

CHAPTER VI. PROCESSING THE DATA FOR PUERTO RICO

BACKGROUND

The 1960 Censuses of Population and Housing for Puerto Rico were carried out as a joint effort of the Bureau of the Census and the Commonwealth of Puerto Rico Planning Board, with local personnel and facilities in Puerto Rico used for most of the work. The content of the censuses was adapted somewhat to fit the special needs and circumstances of the area.

The clerical and machine processing of the data collected in the 1960 censuses of Puerto Rico was carried out in part in Washington and Jeffersonville and mainly in a temporary processing office established for the purpose in the city of San Juan. This was the first time that the census of Puerto Rico was processed in Puerto Rico and largely by Puerto Rican employees.¹

The Puerto Rico Office maintained continuous direct working relations with the Washington headquarters of the Bureau regarding processing procedures, administration and personnel management, time schedules, production controls and quality control, coding and editing problems, and other matters concerning the program.

SPACE, EQUIPMENT, AND SUPPLIES

The Puerto Rico Decennial Operations Office was opened in April 1960. Under the terms of the agreement between the Bureau of the Census and the Puerto Rico Planning Board, it was the responsibility of the Planning Board to provide space for the processing office. Two floors in a newly constructed commercial building, comprising about 13,000 square feet of space, were made available for this purpose. Although this space was highly satisfactory, various building modifications had to be made before the space could be used for the data-processing operations. The existing electrical wiring and power supply in the building had to be augmented to accommodate the electrical power requirements of the IBM machine installation. Shelving for the filing of enumeration books, IBM punchcards, and other processing materials had to be constructed. Additional overhead lighting to provide adequate illumination was necessary, and various other minor alterations were made.

It was found to be cheaper to acquire the furniture and office equipment from surplus stocks of the Census Bureau and other Federal agencies and pay the cost of transportation to Puerto Rico.

¹ A detailed account of the data processing in Puerto Rico is to appear in a separate publication.

The Bureau rented the following IBM equipment:

<u>Equipment</u>	<u>Quantity</u>
Numerical Punch No. 024.....	20
Punch Card Verifier No. 056.....	15
Sorter No. 083.....	2
Collator No. 077.....	1
Electronic Statistical Machine No. 101	4
Duplicating Summary Punch No. 524....	2
Alphabetic Accounting Machine No. 402	1

PERSONNEL

Recruiting and appointment of personnel to positions in the Puerto Rico Office were carried out in accordance with regular Civil Service Commission competitive standards. All persons hired were given temporary appointments for a period of not to exceed 1 year. The average length of employment was about 9 months. An experienced statistician and a machine tabulation supervisor were detailed from the Bureau of the Census in Washington to serve as Chief of the Operations Office and Chief of the Machine Tabulation Section, respectively. With the exception of these two persons, all other positions in the Operations Office were filled by local Puerto Ricans. At various times, as special needs developed, specialists from Washington were sent to Puerto Rico to handle specific problem areas.

Employment varied from 30 to 87 persons during the period of data-processing activity. (See appendix G, table 2).

The Census Bureau used a special pay plan for employees of the Puerto Rico Office. The salaries paid, while considerably lower than those established for Federal employees at the same levels by the Classification Act, were comparable to those paid by private industry in Puerto Rico and by the Puerto Rican government. In general, this proved satisfactory. Recruiting efforts resulted in a plentiful supply of qualified applicants, and attrition rates were low. However, at some of the higher supervisory grades, particularly those requiring machine processing experience, the lower pay scale was an obstacle to recruiting experienced personnel, and also resulted in the loss of several key supervisors who accepted higher paying employment elsewhere after receiving considerable training.

From the point of view of intelligence, education, and willingness, the caliber of personnel hired was extremely high. However, very few persons with experience in the statistical clerical area or with machine processing backgrounds were available. As a result, an intensive training period was required.

CENSUS SCHEDULES

In the enumeration of population and housing in Puerto Rico, every fourth housing unit (and every fourth person in group quarters) was designated a sample unit for

which the enumerator asked a series of supplementary questions in addition to the basic census items. The enumerator collected and recorded the sample data in the same interview in which he obtained the basic census items; this procedure differed from the two-stage enumeration procedure used in most of the United States.

There were two variations of the schedule containing the basic items. The one used in most of Puerto Rico was form 6OPH-1PR. The other, used in the cities of San Juan, Mayaguez, and Ponce, for which data were to be published by city block, was form 6OPH-2PR. The two forms differed only in the inclusion of questions on rent or value of housing unit on the form used in the three cities. The schedule used for sample data (and also for 100-percent data for the 25-percent sample of population and housing units) was form 6OPH-3PR. Some additional, supplemental forms were used for the collection of data, but the information from them was transcribed to form 6OPH-1PR or form 6OPH-2PR. The 6OPH-1PR and 6OPH-2PR were bound into enumeration books. The 6OPH-3PR was a leaflet.

When the enumeration was completed, therefore, the Puerto Rico Operations Office received two or more enumeration books for each ED, containing the basic items of data for 75 percent of the population and housing units on either form 6OPH-1PR or form 6OPH-2PR, and also a group of 6OPH-3PR leaflets containing both the basic data and the sample data for the remaining 25 percent of the population and housing units in the ED.

PROCESSING OPERATIONS

The processing of the data from the censuses of Puerto Rico was done by conventional punchcard machine tabulation rather than by means of the microfilm-FOSDIC-computer complex used for U.S. data. The major operations in processing the data for Puerto Rico were the following:

1. Receipt and check-in of enumeration books
2. General coding, and verification
3. Coding of occupation and industry items, and verification
4. Key-punching the items into punchcards, and verification
5. Tabulating data by means of punchcard machine processes
6. Editing and balancing machine tabulations
7. Posting, checking, and typing tables for publication

The procedures were developed largely on the basis of procedures used in the 1950 censuses of the United States and Puerto Rico, which had been processed by punchcard machine tabulation.

Receipt and Check-In of Enumeration Books

Prior to the enumeration of Puerto Rico, Minor Civil Division (MCD) Sheets were prepared which listed the ED's within individual areas (city, town, village, or barrio) in Puerto Rico.

The MCD Sheets were used by the field offices to insure that each of the ED's was enumerated and that the basic enumeration book and sample enumeration schedules for each ED were submitted to the Operations Office. The field offices transmitted to the Operations Office the enumeration books for an entire municipality at a time along with the corresponding MCD Sheets. The receipt

and check-in operation involved comparing and reconciling any differences between the field office copy of the MCD Sheets and the original MCD Sheets, and checking in each enumeration book by verifying municipality and ED numbers on the enumeration book labels against those on the MCD Sheets.

Coding

After the enumeration books had been checked in, the data were coded. Coding involved preparation of entries on a schedule for transfer to a punchcard. Many codes were simple enough that the punchcard operator could punch them directly from the schedule entries. Other codes, however, were more complicated or required more checking for their determination, so a special coder wrote the code numbers on the schedule prior to punching.

If a coder was unable to code an entry, he prepared a problem referral slip which was given to the supervisor for resolution of the problem.

General coding.—The general coders entered all required codes except those for occupation and industry, and also examined for completeness certain items which had been precoded by the enumerator.

In addition to the items on occupation and industry, the following items were coded and examined:

Basic enumeration schedules:

Items coded:

- Relationship to head of household
- Place of birth

Items examined:

- Marital status

Sample enumeration schedules:

Items coded:

- Relationship to head of household
- Marital status
- Place of birth
- Migration
- Place of employment
- Individual income from various sources

Items examined:

- Birth date relative to April 1946
- Date last worked
- Monthly rent for land
- Rent
- Monthly utility costs

In addition, group quarters were coded to differentiate among institutions, i.e., jails, mental hospitals, etc., and noninstitutions such as boardinghouses and military installations.

The coders used a manual, General Coding Instructions, which gave the codes in full detail, and a summary "Code Card" which listed the most commonly used codes.

General coding was checked by a system of quality control. The work of each coder was completely verified at first, and records were kept on his work. Coders who reached a specified level of accuracy were designated qualified coders, and thereafter only a certain percentage of their work was verified.

Verification of the coding was also verified. The method used (following the pattern used in quality control of the 1950 United States censuses) was to note coding errors before the portfolio of enumeration books went to the

verifier; after the verifier completed his work, the errors he listed were compared with those previously noted.

General coding began June 1, 1960, and lasted 18 weeks. The coders processed the schedules at an average rate of 1,900 lines (persons enumerated) per man-day. (These coding rates cannot be compared with those achieved in editing and coding the FOSDIC schedules in the United States, both because the methods of maintaining the records were dissimilar and because various items were punch-coded in Puerto Rico, i.e., the card puncher translated the enumerator's entries directly into a coded format at the time of punching, whereas the punching operation was eliminated completely in the United States.)

Industry and occupation coding.--Questions on industry, occupation, and class of worker were asked for persons 14 years old and over living in every sample household enumerated in Puerto Rico. The industry and occupation coding clerks coded industry, occupation, and class of worker, and also entered codes for persons serving in the Armed Forces. The coders used a manual, Industry and Occupation Coding Instructions, which gave detailed instructions, an Alphabetical Index of Occupations and Industries, and "Company Name Lists" (one for each of the three standard metropolitan statistical areas and one covering the remaining territory in Puerto Rico) to determine the codes for the schedule entries. The verification procedure used was basically the same as that for general coding.

Industry and occupation coding started in mid-June 1960 and proceeded at an average rate of 1,600 persons coded per man-day (based on the total population of the sample enumeration book). This operation took 14 weeks to complete. (These rates are not comparable with those obtained in the United States, primarily because they are based on the total population in the enumeration book rather than on the number of persons for whom coding was required.)

Punching

After the schedules were coded, information was transferred to punchcards. A card was prepared for each person and for each housing unit. Each column or group of columns on a card represented a characteristic of that person or housing unit. For example, column 23 on the basic population card was for the sex of the person enumerated. On the card for a male, a hole was punched in the space allotted to "1," the code for "male." When the card was tabulated, one male was counted.

Four separate punching operations were involved, consisting of--

1. Approximately 1,735,000 basic population cards, each requiring 32 columns of punching
2. Approximately 620,000 sample population cards, each requiring 79 columns of punching
3. Approximately 398,000 basic housing cards, each requiring 40 columns of punching
4. Approximately 135,000 sample housing cards, each requiring 72 columns of punching

The punchcard operators began work on the basic population cards in mid-June 1960 and completed this operation by October 1, 1960. The basic housing punching operation started a week later and was completed by October 28, 1960. Punching for the sample data began the second week in July 1960 and was completed January 25, 1961.

The punching was verified by a system of quality control. Initially, each punchcard operator's work was completely verified. When he met certain minimum standards of efficiency, he became a "qualified puncher," and thereafter, only a small percentage of his work was verified.

Mechanical Editing

Mechanical editing involved inspection of the punchcards by machine to detect errors in punching, coding, or enumeration. The machines made three types of checks: (1) checks for impossible codes, (2) checks for inconsistent code combinations, and (3) checks of quantitative data for relationships, magnitudes, and arithmetic.

For example, only two entries in the column for "Electric lighting" had meaning--a "1" if the housing unit had electric lighting and a "2" if it did not; any other numbers punched in that column were impossible codes, and the cards having them were withdrawn and corrected.

Mechanical editing greatly simplified examination for errors. For most of the card types, unusual or inconsistent reports could be detected in one run through the 101 Electronic Statistical Machine, and the cards removed for investigation.

The selected cards were machine-listed, and the listed items were examined for possible errors. Other items on the card sometimes explained the entry or indicated the correct one. If they did not, the enumeration book was examined. If changes were necessary, new cards were punched. If no change was necessary, an explanation was entered on the listing sheet. Then, if the figures were questioned in the critical examination of the tabulations, the listing sheets usually provided the explanation.

Tabulating

The data collected by the enumerators reached the tabulating process in the form of punchcards. The tabulating machine read the entries on the cards, added or counted the number of entries in each classification, and printed the results.

The information, which was scattered through the more than 3,000,000 punchcards, had to be funneled into meaningful groups. These groups had been generally determined by the Bureau of the Census staff in consultation with users of the data. The subject specialists put the general requirements in the form of table outlines, which were translated by the machine specialists into tabulation specifications. Those specifications indicated how the cards should be grouped and how the machines should be wired to provide the data for the proposed publications.

The cards punched from the enumeration schedules were sorted mechanically into the groups for which totals were needed. The cards for each municipio were tabulated as a unit, and each tabulation sheet carried separate totals for each group into which the cards had been sorted. Totals for the municipio were obtained by adding the printed totals for each barrio, subbarrio, city, etc. The figures on the tabulation sheets were critically examined by the subject specialists before they were transferred to final publication tables.

All tabulation work for the censuses of population and housing of Puerto Rico, except for the series PC (1)-D population reports (on detailed characteristics) and the housing volume II reports (on metropolitan housing), was completed in the Puerto Rico Operations Office.

These two reports were tabulated in Washington, on Census Bureau punchcard equipment--simultaneously with the tabulation of the other reports in Puerto Rico--in order to make these data available at an earlier date than would otherwise be possible.

Preparation of Tables for Publication

To transform the figures on the tabulation sheets into tables for the final published volumes, four steps were required before printing:

1. Tabulations were reviewed to detect possible machine failure and also errors in enumeration, editing, coding, and punching.
2. Final tables were prepared by hand posting.
3. A machine check was made, when the tables were in final form, to locate arithmetic errors in the original figures and also typing or typesetting errors.
4. Related figures on the tables were checked to see if they were consistent, and the 1960 data were compared with those for 1950.

The result work operations which were carried out in Puerto Rico produced the tables for population report series PC(1)-A--Number of Inhabitants, series PC(1)-B--General Characteristics, and series PC(1)-C--Social and Economic Characteristics; housing vol. I--General Housing Characteristics, and vol. III--Block Statistics; and the population and housing census tract reports. Result work for population series PC(1)-D and housing vol. II, which were tabulated in Washington, was carried out in Jeffersonville.

Manual typing of the final tables and all related operations necessary to prepare the tables for printing were also conducted in Washington.

INFLATING THE SAMPLE: RATIO ESTIMATES

The procedure established for inflation of the sample required specified counts for population and for housing from both the basic and the sample cards. For this purpose, the basic file (representing 75 percent of the population and housing units) was grouped into weighting areas, by municipio and city (selected larger cities) and, in the three largest cities (San Juan, Mayaguez, and Ponce), by barrio.

For each weighting area, the cards were sorted, and counts were obtained, for each of the following classifications:

Male:

- Heads of households
- Other members, 14 years old and over
- Other members, under 14 years old

Female:

- Heads of households
- Other members, 14 years old and over
- Other members, under 14 years old

Housing units:

- Owner occupied
- Renter occupied
- Vacant

While these operations were being carried out, a similar series of operations were being performed on the punchcards for the 25-percent sample of the population and housing units. First, complete clerical editing, coding, and verification operations were performed, then card punching and verification, followed by a mechanical edit covering the same items as in the corresponding edit of the basic cards.

The cards for the sample were grouped into weighting areas, and the cards were sorted into, and counts obtained for, the identical classification established for the basic cards. For each weighting area, the counts obtained were posted on the Ratio Estimate Control Record. At this point, all the required information was available for adjusting the cards for the sample so that each classification would contain an exact 25-percent sample.

It was necessary to provide the sample with a uniform integer weight--in this case 4--in order to permit the cards to be tabulated on 101 (unit count) equipment. The machine would count each card once but the resultant tabulations could be clerically multiplied by 4 to obtain the inflated sample results. For each separate weighting-area classification group, a series of instructions was developed which provided for either the replication or destruction of the punchcards, or the conversion of cards for the sample to basic cards, in such a fashion as to provide a set of sample and basic cards (1) which agreed in their total count with the previously established population and housing control totals and (2) in which each sample card had a uniform weight of 4.