

DOCUMENT RESUME

ED 333 889

IR 053 633

AUTHOR Wallace, Danny P.; And Others
 TITLE Age Analysis of Public Library Collections. Final Report.
 INSTITUTION Louisiana State Univ., Baton Rouge. School of Library and Information Science.
 SPONS AGENCY Department of Education, Washington, DC.
 PUB DATE 90
 CONTRACT R039A90006
 NOTE 135p.
 PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC06 Plus Postage.
 DESCRIPTORS *Decision Making; *Evaluation Criteria; Library Automation; Library Circulation; Library Collection Development; *Library Collections; Obsolescence; *Public Libraries; Statistical Analysis
 IDENTIFIERS *Aging (Library Materials); *Weeding (Library)

ABSTRACT

The use of information regarding the ages of library items is a standard component of many approaches to weeding library collections, and has a long history in the literature of collection management. Current and past approaches to using aging information to make weeding decisions make use of very arbitrary decision criteria. This study examined the actual aging patterns of public library materials as a precursor to the production of a set of collection management procedures based on sophisticated analyses of aging data. The operational objectives of the study were (1) to develop a set of profiles of the aging characteristics of public library collections, based on analysis of a body of diverse public library holdings data; (2) to formulate a set of collection management guidelines based on the aging profiles, with detailed descriptions of appropriate analyses and their interpretation; and (3) to design an easy to use, flexible computer program to aid in the application of the guidelines. This report covers the problem statement and objectives of the study; the background and context of the study; approaches to weeding library collections; methodology; sample; data collected; data analysis; sample application of age data to collection analysis (Profile 1--the 90% retention rate and Profile 2--curve examination); comparison of the percentage and curve examination methods; continuation of the project; and dissemination of project results. Thirteen notes/references are included, and 109 figures are appended. (BBM)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

IR

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it
 - Minor changes have been made to improve reproduction quality
-
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED333889

AGE ANALYSIS OF PUBLIC LIBRARY COLLECTIONS

FINAL REPORT, USDE GRANT #R039A90006

Danny P. Wallace, Associate Dean
Bert R. Boyce, Dean
School of Library and Information Science
Louisiana State University

Kathleen M. Heim, Dean
Graduate School
Louisiana State University

John P. McLain, Programmer
Louisiana Cooperative Extension Service

Donald H. Kraft, Chair
Department of Computer Science
Louisiana State University

R053633

**AGE ANALYSIS OF PUBLIC LIBRARY COLLECTIONS
FINAL REPORT, USDE GRANT #R039A90006**

Contents

Acknowledgments	ii
Abstract	iii
Problem Statement and Objectives	1
The Background and Context of the Study	3
Approaches to Weeding Library Collections	4
Methodology	8
Sample	9
Data Collected	10
Data Analysis	12
Sample Application of Age Data to Collection Analysis	16
Profile 1: The Ninety Percent Retention Rate	17
Profile 2: Curve Examination	17
Comparison of the Percentage and Curve Examination Methods	19
Continuation of the Project	19
Dissemination of Project Results	21
Notes	23
Figures	

Acknowledgments

Any project of this scope necessarily involves the contributions of a large number of people, some of whom devote many hours to the project and some of whom are relied upon for only a few minutes. The authors wish to express their particular appreciation to the following individuals for their assistance in support of this work: Ms. Diane Hudson, who served as graduate research assistant for the project and was responsible for nearly all of the computerized data analysis and generation of graphs; Mr. Benjamin P. Bryant, graduate research assistant to the principal investigator, who was always willing to take on added work to support the project; Ms. Vivian A. Auld, Computer Resource Coordinator for the MSU School of Library and Information Science, who served as trouble shooter and general resource person for a number of research activities; the staff of the OCLC Office of Research, and especially Dr. Martin Dillon, Director, and Mr. Mark A. Crook, Consulting Analyst, without whom the data would not have been supplied and the project would have been impossible.

AGE ANALYSIS OF PUBLIC LIBRARY COLLECTIONS

FINAL REPORT, USDE GRANT #R039A90006

Abstract

The use of information regarding the ages of library items is a standard component of many approaches to weeding library collections, and has a long history in the literature of collection management. Current and past approaches to using aging information to make weeding decisions make use of very arbitrary decision criteria. This study examined the actual aging patterns of public library materials as a precursor to production of a set of collection management procedures based on sophisticated analyses of aging data.

AGE ANALYSIS OF PUBLIC LIBRARY COLLECTIONS

FINAL REPORT, USDE GRANT #R039A90006

Problem Statement and Objectives

The planned removal of obsolete materials from public library collections is essential to the maintenance of quality library service. Although many strategies for identifying obsolete items have been published, none of them is based on an empirical examination of the aging characteristics of library collections. Selection of items for "weeding," or removal from the active collection, is an ongoing problem in public library operations. Library patrons generally need and desire items that are most up-to-date. This is particularly true of works in certain subject areas, such as science, technology, political science, and medicine, but is also true of all areas of the library's collection. Unused items generate clutter on the shelves, and may have a seriously detrimental effect on the patron's assessment of the library and its collection. Space is generally at a premium in public libraries, and unused and unusable items compete for space with high-use, highly useful items.

Current decision making regarding removal of items from public library collections focuses on items that have deteriorated physically, or makes use of arbitrary criteria for evaluating the use of materials or the age at which an item should be removed from the collection. The general objective of the project described here was to develop a more precise, scientific basis for making collection management decisions based on the aging of public library materials.

The project described here should be of specific benefit to public libraries and their patrons. The project makes use of data gathered from automated systems to provide accurate data regarding the aging of library collections, and to develop analytical tools for identifying aging patterns for library materials. Although most automated circulation systems are currently capable of providing management reports pertaining to the ages of circulated and, sometimes, uncirculated materials, they are not currently capable of providing detailed analyses of such aging data. At the same time, librarians are generally not familiar with the basic methods for systematically analyzing the aging of library collections, and are therefore unable to develop effective, sophisticated strategies for utilizing the reports generated by their circulation systems.

The specific operational objectives of the study were

1. to develop a set of profiles of the aging characteristics of public library collections, based on analysis of a body of diverse public library holdings data;
2. to formulate a set of collection management guidelines based on the aging profiles, with detailed descriptions of appropriate analyses and their interpretation;
3. to design an easy to use, flexible computer program to aid in the application of the guidelines.

The objectives encompassed a combination of pure and applied research goals. At a pure research level, the aging patterns of public library collections are of interest simply because they are an unknown. From an applied, action oriented direction, the project was very deliberately aimed at the development of collection management procedures based on age analysis. Success in achieving the project's objectives is the focus of this report.

The Background and Context of the Study

Most experienced librarians are well aware that items in their collection tend to attract less use as they age. In many libraries this factor is used as a rationale for discarding older items, and the process of determining the point at which an item has become so old that it should be discarded is an important collection management issue. This aging process has also been observed in the context of the references included in scholarly publications: most references tend to be to relatively recent publications, and the likelihood of a publication being cited appears to decline over time. This aging process is generally referred to under the bibliometric category of "obsolescence," although the aging of information sources and the obsolescence of a technology or methodology clearly are not directly analogous. When a piece of machinery is said to be obsolescent, there is usually the implication that it has been replaced by a better piece of machinery and is therefore no longer of use. Obsolescence in bibliometrics, however, suggests only that older materials are not used, not that they are no longer useful.

Obsolescence has usually been studied in the context of the circulation of items in a library collection, or of the citation of one body of literature by another. The results of obsolescence studies are quite consistent: when items are ranked according to their age at the time they are used (circulated, requested, cited, etc.), recent items account for a very large proportion of the items used, while very old items receive very little use. The principles underlying the study of obsolescence are also of substantial potential use in library collection management:

If documents are being considered, the interest is probably a practical one in the probability that an item will be required, as a guide to such questions as when to discard older volumes, how long to keep new ones, what sort of retrospective storage and access an information retrieval system should provide, and so on.¹

The relationship between obsolescence and use has not been adequately explored. Although use of the collection is an important goal for any library, age data may well constitute an important independent factor in collection analysis. The idea that older materials may legitimately be either discarded or relegated to some remote storage area is familiar to all librarians. The problem of employing strategies for discarding or moving to secondary storage lies in determining when an item is old enough to be removed from the primary collection, and decisions are frequently made on the basis of *ad hoc* rules of thumb or vague guesses. Determination of the actual patterns with which use of the collection declines over time can help make it possible to make more informed decisions and reduce the potential for making incorrect decisions. The use of systematic obsolescence studies in collection management is at present hampered by the difficulty of gathering appropriate data. The Age Analysis of Public Library Collections project explored the ways in which machine-readable data can be used in analyzing the obsolescence of public library collections.

Approaches to Weeding Library Collections

Weeding library collections has a number of potential benefits, including effective utilization of space, efficient employment of staff and patron time, increasing the appeal of the library's collection, enhancing the reliability of the collection, monitoring the physical condition of materials in the collection, and providing a basis on which to evaluate the strengths and weakness of the collection.² According to Lancaster, "weeding is necessary because the value of an item to a collection changes over time."³ It has also been suggested that effective weeding has a positive effect on circulation, although no conclusive study of the relationship between weeding and circulation has yet been presented.⁴

Traditional approaches to weeding concentrate on evaluation of the physical appearance of library materials, elimination of unnecessary duplication, evaluation of the content or format of materials, use, and aging.⁵

Evaluation based on physical or content considerations requires actual examination of the material under consideration. Systematic examination of an entire library's collection is an overly time-consuming activity, and the number of items actually identified for removal may not justify such expenditure of time. Furthermore, these approaches require a substantial level of judgment. Physical deterioration alone is rarely an appropriate criterion for weeding, since such deterioration may accompany items that are old and rare, inherently fragile items that have been damaged by poor storage and/or handling, or high use items that need to be repaired or replaced rather than weeded. Evaluation of the content of a work is always complex, and requires an intellectual judgment on the part of a skilled evaluator. Because evaluation of content is time-consuming and uncertain, it is of limited use as a primary means for identifying items for weeding.

Use of library materials is a consideration that has long played an important part in collection management. Use is generally measured by circulation, and the fundamental assumption of use studies is that items that are not used should not be retained. The most common approach involves examining circulation for some fixed period of time, and targeting items that did not circulate during that time for removal from the collection.⁶ Although it is impossible to argue against the principle that library materials are acquired and maintained with their use as a primary objective, there are several problems with noncirculation as a primary criterion for weeding. The most obvious problems relate to identifying an appropriate sampling period for gathering circulation data and the balance between circulation and in-house use. The greatest limitation of use analysis is that circulation counts provide no information at all regarding the reasons why items are or are not used. Obsolete items may circulate, while nonobsolete items may not. In some cases, the ab-

sence of current materials may force library patrons to check out material that is out of date. Under other circumstances, the perception that a section of a collection contains mostly old material may discourage circulation of items in the section that are not obsolete.

Current weeding methods based on age generally make use of rather arbitrary criteria for different classes of items. A classic discussion of this approach was McClellan's methodology, which combined age categorizations with physical condition and use considerations.⁷ A typical scheme offers suggestions such as discarding reference works after ten years or when they are superseded, maintaining works in religion and philosophy virtually forever, providing "frequent revision" for the social sciences, and maintaining science and technology materials for no more than ten years.⁸ These age categories appear to have been derived from the collective personal experiences of librarians, and there is no evidence of an empirical basis for them. There appears to be general agreement that libraries should discard old materials, but there is no systematic approach to defining "old." The project described here was designed to provide librarians with a sophisticated set of procedures for determining what proportion of a collection can be considered old, and for identifying materials so old that they are likely candidates for weeding.

In a more systematic approach, some set of characteristics of the collection itself would be used to select "target" items. The sample of target items would then be collected in one place, presumably by a page or other nonprofessional employee, and examined by a librarian or higher level nonprofessional to determine what action should be taken. The actions to be taken could include

- a. removal from the collection (the classic sense of the term "weeding"),
- b. conversion to an alternative format such as microreproduction or an electronic medium,
- c. repair or replacement,
- d. enhancement to encourage use,

- e. relegation to some secondary storage site,
- f. transfer to another library where use or usefulness might be higher, or
- g. retention of the item as is.

Information that could contribute to the decision includes

- a. circulation or other use history,
- b. physical condition of the item,
- c. appraisal of the item's intrinsic value,
- d. evaluation of the item's value to the collection,
- e. language of the item,
- f. number of copies of the item held in the collection,
- g. availability of the item elsewhere,
- h. availability of the item in other formats, and
- i. index or bibliography coverage of the item.

Although this process of examination can be done for a collection as a whole, the time required for doing so is often intimidating. A reasonable estimate of an efficient weeding process might require an average of twenty minutes of professional level time for each item examined. For the smallest library in the study database, with a total collection size of 14,849 items, spending twenty minutes per item on the examination process would require 4,950 person/hours, equivalent to more than two years of professional level effort. For the largest library in the database, with a collection of 98,085 items, more than forty-seven years of effort would be required. This does not account for delays due to misshelving, items in circulation at the time sought for examination, or other delays in locating items. It also does not include the time required for assembling the materials prior to examination. Spending less time per item will accelerate the process and reduce person/hours but will inevitably reduce the quality of the decisions made. Although some libraries conduct

regular, complete inventories for weeding and other purposes, a substantial disruption of normal services is generally involved. For larger libraries, any inventory approach is simply overwhelming.

Methodology

Data gathering and data manipulation are the major problems of a study of the aging characteristics of public library collections. Although gathering data directly from local automated systems for a sample of libraries might be desirable, such an approach presents problems of data compilation and comparability of data formats that would undoubtedly be substantial. As a substitute for data from local systems, this study used data from the OCLC (Online Computer Library Center) union catalog. The original plan of work entailed the following steps:

1. Demographic data from the *American Library Directory* and the *Public Library Data Service Statistical Report* were to be used to develop profiles for ten demographically diverse public libraries. Libraries would be selected on the basis of geographic location, size of collection, population of community and other demographic characteristics. The objective for giving consideration to demographic data was to identify libraries representative of a cross-section of U. S. public libraries. The desired demographic profile called for four libraries with collections of fewer than 20,000 items, three libraries with collections of between 20,000 and 70,000 items, and three libraries with collections of more than 70,000 items. Two libraries were to be located in the Northeast, three in the Midwest, three in the South, and two in the West. Two were to represent nonmetropolitan areas, three to represent central city metropolitan areas, and five to represent suburban metropolitan areas.

2. OCLC, Inc. would match the ten profiles to ten public libraries represented in the OCLC online union catalog, and provide the investigators with a tape copy of the complete holdings data for each of the ten libraries. Data provided to the investigative team would not include identification of the ten libraries, thereby assuring confidentiality.
3. The aging characteristics of the data would be analyzed as a prerequisite to development of a model for incorporating aging data into weeding decisions and design of specific techniques for doing so. The model was to be tested by the development of weeding profiles

The original plan was modified in a number of ways, most of which arose from the selection of OCLC as a source of data. The demographic profiles originally developed proved to be too complex for easy matching to OCLC's database, and were therefore simplified to represent only rough size and geographic area considerations. Even then, drawing data for ten libraries was found to be a greater drain on OCLC resources than was anticipated, and a compromise sample of seven libraries was selected.⁹

Sample

The nature of the sample of seven public libraries is summarized in Table 1. There were no very small libraries, because such libraries are not represented in the OCLC database. There were no very large libraries, due to the problem of selecting data from the database and delivering it in an acceptable format within a reasonable period of time; including larger libraries would have delayed the project further and probably would have exacerbated file conversion problems. In comparison to the desired demographic breakdown, the sample has one fewer library in the under 20,000 holdings category and two

fewer in the more than 70,000 category. The libraries sampled can all be considered to be within the midrange of U. S. public libraries. All regions of the nation except the south were represented. No information on the communities served by the libraries is available, so the distribution among metropolitan, suburban and nonmetropolitan is unknown, since the identities of the libraries are not known to the researchers.

TABLE 1: SAMPLE BREAKDOWN

=====			
LIBRARY	REGION	TOTAL HOLDINGS	USABLE HOLDINGS
=====			
1	Midwest	14,849	12,899
2	Midwest	22,409	20,828
3	Midwest	98,085	96,773
4	Northeast	16,326	14,892
5	Northeast	43,260	43,000
6	West	18,224	18,176
7	West	37,190	36,925
=====			
TOTAL		250,343	243,493

Data Collected

The data provided by OCLC identified 250,343 publications held by the seven libraries. Missing data elements and other flaws in the sample data resulted in a usable database of 243,493 records. Each record consisted of the OCLC identification number for each item, a unique library identification number (a sequential number from one to seven),

the language of the publication, date information from the OCLC MARC fixed field, additional date information from the \$C subfield of field 260, and call number from either the 082 or 092 field (with the 082 field preferred). These data can be summarized as follows.

Language. Sixty-seven language codes were represented in the sample. English language publications were overwhelmingly predominant, accounting for 97.3% of the sample, including 99.4% of Dewey classified items and 90.5% of non-Dewey items. No other language accounted for as much as one percent of the sample. This paucity of foreign language obviates the utility of conducting a comparative analysis of the aging characteristics of publications based on language of publication.

Date. Preference was arbitrarily given to the date from the OCLC fixed field. This differed from the \$C/260 field date in only a very small number of cases. Some editing was required to eliminate obviously bogus dates, including dates in the distant future and dates represented by numbers of less than four digits. No attempt was made to correct these erroneous dates; they were simply recorded as missing data. The oldest date recorded was for a non-Dewey item published in 1800; the oldest Dewey item was published in 1850. Date information was used to calculate ages of items, using 1990 as a base. Ages therefore ranged from zero years to 190 years, with a mean of 8.72. The mean for Dewey items was 8.01 years, and the mean for non-Dewey items was 11.01 years.

Call number. The sample was divided into two major groupings for purposes of analysis: Dewey and non-Dewey. Items with special additions to the call number field, such as "ya" or "x", were consolidated with items not possessing these amendments. Dewey items were analyzed in depth. Non-Dewey items have as yet been analyzed only in a very simple manner due to the complexity of dealing with the variety of non-Dewey representations of the call number field. Counts of Dewey and non-Dewey items and the percentage of Dewey items for each library are shown in Table 2. Dewey items accounted for between 73.8 and 85.1 percent of the collection, with an aggregate percentage of 77.4 per-

cent for the database as a whole. This very high percentage of Dewey items may suggest that the libraries for which data were analyzed have been more assiduous in reporting holdings for the classified portions of their collections than for fiction and other unclassified materials. The libraries are very similarly balanced in terms of where their holdings lie. Pearson correlations among the seven libraries by numbers of items in each Dewey class (broken down to the Dewey tens) yielded coefficients between .92 and .99, indicating that the balance of holdings in each Dewey tens class was very similar for all the libraries in the sample.

TABLE 2: BALANCE OF DEWEY AND NON-DEWEY HOLDINGS

LIBRARY	DEWEY HOLDINGS	NON-DEWEY HOLDINGS	DEWEY PERCENT
1	10,435	2,464	80.9%
2	17,724	3,104	85.1%
3	71,405	25,368	73.8%
4	11,397	3,495	76.5%
5	35,895	7,105	83.5%
6	13,773	4,403	75.8%
7	27,786	9,139	75.2%
TOTAL	188,415	55,078	77.4%

Data Analysis

The major focus of the data analysis was the search for patterns within the relationship between call number ranges and ages of items. Analyses were conducted in a number of ways, including examination of all libraries as a group in addition to individual libraries,

all Dewey items as a group as well as several variations on Dewey ranges, and all non-Dewey items as a group. Non-Dewey items were not further subdivided, and all analysis thus far has dealt with Dewey and non-Dewey items separately rather than in the aggregate.

Patterns were examined by creating statistical profiles of variations relationships between call number and age and by constructing graphic representations of the relationships. The mean, median and maximum age values for groupings by Dewey 100s for each library and for all libraries as a group are shown in Table 2.

As can be seen from Table 2, there are fairly consistent but not especially great variations in the mean and median values. The mean age for all Dewey items ranges from 4.875 years for Library 2 to 14.596 years for Library 5.¹⁰ The median age ranges from four to thirteen years. The mean age for all libraries ranges from 7.056 years for the Dewey 800s to 10.745 for the Dewey 400s; the median ranges from six to eight years. The patterns for individual libraries vary somewhat, but overall the patterns for different libraries and different Dewey classes are remarkably similar. Although the differences in means among the libraries and among the Dewey classes are statistically significant (oneway analysis of variance reveals a probability level of .001 in both cases), it is hard to attach much importance to differences that are numerically rather small. The differences among the mean ages are represented graphically in Figures 1 through 12. These graphs reiterate the consistency of the patterns among libraries and among Dewey classes.

TABLE 3: MEAN, MEDIAN AND MAXIMUM VALUES
DEWEY 100s BY LIBRARIES

DEWEY	LIB 1	LIB 2	LIB 3	LIB 4	LIB 5	LIB 6	LIB 7	ALL LIBS
000s								
MEAN	4.724	5.229	5.967	7.981	14.000	7.581	6.278	7.702
MEDIAN	4.000	5.000	6.000	7.000	13.000	7.000	6.000	7.000
MAX	18.000	35.000	90.000	33.000	93.000	49.000	25.000	93.000
100s								
MEAN	5.171	5.686	6.730	11.068	14.631	8.628	6.947	8.579
MEDIAN	4.000	4.000	5.000	9.000	13.000	6.000	6.000	7.000
MAX	36.000	52.000	52.000	45.000	80.000	40.000	53.000	80.000
200s								
MEAN	6.503	5.633	7.344	12.531	15.345	9.023	6.726	9.123
MEDIAN	4.000	5.000	6.000	9.000	13.000	7.000	6.000	7.000
MAX	107.000	52.000	52.000	53.000	89.000	52.000	41.000	107.000
300s								
MEAN	5.025	4.671	5.556	9.226	14.050	7.039	6.105	7.409
MEDIAN	4.000	4.000	5.000	8.000	13.000	5.000	5.000	6.000
MAX	134.000	89.000	90.000	118.000	81.000	81.000	53.000	134.000
400s								
MEAN	6.793	7.416	9.240	14.831	16.731	13.101	8.960	10.745
MEDIAN	5.000	6.000	7.000	12.000	14.000	9.000	7.000	8.000
MAX	64.000	35.000	90.000	42.000	84.000	92.000	103.000	103.000
500s								
MEAN	5.432	5.925	6.760	10.669	15.402	8.855	6.777	8.410
MEDIAN	4.000	5.000	6.000	8.000	14.000	7.000	6.000	7.000
MAX	36.000	35.000	54.000	64.000	72.000	54.000	52.000	72.000
600s								
MEAN	5.672	4.972	6.011	9.148	14.539	8.135	6.758	8.014
MEDIAN	5.000	4.000	5.000	8.000	13.000	7.000	6.000	7.000
MAX	84.000	90.000	90.000	95.000	68.000	60.000	41.000	95.000
700s								
MEAN	6.211	5.516	6.768	12.668	14.792	8.659	9.451	9.247
MEDIAN	5.000	4.000	6.000	9.000	14.000	7.000	6.000	8.000
MAX	90.000	90.000	90.000	86.000	94.000	84.000	95.000	95.000
800s								
MEAN	4.959	3.988	5.836	6.475	13.918	6.370	6.326	7.056
MEDIAN	4.000	4.000	5.000	5.000	13.000	5.000	5.000	6.000
MAX	92.000	83.000	90.000	90.000	89.000	90.000	94.000	94.000
900s								
MEAN	6.749	5.163	6.224	13.489	15.856	10.838	6.381	8.987
MEDIAN	5.000	4.000	5.000	9.000	14.000	7.000	6.000	7.000
MAX	118.000	109.000	103.000	128.000	140.000	139.000	60.000	140.000
ALL								
MEAN	5.495	4.875	6.145	9.096	14.596	7.998	6.813	8.012
MEDIAN	4.000	4.000	5.000	7.000	13.000	6.000	6.000	6.000
MAX	134.000	109.000	103.000	128.000	140.000	139.000	103.000	140.000

Although the patterns of mean, median and maximum values are of interest, more meaningful information is gained by graphing the entire range of values for some combination of call number range, library and age of item. It is conventional to construct graphs in a cumulative manner. Previous obsolescence studies have tended to reveal the pattern shown in Figure A.¹¹ This curve shows a very rapid rise representing a clustering of relatively young items, falling into a declining rate of cumulation representing the older portion of the collection. Bourne used this common bibliometric distribution to posit the "ninety percent library" as a decision making model for library practice.¹² In age analysis, the selection of some retention percentage can be used to define target items that will be candidates for some kind of weeding action. Those items below the targeted percentage level can be assumed to be contributing members of the collection and will not be considered candidates for weeding, while those items above the targeted percentage level will be examined in some way to determine whether action should be taken to either remove them from the collection or enhance their likelihood of use. Figure A projects lines for the eightieth and ninetieth percentiles. In practice, the desired retention rate would be decided at the local level for an individual library, based on such considerations as availability of shelf space, the focus of collection management efforts, budget allocations, circulation characteristics of the collection or detailed examination of the aging patterns of specific groups of materials in the collection.

One of the striking characteristics of the body of data used for this study is the consistency of patterns. If the entire database is examined, the pattern is very much in keeping with that of Figure A. This pattern is also seen when the collections of individual libraries are considered, as shown in Figure 109. Overall, the pattern shows a rapidly rising line that begins to fall into a flatter line at about seventeen to eighteen years. The same basic patterns prevail for individual Dewey call number ranges, particularly when grouped across all libraries. When call number groups are graphed for individual libraries, the patterns are

somewhat less smooth, but still fall into the general curve of Figure A. The curve tends to be somewhat ragged for the Dewey 000s, 100s, 300s, and 500s but much smoother for the 300s, 600s, 700s, 800s and 900s.

The aging patterns for Dewey ranges (by hundreds) tend to show Library 5 as the oldest, but the anomaly noted above clearly influences this. Library 4 is consistently second oldest, followed by Library 6. Library 7 and Library 3 trade positions for fourth and fifth oldest, depending on call number range. Libraries 1 and 2 trade places for sixth and seventh oldest, with Library 2 most consistently the "newest" collection. This suggests that the libraries' collections do have different characteristics and histories, despite the overall similarity in aging patterns.

The Dewey 400s were consistently the oldest class of materials, representing the greatest mean age for all libraries except Library 7. The Dewey 000s, 300s, and 800s were the youngest classes, although none of them was universally younger than the others.

Sample Application of Age Data to Collection Analysis

As was mentioned earlier, specific application of age analysis to collection management decision making must be based on local objectives and the availability of resources for conducting the analysis. As an example of how age data might be incorporated into selecting target items for decision making, two approaches to targeting will be described. The first makes use of an arbitrary retention percentage of ninety percent; that is, the oldest ten percent of the collection will be examined. The second approach makes use of the point at which the curve for each Dewey major class (Dewey hundreds) makes the transition from the dramatic rise to the flatter tail of the distribution. Library 1 has arbitrarily been chosen for these examples; only the Dewey portion of the collection is included in the examples.

Profile 1: The Ninety Percent Retention Rate

The first method for targeting items for examination consists of looking at the oldest ten percent of items either across the board or in various classes (that is, selecting a ninety percent retention rate). It seems likely that examining items within individual classes will be more useful than an across the board approach, although the results should ultimately be very similar. For Library 1, with a Dewey collection of 10,435 items, the oldest ten percent comprises 1,044 items (all percentages are rounded upward to provide for a slight oversampling). This would involve those items approximately ten years old or older. Table 4 shows the target samples that would be chosen if analysis were conducted at the level of the Dewey hundreds rather than across the board. Analysis at this level results in class age cutoffs from eight to twelve years, and a total target sample of 1,047 items, a miniscule oversampling in comparison to the across the board ten percent examination figure. Assuming that examining each item targeted would require an average of about twenty minutes of professional level staff time, conducting an analysis based on ninety percent retention for Library 1 would require 349 person/hours, or almost nine weeks of effort.

Profile 2: Curve Examination

Examining the transition points in the actual distributions produces substantially different results, also shown in Table 4. The graph of the collection taken as a whole appears to make the transition at about 14 years, with a target sample of 629 items, only six percent of the population of 10,435 items. In most cases, the transition for the ten Dewey classes approximates the point at which there are fewer than ten items of each age; there is an ele-

ment of arbitrariness in this approach that allows for refinement in local decision making. The most obvious result of this approach as applied for this example is that the total number of items to be examined is substantially less. When individual Dewey class distributions were examined, 572 items were targeted, representing 5.5 percent of the Dewey population for Library 1. This would require about 191 person hours, not quite five weeks, for analysis.

TABLE 4: COMPARISON OF METHODS FOR IDENTIFYING TARGET ITEMS

DEWEY RANGE	TOTAL HOLDINGS	90% SAMPLE	AGE CUTOFF	AGE CUTOFF FROM GRAPH	GRAPHIC SAMPLE
000s	308	31	8	11	17
100s	240	24	11	7	46
200s	169	17	10	12	14
300s	1827	183	8	15	62
400s	87	9	10	15	5
500s	507	51	10	12	41
600s	1670	167	10	16	64
700s	1254	126	12	17	67
800s	3212	322	8	17	143
900s	1151	117	12	13	113
TOTAL	10425	1047	10	14	572

Comparison of the Percentage and Curve Examination Methods

The major difference between the selection of a desired retention percentage and examination of the curve of the actual age distribution is the extent to which the latter method can be fine tuned to meet very specific decision making goals. If the objective is to examine as few items as possible while still maintaining a sound, systematic basis for targeting items, the curve examination method, carefully applied, is very attractive. This flexibility is useful for accommodating goals related to collection balance and growth, since individual sections of the collection can be handled separately. Although the retention percentage approach can also be tailored to specific sections of the collection, determining discrete retention percentages for subsections of the collection may be a difficult process and will inherently require an element of truly arbitrary decision making. At the same time, a disadvantage of the curve examination method is that it is substantially less mechanical than the retention percentage approach, and therefore requires substantially more intellectual effort on the part of the professional level employee charged with identifying the target sample. Determining an appropriate, meaningful and useful transition point requires developing a "feel" for the body of data being considered, and also requires being able to relate graphic data to the kind of tabular presentation produced by a spreadsheet.

Continuation of the Project

Like most truly interesting research, the Age Analysis of Public Library Collections project cannot be considered truly complete. It is only the beginning of a long-term research process. The project has turned out to be substantially more ambitious than the

researchers' had realized at the outset. The wealth of the data provided by OCLC is phenomenal, considering its seeming simplicity. Describing and understanding the aging patterns of the data collected have occupied most of the period for which funding was provided, and will continue to be the focus of research efforts for some time. Only general recommendations for library practice can be made at this time; more specific procedures and rules for expediting them will need to follow further analysis of the data. The development of programs for computerized analysis, which was initially viewed as being one product of the study, has yet to be accomplished. The project, although it has revealed much of interest about the age characteristics of public library collections, must be considered a work still in progress. The researchers plan to extend the research to include the following:

1. further, more detailed analysis of the OCLC data, including comprehensive analysis of both the Dewey and non-Dewey records;
2. development of specific recommendations and procedures for utilizing age information in making collection management decisions;
3. application of the age analysis process to an individual library in a case study mode of operation;
4. development of computer programs for age analysis.

It seems evident, based on the analysis conducted up to this time, that studying an individual library is a necessary next step and is a prerequisite to any definitive statement regarding the efficacy of detailed age analysis as a collection management tool. This case study should ideally involve application of a variety of weeding methods, including the more traditional approaches already in general use as well as the retention percentage and curve examination methods described in this report. The data analyzed for this study seem to emphasize the attractiveness of detailed age analysis as a collection management tool. The development of an easy to use computer analysis program capable of working with existing local databases should be fairly simple, assuming that the consistency of the aging

patterns observed for the seven libraries studied is typical. There is no reason to believe that the data examined are not typical, at least of libraries in the size range of those studied. Although the objectives of the project have not been fully achieved, then, they certainly appear to be achievable, given further time, thought and effort. It is the fervent intent of the research team to apply such time, thought and effort and build from the foundation that has been laid to a more comprehensive understanding of the role of age analysis in evaluating library collections. A major benefit of the funded stage of the Age Analysis of Public Library Collections project has been the establishment of a solid base around which further research can be shaped.

Dissemination of Project Results

The work reported here has been disseminated in a number of public forums. Reports of the work in progress have been delivered to the Louisiana State Library Administrative Librarians' Conference (Baton Rouge, December 1990), the Louisiana Library Association Annual Conference (Monroe, March 1990), the Association for Library and Information Science Education Annual Conference (Chicago, January 1990), and the American Library Association Annual Conference (Dallas, June 1989). The project formed the basis for two all-day workshops on "Effective Collection Weeding" given for the Northeast Texas Library System (McKinney and Quitman, October 1990). A talk on the project, its findings and their implications for library practice will be given at the March, 1991 National Conference of the Public Library Association. An article describing the focus of the work appeared in *Public Libraries*, the official journal of the Public Library Association.¹³ It is anticipated that several other articles and presentations will be produced, including at least one article for a major library and information science research journal. Responses to the work have been frequent and positive. Many librarians

have expressed an interest in exploring how age analysis could effectively be used in their individual libraries, and several have volunteered to have their libraries serve as test sites for further study. The research team anticipates selecting one or more of these volunteer libraries as a case study test site for the next stage of the project.

Notes

1. Maurice B. Line and A. Sandison, "Obsolescence' and Changes in the Use of Literature With Time," *Journal of Documentation* 30 (September 1974): 283.
2. Joseph P. Segal, *Evaluating and Weeding Collections in Small and Medium-Sized Public Libraries: The CREW Method* (Chicago: American Library Association, 1980): 4.
3. F. Wilfrid Lancaster, "Obsolescence, Weeding, and the Utilization of Space," *Wilson Library Bulletin* 62 (May 1988): 47.
4. Loriene Roy, "Does Weeding Increase Circulation? A Review of the Related Literature," *Collection Management* 10 (1988): 141 - 156.
5. Stanley J. Slote, *Weeding Library Collections -- II*, Second Revised Edition (CO: 1982) 36 - 38.
6. A discussion of the use of circulation and in-house use analysis can be found in Stanley J. Slote, *Weeding Library Collections: Library Weeding Methods*, 3rd Edition (Englewood, CO: Libraries Unlimited, 1989).
7. A. W. McClellan, "New Concepts of Service," *Library Association Record* 58 (1956): 299 - 305.
8. Wallace John Bonk and Rose Mary Magrill, *Building Library Collections*, Fifth Edition (Metuchen, NJ: Scarecrow Press, 1979): 317.
9. This process of revising the sampling process through interaction with staff at OCLC unfortunately delayed the project by nearly three months.
10. Library 5 represents an unexplained anomaly in that the minimum age of the items recorded for Library 5 is four years, as opposed to zero years for the six other libraries. This may represent a library that truly has not added items to its collection since 1986, but seems more likely to represent a failure to add holdings information to the OCLC database since that year. This effect of this aberration on statistics for the database as a whole has not been compensated for in this report; as a result, such statistics should be interpreted cautiously.
11. See, for instance, R. E. Burton and R. W. Kebler, "The 'Half-Life' of Some Scientific and Technical Literatures," *American Documentation* 11 (January 1960): 18 - 19.
12. Charles P. Bourne, "Some User Requirements Stated Quantitatively in Terms of the 90 Percent Library," in Allen Kent and Orrin E. Taulbee, eds., *Electronic Information Handling* (Washington: Spartan Books, 1965): 93 - 110.
13. Danny P. Wallace, "The Young and the Ageless: Obsolescence in Public Library Collections," *Public Libraries* 29 (March/April 1990): 102 - 105.

Figure A
The Typical Obsolescence Curve

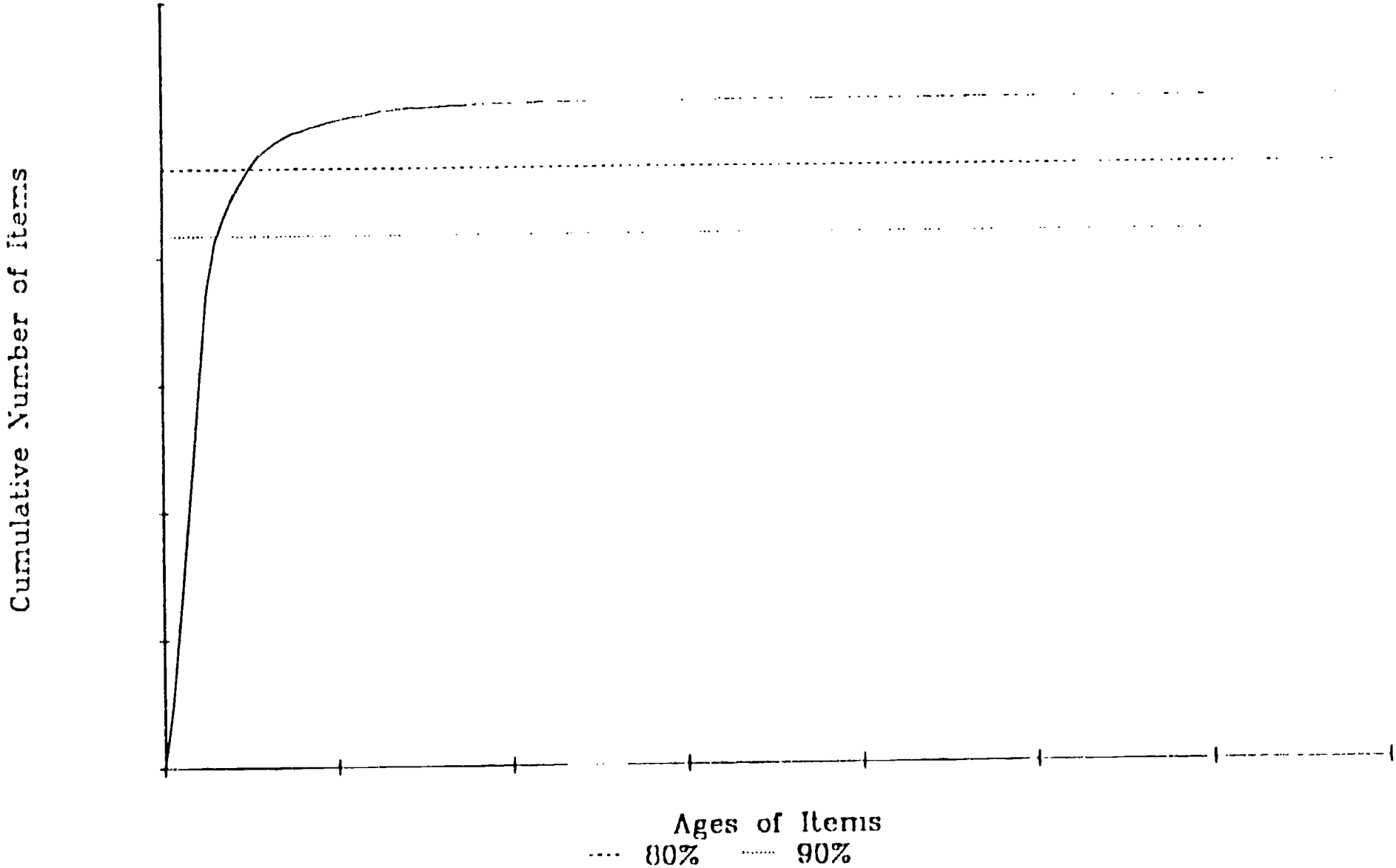


Figure 1

Mean Ages All Libraries

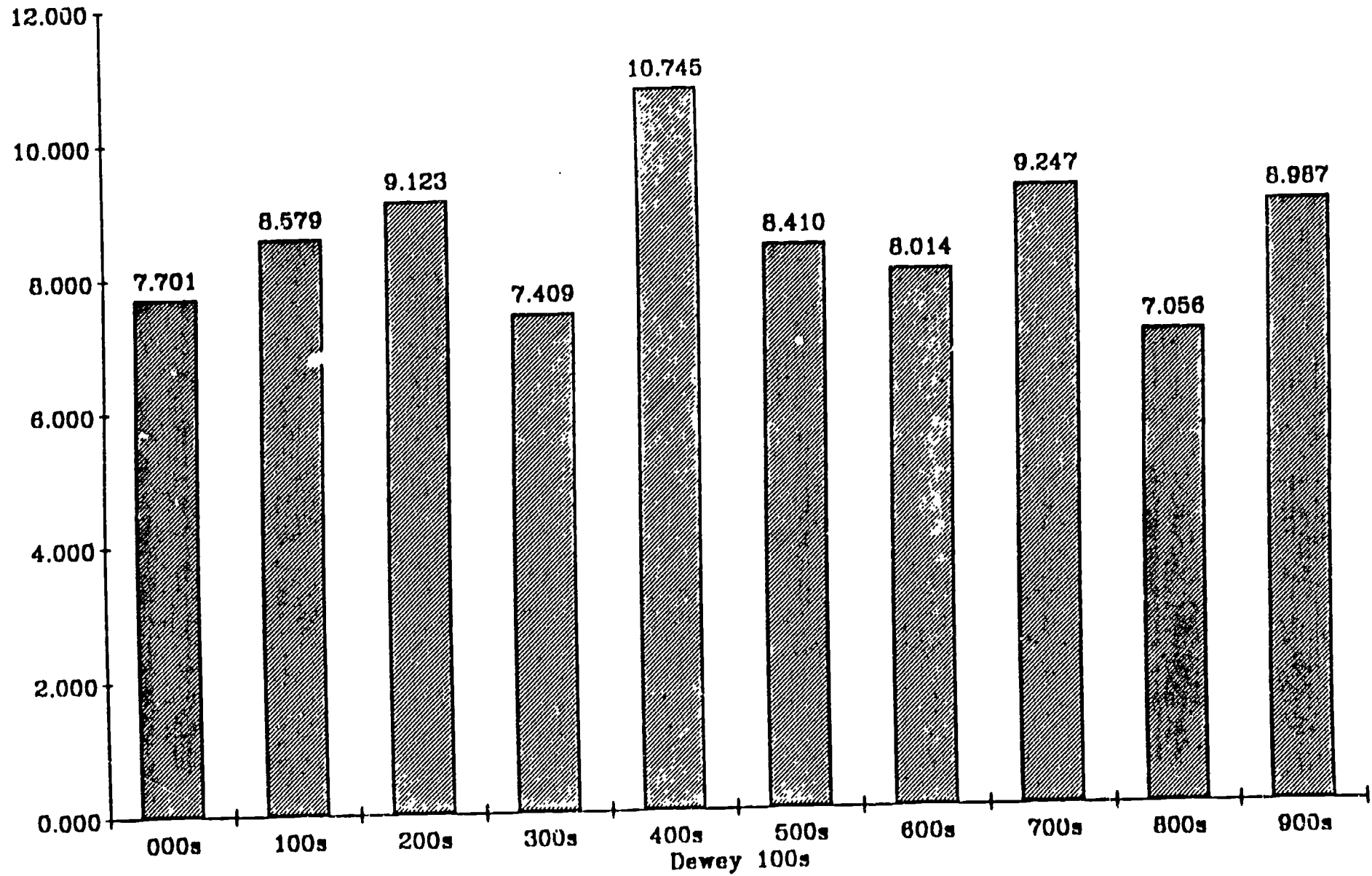
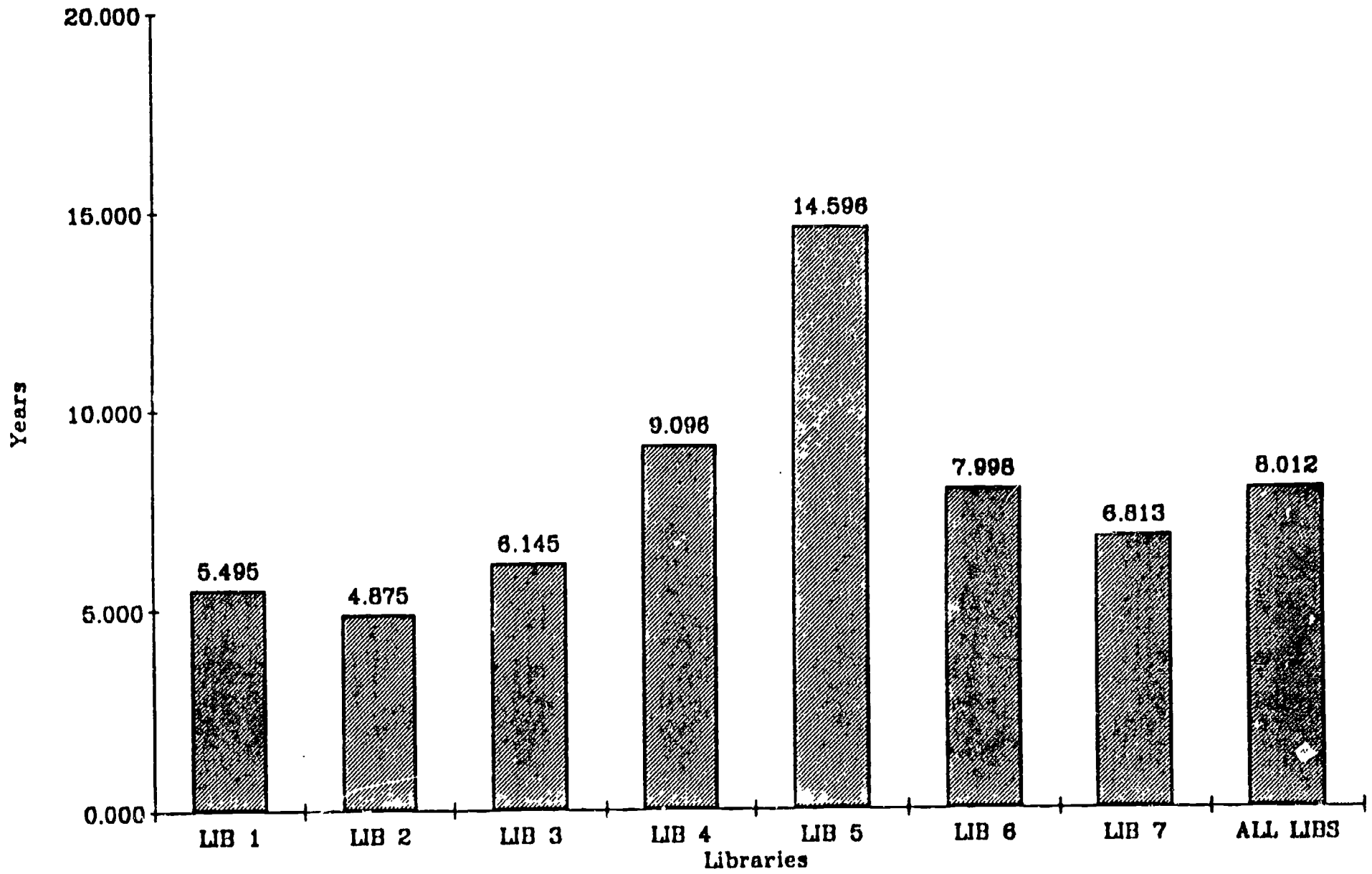


Figure 2

Mean Ages All Deweys



31 A

Figure 3

Mean Ages Dewey 000s

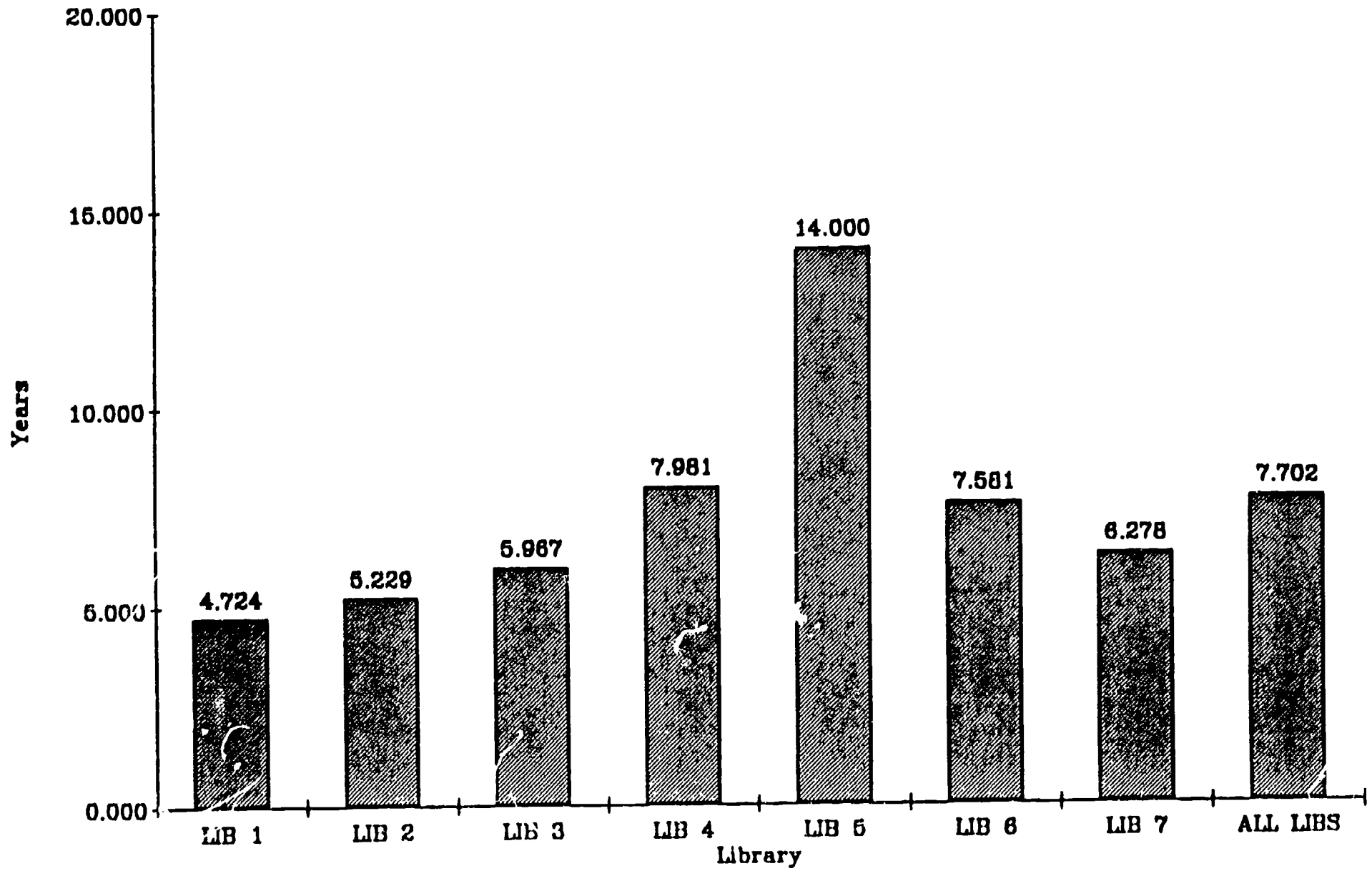


Figure 4

Mean Ages Dewey 100s

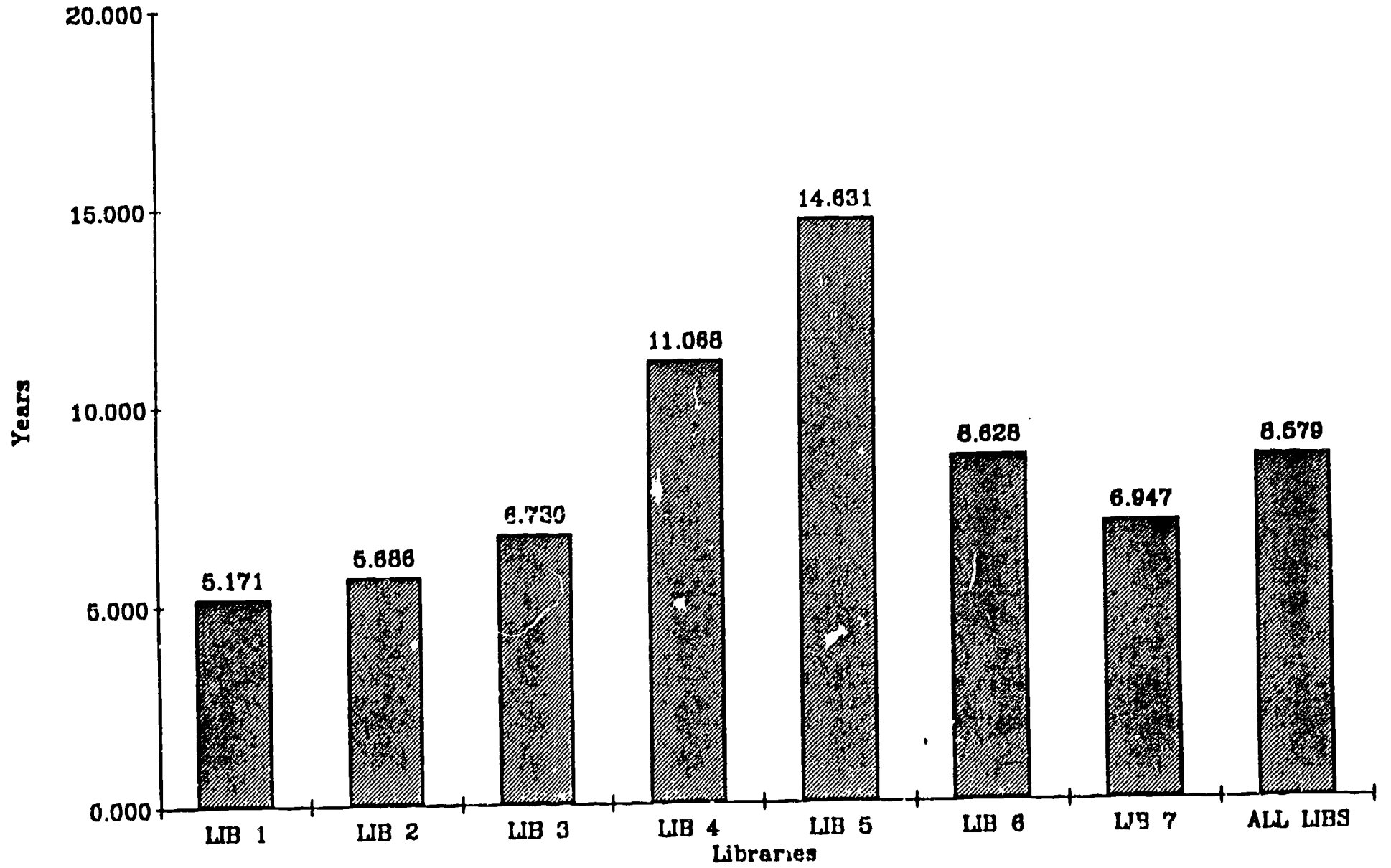
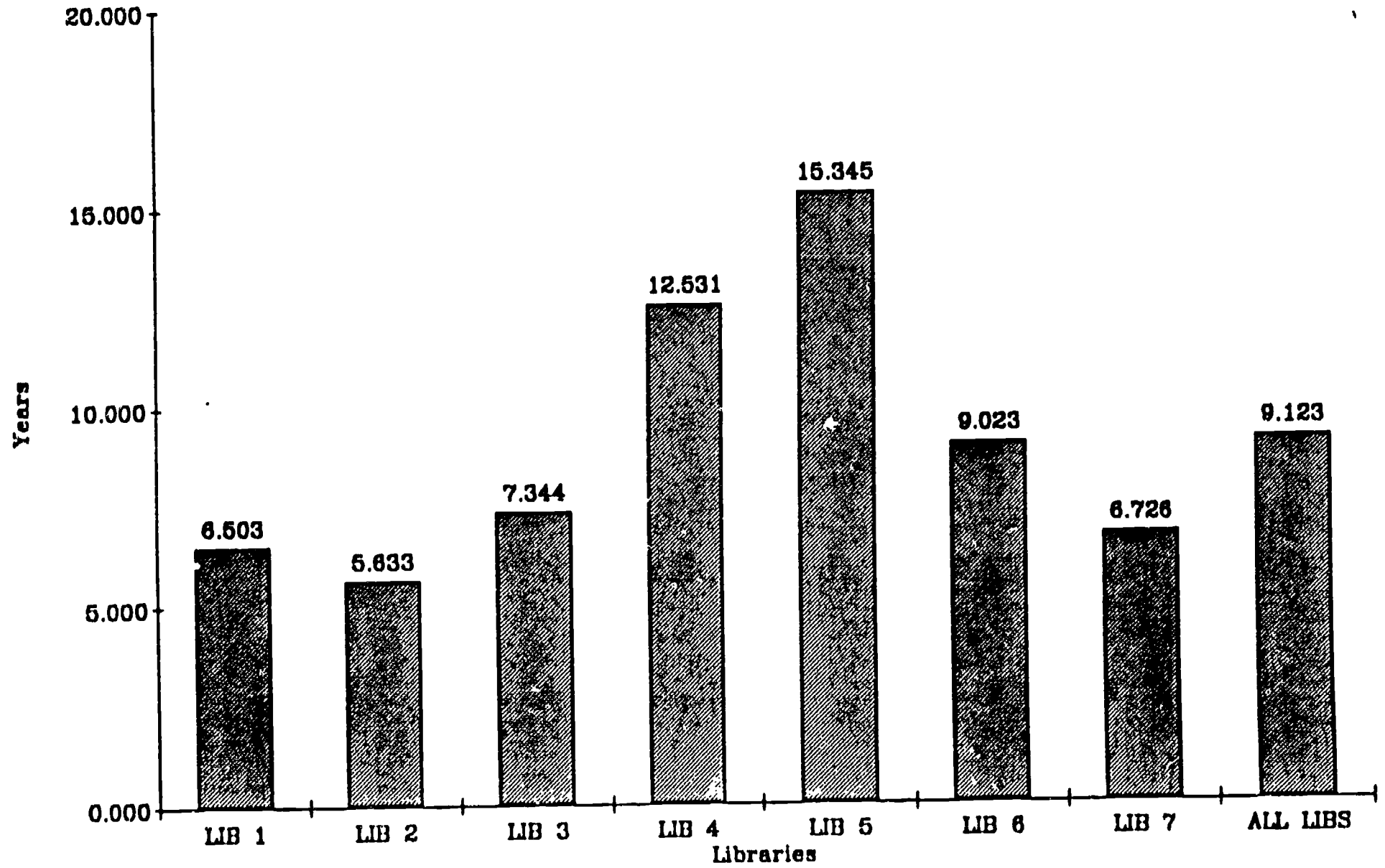


Figure 5

Mean Ages
Dewey 200s



34 A

Figure 6

Mean Ages Dewey 300s

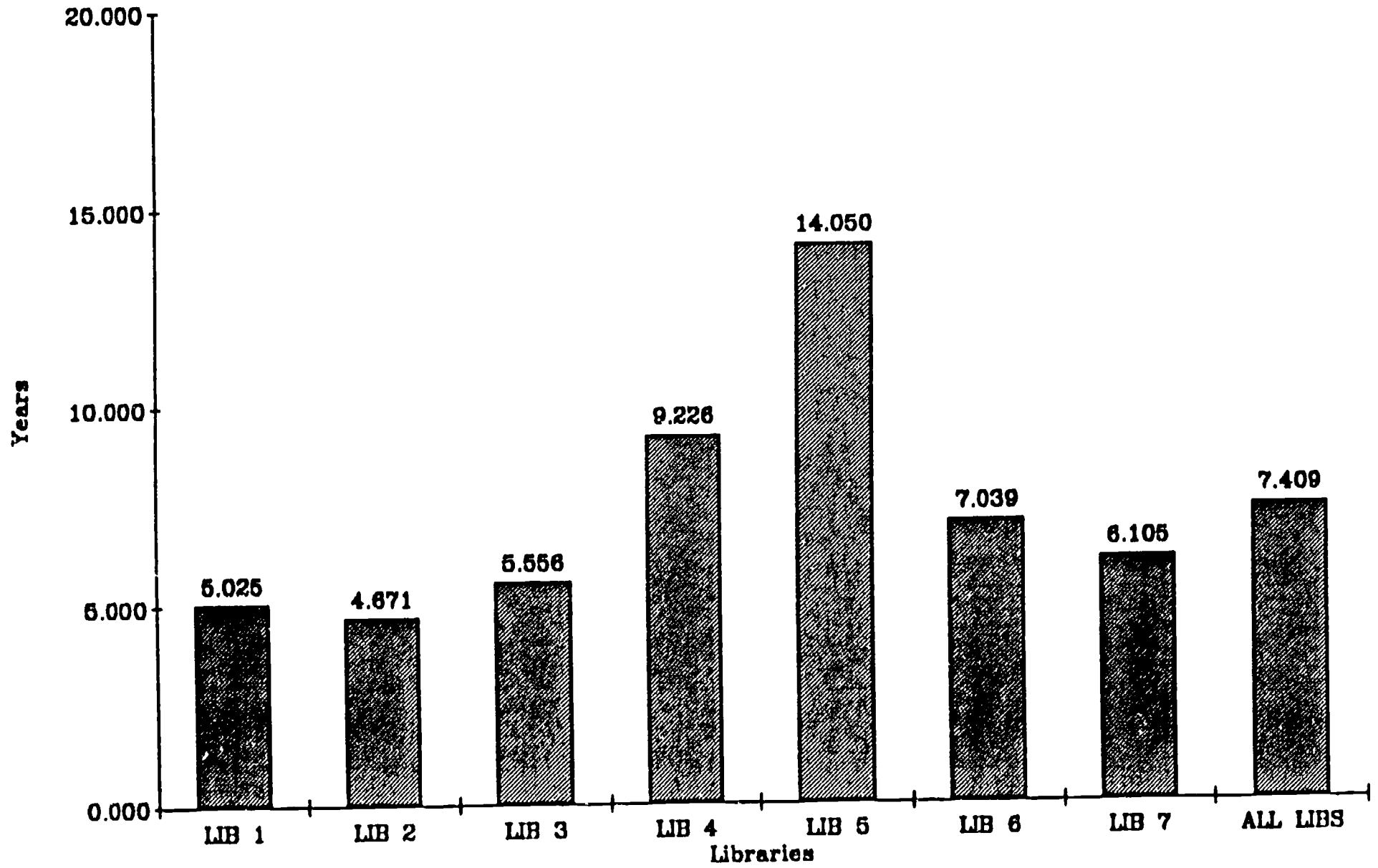


Figure 7

Mean Ages Dewey 400s

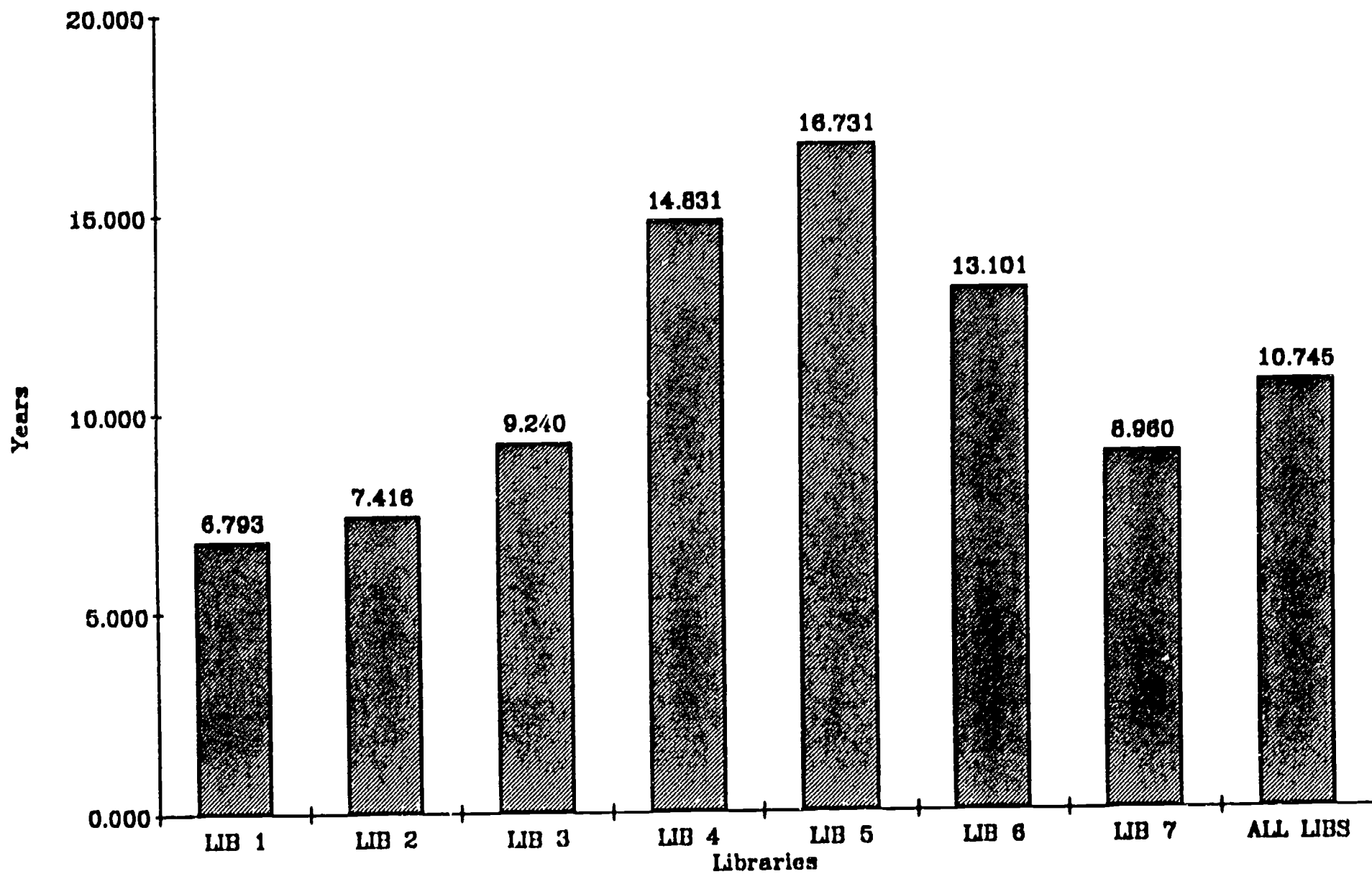
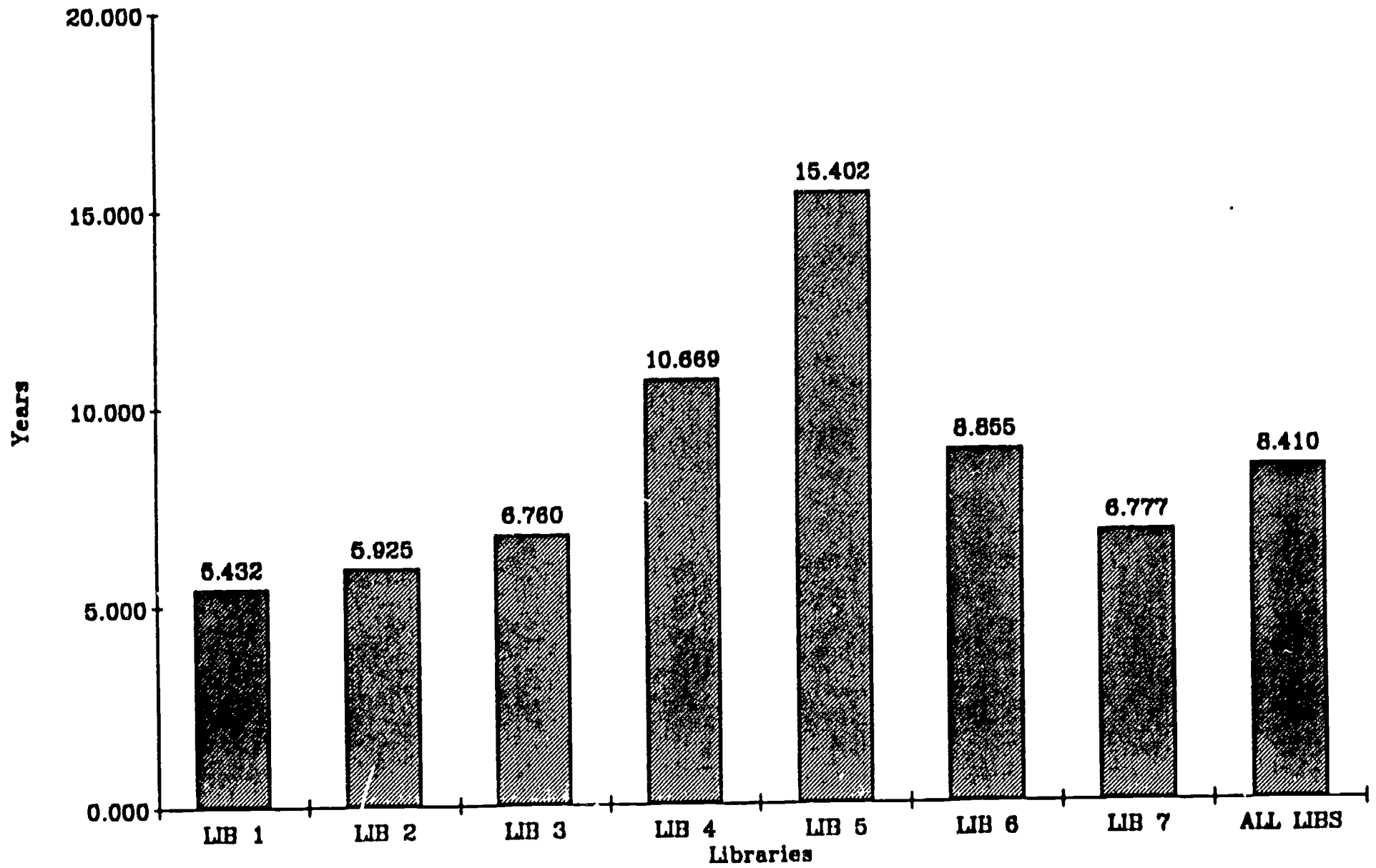


Figure 8

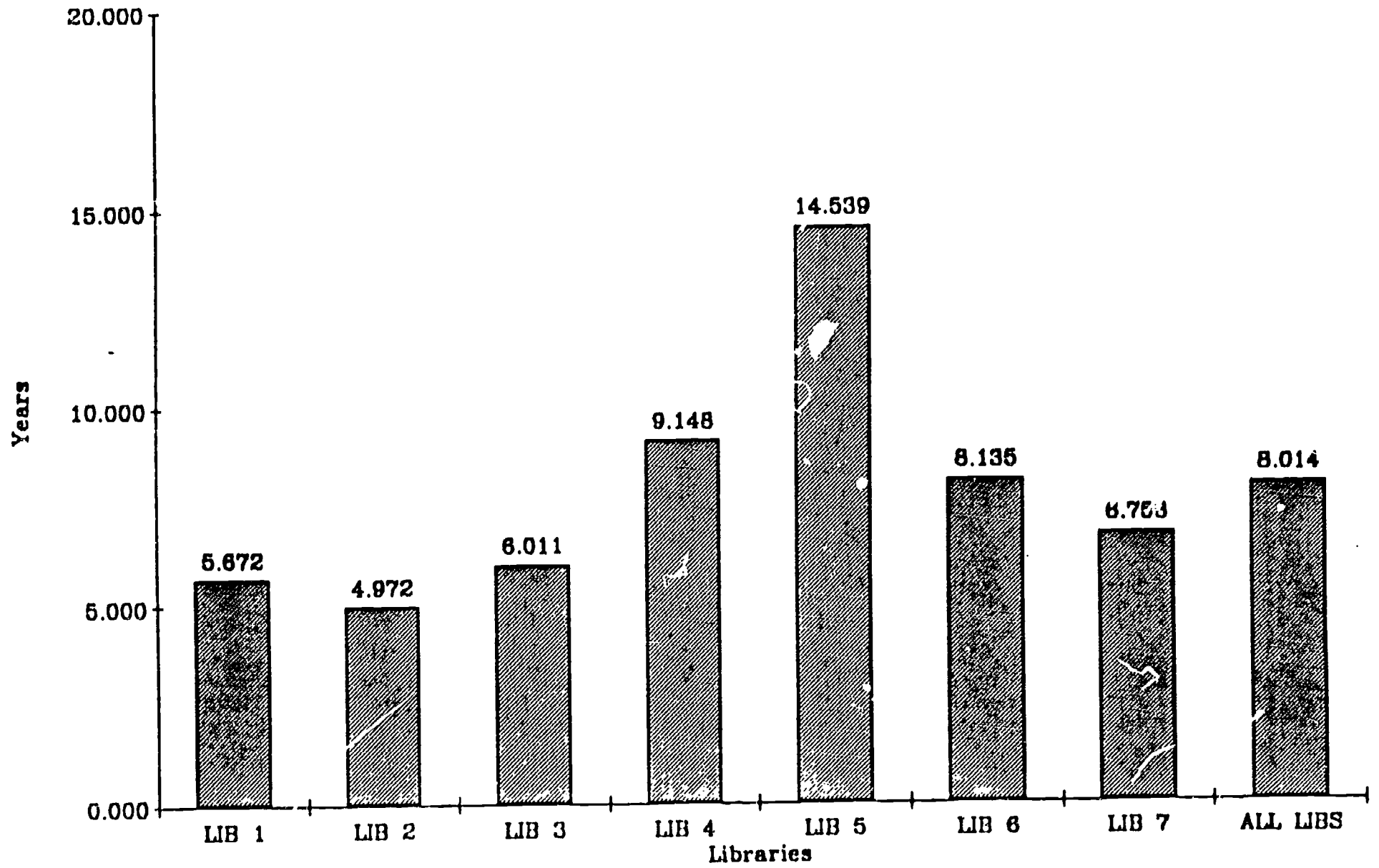
Mean Ages
Dewey 500s



37 A

Figure 9

Mean Ages Dewey 600s



38 A

Figure 41

DEWEY 000s

LIBRARY 2

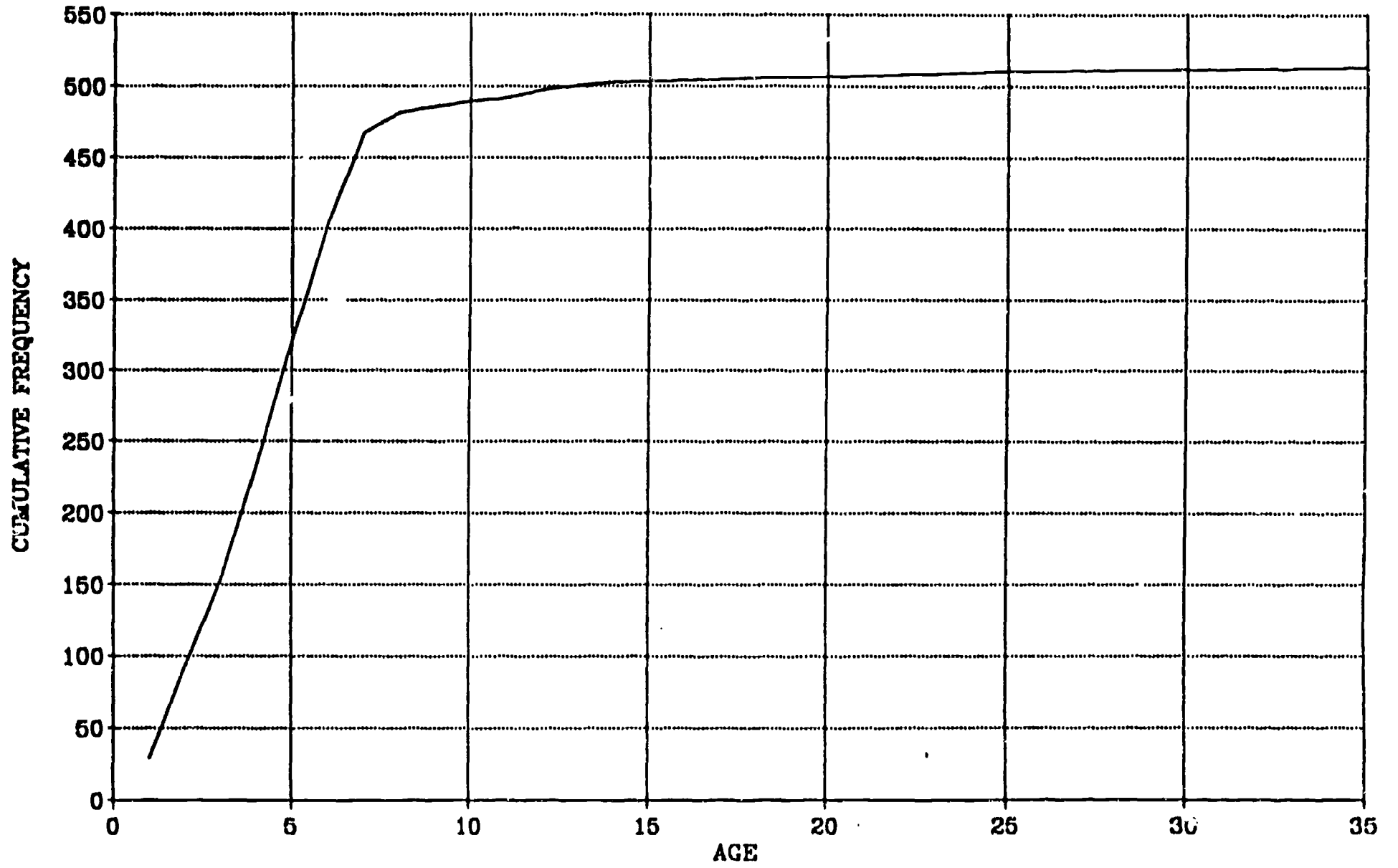
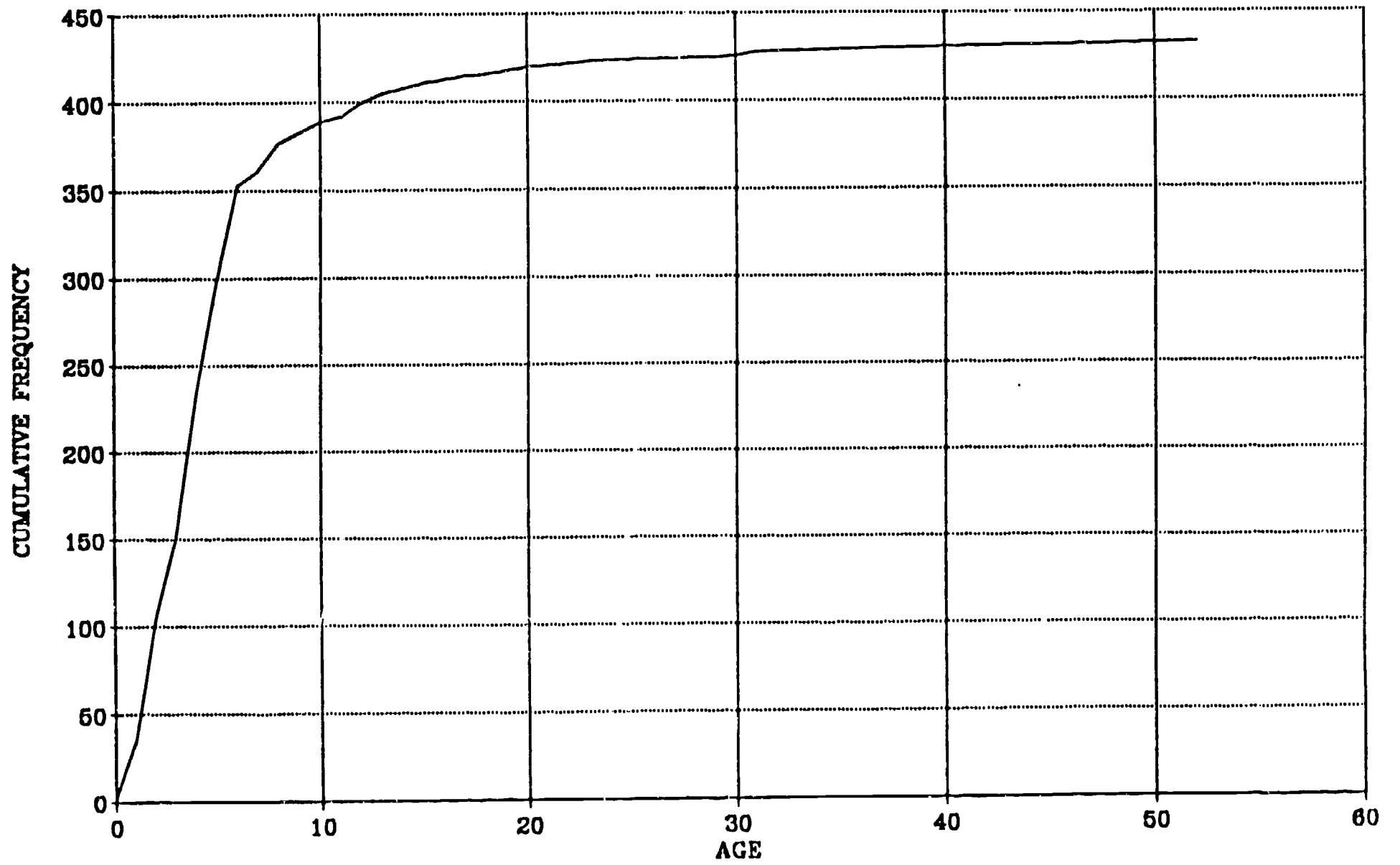


Figure 42

DEWEY 100s

LIBRARY 2



40

40A

Figure 43

DEWEY 200s
LIBRARY 2

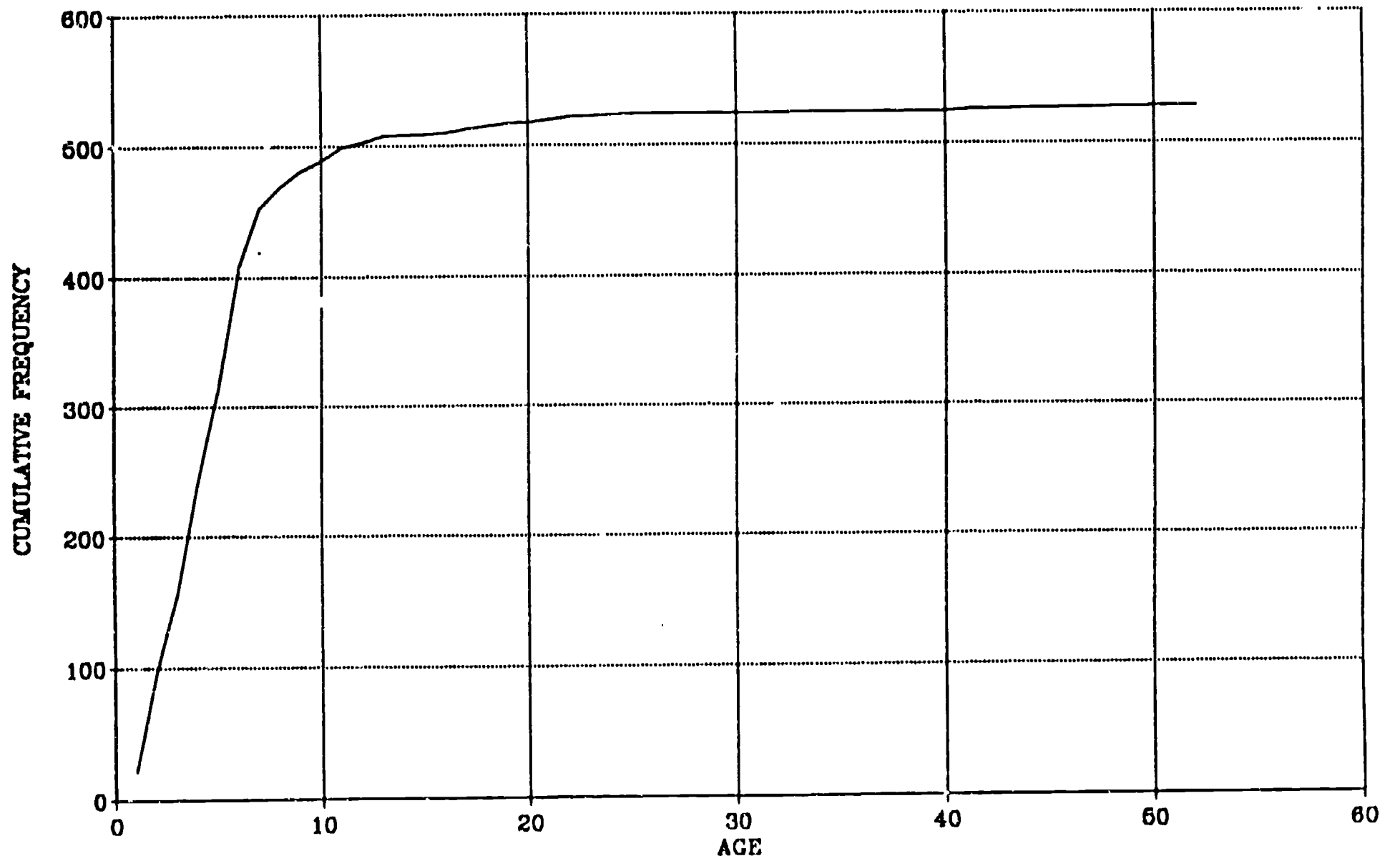


Figure 44

DEWEY 300s

LIBRARY 2

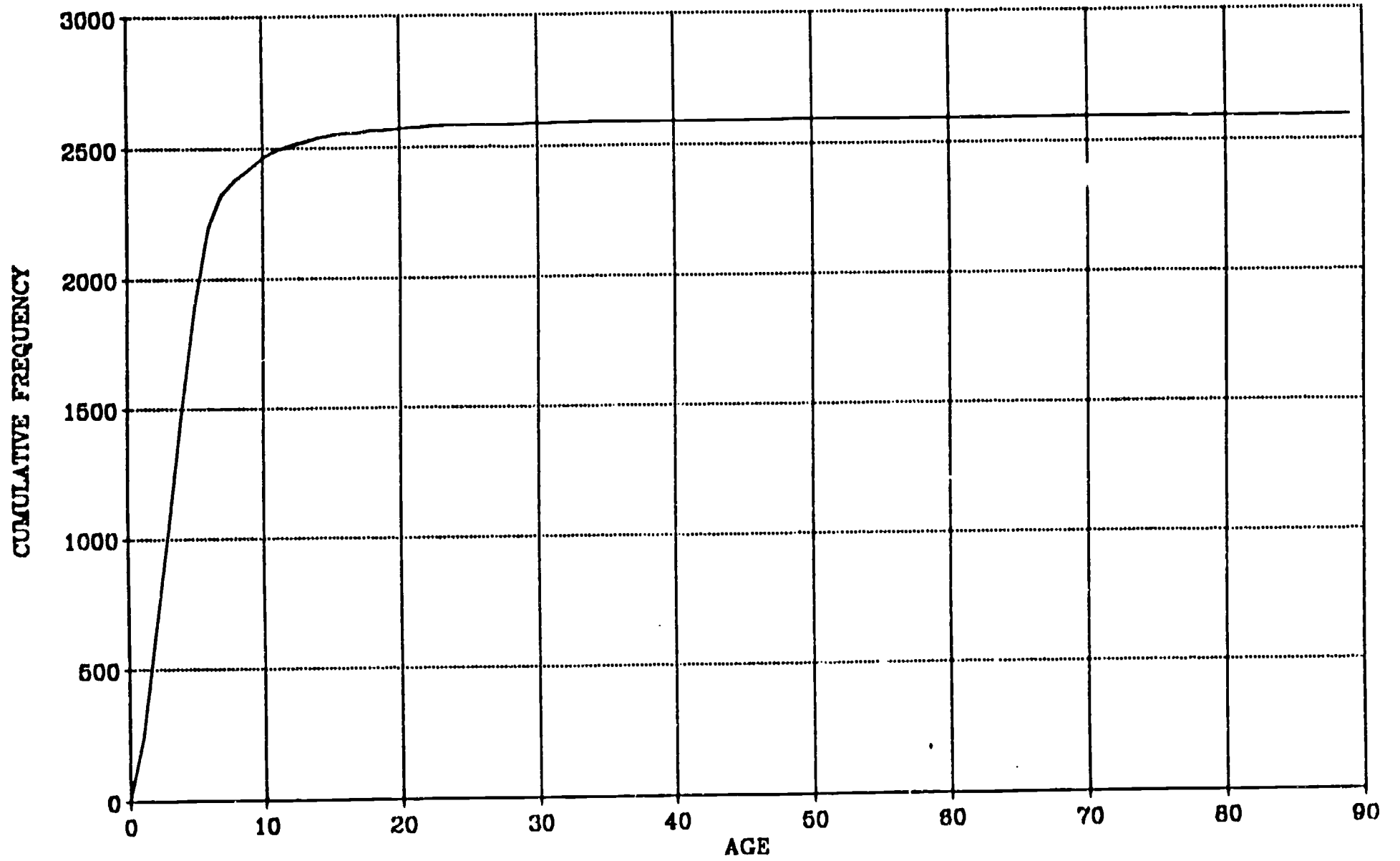


Figure 45

DEWEY 400s

LIBRARY 2

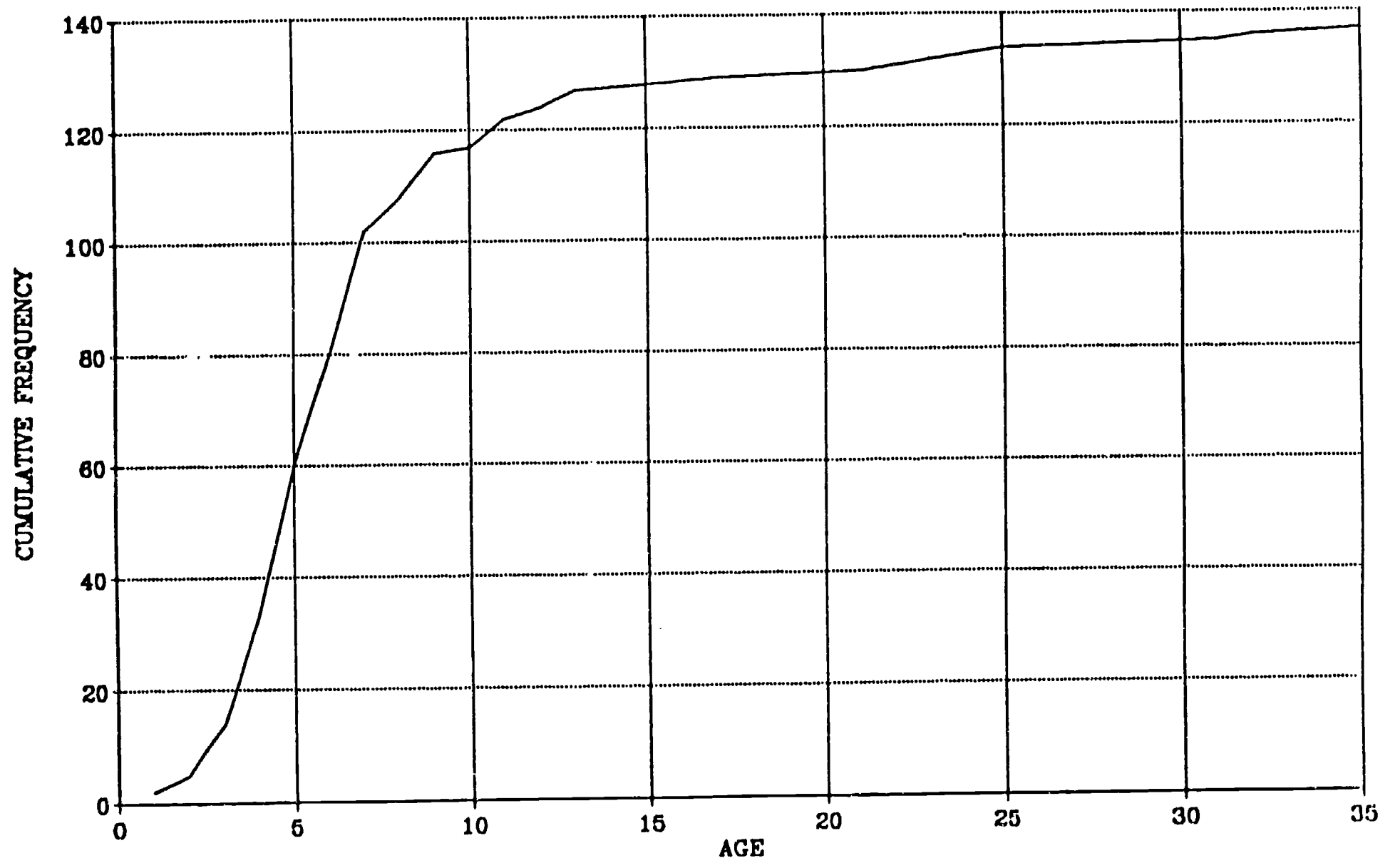


Figure 46

DEWEY 500s

LIBRARY 2

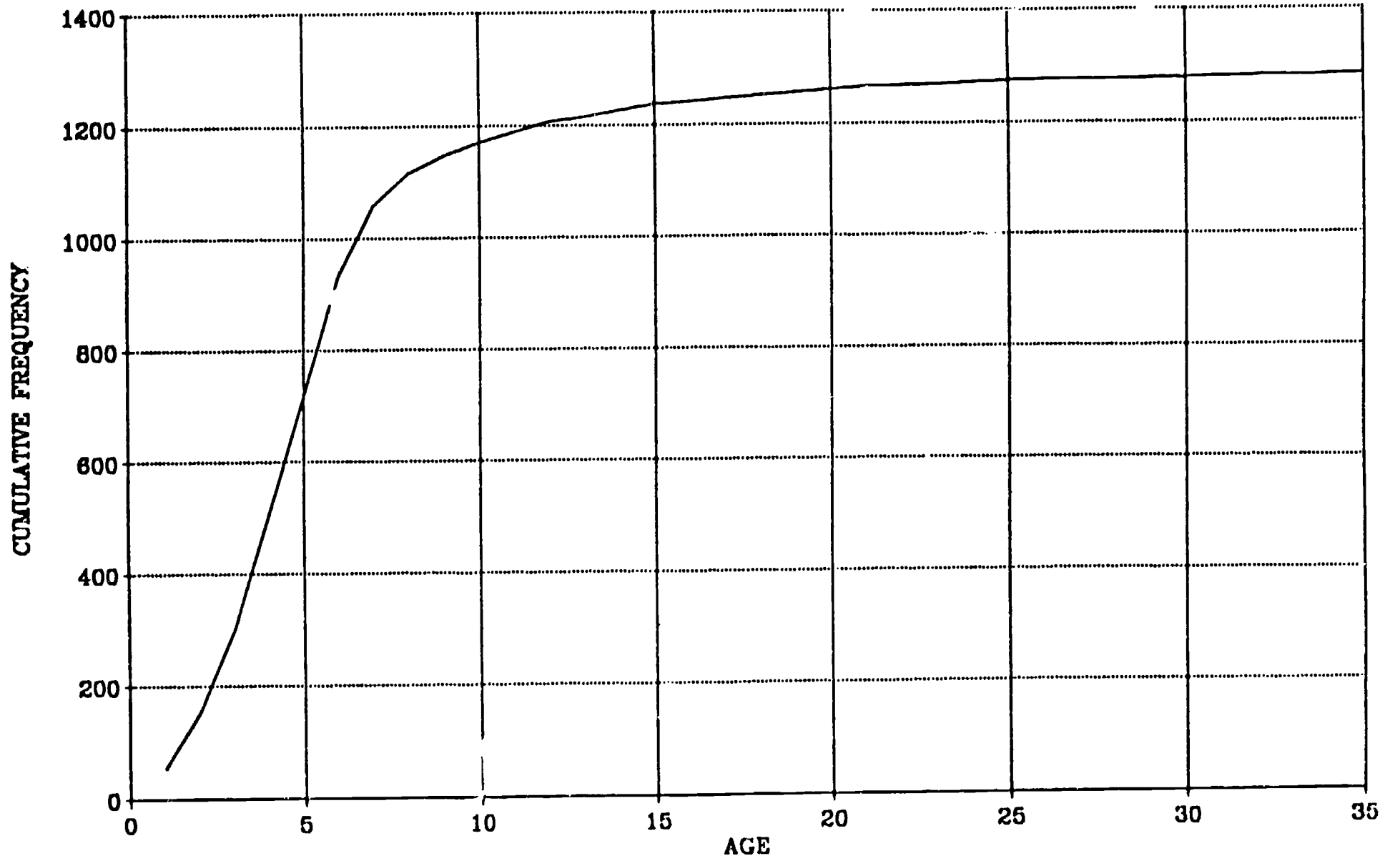


Figure 47

DEWEY 600s

LIBRARY 2

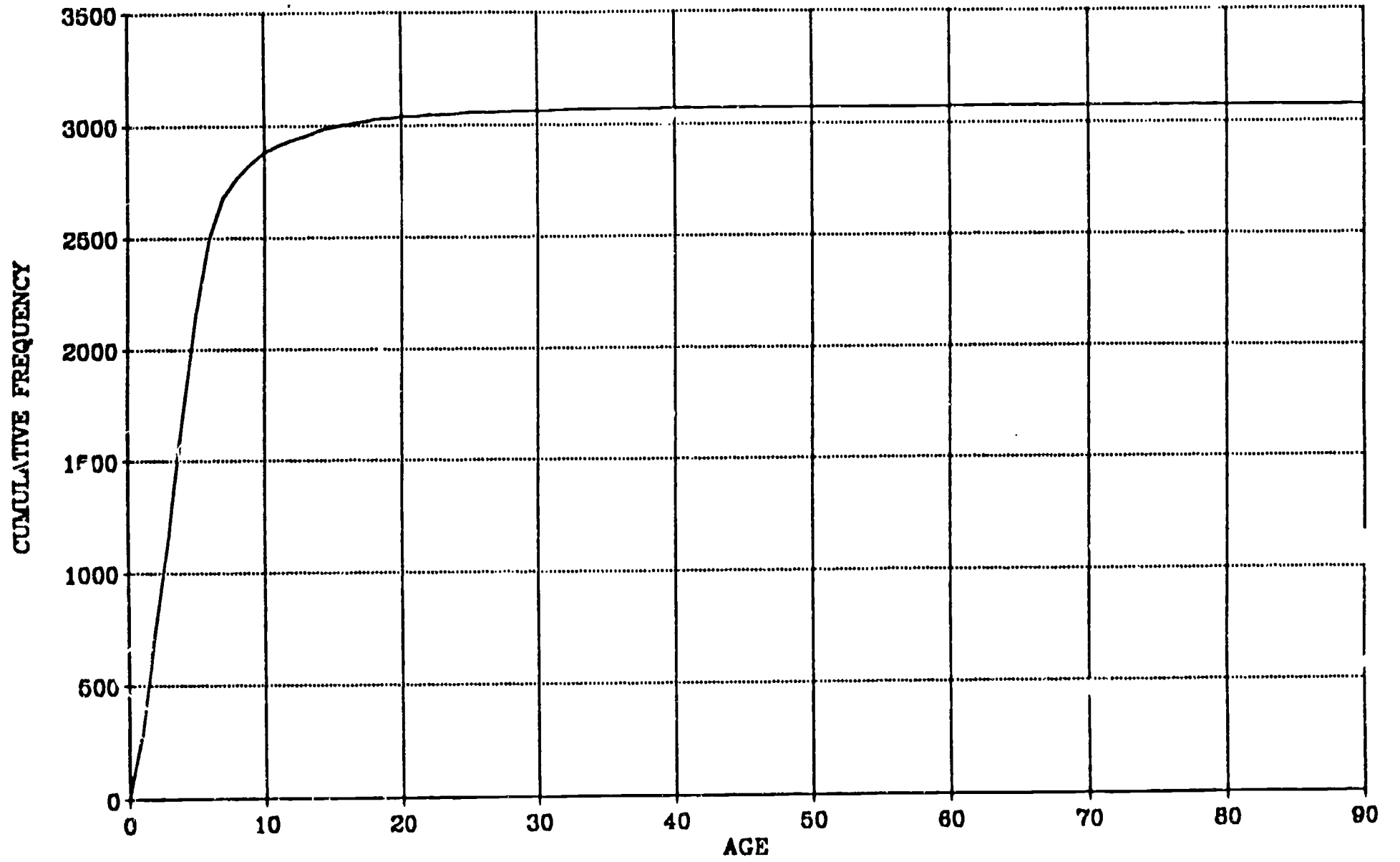


Figure 48

DEWEY 700s

LIBRARY 2

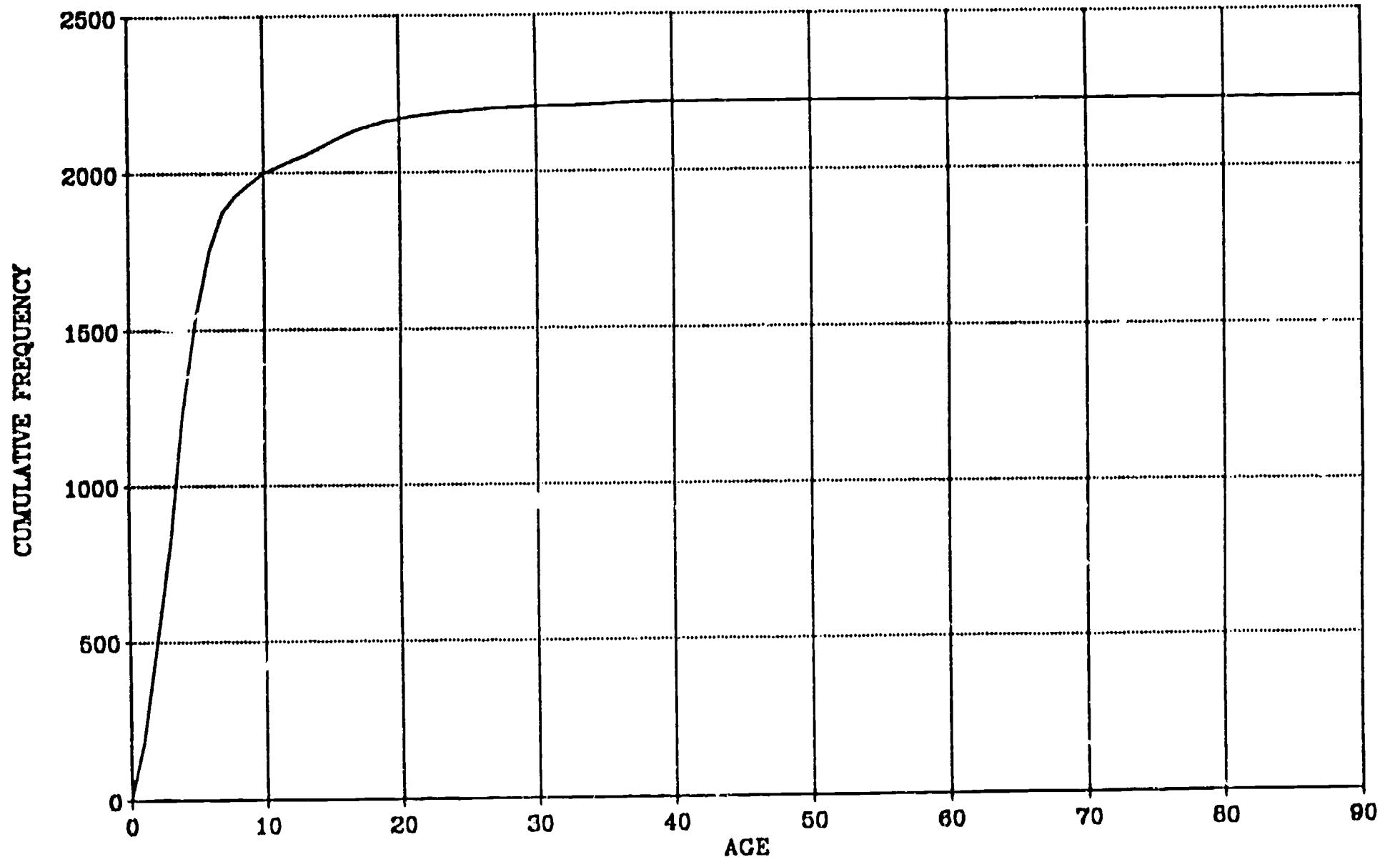
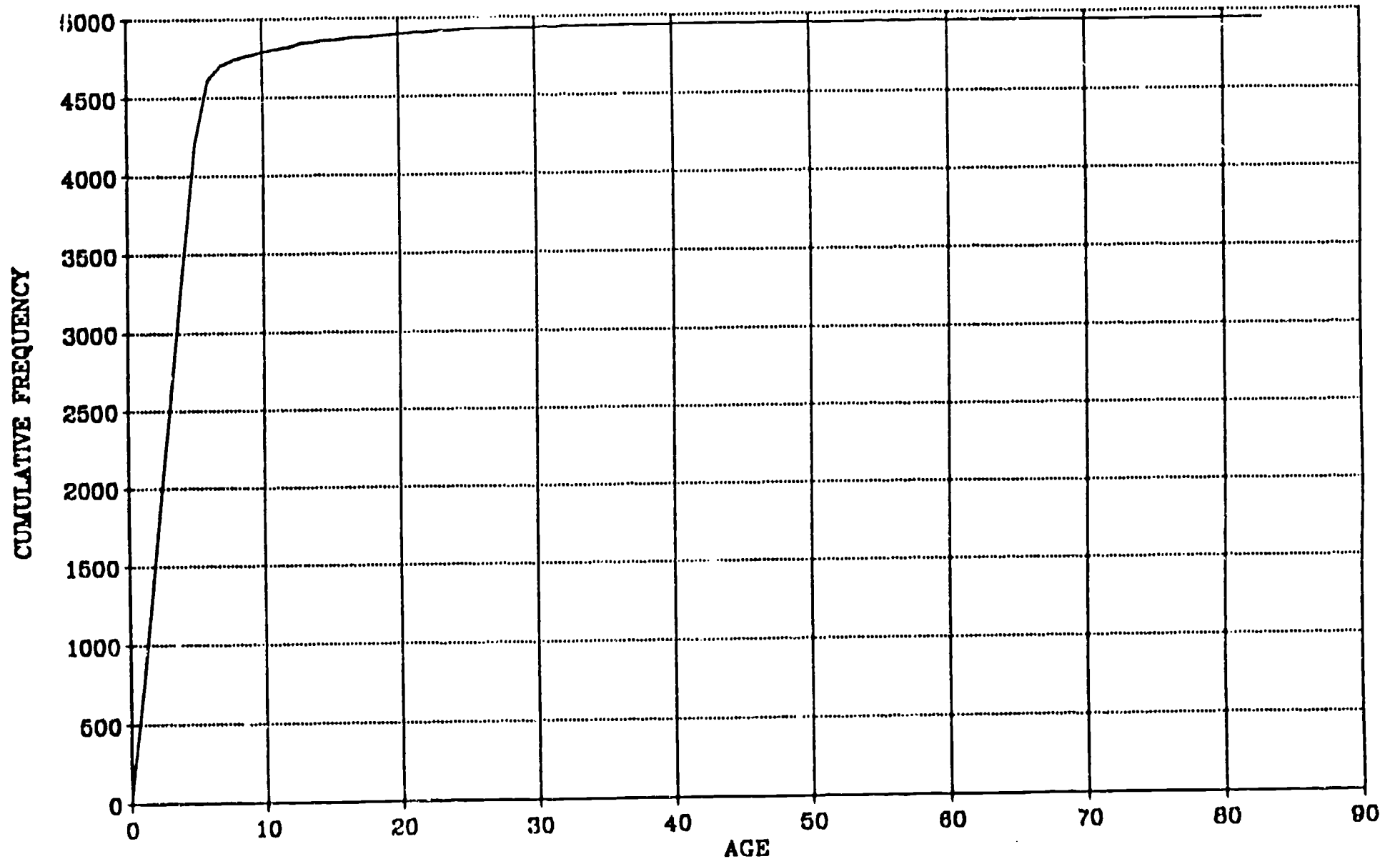


Figure 49

DEWEY 800s

LIBRARY 2

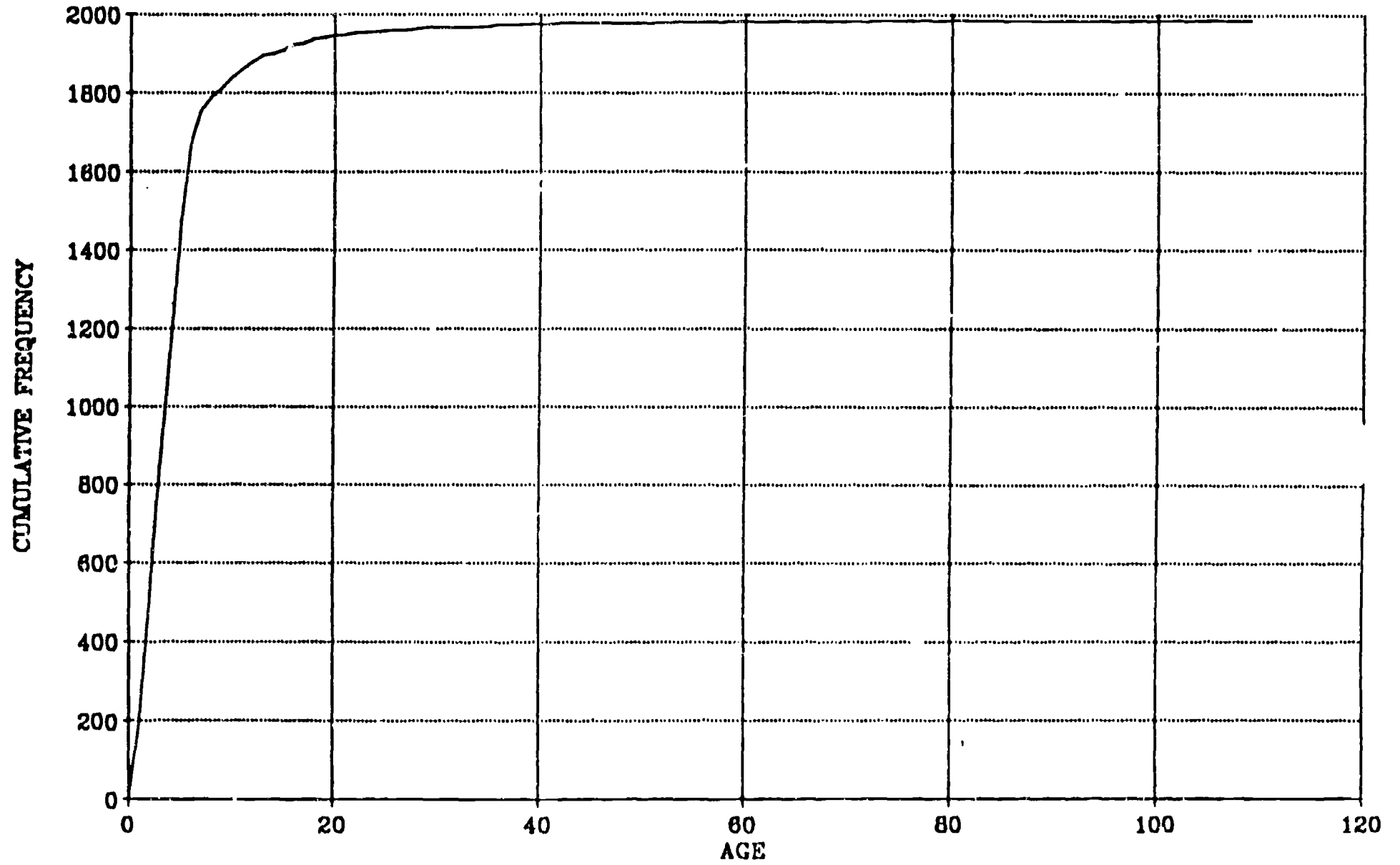


47A

Figure 50

DEWEY 900s

LIBRARY 2

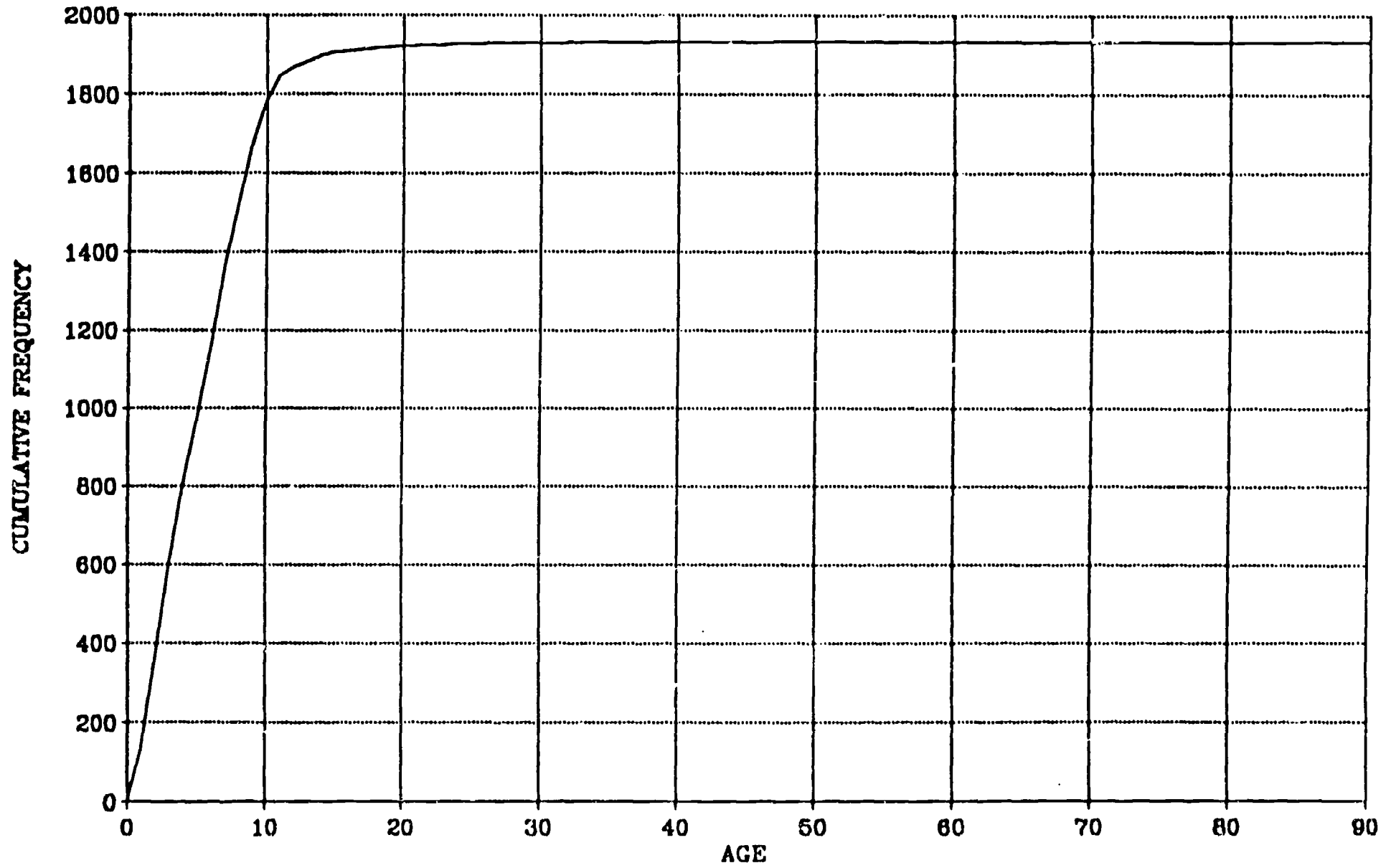


48A

Figure 51

DEWEY 000s

LIBRARY 3

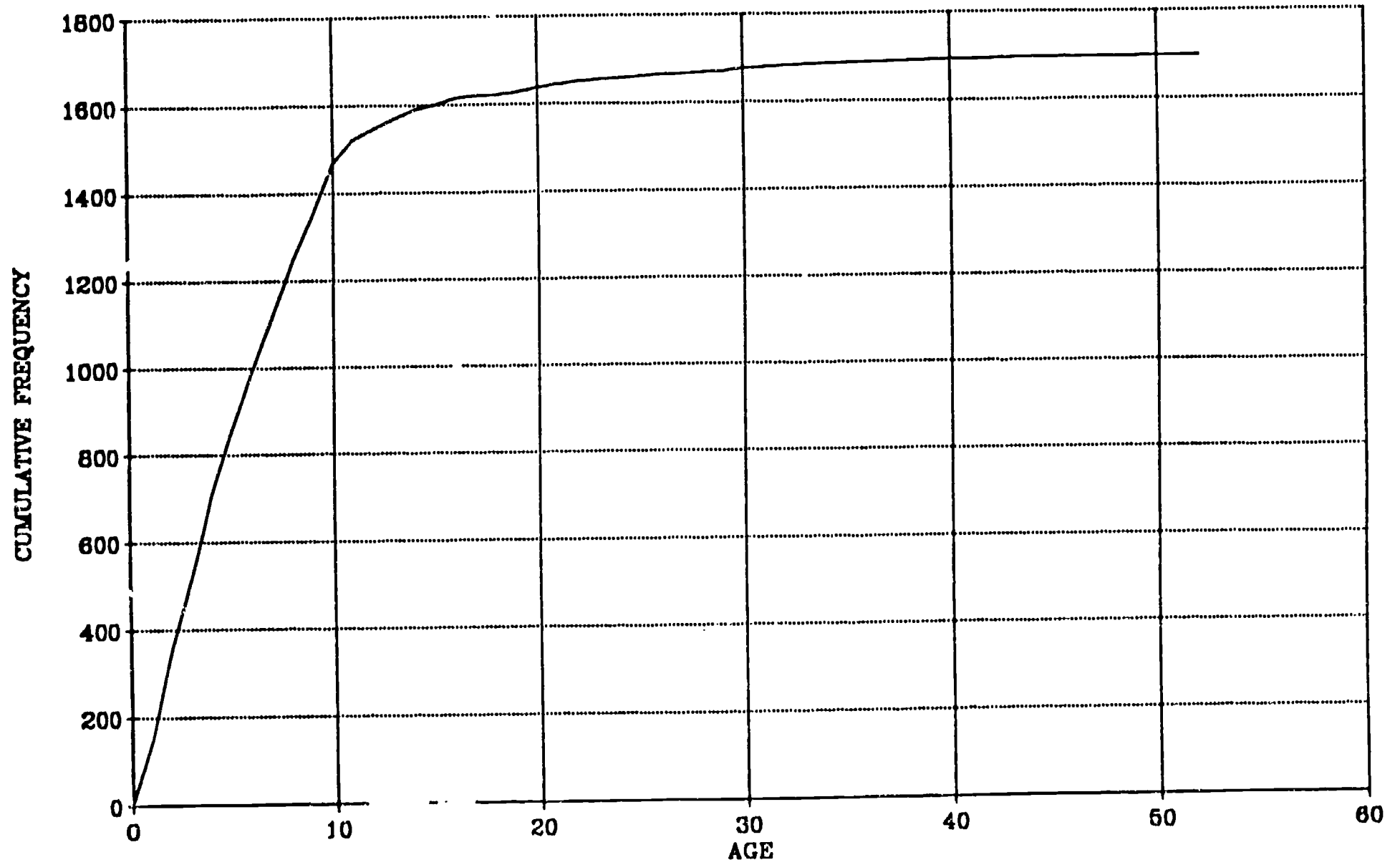


49 A

Figure 52

DEWEY 100s

LIBRARY 3



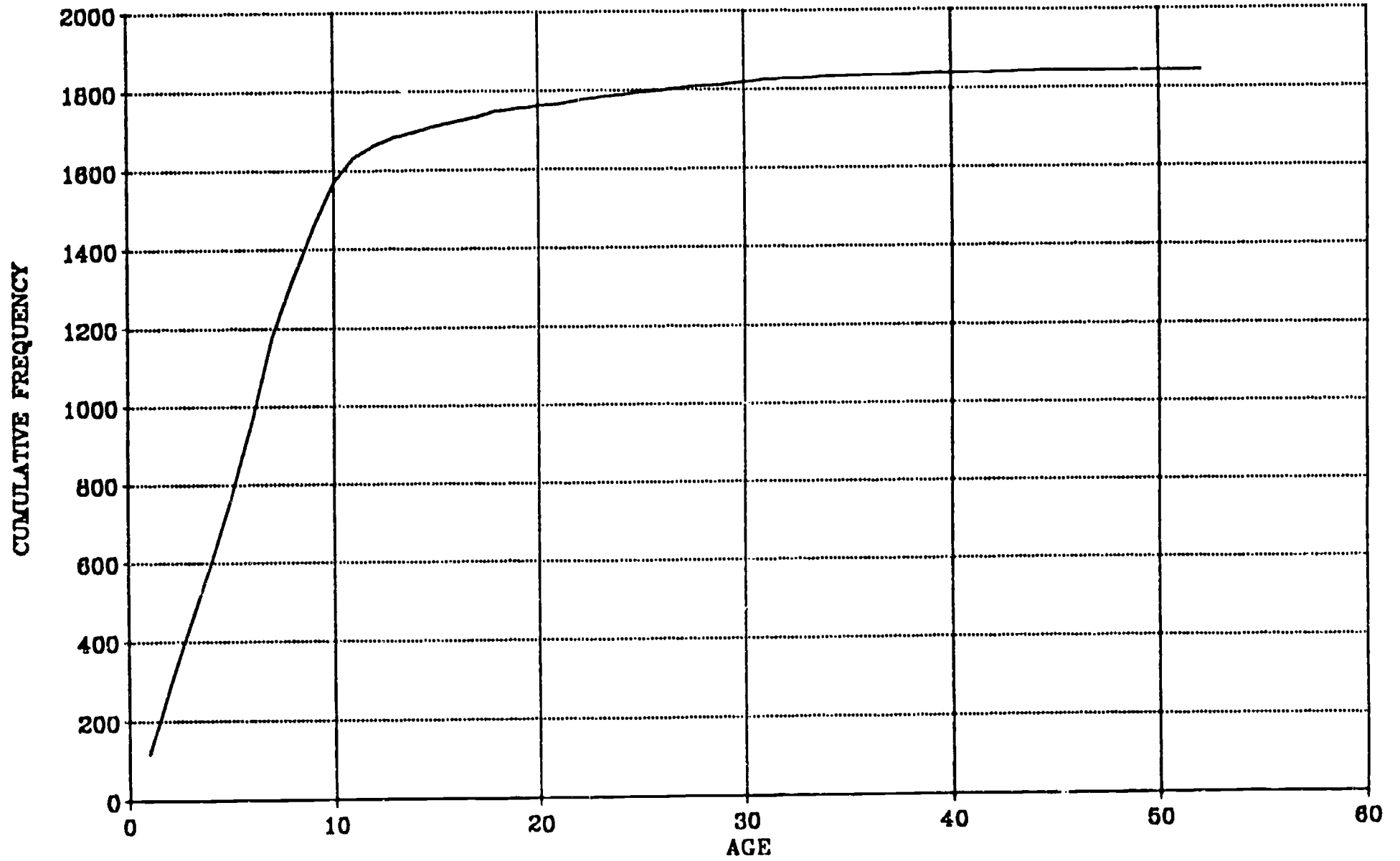
50

50A

Figure 53

DEWEY 200s

LIBRARY 3

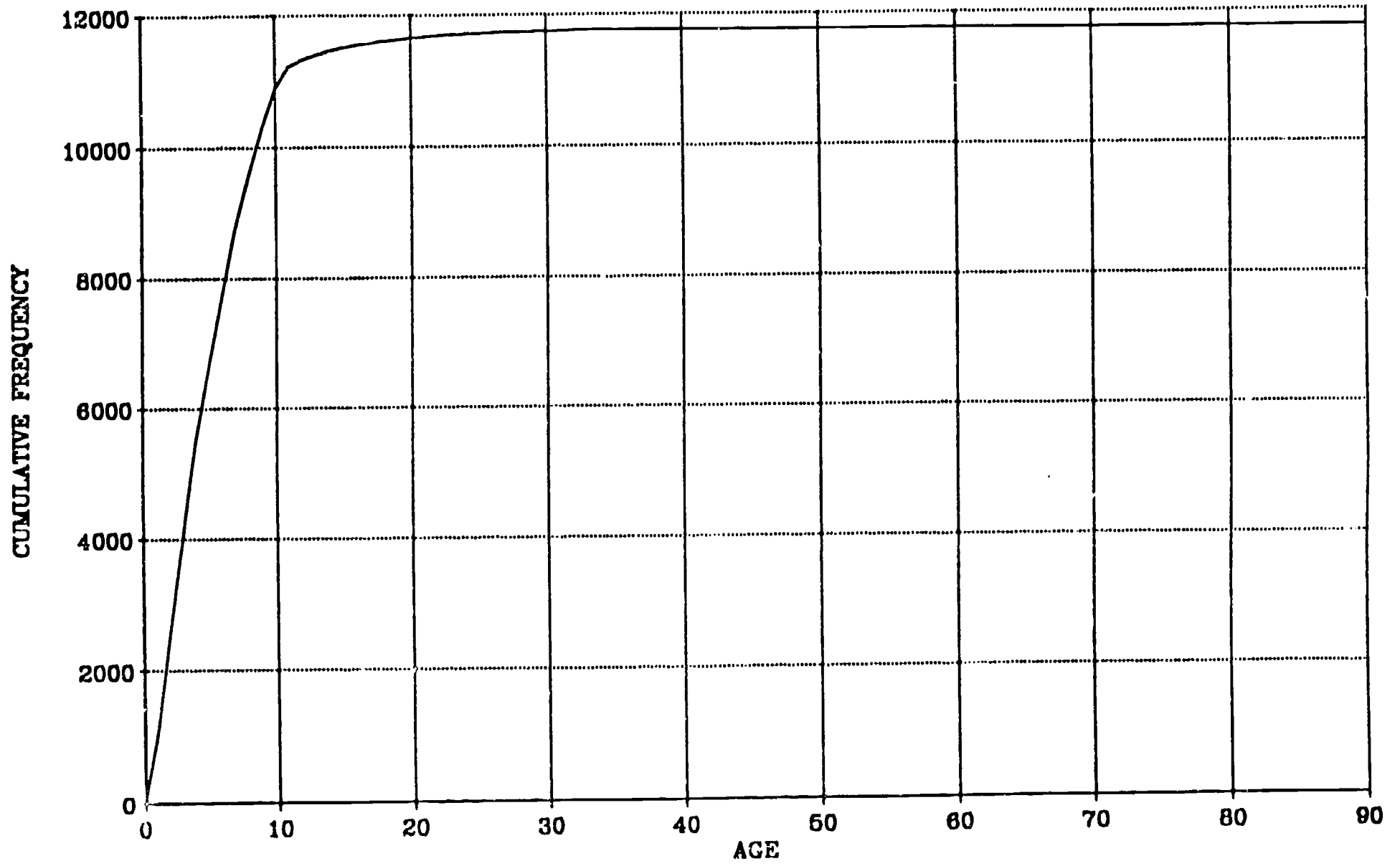


51 A

Figure 54

DEWEY 300s

LIBRARY 3

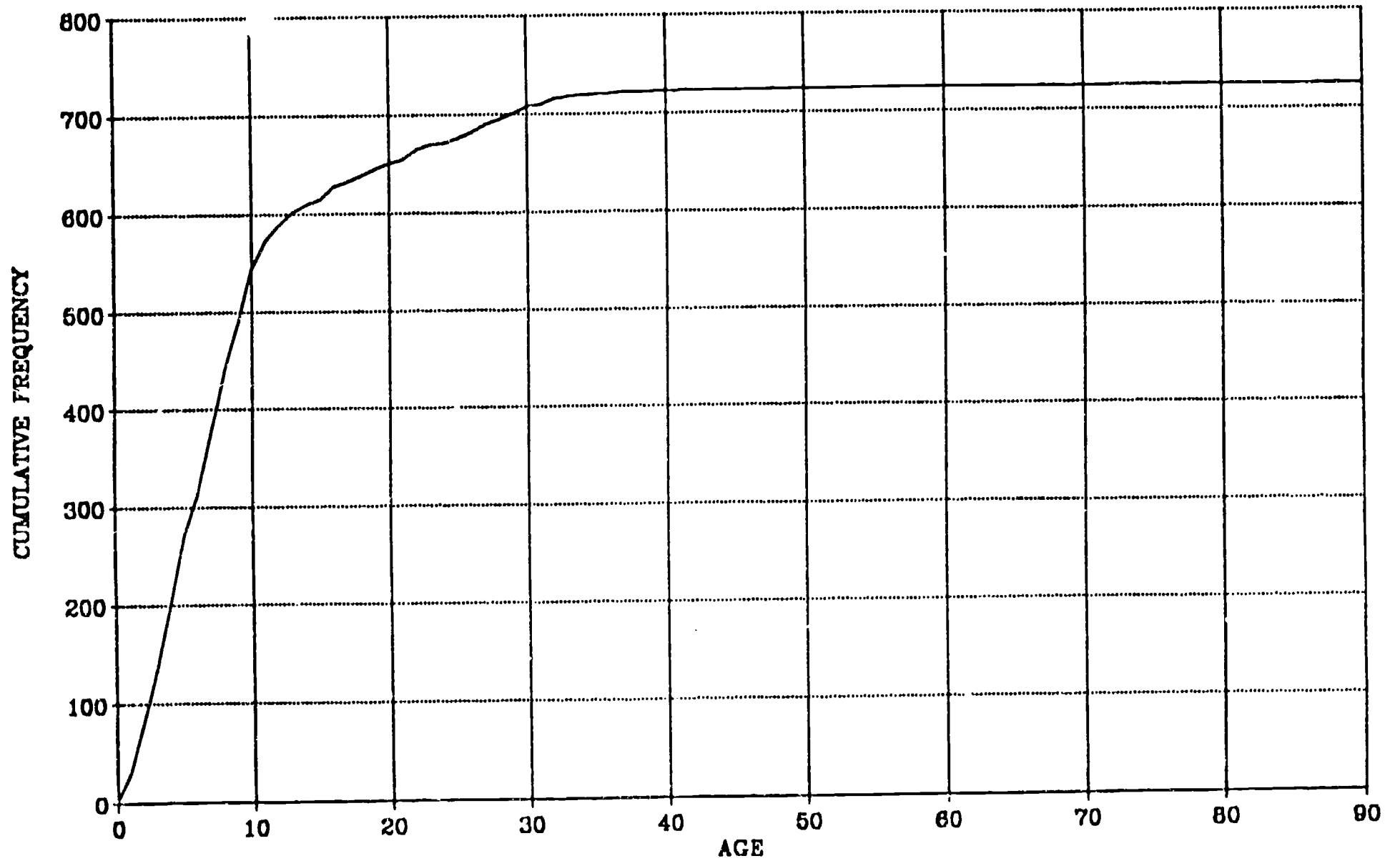


52 A

Figure 55

DEWEY 400s

LIBRARY 3

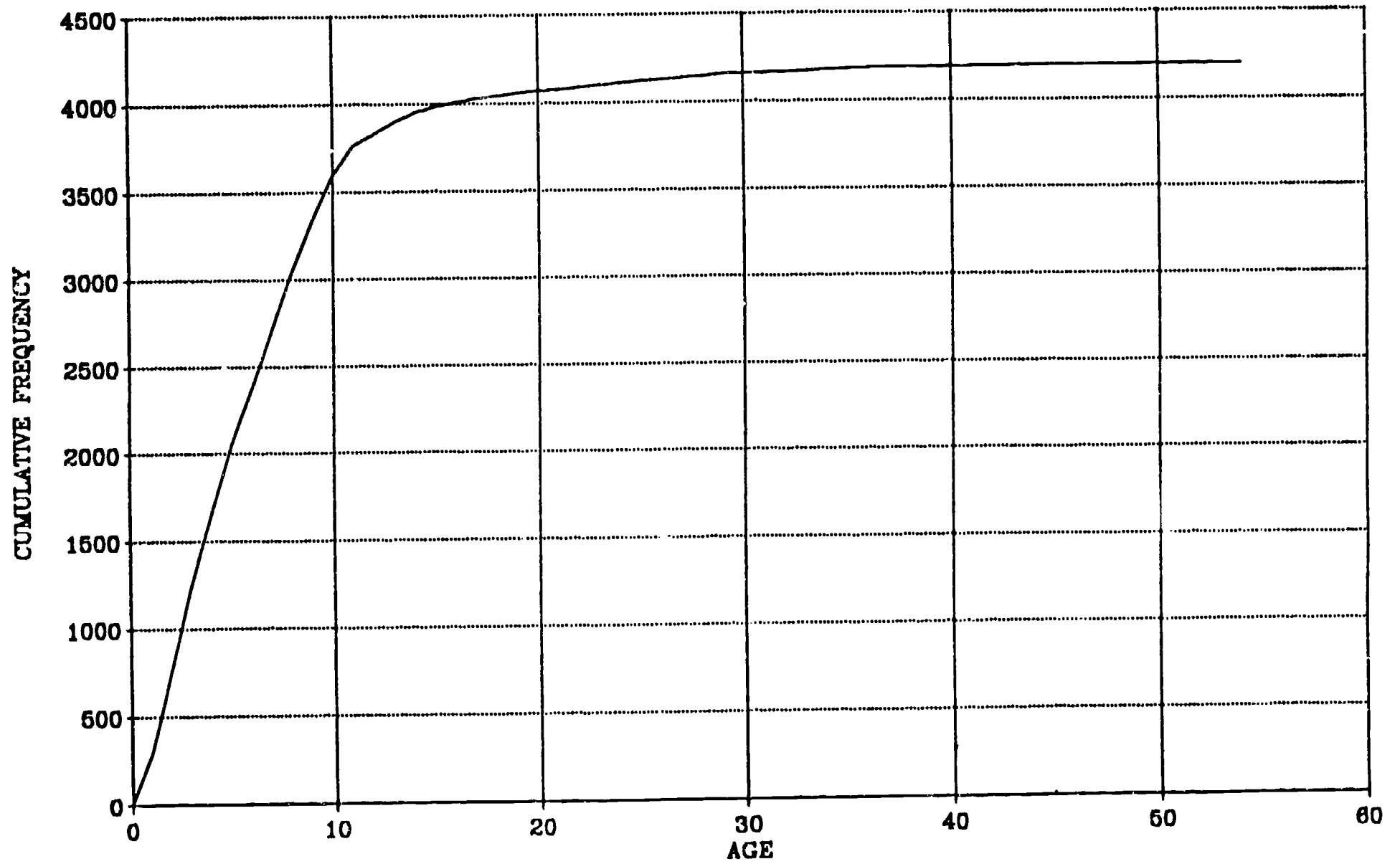


53A

Figure 56

DEWEY 500s

LIBRARY 3



54A

Figure 57

DEWEY 600s

LIBRARY 3

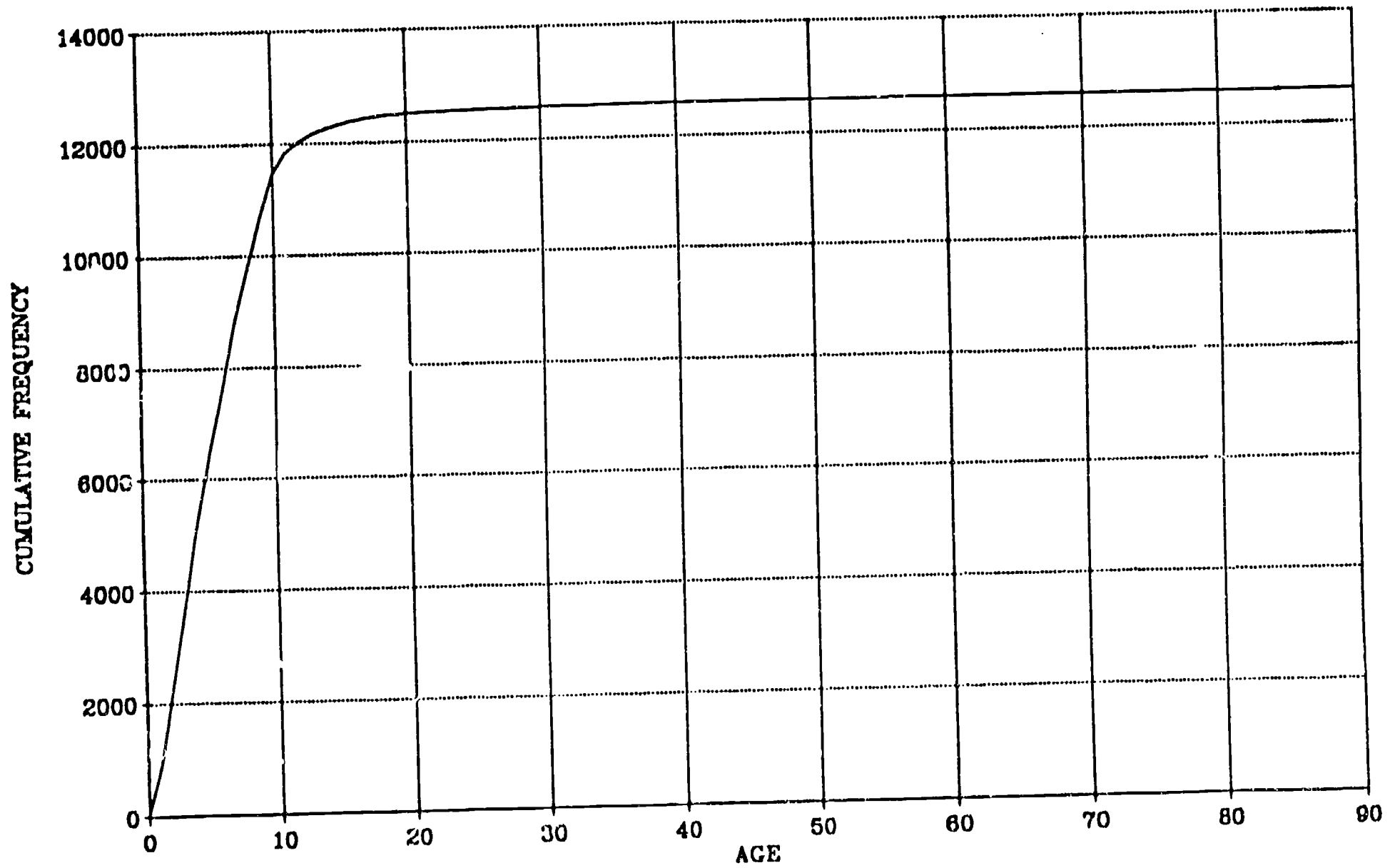
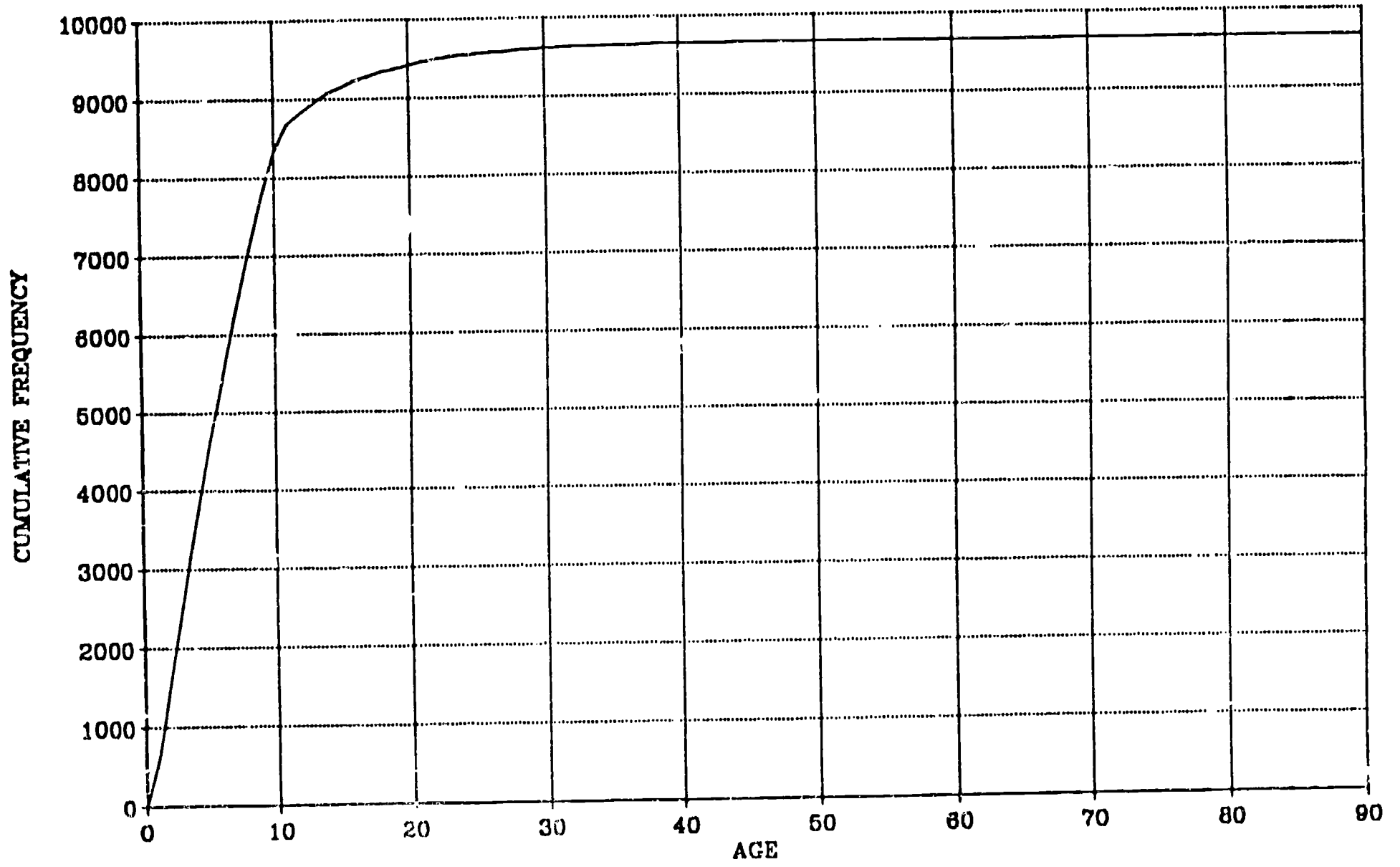


Figure 5c

DEWEY 700s

LIBRARY 3



56 A

Figure 59

DEWEY 800s

LIBRARY 3

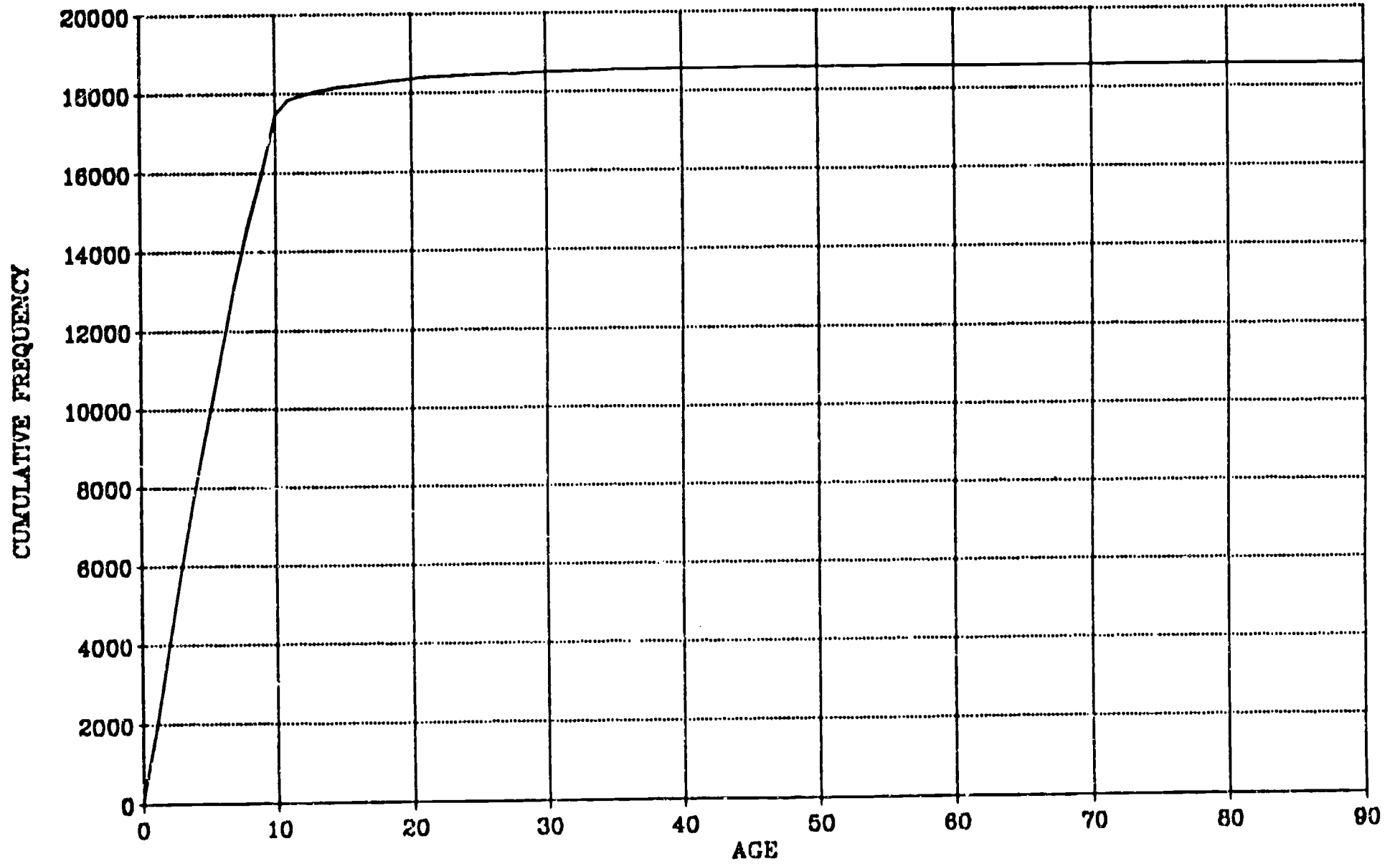
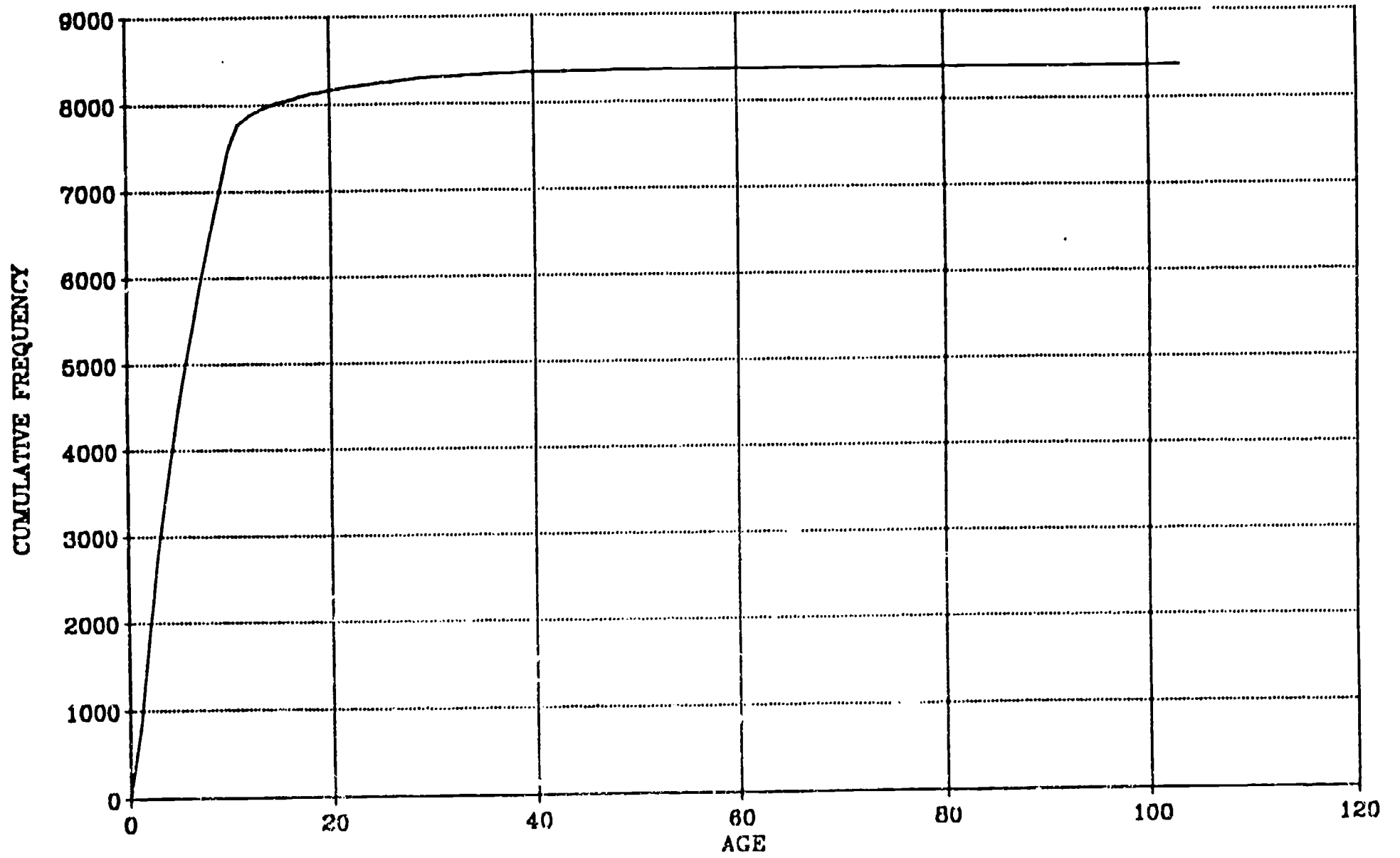


Figure 60

DEWEY 900s

LIBRARY 3



58A

Figure 61

DEWEY 000s

LIBRARY 4

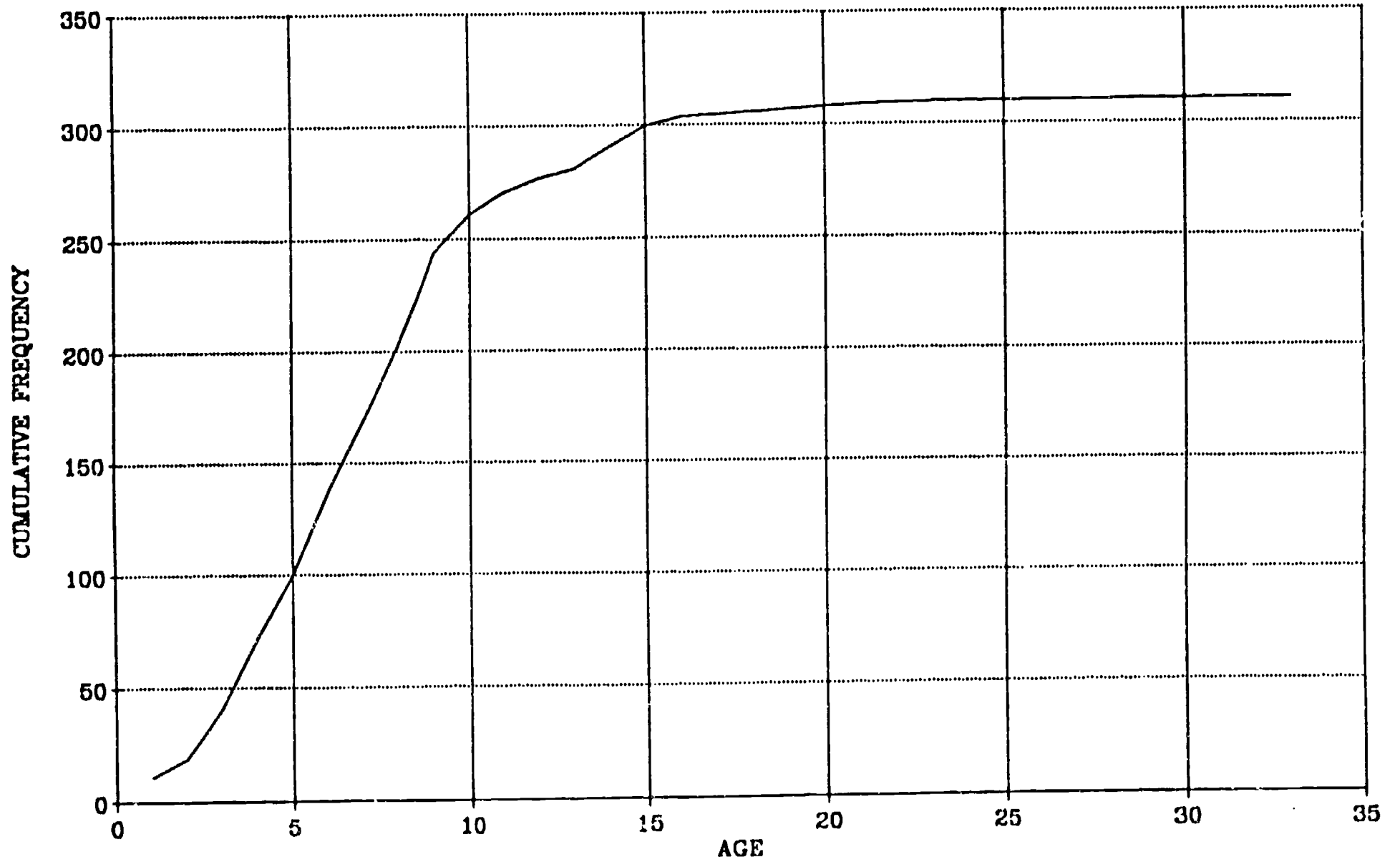
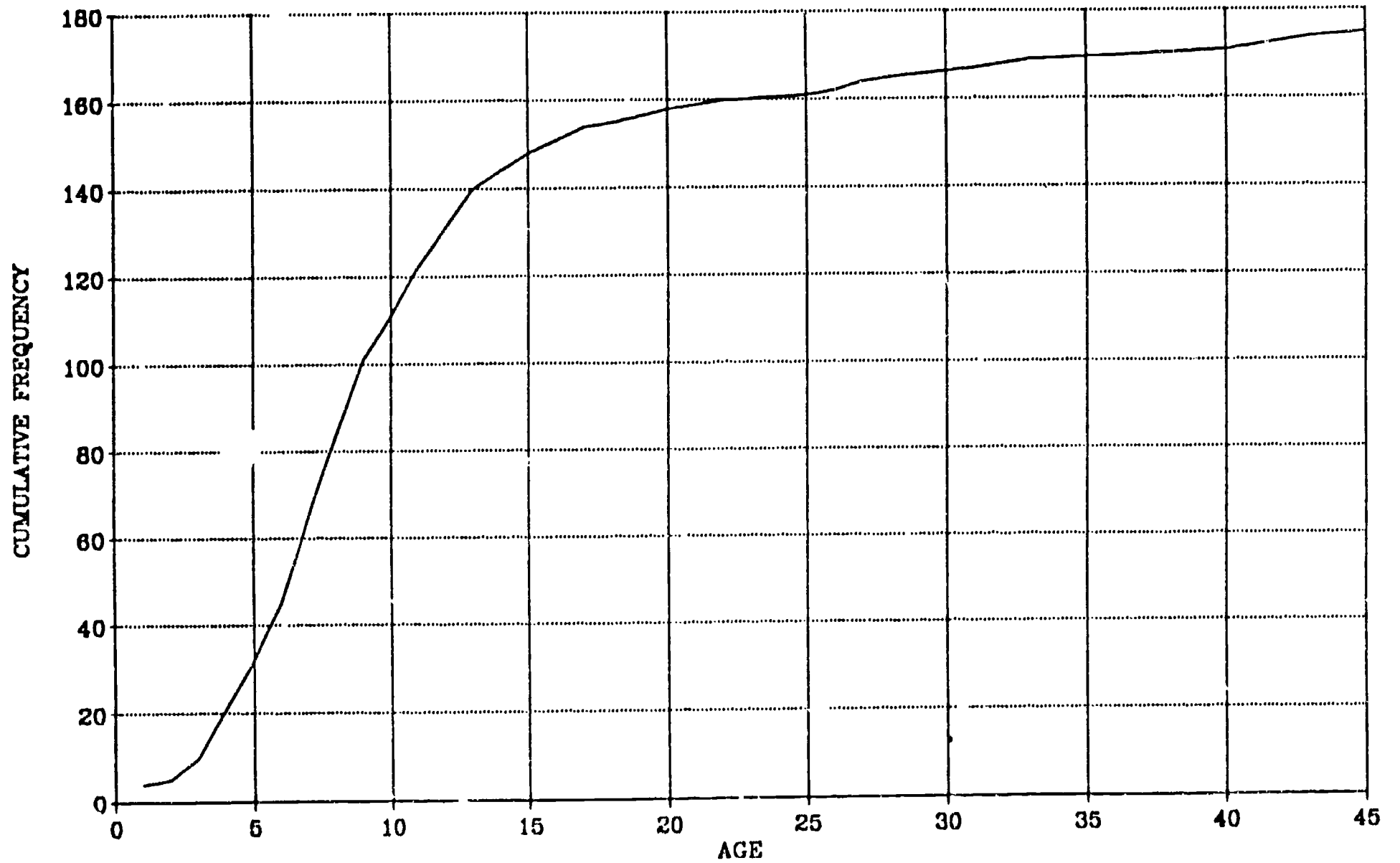


Figure 62

DEWEY 100s

LIBRARY 4



(6)

(6)A

Figure 63

DEWEY 200s

LIBRARY 4

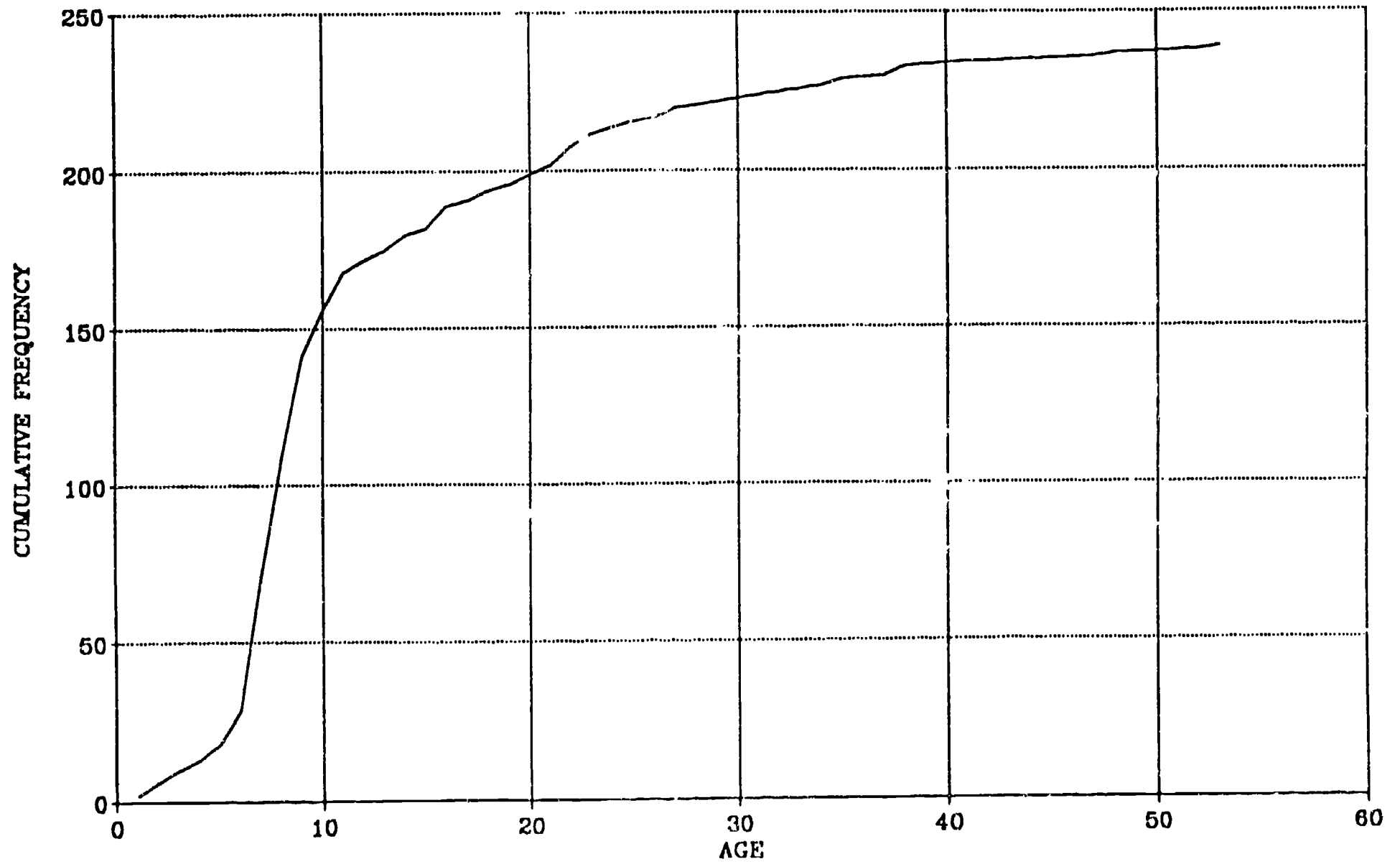


Figure 64

DEWEY 300s

LIBRARY 4

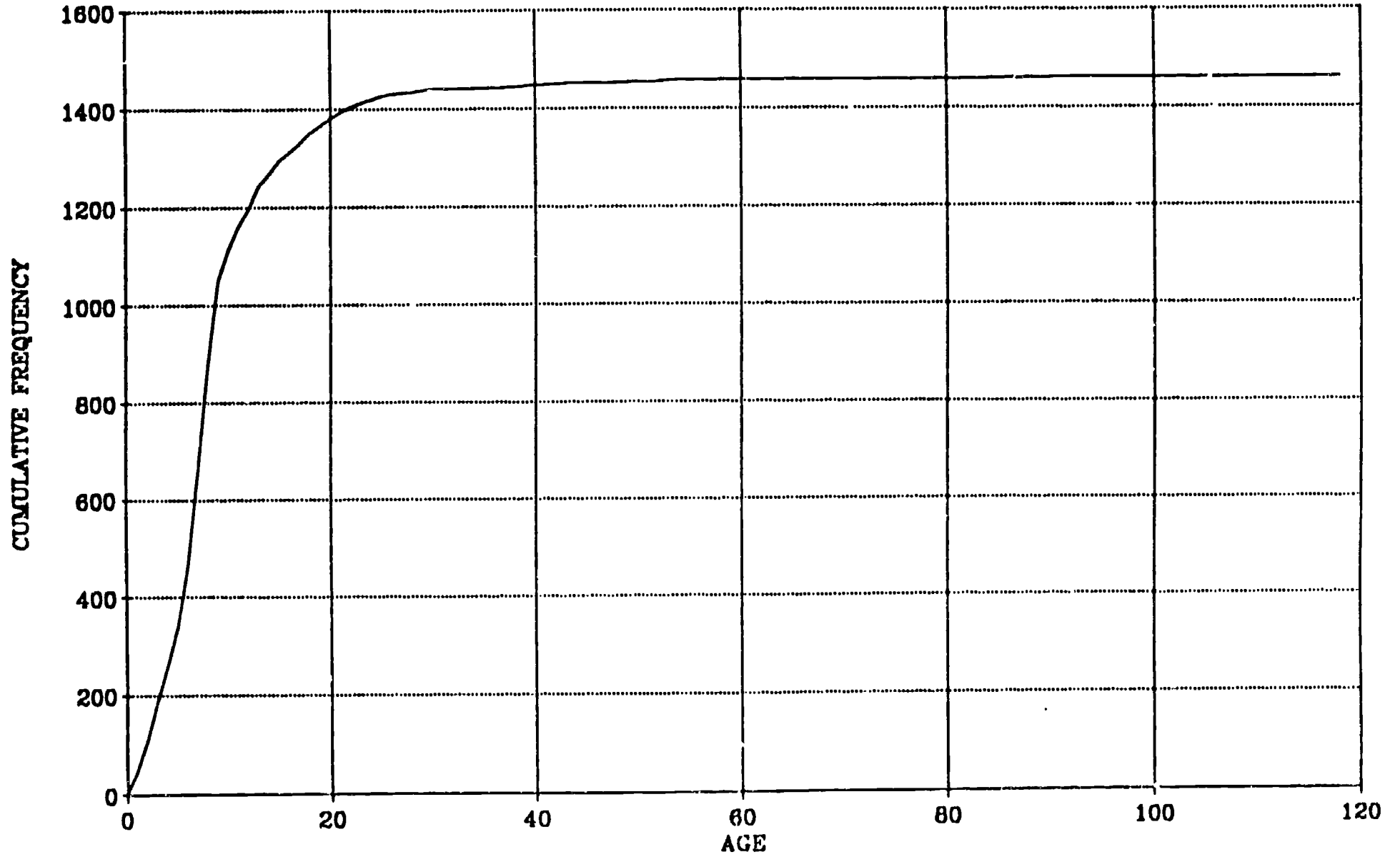


Figure 65

DEWEY 400s

LIBRARY 4

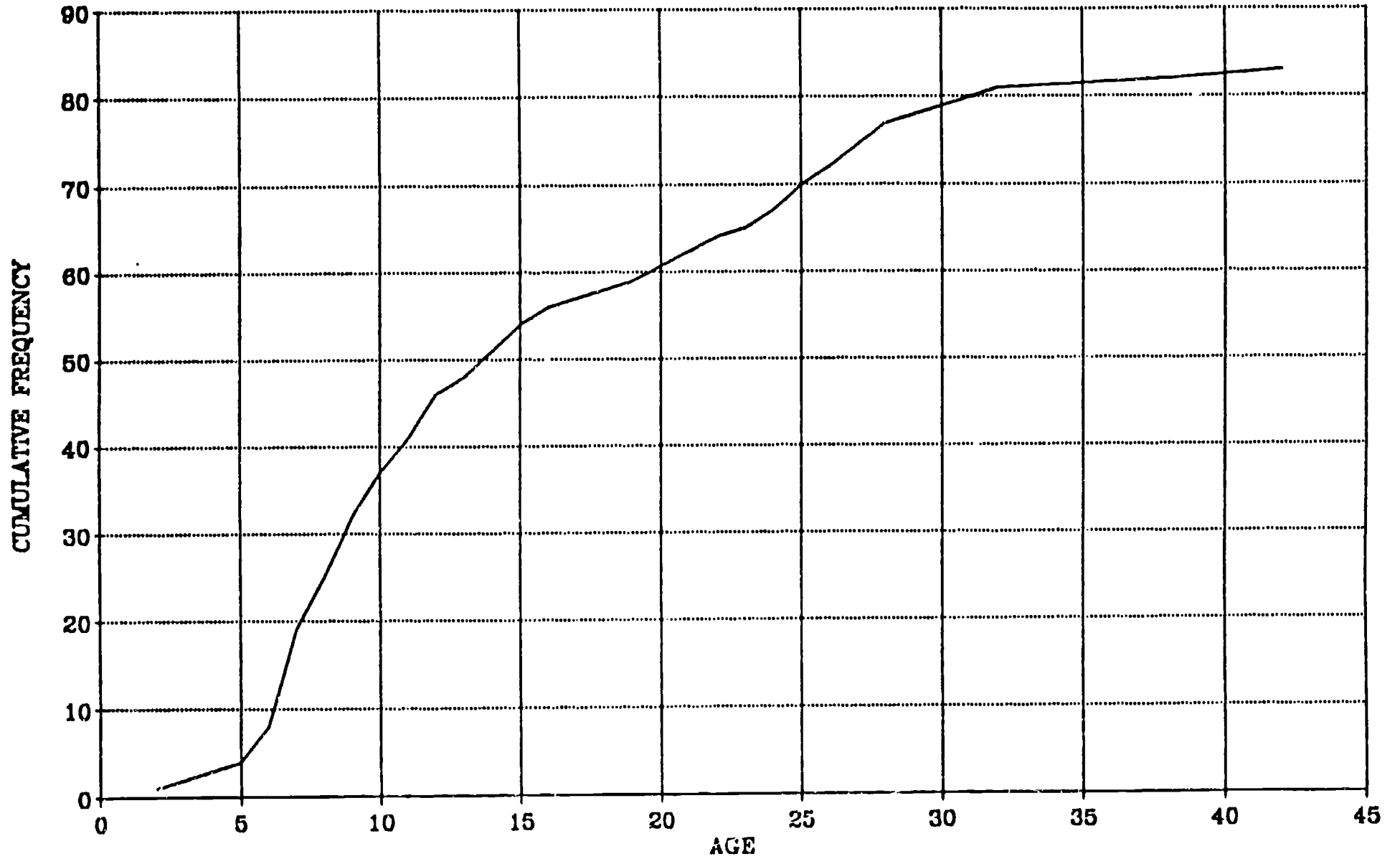


Figure 66

DEWEY 500s

LIBRARY 4

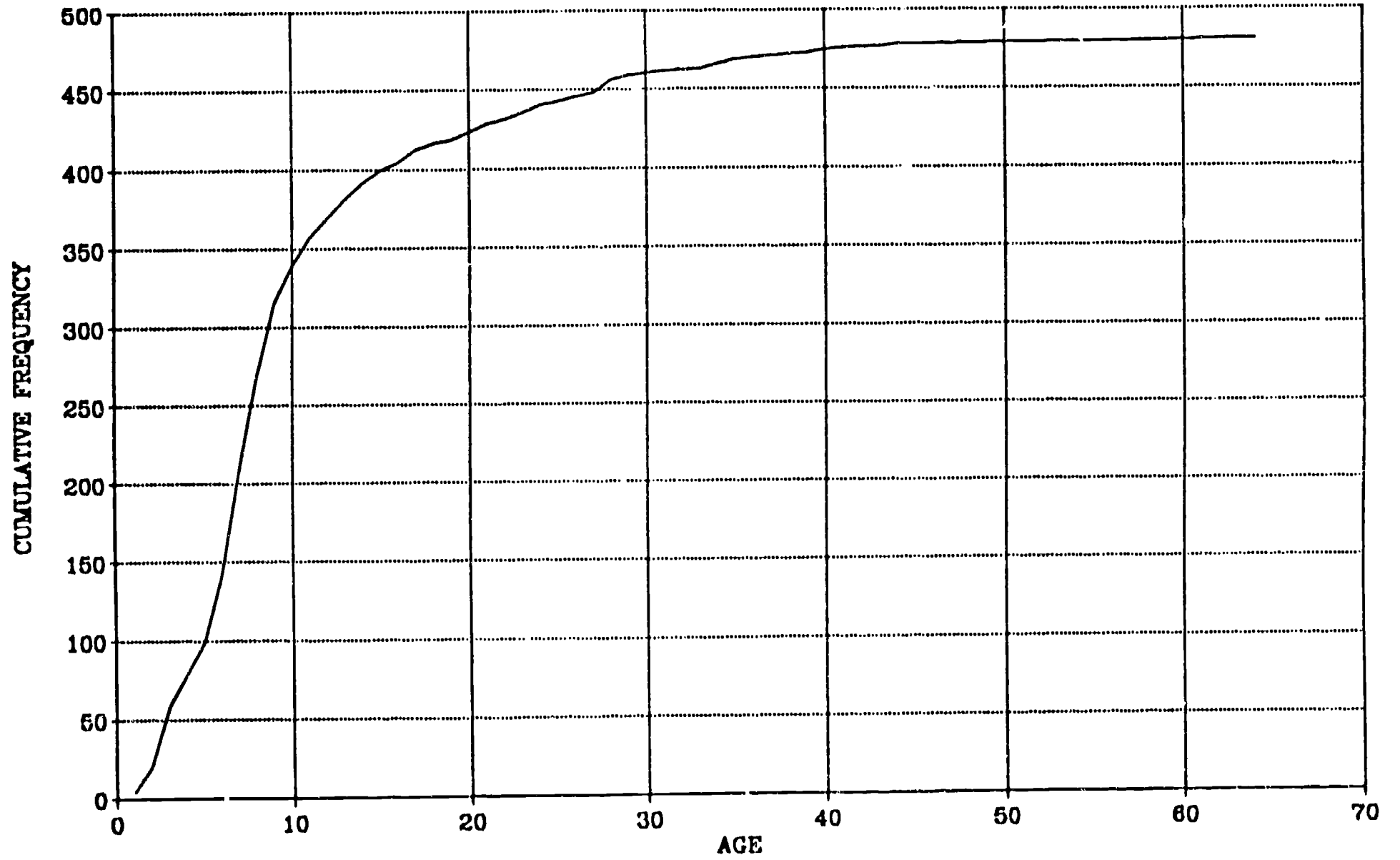


Figure 67

DEWEY 600s

LIBRARY 4

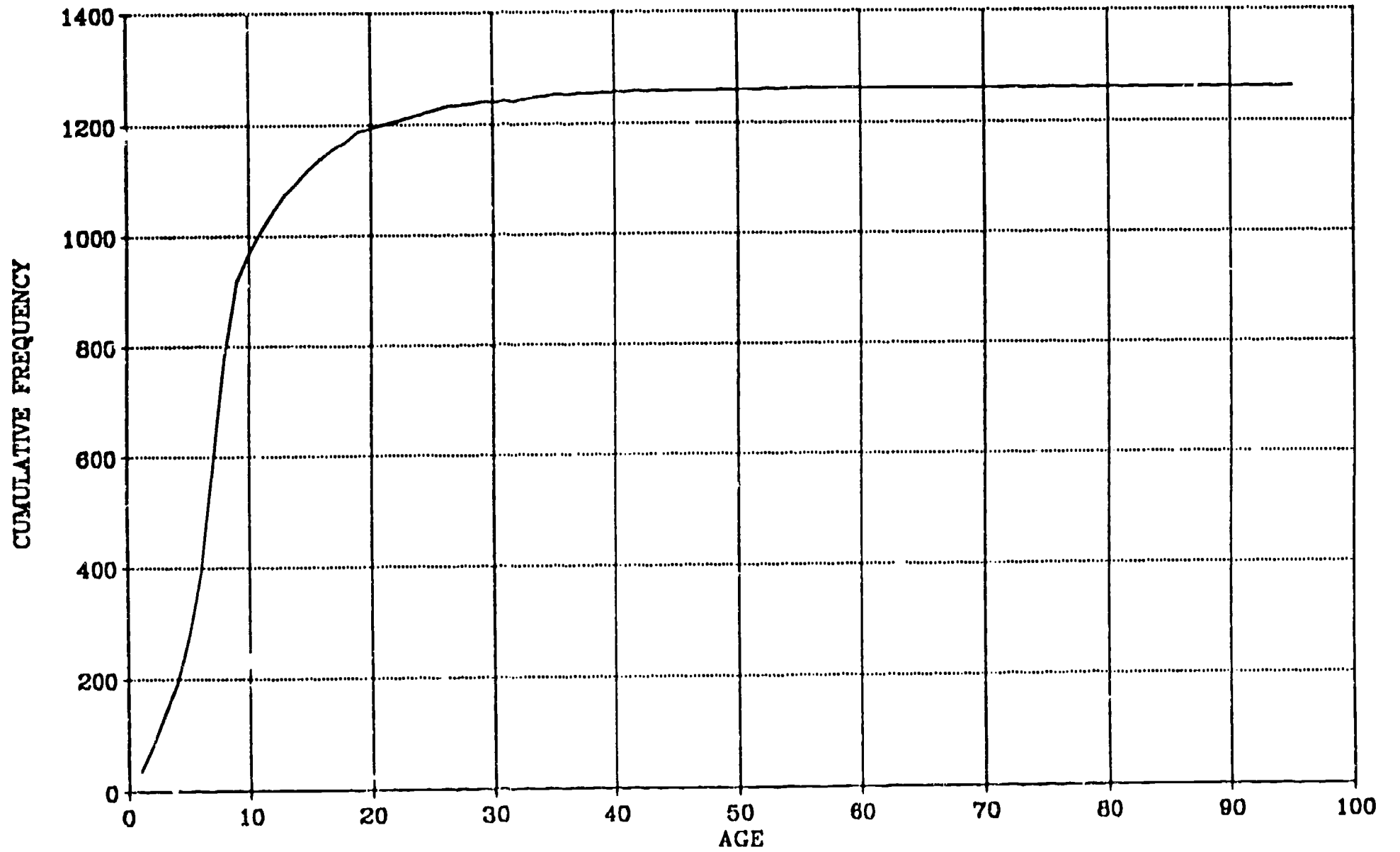


Figure 68

DEWEY 700s

LIBRARY 4

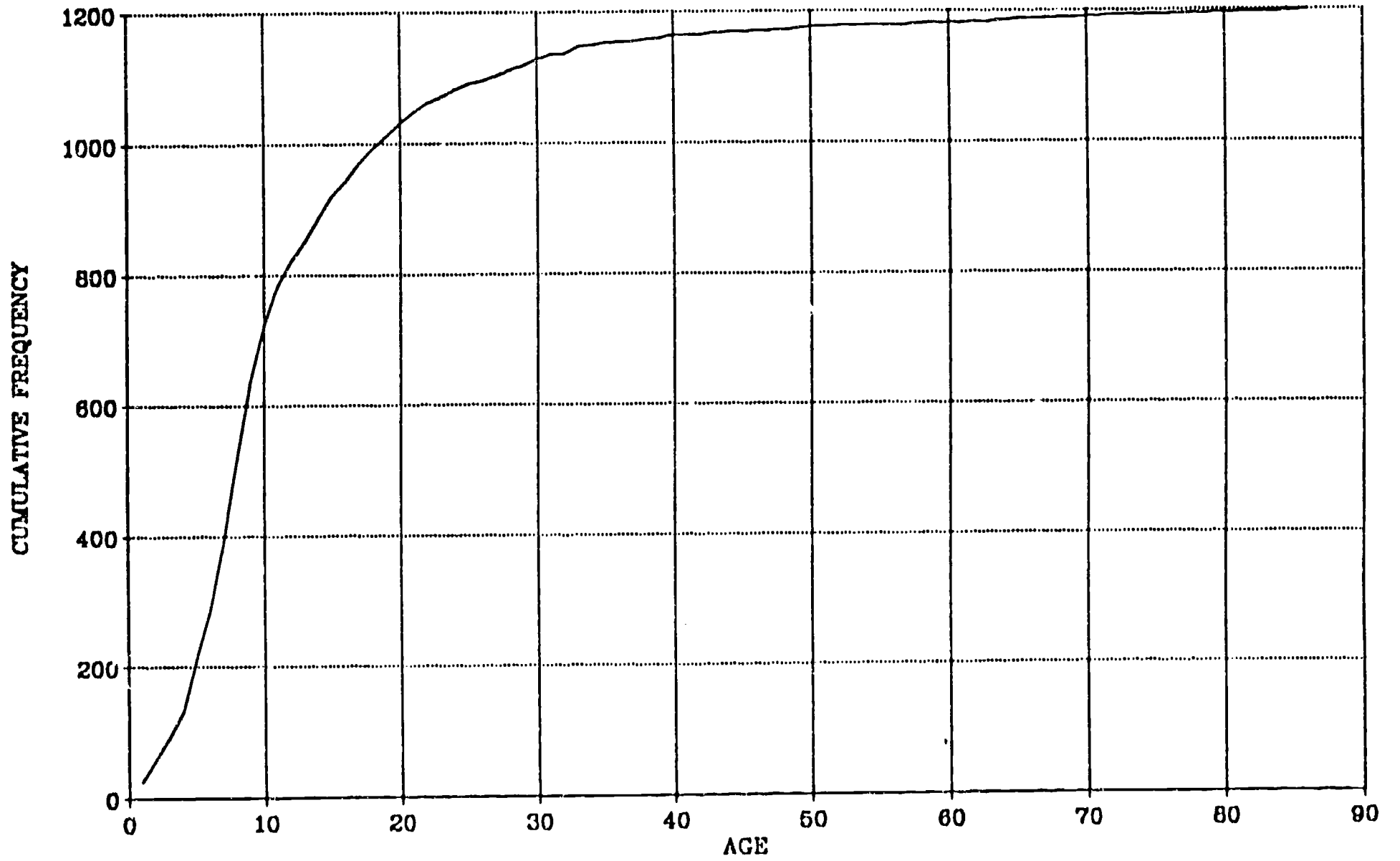


Figure 69

DEWEY 800s

LIBRARY 4

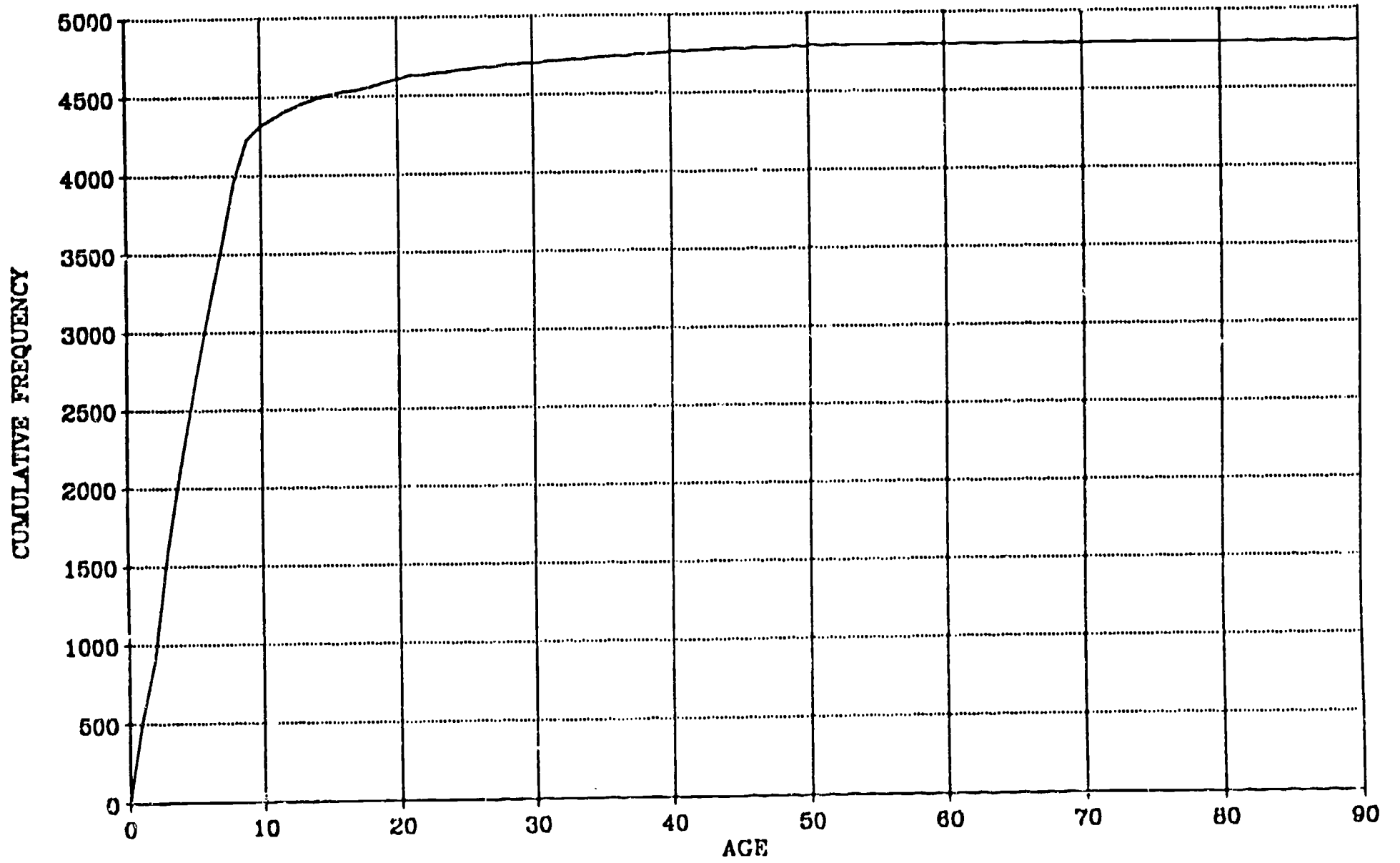


Figure 70

DEWEY 900s

LIBRARY 4

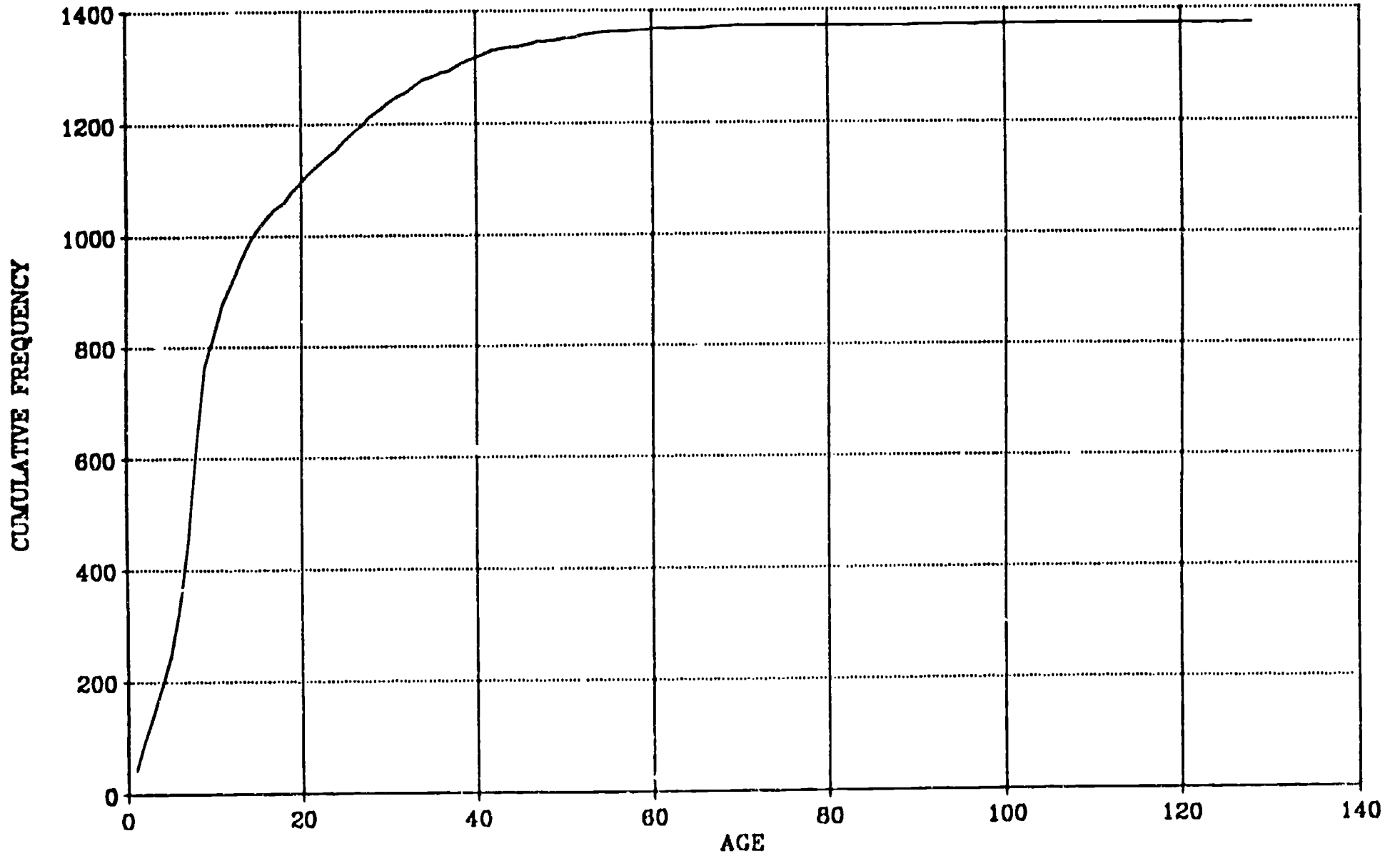


Figure 71

DEWEY 000s

LIBRARY 5

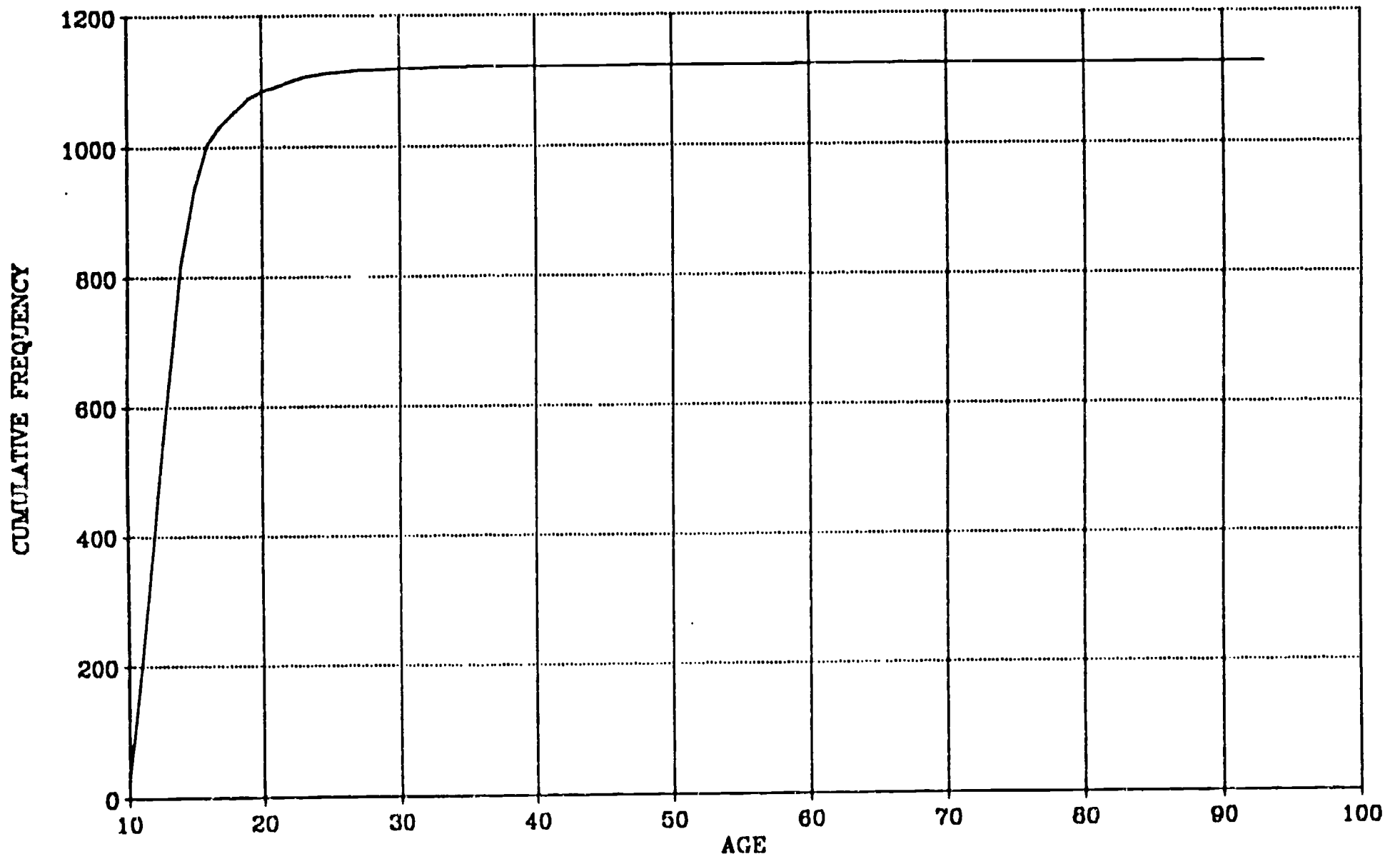


Figure 72

DEWEY 100s
LIBRARY 5

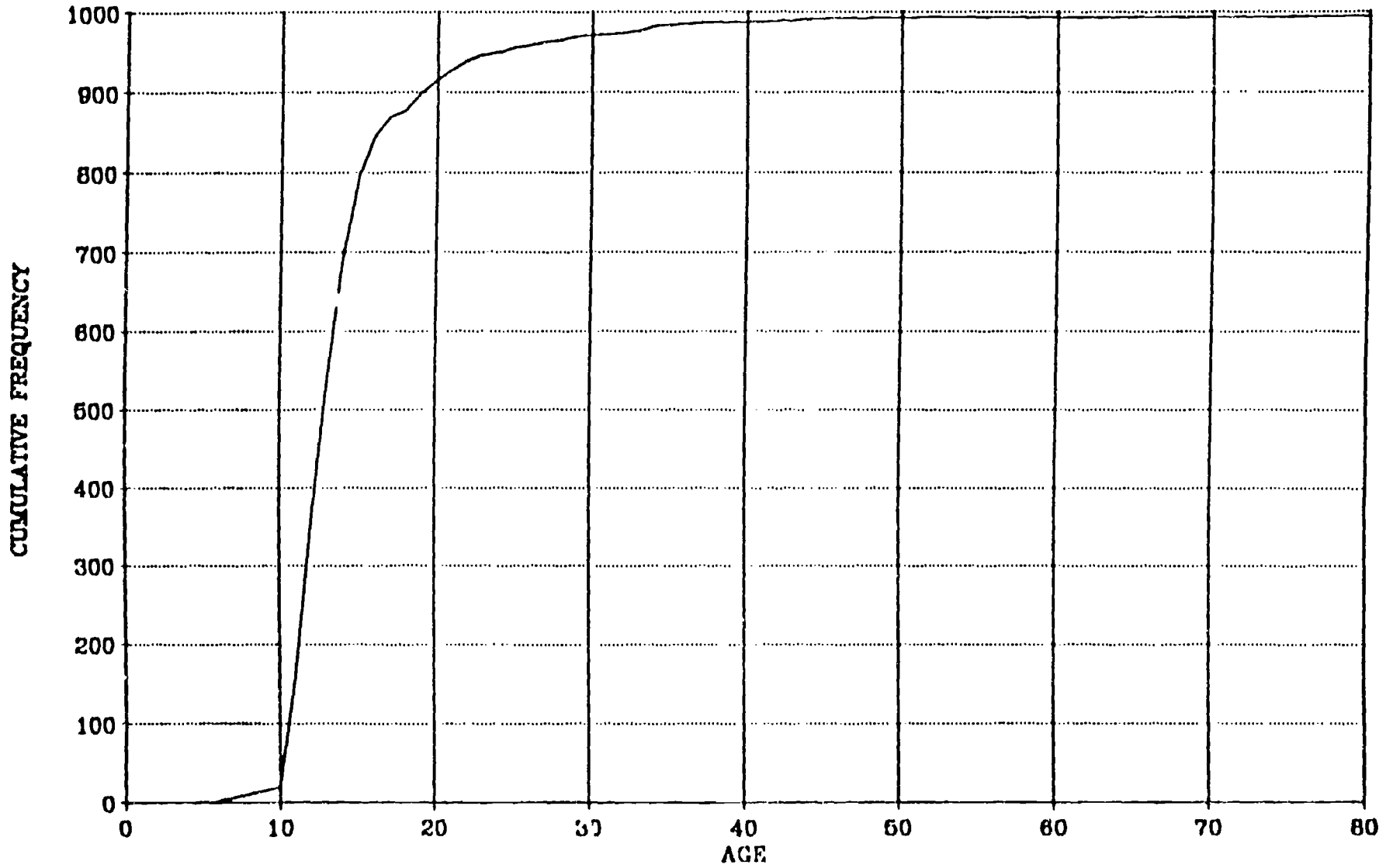


Figure 73

DEWEY 200s

LIBRARY 5

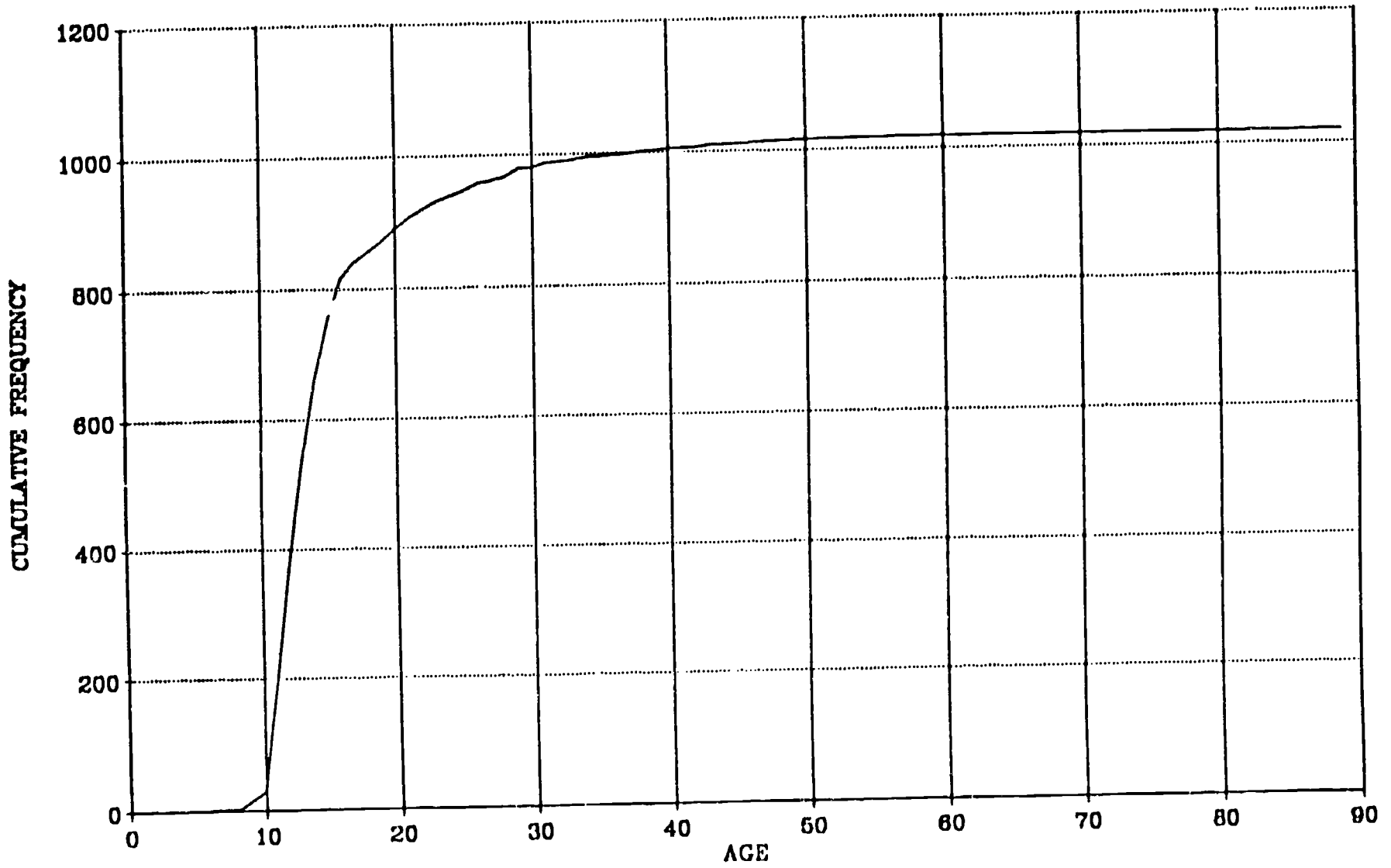


Figure 74

DEWEY 300s

LIBRARY 5

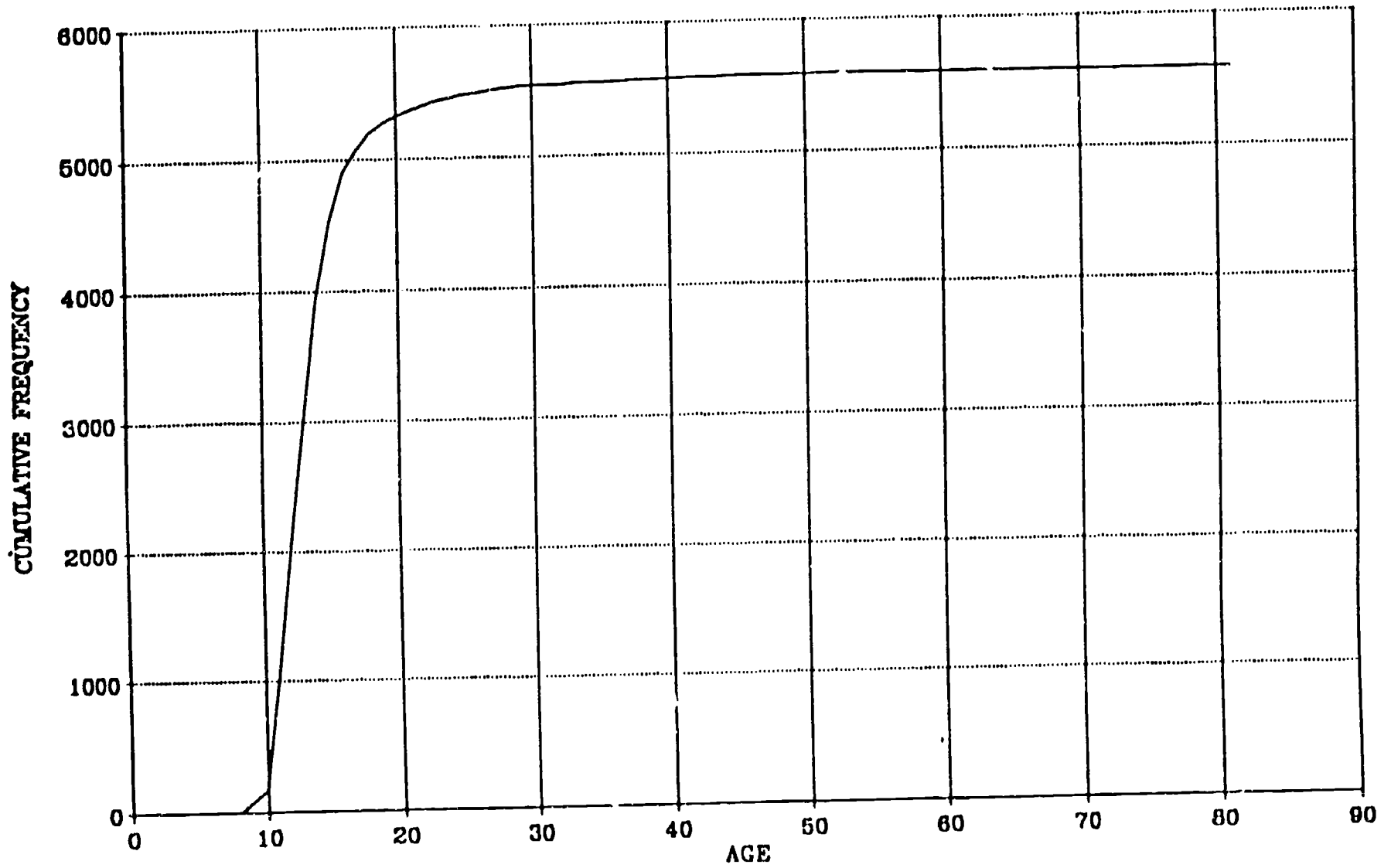


Figure 75

DEWEY 400s

LIBRARY 5

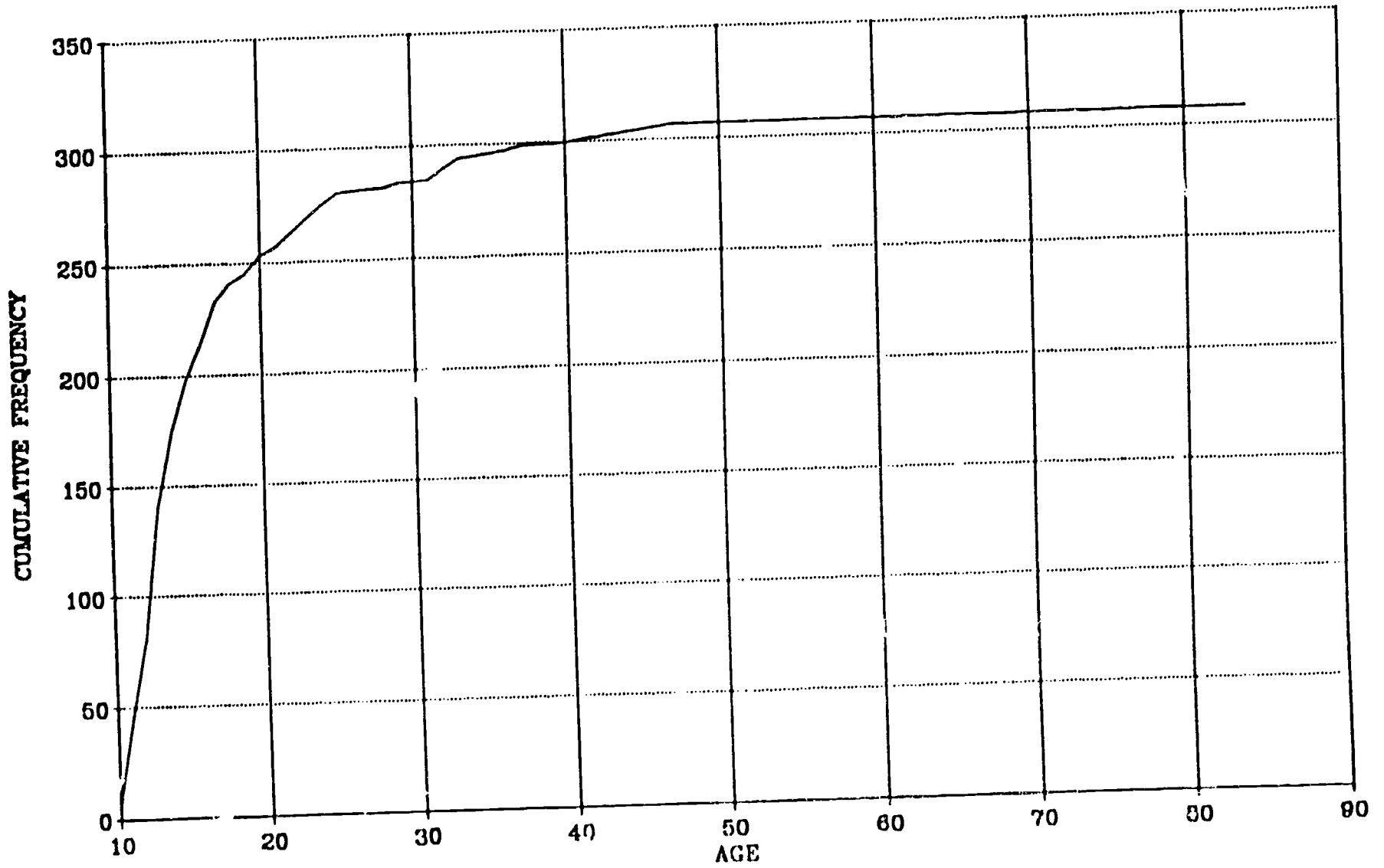


Figure 76

DEWEY 500s

LIBRARY 5

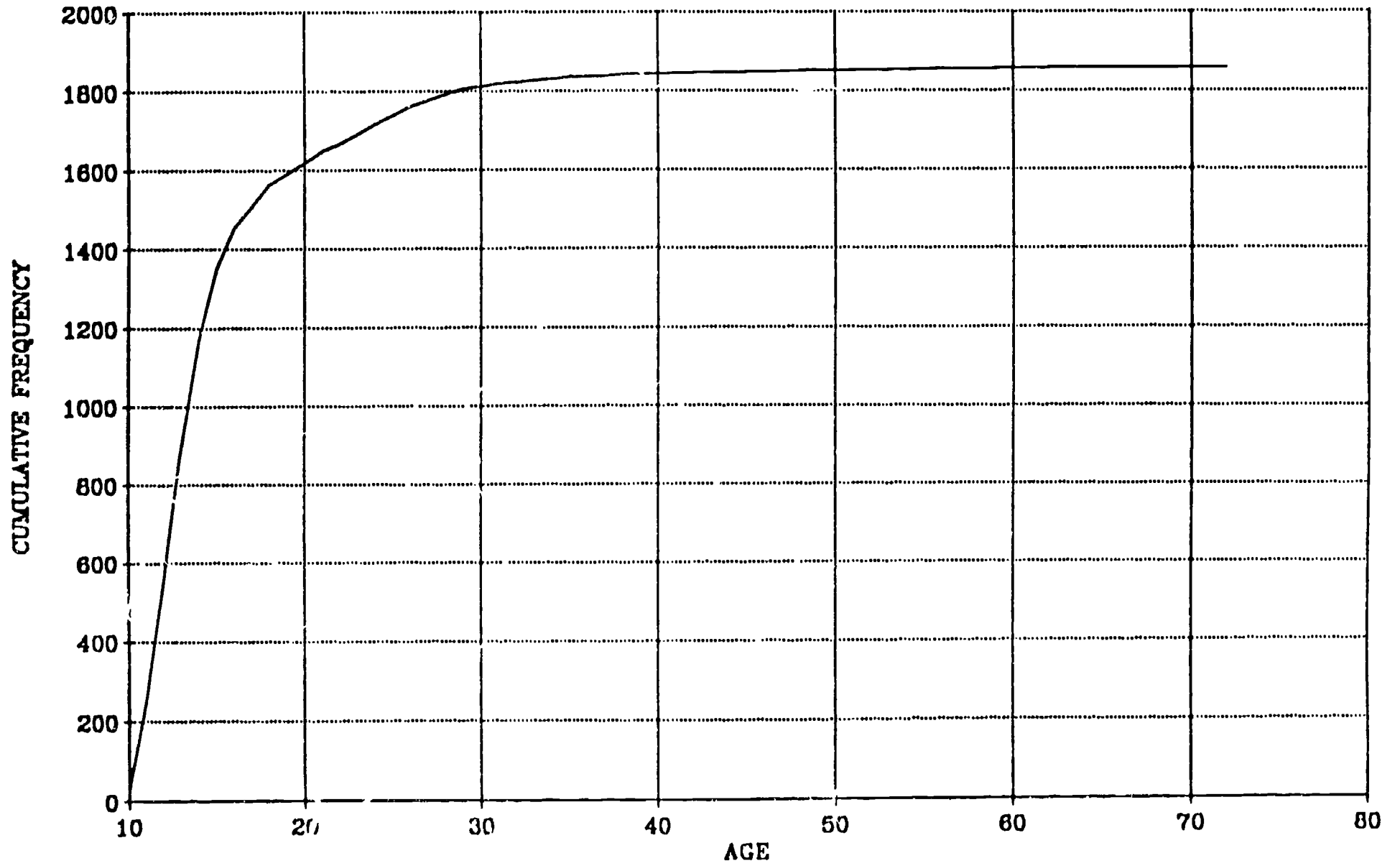


Figure 77

DEWEY 600s

LIBRARY 5

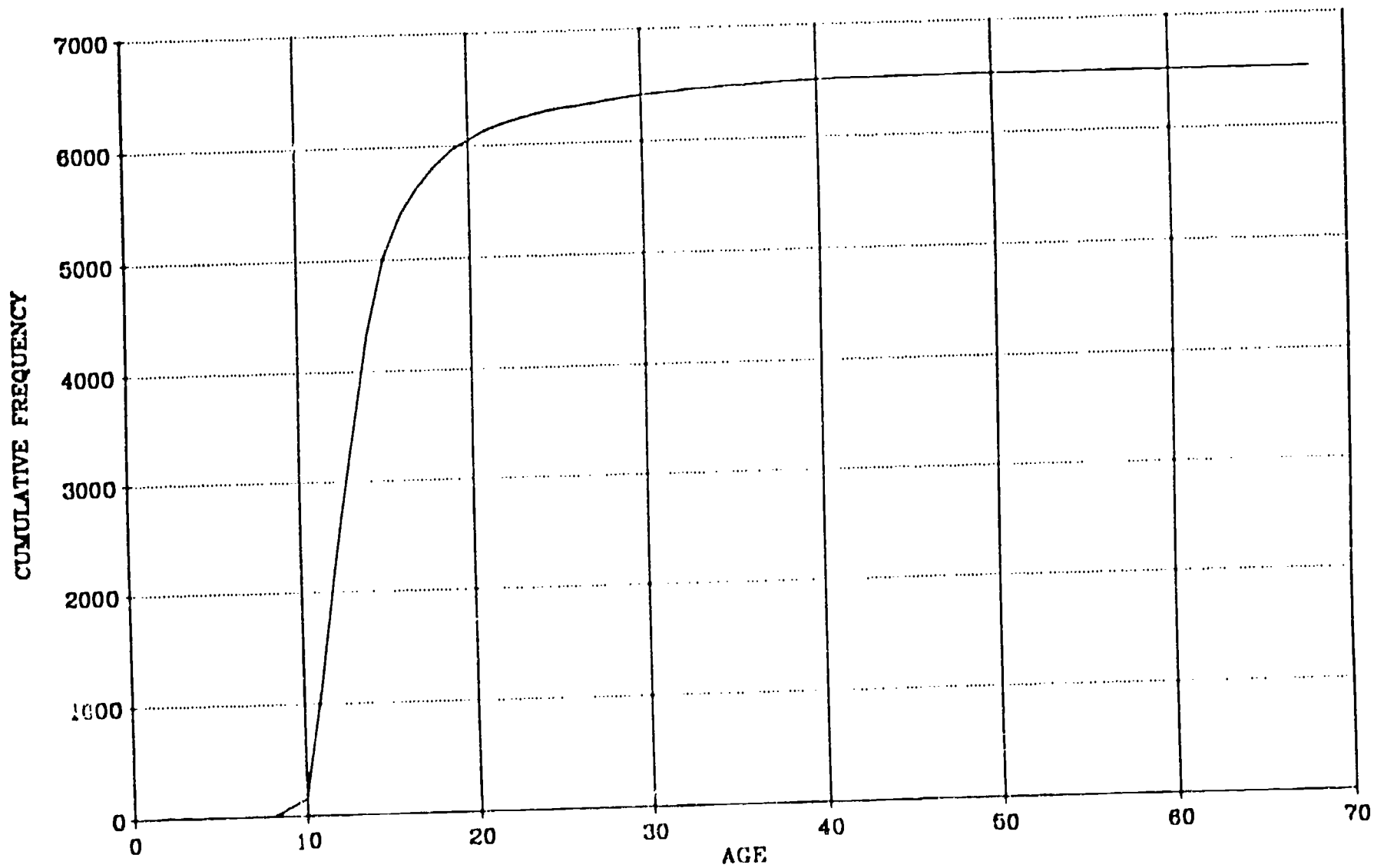


Figure 78

DEWEY 700s

LIBRARY 5

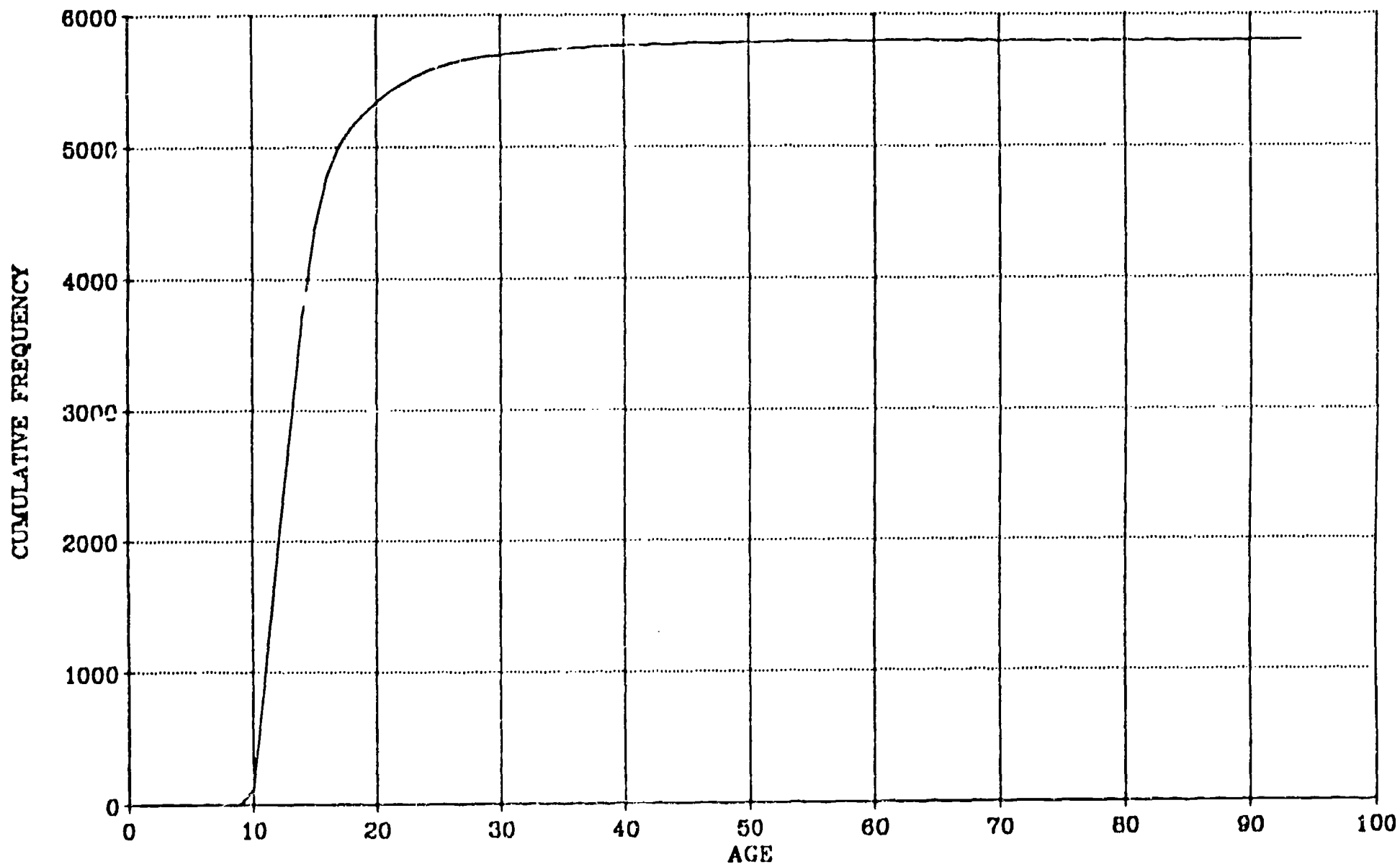


Figure 79

DEWEY 800s

LIBRARY 5

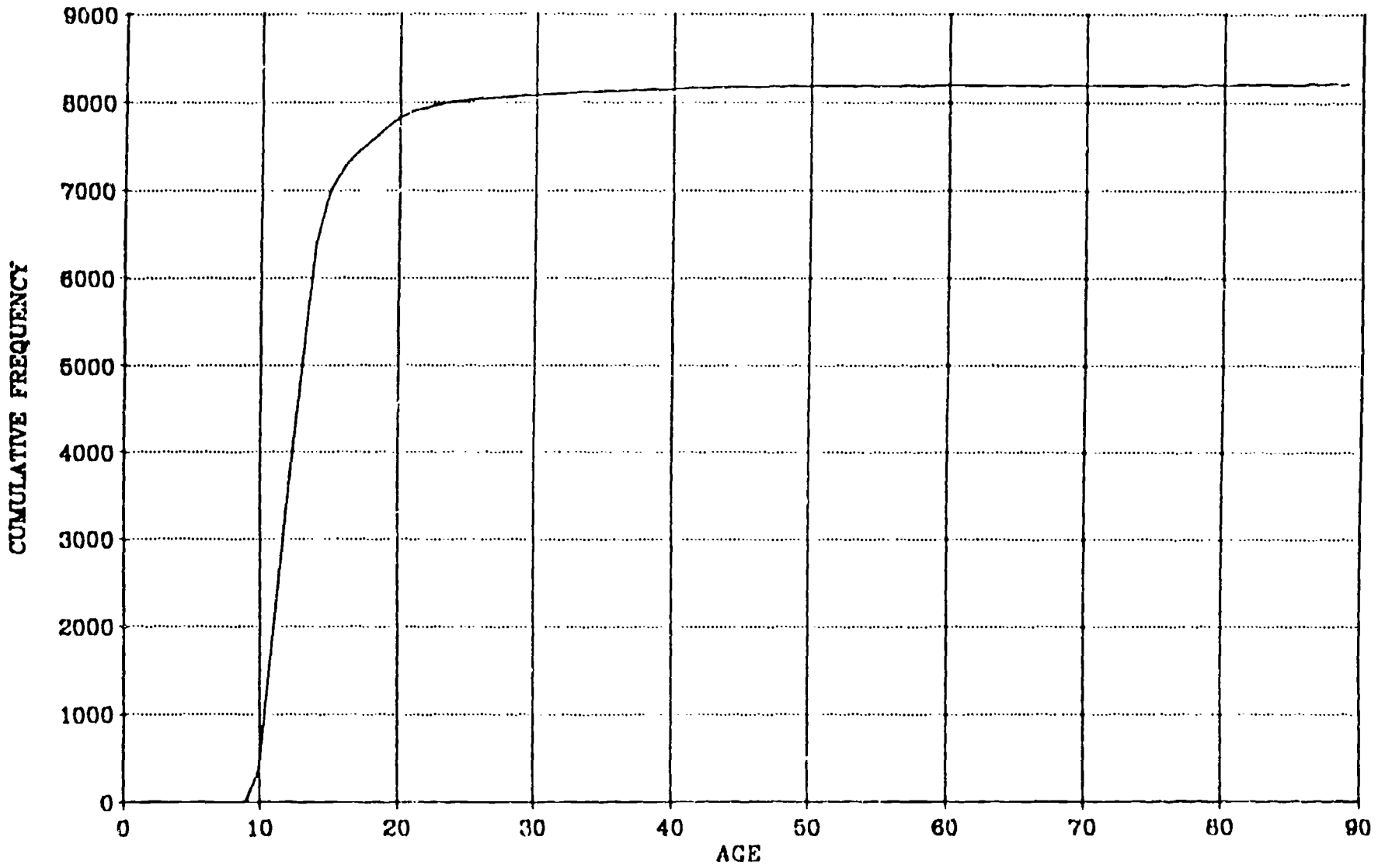


Figure 80

DEWEY 900s

LIBRARY 5

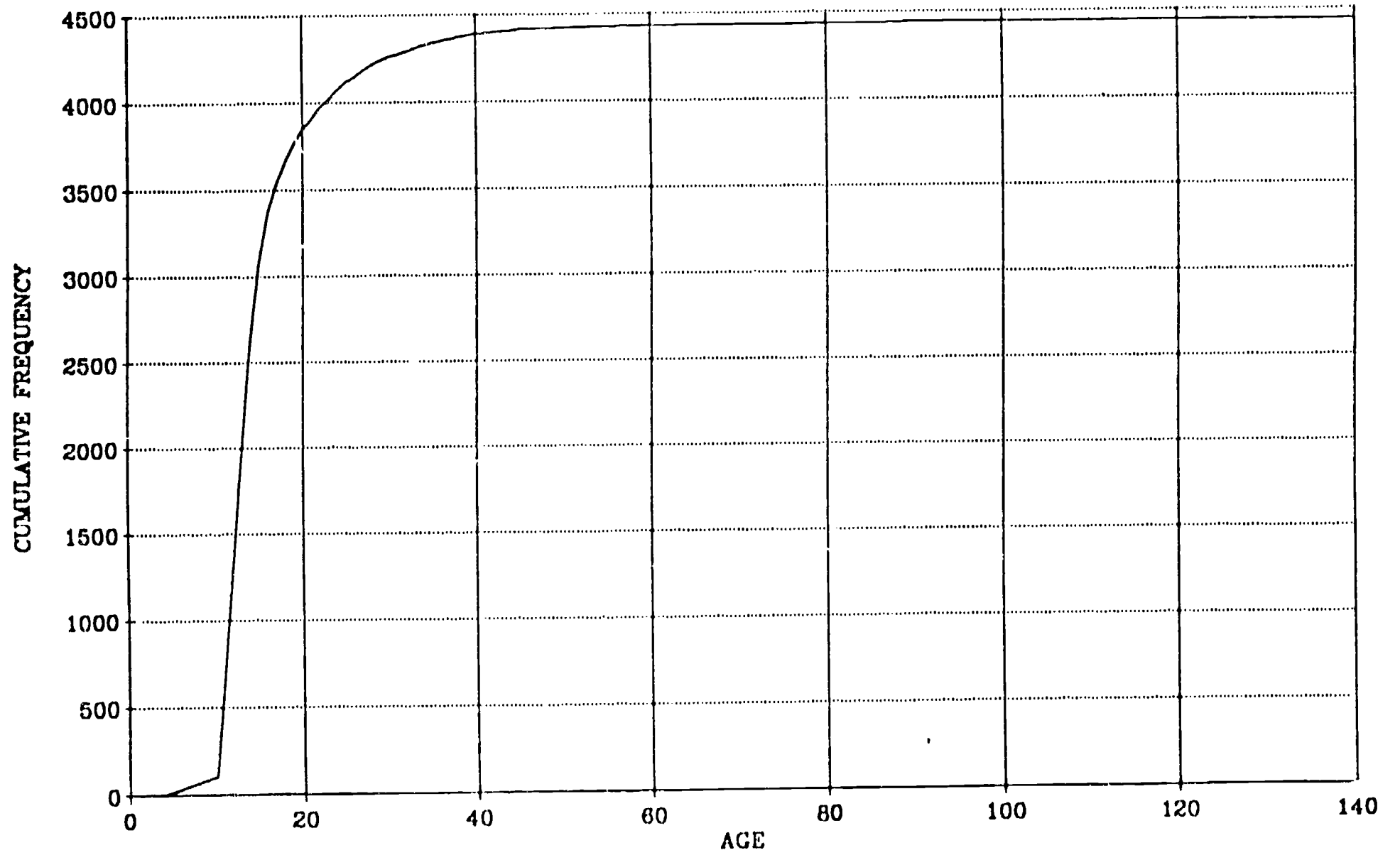


Figure 81

DEWEY 000s

LIBRARY 6

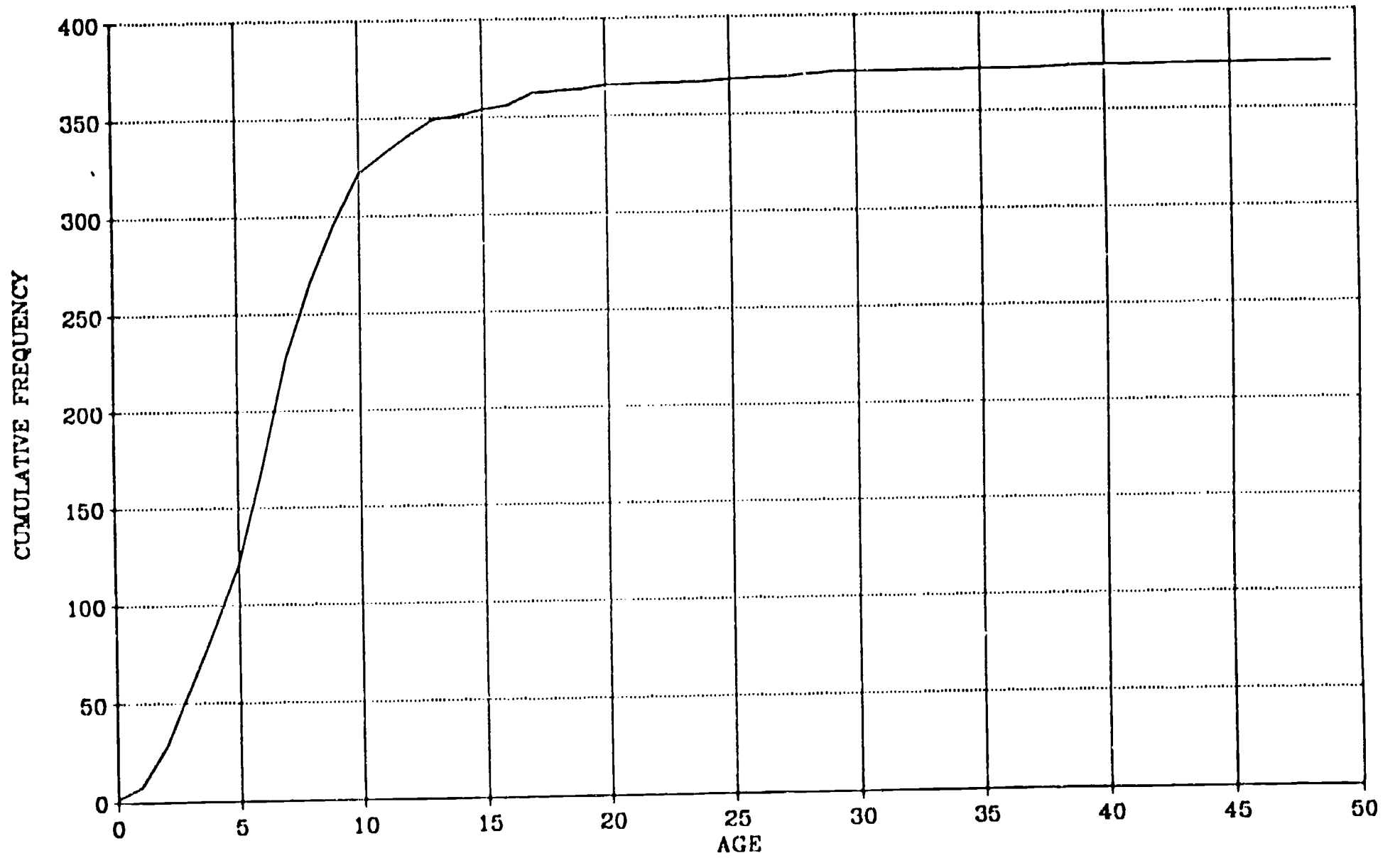


Figure 82

DEWEY 100s

LIBRARY 8

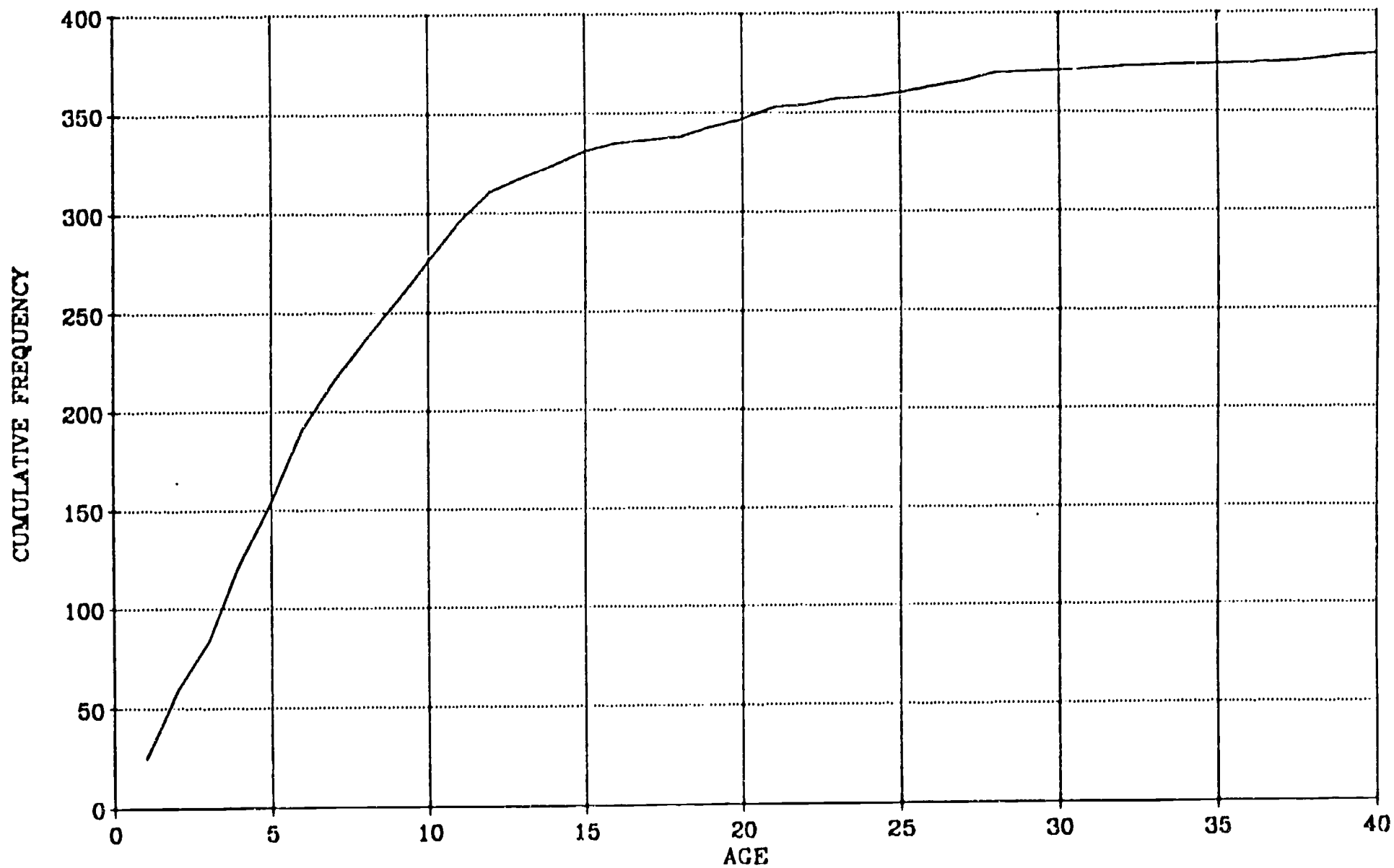


Figure 83

DEWEY 200s

LIBRARY 8

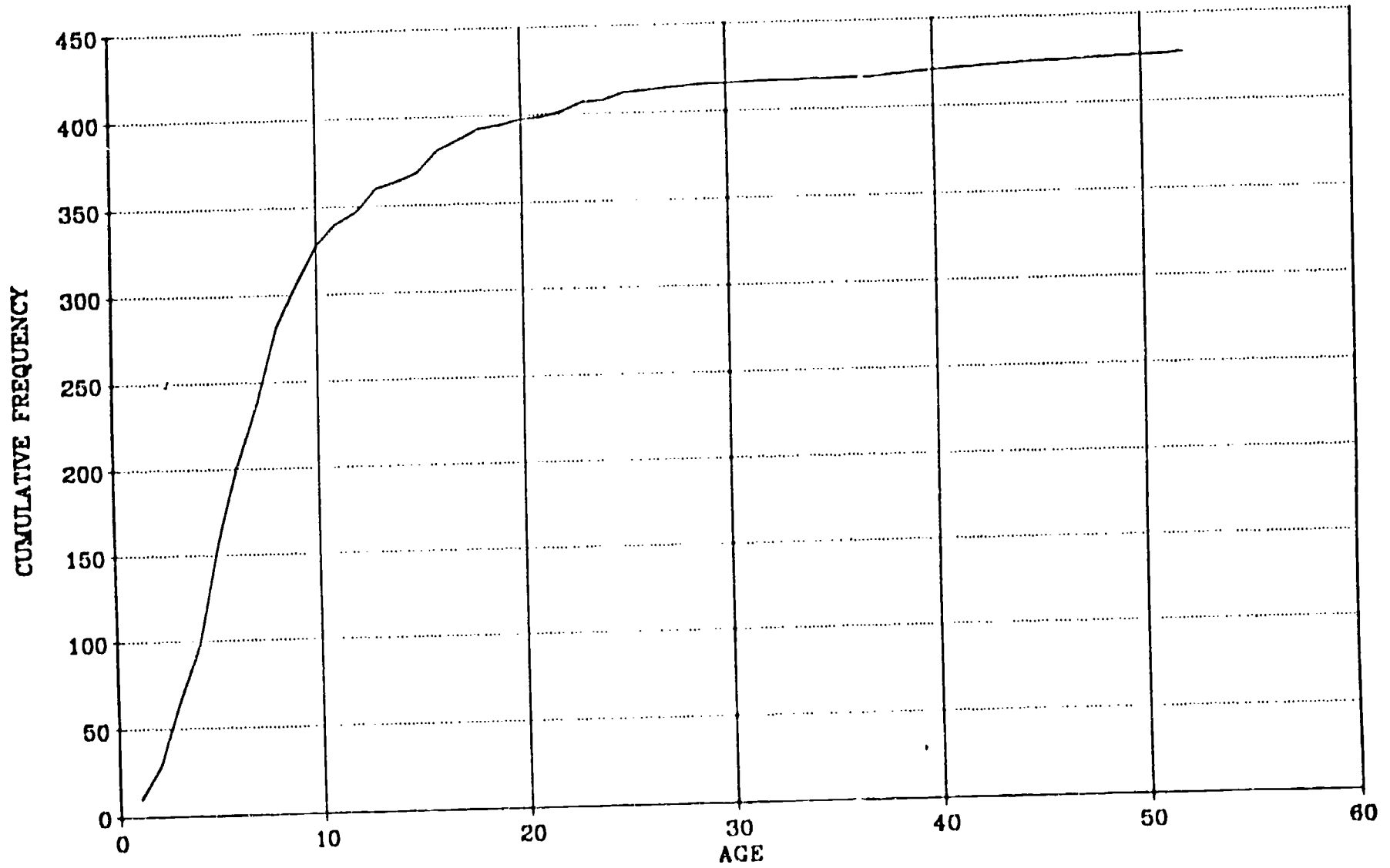


Figure 84

DEWEY 300s

LIBRARY 6

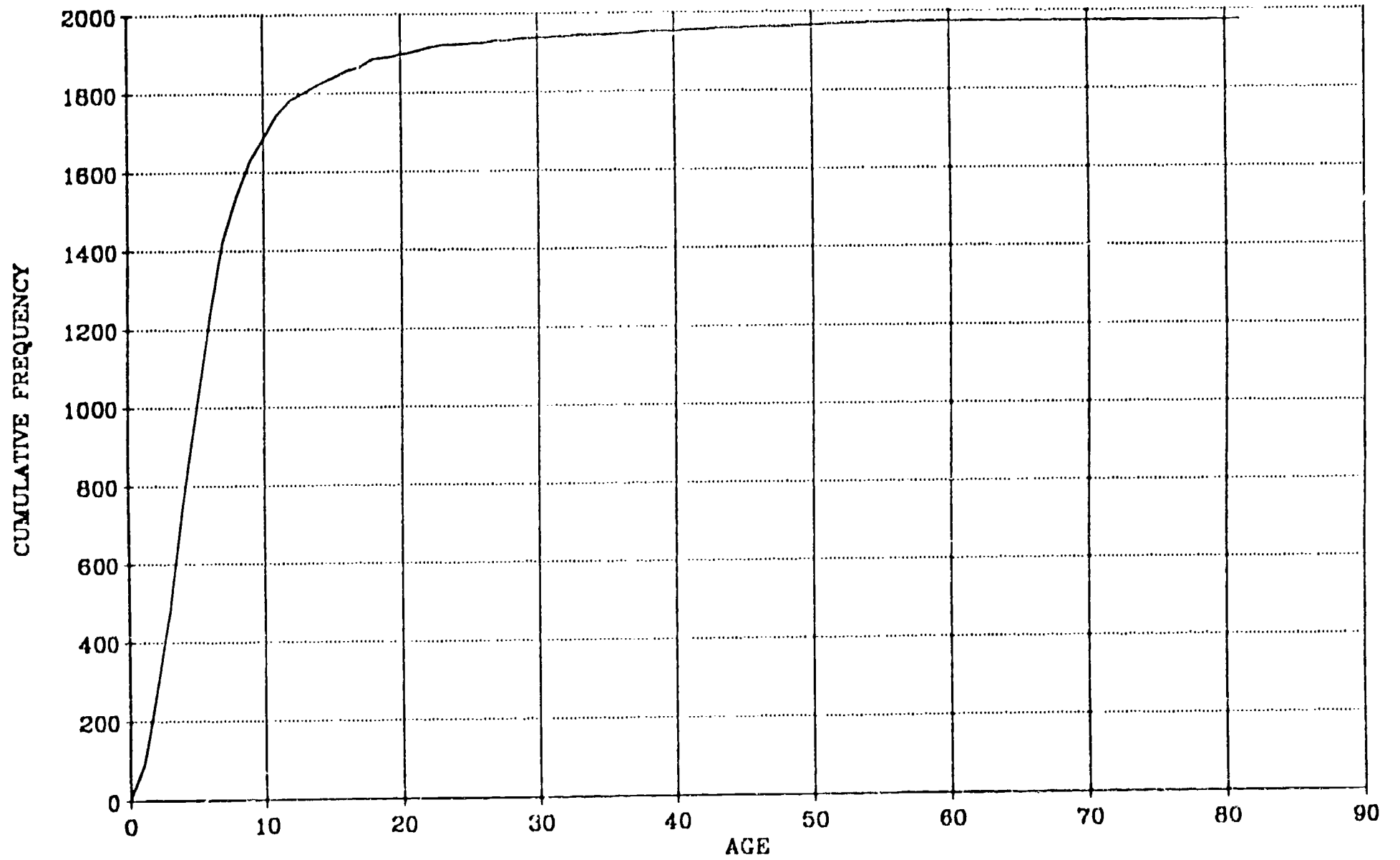


Figure 85

DEWEY 400s

LIBRARY 6

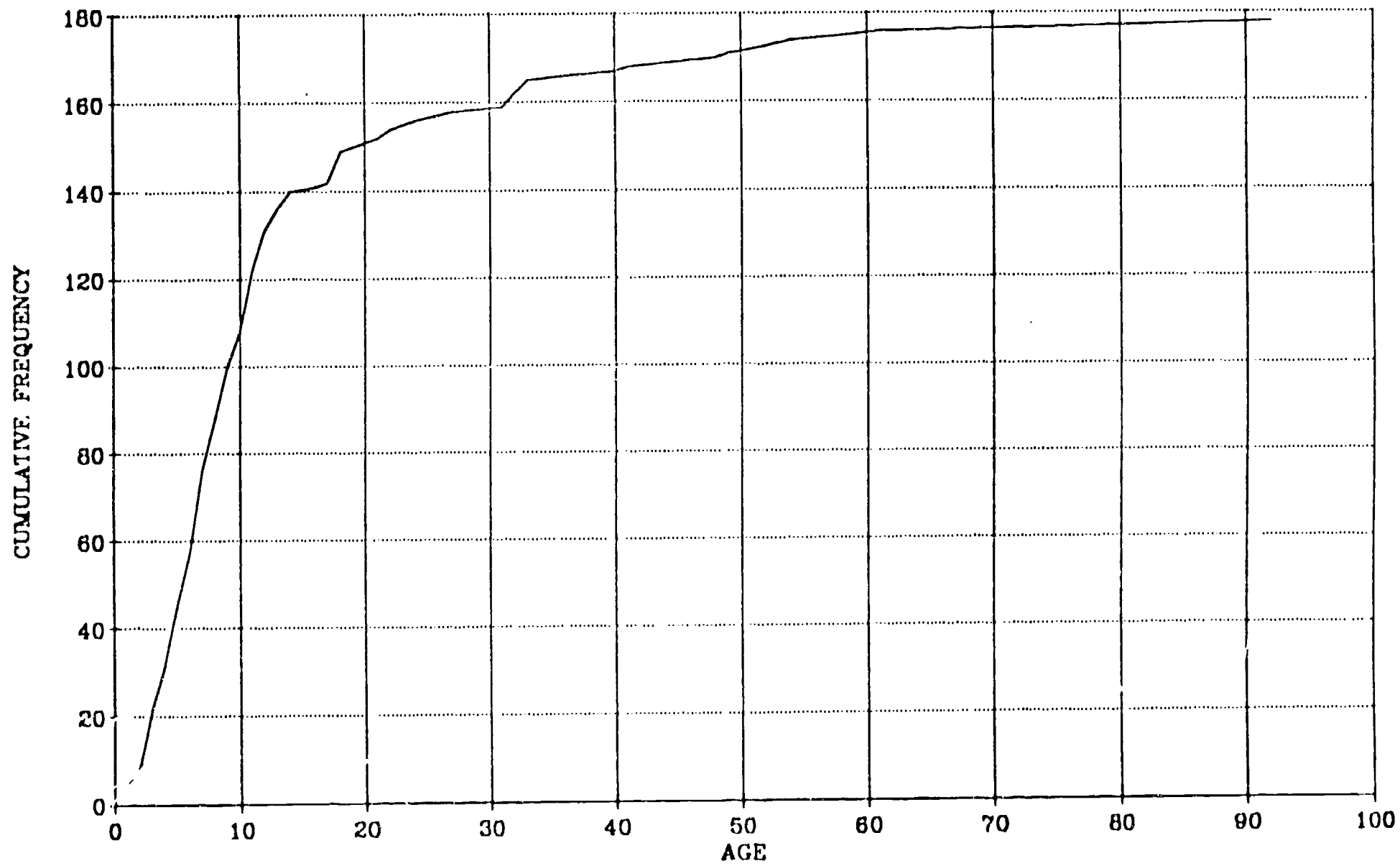


Figure 86

DEWEY 500s

LIBRARY 6

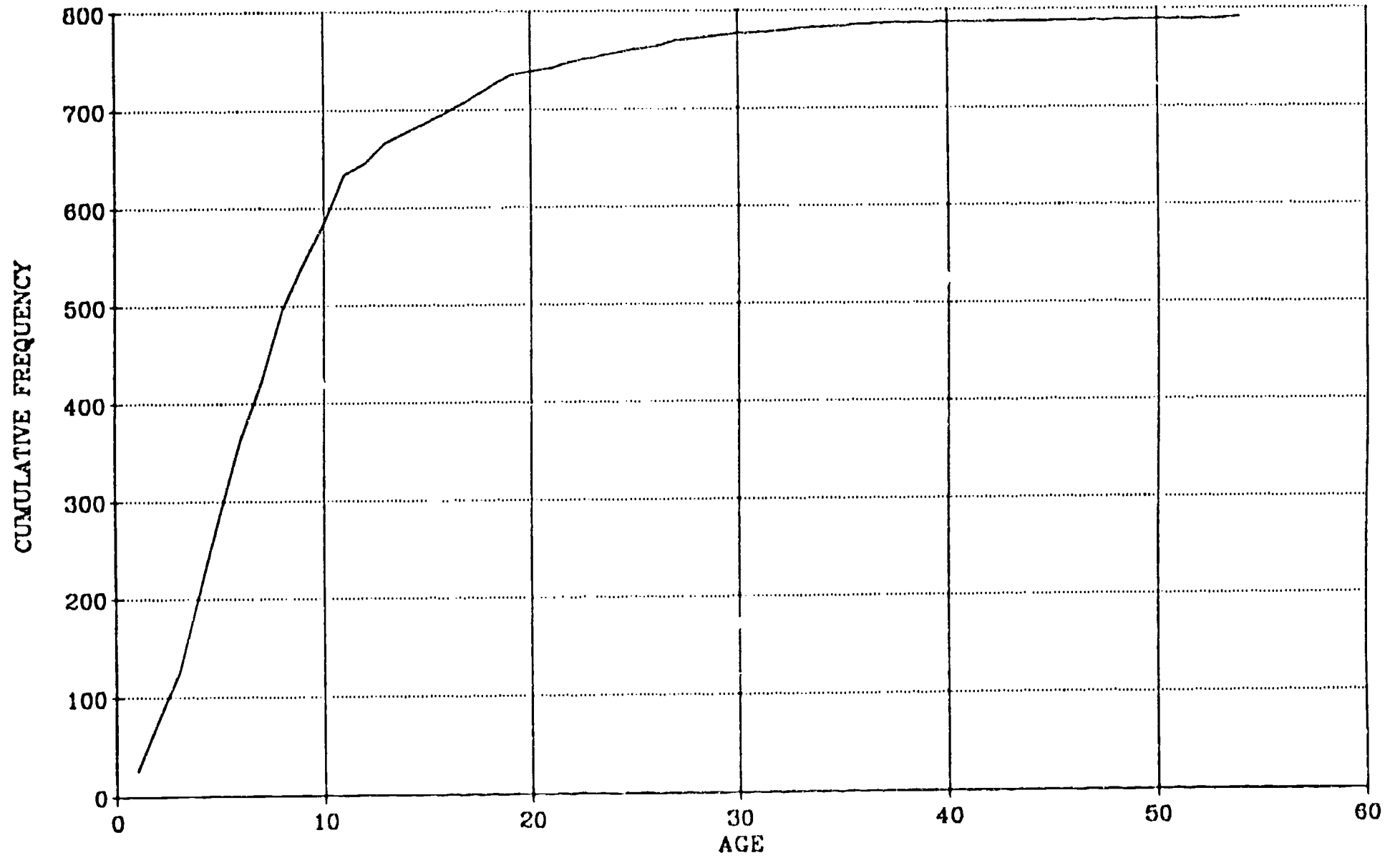


Figure 87

DEWEY 600s

LIBRARY 6

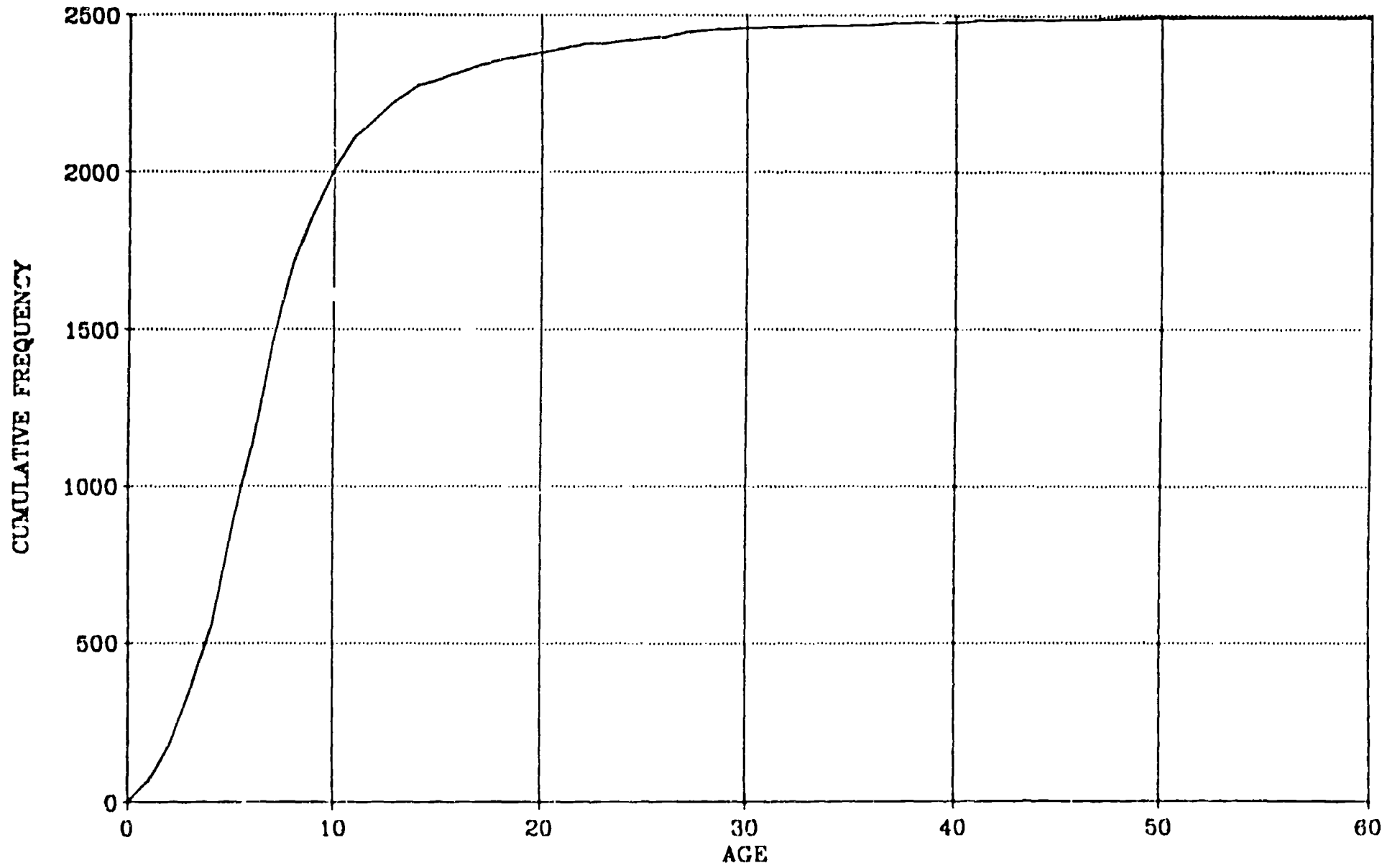


Figure 88

DEWEY 700s

LIBRARY 6

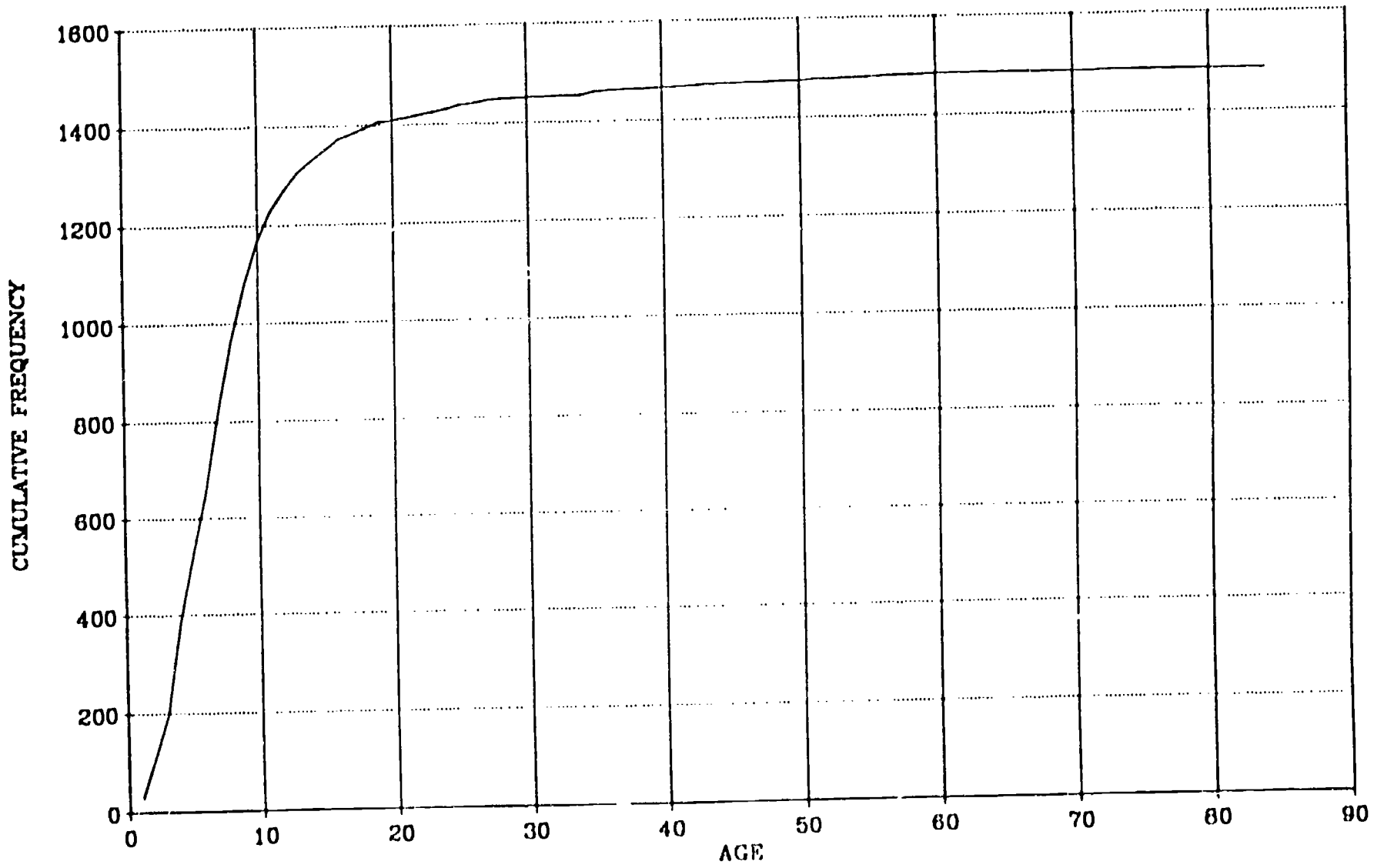


Figure 89

DEWEY 800s

LIBRARY 6

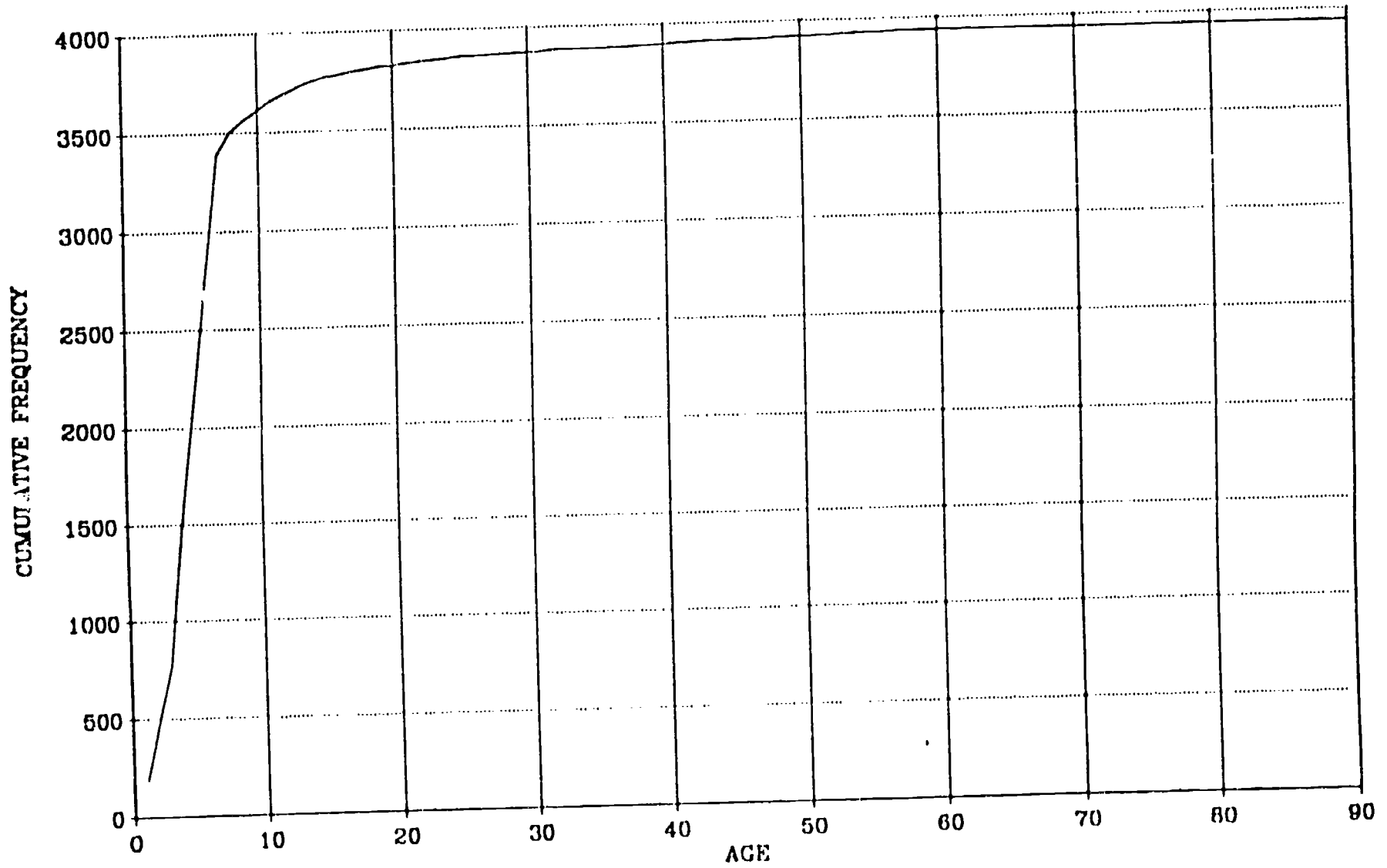


Figure 90

DEWEY 900s

LIBRARY 8

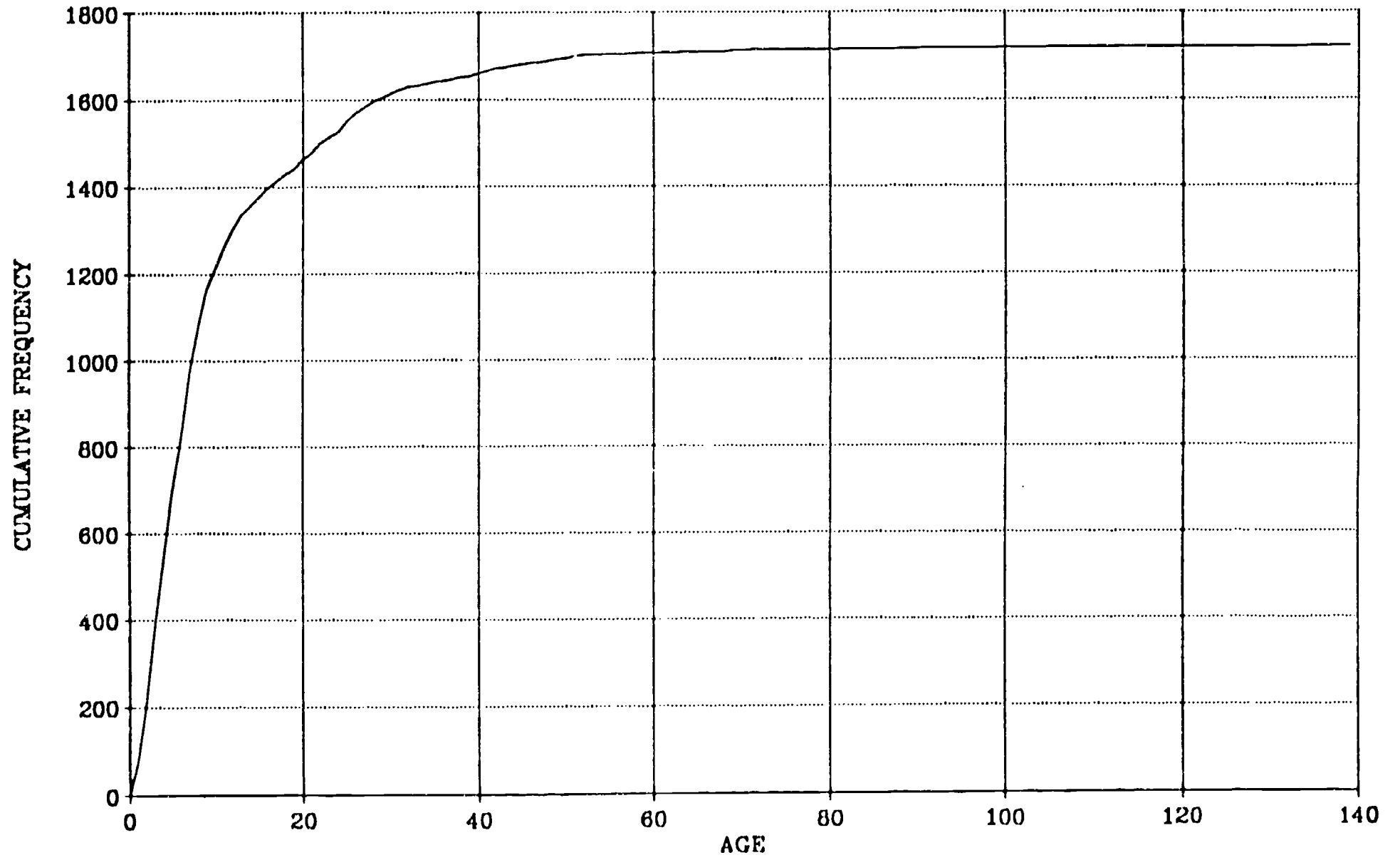


Figure 91

DEWEY 000s
LIBRARY 7

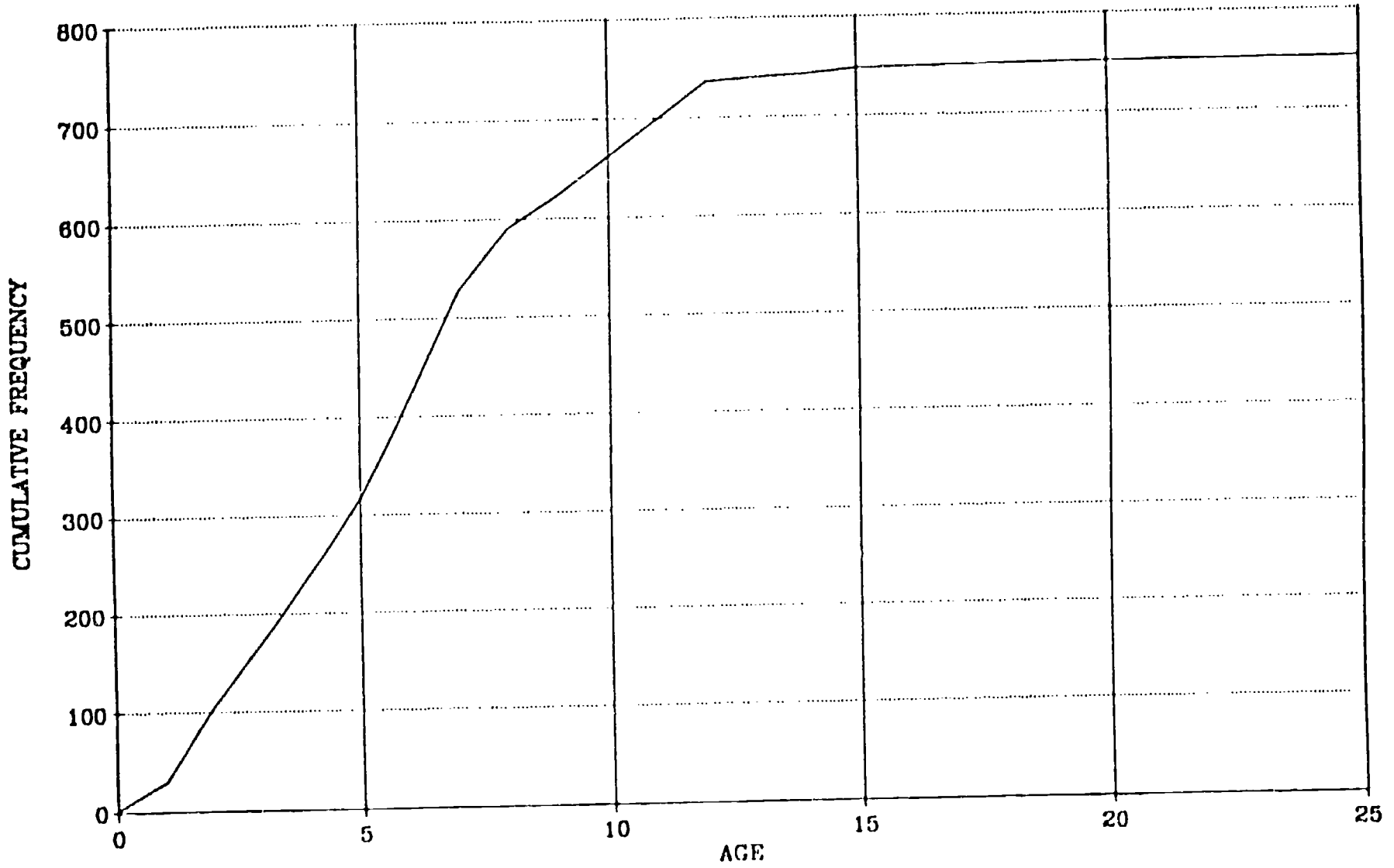


Figure 92

DEWEY 100s
LIBRARY 7

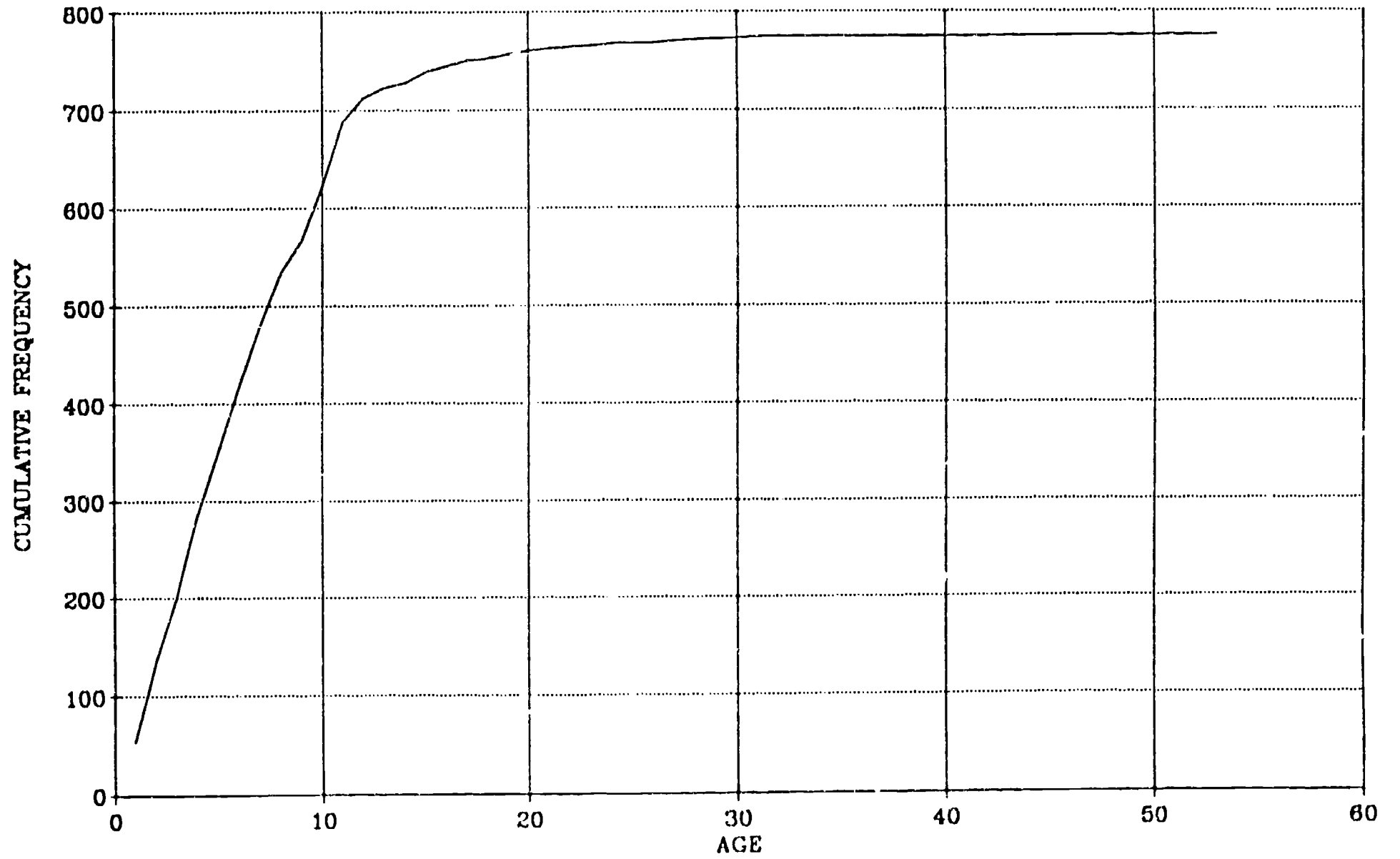


Figure 93

DEWEY 200s
LIBRARY 7

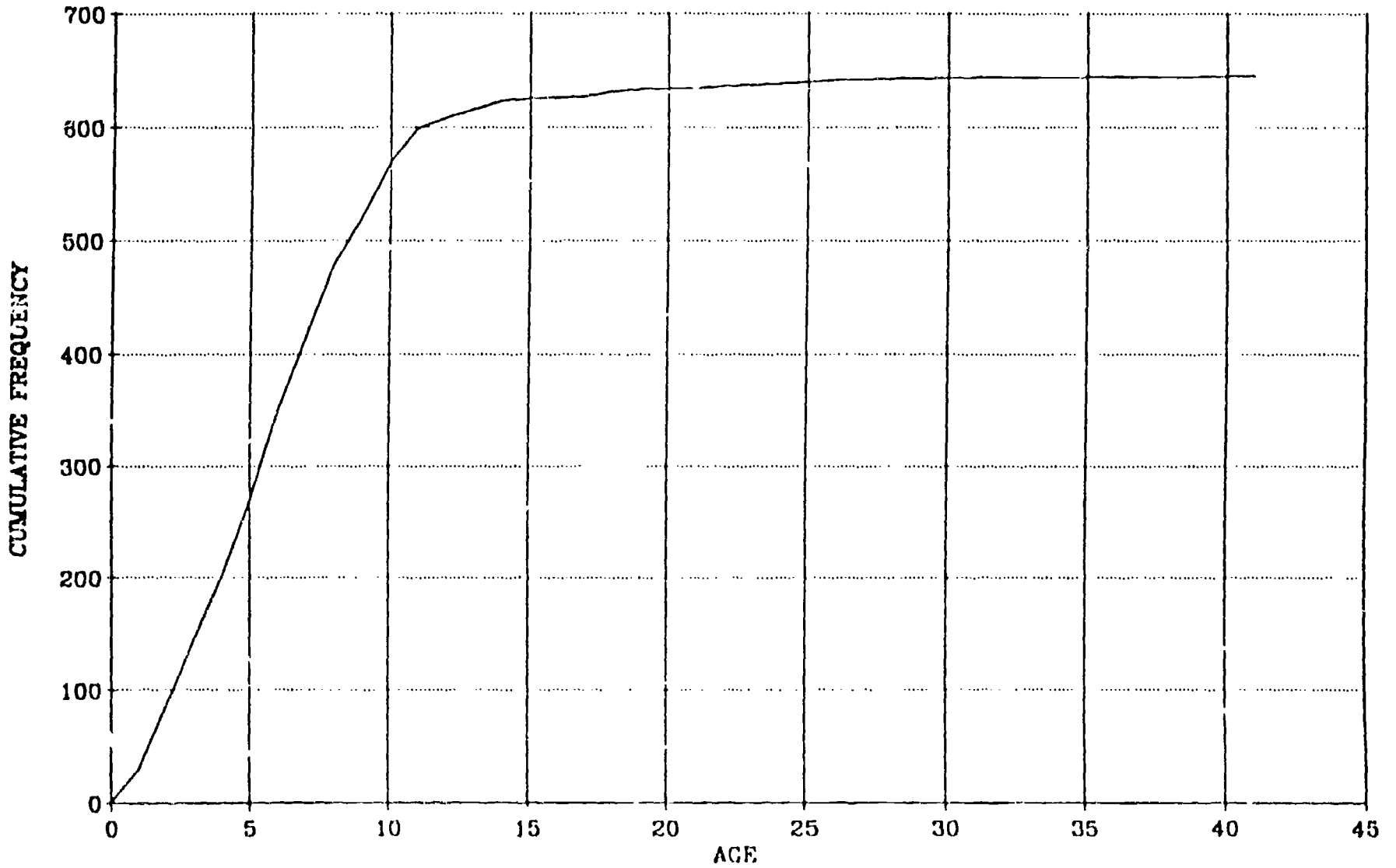


Figure 94

DEWEY 300s
LIBRARY 7

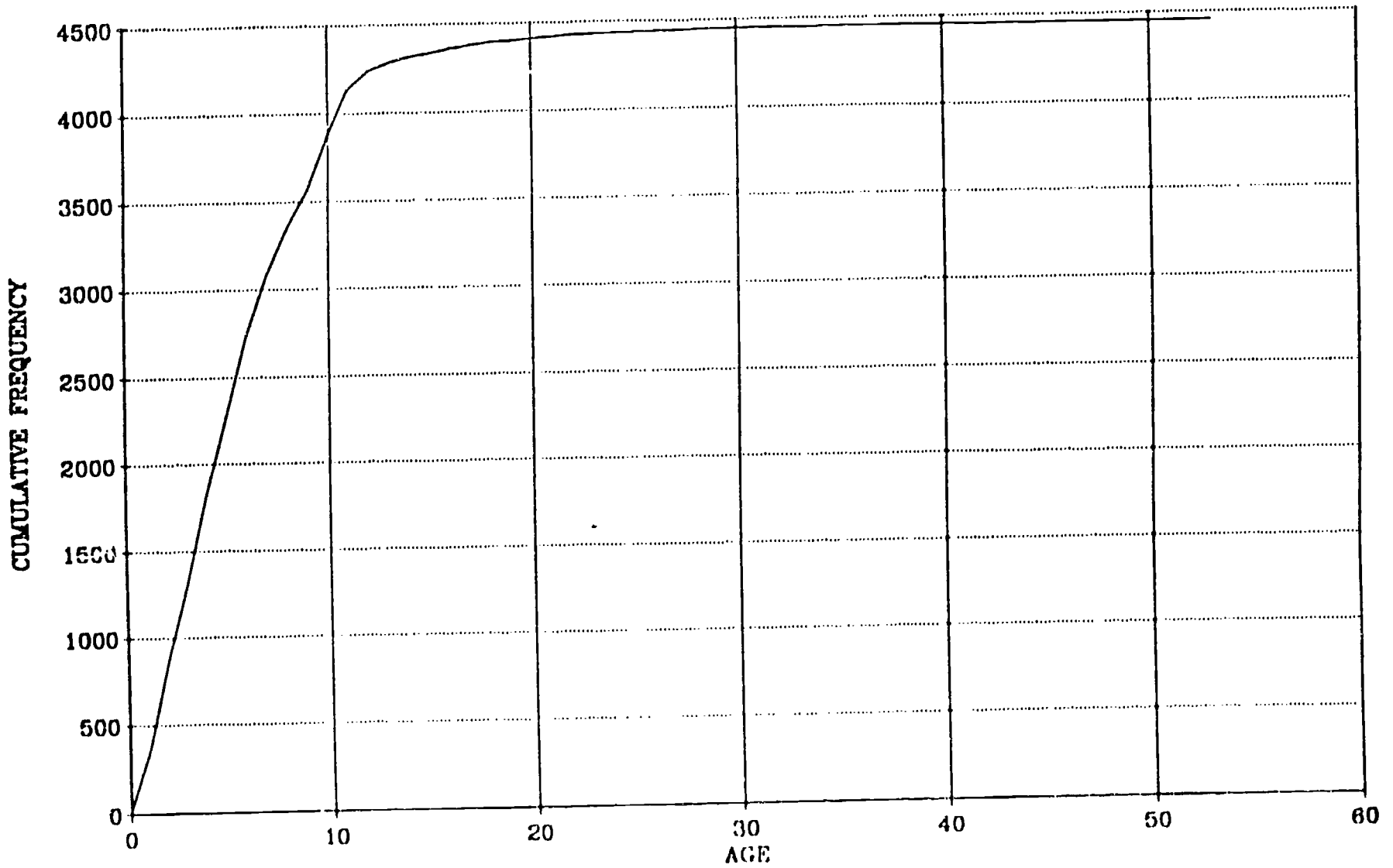


Figure 95

DEWEY 400s
LIBRARY 7

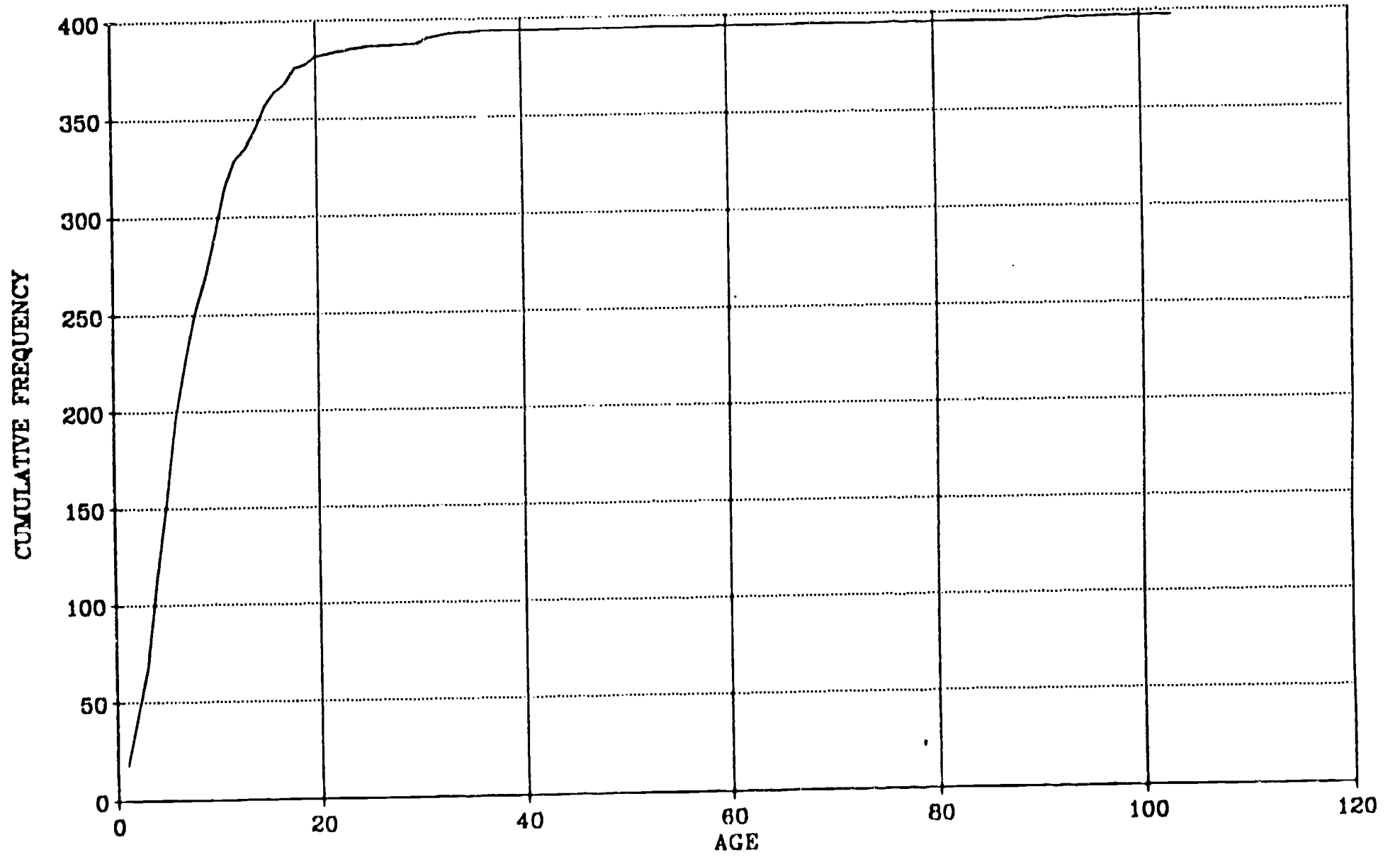


Figure 96

DEWEY 500s

LIBRARY 7

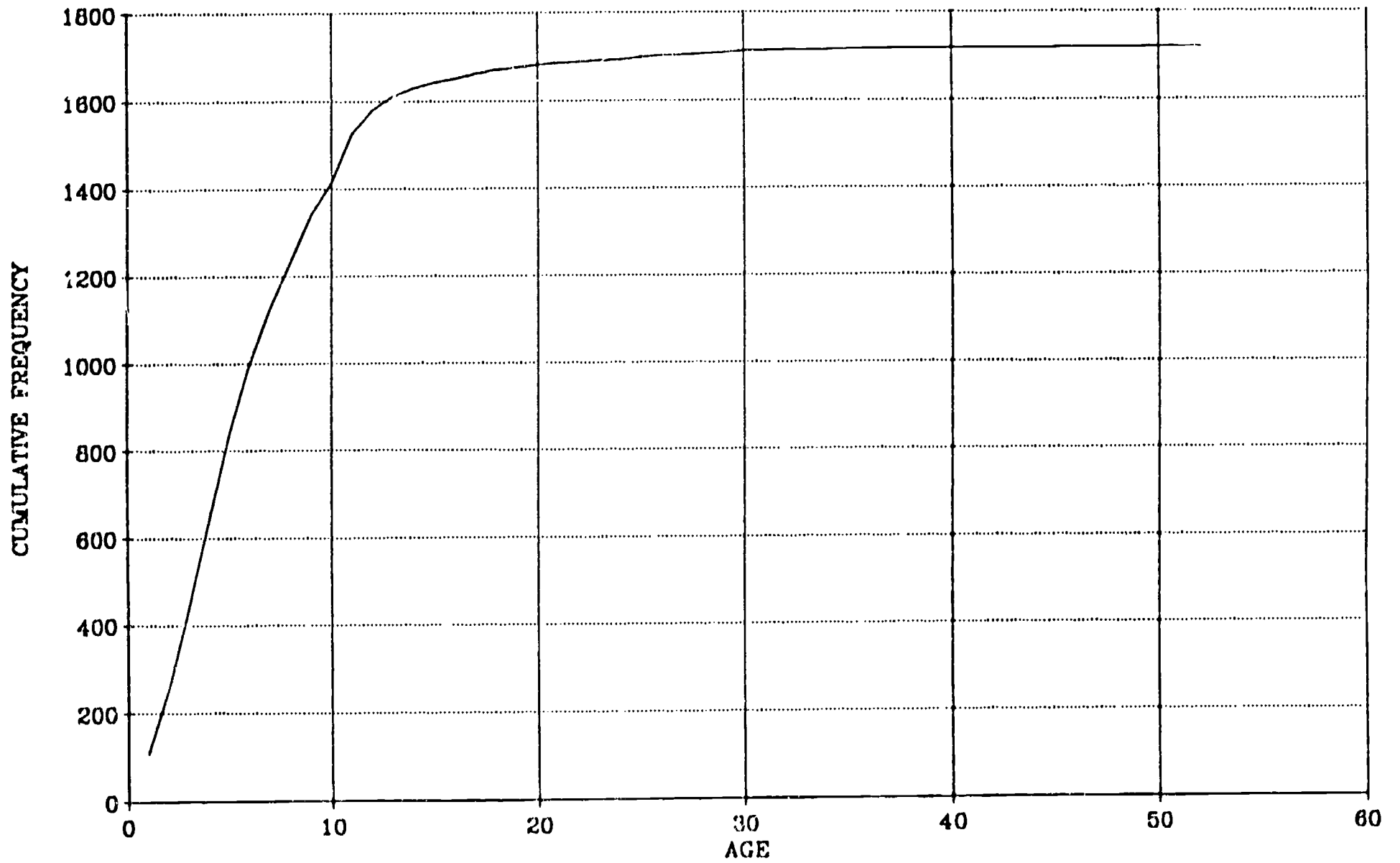


Figure 97

DEWEY 600s

LIBRARY 7

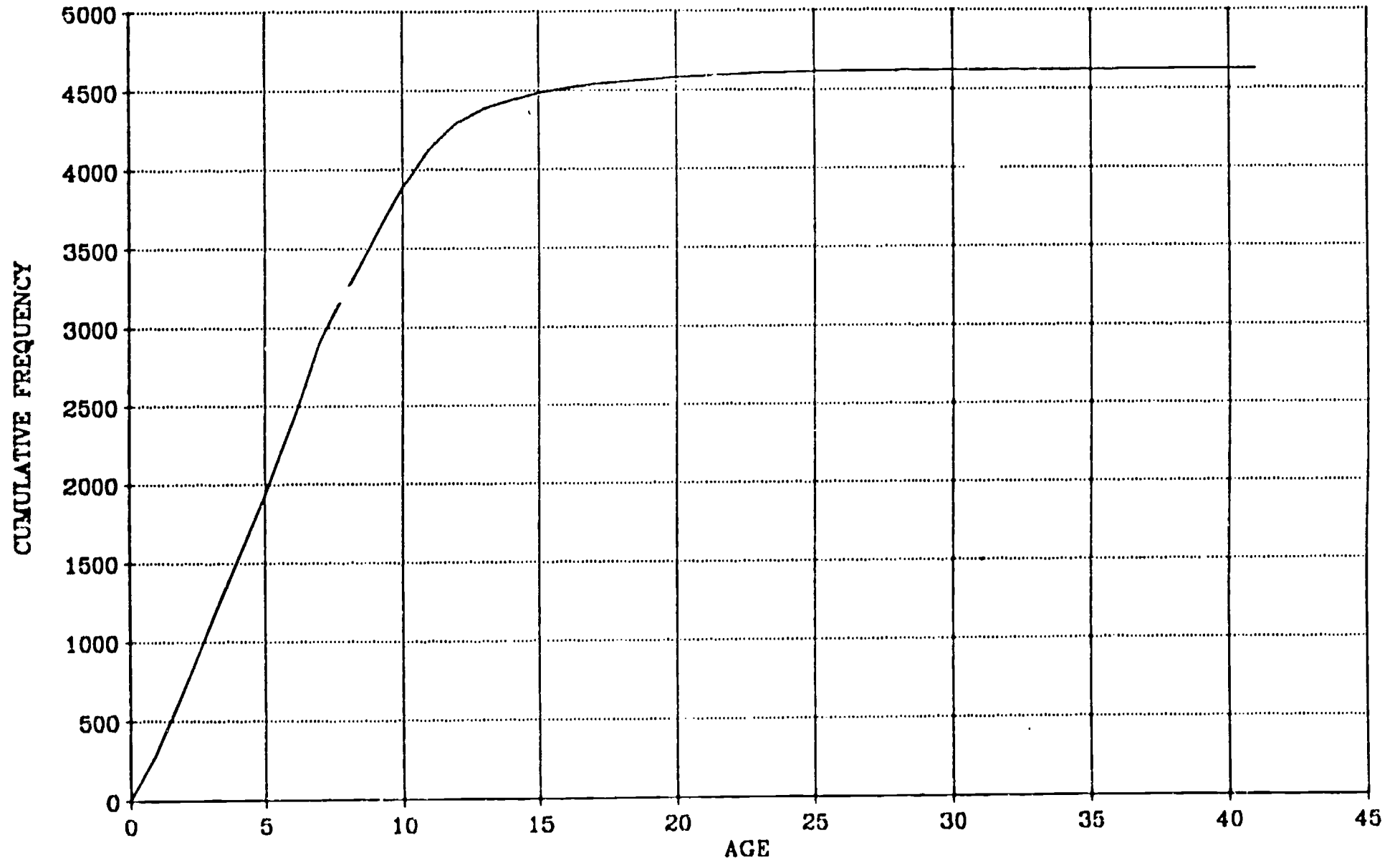


Figure 98

DEWEY 700s
LIBRARY 7

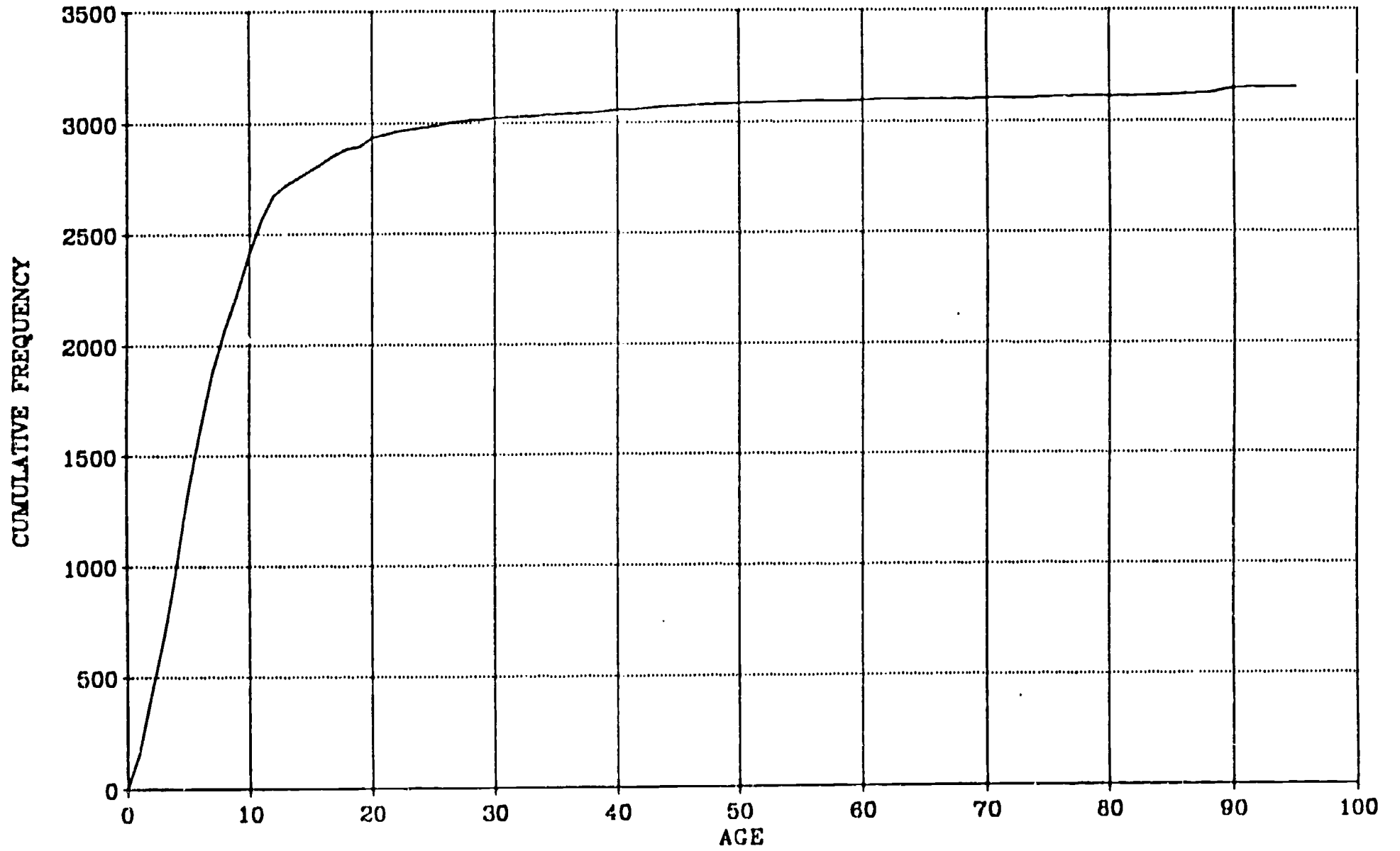


Figure 99

DEWEY 800s

LIBRARY 7

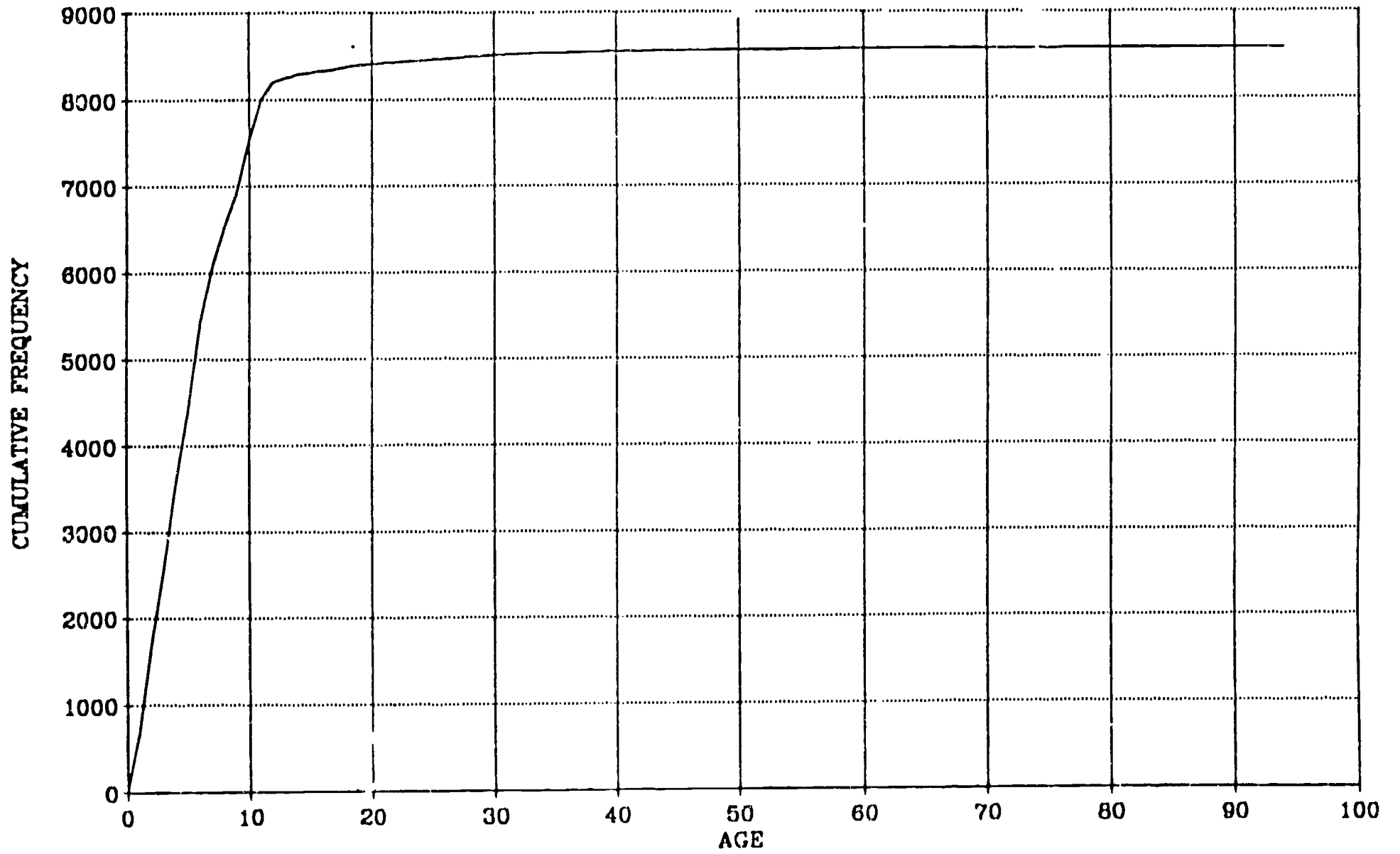


Figure 100

DEWEY 900s
LIBRARY 7

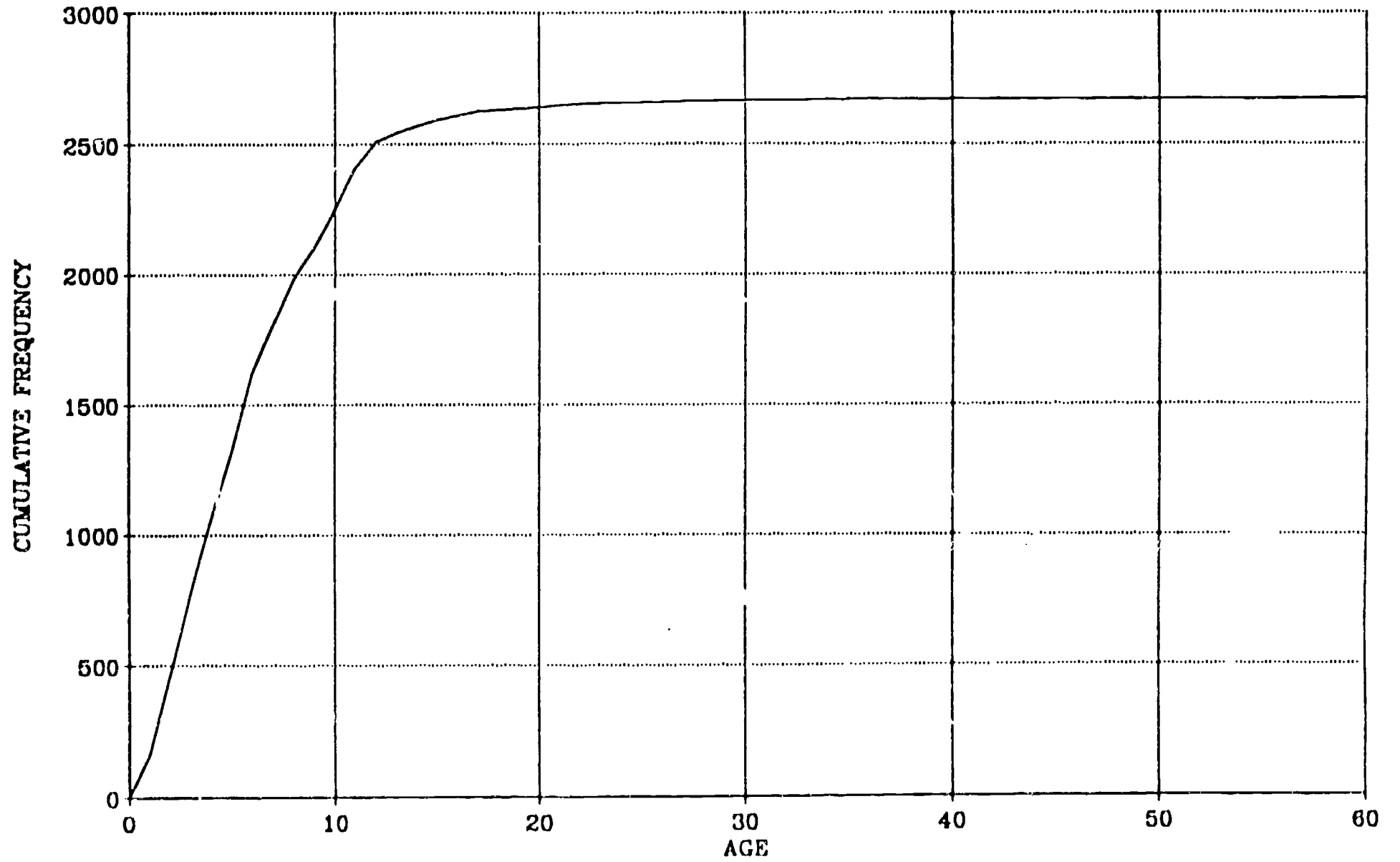


Figure 101

ALL LIBRARIES

NON--DEWEY

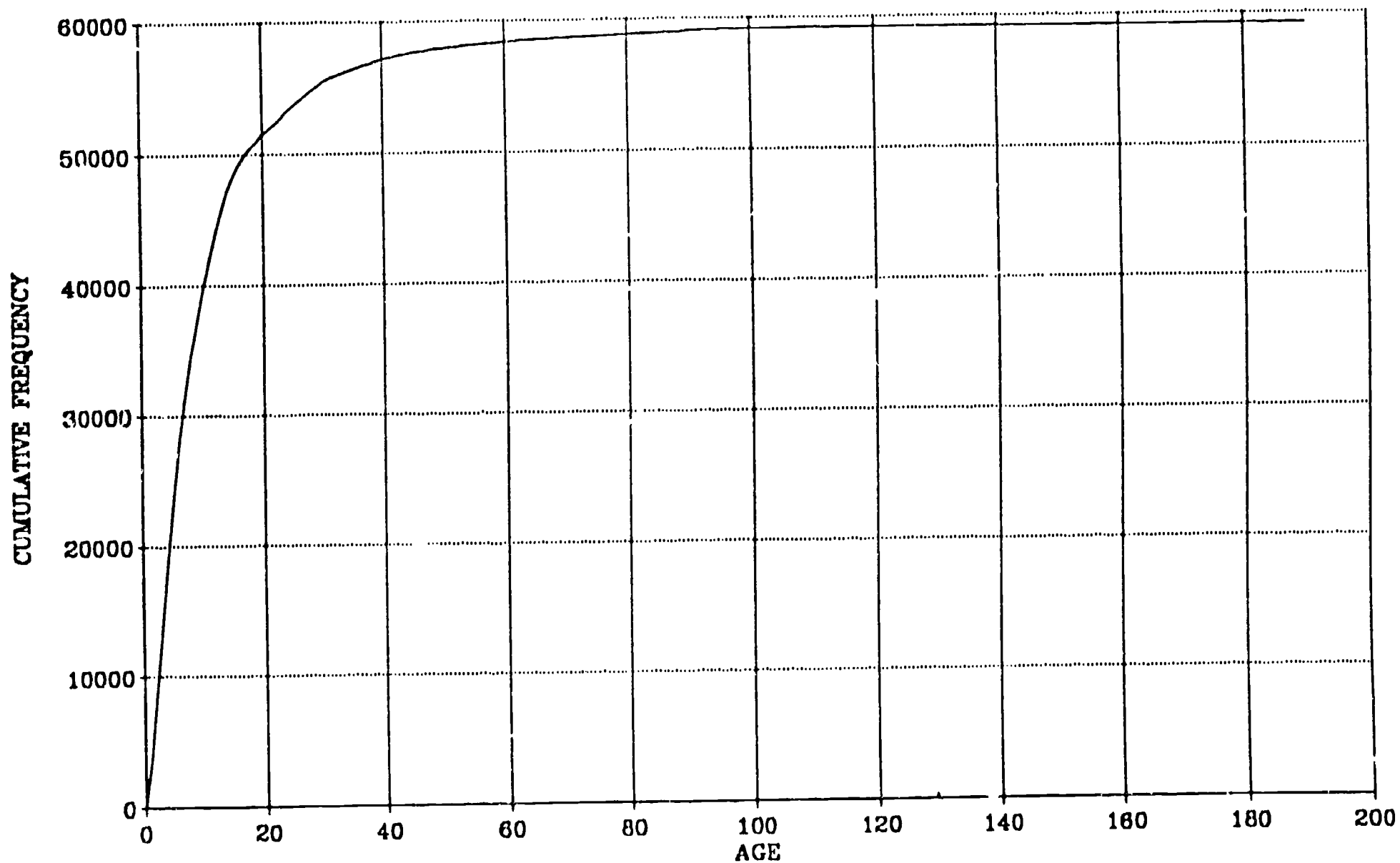


Figure 102

LIBRARY 1

NON-DEWEY

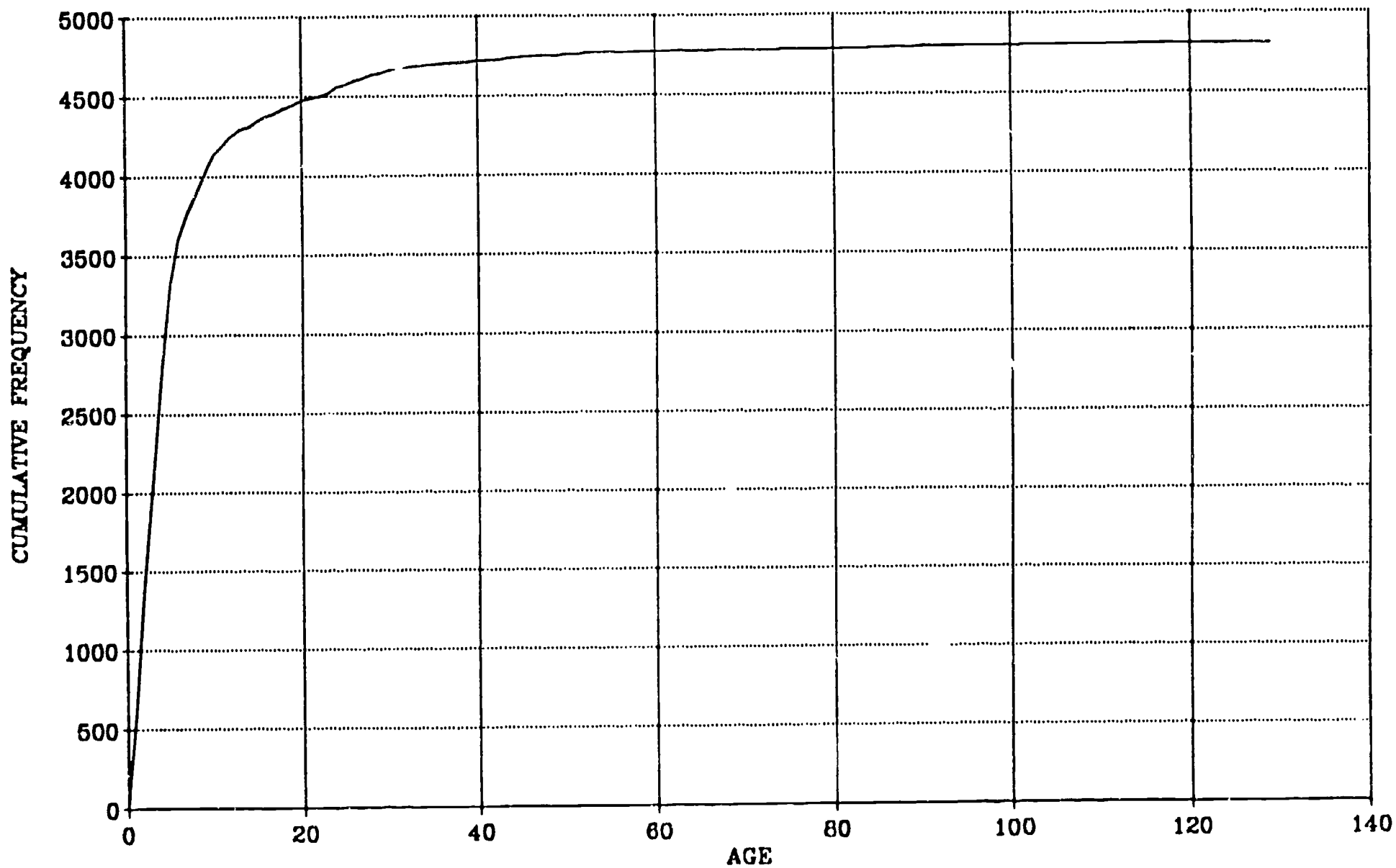


Figure 103

LIBRARY 2

NON-DEWEY

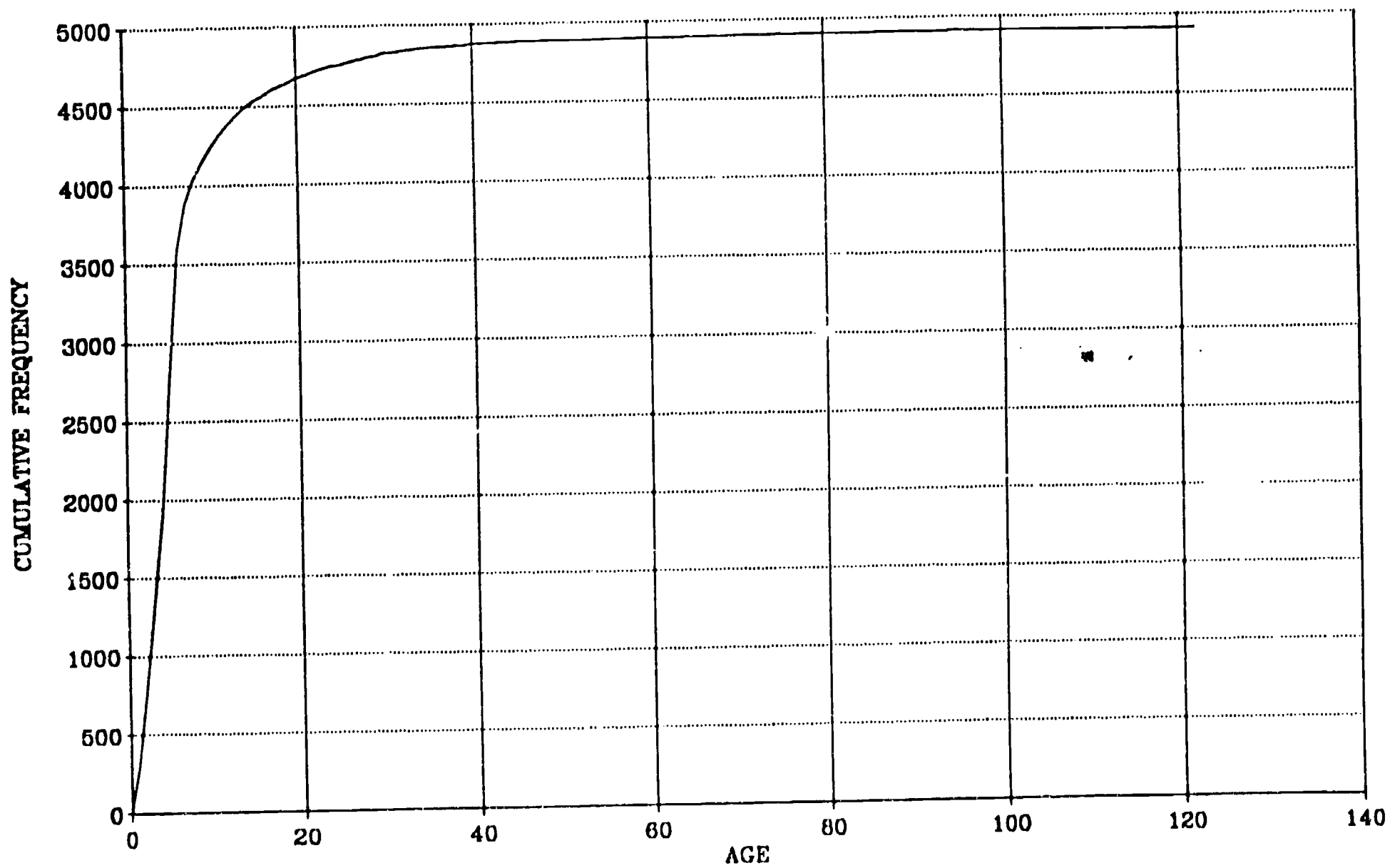


Figure 104

LIBRARY 3

NON-DEWEY

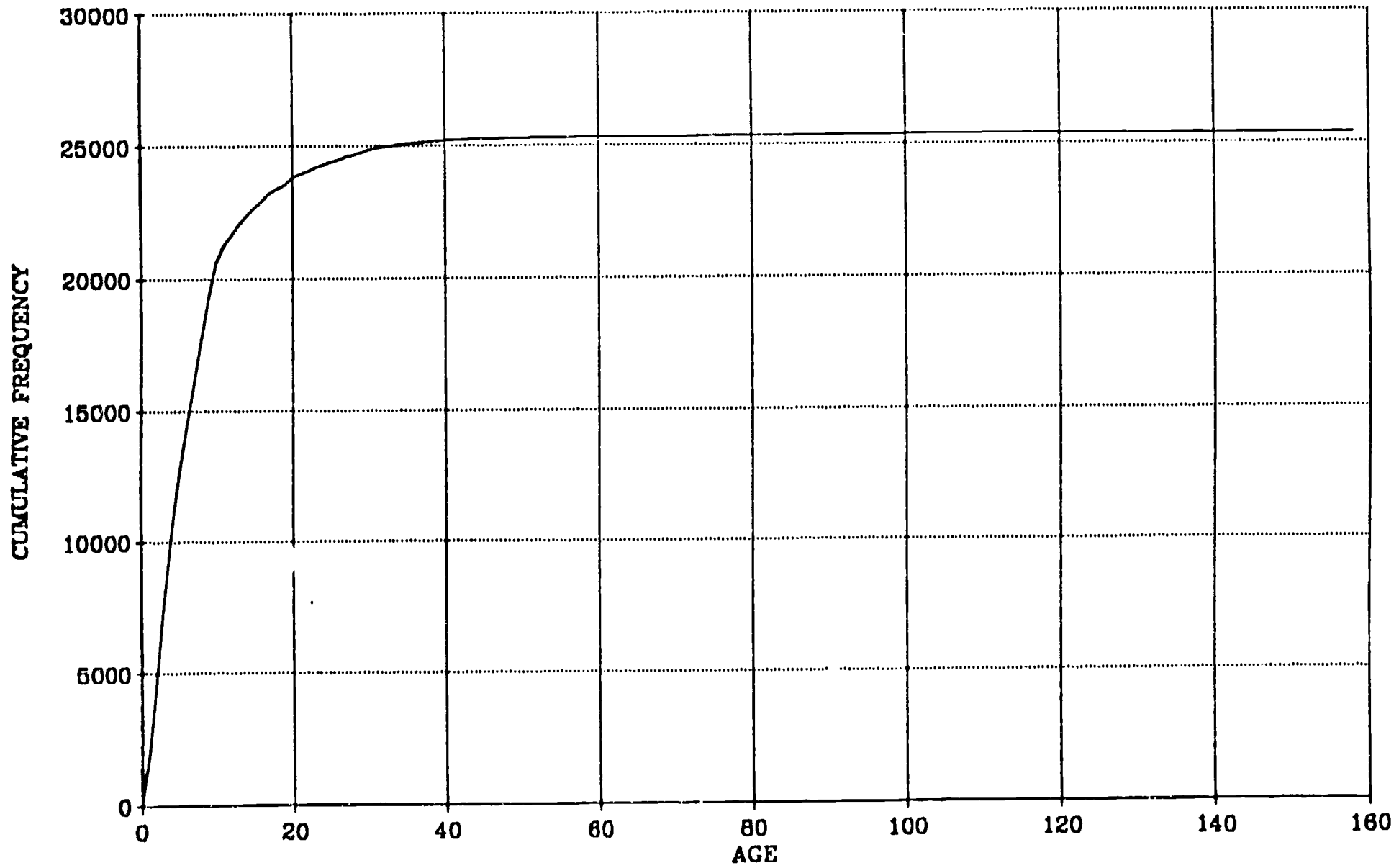


Figure 105

LIBRARY 4

NON-DEWEY

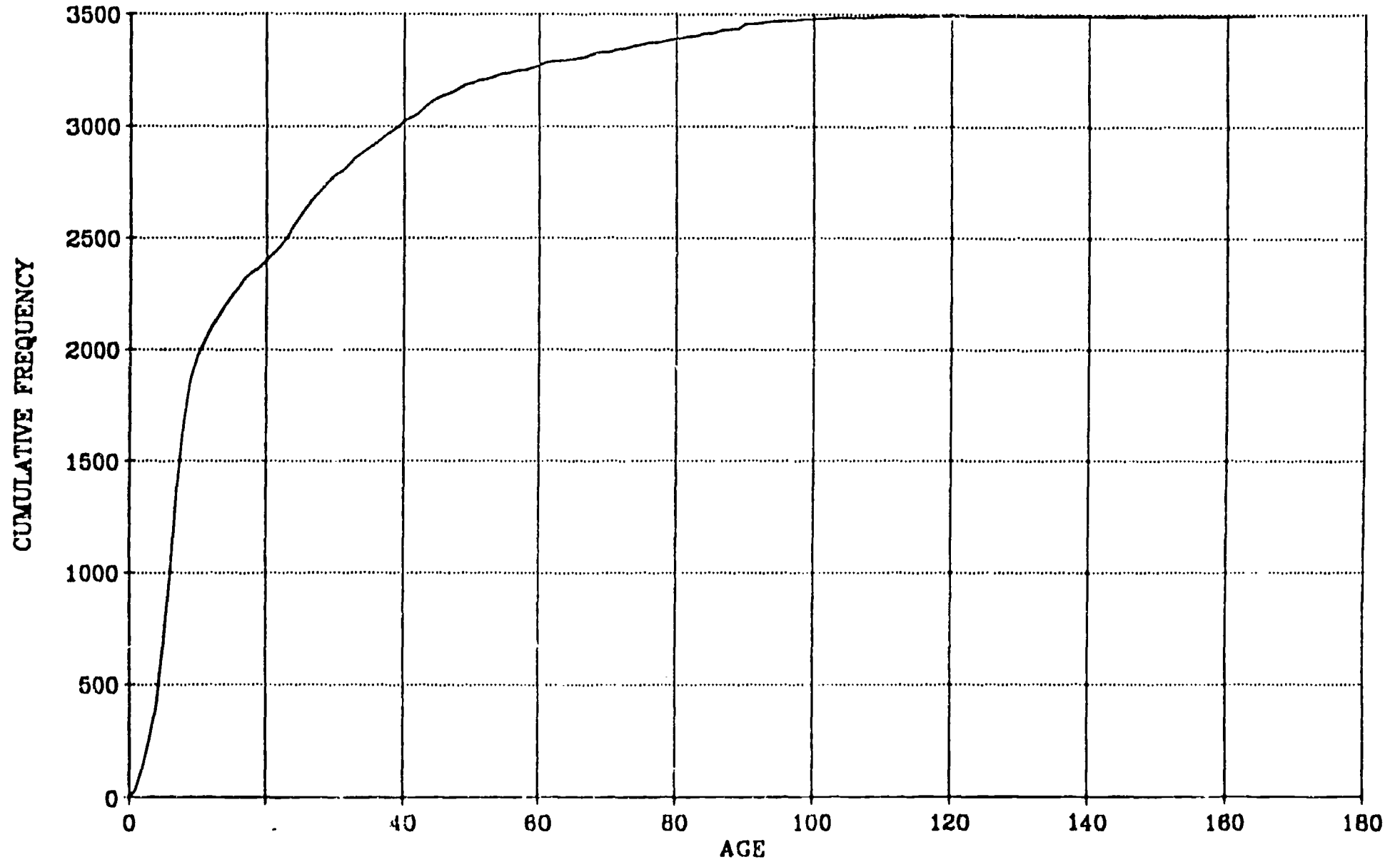


Figure 106

LIBRARY 5

NON-DEWEY

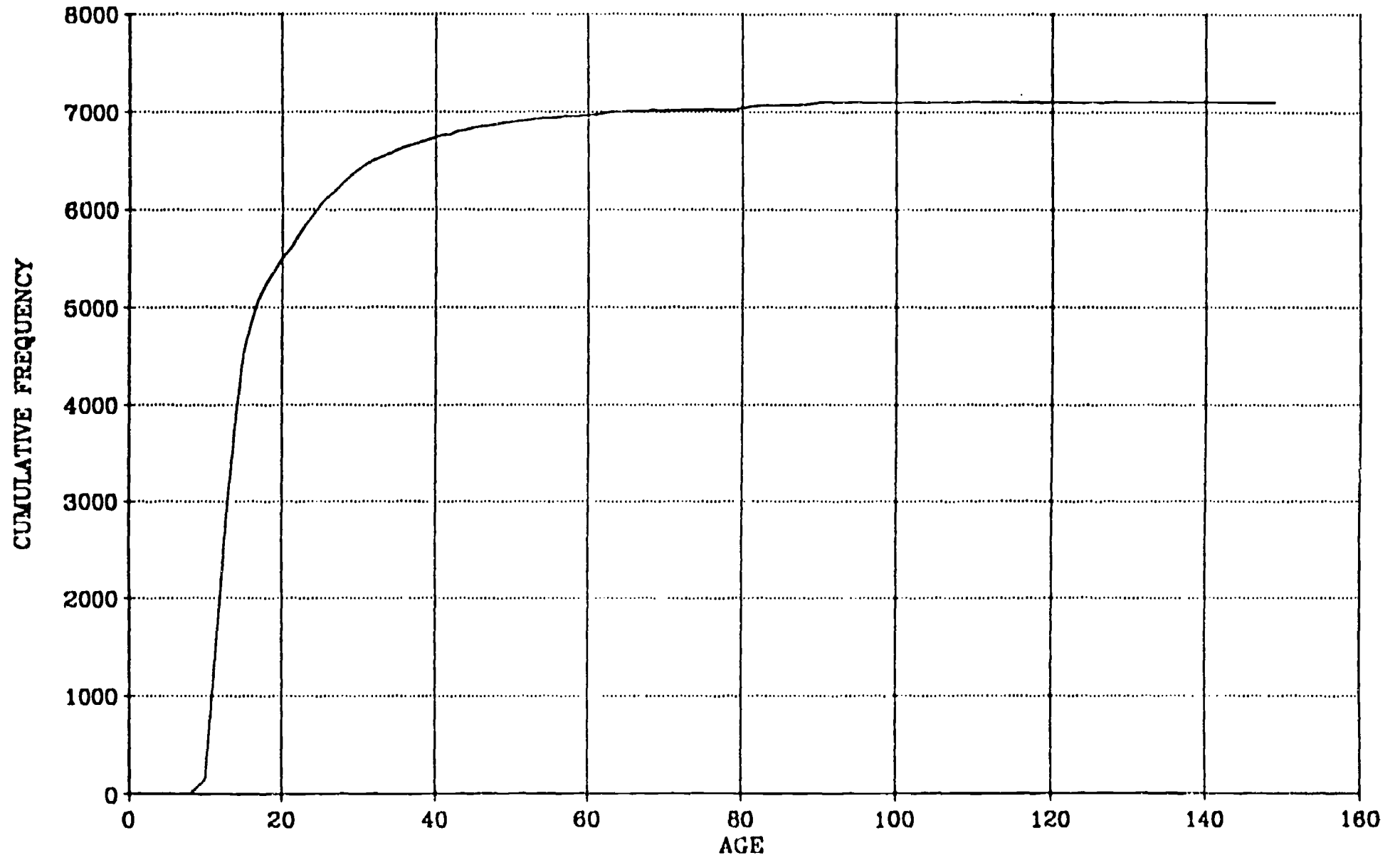


Figure 107

LIBRARY 6

NON-DEWEY

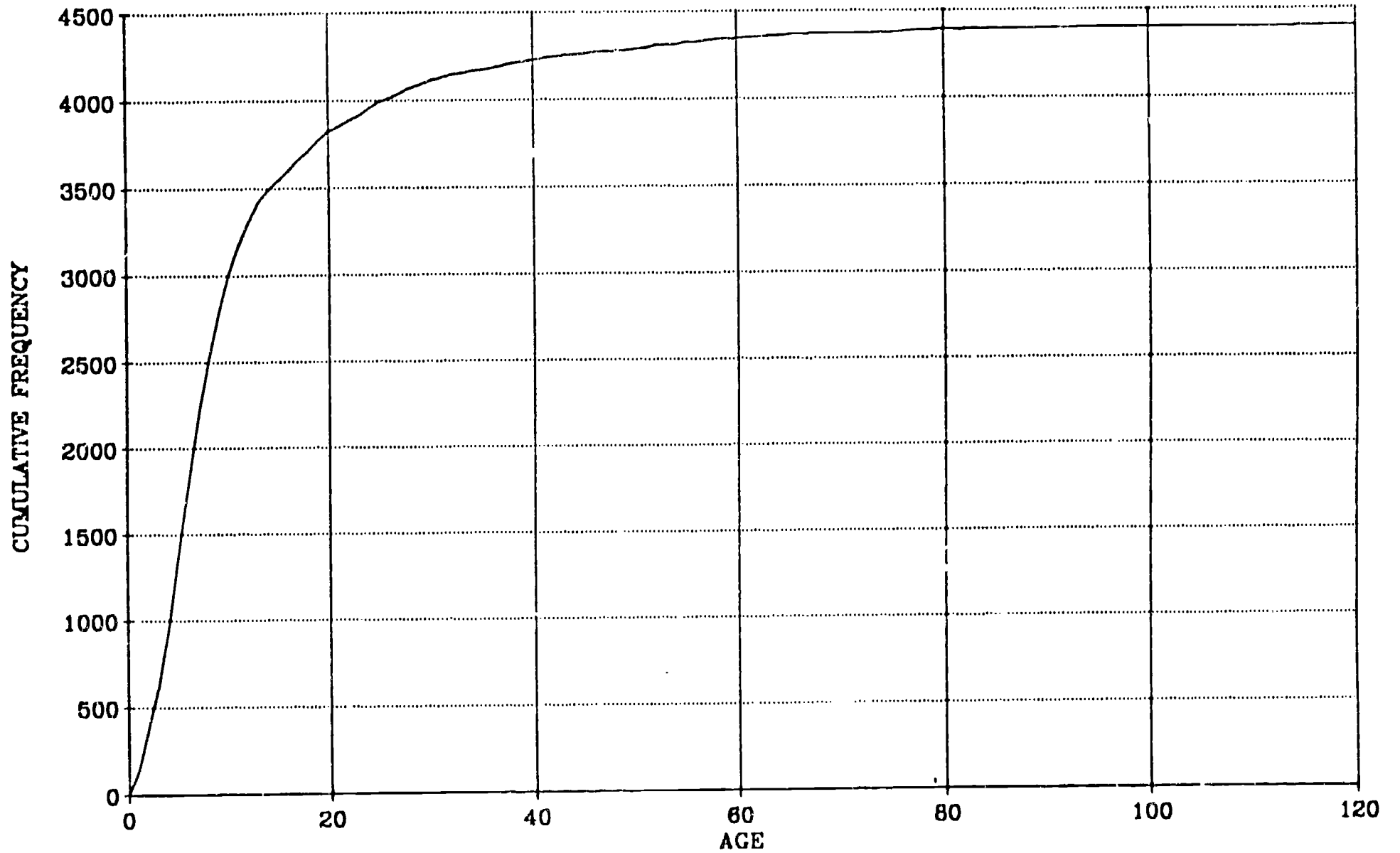


Figure 108

LIBRARY 7

NON-DEWEY

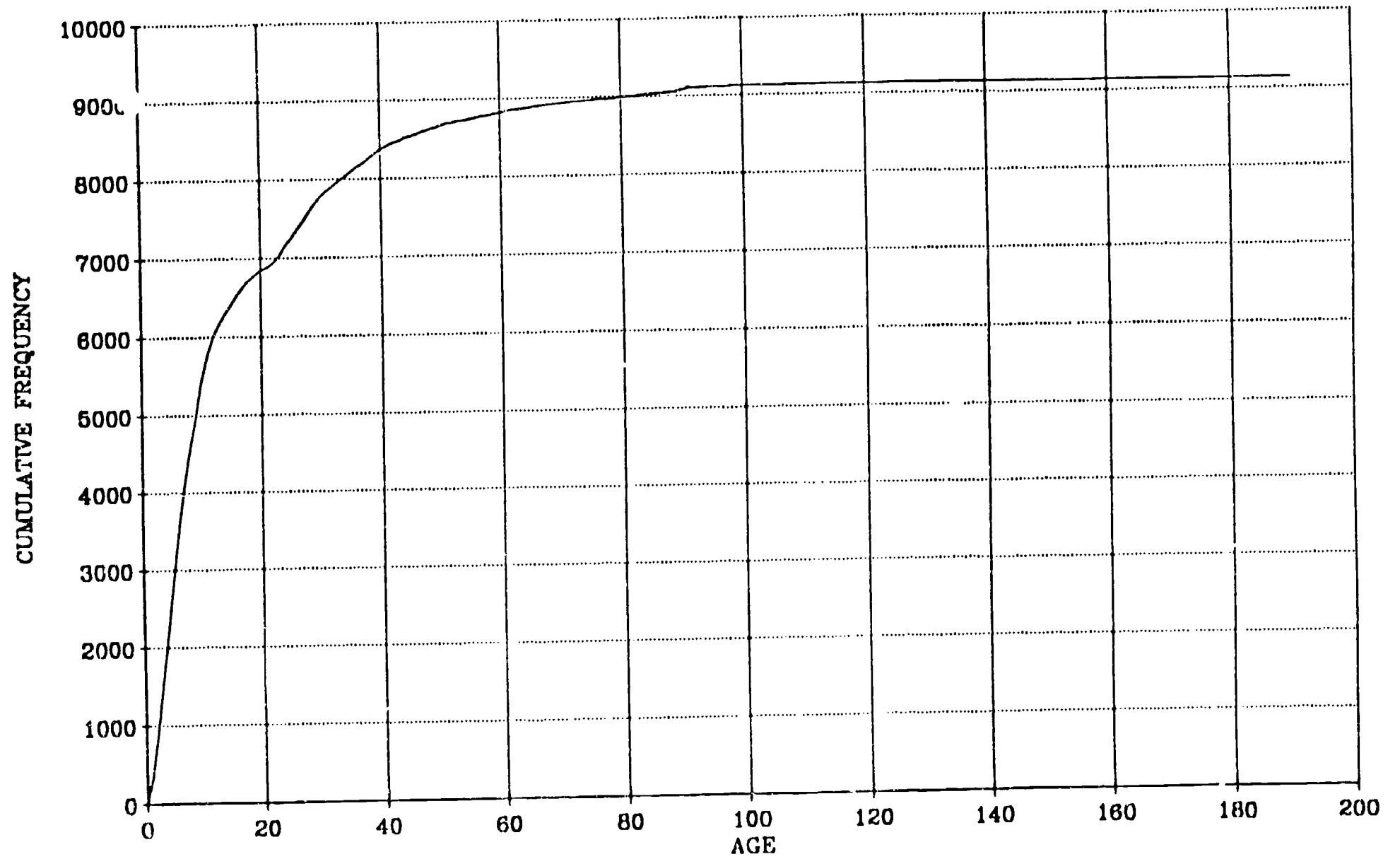


Figure 109
Age Distribution by Library

