Summary

- Multiple data sources
    - Individual-level data weighted to obtain population-level estimates for a given year
    - Population-level measures (life expectancy variables)
- State-level and county-level data
- Race-specific (non-white, white) and sex-specific (female, male) measures
    - e.g., percent of homeowners among non-white population, percent of homeowners among white population
    - e.g., percent of males with a high school diploma, percent of females with a high school diploma
- Additional measures: ratios of race-specific and sex-specific measures
    - e.g., ratio of white homeowner rates to non-white homeowner rates
        - Ratio of 1 means that whites and nonwhites have the same rates of having a high school diploma in the specified unit of analysis
        - Ratios above 1 indicate higher rates among whites; ratios below 1 indicate higher rates among non-whites (e.g., 42% of whites employed and 47.6% of non-whites employed in Pennsylvania in 1994 = a ratio of 0.88).
    - e.g., ratio of male high school diploma rates to female high school diploma rates
        - Ratio of 1 means that males and females have the same rates of having a high school diploma in the specified unit of analysis.
        - Ratios above 1 indicate higher rates among males; ratios below 1 indicate higher rates among females (e.g., 31.2% of males and 32.7% of females who have at least attained a bachelor's degree in Connecticut in 2005 = a ratio of 0.95).

Race Measures

Source Variable Universes

1. IPUMS CPS Annual Social and Economic Supplement Samples
   - EDUC, EMPSTAT, LABFORCE, OFFPOV, OWNERSHP, HHINCOME, SPMNCHILD, SPMPOV, OCC1950, RACE, HISPAN, STATEFIP, COUNTY, YEAR, ASECWT
2. IPUMS USA (Harmonized U.S. Census Data)
   a. LIT, RACE, HISPAN, STATEICP, COUNTYICP, PERWT
3. IPUMS National Historic Geographic Information Systems
   a. DISSIM
4. CDC National Vital Statistics Systems
   a. E0, E1
   b. *Infant mortality variable

Coding scheme for race

Data: IPUMS CPS

- *Variables used*: HISPAN, RACE, YEAR, ASECWT
  - Individuals were considered **white** if their original values were:
    - RACE = 100
    - HISPAN = 0
  - Individuals were considered **non-white** if their original values were:
    - RACE = 200-900
    - HISPAN = 100-612

Data: IPUMS USA

- *Variables used*: HISPAN, RACE, YEAR, ASECWT
  - Individuals were considered **white** if their original values were:
    - RACE = 1
    - HISPAN = 0
  - Individuals were considered **non-white** if their original values were:
    - RACE = 2-6
    - HISPAN = 2-4

Measures

st_perc_white_hsd

st_perc_nonwhite_hsd

cnty_perc_white_hsd

cnty_perc_nonwhite_hsd

Type: Numeric

Description: Percentages of white and non-white individuals with at least high school degree in each year and within each locality.

Availability:

- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: EDUC, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **less than high school** if their original value was:
        - EDUC = 2-72 ([None, preschool, or kindergarten]-[12th grade, diploma unclear])
    - Individuals were considered **at least high school degree** if their original value was:
        - EDUC = 73-125 ([High school diploma or equivalent]-[Doctorate degree])

---

st_race_ratio_hsd

cnty_race_ratio_hsd

Type: Numeric

Variable description: The ratio of the percentage of white individuals with at least a high school degree to the percentage of non-white individuals with at least a high school degree in each year and within each locality..

- *Note*: values are coded as –1 if the percentage of non-white individuals with at least high school is 0.

Availability:

- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_hsd*, *st_perc_nonwhite_hsd, cnty_perc_white_hsd,* and *cnty_perc_nonwhite_hsd*.

---

st_perc_white_home

st_perc_nonwhite_home

cnty_perc_white_home

cnty_perc_nonwhite_home

Type: Numeric

Description: Percentages of white and non-white individuals who own homes in each year and within each locality.

Availability:

- State-level values: 1976-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: OWNERSHP, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **homeowners** if their original value was:
        - OWNERSHP = 10 [Owned or being bought]
    - Individuals were considered **non-homeowners** if their original value was:
        - OWNERSHP = 21 [No cash rent]
        - OWNERSHP = 22 [With cash rent]

---

st_race_ratio_home

cnty_race_ratio_home

Type: Numeric

Description: The ratio of the percentage of white individuals who own their homes to the percentage of non-white individuals who own their homes in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals who own their homes is 0.

Availability:

- State-level values: 1976-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_home, st_perc_nonwhite_home, cnty_perc_white_home,* and *cnty_perc_nonwhite_hom*

---

st_perc_white_employed

st_perc_nonwhite_employed

cnty_perc_white_employed

cnty_perc_nonwhite_employed

Type: Numeric

Description: Percentages of white and non-white individuals who are employed in each year and within each locality.

Availability:
- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:
- *Data*: IPUMS CPS
- *Variables used*: EMPSTAT, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **employed** if their original VARIABLE value was:
        - EMPSTAT = 1 [Armed Forces – In Labor Force – Employed]
        - EMPSTAT = 10 [At work]
        - EMPSTAT = 12 [Has job, not at work last week]
    - Individuals were considered **unemployed** if their original VARIABLE value was:
        - EMPSTAT = 20-36 ([Unemployed]-[NILF, retired])

---

st_race_ratio_employed

cnty_race_ratio_employed

Type: Numeric

Description: The ratio of the percentage of white individuals who are employed to the percentage of non-white individuals who are employed in each year and within each locality.
- *Note*: values are coded as –1 if the percentage of non-white individuals who are employed is 0.

Availability:
- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_employed, st_perc_nonwhite_employed, cnty_perc_white_employed,* and *cnty_perc_nonwhite_employed*

---

st_perc_white_lf

st_perc_nonwhite_lf

cnty_perc_white_lf

cnty_perc_nonwhite_lf

Type: Numeric

Description: Percentages of white and non-white individuals who are in the labor force in each year and within each locality.

Availability:

- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: LABFORCE, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **in the labor force** if their original value was:
        - LABFORCE = 2 [Yes, in the labor force]
    - Individuals were considered **not in the labor force** if their original value was:
        - LABFORCE = 1 [No, not in the labor force]

---

st_race_ratio_lf

cnty_race_ratio_lf

Type: Numeric

Description: The ratio of the percentage of white individuals who are employed to the percentage of non-white individuals who are employed in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals who are employed is 0.

Availability:

- State-level values: 1962-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_lf, st_perc_nonwhite_lf, cnty_perc_white_lf,* and *cnty_perc_nonwhite_lf*.

---

st_median_white_hi

st_median_nonwhite_hi

cnty_median_white_hi

cnty_median_nonwhite_hi

Type: Numeric

Description: The median household income of white and non-white individuals in each year and within each locality.

Availability:

- State-level values: 1968-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: HHINCOME, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - *Note*: Individuals with a negative income (-9999997/-1) were recoded as missing.

---

st_race_ratio_hi

cnty_race_ratio_hi

Type: Numeric

Description: The ratio of the median household income of the white population to the median household income of the non-white population in each year and within each locality.

- *Note*: values are coded as –1 if the median income of the non-white population is 0.

Availability:

- State-level values: 1968-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_median_white_hi, st_median_nonwhite_hi, cnty_median_white_hi,* and *cnty_median_nonwhite_hi*.

---

st_perc_white_apl

st_perc_nonwhite_apl

cnty_perc_white_apl

cnty_perc_nonwhite_apl

Type: Numeric

Description: Percentages of white and non-white individuals who are above the poverty line in each year and within each locality.

Availability:

- State-level values: 1969-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: OFFPOV, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **above the poverty line** if their original value was:
        - OFFPOV = 2 [Above Poverty Line]
    - Individuals were considered **below the poverty line** if their original value was:
        - OFFPOV = 1 [Below Poverty Line]

---

st_race_ratio_apl

cnty_race_ratio_apl

Type: Numeric

Description: The ratio of the percentage of white individuals above the poverty line to the percentage of non-white children above the poverty line in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals above the poverty line is 0.

Availability:

- State-level values: 1969-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_apl, st_perc_nonwhite_apl, cnty_perc_white_apl,* and *cnty_perc_nonwhite_apl.*

---

st_perc_white_childapl

st_perc_nonwhite_childapl

cnty_perc_white_childapl

cnty_perc_nonwhite_childapl

Type: Numeric

Description: Percentages of white and non-white children who are above the poverty line in each year and within each locality.

Availability:
- State-level values: 2010-2022
- County-level values: 2010-2022

Source information:
- *Data*: IPUMS CPS
- *Variables used*: SPMNCHILD, SPMPOV, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered **above the poverty line** if their original value was:
        - SPMPOV = 0 [Above poverty]
    - Individuals were considered **below the povety line** if their original value was:
        - SPMPOV = 1 [Below poverty]

---

st_race_ratio_childapl

cnty_race_ratio_childap

Type: Numeric

Description: The ratio of the percentage of white children above the poverty line to the percentage of non-white children above the poverty line in each year and within each locality.
- *Note*: values are coded as –1 if the percentage of non-white children above the poverty line is 0.

Availability:
- State-level values: 2010-2022
- County-level values: 2010-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_childapl, st_perc_nonwhite_childapl, cnty_perc_white_childapl,* and *cnty_perc_nonwhite_childapl.*

---

st_perc_white_occw

st_perc_nonwhite_occw

cnty_perc_white_occw

cnty_perc_nonwhite_occw

Type: Numeric

Description: Percentages of white and nonwhite individuals working white-collar jobs in each year and within each locality.

Availability:

- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: OCC1950, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered to have **white collar jobs** if their original value was:
        - OCC1950=
            - (000-099) [Professional, Technical]
            - (200-290) [Managers, officials, and proprietors]
            - (400-490) [Sales workers]

---

st_race_ratio_occw

cnty_race_ratio_occw

Type: Numeric

Description: The ratio of the percentage of white individuals working white collar jobs to the percentage of non-white individuals working white collar jobs in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals with white collar jobs is 0.

Availability:

- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:

- Ratios calculated using constructed variables *st_perc_white_occw, st_perc_nonwhite_occw, cnty_perc_white_occw,* and *cnty_perc_nonwhite_occw* in each year and within each locality.

---

st_perc_white_occb

st_perc_nonwhite_occb

cnty_perc_white_occb

cnty_perc_nonwhite_occb

Type: Numeric

Description: Percentage of white and nonwhite individuals working blue-collar jobs in each year and within each locality.

Availability:

- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: OCC1950, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered to have **blue collar jobs** if their original value was:
        - OCC1950=
            - (100-123) - Farmers
            - (500-595) - Craftsmen
            - (600-690) - Operatives
            - (810-840) - Farm laborers
            - (910-970) - Laborers

---

st_race_ratio_occb

cnty_race_ratio_occb

Type: Numeric

Description: The ratio of the percentage of white individuals working blue collar jobs to the percentage of non-white individuals working blue collar jobs in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals with blue collar jobs is 0.

Availability:
- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:
- Ratios calculated using constructed variables *st_perc_white_occb, st_perc_nonwhite_occb, cnty_perc_white_occb,* and *cnty_perc_nonwhite_occb.*

---

st_perc_white_occp

st_perc_nonwhite_occp

cnty_perc_white_occp

cnty_perc_nonwhite_occp

Type: Numeric

Description: Percentages of white and nonwhite individuals working pink-collar jobs in each year and within each locality.

Availability:
- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:

- *Data*: IPUMS CPS
- *Variables used*: OCC1950, RACE, STATEFIP, COUNTY, YEAR, ASECWT
    - Individuals were considered to have **pink collar jobs** if their original value was:
        - OCC1950=
            - (300-390) - Clerical and kindred
            - (700-720) - Service workers (private household)
            - (730-790) - Service workers (not household)

---

st_race_ratio_occp

cnty_race_ratio_occp

Type: Numeric

Description: The ratio of the percentage of white individuals working pink collar jobs to the percentage of non-white individuals working pink collar jobs in each year and within each locality.

- *Note*: values are coded as –1 if the percentage of non-white individuals with pink collar jobs is 0.

Availability:
- State-level values: 1996-2022
- County-level values: 1996-2022

Source information:
- Ratios calculated using constructed variables *st_perc_white_occp, st_perc_nonwhite_occp, cnty_perc_white_occp,* and *cnty_perc_nonwhite_occp* in each year and within each locality.

---

st_perc_white_lit

st_perc_nonwhite_lit

cnty_perc_white_lit

cnty_perc_nonwhite_lit

Type: Numeric

Description: Percentages of white and nonwhite individuals who are literate in each year and within each locality.

Availability:
- State-level values: 1850-1930
- County-level values: 1850-1930

Source information:
- *Data*: IPUMS USA
- *Variables used*: LIT, RACE, STATEICP, COUNTYICP, YEAR, PERWT
    - Individuals were considered to be **literate** if their original value was:
        - LIT=
            - (4) - Yes, literate (reads and writes)

---

st_race_ratio_lit

cnty_race_ratio_lit

Type: Numeric

Description: The ratio of the percentage of white individuals who are literate to the percentage of non-white individuals who are literate in each year and within each locality.

Availability:

- State-level values: 1850-1930
- County-level values: 1850-1930

Source information:

- Ratios calculated using constructed variables *st_perc_white_lit, st_perc_nonwhite_lit, cnty_perc_white_lit,* and *cnty_perc_nonwhite_lit.*.

---

st_perc_white_vote

st_perc_nonwhite_vote

Type: Numeric

Description: The percentage of white and nonwhite individuals who voted in each year and within each locality.

Availability:

- State-level values: 1976-2020
- County-level values: NA

Source information:

- *Data*: IPUMS CPS
- *Variables used*: VOTED, RACE, STATEFIP, YEAR, ASECWT
    - Individuals were identified as voting each year if their original value was:
        - VOTED= 2 [Voted]
    - Individuals were identified as not voting each year if their original value was:
        - VOTED = 1 [Did not vote]

---

st_race_ratio_vote

cnty_race_ratio_vote

Type: Numeric

Description: The ratio of the percentage of white individuals who voted to the percentage of non-white individuals who voted in each year and within each locality.

Availability:

- State-level values: 1976-2020
- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_race_ratio_vote* and *cnty_race_ratio_vote*.

---

st_white_e0

st_nonwhite_e0

Type: Numeric

Description: Life expectancy at birth for white and non-white individuals, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality.

Availability:
- State-level values: 1900-1940
- County-level values: NA

Source information:
- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_e0

Type: Numeric

Description: The ratio of life expectancy at birth for white individuals to life expectancy at birth of non-white individuals in each year and within each locality.

Availability:
- State-level values: 1900-1940
- County-level values: NA

Source information:
- Ratios calculated using constructed variables *st_white_e0* and *st_nonwhite_e0*.

---

st_whitemale_e0

st_nonwhitemale_e0

Type: Numeric

Description: Life expectancy at birth for white and non-white males, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_male_e0

Type: Numeric

Description: The ratio of life expectancy at birth for white males to life expectancy of non-white males at birth in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_whitemale_e0* and *st_nonwhitemale_e0*.

---

st_whitefemale_e0

st_nonwhitefemale_e0

Type: Numeric

Description: Life expectancy at birth for white and non-white females, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_female_e0

Type: Numeric

Description: The ratio of life expectancy at birth for white females to life expectancy of non-white females, in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_whitefemale_e0* and *st_nonwhitefemale_e0*.

---

st_white_e1

st_nonwhite_e1

Type: Numeric

Description: Life expectancy at age one for white and non-white individuals, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality.

Availability:

- State-level values: 1900-1940
- County-level values: NA

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_e1

Type: Numeric

Description: The ratio of life expectancy at age 1 for white individuals to life expectancy at age 1 of non-white individuals, in each year and within each locality

Availability:

- State-level values: 1900-1940

- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_white_e1* and *st_nonwhite_e1*.

---

st_whitemale_e1

st_nonwhitemale_e1

Type: Numeric

Description: Life expectancy at age 1 for white and non-white males, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_male_e1

Type: Numeric

Description: The ratio of life expectancy at age 1 for white males to life expectancy of non-white males at age 1 in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_whitemale_e1* and *st_nonwhitemale_e1*.

---

st_whitefemale_e1

st_nonwhitefemale_e1

Type: Numeric

Description: Life expectancy at age 1 for white and non-white females, calculated using period life table [a(0)=.3, a(1)=.4, a(x)=.5] and period-specific mortality rates [M(x) converted to q(x) using Chiang's formula], in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age, sex, and race-specific death rates

---

st_race_ratio_female_e1

Type: Numeric

Description: The ratio of life expectancy at age 1 for white females to life expectancy of non-white females at age 1 in each year and within each locality.

Availability:

- State-level values: 1939-1999
- County-level values: NA

Source information:

- Ratios calculated using constructed variables *st_whitefemale_e1* and *st_nonwhitefemale_e1*.

***Infant mortality variable*

Type: Numeric

Variable description:

- 1915-1960: Deaths (ages 0-1) per 1,000 live births
- 1969-2001: Probability of survival from age 0 to age 1

Levels: State
Availability:

- *State*: x
- *County*: x

Source information:

- *Data*: CDC NCHS NVSS
- *Variables used*: Age- and race-specific death rates; life table q(x) values

---

dissim_state

Type: Numeric

Variable description: Index of Dissimilarity based on Reardon & Firebaugh (2002); for calculation guidelines, see (https://coascenters.howard.edu/dissimilarity-index-tutorial). The state value represents how the aggregate measure of how much the county-level populations would need to be redistributed to have equivalent proportions of white and non-white residents in every county that match the state average proportion (e.g., proportion of white residents in county = proportion of white residents in state). If value=1, the state has perfect residential segregation for each county between white and non-white residents; if value=0, no residential segregation exists for any counties, and all have group proportions equivalent to the overall state proportion.

Levels: State (aggregated from counties)
Availability:

- *State*: x
- *County*: x

Source information:

- *Data*: IPUMS NHGIS
- *Variables used*:
    - *1920-1940*: White population (AF15001), non-white population (AF15002)
    - *1950*: White population (B1T001), non-white population (B1T002)
    - *1960*: White population (B48001), non-white population (B48002)
    - *1970*: White population (B18AA1970), non-white population (B18AB1970 + B18AC1970 + B18AD1970)
    - *1980*: White population (B18AA1980), non-white population (B18AB1980 + B18AC1980 + B18AD1980)
    - *1990*: White population (B18AA1990), non-white population (B18AB1990 + B18AC1990 + B18AD1990)
    - *2000*: White population (B18AA2000), non-white population (B18AB2000 + B18AC2000 + B18AD2000 + B18AE2000)

- *2010*: White population (B18AA2010), non-white population (B18AB2010 + B18AC2010 + B18AD2010 + B18AE2010)
- *2020*: White population (B18AA2020), non-white population (B18AB2020 + B18AC2020 + B18AD2020 + B18AE2020)

---

dissim_county

Type: Numeric

Variable description: Index of Dissimilarity based on Reardon & Firebaugh (2002); for calculation guidelines, see (https://coascenters.howard.edu/dissimilarity-index-tutorial). The county value represents how the aggregate measure of how much the tract-level populations would need to be redistributed to have equivalent proportions of white and non-white residents in every tract that match the county average proportion (e.g., proportion of white residents in tract = proportion of white residents in county). If value=1, the county has perfect residential segregation for each tract between white and non-white residents; if value=0, no residential segregation exists for any tracts, and all have group proportions equivalent to the overall county proportion.

Levels: County (aggregated from census tracts)
Availability:

- *State*: x
- *County*: x

Source information:

- *Data*: IPUMS NHGIS
- *Variables used*:
  - *1970*: White population (B18AA1970), non-white population (B18AB1970 + B18AC1970 + B18AD1970)
  - *1980*: White population (B18AA1980), non-white population (B18AB1980 + B18AC1980 + B18AD1980)
  - *1990*: White population (B18AA1990), non-white population (B18AB1990 + B18AC1990 + B18AD1990)
  - *2000*: White population (B18AA2000), non-white population (B18AB2000 + B18AC2000 + B18AD2000 + B18AE2000)
  - *2010*: White population (B18AA2010), non-white population (B18AB2010 + B18AC2010 + B18AD2010 + B18AE2010)
  - *2020*: White population (B18AA2020), non-white population (B18AB2020 + B18AC2020 + B18AD2020 + B18AE2020)

Sex/Gender Measures

Source Variable Universes

5. IPUMS CPS Annual Social and Economic Supplement Samples

- EDUC, EMPSTAT, LABFORCE, OFFPOV, OWNERSHP, HHINCOME, OCC1950, SEX, STATEFIP, COUNTY, YEAR, ASECWT

6. IPUMS USA (Harmonized U.S. Census Data)
    a. LIT, SEX, STATEICP, COUNTYICP, PERWT
7. IPUMS National Historic Geographic Information Systems
    a. DISSIM

Coding scheme for sex/gender

Data: IPUMS CPS

- *Variables used*: SEX, YEAR, ASECWT
    - Individuals were considered **male** if their original values were:
        - SEX = 1
    - Individuals were considered **female** if their original values were:
        - SEX = 2

Data: IPUMS USA

- *Variables used*: SEX, YEAR, ASECWT
    - Individuals were considered **male** if their original values were:
        - SEX = 1
    - Individuals were considered **female** if their original values were:
        - SEX = 2

Measures

st_perc_male_employed

st_perc_female_employed

cnty_perc_male_employed

cnty_perc_female_employed

Type: Numeric

Description: Percentages of males and females employed the percent females employed in each year and within each locality.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: EMPSTAT, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered employed if their original EMPSTAT value was:
        - 10 = Employed: At work
        - 12 = Employed: Has job, not at work last week
    - Respondents were considered unemployed if their original EMPSTAT value was:
        - 20 = Unemployed
        - 21 = Unemployed: Experienced worker
        - 22 = Unemployed: New worker

---

st_sex_ratio_employed

cnty_sex_ratio_employed

Type: Numeric

Description: The ratio of the percent of males employed to the percentage of females employed in each year in each locality.

- Note: values are coded as –1 if the percentage of females employed is 0.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_employed, st_perc_female_employed*, *cnty_perc_male_employed* and *cnty_perc_female_employed.*

---

st_perc_male_lf

st_perc_female_lf

cnty_perc_male_lf

cnty_perc_female_lf

Type: Numeric

Description: Percentages of males and females in labor force in each year and within each locality.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: LABFORCE, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered in labor force if their original LABFORCE value was:
        - 2 = Yes, in the labor force
    - Respondents were considered not in the labor force if their original LABFORCE value was:
        - 1 = No, not in the labor force

---

st_sex_ratio_lf

cnty_sex_ratio_lf

Type: Numeric

Description: The ratios of the percent of males in labor force to the percentage of females in labor force in each year in each locality.

- Note: values are coded as –1 if the percentage of females in the labor force is 0.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_lf, st_perc_female_lf, cnty_perc_male_lf, cnty_perc_female_lf*

---

st_perc_male_apl

st_perc_female_apl

cnty_perc_male_apl

cnty_perc_female_apl

Type: Numeric

Description: Percentages of males and females above the poverty line in each year and within each locality.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: OFFPOV, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered above the poverty line if their original OFFPOV value was:
        - 02 = Above poverty line
    - Respondents were considered below the poverty line if their original OFFPOV value was:
        - 01 = Below poverty line

---

st_sex_ratio_apl

cnty_sex_ratio_apl

Type: Numeric

Description: The ratio of the percent of males above the poverty line to the percentage of females in above the poverty line in each year and within each locality.

- Note: values are coded as –1 if the percentage of females above the poverty line is 0.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_apl, st_perc_female_apl, cnty_perc_male_apl, cnty_perc_female_apl.*

---

st_perc_male_hsd

st_perc_female_hsd

cnty_perc_male_hsd

cnty_perc_female_hsd

Type: Numeric

Description: Percentages of males and females with at least a high school degree in each year and within each locality.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: EDUC, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered to have attained a high school diploma if their original EDUC value was:
        - < 72 = Values indicating that the highest level of education completed was less than that which would be required for a high school diploma
    - Respondents were considered to have no attained a high school diploma if their original EDUC value was:
        - >= 72 = Values indicating that the highest level of education completed was greater than or equal to that which would be required for a high school diploma

---

st_sex_ratio_hsd

cnty st_sex_ratio_hsd

Type: Numeric

Description: The ratio of the percent of males who attained a high school diploma to the percentage of females who attained a high school diploma in each year and within each locality..

- Notes: values are coded as –1 if the percentage of females with a high school diploma is 0. Diploma attainment is captured for 1992-2021. 12 years of schooling is captured for 1962-1991, as a large portion of respondents were recorded as "12th grade, diploma unclear" in the source data in these earlier years. The "12th grade, diploma unclear" category was not indicated for any respondents in 1992-2021.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_hsd, st_perc_female_hsd, cnty_perc_male_hsd, cnty_perc_female_hsd.*

---

st_perc_male_bachelors

st_perc_female_bachelors

cnty_perc_male_bachelors

cnty_perc_female_bachelors

Type: Numeric

Description: Percentages of males and females with at least a bachelor's degree in each year and within each locality.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: EDUC, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered to have attained a bachelor's degree if their original EDUC value was:
        - < 110 = Values indicating that the highest level of education completed was less than that which would be required for a bachelor's degree
    - Respondents were considered to have not attained a bachelor's degree if their original EDUC value was:
        - >= 110 = Values indicating that the highest level of education completed was greater than or equal to that which would be required for a bachelor's degree

---

st_sex_ratio_bachelors

cnty_sex_ratio_bachelors

Type: Numeric

Description: The ratio of the percent of males who attained at least a bachelor's degree to the percentage of females who attained at least a bachelor's degree in each year and within each locality.

- Note: values are coded as –1 if the percentage of females with a bachelor's is 0. Degree attainment is captured for 1992-2021. 4 years of college is captured for 1962-1991, as degree attainment was not measured in these earlier years in the source data.

Availability:

- State-level values: 1962-2023
- County-level values: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_bachelors, st_perc_female_bachelors, cnty_perc_male_bachelors, cnty_perc_female_bachelors.*

---

st_perc_male_lit

st_perc_female_lit

cnty_perc_male_lit

cnty_perc_female_lit

Type: Numeric

Description: Percent of males and females who are literate in each year and within each locality.

Availability:

- State-level values: 1850-1930
- County-level values: 1850-1930

Source information:

- Data: IPUMS USA 1% Samples
- Variables used: LIT, SEX, STATEFIP, COUNTY, YEAR, PERWT
    - Respondents were considered literate if their original LIT value was:
        - 4 = Yes, literate (reads and writes)
    - Respondents were considered not literate if their original LIT value was:
        - 1 = No, illiterate (cannot read nor write)
        - 2 = Can't read, can write
        - 3 = Can't write, can read

---

st_sex_ratio_lit

cnty_sex_ratio_lit

Type: Numeric

Description: The ratio of the percent of males who are literate to the percentage of females who are literate in each year and within each locality.

- Note: values are coded as –1 if the percentage of literate females is 0.

Availability:

- State-level values: 1850-1930
- County-level values: 1850-1930

Source information:

- Ratios calculated using constructed variables *st_perc_male_lit, st_perc_female_lit, cnty_perc_male_lit, cnty_perc_female_lit.*

---

st_perc_male_occw

st_perc_female_occw

cnty_perc_male_occw

cnty_perc_female_occw

Type: Numeric

Description: Percentages of males and females who are in white-collar professions in each year and within each locality.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: OCC1950, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered to be in white-collar professions if their original OCC1950 value was:
        - 000-099 = Professional, Technical
        - 200-290 = Managers, officials, and proprietors
        - 400-490 = Sales workers
    - Respondents were considered to be in non-white-collar professions if their original OCC1950 value was:
        - Any other value, except 997 = Not reported/unknown; 999 = Not in universe

---

st_sex_ratio_occw

cnty_sex_ratio_occw

Type: Numeric

Description: The ratio of the percent of males who are in white-collar professions to the percentage of females who are in white-collar professions in each year and within each locality..

- Note: values are coded as –1 if the percentage of females who are in white-collar professions is 0.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_occw, st_perc_female_occw, cnty_perc_male_occw, cnty_perc_female_occw*.

---

st_perc_male_occp

st_perc_female_occp

cnty_perc_male_occp

cnty_perc_female_occp

Type: Numeric

Description: Percentages of males and females who are in pink-collar professions in each year within each locality.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: OCC1950, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered to be in pink-collar professions if their original OCC1950 value was:
        - 300-390 = Clerical and kindred
        - 700-720 = Service workers (private household)
        - 730-790 = Service workers (not household)
    - Respondents were considered to be in non-pink-collar professions if their original OCC1950 value was:
        - Any other value, except 997 = Not reported/unknown; 999 = Not in universe

---

st_sex_ratio_occp

cnty_sex_ratio_occp

Type: Numeric

Description: The ratio of the percent of males who are in pink-collar professions to the percentage of females who are in pink-collar professions in each year within each locality.

- Note: values are coded as –1 if the percentage of females who are in pink-collar professions is 0.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_occp, st_perc_female_occp, cnty_perc_male_occp, cnty_perc_female_occp.*

---

st_perc_male_occb

st_perc_female_occb

cnty_perc_male_occb

cnty_perc_female_occb

Type: Numeric
Description: Percentages of males and females who are in blue-collar professions in each year within each locality.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Data: IPUMS CPS Annual Social and Economic Supplement Samples
- Variables used: OCC1950, SEX, STATEFIP, COUNTY, YEAR, ASECWT
    - Respondents were considered to be in blue-collar professions if their original OCC1950 value was:
        - 100-123 = Farmers
        - 500-595 = Craftsmen
        - 600-690 = Operatives
        - 810-840 = Farm laborers
        - 910-970 = Laborers
    - Respondents were considered to be in non-blue-collar professions if their original OCC1950 value was:
        - Any other value, except 997 = Not reported/unknown; 999 = Not in universe

---

st_sex_ratio_occb

cnty_sex_ratio_occb

Type: Numeric

Description: The ratio of the percent of males who are in blue-collar professions to the percentage of females who are in blue-collar professions in each year within each locality.

- Note: values are coded as –1 if the percentage of females who are in blue-collar professions is 0.

Availability:

- State: 1968-2023
- County: 1996-2023

Source information:

- Ratios calculated using constructed variables *st_perc_male_occb, st_perc_female_occb, cnty_perc_male_occb, cnty_perc_female_occb.*

Source Variable Universes

[exact wording from ipums, but with non-ASEC universes cut out]

IPUMS CPS Annual Social and Economic Supplement Samples

- EDUC
    - 1962, 1964-1979: Persons age 14+ (pre-1968 samples do not include persons under age 14). 1980+: Persons age 15+.
- EMPSTAT
    - 1962-1967: Persons age 14+. 1968-1988: Persons age 14+. 1989+: Persons age 15+.
- LABFORCE
    - 1962-1967: Civilians age 14+. 1968-1987 (ASEC); 1976-1988 (non-ASEC): Civilians age 14+. 1989+: Civilians age 15+.
- OFFPOV
    - 1969-1979: All persons except for secondary individuals under 14. 1980-2022: All persons except for secondary individuals under 15.
- OCC1950
    - 1968-1969: Civilians age 14+ who were employed, looking for employment, or unemployed who had ever worked and were in month in sample 1 or 5. 1970-1975: Civilians age 14+ who were employed or were unemployed, looking, and had worked in the past. 1976-1987: Civilians age 14+ who were employed, unemployed but had worked in the past, or were not in labor force but had worked in the past 5 years and were in month in sample 4 or 8. 1988: Civilians age 14+ who were employed, on layoff, unemployed but had worked in the past, or were not in labor force but had worked in the past 5 years and were in month in sample 4 or 8. 1989-1993: Civilians age 15+ who were employed, on layoff, unemployed but had worked in the past, or not in labor force but had worked in the past 5 years. 1994+: Civilians age 15+ who were employed, on layoff, unemployed but had worked in the past, or not in labor force but had worked in the past year.
- SEX
    - 1962-1967: Persons age 14+. 1968+: All persons.
- STATEFIP
    - All households and group quarters.
- COUNTY
    - All households.
- YEAR
    - All households.
- ASECWT
    - 1962-1967: Persons age 14+. 1968-2022: All persons.
- Note: only certain counties are included. not all states are available in all years
- Note: when no within-universe male or female respondents in a county, the ratio is missing.

IPUMS USA

- LIT
    - 1850-1860: Persons age 20+. 1870-1910: Persons age 10+. Not available in the 1880 100% database. 1910 Puerto Rico: Persons age 5+. 1920-1940: Persons age 10+
- SEX
    - All persons.
- STATEFIP
    - All households and group quarter.
- YEAR
    - All households and group quarters.
- PERWT
    - All persons.