Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 100%. We keep your highest score.

Next item →

Overview

Now that you have explored how businesses use data in the real world, pause for a moment and think about what you are learning. In this self-reflection, you will consider fairness and data use in three example business cases and respond to brief questions with your thoughts.

1. Case Study #1 1 / 1 point

To improve the effectiveness of its teaching staff, the administration of a high school offered the opportunity for all teachers to participate in a workshop. They were not required to attend; instead, the administration encouraged teachers to sign up. Of the 43 teachers on staff, 19 chose to take the workshop.

At the end of the academic year, the administration collected data on teacher performance for all teachers on staff. The data was collected via student survey. In the survey, students were asked to rank each teacher's effectiveness on a scale of 1 (very poor) to 6 (very good).

The administration compared data on teachers who attended the workshop to data on teachers who did not. The comparison revealed that teachers who attended the workshop had an average score of 4.95, while teachers who did not attend had an average score of 4.22. The administration concluded that the workshop

1. **Case Study #1** 1 / 1 point

To improve the effectiveness of its teaching staff, the administration of a high school offered the opportunity for all teachers to participate in a workshop. They were not required to attend; instead, the administration encouraged teachers to sign up. Of the 43 teachers on staff, 19 chose to take the workshop.

At the end of the academic year, the administration collected data on teacher performance for all teachers on staff. The data was collected via student survey. In the survey, students were asked to rank each teacher's effectiveness on a scale of 1 (very poor) to 6 (very good).

The administration compared data on teachers who attended the workshop to data on teachers who did not. The comparison revealed that teachers who attended the workshop had an average score of 4.95, while teachers who did not attend had an average score of 4.22. The administration concluded that the workshop was a success.

Reflection

Consider this scenario:

- What are the examples of fair or unfair practices?
- How could a data analyst correct the unfair practices?

Now, write 2-3 sentences (40-60 words) in response to each of these questions. Enter your response in the text box below.

Unfair Practices: The main unfair practice is self-selection bias. Because workshop attendance was voluntary, more motivated or confident teachers likely signed up. This makes the comparison unfair, as the two groups weren't equivalent from the start. Other factors like experience or class size were also ignored. How a Data Analyst Could Correct This: An analyst could use matching techniques to compare teachers with similar experience and subjects, where one attended and the other didn't. They should also collect and analyze data on other potential factors (e.g., classroom environment) to isolate the workshop's true effect.

Great work reinforcing your learning with a thoughtful self-reflection! In your response, you likely noted that this is an example of unfair practice. It is tempting to conclude—as the administration did—

Great work reinforcing your learning with a thoughtful self-reflection! In your response, you likely noted that this is an example of unfair practice. It is tempting to conclude—as the administration did—that the workshop was a success. However, since the workshop was voluntary and not random, it is not appropriate to infer a causal relationship between attending the workshop and the higher rating.

The workshop might have been effective, but other explanations for the differences in the ratings cannot be ruled out. For example, another explanation could be that the staff volunteering for the workshop were the better, more motivated teachers. This group of teachers would be rated higher whether or not the workshop was effective.

It's also notable that there is no direct connection between student survey responses and workshop attendance. The data analyst could correct this by asking for the teachers to be selected randomly to participate in the workshop. They could also collect data that measures something more directly related to workshop attendance, such as the success of a technique the teachers learned in that workshop.

2. **Case Study #2** 1 / 1 point

An automotive company tests the driving capabilities of its self-driving car prototype. They carry out the tests on various types of roadways—specifically, a race track, trail track, and dirt road.

The researchers only test the prototype during the daytime. They collect two types of data: sensor data from the car during the drives and video data of the drives from cameras on the car.

They review the data after the initial tests. The results illustrate that the new self-driving car meets the performance standards across each of the roadways. As a result, the car can progress to the next phase of testing, which will include driving in various weather conditions.

Reflection

Consider this scenario:

- What are the examples of fair or unfair practices?

2. **Case Study #2** 1 / 1 point

An automotive company tests the driving capabilities of its self-driving car prototype. They carry out the tests on various types of roadways—specifically, a race track, trail track, and dirt road.

The researchers only test the prototype during the daytime. They collect two types of data: sensor data from the car during the drives and video data of the drives from cameras on the car.

They review the data after the initial tests. The results illustrate that the new self-driving car meets the performance standards across each of the roadways. As a result, the car can progress to the next phase of testing, which will include driving in various weather conditions.

Reflection

Consider this scenario:

- What are the examples of fair or unfair practices?
- How could a data analyst correct the unfair practices?

Now, write 2-3 sentences (40-60 words) in response to each of these questions. Enter your response in the text box below.

Unfair Practices: The testing is unfair due to selection bias and undercoverage bias. By only testing during daytime, the team fails to account for different lighting conditions (night, dusk, dawn) that critically impact sensor and camera performance. This creates an incomplete and overly optimistic picture of the car's real-world capabilities. How a Data Analyst Could Correct This: A data analyst should insist on testing across the full range of environmental conditions, including night, rain, and fog. They must ensure the dataset for analysis represents diverse real-world scenarios to provide a fair and accurate assessment of the car's true performance and safety.

Once again, this case study shows an unfair practice. Your response probably mentioned that, although the researchers test the prototype on three different tracks, they only conduct tests during the day.

Conditions on each track may be very different during the day and night and this could change the results significantly. The data analyst should correct this by asking the test team to add in nighttime

Once again, this case study shows an unfair practice. Your response probably mentioned that, although the researchers test the prototype on three different tracks, they only conduct tests during the day.

Conditions on each track may be very different during the day and night and this could change the results significantly. The data analyst should correct this by asking the test team to add in nighttime testing to get a full perspective of how the prototype performs at any time of the day on the tracks.

3. **Case Study #3** 1 / 1 point

An amusement park plans to add new rides to their property. First, they need to determine what kinds of new rides visitors want the park to build. In order to understand their visitors' interests, the park develops a survey.

They decide to distribute the survey near the roller coasters because the lines are long enough that visitors will have time to answer all of the questions. After collecting this survey data, they find that most of the respondents want more roller coasters at the park. They conclude that they should add more roller coasters, as most of their visitors prefer them.

Reflection

Consider this scenario:

- What are the examples of fair or unfair practices?
- How could a data analyst correct the unfair practices?

Now, write 2-3 sentences (40-60 words) in response to each of these questions. Enter your response in the text box below.

Unfair Practices: The primary unfair practice is sampling bias. By surveying only visitors near roller coasters, the park oversamples roller coaster fans and systematically excludes visitors who prefer other areas (like water rides or children's sections). This makes the data unrepresentative of the entire visitor population. How a Data Analyst Could Correct This: A data analyst should design a stratified sampling method to distribute surveys proportionally across all key park areas (entrance, kiddie rides, food courts, etc.). This ensures the sample accurately represents the park's diverse visitor base

They decide to distribute the survey near the roller coasters because the lines are long enough that visitors will have time to answer all of the questions. After collecting this survey data, they find that most of the respondents want more roller coasters at the park. They conclude that they should add more roller coasters, as most of their visitors prefer them.

Reflection

Consider this scenario:

- What are the examples of fair or unfair practices?
- How could a data analyst correct the unfair practices?

Now, write 2-3 sentences (40-60 words) in response to each of these questions. Enter your response in the text box below.

Unfair Practices: The primary unfair practice is sampling bias. By surveying only visitors near roller coasters, the park oversamples roller coaster fans and systematically excludes visitors who prefer other areas (like water rides or children's sections). This makes the data unrepresentative of the entire visitor population. How a Data Analyst Could Correct This: A data analyst should design a stratified sampling method to distribute surveys proportionally across all key park areas (entrance, kiddie rides, food courts, etc.). This ensures the sample accurately represents the park's diverse visitor base and their varied ride preferences.

A thoughtful reflection would have included information about the decision to distribute surveys in places where visitors would have time to respond. Otherwise, it accidentally introduces sampling bias.

The only respondents to the survey are people waiting in line for the roller coasters. This may unfairly bias survey results, because respondents might prefer roller coasters. A data analyst could reduce sampling bias by distributing the survey at the entrance and exit of the amusement park. This would avoid targeting roller coaster fans and provide results from the park's general audience.