
CTDPH Vintage 2022 County Estimates: Methodology

June 2025



Prepared by:

Demographic Analytics Advisors and the Connecticut Department of Public Health

Christopher Dick
Founder and CEO
Demographic Analytics Advisors

Karyn Backus
Epidemiologist 4
Connecticut Department of Public Health

The authors gratefully acknowledge the contributions of the following staff of the CT Data Collaborative and the Connecticut Department of Public Health:

Michelle Riordan-Nold
Executive Director
CT Data Collaborative

Laura E. Hayes, PhD
Data Scientist
Connecticut Department of Public Health

Cynthia Willner
Senior Research Associate
CT Data Collaborative

This publication was supported by funds from the Department of Health and Human Services (HHS) Centers for Disease Control and Prevention (CDC) Notice of Funding Opportunity (NOFO) number CK19-1904, entitled Epidemiology and Laboratory Capacity for Prevention and Control of Emerging Infectious Diseases (ELC) (Grant No. # 6NU50CK000524) with funds made available under the Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020 (P.L. 116-123); the Coronavirus Aid, Relief, and Economic Security Act, 2020 (the “CARES Act”) (P.L. 116-136); the Paycheck Protection Program and Health Care Enhancement Act (P.L. 116-139); and/or the Consolidated Appropriations Act, 2021, Division M – Coronavirus Response and Relief Supplemental Appropriations Act, 2021 (P.L. 116-260). The contents of the publication are solely the responsibility of the authors and do not necessarily represent the views of the HHS.

Suggested Citation:

Dick, C. and Backus, K. (2025). *CTDPH Vintage 2022 County Estimates: Methodology*. Connecticut Department of Public Health, Hartford, CT. www.ct.gov/dph/populationdata

Background

This document explains the methodological process and outcomes of the development of population estimates for Connecticut's historical counties by age, sex, race, and Hispanic origin (ASRH).

In June 2022, the United States Census Bureau and the State of Connecticut announced that the Census Bureau was adopting Connecticut's planning regions as county-equivalent statistical units for future Census data products.¹ The State of Connecticut had requested this change to better support state governance.^{2,3} In 2023, the Census Bureau's Population Estimates Program released the Vintage 2022 annual estimates using planning regions in lieu of Connecticut's historical counties.⁴ Annual population estimates for Connecticut's historical counties were no longer being released by the Census Bureau. *Total* population estimates for Connecticut's historical counties can still be calculated for 2022 and beyond by aggregating town-level total population estimates to the county level, since Connecticut's towns are nested within both its historical counties and planning regions. However, the Census Bureau does not publish town-level population estimates by ASRH. Thus, it is not possible to directly calculate population estimates by ASRH for Connecticut's historical counties from data released by the Census Bureau.

The Census Bureau was at the forefront of the transition to planning regions in 2022, yet broad adoption of the planning regions as county-equivalent geographic units throughout the federal, state, and local data ecosystems lagged. As of the publication of this methodology, many agencies and organizations have yet to adjust their operations to collect data by planning region, which is a necessary precursor for analysis and reporting by planning region. Due to this delay in the adoption of the new geographies across agencies and organizations, many data users, such as those in public health surveillance, continue to require annual population estimates in the historical county geographic boundaries to support ongoing work.

To fill the gap between availability and need, the Connecticut Department of Public Health (CTDPH) worked with Demographic Analytics Advisors (DAA) to develop and release estimates of Connecticut's historical county populations by ASRH for Vintage 2022. The following sections explain how the authors were able to develop reasonable and effective county-based estimates for the State of Connecticut using publicly available data sources, such as the 2010 and 2020 Censuses, annual population estimates from the Census Bureau, estimates from CTDPH, and the modified race file from the 2010 Census. These estimates were purposefully developed in a way that supports future releases of annual county-based demographic estimates until county data are no longer needed.

¹ <https://www.federalregister.gov/documents/2022/06/06/2022-12063/change-to-county-equivalents-in-the-state-of-connecticut>

² <https://portal.ct.gov/-/media/opm/igpp/org/county-equivalency/county-equivalency-request-letter-to-us-census-bureau-81419-signed.pdf?rev=a7d3a649e89a4260af9f9ef68f83dbc6&hash=38BA3051FD8E6ED314A427D4A50FFD93>

³ <https://www.ctdata.org/blog/census-connecticut-county-equivalents-councils-of-government-cogs>

⁴ <https://www.ctdata.org/blog/census-bureau-releases-first-population-estimates-for-connecticuts-county-equivalent-planning-regions>

Methodology

To develop the Vintage 2022 county-level estimates, we sourced three population estimate datasets, listed below:

1. CTDPH Interim July 1, 2020 town ASRH estimates ⁵
2. US Census Bureau Vintage 2022 town total estimates ⁶
3. US Census Bureau Vintage 2022 state characteristics estimates ⁷

Step 0: Development of CTDPH Interim July 1, 2020 town ASRH Estimates

We began the process of developing these estimates by using the CTDPH Interim July 1, 2020 town ASRH Estimates. These estimates use the 2020 Census, 2010 Census Modified Race File, and the Vintage 2021 population estimates from the Census Bureau to create town-level ASRH estimates for July 1, 2020. For further information on these estimates see the methodology statement.⁸

Step 1: Proportional Raking

We began the process of developing these estimates by raking the July 1, 2020 town ASRH estimates to two marginal totals concurrently using iterative proportional fitting. These marginal totals were the Vintage 2022 town totals and the Vintage 2022 state characteristics. Our process used the R package 'mipfp'.⁹ This step started with the town ASRH distribution from the CTDPH interim July 1, 2020 town ASRH estimates and adjusted the town-level demographic characteristics until the distribution matched both the state population distribution by ASRH and the town total populations for July 1, 2022 from the Census Bureau's Vintage 2022 estimates. We used state-level ASRH instead of planning region ASRH because we wanted to keep the estimates closer to the town and county ASRH distributions from the 2020 Census, given that the Vintage 2022 estimates still did not use the 2020 Census race data.

Step 2: Controlled Rounding

After raking the July 1, 2020 data to July 1, 2022, we were left with town ASRH estimates that add up to each of the marginal totals but included fractional numbers. Our next step was to round the raked data matrix while still maintaining the marginal totals from the Vintage 2022 town totals and state characteristics. To do this we wrote an algorithm in R to iteratively match the column and row marginals until each cell within the raked matrix was rounded to a whole number and matched the marginals.¹⁰ In this algorithm, we constrained the rounding process in order to minimize the potential for bias in the final town ASRH estimates.

⁵ <https://portal.ct.gov/dph/health-information-systems--reporting/population/town-population-with-demographics>

⁶ https://www2.census.gov/programs-surveys/popest/datasets/2020-2022/cities/totals/sub-est2022_9.csv

⁷ <https://www2.census.gov/programs-surveys/popest/datasets/2020-2022/state/asrh/sc-est2022-alldata6.csv>

⁸ <https://portal.ct.gov/dph/health-information-systems--reporting/population/town-population-with-demographics>

⁹ <https://cran.r-project.org/package=mipfp>

¹⁰ The code for this step of the process is available upon request.

Step 3: Aggregate to County

Once we had raked and rounded the data, the final step was to aggregate each of the towns into the eight county geographies using the following demographic characteristics:

1. Age (18 age groups: 17 five-year age groups plus an '85 years and above' age group)
2. Sex
3. Race and Hispanic Origin (Hispanic all races, Non-Hispanic White, Non-Hispanic Black, Non-Hispanic American Indian or Alaska Native, Non-Hispanic Asian, Non-Hispanic Native Hawaiian or Pacific Islander, and Non-Hispanic two or more races)

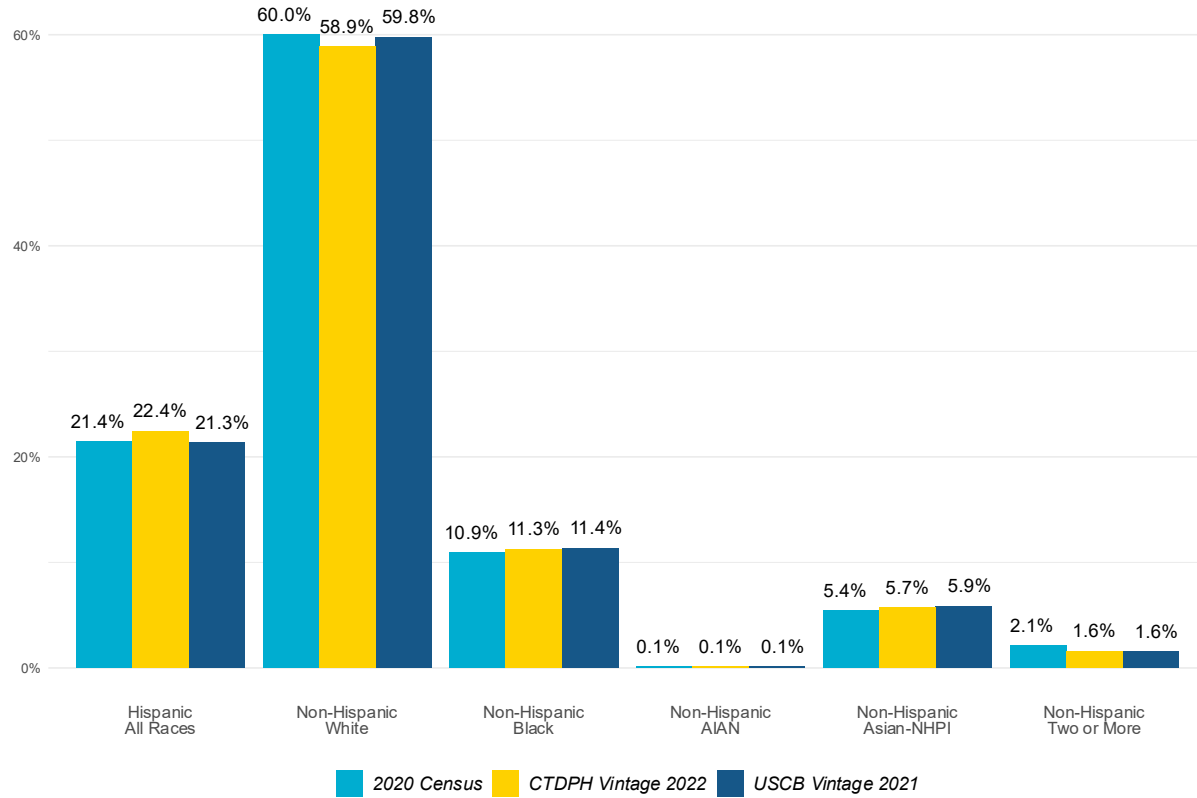
Analysis

Below we have provided details on the outcome of our methodology as compared to both the 2020 Census (with preliminary modified race) and the Vintage 2021 population estimates from the Census Bureau. These two data sources served as valuable comparisons because they use different sources for Hispanic origin and race data – the Vintage 2021 population estimates are still based on race and Hispanic origin from the 2010 Census. Therefore, relative agreement between each of the data sets would confirm a stability that we expect to continue into our estimates.

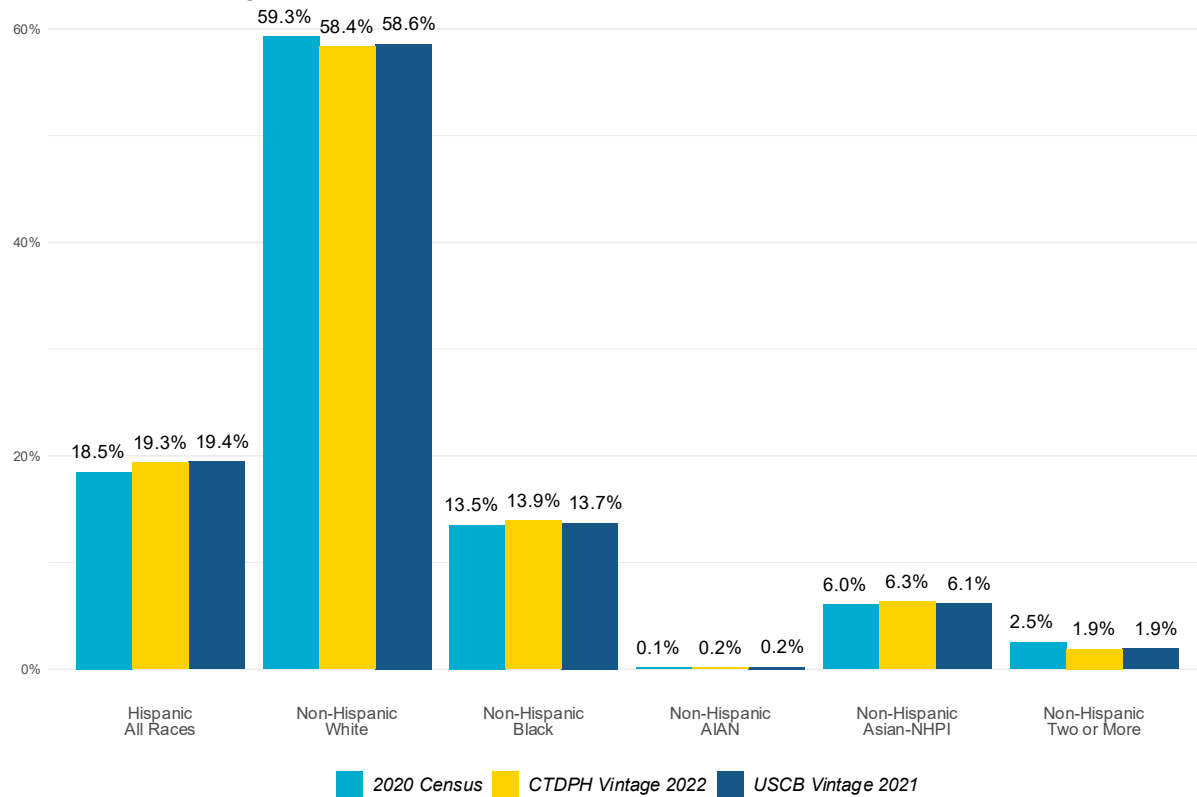
Below we show the county-level charts, with three different datasets compared:

1. The light blue bars labeled as 2020 Census represent the April 1, 2020 Census data after reallocating the Some Other Race category using the 2010 modified race method.
2. The yellow bars labeled as CTDPH Vintage 2022 represent the CTDPH Vintage 2022 county estimates that we have described in this methodology statement.
3. The dark blue bars labeled as USCB Vintage 2021 represent the Vintage 2021 county estimates released by the Population Estimates Program at the United States Census Bureau that do not include race / Hispanic data from the 2020 Census.

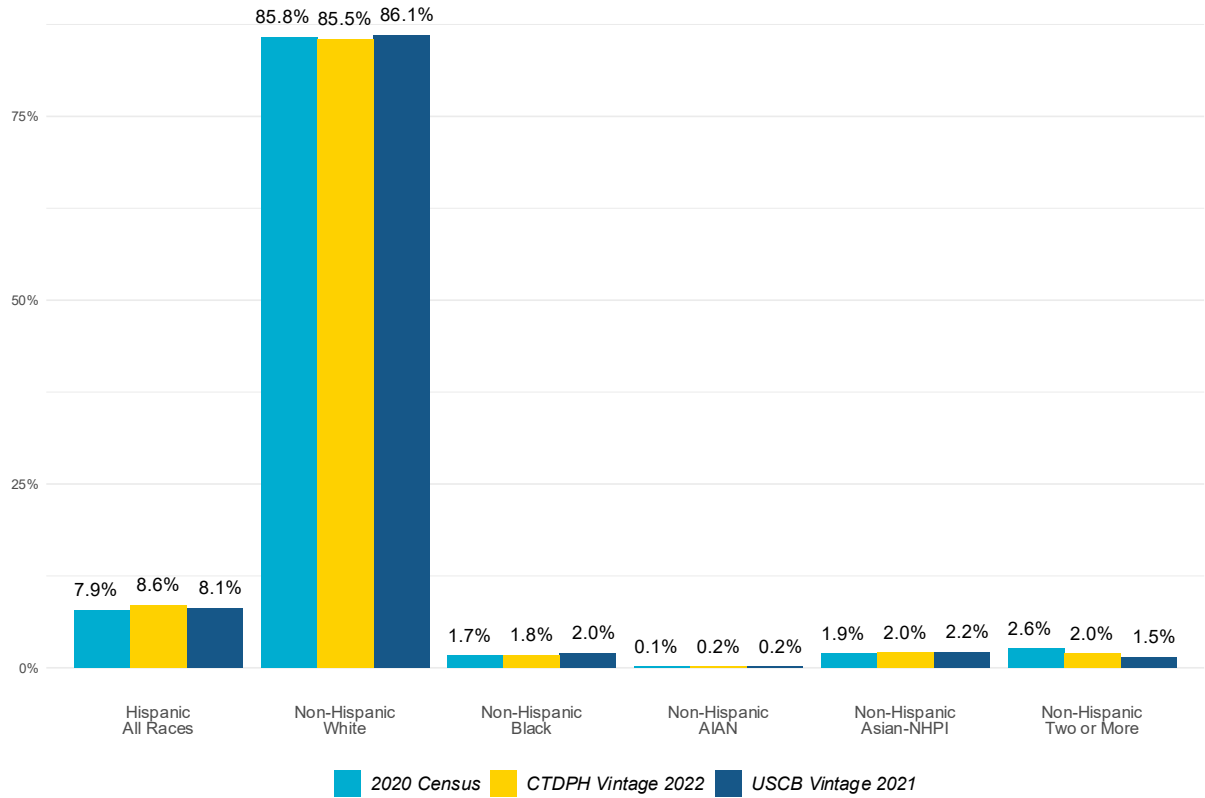
Fairfield County



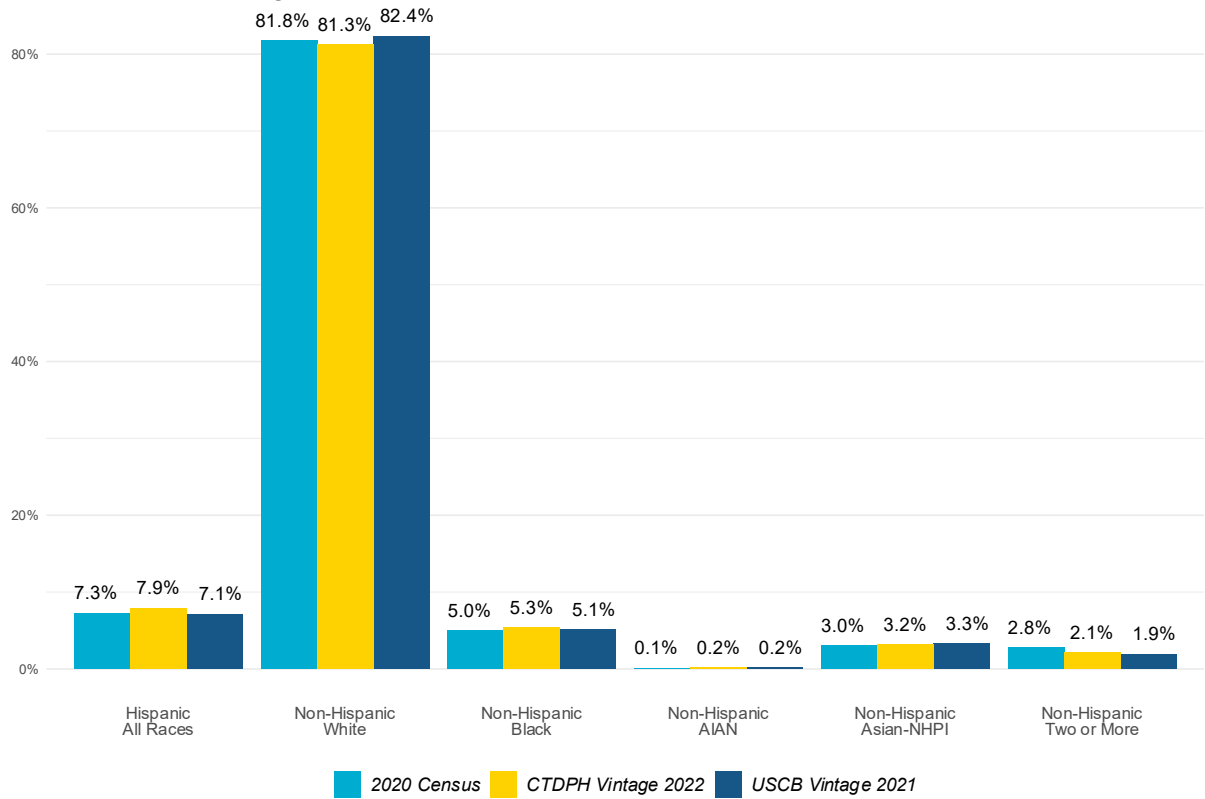
Hartford County



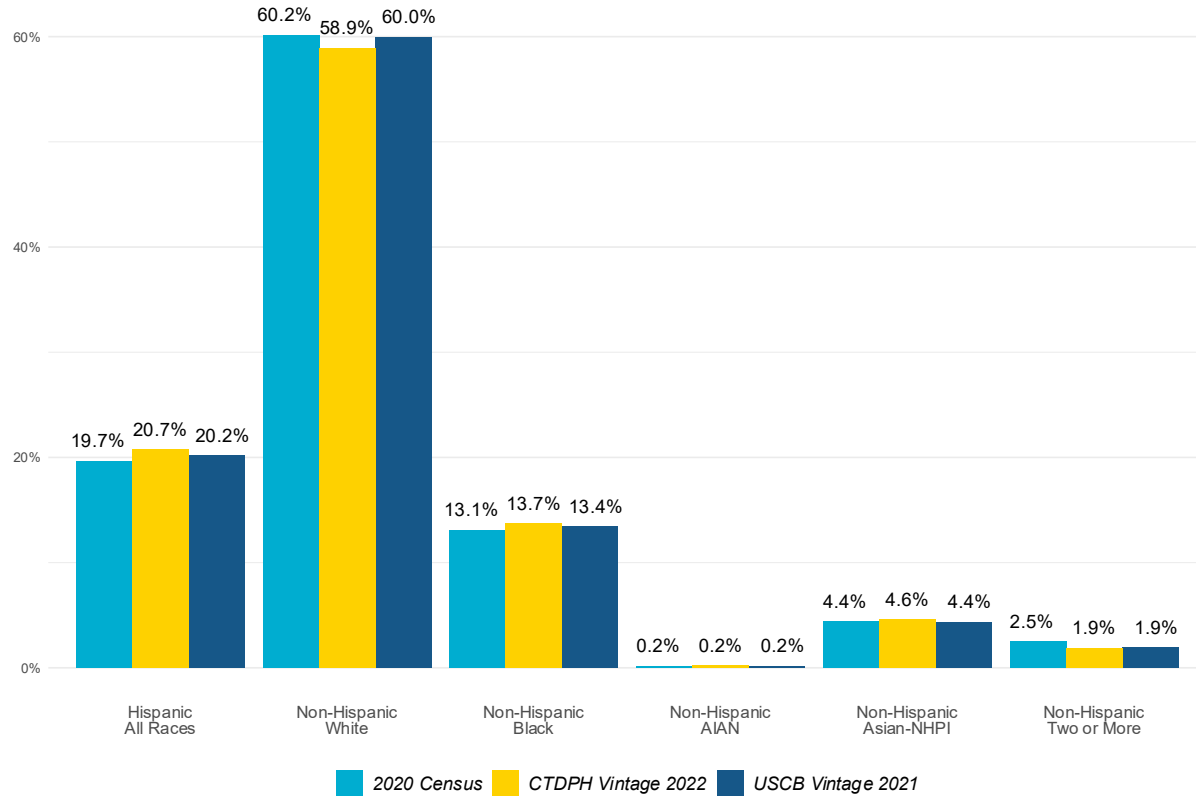
Litchfield County



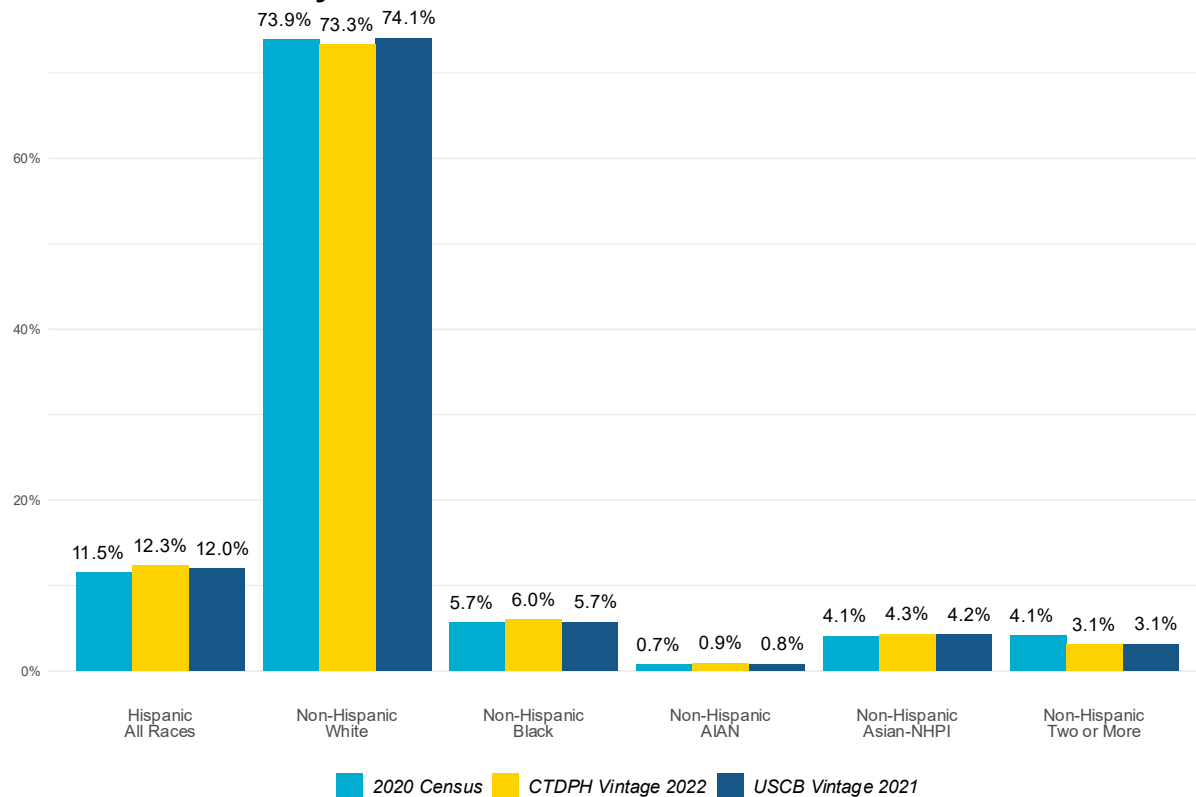
Middlesex County



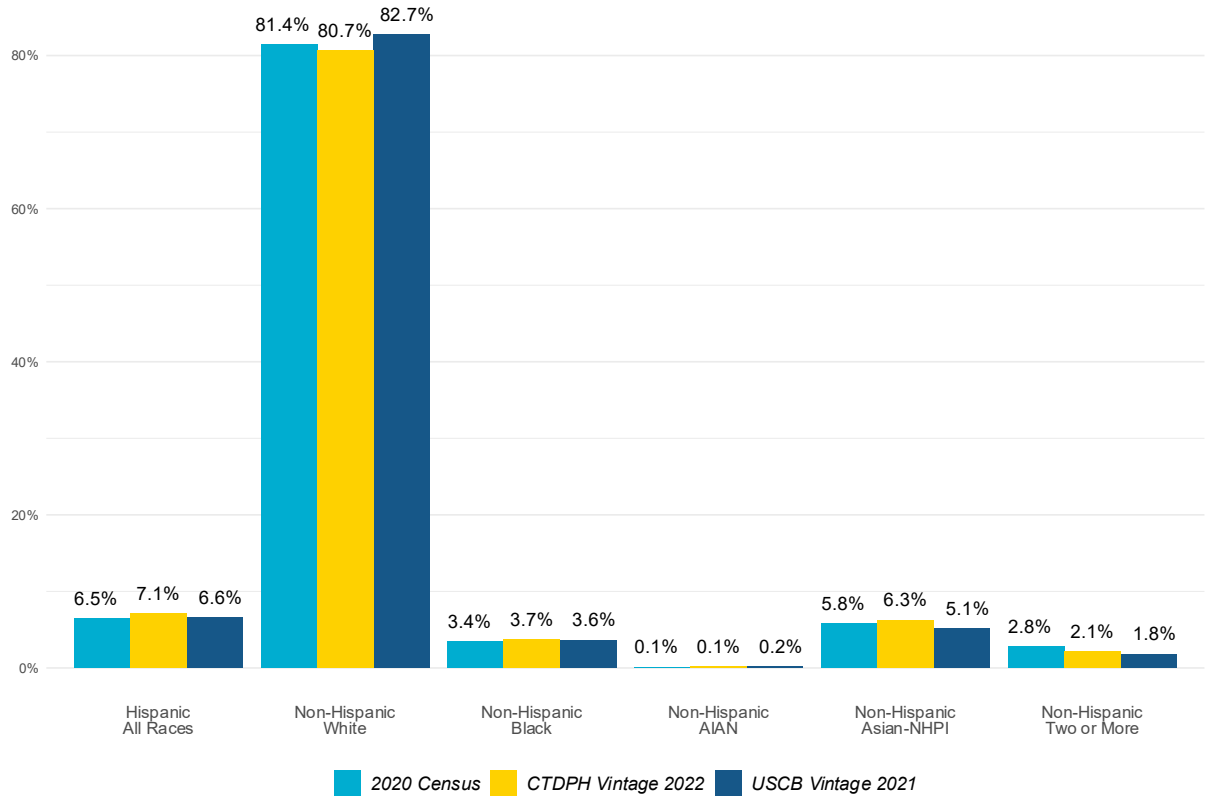
New Haven County



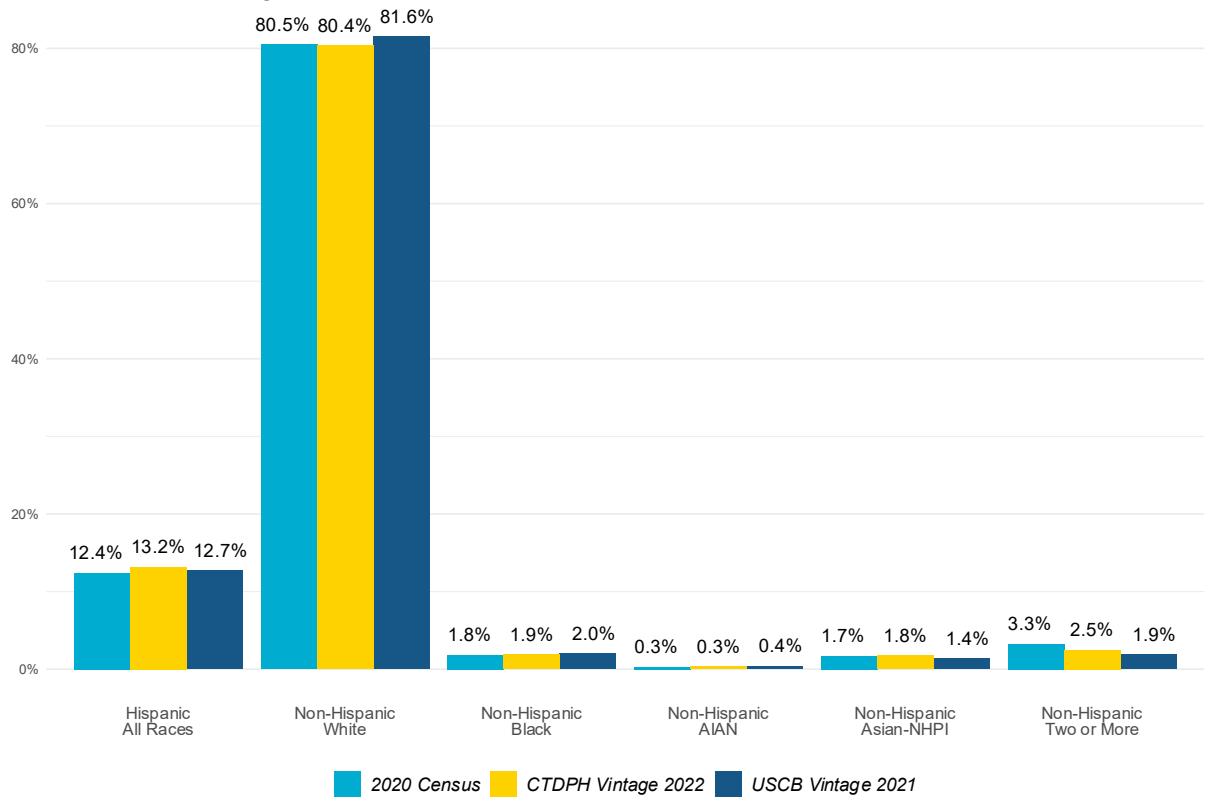
New London County



Tolland County

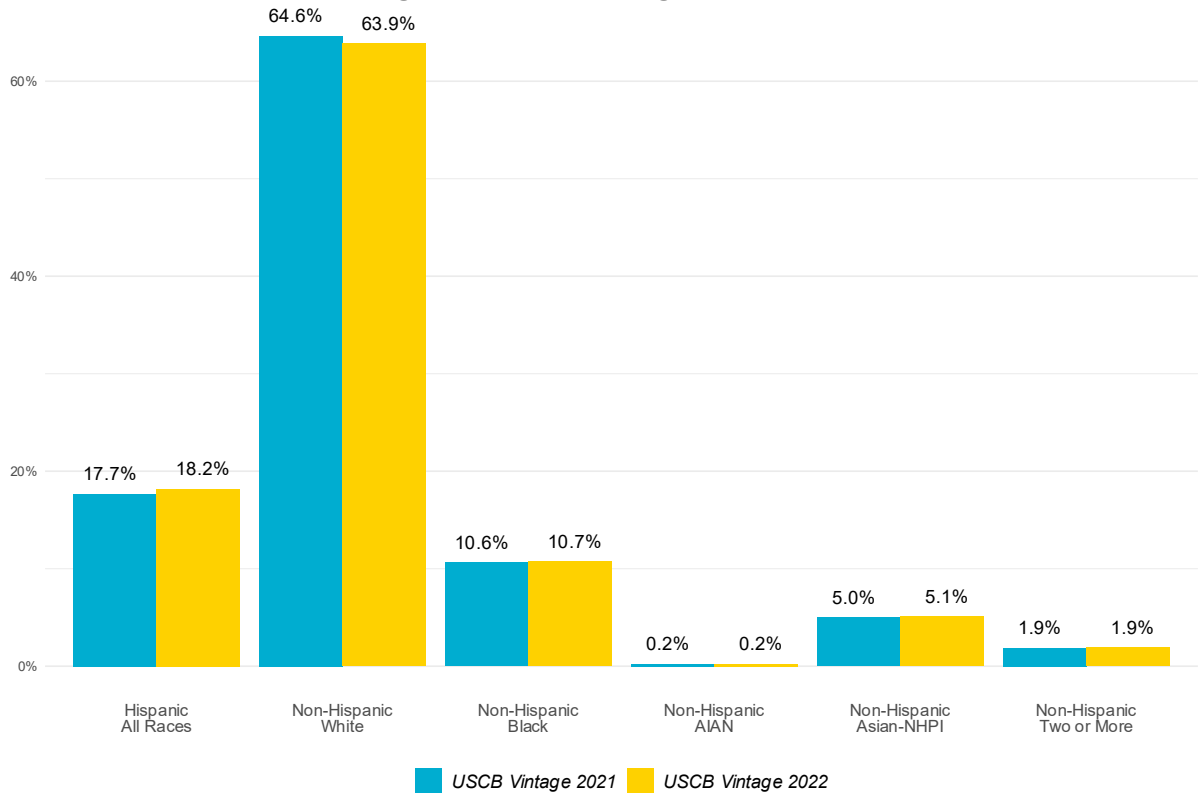


Windham County



Through these series of charts, we show that each county's race / ethnicity distributions were reasonably consistent across the three distinct datasets. We noted two consistent patterns of group differences for all counties. First, the CTDPH Vintage 2022 estimates were slightly lower for non-Hispanic White. This is because, as we see in the chart below, the Vintage 2022 estimates for ASRH at the state level are lower by about 1 percentage point in Vintage 2022 for the non-Hispanic White population than they were in the Vintage 2021 estimates. Second, the 2020 Census estimates were notably higher for non-Hispanic TOM races, which is expected given that the 2020 Census saw a large increase in this population. Although some variation was apparent, we considered the magnitude of difference based on the size of the individual race / ethnicity groups to be within acceptable levels and consider these estimates to be fit-for-use as denominators for calculation of population-based rates in public health.

State Comparison: Vintage 2021 vs Vintage 2022



Conclusion

County-level population estimates with demographic characteristics are needed to support state and national public health use cases as well as broader state and federal governance. While the Census Bureau adopted the planning regions for the Vintage 2022 estimates, many other agencies and organizations have not yet updated their operations to collect and share data by planning region. This created a geographic inconsistency between how public health data for 2022 and later were collected (historical county) and the population estimates by demographic characteristics available for 2022 and later (planning regions).

To meet state and national public health needs, the CTDPH has produced and released Vintage 2022 population estimates based on the historical counties that were used prior to the Census Bureau's adoption of Connecticut's planning regions as county-equivalents. The estimates released using this methodology fill that need in a way that is demographically plausible and that follows historical trends at the county level. Annual estimates may continue to be released each year until the need for historical county estimates no longer exists.