

Korrelasjon er et uttrykk for lineær samvariasjon mellom to variabler. Det er flere måter å beregne korrelasjon på, og den vi bruker heter Pearson korrelasjon. Den er alltid mellom -1 og 1. En korrelasjon på 1 kaller vi perfekt positiv, og korrelasjon på -1 er perfekt negativ. Hvis korrelasjonen er 0, sier vi at det ikke er noen (lineær) samvariasjon. Men det betyr ikke at det ikke finnes en samvariasjon mellom to variabler – det betyr bare at vi kan ikke måle noen lineær samvariasjon. Det at vi finner en korrelasjon nærme 1 eller -1, behøver heller ikke bety at de to variablene henger sammen eller påvirker hverandre. Samvariasjonen kan være tilfeldig.

Dette er korrelasjonen mellom Alder og Alder. Den er perfekt positiv. Alle variabler har perfekt positiv samvariasjon med seg selv

Dette er det nøyaktige **signifikansnivået** – dvs. hvor sikre kan vi være på å ikke ta feil dersom vi påstår at det er en samvariasjon (dvs. korrelasjon) mellom disse to variablene. Jo lavere, desto mer signifikant er korrelasjonen. Denne er ikke signifikant hverken på 1%- eller 5%-nivået, og derfor har heller ikke korrelasjonen asterikser.

Dette er **N** – dvs. hvor mange observasjoner korrelasjonsberegningen er basert på. Denne kan variere fra korrelasjon til korrelasjon i tabellen, avhengig om vi har missing (mangler observasjoner) på noen av variablene våre. Vi ser her at vi har 215 observasjoner av Skostørrelse og Alder som danner grunnlaget for korrelasjonsberegningen.

Correlations

	Alder (oppgi i år)	Skostørrelse	Høyde
Alder (oppgi i år)	Pearson Correlation Sig. (2-tailed) N	1 ,100 216	,121 ,075 216
Skostørrelse	Pearson Correlation Sig. (2-tailed) N	,100 ,145 215	1 ,000 215
Høyde	Pearson Correlation Sig. (2-tailed) N	,121 ,075 216	,638** ,000 215

** . Correlation is significant at the 0.01 level (2-tailed).

Alle tall (korrelasjoner, N og sig) over og til høyre for den grønne streken, er de samme som tallene under og til venstre for denne streken. Dere kan se at ett-tallene danner en diagonal linje.

Dette er **korrelasjonen** mellom Høyde og Skostørrelse: Korrelasjonen er på 0,638 og er den høyeste av alle korrelasjonene i denne tabellen. De to asteriksene (**) ved siden av korrelasjonen forteller oss at korrelasjonen er signifikant på 1%-nivået (dvs. at den er signifikant forskjellig fra 0). Hva asteriksene betyr står angitt under tabellen.

Dette er det nøyaktige **signifikansnivået** – dvs. hvor sikre kan vi være på å ikke ta feil dersom vi påstår at det er en samvariasjon/korrelasjon i populasjonen som vi har hentet dataene våre fra mellom disse to variablene. Jo lavere, desto mer signifikant er korrelasjonen. Denne er veldig signifikant.