

# Lesson 3: Data Science & Analytics

## -Problem solving with Big Data Facilitator Guide

**Objectives** In this lesson, students look at how data is collected and used by organisations to solve problems in the real world. Then students are presented with 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

#### Getting Started (10 min)

- Ways to collect data

#### Main Activity (40 minutes)

- Routz activity - show video
- Nyle activity - show video

#### Wrap-up (15 minutes)

- Brainstorm websites, apps, or companies you use or know about. What data are they collecting?
- Careers reflection and next steps

### Objectives

Students will be able to:

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

### Preparation/materials needed

- Print copies of **PwC Data in the Real World – Activity Guide**
- Prepare projector if you will show videos to the whole class

### Links

- **Video:** Why Are So Many Users Obsessed With Waze?  
<https://studio.code.org/s/pwc/stage/3/puzzle/2>
- **Video:** How Big Data Is Used In Amazon Recommendation Systems To Change Our Lives:  
<https://studio.code.org/s/pwc/stage/3/puzzle/3>

### Vocabulary

- **Big Data** - a broad term for datasets so large or complex that traditional data processing applications are inadequate
- **User Actively Provided Data** - The user actively provides data and understands that their actions are providing data to the company
- **User Passively Provided Data** - The user might not know that their behaviour on the website is generating data to solve a problem.
- **Sensor data** - Sensors automatically collect data about the world. Usually this is not measuring anything about people.

# Teaching guide

*This guide includes a suggested script for the session (in orange). However, please feel free to tailor and adapt this accordingly when delivering.*

## Section 1: Getting started (approx. 10 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?*

- Have students discuss their ideas before sharing with the class. You can record their ideas on the board to refer to later in the class.

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 scene 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.

**Outline** to the class that 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.

## Section 2: Main Activity (approx. 40 min)

Students may complete this activity individually or in pairs. Give students copies of PwC Data in the Real World - Activity Guide.

### Routz case study

**Introduce** students to the "Routz" problem as defined on the first page of the Activity Guide.

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](https://studio.code.org/s/pwc/stage/3/puzzle/2): <https://studio.code.org/s/pwc/stage/3/puzzle/2>

**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.

**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 the own levels on Code Studio. Students also have access to these videos and so if they wish can watch them on their own too.

### Nyle case study

Ask students to move on to the Nyle problem even if they haven't finished working on the Routz case study.

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: <https://studio.code.org/s/pwc/stage/3/puzzle/3>

### 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 behaviour 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 advertisers' 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.

**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.

**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 behaviour online to collect data as well. In fact sometimes sensors like a GPS signal can collect data without you even knowing it.

### Section 3: Wrap-Up (approx. 15 min)

**Reflection:** Today we looked at three companies that collect data to solve problems. 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 behaviour online to collect data as well. In fact sometimes sensors like a GPS signal can collect data without you even knowing it.

Think of some other websites, apps, or companies you use or know about. What form of data are they collecting? How are they using it to solve a problem? **Ask students to share their ideas with their classmates**

Sometimes people do not realize that they are providing data when they use technology. What's one example of data from above that a user may not want to share? Why might a user want to keep this data private? As students to **discuss their ideas together in groups and use the table in their activity guide to jot down ideas.**

#### Employability skills reflection:

What employability skills have you developed in this session? What might your next steps be in learning more and thinking about your career?

The slide outlines multiple employability skills as a prompt for pupils to reflect. Encourage them to pick three and think of three specific examples to support how they have displayed these skills in the session. This is good practice for job applications and interviews - particularly competency based application processes.

#### Want to find out more

If you have time you could click on the links provided, or just point the pupils towards these. It should be made clear that the opportunities outlined are just those at PwC, and pupils should be encouraged to look at other opportunities too.

Link for more info: More information about how PwC uses data and analytics to support clients can be found [here](#)

Link for opportunities in tech:

- Technology Degree Apprenticeships and Data Science Graduate Apprenticeships.  
(<https://www.pwc.co.uk/careers/early-careers/our-programmes/flying-start-degrees/technology.html>)
- School and College Leaver Apprenticeships at PwC  
(<https://www.pwc.co.uk/careers/early-careers/our-programmes/join-us-from-school.html>)

**Disclaimer:** This content is for general information purposes only, and should not be used as a substitute for consultation with professional advisors. Information available here is not to be relied upon as professional advice or for the rendering of professional services. This lesson is a derivative of "Problem Solving with Big Data" from [Code.org's CS Principles Course](#), used under [CC BY- NC-SA 4.0](#). This lesson is licensed under [CC BY-NC-SA 4.0](#) by PwC. [www.pwc.com/us/aypcurriculum](http://www.pwc.com/us/aypcurriculum)

Certain links in this Site connect to other Web Sites maintained by third parties over whom PricewaterhouseCoopers LLP has no control. PricewaterhouseCoopers LLP makes no representations as to the accuracy or any other aspect of information contained in other Web Sites. © 2019 PricewaterhouseCoopers LLP (a UK limited liability partnership). All rights reserved. 'PwC' refers to the UK member firm, and may sometimes refer to the PwC network. Each member firm is a separate legal entity. Please see [www.pwc.com/structure](http://www.pwc.com/structure) for further details.