# Problem 3: Customer satisfaction score

We aim to investigate the factors influencing customer satisfaction at *FashionFancyClothes* ready-to-wear shops. The file `satisfaction.txt` comprises satisfaction scores from 400 *FashionFancyClothes* customers, along with their total purchase amounts, membership duration, age, and preferred store location. We want to analyze how these factors influence customer satisfaction using a linear model of the form:

$$\texttt{score} = \beta_0 + \beta_1 \, \texttt{purch\_amount} + \beta_2 \, \texttt{memb\_duration} + \beta_3 \, \texttt{age} + \epsilon \tag{1}$$

with $\epsilon \sim \mathcal{N}(0, \sigma^2)$

a) Fit the model and report the estimates of the unknown parameters.

b) State and verify the model assumptions.

c) Perform a test of level 5% to verify whether `memb_duration` and `age` have an effect on the satisfaction score.

d) Perform any other statistical tests that you consider useful to reduce the model, and update the estimates of its parameters.

e) Let's now add to the model found in (d) the variable `store`, added as a random intercept. Fit the new model and report the PVRE index.

f) Report the dot plot of the estimated random intercepts. Ignoring the effect of the fixed effect covariates, which store is associated to the highest score?

Upload your results here:
`https://forms.office.com/e/xkWiqD55z7`