



NACIONES UNIDAS

PROPIEDAD DE
LA BIBLIOTECA

C.1

CONSEJO
ECONOMICO
Y SOCIAL



LIMITADO

ST/ECLA/Conf.32/L.9
21 de mayo de 1968

ESPAÑOL
ORIGINAL: INGLES

SEMINARIO SOBRE ORGANIZACION Y
LEVANTAMIENTO DE CENSOS DE POBLACION
Y HABITACION PARA AMERICA LATINA

Organizado por la Comisión Económica para América Latina, la Oficina de Estadística y la Oficina de Cooperación Técnica de las Naciones Unidas, con la colaboración del Instituto Interamericano de Estadística, el Centro Latinoamericano de Demografía y la Dirección de Censos de los Estados Unidos
Santiago de Chile, 20 a 31 de mayo de 1968

ELABORACION ELECTRONICA DE DATOS CENSALES

I. Uso de computadoras en la elaboración de datos censales

* Preparado por la secretaría de la Comisión Económica para Asia y el Lejano Oriente (CEPALO) para el Seminario sobre Organización y Levantamiento de Censos de Población y Habitación de los países de la región de la CEPALO (Bangkok, 24 de noviembre a 1° de diciembre de 1967), y publicado por la CEPALO como documento de las Naciones Unidas E/CN.11/ASTAT/SPHC/L.6 (2 de noviembre de 1967).

INDICE

	<u>Párrafos</u>	<u>Página</u>
I. INTRODUCCION	1- 4	1
II. VENTAJAS DE LA ELABORACION ELECTRONICA.....	5-13	2
III. NECESIDAD DE PLANIFICACION INICIAL	14-16	5
IV. IMPORTANCIA QUE TIENE PARA LA LABOR DE PLANEACION EL EQUIPO UTILIZADO	17-21	7
V. DISEÑO DE LOS SISTEMAS Y PROGRAMACION	22-26	9
VI. CONCLUSION	27-29	11

I. INTRODUCCION

1. Con escasas excepciones, los resultados de la serie de censos de población de 1960 en la región de la CEPALO se tabularon utilizando equipos corrientes de tarjetas perforadas. A juzgar por la rapidez con que el trabajo estadístico de la región se está orientando al uso de computadoras, parece evidente que casi todos los países piensan elaborar electrónicamente los resultados censales de 1970. Las oficinas de estadística de los países que para entonces no dispongan de computadoras tal vez puedan utilizar los equipos de algún país vecino o del centro regional de computación que se ha propuesto para la región de la CEPALO. Por lo tanto, para los fines de organizar y planificar los próximos censos, en el presente trabajo se ha supuesto que en 1970 la elaboración de datos en la región se hará casi exclusivamente con técnicas electrónicas.

2. Para los censos de población de 1960 se preparó una documentación abundante con instrucciones para el uso de métodos de tarjetas perforadas. Cabe destacar que muchos de los principios enunciados en la publicación informativa Manual sobre métodos para la elaboración de datos, publicado conjuntamente por las Naciones Unidas y la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), pueden aplicarse directamente a la elaboración electrónica de datos. Asimismo, los países han preparado algún material útil para la elaboración electrónica de los censos.

3. Entre las actividades planeadas para la serie de censos de población y habitación de 1970, se ha propuesto la realización de seminarios destinados a examinar consideraciones técnicas sobre el uso de computadoras en la elaboración de datos censales. Se ha pensado que en la región de la CEPALO podría realizarse un seminario de esta índole en la primera mitad de 1969. La documentación de esta reunión y de otras similares, así como los servicios de organizaciones internacionales y otras organizaciones con experiencia en la elaboración electrónica de datos de los censos de población, ayudarán a los países de la CEPALO en la preparación de los censos de 1970. En el presente trabajo, por lo tanto, sólo se examinarán los principios y los aspectos generales de la planificación de la elaboración electrónica de los censos de población y habitación.

4. Para muchos países de la región, los censos de 1970 marcarán su iniciación en el uso de computadoras. Por esto tal vez sea útil hacer una breve descripción de las principales características de la elaboración electrónica, especialmente en su aplicación al trabajo censal.

II. VENTAJAS DE LA ELABORACION ELECTRONICA

5. El uso de tarjetas perforadas para registrar los datos estadísticos, y el uso de máquinas selectoras, contadoras y tabuladoras para compilar y resumir la información registrada en las tarjetas respondió en un comienzo a la necesidad de contar con instrumentos adecuados para la tarea masiva de elaborar los datos censales. La computadora electrónica permite aplicar métodos mucho más rápidos, versátiles y precisos que los métodos anteriores en la elaboración de los datos censales. Actualmente, se ha generalizado el uso de equipos electrónicos para la elaboración masiva de programas estadísticos en casi todos los modernos organismos gubernamentales de estadística del mundo.

6. En general, los principales atributos por los que la elaboración electrónica de datos supera ampliamente a los métodos corrientes incluyen los siguientes:

- i) Rapidez muchísimo mayor en las computaciones y transferencias de datos;
- ii) Almacenamiento de memoria que proporciona la capacidad para "recordar" y recuperar con gran velocidad de acceso los datos almacenados en la memoria;
- iii) Capacidad de ejecución secuencial automática de instrucciones;
- iv) Capacidad de tomar automáticamente decisiones lógicas durante la elaboración;
- v) Precisión y autoverificación muy mejorados.

La combinación de estas características de los equipos electrónicos de computación permite resolver problemas complejos de manejo de datos por medio de un proceso continuo que se realiza a velocidades electrónicas; ésto contrasta con el manejo fragmentario de datos en el sistema corriente de tarjetas perforadas. La capacidad para elaborar grandes cantidades de

/datos y

datos y la flexibilidad de los nuevos equipos, marcan una diferencia cualitativa con respecto a dispositivos anteriores, y permiten utilizar nuevas técnicas en las operaciones estadísticas.

7. Por ejemplo, las computadoras suelen usarse con grandes ventajas en las tareas censales de crítica y corrección de los datos, antes de la tabulación o junto con ella. Se puede hacer que la computadora asigne (o impute) valores automáticamente cuando se han omitido datos personales o habitacionales o cuando algunos de los datos anotados (o codificados) son incoherentes o imposibles. La mayoría de las computadoras pueden aplicar a los datos todas las normas que utilizaría todo un grupo de funcionarios para descubrir y conciliar incoherencias y deducir la información que falta. Suele suceder, por ejemplo, que de las características anotadas para alguien, una sea incongruente con las demás. En estas circunstancias, puede aplicarse el procedimiento general de deducir el valor de la característica partiendo de una distribución de esa misma característica en el subgrupo pertinente de la población. Así, si falta el estado civil de un pariente masculino del jefe de hogar, la computadora puede asignarle el estado civil más probable, conforme a la distribución de estados civiles de los parientes masculinos de jefes de hogar, pertenecientes al mismo grupo de edades.

8. Para descubrir datos omitidos o imposibles pueden usarse procedimientos similares. En muchos sentidos, la computadora realiza estas tareas de crítica mucho mejor y más rápido que cuando se hacen por inspección manual. La computadora aplica las normas de crítica por separado a cada persona empadronada, y lo hará con coherencia y uniformidad estrictas; se puede hacer que la computadora lleve un registro de la frecuencia con que se ha aplicado cada regla, dando así una medida de la calidad de los datos empadronados.

9. El uso de computadoras en la elaboración de censos ofrece la posibilidad de obtener estadísticas más detalladas o más informativas. La flexibilidad de las computadoras permite hacer mucho mayor uso de los datos en estadísticas derivadas y otros subproductos de los que se dispone por primera vez. Los coeficientes y relaciones clave que antes no podían obtenerse sin una costosa reelaboración de los datos, o sin una gran cantidad de computaciones

/manuales, ahora

manuales, ahora pueden programarse como subproducto de las tabulaciones electrónicas. Asimismo, los datos pueden reordenarse con más facilidad por más variables de clasificación y en clasificaciones cruzadas más detalladas; utilizando otros métodos, este trabajo tendría un costo prohibitivo en tiempo y dinero. Las computadoras ofrecen también la posibilidad técnica de combinar y relacionar datos censales con datos de otras fuentes que se refieren a la misma área de investigación, de modo que se podrían comparar y combinar datos censales con aquellos de otras fuentes que, por una razón u otra, no pueden recolectarse en el censo. Por ejemplo, los datos del registro civil podrían usarse para actualizar en los períodos intercensales algunas de las características de la población compiladas en el censo.

10. Otra ventaja de la elaboración electrónica de datos censales reside en que a menudo permite preparar directamente los cuadros definitivos para la impresión offset sobre la base de los cuadros preparados por las computadoras. Con los equipos más antiguos de tabulación mecánica, generalmente los resultados no pueden prepararse directamente para su impresión. Comúnmente hay que hacer una cantidad considerable de contabilización a mano basándose en los listados, y luego hay que mecanografiar los cuadros preparados en la forma en que se publicarán. Al preparar los cuadros definitivos directamente para la imprenta, las computadoras eliminan la necesidad de confeccionar un original, de componer tipos y de corregir pruebas; por lo tanto, los resultados definitivos del censo pueden darse a conocer al público muchos meses antes que si se usan métodos convencionales de impresión.

11. Como en la región generalmente escasean el espacio físico y las instalaciones para almacenamiento, las computadoras tienen la ventaja de necesitar una superficie mucho menor que el equipo convencional. En los países que usan sistemas de computadores de cinta magnética, el almacenamiento de los registros censales sólo necesita una fracción del espacio que normalmente se necesita para guardar tarjetas.

12. Por supuesto, la transición al sistema de computadoras tiene algunas desventajas y presenta algunas dificultades. El uso de computadores plantea nuevos problemas, que obligan a organizar la elaboración de datos censales

en forma más sistemática que cuando se usan equipos corrientes de tarjetas perforadas. En la elaboración electrónica, las instrucciones que se dan a la computadora antes de iniciar cualquier operación deben ser muy detalladas y explícitas. Por ejemplo, en los procesos de compaginación de datos, la computadora deberá programarse para prever todas las combinaciones pertinentes de entradas que puedan darse. El no considerar todas las contingencias haría que la computadora se detuviera cada pocos minutos, o rechazara muchísimos registros por alguna deficiencia en los datos que no se previó al programar la computadora. Por esta razón, el personal de programación y supervisión debe ser muy competente y capacitado. Como hay gran demanda de personal con estas calificaciones, y éste es muy escaso, para atraerlo suele ser necesario ofrecer incentivos bajo la forma de primas y subsidios especiales. Las instalaciones para albergar las computadoras deben satisfacer muchas más exigencias que las destinadas a equipos para tarjetas perforadas de tipo convencional, pues deben tener aire acondicionado, sistemas para eliminar el polvo, pisos especiales, etc.

13. Es fundamental reconocer que la disponibilidad de equipos muy perfeccionados de elaboración de datos no es en sí garantía de resultados censales eficientes y exactos. Para que el uso de las computadoras tenga un buen éxito, las operaciones de elaboración censal deberán planificarse con mucho más disciplina y precisión que si se utilizan técnicas más antiguas.

III. NECESIDAD DE PLANIFICACION INICIAL

14. Los métodos y equipos que se usan en la elaboración tienen vinculaciones directas con los objetivos y resultados del censo, e influyen mucho en casi todas sus operaciones. Cada etapa del censo están estrechamente ligada a la operación anterior, y a su vez, influye en las operaciones siguientes. Así, es esencial que la función de elaboración de datos esté representada desde el primer momento en las labores de planificación y organización del censo a través de un miembro integral del grupo que planifica el censo en general.

15. Dentro del grupo de planificación, el representante de la elaboración de datos, participaría activamente en el diseño del cuestionario, para que el diseño definitivo contribuya a una elaboración eficiente. Asimismo, ayudaría a determinar los procedimientos de control, las especificaciones para las tareas de crítica y codificación, las necesidades de tabulación, los programas y otros factores, para que las autoridades de planificación censal puedan considerarlos en forma realista.

16. La introducción de las computadoras en la elaboración censal, por su misma naturaleza, eleva el tiempo necesario para preparar la elaboración estadística. Es fundamental, por lo tanto, que haya una planificación preliminar cuidadosa. En el pasado los encargados de elaborar los datos se han visto a veces obligados a improvisar su planificación detallada. Cuando se usan computadoras, este tipo de planificación no sólo es difícil sino peligrosa para los resultados definitivos. La programación de las computadoras es una labor muy intrincada que debe planificarse con gran eficiencia y una atención minuciosa a los detalles. El costo, esfuerzo y tiempo que se necesitan para modificar los procedimientos electrónicos hacen indeseables tales cambios. Así, para cumplir con los plazos fijados, los planes completos de tabulación y las especificaciones definitivas deben entregarse a los programadores con mucho más anticipación que cuando se usan sistemas de elaboración más antiguos. Las posibilidades de variar los planes de tabulación o de improvisar son muy limitadas cuando los plazos apremian. Esto, como es evidente, indica que la planificación general del censo debe hacerse con mucha anticipación. Muchos países incluidos aquellos que tienen una larga historia de levantamientos de censos, dedican cinco años o más a la planificación activa del censo. Los Estados Unidos, con 180 años de experiencia censal, ha estado trabajando activamente en la planificación de los censos de 1970 desde 1964.

IV. IMPORTANCIA QUE TIENE PARA LA LABOR DE
PLANEACION EL EQUIPO UTILIZADO

17. La capacidad, velocidad y configuración de la computadora que ha de utilizarse para elaborar el censo es, naturalmente, el factor de mayor importancia en relación con el plan de elaboración de datos. El equipo puede consistir, ya en computadoras de poca velocidad que funcionan mediante tarjetas y poseen muy limitada capacidad de almacenamiento, ya en instalaciones de gran capacidad de almacenamiento y mucho equipo auxiliar. Estas consideraciones determinarán frecuentemente el concepto general que presida el plan de levantamiento del censo, con inclusión del diseño del formulario de empadronamiento, de los trabajos sobre el terreno, de las labores de codificación, de la preparación de los datos de entrada, de la compaginación y de la tabulación. La influencia que ejerce el equipo sobre el plan censal en su totalidad se pone de relieve especialmente si se consideran los efectos tan generales que ejercería sobre todas las operaciones censales el uso de las técnicas de barrido óptico de los cuestionarios.

18. En consecuencia, para la planeación es esencial tener, al menos, una idea de qué equipo ha de utilizarse en el levantamiento del censo. A este efecto deben distinguirse, concretamente, los siguientes sistemas: i) tarjetas únicamente; ii) tarjetas y cinta magnética, y iii) tarjetas, cinta magnética y lectoras ópticas.

i) Sistemas que operan únicamente con tarjetas

19. El uso de computadoras que poseen sólo capacidad para entrada y salida de tarjetas habrá de acelerar, desde luego, las labores de compilación estadística, si se compara su rendimiento con el uso de máquinas ordinarias perforadoras de tarjetas, pero, en realidad, representará únicamente una modificación mínima de los métodos previamente utilizados. Con las computadoras se obtiene una mayor velocidad de cálculo y de compaginación de los datos que con las máquinas perforadoras corrientes, pero el volumen del trabajo de perforación de tarjetas que es preciso realizar para preparar las tarjetas de entrada, no varía. Por otro lado, lectoras de tarjetas de alta velocidad (que llegan a leer unas 800 tarjetas por minuto, en tanto que las tabuladoras corrientes sólo admiten unas 150) permiten que la

/computadora haga

computadora haga con mayor rapidez cada pasada y la capacidad de almacenamiento de la memoria hace posible tabular más variables en cada pasada. Esta capacidad de almacenamiento permite igualmente que puedan compaginarse mejor los datos registrados y facilita el descubrir más fácilmente los códigos inválidos, las incongruencias y el grado de lógica de las operaciones. Sin embargo, las operaciones de clasificación, intercalación, etc. pueden continuar efectuándose normalmente con el equipo de perforación.

ii) Sistema de tarjetas y cinta magnética

20. El tipo más frecuente de equipo, en el que probablemente se elaborará la próxima serie de censos de la región, consistirá en un sistema de capacidad media, que admita tarjetas y cinta magnética y posea una impresora, estando dotado posiblemente de equipo auxiliar para almacenamiento de datos en disco o en tambor. Esta instalación, una vez perforadas las tarjetas, pasadas a cinta magnética, compaginadas y corregidas en caso necesario, es capaz de clasificar y tabular los datos a gran velocidad. Sin embargo, también en este caso han de ser tenidos en cuenta en el plan de elaboración los gastos y demoras ocasionados por la perforación manual de los datos antes de que puedan llevarse a la cinta magnética para su elaboración a gran velocidad por la computadora. El indicado tipo de computadora será efectivo, especialmente, a causa de la mayor rapidez de los cálculos y por favorecer la tabulación definitiva de los datos. No obstante, los trabajos de tabulación quizás no puedan iniciarse hasta que se haya obtenido una cinta magnética "limpia", por lo cual ha de darse importancia especial, al planificar el censo, a acelerar las labores de perforación de tarjetas, transferencia de los datos de las tarjetas a la cinta magnética, y compaginación, así como al control de la calidad de esas operaciones.

iii) Sistema de tarjetas, cinta magnética y dispositivos de lectura óptica

21. Los sistemas de barrido óptico consisten en métodos automáticos de perforación de tarjetas o de registro en cinta magnética a partir del formulario censal, y su uso es relativamente reciente. Algunos países de la región de la CEPAL están estudiando la conveniencia de utilizar dichos dispositivos para la confección de los censos de 1970. En este caso, al eliminarse la necesidad de llevar a cabo operaciones de perforación de tarjetas en gran escala, habrá de modificarse concordantemente la planificación

en lo relativo a la asignación de recursos, aunque por otra parte será necesario emplear recursos y esfuerzos adicionales para obtener la necesaria calidad del trabajo sobre el terreno, que garantice un elevado porcentaje de formularios censales capaces de lectura electrónica. Esta lectura puede efectuarse mediante equipo lector de marcas o mediante dispositivos de identificación de caracteres. Sin embargo, aunque se han logrado progresos alentadores en estos últimos dispositivos, las técnicas correspondientes no se han desarrollado aún hasta el punto de poder ser generalmente utilizados para levantar censos de población. Los equipos de identificación de caracteres capaces de leer con precisión suficiente los símbolos escritos a mano o no normalizados que se usan en los trabajos censales, no son aún adecuados para su empleo general, si bien han sido y pueden ser utilizados con éxito los equipos capaces de leer las marcas en función de su posición. Es de desear que se obtengan experiencias más amplias con esas técnicas en aquellos sectores en que sea más fácil el control del trabajo sobre el terreno.

V. DISEÑO DE LOS SISTEMAS Y PROGRAMACION

22. El poder hacer a máquina las labores de compaginación y tabulación del censo exige un gran esfuerzo de diseño del sistema aplicable y una intensa labor de programación para la computadora. En todos los países existe una tendencia general a subestimar la magnitud del trabajo que requieren la programación y la instalación de los sistemas de cálculo electrónico. Dada la interacción de la elaboración de datos mediante computadora con prácticamente todos los demás elementos del censo, es esencial preparar de antemano la programación de las computadoras y comprobarla cuidadosamente, ya que no basta con que los programas estén preparados en el momento en que se proyecta comenzar la labor principal de elaboración de datos.

23. La mayor dificultad de la programación consiste en determinar los procedimientos de compaginación, esto es, en dar capacidad a las computadoras para que comprueben las fallas lógicas, tales como las combinaciones improbables de la codificación. Suele ser más difícil de lo que se cree

/la traducción

la traducción de las normas de compaginación en pasos lo suficientemente sencillos y exactos para que la computadora siga las instrucciones. Es especialmente importante comprobar con mucha antelación esos programas mediante la utilización de datos reales. La compaginación electrónica obliga a descubrir y corregir las omisiones de respuestas y las incongruencias, en cuanto sea posible, mediante operaciones efectuadas por la propia computadora. Si se hacen en fecha temprana pasadas de prueba para ensayar los programas de compaginación, será posible descubrir otras fases de las operaciones censales que puedan dar lugar a problemas de compaginación. Así se pone de relieve la necesidad de adoptar medidas correctivas, por ejemplo en la redacción del formulario, en las instrucciones que se dan a los empadronadores, en la codificación, en la labor de perforación de tarjetas o en la supervisión.

24. Si no se lleva a cabo una comprobación temprana y a fondo de toda la fase de elaboración de datos, podría suceder que la computadora, al ejecutar las operaciones reales, rechace un número excesivo de datos por una u otra causa, lo que, a su vez, crea el problema de tener que volver a registrar los datos rechazados y corregidos a consecuencia de la operación de compaginación. En consecuencia, se tardará más de lo proyectado y se necesitará más tiempo de la computadora que el que se pensaba utilizar.

25. La previa comprobación de la totalidad de las operaciones de la computadora, esto es, de los programas de compaginación y tabulación, permitirá también que se obtengan estimaciones mucho más exactas de las horas de computadora que se precisan. Con frecuencia han sido muy grandes los márgenes de error de las estimaciones basadas por entero en factores teóricos, lo que tiene importancia cuando las instalaciones están muy sobrecargadas de trabajo, y ha de contarse con un plan de operaciones de gran exactitud.

26. Las tareas de programación exigen mucho tiempo y la capacidad de las nuevas computadoras suele quedar disminuida por la falta de personal debidamente calificado. Las computadoras exigen un personal bien preparado, el que, para formarse suele necesitar hasta todo un año de instrucción. Por ello es de gran importancia que en el plan del censo se calcule un tiempo suficiente para que los programadores se capaciten con la suficiente antelación para terminar los programas de compaginación y tabulación, y ensayarlos debidamente, con anterioridad a la fecha de determinación del empadronamiento, a fin de comenzar inmediatamente después las tareas de elaboración de datos.

VI. CONCLUSION

27. El aumento del rendimiento y de la flexibilidad de la presente generación de computadoras, en relación con la precedente, situará a los estadísticos que elaboren el censo de 1970 en una posición muy ventajosa en relación con los que levantaron el censo de 1960. El nuevo equipo mejorará notablemente las técnicas censales de los países de la región, de igual modo que las mejoró hace unos años el empleo de las tarjetas perforadoras. A fin de usar eficazmente este equipo, ha de integrarse el diseño de la elaboración con los demás aspectos del diseño general del censo. Esta integración exige que se logre el equilibrio entre las tareas de supervisión, control de la calidad, compaginación y capacidad de rectificar los errores que se descubran, ya que en otro caso podrían presentarse problemas muy graves.

28. Debemos insistir, por lo tanto, en que se termine la programación de las computadoras bastante antes de proceder a la elaboración de los datos, y en que se comprueben cuidadosamente los datos que se obtengan mediante los adecuados ensayos. Aun en el caso de una vez efectuada esta comprobación no se modifique el diseño del sistema de cálculo electrónico, probablemente se conseguirá mejorar la supervisión y los controles de calidad de las etapas fundamentales de levantamiento del censo.

29. Los directores y planeadores de los censos de la región de la CEPALO, al aplicar de manera general las computadoras a la serie de censos de población de 1970, y si planifican cuidadosamente todas las tareas de elaboración, podrán confiar razonablemente en reducir el intervalo existente entre la reunión de datos y la publicación, por lo menos en una mitad del tiempo que fue necesario para los censos de 1960.

