✓ **Herzlichen Glückwunsch! Sie haben bestanden!**
ZUM BESTEHEN 80 % oder höher

[Lernen Sie weiter]

BEWERTUNG
90 %

Course challenge

NEUESTE EINREICHUNGSBEWERTUNG
90%

---

1. **Scenario 1, questions 1-5** `1 / 1 Punkten`

   You are a data analyst at a small analytics company. Your company is hosting a project kick-off meeting with a new client, Meer-Kitty Interior Design. The agenda includes reviewing their goals for the year, answering any questions, and discussing their available data.

   Before the meeting you review the About Us tab on their website and their business plan, linked below:

   | 📄 **Meer-Kitty Interior Design About Us Page.pdf** PDF-Datei | ↗ |

   | 📄 **Meer-Kitty Interior Design Business Plan.pdf** PDF-Datei | ↗ |

   Meer-Kitty Interior Design has two goals. They want to expand their online audience, which means getting their company and brand known by as many people as possible. They also want to launch a line of high-quality indoor paint to be sold in-store and online. You decide to consider the data about indoor paint first.

   You may click the link to create a copy of the dataset: Kitty Survey Feedback. Please refer to the **Meer-Kitty survey feedback** tab.

   Or, if you don't have a Google account, download the template directly from the attachment below.

   | Kitty Survey Feedback - Meer-Kitty survey feedback.csv |

   You are pleased to find that the available data is aligned to the business objective. However, you do some research about confidence level for this type of survey and learn that you need at least 120 unique responses for the survey results to be useful. Therefore, the dataset has two limitations: First, there are only 40 responses; second, a Meer-Kitty superfan, User 588, completed the survey 11 times.

   As the survey has too few responses and numerous duplicates that are skewing results, what are your options? Select all that apply.

   ☐ Remove the duplicates from the data and proceed with analysis.

   ☐ Locate another dataset about indoor paint.

   ☑ Talk with stakeholders and ask for more time.

   > ✓ **Richtig**
   > With numerous duplicates, the best option is to talk with stakeholders and ask for more time. Then, you can repeat the survey in order to create a new, improved dataset.

   ☑ Repeat the survey in order to create a new, improved dataset.

   > ✓ **Richtig**
   > With numerous duplicates, the best option is to talk with stakeholders and ask for more time. Then, you can repeat the survey in order to create a new, improved dataset.

2. **Scenario 1 continued** `1 / 1 Punkten`

   During the meeting, you also learn that Meer-Kitty videos are hosted on their website. For each product offered, there is an accompanying video for customers to learn more. So, more views for a video suggests greater consumer interest.

   Your goal is to identify which videos are most popular, so Meer-Kitty knows what topics to explore in the future. Unfortunately, Meer-Kitty has just three months of data available because they only recently launched the videos on their site.

   Without enough data to identify long-term trends about the video subjects that people prefer, what should you do?

   ○ Tell the client you're sorry, but there is no way to meet their objective.

   ○ Watch the videos and use your gut instinct to identify which are most successful.

   ● Find an alternate data source that will still enable you to meet your objective.

   ○ Move ahead with the data you have to determine the top video subjects.

   > ✓ **Richtig**
   > Without enough data to identify long-term trends, one option is to find an alternate data source that will still enable you to meet your objective. In this case, you could find data from a similar company and learn about its consumer interest and trends.

3. **Scenario 1 continued** `0 / 1 Punkten`

   Now that you've identified some limitations with Meer-Kitty's data, you want to communicate your concerns to stakeholders. In addition to insufficient video trend data, your main concern with the indoor paint survey is that the data isn't representative of the population as a whole.

   Clearly, one particular respondent, the superfan, is overrepresented. This is an example of margin of error.

   ● True

   ○ False

   > ❗ **Falsch**
   > Review the section on sample probability for a refresher.

4. **Scenario 1 continued**   `1 / 1 Punkten`

The stakeholders understand your concerns and agree to repeat the indoor paint survey. In a few weeks, you have a much better dataset with more than 150 responses and no duplicates.

You may click the link to create a copy of the dataset: Kitty Survey Feedback. Please refer to the **New Meer-Kitty survey feedback** tab.

Or, if you don't have a Google account, download the template directly from the attachment below.

Kitty Survey Feedback - New Meer-Kitty survey feedback.csv

You notice that questions 4 and 5 are dependent on the respondent's answer to question 3. So, you need to determine how many people answered Yes to question 3, then compare that to responses to questions 4 and 5. That way, you will know if questions 4 and 5 have any nulls.

You decide to use a spreadsheet tool that changes how cells appear when they meet a certain value — in this case, the word Yes. You are using VLOOKUP.

○ True

◉ False

> ✓ **Richtig**
> To change how cells appear when they meet a certain value, use conditional formatting.

---

5. **Scenario 1, continued**   `1 / 1 Punkten`

You continue cleaning the data. You use tools such as remove duplicates and COUNTIF to ensure the dataset is complete, correct, and relevant to the problem you're trying to solve. Then, you complete the verification and reporting processes to share the details of your data-cleaning effort with your team.

While reviewing, your team notes one aspect of data cleaning that would improve the dataset even more. They point out that the new survey also has a new question in Column G: "What are your favorite indoor paint colors?" This was a free-response question, so respondents typed in their answers. Some people included multiple different colors of paint. In order to determine which colors are most popular, it will be necessary to put each color in its own cell.

What spreadsheet function enables you to put each of the colors in Column G into a new, separate cell?

○ Delimit

○ MID

○ Divide

◉ SPLIT

> ✓ **Richtig**
> To put each of the colors in Column G into a new, separate cell, use SPLIT. SPLIT is a spreadsheet function that divides text around a specified character and puts each fragment into a new, separate cell.

---

6. **Scenario 2, questions 6-10**   `1 / 1 Punkten`

You've completed this program and are interviewing for a junior data scientist position. The job is at B.Spoke Market Research, a company that analyzes market conditions using customer surveys and other research methods. The detailed job description can be found below:

📄 **C4 B.Spoke Market Research Job Description.pdf**
PDF-Datei

So far, you've had a phone interview with a recruiter and you've secured a second interview with the B.Spoke team. The recruiter's email can be found below:

📄 **C4 S2 Email from Recruiter.pdf**
PDF-Datei

You arrive 15 minutes early for your interview. Soon, you are escorted into a conference room, where you meet Jodie Choi, the data science lead. After welcoming you, the behavioral interview begins.

For your first question, your interviewer wants to learn about your experience with spreadsheets. She says: Sometimes the team needs data that is stored in different spreadsheets. So, we use a spreadsheet function to find the information we need.

There is a spreadsheet function that allows a data analyst to search for a value in the first column of a given range and return the value of a specified cell in the row in which it is found. What function allows you to complete these tasks?

○ COUNTIF

○ RETURN

○ SEARCH

◉ VLOOKUP

> ✓ **Richtig**
> VLOOKUP searches for a value in the first column of a given range and returns the value of a specified cell in the row in which it is found.

---

7. **Scenario 2, continued**   `1 / 1 Punkten`

Next, your interviewer wants to know more about your understanding of tools that work in both spreadsheets and SQL queries. She explains that the data her team receives from customer surveys sometimes has many duplicate entries.

She says: Spreadsheets have a great tool for that called remove duplicates. But when writing a SQL query, what command should you include in your SELECT statement to remove duplicates?

○ DIVERSE

○ DIFFERENT

○ DISCRETE

◉ DISTINCT

> ✓ **Richtig**
> To remove duplicates in a SQL query, include DISTINCT in your SELECT statement.

8. **Scenario 2, continued**

Now, your interviewer explains that the data team usually works with very large amounts of customer survey data. After receiving the data, they import it into a SQL table. But sometimes, the new dataset imports incorrectly and they need to change the format.

She asks: What function would you use to convert data in a SQL table from one datatype to another?

◉ CAST

◯ CONVERT

◯ CHANGE

◯ COALESCE

> ✓ **Richtig**
> The CAST function is used to convert data in a SQL table from one datatype to another.

9. **Scenario 2, continued**

Next, your interviewer explains that one of their clients is an online retailer that has a vast inventory. She has a list of items by name, color, and size. Then, she has another list of the price of each item by size, as a larger item sometimes costs more. The stakeholder needs one list of all items by name, color, size, and price.

She then says: Could you use the CONCAT function to add strings together to create new text strings?

◉ Yes

◯ No

> ✓ **Richtig**
> To add strings together to create new text strings, use the CONCAT function.

10. **Scenario 2, continued**

For your final question, your interviewer explains that her team often comes across data with extra spaces.

She asks: Which SQL function enables you to eliminate those extra spaces for consistency?

◯ LENGTH

◯ SUBSTR

◯ LEN

◉ TRIM

> ✓ **Richtig**
> To eliminate extra spaces for consistency, use the TRIM function.