3 Sentralmål og spredning

Overskriftene viser til delkapitlene i boka. Først kommer en oppsummering av hele kapittelet. Under hver overskrift står gjennomregnede eksempler med løsning, og deretter oppgaver du skal løse uten hjelpemidler. Fasiten står bakerst.

Formelark

SENTRALMÅL

Gjennomsnitt

$$\bar{x} = \frac{\text{sum}}{n}$$

Median Midterste verdi i sortert rekkefølge, plass $\frac{n+1}{2}$.

Fra frekvenstabell

$$\bar{x} = \frac{\sum x \cdot f}{n}$$

Klassedelte data

$$\bar{x} \approx \frac{\sum m \cdot f}{n}$$

- Høyreskjev: $\bar{x} >$ median. Venstreskjev: $\bar{x} <$ median.
- Medianen er robust mot uteliggere.

SPREDNING

Variasjonsbredde

$$\text{største} - \text{minste}$$

Kvartilbredde

$$Q_3 - Q_1$$

Q_1 og Q_3: medianen av nedre og øvre halvdel.

Uteligger Utenfor $Q_1 - 1{,}5 \cdot (Q_3 - Q_1)$ og $Q_3 + 1{,}5 \cdot (Q_3 - Q_1)$.

Standardavvik

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

- Femtallsoppsummering: minste, Q_1, median, Q_3, største.
- Del på n i stedet for $n - 1$ når dataene er hele populasjonen (σ).

VELGE OG SAMMENLIKNE

Velg etter form Symmetrisk: gjennomsnitt og standardavvik. Skjev eller uteliggere: median og kvartilbredde.

Velg etter formål Typisk: ofte median. Total: $\bar{x} \cdot n$.

Sammenlikne grupper Sentrum, spredning, form og uteliggere, i konteksten. Forskjell i forhold til spredningen.

- En forskjell i sentrum betyr ikke at alle i den ene gruppen er høyere.

3.1.1 Gjennomsnitt og median

EKSEMPEL *Søvn i klassen*

De 24 elevene i 1.1.1 sov til sammen 169 timer natt til tirsdag. Hva var gjennomsnittet?

Løsning

$$\bar{x} = \frac{169}{24} \approx 7{,}04 \text{ timer}.$$

Elevene sov i gjennomsnitt omtrent sju timer. ∎

EKSEMPEL *Median med oddetall og partall*

Finn medianen.

a) $7, 3, 9, 5, 4$

b) $12, 8, 15, 10, 20, 9$

Løsning

a) Sortert: $3, 4, 5, 7, 9$. Med $n = 5$ er medianen nummer $\frac{5+1}{2} = 3$, altså 5.

b) Sortert: $8, 9, 10, 12, 15, 20$. Med $n = 6$ er de to midterste 10 og 12, og medianen er $\frac{10+12}{2} = 11$. ∎

EKSEMPEL *Lønn i en liten bedrift*

En bedrift har seks ansatte med månedslønn (i tusen kroner): $38, 41, 43, 45, 47, 180$. Den høyeste er sjefen.

a) Regn ut gjennomsnitt og median.

b) Hvilket tall gir best inntrykk av hva en vanlig ansatt tjener?

Løsning

a) $\bar{x} = \frac{394}{6} \approx 65{,}7$ tusen. Medianen er $\frac{43+45}{2} = 44$ tusen.

b) Medianen. Fem av seks tjener under 50 tusen, men sjefens lønn drar gjennomsnittet opp til nesten 66 tusen, som ingen av de vanlige ansatte tjener. ∎

3.1 Finn gjennomsnitt og median.

a) $4, 8, 6, 2, 10$

b) $3, 9, 5, 7, 11, 1$

3.2 Antall SMS-er sendt i går av 8 elever: $15, 3, 22, 8, 40, 12, 5, 9$. Finn gjennomsnitt og median.

3.3 Hvilken plass i sortert rekkefølge har medianen når $n = 25$? Og når $n = 40$?

3.4 Seks elever har lommepenger (kr per måned): $200, 300, 300, 400, 500, 2500$.

a) Finn gjennomsnitt og median.

b) Hvilket mål beskriver en typisk elev best?

3.5 Gjennomsnittet av fem tall er 12. Fire av tallene er 10, 11, 13 og 15. Hva er det femte?

3.6 En klasse på 25 elever har gjennomsnittshøyde 170 cm. En ny elev på 195 cm begynner. Hva blir gjennomsnittet?

3.7 Er gjennomsnittet trolig større eller mindre enn medianen?

a) Antall følgere på Instagram blant elevene.

b) Antall riktige på en lett prøve (de fleste får nesten alt riktig).

3.8 Hva betyr det at medianen er robust mot uteliggere?

3.9 Folketallet i de 357 kommunene i Norge har gjennomsnitt 15 763 og median 5 447 (SSB, 2026). Forklar hvorfor de er så forskjellige.

3.10 Kan gjennomsnittet av antall søsken være 1,4? Forklar.

3.11 Lag et datasett med fem tall der gjennomsnittet er større enn medianen.

3.12 Hva blir utskriften?

Python

```
data = [7, 3, 9, 1, 5]
s = sorted(data)
print(s[len(s) // 2])
print(sum(data) / len(data))
```

3.13 Hva blir utskriften?

Python

```
data = [10, 20, 30, 40, 1000]
print(sum(data) / len(data))
print(sorted(data)[2])
```

3.14 En avis skriver at «gjennomsnittsnordmannen» har en formue på flere millioner kroner, men de fleste har mye mindre. Forklar hvordan det kan stemme, og hvilket mål avisen heller burde brukt.

3.15 Gjennomsnittet av 10 tall er 8. Ett av tallene endres fra 5 til 25. Hva blir det nye gjennomsnittet?

3.1.2 Sentralmål fra tabeller

EKSEMPEL *Medianen for søsken*

Finn medianen for antall søsken.

Løsning

Søsken	Frekvens	Kumulativ frekvens
0	4	4
1	8	12
2	5	17
3	2	19

Søsken	Frekvens	Kumulativ frekvens
4	1	20

Den kumulative frekvensen passerer 10 ved verdien 1 (den er 12 der). Både elev nummer 10 og 11 i sortert rekkefølge har 1 søsken, så medianen er 1. ∎

EKSEMPEL *Gjennomsnittsalderen i Norge*
Tabellen viser befolkningen i Norge 1. januar 2026 i tiårsklasser, i tusen personer (SSB, tabell 07459). Anslå gjennomsnittsalderen, og finn medianklassen.

Alder	Tusen personer
$[0, 10)$	571
$[10, 20)$	668
$[20, 30)$	703
$[30, 40)$	792
$[40, 50)$	727
$[50, 60)$	749
$[60, 70)$	635
$[70, 80)$	498
$[80, 90)$	239
$[90, 100)$	44
$[100, 110)$	1
Sum	5627

Løsning
Summen av klassemidtpunkt ganger frekvens er

$$5 \cdot 571 + 15 \cdot 668 + 25 \cdot 703 + \cdots + 105 \cdot 1 = 235\,305,$$

så

$$\bar{x} \approx \frac{235\,305}{5627} \approx 41{,}8 \text{ år}.$$

Kumulativ frekvens: 571, 1239, 1942, 2734, 3461, ... Den passerer 2813,5 i klassen $[40, 50)$, som er medianklassen.

Med alderen til hver enkelt (fra SSBs tabell med ettårig alder) blir gjennomsnittet 41,8 år og medianen 41 år. Anslaget fra klassene traff svært godt. ∎

3.16 Antall barn i 20 familier:

Barn	0	1	2	3	4
Frekvens	3	5	7	4	1

a) Regn ut gjennomsnittet.

b) Finn medianen.

3.17 Karakterer på en prøve:

Karakter	1	2	3	4	5	6
Frekvens	1	3	6	8	5	2

a) Hvor mange elever er det?

b) Regn ut gjennomsnittet.

c) Finn medianen.

3.18 Lag en kolonne med kumulativ frekvens for frekvensene $5, 9, 12, 8, 6$.

3.19 Finn klassemidtpunktene for klassene $[0, 4)$, $[4, 8)$ og $[8, 12)$.

3.20 Høyden til 50 elever:

Høyde (cm)	Frekvens
$[150, 160)$	6
$[160, 170)$	15
$[170, 180)$	18
$[180, 190)$	9
$[190, 200)$	2

a) Anslå gjennomsnittet.

b) Finn medianklassen.

3.21 Hvorfor er gjennomsnittet fra klassedelte data bare et anslag?

3.22 Befolkningen i Norge i 2026 er omtrent 5 627 000. Halvparten er omtrent 2 814 000. Den kumulative frekvensen er 2 734 000 ved 40 år og 3 461 000 ved 50 år. I hvilken klasse ligger medianalderen?

3.23 I en frekvenstabell er $n = 30$. Den kumulative frekvensen er 14 for verdien 3 og 22 for verdien 4. Hva er medianen?

3.24 Tim regner ut gjennomsnittet fra en frekvenstabell ved å legge sammen frekvensene og dele på antall verdier. Hva er feil?

3.25 Antall timer skjermtid per dag for 40 elever:

Timer	Frekvens
$[0, 2)$	4
$[2, 4)$	12
$[4, 6)$	14
$[6, 8)$	7
$[8, 10)$	3

Anslå gjennomsnittet.

3.26 Hva blir utskriften?

Python

```
x = [1, 2, 3]
f = [5, 3, 2]
print(sum(a * b for a, b in zip(x, f)) / sum(f))
```

3.27 Hva blir utskriften?

Python

```
f = [3, 5, 7, 4, 1]
kum = 0
for a in f:
    kum = kum + a
    print(kum)
```

3.28 En tabell over inntekt har siste klasse «1 million og mer». Hvorfor er det vanskelig å anslå gjennomsnittet?

3.29 I en frekvenstabell med $n = 40$ har verdien 5 den kumulative frekvensen 20, og verdien 6 har 27. Hva er medianen?

3.30 Gjennomsnittsalderen i en forening er 34 år for 30 medlemmer. Ti nye medlemmer med gjennomsnittsalder 54 år melder seg inn. Hva blir gjennomsnittsalderen?

3.2.1 Variasjonsbredde og kvartiler

EKSEMPEL *Kvartiler med oddetall og partall*

Finn kvartilene og kvartilbredden.

a) 3, 7, 8, 12, 15, 18, 21

b) 2, 5, 6, 9, 11, 14, 15, 20

Løsning

a) Medianen er 12. Nedre halvdel: 3, 7, 8, så $Q_1 = 7$. Øvre halvdel: 15, 18, 21, så $Q_3 = 18$. Kvartilbredden er $18 - 7 = 11$.

b) Medianen er $\frac{9+11}{2} = 10$. Nedre halvdel: 2, 5, 6, 9, så $Q_1 = \frac{5+6}{2} = 5{,}5$. Øvre halvdel: 11, 14, 15, 20, så $Q_3 = \frac{14+15}{2} = 14{,}5$. Kvartilbredden er $14{,}5 - 5{,}5 = 9$.

■

EKSEMPEL *Søvnen i klassen*

De 24 søvnverdiene fra 1.1.1, sortert:

4,5 5 5,5 6 6 6,5 6,5 6,5 6,5 7 7 7
7 7 7,5 7,5 7,5 7,5 8 8 8 8,5 9 9,5

Finn medianen, kvartilene og kvartilbredden.

Løsning

Medianen er gjennomsnittet av nummer 12 og 13: $\frac{7+7}{2} = 7$.

Q_1: nummer 6 og 7 i første rad er 6,5 og 6,5, så $Q_1 = 6{,}5$.

Q_3: nummer 6 og 7 i andre rad er 7,5 og 8, så $Q_3 = 7{,}75$.

Kvartilbredden er $7{,}75 - 6{,}5 = 1{,}25$ timer. Den midterste halvdelen av klassen sov mellom 6,5 og 7,75 timer.

■

3.31 Finn variasjonsbredden for 12, 7, 25, 18, 3, 14.

3.32 Finn medianen, kvartilene og kvartilbredden for 4, 7, 9, 10, 12, 15, 18.

3.33 Finn kvartilene og kvartilbredden for 2, 3, 5, 8, 9, 11, 13, 16.

3.34 Finn kvartilene for 21, 15, 30, 18, 26, 12, 24, 19, 28.

3.35 Hvor stor del av observasjonene ligger mellom Q_1 og Q_3?

3.36 Hvorfor er kvartilbredden et bedre spredningsmål enn variasjonsbredden når det finnes uteliggere?

3.37 Hva er Q_1, medianen og Q_3 uttrykt som persentiler?

3.38 En elev fikk vite at resultatet hennes på en nasjonal prøve var på 80-persentilen. Hva betyr det?

3.39 I et datasett er $Q_1 = 14$ og kvartilbredden 9. Hva er Q_3?

3.40 Datasettet 5, 6, 8, 9, 10, 12, 14 får en ny verdi, 60. Hva skjer med variasjonsbredden og kvartilbredden?

3.41 Hva blir utskriften?

Python

```
data = [3, 9, 4, 7, 12]
print(max(data) - min(data))
```

3.42 Hva blir utskriften?

Python

```
s = sorted([8, 2, 6, 4, 10, 12])
nedre = s[:3]
ovre = s[3:]
print(nedre, ovre)
print(nedre[1], ovre[1])
```

3.43 Lag et datasett med 7 tall der variasjonsbredden er 20 og kvartilbredden er 2.

3.44 Hvorfor kan Python gi litt andre kvartiler enn du får for hånd?

3.45 To klasser har samme median, 4, på en prøve, men kvartilbredde 1 og 3. Beskriv forskjellen.

3.2.2 Boksplott

EKSEMPEL *Eleven som trente mye*

En tolvte elev trente 40 timer. Nå er $Q_1 = 3$ og $Q_3 = 9$ (fra 3.2.1). Er 40 en uteligger? Hvor går værhåret?

Løsning
Kvartilbredden er $9 - 3 = 6$. Øvre grense: $9 + 1{,}5 \cdot 6 = 18$. Nedre grense: $3 - 9 = -6$.

$40 > 18$, så 40 er en uteligger og tegnes som et punkt. Det høyre værhåret går ut til 14, som er den største verdien under grensen. ∎

3.46 Finn femtallsoppsummeringen og tegn et boksplott for $3, 5, 6, 8, 9, 11, 12, 14, 20$.

3.47 Et datasett har $Q_1 = 40$ og $Q_3 = 56$. Finn grensene for uteliggere. Er 15 og 85 uteliggere?

3.48 Datasettet $2, 4, 5, 5, 6, 7, 8, 9, 25$ har $Q_1 = 4{,}5$ og $Q_3 = 8{,}5$. Er 25 en uteligger? Hvor går det høyre værhåret?

3.49 Hva forteller det om formen at medianen ligger nær Q_1 og det høyre værhåret er langt?

3.50 Hvor stor del av dataene ligger omtrent i

a) boksen,

b) det venstre værhåret?

3.51 To boksplott viser lønn i to bedrifter. Bedrift A har median 45 000 og kvartilbredde 5000. Bedrift B har median 42 000 og kvartilbredde 15 000. Sammenlikn.

3.52 Kommunene i Nordland har median 2024 og $Q_3 = 7665$ innbyggere (SSB, 2026). Hva kan du si om kommunene med over 7665 innbyggere?

3.53 Hva viser et boksplott ikke?

3.54 Kari tegner værhårene ut til grensene $Q_1 - 1{,}5 \cdot$ kvartilbredde og $Q_3 + 1{,}5 \cdot$ kvartilbredde. Hva er feil?

3.55 Tegn boksplott under hverandre og sammenlikn: A har femtallsoppsummering $10, 15, 18, 20, 24$ og B har $8, 12, 18, 26, 30$.

3.56 Hva blir utskriften?

Python

```
q1 = 12
q3 = 20
iqr = q3 - q1
print(q1 - 1.5 * iqr, q3 + 1.5 * iqr)
```

3.57 Hva blir utskriften?

Python

```
data = [1, 5, 6, 7, 8, 9, 30]
grense = 9 + 1.5 * (9 - 5)
print([x for x in data if x > grense])
```

3.58 Lag et datasett med 7 verdier som har én uteligger. Vis at det er en uteligger.

3.59 Hvorfor er boksplott praktiske når man skal sammenlikne mange grupper?

3.2.3 Standardavvik

EKSEMPEL *Standardavviket for skytter A*
Regn ut standardavviket for skytter A: $4, 6, 7, 8, 10$.

Løsning
$\bar{x} = \dfrac{35}{5} = 7$.

x	$x - \bar{x}$	$(x - \bar{x})^2$
4	-3	9
6	-1	1
7	0	0
8	1	1
10	3	9
Sum	0	20

$$s^2 = \frac{20}{5-1} = 5, \qquad s = \sqrt{5} \approx 2{,}24 \ \text{cm}.$$

∎

EKSEMPEL *Standardavviket for skytter B*
Regn ut standardavviket for skytter B: $1, 4, 7, 10, 13$. Hvem er mest presis?

Løsning
$\bar{x} = 7$. Kvadrerte avvik: $36 + 9 + 0 + 9 + 36 = 90$.

$$s^2 = \frac{90}{4} = 22{,}5, \qquad s = \sqrt{22{,}5} \approx 4{,}74 \ \text{cm}.$$

Skytter A har mindre standardavvik (2,24 mot 4,74 cm), og er mest presis: pilene ligger tettere rundt gjennomsnittet. ∎

3.60 Regn ut standardavviket for $2, 4, 6, 8, 10$.

3.61 Regn ut standardavviket for $5, 7, 7, 9$.

3.62 Hvorfor er summen av avvikene $x - \bar{x}$ alltid 0?

3.63 Variansen er 16. Hva er standardavviket?

3.64 Hva er standardavviket til $9, 9, 9, 9, 9$? Forklar.

3.65 Hvilken enhet har standardavviket for høyde målt i cm? Og variansen?

3.66 En prøve har gjennomsnitt 30 poeng og standardavvik 6. Omtrent hvor mange av 30 elever ligger mellom 24 og 36 poeng, hvis fordelingen har én topp og er omtrent symmetrisk?

3.67 Når deler du på n og når på $n - 1$ i standardavviket?

3.68 To skyttere har standardavvik 1,5 cm og 4,2 cm for avstanden fra midten. Hvem er mest presis?

3.69 Alle verdiene i et datasett med standardavvik 3 økes med 5. Hva blir standardavviket? Hva om de ganges med 2 i stedet?

3.70 Regn ut standardavviket for temperaturene $-2, 0, 1, 3, 3$ (°C).

3.71 Hva blir utskriften?

Python

```
data = [2, 4, 6]
m = sum(data) / len(data)
print([(x - m) ** 2 for x in data])
```

3.72 Hva blir utskriften?

Python

```
data = [2, 4, 6]
m = sum(data) / len(data)
s2 = sum((x - m) ** 2 for x in data) / (len(data)
- 1)
print(s2, s2 ** 0.5)
```

3.73 Lise har regnet ut variansen 9 og skriver «standardavviket er 9». Hva er feil?

3.74 Lag to datasett med fire tall og gjennomsnitt 10, der det ene har standardavvik 0 og det andre et standardavvik større enn 5.

3.3.1 Velge og begrunne mål

EKSEMPEL *Hvilke mål?*

Hvilket sentralmål og spredningsmål vil du bruke? Begrunn.

a) Høyden til 200 elever på Vg2.

b) Hvor mange minutter elevene bruker på skoleveien, når noen få har over en time.

c) Prisen på leiligheter solgt i en by.

Løsning

a) Gjennomsnitt og standardavvik: høyde er omtrent symmetrisk.

b) Median og kvartilbredde: høyreskjev med uteliggere.

c) Median og kvartilbredde: boligpriser er høyreskjeve, med noen svært dyre boliger. ■

EKSEMPEL *Kantinebudsjettet*

400 elever bruker penger i kantina. De fleste bruker 0 –100 kr i uka, men noen bruker over 300 kr. Medianen er 60 kr og gjennomsnittet 85 kr per uke.

a) Hvilket tall beskriver en typisk elev?

b) Kantina vil anslå inntekten per uke. Hvilket tall skal den bruke, og hva blir anslaget?

Løsning

a) Medianen, 60 kr: halvparten bruker mindre, halvparten mer.

b) Gjennomsnittet, fordi inntekten er en sum: $85 \cdot 400 = 34\,000$ kr. Med medianen ville kantina anslått 24 000 kr, og bommet med 10 000 kr, fordi de som bruker mye, ikke ville blitt regnet med. ■

3.75 Hvilke mål hører sammen?

a) Gjennomsnitt og ...

b) Median og ...

3.76 Hvilket sentralmål og spredningsmål vil du bruke? Begrunn.

a) Vekten til nyfødte.

b) Antall timer elever spiller dataspill i uka.

c) Prisen på brukte biler til salgs på nett.

3.77 Hvilket sentralmål passer for favorittfarge? Hvorfor kan du ikke bruke gjennomsnitt eller median?

3.78 En skole har 520 elever. I gjennomsnitt kjøper hver elev 1,8 boller i uka i kantina. Hvor mange boller må kantina regne med å selge per uke?

3.79 I en bydel er medianinntekten 520 000 kr og gjennomsnittsinntekten 780 000 kr. Hva sier det om fordelingen?

3.80 En arbeidsgiver vil vise at lønningene er høye, og en fagforening at mange tjener lite. Hvilket mål vil hver av dem trolig velge når lønnsfordelingen er høyreskjev?

3.81 Hvorfor er det lurt å se på en graf av dataene før du velger mål?

3.82 Medianinntekten etter skatt for husholdningene i kommunene har gjennomsnitt 663 243 kr og median 657 600 kr (SSB, 2024). Hva sier det om formen?

3.83 Hva er galt med å oppgi «median 45 minutter og standardavvik 38 minutter»?

3.84 En kommune vil vite hvor mye vann innbyggerne bruker til sammen. Forbruket er høyreskjevt. Hvilket sentralmål trenger kommunen, og hvorfor?

3.85 Hva blir utskriften?

Python

```
data = [3, 4, 4, 5, 30]
s = sorted(data)
print(sum(data) / len(data), s[2])
```

3.86 Hva blir utskriften?

Python

```
snitt = 85
n = 400
print(snitt * n)
```

3.87 Folketallet i kommunene er sterkt høyreskjevt. Hvilke mål ville du brukt for å beskrive en typisk kommune og spredningen?

3.88 Gi et eksempel der gjennomsnittet er det beste målet selv om fordelingen er skjev.

3.89 Hvorfor er medianen av medianinntektene i kommunene ikke det samme som medianinntekten for alle husholdninger i Norge?

3.3.2 Sammenlikne grupper

EKSEMPEL *Søvn i to klasser*

I klasse A sov elevene i gjennomsnitt 7,0 timer med standardavvik 1,2 timer. I klasse B sov de i gjennomsnitt 7,4 timer med standardavvik 1,1 timer. Begge fordelingene er omtrent symmetriske. Er forskjellen stor?

Løsning

Forskjellen er 0,4 timer, omtrent en tredjedel av et standardavvik. Søvnen varierer mye mer innenfor hver klasse enn mellom klassene, så fordelingene overlapper nesten helt. Mange elever i A sover lenger enn mange i B. Forskjellen er liten, og kan også skyldes tilfeldigheter, slik vi så i 1.1.1. I kapittel 9 og 10 lærer vi å vurdere det mer presist. ∎

3.90 Hvilke tre ting bør du sammenlikne når du sammenlikner to grupper?

3.91 To klasser fikk disse resultatene på en prøve (poeng):

	Klasse A	Klasse B
Gjennomsnitt	32	35
Standardavvik	4	9

Sammenlikn klassene.

3.92 Medianinntekten etter skatt i kommunene (SSB, 2024) er 620 500 kr i Innlandet og 708 000 kr i Vestland. Hvor stor er forskjellen i prosent av Innlandet?

3.93 To grupper har medianer 40 og 46. Gruppe A har $Q_3 = 44$, og gruppe B har $Q_1 = 43$. Overlapper boksene?

3.94 Hvorfor er det ikke nok å sammenlikne gjennomsnittene?

3.95 To grupper har gjennomsnitt 100 og 106. Er forskjellen stor hvis standardavviket i begge er

a) 3?

b) 30?

3.96 Femtallsoppsummering for skolevei (minutter) i to klasser:

	Klasse A	Klasse B
Minste	5	2
Q_1	10	8
Median	15	12
Q_3	20	25
Største	35	60

Tegn boksplottene og sammenlikn.

3.97 En elev skriver: «Gutter er høyere enn jenter.» Hvordan bør setningen skrives for å være riktig?

3.98 To grupper har svært ulik størrelse, 8 og 200. Hva bør du passe på når du sammenlikner dem?

3.99 Hvilke grafer passer for å sammenlikne en numerisk variabel i tre grupper?

3.100 Hva blir utskriften?

Python

```
a = [12, 15, 18]
b = [10, 15, 26]
for g in [a, b]:
    print(sum(g) / len(g), max(g) - min(g))
```

3.101 Hva blir utskriften?

Python

```
grupper = {"A": [3, 5, 7], "B": [1, 5, 9]}
for navn in grupper:
    g = grupper[navn]
    print(navn, sorted(g)[1])
```

3.102 Klasse A har gjennomsnittskarakter 4,2 og klasse B 3,9. Rektor sier at A har bedre undervisning. Gi to grunner til at det ikke nødvendigvis stemmer.

3.103 Skriv en sammenlikning på tre setninger: Gruppe A har median 8, kvartilbredde 2, symmetrisk. Gruppe B har median 8, kvartilbredde 6, høyreskjev.

3.104 Hva betyr det at to fordelinger overlapper mye?

3.4.1 Undersøkelse: Hvem er typisk?

3.105 Formuler et statistisk spørsmål om folketallet i kommunene som kan besvares med sentralmål og spredningsmål.

3.106 Hvilke mål og grafer vil du bruke for å sammenlikne folketallet i kommunene i to fylker? Begrunn.

3.107 Medianinntekten etter skatt (SSB, 2024) for de 178 største kommunene har median 671 000 kr, og for de 179 minste 638 800 kr. Hvor stor er forskjellen?

3.108 Kan undersøkelsen over vise at det å bo i en stor kommune gir høyere inntekt? Begrunn.

3.109 Hvorfor er sekundærdata fra SSB godt egnet til denne undersøkelsen? Nevn også en ulempe.

3.110 En gruppe oppgir gjennomsnittlig folketall i kommunene i et fylke som «8 249,97». Hva er problemet?

3.111 Hva bør en rapport fra en slik undersøkelse inneholde?

3.112 Hva blir utskriften?

Python

```
folk = [700, 2000, 5400, 12000, 290000]
s = sorted(folk)
print(s[2], round(sum(folk) / len(folk)))
```

3.113 Hvordan kan du gjøre store og små kommuner til en kategorisk variabel? Hvilken grense er naturlig?

3.114 Skriv en ærlig konklusjon: Medianen av medianinntektene er 744 100 kr i Rogaland og 607 750 kr i Finnmark.

Fasit

3.1.1 Gjennomsnitt og median

3.1
a) $\bar{x} = 6$, median 6

b) $\bar{x} = 6$, median $\frac{5+7}{2} = 6$

3.2
$\bar{x} = \frac{114}{8} = 14{,}25$. Sortert:

$3, 5, 8, 9, 12, 15, 22, 40$, median

$\frac{9+12}{2} = 10{,}5$.

3.3
$n = 25$: plass 13. $n = 40$: gjennomsnittet av plass 20 og 21.

3.4
a) $\bar{x} = \frac{4200}{6} = 700$ kr, median $\frac{300+400}{2} = 350$ kr.

b) Medianen. Én elev med 2500 kr drar gjennomsnittet opp, og fem av seks får under 700 kr.

3.5
$5 \cdot 12 - 49 = 11$

3.6
$\frac{25 \cdot 170 + 195}{26} = \frac{4445}{26} \approx 171{,}0$ cm

3.7
a) Større (høyreskjev: noen få har svært mange følgere).

b) Mindre (venstreskjev).

3.8
Den endrer seg lite eller ikke i det hele tatt når en ekstrem verdi legges til eller endres, fordi den bare avhenger av hvilken verdi som er i midten.

3.9
Fordelingen er sterkt høyreskjev: de fleste kommunene er små, men noen få er svært store (Oslo har over 700 000). De store kommunene drar gjennomsnittet opp, mens medianen bare er verdien i midten.

3.10
Ja. Gjennomsnittet trenger ikke være en verdi som finnes i dataene. 1,4 betyr at summen av søsken delt på antall elever er 1,4.

3.11
For eksempel $1, 2, 3, 4, 20$: median 3, gjennomsnitt 6.

3.12
```
5
5.0
```

3.13
```
220.0
30
```

3.14
Formue er sterkt høyreskjev: noen få har enorme formuer som drar gjennomsnittet langt opp. Medianen viser hva en typisk person har.

3.15
Summen øker med 20, fra 80 til 100. Nytt gjennomsnitt: 10.

3.1.2 Sentralmål fra tabeller

3.16
a) $\frac{0+5+14+12+4}{20} = \frac{35}{20} = 1{,}75$

b) Kumulativ frekvens: $3, 8, 15, 19, 20$. Nummer 10 og 11 har 2 barn. Median 2.

3.17
a) 25

b) $\frac{1+6+18+32+25+12}{25} = \frac{94}{25} = 3{,}76$

c) Kumulativ frekvens: $1, 4, 10, 18, \ldots$. Nummer 13 har karakter 4. Median 4.

3.18
$5, 14, 26, 34, 40$

3.19
2, 6 og 10

3.20
a) $\frac{155 \cdot 6 + 165 \cdot 15 + 175 \cdot 18 + 185 \cdot 9 + 195 \cdot 2}{50} = \frac{8610}{50} = 172{,}2$ cm

b) Kumulativ frekvens: $6, 21, 39, \ldots$. Halvparten er 25, så medianklassen er $[170, 180)$.

3.21
Vi kjenner ikke de nøyaktige verdiene, bare hvor mange som er i hver klasse. Vi later som om alle har klassemidtpunktet.

3.22
$[40, 50)$

3.23
Nummer 15 og 16 har begge verdien 4, så medianen er 4.

3.24
Han må gange hver verdi med frekvensen, summere, og dele på antall observasjoner n (summen av frekvensene).

3.25
$\frac{1 \cdot 4 + 3 \cdot 12 + 5 \cdot 14 + 7 \cdot 7 + 9 \cdot 3}{40} = \frac{186}{40} = 4{,}65$ timer

3.26
```
1.7
```

3.27
```
3
8
15
19
20
```

3.28
Klassen har ingen øvre grense, så vi vet ikke hvilket klassemidtpunkt vi skal bruke, og noen kan ha svært høy inntekt som påvirker gjennomsnittet mye.

3.29
Nummer 20 har verdien 5 og nummer 21 har verdien 6. Medianen er 5,5.

3.30
$\frac{30 \cdot 34 + 10 \cdot 54}{40} = \frac{1560}{40} = 39$ år

3.2.1 Variasjonsbredde og kvartiler

3.31
$25 - 3 = 22$

3.32
Median 10. $Q_1 = 7$ (median av $4, 7, 9$), $Q_3 = 15$ (median av $12, 15, 18$). Kvartilbredde 8.

3.33
Median 8,5. $Q_1 = \frac{3+5}{2} = 4$, $Q_3 = \frac{11+13}{2} = 12$. Kvartilbredde 8.

3.34
Sortert: $12, 15, 18, 19, 21, 24, 26, 28, 30$.

Median 21. $Q_1 = \frac{15+18}{2} = 16{,}5$,

$Q_3 = \frac{26+28}{2} = 27$.

3.35
Omtrent halvparten.

3.36
Kvartilbredden bygger på den midterste halvdelen av dataene og påvirkes ikke av ekstreme verdier. Variasjonsbredden bygger på de to mest ekstreme verdiene.

3.37
25-, 50- og 75-persentilen.

3.38
Omtrent 80 % av elevene fikk like godt eller dårligere resultat, og omtrent 20 % fikk bedre.

3.39
23

3.40
Før: variasjonsbredde 9, $Q_1 = 6$, $Q_3 = 12$, kvartilbredde 6. Etter ($n = 8$): variasjonsbredde 55, $Q_1 = \frac{6+8}{2} = 7$, $Q_3 = \frac{12+14}{2} = 13$, kvartilbredde 6. Variasjonsbredden øker kraftig, kvartilbredden endres ikke.

3.41
```
9
```

3.42
```
[2, 4, 6] [8, 10, 12]
4 10
```

3.43
For eksempel $0, 9, 10, 10, 10, 11, 20$: variasjonsbredde 20, $Q_1 = 9$, $Q_3 = 11$, kvartilbredde 2.

3.44
Det finnes flere metoder for å regne ut kvartiler. Python bruker en annen metode enn halvdelsmetoden. For store datasett er forskjellen liten.

3.45
Den midterste halvdelen av klassen med kvartilbredde 1 har karakterer tett rundt 4. I den andre klassen er den midterste halvdelen mye mer spredt.

3.2.2 Boksplott

3.46
Minste 3, $Q_1 = 5{,}5$, median 9, $Q_3 = 13$, største 20. Boks fra 5,5 til 13 med strek ved 9, værhår til 3 og 20.

3.47
Kvartilbredde 16. Grenser: $40 - 24 = 16$ og $56 + 24 = 80$. 15 og 85 er begge uteliggere.

3.48
Kvartilbredde 4, øvre grense $8{,}5 + 6 = 14{,}5$. 25 er en uteligger. Værhåret går til 9.

3.49
Fordelingen er høyreskjev.

3.50
a) Halvparten b) En fjerdedel

3.51
A har høyere typisk lønn og mye mindre spredning. Lønningene i B varierer mye mer, så noen i B tjener trolig mer enn de fleste i A, og andre mye mindre.

3.52
Omtrent en fjerdedel av kommunene i Nordland har over 7665 innbyggere.

3.53
Antall observasjoner, de enkelte verdiene, og om fordelingen har flere topper.

3.54
Værhårene skal gå ut til den mest ekstreme verdien som ligger innenfor grensene, ikke ut til selve grensene.

3.55
Samme median (18). B har mye større spredning (kvartilbredde 14 mot 5). A har medianen nær Q_3 og langt venstre værhår (litt venstreskjev). B er omtrent symmetrisk.

3.56
```
0.0 32.0
```

3.57
```
[30]
```

3.58
For eksempel $4, 5, 6, 7, 8, 9, 30$: $Q_1 = 5$, $Q_3 = 9$, kvartilbredde 4, øvre grense 15. $30 > 15$.

3.59
Hvert boksplott tar lite plass, de kan stå under hverandre på samme akse, og sentrum, spredning, form og uteliggere kan sammenliknes direkte.

3.2.3 Standardavvik

3.60
$\bar{x} = 6$. Kvadrerte avvik: $16 + 4 + 0 + 4 + 16 = 40$. $s = \sqrt{\frac{40}{4}} = \sqrt{10} \approx 3{,}16$.

3.61
$\bar{x} = 7$. Kvadrerte avvik: $4 + 0 + 0 + 4 = 8$. $s = \sqrt{\frac{8}{3}} \approx 1{,}63$.

3.62
Gjennomsnittet er balansepunktet: de positive og de negative avvikene veier nøyaktig like mye.

3.63
4

3.64
0. Alle verdiene er lik gjennomsnittet, så det er ingen spredning.

3.65
Standardavviket: cm. Variansen: cm^2.

3.66
Omtrent to tredjedeler, altså rundt 20 elever.

3.67
$n - 1$ for et utvalg (vanligst), n når dataene er hele populasjonen.

3.68
Skytteren med 1,5 cm: pilene ligger tettere.

3.69
Økes med 5: fortsatt 3. Ganges med 2: 6.

3.70
$\bar{x} = 1$. Kvadrerte avvik: $9 + 1 + 0 + 4 + 4 = 18$. $s = \sqrt{\frac{18}{4}} = \sqrt{4{,}5} \approx 2{,}12$ °C.

3.71
```
[4.0, 0.0, 4.0]
```

3.72
```
4.0 2.0
```

3.73
Standardavviket er kvadratroten av variansen: $s = 3$.

3.74
For eksempel $10, 10, 10, 10$ ($s = 0$) og $1, 4, 16, 19$ ($\bar{x} = 10$, kvadrerte avvik $81 + 36 + 36 + 81 = 234$, $s = \sqrt{78} \approx 8{,}8$).

3.3.1 Velge og begrunne mål

3.75
a) standardavvik b) kvartilbredde

3.76
a) Gjennomsnitt og standardavvik (omtrent symmetrisk).
b) Median og kvartilbredde (høyreskjev: mange spiller lite, noen svært mye).
c) Median og kvartilbredde (høyreskjev med noen svært dyre biler).

3.77
Typetallet. Favorittfarge er kategorisk uten rekkefølge, så verdiene kan verken legges sammen eller sorteres.

3.78
$1{,}8 \cdot 520 = 936$ boller.

3.79
Gjennomsnittet er mye høyere enn medianen, så fordelingen er høyreskjev: noen få har svært høy inntekt.

3.80
Arbeidsgiveren: gjennomsnittet (høyest). Fagforeningen: medianen (lavere).

3.81
Grafen viser formen og eventuelle uteliggere, som avgjør om gjennomsnitt og standardavvik eller median og kvartilbredde passer best.

3.82
Gjennomsnitt og median er nesten like, så fordelingen er omtrent symmetrisk.

3.83
Målene er blandet. Median hører sammen med kvartilbredde, og gjennomsnitt med standardavvik.

3.84
Gjennomsnittet, fordi totalen er gjennomsnittet ganget med antall innbyggere. Medianen ville gitt for lav total.

3.85

```
9.2 4
```

3.86

```
34000
```

3.87
Median og kvartilbredde.

3.88
Når vi skal regne ut en total, for eksempel samlet strømforbruk, samlet inntekt i en kantine eller hvor mye avfall en kommune må hente.

3.89
Hver kommune teller likt i medianen av kommunene, uansett hvor mange husholdninger den har. Store kommuner får for liten vekt.

3.3.2 Sammenlikne grupper

3.90
Sentrum, spredning og form (med uteliggere).

3.91
B har høyere gjennomsnitt (3 poeng), men mye større spredning. Forskjellen i gjennomsnitt er liten i forhold til spredningen i B, så fordelingene overlapper mye. A er mer jevn.

3.92
$$\frac{87\,500}{620\,500} \cdot 100\,\% \approx 14{,}1\,\%$$

3.93
Ja, litt: Boksen til A går opp til 44, og boksen til B starter på 43.

3.94
Gjennomsnittet sier ikke noe om spredning eller form. To grupper kan ha samme gjennomsnitt, men svært forskjellig spredning, eller ulik form.

3.95
a) Ja, forskjellen er to standardavvik.
b) Nei, forskjellen er en femdel av et standardavvik.

3.96
A har høyere median (15 mot 12), men B har større spredning (kvartilbredde 17 mot 10). B er høyreskjev med noen svært lange skoleveier. A er omtrent symmetrisk.

3.97
For eksempel: «Gutter er i gjennomsnitt høyere enn jenter, men fordelingene overlapper mye, så mange jenter er høyere enn mange gutter.»

3.98
Sentral- og spredningsmålene i den lille gruppen er mye mer usikre. Det bør nevnes, og man bør være forsiktig med å konkludere.

3.99
Boksplott under hverandre på samme akse, eller histogrammer/prikkplott over hverandre med samme skala.

3.100

```
15.0 6
17.0 16
```

3.101

```
A 5
B 5
```

3.102
For eksempel: Forskjellen kan skyldes tilfeldig variasjon (særlig hvis spredningen er stor); klassene kan være forskjellige på andre måter (forkunnskaper, motivasjon); det er en observasjonsstudie, ikke et eksperiment.

3.103
For eksempel: Gruppene har samme median, 8. Gruppe B varierer mye mer, med kvartilbredde 6 mot 2. B er høyreskjev med noen høye verdier, mens A er symmetrisk.

3.104
Mange verdier i den ene gruppen ligger i samme område som verdier i den andre, så selv om sentrum er forskjellig, er det mange i den «lave» gruppen som har høyere verdier enn mange i den «høye».

3.4.1 Undersøkelse: Hvem er typisk?

3.105
For eksempel: «Hvor stor er en typisk kommune i Nord-Norge, og varierer størrelsen mer der enn i Sør-Norge?»

3.106
Boksplott på samme akse, og median og kvartilbredde, fordi folketallet er sterkt høyreskjevt med uteliggere.

3.107
32 200 kr.

3.108
Nei. Det er en observasjonsstudie, så den viser en sammenheng. Andre forhold, som næringsliv, alder på innbyggerne og utdanning, kan forklare forskjellen.

3.109
Fordel: data for alle kommunene, samlet inn på samme måte av en etterrettelig kilde. Ulempe: Vi har bare medianen for hver kommune, ikke dataene for hver husholdning, og definisjonene er bestemt av SSB.

3.110
Falsk presisjon: folketall oppgis i hele personer, og for et typisk folketall holder det å runde av, for eksempel til 8200. Gjennomsnittet passer dessuten dårlig for en skjev fordeling.

3.111
Spørsmål, datasett og kilde, valg av mål og grafer med begrunnelse, tabell og grafer, konklusjon i konteksten, og begrensninger (for eksempel at hver verdi er en median for en kommune, og at det er en observasjonsstudie).

3.112

```
5400  62020
```

3.113
Del kommunene ved medianfolketallet: over medianen er «store», resten «små». Da blir gruppene omtrent like store.

3.114
For eksempel: «Kommunene i Rogaland har typisk høyere medianinntekt etter skatt for husholdningene enn kommunene i Finnmark, omtrent 744 000 mot 608 000 kr i 2024. Hver kommune teller likt, uansett størrelse.»