Exercises for the course

"An introduction to R"

Exercise session Data visualization and graphics: Monday - March 9, 2020

Exercise 1: *Making simple plots in base R*

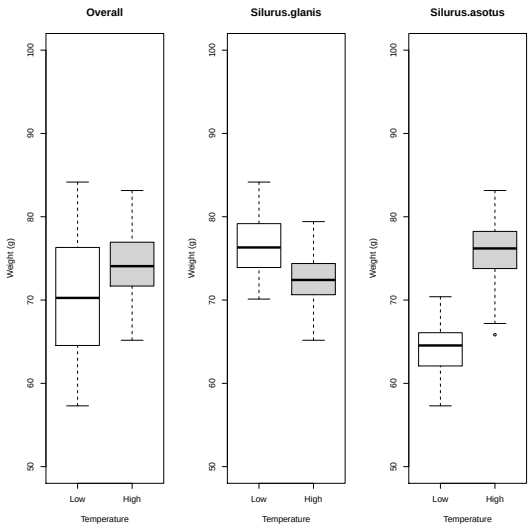
a) Download the data file `Catfish_2.csv` from the web page and import it into R. Get an overview of the data by using `str()` and `head()`. Plot a histogram of the variable Weight. Use a colour of your choice and set the number of bins to 50.

b) The previous plot contains the weight measurements of both month March and April. Plot a histogram that displays only the values for April. As you wish to add a density curve make sure that probability densities, instead of frequencies, are displayed. Add a blue density line with a width of 1.5 to the plot.

c) Make a scatterplot and plot the length against the weight of each fish. Change the label of the x-axis to "Length (cm)" and the label of the y-axis to "Weight (g)".

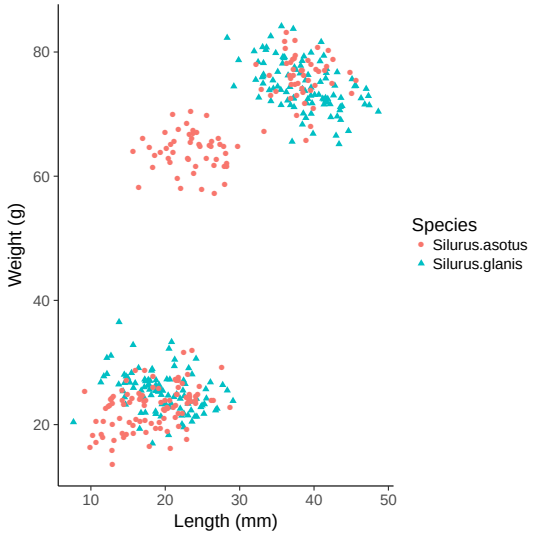
Exercise 2: *Multi-panel figures*

Make a subset (Catfish_April) that includes only the measurements for the month April and produce a figure, which resembles the multi-panel figure below. Use the command
`boxplot(Weight ~ Temperature, data = Catfish_April, xlab = "Temperature", names=c("Low","High"), col=c("white", "lightgrey"), . . .)`
to produce the left boxplot and adapt the above command to produce the other boxplots. In the data argument you can use Catfish_April[] and logical operators(e.g. ==) to choose a specific species for plot two and three. You can change the ratio of height and width of your multi-panel figure by changing the size of your plotting window.

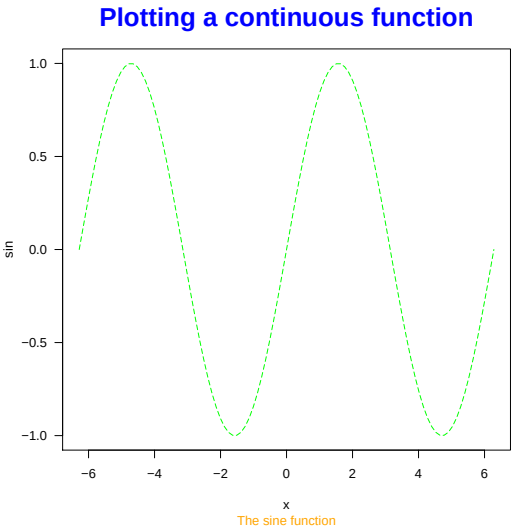


Exercise 3: *Graphics with ggplot2*

Try to reproduce the following plot using the package `ggplot()`. You might have to install and load `ggplot2` first. For the plot, the point size is set to 2, and the classic theme is used with a base_size = 17. For more help visit the `ggplot2` documentation website http://docs.ggplot2.org. Under `geom_point` you will find several examples that will help you to make this plot.



Exercise 4: *Plotting a continuous function* Use `plot()` to produce the following plot of the sine function on the interval $[-2{\cdot}\pi, 2{\cdot}\pi]$:



The text of the main title (`main =`) is magnified (`cex.main =`) by a factor 2. The line type is `'longdash'`. The colours used in the plot are 'blue', 'green' and 'orange'. Furthermore you need the option `las =` to obtain horizontal axis labels.

Exercise 5: *We learned that there are many pre-defined data sets you can use in R. The data frame* `cars` *is one of those and gives the distance taken to stop from a certain speed. Note that the data were recorded in the 1920s.* If `cars` is not known, then say `data(cars)` in order to load it. Plot `dist` as a function of `speed` without titles and without axes (`axes = FALSE`). Add a line

with slope 4 and intercept -17.5. Furthermore add axes with the command `axis()`. Specify the positions of the ticks with `at =`. 'Pretty' positions of the ticks are generated with `pretty(speed)` and `pretty(dist)`, respectively. You might want to draw the axis into the margins by one line (`line = 1`). Add a legend with `legend()`. The text in the legend is generated with the command `expression()`. Furthermore add a main title and a subtitle with the command `title()`.