

broadcast

COVID-19 Data Project:

Health Equity Handbook

Guide for Health Equity Data Interns

06.11.2020

Created by Sabine Benzakour, Audrey Caprio, Jasmine Johnson, Kazi Maisha, & Mel Terry

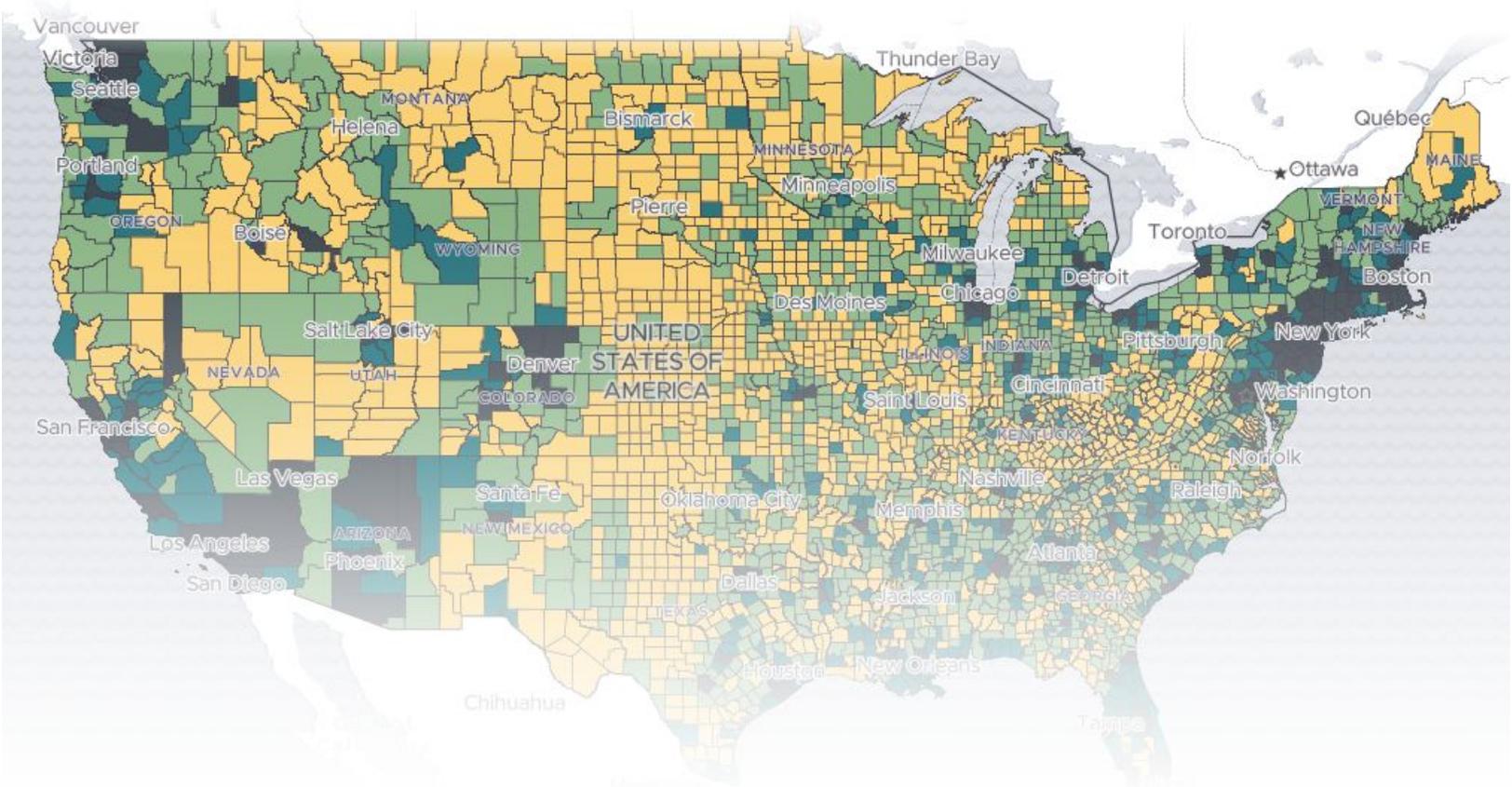


Table of Contents

Overview

About BroadStreet

Growth of The COVID-19 Data Project

Quick Start Guide

- Definitions

- County Dashboard Evaluation

- Data Entry Roles, Method, & Tips

Definitions

Region and Block System

County Dashboard Evaluation

Health Equity Data Track

- Goal

- Data

- Instructions/Links

Data Entry

- Health Equity Team Roles

 - Health Equity Team Leads

 - Health Equity Acting Leads

 - Health Equity Quality Assurance Leads

 - Health Equity Data Entry Interns

- Health Equity Data Entry Spreadsheet

- Data Sources

- Data Collection & Input

 - State-Specific Information

Data Entry Challenges

- Varying Reporting Methodologies

 - Percentages vs. Raw Numbers

 - Race and Ethnicity Categorization

- Untimely Data Releases

- Issues with Disease Reporting Systems

Frequently Asked Questions

References

Overview

Thank you for joining us for The COVID-19 Data Project. To our knowledge, this is the first time in history that a research-ready, longitudinal dataset has been created for a pandemic in near real-time. The primary goal is to collect county-level COVID-19 cases and deaths for the U.S. since the onset of the outbreak in North America. Being a part of this team is a unique opportunity to experience the unfolding of a rapidly spreading infection into cities and across communities. Through keeping vigil over the numbers, we are witness to the consequences of both action and inaction. We will examine what we see and will share it in the public domain to educate and inform ourselves, our families, our communities, and researchers for years to come. We are glad to have you.

~ *Tracy Flood*
CEO and Cofounder of BroadStreet

About BroadStreet

[BroadStreet](#) is a data and software company for social good. We believe that all people deserve access to the best information possible. BroadStreet builds our software for those working towards community change. We strive to ensure that data is accessible, affordable, and delightful to explore. BroadStreet has thousands of members all over the country. Members include students, non-profit leaders, and executives of multi-state and national organizations. We are proud of the range of our members.

BroadStreet was founded in the midwest by three co-founders: James Walters, Tom Schmitt, and Tracy Flood. The three co-founders met at a [Milwaukee Open Civic Data Initiative](#). The midwestern roots and collaborative, open-data values have served as the foundation of BroadStreet. Upon this foundation, we build software in an iterative and continuous way, always improving to meet the needs of our members.

Growth of The COVID-19 Data Project

The COVID-19 Data Project began as an in-house project by BroadStreet to collect county-level coronavirus case (#) and death (#) totals. As the pandemic spread across North America, 50 new Emergency Data Relief Interns were recruited to the team. This number doubled a week later as we grew to meet the expanding need for research-quality datasets. Now we have over 300 interns working in various data entry roles and special projects.

Quick Start Guide

Definitions

- Health equity is a subset of equity and addresses issues around fair opportunities to attain full health potential
- Race is a person's self-identification with one or more social groups
- Ethnicity is defined as a category of people who identify with each other on the basis of similarities such as cultural traditions, history, ancestry, language, and nation

County Dashboard Evaluation

- On a regular basis, new interns will help conduct a county dashboard evaluation to identify counties/parishes/territories that are newly reporting health equity data
- To complete your assigned counties, google the county/use the NACCHO link provided, input the appropriate number (**0**, **1**, or **2**) into the spreadsheet cell, and add the website to the spreadsheet in the appropriate column

Data Entry Roles, Method, & Tips

- Within the Health Equity track, there are four roles: Team Leads, Acting Team Leads, Quality Assurance Leads, & Data Entry Interns
- Data entry interns store their COVID-19 data in the [Health Equity Data Entry spreadsheet](#) which contains their county assignment and tabs with each month's data in it
- Tabs with data entry in it will contain the county name, county website to find the data, and each day broken down into sub-columns based on racial and ethnic categories
- Data sources will be listed next to each county
 - IMPORTANT: Use county-specific data NOT state-specific data
- Tips
 - When no data is reported for a category of race/ethnicity, use a "-" not a 0
 - If data is only reported as a percent, multiply it by the total number of cases
 - Use notes NOT comments to mark any discrepancies in data
 - If data includes residential and non-residential #s, only report residential #s

Data Input Flowchart:



Definitions

Race is a social grouping of people who have similar physical or social characteristics that are generally considered by society as forming a distinct group¹

Ethnicity is a social group that shares a common and distinctive culture, religion, language, or the like²

Socioeconomic status is the social standing or class of an individual or group which is often measured as a combination of education, income, or occupation³

Equity is the absence of avoidable, unfair, or remediable differences among groups of people, whether those groups are defined socially, economically, demographically or geographically or by other means of stratification⁴

Health equity implies that everyone should have a fair opportunity to attain their full health potential and that no one should be disadvantaged from achieving this potential⁴

Health disparity is a particular type of health difference that is closely linked with social or economic disadvantage⁵

Region and Block System

States are organized into six regions. As of right now, only certain states and counties report health equity data, but the Department of Health and Human Services has mandated the reporting of health equity data by the county level for each state starting August 1st.⁶ As of September 26th, 2020, 1,057 counties are currently reporting health equity data.

County Dashboard Evaluation

The first responsibility of an intern on the Health Equity track is to determine if their [assigned U.S. counties](#) report COVID-19 case counts by race and ethnicity, and to state whether they do or do not in the [UPDATED Equity Dashboard spreadsheet](#). County-level COVID-19 counts can be located on [state health department websites](#), county websites, and county health department websites (which can be found using Google or [the National Association of County Health Officials \(NACCHO\)](#)).

To report their findings, interns enter data into the Equity Dashboard spreadsheet. This spreadsheet contains tabs that are named after U.S. regions, such as the “Southeast Counties” tab or “River Valley Counties” tab. Each of these tabs contain the following columns and subcolumns:

- *Volunteer* - Name of intern
- *County, State* - County and state
- *Link to county website (if COVID info)* - Website that contains COVID-19 data by race and ethnicity for a specific county or state
- *Demographics of Those Tested*
 - *Positive Cases by Race* - Status of whether or not there is data reported for positive case counts by race for a specific county
 - *Negative Cases by Race* - Status of whether or not there is data reported for negative case counts by race for a specific county
 - *Total Cases by Race* - Status of whether or not there is data reported for positive + negative case counts by race for a specific county
 - *Positive Cases by Ethnicity* - Status of whether or not there is data reported for positive case counts by ethnicity for a specific county
 - *Negative Cases by Ethnicity* - Status of whether or not there is data reported for negative case counts by ethnicity for a specific county
 - *Total Cases by Ethnicity* - Status of whether or not there is data reported for positive + negative case counts by ethnicity for a specific county

Under each subcolumn, interns use the following codes to report the status of each county:

- 0 if the data is not available
- 1 if county-level data is available
- 2 if the data is incomplete or outdated

As of September 2020, these tabs only contain the following columns:

- *Volunteer*
- *County/State*
- *Link to county website (if COVID info)*
- *Demographics of Those Tested*
 - *Positive Cases by Race*
 - *Positive Cases by Ethnicity*
- *Additional Information*

Data Entry

As stay-at-home orders expire throughout North America, it is essential that we examine COVID-19 data by race and ethnicity at the county level. According to the [Centers for Disease Control and Prevention \(CDC\)](#), communities of color are being affected disproportionately by COVID-19 due to economic and social conditions such as residential segregation, systemic inequalities, and the inaccessibility of quality health care.⁷ Collecting and analyzing county-level COVID-19 data by race and ethnicity helps us to address the needs of the most vulnerable communities; this, in turn, benefits every community affected by COVID-19.

Health Equity Team Roles

The health equity team is divided into four distinct roles: Health Equity Team Lead, Health Equity Acting Lead, Health Equity Quality Assurance Lead, and Health Equity Data Entry Interns.

Health Equity Team Leads

Team leads assist the program manager in building and organizing the Health Equity track. Additionally, they encourage a sense of community between the data entry interns, support them, and assess the quality of the data they enter into the [Health Equity Data Entry spreadsheet](#).

Health Equity Acting Leads

Acting leads assist their respective team leads in building and organizing the Health Equity Track. They perform the roles of the team lead as an understudy. When a team lead may not be able to act on their roles for a certain amount of time, the acting lead then acts as the interim team lead. Besides maintaining this flexible position, acting leads are responsible for communicating with data entry interns and fostering a sense of community in the team.

Health Equity Quality Assurance Leads

Quality assurance leads check for the accuracy of the data that is entered by data entry interns in the [Health Equity Data Entry spreadsheet](#). They also check for weird trends in data, which include the daily number of cases substantially decreasing in a day and the daily number of cases not being updated after a day. In this manner, they ensure the quality and accuracy of the data collected.

Health Equity Data Entry Interns

Data entry interns collect county-level COVID-19 data by race and ethnicity. Specifically, they obtain confirmed COVID-19 case counts by race (White, Black/African American, Asian, American Indian/Alaska Native, and Native Hawaiian/Pacific Islander) and ethnicity (Hispanic and non-Hispanic) at the county level.

Each intern is assigned approximately 15 counties to collect data for. After obtaining the data, interns enter the case counts into the [Health Equity Data Entry spreadsheet](#).

Health Equity Data Entry Spreadsheet

Data entry interns store their COVID-19 data in the [Health Equity Data Entry spreadsheet](#).

The “County Assignments” tab contains the following columns:

- *[unlabeled]* - Team leads contributing to the Health Equity track
- *Team Member* - Data entry interns contributing to the Health Equity track
- *Assignment* - County assignments for each intern
- *County/State Updates* - Notes for performing data entry for specific counties and states

	A	B	C	D
1		Team Member	Assignment	County/State Updates:
11	Group 2/ Aisha	Brittney Boakye		Ohio Notes:
12		Christine Zhu		- https://coronavirus.ohio.gov/wps/por
13		Tara Pinto		-Update time: 2 PM EST
14		Lexie Adams		Indiana Notes:
15		Mahlet Shenkute	https://docs.google.com/spreadsheets/d/1JXK88OyTHt8uHbrn1sT5Xvw308JFkmic2UijQcuC4CA/edit#gid=1340191497	- https://www.coronavirus.in.gov/
16		Melat Ali		-Update Time: 12 PM EST
17		Khushali Vashi		
18		Jonas Zhang		
19		Jeanne Moua		
20		Kristina Var		
21		Ting Ting Juan		
22				

Each month out of the year has its own tab, such as the “June Data” tab or “July Data” tab, and each of these tabs contain the following columns and subcolumns:

- *County/State Website w/ COVID Info* - county and state websites where interns collect data
- *Each day of the month has its own column (formatted dddd, m/dd/yy) and contains the following subcolumns:*
 - *Race*
 - *White*
 - *Black/African American*
 - *Asian*
 - *American Indian/Alaska Native*
 - *Native Hawaiian/Pacific Islander*
 - *2+ races*
 - *Other*
 - *Unknown*
 - *Ethnicity*
 - *Hispanic (all races)*
 - *Non-Hispanic*
 - *Not Specified*

As of September 2020, these tabs contain the following new columns:

- *Total (stated on website)*
 - *Total Daily Case Count*
- *Calculated Totals*
 - *Race Total*
 - *Ethnicity Total*

	A	B	C	D	E
1		County/State Website w/ COVID Info	Wednesday, 7/1/20		
2			Race		
3			White	Black/African / Asian	
263	Adams County, Ohio	https://coronavirus.ohio.gov/wps/portal/	20	0	0
264	Allen County, Ohio	https://coronavirus.ohio.gov/wps/portal/	213	60	0
265	Ashland County, Ohio	https://coronavirus.ohio.gov/wps/portal/	54	0	0
266	Ashtabula County, Ohio	https://coronavirus.ohio.gov/wps/portal/	297	58	0
267	Athens County, Ohio	https://coronavirus.ohio.gov/wps/portal/	35	0	2
268	Auglaize County, Ohio	https://coronavirus.ohio.gov/wps/portal/	82	0	0
269	Belmont County, Ohio	https://coronavirus.ohio.gov/wps/portal/	365	74	3
270	Brown County, Ohio	https://coronavirus.ohio.gov/wps/portal/	45	0	0
271	Butler County, Ohio	https://coronavirus.ohio.gov/wps/portal/	631	210	103
272	Carroll County, Ohio	https://coronavirus.ohio.gov/wps/portal/	31	0	0
273	Champaign County, Ohio	https://coronavirus.ohio.gov/wps/portal/	19	0	0
274	Clark County, Ohio	https://coronavirus.ohio.gov/wps/portal/	267	98	1

F	G	H	I	J	K	L	M
					Ethnicity		
American Indian/Alas	Native Hawaiian/Pac	2+ races	Other	Unknown	Hispanic (all r	Non-Hispanic	Not Specified
0	0	1	0	0	0	20	1
0	0	11	6	7	9	268	20
0	0	0	0	1	1	48	6
0	1	8	21	36	22	318	81
0	0	0	0	0	2	32	3
0	1	0	8	8	5	79	15
1	0	35	39	17	9	416	109
0	0	0	1	7	1	38	14
3	7	51	203	240	314	764	370
0	0	1	0	19	0	29	22
0	0	0	1	18	0	17	21
0	0	14	286	73	313	364	62

Data Sources

Data entry interns collect COVID-19 case counts by race and ethnicity from [state health department websites](#) if the counts are broken down by county. If the counts are not broken down by county, interns collect them from county websites or county public health department websites.

The websites that interns currently use to collect data can be found under the “County/State Website w/ COVID Info” column of the [Health Equity Data Entry spreadsheet](#).

Note: “County/State Website w/ COVID Info” columns are found under every tab of the [Health Equity Data Entry spreadsheet](#) except for the “County Assignments” tab.

Data Collection & Input

To collect data for a specific county, click on the link provided in the “County/State Website w/COVID info” column next to that county and locate its COVID-19 data reported by race and ethnicity. These websites will look different for different states. While some states use one website to report their county-level data, other states rely on individual counties to report their data on their own websites. See the State-Specific Information section below for notes specific to each state.

	A	B
1		County/State Website w/ COVID Info
2		
3		
263	Adams County, Ohio	https://coronavirus.ohio.gov/wps/portal/
264	Allen County, Ohio	https://coronavirus.ohio.gov/wps/portal/
265	Ashland County, Ohio	https://coronavirus.ohio.gov/wps/portal/
266	Ashtabula County, Ohio	https://coronavirus.ohio.gov/wps/portal/
267	Athens County, Ohio	https://coronavirus.ohio.gov/wps/portal/
268	Auglaize County, Ohio	https://coronavirus.ohio.gov/wps/portal/
269	Belmont County, Ohio	https://coronavirus.ohio.gov/wps/portal/
270	Brown County, Ohio	https://coronavirus.ohio.gov/wps/portal/
271	Butler County, Ohio	https://coronavirus.ohio.gov/wps/portal/
272	Carroll County, Ohio	https://coronavirus.ohio.gov/wps/portal/
273	Champaign County, Ohio	https://coronavirus.ohio.gov/wps/portal/
274	Clark County, Ohio	https://coronavirus.ohio.gov/wps/portal/

It is important to navigate and familiarize yourself with the website the first time you enter data. When collecting data, **always check its timestamp and confirm that it has been updated that day**. Most states update by 5pm, but some states will update twice in one day, such as at 10am and 5pm. Information about when each state updates their data is also included below.

Once you have found the race and ethnicity data and checked its timestamp to ensure it is current, copy the data you find over into the spreadsheet. Include the numbers for all races and ethnicities that are reported. Some counties will report when a race or ethnicity has no cases as “0” so input that as “0”. If a county does not report a number of cases for a specific race or ethnicity, use a dash (-) in that cell on the spreadsheet.

		H	I	J	K
1	No data was reported for the "2+ races" category				
2					Ethnicity
3		2+ races	Other	Unknown	Hispanic (all i
63	Brevard County, Florida	-	127	531	349
64	Broward County, Florida	-	1337	2205	5724
65	Calhoun County, Florida	-	2	11	1
66	Charlotte County, Florida	-	16	87	55
67	Citrus County, Florida	-	17	42	28
68	Clay County, Florida	-	67	150	69
69	Collier County, Florida	-	355	1535	2034
70	Columbia County, Florida	-	38	8	41
71	DeSoto County, Florida	-	24	76	399
72	Dixie County, Florida	-	0 cases were reported for those of Hispanic descent	1	0
73	Duval County, Florida	-		2170	652
74	Escambia County, Florida	-		320	67
75	Flagler County, Florida	-	11	11	46
76	Franklin County, Florida	-	0	0	0

The following are tips for entering data:

- Use cumulative counts, NOT daily counts, if both options are available.
- If residential and non-residential counts are listed, only use residential data.
- If a county lists data as “multiracial” or “biracial”, then input those numbers into the “2+ races” column. Include a note (NOT a comment) to indicate that the data is reported as “mulitracial” or “biracial.”
- If a county lists data as “refused to answer,” then report this data in the “other” column and add a note.
- If a county reports data as a percentage, multiply it by the total number of cases to convert it to the raw data number.

Note: It is VERY important to make sure you collect data each day to capture that data before it is written over at the next update. Each team lead will set a deadline each day to have data entered by. If you have concerns about meeting that deadline, please contact your team lead.

State-Specific Information

Please refer to the information below for state-specific information related to each state's website and reporting windows. This information is current as of July 15th, 2020.

In order to see an updated list of state-specific strategies, please go to [State-Specific Strategies](#).

Florida	Updates at 11am EST
	Florida's data by race ONLY works on the First Florida County (Alchua), so use that link for all counties
	Search county with the state link on their search bar in the upper right hand corner
	State Link: https://experience.arcgis.com/experience/96dd742462124fa0b38ddedb9b25e429
Georgia	All counties have individual links, so make sure to check timestamps for daily updates
Illinois	Updates at 4pm EST daily
Indiana	All counties update at 12pm EST
Louisiana	All parishes update at 2pm EST once weekly on Wednesdays
North Carolina	All counties update at 12pm EST; (accept occasionally on the weekends they do not update until 3~4pm EST)
	Data is reported in percentages but if you hover over the bars on the graph it presents as the count (the percentage and count do not match up), so the count data is what we record for NC by hovering over each bar in the graph

Ohio	All counties update at 2pm EST
Pennsylvania	Lehigh County updates at around 1pm EST
West Virginia	All counties are on one website: https://dhhr.wv.gov/COVID-19/Pages/default.aspx
	Use the “County Summary” tab to find data and the dropdown menu to scroll down and navigate between counties
	Check each county individually
	All county data updates at 10am EST

Data Entry Challenges

Varying Reporting Methodologies

Percentages vs. Raw Numbers

Some counties report COVID-19 cases as percentages (%) of total daily cases while others report their cases as raw counts (#). This can lead to misrepresentation of the distribution of cases by race/ethnicity.

Race and Ethnicity Categorization

Race and ethnicity are independent of each other, yet some counties combine racial and ethnic categories under one chart or table; this creates a challenge in interpreting their data. For example, when “Hispanic” is treated as a racial category, it is unclear whether Black Hispanics with positive cases of COVID-19 are counted as Black, Hispanic, or both. Some counties will only classify their data into two to four distinct racial categories, such as “White,” “Black/African American,” “Other,” and “Unknown.” This results in a lack of data representative of other racial categories, such as “Asian” and “American Indian or Alaska Native.”

Untimely Data Releases

COVID-19 case counts are collected daily. When counties only update their data on specific days (e.g., weekdays), gaps are unavoidably created. This complicates our ability to observe daily trends for those specific counties. Some data are not reflective of a 24-hour time period because some counties release new data either multiple times within a 24-hour time-period or less frequently than every 24 hours.

Issues with Disease Reporting Systems

Due to the large volume of COVID-19 tests being processed and technical issues arising with disease reporting systems, we have experienced additional gaps in the data. In August, California experienced complications with its electronic laboratory reporting system. This led to a backlog of test results and in state and county health departments underreporting daily new COVID-19 cases.

Frequently Asked Questions

Where do I find my county assignments?

Your county assignments should be located in the first tab of the spreadsheet called “County Assignments.” If you cannot locate your county assignments there, please contact your team lead.

My county lists daily numbers and cumulative numbers. Which set of numbers do I use?

For the purposes of this project, we are using cumulative cases, NOT daily cases.

Where do I count data that my county lists as “multiracial” or “biracial”?

Input this data under the “2+ races” category. This is separate from data that is listed as “other” or “unknown,” which have their own respective columns. When inputting this data, right click on the cell and press “insert note” to add a note to the cell. DO NOT press “comment” to add notes, as these do not transfer over to new spreadsheets.

Where do I count data that my county lists as “refused to answer?”

Input this data under the “other” category and add a note to the cell indicating that the number includes data listed as “refused to answer.”

What do I do if the data on my county's website is in percentages?

When data is listed as a percentage, multiply it by the total number of cases for that county and round to the nearest whole number. For example, if the website states that 23% of the people who tested positive for COVID-19 are Asian and there are 150 total cases, then you would multiply 0.23 by 150 and get 34.5. Then, you would round up to 35 and enter this into the "Asian" column.

What do I do if my county does not list a specific race or ethnicity?

If a county reports **absolutely no data** for a specific race or ethnicity, then input a dash (-) into the cell for that category. It is **very important to NOT put 0 here** since the counties are technically not reporting numbers for those categories. Only list 0 if the county is reporting 0 cases for a specific race or ethnicity.

What do I do if cases are separated into residential and non-residential tallies?

When cases are separated into residential and non-residential tallies, only include the residential numbers in the spreadsheet. It is important to avoid counting non-residential numbers as part of that county because it can cause a case to be double-counted. Many people often travel outside of their county of residence to seek medical care, especially in more rural areas. This could mean that many cases were diagnosed in a specific county at their hospital, but those cases are "housed," followed-up, and reported as their home county.

What do I do if I want to make a note on a data cell in the spreadsheet?

To make a note on the spreadsheet, enter the data into the cell, right-click on that cell, and click on "Insert note." DO NOT press "Comment" because comments do not transfer over when the spreadsheet is copied.

What do I do if I have a question about the data I am seeing?

Questions can be directed towards team leads, and if they are not sure of the answer, they will take your question to the health equity track lead. Also, it is important to utilize your peers in your groups to answer questions and build community.

References

1. Barnshaw, J. Race. In Schaefer, Richard T., ed. *Encyclopedia of Race, Ethnicity, and Society*. 1. Thousand Oaks, CA: SAGE Publications; 2008:1091.
2. Dictionary.com. Ethnicity. Dictionary.com website. Accessed August 20, 2020. <https://www.dictionary.com/browse/ethnicity>
3. American Psychological Association. Socioeconomic status. American Psychological Association website. Accessed September 10, 2020. <https://www.apa.org/topics/socioeconomic-status>
4. World Health Organization. Health equity. World Health Organization website. Accessed August 22, 2020. https://www.who.int/topics/health_equity/en/
5. The Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020: Recommendations for the Framework and Format of Healthy People 2020. Healthy People 2030 website. October 28, 2008. Accessed September 17, 2020. https://www.healthypeople.gov/sites/default/files/PhaseI_0.pdf
6. Weiland N, Mandavilli A. Trump Administration Sets Demographic Requirements for Coronavirus Reports. *The New York Times*. June 4, 2020. Accessed September 15, 2020. <https://www.nytimes.com/2020/06/04/us/politics/coronavirus-infection-demographics.html>
7. Centers for Disease Control and Prevention. Health Equity Considerations and Racial and Ethnic Minority Groups. Centers for Disease Control and Prevention website. Accessed September 15, 2020. <https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html>