

Presentasjon og analyse



Målet for opplæringen er at eleven skal kunne:

- Bruke og vurdere valg av passende sentralmål og spredningsmål for statistisk datamateriale
- Analysere og presentere funn i datasett fra lokalsamfunn og media.

Informasjon

På ungdomskolen har du sannsynligvis vært med på å samle inn informasjon, kanskje gjennom en spørreundersøkelse eller ved et forsøk. Du har kanskje spurt dine medelever om hvor mange søsken de har, eller hvor mange timer de bruker på skjerm. Kanskje har du registrert hvor mange biler som passerer skolen i et bestemt tidsrom.



Hvert enkelt svar fra spørreundersøkelsen eller hvert enkelt resultat fra forsøket kalles en **observasjon**, og antallet som svarer eller antall resultater fra undersøkelsen kalles **sum observasjoner**.

Informasjonen hver enkelt observasjon gir kalles **data**, og alle dataene samlet kalles **datamateriale**. I dette kapittelet skal vi først se på hvordan vi kan **presentere** informasjonen fra et datamateriale. Deretter skal vi se hvilke **analyser** vi kan gjøre av et datamateriale.

Tenk deg at læreren spør klassen om hvor mange transportmidler hver enkelt elev brukte for å komme til skolen i dag. Da vil hvert enkelt svar være en **observasjon** og antall elever som var med på undersøkelsen vil være **sum observasjoner**.

Hva hver enkelt elev svarer vil være **data**, mens alle svarene samlet vil bli undersøkelsens **datamateriale**. Det er dette **datamaterialet** som vi enten kan **presentere** eller **analysere**

Eksempel:

En taxisjåfør registrerte antall turer hver dag en uke i desember. Her blir **antall observasjoner** 7.

Sjåføren registrerte følgende observasjoner fra mandag til søndag:

14 - 17 - 12 - 21 - 29 - 37 - 14



Presentasjon – tabell og diagram

Når vi har skaffet oss et datamateriale kan vi lage en presentasjon av informasjonen i datamaterialet. Det er ryddig å først systematisere dataene i en **tabell**, og det kan være lurt å gjøre dette i ExCel:

	A	B
1	Dag	Antall turer
2	Mandag	14
3	Tirsdag	17
4	Onsdag	12
5	Torsdag	21
6	Fredag	29
7	Lørdag	37
8	Søndag	14

Tabell er ikke alltid den beste måten å presentere informasjon på, spesielt ikke dersom det er mye informasjon som skal presenteres. I slike tilfeller kan vi bruke mer visuelle hjelpemidler, for eksempel en graf eller et **diagram**.

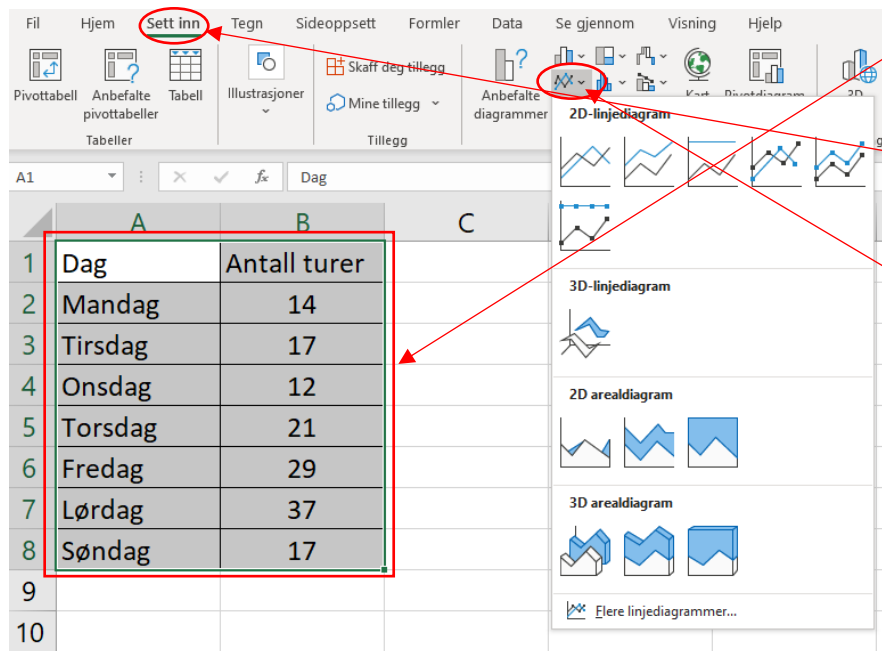
Det finnes tre typer diagrammer du bør kjenne til og når de kan brukes:

Type	Linjediagram	Søylediagram	Sektordiagram
Brukes når vi ønsker å vise	Utvikling over tid	Forskjellen mellom dataene. Her trenger ikke alle data være med.	Andel (gjerne prosent) av datamaterialet
Eksempel			

På de neste sidene viser vi hvordan du kan lage slike diagrammer i ExCel.

Linjediagram:

Dersom vi ønsker å vise utvikling i antall turer gjennom uka, kan vi bruke et



1. Marker tallene i tabellen.

2. Trykk på «Sett inn».

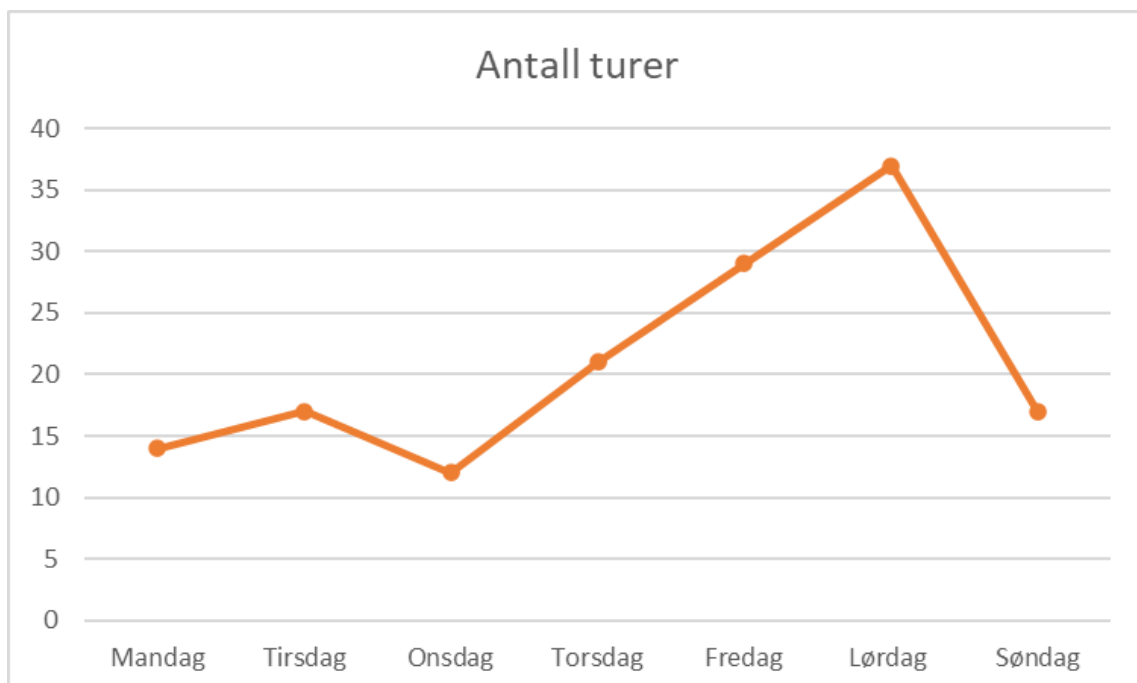
3. Trykk på «Sett inn linje- eller arealdiagram».

4. Velg det diagrammet du ønsker.

Dag	Antall turer
Mandag	14
Tirsdag	17
Onsdag	12
Torsdag	21
Fredag	29
Lørdag	37
Søndag	17

Linjediagram.

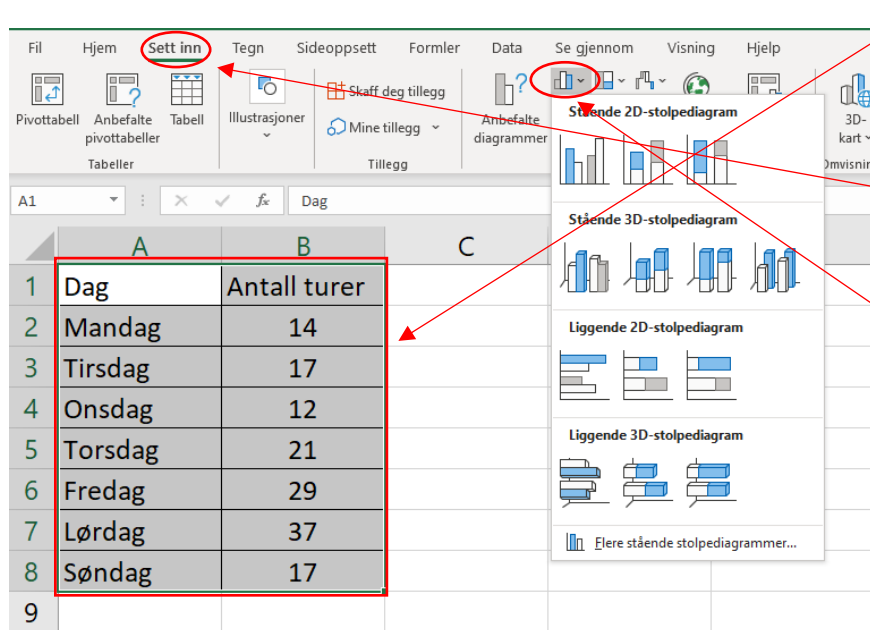
Utviklingen i antall turer gjennom uka blir slik:



Klarer du å lage dette diagrammet i Excel?

Søylediagram:

Dersom vi ønsker å vise forskjellen i antall turer gjennom uka, kan vi bruke et søylediagram.



1. Marker tallene i tabellen.

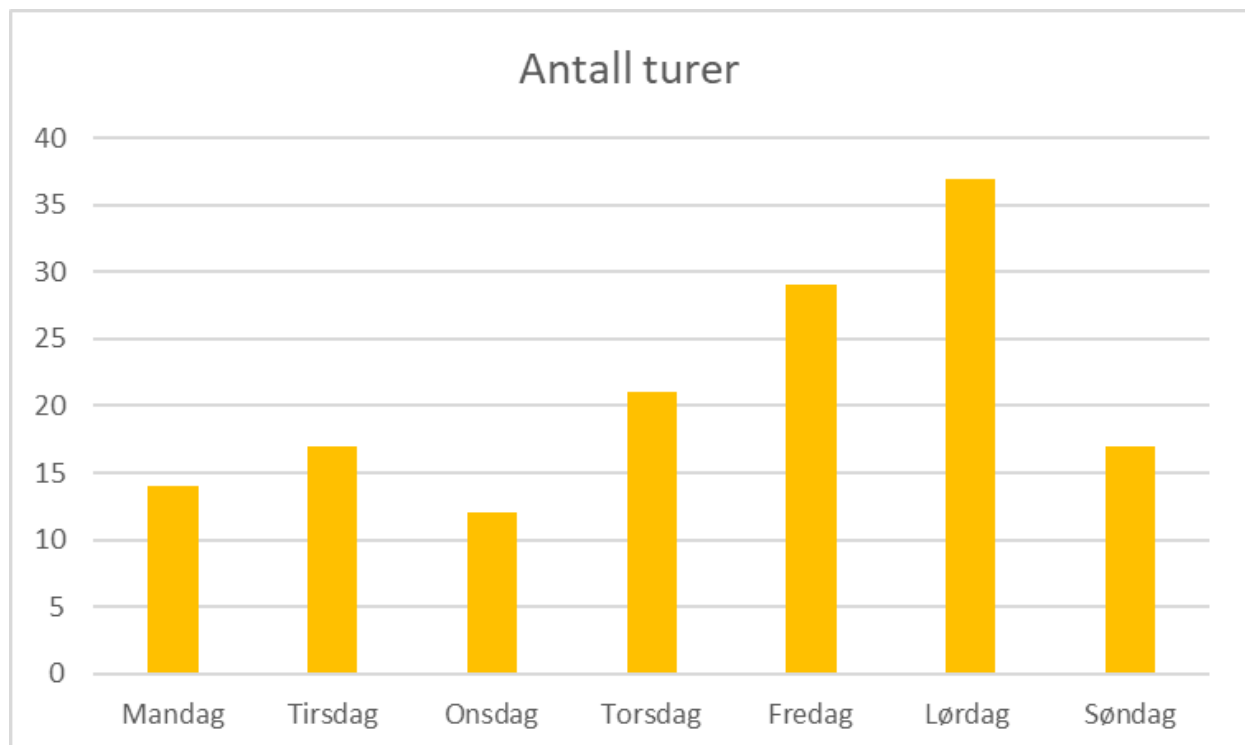
2. Trykk på «Sett inn».

3. Trykk på «Sett inn stående eller liggende stolpediagram».

4. Velg det diagrammet du ønsker.

Dag	Antall turer
Mandag	14
Tirsdag	17
Onsdag	12
Torsdag	21
Fredag	29
Lørdag	37
Søndag	17

Antall turer fra mandag til søndag blir slik:



Klarer du å lage dette diagrammet i ExCel?

Sektordiagram:

Dersom vi ønsker å vise den prosentvise fordelingen av antall turer, kan vi bruke et **sektordiagram**.

The screenshot shows the Excel interface with the 'Sett inn' menu open. The '2D-sektordiagram' options are visible, including '2D-sektordiagram', '3D-sektordiagram', and 'Hjuldiagram'. The data table is highlighted with a red box.

	A	B	C	E
1	Dag	Antall turer		
2	Mandag	14		
3	Tirsdag	17		
4	Onsdag	12		
5	Torsdag	21		
6	Fredag	29		
7	Lørdag	37		
8	Søndag	17		
9				
10				

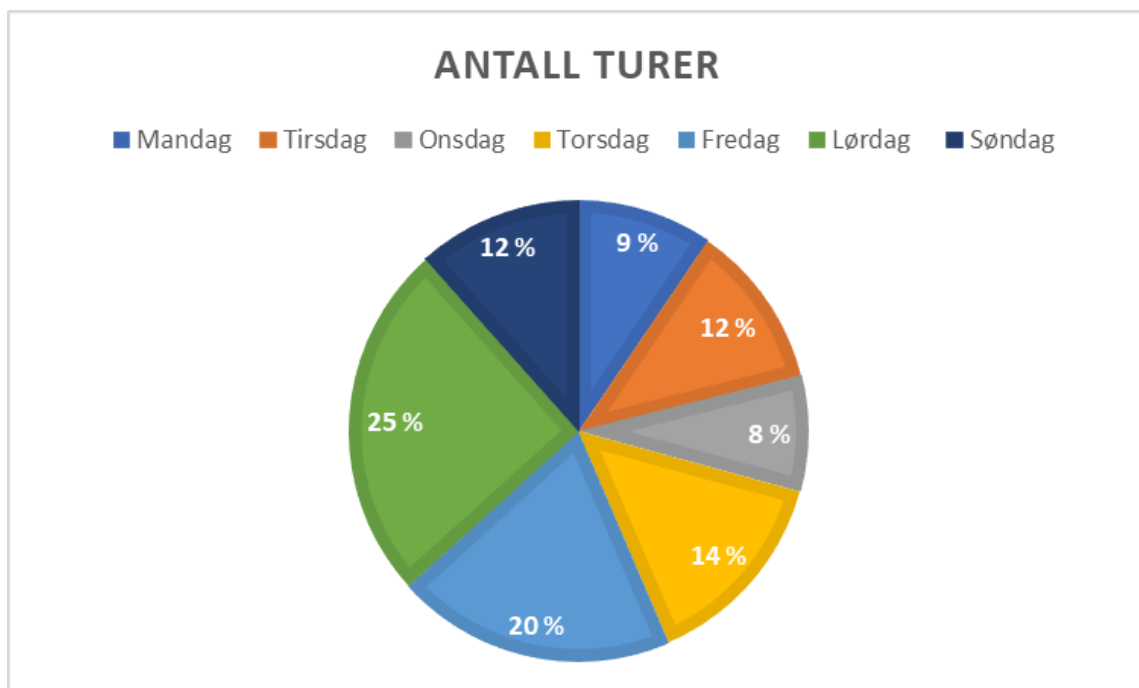
1. Marker tallene i tabellen.

2. Trykk på «Sett inn».

3. Trykk på «Sett inn sektor- eller hjuldiagram».

4. Velg det diagrammet du ønsker. Husk å inkludere prosent.

Den prosentvise fordelingen av antall turer blir slik:



Klarer du å lage dette diagrammet i Excel?

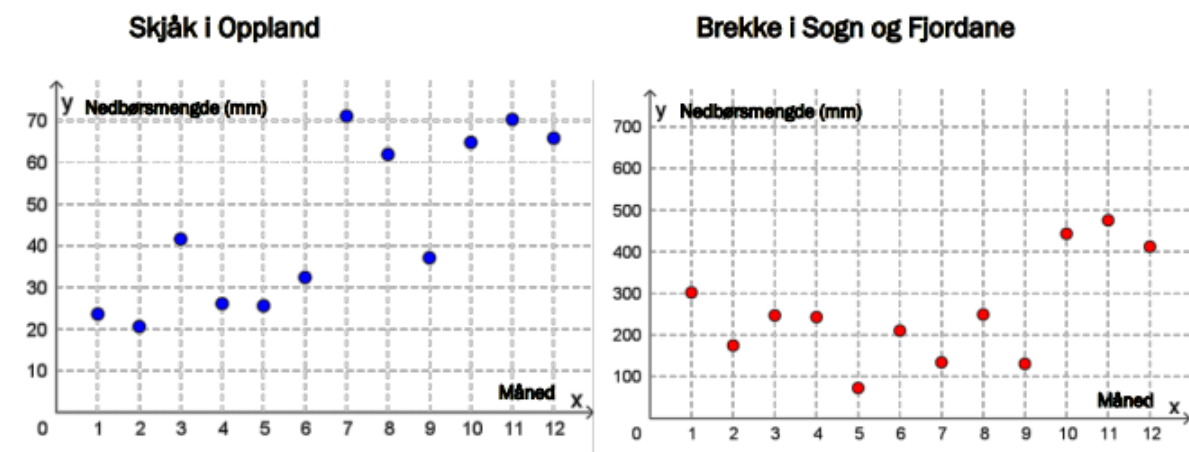
Oppgave 1

By	Antall innbyggere
São Paulo	11 968 000
Mexico City	8 919 000
Lima	8 894 000
New York	8 550 000
Bogotá	7 862 000
Rio de Janeiro	6 477 000
Santiago	5 507 000
Los Angeles	3 972 000
Caracas	3 290 000
Buenos Aires	3 054 000
Salvador	2 921 000
Brasília	2 914 000
Toronto	2 826 000
Chicago	2 721 000

Tabellen ovenfor viser hvor mange innbyggere det er i hver av de 14 største byene i Sør- og Nord-Amerika.

Fremstill innbyggertallene i et passende diagram. Begrunn valg av diagram.

Oppgave 2



Ovenfor ser du to diagram. Diagrammene viser nedbørsmengde hver måned i 2017 i Skjåk og Brekke.

Fremstill den månedlige utviklingen i nedbørsmengde for hvert av stedene i et felles diagram. Begrunn valg av diagram.

Oppgave 3

Tabellen nedenfor viser en oversikt over høydene til elevene ved en skole.

Høyde i cm	Frekvens
$[150, 160)$	10
$[160, 170)$	30
$[170, 180)$	50
$[180, 200)$	10

Fremstill resultatet i et diagram. Diagrammet skal vise den prosentvise fordelingen mellom de ulike høydene.

Begrunn valg av diagram.

Presentasjonsoppgave

Velg deg et parti fra din lokale valgkrets. Lag en presentasjon av partiets resultater ved de siste kommune- og stortingsvalgene.

Velg ulike datamateriale, slik at du får vist at du behersker både linje-, søyle- og sektordiagram.

Analyse

Når vi har skaffet oss et datamateriale kan vi gjøre mer enn å bare presentere resultatene. Vi kan også gjøre noen analyser av informasjonen. Nedenfor har vi listet opp noen spørsmål som kan være naturlig å stille til datamaterialet fra taxi-sjåføren.

Hvilket resultat er i midten?

Dersom vi skal finne hva som er i midten må vi først sette resultatene i rekkefølge, for eksempel fra lavest til høyest:

12 - 14 - 17 - 17 - 21 - 29 - 37

Her ser vi at 17 turer er det resultatet i midten. Dette kalles for øvrig for **median**.

Spørsmål til diskusjon: hva om det er 2 i midten? Når inntreffer dette?

Hva er det vanligste resultatet?

Er det et resultat som kommer oftere enn andre? I datamaterialet til taxi-sjåføren ser vi at 17 turer er det resultatet som forekommer oftest. Dette kalles for øvrig for **typetall**.

Spørsmål til diskusjon: hva om det er flere observasjoner som forekommer oftest? Hva om ingen observasjoner forekommer flere ganger?

Hva om alle resultatene hadde vært like?

Tenk om sjåføren kunne fordelt turene slik at det ble kjørt like mange turer hver dag, istedenfor mange turer noen dager og få turer andre dager? Dette kalles for øvrig for **gjennomsnitt**.

I så fall må vi først finne ut hvor mange turer sjåføren kjørte til sammen. Deretter må vi fordele disse turene på antall dager. Dette kan skrives slik:

$$\text{Gjennomsnitt} = \frac{\text{sum data}}{\text{antall observasjoner}} = \frac{\text{sum turer}}{\text{antall dager}} = \frac{147}{7} = 21$$

Spørsmål til diskusjon: for hvilke typer undersøkelser er det ikke mulig å regne gjennomsnitt?

Disse tre analysene kalles for **sentralmål**. Blir du bedt om å finne **sentralmålene** til et datamateriale er det disse analysene du skal gjøre.

Hva om vi legger sammen resultatene underveis?

Det kan kanskje være interessant å vite hvor mange turer sjåføren har kjørt fra mandag til onsdag, eller fra mandag til fredag. Dette kalles for øvrig **kumulativ frekvens**. Kumulativ kommer av ordet akkumulere, som betyr å samle opp.

Vi kunne naturligvis skrevet det slik:

Kumulativ frekvens for mandag: antall turer mandag.

Kumulativ frekvens for tirsdag: antall turer mandag + tirsdag

Kumulativ frekvens for onsdag: antall turer mandag + tirsdag + onsdag

osv...

men det er mer fornuftig å gjøre dette i en tabell:

	A	B	C
1	Dag	Antall turer	Kumulativ frekvens
2	Mandag	14	14
3	Tirsdag	17	31
4	Onsdag	12	43
5	Torsdag	21	64
6	Fredag	29	93
7	Lørdag	37	130
8	Søndag	17	147

Spørsmål til diskusjon: nedenfor har vi vist formlene som ble brukt for å regne ut den kumulative frekvensen for dette datamaterialet. Hvordan er regnearket bygd opp?

	A	B	C
1	Dag	Antall turer	Kumulativ frekvens
2	Mandag	14	=B2
3	Tirsdag	17	=C2+B3
4	Onsdag	12	=C3+B4
5	Torsdag	21	=C4+B5
6	Fredag	29	=C5+B6
7	Lørdag	37	=C6+B7
8	Søndag	14	=C7+B8

Hvor stor (prosent)andel utgjør hvert resultat?

Det kan kanskje være interessant for sjåføren å vite hvor mange prosent av turene som ble kjørt på hver av dagene. Vi har tidligere sett at vi kan finne dette ved å lage et sektordiagram, men det kan også gjøres ved regning.

Dette kalles **relativ frekvens**, og dersom vi legger sammen prosentene underveis kalles dette **kumulativ relativ frekvens**. Dette er også fornuftig å gjøre i en tabell:

	A	B	C	D
1	Dag	Antall turer	Relativ frekvens	Kumulativ relativ frekvens
2	Mandag	14	9,5 %	9,5 %
3	Tirsdag	17	11,6 %	21,1 %
4	Onsdag	12	8,2 %	29,3 %
5	Torsdag	21	14,3 %	43,5 %
6	Fredag	29	19,7 %	63,3 %
7	Lørdag	37	25,2 %	88,4 %
8	Søndag	17	11,6 %	100,0 %
9	Sum turer	147	100,0 %	

Spørsmål til diskusjon: nedenfor har vi vist formlene som ble brukt for å regne ut den relative frekvensen for dette datamaterialet. Hvordan er regnearket bygd opp?

	A	B	C
1	Dag	Antall turer	Relativ frekvens
2	Mandag	14	=B2/\$B\$9
3	Tirsdag	17	=B3/\$B\$9
4	Onsdag	12	=B4/\$B\$9
5	Torsdag	21	=B5/\$B\$9
6	Fredag	29	=B6/\$B\$9
7	Lørdag	37	=B7/\$B\$9
8	Søndag	17	=B8/\$B\$9
9	Sum turer	=SUMMER(B2:B8)	=B9/\$B\$9

Er det stor forskjell på resultatene?

Hvilken dag kjører sjåføren færrest turer? Hvor mange turer kjører sjåføren på den travleste dagen? Hvor stor er forskjellen mellom det høyeste og det laveste antall turer? Dette kalles for øvrig **variasjonsbredde**, og er en del av det som kalles **spredningsmål**.

Vi ser at sjåføren kjører 37 turer på den travleste dagen, og 12 turer på den roligste dagen. Vi kan dermed regne ut **variasjonsbredden** slik:

Variasjonsbredde = høyest resultat - lavest resultat = 37 turer - 12 turer = 25 turer.

Oppgave 4

Nedenfor ser du karakterene til en elev på vurderinger i første termin på VG1:

3 - 4 - 2 - 4 - 5 - 2 - 3 - 6 - 3 - 4 - 5 - 3

- Hvilken karakter er den midterste karakteren til denne eleven?
- Hva er den vanligste karakteren denne eleven har fått?
- Hvor høy er gjennomsnittskarakteren til denne eleven?
- På hvor mange prosent av vurderingene fikk eleven karakteren 5?
- Hvor stor var forskjellen på den høyeste og den laveste karakteren for denne eleven?

Oppgave 5

Vi spurte 8 elever hvor mye penger de hadde brukt i kantina i storefri. Nedenfor finner du svarene de ga (i kroner):

55, 70, 45, 60, 130, 50, 65 og 70

- Hvor stor var forskjellen i pengebruk mellom den som brukte mest og den som brukte minst?
- Hvor mye brukte hver av elevene i gjennomsnitt?
- Hva er midtpunktet til dette datamaterialet?
- Hvor mange prosent av elevene brukte mer penger enn gjennomsnittet? Hvor mange prosent av elevene brukte mindre penger enn gjennomsnittet?

Oppgave 6

Finn sentralmål og spredningsmål til datamaterialet i oppgave 1.

Analyse i ExCel

ExCel kan forenkle analysearbeidet for oss dersom vi kjenner kommandoene. Nedenfor finner du en oversikt over hvordan du kan bruke ExCel til å finne **sentralmål** og **spredningsmål** til et datamateriale.

Dersom du skal skrive inn kommandoen gjør du det i følgende rekkefølge:

1. Begynn med å skrive = [kommandoen]
2. Dobbeltklikk på kommandoen som kommer opp.
3. Marker tallene du ønsker at ExCel skal analysere
4. Trykk «Enter»

Hvilken analyse	Kommando
Gjennomsnitt	=gjennomsnitt(datamaterialet)
Median	=median(datamaterialet)
Typetall	=modus(datamaterialet)
Variasjonsbredde	=maks(datamaterialet) - min(datamaterialet)
Standardavvik	=stdav.p

Standardavvik er et **spredningsmål**. Ved å regne ut **standardavviket** til et **tallmateriale** sammenligner vi hver enkelt **observasjon** med **gjennomsnittet**, og **standardavviket** vil være en samlet vurdering av denne forskjellen. Det betyr at jo mer hver enkelt **observasjon** avviker fra **gjennomsnittet**, jo høyere blir **standardavviket**. Det motsatte gjelder også: jo høyere **standardavvik**, jo større spredning blant **observasjonene**.

Merk: disse kommandoene fungerer kun når datamaterialet er skrevet som en liste med tall, slik det er gjort i eksempelet med taxi-turer.

Dersom du skal finne **sentral- og spredningsmål** når data er samlet i kategorier, som i oppgave 3 må dette løses på en annen måte. Dette skal du lære senere dette skoleåret.

ExCel-analyse av taxi-turene:

	A	B
1	Dag	Antall turer
2	Mandag	14
3	Tirsdag	17
4	Onsdag	12
5	Torsdag	21
6	Fredag	29
7	Lørdag	37
8	Søndag	17
9	Sum	147
10		
11	Median	17
12	Typetall	17
13	Gjennomsnitt	21
14	Variasjonsbredde	25
15	Standardavvik	8,3

	A	B
1	Dag	Antall turer
2	Mandag	14
3	Tirsdag	17
4	Onsdag	12
5	Torsdag	21
6	Fredag	29
7	Lørdag	37
8	Søndag	17
9	Sum	=SUMMER(B2:B8)
10		
11	Median	=MEDIAN(B2:B8)
12	Typetall	=MODUS(B2:B8)
13	Gjennomsnitt	=GJENNOMSNIITT(B2:B8)
14	Variasjonsbredde	=B7-B4
15	Standardavvik	=STDAV.P(B2:B8)

Legg merke til at vi forklarer hvilke analyser vi utfører.

Oppgave 7

Bruk datamaterialet du finner i oppgave 2, og gjør en analyse av sentralmål og spredningsmål ved hjelp av ExCel.

Spørsmål til diskusjon: kommandoer er ment for å forenkle arbeidet. Er det noen av kommandoene som fremstår som en mer tungvint metode enn å utføre analysen selv? Kan dette variere ut fra størrelsen på datamaterialet?

Presentasjonsoppgave

Ta frem presentasjonen av ditt lokale parti. Bruk ExCel til å gjøre så mange analyser som mulig av tallmaterialet.

I presentasjonen skal du inkludere relevante kommentarer.

En eksamensoppgave

Steffen bruker en app for å samle data om sykkelturene sien.

Han setter dataene i en tabell.

Tabellen inneholder to typer opplysninger:

1. Gjennomsnittsfart for hver kilometer
2. Antall meter stigning for hver kilometer

Vennene var imponerte, og det hadde blant annet disse kommentarene og spørsmålene til Steffen:

«Wow, Steffen. Det ble mange mil. Hvor lang tid brukte du?»

«Oj, det ble høyt etter hvert. Hvor mange høydemeter ble det i alt?»

«Skal si du holdt bra gjennomsnittsfart»

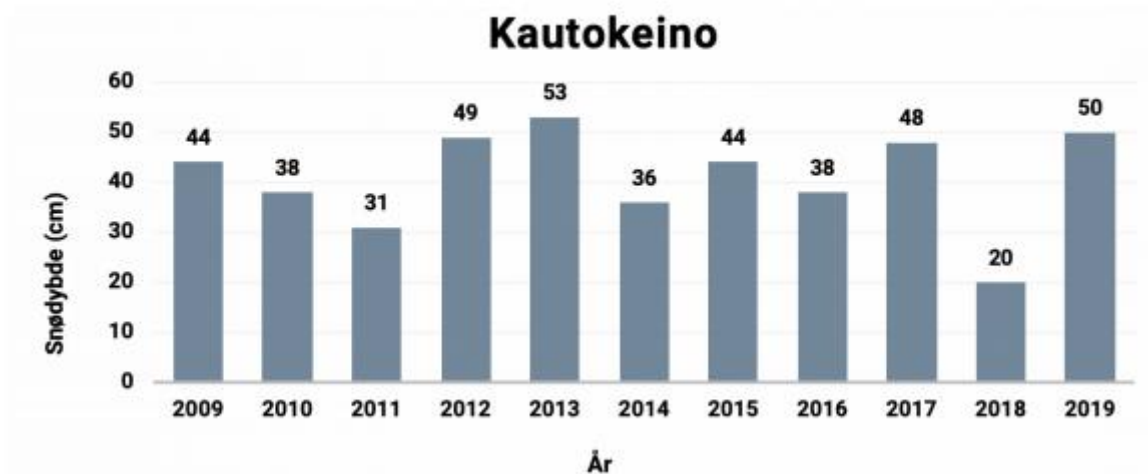
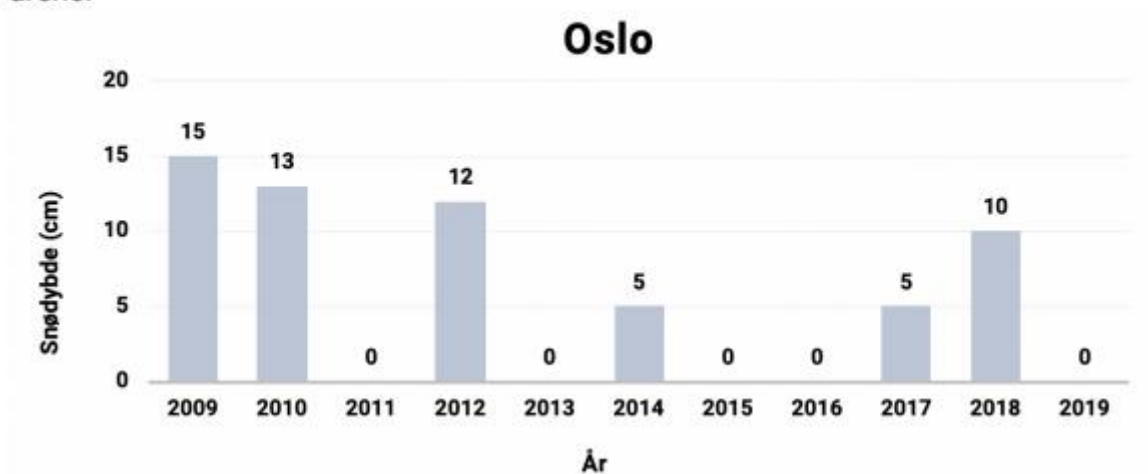
«Litt av ei løype, det går jo opp og ned hele tiden».

Bruk tabellen nedenfor og gjør beregninger, lag diagrammer og gi en beskrivelse av sykkelturen.

Kilometer (Første, andre ...)	Gjennomsnittsfart denne kilometeren (km/h)	Stigning denne kilometeren (m)
1.	22,92	27
2.	23,66	25
3.	17,26	65
4.	23,83	13
5.	36,84	-37
6.	38,02	-42
7.	17,36	-29
8.	23,38	-25
9.	24,34	10
10.	23,03	19
11.	16,76	26
12.	15,07	17
13.	14,69	7
14.	17,38	13
15.	24,70	-13
16.	15,67	47
17.	17,08	-7
18.	14,81	39
19.	15,06	42
20.	15,63	44

En eksamensoppgave

Diagrammene nedenfor viser snødybden i Oslo og i Kautokeino julaften de 11 siste årene.



- a) Bestem gjennomsnittet og standardavviket for snødybdene i Oslo og for snødybdene i Kautokeino.

Etter å ha regnet ut gjennomsnittet for Oslo og Kautokeino kom Isak med følgende påstand:

«Siden gjennomsnittet for Kautokeino ble høyere enn gjennomsnittet for Oslo, må standardavviket for Kautokeino også bli høyere enn standardavviket for Oslo. Det er alltid slik at det datamaterialet som har høyest gjennomsnitt, også har høyest standardavvik.»

- b) Er påstanden riktig? Begrunn svaret ditt.

En eksamensoppgave

Nedenfor ser du en liste over skuespillere som har vunnet Oscar for beste kvinnelige hovedrolle de siste 20 årene.

År	Navn	Alder
2018	Olivia Colman	44
2017	Frances McDormand	60
2016	Emma Stone	28
2015	Brie Larson	26
2014	Julianne Moore	54
2013	Cate Blanchett	44
2012	Jennifer Lawrence	22
2011	Meryl Streep	62
2010	Natalie Portman	29
2009	Sandra Bullock	45

År	Navn	Alder
2008	Kate Winslet	33
2007	Marion Cotillard	32
2006	Helen Mirren	61
2005	Reese Witherspoon	29
2004	Hilary Swank	30
2003	Charlize Theron	28
2002	Nicole Kidman	35
2001	Halle Berry	35
2000	Julia Roberts	33
1999	Hilary Swank	25

- a) Bestem gjennomsnittet, medianen, variasjonsbredden og standardavviket for alderen til de kvinnelige prisvinnerne disse 20 årene.

Tabellen nedenfor viser tilsvarende verdier for skuespillere som har vunnet Oscar for beste mannlige hovedrolle de siste 20 årene.

Gjennomsnitt	44 år
Median	43,5 år
Variasjonsbredde	31 år
Standardavvik	8 år

- b) Hva kan du si om aldersfordelingen blant de mannlige skuespillerne sammenliknet med de kvinnelige ut fra disse verdiene og resultatene fra oppgave a)?



Behov for å samle observasjoner

Frem til nå har vi arbeidet med datamateriale som inneholder relativt få observasjoner. Når antall observasjoner er relativt lavt går det fint å lage en liste med alle verdiene. Dersom antall observasjoner øker, vil også lista med verdier øke. Til slutt vil lista bli så lang at den blir uoversiktlig.

Dersom et datamateriale består av mange like observasjoner, kan det være hensiktsmessig å samle observasjonene i **kategorier**.

Dersom et datamateriale inneholder mange observasjoner som er nesten like, kan det være hensiktsmessig å samle observasjonene i **grupper** eller **klasser**.

Analyse av kategorier

En lærer spurte klassen hvor mange transportmidler elevene vanligvis bruker frem og tilbake til skolen. Læreren registrerte følgende svar, skrevet i stigende rekkefølge:

0 0 0 0 1 1 1 1 2 2 2 2 2 2 2 2 3 4 4 4 4 4 4 4 5 5 5 6 6

Læreren systematiserte svarene i tabellen nedenfor:

X kategori/verdi	f frekvens
Antall transportmidler	Antall elever
0	4
1	4
2	8
3	1
4	9
5	2
6	1

Når flere av svarene er like kan de samles i kategorier. Dette gir en mer ryddig presentasjon av resultatene, men medfører en utfordring når vi skal utføre en analyse over sentra- og spredningsmål.

I kategori-kolonnen (**X**) skrives hvilke svar som har blitt registrert.

I frekvens-kolonnen (**f**) skrives hvor mange som har gitt dette svaret.

Læreren ønsket deretter å gjøre en analyse av datamaterialet, og læreren ønsket å gjøre dette ut fra informasjonen presentert i tabellen.

Variasjonsbredde

Variasjonsbredde betyr forskjellen på høyeste og laveste kategori/verdi. I dette tilfellet blir det forskjellen på flest og færrest transportmidler.

Vi ser i tabellen at det høyeste svaret var 6 transportmidler, mens det laveste svaret var 0 transportmidler. Forskjellen mellom disse svarene blir: $6 - 0 = 6$.

Variasjonsbredden i datamaterialet er dermed 6.

Median

Median er det **verdien** til det midterste svaret i en ordnet rekkefølge. I dette tilfellet har det blitt avgitt 29 svar og svaret i midten blir dermed svar nummer 15. Hva står det på dette svaret?

Vi ser at de fire første svarene er 0, deretter er det fire som har svart 1. Dette blir til sammen 8 svar. Videre er det åtte som har svart 2, og til sammen utgjør dette de 16 første svarene.

Læreren var ute etter svar nummer 15, og ut fra opptellingen ovenfor skjønner læreren at svar nummer 15 var 2 transportmidler.

Medianen i tallmaterialet er dermed 2.

Spørsmål til drøfting: hva blir medianen dersom det er to svar i midten, og disse havner i hver sin kategori?

Typetall

Typetall er den kategorien/verdien med høyest frekvens. I dette tilfellet betyr det hvilket svar som ble gitt flest ganger.

I tabellen ser vi at 4 transportmidler ble svart 9 ganger, og dette svaret er det som ble avgitt flest ganger.

Typetallet i datamaterialet er dermed 4.

Spørsmål til diskusjon: hva blir typetallet dersom flere kategorier har den høyeste frekvensen?

Gjennomsnitt

Gjennomsnitt regnes ved å dividere **sum data** på **antall observasjoner**.

I dette tilfellet blir regnestykket: $\frac{\text{sum transportmidler}}{\text{antall elever}}$.

Det var 29 elever som svarte på undersøkelsen, men hvor mange transportmidler brukte klassen til sammen? For å finne ut dette må vi regne ut antall transportmidler fra hver kategori, og legge sammen disse svarene:

0 transportmidler · 4 = 0 transportmidler

1 transportmidler · 4 = 4 transportmidler

2 transportmidler · 8 = 16 transportmidler

3 transportmidler · 1 = 3 transportmidler

4 transportmidler · 9 = 36 transportmidler

5 transportmidler · 2 = 10 transportmidler

6 transportmidler · 1 = 6 transportmidler

Til sammen benytter klassen 75 transportmidler frem og tilbake til skolen. I gjennomsnitt blir dette:

$$\frac{\text{sum transportmidler}}{\text{antall elever}} = \frac{75 \text{ transportmidler}}{29 \text{ elever}} = 2,6 \text{ transportmidler per elev.}$$

Gjennomsnittet i datamaterialet er dermed 2,6.

Oppgave 8

En lærer spurte klassen hvor mange transportmidler elevene vanligvis bruker frem og tilbake til skolen. Læreren registrerte svarene i tabellen nedenfor:

Antall transportmidler	Frekvens	X · f	Kumulativ frekvens
0	4		
1	3		
2	8		
3	1		
4	6		
5	3		
6	2		
Sum			

Finn sentral- og spredningsmål til datamaterialet presentert i tabellen ovenfor.

Oppgave 9

En klasse oppnådde følgende karakterer på de tre første matematikkprøvene:

Karakter	Frekvens første prøve	Frekvens andre prøve	Frekvens tredje prøve
1	0	2	1
2	3	4	4
3	5	7	6
4	8	5	4
5	4	3	4
6	1	0	2

Gjør en analyse av minst en av matteprøvene fra tabellen ovenfor.

En eksamensoppgave (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

- Hvor mange prosent av elevene fikk karakteren 1 eller 2?
- Bestem mediankarakteren.
- Bestem gjennomsnittskarakteren.

En eksamensoppgave (del 1)

Tone spiller håndball. Tabellen nedenfor viser den kumulative frekvensen for antall mål hun skåret i de seks første kampene hun spilte.

Kamp nummer	Kumulativ frekvens for antall mål Tone skåret
1	4
2	12
3	15
4	21
5	25
6	30

- Hvor mange mål skåret Tone i kamp nummer 4?
- Hvor mange mål skåret Tone i gjennomsnitt per kamp?

Analyse av kategorier i ExCel

En analyse av et datamateriale, også når informasjonen er kategorisert, kan gjøres raskere og mer oversiktlig i ExCel.

Nedenfor har vi gjennomført analysen av lærerens undersøkelse av antall transportmidler elevene i klassen bruker for å reise frem og tilbake til skolen:

	A	B	C	D
1	<i>Kolonnen brukes til å finne:</i>			
2	<i>Variasjonsbredde</i>	<i>Typetall</i>	<i>Median</i>	<i>Gjennomsnitt</i>
3	X Antall transportmidler	f Antall elever	Kumulativ frekvens	X · f
4	0	4	4	0
5	1	4	8	4
6	2	8	16	16
7	3	1	17	3
8	4	9	26	36
9	5	2	28	10
10	6	1	29	6
11	Sum	29		75
12				
13	Median		2	
14	Typetall		4	
15	Gjennomsnitt		2,6	
16	Variasjonsbredde		6	

	A	B	C	D
1	<i>Kolonnen brukes til å finne:</i>			
2	<i>Variasjonsbredde</i>	<i>Typetall</i>	<i>Median</i>	<i>Gjennomsnitt</i>
3	X Antall transportmidler	f Antall elever	Kumulativ frekvens	X · f
4	0	4	=B4	=A4*B4
5	1	4	=C4+B5	=A5*B5
6	2	8	=C5+B6	=A6*B6
7	3	1	=C6+B7	=A7*B7
8	4	9	=C7+B8	=A8*B8
9	5	2	=C8+B9	=A9*B9
10	6	1	=C9+B10	=A10*B10
11	Sum	=SUMMER(B4:B10)		=SUMMER(D4:D10)
12				
13	Median	=A6		
14	Typetall	=A8		
15	Gjennomsnitt	=D11/B11		
16	Variasjonsbredde	=A10-A4		

Klarer du å lage denne tabellen i ExCel?

Oppgave 11

Liverpool FC		Newcastle United FC	
Antall mål per kamp	Frekvens	Antall mål per kamp	Frekvens
0	8	0	14
1	14	1	13
2	7	2	7
3	4	3	2
4	3	4	0
5	1	5	1
6	1	6	1

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

Bestem gjennomsnittet, median, typetall og variasjonsbredden for antall skårede mål per kamp for begge klubbene.

Oppgave 12

I løpet av første termin hadde en klasse gjennomført 3 prøver i matematikk. Læreren hadde registrert følgende karakterer på prøvene:


Karakter	Frekvens første prøve	Frekvens andre prøve	Frekvens tredje prøve
1	0	2	1
2	3	4	4
3	5	7	6
4	8	5	4
5	4	3	4
6	1	0	2

Lag en analyse av de tre prøvene, og kommenter resultatene.

Analyse av data i grupper

Noen undersøkelser produserer et datamateriale hvor mange av observasjonene er nesten like, men ikke helt identiske. Dersom vi skal presentere resultater fra slike undersøkelser kan det være hensiktsmessig å samle relativt like observasjoner i grupper. I slike grupper må vi spesifisere nederste og øverste verdi, slik at det er tydelig hvilke observasjoner som skal registreres. Vi bruker gjerne formuleringene «fra og med» og «opp til». Husk: en observasjon skal ikke registreres i 2 ulike grupper.

Tenk deg at skolen ønsker å kjøpe skolegenser til alle elevene på skolen, og at disse genserne er tilgjengelig i størrelsene S, M, L og XL. Skolens innkjøpsansvarlig antar at størrelsene avgjøres av høyden til elevene, og sendte følgende bestillingskjema:

	X Høyde i cm	Størrelse	f frekvens
	[140 - 150>	S	
	[150 - 165>	M	
	[165 - 184>	L	
	[184 - 200>	XL	

Lærerne for vg2 registrerte følgende høyder på elevene sine, målt i hele cm og ordnet i stigende rekkefølge:

142, 144, 144, 147, 148, 148, 150, 151, 151, 153, 154, 156, 156, 156, 156, 157, 157, 159, 159, 159, 160, 161, 163, 163, 164, 164, 165, 165, 165, 166, 166, 167, 167, 167, 168, 168, 169, 169, 170, 170, 170, 171, 171, 172, 172, 172, 173, 173, 173, 173, 174, 175, 175, 176, 177, 178, 178, 179, 179, 179, 180, 180, 180, 181, 181, 181, 182, 182, 184, 184, 184, 185, 185, 186, 186, 187, 187, 188, 189

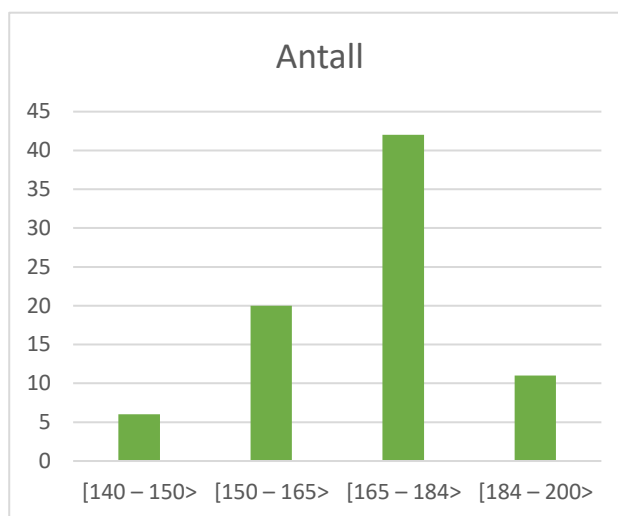
Spørsmål til diskusjon: i hvilken gruppe skal personer som måler 150, 165 og 184 plasseres?

Spørsmål til diskusjon: hvordan ville et søylediagram laget av datamaterialet ovenfor sett ut?

Kan du fullføre bestillingen til lærerne for vg2?

Lærerne fra vg2 returnerte følgende bestilling:

X Høyde i cm	Størrelse	f frekvens
[140 - 150>	S	6
[150 - 165>	M	20
[165 - 184>	L	42
[184 - 200>	XL	11



Da innkjøpsansvarlig mottok bestillingen kom det et ønske om å finne gjennomsnittshøyden og medianhøyden til elevene på påbygg. Innkjøpsansvarlig har ikke informasjon om høyden til hver enkelt elev, og denne analysen må gjøres på bakgrunn av datamaterialet presentert i tabellen ovenfor.

Gjennomsnitt i et gruppert materiale

Når informasjonen er presentert i et gruppert materiale vet vi kun hvor mange elever som er registrert i hver gruppe, men vi har ingen mulighet til å finne informasjon om hver enkelt elevs høyde. Dermed må vi anta at elevenes høyde fordeler seg slik innenfor en gruppe at vi kan tillegge alle elevene i den gruppa en høyde som er midt i gruppa. Dette kalles gruppas midtpunkt.

Det er registrert 6 elever i gruppa [140 - 150>. Når vi skal regne den samlede høyden til disse 6 elevene antar vi at alle 6 elevene er 145 cm høye, fordi 145 cm er midt mellom 140 cm og 150 cm. Tilsvarende gjelder for de andre gruppene.

Deretter blir regningen identisk med regningen på side 35. Vi har valgt å utføre regningen ved hjelp av Excel:

	A	B	C	D	E
1	Høyde		Xm	f	
2	Fra og med	Opp til	Midtpunkt	Frekvens	Xm · f
3	140	150	145	6	870
4	150	165	157,5	20	3150
5	165	184	174,5	42	7329
6	184	200	192	11	2112
7	Sum			79	13461
8					
9	Gjennomsnitt	170			

	A	B	C	D	E
1	Høyde		Xm	f	
2	Fra og med	Opp til	Midtpunkt	Frekvens	Xm · f
3	140	150	=(A3+B3)/2	6	=C3*D3
4	150	165	=(A4+B4)/2	20	=C4*D4
5	165	184	=(A5+B5)/2	42	=C5*D5
6	184	200	=(A6+B6)/2	11	=C6*D6
7	Sum			=SUMMER(D3:D6)	=SUMMER(E3:E6)
8					
9	Gjennomsnitt	=E7/D7			

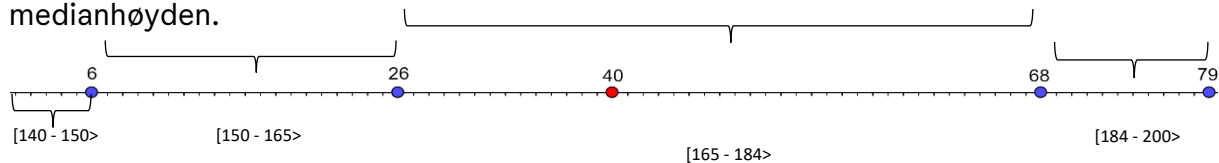
Spørsmål til diskusjon: hvorfor har vi brukt parenteser i utregningen i kolonne C?

Median i et gruppert materiale

Informasjonen i tabellen forteller at det er 79 elever på vg2. Medianhøyden er høyden til elev nummer 40 i ordnet rekkefølge.

Steg 1 – i hvilken gruppe finner vi medianhøyden?

Ved å bruke kumulativ frekvens finner vi at elev nummer 40 er i gruppen med høyde [165 – 184>. Dette kan også visualiseres ved en tallinje. Tallene skrevet med blått markerer høyeste elev i hver gruppe. Tallet skrevet med rødt markerer eleven med medianhøyden.



Steg 2 – hvor i gruppa finner vi medianhøyden?

For å kunne si noe mer spesifikt om medianhøyden må vi gjøre en antakelse. Vi må anta at høyden til elevene er fordelt jevnt utover gruppa. Tallinjen ovenfor viser at medianhøyden sannsynligvis finnes i gruppas nedre halvdel, og nærmere 165 cm enn 184 cm.

Steg 3 – kan vi anslå en tilnærmet verdi?

Gruppa hvor vi finner medianen har en bredde på 19 cm. Antall elever som fikk registrert høyden sin i denne gruppa er 42. Medianens nummer i gruppa er 14. Vi antar at elevenes høyde er jevnt fordelt utover gruppa. I vårt tilfelle betyr det at gruppebreddens 19 cm fordeles jevnt over høyden til de 42 elevene som ble registrert i gruppa. Det betyr at vi tenker at hver elev er 0,45 cm høyere enn forrige elev.

Gruppens nedre grense er 165 cm. I og med at medianhøyden er nummer 14 i denne gruppa, betyr det at høyden har vokst med 0,45 cm fjorten ganger fra 165 cm. Vi kan dermed utføre følgende regnestykke:

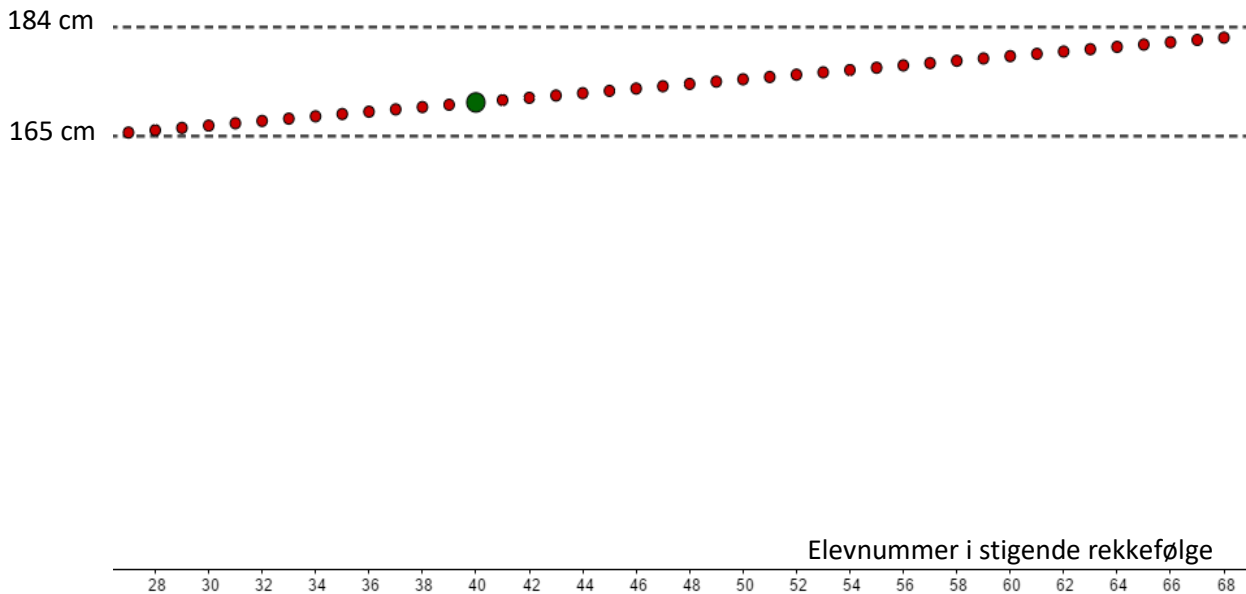
$$165 \text{ cm} + 0,45 \text{ cm} \cdot 14 = 165 \text{ cm} + 6,3 \text{ cm} = 171 \text{ cm}.$$

Generelt kan du regne ut medianen i gruppert materiale slik:

$$\text{Median} = \text{nedre grense} + \frac{\text{gruppebredden}}{\text{antall observasjoner i gruppa}} \cdot \text{medianens nummer i gruppa}$$

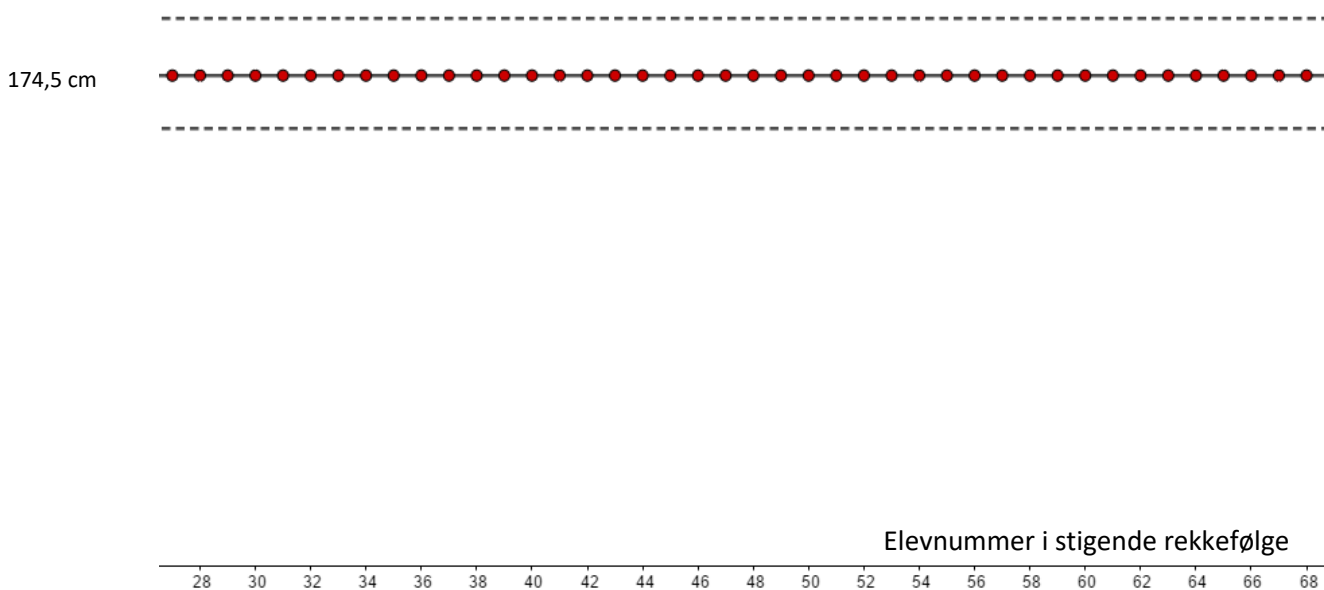
Visualisering av fordeling av høyder i gruppa [165 - 184>, median:

Ved å kun studere tabellen på side 41 har vi ingen mulighet til å vite høyden til hver enkelt elev. Vi antar dermed at høyden til den laveste eleven er 165 cm og at høyden til de andre elevene vokser jevnt opp mot 184 cm. De 42 elevene med høyde registrert i denne gruppa kan dermed visualiseres slik, hvor høyden på hver enkelt elev er markert med et rødt punkt:



Visualisering av fordeling av høyder i gruppa [165 - 184>, gjennomsnitt:

Ved å kun studere tabellen på side 41 har vi ingen mulighet til å vite høyden til hver enkelt elev. Vi antar dermed at høyden til alle elevene i denne gruppa er identisk med gruppas midtpunkt. Dette kan visualiseres slik:



Oppgave 13

I oppgave 10 viste vi en oversikt over hvor mange ganger hver enkelt elev på Hellerud måtte være i karantene. Nedenfor finner du en oversikt over hvor mange dager hver enkelt elev har vært i karantene.

X Antall dager i karantene	f Frekvens
[0 - 5>	75
[5 - 10>	235
[10 - 20>	100
[20 - 30>	150
[30 - 35>	25

Lag en presentasjon og en analyse av datamaterialet. Du velger selv hvilket diagram du ønsker å bruke. I analysen skal du presentere relevante sentralmål.

Oppgave 14

Tabellen nedenfor viser gruppering av årsinntekt for lønsmottakere i en tenkt kommune i Norge. Inntekten er oppgitt i antall 1 000 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

Finn gjennomsnittslønnen og medianlønnen for lønsmottakere i denne kommunen.

Histogram

Når vi skal presentere resultatene fra et klassesdelt 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** regner vi ved å dividere **frekvens** på **klassebredde**.

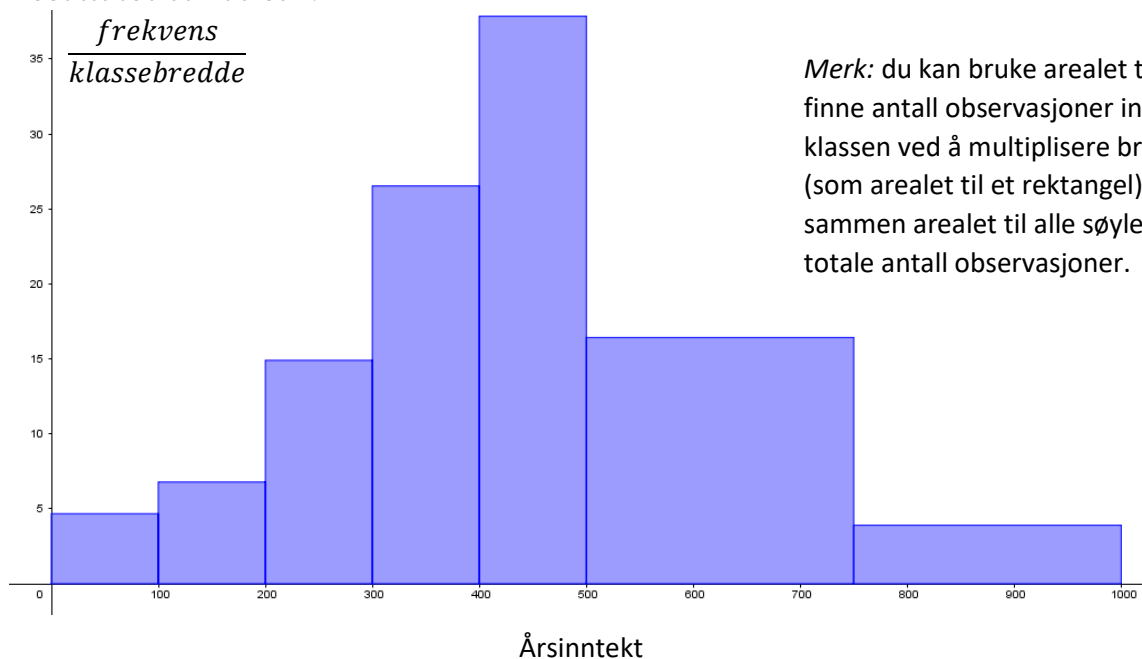
$$\frac{\text{frekvens}}{\text{klassebredde}} = \text{søylehøyde}$$

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:



Merk: du kan bruke arealet til en søyle for å finne antall observasjoner innenfor den klassen ved å multiplisere bredden og høyden (som arealet til et rektangel). Legger du sammen arealet til alle søylene finner du det totale antall observasjoner.

En eksamensoppgave (del 1)

Lise har hentet inn informasjon om høyden på husene i området der hun bor.

Høyde (meter)	Antall hus
$[3,5)$	2
$[5,7)$	8
$[7,9)$	10
$[9,13)$	8

- Bestem gjennomsnittshøyden på husene i området der Lise bor.
- Framstill datamaterialet i et histogram.

En eksamensoppgave (del 1)

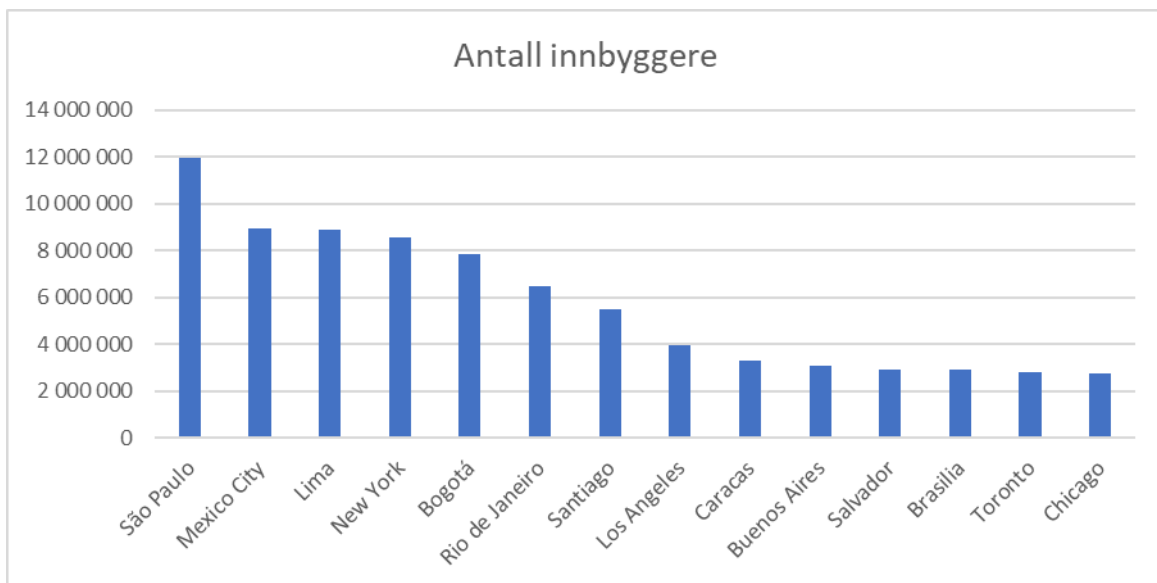
Tabellen nedenfor viser en oversikt over høydene til elevene ved en skole.

Høyde i cm	Frekvens
$[150,160)$	10
$[160,170)$	30
$[170,180)$	50
$[180,200)$	10

- Bestem gjennomsnittshøyden til elevene ved skolen.
- Lag et histogram som viser fordelingen av høydene.

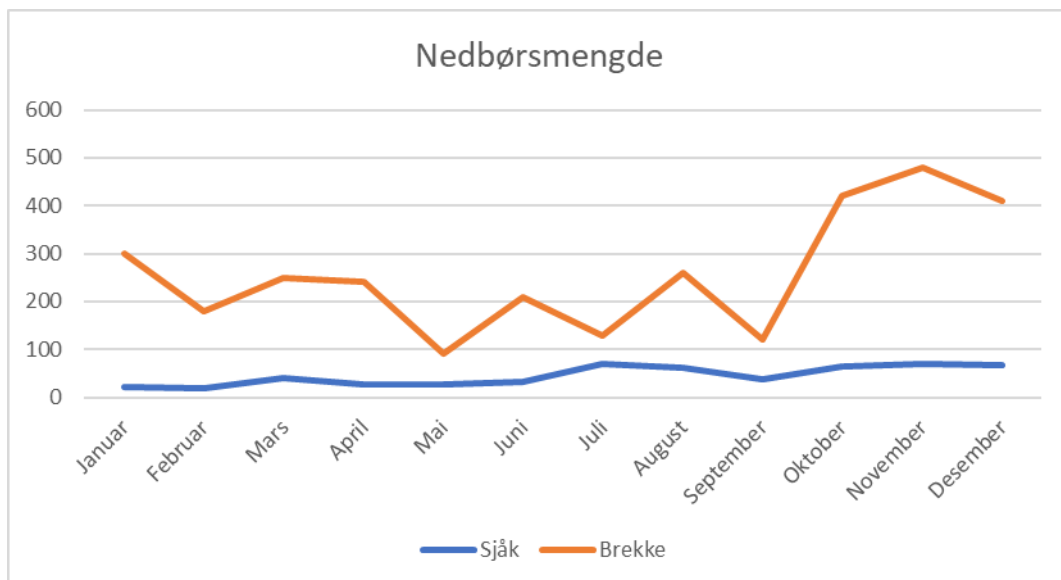
Løsningsforslag

Oppgave 1



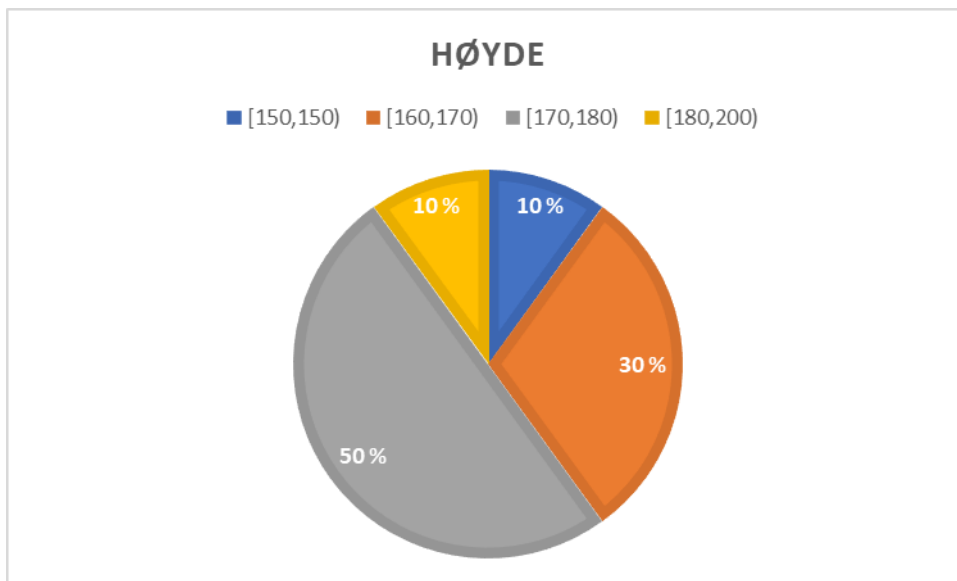
Vi valgte søylediagram, fordi tabellen kun viser et utvalg av byene i Sør- og Nord-Amerika.

Oppgave 2



Utvikling over tid er mest hensiktsmessig å vise gjennom et linjediagram.

Oppgave 3



Et sektordiagram er mest hensiktsmessig å benytte når den prosentvise fordelingen skal vises.

Oppgave 4

- Det er to karakterer i midten (3 og 4). Medianen blir da 3,5.
- Denne eleven fikk flest 3-ere.
- Gjennomsnittskarakteren er 3,7.
- Eleven fikk karakteren 5 på omtrent 17 % av prøvene.
- Den beste prøven var 4 karakterer bedre enn den dårligste.

Oppgave 5

Vi spurte 8 elever hvor mye penger de hadde brukt i kantina i storefri. Nedenfor finner du svarene de ga (i kroner):

55, 70, 45, 60, 130, 50, 65 og 70

- Den som brukte mest penger brukte 85 kroner mer enn den som brukte minst.
- I gjennomsnitt brukte elevene omtrent 68 kroner.
- Midtpunktet til datamaterialet er 62,50 kroner.
- 37,5 % av elevene brukte mer enn gjennomsnittet, mens 62,5 % av de spurte brukte mindre enn gjennomsnittet.

Oppgave 6

I gjennomsnitt bor det 5 705 357 mennesker i hver av byene.

Medianantall innbyggere er 4 739 500.

Det bor 9 247 000 flere innbyggere i den største enn i den minste av byene på lista.

Ingen av byene har likt innbyggertall.

Oppgave 7

	A	B	C
1	Måned	Sjåk	Brekke
2	Januar	22	300
3	Februar	20	180
4	Mars	41	250
5	April	28	240
6	Mai	26	90
7	Juni	32	210
8	Juli	71	130
9	August	62	260
10	September	38	120
11	Oktober	65	420
12	November	70	480
13	Desember	66	410
14			
15	Gjennomsnitt	45,1	257,5
16	Median	39,5	245
17	Typetall	#I/T	#I/T
18	Variasjonsbredde	51	390
19	Standardavvik	19,3	120,0

	A	B	C
1	Måned	Sjåk	Brekke
2	Januar	22	300
3	Februar	20	180
4	Mars	41	250
5	April	28	240
6	Mai	26	90
7	Juni	32	210
8	Juli	71	130
9	August	62	260
10	September	38	120
11	Oktober	65	420
12	November	70	480
13	Desember	66	410
14			
15	Gjennomsnitt	=GJENNOMSNIIT(B2:B13)	=GJENNOMSNIIT(C2:C13)
16	Median	=MEDIAN(B2:B13)	=MEDIAN(C2:C13)
17	Typetall	=MODUS(B2:B13)	=MODUS(C2:C13)
18	Variasjonsbredde	=MAKSA(B2:B13)-MIN(B2:B13)	=MAKSA(C2:C13)-MIN(C2:C13)
19	Standardavvik	=STDAV.P(B2:B13)	=STDAV.P(C2:C13)

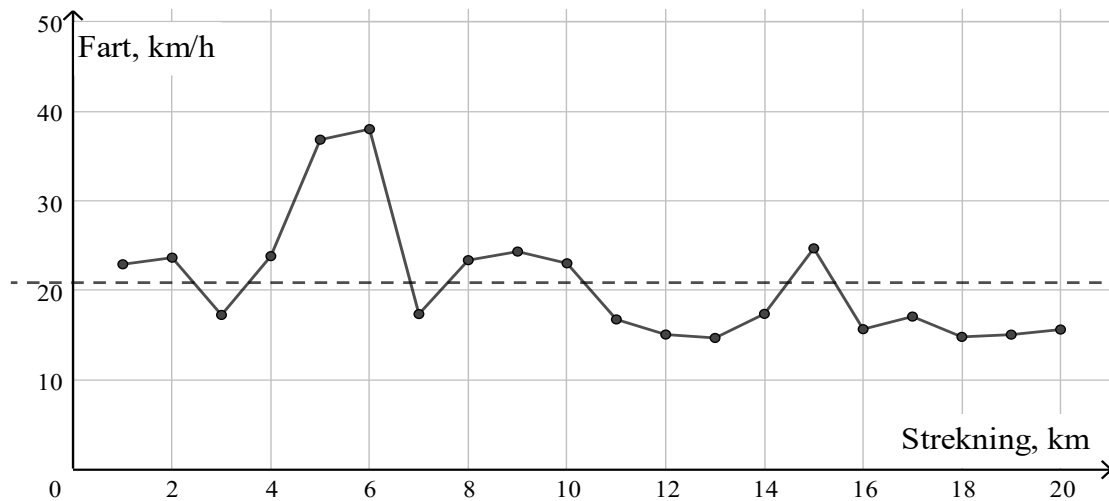
Eksamensoppgave side 174

Jeg skrev inn tallene i hver sin kolonne i regnearket i GeoGebra. Jeg laget lista «fart» av tallene for fart og brukte funksjonen «gsnitt» på lista «fart» for å finne gjennomsnittsfarten på turen. Gjennomsnittsfarten ble 20,87 km/h.

Tiden brukt på turen blir da

$$\frac{20 \text{ km}}{20,87 \text{ km/h}} = 0,958 \text{ h}$$
$$= 0,958 \cdot 60 \text{ min} = 57,50 \text{ min} = 57 \text{ min } 30 \text{ s}$$

Jeg brukte verktøyet «Polylinje» på kolonnene for km og fart og fikk tegnet en fartsprofil for turen, se nedenfor. Gjennomsnittsfarten på 20,87 km/h er tegnet inn som en stiplet linje.



Vi ser at farten har gått både opp og ned, men det gikk saktere i siste halvdel av turen enn i første.

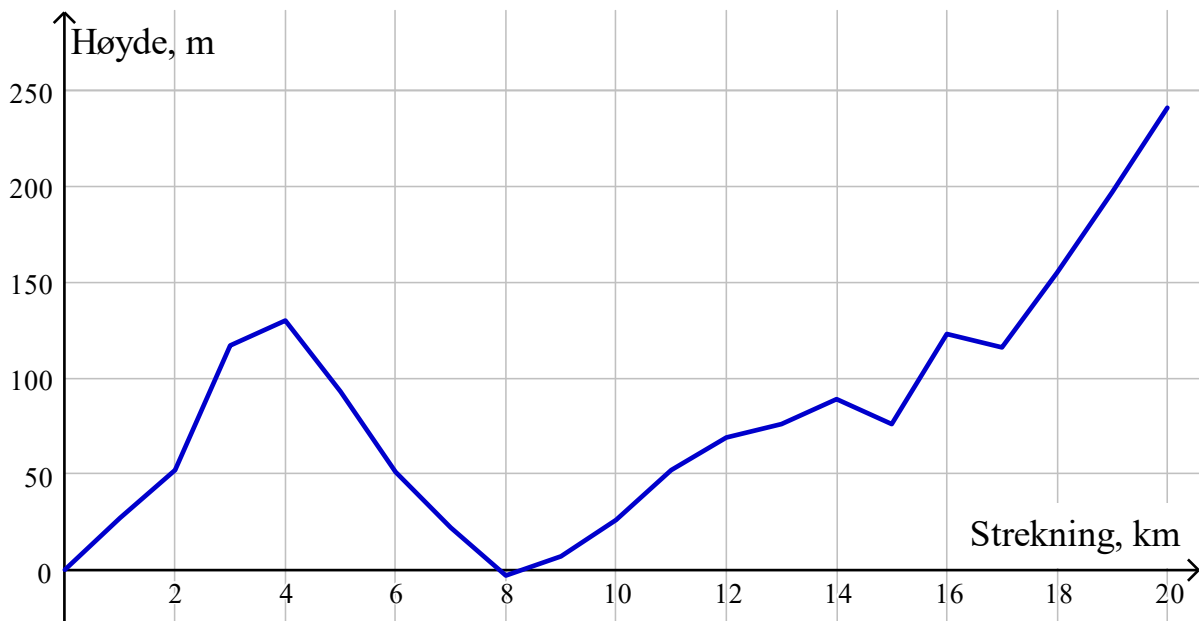
Jeg laget en ny kolonne i regnearket med de akkumulerte tallene for stigning, altså at tallene for hver km er lik summen av alle stigningene på kilometerne som alt er syklet.

	A	B	C	D
1	Kilometer	Gjennomsnittsfart denne kilometeren (km/h)	Stigning denne kilometeren (m)	Akkumulert stigning (m)
2	1	22.92	27	27
3	2	23.66	25	52
4	3	17.26	65	117
5	4	23.83	13	130

	A	B	C	D
1	Kilometer	Gjennomsnittsfart denne kilometeren (km/h)	Stigning denne kilometeren (m)	Akkumulert stigning (m)
2	1	22.92	27	C2
3	2	23.66	25	D2 + C3
4	3	17.26	65	D3 + C4
5	4	23.83	13	D4 + C5

Jeg brukte verktøyet «Polylinje» på kolonnene for kilometer og for akkumulert stigning og fikk tegnet en høydeprofil for turen, se nedenfor. Jeg lot turen starte på høyde null og la derfor inn punktet (0, 0) i lista for polylinja for at den skulle starte i origo. (Punktene U, V, osv. ble dannet automatisk av kommandoen «Polylinje».)

● høyde = Polylinje {(0, 0), U, V, W,



Vi ser at turen går mest oppover. Det er hovedsakelig bare fra 4 km til 8 km at det går nedover. Vi ser også at der det går nedover, er farten stor, og motsatt.

Kommentar:

Det går også an å bruke regneark i Excel til å løse denne oppgaven. Det er til og med mulig å tegne og skrive for hånd og ta bilde av det med webkameraet, dersom PC-en du skal levere inn eksamensoppgaven med, har webkamera.

Det er flere ting som kan regnes ut eller diskuteres, for eksempel gjennomsnittsstigningen per km eller gjennomsnittshøyden. Vi kan for eksempel undersøke om det er slik overalt at jo brattere oppover det er, jo saktere sykler han.

Eksamensoppgave side 174

Bruker Excel.

	A	B	C
1	Snødybde (cm) på julaften		
2	År	Oslo	Kautokeino
3	2009	15	44
4	2010	13	38
5	2011	0	31
6	2012	12	49
7	2013	0	53
8	2014	5	36
9	2015	0	44
10	2016	0	38
11	2017	5	48
12	2018	10	20
13	2019	0	50
14	Gjennomsnitt	5,45	41,00
15	Std.av.	5,73	9,24

b)

Påstanden er ikke riktig. Standardavviket sier noe om spredningen i tallmaterialet. Vi kan ha et datamateriale med høyt gjennomsnitt, men med mange tilnærmet like verdier; da vil standardavviket være lite selv om gjennomsnittet er høyt. Omvendt kan det være et datamateriale med mange veldig forskjellige verdier, som gir et høyt standardavvik uavhengig av gjennomsnittet.

Eksamensoppgave side 176

a)

Bruker funksjonene i Excel og får følgende statistikk:

B	C	D	E	F
Alder i år				
44				
60				
28	Gjennomsnitt	37,75	år	
26	Eldst	62	år	
54	Yngst	22	år	
44	Variasjonsbredde	40	år	
22	Median	33	år	
62	Standardavvik	12,6984251	år	
29				
45				
33				
32				
61				
29				
30				
28				
35				
35				
33				
25				

b)

Mennene er 6-7 år eldre enn damene, når de vinner en Oscar. Begge spredningsmålene forteller oss også at det er en mindre aldersvariasjon blant menn.

Oppgave 8

Antall transportmidler	Frekvens	$X \cdot f$	Kumulativt
0	4	0	4
1	3	3	7
2	8	16	15
3	1	3	16
4	6	24	22
5	3	15	25
6	2	12	27
Sum	27	73	

Analyse:

Typetall: 2 transportmidler

Variasjonsbredde: 6 transportmidler

Gjennomsnitt: 2,7 transportmidler

Median: 2 transportmidler

Oppgave 9

Analyse:	Første prøve	Andre prøve	Tredje prøve
Typetall	4	3	3
Variasjonsbredde	4	4	5
Gjennomsnitt	3,8	3,1	3,6
Median	4	3	3

Eksamensoppgave side 180

- 25 % fikk karakteren 1 eller 2
- Mediankarakteren var 3
- Gjennomsnittskarakteren var 3,2

Eksamensoppgave side 181

- Tone skåret 6 mål i kamp nr. 4
- Tone skåret 5 mål i gjennomsnitt per kamp

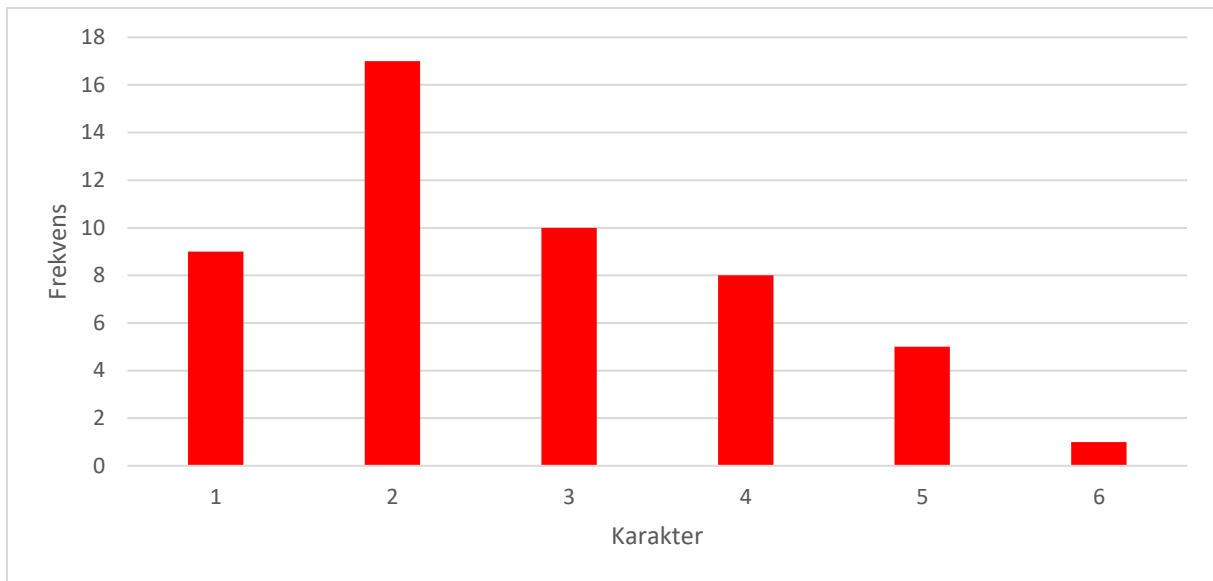
Oppgave 10

a)

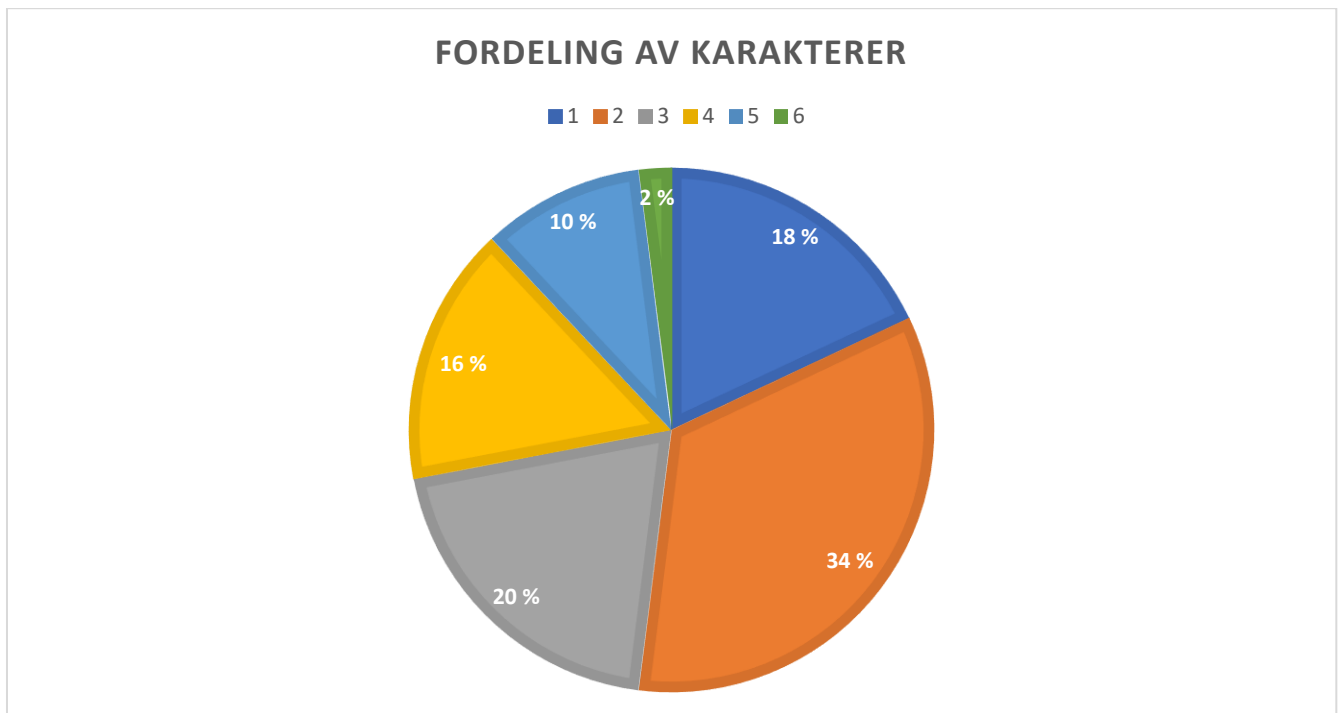
X Karakter	f Frekvens
1	9
2	17
3	10
4	8
5	5
6	1
Sum	50

b) Presentasjon av resultatet:

Antall:



Prosentvis fordeling av karakterer:



Oppgave 11

	A	B	C	D	E	F	G
1		Liverpool			Newcastle		
2	Ant. Mål	Frekvens	Sum mål	Kumulativt	Frekvens	Sum mål	Kumulativt
3	0	8	0	8	14	0	14
4	1	14	14	22	13	13	27
5	2	7	14	29	7	14	34
6	3	4	12	33	2	6	36
7	4	3	12	36	0	0	36
8	5	1	5	37	1	5	37
9	6	1	6	38	1	6	38
10	Sum	38	63		38	44	
11							
12	Analyse:	Liverpool	Newcastle				
13	Typetall	1	0				
14	Variasjonsbredde	6	6				
15	Gjennomsnitt	1,7	1,2				
16	Median	1	0,5				

	A	B	C	D	E	F	G
1		Liverpool			Newcastle		
2	Ant. Mål	Frekvens	Sum mål	Kumulativt	Frekvens	Sum mål	Kumulativt
3	0	8	=A3*B3	=B3	14	=A3*E3	=E3
4	1	14	=A4*B4	=D3+B4	13	=A4*E4	=G3+E4
5	2	7	=A5*B5	=D4+B5	7	=A5*E5	=G4+E5
6	3	4	=A6*B6	=D5+B6	2	=A6*E6	=G5+E6
7	4	3	=A7*B7	=D6+B7	0	=A7*E7	=G6+E7
8	5	1	=A8*B8	=D7+B8	1	=A8*E8	=G7+E8
9	6	1	=A9*B9	=D8+B9	1	=A9*E9	=G8+E9
10	Sum	=SUMMER(B3:B9)	=SUMMER(C3:C9)		=SUMMER(E3:E9)	=SUMMER(F3:F9)	
11							
12	Analyse:	Liverpool	Newcastle				
13	Typetall	=A4	=A3				
14	Variasjonsbredde	=A9-A3	=A9-A3				
15	Gjennomsnitt	=C10/B10	=F10/E10				
16	Median	=A4	=(A3+A4)/2				

Oppgave 12

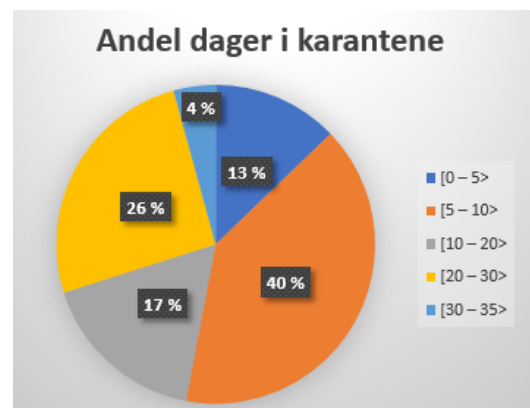
	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Første prøve				Andre prøve				Sum verdier			
2	Karakter	Frekvens	Sum verdier	Relativt	Kumulativt	Frekvens	Sum verdier	Relativt	Kumulativt	Frekvens	Sum verdier	Relativt	Kumulativt
3	1	0	0	0 %	0	2	2	10 %	2	1	1	5 %	1
4	2	3	6	14 %	3	4	8	19 %	6	4	8	19 %	5
5	3	5	15	24 %	8	7	21	33 %	13	6	18	29 %	11
6	4	8	32	38 %	16	5	20	24 %	18	4	16	19 %	15
7	5	4	20	19 %	20	3	15	14 %	21	4	20	19 %	19
8	6	1	6	5 %	21	0	0	0 %	21	2	12	10 %	21
9	Sum	21	79	100 %		21	66	100 %		21	75	100 %	
10													
11	Analyse:	Første prøve	Andre prøve	Tredje prøve									
12	Typetall	4	3	3									
13	Variasjonsbredde	4	4	5									
14	Gjennomsnitt	3,8	3,1	3,6									
15	Median	4	3	3									
16													
17	I tillegg kan det være interessant å sammenligne hvor mange prosent av elevene som fikk hver av karakterene på hver av prøvene.												
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													
34													
35													
36													
37													
38													
39													
40													
41													
42													
43													
44													
45													
46													
47													
48													
49													
50													
51													
52													
53													
54													
55													
56													
57													
58													
59													
60													
61													
62													
63													
64													
65													
66													
67													
68													
69													
70													
71													
72													
73													
74													
75													
76													
77													
78													
79													
80													
81													
82													
83													
84													
85													
86													
87													
88													
89													
90													
91													
92													
93													
94													
95													
96													
97													
98													
99													
100													

Oppgave 13

Antall registrerte i hver gruppe



Prosentvis fordeling av dager i karantene



	A	B	C	D	E	F
1	X		f			
2	Antall dager i karantene		Frekvens	X_m	X_m · f	Kumulativ frekvens
3	F.o.m	til				
4	0	5	75	2,5	187,5	75
5	5	10	235	7,5	1762,5	310
6	10	20	100	15	1500	410
7	20	30	150	25	3750	560
8	30	35	25	32,5	812,5	585
9	Sum		585		8012,5	
10						
11	Gjennomsnitt	14				
12	Median					
13	Steg 1	Er i gruppe [5 - 10>				
14	Steg 2	Er sannsynligvis høyt i gruppe [5 - 10>				
15	Steg 3	9,8				

	A	B	C	D	E	F
1	X		f			
2	Antall dager i karantene		Frekvens	X_m	X_m · f	Kumulativ frekvens
3	F.o.m	til				
4	0	5	75	$=(A4+B4)/2$	$=D4*C4$	$=C4$
5	5	10	235	$=(A5+B5)/2$	$=D5*C5$	$=F4+C5$
6	10	20	100	$=(A6+B6)/2$	$=D6*C6$	$=F5+C6$
7	20	30	150	$=(A7+B7)/2$	$=D7*C7$	$=F6+C7$
8	30	35	25	$=(A8+B8)/2$	$=D8*C8$	$=F7+C8$
9	Sum		$=SUMMER(C4:C8)$		$=SUMMER(E4:E8)$	
10						
11	Gjennomsnitt	$=E9/C9$				
12	Median					
13	Steg 1	Er i gruppe [5 - 1				
14	Steg 2	Er sannsynligvis				
15	Steg 3	$=A5+(5/C5*225)$				

Oppgave 14

	A	B	C	D	E	F
1	X		f			
2	Inntekt		Frekvens	X _m	X _m · f	Kumulativ frekvens
3	F.o.m	til				
4	0	100	467	50	23350	467
5	100	200	678	150	101700	1145
6	200	300	1490	250	372500	2635
7	300	400	2653	350	928550	5288
8	400	500	3785	450	1703250	9073
9	500	750	4106	625	2566250	13179
10	750	1000	987	875	863625	14166
11	1000	5000	45	3000	135000	14211
12	Sum		14211		6694225	
13						
14	Gjennomsnitt	471				
15	Median	Person nummer 7 106				
16	Steg 1	Er i gruppe [400 - 500>				
17	Steg 2	Er omtrent midt i gruppe [400 - 500]				
18	Steg 3	448				

	A	B	C	D	E	F
1	X		f			
2	Inntekt		Frekvens	X _m	X _m · f	Kumulativ frekvens
3	F.o.m	til				
4	0	100	467	$=(A4+B4)/2$	$=D4*C4$	$=C4$
5	100	200	678	$=(A5+B5)/2$	$=D5*C5$	$=F4+C5$
6	200	300	1490	$=(A6+B6)/2$	$=D6*C6$	$=F5+C6$
7	300	400	2653	$=(A7+B7)/2$	$=D7*C7$	$=F6+C7$
8	400	500	3785	$=(A8+B8)/2$	$=D8*C8$	$=F7+C8$
9	500	750	4106	$=(A9+B9)/2$	$=D9*C9$	$=F8+C9$
10	750	1000	987	$=(A10+B10)/2$	$=D10*C10$	$=F9+C10$
11	1000	5000	45	$=(A11+B11)/2$	$=D11*C11$	$=F10+C11$
12	Sum		$=SUMMER(C4:C11)$		$=SUMMER(E4:E11)$	
13						
14	Gjennomsnitt	$=E12/C12$				
15	Median	Person nummer 7 1				
16	Steg 1	Er i gruppe [400 - 50				
17	Steg 2	Er omtrent midt i gr				
18	Steg 3	$=A8+(100/C8*1818)$				

Eksamensoppgave side 191

a)

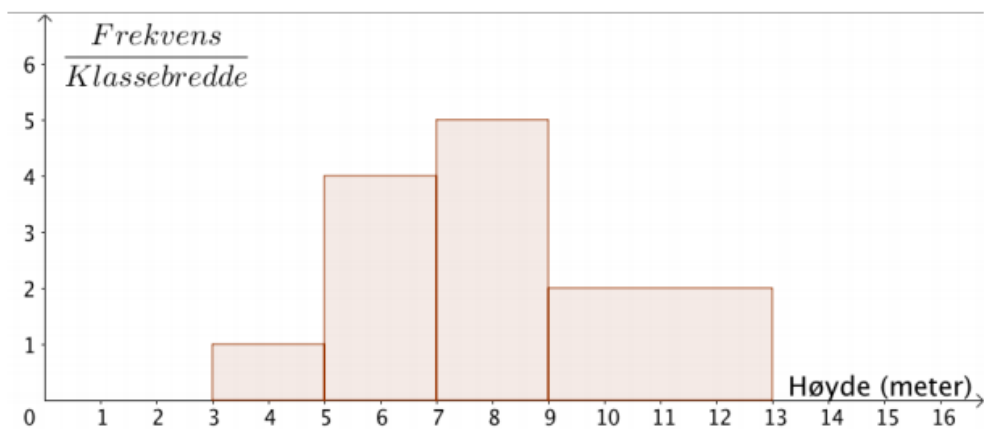
$$\frac{4 \cdot 2 + 6 \cdot 8 + 8 \cdot 10 + 11 \cdot 8}{2 + 8 + 10 + 8} = \frac{8 + 48 + 80 + 88}{28} = \frac{224}{28} = \frac{112}{14} = \frac{56}{7} = 8$$

Gjennomsnittshøyden på husene i området der Lise bor er 8 meter.

b) Regner ut histogramhøydene, som er gitt ved $\frac{\text{Frekvens}}{\text{Klassebredde}}$

$$\frac{2}{2} = 1 \quad \frac{8}{2} = 4 \quad \frac{10}{2} = 5 \quad \frac{8}{4} = 2$$

Tegner histogrammet:



Eksamensoppgave side 191

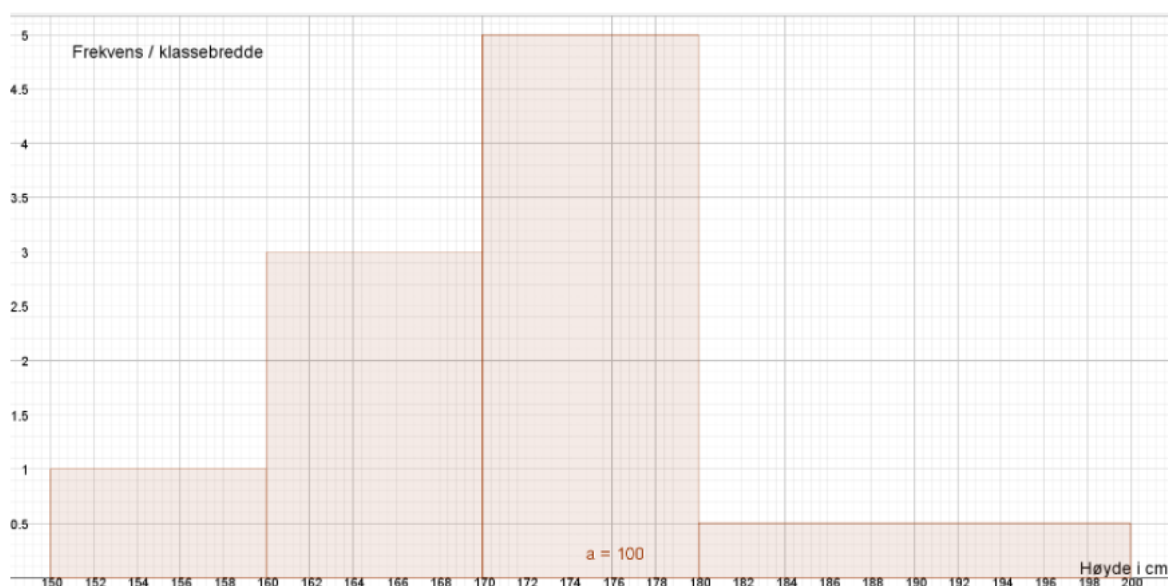
a)

Høyde i cm	Klassemidtpunkt, x_m	Frekvens, f	$f \cdot x_m$
[150, 160)	155	10	1550
[160, 170)	165	30	4950
[170, 180)	175	50	8750
[180, 200)	190	10	1900
Sum		100	17150

Gjennomsnitt: $\frac{17150}{100} = 171,5 \text{ cm}$

Gjennomsnittshøyden til elevene ved skolen er 171,5 cm.

Høyde i cm	Klassebredde, b	Frekvens, f	Histogramhøyde, $\frac{f}{b}$
[150, 160)	$160 - 150 = 10$	10	$\frac{10}{10} = 1$
[160, 170)	$170 - 160 = 10$	30	$\frac{30}{10} = 3$
[170, 180)	$180 - 170 = 10$	50	$\frac{50}{10} = 5$
[180, 200)	$200 - 180 = 20$	10	$\frac{10}{20} = 0,5$



PS: du må tegne histogrammet for hånd, siden dette er del 1.

	A	B		
	Uke	Antall bekreftet smittet av Covid-19		
1				
2	40	742		
3	41	1 072		
4	42	915		
5	43	1 096		
6	44	3 402		
7	45	4 162		
8				
9	Gjennomsnitt	1898	9	Gjennomsnitt =GJENNOMSNITT(B2:B7)
10	Median	1084	10	Median =MEDIAN(B2:B7)
11	Variasjonsbredde	3 420	11	Variasjonsbredde =MAKSA(B2:B7)-MIN(B2:B7)