

Documento de Referencia N° 4

Septiembre de 1989

ORIGINAL: ESPAÑOL

C E P A L

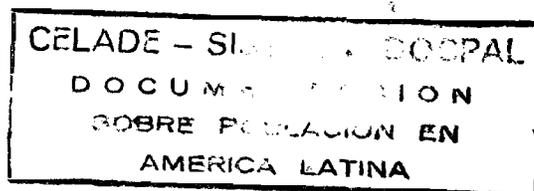
Comisión Económica para América Latina y el Caribe

Seminario Sistemas Computacionales para el Acceso de
Usuarios a la Información Censal

Santiago, 12 al 15 de septiembre de 1989

LA DIFUSION DE INFORMACION DEL CENSO DE 1990
DE LOS ESTADOS UNIDOS DE NORTEAMERICA */

*/ Este documento fue preparado por la Oficina de Censos de los Estados Unidos de Norteamérica, en el marco del Proyecto sobre Apoyo al Diseño y Preparación a la Ronda de los Censos del 90, administrado por la Comisión Económica para América Latina y el Caribe, CEPAL y financiado por el Fondo de las Naciones Unidas para actividades en materia de población.



La Difusión de Información del Censo de 1990 de los Estados Unidos

Presentación de JOHN KAVALIUNAS, Jefe de la Oficina de Adiestramiento, División de Servicios a los Usuarios de Datos, Oficina del Censo de los Estados Unidos

Introducción

En 1990, los Estados Unidos celebrarán el bicentenario de un acontecimiento arraigado en la Constitución. Según indicado en este documento, el nuevo país levantó un censo con el propósito primordial de determinar la representación popular en el Congreso. Cada 10 años, sin interrupción por 200 años, el censo ha documentado el aumento y el movimiento de la población americana. Además, preguntas adicionales a través de las décadas, han medido cambios en la sociedad americana y el censo hoy en día provee no sólo información para uso gubernamental, sino también información muy útil para la empresa privada, autoridades locales, investigadores y un sinnúmero de usuarios.

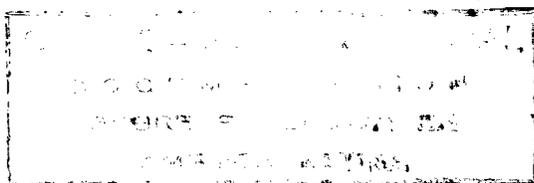
Quiero hablar hoy de nuestro programa de difusión para el censo de 1990. Para evaluar y mejorar nuestros productos, hemos consultado a muchos de los usuarios de nuestros datos. Al fin y al cabo, los productos no son tan diferentes de los del censo del 80. Pero, hemos introducido un nuevo medio --el compact disc-- y pretendemos poner más énfasis en el acceso directo. Hemos cambiado el método de protección de la confidencialidad de información y creemos que los usuarios tendrán información útil más accesible. Además tenemos una sólida red de distribuidores secundarios que nos ayudarán a difundir la información y a asistir a los usuarios primarios.

La Planificación de los Productos

Durante 1984 y 1985, llevamos a cabo 65 reuniones locales en las capitales estatales y en otras ciudades importantes con los usuarios de nuestra información. Se incluye en este grupo a investigadores, planificadores, bibliotecarios, periodistas, y otros. Les preguntamos sobre la utilidad de la información brindada y sobre el nivel de desagregación prospectivo necesario. Además, realizamos 10 conferencias en varias partes del país sobre los medios de difusión y el contenido de informes y cintas. Nos dijeron que no les gustaba mucho la microficha, y aunque debíamos explorar los nuevos medios de difusión, las publicaciones son y continuarán siendo el medio preferido tanto para el usuario novicio como para el experimentado.

Los usuarios también abordaron el tema de la falta de información para algunos grupos o sectores debido a la utilización de ciertos métodos para mantener la información confidencial.

Empezamos hace poco otro proyecto que le permite a los usuarios repasar las especificaciones de las tablas propuestas, o sea el contenido de las cintas y de los



informes. Mandamos estos bosquejos a un grupo de aproximadamente 410 organizaciones y usuarios particulares.

Producción de los Productos

En los Estados Unidos, se levanta el censo por correo. Mas del 90% de los hogares reciben un formulario por correo unos días antes del Día Censal, el 1 de abril, el cual debe ser devuelto por correo. En las áreas más rurales o donde no existe una dirección exacta, un enumerador hace una visita personal al hogar.

Hay unas preguntas básicas en el formulario sobre cada persona y cada vivienda. Esta es el formulario corto dirigido al 100% de la población. Una muestra del 19% recibe un formulario más extenso con preguntas adicionales sobre educación, ingresos, ocupación, y gastos mensuales de la vivienda. Por primera vez, realizaremos un muestreo diferencial de acuerdo con la población del área geográfica estimada, o sea una muestra del 12 al 50 por ciento.

Primero calculamos las respuestas a las preguntas del 100% y preparamos los archivos y los informes. Después, desarrollamos los productos basándonos en la muestra.

Hay unos productos que contienen solamente información del 100 por ciento, y otros que contienen solamente información de la muestra. Hay unos que combinan las dos fuentes de información. No publicamos información de la muestra para áreas pequeñas.

Primero preparamos las cintas con los archivos electrónicos. Generalmente, los archivos electrónicos están disponibles al público unos dos o tres meses antes de las publicaciones. Las microfichas tardan más tiempo.

Los Productos

Nuestra oficina publica de 2,500 a 4,500 volúmenes por año. El censo de 1980 produjo 1,400 informes, con unas 300,000 páginas, o sea una cantidad de informes tan grande que si se pusiera uno encima del otro, la torre llegaría hasta el 5o piso.

En 1965 se introdujo la cinta magnética como medio de difusión de la información del censo del 60. En los últimos años el uso de nuevos medios electrónicos se estableció firmemente como un medio de difusión. En 1984 añadimos dos medios nuevos: el servicio de datos en línea y los disquetes.

Aunque había mucha información en cinta magnética, los archivos nuestros eran tan grandes, que eran pocos los usuarios que tenían máquinas y recursos necesarios para procesar la información. No había mucha información en un disquete, y para descargar un archivo grande de una base de datos en línea se necesitaban horas o días. Además los usuarios tales como bibliotecas, autoridades locales, y aún nuestras oficinas regionales, no tenían acceso a la información. En 1986, comenzamos a evaluar una

nueva tecnología que nos pareció ser muy apropiada: el CD-ROM (Compact Disk-Read Only Memory).

En un disco cabe la información de 3 a 4 carretes de cinta. Los usuarios con una computadora personal y equipo para leer el disco pueden tener acceso a información que antes solamente estaba disponible en forma de montones de listados. Para evaluar su aceptación en el mercado, desarrollamos un disco experimental con datos de una variedad de censos y encuestas. La demanda de este producto fue tan grande -- el año pasado vendimos 223 copias, y distribuimos gratis unos 100 copias más -- que incorporamos este medio en nuestros planes de difusión de la información de los censos económicos de 1987 y del censo de población y vivienda de 1990.

Archivos Electrónicos

Proveemos 15 veces más información censal en cinta magnética que en los informes. Por ejemplo, en una publicación puede aparecer el porcentaje de personas menores de 18 años de edad y el porcentaje de personas de más de 65 años de edad. El archivo del cual se extrae esta información nos da el número de personas por edad y por sexo en grupos quinquenales de edad.

Desarrollamos cuatro series de archivos electrónicos o STFs (Summary Tape Files-Cintas de Archivos Sumarios): dos con información del 100 por ciento, y dos de la muestra. Cada serie consiste de 3 o más subarchivos. Las tablas dentro de cada serie son iguales, lo que varía es la geografía. (Vea el cuadro más abajo.) Por ejemplo: STF 3 contiene 300 tablas de información de la muestra. El archivo 3A contiene información para grupos de manzanas, sectores censales, ciudades, condados y para el estado. Hay un archivo de 1 a 3 carretes de cinta de alta densidad para cada estado.

El archivo 3B contiene las mismas tablas para cada código postal y es un archivo de 10 carretes. El STF 3B del censo del 80 fue uno de los archivos más populares y uno de los que más se compró.

El archivo 3C es un archivo sumario con información para la nación, todos los estados, condados, y ciudades grandes.

También hay archivos en esta serie para distritos legislativos.

Archivos de Microdatos

En 1980 uno de los productos más populares y más comprados fue la serie de archivos de microdatos, esto es, un archivo con las respuestas verdaderas de un porcentaje de las viviendas que recibieron el formulario censal no básico. Es un producto útil porque permite que el usuario haga las tabulaciones cruzadas según las necesite. Así, si un investigador quiere estudiar las características de, digamos, profesionales de la salud, puede recorrer este archivo, sacar los registros de personas con esta ocupación, y

Programa de Cintas Para Computadora del Censo de 1990

<u>Archivo</u>	<u>Nivel Geográfico</u>	<u>Contenido</u>
STF 1A	Grupos de manzanas	Información básica (Preguntas del 100%)
STF 1B	Manzanas	Información básica
STF 1C	Nación, estados, condados/ciudades grandes	Información básica
STF 1D	Distritos legislativos	Información básica
STF 2A	Sectores censales	Tabulaciones cruzadas de la información básica
STF 2B	Condados/ciudades	Tabulaciones cruzadas
STF 2C	Nación, estados, condados/ciudades grandes	Tabulaciones cruzadas
STF 3A	Grupos de manzanas	Información de la muestra
STF 3B	Códigos postales	Información de la muestra
STF 3C	Nación, estados, condados/ciudades grandes	Información de la muestra
STF 3D	Distritos legislativos	Información de la muestra
STF 4A	Sectores censales	Tabulaciones cruzadas de la información de la muestra
STF 4B	Condados/ciudades	Tabulaciones cruzadas de la información de la muestra
STF 4C	Nación, estados, condados/ciudades grandes	Tabulaciones cruzadas de la información de la muestra

hacer sus propios cálculos de nivel de educación, nivel de ingresos, y características de la familia, o lo que necesite.

Tomamos medidas para que nadie pueda identificar personas o viviendas. Quitamos toda clase de identificación personal como nombre, apellido, número de identificación, y dirección. Cada registro está asignado a un área preidentificada de más de 100,000 habitantes. El área puede ser una ciudad, un condado, o grupos de condados que tienen este número de habitantes. También, repasamos los registros para ver si hay valores extraños, por ejemplo ingresos, edad, etc., que puedan inadvertidamente identificar a alguna persona o sector.

En 1980 proporcionamos tres archivos de microdatos:

El Archivo A contenía una muestra del 5% e identificaba a condados o grupos de condados

El Archivo B contenía una muestra del 1%; los registros llevaban la identificación del área metropolitana

El Archivo C contenía también una muestra del 1% y los registros llevaban un código que identificaba a las áreas urbanas y rurales.

Este producto fue lo que más se vendió durante la década--alrededor de 1,250 copias. Existen archivos de microdatos para los censos de 1970, 1960, 1950 y 1940. Es un producto muy útil para profesores y otros investigadores que quieren estudiar los cambios en la sociedad americana y la relación entre las variables socioeconómicas.

Para el censo de 1990, planeamos desarrollar dos archivos, similares a los Archivos A y B de 1980.

CD-ROM

Ya se produjeron varios discos con datos estadísticos y pretendemos poner los STFs 1 y 3, inclusive los subarchivos, en este medio. Sólo estos dos archivos requieren alrededor de 30 discos.

Bases de Datos de Acceso Directo

Hay muchas bases de datos disponibles a todos los organismos y personas con computadora personal y línea telefónica.

Aunque la oficina del censo no permite acceso directo a su base de datos, para satisfacer las demandas de los usuarios desarrollamos y mantenemos una base de datos por intermedio de dos empresas privadas que proveen un sistema mundial de datos en línea: DIALOG y COMPUSERV.

La base de datos -- CENDATA-- actualmente es pequeña, con un tamaño de 145 megabytes. Cada día cargamos al sistema nuevos anuncios de prensa, archivos pequeños nacionales, y las estadísticas corrientes más importantes, como las de comercio exterior. El sistema tiene alrededor de 3,000 usuarios por mes. Por ser tan grandes los archivos del censo de población y vivienda, no podemos poner en el sistema todos los archivos del censo del 90. Sin embargo, proponemos poner extractos de archivos con información nacional, estatal, y para áreas más pequeñas.

Mantenemos también un sistema de fichero electrónico para el uso de los organismos estatales que participan con nosotros en un programa de difusión secundaria. Este sistema reside en una computadora personal en nuestras oficinas y contiene información administrativa como también archivos pequeños con información estatal. El sistema recibe casi 900 llamadas al mes. Para el censo de 1990, tenemos planes de usar este sistema para anunciar la disponibilidad de archivos, informes, y otros productos, como también para proveer archivos pequeños e información actual.

Disquetes

Actualmente ofrecemos pocos productos en este medio. Sin embargo en 1988 por primera vez vendimos más disquetes que cintas. Por su capacidad, los disquetes no son un medio apropiado para la difusión de archivos grandes. No tenemos planes para proveer información del censo de 1990 en este medio, a no ser que aparezcan unos archivos pequeños extraídos de los archivos en cinta o CD-ROM.

Protección de la Información

Se repasa toda la información disponible al público, bien sea en informes o en medios electrónicos, para proteger la confidencialidad de las respuestas.

En los últimos censos aplicamos un sistema que llamamos de supresión. Esto es, si no hay suficientes personas o viviendas en un grupo con tales o cuales características, no publicamos la información. Para la información básica (las preguntas del 100%), el número era de 5 para las viviendas u hogares y de 15 para las personas. Para la información obtenida de la muestra, usamos 10 y 30.

Por ejemplo, si en una manzana hay 6 viviendas, proveemos datos sobre lo perteneciente a las viviendas, el valor o alquiler, el número de cuartos, etc. Pero si en la misma manzana solo hay 12 personas, o sea, menos de 15, se provee solamente el número total de personas, y no publicamos información sobre sus características.

También se aplicó un sistema de supresión complementaria, esto es para prevenir que no se pueda obtener un determinado dato por medio de una substracción. Por ejemplo, si hay un sector censal con 4 familias japonesas y 20 familias negras, no publicamos las características ni de las familias japonesas ni de las familias negras.

Los usuarios reclamaron que empleando este método, perdieron mucha información.

Investigamos varios métodos, incluyendo métodos usados por otros países para proveer información y, al mismo tiempo, proteger la privacidad del informante. En fin, el método propuesto se llama validación de la confidencialidad.

Seleccionaremos una muestra pequeña de los archivos internos. Aparearemos los registros seleccionados a otros registros con características semejantes en una área diferente. Los registros apareados tendrán unas características básicas en común: tamaño de hogar, raza, origen hispanico, número de personas con más de 18 años de edad, etc. Cambiaremos las demás características.

Los estadísticos nos aseguran que este sistema protegerá la confidencialidad como también proveerá información importante para los usuarios.

Informes, Microfichas y Listados

Los informes publicados se derivan de los archivos, o son, se puede decir, una muestra de toda la información que hay. Al igual que en las cintas, tenemos informes que contienen la información básica (las preguntas del 100 %) y otros que tienen información de la muestra. Hay otras series que combinan información de las dos fuentes. Tenemos series que tratan solamente de las características de la población; otras de las características de la vivienda; y aún otras que combinan información sobre la población y vivienda. Generalmente, en cada serie, hay un informe para cada estado y un volumen que es un resumen de la información para la nación. Pretendemos publicar del censo de 1990, más información por tema, esto es, un archivo y uno o dos informes relacionados con un asunto de interés social--por ejemplo, ingresos, características de los ancianos, características de casas rodantes, etc. Proporcionaremos información a nivel estatal y a nivel de ciudad/condado grande.

Desde 1968 hemos producido microfichas de cada informe publicado por nuestra oficina. Esto es un gran servicio para los usuarios que necesitan información histórica cuando no se tiene disponible la publicación porque está agotada. Para el censo de 1980, se introdujo la microficha como producto en sí mismo. Hay productos que existen solamente en microficha y también proporcionamos copias de listados en microfichas. Este último producto proporcionó por primera vez acceso a la información que existía antes solamente en forma electrónica. Continuaremos este programa en 1990.

Productos Geográficos

Para el censo de 1990, la Oficina del Censo designó manzanas para todo el territorio nacional, entre 6 y 7 millones de manzanas. Para organizar y clasificar esta geografía, desarrollamos un sistema de información geográfica llamado TIGER. La sigla representa: sistema de referencia y codificación topológicamente integrado. Debido al interés suscitado por este nuevo sistema, los usuarios encontrarán muchas aplicaciones. Cada registro en el sistema representa un segmento de una calle y contiene códigos para las áreas censales en los dos lados de la calle, como también la

latitud y la longitud del segmento. Los usuarios pueden entonces usar el sistema para cartografía computerizada, para calcular la distancia entre dos puntos, y otras aplicaciones. Ya están disponibles los archivos que son organizados por condado. Este sistema da por primera vez un inventario de la red nacional de calles y carreteras.

Otros Productos

Desarrollaremos también guías para los usuarios, documentación técnica para cada archivo, una historia de los procedimientos del censo, y otros manuales, guías de uso, programas de adiestramiento y folletos. Se propone también un nuevo producto: documentación técnica en disquete.

Difusión de la Información

La Oficina del Censo provee información directamente al público, pero más recientemente, hemos utilizados intermediarios. Tenemos en los Estados Unidos una red muy grande de distribuidores secundarios que está formada por bibliotecas, gobiernos estatales y municipales, y la empresa privada.

Por ley, la mayoría de nuestros informes se venden a través de la Imprenta del Gobierno. Pero la información en medios electrónicos la vendemos nosotros mismos.

Por más de 100 años, la Imprenta del Gobierno ha mandado copias de todo lo que imprime el gobierno a un grupo de bibliotecas selectas. Existe una red de más de 1,300 bibliotecas, incluyendo las más grandes bibliotecas públicas y universitarias. Hay también un grupo de 90 bibliotecas que reciben directamente de la Oficina del Censo, todos los informes nuestros. Hace poco tiempo, estas bibliotecas han empezado a tener acceso a la información electrónica. Por medio de la Imprenta del Gobierno, las bibliotecas repositorias pueden pedir los CD-ROMs de nuestra oficina.

Reconociendo la necesidad de facilitar acceso a información electrónica, desarrollamos un programa de divulgación con otros intermediarios públicos. Entramos en convenio con gobiernos estatales para proveer a una agencia estatal, designada por el gobernador, toda la información que producimos para ese estado. El estado, a su vez, tiene el deber de procesar y difundir la información, de informar al público, y de establecer otras fuentes de disponibilidad en el estado. Hoy día, el programa funciona en todos los 50 estados, el Distrito Federal, Puerto Rico, las Islas Vírgenes, y la Isla de Guam, y cuenta con más de 1,400 agencias estatales, universidades, bibliotecas, cámaras de comercio y juntas locales de planificación. Calculamos que la red responde a más de medio millón de solicitudes al año. Y el alcance del programa ha superado nuestros planes iniciales. Con la experiencia ganada al difundir la información, las agencias nos han ayudado a planificar los productos y a promover el censo.

Existen también varias empresas privadas que sirven como distribuidores secundarios. Estas compañías y consultores compran los datos nuestros, agregan valor, y los venden. Por ejemplo, desarrollan perfiles de áreas designadas por el cliente, hacen estimaciones y proyecciones para áreas pequeñas, o crean productos a medida para computadoras personales. Para fomentar esta actividad y proveer un servicio al público, mantenemos una lista de todas las compañías con información sobre los servicios que ofrecen; referimos a éstas muchas veces las solicitudes que nosotros no podemos atender. También, les mandamos gratis copias de nuestros boletines mensuales, el catálogo anual, y cualquier otra información ocasionalmente.

Resumen

Creemos que el Censo Bicentenario de los Estados Unidos será el más completo que se haya levantado en el país. Los productos que se desarrollarán después del censo responderán a las necesidades de los usuarios. Además de los productos tradicionales--informes, microfichas, y cintas--ofreceremos más información en nuevos medios, principalmente CD-ROM. Ya tenemos instalados en el país, una sólida y gran red de distribuidores secundarios que nos ayuda a atender las demandas y solicitudes del público.