

Velkommen!

Statistikk for ingeniører

Digital plenumstime UKE 9

Vi starter klokka 15:15

Skriv til oss i chat om du har spørsmål underveis

Det blir gjort opptak av denne timen :)

Thea Bjørnland, Institutt for matematiske fag, NTNU (Trondheim)

Agenda

Praktisk informasjon om tellende prosjektoppgave

Forberedende prosjekt:

- Gjennomgang av diskusjonsoppgaven

- og typiske 'feil' som er relevante for prosjektoppgaven

Litt om lineær regresjon

Tellende prosjektoppgave i statistikk

Prosjektoppgaven er en egen delvurdering og teller 30% av endelig karakter
(individuell eksamen teller 70%)

Karakter prosjekt og eksamen  

	E eksamen	D eksamen	C eksamen	B eksamen	A eksamen
E prosjekt	E	D	D	C	B
D prosjekt	E	D	C	C	B
C prosjekt	D	D	C	B	B
B prosjekt	D	C	C	B	A
A prosjekt	D	C	B	B	A

Det er nok mest vanlig med et prosjekt som ikke endrer eksamenskarakter, eller at prosjektet trekker opp

Tellende prosjektoppgave i statistikk

Det er 3-6 studenter i hver gruppe. Gruppene dere skal levere prosjektet i er de samme gruppene som ble etablert i forprosjektet

Prosjektet leveres gruppevis i **Inspera**

Frist for innlevering er 18. november klokka 12:00 i Inspera.

Det kan *ikke* gis utsettelse på innleveringsfristen.

Alle studenter i samme gruppe får samme bokstavkarakter.

Merk NTNUs regelverk for klage på karakter:

Ved klage på karakterfastsettingen av gruppearbeid, der det gis en felles karakter, klager du individuelt. En eventuell endring etter klage vil kun gjeld for den/de gruppemedlemmene som har klaget.

Tellende prosjektoppgave i statistikk

Prosjektmoduler 

Informasjon om
gruppearbeid

Prosjektgru
påmelding

Forprosjek

← Åpnes på fredag for studenter



Prosjektmoduler i statistikk

Her kan du klikke deg inn i innholdsmappa for din prosjektmodul; ISTx1001, ISTx1002 eller ISTx1003.

Ser du ikke riktig innholdsmappe her? Ta kontakt med thea.bjornland@ntnu.no.



ISTx1001 Industriell statistikk

Tilgjengelighet: Elementet er tilgjengelig, men det kan hende noen studenter eller grupper ikke har tilgang til det.
Aktivert: Tilpasset visning



ISTx1002 Usikkerhet og støy i målinger

Tilgjengelighet: Elementet er tilgjengelig, men det kan hende noen studenter eller grupper ikke har tilgang til det.
Aktivert: Tilpasset visning



ISTx1003 Statistisk læring og data science

Tilgjengelighet: Elementet er tilgjengelig, men det kan hende noen studenter eller grupper ikke har tilgang til det.

Du ser bare den modulen
som du hører til i

Tellende prosjektoppgave i statistikk

ISTx1001

Faglærer:
Thea Bjørnland

ISTx1002

Faglærer:
Bjørn Olav Hogstad

ISTx1003

Faglærer:
Sara Martino

Alle moduler: Digital undervisning over zoom

NB: Nye zoom-lenker

Veiledning i øvingstimer, fysisk og/eller digitalt (zoom og forum),
med en del lokale forskjeller

-> Følg med på sidene for prosjektet OG campus-sidene i Blackboard

Første undervisningstime mandag 21. oktober klokka 15:15 - 16:00

Prosjektperioden varer i 4 uker

I alle 3 prosjektene kan dere starte arbeidet den første uka

Typiske feil fra forberedende prosjekt

En god del grupper leverte ikke alle filer som ble bedt om.

Dette er ok for forprosjekt, men dumt å gjøre slike feil i prosjektet

En god del studenter / grupper fulgte ikke med på frister osv.

Husk ingen utsettelse av frist på tellende prosjektarbeid

Innlevering i Inspira 18. november klokka 12:00 er endelig ,
akkurat som ved en eksamen

Dersom det er noe i oppgaveteksten dere synes er uklart:

Oppsøk veiledning og/eller bruk forumet!

Diskusjonsoppgave i forprosjekt

Gruppediskusjon om et stokastisk forsøk

Tenk dere dette stokastiske forsøket: En person velges tilfeldig fra NTNUs studenter og ansatte. Vi registrerer skostørrelse og høyde (i cm). Til dette stokastiske forsøket kan vi tilegne to stokastiske variabler: X for skostørrelse og Y for høyde.

Hvilke egenskaper tror dere de stokastiske variablene X og Y vil ha? Diskuter verdimenge, diskret/kontinuerlig, forventningsverdi og standardavvik, kovarians og formen på sannsynlighetsfordelingene til X og Y. Skriv (kortfattet) ned deres tanker, gjerne med en håndtegnet skisse av fordelingene.

Verdimengde

$X: 35, 36, 37, \dots, 45 ?$
 $37 \approx 1/3 \quad 40,5 ?$

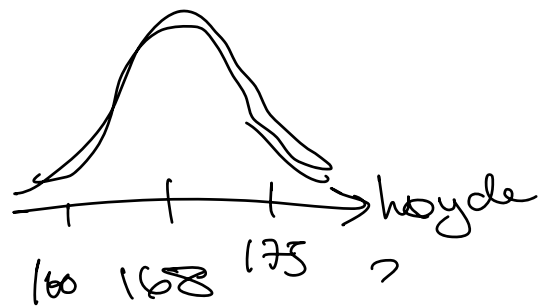
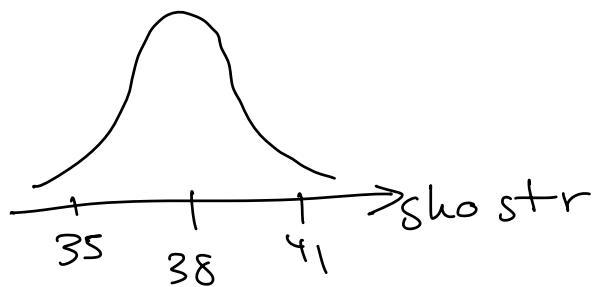
$Y: \text{høyde}$

kontinuerlig $(0, \infty)$
 $(150, 210) ?$

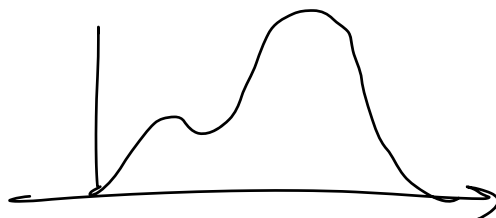
↳ kan modellere X
som kontinuerlig
↳ diskret også OK.

Sannsynlighetsfordelinger:

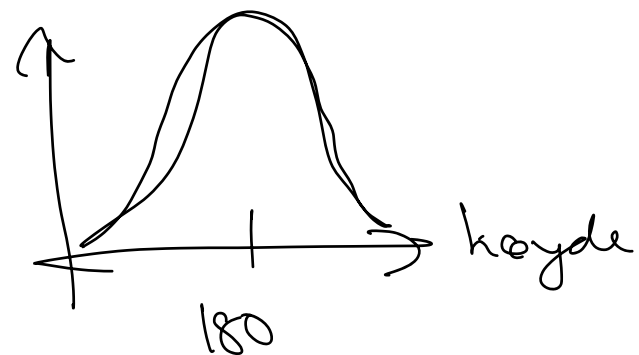
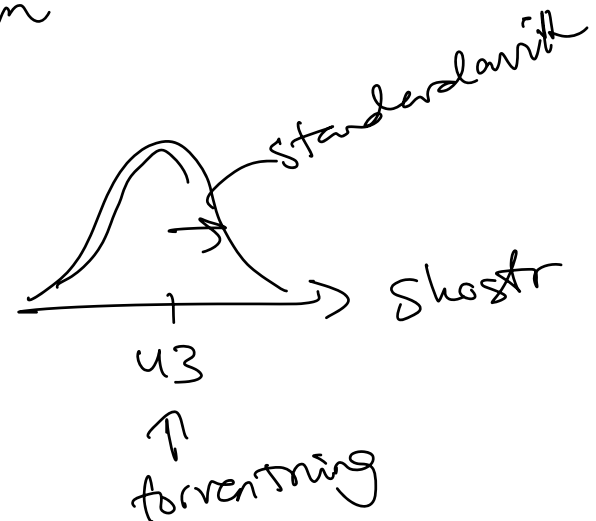
kvinne



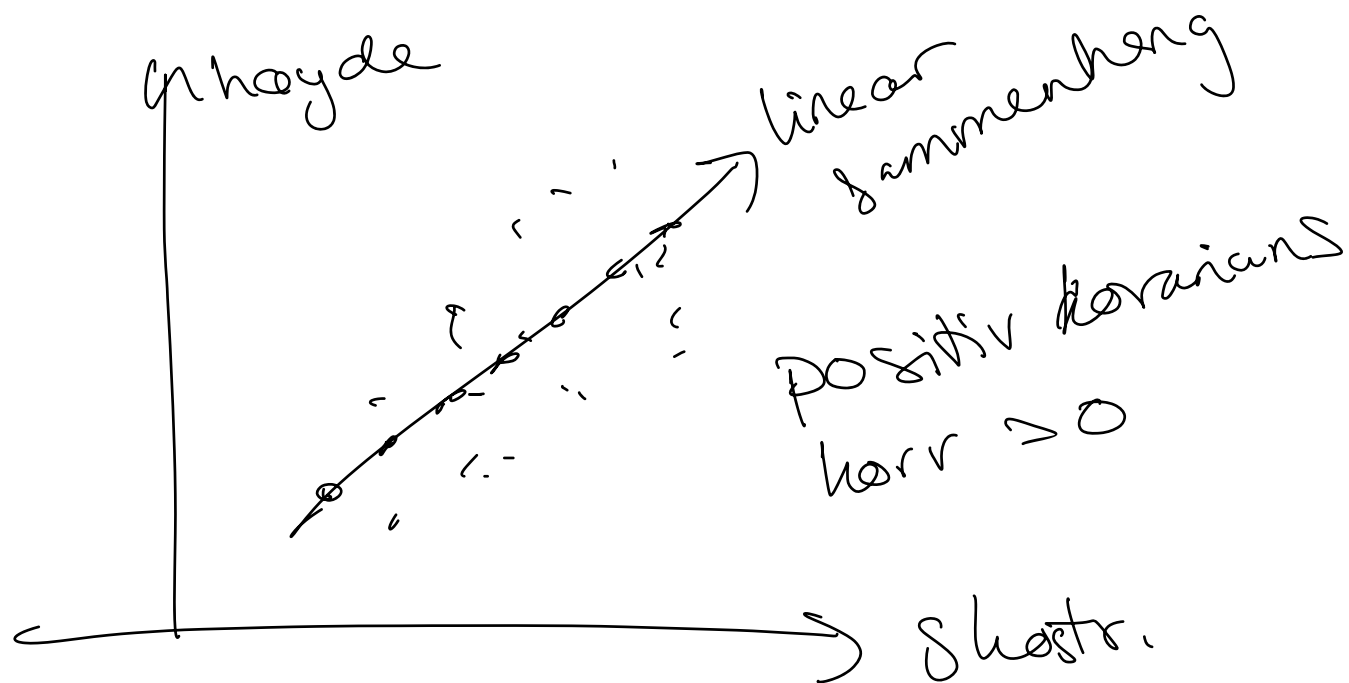
Ingeniører :

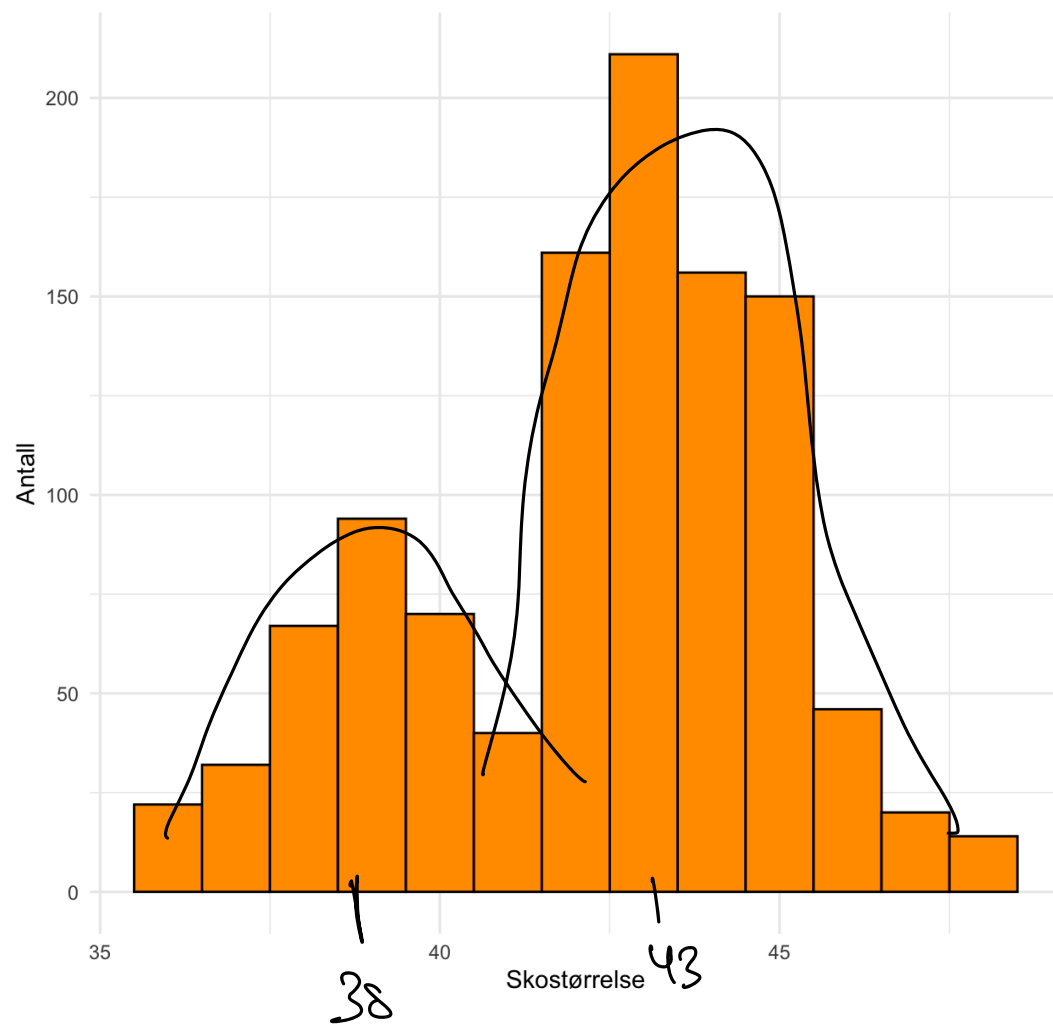


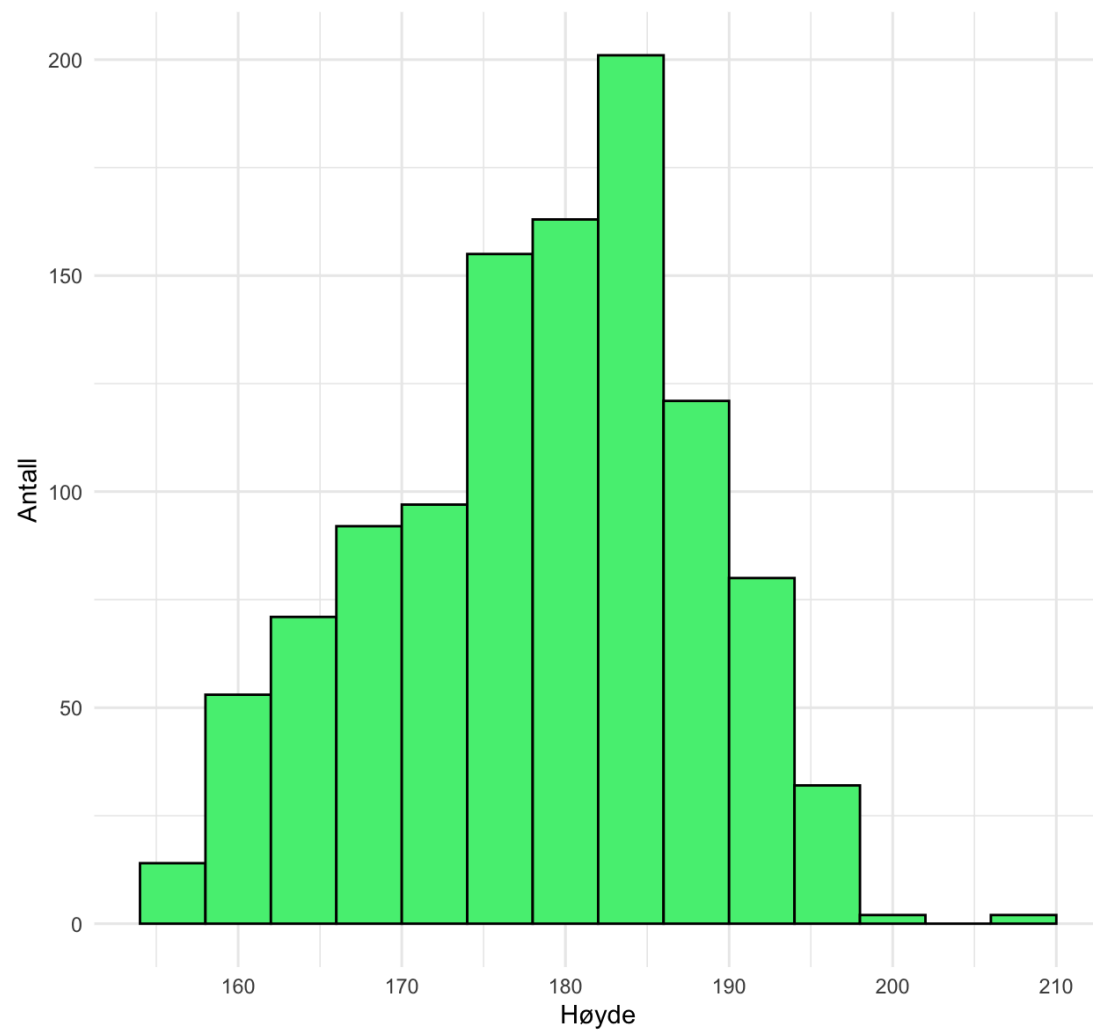
Mann

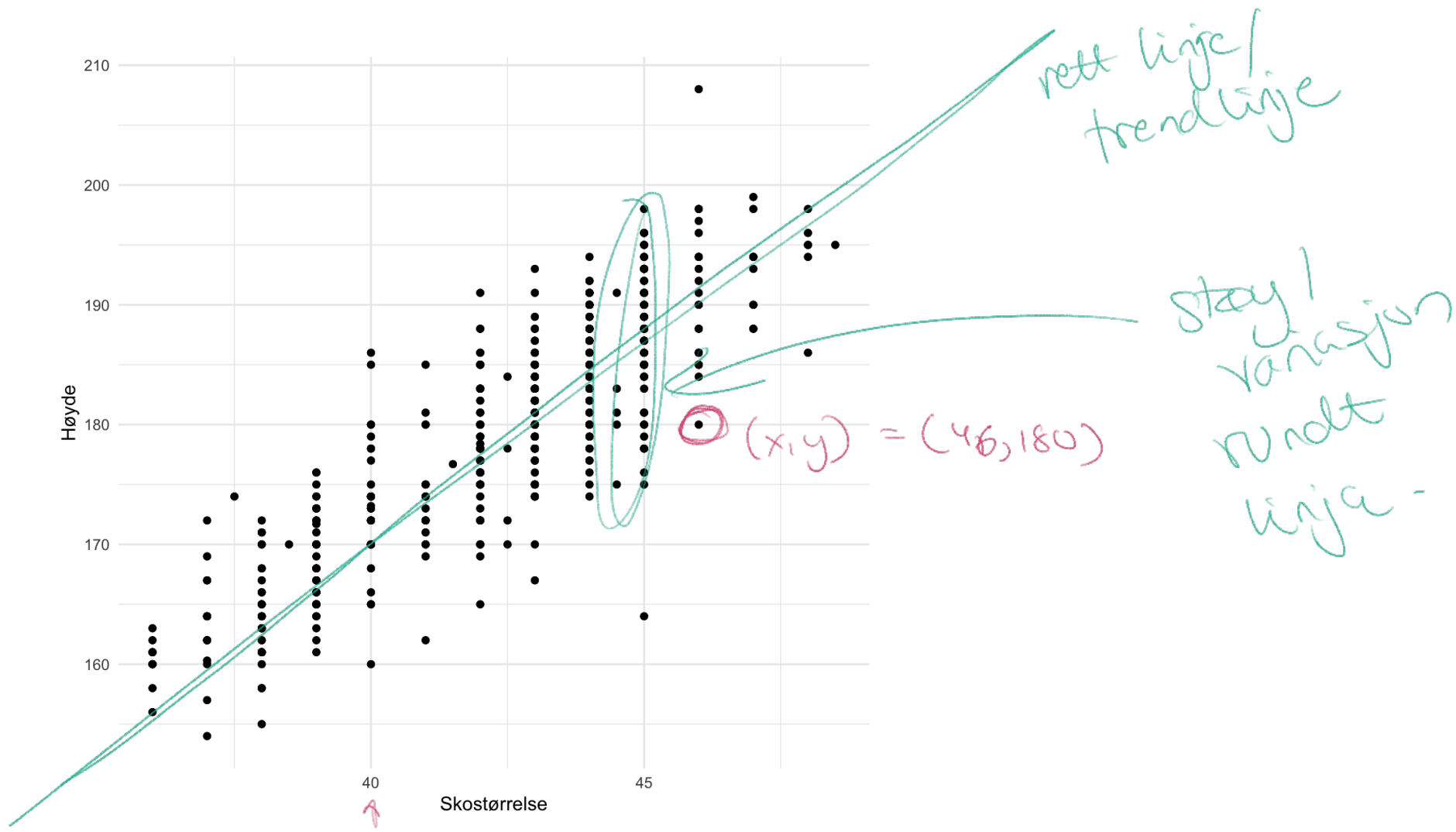


Kovarians? Korrelasjon

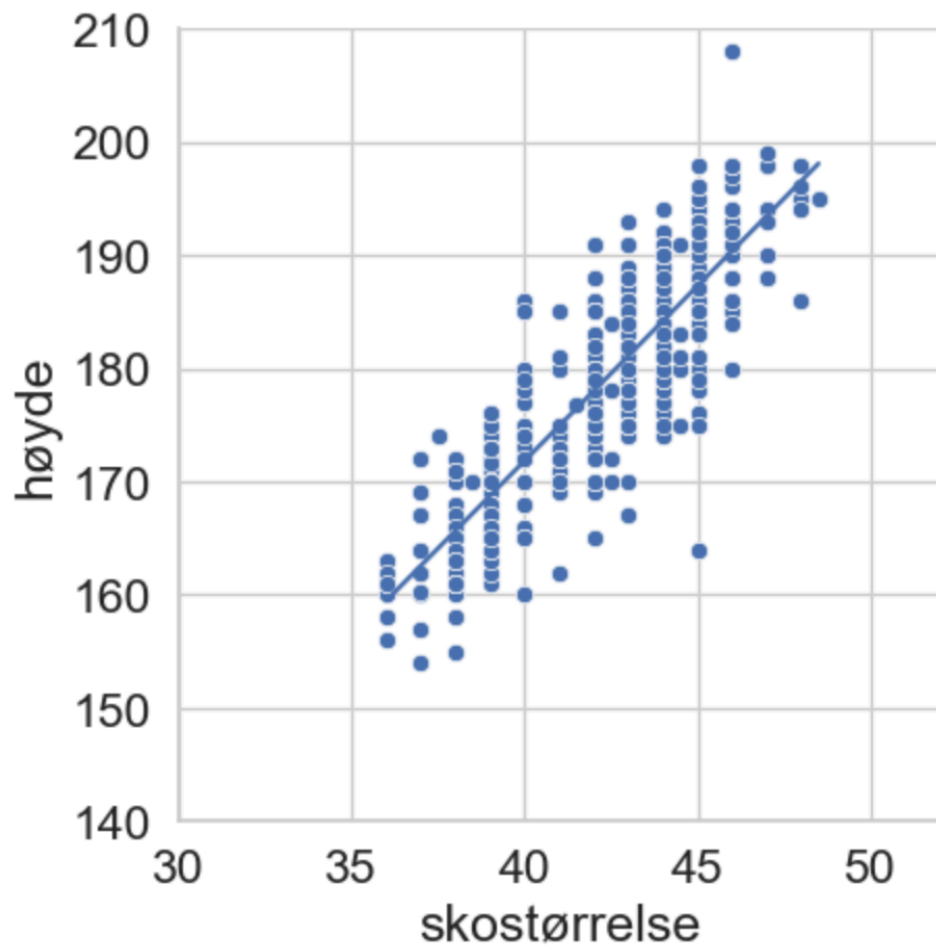








Korrelasjon = 0.86



$y \sim x$

```
modell = smf.ols('høyde~skostørrelse', data=df).fit()
modell.summary()
```

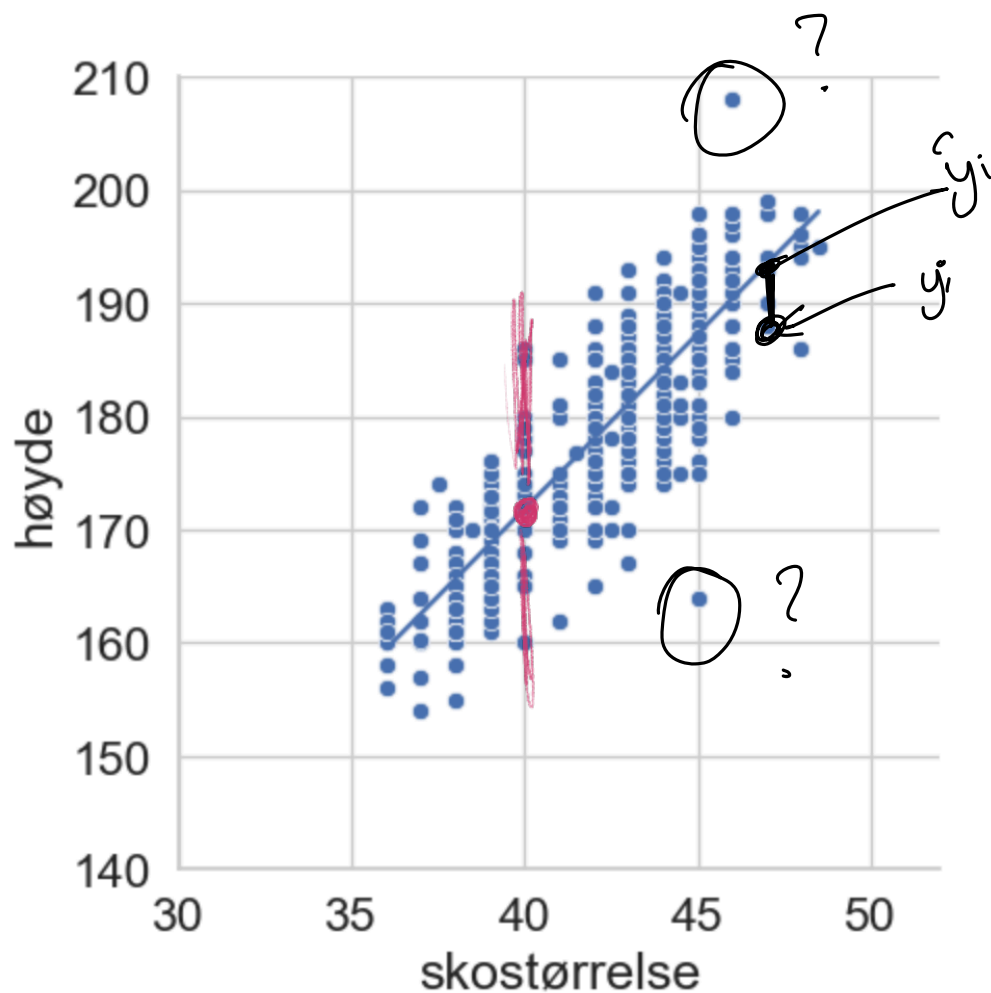
OLS Regression Results

R^2

Dep. Variable:	høyde	R-squared:	0.746
Model:	OLS	Adj. R-squared:	0.746
Method:	Least Squares	F-statistic:	3180.
Date:	Mon, 14 Oct 2024	Prob (F-statistic):	4.94e-324
Time:	09:07:54	Log-Likelihood:	-3233.8
No. Observations:	1083	AIC:	6472.
Df Residuals:	1081	BIC:	6481.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	48.3120	2.321	20.812	0.000	43.757	52.867
skostørrelse	3.0882	0.055	56.392	0.000	2.981	3.196

Omnibus:	27.196	Durbin-Watson:	1.860
Prob(Omnibus):	0.000	Jarque-Bera (JB):	51.188
Skew:	-0.145	Prob(JB):	7.67e-12
Kurtosis:	4.025	Cond. No.	676.



	coef	std err	t	P> t	[0.025	0.975]
Intercept	48.3120	2.321	20.812	0.000	43.757	52.867
skostørrelse	3.0882	0.055	56.392	0.000	2.981	3.196

`np.sqrt(modell.mse_resid)`

4.787299286887585

$\hat{\beta}_0$

$\hat{\beta}_1$

σ

$$y = 48,31 + 3,09x$$

$$\sigma \approx 4,79 \text{ cm}$$

Lineær regresjon

← høyde

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

← skostørrelse

↑ 40

← støy, feilteori

$$\varepsilon \sim N(0, \sigma)$$

↑ summe overalt

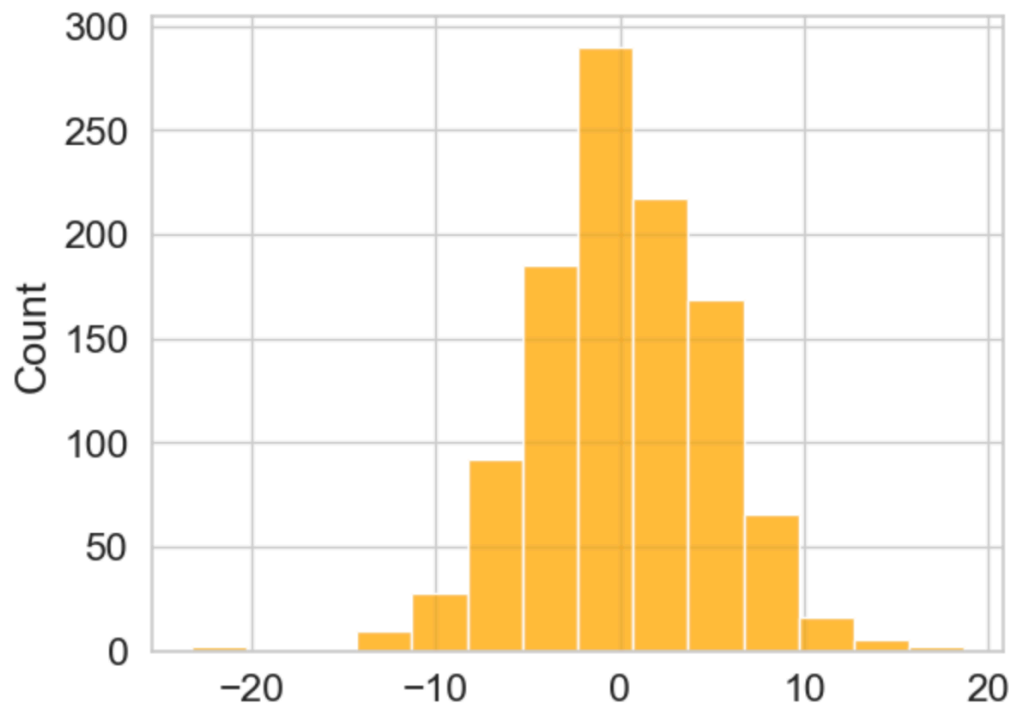
Gult opppe:

$$SD(\hat{\beta}_0) \text{ og}$$

$$SD(\hat{\beta}_1)$$

↑ se formelark.

```
sns.histplot(modell.resid, binwidth = 3, color = 'orange')
plt.show()
```



↑
Ser fint og
normalfordelt
ut!

Residual:

$$\hat{\epsilon}_i = \epsilon_i = y_i - \hat{y}_i$$

$n = 1083$ (?)
realisasjoner
av $\epsilon \sim N(0, \sigma)$
↑ antagelse

QQ-plot

