Demographic trends in New York State

JAN VINK

CORNELL PROGRAM ON APPLIED DEMOGRAPHICS

Cornell University

Program on Applied Demographics

CORNELL POPULATION CENTER

Overview

- Census
  - Why and how?
  - Census quality
  - Interpreting Census 2020 data
    - Disclosure Avoidance
    - Changes in race/ethnicity data collection and processing
  - Census 2020 results
- Population Estimates & Projections
  - Background and results
- Implications for the American Community Survey

Census

- Every 10 years, reference date April 1st
- Purpose
  - Apportionment of House seats
  - Redistricting
- Other uses
  - Distribution of funds
  - Planning
  - Demographic change

- Questions asked:
  - How many people in the household?
  - For every person:
    - Age, sex, Hispanic Origin, race, relationships with householder?
  - Is the house owned or rented?
- Household relations used to tabulate statistics on household types and size, etc.

Census data collection

Mission: "Count everybody once, only once and in the right place"

Step 1: Start with address frame

Step 2: Alert everyone how they can get counted (address linked to IDs)

- Type of enumeration areas
  - Update/Leave
  - Mail-out/mail back
    - Internet first, Internet choice, multi-lingual in some areas

Step 3: Collect responses over internet, mail and phone (ID not necessary)

Census data collection

Step 4: Non-response follow-up

- Use of administrative data for efficiency
- Use of proxy interviews

Step 5: Editing of responses and filling gaps

- Un-duplication
- Editing rules
- Item non-response
- Whole household imputations

Census quality

Quality of data collection

- Self-response rates
- Imputation rates
- Address deletions

Etc.

Quality of data collected

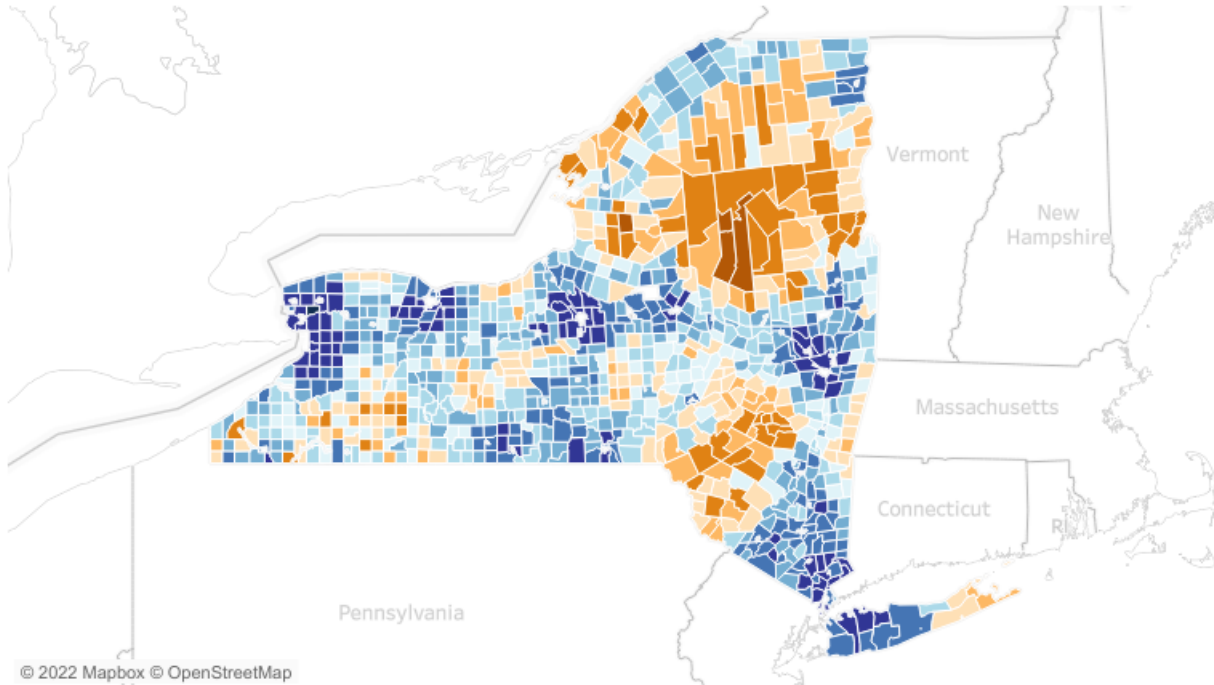
- Post Enumeration Survey
  - Independent Survey that is linked back to Census responses
- Demographic Analyses
  - Independent estimate of US population by single years of age and some race detail

Self-Response by Township

This map features self-response rates from households that responded to the 2020 Census online, by mail, or by phone.

New York
Self-Response
64.2%

Adams, New York
Self-Response
62.4%



Select Mode
Total

Select State
New York

Select Town
All

Geographies

- Return to State
- County
- City
- Congressional District
- Town and Township
- Tribal Area

Self-Response Rate (%)

0-15 | 16-30 | 31-40 | 41-50 | 51-56 | 57-62 | 63-68 | 69-74 | 75-85 | 86-100

Adams, New York Self-Response Rate

Historical Data

Email us at responseratemap@census.gov
Responses received as of 1/28/2021.
Click here for technical details

Shape your future START HERE >
United States Census 2020

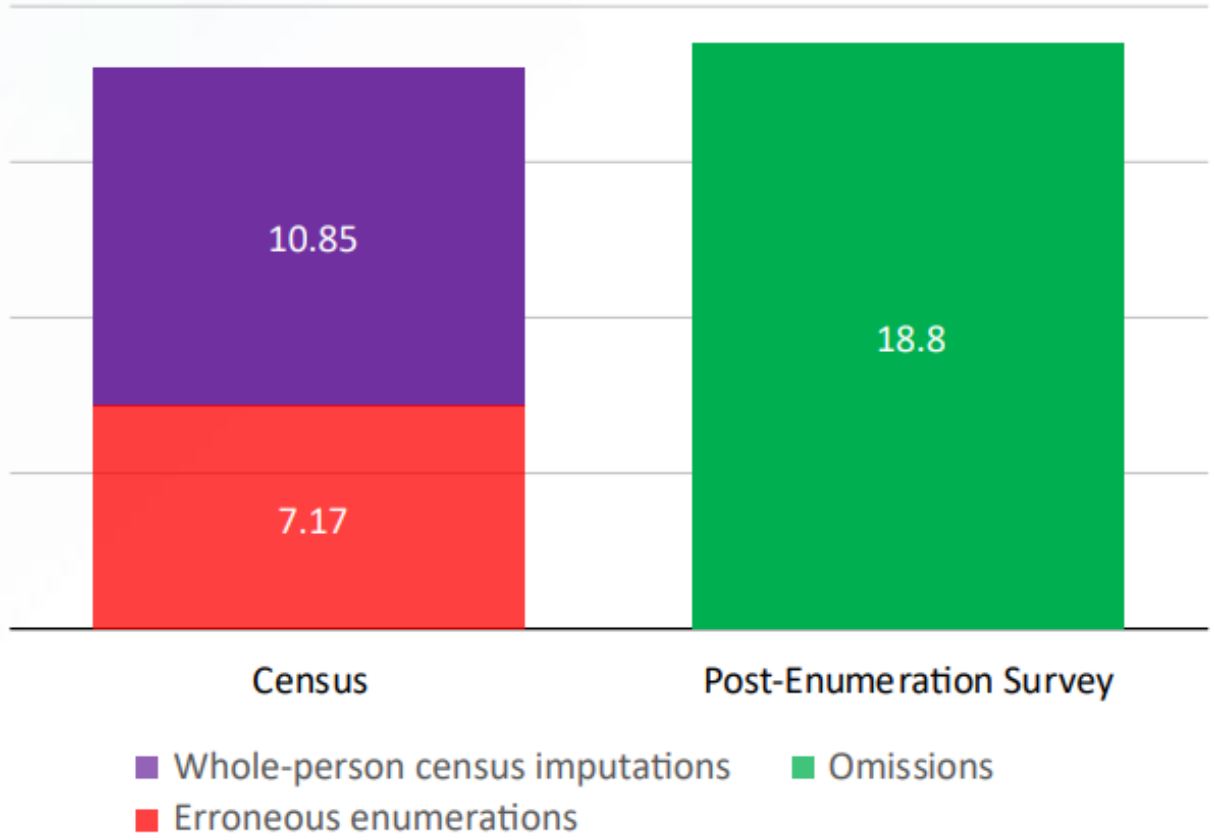
Census process quality indicators

Table 1: 2020 Census Operational Quality Metrics*

	2020		2010	
	United States	New York	United States	New York
Self-response rate (overall)	70%	68%	63%	62%
ID Processing	63%	59%	63%	62%
Non-ID Processing	8%	10%	<0.5%	<0.5%
Enumerated by Non-Response Follow-Up (NRFU)	27%	29%	32%	32%
Percent Count Imputed	< 0.5%	1%	<0.5%	<0.5%

*Metrics include all occupied and vacant housing units and exclude observations deleted during processing, and group quarters

Components of the Census Coverage in 2020 Level (In millions)

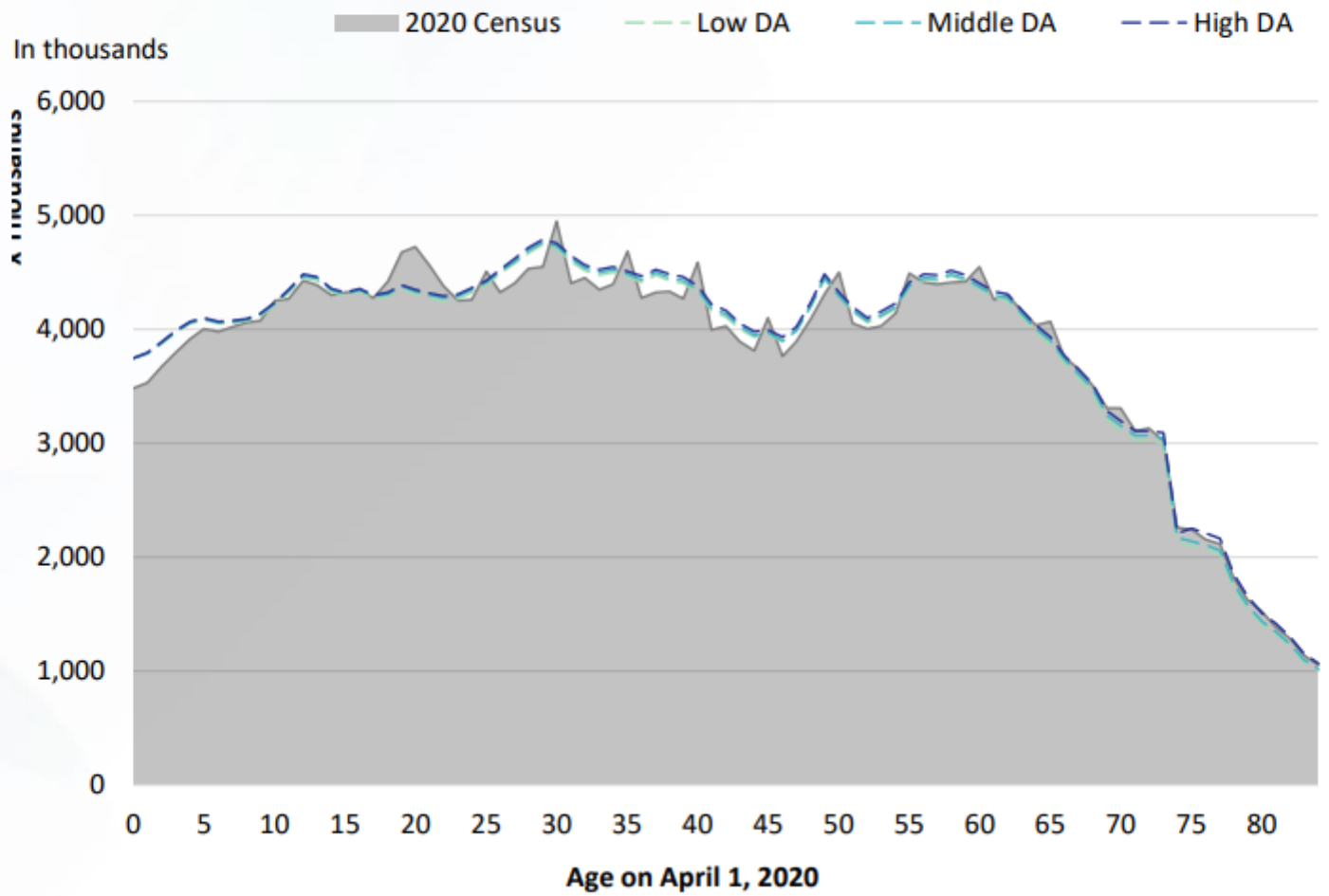


United States Census 2020

Percentage Net Coverage Error by Race and Hispanic Origin 2010 and 2020

Race or Hispanic Origin	2010	2020
Total	0.01	-0.24
Race alone or in combination		
White	0.54*	0.66*
Non-Hispanic White alone	0.83*	1.64*
Black or African American	-2.06*	-3.30*
Asian	0.00	2.62*
American Indian or Alaska Native	-0.15	-0.91*
On Reservation	-4.88*	-5.64*
American Indian Areas Off Reservation	3.86	3.06
Balance of the United States	0.05	-0.86*
Native Hawaiian or Other Pacific Islander	-1.02	1.28
Some Other Race	-1.63*	-4.34*
Hispanic or Latino	-1.54*	-4.99*

Resident Population by Single Year of Age in the 2020 Census and 2020 Demographic Analysis (DA) Estimates

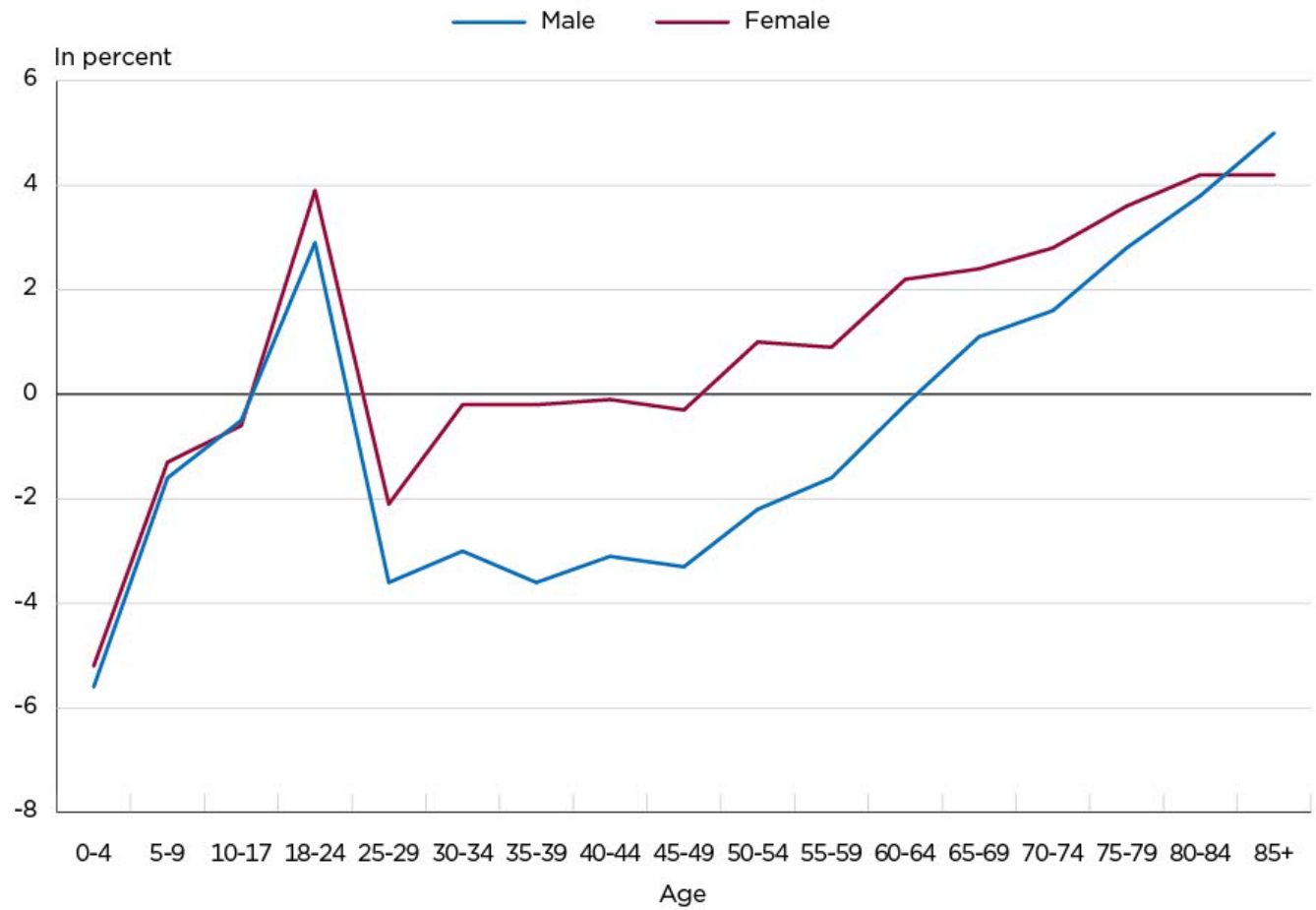


Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

United States Census 2020

Figure 3.

2020 Demographic Analysis Middle Series Net Coverage Error Estimates for Selected Age Groups by Sex: April 1, 2020



Note: 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.

Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

Census data products

Data products

- Apportionment data — April 26'th 2021
  - State totals (20,201,249 residents, 26 seats in House)
- Redistricting data — August 12'th 2021
  - Voting age, race/Hispanic origin by Census Block
- Demographic and Housing Characteristics (DHC) — May 2023
  - Detailed age, sex, household type, etc.
  - Balance between demographic detail and geographic detail
- Detailed DHC — TBD
  - Detailed race
  - Family relations

Disclosure avoidance system

The Census is regulated by title 13, U.S. Code

One of the provisions forbids the Bureau to disclose or publish any private information that identifies an individual

A Disclosure Avoidance System avoid release of private information by:

- Not releasing all the data collected
  - Suppression
- Releasing data that is slightly different from what is enumerated
  - Swapping (1990 – 2010)
  - Differential Privacy (new in 2020)

Differential Privacy

Every piece of data adds a quantifiable amount of information and with that a risk of disclosure

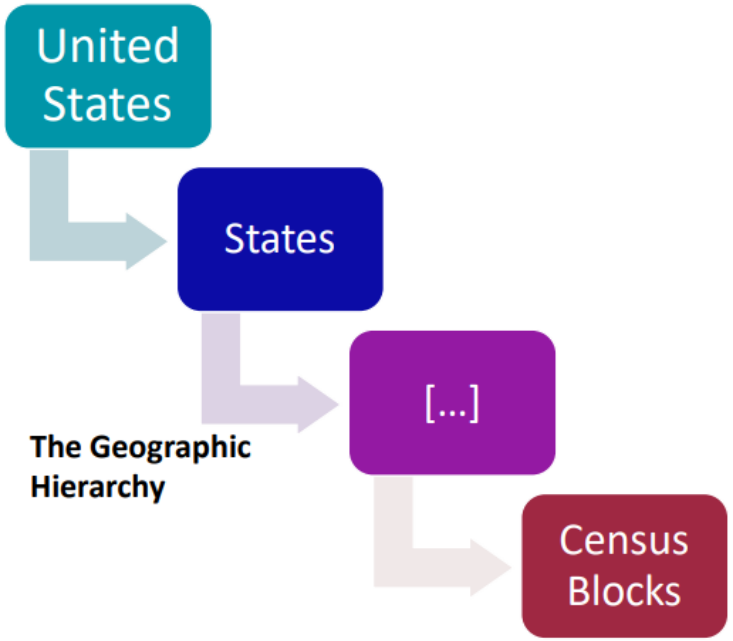
Adding noise to the data limits that risk of disclosure

How much noise?

- Too much noise can impact the accuracy and usability
- Too little noise can impact the disclosure risk

Top Down Algorithm

Census Bureau's implementation of a differential privacy framework



2020 DAS: Splitting person data from housing unit data

Person variables
Age
Sex
Race
Ethnicity
Relationship
Household/GQ type

Housing unit variables
Occupancy status
Tenure
Race/ethnicity Hholder
Age HHolder
Household type
Household size

Example noise: My block - 2020

My own count

- 4 Occupied houses
- 7 NH White adult + 4 NH White youth

Published PL94-171 (noise added)

- 4 Occupied houses
- 4 NH White adults + 6 NH White youth + 3 NH Asian youth

Top Down Algorithm analyses

Census Bureau rereleased 2010 Census counts with new DAS

Less accuracy for small counts

- Areas with few people
- Unique populations

Most error is random and washes out when aggregating

Many small areas with impossible or improbable counts

How will Differential Privacy affect data in the 2020 Census?

The Census Bureau has changed its privacy protection method to Differential Privacy. Instead of other methods formerly used, such as swapping, this method uses an algorithm called the Top-Down Algorithm (TDA) to randomly inject error into data, making it harder for computers to identify personal data. Almost all data in the redistricting product is subject to the TDA. Much of it can still be considered reliable, while other data points should be used with caution. Below is a guide produced by the Maine State Data Center for using data from the 2020 Redistricting Data product.

Green Light

This data is highly reliable. It has low levels of error and can be used as usual.

Green light variables:

- State-level total population, all other demographic & housing variables
- County-level total population, race & ethnicity except for Native Hawaiian and Pacific Islander, all other housing & Group Quarters variables
- Total population in medium and large cities & towns (>500 population)
- Total Population in Census Tracts
- Number and type of occupied group quarters units at the block level
- Total housing units at the block level
- Any data point at the county or county subdivision/place level with a count of at least 250
- Any data point at the Tract level with a count of at least 500

Yellow Light

This data is slightly less reliable, but can still be used with caution. Understand that these variables may have moderate error.

Yellow light variables:

- Race and ethnicity in large towns & cities (>5,000 population)
- Total population in small towns & cities
- Any cell at the county or county subdivision level that has a count between 60-250
- Any data point at the Tract level with a count between 250-500

Red Light

This data is subject to high levels of error. Use these variables with extreme caution or consider aggregating them together to mitigate error.

Red light variables:

- Block-level data
- Race and ethnicity data in small towns
- Total population in county subdivisions with fewer than 60 population
- Any statistic that is divided across tables; for example, persons per household
- Any data point at the county subdivision level that has a count of less than 60
- Any data point at the tract level with a count of less than 250

Impossible and improbable blocks (NY 2020)

	2010		2020	
	Count	**% of all**	**Count**	**% of all**
Non empty blocks	250,070		233,182	
Households (occupied houses) and household population				
Household population > 0, but occupied houses = 0	Impossible in 2010		14,276	6.10%
Household population = 0, but occupied houses > 0			1,834	0.80%
PPH > 10	53	0.00%	4,510	1.90%
Youth only				
Only 0-17	21	0.00%	2,808	1.20%
Without GQ and only 0-17	1	0.00%	2,795	1.20%

Census results

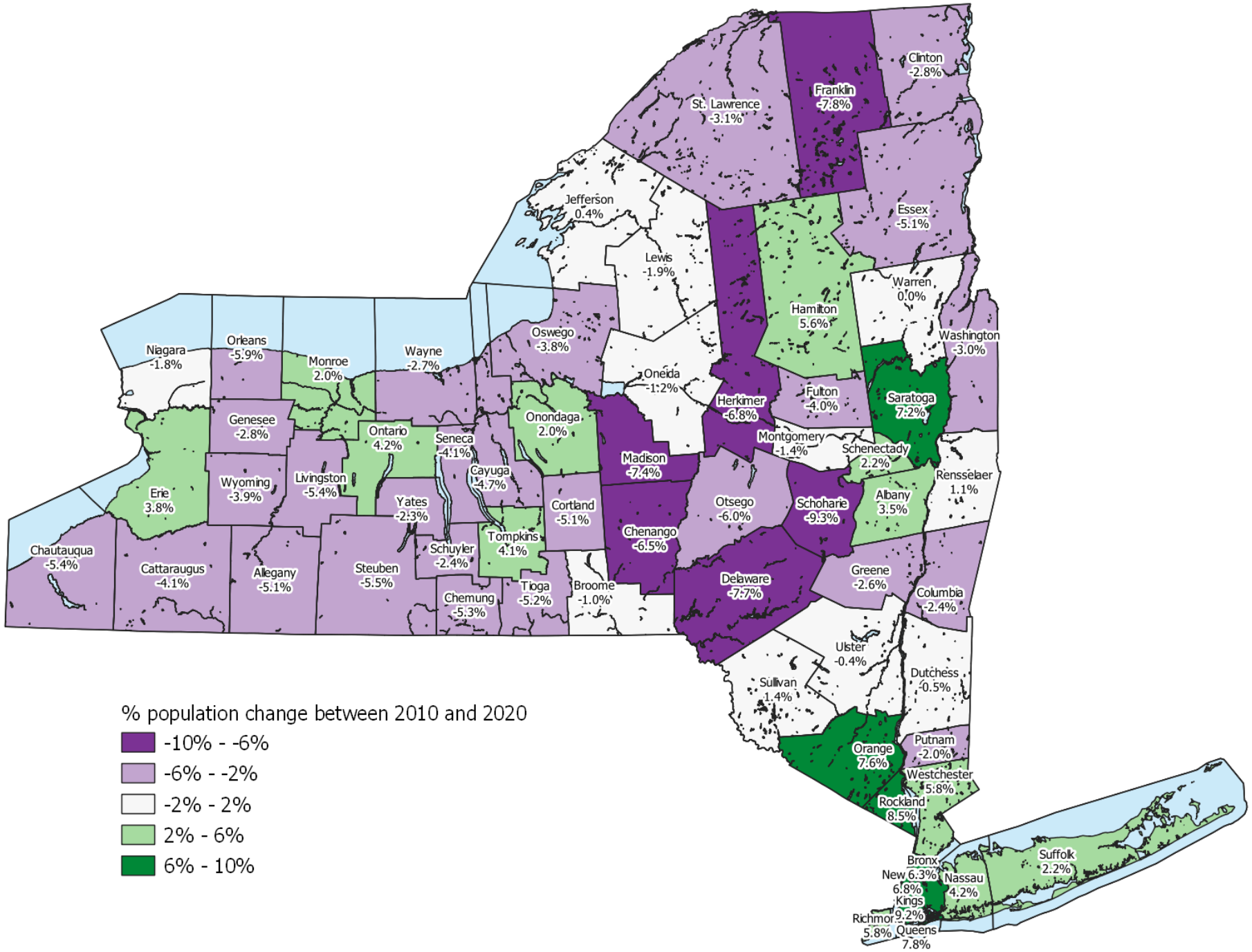
CHANGE IN POPULATION

CHANGE IN HOUSING UNITS

DIVERSITY

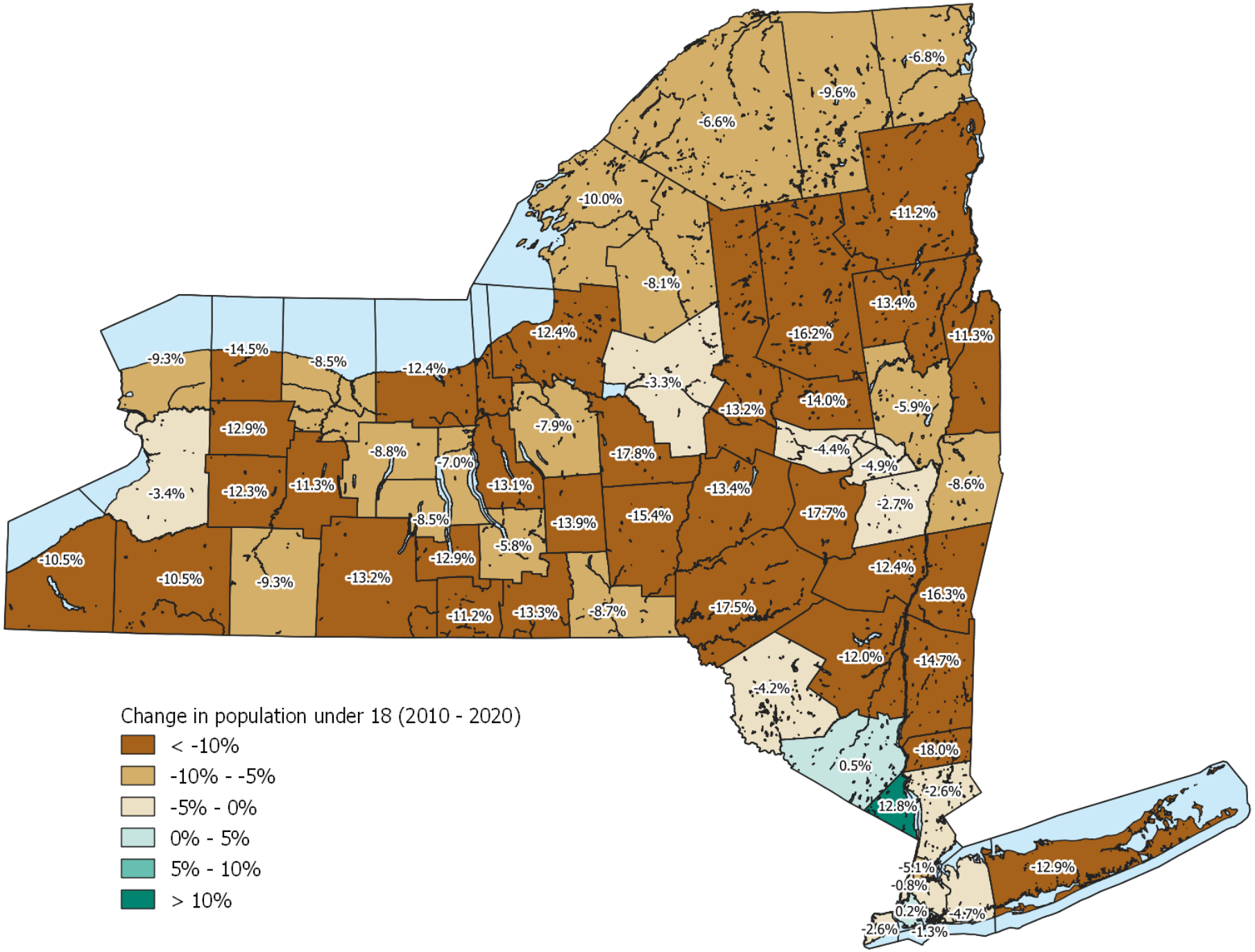
Change in population

	Total population		2010-2020 Change	
	2010	2020	Count	%
Capital Region	1,079,207	1,106,088	26,881	2.5%
Central New York	791,939	785,114	-6,825	-0.9%
Finger Lakes	1,217,156	1,222,868	5,712	0.5%
Long Island	2,832,882	2,921,694	88,812	3.1%
Mid-Hudson	2,290,851	2,398,150	107,299	4.7%
Mohawk Valley	500,155	483,358	-16,797	-3.4%
New York City	8,175,133	8,804,190	629,057	7.7%
North Country	433,193	421,694	-11,499	-2.7%
Southern Tier	657,909	640,036	-17,873	-2.7%
Western New York	1,399,677	1,418,057	18,380	1.3%
New York State	**19,378,102**	**20,201,249**	**823,147**	**4.2%**


Clinton
-2.8%
Franklin
-7.8%
St. Lawrence
-3.1%
Essex
-5.1%
Jefferson
0.4%
Lewis
-1.9%
Warren
0.0%
Hamilton
5.6%
Washington
-3.0%
Oswego
-3.8%
Niagara
-1.8%
Orleans
-5.9%
Monroe
2.0%
Wayne
-2.7%
Oneida
-1.2%
Herkimer
-6.8%
Fulton
-4.0%
Saratoga
7.2%
Genesee
-2.8%
Onondaga
2.0%
Ontario
4.2%
Seneca
-4.1%
Montgomery
-1.4%
Schenectady
2.2%
Madison
-7.4%
Cayuga
-4.7%
Wyoming
-3.9%
Livingston
-5.4%
Rensselaer
1.1%
Erie
3.8%
Yates
-2.3%
Otsego
-6.0%
Schoharie
-9.3%
Albany
3.5%
Cortland
-5.1%
Tompkins
4.1%
Chenango
-6.5%
Chautauqua
-5.4%
Schuyler
-2.4%
Greene
-2.6%
Cattaraugus
-4.1%
Allegany
-5.1%
Steuben
-5.5%
Broome
-1.0%
Delaware
-7.7%
Columbia
-2.4%
Tioga
-5.2%
Chemung
-5.3%
Ulster
-0.4%
Dutchess
0.5%
Sullivan
1.4%
Putnam
-2.0%
Orange
7.6%
Westchester
5.8%
Rockland
8.5%
Bronx
6.3%
New
6.8%
Nassau
4.2%
Suffolk
2.2%
Kings
9.2%
Richmor
5.8%
Queens
7.8%
% population change between 2010 and 2020
-10% - -6%
-6% - -2%
-2% - 2%
2% - 6%
6% - 10%

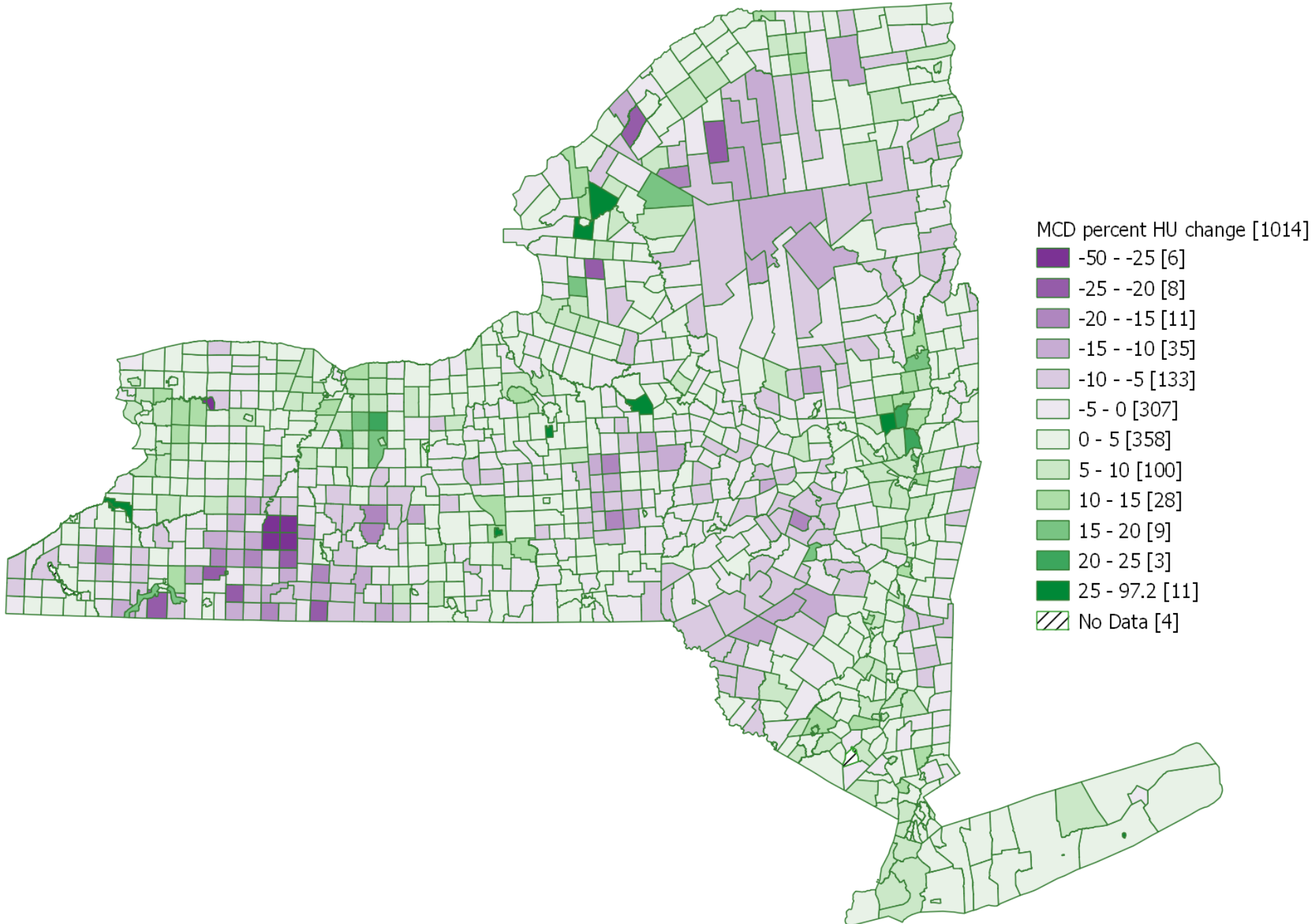
Population under 18

	Population under 18		2010-2020 Change	
	2010	2020	Count	%
Capital Region	229,116	213,267	-15,849	-6.9%
Central New York	179,192	160,609	-18,583	-10.4%
Finger Lakes	273,752	247,750	-26,002	-9.5%
Long Island	669,250	608,398	-60,852	-9.1%
Mid-Hudson	561,120	547,753	-13,367	-2.4%
Mohawk Valley	107,952	99,406	-8,546	-7.9%
New York City	1,768,111	1,740,142	-27,969	-1.6%
North Country	95,132	86,984	-8,148	-8.6%
Southern Tier	137,047	121,770	-15,277	-11.1%
Western New York	304,257	287,035	-17,222	-5.7%
New York State	**4,324,929**	**4,113,114**	**-211,815**	**-4.9%**

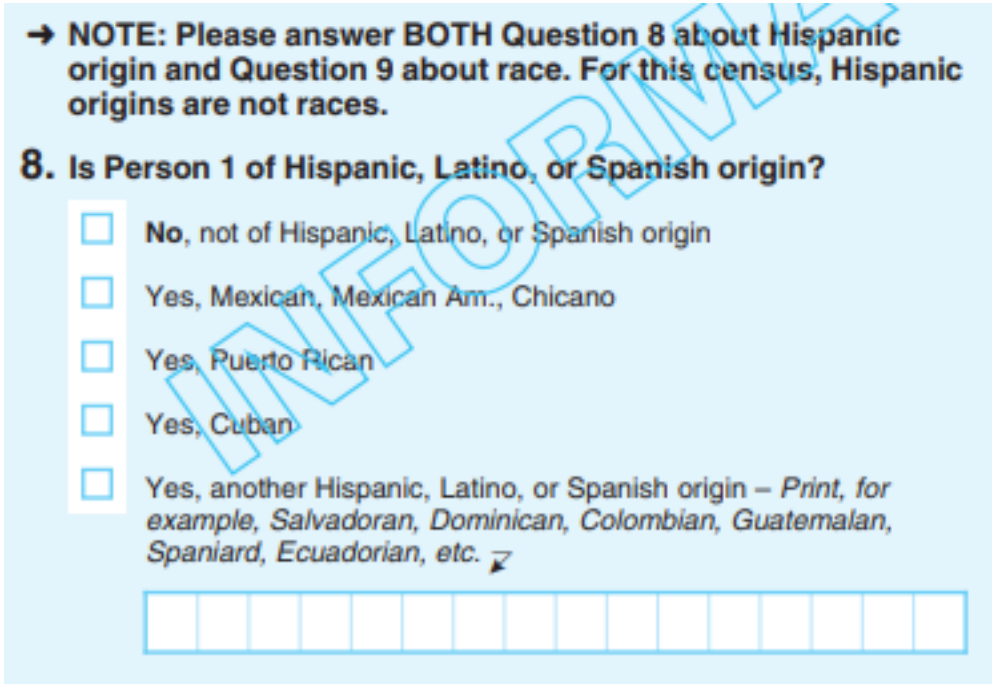

-6.8%
-9.6%
-6.6%
-10.0%
-11.2%
-8.1%
-13.4%
-12.4%
-16.2%
-11.3%
-9.3%
-14.5%
-8.5%
-12.4%
-3.3%
-14.0%
-5.9%
-13.2%
-12.9%
-7.9%
-4.4%
-8.8%
-17.8%
-4.9%
-7.0%
-8.6%
-3.4%
-12.3%
-11.3%
-13.1%
-13.4%
-2.7%
-8.5%
-13.9%
-15.4%
-17.7%
-5.8%
-10.5%
-12.9%
-12.4%
-10.5%
-9.3%
-13.2%
-16.3%
-11.2%
-13.3%
-8.7%
-17.5%
-12.0%
-14.7%
-4.2%
-18.0%
0.5%
-2.6%
12.8%
-12.9%
-5.1%
-0.8%
0.2%
-4.7%
-2.6%
-1.3%
Change in population under 18 (2010 - 2020)
< -10%
-10% - -5%
-5% - 0%
0% - 5%
5% - 10%
> 10%

Housing Units

	Housing units		2010-2020 Change	
	2010	2020	Count	%
Capital Region	505,621	532,119	26,498	5.2%
Central New York	344,778	353,838	9,060	2.6%
Finger Lakes	528,490	549,758	21,268	4.0%
Long Island	1,038,331	1,055,672	17,341	1.7%
Mid-Hudson	901,589	939,205	37,616	4.2%
Mohawk Valley	237,194	234,485	-2,709	-1.1%
New York City	3,371,062	3,618,635	247,573	7.3%
North Country	220,702	224,262	3,560	1.6%
Southern Tier	307,071	311,336	4,265	1.4%
Western New York	653,265	668,756	15,491	2.4%
New York State	**8,108,103**	**8,488,066**	**379,963**	**4.7%**


MCD percent HU change [1014]
-50 - -25 [6]
-25 - -20 [8]
-20 - -15 [11]
-15 - -10 [35]
-10 - -5 [133]
-5 - 0 [307]
0 - 5 [358]
5 - 10 [100]
10 - 15 [28]
15 - 20 [9]
20 - 25 [3]
25 - 97.2 [11]
No Data [4]

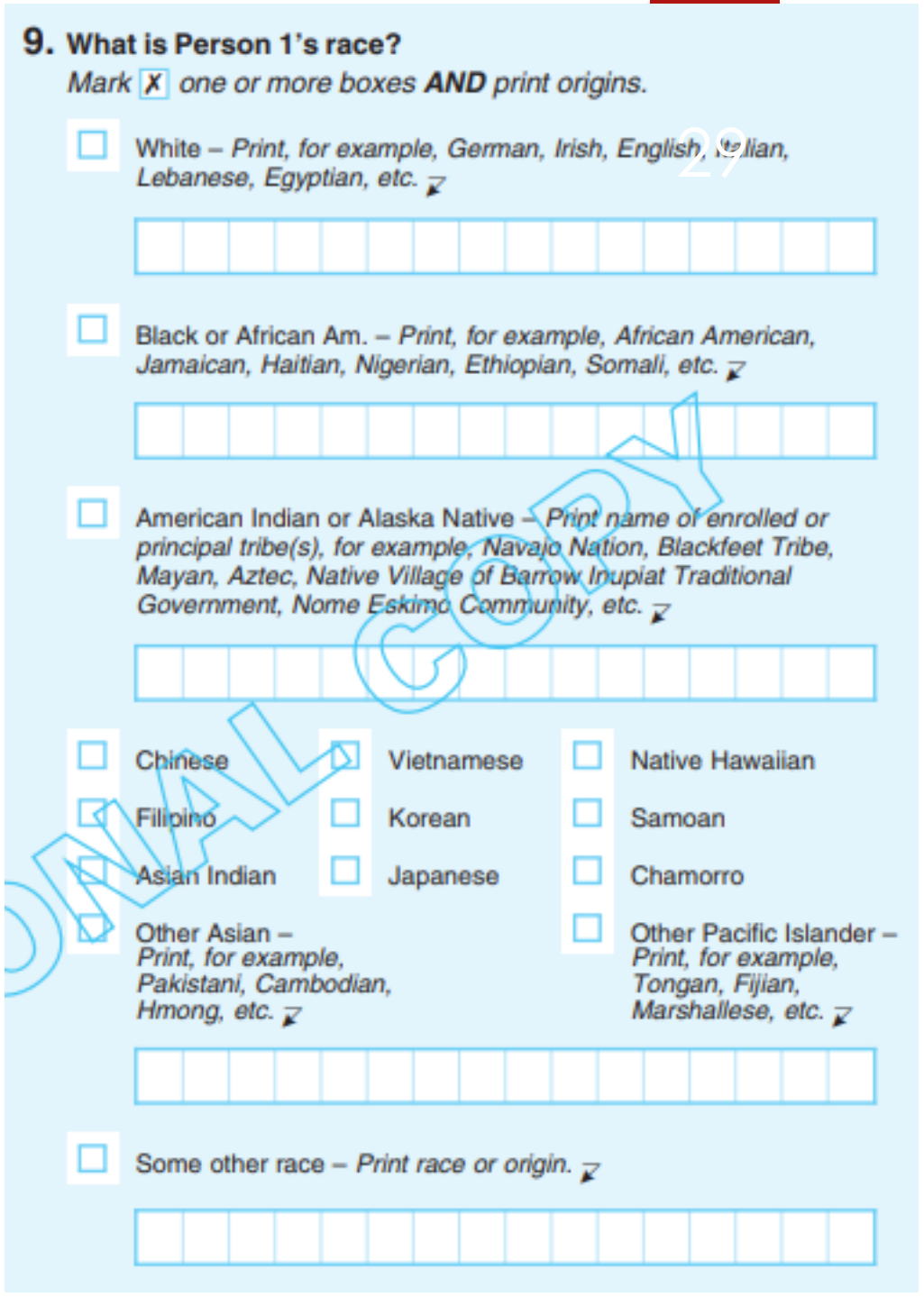
Race/ethnicity questions



➜ **NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.**

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- ☐ **No**, not of Hispanic, Latino, or Spanish origin
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, another Hispanic, Latino, or Spanish origin – *Print, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian, etc.*



9. What is Person 1's race?

Mark ☒ one or more boxes ***AND*** *print origins.*

- ☐ White – *Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.*
- ☐ Black or African Am. – *Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.*
- ☐ American Indian or Alaska Native – *Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.*
- ☐ Chinese
- ☐ Filipino
- ☐ Asian Indian
- ☐ Vietnamese
- ☐ Korean
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Chamorro
- ☐ Other Asian – *Print, for example, Pakistani, Cambodian, Hmong, etc.*
- ☐ Other Pacific Islander – *Print, for example, Tongan, Fijian, Marshallese, etc.*
- ☐ Some other race – *Print race or origin.*

Race/ethnicity processing

Office of Management and Budget sets race standards

- 1997 race standards based on origin, e.g. persons from Western Europe are classified as White
- Detailed list of origins is used to process Census responses

If race based on origin and self-identified race do not correspond, response gets race based on origin added to the response

- E.g. somebody checking only black and answering Canadian as origin will have white added as a race response and will be tallied as multi-racial (black and white)

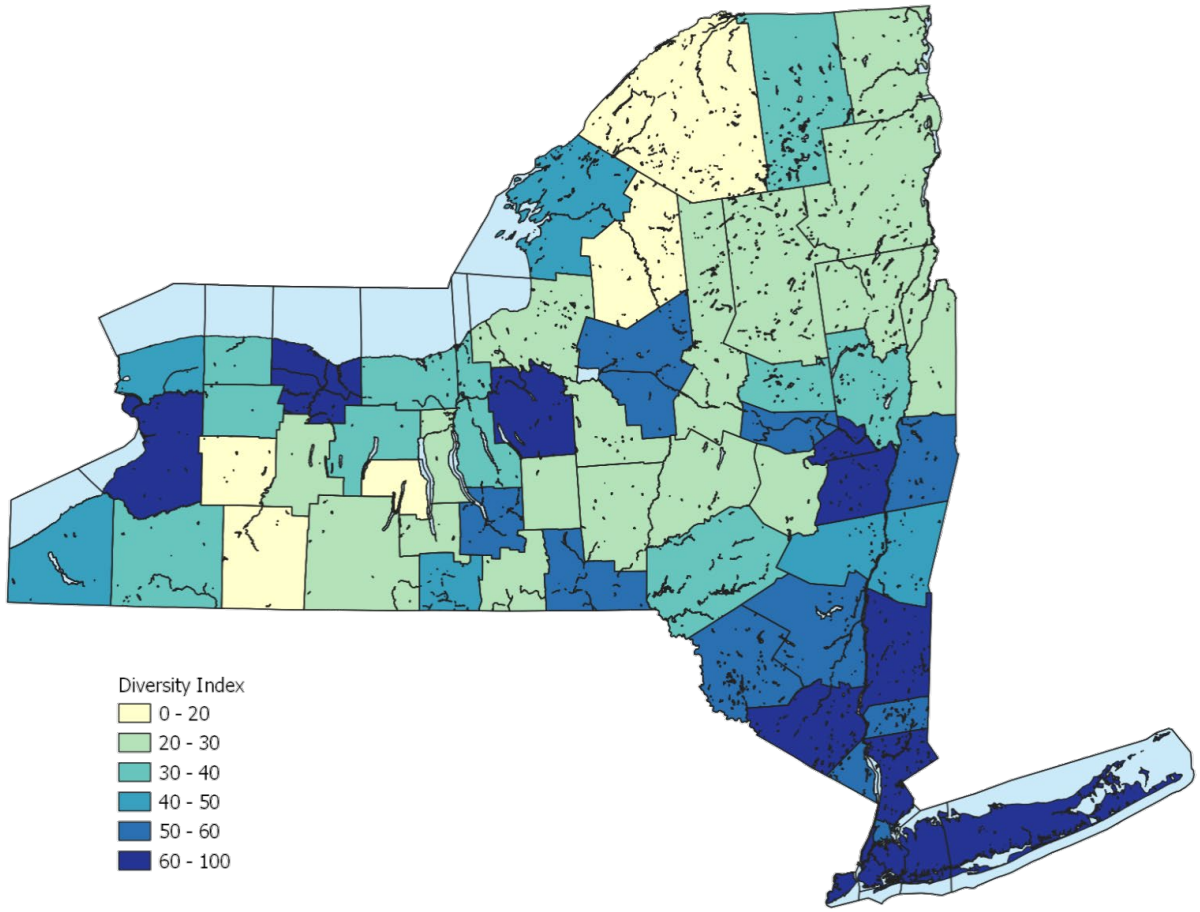
Changing diversity

Comparison over time complicated because:

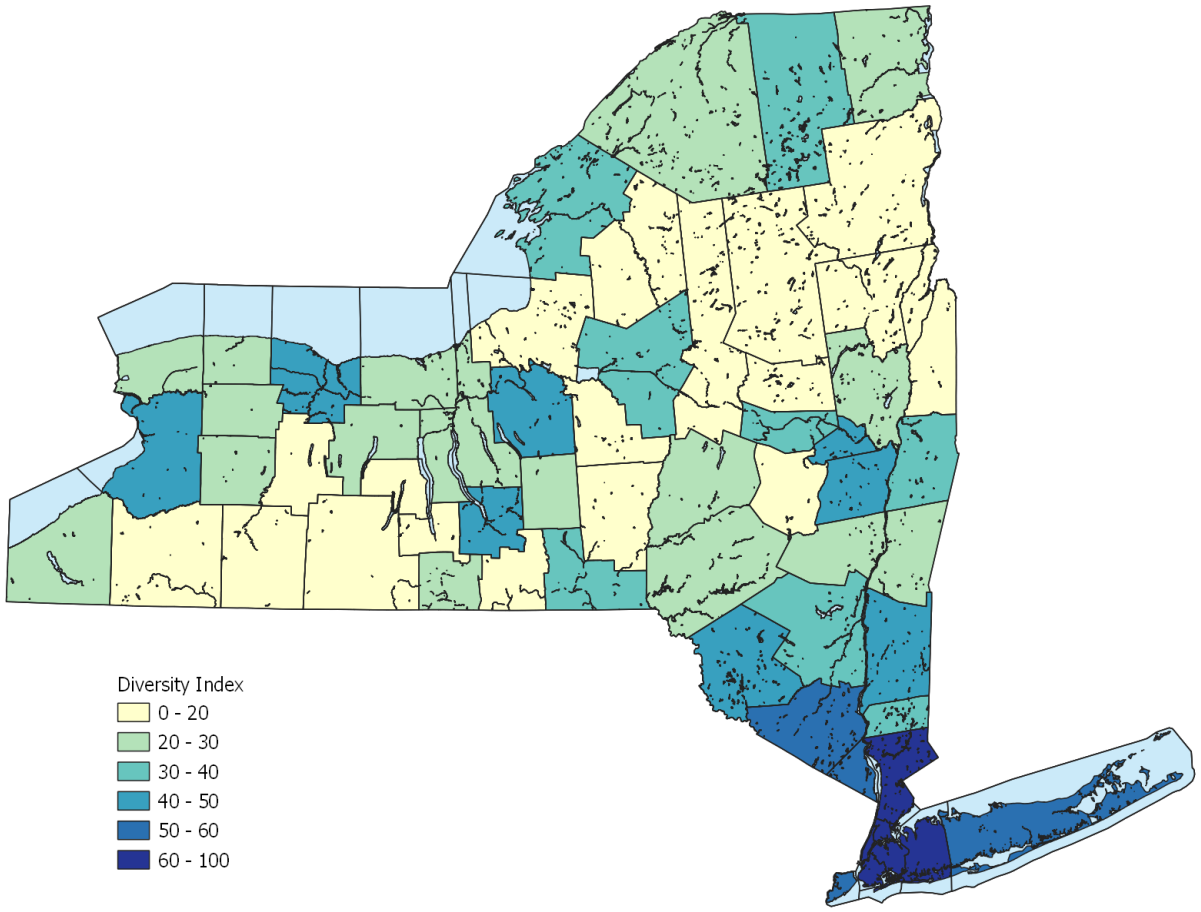
- Coverage differences
- Format of the question changed
- Processing of the open-ended answers changed
- Personal attitudes and race identity changed

	2020	2010	Change
Some other race	2,210,633 (11%)	1,441,563 (7%)	+53%
Two or more races	1,767,463 (9%)	585,849 (3%)	+202%

Diversity Index



Population under 18



Population 18 and over

Estimates and projections

Cohort component method

$$Pop_{t+1} = Pop_t + Births\ - Deaths + Migr_{in} - Migr_{out}$$
$$Age_{t+1} = Age_t + 1$$

Components of change:

- In **estimates** based on Vital Statistics, IRS, Medicare, ACS, etc.
- In **projections** based on assumptions, most often recent observations kept constant

Estimates

Census Bureau produces annual estimates

- County level and above
  - Components of change
  - Total population
  - By age, sex, race, Hispanic origin
- Sub county (cities, towns, villages)
  - Total population

Based on last Census (including corrections) and estimates of change added

- Estimates this year based on ‘blended base’

Projections

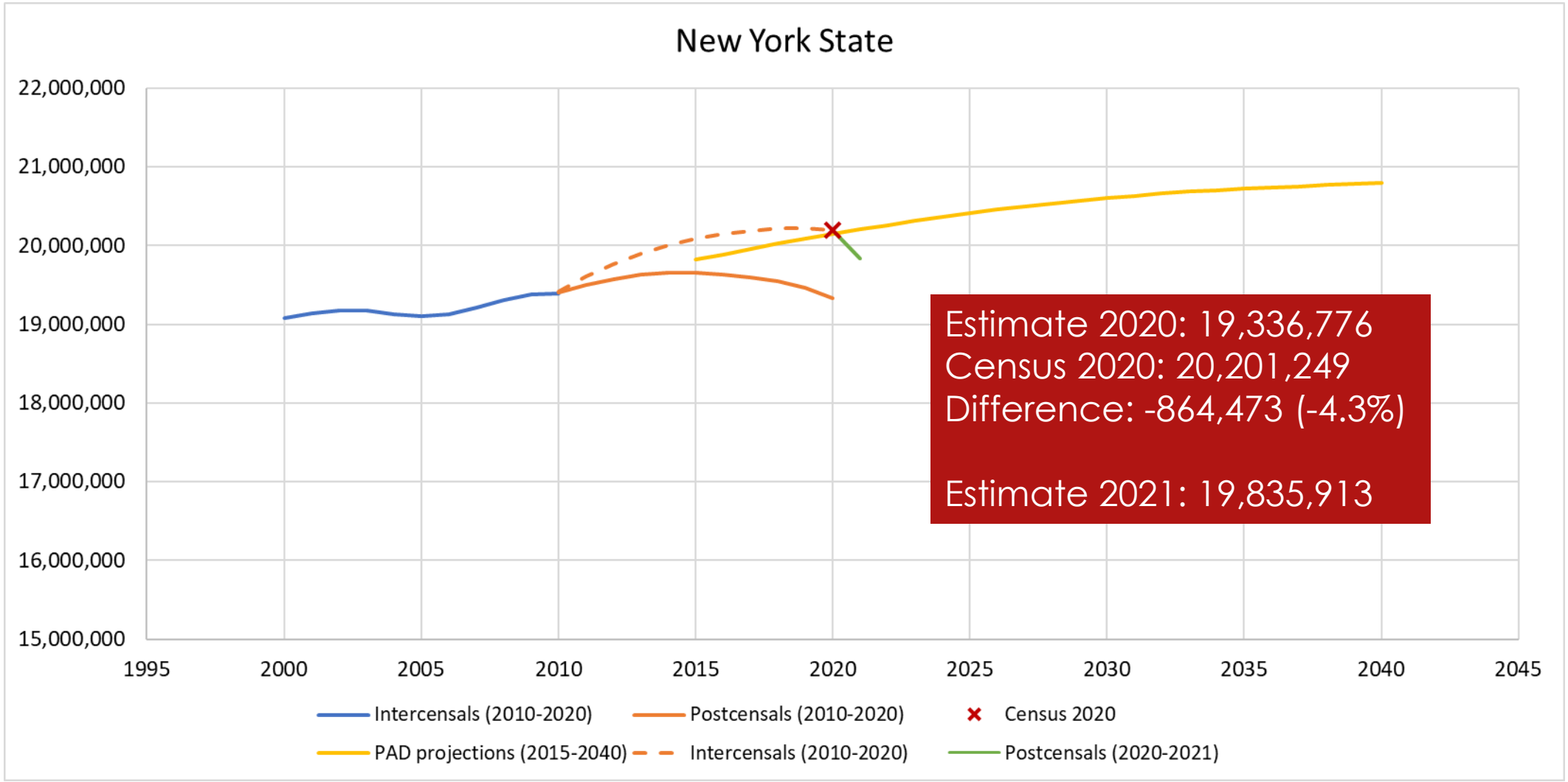
Cornell Program on Applied Demographics produces County projections (most recent produced in 2018)

- Age, sex
- 2015 – 2040

Plans:

- 2020 – 2050 Preliminary state projections produced Fall 2022
- 2020 – 2050 County projections released Summer/Fall 2023

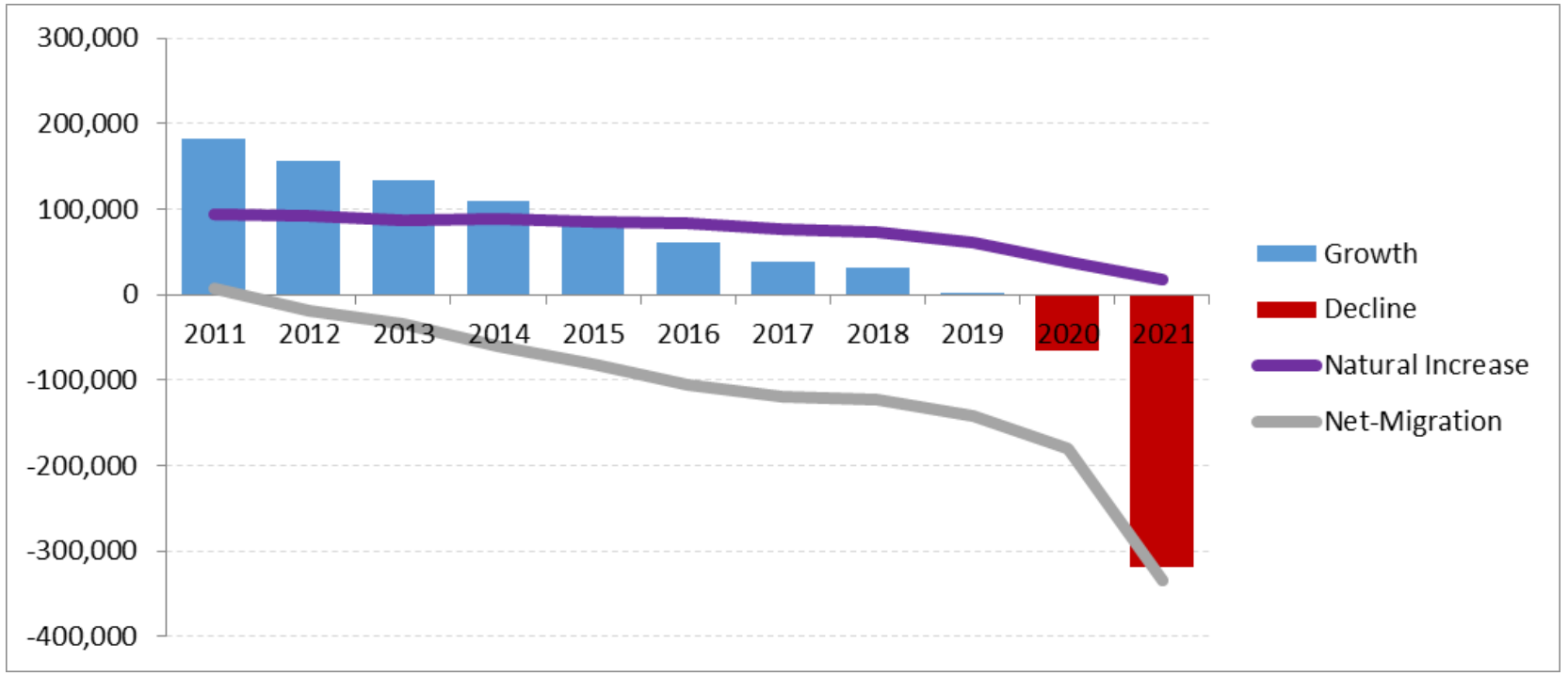
Estimates & projections

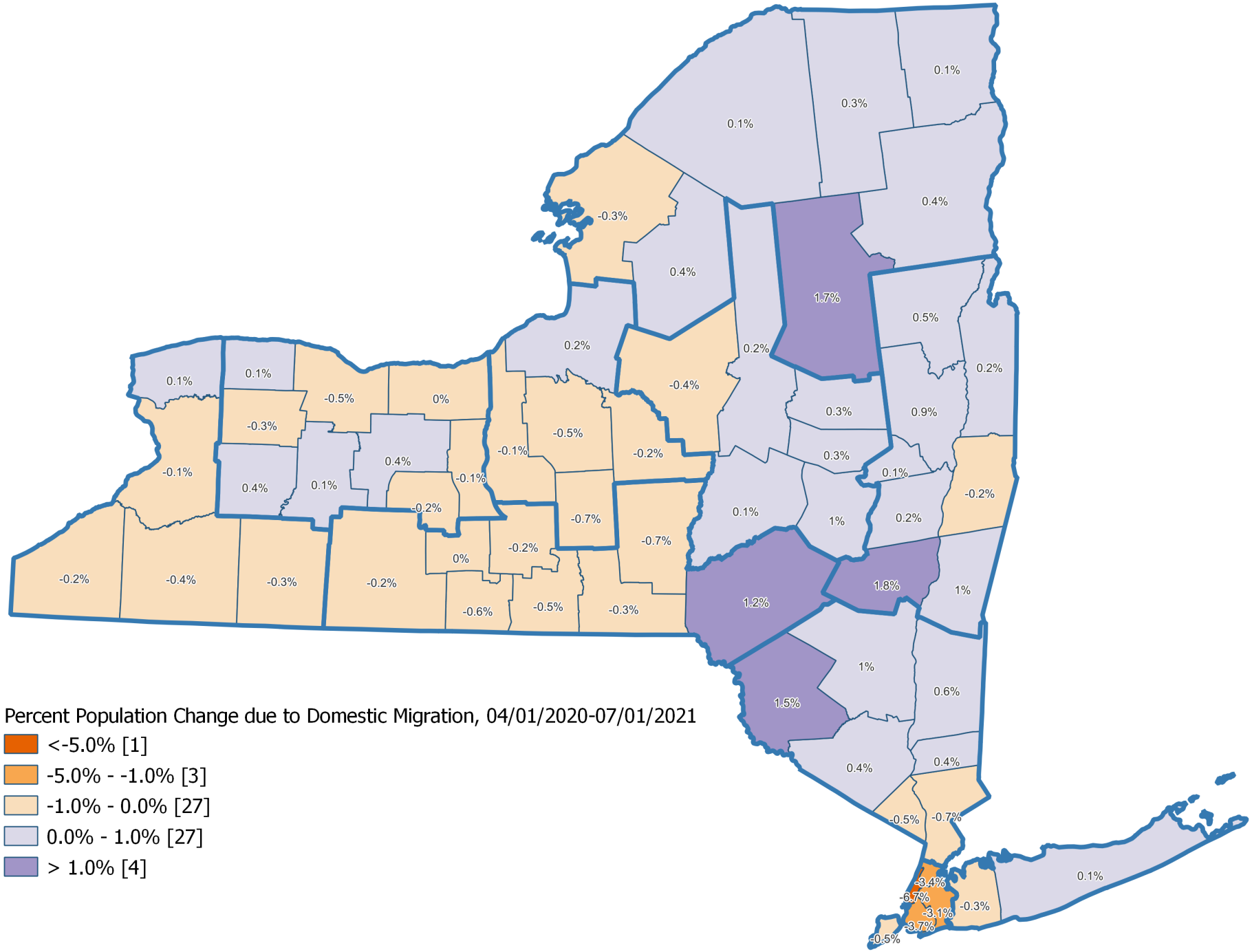


Estimated change since Census day

			Change between Census 2020 and 2021					
			Difference		Due to Natural Increase		Due to Net-Migration	
	Census 2020	Estimate 2021	Count	%	Count	Rate	Count	Rate
New York State	**20,201,249**	**19,835,913**	**-365,336**	**-1.8%**	**25,796**	**0.1%**	**-387,397**	**-1.9%**
Capital Region	1,106,088	1,106,274	186	0.0%	-3,049	-0.3%	3,145	0.3%
Central New York	785,114	780,472	-4,642	-0.6%	-1,169	-0.1%	-3,570	-0.5%
Finger Lakes	1,222,868	1,217,005	-5,863	-0.5%	-2,363	-0.2%	-3,637	-0.3%
Long Island	2,921,694	2,917,251	-4,443	-0.2%	-872	-0.0%	-4,011	-0.1%
Mid-Hudson	2,398,150	2,399,452	1,302	0.1%	4,650	0.2%	-3,566	-0.1%
Mohawk Valley	483,358	480,871	-2,487	-0.5%	-1,903	-0.4%	-612	-0.1%
New York City	8,804,190	8,467,513	-336,677	-3.8%	38,564	0.4%	-370,153	-4.2%
North Country	421,694	420,358	-1,336	-0.3%	-667	-0.2%	-727	-0.2%
Southern Tier	640,036	635,042	-4,994	-0.8%	-2,633	-0.4%	-2,422	-0.4%
Western New York	1,418,057	1,411,675	-6,382	-0.5%	-4,762	-0.3%	-1,844	-0.1%

New York State – estimated annual change




0.1%
0.3%
0.1%
0.1%
-0.3%
0.4%
0.4%
1.7%
0.5%
0.2%
0.2%
-0.4%
0.2%
0.1%
0.1%
-0.5%
0%
-0.3%
0.3%
0.9%
0.4%
-0.5%
-0.1%
-0.2%
0.3%
-0.1%
0.4%
0.1%
-0.1%
0.1%
-0.2%
-0.2%
1%
0.2%
-0.7%
-0.7%
0%
-0.2%
1%
1.8%
-0.2%
-0.4%
-0.3%
-0.2%
1.2%
-0.6%
-0.5%
-0.3%
1%
1.5%
0.6%
0.4%
0.4%
-0.5%
-0.7%
-3.4%
-6.7%
-3.1%
-0.3%
-3.7%
-0.5%
0.1%
Percent Population Change due to Domestic Migration, 04/01/2020-07/01/2021
<-5.0% [1]
-5.0% - -1.0% [3]
-1.0% - 0.0% [27]
0.0% - 1.0% [27]
> 1.0% [4]

American Community Survey

Ongoing national survey

Sample size: around 2 million households annually + 150K Group Quarters

Data collected on a wide variety of topics

- Demographic
- Economic
- Social
- Housing

Reported annually: 1 yr ACS for areas 65K+, 5 yr ACS for all areas

ACS

The 2020 1-year ACS only published as an experimental product with only state estimates for a limited set of tables

Responses weighted using population estimates as controls

- For the 2020 1-year ACS and 5-year ACS the Vintage 2020 population estimates were used as control
  - This means Census results were not taken into account
    - New York State totals are much lower than Census counts
- For 2021ACS products the Census Bureau was planning to use an intercensal estimate
  - A full intercensal estimate can only produced after DHC publication
    - Provisional intercensal?

Resources

Background

- Census 2020
  - https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-main.html
- Population estimates
  - https://www.census.gov/programs-surveys/popest.html
- ACS
  - https://www.census.gov/programs-surveys/acs

Data

- Census Bureau: https://data.census.gov
- Cornell PAD: https://pad.human.cornell.edu/

Questions?

Jan Vink
Email: jkv3@cornell.edu

Cornell PAD
Email: padinfo@cornell.edu
Twitter: @PADCornell