

E L M A R C O M U E S T R A L  
=====

Autor: Carlos CAVALLINI

## EL MARCO MUESTRAL

---

1. Sistemas de Información. Las dificultades más imperiosas y urgentes, que han surgido a partir de mediados de la década del 70 en la mayoría de las Oficinas Nacionales de Estadística de América Latina, han estado asociadas, en grado significativo, por un lado, con el masivo y progresivo requerimiento de información, cada vez más sofisticada, que demandan los grupos, en creciente aumento, de usuarios, y por el otro, con las inflexibles limitaciones presupuestarias a que han sido sometidas dichas Oficinas.

Esta necesidad de información para la sociedad, que se pide satisfacer con menos disponibilidad de recursos y manteniendo aceptables niveles de confiabilidad, en cierta forma está amenazando los equilibrios tradicionales de trabajo dentro de los cuales se desarrollaba el orden estadístico.

Se arguye, y en cierta forma es ya un desafío casi irreversible, que la superación de este conflicto se encontrará a través de una transformación tecnológica de los métodos que actualmente se utilizan.

Estas transformaciones, a las cuales se deberán someter las Entidades Estadísticas para poder estabilizar con éxito las importantes funciones que les competen, deberá buscar sus bases, principalmente, en el fundamental papel que están desempeñando los Sistemas de Información.

2. La estadística, como la debemos comprender en su concepto más amplio, forma parte de la tetrada, que conjuntamente con la cibernetica, la temática y la taxonomía, apuntalan el enorme desarrollo que están teniendo las ciencias de la información.

Es a través de la óptima interconectabilidad de estos cuatro patrones técnicos que las Oficinas de Estadística lograrán garantizar la plena capacidad para mantener, a bajo costo, depósitos de datos, que permitan una transmisión fácil de la información requerida.

3. El concepto de Sistema ha invadido todos los campos de la ciencia y aún ha penetrado en el pensamiento y en el habla coloquial. Profesiones y ocupaciones hasta hace poco desconocidas, llevan su nombre, como proyecto de sistemas, análisis de sistemas, ingeniería de sistemas y así por el estilo. En su acepción general se lo liga a un conjunto de reglas o principios que regulan ordenadamente ciertas cosas para contribuir a un determinado objetivo. En una consideración matemática, un sistema puede ser definido como un complejo de elementos interactuantes donde deban hacerse tres tipos de distinciones, a saber: i) de acuerdo con su número; ii) de acuerdo con sus aspectos y iii) de acuerdo con las relaciones entre elementos. (1)

Por supuesto que existen diferencias considerables y manifiestas entre los distintos sistemas, sean estos propios de la física, o de la biología o de las ciencias sociales, pero todos concuerdan en que el proceso de su funcionamiento no es completamente accidental, sino que sigue regularidades o leyes que pueden ser determinadas.

4. Los Sistemas de Información, punto de nuestro interés, considerados como una red flexible de los servicios existente para dar una respuesta adecuada al crecimiento de la información, están jugando un papel

(1) Ludwig von BERTALANFFY, "Teoría general de los Sistemas", Fondo de Cultura Económica, Madrid, 1976

fundamental en el desarrollo de las naciones. Existe un estudio sobre la posibilidad de establecer un sistema mundial de información científica, UNISIST(2); El hecho más sobresaliente de encarar la transferencia de la información, principal objetivo de las Oficinas de Estadística, a través de un Sistema de Información constituirá, además de un aumento de capacidad, una mejor utilización de los recursos disponibles. Con la actual tendencia de los costos crecientes en el tratamiento de los datos, la información estadística se podrá convertir en un lujo que pocos usuarios se podrán permitir, a pesar de que esta necesidad de conocimiento se hace día a día más solicitada. Si no se revén los procesos vigentes, considerando los nuevos medios metodológicos y técnicos, que coordinados soporten a un Sistema de Información, el acceso, que debiera ser económico, fácil y rápido, y a una información confiable, por parte de una comunidad en constante crecimiento, se hará cada vez más difícil.

5. El estudio, que las Oficinas Nacionales de Estadística debieran llevar a cabo, para desarrollar un Sistema Integrado de Información Estadística, será beneficioso, no en sentido estrictamente financiero, si no, además, en eficiencia y en capacidad. El ahorro de dinero por unidad de medida de transferencia de información, puede ser sólo una de las motivaciones posibles para establecer un Sistema de Información. Por otra parte, este ahorro se logrará luego de un período inicial durante el cual deberán, en algunos casos, efectuarse gastos suplementarios.

Los elementos básicos para generar un sistema integrado estarán conformados y regulados por las disciplinas que hacen, principalmente, a las siguientes partes:

(i) Cibernética, cuya teoría conviene para describir la estructura formal de la transferencia de la información, mediante diagramas de bloques.

(2)UNISIST. Informe del estudio sobre la Posibilidad de Establecer un Sistema Mundial de Información Científica realizado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, y el Consejo Internacional de Uniones Científicas - UNESCO.

y de flujo, que hacen al método logicial.

ii) Telemática, que a diferencia de la electricidad, no transmite una corriente inerte, sino información. Principalmente vincula las computadoras con los archivos de datos, y ~~conforma~~ la red de entrada y salida de los mismos.

iii) Taxonomía, que tratará, como ciencia de la clasificación, de los problemas del ordenamiento y del dato símbolo, que como tal carece de significado, pero que dependerá de la lógica del proceso para transformarse en información.

iv) Estadística, cuya metodología debe contribuir a la construcción de los modelos y de las funciones paramétricas que soportará al Sistema, así como a la generación de diseños que se adecuen a las leyes que gobernarán al mismo.

6. Un sistema de información debe estar al alcance de todos los usuarios y paulatinamente, a través de un proceso evolutivo, se debe ir desarrollando, aprovechando plenamente los recursos y tendencias existentes. Una condición esencial para la creación del Sistema, es que todos los participantes potenciales que participen del mismo estén convencidos, no sólo de que su filosofía es sana, sino también, de que el presente esquema constituye una proposición realizable y ventajosa, que permitirá mantener en alto el nivel de producción de estadísticas oportunas y necesarias.

En el período inicial, las dimensiones del Sistema pueden estar determinadas estrictamente, en relación con un programa limitado de prioridades. Por ejemplo, creación de archivos de información y de marcos muestrales; de modelos paramétricos para la selección de muestras, para la entrada y salida de información, para el cálculo de estimaciones y tabulación de resultados, etc. En función de estos programas se

podrán establecer, desde un punto de vista funcional, los puntos focales, o terminales, para el tratamiento, distribución y consulta de la información, dado que las organizaciones gubernamentales, generalmente, tendrán una participación activa y creciente en el planeamiento y conducción del Sistema, considerando que el mismo será de importancia nacional para el progreso económico, estos puntos focales no sólo estarán ubicados dentro del área de la Oficina de Estadística, sino, que podrán situarse en otros organismos intergubernamentales, que así lo requieran, para sus actividades específicas. Asimismo, en un estudio de avance, las organizaciones no gubernamentales podrán ser consideradas para tener cabida dentro del Sistema.

Estos clientes periféricos, por así llamarlos, podrán estar dotados de "terminales inteligentes", que son pequeñas unidades de proceso, de potencia y capacidad limitadas, pero dotadas de capacidad de tratamiento autónomo. La importancia de estas terminales, es que llevan a la periferia una parte del tratamiento, aliviando a la unidad central de procesamiento de una parte del trabajo.

7. Por otro lado, y dado que muchos países ya cuentan con determinados modelos de sistemas, como son los sistemas de información académica, de información financiera, etc. se deberá estudiar la posibilidad de interconectar al conjunto de todos los Sistemas y reducir el conjunto a dimensiones más razonables.

De esta manera, toda la información que tiene un país, y que en sí no es nada, pues la información es información cuando alguien resulta informado, podrá activarse, interrelacionarse y ser de utilidad para una sociedad cada vez más ávida de conocimientos.

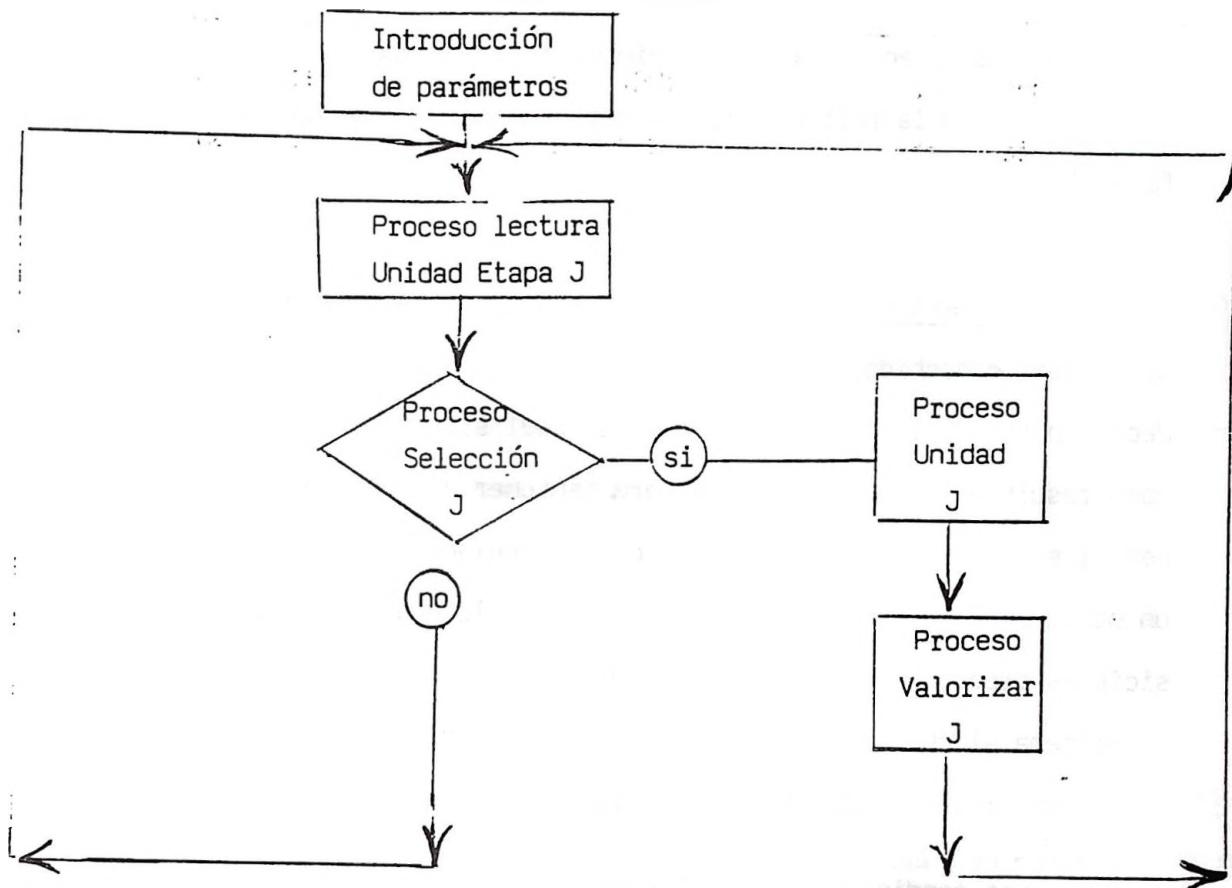
## 8. La Estadística como Soporte de los Sistemas de Información.

Un sistema Integrado de Información Estadística, estará basado en un conjunto de reglas y principios que regirán al mismo. (3) Una de las principales ramas del Sistema será aquélla que tenga relación con los programas de encuestas de hogares, dado que permitirá mantener la coordinación de dichos programas, con el consiguiente beneficio de obtener una integración mayor de información y por ende, de resultados. Esta integración, a través del tiempo y del espacio, llevará además, a la construcción de series cronológicas.

Uno de los archivos del Sistema estará constituido por el archivo metodológico, o soporte del sistema, el cual contendrá, básicamente, los modelos estadísticos y los programas de utilización.

Los modelos estadísticos principales estarán referidos, en primer lugar, i) a la selección de unidades muestrales; ii) al cálculo de las estimaciones y iii) al análisis de las variaciones. Los mismos, obedecerán al tipo de modelo paramétricos, que se activan en función de determinados parámetros y de determinados supuestos, que determinan al modelo. Un diagrama en bloque, para la selección de unidades muestrales en un número variable de etapas, puede responder, por ejemplo, a la siguiente lógica.

(3) Carlos CAVALLINI, "Principios de un Sistema Integrado de Procesamiento Muestral", E/CEPAL/R.338, 27 de mayo de 1983



Se deduce, que previamente al diseño logicial de los modelos paramétricos, se deberá investigar el isomorfismo de conceptos y leyes entre los distintos procesos que estructuran a las diversas etapas. Este estudio permitirá, además de condensar la programación, obtener provechosas transferencias de un proceso a otro, minimizar la repetición de esfuerzo teórico en diferentes etapas y estimular el desarrollo de modelos que se adecuen a estos principios.

9. Dato símbolo otro punto importante ha de corresponder a la fijación de los valores simbólicos que identificarán a las distintas unidades muestrales, a los campos y subcampos en que aquella se subdivisiona, a los códigos de variables y estratos, a las especificaciones de los distintos indicadores socioeconómicos y estadísticos, etc. Toda percepción exterior entra al sistema en forma simbólica y la creación de estos simbolismos debe ir munida de una ágil flexibilidad, que permite interca-

laciones sin alterar su correspondencia biunívoca dentro del sistema, y llegado el caso, entre sistemas. Símbolos inadecuados así como procesos inadecuados son la raíz de los más serios problemas de un sistema de información.

10. Retroalimentación. Por otro lado, todo proceso de un sistema debe estar concebido, de manera tal, que el sistema se autoregule. Es decir, parte de la información que sale del sistema es remitida de nuevo, como resultado de una respuesta para mantener a determinadas condiciones vigentes. Esta retroalimentación, o regulación homeostática, juega un papel importante en la concepción de un Sistema, y se basa en disposiciones preestablecidas. Así, por ejemplo, la información que provee el sistema al operar a una determinada variable, vuelve al mismo, con nuevos valores que actualizan a la variable.

Las condiciones que deben prevalecer para mantener regulado al Sistema, deben estar basadas, principalmente, en modelos estadísticos diseñados para tal fin.

11. Archivos. Dentro de todo este método logicial del Sistema, los archivos de información básica ocupan un lugar principal. El uso de los términos "Archivo de Datos", "Banco de Datos" y "Sistema de Base de Datos", a veces crea confusión, y otras, sus conceptos son interpretados de distintas maneras. Enrique Dieulefait (4), los define de la siguiente forma:

i) Archivo de Datos: Es un conjunto de observaciones manejables mediante un sistema de computación definido en base a una definición única, y cuyos elementos corresponden a distintos períodos o a distintos casos en un mismo

(4) Dieulefait, Enrique. Est. Mat. "Banco de Datos. Un Sistema de Información" Colegio de Graduados en Estadística. Buenos Aires. Documento presentado en el VI Coloquio Argentino de Estadística.

periodo

ii) Base de Datos: Es un conjunto de Archivo de Datos convenientemente estandarizados y perfectamente identificados.

iii) Banco de Datos: Es un sistema capaz de manipular en forma simple y económica, la información estadística contenida en la Base de Datos, permitiendo el análisis de la misma de acuerdo con modelos de tratamiento preterminados, sin necesidad de que cada consulta implique la confección de un programa especialmente orientado a las particulares necesidades del problema.

La organización de estos archivos, tanto para la información primaria, como para la información elaborada y para las funciones que gobernarán al mismo, deberán basarse en criterios que satisfagan los requerimientos estadísticos, las modalidades de la computación y las necesidades de los usuarios. Los estudios interrelacionados que podrán hacerse utilizando la información de los diferentes archivos, es uno de los mayores beneficios que el Sistema otorga.

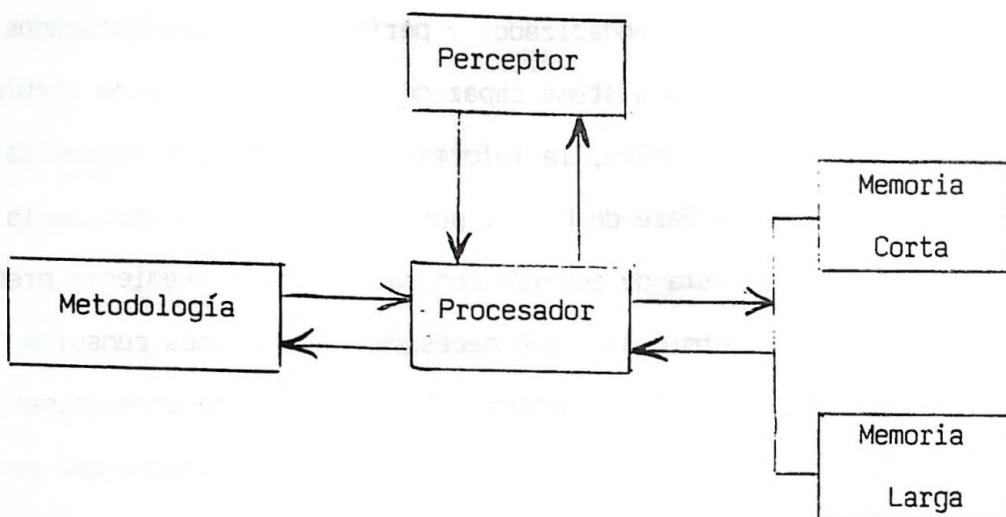
12. En forma general, el proceso de un Sistema de Información consta de los siguientes componentes principales.

i) El aparato perceptor. Toma la información y la traduce a su lenguaje simbólico, o recibe la información simbólica y la traduce al lenguaje humano.

ii) El procesado central. Procesa la información simbólica.

iii) La Memoria. Almacena, reagrupa y regula los datos simbólicos. La memoria corta es un componente del procesador central. La memoria larga está compuesta por archivos externos al procesador central.

iv) El aparato metodológico. Es un repertorio de programas simbólicos para operar en los registros.



13. El futuro de la transferencia de información, a través de los Sistemas de Información, contribuirá en gran medida, a resolver en una forma sustancial, el receso operativo que actualmente se vislumbra en muchas Entidades Estadísticas de América Latina. Las máquinas cada vez más baratas y el avance metodológico cada vez más sofisticado, deben ser aprovechados y regulados bajo normas organizativas, que conlleven a un aumento de la producción estadística y que ésta pueda desarrollarse a través de una reducción de los recursos por unidad producida y dentro de los costos presupuestados.

George J. Stigler, premio Nobel de Economía 1982, viene desarrollando, como trabajo prioritario, la denominada "economía de la información". La misma concierne a la cantidad adecuada de información que deben disponer y que han de requerir las personas, tanto sean hombres de negocio o consumidores en general, para estar perfectamente informados del mercado y poder desarrollar sus actividades. La información es tratada como informados del mercado y poder desarrollar sus actividades. La información es tratada como cualquier otra utilidad y el efecto que la misma tendrá en la conducta de la sociedad estará en función del precio y de

la confiabilidad con que ésta sea suministrada.

14. El Marco Muestral. En el diseño de una investigación en base a muestras de hogares, sea esta enumeración completa o no, el Marco Muestral constituye un eslabón de principal importancia. En rigor, representa al conjunto de unidades muestrales de selección que cubre todo el ámbito de la investigación. Así, por ejemplo, en una investigación que se desee realizar sobre los hogares particulares en un determinado espacio geográfico, el Marco deberá contener a todos los hogares de dicho espacio, con la correspondiente ubicación geográfica de cada uno de esos hogares, de tal manera, que los hogares seleccionados para su medición, sean fácilmente ubicados. Si el Marco no contempla a la lista real de hogares, los hogares faltantes en dicha lista no tendrán posibilidad de ser seleccionados, y además, el factor de expansión para ser aplicado a los resultados muestrales estará subevaluado. Asimismo, hogares que figuren en la lista cuando en realidad no debieran figurar, causará un efecto inverso.

Dado que en la práctica no es posible, o es difícil, contar con una lista actualizada de hogares, el Marco podrá estar configurado por un conjunto de agregados de hogares que abarquen todo el área bajo estudio. Generalmente, estas unidades que agrupan los hogares son áreas geográficas, llámense manzanas, o sectores, o segmentos, o áreas de empadronamiento, etc. En estos casos, lo que se acostumbra hacer, es, primero, seleccionar un área geográfica, y luego, dentro del área seleccionada, seleccionar el número de hogares necesarios.

En determinadas ocasiones los Marcos Muestrales deben obedecer a características especiales. Por ejemplo, en una investigación sobre presupuesto de familias obreras, podrá ser más efectivo contar con un

Marco de empresas, establecimientos, industrias, etc. los cuales, una vez seleccionados, podrán proveer las listas del personal obrero. Si bien, en estos casos, algunos hogares pueden tener una mayor probabilidad de integrar la muestra, la metodología y el trabajo operativo a realizar, para otorgarle a cada hogar igual probabilidad, es de un costo ínfimo comparado con el costo que representaría construir un Marco adecuado para esta determinada investigación.

15. Cuando se posee más de un Marco, por ejemplo, uno demográfico, otro agropecuario, otro industrial, etc. se debe tratar de mantener una asociación funcional entre aquellas unidades similares de los distintos Marcos, con lo cual se acrecentará el valor aditivo de dichas unidades. Estos Marcos, ligados a su vez con la información que se vaya obteniendo de las distintas investigaciones, podrán irse integrando a una Base de Datos, la cual, unida a los complejos sistemas de computación, con que actualmente cuentan la mayoría de las Oficinas Nacionales de Estadística de América Latina, permitirá un mayor número de cruzamiento de variables en un tiempo reducido. Todo esto, ha de significar una sensible reducción en los costos por unidad de información que se disponga para el usuario.

16. En general, siempre ha habido Marcos Muestrales para la selección de muestras de hogares. Los censos de población y habitación han sido los ricos contribuyentes para este fin. En un comienzo, podríamos decir durante la década del 40, era común que el Marco estuviese constituido por el conjunto de carpetas censales. Estas carpetas agrupaban el trabajo realizado por un empadronador censal, y constituían las unidades muestrales de selección de primera etapa. Dentro de las carpetas seleccionadas, se seleccionaban, luego, los hogares.

trabajan en base a tarjetas perforadas, fue la tarjeta sumaria de hogares lo que comenzó a utilizarse como una unidad de selección primaria. Posteriormente, a partir de la década del 60, irrumpen las computadoras en base a programas almacenados, que crean toda una nueva disposición de trabajo, sobre todo, dada la alta velocidad de procesamiento y su gran capacidad de archivo. Se promueve, así, la creación de Marcos Muestrales Computarizados.

17. Marco Muestral Computarizado. El Marco Muestral Computarizado, para la selección de muestras de hogares, se genera, principalmente, con el censo de población, o más aún, con el pre-censo de población, pues será este último el que servirá de marco para levantar al censo.

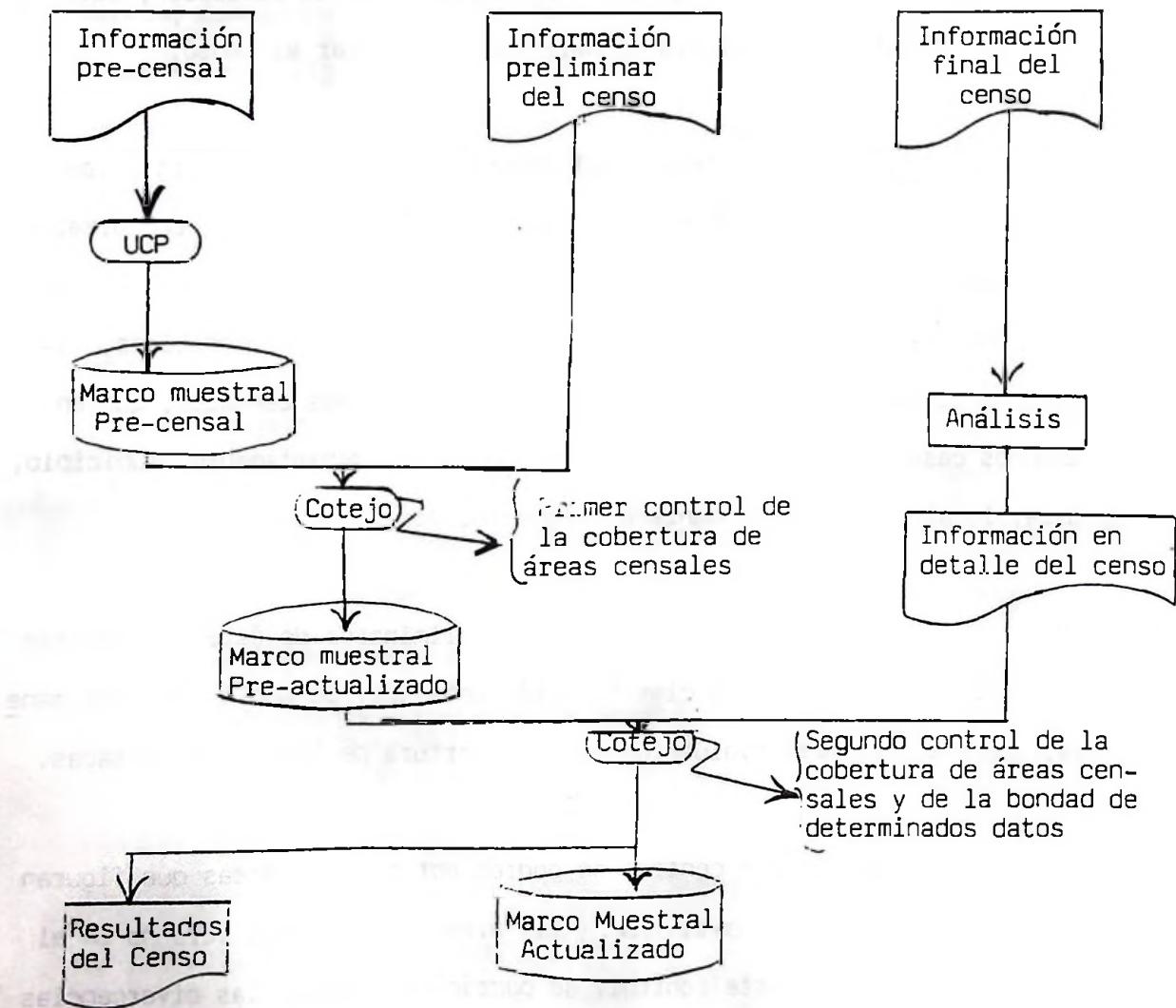
La información pre-censal deberá contener, entre otros, las áreas censales que deberán ser censadas, y que, adicionadas, conformarán la superficie de la región o del país adonde se llevará a cabo el censo. Esta información, será conveniente archivarla en cintas magnéticas, siguiendo el ordenamiento clásico de las organizaciones censales, que en algunos casos podría ser, por región, provincia, departamento, municipio, área, localidad, sector, manzana, segmento, etc.

18. Levantado el censo, los datos preliminares de éste, ingresarán a una cinta, bajo la misma clasificación anterior. Se podrá, de esta manera, hacer una primera evaluación de la cobertura de las áreas censadas.

De este primer control se podrán obtener, i) áreas que figuran en el pre-censo pero no en el censo, y ii) áreas en el censo pero no en el pre-censo. En base a este control, se podrán conciliar las divergencias y se obtendrá un primer Marco Muestral actualizado.

19. Es de hacer notar la importancia que reviste este primer cotejo Pre-Censo vs Censo. El mismo, no asegura la bondad de la información, pero sí, otorga un grado de confiabilidad sobre el barrido de la cobertura censal, el cual, de no realizarse, podrá traer, a la larga, serios tropiezos, que aumentarán sustantivamente el costo para poder subsanarlos.

20. La segunda evaluación se puede realizar cuando se haya ingresado la información en detalle del censo. En este caso se puede cotejar nuevamente la cobertura censal, y además, cotejar ciertos datos de la información preliminar con la información final. Luego de este segundo control, se obtendrá el Marco Muestral actualizado.



21. Definición del Marco Muestral Computarizado. De acuerdo con lo dicho, podemos definir al Marco Muestral Computarizado como a un archivo de información proveniente de una determinada investigación, en este caso, el Censo de Población y Vivienda para ser utilizado, en forma individual o asociado con otros Archivos, en la selección de muestras y en obtención de resultados de aquellas variables, simples o cruzadas, que tiene almacenadas. Esta información debe estar archivada en forma organizada, obedeciendo a una taxonomía regida por un conjunto de reglas y principios que se deben determinar. El Archivo o Base de Datos en el caso de contarse con más de un Archivo, así construido, configurará, asimismo, un punto de arranque para la creación de un Sistema de Información.

22. Unidades Muestrales de Selección que debe Contener el Marco.

Se deberá estudiar la posibilidad de ingresar al Marco todas las Unidades Muestrales de Selección que el Censo pueda proveer. Estas Unidades depen derán de la división política del país y de las creadas para el operativo censal. En principio, se pueden mencionar como las más comunes, las siguien tes: País, región, estado, provincia, departamento, distrito, comuna, cantón, municipio, área (urbana o rural), entidad, territorio federal, gran conglomerado, ciudad, localidad, pueblo, centro poblado, población diseminada, sector, sección, manzana, polígono, segmento, etc.

23. En algunos casos, puede existir una inversión de jerarquía con algunas de estas Unidades. Por ejemplo, puede ocurrir que una manzana contenga a varios segmentos, y otras veces, que un segmento abarque a más de una manzana. En estos casos, si el segmento viene dado como una unidad de jerarquía mayor, con respecto a la unidad manzana, el Marco reflejará este hecho repitiendo el código de la manzana para los distintos códigos del segmento, tal como se muestra en el siguiente ejemplo ficticio.

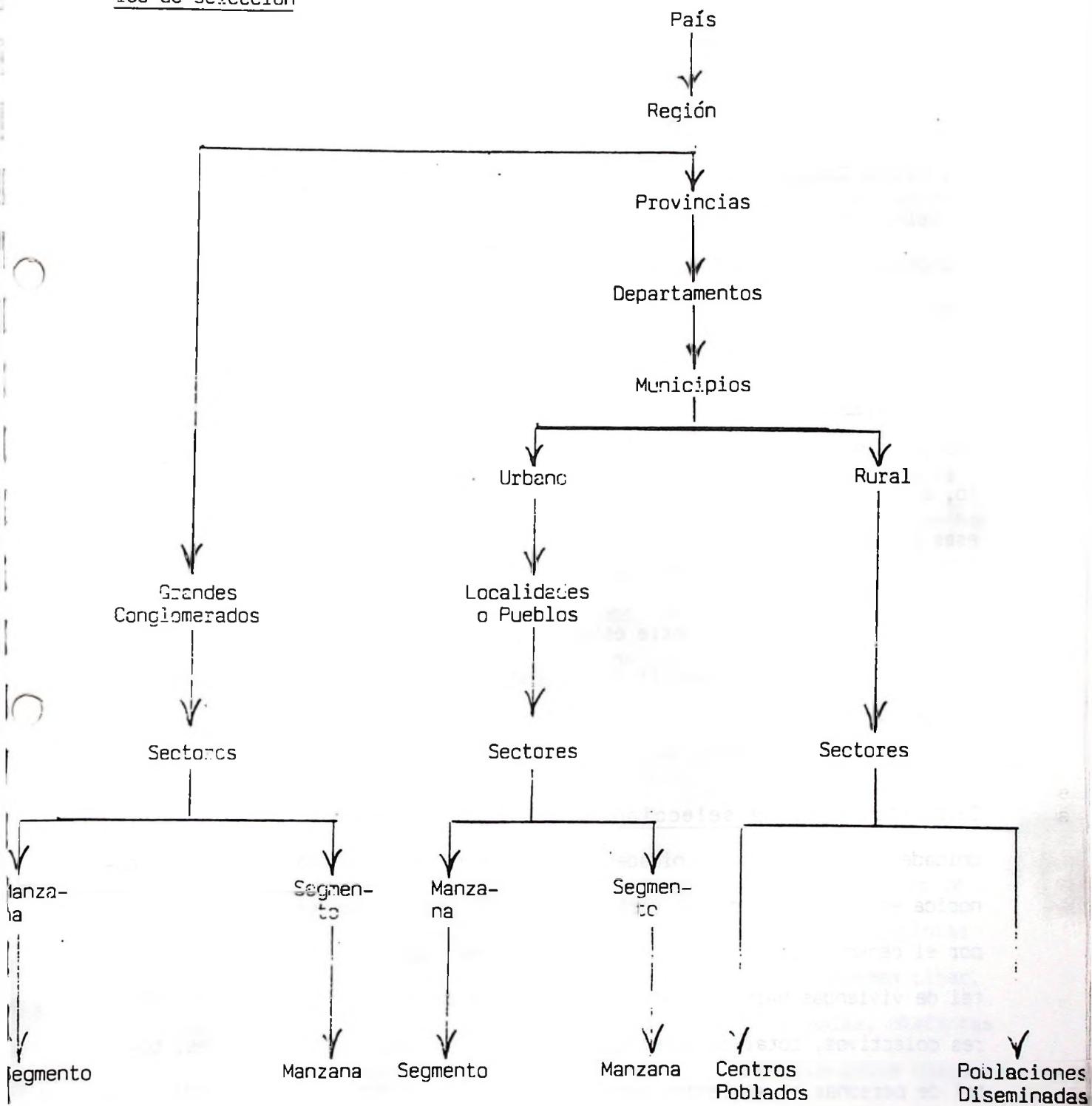
Segmento	Viviendas		Manzana	Viviendas	
	Nº	Parcial	Acumulado	Nº	Parcial
1	41	41	1	20	20
			2	14	34
			3	7	41
2	31	72	4	20	61
			5	11	72
3	15	87	6	15	87
4	6	93	6	6	93

En este ejemplo, la manzana 6 del segmento 3 es la misma manzana 6 del segmento 4. Si se selecciona una muestra de manzanas y resulta seleccionada, ya sea la manzana 6 del segmento 3 o la manzana 6 del segmento 4, se habrá seleccionado a la manzana 6, que contiene a los segmentos 3 y 4 con un total de 21 viviendas.

24. Por otro lado, un caso similar aunque de magnitud distinta, puede ocurrir cuando alguna de estas Unidades pertenece a más de una Unidad de otro orden. Es el caso de los "grandes conglomerados" que pertenecen, una parte a una entidad política y otra parte a otra entidad política. Por ejemplo, la ciudad de Monterrey, en México, forma parte de 3 Estados, la ciudad de Caracas, en Venezuela, se reparte en 2 Estados, el Área Metropolitana de Asunción comprende a la ciudad de Asunción y 10 Distritos colindantes, etc. En estos casos, puede ser conveniente, sobre todo desde el punto de vista práctico, que la información se registre más de una vez en el Archivo, una para el "gran conglomerado" y otra para la entidad política respectiva. La facilidad operativa que otorgará esta forma de proceder, será beneficiosa en relación con el aumento físico del Marco, en especial, teniendo en cuenta la alta densidad con que trabajan los actuales archivos de computación.

25. En el siguiente flujo de ramales se muestra, en forma sencilla, la distribución jerárquica de ciertas Unidades Muestrales de Selección utilizados en algunos censos de población.

Distribución jerárquica  
de las unidades muestra-  
les de selección



26. Diseño de la Unidad Muestral. Los distintos campos que deberá contener la Unidad Muestral, las posiciones de cada campo y los códigos que se utilizarán, deberán ser objeto de un estudio minucioso, en función de los datos que se puedan obtener del censo, de las futuras investigaciones proyectadas y de la integración del Archivo a una Base de Datos.

Entre los campos principales de la Unidad que deberán considerarse, están: Código de la Unidad, código de la ubicación geográfica, variables de selección, características de estratificación, indicadores estadísticos y económicos, factores de costo y accesibilidad, fecha de actualización, códigos cartográficos y datos de control.

26.1. Códigos de la Unidad y de la Ubicación Geográfica. Las posiciones del campo identificarán a la Unidad Muestral, indicando si es un segmento, o una manzana, o un sector, o un pueblo, etc. Los valores grabados en esas posiciones indicarán el número de la Unidad y su ubicación geográfica.

La relación entre estos códigos y la Unidad deberá ser biunívoca. Además, estos Códigos, permitirán la generación de archivos de sumarias, como se verá en párrafo 37.

26.2. Variables de selección. Las variables que deberán contener las Unidades, para que estas Unidades sean seleccionadas con una probabilidad conocida en función de dichas variables, serán, principalmente, las provistas por el censo. Entre las más importantes podemos citar: total de viviendas, total de viviendas particulares, total de viviendas colectivas, total de hogares colectivos, total de personas, total de hombres, total de mujeres, total de personas en viviendas particulares, total de hombres en viviendas particulares, total de mujeres en viviendas particulares, total de personas en

viviendas colectivas, total de hombres en viviendas colectivas, total de mujeres en viviendas colectivas, total de viviendas ocupadas, total de viviendas desocupadas, total de viviendas en construcción, total de viviendas propias, total de viviendas alquiladas, total de viviendas en otras condiciones de tenencia, total de cuartos de los hogares particulares, total de cuartos de los hogares colectivos, total de personas de: 5-15 años; 15-25 años; 25-55 años; 55 y más años, total de varones de: 5-15 años; 15-25 años; 25-55 años; 55 y más años, total de mujeres de: 5-15 años; 15-25 años; 25-55 años; 55 y más años, total de personas nacidas en el país, total de hombres nacidos en el país, total de mujeres nacidas en el país, etc.

26.3. Características de estratificación. Además de los códigos de ubicación geográfica, que permitirán estratificar a las Unidades geográficamente, el censo provee una cantidad de características que relacionadas entre sí, suministrarán una serie de coeficientes que serán de suma utilidad para la formación de estratos. Por ejemplo, por Unidad, se pueden obtener tasas de nacimiento, de fecundidad, de población económicamente activa, de población económica inactiva, de obreros, de empleados, de trabajadores por cuenta propia, de servicio doméstico, de patronos, de sueldos y salarios, de ocupaciones, de analfabetismo, de incapacitados o "minusválidos", de estudios realizados, de masculinidad, de tipos de vivienda, de condiciones de ocupación, de áreas de producción, etc.

26.4. Indicadores estadísticos y económicos. Estos indicadores se podrán ir construyendo a medida que se vayan llevando a cabo las distintas investigaciones basadas en muestras de hogares. Entre ellos se pueden citar, por ejemplo, la correlación intraclasa para determinadas variables, distintas tasas de colaboración de la población, coeficientes de relación entre distintos atributos, etc.

26.5. Factores de costo y accesibilidad. Si bien, hasta el presente, la mayoría de los censos no han previsto la obtención de estos factores, se hace necesario contar con los mismos dentro de un Marco, entre cuyos fines principales, está el de contribuir a la reducción de los costos por niveles de confiabilidad. Entre estos factores se pueden mencionar, sobre todo para el área rural, distancia en kilómetros desde el centro poblado más cercano para llegar a la unidad, tiempo en horas para recorrer esa distancia, estado de los caminos, medios de transporte, épocas de transitabilidad, costo del transporte, tipos de hospedaje, costo del hospedaje, cota de la unidad, etc.

Muchos de estos valores se podrán conjeturar, con los técnicos conocedores del terreno, en la Oficina, y otros, se irán obteniendo a través de las distintas investigaciones que se vayan realizando.

26.6. Fecha de actualización. Cada vez que una Unidad sea actualizada, sea parcial o total, se registrarán las fechas de actualización correspondientes. La antigua información se transferirá a un Archivo Histórico.

26.7. Códigos cartográficos. En este campo se registrarán los códigos de la respectiva cartografía de la Unidad.

26.8. Datos de control. Corresponden a artificios que permitan conocer la validación de la Unidad, por ejemplo, el registro de un total de todos los valores numéricos de la Unidad, etc.

27. Las Unidades Muestrales así diseñadas y con la información correspondiente, conformarán un Marco Muestral que permitirá para los futuros trabajos de investigación, poder establecer diseños que contemplen la formación de estratos socio-económicos que redunden en una disminución sustancial de los costos, por resultado obtenido, manteniendo una precisión dada.

28. Sub-Marcos Muestrales. Por razones prácticas de utilización del Marco Muestral, convendrá, no tener un Marco Muestral único, sino, una serie de subconjuntos del Marco Muestral, Sub-Marcos, que unidos, configuren el Marco Muestral del país. Para ello, será beneficioso estudiar sobre la división más práctica y conveniente que optimice su utilización. Por ejemplo, se podrán obtener Sub-Marcos por provincia, por gran conglomerado, por área, etc.

29. Simulación del Marco Muestral. A medida que se vayan realizando las distintas etapas que componen al Marco Muestral, éstas se deberán ir simulando con datos ficticios, primero en el papel, de acuerdo con la lógica establecida para grabar y leer información, y luego en la computadora. Se deberán contemplar, en la medida de las posibilidades, todos los casos posibles de presentación, forma de detectar las anomalías y manera de corregirlas.

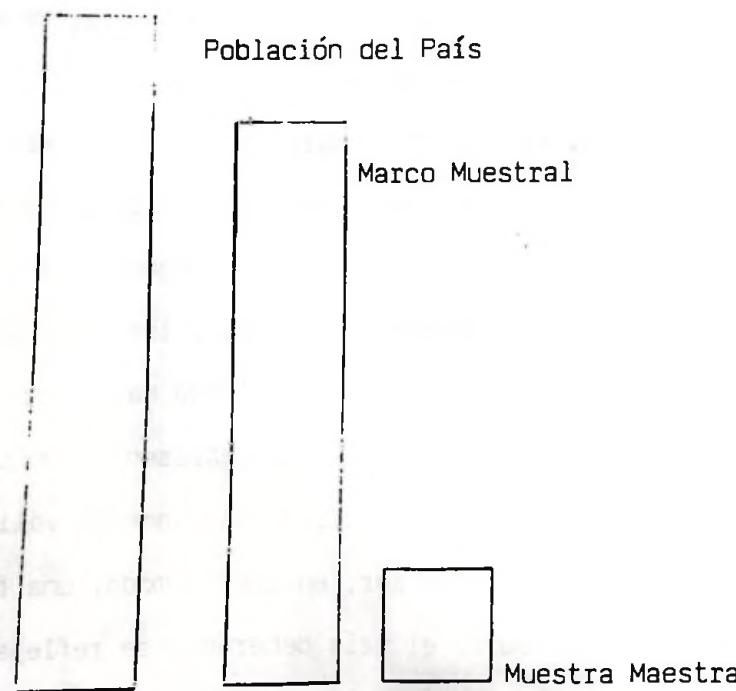
30. Actualización del Marco Muestral. El Marco Muestral representa a la población del país. En este caso, la palabra población se refiere a la población de personas, a la población de viviendas, a la población de manzanas, a la población de sectores, etc. Las investigaciones estadísticas que se realizan durante el período intercensal, 10 años aproximadamente, deberán utilizar, para consultar a estas poblaciones del país, muestras de unidades seleccionadas de este Marco. Por tanto, si el Marco Muestral no representa a estas poblaciones, ya sea por ser anacrónico, o tener unidades equivocadas, o faltarle unidades, etc., las inferencias que se obtengan de estas investigaciones adolecerán de esos mismos defectos. Es decir, los resultados que se obtengan de muestras seleccionadas de un determinado Marco serán resultados para ese Marco. De allí, que si el Marco no representa confiablemente al país, los resultados obtenidos no serán significativamente válidos para el país. Por ello, el Marco Muestral, debe ser, en cierto modo, una fotografía del país, y las variaciones que sufra el país deberán irse reflejando en el

Marco. Las muestras que se seleccionen de un Marco así construido y así conservado, serán muestras que representarán al país con un alto grado de confiabilidad.

Mantener un Marco actualizado en su totalidad, es equivalente a mantener un censo actualizado, lo cual representa un costo excesivo. Lo que se sugiere hacer, por razones prácticas y de costo, es seleccionar una Muestra Maestra del Marco y mantener actualizada esta Muestra Maestra. Las muestras que se seleccionen para las distintas investigaciones serán, por ende, seleccionadas de la Muestra Maestra.

El camino a seguir, que se recomienda, es el siguiente:

- i) Construcción del Marco Muestral en base al Censo de Población y Vivienda,
- ii) Evaluar al Censo, lo cual es equivalente a evaluar al Marco Muestral,
- iii) Seleccionar del Marco Muestral una Muestra Maestra,
- iv) Mantener actualizada la Muestra Maestra



La diferencia entre la población del país y el Marco Muestral se conocerá a través de la evaluación del censo. Por otro lado, en determinados períodos, se deberá hacer una evaluación de la Muestra Maestra, para conocer si esta Muestra Maestra sigue representando al Marco y si este Marco sigue representando a la población de estudio con confiabilidad. Para ello, se deberá contar con determinados indicadores, que permitan ir conociendo el grado de variación, por lugar geográfico, que registren las investigaciones llevadas a cabo, como por ejemplo, tasa de crecimiento poblacional, permisos de construcciones, zonas de desarrollo, etc.

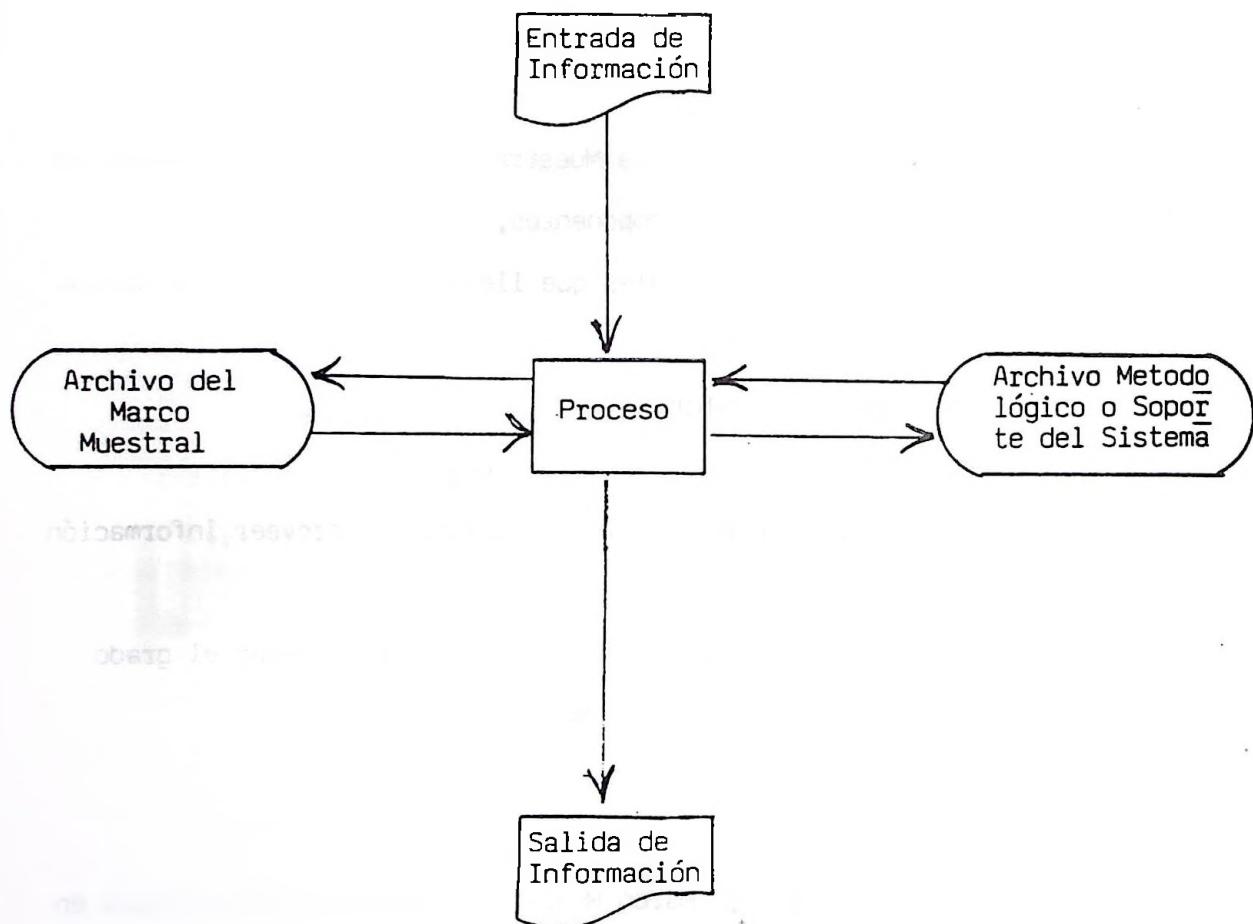
Para mantener actualizada la Muestra Maestra se puede adoptar un modelo que utilice los siguientes componentes,

- i) Investigaciones muestrales que lleve a cabo la Oficina Nacional de Estadística,
- ii) Investigaciones ad-hoc,
- iii) Denuncias,
- iv) Coordinación con otras oficinas que puedan proveer información útil para dicho fin,
- v) Evaluación, en determinados períodos, para conocer el grado de representatividad de la Muestra Maestra,
- vi) Actualización cartográfica.

33. Archivo Metodológico. El Marco Muestral computarizado configura en sí, un pequeño sistema de información. Es, en rigor, un archivo de datos, que se trata de mantener actualizado, y que debe poseer la flexibilidad de ser consultado en forma fácil, rápida, confiable y económica. Por ello, se debe ir tratando, de que los modelos estadísticos y matemáticos que se vayan desarrollando, se deben ir programando como funciones paramétricas. Estas funciones así construidas, con sus correspondientes denominaciones, serán

registradas en un Archivo Metodológico, del cual podrán ser llamadas para su utilización. Este Archivo Metodológico, que también recibe el nombre de Soporte del Sistema, comprenderá, asimismo, a los programas de entrada y salida del sistema, cualesquiera sea su función.

Gráficamente, la comunicación del sistema del Marco Muestral Computarizado es,



34. Diccionario del Marco Muestral. Cada programa que se realice, así como los componentes del mismo, sean estos subrutinas, diseños de registros etc., y cada función, modelo, etc. que se utilice en la construcción del

Marco, tendrá un nombre particular, biunívoco, que se registrará en un Diccionario con su descripción, definición, diseño, tamaños de registro, etc.

### 35. Personal a Colaborar en la Construcción del Marco Muestral.

Para participar en la construcción metodológica y práctica del Marco Muestral, así como para pulsar las variaciones del país y sugerir las acciones a tomar para mantener actualizado el Marco y, por ende, la cartografía, será conveniente la formación de un equipo de personas, preferentemente con dedicación exclusiva. Estas personas deberán estar en capacidad de diseñar los distintos tipos de investigaciones necesarias para satisfacer los componentes mencionados en párrafo 32. Si bien, la formación de este equipo ha de depender tanto de la estructura organizativa de la Oficina como de los recursos materiales con que la misma cuenta, se deberá contemplar la posibilidad de integrar el equipo con personal de la oficina de muestreo, de la oficina de computación y de la oficina de cartografía. El personal de la oficina de muestreo deberá poseer amplios conocimientos en estadística, de manera tal, de sugerir las funciones paramétricas que integrarán el archivo metodológico. El personal de computación, con preferencia a nivel de analista de sistemas, serán los encargados de traducir al lenguaje de la máquina la lógica para la construcción del Marco Muestral y de las fórmulas estadísticas que comenzarán a regir al mismo. La persona de cartografía deberá conocer la organización cartográfica de la Oficina y el estado actual de la cartografía, dado que cada Unidad Muestral, teóricamente, deberá tener su correspondiente croquis actualizado.

Una vez establecido el equipo, convendrá, previo a la iniciación de las tareas, que el mismo sea asesorado en la filosofía del Marco Muestral y en los lineamientos generales a los cuales se deberá ceñir, con vistas a una futura integración del mismo a una Base de Datos, para luego trabajarlo

en profundidad.

36. Generación del Marco Muestral Computarizado. En función de la infraestructura existente y de la configuración que se adopte en dimensionar al Marco, se establecerán las reglas y principios que lo gobernarán. Teniendo en cuenta la situación promedio de las Oficinas Nacionales de Estadística en América Latina, se pueden sugerir las siguientes etapas de trabajo para la construcción del Marco Muestral.

1a. etapa

- Definición y objetivos del Marco Muestral
- Definición de las Unidades Muestrales que integrarán el Marco Muestral.
- Requerimientos.
  - Personal (formación del equipo de trabajo)
  - Utiles de trabajo
  - Lugar físico de trabajo
  - Financiamiento
  - Sistema de computación
- Coordinación
  - Con los distintos sectores de la Oficina Nacional de Estadística.
  - Con los principales usuarios.

2a. etapa

- Diagramación lógica de la construcción del Marco Muestral
- Definiciones
  - Diseño de las Unidades Muestrales
  - Códigos
  - Variables de selección

- Características de estatificación
- Códigos cartográficos
- Datos de Control
- Comienzo del Diccionario

#### 3a. etapa

- Simulación manual de la lógica establecida con datos ficticios.
  - Carga del Marco Muestral
  - Verificación de controles
  - Correcciones
- Generación de archivos sumarios

#### 4a. etapa

- Programa de computación
- Simulación mecánica con datos ficticios
  - Grabación del Marco Muestral
  - Validación
  - Correcciones
  - Lectura del Marco Muestral

#### 5a. etapa

- Comienzo de los programas de las funciones básica paramétricas.
  - Tamaños muestrales
  - Generación de números aleatorios
  - Selección de muestras
  - Estimadores

6a. etapa

- Comienzo de la Construcción del Marco Muestral basado en el Censo de Población y Vivienda.
- Verificación de los programas de grabación y lectura.

7a. etapa

- Creación de los Sub-Marcos

8a. etapa

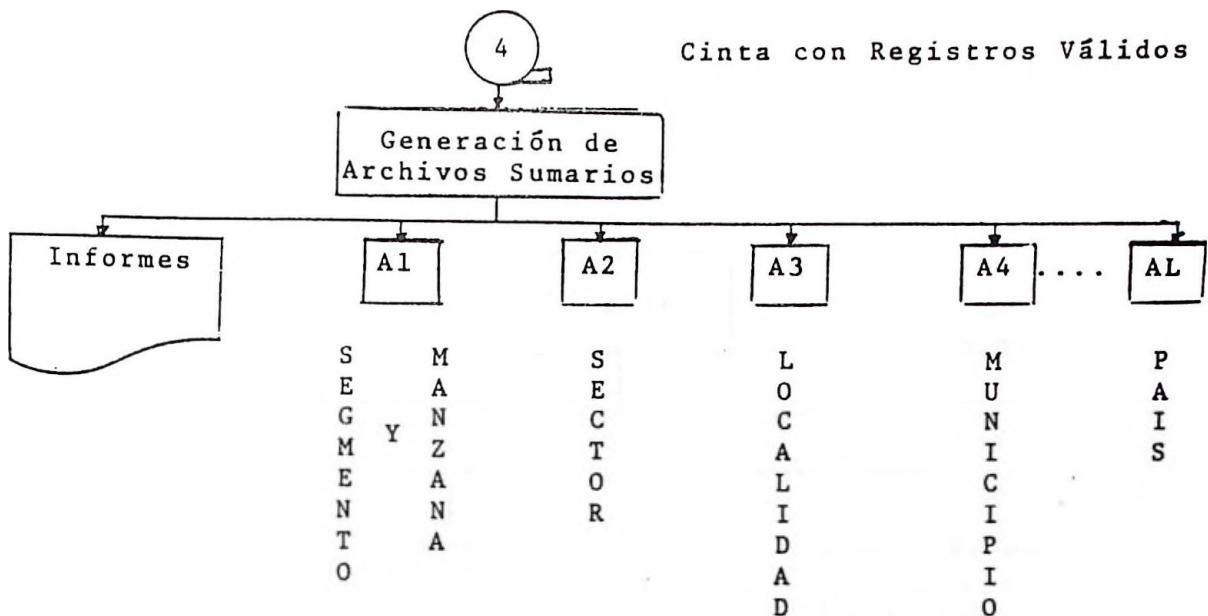
- Métodos de actualización y construcción del Archivo Histórico.

9a. etapa

- Desarrollo de nuevas rutinas
- Construcción de nuevos archivos interrelacionados entre sí y con el Marco Muestral.

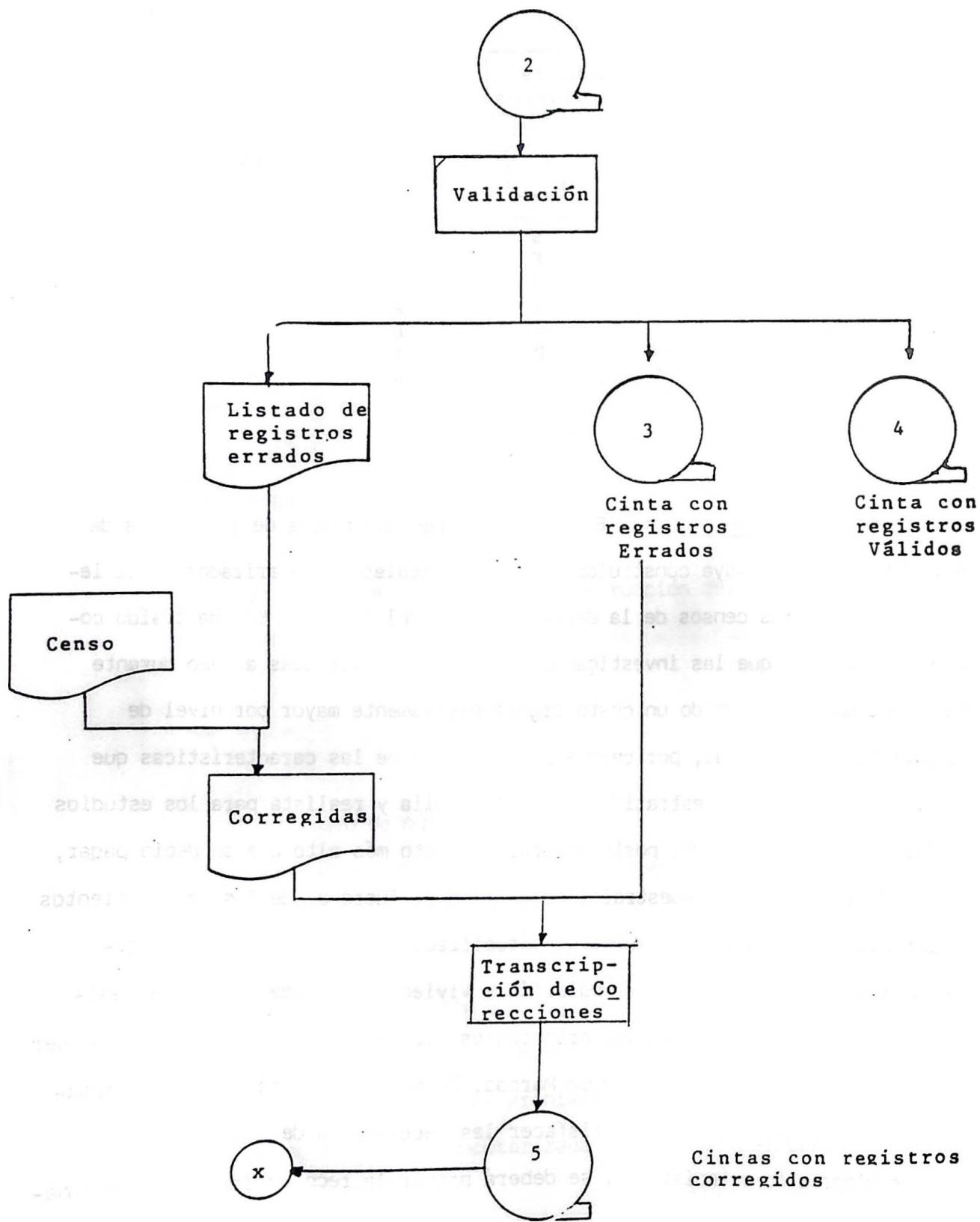
37. Flujograma. El siguiente flujograma, trata de mostrar, a grandes rasgos, la lógica de i) transcripción y ordenamiento; ii) validación y corrección de errores y iii) generación de archivos; que en algunos casos podrá emplearse para la construcción computarizada del Marco Muestral.

## iii) Generación de Archivos Sumarios

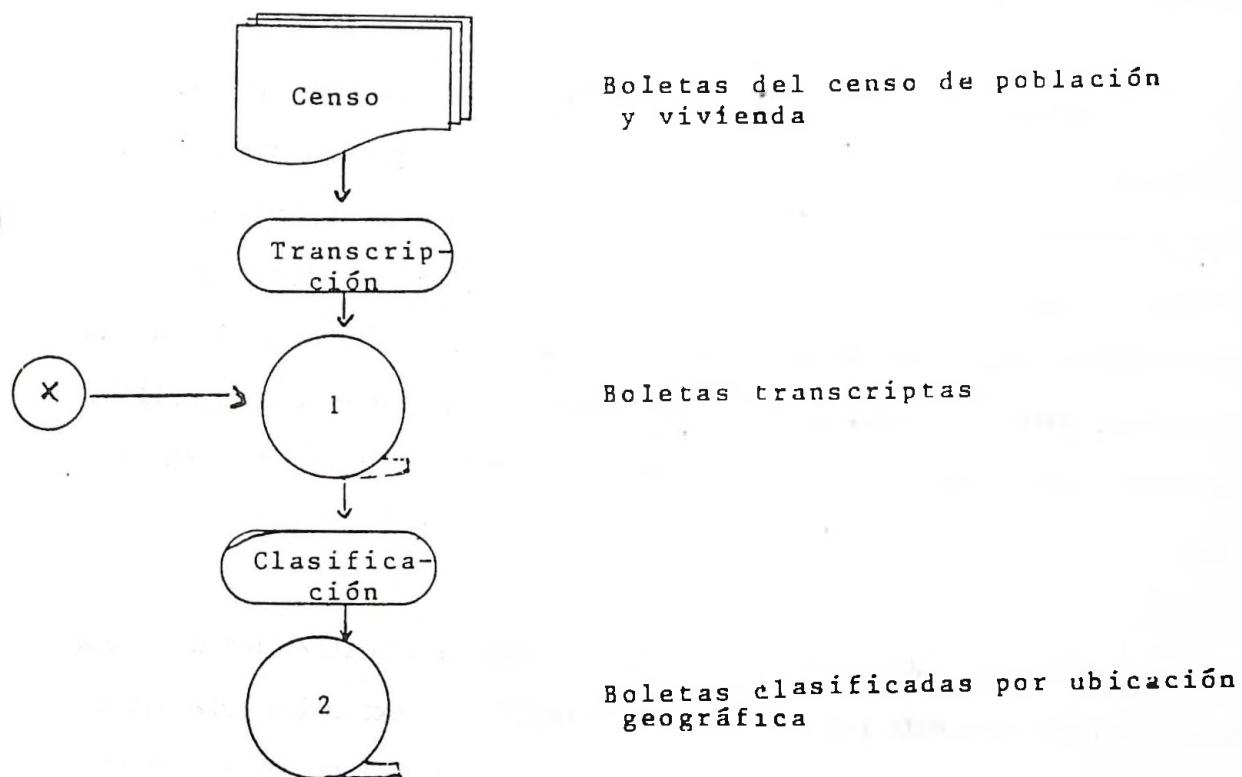


38. Beneficios Estimados. El hecho, de que la mayoría de los países de América Latina, no haya construido Marcos Muestrales Computarizados en el levantamiento de los censos de la década del 60, del 70 y del 80, ha traído como consecuencia, que las investigaciones muestrales llevadas a cabo durante esas décadas hayan tenido un costo significativamente mayor por nivel de eficiencia. En especial, por carecerse de muchas de las características que hubiesen permitido una estratificación más amplia y realista para los estudios realizados. Por otro lado, posiblemente, el costo más alto que se debió pagar, por carecerse de Marcos Muestrales actualizados, fuera el de los levantamientos de los Pre-Censos. Un Marco Muestral actualizado debe ser la base del pre-Censo para el próximo Censo de Población y Vivienda. No debe olvidarse, asimismo, que uno de los objetivos principales que tiene todo Censo, es el de servir de base para la construcción de Marcos. De ahí, que además de la información que recoge el Censo para satisfacer las necesidades de los usuarios como consumidores de estadísticas, se deberá prever la recolección de determina-

## ii) Validación y Corrección de Errores



## i) Transcripción y Ordenamiento



dos valores, que no aumentan el costo ni el tiempo del operativo censal, pero que sí influirán positivamente en la disminución de los costos de las futuras investigaciones estadísticas. Por ejemplo, registrar los valores de accesibilidad al segmento censal o área del empadronador. Este valor, el empadronador lo puede registrar en su hoja resumen, si así lo desea, en el día anterior al operativo censal, cuando llena otras formalidades de su labor.

39. Supongamos, para fijar ideas, que en un país de unos 10 millones de personas, el costo de un Censo, que aproximadamente se puede estimar en una hora de remuneración per cápita (5), haya sido de unos 10 millones de horas. Podemos conjeturar, además, en base a las experiencias realizadas, que el 30 por ciento de este costo es insumido por el Pre-Censo. Es decir, el costo del Pre-Censo, para este caso en particular, se puede estimar en unos 3 millones de horas, o sea, una 25.000 horas por mes en un período intercensal de 10 años.

Por otro lado, suponiendo que el equipo de actualización del Marco Muestral esté compuesto por 3 técnicos con dedicación exclusiva para dicho trabajo, y que cada técnico, en promedio, cueste mensualmente el equivalente a 2000 horas, la erogación mensual sería de unas 6000 horas. La diferencia de 19.000 horas representaría así, un ahorro de más del 75 por ciento del costo del Pre-Censo, si se cuenta con un Marco actualizado que reemplace al Pre-Censo. Pero aún, en el caso de que el Marco Muestral sólo actualizase un 50 por ciento del país, el ahorro sería significativo.

40. Asimismo, se debe imputar como ahorro, el hecho de que se evitará, en muchos casos, actualizar en el terreno a las Unidades Primarias de Selección, en las Encuestas de Hogares, lo cual, en 10 años representa un presupuesto sustancial.

41. Otros beneficios importantes que arrojará un Marco Muestral así constituido, vendrán dados por i) el mejoramiento de la representatividad de las muestras; ii) una mayor posibilidad de estratificación; iii) mayores elementos para poder analizar y mejorar los diseños muestrales y iv) obtención de distintos resultados de la información almacenada.