Ethics and Privacy

Students consider ethical issues and privacy in the context of data science.

Prerequisites	Introduction to Computational Data Science
Relevant Standards OK NGSS K12CS CSTA	*Select one or more standards from the menu on the left (⌘-click on Mac, Ctrl-click elsewhere).*
Lesson Goals	Students will be able to... • Describe ethical and privacy considerations when it comes to data science
Student-facing Lesson Goals	• Let's discuss ethical concerns surrounding data science.
Materials	• Lesson Slides (Google Slides) • Computer for each student (or pair), with access to the internet • Student workbook, and something to write with
Preparation	• Make sure all materials have been gathered • Decide how students will be grouped in pairs
Supplemental Resources	

Language Table

Types	Functions	Values
Number	`num-sqrt, num-sqr, mean, median, modes`	`4, -1.2, 2/3`
String	`string-repeat, string-contains`	`"hello", "91"`
Boolean	`==, <, <=, >=, string-equal`	`true, false`
Image	`triangle, circle, star, rectangle, ellipse, square, text, overlay, bar-chart, pie-chart, bar-chart-summarized, pie-chart-summarized, histogram, scatter-plot, lr-plot`	
Table	`count, .row-n, .order-by, .filter, .build-column`	

Case Studies

40 minutes

Overview

Students break into groups and read one of three case studies, each dealing with a different issue in Data Science. They discuss the implications of each, then share back.

Launch

> "With great power comes great responsibility"

During World War 2, scientists were engaged in a race to develop new weapons, more powerful than anything the world had ever seen. While the immediate goal was "win the war", many of the scientists realized that the weapons they were

had ever seen. While the immediate goal was "win the war", many of the scientists realized that the weapons they were developing could be used for all sorts of things *after the war was over* - and not all of them were good.

With tech companies hiring Data Scientists at a staggering rate and collecting massive datasets on users for those scientists to mine, there's a *new* arms race happening right now. Search engines tailor their results based on what they know about the customer doing the search, and social media networks want to recommend friends based on what they know about all of us. Both of these goals require building profiles on everyone, figuring out what their preferences are and where they tend to spend their time. They might require figuring out whether each of us is male or female, more likely to go to a sports game or a play, or about to buy a dishwasher or a television.

But these datasets and profiles could be used for far more than that. What if the FBI used them to try and figure out who is likely to commit a crime, or a company tries to learn their employees' religion or sexual orientation?

As they build ever-more sophisticated models based on ever-more accurate datasets, Data Scientists need to think about the ethics of what they're doing as well!

Investigate

Divide the class into groups of 3-4, and assign each group a different case study. Have each group choose one person to share back with the class.

- How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did (Forbes)
- Facebook 'likes' can reveal your secrets (CNN)
- Algorithmic Bias in Criminal Sentencing (Propublica)

(**Note:** The third article is quite long, but only the first half is needed for students to complete this activity.)

Have students complete Case Study: Ethics, Privacy, and Bias (Page 82).

Synthesize

Give students time to discuss and share back. Encourage students to share back differing views on the articles.

What are some commonalities and differences among the issues raised by these articles?

OPTIONAL: Can the class come up with a list of "Rules for Ethical Data Science"?

Going Deeper

1) For homework, have students write arguments in support of a randomly-chosen side of each case study. Select twelve students (two for each side of all three case studies), and have them debate in front of the class. Each side gets to make "opening" and "closing" arguments, and they take turns so that the closer for each side can respond to what the other side said. Then have the class vote on who was most convincing.

2) For homework, have students find their own articles about ethical issues in data science and write a one-page essay defending one side of it.