

## Using NEO CANDO -- General Information and Precautions

---

Most of the data in NEO CANDO are updated annually. The Census data are the exception. The Census of the Population occurs every ten years. Many of the data are available from 1990 to the most current year. However, this will vary by data source. Only census, vital statistics and HMDA data are available for the 16 counties outside Cuyahoga County. The following information provides the user with some general information and caveats regarding the data in NEO CANDO. However, we encourage the user to read the data documentation for each of the data sources we provide to understand how we calculate the indicators, precautions and caveats about using the data, and how to appropriately cite the data from NEO CANDO.

### Table of Contents ([links](#))

[Infrequent event problem](#)

[Rate versus Count](#)

[Ratio](#)

[Median](#)

[Population counts used in denominators](#)

[Population in non-census years](#)

[Geocoding](#)

[Geographic reference maps](#)

---

### Infrequent event problem

There are some events that occur infrequently (infant deaths, homicides), and the user should be cautioned in using these numbers to calculate annual rates at the neighborhood or census tract level. The fluctuation from year to year could be an anomaly rather than a trend. In these cases, it is suggested that a three-year average, or even five-year average, be used to calculate the rate to "smooth" any fluctuations that are not meaningful. For example, the infant mortality rate is an example where the rate may be unreliable due to the small number of infant deaths. To illustrate calculating a three year average:

| <u>Year</u>  | <u>Infant deaths</u> | <u>Live births</u> |
|--------------|----------------------|--------------------|
| 1994         | 15                   | 200                |
| 1995         | 6                    | 205                |
| 1996         | 4                    | 199                |
| <b>TOTAL</b> | <b>25</b>            | <b>604</b>         |

Three year average =  $25/604 * 1000 = 41$  infant deaths per 1000 live births (infant mortality rate)

In order to compare the data over time, choose the three years prior to 1994 and calculate a three-year average for this time period also.

### Rate versus Count

Most data in NEO CANDO are reported in counts and rates. To determine which is more appropriate for your use, consider the following:

A count would be used when you want to determine the magnitude of a problem/issue that needs to be addressed. For example, it would be important to know the actual number of births to first time mothers or teen mothers to determine the need for a nurse visitation program geared toward these new mothers.

A **rate** would be used when comparing across geographic areas or over time periods. A rate measures the probability of an event occurring in a particular area during a particular time period. Using the above example, a rate might be used to determine the geographic areas that are more likely to have a higher rate of births to teens or first time mothers in order to target the nurse visitation program to areas of greater need.

A **percent** is a rate per 100. Poverty rates are per 100 population. Poverty rate is the same as percent poor. Rates per 1,000 are commonly used when reporting vital statistics related indicators. Rates per 100,000 are commonly used when reporting crime related statistics. The "per" number is called the multiplier. When calculating a rate you need 3 pieces of information – the numerator, the denominator and the multiplier. For example, to calculate the poverty rate we need to know the number of persons living below poverty (numerator), the total number of persons for whom poverty status is determined (population living below poverty + population living above poverty) (the denominator) and the "100".

**It will be clear in NEO CANDO what indicators are rates – the multiplier (1,000 or 100,000) or the word "rate" is included in the indicator name.**

### Ratio

In NEO CANDO, we do have one ratio – the adult/child ratio. The adult/child ratio simply divides the number of adults by the number of children to determine how many adults there are per child. If there are 100 adults and 50 children, the ratio would be 2 adults for every child – 100 divided by 50 = 2.

### Median

Medians are used to report income, rent and housing value in CANDO. The median represents the middle value in a distribution. The median divides the total frequency into two equal parts. Half of the distribution has a value above the median and half of the distribution has a value below the median.

When comparing indicators across geographic areas or time periods, it is important to consider structural, geographic and economic variations that may contribute to the differences in the rate.

### Population counts used in denominators

Denominators in rates are usually, but not always, the residential population. Rates are also computed using other denominators, such as teenage female population when computing the teen birth rates. Census tracts, in general, have a low population, and we caution the user in interpreting rates at this geographic level. Three-year averages or aggregating tracts could be used to avoid the low population issue. This is particularly important when calculating rates based on specific population groups rather than on the entire population. We consider tracts with a population less than 100 to be non-residential, and therefore consider rates calculated using these tracts to be unreliable.

The neighborhoods of Downtown and Industrial Valley are non-residential areas and rates calculated based on these areas should be viewed cautiously.

### Population in non-census years

The population of the census tract or neighborhood is used to calculate many of the rates within NEO CANDO. The Census of the Population occurs every ten years on the decade. In order to calculate the population between census years, we use linear interpolation and extrapolation techniques. We are unable to release the tract level population estimates. However, the annual population estimates for the counties and places as a whole are available from the Census Bureau.

## Geocoding

Much of the data the Center receives is at the address level. In order to determine what census tract the address is located in, the Center geocodes the data. Geocoding assigns latitude and longitude, 2000 census tracts and other geographic identifiers, such as zip code and municipality, to valid addresses. Not all addresses can be geocoded. The main reason addresses do not geocode is that the address range does not exist. In these cases, the address cannot be assigned to a specific tract or neighborhood; therefore, we put them in the UNKNOWN category.

In order to assign the Statistical Planning Area (SPA), or neighborhood, to the census tracts in Cleveland, we have a correspondence file that contains all of the census tracts and the neighborhood associated with each census tract. Each neighborhood in the City of Cleveland and each suburban municipality consist of a number of census tracts. The number of census tracts in each neighborhood/municipality varies depending on population size. For a list of the census tracts in each neighborhood [click here](#).

The 1990 data in NEO CANDO have been put into the 2000 census tract boundaries to allow for comparisons between the 1990 and 2000 Censuses. Some census tract boundaries and numbers change from Census year to Census year. In order to accurately compare the data over time, the data must be in the same geographic boundaries. For those census tracts that changed between the 2 Census years, we used GIS techniques to determine the proportion of the 1990 tract population that lived in the corresponding 2000 tract.

## Geographic reference maps

The Center has created geographic reference maps showing the cities, townships and villages within each of the 17 counties included in NEO CANDO. Within Cuyahoga County we also have maps for each of the 36 statistical planning areas in the City of Cleveland. The maps also include the census tracts that are within or cross the neighborhoods, cities, townships or villages. To view these geographic reference maps [click here](#).