Check Your Assumptions

Unit 4 Lesson 3 (U4L3)

For the Students

- Digital Divide and Checking Assumptions - Activity Guide

Table of Contents

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

Video: Google Flu Trends

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

***IN YOUR NOTES -** reply to the following prompt

Prompt (2-min): What are the potential beneficial effects of using a tool like Google Flu Trends?

-Share with your group

-Share out as a class

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

***Go to Google Classroom and select the following article (6-8 minutes):**

Wired Magazine - What Can We Learn from the Epic Failure of Google Flu Trends

***Thinking Prompt:**

While reading, think about the following:

-Why did Google Flu Trends (GFT) eventually fail?

-What assumptions did they make about their data or their model that ultimately proved not to be true?

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

***Discussion: Why did it fail? What did they assume?**

- Google Flu Trends worked well in some instances but often over-estimated, under-estimated, or entirely missed flu outbreaks. A notable example occurred when Google Flu Trends largely missed the outbreak of the H1N1 flu virus.
- Just because someone is reading about the flu doesn't mean they actually have it.
- Some search terms like "high school basketball" might be good predictors of the flu one year but clearly shouldn't be used to measure whether someone has the flu.
- In general, many terms may have been good predictors of the flu for a while only because, like high school basketball, they are more searched in the winter when more people get the flu.
- Google began recommending searches to users, which skewed what terms people searched for. As a result, the tool was measuring Google-generated suggested searches as well, which skewed results

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

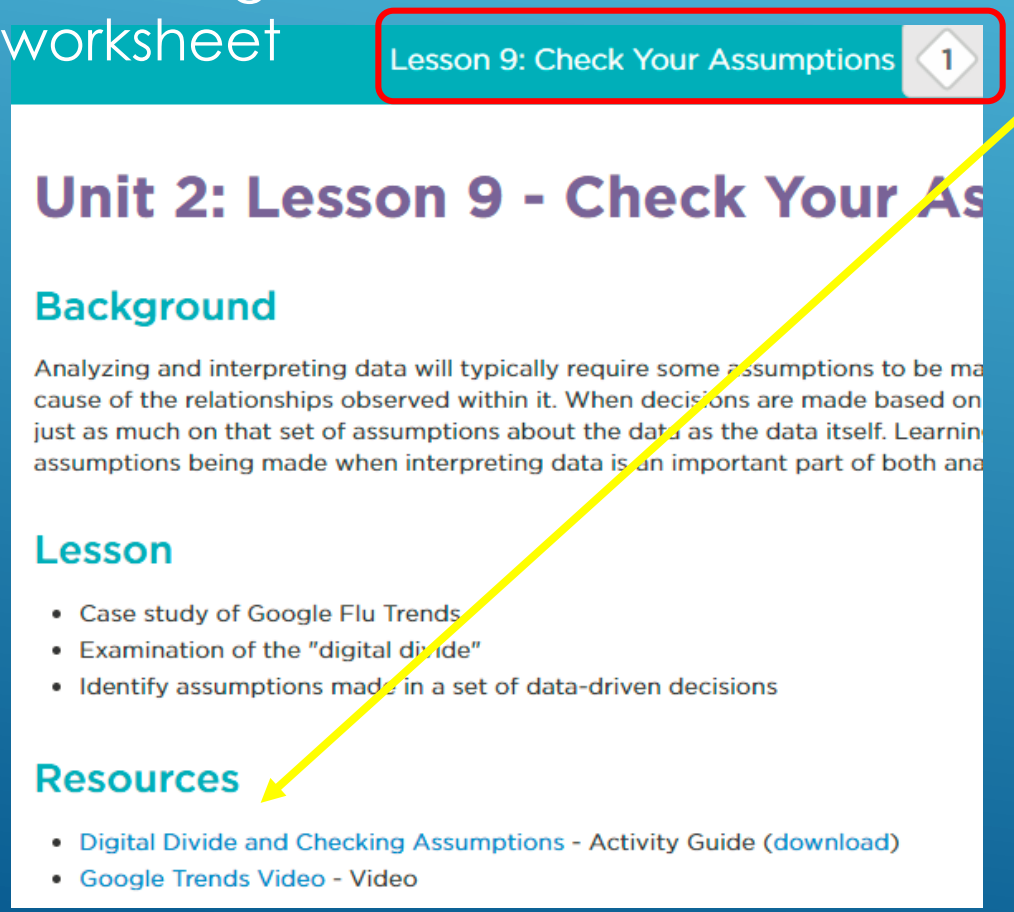
* The amount of data now available makes it very tempting to draw conclusions from it. There are certainly many beneficial results of analyzing this data, but we need to be very careful. To interpret data usually means making key assumptions. If those assumptions are wrong, our entire analysis may be wrong as well. Even when you're not conducting the analysis yourself, it's important to start thinking about what assumptions other people are making when they analyze data, too.

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

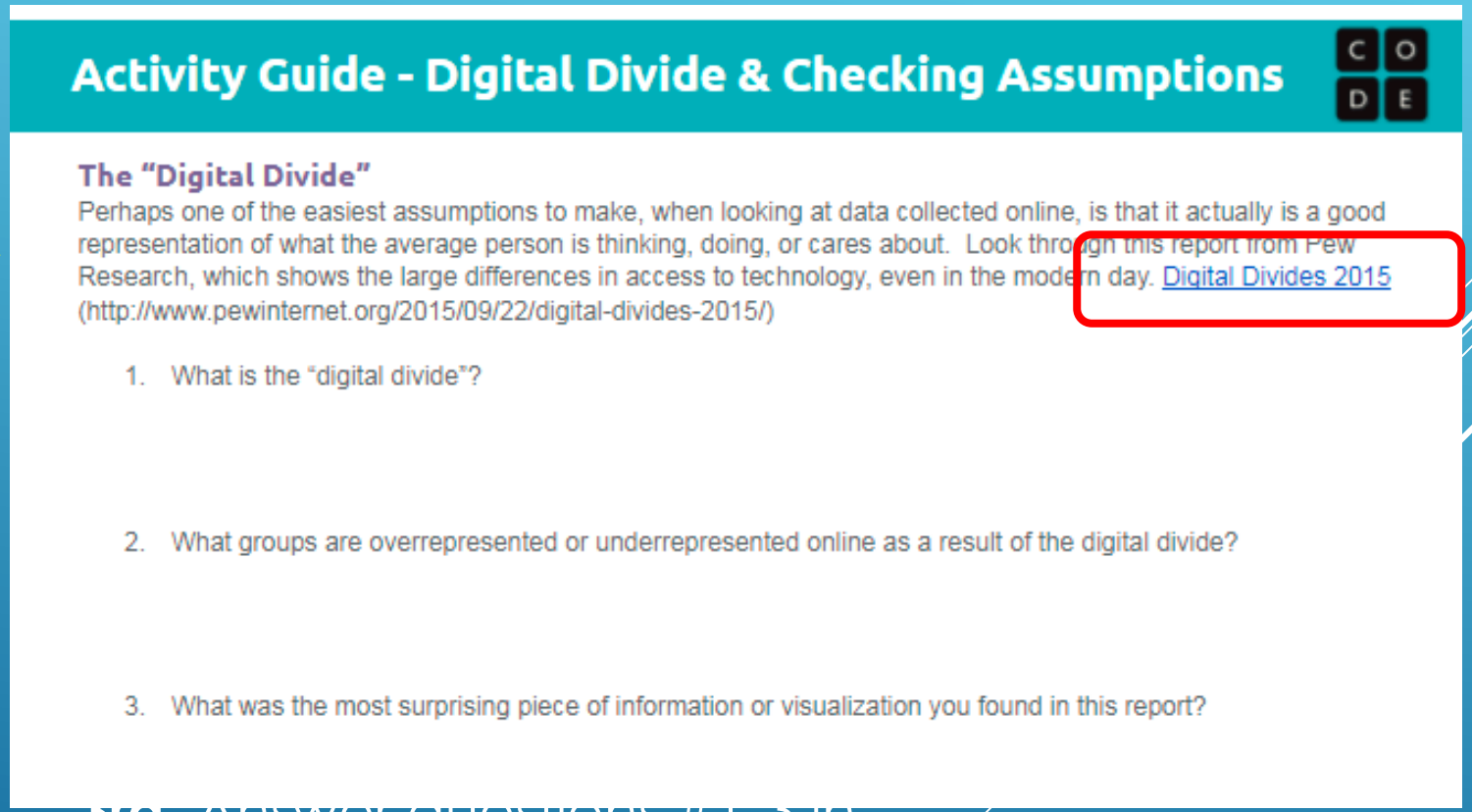
***Activity (25-minutes total)**

***Part 1: The Digital Divide**

1st: Navigate to the online worksheet



2nd: Click on "Digital Divides 2015," and examine the slideshow



3rd: Answer questions #1-3 in groups

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

*Activity (continued)

*Part 2: Checking Your Assumptions

1st: Read through your assigned scenario

2nd: In your groups, respond to the three questions on the back about your given scenario

Identifying Assumptions in Data Analysis

When you use data to make decisions, you need to be careful to **identify your assumptions** and **reflect on how those assumptions impact your analysis.**

Pick one of the scenarios below. With a partner, respond to the questions you find there about the assumptions made to conduct that analysis.

	Scenario	Data	Decision
#1	A city would like to more efficiently locate potholes that need to be filled.	The city builds an app that allows residents to report potholes from their smartphone.	The city will use this app as the primary method of identifying potholes.
#2	A news agency would like to predict the outcome of a coming election.	A social media company (e.g., Twitter) keeps track of how many times each candidate has been mentioned on the platform.	The news agency will use this data to predict the outcome of the election.
#3	A state government is trying to determine which issues are most important for the upcoming year.	The government creates an online survey where citizens can vote for the issue they care most about.	The government will use the results of this survey to help prioritize issues in the coming year.
#4	A chef is deciding where to open a new restaurant in the city.	A restaurant reviews website keeps track of the areas of the city that receive the most restaurant reviews.	The chef will use the data to choose the location for opening a new restaurant.

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

***Wrap-Up**

-Class discussion of your Activity Guide responses

Goal: for us to have a group understanding of what kinds of assumptions are being made to interpret the data

**Possible Assumptions:

- The data collected is representative of the population at large (e.g., ignoring the "digital divide").
- Activity online will lead to activity in the real world (e.g., people expressing interest in a candidate online means they will vote for him or her in real life).
- Data is being collected in the manner intended (e.g., ratings are generated by actual customers, instead of business owners or robots).
- Many other assumptions regarding data are possible

UNIT 4 LESSON 3 (U4L3)
CHECK YOUR ASSUMPTIONS

*