**Opgave – lineær regression, samfundsfag**

**Opstil 1 lineær regression. Hvis I har flere, så vælg 1. Fortolk regressionen og begrund sammenhængen teoretisk. Forhold jer kritisk til den – giver det i jeres tilfælde mening at opstille det som en lineær model?**

**Gør således:**

• Du skal konstruere og tolke et diagram med lineær regression med angivelse af ligning og R 2. Diagrammet skal forsynes med overskrift og aksetitler med måleenhed. Se hvordan her: <https://sites.google.com/view/excelisamfundsfag/lektion-7-line%C3%A6r-regression>

• Med afsæt i det konstruerede diagram skal du udlede, hvad tendenslinjens ligning (hældningskoefficienten), R2 -værdien og punkternes beliggenhed i diagrammet viser om sammenhængen mellem de to variable i diagrammet:

o Hældningskoefficienten i ligningen tolkes som, at hvis variablen på x-aksen stiger med 1, så ændres variablen på y-aksen med hældningskoefficienten. Eksempel: I ligningen y = -0,19x + 25,459 er hældningskoefficienten -0,19: Når x stiger med 1, falder y med 0,19 procentpoint.
Vær opmærksom på rigtig brug af måleenhederne.

o Punkternes beliggenhed i diagrammet: Hvad man kan udlede af punkternes beliggenhed i diagrammet om sammenhængen mellem de to variable? Ligger punkterne meget tæt på tendenslinjen, mere spredt omkring tendenslinjen som et bredt bælte, eller helt spredt som i en bisværm? Ligger punkterne på en måde, som er udtryk for lineær regression? Jo tættere punkterne ligger på tendenslinjen, jo højere bliver R2 .
Er der nogle punkter, der afviger meget fra tendenslinjen (outliers), og hvad kan det skyldes? Samler punkterne sig i nogle grupper, og hvad er det udtryk for?

o R2 viser forklaringsgraden, som udtrykker, hvor stor en del af variationen i den afhængige variabel y, der kan forklares/beskrives af den uafhængige variabel x. Hvis R2 fx er 0,39, kan prisen på en pakke cigaretter (x) forklare 39% af variationen i andel rygere i landene (y), mens de resterende 61% af variationen må forklares ud fra andre årsager. R2 skal tolkes i sammenhæng med punkternes beliggenhed i diagrammet.

o Den faglige sammenhæng mellem de to variable: Man skal anvende viden til forklare, hvilken faglig sammenhæng der er mellem de to variable. Så begrund den fundne sammenhæng teoretisk til sidst.

I alt fylder afleveringen ca. 1 s.

I kan læse mere her: <https://xn--vlgeradfrdogstatistik-f3bh.ibog.forlagetcolumbus.dk/?id=202>