# Homework 4

### Introduction to Data Science (QR2)

due: **10/8/25**

Instructions: Please submit solutions on Populi. Only a knitted pdf of an `Rmarkdown` file will be accepted.

**Problem 1:** Write a short paragraph on a topic you found interesting in the readings related to prediction. What sparked your interest and what did you learn?

**Problem 2:** In this problem, you will investigate the impact of faculty attractiveness on teaching evaluation score. Yes, this has been rigorously studied! In your Rstudio, load in the data `beautydata.csv` and fit a simple linear regression model of `CourseEvals` on `BeautyScore`. `CourseEvals` are on the standard Likert scale, and `BeautyScore` is based on a panel of undergraduate students rating headshots of teachers. A couple suggestions for code:

```
# load in the data
beautydata = read.csv(''beautydata.csv'')
# fit a model, note the similarity of lm() and model_train() in our computing tutorials.
modelfit = lm(CourseEvals ~ BeautyScore, data = beautydata)
# print out a summary of the fitted model!
summary(modelfit)
# Include extra explanatory variables on the right-hand-side, like gender and english-speaker.
modelfit = lm(CourseEvals ~ BeautyScore + female + nonenglish, data = beautydata)
```

a. Interpret all coefficient estimates in the regression outputs. What do you conclude about the effect of attractiveness on evaluations, and why do you think this is the result?

b. Include an additional variable in your model that is the product of gender and beauty score, like: `female*BeautyScore`. How do your coefficient estimates change? Why do they change and what is their new meaning?