

STK1000 :

Les siste blogg innlegg på [www, statistikk.no](http://www.statistikk.no) som forberedelse til forelesningen.

Mandag 28/8

Fortsatt fra sist: ulike typer data og hvordan vi oppsummerer dem

Boka: Ch 1.1-1.3

R & RStudio: Appelsindataoppgave

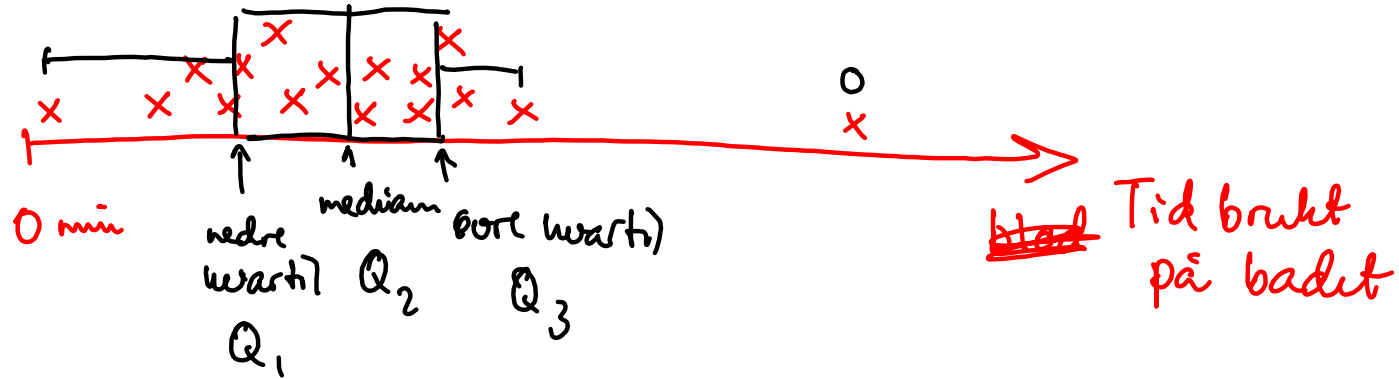
---

Smu deg til sidemargen og forklar hva som menes med kategoriske data og numeriske data, diskret og kontinuerlige data.

Tabeller med antall (%)  
Søylediagram / kakediagram

histogram, boksploTT

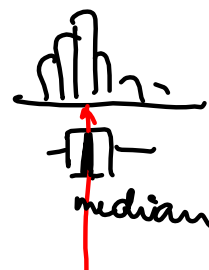
Lag 16 kryss, og gi arket til sidemannen som skal  
 løse et boksplokk :



Kontinuerlige data:  
skjevt eller symmetrisk?



SKJEVT



gjennomsnitt

SYMM

mål eller oppsummeringstall for det typiske

✓ median

~~gjennomsnitt~~

✓ median

✓ gjennomsnitt

og for varianjer/spredning

✓ min, maks  
maks - min = range

✓ Q<sub>1</sub>, Q<sub>3</sub>

Q<sub>3</sub> - Q<sub>1</sub> = IQR

✓ min, maks  
range

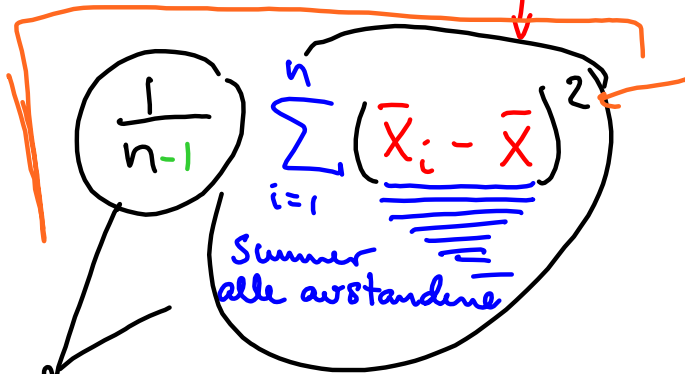
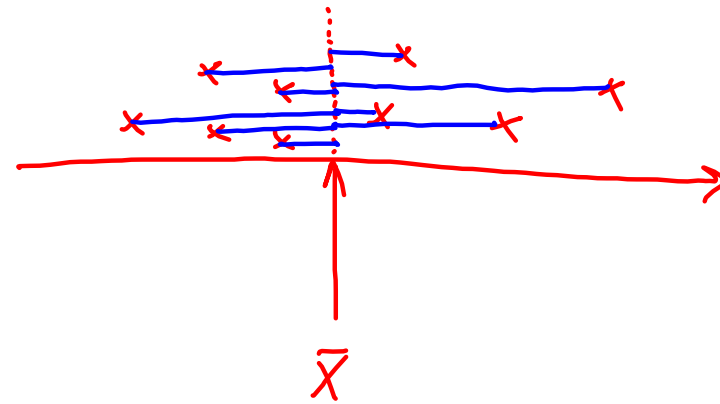
✓ Q<sub>1</sub>, Q<sub>3</sub>

Q<sub>3</sub> - Q<sub>1</sub> = IQR

~~SD = "gjennomsnittlig avvik fra g.<sub>1</sub> av"~~

SD = "gjennomsnittlig avvik fra gjennomsnittet"

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$



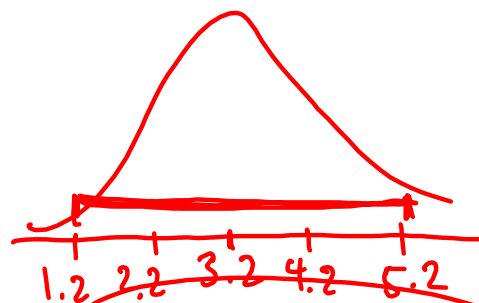
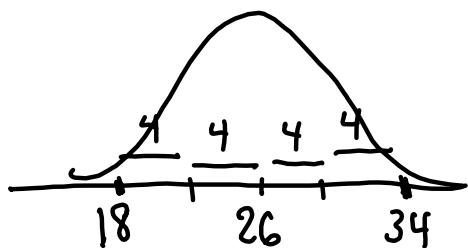
nesten gjennomsnittlig kvadrert avvik fra gjennomsnittet.

Detta bør dere tenke om SD:

SD er en måte å oppsummere variasjon på i symmetriske data uten ekstremverdier

Hvis vi tar  $\bar{x} \pm 2 \cdot SD$ , vil det inneholde ca 95% av alle målingene (Forklaring kommer senere)

Noen sier til meg at i deres data er gj. om BM  $\pm$  på 26, og  $SD=4$ . Hva tenker jeg?



Jeg sier: Tykkelse på berglag er 3.2 m i gj. om.  $SD=1m$

Tegn fordelingen sammen med sidemannen.

↓  
 $\approx$  Symmetrisk fordeling

# FORSKNINGS PROSESSEN (Ch 3.1 - 3.3)

Bakgrunn i et fagfelt

↓  
Problemstilling (Bachelor, Master, PhD)

↓  
Plan for å bevare PS  
Forsknings design

Experimentelle design  
Lab  
Intervensjon

Observasjonelle design