

✓ **Congratulations! You passed!**

Grade  
received **100%**

Latest Submission  
Grade **100%**

To pass 80% or  
higher

[Go to next item](#)

**1. Scenario 1, questions 1-5**

1 / 1 point

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 File



**Meer-Kitty Interior Design Business Plan.pdf**  
PDF File

**As the survey has too few responses and numerous duplicates that are skewing results, you decide to repeat the survey in order to create a new, improved dataset. What is your first step?**

- ☐ Delete all of the data from the current, skewed survey.
- ☒ Talk with stakeholders, explain the new timeline, and ask for approval.
- ☐ Write new, improved survey questions.
- ☐ Find a survey tool that only allows someone to complete the survey once.



**Correct**

Before repeating the survey, it's necessary to talk with stakeholders, explain the new timeline, and ask for approval.

**2. Scenario 1 continued**

1 / 1 point

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.
- ☒ 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.
- ☐ Watch the videos and use your gut instinct to identify which are most successful.



**Correct**

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

1 / 1 point

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. What does this situation describe?**

- ☒ Sampling bias
- ☐ Confidence level
- ☐ Margin of error
- ☐ Statistical significance

✓ Correct

This situation describes sampling bias. Sampling bias occurs when a sample isn't representative of the population as a whole.

4. Scenario 1 continued

1 / 1 point

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.

To use the template for the survey feedback, click the link below and select "Use Template."

Link to template: [Kitty Survey Feedback](#)

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



**Kitty Survey Feedback - New Meer-Kitty survey feedback**

CSV File

If you are using the template, please refer to the **New Meer-Kitty survey feedback** tab. 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 contain the word Yes. Which tool do you use?**

- ☒ Conditional formatting
- ☐ Filtering
- ☐ Data validation
- ☐ CONCATENATE

✓ Correct

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

#### 5. Scenario 1, continued

1 / 1 point

You have finished cleaning the data to ensure it 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.

Your team notes one aspect of data cleaning that would help improve the dataset. 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.

**You use a spreadsheet function to divide the text strings in Column G around the commas and put each fragment into a new, separate cell. In this example, what are the commas called?**

- ☐ Substrings
- ☐ Partitions
- ☒ Delimiters
- ☐ MIDs



**Correct**

The commas are delimiters, which are characters that indicate the beginning or end of a data item.

#### 6. Scenario 2, questions 6-10

1 / 1 point

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 File

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 File

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.

**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?**

- ☐ SEARCH
- ☐ COUNTIF
- ☒ VLOOKUP
- ☐ RETURN



**Correct**

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 point

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?**

- ☒ DISTINCT
- ☐ DIFFERENT
- ☐ DISCRETE
- ☐ DIVERSE

✓ Correct

To remove duplicates in a SQL query, include DISTINCT in your SELECT statement.

8. Scenario 2, continued

1 / 1 point

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: Is there a command or function that converts data in a SQL table from one datatype to another? You respond: Yes, it's the CAST function.**

- ☒ True
- ☐ False

✓ Correct

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

9. Scenario 2, continued

1 / 1 point

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 client needs one list of all items by name, color, size, and price.

**She then asks: If you were to use the CONCAT function to complete this task, what would it enable you to do?**

- ☒ Create a unique key to tell products apart
- ☐ Search for and return missing products in inventory
- ☐ Create a new product database table
- ☐ Clean the product identifier text strings

✓ Correct

Using the CONCAT function to combine each string into a single text string would enable you to create a unique key. You can use the key to tell products apart and count them more easily.

10. Scenario 2, continued

1 / 1 point

For your final question, your interviewer explains that her team often uses the TRIM function when writing SQL queries.

**She asks: What is the TRIM function used for in SQL?**

- ☐ To return the smallest numeric value from a list
- ☐ To shorten the list of results
- ☒ To eliminate extra leading or trailing spaces
- ☐ To eliminate null values

✓ Correct

The TRIM function is used to eliminate extra leading or trailing spaces.