MEMORANDO

Para Información de :	*Producido por:*
INTEGRANTE DEL COMITÉ EJECUTIVO DE ENCUESTAS A HOGARES	COORDINADOR DEL COMITE DE ENCUESTAS A HOGARES DIRECTOR DE LA DIRECCION DE ENCUESTAS A HOGARES BUENOS AIRES,

ASUNTO:

Adjunto el informe "Nuevo Diseño Muestral en el Aglomerado Gran Buenos Aires".

Resalto la subdivisión de radios censales 91 a fin de convertirlos en unidades de primera etapa de muestreo. Sería eficiente utilizar esta subdivisión para la cartografía del Censo del 2000.

Lic. CLYDE CHARRE DE TRADUCHI
Dirección de Encuestas a Hogares
I.N.D.E.C.

Director del INDEC 12/4/99.
Dr. Hector Montero
Director de la Dirección de Estudios de Ingresos y Gastos de los Hogares
Lic. Carmen Dopico
Director de la Dirección Nacional de Recursos Humanos y Organización 9.4.
Lic. Ana Maria Edwin
Director de la Dirección Nacional de Metodología Estadística, Tecnología y Coordinación del Sistema Estadístico Nacional
Sr. Patricio Carranza
Director de la Dirección Nacional de Estadísticas sociales y de Población
Lic. Alejandro Giusti
Director de la Dirección Asistente de Estadísticas Sociodemográficas y Económicas
Lic. Samuel Goldberg
Director de la Dirección de Metodología Estadística
Lic. Norma Pizarro
Director de la Dirección de Difusión
Lic. Leonardo Piatti
Coordinador del Marco de Muestreo Nacional Urbano
Lic. Sara Novaro

LIC. CYNTHIA POK.

C. E. S.
Centro Estadístico de Servicios
DIRECCION DE DIFUSION
I.N.D.E.C.

23 MAY 2000

NUEVO DISEÑO MUESTRAL EN EL AGLOMERADO GRAN BUENOS AIRES

Descripción de la metodología utilizada en la elaboración del marco de muestreo

Características generales del diseño de la muestra

.- Ambito geográfico:

- Coincide con lo que definió el Censo Nacional de Población y Vivienda 1991 (Cen'91), es decir, Capital Federal y los Partidos del Conurbano (Partidos). A estos últimos se los puede clasificar en tres grupos:

- Primer grupo: son aquellos cuya superficie integra completamente el aglomerado censal y por ende de la E.P.H.:

  General Sarmiento
  Moron
  Lomas de Zamora
  Quilmes
  Lanús
  General San Martín
  Tres de Febrero
  Avellaneda
  San Isidro
  Vicente López

- Segundo Grupo: son los que su superficie integra parcialmente el aglomerado por tener definidos radios de tipo rurales y/o mixtos, o como en el caso de Tigre y San Isidro por no incluir el sector Islas:

  La Matanza
  Almirante Brown
  Merlo
  Moreno
  Esteban Echeverría
  Tigre
  Florencio Varela
  Berazategui
  San Fernando

- Tercer Grupo: los no comprendidos en lo que tradicionalmente se denomina Gran Buenos Aires (tampoco se los incluía en el anterior diseño muestral EPH), cuya superficie integra parcialmente el aglomerado:

  Escobar
  Pilar
  San Vicente
  General Rodriguez
  Marcoz Paz
  Cañuelas

Aunque el Censo'91 incluye en este último grupo el barrio Ruta Sol (Partido de La Plata), la EPH no lo toma.

Durante el transurso de la elaboración del marco muestral , algunos partidos han sufrido modificaciones, ya sea por división, cesión o incorporación de territorios entre ellos. Puesto que el punto de partida de este diseño fue el Censo'91, los nombres de los nuevos partidos se considerarán cuando se disponga de toda la información necesaria en forma definitiva y cuando esté completamente implementada.

- Tipo de muestreo:

Se trata de una muestra bietápica, estratificada,donde las unidades de primera etapa (UPE) son los radios censales o subdivisiones de los mismos y las unidades de segunda etapa (USE) son las viviendas.

Elaboración del marco muestral

Características generales:

El INDEC, en conjunto con las Direcciones Provinciales Estadísticas, y con fines exclusivamente operativos, divide a los departamentos en fracciones , que son extensiones geográficas con aproximadamente entre 5 y 7 mil viviendas . Estas fracciones a su vez se dividen en radios censales, compuestos de alrededor de 300 viviendas cada uno.

En el caso de la Capital Federal, se encuentra dividida en 21 distritos escolares (divisiones administrativas gubernamentales), los que a su vez son divididos por el INDEC en fracciones y radios censales.

Puesto que la selección de las UPE se realizó en forma independiente en Capital por un lado y en Partidos por otro; y que las características que presentan cada uno de los dos dominios son diferentes, se elaboraron dos marcos muestrales.

Marco Muestral de Partidos:

Se comenzó trabajando con la información del Cen'91 a nivel de radio censal, debido a que éstos constituyen un agrupamiento de viviendas del que se dispone de todos los datos relevados por el Cen'91 y de su correspondiente cartografía.

El Cen'91 clasificó a estos radios de acuerdo a sus características en urbanos, rurales y mixtos.

Para obtener un diseño muestral eficiente, se trabajó en la elaboración de un Marco Muestral cuyas principales características fueran:

- Tamaño de UPE (medido en cantidad de viviendas) lo mas parejo posible evitando una excesiva variación de esta característica.
- Extensión razonable en terreno, evitando grandes distancias dentro de cada UPE.
- Distribución uniforme de viviendas por manzana dentro de cada UPE, evitando sectores de muy baja densidad de viviendas.

Con el objeto de lograr estas características, se analizó la distribución de la cantidad de viviendas de los radios censales y se hicieron simulaciones de división de los mismos con el objeto de determinar la cantidad de viviendas a fijar como límite para dividir todo radio censal que sobrepase dicho valor.

Finalmente se decidió fijar los siguientes límites: 550 viv. para los radios de los tradicionales 19 partidos, y 600 para los del Resto.

Como consecuencia de esto, cualquier radio censal que sobrepasara estos valores, fue dividido en por lo menos dos **subgrupos**.

Para los radios que aun teniendo menos de 550 viviendas presentaran una extensión geográfica que hubiese dificultado las tareas de campo fue necesario una división en dos o más **subgrupos**.

En los radios que presentaban una densidad de viviendas despareja, esto es sectores con bastantes viviendas por manzana y sectores con muy poca cantidad de viviendas por manzana; se "recortó" el radio, es decir se formaron subgrupos. Se incluyeron en el marco del cual se seleccionarían las UPE a los del primer caso, mientras que al subgrupo con baja densidad se lo incluyó en un marco complementario llamado **Crecimiento Potencial** del cual se explicará mas adelante.

Cuando un radio censal contaba con menos de 150 viviendas se lo anexaba a otro radio lindante y con similares características demográficas.

Se realizaron operativos en campo con el objeto de verificar cartografía en zonas de borde, realizar conteos en zonas que se presumía crecimiento, verificar límites dudosos, etc.

Se contó entonces con diversas fuentes de información para la toma de decisiones:

- Archivos informáticos a nivel de radio y manzana generados a partir de archivos censales .
- Cartografía censal a nivel de radio y fracción.
- Areas EPH del diseño anterior. (en. base al Cen'80).
- Informes elaborados por personal que salió a realizar tareas en campo.

En total se dividieron y/o recortaron 595 radios censales dando lugar a nuevas unidades primarias y/o zonas de crecimiento potencial.

La siguiente tabla muestra la cantidad de radios censales por partido, clasificados según el Cen'91 en urbanos, mixtos y rurales y la cantidad de UPE definidas después de haber realizado las tareas de división y/o recorte.

Partido	Radios Urbanos	Radios Mixtos	Radios Rurales	Total de Radios	Eliminados por rurales	Total de radios a trabajar	UPE definidas	Crecimiento Potencial
Almirante Brown	299	2	2	303	2	301	316	12
Avellaneda	346	0	0	346	0	346	348	6
Berazategui	187	0	3	190	3	187	199	21
Cañuelas	4	0	0	4	0	4	2	2
Escobar	92	10	14	116	14	102	87	26
Esteban Echeverria	170	6	13	189	13	176	197	38
Florencio Varela	144	6	3	153	3	150	176	30
General Rodriguez	29	6	16	51	16	35	45	8
General San Martín	362	0	0	362	0	362	360	10
General Sarmiento	424	0	0	424	0	424	460	36
La Matanza	790	5	0	795	0	795	847	47
Lanús	431	0	0	431	0	431	441	2
Lomas de Zamora	487	0	0	487	0	487	488	12
Marcos Paz	19	6	7	32	7	25	26	3
Merlo	215	4	2	221	2	219	278	26
Moreno	148	3	2	153	2	151	212	34
Morón	513	0	0	513	0	513	532	10
Pilar	78	23	17	118	17	101	90	37
Quilmes	479	0	0	479	0	479	477	17
San Fernando	129	0	0	129	0	129	115	4
San Isidro	251	0	0	251	0	251	256	0
San Vicente	49	5	24	78	24	54	53	8
Tigre	170	0	17*	187	17*	170	182	36
Tres de Febrero	324	0	0	324	0	324	325	2
Vicente López	306	0	0	306	0	306	311	0
Total Partidos	6446	76	120	6642	120	6522	6823	427

(*) Sector Islas

A continuación se pueden observar las principales características de la distribución de la cantidad de viviendas (t_viv) por radio censal antes de llevar a cabo las tareas de división,recorte y anexiones. Se trabajó sin los radios rurales debido a que fueron excluídos en su totalidad, por lo tanto la distribución se analizó para 6522 radios.

Statistics

T_VIV

N	Valid	6522
	Missing	0
Mean		371,17
Median		358,00
Std. Deviation		127,40
Minimum		0
Maximum		1912
Percentiles	5	210,00
	10	246,00
	20	287,00
	25	301,75
	30	314,00
	40	336,00
	50	358,00
	60	381,00
	70	407,00
	75	422,00
	80	442,00
	90	503,00
	95	576,85

Se puede observar la gran dispersión de la cantidad de viviendas, a pesar de que el tamaño al que apuntaba el Cen'91 era de 300.

Se observa también por que se eligieron los cortes en 550 viviendas para la mayoría de los partidos. De haber decidido un límite de 300, se deberían haber dividido el 75% de los radios. Esto se explica observando la tabla anterior: el percentil 25 es aproximadamente 300, es decir que el 25% de los radios tienen 300 viv. o menos.

En el gráfico siguiente (llamado de tallos y hojas) se puede apreciar la forma de la distribución y la cantidad de valores que son considerados extremos.
Notar que en este tipo de gráficos, no solo se puede apreciar la forma de la distribución, sino que se pueden leer los valores de la misma.

```
T_VIV Stem-and-Leaf Plot

 Frequency     Stem &  Leaf

    93,00 Extremes    (=<122)
    20,00        1 .  23
    33,00        1 .  45
    44,00        1 .  6677
    70,00        1 .  888999
   133,00        2 .  00000111111
   180,00        2 .  22222233333333
   276,00        2 .  444444444455555555555
   291,00        2 .  6666666666677777777777
   439,00        2 .  88888888888888888999999999999999999
   543,00        3 .  000000000000000000001111111111111111111111
   580,00        3 .  22222222222222222222233333333333333333333333
   600,00        3 .  44444444444444444444445555555555555555555555555
   573,00        3 .  66666666666666666666667777777777777777777777777
   519,00        3 .  88888888888888888888999999999999999999
   452,00        4 .  00000000000000000001111111111111111
   344,00        4 .  22222222222222233333333333
   293,00        4 .  444444444455555555555
   214,00        4 .  6666666667777777
   151,00        4 .  88888899999
   118,00        5 .  000001111
    95,00        5 .  2222333
    75,00        5 .  445555
    67,00        5 .  66677
    46,00        5 .  889
     3,00        6 .  &
   270,00 Extremes    (>=602)

Stem width:   100
Each leaf:     13 case(s)
& denotes fractional leaves.
```

Se puede apreciar que la misma tiene una forma aproximadamente normal, con una media de 371 viv. por radio, pero que presenta 383 valores extremos (93 que están por debajo de 122 y 270 radios por encima de 602).
Es de destacar que si bien la media se encuentra cerca de las 300 viv. por radio, lo hace a expensas de demasiados valores extremos.
Su amplitud es de 1942, y el desvío estándar es de 127,4.

Una vez efectuadas las tareas de división, recorte y anexiones, en el conjunto de los 6522 radios, se obtuvo un marco muestral conformado por 6823 UPE[1].
Los siguientes cuadros y gráficos, muestran las características en cuanto a la cantidad de viviendas (VF que significa Viviendas Finales) de las UPE:

[1] Unidad de Primera Etapa de selección

Statistics

VF

N	Valid	6823
	Missing	0
Mean		350,20
Median		347,00
Mode		332
Std. Deviation		80,00
Variance		6399,42
Range		681
Minimum		135
Maximum		816
Sum		2389381
Percentiles	5	220,00
	10	249,00
	20	285,00
	25	297,00
	30	307,00
	40	327,00
	50	347,00
	60	368,00
	70	390,00
	75	403,00
	80	417,00
	90	457,00
	95	490,00

```
VF Stem-and-Leaf Plot

 Frequency     Stem &  Leaf

     1,00 Extremes    (=<135)
    48,00        1 .  455
    49,00        1 .  6677
    71,00        1 .  88999
   166,00        2 .  00000111111
   208,00        2 .  22222233333333
   314,00        2 .  44444444455555555555
   362,00        2 .  666666666667777777777777
   579,00        2 .  88888888888888888999999999999999999999
   680,00        3 .  00000000000000000000001111111111111111111111
   698,00        3 .  22222222222222222222222333333333333333333333
   671,00        3 .  44444444444444444444445555555555555555555555
   621,00        3 .  66666666666666666666677777777777777777777777
   557,00        3 .  88888888888888888888999999999999999999
   482,00        4 .  00000000000000000011111111111111
   371,00        4 .  22222222222222333333333333
   292,00        4 .  44444444445555555555
   220,00        4 .  666666667777777
   158,00        4 .  88888899999
   125,00        5 .  000001111
    95,00        5 .  2222333
    34,00        5 .  445
     2,00        5 .  &
    19,00 Extremes    (>=562)

Stem width:    100
Each leaf:      15 case(s)


& denotes fractional leaves.
```

Se puede observar que la distribución sigue siendo normal, pero que la media es más baja que la anterior (ahora es 350) y se logró reducir la cantidad de valores extremos: ahora hay 1 con menos de 135 y 19 con más de 562.

A fin de comparar la variabilidad de ambas distribuciones, se muestran los coeficientes de variación de ambas: 34,3% para la primera y 22,8% para la última.

Marco Muestral de Capital Federal

Estos son los principales estadísticos descriptivos de la distribución de los tamaños de los radios censales (en su totalidad urbanos), correspondientes a los 21 distritos de la Capital Federal:

Statistics

T_VIV

N	Valid	3410
	Missing	0
Mean		350,12
Median		322,00
Mode		268[a]
Std. Deviation		135,07
Variance		18244,73
Range		1270
Minimum		0
Maximum		1270
Sum		1193900
Percentiles	5	200,00
	10	222,00
	20	249,00
	25	263,00
	30	277,00
	40	300,00
	50	322,00
	60	350,00
	70	382,00
	75	400,00
	80	428,00
	90	511,90
	95	613,45

a. Multiple modes exist. The smallest value is shown

```
T_VIV Stem-and-Leaf Plot

 Frequency     Stem &  Leaf

    16,00 Extremes    (=<42)
    11,00        0 .  &
    27,00        1 .  4&
   114,00        1 .  56788999
   515,00        2 .  00001111112222222333333344444444444
   674,00        2 .  55555555566666666777777777888888889999999999
   672,00        3 .  00000000001111111111222222222233333334444444
   525,00        3 .  55555555666666677777778888889999999
   283,00        4 .  0000011122223333444
   198,00        4 .  555666778899
   116,00        5 .  001122334
    74,00        5 .  56789
     4,00        6 .  &
   181,00 Extremes    (>=606)

Stem width:    100
Each leaf:      15 case(s)


& denotes fractional leaves.
```

Se puede observar que la distribución presenta 54 radios con menos de 150 viviendas. Cada uno de los mismos fue verificado en campo a los efectos de determinar las características de los mismos, y efectuar conteo de viviendas para decidir su inclusión en el marco muestral o en el de crecimiento Potencial.

Se dividieron los radios con más de 600 viviendas , excepto aquellos conformados por una sola manzana, que de hecho eran la mayoría.

A continuación se muestran las características del marco muestral de la Capital una vez efectuado las divisiones y tomadas las decisiones con los menores de150:

Statistics

VF

N	Valid	3401
	Missing	0
Mean		353,99
Median		324,00
Mode		278
Std. Deviation		132,25
Variance		17489,02
Range		1151
Minimum		119
Maximum		1270
Sum		1203913
Percentiles	5	207,00
	10	225,00
	20	253,00
	25	266,00
	30	278,60
	40	302,00
	50	324,00
	60	352,00
	70	383,00
	75	402,00
	80	430,00
	90	513,80
	95	618,80

```
VF Stem-and-Leaf Plot

 Frequency     Stem &  Leaf

     1,00        1 .  &
      ,00        1 .
    20,00        1 .  455
    32,00        1 .  6677
    75,00        1 .  8888999999
   153,00        2 .  000000000111111111111
   211,00        2 .  22222222222222233333333333333
   279,00        2 .  44444444444444444444555555555555555555
   267,00        2 .  66666666666666666677777777777777777777
   283,00        2 .  88888888888888888889999999999999999999
   307,00        3 .  00000000000000000000011111111111111111111
   252,00        3 .  22222222222222222223333333333333333333
   256,00        3 .  44444444444444444455555555555555555555
   218,00        3 .  66666666666666667777777777777777
   181,00        3 .  88888888888999999999999999
   123,00        4 .  000000000001111111
   114,00        4 .  2222222233333333
    87,00        4 .  444444555555
    89,00        4 .  6666666677777
    73,00        4 .  8888899999
    52,00        5 .  0000111
    46,00        5 .  222333
    32,00        5 .  44455
    27,00        5 .  6777
    32,00        5 .  88899
     5,00        6 .  0
   186,00 Extremes    (>=606)

Stem width:   100
Each leaf:       7 case(s)
```

Se observa que en el caso de Capital, no se produjo ningún cambio importante en la distribución debido a que no se hicieron gran cantidad de divisiones.
Las medias y los coeficientes de variación (38.6% y 37.3% respectivamente) son practicamente iguales.

Resumiendo:

De todas estas subdivisiones de radios se cuenta con la cartografía correspondiente y el detalle de las manzanas involucradas. En el único caso donde no se dispone de esta información (salvo para las áreas seleccionadas por la EPH) es en los partidos del Resto.

Para la EPH, sería importante que se tenga en cuenta para el próximo Censo Nacional de Población y Viviendas esta división, en el caso de que se formen nuevos radios a partir de los ya existentes.

Esto permitiría estudiar las carácterísticas del marco muestral de GBA a la luz del Censo'2000.

Crecimiento potencial

Todos aquellos radios o subgrupos de los mismos que tienen una baja densidad de viviendas particulares habitadas forman parte del marco muestral, pero incluídos dentro de un conjunto denominado "crecimiento potencial".

Estos radios o subgrupos pertenecientes a "crecimiento potencial" tienen asignado un grupo de rotación, a fin de que sean revisados en ocasión de la actualización de listado del grupo de rotación al cual pertenecen a fin de detectar variaciones con respecto a la cantidad de viviendas que tenían al momento de implementar el diseño.

El objetivo de este control a través de este marco complementario es poder incorporar, si fuese necesario, nuevas zonas al marco muestral original con su correlato muestral, y así mantener actualizadas las estimaciones provenientes de la Encuesta.

A continuación se muestra la cantidad de zonas de crecimiento potencial en cada uno de los Partidos y en Capital:

Partido	Crecimiento Potencial
Almirante Brown	12
Avellaneda	6
Berazategui	21
Cañuelas	2
Escobar	26
Esteban Echeverria	38
Florencio Varela	30
General Rodriguez	8
General San Martín	10
General Sarmiento	36
La Matanza	47
Lanús	2
Lomas de Zamora	12
Marcos Paz	3
Merlo	26
Moreno	34
Morón	10
Pilar	37
Quilmes	17
San Fernando	4
San Isidro	0
San Vicente	8
Tigre	36
Tres de Febrero	2
Vicente López	0
Total Partidos	427
Capital Federal	13
TOTAL GBA	440

En el caso de los Partidos, también se han incluído en este marco de Crecimiento Potencial, aquellas zonas en las que el porcentaje de viviendas desocupadas superase el 50% según el Censo'91 o debido a que se disponía de alguna otra infomación adicional que indicara una zona de viviendas de fin de semana o de Country Clubs donde las familias no residen en forma permanente.

En Capital Federal se han incluído en este marco complementario a zonas tales como Puerto Madero y la zona del Autódromo.