Government. Before the advent of the computer, the National Archives were concerned primarily either with the basic original records or documents obtained by the Federal agencies, or with the analytic or statistical end products. The problems of intermediate worksheets and data in semiprocessed form were left largely to the discretion of the individual agencies involved. Thus, for example, with respect to the corporate tax records of the Internal Revenue Service, the National Archives has preserved in warehouses bales of tax returns filed by corporations going back to 1909. In addition, National Archives has also preserved the statistical tabulations of tax returns. With the development of machine readable data, however, it is becoming increasingly obvious that bodies of information in machine readable form which are intermediate between the original records obtained by a Federal agency and the final statistical tabulations may be more worth preserving than the original records themselves. There is a growing recognition by the National Archives of this fact. The committee was very much impressed by the active interest which the staff of the National Archives showed in this problem. However, again the problem is so vast that it may require completely new procedures and policies in the future.

IMPACT OF THE COMPUTER ON DATA PROCESSING

Data processing methods have undergone a systematic evolution which has had far-reaching implications for the Federal statistics system ever since the original punchcard equipment was introduced. Early computers were to some degree a logical extension of this punchcard equipment. Although the Univac Model I pioneered by the Bureau of the Census in the early 1950's represented a monumental step forward, it was only the modest beginning of what has turned out to be a completely new technology. Each succeeding generation of computers incorporates improvements in the size of memory, the speed of computation, and the density of data storage on tape such that the capacity and speed of operate have been increased many times over. By now the technological revolution has become so great that a reexamination of the organization of the Federal statistical system is urgently needed.

Increase in efficiency

From the outset, the computer, like other forms of automation, has reduced the amount of labor required in the processing of data. Before their introduction, a large organization of clerks and punchcard machine operators was needed to handle the huge volume of punchcards required for any substantial statistical operation. Sorting tabulating, and computing were relatively lengthy processes. Even for minimal tabulations a great many steps were requested. It is true, of course, that the computer has made necessary the development of specialists who could write programs for data processing, but once a program is written and proved out, it can be used to process large masses of information rapidly and with a small staff.

Reduction in processing time

Equally important, the time required for data processing has also been substantially reduced. Operations which formerly took 7 to 8 months to carry out now have been reduced to a matter of weeks. In the processing of the 1960 population census, the time required for certain steps was reduced from several years to several months. This shortening of time has not only meant an increase in efficiency in terms of overhead and other fixed elements in the program, but it has also resulted in making important information available more promptly. This reduction of the timelag between the collection of information and its availability greatly affects the usefulness of the information.

Improvement in data quality

The computer has also made possible new kinds of analysis which could not have been done before because of the cost and time required to carry out the necessary computations. First, it has become possible to examine and edit much more carefully than was possible heretofore. Computers can "wash" the information, and find inconsistencies which would have gone unnoticed in hand editing. Editing instructions to test the reasonableness of the basic information can be built into the processing programs. Thus, in the case of census data for manufacturing establishments, the computer can spot errors in reporting wage bills and manhours by computing average hourly earnings. Where the resulting figures are outside a reasonable range, the original information can be questioned. Other kinds of inconsistencies can be tested in a similar way, and for