

## Opgave 1

I [dette datasæt](#) fremgår løndata for 200 medarbejdere hos en virksomhed. Det antages, at sammenhængen mellem løn og anciennitet er lineær.

- i) Lav lineær regression på datasættet
- ii) Brug residualerne for den lineære regression til at afgøre, om sammenhængen mellem anciennitet og løn er lineær
- iii) Bestem den forventede løn efter 10 år.
- iv) Bestem sandsynligheden for, at lønnen efter 10 år er mellem 39000 og 41000 månedligt.
- v) **Svær:** Bestem et 95%-konfidensinterval for lønnen efter 5 år.

## Opgave 2

En gruppe af tomatentusiaster har en antagelse om, at en bestemt sort af tomatplanter bliver ved med at vokse i august. De har derfor målt vægten af deres tomatplanter i løbet af august måned. Vægten af planterne fremgår af [dette datasæt](#).

- i) Lav lineær regression på datasættet.
- ii) Brug residualerne for regression til at afgøre, om sammenhængen mellem forløbne dage og vægt er lineær
- iii) Bestem et 95%-konfidensinterval for hældningen  $a$ , og brug dette til at afgøre, om vi kan konkludere, at planterne stadig vokser.

## Opgave 3

En virksomhed producerer slikposer i størrelserne 50, 100, 150, 200 og 250 gram. De har målt variationen på posens forventede vægt og den reelle vægt. Resultatet af deres undersøgelse fremgår af [dette regneark](#).

- i) Lav lineær regression på datasættet.
- ii) Brug residualerne for regressionen til at afgøre, om sammenhængen mellem den forventede vægt og variationen fra den forventede vægt er lineær.
- iii) Bestem et 95%-konfidensinterval for hældningen af den lineære sammenhæng.
- iv) Tegn tre lineære funktioner i samme koordinatsystem; én for hver af  $a$ -værdierne:

- Den nedre grænse for konfidensintervallet,
  - $\alpha$ -værdien, som regressionen giver,
  - Den øvre grænse for konfidensintervallet.
- v) **Svær:** Bestem sandsynligheden for, at en pose, der bør veje 200 gram har en vægt på mellem 180g og 220g