

Mapping Alaska Along With US Counties

Adeline J. Wilcox, US Census Bureau, Washington, DC

OFFICIAL STATEMENT

SAS software does not meet the standards for mapping software followed by the US Census Bureau's Geography Division.

ABSTRACT

I share my experience with mapping Alaska using the SAS® map data sets named COUNTY, COUNTIES and USCOUNTY. Some differences among these three data sets will be summarized. I report difficulties mapping Alaska on the same page with the rest of the United States and describe how deleting data from the Alaskan county equivalent named Aleutians West Census Area improves the appearance of maps made with COUNTIES. Changes in US county FIPS codes encountered while using SAS to analyze Census 2000 data will be described. SAS/GRAPH mapmakers and those who use FIPS codes with longitudinal geographic area data may find this information helpful. SAS users interested in changes in US county geography before the 1990 census will want to look at the ANOMALY data set distributed with SAS/GRAPH Version 8.1. SAS users who want documentation on changes in US county FIPS codes between 1990 and 2000 can find them tabled in this paper.

INTRODUCTION

SAS/GRAPH Version 8.1 offers three map data sets for creating maps displaying counties of the United States. You can study these data sets by looking at the SAS map metadata stored in the SAS data set named METAMAPS. You will probably find METAMAPS in the same library where your SAS map data sets are. SAS Technical Report P-196 advises us that "Many sites automatically assign a libref of MAPS to the SAS data library that contains the SAS/GRAPH map data library." Running PROC DATASETS with the CONTENTS statement WHERE COUNTRY='UNITED STATES', TYPE='GRAPH', and VAR1='COUNTY' lists four SAS data sets. Three of these are map data sets; the fourth is a template data set shared by the other three. The map data sets are named COUNTIES, COUNTY, and USCOUNTY and the template data set is named CNTYNAME.

While SAS Technical Report P-196 tells us that the SAS map data set named COUNTIES will produce a map displaying the 48 states, Hawaii and Puerto Rico, SAS Technical Support told me that, beginning with Version 7, this data set also contains the coordinates for the county boundaries of Alaska. Neither COUNTY nor USCOUNTY contain the county boundaries of Alaska.

At the US Census Bureau, the general term county equivalent includes geographic entities such as boroughs, census areas, parishes, municipios, and cities independent of any county. Baltimore City, Maryland is one such independent city. Baltimore City, MD and Baltimore County, MD are two separate and distinct geographic entities. Alaska does not have any counties; only boroughs and census areas.

PROJECTING MAP DATA SETS

In his mathematical treatment of graphical coordinates, Wilkinson identifies map projections as a continuous transformation (Wilkinson, 1999). One of these transformations projects the distribution on the sphere onto the plane.

Version 8 SAS/GRAPH documentation tells us that the map data sets COUNTIES and COUNTY are among the five map data sets containing unprojected values of longitude and latitude. Consequently, you must GPROJECT these data sets before GMAPping them. PROC GPROJECT transforms the spherical coordinates in the unprojected map data set to Cartesian coordinates in the projected map data set it outputs.

```
libname mine '~/nesug01/';
proc gproject data=maps.counties
out=mine.projec52;
id state county;
run;
quit;
```

Once you have PROJECTed the SAS map data set COUNTIES, you can run PROC GMAP and output an outline map.

OUTLINE MAPS

It is helpful to begin mapping work by producing an outline map. Having an outline map to refer to will aid in identification of missing geographic entities. When run, this code will output an outline map of COUNTIES.

```
proc gmap map=nesug01.projec52
data=nesug01.projec52;
pattern value=mempty;
id state county;
chorosegment/discrete
coutline=black nolegend;
run;
quit;
```

FIPS CODES

Mapping US county level data requires an understanding of Federal Information Processing System (FIPS) codes. Each state's two-character FIPS code uniquely identifies it. US counties can be uniquely identified by their five-character FIPS code, of which the first two characters are the state FIPS code. SAS users are referred to <http://www.census.gov/population/estimates/fips.txt> (visited 23Apr2001). In SAS/GRAPH GMAPping of the US, FIPS codes are used as "identification variables that uniquely identify the unit areas in a map" (SAS Institute, 1990).

THE TEMPLATE DATA SET

The SAS Version 8.1 template data set for the map data sets delineating U.S. counties is named CNTYNAME. It contains state and county FIPS codes. SAS Technical Report P-196 describes how to use template data sets to build response data sets. CNTYNAME contains 3,143 observations and reflects US county geography as of 1994. See the table. The data set you plan to map, called the response data set, must have a unique state and county FIPS code on each observation. If you do not have a one-to-one correspondence between geographic entities in your data and the template data set, you need to know which geographic entities are missing from your data set. Page 739 of the Version 8 SAS/GRAPH documentation guides both the SAS user mapping with missing data and the one who wishes to display only part of a map data set on the output map.

THE RESPONSE DATA SET

Response data sets are described in *SAS/GRAPH® Software: Reference Version 8*. Besides the identification variables matching those in the template data set, response data sets contain at least one response variable which you want to present on your map.

To map 1990 census data, counties created after 1990 and before 1994 must be added to the response data set. Fortunately, both Yukon-Koyukuk and Southeast Fairbanks Census Areas have the same response value for the response variable in my response data set from the 1990 census. So I created observations for the new boroughs of Denali (068) and Yakutat (282) with code similar to

```
data createco;
length state $ 2 county $ 3 response $ 1;
input state county response;
lines;
02 068 0
02 282 0
;
```

Because the Census Area from which Yakutat Borough was formed, Skagway-Yakutat-Angoon (231), was subsequently renamed Skagway-Hoonah-Angoon and recoded (232), I used a statement like

```
if state='02' and county='231' then
county='232';
```

to update the 1990 census response data to the template data set.

To map geographic data as of 01Jan2000, counties created, dissolved, or recoded after 1994 and before 2000 must be mapped back to 1994 geography. Observations for the now-dissolved counties in Montana and Virginia should be created. Using the response value for Halifax County, VA (083) I wrote code similar to:

```
data createco;
length state $ 2 county $ 3 response $ 1;
input state county response;
lines;
51 780 1
30 113 1
;
```

This code creates response observations for the formerly independent city of South Boston, VA and the former county of Yellowstone National Park, MT (113). Luckily, both Gallatin County, MT (031) and Park County, MT (067) had the same values of the response variable in my Census 2000 response data set.

When Dade County, FL (025) renamed itself Miami-Dade County in 1997, it received a new FIPS code (086). The following SAS statement will map Miami-Dade County back to Dade County:

```
if state='12' and county='086' then
county='025';
```

THE US COUNTY MAP DATA SET

On the USDA's National Agricultural Statistics Service Web site, I found a response data set suitable for mapping with the USCOUNTY map data set. Despite its "All States" description, USDA's 1998 All Dry Hay data set contains no data for AK, CA, TX, FL, and several other states. When the ALL option is used on the PROC GMAP statement as

```
proc gmap map=maps.uscounty data=hay all;
id state county;

choro Yield/discrete coutline=black
legend=legend1;
format Yield hayfmt.;
run;
quit;
```

all 50 states are displayed on the map. See Figure 1. Note that the USCOUNTY map can show at most one response value for Alaska because it contains no county boundaries for that state. USCOUNTY

is the only one of the three county SAS map data set which presentably displays the entire US on one page.

For USDA's dry hay data set, it does not make subject matter sense to assign hay yield to South Boston City. Anyway, the independent city county equivalents in the state of Virginia appear only as specks on the USCOUNTY map. The Montana county equivalent named Yellowstone National Park is actually visible on the USCOUNTY map as a narrow L-shaped area with a spur at its heel. The way I have grouped the hay yield data prevents me from displaying a response value for Yellowstone National Park because the hay yield values for Gallatin County and Park County fall into different categories.

I programmed both Figures 2 and 3 using the USCOUNTY map data set. Notice I deleted Alaska from both. For the 1990 census and Census 2000 respectively, these maps display the counties and county equivalents in which housing units, before nonresponse followup, were enumerated entirely by mailing back census forms.

THE COUNTY MAP DATA SET

The outline map of this map data set displayed the counties of the 48 coterminous states and DC in the lower right quadrant of a landscape page, Alaska without county boundaries in the upper left quadrant, and Hawaii in the lower left corner. Deleting Alaska from the input data set for PROC GPROJECT changed the PROC GMAP output very little. The COUNTY map data set appears to offer no features not found in either the USCOUNTY or COUNTIES map data sets.

THE COUNTIES MAP DATA SET

My outline map of the entire COUNTIES map data set dismayed me. A shrunken continental US drifted towards the far left of the landscape page. Like it, Alaska was rotated 90°. Tiny islands floated in three of the corners. Reference to an atlas identified one group of islands as the westernmost Aleutian Islands, aswim halfway across the page from the rest of the state of Alaska.

Only by deleting all noncontiguous areas from the map data set before projection, can a presentation-quality map of the 48 coterminous states and DC be created. Removing all islands and Alaska produces a very detailed map on which the Florida Keys can be seen. Neither USCOUNTY nor COUNTY show the Florida Keys.

For mapping Alaska, I found the COUNTIES map data set useful. Except for Aleutians West Census Area, Figure 4 shows the rest of the Census 2000 data that could not be presentably displayed in Figure 3. I programmed the map in Figure 4 by deleting the 48 coterminous states, DC, HI, Puerto Rico, and the county equivalent named Aleutians West Census Area. The result is not particularly attractive. In the 1990 census, only the southernmost of the three shaded county equivalents in Alaska was entirely mailback type of enumeration.

There may be a way for SAS users with excellent PROC GREPLAY skills to create a presentation-quality map of all US counties. Until a SAS map data set that keeps all of the Aleutian Islands together and contains county boundaries for Alaska comes along, it does not seem worth pursuing.

COUNTY BOUNDARY CHANGES

Information on county boundary changes since 1990 can be found at

<http://www.census.gov/population/www.estimates/geonotes/> (Visited 26Jun2001) Many boundary changes between counties and independent cities occurred in Virginia but they are all too small to be visible when the output is printed on 8.5 x 11 paper. In Maryland, prior to 01Jul1997, the city of Takoma Park had been split between Prince George's

County and Montgomery County. Effective that date, all of Takoma Park city is included in Montgomery County. Again, this boundary change is too small to be visible on any of the three US county map data sets. In Alaska, annexation of part of Skagway-Hoonah-Angoon Census Area by Juneau Borough took place effective 01Jan1994. It is not known if this change was made to the SAS Version 8.1 COUNTIES map data set.

CONCLUSION

SAS users need to recognize that US county geography is not fixed. Changing FIPS codes can confuse the naive programmer. PROC GMAPpers need to be familiar with current FIPS codes and know the vintage of their response data sets and SAS map and template data sets.

REFERENCES

SAS Institute Inc., SAS® Technical Report P-196, *SAS/GRAPH® Software: Map Data Sets, Release 6.06*, Cary, NC: SAS Institute Inc., 1990. 165 pp.

SAS Institute Inc., *SAS/GRAPH® Software: Usage, Version 6, First Edition*, Cary, NC: SAS Institute Inc., 1991. 887 pp.

SAS Institute Inc., *SAS/GRAPH® Software: Reference Version 8*, Cary, NC: SAS Institute Inc., 1999. 1240 pp.

Wilkinson, L., *The Grammar of Graphics*, New York: Springer-Verlag, 1999

ACKNOWLEDGEMENTS

Stephen Ash, Mike Fournier, and Catherine Miller of the US Census Bureau and Marcia Surratt of SAS Institute, Inc. contributed factual information to this paper. Mike Zdeb recommended the HWPSL009 font for the maps.

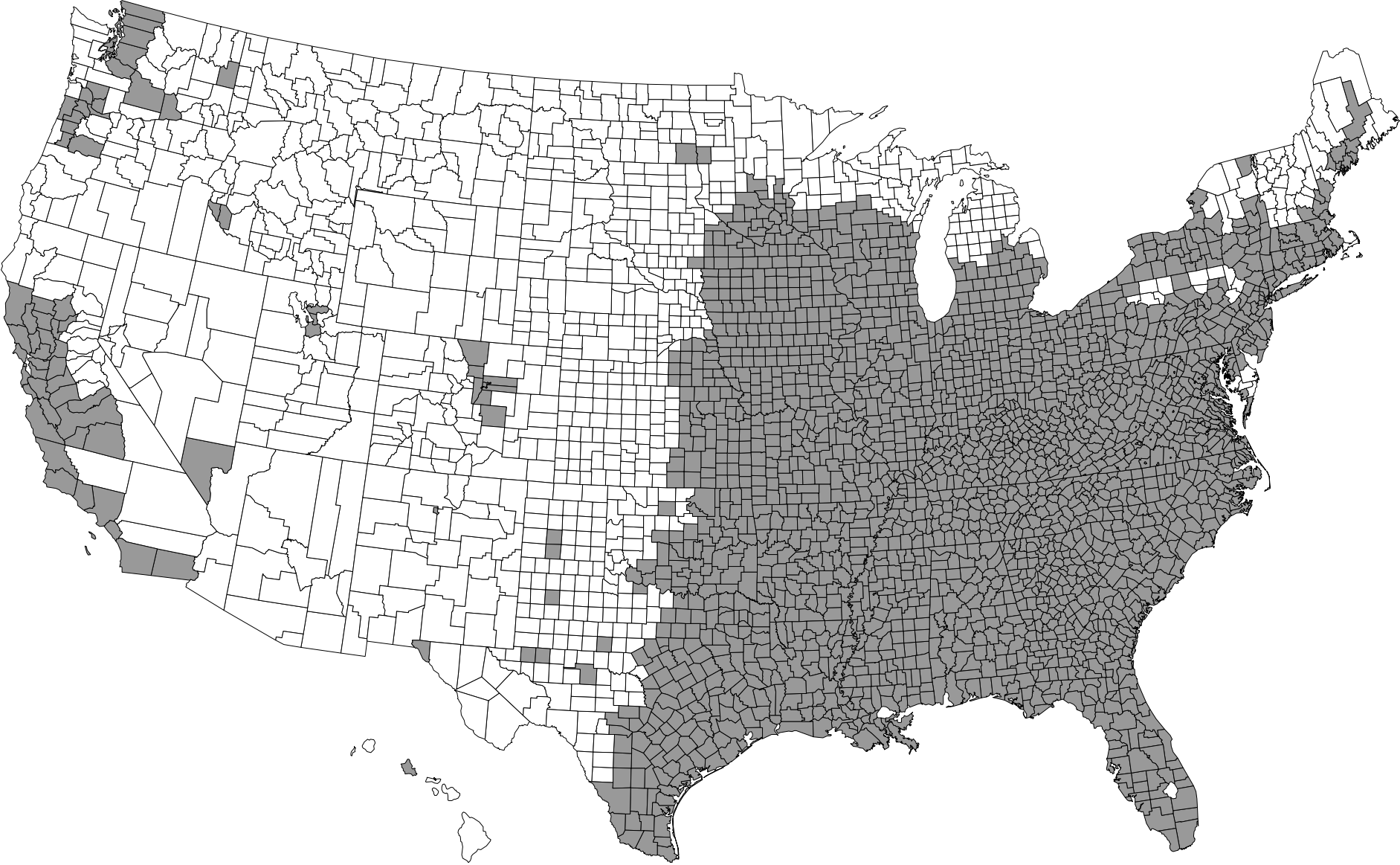
CONTACT INFORMATION

Adeline Wilcox
 US Census Bureau
 Planning, Research and Evaluation Division
 SFC-2, Mod 2, Room 1805
 Washington, DC 20733-9200

Adeline.J.Wilcox@census.gov
 (301) 457-4142 (Please do not leave voice mail)

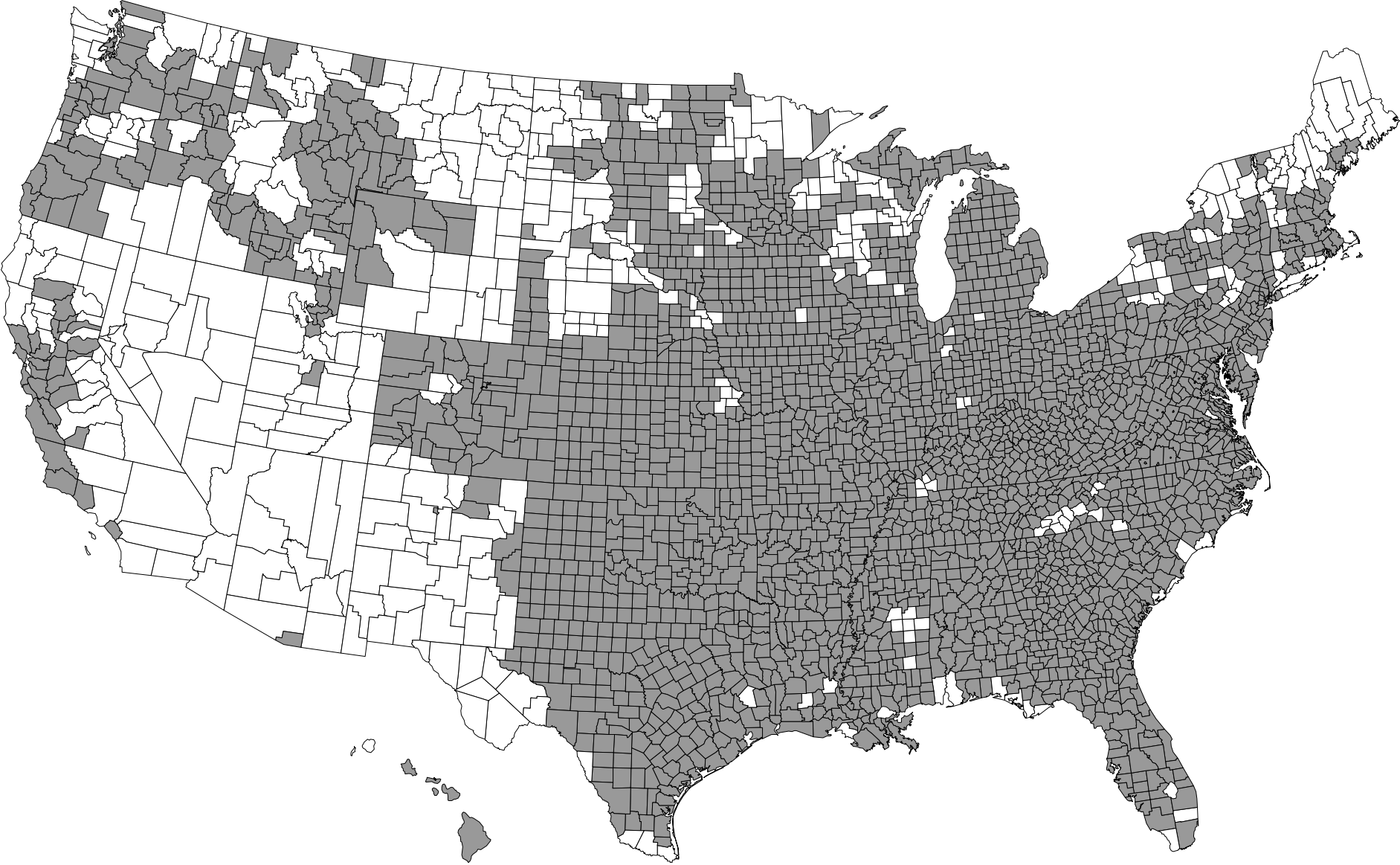
SUMMARY OF INTERCENSAL CHANGES IN WHOLE US COUNTIES AND COUNTY-EQUIVALENTS			
DATE	STATE	CHANGE	TOTAL
01Apr1990			3,141
07Dec1990	AK	New Borough Denali (068) formed from part of Yukon-Koyukuk (290) and part of Southeast Fairbanks (240)	3,142
22Sep1992	AK	New Borough Yakutat (282) formed by removing territory from Skagway-Yakutat-Angoon (231) and renaming the latter Skagway-Hoonah-Angoon (232)	3,143
30Jun1995	VA	Independent city of South Boston (780) dissolved into Halifax County (083)	3,142
07Nov1997	MT	Yellowstone National Park (113) divided between Gallatin County (031) and Park County (067)	3,141
13Nov1997	FL	Dade County (025) renamed Miami-Dade (086)	3,141
01Apr2000			3,141

Figure 2. 1990 Census Counties of the 48 Coterminous States, DC and Hawaii



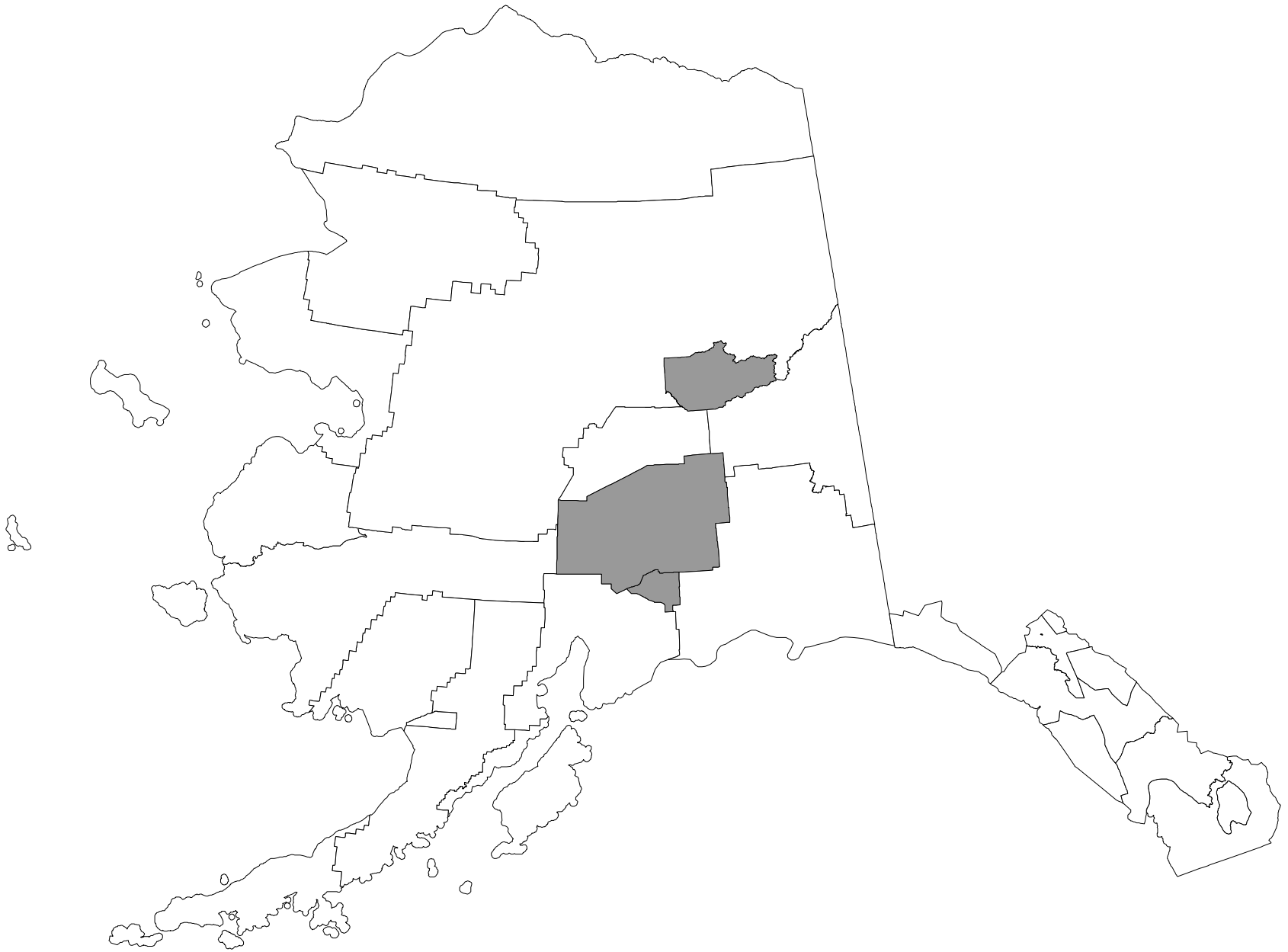
Type of Enumeration Entirely Mailback No Yes

Figure 3. Census 2000 Counties of the 48 Coterminous States, DC and Hawaii



Type of Enumeration Entirely Mailback No Yes

Figure 4. Census 2000 Alaska Boroughs and Census Areas



Type of Enumeration Entirely Mailback No Yes
Aleutians West Not Mailback and Not Shown