

Secondary data

Secondary data

- Any data collected for another purpose
 - Data may be reclassified or other manipulated before becoming available
- Normally collected by a government
 - Private firm secondary data likely propriety or confidential

Secondary data

- If from a government source, data collection is typically extensive and robust

The screenshot shows a document from the U.S. Census Bureau. At the top left is the 'U.S. Census Bureau' logo and the '2000 Census' logo. The main heading is 'Evaluations'. Below this, there are several sections of text, including 'Census 2000 Testing, Experimentation, and Evaluation Program Summary Results' and 'Census 2000 Topic Reports and Evaluation Reports'. The text is small and dense, but it appears to be a detailed report on the evaluation of the 2000 Census process.

Secondary data

- Because of assured confidentiality, data may already be grouped into units
 - May not be able to see the raw data
 - Ecological fallacy – using data to infer that an entire sample/population has that trait
 - Two variables combined show a population characteristic that assumes too much

Uses of secondary data - humans

- Provide demographic data
 - Demographics: population characteristics
 - Typically includes gender, race, age, income, home ownership, education level
 - Typically does not include political views or any opinions

Types of secondary data

- Human – population data
 - Census – a systematic survey of the characteristics of a country's population
 - Most countries administer them
 - UN recommends every 10 years
 - Used in India, Rome, and Egypt 2000+ years ago
 - US: mandated by the Constitution (Article 1) to allocate Congressional seats

Types of secondary data - Census

- US Census: given every 10 years – most years it is a sample
- Full population count every 100 years
- Surveys always occurring with US
 - Population estimates and projections
- Data used to determine funding allocation
 - Affects education units, public health, public safety, transportation, etc...

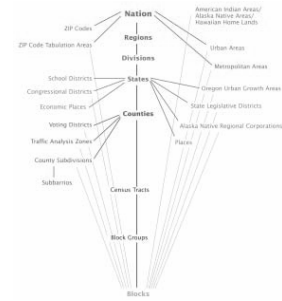
Types of secondary data - Census

- Census data: one of the most used sets of geographical data
- Each household is asked demographic information
 - Individual information is kept confidential for 72 years
 - Data is grouped into different spatial levels

Types of secondary data - Census

- Regions: groups of divisions (4 regions)
- Divisions: group of states (9 divisions)
- States: includes DC and Virgin Islands
- Counties: includes other designations like boroughs and parishes
- Tracts
- Groups
- Blocks

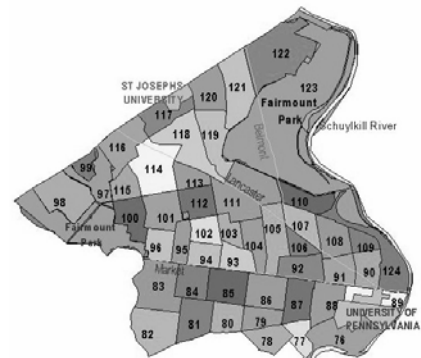
Types of secondary data - Census



Census tracts

- A small, relatively permanent statistical subdivision of a county or statistically equivalent entity.
- Generally contain between 1,000 and 8,000 people.
- Intended to be stable over many decades, so they generally follow relatively permanent visible features. However, they may follow governmental unit boundaries and other invisible features in some instances; the boundary of a state or county is always a census tract boundary.

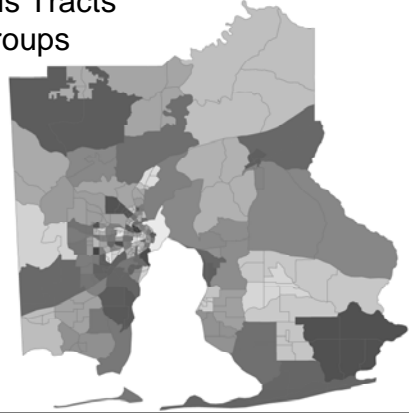
Tracts



Block group

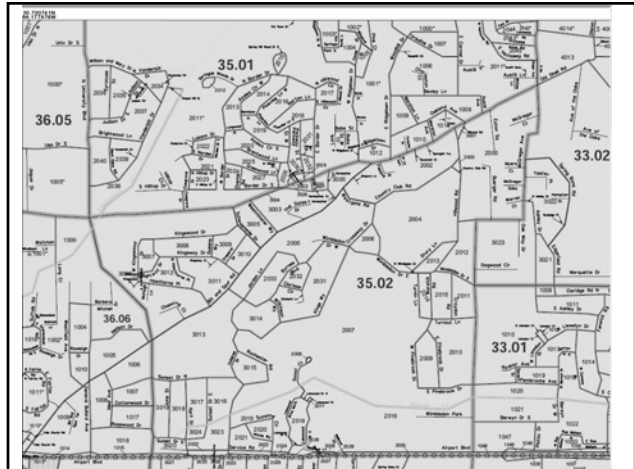
- A group of census blocks.
- Generally contain between 300 and 3,000 people, with an optimum size of 1,500 people.
- May be the smallest set of data available for a region.

Mobile & Baldwin County Census Tracts & Block Groups



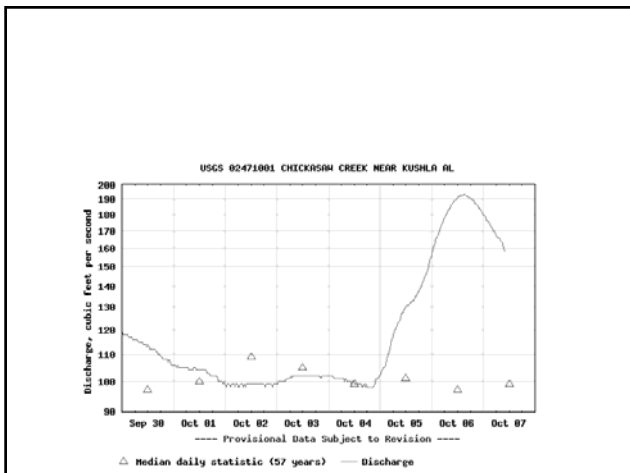
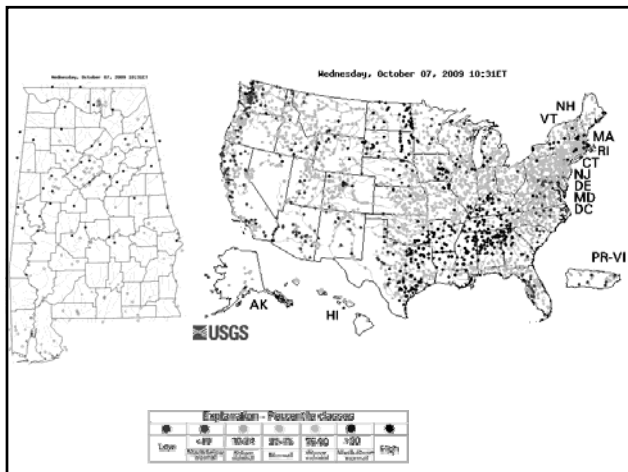
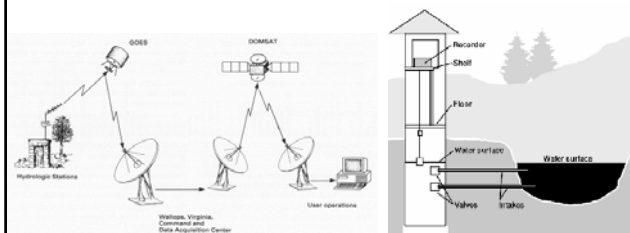
Census Blocks

- An area bounded on all sides by visible and/or nonvisible features shown on a map prepared by the Census Bureau.
- A block is **the smallest** geographic entity for which the Census Bureau tabulates decennial census data.



Gaging Station

- Secure location along a water body that automatically records discharge
- Usually not on rivers with tidal flow



Date / Time	Gage height, feet	Dis-charge, ft ³ /s
09/30/2009 00:00	3.43	120
09/30/2009 00:30	3.41	118
09/30/2009 01:00	3.41	118
09/30/2009 01:30	3.41	118
09/30/2009 02:00	3.41	118
09/30/2009 02:30	3.41	118
09/30/2009 03:00	3.40	117
09/30/2009 03:30	3.40	117
09/30/2009 04:00	3.40	117
09/30/2009 04:30	3.40	117
09/30/2009 05:00	3.39	116
09/30/2009 05:30	3.39	116
09/30/2009 06:00	3.38	116
09/30/2009 06:30	3.38	116
09/30/2009 07:00	3.38	116
09/30/2009 07:30	3.37	115
09/30/2009 08:00	3.37	115
09/30/2009 08:30	3.37	115
09/30/2009 09:00	3.37	115
09/30/2009 09:30	3.37	115
09/30/2009 10:00	3.36	114
09/30/2009 10:30	3.36	114
09/30/2009 11:00	3.36	114
09/30/2009 11:30	3.35	113
09/30/2009 12:00	3.36	114
09/30/2009 12:30	3.35	113
09/30/2009 13:00	3.34	112

Water data

- Uses: drought, changes in land surfaces, floods
- In some cases, there is a long record of data
 - Data collection can be discontinued with little notice
- Disadvantages: not there to actually collect data