Course challenge

LATEST SUBMISSION GRADE

90%

---

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

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.

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 - Meer-Kitty survey feedback.csv

When you refer to the **Meer-Kitty survey feedback** tab, 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, 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.

○ Find a survey tool that only allows someone to complete the survey once.

○ Write new, improved survey questions.

◉ Talk with stakeholders, explain the new timeline, and ask for approval.

> ✓ **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?

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

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

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

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

> ✓ **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. This means the data doesn't represent the population as a whole.

When surveying people for Meer-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.

☑ Increase sample size

> ✓ **Correct**
>
> 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 data from only one source

☐ Use data that keeps updating

☑ Use random sampling

✓ **Correct**

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**

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

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. When using this tool, what is the word Yes?

○ The value in a CONCATENATE range

○ The value in the COUNTA range

○ The value in a VLOOKUP statement

⦿ The value in a conditional formatting rule

✓ **Correct**

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

5.

1 / 1 point

Scenario 1, continued

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

◉ True

○ False

✓ **Correct**

To divide the text strings in Column G 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 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.

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?

◉ VLOOKUP

○ RETURN

○ COUNTIF

○ SEARCH

✓ **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

○ DIVERSE

○ DIFFERENT

○ DISCRETE

✓ **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 SQL function that can convert data types such as currency, dates, and times in a SQL table?

◉ Yes, data types including currency, dates, and times can be converted.

○ No, only currency can be converted.

✓ **Correct**
The CAST function is used to convert currency, dates, and times in a SQL table from one datatype to another.

---

9. **Scenario 2, continued**   `0 / 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 stakeholder needs one list of all items by name, color, size, and price.

She then says: In situations such as this one, could you use the CONCAT function to add strings together to create new text strings?

○ Yes

◉ No

> ✕ **Incorrect**
>
> Review <u>the video on SQL data cleaning functions</u> for a refresher.

10. **Scenario 2, continued**

`1 / 1 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.

◉ True

○ False

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