

Data Visualization Challenge

R Setup in Cursor + Penguins and Simpson's Paradox

Data Viz Challenge - Quick Start Guide

Challenge Overview

In this challenge, you'll set up R inside Cursor (VS Code-compatible), get comfortable with terminals, shells, and command lines, and then prepare to build a two-panel or two-facet ggplot using the Palmer Penguins dataset to illustrate Simpson's paradox.

By the end of Part 1, you will be able to run R code from Cursor and render R chunks in Quarto. Plotting tasks are scaffolded with placeholders you will complete later.

Part 1: R in Cursor Setup

Follow these steps (adapted from the official R extension guidance) to get R working smoothly in Cursor.

1. Install R

- Download and install R (version 3.4.0) from [CRAN R downloads](#)
- During installation on Windows, select the option to **Write R path to the registry** (recommended)

2. Install the R extension in Cursor

- Open Extensions: `Ctrl+Shift+X`
- Search for "R" (publisher: REditorSupport) and install the extension named "R"
- Optional but helpful extensions:
 - "Quarto" (for .qmd editing and preview)
 - "VSCoDe-R-Debugger"

3. Install languageserver (enables R language features)

You can install directly from an R session. In Cursor:

- Open Command Palette: **Ctrl+Shift+P** → type “R: Create R Terminal” → Enter
- In the R console that appears, run:

```
install.packages("languageserver")
```

4. Recommended tools for a better R experience

- radian (modern R console): see [radian project page](#)
- httpgd (SVG/websocket graphics device for the plot viewer):

```
install.packages("httpgd")
```

After installing httpgd, the R extension plot viewer can display SVG plots served by httpgd.

5. Verify R works in Cursor

- 1) Create a new file `hello.R` and add a couple lines:

```
summary(cars)  
mean(cars$speed)
```

- 2) With the editor caret on a line, press **Ctrl+Enter** to send the current line to the active R terminal.

- 3) If Cursor can't find R automatically on Windows, set the R path:

- Open Settings: **Ctrl+,**
- Search for “Rterm”
- Set R > Rterm: Windows to something like `C:\\Program Files\\R\\R-4.4.1\\bin\\x64\\R.exe` (adjust version/path as installed)

6. Running R chunks in Quarto (.qmd)

With the R extension (and Quarto extension installed), you can execute R code chunks inside Quarto documents.

- Create or open a .qmd file and insert an R chunk:

```
R.version.string
```

- Use the chunk toolbar (Run/Run Above/Run All) to execute. If the toolbar is hidden, place the cursor in the chunk and use **Ctrl+Shift+P** → “Quarto: Run Cell” or send lines with **Ctrl+Enter**.

Part 2: Terminals, Shells, and Command Lines

Understanding the differences helps when switching between the system terminal and the R terminal.

- Terminal (or terminal emulator): The window/app that hosts a shell (e.g., Cursor’s Integrated Terminal)
- Shell: The command interpreter you run in a terminal
 - Examples (Windows): PowerShell, `cmd.exe`
 - Examples (macOS/Linux): `bash`, `zsh`
- Command-line interface (CLI): The programs you run inside the shell (e.g., R, `quarto`, `git`)

In Cursor: - Open the integrated terminal: **Ctrl+``** (backtick) - Create an R terminal: **Ctrl+Shift+P** → "R: Create R Terminal" - Use the R terminal for R commands; use the system shell (PowerShell/bash) for tools like `git` and `quarto`

Part 3: Penguins + Simpson’s Paradox (Placeholders)

You’ll use Palmer Penguins to demonstrate how a relationship can reverse when stratified (Simpson’s paradox). For now, set up the project and confirm packages install.

A. Install and load packages

```
install.packages(c("palmerpenguins", "ggplot2"))

library(palmerpenguins)
library(ggplot2)
data(penguins)
```

B. Single overall trend (placeholder)

```
# TODO: Create a single scatterplot with a smooth trend line
# Example variables: bill_length_mm vs flipper_length_mm (or body_mass_g)
# ggplot(penguins, aes(x = bill_length_mm, y = flipper_length_mm)) +
#   geom_point(alpha = 0.7) +
#   geom_smooth(method = "lm")
```

C. Stratified view via facets (canonical)

```
# TODO: Facet by species (or island) to reveal group-wise trends
# ggplot(penguins, aes(x = bill_length_mm, y = flipper_length_mm, color = species)) +
#   geom_point(alpha = 0.7) +
#   geom_smooth(method = "lm", se = FALSE) +
#   facet_wrap(~ species)
```

D. Two-panel composition (canonical alternatives)

```
# TODO: Show overall vs stratified side-by-side
# Option 1: Use facets as above (keeps in one ggplot)
# Option 2: Create two separate ggplots and combine with patchwork or gridExtra

# install.packages("patchwork")
# library(patchwork)

# p_overall <- ggplot(...) + ...
# p_facet   <- ggplot(...) + ... + facet_wrap(~ species)
# p_overall + p_facet
```

E. Brief write-up (placeholder)

- Explain how the overall relationship differs from the within-group relationships
 - Identify the grouping variable that flips or alters the trend (species or island)
-

Submission Checklist

- ☐ R installed (Windows: path written to registry)
 - ☐ R extension installed in Cursor (plus Quarto extension)
 - ☐ `languageserver` installed
 - ☐ `httpgd` installed (recommended)
 - ☐ Able to send lines from `hello.R` to the R terminal
 - ☐ Able to run an R chunk inside a `.qmd`
 - ☐ Packages `palmerpenguins` and `ggplot2` installed
 - ☐ Placeholder plots and notes added to this document
-

Need Help?

- R extension documentation: [VS Code R Extension Wiki](#)
- Quarto docs: [Quarto Guide](#)
- General tip: search for “VS Code” results—most answers apply to Cursor as well