

Final exam — Sample 1

DATA 412/612

Practice exam, questions only. Cumulative through December 9 (weeks 1–14). The newest material is lists and purrr (map, map2, keep, and split-plus-map). Earlier weeks stay in scope.

Put your answers in an R Markdown document and keep the code visible (`knitr::opts_chunk$set(echo = TRUE)` in the setup chunk). You may use your notes, the R help files, and the internet. Do not discuss the exam with anyone else. Load the tidyverse.

1. **Lists** (6 pts). Build a named list `fit` with components `n = 10` and `center = c(1, 2)`. Extract `center` so the result is the numeric vector, not a length-one list. Then replace `n` with 12 without rebuilding the list.
2. **map** (8 pts). Keep only the numeric columns of `mpg` and, with `map_dbl()`, return the mean of each. Pass `na.rm = TRUE`. The function argument must be a function, not a call.
3. **map2** (8 pts). `mu` is `c(0, 10, 20)` and `sigma` is `c(1, 1, 2)`. Use `map2()` and `rnorm` to draw 5 values from each pair. Set the seed to 612 first. Return a list of three numeric vectors.
4. **Split, then map** (10 pts). Split `mpg` by `drv`. On each piece, fit `hwy` on `displ` and return the slope on `displ`. The result should be a named numeric vector, one slope per drive.
5. **One pipeline from earlier weeks** (8 pts). Load `nycflights13`. Add the airline name onto `flights`. For each name, report the number of flights and the mean arrival delay, ignoring missing delays. Keep names with at least 1000 flights. Make a bar plot of that mean delay, with the names ordered by the mean. Title the plot.