

Midterm 1 — Sample 2

DATA 412/612

Practice exam, questions only. Cumulative through October 14 (weeks 1–7): the Quiz 1 material, plus `rowwise()` and `across()`, `case_when()`, `readr`, exploratory plots, and `tidyr`.

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. **A function** (8 pts). Write `finite_mean(x)`. Drop every `NA` and every non-finite value, then return the mean of what remains. If nothing remains, return `NA`. Use a pipe inside the function. Try it on `c(1, NA, Inf, 3, NaN)`.
2. **`case_when()` and `distinct()`** (8 pts). In `mpg`, add `drive`: "front" when `drv` is "f", "rear" when it is "r", and "four" when it is "4". Return the distinct pairs of `class` and `drive`, and the number of cars in each pair.
3. **Import** (10 pts). Read the text below with `read_csv()`. Keep `id` as character, including the leading zeros. Parse `when` as a date. Leave `amount` numeric. Show the problems `readr` reports, and print the tibble after a successful parse of the rows that do have a real date.

```
id,when,amount
007,2024-01-05,12
008,2024-02-11,NA
009,not-a-date,4
```

4. **Separate and widen** (10 pts).
 - a. Split `place` into `city` and `state`.
 - b. Turn `long` into a table with one row per city and one column per quarter.

```
places <- tibble(place = c("Austin, TX", "Boise, ID", "Tampa, FL"))

long <- tibble(
  city = c("Austin", "Austin", "Boise", "Boise"),
  quarter = c("Q1", "Q2", "Q1", "Q2"),
  sales = c(10, 12, 7, 9)
)
```

5. **A faceted plot** (4 pts). From `mpg`, plot `displ` against `hwy`, with one panel per `drv`. Color is not required. Title the figure.