

WELCOME TO:



A More Accurate Denominator:
the Benefit of Better Census
Annual Population Estimates

AAPOR offers webinar sponsorship opportunities to organizations that would like to support AAPOR's online education program.

In addition, AAPOR has an institutional subscription for greater access to the education webinar recording library.

For more information, contact;
Lailah Johnson, AAPOR Program Manager
ljohnson@aapor.org

Join us for the Next Webinar in the 2024 Series:

MMS push-to-web to reach historically under-represented subgroups: Recent results and how to start using this modality

Presenters: Joel LaLone & Matt Hayes

August 22, 2024

1:00 PM – 2:00 PM ET



Please enter your questions in the Q&A box at the bottom of your screen.



Please complete the webinar survey immediately following the session.



TODAY'S PRESENTER

Luke Rogers



A More Accurate Denominator: The Benefit of Better Census Annual Population Estimates

Luke Rogers, PhD

Senior Research Scientist for Estimates Development and Improvement

Christine Hartley, PhD

Assistant Division Chief for Estimates and Projections
Population Division

American Association for Public Opinion Research Webinar

July 25th, 2024

The Population Estimates Program

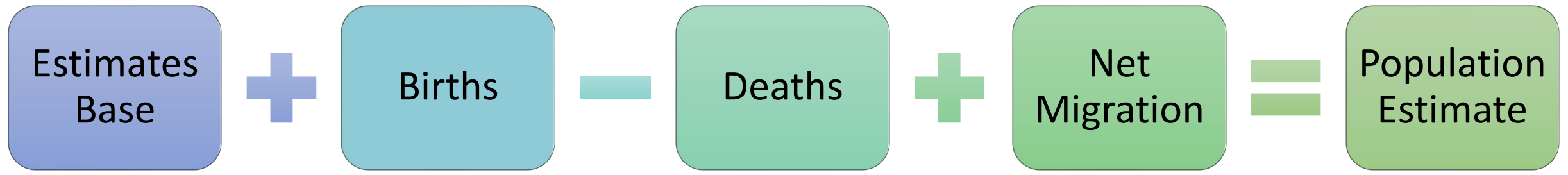
- The Population Estimates Program (PEP) disseminates official measures of population and housing units between decennial censuses
- Developed as per Title 13, U.S. Code, Section 181
- Use cases include:
 - Population controls and denominators
 - Academic and business research
 - Program planning in the public and private sectors
- Time series are released annually featuring data for the date of the last census through the vintage year, which represents the latest year of estimates available
- Vintage 2023 is the current series
 - Contains estimates for April 1, 2020 to July 1, 2023



Availability of Population and Housing Unit Data Products: Vintage 2023

Release Date*	Product	Geography
December 2023	Population totals, components of change, and voting-age population	Nation, states, and Puerto Rico Commonwealth
March 2024	Population totals and components of change	Metropolitan and micropolitan statistical areas, counties, and Puerto Rico municipios
April 2024	Population by age and sex	Nation
May 2024	Population totals--	Cities and towns (incorporated places, minor civil divisions)
	Housing unit totals--	Nation, states, and counties
June 2024	Population by age, sex, race, and Hispanic origin--	Nation, states, and counties
	Population by age and sex--	Metropolitan and micropolitan statistical areas, Puerto Rico Commonwealth and municipios

The Cohort-Component Method for the Nation, State, and Counties



- The starting point for each vintage of population estimates
- Over recent decades, the decennial census typically provided all the necessary detail for the estimates base

- Data on vital records (birth and death certificates) from the National Center for Health Statistics and the Federal-State Cooperative for Population Estimates

- Net domestic migration within the United States, based on administrative records
- Net international migration between the United States and elsewhere, based primarily on the American Community Survey

- Estimates are produced as of July 1 for each year of the time series (also known as a “vintage”)
- Each year, we begin at the estimates base and produce a new vintage of estimates

Developing the Post-2020 Estimates Base

- The most recent decennial census typically forms the estimates base
 - Still applies for the subcounty and housing unit estimates methods
- Vintage 2021 was more complicated for counties and higher levels of geography
 - Data availability was delayed due to the impact of COVID-19 and updated disclosure avoidance requirements
- A method was developed to build an estimates base using alternative data sources

2020 Census: COVID-19 Presents Delays and Risks to Census Count

GAO-20-551R

Published: Jun 09, 2020. Publicly Released: Jun 09, 2020.

The New York Times

Knocked Off Track by Coronavirus, Census Announces Delay in 2020 Count

Knocked Off Track by Coronavirus, Census Announces Delay in 2020 Count ... 15, to Oct. 31, and would begin reopening its field offices – which have been...

Jul 28, 2020

NBC Chicago

US Census Bureau Suspends Field Operations on Virus Concerns

A week after starting its 2020 count, the U.S. Census Bureau on Wednesday suspended field operations for two weeks out of concern about the health and...

Mar 18, 2020

PBS

[How the coronavirus pandemic has affected the 2020 census](#)

But the bureau has suspended all in-person census surveys until further notice, and has stalled field operations until at least April 15. And those parts of the...

Apr 1, 2020

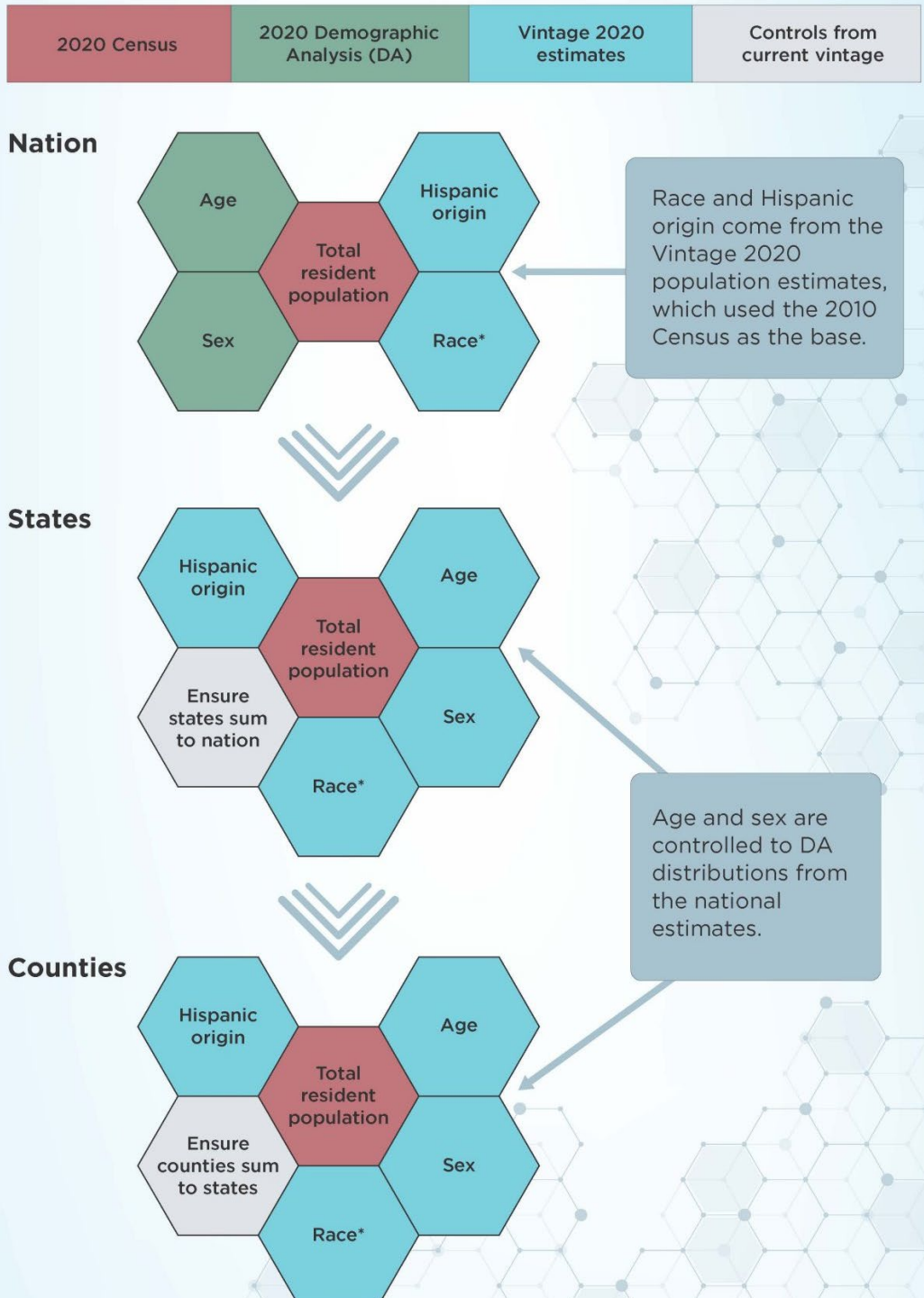
// Census.gov / By Decade / 2020 Census Decade / 2020 Census Program Management / 2020 Census Operational Adjustments Due to COVID-19

2020 Census Operational Timeline and Adjustments Due to COVID-19

share   

Conducting a decennial census involves extensive planning and complex coordination across different operations. In light of the COVID-19 pandemic, the U.S. Census Bureau adapted and shifted some of the 2020 Census operations to protect the health and safety of the public and Census Bureau staff who interacted with them. These adjustments were made to fulfill our Constitutional mandate to count the nation's population and deliver the results to the President and Congress.

Vintage 2021 Blended Base Approach

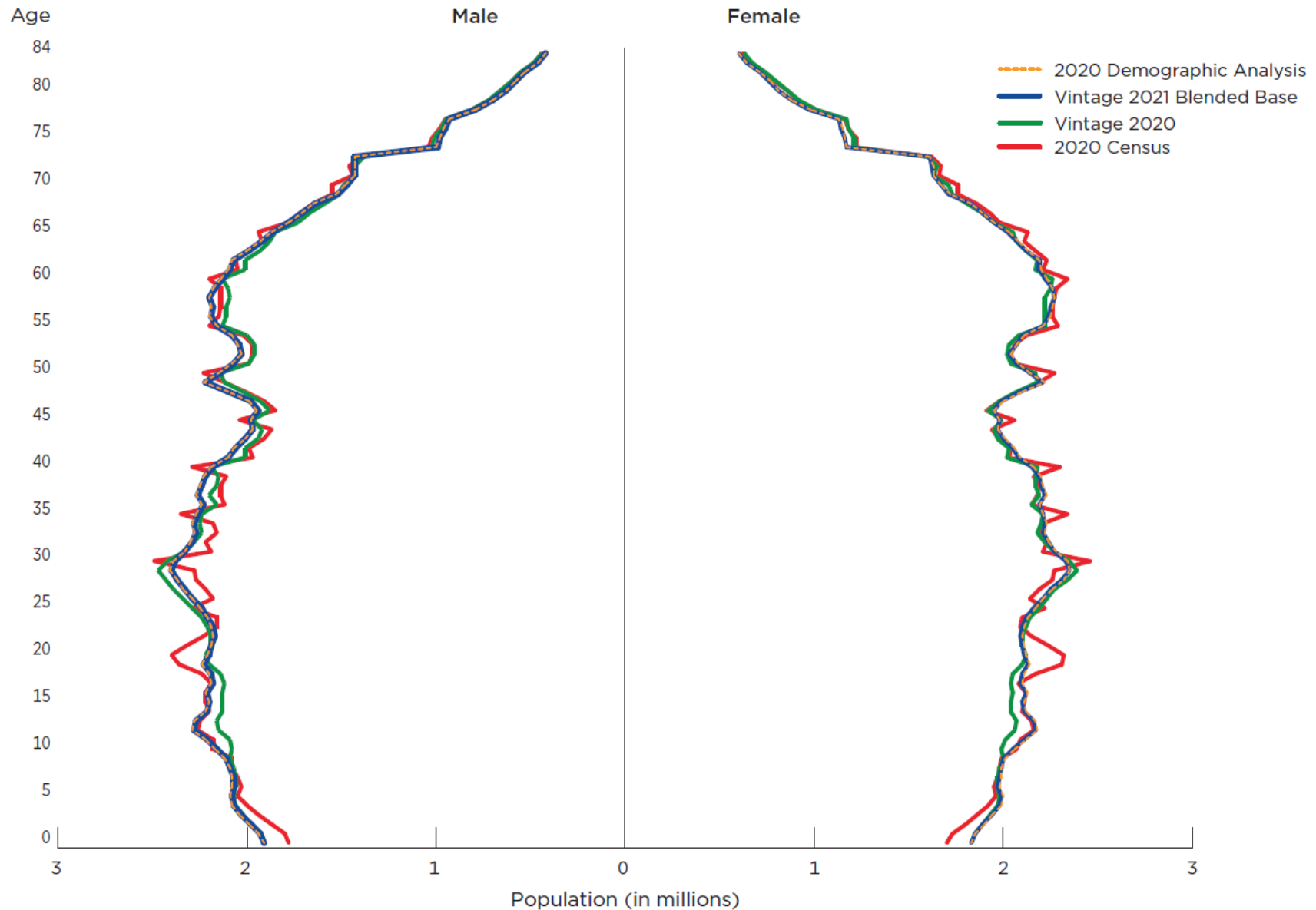


Impact on Vintage 2021

Introducing additional data sources into the estimates base:

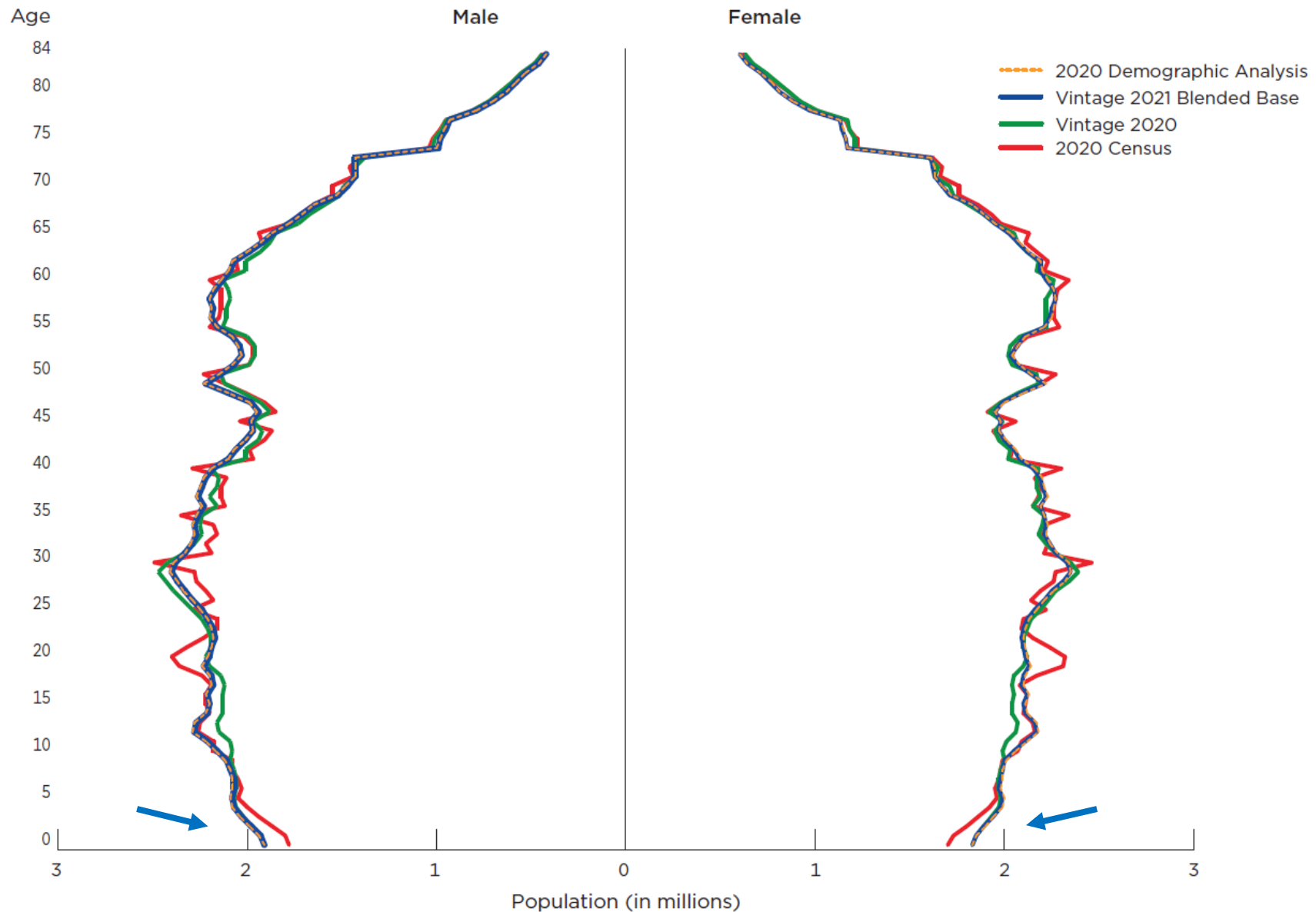
- Resulted in a smoother age/sex distribution than the 2020 Census counts.
- Had a mitigating effect on the persistent undercount of young children in the census, particularly 0 to 4-year-olds (undercounted in the 2020 Census) and 10 to 14-year-olds (undercounted in the 2010 Census and aged forward in the estimates methodology)

Vintage 2021 Blended Base, Base Inputs, and 2020 Census Data by Sex for Ages 0-84: April 1, 2020



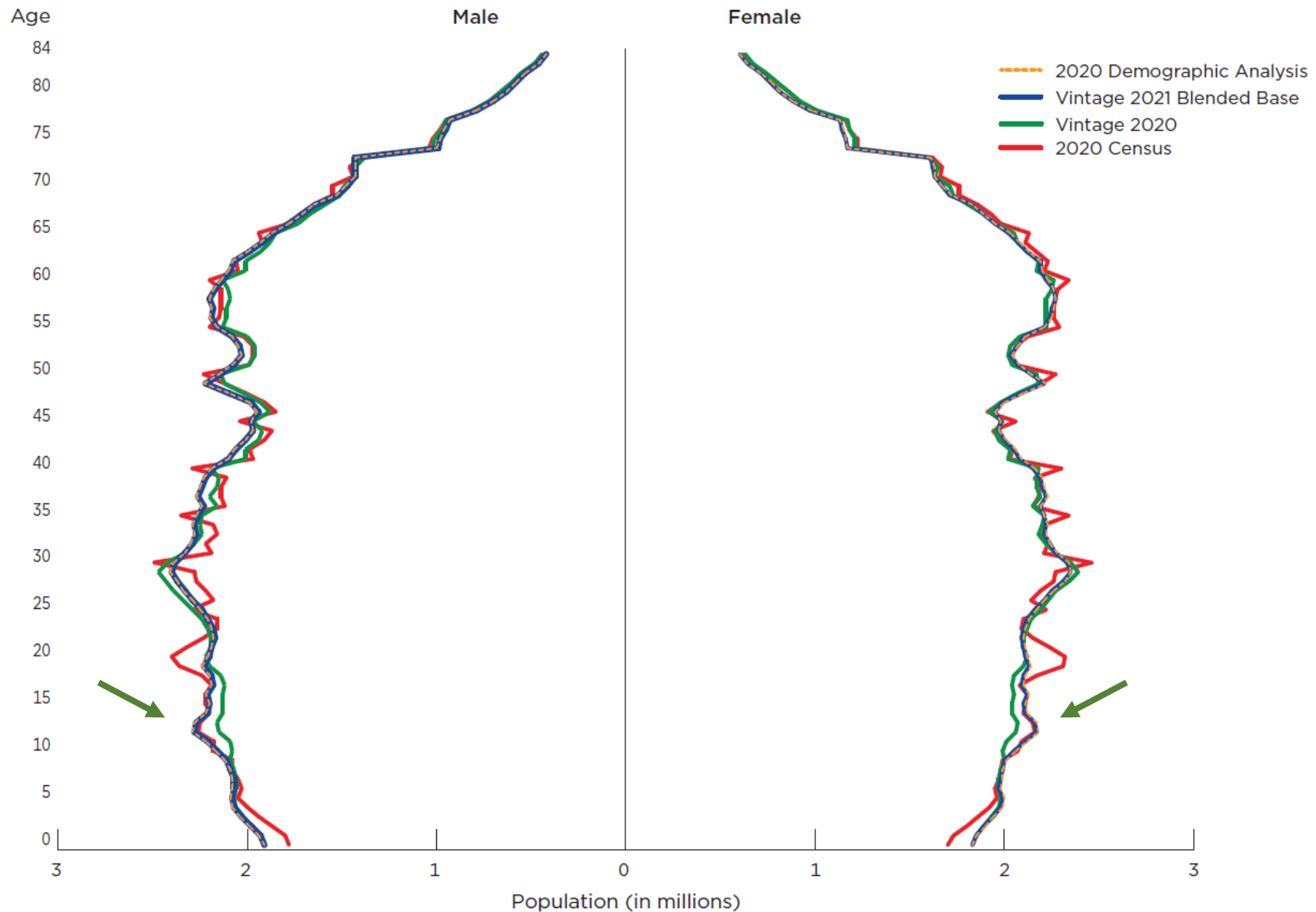
Note: These 2020 Census data by age and sex represent a special tabulation of the 2020 Census with confidentiality protections applied using the 2020 Census Disclosure Avoidance System. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. (DRB clearance number CB-FY22-DSEP-001.)
 Source: U.S. Census Bureau, 2020 Decennial Census; 2020 Demographic Analysis; Vintage 2020 and 2021 Population Estimates.

Vintage 2021 Blended Base, Base Inputs, and 2020 Census Data by Sex for Ages 0-84: April 1, 2020



Note: These 2020 Census data by age and sex represent a special tabulation of the 2020 Census with confidentiality protections applied using the 2020 Census Disclosure Avoidance System. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. (DRB clearance number CB-FY22-DSEP-001.)
 Source: U.S. Census Bureau, 2020 Decennial Census; 2020 Demographic Analysis; Vintage 2020 and 2021 Population Estimates.

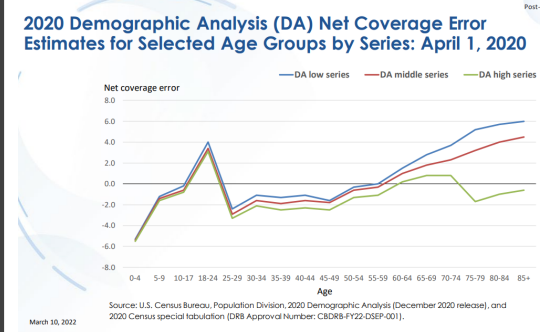
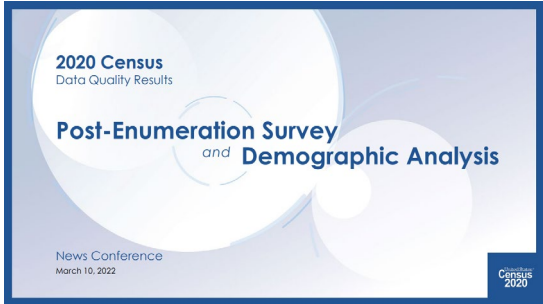
Vintage 2021 Blended Base, Base Inputs, and 2020 Census Data by Sex for Ages 0-84: April 1, 2020



Note: These 2020 Census data by age and sex represent a special tabulation of the 2020 Census with confidentiality protections applied using the 2020 Census Disclosure Avoidance System. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. (DRB clearance number CB-FY22-DSEP-001.)
 Source: U.S. Census Bureau, 2020 Decennial Census; 2020 Demographic Analysis; Vintage 2020 and 2021 Population Estimates.

For Immediate Release: Thursday, March 10, 2022

Census Bureau Releases Estimates of Undercount and Overcount in the 2020 Census



Post-Enumeration Survey and Demographic Analysis Help Evaluate 2020 Census Results

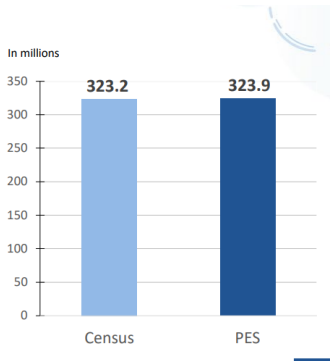
March 10, 2022 – The U.S. Census Bureau released results today from two analyses about the quality of the 2020 Census counts. While both showed the strength of the count for the total U.S. population, each analysis revealed that the 2020 Census overcounted or undercounted various demographic groups.

"Today's results show statistical evidence that the quality of the 2020 Census total population count is consistent with that of recent censuses. This is notable, given the unprecedented challenges of 2020," said Director Robert L. Santos. "But the results also include some limitations – the 2020 Census undercounted many of the same population groups we have historically undercounted, and it overcounted others."

The two analyses are from the Post-Enumeration Survey (PES) and Demographic Analysis Estimates (DA) and estimate how well the 2020 Census counted everyone in the nation and in certain demographic groups. They estimate the size of the U.S. population and then compare those estimates to the census counts.

Coverage of the Census

Census count of the population	323,200,000
Estimate of the population from the PES	323,900,000
Net coverage error	-782,000
Percent net coverage error	-0.24 %



Note: Numbers do not sum to a total due to rounding.

Coverage Concerns

- Following the release of census coverage measures, concerns about decennial census quality became apparent
- Many looked to the population estimates, and particularly the blended base, as a potential solution
 - Stakeholders, advisory groups, census leadership, media

Demographic Analysis Method

Population = Births - Deaths + Immigration - Emigration



The Base Evaluation and Research Team (BERT)

- BERT is made up of subject-matter experts in the areas of
 - Population estimates
 - Age and sex statistics
 - Coverage measurement
 - Race and ethnicity
 - Demography
 - Disclosure avoidance
- Findings from BERT research inform decisions about what 2020 Census data are used in the blended base
- BERT is working with experts from other Census projects, such as Frames, to maximize the impact of research efforts

BERT Progress

- The team has been meeting since March 2022
- Research has been focused on analyzing 2020 Census data and coverage measures
 - Comparisons to benchmark data
 - Evaluations of change since the 2010 Census
 - Assessment of demographic plausibility
- Primary subject areas have been:
 - Age and sex
 - Hispanic origin
 - The 2020 Post-Enumeration Survey (PES)

Preliminary Findings: Age and Sex

Random Samplings Blog

Age Heaping in the 2020 Census Demographic and Housing Characteristics File (DHC)

May 25, 2023

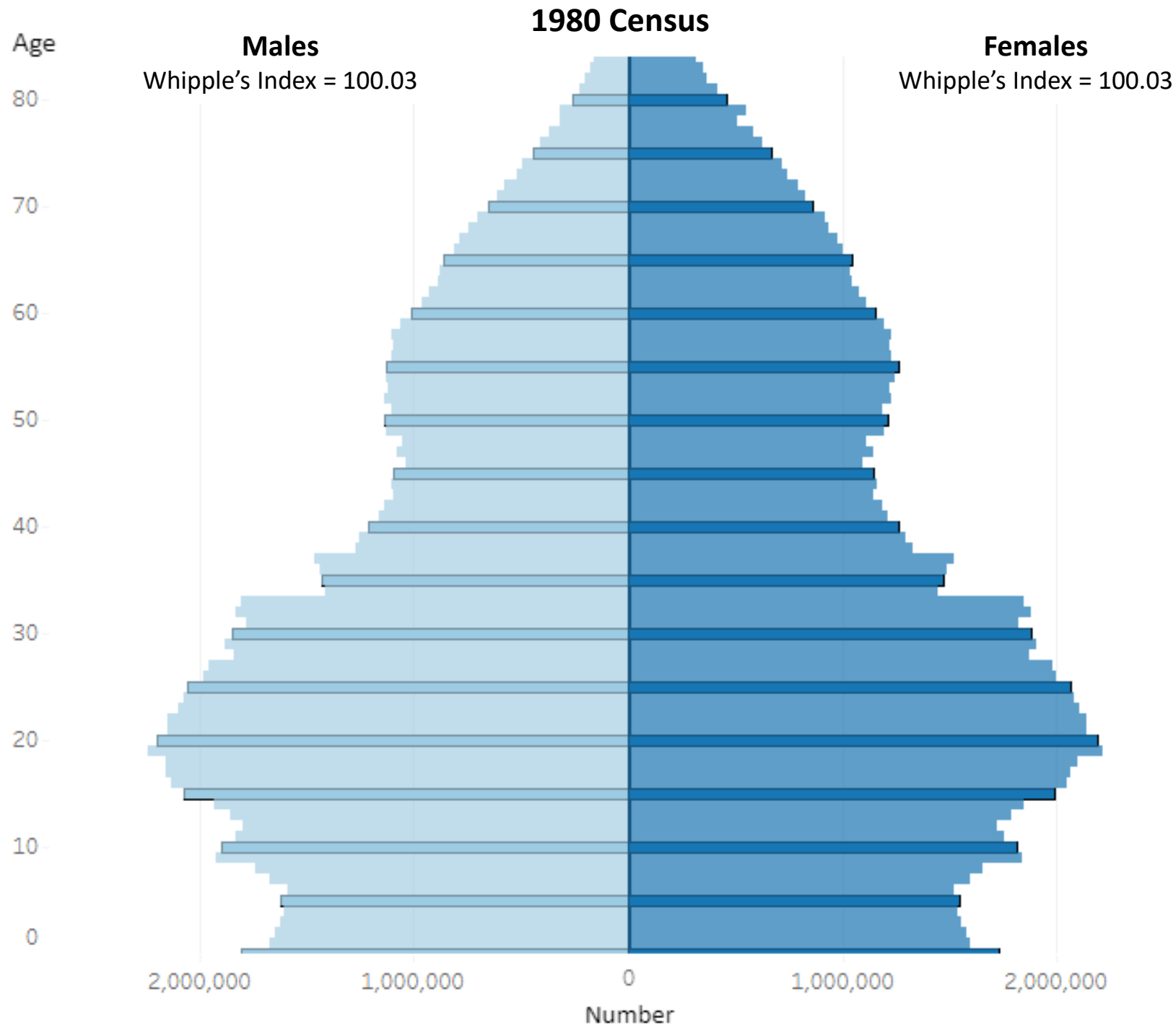
While age heaping is an issue in all decennial census and household survey data, the 2020 Census shows patterns of age heaping that are more pronounced.

Written by: Eric Jensen, Senior Advisor for Population Estimates and Coverage Measurement;
Andrew Roberts, Chief of the Sex and Age Statistics Branch;
Luke Rogers, Senior Advisor for Estimates Development and Improvement

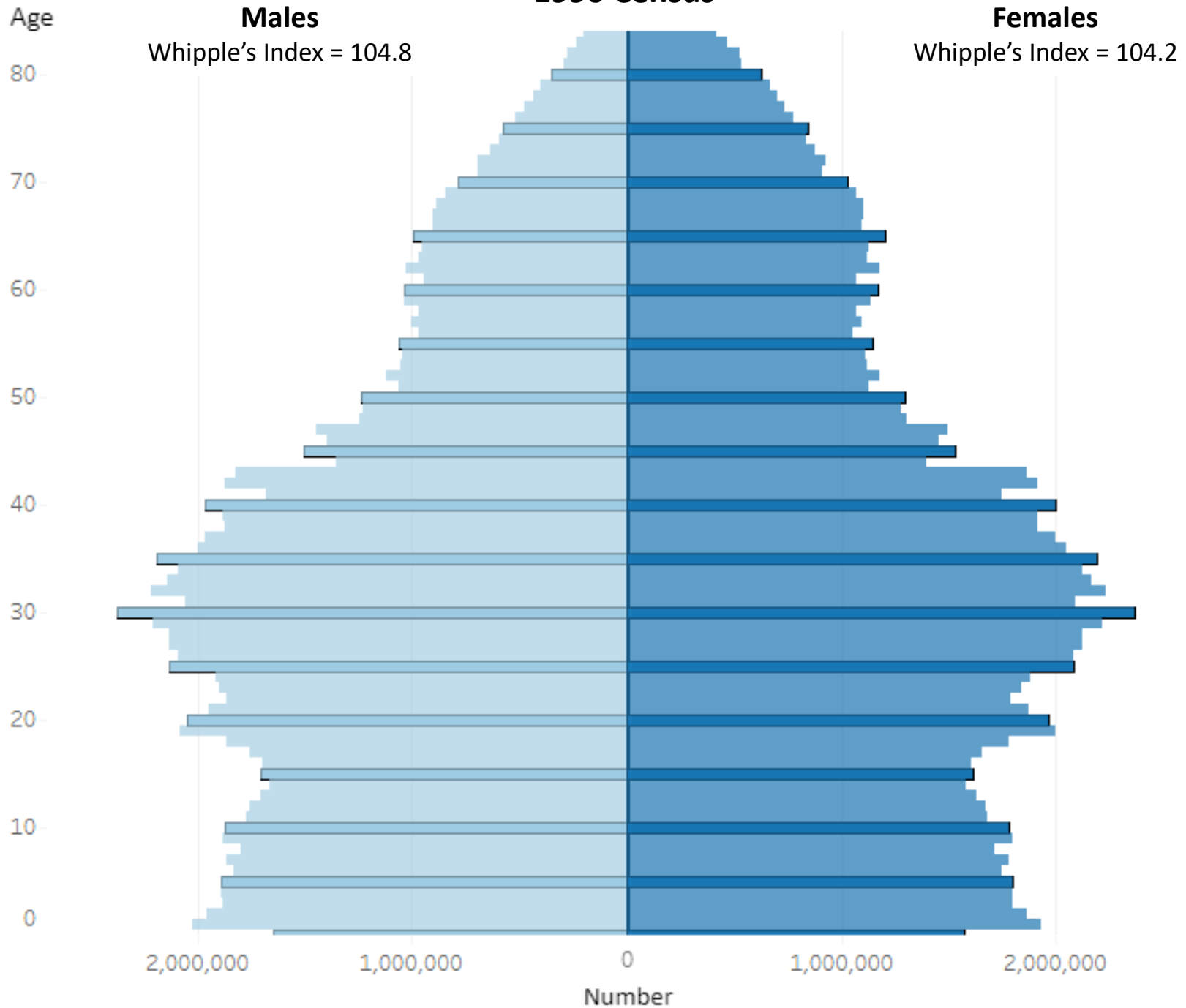
<https://www.census.gov/newsroom/blogs/random-samplings/2023/05/age-heaping-2020-census-dhc.html>

Whipple's Index Interpretation Guidance

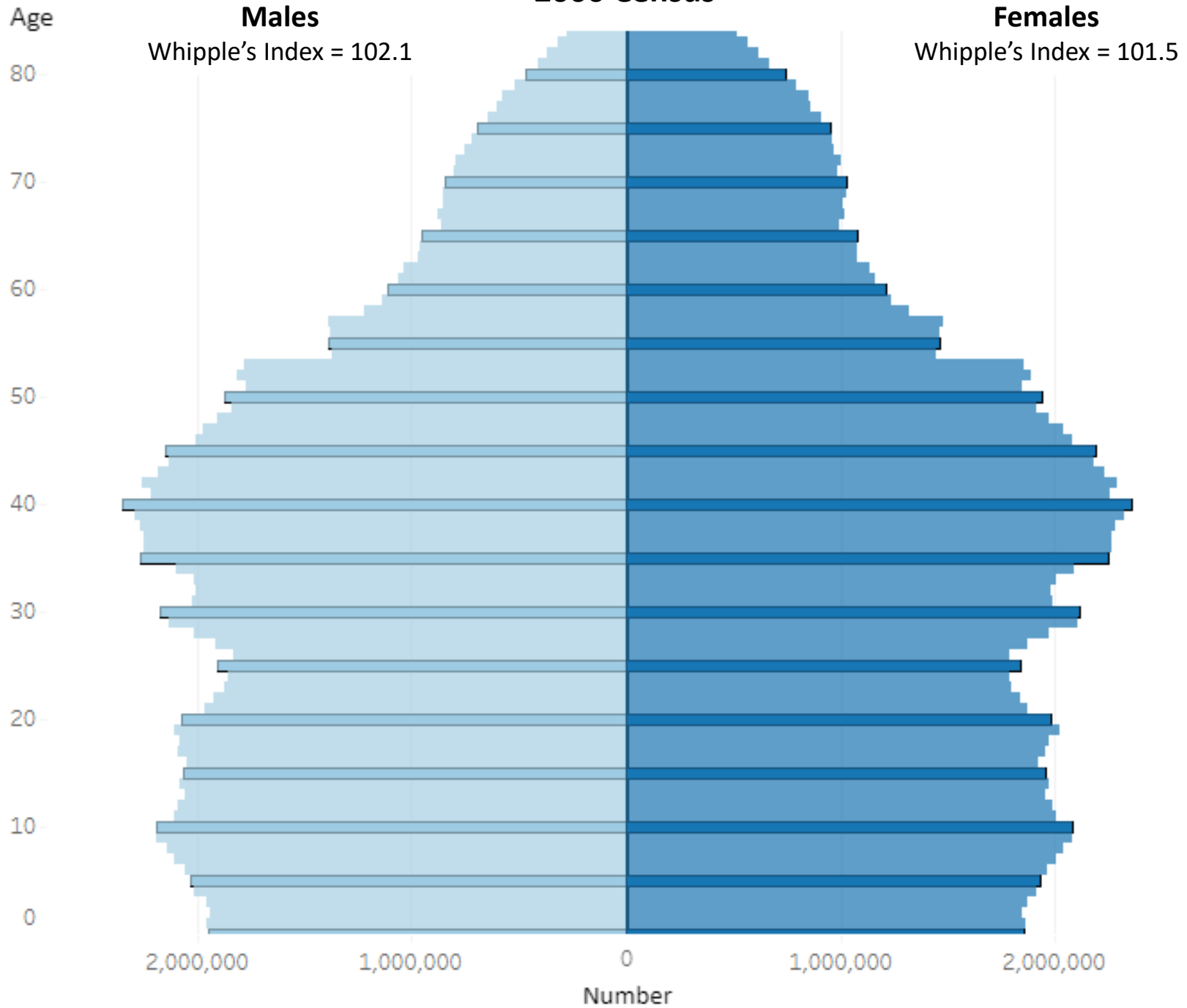
Whipple's index	Quality of data	Deviation from perfect
0	No ages ending in 0 and 5	NA
100	No preference for ages ending in 0 and 5	NA
< 105	Highly accurate	< 5%
105–109.9	Fairly accurate	5–9.99%
110–124.9	Approximate	10–24.99%
125–174.9	Rough	25–74.99%
≥ 175	Very rough	≥ 75%
500	All ages end in 0 and 5	NA

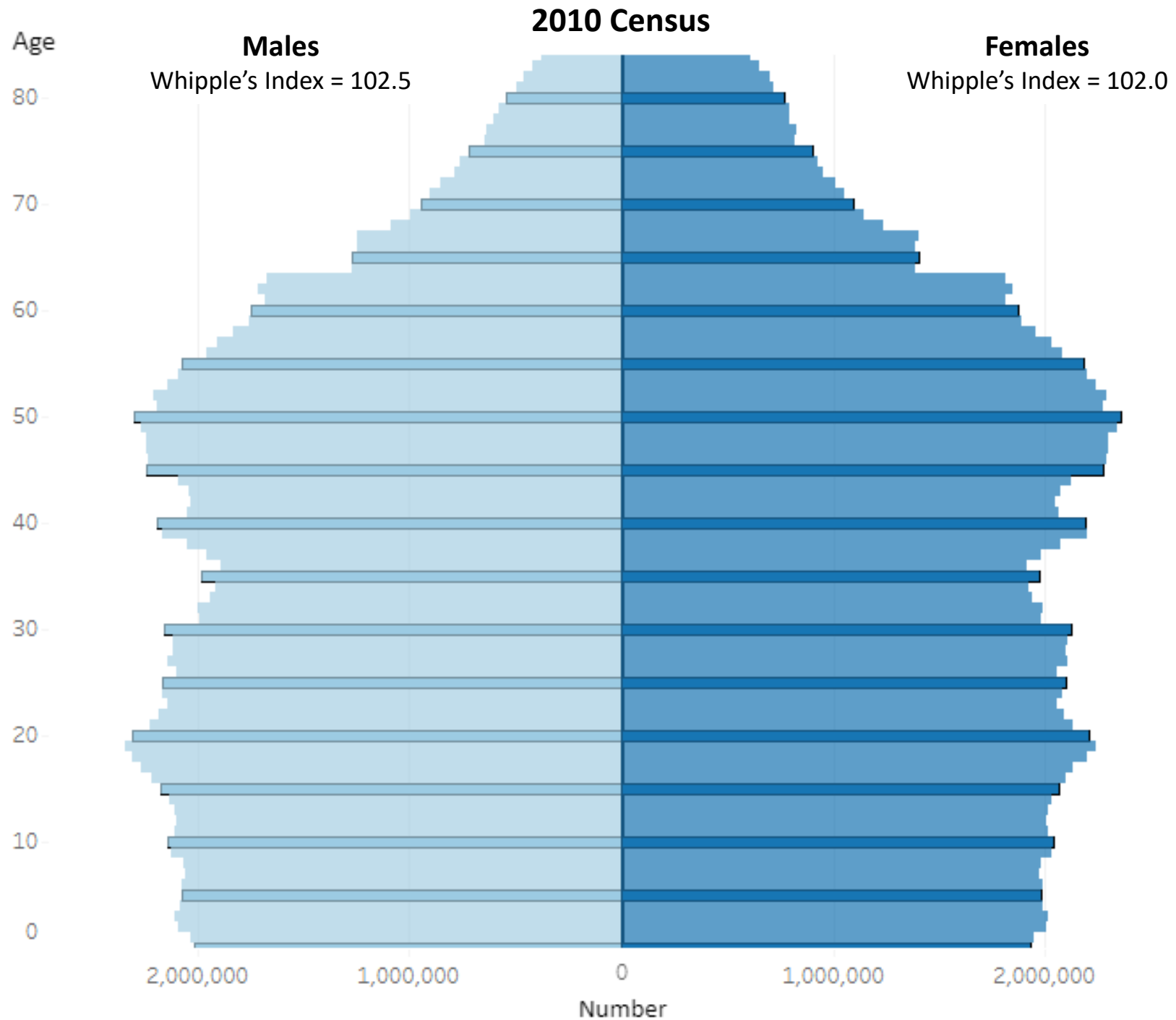


1990 Census



2000 Census





2020 Census

Age

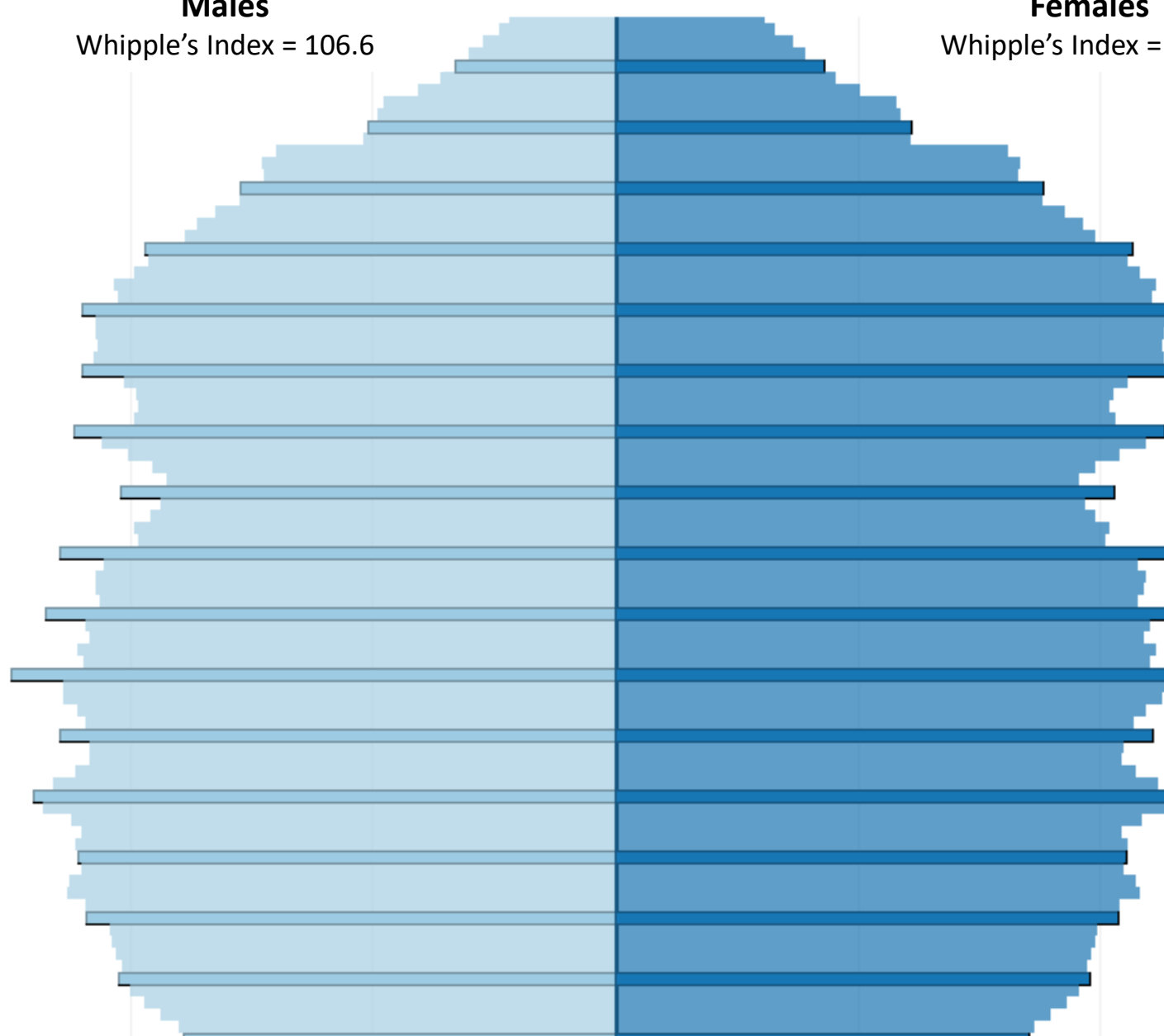
Males

Whipple's Index = 106.6

Females

Whipple's Index = 105.3

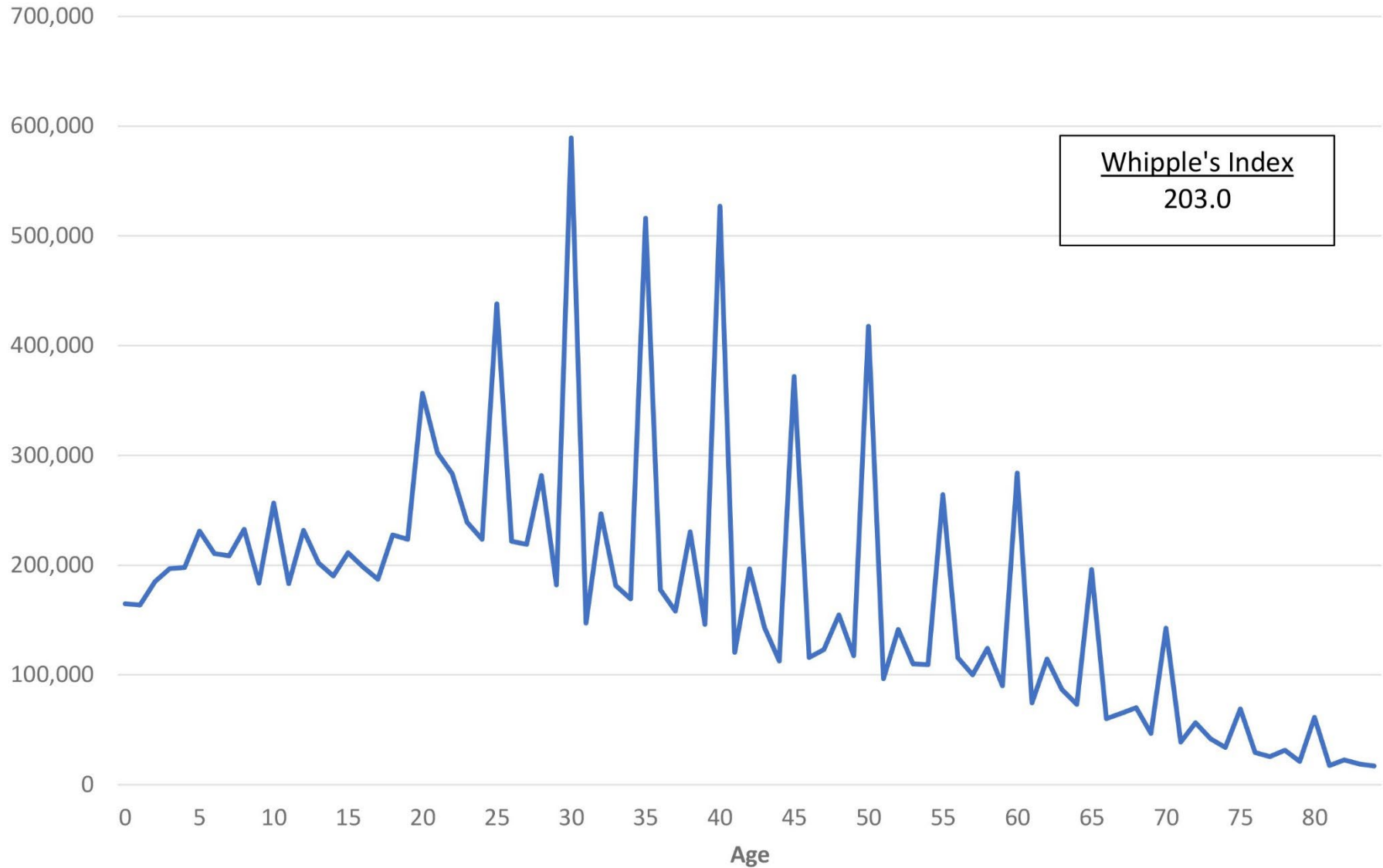
80
70
60
50
40
30
20
10
0



2,000,000 1,000,000 0 1,000,000 2,000,000

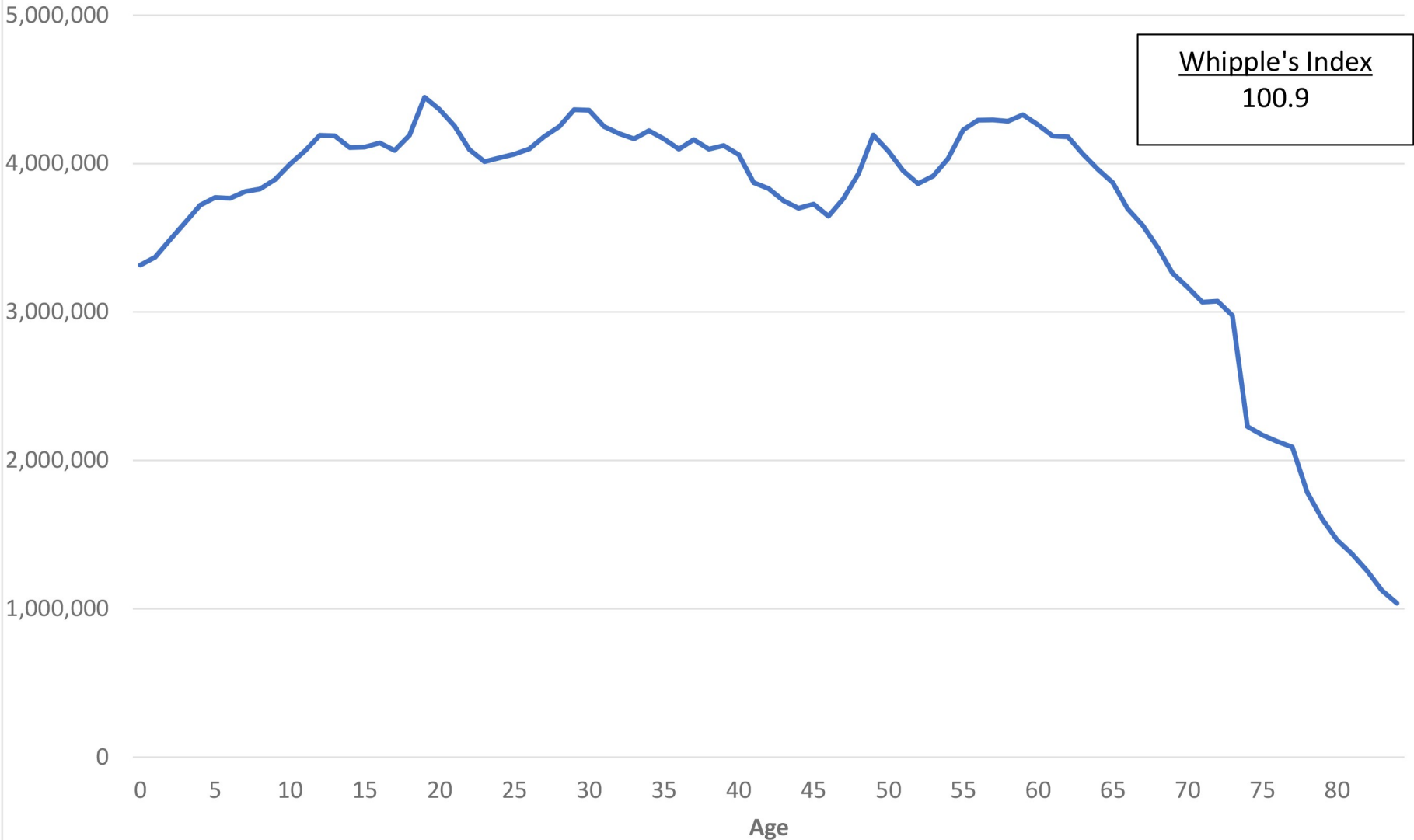
Number

Age Distribution for the Household Population Only Reporting Age in the 2020 Census



Source: U.S. Census Bureau. 2020 Census, special tabulation (DRB Approval Number: CBDRB-FY23-0314).

Age Distribution for the Household Population Reporting Age and Date of Birth in the 2020 Census



Source: U.S. Census Bureau. 2020 Census, special tabulation (DRB Approval Number: CBDRB-FY23-0314).

Next Steps: Age and Sex

- Continue investigating potential causes for the increase in heaping in the 2020 Census relative to previous censuses with an eye toward improving the 2030 Census data collection
 - Assessing possible patterns by race, ethnicity, and sex
 - Examining the influence of imputations
- Investigate sex data quality more closely to determine if correction is needed to incorporate 2020 Census sex data into the population estimates base
- Release 2020 Census “modified race” file (in development) with smoothed age and explore impact of incorporating these data into the estimates base

BERT Progress: Hispanic Origin

- BERT sought to investigate the reasonableness of the 2020 Census Hispanic origin data for use in the blended base by:
 - Examining the impact of changes to the ethnicity question
 - Benchmark analyses of the Hispanic household and group quarters populations by age and sex
 - Cohort change analysis from 2010 to 2020
 - Reviewing differences in the “As Reported” Hispanic population and CEF total resident Hispanic population

Next Steps: Hispanic Origin

- Continue benchmark analysis of Hispanic population by age and sex, using the forthcoming "modified race" file that includes the correction for age heaping
- Identify and investigate any discernible geographic patterns of discrepancy between the 2020 Census and benchmarks
- Explore alternative datasets that would be suitable for identifying under- or overcounts and potentially informing an adjustment

BERT Progress: Coverage

- Coverage experts, including PES staff, assessed the robustness of the PES results for the purposes of adjusting the population estimates base
- After a year and a half of researching the issue, BERT concluded:
 - The PES was designed to estimate the coverage of the census by independently surveying a sample of the population. It was not, however, designed to adjust the annual population estimates
 - The Census Bureau should continue to use the PES to help inform the strengths and limitations of the 2020 Census counts as well as products derived from the 2020 counts, such as the population estimates, but not use the 2020 PES results to adjust the population estimates base
 - It's essential to start planning early for a coverage measurement program that can adjust the population estimates base if there are large coverage errors in the 2030 Census

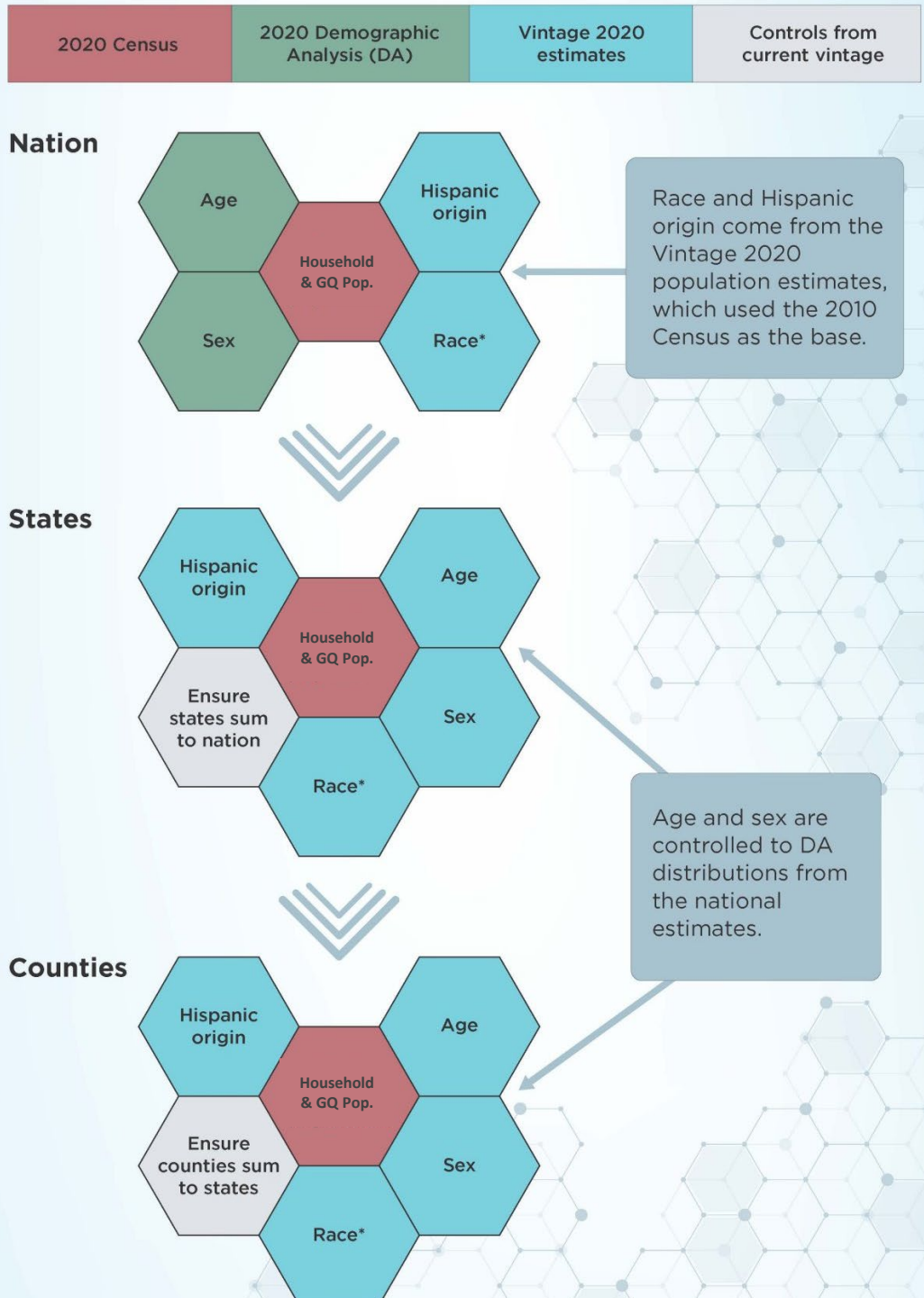
Next Steps: Coverage

- Demographic Analysis
 - Once the 2020 Census Modified Race File is ready, we will release estimates of net coverage error by age and sex for the Black / non-Black population and Hispanic / non-Hispanic (ages 0 to 29) populations
 - We will research the potential to include the results in the blended base
- PES
 - Continue to explore the use of administrative data to develop estimates of net coverage error for the 2030 Census
 - This work is ongoing as part of the Continuous Count Study
 - As this work progresses, we may be able to incorporate the results into BERT
 - Everything is still in its early stages

Upcoming Research: Group Quarters

- Findings from all subject areas indicated distinct differences in the GQ population compared to the household population, so further GQ research will take place as a dedicated subject area across characteristics
- Investigate any potential issues with sex reporting for GQs
- Continue benchmark comparison of GQ population by Hispanic origin and by race with a focus on GQ type
- Explore how to handle GQs when using administrative data in the dual-system estimation approach for estimating net coverage error

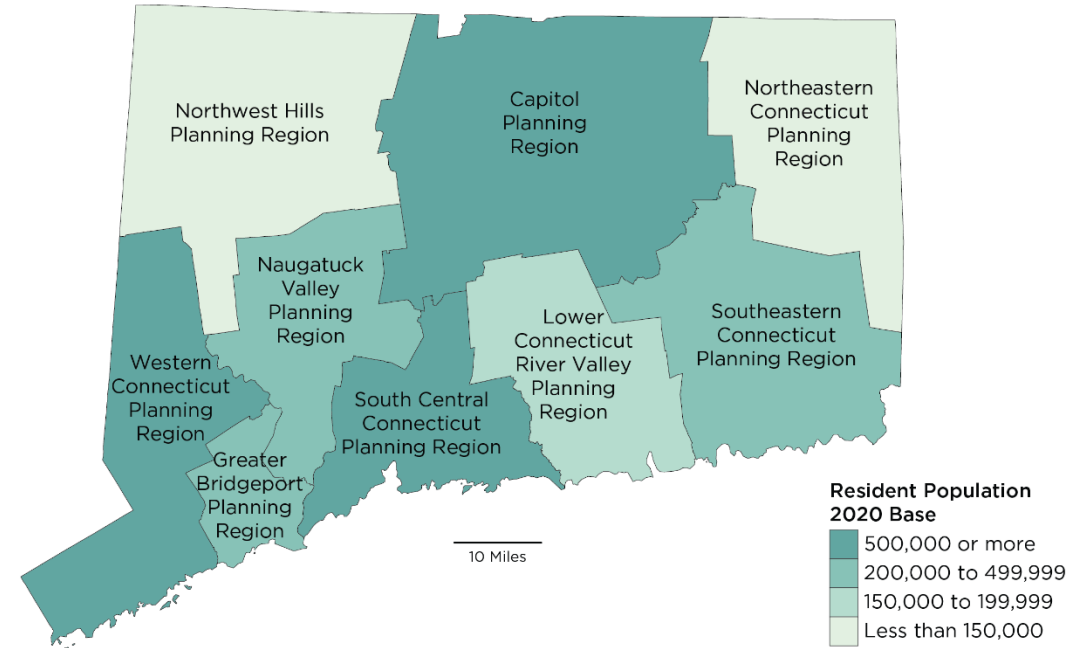
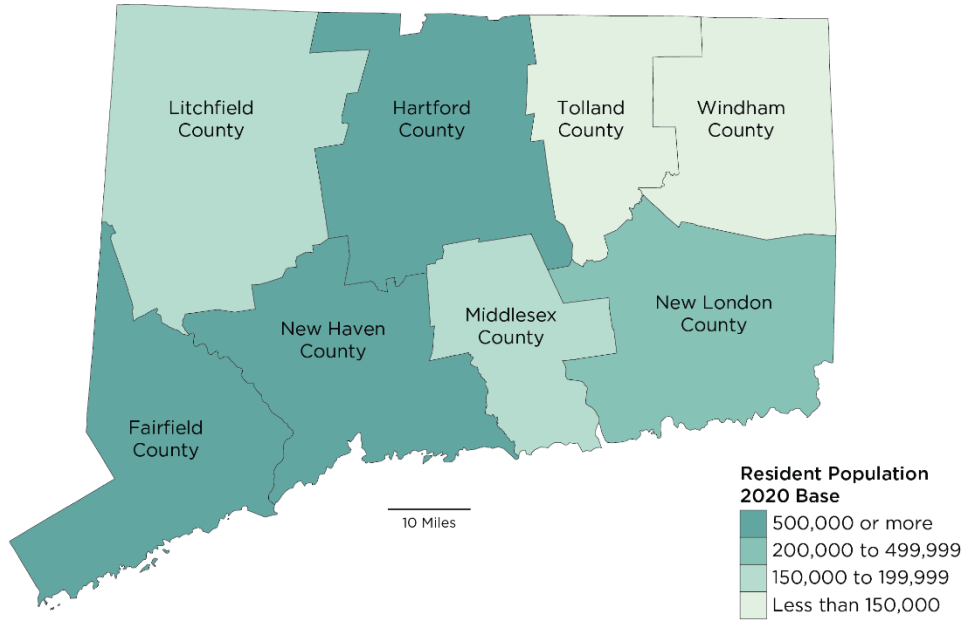
Vintage 2022 Blended Base Approach



Impact on Vintage 2022

- Incorporating subcounty GQ and HH data from the 2020 Census enabled us to:
 - Capture geographic boundary changes, such as in Connecticut (see maps)
 - Incorporate GQ change from April 2020 to July 2022

Impact on Vintage 2022 (cont.)



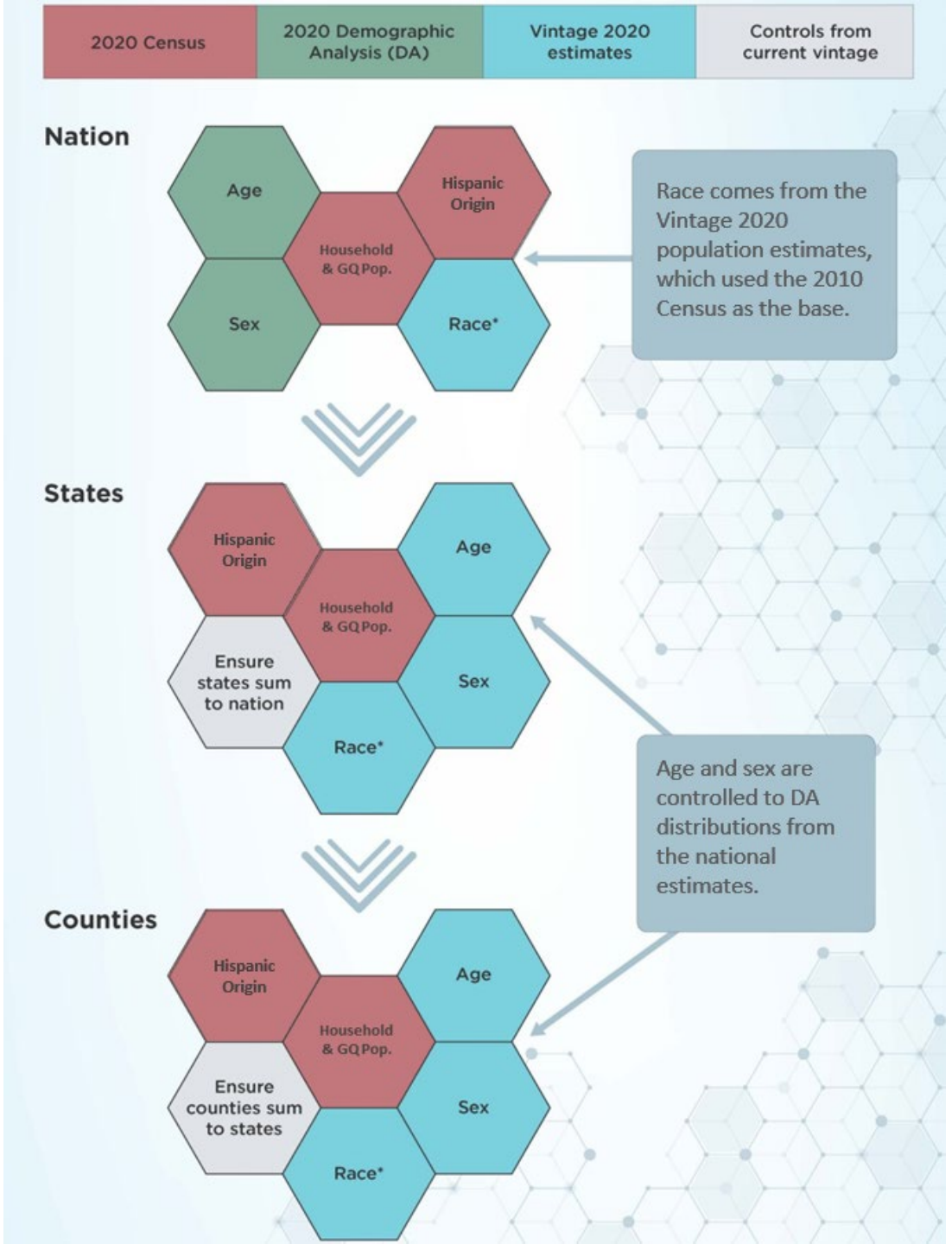
Impact on Vintage 2022 (cont.)

- We also enhanced our method of processing GQ data. The GQ population increased to better approximate Census population counts, as seen below:

Population Estimates and Differences by Vintage and Universe: April 1, 2020

	Population Universe		
	Resident	Household	Group Quarters
Vintage 2021	331,449,281	323,688,422	7,760,859
Vintage 2022	331,449,520	323,210,615	8,238,905
Numeric Difference	239	-447,807	478,046
Percent Difference	0.0%	-0.1%	6.2%

Vintage 2023 Blended Base Approach



Impact on Vintage 2023

Processing improvements included:

- A more effective method for estimating the GQ population
- An approach to avoid applying demographically implausible GQ distributions from Vintage 2020 if donor populations were too small
- These caused broad changes to the GQ population

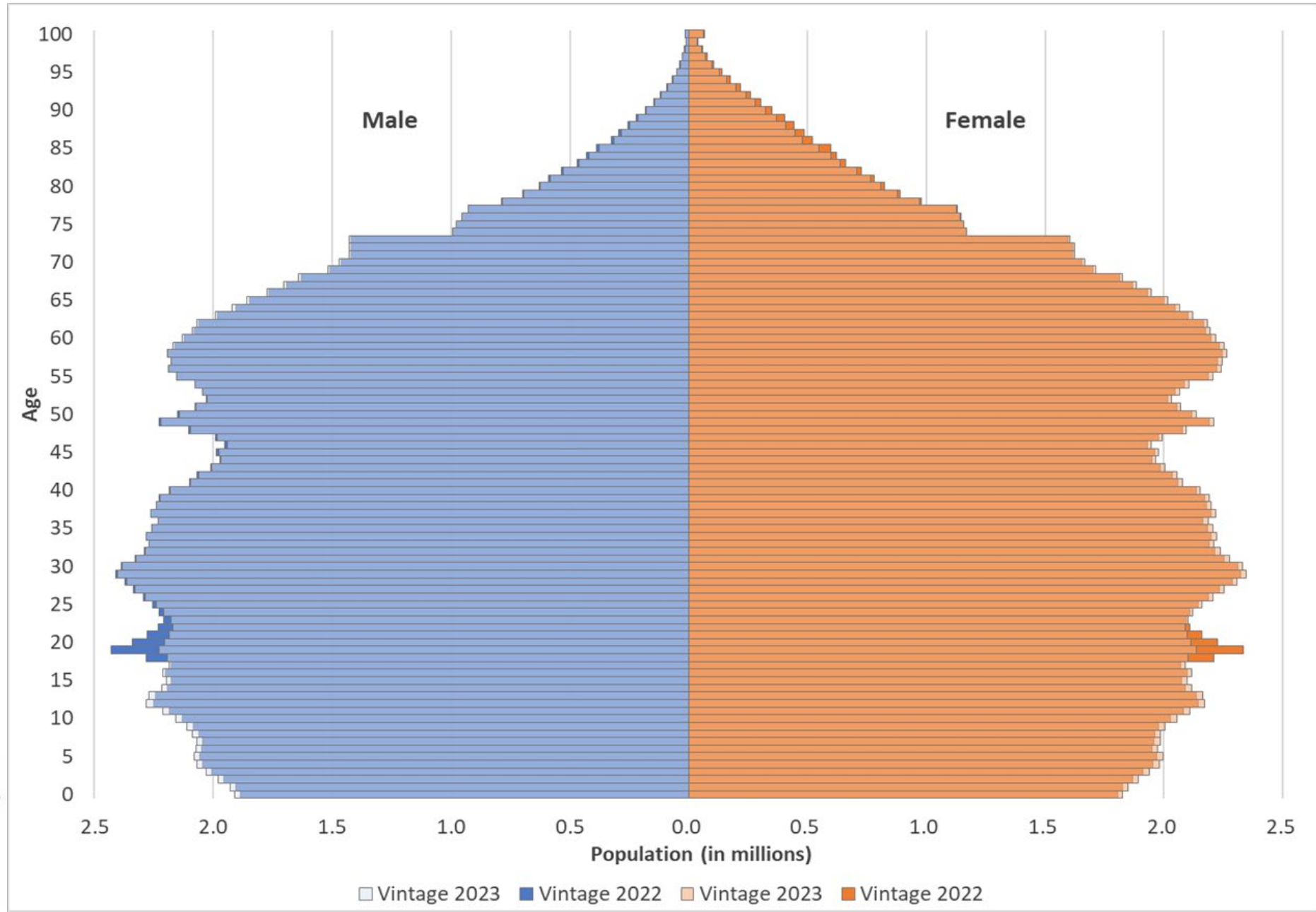
Degree of Population Change for Counties between Vintage 2022 and Vintage 2023 by Universe: April 1, 2020

Absolute Percent Difference	Change in Resident Population			Change in Household Population			Change in Group Quarters Population		
	Gain	Loss	None	Gain	Loss	None	Gain	Loss	None
5.00% or more	1	0	0	0	0	0	54	43	0
2.50 to 4.99%	1	0	0	0	0	0	82	84	0
1.00 to 2.49%	2	0	0	1	0	0	220	196	0
0.00 to 0.99%	1,408	1,437	295	1,320	1,394	429	927	960	511

Impact on Vintage 2023 (cont.)

- Incorporating county-level Hispanic origin resident population totals from the 2020 Census decreased the population among older ages and increased the population among younger ages.

Estimates of Resident Population in Vintage 2022 and Vintage 2023 by Age and Sex: April 1, 2020



Impacts Thus Far

- Original iteration:
 - Resulted in a smoother age/sex distribution than the 2020 Census counts
 - Had a mitigating effect on the persistent undercount of young children in the census, particularly 0 to 4-year-olds (undercounted in the 2020 Census) and 10 to 14-year-olds (undercounted in the 2010 Census and aged forward in the estimates methodology)
- Subsequent iterations:
 - Captured geographic updates
 - Improved the processing of group quarters population data
 - Incorporated resident population by Hispanic origin from the 2020 Census

BASE POPULATION UPDATES BY VINTAGE AND DATA SOURCE: 2020 TO 2023

Vintage	Geographic Updates	CQR	PCGQR	Most Recent Census (Nation, States, and Counties)						2020 Demographic Analysis (National)						Vintage 2020 (Nation, States, and Counties)						
				Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin	Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin	Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin	
2020	X	X		X	X	X	X	X	X													
2021	Subcounty & HU only				X							0 - 85+	X				X		X	X	X	X
2022	X	X		X	X							0 - 100+	X				X		X	X	X	X
2023	X	X	X	X	X				X			X	X				X		X	X	X	X

Lessons Learned

The blended base approach is complex.

Once the base population becomes adaptive, the resulting possibilities are numerous.

A “gold standard” can evolve.

Thank you.

Questions?

Coordination, Dissemination, and Outreach Branch - pop.cdob@census.gov

Base Evaluation and Research Team - pop.bert@census.gov

<https://www.census.gov/popest>