

## 3.2: Grades and t-shirts- ranked data

*Ranked* (or *ordinal*) data do not come directly from measurements and do not easily correspond to numbers.

For example, quality of mattresses could be estimated with some numbers, from bad (“0”), to excellent (“5”). These assigned numbers are a matter of convenience. They may be anything. However, they maintain a relationship and continuity. If we grade the most comfortable one as “5”, and somewhat less comfortable as “4”, it is possible to imagine what is “4.5”. This is why many methods designed for measurement variables are applicable to ranked data. Still, we recommend to treat results with caution and keep in mind that these grades are arbitrary.

By default, R will identify ranked data as a regular numerical vector. Here are seven employees ranked by their heights:

Object `rr` is the same numerical vector, but numbers “1”, “2” and “3” are not measurements, they are ranks, “places”. For example, “3” means that this person belongs to the tallest group.

Function `cut()` helps to make above three groups automatically:

Result is the *ordered factor* (see below for more explanations). Note that `cut()` is *irreversible* operation, and “numbers” which you receive are not numbers (heights) you start from:

Ranked data *always* require nonparametric methods. If we still want to use parametric methods, we have to obtain the measurement data (which usually means designing the study differently) and also check it for the normality. However, there is a possibility to re-encode ranked data into the measurement. For example, with the appropriate care the color description could be encoded as red, green and blue channel intensity.

Suppose, we examine the average building height in various cities of the world. Straightforward thing to do would be to put names of places under the variable “city” (nominal data). It is, of course, the easiest way, but such variable would be almost useless in statistical analysis. Alternatively, we may encode the cities with letters moving from north to south. This way we obtain the ranked data, open for many nonparametric methods. Finally, we may record geographical coordinates of each city. This we obtain the measurement data, which might be suitable for parametric methods of the analysis.

---

This page titled 3.2: Grades and t-shirts- ranked data is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by Alexey Shipunov via source content that was edited to the style and standards of the LibreTexts platform.