



Free Data Sources

CAS Ratemaking Seminar
March, 2008

Mark Florenz, ACAS, MAAA

Contents

Geography of the Census

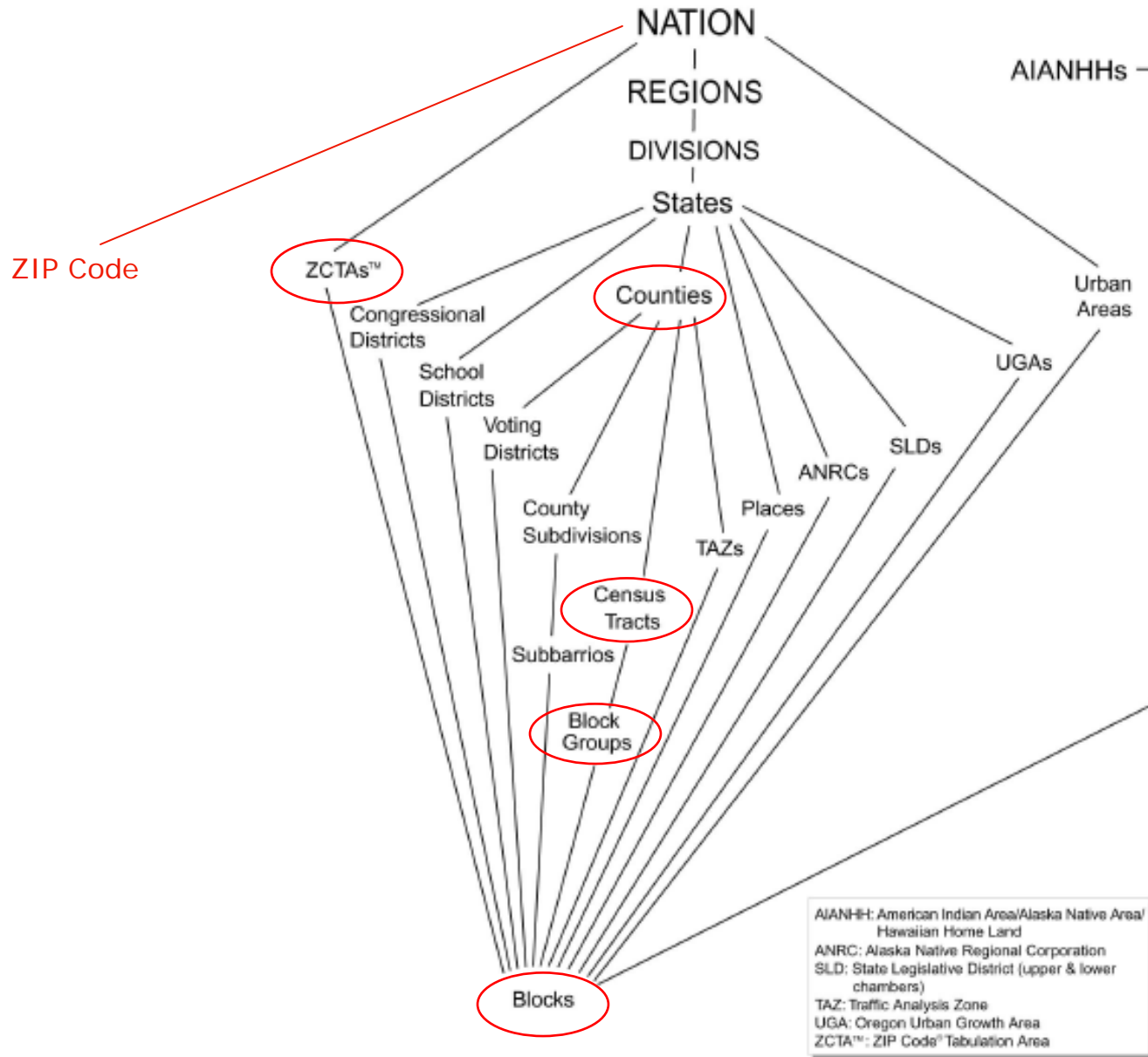
ZIP Codes

Matching Census & Policy Data

Appendix

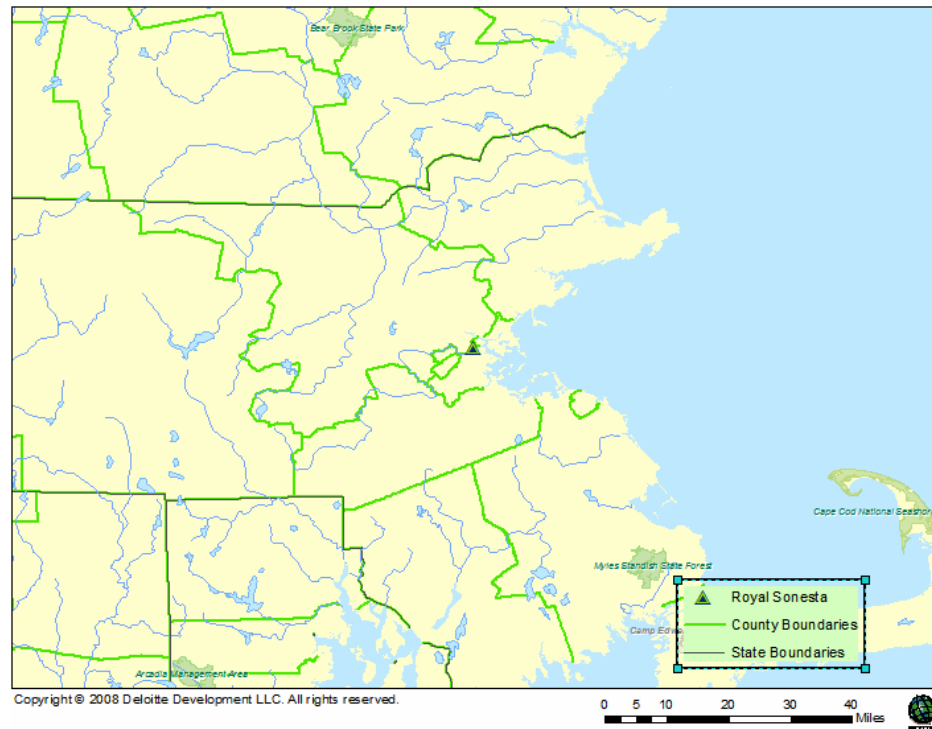
Geography of the Census

Geographic Hierarchy



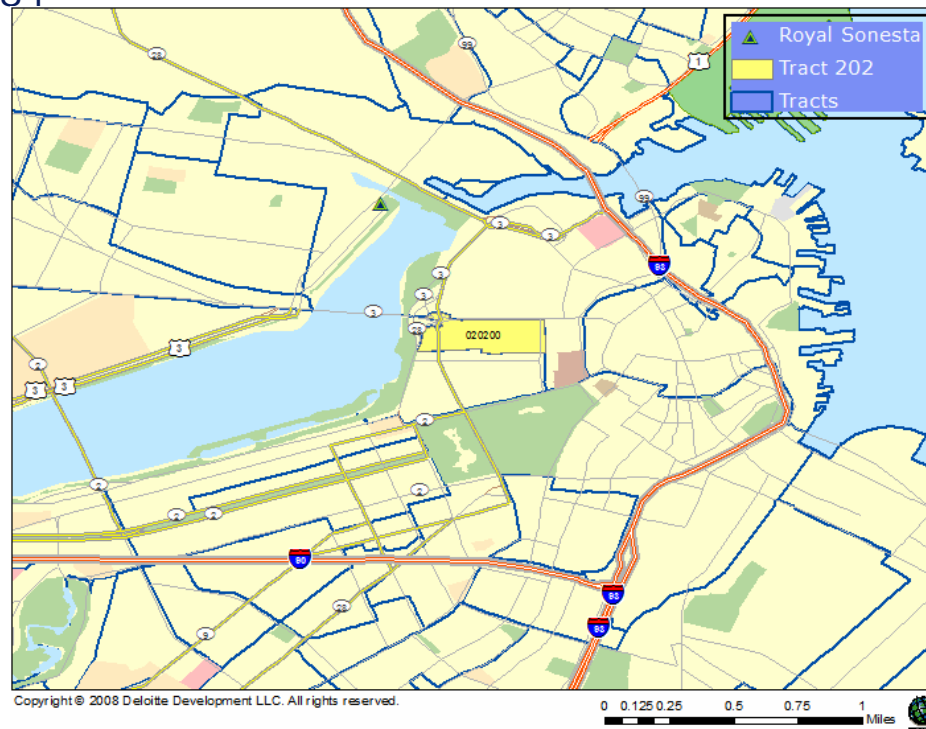
Counties (or the statistical equivalent)

- Generally stable boundaries, primary subdivision of most states
- Established by state law or other administrative action
- Counties in 48 States; Independent Cities in 4 States; Parishes in Louisiana; Cities and Boroughs in Alaska
- Suffolk County
 - Population = 689,807
 - Median Family Income = 44,361
 - Percent Irish = 16%
 - Percent Italian = 11%



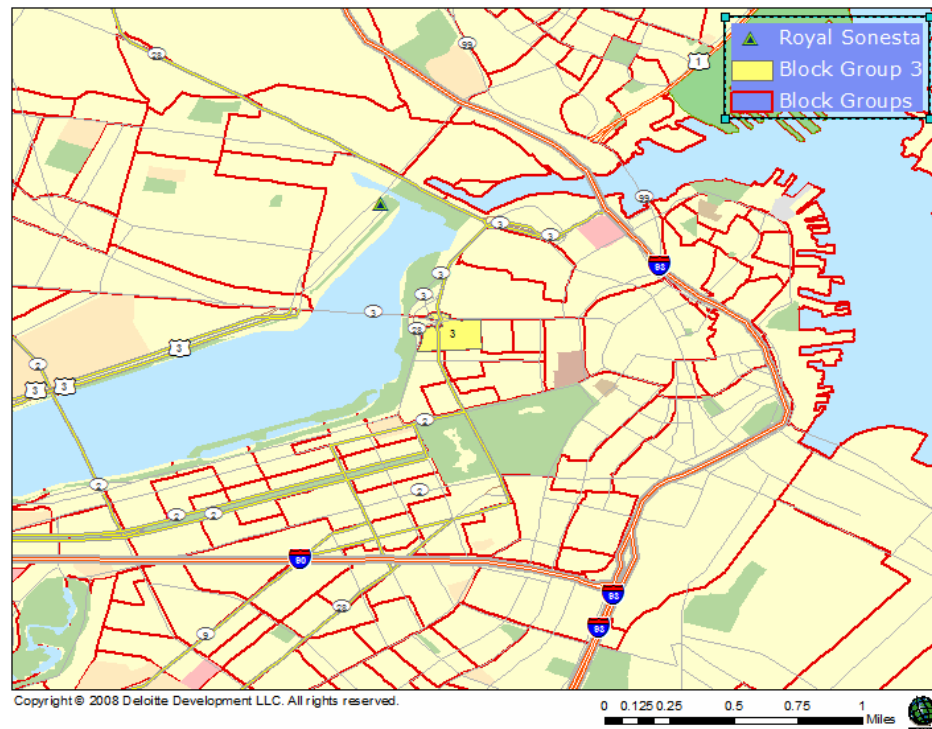
Census Tracts

- Census tracts are small, relatively permanent statistical subdivisions of a county or statistical equivalent
- Designed to have homogenous population characteristics, economic status, and living conditions
- Size: optimally 4,000 people, range between 1,000 and 8,000
- Census Tract 202
 - Population = 3,635
 - Median Family Income = 101,784
 - Percent Irish = 17%



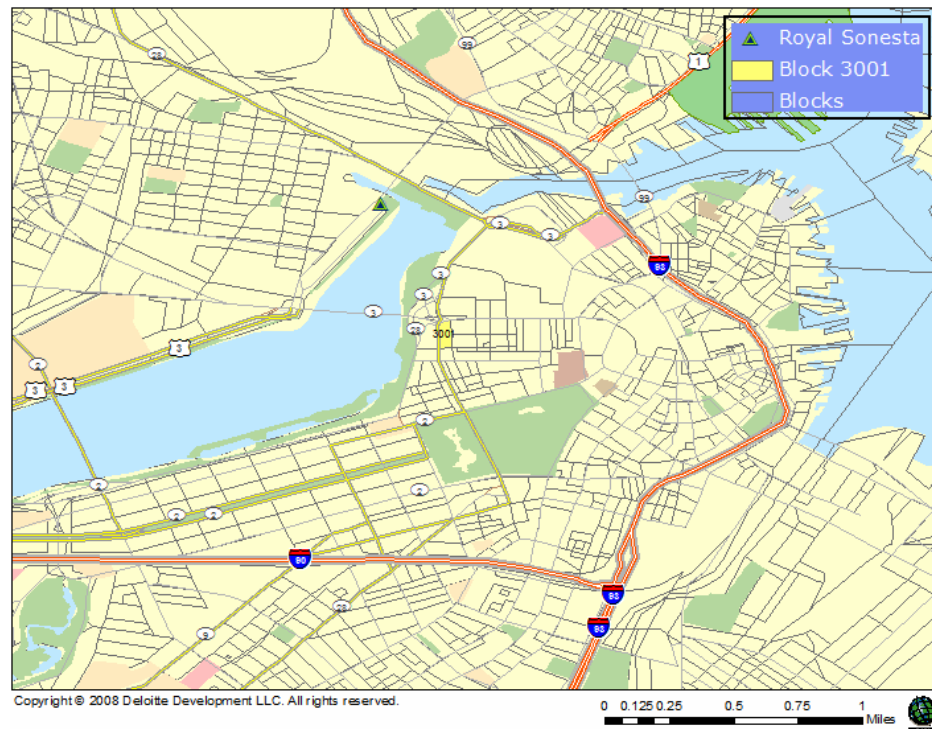
Block Groups

- Cluster of census blocks sharing the same first digit
- Smallest areas for which sample data available
- Most block groups delineated by local participants
- Size: optimally 1,500 people, range between 300 to 3,000
- Block Group 3
 - Population = 1,155
 - Median Family Income = 140,933



Blocks

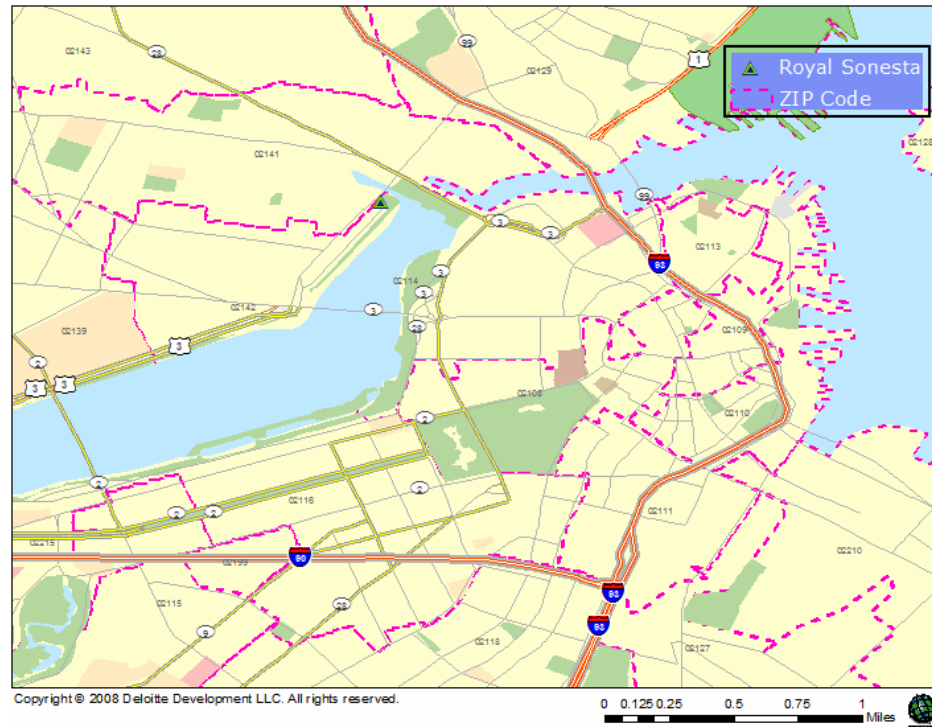
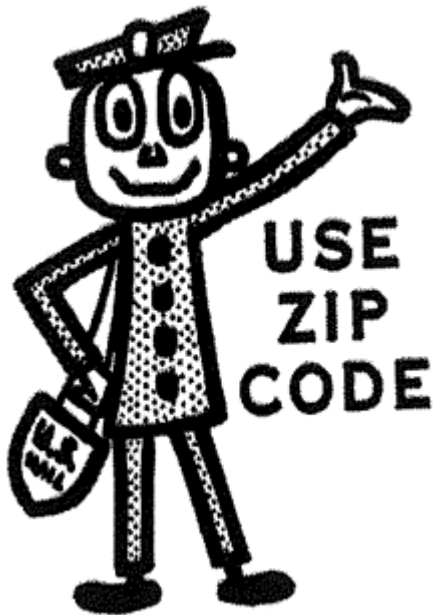
- Smallest units of data tabulation, cover the entire nation
- Do not cross census tracts or counties
- Generally bounded by visible features and legal boundaries
- Size: average about 100 people
- Block 3006
 - Population = 458



ZIP Codes

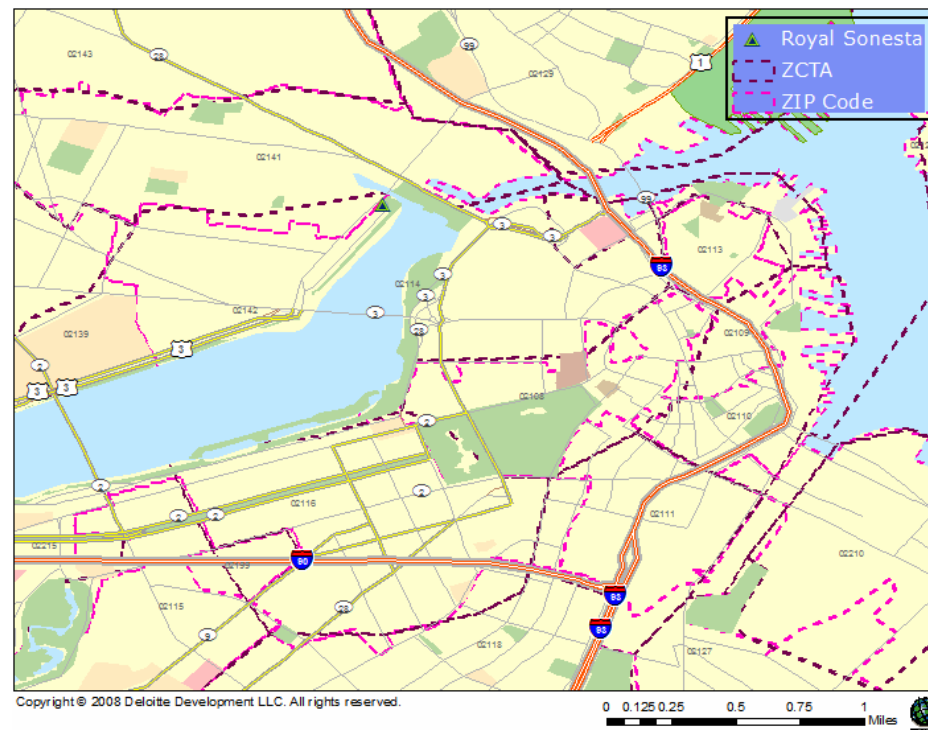
ZIP Codes

- Category for grouping mailing addresses
- First digit represents a group of states, digits 2-3 represent a region in that group, digits 4-5 represent a group of delivery addresses
- Not spatially designed (i.e. there are no ZIP code boundaries)
- Cross state lines, county lines, census tracts,...
- ZIP code definitions change over time, names do not
- ZIP = Zone Improvement Plan

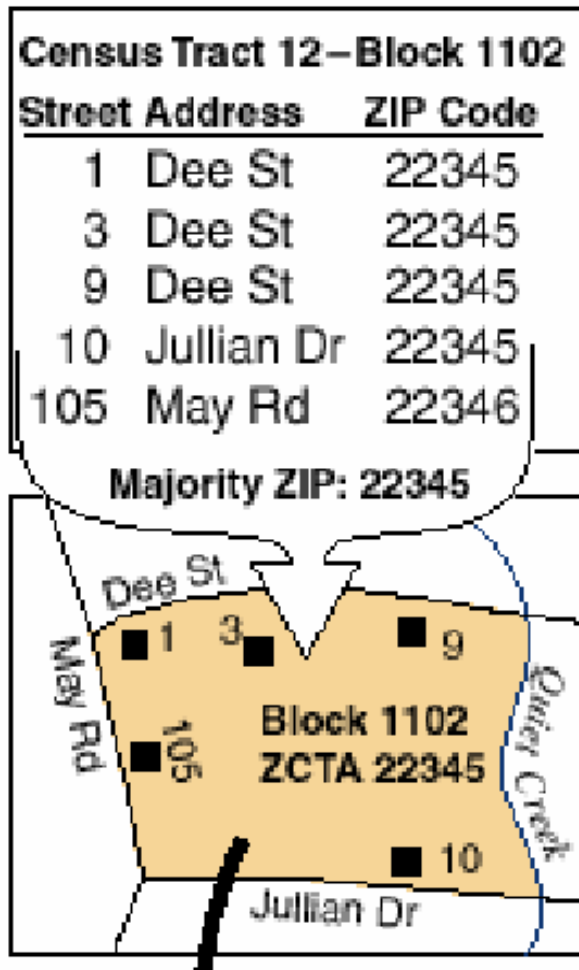


ZIP Code Tabulation Areas (ZCTAs)

- Approximate area representations of USPS ZIP Code service areas
- ZCTAs are aggregations of census blocks that have the same predominant ZIP Code
- ZCTAs do not precisely depict ZIP Code delivery areas, and do not include all ZIP Codes used for mail delivery.
- ZCTA replaces ZIP Code tabulation in 2000 census



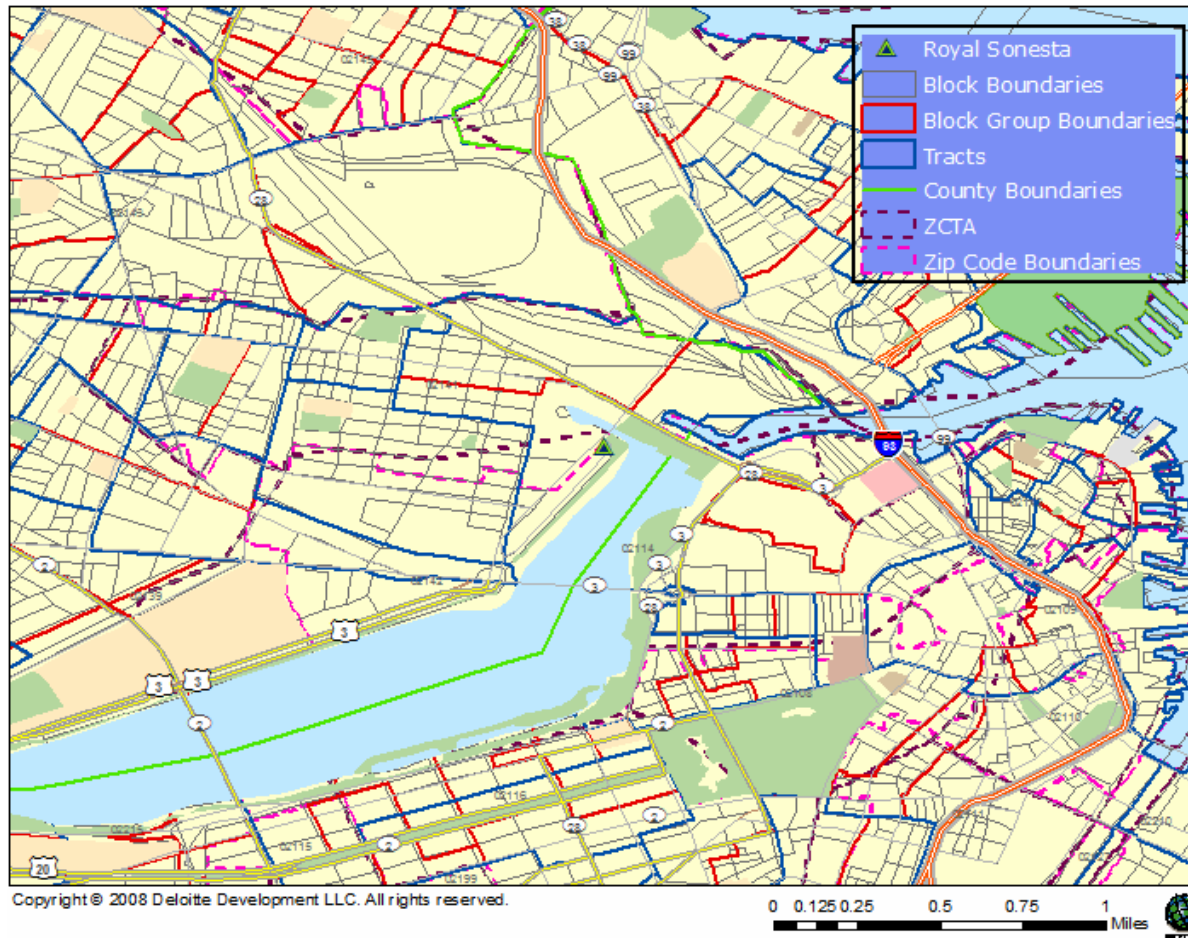
How ZIP Code Tabulation Areas Are Created



Determine the majority ZIP Code for each census block with addresses.

Recap

- Blocks: Smallest Area of Aggregation
- Block Groups: Most Census data released starting at this level
- Census Tracts: Roughly 4,000 people in each
- ZIP Codes: Postal Delivery areas



Matching Census & Policy Data

Geographic Codes

Types of Codes

- FIPS - Federal Information Processing Standards codes are assigned by the National Institute of Standards and Technology (NIST) for a variety of geographic entities, notable county and state.
- Census - are assigned by the U.S. Census Bureau for a variety of geographic entities
- USPS – are assigned by the United States Postal Service for states and mail delivery routes

Commonly Used Codes:

- State – FIPS code: two-digit number; or USPS abbreviation: two-digit character
- County – FIPS code: three-digit number
- Census Tract – Census code: four-digit basic number which may have a two-digit suffix (e.g. 6059.02)
- Block Group – Census code: one-digit number
- Block – Census code: four-digit number
- ZIP Code – USPS: five-digit number which may have a four-digit suffix
- ZCTA Code – Census code: five-digit number (three-digits in rare instances)

Method 1 – Higher Cost, Higher Precision

Geo-code policy address and match external data at the lowest level

Advantages:

- Uses demographically similar population areas
- More stable and closer to individual level

Disadvantages:

- Geo-coding process has cost in terms complexity and \$\$\$
- Detailed and accurate addresses not always available for geo-coding process

Method 2 – Lower Cost, Lower Precision

Translate external data sources to ZIP code level

Advantages:

- ZIP code is generally available in policyholder data
- ZIP code is generally familiar to all stakeholders
- Mappings to ZIP code can be found for little to no money

Disadvantages:

- Lack of precision in mappings
- Lack of consistency over time

Appendix

Further Reading

Reference Resources for Understanding Census Bureau Geography

- <http://www.census.gov/geo/www/reference.html>

The ZIP Code Resources Page:

- <http://mcdc2.missouri.edu/webrepts/geography/ZIP.resources.html>

Federal Information Processing Standards (FIPS) codes

- <http://www.census.gov/geo/www/fips/fips.html>

American Factfinder

- http://factfinder.census.gov/home/saff/main.html?_lang=en