

# Simple data analysis in R

math612.0 A1: From numbers through algebra to calculus and linear algebra

Gunnar Stefansson (editor) with contributions from very many students

March 7, 2022

# Entering data; dataframes

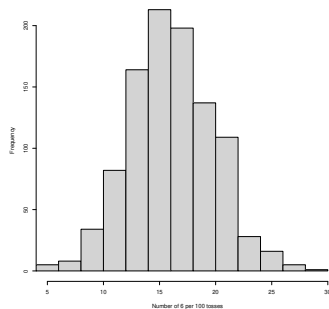
Several methods exist to enter data into R:

- 1 Enter directly: `x<-c(4,3,6,7,8)`
- 2 Read in a single vector: `x<-scan("filename")`
- 3 Use: `x<-read.table("file address")`

Example: Three methods to enter data into R.

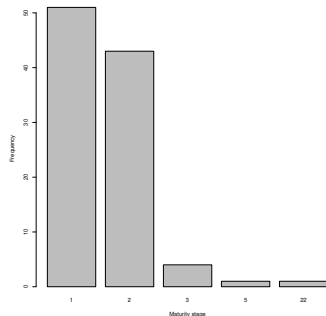
# Histograms

A histogram is a graphical display of tabulated frequencies, shown as bars.  
In R use the command: `hist()`



# Bar Charts

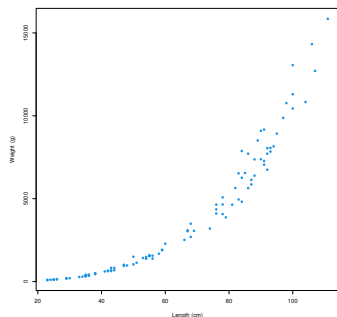
The bars in a bar chart usually correspond to frequencies in categories and are therefore kept apart.



# Mean, standard error, standard deviations

# Scatter plots and correlations

If we have paired explanatory and response data we are often interested in seeing if a relationship exists between them. To do this, we first plot the data in a scatter plot.



**Figure:** Figure: Scatter plot showing the length-weight relationship of fish species "X". Data source : Marine Resource Institution - Iceland.

**Copyright** 2021, Gunnar Stefansson (editor) with contributions from very many students

This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.