Penguin Measurements Dashboard

Oliwia Wójcicka

February 23, 2026

1 Project Overview

This project is a basic R Shiny web application for interactive exploration of penguin body measurements.

The source code for the project is available in the GitHub repository: GitHub Repository.

The dataset used is the Palmer Penguins dataset (`palmerpenguins::penguins`), converted to a data frame and cleaned to ensure stable plotting. Figure 1 shows the overall appearance of the application, including its main layout with a sidebar for user controls and a central panel for data visualisation. This interface allows users to adjust parameters interactively and immediately observe the resulting changes in the plot.

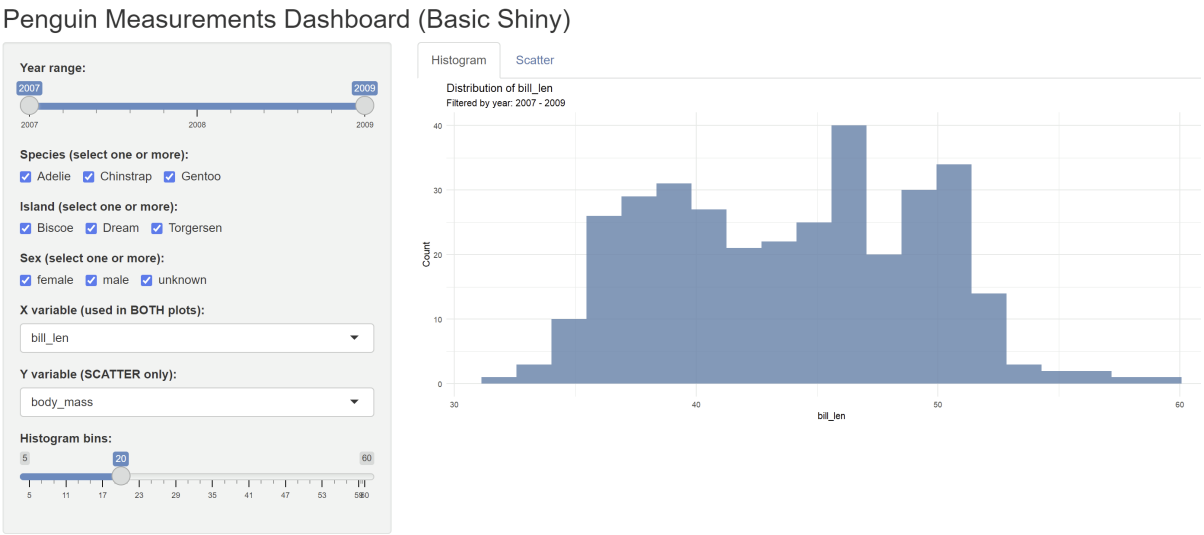

Figure 1: Application layout with sidebar controls

2 Dataset Description

The dataset is loaded in R using:

```
data <- as.data.frame(penguins)
```

The data frame used in this project contains the following columns:

- **species**: penguin species (categorical), e.g., Adelie, Chinstrap, Gentoo
- **island**: island of observation (categorical), e.g., Biscoe, Dream, Torgersen
- **bill_len**: bill length (numeric, mm)

- **bill_dep**: bill depth (numeric, mm)
- **flipper_len**: flipper length (numeric, mm)
- **body_mass**: body mass (numeric, g)
- **sex**: sex category (categorical: female/male; missing values handled as "unknown")
- **year**: year of measurement (integer)

3 Data Preprocessing

To avoid errors and keep the app simple, rows with missing values in the key numeric measurement columns are removed:

`bill_len, bill_dep, flipper_len, body_mass`

Additionally:

- `species` and `island` are converted to factor variables,
- `sex` is converted to character, missing values are replaced with `"unknown"`, then converted back to factor,
- `year` is converted to integer to support a year range slider.

This preprocessing ensures filtering and plots work reliably in Shiny.

4 User Controls (Reactive Inputs)

All controls in the sidebar are reactive inputs. Any user change triggers automatic recalculation of the filtered dataset and refreshes the plots.

Controls

- **Year range slider**: filters observations between selected years (e.g., 2007–2009).
- **Checkbox filters (multiple selection)**: *species*, *island*, and *sex*.
- **X variable selector (used in BOTH plots)**: chooses the numeric feature for the histogram and the X-axis in the scatter plot.
- **Y variable selector (SCATTER only)**: chooses the numeric feature for the Y-axis of the scatter plot.
- **Histogram bins slider**: controls the number of bins and the granularity of the histogram.

5 Reactive Logic

In the server part, the app creates a reactive object `filtered_data()`:

- it starts from the cleaned dataset,
- applies filters in this order: **year range** → **species** → **island** → **sex**,
- returns a filtered data frame used by both plots.

If filtering results in zero rows, the application displays a message ("No data left after filtering...") instead of crashing.

6 Dynamic Outputs (Plots)

6.1 Histogram

The histogram shows the distribution of the selected **X variable** within the currently filtered subset.

- X-axis: chosen numeric measurement (e.g., `bill_len`)
- Y-axis: count of observations per bin
- Number of bins: controlled by the histogram bins slider

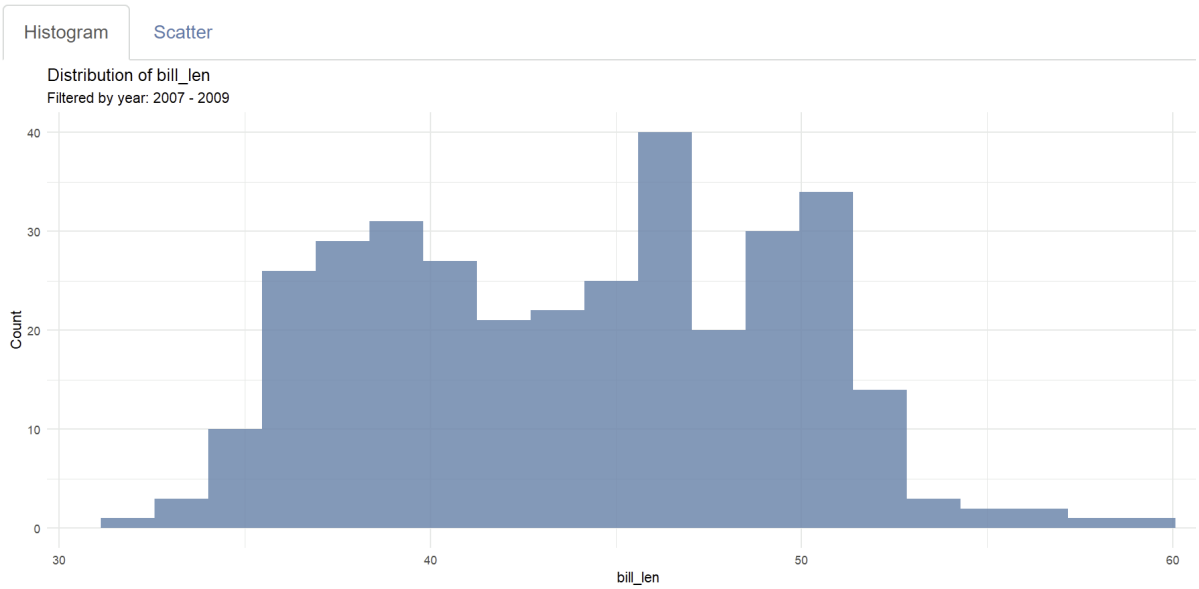

Figure 2: Histogram of the selected X variable after applying filters

6.2 Scatter Plot

The scatter plot shows the relationship between the selected **X** and **Y** variables.

- X-axis: selected X variable
- Y-axis: selected Y variable (scatter only)
- Color: points are colored by species to compare groups

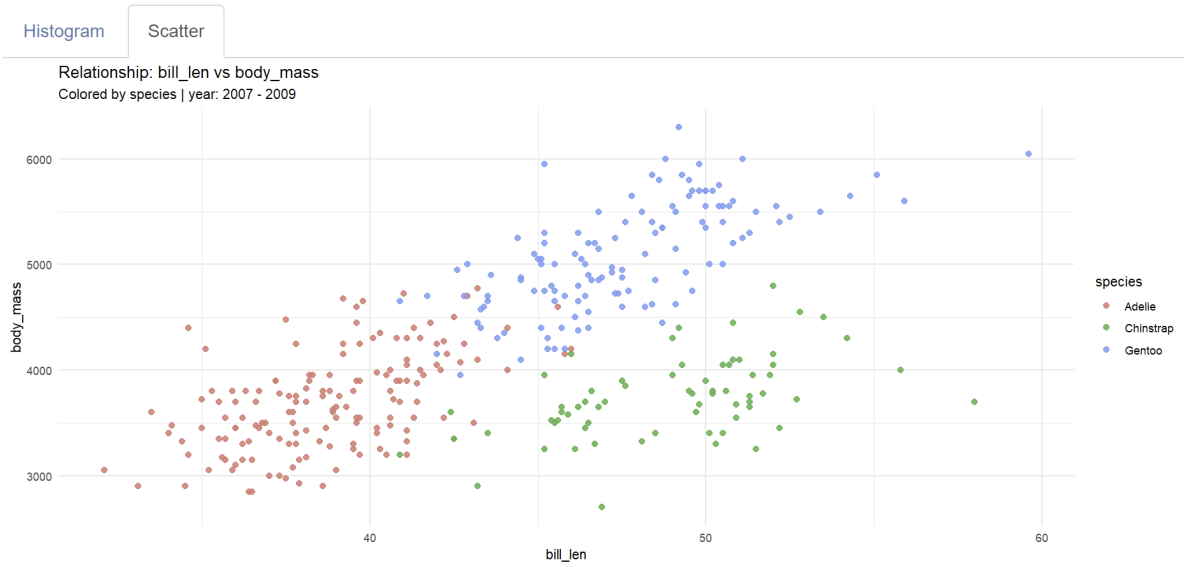

Figure 3: Scatter plot of X vs Y, colored by species

7 Conclusion

This Shiny application demonstrates a basic reactive dashboard for penguin measurements. Users can filter the dataset through sidebar controls and instantly see updated visualizations. The project meets the assignment requirements by implementing a sidebar layout, reactive inputs, and dynamic plot outputs.

8 Use of AI Assistance

ChatGPT was used for assistance with English translation and LaTex report formatting.