CRP 4080: Introduction to Geographic Information Systems for planners

Lecture 6: Census Data Selection & preparation

Wenzheng Li, Ph.D.
City and Regional Planning
Fall 2024

Outlines

Understand Basic Census Geography

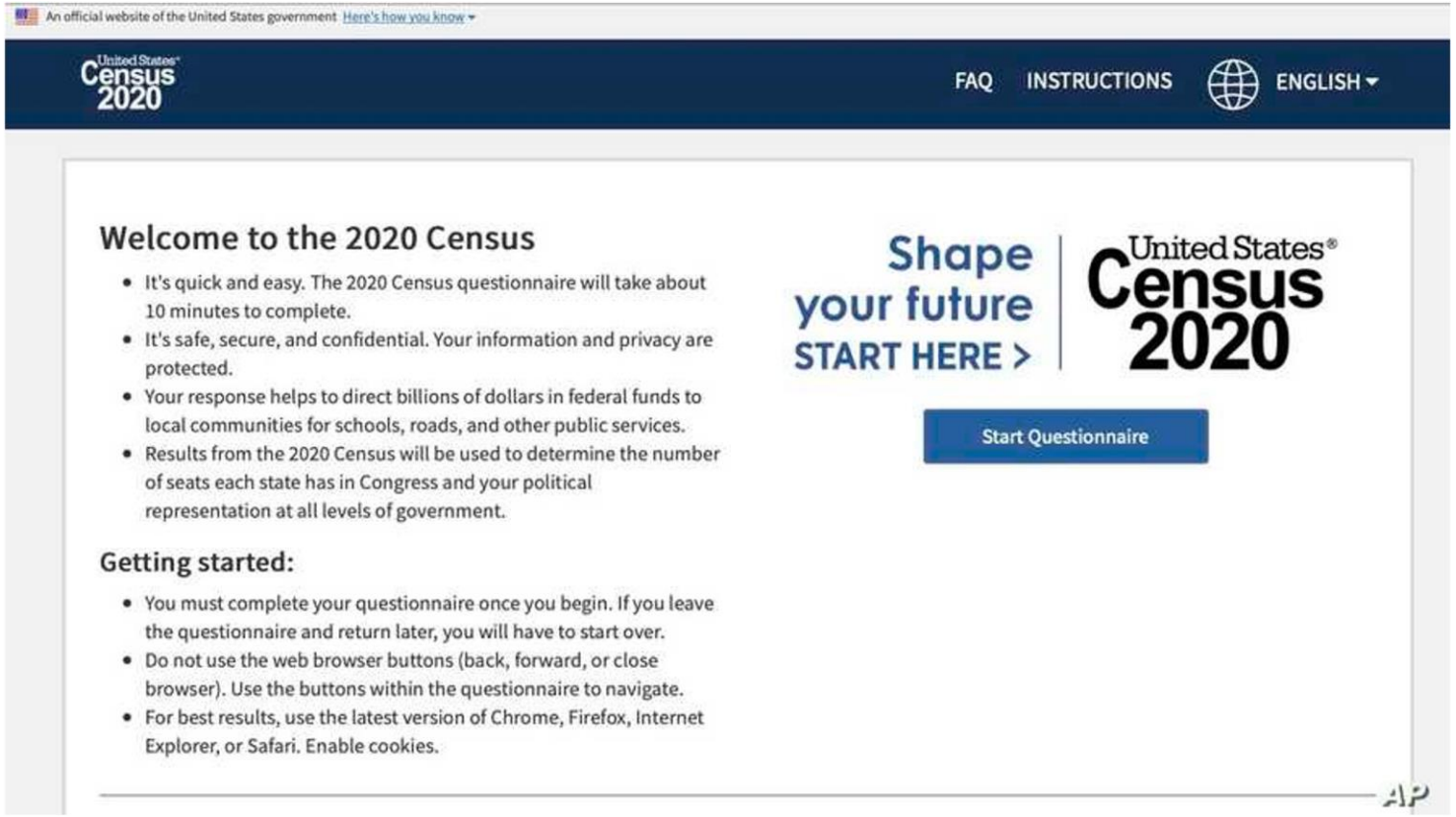
- Census tract, Block, Block Group
- Other local designations: Places, county subdivision

Data sources: CUGIR, Census Bureau, Social Explore

How to find GIS data: Join-table

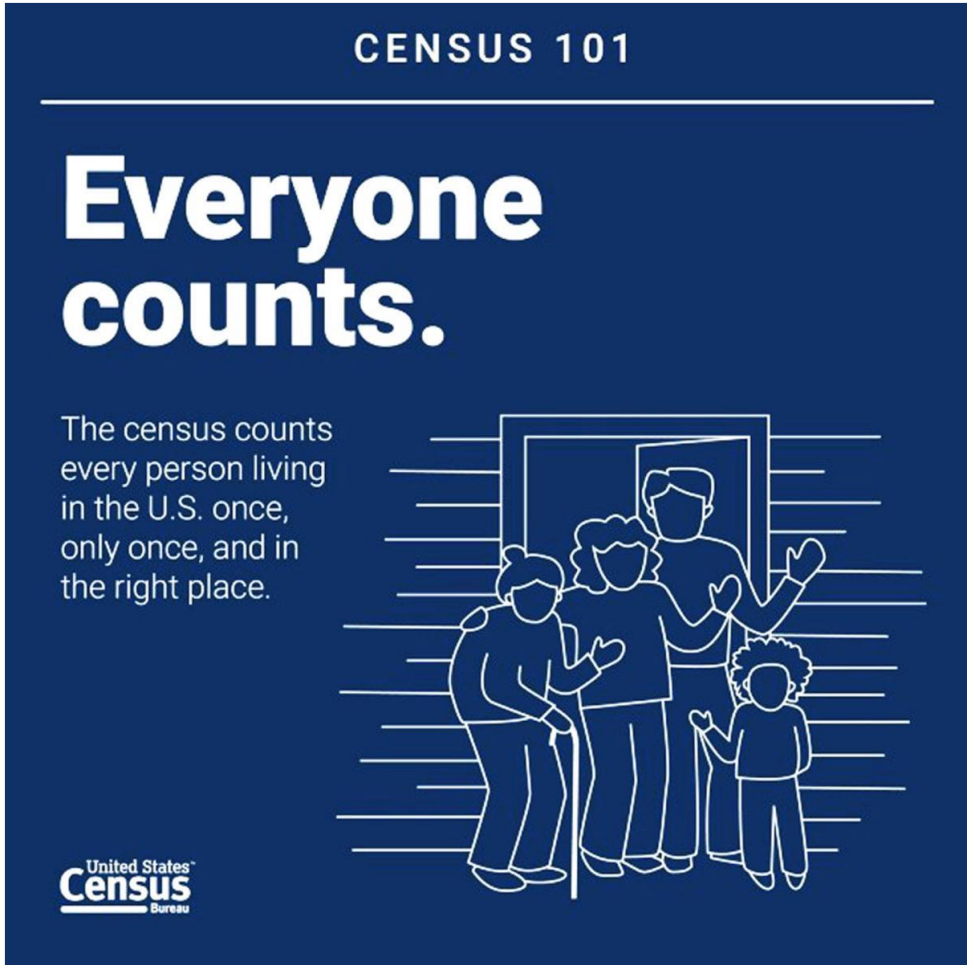
The U.S Census

Beginning March 12, 2020, the public was invited to respond to the census at my2020census.gov. The 2020 Census was the first U.S. census to invite people to respond online.



Decennial Census

- Takes place every 10 years
- 100% count of the population
- Written in US constitution, Article 1 section 2
- Since 1790, it determines the number of seats in the US house of representatives
- Data published down to census block level



Uses of Census Information

- Determines the number of seats each state has in the U.S. House of Representatives
- Defines congressional and state legislative districts, school districts and voting precincts
- Determines how federal funding are spent on infrastructure, programs and services each year
- Provides government the information to offer public services, and provides business and nonprofit organization with critical information for planning decisions
- Provides population benchmarks for nearly every other United States survey

Understanding Census Geography

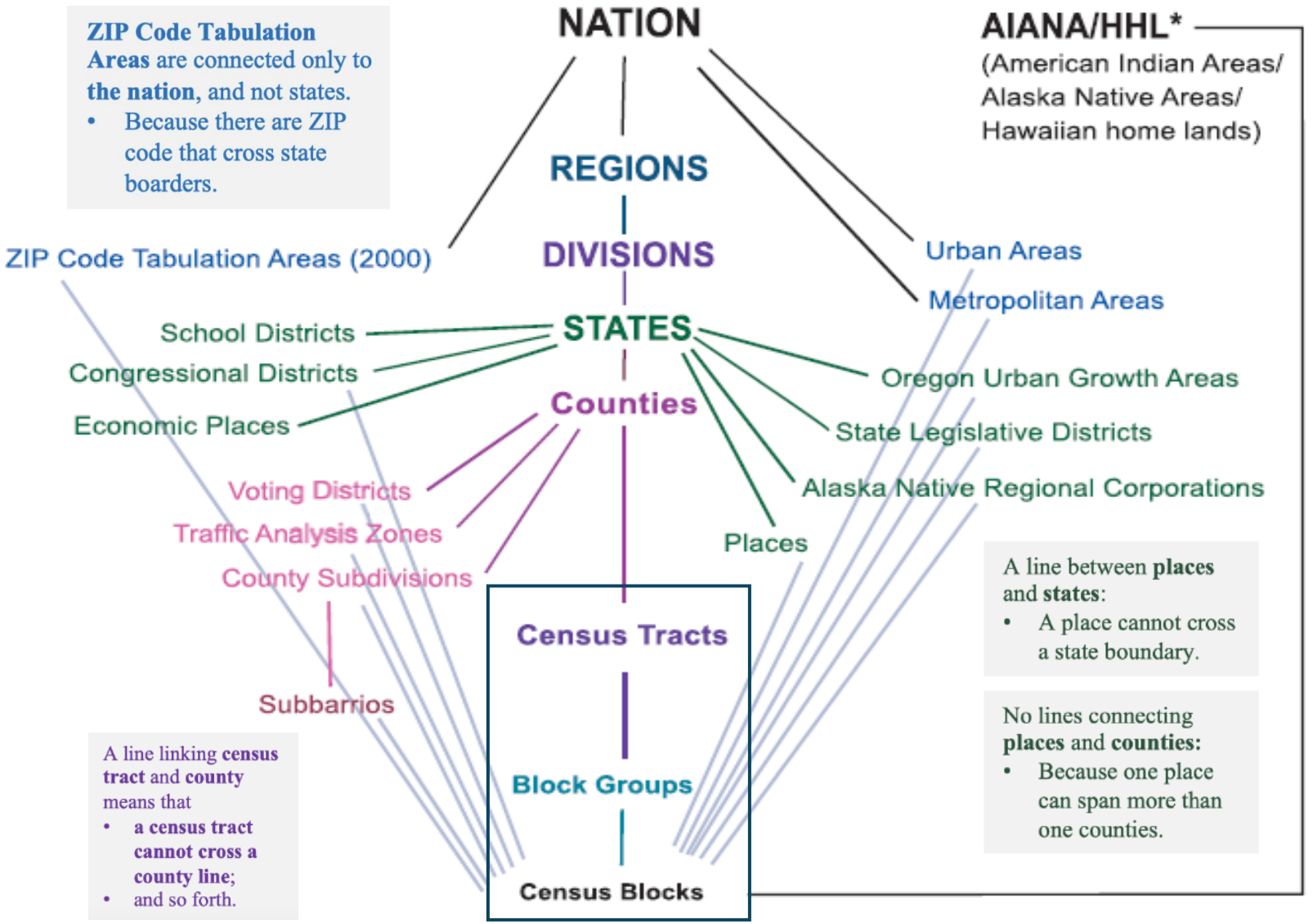
- Nation
- Region
- Division
- State
- County
- County Subdivision
- Census Tract
- Block Group
- Place
- Alaska native regional corporation
- Native Area
- Tribal Subdivision
- American Indian Area/Alaska Native Area (Reservation or Statistical Entry Only)
- American Indian Area (Off-Reservation Trust Land Only/Hawaiian Home Land)
- Tribal Census Tract
- Core Based Statistical Area
- Metropolitan Division
- Combined Statistical Area
- Combined NECTA
- NECTA
- Urban Area
- Congressional District
- State Senate District
- State House District
- PUMA
- ZIP Code Tabulation Area
- School District (Elementary)
- School District (Secondary)
- School District (Unified)

Geographic Entities

The geographic entities are divided into two types: legal/administrative areas, and statistical areas.

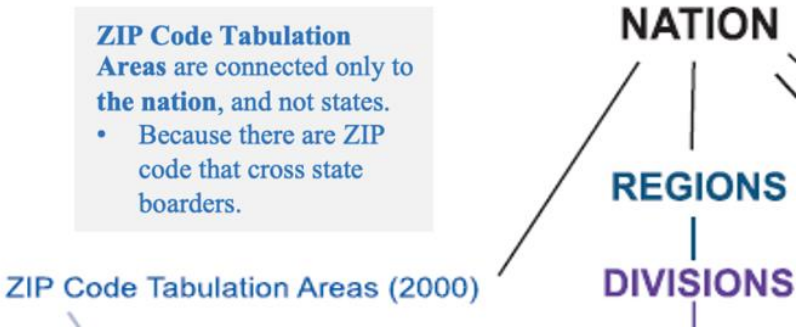
Legal/Administrative Areas	Statistical Areas
Nation	Regions
States	Divisions
Counties	Census County Divisions
Minor Civil Divisions	Census Designated Places
Incorporated Places	Metropolitan and Micropolitan Statistical Areas
Congressional Districts	Urban / Rural Areas
School Districts	Census Tracts
Voting Districts	Block Groups
ZIP Code Tabulation Areas	Public Use Microdata Areas

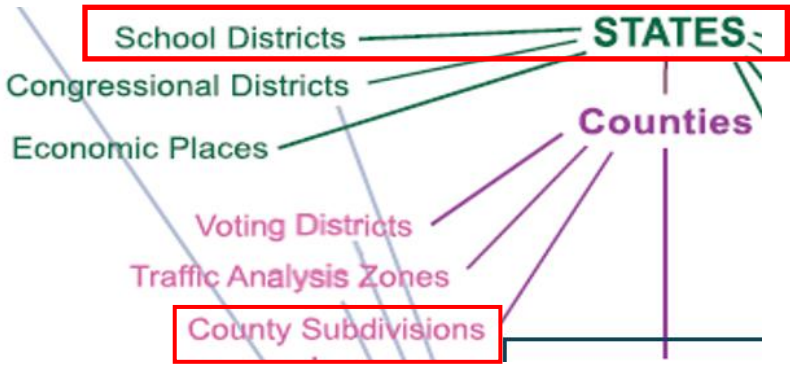
Census Geographic Hierarchy



Examples of the Relationships on the Hierarchy.

- ZIP Code Tabulation Areas (ZCTAs) are based on the U.S. Postal Services ZIP Codes and must **fall within the national boundary only**. In a few cases, ZCTAs can cross into bordering states.
- School districts must **fall within each state**. States are responsible for updating their boundaries, and districts may cross county and place boundaries.
- County subdivisions, as the name suggests, **must fall within the county**. Many county subdivision names repeat throughout the nation and throughout the same state, so it is important you know which county you are working in. For example, in 2010, Beaver was used as the name of 45 different county subdivisions.

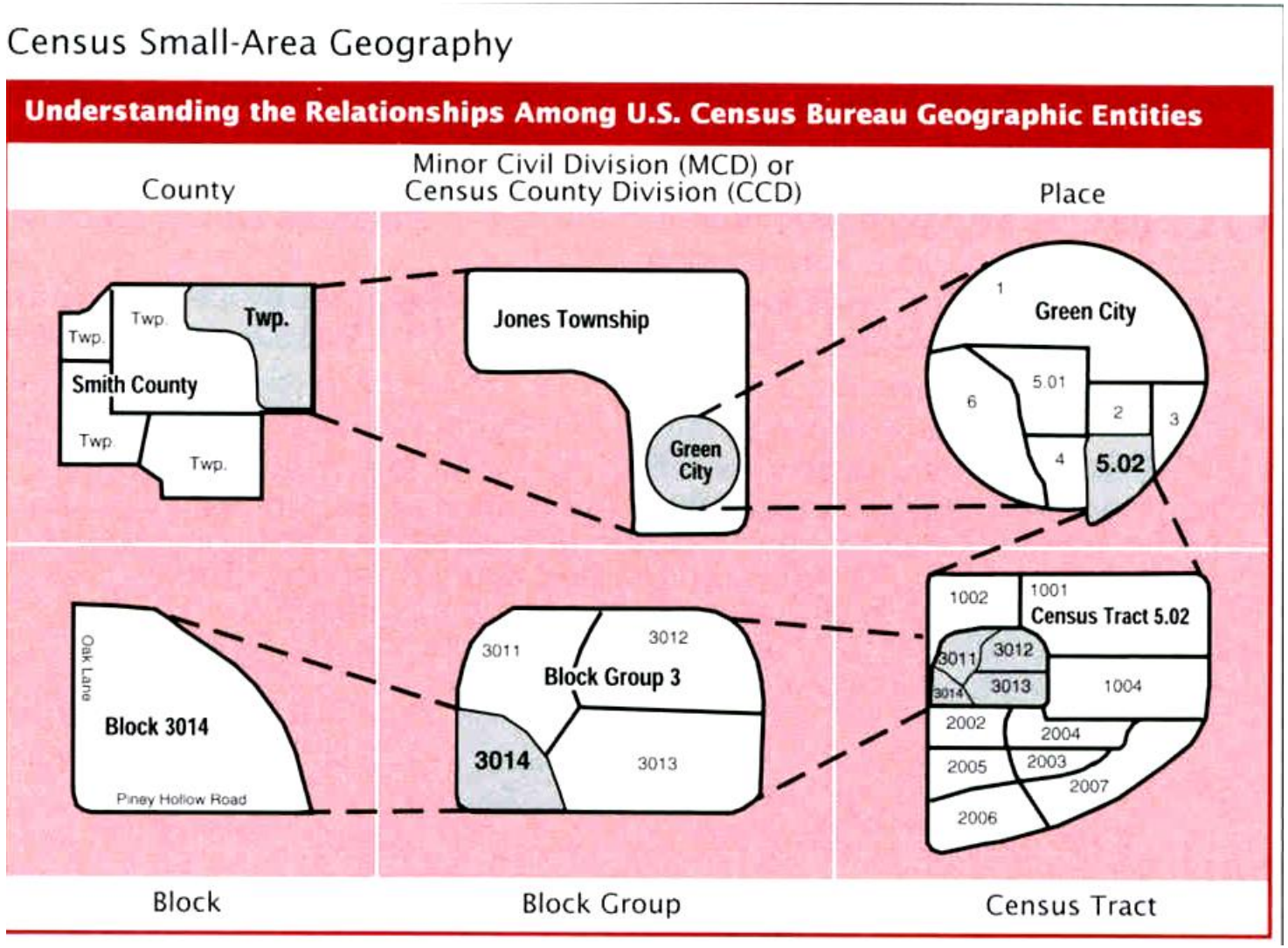




Examples of the Relationships on the Hierarchy.

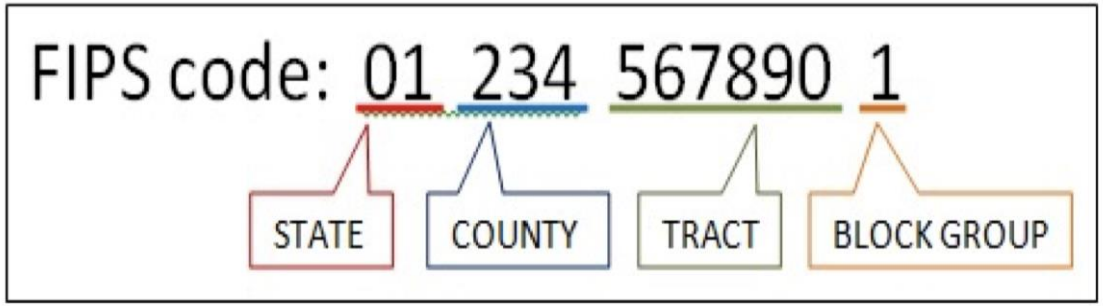
- Census tracts **must stay within a county and therefore a state**.
- Block groups **must stay within each census tract**, so they must also stay within a county and state.
- Blocks fall **within everything**! They are the building blocks for all other geographies and therefore nest within all other geographies.

Small-Area Geography Overview



FIPS Code: Federal Information Processing Standards (FIPS)

STATE	COUNTY	TRACT	BLOCK	STFID_1
36	109	000100	1000	361090001001000
36	109	000100	1001	361090001001001
36	109	000100	1002	361090001001002
36	109	000100	1003	361090001001003
36	109	000100	1004	361090001001004
36	109	000100	1005	361090001001005
36	109	000100	1006	361090001001006
36	109	000100	1007	361090001001007
36	109	000100	1008	361090001001008
36	109	000100	1009	361090001001009
36	109	000100	1010	361090001001010
36	109	000100	1011	361090001001011
36	109	000100	1012	361090001001012
36	109	000100	1013	361090001001013
36	109	000100	1014	361090001001014
36	109	000100	1015	361090001001015
36	109	000100	1016	361090001001016
36	109	000100	1017	361090001001017
36	109	000100	1018	361090001001018
36	109	000100	1019	361090001001019
36	109	000100	1020	361090001001020



Q: what is the unique FIPS code for Tompkins County?

Blocks, Block Groups and Census Tracts

Blocks:

the smallest geographic census unit

- Cover the entire nation
- Do not cross census tracts or counties
- Block numbers may differ from census to census
- Size: average about 100 people
- Generally bounded by visible features and legal boundaries

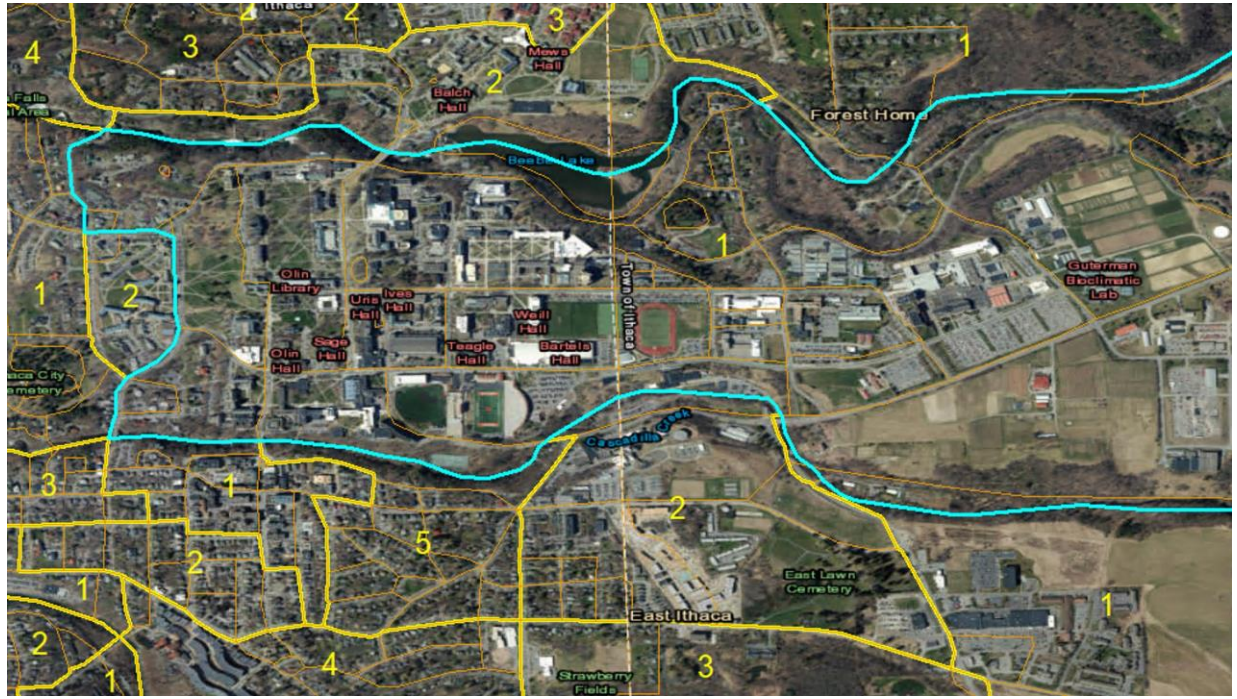
Sibley Hall: Census Block 1035 - 361090003001035

Blocks, Block Groups and Census Tracts

Block groups:

Combination of Census Blocks and are also a subdivision of Census Tracts

- Made up of on average 40 Census Blocks.
- Smallest areas for which all sample data available (American Community Survey data)
- Size: 600 to 3000 people



Blocks, Block Groups and Census Tracts

Census Tract

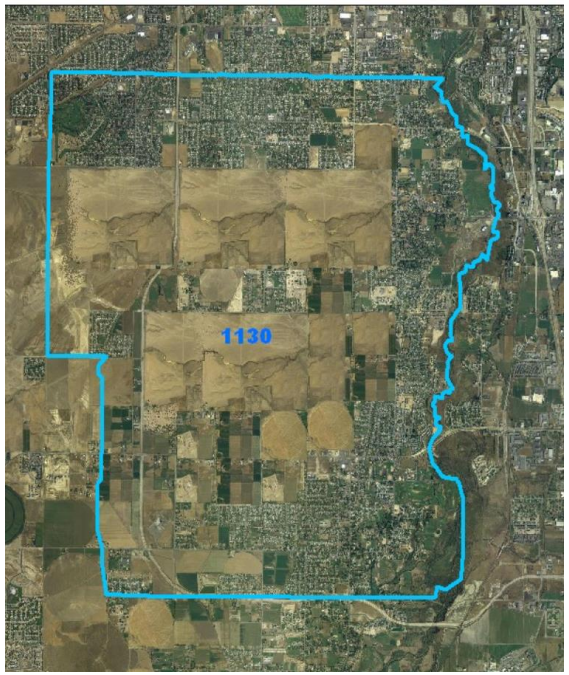
Small, relatively **permanent, consistent** statistical divisions of a county

- Boundaries cannot across county line
- Size: Optimally contains 4,000 people; range between 1,200 and 8,000
- Approximately 65,000 census tracts in U.S.
- Tracts are designed to be fairly homogeneous with respect to demographic and economic conditions when they are first established.
- When a census tract experiences growth and the internal population grows beyond 8,000 persons, the tract is split up. This review and revision process is conducted every decade with collaboration from local planning agencies.

Census Tract

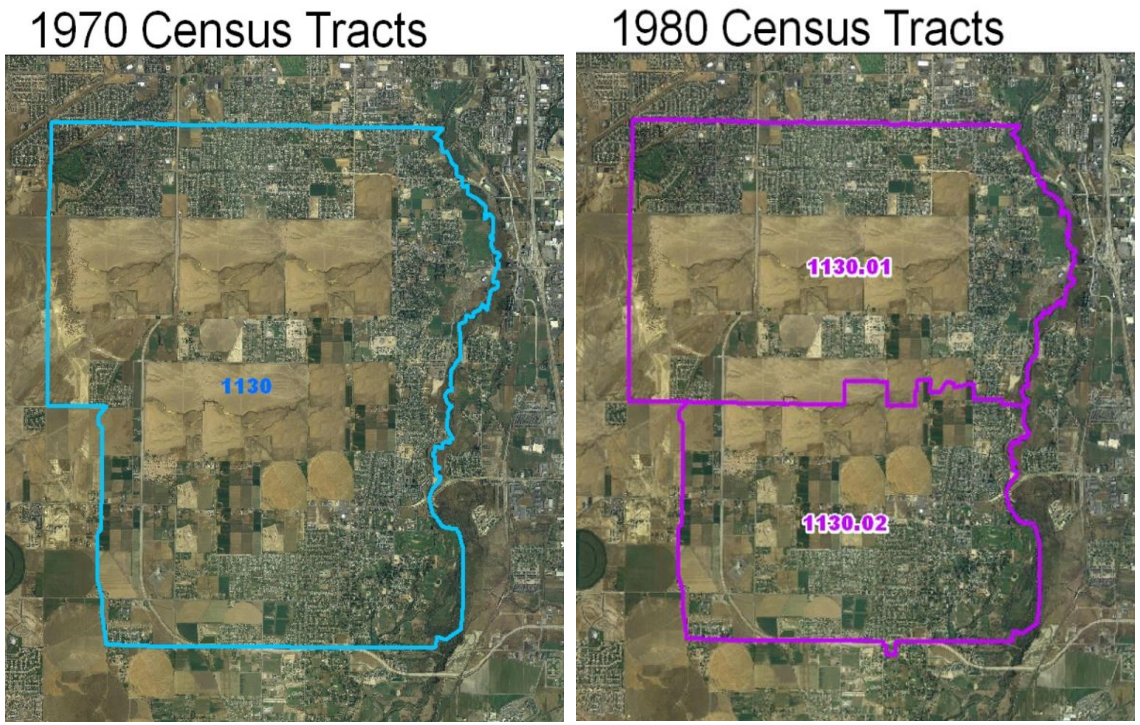
- an example of an area where population continues to grow and therefore the census tract is split over the decades.
- South Jordan City, Utah

1970 Census Tracts



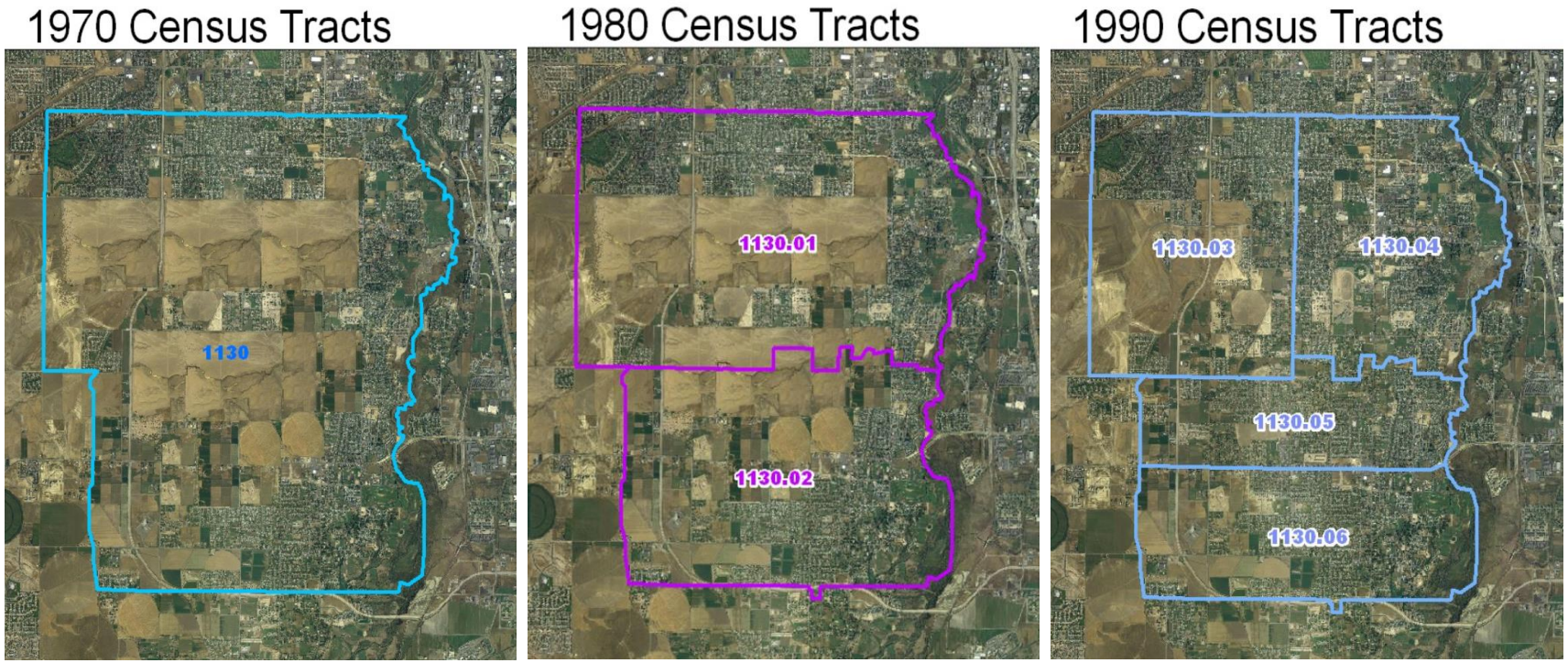
Census Tract

- an example of an area where population continues to grow and therefore the census tract is split over the decades.
- South Jordan City, Utah



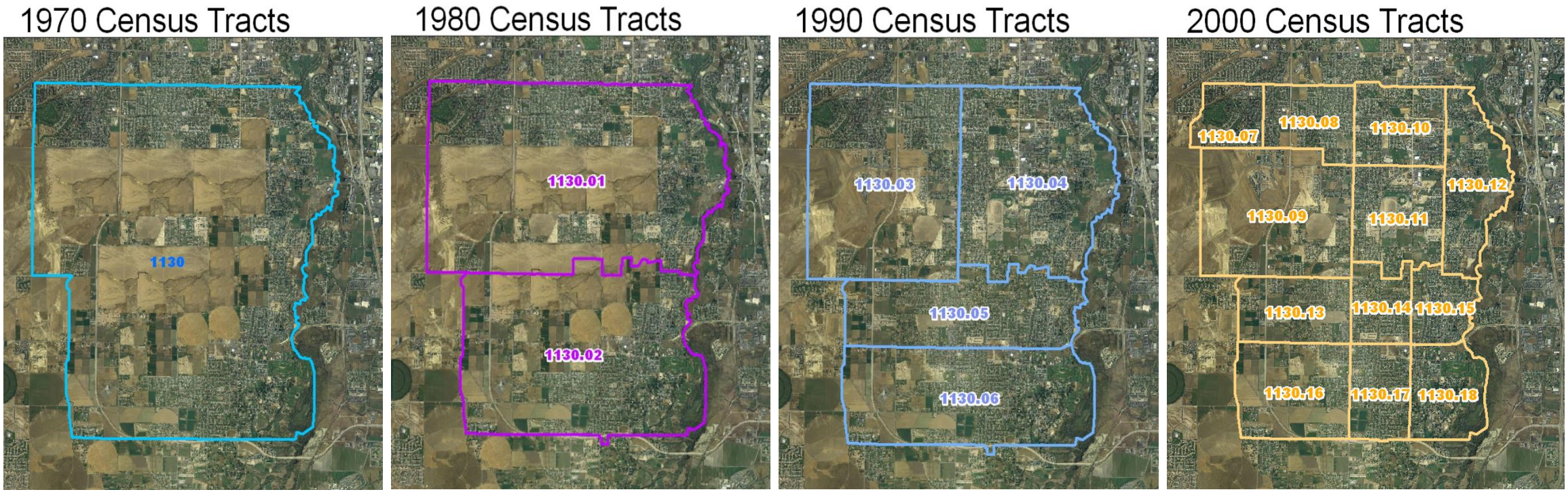
Census Tract

- an example of an area where population continues to grow and therefore the census tract is split over the decades.
- South Jordan City, Utah



Census Tract

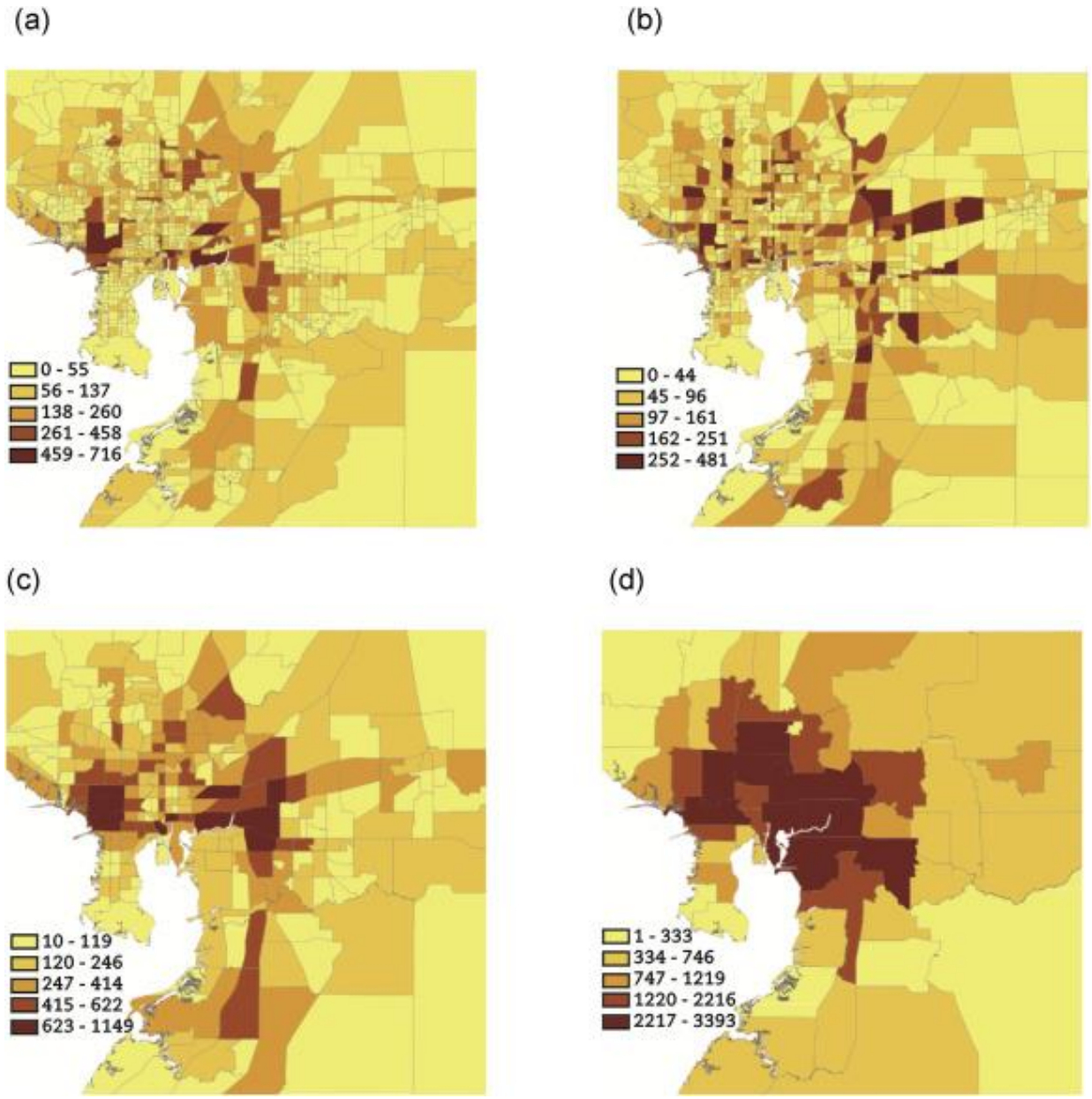
- an example of an area where population continues to grow and therefore the census tract is split over the decades.
- South Jordan City, Utah



The Modifiable Areal Unit Problem (MAUP)

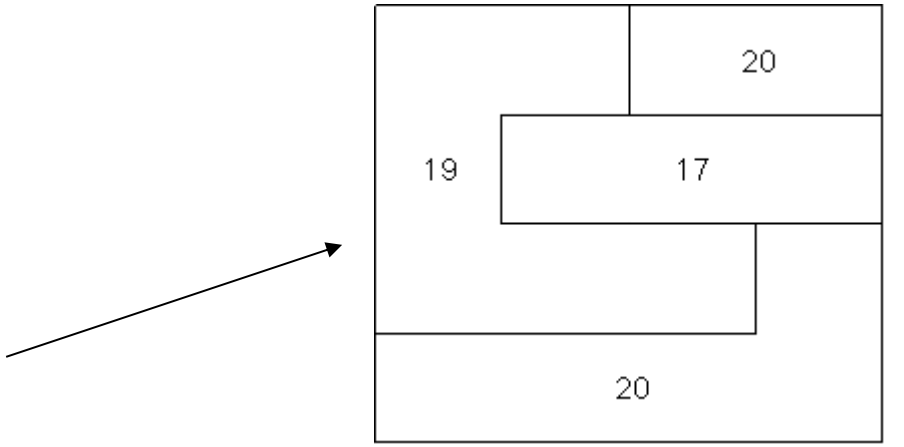
(a) block group.
(b) Traffic analysis zone
(c) census tract.
(d) zip code

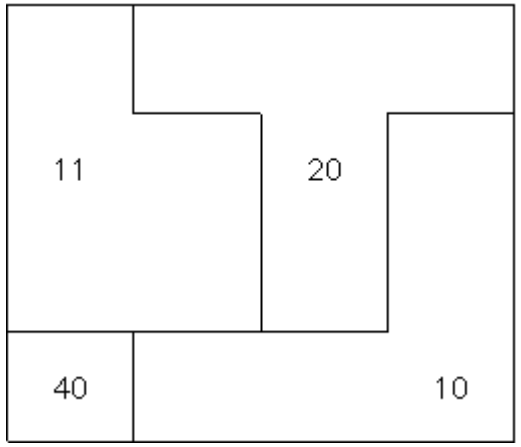
when point-based measures of spatial phenomena are aggregated into districts, for example, population density or illness rates. The resulting summary values (e.g., totals, rates, proportions, densities) are influenced by both the shape and scale of the aggregation unit



MAUP

10	20	20	20
5	10	20	10
15	15	20	10
40	10	10	10





<u>Aggregation (zonation) effect:</u> differences which occur when we group zones of similar sizes in different spatial arrangements.

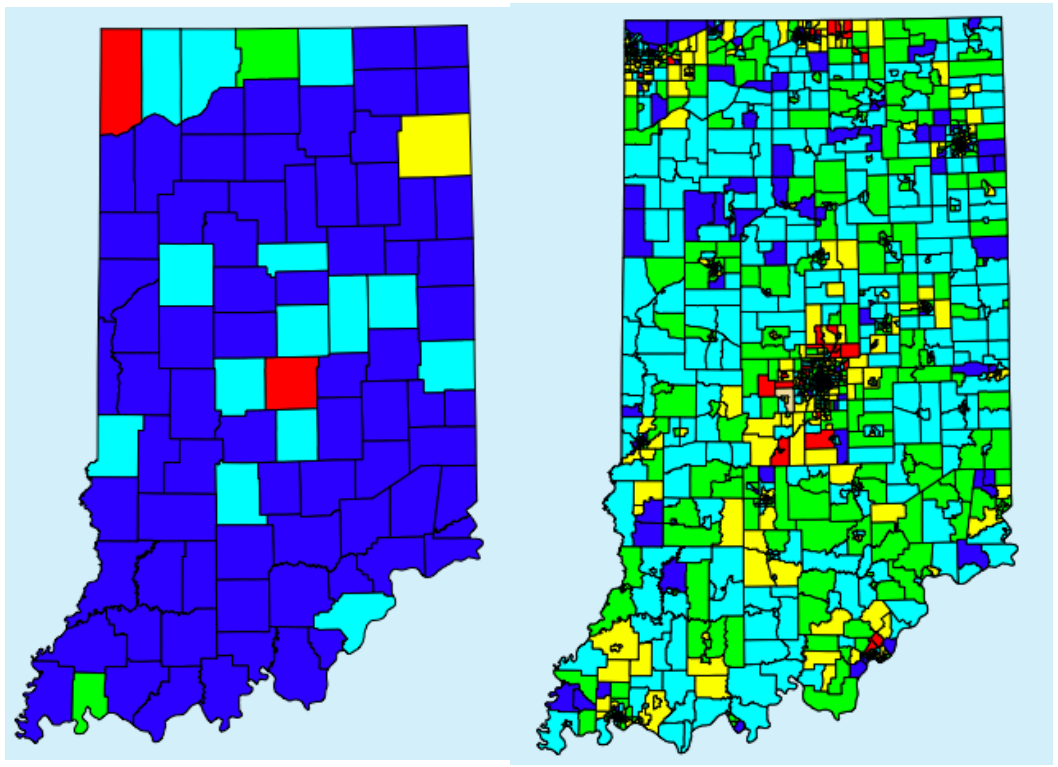
Perceived...

Average income:
$30,000

Real

Average income:
$10,000

Average income:
$50,000



Scale effect: tendency for different results to be obtained from the same set of data when that data is grouped at different levels of spatial resolution

MAUP

1. Are the results independent of the units being used, or do they depend entirely on the choice of spatial units used in the analyses?

2. How can we compare the results of studies from one region to another, when simply changing the spatial units used in both studies could dramatically alter the conclusions reached?

Other local designations: Place

- United States Census Bureau defines a **place** as a concentration of population which has a name, is locally recognized, and is not part of any other place.

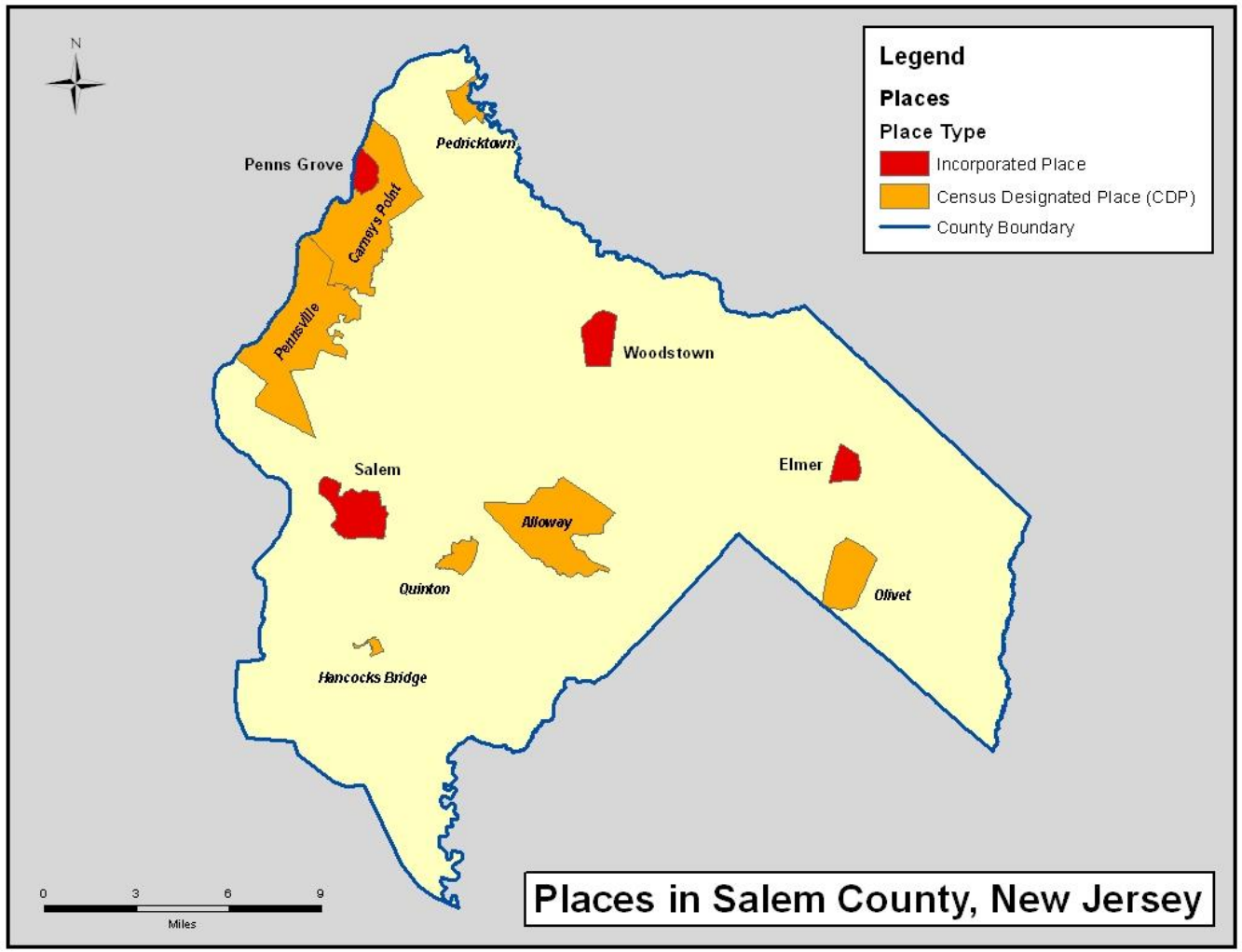
Incorporated Places:

- Legally bounded entity
- Referred to as cities, boroughs, or villages, depending on the state and having legally prescribed limits, powers, and functions.

Census Designated Places (CDPs):

- Statistical entity, Boundaries have no legal status
- Area with a concentration of population, housing, and commercial structures identifiable by name;
- Unincorporated counterparts

Places - Salem County, New Jersey



County Subdivisions

- Minor Civil Divisions (MCDs)
  - Legally bounded entity: townships, towns, etc.
  - Sub-county entities in 29 states, the District of Columbia, and Puerto Rico
  - May have a local government with elected officials

- Census County Divisions (CCDs)
  - Statistical entity
  - Sub-county units that have stable boundaries and recognizable names in 21 states
  - No minimum or maximum population guidelines

Minor Civil Divisions (MCDs):

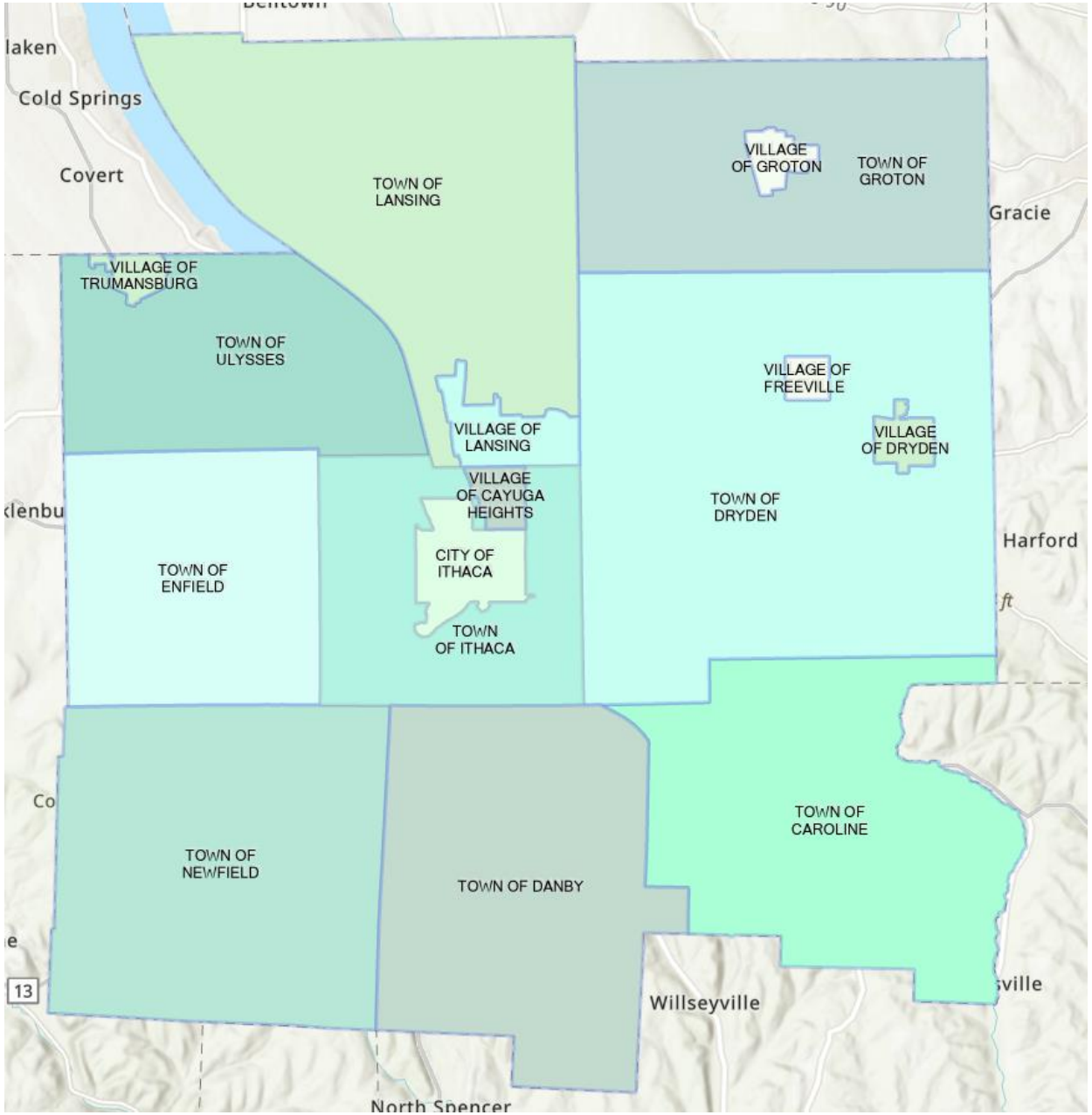
Do NOT have to follow minimum or maximum population guidelines (Townships, towns). May have a formal government with elected officials

Area Municipal Governments

Local Municipal Governments

Located within Tompkins County are the City of Ithaca; the Towns of Ithaca, Caroline, Danby, Dryden, Enfield, Groton, Lansing, Newfield, and Ulysses; and the Villages of Cayuga Heights, Dryden, Freeville, Groton, Lansing, and Trumansburg.

- City of Ithaca
- Town of Ithaca
- Village of Cayuga Heights
- Town of Caroline
- Town of Danby
- Town of Dryden
- Village of Dryden
- Village of Freeville
- Town of Enfield
- Town of Groton
- Village of Groton
- Town of Lansing
- Village of Lansing
- Town of Newfield
- Town of Ulysses
- Village of Trumansburg

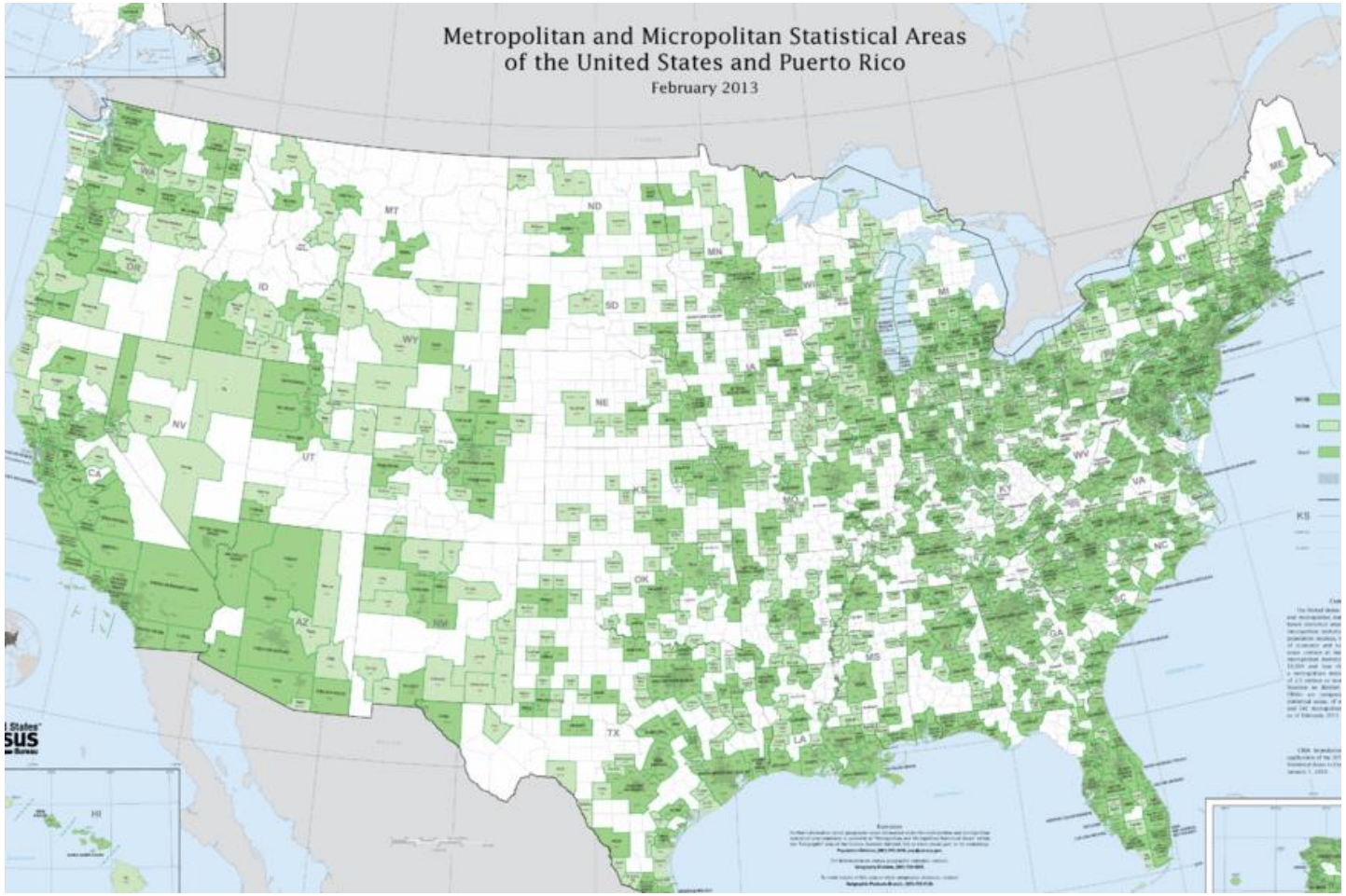


Core based Statistical Area (CBSA)

- consists of one or more counties (or equivalents) anchored by an urban center plus adjacent counties that are socioeconomically tied to the urban center by commuting

- *Metropolitan Statistical Area (MSAs):* urban core population of 50,000+ (at least one county)
- *Micropolitan Statistical Area:* urban core population of 10,000-50,000. typically consists of one county.

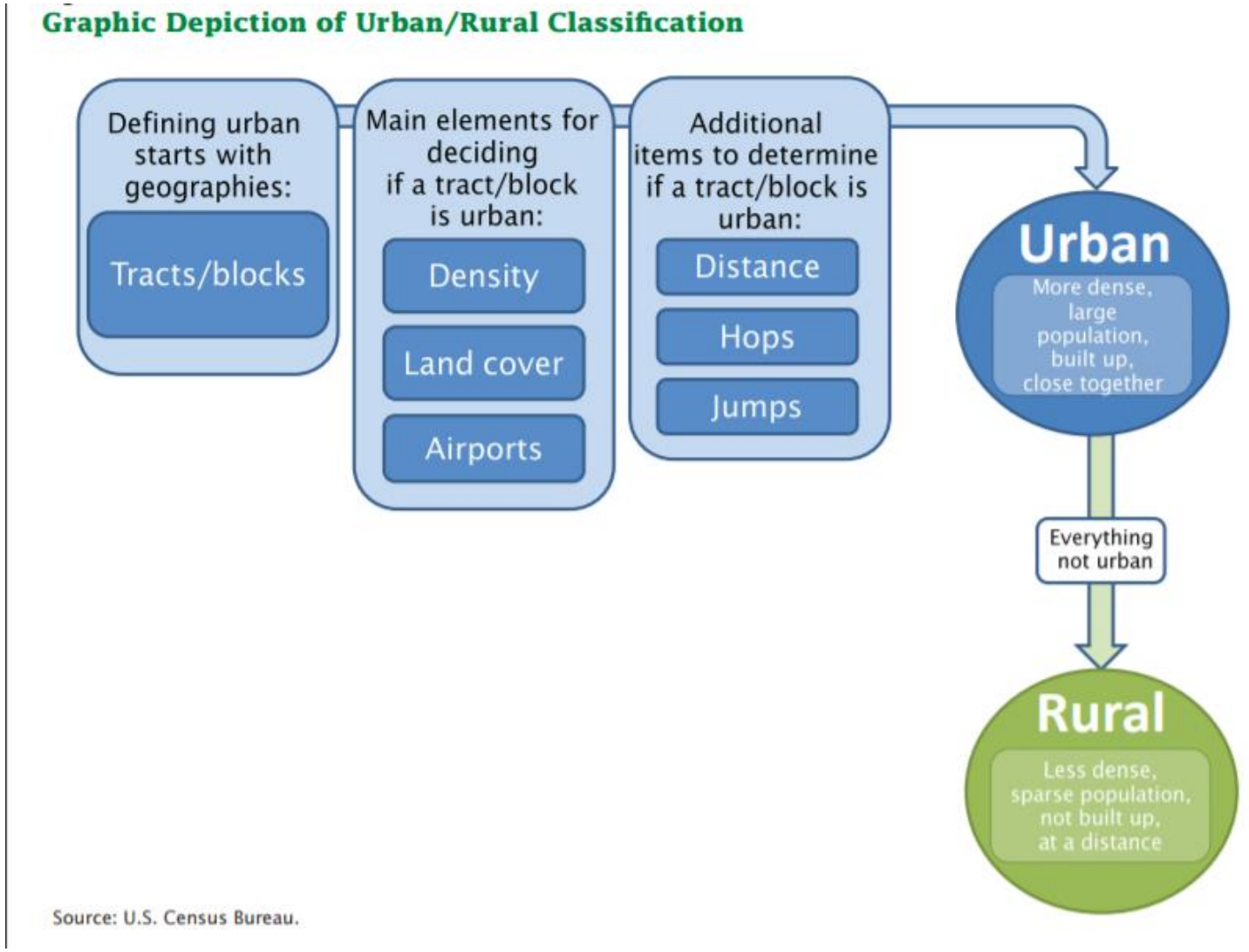
- 935 core-based statistical areas
- 393 metropolitan statistical areas
- 542 micropolitan statistical areas



Urbanized Areas

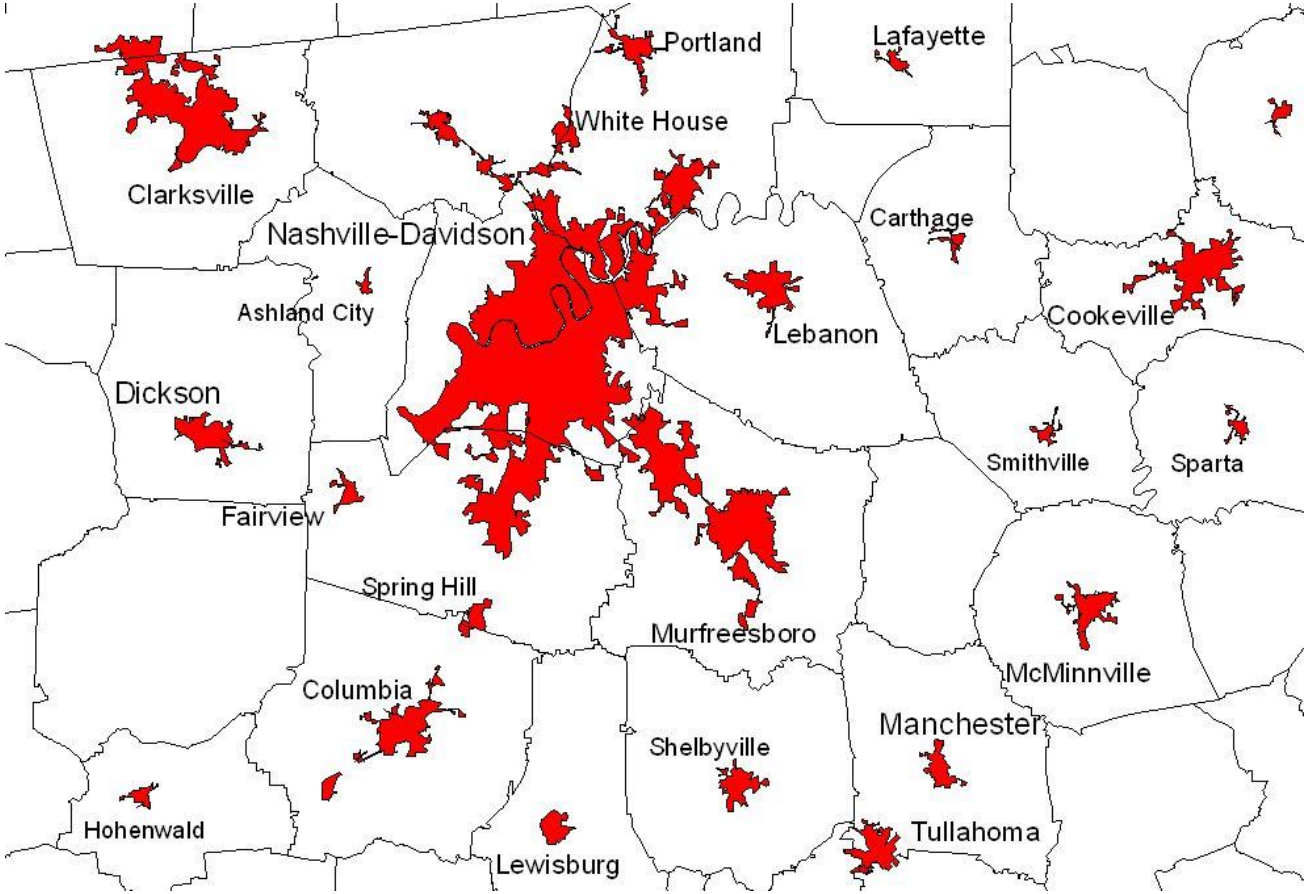
Two types of urban areas: urbanized areas and urban clusters.

- Urbanized areas are areas with 50,000 or more people.
- Urban clusters are areas with at least 2,500 but fewer than 50,000



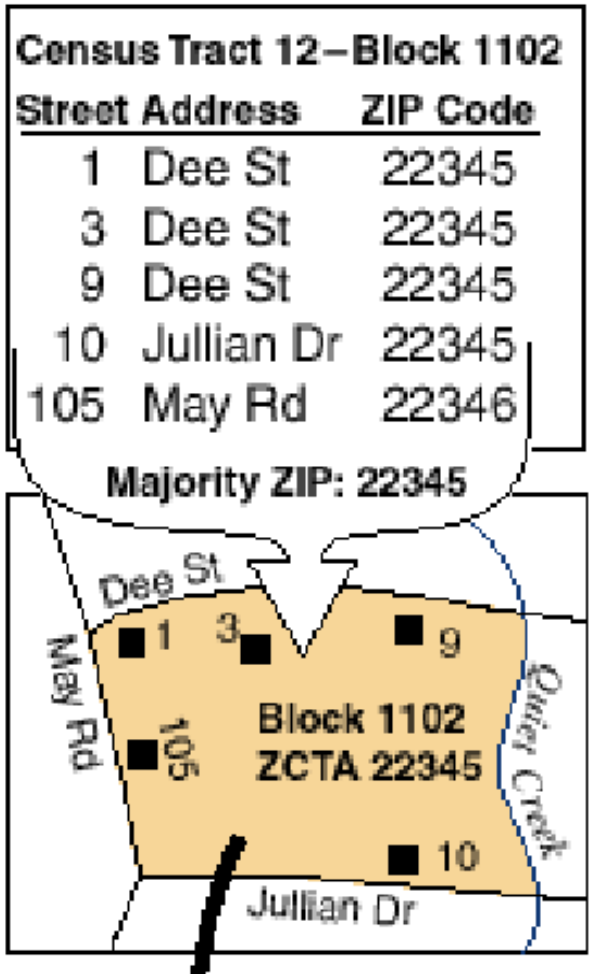
Urbanized Areas

- defined at census block and block group levels, starting with a core area with a density of at least 1,000 people per square mile.
- Geographic entities (census tracts, counties, metropolitan areas) often "split" between urban and rural territory
- “rural" consists of everything else.



ZIP Code Tabulation Areas (ZCTAs)

- Approximate area representations of USPS ZIP Code service areas based on Census blocks
- To address difficulties in mapping USPS ZIP Codes



Finding Census Data

To undertake census data analysis and management, we need to do three things:

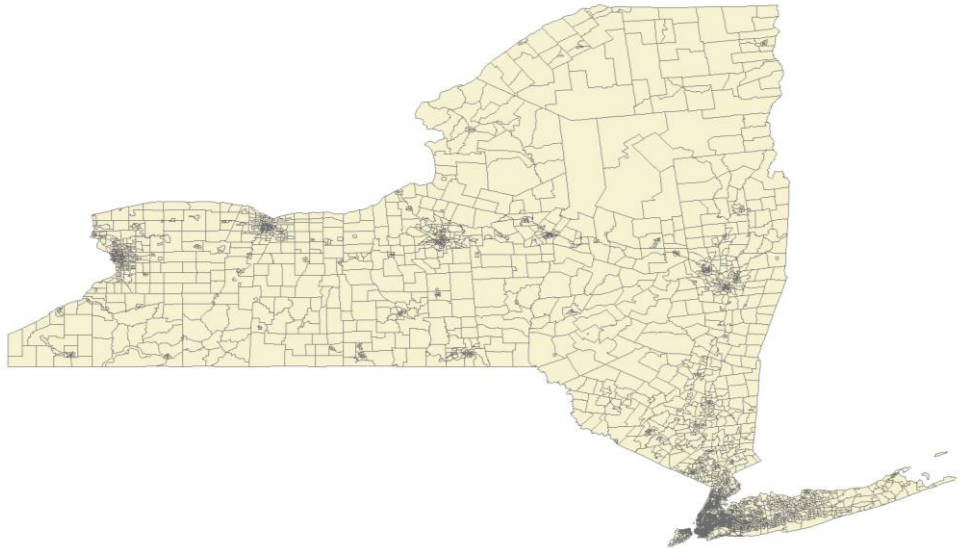
- A source for the boundary files (geographic data)
- A source for the census (attribute) data
- A unique identifier to join the non-spatial table with the spatial boundary

ny_census_tracts

FID	Shape	AREA	PERIMETER	TR36_D00_	TR36_D00_I	STATE	COUNTY	TRACT	NAME	LSAD	LSAD_TRANS
0	Polygon	0.00802	0.45858	2	1	36	089	9901	9901	TR	
1	Polygon	0.01738	0.5601	3	2	36	019	1001	1001	TR	
2	Polygon	0.05585	1.01048	4	3	36	019	1003	1003	TR	
3	Polygon	0.00617	0.42726	5	4	36	033	9913	9913	TR	
4	Polygon	0.07086	1.22731	6	5	36	019	1004	1004	TR	
5	Polygon	0.03171	0.95298	7	6	36	033	9502	9502	TR	
6	Polygon	0.03749	0.90102	8	7	36	033	9501	9501	TR	
7	Polygon	0.00749	0.45325	9	8	36	089	9903	9903	TR	
8	Polygon	0.01883	0.62928	10	9	36	089	9904	9904	TR	
9	Polygon	0.00099	0.174	11	10	36	089	9902	9902	TR	
10	Polygon	0.05491	1.25146	12	11	36	089	9907	9907	TR	
11	Polygon	0.01806	0.61533	13	12	36	019	1002	1002	TR	
12	Polygon	0.01701	0.59964	14	13	36	089	9906	9906	TR	
13	Polygon	0.03294	0.74008	15	14	36	089	9905	9905	TR	
14	Polygon	0.01669	0.67767	16	15	36	033	9504	9504	TR	
15	Polygon	0.01359	0.61199	17	16	36	033	9505	9505	TR	
16	Polygon	0.03818	0.91312	18	17	36	033	9503	9503	TR	
17	Polygon	0.07209	1.44425	19	18	36	033	9506	9506	TR	
18	Polygon	0.02048	0.7715	20	19	36	019	1006	1006	TR	
19	Polygon	0.03336	0.7603	21	20	36	089	9913	9913	TR	
20	Polygon	0.06886	1.35987	22	21	36	089	9924	9924	TR	
21	Polygon	0.10628	1.61375	23	22	36	033	9507	9507	TR	
22	Polygon	0.01752	0.62327	24	23	36	089	9910	9910	TR	
23	Polygon	0.00868	0.58524	25	24	36	019	1008	1008	TR	

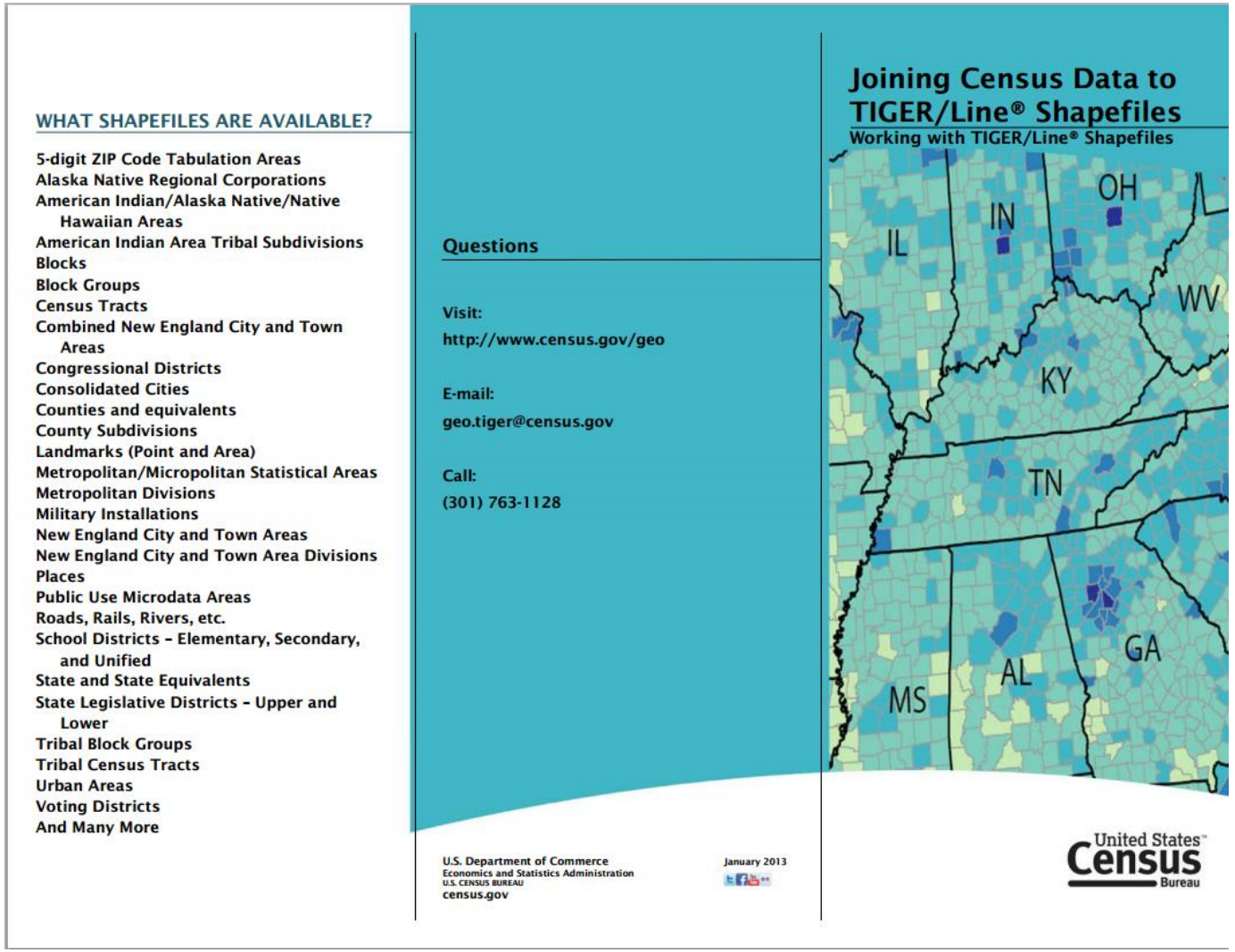
1 (0 out of 4908 Selected)

ny_census_tracts



Geographic Products (boundary file): The TIGER Data Base

- Large database of geographic features, such as:
  - School districts
  - Railroads
  - Rivers
  - Landmarks
  - Legal boundaries
  - Census statistical boundaries
- Cheap, widely used, updated by the Census



WHAT SHAPEFILES ARE AVAILABLE?

5-digit ZIP Code Tabulation Areas
Alaska Native Regional Corporations
American Indian/Alaska Native/Native Hawaiian Areas
American Indian Area Tribal Subdivisions
Blocks
Block Groups
Census Tracts
Combined New England City and Town Areas
Congressional Districts
Consolidated Cities
Counties and equivalents
County Subdivisions
Landmarks (Point and Area)
Metropolitan/Micropolitan Statistical Areas
Metropolitan Divisions
Military Installations
New England City and Town Areas
New England City and Town Area Divisions
Places
Public Use Microdata Areas
Roads, Rails, Rivers, etc.
School Districts – Elementary, Secondary, and Unified
State and State Equivalents
State Legislative Districts – Upper and Lower
Tribal Block Groups
Tribal Census Tracts
Urban Areas
Voting Districts
And Many More

Questions

Visit:
http://www.census.gov/geo

E-mail:
geo.tiger@census.gov

Call:
(301) 763-1128

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

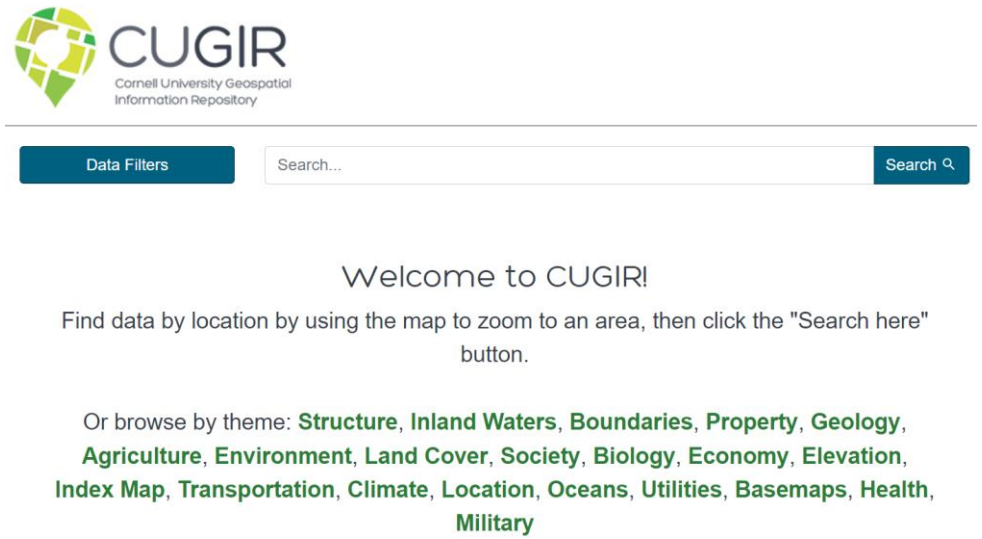
January 2013

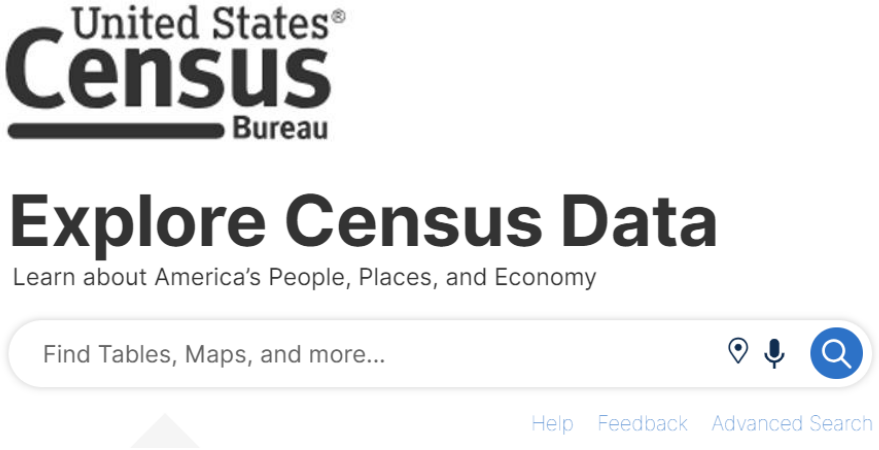
Joining Census Data to TIGER/Line® Shapefiles
Working with TIGER/Line® Shapefiles

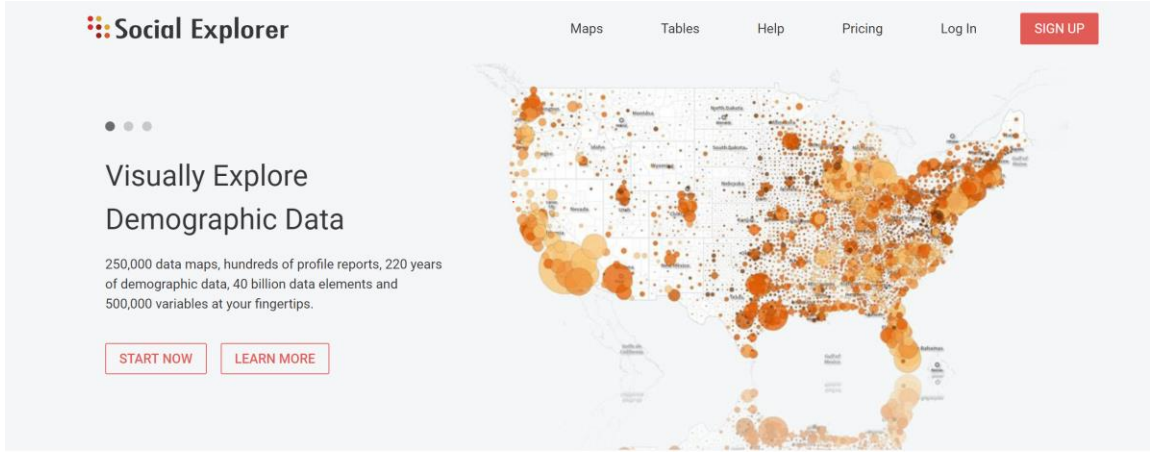
United States Census Bureau

We will explore 3 sources of census data:

- CUGIR: Cornell University Geospatial Information Repository; a state geospatial repository
- US Census: the Federal census site
- Social Explorer: a private vendor







American Community Survey (ACS)

1 year: 12 months of collected data. *Example:* 2019 ACS 1-year estimates. *Date collected between:* January 1, 2019 and December 31, 2019

- provides annual estimates for all states, cities, counties, metropolitan areas, and population groups of 65,000 people or more.
- Most current, less reliable
- Annually released: 2005-present

5 year: 60 months of collected data. *Example:* 2015-2019 ACS 5-year estimates. *Date collected between:* January 1, 2015 and December 31, 2019

- Data for all areas; but no data at the block level
- Most reliable, least current
- Annually released 2009- present

ACS estimates are less precise than the comparable estimates from Census decennial census years.

Public Use Microdata Sample

- Public Use Microdata Sample (PUMS) files enable data users to create custom estimates and tables, free of charge, that are not available through ACS pre-tabulated data products. The ACS PUMS files are a set of records from ***individual people or housing units***, with disclosure protection enabled

- They are NOT Summary data (ie counts, averages and medians for specific geographic areas)

- Public Use Microdata Areas (PUMAs) are non-overlapping, statistical geographic areas that partition each state or equivalent entity into geographic areas containing no fewer than 100,000 people each.

Joining attribute data and boundary files

- We join a shapefile and an attribute table through a unique identifier field that is common to both datasets

- GEOID: a concatenated field containing State, County, Tract, and Block numbers (block group and block).

  For example; 361090001001001

  36 109 000100 1 001

  - 36 = New York State
  - 109 = Tompkins County
  - 000100 = Tract Number
  - 1001 = Block Number
    - 1 = Block group
    - 001 = Block

Other sources of geographic boundary data:

Municipalities may have:

- Boroughs
- Police precincts
- City council districts
- Health Areas
- Historic districts