

Unit 3: Data Science Practices

Lesson 3.3: Graphs and Figures for One Variable

In this lesson, students will learn about different univariate graphs, namely the boxplot, bar chart, histogram, and dotplot.

Duration: 90 minutes

Objective: By the end of this lesson, students will know how to create boxplots, histograms, bar charts, and dotplots in CODAP. Students will also be able to identify when each kind of graph is appropriate.

Lesson Walkthrough Video: [Unit 3 Lesson 3 - Teacher Walkthrough](#)

CSTA Standards in this Lesson

Identifier	Concept	Subconcept	Standards
3A-DA-11	Data & Analysis	Collection Visualization & Transformation	Create interactive data visualizations using software tools to help others better understand real-world phenomena.
3A-DA-12	Data & Analysis	Inference & Models	Create computational models that represent the relationships among different elements of data collected from a phenomenon or process.
3A-IC-24	Impacts of Computing	Culture	Evaluate the ways computing impacts personal, ethical, social, economic, and cultural practices.

Lesson activities

Warmup (10 min)

(CSTA standards in this activity: 3A-DA-11)

- Students examine a graph from March 30, 2023, regarding the prices of [Eggs](#), and are asked to consider the following questions:
 - What kind of graph is this?
 - What variable(s) do you see represented? Is there more than one?
 - What do you notice? What do you wonder?
 - What story does this graph portray?

Seal Tracking Video (10 min)

(CSTA standards in this activity: 3A-IC-24)

- Students watch the seal tracking video to learn some background on seals as a protected species and how scientists track them.
- After the video, ask students:
 - Why is this work important?
 - What could we use this data for?
 - What kinds of data do you think we collect from the seals?

Seal Tracking Data (40 min)

(CSTA standards in this activity: 3A-DA-11, 3A-DA-12)

- Open the [Four Seals](#) example in CODAP.
 - What kind of data is this? What's going on in this dataset?
 - Drag "animal_id" onto the map and see what happens.
 - What does this help you understand?
- Drag "animal_id" onto the y axis and "depth" onto the x axis to create a **comparative dotplot**. You should see data divided by seal.
 - What does each dot represent? Why are there four dotplots? What does it mean when lots of dots are stacked on top of each other?
 - Which seal would you say tends to swim deepest? (You may want to click the Ruler icon, and check a box to show Mean, Median, or other metrics.)
 - Look at the different shapes of the four dotplots. What is unusual about seal 546? How would you describe the shape of seal 541's distribution?
- Click the Ruler icon and check "**Box Plot**."
 - **Teacher:** A box plot summarizes data from a number of observations. The box marks the **median** (in the very middle), the **first quartile** (or 25th percentile of the data, or median of the lower half of the data) on the left side of the box, and the **third quartile** (or 75th percentile of the data, or median of the higher half of the data) on the right side of the box.
 - Where do the "whiskers" of the box plot end?
 - Can you see individual data values in the box plot like you can on the dotplot?
 - Go back into the Ruler icon menu and click "Show Outliers." Are there any outliers? How do you think the program makes the decision to mark certain points as outliers or not?
 - **Teacher:** Outliers might be marked by being more than 2.5 standard deviations from the mean, or more than 1.5 **interquartile ranges** (third quartile - first quartile) past either quartile (1.5 IQRs below Q1 or 1.5 IQRs above Q3).
- Now, create a new graph. Drag "speed" onto the x-axis to create a single dotplot.

- What **shape** does this dotplot have? Is it symmetric? Skewed to the left? Skewed to the right? (**Teacher** note: this plot should be skewed right, because the tail goes off to the right.)
- Where would you say the **center** of the speed variable rests? (You might use the Ruler icon to find the Mean or Median. Are they the same? Different?)
- How much **spread** is there to this data? (You might use the Ruler icon to add the standard deviation or IQR.) How much does the average speed vary from the mean speed?
- Still working with the new speed graph, go to the Graph icon and “Group Into Bins,” then go back to the Graph icon and “**Fuse Dots Into Bars.**”
 - What you see now is called a **histogram** – a summary plot for a set of data that organizes data points into “bins.”
 - Can you see individual data points in a histogram?
 - Why would a histogram be useful?

Activity: Bar Chart vs. Histogram (20 min)

(CSTA standards in this activity: 3A-DA-11)

- Review the slides on examples of bar charts and histograms.
- Discuss the situations listed on slide 12, and decide as a class whether a histogram or bar chart would be more appropriate in each case. This could be a nice chance to leverage **debate** in class!

Exit Ticket (10 min)

(CSTA standards in this activity: 3A-DA-11)

- In CODAP, students create a new graph by dragging temperature onto the x-axis and answer the following questions:
 - What kind of graph is this?
 - Dotplot!
 - Would this data be displayed with a histogram or a bar chart?
 - Histogram (the x-axis is a continuous variable, temperature)
 - What is the **mean** temperature? (Hint: use the Ruler icon!)
 - 13.13

Assessment:

Assess student understanding through participation in class discussions and class activities.

