

Departamento Administrativo Nacional de Estadística



**Dirección de Metodología y Producción
Estadística - DIMPE**

**Metodología del Diseño Muestral
Encuesta de Microestablecimientos
MICRO**

Octubre 2007

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 1 FECHA: 23-10-07
ELABORÓ: Diseño Muestral	REVISÓ: Coordinador Diseño Muestral	APROBÓ : Director DIMPE

1. INTRODUCCIÓN

Para la proposición de un diseño muestral a implementar en cualquier investigación, es conveniente tener en cuenta tres grupos de aspectos importantes, los contextuales, operativos y teóricos. Estos primeros hacen referencia a los objetivos del estudio y al problema en la investigación, los segundos implican una evaluación de recursos disponibles para la etapa de realización de la encuesta y los últimos, compromete el conjunto de condiciones teóricas y conceptuales, relacionadas directamente con la teoría de muestreo. A continuación se presenta el diseño muestral de la encuesta de Microestablecimientos de comercio, servicios e industria.

2. PLANIFICACION

2.1. OBJETIVOS

Objetivo General.

Determinar la evolución en el corto plazo (trimestre) y la estructura (anual) en el mediano plazo de la actividad de los microestablecimientos en Colombia en los sectores de industria, comercio y servicios.

Objetivos Específicos.

- Medir la actividad económica en los establecimientos de industria, comercio y servicios con no más de 10 personas ocupadas.
- Proveer periódicamente estadísticas e indicadores que permitan monitorear la evolución del empleo, el valor de la producción, las ventas y los ingresos de los microestablecimientos.
- Presentar anualmente estadísticas de estructura de la actividad de los microestablecimientos.
- Complementar la información de los sectores de industria, comercio y servicios que no es captada por la Encuesta Anual Manufacturera, la Muestra Mensual Manufacturera, la Encuesta Anual de Comercio, la Muestra Mensual de comercio al por menor, la Encuesta Trimestral de Grandes Almacenes e Hipermercados y la Encuesta Anual de servicios.
- Complementar y actualizar los marcos de lista para las investigaciones económicas.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 2 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------

- Generar la información básica para la promoción del desarrollo de la microempresa.

2.2. CONCEPTOS BÁSICOS

2.3. GRADO DE PRECISIÓN

Para el cálculo del tamaño de muestra, tanto de comercio como de industria, se determinó un nivel de precisión global deseado (error estándar relativo o coeficiente de variación (CV), entre el 5 y el 10 para las variables de interés: ventas y personal ocupado; este CV permite presentar la información a nivel estrato.

2.4. PERÍODO DE REFERENCIA.

Mes anterior al operativo de campo para evolución
 Año anterior al operativo para estructura

2.5. PERÍODO DE RECOLECCIÓN

Encuesta trimestral para evolución.
 Encuesta anual para estructura.

3. DISEÑO ESTADÍSTICO

En esta etapa se construye el plan muestral que debe contener:

3.1. GENERALIDADES

La implementación de la Encuesta Nacional de Microestablecimientos se realizó a partir del año 2001, con la información referida a la estructura de estos microestablecimientos para el año 2000. Desde el primer trimestre del 2002 se viene realizando en forma continua para obtener información de evolución trimestral y, por supuesto, de estructura del año anterior. Para el cuarto trimestre del 2003 la cobertura se redujo a las 14 áreas metropolitanas del país. Para el cuarto trimestre de 2006 la cobertura del estrato 1 se amplió de 14 a 24 áreas metropolitanas de país, conformadas por: Bogotá, Cali, Medellín, Barranquilla, Cartagena, Montería, Manizales, Pereira, Bucaramanga, Villavicencio, Pasto, Cúcuta, Ibagué, Neiva, Tunja, Florencia, Popayán, Valledupar, Quibdó, Riohacha, Santa Marta, Armenia, Sincelejo y San Andrés. Y a partir del primer trimestre de 2007 la cobertura de la encuesta se retoma a nivel nacional.

A continuación se propondrá una nueva estratificación, fusionando los estratos 2 y 3, la razón para proceder en la unión de éstos estratos radica en la homogeneidad que existe

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 3 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------

entre ellos y en la conformación de un nuevo marco de áreas a partir de la información económica del Censo General 2005.

3.2 UNIVERSO

Todos los establecimientos de industria, comercio y servicios que tienen menos de 10 personas ocupadas, ubicados en las cabeceras municipales de todos los municipios del país.

POBLACIÓN OBJETIVO

Todos los establecimientos de industria, comercio y servicios que tienen menos de 10 personas ocupadas, ubicados en las cabeceras municipales de todos los municipios del país. Se excluyen los puestos móviles.

En el sector de comercio se tendrán en cuenta los establecimientos de comercio al por menor, el comercio al por menor de partes y accesorios de vehículos automotores, los talleres de mantenimiento y reparación de automotores y el comercio, mantenimiento y reparación de motocicletas y de sus partes, piezas y accesorios.

En los servicios los establecimientos que se dedican al alojamiento y expendio de comidas y bebidas, agencias de viaje, de empleo y de publicidad y las empresas de seguridad privada.

3.3 VARIABLES DE ESTUDIO

EVOLUCION	ESTRUCTURA
Valor total de producción, ventas o ingresos	Valor total de producción, ventas o ingresos
Personal ocupado y tipo de contratación	Gastos operacionales
Actividad económica	- Salarios, sueldos y prestaciones sociales
Número de establecimientos por actividad económica	- Costo de las mercancías (comercio), materias primas, insumos y empaques (industria)
Tiempo de funcionamiento	- Servicios públicos
Organización jurídica	- Arrendamientos
Tipo de emplazamiento	- Costo de venta de alimentación y bebidas (servicios)
	- Costos de producción de bienes o servicios
	- Gastos de comunicaciones

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 4 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	------------------------------------------------------------------------

EVOLUCION	ESTRUCTURA
	- Gastos en mantenimiento y reparación
	- Gastos en impuestos y otros
	Valor de los activos
	Área del establecimiento

3.4 COBERTURA

La cobertura de la investigación es nacional urbana (total cabeceras municipales del país).

3.5 UNIDADES ESTADÍSTICAS

Unidad de Muestreo: En la primera etapa de la investigación la unidad de muestreo es el Municipio, para el caso de las grandes ciudades no se habla de municipio seleccionado sino de área metropolitana completa, por ejemplo en el caso de Medellín, al ser seleccionada el área incluirá automáticamente a Bello, Envigado, Itagüí, Copacabana, La Estrella, Caldas, Girardot, Barbosa y Sabaneta.

En la segunda etapa la unidad de muestreo es la manzana cartográfica o la *sección*, definida como conjunto de manzanas que pertenezcan a una misma sección cartográfica DANE y que tengan menos de 10 comercios y menos de 5 industrias y que además no pertenezcan a los subestratos 1, 2 o 4; esta unidad dependerá del subestrato en el cual se esté haciendo la selección.

Finalmente en la tercera etapa la unidad de muestreo es el establecimiento industrial, comercial o de servicios.

Unidad de Observación: Las unidades finales de las cuales se obtendrá información son los establecimientos industriales, comerciales y de servicios.

Unidad de análisis: Microestablecimientos económicos.

3.6 MARCO ESTADÍSTICO

El marco de ésta investigación es de áreas, conformado por la base cartográfica, a nivel manzana, de las cabeceras de los 1116 municipios del país.

Cobertura geográfica: Cabeceras municipales 1116 municipios del país.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 5 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------


Fecha de construcción: El marco se encuentra en proceso de actualización a nivel muestra de UPM y USM. En el estrato 1 se realizó la respectiva homologación cartográfica de 2001 a 2004, quedando pendiente la actualización de Bogotá y área metropolitana de Medellín, que mantienen cartografía 1996 y 2001 respectivamente. En los demás estratos, el proceso se está realizando con la información económica del Censo General 2005 y con las bases de homologaciones que contienen la verificación de código DANE versus código censo y AG, proporcionadas por la DIG.

Fuente: El marco muestral en principio se construyó a partir del marco económico de áreas correspondiente a 184 municipios del país, referenciado en las investigaciones “Construcción de marcos de áreas para las investigaciones económicas”, “Censo económico de Cundinamarca” y “Censo Económico”, realizadas por el DANE en 1996 y 1999, respectivamente. En la actualidad, como ya se mencionó, el marco se encuentra en proceso de actualización con la información del Censo General 2005 de donde se filtran los municipios correspondientes a la cobertura de la encuesta.

Método de actualización. En general, los procesos de mantenimiento requeridos se efectúan a nivel de unidades y de información auxiliar. En el caso particular del Marco de Áreas para las Investigaciones Económicas, por tratarse de manzanas cartográficas que representan áreas geográficas, el mantenimiento a nivel de unidad se refiere a la homologación de códigos de una a otra cartografía para las manzanas existentes, la incorporación de nuevos desarrollos y, la eliminación de las zonas demolidas o que ya no pertenecen al área urbana de los municipios; también contempla la eliminación de aquellas zonas que aparezcan repetidas en el marco, por problemas de codificación. A nivel de Información auxiliar, el proceso de mantenimiento considera la actualización de la información correspondiente a cantidad de establecimientos, la adecuación de dicha información para responder a las necesidades de los distintos usuarios y, eventualmente, la inclusión de nuevas variables que enriquezcan el marco.

El marco de Áreas para las Investigaciones Económicas, tiene como unidad la manzana cartográfica DANE. La identificación de ésta se hace a través de un código de 22 dígitos compuesto por la siguiente información:

Campo	Posición	Descripción
cod_dpto	1 - 2	Código del Departamento en que se encuentra la manzana
cod_mun	3 - 5	Código del Municipio en que se encuentra la manzana
clase	6	Indicador de si la manzana pertenece al área urbana o rural de municipio. Para rural = 0 y para urbano =1
sect_rural	7 - 9	Sector rural en que se encuentra la manzana. Para áreas urbanas = 000
secc_rural	10 – 11	Sección rural en que se encuentra la manzana. Para áreas urbanas = 00

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 6 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------

cent_pob	12 – 14	Centro poblado al que pertenece la manzana. Para áreas urbanas = 000
sector	15 – 18	Sector urbano en que se encuentra la manzana
seccion	19 – 20	Sección urbana en que se encuentra la manzana
manzana	21 - 22	Código de la manzana

Este código es asignado especialmente para el DANE por la DIG y usualmente, no coincide con los códigos asignados por otras entidades como Catastro Distrital; esto hace que el DANE tenga su propia cartografía y libertad para modificarla y actualizarla conforme a sus posibilidades. En la medida en que se producen cambios geográficos y físicos en las ciudades, la DIG se encarga de actualizar su cartografía, con el fin de garantizar la idoneidad del material que permite a los grupos operativos de la Entidad movilizarse en campo. Una conjunción importante de cambios hace que se genere una nueva cartografía para determinado municipio.

En este caso, para efectuar el proceso de migración de una a otra cartografía, la DIG se encarga de elaborar una tabla de homologación, esto es, un archivo magnético que contiene el código asignado a cada una de las manzanas del municipio en la antigua cartografía y el o los códigos nuevos asignados. En el caso de los nuevos desarrollos geográficos, esta tabla los incorpora mediante la adición de registros y la habilitación de nuevos códigos.

Para llevar a cabo el mantenimiento en este aspecto, es necesario identificar la base cartográfica aprobada por el DIG más reciente en cada uno de los municipios pertenecientes al Marco de Áreas. A la fecha, Bogotá tiene cartografía 96 como la más reciente, el área metropolitana de Medellín mantiene cartografía 2001 y el resto de municipios, cuentan con cartografía 2004. Lo ideal es que la base cartográfica del marco sea siempre la más reciente posible.

Una vez hecha esta identificación, debe solicitarse a la DIG las tablas de homologación correspondientes. Es importante efectuar una revisión de integridad a las tablas de homologación previniendo, la presencia de celdas vacías y verificando la integridad del código que identifica cada manzana.

En las tablas de homologación suministradas por la DIG, las novedades cartográficas tales como reenumeración, fusión, división, creación y eliminación de manzanas pueden distinguirse como se muestra a continuación. Denótese COD_93 el código de manzana en la cartografía vieja y COD_XX el nuevo código asignado a ésta en la cartografía más reciente:

- **Manzana sin novedad:** Existe un solo registro – fila – para ella y ambos códigos

coinciden.

CODIGO_93	CODIGO_XX
0500110000000004020512	0500110000000004020512

- **Manzana renumerada:** Existe un solo registro – fila – para ella, ambos códigos son diferentes de cero pero no coinciden en las dos cartografías.

CODIGO_93	CODIGO_XX
0500110000000004020512	0500110000000004020518

- **Fusión de manzanas:** Un mismo código de la cartografía nueva corresponde a más de un código de la cartografía vieja.

CODIGO_93	CODIGO_XX
0500110000000004020512	0500110000000004020519
0500110000000004020513	0500110000000004020519
0500110000000004020514	0500110000000004020519

- **División de una manzana:** Un mismo código de la cartografía vieja corresponde a más de un código de la cartografía nueva.

CODIGO_93	CODIGO_XX
0500110000000004020512	0500110000000004020520
0500110000000004020512	0500110000000004020521
0500110000000004020512	0500110000000004020522

- **Creación de una manzana:** Como no existe en la cartografía antigua, el código correspondiente a ésta es 0.

CODIGO_93	CODIGO_XX
000000000000000000000000	0500110000000004020525

- **Eliminación de una manzana:** Como no existe en la cartografía nueva, el código correspondiente a ésta es 0.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 8 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	------------------------------------------------------------------------

CODIGO_93	CODIGO_XX
0500110000000004020512	0000000000000000000000

Note que todas las manzanas pertenecientes al marco deben estar presentes en la tabla de homologación: Si aún existen, deben tener un código asignado en la nueva cartografía y si han sido demolidas, deben aparecer con código nuevo igual a 0; es necesario llevar a cabo esta verificación antes de proceder con la actualización. En el otro sentido, no todos los códigos en la cartografía nueva tienen que estar presentes en el marco construido sobre cartografía 93.

Para efectuar la homologación se recomienda tener en cuenta los siguientes puntos:

1. Cree una copia del marco y haga las modificaciones sobre ésta, no sobre el marco original. Debe respetarse la integridad del marco antiguo como referente histórico.
2. Trate simultáneamente las manzanas que presentan una misma novedad cartográfica y trate por separado las distintas novedades. Esto permite que los procedimientos efectuados para la homologación sean adecuaciones de un estándar para responder a situaciones particulares.
3. Cree un archivo nuevo al que adicionar las manzanas, paulatinamente, conforme se resuelva cada una de las novedades.
4. Guarde una variable que indique qué tipo de novedad presentó la manzana en el momento de la homologación. Ocasionalmente, las tablas de homologación presentan errores que son detectados fácilmente en campo y guardar la novedad facilita el proceso de corrección de estos errores en el marco y su reporte a la DIG.
5. Recuerde que, adicional a la homologación cartográfica, es necesario incluir en el marco la información existente que corresponde a cantidad de establecimientos. En el marco con que se cuenta, dicha información tiene un identificador que permite saber su procedencia y con ello, la fecha de su actualización; es importante que esta variable se conserve.
6. La información que se incluya en cantidad de establecimientos depende de la novedad cartográfica que la manzana presente, por esto es tan importante que la homologación se haga en forma simultánea para manzanas con la misma novedad, y paulatinamente para las distintas novedades. Para el tratamiento cada una de ellas, se recomienda:
 - Manzanas sin novedad: Traslade la información de establecimientos y el identificador de procedencia de un marco al otro sin alteraciones.
 - Manzanas reenumeradas: Traslade la información de establecimientos y el identificador de procedencia de un marco a su lugar correspondiente en el otro sin alteraciones.
 - Fusión de manzanas: Sume la información correspondiente a las manzanas que participan en la fusión y actualice como dato el total. Coloque como identificador de procedencia de la información, el que corresponda según el marco original.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 9 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	------------------------------------------------------------------------

- División de manzanas: Tome la cantidad de establecimientos de la manzana en cuestión y divídala en partes iguales para asignarla a cada una de las manzanas que resultan de la división. Coloque como identificador de procedencia de la información el correspondiente.
- Creación de manzanas: Coloque como información de cantidad de establecimientos, valores 0. Como identificador de procedencia coloque la letra N, señalando que la manzana es nueva y el 0 no es un dato necesariamente cierto.
- Eliminación de manzanas: Simplemente no las incluya en el nuevo marco.

Adicionalmente a las novedades consideradas, es posible encontrar mezcla de ellas, por ejemplo, manzanas con características de fusión y división:

Caso 1.

CODIGO_93	CODIGO_XX
0500110000000004020502	0500110000000004020505
0500110000000004020503	0500110000000004020505
0500110000000004020505	0500110000000004020506
0500110000000004020505	0500110000000004020507
0500110000000004020505	0500110000000004020508

En este caso, las manzanas 02 y 03 probablemente fueron demolidas para construir las manzanas 06, 07 y 08 en esa área de terreno. Note que posiblemente, la manzana 05 no aparezca en la cartografía 93 y que no es posible determinar cuál o cuáles manzanas nuevas se encuentran en el área que correspondía a cada una de las manzanas 02 y 03, por esto se piensa en esta novedad como una fusión y división simultánea.

La inclusión de información para este tipo de novedad debe dar prioridad a la fusión y luego a la división esto es, sumar primero la cantidad de establecimientos de las manzanas presentes en la fusión y luego, dividir esta cantidad en partes iguales para distribuirla entre las manzanas resultantes de la división.

Caso 2.

CODIGO_93	CODIGO_XX
0500110000000004020508	0500110000000004020510
0500110000000004020510	0500110000000004020510
0500110000000004020508	0500110000000004020515
0500110000000004020508	0500110000000004020516

En el segundo caso, una parte de la manzana 08 se fusionó con la 10 y el trozo restante fue dividido en dos manzanas: la 15 y la 16, inexistentes en la cartografía 93. Para asignar la información correspondiente a cantidad de establecimientos, se recomienda dar

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 10 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

prioridad a la división de la manzana 02 en tres partes y luego, a la fusión de una de éstas con la manzana 03, esto es, dividir la información reportada por la 02 en tres partes iguales, asignar una de ellas a la 15, otra a la 16 y la restante, sumarla con la reportada por la 03 y asignarla a esta última.

Para solucionar inconvenientes con novedades de este tipo, es útil consultar a la DIG una explicación viable de esta novedad y, con base en ésta, generar una distribución de la cantidad de establecimientos que respete en lo posible, las áreas geográficas en cuestión. En todos estos casos, debe darse prioridad a las manzanas con alto número de establecimientos.

7. Finalmente, se recomienda homologar las manzanas agrupándolas en el siguiente orden: Manzanas sin novedad, Manzanas Renumeradas, Fusiones, Divisiones, Manzanas Nuevas, Otras Novedades.
8. Si existen manzanas en el marco que han dejado de pertenecer a la zona urbana del municipio, deben ser eliminadas.

Una vez concluido el proceso de homologación cartográfica y de inclusión de la información auxiliar existente a nivel de manzana, es conveniente adicionar la información disponible a nivel de municipio – entre ella, una variable que indique a qué cartografía corresponde cada uno luego de la homologación – y, estratificar el marco tanto a nivel de municipio como de manzana, conforme a los criterios metodológicos de la Encuesta Nacional de Microestablecimientos de Comercio, Servicios e Industria.

De igual manera, para mantener la memoria histórica del proceso y garantizar su verificabilidad, debe guardarse en un CD el marco antiguo, el marco nuevo, las tablas de homologación utilizadas y el o los programas SAS que fueron utilizados para efectuar la homologación, debidamente documentados.

Como revisión de calidad al nuevo marco, es útil llevar a cabo un análisis estadístico sobre la presencia de novedades cartográficas por municipio, y comparar, tanto la estructura de estratificación de ambos marcos, como la distribución por municipio y general, de la información auxiliar contenida en él. Note que, siempre y cuando las modificaciones cartográficas no sean muy drásticas, la distribución de las variables que corresponden a información auxiliar a nivel de manzana, no deben variar sustancialmente.

El éxito en todo este proceso estriba en el cuidado que se tenga sobre el manejo de las tablas de homologación y sus cruces con el marco; se requiere máxima precaución, conocimiento de la estructura del marco y de las tablas de homologación y, claridad y consistencia en la toma de decisiones frente a cada uno de los casos particulares, para que el mantenimiento del marco a nivel cartográfico sea un éxito.

Elementos que lo componen. La última actualización de marco se encuentra en medio magnético y su acceso se encuentra en la ruta <\\systema20\\Microest\\Marco\\Actualizacion\\I2007>.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 11 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

Diseño del registro.

Tabla 1 Especificación variables de ubicación e identificación del marco muestral

Variable	Tamaño	Clase	Descripción
Depmun	5	Identificación y ubicación	Código del departamento y el municipio
Clase	1	Identificación y ubicación	Código de área rural (0) o urbana (1)
Sect_rural	3	Identificación y ubicación	Código del sector rural = 000
Secc_rural	2	Identificación y ubicación	Código de la sección rural=00
Cent_pob	3	Identificación y ubicación	Código del Centro poblado=000
Sector	4	Identificación y ubicación	Código del sector cartográfico urbano
Sección	2	Identificación y ubicación	Código de la sección cartográfica urbana
Manzana	3	Identificación y ubicación	Código de la manzana cartográfica urbana
Industria	8	Auxiliar	Cantidad de establecimientos industriales por manzana
Comercio	8	Auxiliar	Cantidad de establecimientos comerciales por manzana
Servicio	8	Auxiliar	Cantidad de establecimientos de servicios por manzana

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 12 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Problemas del marco: Tradicionalmente, uno de los principales inconvenientes para las investigaciones de tipo económico que se llevan a cabo en el DANE, ha sido la imposibilidad de contar con marcos estadísticos actualizados, en donde sea posible identificar claramente y ubicar los establecimientos comerciales, industriales y de servicios en funcionamiento en el país.

Si bien este marco tiene una gran cantidad de usuarios potenciales, su más importante usuario y contribuyente hasta la fecha, ha sido la Encuesta Nacional de Microestablecimientos de Comercio, Servicios e Industria. Como ya se mencionó, estas unidades con diez o menos personas ocupadas por establecimiento, presentan una gran dinámica además de una muy baja cultura de registro, siendo difícil su estudio empleando marcos de lista.

Otro problema que presentaba el marco era la desactualización cartográfica, inconveniente que se está resolviendo con el proceso de actualización del marco. Para el estrato 1 se realizaron las respectivas homologaciones a cartografía 2004, quedando pendiente las homologaciones de Bogotá y área metropolitana de Medellín que permanecen con cartografía 1996 y 2001, respectivamente. En los estratos restantes se viene realizando la conformación del marco con la información económica de Censo General 2005, donde, no se tiene información de 55 municipios de 1072 que conforman los estratos 2, 3 y 4. Además con la base del censo se estableció la cantidad de comercios, servicios e industrias y posterior estratificación del marco, para municipios disponibles. Con esta información se incurre en subcobertura a nivel de USM, ya que se cuenta sólo con la información de manzanas que tienen por lo menos un establecimiento. La subcobertura se ha resuelto a nivel de muestra de UPM para los dos primeros trimestres de 2007, mediante las homologaciones realizadas por la DIG. Por tanto, el marco se actualizará por completo cuando se hayan realizado las respectivas homologaciones para todos los municipios del país y se actualice la información económica del Censo 2005.

3.7 PARÁMETROS A ESTIMAR


La característica de la población objetivo es estimar el total y sus varianzas estimadas y se presentan a continuación con mas detalle en el diseño muestral.

3.8 DISEÑO MUESTRAL

3.8.1. Método de muestreo:

El método de muestreo para éste estudio es Probabilístico Multietápico Estratificado de conglomerados para algunos estratos y de elementos para otros.

3.8.2. Justificación:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 13 FECHA: 23-10-07</p>
----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

3.8.2.1 Probabilístico:

Las unidades de muestreo en cada una de las etapas (municipios / áreas metropolitanas, manzanas, *secciones*¹ y establecimientos) tienen una probabilidad de selección conocida y superior a cero de ser seleccionadas.

3.8.2.2 Multietápico:

Dado que no existe un marco actualizado de todos los microestablecimientos económicos del país, es necesario llegar a ellos a través de unidades muestrales que los contengan, como municipios o manzanas cartográficas. Debido a los altos costos que implicaría la selección directa de manzanas, el diseño cubre las siguientes etapas:

- Primera etapa: selección de municipios ó áreas metropolitanas (UPM).
- Segunda etapa: selección de manzanas o *secciones* (USM), al interior de las UPM's seleccionadas en la etapa anterior.
- Tercera etapa: selección de establecimientos económicos (UFM) dentro de las USM's.

3.8.2.3. Estratificado:

Se consideran particiones poblacionales, definidas como conjuntos de elementos cuya unión conforma el universo, que no se traslapan y donde todos son diferentes de vacío, esto tratando de lograr homogeneidad dentro de ellas y heterogeneidad entre ellas.

En general se estratifica por las siguientes razones:

- Necesidad del usuario de contar con estimadores individuales para cada estrato
- Diferencia de disponibilidad de información auxiliar que permite aplicar, en algunos estratos diseños proporcionales al tamaño o estimadores más sofisticados.
- Diferencia de características operativas, vale decir de costos, tasas y velocidad de respuesta, distancias, aspecto de seguridad, etc., que hacen que el diseño que es eficiente en un estrato, no lo sea en otro.
- Concentración y distribución de la característica en estudio, que permite reducir considerablemente la varianza del estimador, al ampliar la muestra en aquellos estratos en los que se obtiene mayor ganancia en eficiencia y reducirla en los que la pérdida no es muy grande.

En éste diseño en particular se estratificó en la primera, a continuación veremos la construcción de la estratificación para cada etapa.

¹ La sección corresponde a las secciones construidas a partir de sustraer la sección cartográfica original, de las manzanas que se encuentran en otro estrato.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 14 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

⌚ Primera etapa:

En ésta etapa las unidades son los municipios del país.

Justificación:

- ★ Interés en tener cobertura geográfica similar para las diferentes investigaciones económicas realizadas en el DANE.
- ★ Posibilita hacer más eficiente el diseño muestral, logrando disminución en la varianza de las variables de interés en el estudio y por ende mejores CV's.

- ★ Otro punto muy importante en la decisión de estratificar los municipios es la optimización de la utilización del recurso de logística de campo, al poder diseñar mejor las rutas y hacer más eficiente el trabajo de recolección.

Criterios:

Se crearon 4 estratos para los municipios, así:

- ★ Estrato 1: Incluye las 24 áreas metropolitanas cubiertas.
- ★ Estrato 2: Incluye los municipios restantes para los cuales se levantó información durante el estudio "Construcción de marco de áreas para las investigaciones económicas", con ICF mayor a 1000
- ★ Estrato 3: Municipios con Índice de Complejidad Funcional Urbana (ICF) en el intervalo [150,1000].
- ★ Estrato 4: Municipios con ICF menor que 150.

Tabla 2 Municipios incluidos en las áreas metropolitanas del estrato 1

Depmun	Departamento	Municipio
11001	Bogotá D.C.	Bogotá D.C.
25754	Cundinamarca	Soacha
05001	Antioquia	Medellín
05079	Antioquia	Barbosa
05088	Antioquia	Bello
05129	Antioquia	Caldas
05212	Antioquia	Copacabana
05266	Antioquia	Envigado
05308	Antioquia	Girardotá
05360	Antioquia	Itagüí
05380	Antioquia	La Estrella

METODOLOGIA DEL DISEÑO MUESTRAL

ENCUESTA DE MICROESTABLECIMIENTOS MICRO

CÓDIGO: DM-MICRO-DIM-01
 VERSIÓN : 3
 PÁGINA: 15
 FECHA: 23-10-07

05631	Antioquia	Sabaneta
08001	Atlántico	Barranquilla
08758	Atlántico	Soledad
13001	Bolívar	Cartagena
15001	Boyacá	Tunja
17001	Caldas	Manizales
17873	Caldas	Villamaría
18001	Caquetá	Florencia
19001	Cauca	Popayán
20001	Cesar	Valledupar
23001	Córdoba	Montería
27001	Chocó	Quibdó
41001	Huila	Neiva
44001	La Guajira	Riohacha
47001	Magdalena	Santa Marta
50001	Meta	Villavicencio
52001	Nariño	Pasto
54001	Norte De Santander	Cúcuta
54405	Norte De Santander	Los Patios
54874	Norte De Santander	Villa Del Rosario
63001	Quindío	Armenia
66001	Risaralda	Pereira
66170	Risaralda	Desquebradas
66400	Risaralda	La Virginia
66682	Risaralda	Santa Rosa De Cabal
68001	Santander	Bucaramanga
68276	Santander	Floridablanca
68307	Santander	Girón
68547	Santander	Piedecuesta
73001	Tolima	Ibagué
76001	Valle Del Cauca	Cali
76892	Valle Del Cauca	Yumbo
88001	Departamento A	San Andrés

⌚ Segunda etapa:

Al interior de cada uno de los estratos conformados en la primera etapa se construyó una estratificación diferente.

- Estrato 1:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 16 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Compuesto por las 24 áreas metropolitanas, para la subestratificación se consideraron los siguientes criterios:

- ★ Aumentar la eficiencia de los estimadores, disminuyendo la variabilidad del total de ventas y personal ocupado entre las USM. Dado que existe alta correlación entre estos valores y la cantidad de establecimientos (tamaño) por USM, se homogenizó el tamaño de las USM en cada subestrato.
- ★ Optimizar el trabajo de recolección de información en campo, en términos de tiempo y desplazamiento.

Utilizando la información auxiliar del marco en éste estrato, se realizaron varios ejercicios para ver la distribución de frecuencias de los establecimientos, por manzana, analizando, básicamente, la cantidad de establecimientos industriales y comerciales, que finalmente se utilizó como criterio de estratificación.

Se decidió que algunos de los estratos para los dos tipos de establecimientos fueran diferentes, debido a que las densidades medias de comercio e industria son complementarias, es decir, donde hay mediana concentración de industrias hay pocos comercios y viceversa; ésta situación generaba que al tomarlos conjuntamente para realizar el cálculo del tamaño de muestra haciendo uso de una de las dos variables, (densidad de comercios o de industrias), se hiciera una sobreestimación del número de industrias o una subestimación de los comercios, respectivamente. Por tal razón los estratos resultaron del cruce de rangos de las densidades de acuerdo a que fueran altas, medias o bajas.


Subestrato 1:

Las unidades secundarias de muestreo (USM) en este subestrato son las manzanas cartográficas de alta densidad de comercio o industria, más de 70 comercios o más de 20 industrias, según el marco de áreas, que son comunes para la selección de muestra de comercio, industria y servicios.

El muestreo aquí es de elementos ya que la unidad final de muestreo es el establecimiento comercial, industrial o de servicios.

Subestrato 2:

Este subestrato es diferente en su construcción para industria que para comercio y servicios, debido a que es de concentración media de establecimientos (ver Cuadro 1). Ésta división se fundamenta en los gráficos de asimetría de las densidades de comercio e industria.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 17 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

La USM aquí es la manzana cartográfica. Dado que la unidad final de muestreo son los establecimientos, se dice que el muestreo es de elementos.

Subestrato 3:

Al igual que el subestrato 1, éste también es común para la selección muestral de Comercio e Industria, debido a que es el de más bajas densidades de establecimientos de los dos tipos.

La USM son *secciones*, según la definición ya mencionada.

Este es un muestreo de conglomerados, dado que las unidades finales son las *secciones* y se hace censo de todos los establecimientos industriales, comerciales y de servicios que existan en ella.

Subestrato 4:

Éste subestrato tiene construcción diferente (límites) para comercio y para industria (Cuadro 1), las USM son las manzanas dentro de las cuales se hace censo, por lo tanto el muestreo es de conglomerados. Su creación se justifica en los gráficos presentados anteriormente.

Los límites de los estratos se representan en el Cuadro 1, tanto para industria como para comercio.

Cuadro 1. Subestratos según densidad de establecimientos

		COMERCIO								
		> 70			[11 - 70]			<=10		
SERVICIOS		>70	(40 - 70]	<= 40	>70	(40 - 70]	<= 40	>70	(40 - 70]	<= 40
INDUSTRIA	> 20	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
	(5 - 20]	1 1	1 1	1 1	1 1	2 2	2 2	1 1	2 2	4 2
	<=5	1 1	1 1	1 1	1 1	2 4	2 4	1 1	2 4	3 3

Com y Ser

Industria

■ Com y Ser
■ Industria

- Estrato 2 :

En éste estrato se encuentran 236 municipios del país, entre ellos las capitales de departamento que no pertenecen al estrato 1 y algunos otros como: Palmira, Sogamoso, Girardot.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 18 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Se realizó subestratificación, básicamente con la misma justificación, que en el estrato 1:

- ★ Disminuir la variabilidad en las densidades de establecimientos por manzanas, controlando los corredores comerciales de los municipios y diferenciándolos claramente de las zonas residenciales.
- ★ Optimizar el trabajo de recolección de información en campo, al focalizar el esfuerzo de los recolectores de acuerdo a la importancia de cada estrato y sus establecimientos.

Los criterios de subestratificación se determinaron con base en el análisis de la información del marco de áreas de los municipios actualizado parcialmente con el Censo General 2005.

En éste estrato no se hará uso de doble estratificación para comercio e industria, dado que la diferencia entre las densidades no es tan alta.

Subestrato de inclusión forzosa:

Las USM son las manzanas de alta densidad de comercios, servicios o industrias (corredores comerciales). El muestreo aplicado es de elementos ya que la unidad final de muestreo son los establecimientos.

Subestrato de inclusión probabilística:

Las USM son las manzanas de baja densidad de comercios, servicios o industrias (residenciales). El muestreo aplicado es de conglomerados ya que se hace censo de todos los establecimientos industriales, comerciales y de servicios, en las manzanas seleccionadas

• Estrato 3:

En éste estrato se encuentran 835 municipios del país, entre ellos: Yolombó, Apulo, Anapoima, Santa Rosa de Viterbo.

En éste estrato no hay subestratificación de las manzanas, ya que son municipios muy pequeños y con poca actividad económica razón por lo cual se decidió realizar censo de establecimientos.

3.8.1 Tamaño de muestra

Para el cálculo del tamaño de muestra, tanto de comercio como de industria, se determinó un nivel de precisión global deseado (error estándar relativo o coeficiente de variación (CV)), entre el 5 y el 10 para las variables de interés: ventas, personal ocupado y total de establecimientos; este CV permite presentar la información a nivel estrato.

Al igual que los demás procedimientos realizados en ésta metodología el tamaño de muestra se calcula por etapas: primero municipios, luego manzanas o *secciones* y

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 19 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

finalmente establecimientos, dependiendo del estrato y el subestrato en los cuales se esté trabajando.

En general la determinación del tamaño de muestra se realiza por simulación siguiendo la metodología de Särndal (1992), en la cual se fija un CV y a partir de éste se halla el tamaño de muestra despejando n de la fórmula de la varianza del estimador según el diseño utilizado en cada caso.

$$CV = \left(\frac{\sqrt{V(\hat{t}_y)}}{\hat{t}_y} \right) * 100$$

Como el diseño es estratificado, al fijar el tamaño de muestra dentro de cada estrato se busca que éste permita alcanzar el CV fijado para el estudio. La estructura de estratificación cumple los principios de independencia e invarianza por lo cual en la formula para el CV de éste estudio, la varianza estimada del estimador resulta de sumar las obtenidas en cada estrato y con un procedimiento análogo se llega al estimador del total.

Fuentes de información:

- Comercio, servicio e industria:

Para el cálculo del tamaño de muestra de comercio, servicio e industria se hizo uso de la información de ventas obtenida de la Encuesta de microestablecimientos por áreas, realizada en los cuatro trimestres del año 2002.

3.8.1.1. Cálculo del Tamaño de Muestra:

Estrato1:

Para el cálculo del tamaño de muestra, se determinó un intervalo de oscilación del CV entre el 5 y el 10, para las variables de interés. En la primera etapa de la investigación se incluyeron las 24 áreas metropolitanas y dentro de cada una los subestratos ya mencionados. Dentro de cada uno de éstos subestratos se ajustaron el total y su varianza, estratificando la información disponible de la Encuesta de microestablecimientos

En el caso de la varianza el ajuste se hizo teniendo en cuenta el tamaño, en términos de USM del subestrato en cada área metropolitana; para el total se manejo la misma idea de ajuste por lo cual fue necesario estimar el promedio de las variables de interés en cada subestrato según la USM, luego se ponderó por el número de USM's en el subestrato. El

	<p>METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 20 FECHA: 23-10-07</p>
----------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

único criterio de ajuste para el total de las variables de interés en el área es que no exceda el del área inmediatamente anterior en importancia.

Se decidió hacer el siguiente ejercicio para poder determinar el tamaño de muestra:

- Bajo el supuesto de estratificar y aplicar al interior de cada estrato MAS, como alternativa para observar el tamaño de muestra óptimo, se utilizó la asignación de Neyman (Bautista, 1996) con miras a distribuir en los subestratos el tamaño de muestra determinado inicialmente. Las variables auxiliares fueron la densidad de comercio más servicio y la densidad de industria por manzana, que es más precisa para las 24 áreas metropolitanas ya que la genera el marco de áreas.

En las Tablas 3, 4 y 5 se presentan los tamaños de muestra definitivos para las 24 áreas metropolitanas según el subestrato y categoría.

Tabla 3 Tamaño de muestra para industria áreas metropolitanas por USM's.



METODOLOGIA DEL DISEÑO MUESTRAL

ENCUESTA DE

MICROESTABLECIMIENTOS

MICRO

CÓDIGO: DM-MICRO-DIM-01
VERSIÓN : 3
PÁGINA: 21
FECHA: 23-10-07

ÁREA METROPOLITANA		INDUSTRIA			
		SUBESTRATO 1	SUBESTRATO 2	SUBESTRATO 3	SUBESTRATO 4
Medellín	POBLACIÓN	95	271	1269	382
	MUESTRA	95	28	25	20
B/quilla	POBLACIÓN	45	53	785	292
	MUESTRA	45	10	30	10
Bogotá	POBLACIÓN	215	479	2863	2005
	MUESTRA	215	22	50	20
Cartagena	POBLACIÓN	38	4	420	105
	MUESTRA	38	4	30	16
Tunja	POBLACIÓN	5	1	102	55
	MUESTRA	5	1	15	10
Manizales	POBLACIÓN	17	4	235	92
	MUESTRA	17	4	30	14
Florencia	POBLACIÓN	2	0	83	33
	MUESTRA	2	0	15	20
Popayán	POBLACIÓN	3	0	151	65
	MUESTRA	3	0	15	12
Valledupar	POBLACIÓN	4	1	192	51
	MUESTRA	4	1	15	10
Montería	POBLACIÓN	4	0	181	58
	MUESTRA	4	0	20	20
Quibdó	POBLACIÓN	1	1	34	23
	MUESTRA	1	1	12	10
Neiva	POBLACIÓN	10	0	234	68
	MUESTRA	10	0	25	10
Riohacha	POBLACIÓN	14	0	88	551
	MUESTRA	14	0	15	25
Santa Marta	POBLACIÓN	23	3	243	61
	MUESTRA	23	3	20	12
V/cencio	POBLACIÓN	11	13	260	99
	MUESTRA	11	13	25	15
Pasto	POBLACIÓN	16	22	160	139
	MUESTRA	16	10	18	14
Cúcuta	POBLACIÓN	14	8	503	113
	MUESTRA	14	8	25	15
Armenia	POBLACIÓN	11	7	194	109
	MUESTRA	11	7	23	15
Pereira	POBLACIÓN	28	39	324	216
	MUESTRA	28	12	30	15
B/manga	POBLACIÓN	38	29	498	304
	MUESTRA	38	10	27	25
Sincelejo	POBLACIÓN	2	0	158	51
	MUESTRA	2	0	20	15
Ibagué	POBLACIÓN	11	11	244	87
	MUESTRA	11	11	20	17
Cali	POBLACIÓN	49	70	925	396
	MUESTRA	49	20	50	25
San Andrés	POBLACIÓN	9	0	18	31
	MUESTRA	9	0	12	10

Tabla 4 Tamaño de muestra para comercio más servicios áreas metropolitanas por

ÁREA METROPOLITANA		COMERCIO + SERVICIO			
		SUBESTRATO 1	SUBESTRATO 2	SUBESTRATO 3	SUBESTRATO 4
Medellín	POBLACIÓN	95	501	1269	152
	MUESTRA	95	18	30	20
B/quilla	POBLACIÓN	45	329	785	16
	MUESTRA	45	20	40	8
Bogotá	POBLACIÓN	215	2127	2863	357
	MUESTRA	215	20	50	20
Cartagena	POBLACIÓN	38	106	420	3
	MUESTRA	38	15	30	3
Tunja	POBLACIÓN	5	56	102	0
	MUESTRA	5	15	12	0
Manizales	POBLACIÓN	17	93	235	3
	MUESTRA	17	15	30	3
Florencia	POBLACIÓN	2	33	83	0
	MUESTRA	2	10	10	0
Popayán	POBLACIÓN	3	65	151	0
	MUESTRA	3	12	12	0
Valledupar	POBLACIÓN	4	52	192	0
	MUESTRA	4	10	12	0
Montería	POBLACIÓN	4	58	181	0
	MUESTRA	4	13	23	0
Quibdó	POBLACIÓN	1	24	34	0
	MUESTRA	1	10	12	0
Neiva	POBLACIÓN	10	68	234	0
	MUESTRA	10	12	23	0
Riohacha	POBLACIÓN	14	551	88	0
	MUESTRA	14	25	13	0
Santa Marta	POBLACIÓN	23	63	243	1
	MUESTRA	23	17	18	1
V/cencio	POBLACIÓN	11	109	260	3
	MUESTRA	11	20	30	3
Pasto	POBLACIÓN	16	158	160	3
	MUESTRA	16	15	18	3
Cúcuta	POBLACIÓN	14	117	503	4
	MUESTRA	14	12	20	4
Armenia	POBLACIÓN	11	115	194	1
	MUESTRA	11	10	18	1
Pereira	POBLACIÓN	28	242	324	13
	MUESTRA	28	12	30	13
B/manga	POBLACIÓN	38	325	498	8
	MUESTRA	38	20	25	4
Sincelejo	POBLACIÓN	2	51	158	0
	MUESTRA	2	10	12	0
Ibagué	POBLACIÓN	11	94	244	4
	MUESTRA	11	13	25	4
Cali	POBLACIÓN	49	415	925	51
	MUESTRA	49	25	50	10
San Andrés	POBLACIÓN	9	31	18	0
	MUESTRA	9	10	8	0

USM's.

Tabla 5 Tamaño de muestra por subestrato y categoría para comercio, servicios e industria áreas metropolitanas por USM's.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 23 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

Área	Subestrato	Categoría	n(maniz)	n(secc)	Área	Subestrato	Categoría	n(maniz)	n(secc)	Área	Subestrato	Categoría	n(maniz)	n(secc)
Medellín	1	1	95		Valledupar	1	1	4		Cúcuta	1	1	14	
	2	2	18			2	2	10			2	2	12	
	3	3	28			3	3	1			3	3	8	
	4	2	375,6	30		4	2	0			4	2	377,75	25
	4	3	20			4	3	10	15		4	3	4	
B/quilla	1	1	45		Montería	1	1	4		Armenia	1	1	11	
	2	2	20			2	2	13			2	2	10	
	3	3	10			3	3	0			3	3	7	
	4	2	608,4	40		4	2	414	23		4	2	325,22	23
	4	3	8			4	3	0			4	3	1	
Bogotá	1	1	215		Quibdó	1	1	1		Pereira	1	1	28	
	2	2	20			2	2	10			2	2	12	
	3	3	22			3	3	1			3	3	12	
	4	2	850	50		4	2	231,12	12		4	2	520,2	30
	4	3	20			4	3	0			4	3	13	
Cartagena	1	1	38		Neiva	1	1	10		B/manga	1	1	38	
	2	2	15			2	2	12			2	2	20	
	3	3	4			3	3	0			3	3	10	
	4	2	480	30		4	2	411,25	25		4	2	378	27
	4	3	3			4	3	0			4	3	4	
Tunja	1	1	5		Riohacha	1	1	14		Sincelejo	1	1	2	
	2	2	15			2	2	25			2	2	10	
	3	3	1			3	3	0			3	3	0	
	4	2	243,75	15		4	2	151,5	15		4	2	342,4	20
	4	3	0			4	3	0			4	3	0	
Manizales	1	1	17		Santa Marta	1	1	23		Ibagué	1	1	11	
	2	2	15			2	2	17			2	2	13	
	3	3	4			3	3	3			3	3	11	
	4	2	474,9	30		4	2	309,8	20		4	2	425	25
	4	3	3			4	3	1			4	3	4	
Florencia	1	1	2		V/cencio	1	1	11		Cali	1	1	49	
	2	2	10			2	2	20			2	2	25	
	3	3	0			3	3	13			3	3	20	
	4	2	222,45	15		4	2	513,3	30		4	2	900	50
	4	3	0			4	3	3			4	3	10	
Popayán	1	1	3		Pasto	1	1	16		San Andrés	1	1	9	
	2	2	12			2	2	15			2	2	10	
	3	3	0			3	3	10			3	3	0	
	4	2	249,75	15		4	2	298,8	18		4	2	201,6	12
	4	3	0			4	3	3			4	3	0	
			12					14		Tamaño Muestra Total = 11203				

Estrato 2:

El trabajo realizado con los 236 municipios que pertenecen a este estrato, para el cálculo de las varianzas y totales estimados, estuvo totalmente basado en la información disponible del Censo 2005 para 234 de ellos y de los cuales se determinó el total de establecimientos. Posteriormente la subestratificación se realizó mediante el método de Hidroglou.

Teniendo éstos valores se procedió al cálculo del tamaño de muestra para las UPM's a partir un diseño MAS.

El diseño a aplicar en este estrato será PPT pero el tamaño de muestra se calculó asumiendo MAS, buscando corregir una posible sobreestimación originada en la construcción misma del estrato y la poca disponibilidad de información.

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 24 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

Se determinó tomar 12 municipios que nos permiten tener un CV's: 12.5% para el total de establecimientos.

Luego se procedió a calcular los tamaños de muestra para las USM's en el subestrato de inclusión probabilística. Se determinó que el mayor peso a la varianza total se encuentra en las UPM's con un aporte del 94%, por lo tanto el tamaño de muestra de las USM's no influirá en el CV. Se seleccionarán como mínimo 3 secciones para los municipios con menos de 10 secciones y un máximo de 10 secciones para los municipios con más de 100 secciones.

Estrato	E2	Peso
Varianza UPMs	4038329880	94,49%
Varianza USMs	235706293	5,51%
Varianza Total	4274036173	100,00%
CV	12,88%	

Estrato 3:

En éste estrato sólo se hace cálculo de tamaño para UPM's debido a que dentro de éstas se realiza censo de establecimientos. Para determinar los totales y las varianzas se utilizó al igual que en el estrato anterior la información del económica del Censo General 2005 filtrando los municipios que pertenecen a éste estrato para hacer las estimaciones. El tamaño de muestra final fue 15 municipios, con base en criterios operativos, no se presentan CV's debido a que son mayores al 15%.

Debido a que no se puede determinar el aporte de cada estrato al CV global, puesto que la variable utilizada para determinar los tamaños de muestra en el estrato 1 fue el total de ventas y para los estratos 2 y 3, el total de establecimientos, se presentan los CV solamente por estrato.

Estrato	TOTAL POR ESTRATO		
	E1	E2	E3
Varianza Estimada	4,83342E+32	4291555309	996293863
CV	5,34%	12,88%	>15%

3.8.2 Método de selección

El diseño muestral específico aplicado para hacer la selección de las unidades muestrales esta determinado por la etapa y el estrato en que se está trabajando.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 25 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Durante éste capítulo se presentarán los métodos utilizados y el procedimiento de selección para cada estrato, en cada etapa del diseño.

3.8.2.1. Estrato 1:

⌚ Primera etapa:

Dado que por definición las áreas metropolitanas son de inclusión forzosa (IF), no aplica la utilización de un diseño específico ni de un método de selección.

⌚ Segunda etapa:

A continuación se enuncian los métodos para cada subestrato.

Subestrato 1:

Las UPM's son manzanas cartográficas con IF debido a que son de alta densidad de establecimientos.

La selección de establecimientos (tercera etapa del muestreo), dentro de éste estrato se realiza utilizando Muestreo Aleatorio Simple (MAS) sin reemplazamiento, de los establecimientos levantados en cada manzana durante el proceso de enlistamiento; la selección se hace por medio del algoritmo de Fan Muller (Anexo 1) y los criterios de selección de establecimientos se presentan en el Anexo 1.1.

Subestrato 2:

Las UPM's son manzanas cartográficas seleccionadas haciendo uso del diseño de Probabilidad Proporcional al Tamaño (PPT) con reemplazamiento. Este diseño resulta ser más eficiente que el MAS en poblaciones con distribuciones asimétricas, como se vio en las gráficas presentadas anteriormente para Bogotá y Soacha, como es el caso de las densidades de establecimientos comerciales, industriales y de servicios, la selección de las manzanas se hace utilizando el método Acumulativo Total (Anexo 2).

Dentro de éste subestrato también aplica la selección de establecimientos, y se realiza usando MAS de los establecimientos levantados en cada manzana durante el proceso de enlistamiento y para la selección también se hace uso del algoritmo de Fan Muller.

Subestrato 3:

Las UPM's son las *secciones*, usando la definición dada anteriormente. El diseño aplicado es MAS. La selección de las *secciones* a visitar se hace utilizando el algoritmo de Fan Muller.

	<p>METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 26 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Aquí el diseño es bietápico, ya que no hay selección de establecimientos sino que se hace censo en las *secciones* seleccionadas en la muestra.

Subestrato 4:

Las UPM's son manzanas cartográficas. El diseño utilizado es MAS, y la selección se hace usando el algoritmo de Fan Muller.

Este también es un estrato bietápico debido a que no hay selección de establecimientos sino que se hace censo de todos los que se encuentren en las manzanas seleccionadas.

3.8.2.2. Estrato 2:

⌚ Primera etapa:

El diseño utilizado para la selección de los municipios de éste estrato es PPT, haciendo uso de la información disponible en el marco de áreas. La selección se hará utilizando el método Acumulativo total.

⌚ Segunda etapa:

Subestrato de inclusión forzosa:

Las USM son manzanas cartográficas, clasificadas como de inclusión forzosa, razón por la cual no aplica diseño de muestreo ni método de selección.

Dentro de éste estrato se realizará tercera etapa de muestreo, haciendo uso del diseño MAS; la selección de los establecimientos será por medio del algoritmo de Fan Muller y los criterios de selección de establecimientos se presentan en el Anexo 1.2.

Subestrato probabilístico:

En éste subestrato las USM son secciones cartográficas. El diseño utilizado para la selección de la muestra es MAS de las secciones de baja concentración económica y la selección se hará con el algoritmo de Fan Muller.

Dentro de éste estrato se realizará tercera etapa de muestreo, haciendo uso del diseño MAS; la selección de los establecimientos será por medio del algoritmo de Fan Muller y los criterios de selección de establecimientos se presentan en el Anexo 1.2.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 27 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

3.8.2.3. Estrato 3:

 Primera etapa:

El diseño muestral utilizado en éste estrato es MAS. Para la selección de cada municipio se emplea el algoritmo de Fan Muller.

No hay más etapas dentro de éste estrato, debido a que se considera que éstos municipios son tan pequeños que se puede hacer censo de todos los establecimientos en las manzanas sin incrementar, sensiblemente, en tiempo y costos el trabajo de campo.

3.8.3. Metodología de estimación

Estimador

A continuación presentaremos los estimadores del total y sus varianzas estimadas; dado que la estructura de estratificación cumple los principios de independencia e invarianza la varianza estimada del estimador resulta de sumar las obtenidas en cada estrato y con un procedimiento análogo se llega al estimador del total.

3.8.3.1. Estrato 1:


Los resultados se presentarán teniendo en cuenta los subestratos existentes en éste estrato.

Subestrato 1

Estimador del total de la variable de interés:

$$\hat{t}_{p1} = \sum_{k=1}^{n_k} \frac{N_{kC}}{n_{kC}} \sum_{e=1}^{n_{kC}} y_{keC}$$

Estimador de la varianza:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 28 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

$$\hat{v}(\hat{t}_{p1}) = \sum_{k=1}^{n_k} \frac{N_{keC}}{n_{keC}} (N_{keC} - n_{keC}) s_{y_{keC}}^2$$

donde

$$s_{y_{keC}}^2 = \frac{1}{n_{keC} - 1} \sum_{e=1}^{n_{keC}} (y_{keC} - \bar{y}_{keC})^2$$

Subestrato 2

Estimador del total de la variable de interés:

$$\hat{t}_{p2} = \sum_{k=1}^{n_k} \left(\frac{\sum x_k}{m x_k} \right) \left(\frac{N_{kC}}{n_{kC}} \right) \sum_{e=1}^{n_{kC}} y_{keC}$$

Estimador de la varianza:

$$\hat{v}(\hat{t}_{p2}) = \frac{1}{m} * \frac{1}{m-1} \sum_{k=1}^m \left[\left(\frac{\sum x_k}{x_k} \right) \hat{t}_{kC} - \hat{t}_{p2} \right]^2$$

Subestrato 3

Estimador del total de la variable de interés:

$$\hat{t}_{p3} = \sum_{j=1}^{n_j} \frac{N_j}{n_j} \sum_{e=1}^{N_{jC}} y_{jeC}$$

Estimador de la varianza:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 29 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

$$\hat{v}(\hat{t}_{p3}) = \frac{N_j}{n_j} (N_j - n_j) s_{yjC}^2$$

donde

$$s_{yjC}^2 = \frac{1}{n_j - 1} \sum_{j=1}^{n_j} (y_{jC} - \bar{y}_c)^2$$

Subestrato 4

Estimador del total de la variable de interés:

$$\hat{t}_{p4} = \sum_{k=1}^{n_k} \frac{N_k}{n_k} \sum_{e=1}^{N_{kC}} y_{keC}$$

Estimador de la varianza:

$$\hat{v}(\hat{t}_{p4}) = \frac{N_k}{n_k} (N_k - n_k) s_{ykC}^2$$

donde

$$s_{ykC}^2 = \frac{1}{n_k - 1} \sum_{k=1}^{n_k} (y_{kC} - \bar{y}_c)^2$$

Total estrato 1:

Estimador del total:

$$\hat{t}_{ES1} = \sum_{p=1}^{24} \sum_{h=1}^4 t_{ph}$$

Estimador de la varianza:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 30 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

$$\hat{v}(\hat{t}_{ES1}) = \sum_{p=1}^{24} \sum_{h=1}^4 \hat{v}(\hat{t}_{ph})$$

3.8.3.2. Estrato 2:

Subestrato de inclusión forzosa:

Estimador del total:

$$\hat{t}_{11} = \sum_{k=1}^{n_k} \hat{t}_{kC} \quad , \text{ donde } \quad \hat{t}_{kC} = \frac{N_{kC}}{n_{kC}} \sum_{e=1}^{n_{kC}} y_{keC}$$

Estimador de la varianza:

$$\hat{v}(\hat{t}_{11}) = \frac{1}{p} * \frac{1}{p-1} \sum_{l=1}^p \left[\left(\frac{\sum x_l}{x_l} \right) \hat{t}_{kC} - \hat{t}_{11} \right]^2$$


Subestrato de inclusión probabilística:

Estimador del total:

$$\hat{t}_{12} = \sum_{j=1}^{n_j} \left(\frac{\sum x_j}{\frac{SEIP}{mx_j}} \right) y_{jC}$$

donde y_{jC} es el total de la variable de interés en la sección j

Estimador de la varianza:

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 31 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

$$\hat{v}(\hat{t}_{l2}) = \frac{1}{p} * \frac{1}{p-1} \sum_{l=1}^p \left[\left(\frac{\sum x_l}{x_l} \right) y_{jc} - \hat{t}_{l2} \right]^2$$

 **Total estrato 2:**

Estimador del total:

$$\hat{t}_{ES2} = \frac{1}{p} \sum_{l=1}^p \left(\frac{\sum x_l}{px_l} \right) \sum_{h=1}^2 \hat{t}_{lh} \quad \text{donde} \quad h: \begin{cases} 1 & \text{SEIF} \\ 2 & \text{SEIP} \end{cases}$$

Estimador de la varianza:

$$\hat{v}(\hat{t}_{ES2}) = \sum_{l=1}^p \sum_{h=1}^2 \hat{v}(\hat{t}_{lh})$$


3.8.3.3. Estrato 3:

Estimador del total:

$$\hat{t}_{ES4} = \frac{N_i}{n_i} \sum_{i=1}^{n_i} y_i$$

Estimador de la varianza:

$$\hat{v}(\hat{t}_{ES4}) = \frac{N_i}{n_i} (N_i - n_i) s_{y_i}^2 \quad \text{donde} \quad s_{y_i}^2 = \frac{1}{n_i - 1} \sum_{i=1}^{n_i} (y_i - \bar{y})^2$$

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 32 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

3.8.3.4. Estimadores globales:

Total:

$$\hat{t}_y = \hat{t}_{ES1} + \hat{t}_{ES2} + \hat{t}_{ES3}$$

Varianza:

$$\hat{v}(\hat{t}_y) = \hat{v}(\hat{t}_{ES1}) + \hat{v}(\hat{t}_{ES2}) + \hat{v}(\hat{t}_{ES3})$$

A continuación se presenta las probabilidades de selección para cada estrato y cada etapa, para facilitar el análisis se presenta en la siguiente tabla:

Tabla 6 Probabilidades de selección y de inclusión según Estrato y Etapa de muestreo

ESTRAT O	ETAPA 1	ETAPA 2					ETAPA 3							
		SE1	SE2		SE3	SE4	SE1		SE2		SE3		SE4	
1	1	1	C	I	$\frac{n_j}{N_j}$	$\frac{n_k}{N_k}$	C	I	C	I	C	I	C	I
			$\frac{x_k}{\sum_{SE2} x_k}$	$\frac{x_k}{\sum_{SE2} x_k}$			$\frac{n_{kC}}{N_{kC}}$	$\frac{n_{kI}}{N_{kI}}$	$\frac{n_{kC}}{N_{kC}}$	$\frac{n_{kI}}{N_{kI}}$	1	1	1	1
2	$\frac{x_j}{N_j}$ $\sum_{l=1} x_l$	SEIF			SEIP		SEIF				SEIP			
		1			C	I	C		I		C		I	
					$\frac{n_j}{N_j}$	$\frac{n_j}{N_j}$	$\frac{n_{kC}}{N_{kC}}$		$\frac{n_{kI}}{N_{kI}}$		$\frac{n_{kC}}{N_{kC}}$		$\frac{n_{kI}}{N_{kI}}$	
3	$\frac{n_i}{N_i}$	1					1							

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 33 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

x_l : Total de establecimientos de comercio, servicios e industria en el municipio l en el estrato 2 $l = 1, 2, \dots, 43$

n_i : Número de municipios seleccionados en el estrato i $i = 2$ o 3

N_i : Número total de municipios en el estrato i $i = 2$ o 3

x_k : Número de establecimientos de comercio o de industria en la manzana k

n_j : Número de secciones seleccionadas en el subestrato 3 o estrato 2

N_j : Número total de secciones en el subestrato 3 o estrato 2

n_k : Número de manzanas seleccionadas en el subestrato que corresponda

N_k : Número total de manzanas en el subestrato que corresponda

N_{kc} : Número total de establecimientos enlistados de comercio y servicios en la manzana k

N_{kl} : Número total de establecimientos enlistados de industria en la manzana k

n_{kc} : Número de establecimientos de comercio seleccionados en la manzana k

n_{kl} : Número de establecimientos de industria seleccionados en la manzana k

C: Comercio I: Industria SE: Subestrato SEIF: Subestrato de inclusión forzosa SEIP: Subestrato de inclusión probabil

Notas:

$$n_{kc} = [N_{kc} * R] + 1$$

$$n_{kl} = [N_{kl} * S] + 1$$

$[N_{ke} * R] + 1$ esta ecuación define el tamaño de muestra a seleccionar (aproximado en todos los casos por exceso), los corchetes cuadrados hacen referencia a la función parte entera.

El porcentaje de selección de UFM's se encuentra en los anexos 1.1 y 1.2.

Factores de expansión

En la siguiente tabla se presentan los factores de expansión según la etapa de muestreo, así como el total; las convenciones utilizadas son las mismas que en las probabilidades de selección.

Tabla 7 Factores de Expansión según estrato y etapa de muestreo

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 34 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

ESTRAT O	ETAPA 1	ETAPA 2				ETAPA 3							
		SE1	SE2		SE3	SE4	SE1		SE2		SE3		SE4
1	1	1	C	I	$\frac{N_j}{n_j}$	$\frac{N_k}{n_k}$	C	I	C	I	C	I	C I
			$\frac{\sum x_k}{mx_k}$	$\frac{\sum x_k}{mx_k}$			$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$	$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$	1	1	1 1
2	$\frac{\sum x_l}{p x_l}$	SEIF		SEIP		SEIF				SEIP			
		1		C	I	C	I	C	I	C	I	C	I
				$\frac{N_j}{n_j}$	$\frac{N_j}{n_j}$	$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$	$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$	$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$	$\frac{N_{kC}}{n_{kC}}$	$\frac{N_{kI}}{n_{kI}}$
4	$\frac{N_i}{n_i}$	1				1							

m : Número de manzanas seleccionadas

p : Número de municipios seleccionados


Nota : aquí aplican las convenciones presentadas en la Tabla 15

3.9 MANTENIMIENTO DE LA MUESTRA

La actualización de la información correspondiente a cantidad de establecimientos por sector económico proviene, en el caso de la Encuesta, de los archivos de Enlistamiento generados por cada operativo trimestral. Por esta razón es importante llevar a cabo un proceso minucioso de depuración de éstos, en busca de errores de codificación de manzanas, registros repetidos (bien sea a nivel de manzana o a nivel de establecimiento) y errores de clasificación en cuanto al sector económico.

Si bien en los mencionados archivos se encuentran conteos resumen por sector económico, por precaución, se recomienda que los conteos de cantidad de establecimiento se hagan directamente sobre el archivo a nivel de establecimientos, agrupando por manzana. En ocasiones, los conteos resumen, que son utilizados por los programas de captura para controlar el número de registros ingresados por manzana, vienen contaminados y es preferible llevar a cabo los cálculos nuevamente.

En cuanto a la periodicidad de la actualización, es necesario considerar tres casos distintos: De un lado están las manzanas de alta concentración económica, que son visitadas en panel todos los trimestres, de otro las manzanas de baja concentración, que son visitadas de manera ocasional y del otro, las manzanas que presentan cambios tan importantes en su concentración, como para migrar de uno de los grupos anteriormente considerados al otro. La indicación general es que las actualizaciones para el primer grupo se lleven a cabo cada dos trimestres, previa verificación de la consistencia de la información recolectada en los dos operativos consecutivos, la del segundo grupo, trimestralmente, salvo para aquellos casos que ameritan un cambio de estrato según la estratificación de la Encuesta y para el tercer grupo, trimestralmente, previa verificación en DANE Central de la identificación cartográfica de la manzana en cuestión. En éste

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 35 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

último caso, es recomendable también, que en el siguiente operativo se efectúe un nuevo recuento de la misma manzana para corroborar que la información actualizada en el marco sea consistente.

Antes de efectuar los conteos por sector económico que serán actualizados en el marco, es necesario considerar por aparte los establecimientos que aparecen como desocupados (E211=3) puesto que para ellos, el sector económico asignado no es real. La instrucción impartida en campo es que, al no poder asignar un sector económico a estos establecimientos, sean asignados al sector predominante en esa cara de manzana, pudiendo producir una sobre valoración del verdadero número de establecimientos por sector.

Para efectos de la actualización se recomienda ignorar este hecho, para los casos en que el número de establecimientos desocupados no supere el 15% del número total de establecimientos de la manzana, en cada uno de los sectores económicos. Cuando se presente una proporción mayor de establecimientos desocupados, es preferible aguardar un periodo y efectuar un nuevo recuento en la manzana, para garantizar la certeza de la información que el marco contiene. Manzanas con esta particularidad así como aquellas que migran de uno a otro estrato por cambiar drásticamente de concentración, deben ser enviadas como trabajo adicional a las regionales en el siguiente trimestre a su detección, para que se efectúen en ellas nuevos recuentos.

3.10 INDICADORES DE CALIDAD

La necesidad de un ajuste en los factores de expansión surge, básicamente, de la no respuesta que puede presentarse en cada una de las etapas del diseño muestral; existen otras formas de solucionar el problema de pérdida de muestra, destacándose entre ellas la imputación como una muy importante alternativa. Sin embargo, si no se cuenta con información para llevar a cabo una buena imputación y los porcentajes de pérdida de muestra se mantienen bajos – por ejemplo, por debajo del 5% – un ajuste puede subsanar el inconveniente sin mayores repercusiones.

Es importante aclarar que una modificación en los factores de expansión usualmente genera sesgo en los estimadores que se emplean para la generación de cifras. Los procedimientos de ajuste parten de suponer que el comportamiento de las unidades para las que no se tiene información es igual al de las unidades para las que sí se tiene, supuesto no verificable. En la medida en que el comportamiento real de las unidades sin información difiera del que se ha observado en las otras, las estimaciones producidas se alejarán del verdadero valor del parámetro; adicionalmente, los estimadores que se utilizan en la Encuesta son insesgados o aproximadamente insesgados, al modificar los factores de expansión es posible que estas propiedades se pierdan.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 36 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Hay unos pocos casos particulares en que es posible utilizar una herramienta alternativa para subsanar el problema de pérdida de información que se presenta en la Encuesta. Cuando se recolecta información de Estructura (anual) y de Evolución (mensual) y un establecimiento que puede rendir para ambas contesta la de Evolución y rechaza la de Estructura, es posible utilizar su información de Evolución para imputar la correspondiente a Estructura. De igual manera, cuando se produce un rechazo en bloque de los comerciantes de alguna manzana de alta densidad que es visitada todos los trimestres, es posible utilizar su información anterior para hacer una imputación de la nueva, en lugar de efectuar un ajuste de factores.

En qué casos se presenta la necesidad de un ajuste? En la primera etapa, en la selección de municipios, puede requerirse si por problemas de orden público es imposible acceder a un municipio, hecho que sucede con mayor frecuencia en los municipios de estratos 3 y 4. En la segunda etapa, en la selección de manzanas o secciones*, puede ser necesario un ajuste si ha sido imposible obtener información para algunas de ellas, bien sea por problemas de seguridad, imposibilidad de ubicarlas en cartografía o rechazo en bloque de los comerciantes de alguna manzana de alta densidad.

En cuanto a la última etapa, la de selección de establecimientos, se han detectado dos situaciones en las que se produce pérdida de muestra: En primer lugar, por motivos de organización y presupuesto, en algunas ocasiones es imposible completar el proceso de entrevista en las fuentes que han sido seleccionadas, existe un cierto número de establecimientos que pese a pertenecer a la muestra, no tienen formulario grabado en el archivo de Formularios. En segundo lugar, se produce pérdida de muestra cuando, a pesar de la insistencia del equipo de campo, el establecimiento seleccionado rechaza la entrevista, o se encuentra cerrado en todas las revisitas, o el emplazamiento aparece desocupado aunque en el enlistamiento no lo estaba. En estos casos, aparece grabado un formulario en el mencionado archivo, con resultado de encuesta R, C o D, según corresponda.

Con respecto a los establecimientos Fuera de Universo en el momento de la entrevista y los considerados como Nuevos, no son evaluados como pérdida de muestra; ninguno de ellos tiene información que rendir para el periodo de referencia que es de interés en la Encuesta.

En qué consiste el ajuste? Como se mencionó anteriormente, el procedimiento de ajuste aplicado en la Encuesta Nacional de Microestablecimientos de Comercio, Servicios e Industria consiste, fundamentalmente, en una distribución de los factores de expansión de los elementos que son considerados pérdida de muestra, entre aquellos que no lo son. La elección de cuáles son los elementos que se consideran pérdida de muestra se ha detallado en los párrafos anteriores y la de quiénes son aquellos que han de recibir parte de ese factor, depende de la etapa de diseño en que se produzca el ajuste. En todo caso, se procura que la distribución del factor de una unidad se haga, entre las unidades que participaron con ella en el proceso de selección.

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 37 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Dicho de otra manera, el ajuste de factores de un establecimiento enlistado como comercio, con resultado de encuesta Rechazo, se efectúa entre los demás establecimientos enlistados como comercio dentro de esa manzana. El ajuste de factores de una manzana que no fue encontrada en cartografía se lleva a cabo distribuyendo su factor de expansión entre las demás manzanas de su subestrato de diseño y ciudad (en el caso de subestratos 1, 2 y 4 y entre las demás de su sección* en el caso del subestrato 3). El factor de expansión de un municipio de estrato 4, perdido por problemas de orden público, se distribuye entre los demás municipios seleccionados como muestra en el estrato 4.

El ajuste de factores de expansión se lleva a cabo de atrás hacia delante, es decir, de la última etapa a la primera, obteniéndose un factor para cada etapa. Dicho factor puede variar de estrato en estrato para la primera etapa de selección, dentro de cada ciudad y subestrato, por sector económico para la segunda etapa y, dentro de cada manzana, por sector económico para la tercera etapa.

La razón por la cual el factor de ajuste para la segunda etapa puede variar para cada subestrato y ciudad, por sector económico es que, si por ejemplo, en una manzana todos los establecimientos de comercio entrevistados tienen resultado de encuesta cerrado, rechazo o desocupado, no existe un establecimiento para distribuir este factor. En este caso, la manzana puede figurar como sin respuesta para el sector comercio, pero no así para servicio o para industria, por eso es posible que requiera ser ajustada solamente para esta actividad. Por esta misma razón es necesario efectuar el ajuste desde la última etapa hasta la primera, para verificar los casos en que no hay una unidad que pueda actuar como receptora de los factores de las unidades sin información.

En cada etapa, para efectuar el ajuste, se multiplica el factor de expansión original de la unidad por un factor de ajuste por pérdida de muestra que, para el caso de la tercera etapa corresponde a

$$faj3_{ij} = \frac{n_{ij} - fu_{ij}}{n_{ij} - fu_{ij} - rcd_{ij} - (n_{ij} - form_{ij})}$$

con n_{ij} la cantidad de establecimientos seleccionada para la i-ésima actividad en la j-ésima manzana, fu_{ij} la cantidad de establecimientos fuera de universo encontrada en los formularios de la i-ésima actividad en la j-ésima manzana, rcd_{ij} la cantidad de formularios cerrados, rechazos o fuera de universo para la i-ésima actividad en la j-ésima manzana y $form_{ij}$ la cantidad de formularios encontrados para la i-ésima actividad en la j-ésima manzana. Como puede observarse, el factor de ajuste de la tercera etapa puede cambiar de manzana en manzana y dentro de una misma manzana, de una actividad de enlistamiento en otra.

Para el ajuste de la segunda etapa, el factor por el que se multiplica se convierte en

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 38 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

$$faj2_{ikl} = \frac{n_{kl}}{n_{kl} - r_{kl} - per_{ikl}}$$

con n_{kl} la cantidad de manzanas seleccionada en el k-ésimo subestrato de la l-ésima ciudad, r_{kl} la cantidad de manzanas que es necesario ajustar para todas las actividades y per_{ikl} la cantidad de manzanas que deben ajustarse para la i-ésima actividad en el k-ésimo subestrato de la l-ésima ciudad.

Como una de las causas del ajuste de manzanas es que todos los formularios de cierta actividad tienen como resultado de encuesta cerrado, rechazo o desocupado, el factor de ajuste de manzanas puede cambiar no solo para distintas combinaciones de subestrato y ciudad sino que, además, dentro de un mismo subestrato y ciudad puede tomar 3 valores diferentes: uno para cada actividad de enlistamiento.

En el estrato uno de municipios – Áreas Metropolitanas – el ajuste se hace no dentro de cada ciudad – subestrato sino dentro de cada combinación Área Metropolitana – subestrato. En los subestratos tres de los estratos uno y dos, como la unidad de selección es la sección* en lugar de la manzana, se ajustan las manzanas al interior de su sección*. Si llegase a perderse una sección* completa debería ser ajustada por las demás secciones* del municipio o Área metropolitana según corresponda. Esto es infrecuente.

Finalmente, si se perdiera un municipio por las causas anteriormente mencionadas, el factor que correspondería es

$$faj1_m = \frac{n_m}{n_m - per_m}$$

con n_m el número de municipios seleccionados y per_m el número de municipios que no reportan información.

Solamente pueden ajustarse municipios de estratos tres y cuatro, y en ningún caso puede ajustarse más del 10% del tamaño de muestra seleccionado. Previsiones similares deben tenerse en el ajuste de cada una de las etapas.

Una vez contruidos los factores de ajuste para cada una de las etapas, se recomienda efectuar las multiplicaciones correspondientes con los factores de expansión básicos, para trabajar con ellos en la etapa de procesamiento. Vale la pena resaltar que, utilizando esta metodología, los tamaños de muestra y de universo n y N que ingresan a la estimación de varianza son los originales, independientemente del ajuste efectuado.

Ahora, es necesario un paso adicional previo al proceso de estimación de varianza. Como se han ajustado los factores, se supone que la estimación que se efectúa es una

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 39 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------

estimación por dominios, sobre el dominio de los respondientes; por esta razón, es necesario que en el archivo de formularios que entra al procesamiento, estén todos los registros previstos conforme al esquema de muestreo utilizado, es decir, no pueden eliminarse los establecimientos con resultado de encuesta Cerrado, Rechazo o Desocupado, aunque hayan sido ajustados; deben ingresar como pertenecientes al dominio de No Respondientes, con valores 0 en las variables de procesamiento.

De igual manera, los formularios que no se encuentran en el archivo pero debieran estar, deben ser creados en las manzanas correspondientes y con la actividad de enlistamiento respectiva. También debe adicionarse un registro por cada manzana ajustada, para que ingrese a la estimación de varianza de la segunda etapa y en caso de haber ajustado municipios, un registro también por cada uno de ellos.

Es importante anotar que el ingreso de estos denominados Registros Cero, tiene un impacto menor en la estimación de varianza, pero no debe tener ningún efecto en el proceso de estimación de parámetros. Como se efectúa una estimación de dominios, los valores cero de estos registros no suman en los totales de las variables, pero tampoco entran en los conteos; es decir, si por ejemplo se está procesando el valor promedio de ventas por establecimiento, un cero de uno de estos registros no sumará en el total de ventas pero, por no pertenecer al dominio de los respondientes, tampoco contará en el número de establecimientos sobre el que se dividirá este total.

4. BIBLIOGRAFIA

Leonardo Bautista S. Diseños de Muestreo Estadístico. Universidad Nacional de Colombia. 1998.

Metodología del Diseño Muestral Etapa de Evolución III Trimestre Encuesta de Microestablecimientos – Micro 2004.

Lineamientos de Ajuste de Factores de Expansión. Encuesta de Microestablecimientos de Comercio, Servicio e Industria. Abril de 2003.

Lineamientos para el Mantenimiento de Marco de Áreas para las Investigaciones Económicas. Abril de 2003.

ANEXO 1

ALGORITMO DE FAN MULLER

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 40 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

Sea N : Tamaño del universo y sea n : Tamaño de la muestra a seleccionar.

Algoritmo:

- 1.1 Genere un número aleatorio λ_1 entre (0,1)
- 1.2 Decida si $\lambda_1 < n/N$, de ser así el primer elemento pertenece a la muestra.
- 1.3 Haga n_2 igual a 1 si el primer elemento fue seleccionado, 0 si no.
- 1.4 Determine C_2 igual a $(n-n_2)/(N-1)$

- 2.1 Genere un número aleatorio λ_2 entre (0,1)
- 2.2 Decida si $\lambda_2 < C_2$, de ser así el elemento 2 pertenece a la muestra.
- 2.3 Haga n_3 igual a la cantidad de elementos seleccionados en los dos ensayos anteriores.
- 2.4 Determine C_3 igual a $(n-n_3)/(N-1)$

En general,

1. Genere un número aleatorio λ_k entre (0,1)
2. Decida si $\lambda_k < C_k$, de ser así el elemento k pertenece a la muestra.
3. Haga n_k igual a la cantidad de elementos seleccionados en los $k-1$ ensayos anteriores.
4. Determine C_k igual a $(n-n_k)/(N-k+1)$
5. El proceso se detiene cuando $n_k=n$

ANEXO 1.1

INSTRUCCIONES PARA LA SELECCIÓN DE ESTABLECIMIENTOS E1

	METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO	CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 41 FECHA: 23-10-07
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------------------

EST.	SUB.	CAT.	COMERCIO	SERVICIOS	INDUSTRIA
1	1	1	Si $N_C \leq 5$ entonces $n_C = N_C$ Si $5 < N_C \leq 50$ entonces $n_C = 5$ Si $N_C > 50$ entonces $n_C = \lceil 0.1 \times N_C \rceil$	Si $N_S \leq 3$ entonces $n_S = N_S$ Si $3 < N_S \leq 30$ entonces $n_S = 3$ Si $N_S > 30$ entonces $n_S = \lceil 0.1 \times N_S \rceil$	Si $N_I \leq 12$ entonces $n_I = N_I$ Si $N_I > 12$ entonces $n_I = 12$
1	2	2	Si $N_C \leq 3$ entonces $n_C = N_C$ Si $3 < N_C \leq 30$ entonces $n_C = 3$ Si $N_C > 30$ entonces $n_C = \lceil 0.1 \times N_C \rceil$	Si $N_S \leq 5$ entonces $n_S = N_S$ Si $5 < N_S \leq 50$ entonces $n_S = 5$ Si $N_S > 50$ entonces $n_S = \lceil 0.1 \times N_S \rceil$	0
1	2	3	0	0	Si $N_I \leq 12$ entonces $n_I = N_I$ Si $N_I > 12$ entonces $n_I = 12$
1	3	1	Si $N_C \leq 5$ entonces $n_C = N_C$ Si $N_C > 5$ entonces $n_C = 5$ (Tabla de selección en campo)	Si $N_S \leq 5$ entonces $n_S = N_S$ Si $N_S > 5$ entonces $n_S = 5$ (Tabla de selección en campo)	$n_I = N_I$
1	4	2	Si $N_C \leq 5$ entonces $n_C = N_C$ Si $N_C > 5$ entonces $n_C = 5$	Si $N_S \leq 5$ entonces $n_S = N_S$ Si $N_S > 5$ entonces $n_S = 5$	0
1	4	3	0	0	$n_I = N_I$

Los cuadros señalados en azul representan las combinaciones de estrato, subestrato y categoría donde la selección se hace en campo empleando una tabla de selección de muestra previamente diseñada por el equipo y por tanto no se utiliza el programa de selección.

ANEXO 1.2

INSTRUCCIONES PARA LA SELECCIÓN DE ESTABLECIMIENTOS EN EL ESTRATO 2 Y 3

COMERCIO	SERVICIOS	INDUSTRIA
Sea N_C la cantidad de comercios en una manzana y n_C indica la cantidad de establecimientos de la actividad que se deben seleccionar en la manzana. Si $N_C \leq 5$ entonces $n_C = N_C$ Si $5 < N_C \leq 50$ entonces $n_C = 5$ Si $N_C > 50$ entonces $n_C = \lceil 0.1 \times N_C \rceil$	Sea N_S la cantidad de servicios en una manzana y n_S indica la cantidad de establecimientos de la actividad que se deben seleccionar en la manzana. Si $N_S \leq 5$ entonces $n_S = N_S$ Si $5 < N_S \leq 50$ entonces $n_S = 5$ Si $N_S > 50$ entonces $n_S = \lceil 0.1 \times N_S \rceil$	$n_I = N_I$
Si $N_C \leq 7$ entonces $n_C = N_C$ Si $N_C > 7$ entonces $n_{CS} = 7$	Si $N_S \leq 7$ entonces $n_S = N_S$ Si $N_S > 7$ entonces $n_{CS} = 7$	$n_I = N_I$

	<p align="center">METODOLOGIA DEL DISEÑO MUESTRAL ENCUESTA DE MICROESTABLECIMIENTOS MICRO</p>	<p>CÓDIGO: DM-MICRO-DIM-01 VERSIÓN : 3 PÁGINA: 42 FECHA: 23-10-07</p>
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------

ANEXO 2

ALGORITMO ACUMULATIVO TOTAL

Sea N el número de manzanas en el estrato y sea m : el número de manzanas a seleccionar.

Sea $P(j)$ tal como se definió anteriormente, tal que:

1. $P(1), P(2), \dots, P(N)$ son conocidos y $P(j) > 0 \quad \forall \quad j, j = 1, \dots, N$
2. $\sum_{j=1}^N P(j) = 1$

Algoritmo:

1. Construya la variable $P(j)$ acumulada
2. Para cada j genere la variable intervalo de selección: en la cual el $P(j)$ acumulado es el límite superior y el límite inferior esta dado por $P(j-1)$ acumulado.
3. Genere m números aleatorios (λ_k) , en el intervalo $[0,1]$.
4. Si λ_k pertenece al intervalo $[P(j-1), P(j)]$ entonces el elemento j pertenece a la muestra.