

Exercise 1 - vectors and data frames

- The following table gives the volume, area, length and maximum and mean depths of some Scottish lakes[1]. Create vectors, holding the lake's name and all the parameters and build a dataframe called `scottish.lakes` from the vectors

- evaluate the highest and lowest volume and area lake
- order the frame with respect to the area and determine the two largest area lakes
- by summing up the areas occupied by the lakes, determine the area of Scotland covered by water

message function -> print a string
 paste/paste0 function -> put everything together and the function makes a string
 print function -> used for functions usually

Loch	Volume [km ³]	Area [km ²]	Length [km]	Max. depth [m]	Mean depth [m]
Loch Ness	7.45	56	39	230	132
Loch Lomond	2.6	71	36	190	37
Loch Morar	2.3	27	18.8	310	87
Loch Tay	1.6	26.4	23	150	60.6
Loch Awe	1.2	39	41	94	32
Loch Maree	1.09	28.6	20	114	38
Loch Ericht	1.08	18.6	23	156	57.6
Loch Lochy	1.07	16	16	162	70
Loch Rannoch	0.97	19	15.7	134	51
Loch Shiel	0.79	19.5	28	128	40
Loch Katrine	0.77	12.4	12.9	151	43.4
Loch Arkaig	0.75	16	19.3	109	46.5
Loch Shin	0.35	22.5	27.8	49	15.5

Exercise 2 - Athletics world's races

- The following Web server [2] contains an updated list of all the best timings in official 100 m athletics races.
- open the results on the Web, reported as a table and create a `tibble` that contains all the results
- to parse the Web data, use the `rvest` and `tidyverse` packages and the following code:

```
library(rvest)
library(tidyverse)
men100m_html <- read_html("http://www.alltime-athletics.com/m_100ok.htm")
men100m_html |> html_nodes(xpath = "//pre") |> html_text() -> men100m_list
men100m_tbl <- read_fwf(men100m_list)
                    specify an object, not a file
```

- the resulting `tibble` has all data represented as "characters". Explore the `tidyverse` packages and manipulate the `tibble` to add a name of the columns and to convert the columns to the proper format
- once the `men100m_tbl` is properly formatted, using the `ggplot2` package, realize the following plots:
 - (1) evolution of the fastest time as a function of date of the races
 - (2) which country had the most fastest runners for each year ?
- perform the same analysis for the women 100m races and compare the results obtained. Which country has the fastest number of men and women ?

References

- [1] Lakes of Scotland: https://en.wikipedia.org/wiki/List_of_lochs_of_Scotland
- [2] Track, Field, and Athletics: all-time, performances https://www.alltime-athletics.com/m_100ok.htm