Data analyst job opportunities
The importance of fair business decisions
Optional: Exploring your next job
Weekly challenge 5
Course challenge
Reading: Test-taking strategies 10 min
Quiz: *Course challenge* 10 questions
Video: Congrats! Course wrap-up 3 min
Reading: Coming up next... 10 min

Course challenge

Quiz • 50 min

Congratulations! You passed!

Grade received 97.50%
Latest Submission Grade 97.50%
To pass 80% or higher

Go to next item

Review Learning Objectives

1. Scenario 1, question 1-5 — 1 / 1 point

Submit your assignment
Due Jul 9, 11:59 PM +08 Attempts 3 every 24 hours

Receive grade
To Pass 80% or higher

Your grade 97.50%

View Feedback
We keep your highest score

Try again

You've just started a new job as a data analyst for a midsized pharmacy chain with 38 stores in the American Southwest. Your supervisor shares a new data analysis project with you.

She explains that the pharmacy is considering discontinuing a bubble bath product called Splashtastic. Your supervisor wants you to analyze sales data and determine what percentage of each store's total daily sales come from that product. Then, you'll present your findings to leadership.

You know that it's important to follow each step of the data analysis process: ask, prepare, process, analyze, share, and act. So, you begin by defining the problem and making sure you fully understand stakeholder expectations.

One of the questions you ask is where to find the dataset you'll be working with. Your supervisor explains that the company database has all the information you need.

Time Details Report an issue

Next, you continue to the prepare step. You access the database and write a query to retrieve data about Splashtastic. You notice that there are only 38 rows of data, representing the company's 38 stores. In addition, your dataset contains five columns: Store Number, Average Daily Customers, Average Daily Splashtastic Sales (Units), Average Daily Splashtastic Sales (Dollars), and Average Total Daily Sales (All Products). You decide to use a spreadsheet to work with the data because you know that spreadsheets work well for processing and analyzing a small dataset, like the one you're using.

Fill in the blank: To get the data from the database into a spreadsheet, you would first _____ the data as a .CSV file, then import it into a spreadsheet.

- (●) download
- ( ) print
- ( ) copy and paste
- ( ) email

Correct

2. Scenario 1 continued — 0.75 / 1 point

You've downloaded the data from your company database and imported it into a spreadsheet. To use the dataset for this scenario, click the link below and select "Use Template."

Link to template: Course Challenge - Scenario 1

OR

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

Course Challenge Dataset - Scenario 1 - Scenario 1_ Pharmacy Data - Part 1
CSV File

Now, it's time to process the data. As you know, this step involves finding and eliminating errors and inaccuracies that can get in the way of your results. **While cleaning the data, you notice there's missing data in one of the rows. What might you do to fix this problem? Select all that apply.**

- [x] Delete the row with the missing data point

This should not be selected
Review the section on the data analysis process for a refresher.

- [x] Ask you supervisor for guidance

Correct

- [x] Ask a colleague on your team how they've handled similar issues in the past

Correct

- [ ] Sort the spreadsheet so the row with missing data is at the bottom

3. Scenario 1 continued — 1 / 1 point

Once you've found the missing information, you analyze your dataset.

During analysis, you create a new column F. At the top of the column, you add the attribute Average Percentage of Total Sales - Splashtastic. Select the correct definition for an attribute.

- (●) A characteristic or quality of data used to label a column
- ( ) A headline or subhead
- ( ) All of the characteristics of something contained in a table
- ( ) An observation of data within a column

Correct

4. Scenario 1 continued — 1 / 1 point

Next, you determine the average total daily sales over the past 12 months at all stores. The range that contains these sales is E2:E39. **Identify the correct way to write your function.**

- ( ) =AVERAGE(E2,E39)
- ( ) =AVERAGE(E2-E39)
- (●) =AVERAGE(E2:E39)
- ( ) =AVERAGE(E2+E39)

Correct

5. Scenario 1 continued — 1 / 1 point

You've reached the share phase of the data analysis process. What can you do in this phase to share the Splashtastic sales insights you've discovered?

- ( ) Present your findings to customers.
- ( ) Establish a repository for the data.
- ( ) Revisit the analyze phase.
- (●) Present your findings to stakeholders.

Correct

6. Scenario 2, questions 6-10 — 1 / 1 point

You've been working for the nonprofit National Dental Society (NDS) as a junior data analyst for about two months. The mission of the NDS is to help its members advance the oral health of their patients. NDS members include dentists, hygienists, and dental office support staff.

The NDS is passionate about patient health. Part of this involves automatically scheduling follow-up appointments after crown replacement, emergency dental surgery, and extraction procedures. NDS believes the follow-up is an important step to ensure patient recovery and minimize infection.

Unfortunately, many patients don't show up for these appointments, so the NDS wants to create a campaign to help its members learn how to encourage their patients to take follow-up appointments seriously. If successful, this will help the NDS achieve its mission of advancing the oral health of all patients.

Your supervisor has just sent you an email saying that you're doing very well on the team, and he wants to give you some additional responsibility. He describes the issue of many missed follow-up appointments. You are tasked with analyzing data about this problem and presenting your findings using data visualizations.

An NDS member with three dental offices in Colorado offers to share its data on missed appointments. So, your supervisor uses a database query to access the dataset from the dental group. The query instructs the database to retrieve all patient information from the member's three dental offices, located in zip code 81137.

The table is dental_data_table, and the column name is zip_code. You have written the following query, but received an error when it ran. What is the clause that will correct this query?

```
SELECT *
FROM dental_data_table
WHERE dental_data_table = 81137
```

- (●) WHERE zip_code = 81137
- ( ) WHERE = 81137
- ( ) WHERE zip_code 81137
- ( ) WHERE_zip_code = 81137

Correct

7. Scenario 2 continued — 1 / 1 point

The dataset your supervisor retrieved and imported into a spreadsheet includes a list of patients, their demographic information, dental procedure types, and whether they attended their follow-up appointment. To use the dataset for this scenario, click the link below and select "Use Template."

Link to template: Course Challenge - Scenario 2

OR

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

Course Challenge Dataset - Scenario 2
CSV File

The patient demographic information includes data such as age, gender, and home address. You review the demographic data, paying particular attention to geography. **What geographic aspect of the data may negatively impact fairness?**

- (●) The patients all live in the same zip code.
- ( ) The patients all live in houses.
- ( ) The patients all live in the same city.
- ( ) The patients all live in the same country.

Correct

8. Scenario 2 continued — 1 / 1 point

As you're reviewing the dataset, you notice that there are a disproportionate number of senior citizens. So, you investigate further and find out that this zip code represents a rural community in Colorado with about 800 residents. In addition, there's a large assisted-living facility in the area. Nearly 300 of the residents in the 81137 zip code live in the facility.

You recognize that's a sizable number, so you want to find out if age has an effect on a patient's likelihood to attend a follow-up dental appointment. You analyze the data, and your analysis reveals that older people tend to miss follow-ups more than younger people.

So, you do some research online and discover that people over the age 60 are 50% more likely to miss dentist appointments. Sometimes this is because they're on a fixed income. Also, many senior citizens lack transportation to get to and from appointments.

With this new knowledge, you write an email to your supervisor expressing your concerns about the dataset. He agrees with your concerns, but he's also impressed with what you've learned and thinks your findings could be very important to the project. He asks you to change the business task. Now, the NDS campaign will be about educating dental offices on the challenges faced by senior citizens and finding ways to help them access quality dental care.

Changing the business task involves defining the new question or problem to be solved.

- (●) True
- ( ) False

Correct

9. Scenario 2 continued — 1 / 1 point

You continue with your analysis. In the end, your findings support what you discovered during your online research: As people get older, they're less likely to attend follow-up dental visits.

But you're not done yet. You know that data should be combined with human insights in order to lead to true data-driven decision-making. So, your next step is to share this information with people who are familiar with the problem professionally. They'll help verify the results of your data analysis.

Fill in the blank: Subject matter experts are people who are familiar with a problem. They can help by identifying inconsistencies in the analysis, offering insights into the business problem, and _____.

- (●) validating the choices being made
- ( ) collecting data relevant to the business problem
- ( ) redefining the business problem
- ( ) creating a presentation with the data

Correct

10. Scenario 2 continued — 1 / 1 point

The subject-matter experts are impressed by your analysis. The team agrees to move to the next step: data visualization. You know it's important that stakeholders at NDS can quickly and easily understand that older people are less likely to attend important follow-up dental appointments than younger people. This will help them create an effective campaign for members.

It's time to create your presentation to stakeholders. It will include a data visualization that demonstrates the lifetime trend of people being less likely to attend follow-up appointments as they get older.

Fill in the blank: The type of chart that would be most effective in visualizing this is a _____.

- (●) line chart
- ( ) doughnut chart
- ( ) bar chart
- ( ) pie chart

Correct