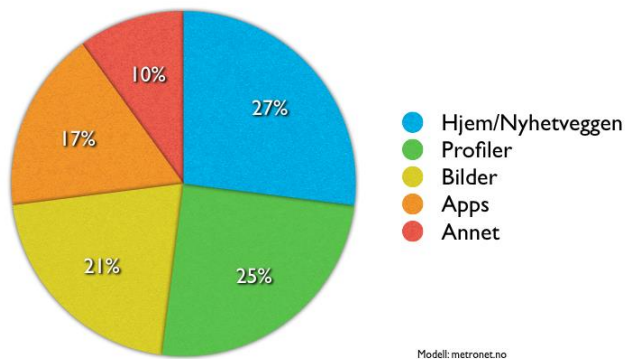


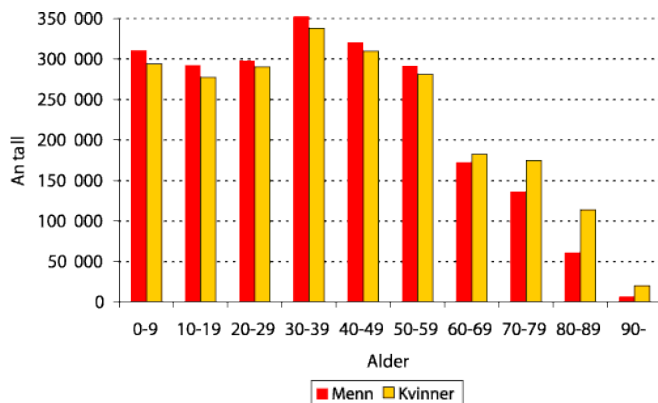
Kapittel 5. Statistikk

Tid brukt på Facebook: fordelt på Innhold



Dette kapitlet handler blant annet om:

- Beregne gjennomsnitt og andre sentralmål.
- Framstille data i frekvenstabeller.
- Beregne standardavvik og andre spredningsmål.
- Framstille data i søyle-, sektor- og andre typer diagrammer.
- Bruke Excel til å gjøre statistiske beregninger.



Mål for kapittel 5. Statistikk



Kompetansemål

Mål for opplæringen er at eleven skal kunne

- planlegge, gjennomføre og vurdere statistiske undersøkelser
- beregne og drøfte sentralmål og spredningsmål
- beregne og gjøre rede for kumulativ og relativ frekvens, representere data i tabeller og diagrammer og drøfte ulike dataframstillinger og hvilket inntrykk de kan gi
- gruppere data og beregne sentralmål for et gruppert datamateriale
- bruke regneark i statistiske beregninger og presentasjoner

Læringsmål

Etter at du har arbeidet med dette kapitlet skal du sette kryss i de boksene som tilhører de læringsmålene du har oppnådd. Det er viktig at du er ærlig og at du ikke krysser i de boksene som du føler at du ikke kan. På den måten vet du på hvilket område du må forbedre deg.

Etter dette kapitlet vet jeg

- hvordan jeg setter data inn i en tabell med frekvens, relativfrekvens, kumulativ frekvens og relativ kumulativ frekvens
- hvordan jeg tegner diagrammer for hånd
- hvordan jeg tegner diagrammer i Excel
- hvordan jeg finner sentralmål: gjennomsnitt, typetall og median
- hvordan jeg finner spredningsmål: variasjonsbredde og standardavvik
- hvordan jeg kan finne sentralmål og spredningsmål ved hjelp av Excel
- hvordan jeg finner gjennomsnitt i et klassesdelt materiale
- hvordan jeg kan finne hvilken klasse medianen er i et klassesdelt materiale

Etter dette kapitlet kan jeg forklare

- hvilken informasjon relativ frekvens og kumulativ frekvens gir
- hva ulike sentralmål forteller om et datamateriale
- hva ulike spredningsmål forteller om et datamateriale
- hva vi mener med et klassesdelt materiale

Etter dette kapitlet kan jeg vurdere og

- argumentere rundt og sammenligne praktiske situasjoner ut i fra tallverdien til sentralmål og spredningsmål
- velge hensiktsmessige hjelpemidler når jeg jobber med statistiske beregninger

Utforskende oppgave – Statistisk undersøkelse

I denne oppgaven skal du gjøre en undersøkelse blant elevene på skolen.

Spør elever i klassen om hvor mange skjermer (telefon, pc, tv, nettbrett etc.) de har hjemme.

Registrer svarene i tabellen under

Elev nr	Antall skjermer	Elev nr	Antall skjermer
1		26	
2		27	
3		28	
4		29	
5		30	
6		31	
7		32	
8		33	
9		34	
10		35	
11		36	
12		37	
13		38	
14		39	
15		40	
16		41	
17		42	
18		43	
19		44	
20		45	
21		46	
22		47	
23		48	
24		49	
25		50	

Hvordan kan du fremstille dette resultatet på en oversiktlig måte? Vis på neste side.

Hvor mange elever har 5 skjermer eller mindre hjemme?

Hvor mange prosent av elevene har 6 skjermer hjemme?

x Antall skjermer	f Frekvens	Relativ frekvens	Kumulativ frekvens	Relativ kumulativ frekvens	$x \cdot f$ Sum skjermer
Sum					

1. Hva er statistikk?

Statistikk handler om å trekke informasjon ut av et *datamateriale* og å framstille materialet på oversiktlige måter. Et datamateriale består av mange tall, og hvert tall kaller vi gjerne en *observasjon*. Eksempler på datamateriale:

- Standpunktkarakterene til alle elevene i 2P som var oppe til eksamen
- Høydene til alle som er på militærsesjon et år
- Antall mål en bestemt fotballspiller har scoret i hver kamp han har spilt
- Maksimumstemperaturen på Blindern hver dag i 2013

Statistikk deles inn i to hovedområder: **presentasjon** og **analyse**.

Presentasjon handler om å presentere observasjonene på en oversiktlig måte, enten via **frekvenstabell** eller **diagram**. Det er fire diagram du skal kunne lage: **stolpe/søylediagram**, **linjediagram**, **sektordiagram** og **histogram**.

Analyse går ut på å fortelle noe om observasjonene. Dette deles i to: **sentralmål (typetall, median og gjennomsnitt)** og **spredningsmål (variasjonsbredde og standardavvik)**.

Vi bruker karakterene i to 2P-grupper med tilsammen 50 elever som eksempel:

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

Dette ser uoversiktlig ut. Vi framstiller derfor tallene i en *frekvenstabell* og som et *diagram*.

2. Frekvenstabeller

2.1 Frekvens

En *frekvenstabell* viser hvor mange ganger hver dataverdi forekommer. I datamaterialet ovenfor er det seks dataverdier (de seks mulige karakterene), og det er 10 elever som har fått karakteren 3. Vi sier at *frekvensen* til karakteren 3 er lik 10. Andre ord for frekvens er *hyppighet* og *antall*.

Frekvenstabellen blir slik:

Karakter	Frekvens
1	9
2	16
3	11
4	8
5	5
6	1
Sum	50

Vi bør regne ut summen av alle seks frekvensene og sjekke at den blir lik antall observasjoner (her 50).

2. 1 Relativ frekvens

Ofte er det mer opplysende å finne ut hvor stor *del* av datamaterialet hver frekvens utgjør. Det oppgir vi i prosent og kaller det *relativ frekvens*.. Her er et eksempel på utregning:

$$\text{Relativ frekvens for karakteren 4: } \frac{8}{50} = 0,16 = 16 \%$$

Her er tabellen en gang til hvor vi har tatt med relative frekvenser:

Karakter	Frekvens	Relativ frekvens
1	9	18 %
2	16	32 %
3	11	22 %
4	8	16 %
5	5	10 %
6	1	2 %
Sum	50	100 %

Sjekk at summen av de relative frekvensene blir 100 %.

Oppgave 1

Fyll ut den relative frekvensen for antall skjermer elever har hjemme.

2. 2 Kumulativ frekvens

Det kan være interessant å slå sammen flere frekvenser. For eksempel kan det være nyttig for en lærer å vite hvor mange elever som har fått karakteren 1, 2 eller 3. Dersom vi legger sammen flere frekvenser kalles dette kumulativ frekvens.

Tabellen under er utvidet slik at den også viser kumulativ frekvens

Karakter	Frekvens	Relativ frekvens	Kumulativ frekvens
1	9	18 %	9
2	16	32 %	25
3	11	22 %	36
4	8	16 %	44
5	5	10 %	49
6	1	2 %	50
Sum	50	100 %	Skal ikke summeres

Dette betyr for eksempel at 36 av de 50 elevene fikk karakteren 3 eller lavere.

Oppgave 2

a) Hvorfor står det «skal ikke summeres» i den nederste ruta i kolonnen til kumulativ frekvens?

b) Fyll ut den kumulative frekvensen for antall skjermer elever har hjemme.

2. 3 Relativ kumulativ frekvens

Ønsker vi å vise den kumulative frekvensen i prosent kalles dette relativ kumulativ frekvens. Dette regnes ut på samme måte som relativ frekvens, bortsett fra at vi bruker den kumulative frekvensen.

Tabellen under er utvidet slik at den også viser relativ kumulativ frekvens.

Karakter	Frekvens	Relativ frekvens	Kumulativ frekvens	Relativ kum. fr.
1	9	18 %	9	18 %
2	16	32 %	25	50 %
3	11	22 %	36	72 %
4	8	16 %	44	88 %
5	5	10 %	49	98 %
6	1	2 %	50	100 %
Sum	50	100 %	Skal ikke summeres	

Dette betyr for eksempel at 72 % av de 50 elevene fikk karakteren 3 eller lavere.

Oppgave 3

Fyll ut den relative kumulative frekvensen for antall skjermer elever har hjemme.

2. 4 Typetall

Typetall er den observasjonen som forekommer flest ganger. I denne undersøkelsen vil det si hvilken karakter som forekommer oftest.

Dette finner vi ved å se på frekvenskolonnen. Det er flest som har fått karakteren 2. Det vil si at typetallet til denne undersøkelsen er 2.

Oppgave 4

Finn typetallet til antall skjermer elever har hjemme.

2. 5 Median

Medianen til et datamateriale er den observasjonen som står i midten når vi har sortert tallene i rekkefølge. Når antall observasjoner er et partall vil det være to observasjoner i midten.

Dersom disse er ulike må du legge sammen observasjonene og dividere på 2.

Vi finner hvilken observasjon som er i midten ved å bruke formelen

$$\frac{\text{Ant. observasjoner} + 1}{2}$$

Dersom svaret blir et desimaltall er antall observasjoner et partall, og det vil derfor være to observasjoner i midten.

Vi kunne ha laget en liste med de 50 karakterene, og noen velger å gjøre dette på prøver. Det kan imidlertid ta ganske lang tid, så det finnes heldigvis en raskere metode.

Mediankarakteren er den karakteren som står i midten når karakterene er skrevet i rekkefølge. I dette tilfellet er det 50 observasjoner

Observasjonen i midten vil derfor være nummer $\frac{50+1}{2} = 25,5$. Vi er derfor ute etter observasjon nummer 25 og 26. I kolonnen kumulativ frekvens ser vi at observasjon nummer 25 er karakter 2. Dermed må observasjon nummer 26 være karakter 3. For å finne medianen må vi legge dem sammen og dividere på 2:

$$\text{Median} = \frac{2+3}{2} = 2,5$$

Vi ville fått det samme svaret om vi hadde skrevet en liste med alle 50 karakterene, og det er bedre å gjøre det enn å ikke besvare oppgaven!

Oppgave 5

- a) For hvilke typer undersøkelser vil det ikke være mulig å finne medianen?
- b) Finn medianen til antall skjermer elever har hjemme.

2. 6 Variasjonsbredde

Variasjonsbredde viser forskjellen mellom den høyeste og den laveste observasjonen.

I denne undersøkelsen finner vi det ved å se på kolonnen for karakterer. Den høyeste karakteren er 6, og den laveste er 1.

$$\text{Variasjonsbredde} = 6 - 1 = 5.$$

Oppgave 6

- a) Hva ville variasjonsbredden vært dersom ingen elever fikk karakteren 1?
- b) For hvilke typer undersøkelser vil det ikke være mulig å regne ut variasjonsbredde?
- c) Finn variasjonsbredden til antall skjermer elever har hjemme

2.7 Gjennomsnitt

Med gjennomsnitt tenker vi at vi fordeler verdien til alle observasjonene slik at alle observasjonene får den samme verdien. Som en formel skrives det slik:

$$\frac{\text{sum verdier}}{\text{antall observasjoner}}$$

I dette datamaterialet tenker vi at alle skulle fått den samme karakteren. Vi må derfor legge sammen alle karakterene og dele på antall observasjoner (50).

Vi kunne ha laget en liste med de 50 karakterene, og noen velger å gjøre dette på prøver. Det kan imidlertid ta ganske lang tid, så det finnes heldigvis en raskere metode.

For å løse oppgaven raskere trenger vi en kolonne til, hvor vi multipliserer verdien karakterene med frekvensen. Da blir tabellen slik:

x Karakter	f Frekvens	Relativ frekvens	Kumulativ frekvens	Relativ kum. fr.	x · f Verdier
1	9	18 %	9	18 %	9
2	16	32 %	25	50 %	32
3	11	22 %	36	72 %	33
4	8	16 %	44	88 %	32
5	5	10 %	49	98 %	25
6	1	2 %	50	100 %	6
Sum	50	100 %			137

Vi har nå regnet ut at verdien til alle observasjonene er 137. For å finne gjennomsnittskarakteren må vi dividere dette på antall observasjoner (50).

$$\text{Gjennomsnitt} = \frac{137}{50} = 2,74 \approx 2,7$$

Oppgave 7

- For hvilke typer undersøkelser vil det ikke være mulig å finne gjennomsnittet?
- Hvor mange skjermer har hver elev i gjennomsnitt hjemme?

Oppgave 8

En klasse gjorde en undersøkelse på hvor mange søsken hver enkelt elev hadde. De samlet resultatet til følgende frekvenstabell:

x Ant. søsken	f Ant. elever				
0	2				
1	3				
2	5				
3	4				
4	4				
5	2				
Sum					

- Gjør nødvendige beregninger for å finne ut antall elever i denne klassen. Hvor mange søsken har flest av elevene? Med hvor mye varierte antall søsken i denne klassen?
- Utvid tabellen slik at du finner ut hvor mange prosent av elevene som hadde 3 søsken.
- Utvid tabellen slik at du finner ut medianen til dette datamaterialet.
- Utvid tabellen slik at du finner ut hvor mange prosent av elevene som hadde 3 søsken eller færre.
- Utvid tabellen slik at du finner ut hvor mange søsken hver elev hadde i gjennomsnitt.

Oppgave 9

En lærer målte temperaturen ved skolestart hver dag i mars, og presenterte resultatet gjennom tabellen nedenfor:

x Temperatur	f Ant. dager				
-1	3				
0	3				
1	4				
2	10				
Sum					

Gjør nødvendige beregninger og finn median, typetall, gjennomsnitt og variasjonsbredde til tallmaterialet ovenfor.

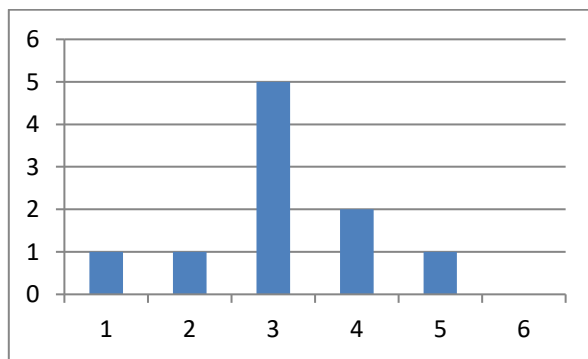
2.8 Standardavvik

Ofte gir ikke variasjonsbredden et godt bilde av spredningen. For eksempel kan vi tenke oss en 2P-gruppe hvor nesten alle elevene fikk karakter 3, mens en elev fikk 1 og en fikk 6. I en annen gruppe fikk to elever 1, fire fikk 2, fem fikk 3, fire fikk 4, tre fikk 5 og to fikk 6.

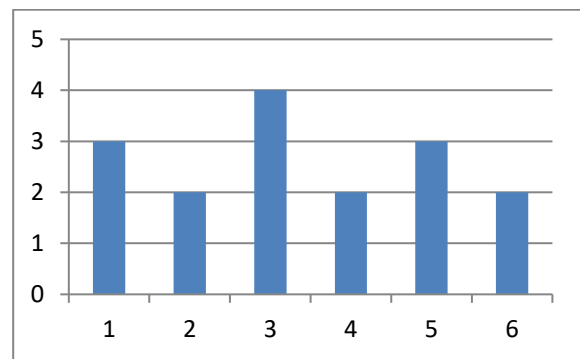
Variasjonsbredden er 5 i begge gruppene, men det er likevel naturlig å si at det er mer spredning i karakterene i den andre gruppen.

Standardavvik er et spredningsmål som sier noe om hvor “bred” en fordeling er. Det venstre diagrammet nedenfor viser et datamateriale med lite standardavvik og det høyre et med stort standardavvik.

Den nøyaktige definisjonen av standardavvik er litt vanskelig, så vi tar den ikke med her. Heldigvis er det bare aktuelt å beregne standardavvik i del 2-oppgaver, og da kan vi gjøre det ved å bruke regnearket Excel slik det er forklart i neste delkapittel.



Lite standardavvik.



Stort standardavvik.

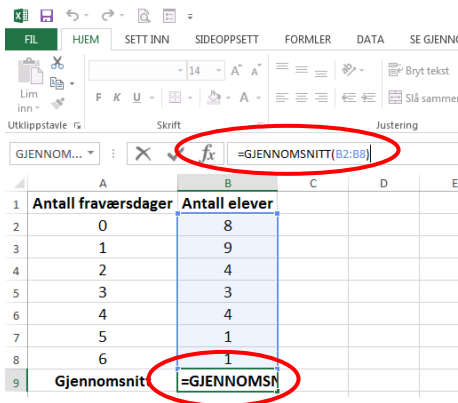
3. Beregning av sentralmål og spredningsmål i Excel

3.1 Når tallmaterialet er presentert i en liste

De sentral- og spredningsmål i 2P som kan være aktuelle å beregne i Excel er gjennomsnitt, median, typetall og standardavvik.

Viktig: alle formler og kommandoer må begynne med likhetstegn (=).

- I. Klikk på cellen som skal inneholde sentralmålet og sett inn et likhetstegn (=).
- II. Skriv inn ønsket kommando med parentestegn: **Gjennomsnitt(dataområde)** eller **Median(dataområde)** eller **Modus(dataområde)**, modus er det samme som typetall eller **STDAV.P.(dataområde)**, standardavvik.
- III. Bruk musepekeren og marker dataområdet til kommandoen og trykk *Enter*.



- Ikonet f_x på verktøylinjen er en hurtigfunksjon. Da får du likhetstegn inn i cellen. Og tilgang på alle Excel kommandoer i et nytt vindu.
- Vi kan bytte mellom å se formler og beregninger i Excel ved å taste (**Ctrl + `**). Apostrofe = (Shift + tasten bak pluss) **eller** Velg *fane-formler* deretter *verktøyikon-Vise formel*.
- På en prøve må du legge ved utskrift av både resultater og formler med navn på rader og på kolonner.

Eksempel 8

Vi har målt høyden til 7 jenter. Høydene i cm er: 177,164, 170, 168, 172, 161, 169.

Bildet under viser hvordan vi finner medianen, gjennomsnittet og standardavviket til disse høydene i Excel. Til venstre ser vi formlene som må skrives inn og til høyre hvordan resultatene blir.

	A	B
1		
2		Høyde i cm
3		177
4		164
5		170
6		168
7		172
8		161
9		169
10		
11	Antall:	= ANTALL(B3:B9)
12	Gjennomsnitt:	=GJENNOMSNIITT(B3:B9)
13	Standardavvik:	=STDAV.P(B3:B9)
14	Median:	=MEDIAN(B3:B9)

	A	B
1		
2		Høyde i cm
3		177
4		164
5		170
6		168
7		172
8		161
9		169
10		
11	Antall:	7
12	Gjennomsnitt:	168,7
13	Standardavvik:	4,8
14	Median:	169

Formelen for antall i B11 trenger vi egentlig ikke her. Den er bare tatt med for å vise at Excel lett kan telle opp antall celler i et område; her antall høyder.

Oppgave 10

Vi spurte 8 elever i kantina om hvor mye de hadde handlet for. De svarte at de hadde brukt følgende beløp (i kroner): 55, 70, 45, 60, 80, 50, 65 og 70.

Bruk Excel til å finne medianen, gjennomsnittet, typetallet og standardavviket til tallmaterialet.

3.2 Når tallmaterialet er presentert i en tabell

x Karakter	f Frekvens
1	9
2	16
3	11
4	8
5	5
6	1

Dersom du blir bedt om å finne gjennomsnittskarakteren og standardavviket til tallmaterialet presentert i tabellen til venstre fungerer ikke formlene vist ovenfor. På de neste sidene vises hvordan det må gjøres.

Gjennomsnitt:

	A	B	C	D
		x karakter	f frekvens	x · f
1				
2		1	9	9
3		2	16	32
4		3	11	33
5		4	8	32
6		5	5	25
7		6	1	6
8		Sum	50	137
9				
10	Gjennomsnitt (\bar{x})	2,7		

	A	B	C	D
		x karakter	f frekvens	x · f
1				
2		1	9	=B2*C2
3		2	16	=B3*C3
4		3	11	=B4*C4
5		4	8	=B5*C5
6		5	5	=B6*C6
7		6	1	=B7*C7
8		Sum	=SUMMER(C2:C7)	=SUMMER(D2:D7)
9				
10	Gjennomsnitt (\bar{x})	=D8/C8		

Vi må gjøre de samme beregningene som når vi regner gjennomsnitt for hånd. Se på formelutskriften til høyre, og se om du forstår hvilke regneoperasjoner som er utført.

Standardavvik:

	A	B	C	D	E
		x karakter	f frekvens	x · f	$(x - \bar{x})^2 \cdot f$
1					
2		1	9	9	27,2484
3		2	16	32	8,7616
4		3	11	33	0,7436
5		4	8	32	12,7008
6		5	5	25	25,538
7		6	1	6	10,6276
8		Sum	50	137	85,62
9					
10	Gjennomsnitt (\bar{x})	2,7			
11					
12	Standardavvik	1,3			

1. Vi regner ut avviket for hver enkelt karakter ved å regne: $(\text{karakter} - \text{gjennomsnitt})^2 \cdot \text{frekv.}$

2. Vi legger sammen alle avvikene.

(formler neste side)

3. Vi regner $\sqrt{\frac{\text{sum avvik}}{\text{sum frekvens}}}$

	A	B	C	D	E
1		karakter	frekvens	$x \cdot f$	$(x - \bar{x})^2 \cdot f$
2		1	9	=B2*C2	=(B2-\$B\$10)^2*C2
3		2	16	=B3*C3	=(B3-\$B\$10)^2*C3
4		3	11	=B4*C4	=(B4-\$B\$10)^2*C4
5		4	8	=B5*C5	=(B5-\$B\$10)^2*C5
6		5	5	=B6*C6	=(B6-\$B\$10)^2*C6
7		6	1	=B7*C7	=(B7-\$B\$10)^2*C7
8		Sum	=SUMMER(C2:C7)	=SUMMER(D2:D7)	=SUMMER(E2:E7)
9					
10	Gjennomsnitt (\bar{x})	=D8/C8			
11					
12	Standardavvik	=ROT(E8/C8)			

Oppgave 11

Bruk Excel til å finne gjennomsnitt og standardavvik for tallmaterialet i oppgave 8 og 9.

4. Klassedelt (gruppedelt materiale)

Når det er mange ulike dataverdier er det upraktisk å ta med alle i en frekvenstabell. Et eksempel er inntektsstatistikk hvor inntekten kan ha tusenvis av forskjellige verdier. Da grupperer vi verdiene i *klasser* (*grupper*) og finner frekvensen for hver klasse.

Her er et eksempel på en klassedelt frekvenstabell som viser inntekten til lønsmottakerne i en tenkt kommune. Inntektene er oppgitt i tusener av kroner.

Inntekt	Frekvens
[0, 100>	467
[100, 200>	678
[200, 300>	1490
[300, 400>	2653
[400, 500>	3785
[500, 750>	4106
[750, 1000>	987
[1000, 5000>	45
	$N = 14211$

Dette betyr for eksempel at det var 678 lønsmottakere som hadde en inntekt *fra og med* 100 000 kr og *inntil* 200 000. En inntekt på nøyaktig 200 000 kr faller i neste klasse, nemlig [200, 300>. Nederst i frekvenstabellen har vi regnet ut det totale antall lønsmottagere (N).

Legg merke til alle klassene ikke behøver å ha samme bredde. Bredden til klassen [300, 400> er 100, mens klassebredden til [1000, 5000> er 4000.

4.1 Gjennomsnitt i klassedelt materiale

Fordi vi ikke kjenner alle verdiene i et klassedelt materiale, går det ikke an å finne en *nøyaktig* verdi for gjennomsnittet. Det beste vi kan gjøre er å anta at *alle* verdiene i en klasse er lik verdien *midt i klassen*. Da utvider vi tabellen og regner slik:

Inntekt	Frekvens f	Midtpunkt x_m	$x_m \cdot f$
[0, 100>	467	50	23350
[100, 200>	678	150	101700
[200, 300>	1490	250	372500
[300, 400>	2653	350	928550
[400, 500>	3785	450	1703250
[500, 750>	4106	625	2566250
[750, 1000>	987	875	863625
[1000, 5000>	45	3000	135000
	$N = 14211$		Sum = 6694225

Gjennomsnittet blir da $6694225 : 14211 = 471$, som tilsvarer 471 000 kr.

4.2 Median i klassedelt materiale

Heller ikke medianen er det mulig å finne *nøyaktig* i et klassedelt materiale. For å finne en tilnærmet verdi lager vi først en tabell som viser kumulativ frekvens:

Inntekt	Frekvens	Kumulativ frekvens
[0, 100>	467	467
[100, 200>	678	1145
[200, 300>	1490	2635
[300, 400>	2653	5288
[400, 500>	3785	9073
[500, 750>	4106	13179
[750, 1000>	987	14166
[1000, 5000>	45	14211
	$N = 14211$	

Hvis alle inntektene var ordnet i stigende rekkefølge, ville medianen være inntekten på plass nr. $14212 : 2 = 7106$. Vi ser av kolonnen med kumulativ frekvens at denne inntekten ligger i klassen [400,500>. For å beregne en best mulig verdi for medianen må vi anta at alle de 3785 inntektene i denne klassen ligger *jevnt fordelt* innenfor klassen.

Oppgave 12

Vi målte høyden til elevene på VG2. Resultatet finner du i tabellen nedenfor.

Høyde cm	Frekvens
[150 – 170>	38
[170 – 180>	34
[180 – 200>	18

Finn medianhøyden og gjennomsnittshøyden til elevene

5. Diagrammer

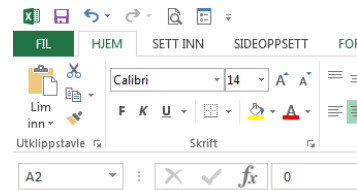
5.1 Søylediagram

Tallene i en frekvenstabell kan vi også framstille i et *diagram*. Det er mest aktuelt å gjøre dette i del 2 - oppgaver, og da kan vi bruke regnearket Excel. Men du bør også kunne tegne et diagram på papir i del 1.

Det er noen små forskjeller på menyer og kommandoer mellom Excel på Windows og Excel på Mac. Teksten beskriver Windows-versjonen. Hvis du bruker Mac får du et ark av læreren som beskriver forskjellene på de to versjonene.

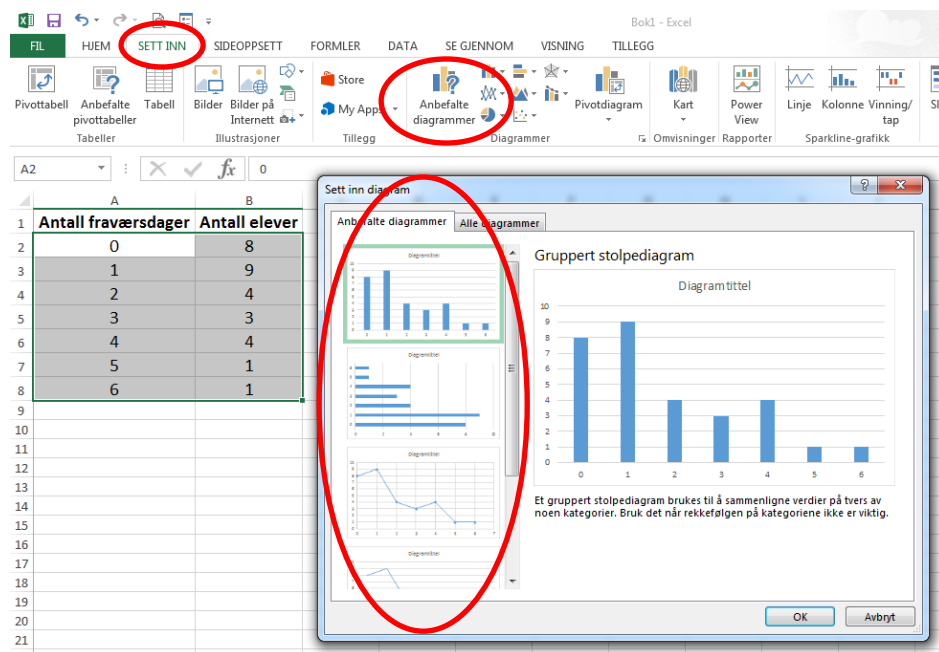
Vi bruker en tabell over elevfravær som eksempel:

Legg inn dataene som du vil lage søylediagram av i Excel. Merk ut med pekeplate eller mus dataområdet. Ikke ta med eventuell overskrift.



	A	B
1	Antall fraværsdager	Antall elever
2	0	8
3	1	9
4	2	4
5	3	3
6	4	4
7	5	1
8	6	1
9		
10		

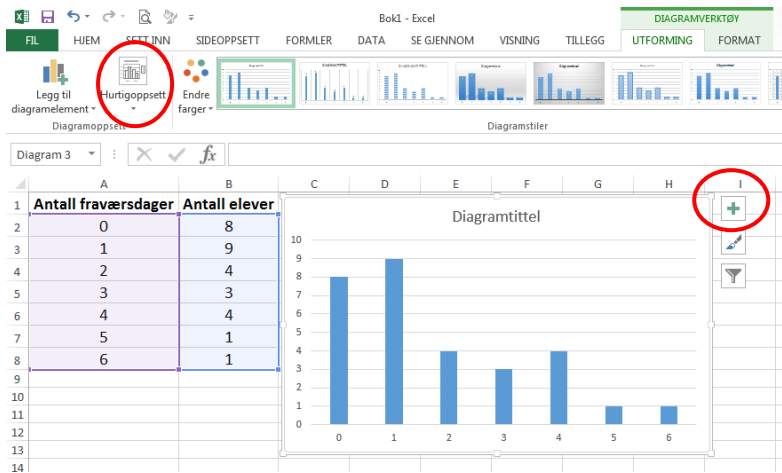
Velg fanen *Sett inn* og *Anbefalte diagrammer*. I nytt vindu velg ønsket diagram og klikk ok:



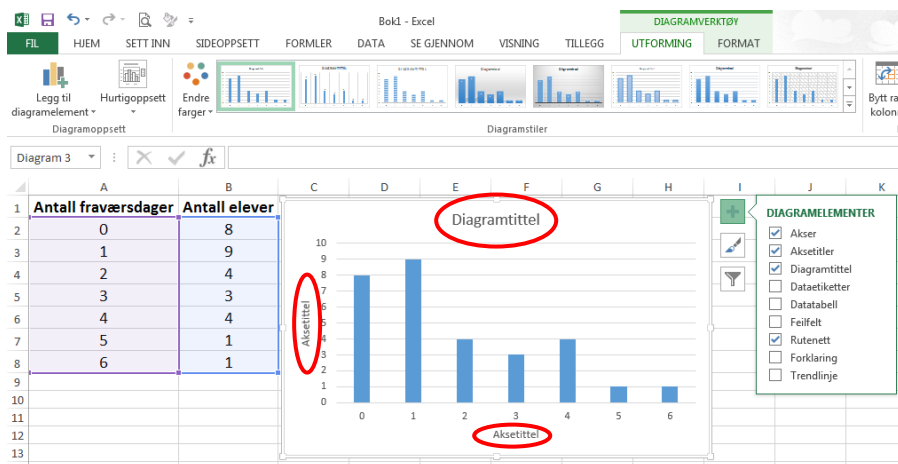
The screenshot shows the Excel interface with the 'SETT INN' tab selected on the ribbon. The 'Anbefalte diagrammer' button is circled in red. Below, the 'Sett inn diagram' dialog box is open, also with the 'Anbefalte diagrammer' tab selected. A 'Gruppert stolpediagram' is chosen, and its preview is shown. The dialog box includes an 'OK' button.

Formatering av diagrammet, diagramtittel og aksetitler:

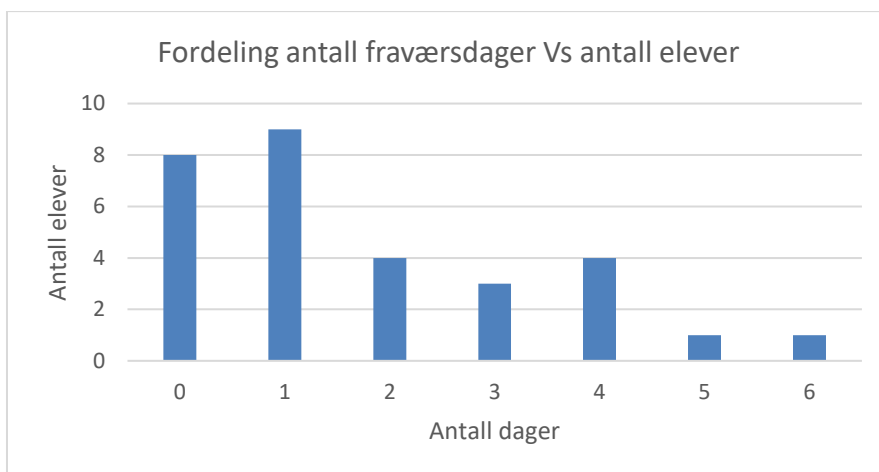
Velg enten fanen *hurtigoppsett* eller + tegnet til høyre på diagram vinduet.



Skriv inn ønsket tekst inn i feltene **diagramtittel** og **aksetittel**:



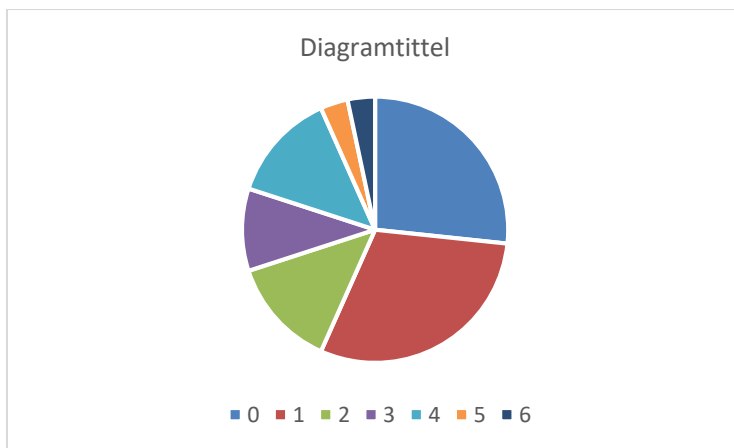
Diagrammet ferdig:



Av og til vil du lage diagrammer hvor det er mer enn en kolonne for hver dataverdi. Et eksempel kan være fraværet i flere klasser som skal sammenlignes. Da merker du bare ut alle kolonnene det skal lages søyler av. Resten blir som før.

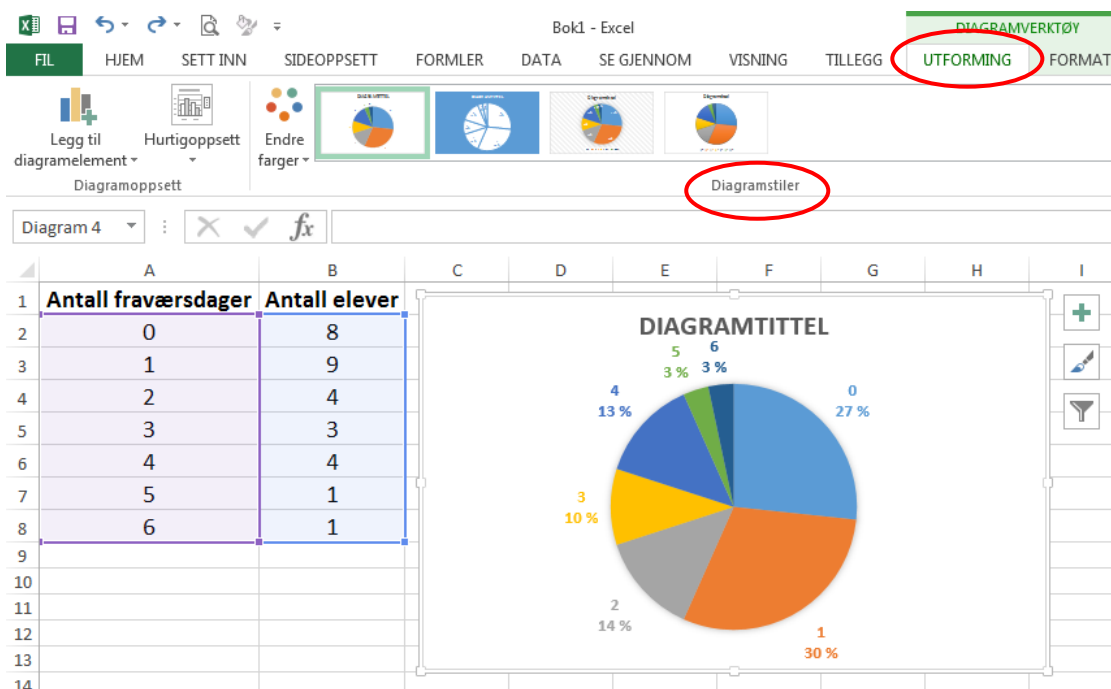
5.2 Sektordiagram

I mange tilfeller hvor man ikke har for mange dataverdier, er det vanlig å lage et *sektor-diagram* (populært kalt “kakediagram”). Vi går fram på samme måte som for et stolpediagram, men velger *Sektor* istedenfor *Stolpe*. Med tallene fra forrige eksempel får vi da dette diagrammet:

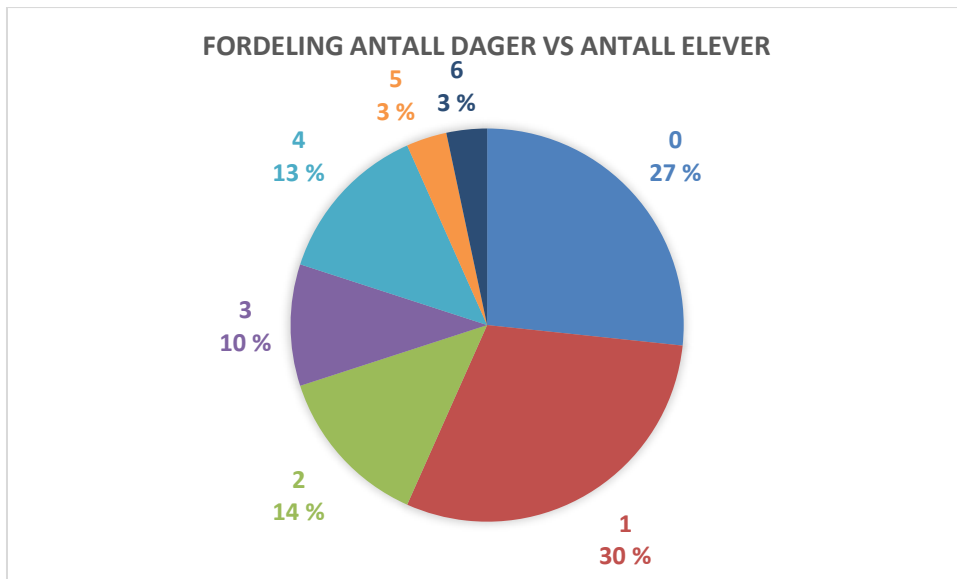


Det er en sektor for hver av de sju verdiene til antall fraværsdager. Disse verdiene er fargekodet, og koden står til høyre. I svart-hvitt er det vanskelig å se hva som er hva.

Vise den relative frekvensen i prosent for hver av sektorene. Velg fanen *utforming* og velg et av *diagramstilene*:



Til slutt skriver du inn en passende diagramtittel.



Det er mulig å lage mye “pynt” på diagrammene. Ikke bruk tid på det, i hvert fall ikke på prøver!

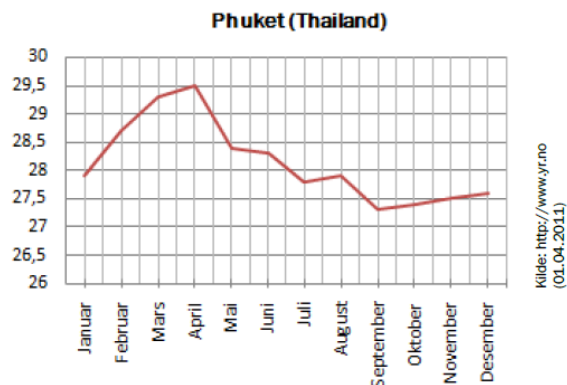
Et diagram skriver du ut på vanlig måte etter å ha klikket på det. Da får du bare med diagrammet, ikke hele regnearket.

Oppgave 13

Framstill frekvenstabellen på side 156 som et søylediagram og som et sektordiagram i Excel.

5.3 Linjediagram

Linjediagrammer brukes nesten bare for å vise hvordan noe utvikler seg over tid. Det betyr at på vannrett akse har vi som regel timer, dager, uker, måneder eller år. I Excel lager du et linjediagram på samme måte som et stolpediagram



5.4 Tegne sektordiagram for hånd

I del 1 kan du hende at du blir bedt om å tegne et sektordiagram på papir. Da vil det bare være noen få sektorer, og det vil være tall som skal være mulig å håndtere uten å være veldig god i hoderegning (for å tegne et bra sektordiagram trenger du passer, gradskive og linjal).

Eksempel

En del mennesker ble spurt om de var fornøyd med regjeringen. 50 % svarte “ja”, 25 % svarte “nei” og 25 % svarte “vet ikke”. (Dette er altså de relative frekvensene for dataverdiene.)

Vi vil lage et sektordiagram som illustrerer svarene. Vi regner da ut hvor mange grader hver av de tre sektorene må fylle. Hele sirkelen utgjør 360° .

Ja-sektoren må fylle $360^\circ \cdot 0,5 = 180^\circ$

Nei-sektoren må fylle $360^\circ \cdot 0,25 = 90^\circ$.

Vet ikke-sektoren må fylle $360^\circ \cdot 0,25 = 90^\circ$.

Vi lager en litt stor sirkel med passer og bruker gradskive til å lage de tre sektorene med riktig gradtall. Til slutt skriver vi passende tekst i hver sektor. Resultatet skal være omtrent slik:

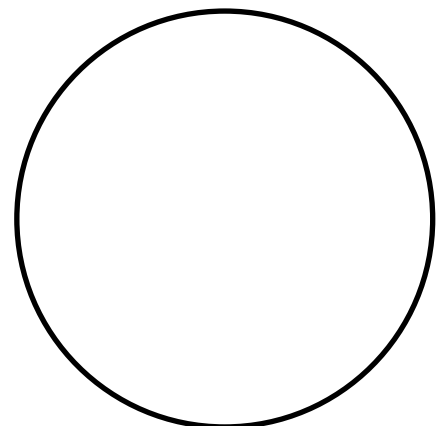


Tallene på del 1 er som regel gitt slik at sirkelen kan deles i relativt enkle brøker.

Oppgave 14

To klasser på til sammen 60 elever skal ha aktivitetsdag. 20 elever ønsker langrenn, 15 slalåm, 15 aking og 10 fottur. Lag et sektordiagram nedenfor (og uten kalkulator) som illustrere denne svarfordelingen.

Aktivitet	Grader
Langrenn	
Slalåm	
Aking	
Fottur	



5.5 Histogram

Når vi skal presentere resultatene fra et klassedelt materiale i et diagram, bruker vi histogram. Histogram ser ut som et søylediagram hvor søylen er tegnet inntil hverandre. Imidlertid er det størrelsen (arealet) på søylen som viser antall observasjoner (og ikke høyden, som ved et søylediagram). Vi må derfor regne ut både **klassebredden** og **søylehøyden**.

Klassebredden er forskjellen mellom nedre og øvre grense i klassen.

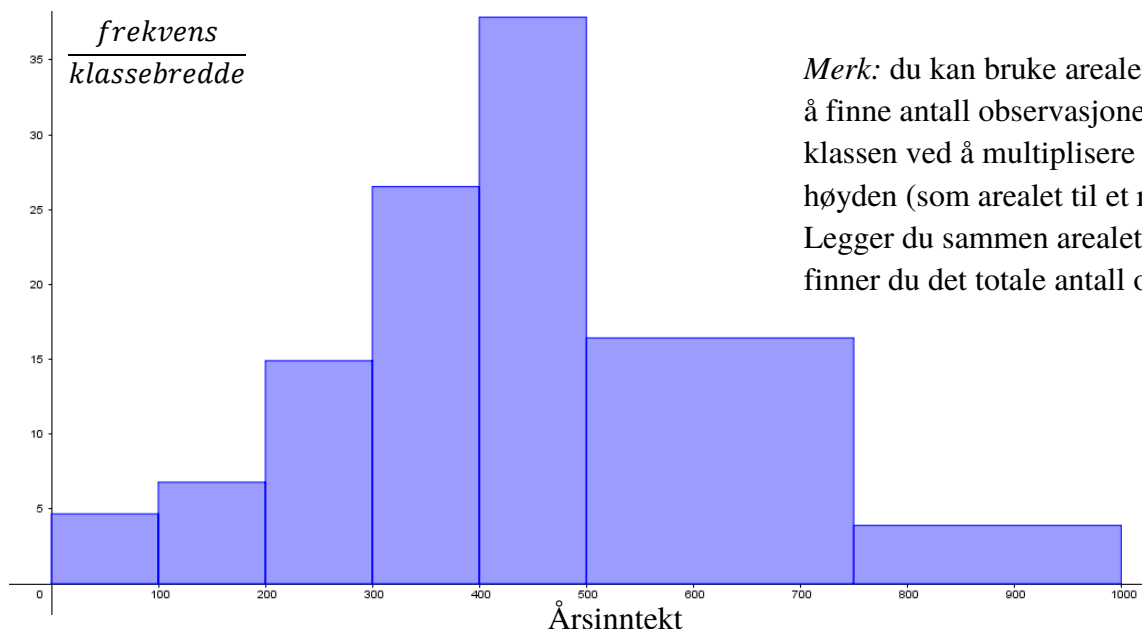
Søylehøyden er $\frac{\text{frekvens}}{\text{klassebredde}}$

Eksempel:

Fremstill resultatene fra undersøkelsen om årsinntekt i et histogram.

Inntekt	Frekvens	Klassebredde	Søylehøyde
[0, 100>	467	100	4,67
[100, 200>	678	100	6,78
[200, 300>	1490	100	14,9
[300, 400>	2653	100	26,53
[400, 500>	3785	100	37,85
[500, 750>	4106	250	16,43
[750, 1000>	987	250	3,9

Resultatet blir da slik:



Oppgave 15

Tegn et histogram til datamaterialet i oppgave 12

Utfyllingsoppgave – Diagrammer

I denne oppgaven skal vi utforske bruk av ulike diagrammer og se hvordan vi kan tegne dem for hånd og i Excel

Del 1

På eksamen kommer det ofte oppgaver hvor du skal tegne et «hensiktsmessig» diagram.

Fyll inn tabellen under slik at du får en oversikt over de ulike diagramtypene og når det er hensiktsmessig å bruke dem

Diagramtype	Skisse av diagramtypen	Når bruker vi denne diagramtypen?
Stolpediagram (søylediagram)		
Sektordiagram (kakediagram)		
Linjediagram (kurvediagram)		

Del 2

Karakter	Frekvens
1	9
2	17
3	10
4	8
5	5
6	1

Tabellen viser karakterfordelingen på en matematikkeksamen på en skole.

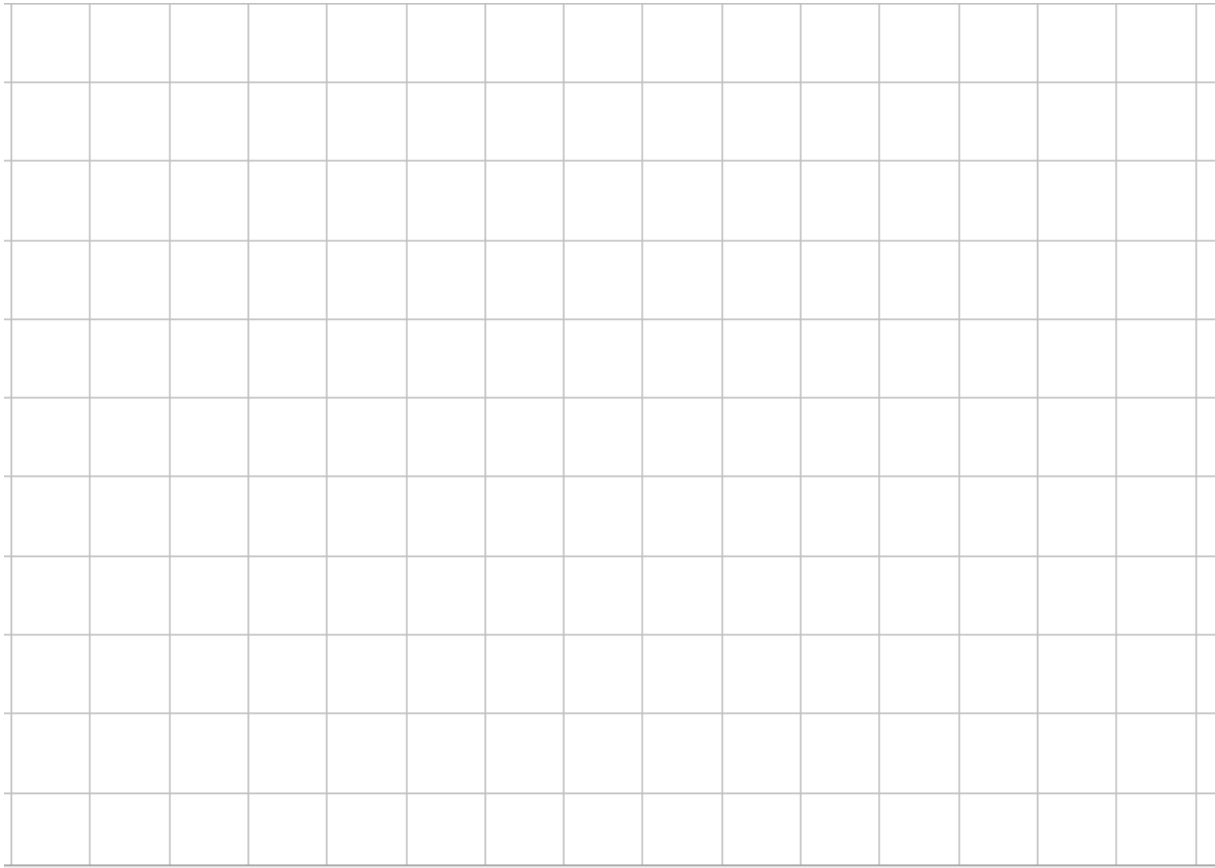
- a) Hvorfor vil sektordiagram kunne være hensiktsmessig å bruke her?
- b) Kunne man brukt et stolpediagram istedenfor? Begrunn svaret
- c) Hvorfor vil ikke et linjediagram være hensiktsmessig i denne situasjonen?
- d) Tegn et sektordiagram i Excel

(Lim inn Excelbildet her)

e) Fyll ut tabellen og tegn sektordiagrammet for hånd

Karakter	Frekvens	Gradtall
1		
2		
3		
4		
5		
6		

d) Tegn diagrammet for hånd.



e) Tegn diagrammet i Excel.

(Lim inn Excelbildet her)

Del 4

År	Antall kamper
1986	1
1987	2
1988	4
1989	10
1990	9
1991	6
1992	4
1993	9
1994	11
1995	11
1996	4

Tabellen viser antall spilte landskamper for en fotballspiller i perioden 1986-1996.

- Hvorfor er et linjediagram mest hensiktsmessig å bruke her?
- Tegn et linjediagram for hånd



c) Tegn et linjediagram i Excel.

(Lim inn Excelbildet her)

Del 5

I media kan vi av og til finne diagrammer som er ikke sier noe feil, men som fremstiller data på en måte som kan misforstås eller gi feil inntrykk. Kan du finne et slikt eksempel og lime inn her?

Hvorfor tror du diagrammet er laget på denne måten?

Eksamensoppgaver. Løsningsforslag finner du på ndla.no eller matematikk.net

V15 Oppgave 1 (del 1)

Dag	Temperatur
Mandag	4 °C
Tirsdag	10 °C
Onsdag	12 °C
Torsdag	5 °C
Fredag	6 °C
Lørdag	

Tabellen ovenfor viser hvordan temperaturen har variert i løpet av noen dager.

Hva må temperaturen være på lørdag dersom medianen av målingene skal bli 7 °C?

V15 - Oppgave 6 (del 1)

Alder	Frekvens
[20,30)	10
[30,40)	20
[40,50)	30
[50,70)	40

Tabellen ovenfor viser aldersfordelingen for lærerne ved en skole.

- Bestem gjennomsnittsalderen for lærerne ved skolen.
- Lag et histogram som viser aldersfordelingen for lærerne.
- Utvid tabellen med en kolonne som viser relativ frekvens, og en kolonne som viser kumulativ frekvens.

V15 - Oppgave 3 (del 2)

Tallene nedenfor viser temperaturen målt i grader celsius klokka 16 den 30. juni de siste 20 årene i by A.

20 18 20 19 19 21 20 22 22 18 17 18 22 19 21 20 22 22 21 17

- Bruk regneark til å bestemme gjennomsnitt og standardavvik for datamaterialet.
(Oppgaven fortsetter på neste side)

Tilsvarende data er samlet inn i by B. Gjennomsnittet her er 20,8 °C, og standardavviket er 3,4 °C.

Noen planlegger et større utearrangement 30. juni neste år og er avhengige av varmt vær. Arrangementet skal finne sted enten i by A eller i by B.

b) Hvilket råd vil du gi arrangørene ut fra de oppgitte dataene?

H15 - Oppgave 6 (del 1)

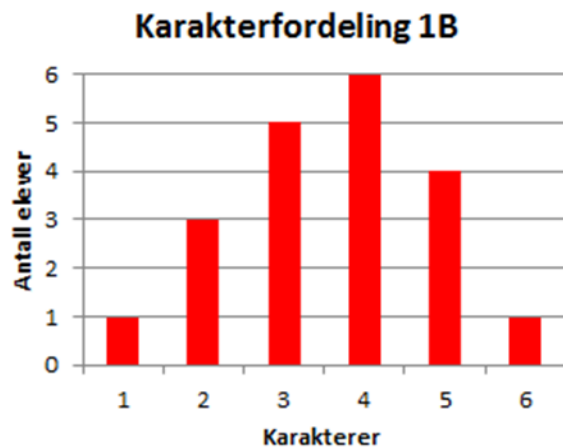
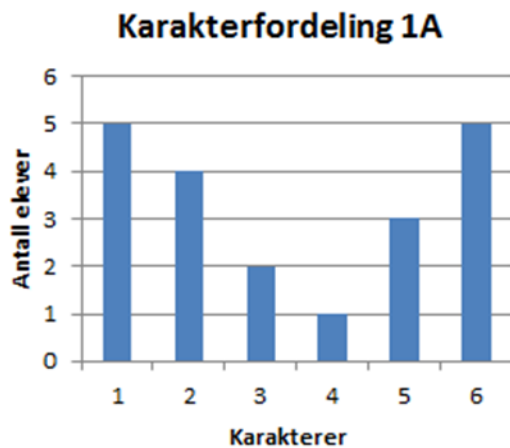
Alder	Bedrift A Frekvens	Bedrift B Frekvens
[20,40)	52	35
[40,60)	36	45
[60,70)	12	20
Sum	100	100

Hver av de to bedriftene A og B har 100 ansatte. Tabellen ovenfor viser aldersfordelingen for de ansatte i bedriftene.

a) I hvilken bedrift er medianalderen lavest? Grunngi svaret.

b) Bestem gjennomsnittsalderen for de ansatte i bedrift B.

H15 - Oppgave 9 (del 1)



(Oppgaven fortsetter på neste side)

Diagrammene på forrige side viser hvordan karakterene i klasse 1A og 1B fordelte seg ved forrige matematikkprøve.

- Bestem gjennomsnittskarakteren i hver av de to klassene.
- I hvilken klasse er standardavviket for karakterfordelingen størst? Grunngi svaret.
- Bestem den kumulative frekvensen for karakteren 3 i hver av de to klassene.
- Bestem den relative frekvensen for karakteren 6 i hver av de to klassene.

H15 - Oppgave 1 (del 2)

	Oslo	Østlandet for øvrig	Sør-Norge	Vestlandet	Midt-Norge	Nord-Norge
Elektriske ovner	59.7%	36.6%	30.4%	32.4%	34.2%	36.5%
Varmepumpe	8.3%	18.0%	25.5%	33.2%	33.8%	27.7%
Vannbåren varme	7.3%	9.3%	8.1%	6.4%	4.0%	5.0%
Sentralvarme	12.6%	4.8%	3.1%	2.8%	0.4%	7.5%
Vedfyring	5.3%	26.8%	28.3%	19.9%	19.6%	22.0%
Annet eller vet ikke	6.8%	4.5%	4.6%	5.3%	8.0%	1.3%

Tabellen ovenfor gir en oversikt over de viktigste varmekildene for husstander i ulike deler av Norge.

Bruk regneark til å lage **ett** diagram der du presenterer opplysningene i tabellen på en oversiktlig måte.

H15 - Oppgave 4 (del 2)



Figuren til venstre viser sterkeste middelvind ulike steder i Sør-Norge under ekstremværet «Nina» i januar 2015.

(Oppgaven fortsetter på neste side)

Vi lar den røde streken være skillet mellom Vestlandet og Sør-Østlandet.

- Bruk regneark til å bestemme gjennomsnitt og standardavvik for sterkeste middelvind på Vestlandet og sterkeste middelvind på Sør-Østlandet.
- Hva forteller svarene i oppgave a) om sterkeste middelvind på Vestlandet sammenliknet med sterkeste middelvind på Sør-Østlandet?

V16 - Oppgave 1 (del 1)

Dato	Temperatur
01.03	2°C
02.03	0°C
03.03	--4°C
04.03	--6°C
05.03	2°C
06.03	6°C

Guro målte temperaturen utenfor hytta de seks første dagene i mars. Se tabellen ovenfor.

Bestem variasjonsbredden, gjennomsnittet og medianen for temperaturmålingene.

V16 - Oppgave 5 (del 1)

Alder	Frekvens
[0,10)	40
[10,20)	20
[20,30)	60
[30,50)	20
[50,60)	20
[60,80)	40
Sum	200

Tabellen ovenfor viser aldersfordelingen for de 200 personene som bor i blokk Z på Tirilltoppen.

- Lag et histogram som viser aldersfordelingen for personene som bor i blokk Z.
- Bestem gjennomsnittsalderen for personene som bor i blokka.

(Oppgaven fortsetter på neste side)

Aurora bor i blokk Z. Hun er 32 år. Hun vet at de yngste i blokka er nyfødte, og at den eldste er 79 år. Hun påstår derfor at hennes alder er lavere enn medianverdien.

c) Vurder om Auroras påstand er riktig.

V16 - Oppgave 9 (del 1)

Antall land	Frekvens	Relativ frekvens	Kumulativ frekvens
$[1,6)$	5		
$[6,11)$			15
$[11,16)$	2	0,1	
$[16,21)$			19
$[21,26)$			20

Ole har undersøkt hvor mange land hver elev i en 2P-gruppe har besøkt. Han har satt opp en tabell. Ovenfor ser du noen av tallene i tabellen.

Tegn av tabellen, gjør beregninger, og fyll inn tallene som mangler.

V16 - Oppgave 1 (del 2)

Ved en skole er det 440 elever. Elevene blir spurt om hvor ofte de bruker sykkelhjelm. Tabellen nedenfor viser resultatene.

Alltid	88
Nesten alltid	176
Noen ganger	110
Aldri	22
Sykler ikke	44

Bruk regneark til å lage et sektordiagram som illustrerer opplysningene i tabellen ovenfor. Det skal gå klart fram av diagrammet hvor mange prosent hver sektor utgjør.

V16 - Oppgave 2 (del 2)

Hans og Grete går til Høgfjell hver dag. Nedenfor ser du hvor mange minutter Hans har brukt på hver tur de to siste ukene.

25 30 26 24 32 25 27 30 28 31 24 35 32 33

a) Bestem gjennomsnitt og standardavvik for datamaterialet.

Grete har i gjennomsnitt brukt like lang tid som Hans per tur de siste 14 dagene, men standardavviket hennes er 1,2.

b) Hva kan du ut fra dette si om tidene Grete har brukt på turene, sammenlignet med tidene Hans har brukt?




H16 - Oppgave 7 (del 1)

Sondre solgte frukt og grønnsaker på torget 20 lørdager i løpet av 2016. Hver av de 20 lørdagene skrev han opp hvor mange kunder han hadde. Han laget også en tabell. Tabellen ser du nedenfor, men her mangler noen av tallene Sondre satte inn.

Antall kunder	Frekvens	Relativ frekvens	Kumulativ frekvens
$[0, 50)$		0,05	
$[50, 100)$			6
$[100, 150)$	8		
$[150, 200)$			20

a) Tegn av tabellen ovenfor, og fyll inn tallene som mangler.
Gjør beregninger eller forklar hvordan du tenker.

Nedenfor ser du listen der Sondre har skrevet opp hvor mange kunder han hadde hver av de 20 lørdagene. Tre av tallene er skjult under flekker.

116 100  125 185 125 150  60 75
120 50 140 175  88 133 96 105 169

b) Foreslå tre *mulige* tall som kan stå under de tre flekkene slik at de 20 verdiene ovenfor gir resultatene i tabellen.

H16 - Oppgave 9 (del 1)

Tabellen nedenfor viser hvor mange poeng hver av de 30 elevene i en 2P-gruppe fikk på en matematikkprøve.

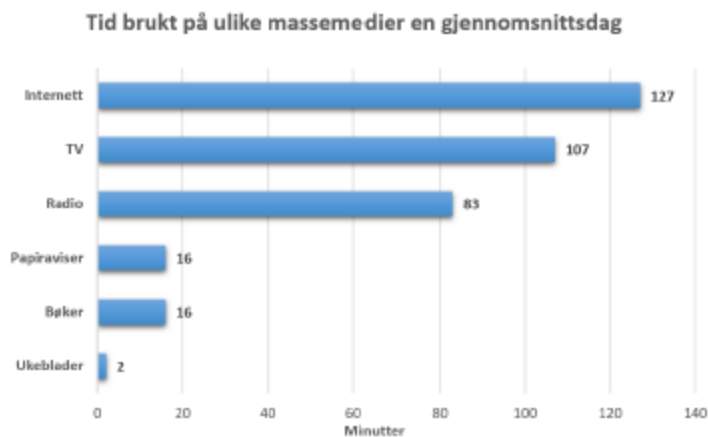
Poeng	Antall elever (frekvens)
[0, 5)	4
[5, 10)	12
[10, 15)	10
[15, 20)	0
[20, 25)	4

- a) Bestem gjennomsnittet for det klassedelte datamaterialet.

Per var en av elevene som hadde prøven. Han fikk 10 poeng og påstår at han var blant den beste halvdel av elevene.

- b) Kan Per bruke medianen for datamaterialet til å begrunne påstanden sin?

H16 - Oppgave 1 (del 2)



Diagrammet ovenfor viser hvor mange minutter personer i Norge brukte på ulike massemedier en gjennomsnittsdag i 2015.

Lag et sektordiagram som viser hvor stor andel av tiden som ble brukt på hvert av de ulike massemediene.

H16 - Oppgave 4 (del 2)

Klasse 2A har hatt matematikkprøve. De 15 elevene i klassen fikk disse poengsummene:

31 8 24 9 24 14 26 13
26 4 13 27 12 20 28

- a) Bestem gjennomsnittet og medianen for poengsummene.

Påstand 1: Gjennomsnittet er 7,0 % lavere enn medianen.
Påstand 2: Medianen er 7,5 % høyere enn gjennomsnittet.

- b) Avgjør om hver av påstandene ovenfor er riktige.
c) Bestem standardavviket for poengsummene.

15 elever fra klasse 2B har hatt samme prøve. I denne klassen ble gjennomsnittet for poengsummene 18,6 og standardavviket 5,9.

- d) Hva kan du si om poengsummene i 2B sammenliknet med poengsummene i 2A?

V17 - Oppgave 1 (del 1)

I en klasse er det 16 elever. Tabellen nedenfor viser hvor mange søsken de 16 elevene har.

Antall søsken	Frekvens
0	5
1	6
2	2
3	2
4	1

Bestem gjennomsnittet, medianen, typetallet og variasjonsbredden.

V17 - Oppgave 6 (del 1)

Et år deltok 1 000 elever i en konkurranse. Besvarelsene ble vurdert, og lærerne laget en tabell. Tabellen ser du nedenfor, men her mangler noen av tallene lærerne satte inn

Poengsum	Frekvens	Relativ frekvens	Klassemidtpunkt
$[0, 30)$	100		
$[30, 50)$			
$[50, 70)$		0,6	
$[70, 100)$	200		

- a) Tegn av tabellen ovenfor, og fyll inn tallene som mangler.
- b) Bestem gjennomsnittlig poengsum for elevene som deltok i konkurransen.

Et annet år deltok 3 525 elever i konkurransen. Tabellen nedenfor viser poengfordelingen

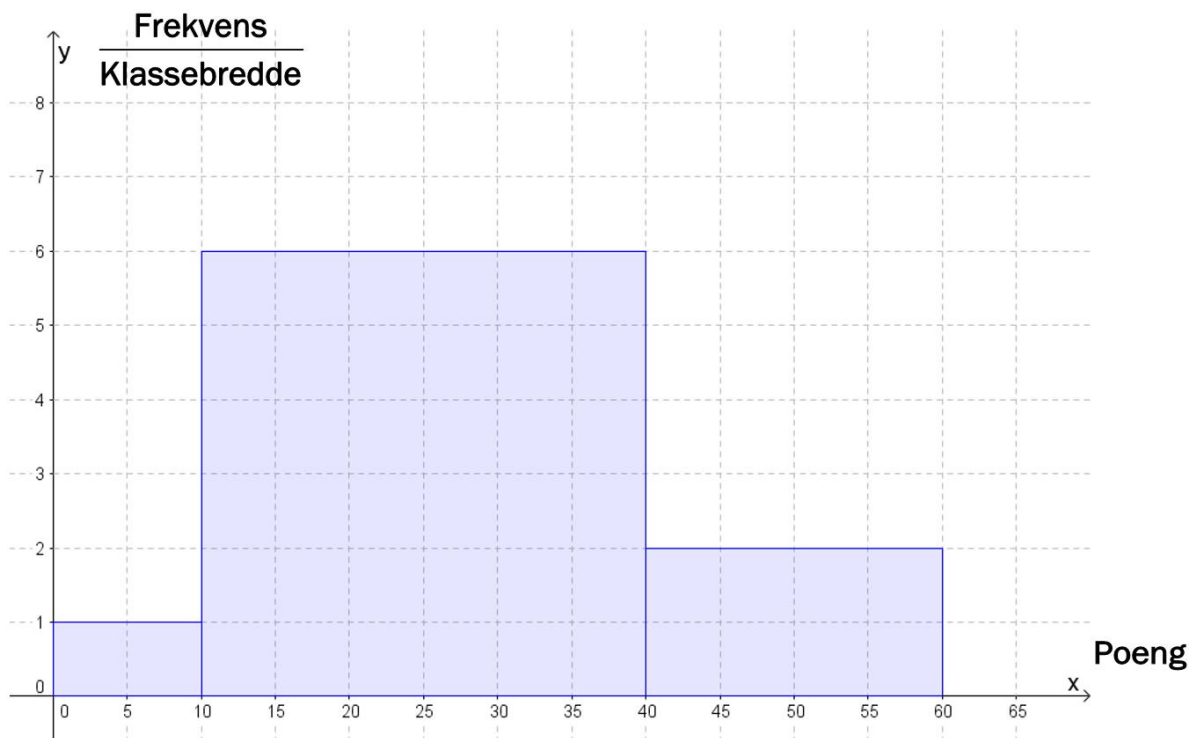
Poengsum	Frekvens
$[0, 30)$	563
$[30, 50)$	700
$[50, 70)$	2000
$[70, 100)$	262

- c) Bestem medianen for poengsummene til elevene som deltok i konkurransen dette året.

V17 - Oppgave 5 (del 2)

Ved en skole kom alle elevene som hadde valgt 2P, opp til skriftlig eksamen. Histogrammet på neste side viser poengfordelingen

(Oppgaven fortsetter på neste side)



a) Vis at det til sammen var 230 elever i 2P-gruppene.

b) Bestem gjennomsnittlig poengsum for elevene.

V17 - Oppgave 8 (del 2)

Liverpool FC	
Antall mål per kamp	Frekvens
0	8
1	14
2	7
3	4
4	3
5	1
6	1

Newcastle United FC	
Antall mål per kamp	Frekvens
0	14
1	13
2	7
3	2
4	0
5	1
6	1

(Oppgaven fortsetter på neste side)

Tabellene ovenfor viser hvor mange mål Liverpool FC og Newcastle United FC skåret per kamp i sesongen 2015–2016

a) Bestem gjennomsnittet og medianen for antall skårede mål per kamp for begge klubbene.

b) Bestem standardavviket for antall skårede mål per kamp for begge klubbene.

Hva forteller dette oss?

H17 - Oppgave 1 (del 1)

Tabellen nedenfor viser karakterfordelingen ved en skole ved norskeksamen våren 2017.

Karakter	Antall elever
1	3
2	12
3	25
4	12
5	6
6	2

d) Hvor mange prosent av elevene fikk karakteren 1 eller 2?

e) Bestem mediankarakteren.

f) Bestem gjennomsnittskarakteren.

H17 - Oppgave 4 (del 1)

Et idrettslag har 240 medlemmer. Idrettslaget har fire forskjellige aktivitetsgrupper.

Medlemmene fordeler seg slik:

Aktivitetsgruppe	Antall medlemmer
Langrenn	60
Hopp	40
Freestyle	80
Alpint	60

Gjør beregninger og lag et sektordiagram som viser fordelingen av medlemmene på de ulike gruppene.

Det skal gå klart fram hvor mange grader hver av sektorene i diagrammet er på.

H17 - Oppgave 3 (del 2)

Maskin A og maskin B fyller vann på flasker. I hver flaske skal det være 500 mL vann.

Anders måler hvor mye vann det er i 20 av flaskene fra maskin A. Nedenfor ser du resultatene.

510 mL	490 mL	495 mL	480 mL	520 mL
500 mL	504 mL	508 mL	501 mL	516 mL
498 mL	485 mL	499 mL	502 mL	515 mL
505 mL	497 mL	500 mL	493 mL	516 mL

- a) Bestem gjennomsnittet og standardavviket for antall mL vann i disse 20 flaskene.

Anders måler også hvor mye vann det er i 20 flasker fra maskin B. Han regner ut at gjennomsnittet er det samme som for maskin A men at standardavviket er 2,5 mL.

- b) Hva kan vi si om de 20 flaskene fra maskin B sammenliknet med de 20 flaskene fra maskin A ut fra disse beregningene?

H17 - Oppgave 5 (del 2)

I en klasse på Vg2 idrettsfag er det 30 elever. Tabellen nedenfor viser hvor mye elevene trener utenom skoletiden i løpet av en uke.

Antall minutter	Antall elever	Kumulativ frekvens	Relativ frekvens	Kumulativ relativ frekvens
[0, 60)	3			
[60, 180)	6			
[180, 300)	12			
[300, 420)	6			
[420, 540)	3			

- c) Tegn av tabellen i besvarelsen din, og fyll inn verdier for kumulativ frekvens, relativ frekvens og kumulativ relativ frekvens.

(Oppgaven fortsetter på neste side)

- d) Lag et histogram som viser hvor mye elevene trener utenom skoletiden.
- e) Bestem gjennomsnittet for det klassesdelte datamaterialet.
- f) Bestem medianen for det klassesdelte datamaterialet.

V18 - Oppgave 1 (del 1)

Markus og vennene hans spiller kort. Nedenfor ser du hvor mange poeng Markus fikk i hver av de siste åtte rundene.

Runde	Poengsum Markus
1	20
2	-15
3	5
4	15
5	-8
6	-3
7	-24
8	30

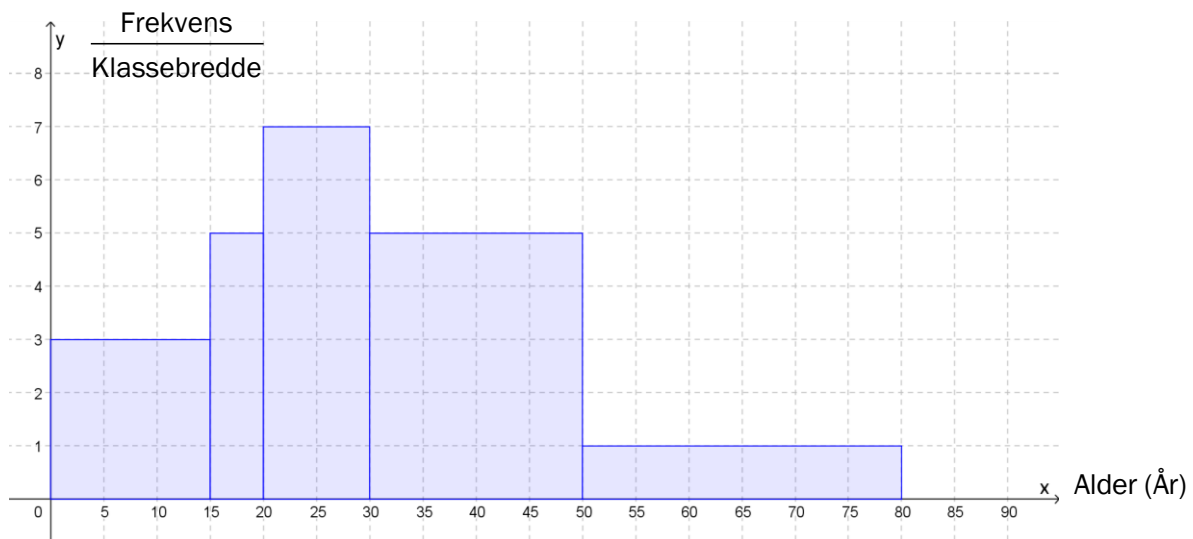
Bestem variasjonsbredden og gjennomsnittet for poengsummene.

V18 - Oppgave 5 (del 1)

Per og Kari vil lage et diagram som viser aldersfordelingen til innbyggerne i et boligområde. De diskuterer om de skal bruke et histogram eller et søylediagram.

Ut fra opplysningene de har fått, lager Per histogrammet nedenfor. Innbyggerne er delt inn i fem aldersgrupper.

(Oppgaven fortsetter på neste side)



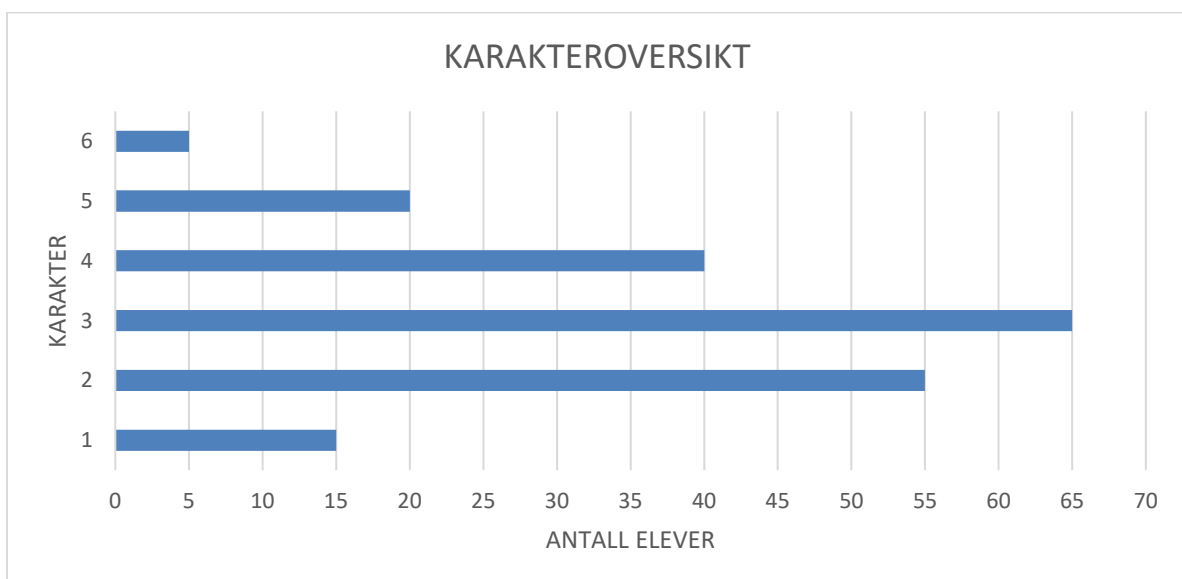
a) Hvor mange personer bor i boligområdet?

Kari lurer på om et søylediagram vil være bedre egnet.

b) Lag et søylediagram som viser hvor mange personer det er i hver aldersgruppe.

Mener du et søylediagram eller et histogram er best egnet til å illustrere dette datamaterialet?

V18 - Oppgave 7 (del 2)



Diagrammet ovenfor viser karakterfordelingen ved en matematikkeksamen et år.

(Oppgaven fortsetter på neste side)

- a) Hvor mange prosent av elevene fikk karakteren 4 eller bedre?
- b) Lag et regneark som vist nedenfor. Legg inn verdier i de hvite cellene og formler i de blå cellene. Bruk regnearket til å bestemme gjennomsnittskarakteren og standardavviket til karakterfordelingen.

	A	B	C	D
1	Karakter	Frekvens		
2	x	f	$x \cdot f$	$(x - \bar{x})^2 \cdot f$
3	1			
4	2			
5	3			
6	4			
7	5			
8	6			
9	Sum			
10				
11	Gjennomsnitt (\bar{x})			
12				
13	Standardavvik			

Året etter var det 180 elever som hadde eksamen. Gjennomsnittskarakteren dette året var 3,25.

- c) Hva var gjennomsnittskarakteren dersom vi ser disse to årene under ett?