Lesson: 1.1 What is Data Science?

In this lesson, students will learn what data science is, what a data scientist does, and the different types of questions that can be asked about data. Students will learn that statistical questions include computations or finding a relationship or pattern.

Objective

Students will be able to:

Recognize and formulate statistical questions

Think critically about data and its sources

Activities

These are all the activities included in the lesson.

Activity

1.1.1 What is Data Science?

1.1.2 What is Data Science?

1.1.3 What Does a Data Scientist Do?

1.1.4 What Does a Data Scientist Do?

1.1.5 Statistical Questions

1.1.6 Lookup, Compute, Relate

1.1.7 Exploring Data I

1.1.8 Exploring Data II

1.1.9 Exploring Data III

1.1.10 Exploring Data IV

1.1.11 My Interests and Goals

Solution References

Refer to the solution reference for a more detailed look at exercise solutions.

Solution Reference

1.1.2 What is Data Science?

1.1.4 What Does a Data Scientist Do?

1.1.5 Statistical Questions

1.1.6 Lookup, Compute, Relate

Problem Guides

Refer to the problem guides for a more in-depth look at this lesson's problems.

Problem Guide

1.1.4 What Does a Data Scientist Do?

1.1.5 Statistical Questions

1.1.6 Lookup, Compute, Relate

1.1.8 Exploring Data II

1.1.10 Exploring Data IV

1.1.11 My Interests and Goals

Vocabulary

These are the key terms for this lesson.

Term Definition

data science

The process of learning about the world using data and computation.

statistical question

A question that could have a variety of different answers.

Planning Notes

This is a longer lesson and may span between two class periods.

Allow time for discussions and for answering questions that arise during the lesson. This will help gauge what your students already know, or don’t know, about data, statistics and coding.

Explore the data tour in the Exploring Data activity before class. If you have a student who belongs to a country listed in the tour, consider asking them to share their experience of the data fact that is shared in the tour.

Teaching and Learning Strategies

Lesson Opener:

Have students brainstorm and write down answers to the discussion questions listed below. Students can work individually or in groups/pairs. Have them share their responses. [5 mins]

Activities:

Watch the lesson video and complete the corresponding quiz to check for understanding. [10 mins]

Watch What Does a Data Scientist Do? connection video and answer the corresponding free-response activity. [10 mins]

Allow time for a discussion of imposter syndrome. This would be a good time to instill confidence in your students and remind them that failure is a part of the learning process and should be expected and valued.

Complete the Statistical Questions free-response activity. [10 mins]

Students can be paired up or put into small groups to discuss the questions in this activity.

Complete the Lookup, Compute, Relate free-response activity. [15 mins]

Have students explain why their computer and related questions qualify as such.

Complete the Exploring Data I connection site and corresponding free-response activity. [15 min]

Bring together the class as a whole to share and discuss their findings and thoughts about the data in this activity.

Complete the Exploring Data III connection article and corresponding free-response activity. [15 mins]

The questions in this activity are especially important to discuss. Can data always be trusted? Lead a discussion with your students to see their thoughts on the matter.

Complete the My Interests and Goals free-response activity. [5-10 mins]

The responses to these questions can be helpful in pairing or grouping students based on their shared interests later in the course.

Lesson Closer:

Have students reflect and discuss their responses to the end-of-class discussion questions. There is an option for students to record their own data for a 24-hour period and then answer the discussion questions based on their own data. [5 mins]

Prior Knowledge

No prior knowledge is needed.

Video Slides

Discussion Questions

Beginning of Class:

In your own words, explain what data is.

Sample Response: Data is information, or observations, that have been gathered and recorded.

Where does data come from?

Sample Response: Data can come from cell phones, computers, school records, surveys, etc.

Give an example of data that is being collected about you.

Sample Response: School grades, attendance, doctor’s records, social media posts, etc.

End of Class:

Think of your normal day-to-day routine. What information or data could potentially be collected electronically based on your activities during the day?

Sample Response: Screen time, social media comments/posts, GPS data, purchases made (not cash), shows watched on Netflix, songs played on Spotify.

Optional: Keep a data diary for 24 hours. Write down everything that you do that could provide someone with electronic data about you.

What concerns might you have about data collected about yourself?

Sample Response: I am concerned that someone could find my location or steal my account information.

What good examples of data use could result from the data that is collected about yourself?

Sample Response: Spotify can recommend song suggestions. Netflix knows which type of shows I like to watch.

Modification: Advanced

Have students explore a subset of data science (artificial intelligence, predictive analysis, reporting) further and present their findings to the class.

Modification: Special Education

Have students keep a list of keywords that they find are usually associated with statistical questions such as “typically”, “on average”, “normally”, etc.

Modification: English Language Learners

Provide students with a list of keywords that they find are usually associated with statistical questions such as “typically”, “on average”, “normally”, etc.