

Introduction to ggplot

Intro R seminar homework Week 4

The Assignment

02 February 2020

This assignment requires use of the `iris` dataset, which is a default dataset available in R. It also requires use of `ggplot` functions, which are not available in base R. Load the required package.

Data structure

Show the structure of the `iris` data. Which variables are continuous? Which are categorical?

Scatterplot

Simple defaults

Using the `iris` data, create a simple scatterplot with `Sepal.Length` against `Sepal.Width`. Default plot options are fine, but if you're feeling saucy, you can change the theme, or point shape, or anything else.

Add a categorical variable

Now add information to the scatterplot with `Sepal.Length` against `Sepal.Width`. Plot the points by the categorical variable, `Species`.

Plotting linear relationships

Add `smoothed` trendlines to show the linear relationship between sepal width and sepal length for each species. Ensure lines and points have the same color within each species.

Facet by categorical variable

Display the relationship between sepal width and sepal length within the categorical variable `Species` in three side-by-side panes, instead of within a single graph pane. Tweak your `aes()` to ensure that the categorical variable is not accounted for twice in the final graph.