

### Focus on integrity

### Data integrity and analytics objectives

### Overcoming the challenges of insufficient data

## Testing your data

Consider the margin of error

### Weekly challenge 1

 **Reading:** Glossary: Terms and definitions  
10 min

 **Quiz:** Weekly challenge 1  
8 questions

✔ Congratulations! You passed!

Grade received 100% Latest Submission Grade 100% To pass 80% or higher

Quiz • 40 min

**Go to next item**

### Review Learning Objectives

1. Fill in the blank: Data \_\_\_\_\_ refers to the accuracy, completeness, consistency, and trustworthiness of data throughout its life cycle. 1 / 1 point

☒ Submit your assignment

**Due** Jun 18, 11:59 PM +08 **Attempts** 3 every 24 hours

- ☐ replication
- ☒ integrity
- ☐ analysis **✓ Receive grade**
- To Pass 80% or higher**
- ☒ **Correct**

**Try again**

**Your grade**  
**100%**

**View Feedback**

2. What is one potential problem associated with data manipulation that analysts must be aware of? 1 / 1 point

- ☐ Data manipulation can make a dataset easier to read.
- ☒ Data manipulation can introduce errors.
- ☐ Data manipulation can separate a dataset among different locations.
- ☐ Data manipulation can help organize a dataset.

3. You are working for a global technology company. You have a dataset with the company's total cell phone sales by country from 2015 to present. Based on the data you have, what questions are you able to answer?

- ☐ What was the effect on sales when a new phone model was launched?
- ☐ What was the effect on sales when new phone features were introduced?
- ☐ What are the mean cell phone sales for each country since 2010?
- ☒ What countries have the most cell phone sales in the past three years?


4. A data analyst is given a dataset for analysis. To use the template for this dataset, click the link below and select "Use Template."

Link to template: [June 2014 Invoices](#) 

**OR**

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

June 2014 Invoices - Sheet1

 **June 2014 Invoices - Sheet1**  
CSV File

What rows have duplicate data?

- ☐ Rows 14 and 15  
☐ Rows 3 and 5  
☐ Rows 1 and 2  
☒ Rows 8 and 9

5. A data analyst is working on a project around a national supply chain. They have a dataset with lots of relevant data from about half of the country. However, they decide to generate new data that represents the entire nation. What type of insufficient data does this scenario describe?

- ☐ Data that keeps updating
- ☐ Outdated data
- ☐ Data from only one source
- ☒ Geographically limited data

6. A clothing manufacturer wants to learn more about why their consumers have purchased the brand's products. How should this manufacturer conduct their survey?

- ☒ Send the survey to a representative sample of their customers
- ☐ Send the survey to random people who buy clothes
- ☐ Send the survey to their least frequent customers
- ☐ Send the survey to customers who have purchased more than one product

7. A restaurant wants to gather data about a new dish by giving out free samples and asking for feedback. Who should the restaurant give samples to? 1 / 1 point

- ☐ Diners who are willing to pay for the samples
- ☐ Diners who spend the most money on their meal
- ☒ All diners
- ☐ Selecting diners at random

8. Which of the following processes helps ensure a close alignment of data and business objectives? 1 / 1 point

- ☒ Maintaining data integrity
  - ☐ Transferring data multiple times
  - ☐ Completing data replication
  - ☐ Having data update automatically during analysis
- ☒ **Correct**