

# Collecting, Analyzing, and Presenting Geographic Information in Survey Data

Dave Stinchcomb

Surveillance Research Program  
Division of Cancer Control and Population Sciences  
National Cancer Institute

May 4, 2007

## Introduction

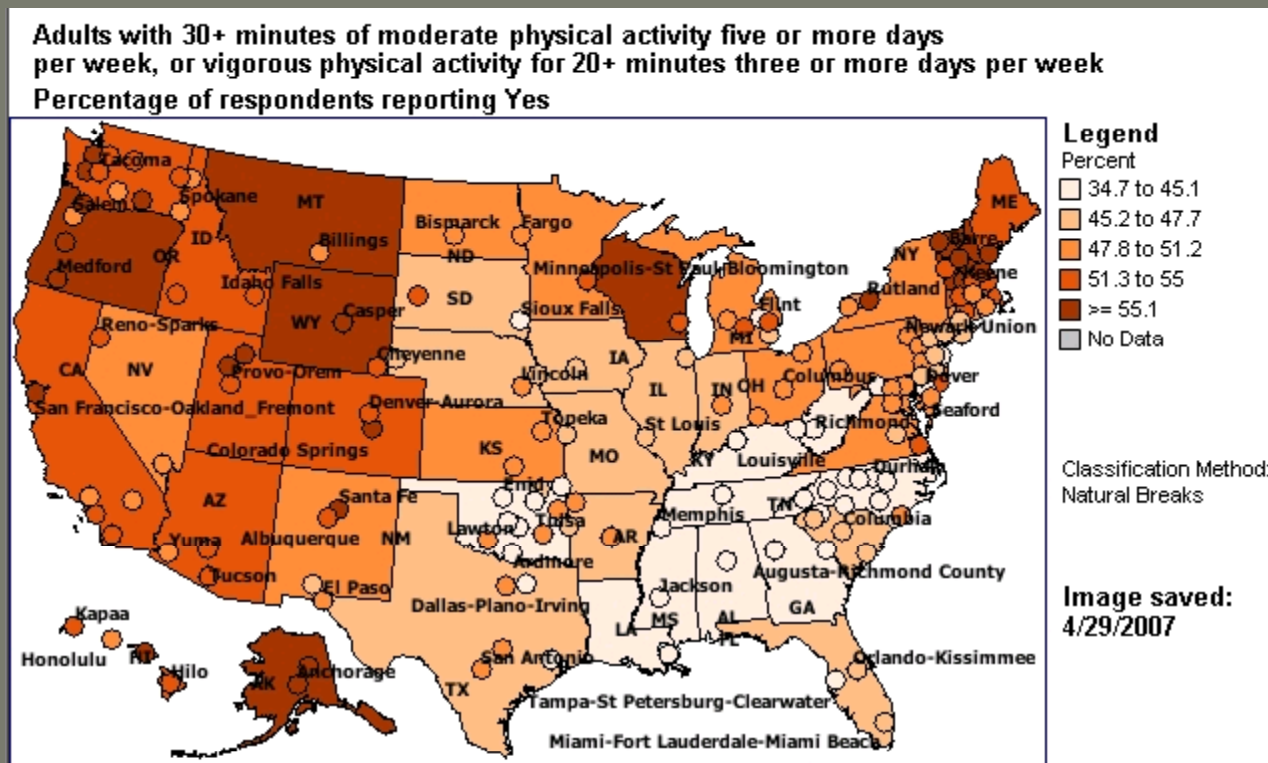
- Geographic information in survey data
  - Respondent's residence, workplace, space-time paths, etc.
  - Ability to stratify results by geographic area
    - Region, state, county, etc.
  - Enable spatial data linkages (via GIS, e.g.)
- Three examples – collecting, analyzing, presenting:
  - California Health Interview Survey (CHIS)
  - Behavioral Risk Factor Surveillance System (BRFSS)
  - Health Information National Trends Survey (HINTS)

## Example 1 – CHIS

- Collects respondent's residence location
- Hierarchical series of information sources:
  - Mailing address: “Is this where you live”?
  - Ask for street address of residence
  - Ask for nearest intersection
  - Mailing address ZIP code
- Geocoding match rates (CHIS 2003):
  - Street address: 85.9%
  - 9-digit ZIP code: 1.1%
  - 7-digit ZIP code: 0.4%
  - 5-digit ZIP code: 12.6%:

## Example 2 – BRFSS

- Large national phone-based survey
- Results available by state and MSA:

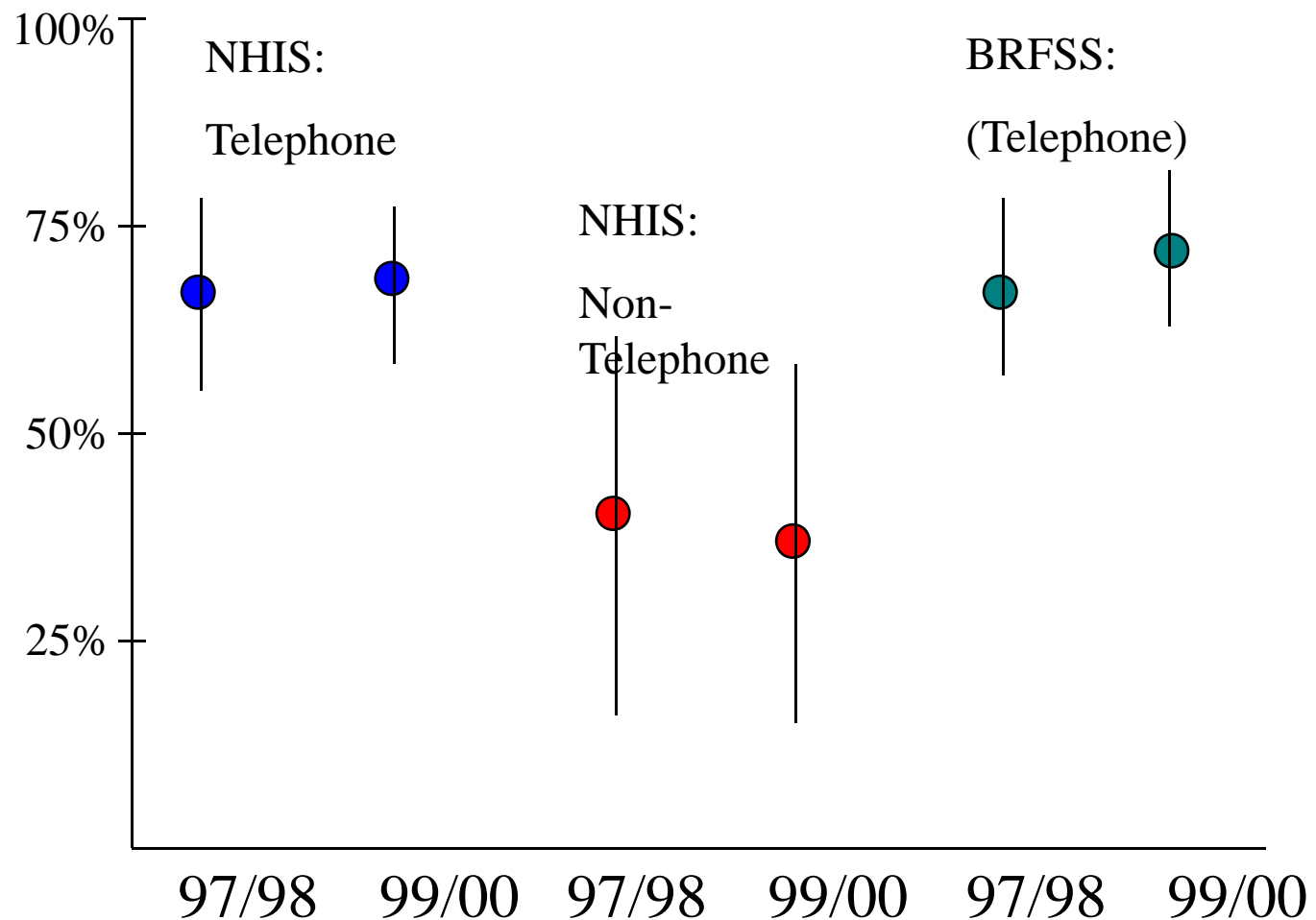


# BRFSS with NHIS

- Statistical modeling project to combine BRFSS with the National Health Interview Survey (NHIS)
  - NHIS: large national in-person survey
- NCI, NCHS, Univ. of Michigan, Univ. of Pennsylvania
- Still in research phase
- Two main goals:
  - Correct for telephone-based survey bias
  - Improve geographic detail
- County-level model with:
  - NHIS response: households with phones
  - NHIS response: households without phones
  - BRFSS response
- Model covariates include:
  - Demographics
  - Socio-economic factors, crime rate
  - Population density, urban/rural, commuting

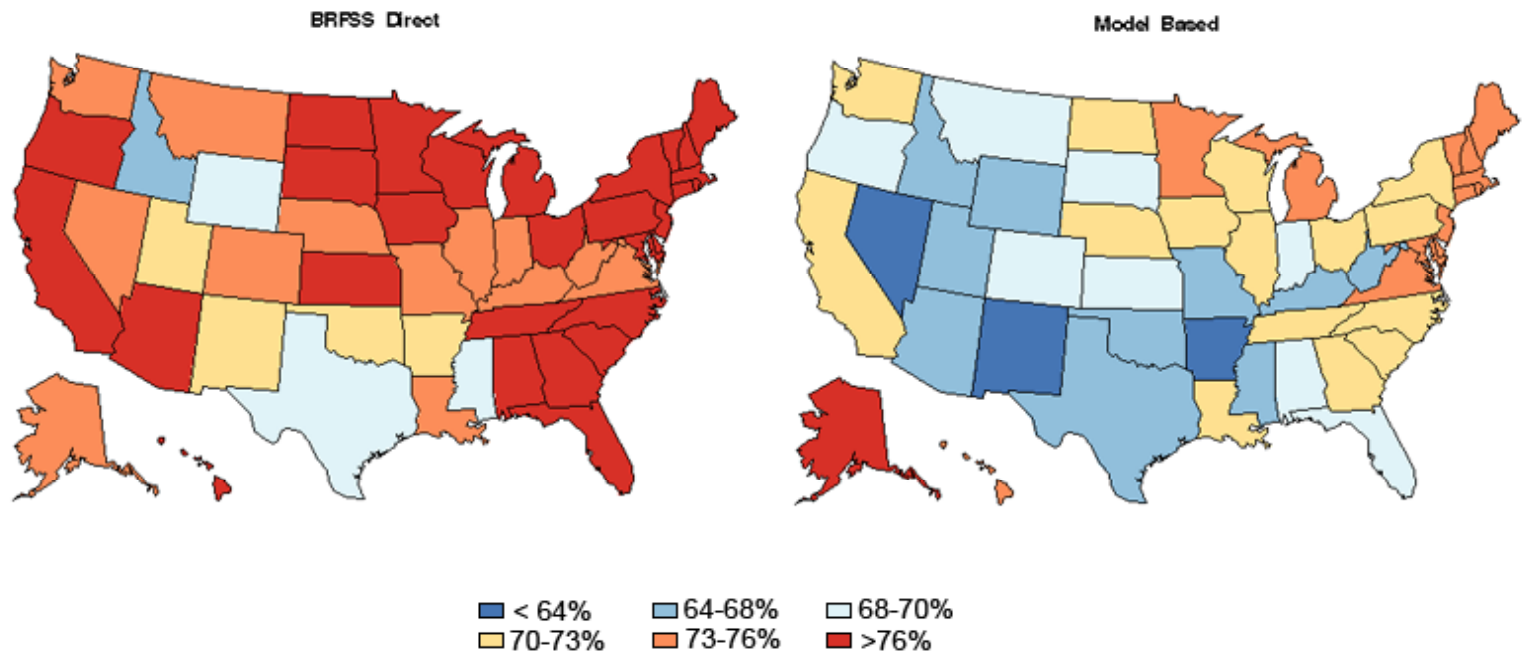
# BRFSS with NHIS – example

**Mammography Estimates from BRFSS and NHIS  
(% mammography within the past 2 years – age 40+)**



# BRFSS with NHIS – example

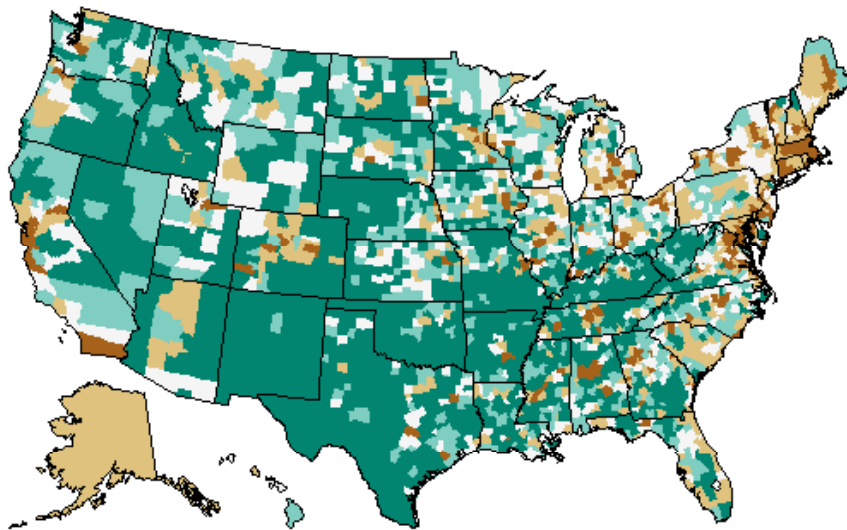
% Mammography within 2 Years, Females (40+), 2000-2003



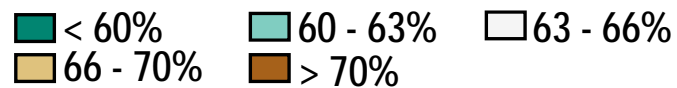
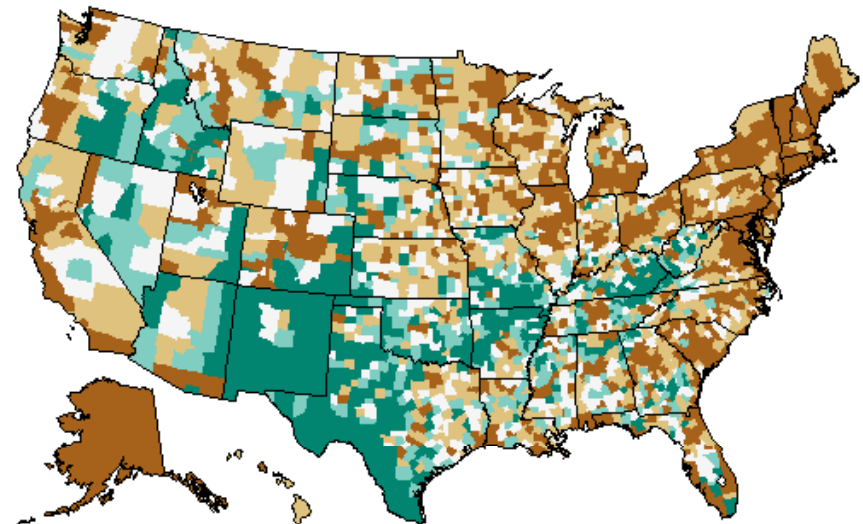
# BRFSS with NHIS – example

## % Mammography within 2 Years, Females (40+)

1997 - 1999



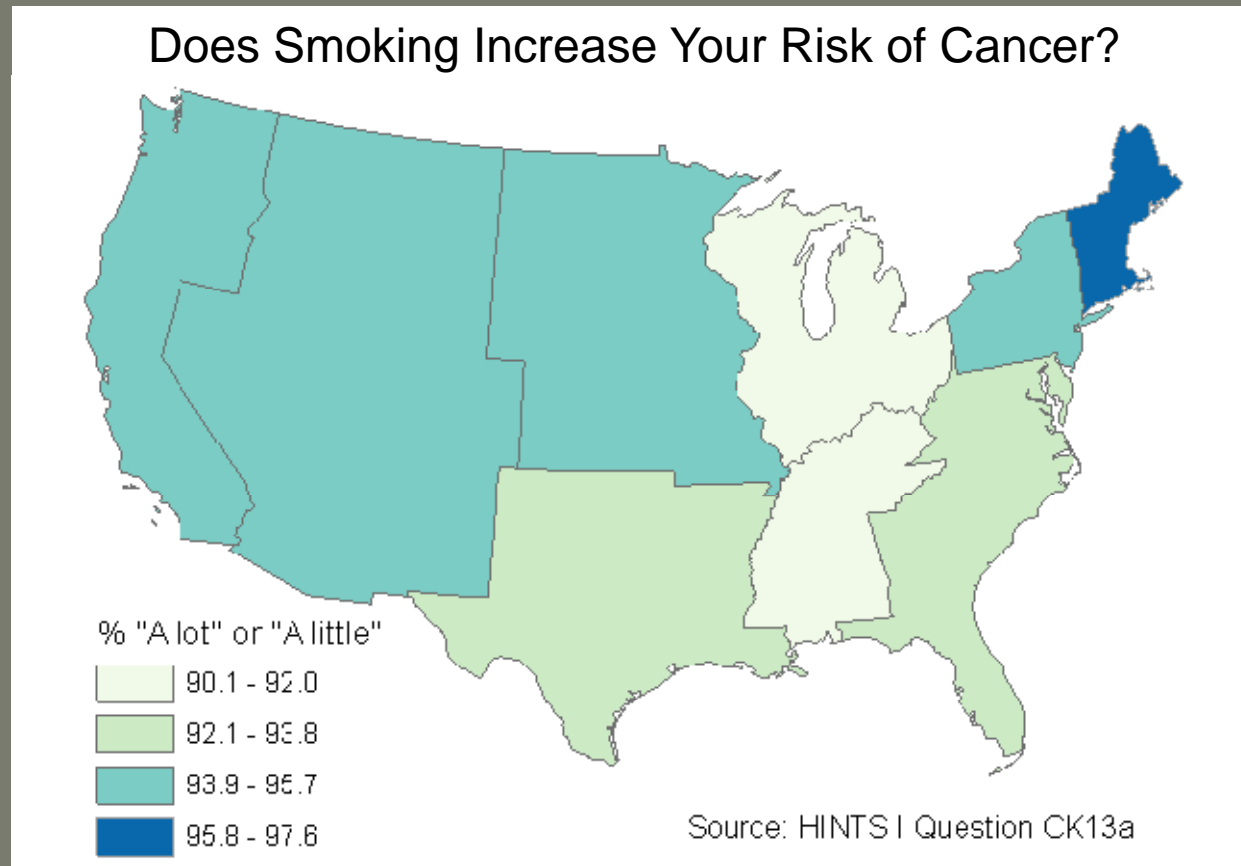
2000 - 2003





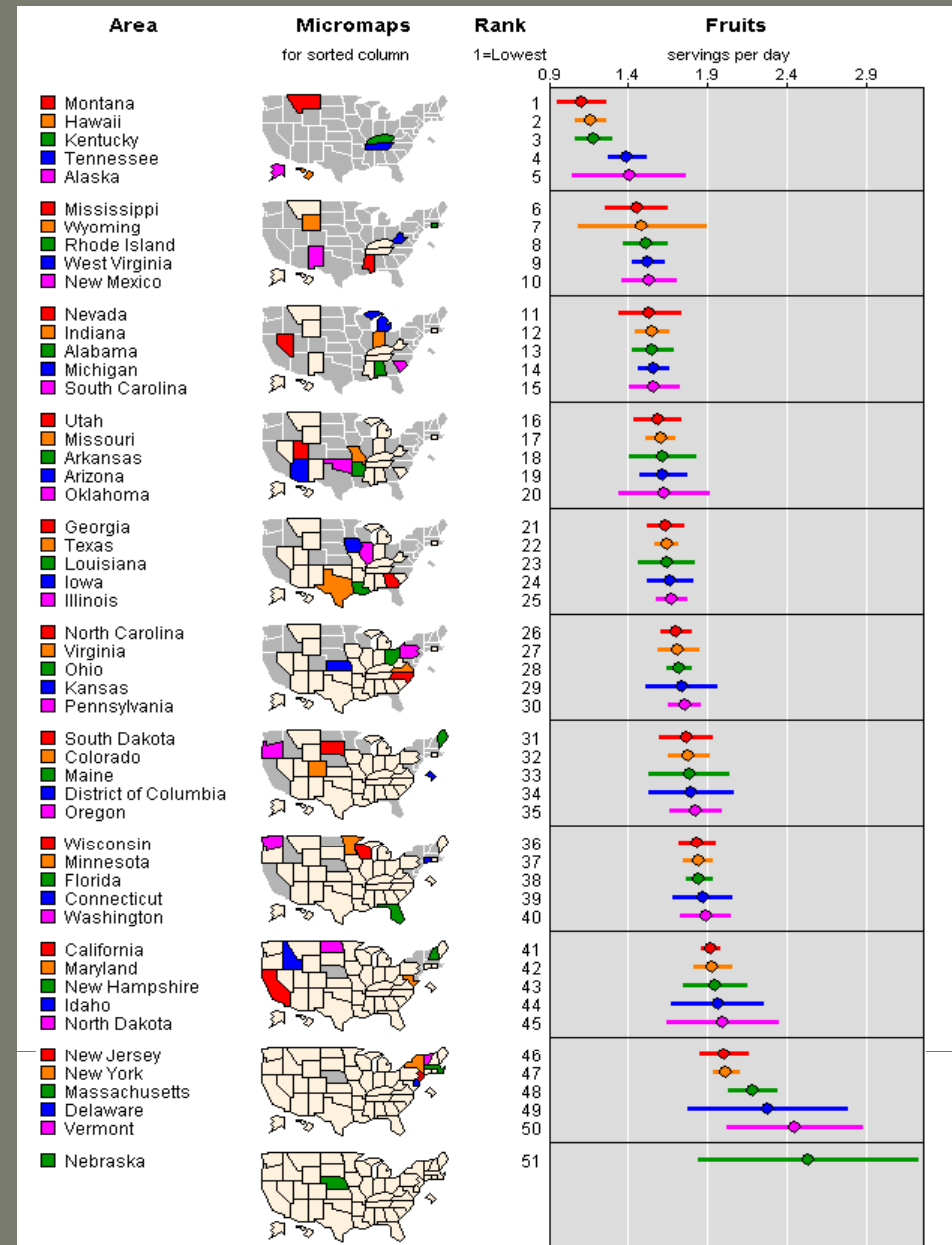
## Example 3 – HINTS

- Relatively small national survey
  - About 6,000 samples from 50 states
  - Supports reporting by Census Division:



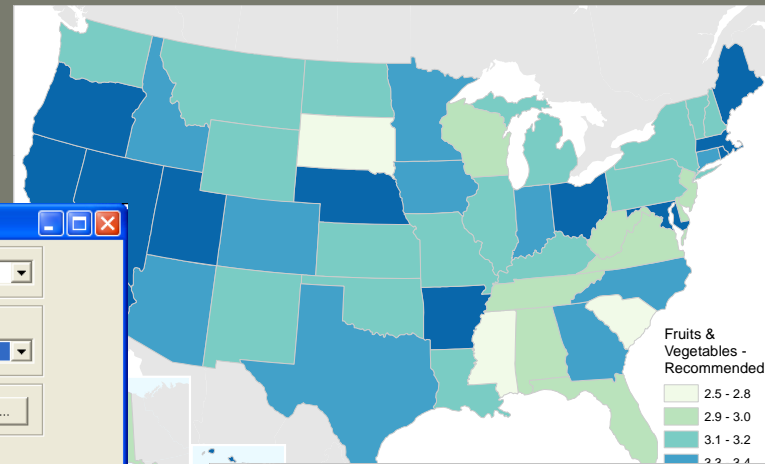
# HINTS by State

- Some states have few samples – wide confidence intervals:



# Smoothing HINTS State Data

- Headbang weighted smoothing
  - Borrows information from neighbors
  - Weighted by sample size



Head-Bang! for ArcMap(TM) - 1.1.2

Input Layer: usstate\_HINTS\_merged

Variable Names: Variable to Smooth: fv\_mean, Weight: fv\_n

Output Filename: [Browse...]

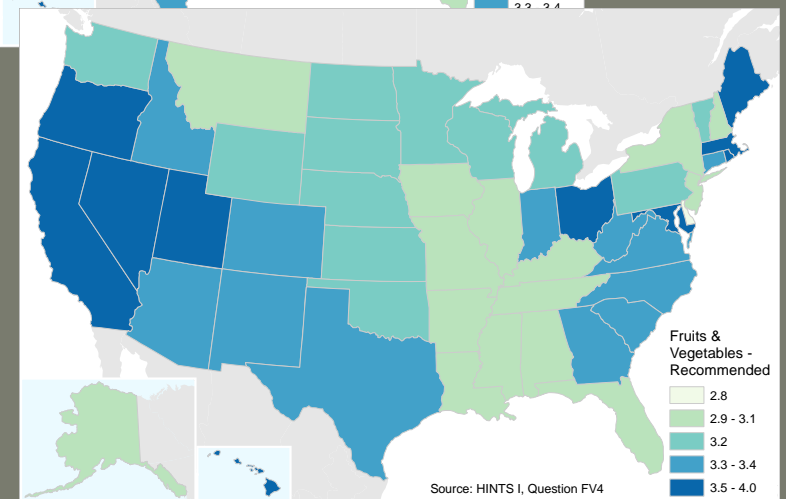
Centroids:  Custom Centroids  
X Centroid: [ ]  
Y Centroid: [ ]

Parameters: Nearest Neighbors (NN): 10, Iterations (Niter): 10, Triples (Ntrip): 7, Angle in Degrees (ThetaStar): 135

Smoothed Count:  Calculate Smoothed Count, Population: [ ]

Merge Options: Variable: [ ],  Text,  Numeric

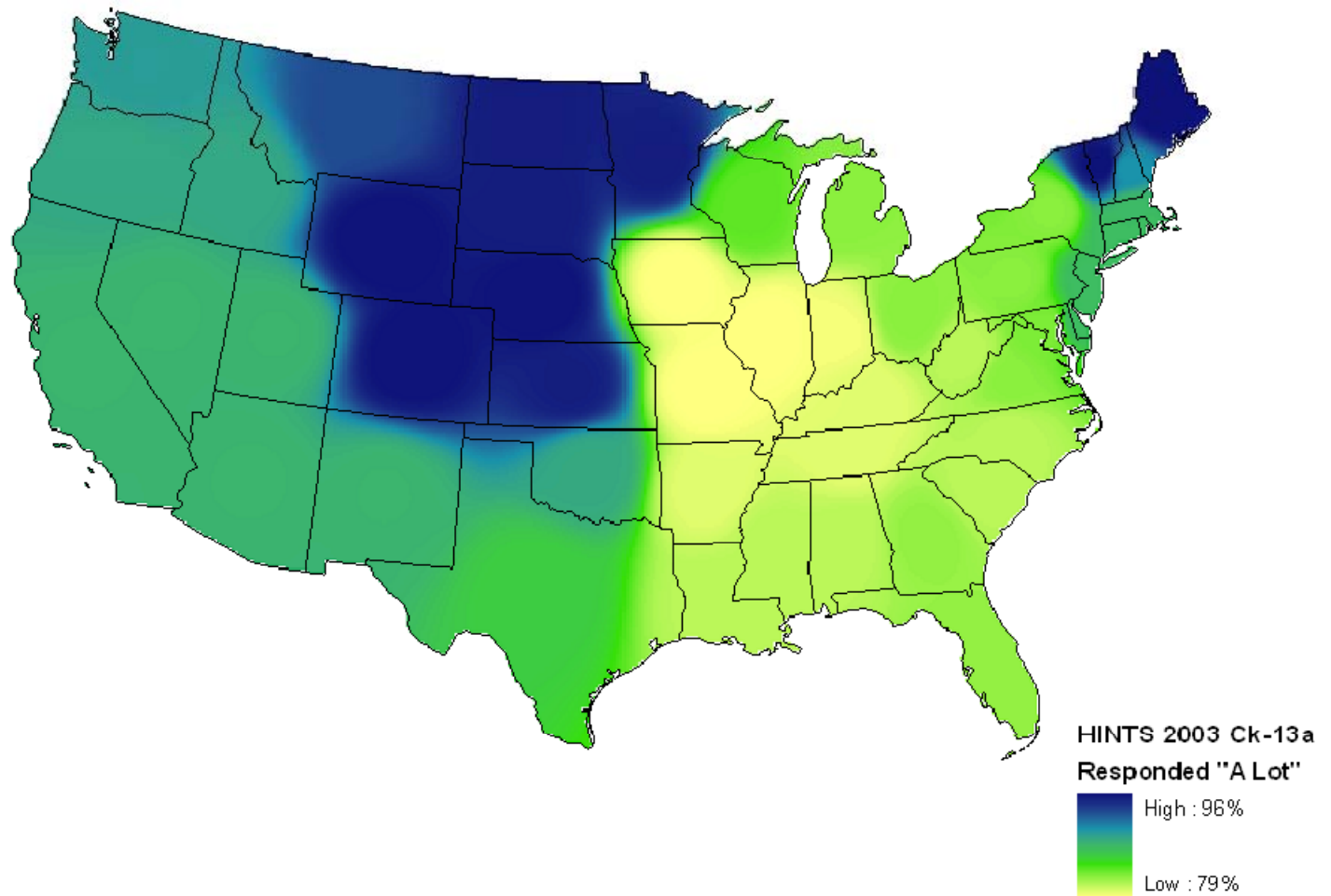
[About Headbang](#) [Help](#) [Smooth]



<http://srab.cancer.gov/headbang/>

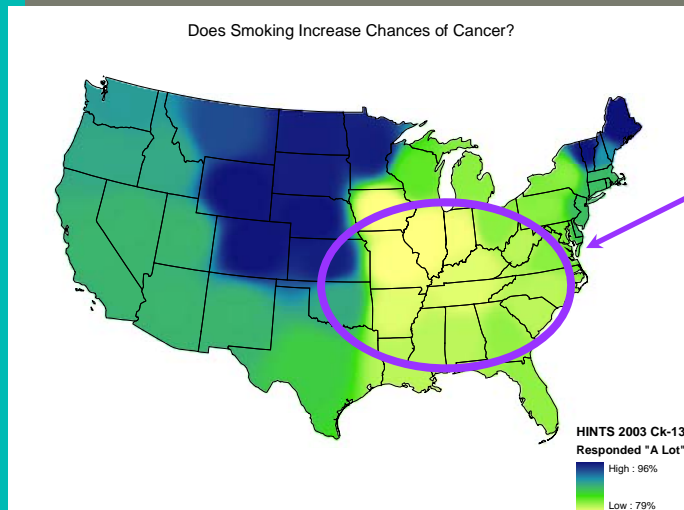
# Convert to a Continuous Surface

- “Weather map” style (an isopleth map)
- Avoids transitions at state boundaries



# HINTS Knowledge Maps

- Maximize geographic information for communication planning
  - Not constrained by Census Divisions
  - Can show more geographic detail where there are more samples



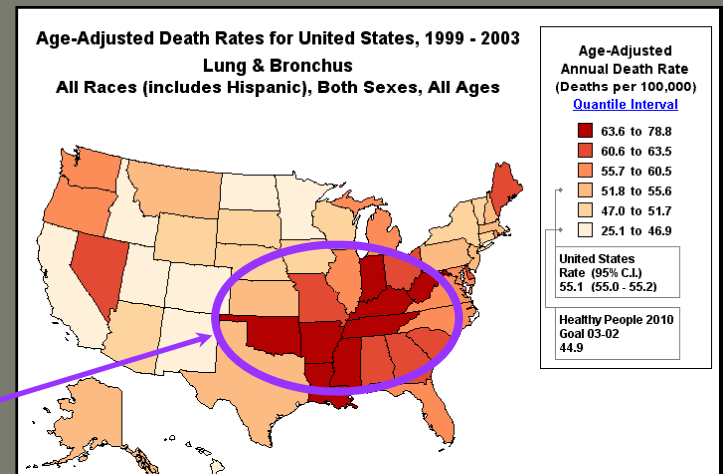
HINTS 2003

"Smoking leads to lung cancer"

Low belief in risk of smoking

State Cancer Profiles  
Lung Cancer Mortality

High lung cancer  
mortality rates



## Conclusions

- Geographic information in survey data
  - Differences in collected information
  - It is possible to collect high quality geographic information about respondents (CHIS)
  - Can augment geographic information by combining with other data (BRFSS/NHIS)
  - Can use smoothing and isopleth mapping to maximize visualization (HINTS)
- Provide the best quality data for public health communication planning
- Can link knowledge/beliefs, behavior, and health outcomes



# Thank You

Dave Stinchcomb  
StinchcD@mail.nih.gov  
301-594-7251