Carl Vinson
Institute of Government
UNIVERSITY OF GEORGIA

# Methodology

# 2024 Long-Term Population Projections

Released: July 2024

# Introduction

The Applied Demography Program at the University of Georgia's Carl Vinson Institute of Government produced resident population projections for Georgia's counties in single-year intervals for the period 2024 through 2060, a series of projections referred to as the 2024 Long-Term Population Projections. This report describes the methodology, assumptions, and data sources used to produce these population projections. The Institute of Government used a cohort-component projection technique to project resident population by age, race and ethnicity, and sex for the state of Georgia and its counties. Demographers and data scientists relied on publicly available data from state and federal agencies, regional expertise, and current research on US population and demographic trends to produce these population projections.

These projections, like all projections, involve the use of certain assumptions about future events that may or may not occur. Users should be aware that although these projections were prepared using standard methodologies and extensive attempts were made to account for existing demographic patterns, they may not accurately project the future population of the state or any specific county. The projections are based on historical trends and current estimates. These projections should be used with full awareness of the inherent limitations of population projections and with knowledge of the procedures and assumptions described in this report.

Population projections serve as a resource for local and state decision-makers as these leaders plan for the future. Just as exogenous factors such as federal policy can influence international migration, and global challenges such as a worldwide pandemic can influence all three components of population change, decisions made today by Georgia's leaders can influence the population trajectory. The Applied Demography Program at the Carl Vinson Institute of Government regularly refines model assumptions and parameters to incorporate the latest data and research into these projections.

# Projection Methodology

## OVERVIEW

The projections in this report were produced using a cohort-component projection method. The institute used the Census Bureau's Vintage 2023 Population Estimates as the base population. The cohort populations included single year of age, four race/ethnic groups, and two sex groups for each of Georgia's 159 counties. The components of population change include fertility, mortality, and migration. For each year in the projection period, the survived population is advanced one year of age using projected age-specific survival rates and net migration rates.

## COHORT-COMPONENT PROJECTION METHOD

In the cohort-component method, separate rates for the components of population change (fertility, mortality, and migration) are projected for each population cohort (age, race and ethnicity, and sex).

Cohorts refer to persons with one or more common characteristics, in this case: age, race or ethnicity, and sex. Components refer to the components of population change: births, deaths, and migration. These projections of components for each cohort are then combined as follows:

$$P_{t2} = P_{t1} + B_{t1-t2} - D_{t1-t2} + M_{t1-t2,}$$

where $P_{t2}$ = the population projected at future date $t_1 - t_2$,

$P_{t1}$ = the population at the base year $t_1$,

$B_{t1-t2}$ = the number of births that occur during the interval $t_1 - t_2$,

$D_{t1-t2}$ = the number of deaths that occur during the interval $t_1 - t_2$, and

$M_{t1-t2}$ = the net migration that takes place during the interval $t_1 - t_2$.

Appropriately adapting the cohort-component technique involves four steps:

1. Select a baseline set of cohorts for the projection area or areas of interest for the base time period.
2. Determine migration, survival, and fertility measures for each cohort for the baseline time period.
3. Determine a method for projecting trends in fertility, survival, and migration rates over the projection period.
4. Apply the rates to the baseline population and component trends for the projection period.

## POPULATION DATA AND BASELINE RATES

### Base Population

The base population is the estimate used as the launch point in the projection model. The Census Bureau's Vintage 2023 population estimates for Georgia's counties with characteristics of age, race and ethnicity, and sex were used as the baseline population for the 2024 Long-Term Population Projections. Institute faculty and staff modified the 2023 population characteristics; specifically, the population was grouped into single-year age groups (0 to 84, and age 85 and over), two sex groups (male and female), and four racial and ethnic groups (non-Hispanic White, non-Hispanic Black or African American, non-Hispanic Other, and Hispanic).

Faculty and staff at the institute developed 2010-2020 intercensal estimates. Intercensal estimates are a set of population estimates which reconcile the existing postcensal population estimates and the subsequent decennial census counts. The Census Bureau has yet to release official intercensal estimates. Without the intercensal adjustments, postcensal population estimates do not align with any subsequent censuses or any future postcensal estimates, creating an artificial gap between vintages of postcensal estimates. This gap undermines the quality of any time-series population research or modeling that bridges a decade. Pending the release of official intercensal estimates, we produced independent intercensal estimates using the Das Gupta error closure methods, the same set of methods used by the Census Bureau. After selecting one of six Das Gupta methods, calculating new intercensal estimates, and applying the age, sex, and race distribution from the Vintage 2019 postcensal estimates to our new estimates, we were able to construct a set of independent set of adjusted population estimates for 2010-2019 broken down by single-year age, race, and sex.

### Fertility Data and Rates

Birth data were collected from the Georgia Department of Public Health's (DPH) Office of Health Indicators and Planning (OHIP) Online Analytical Statistics Information System (OASIS)[i]. The data include the age (ages 10-55), race, sex, birth order event, and county of residence of the mother from 2000 to 2023. Institute faculty and staff modified the race and ethnicity characteristics into the same four groups used in the base population. Historical (2000-2023) age-specific fertility rates (ASFR) for each birth order event across the four racial and ethnic groups were calculated using the internal Intercensal Estimates developed by the Institute of Government.

Age-specific fertility rates (ASFR) for each birth order event across four racial and ethnic groups were projected for each year in the projection period using Gaussian Process Regressions. These rates were projected at the state level and applied the base population in each county.

### Mortality Data and Rates

Death data were collected from the Georgia Department of Public Health's (DPH) Office of Health Indicators and Planning (OHIP) Online Analytical Statistics Information System (OASIS)[i]. The data include single year of age (0-100, and 100 and over), race and ethnicity, sex, and primary county of residence for years 2000 to 2023. Institute faculty and staff modified the race and ethnicity characteristics of the mortality data into the same four groups used in the

base population. Historical (2000-2023) age, race and ethnicity, and sex specific mortality rates were calculated using the internal Intercensal Estimates developed by the institute.

Age, race and ethnicity, and sex-specific Mortality Rates were projected for each year in the projection period using Gaussian Process Regressions. These rates were projected at the state level and applied the base population in each county.

After evaluating multiple scenarios, institute researchers determined removing mortality data from 2020, 2021, and 2022, due to the elevated number of deaths during those years, produced a more stable projection of long-term mortality rates. These data were removed from the final mortality rate projection model.

### Migration Data and Rates

Institute of Government faculty and staff calculated residual net migration rates and schedules from the 2000-2023 population estimates to produce an age, race and ethnicity, and sex structure at the county level and at the Georgia Department of Community Affairs (DCA) regional level. Historical (2000-2023) age, race and ethnicity, and sex-specific migration rates were calculated using the internal Intercensal Estimates developed by the Institute of Government.

Age, race and ethnicity, and sex-specific migration rates were projected for each year in the projection period using Gaussian Process Regressions. Rates were calculated at the county level and region level. County and region migration rates were evaluated for each county and cohort combination.

A review panel of Institute of Government experts evaluated county and region migration rates for each county cohort combination. In some cases, county or migration rates were adapted from other jurisdictions outside of the DCA region as appropriate based on comparable geographic and demographic characteristics.

### Group Quarters Data

A “group quarters” is a place where people live or stay in a group living arrangement that is owned or managed by an organization providing housing and/or services for the residents. Examples include nursing homes, university housing, and prisons. The 2024 projections incorporate stochastic simulations for the group quarters population.

The base group quarters population file provides estimates of the total group quarters population by county for 2020, broken down by 5-year age groups, race, and sex, for use as a reference in projecting future group quarters populations. This population file was created from two component census tables, both descending from the 2020 Census Demographic and Housing Characteristics File (DHC):

1. A combined version of tables PCT18A - PCT18H: Group Quarters Population by Sex by Age by Major Group Quarters Type (All Races)

2. Table PCO1: Group Quarters Population by Sex by Age

Table 1 was used as the base table, as it had a total estimate closer to the official Census Bureau total. Each 5-year age group in Table 2 was classified into one of the three larger age groups from Table 1. Then, in Table 1, the total population by sex was calculated for each of the three larger age groups. The individual race estimates were then divided by the sex-age group totals to find the proportion of each age group accounted for by each race group.

The tables were then combined based on county, sex, and the three larger age groups from Table 1, and new estimates were calculated for each five-year age group using the race proportions calculated earlier, rounding down to ensure only whole persons were counted. This process is summarized by the equation below. It occurred four times for each sex and five-year age group combination and once for each race group.

*New Estimate = Old Estimate * (Race total for large age group/Total for large age group)*

## PROJECTION APPROACH AND MODEL

The Institute of Government's projection method relies on a bottom-up approach. Institute of Government faculty and staff produce a state projection by reviewing the projections by age, race and ethnicity, and sex of all of the 159 counties in Georgia. This review and analysis incorporate faculty and staff expertise on local trends and economic development in Georgia. The state-level projection is cross-validated with other sources, including independent projections, neighboring-state projections, and simple autoregressive integrated moving-average methods.

### Projecting Rates and Gaussian Process Regression

Institute data scientists projected fertility, mortality, and migration rates using Gaussian process regression (GPR). GPRs are machine learning algorithms used for regression (i.e., predicting fertility rates and housing prices) and classification problems (i.e., identifying in which categories an observation belongs).

GPRs are flexible non-parametric machine learning models commonly used for modeling spatial and time series data. Deploying GPRs to project population component rates enables data scientists and demographers to consider multiple model specifications that account for long-term (historical) trends and adjusting for seasonality or recent trends.

[i] oasis.state.ga.us