DOCUMENT RESUME

ED 078 816 JC 730 153

AUTHOR Landini, Albert; Bannister, R. Wayne
TITLE Student Residence Locations and Associated Census Data. Interim Report.
PUB DATE Jun 73
NOTE 19p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Census Figures; Community Colleges; *Computer Programs; *Data Bases; Demography; Disadvantaged Groups; Handicapped; *Maps; Post Secondary Education; *Student Enrollment; Technical Reports

ABSTRACT

The primary purpose of this project was to build a data base that would aid in identifying disadvantaged persons and later in identifying handicapped persons in the services areas of the Los Angeles Community College District. Five main steps were involved: (1) finding a machine readable file of all enrolled students in the District; (2) defining the District in terms of consus tracts included within it; (3) obtaining and installing the necessary software and reference files at the District's computer center and preparing special computer programs; (4) preparing the geographic data for input to the computer mapping program; and (5) using specialized software to add census tract identifiers to each student record. Several technical tools were produced and/or installed for the District: L.A. Co. ACG, a geographic base file; ADMATCH, an address matching program; SYMAP, a computer mapping program; base map of district outline on county census tract map; listing of all 1970 census tracts in county and in District; "An Outline" L.A. Co.; "B Data Points" county 1970 census tract centroids; "C Otolegend" District outline; ten student enrollment maps; eleven census data maps; student enrollment report; and census data report. (KM)

FILMED FROM BEST AVAILABLE COPY

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

ED 078816

STUDENT RESIDENCE LOCATIONS AND ASSOCIATED CENSUS DATA

INTERIM REPORT

by

Albert Landini
Project Director

Community College Data Base Team

Mark Oleesky
Ralph Carlson
Fred Ige
Steve Haggerty
Marie Sanchez
Charles Montgomery

R. Wayne Bannister
Technical Services Supervisor

JC 730 153

June 1973

In cooperation with the Los Angeles Community College District Institutional Research Council: Dr. Arthur Cherdack, President; and Dr. Ben K. Gold, Project Coordinator.

TABLE OF CONTENTS

NOTE: Appendix material excluded due to marginal reproducibility. This material is available from Ben K. Gold, Director of Research, Los Angeles City College, Los Angeles, California 90029

STUDENT RESIDENCE LOCATIONS AND ASSOCIATED CENSUS DATA

Conversations with faculty and staff of the Los Angeles Community College District who were interested in institutional research, revealed their growing awareness of geographically based information systems and how they could be used for a wide range of information processing tasks to aid the educational decision maker. This report deals with one of those tasks; the building of an initial District-wide data base from which to conduct an investigative evaluation of the areal distribution of disadvantaged students and their associated socio-economic and demographic characteristics.

I. INTRODUCTION

Preliminary investigation of the District's planning procedures for disadvantaged students indicated that they were somewhat confined by the organizational and student body requirements of each of the eight individual colleges. Efforts were being made by each school, independently from one another, to respond in a more direct fashion to perceived community needs, but systematic methods seemed lacking in ascertaining those needs on a District level. In fact little documentation was available substantiating that the community colleges completely understood the socio-economic make-up of all residents in their individual service areas, particularly those outside their immediate client group. Without such information they were

unable to respond totally to those peoples academic needs, and human and social wants. The system described here represents a first effort to produce those kinds of socio-economic information to enable appropriate responses to the community by the individual colleges or total District.

II. PURPOSE

The primary purpose of this project was to build a data base that would specifically aid in identifying disadvantaged persons and later be of use in identifying handicapped persons as defined in the California State Plan for Vocational Education, 1969, who resided within the service areas of the eight colleges that comprised the Los Angeles Community College District. The Vocational Education Plan placed the responsibility for identifying those persons directly with the local educational agency.

The definition of handicapped persons was limited to including those individuals who were mentally retarded, hard of hearing, deaf, speech imparied, visually handicapped, emotionally disturbed, crippled, or other health impaired persons who because of their handicap could not succeed in vocational education programs. It was deemed that at this time these people as a group were best identified through more traditional processes of survey and interview; or through the use of data files other

than those available to the project at this time.

However, disadvantaged persons were defined as those having academic, socio-economic, cultural, or other handicaps that prevented them from succeeding in vocational education programs designed for persons without such handicaps, and who for that reason required specially designed educational programs or related services. The term included persons whose needs for such special programs or related services resulted from poverty, neglect, delinquency, or cultural or linguistic isolation from the community at large, but did not include physically or mentally handicapped persons. (emphasis added).

The products of this project would allow the educational decision maker to make some judgements about the number of possible disadvantaged students that his school was serving, and the nature of the disadvantaged students being served. Visual correlations of the spatial distributions of student enrollment patterns with certain census variables would allow the concerned college staff to make some judgement as to how well disadvantaged persons were being served by the District through its individual colleges. Additionally the necessary data was developed for later use in statistical tests of associations and differences to verify some of the hypotheses developed in the initial visual examination of the mapped data.

III. METHODOLOGY

To achieve the stated purpose, several detailed steps were undertaken. They are described here in general terms, but more detailed technical memoranda were produced throughout the project to document the process for future replication. Those technical memoranda are not included in this report.

The first step in the process was to identify if a machine readable file of all enrolle students for each school in the District existed. It was found that such files did exist and were already collapsed to a total (summary) file of all students, for all colleges, for the entire District. Each record on this file had two pieces of information critical to the project: (1) the student's house address, and (2) the school attended.

Next it was necessary to areally define the District in terms of census tracts included within it. This was important as our purpose was to later produce a report of the number of students living in each census tract who attended a given college. This laborious task of identifying the census tracts making up the District was achieved by transfering the District boundary from a series of County Registrar of Voters maps to a census tract map for the entire Los Angeles County. Once this was completed a tabular report (not included here) was produced listing all of the tracts falling within the District or that portion of those that were split by the boundary line.

A third step required that the necessary software and reference files be obtained and installed at the District's computer center that allowed us to add census tract identifiers to each student record, and to produce computer drawn maps of our data. Additionally, special computer programs were also prepared to produce reports from machine readable census data files and the tracted student enrollment data for later tabulation purposes.

In the fourth step the College Data Base Team faced the task of preparing the geographic data for input to the computer mapping program. To do this required that they identify a number of points that made up the outline of the entire County. Using hand digitizing methods, x and y coordinates were produced for this County outline. Similar procedures were followed to produce x-y coordinates outlining the District boundary within the County, and the visual centers (centroids) for each of the census tracts in the entire County.

The final step required the use of specialized software to add census tract identifiers to each student record. The resultant data along with selected census data was mapped and tabular reports prepared. This was followed by a final report prepared for presentation and documentation purposes.

IV. TECHNICAL TOOLS USED

In the development of the data base reported on here, three

specific technical tools were used: (1) L.A.Co. ACG, a geographic base file; (2) ADMATCH, an address matching program; and (3) SYMAP, a computer mapping program. These technical tools were largely outgrowths of efforts surrounding the 1970 Census. They were designed to enable more effective use of census and local data by local government officals and to assist them in their decision making process. Such tools have had increasing importance: First, because 1970 Census data was available on Census Summary Computer Tapes; and Second, because governmental agencies are more and more maintaining an increasing amount of their records in computer form or in a form easily converted for use on computers. This means that a wealth of information valuable to research professionals is or will be computerized.

L.A.Co. ACG. The Los Angeles County Address Coding Guide is basically a computer readable urban map. That is, characteristic s on a map, such as street intersections and street address ranges have been recorded onto computer tape. Also included in this geographic base file are so-called "geographic codes". These are always such codes as census tract and census block and may include special area codes such as school attendance areas, or community planning districts. These geographic codes are provided for each street covered by the geographic base file. ACG is simply a method for producing such a file.

An Address Coding Guide (ACG) geographic base file can be used with any local data records that are in computer form (cards or tape) and that have addresses on the individual records. An ACG provides the base for assigning geographic codes to local data records. With ADMATCH, a computer tool described later, codes from an ACG can be transferred to data records containing street addresses. This means that data can be aggregated to any geographic area desired, if that area identifier has been coded in to the ACG. Thus Los Angeles Community College District enrollment figures for each college can be studied at the same level of aggregation after ADMATCHing. In addition since the census tract was chosen as the unit of aggregation, these enrollment figures can be compared to census data at the tract level, and more importantly to other local data also aggregated to the census tract level.

ADMATCH. This is a package of user-oriented computer programs and documentation designed to assist in the assignment of geographic codes to computerized data records containing street addresses. Geographic codes for areas such as census tracts, school districts and traffic zones can be readily assigned to data records in various organizational files. The ADMATCH system was developed because the assignment of geographic codes to data records through address matching has become a very important step, not only in creating urban and municipal information systems, but also in many ongoing planning and evaluation studies.

The customary method of address matching consists of the manual process of locating the address of a data collection form on a map or street index and recording the appropriate geographic code on the form. However, this tedious and error prone method has been largely replaced by the advent of machine matching techniques. ADMATCH, as in hand matching, also requires a reference source similar to a map or street index. However, the reference source for ADMATCH must be computer readable. The Census Bureau, assisted by local planning agencies in most of the Nation's metropolitan areas, has prepared reference sources that can be used with ADMATCH. An ACG geographic base file, described earlier, is one such reference source. Locally prepared street indexes with address ranges and geographic codes can also be used as references files. ADMATCH links the data records to the reference file records by matching the address on the data record to the address range in the reference file. When a match occurs, a geographic code from the reference file is attached to the data record.

Thus the subject of this paper is a perfect example of address matching in which census tract codes were assigned to student enrollment records. This data was then aggregated by tract for each college as an aid to determining individual college service areas and relationships to census information pertaining to disadvantaged students.

SYMAP. SYMAP is the best known, most comprehensive, and most widely used of the many available computer mapping programs. It was developed by the Harvard Laboratory of Computer Graphics and Spatial Analysis. The program enables three basic types of maps to be produced which, essentially, allow for three ways of shading areas on a map. Mathematical operations can also be performed in conjunction with SYMAP. In addition, SYMAP can provide textual definitions of geographic areas and show actual data values on a map.

The display of various types of data on maps can be especially useful for understanding the geographic implications of data for a heterogeneous area, whether it be for a college service area, or an entire city or region. Planners and analysts often times find it difficult to see differences in geographic data when it is simply presented in masses of tabular printout. On maps, however, large amounts of data can be displayed in varying patterns, and as a result, spatial relationships and trends can become immediately apparent.

Graphic indications of general geographic patterns and trends are useful for planning and developing guidelines for further statistical analysis, as well as for administrative and public information purposes because of their visual impact and clarity. For example, coventional tabulations of various neighborhood socio-economic indicators presented at a public meeting may be ineffective and almost meaningless. Lines of bar charts,

ordinarily an effective means of graphic display, become too confusing when more than two or three service areas are simultaneously represented. But use of a few well chosen data maps, displaying the same indicators, can make the same comparisons immediately evident.

Until recently, data maps have been produced only by draftspersons. However, with the advance of computer technology, the quick and efficient production of various types of data maps drawn by computer is now not only possible but relatively easy and inexpensive. Four elements are involved in any computer mapping process:

- Selecting and specifying data to be mapped.
- Linking the data file to a geo-base file.
- Manipulating and organizing data to equipment and program constraints.
- Producing a special features overlay to emphasize certain aspects of the data.

These elements were all integrated into the maps shown in this report.

V. MAPS OF STUDENT DATA BY CENSUS TRACT*

Ten maps of student enrollment data were produced. One for each college illustrating the number of students residing in each census tract that attended a particular college. A ninth map was a map of all students who were enrolled in Instructional Television (ITV). The tenth map was in a sense a summary map showing the total number of students residing in any census tract regardless of the school or program they attended.

*See note in table of contents

These maps were intended to show the areal extent of the individucal college's and total District service area(s), in a sense they were a graph of any one school's ability to draw students in. The maps provide a means of comparing the relative service areas for any two or more colleges. They allowed the judging of existing physical and socio-economic attributes surrounding the college that facilitated or hindered the size of enrollment.

The Student Data Maps produced were:

Maps of Student Data by Census Tract

1. Students Attending Pierce College.
2. Students Attending Valley College.
3. Students Attending Los Angeles City College.
4. Students Attending Trade Tech.
5. Students Attending East Los Angeles College
6. Students Attending Southwest Los Angeles College.
7. Students Attending West Los Angeles College.
8. Students Attending Harbor College.
9. Students Attending Instructional Television.
10. Students Attending Los Angeles Community College District Institutions.

VI. MAPS OF CENSUS DATA BY CENSUS TRACT

The 1970 Census information was released in both traditional printed form and in the form of computer tapes. On these tapes (1st, 2nd, 3rd, and 4th count) were li terally thousands of data items that could be associated with each census tract and that in some way described the nature of the population and housing in those tracts.

From that vast universe of itemsthe Data Base Team selected eleven that seemed most critical to many people from the various colleges interested in disadvantaged students. As such they did not represent a final compendium of data but rather served as pointers to the types of data that can be associated with student enrollment data at the census tract level.

Census data decays over time and because of this it was anticipated that future replications of this project will see census data replaced with other socio-economic data collected at the local level that is also aggregated to the census tract. Numerous governmental and private agencies have been aware of this decay potential and have banded themselves together in informal and formal relationships to facilitate the future interchange of information.

The Census Data Maps produced were:

Maps of Census Data by Census Tract

1. Race Data
   a. Percent White Population 1970.
   b. Percent Black Population 1970.
   c. Percent Spanish Surname Population 1970.
   d. Percent Oriental Population 1970.
   e. Percent "other" Population 1970.

2. Age Data
   a. Percent of Population Aged 19-25 in 1970
   b. Median Age of Population 1970.
   c. Percent of Population Aged 62 and Over 1970.

3. Income Data
   a. Median Family Income 1970.
   b. Percent of Population Below Poverty Level 1970.

4. Education Data
   a. Median Years of Education Completed for Persons 25 and Over 1970.

VII. TABULAR REPORTS

Two tabular reports were also produced listing the raw information used in producing the data maps. These reports were anticipated as being important sources of information for those persons wishing to do more detailed analysis of student characteristics for individual college service areas. The data used to produce these reports was presented in machine readable form to representatives of the Los Angeles Community College District Institutional Research Council. The reports produced were:

Tabular Reports

1. Number of Students Enrolled by Census Tract by Individual College.
2. Number of Persons by Census Tract by Each of Eleven Data Items.

In using the material from these two reports, it was believed that the researcher would be accustomed to the consideration that data processing of this magnitude is subject to some error. In the report of student enrollment by census tract by college, error could have resulted from faulty information in either the student data file or the ACG reference file. Statistics generated in the matching process indicated that we were approximately 80% accurate in our assignment of census codes

to student records. The remaining 20% of student records were apportioned to census tracts within colleges by a factor computed from known total enrollments for each college.

Error in the census data was the result of Census Bureau collection, processing, or more likely the enforcement of census suppression rules. These rules are developed and enforced to minimize the liklihood of revealing information about any individual.

VIII. INTERPRETATION GUIDELINES

Staring at data is a long term avocation of institutional researchers. In the case of map interpretation it can become a full-time vocation. In these maps and reports are the basic tools for testing postulated hypotheses growing out of the initial map investigation. These hypotheses will grow out of questions resulting from viewing the maps.

Some basic comparisons are suggested as starters. First, look at the student enrollment map for the college most interesting to you. How does the distribution of students compare to neighboring colleges, the District boundary, and the various elements represented in the census data maps? Are there any clumpings of students? Are there areas where there are no students, and you would expect or desire some?

Some additional suggestions could be posed as tenative research questions of particular interest to those dealing with the disadvantaged:

Federal Fund Raisers can ask:

* What is the real service area of the college?
* What are the socio-economic characteristics that can be associated with its students?
* How many of the students, actual or potential, live in specially funded areas such as model cities and model neighborhoods?

Curriculum Designers might want to know:

* How many disadvantaged students living outside the normal service area travel to this institution for special courses?
* What are the characteristics of these students exerting extra effort?
* How well does the college's curriculum relate to that of high schools producing most of its disadvantaged students?

Recruitment Counselors may want to consider:

* From what part of the real service area do most of the disadvantaged students come from?
* What classes do these disadvantaged students take?
* Where should we look to find more of these disadvantaged students?

School Planners concerned with Community Impact could question:

* How well does the service area relate to other special governmental areas?
* How often does the college coordinate community impact programs with those initiated by other agencies for a given area?
* How well does the college respond to stated and measured community needs?

IX. CONCLUSIONS

This project was a first step in building a District wide data base for the Los Angeles Community College District. The

primary mission was to produce maps, reports and data that would aid educational decision makers in identifying the residential location of disadvantaged students and their proximity and relationship to individual college service areas and existing student enrollments.

In fulfilling this mission the Community College Data Base Team produced and/or installed several technical tools for the District under the auspices of their Institutional Research Council. Those tools were:

- L.A.Co. ACG address reference file
- ADMATCH
- SYMAP
- Base Map of District Outline on County Census Tract Map
- Listing of all 1970 census tracts in County and in District
- "A Outline" L.A.Co.
- "B Data Points" County 1970 census tract centroids
- "C Otolegend" District outline
- Ten student enrollment maps
- Eleven census data maps
- Student enrollment report
- Census data report

These tools were produced and/or delivered in anticipation that they would provide the District institutional researchers with the necessary means to conduct systematic and more extensive research into the dynamics of the disadvantaged. These tools also provided the District with a means of producing time series data about the changing nature of college service areas and student characteristics.

UNIVERSITY OF CALIF.
LOS ANGELES

JUL 18 1973

CLEARINGHOUSE FOR
JUNIOR COLLEGE
INFORMATION

Sample Map from Appendix I.

STUDENTS LOS ANGELES COMMUNITY COLLEGE DISTRICT