

Indiana Library Statistics Project: A Report

Edward A. Stockey

Background

The Library Statistics Project, a joint effort of the Extension and Data Services Divisions of the Indiana State Library has two primary goals:

1. Reduce the costs involved in producing the *Annual Library Statistics Report*.
2. Provide improved analysis of the data which is being collected.

In achieving these goals it was realized that two separate but related tasks would have to be accomplished.

1. Generate the *Annual Report* from the data input.
2. Create a data base from which customized analyses and reports could be produced.

Because the project would require computer programs on a continuing basis, it was decided that rather than employ or contract for programming, a statistical applications package would be utilized.

Edward A. Stockey is Head of the Data Services Division at the Indiana State Library. He received a BA in Psychology from Indiana University, and did advanced studies in Classical Archeology at the University of Pennsylvania. He received his MLS from Drexel University. Prior to working at the State Library, he was manager of INCOLSA's Regional Processing Center.

Basically, a statistical analysis applications package is a set of computer programs which can be used by a non-programmer to analyze data, generate reports, etc., without the necessity of writing instructions in a higher level computer language such as FORTRAN, PL1, etc. Typically, once data has been input into the computer, very brief one or two line instructions will generate analyses and/or reports that might normally require hundreds or even thousands of programming instructions. For the Library Statistics Project, the Statistical Analysis System (SAS) applications package, available on the state's computers, was chosen.

SAS is very flexible, simple to use, and provides the following capabilities:

- * retrieval
- * transformations
- * manipulations
- * maintenance
- * report writing
- * printer graphics
- * data reduction and summarization
- * statistical analysis

SAS consists of a data-handling language and a library of procedures that work together as a system as is shown in Figure 1.¹

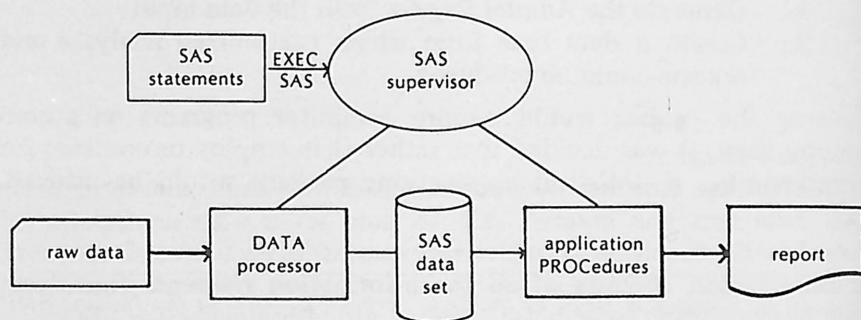


Figure 1

Using SAS, the Project consisted of three steps:

- * Data input
- * Creation of a library statistics data base
- * Report production

Data Input

Individual public library data for the year 1980 was keyed onto magnetic diskettes from each library's Annual Report Form. Individual library data for 1979, already in magnetic form (tape) was also input into the computer. Finally, summary statewide public library statistics for the years 1970-78 was keyed onto diskettes and input. Initial analysis of the 1980 data indicated several discrepancies and errors. Therefore, a Data Correction Form was produced. Basically, the Data Correction Form was a printout of a library's data as it was initially reported and input into the system. Libraries were asked to verify and/or correct the data and return it to the State Library. Upon receipt, the corrections were made to the original data. The philosophy behind the data input phase of the project was that the individual library is responsible for ensuring that the information submitted to the State Library is correct and accurate. The State Library, however, assumes responsibility for verifying that the data is correctly input into the computer and reported.

Data Base Creation

After the data has been input, a SAS data base containing three SAS data sets was created. A SAS data set is a sequential file of records with the same variables represented in each record. A record or observation consists of all the information elements (variables) such as circulation, telephone number, etc., for a particular library or in the case of statewide summary data - year. The data base creation process is shown in Figure 2.

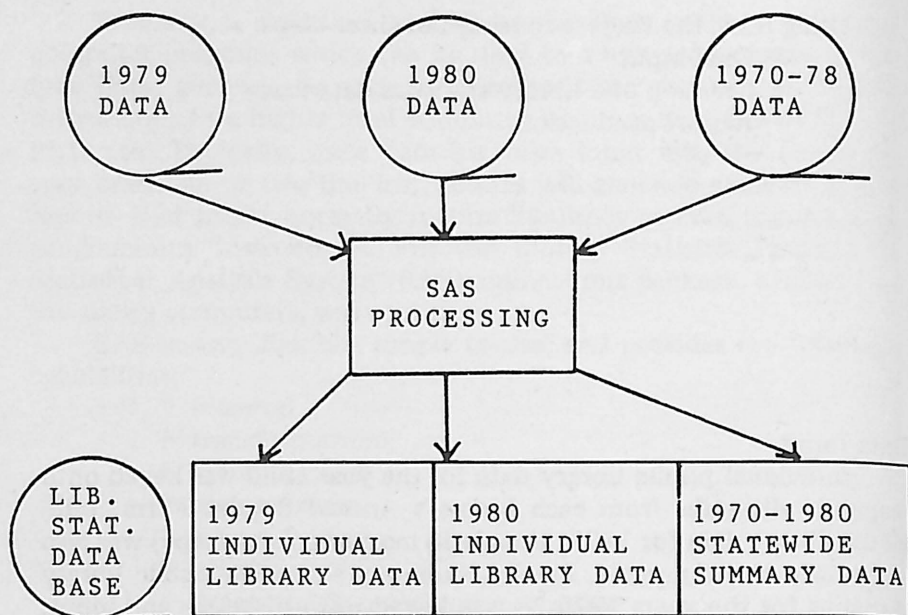


Figure 2

Report Production

Using the Library Statistics Data base, a computer print tape was produced. This tape was forwarded to a computer service bureau which utilized a Xerox 9700 non-impact printer (laser) to print the *Statistics of Indiana Libraries - 1980*. For 1981, the report forms themselves, as well as the Annual Report will be computer-generated. In addition to the Annual Report, a customized, individual report will be produced to assist libraries in their budget preparation efforts.

Conclusion

The use of a statistical analysis applications package, SAS, has provided a very efficient and cost effective method for the creation of a library data management and reporting system. As with all projects involving statistical analysis, manual and automated, the major problems involve data integrity and comparability. The State Library has created a statistics committee to assist in simplifying and standardizing data collection as well as specifying the types of analyses and/or reports that would be most helpful to a library's decision making process.

Notes

¹ SAS Institute, *SAS VIEWS*, Cary, North Carolina: SAS Institute, 1-2.