

In the following exercises you will work with the diamonds dataset which can be found in ggplot2 library. It contains information about the prices and a series of other attributes for about 54,000 diamonds. You will have to perform basic data manipulation, extract summary statistics, and visualize interesting elements of this dataset, without loading any library.

The variables you will find in diamonds are:

- price: the price in US dollars
- carat: weight of the diamond
- cut: quality of the cut (Fair, Good, Very Good, Premium, Ideal)
- color: diamond color, from D (best) to J (worst)
- clarity: a measurement of how clear the diamond is (I1 (worst), SI2, SI1, VS2, VS1, VVS2, VVS1, IF (best))
- x: length in mm
- y: width in mm
- z: depth in mm
- depth: total depth percentage = $z / \text{mean}(x, y)$
- table: width of top of diamond relative to widest point

You are expected to try all the exercises and within 1 week deliver:

- a R script with the code you used
- a report (as text file) presenting and commenting on your results (plus any graphs you might create)