DATA FILES FOR COMPUTERIZED CARTOGRAPHY
FROM THE U.S. CENSUS BUREAU

Larry W. Carbaugh
U.S. Bureau of the Census

As a result of the 1970 census, the Bureau of the Census has developed and re-
leased hundreds of machine-readable data files describing the characteristics of the
population of the United States. In addition, other machine-readable files have
been developed which describe the location and extents of the areas from which these
statistics were gathered. The release of such files provides researchers with the
means of analyzing the characteristics of the populations and the spatial relation-
ships of these characteristics through the use of the computer. The purpose of this
paper is to outline the contents of these files, highlight some computer display
programs developed by the Bureau, and to provide some comments on the distribution
and support problems associated with release of machine-readable data.

DATA FILES

Because of the cost and volume of data, the Bureau of the Census cannot publish
all the data summaries that it tabulates, nor can it provide data for all the areas
recognized in a census. However, by producing the summaries on computer tape, the
Bureau can release large amounts of data for many geographic areas.

In general, the files that are available from the Census Bureau contain data
summarized by geographic area. Data are provided for the political subdivisions of
the United States such as States, counties, minor civil divisions, and places, as
well as for other areas such as blocks, census tracts and census enumeration dis-
tricts. The data presented include characteristics of the population of each area
such as age, sex, race, family income, educational attainment, employment status,
occupation, and industry. Also included are data on the housing stock such as va-
cancy status, tenure, plumbing facilities, rent and value. A complete description
of the items contained on the data files is presented in the 1970 Census User's
Guide Parts I and II, which is available from the Bureau of the Census for $5.25.

POINT OR AREA DESCRIPTION FILES

In addition to the computer tapes containing data, the Bureau maintains tapes
of geographic reference files. These files contain identifiers such as latitude and
longitude coordinates which locate the centers or the boundaries of the areas pre-
sented on the data files mentioned above. They include:

U.S. Centers of Population - a one-tape file containing the latitude-longitude
coordinates for the 1970 census population centers of counties in the U. S.

<u>Master Enumeration District List</u> (MEDList) – contains latitude-longitude coordinates for each enumeration district (ED) identified in the 1970 census. Each ED record also contains codes for all higher level areas. Nine reels of tape.

<u>DIMECO</u> – contains latitude-longitude coordinates describing the boundaries of all counties in the United States. One reel of tape.

<u>GBF/DIME</u> – a "computerized map" for the urbanized portion of 200 metropolitan areas, containing latitude-longitude and state plane coordinates for the intersections of all map features (i.e., streets, railroads, streams, etc.). Codes identifying census tracts and blocks are contained on these files.

<u>SMSA census tract boundaries</u> – files containing latitude-longitude coordinates describing the boundaries of census tracts in all standard metropolitan statistical areas recognized in the 1970 census.

COMPUTER PROGRAMS

In an effort to extend the usefulness of these census files, the Bureau has also developed several computer programs to display geographically identified data. They include a printer-oriented computer mapping system, GRIDS, and a line printer-oriented program, CENPLOT. In addition, the Census Bureau has initiated a clearinghouse for information on computer tapes designed to process 1970 census products such as the geographic reference files. Further information about this clearinghouse activity can be obtained from the Data User Services Division, Bureau of the Census, Washington, D.C. 20233.

OBTAINING THE DATA FILES

In general, the computer files mentioned above are available at a standard cost of $80 per tape reel. Technical documentation is included with each purchase of a file, but it can also be purchased separately for $3 per file. Additional information on the tape files available from the Census Bureau can be obtained from: Chief, Customer Services Branch, Data User Services Division, Bureau of the Census, Washington, D.C. 20233.

DISTRIBUTION OF COMPUTERIZED DATA AND SOFTWARE

Before I close, I would like to comment on the development of computerized data files and programs and some of the problems associated with their distribution. While the development of machine-readable data and computer programs has increased the capability of researchers to analyze the complexity of today's world, the release of such data files carries with it an obligation to provide some form of assistance to users of the files. This obligation includes the development of adequate documentation, notification of errors or updates, and, in the case of programs, some installation consultation. This obligation is further complicated by the wide variety of computing equipment now available. Therefore, unless this obligation is carefully considered beforehand, and adequate staff and proper materials are prepared, users of such data files will become extremely frustrated and the usefulness of the data will be diminished.