

Undersøkelse 1

I en spørreundersøkelse har vi spurt 15 personer hvor mange epler de spiser i løpet av en uke. Svarene var disse:

7 4 5 4 1 0 6 5 4 8 1 6 8 0 14

Undersøkelse 2

300 personer ble spurt hva deres favorittdyr er.

- 46 personer svarte tiger
- 23 personer svarte løve
- 17 personer svarte krokodille
- 91 personer svarte hund
- 72 personer svarte katt
- 51 personer svarte andre dyr

10 personer testet hvor mange sekunder de kunne holde pusten. Resultatene ble disse:

47 124 61 38 97 84 101 79 56 40

Undersøkelse 3

Hentet fra medienorge.uib.no.

År	2009	2010	2011	2012	2013	2014
mobiltelefoner	2 365	2 500	2 250	2 200	2 400	2 100
uten smartfnk.	1 665	1 250	790	300	240	147
med smartfnk.	700	1 250	1 460	1 900	2 160	1 953

Statistikk handler grovt sett om to ting; å *presentere* og å *tolke*. For begge disse formålene har vi noen verktøy som vi skal vise ved hjelp av denne undersøkelsen.

0.1 Presentasjonsmetoder

I statistikk betyr *data* informasjon som er hentet inn, og *datasettet* er samlingen av disse dataene. I vårt eksempel på side 1 dataene er tallfølgen datasettet vårt. Skal vi presentere vår undersøkelse, bør vi vise datasettet slik at det er lett for andre å hva vi har funnet.

Frekvenstabell

I en frekvenstabell setter man opp dataene i en tabell som viser hvor mange ganger, *frekvensen*, hver unike data dukker opp.

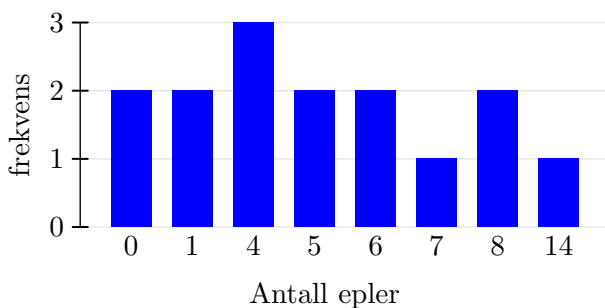
Undersøkelse 1

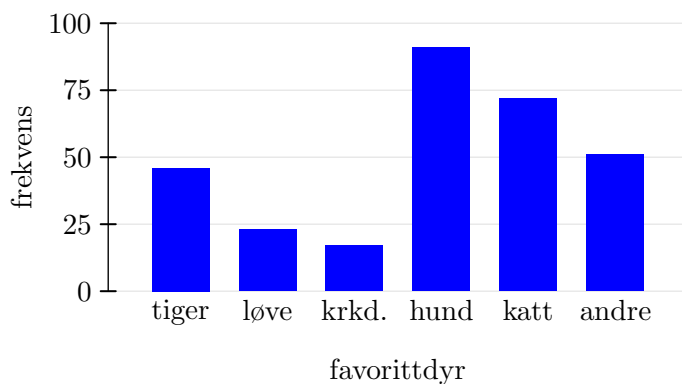
I vår undersøkelse har vi to 0, to 1, én 2, tre 4, to 5, to 6, én 7, én 8 og én 14. I en frekvenstabell skriver vi da

Antall epler	Frekvens
0	2
1	2
4	3
5	2
6	2
7	1
8	2
14	1

Søylediagram (stolpediagram)

Med et søylediagram presenterer vi dataene med søyler som viser frekvensen:

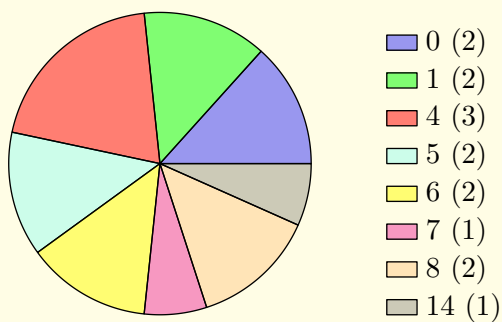




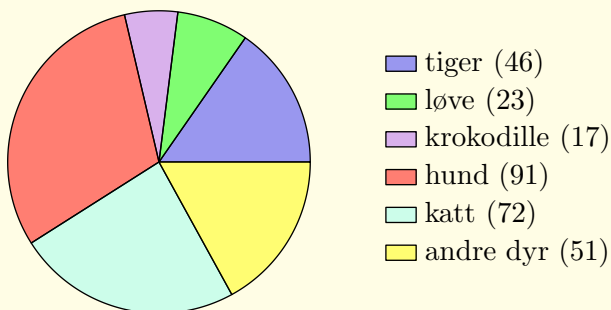
Sektordiagram (kakediagram)

I et sektordiagram vises frekvensene som sektorer av en sirkel:

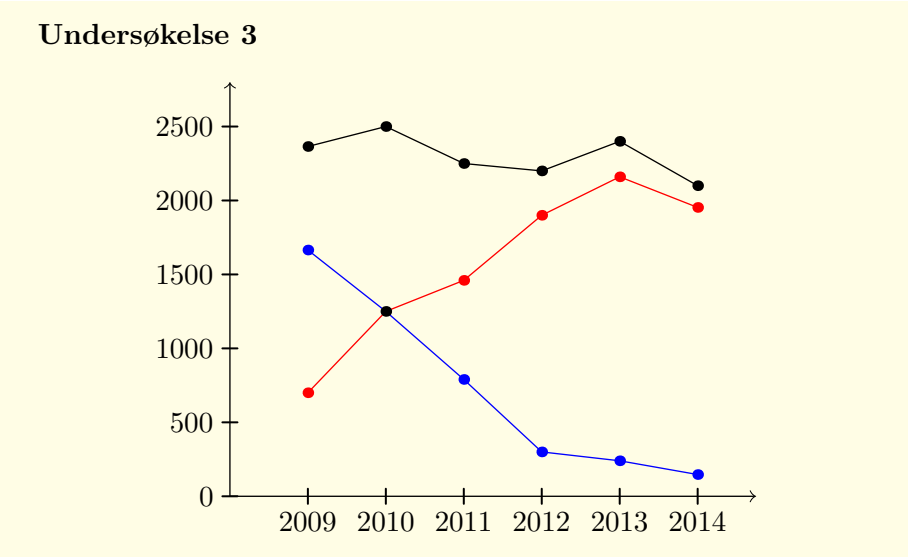
Spørreundersøkelse 1



Spørreundersøkelse 2



Linjediagram



0.2 Sentralmål

I datasett vil det ofte være svar som er helt eller veldig like, og som gjentar seg. Dette betyr at vi kan si noe om hva som gjelder for mange. De matematiske begrepene som forteller noe om dette kalles *sentralmål*.

Typetall

0.1 Typetall

Typetallet er svaret som opptrer flest ganger i datasettet.

Undersøkelse 1

I datasettet er det tallet 4 som opptrer flest (tre) ganger. Dette kan vi se både fra selve datasett på s ?? eller frekvenstabellen på s ??, søylediagrammet på s ?? eller sektordiagrammet ??.

4 er altså typetallet.

Gjennomsnitt

Når et datasett består av svar i form av tall kan vi finne summen av svarene. Når vi spør oss hva *gjennomsnittet* er, spør vi om dette:

”Hvis alle svarene var like, men summen den samme, hvilken verdi måtte alle svarene da ha hatt?”

Dette er jo ingenting annet enn divisjon (se [MB](#), s. 23):

0.2 Gjennomsnitt

$$\text{gjennomsnitt} = \frac{\text{summen av svarene fra datasettet}}{\text{antall svar}}$$

Datasett 2

Vi summerer svarene fra datasettet, og deler med antall svar:

$$\begin{aligned}\text{gjennomsnitt} &= \frac{47 + 124 + 61 + 38 + 97 + 84 + 101 + 79 + 56 + 40}{10} \\ &= \frac{727}{10} \\ &= 72,7\end{aligned}$$

Altså, i gjennomsnitt holdt de 10 deltakerne pusten i 72,7 sekunder.

Datasett 1

Metode 1

$$\begin{aligned}\text{gjennomsnitt} &= \frac{7 + 4 + 5 + 4 + 1 + 0 + 6 + 5 + 4 + 8 + 1 + 6 + 8 + 0 + 14}{15} \\ &= \frac{73}{15} \\ &\approx 4.87\end{aligned}$$

Metode 2

Vi utvider frekvenstabellen fra side ?? for å finne summen av svarene fra datasettet (vi har også tatt med summen av frekvensene):

Antall epler	Frekvens	antall · frekvens
0	2	$0 \cdot 2 = 0$
1	2	$1 \cdot 2 = 2$
4	3	$4 \cdot 3 = 12$
5	2	$5 \cdot 2 = 10$
6	2	$6 \cdot 2 = 12$
7	1	$7 \cdot 1 = 7$
8	1	$8 \cdot 1 = 8$
14	1	$14 \cdot 1 = 14$
sum	15	73

Nå har vi at

$$\begin{aligned}\text{gjennomsnitt} &= \frac{73}{15} \\ &\approx 4,87\end{aligned}$$

Altså, i gjennomsnitt spiser de 15 respondentene 4,87 epler i uka.

Median

0.3 Median

Medianen er tallet som ender opp i midten av datasettet når det rangeres fra svar med lavest til høyest verdi.

Hvis datasettet har partalls antall svar, er medianen gjennomsnittet av de to svarene i midten (etter rangering).

Undersøkelse 1

Vi rangerer datasettet fra lavest til høyest svar:

0 0 1 1 4 4 4 5 5 6 6 7 8 8 14

Tallet i midten er 5, altså er medianen 5.

Undersøkelse 2

Vi rangerer datasettet fra lavest til høyest svar:

38 40 47 56 61 79 84 97 101 124

De to tallene i midten er 61 og 79. Gjennomsnittet av disse er

$$\frac{61 + 79}{2} = 70$$

Altså er medianen 70.