

Course challenge

- Reading: Glossary: Terms and definitions
- View: Get ready for the course challenge
- Quiz: Course challenge: Questions
- View: Congratulations!
- Reading: Coming up next...

0 Congratulations! You passed!

Grade received **Course challenge** To pass 80% or higher Go to next item

Review Learning Objectives

1. Scenario 1, questions 1-5 6.6 / 10 point

You are a data analyst at a retail analytics company. Your company is hosting a project kick-off meeting with a new client, **Heer-Kitty Interiors**. The client includes reviewing their goals for the year, assessing any questions, and creating the feedback database. (over 24 hours)

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

**Heer-Kitty Interiors Design About Us Page.pdf**  
PDF File

**Heer-Kitty Interiors Design Business Plan.pdf**  
PDF File

Heer-Kitty Interiors Design has two goals. They want to expand their online audience, which means getting their content and brand known by as many people as possible. They also want to launch a line of high-quality indoor plants to be sold in stores and online. You decide to consider the client about indoor gardening.

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

Link to template: [Heer-Kitty Interiors Feedback](#)

OR

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

**Kitty Survey Feedback - Heer-Kitty survey feedback**  
CSV File

When you refer to the **Heer-Kitty survey feedback** file, you are pleased to find that the available data is aligned to the business objective. However, you do come across 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 90 responses, second, a Heer-Kitty specific, user ID, has completed the survey 21 times.

**As the survey has two data responses and numerous duplicates that are slowing results, what are your options? Select all that apply.**

- ☒ **Talk with stakeholders and ask for more time.**
- ☐ **Cancel**  
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.**
- ☐ **Cancel**  
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.
- ☒ **Locate another dataset about indoor plants.**
- ☐ **This should not be selected**  
Review [the question and answer data in business objectives](#). for a refresher.
- ☒ **Remove the duplicates from the data and proceed with analysis.**
- ☐ **This should not be selected**  
Review [the question and answer data in business objectives](#). for a refresher.

2. Scenario 1 continued 6.6 / 10 point

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

Your goal is to identify which videos are most popular on Heer-Kitty because what topics to explore in the future. Unfortunately, Heer-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 are your available options? Select all that apply.**

- ☒ **Find an alternate data source that will still enable you to meet your objective.**
- ☐ **Cancel**  
Without enough data to identify long-term trends, one option is to talk with stakeholders and ask to adjust the objective. You could also ask to wait for more data and provide an updated timeline. Or you could find an alternate data source that will still enable you to meet your objective.
- ☒ **Talk with Heer-Kitty stakeholders and ask to adjust the objective.**
- ☐ **Cancel**  
Without enough data to identify long-term trends, one option is to talk with stakeholders and ask to adjust the objective. You could also ask to wait for more data and provide an updated timeline. Or you could find an alternate data source that will still enable you to meet your objective.
- ☒ **Ask to wait for more data and provide Heer-Kitty with an updated timeline.**
- ☐ **Cancel**  
Without enough data to identify long-term trends, one option is to talk with stakeholders and ask to adjust the objective. You could also ask to wait for more data and provide an updated timeline. Or you could 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.**
- ☐ **This should not be selected**  
Review [the question and answer data in business objectives](#). for a refresher.

3. Scenario 1 continued 6.6 / 10 point

After the your identified some limitations with Heer-Kitty's data, you want to communicate your concerns to stakeholders. In addition to confirming video trends, your main concern with the indoor plant survey is that the dataset's representation of the population as a whole.

Clearly, one particular respondent, the supervisor, is overrepresented. This means the data doesn't represent the population as a whole.

**When reviewing sample for Heer-Kitty in the future, what are some best practices you can use to address some of the issues associated with sampling bias? Select all that apply.**

- ☒ **Use data from only one source**
- ☐ **This should not be selected**  
Review [the question and answer data in business objectives](#). for a refresher.
- ☒ **Use data that keeps updating.**
- ☐ **This should not be selected**  
Review [the question and answer data in business objectives](#). for a refresher.
- ☒ **Increase sample size**
- ☐ **Cancel**  
To address some of the issues associated with sampling bias, random sampling can help. With random sampling, analysts can select a sample from a population so that every possible type of the sample has an equal chance of being chosen. In addition, by increasing sample size, you're more likely to survey part of a population that is representative of the whole.
- ☒ **Use random sampling**
- ☐ **Cancel**  
To address some of the issues associated with sampling bias, random sampling helps select a sample from a population so that every possible type of the sample has an equal chance of being chosen. In addition, by increasing sample size, you're more likely to survey part of a population that is representative of the whole.

4. Scenario 1 continued 6 / 10 point

The stakeholders understood your concerns and agree to repeat the indoor plant survey. In a few weeks, you have a much better dataset with more than 120 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 Heer-Kitty survey feedback**  
CSV File

If you are using the template, please refer to the **New Heer-Kitty survey feedback** tab located at the bottom of the page. You notice that questions 4 and 5 are duplicated in 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. Therefore, you will need a spreadsheet function to divide the total strings in Column C around the commas and put each fragment into a new, separate cell. You are using the **SPLIT** function.

- You decide to use a spreadsheet function to divide the total strings in Column C around the commas and put each fragment into a new, separate cell. You are using the SPLIT function.**
- ☐ **The value is a CONCATENATE statement**
- ☒ **The value is a conditional formatting rule**
- ☐ **The value is the COUNTA range**
- ☐ **The value is a CONCATENATE range**
- ☐ **Cancel**  
To change how cells appear when they meet a certain value, use conditional formatting.

5. Scenario 1, continued 6 / 10 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 views your aspect of data cleaning that would help improve the dataset. They point out that the new survey data has a new question in column 5: "What are your favorite indoor plants?" They have three response questions, so respondents typed in their answers. Some people included multiple different colors of plants to determine which colors are most popular. In the responses to each color, you have a new cell.

**You decide to use a spreadsheet function to divide the text strings in Column C around the commas and put each fragment into a new, separate cell. You are using the SPLIT function.**

- ☒ **Yes**
- ☐ **No**
- ☐ **Cancel**  
To divide the text strings in Column C around the commas and put each fragment into a new, separate cell, you 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 1-10 6 / 10 point

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

**CR 8 Spalte 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. Spalte team. The recruiter's email can be found below:

**CR 11 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 Julie Chen. The discussion lasts 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?**

- ☐ **INDEX**
- ☐ **RETURN**
- ☒ **VLOOKUP**
- ☐ **COUNTIF**
- ☐ **Cancel**  
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 6 / 10 point

Next, your interviewer explains that one of their clients is a coffee retailer that has a need 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 to duplicate each as this one, could you use the CONCAT function to add strings together to create one new string?**

- ☐ **Yes**
- ☐ **No**
- ☐ **Cancel**  
To remove duplicates in a SQL query, the data team can include DISTINCT in the SELECT statement.

8. Scenario 2, continued 6 / 10 point

Now, your interviewer explains that the data team usually works with very large amounts of customer survey data, after reviewing the data, they import it into a SQL table. But sometimes, the new dataset imports incorrectly and you're left with the data error.

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

- ☐ **CONJUNCT**
- ☐ **CONJUNCT**
- ☐ **CONJUNCT**
- ☒ **CAST**
- ☐ **Cancel**  
The CAST function is used to convert data in a SQL table from one datatype to another.

9. Scenario 2, continued 6 / 10 point

Next, your interviewer explains that one of their clients is a coffee retailer that has a need 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 to duplicate each as this one, could you use the CONCAT function to add strings together to create one new string?**

- ☒ **Yes**
- ☐ **No**
- ☐ **Cancel**  
To add strings together to create new text strings, use the CONCAT function.

10. Scenario 2, continued 6 / 10 point

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

**She asks: Which function would enable you to eliminate those extra spaces? You respond: To eliminate extra spaces for consistency, use the TRIM function.**

- ☒ **Yes**
- ☐ **No**
- ☐ **Cancel**  
To eliminate extra spaces for consistency, use the TRIM function.