

Lesson 3: Data science & analytics – Problem solving with big data



Overview

In this lesson, students look at how data is collected and used by organizations to solve problems in the real world. Then students are presented two scenarios that could be solved using data and brainstorm the types of data they would want to solve them 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 two 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.



Agenda

Warm Up (3 min)

Activity (30 min)

- Routz
- Nyle

Wrap Up (10 min)

- Discussion
- Career Discussion



Objectives

Students will be able to:

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



Preparation

- Print copies of **PwC Data in the Real World – 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..

- **PwC Data in the Real World – Activity Guide**



Vocabulary

- **Big Data** - a broad term for datasets so large or complex that traditional data processing applications are inadequate.



Teaching guide

Warm Up (3 min)

Prompt

There are lots of ways that apps, companies, or governments might collect data. What ways to collect data are you already aware of?

Discuss: Have students brainstorm their ideas before sharing with the class. You can record their ideas on the board to refer back to later in the class.

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.

Activity (30 min)

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

Distribute: Give students copies of **PwC Data in the Real World - Activity Guide - Activity Guide**



Data in the Real World

Routz

Introduce students to the "Routz" problem as defined 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 Waze yet. If students finish early, ask them to think of other types of data and how they could be collected.

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

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.

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. After students have shared in their pairs, introduce the **Amazon video**.

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.

Wrap Up (10 min)

Discussion

Reflection: 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 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.



Teaching tip

Accessing Videos: It's recommended that the class watch the videos in this lesson together on a large screen or projector. They can be found in their own levels on Code Studio. Students also have access to these videos and so if they wish can watch them on their own too.



Remarks

Today we saw some examples of different sources of data that real-world apps and websites use to solve problems. Websites often ask you directly for data, but they might record your behavior online to collect data as well. In fact sometimes sensors like a GPS signal can collect data without you even knowing it.

Career discussion

Introduce yourself and your career:

- What do you work, what do you do, and what do you love most about your job?
- What or who inspired you?
- How did you get interested in computer science?
- Did you have a mentor?
- Share a story about how tech affects everyone
- Consider showing the inspirational **Data and Medicine - Video**.

Ask the students questions and leave time for Q&A.

- What jobs are they interested in, what are their favorite tech gadgets or apps, and how do they think they are built?
- Do the students have any questions for you?



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.



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.



This curriculum is available under a Creative Commons License (CC BY-NC-SA 4.0).

If you are interested in licensing Code.org materials for commercial purposes, [contact us](#).

This lesson is a derivative of "Problem Solving with Big Data" from Code.org's CS Discoveries Course, used under CC BY-NC-SA 4.0. This lesson is licensed under CC BY-NC-SA 4.0 by PwC.

This content is for general information purposes only, and should not be used as a substitute for consultation with professional advisors.

© 2020 PwC. All rights reserved. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.