

Innovations in Integrating Data at the U.S. Census Bureau

**ANCSDAAP Population Census Conference
March 2009**

**Nancy M. Gordon
Associate Director for Strategic
Planning and Innovation
U.S. Census Bureau**

U S C E N S U S B U R E A U

Dissemination Challenges

Present information that meets users' needs because it is:

- Relevant
- Timely
- Accurate
- Accessible

Leverage existing sources of information to create new and better products at modest cost

Innovations at the U.S. Census Bureau: Integrating Data

TheDataWeb: Software engineering that enables analysts to access, use, and integrate data from several sources at the same time

DataFerrett: Analytic tools used to create graphs, maps, and tables based on data in TheDataWeb

HotReports: Web sites that present information to non-technical decision makers using Internet Web 2.0 interactive presentation technology

HotReports are created by using DataFerrett tools operating on data files in TheDataWeb

TheDataWeb: Infrastructure for Integrating Data from Multiple Sources

Open source system – handles different kinds of data and datasets in the same framework

Stores data one time, uses them many times

Empowered by statistical intelligence

- Documentation, statistical usage rules, data integration rules

Software infrastructure that enables DataFerrett and HotReports to work

DataFerrett: Tools for Integrating Data

Targets sophisticated analysts who can access and integrate the public-use files in TheDataWeb

A group of advanced Web-based analytic tools for sophisticated users

- Data manipulation, including advanced tabulations
- Mapping and business graphics using statistical rules
- Statistical package, including regressions, later this year

HotReports: Presentation Technology for TheDataWeb

Targets decision makers who have limited time
and statistical background

- Focuses on one topic using information from many data sources
 - Such as economic development or emergency response
- Brings together relevant variables for local areas
 - Maps, charts, simple tables, and short text boxes
- Supports decision making through guided use of statistical data

Change County:

[Select Another State](#)

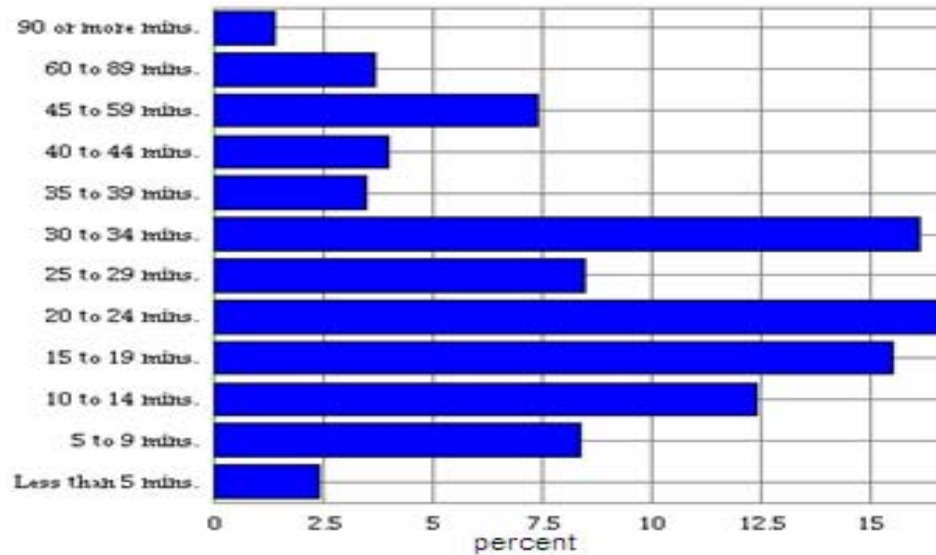
[Printer Friendly](#)

Wayne County, Michigan

[Summary](#) [Overview](#) [Economics](#) [Demographics](#) [Housing](#) [Transportation](#) [Community](#)

Transportation Overview

Percent of Commuters by Travel Time to Work



Wayne County, Michigan

**Data Source: 2005 American Community Survey

Commuting Time

747,122 workers commuted to jobs in **Wayne County, Michigan**, taking on average **24.7** minutes each way.

**Data Source: 2005 American Community Survey
[information on data source, confidentiality, protection, sampling error, nonsampling error, and definitions.](#)



[Access 'On The Map'](#) to explore where workers live in relationship to where they work.

Innovations at the U.S. Census Bureau: Local Employment Dynamics

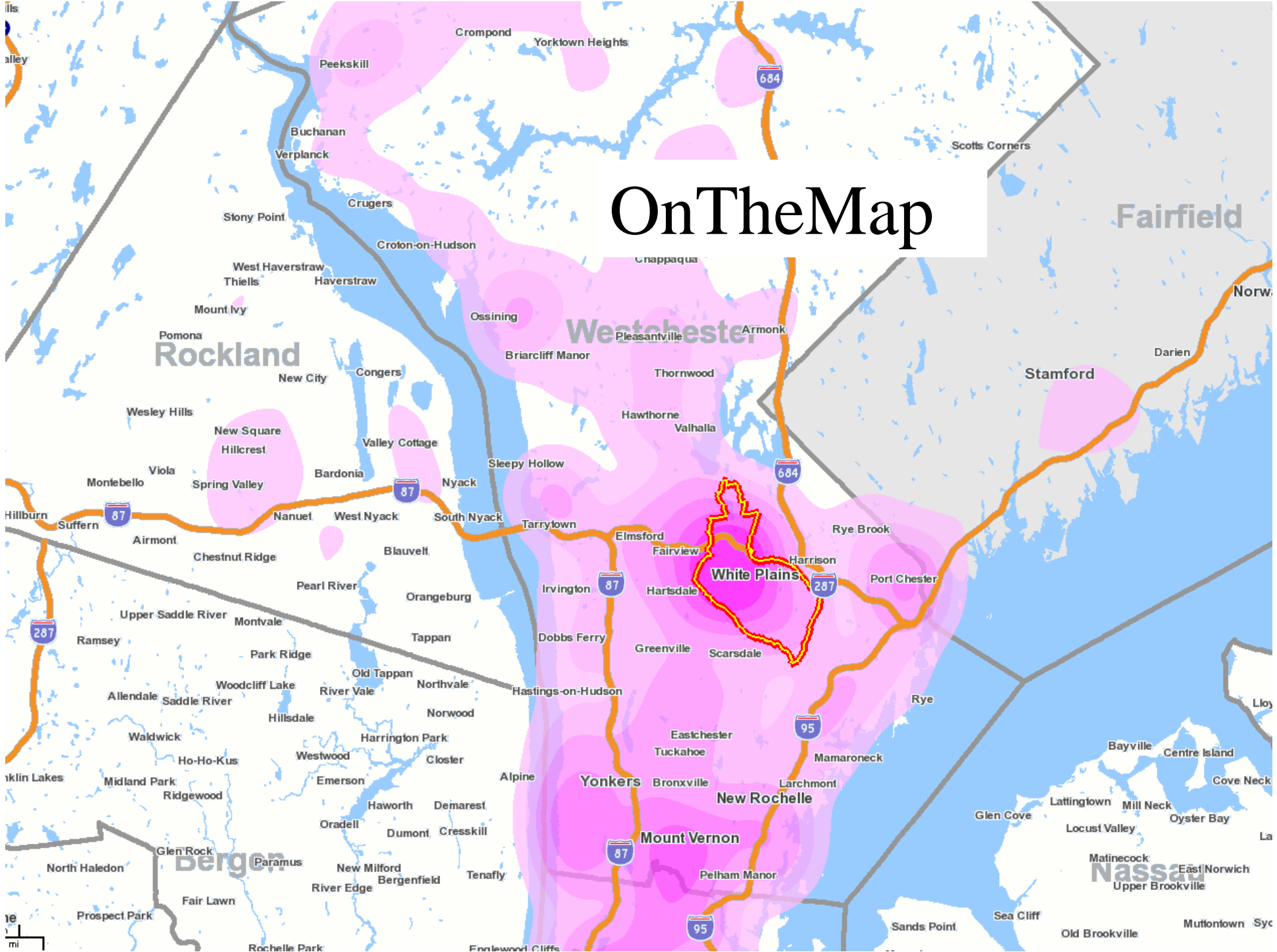
A collaboration, with 48 of the 50 states so far

- Integrates workers' characteristics with their employers' characteristics
- Information available quarterly for small geographic areas and by industry
- Pre-defined and user-defined tables

OnTheMap software

- Easy visualization of massive amounts of data
 - On where workers live and their commuting patterns
- Strict protection of respondents' confidentiality

OnTheMap



Challenges in Integrating Data

Getting the metadata about the original data files into one consistent format

Obtaining electronic access to data at each source

Addressing IT security issues

Addressing the perception of having too much information in one place

Benefits from Integrating Data

Provide decision makers with critical information never before available

- Such as the number of jobs created by new and expanding firms and the number of jobs lost as firms contract or go out of business
- Not just the net change in employment

In other words, create new information in a cost effective way, by leveraging what is already available