NACIONES UNIDAS

CONSEJO
ECONOMICO
Y SOCIAL

LIMITADO
ST/ECLA/Conf.32/L.22
24 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

EL PAPEL DEL MUESTREO EN LOS CENSOS DE POBLACION SU EFECTO SOBRE LA EXACTITUD Y LA OPORTUNIDAD

Por

Joseph Waksberg, Jefe de la División de Métodos Estadísticos
Oficina del Censo de los Estados Unidos

DEPARTAMENTO DE COMERCIO DE LOS ESTADOS UNIDOS
Oficina del Censo
Washington, D. C. 20233

Abril 21 de 1967

EL PAPEL DEL MUESTREO EN LOS CENSOS DE POBLACION SU EFECTO SOBRE LA EXACTITUD Y LA OPORTUNIDAD

Por

Joseph Waksberg, Jefe de la
División de Métodos Estadísticos
Oficina del Censo de los Estados Unidos

I. Introducción

En los tiempos modernos, el levantamiento de los censos de población tiene esencialmente dos objetivos. El primero es obtener la cuantía total de la población, en relación, generalmente, con pequeñas áreas individuales del país. El segundo es reunir información - y hacerla útil y disponible - sobre características económicas y sociales escogidas de la población. El número de áreas que contienen tabulaciones pormenorizadas de tales características es generalmente más limitado que el de las utilizadas para el recuento de la población, aunque, en relación con áreas muy pequeñas, se desean con frecuencia medidas resumidas de algunas características. Por otra parte, la naturaleza de las áreas tiende a cambiar de un país a otro.

Este estudio se refiere exclusivamente al segundo objetivo, o sea, las estadísticas relativas a las características económicas y sociales de la población. Se discuten algunos aspectos concernientes a la oportunidad de la publicación de los resultados, el efecto que las largas demoras en la publicación de dichos resultados produce en nuestra comprensión de los cambios que ocurren; se describen métodos para acelerar la preparación de los informes, y se sugiere un programa para ampliar el empleo del muestreo en los censos de los países en desarrollo con el fin de reducir las demoras.

Los datos censales corresponden, desde luego, a los programas de acción y a la investigación científica; para ello se requieren datos razonablemente extensos y resultados prontos. La mayor parte de los países del mundo se han dado cuenta, en la actualidad, de la importancia de levantar periódicamente censos de población y han tomado las providencias tendientes a conducirlos sobre bases de regularidad. Sin embargo, por variadas razones, en muchos casos encuentran dificultades para producir estos resultados dentro de un "lapso" razonable. (Entendemos por "lapso" el tiempo que transcurre entre el levantamiento del censo y la producción o publicación de los resultados.) Las largas demoras reducen seriamente el valor práctico de los resultados, a más de que, generalmente, son los países de programas estadísticos intercensales mínimos los que tienen las demoras más grandes. Zarkovich presenta algunos casos extremos de retardo en las publicaciones del censo.[1/] Por ejemplo, todas las tabulaciones planeadas para el Censo de 1921 en Yugoeslavia no se habían completado cuando se levantó el Censo de 1931. El procesamiento del nuevo Censo era todavía incompleto cuando, en 1941, el país se encontró envuelto en la guerra y fué destruído el material censal. La difusión del uso y la disponibilidad de equipo automático de tabulación en los últimos años ha mejorado la situación, pero todavía son comunes serias demoras. Por ejemplo, de los 17 países latinoamericanos cuyos datos de los recientes censos están disponibles, al menos se publicó una serie de resultados finales en 12 casos. Para estos 12, el lapso promedio requerido para la primera publicación fué de alrededor de los 3 años; parece probable que algunos países necesitarán 5 años o más para producir los datos. (Ver Tabla 1.) De los 5 países que no han producido todavía los datos finales, en 2 casos ha transcurrido ya un lapso de 6 años y medio.

¿Qué se puede hacer para reducir las largas demoras provenientes esencialmente de la imprevisiva carga de trabajo implicada en la edición, codificación, perforación y tabulación de los datos? La utilización del

---

1/ S. S. Zarkovich. Métodos de Muestreo y Censos. Organización de las Naciones Unidas para la Agricultura y la Alimentación, 1965. p. 87.

más sofisticado equipo y técnicas de tabulación se encarga solamente de parte del problema. Es teóricamente posible concebir un proceso educativo que inclinará a un administrador a consagrar suficiente presupuesto, mano de obra y otros recursos al censo, a fin de que la totalidad del censo se procese tan rápidamente como parezca deseable. Teniendo en cuenta la competencia por tales recursos entre varios de los programas urgentes en los países en desarrollo, es improbable que ésto sea factible en muchas situaciones prácticas. Tanto más, cuanto no es cierto que sea necesario.

Una alternativa realista es reducir la carga de trabajo del procesamiento de los datos mediante la utilización del muestreo; lo cual se puede obtener por medio del desarrollo de dos tipos de programa. El primero consiste en utilizar muestras para un programa de informes preliminares, para proseguir luego con los informes finales basados en el censo total. Esto contribuirá, normalmente, a la utilización de preferencia de una muestra pequeña.

El segundo es abandonar la noción del tratamiento del 100 por ciento de los datos, con excepción de una mínima información necesaria para áreas muy pequeñas, y basar la mayor parte de las estadísticas en muestras. Realmente se requieren grandes muestras para tales fines, pero una muestra de un tamaño entre el 25 y el 10 por ciento de la población elimina del 75 al 90 por ciento del trabajo de producción. (Desde luego, es posible también combinar estos dos programas, de tal manera que se produzcan los datos preliminares partiendo de una muestra pequeña seguida de una muestra grande para los datos finales.)

II. Experiencias con las Tabulaciones Preliminares

Acortamientos en el Lapso: En años recientes, la producción de los datos preliminares tomados de una muestra de un censo llegó a ser razonablemente común. De los 17 países latinoamericanos examinados, 13 desarrollaron tales programas. Entre los países que tenían programa de muestra hubo variaciones en cuanto al tamaño de la muestra pero todos tuvieron la meta común de acortar el lapso de las publicaciones.

En muchos países se alcanzó esta meta en forma substancial. En la Tabla 1 se relacionan los países, las fechas de los censos más recientes,

las fechas de publicación de los resultados preliminares y fechas semejantes de las publicaciones finales. El análisis del lapso resultará un poco imperfecto ya que no se dispone del mes de publicación y, para los períodos cortos considerados, el tiempo de un mes puede producir algunas diferencias. No obstante, el acortamiento del lapso es básicamente claro. En 13 países el lapso promedio entre el censo y los resultados preliminares fué de 2 años aproximadamente. En muchos países el acortamiento del lapso fué, por lo menos, de 1 ó 2 años. Los dos países que gastaron más de 3 años en producir los resultados preliminares no han dado todavía las cifras finales, más o menos a los 6 años y medio después del censo.

En otras partes del mundo se pueden presentar ejemplos semejantes de acortamiento del lapso. En el Censo de Población de los Estados Unidos en 1950, los datos preliminares correspondientes a la Nación y a los Estados más grandes (basados en un 0.1 por ciento de muestra) fueron dados a la luz en febrero-abril de 1951. Los informes finales, comparables a las estadísticas del censo total de los Estados Unidos, no se completaron hasta 1953. En octubre de 1950 se levantó el Censo del Japón. Los datos preliminares de la muestra del 1 por ciento se entregaron en mayo y junio de 1951. Areas más detalladas, basadas en una muestra del 10 por ciento, vieron la luz en la primavera de 1952. Cuando el Censo de 1965, Japón produjo los datos preliminares, igualmente de una muestra del 1 por ciento, exactamente al año del Censo. Yugoeslavia produjo resultados de la muestra en agosto-noviembre de 1954 para un Censo levantado en abril de 1953 - inmensa distancia de la aciaga experiencia de sus censos inmediatamente anteriores. Indonesia publicó datos de su Censo de octubre de 1961 en diciembre de 1963, basados en muestra del 1 por ciento. No se han entregado los datos finales. El Censo de Korea del Sur se levantó en junio de 1961. Características adicionales económicas y sociales se publicaron en julio de 1962. Los resultados finales no estuvieron disponibles sino hasta octubre de 1963. Un grupo de países africanos utilizó también programas de muestra para censos recientes.

Al menos en algunos de estos países, no se reconoció la necesidad de un programa para producir datos preliminares hasta cuando ya estaba en marcha el censo. Es verosímil que si en los planes iniciales se hubiera

incluído tal programa, el lapso se hubiera acortado más aún.

Exactitud de los Resultados Preliminares: Muchos de los informes de los datos preliminares incluyen tablas de errores de muestreo. Los errores de muestreo son, desde luego, una función del tamaño de la muestra utilizada para elaborar el muestreo. En la mayoría de los países se escogieron muestras adecuadas para impartir alta exactitud a las estadísticas nacionales. En muchos casos, y para subdivisiones principales del país, se produjeron también datos útiles. En general, las estadísticas del país, tomado en su totalidad, estuvieron sujetas a errores triviales justificados. Como era de esperar, los errores alcanzaron mayor importancia en las subdivisiones geográficas. Por ejemplo, en el Censo del Japón de 1965, la estimación del número de personas de menos de cinco años de edad en todo el Japón, basada en la muestra preliminar del uno por ciento, tiene un coeficiente de variación debido al muestreo de alrededor de un tercio del uno por ciento. Se publicaron también, para siete grandes ciudades, las distribuciones preliminares por grupos de edad. La escala del coeficiente de variación del número de personas menores de 5 años en las ciudades varió de un mínimo del uno por ciento hasta cerca del tres y medio por ciento. En otros programas se presentaron ejemplos semejantes.

Las tablas de errores de muestreo no reflejan necesariamente todos los elementos que afectan la exactitud de los datos de la muestra. Consideración importante es la capacidad de controlar el muestreo y otras operaciones, a fin de que la muestra represente correctamente el censo total. Para examinar qué tanto éxito alcanzaron en dicho control varios países es provechoso comparar los datos preliminares publicados, y los datos finales. Las Tablas 2 - 5 contienen tales comparaciones en lo que respecta a Venezuela, Costa Rica, El Salvador y los Estados Unidos. Estos no fueron seleccionados como una muestra de todos los países. Hay países cuyos datos resultaron convenientemente disponibles. Las características muestran que los renglones se escogen aleatoriamente de entre las clases de datos que parecen ser particularmente importantes. Los datos preliminares de Venezuela se basaron en una muestra del 4 por ciento en Caracas y del 1 por ciento en el balance del país. Las muestras de Costa Rica y El Salvador fueron esencialmente del 5 por ciento. La muestra de los

Estados Unidos fué del 0.1 por ciento.

En todos los casos, los datos preliminares estuvieron muy cercanos a las cifras finales. Es difícil vislumbrar cualquier deducción relacionada con la situación social y económica que se pueda obtener de las cifras finales, que no esté clara en las bases de la información preliminar. Además, para la mayor parte de los propósitos, cuando se publican los datos finales, son ellos solamente una tosca aproximación de la situación actual. La superimposición de un insignificante error de muestreo parecería tener únicamente un efecto menor sobre el error total.

III. Extensión del Uso del Muestreo

Si uno mira los datos de un país en conjunto, parece innecesario molestarse en utilizar la totalidad de las cifras finales, una vez que estén disponibles las cifras preliminares. Solamente cuando uno se interesa por los datos de una área pequeña resultan inadecuadas las pequeñas muestras. Por éso, el interrogante natural es si una extensión del uso del muestreo no producirá datos inadecuados para las áreas pequeñas aunque se obtenga un acortamiento del lapso. Dado que el tiempo transcurrido entre la recolección de los datos y la publicación de los resultados depende en gran parte de la carga de trabajo del procesamiento, cualquier reducción significativa en dicha carga de trabajo traería la correspondiente reducción en el lapso.

Se plantean dos interrogantes al considerar este programa: (1) Hasta qué punto las demoras afectan seriamente la utilización de los datos; (2) La disminución en la exactitud es lo suficientemente grande como para hacer inaconsejables las estadísticas de muestra para las pequeñas áreas. Examinemos primero la consideración referente al lapso.

Efecto de las demoras de la publicación en la utilización de los datos: No es fácil cuantificar el efecto de la demora en la producción de los datos estadísticos sobre el análisis de la situación actual. El error promedio a lo largo del tiempo debe considerarse una aproximación general resultante de utilizar la información del censo más reciente como una constante durante una década. Un simple modelo sería considerar la "verdad" durante la década 1955-1964 como representada por una tendencia lineal desde el Censo de 1950

hasta los resultados de 1960. De cualquier manera, debido al largo tiempo gastado en procesar los datos del censo, la única información de que dispone un analista durante este período de tiempo es el Censo de 1950. Así, en promedio, el analista utiliza datos con 9 años de antigüedad.

Se usarán los datos de Venezuela como ilustración de lo que pueden significar tales errores. Usamos términos relativos más bien que absolutos, dado que es irreal suponer que un analista utilizará los datos censales sin procurar hacer un intento de ajuste de la creciente población.

El promedio anual de errores en la década ha sido calculado sobre dos bases. La primera supone que los Censos de 1950 y 1960 fueron ejecutados sobre bases del 100 por ciento y, consecuentemente, no hay errores de muestreo. Sin embargo, se considera que los datos de 1960 no están disponibles sino hasta después de 1964. La segunda supone que tanto el Censo de 1950 como el de 1960 fueron ejecutados sobre una base de muestra, y que el tamaño de los errores de muestreo son del orden de magnitud de los errores que se muestran en la Tabla 2. Así, hay un componente adicional de error durante toda la década. Con todo, la segunda base supone también que los datos de muestra están disponibles para el final de 1962, de modo que los dos últimos años no están sujetos a los mismos errores de extrapolación que los datos del censo total.

La Tabla 6 muestra tales errores para las estadísticas seleccionadas. Los resultados enseñan que si el muestreo reduce a la mitad el tiempo que se gasta en la publicación, y si uno supone que los datos censales se utilizan constantemente a lo largo de una década, entonces los errores actuales producidos al utilizar los datos serán en promedio menores con los datos de muestra que con el censo total.

La razón para ello es que el muestreo introduce solamente modestas adiciones al error total, mientras que en los últimos años de la década el error resultante de la utilización de los datos del Censo de 1950 es considerable, y reemplazando los datos estadísticos de 1950 por los de 1960 se reduce apreciablemente el promedio de error.

Desde luego, el análisis descrito arriba depende de los métodos utilizados para extrapolar los datos de 1950. Con técnicas más sofisticadas

se puede reducir probablemente el promedio de errores. Por otra parte, los errores de muestreo utilizados en el análisis son los que se presentan como consecuencia de la muestra relativamente pequeña usada para las tabulaciones preliminares (4 por ciento en Caracas y 1 por ciento en el balance de Venezuela). Desde luego, la utilización de una muestra grande reducirá significativamente los errores de muestreo: por ejemplo, una muestra del 20 ó 25 por ciento. Además, con una muestra grande, los datos de áreas pequeñas quedarán sujetos solamente a moderados errores de muestreo. Sospecho que los cambios en el curso de una década en áreas locales serán verosímilmente mucho mayores que los que pueda presentar la totalidad del país, y superiores los beneficios del muestreo.

Exactitud de las Muestras en Comparación con el Censo Total: A partir del Censo de 1950, la Oficina del Censo de los Estados Unidos ha confiado cada vez más en el muestreo para la recolección y tabulación de los datos básicos de los censos. La opinión predominante en la Oficina es que lejos de tener un efecto deletéreo sobre la calidad, el muestreo probablemente ha mejorado las estadísticas. Debido a la reducción en los costos de recolección y tabulación dispone de mayores recursos para dedicarlos a otros aspectos que afectan la calidad, tales como detenido adiestramiento y supervisión de los empadronadores, control de la calidad de trabajo de campo, conducción de la investigación sobre metodología hacia la mejora de otros procedimientos, etc. Nuestras evaluaciones de la calidad de censos recientes indican que, para muchísimos tipos de datos, estas mejoras compensan la adición de la varianza de muestreo de modo que los errores totales son menores que si se hubieran utilizado técnicas recientes en censos completos. Solamente en áreas pequeñas, tales como ciudades y espacio pequeños, el muestreo produce disminución mensurable en la exactitud, pero aun aquí las reducciones son realmente modestas.

Parece que otros países han hecho el mismo enfoque. El año pasado Inglaterra se levantó una muestra del 10 por ciento del censo. La tabulación del Censo Japonés de 1965, aparentemente le pondrá mucha confianza a una muestra del 20 por ciento. En el más reciente censo de Israel, muy ceñido al modelo de los Estados Unidos, se indagó el 100 por ciento de los datos solamente en un limitado número de renglones, y el balance se basó

en la muestra. Un número de países europeos utilizó las muestras para algunas de sus detalladas tabulaciones, a fin de reducir los costos de tabulación y codificación. Resulta interesante observar que el uso más extenso del muestreo se llevó a cabo en los países más industrializados y ricos. Los países que tienen más facilidades para proporcionar actividades estadísticas detalladas están confiando en la muestra para obtener resultados más pronto y reducir los costos.

Es conveniente revisar el razonamiento que conduce al juicio de que el muestreo no aumenta apreciablemente los errores en los datos, sin tener en cuenta el efecto de la reducción en el lapso para publicar los datos. Al analizar los efectos de la utilización de las muestras, es falso considerar que los censos completos están exentos de error y que las diferencias entre un censo completo y un muestreo reflejan las diferencias entre la verdad y la aproximación. Se reconoce ahora generalmente que en las mediciones, en general, es inalcanzable la perfección, y el campo económico y social no es precisamente la excepción. Las respuestas o mediciones de tales asuntos se obtienen solamente con diferentes grados de exactitud. A ésto se sobrepone el hecho de que frecuentemente los datos del censo están incluídos en análisis en los que las definiciones censales no son exactamente lo que requieren los usuarios, sino que se utilizan porque no hay otras fuentes de información disponibles.

En el análisis del aprovechamiento de los datos censales y de los requerimientos de exactitud que implica dicha utilización, es necesario considerar los totales errores cuadrados, promedios estimados de las estadísticas como base apropiada para el juicio, y no exactamente los errores de muestreo. La información disponible indica que para muestras del orden del 25 por ciento de la población, la varianza del muestreo es frecuentemente un componente pequeño del error total.[2/] Además, ésto refleja

---

2/ Ver Hansen, Morris H.; Hurwitz, Wm. N.; y Bershad, Max A.: Medición de Errores en Censos y Encuestas, Bulletin, International Statistical Institute, 38 (2), 1961. También, La Exactitud de las Estadísticas Censales Con y Sin Muestreo, Bureau of the Census Technical Paper No. 2, Washington, D. C., 1960.

solamente el efecto de la medida de los errores, e ignora el componente adicional del error no estreable resultante de los datos extrapolados fuera de tiempo, y acomodado en las definiciones.

Las experiencias del reciente Censo de los Estados Unidos presentan relevantes datos a este propósito. La Oficina del Censo de los Estados Unidos dirigió amplias investigaciones en unión con los Censos de Población y Vivienda de 1950 y 1960 para suministrar medidas de respuesta de sesgo y varianza en las operaciones censales. Estos componentes de errores surgen de las diferencias de interpretación por parte de los empadronadores y también de quienes responden a las preguntas que se hacen, de la carencia de conocimiento de la respuesta correcta de quienes responden, de los errores aleatorios o malentendidos de las instrucciones por parte de los codificadores, de los errores de transcripción, y de factores similares. En un censo enumerativo hay también componentes de varianza altamente significativos resultantes de las actividades de los empadronadores - cada empadronador puede introducir ciertos tipos de errores de mayor o menor consistencia o correlación en el trabajo.

Utilizando los resultados de estos estudios empíricos como estimación de la respuesta de sesgo y varianza (tanto la varianza de simple respuesta como la que proviene del efecto de correlación del empadronador) la estimación se prepara del total error cuadrado promedio que resultaría del uso de diferentes tamaños de muestra en un censo. La Tabla 7 presenta tales datos, basados en la experiencia de 1950. La respuesta de sesgo y varianza tiende a variar grandemente entre distintos renglones recolectados en el censo, y también de una a otra área, y depende en parte de la capacidad de los empadronadores empleados. Por tal razón la Tabla 7 presentará una vista general de órdenes de magnitud de los errores totales, más bien que una estimación ceñida. Es conveniente considerar las implicaciones de esta Tabla. Por ejemplo, para una característica que describe cerca de 500 personas en una área de 2500 personas, el aumento en la raíz cuadrada del error total promedio resultante de la variabilidad del muestreo es solamente de alrededor del 25 por ciento. Para áreas mayores y números más grandes, el error relativo adicional introducido por el muestreo es todavía más pequeño. Por ejemplo, una estimación de 10.000 en una área de 50.000 tiene un aumento de sólo 5 por ciento en la raíz cuadrada del

error medio a través del muestreo. Por consideraciones de esta naturaleza, se llegó a la decisión de introducir el muestreo en una mayor extensión en el Censo de Población y Vivienda de 1960 que en cualquiera de los anteriores, y utilizar las economías en fines que parezcan más productivos. La experiencia actual presenta al muestreo en una posición todavía más favorable que la señalada en la Tabla 7.

Dicha Tabla supone un coherente nivel de sesgo y varianza de respuesta, independiente del método de recolección. En la práctica, hay una reducción en el origen de los errores provenientes de las mejoras en los procedimientos y un mayor énfasis en el campo de la supervisión y control, lo cual es posible debido a los menores costos que resultan de la utilización del muestreo. Por éllo, el uso del muestreo hace más dignas de confianza a muchas estadísticas publicadas, que si se hubieran obtenido de otro modo.

Los problemas asociados con errores de respuesta y no relacionados en otras encuestas y censos han sido objeto de muchos estudios e informes en los últimos años. Al revisarlos, es difícil eludir el juicio de que el anterior análisis es razonablemente típico, y que el uso de grandes muestras tiene solamente un efecto pequeño en la exactitud de las estadísticas, aun para áreas pequeñas como terrenos y ciudades pequeñas. En general, el factor dominante que afecta la confiabilidad es el error de informe.

IV. Recomendaciones para un Programa Censal

Creo que sería provechoso sugerir una considerable expansión en el uso de las muestras al planear un censo. Las ventajas para las naciones menos industrializadas pueden ser todavía mayores que para otras más desarrolladas. Los dos principales aspectos de ésto son los siguientes:

A. Recolección de la muestra o procesamiento de los datos básicos.

1. Habrá un estudio sobre el valor real de los recursos disponibles y de su efecto sobre el lapso entre la obtención de las estadísticas y la publicación de los resultados; los objetivos del censo, en particular los relacionados con la necesidad de datos de áreas pequeñas, y la calidad de los datos accesibles, incluso

en las mejores circunstancias.

2. Solamente un mínimo sencillo de datos se considerará como base 100 por ciento - por ejemplo, edad, sexos, y posiblemente otras dos o tres particularidades importantes necesarias para las posibles áreas mínimas.

3. Para las estadísticas restantes, se establecerá una mínima tasa de muestra que reuna los objetivos estadísticos de las áreas pequeñas y que se puedan procesar en forma razonablemente rápida. Las unidades del muestreo serán pequeñas unidades tales como hogares, con disposiciones especiales para el muestreo en sitios que normalmente no están preparados para habitación, tales como instituciones.

4. El método de selección de la muestra debe ser susceptible de supervisión y control razonablemente ajustados. Si es necesario, puede ser conveniente considerar la recolección de toda la información sobre la base del 100 por ciento, pero para procesar solamente una muestra seleccionada en las oficinas.

5. Algunos de los recursos ahorrados por la utilización del muestreo deben dedicarse a adiestramiento, supervisión inmediata, más tabulaciones, investigación en los sistemas para mejorar metodológicamente el censo, etc.

6. Las tabulaciones del 100 por ciento de los datos deben ser simples, de modo que no se distraigan demasiados esfuerzos del procesamiento de la muestra de información. Se debe tener en cuenta la restricción de los cómputos del 100 por ciento, de manera que se puedan producir con sencillas cuentas anuales, posiblemente en la oficina del campo.

B. Tabulaciones Preliminares

1. Para decidir si son necesarios o no datos preliminares se debe tener en cuenta la disponibilidad de datos de muestra intercensal de otras encuestas.

2. Hay que reconocer que el programa es provisional, y la muestra se debe destinar para suministrar datos únicamente para el país como un total y posiblemente para unas pocas subdivisiones

principales.

3. Los planes para los datos preliminares, métodos de muestreo, etc., se deben determinar por anticipado, de manera que el programa pueda avanzar rápidamente.

4. Debe observarse que una ventaja - subproducto de las tabulaciones preliminares - es la de que sirven como prueba de los procedimientos y programas. El programa suministra también una muestra conveniente que se puede utilizar para análisis especiales.

Tabla 1: Fechas de Publicación de las Características de la Población Correspondiente a los más Recientes Censos en América Latina

País	Fecha del Censo	Fechas de publicación de los datos preliminares basados en la muestra	Fechas de publicación final, datos 100%
Argentina	Sept. 1960	1963	1965
Brazil	Sept. 1960	1965	No
Chile	Nov. 1960	1962-1963	1964
Colombia	Jul. 1964	No	No
Costa Rica	Abril 1963	1964	1966
Ecuador	Nov. 1962	1964	1964
El Salvador	Mayo 1961	1962	1965
Guatemala	Abril 1964	1966	No
Honduras	Abril 1961	1962	1964
México	Junio 1960	No	1962-1964
Nicaragua	Mayo 1962	1964	1964
Panamá	Dic. 1960	No	1962-1965
Paraguay	Oct. 1962	No	1965
Perú	Jul. 1961	1963	1966
Rep. Dominicana	Agto. 1960	1966	No
Uruguay	Oct. 1963	1964	No
Venezuela	Feb. 1961	1963	1966

Fuente: Tabulación inédita elaborada por el Instituto Interamericano de Estadística

Tabla 2: Estadísticas Seleccionadas de Población Publicadas en la Muestra Preliminar y en los Informes Finales; Censo de Venezuela de 1961

(Los datos preliminares se basan en una muestra de hogares que abarca el 4 por ciento de los hogares en Caracas y el 1 por ciento del balance de Venezuela. Los datos preliminares fueron entregados en 1963; los datos finales en 1966)

Renglón	Muestra Preliminar	Informe Final
Total hombres	3.824.000	3.822.000
Total mujeres	3.700.000	3.702.000
Hombres menores de 1 año	148.000	156.000
Mujeres menores de 1 año	153.000	151.000
Hombres 25-29 años	286.000	283.000
Mujeres 25-29 años	268.000	268.000
Hombres de 85 años y más	4.675	5.808
Mujeres de 85 años y más	13.850	11.494
Hombres jefes de familia	1.017.000	1.009.000
Mujeres jefes de familia	337.000	335.000
Hombres que residen en habitación especial	208.000	224.000
Mujeres que residen en habitación especial	90.000	94.000
Hombres nunca casados, 15 años y más	964.000	944.000
Mujeres casadas con su consentimiento	413.000	412.000
Hombres económicamente activos	1.957.000	1.929.000
Hombres sin empleo	287.000	274.000
Hombres no incluídos entre los obreros	600.000	567.000
Mujeres no incluídas entre los obreros	2.044.000	2.007.000

Tabla 3: Estadísticas Seleccionadas de Población como se Publicaron en la Muestra Preliminar y en los Informes Finales; Censos de Costa Rica, 1963.

(Los datos preliminares se basan en el 5 por ciento de la muestra de hogares. Los datos preliminares fueron entregados en 1964; los datos finales en 1966)

Renglón	Muestra preliminar	Informe final
Hombres de menos de 1 año	25.300	26.000
Mujeres de menos de 1 año	26.100	25.200
Hombres de 25-29 años	41.900	41.300
Mujeres de 25-29 años	43.000	43.500
Hombres de 75 años y más	7.478	7.383
Mujeres de 75 años y más	7.553	7.822
Jefes de hogar	231.300	231.200
Esposas de jefes de hogar	171.200	170.700
Hombres alfabetas de 10 años y más	369.800	371.300
Mujeres alfabetas de 10 años y más	371.900	374.200
Hombres analfabetas de 10 años y más	61.200	60.900
Mujeres analfabetas de 10 años y más	65.500	63.600
Empleados	371.500	367.800
Sin empleo	28.400	27.500

Tabla 4: Estadísticas Seleccionadas de Población como se Publicaron en la Muestra Preliminar y el Informe Final. Censo de El Salvador de 1961

(Los datos preliminares se basen en la muestra del 5 por ciento de los hogares y en la muestra del 10 por ciento de las habitaciones especiales. Los datos preliminares fueron entregados en 1962; los datos finales en 1965)

Renglón	Muestra preliminar	Informe final
Hombres menores de 1 año	50.800	49.100
Mujeres menores de 1 año	48.900	48.100
Hombres de 25-29 años	81.800	80.900
Mujeres de 25-29 años	93.800	91.600
Hombres de 75 años y más	12.900	13.000
Mujeres de 75 años y más	17.100	16.200
Hombres nunca casados, 14 años y más	308.900	310.600
Mujeres viudas	47.000	45.100
Empleados	763.400	766.000
Sim empleo	43.200	41.100

Tabla 5: Características Seleccionadas de la Población como se Publicaron en la Muestra Preliminar y en los Informes Finales, Censo de Estados Unidos en 1950

(Datos preliminares basados en la muestra del 0.1 por ciento de las personas. Los datos preliminares fueron entregados en 1951; los datos finales en 1952 - 1953)

Renglón		Muestra preliminar	Informe final
TOTAL DE ESTADOS UNIDOS			
Nacimientos			
Blancos nativos		125.068.000	125.191.000
Blancos nacidos en el exterior		10.147.000	10.184.000
Estado Civil (personas de 14 años y más)			
Hombres:	Solteros	14.398.000	14.518.000
	Casados	37.238.000	37.400.000
	Viudos o Divorciados	3.287.000	3.394.000
Mujeres:	Casadas	37.338.000	37.504.000
	Solteras	11.623.000	11.454.000
	Viudas o Divorciadas	8.030.000	8.084.000
Residencia hace 1 año			
La misma casa		119.677.000	119.190.000
Diferente casa, el mismo país		16.175.000	16.476.000
Diferente país o en el exterior		9.175.000	9.075.000
Ejército Nacional		2.337.000	2.422.000
Obreros (personas de 14 años y más)			
Total:	Obreros civiles	58.735.000	59.072.000
	Empleados	55.843.000	56.239.000
	Sin empleo	2.892.000	2.832.000
	Total no obreros	52.323.000	52.300.000
	Mujeres no obreras	40.668.000	40.542.000
No blancos:	Obreros civiles	5.848.000	6.078.000
	Empleados	5.355.000	5.602.000
	Sin empleo	492.000	476.000
	Total no obreros	4.808.000	4.876.000
	Mujeres no obreras	3.440.000	3.599.000

Tabla 5: Continuación

Renglón		Muestra preliminar	Informe final
NEW JERSEY			
Estado Civil (personas de 14 y más años)			
Hombres:	Solteros	477.000	484.000
	Casados	1.185.000	1.252.000
	Viudos o divorciados	110.000	103.000
Mujeres:	Solteras	457.000	412.000
	Casadas	1.236.000	1.259.000
	Viudas o divorciadas	280.000	260.000
Residencia			
La misma casa		4.126.000	4.118.000
Diference casa, el mismo país		341.000	325.000
Diferente país o en el exterior		234.000	225.000
Ejército Nacional		56.000	64.000
Obreros			
Obreros civiles		2.010.000	2.068.000
Empleados		1.900.000	1.963.000
Sin empleo		110.000	105.000
Total no obreros		1.692.000	1.668.000
Mujeres no obreras		1.349.000	1.314.000

Tabla 6: Comparación del Promedio Anual de Errores de un Censo Completo y de un Censo de Muestra, Estadísticas Seleccionadas de Venezuela

(Para el promedio de errores se supone que solamente están disponibles los datos del censo completo de 1950, mientras que los datos de muestra están disponibles durante los últimos dos años. Los promedios se calcularon como promedio de las diferencias anuales entre las cifras de 1950 y 1960, y las tendencias lineales desde 1950 hasta 1960)

Renglón	Datos según se informa en los Censos		Cambio del promedio anual	Promedio de errores 1955 - 1964	
	1950	1960		Censo completo	Muestra
Porcentaje de población menor de 1	3.6	4.1	+0.05	.45	.42
Porcentaje de población 25.29	8.0	7.3	-0.07	.63	.52
Porcentaje de hogares encabezados por una mujer	24.7	24.9	+0.02	.18	.18
Tamaño promedio del hogar	5.7	5.6	-0.01	.09	.11
Porcentaje de hombres de 10 y más años en el obrerismo	80.6	77.3	-0.33	2.97	2.90
Porcentaje de todas las mujeres no obreras	57.1	54.8	-0.23	2.07	1.90

Tabla 7. Errores esperados (RCEM) de las repeticiones de casilla estimadas por los renglones individuales basados en un censo completo, y en una muestra del 25 por ciento de los hogares

Repetición de la casilla	Area de 2.500 personas que tiene RCEM basado en --		Repetición de la casilla	Area de 10.000 personas que tiene RCEM basado en --		Repetición de la casilla	Area de 50.000 personas que tiene RCEM basado en --	
	Censo completo	Muestra del 25 por ciento		Censo completo	Muestra del 25 por ciento		Censo completo	Muestra del 25 por ciento
12	7	10	50	14	20	250	34	46
50	11	19	200	30	40	1.000	85	105
125	22	31	500	52	67	2.500	180	200
500	49	62	2.000	140	160	10.000	620	650
1.250	89	102	5.000	320	330	25.000	1.520	1.530

Nota 1: Los cálculos suponen una respuesta relativa de sesgo del 6 por ciento y una varianza de respuesta igual a la varianza de muestreo para una muestra del 25 por ciento de los hogares.

Nota 2: La exactitud de los resultados (repeticiones de casilla) se mide por una cierta clase de promedio de los errores actuales que puedan ocurrir, la raíz cuadrada del error medio (RCEM). Una regla útil de trabajo será suponer que aproximadamente dos tercios de todos los resultados de un censo o de una muestra discreparán de sus verdaderas repeticiones de casilla no más que por la raíz cuadrada del error medio (RCEM)

FUENTE: La Exactitud de las Estadísticas Censales Con y Sin Muestreo, Oficina del Censo de los Estados Unidos, Estudio Técnico No. 2, tabla 15.

REFERENCIAS

(1) Eckler, A. Ross y Hurwitz, Wm. N., "Respuesta, Varianza y Sesgos en Censos y Encuestas", Bulletin, International Statistical Institute, 36 (2). 1958.

(2) Hansen, Morris H.; Hurwitz, Wm. N.; y Bershad, Max A.: "Errores de Medición en Censos y Encuestas". Bulletin, International Statistical Institute 38 (2) 1961.

(3) Hansen, Morris H.; Hurwitz, Wm. N.; y Pritzker, Leon, "La Estimación e Interpretación de las Grandes Diferencias y la Varianza de Simple Respuesta". Contribuciones a las Estadísticas Presentadas al Prof. P. C. Mahalanobis, Pergamos Press, Calcuta, 1964.

(4) Hanson, Robert H. y Marks, Eli S., "La Influencia del Entrevistador en la Exactitud de los Resultados de la Encuesta", Journal of Ameircan Statistical Association, Vol. 53, No. 283, Sept. 1958.

(5) Taeuber, Conrad y Hansen, Morris H., "Una Evaluación Preliminar de los Censos de Población y Vivienda 1960", Bureau of the Census Technical Paper No. 13, Washington, D. C., 1965.

(6) Waksberg, Joseph y Hanson, Robert H., "Aplicaciones del Muestreo en los Censos de Población y Vivienda", Bureau of the Census Technical Paper No. 13, Washington, D. C., 1965.

(7) Zarkovich, S. S., "Calidad de los Datos Estadísticos", Food and Agriculture Organization of the United Nations, Roma, 1966.

(8) Zarkovich, S. S., "Métodos de Muestreo y Censos", Food and Agriculture Organization of the United Nations, Roma, 1965.

(9) U. S. Bureau of the Census, "Muestreo en el Censo de Población y Vivienda de 1950", Bureau of the Census Working Paper No. 4, Washington, D. C., 1956.

(10) U. S. Bureau of the Census, "La Exactitud de las Estadísticas del Censo Con y Sin Muestreo", Bureau of the Census Technical Paper No. 2, Washington, D. C., 1960.

(11) U. S. Bureau of the Census, "La Encuesta de Empadronamiento por Correo", Bureau of the Census Technical Paper No. 4, Washington, D. C., 1960.

(12) U. S. Bureau of the Census, "Programa de Evaluación e Investigación de los Censos de Población y Vivienda de los Estados Unidos", Series ER60, Nos. 1 a 6, Washington, D. C., 1963-1965.