Welcome to the *All of Us* Data Browser Scavenger Hunt!

A data sandbox allows users to interact with data in an isolated and controlled environment without impacting or risking changing the data itself. In this training, the sandbox is the publicly available Data Browser for the *All of Us* Research Program.

Using a scavenger hunt model, we will explore and visualize data in the data browser, providing an interactive overview of the data types and potential questions that can surface using *All of Us* datasets.

This activity was adapted from the Hunter College Data Sandbox.

To get started, go to: http://www.researchallofus.org/nlm and use the Data & Tools tab at the top of the page.

1. What types of data sources are collected in the *All of Us* Data Set?

    Answer: Electronic Health Records, Biosamples & Bioassays, Surveys, Physical Measurements, Mobile Health Data (Wearable Devices)

    Navigation: Data & Tools → Data Sources

2. How many participants are enrolled in your state?

    Answer: Varies by state.

    Navigation: Data & Tools → Data Snapshots →   Scroll down to Geography for the interactive map

3. Where can you find the survey questions that participants respond to?

    Answer: The Survey Explorer

    Navigation: Data & Tools → Survey Explorer
    Now use the Data Browser: https://databrowser.researchallofus.org

4. During the past 12 months, what percentage of people were told by a healthcare provider or doctor's office that they did NOT accept their healthcare coverage?

Answer: 87.44% (166,320) of people were told their insurance was not accepted by a healthcare provider.

Navigation: Data Browser →   Health Care Access & Utilization Survey →   Scroll to and click: DURING THE PAST 12 MONTHS, were you told by a health care provider or doctor's office that they did not accept your health care coverage?

Search Across Data Types Keyword Search: Healthcare Coverage

5. What percentage of participants reported using marijuana at any point in their lifetime?

Answer: 49.67% (197,800) of participants answered Which Drugs Used: Marijuana Use

Navigation: Data Browser →   Lifestyle Survey →   Scroll to and click: In your LIFETIME, which of the following substances have you used?

Search Across Data Types Keyword Search: Marijuana

6. What was the top answer participants gave for the highest grade or year of school completed?

Answer: 25.44% of participants answered 1–3 years of college

Navigation: Data Browser → The Basics → Scroll to and click: What is the highest grade or school you completed?

Search Across Data Types Keyword Search: School Completed

7. What was the top response participants gave for "My neighborhood has several free or low-cost recreation facilities, such as parks, walking trails, bike paths, recreation centers, playgrounds, public swimming pools, etc."

    Answer: 33.48% of participants answered Strongly Agree

    Navigation: Data Browser → Social Determinants of Health → Scroll to and click: My neighborhood has several free or low-cost recreation facilities, such as parks, walking trails, bike paths, recreation centers, playgrounds, public swimming pools, etc. Would you say that you...

    Search Across Data Types Keyword Search: Recreation

8. What percentage of participants reported having a mood disorder?

    Answer: 34.94% or 89,020 with a mood disorder

    Navigation: Data Browser → Conditions → Keyword Search "Mood Disorder"

    Search Across Data Types Keyword Search: Mood Disorder

9. In July 2020, what was the highest reported answer for "How happy are you?"

    Answer: Moderately Happy (20,700; 51%)

    Navigation: Data Browser →   COVID-19 Participant Experience (COPE) Survey → Scroll down to General Well-Being →   Click: Choose the answer that best describes how you felt in the past month. In general, how happy are you?

    Search Across Data Types Keyword Search: Happy

10. How many genomic variants are found in the Breast Cancer gene BRCA2? *Hint: In the Genomics section, use the SNP/Indel Variants Search*

    Answer: 34,873

    Navigation: Data Browser → Under Genomics: SNP/Indel Variants → Search by gene, variant, rs number or genomic region "BRCA2"

    Search Across Data Types Keyword Search: BRCA2

11. What is the participant count for records of any Fitbit data by 2022?

Answer: 15,620 had any records for Fitbit data by 2022

Navigation: Data Browser → Under Measurements and Wearables: Fitbit → Select "Any Fitbit Data" at top left → Hover over 2022

Search Across Data Types Keyword Search: Fitbit

12. What percentage of participants have been proscribed oxycodone (oxycodone hydrochloride 5 MG Oral Tablet)?

Answer: 22.76% of participants

Navigation: Data Browser → Drug Exposures → Keyword Search "Oxycodone"

Search Across Data Types Keyword Search: Oxycodone

13. What percentage of participants have had a blood draw recorded on their EHR (collection of venous blood by venipuncture)?

Answer: 40.14% of participants

Navigation: Data Browser → Procedures → Keyword search "Venipuncture"

Search Across Data Types Keyword Search: Venipuncture

14. What is the highest hemoglobin value (Mass/Volume) in blood for men versus women?

Answer: 14–17 for males and 11–14 for females

Navigation: Data Browser → Labs & Measurements → Keyword search "Hemoglobin"

Search Across Data Types Keyword Search: Hemoglobin

15. For participants who rate their satisfaction with social activities and relationships as VERY GOOD, what is the age range with the highest count?

   Answer: 60–69, 27,440 participants in this age group answered VERY GOOD

   Navigation: Data Browser →Overall Health Survey → In general, how would you rate your satisfaction with your social activities and relationships? → Select Social Satisfaction: Very Good → Click on the graph icon → Select Age when Survey was Taken

   Search Across Data Types Keyword Search: Relationships

16. Now that you have explored each data type in the browser share with the group an interesting observation you made using the data browser.

ORAU
Further. Together.