

Contenido

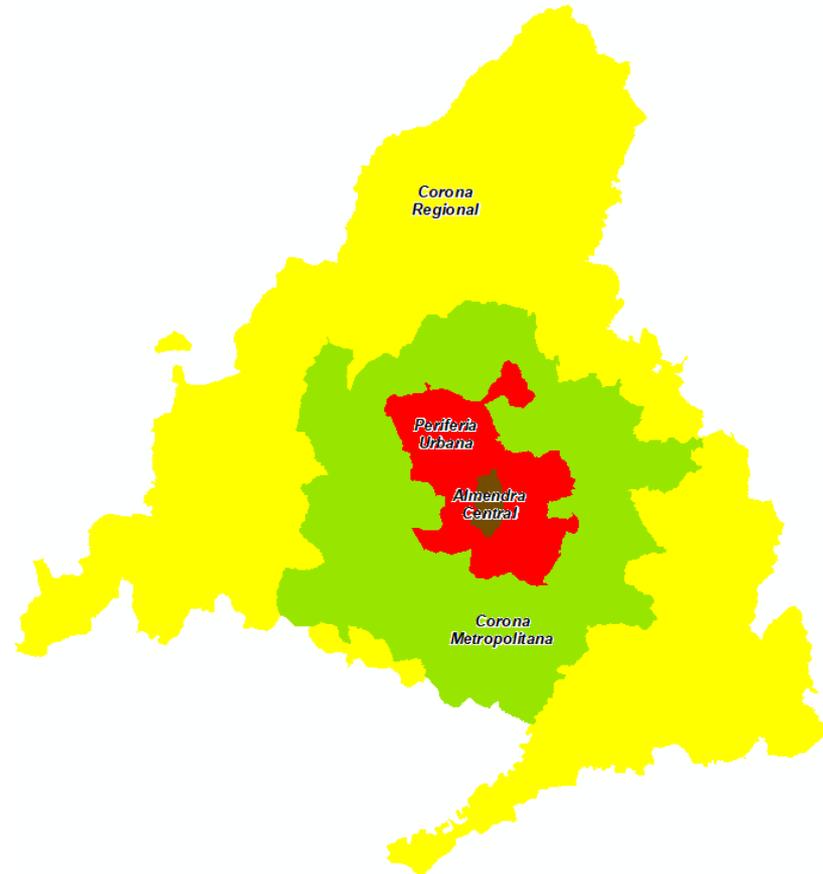
1. Planteamiento para el diseño
2. Diseño
3. Selección muestra
4. Unión de los dos campos, generación de los factores de expansión (hogares, personas, viajes)
5. Contraste resultados con datos sociodemográficos
6. Nivel de error para la CM de la variable viajes
7. Conclusiones

Encuesta de Movilidad 2018 ¿Cómo te mueves?



CAPI 13.000 Hogares

CATI 50.000 Individuos



Universo: hogares/familias residentes empadronadas en Madrid.

Marco de referencia: el Padrón de la Comunidad de Madrid.

Criterios de muestreo:

- Población a encuestar distribuida de forma proporcional a la población residente en las coronas A, B y C.
- Dentro de las coronas A y B, proporcional a la población de cada distrito en Madrid y a los municipios en la corona B, lo que también es aplicable a Aranjuez en la corona C.
- Para determinar el tamaño de la muestra en la Corona C se tuvieron en cuenta aspectos relativos a la localización, accesibilidad, oferta, etc..



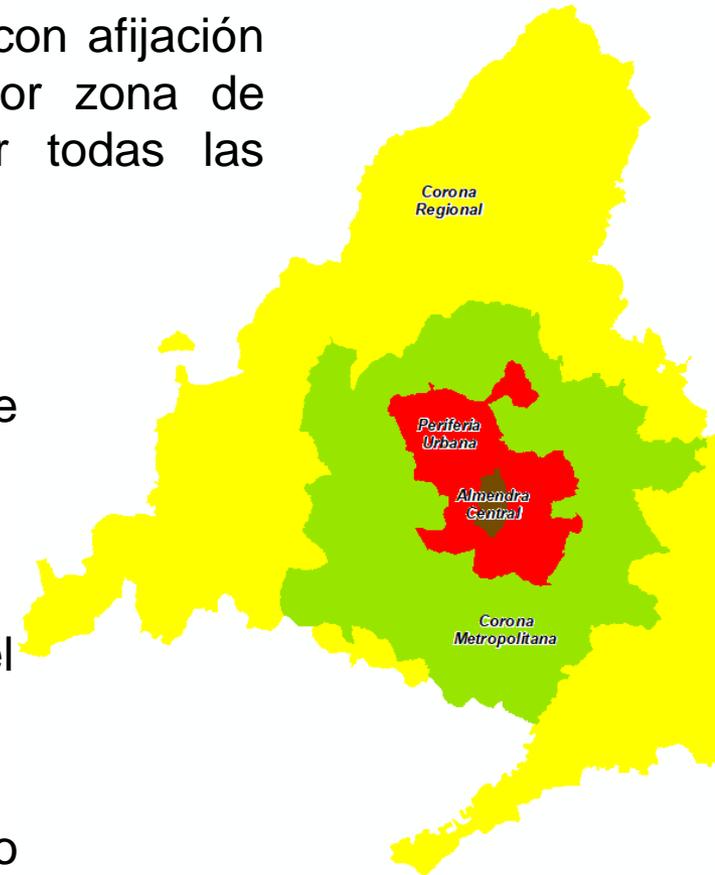
Por lo tanto y bajo estos aspectos, para la selección muestral se propuso un muestreo sistemático con afijación proporcional y con representación mínima por zona de transporte, con el objeto de poder acercar todas las casuísticas de las mismas.

Selección de la muestra

Estratificación por Coronas y zonas de transporte.

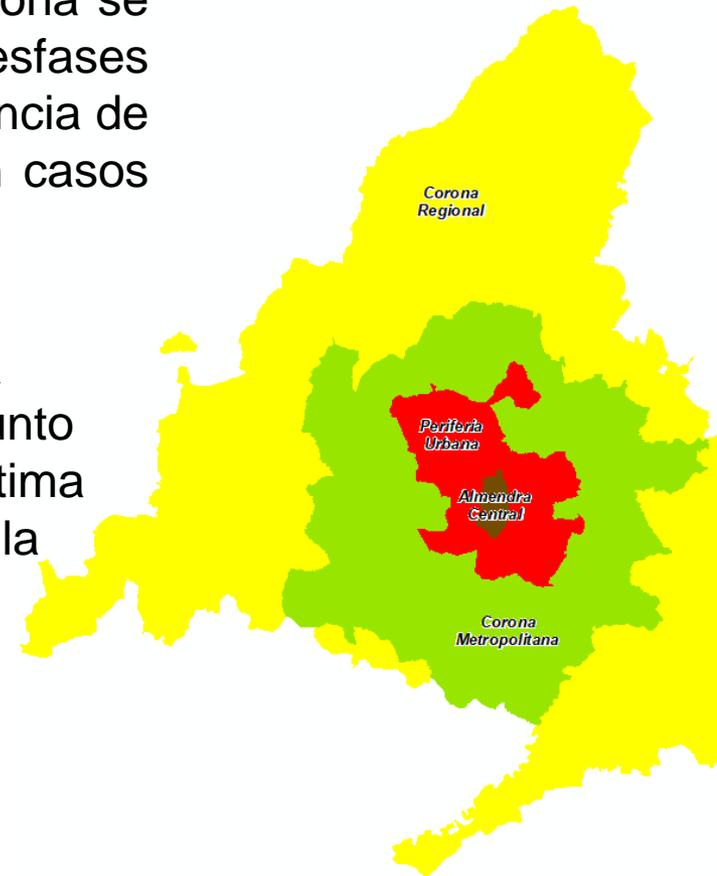
Afijación proporcional al número de hogares/viviendas en cada estrato con un mínimo óptimo de 7 unidades por zona y el resto asignado de forma proporcional al tamaño de la zona.

En cada zona se aplica un muestreo sistemático proporcional al tamaño del hogar con arranque aleatorio.



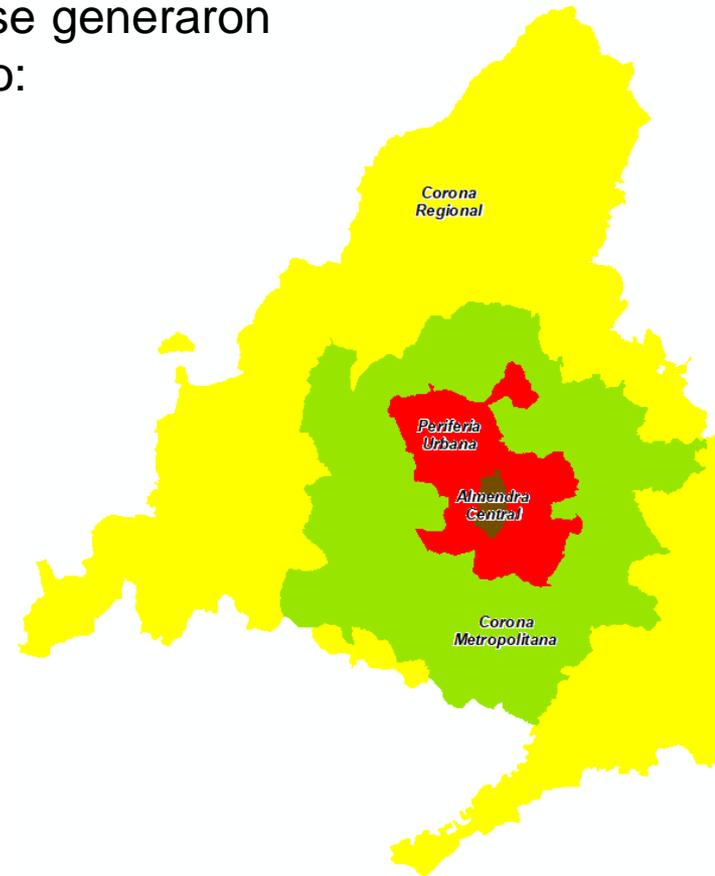
El número mínimo de encuestas por cada zona se planteó con el fin de corregir los posibles desfases de la proporcionalidad pura sin perder la esencia de la misma y evitando muestras ineficaces en casos extremos.

La muestra seleccionada se compuso de una muestra titular de 13.000 unidades y un conjunto de muestras suplentes que asegurasen la óptima consecución de un 100% de cobertura sobre la muestra titular

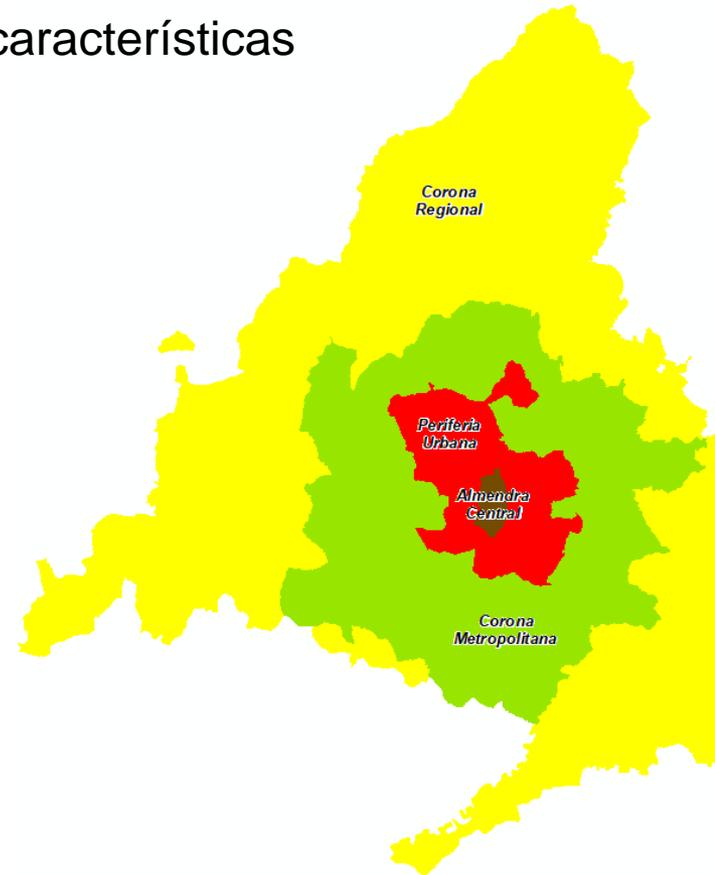


Las sustituciones se plantearon con un criterio de máxima similitud de la muestra original. Para lo cual se generaron 5 tipologías de hogares en función del tamaño:

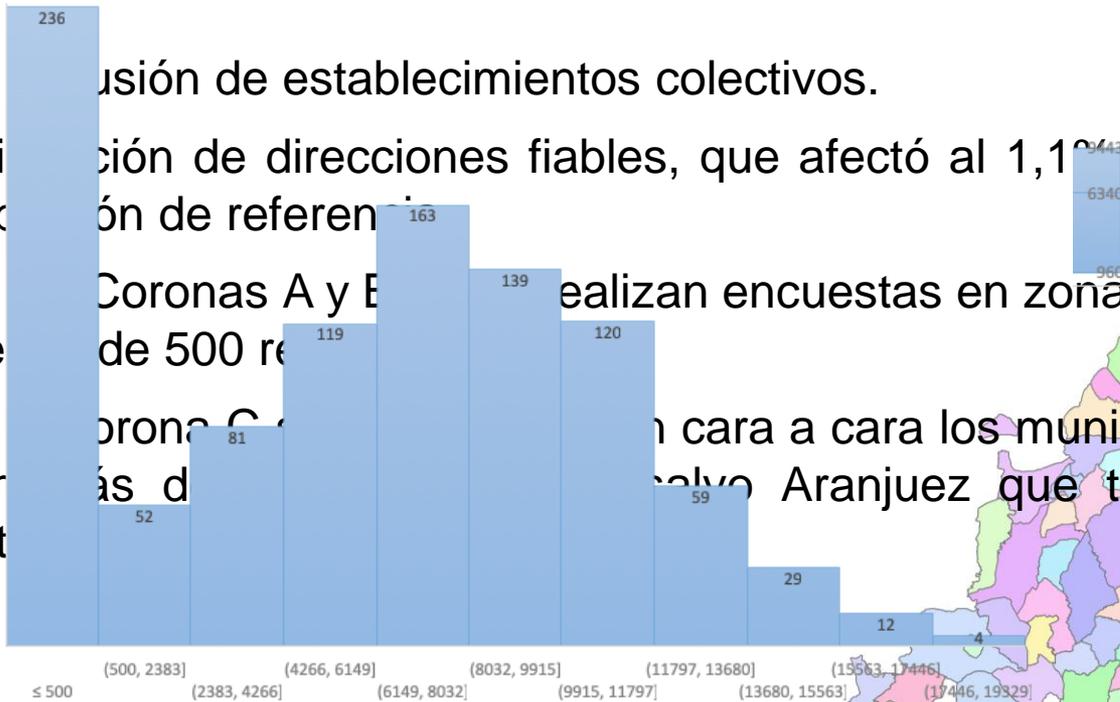
1. Hogares unipersonales
2. Hogares con dos miembros
3. Hogares con tres miembros
4. Hogares con 4 miembros
5. Hogares de 5 a 10 miembros



Para la definición del Marco de muestreo hay que matizar la población objetivo para ser seleccionada en la muestra del estudio, para ello se definieron una serie de características que debía de cumplir el marco:

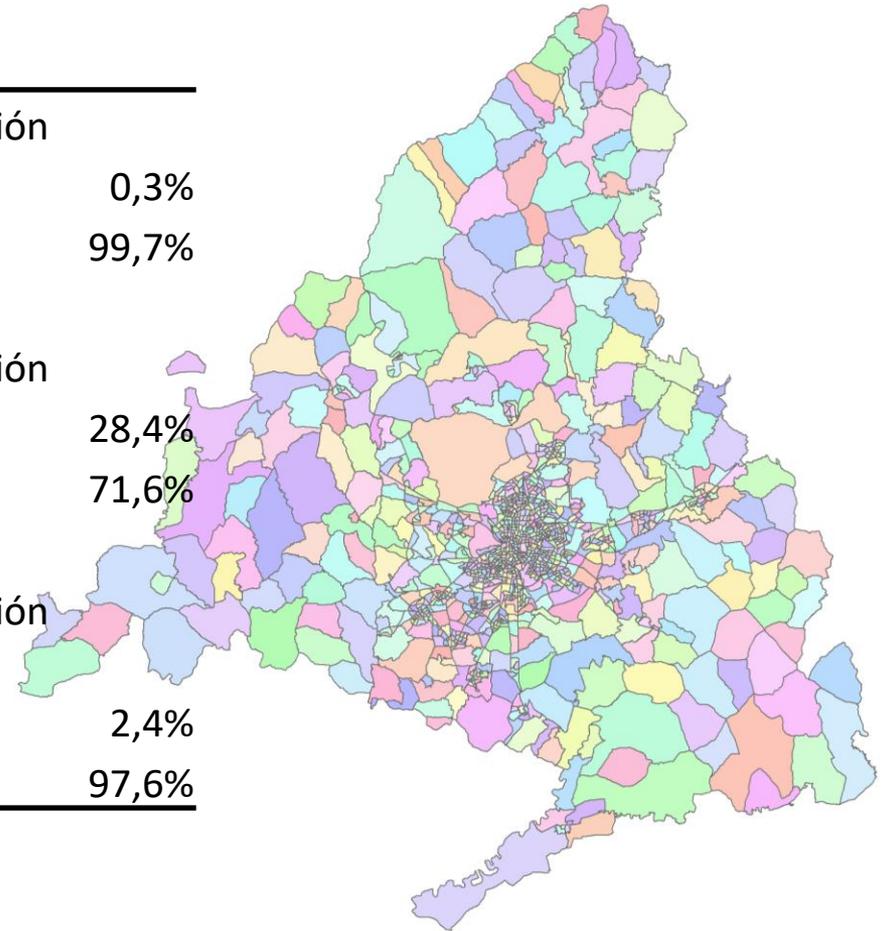


- No inclusión de establecimientos colectivos.
- Asignación de direcciones fiables, que afectó al 1,1% de la población de referencia.
- En Coronas A y B se realizan encuestas en zonas con un número de 500 viviendas.
- En Corona C se realiza encuesta cara a cara los municipios con más de 500 viviendas salvo Aranjuez que tendrá un tratamiento especial.



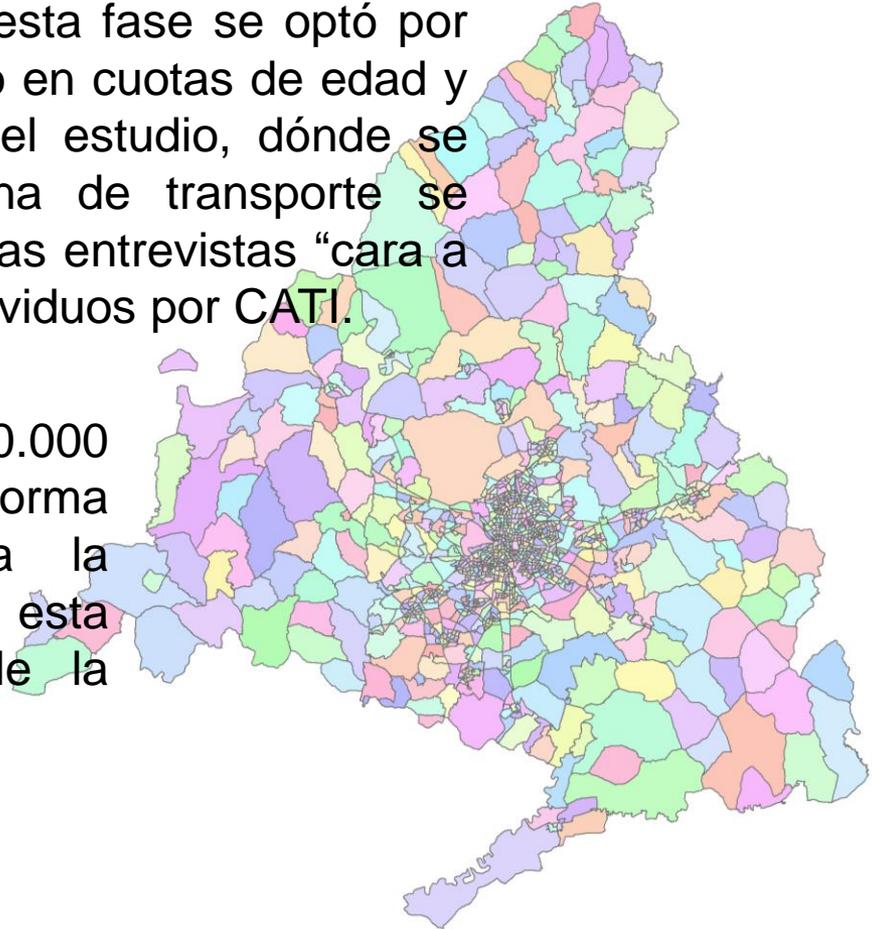
Con todo esto la población que entró a formar parte del marco muestral fue de 6.293.944 residentes lo que supuso un 95,5 % del total de la población recogida en el Padrón.

Zonas		Población	
< 500	23,3%	0,3%	
>500	76,7%	99,7%	
Municipios		Población	
< 5000	70,8%	28,4%	
>5000	29,2%	71,6%	
Zonas		Población	
No muestreadas	28,9%	2,4%	
Incluidas	71,1%	97,6%	



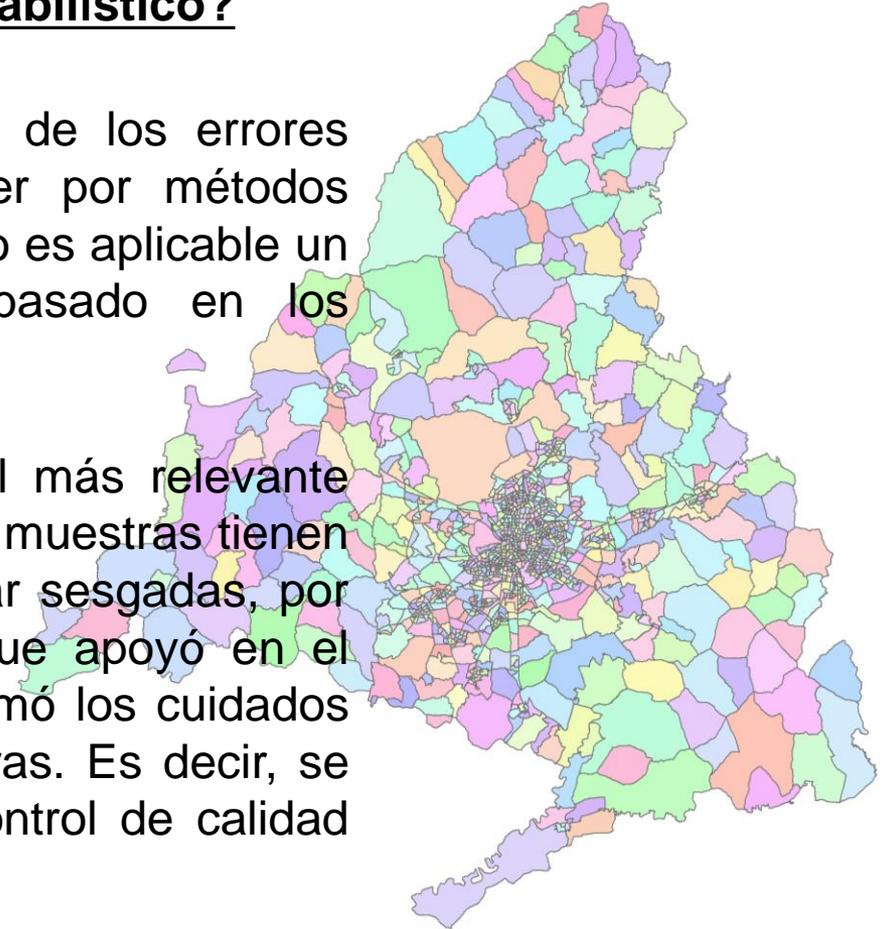
Para la realización del muestreo en esta fase se optó por un muestreo *No probabilístico* basado en cuotas de edad y sexo, respetando el requerimiento del estudio, dónde se consideró que dentro de cada zona de transporte se mantendrá la proporcionalidad entre las entrevistas “cara a cara a familias” y las entrevistas a individuos por CATI.

Así que el reparto de las 50.000 entrevistas a individuos se hizo de forma proporcional y complementaria a la muestra domiciliaria y cumpliendo de esta manera el objetivo de refuerzo de la misma.



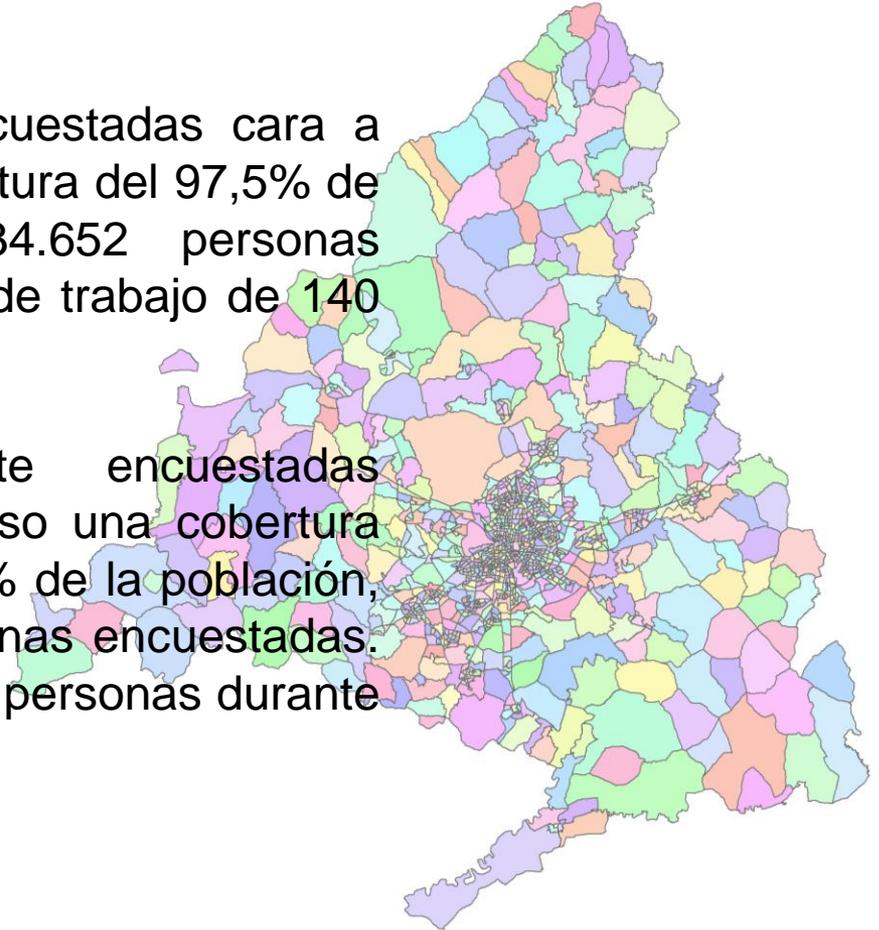
¿Qué implica un muestreo no probabilístico?

- Lo primero, es que el cálculo de los errores muestrales se tiene que hacer por métodos indirectos, lo que supone que no es aplicable un método clásico de cálculo basado en los muestreos probabilísticos.
- El segundo aspecto y quizá el más relevante para la operación, es que estas muestras tienen una alta susceptibilidad de estar sesgadas, por eso, el equipo de muestreo que apoyó en el desarrollo de los trabajos extremó los cuidados en el análisis de dichas muestras. Es decir, se realizó un seguimiento y un control de calidad específico de dichas encuestas.



Magnitudes

- 812 zonas de transporte encuestadas cara a cara, lo que supuso una cobertura del 97,5% de la población objetivo y 34.652 personas encuestadas. Con un equipo de trabajo de 140 personas durante 4 meses.
- 948 zonas de transporte encuestadas Telefónicamente, lo que supuso una cobertura global del estudio de un 99,9% de la población, para un total de 50.412 personas encuestadas. Y un equipo de trabajo de 108 personas durante 4 meses.



Unión de los dos campos y factores de expansión

- Dos muestras diferenciadas, Marco, Universo, diseño muestral, método de encuestación, a quién se encuesta...

Elevación CAPI

Elevación CATI

Análisis de agregabilidad:
Sesgos
Diferencias de resultados
Análisis exploratorio de la variables de clasificación y test Chi cuadrado

Fusión de las muestra: método basado en Granger y Newbold

$$\widehat{X}_t = \sum_{h=1}^L W_h \widehat{X}_h$$

$$\begin{aligned} \widehat{V}(\widehat{X}_t) &= \widehat{V}\left(\sum_{h=1}^L W_h \widehat{X}_h\right) \\ &= \sum_{h=1}^L W_h^2 (1 - f_h) \frac{\widehat{S}_h^2}{n_h} \end{aligned}$$

Validación de la calidad de la información

Distribución por sexos			
	Frecuencia	EDM2018	Padrón2017
1. Hombre	2.980.894	47,7	47,8
2. Mujer	3.263.086	52,3	52,2
Total	6.243.980	100,0	100,0

Distribución por edades			
	Frecuencia	EDM2018	Padrón2017
de 4 a 17	940.921	15,1	15,2
de 18 a 25	515.307	8,3	8,1
de 26 a 64	3.683.775	59,0	59,0
65 o más	1.103.976	17,7	17,8
Total	6.243.980	100,0	100,0

Distribución por tamaño del hogar			
	Frecuencia	EDM2018	Padrón2017
1,00	640.236	10,3	11,9
2,00	1.513.743	24,2	22,3
3,00	1.525.147	24,4	22,3
4,00	1.750.443	28,0	24,4
5,00 o más	814.411	13,0	19,0
Total	6.243.980	100,0	100,0

Validación de la calidad de la información

Distribución por nivel de estudios*			
	Frecuencia	EDM2018	EPA 2017
1. Menos que Primaria	660.459	10,6	4,3
2. Educación Primaria	878.675	14,1	10,4
3. Primera etapa de Educación Secundaria y similar	841.541	13,5	22,6
4. Segunda etapa de Educación Secundaria y similar	1.291.271	20,7	19,1
5. Educación postsecundaria no superior	232.502	3,7	5,7
6. Enseñanzas de Formación Profesional de grado superior y equivalentes + 7. Grados universitarios/licenciaturas/másteres y enseñanzas de doctorado.	599.029	37,5	38,0
Total	6.243.980	100,0	100,0
* hay que tener en cuenta que la muestra EDM2018 incluye edades de 4 a 16 años mientras que la muestra EPA no.			

Validación de la calidad de la información

Distribución por situación laboral*			
	Frecuencia	EDM2018	EPA 2017
1. Trabaja + 2. Trabaja y estudia	2.841.885	45,5	54,0
3. Jubilado / Retirado/ Pensionista	1.201.634	19,2	20,5
4. Parado, ha trabajado antes	463.344	7,4	8,3
5. Parado, busca primer trabajo	24.265	0,4	0,9
6. Estudiante	1.313.134	21,0	6,7
7. Cuidado de familiares +8. Trabajo doméstico no remunerado	314.422	5,0	8,7
9. Otra situación	85.295	1,4	0,9
Total	6.243.980	100,0	100,0

* hay que tener en cuenta que la muestra EDM2018 incluye edades de 4 a 16 años mientras que la muestra EPA no.

Distribución por sector de actividad de los entrevistados activos			
	Frecuencia	EDM2018	EPA 2017
1. Educación pública y privada + 2. Salud privada o pública y servicios sociales + 3. Administración pública	676.827	23,8	22,0
4. Otros Servicios	1.826.327	64,3	64,0
5. Industria	173.156	6,1	8,8
6. Construcción	143.555	5,1	5,1
7. Agricultura	5.805	0,2	0,1
8. Ns/Nc	16.217	0,6	
Total	2.841.885	100,0	100,0

Errores de muestreo

		Número de viajes por individuo					
		Media	Error estándar de la media	Desviación estándar	Recuento	95,0 CL inferior para media	95,0 CL superior para media
Total		2,548	,006	1,632	85064	2,537	2,559
Sexo	Hombre	2,546	,008	1,545	40610	2,531	2,561
	Mujer	2,550	,008	1,708	44454	2,535	2,566
Edad	4-17	2,517	,010	1,157	12819	2,497	2,537
	18-25	2,412	,016	1,304	7020	2,382	2,443
	26-64	2,695	,008	1,727	50185	2,680	2,710
	65 ó más	2,149	,014	1,716	15040	2,121	2,176
¿Tiene Carné de conducir?	No	2,275	,008	1,440	33153	2,259	2,290
	Sí	2,723	,008	1,722	51911	2,708	2,738
Nivel de estudios	Bajo	2,341	,008	1,517	32433	2,325	2,358
	Medio	2,574	,010	1,645	28920	2,555	2,593
	Alto	2,800	,011	1,729	23712	2,778	2,822
Ocupación	Trabaja	2,710	,008	1,551	38716	2,695	2,725
	Jubilado	2,244	,014	1,782	16370	2,217	2,271
	Parado	2,610	,026	2,086	6643	2,560	2,660
	Estudiante	2,513	,009	1,219	17889	2,495	2,530
	Otros	2,356	,028	2,069	5445	2,301	2,411

Los errores de muestreo reales sobre la variable número de viajes se calculan como la raíz de la varianza del estimador para cada uno de los atributos que estamos estimando.

Errores de muestreo

		Número de viajes Totales				
		Suma	Error estándar de suma	95,0 CL inferior para suma	95,0 CL superior para suma	N total
Total		15.847.266	34.948	15.778.768	15.915.765	6.243.980
Sexo	Hombre	7.563.078	22.853	7.518.286	7.607.870	2.980.894
	Mujer	8.284.188	26.441	8.232.364	8.336.013	3.263.086
Edad	4-17	2.360.270	9.612	2.341.431	2.379.110	940.921
	18-25	1.056.628	8.021	1.040.907	1.072.349	515.307
	26-64	10.066.158	28.400	10.010.494	10.121.822	3.683.775
	65 ó más	2.364.210	15.446	2.333.935	2.394.484	1.103.976
¿Tiene Carné de conducir?	No	5.508.575	19.247	5.470.851	5.546.299	2.433.516
	Sí	10.338.691	28.793	10.282.257	10.395.126	3.810.464
Nivel de estudios	Bajo	5.553.565	20.057	5.514.253	5.592.877	2.380.675
	Medio	5.434.832	20.529	5.394.595	5.475.069	2.122.802
	Alto	4.858.869	19.545	4.820.561	4.897.178	1.740.503
Ocupación	Trabaja	7.669.678	22.397	7.625.780	7.713.577	2.841.885
	Jubilado	2.689.489	16.739	2.656.681	2.722.298	1.201.634
	Parado	1.268.861	12.480	1.244.400	1.293.322	487.609
	Estudiante	3.280.463	11.970	3.257.001	3.303.924	1.313.134
	Otros	938.775	11.208	916.807	960.743	399.717

Conclusiones

- Una operación de campo de muy alta complejidad, tanto por el volumen de las encuestas como por el diseño de la operación.
- El esfuerzo se ha centrado en obtener un alta representatividad y evitar sesgos en los resultados.
- Los análisis de los resultados de la muestra nos evidencian unos resultados de muy alta calidad.
- Los resultados han sido muy satisfactorios, con 85064 individuos encuestados, puede que sea el estudio más potente que exista sobre la Comunidad de Madrid más allá de la propia utilidad en cuanto a la movilidad.

edM2018

Muchas gracias

26 febrero 2020

pcortinas@cee.uned.es