

Lesson 3: Data Science & Analytics

-Problem solving with Big Data - Activity Guide

Name.....

Date.....

Big Data is a broad term for datasets so large or complex that traditional data processing applications are inadequate. As a class we're going to review some business case studies and think about how these companies could use big data effectively.

Routz

Routz is a company that helps people navigate traffic. It has maps of locations all around the world and knows real-time information about traffic conditions. When you type in a destination you want to go, it will give you the best route it can find.

Task: Routz wants to improve its data to provide even better directions. Your job is to help the company decide what data it could collect to improve those directions. Have a look at the questions below and put your ideas in the 'Answer' box.

Question	Answer
What data should Routz collect to improve its recommendations?	
Where could it get that data?	
Watch the video about Waze, and explain how it solves this problem.	

Nyle

Task: Nyle is a company that sells all kinds of products online. Users on the site can search for and order products that are interesting to them. Advertisers would like to advertise their products on Nyle to increase sales. Ideally they would only show ads to people who are likely to buy their products.

Have a go at answering the three questions below:

Question	Answer
What kind of information should Nyle collect about its users to make sure ads are relevant to their customers?	
How might it collect this information?	
Watch the video about Amazon and explain how it solves this problem	

Reflection/wrap up

There are many ways that data can be collected to solve problems by different companies. Fill in as many examples from your discussions today that fall into each of the categories below.

Form of data collection	Example of this data collection
<p>User actively provided: The user actively provides data and understands that their actions are providing data to the company</p>	
<p>User passively provided: The user might not know that their behaviour on the website is generating data to solve a problem</p>	
<p>Sensor data: Sensors automatically collect data about the world. Usually this not measuring anything about people.</p>	

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? **Discuss your ideas together in groups and use the box below to brainstorm any ideas.**

Notes:

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.

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 for further details.