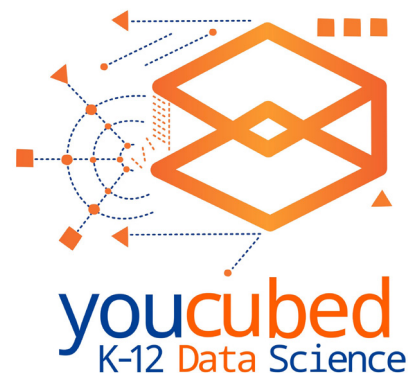


Unit 2: Working With Data Analysis Tools



Introduction:

In this lesson, students will be introduced to CODAP, a data visualization and analysis tool that will be used throughout the data science units. One way to become familiar with CODAP and also see a version of this lesson taught is to take Jo Boaler’s online course that accompanies these lessons. <https://www.youcubed.org/21st-century-teaching-and-learning/>

CODAP has many tutorials and they can be found here: <https://codap.concord.org/help/>

Here students will explore and analyze health data and extract insights to share with their class. They will be invited to investigate data they are interested in and practice communicating their findings, explaining how they know what they know, and embracing the mistakes and successes along the way. Because this lesson introduces a technology that might be new for both students and teachers, this is an opportunity for the whole learning community to be curious and push the boundaries of their understanding.

Each lesson in the unit will start with a data talk.

Lesson 1: Exploring Data with CODAP

Time	Activity	Materials
10 min	Data talk:	Steph Curry graph (https://fivethirtyeight.com/features/stephen-curry-is-the-revolution/)
20 min	CODAP exploration: Explore health data in CODAP as a class. The teacher can control the computer, students can direct with questions and curiosities.	CODAP health data (https://codap.concord.org/app?di=https://codap.xyz/plugins/nhanes15/nhanes.html) Note that the data includes attributes relating to sexual health. Decide if this is appropriate for your students - there are other attributes that give plenty of choice for students to work with. If you do not want to use the health data this is different dataset that can also work well: data portal from California residents https://codap.concord.org/releases/app/dg/en/cert/index.html#shared=82989

20 min	<p>Student CODAP exploration: Students explore CODAP on their own - either health or California residents.</p>	<p>CODAP health data (https://codap.concord.org/app?di=https://codap.xyz/plugins/nhanes15/nhanes.html)</p> <p>Another good option: CODAP Californians https://codap.concord.org/app/static/dg/en/cert/index.html#shared=82989</p>
5 min	<p>Reflection/journal entry:</p> <ul style="list-style-type: none"> • How does the CODAP tool help with our data filled world you learned about in the last unit? • What did CODAP allow you to see or learn? • What are you excited about? 	<ul style="list-style-type: none"> • Journal

Data talk:

Today’s data talk is a graph illustrating Steph Curry’s shooting efficiency. We chose this visual because it touches on something that our students were familiar with, basketball, and visualizes it in a new and interesting way.

For example, some noticings could be:

- The size of the hexagons indicates the frequency with which he shoots from that spot.
- I wonder why one of his most effective zones is right next to his least effective zone (top left).

This Data Talk is meant to show students yet another creative way of representing data and encouraging them to further develop their data literacy skills.

CODAP exploration:

Guide the students through an introduction to CODAP. Teachers should explore the tool on their own before bringing it to the class. If you already took the DS Online Course (<https://www.youcubed.org/21st-century-teaching-and-learning/>), you will have had a chance to explore the software. If you haven’t, we encourage you to check out our new online course!

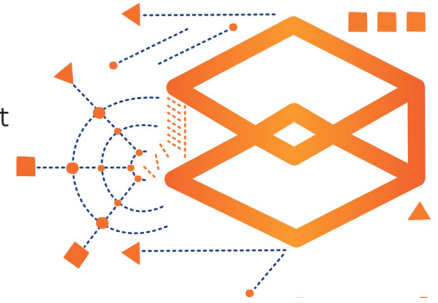
Additionally, CODAP has some great tutorials you can learn more from here: <https://codap.concord.org/help/>

Note that the data includes attributes relating to sexual health. Decide if this is appropriate for your students - there are other attributes that give plenty of choice for students to work with. We recommend looking at all column headings in any data set before you introduce it to your students.

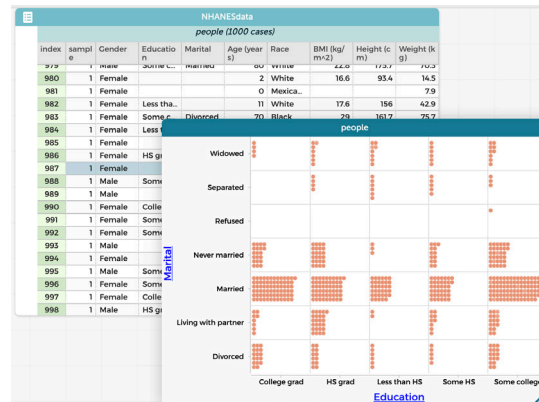
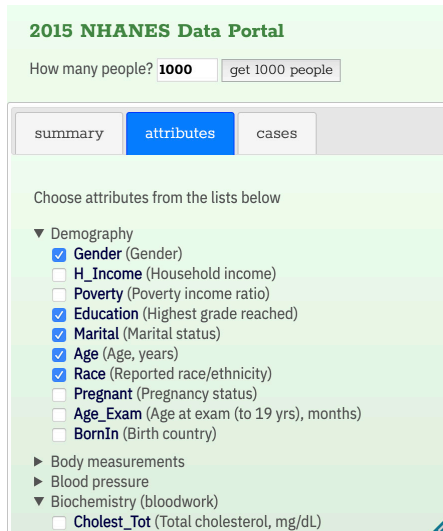
If you do not want to use the health data here is a different dataset that can also work well: 2013 California ACS <https://codap.concord.org/app/static/dg/en/cert/index.html#shared=82989>

Most importantly, however, you do not need to be a CODAP expert to introduce it to your class. This is a great moment to model uncertainty, question-asking, and a willingness to engage with new technology. You can all learn together!

As a class, you will explore CODAP's health dataset. In order to be on the same page as you learn to use the software, you can share your computer screen with the software up, and students can ask questions they have about the data, to be answered together.



First, you will have to select what data you would like to look at in the Data Portal (see image below). The tool will get a random sample of the size of your choosing of the (very large) 2015 NHANES data with the attributes that you select.



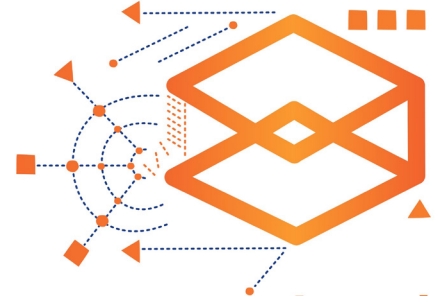
Afterward, you can experiment with displaying these different attributes and how they interact. First, you might ask students to pick an attribute they would like to explore and put that in a graph where students can look at it and notice some features of the distribution. You can show some of CODAP's capabilities like displaying the mean of a distribution. Then you might ask if they wonder anything about how that variable relates to other variables and bring those in either in the second axis or as color-coding (or both).

The goal is to show the students the basic capabilities that CODAP has. You can control the screen, but the exploration should be student-driven. If a student asks a question that you do not know the answer to, try investigating how to answer it and find the answer together (Google it!). This is something we want to encourage students to do, so modeling it early is ideal.

Student CODAP exploration:

Give students an opportunity to explore CODAP on their own. In pairs or small groups, they can explore the same data set and perform similar data manipulations to the ones performed for the whole class themselves, looking at attributes they find the most interesting.

Encourage students to take notes about interesting things they find, tools they discover and any wonderings they have about their data exploration. In the next class they will delve deeper into the data, working in groups to find interesting things to share in a data graph.



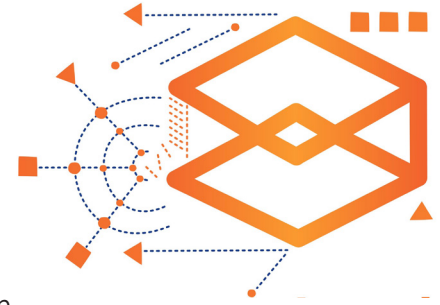
Look-fors:

- **Are students exploring and finding their own tools?** Encourage students to be explorers within the data set. If they feel they have “messed it up” they can refresh their screen and start over. Encourage them to persevere and celebrate mistakes.
- **Are students making interesting discoveries about CODAP tools?** Consider pausing class and asking students to share interesting tools or strategies they have found. You may want to start a class list of valuable tools and strategies.

Reflection/journal entry:

- How does the CODAP tool help with our data filled world you learned about in the last unit?
- What did CODAP allow you to see or learn?
- What are you excited about?

Lesson 2: Telling a Story with a Visual



Introduction:

In this activity students continue their exploration of the health data set in CODAP. Their goal today is to make a data visual representing 3 findings from the data set. Their findings should be unique, interesting and not something obvious.

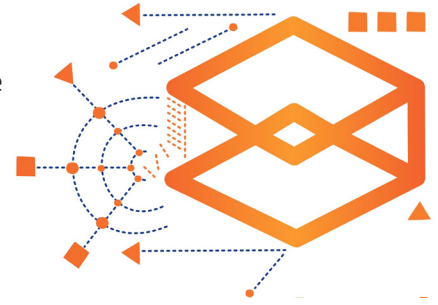
Time	Activity	Materials
10 min	Data talk:	Hot dog eating contest graph (https://fivethirtyeight.com/features/competitive-hot-dog-eaters-have-made-america-great-again/)
30 min	Data graph activity: <ul style="list-style-type: none"> Ask students to work in groups and explore health or another data set you have chosen using CODAP. Ask students to make a data display to illustrate their findings for the class. These could be some form of display on chart paper or through google docs, slides or another digital tool. Expect all of the students' presentations to look different, with different questions and findings. This will add to the richness of the discussion that follows. 	<ul style="list-style-type: none"> Chart paper Colored pens or pencils Rulers or scissors The data and visualizations could be made in a google doc or as a slide presentation
20 min	Student presentations: <ul style="list-style-type: none"> Guiding questions for discussions are: "How do you know this is true?" Is the data graph clear? Does the visual contain all the information needed for you to understand the graph? 	
10 min	Reflection/journal entry: <ul style="list-style-type: none"> What did you learn from another group's presentation? What was something fun or interesting you discovered? 	<ul style="list-style-type: none"> Journal

Data talk :

For the second data talk in this unit, show students the graph of the results of Nathan's hot dog eating contest. The graph we chose has overlapping data points that can be hard to read. We also focused on the information around the graph, like title, axis labels, and color-coding. This sparked a discussion that students kept in mind as they shared graphs of their insights in the later part of the lesson.

For example, some noticings could be:

- I wonder why the winners of the last 20 years eat so much more than the winners of the 70s and 80s?
- The transparent circles make it easier to see lots of overlapping data points.



Data graph activity:

Students will explore the CODAP health data in groups again, but now with the goal of making a poster of the 3 most interesting facts they draw from the data. Students can choose any sampling of attributes they want from the data.

Students should make sure everyone in their group knows why the findings are true. Encourage students to not just choose the first 3 facts they find, but to choose samples and look for insights that truly interest them. Each interesting finding should have an accompanying visual - a graph or a drawing that represents the finding. For example, students graphing the relationship between age and height could extract the insight that kids grow much more quickly than adults. Remind students that data science is about illustrating interesting findings from data. It is not about sharing obvious patterns in data.

Student presentations:

Finally, each group should present their 3 most interesting findings from the CODAP health dataset to the class. Encourage students to explain how they know each of their findings are true, and to share their process and reasoning. Consider leaving time for Q&A from the students after each presentation. If students can't answer the questions asked, this is a great moment to model comfort with uncertainty. Share that "I don't know" is a perfectly reasonable answer, and encourage students to make conjectures about what they think the answer might be or how they might find it.

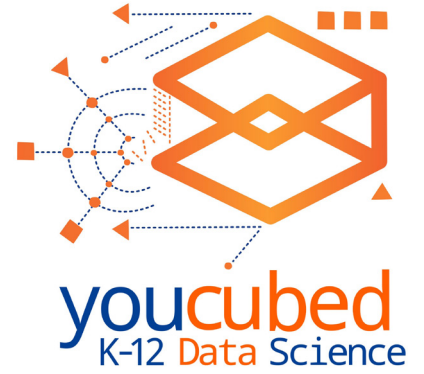
Look-fors:

- **How do students select an insight?** Students will likely explore many combinations of the attributes they sampled, but how do they find insights from them or select those insights as "interesting" (or not)? Discussing each of the graphs they create in their small groups is a powerful tool for students to find insights. They will have developed some experience in analyzing graphs from the earlier Data Talks, and treating each of their own graphs as little data talks with their group might help them draw conclusions. If students are stuck, you might encourage them to have these conversations with questions such as "Do you see any patterns in this data? What does that mean?" or "Is there anything that surprises you about this data? Did you expect it to look like this when you chose to visualize it?"
- **How are students justifying their insights?** Students should be relying on the data to justify their insights, not personal experience or conjecture. The guiding question of the final presentation should be "How do you know this is true?", asking students to show the data they are drawing upon to support their conclusions. If students are struggling to justify a claim with the data at hand, you could ask them questions such as "What is a claim that you can make from just the data at hand?" or "What kind of data would they need access to in order to justify that claim?"

Reflection/journal entry:

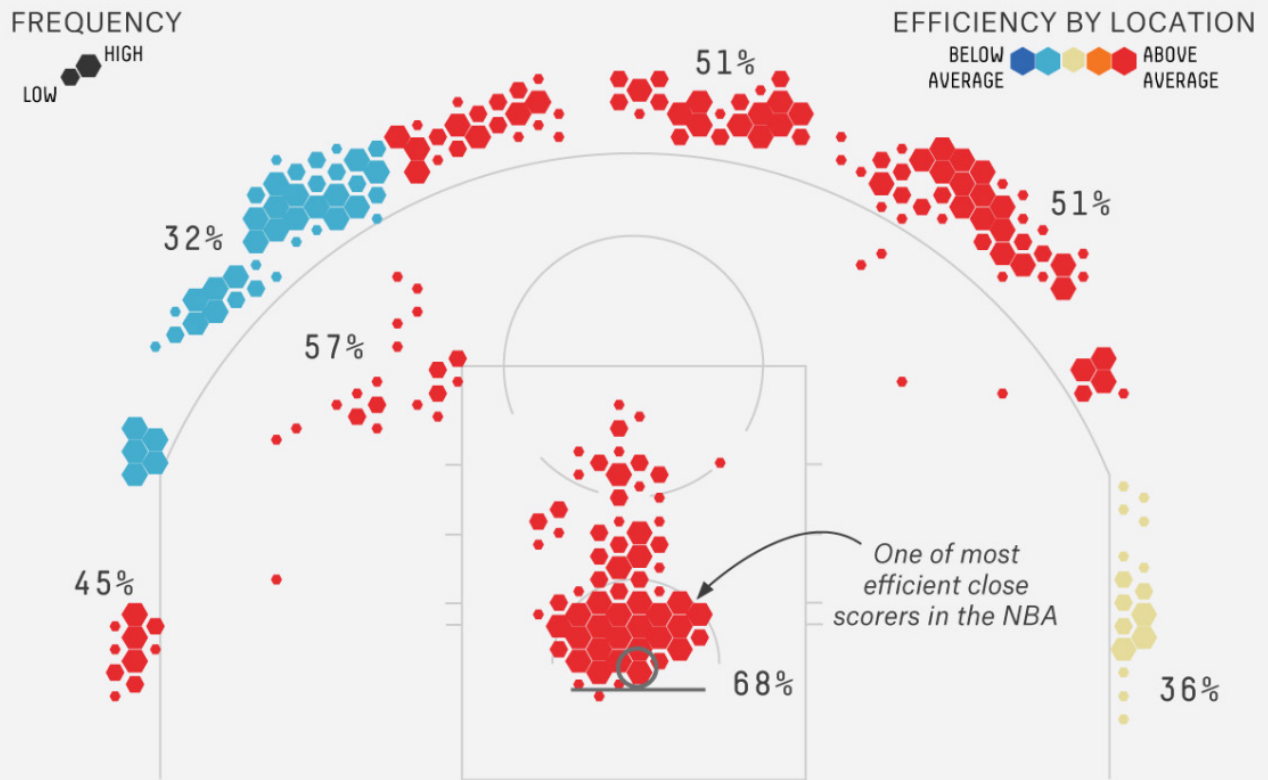
What did you learn from another group's presentation? What was something fun or interesting you discovered?

Data Talk



Stephen Curry Is One Of The Best

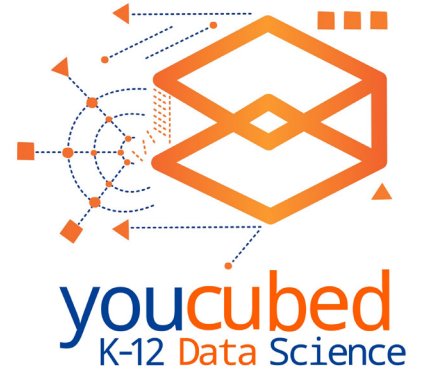
All of his shots, 2015-16 regular season



KIRK GOLDSBERRY

SOURCE: NBA

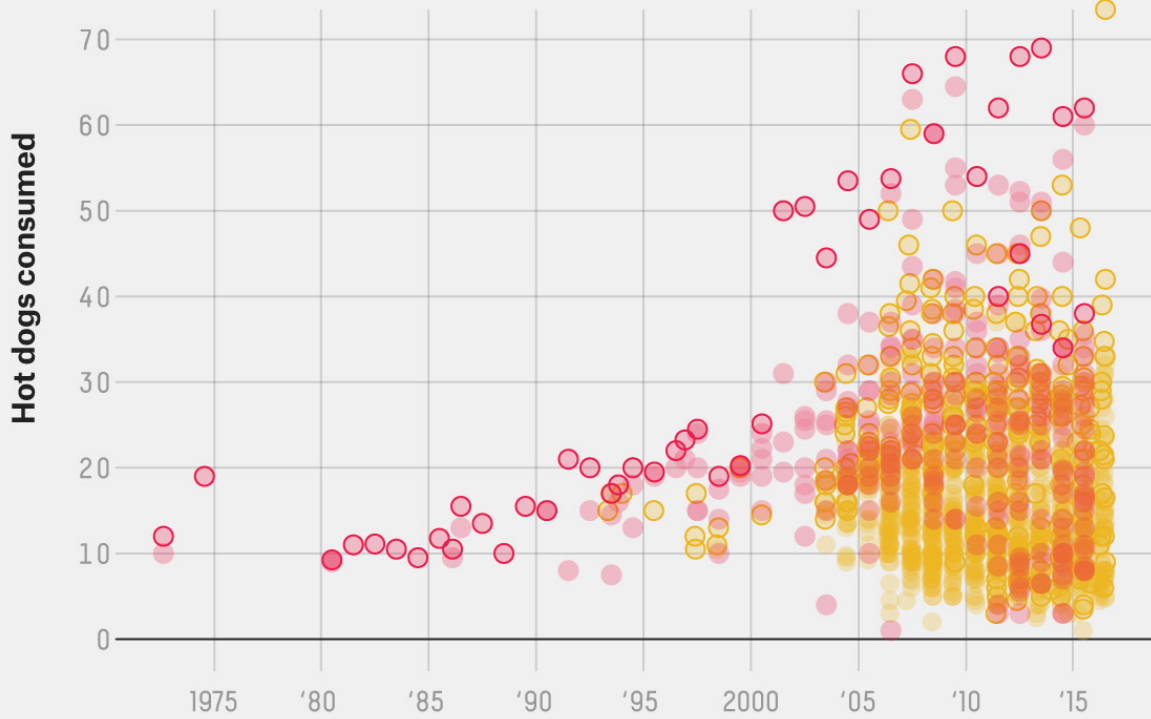
Data Talk



“Think we may wanna write this stuff down?”

Available Nathan’s Hot Dog Eating Contest results

○ FINALS WINNER ● FINALS LOSER ○ QUALIFIER WINNER ● QUALIFIER LOSER



FIVETHIRTYEIGHT

SOURCE: EATFEATS