Reading and Writing Data

Exercise 1: Importing sparrow data

a) Download the data file Sparrows.csv. This file contains Weight and WingLength measurements of three treatment groups.

Use **read.table(file = "Sparrows.csv", header = TRUE)** to import the data into the object named Sparrows. Afterwards, use the **View()** function to display the data.

b) Something must have gone wrong! All variables are combined in one big column. Can you figure out why R does not import the data set correctly? Write a short answer.

Hint: Consult the help page ?read.table and find out the default for the argument 'sep'.

c) Fix this issue by specifying the 'sep' argument in read.table() and import the sparrow data again. Use View() to check if the import worked correctly this time.

Exercise 2: Get an overview

a) Get an overview of the sparrow data with the command str().
b) Do you know another function that provides an overview and basic statistics of each variable? Use the function to return the minimum, median, mean, and maximum for Weight and WingLength measurements.
c) Display the first 10 and the last 15 rows of the large 'Sparrow' data set.

Exercise 3: Checking and cleaning data frames

a) If you see a couple of extra columns (e.g. X, X.1) that have been accidently included and only contain NAs, remove these by indexing with square brackets.
b) During data entry, three rows have been entered twice. Which are the duplicate rows? Remove the duplicate rows from the data frame.

Hint: it might be faster if you incorporate which().

c) Display the levels of the factor 'Treatment'?

Exercise 4: Missing values

a) Find out which rows in the variable 'WingLength' contain NAs

Hint: it might be faster if you incorporate which().

b) Replace NAs with the values 26, 16,30, 30.5,30.5 and 29 (in this order). Use which(is.na()) to check if your fix worked.

Exercise 5: Sorting data frames

a) Sort the Sparrow data frame by ‘Weight’ and ‘WingLength’ column, and create a new object Sparrows_sorted.
b) Show the individuals with the six longest WingLength. To do that, use head() and order(), if possible in the same command line.

Exercise 6: Export data

a) Save the ordered table (by ‘Wing’ and ‘WingLength’ column) into a comma separated file named ‘Sparrows_sorted’ on your Desktop.
b) Save your script and remove all objects from the working environment.