

Beskrivende statistikk: Datatyper og visualisering

**Thea Bjørnland
Institutt for matematiske fag
NTNU**

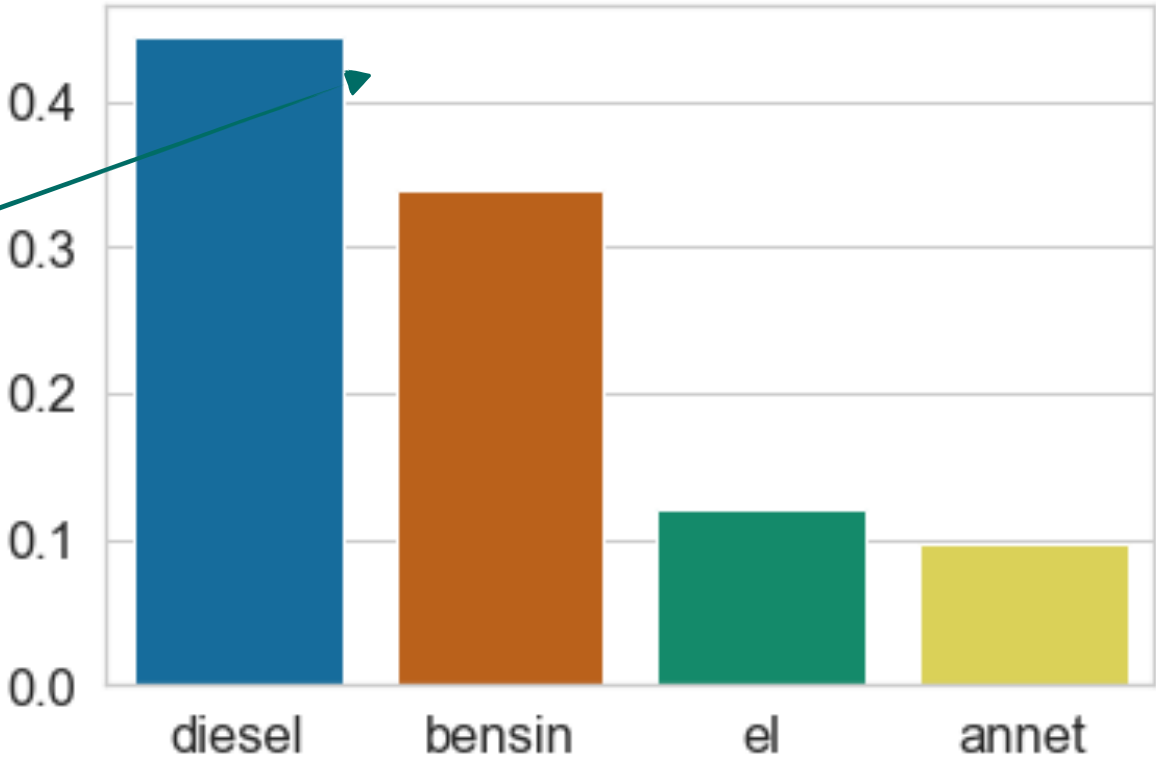
Kategoriske data

Observasjoner som kan deles inn i *kategorier*

Tabell: Personbiler 2020

Drivstoff	Antall biler	Andel (%)
Bensin	950 131	33,8 %
Diesel	1 246 671	44,4 %
El	340 002	12,1 %
Annet	272 314	9,7 %

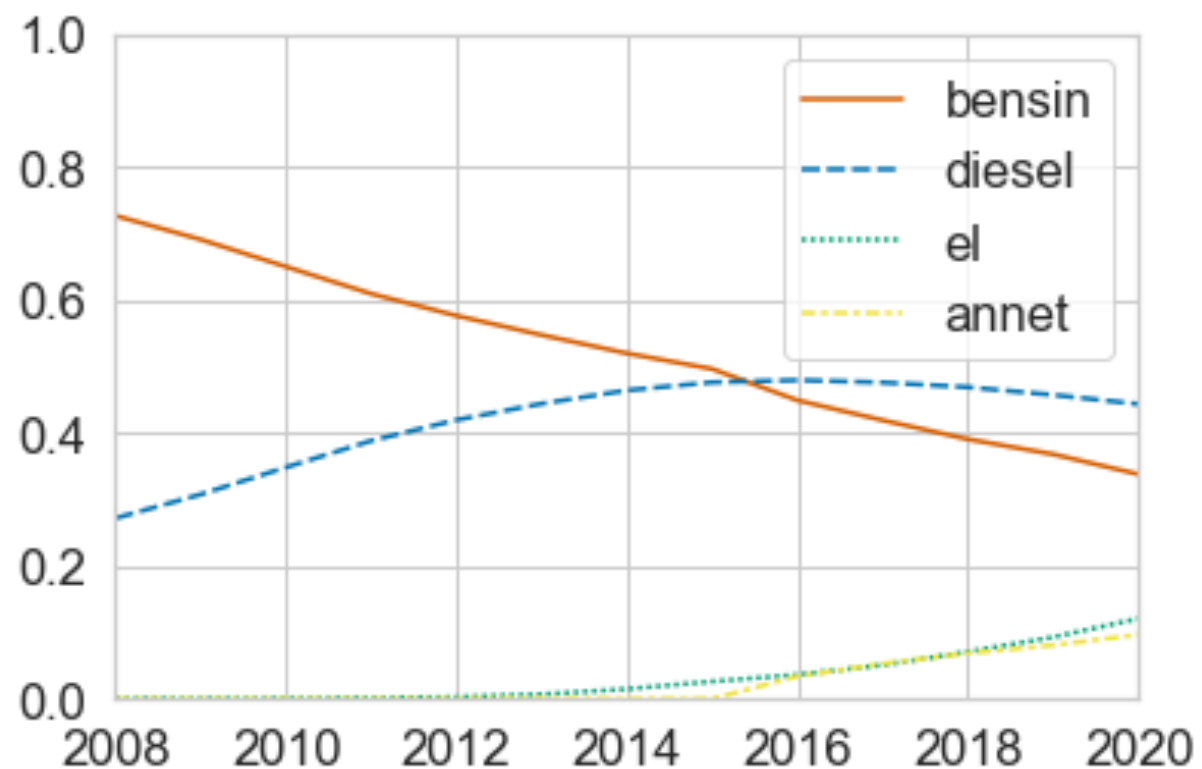
Figur: Stolpediagram for drivstofftype i 2020



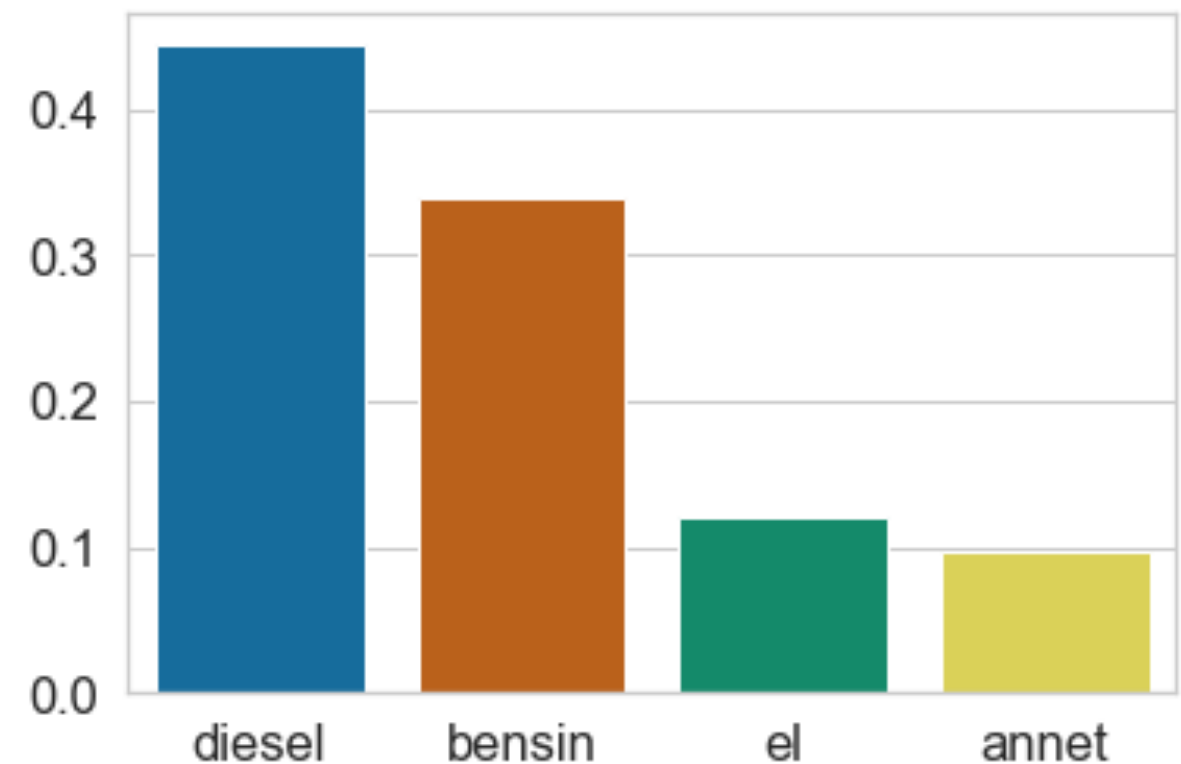
Kilde: <https://www.ssb.no/statbank/list/>

Kategoriske data: utvikling over tid

Figur: Utvikling av andel biler over tid



Figur: Stolpediagram for drivstofftype i 2020



Diskrete data

Observasjoner som kan ta *diskrete tallverdier*, f.eks 0, 1, 2, osv.

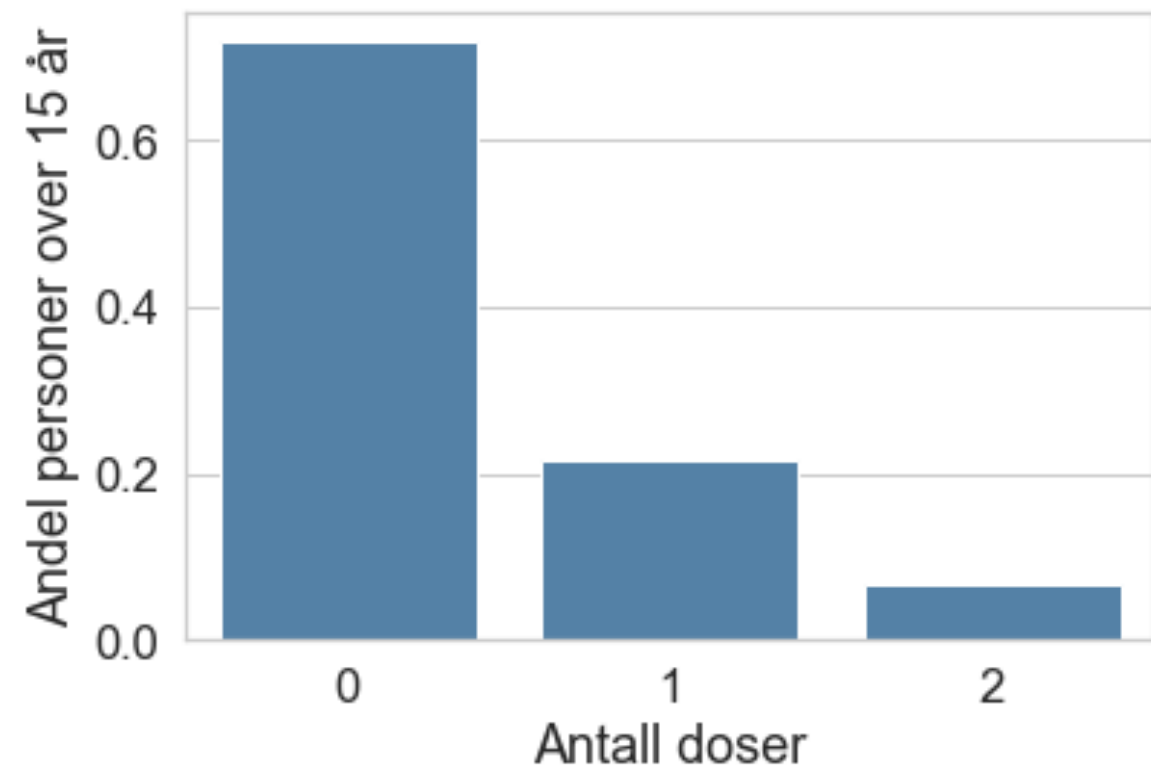
Tabell: Vaksinedoser per 14. april 2021

Antall doser	Antall personer over 15 år	Andel
0	3 207 216 *	0.72
1	961 638	0.21
2	298 857	0.07

Kilder: <https://statistikk.fhi.no/sysvak>

* Befolkningstall (over 15 år) per 1. januar 2021 fra SSB:
<https://www.ssb.no/befolkning/statistikker/folkemengde/aar-per-1-januar>

Figur: Stolpediagram over andel vaksinedoser



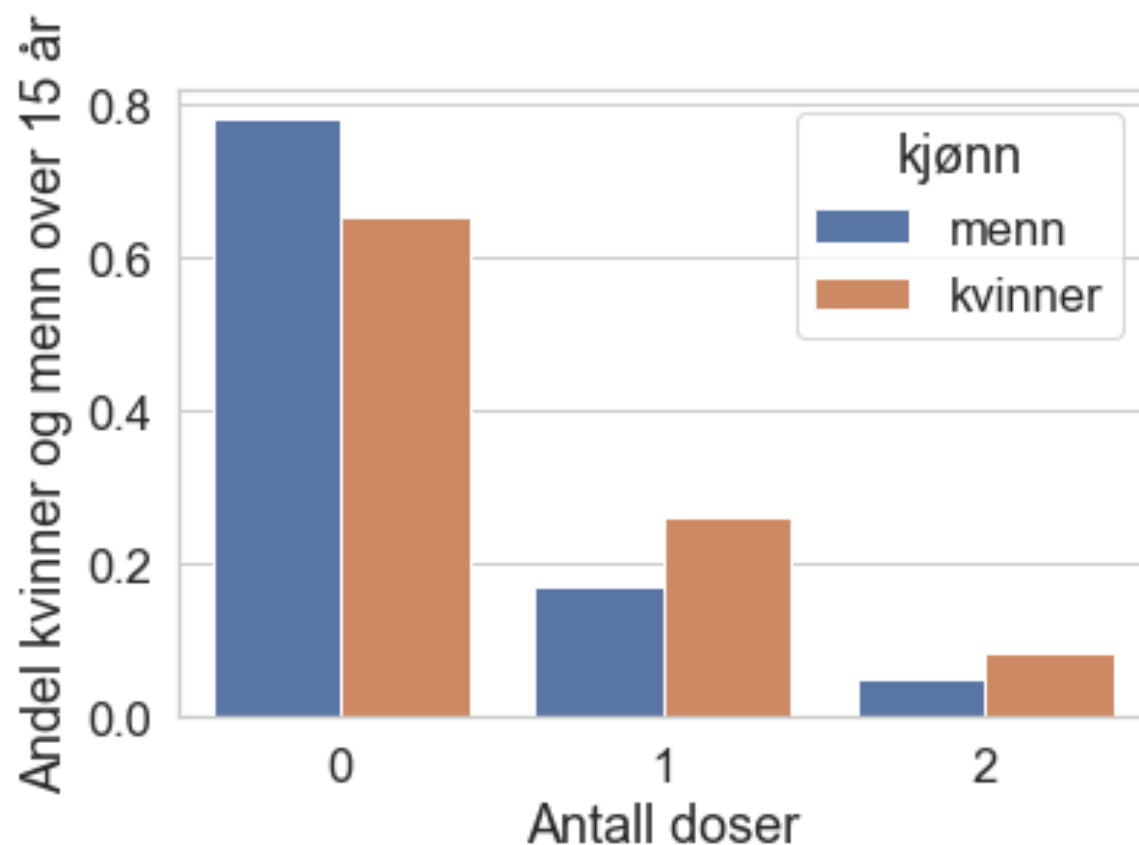
Median = 0 vaksinedoser

Gjennomsnitt $\bar{x} = 0,35$ doser

Diskrete data: delt inn i kategorier

Observasjoner som kan ta *diskrete tallverdier*, f.eks 0, 1, 2, osv.

Figur: Stolpediagram over andel vaksinedoser for menn og kvinner innen 14. april 2021



Kilder: <https://statistikk.fhi.no/sysvak>

<https://www.ssb.no/befolkning/statistikker/folkemengde/aar-per-1-januar>

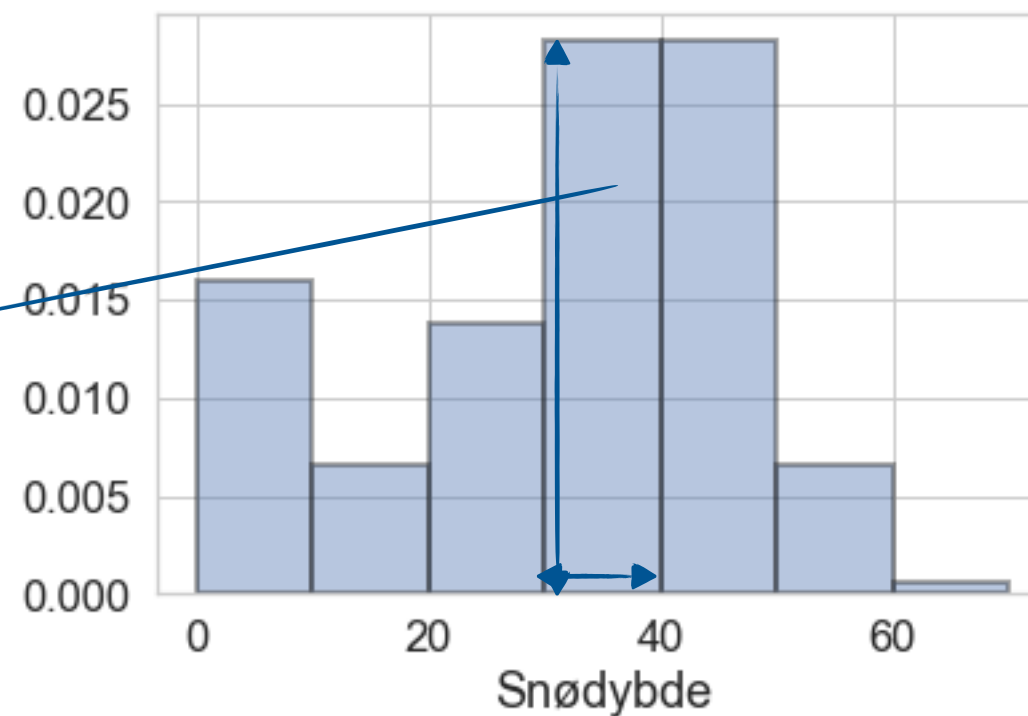
Kontinuerlige data

Observasjoner som kan gjøres på en *kontinuerlig skala*

Tabell: Daglig målt snødybde i Biri, Gjøvik, 1. november 2019 til 31. april 2020

Snødybde	Antall dager	Andel
[0, 10)	29	0,16
[10, 20)	12	0,07
[20, 30)	25	0,14
[30, 40)	51	0,28
[40, 50)	51	0,28
[50, 60)	12	0,07
[60, 70)	1	0,01

Figur: Histogram over snødybdemålinger



Dataene er lastet ned fra
seklima.met.no

Kontinuerlige data

Observasjoner som kan gjøres på en *kontinuerlig skala*

Tabell: Daglig målt snødybde i Biri, Gjøvik, 1. november 2019 til 31. april 2020

Snødybde	Antall dager	Andel
[0, 10)	29	0,16
[10, 20)	12	0,07
[20, 30)	25	0,14
[30, 40)	51	0,28
[40, 50)	51	0,28
[50, 60)	12	0,07
[60, 70)	1	0,01

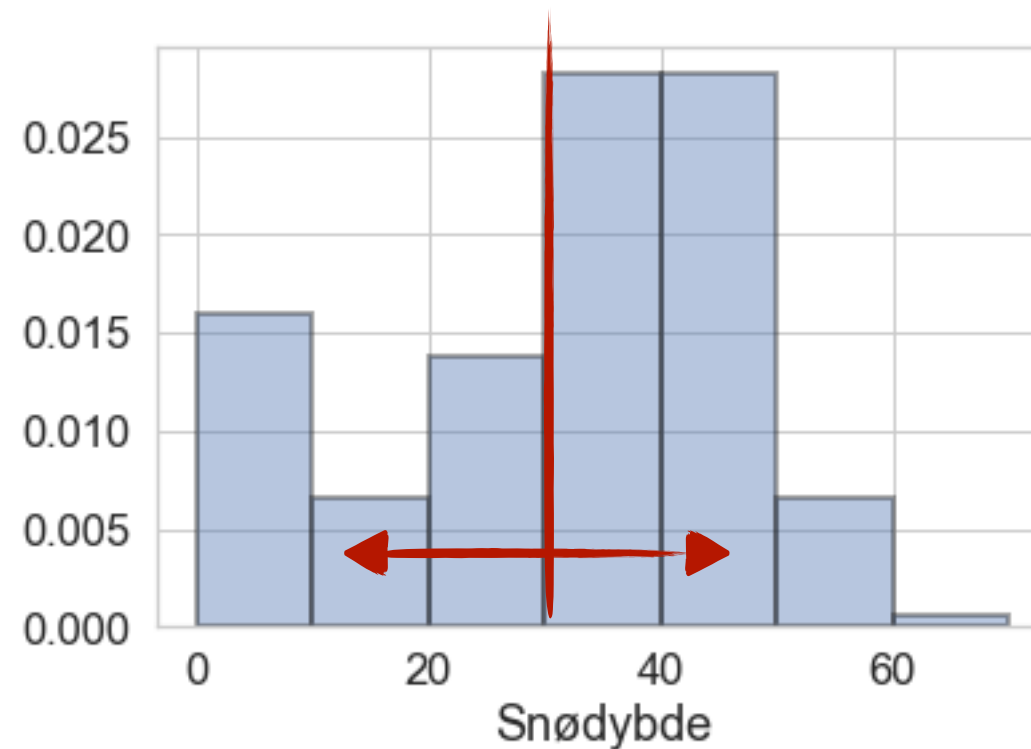
Dataene er lastet ned fra
seklima.met.no



Median = 34,0 cm

Kvartilbredde = 25,0 cm

Figur: Histogram over snødybdemålinger



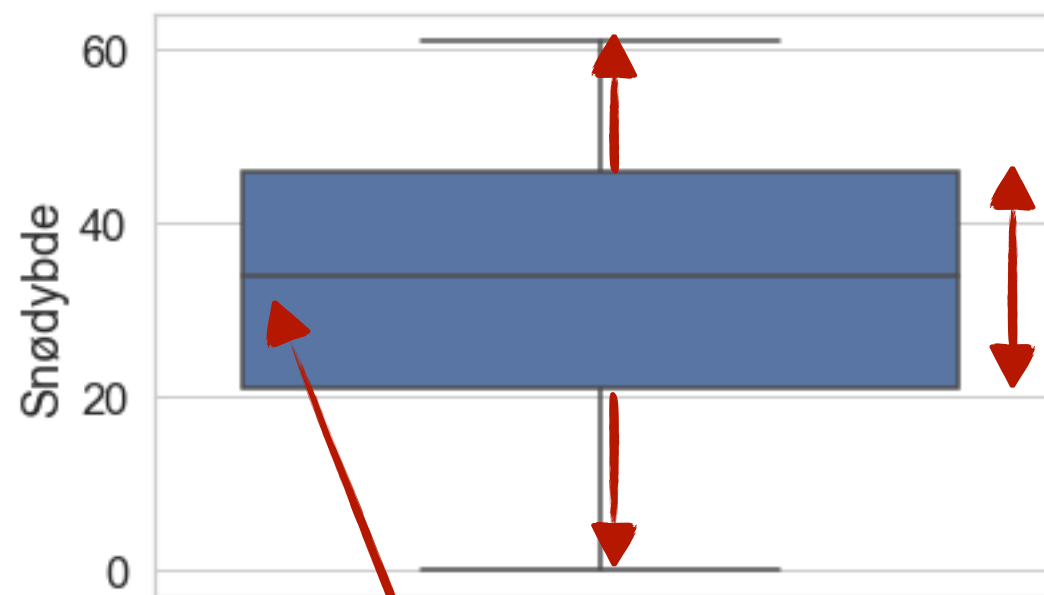
Gjennomsnitt $\bar{x} = 31,3$ cm

Standardavvik $s = 16,4$ cm

Kontinuerlige data

Observasjoner som kan gjøres på en *kontinuerlig skala*

Figur: Boksplott over snødybdemålinger



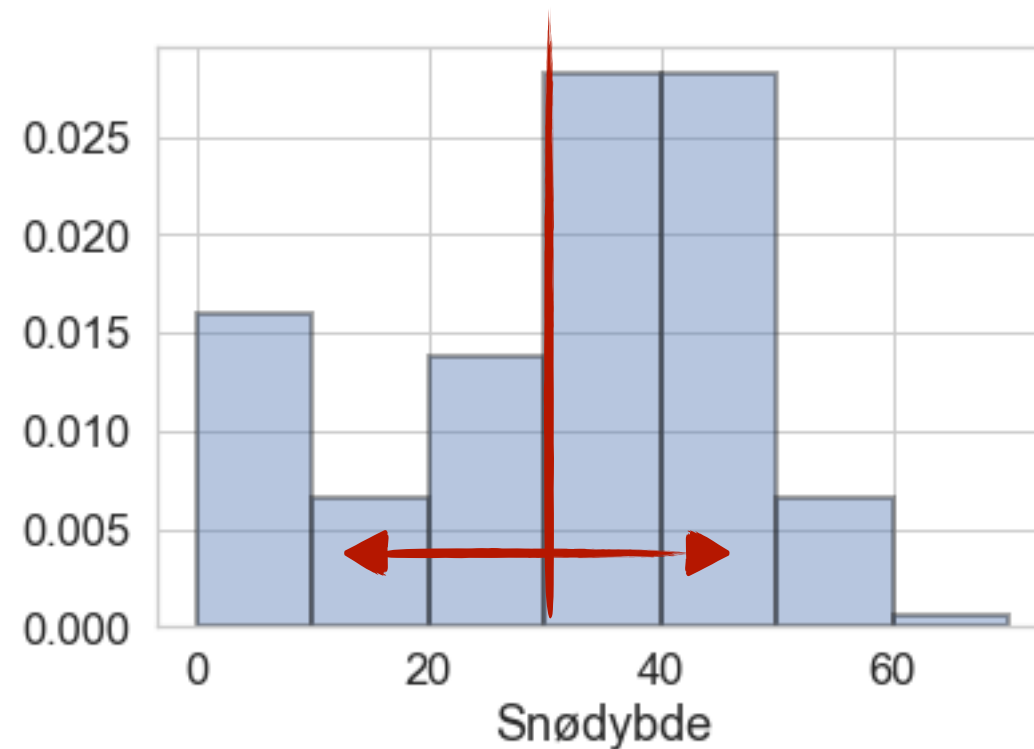
Dataene er lastet ned fra
seklima.met.no



Median = 34,0 cm

Kvartilbredde = 25,0 cm

Figur: Histogram over snødybdemålinger



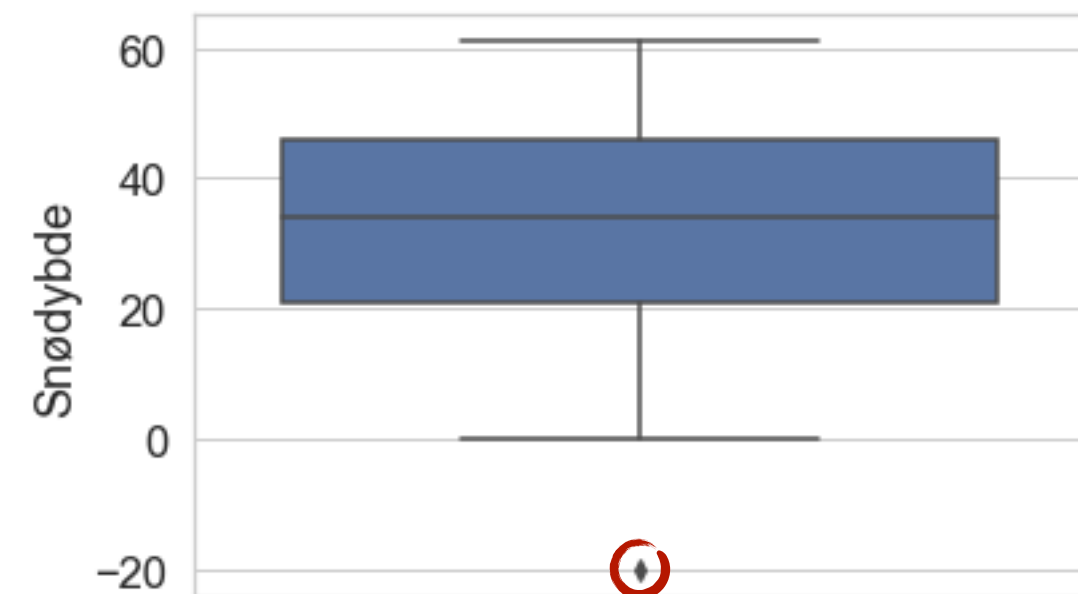
Gjennomsnitt $\bar{x} = 31,3$ cm

Standardavvik $s = 16,4$ cm

Kontinuerlige data: Uteliggere i boksplott

Observasjoner som kan gjøres på en *kontinuerlig skala*

Figur: Boksplott over snødybdemålinger



Dataene er lastet ned fra
seklima.met.no



Husk!

- “uteliggere” må aldri fjernes bare fordi de er uteliggere!
- Vi må aldri “pynte på” et datasett dersom vi ikke er helt sikre på at vi har en feil eller ugyldig observasjon