Glossary

**Characteristic carriers** All students who took part in the YOU-PB study are considered characteristic carriers.

**Case** Each individual row in the table corresponds to a student who took part in the survey. This is also called a case. Here, one characteristic carrier corresponds to one case.

**Scope** The number of all cases denotes the scope of the data set, which is often described as **n.**

**Characteristic/variable** The information provided by the respondent is noted in characteristics. In a survey, a characteristic name symbolizes the question behind it. A distinction is made between categorical and numerical characteristics. Characteristics are, for example, gender, age, laptop use, etc.

**Characteristic value/expression** The values that a statistical variable can take on are called characteristic values.

**Absolute frequency** The absolute frequency indicates how often a characteristic occurs in the data set.

**Relative** frequency The relative frequency indicates the proportion of the absolute frequency of a characteristic expression in the total number n of characteristic carriers (whereby only those who have made a statement about a characteristic are taken into account). The relative frequency can be noted as a decimal point (between 0 and 1) or as a percentage (between 0% and 100%).

**Categorical characteristic** In the case of a categorical characteristic, the characteristics (=responses of the respondents) can be divided into specific categories. For example, *gender is* a simple form of a categorical characteristic because it only has two possible characteristics in the data set: male and female. Categorical characteristics can also have many more different values, for example the characteristic laptop\_use has seven values that indicate how often the laptop is used by the characteristic holder: daily, several times a week, once a week, once a fortnight, several times a month, less often, never.

**Percentage** As groups are often compared with each other, it is important to distinguish how large the respective groups are (e.g. number of boys and number of girls). If the groups are of different sizes, comparisons should be made using percentages.

**Summarizing** If a characteristic has several values, it can be helpful to first reduce these to two groups or to summarize several values so that two groups are created. However, there is no right and wrong here. For example, you could put everyone who says they never read online magazines in the "never readers" group and everyone else in the "general readers" group. Alternatively, you could group all those who read online magazines between daily and at least once a week as "frequent readers" and all others as "infrequent readers". Other groupings are also possible. Groupings should always be justified (e.g. in terms of content).

**Numerical characteristic** In the case of a numerical characteristic, the values are numbers (and nothing else!) and can therefore be ordered by size in particular. Examples of numerical characteristics are, for example, height or number\_accounts.

**Range** The difference between the largest value (maximum) and the smallest value (minimum) of a distribution is called the range.

**Maximum** A maximum can only be specified for a numerical characteristic. It denotes the largest occurring value of a distribution.

**Minimum** A minimum can only be specified for a numerical characteristic. It denotes the smallest occurring value of a distribution.

**Arithmetic** mean The average is usually colloquially referred to as the arithmetic mean. The arithmetic mean is calculated from the sum of all values of a characteristic divided by the number of cases.

**Median** As an alternative to the arithmetic mean, you can consider the median, which is the value in the middle of an ordered set of data. Another word for the median is "central value". The median is robust against outliers, whereas the arithmetic mean is not. Especially with skewed distributions, the median is more suitable for describing the "average".