

BioR — Class 3 Live-Script

ggplot2: data visualization from tidy biological tables

Setup

1) Open the BioR RStudio Project (same structure as Class 1):

- data/ scripts/ outputs/ outputs/figures/
- 2) Put the dataset file in: data/BioR_Class3_dataset.csv
- 3) Start a new script: scripts/03_ggplot2_practice.R

Dataset

Columns:

- sample_id (site+treatment+rep)
- site
- treatment (CTRL/HEAT)
- rep (R1–R3)
- day (0,7,14,21)
- date
- temperature (°C)
- biomass (g m⁻²)

Starter code (copy/paste)

```
library(tidyverse)
library(here)

dat <- readr::read_csv(here("data", "BioR_Class3_dataset.csv"))

# Quick check
dplyr::glimpse(dat)
```

Exercises

E1 — Scatter + model trend (core pattern)

- Make a scatter plot of biomass vs temperature.
- Map treatment to color.
- Facet by site.
- Add a linear trend (`lm`) with confidence band.
- Add axis labels with units.

E2 — Time series with correct grouping

- Plot biomass over day.
- Use lines + points.
- Color by treatment.
- IMPORTANT: group by rep (or sample_id) so lines connect correctly.
- Facet by site.

E3 — Distribution at a single time point

- Filter to day == 14 and compare biomass by treatment.
- Use a boxplot.
- Add jittered points.
- Make CTRL appear before HEAT on the x-axis.

E4 — Polish (theme + titles + legend)

- Pick one of your plots and polish it:
- Add a title and a short subtitle.
- Use `theme_minimal()` or `theme_bw()` and adjust text size.
- Move the legend to the top.

E5 — Export correctly

- Save two figures to outputs/figures/:
- PNG at 300 dpi (e.g., width 7, height 4)
- PDF (vector) with the same size.
- Use `here()` for paths.

Optional extensions (if we finish early)

- Use a \log_{10} y-scale and update the y-axis label accordingly.

- Add separate trend lines per treatment (already happens if color is mapped) — verify it.
- Create a small-multiples figure with `facet_grid(treatment ~ site)`.
- Make a multi-panel figure (patchwork) combining E1 and E2.

Common pitfalls (quick fixes)

No legend appears: You likely set color outside `aes()`. Use `aes(color = treatment)` for mapping.

Lines connect the wrong points: You forgot `group = rep (or sample_id)` in `aes()`.

Weird ordering on x-axis: Convert to factor and set levels: `factor(treatment, levels = c('CTRL','HEAT'))`.

'object not found': Column name typo; check `names(dat)` or `glimpse(dat)`.

'+' errors: Put `+` at the END of the line in ggplot code, not the beginning.

Submission checklist

- `scripts/02_ggplot2_practice.R` runs top-to-bottom.
- Two figures saved in `outputs/figures/` (PNG + PDF).
- Figures have axis labels (with units) and a readable theme.