

Exercise: How much do you remember from Quantitative Reasoning?

YSC2210 - DAVis with R

Michael T. Gastner

14 January 2022

The purpose of this exercise is to remind you of the R and RStudio skills you learned in Quantitative Reasoning. This course does *not* directly build on preexisting knowledge. However, when you return to this exercise at the end of this course, you will be able to contrast the basic methods we used in Quantitative Reasoning with the advanced methods that you will have learned in the meantime. For this exercise, do not feel ashamed to tackle the problems in a relatively basic way.

1 Passengers and crew on the Titanic

The data at https://michaelgastner.com/data_for_QR/titanic.csv is a spreadsheet with information about passengers and crew on board of the Titanic during its maiden voyage in 1912, which ended tragically when the ship hit an iceberg and sank. I found these data on the website of the “Encyclopedia Titanica”.¹

There are eleven columns:

- family name
- given name
- gender
- age
- class
- whether the passenger survived
- ticket number
- number of passengers travelling with this ticket
- three columns that indicate the ticket price in British pounds, shillings and pence

In this exercise, we conduct some basic analysis of these data and visualise them with the methods you learned in Quantitative Reasoning.

- Create a project titled `titanic`.
- Download the spreadsheet and save the file in the project directory. Also start a new R script called `titanic.R` in this directory.
- The spreadsheet file has the extension `.csv`, which stands for “comma-separated values”. What is the reason for this name? Include your answer as a comment in your R script.
- Import the data as a data frame called `titanic`. Include the import command in `titanic.R`.
- How many rows and columns does `titanic` have? Here and in the following tasks, use R to find the answer and include the answer as a comment in your script.
- What are the unique values in the column `class`?
- How many people were in each class?

¹Encyclopedia Titanica, URL: <https://www.encyclopedia-titanica.org/>. Accessed on 2020-07-26.

- (h) Create a subset of `titanic` called `sec_class` that contains only second-class passengers.
- (i) What is the median age of second-class passengers?
- (j) How many second-class passengers were exactly 30 years of age?
- (k) How many of them were female?
- (l) Append a column `price` to `sec_class` with the total ticket price for each second-class passenger. The total ticket price can be calculated with the formula below:

$$\text{Total} = \text{Pounds} + \frac{\text{Shillings}}{20} + \frac{\text{Pence}}{240} .$$

- (m) How many people in each class died and survived?
- (n) Represent these data as a bar chart (figure 1).

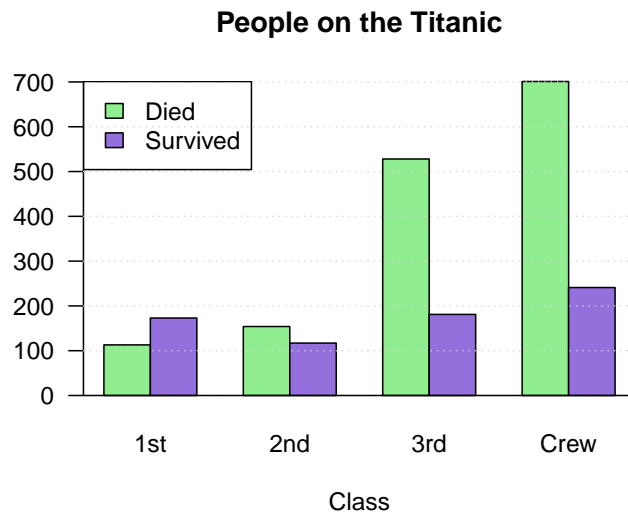


Figure 1: For exercise 1(n), make a bar chart similar to this figure.