

# Forprosjekt

## Formål

- Bli kjent med gruppene deres
- Øve dere på praktiske ferdigheter
- Reflektere rundt egenskaper til stokastiske variabler

## Leveranser

Forprosjektet består av tre obligatoriske leveranser:

1. Gruppekонтракт (se forslag til mal)
2. Maks 1 sides diskusjon om det stokastiske forsøket
3. En excel- eller csv-fil med gruppas datasett
4. En PDF-fil som viser et kryssplott og regresjonslinje fra dataanalysen i Python

## 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.

## Innsamling, registrering og behandling av data

Bruk notatboken i JupyterHuben og ta utgangspunkt i Excel-fila som undervisningsteamet har startet på. Legg til skostørrelsen og høyden til hvert av deres gruppemedlemmer. Skriv Python-kode for å

- lese inn Excel-fila
- finne en rett linje (regresjonslinje)  $y = a + bx$  der  $x$  er skostørrelse og  $y$  er høyde
- plotte dataene og regresjonslinja og lagre figuren som PDF

Vi anbefaler å bruke notatbøker fra uke 1 for å finne Python-kode for å gjennomføre data-analysen. Men dere kan også bruke Google-søk eller prøve dere med ChatGPT, Bard, Bing, etc for å få forslag til egnet Python-kode. Dere skal ikke levere jupyter-notatboka til oss, kun resultatet av analysen deres.