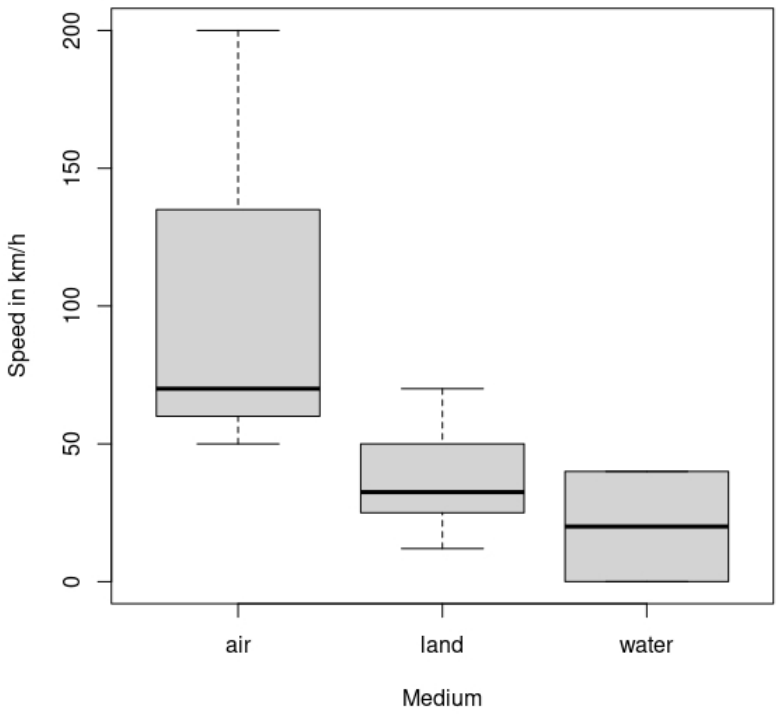
DAT-B-INFER Data Science 2 - Inference
Practical 1: Using R

1. Working with dataframes.

   (a) Create a dataframe called `animals` with columns names and values as follows; make sure `medium` is a *factor*. We will be using this data frame again later:

	species	speed	medium	weight
1	swift	200.00	air	0.02
2	falcon	70.00	air	0.70
3	goose	70.00	air	2.20
4	starling	50.00	air	0.05
5	cheetah	70.00	land	50.00
6	horse	50.00	land	450.00
7	hippo	30.00	land	2500.00
8	man	25.00	land	80.00
9	squirrel	12.00	land	0.60
10	cat	35.00	land	4.00
11	shark	40.00	water	600.00
12	seahorse	0.02	water	0.10

   (b) Create the following plot in R.



(c) Select all animals that live in water from the data.frame.

(d) Write a piece of code that finds how many animals in the data.frame belong to each of the three media.

2. Functions

(a) Write a function called `skewed` that calculates the skewness index ($I = 3(\bar{x} - p_{50})s$) of the data set where $\bar{x}$ is the mean of the data, p_{50} is the median and s is the standard deviation. It should return either "The data is significantly skewed" if the index is greater than 1 and "The data is not significantly skewed" otherwise.

(b) Generate 100 observations, x, that are standard normally distributed and use your function to determine if the data are skewed. Verify the result visually.

(c) Find $y = x^2$ and check whether your resulting data are skewed.

3. R packages

(a) Write functions that convert temperatures from F to C and vice versa. The formula to obtain degrees C from degrees F is:

$$C = (F - 32) * 5/9$$

(b) Create an R package `temp_convert`. Make sure that you include all necessary information so that another user is able to utilize your package.