

Using R to Co-Create an Inclusive Data Analysis Approach with the HBCU Health Equity Data Consortium

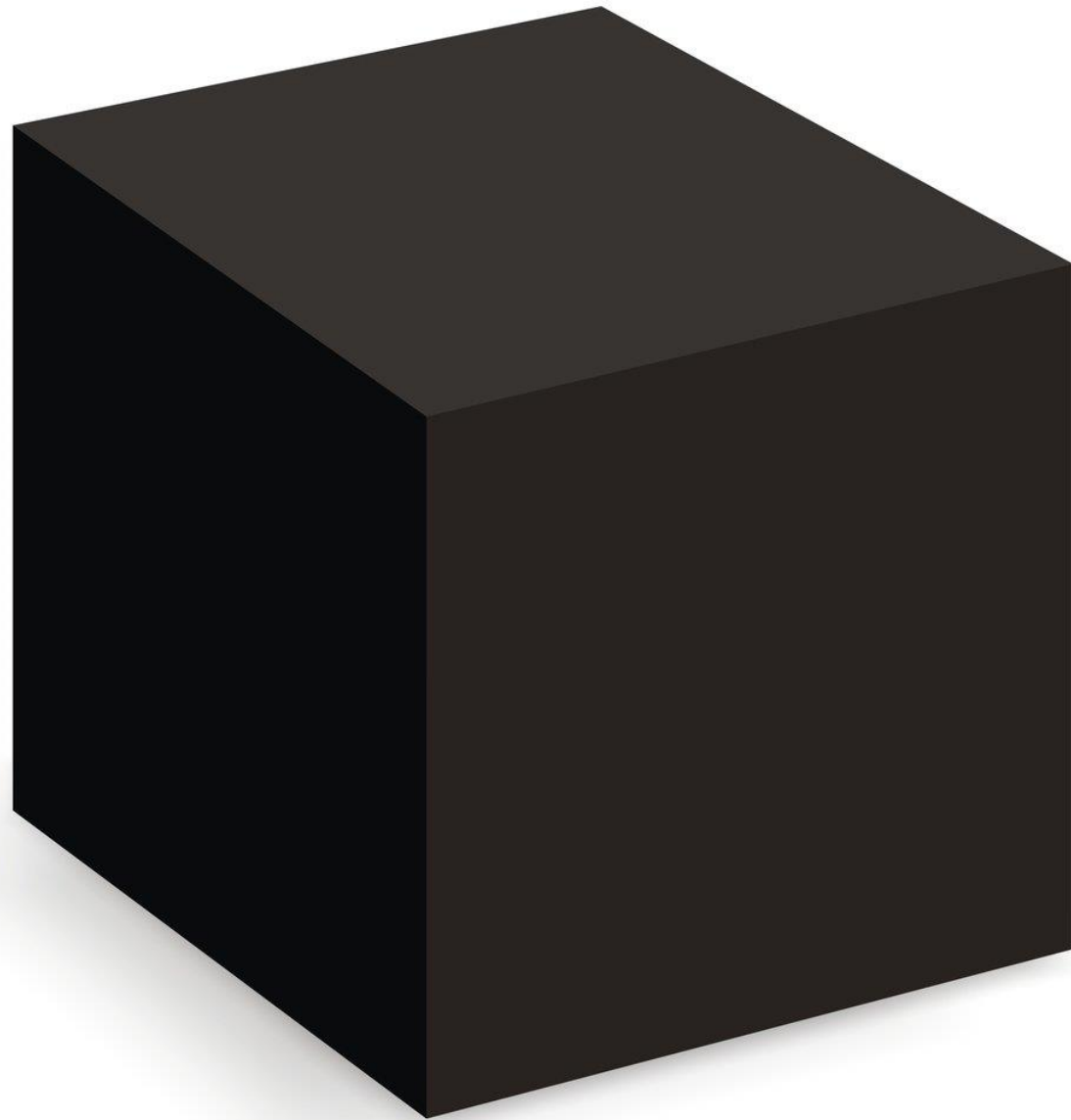
Lois Adler-Johnson

Data Analyst & Advisor @ North Carolina Institute for Public Health



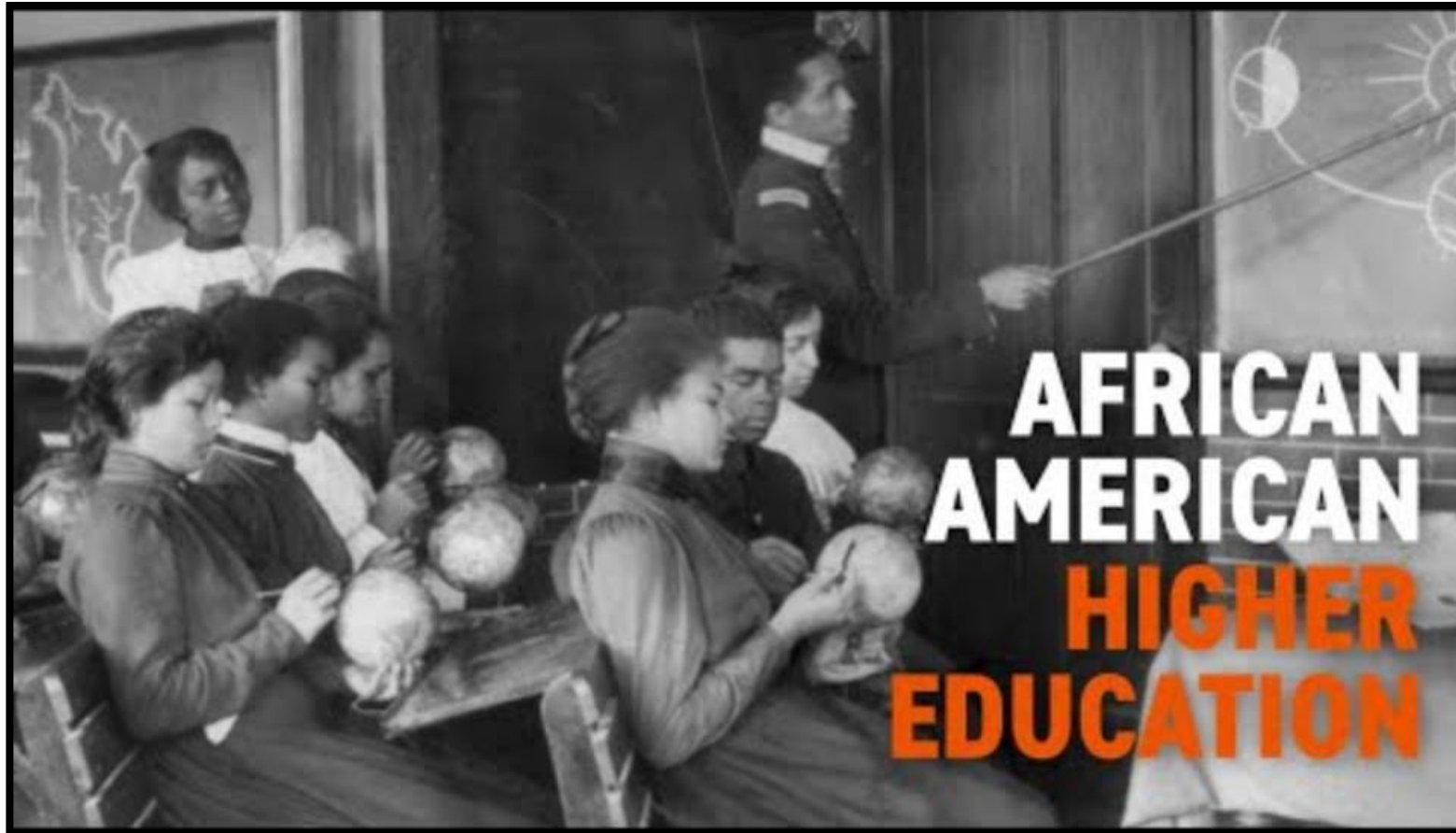
GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

North Carolina Institute for Public Health



HBCU Health Equity Data Consortium

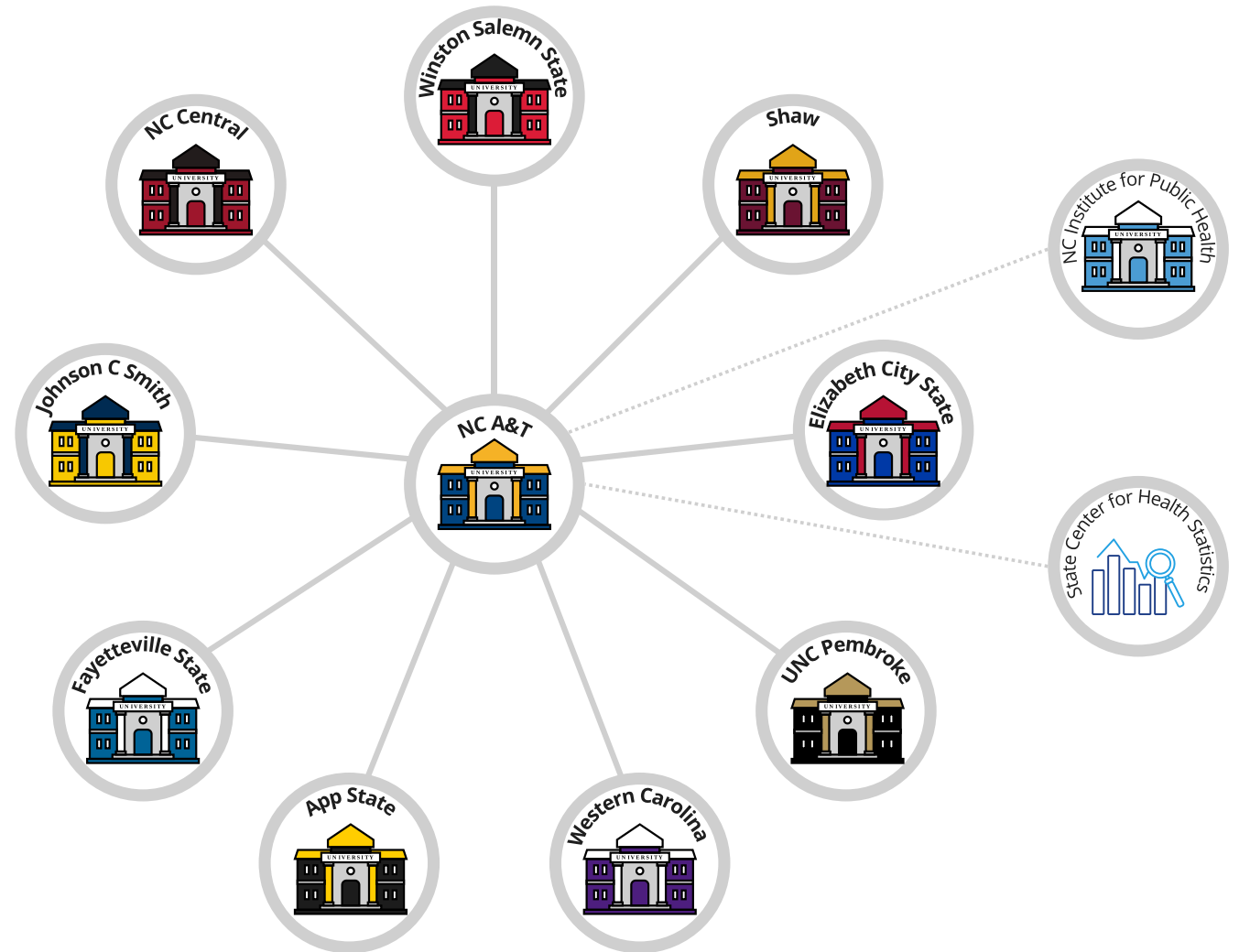
Historical Context




[African American Higher Education by Black History in Two Minutes](#)

HBCU HEDC Background

- Formed in 2022
- Emphasizes equity approaches to:
 - community engagement
 - minority serving institution infrastructure development
 - reciprocal data literacy
 - health disparity research
 - data equity for historically marginalized populations



COVID-19 Impact Survey



COVID-19 IMPACT SURVEY

WE NEED TO HEAR FROM YOU!


COMPLETE YOURS TODAY!

SURVEY CLOSES ON 12/01/23!

Scan the QR code or visit
this link to complete your
consent form and your
survey today!

<https://tinyurl.com/nc-covidimpact>


SCAN ME



The HBCU Health Equity Data Consortium is conducting the **COVID-19 Impact Survey** to better understand how the COVID-19 pandemic has impacted households in North Carolina.

Time's running out! Eligible participants who complete the survey by 12/01/2023 will receive a \$30 gift card!

HBCU Health Equity Consortium



COVID-19 Impact Survey

Eligibility

We first need to ask some questions to ensure you are eligible to participate.

S1. Are you at least 18 years of age?

- ☐ Yes
- ☐ No

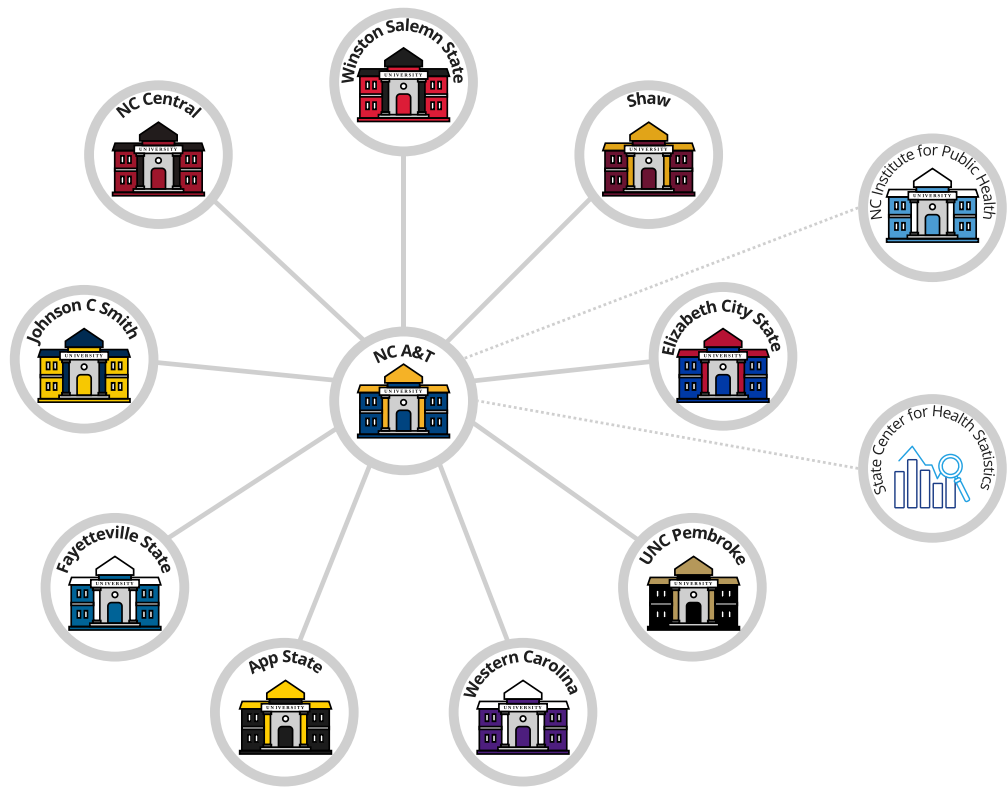
S2. Has anyone in your household already responded to this survey? We are asking for only 1 response per household.

- ☐ Yes
- ☐ No
- ☐ Don't Know

S3. Is your primary residence in NC Medicaid Region 2, 3, or 6? See map below.

- ☐ Yes, my primary residence is in region 2 (green)
- ☐ Yes, my primary residence is in region 3 (orange)
- ☐ Yes, my primary residence is in region 6 (purple)
- ☐ I do not live in one of the highlighted regions





**COVID-19
IMPACT SURVEY**
WE NEED TO HEAR FROM YOU!

COMPLETE YOURS TODAY!

SURVEY CLOSES ON 12/01/23!

Scan the QR code or visit this link to complete your consent form and your survey today!

<https://tinyurl.com/nc-covidimpact>

SCAN ME

The HBCU Health Equity Data Consortium is conducting the **COVID-19 Impact Survey** to better understand how the COVID-19 pandemic has impacted households in North Carolina.

Time's running out! Eligible participants who complete the survey by 12/01/2023 will receive a \$30 gift card!

HBCU Health Equity Consortium

Data Analysis Workgroup





Workgroup Formation

- Developed brief participation survey to identify:
 - leads from **each** institution
 - how folks would like to engage
- Convened at biweekly meetings throughout the **entire** project

HBCU Health Equity Data Consortium: Data Analysis Workgroup Participation Survey

Complete this survey to indicate your preferred level of involvement with the data analysis workgroup.

lois.c.johnson@gmail.com [Switch account](#) 

 Not shared

* Indicates required question

Data Analysis Workgroup Participation

While we invite everyone to participate in our data analysis workgroup, we'd like to identify one data analysis lead from each member and ancillary institution. We think this will ensure representation of all institutions and make space for consortium-wide contributions to the process of analyzing and interpreting the results of the COVID-19 Impact Survey. **Please work within your institution to identify your data analysis lead prior to completing this survey.**

How would you like to be involved in the data analysis workgroup? Check all that * apply.

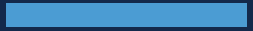
- ☐ Join biweekly data workgroup meetings
- ☐ Participate in discussions regarding analysis methods and data interpretation
- ☐ Prepare data for data analyses (i.e., data dictionary, merging raw data and phone data, data cleaning, and variable creation)
- ☐ Conduct data analyses (i.e., descriptive statistics, bivariate analyses, and multivariate analyses)
- ☐ I would not like to participate in the data analysis workgroup.
- ☐ Other: _____

[Submit](#) [Clear form](#)

Choosing & Using R

- Variety of preferred, mostly licensed programming languages among workgroup members
- Free and open-source!
- Literate programming
- Research reproducibility
- Volume of self-learning resources

**Snippets of our work
together**



Developed Shared Code

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for data import, cleaning, and preliminary figures. The code includes comments and functions for importing data, cleaning it, and creating preliminary figures.
- Environment Pane:** Shows the current environment with variables like `analysis_file...`, `phase1_clean`, and `phase2_clean`, each with 573 variables.
- Files Pane:** Lists files in the project, including `01_import-data.R`, `02_data-cleaning.R`, and `03_preliminary-figures.R`.
- Console:** Shows the output of the R code, indicating that the data is clean.

```
8  
9 # IMPORT DATA ####  
10 source("Code/01_import-data.R")  
11 rm(eligible, postcard_responses, potential_spam)  
12  
13 # CLEAN DATA ####  
14 analysis_file_clean <- analysis_file %>%  
15   mutate(q01=ifelse(as.numeric(q01)>99, as.numeric(NA), as.numeric(q01)))  
16  
17 # Individual Demographics ####  
18  
19 # > Region ####  
20 mutate(region=case_when(  
21   str_detect(s3, "1") ~ "1",  
22   str_detect(s3, "2") ~ "2",  
23   str_detect(s3, "3") ~ "3",  
24   str_detect(s3, "4") ~ "4",  
25   str_detect(s3, "5") ~ "5",  
26   str_detect(s3, "6") ~ "6",  
27   str_detect(s3, "I do not") ~ "I do not live in one of the highlig  
28   TRUE ~ as.character(NA))) %>%  
29  
30 # > Age ####  
31 mutate(age_group=case_when(  
32   q01 <= 19 ~ "18-19",  
33   q01 >= 20 & q01 <=29 ~ "20-29",  
34   q01 >= 30 & q01 <=39 ~ "30-39",  
35   q01 >= 40 & q01 <=49 ~ "40-49",  
36   q01 >= 50 & q01 <=59 ~ "50-59",  
37   q01 >= 60 & q01 <=69 ~ "60-69",  
38   q01 >= 70 & q01 <=79 ~ "70-79",  
39   q01 >= 80 ~ "80 or over",  
40   TRUE ~ as.character(NA))) %>%  
41 # Order  
42 mutate(age_group=factor(age_group, levels= c(  
43   "18-19", "20-29", "30-39", "40-49", "50-59", "60-69", "70-79", "80 or  
44  
45 # > Years in Residence  
46 mutate(years_in_residence = as.numeric(q02)) %>%  
47  
48  
13:1 # CLEAN DATA -
```

Environment Pane:

Variable	Obs.	Variables
analysis_file...	obs.	of 573 variables
phase1_clean	obs.	of 573 variables
phase2_clean	obs.	of 573 variables

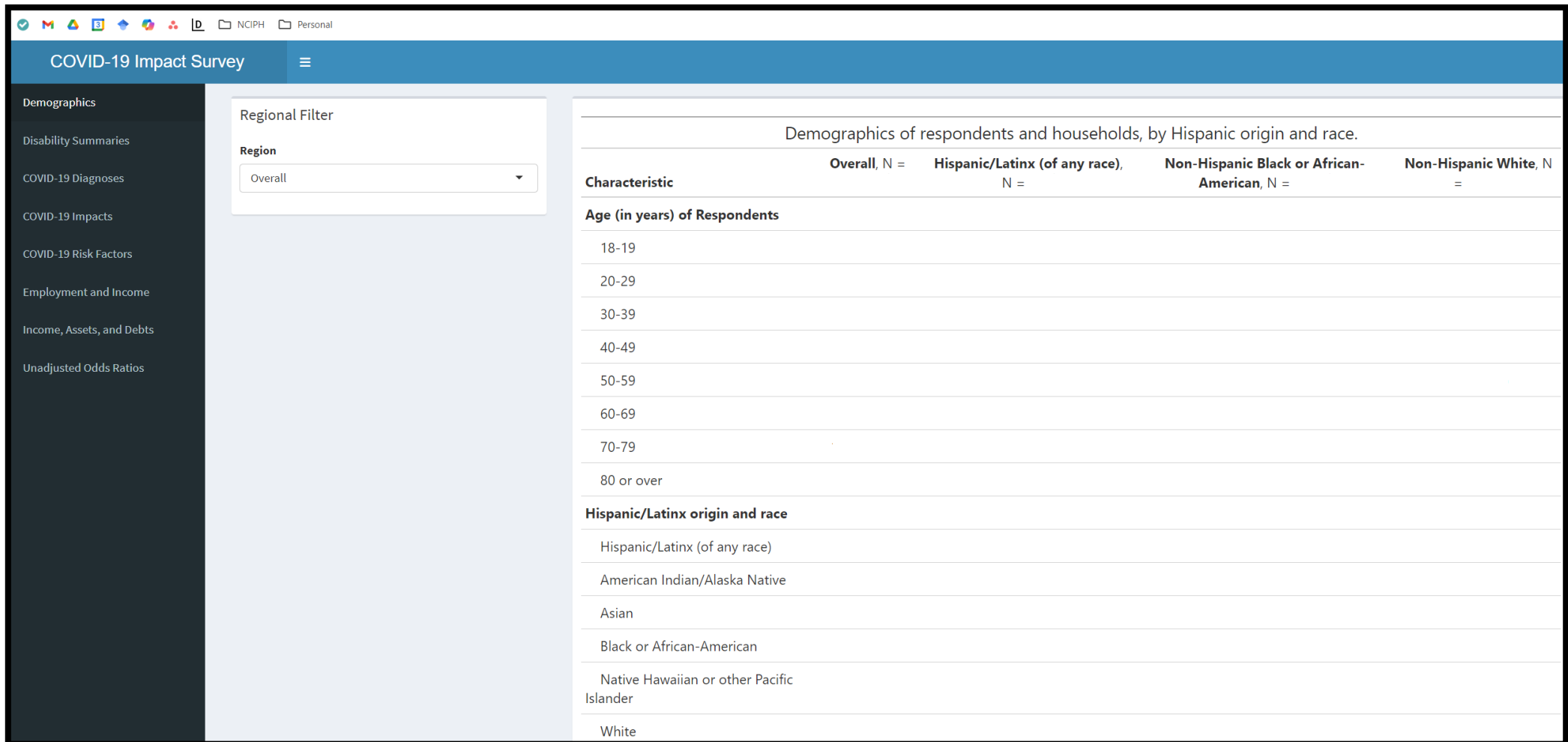
Files Pane:

Name	Size	Modified
..	11.9 KB	Jun 28, 202
00_libraries-and-constants.R	1.5 KB	Apr 25, 202
01_import-data.R	4.5 KB	Mar 28, 202
01p2_import-data.R	4.2 KB	Nov 16, 202
02_data-cleaning.R	60 KB	Apr 24, 202
02p2_data-cleaning.R	34.8 KB	Oct 26, 202
03_preliminary-figures.R	2.6 KB	Apr 18, 202
04_acs-datapull.R	6.6 KB	Jan 22, 202
05_covid19-missingness.R	2.2 KB	Aug 11, 202
06_summary-tables (dissemination).R	13.1 KB	Jul 3, 2024
06_summary-tables_county.R	8.1 KB	Apr 25, 202
06_summary-tables.R	12.7 KB	Mar 5, 2024

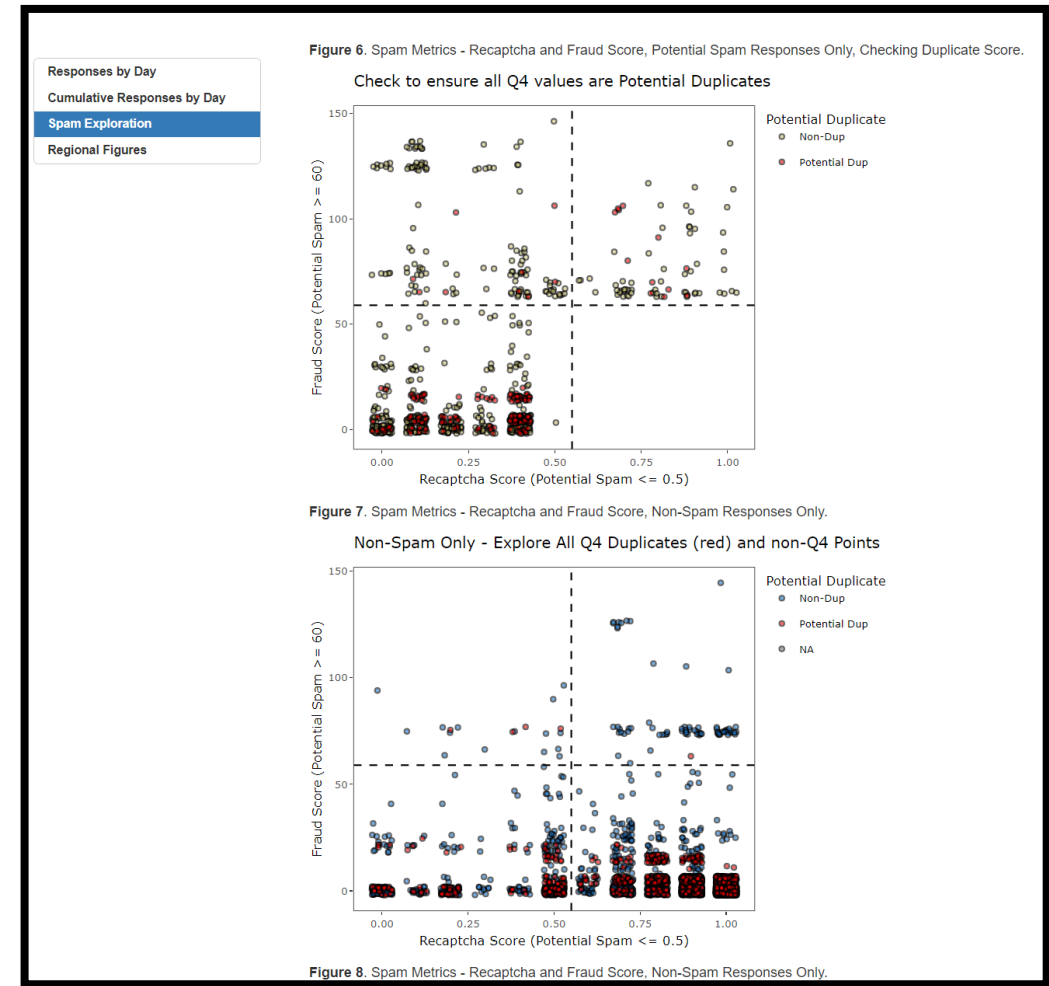
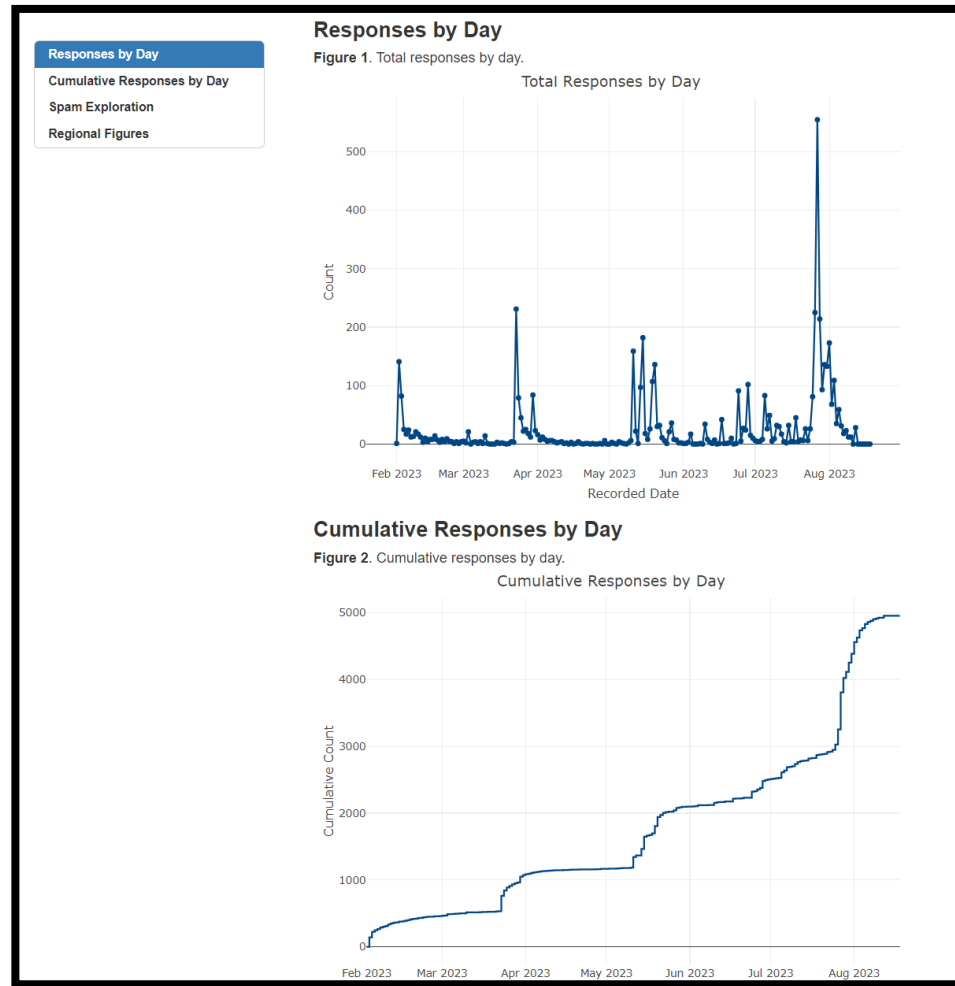
R Workshops



Shiny Dashboard



Markdown



Quarto

COVID-19 Diagnoses, Risks, and Impacts

Table 3. COVID-19 diagnoses, risks and impacts, by Hispanic origin and race (Overall).

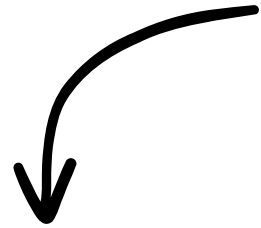
Characteristic	Overall, N =	Hispanic/Latinx (of any race), N =	Non-Hispanic Black or African-American, N =	Non-Hispanic White, N =
COVID-19 Diagnoses				
Household COVID-19 diagnoses				
Household with no COVID-19 diagnosis				
Household with COVID-19 diagnosis				
Severity of COVID-19 illness among HH members who tested positive				
Mild illness				
Moderate illness				
Severe illness but not hospitalized				
Severe illness requiring hospitalization				
Severe illness resulting in death				
COVID-19 symptoms that last longer than 4 weeks				
Co-morbidities				
Co-morbidities among household members				
None				

Table of contents

- Overall
- Demographics
- Disability Summaries
- [COVID-19 Diagnoses, Risks, and Impacts](#)
- Income, Assets, Debts, and Wealth
- Region 2
- Region 3
- Region 6

Impact and Recommendations

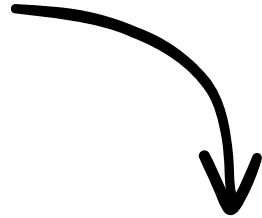
Shared code
+ training



Transform
raw results

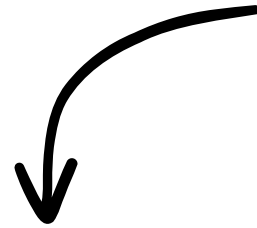


Generate
descriptive
statistics



Conduct
univariate
regression

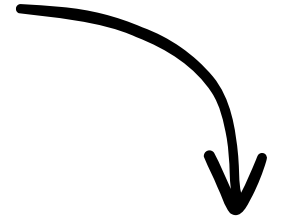
R Markdown,
Quarto, and
Shiny output



Make coding
& analysis
decisions

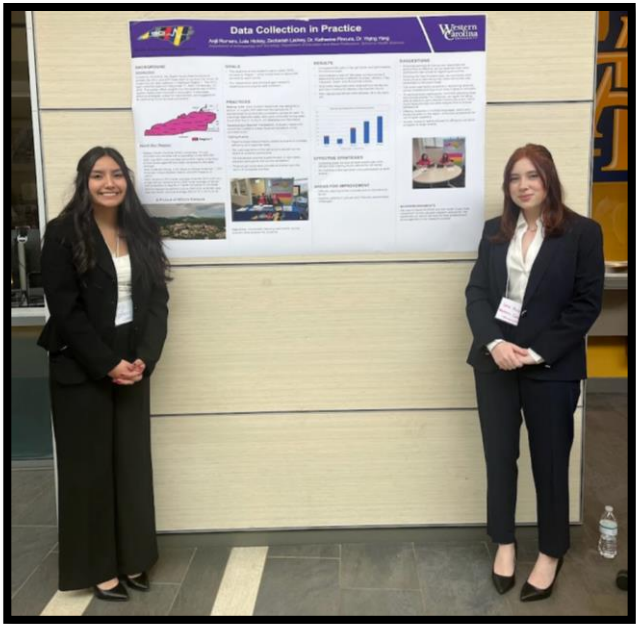


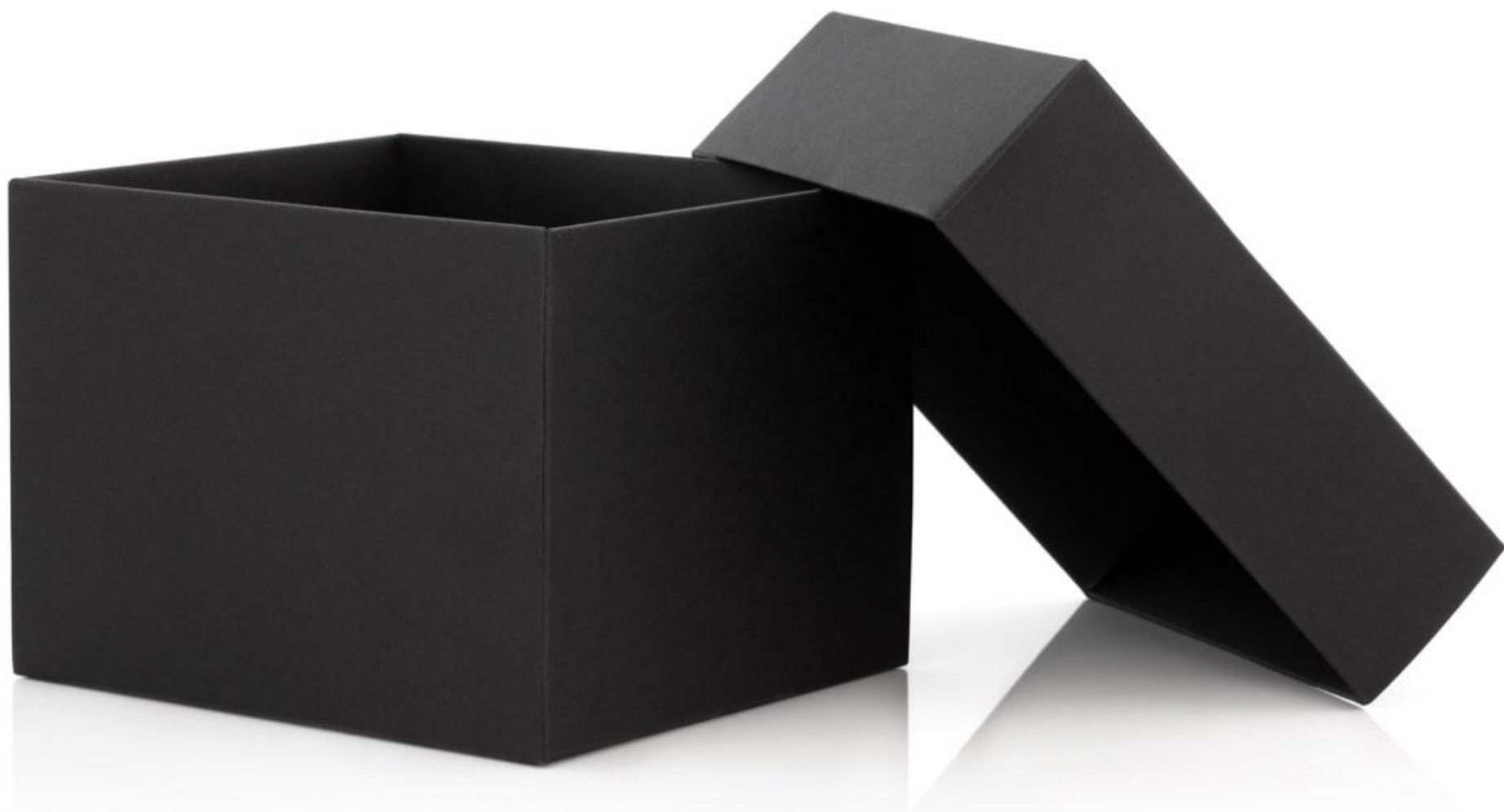
Brainstorm
exploratory
analyses



Disseminate
key findings

Student Event





Partnership



Recommendations

Break out of our
silo

Demystify data
science

Promote data
democratization

Form data
workgroups
tailored to project
context

Engage
participants
throughout the
data lifecycle



Data Equity Resources



Acknowledgements

CONSORTIUM HUB & CONSULTANTS

North Carolina A&T State University

Brandy Baldwin

Nia Bright

Breyana Davis

SydKni P. Haley

Erica McAlpine

Jaela Melton

Ayanna Proctor

Dr. Checo Rorie

★ **Dr. Ashley Sanderlin**

Dr. Miriam Wagner

Dr. Padonda Webb

NC Institute for Public Health

Lois Adler-Johnson

Adrienne Hall

Jaquayla Hodges

Destiny James

Lauren Jones

Dr. John Wallace

Samuel Dubois Cook Center on Social Equity

Dr. William A. Darity, Jr.

NC State Center for Health Statistics

Willysha Jenkins

Reagan Jones

Vincent Lella

DeNita Murdock-Nash

Jovan Slaughter

Joseph Watkins

PHASE 1 INSTITUTIONS

Elizabeth City State University

★ **Dr. Scott Bradshaw**

Dr. Anthony Emekalam

Johnson C. Smith University

Dr. Karen Butler

Dr. Shayna Duncan

Dr. Melvin Herring

★ **Dr. Sabina Otienoburu**

Winston-Salem State University

Marian Anderson-Booker

Dr. Hartensia Davis-Bailey

Dr. Jill Keith

Dr. LaVie Leasure

Dr. Tennille Presley

Dean Leslee Shephard Battle

★ **Dr. Cynthia Williams Brown**

PHASE 2 INSTITUTIONS

Appalachian State University

Dr. Margaret Brown

Dr. Adam Hege

★ **Dr. Martie Thompson**

Shaw University

Dr. Anderton-Georgie

Dr. Dorothy C. Browne

★ **Dr. Melvin Jackson**

Chetna Kumari

Shelby Palmer

Dr. Kimberly Raiford

Dr. Jerome Wilson

North Carolina Central University

★ **Dr. Nicole Diggs**

Dr. Irene Doherty

Fayetteville State University

Dr. Stephanie Hodges

★ **Dr. Sherry Leviner**

Dr. Christy Swinson

The University of North Carolina at Pembroke

Dr. Jacqueline Davis

Dr. Eva Skuka

★ **Dr. Joe West**

Western Carolina University

Dr. Terri Durbin

Dr. Katie Pincura

★ **Dr. Yiqing Yang**



Health Equity Data Consortium

Bold = Workgroup member

★ **= Workgroup Lead**

Thank You

Lois Adler-Johnson

Data Analyst & Advisor

lois.johnson@unc.edu

[linkedin.com/in/loiscjohnson](https://www.linkedin.com/in/loiscjohnson)



GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

North Carolina Institute for Public Health