

CONTENTS

CHAPTER 13. Research and Assistance in Data Use

(Page numbers here omit the chapter prefix, 13 – , which appears as part of the number of individual pages. The prefix indicates the location of the material in the final consolidated edition of the **Procedural History**.)

	Page
Introduction	1
Census Use Studies	2
The New Haven Census Use Study	2
Preparations	2
Basic Program and Resources	3
Record Matching	4
Health Information System	4
Computer Mapping Research	4
Development of the Dual Independent Map Encoding (DIME) System	8
Address Coding Guides (ACG's)	8
Establishing a Geographic Base File	8
Development of the DIME File	9
Southern California Regional Information Study	11
Mandatory Work Projects	11
Optional Work Projects	11
Record Matching	12
Resource Application	12
Computer Mapping	12
Census Data Processing Programs	12
Data User Services	13
Summary Tapes	13
Public-Use Samples	15
Geographic Reference Materials	15
Computer-Generated Reports	16
Publications	19
Census User Conferences	20
Summary Tape Processing Center Program	22
Summary Tape Information Library Program	22
Users' Service Staff	22
Reference Retrieval	23
Bibliography	24

Chapter 13. RESEARCH AND ASSISTANCE IN DATA USE

INTRODUCTION

The Bureau of the Census has prepared printed reports as the primary data product of each census it conducted. In addition, since its establishment as a permanent agency in 1902, the Bureau also has made available summaries of data which are not included in the published reports. Title 13 of the United States Code, the statute under which censuses have been taken since 1954, authorized "special" tabulations from basic census records and recognized the necessity for the Bureau to perform some kind of rudimentary data-access service. The computerization of the 1960 censuses greatly increased the Bureau's ability to deal with special requests: For the first time it was possible to generate for internal use a much larger set of tabulations than was to be published, and to provide unique tabulations of much greater variety for outside users. As proliferation of electronic computing machinery gave users of summary census data the capacity for processing and using much more data, their requests for special tabulations increased as well. In 1963 a Bureau staff member was assigned to coordinate all users' requests, and in 1966 a program was initiated to improve accessibility to census data. A principal feature of this program was the establishment of a central documentation office to identify, index, define, and maintain the documentation for selected tape files.

During the period 1963-1966 Census Bureau officials and advisory groups composed of persons from outside the Bureau also gave attention to the ways in which 1960 census data were being used and how adequate a resource they were proving to be. A Task Force on Uses of Census Statistics was appointed by the Director of the Bureau of the Census in 1963. As a result of the findings of the Task Force and recommendations made by the Census Advisory Committee of the American Statistical Association, a formal program of research on data uses was launched in the spring of 1964. The focus of this effort extended beyond investigation of specific uses to include (1) evaluation of the appropriateness of available census data in relation to the needs of the users, and (2) the effectiveness of promotion and dissemination of data to potential users.

To advise the Census Bureau on the needs for small-area data, the Secretary of Commerce established the Census Advisory Committee on Small Areas in the spring of 1965. This committee, composed of planners, research specialists, and Federal agency and other governmental representatives, provides guidance in matters relating to census data for small areas. One of the most apparent needs identified through this effort was the users' desire for additional data tabulated by the smallest feasible geographic unit, particularly the individual city block in metropolitan areas.

Soon after being formed, the Committee recommended that the Bureau establish a research and development

project in a metropolitan area to explore the needs for small-area data, and methods for using them, in advance of the 1970 census. The Bureau accepted this suggestion and, in January 1966, began to develop plans for such a study (called the Census Use Study) with the assistance of interested Federal agencies. This study, which began in New Haven, Conn., in 1966 (see pp. 2-5.) and was continued at Bureau headquarters in Suitland, Md., indicated that census users' requirements for data went beyond the simple need for statistics; tools and techniques were necessary to relate these statistics to local data and to display the results graphically or in some other form for interpretation. The Bureau's research and development efforts were expanded through the Southern California Regional Information Study (SCRIS) in Los Angeles, Calif., in 1969 (see pp. 11-12. for description), and the Unified Statistical Evaluation Study (USES) in Indianapolis, Ind., in 1971. (The objective of the USES was to develop the use of census tools and techniques in the eight-county Indianapolis area.) Special attention was given to the use of computers, and systems were developed for geographic coding of locations and data; for matching, comparing, and analyzing records from various sources; for generating maps by computer, and for displaying data on these maps.

In 1966, also, one member of the Bureau staff was given primary responsibility for the improvement of data access and for refinement of a data delivery system for the 1970 census. While many parts of the Bureau contributed to this effort,¹ two subsequently created offices were charged with special responsibilities in assuring that 1970 census data users would have satisfactory access to census data and the information and tools necessary to use them. They were the Data Access and Use Laboratory and the Central Users' Service.

The Data Access and Use Laboratory (DAUL), organized in October 1967, carried out a variety of communications, product development, and documentation activities which were associated primarily with the new, computer-related products planned by the Bureau. It established contact with hundreds of potential users of 1970 census

¹The many complexities of data content and delivery in connection with the 1970 census and the necessity for formal inter-divisional coordination within the Bureau were recognized as planning progressed. In response, the Tabulation and Data Delivery System Planning Group ("Tab Group") was organized in November 1967 with representatives of the Population, Housing, Geography, Systems, Statistical Methods, and Field Divisions, and the Data Access and Use Laboratory. The Tab Group's responsibilities included the following: (1) establishing the components of the Bureau's tabulation and data delivery system; (2) setting deadlines for the delivery of specifications and finished products from the appropriate divisions; and (3) determining priorities for producing statistics in both published and unpublished forms. The Tab Group made recommendations on such basic issues as the extent of data to be reported in the first count tape series, the file structure of the fourth count tape series, and rules for suppression of data that might allow identification of individual persons, households, or housing units.

summary tapes well in advance of the census. DAUL was a major contributor to the planning and development of the 1970 census data delivery system; it assisted in the preparation of technical documentation of the summary tape files, promoted plans for microfilm copying of tape data summaries, and coordinated and developed the 1970 public use sample. As they related to the use of 1970 census materials, these activities are described on pp. 13-23.

The Central Users' Service, established in the fall of 1969 and since 1972 known as the Users' Service Staff, had as its principal mission arrangements for the filling of customer orders for data in any form other than printed volumes, i.e., special tabulations, maps, computer programs, summary tapes, and other data in machine-readable form. This service is described in greater detail on pp. 22-23.

From the mid-1960's on, especially, Bureau staff members participated in many conferences of national and regional associations of professional planners and

statisticians, as well as of census users, which pertained to the 1970 census and its products, and frequently presented papers on the Bureau's research and data-delivery efforts. A selection of these papers appears in the bibliography on p. 25. Consultation became frequent with other Federal agencies and with representatives of foreign countries engaged in development of computer graphics and geographic coding. Beginning in 1969 a number of reports and manuals were published by the Bureau and by the SCRIS describing the various census use study projects and the results of applying the techniques which had been developed. These also are listed in the bibliography.

The Bureau's principal efforts to carry out research in the area of small-area data and its applications, and to assist census users, are described in the following sections of this chapter. The methods which the Bureau developed to allocate its own data to small areas through geographic coding, and to publish the data in machine-readable form, are subjects treated in other parts of the procedural history.

CENSUS USE STUDIES

THE NEW HAVEN CENSUS USE STUDY

Preparations

It was decided in 1966, while plans were going forward for a full-scale test of proposed 1970 census procedures, that the best site for this test and for the first census use study would be a standard metropolitan statistical area (SMSA) of not more than 500,000 inhabitants, containing a central city of at least 100,000 persons, and convenient to Washington, D.C. Potential sites for the use study were suggested and evaluated by interested Federal agencies and by the Census Advisory Committee. The area selected as most suitable for both the test and the study was the New Haven, Conn., SMSA.

The census pretest would provide a body of current census data comparable to that which would be available from the 1970 census. Observation of local uses of these census data was expected to be helpful in evaluating the eventual impact and utility of the 1970 census program. In February 1966 the mayor of New Haven and representatives of other local agencies in the New Haven area were invited to cooperate in the research project. These local officials endorsed the project and pledged their cooperation. The pretest was conducted in New Haven in April 1967. In this pretest, basic population and housing statistics were gathered for all households in the New Haven SMSA; additional questions were asked at every fourth housing unit. (For details, see, U.S. Bureau of the Census, Pretests and Dress Rehearsals of the 1970 Census of Population and Housing, Working Paper No. 32 (1972).) Tabulation plans for the pretest census were disseminated through meetings held in New Haven with local organizations shortly after the Census Use Study began operating in September 1966. These meetings were designed to assist local agencies in determining how they could use census data in their operations and how to formulate requests for data tabulations. The meetings revealed that many of the

potential data users, particularly smaller local agencies, had little knowledge of the nature, use, or format of census data.

The meetings further indicated that many users' expressed needs for census data could not be met by standard Census Bureau tabulations, either in published or unpublished form. Users repeatedly stated that tabulations tailored to their specialized needs would be a significant contribution to the specific information needed to analyze urban problems. Based on these findings in New Haven, the Census Use Study staff undertook a number of new tabulation programs.

Participating agencies.--Since the Census Use Study was originally conceived as a cooperative effort among Federal, State, and local interests, several interested agencies agreed to provide the major share of the financial support, direction, and evaluation. The cost of the pretest census, considered a vital part of the preparation for the 1970 decennial census, was borne by the Bureau of the Census. A consortium of Federal agencies (the Economic Development Administration of the Department of Commerce; the Office of Civil Defense of the Department of Defense; the Department of Health, Education, and Welfare; the Department of Housing and Urban Development; and the Department of Transportation) funded the ensuing research and development activities. The city of New Haven provided substantial office space, staff personnel, computer time, and equipment and facilities to the project.

Several State and local agencies provided funds or services in return for tabulations. Over 40 separate participants from the New Haven area were connected with the use study, and in addition a number of private research organizations also contributed advice and talent. Computer programming assistance and processing time were provided by several university-affiliated and

private contractors, and a university laboratory assisted in the mapping experiments.

Staffing and project development.--The use study staff in New Haven was comprised of 14 professional and clerical employees, supplemented by four computer programmers loaned by a local redevelopment agency. As use study activities expanded and technical developments accelerated, an office also was established at Census Bureau headquarters in Suitland, Md. This office was assigned the primary task of testing some of the New Haven products in other cities. Additional technical refinements for the program were developed in Suitland and the production of reports was directed from there. The office in New Haven was closed in the summer of 1969, although a part-time Census Bureau employee continued to work from his own home in New Haven after that time on the New Haven Health Information System (see below) to complete this project.

Basic Program and Resources

In March and September 1967, two surveys were conducted in which approximately 60 organizations in the SMSA were identified as users of census statistics. Preliminary contacts with these organizations involved orientation seminars and personal visits designed to describe the basic purposes of the use study and to gather information about each organization. This information enabled the staff to evaluate the data interests and needs of agencies, so that an effective work program could be developed. Additional data-user surveys conducted during the duration of the use study provided feedback to the staff on the effectiveness of their research activities, and also contributed further to the Census Bureau's knowledge of the needs of the census data users.

Following the first data-user survey, the basic program outlined below was established:

1. Exploration of current uses and probable future needs for census data in existing local, State, and Federal programs.
2. Development of a system to assist in inter-relating census data with local and State data to meet specific needs. Initial contacts had revealed that many agencies wished to inter-relate such data, but were hindered by the differing definitions of geographic areas used by the agencies in collecting and tabulating the data.
3. Investigation of the possibilities for cooperative data collection by the Census Bureau and other Federal, State, and local agencies, with special emphasis on preparation of guidelines and development of uniform terms and procedures.
4. Study of the adequacy of census data tabulations for local use, with particular attention to the level of detail and the media of presentation (e.g., magnetic tape, printout).
5. Research on the feasibility of developing computer-mapping techniques for local agency display of census or other data.

6. Analysis of local use of census data, and development of procedures which could be incorporated into local community programs to facilitate additional use, or improved use, of census data.

The local agencies were encouraged to develop research projects which utilized census information. These projects were intended to help the users become better acquainted with census data as well as to identify needs that could not be fulfilled unless census procedures were modified. The use study staff also analyzed the extent to which local agency needs in New Haven were similar to those of other cities. The work program of the use study reflected this dual orientation towards national as well as local needs, so that the programs and procedures to be developed would have broad application and usefulness.

The information that was offered for tabulations came from two sources: First, from the local school census and from the records of local municipal agencies and companies; second, from the special census of the New Haven SMSA.

Data-user projects involved five general stages of work by the use study staff: (1) initial meetings with local agencies and the data-user survey; (2) preparation of final specifications for individual projects developed in cooperation with the agencies; (3) in some cases, developing special surveys to collect the data; (4) data processing, involving computer programming and production of tabulations; and (5) delivery of the tabulations and a followup survey of the agencies to determine their reactions and planned uses for the data.

Some of the planned uses of the data reported to the Study staff in the project development meetings were:

1. Education planning--school districting, school program planning, analysis of local school censuses, and general socioeconomic studies of the community.
2. Health planning--definition of areas for proposed intensive medical care units and child day care centers, examination of the location and nature of child health problems, analysis of characteristics of areas associated with deficient birth weights and child disease, general survey of family health needs, establishment of a data base for a regionwide hospital planning study, and identification of areas with unmet needs for health service.
3. Transportation planning--development of comprehensive traffic models, studies of locations of generators of heavy traffic, studies of parking facilities, estimates of highway needs for shopping trips and trips to work, and studies of drivers' characteristics.
4. Housing and redevelopment planning--study of the records of complaints to the New Haven City Division of Neighborhood Improvement, housing demand-supply models, study of housing changes from 1960 to 1967 on a small-area basis, study of the population in different types of government-supported housing, general studies for renewal

areas and Community Action Program areas, and study of migration into the city.

Two special surveys were conducted in the New Haven SMSA during the autumn of 1967, soon after the census pretest in New Haven, in order to combine the special-purpose data obtained by the surveys with the census data on an area basis. One, concerned with family health, produced some data for use in a health information system (see below), and the other, an area travel survey, was an attempt to apply small-scale survey techniques to transportation planning.

Record Matching

It was found that local agencies were manually coding their own data to census areas in order to relate their various machine-readable data files to each other and to census data. Agencies also applied geographic codes to their data for local administrative and planning areas, such as community action program areas, school districts, and traffic zones. Research in the Census Bureau indicated that such coding could be assisted by computer. A system called ADMATCH/DOS was developed jointly by the Census Bureau and a private contractor for the New Haven Census Use Study; it was used on a computer with a disk operating system (DOS) to match street addresses with geographic codes for census tracts and blocks contained in the Bureau's address coding guide for New Haven. (For further development of record matching techniques outside the New Haven Census Use Study, see p. 12.)

Health Information System

Late in 1966, the Connecticut State Department of Health, the Children's Bureau of the U.S. Department of Health, Education, and Welfare, and the Census Bureau began the development of a maternal and child health information system for New Haven. The project involved linking by computer approximately 300 data summaries from five independent machine-readable data files for 107 of the 120 block groups (i.e., enumeration districts) in New Haven, and subjecting the summaries to various analyses. The five data files were the following:

1. The complete-count population and housing census statistics from the special census of New Haven.
2. The 25-percent sample data from the same census.
3. Birth records for the city of New Haven for the calendar years 1967 and 1968, together with certain infant and fetal death records.
4. Approximately 1,900 hospital obstetrical records for 1966-1967.
5. Data from a family health survey conducted in New Haven by the use study staff in the fall of 1967. This survey was used to gather data in maternal and child health subject-matter areas not covered by the files 1-4 above. The subjects covered in the family health survey were as follows:
 - (1) utilization of community health resources and

health manpower for the surveillance of health; (2) health status--the distribution of morbidity and the extent to which normal activities of the population were impaired; (3) childbearing expectation as a measure of potential population growth; and (4), the need for and utilization of day-care services.

These five sources of information were linked into a system by geographically coding the addresses to block-side location. In the case of the census data, the housing units were identified by unique address serial numbers; the addresses of the other records were coded through the ADMATCH address matching system. For purposes of geographic analysis, the coded data were related to the geographic base (DIME) file (see p. 9). This allowed mechanical identification of health data by geographic coordinates, calculation of areas, and aggregation of data to those areas in forms capable of being printed out by computer either graphically, in tabulations, or as maps. In the health information system, as in others involving the use of basic census records, the confidentiality of the census questionnaires was fully maintained.

The Bureau continued the development of this system for use with 1970 census data in New Haven, Los Angeles (see p. 11), and Suitland; the system was adopted wholly or in part by a number of agencies in all areas of the United States. In 1972 a long-term project entitled "Development of a System of Health, Social and Resource Indicators" was begun at the Census Bureau in conjunction with the U.S. Office of Economic Opportunity.

Computer Mapping Research

During the progress of the Bureau's program in New Haven, experimentation with computer mapping techniques was carried out simultaneously with the various surveys that were conducted. This experimentation brought about a new geographic coding system called Dual Independent Map Encoding (DIME) (see p. 8ff).

The use study experimented with two general types of computer-mapping techniques, those associated with "character printing" equipment and those associated with "line drawing" equipment with the following objectives:

1. Developing methods for relating census data and local records to geographic coordinates.
2. Experimenting with construction of statistical graphs, charts, and maps by computer, using both local and census data.
3. Developing a package of computer programs and instructions for producing maps.

Three distinct tasks were formulated to further these objectives:

1. Constructing and refining a base file containing all geographic identifiers necessary for computerized coding of geographic data to any given location covered by the file.
2. Testing available mapping packages.
3. Developing new mapping techniques for use with existing graphics hardware.

These investigations, extending for a period of more than a year, led to the development of tools designed to fulfill a number of geographic data coding and mapping requirements of local data users.

The computer mapping process.--Map specifications often required modification to fit the programs, computers, display devices, and data which were available in a particular situation. About 80 percent of the use study staff's time was spent on the preparatory stages of the mapping activities, and only 20 percent of the time in actually testing programs and equipment and in producing the maps themselves. Mapping experiments performed at the university laboratory and elsewhere, using records of the New Haven Department of Health, produced maps displaying the incidence of births.

A variety of experiments were performed on a file of 29,000 punchcards containing data on time spent by visiting nurses at different locations in the city. The use of the ADMATCH system and a modification of the DIME file made possible the assignment of individual coordinates to 98.6 percent of the data records in less than 2 hours. Data values in the file were tallied to census tract and block, as well as to grid cells representing 1,000-foot squares. The aggregated visiting-time data were then mapped using a number of the mapping techniques tested, namely, SYMAP (see fig. A), MAP 01, a two-dimensional histogram program, and a dot map program (see fig. B). A three-dimensional physical model showing the data distribution was also built. In all these experiments, the mapping process proved to be simple because of the ease with which the geographic codes were assigned to each data record.

A useful byproduct of the many matches performed on local data was the discovery of errors and omissions in address information in the DIME file. These errors were corrected with file update programs.

Selecting appropriate data to be mapped.--Two criteria determined the appropriateness of data used in the Bureau's computer-mapping experiments: First, that the

data could be geographically indexed by address or coordinates; second, when census data were used, that the confidentiality standards of the Census Bureau were maintained, thus guarding against the possible identification of individuals or families. When incidence maps were produced as part of the use study, nonconfidential local records were used, and only summary data from the census were used in mapping.

Items displayed included percentages and densities in the total population of females, persons and families of races other than white, large families, preschool children, and adults 65 years of age and older. Thus one map might show, for each block in a town, the number of persons of races other than white in the block expressed as a percent of the total population of the block, while another map might show the number of preschool children per acre in each block.

Data user surveys.--The two initial data-user studies conducted in New Haven (see p. 3) indicated that most respondents did not have well-defined mapping programs. Of those which did use maps regularly, however, a substantial majority thought that the lack of maps weakened the Census Bureau's own publication program.

After the computer maps representing the techniques tested by the Bureau had become available, projects were launched to ascertain user responses and evaluation of the maps produced. A series of interviews of personnel of 23 local planning and administrative organizations was conducted by a representative of one of the use study participants, and an urban atlas with a related questionnaire was distributed to over 100 persons throughout the country who were believed to be frequent users of data maps. In almost all cases, respondents agreed that computer mapping technology could assist them in administration and planning, but that the cost of obtaining the maps, in terms of both time and money, would be the major criterion they would apply. (For the Bureau's development of computer mapping systems outside the New Haven Census Use Study, see p. 12.)

**Figure A.--SYMAP Shading Map Showing Grid Cell Levels--
Adjusted Hours Spent by Visiting Nurses Association**

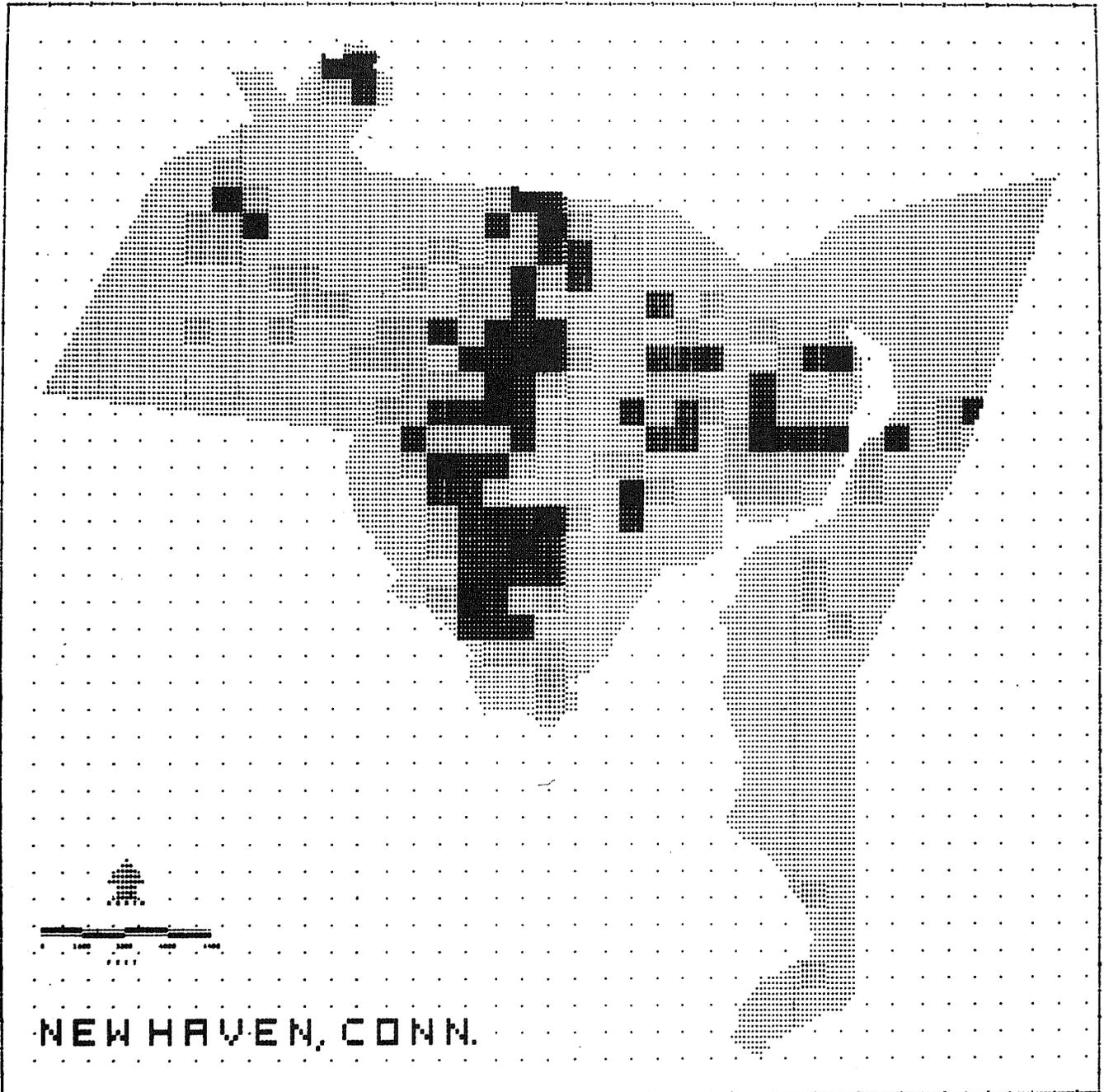
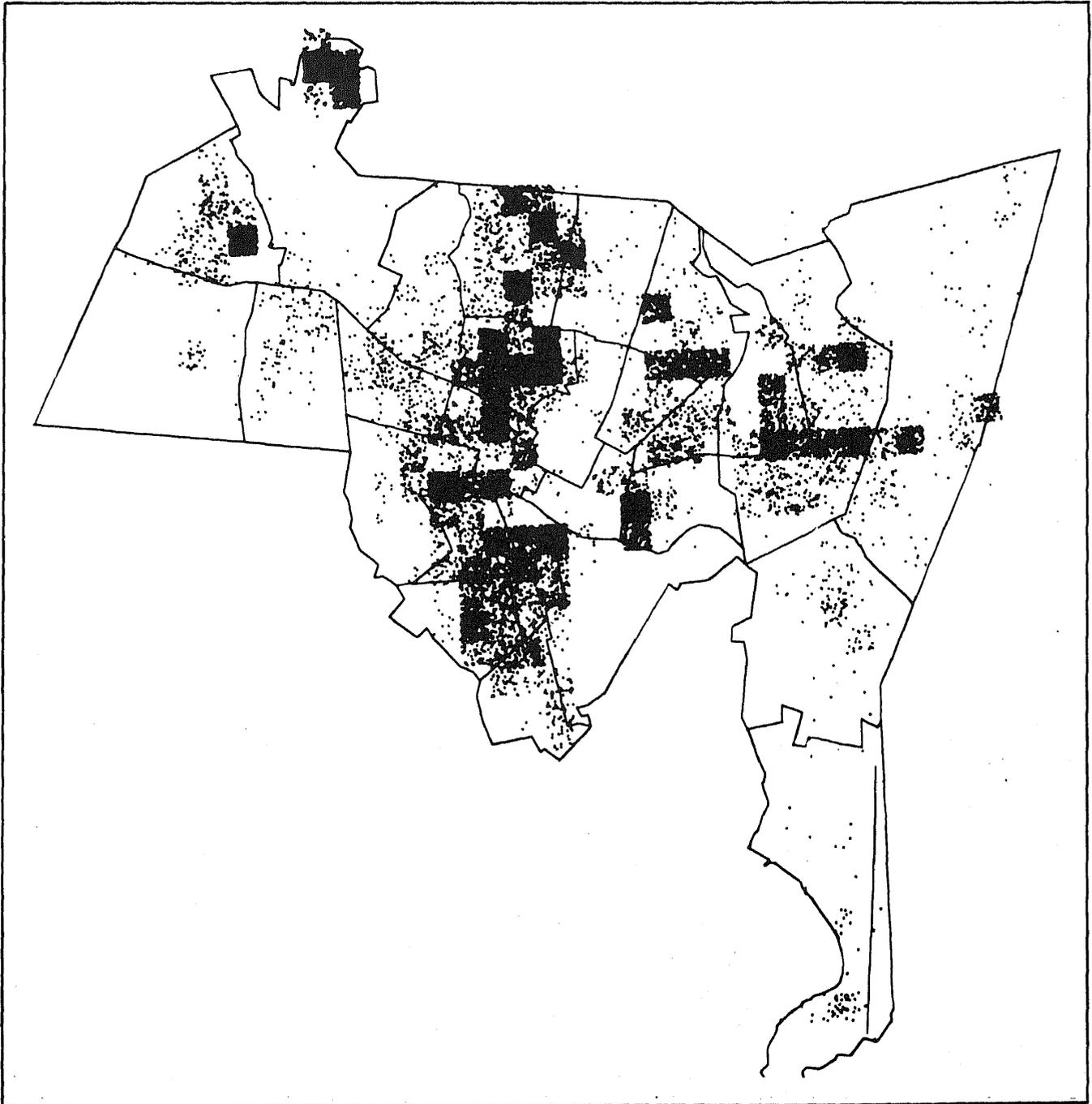


Figure B.--Advanced CRT Outline Map Showing Grid Cell Levels--Adjusted Hours Spent by Visiting Nurses Association



DEVELOPMENT OF THE DUAL INDEPENDENT MAP ENCODING (DIME) SYSTEM

Address Coding Guides (ACG's)

Fundamental to urban planning and analysis is the geographic location of activities, facilities, and conditions throughout the city. Precise geographic identification is particularly important for the analysis of data about small areas, such as groups of blocks comprising police beats, school attendance areas, health districts, and other locally designated special areas. Recognizing this need of local officials, and because one of the requirements of the 1970 census was coding to specific locations, the Census Bureau developed a program to prepare address coding guides (ACG's) for the urbanized areas of 145 SMSA's which were to be enumerated by the mail-out/mail-back procedure in 1970. (The Bureau's geographic work is described in chapter 3.) Each record in the ACG describes a block face (side) by the street name, the first and last possible addresses along the block face, various geographic codes such as census tract, block, municipality, State, etc. An ACG was compiled for the urbanized portion of the New Haven SMSA for use in the 1967 pretest.

Establishing a Geographic Base File

In the fall and winter of 1966-1967, Bureau staff members examined the New Haven ACG as a possible tool for use in computer mapping and in matching various local record files with each other and with census records (fully maintaining the confidentiality requirements in the use of census data). This file, however, did not include grid coordinates, which are the reference points necessary to produce graphic displays or computer-generated maps of geographic areas composed of city blocks. Blocks usually are represented as being comprised of vertical and horizontal lines, such as those on a sheet of graph paper, from which linear measurements can be made for some point or points of reference. Therefore, one of the first tasks of the use study was to create a suitable file incorporating street addresses, census tract and block numbers, complete topology of the metropolitan maps, and grid coordinates. This new approach to developing a geographic base file was later called Dual Independent Map Encoding (DIME) (see p. 9).

During the initial attempts to expand the New Haven ACG into a more useful geographic base file, the first difficulties were encountered with digitizing. (Digitizing is the mechanical process of assigning geographic coordinates to points on a map so that these coordinates may be incorporated into the geographic base file. After nodes and node numbers have been marked at street intersections and other geographic points on a map, the map is aligned on the bed of an electronic coordinate locator, or digitizer. The cursor of the machine is first positioned over a reference point on the map from which all X and Y coordinates are derived. The cursor is then moved by hand from node to node. For each node, the node number and map reference number is recorded onto a machine-readable record along with the X and Y coordinates (in hundredths of inches) calculated automatically by the digitizer. The resultant file of node numbers and associated coordinates then can be merged

with the ACG street data to provide a DIME file.) For the typical intersection (e.g., two streets crossing), eight readings had to be made on the digitizer, one for each of the block side ends forming the intersection. (Eight readings had to be made because the file, as originally created, provided no way of identifying each block side with relation to an intersection.) When such an intersection was later plotted mechanically, plotted points at the intersection were usually shown in slightly differing locations. This discrepancy was due to variability introduced by positioning the cursor over the same point eight different times, and also to error introduced by mechanical slippage in the digitizer itself. In many cases the block sides could not be located with certainty on the maps.

Since the ACG was a file of streets only, nonstreet features existing in the coded area such as railroad tracks, rivers, lake shores, and municipal boundaries were not included in the file and therefore could not be digitized. The absence of nonstreet features which formed block boundaries made it impossible to plot block maps or to calculate the area of blocks. This absence also limited the usefulness of a plotted map for display or analysis.

A related problem was the inability to digitize curved streets accurately. In the ACG, each block side was coded from intersection to intersection without regard to curves that might occur along the block side. As a result, when a curved block side was digitized and plotted, it was shown as a straight line. This often led to plotted curved streets incorrectly crossing other plotted streets and depicting nonexistent intersections.

A substantial number of coding errors and omissions were detected in the New Haven ACG as it was used in the use study. The discovery of these errors not detected by the then-existing ACG clerical and computer edits, and the coordinate assignment problems discussed above, led to the conclusion that alternative approaches should be explored. Although the ACG was sufficiently accurate for the geographic coding of census questionnaires, the Bureau staff concluded that improving the ACG itself for other purposes such as computer mapping would not be practical, and that a new system was needed.

The task of developing a new system was referred to the Technical Steering Group² of the Census Use Study, which established a subcommittee in May 1967 to explore other methods of automated geographic coding, to conduct original research if necessary, and to develop a geographic base file which could be used efficiently for computer mapping and geographic analysis.

Theoretical base.--The subcommittee first considered existing method of creating geographic base files. It then proposed construction of a geographic base file using graph theory as the conceptual framework.

Each street, river, railroad track, municipal boundary, or other map feature could be considered as one or more straight-line segments. Curved lines could be divided into a series of straight-line segments. When features intersect or when straight-line segments change direction, vertices are formed. Using a concept derived from graph

²The Technical Steering Group had oversight of the Census Use Study, and was composed of members of the Census Advisory Committee on Small Areas (see p. 1).

theory, a method was devised for creating a geographic base file where each vertex (intersection or node), as well as each line segment (the line between two nodes), and the intervening enclosed area, could be uniquely identified. Also, each could be identified in terms of its place within the entire region encompassed by the file and in relation to adjacent nodes, line segments, and enclosed areas. The entire region encompassed by the geographic base file could be viewed as a series of interrelated nodes, lines, and enclosed areas. Each line segment could be identified by its name as taken from the source map, e.g., Main Street or Green River, and by the node numbers at each end of the line segment. Nodes were numbered sequentially and uniquely. Enclosed areas were numbered using a systematic series of block numbers or other areal identifiers; they could also be described by the nodes or line segments which bound them.

The creation of such a file of line segments, nodes, and enclosed areas allowed for the editing of the file by computer, using an algorithm derived from graph theory. When coded, the three elements--segments, nodes, and enclosed areas--could be formed by the computer into separate incidence matrices, e.g., line segment-node, line segment-enclosed area, and enclosed area-node. The relationship established within these incidence matrices provided for the development of the computer editing system.

Such a file, as a description of a map, then could be digitized and plotted by computer to produce a complete replica of the source map. This file was given the name DIME, for the Dual Independent Map Encoding process involved. (The term refers to the fact that each boundary segment in the network was described by specifying its two end nodes and its right and left blocks. With each node and block uniquely numbered, the computer could then

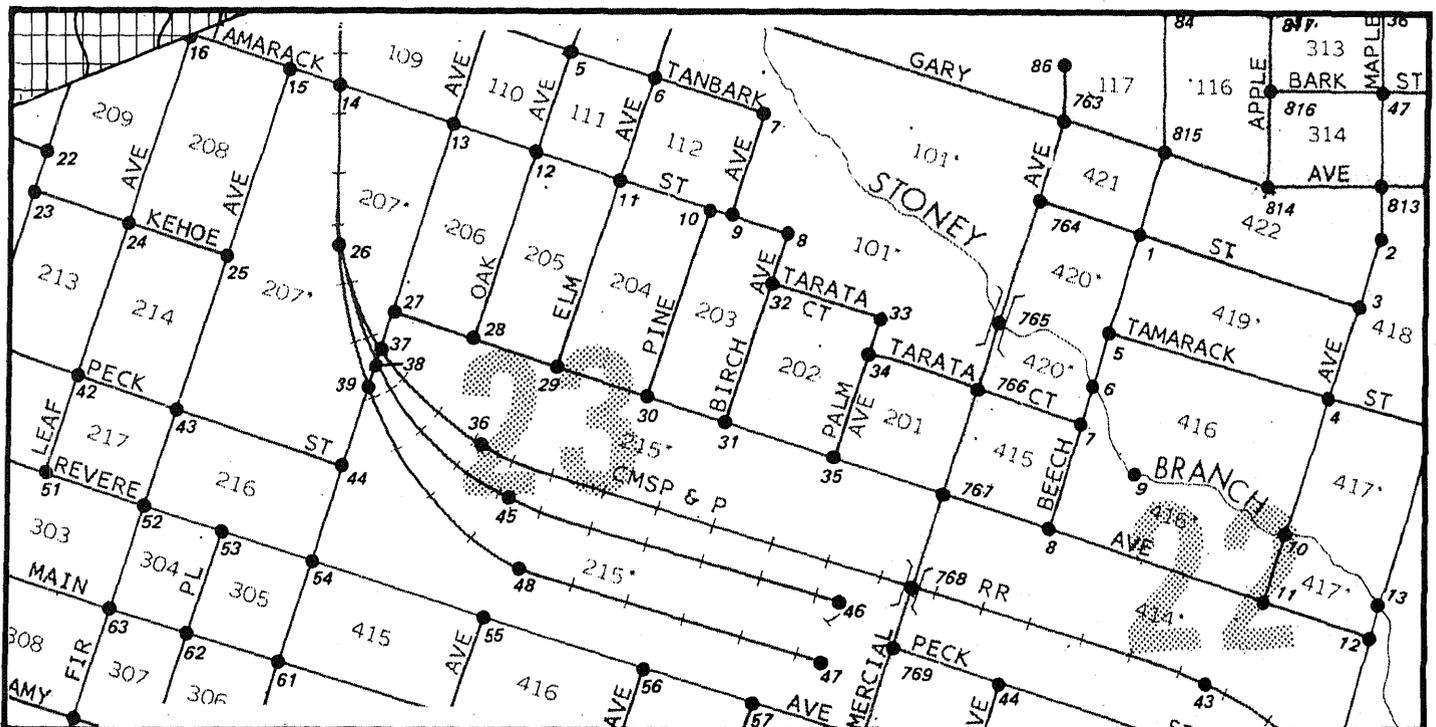
construct two independent networks and match them to make certain that the existing network was completely represented and that all land was accounted for.)

Development of the DIME File

From this theoretical base, the staff developed procedures and created a DIME file for New Haven. The same reference sources used to create the ACG--Census Bureau Metropolitan maps and local address maps or listings--were employed. Only one major change was made to the Metropolitan maps prior to DIME coding--node points were placed on the map and then were assigned unique identification numbers (see fig. C). The coding form used was a standard worksheet (similar in format to that used for the ACG) designed so that the data on the worksheet could be transcribed to punchcards.

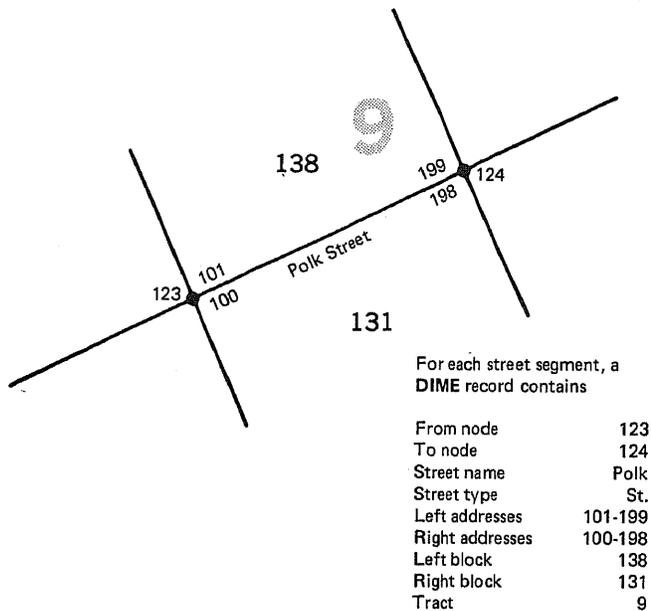
While the ACG was constructed on a block-side basis, the DIME file was constructed on a street-segment basis. Each ACG record contained the appropriate codes for one side of a street between two intersections on that side of the street. Each DIME segment record contained the appropriate codes for both sides of a street between two nodes (see fig. D). After the local coding operation was completed, the file was edited using DIME edit techniques; rejected records were corrected and reinserted in the file. The coding maps were digitized and the nodes and coordinates from this process were inserted into the DIME file. Test maps were prepared on a local map-plotting machine; coordinate errors detected on the maps were corrected and the file again was updated. If coordinates were improperly assigned to nodes or vice versa, a correction was made; if block faces were changed prior to the compilation of the data the file was updated and the maps were changed. The file was then considered complete.

Figure C.—Numbering the nodes.



A node is a point on the map where streets (or other features which are to be represented) intersect, end, or sharply curve.

Figure D.—Contents of a DIME Record.



Testing.--During the winter of 1967-1968 a DIME file was created for the city of Roanoke, Va., using the Bureau's draft instruction manuals and worksheets. In the spring of 1968 coding procedures were tested at Binghamton, N.Y., where a DIME file was being compiled by the Broome County Planning Board. The manuals and worksheets used at Roanoke were revised, and a new coding manual was introduced for the Binghamton test. Areal coverage for this test was extended beyond the central city to the county boundary. After training, local personnel completed the coding operations with only nominal monitoring. The Bureau's Metropolitan maps were used for the urbanized portion of the county, and county highway maps were used for the remainder.

Another test leading to the development of the DIME computer program package, similar to the test in Binghamton, was conducted in the Greenville, S.C., SMSA by the Bureau and a local planning commission later in 1968. Geographic coverage was limited to the city postal delivery portion of the urbanized area of the SMSA, as in the ACG program. Prior to the test, the coders' and supervisors' manuals and the coding and administrative forms were refined in accordance with the findings of the previous tests. However, this test was designed not only to evaluate the DIME computer program package but also to develop materials which could be used to create geographic base files for SMSA's to be enumerated by conventional (or nonmail) means in 1970, rather than by the mail-out/mail-back technique.

Concurrent with the Greenville test, an experiment was conducted to examine the feasibility of merging an independently created DIME file without addresses for Madison, Wis., with the ACG for Madison. The DIME file was created and edited by Bureau staff; the ACG file was coded by a local agency and processed as part of the standard ACG program. The Metropolitan maps were used as the base for creating both files. Theoretically, it should have been possible to merge these two files into an ACG/DIME file. However, a substantial number of records could not be merged. For example, of the 2,345 block

faces which could be compared, there was at least one difference between the two files in 1,104 cases. An analysis of the differences revealed that 19 percent were due to street name spelling, 3 percent to errors in the DIME file, 47 percent to errors in the ACG, 9 percent to errors in both files, and 22 percent to ambiguities in the Metropolitan maps. Reconciling the various differences took more time than was feasible, and the experimental merging of separate ACG and DIME files by computer was discontinued.

However, recognition of the desirability of inserting DIME features into the ACG's increased as it became apparent that the ACG's could contain a considerable number of errors that were not detectable by existing ACG edits, and further research was undertaken. Inserting unique, computer-generated node numbers into a specially prepared ACG was attempted, but this experiment was only partially successful and was abandoned.

In early 1969, the Bureau staff investigated the possibility of adding DIME features, i.e., node number and left/right parity status (specifying that even numbers are on the right and odd numbers on the left, or vice versa), to a specially prepared ACG block side file by clerical procedures.

The following procedures were developed for adding these features:

1. Inserting nodes and node numbers on Metropolitan maps as in DIME, i.e., with all intersections and points along curved features numbered uniquely.
2. Preparing a specially formatted worksheet printout version of the ACG, allowing for the clerical insertion of node numbers and left/right parity status for each block side in the file.
3. Coding additional records for nonstreet features (which were not coded in the ACG) and for block sides which had to be split into two or more pieces because of curves or other reasons.
4. After coding was completed, keypunching the added codes (and unique identification numbers for each record) and merging them into the original ACG file. The resulting "segment side" records were then merged to form segment records.
5. Editing the resulting file and correcting the edit rejects, using the DIME edits.
6. Digitizing the Metropolitan maps used in coding to create a file of node records--each node record containing the node number and its coordinates.
7. Merging the node record file with the DIME file.

This technique was tested under field conditions, using the ACG for a small number of census tracts in Madison, Wis. Testing showed that the system was workable and sufficiently economical to warrant its use in a nationwide program. The Bureau accepted it as its program for inserting DIME features into existing ACG's after the 1970 Census.

The DIME method was used in the 1970 census, however, to create geographic base files for coding place-of-work data for 80 SMSA's in which the census was taken by "conventional" methods rather than by the mail-out/mail-back procedure. (For details, see chapter 3.)

Several foreign countries (France, Israel, and the Federal Republic of Germany, for example) have used the DIME clerical and computer program packages to create DIME files for their administrative records programs and national censuses. Arrangements have been made for the interchange of information through several international conferences during the 1970's.

SOUTHERN CALIFORNIA REGIONAL INFORMATION STUDY

The Southern California Regional Information Study (SCRIS) was an expanded version of the New Haven Census Use Study program, and covered the California counties of Los Angeles, Orange, Riverside, and San Diego, a geographic area of 16,274 square miles.

The SCRIS activities were sponsored by the Southern California Association of Governments (SCAG) and the Bureau of the Census. The project also was supported by municipal, regional, State, and Federal government organizations, and by private and academic institutions. The Census Bureau served as principal liaison between SCRIS and interested Federal departments and agencies, while SCAG was responsible for local liaison and for assisting in evaluating study projects proposed by local agencies.

The SCRIS office was opened with three professional staff members in July 1969, in temporary quarters provided by and shared with the Los Angeles City Planning Department. In August, a permanent office was established in Los Angeles. Management of SCRIS was assumed by the County of Los Angeles in July 1972.

The counties selected for SCRIS presented a broad spectrum of testing opportunities, as their areas included industrial, agricultural, and mineral uses of the land, as well as average population densities ranging from 1,816 persons per square mile in predominantly urban Orange County to 64 per square mile in Riverside County, which contains large areas of desert. Los Angeles County, which is 98.7 percent urban, has a population of approximately 7.0 million people, of whom about 2.8 million live in the city of Los Angeles. The other 76 cities in the county range in size from Long Beach, with nearly 360,000 people to Vernon, with approximately 260 people. These cities are autonomous, and their boundaries in many cases are interlaced and complicated. For example, the city of Industry (with a population of less than 750) covers a total area of 10.8 square miles, but extends in one direction for almost 14 miles and adjoins seven different cities at various points.

Mandatory Work Projects

The first phase of the SCRIS work activities consisted of four major "mandatory" work projects, i.e., projects

essential to the study, which were conducted between 1969 and the end of 1970. In order of priority, these projects were the following:

1. To complete ACG/DIME files for the study area.
2. To create geographic base files from these by inserting coordinates into them, and to experiment with the use of these coordinates for data retrieval and graphic display.
3. To develop updating and maintenance procedures for the geographic base files.
4. To develop a general-purpose system for linking local geographic location identifiers (such as school districts) associated with various data to the geographic base file and to each other (in order to link economic and demographic data on a common-area basis), and to provide the capability for displaying the data by computer graphic techniques.

Optional Work Projects

The second phase of the SCRIS work activities was carried on concurrently with the mandatory projects, and consisted of "optional" work projects indicated by local interests. These optional projects were evaluated by SCRIS and SCAG, and were conducted with technical assistance from the Census Bureau where necessary. The optional projects were concerned mainly with the following:

1. Application of the geographic base file techniques to local situations.
2. Investigation of ways to increase the detail in the base files by adding such information as land contours, utility lines, etc.
3. Conduct of local surveys using the geographic base file as a sampling frame, and evaluation of these surveys for possible national application.

In one example of an "optional" work project, SCRIS assisted the Orange County Planning Department (later the Orange County General Planning Program) in its appraisal and analysis of housing conditions in the county by developing a geocoding system for the entire county. The county planners also requested that the system be able to provide detailed information at the block level for the county's nonurbanized area, which has a fast-growing population.

Experience was gained in handling typical coding problems such as using maps with differing scales and overlapping coverage and determining boundaries of the ACG's (which were based in part on the extent of post office city delivery service).

In August 1971 SCRIS began work on two three-phase projects sponsored jointly with the Los Angeles County Departments of Health and Mental Health. These were as follows:

1. Preparation of population profiles of the respective departments' service areas, using special census tabulations.

2. Establishment of service rates by combining census and local data records.
3. Development of projections from these rates for future needs by service areas.

Several SCRIS staff members were appointed to a county task force on the replacement of a major hospital destroyed in the 1971 Southern California earthquake. The SCRIS representatives provided data and assisted in formulating recommendations.

Record Matching

In addition to producing a translation of the ADMATCH/DOS computer program package (see p. 4) suitable for use with equipment requiring a language different from the original, SCRIS developed a system called ADMATCH/OS for larger computers where greater flexibility and faster operating speeds were possible, and a program called DIME-ADMATCH Record Conversion (DARC) for modifying ACG-DIME files so they could be used as input for an ADMATCH program.

During the course of experiments using ADMATCH in transportation studies, the need to match street intersections, major traffic generators, and individual names, etc., led to the development of computer program packages designed to provide generalized record linkage. The first of these packages was a Universal Matcher (UNIMATCH), compatible with the ADMATCH system but also designed to supersede the ADMATCH Matcher program. In UNIMATCH the user selected the fields to be compared, the comparisons to be made (e.g., character, numeric, or parity comparisons in fixed fields or intervals), and the subsequent actions to be taken. UNIMATCH included record editing capabilities that allowed matching of fixed format records without separate preprocessing. Development of a program called UNISTAN was begun in 1972 to replace the ADMATCH Preprocessor with a more flexible and generalized method of standardizing records. Its main purpose was to provide the user with the capability of standardizing fields subject to uncertainty and variability, such as house addresses, names, etc. After record standardization, the standardized files could be processed through the UNIMATCH programs.

Resource Application

In 1971 the Census Bureau and SCRIS began developing a generalized system for determining the service areas for a set of facility locations (for example, fire stations, repair shops, etc.). This system, called Computerized Resource Allocation Model (CRAM), was a refinement and extension of the Network Allocation of Population to Shelter (NAPS) system developed in connection with a Census Bureau contract with the Office of Civil Defense. The CRAM system used a DIME file as its network base, and demand and facility capacity as input. It applied a shortest-path algorithm to the DIME file; the DIME file represented a network of interconnected links, with the nodes in the DIME file as the connections. The resulting analysis could be printed out graphically in the form of service area boundaries, facility locations, transportation routes, or similar recommendations.

Computer Mapping

Following earlier studies of computer mapping systems begun in 1967 by the New Haven Census Use Study (see p. 4), SCRIS investigated and tested several of the systems available in 1970. Most of those tested required extensive programming and data preparation, as well as sophisticated computer facilities. In order to produce inexpensive printed maps quickly and easily, the Census Bureau and SCRIS began developing a generalized computer graphics system called GRIDS (Grid Related Information Display System), written in a standard computer language but including a computer mapping language as well. GRIDS was designed to operate on computers of various sizes and complexity, the smallest being one with 32,000 bytes (units of computer storage). GRIDS' finest unit of resolution was one grid cell; the system produced maps by dividing the area to be mapped into a network of rectangular grid cells similar to those shown in fig. A (p. 6). A grid cell could be as small as one printed character or as large as 55 x 55 characters.

GRIDS accepted as many as eight input variables and two coordinates to be mapped, provided for manipulation of the data in many ways, and produced up to five different maps for each computer run. It produced three types of maps: (1) shaded maps, in which data values were represented by overprinted characters of varying darkness; (2) density maps, in which a character was printed randomly throughout each grid cell, with more characters representing higher values; and (3) value maps, in which the actual data value was printed in each cell.

SCRIS developed a program called DACS (DIME Area-Centroid System), designed to calculate areas and locate centroids of blocks, census tracts, and other areas defined in a DIME or geographic base file. The centroid locations were used primarily as input to the GRIDS and other map-generating computer packages; the area calculations were useful in spatial analysis or in resource allocation programs.

SCRIS also experimented with the use of a cathode ray tube photographic reproduction device operated by computer. In 1971-1972 SCRIS modified the software system associated with this device to improve its mapping capabilities and began developing a mapping system tailored for use with DIME files.

Census Data Processing Programs

In 1970-1972 SCRIS developed a series of programs called STRIP (Summary Tape Retrieval and Information Processors). These programs allowed users to select census summary tape records by geographic code and to manipulate the data in various ways. The STRIP programs included CATS (Columnar Aggregation and Tabulation System); a First Count Reporter to produce summaries of selected data from the first count census summary tape; and FOURUSE, a package of programs for use with the fourth count census summary tapes.

DATA USER SERVICES

Summary Tapes

A "summary tape" contains tallies, in machine-readable form, of the responses to inquiries on the census questionnaires, summarized by census block, tract, enumeration district (ED), or other geographic area, in far more detail than appears in the printed reports.

Following the 1960 census enumeration, summary tapes were created from the census basic record tapes (those containing the data by individual household) as one of the processing steps in producing the printed reports. These summary tapes were made available to the public after review to make certain that confidentiality rules were observed. Any data which might reveal the identity of a specific individual, household, or housing unit was "suppressed," usually by combining the relevant data with those for other areas and reporting them together. Approximately 70 census users purchased these tapes, and some provided copies to other users. However, the 1960 summary tapes were recorded in a combination of binary, binary-coded decimal, and excess three codes suitable to the Bureau's needs, so that both the tapes and their accompanying documentation had to be converted to other formats compatible with users' computer equipment. This conversion frequently was complex and error-prone.

The demand for summary tapes increased during the 1960's, however, because the tapes contained detail not available in the printed reports and because more users had computer equipment available. The Bureau therefore developed plans and procedures for issuing 1970 census summary tapes that were oriented insofar as feasible toward users' needs. These needs had been determined to be (1) tapes compatible with users' computer equipment; (2) adequate documentation, technical assistance, and education for potential users; and (3) the maximum possible amount of data and small-area geographic detail. As more of the country would be enumerated by block and tract in 1970 than in 1960, the amount of detailed data available for publication would be correspondingly greater; further, more data would be cross-tabulated, especially by race. As the summary tapes made available for purchase would contain the same tabulations as the Bureau's work files (the statistics being processed to produce the printed reports), it would only be necessary to run these tabulations through computer programs which would suppress any data that might violate confidentiality rules, and to perform the necessary conversions to make the tapes usable on computers outside the Bureau. Thus users usually could have summary data in their possession in advance of the printed reports, and in far greater detail.

One of the dress rehearsals for the 1970 census, conducted in Dane County, Wis. (Madison SMSA), in 1968 provided the census basic record tapes for both complete-count and sample population and housing data in the format decided upon for 1970. The Bureau used these tapes to test its data-delivery systems for 1970 and also to produce six different summary tapes for users. These summary tapes, known as the "test tapes," were released for sale between January 1969 and September 1970 so that users could begin planning

and testing their own systems and procedures for analyses incorporating the 1970 census data along with their own.

Summary tapes from the 1970 census were produced in several series, or "counts."³ The first three counts contained only 100-percent (sometimes called "complete-count") data derived from the responses to the census questions asked for all households and housing units. The fourth, fifth, and sixth count summary tapes present tabulations of the full range of census data, including 20-, 15-, and 5-percent sample population and housing characteristics. Except for the fifth count, for ZIP code area, the tapes were prepared on a State-by-State basis by count and, if applicable, by file⁴ within count. The first summary tapes, the first counts for the States of Montana and Wyoming, were issued in September 1970, and the comparable printed reports were issued in November 1970. For the coverage and content of the 1970 census summary tapes compared with the printed reports derived from them, see fig. E.

The Bureau's 1970 census summary tapes (prepared from the basic record tapes and used, as in 1960, for processing the data selected for the printed reports) were converted in a special operation to master tapes compatible with most computer hardware used outside the Bureau. Copies then were prepared from the master tapes in 7-channel, 556- or 800-cpi, or in 9-channel, 800-cpi format, according to the requesting user's needs. The user was charged for the physical tape reel, the processing necessary for copying the master tapes, documentation, and postage and handling. Routine orders were accepted on this basis for the tapes either for complete or partial files, as they had been prepared (by count for each State and, if applicable, by file within count).

There were some users who wanted the same level of geographic and statistical detail as was available on the summary tape, but who did not want to work with the computer tapes. Accordingly, the Bureau prepared microfilm copies of the first count tapes. A document recorder was used to "read" the summary tape according to specific formatting instructions and to arrange the tape's contents for photographing. To minimize the amount of film and programming required, the microfilm frames were essentially arrays of numbers organized into lines and columns. Each frame represented data for one particular geographic area, such as an enumeration district or county. Coded geographic identification appeared at the top of each frame. There was no

³The word "count" refers to a set of specified data tabulations (tallies) for certain kinds of geographic units; for example, there are 55 separate tabulations and cross-tabulations of 100-percent population and housing data for each enumeration district or block group, and for counties and other selected areas in the "first count." Numbering of the counts--first through sixth--refers to the computer run(s) through which the census basic record tape is put to obtain the specified data and the order in which the tabulations were released.

⁴The word "file" refers to a collection of records at a particular level of tabulation. For instance, File A of the first count summary tape contains the data by enumeration district, whereas File B contains the same data summed by place, county, etc.

descriptive text on the microfilm itself; purchasers were provided with documentation similar to that used with the tapes.

To satisfy the needs of users who had access to computers but lacked adequate programming support to develop the software necessary to utilize the summary

tapes, the Census Bureau developed computer programs to select and display the desired tabulations for geographic areas of interest from the first five counts. These programs, called DAULIST (Data Access and Use List) 1, 2, 3, 4, and 5, were written in standard machine languages usable on most small-scale computers and were made available as standard Bureau products.

Figure E. Summary Tape Files and Printed Reports Derived From These Files

Summary tape file and printed report derived from the tape file	Geographic coverage	Subject content
TAPE: <u>First Count</u> , File A No reports derived from this file	ED's or blockgroups	400 complete-count population and housing data cells ¹
TAPE: <u>First Count</u> , File B REPORT: <u>Population</u> , Volume I, Chapter A	States, counties, congressional districts, MCD's (CCD's), places Same as above	400 complete-count cells (same as File A) Population totals only
TAPE: <u>Second Count</u> , File A REPORT: <u>Population and Housing</u> , Census Tract Reports	Census tracts Census tracts	3,500 complete-count population and housing data cells 190 complete-count population and housing cells
TAPE: <u>Second Count</u> , File B REPORT: <u>Population</u> , Volume I, Chapter B REPORT: <u>Housing</u> , Volume I, States and Small Areas, Chapter A	States, SMSA's and their components, counties, urbanized areas, MCD's (CCD's), places Same as above Same as above	3,500 complete-count population and housing data cells Extracts of the above Extracts of the above
TAPE: <u>Third Count</u> file REPORT: <u>Housing</u> , Volume III, Block Statistics	City blocks City blocks	250 complete-count population and housing data cells 25 complete-count population and housing data cells
TAPE: <u>Fourth Count</u> , File A REPORT: <u>Population and Housing</u> , Census Tract Reports	Census tracts Census tracts	13,000 sample population and housing data cells Limited extracts of the above
TAPE: <u>Fourth Count</u> , File B (MCD/CCD) and File C (other areas) REPORT: <u>Population</u> , Volume I, Chapter C REPORT: <u>Housing</u> , Volume I, States and Smaller Areas, Chapter B	States, SMSA's and component, urbanized areas, counties, MCD's (CCD's), places States, SMSA's, counties, urbanized areas, larger places States, SMSA's counties, urbanized areas, larger places	13,000 (File B) or 27,000 (File C) sample population and housing data cells Extracts of the above Extracts of the above
TAPE: <u>Fifth Count</u> No reports derived from this file	5-digit ZIP code areas in SMSA's; 3-digit ZIP code areas elsewhere	900 sample population and housing data cells
TAPE: <u>Sixth Count</u> REPORT: <u>Population</u> , Volume I, Chapter D REPORT: <u>Housing</u> , Volume II, Metropolitan Housing characteristics	States, SMSA's, metropolitan counties, cities with 50,000+ population, central cities, and, for housing only, nonmetropolitan counties with 50,000+ population Same as above, excluding nonmetropolitan counties with 50,000+ population Same as above, including nonmetropolitan counties with 50,000+ population	150,000 sample population cells 110,000 sample housing cells Extracts of the above Extracts of the above

NOTE: Summary tapes associated with special subject reports also were released.

¹A "data cell" represents a single summary of a tabulation unit (e.g., persons or housing units) having a certain characteristic, such as an age group or type of tenure. For example, a tabulation of total population by single years of age from "under 1" to "100 or more" would constitute 101 data cells.

Public-Use Samples

"Public-use samples" are samples of the Bureau's basic census records for individual persons, households, and housing units, with identifying information (such as names, addresses, and detailed geography) removed to preserve the confidentiality of the records. These data are made available to the general public on computer tapes on which responses to the items on the census sample questionnaires are given as numeric codes.

1960--1-in-1,000 and 1-in-10,000 samples.--The creation of a publicly available sample of basic data from the census was initially proposed by the Bureau's Population Technical Advisory Committee. The resultant project was cosponsored by the Bureau and the Population Council.

The first public-use samples were released on tape and punchcards in 1963; they consisted of a 1-in-1,000 sample of the 1960 population census, all taken from the 25-percent sample (in which the most detailed characteristics were collected) and a 1-in-10,000 sample which was a random subsample of the 1-in-1,000 sample. The 1-in-1,000 public-use sample included the records of the population characteristics of approximately 180,000 individuals who had been enumerated in the 25-percent sample of the 1960 census, as well as certain characteristics of the housing units in which they lived, and selected characteristics of their households, families, and sub-families (if any). Thus the 180,000 individuals comprised a 0.1-percent sample of the 1960 population of the United States. These data were identified only by broad geographic regions of the country.

These public-use samples were furnished to about 85 purchasers (some of whom distributed copies), enabling users in universities, government agencies, and private research organizations to prepare analytical tabulations of the characteristics of the U.S. population and occupied housing units. Researchers used the samples in their studies of family relationships in minority populations, characteristics of working mothers, migration to urban centers, and many other subjects. The samples also were used as training material in the social science departments of many colleges and universities.

In 1968 the Census Bureau solicited the reactions and suggestions of users of the samples. The response indicated the desirability of greater geographic detail than the regional designations on the 1-in-1,000 sample. Similarly, users interested in studying subgroups of the population, such as minority races, found these insufficiently represented in the 1-in-1,000 sample and requested a larger sample size.

1960 and 1970--1-in-100 samples.--As a result of the above study a new series of public-use samples was prepared from the basic records of the 1960 and 1970 censuses, representing an expansion in content and sample size as compared with the earlier 1960 samples. In order to facilitate time series comparisons, the new 1960 samples were designed to parallel the 1970 samples in form, content, and geographic area. The newer public-use samples identified smaller geographic areas, some as small as 250,000 population. In addition, the primary sample size was 1-in-100, or one sample unit (household, vacant unit, or individual person in group quarters)

for every 100 such units in the country. For each housing unit, information was provided about the housing unit itself as well as the characteristics of each person in it. The samples were self-weighting, i.e., each person or household was to be assigned a weight of 100 so that a user needed only to multiply his computations of public-use sample records by 100 in order to estimate the frequency of a particular characteristic for the entire population.

Public-use samples of the 15-percent and 5-percent data for 1970 were prepared separately, since the two 1970 sample questionnaires contained some different questions. Each 1-in-100 sample of the total population actually contained either one-fifteenth of all 15-percent basic records or one-fifth of all 5-percent basic records, including slightly over 2 million persons in each case.

Six different public-use samples were drawn from the basic records of the 1970 census, three from the 15-percent sample and three from the 5-percent sample, one each for three types of geographic areas, as follows:

1. County group samples.--These files were organized and identified by particular groups of counties. Every SMSA having a population over 250,000 was identified and, whenever possible, subdivisions comprising one or more counties and having a population over 250,000 also were identified. Outside these large SMSA's, groups of economically interrelated counties were identified. State boundaries were not necessarily observed in determining county groups; county groups were defined on a map which was part of the technical documentation.

2. State samples.--These tape files were organized and identified by State. In addition to the State code, three geographic codes were used where appropriate: urban, rural, metropolitan/nonmetropolitan, and central city/noncentral city. One or more of these codes were omitted when inclusion would result in the identification of an area with 250,000 or less population. For instance, no urban/rural designation was given for sample units from Alaska, Delaware, Hawaii, Nevada, Rhode Island, Utah, Vermont, or Wyoming.

3. Neighborhood characteristics samples.--These files were organized by urbanized and rural areas identified only by size and geographic region of the country. In addition to the types of information on housing units and persons that appear in the other public-use samples, these files contain census characteristics for the specific neighborhood in which the reported persons lived. The specific neighborhoods were not identified, but represented areas approximately the size of a census tract (about 4,000 people).

Both 1-in-1,000 and 1-in-10,000 subsamples for the Nation were prepared from each 1-in-100 sample.

Geographic Reference Materials

The only geographic tool available to users of the 1960 census summary tapes which allowed them to relate the tabulations to their geographic locations was the GICS (Geographic Identification Code Scheme). The

GICS was a printed list by State of the geographic codes used in the census, from the State level down to place, plus the range of enumeration district numbers for each place.

Between February and November 1970 the Bureau developed and issued two lists of the 1970 geographic codes. These were the following:

1. MEDList (Master Enumeration District List), a listing on magnetic tape (also available on microfilm and enlarged paper copies of the microfilm) by State of the geographic codes from the State level to the individual census tract and ED or block group level. The list also included 1970 population and housing counts for each ED or block group, based on first-count tabulations. (See fig. F for an example of the MEDList as displayed on microfilm.)

2. GACI (Geographic Area Code Index), a truncated version of the MEDList which provided the geographic codes down to the level of place for those users not working with a finer level of detail. The entire United States GACI file was contained on a single reel of tape or one roll of microfilm; enlarged paper copies of the microfilm also were made available.

In the spring of 1972 the Bureau also issued the GACI as a paperback reference book bound in four separate parts, one for each region (Northeast, North Central, South, and West), entitled 1970 Geographic Identification Code Scheme.

While maps have been an integral part of the Bureau's publication program for many years, the Bureau began describing in detail the contents and coverage of its available map series only within the last decade. Information about the principal maps needed by census users to identify geographic areas for which data were compiled on the summary tapes had been contained in scattered publications and order forms. In the spring of 1972 the Bureau for the first time issued a consolidated inventory of its metropolitan, county, place, and vicinity maps. (Vicinity maps cover a total of about 125,000 blocks in over 900 areas outside the urbanized areas of cities with 50,000 or more inhabitants, for which the Bureau undertook to compile block statistics.) This two-volume inventory, which indicates the maps available for approximately 20,000 places listed alphabetically by county within State, was provided to Department of Commerce field offices, Census Bureau data collection centers, and recognized summary tape processing centers for users' reference.

For the census user who wished to associate local data records, such as those for school enrollment, hospital admissions, traffic generators, etc., with census geography by means of computer, the Bureau provided several basic tools. These were the following:

1. ACG's (address coding guides), magnetic tapes containing geographic codes for approximately 6.5 million block sides for the Post Office city delivery areas in 145 SMSA's. The ACG's were developed between 1965 and 1969.

2. DIME (Dual Independent Map Encoding) files (also known as GBF's--geographic base files), magnetic tapes containing geographic codes for approximately

9 million street segments in 196 SMSA's. These files were developed between 1969 and 1971 (see p. 8ff). Both the ACG's and the DIME files contained essentially the same information-- street names, address number ranges, census block numbers, census tract numbers, and other geographic codes. The DIME files contained several codes (for internal control and for mapping) not appearing in the ACG's as well as codes for nonstreet features such as rivers, railroad tracks, etc., appearing on the Bureau's Metropolitan maps. The ACG's and the DIME files could be used for geographic identification of data; the DIME files could be used for computer mapping and area calculation as well. Two computer programs, called TGUIDE (tract guide) and DARTS (DIME Address Range Telescoping System), were developed in 1972 to create smaller tract-level coding guides for users not interested in block-side detail.

3. ADMATCH, a computer package allowing the user to relate specific addresses and local data for those addresses to the census data on the summary tapes by matching the addresses either to a DIME file (after the DIME records had been "split" into block side records through the use of NICKLE (a computer program developed by the Bureau in 1971-72) or the SCRIS' DARC program (see p. 12), or to an ACG. The ADMATCH/DOS package for disk operating computer systems was developed in 1967 (see p. 4) and the ADMATCH/OS package for larger computers was developed in 1970-1972 (see p. 12).

For computer programs developed for record matching, resource application, and for local manipulation of summary tapes, see pp. 4 and 12.

Computer-Generated Reports

In conjunction with most of its major publishing activities during and after the 1960 census, the Bureau prepared brief, textual reports which summarized current and historical statistics for large areas. Press releases and analytic reports were issued for each State, for groups of States, and for the Nation as a whole. Few of these 1960 analytic reports were prepared for areas smaller than a whole State, even though population data were tabulated for areas as small as a single city block in a metropolitan area. It was impractical to prepare by hand textual reports at the county level, or even to prepare a one-page press release for each of the 3,135 counties in the United States.

During the 1960's, techniques were developed for using a computer to generate small-area analytical reports quickly and inexpensively. In 1965 the Bureau used the computer for the first time to generate press releases which reported data from the 1963 Census of Business. The releases provided data about the counties and larger areas for local news outlets.

New Haven tract reports--In April 1967 an enumeration of the New Haven, Conn., SMSA was conducted as a "pretest" of data-collection techniques for the 1970 census. This pretest also provided a data base for a cooperative effort called the New Haven Use Study (for details, see p. 2) in which research into methods to combine census data and local information for a variety of local applications was coordinated. The data needs of

Figure F. Example of MEDList, Displayed on Microfilm

1970 CENSUS OF POPULATION AND HOUSING
MASTER ENUMERATION DISTRICT LIST (MEDList)

Michigan		1970 counts																											
State	1970	1960	County	County of tab.	CC	MCD/CCD	Place			Urbanized area	Tracted area	Univ. area		SEA	ESR	CBD	Area name	Tract		Blk. grp.	ED		Urban/rural	Ward	Cong. dist.	1970 counts			
							Code	Desc.	Size			SCA	SMSA					Prefix	Code		Basic	Suffix				Code	Suffix	Housing	Pop.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
	26	34	121		1												MUSKEGON									49831	149493		
	26	34	121		1	090					5320						Ravenna TWP	0029									650	2102	
	26	34	121		1	090	2275	4	02		5320			1	34062	06	050	1	09	0442				1	09	09	250	801	
	26	34	121		1	090		7			5320			1	34062	06	050	1	09	0443				1	09	09	400	1301	
	26	34	121		1	095		4	06		5320	5320					Roosevelt Park City									633	2578		
	26	34	121		1	095	2345				5320						Roosevelt Park City	0022		1	2			0	09	09	633	2578	
	26	34	121		1	110		7			5320						White Hall TWP			3						200	885		
	26	34	121		1																					312	1175		
	26	34	121		1																					121	518		
	26	34	121		1																					581	1930		
	26	34	121		1																					30	90		
	26	34	121		1																					40	125		
	26	34	121		1																					72	150		
	26	34	121		1																					60	130		
	26	34	121		1																					165	625		
	26	34	121		1																					214	810		

Note: Explanation of column heading abbreviations: Federal Standard County; County of Tabulation; Central County Code (CC); Minor Civil Division (MCD)/Census County Division (CCD); Place Description; Standard Consolidated Area (SCA); Standard Metropolitan Statistical Area (SMSA); Universal Area Code; State Economic Area (SEA); Economic Sub-Region (ESR); Central Business District (CBD); Block Group; Enumeration District (ED); Urban/Rural; Congressional District.

the local agencies varied widely; some users wanted a single analytic report for the characteristics of the SMSA as a whole, others wanted a report with data for the city of New Haven only, while still others wanted statistical tables covering selected topics in particular census tracts or in the suburban areas. It was finally decided to produce a 12-page narrative text on the computer for each of the 77 census tracts in the New Haven SMSA. By having the computer generate analytical texts, not only were the greatest number of users provided with most of the data they needed in a standardized form, but users also found the reports easier to understand and use than printouts of statistical tables. Some agencies wanted only a few of these New Haven tract reports, but many users wanted all 77 for area-wide analyses. Special tabulations were provided to fill any data needs which could not be met by using the tract reports.

Bureau staff members planned the basic content of the reports and wrote the master text. The actual computer programming was provided by a commercial firm with experience in creating computer-generated reports. The population information in the reports included data on median age, sex and marital distribution, racial composition, and the household and family composition of each tract. Also included were analytical measures, such as a school-load factor and a dependency ratio. Housing information included the total housing stock, type of structure, average number of rooms per housing unit, median rent, and median house value. Analytic measures for the housing data included an index of overcrowding and an estimate of housing quality.

1970 census press releases.--Based on the experience gained from the New Haven tract reports and from the Business Census, planning began for production of computer-generated press releases to announce the results of the 1970 Census of Population and Housing. The final format was developed by an intra-Bureau committee chaired by the Chief of the Public Information Office.

The press releases were designed to be sent to newspapers and radio and television stations in each county, providing summaries of the census information about that county. The releases were generated in the computer by use of an auxiliary program during the actual processing of the 1970 census data, and approximately 15,000 releases were sent out. In order to match the recipients' addresses with their appropriate counties, the Bureau's news media address file was coded to conform with the FIPS (Federal Information Processing Standards) codes for States, counties, and SMSA's. Along with the press release for the particular county, each mailing package contained two advance census reports for the State: one report (Series PC(V2)) provided selected population characteristics for the entire State and each constituent county, the other report (Series HC(V1)) provided comparable housing data. Many newspapers requested additional copies, or wanted press releases on adjacent counties.

The first press release packages were mailed in August 1970, 4 months after collection of data began. The flow continued into 1971, accelerating in February, and ended March 3 with the reports for California, Illinois, and Texas.

Each press release covered about 40 items of basic information about an entire county's population and its housing. Population data included such items as age distribution, racial composition, and the percentage of population living in rural areas and of that living in urban places. Growth or decline in population between 1960 and 1970 was assessed for both the county and for the State as a whole. The population also was classified by the number of persons living in households, the number of persons living alone or with roommates, and the number of persons living in dormitories, military barracks, or other institutional housing.

The housing information included the number of owner-occupied housing units and the average value of such units, the number of renter-occupied units and the average rent per unit, and a count of all vacant housing units. Because of the interest of planners and other researchers, the housing which was in use year round was categorized by how much of it was in single-family homes, mobile homes or trailers, or housing units in multiunit structures. Two analytic measures were included: One cited the percentage of the housing units in the county in which there were more occupants than rooms (often an indicator of "overcrowded" conditions). The other measure was the number of housing units lacking some or all standard plumbing facilities. This could be used as an indirect measure of housing quality. All of the above information was based upon the 100-percent questions asked of the entire population.

The computer was programmed to select the most important statistics about an area, and to incorporate them in an explanatory text; the meaning of each number was made clear from the context of the sentence in which it appeared. A separate computer-generated report was produced for each particular area.

The computer selected the appropriate phrases from the alternatives in the "master text," and then printed them out along with the relevant data for the county. For example, a sentence might start out, "Between 1960 and 1970, the population...". The next word could either be "rose" or "fell," depending upon the data for the particular county. Phrases or even sentences which were irrelevant were excluded. Thus, if a county was totally rural, a sentence discussing the urban population would not be included in the report for that county, although such a sentence normally would be used. There was also provision for including additional sentences which took note of special facts about the particular county, such as a very high average-income level or an unusual ratio among the racial groups in the county population. The text also included analytic measures, such as ranks, percentages, and statistical averages. Comparisons were made with comparable statistics for larger areas (county-State comparisons, for example), or with earlier statistics for the same areas (such as percent changes from 1960 to 1970).

The entire computer news release project required approximately 26 hours of computer time and about 80 hours of high-speed printer time, or about \$0.48 for running each release. (Other costs, such as those for staff services, were included in various divisions' accounts.)

Congressional district profiles.--While the Census Bureau supplied congressional district tabulations on

the first-count summary tapes, these statistics and the ones in the congressional district data reports (described in chapter 10) were essentially tables of 100-percent data, to be analyzed by the individual reader. An ad hoc Committee on Congressional District Data recommended in September 1970 that analytical reports should be issued to describe the new congressional districts, after a State had redistricted, to reflect the population shifts during the 1960's. As soon as possible after a State was redistricted the computer generated a narrative profile of each of the districts based on 1970 population and housing data retabulated to reflect the new district boundaries. These computer-generated analytical reports, called Congressional District Profiles, were prepared for the use of members of Congress, State officials, and other interested persons as soon as the data could be retabulated to reflect the revised district boundaries.

The reports averaged about 25 pages each, and covered a full range of the 100-percent and sample population and housing characteristics tabulated for 1970. In addition, each district report contained comparable data for the entire State, and the district's ranking among the other congressional districts within the State on such factors as ethnic origin, racial composition, education level, median income, labor force characteristics, and a number of housing characteristics. Profiles were issued in the spring and summer of 1972 for 429 of the 435 congressional districts in the 93d Congress; none were issued for the six single-district States (Alaska, Delaware, Nevada, North Dakota, Vermont, and Wyoming). The Bureau furnished copies for each congressman; the National Technical Information Service (NTIS) of the Department of Commerce handled public distribution of duplicated and microform copies.

Central City-SMSA socioeconomic profiles.--In the latter part of 1972 the Bureau produced, primarily for the use of city officials, approximately 200 computer-generated profiles for those SMSA's containing central cities with populations between 100,000 and 1 million inhabitants. The content and text of each profile was based on the tabulations contained in the fifth count summary tapes (900 sample population and housing data cells for 5-digit ZIP code areas). These profiles were similar in size, format, and method of distribution to the Congressional District Profiles described above, except that a two-page press release, containing excerpts from the text, was provided for each socioeconomic profile as well.

Publications

During the 1960's the Bureau began publishing on a regular or occasional basis a number of pamphlets, memorandums, and newsletters particularly for census users. No comparable publications had been available in connection with the 1960 or earlier censuses. The principal ones inaugurated are as follows:

Small-Area Data Notes.--SAD Notes are monthly bulletins which report current Census Bureau activities likely to be of interest to users of data for counties and smaller areas. Contributions from readers and items dealing with projects outside the Bureau are occasionally included. Publication was begun in January 1966, with the title Small-Area Data Activities, through the co-

operative efforts of the Census Advisory Committee on Small Areas and the Bureau. At the time stress was placed on the bulletin's potential role as an "information exchange," reporting on all types of programs likely to be of interest to small-area data users, and heavily dependent upon the users' own contributions. Early issues reflected this information-exchange function but, as decennial census activity increased, emphasis later was placed on reporting programs relating to the 1970 census. Three issues of Small-Area Data Activities were published in 1966, two in 1967, and four in 1968. Under a new title, Small-Area Data Notes, three issues were published in 1969, 12 in 1970, and on a monthly basis thereafter. Distribution (about 13,000 copies per issue at the height of the data-processing period) reached urban planners, local officials, and many other small-area data users throughout the country who requested copies.

Data Access Descriptions.--Data Access Descriptions (DAD's) are brief papers published on an occasional basis in the form of pamphlets. The series was started in late 1966. DAD's generally have been oriented in subject content toward programs of concern to data users whose requirements go beyond the Bureau's printed reports. They included descriptions of the contents of the census summary tape series, census geographic aids, place-of-work data to be derived from the 1970 census, and several other access-related activities. In addition, several DAD's were used to report in advance the planned content of selected printed reports, so that data users could determine whether or not these reports could adequately meet their data requirements. Subscribers to Small-Area Data Notes also received these pamphlets.

Summary Tape User Memoranda.--In spite of their relative frequency of issue and broad subject-area coverage, Small-Area Data Notes and Data Access Descriptions were found to be inadequate for frequent and quick communication with users who had a serious interest in census data on computer tape. To fill this communications requirement, a memoranda series was begun in November 1967; issues later were given the title, Summary Tape User Memoranda. Each memorandum dealt with a specific topic, and could be printed and distributed more quickly than the larger publications. The first few memoranda went to about 200 persons who had an interest in helping organize workshops on the census and in establishing user groups to cooperate in acquiring and working with summary tapes. However, within a year the mailing list grew by several hundred, and by the summer of 1971 it had reached approximately 9,000 as other persons requested copies on a regular basis. About 90 memoranda were issued between the fall of 1967 and the spring of 1972, and covered such subjects as the test reels for the summary tape series, census maps, computer programs available, and related topics.

1970 Census Users' Guide.--In order to provide a comprehensive instructional and reference guide to the 1970 census, preparation of the 1970 Census Users' Guide, Parts 1 and 2 (see below), began in 1968; the Guide was published in its final form in October 1970. Two draft versions, prepared during 1968-1969, received substantial public distributions. The first, printed in September 1968, was sent to the relatively few users who at that time had indicated that they anticipated using 1970 census summary tapes. Following comments and

suggestions from these users, a second draft of the Guide was published in April 1969, and a total of 3,500 copies were sold by the U.S. Superintendent of Documents. In addition, the Census Bureau distributed about 1,000 copies to persons who had requested that their names be included on the mailing list of potential summary tape users. The final 1970 Census Users' Guide was made available to the public in advance of the first-count summary tapes for most States, yet it was completed late enough (October 1970) so that most descriptions of Bureau 1970 programs and services were accurate. Only changes in product-availability dates and some minor adjustment in technical documentation descriptions required correction notices; these notices were printed in Small-Area Data Notes. By the spring of 1972 approximately 8,500 copies of Part 1 and 6,500 copies of Part 2 had been sold. Part 1 included information on the collection and processing of the 1970 census data and the various media in which the results would be available, a dictionary of census concepts, and a glossary of census terms. Part 2 contained the documentation specifically related to the use of the summary tapes and the address coding guide.

Other.--Other publications aimed at the education of census users were in the form of (1) a speaker's reference guide for use by persons outside the Bureau, such as Department of Commerce field staff members, who were called upon to make presentations about the 1970 census; and (2) papers and articles prepared for delivery before census users' groups or publication in appropriate journals (see bibliography, p. 25).

Census User Conferences

Efforts to furnish data users with information they needed to effectively use the results of the 1970 census included the Bureau's participation in numerous conferences around the country, where its primary role was to encourage widespread participation and to provide speakers and descriptive materials.

Initial planning for the conference program began in 1966 when the idea of "Tally Tape User Workshops" (referring to what are now called "summary tapes") was put forward. The planned objectives of the program, which remained largely unchanged in subsequent years, were as follows: (1) to furnish information on census tape files and related materials, and to permit advance planning for tape use, (2) to encourage cooperation among potential tape users, (3) to exchange ideas about using the data, and (4) to provide feedback to the Bureau concerning possible problems and improvements.

The only significant modification in the general plan suggested in 1966 was reliance upon a conference rather than a workshop approach. This was mainly the result of attendance by larger numbers of data users and a lower proportion of technically oriented users than had been anticipated.

Bureau staff members participated in 140 conferences for census data users during the 4 years, 1968-1971 (see fig. G); 73 conferences were held during the census year, 1970, alone. Over 14,000 people attended these conferences, which were held in all the States, the District of Columbia, and Puerto Rico. There were also three series of conferences for Federal officials which are described below.

Early conference experience.--The possibility of a widespread conference program to assist census data users was first publicly suggested in the September 1967 issue of Small-Area Data Activities. By January 1968, 27 potential conference sponsors had contacted the Bureau, and arrangements were well under way for a prototype summary tape user workshop. The prototype workshop, held in Madison, Wis., in May 1968, was sponsored jointly by a university, a computer-processing center, and the Census Bureau. Over 100 persons, about one-third of whom were from outside Wisconsin, attended the 2-day conference. The first day was spent discussing such topics as the uses of the 1960 summary tapes, content of the 1970 tapes, and the geographic codes for the data. The second day was devoted to a critique of the first day's program. Resulting decisions with regard to the conference program were the following: (1) to rely on local sponsors to schedule and arrange conferences, (2) to encourage attendees to read basic materials in advance, and (3) to reduce the amount of technical detail discussed. During the remainder of 1968 representatives of the Census Bureau participated in a few additional conferences wholly or partly concerned with planning for the use of 1970 census data. However, the users' conferences program was not fully under way until 1969, when 33 conferences were held.

Conference organization.--The Census Bureau relied on local organizations, such as planning commissions, universities, chambers of commerce, and Department of Commerce field offices to sponsor the census users' conferences. Of the 140 conferences held, 97 were sponsored by local organizations and 43 by Commerce field offices (sometimes with cosponsorship by local groups). Expenses, including Census Bureau speakers' travel costs, were paid locally, often by charging a registration fee. However, in the case of each conference sponsored by the Department of Commerce, the Census Bureau absorbed the travel expense for one of its representatives.

Conference programs generally included three principal elements: (1) discussion by a Census Bureau representative of census enumeration procedures and processing methods, data products, and geographic tools, (2) presentations by local speakers on data use, and (3) discussion about formation of a summary tape processing center or the services of existing centers. Following the 1968-1971 conferences in which the Bureau participated, it was assumed that data users would rely increasingly upon local expertise, particularly that of personnel of summary tape processing centers, for explanations of technical documentation and advice on data use. The Census Bureau continued its policy of cooperating with local conference sponsors when requested.

Special conference series.--The Bureau sponsored three special conference series designed to assist personnel in other Federal agencies. The first series was for selected employees of the 42 Department of Commerce field offices. Four conferences, each lasting 2 days, were held during December 1969 and January 1970 in Atlanta, Ga., Detroit, Mich., Los Angeles, Calif., and Washington, D.C. Those attending heard presentations from Census Bureau speakers and were given detailed materials. The meetings helped prepare field office employees to answer the increasing number of questions

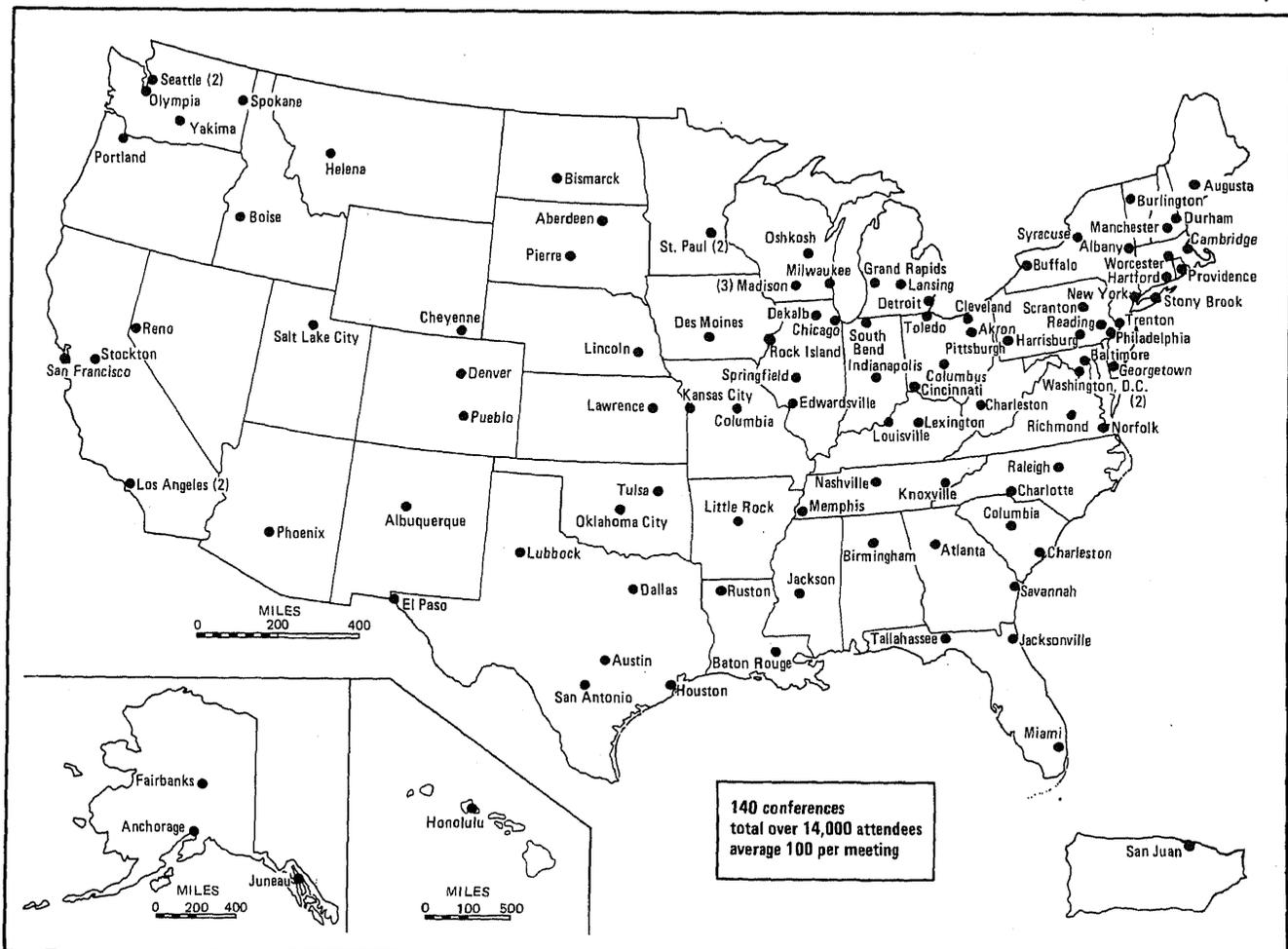
about the 1970 census directed to their offices and to conduct seminars on the census themselves. The second series was planned for interested Federal employees in Washington, D.C. Eleven conferences were held between November 1970 and April 1971--three for congressional staff members and eight for personnel of the Federal Home Loan Bank Board and the Departments of Labor, Housing and Urban Development, Commerce, and Health, Education, and Welfare. The programs, normally a half day in length, included both introductory information and specialized presentations relevant to the interests of the particular agencies (e.g., employment and occupation statistics for Department of Labor programs). A third series of special conferences was conducted in October and November 1971 in the 10 cities designated as Federal regional centers (Atlanta, Ga., Boston, Mass., Chicago, Ill., Dallas, Tex., Denver, Colo., Kansas City, Mo., New York, N.Y., Philadelphia, Pa., San Francisco, Calif., and Seattle, Wash.). These conferences were primarily for employees of Federal regional offices, and covered all Census Bureau programs rather than the decennial census alone.

Intergovernmental seminars.--In 1966 the Bureau established an intergovernmental services program for State and local government officials. This action was taken in response to the recommendations of the 1965

National Governors' Conference and the 1966 National Conference on Comparative Statistics. As a major part of the program the Census Bureau, in consultation with the Office of Management and Budget, developed a series of 5-day seminars which focused on the kinds of statistical information and services available from the Census Bureau and other Federal agencies, with emphasis on State and small-area data--including data from the 1970 census. Attendance at these seminars was by invitation of State, county, and city governmental units; regional, county, and city planning agencies; and university research centers. As of June 1972, 18 seminars had been held, with a total of 460 participants representing 301 organizations: 111 participants from 48 State governments, 159 from 97 cities, 52 from 42 counties, 116 from 95 planning agencies, 14 from 11 university research centers, and 8 from other types of organizations.

DIME workshops.--A series of regional workshops for planners, statisticians, systems analysts, and other professional users of DIME geographic base files was organized in 1972; the first two workshops were held that year in Indianapolis, Ind. (June 4-9) and Atlanta, Ga. (October 29-November 3). These workshops were sponsored jointly by the Bureau of the Census, the Department of Housing and Urban Development, the Defense Civil Preparedness Agency, and a national foundation.

Figure G. Cities Where 1970 Census User Conferences Were Held (1968-1971)



Summary Tape Processing Center Program

The potential role of organizations outside the Census Bureau in helping to meet the requirements of census data users was recognized by the Bureau several years in advance of the 1970 census. Purchasers of the limited tape series released following the 1960 census often made data and data services available to persons outside their own organizations. Their expertise and resources thereby were shared, and more effective and extensive use of census data resulted.

During the early 1960's the Bureau had occasional contact with several organizations, primarily universities and State government agencies, which used 1960 census data on tape. By mid-decade, following internal discussion and communication with a number of potential census summary tape users, it appeared that more active encouragement of user cooperation on the part of the Bureau was necessary. In 1967 serious data users were urged to volunteer as potential sponsors of summary tape user workshops. During 1968 such sponsors were given the title, "1970 Census Summary Tape Coordinators," and their roles were expanded to include conference planning and the possible establishment of processing centers.

A summary tape processing center program was launched formally in October 1968. The extent and nature of the resources and services of processing centers were expected to vary, as would their policies regarding clientele and charges. The Bureau had no requirements for the operation of any center, but it would recognize centers which generally provided any of the following kinds of services: (1) preparation of summary tape copies for other users; (2) provision of an information exchange for users on subjects such as display programs, problems in tape use, and new applications of census data; (3) maintenance of general computer programs prepared by the Bureau or by others, which would be available to local users; and (4) development or purchase of special programs, and advice to users of the availability of these programs. On its part, the Bureau would give priority to tape orders from recognized centers, furnish technical information to them, and advise census tape users of the location and services of processing centers. In addition, the Bureau furnished the centers with copies of the 1970 Census Users' Guide, the occasional informational memorandums, and consultation in meetings of processing center representatives. Technical information and consultation were made available to individual centers on the same basis as to any other data user.

The first summary tape processing center was recognized in January 1969; by mid-1972 there were 166 such organizations throughout the country. Of this number, 34 were government related, 53 were associated with universities, and 79 were private firms (some with several branch locations). These centers served an estimated 6,000 clients, with government centers accounting for about half of these and the remainder divided equally between universities and commercial organizations.

The most notable degree of interaction among processing centers occurred in connection with tape copying; there also was cooperation in developing computer pro-

grams for data display and manipulation and in exchanging tape data, such as that among the eight universities which make up the Association of Southeastern Summary Tape Processing Centers. Most centers limited their holdings to one or two States, representing the area from which most requests for their services were expected. In a few cases, centers had data for only part of a State; councils of governments which were recognized as processing centers, for instance, sometimes maintained data only for areas under the jurisdiction of their members. Services available ranged from tape copying and printed display of data to complex statistical analysis and consultation on data use. These differences are a reflection of such considerations as center objectives, personnel capability, computer facilities, and competing responsibilities.

Summary Tape Information Library Program

Since the earliest censuses, libraries have played an important role in making census data widely available, and most libraries of any size carry at least some of the Census Bureau's printed reports. A "Summary Tape Depository Library" program was initiated in March 1969 to assist interested libraries in making information about census summary tapes available to their patrons. In 1970 the name was changed to that of "Summary Tape Information Library" program to reflect this purpose. The Bureau issued occasional memorandums specifically oriented to library needs and furnished free copies of selected materials, such as the 1970 Census Users' Guide. Approximately 85 libraries were involved in this program.

Users' Service Staff

In the early 1960's the Bureau began to receive a number of inquiries about statistics that might be available from the 1960 Censuses of Population and Housing. Many of these requests concerned data that were not to be tabulated or published in the printed reports from the census. Others dealt with the possibility of obtaining census data in machine-readable form or obtaining special purpose statistics that could be provided only by retabulating the responses on the basic census questionnaires. Initially, one staff member was charged with examining each request and determining if the expressed data need could be met or even approximated with materials from the census results, and advising the inquirer accordingly. In some cases tabulated, but unpublished, census statistics could be made available at the nominal cost of clerically posting some statistical tables or photocopying some existing worksheet information. In other cases it was necessary to advise the requester that to furnish the data he needed would require the Bureau to write a special computer program and retabulate through the computer the tape reels containing the basic census records for that area of the United States for which the special statistics were needed.

As the number and complexity of such requests increased toward the mid-1960's, the staff assigned to handle these inquiries was enlarged. Then, in the latter half of the decade, as the Bureau entered the preparatory stages of the 1970 decennial census, outside inquiries about 1960 data waned and the special requests staff was reduced accordingly.

In the fall of 1969, in line with efforts to increase the accessibility and use of data products that would be developed from the results of the 1970 census, a new office, the Central Users' Service, was organized. Its first activities included the preparation and delivery of copies of test versions of the 1970 summary tapes based on the 1968 Special Census of Dane County, Wis. Gradually the Central Users' Service's responsibilities were expanded to include dissemination of copies of maps, ACG's (address coding guides), GBF's (geographic base files), DAUList summary tape display programs, the ADMATCH computer program, etc., and the coordination of requests for special tabulations of the basic 1970 census records.

This unit, later called the Users' Service Staff, also coordinated the actions of the Bureau's operating divisions which participated in providing users with census materials or special cost and time estimates; arranged subject-matter assistance for tabulation orders requiring special demographic, geographic, or statistical consultation; and combined requests for similar materials, where such action was advantageous to the involved parties.

Reference Retrieval

With increasing numbers of requests from census users, it became apparent that the need for quick, accurate, and complete responses to questions about the availability and location of items in the approximately 2,000 tables of population and housing census data could not be met satisfactorily by the human memory.

Based on proposals advanced in 1966 but not budgeted until 1969, development began in June 1970 of a time-sharing computer system called RIACT (Retrieval of Information About Census Tables). As the name of the system implies, the RIACT unit of information was the table (or tabulation), specifically the frequency table. The frequency table has unique characteristics which make possible standardized descriptions of individual tables. (Structurally, frequency tables have two basic

components: (1) the universe of the tabulations--the total number of units to be distributed, e.g., persons or housing units, and (2) the stratifier(s)--the variable attribute(s)--age, sex, etc.--into whose categories the universe is distributed.)

The user consulted available printed lists of universes and stratifiers and the associated codes for the purpose of describing the desired tabulation(s) and submitted the description to the Bureau's computer by remote terminal console. The RIACT system then searched the file and indicated which tables might be used to provide the required data and whether these tables were available. Although in 1971 the RIACT system was operational for frequency tables appearing in the first through sixth count summary tapes and for the printed census reports, a final assessment indicated that it was not useful for the majority of census users, and further development was halted. The RIACT system was replaced in 1972-73 by two printed indexes, one covering the first through sixth count 1970 census summary tapes, and the other covering selected printed reports (Population Volume I; Housing Volumes I, II, and III; and the Census Tract Reports). (A preliminary index covering the first through fifth count summary tapes was issued in the spring of 1972.) In these computer-generated indexes each subject variable was listed alphabetically, with the number of categories in the variable following in parentheses. The universe of the tabulation, i.e., the type of units (persons, families, etc.) being counted, was then indicated. Opposite the subject entry was an abbreviated reference to the tape or report in which that particular subject could be found, along with the appropriate table number. In the entry example below, 4-P #57 indicated the fourth count population tapes, table 57:

Marital status (2) by own
children of selected ages,
presence of (3), by em-
ployment and labor force
status (2)

Universe: Persons,
female, 16 years old
and over

4-P #57

BIBLIOGRAPHY

- Barraclough, Robert E. "Geographic Coding." In: Federal Statistics, Report of the President's Commission. Washington, D.C., U.S. Govt. Print. Off., 1971, Vol. II, pp. 221-295.
- Beresford, John C. "Census Results As a Resource for Bank Market Research." 22 pp. (Paper presented at the 52nd Annual Convention of the Bank Public Relations and Marketing Association, Denver, Colo., October 9, 1967.)
- ."1970 Census Data Access Considerations." 15 pp. (Paper presented at the Federal Statistics Users' Conference, Washington, D.C., October 20, 1967.)
- Corbett, James P., and George L. Farnsworth. "Theoretical Basis of Dual Independent Map Encoding (DIME)." 10 pp. (Paper presented at the Conference of the Urban and Regional Information Systems Association, New Orleans, La., September 7-11, 1971.)
- Daly, Joseph F., and Robert B. Voight. "The 1970 Census Address Coding Guide Improvement Program." 9 pp. (Paper presented at the 1969 Annual Conference of the Urban and Regional Information Systems Association, Los Angeles, Calif., September 5, 1969.)
- Dolleck, Sol, and Herman H. Fasteau. "Computerized Geographic Coding." Data Processing Magazine, Vol. VIII, No. 10, October 1966, pp. 40-42.
- Dolleck, Sol; Herman H. Fasteau; and George Minton. "Development of an Automated Geographic Coding System." 62 pp. (Prepared for a meeting of the Census Advisory Committee of the American Statistical Association, November 1965.)
- Farnsworth, George L. "Current Developments in Dual Independent Map Encoding." 11pp. (Paper delivered at the Annual Conference of the Urban and Regional Information Systems Association, Clayton, Mo., September 6, 1968.)
- ."A Geographic Base File for Urban Statistical Analysis." 8 pp., attachments. (Paper presented to the Committee on Small Area Statistics at the Annual Meeting of the American Statistical Association, Pittsburgh, Pa., August 23, 1968.)
- ."Spatial Analysis." 10 pp. (Paper delivered at the Annual Conference of the Urban and Regional Information System Association, Louisville, Ky., September 1970.)
- Fasteau, Herman H. "The State of Automated Geographic Coding of Addresses." 9 pp. (Prepared for the Census Advisory Committee on Small Areas, March 1966.)
- Fay, William T. "The Geography of the 1970 Census, A Cooperative Effort." 11 pp. (Paper presented at the Annual Meeting of the American Society of Planning Officials, Philadelphia, Pa., April 18, 1966.)
- ."Problems of Area Identification." 8 pp. (Paper presented at the Fourth Annual Conference of the Council of Social Science Data Archives, UCLA, Los Angeles, Calif., June 15, 1967.)
- , and Robert L. Hagan. "Computer-Based Geographic Coding for the 1970 Census." In: U.S. Bureau of the Census. Needs and Plans for the 1970 Censuses. Census Tract Conference Papers, Series GE-40, No. 2. Washington, D.C., 1967, pp. 27-33.
- Garland, Michael G. "Geographic Tools: Aids for Data Analysis." Public Management. Vol. 52, No. 9, September 1970, pp. 18-20.
- ."New Features of the 1970 Census." 8 pp. (Paper presented at the 1970 American Institute of Planners Confer-In, Minneapolis-St. Paul, Minn., October 19, 1970.)
- Gura, Benjamin. "Program of the Data Access and Use Laboratory, U.S. Bureau of the Census." U.S. Statistical Reporter, No. 71-1, July 1970, pp. 1-7.
- ."The 1970 Census: What It Will Do." Public Management, Vol. 52, No. 9, September 1970, pp. 14-16.
- Jaro, Matthew A. "UNIMATCH: A Computer System for Centralized Record Linkage Under Conditions of Uncertainty." 25 pp. (Paper presented at the Spring Joint Computer Conference, Atlantic City, New Jersey, May 16-18, 1972.)
- "Latitude and Longitude Coordinates." Census Tract Memo, No. 7, May 17, 1965. p. 2.
- Lawrence, Charles B., Jr. "Developments in Geographic Coding in the Bureau of the Census." (Summary of remarks at the Round-Table Discussion on Uniform Coding in Federal Statistics, Federal Statistical Users Conference, Washington, D.C., October 2, 1964.)
- Meyer, Morton A. "Census Bureau's Geographic Base (DIME) Files." 21 pp., appendix. (Paper presented at the Annual Meeting of the Association of American Geographers, Kansas City, Mo., April 24, 1972.)
- ."Update and Maintenance of Census Bureau Geographic Base (DIME) Files." 7 pp., appendix. (Paper presented at the conference on "Geographic Base File System--Uses, Maintenance, Problem Solving," Arlington, Tex., November 16-17, 1971.)

- Michigan State University. Institute for Community Development and Services. Automatic Data Processing, Its Application to Urban Planning. Edited by Richard D. Duke. East Lansing, Mich., Michigan State University, 1961. 112 pp.
- Minton, George. "Inspection and Correction Error in Data Processing." Journal of the American Statistical Association, Vol. 64, No. 328, December 1969, pp. 1256-1275.
- O'Donnell, Gerald L. "A Census Mini-Glossary." Public Management, Vol. 52, No. 9, September 1970, p. 17.
- Smith, Caby C., and Marvin S. White, Jr. "Geocoding Techniques Developed by the Census Use Study." In: AFIPS Conference Proceedings, Vol. 33, pp. 609-617. (Paper presented at the Fall Joint Computer Conference sponsored by the American Federation of Information Processing Societies, Inc., Las Vegas, Nev., November 16-17, 1971.)
- Southern California Regional Information Study. Census Use Study.
1. Computer Graphics. 1970. 22pp.
 2. ACG DIME Updating System, A First Look. 1970. 26 pp.
 3. Census Data 1970, Characteristics of Cities and Unincorporated Places, San Bernardino County. 1971. 86 pp.
 4. ACG DIME Updating System, An Interim Report. 1970. 55 pp.
 5. 1970 Census Data, Characteristics of Cities and Unincorporated Places, Los Angeles County. 1971. 277 pp., supplements and maps.
 6. Population and Housing Data, City of Los Angeles. 1971. 165 pp.
 7. Background for Planning 1970. 1971. xxi, 302 pp.
 8. ACG DIME Updating System, The Long Beach California Experience. 1971. 68 pp.
 9. 1970 Census Data for Mental Health Areas. 1971. 78 pp.
 10. Research Notes: 1970-1971. 1972. 88 pp.
 11. 1970 Census Selected Social and Economic Characteristics for Cities and Unincorporated Places in Los Angeles County. 1972. 238 pp.
 12. DIME Updating System, UPDIME: A User's Manual, Los Angeles County. 1972. 90 pp.
- Systems Development Corporation. A Geographic Base File for Urban Data Systems. [Santa Monica, Calif., 1969.] 24 pp.
- Urban and Regional Information Systems Association. Urban and Regional Information Systems for Social Programs. Edited by John E. Rickert. N. p., n.d., 307 pp. (Papers presented at the Fifth Annual Conference of the Urban and Regional Information Systems Association, Garden City, N.Y., September 7-9, 1967.) Includes the following:
- "The Development of a Geographic Base File and Its Uses for Mapping," by Donald F. Cooke and William H. Maxfield, pp. 207-218.
- "Experiments in Mapping with a Geospace Plotter," by R. G. Loomis and J. J. Lorenzo, pp. 219-233.
- "Automatic Plotting and Digitizing," by Ernest Nussbaum, pp. 234-261.
- "The New Haven Census Use Study--A General Description," by Caby C. Smith, pp. 276-285.
- "Technical Aspects of the New Haven Census Use Study," by Robert B. Voight, pp. 299-303.
- Ninth Conference, New Orleans, La., 1971. Geocoding - 71. [Cambridge, Mass.] Urban and Regional Information Systems Association and Urban Data Processing, Inc., 1971. 130 pp. (Papers from the Working Session on Geographic Base File Developments.)
- U.S. Bureau of the Census. ADMATCH Users Manual. Census Use Study. Washington, D.C., 1970. x, 61 pp.
- Census Use Study.
1. General Description. 1970. 27 pp.
 2. Computer Mapping. 1969. 44 pp.
 3. Data Tabulation Activities. 1970. 23 pp.
 4. The DIME Geocoding System. 1970. 46 pp.
 5. Data Interests of Local Agencies. 1970. 92 pp.
 6. Family Health Survey. 1969. 41 pp.
 7. Health Information System. 1969. 67 pp.
 8. Data Uses in Health Planning. 1970. 48 pp.
 9. Data Uses in Urban Planning. 1970. 28 pp.
 10. Data Uses in School Administration. 1970. 30 pp.
 11. Area Travel Survey. 1970. 43 pp.
 12. Health Information System--II. 1971. 306 pp.
 13. Computer Resource Allocation Model (CRAM). To be published.
 14. Geocoding with ADMATCH, A Los Angeles Area Experience. 1971. 23 pp.
 15. The DIME File--How to Use It. To be published.
- Data Access Description.
11. Inclusion of Transient Persons in the 1970 Decennial Census. Series CEP-3. December 1969. 2 pp.
 12. 1970 Census Geography. Series CG-1. December 1969. 15 pp.
 13. First Count Summary Papers From the 1970 Census of Population and Housing. Series CT-2 (Rev.). March 1970. 8 pp.
 14. Items Contained in the 1970 Census of Population and Housing. Series CEP-1 (Rev.). March 1970. 16 pp.
 15. Contract Block Statistics Program. Series CG-2. April 1970. 10 pp.
 18. General Information About Summary Tapes. Series CT-1 (Rev.). June 1970. 6 pp.
 19. Printed Reports From the 1970 Census--Housing, Volume II. Series CEP-5. August 1970. 9 pp.
 20. Availability of Place-of-Work Data in the 1970 Census. Series CEP-2 (Rev.). November 1970. 10 pp.
 21. Census Bureau Unpublished Data and Special Services: Policy and Delineation. Series PA-1 (Rev.). December 1970. 5 pp.

22. Fourth Count Summary Tapes From the 1970 Census of Population and Housing. Series CT-5 (Rev.). March 1971. 30 pp.
23. Fifth Count (ZIP Code) Summary Tapes From the 1970 Census of Population and Housing. Series CT-6. March 1971. 10 pp.
24. Public-Use Samples of Basic Records From the 1960 and 1970 Censuses. Series CT-8. May 1971. 13 pp.
25. Third Count (Block) Summary Tapes From the 1970 Census of Population and Housing. Series CT-4 (Rev.). July 1971. 7 pp.
26. Second Count Summary Tapes From the 1970 Census of Population and Housing. Series CT-3 (Rev.). December 1971. 14 pp.
27. Printed Reports From the 1970 Census of Population and Housing. Series CEP-4 (Rev.). January 1972. 18 pp.
28. Delineation of Problem Housing Areas. Series CEP-6. May 1972. 9 pp.
29. Low-Income Data From the 1970 Census. Series CEP-7. May 1972. 15 pp.
30. Sixth Count Summary Tapes From the 1970 Census of Population and Housing. Series CT-7. (To be issued.)
- DAUList . . . User Instructions and Program Documentation. Washington, D.C., 1971-72.
- DAUList 1. /Display Programs for First Count Summary Tape/ 1971. 110 pp.
- DAUList 2. /Display Programs for Second Count Summary Tape/ 1971. 74 pp.
- DAUList 3. /Display Programs for Third Count Summary Tape/ 1971. 104 pp.
- DAUList 4. Population. Programs for Fourth Count Summary Tape 1972. 85 pp.
- DAUList 4. Housing Programs for Fourth Count Summary Tape 1972. 85 pp.
- DAUList 5. /Display Programs for Fifth Count Summary Tape/ 1972. 48 pp.
- DIME Clerical Procedures Manual, Preliminary Manual, Preliminary Report. Washington, D.C., 1970. 87 pp.
- The DIME Editing System, Preliminary Report. Washington, D.C., 1970. 107 pp.
- Geographic Base Files--Plans, Progress, and Prospects. Conference Proceedings, Jacksonville, Fla., April 1-2, 1971. Computerized Geographic Coding, Series GE-60, No. 2. Washington, D.C., 1971. 93 pp.
- Geographic Base Files--Uses, Maintenance, and Problem Solving. Conference Proceedings, Arlington, Tex., November 16-17, 1971. Computerized Geographic Coding, Series GE-60, No. 3. Washington, D.C., 1972. 70 pp.
- GRIDS: A Computer Mapping System. Census Use Study. Washington, D.C., 1972. xviii, 184 pp.
- Index to 1970 Census Maps. /Washington, D.C., 1972./ 2 vols. States Individually paged.
- Index to 1970 Census Summary Tapes: Counts 1 thru 5. /Washington, D.C., 1972./ 54 pp.
- Index to 1970 Census Summary Tapes: Counts 1 thru 6. (To be issued.)
- Index to Selected 1970 Census Reports. (To be issued.)
- New Uses of Census Resources. The Southern California Use Study and a Related Paper. Census Tract Papers, Series GE-40, No. 7. Washington, D.C., U.S. Govt. Print. Off., 1971. 52 pp.
- "Grid-Related Information Display System (GRIDS): Key to Instant Mapping of Local and Census Data," by Matthew A. Jaro, pp. 35-36.
- "DIME Applications and the Computerized Resource Allocation Model," by George L. Farnsworth, pp. 37-39.
- "ADMATCH, A Computer Tool for Urban Studies," by A. H. Rosenthal, pp. 40-43.
- 1970 Census--And You. By Gary Young. (Revised, February 1972.) Washington, D.C., 1972. 11 pp.
- 1970 Census Users' Guide. 2 vols. Washington, D.C., U.S. Govt. Print. Off., 1970. 537 pp.
- OS ADMATCH: An Address Matching System. Census Use Study. Washington, D.C., 1972. x, 66 pp.
- Papers Presented at the Conference on Small-Area Statistics, American Statistical Association, Pittsburg, Pa., August 23, 1968, and Related Paper. Census Tract Papers, Series GE-40, No. 5. Washington, D.C., 1969. 61 pp.
- Public-Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation. Washington, D.C., 1972. 203 pp.
- Small-Area Data Activities. Vol. 1, No. 1, January 1966 - Vol. 3, No. 4, December 1968.
- Small-Area Data Notes. Vol. 4, No. 1, September 1969-
- Summary Tape User Memoranda. No. 1-, November 1967-
- Use of Address Coding Guides in Geographic Coding. Case Studies. Conference Proceedings, Wichita, Kans., November 19-20, 1970. Computerized Geographic Coding, Series GE-60, No. 1. Washington, D.C., 1971. 75 pp.
- USES Report No. 1. Census Fourth Count Summary Tape: Selected Socio-Economic Characteristics of Marion County, Indiana. Census Use Study. Washington, D.C., 1972. 400 pp.
- U.S. Urban Renewal Administration. Using Computer Graphics in Community Renewal. Community Renewal Program Guide No. 1. Washington, D.C., 1963. Various paging.