

Mål for kapittel 6. 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

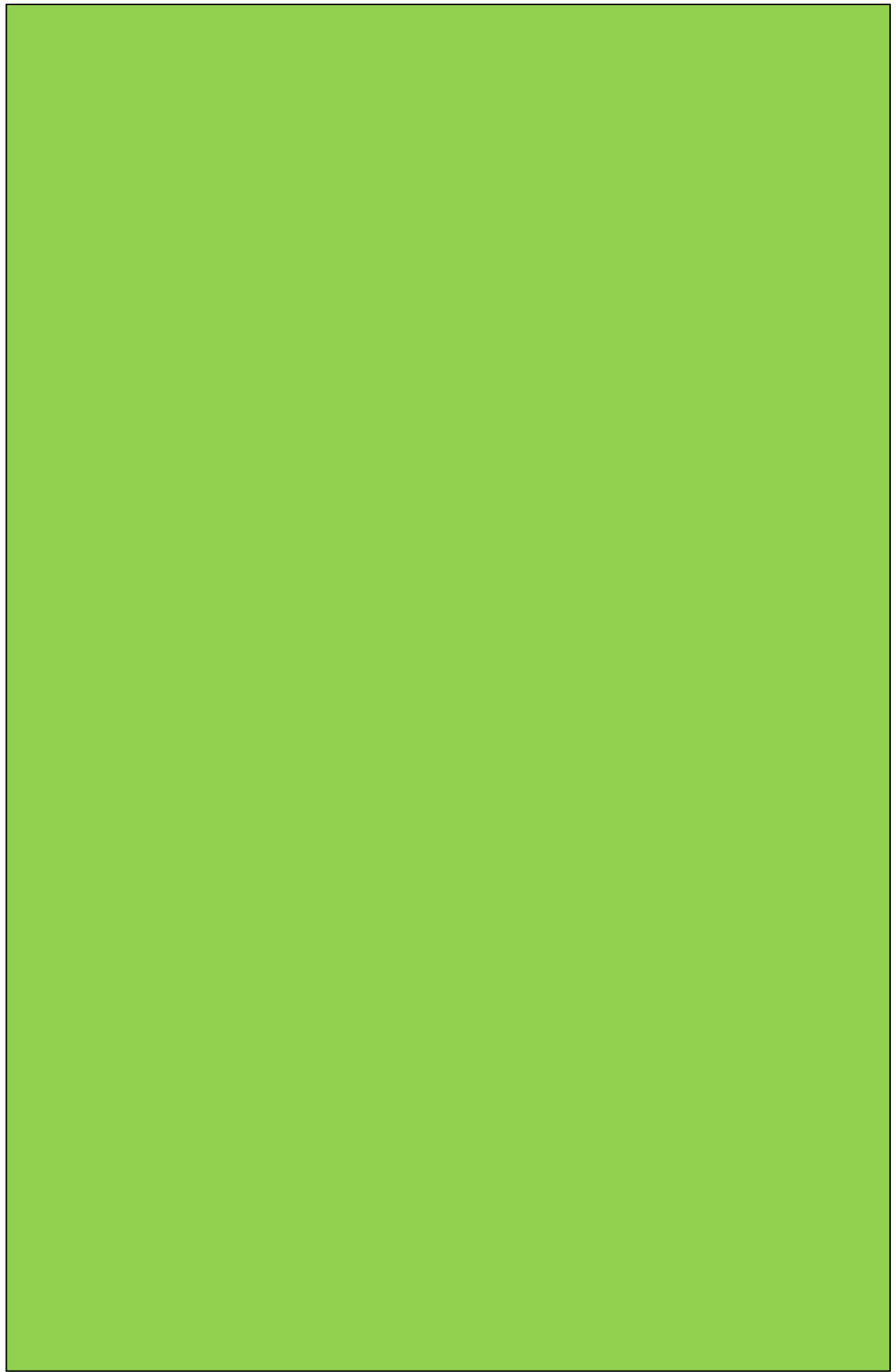
Utforskende oppgave – Statistisk undersøkelse

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

Spør elever i kantina en dag om hvor mange filmer de ser i løpet av en uke.

Registrer svarene i tabellen under

Elev nr	Antall filmer	Elev nr	Antall filmer
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	
--		--	



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

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

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. Et annet ord for frekvens er *hyppighet*.

Frekvenstabellen blir slik:

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

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

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	17	34 %
3	10	20 %
4	8	16 %
5	5	10 %
6	1	2 %

Sjekk at summen av de relative frekvensene blir 100 %.

Tabellen under er utvidet slik at den også viser *kumulativ frekvens* og *relativ kumulativ frekvens*. Ordet kumulativ betyr “oppsamlet”.

Karakter	Frekvens	Relativ frekvens	Kumulativ frekvens	Relativ kum. fr.
1	9	18 %	9	18 %
2	17	34 %	26	52 %
3	10	20 %	36	72 %
4	8	16 %	44	88 %
5	5	10 %	49	98 %
6	1	2 %	50	100 %

Dette betyr for eksempel at 36 av de 50 elevene fikk 3 eller dårligere og at dette utgjør 72 % av elevene.

Oppgave 1

Løs denne oppgaven uten kalkulator.

Noen elever ble spurt om hvor mange Pcer, nettbrett og mobiler det til sammen var i familien. De ga følgende svar:

3 5 6 4 4 7 4 4 7 5

Framstill disse dataene i en tabell som viser frekvenser, kumulative frekvenser, relative frekvenser og relative kumulative frekvenser.

3. Diagrammer

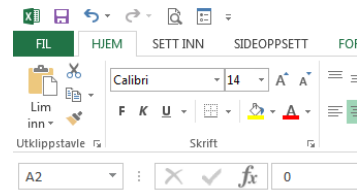
3.1 Søylediagram

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.

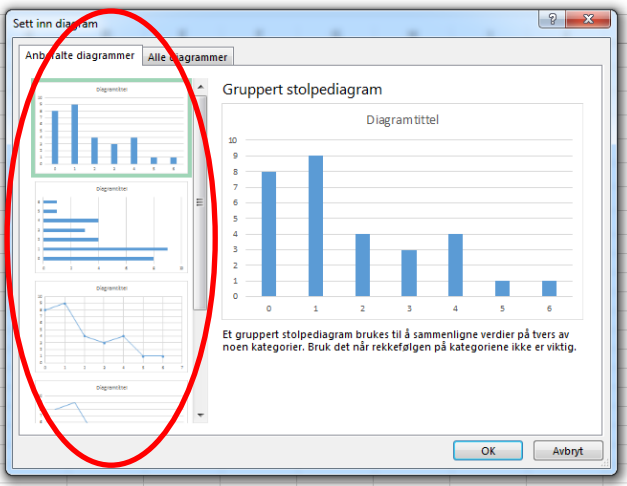

Vi bruker en tabell over elevfravær som eksempel:

Legg inn dataene som du vil lage søyle-diagram 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		

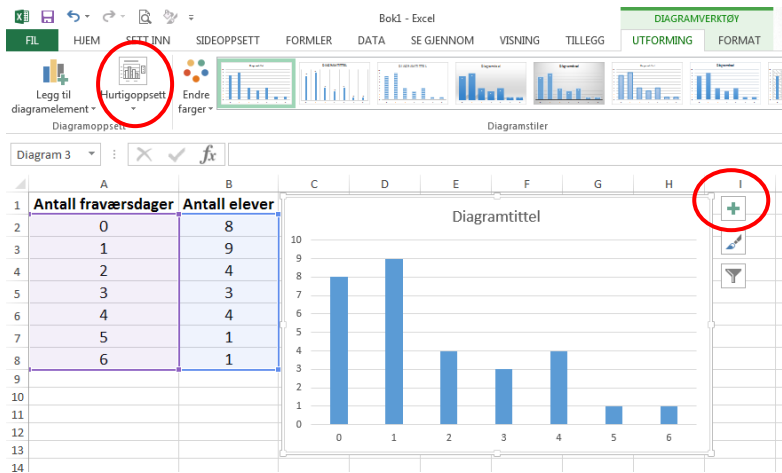
Velg fanen *Sett inn* og *Anbefalte diagrammer*. I nytt vindu velg ønsket diagram og klikk ok:



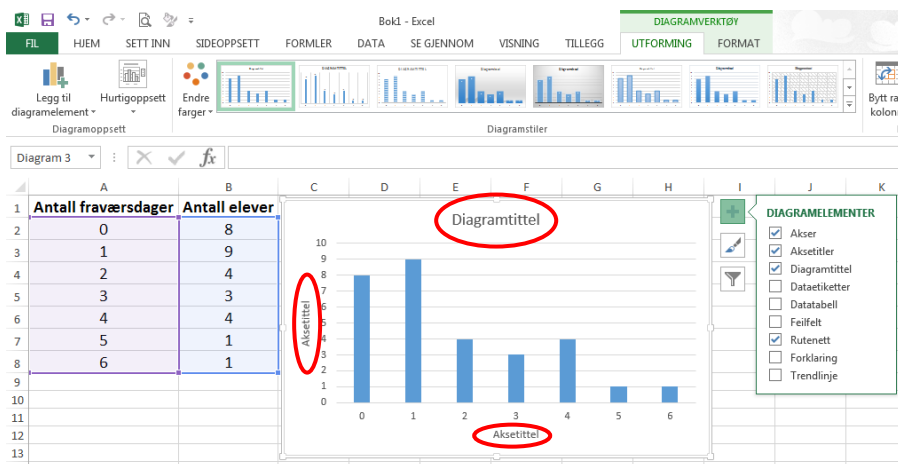
The dialog box 'Sett inn diagram' is open, showing a preview of a grouped bar chart. The chart has a title 'Gruppert stolpediagram' and a subtitle 'Diagramtittel'. The x-axis is labeled with values 0 through 6, and the y-axis is labeled with values 0 through 10. A note at the bottom of the dialog states: 'Et gruppert stolpediagram brukes til å sammenligne verdier på tvers av noen kategorier. Bruk det når rekkefølgen på kategoriene ikke er viktig.'

Formatering av diagrammet, diagramtittel og aksetitler:

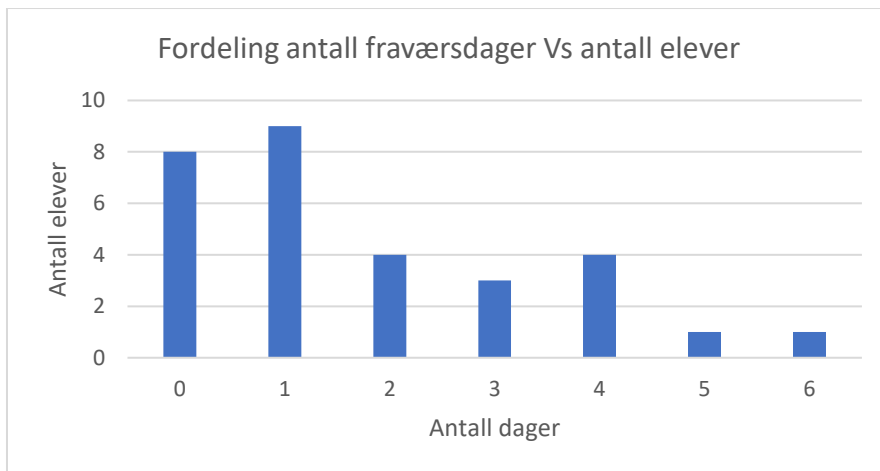
Velg enten fanen *hurtigoppsett* eller + tegnet til høyre på diagram vinduet.



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



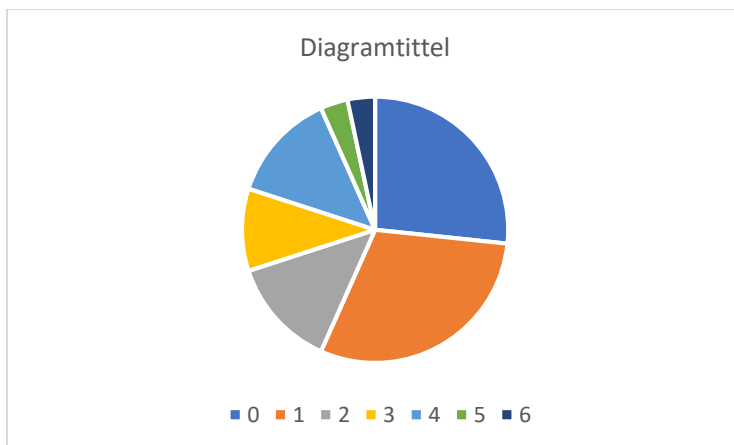
Diagrammet ferdig:



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.

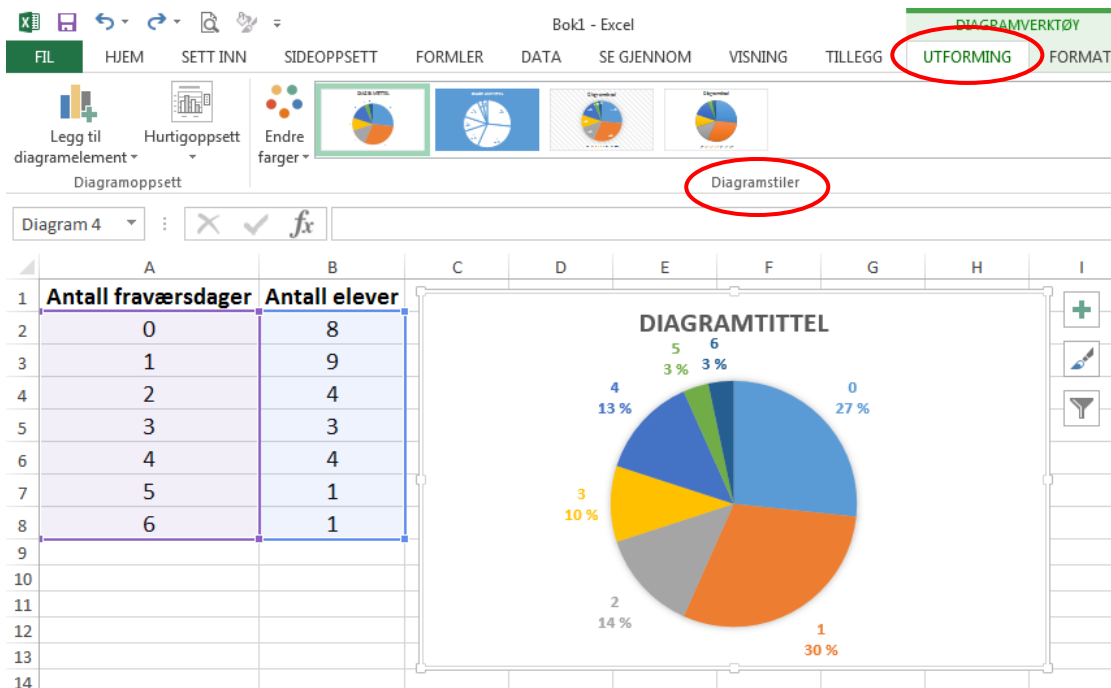
3.2 Sektordiagram

I mange tilfelle 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:

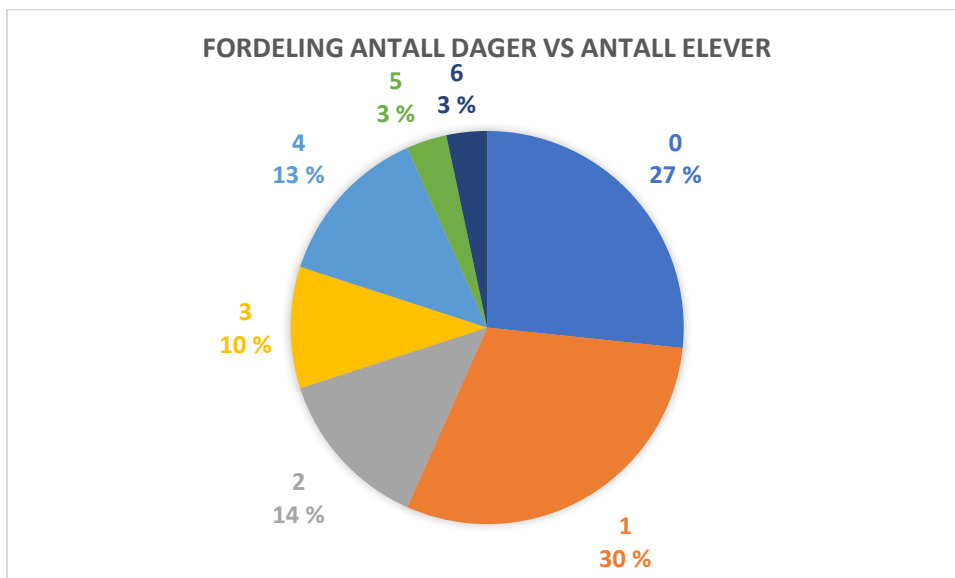


Det er en sektor for hver av de sju verdiene til antall fraværsdager. Disse verdiene er fargekodet, og koden står til høyre. I svart-hvitt er det vanskelig å se hva som er hva.

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.

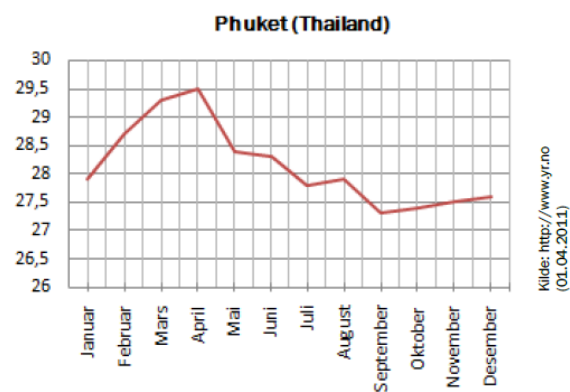
Oppgave 2

Framstill frekvenstabellen i oppgave 1 som et søylediagram og som et sektordiagram.

3.3 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.





3.4 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 1

En del mennesker ble spurt om de var fornøyd med regjeringen. 60 % svarte “ja”, 30 % svarte “nei” og 10 % svarte “vet ikke”. (Dette er altså de relative frekvensene for dataverdiene.)

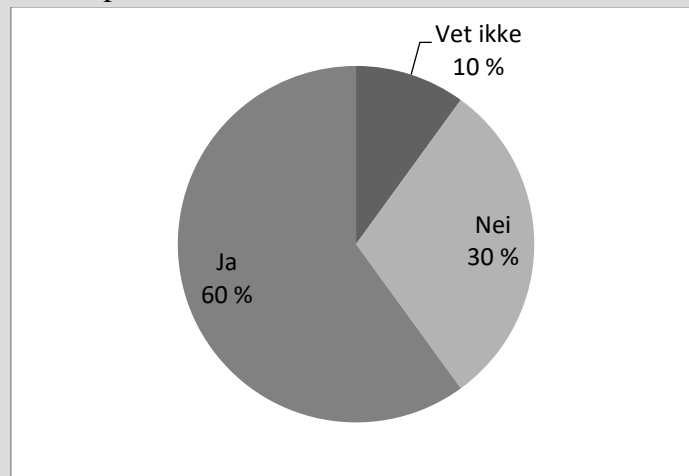
Vi vil lage et sektordiagram som illustrerer svarene. Vi regner da ut hvor mange grader hver av de tre sektorene må fylle. Hele sirkelen utgjør 360° .

Ja-sektoren må fylle $360^\circ \cdot 0,6 = 216^\circ$.

Nei-sektoren må fylle $360^\circ \cdot 0,3 = 108^\circ$.

Vet ikke-sektoren må fylle $360^\circ \cdot 0,1 = 36^\circ$.

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:



Oppgave 3

To klasser på 60 elever skal ha aktivitetsdag. 12 elever ønsker langrenn, 22 slalåm, 18 aking og 8 fottur. Lag et sektordiagram på papir (og uten kalkulator) som illustrerer denne svarfordelingen.

Tips: Legg merke til at $360/60 = 6$. En elev svarer altså til 6 grader.

Utfyllingsoppgave – Diagrammer

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

Oppgave 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 (kurvedigram)		
-----------------------------------	--	--

Oppgave 2

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

Tabellen viser karakterfordelingen på en matematikkeksamen på en skole.

- Hvorfor vil sektordiagram kunne være hensiktsmessig å bruke her?
- Kunne man brukt et stolpediagram istedenfor? Begrunn svaret
- 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		

Oppgave 3

Antall fraværsdager	Antall elever
0	8
1	9
2	4
3	3
4	4
7	1
9	1
10	1
11	1
12	1
15	1

Tabellen viser fraværsdager for en klasse på Vg2.

- a) Hvorfor er det ikke hensiktsmessig å bruke et sektordiagram her?

- b) Hvorfor er stolpediagram det mest hensiktsmessige?

- c) Hva må du være oppmerksom på her når du skal tegne stolpediagrammet?

d) Tegn diagrammet for hånd.



e) Tegn diagrammet i Excel.

(Lim inn Excelbildet her)

Oppgave 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.

- a) Hvorfor er et linjediagram mest hensiktsmessig å bruke her?

- b) Tegn et linjediagram for hånd

c) Tegn et linjediagram i Excel.

(Lim inn Excelbildet her)

Oppgave 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?

4. Sentralmål

Et *sentralmål* er et tall som viser hovedtendensen i et datamateriale. De vanligste sentralmålene er *gjennomsnitt*, *median* og *typetall*.

4.1 Gjennomsnitt

Vi finner *gjennomsnittet* av et datamateriale ved å legge sammen alle observasjonene og dividere med antall observasjoner.

Eksempel 2

Høydene til 10 gutter (i cm) er 187, 182, 175, 184, 173, 180, 182, 177, 171, 186.

Vi regner ut summen av høydene: $187 + 182 + 175 + 184 + 173 + 180 + 182 + 177 + 171 + 186 = 1797$.

Gjennomsnittet er $1797 \text{ cm} / 10 = 179,7 \text{ cm}$.

Oppgave 4

Noen elever ble spurt om hvor mange PCer, nettbrett og mobiler det til sammen var i familien. De ga følgende svar:

3 5 6 4 4 7 4 4 7 5

Finn gjennomsnittlig antall "datadingser" i disse familiene.

Oppgave 5



- Kan en gjennomsnittshøyde være større enn *alle* høydene som er brukt for å regne ut gjennomsnittet?
- Kan *alle* høydene være lik gjennomsnittshøyden?

Eksempel 3

Hvis datamaterialet er satt i en frekvenstabell, finner vi gjennomsnittet slik:

Karakter	Frekvens	Karakter * Frekvens
1	9	9
2	17	34
3	10	30
4	8	32
5	5	25
6	1	6
	Sum = 50	Sum = 136

Dette viser at summen av alle 50 karakterene er 136.

Gjennomsnittskarakteren er $136 / 50 = 2,72$

Oppgave 6

Taiba spiller håndball. Hun har laget en tabell som viser antall mål hun har scoret i kampene. Tabellen viser for eksempel at i 5 av kampene har hun scoret 2 mål i hver kamp. Frekvensene er her antall kamper hvor hun har scoret 0 mål, 1 mål, osv.

Antall mål	Antall kamper
0	2
1	3
2	5
3	7
4	8
5	6
6	2

- Hvor mange kamper har hun spilt?
- Lag en ekstra kolonne og finn ut hvor mange mål hun har scoret tilsammen på disse kampene.
- Hvor mange mål har hun scoret i gjennomsnitt per kamp?

4.2 Median

Medianen i et datamateriale er den midterste verdien når materialet er sortert i stigende rekkefølge.

Eksempel 4

Hva er medianen for karakterene 4, 2, 2, 5 og 1?

Vi ordner dem i stigende rekkefølge: 1, 2, **2**, 4, 5. Da ser vi at medianen er 2.

Eksempel 5

Hva er medianen for karakterene 4, 2, 2, 5, 1 og 3?

Vi ordner dem i stigende rekkefølge: 1, 2, **2, 3**, 4, 5. Da ser vi at det er ikke noe tall nøyaktig i midten. I slike tilfelle er medianen definert som gjennomsnittet av de to verdiene nærmest

midten. Derfor er medianen $\frac{2+3}{2} = \frac{5}{2} = 2,5$.

Oppgave 7

a) Finn medianen i dette datamaterialet, som viser antall bøker noen tilfeldige elever hadde lest i 2012: 6, 2, 1, 0, 4, 1, 7, 0, 3.

b) Finn medianen i dette datamaterialet, som viser antall fraværsdager for noen elever: 2, 4, 0, 0, 1, 2, 3, 1, 4, 1, 2, 0.

Hvis datamaterialet er ordnet i en frekvenstabell, er medianen den minste dataverdien hvor den relative kumulative frekvensen er større enn eller lik 50 %.

Karakter	Frekvens	Kumulativ frekvens	Relativ kum. fr.
1	9	9	18 %
2	17	26	52 %
3	10	36	72 %
4	8	44	88 %
5	5	49	98 %
6	1	50	100 %

Her er medianen lik 2.

4.3 Typetall

Typetallet i et datamateriale er den dataverdien som forekommer flest ganger (“er typisk for”) i datamaterialet.

Eksempel 6

Typetallet i karaktereksemplet ovenfor er 2, fordi det er flest elever som har fått 2 (17 elever)

Oppgave 8

Finn typetallet i datamaterialet i oppgave 1.

5. Spredningsmål

Et *spredningsmål* er et tall som sier noe om hvor stor spredning det er på tallene i et datamateriale.

5.1 Variasjonsbredde

Det enkleste spredningsmålet er *variasjonsbredde*.

Variasjonsbredden i et datamateriale er forskjellen mellom største og minste verdi i materialet.

Eksempel 7

I oppgave 7a hadde vi tallene 6, 2, 1, 0, 4, 1, 7, 0, 3, som var antall leste bøker i 2012. Største verdien er 7 bøker og minste er 0 bøker. Variasjonsbredden er 7 bøker – 0 bøker = 7 bøker.

Oppgave 9

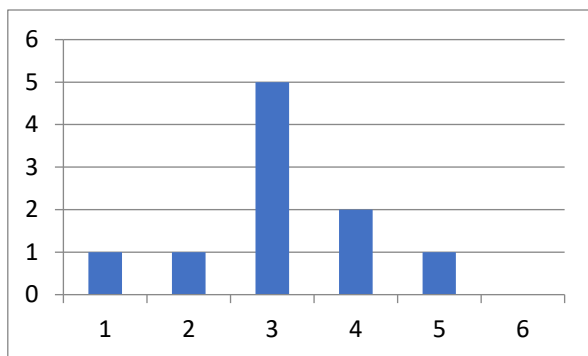
Åtte jenter løper 100 m. Tidene i sekunder er 15,6 17,8 14,4 18,2 16,3 14,9 15,8 17,1. Finn variasjonsbredden.

5.2 Standardavvik

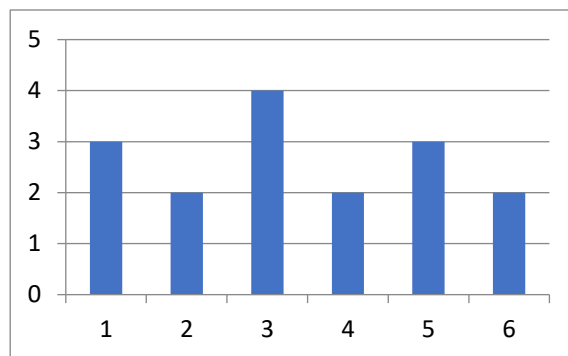
Ofte gir ikke variasjonsbredden et godt bilde av spredningen. For eksempel kan vi tenke oss en 2P-gruppe hvor nesten alle elevene fikk karakter 3, mens en elev fikk 1 og en fikk 6. I en annen gruppe fikk to elever 1, fire fikk 2, fem fikk 3, fire fikk 4, tre fikk 5 og to fikk 6. Variasjonsbredden er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

6. Beregning av sentralmål og spredningsmål i Excel

De sentral- og spredningsmål i 2P som kan være aktuelle å beregne i Excel er gjennomsnitt, median, typetall og standardavvik. Men husk at gjennomsnitt og median med enkle tall må du også kunne beregne uten hjelpemidler.

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*.

	A	B	C	D	E
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	Gjennomsnit	=GJENNOMSNI			

- 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.

7. 🧠 **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.

7.1 Gjennomsnitt i klassesdelt materiale

Fordi vi ikke kjenner alle verdiene i et klassesdelt 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.



7.2 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.

Eksamensoppgaver statistikk

E1

(Eksamen vår 2010, Del 1)

I fjor kom alle elevene i en matematikkgruppe opp til eksamen i 2P. De oppnådde disse resultatene:

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

- 1) Lag en tabell som viser frekvens og kumulativ frekvens.
- 2) Finn medianen og gjennomsnittet for datamaterialet.

E2

(Eksamen høst 2016, Del 2, 6p)

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?

E3

(Eksamen høst 2016, Del 1, 3p)

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?

E4

(Eksamen vår 2011, Del 1)

Nedenfor ser du hvor mange mål som ble scoret i fotballkampene mellom Rosenborg og Brann i Eliteserien i årene fra 2005 til 2009:

5 5 0 4 3 5 2 0 2 2

- 1) Finn gjennomsnittet og medianen for dette datamaterialet.
- 2) Sett opp resultatene i en tabell. Tabellen skal vise frekvens og kumulativ frekvens.
- 3) Hva er den kumulative frekvensen for to mål, og hva betyr dette?

E5

(Eksamen vår 2016, 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 første seks dagene i mars. Se tabellen ovenfor.

Bestem variasjonsbredden, gjennomsnittet og medianen for temperaturmålingene.

E6




(Eksamen høst 2016, Del 1, 4p)

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.

E7

(Eksamen høst 2012, Del 1)

Alle som går på tur til Pollfjell, skriver navnet sitt i boka som ligger i postkassen på toppen av fjellet. Nedenfor ser du hvor mange som har skrevet seg inn i boka hver uke de 12 siste ukene.

6 12 20 4 10 15 5 12 8 12 18 10

Bestem gjennomsnittet, medianen, typetallet og variasjonsbredden for dette datamaterialet.

E8

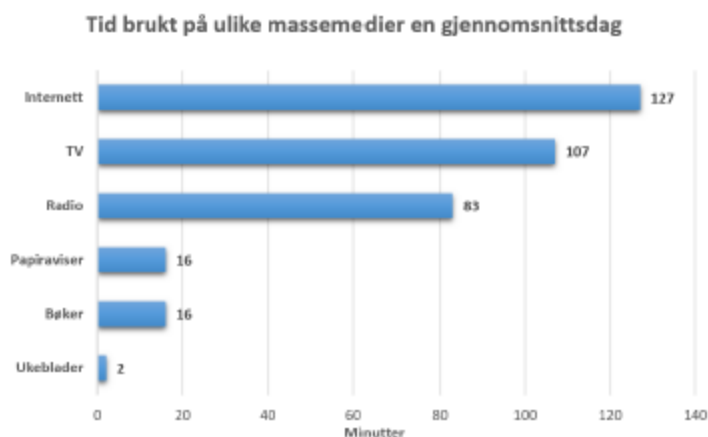
(Eksamen vår 2012, Del 1)

Antall datamaskiner	Antall elever
1	3
2	4
3	3
4	6
5	2
6	2

20 elever blir spurt om hvor mange datamaskiner de har hjemme. Se tabellen ovenfor. Finn variasjonsbredden, typetallet, medianen og gjennomsnittet.

E9

(Eksamen høst 2016, Del 2, 2p)



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.

E10

(Eksamen høst 2012, Del 2)

År	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Antall kamper per år	1	2	4	10	9	6	4	9	11	11	4
Antall mål per år	0	0	2	3	3	2	0	5	1	4	0

Tabellen ovenfor viser hvor mange landskamper Jan Åge Fjørtoft spilte, og hvor mange mål han skåret per år i perioden 1986–1996.

- a) Hvor mange mål skåret Fjørtoft i gjennomsnitt per kamp i denne perioden?
I hvilket år skåret han flest mål per kamp?
- b) Tegn av tabellen nedenfor i besvarelsen din, og fyll inn tallene som mangler.
Hva er den kumulative frekvensen for to mål per år, og hva forteller dette svaret?

Antall mål per år	Frekvens	Kumulativ frekvens
0		
1		
2		
3		
4		
5		

E11

(Eksamen vår 2013, Del 2)

Stortinget ved starten av perioden 2009–2013	
Parti	Antall representanter
Arbeiderpartiet	64
Fremskrittspartiet	41
Høyre	30
Sosialistisk Venstreparti	11
Senterpartiet	11
Kristelig Folkeparti	10
Venstre	2

a) Lag et sektordiagram som illustrerer opplysningene gitt i tabellen ovenfor.

Stortinget ved starten av perioden 2009–2013		
Parti	Antall kvinner	Antall menn
Arbeiderpartiet	32	32
Fremskrittspartiet	10	31
Høyre	9	21
Sosialistisk Venstreparti	3	8
Senterpartiet	7	4
Kristelig Folkeparti	4	6
Venstre	2	0

b) Lag et passende diagram som illustrerer opplysningene gitt i tabellen ovenfor.


E12

(Eksamen høst 2008, Del 2)

Lengdehopp er en gren av friidrett som går ut på å hoppe så langt man kan i et hopp. I konkurranser har man som regel tre hopp, der det beste hoppet teller.

Anna og Petra konkurrerer om å kvalifisere seg til lengdehoppkonkurransen i et friidrettsstevne. De får ti hopp hver, og den beste av dem er kvalifisert til konkurransen. Her er resultatene (oppgitt i meter) fra kvalifiseringen:

Hopp	1	2	3	4	5	6	7	8	9	10
Anna	5,10	5,45	5,92	4,10	5,23	5,32	5,89	4,91	4,37	5,42
Petra	5,44	5,80	5,67	5,74	5,72	5,04	5,73	5,53	5,59	5,83

- Finne gjennomsnitt og median for hver av de to jentenes resultater.
- Finne variasjonsbredde og standardavvik for hver av de to jentenes resultater.
-  Foreta en vurdering av jentenes resultater og det du fant i a) og b), og argumenter for hvem du synes skal bli kvalifisert.

E13

(Eksamen høst 2010, Del 2)

I klasse 1A er det 12 jenter og 12 gutter. Nedenfor ser du hvor mange timer de bruker på lekser hver uke.

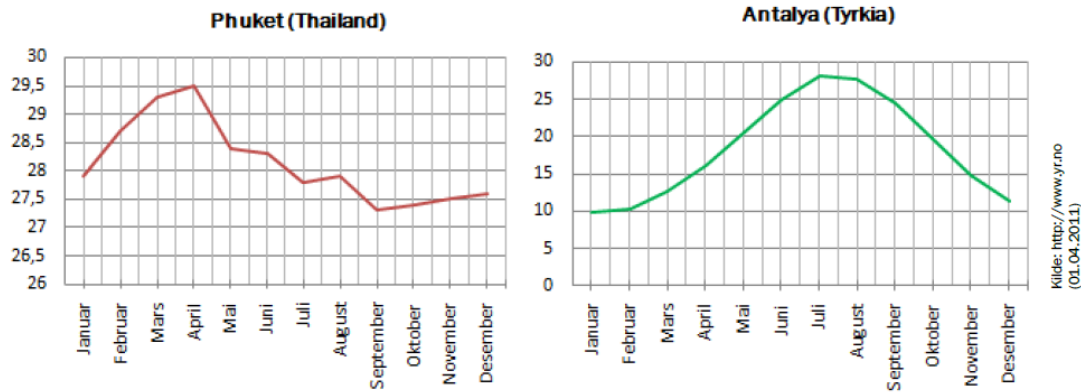
Jentene: 7, 5, 5, 7, 7, 6, 8, 8, 5, 4, 6, 10

Guttene: 2, 5, 6, 7, 9, 6, 4, 9, 12, 2, 13, 3

Bruk ulike sentral- og spredningsmål og gjør rede for hva dette datamaterialet viser om jentenes og guttenes arbeidsvaner i denne klassen.

E14

(Eksamen vår 2012, Del 2)




Ovenfor ser du linjediagram som viser gjennomsnittstemperaturen per måned ved to kjente feriesteder.

- Bruk diagrammene og lag en tabell som viser gjennomsnittstemperaturen per måned for hvert av de to stedene Phuket og Antalya.
- 1) Finn gjennomsnittstemperaturen per år for hvert av de to stedene.
 - 2) Finn standardavviket for temperaturene i a) for hvert av de to stedene.

Jon påstår at det er godt samsvar mellom diagrammene og resultatene fra b).

Asbjørn er enig, men mener at diagrammene lett kan tolkes feil.

-  Forklar hvorfor diagrammene lett kan tolkes feil. Hvordan kunne diagrammene vært laget for å unngå dette?

E15

(Eksamen vår 2010, Del 2)

År	Antall laks	Totalvekten av laksen (kg)	Gjennomsnittsvekten for laksen (kg)
2000	3447	14550	4,22
2001	5694	24218	4,25
2002	3142	1)	5,51
2003	4401	22425	5,10
2004	2928	17979	2)
2005	5783	28144	4,87
2006	6507	26695	4,10
2007	3555	21074	5,93
2008	4782	28903	6,04
2009	3916	24361	6,22

Tabellen ovenfor viser hvor mange laks, totalvekten av laksen og gjennomsnittsvekten for laksen som er fanget i elva Gaula i Sør-Trøndelag de siste ti årene.

- Hvilke tall skal stå i tabellfeltene som er merket 1) og 2)?
- Lag et passende diagram som viser hvor mange laks som er fanget i Gaula per år de siste ti årene.
- Finn gjennomsnittet av og standardavviket for totalvekten av laksen fanget i Gaula per år de siste ti årene.

E16

(Eksamen høst 2011, Del 2)

a) Finn median, gjennomsnitt og standardavvik for tallmengden:

2 5 21 15 17 5 9 19 10 14 7 3 2 11 13

Vi dobler alle tallene i tallmengden og får:


4 10 42 30 34 10 18 38 20 28 14 6 4 22 26

b) Finn median, gjennomsnitt og standardavvik for denne tallmengden.
Sammenlikn med resultatene fra a) og kommenter.

Berit får en idé og setter opp tabellen nedenfor.





















Tallmengde 1 15 tall	2	5	21	15	17	5	9	19	10	14	7	3	2	11	13
Tallmengde 2 De 15 tallene doblet	4	10	42	30	34	10	18	38	20	28	14	6	4	22	26
Tallmengde 3 De 15 tallene tredoblet	6	15	63	45	51	15	27	57	30	42	21	9	6	33	39
Tallmengde 4 De 15 tallene firedoblet	8	20	84	60	68	20	36	76	40	56	28	12	8	44	52

Hun beregner median, gjennomsnitt og standardavvik for hver av tallmengdene og påstår at hun har funnet regler som sier noe om hvordan medianen, gjennomsnittet og standardavviket endrer seg når tallene i en tallmengde dobles, tredobles, firedobles osv.

c)  Formuler disse reglene, og gi en begrunnelse for at de er riktige.

E17


(Eksamen høst 2012, Del 2)

Nr	Resultat	Utøver	Land	Verdensdel
1	8,47	Christian Reif	 Tyskland	Europa
2	8,46	Dwight Phillips	 USA	Nord-Amerika
3	8,40	Fabrice Lapierre	 Australia	Oceania
4	8,35	Alain Bailey	 Jamaica	Sør-Amerika
5	8,33	Chris Noffke	 Australia	Oceania
6	8,30	Irving Saladino	 Panama	Sør-Amerika
7	8,27	Ndiss Kaba Badji	 Senegal	Afrika
8	8,27	Eusebio Cáceres	 Spania	Europa
9	8,25	Pavel Shalin	 Russland	Europa
10	8,24	Salim Sdiri	 Frankrike	Europa
11	8,24	Kafétien Gomis	 Frankrike	Europa
12	8,23	Godfrey Khotso Mokoena	 Sør-Afrika	Afrika
13	8,23	Christopher Tomlinson	 England	Europa
14	8,22	Tommi Evilä	 Finland	Europa
15	8,22	Tyrone Smith	 Bermuda	Sør-Amerika
16	8,22	Greg Rutherford	 England	Europa
17	8,21	Morten Jensen	 Danmark	Europa
18	8,20	Wilfredo Martínez	 Cuba	Sør-Amerika
19	8,20	Trevell Quinley	 USA	Nord-Amerika
20	8,19	Christian Taylor	 USA	Nord-Amerika

Ovenfor ser du verdensstatistikken fra 2010 for øvelsen lengdehopp for menn.

- Lag et sektordiagram som viser hvordan de 20 utøverne fordeler seg mellom de ulike verdensdelene.
- Finn gjennomsnittslengden og standardavviket for resultatene til de 20 utøverne.

Sondre har funnet resultatene for utøverne som står som nummer 21–40 på verdensstatistikken. Standardavviket for resultatene til disse 20 utøverne er tilnærmet lik 0,0258.

-  Hva forteller dette om resultatene til utøverne som står som nummer 21–40, i forhold til resultatene til de 20 beste utøverne?

E18

Eksamen vår 2012, Del 2)

En dag gjorde klasse 1A et forsøk i naturfagtimen. Seks elever slapp hver sin stålkule fra 1 m høyde og målte tiden det tok før kule traff bakken. Resultatene ser du i tabellen nedenfor.



Elev	1	2	3	4	5	6
Tid (sekunder)	0,46	0,45	0,47	0,44	0,52	0,46

a) Bestem gjennomsnittet og standardavviket for måleresultatene.

Klassen la merke til at elev nummer 5 målte en større falltid enn de andre. Mange mente at dette resultatet måtte skyldes målefeil, og at det derfor burde forkastes.

Da ga fysikklærer Strøm dem denne regelen:

«Når vi har seks målinger, kan vi forkaste et måleresultat dersom det ligger mer enn 1,4 standardavvik fra gjennomsnittet.»

- b)  Finn ut om måleresultatet til elev nummer 5 kan forkastes dersom vi bruker regelen ovenfor.
- c)  Bestem gjennomsnittet og standardavviket for de fem andre måleresultatene. Hvordan har gjennomsnitt og standardavvik endret seg? Virker dette rimelig? Forklar.

E19

(Eksamen vår 2013, Del 1)

En kveld kjørte en taxisjåfør 10 turer.

Nedenfor ser du hvor mange passasjerer han hadde med på hver av turene.

1 5 3 3 5 2 1 4 1 2

- a) Bestem medianen, gjennomsnittet og typetallet for dette datamaterialet.
- b) Sett opp en tabell som viser frekvens og kumulativ frekvens for antall passasjerer på turene.

E20









(Eksamen vår 2013, Del 2)

Tabellene nedenfor viser resultatene for de åtte beste utøverne på 1500 m skøyter for menn under OL i 1968 og under OL i 2010.

OL 1968

Plass	Utøver	Land	Tid (sekunder)
1	Kees Verkerk	 Nederland	123,4
2	Ivar Eriksen	 Norge	125,0
2	Ard Schenk	 Nederland	125,0
4	Magne Thomassen	 Norge	125,1
5	Johnny Höglin	 Sverige	125,2
6	Bjørn Tveter	 Norge	125,2
7	Svein-Erik Stiansen	 Norge	125,5
8	Eduard Matusevitsj	 Sovjetunionen	126,1

OL 2010

Plass	Utøver	Land	Tid (sekunder)
1	Mark Tuitert	 Nederland	105,57
2	Shani Davis	 USA	106,10
3	Håvard Bøkko	 Norge	106,13
4	Ivan Skobrev	 Russland	106,42
5	Mo Tae-bum	 Korea	106,47
6	Chad Hedrick	 USA	106,69
7	Simon Kuipers	 Nederland	106,76
8	Mikael Flygind Larsen	 Norge	106,77

- Hvor mange prosent sank vinnertiden med fra 1968 til 2010?
- Bestem gjennomsnittstiden for de åtte beste i 1968 og for de åtte beste i 2010.
- Bestem standardavviket for de to tallmaterialene.
Hvorfor er standardavviket større i 1968 enn i 2010?

E21

(Eksamen vår 2011, Del 1)

I en 2P-gruppe er det 10 elever. Læreren har undersøkt hvor mye tid elevene bruker på matematikkleksene i løpet av en uke.

Resultatene er gitt i tabellen nedenfor.

Antall minutter	Antall elever
$[0,30)$	1
$[30,60)$	3
$[60,120)$	5
$[120,240)$	1

Finn gjennomsnittet for dette grupperte datamaterialet.

E22

(Eksamen høst 2012, Del 1)

Tabellen nedenfor viser hvor mye penger hver av de 10 elevene i en 2P-gruppe bruker i kantinen i løpet av en uke.


Kroner	Antall elever
$[0,50)$	1
$[50,100)$	5
$[100,150)$	1
$[150,200)$	3


Gjør beregninger og avgjør om gjennomsnittet er større enn medianen for dette datamaterialet.



E23

(Eksamen vår 2011, Del 2)

Politiet har gjennomført fartskontroller på to veistrekninger. Den ene veistrekningen har fartsgrense 50 km/h og den andre 80 km/h. Nedenfor ser du resultatene fra hver av de to kontrollene.

Fartsgrense 50 km/h 	
Fart	Antall biler
[45,50)	25
[50,55)	26
[55,60)	23
[60,65)	3
[65,70)	2
[70,75)	1

Fartsgrense 80 km/h 	
Fart	Antall biler
[70,75)	7
[75,80)	43
[80,85)	17
[85,90)	8
[90,95)	0
[95,125)	5

- Presenter dataene fra tabellene ovenfor i hvert sitt stolpediagram.
-  Hvor mange prosent av bilførerne kjører 10 % eller mer over fartsgrensen i hver av de to kontrollene?
- Finn gjennomsnittsfarten til bilene i hver av de to kontrollene.
- Hvor mange prosent over fartsgrensen er gjennomsnittsfarten til bilene i hver av de to kontrollene?
-  Bruk svarene i a), b), c) og d) til å vurdere om bilførerne kjører mest lovlydig på veistrekningen med fartsgrense 50 km/h eller på veistrekningen med fartsgrense 80 km/h.

E24

(Eksamen vår 2012, Del 1)

Nedenfor ser du hvor mange tekstmeldinger hver av de 20 elevene i en 2P-gruppe sendte i løpet av en uke:

4 88 69 21 66 8 16 57 86 21 37 22 78 27 28 44 42 71 82 95

- 1) Grupper datamaterialet i klasser med bredde 20. La den første klassen starte med 0.

I hvilken klasse ligger medianen?

- 2) Finn gjennomsnittet i det klassesdelte materialet.

E25

(Eksamen vår 2013, Del 1)

Tabellen nedenfor viser inntektene til personene i et borettslag.

Inntekt (i 1000 kroner)	Antall personer
[300, 400)	20
[400, 500)	20
[500, 700)	10

Bestem gjennomsnittsinntekten til personene i borettslaget.

E26

(Eksamen høst 2011, Del 1)

Politiet har gjennomført en farts kontroll i 30 km-sonen utenfor skolen.

Resultatene er gitt i tabellen nedenfor.

Fart (km/h)	Antall biler
[20,30)	20
[30,40)	20
[40,50)	10



Finn gjennomsnittsfarten.

E27

(Eksamen vår 2008, Del 2)

Et nytt leilighetskompleks, UTSIKTEN, er fylt opp med nye beboere. En oversikt viser følgende aldersfordeling:

Alder	0–19 år	20–39 år	40–59 år	60–79 år
Frekvens	17	29	51	23

- Hvor mange personer bor i leilighetskomplekset?
- Tegn et søylediagram som viser aldersfordelingen.
- Forklar hvorfor medianen må ligge i intervallet 40–59 år.
- Regn ut gjennomsnittsalderen ut fra det klassesdelte materialet.

E28

(Eksamen 2P-Y vår 2015, Del 1)

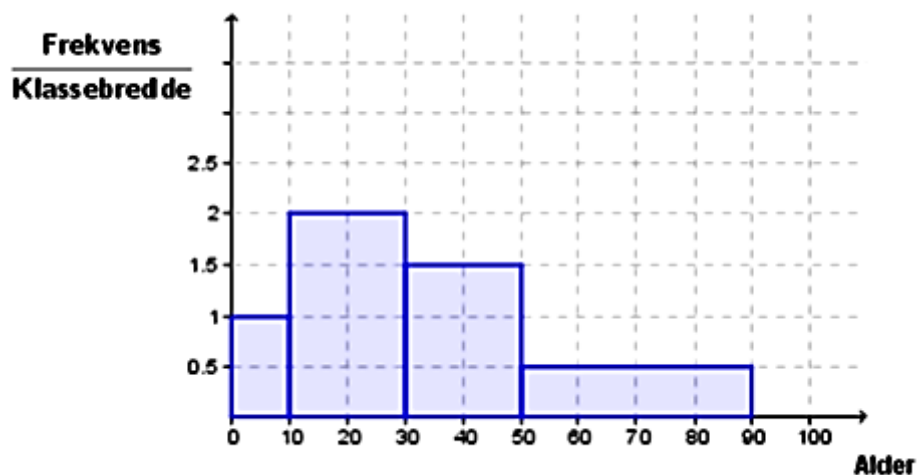
Alder	Frekvens
[20,30)	10
[30,40)	20
[40,50)	30
[50,70)	40

Tabellen over 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

E29

(Eksamen 2P-Y høst 2014, Del 1)



Histogrammet ovenfor viser aldersfordelingen blant de besøkende på en kinoforestilling.

- Forklar at det var 30 besøkende mellom 30 og 50 år.
- Hvor mange prosent av de besøkende var mellom 0 og 10 år?
- Bestem gjennomsnittsalderen blant de besøkende.