PROBLEMS IN SPATIAL DATA ANALYSIS

Deane Merrill, Lawrence Berkeley Laboratory

ABSTRACT

A common problem in dealing with socio-economic, environmental, demographic, and health data is the need to combine data from differently defined geographic entities. For example, 1970 and 1980 Census data are not directly comparable even at the county level, due to changes in geographic definitions. More drastic problems occur in combining data from different government agencies.

In most integrated data systems, either the problem is ignored, or data are forced into consistency with undesirable effects: one must either aggregate to larger areas with loss of geographic detail, or disaggregate to smaller areas under arbitrary assumptions. A unique solution to this dilemma has been implemented in SEEDIS, the Socio-Economic Environmental Demographic Information System at Lawrence Berkeley Laboratory (LBL).

In SEEDIS, some 70 geographic levels (e.g., 1970 or 1980 counties) are defined, corresponding to archived data files. If efficiency considerations are ignored, each file needs to be stored only at the most detailed level for which the data are complete. Geocode correspondence files provide the information required to transform data from any level to any other level. Inevitably, disaggregation requires ad hoc proportionality assumptions; different assumptions are suitable for different applications. In SEEDIS, unlike other systems, these assumptions are under the user's control.

The SEEDIS geocode conversion files and proxy variable files can be used to aggregate and disaggregate arbitrary data files, either within or outside SEEDIS. Other files now being developed describe other dimensions of the data, for example industrial or occupational classification.

THE NEED FOR GEOGRAPHIC COMPARABILITY

A difficult problem faces the analyst who tries to combine and analyze data from a number of different sources: in general the geographic units of his/her various data sets do not correspond. Fortunately, Federal Information Processing System (FIPS) standards are being used increasingly within the Federal government. Although FIPS standards can resolve differences of definition themselves changed slightly from 1960 to 1970. Should one say that a city's population has increased by 30% if the increase is due merely to an enlargement of the incorporated area? Such an increase would be important to a city planner concerned with tax revenues, but not to an environmentalist or demographer concerned with changes in population density. In this paper we focus our interest on the more difficult second case, where one wants to study various aspects of the physical world, regardless of political nomenclature.

Methods are needed for easily comparing data sets which describe slightly different geographic entities, for example 1970 or 1980 Census counties. A related problem is the need to aggregate data to larger geographic areas (for example states to Federal regions) either to achieve meaningful sample sizes or for summary display purposes.

SOME APPROACHES TO THE COMPARABILITY PROBLEM

Analysts and government agencies have adopted various schemes in an attempt to resolve the problems of geographic incompatibility.

The simplest approach (not a solution) is to simply present the data for the particular geographic entities to which they pertain. For example, the Bureau of the Census provides 1970 data for 1970 counties and 1980 data for 1980 counties. The conscientious user soon notices that the county codes in the 1970 and 1980 Censuses are slightly different from each other, and in both cases are slightly different from standard FIPS codes. The FIPS definitions themselves changed slightly between 1970 and 1980.

Another approach (again not a solution) is to match areas by name, or by geographic in the Census Bureau's County Data Book, differences between county definitions in the 1960 and 1970 Census are ignored. In the cases where county boundary changes occurred, comparisons of 1960 and 1970 data give erroneous and misleading results.

A third approach was adopted for Oak Ridge National Laboratory's Geocology Data Base [1], a collection of diverse county-level environmental data designed for use with the Statistical Analysis System (SAS). The authors defined a standard set of 3071 county equivalents, which are aggregates of counties as defined in several major data sets. For example, each indepen-
dent city in Virginia is combined with one of the coun-
tries adjacent to it. By simple aggregation, the Geoco-
logy authors calculated (or estimated) all data in terms
of the 3071 standard county equivalents. The Geoeco-
logy user is thus relieved of the onerous task of achiev-
ing comparability, but at a price; finer geographic
detail originally present in some of the files has been
lost. Also, data from a few larger areas had to be disag-
gregated to the standard Geocology county units,
requiring arbitrary proportionality assumptions. Like-
wise, disaggregation of data from newer geographic
entities like 1980 Census counties will require similar
proportionality assumptions.

A fourth approach, rather opposite to that of Geoeco-
logy, was adopted for the early results of the Lawrence
Berkeley Laboratory's PAREP (Populations at Risk to
Environmental Pollution) project [2]. In the PAREP pro-
ject, data were disaggregated to provide estimates for
county pieces, i.e. the largest subcounty units that can
be aggregated to form the various county units of any
of the original files. Although the user loses none of the
detail present in the source data files, a serious draw-
back of this approach is its dependence on propor-
tionality assumptions out of the user’s control.

The next section of this paper describes a fifth
approach, which has been implemented in SEEDIS, the
Lawrence Berkeley Laboratory’s Socio-Economic
Environmental Demographic Information System [3].
General features of SEEDIS are described in the final
section of this paper.

THE SEEDIS APPROACH TO
THE COMPARABILITY PROBLEM

Fundamental to SEEDIS is the concept of geographic
level, i.e. the geographic detail of the data in question.
Most geographic information systems, for example
UPGRADE [4] and DIFMAP [5], define several geographic
levels, usually including nations, states, counties, and a
few others. For mapping purposes, geographic base
files (GBF) are provided, which describe the boundaries
of each geographic unit at each level. Presently,
SEEDIS defines 77 different levels of geographic detail;
32 of these have associated GBF’s for choropleth or
symbol mapping.

In Table 1 are listed the 77 geographic levels in SEEDIS
as of February 1982. “Level” is the name used by
SEEDIS; “Description” briefly describes the geographic
entities of the level; “Year” is a year for which the enti-
ties are defined; and “Units” is the number of geo-
graphic entities at that level. “Map” indicates the avail-
ability of a GBF for mapping; “p” specifies polygon
boundaries for shaded choropleth mapping; “s” speci-
fies point locations for symbol mapping; “c” speci-
sifies the absence of a GBF. Most geographic levels
are defined for the United States, including the terri-
tories of American Samoa, Guam, Puerto Rico, and Vir-
gin Islands. The NATION80 level provides the 233
nations of the world, according to 1980 PIPS definitions.

SEEDIS, unlike other systems, explicitly distinguishes
between minor variants of geographic levels, for exam-
ple 1970 and 1980 Census counties. Separate levels
describe Bureau of Economic Analysis Areas as defined
in 1959 and 1977, or Standard Metropolitan Statistical
Areas (SMSA’s) as defined in six different years. Each
SEEDIS level corresponds to at least one data file
installed in the system; every data file is installed at
least at its own most detailed geographic level. A file
can be installed at several different levels; for example
Summary Tape File 1A (STF1A) of the 1980 Census of
Population is installed at ten different levels; namely,
those provided by the Bureau of the Census plus some
others obtained by aggregation.

Of course it is impractical to archive every data file at
every geographic level. SEEDIS achieves comparability
among different files through a set of detailed geocode
correspondence files. These files, still being completed,
describe each geographic level in terms of the units of
other levels. Where entities mutually overlap, for exam-
ple 1970 and 1980 Census counties in Alaska, subcounty
entities (1970/1980 county pieces) are defined whose
membership in each 1970 or 1980 county is unique.

The SEEDIS user obtains data from an archived data-
base at its original geographic level, then he/she
transforms the extracted data to any other desired
level for comparison with other data. As an example,
suppose one wishes to map, at the Census tract level,
percentage change in population between 1970 and
1980. As indicated in Table 1, SEEDIS provides polygon
mapping at the 1970 census tract level (the GBF was
produced at LDL for the Census Bureau). At least four
private companies have produced proprietary GBF’s of
1960 Census geography, but none is available in the
public domain.

In SEEDIS, 1960 Census population can be obtained for
1980 Census tracts; the data are aggregated or disag-
gregated (usually the former in this case) to the level of
1970 tracts: then 1970 population is extracted, the per-
centage change is calculated, and the map is drawn.
Where 1970 and 1980 tracts mutually overlap, an
automatic intermediate step involving partial disaggre-
gation to the level of 1970/1980 tract pieces. Special
software, still being developed, is available in SEEDIS to
make this task as simple as possible for the user.

In aggregating or disaggregating, SEEDIS must distin-
uish between additive data, such as land area or labor
force, and non-additive data, such as population density
or per capita income. Four cases are considered:
(1) aggregation of additive data involves simple addition;
(2) disaggregation of additive data assumes propor-
tionality with an proxy variable such as land area or
population, which is selected from a menu by the user;
(3) aggregation of non-additive data is an average,
weighted by a user-selected variable; (4) disaggregation
of non-additive data assumes the same data value
throughout the entire larger area.

IMPLEMENTATION OF DATA
AGGREGATION AND DISAGGREGATION

At the present time, data aggregation and disaggrega-
tion in SEEDIS are crudely implemented for most but
not all geographic levels. The user interface is primi-
tive and the choice of proxy variables is limited. The
implementation is somewhat level-specific, and
response is slow. Present development is directed
toward elimination of all these problems.

Later, default weights and proxy variables will be
specified for all 27,000 data elements in SEEDIS. (A
data element is an attribute stored for each geographic
entity in a database.) At the same time, the additivity
or non-additivity of each data element will be recorded,
so that less user intervention will be required. The discriminating user will still be free to choose his/her own weighting and proxy variables, when desired; these can be supplied by the user or selected from any database in SEEDIS.

The concept of geographic level will be generalized to other data dimensions, including time, race, age, industry, occupation, and cause of death. Hierarchical descriptions, cross-level correspondence files, and proxy variable files already exist for 1933, 1967, and 1972 Standard Industrial Classification (SIC) codes, 1970 and 1980 Census industry codes, and 1972 input-output industry codes. Similar descriptions are available for the different revisions of the International Classification of Diseases and Accidents (ICDA) codes, and the various occupation codes used by different government agencies. Geography, the one dimension already studied in detail, is the most complex.

It is interesting to note that additivity can apply in one dimension and not in another. For example, population is additive over geography but not time; inches of rainfall are additive over time but not geography; population density is additive over neither geography nor time, and industrial emissions are additive over both geography and time.

INTERPOLATION OF POINT DATA

Our discussion so far has considered the transformation of areal data from one set of geographic units to another. The geographic units may be represented as polygons which cover the area being studied. The methods described so far are not appropriate for the data measured at discrete sampling points, such as weather data, air or water quality data, or geologic survey data.

Theoretical methods for interpolating point measurements are extensively discussed in the literature. A comprehensive overview is provided by Ripley [6], who identifies five general methods for the smoothing and interpolation of point data: trend surfaces, moving averages, tessellations and triangulations, stochastic process prediction (kriging), and contouring. One of these methods, a moving average, has been implemented in SEEDIS in order to allow the user to easily estimate air quality at any desired geographic level. Input to the model is a data file of 1974-1976 air quality monitoring stations, along with their latitude and longitude coordinates. The model is a generalization of the method used in the PAREP project [2] to produce air quality estimates at the county and tract level [7,8]. The county level estimates have been widely circulated and are included in the Geocology data base [1].

Numerous studies in the last decade have attempted to relate human health to air quality, using, for example, mortality rates and air quality estimates at the SMSA, county, or tract level. Such applications are concerned only with long-term effects on a mobile population, so extreme detail in time and space is unnecessary. The same model might be used to estimate, say, crop exposure to acid rain, or long-term changes in population density.

In a moving average model, a data value at any point is estimated as a weighted average of values measured at nearby points. The weight \( w(d) \) is a decreasing function of the distance \( d \); according to Ripley [6], common choices of \( w(d) \) are \( d^{-a} \), \( e^{-ad} \), and \( e^{-x^a} \), where \( a \) and \( r \) are positive constants. In the SEEDIS model, \( w(d) \) is of the third form, specifically:

\[
w(d) = e^{-d^a/d^2}
\]

The choice of \( w(d) \) was dictated by three criteria: (1) the estimated function should be smooth in the vicinity of the measured points; (2) the estimated function need not pass exactly through all the measured points; (3) the area integral of the estimated function should be finite, so that distant points can be ignored in the calculation.

An estimate of air pollution is obtained as a weighted average at any desired point, for example at the population centroid of a county or tract. Reasonable values of \( d_0 \) are dictated by data availability and consistency, by the variability of air quality over time, and by the mobility of the population. Studies are continuing at LBL to determine the validity of the model.

For the PAREP county level analysis [7] and the Geocology data base [1], estimates were calculated at county population centroids with \( d_0 = 20 \) kilometers. For the PAREP tract level analysis [8], estimates were calculated at tract centroids with \( d_0 = 10 \) kilometers. In the interactive SEEDIS model, the choice of centroids and the choice of \( d_0 \) are left to the user. For an area larger than a county, estimation at a single population centroid is not appropriate. In this case, a population-weighted average of county values should be calculated, using the procedure described earlier for aggregation of non-additive data.

SEEDIS OVERVIEW

SEEDIS, the Socio-Economic Environmental Demographic Information System, is an interactive data management and analysis system under development by Lawrence Berkeley Laboratory, a Department of Energy facility administered by the University of California. SEEDIS embodies some 60 person-years of integrated development under funding by the Department of Energy, the Department of Labor, and other government agencies.

SEEDIS provides researchers with easy access to data files installed in the system, including the 1970 Census of Population and several hundred other socioeconomic, environmental, demographic, health- and energy-related files. The 1980 Census is partially installed, and the rest will be available a week or two after its release from the Bureau of the Census. At LBL, the data base on tape in an automatic tape library (ATL) includes 2.5 billion data values, roughly 25 billion bytes or 250 tapes at 6250bpi. The archive will double with the addition of the 1990 Census. The most frequently used data, 85 million data values, are presently installed for rapid interactive access in a disk-based system. Later this year, caching mechanisms will be implemented to provide interactive access to the entire tape-based data inventory.

SEEDIS operates on a network of nine DEC VAX-11/780 computers, located in the San Francisco Bay Area, in the State of Washington in Washington DC, and in North Carolina. Other government agencies having similar equipment could link into the SEEDIS network or operate independent SEEDIS systems. For automatic
access to the LBL data archive, a DEENET link to one of
the nodes already in the network would be required.
Data can be archived anywhere in the network, except
for response time, SEEDIS at every node behaves as if
all the data were locally stored.

Users can easily extract data required for specific
applications. These data can be analyzed within SEEDIS
or exported in a simple self-describing format to other
computers. Conversely, users can load data from exter­
nal sources for analysis and graphic display in SEEDIS.
A variety of terminals, plotters, and film recording
devices, both black-and-white and color, are sup­
ported. SEEDIS is user-friendly and is completely
described on-line. Extensive printed documentation is
available.

SEEDIS, as part of a publicly funded research program,
is in the public domain. The State Data Program and
Survey Research Center (SDP/SRC) on the University of
California's Berkeley campus can provide standard
Census data reports and more specialized data extract­
sion services at cost. In the near future, it is hoped
that such services, as well as tape copies of SEEDIS
for installation on a DEC VAX-11/780 computer, will be
available through the National Energy Software Center
and the National Technical Information Service.

Even for installations not having a DEC VAX computer,
portions of SEEDIS will be useful in other geographic
information systems. The geocode description files, the
cross-level correspondence files, and the proxy variable
files used for aggregation and disaggregation are all
stored in self-describing ASCII format. Data dictionaries
are available for major public data files, including the
1974 Census of Agriculture, the 1977 County and City
Data Book, the 1977 Area Resource File, and the 1980
Census Summary Tape File 1. Geographic base files
used for mapping, although in a VAX binary format, can
be readily converted with a simple FORTRAN program to
any desired format.

Further information about SEEDIS, and copies of a 26-
page summary overview document [3] can be obtained
from: Ilona Einowski, Data Librarian, SDP/SRC, 2538
Channing Way, University of California, Berkeley CA
94720; Tel (415) 642-6571.

REFERENCES
1. Olson, H.L., Emerson, C.J., and Nungesser, M.K.;
Geocology: A County-Level Environmental Data Base
for the Conterminous United States; Oak Ridge National
2. Sacks, S.T., Selvin, S., and Merrill, D.W.; Building a
United States Data Base: Populations at Risk to Environ­
mental Pollution; Lawrence Berkeley Laboratory Report
LBL-9638; October 1979.
3. McCarthy, J.L., Merrill, D.W., Marcus, A., Benson, W.H.,
Gey, F.C., Holmes, H., and Quong, C.; SEEDIS: The
Socio-Economic Environmental Demographic Information
System; Lawrence Berkeley Laboratory Report
PUI-4564; September 1981 (draft).
UPGRADE User’s Manual (version 4.0), Sigma Data Com­
puting Corporation, January 1990.
5. Zimmerman, E.K.; The Evolution of the Domestic
Information Display System: Toward a Government Pub­
lic Information Network; Review of Public Data Use Vol.
4, No. 1; June 1980.
6. Ripley, B.D.; Spatial Statistics; John Wiley and Sons,
1981.
7. Selvin, S., Merrill, D., Kwok, L., and Sacks, S.; Ecologic
Regression Analysis and the Study of the Influence of
Air Quality on Mortality; Lawrence Berkeley Laboratory
Report LBL-12217; October 1981; submitted to Journal
of Environmental Economics and Management.
8. Selvin, S., Sacks, S., Merrill, D.W., and Winkelman, W.;
The Relationship Between Cancer Incidence and Two
Pollutants (Total Suspended Particulate and Carbon
Monoxide) for the San Francisco Bay Area; Lawrence
Berkeley Laboratory Report LBL-10847; June 1960.

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Description</th>
<th>Year</th>
<th>Units</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRA</td>
<td>Bureau of Economic Analysis regns</td>
<td>1977</td>
<td>9</td>
<td>p</td>
</tr>
<tr>
<td>BECHT</td>
<td>Bechtel energy model regions</td>
<td>1978</td>
<td>14</td>
<td>p</td>
</tr>
<tr>
<td>CENSUS</td>
<td>Census regions</td>
<td>1970</td>
<td>9</td>
<td>p</td>
</tr>
<tr>
<td>COAL</td>
<td>Coal supply regions</td>
<td>1970</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>FED</td>
<td>Federal regions</td>
<td>1970</td>
<td>10</td>
<td>p</td>
</tr>
<tr>
<td>NATION80</td>
<td>Nations, PIPS definitions</td>
<td>1960</td>
<td>233</td>
<td>p</td>
</tr>
<tr>
<td>NFC</td>
<td>Nat. Petroleum Council oil &amp; gas regns</td>
<td>1976</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>PAD</td>
<td>Petroleum district refinery regions</td>
<td>1970</td>
<td>7</td>
<td>p</td>
</tr>
<tr>
<td>WATER</td>
<td>Water Resources Council regions</td>
<td>1972</td>
<td>22</td>
<td>p</td>
</tr>
<tr>
<td>Level</td>
<td>Level Description</td>
<td>Year</td>
<td>Units</td>
<td>Map</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>AQCR</td>
<td>EPA air quality control regions</td>
<td>1972</td>
<td>247</td>
<td>P</td>
</tr>
<tr>
<td>BEA69</td>
<td>Bureau of Economic Analysis areas</td>
<td>1969</td>
<td>173</td>
<td>P</td>
</tr>
<tr>
<td>BEA77</td>
<td>Bureau of Economic Analysis areas</td>
<td>1977</td>
<td>165</td>
<td>P</td>
</tr>
<tr>
<td>BLS</td>
<td>BLS labor market areas</td>
<td>1979</td>
<td>437</td>
<td>P</td>
</tr>
<tr>
<td>NECMA77</td>
<td>New England county metro areas</td>
<td>1977</td>
<td>270</td>
<td>P</td>
</tr>
<tr>
<td>NECMA79</td>
<td>New England county metro areas</td>
<td>1979</td>
<td>275</td>
<td>P</td>
</tr>
<tr>
<td>NECMA81</td>
<td>New England county metro areas</td>
<td>1981</td>
<td>310</td>
<td>P</td>
</tr>
<tr>
<td>PUS70</td>
<td>Public Use Sample county groups</td>
<td>1970</td>
<td>408</td>
<td>P</td>
</tr>
<tr>
<td>SCSA79</td>
<td>Standard consolidated stat areas</td>
<td>1979</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>SCSA81</td>
<td>Standard consolidated stat areas</td>
<td>1981</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>SMSA71</td>
<td>Standard metropolitan stat areas</td>
<td>1971</td>
<td>246</td>
<td>S</td>
</tr>
<tr>
<td>SMSA73</td>
<td>Standard metropolitan stat areas</td>
<td>1973</td>
<td>267</td>
<td>S</td>
</tr>
<tr>
<td>SMSA75</td>
<td>Standard metropolitan stat areas</td>
<td>1975</td>
<td>276</td>
<td>S</td>
</tr>
<tr>
<td>SMSA77</td>
<td>Standard metropolitan stat areas</td>
<td>1977</td>
<td>283</td>
<td></td>
</tr>
<tr>
<td>SMSA79</td>
<td>Standard metropolitan stat areas</td>
<td>1979</td>
<td>288</td>
<td>S</td>
</tr>
<tr>
<td>SMSA81</td>
<td>Standard metropolitan stat areas</td>
<td>1981</td>
<td>320</td>
<td></td>
</tr>
<tr>
<td>WRASA</td>
<td>Water Res Council aggregate subareas</td>
<td>1972</td>
<td>140</td>
<td>P</td>
</tr>
<tr>
<td>WRSA</td>
<td>Water Resources Council subareas</td>
<td>1972</td>
<td>322</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Description</th>
<th>Year</th>
<th>Units</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEAPART77</td>
<td>STATE/BK/AQCR/NECMAB pieces</td>
<td>1977</td>
<td>585</td>
<td>P</td>
</tr>
<tr>
<td>STAQCR</td>
<td>STATE/AQCR pieces</td>
<td>1972</td>
<td>314</td>
<td>P</td>
</tr>
<tr>
<td>STBEA69</td>
<td>STATE/BEA69 pieces</td>
<td>1969</td>
<td>290</td>
<td>P</td>
</tr>
<tr>
<td>STBEA77</td>
<td>STATE/BEA77 pieces</td>
<td>1977</td>
<td>285</td>
<td>P</td>
</tr>
<tr>
<td>STLMA</td>
<td>STATE/LMA pieces</td>
<td>1979</td>
<td>432</td>
<td>P</td>
</tr>
<tr>
<td>STNECMA77</td>
<td>STATE/NECMAB pieces</td>
<td>1977</td>
<td>383</td>
<td>P</td>
</tr>
<tr>
<td>STNECMA79</td>
<td>STATE/NECMAB9 pieces</td>
<td>1979</td>
<td>368</td>
<td>P</td>
</tr>
<tr>
<td>STNECMA81</td>
<td>STATE/NECMAB1 pieces</td>
<td>1981</td>
<td>404</td>
<td>P</td>
</tr>
<tr>
<td>STPUS70</td>
<td>STATE/PUS70 pieces</td>
<td>1970</td>
<td>335</td>
<td>P</td>
</tr>
<tr>
<td>STSCSA81</td>
<td>STATE/SCSA81 pieces</td>
<td>1981</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>STSMSA71</td>
<td>STATE/SMSA71 pieces</td>
<td>1971</td>
<td>334</td>
<td>S</td>
</tr>
<tr>
<td>STSMSA73</td>
<td>STATE/SMSA73 pieces</td>
<td>1973</td>
<td>365</td>
<td>S</td>
</tr>
<tr>
<td>STSMSA75</td>
<td>STATE/SMSA75 pieces</td>
<td>1975</td>
<td>373</td>
<td>S</td>
</tr>
<tr>
<td>STSMSA77</td>
<td>STATE/SMSA77 pieces</td>
<td>1977</td>
<td>381</td>
<td></td>
</tr>
<tr>
<td>STSMSA79</td>
<td>STATE/SMSA79 pieces</td>
<td>1979</td>
<td>386</td>
<td>S</td>
</tr>
<tr>
<td>STSMSA81</td>
<td>STATE/SMSA81 pieces</td>
<td>1981</td>
<td>423</td>
<td></td>
</tr>
<tr>
<td>STWATER</td>
<td>STATE/WATER pieces</td>
<td>1972</td>
<td>110</td>
<td>P</td>
</tr>
<tr>
<td>STWRSA</td>
<td>STATE/WRSA pieces</td>
<td>1972</td>
<td>237</td>
<td>P</td>
</tr>
<tr>
<td>STWRSA</td>
<td>STATE/WRSA pieces</td>
<td>1972</td>
<td>368</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Description</th>
<th>Year</th>
<th>Units</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD66</td>
<td>88th Congr Congressional Districts</td>
<td>1986</td>
<td>436</td>
<td></td>
</tr>
<tr>
<td>CD61</td>
<td>91st Congr Congressional Districts</td>
<td>1989</td>
<td>436</td>
<td></td>
</tr>
<tr>
<td>CD64</td>
<td>94th Congr Congressional Districts</td>
<td>1976</td>
<td>436</td>
<td>P</td>
</tr>
<tr>
<td>CD65</td>
<td>95th Congr Congressional Districts</td>
<td>1976</td>
<td>436</td>
<td>P</td>
</tr>
<tr>
<td>CD66</td>
<td>96th Congr Congressional Districts</td>
<td>1976</td>
<td>276</td>
<td></td>
</tr>
<tr>
<td>LMPM</td>
<td>Labor Market Prog Model areas</td>
<td>1979</td>
<td>1280</td>
<td></td>
</tr>
<tr>
<td>BEA70</td>
<td>State economic areas</td>
<td>1970</td>
<td>510</td>
<td>P</td>
</tr>
<tr>
<td>STATE</td>
<td>States and territories</td>
<td>1970</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>Level Description</td>
<td>Year</td>
<td>Units</td>
<td>Map</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td><strong>County Levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNTY70BO</td>
<td>COUNTY/COUNTY70 pieces</td>
<td>1980</td>
<td>3226</td>
<td>p</td>
</tr>
<tr>
<td>COUNTY</td>
<td>Counties, U.S. Census Bureau</td>
<td>1970</td>
<td>3225</td>
<td>p</td>
</tr>
<tr>
<td>COUNTY90</td>
<td>Counties, U.S. Census Bureau</td>
<td>1980</td>
<td>3253</td>
<td>p</td>
</tr>
<tr>
<td>MSP</td>
<td>Counties, Johns Hop Mort Surv Prog</td>
<td>1970</td>
<td>3075</td>
<td>p</td>
</tr>
<tr>
<td>NCHS</td>
<td>Counties, Nat Center Health Stat</td>
<td>1970</td>
<td>3022</td>
<td>p</td>
</tr>
<tr>
<td>NC1</td>
<td>Counties, National Cancer Inst</td>
<td>1970</td>
<td>3061</td>
<td>p</td>
</tr>
<tr>
<td><strong>Place and MCD Levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCD70</td>
<td>Census Minor Civil Divisions</td>
<td>1970</td>
<td>35196</td>
<td>s</td>
</tr>
<tr>
<td>MCD80</td>
<td>Census Minor Civil Divisions</td>
<td>1980</td>
<td>35197</td>
<td>-</td>
</tr>
<tr>
<td>MCDPT83</td>
<td>MCD80/PLACE80 pieces</td>
<td>1980</td>
<td>53032</td>
<td>-</td>
</tr>
<tr>
<td>PLACE</td>
<td>Places, Census, pop over 1000</td>
<td>1970</td>
<td>11970</td>
<td>s</td>
</tr>
<tr>
<td>PLACE80</td>
<td>Places, Census</td>
<td>1980</td>
<td>19144</td>
<td>-</td>
</tr>
<tr>
<td>PLBLS</td>
<td>Places, Bureau of Labor Statistics</td>
<td>1979</td>
<td>1585</td>
<td>-</td>
</tr>
<tr>
<td>PLEPA</td>
<td>Places, Env Protection Agency</td>
<td>1972</td>
<td>7745</td>
<td>s</td>
</tr>
<tr>
<td>TOWNSHIP</td>
<td>Bureau Labor Statistics townships</td>
<td>1979</td>
<td>112</td>
<td>-</td>
</tr>
<tr>
<td><strong>Small Subcounty Levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AQMS</td>
<td>EPA air qual monitoring stations</td>
<td>1976</td>
<td>8625</td>
<td>s</td>
</tr>
<tr>
<td>AQMSLOC</td>
<td>EPA air qual monitoring locations</td>
<td>1976</td>
<td>5777</td>
<td>s</td>
</tr>
<tr>
<td>EGED70</td>
<td>Census enum dists &amp; block groups</td>
<td>1970</td>
<td>249189</td>
<td>s</td>
</tr>
<tr>
<td>EDBG80PT</td>
<td>EDBG80/URBARURL/CD86 pieces</td>
<td>1980</td>
<td>322000</td>
<td>-</td>
</tr>
<tr>
<td>PREDDBG80</td>
<td>EDBG80/CD86 pieces</td>
<td>1980</td>
<td>319105</td>
<td>-</td>
</tr>
<tr>
<td>TRACT</td>
<td>Census Tracts</td>
<td>1970</td>
<td>34869</td>
<td>p</td>
</tr>
<tr>
<td>TRACT80</td>
<td>Census Tracts</td>
<td>1980</td>
<td>48475</td>
<td>-</td>
</tr>
<tr>
<td>TRACT80PT</td>
<td>MCD80/PLACE80/TRACT80 pieces</td>
<td>1980</td>
<td>75000</td>
<td>-</td>
</tr>
</tbody>
</table>