

**Departamento Administrativo
Nacional de Estadística**



Dirección de Censos y Demografía - DCD

**Diseño Conceptual y Metodológico
Censo General 2005 - CGRAL**

Abril 2005

	DISEÑO CONCEPTUAL Y METODOLÓGICO		CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 2 FECHA: 15-04-05
	CENSO GENERAL 2005 – CGRAL		
ELABORÓ: Equipo CM	REVISÓ: Coordinador CM	APROBÓ: Coordinador CGRAL	

CONTENIDO

Introducción.....	3
I. DEFINICIÓN DE CONCEPTOS.....	5
II. DISEÑO ESTADÍSTICO	9
2.1 Generalidades.....	9
2.2 Universo	9
2.3 Unidades Estadísticas	9
2.4 Marco Estadístico	9
2.4.1 Modelo Estadístico para la Definición de los Grupos de Conglomerados Homogéneos de Municipios.....	13
2.4.2 Análisis de Conglomerados.....	13
2.4.3 Resultados del modelo.....	15
2.4.4 Conformación de los Grupos de Conglomerados.....	19
2.5 Diseño Encuesta Cocensal.....	29
III. RESTITUCIÓN A UN ÚNICO MOMENTO ESTADÍSTICO.....	31
IV. CONTENIDO TEMÁTICO.....	36
4.1 Censo de Población y Vivienda.....	37
4.2 Unidades Económicas y Agropecuarias	40
4.3 Entorno Urbanístico.....	40
V. NORMAS PARA EL MANEJO DE LA INFORMACIÓN CENSAL DE POBLACIÓN Y VIVIENDA.....	40
ANEXO: INDICADORES DE CONTROL DE CALIDAD Y EVALUACIÓN	49



DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 3
FECHA: 15-04-05

Introducción

El Departamento Administrativo Nacional de Estadística – DANE como rector de las políticas de información estratégica para la toma de decisiones en todos los ámbitos de la sociedad, es la entidad oficial responsable de diseñar, implementar y realizar transferencia tecnológica de las metodologías que garanticen su generación y uso con los mayores estándares de calidad, cobertura y oportunidad para satisfacer las necesidades de los usuarios finales.

Este deber se ve aun mas, si se considera la fuerte dinámica de los cambios sociales, políticos, económicos que en la última década viene experimentando el país, lo cual ha determinado un rezago importante en información diagnóstica que permita el direccionamiento adecuado de las políticas públicas. En el contexto local, éste déficit de información es aun mas extremo, determinando grandes falencias en las administraciones para el direccionamiento adecuado de la inversión que permita con una visión de largo plazo garantizar un desarrollo económico sustentable y sostenible para lograr el mejoramiento de la calidad de vida de las personas. Por lo anterior, la información no solo es necesaria para agilizar y mejorar procedimientos, sino que tiene un gran valor estratégico en el proceso de toma de decisiones y de seguimiento de la política pública.

En este mismo orden de ideas, se ha identificado que uno de los principales obstáculos para el adecuado desarrollo de un proceso de descentralización es la baja capacidad de gestión de los gobiernos locales, motivado en su gran mayoría en la falta de información diagnóstica local que permita la evaluación y prospección de las variables básicas que determinan el mejoramiento económico y social de las comunidades.

Por lo anterior y, teniendo en cuenta que la mayoría de la información estratégica que se produce se realiza a través de investigaciones basados en diseños muestrales, en las cuales el problema está en su desagregación, se requiere el desarrollo de un gran macroproyecto que permita obtener información diagnóstica de todos los entes territoriales sobre su unidad básica, las personas, y las características donde viven. Este tipo de proyecto solo se puede lograr a través de los Censos de Población y Vivienda para cada una de las unidades territoriales, obteniéndose información estratégica que permita caracterizar a cada una de ellas e identificar los principales déficit que ocasionan problemas de pobreza, desplazamiento, desarrollo económico, crecimiento demográfico, servicios públicos, vivienda y en el marco de la normatividad la asignación de recursos fiscales básicos para el desarrollo de proyectos de grandes impactos en lo local y regional.

El primer eslabón en la estructuración de un Proyecto Censal Confiable, es la realización de un diseño conceptual y metodológico robusto desde todo punto de vista, el cual se constituye en la bitácora para el desarrollo e implementación de procesos básicos de éxito como la realización de presupuestos equilibrados, operativos de campo con todos los

 DANE	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 4 FECHA: 15-04-05
---	--	---

soportes necesarios y el diseño de toda la plataforma tecnológica que permita el acceso adecuado y oportuno de la información censal.

En este documento se presenta cada uno de los elementos conceptuales y metodológicos que se definen en los párrafos anteriores, así como el contenido temático del Censo General. En cuanto a la parte de Población y Vivienda, se definen los criterios para definir el contenido del cuestionario básico, así como las normas de tipo estructural, de consistencia y validación general e imputación necesarias para el desarrollo del software del cuestionario inteligente para las DMC así como para el proceso de depuración final de la información.

I. DEFINICIÓN DE CONCEPTOS.

Para la adecuada realización del Censo de General se debe manejar una serie de conceptos con el fin de garantizar la calidad de la información y su adecuado uso por parte de los diferentes usuarios.

Como el censo de Población y Vivienda tiene tres unidades de observación que permiten la estructuración adecuada de la información censal: Unidad de Vivienda, hogar particular y las personas residentes, todas ellas referidas a una unidad geográfica, el manejo adecuado de cada una de ellas, en cada uno de los procesos del Proyecto censal, son la garantía para obtener la información estratégica para cada uno de los entes territoriales. Para su identificación se tienen los siguientes conceptos fundamentales, de cuya adecuada aplicación depende la calidad de la información censal.

- **Edificación:** Es una construcción, independiente y separada, compuesta por una o más unidades.
 - ✓ **Independiente**, porque tiene acceso directo desde la vía pública, caminos, senderos o espacios de circulación común.
 - ✓ **Separada**, porque generalmente tiene paredes que la delimitan y diferencian de otras construcciones.
- **Unidad:** Es un espacio independiente y separado que constituye parte, o la totalidad de una edificación cuyos usos pueden ser económico, vivienda o mixto
 - ✓ **Económico:** cuando está destinado o está siendo utilizado para la industria, el comercio o los servicios;
 - ✓ **Industria:** Es la transformación física y/o química de materiales y componentes en productos nuevos, ya sea que el trabajo se efectúe con máquinas o a mano, en una fábrica o a domicilio, que los productos se vendan al por mayor o al por menor.
 - ✓ **Comercio:** Compra y venta al por mayor y al por menor de mercancías (bienes que van al mercado), nuevas y usadas y que no son sometidas a transformaciones en su naturaleza intrínseca durante las actividades inherentes a la comercialización, transporte, almacenaje, empaque, reempaque, embalaje, etc., o en aquellas a adecuar, exhibir, presentar o promocionar las mercancías objeto de la venta.
 - ✓ **Servicios:** Son productos heterogéneos generados cada vez que son solicitados, sobre los cuales no recaen derechos de propiedad por parte del usuario y no pueden ser negociados por separado de su producción ni pueden ser transportados ni almacenados.
 - ✓ **Actividad Económica Asociada a los Hogares:** Cuando en una vivienda se comparten los espacios propios del hogar para desarrollar con regularidad una actividad comercial, de servicios o industrial que genere ingresos.
 - ✓ **Actividad Agropecuaria Asociada a la Vivienda:** Solo es aplicable en Clase 3 (rural disperso). Se define de la siguiente forma: La actividad agropecuaria se identifica y ubica al interior de los predios rurales o fincas con vivienda, hogares y personas con residencia habitual. Para los fines del Censo General 2005 se ha adoptado esta categoría **como unidad de observación**, definida así: tierra donde

se ubica la vivienda rural con hogares y personas residentes habituales y además ocupada por cultivos agrícolas; transitorios y permanentes, cultivos forestales, forrajes o pastos y malezas y rastrojos; animales de crianza como bovinos, equinos, porcinos, cabras, ovejas, asnos, aves de corral; y cultivo de peces y camarones, que pertenecen a una unidad física y geográfica mayor denominada FINCA. Para fines censales no se tiene en cuenta el de tipo de administración, titulación, tenencia o tamaño de la unidad.

- ✓ Vivienda: cuando está habitado o destinado para ser habitado por personas;
- ✓ Mixto: cuando se combinan los dos usos: económico y vivienda;
- **Unidad de Vivienda**: Es un espacio independiente y separado con áreas de uso exclusivo, habitado o destinado a ser habitado por una o más personas. Con el fin de facilitar la cuantificación del stock de unidades de vivienda y el análisis de la información. para efectos censales se tiene la siguiente tipología:
 - ✓ **Casa**: Es la edificación constituida por una sola unidad cuyo uso es el de vivienda, con acceso directo desde la vía pública o desde el exterior de la edificación. El servicio sanitario y la cocina pueden estar o no dentro de ella. Para efectos censales, también se consideran casas aquellas en donde el garaje, la sala o alguna habitación se destina para uso económico.
 - ✓ **Casa Indígena**: Es la edificación constituida por una sola unidad cuyo uso es el de vivienda, construida según la costumbre de cada grupo étnico conservando la estructura tradicional.
 - ✓ **Apartamento**: Es una unidad de vivienda, que hace parte de una edificación, en la cual hay otra(s) unidad(es) que generalmente es (son) de vivienda. Tiene acceso directo desde el exterior o por pasillos, patios, corredores, escaleras o ascensores. Dispone de servicio sanitario y cocina en su interior.
 - ✓ **Cuarto**: Es una unidad de vivienda, que hace parte de una edificación y que dispone de uno o más espacios. Tiene acceso directo desde el exterior o por pasillos, patios, zaguanes, corredores u otros espacios de circulación común. En general carece de servicio sanitario y cocina en su interior, o sólo dispone de uno de estos dos servicios.
 - ✓ **Otro tipo de Vivienda**: Es un espacio adaptado para vivienda, donde en el momento de la entrevista habitan personas. Generalmente carece de servicio sanitario y cocina, tales como: los vagones de trenes, los contenedores, las embarcaciones, las tiendas de gitanos, las cuevas, los puentes, las casetas .
- **Residente Habitual**: Es la persona que vive permanentemente o la mayor parte del tiempo en una vivienda aunque en el momento de la entrevista se encuentre ausente. También son residentes habituales, porque no tienen residencia habitual en otra parte:
 - ✓ Las personas que se encuentran ausentes por motivos especiales como: vacaciones, cursos de capacitación, viajes de negocio; siempre y cuando la ausencia sea de 6 meses o menos.

- ✓ Los agentes viajeros, marinos mercantes.
 - ✓ Las personas secuestradas sin importar el tiempo de ausencia.
 - ✓ Los enfermos internados en hospitales o clínicas sin importar el tiempo de ausencia,
 - ✓ Las personas desplazadas, sin importar el tiempo de permanencia en el hogar que se está entrevistando.
 - ✓ Las personas detenidas temporalmente en inspecciones de policía.
- **Hogar:** Este concepto que se presenta a continuación es el resultado de un consenso entre todas las Direcciones Técnicas del DANE¹, con lo cual se busca la homologación de las diferentes investigaciones con el censo, teniendo en cuenta que éste conforma los marcos para las diferentes investigaciones. Igualmente se realizó una homologación con el utilizado por la CAN², el cual tiene algunas diferencias con el contexto colombiano El concepto de Hogar formulado fue el siguiente:

Es la persona o grupo de personas, parientes o no, que ocupan la totalidad o parte de una vivienda; atienden necesidades básicas, con cargo a un presupuesto común y generalmente comparten las comidas.
 - **Jefe o Jefa de Hogar:** Es el residente habitual reconocido como tal por los demás miembros de su hogar.
 - **Informante en el Hogar:** Es cada una de las personas residentes habituales mayores de 16 años. En el caso de no estar presente y no poderse conseguir en las visitas, la información la debe dar el Jefe(a) del Hogar, o su cónyuge o una persona residente mayor de 18 años que conozca la información de la persona. Para los menores de 18 años la debe dar el Jefe(a) del Hogar o su cónyuge.
 - **Lugar Especial de Alojamiento – L.E.A.:** Es una institución en la cual viven (duermen) colectivamente un grupo de personas generalmente no parientes.

Adicionalmente existen otros conceptos básicos en el desarrollo del censo. Unos están relacionados con el contenido temático los cuales se pueden ver en el Anexo sobre Glosario de términos. Otros se relacionan con la modificación que el DANE ha realizado en cuanto a la ejecución del censo general 2005, conceptos que por su relevancia metodológica se enuncian a continuación:

- **Conglomerado:** Conjunto de municipios contiguos que conforman un área geográfica homogénea en los aspectos socio-económicos, culturales y demográficos.
- **Unidad de Barrido:** área geográfica que se puede censar con un equipo de seis encuestadores y un supervisor durante un día.

¹ El DIRPEN como Dirección Técnica encargada de la estandarización de los Sistemas Estadísticos oficializa dicho concepto

² Es la persona o conjunto de personas sean parientes o no que ocupen en su totalidad o en parte de una vivienda; comparten al menos las comidas principales y atienden en común otras necesidades básicas, con cargo a un presupuesto común



DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 8
FECHA: 15-04-05

- **Período Censal:** tiempo de duración de la recolección de la información para uno o varios conglomerados.
- **Momento censal:** Es la 00:00:00 horas del día de inicio del operativo de recolección para cada conglomerado o grupo de conglomerados.
- **Población censada:** Es el cantidad de hombres y mujeres por edad que se obtiene para cada municipio durante el período censal definido para la recolección.
- **Momento Estadístico Único:** Fecha única de referencia definida para restituir la información de todos los municipios. Dado que cada grupo de conglomerados tiene momentos censales diferentes es necesario, restituir la información censal a una única fecha de referencia, que permita restituir la información a niveles geográficos agregados, departamento, región y nacional. Para este efecto, se realizara un proceso de conciliación de acuerdo a las características de cada conglomerado y partiendo de la ecuación compensadora se llevan todos los municipios a la fecha más próxima a la recolección de los grupos de conglomerados de mayor dinámica. En este orden de ideas, se toma como fecha de referencia única para la información censal el momento censal de los conglomerados C4 y C5³.

³ se adopta como fecha la mediana de dicho período de recolección, correspondiendo el 11 de noviembre.

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 9 FECHA: 15-04-05
---	--	---

II. DISEÑO ESTADÍSTICO

2.1 Generalidades.

El Censo General 2005 consta de: censo de población, de hogares y de viviendas; la conformación de marco estadístico de unidades económicas y de unidades agropecuarias asociadas a vivienda. El censo de población y de vivienda se diseñó con una encuesta básica para toda la población y una encuesta por muestreo probabilístico para profundizar en algunos temas. Igualmente, como el censo se ejecuta durante 6 meses, se clasificaron los municipios en 5 conglomerados según movilidad con el fin de minimizar el sesgo por migración y censar primero los conglomerados con menos movilidad y continuar con los de movilidad moderada hasta censar de último los de movilidad severa.

2.2 Universo

El universo, que coincide con la población objetivo, está constituido por el área total del país 114´174.800 hectáreas, conteniendo las áreas agropecuarias, como las no agropecuarias, las urbanas, las semiurbanas, los parques naturales, los bosques, las áreas erosionadas y cuerpos de aguas.

2.3 Unidades Estadísticas

Vivienda

Hogar

Personas residentes habituales

Unidad económica

Unidad agropecuaria

2.4 Marco Estadístico

La adecuación del marco estadístico en conglomerados es uno de los aspectos metodológicos de mayor importancia para el desarrollo adecuado del censo de Población y Vivienda. Su importancia radica en el hecho que permite minimizar los sesgos factibles por factores determinantes de la dinámica de la población residente en cada una de las

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 10 FECHA: 15-04-05
---	--	--

unidades geopolíticas en las cuales está dividido el país, como es el caso de la movilidad espacial⁴ y del crecimiento vegetativo o natural⁵.

La realización del proceso de conformación de conglomerados, se estructuró en las siguientes fases:

- **Fase 1:** Estructuración de la información municipal con variables de migración, fecundidad, mortalidad, desplazamiento forzado y densidad de población, todas ellas sintomáticas de desarrollo.
- **Fase 2:** Desarrollo de los procesos estadísticos⁶ para identificar las variables que discriminen adecuadamente, para minimizar los sesgos provenientes de la movilidad a nivel municipal.
- **Fase 3:** Desarrollo de los procesos multivariados⁷ para la conformación de los conglomerados homogéneos en cuanto a movilidad. En este proceso se incluyen factores condicionantes como:
 - ✓ Los municipios de la Orinoquía, Amazonia y los del andén Pacífico del Chocó, Valle, Cauca y Nariño (censo por rutas), no se incluyen en el análisis de conglomerados;
 - ✓ Municipios nuevos de los cuales no se tiene información suficiente, se deben incluir en los conglomerados de los municipios de donde fueron segregados;
 - ✓ Bogotá como ciudad región debe estar en el mismo conglomerado de los municipios de mayor influencia por su cercanía;
 - ✓ Las principales áreas metropolitanas o ciudades menores, deben estar con su área de influencia, lo cual implica la identificación de estos centroides.
 - ✓ Se tendrá en cuenta la clasificación subregional que tienen las diferentes Sedes y Subsedes del DANE, lo cual determina una dinámica propia en cuanto a la clasificación de los centroides.
- **Fase 4:** Conformación final de los Grupos de Conglomerados, los cuales a sus características determinan el orden en el tiempo en el cual el censo debe ser ejecutado en cada conglomerado y la formulación final de un modelo matemático de restitución de la población, a un momento estadístico único.

Teniendo en cuenta las características de cada una de las diferentes regiones, la ubicación de los centroides y su categoría, la movilidad poblacional sea esta aleatoria u motivada por factores exógenos, se generan interacciones entre municipios de diferentes regiones o a su interior.

El considerar este tipo de dinámica es el principal factor para garantizar que el sesgo por movilidad tanto en los niveles y estructuras de la información censal se minimice. Los grandes centroides como son Bogotá, Cali, Medellín y Barranquilla determinan

⁴ Como es el caso de la migración de la población motivada por factores educativos, laborales, ambientales, entre otros, así como los determinados por razones de preservar la vida debido a las amenazas de grupos armados fuera de la Ley.

⁵ Determinado por las defunciones y nacimientos de cada unidad geopolítica.

⁶ Método de Componentes Principales.

⁷ Análisis factorial jerárquico.



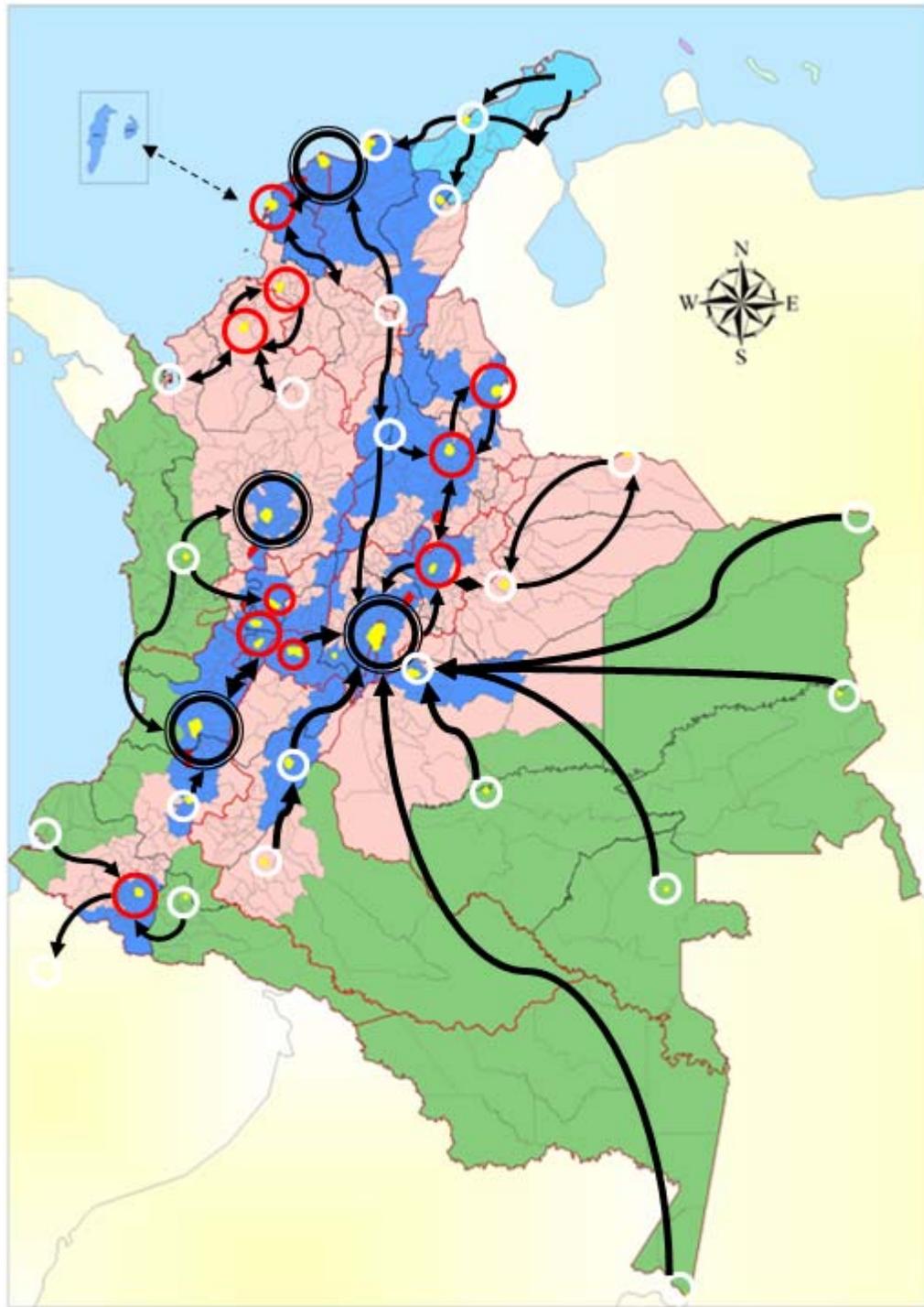
**DISEÑO CONCEPTUAL Y
METODOLÓGICO
CENSO GENERAL 2005 – CGRAL**

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 11
FECHA: 15-04-05

corredores en donde se encuentran centroides sub - regionales importantes para los procesos en cadena de los movimientos migratorios. Es el caso de Pereira, Manizales, Armenia, Ibagué en la parte centro occidental; Popayán, Pasto y Neiva en la parte sur; Villavicencio en la parte oriental; Barrancabermeja, Bucaramanga y Cúcuta en parte centro oriental; Valledupar Cartagena, Barranquilla, Santa Marta y Riohacha en la parte norte del país.

Se identifica igualmente un eje importante que por sus características agroindustriales y mineras se genera una dinámica interna en dichas sub regiones. Es el caso del eje Uraba – Montería – Sincelejo con el área de influencia hasta el norte del departamento de Antioquia: Igualmente se tiene el eje Arauca – Yopal de importante impacto agropecuario y petrolero.

Esto se puede observar en el siguiente mapa.



	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 13 FECHA: 15-04-05
---	--	--

2.4.1 Modelo Estadístico para la Definición de los Grupos de Conglomerados Homogéneos de Municipios.

El objetivo de éste ejercicio es, conformar, a partir de datos estadísticos de muy variada naturaleza, agrupaciones de municipios que presenten, de conformidad con las estadísticas disponibles para cada uno de ellos condiciones similares que los aglutine en grupos homogéneos, con el fin de utilizar dichas agrupaciones como parte de la estrategia para realizar el Censo General 2005 en el territorio colombiano.

En particular se plantean dos fines primordiales. En primer lugar, se pretende una aproximación a la caracterización de la situación socio-demográfica de los municipios colombianos en la actualidad. Complementariamente, se desea obtener una tipología municipal que ilustre el análisis y permita avanzar en el establecimiento de una clasificación estadística basada en aspectos socio-demográfica para fines censales.

Dado un conjunto de individuos (en este caso municipios) y teniendo de cada uno de ellos una información en forma de variables (indicadores parciales), el análisis de conglomerados trata de clasificarlos en grupos (no determinados "a priori") de manera que los individuos pertenecientes a un grupo sean, con respecto a la información que se dispone, tan similares como sea posible y que cada grupo se distinga del resto lo máximo que se pueda en términos de distancias multivariantes.

La información que se ha utilizado para la conformación de conglomerados, proviene de las siguientes fuentes:

- ✓ Información sobre evolución de la población municipal proporcionada por los censos nacionales de población y vivienda de los años 1985 y 1993
- ✓ Las estadísticas de nacimientos y defunciones, ocurridas en el país clasificadas de conformidad con el lugar de residencia de la madre o residencia habitual del fallecido.
- ✓ Cifras de proyecciones de población municipales elaboradas por el DANE.
- ✓ Información sobre educación específicamente matrícula primaria y secundaria.
- ✓ Desplazamiento según información suministrada por la red de Solidaridad Social
- ✓ Información sobre el Número de electores inscritos
- ✓ Valor de ingresos tributarios municipales

2.4.2 Análisis de Conglomerados

El análisis de conglomerados ("cluster analysis") se apoya en una serie de técnicas (fundamentalmente algoritmos), que tienen por objeto la búsqueda de grupos similares de individuos que se engloban en conglomerados lo más homogéneos posibles. Estos no son conocidos de antemano sino sugeridos por la información que contienen los datos.

Fue necesario realizar pruebas de los distintos métodos y técnicas existentes hasta obtener un resultado que se ajustase adecuadamente a los objetivos buscados. A lo largo del proceso resultó evidente la conveniencia de aplicar un método jerárquico, tanto por la adecuación de los datos, por la necesidad de controlar el proceso y porque se deseaba obtener una tipología coherente a varios niveles.

Como algoritmo de clasificación, tras las pruebas pertinentes, se eligió el modelo de **Ward**. Este método, define la distancia entre dos grupos como la distancia euclidiana al cuadrado respecto a la media de las observaciones. En cada paso los conglomerados resultantes son aquellos que presentan menor incremento en la suma global de distancias al cuadrado dentro de los posibles. Además es una técnica bastante sensible a la presencia de outliers, por lo que conviene eliminar los casos extremos del análisis y tratarlos como grupos de un solo individuo.

En este tipo de análisis hay que ser especialmente cuidadoso en la elección de las variables que van a caracterizar a cada individuo y que sirven de base para realizar agrupaciones

Una vez seleccionadas las variables existen tres problemas fundamentales para su tratamiento conjunto:

- ✓ Las variables están medidas en diferentes unidades y presentan variabilidad muy distinta.
- ✓ Normalmente están altamente correlacionadas, lo que supone redundancia en la información.
- ✓ Hay un número demasiado grande de variables (dimensiones) lo que dificulta el tratamiento y la interpretación del análisis.

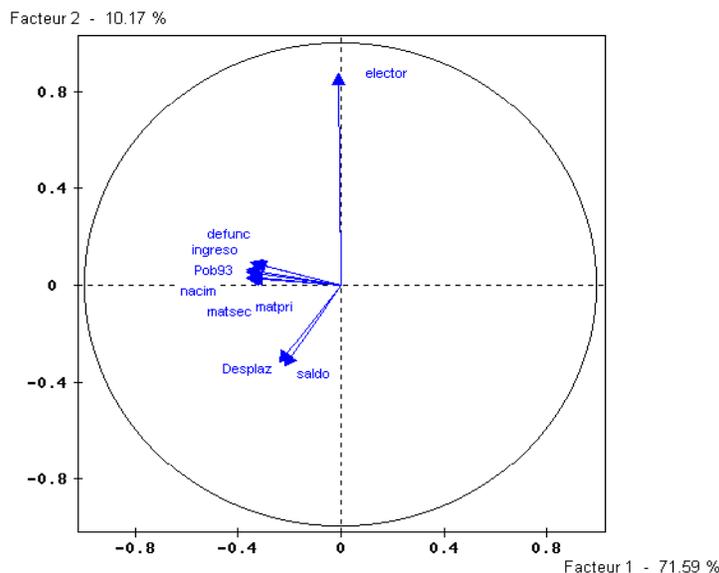
Para solucionar el primer problema, las variables de entrada han sido normalizadas. Para solucionar los otros dos problemas se ha acudido al análisis de factores principales, que permite, además, analizar con precisión el auténtico significado de cada indicador parcial y orienta sobre como plantear algunos indicadores sintéticos.

Antes de utilizar la medida de distancia, se debe encontrar el conjunto de variables que mejor representen el concepto de similaridad. (Concepto con el cual se busca el máximo de homogeneidad interna de los conglomerados). Para este caso el interés recae en tomar 1021 municipios de Colombia y agruparlos en cuatro cluster, (Excluyendo los 98 municipios que se tiene previsto que serán censados bajo una estrategia de operativo por rutas)

Las variables que mejor representan el concepto de similaridad, bajo el estudio fueron:

- ✓ Matricula primaria
- ✓ Matricula secundaria
- ✓ Nacimientos
- ✓ Defunciones

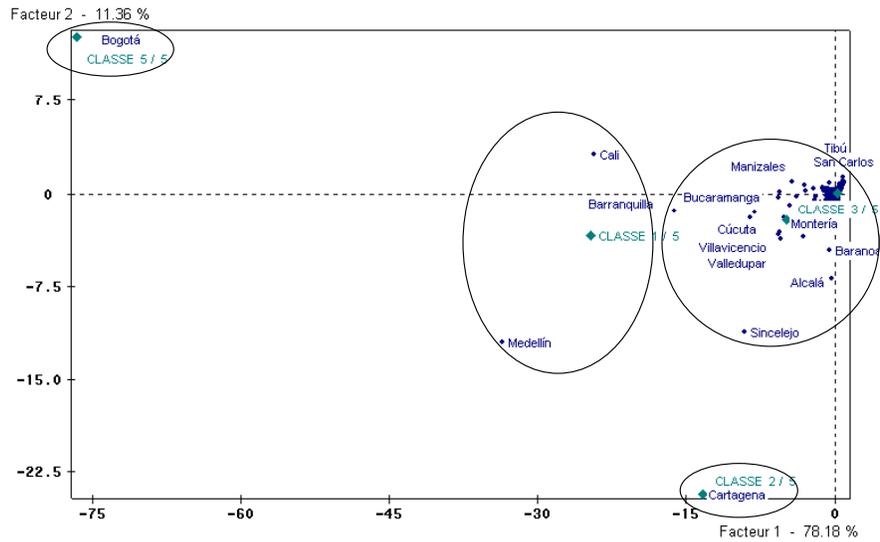
- ✓ Número de predios
- ✓ Número de electores inscritos
- ✓ Valor de ingresos tributarios municipales.



2.4.3 Resultados del modelo

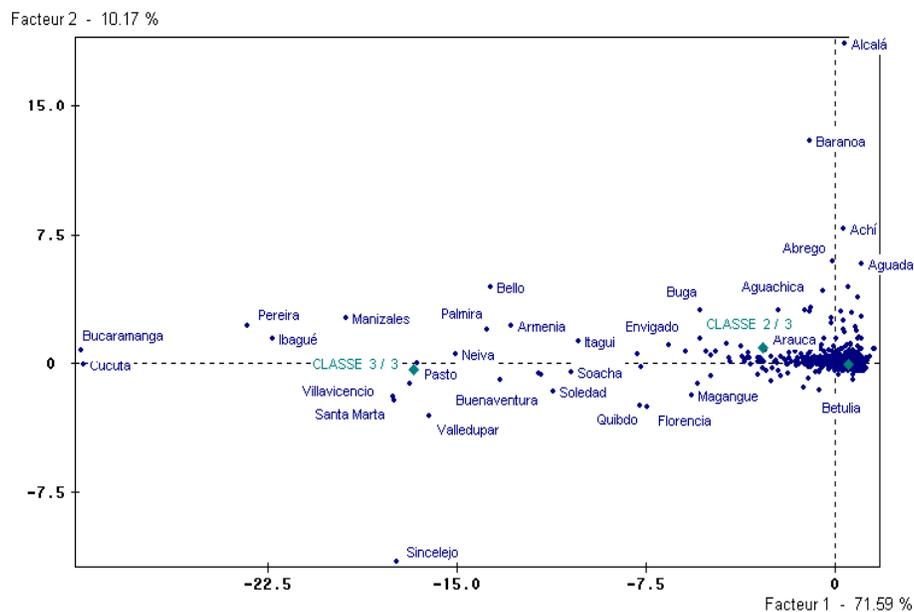
Para la conformación de los conglomerados que se espera definir para el levantamiento censal de los 1021 municipios se decidió utilizar el programa SPAD.

El primer resultado obtenido se muestra en el siguiente gráfico, donde claramente se pueden ver a partir de las variables matrícula primaria, matrícula secundaria, nacimientos, defunciones, número de predios, valor de ingresos tributarios municipales, número de electores inscritos, los 1021 municipios colombianos con información en todas sus variables pueden organizarse en 4 posibles conglomerados:



El método muestra que los municipios de Bogotá y Cartagena son entidades atípicas, dentro de dicho conjunto y esto en términos de uso para fines de operativos censales significaría que para el propósito del censo deberían, conformar cada uno de ellos un conglomerado. El tercero estaría integrado por los municipios de Cali, Barranquilla y Medellín; el cuarto por el resto de municipios.

Dejando de lado los municipios atípicos e incluyendo en el modelo la información sobre la variable desplazamiento interno de la población, captada en cada municipio, se obtuvieron los resultados que se indican a continuación:



	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 17 FECHA: 15-04-05
---	--	--

De acuerdo con éste último análisis se logró una clasificación del conjunto de municipios del país identificando las siguientes agrupaciones:

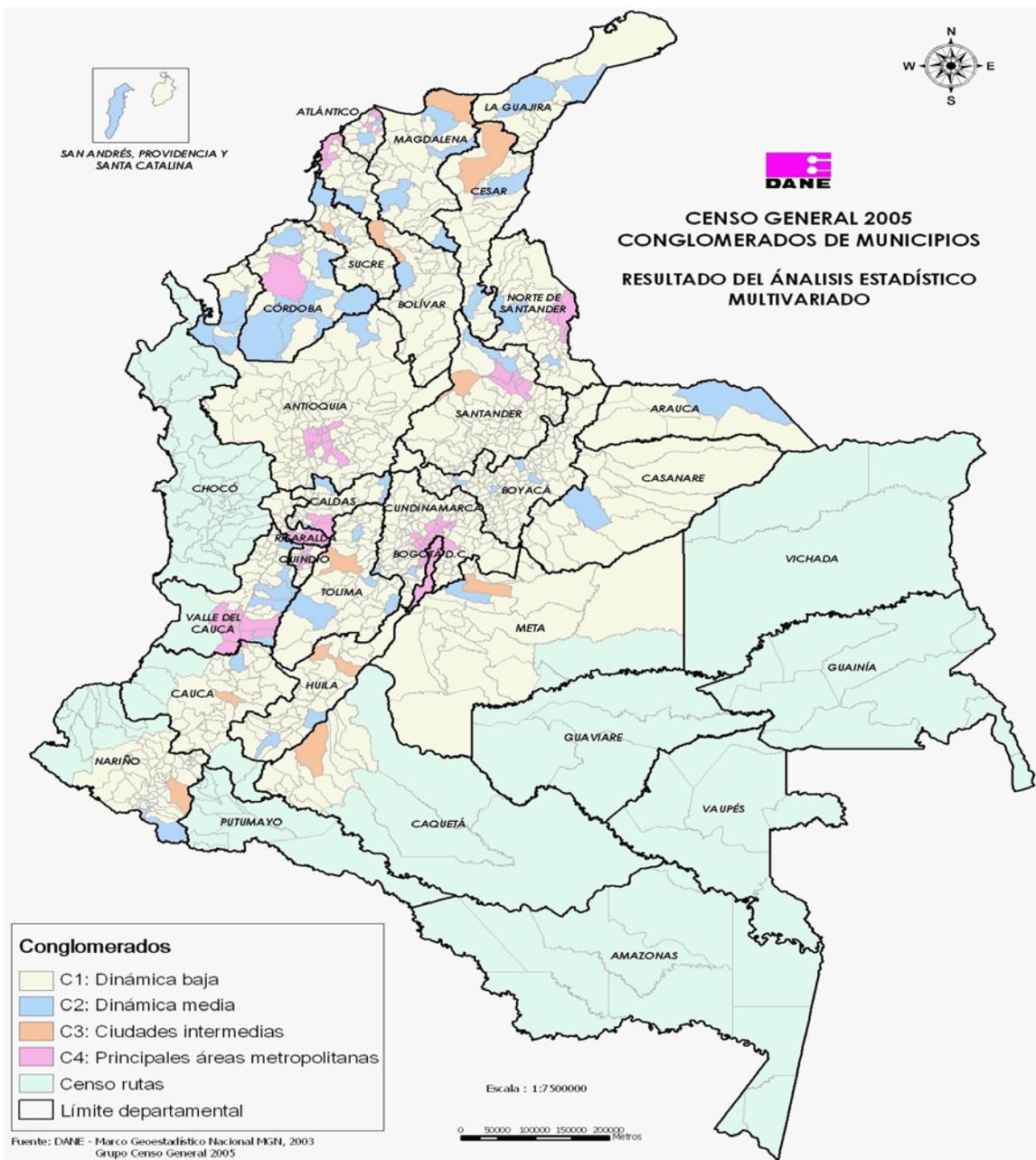
- ✓ **Grupo 1.** Finalmente, un grupo en el cual se ubican 883 municipios que expulsan y reciben población en pequeña escala. Son sin duda los municipios más pequeños del país, y por supuesto, el volumen de información de las variables sintomáticas fluctúan mucho y su calidad es poco confiable. Para este grupo de municipios no se posee una herramienta estadística que pueda medir la expulsión de población.
- ✓ **Grupo 2.** Este conjunto esta conformado por 55 municipios con comportamientos ambivalentes, esto es, que presentan condiciones de expulsión y recepción de población desplazada que en buena medida se compensa con un efecto final negativo de flujo de población. Son municipios mas pequeños que el anterior grupo, cuyos datos de nacimiento, defunciones y matriculas solo alcanzan un tercio de los del grupo anterior.
- ✓ **Grupo 3.** Un conjunto adicional de 36 municipios de tamaño intermedio dentro de los cuales están las capitales departamentales y municipios aledaños a estas. Este grupo se caracteriza por que los municipios son claramente receptores y poco expulsores lo que puede pensarse que se debe a procesos de retorno de migrantes. Presentan altos volúmenes de matriculas primaria y secundaria. Este grupo contiene al 18.9% de la población.
- ✓ **Grupo 4.** Conformado por un conjunto de municipios receptores de población los cuales están más afectados por la migración interna y fundamentalmente por efectos del impacto de la población desplazada. En esta categoría, están los municipios de Bogotá, Medellín Cali y Barranquilla, que presentan un comportamiento claro en cuanto a su condición de receptores de desplazados. Medellín y Cali se caracterizan por presentar elevados volúmenes de muertes.

Estos municipios son centroides⁸ de otros municipios aledaños a ellos, llamados ciudades dormitorio”, operacionalizado con base en criterios de distancia y tiempo de desplazamiento entre los municipios y la ciudad central; de tal forma que éste conglomerado esta conformado por 47 municipios de la siguiente manera:

Bogotá es centroide de 17 municipios,
Medellín es centroide de 13 municipios,
Barranquilla es centroide de 4 municipios y
Cali es centroide de 9 municipios.
Este grupo contiene el 36% de la población colombiana.

⁸ Un municipio se define como Centroide en el sentido que este actúa como municipio aglutinador de un conjunto de municipios que se ubican en su contorno espacial

Los distribución de los grupos de conglomerados resultante del modelo se pueden observar en el mapa siguiente, lo cual es consistente con los movimientos de migración interna, ligados a cambio de residencia por factores selectivos debidos a trabajo, estudio y ambientales. Igualmente es coherente, con los flujos de áreas afectadas por el conflicto armado que determina movimientos forzados hacia áreas urbanas mayores.



2.4.4 Conformación de los Grupos de Conglomerados

A partir de los resultados obtenidos mediante procesos de análisis multivariado con variables⁹ sintomáticas que evidencian una serie de centroides de “atracción” ubicados regionalmente, se obtiene un primer escenario de la dinámica población de los entes territoriales los cuales se mostraron en la sección anterior.

Este resultado estadístico define cuatro grupos de conglomerados conformados por todas las áreas metropolitanas¹⁰; otro grupo conformado por todas las ciudades intermedias la mayoría de las cuales son capitales de departamento y municipios que por su posición regional son centros importantes en la parte comercial y de servicios¹¹; un grupo de dinámica media conformado por municipios con características turísticas regionales y por último un grupo en los cuales se encuentra un gran número de municipios que desde el punto de vista demográfico se pueden considerar con una dinámica baja. En la tabla siguiente se pueden observar la distribución en cuanto a población y número de municipios.

Cantidad y porcentaje de población según grupos de conglomerados			
GRUPO	Población 2005	%	N° municipios
Grupo 4	21.823.783	47,4%	72
Grupo 1	14.093.164	30,6%	883
Grupo 2	4.659.773	10,1%	55
Grupo 3	3.473.828	7,5%	11
Censo rutas	1.994.563	4,3%	98
TOTAL	46.045.111	100,0%	1119

En la fase siguiente se realizó una serie de procedimientos en las cuales se involucra información sobre distancias a estos centroides de municipios del área de influencia, características regionales identificadas por las sedes y subsedes del DANE y aspectos de tipo operativo que tienen que ver con rendimientos y número de encuestadores por municipio¹². Con esta información, se realizó un balance entre grupos de conglomerados, lo cual definió la reubicación de municipios controlando que tuviese características demográficas homogéneas del grupo al cual es trasladado.

El desarrollo de éste proceso, determina la necesidad de abrir algunos de los conglomerados, encaminadas a facilitar el proceso operativo en especial, la utilización de los computadores de bolsillos o DMC manteniendo los mismos

⁹ Nacimientos, defunciones, asistencia escolar, población infantil con prevalencia en polio y triple viral, asistencia escolar por niveles, número de votantes, usuarios de energía eléctrica, beneficiarios SISBEN, número de predios, densidad de población, inmigración censo 1993, desplazados.

¹⁰ Incluye Bogotá D.C. Son ciudades de un alto desarrollo lo cual determina una dinámica alta.

¹¹ Dinámica media – alta.

¹² Urbano y rural.

principios de homogeneidad definidos en el modelo estadístico dando como resultado la siguiente clasificación:

- Un conglomerado en el cual está Bogotá D.C. y 24 municipios de su área de influencia que están altamente correlacionados con dicho centroide.
- El grupo 1 se abre en dos, con lo cual se obtiene un grupo de conglomerados con un total de municipio bajo, el cual por sus características de dinámica poblacional sería el primer grupo a realizar el operativo censal o dadas las condiciones operativas, esto grupos se pueden realizar de forma traslapada con los otros grupos de conglomerados.

En el cuadro siguiente se puede observar la nueva clasificación:

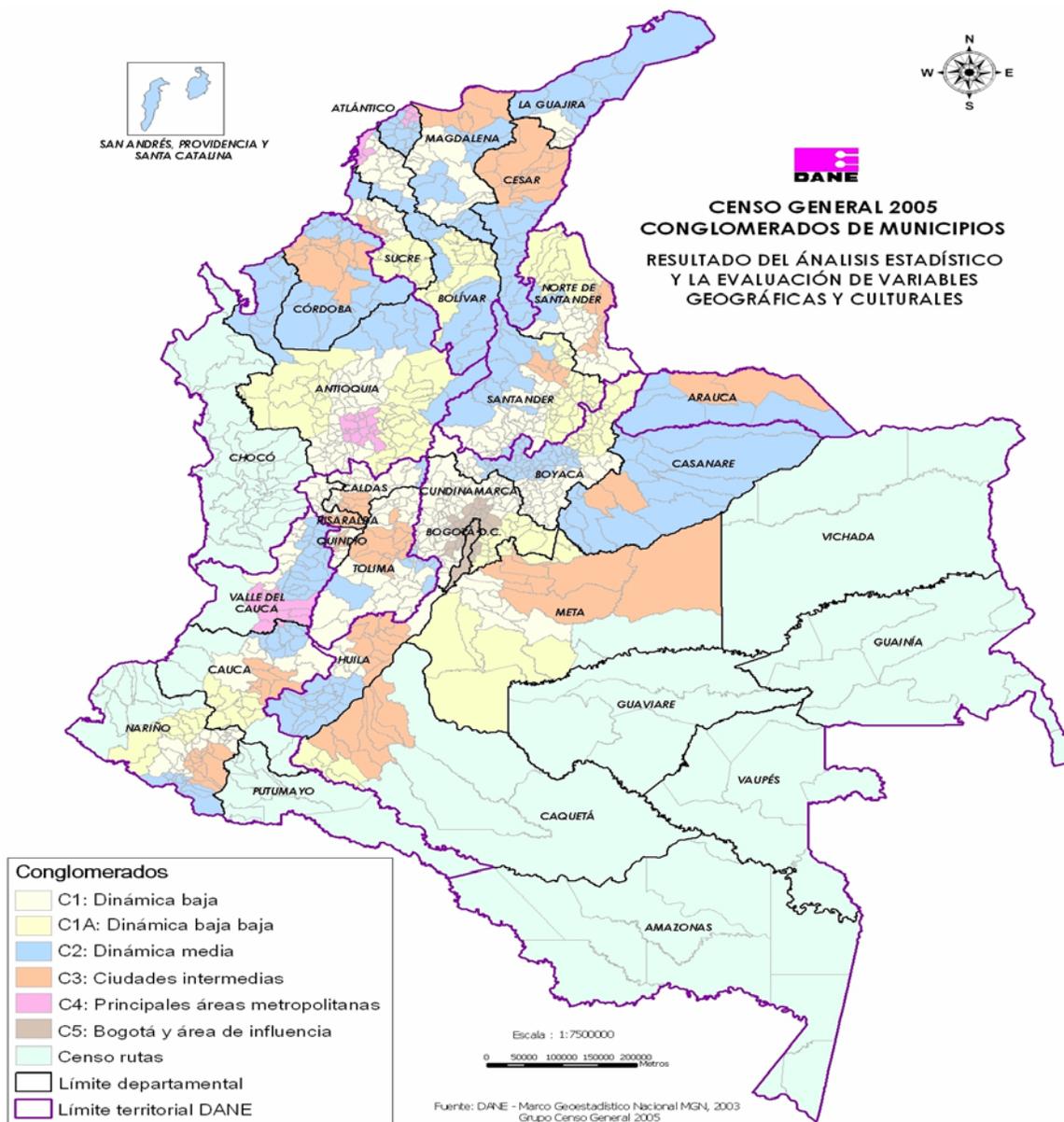
Cantidad y porcentaje de población según conglomerado					
GRUPO	Dinámica	Población 2005	%	Nº municipios	No. Encuestadores
C5: Dinámica Especial*	Especial	8.331.279	18,1%	25	6346
C4: Dinámica Alta**	Alta	9.843.669	21,4%	38	6883
C3: Dinámica Media - Alta***	Media-Alta	9.538.309	20,7%	116	7032
C2: Dinámica Media	Media	7.973.547	17,3%	259	7568
C1A: Dinámica Baja	Baja	5.868.561	12,7%	402	6912
C1: Dinámica Baja -Baja	Baja-Baja	2.495.183	5,4%	181	2957
Censo rutas		1.994.563	4,3%	98	2338
TOTAL		46.045.111	100,0%	1119	40036

* Bogotá y Area de Influencia

** Principales Areas Metropolitanas

*** Ciudades Intermedias

En el mapa siguiente se presentan los grupos de conglomerados, los cuales se ejecuta el censo desde el amarillo mas tenue hasta finalizar en Bogotá D.C. y su área de influencia.



De acuerdo con todo lo anterior, desde el punto de vista técnico y, considerando que el objetivo es minimizar los sesgos que se puedan introducir en la estructura y niveles de la información censal por factores de dinámica poblacional, la implementación operativa del censo de Población y Vivienda debe tener en cuenta lo siguiente:

- La recolección es simultánea en todos los conglomerados, ubicados en las diferentes regiones del país, que conforman cada uno de los diferentes grupos.

- Al interior de cada conglomerado, la recolección se realiza de forma simultánea en todos los municipios que lo conforman. El momento censal es la 00.00 horas del día del inicio
- La distribución espacial y ejecución del censo de los conglomerados se realiza desde el grupo de más baja dinámica hacia los de más alta. Sin embargo, por necesidades operativas u optimización del uso de las DMC, en estos conglomerados se puede realizar la recolección censal traslapándolos con los otros conglomerados.

Este proceso garantiza que el censo se realiza en el mismo sentido de los movimientos poblacionales, con lo cual en alguno de los centroides de atracción la persona o grupos de personas serán censadas. Un segundo problema que introduce sesgos en la información de población es la referida en los “contraflujos” lo cual puede determinar que un volumen de población no sea captada en el censo, dado que al realizarse en un municipio estas personas se muevan hacia municipios ya censados.

Como se puede observar en el mapa siguiente, los movimientos migratorios tienen características regionales muy particulares, los cuales se pueden clasificar como procesos encadenados que van de áreas “de menos oportunidades de desarrollo” hacia “centroides” de orden subregional, regional y de efecto nacional.

Lo anterior puede ser solucionado de dos maneras: La primera es mediante la realización de una campaña de sensibilización fuerte, en la cual se induzca a las personas no censadas a reportarlo en las oficinas del censo en cada municipio.. Una segunda forma, la cual desde el punto de vista técnico es la más adecuada, es realizar la recolección en un período menor a los doce meses en el cual se pueda traslapar la recolección entre conglomerados de baja y media dinámica, dejando los de mayor intensidad para el final del período.

Lo anterior significa que para minimizar el sesgo por movilidad y crecimiento natural la recolección de la información censal debe realizarse en el siguiente orden:

Primero el C1; luego el C1A; luego el C2; luego el C3; luego el C4 y por último el C5.

En cuanto a los tiempos para su ejecución, se deben tener en cuenta la existencia de meses o semanas muertas en los cuales la recolección en las cabeceras municipales no debe realizarse, porque implicaría altas probabilidades de la existencia de sesgos por factores de la dinámica poblacional, con efectos significativos sobre las estructuras por edad y sexo de los entes territoriales. A continuación se presenta la periodicidad de recolección de la información censal.

La estrategia de recolección para minimizar estos sesgos debe considerar los siguientes aspectos:

- ✓ Se requiere un número de días para facilitar el traslado de las máquinas a los municipios del siguiente grupo de conglomerados.
- ✓ La capacitación debe realizarse con la suficiente antelación para garantizar la adecuada formación de los encuestadores.

- ✓ Se deben realizar procesos para cada municipio con el fin de estimar con la información censal una serie de indicadores estratégicos que permitan evaluar la calidad de la información (además de la cobertura).
- ✓ Se debe generar una línea de base para cada municipio con los indicadores estratégicos sobre fecundidad, mortalidad, estructuras, movilidad, índices como relación niños mujer, masculinidad, Myers, entre otros para definir los ajustes necesarios en cada municipio y garantizar así la validez de la información.

Adicionalmente, se debe considerar que entre el 22 de mayo de 2005 y 22 de mayo de 2006, existen una serie de meses “muertos”, en los cuales sería riesgoso la realización de la recolección porque se incrementaría la parte de revisitas y, lo más importante se aumenta la probabilidad de que un gran número de personas no sean censadas, períodos que se puede observar en el gráfico siguiente:

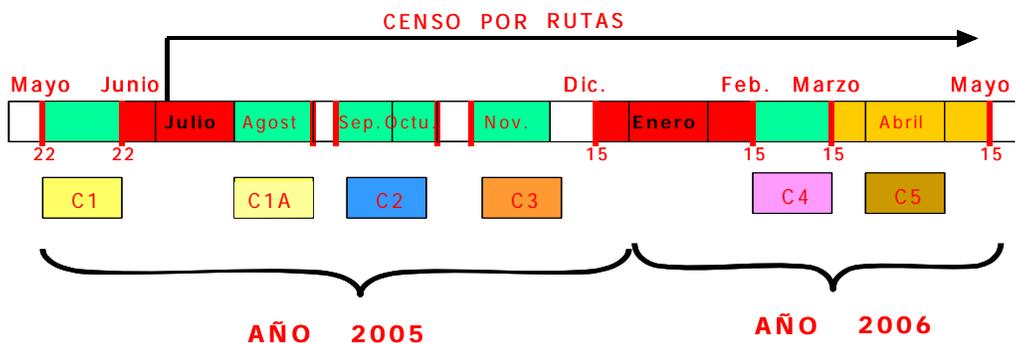


Los “colores” tienen los siguientes significados:

- Período apropiado para realizar la recolección
- Período NO apropiado para realizar la recolección por vacaciones
- Período de RIESGO para realizar la recolección por ÉPOCA de Elecciones
- Período para transición

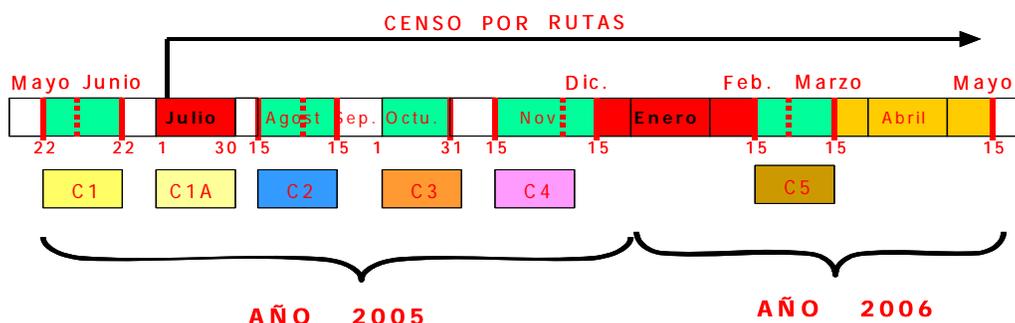
Bajo este orden de ideas , se pueden considerar diferentes estrategias de recolección de la información teniendo en cuenta la caracterización por conglomerados que va, desde utilizar todo el periodo con recolección consecutiva a, lapsos menores de tiempo y recolección con traslape de conglomerados.

Opción 1: Sería realizar en forma consecutiva la recolección iniciando en los conglomerados de más baja dinámica y terminando con Bogotá y su área de influencia como se observa en el siguiente gráfico.



Con éste esquema las principales áreas metropolitanas, así como Bogotá y su área de influencia se haría la recolección entre el 15 de febrero y el 15 de marzo de 2006, proceso que estaría en período de elecciones de parlamentarios y presidenciales, cuyos efectos sobre la credibilidad de la información resultante sería nefasta, además del hecho que va en contra de las recomendaciones de Organizaciones como las Naciones Unidas.

Opción 2: Realización de la recolección en forma consecutiva, pero dejando solo Bogotá y su área de influencia para realizar la recolección entre el 15 de febrero y el 15 de marzo de 2006 y, el resto de conglomerados, se haría en el 2005. Bajo esta estrategia implica utilizar parte de los meses de junio y julio para la ejecución del censo en los conglomerados de más baja dinámica, teniendo en cuenta que dicho período de vacaciones no tiene las mismas dimensiones de desplazamientos a sitios de vacaciones por parte de los miembros de un hogar como las de fin de año. En el gráfico siguiente se puede observar dicha estrategia.



Sin embargo la realización de la recolección del conglomerado 5 en el 2006, continúa con la problemática de las elecciones incrementado con el hecho de ser el centroide principal y de mayor incidencia en lo político, factores que pueden afectar la calidad de la información de dicho conglomerado.

Opción 3: Esta estrategia implica realizar la recolección en su totalidad en el año 2005 de acuerdo con el calendario de meses óptimos y el resto de tiempo, hasta mayo de 2006, para todo el proceso de consolidación, ajustes por cobertura, restitución de población vivienda y hogar, conformación de archivos censales finales y producción de tabulados. Para logra éste propósito es necesario que la recolección de algunos grupos de conglomerados se realicen de forma simultanea o se traslapen en algún momento, teniendo en cuenta el traslado de PDA, la sensibilización y la capacitación del personal operativo en cada municipio. Esta alternativa se puede observar en la siguiente gráfica.



La viabilidad técnica de la tercera opción está soportada por lo siguiente:

- ✓ Los ajustes a realizar en la restitución a un momento estadístico único son mínimos, considerando el lapso de 4 meses en la recolección. El ajuste mayor sería sobre los 18 municipios seleccionados para realizar la recolección en el mes de mayo.
- ✓ Al realizar de forma simultánea algunos de los grupos de conglomerados o con traslape, se está minimizando el sesgo por el efecto de los contraflujos.

Al realizar de forma traslapada los conglomerados de baja – baja y baja dinámica con los otros conglomerados, como lo muestra la última gráfica, implica una reordenación de dichos conglomerados para mantener el principio demográfico de que los “centroides” deben realizarse simultáneamente con su “área de influencia”, básico para garantizar la minimización de los sesgos por movilidad. Lo anterior significa, que el grupo de municipios de baja dinámica, grupo C1 y C1A, deben ser redistribuidos en dos grandes grupos bajo los siguientes criterios:

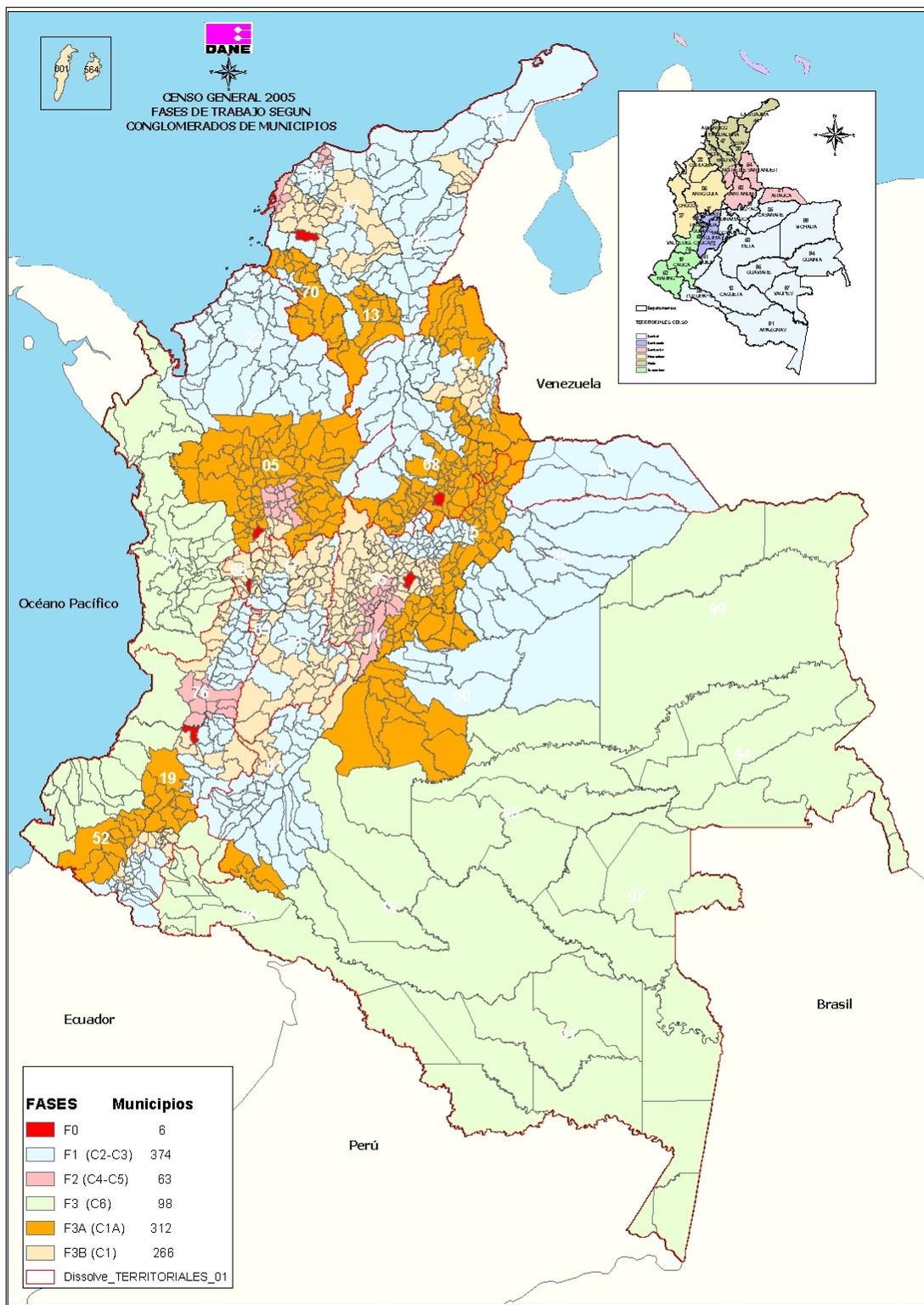
- ✓ El tamaño de los dos grupos deben estar balanceados en cuanto a número de entrevistadores, para garantizar la utilización de las PDA.
- ✓ Deben ser distribuidos de tal manera que concuerden con el área de influencia de los grupos de conglomerados C2 y C3, primer gran grupo a realizarse la recolección censal, y con el C4 y C5 el último a realizarse. Esta estrategia fortalece el diseño con lo cual se puede garantizar que los efectos en la información por los “contraflujos” tiene una probabilidad muy baja.

La recolección es a partir del 22 de mayo en un número reducido de municipios (uno por territorial DANE), del grupo de conglomerados de dinámica baja- baja, baja y media, lo cual permitirá realizar los ajustes finales a todos los procesos. Es importante tener en cuenta que si las correcciones o ajustes resultantes afectan la comparabilidad de la información con el resto de municipios, estos municipios deben tener la recolección en el momento censal que le corresponde a los respectivos conglomerados que pertenecen. La selección de estos 6 o los requeridos como municipios de inicio se realizan bajo los siguientes condicionantes:

- ✓ Deben tener la cartografía completamente adecuada para la recolección.

- ✓ El municipio no debe ser un centroide.
- ✓ Se deben seleccionar municipios que en la parte rural esté ubicado un resguardo indígena, con el fin de poder observar los procesos operativos en éste tipo de territorios.
- ✓ Deben ser municipios con características en la parte rural, en cuanto a tiempos de desplazamiento por tamaño de superficie, que permita probar el diseño operativo de regreso diario a la cabecera municipal y la parte de sincronización de información.
- ✓ Debe considerarse municipios en los cuales se aplique cuestionario a papel.
- ✓ Deben tener características disímiles para probar la capacitación.
- ✓ Distribución de municipios igual para cada una de las Territoriales DANE.
- ✓ No debe pertenecer a conglomerados de dinámica media, media – alta, alta y especial
- ✓ No estar cerca de capitales o ciudades intermedias o centroides.
- ✓ Debe constituir un grupo homogéneo en cuanto a características y proximidad.

En este orden de ideas, se distribuyen los municipios de dinámica “baja – baja” y “baja” de acuerdo a su ubicación respecto de los “centroides” en cada departamento e incluso se considera en varios casos varios departamentos, manteniendo el principio básico del modelo que la recolección se hace en el mismo “sentido de la movilidad espacial”. En el mapa siguiente se puede observar dicho esquema.



	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 28 FECHA: 15-04-05
---	--	--

Opción 4: Dado los inconvenientes presentados en el proceso licitatorio para la adquisición de las PDA, así como los parámetros definidos por la Dirección General del DANE y las restricciones presupuestales¹³, esta estrategia implica realizar la recolección para un grupo de conglomerados en el 2005 y, otro grupo en febrero de 2006.

Se presenta un caso especial con la Etnia Wayuu, cuyo territorio corresponde a la media y alta Guajira, grupo que se caracteriza por una alta movilidad dependiendo del comportamiento de las lluvias y de sus labores en el mantenimiento de su ganado, lo cual genera una alta movilidad entre Colombia y Venezuela, así como hacia el resto de municipios de la baja Guajira.

Considerando los aspectos anteriores y con el fin de minimizar los sesgos que se puedan introducir en la estructura y niveles de la información censal por factores de dinámica poblacional, la implementación operativa del censo de Población y Vivienda bajo esta estrategia debe tener en cuenta los aspectos básicos definidos en el modelo metodológico:

1. La recolección es simultánea en todos los conglomerados, ubicados en las diferentes regiones del país, que conforman cada uno de los diferentes grupos.
2. Al interior de cada conglomerado, la recolección se realiza de forma simultánea en todos los municipios que lo conforman. El momento censal es la 00.00 horas del día del inicio de la recolección en cada grupo.
3. La distribución espacial y ejecución del censo de los conglomerados se realiza desde el grupo de más baja dinámica hacia los de más alta.

La necesidad de realizar parte de la recolección en febrero de 2006 y considerando aspectos como elecciones y el período de festivos regionales, se afecta el punto tres lo cual genera las siguientes limitaciones:

- Por las características del primer trimestre de 2006, los conglomerados de mayor dinámica no se pueden realizar dada las implicaciones del proceso electoral.
- La alta influencia en gran parte de la Región Atlántica, en especial del departamento de Bolívar hacia el norte, del carnaval de Barranquilla, implica que toda esa subregión no se puede realizar en el 2006.
- Se debe tener en cuenta la dirección de los movimientos poblacionales hacia los principales centroides, con el fin de minimizar el efecto de los contraflujos.
- El grupo de municipios que se realizan en el 2006 deben ser en su mayoría de los conglomerados baja – baja y baja con características geográficas complejas o distantes de los principales centroides.
- Se debe tener en cuenta para esta selección las dinámicas propias existentes al interior de ciertas regiones del país como es el caso del sur del departamento de Bolívar, sur del Cauca, el eje Arauca – Casanare, occidente de Nariño así como el eje Uraba Antioqueño – Córdoba – Sucre.

¹³ Utilización mínima de papel; Máximo 14300 PDA

- La restitución al momento estadístico único del grupo de municipios realizados en el 2006 implica el desarrollo de procesos de conciliación con el fin de obtener los parámetros adecuados para realizar la retrepoyección, aspecto que requiere tiempo. Este punto es importante tenerlo en cuenta, considerando la fecha del 22 de mayo de 2006 definido como entrega oficial de los datos censales, dado que además de dichos procesos de conciliación se deben realizar todos los de depuración de la información censal.

Considerando todo lo anterior, y con el fin de garantizar los principios básicos del modelo soportado por los conglomerados, así como los ajustes para contrarrestar el posible efecto de no realizar la recolección desde los grupos de más baja a la de más alta dinámica, se definen los siguientes parámetros adicionales el modelo de recolección:

- a. Realizar la recolección de todo el departamento de la Guajira y los Resguardos de la Sierra nevada de Santa Marta entre el 1 y el 14 de septiembre de 2005. Con esto se garantiza los siguientes aspectos: evitar el efecto de las lluvias en la media y alta Guajira y realizar simultáneamente toda el área de influencia de los centroides regionales de Riohacha, Uribia y Maicao con lo cual se minimiza para esta subregión los sesgos por movilidad.
- b. Realizar la recolección de los conglomerados de media, media – alta y especial¹⁴ entre el 8 de octubre y el 15 de diciembre de 2005. Considerando estos grandes centroides, se realizan simultáneamente todos los municipios de sus áreas de influencia así como los que definen los corredores de movilidad poblacional.
- c. Realizar la recolección entre el 1 y el 28 de febrero de 2006 los municipios más alejados de los centroides principales, así como los de los ejes Arauca – Casanare y Uraba Antioqueño – Córdoba – Sucre. Estos últimos se incluyen considerando que los centroides de Uraba, Montería, Sincelejo y, Arauca, Yopal generan en cada subregión una dinámica muy propia por lo cual se debe realizar la recolección simultáneamente con sus respectivas áreas de influencia.

2.5 Diseño Encuesta Cocensal

La Encuesta Cocensal, se llevará a cabo en forma simultanea con el censo de población en una muestra probabilística de los hogares censados, con el objeto de ampliar el temario incluido en el censo.

Las características y tamaño de la muestra requerida dependen en forma importante del nivel de desagregación de los resultados, de la magnitud y variabilidad de los indicadores por ser estimados, y de la precisión deseada en las estimaciones. Para esta descripción, se parte de la necesidad de estimaciones a nivel municipal, total, cabecera y resto, y para grandes localidades o grupos de pequeñas.

¹⁴ Bogotá D.C. y 24 municipios cercanos.



**DISEÑO CONCEPTUAL Y
METODOLÓGICO
CENSO GENERAL 2005 – CGRAL**

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 30
FECHA: 15-04-05

La selección de la muestra es sistemática dentro de los hogares censados, con una fracción de muestreo aplicada rutinariamente en cada municipio (uno de cada dos, o de cada tres, etc.).

La fracción de muestreo varía de municipio a municipio, y está condicionada por el tamaño de la muestra requerida para estimaciones confiables y precisas a este nivel (municipal) y por el tamaño del municipio.

El tamaño de la muestra varía aproximadamente, entre 500 y 1800 hogares por municipio, dependiendo del tamaño de éstos, para estimaciones globales de fenómenos con una probabilidad de ocurrencia de 10% o más, un error estándar relativo del 7.5% o menos, y un nivel de confianza del 95%. Para las grandes ciudades también se están considerando estimaciones separadas para localidades o comunas de más de 150.000 habitantes, o grupos de los de tamaño inferior.

La relación entre tamaño de la muestra y tamaño del universo no es lineal, para un mismo error estándar. Ello quiere decir que el tamaño de la muestra crece en forma no proporcional con el aumento del tamaño del municipio, hasta cierto nivel, a partir del cual se vuelve constante.

Se prevé entonces, que la muestra de municipios muy pequeños, será una alta proporción de su población, y que la fracción de muestreo se reducirá progresivamente con el incremento del tamaño del municipio, es decir, disminuirá la proporción de la muestra. Las varias fracciones de muestreo entre municipios, o lo que es lo mismo, las diferentes probabilidades de selección, implican la necesidad de ponderar los resultados de cada hogar y persona por el recíproco de su probabilidad, para la combinación de resultados de dos o más municipios, con el fin de evitar el sesgo de estimación asociado con dicha variación de las probabilidades.

Ver documento "Metodología estadística de diseño y estimación de la encuesta cocensal".

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 31 FECHA: 15-04-05
---	--	--

III. RESTITUCIÓN A UN ÚNICO MOMENTO ESTADÍSTICO.

Como se describe en el punto anterior, cada grupo de conglomerados tendrá un *período censal* diferente y, por lo tanto, un *momento censal* al inicio de la recolección¹⁵. Esto significa que la información censal para cada uno de los municipios del país tienen diferentes tiempos de referencia, lo cual no permite la obtención de los acumulados departamentales, regionales y nacionales tanto de vivienda, hogares y población.

Lo anterior requiere llevar la información básica de población, vivienda y hogares a un único momento estadístico, para lo cual debe tenerse en cuenta las siguientes consideraciones:

- ✓ Debe ubicarse cercano al momento censal del grupo de conglomerados de más dinámica poblacional, lo cual significa que dichos municipios que lo conforman no tendrán ningún tipo de ajuste.
- ✓ Con el fin de evitar problemas con los datos de población por parte de los entes territoriales, teniendo en cuenta que estos conocen el dato observado, es mejor llevar hacia delante dichos valores y no retroproyectar.
- ✓ Se proyectan los totales de población de cada municipio y se mantienen las estructuras por edad y sexos constantes.
- ✓ La proyección de estos totales, se realiza considerando la ecuación compensadora, de la cual se obtiene un factor de crecimiento para el período dado por factores de movilidad espacial y crecimiento natural.
- ✓ Con las preguntas de control se pueden estimar un factor adicional por conglomerado, que se aplicaría a la estructura de la población, determinado por la información obtenida por las preguntas del módulo de control sobre el volumen de personas que se trasladan a municipios ya censados y, que no fue censado en el municipio de origen.
- ✓ Se realiza un ajuste a la estructura de población cuando se detecten personas duplicadas. Para esto se debe realizar por sistemas un “cotejo” en el cual se toma la información de las variables de control y se cruza con el nombre, edad, fecha de nacimientos, sexo, relación de parentesco, estado conyugal y nivel educativo. En el caso de darse éste hecho para cada municipio y, entre municipios, se debe tener en cuenta las siguientes situaciones:
 - a. Si la persona se mueve al interior de la cabecera o parte rural de un mismo municipio y se detecta en más de una ocasión, se contabiliza una sola vez y, para efectos del censo se deja en el último lugar en que fue censado.
 - b. Si la persona se mueve de cabecera a resto o resto a cabecera del mismo municipio se contabiliza en el último lugar donde fue censado y se borra del lugar inicial.
 - c. Si la persona se mueve entre municipios, se deja en el último y se borra del resto.

¹⁵ Ver definiciones del punto 5.1 de éste capítulo.

- ✓ Se calculan las tasas de jefatura por municipio y de acuerdo con su dinámica se pueden utilizar para la proyección de los hogares y viviendas. De acuerdo con la dinámica, se puede considerar que aquellos grupos de conglomerados clasificados como baja y baja – baja, dichos volúmenes en un corto plazo como es el tiempo del desarrollo del censo se mantienen constantes.
- ✓ Se utilizarán los indicadores demográficos que permitieron caracterizar cada conglomerado, así como otros adicionales, con los cuales se debe realizar el monitoreo sobre la calidad de la información en cada municipio. Los indicadores adicionales son: Índice de masculinidad, relación niños mujer, tasa bruta de natalidad, tasa de fecundidad general, tasa bruta de mortalidad, estructuras por edad y sexo. Igualmente se obtendrán volúmenes estimados de población a medida que avanza el censo en cada municipio, valor que se obtiene así:
 - a. Se toma el valor de población de los primeros dos días y bajo el supuesto que los rendimientos son los mismos se proyecta linealmente la población para los días saltantes para completar el período censal del conglomerado.
 - b. Al cuarto día se toma el acumulado y se proyecta para el período censal faltante del conglomerado. Así sucesivamente a medida que avanza el censo.
 - c. Cada estimación se compara con los indicadores contenidos en la Base Maestra y se analiza la consistencia. El primer nivel de análisis lo realiza el responsable del municipio y debe reportarse al DANE Central.

Teniendo en cuenta los puntos anteriores, así como la opción 3 definida con los conglomerados, el momento estadístico se ubicará en la fecha correspondiente al momento censal de los dos últimos conglomerados que inician el 20 de octubre de 2005 y que corresponden a los de más alta dinámica poblacional. Esto significa que el grupo de conglomerados C4 y C5 no requieren ningún tipo de ajuste.

A continuación se presenta la formulación matemática de la restitución de la población a un momento estadístico único.

El proceso de cambio poblacional y sus componentes de un municipio z ($z = 1, 2, 3, \dots, 1119$), perteneciente a un conglomerado j ($j = 1, 2, 3, 4, 5$ y 6), se resume en la denominada *Ecuación Compensadora* que se presenta a continuación

$$N_{(t+n)}^{(z,j)} = N_{(t)}^{(z,j)} + B_{(t,t+n)}^{(z,j)} - D_{(t,t+n)}^{(z,j)} + I_{(t,t+n)}^{(z,j)} - E_{(t,t+n)}^{(z,j)}$$

Donde :

$N_{(t+n)}^{(z,j)}$: es la población del municipio en el momento $t+n$

$N_{(t)}^{(z,j)}$: es la población en el momento censal del conglomerado j .

$B_{(t,t+n)}^{(z,j)}$: son los nacimientos del período t a $t+n$.

$D_{(t,t+n)}^{(z,j)}$: son las defunciones del período t a t+n.

$I_{(t,t+n)}^{(z,j)}$: son los inmigrantes a z en el período t a t+n.

$E_{(t,t+n)}^{(z,j)}$: son los emigrantes de z en el período t a t+n.

Si $\bar{N}_{(t+n)}^{(z,j)}$ es la población media del municipio en el período t a t+n se tiene:

$$\frac{N_{(t+n)}^{(z,j)} - N_{(t)}^{(z,j)}}{N_{(t+n)}^{(z,j)}} = \frac{B_{(t,t+n)}^{(z,j)}}{N_{(t,t+n)}^{(z,j)}} - \frac{D_{(t,t+n)}^{(z,j)}}{N_{(t,t+n)}^{(z,j)}} + \frac{I_{(t,t+n)}^{(z,j)} - E_{(t,t+n)}^{(z,j)}}{N_{(t,t+n)}^{(z,j)}}$$

Ecuación que es equivalente a:

$$r_{(t+n)}^{(z,j)} = TBN_{(t)}^{(z,j)} - TBM_{(t,t+n)}^{(z,j)} + TNM_{(t,t+n)}^{(z,j)}$$

Donde:

$r_{(t+n)}^{(z,j)}$ tasa de crecimiento del municipio z en el período t a t+n.

Como aportes relativos a este crecimiento están:

$TBN_{(t)}^{(z,j)}$ tasa bruta de natalidad del municipio z en el período t a t+n.

$TBM_{(t,t+n)}^{(z,j)}$ tasa bruta de mortalidad del municipio z en el período t a t+n.

$TNM_{(t,t+n)}^{(z,j)}$ tasa neta de migración del municipio z en el período t a t+n.

Con este crecimiento relativo estimado para cada municipio del conglomerado j, se puede proyectar entonces los totales poblacionales al momento estadístico único de la siguiente forma:

Si la proyección es hacia delante, el total en el momento t+n (t el momento censal del conglomerado j), es:

$$PT_{(t+n)}^{(z,j)} = PT_{(t)}^{(z,j)} * EXP\left(r_{(t,t+n)}^{(z,j)*\phi}\right)$$

Donde:

$PT_{(t+n)}^{(z,j)}$ es la población proyectada del municipio z del conglomerado j en el momento t+n

$PT_{(t)}^{(z,j)}$ es la población total obtenida del censo en el momento censal t.

ϕ es el tiempo, en días, entre el momento censal t y el momento estadístico t+n para el municipio z del conglomerado j.

Si la proyección es hacia atrás, el total en el momento t+n es:

$$PT_{(t+n)}^{(z,j)} = PT_{(t)}^{(z,j)} * \left\{ EXP\left(r_{(t,t+n)}^{(z,j)*\phi}\right) \right\}^{-1}$$

Para obtener la población por estructura según edad simple y sexo, como se supone que en el periodo definido por ϕ es relativamente corto, se tiene:

Si θ_m y θ_h es la estructura por sexo de mujeres y hombres respectivamente y, β_x es la distribución porcentual por edad x y sexo, entonces la población por edad y sexo del municipio z del conglomerado j es:

Para la estructura de mujeres por edad simple:

$$PT_{(t+n,x,m)}^{(z,j)} = PT_{(t+n)}^{(z,j)} * \left(\theta_m * \beta_{(x,m)} \right)$$

Para la estructura de hombres por edad simple:

$$PT_{(t+n,x,h)}^{(z,j)} = PT_{(t+n)}^{(z,j)} * \left(\theta_h * \beta_{(x,h)} \right)$$

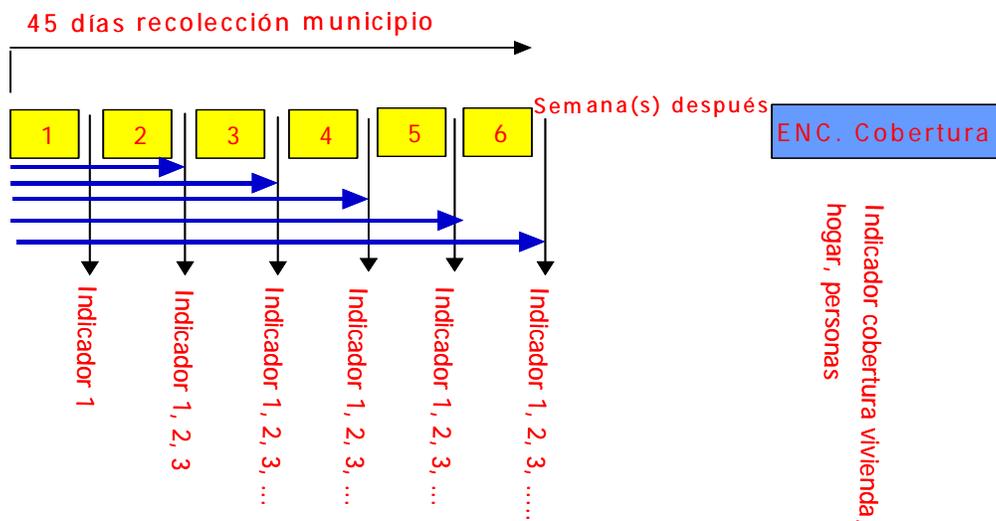
Con el fin de realizar el monitoreo para garantizar la consistencia y calidad de la información, se debe construir una Tabla de Indicadores que forma parte del Sistema de Monitoreo y Control, en la cual debe estar para cada municipio toda la información existente sobre población, nacimientos, defunciones, predios, usuarios de servicios públicos, así como los indicadores estratégicos mencionados en la formulación anterior.

Se debe tener en cuenta la población observada en el censo de 1993, así como las proyecciones para cada uno de los municipios¹⁶, con los cuales se construye intervalos de confianza con el fin de determinar los límites inferior y superior en el cual se puede ubicar la estimación de la población. Esta estimación se realiza en las cabeceras considerando que solo para dicha área se puede contar con la información censal de manera oportuna al ser captada a través de los DMC¹⁷. El proceso de estimación se realiza tomando la información por días o semanalmente, dependiendo del tamaño de la cabecera y con un factor de expansión se estima la población que se obtendría al final del período de recolección

¹⁶ Tanto para cabecera como resto

¹⁷ Dispositivo Móvil de Captura

Esta estimación, dependiendo de su ubicación dentro del intervalo, determina una serie de alarmas que pueden estar relacionadas con problemas de cobertura, manejo de los conceptos de residente habitual entre otros, con lo cual se hacen las respectivas correcciones en terreno. Esta información, se contrastan con algunos indicadores como el Índice de Masculinidad¹⁸ o la Relación Niños Mujer¹⁹ los cuales pueden explicar efectos de migración, sobre-mortalidad masculina, mortalidad infantil, omisión en niños(as) entre otros. Ver anexo donde están algunos de estos indicadores



Como se puede observar en la gráfica anterior, este proceso se realiza acumulando la información, con lo cual se pueden obtener otros indicadores estratégicos, de acuerdo a los conglomerados a que pertenecen, información que debe ser manejada no solo a nivel central sino además a nivel local y a nivel de las Direcciones Territoriales del DANE.

Estos indicadores son de gran utilidad en el caso de utilizar en las cabeceras un control de cobertura con verificación y corrección de errores, garantizando no solo un alto índice de cobertura de unidades de observación, sino además una adecuada consistencia y completez estadística de la información tanto del cuestionario básico como del ampliado.

Otros aspectos importantes a tener en cuenta es que con la información de “control de movilidad”, se pueden construir indicadores que permiten la cuantificación del efecto de la movilidad e identificar el número de casos de “personas duplicadas”, los cuales si los valores resultantes son significativos (más de 5%), deben ser considerados en la parte de restitución.

¹⁸ Este índice es un indicador de estructura de la población y muestra las diferencias en la composición por sexo y su fórmula es $IM = (\text{Total hombres}) / (\text{Total mujeres}) * 100$

¹⁹ Este indicador de fecundidad es de mucha utilidad cuando no existe información o poblaciones pequeñas. Expresa de manera simple las condiciones de fecundidad así como el efecto de la mortalidad infantil y juvenil. Su fórmula de cálculo es $RNM = (\text{Niños de } 0 - 4) / (\text{Mujeres en edad reproductiva})$

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 36 FECHA: 15-04-05
---	--	--

IV. CONTENIDO TEMÁTICO

El censo, por su carácter de universalidad poblacional y geográfica, así como por la diversidad e integralidad de las temáticas que se buscan cubrir, constituye la base fundamental e insustituible de conocimiento acerca de las condiciones demográficas, económicas, sociales y culturales de la población del país.

La temática definida para obtener información estratégica a través del censo general, está construida a partir de las siguientes premisas:

- Para la definición de los temas censales se tienen en cuenta las necesidades de información que tiene el país, la actualización de los datos existentes, su comparabilidad nacional (con censos anteriores) e internacional (con otros países del mundo).
- Idoneidad de las preguntas con lo cual se garantiza la calidad de la información.
- Para la recolección de la información de la temática censal se utilizarán cuestionarios inteligentes desarrollados bajo un software especializado para Computadores de Bolsillo o PDA. También se utilizarán, en casos especiales, cuestionarios diseñados a papel.
- Cada variable debe tener asociadas variables geográficas que permitan por un lado la georeferenciación de la información y la producción de tabulados o cruces de variables de acuerdo con las necesidades de los usuarios.
- Garantizar la información básica para el cálculo de NBI, migración, mortalidad y fecundidad, así como las que permiten la obtención de los volúmenes de vivienda, hogares y personas. Esta última según estructura por edad, sexo y grupos étnicos²⁰.
- Caracterización de las unidades de vivienda para cada unidad geográfica.
- Caracterización de la población para cada unidad geográfica.
- Construcción o actualización de marcos para la realización de encuestas sobre temas sociales, demográficos y económicos.
- Se utilizará un solo cuestionario considerando que el levantamiento del censo de población y vivienda será de forma simultánea y así garantizar la relación entre cada unidad de vivienda y los datos relativos a sus residentes habituales.
- Se define un contenido Básico que debe ser aplicado a todo el universo de cada municipio y, uno Ampliado para ser aplicado a una muestra representativa en cada municipio (cabecera y resto), para cada uno de los contenidos temáticos estratégicos como es el caso de las variables para el cálculo del NBI.
- Se utilizarán cuestionarios básicos para variables sobre unidades económicas y agropecuarias para la conformación de marcos.
- El contenido temático se constituye en el insumo básico para la estructuración del Sistema Estadístico Básico de Colombia.

²⁰ Indígenas según cada etnia o pueblo y población negra.

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 37 FECHA: 15-04-05
---	--	--

4.1 Censo de Población y Vivienda

El contenido temático de los cuestionarios censales, están definidos de acuerdo con los requerimientos institucionales y técnicos en cuanto a la necesidad de representatividad en pequeñas áreas de la información básica censal. En efecto, es importante considerar una serie de condicionantes que definen los contenidos del formulario básico, el cual será aplicado a todo el universo. Estos parámetros son los siguientes:

- La prevalencia de muchos fenómenos requieren muestras muy grandes con costos beneficios altos, que determinan su obligatoriedad de inclusión en el Básico. Es el caso de las variables que permiten medir los índices de NBI e ICV, así como población por edad y sexo, migración, desplazamiento, fecundidad y mortalidad infantil.
- Aspectos de tipo legal como es el caso de la información que debe certificar el DANE.
- Información que por sus características, se constituyen en información estratégica para el establecimiento de líneas de base en los procesos de planeación, seguimiento y evaluación de la gestión pública local.
- Dadas las frecuencias o totales que se requieren en subpoblaciones, caso de información por sexo y edad, para obtener información robusta se requiere su aplicación al total del universo en cada municipio. Este es el caso de la información de migración, desplazados, nivel y grado alcanzado en educación así como fecundidad y mortalidad infantil.

Bajo estos dos parámetros, las Temáticas para cada uno de los cuestionarios son las siguientes:

Cuestionario Básico

Se aplica a todo el universo, cabecera y resto, de cada uno de los municipios colombianos con un total de 28 preguntas. Su contenido es el siguiente:

- ***Identificación:*** Es un capítulo conformado por 8 ítems que permite relacionar la información censal en cuanto a vivienda, hogar y personas a nivel de cada ente territorial (cabecera, centro poblado, rural disperso), así como la parte de territorialidad para resguardos indígenas y territorios colectivos de comunidades negras.
- ***Vivienda y hogar:*** Contiene dos preguntas con las cuales se puede cuantificar el inventario de las viviendas según la topologías definida para el censo (Casa, apartamento, cuarto, vivienda de inquilinato y otro tipo de vivienda) y, a nivel de hogar la parte de emigración internacional.

- Listado de personas que conforman el hogar: Bajo el concepto de Residencia Habitual²¹ y ordenado según la relación de parentesco, permite realizar un control previo sobre el tamaño del hogar. Adicionalmente se incluyen dos controles con el fin de detectar personas que hayan sido excluidas y que son residentes habituales del hogar²² y; en el caso contrario minimizar el sesgo por incluir personas no miembros del hogar como es el caso de los estudiantes que pasan la mayor parte del tiempo en la ciudad donde realizan dicha actividad. Se incluye una pregunta sobre “defunciones en el último año” el cual es de suma importancia para la evaluación de los registros de defunciones.
- Datos de la Población Residente: se incluyen preguntas con el propósito de realizar controles para evitar omisiones o duplicidad. Las preguntas están adicionalmente controladas para realizarse de acuerdo con la edad y el sexo de la persona. La temática es la siguiente: sexo, edad, relación de parentesco con el jefe(a), pertenencia étnica²³, discapacidad, migración²⁴, desplazamiento forzoso en los últimos cinco años²⁵, educación²⁶, fecundidad y mortalidad²⁷, variables que por sus frecuencias a nivel de edad y sexo requieren ser aplicadas a toda la población.
- Preguntas de Control: Se incluyen una serie de preguntas que permiten identificar: si es residente habitual de el hogar censado o de otro; si ya fue censado en otro lugar del mismo municipio o en otro municipio. Igualmente se incluye una pregunta que permite controlar si existe asociado a cada hogar una actividad económica, la cual determina si es necesario la aplicación del cuestionario adicional sobre industria, comercio o servicios, así como en el caso de la parte rural si adicionalmente se presenta actividad agropecuaria.

Cuestionario Ampliado

Este instrumento se aplica a una muestra representativa para cada uno de los municipios²⁸ del país con un total de 80 preguntas. En el caso de los municipios muy pequeños, así como en resguardos indígenas y territorios colectivos de comunidades negras, éste cuestionario se aplicará a todo el universo. El contenido temático incluye lo del cuestionario básico más los siguientes temas:

- Vivienda: Es un capítulo con un total de 10 preguntas que permite la caracterización de la vivienda; estimar déficit cualitativo y cuantitativo, así como las variables requeridas, en cuanto a vivienda, para el cálculo de I NBI e ICV.

²¹ Es la persona que vive permanentemente o la mayor parte del tiempo en una vivienda aunque en el momento del censo se encuentre ausente.

²² Las personas tienden a omitir como residentes los niños menores, ancianos, personas internadas en clínicas, las secuestradas o las que están en ese momento en vacaciones.

²³ Para la población negra se mide desde dos aspectos: por autoreconocimiento fenotípico y cultural.

²⁴ Permite medir migración interna e internacional. La primera si es desde la cabecera o parte rural de un ente territorial.

²⁵ Identifica el año y la procedencia.

²⁶ Alfabetismo, asistencia y último año aprobado.

²⁷ Se aplica solo a las mujeres de 12 años o más.

²⁸ Cabecera, resto y algunas ciudades a nivel de localidad.

- Hogar: Es un capítulo con un total de 15 preguntas en los cuales se incluyen variables para la estimación de pobreza subjetiva y emigración internacional.
- Listado de personas que conforman el hogar: Contiene 6 preguntas. Adicionalmente se pregunta sobre aporte a los gastos del hogar, lo cual permite por primera vez realizar análisis de la información con perspectiva de género.
- Datos de la Población Residente: Este capítulo contiene 43 preguntas, algunas corresponden a temáticas nuevas como pobreza medida desde la perspectiva de consumo de alimentos, salud y seguridad social entre otros. Se complementa la parte de educación al indagar sobre causa de inasistencia escolar, idioma extranjero, uso del computador y hábitos de lectura. Igualmente en la parte económica se indaga sobre el sitio donde realiza el trabajo, dirección del trabajo, categoría ocupacional e ingreso. De acuerdo con una reunión técnica con funcionarios del Ministerio de Agricultura se incluye una pregunta a las mujeres de 12 años o más que viven en la parte rural sobre la actividad económica principal que desarrolla la mujer.
Preguntas de control: son idénticas a las del cuestionario básico.

Cuestionario Básico para los Lugares Especiales de Alojamiento - LEA

Este cuestionario es necesario para poder captar la información básica de los residentes habituales de cada ente territorial, que por características especiales de trabajo, estudio, religiosas o de reclusión, cumplen con el criterio de Jure del censo de población y Vivienda.

Lo anterior significa que la población por edad y sexo de cada uno de los diferentes municipios está conformado por todas las personas censadas tanto en los cuestionarios aplicados a los hogares particulares como en los LEA que se encuentren en su parte urbana y rural.

El contenido temático para la población residente en los LEA no solo debe captar las estructuras por edad y sexo en cada uno de los tipos de LEA²⁹, sino además las características básicas que se obtiene para la población residente en hogares particulares. Igualmente se incluyen algunas variables que permiten caracterizar las diferentes instituciones.

Nota: Por las características existentes en cada municipio, dado que muchos de estos LEA están ubicados en lugares específicos, el personal que aplica este cuestionario debe ser diferente al de hogares particulares. Una actividad que debe realizar el personal de línea es identificar con antelación los LEA que existan en cada municipio, e incluso aplicarlos antes del censo de población y vivienda.

En el documento anexo "**Contenido Temático del Censo General 2005**" se pueden observar las estructuras de las diferentes preguntas que contiene cada tema en cada uno de los cuestionarios.

²⁹ Cárceles o Centros de reclusión, albergues u orfanatos, asilos, conventos o seminarios, internados de estudio, cuarteles, campamentos de trabajo, habitantes de la calle, albergues de desplazados, prostíbulos entre otros.

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 40 FECHA: 15-04-05
---	--	--

4.2 Unidades Económicas y Agropecuarias

El contenido de estos dos cuestionarios se coordinó con los grupos encargados del tema llegándose a unos contenidos básicos que se pueden ver al detalle en el Anexo.

4.3 Entorno Urbanístico

este es un tema nuevo dentro de proceso censal con el cual se amplía de forma significativa el uso de la información, al poderse relacionar los datos de población y sus características habitacionales con el entorno urbanístico que rodea cada unidad de vivienda.

Sin embargo, es necesario tener cuidado con el proceso operativo de recolección para poder garantizar que la información que se obtiene es pertinente.

Desde el punto de vista metodológico, la unidad de observación es el “lado de manzana”, teniendo en cuenta que operativamente toda la información esta “amarrada” al AG. De acuerdo con esto, el cuestionario debe ser “único por lado de manzana” y aplicar tanto los criterios de hábitat como de características físicas de las vías, andenes etc.

Para lograr operacionalizar este proceso es necesario identificar si cada AG es de uso “residencial”³⁰ o “no residencial”. Sólo a las AG de uso residencial se les aplica los criterios de hábitat.

V. NORMAS PARA EL MANEJO DE LA INFORMACIÓN CENSAL DE POBLACIÓN Y VIVIENDA

Para la conformación de los archivos del censo general 2005, tanto en la captura de la información como en la consolidación final se deben considerar una serie de aspectos de tipo técnico con el fin de garantizar la consistencia de la información y por lo tanto su posterior uso por parte de los diferentes usuarios.

En el desarrollo de un proyecto censal se debe considerar una fase importante para la obtención de información con calidad para un adecuado uso por parte de los usuarios, en la cual se realiza la depuración de los datos censales.

Esta fase de depuración, en el modelo del Censo General tiene las siguientes características:

1. El diseño lógico y físico de las Bases de Datos que contienen los datos censales.
2. Las normas de consistencia que deben implementarse en los DMC para la recolección inteligente de los datos censales.

³⁰ Si en la AG existe como mínimo una vivienda.

3. Las normas de validación de la consistencia e imputación del dato. Este proceso es delicado y por lo tanto se debe realizar una vez se tenga todos los datos consolidados por municipio, dado que se requiere evaluar el efecto de la imputación sobre la estructura de los datos.

En la generación de las herramientas computacionales, así como en el diseño físico y lógico de los archivos censales se deben tener en cuenta las siguientes normas:

- **Las de tipo estructural:** Consiste en la verificación y corrección de las variables estructurales que vienen en el formulario censal, las cuales garantizan que la información de cada una de las unidades de observación pueda ser georreferenciada y alimentar la conformación de los diferentes tabulados o cruces de variables para el análisis de la información. Estas normas tienen que ver con:

VARIABLES DE IDENTIFICACIÓN que definen la pertenencia del cuestionario a una unidad geográfica y son: Departamento, Municipio (cabecera, centro poblado, resto rural, manzana) y que está contenida en la AG; número de la vivienda; número del hogar; número del cuestionario; número del formulario adicional³¹.

VARIABLES DE CONTENIDO constituidas por las variables de cada capítulo que son básicas por cada unidad de observación. En el caso del censo de Vivienda y Población son las siguientes: Condición de ocupación de la vivienda; número de la persona; sexo; edad; parentesco; nivel educativo y años aprobados; estado civil.

- **Las de Consistencia:** Son normas que dan las pautas para hacer la verificación de la información censal. Los datos consistentes son aquellos que están dentro del rango permitido para la variable y que además mantienen una relación lógica con otras variables dentro del mismo capítulo o con variables de otros capítulos.

A continuación se definen las normas generales en cada uno de los tipos enunciados anteriormente:

1. Estructurales para el Censo de Población y Vivienda.

En el documento anexo "**Normas para el Control de Estructura**" se detallan todos los procesos que se requieren tener en cuenta.

2. Consistencia Interna del Censo de Población y Vivienda

Las normas de consistencia que se deben tener en cuenta para realizar de forma automática en el momento de la recolección cuando se utiliza DMC, deben ser las más generales que no retarden demasiado la recolección.

³¹ Para el caso de los cuestionarios a papel. Para los formularios inteligentes, el sistema debe asignar un consecutivo

A continuación se presentan algunas normas de consistencia generales de los capítulos de vivienda, hogar y personas, considerando que los contenidos de los cuestionarios no están aún cerrados oficialmente.

- En las variables de municipio, departamento y país deben ser códigos válidos.
- El total de hogares existentes en la vivienda debe ser igual al conteo de los registros de hogar para dicha vivienda.
- Si la pregunta de Condición de ocupación de la vivienda es “Ocupada con todas las personas ausentes”, en la pregunta de Total de hogares en la vivienda debe ser 01.
- Si el resultado de la entrevista es vivienda “Desocupada”, en la pregunta de Total de hogares en la vivienda debe ser 00.
- Si en tipo de vivienda viene Apartamento o Tipo “cuarto” y el material de paredes es Zinc, tela, cartón, latas, desechos, etc., o sin paredes, el tipo de vivienda es Casa.
- Si en tipo de vivienda viene Apartamento y todos los hogares de la vivienda dicen que “No tienen servicio sanitario”, la vivienda es Tipo “Cuarto”.
- Si la vivienda no tiene servicio de energía pero alguno de los hogares de la vivienda cocina con energía eléctrica, o cuenta con: TV o con aire acondicionado, entonces la vivienda sí cuenta con energía eléctrica.
- Si existe un sólo hogar en la vivienda, el uso del servicio sanitario es exclusivo del hogar.
- Si en la vivienda existe más de un hogar y utilizan algún tipo de servicio sanitario, el uso del servicio sanitario es compartido con otros hogares.
- Si en el hogar no cocinan, la pregunta de dónde proviene el agua para preparar los alimentos debe estar en blanco.
- Si en el hogar cocinan y cuenta con servicio de acueducto entonces de dónde proviene el agua para preparar los alimentos debe ser acueducto.
- Las preguntas de edad, sexo y parentesco no deben venir en blanco.
- Toda mujer de 12 años o más debe tener información de fecundidad.
- Si en la edad de la persona viene 00 en la pregunta de migración en los últimos cinco años debe ser que la persona no había nacido y debe estar en blanco la pregunta sobre lugar donde vivía.
- Los códigos de departamento y municipio asociados a la AG son iguales al nombre del municipio y departamento de la opción 3 “en otro municipio”, el dato es erróneo y debe cambiarse por la opción 2 “en éste municipio”.
- Los niveles aprobados en educación deben ser consistentes con la edad (se debe construir tabla de control).
- Si en parentesco viene “Empleado(a) doméstico(a)”, en “la mayor parte del tiempo de la semana pasada” debe ser código 1 o 2; en la pregunta de

ocupación debe venir código relacionado con oficios en el hogar; en categoría ocupacional debe venir marcado el código 4 y en la pregunta sobre actividad económica debe escribir hogar particular con empleada doméstica.

- El total de hijos sobrevivientes, suma de total hombres y mujeres de la pregunta 75, no debe ser mayor al de total hijos tenidos vivos.
- Se deben construir tablas de homologación para los procesos de imputación de sexo, edad, nivel educativo y estado conyugal.

Las siguientes normas se definen de manera específica para cada uno de los temas:

CAPÍTULO II. DATOS DE LA VIVIENDA

- Este capítulo sólo debe estar diligenciado para el hogar 01 de cada vivienda.
- Este capítulo debe estar sin diligenciar en los formularios adicionales, cuando el hogar es el 01 de la vivienda y tenga más de 16 personas.
- Si en pregunta “material predominante de las paredes exteriores” hay marcación en código 6 (caña, esterilla, otro vegetal) ó código 7 (zinc, tela, cartón, latas, desechos, plástico), entonces, el tipo de vivienda debe ser código 1 (casa).
- Si en pregunta “tipo de vivienda” hay marcación en código 5 (otro tipo de vivienda), entonces, en pregunta “material predominante de las paredes exteriores” el código de respuesta debe ser 8 (zinc, tela, cartón, latas, desechos, plástico) ó código 8 (sin paredes) y en pregunta 8 (material de pisos) el código debe ser 6 (tierra, arena).
- Si en pregunta “material de paredes” el código de respuesta es 6 (caña, esterilla, otros vegetales), 7 (zinc, tela, cartón, latas) ú 8 (sin paredes), entonces, en pregunta 8 (material de pisos) el código de respuesta debe ser 3, 4 ó 5 (madera burda, tabla, tablón, otro vegetal, cemento, gravilla, tierra, arena).
- Si en pregunta “tipo de vivienda” el código de respuesta es 5 (otro tipo de vivienda), entonces, en la pregunta 10 (servicios públicos), el único servicio que puede tener marcación en código 1 es el servicio de energía eléctrica.
- Si en pregunta “tipo de vivienda” el código de respuesta es 2 (apartamento), entonces, en pregunta “servicio sanitario”, el código de respuesta debe ser 1, 2 ó 3 (inodoro conectado a alcantarillado, inodoro conectado a pozo séptico ó inodoro sin conexión, letrina, bajamar) y en pregunta “sitio de preparación de alimentos”, la marcación debe ser en código 1 (en un cuarto usado sólo para

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 44 FECHA: 15-04-05
---	--	--

cocinar), ó en código 3 (en una sala comedor con lavaplatos) ó en código 6 (En ninguna parte no preparan alimentos).

CAPÍTULO III. DATOS DEL HOGAR

- Este capítulo debe venir sin diligenciar en los formularios adicionales.
- El número de cuartos de la preguntas “cuartos de que dispone el hogar” debe ser menor o igual a los cuartos de la pregunta “cuartos que dispone la vivienda” en el capítulo I
- El número de cuartos de pregunta “cuartos usados Exclusivamente para dormir” , debe ser menor o igual al número de cuartos de pregunta “cuartos de que dispone el hogar”.
- En la pregunta “cuartos de que dispone el hogar” el número debe estar entre 1 y 20.
- En pregunta “cuartos usados exclusivamente para dormir” el número debe estar entre 1 y 20.
- Si en la pregunta “tipo de vivienda”, el código marcado es 3 (cuarto) ó 4 (otro tipo de vivienda), entonces, en la pregunta “cuartos de que dispone el hogar” el número de cuartos no puede ser mayor de 3.
- Si en pregunta “tipo de sanitario” el código de respuesta es 1 (inodoro conectado alcantarillado), entonces, en pregunta “servicios de la vivienda” debe tener marcación en código 1 para alcantarillado.
- Si en pregunta “tipo de vivienda”, el código de respuesta es 5 (otro tipo de vivienda), entonces, en pregunta “sitio de preparación de alimentos”, la marcación debe corresponder a códigos 2, 4, 5 o 6.
- Si en la pregunta “conexión a servicios públicos”, la vivienda no tiene energía eléctrica (código 2), entonces, en pregunta “combustible para cocinar” la respuesta debe ser diferente de código 1 (electricidad).
- Si en pregunta “agua para preparar alimentos”, la marcación corresponde a código 1 (acueducto), entonces en pregunta “servicios públicos”, acueducto debe tener marcación en código 1.

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 45 FECHA: 15-04-05
---	--	--

CAPITULO IV. LISTADO DE LAS PERSONAS QUE CONFORMAN EL HOGAR

F1. Población residente

- El total de personas en el hogar debe ser igual al número de personas entrevistadas.
- El número de orden registrado en pregunta “mayor aportante”, debe corresponder a uno de los números de orden registrados en pregunta “listado de las personas que conforman el hogar”. En algunos casos el número puede ser “00” ó “98”
- El número de orden de la persona registrada en pregunta “mayor aportante”, debe corresponder a una persona que en pregunta “edad” tenga 12 años o más.
- La edad válida para ser Jefe(a) de hogar es: Para hombres es 14 años o más; Mujer 12 años o más.
- El número de orden que le corresponde a cada persona en el listado de personas que conforman el hogar, debe ser el mismo que se registra en pregunta “orden “ para esa persona.
- Si existe la fecha de nacimiento, esta debe ser consistente con la edad declarada.
- Si en el hogar hay cónyuge, esposo(a), compañero(a) o pareja (pregunta “parentesco con el jefe del hogar = 2”), entonces, esa persona debe ser de sexo contrario al del jefe. Durante la toma de información debe aparecer una ventana de verificación, pues puede haber error en la asignación de parentesco o en el sexo. Si se corrobora que el sexo es el mismo del jefe, para parentesco 2, deje así.
- En todo hogar la persona con número de orden 01 en el listado de las personas que conforman el hogar, debe tener en pregunta “número de orden de la persona” código 01 y en pregunta “parentesco” código 1 (jefe o jefa de hogar).
- En el hogar siempre debe haber una persona en pregunta “parentesco” con código 01 (jefe) y éste debe ser único.
- En el hogar sólo debe haber una persona en pregunta “parentesco”, con código 2 (esposo(a), compañero(a), pareja y este debe corresponder al número de orden de persona 02.



DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 46
FECHA: 15-04-05

- Si en pregunta “sexo” la respuesta corresponde a código 1 (hombre) y la edad es menor o igual a 13 años, entonces, en pregunta “estado conyugal” el código de respuesta debe ser 5 (soltero).
- Si en pregunta “sexo” la respuesta corresponde a código 2 (mujer) y la edad es menor o igual a 11 años, entonces, en pregunta “estado civil” el código de respuesta debe ser 5 (soltera).
- De acuerdo con la edad de la persona, deben cumplirse las siguientes premisas:

EDAD (años)	SEXO	PARENTESCO	FILTROS (en blanco)	MIGRACIÓN	MÁXIMO AÑO APROBADO		
00	Hombre ó mujer	Hijo, nieto, hermano, otro pariente, otro no pariente,	F2, F3 y F4	P40=1	En blanco		
01				P40=1			
02				P40=1			
03						P40=1	01
04					F3 y F4	P40=1	02
05						P40<>1	04
06						P40<>1	05
07						P40<>1	06
08						P40<>1	07
09					P40<>1	08	
10				Hijo, nieto, hermano, otro pariente, empleada domestica, otro no pariente,	F4	P40<>1	09
11		P40<>1	10				
12 y más	Hombre	Cualquier parentesco	F4	P40<>1	Según tabla en pregunta de “ultimo nivel aprobado”		

Adicionalmente las siguientes relaciones:

- $(\text{Edad del jefe menos edad del hijo}) \geq 14$
- $(\text{Edad del jefe menos edad del nieto}) \geq 30$
- $(\text{Edad del padre menos edad del jefe}) \geq 14$

Migración:

- Los códigos deben ser válidos de acuerdo a la divipola o la de países.
- Si la edad es menor a 5 años entonces en la pregunta “donde vivía hace cinco años” debe ser la opción “no había nacido”
- Si el departamento y municipio son iguales al de la AG:



DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL

CÓDIGO: CM-CGRAL-TEM-01
VERSIÓN : 03
PÁGINA: 47
FECHA: 15-04-05

- Si es la pregunta de nacimiento, entonces el código es 1.
- Si es en la pregunta de residencia hace cinco años entonces el código es 2.

Desplazamiento:

- Para que el dato sea válido se debe cumplir: 2005 menos año de desplazamiento debe ser menor que la edad
- Códigos de municipio y departamento de acuerdo con la divipola.
- El año es válido si 2005 menos año desplazamiento es \leq a cinco.

Pertenencia étnica:

- Si la persona se autorreconoce como indígena se le debe preguntar “pueblo” y si “habla la lengua”
- Si la persona se autorreconoce como “ROM” debe tener dato de si “habla la lengua”.
- Cualquiera de las otras opciones no debe tener información en las preguntas de “pueblo” y “lengua”

Datos sobre pobreza:

- Si en pregunta “por falta de dinero no comió”, la respuesta es código 1, entonces el campo para “cuántos días” debe estar diligenciado y sólo aceptar números de 1 a 7
- Si en pregunta “alimentos consumidos” tiene marcación en “Sí”, necesariamente debe venir diligenciado el campo de número de días con valores de 1 a 7.

Educación:

- La edad de la persona, menos el total de años de estudio (sumatoria de años aprobados contados a partir de básica primaria hasta el último año de estudios aprobado), debe ser mayor o igual a 5.
- Si en pregunta “lugar de uso del computador”, hay marcación en código 1 “en ésta vivienda”, necesariamente en pregunta “tenencia de bienes” computador debe tener marcación mínimo “01”.
- Si la persona tiene entre 3 y 5 años entonces, la única respuesta posible pregunta “en la última semana utilizó el computador” es la alternativa “Entretenimiento”
- Si la persona en la pregunta “lugar donde utiliza el computador” es la opción “en el sitio donde estudia” entonces la persona debe estar asistiendo a un establecimiento educativo

	DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL	CÓDIGO: CM-CGRAL-TEM-01 VERSIÓN : 03 PÁGINA: 48 FECHA: 15-04-05
---	--	--

- Para realizar el proceso de validación de consistencia de la variable “Ultimo año aprobado”, se tiene que tener en cuenta la edad de las personas, para lo cual se consideran los siguientes años de estudio que como máximo puede tener una persona de acuerdo con su edad³²

Económicas:

- Si en parentesco viene “Empleado(a) doméstico(a)”, en “Durante la semana pasada” debe ser código 1 o 2; en la pregunta de ocupación debe venir código relacionado con “oficios en el hogar”; en categoría ocupacional debe venir marcado el código 4 y en la pregunta sobre actividad económica debe escribir “hogar particular con empleada doméstica”.

Fecundidad y mortalidad:

- El total de hijos sobrevivientes, suma de total hombres y mujeres, no debe ser mayor al de total hijos tenidos vivos.
- Para mujeres de 12 a 15 años de edad, la diferencia entre el año censal y el año declarado³³ de “nacimiento del último hijo”, más 11 debe ser **menor o igual** que la edad de la mujer.
- Para mujeres de 51 a 54 años de edad, el año de nacimiento del “último hijo” debe ser **menor o igual** que la diferencia entre el año censal y la edad de la mujer menos 50.
- El año declarado del nacimiento del “primer hijo nacido vivo” de una mujer de 12 años o más, debe ser consistente con su edad, así: la diferencia entre el “año censal” y el “año declarado”, más 11 **((2005 menos año) mas 11)**, debe ser **menor o igual** que la “edad de la mujer”;

Independiente que se realice la captura de los datos censales con DMC y, considerando que en algunos casos la recolección es con “formulario a papel”, el archivo censal final debe ser depurado, lo cual significa realizar el proceso de “Validación de Consistencia e Imputación”. En el documento “**Especificaciones para el Control de Validación de Consistencia e Imputación**” se indican para cada una de las preguntas de los Módulos de Vivienda, Hogar y Personas la estructura de la validación y/o la imputación.

³² Ver documento anexo “Normas para el control de Validación de Consistencia e Imputación”

³³ Si viene marcada la opción “6” (2000 o antes), asimile como año declarado “2000”.



DISEÑO CONCEPTUAL Y METODOLÓGICO CENSO GENERAL 2005 – CGRAL

CÓDIGO: CM-CGRAL-TEM-01
 VERSIÓN : 03
 PÁGINA: 49
 FECHA: 15-04-05

ANEXO: INDICADORES DE CONTROL DE CALIDAD Y EVALUACIÓN

Cod. ^a	Indicador	Tipo ^b	Unidad
Vivienda			
1	Total Viviendas	v	
1	Tasa de incremento de viviendas	t	
2	Total Viviendas según tipología	v	
1 - 2	Total viviendas según condición de ocupación	v	
1 - 2	Porcentaje de viviendas ocupadas	p	
1 - 2	Porcentaje de viviendas desocupadas	p	
Hogares			
1	Total hogares	v	
2	Total hogares por tipo de servicios	v	
2	Usuarios de servicios públicos según tipología	v	
1	Cobertura de servicios públicos	p	
2	Número de cuartos por hogar	v	
2	Cocina o cuarto para cocinar por hogar	v	
2	Hogares según fuente de agua para preparar alimentos	v	
2	Servicio sanitario por hogar	v	
2	Hogares con exclusividad de servicio sanitario	v	
2	Hogares según forma eliminación de la basura	v	
Personas			
1	Total personas	v	
2	Personas por hogar	p	
2	Personas por vivienda	p	
1 - 2	Densidad de población	t	
1 - 2	Crecimiento intercensal de la población	t	
1 - 2	Población en cabecera y en resto	p	
1	Personas por edad - 0 - 4 años	v	
1	Personas por edad - 5 - 25 años	v	
1	Personas por edad - 26 o más años	v	
1	Personas por sexo	v	
1 - 2	Tasa de natalidad	t	
1 - 2	Tasa general de Fecundidad	t	
2	Tasa global de fecundidad	t	
1 - 2	Relación niños - mujer	r	
2	Tasa bruta de mortalidad	t	
2	Tasa de mortalidad infantil	t	
1 - 2	Índice de masculinidad	t	
1 - 2	Asistencia escolar	v	
1 - 2	Alfabetismo	t	
1 - 2	Tasa de alfabetas	t	
2	Total personas sin educación según edad	v	
2	Total de personas según Nivel de educación	v	
2	Porcentaje de personas en educ. superior	p	
2	Total personas según tipo de establec. Educ. donde asiste	v	
2	Total personas según tipo de Discapacidad	v	
2	Proporción de personas con discapacidad	p	
1 - 2	Total personas por Etnia o Cultura indígena	v	
1 - 2	Total de personas nativas	v	
1 - 2	Total personas inmigrantes	v	
1 - 2	Tasa de inmigración	t	
1 - 2	Total personas extranjeras	v	
2	Proporción de personas extranjeras	p	
1 - 2	Total personas desplazadas	v	
2	Proporción de personas desplazadas	p	
2	Total de personas según Condición de Actividad	v	
2	Proporción de personas que estudian	p	
2	Proporción de personas que trabajan	p	
Económicas			
1 - 2	Porcentaje de Población económicamente activa PEA	p	
1 - 2	Porcentaje de Población en edad de trabajar PET	p	
2	Total de personas por Rama de Actividad	v	
2	Total de personas según ocupación	v	
2	Total de personas según Categoría ocupacional	v	