Course of Study Bachelor Computer Science	Exercises Statistics WS 2022/23
Sheet I - Solutions	

1 Descriptive Statistics - Variables

1. Are the following variables qualitative or quantitative?

   (a) Body height
   (b) Hair color
   (c) Temperature in Celsius
   (d) Temperature in Kelvin
   (e) Number of bottles of wine in a student's flat
   (f) Birthday

   Answer:

   (a) Quantitative
   (b) Qualitative
   (c) Quantitative
   (d) Quantitative
   (e) Quantitative
   (f) Quantitative

2. Which scales should be used for the following variables?

   (a) Body height
   (b) Hair color
   (c) Temperature in Celsius
   (d) Temperature in Kelvin
   (e) Number of bottles of wine in a student's flat
   (f) Birthday

   Answer:

   (a) Ratio

(b) Nominal

(c) Interval

(d) Ratio

(e) Ratio

(f) Interval

3. It is possible to transform a variable "downwards", from a scale with more information contained, to a scale with less information contained. Give an example for the variable *Price for a bottle of wine* for the transformation from a ratio to an ordinal scale.

   Answer:

   - Ratio scale: Price measured in euro.
   - Ordinal scale: Price mesured as cheap, normal, expensive.

4. Is it possible to transform a variable "upwards", from a scale with less information contained, to a scale with more information contained? Give an example (showing if it is possible or not)!

   Answer: No, you cannot gain more information about a variable by transforming the variable from one scale to another. Example: Assume that the variable *Price for a bottle of wine* originally has an ordinal scale. This means that you know for each bottle if it is cheap, normal or expensive. Given **only** this information for each bottle, do you know the price in euro (ratio scale) for the bottles? No!

5. Consider the question of describing students attitudes towards to legalisation of Marihuana, what proportion of them wants to legalize the drug and whether this proportion differs by gender and age.

   (a) Which data collection method is most suitable here: survey or experiment?

   (b) How could you capture the attitudes towards legalisation in a single variable?

   (c) Which variables are needed to answer the questions? Describe the type and the scale of the variables.

   (d) How would an appropriate data set look? Try to describe the question in more details.

   Answer:

(a) Survey: The information would be obtained via a questionaire given to a sample of students.

(b) There are different options to ask the students attitudes:

- simply ask: "What do you think about legalisation?"
  Problem: Capturing long answers in a variable attitutde may make it difficult to summarize and distil the information obtained.
- Common way: translate it into a score
  One could for example ask 5 "yes/no" questions which relate to attitudes towards legalisation like "Do you believe that legalisation would endanger the health of young people?", "Do you think legalization would encourage the entry into harder drugs?", ... The number of answers showing a positive attitude can be summed up. Thus the answers of each student can be summarized on a scale from 0 to 5.

(c) Needed variables are:

- Attitude: quantitative variable, ordinal scale
- Legalise: binary ("yes/no") variable capturing wether the student agrees to legalize Marihuana. This is qualititaive variable with nominal scale.
- Gender: qualititaive variable with nominal scale
- Age: quantitative (continous) variable with ratio scale.

(d) A data set might look as:

Student	A1	...	A5	Attitude	Legalize	Gender	Age
1	yes	...	no	3	yes	male	22
2	no	...	yes	2	no	female	25
⋮	⋮	...	⋮	⋮	⋮	⋮	⋮

A1, ..., A5 refer to variables capturing attitudes towards legalisiation and "Legalise" is the score variable summarizing these questions.
More detailled questions:

- What is the average attitude towards legalisation among students and how much does ist vary?
- What percentage of students answer "yes" when asked to legalise Marihuana?
- What is difference in the proportion calculated above when stratified by gender?

- What is average of those students who support the legalisation compared with the average age of those students who do not support the legalisation?