OnTheMap: 10 Years of Distributing Data in a Web GIS

Matthew Graham

Product Development & Quality Assurance Branch

LEHD Program, Center for Economics Studies

U.S. Census Bureau

November 18, 2016

Desiderata

- Disclaimer: Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

Background

- LEHD Program uses admin. and survey data to produce public-use data and tools focused on employment and the labor market.

- Data include Quarterly Workforce Indicators (QWI) and LEHD Origin-Destination Employment Statistics (LODES), and Job-to-Job Flows (J2J).

- Tools include OnTheMap, OnTheMap for Emergency Management, and QWI Explorer.

- More info: http://lehd.ces.census.gov/

What We Aim To Do

- Grow the user base for the data
  - Tools should work well and be as simple as possible.

- Expand usage within the user base
  - Focus on *analyses* of data rather than raw data tables.

- Increase understanding of the data
  - Provide multiple levels of information/learning.

LODES Data Description

- Origin-Destination data on jobs
  - 130m jobs, annual release (13 years of data)
  - Geography
    - Home & Work Census Blocks (11m)
  - Firm Characteristics
    - Ownership (3), NAICS Sector (20), Firm Age (5), Firm Size (5)
  - Person/Job Characteristics
    - Age (3), Earnings (3), Race (6), Ethnicity (2), Sex (2), Education (4)
  - **Large & Sparse (1.3×10^8 / 1.6×10^{20})**

Application Challenges

- Database
  - OD nature requires double-indexing
  - No transactional requirements: "Write once, read a million times" led to indexed flat file solution
  - Key tradeoff: Size → Speed (denormalization)
- Flexible, block-based aggregation to larger areas
  - No split blocks
  - Block approximations based on internal points
  - Support automated tabulation for 18 higher level geographies (Places, County Subs, ZCTAs, School & Legislative Districts, etc.) – "Block Crosswalk"

User Interface

- Basic GIS tools for flexible selection
  - Select & buffer existing boundary
  - Freehand points, lines, polygons and select or buffer
  - Import shapefiles & KMLs (not **too** complicated)
  - Unions, intersections, subtractions
- Other considerations
  - Easy to navigate and clean GUI, but feature-full
  - Search- *and* map-based
  - Presents statistics and geography as integrated analyses
- Fast. Fast. **Fast.**

OnTheMap Live Examples

- http://onthemap.ces.census.gov

Technology Stack

- OS: Red Hat Linux

- Storage: PostgreSQL & Indexed Flat Files

- Processing: Python

- Display: Javascript

- Mapping: Mapserver

- Geography: TIGER + OSM Streets

Thank You!

- Longitudinal Employer-Household Dynamics
  - http://lehd.ces.census.gov

- OnTheMap
  - http://onthemap.ces.census.gov

- Questions and Feedback
  - CES.OnTheMap.Feedback@census.gov
  - matthew.graham@census.gov

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