

Homework 8

INTRODUCTION TO DATA SCIENCE (QR2)

due: 11/7/25

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

Problem 1: This question is related to your project.

1. Run a simulation to produce the “Top decile” bottom row of Table 5.3 in the [SFFA expert report](#). Comment on the racial breakdown of this random lottery among top decile applicants to the actual racial breakdown in the admitted class. The `sample()` function will be useful.
2. Fit and interpret a regression model with your data. Please be sure to write out the model you are fitting and how it pertains to your question of interest. Remember that `glm(..., family = “binomial”)` is used for logistic regression and `lm()` is used for standard linear regression.