

Final exam — Sample 2

DATA 412/612

Practice exam, questions only. Cumulative through December 9 (weeks 1–14). The newest material is lists, loops, and **purrr** (**pmap**, **keep**, **discard**). Earlier weeks stay in scope, including a report chunk.

Put your answers in an R Markdown document and keep the code visible except where a chunk option below says to hide it. You may use your notes, the R help files, and the internet. Do not discuss the exam with anyone else. Load the tidyverse.

1. **A loop and a map** (8 pts). **bases** is `c(2, 3, 4)`. With a pre-allocated **for** loop, fill a numeric vector with the square of each entry. Then do the same job with `map_dbl()`. The two results should match.
2. **pmap** (6 pts). Using the same means and standard deviations as `mu = c(0, 10)` and `sigma = c(1, 2)`, draw 4 normal values from each pair with `pmap()` and a list of arguments. Set the seed to 612 first.
3. **keep and discard** (6 pts). From `mpg`, **keep** the numeric columns and report how many there are. Then **discard** every column that is not a character column, and print the names that remain.
4. **Tidy, then test** (12 pts). Reshape `wide` to a `year` column and a `sales` column. `year` is numeric. Drop the missing sales. Turn `store` into a factor ordered by total sales, largest first. Run a two-sample `t.test()` of `sales` for store A against store B. State the null hypothesis in one sentence.

```
wide <- tibble(
  store = c("A", "B", "A"),
  y2022 = c(5, 9, NA),
  y2023 = c(7, 4, 6)
)
```

5. **The report around that test** (8 pts). Put the test in a chunk labeled `store-test` with `echo = TRUE`. Put the reshape in a setup-style chunk with `include = FALSE`. Send the store totals to `knitr::kable()` with a caption. In YAML, set a bibliography file `refs.bib`, and in the text cite `wickham2017r4ds` in parentheses.