

Opgave 333

Læs opgaverne 331 og 332, hvis du ikke allerede har regnet dem.

Er der belæg for at påstå følgende:

- At mekanikeren har mindst én kunde over 70 år?
- At mekanikeren ikke har kunder under 20 år?
- At ingen af mekanikerens kunder kører i en bil, der er over 23 år gammel?
- At mekanikerens unge kunder kører i gamle biler?

Opgave 334

Et nystartet ungdomshus afholder forskellige koncerter og undersøger aldersfordelingen af koncertgæsterne:

Alder i år	10-15	15-20	20-25	25-30
Procent	31%	42%	18%	9%

- Bestem de kumulerede frekvenser.
- Tegn en sumkurve.
- Bestem kvartilsættet.
- Gør rede for, hvad medianen fortæller om ungdomshusets koncertgæster.

Opgave 335

En dag med strålende solskin i juli måned kom der 319 tilskadekomne ind på en skadestue i Danmark. Skemaet herunder viser de tilskadekomnes alder.

Alder i år]0;20]]20;40]]40;60]]60;80]
Antal	163	89	46	21

- Tegn en sumkurve og aflæs kvartilsættet.
- Tegn et histogram.

En regnfuld dag i januar måned kom der på samme skadestue 240 tilskadekomne.

Alder i år]0;20]]20;40]]40;60]]60;80]
Antal	68	47	46	79

- Tegn en sumkurve og aflæs kvartilsættet.
- Tegn et histogram.
- Sammenlign de to observationssæt.
- Hvorfor er de forskellige?
- Hvorfor har skadestuer interesse i denne slags statistik?

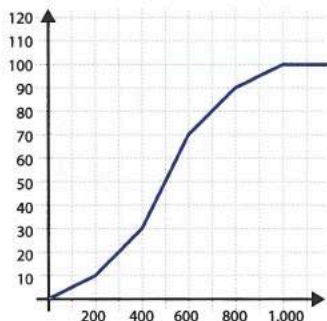
Opgave 336

En hyppighedstabel har typisk fire kolonner:

1. Observationsinterval
2. Intervalhyppighed
3. Intervalfrekvens
4. Kumuleret frekvens

En af kolonnerne kan ikke udfyldes på baggrund af sumkurven alene, men de andre tre kan.

- Tegn en hyppighedstabel og udfyld, hvad du kan, på baggrund af sumkurven.



Opgave 337

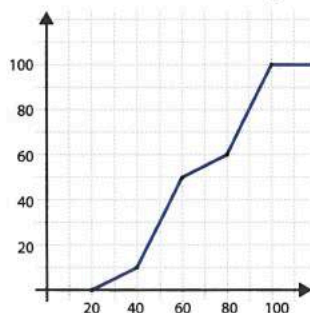
- Opstil en hyppighedstabel med intervalfrekvens og kumuleret intervalfrekvens for observations-sættet i tabellen:

Observation]0;3]]3;6]]6;9]]9;12]]12;15]
Hyppighed	2	5	13	7	1

- Tegn sumkurven.
- Aflæs kvartilsættet for observationerne, idet du viser, hvordan du aflæser det, ved at sætte stiplede pile på sumkurven.
- Aflæs på sumkurven, hvor mange procent af observationerne, der er større end 10.

Opgave 338

- Tegn denne sumkurve af på kvadreret papir.
- Tegn en hyppighedstabel og udfyld rækkerne: observationsintervaller og kumuleret frekvens.



- Kan kolonnen med frekvens udfyldes?
- Kan kolonnen med hyppighed udfyldes?

Opgave 339

- Tegn kurven herunder af på kvadreret papir, og skriv tal på akserne.
- Tegn en hyppighedstabel og udfyld, hvad du kan, på baggrund af sumkurven.

