Possible tasks for reviewing the learning content of the "Data detectives at work" lesson series

This document contains possible solutions to the tasks. Below the tasks you will find information for a possible evaluation and an estimated time required to solve the task.

In addition, a subjective assessment of the level of difficulty is given (easy - medium - difficult).

The tasks are intended as a "quarry". You can take tasks from here and put them together as you wish for a class assignment or a test or let the different formats inspire you to create your own tasks. **If a test is planned for a school lesson (45 minutes), you should never use all the tasks!**

Contents

[Task 1: Matching (easy) 2](#_Toc70076003)

[Task 2: Row and column percentages (medium) 3](#_Toc70076004)

[Task 3: Reading multi-field tables (difficult) 6](#_Toc70076005)

[Task 4: Variables (easy) 8](#_Toc70076006)

[Task 5: Distribution comparison (medium) 9](#_Toc70076007)

[Task 6: The data analysis cycle (medium) 10](#_Toc70076008)

### Task 1: Matching (easy)

1. Match the following terms correctly and connect them with dashes!

|  |  |
| --- | --- |
| Numerical characteristic  Categorical characteristic  Characteristic value | Smartphone\_Possession  Hair color  Age  blond  15 years  Gender  Male |

Evaluation suggestion 1a: 0.5 points per correct assignment, max. 3 points in total

Time required: 2 minutes

1. Name another numerical characteristic: E.g. size  
   And two possible characteristics: E.g. 156 cm (1)  
    E.g. 134 cm (2)
2. Name another categorical characteristic: E.g. eye color  
   And two possible characteristics: E.g. blue (1)  
    E.g. brown (2)

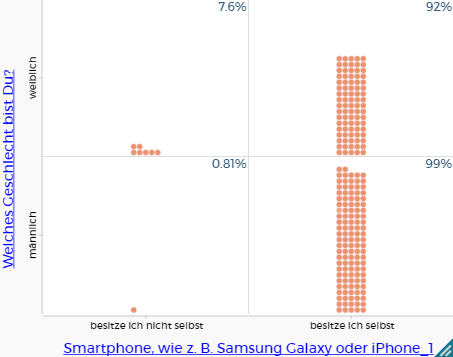
**Evaluation suggestion** 1b: 1 point per correct characteristic, 0.5 points per correct characteristic, max. 4 points

|  |  |
| --- | --- |
| Task | Max. Points |
| 1a | 3 |
| 1b | 2 |
| 1c | 2 |

**Time required**: 5 minutes

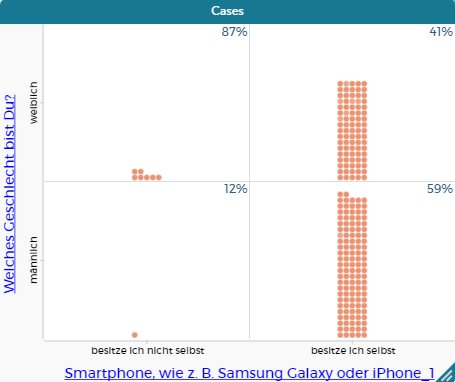
### Task 2: Row and column percentages (medium)

1. Fill in the blanks using the diagram.

92% of the girls surveyed own their own smartphone.

99% of the boys surveyed own their own smartphone.

This means that in this data set, boys (girls or boys?) are more likely to have their own smartphone than girls  
 (girls or boys?).

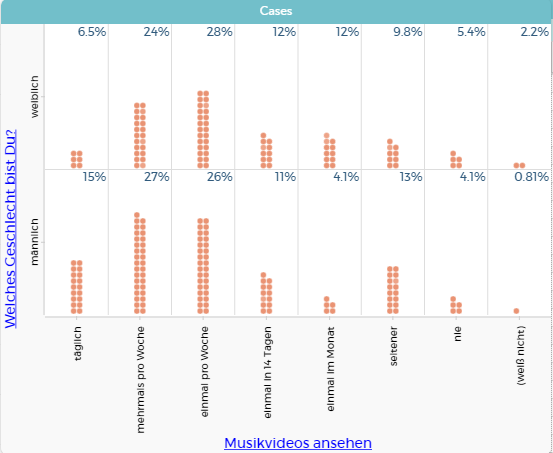
1. Fill in the blanks using the diagram.

41% of smartphone owners are girls.

59% of smartphone owners are boys.

1. Why can it happen that, for example, in the graph in b) the percentages in the first column do not add up to 100%?

The result is 99%. This can happen due to the display in CODAP because the values are rounded. If both figures are rounded down, something is missing when adding them up.

1.   
   Overall, 30.5% of the girls surveyed watch music videos daily or several times a week.

28% of girls watch music videos once a week (how often?).

17.4% of the girls surveyed watch music videos less than once a month.

What is the situation among boys? Overall, 42% of the boys surveyed watch music videos daily or several times a week.

11% of the boys surveyed watch music videos once every 14 days (how often?).

17.91% of the boys surveyed watch music videos less than once a month.

1. Fill in and justify: This means that in this data set, boys (girls or boys?) tend to watch music videos more often than girls (girls or boys?).  
   Justification: *ATTENTION: Several justifications are possible here!*

7.6% of girls say they never (or don't know) watch music videos, but only 4.91% of boys say the same. Therefore, overall, boys watch music videos more often than girls. 42% of boys say they watch music videos at least several times a week, but only 30.5% of girls.

**Assessment suggestion** 2a, 2b, 2d): 0.5 points per correctly completed gap.

2c: Correct reasoning results in 1 point

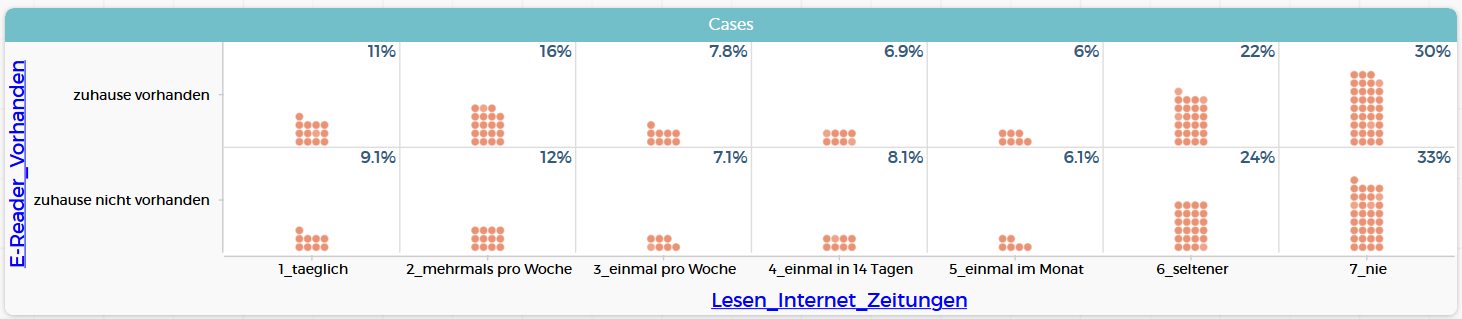
2e: 0.5 points per correctly completed gap, appropriate explanation: Corresponding proportion of boys is described: 1 point, corresponding proportion of girls is described: 1 point

Max 10 points

|  |  |
| --- | --- |
| Task | Max. Points |
| 2a | 2 |
| 2b | 1 |
| 2c | 1 |
| 2d | 3 |
| 2e | 3 |

**Time required**: 10 minutes

### Task 3: Reading multi-field tables (difficult)



1. Look at the graph. For the following statements, mark whether they are true, false or cannot be assessed on the basis of the data.

|  |  |  |  |
| --- | --- | --- | --- |
| **Statement** | **true** | **wrong** | **cannot be judged** |
| Those who have an e-reader at home read newspapers online more frequently (i.e. at least once a week) than those who do not have an e-reader at home. | X |  |  |
| 11% of those who have an e-reader at home say they read newspapers online every day. | X |  |  |
| 33% of those who say they never read newspapers online do not have an e-reader at home. |  | X |  |
| Of all respondents, 20.1% read newspapers on the Internet every day. |  | X |  |
| Those who say they read newspapers online once a week also read printed newspapers at least once a week. |  |  | X |
| 6% of all respondents read newspapers online once a month and have an e-reader at home. |  | X |  |

1. Formulate a statistically correct statement for the diagram above to explain the value 16%.

Of those who have an e-reader at home, 16% say they read newspapers online several times a week.

*Attention: Pay close attention to the wording here. Reference must be the characteristic "E-Reader\_Existing".*

1. Let's look at the "middle group", i.e. those who read newspapers online at least once a month and at most once a week. What percentage of those who have an e-reader at home belong to this group? Calculate.

6% + 7,8% + 6,9% = 20,7%

**Assessment suggestion** 3a: 1 point per correctly placed cross (justified by the difficulty of the tasks)

3b: Correct formulation results in 2 points.

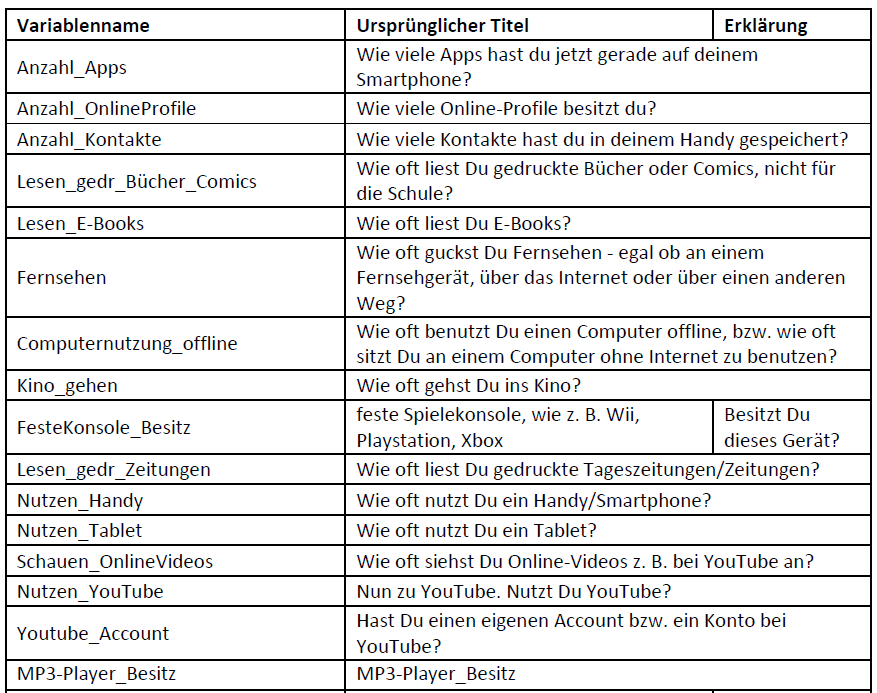
3c: 1 point

Max 9 points

|  |  |
| --- | --- |
| Task | Max. Points |
| 3a | 6 |
| 3b | 2 |
| 3c | 1 |

**Time required**: 15 minutes

### Task 4: Variables (easy)



Here you can see a section of the variable list for a data record.

Which variables do you need to answer the following questions? Please note that some questions require more than one variable. Write down the variable names that are needed to answer questions a)-d).

1. Are those who use YouTube more likely to have more apps on their smartphone or those who do not use YouTube?   
   Use\_YouTube, number\_of\_apps
2. Do those who have an MP3 player tend to go to the movies?  
   go\_to\_the\_cinema, MP3\_player\_ownership
3. What percentage use their cell phone or smartphone at least once a week?  
   Use\_mobile\_phone
4. Are there more respondents who have a fixed games console and frequently use a tablet or more respondents who do not have a fixed games console and frequently use a tablet?  
   Fixed\_console\_ownership, use\_tablet

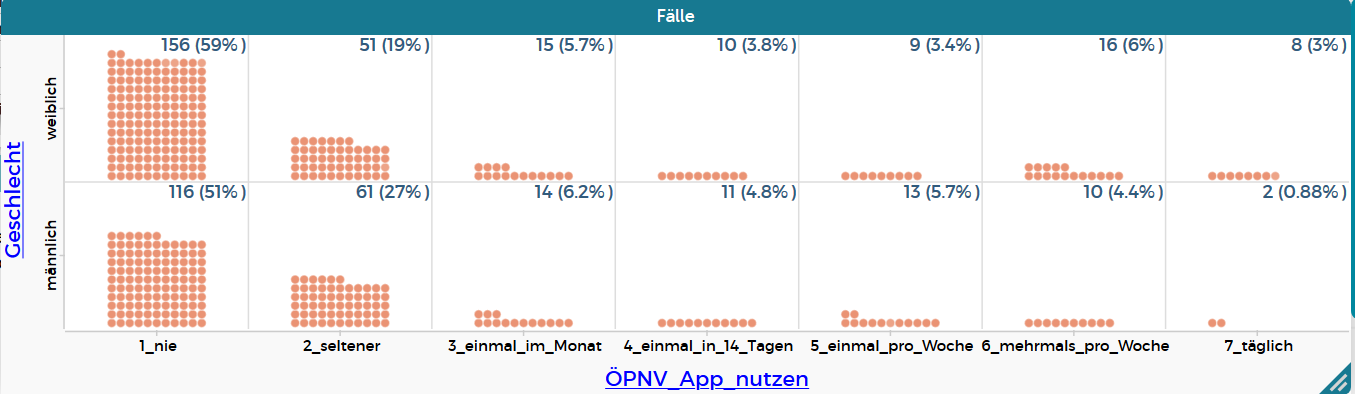
**Evaluation suggestion** 0.5 points per correct feature

Max 3.5 points

|  |  |
| --- | --- |
| Task | Max. Points |
| 4a | 1 |
| 4b | 1 |
| 4c | 0,5 |
| 4d | 1 |

**Time required**: 7 minutes

### Task 5: Distribution comparison (medium)



Tom looks at the two distributions and claims: "Girls use a public transport app more often". Comment on this statement and formulate more precisely than Tom how girls and boys differ in their use of a public transport app. Identify two user groups and use the corresponding values for your explanation.

The statement is absolutely false (*rating: 1 point).* There are boys who use the app more often than girls. The statement tends to be true.

Two user groups can be identified: The frequent users and the infrequent users (*rating: 1 point)*. Frequent users are those who use the public transport app daily, several times a week, once a week or once every 14 days (*rating: 1 point*). The infrequent users are those who use the app once a month, less often or never (*rating: 1 point*).

Among the frequent users, 8+16+9+10=43 girls (*or corresponding percentages, rating: 1 point*) and 2+10+13+11=36 boys, i.e. 7 more girls than boys (*rating: 1 point*). The statement is correct here.

However, there are also more girls than boys among the low users (rating: 1 point):   
Boys as low users: 116+61+14=191   
Girls as low users: 156+51+15=222 (*rating: 1 point*)  
The statement is therefore not correct for this subgroup.

*ATTENTION: When making comparative statements, make sure that a comparison is really being made (more than, less than, etc.). A mere statement of values (those have this much, the others have this much) is not a comparison and should lead to points being deducted.*

**Valuation proposal** See text

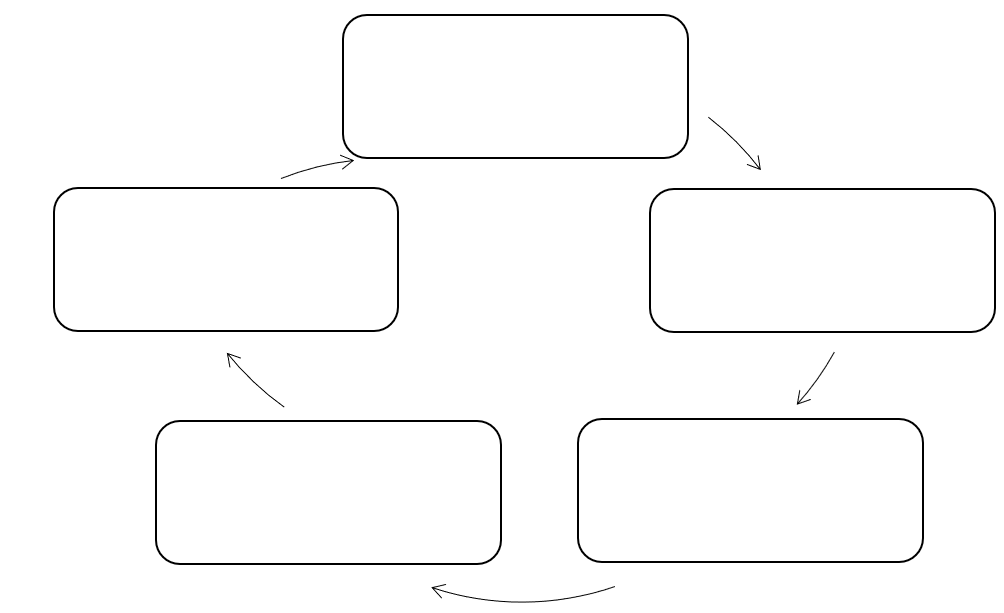
Max 8 points

**Time required**: 20 minutes

### Task 6: The data analysis cycle (medium)

1. Enter the following terms that belong to a data analysis project in a meaningful order in the diagram!

Analysis, data, conclusion, problem, plan



Analysis

Data

Plan

Problem

Conclusion

1. Briefly describe what you do in the "Analysis" phase.

In the "Analysis" phase, data is analyzed, e.g. using visualizations or calculations. calculations.

1. Briefly describe what you do in the "Problem" phase.

In the Problem phase, for example, the questions that you want to investigate are set up.

1. Name two aspects of a good statistical question.

E.g. (further possible answers can be found in session 4, slide 21)

- Yes/no answers are avoided

- The question can be answered with the given data

**Evaluation suggestion** 6a: 0.5 points per correct entry

6d) 1 point per correct entry

Max 6 points

|  |  |
| --- | --- |
| Task | Max. Points |
| 6a | 2 |
| 6b | 1 |
| 6c | 1 |
| 6d | 2 |

**Time required**: 10 minutes