

# Lesson 14: Problem Solving with Big Data

## Overview

**Question of the Day: How is our data collected and why is it useful?**

In this lesson, students look at how data is collected and used by organizations to solve problems in the real world. Students are presented with three scenarios that could be solved using data and brainstorm the types of data they would need and how they could collect the data. Each problem is designed to reflect a real-world service that exists. After brainstorming, students watch a video about a real-world service and record notes about what data is collected by the real-world service and how it is used. At the end of the lesson, students record whether data was provided actively by a user, was recorded passively, or is collected by sensors.

## Purpose

In this lesson students see three examples of how the data problem solving process is used to solve real-world problems. This lesson expands the types of problems students think of as data problems and helps them to relate what they know about data to their real world experiences with common Internet services. The examples also provide an opportunity to reflect on the fact that in their own lives they are intentionally and unintentionally producing data that companies collect and use.

## Assessment Opportunities

1. **Give examples of how data is collected from sensors and tracking user behavior.**

Activity Guide, page 2: There should be multiple examples of tracking user behavior and sensor data in the second and third columns of the chart.

2. **Determine data that would be helpful in solving a problem, and how that data could be collected.**

Activity Guide: On each page, there should be reasonable descriptions of relevant data and how it could be collected. These do not need to be the same as the data actually collected by the companies, but should be relevant to the given problem and possible to collect.

## Objectives

Students will be able to:

- Determine data that would be helpful in solving a problem, and how that data could be collected.
- Distinguish between data that users intentionally and unintentionally produce.
- Give examples of how data is collected from sensors and tracking user behavior.

## Preparation

- Print copies of the activity guide
- Prepare projector if you will show videos to the whole class

## Links

**Heads Up!** Please make a copy of any documents you plan to share with students.

For the teachers

- **CSD Unit 5 - Data & Society** - Slides

For the students

- **Amazon and Data** - Video ([Download](#))
- **Data in the Real World** - Activity Guide
- **Netflix and Data** - Video ([Download](#))
- **Waze and Data** - Video ([Download](#))

### 3. Distinguish between data that users intentionally and unintentionally produce.

Activity Guide, page 2: The first and second columns should give multiple examples of data that has been actively and passively collected.

## Standards

Full Course Alignment

### CSTA K-12 Computer Science Standards (2017)

- **IC** - Impacts of Computing

## Agenda

#### Warm Up (5 minutes)

Journal

#### Activity (35 minutes)

Web Pix

Routz

Nyle

Reflection

#### Wrap Up (5 minutes)

Journal

## Teaching Guide

### Warm Up (5 minutes)

#### Journal

**Prompt:** In the last lesson, we saw how data could be collected through questions and used to make a recommendation. There are other ways our data can be collected that are happening all around us every day. What are some ways that apps, companies, or governments could be collecting data about you?

**Discuss:** Have students journal individually, then share with a neighbor, and finally discuss as a whole class. You can record their ideas on the board to refer back to later in the class.

#### Discussion Goal

**Goal:** This is primarily a brainstorm. Some students may bring a lot more prior knowledge than others to the class and at this point they haven't been explicitly taught anything about data collection outside of surveys. Aim primarily just to get ideas out and set the stage for the lesson. You're aiming to change focus from surveys that 10-20 people take to the vast amounts of data they might know is collected by modern technological tools.

#### Remarks

Great work. Today we're going to look at some real-world examples of how data is collected to solve problems. Keep an eye out for these ideas and think about whether you're seeing any new ones.

problems. Keep an eye out for these ideas and think about whether you're seeing any new ones.

**Question of the Day:** How is our data collected and why is it useful?

## Activity (35 minutes)

**Group:** Students may complete this activity individually or in pairs.

**Distribute:** Give students copies of activity guide

### Data in the Real World

#### Web Pix

**Web Pix:** Read the introduction to Web Pix on the first page. Give students a few minutes to write down their ideas and/or share them with a partner. Let students know that they will **not** be able to answer the question about Netflix yet. If students finish early, ask them to think of other types of data and how they could be collected.

**Share Out:** Allow students to quickly share out some ideas, then introduce the Netflix video.

#### Remarks

Many companies, such as YouTube, Facebook and Netflix, recommend videos and posts to users based on the data that they have collected about them. We're going to watch a short video about how Netflix does this. Pay attention to how Netflix collects data to help it recommend good videos.

**Display:** Watch the Netflix video



**Video: Netflix and Data**

#### Teaching Tip

**Accessing Videos:** These videos are embedded in the slides for this lesson and can also be downloaded from within this lesson plan. They can also be found in their own levels on Code Studio so students have access to these videos if they wish to watch them on their own or if they are absent for this lesson.

**Discuss:** Have students answer the third question about Netflix. Allow students to share with a partner, then discuss with a group the types of data that Netflix collects to help it make recommendations.

#### Discussion Goal

**Goal:** Students should understand that Netflix bases its suggestions not only on the survey data that users create, but also on the behavior of users as they use the site, such as what they watch and the ratings they give various videos.

#### Routz

**Routz:** Ask students to move on to the Routz problem. Again, they should take a few minutes to work individually or in pairs on the first two problems.

**Share Out:** Allow students to quickly share out some ideas, then introduce the Waze video.

**Display:** Watch the Waze video.



### Video: Waze and Data

**Discuss:** Allow students to share with a partner, then lead a short discussion on the types of data that Waze collects to help it find the best route.

#### Discussion Goal

**Goal:** Students should note that while some data is being intentionally added by the users, Waze also collects GPS/location data on each user without the user's active input. Sensor data, such as time, temperature, and location, are automatically collected by computers without a user actively adding them.

## Nyle

**Nyle:** Ask students to move on to the Nyle problem. Again, they should take a few minutes to work individually or in pairs on the first two problems.

**Share Out:** Allow students to quickly share out some ideas, then introduce the Amazon video.

**Display:** Watch the Amazon video.



### Video: Amazon and Data

**Discuss:** Allow students to share with a partner, then lead a short discussion on the types of data that Amazon collects to help it find the best products to suggest.

#### Discussion Goal

**Goal:** There are a few different aspects of the Amazon video that may be of interest.

- Just looking at something online produces data that can be used by advertisers or others.
- Amazon decides what you might buy by looking at similar users and using their behavior to predict yours
- Different types of data, such as "clicks", "likes" and "purchases", may be weighted differently
- The user's needs are not the priority. The advertiser's needs are.

Any of these topics are relevant to the lesson, but the most important thing for students to realize is that they are not always aware when they are producing data.

## Reflection

### Remarks

We've found lots of different ways that the computer can collect data. In general, data can be collected automatically from sensors, as the Waze app collects GPS data; it can be collected from users who are intending to produce data, as when people rate Netflix videos; and it can be collected from users who are not intended to produce data, as when Amazon records which products you look at.

**Reflection:** Ask students to fill out the chart at the bottom of the page that categorizes the different types of data. Allow students to compare their answers with a partner.

## Wrap Up (5 minutes)

### Journal

**Prompt:** Today we looked at three companies that collect data to solve problems. Brainstorm some other websites, apps, or companies you use or know about. What data are they collecting? How are they using it to solve a problem?

**Discuss:** Ask students to share their ideas with their classmates.

#### Discussion Goal

**Goal:** This prompt is meant to help students make connections between their personal experiences with data collecting services and what they have learned in this lesson. The goal here is to connect the examples students saw in today's lesson to other apps, websites, or services that they may be aware of. If you need to prompt students you might suggest they think of social media websites, media websites, useful apps they or their family uses, etc.

### Remarks

Today we saw some examples of different sources of data that real-world apps and websites use to solve problems. This also means these websites are collecting lots and lots of data - usually this is too much for a human to analyze! In the next lesson, we'll look at how we can help computers make decisions and learn from all of this data.