



ENCUESTA SOCIAL 2018. HOGARES Y MEDIO AMBIENTE EN ANDALUCÍA ERRORES DE MUESTREO Y DE FALTA DE RESPUESTA

Introducción.....	2
1.- Características de la encuesta.....	3
1.1. Objetivos.....	3
1.2. Ámbitos de la encuesta.....	3
1.3. Marco de la encuesta.....	4
1.4. Diseño de la muestra.....	4
1.5. Trabajos de campo.....	5
2.- Errores de muestreo.....	5
3.- Errores ajenos al muestreo: análisis de la no respuesta.....	7
3.1. Encuestas realizadas según canal.....	8
3.2. Encuestas no realizadas: tasas de respuesta y análisis de las incidencias finales.....	9
3.3. Muestra teórica y efectiva.....	13
4.- Análisis de la eficacia y eficiencia de los trabajos de campo.....	15



Introducción

El desarrollo del Plan Estadístico de Andalucía 2013- 2020 ha renovado el compromiso del Instituto de Estadística y Cartografía de Andalucía con los principios fundamentales de la estadística pública, entre los que se encuentran la confianza en la independencia profesional y en la integridad, responsabilidad y transparencia del Sistema Estadístico y Cartográfico de Andalucía, para mantener y mejorar la calidad de sus estadísticas. El Plan Estadístico de Andalucía 2013-2020 se ajusta a las recomendaciones contenidas en el Código de Buenas Prácticas de las Estadísticas Europeas, adoptado por los miembros del Sistema Estadístico de la Unión Europea el 28 de septiembre de 2011, incorporando a los principios profesionales contemplados en la Ley 4/1989, de 12 de diciembre, de Estadística de la Comunidad Autónoma de Andalucía, otros nuevos que refuerzan la línea de transparencia y calidad de la actividad estadística pública en Andalucía.

En relación con el Código de Buenas Prácticas de las Estadísticas Europeas cabe destacar el principio 12 (precisión y fiabilidad), en el que se declara que las estadísticas europeas deben reflejar la realidad de forma precisa y fidedigna, y que para ello se deben medir y documentar sistemáticamente los errores de muestreo y los que no son de muestreo con arreglo al marco de los componentes de calidad del Sistema Estadístico Europeo.

A este respecto, conviene remarcar que en toda investigación que se realice en una población a través de una encuesta por muestreo es inevitable que surjan errores debido a la existencia de discrepancias entre las afirmaciones y mediciones recogidas en la encuesta y la realidad. Estos errores se suelen dividir en dos tipos:

- Errores de muestreo: se presentan como consecuencia de inferir resultados de un subconjunto de la población a la población entera. Estos errores decrecen al aumentar el tamaño de muestra y pueden estimarse con los datos que se obtienen en la encuesta si se trata de un muestreo probabilístico.
- Errores ajenos al muestreo: engloban al resto de errores y surgen principalmente por los errores y deficiencias que se producen durante el desarrollo y la ejecución de los procedimientos utilizados para realizar la encuesta, por ejemplo definiciones concebidas erróneamente, fracaso al obtener la respuesta de las unidades muestrales, etc.

En este contexto, el Instituto de Estadística y Cartografía de Andalucía impulsa los procesos de adaptación de su producción estadística a los principios de calidad, por lo que, en particular, calcula y publica los errores de muestreo de las variables más representativas de sus encuestas, así como la información de los errores ajenos al muestreo que afectan a los niveles de respuesta.

En relación a los errores ajenos al muestreo, los datos se mostrarán segmentados según una serie de variables demográficas y de territorio. El objetivo es mostrar cómo se distribuyen entre la población objeto de estudio los problemas de no respuesta y de este modo poder localizar los sectores y ámbitos de la población donde ha sido más complicado el proceso de encuesta. Las variables son las siguientes:



- Sexo
- Grupos de edad
- Provincia.
- Grado de urbanización
- Nivel de estudios

1.- Características de la encuesta

1.1. Objetivos

Los objetivos principales de la *Encuesta Social 2018. Hogares y Medio Ambiente en Andalucía* son:

1. Investigar los hábitos, pautas de consumo y actitudes de los hogares en relación con el medio ambiente.
2. Estudiar el equipamiento de las viviendas y el uso que hacen del mismo en relación con los diferentes aspectos del medio ambiente, al objeto de que se puedan formular políticas medioambientales relacionadas con los hogares (ahorro energético, reducción del consumo de energía, separación de basuras, etc.).
3. Medir la evolución de estos fenómenos en la última década comparando los resultados con los obtenidos en Encuesta Social 2008: Medio Ambiente y Hogares en Andalucía.

1.2. Ámbitos de la encuesta

Ámbito territorial

La encuesta se extiende a todo el territorio de la Comunidad Autónoma de Andalucía.

Ámbito poblacional

La población objeto de la investigación es la población residente en viviendas familiares de Andalucía en las que reside al menos un individuo de 16 y más años.

No se consideran los llamados hogares colectivos (hospitales, residencias, cuarteles, etc.). Sí se incluyen, sin embargo, las familias que, formando un grupo independiente, residen en dichos establecimientos colectivos (por ejemplo, el director o el conserje del centro).

Ámbito temporal

El trabajo de campo se ha realizado entre los días 13 de mayo y 12 de julio de 2019.

1.3. Marco de la encuesta

El marco de población utilizado para extraer la muestra procede de la Base Longitudinal de Datos de Población de Andalucía a fecha 1 de julio de 2018. De este marco, formado por personas residentes en Andalucía con más de 15 años de edad que residieran en una vivienda no colectiva (número de miembros del hogar menor que 9), se extrajo la muestra. Posteriormente, esta muestra se cruzó con la información procedente de la Base de Datos de Usuarios (BDU) del Sistema Sanitario Público de Andalucía para obtener los números de teléfono de las unidades muestrales seleccionadas.

1.4. Diseño de la muestra

Tipo de muestreo: unidades muestrales y estratificación

Se ha utilizado un muestreo aleatorio simple estratificado por provincia y grado de urbanización (24 zonas), como unidades de selección se han tomado las personas residentes en estos estratos. En la siguiente tabla se observa la distribución por estos dos estratos.

	Zona Urbana	Zona de densidad intermedia	Zona Rural
Almería	34,3%	51,6%	14,1%
Cádiz	68,5%	27,4%	4,1%
Córdoba	41,0%	44,0%	15,0%
Granada	35,4%	42,6%	22,0%
Huelva	27,5%	53,6%	18,9%
Jaén	26,4%	50,4%	23,2%
Málaga	55,8%	29,4%	14,8%
Sevilla	56,4%	38,0%	5,6%

Tamaño de la muestra

El tamaño efectivo de la muestra ha sido de 3.005 personas. El tamaño teórico de la muestra era de 5.000 personas. No se ha utilizado reemplazo de muestra.

Afijación

Entre las provincias se usa una afijación de compromiso entre la uniforme y la proporcional al tamaño de cada provincia. En las zonas la afijación es proporcional al tamaño poblacional de cada una dentro de la misma provincia.

Selección de la muestra

La selección de la muestra se ha realizado de tal forma que dentro de cada estrato cualquier individuo tenga la misma probabilidad de ser seleccionado, es decir, se tengan muestras autoponderadas dentro de cada estrato. La probabilidad de selección del individuo i en el estrato h será:

$$P_{i,h} = \frac{n_h}{N_h}$$

donde:

n_h es el tamaño del estrato h en la muestra

N_h es el tamaño del estrato h en la población

1.5. Trabajos de campo

El trabajo de campo se ha realizado entre el 13 de mayo al 12 de julio de 2019. El método de recogida utilizado en la encuesta realizada por el IECA ha sido multicanal. La información ha sido recogida mediante entrevista telefónica asistida por ordenador (CATI) y también se ofreció la posibilidad de suministrar la información a través de Internet (CAWI). La administración, control y seguimiento de la recogida de información se ha efectuado a través de PI@teA, la plataforma de recogida de encuestas del IECA.

2.- Errores de muestreo

Para el cálculo de los errores de muestreo de las principales características investigadas se utiliza el método indirecto Bootstrap, basado en el remuestreo a partir de pseudopoblaciones generadas según Booth et al (1994)¹.

Este método es utilizado, por ejemplo, por la Oficina de Estadística de Holanda en Kuijvenhoven and Scholtus (2011)². Se basa en la generación de pseudopoblaciones a partir de

¹ Booth J. G., Butler R. W. and Hall P. (1994), 'Bootstrap Methods for Finite Populations', Journal of the American Statistical Association 89, pp. 1282–1289.

² Kuijvenhoven, L and Scholtus, S. (2011): Bootstrapping Combined Estimator based on Register and Sample Survey Data. Discussion paper 20112 Statistics Netherlands.

la muestra, donde cada unidad muestral se repite d_k veces, sean $d_k = \frac{1}{\pi_k}$ los pesos de diseño calculados como la inversa de la probabilidad de sección de la unidad muestral. Posteriormente, se aplica el diseño muestral original a las pseudopoblaciones generadas, y se calcula el estimador original a cada muestra Bootstrap obtenida.

Estos son los 4 pasos que componen el algoritmo:

1. Dado que los pesos de diseño pueden contener decimales, una forma estocástica de redondearlos sería a partir de $d_k = \lfloor d_k \rfloor + \varphi_k$ (con $\varphi_k \in [0,1)$). Por tanto se redondearía hacia abajo $\delta_k = \lfloor d_k \rfloor$ con probabilidad $1-\varphi_k$ y hacia arriba $\delta_k = \lfloor d_k \rfloor + 1$ con probabilidad φ_k . Posteriormente, se genera una pseudopoblación \hat{U} replicando δ_k veces cada elemento k de la muestra original s
2. Se extrae una muestra s^i de \hat{U} con el diseño muestral original, y se obtienen los pesos Bootstrap siguiendo el mismo procedimiento que para la muestra completa: calculando los pesos iniciales como inversa de la probabilidad de selección $d_k^i = \frac{1}{\pi_k}$ y aplicando el mismo procedimiento de ajustes (corrección por falta de respuesta y calibración) que se aplicaron a los pesos de la muestra completa para obtener los pesos Bootstrap $w_{k,b}^i$. Finalmente obtenemos el estimador $\hat{\theta}_b^i$ para cada muestra Bootstrap con los pesos obtenidos tras los ajustes $w_{k,b}^i$
3. El paso 2 se repite C veces (C=1000 veces en nuestro caso) y se obtienen $\hat{\theta}_1^i, \dots, \hat{\theta}_C^i$. Se calcula:

$$v_{boot}^b = \frac{1}{C-1} \sum_{c=1}^C (\hat{\theta}_c^i - \bar{\theta}^i)^2,$$

$$\bar{\theta}^i = \frac{1}{C} \sum_{c=1}^C \hat{\theta}_c^i$$

4. Los pasos 1 a 3 se repiten B veces (B=1 en nuestro caso) y se obtienen $v_{boot}^1, \dots, v_{boot}^B$. Finalmente la varianza Bootstrap del estimador $\hat{\theta}$ es dada por:

$$v_{boot}(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^B v_{boot}^b$$

En la página web del IECA se publican los errores de muestreo de algunas de las principales variables objeto de estudio:

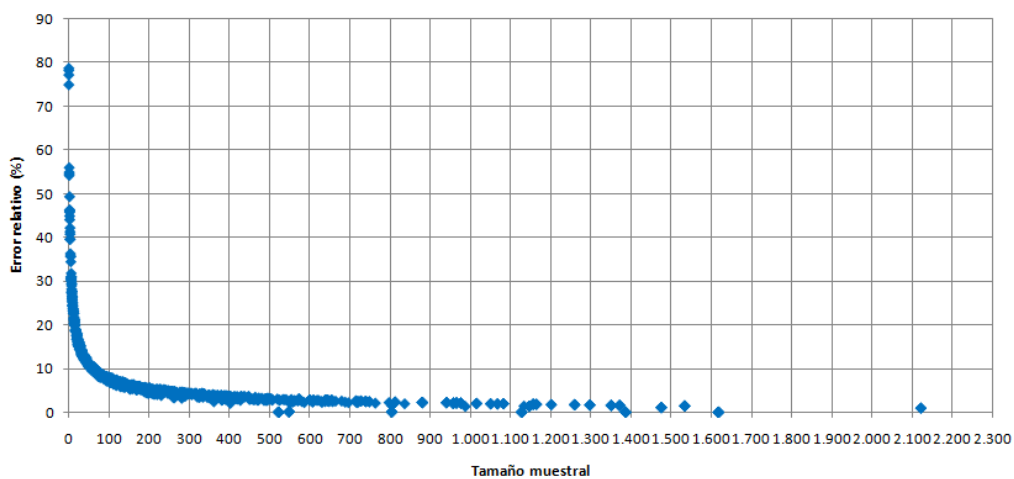
<https://www.juntadeandalucia.es/institutodeestadisticaycartografia>

Las cifras se muestran en términos de errores relativos en porcentaje (coeficiente de variación).

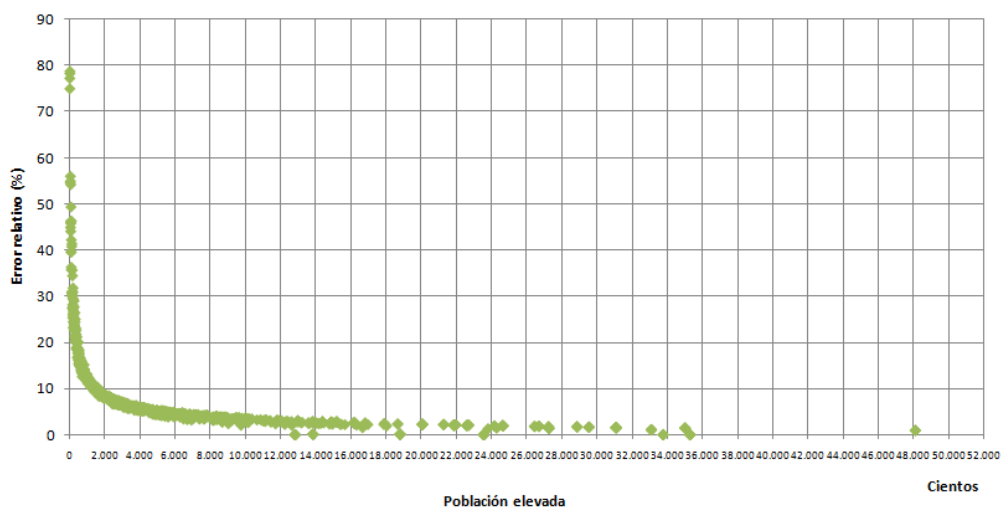
$$\widehat{CV}(\hat{\theta}) = \frac{\sqrt{v_{boot}(\hat{\theta})}}{\hat{\theta}} \cdot 100$$

En los siguientes gráficos se muestran la distribución de los errores relativos calculados en relación con el tamaño de muestra y la población elevada:

Errores de muestreo según tamaño muestral



Errores de muestreo según población elevada



Según los errores calculados, aquellas estimaciones obtenidas con menos de 20 unidades muestrales están sujetas a una alta variabilidad, con errores de muestreo relativos superiores al 20% en la mayoría de los casos. Por tanto estas estimaciones carecen de representatividad estadística y no se han mostrado en la tabulación de los resultados.

3.- Errores ajenos al muestreo: análisis de la no respuesta

El estudio de los errores ajenos al muestreo engloba numerosos elementos.

Biemer y Liberg (2003) identifican cinco componentes del error de no muestreo: Error de especificación, Error de medida, Error de proceso, Error de cobertura o marco y Error debido a la no respuesta.

- Error de especificación: Los errores de especificación ocurren si los datos recogidos presentan inconsistencias con respecto a los objetivos fijados en la encuesta. Este es el caso que se da cuando, por ejemplo, el entrevistado interpreta alguna pregunta de manera distinta a como se fijó en un principio.
- Error de medida: Los errores de medida o de respuesta se producen cuando existen diferencias entre el valor recogido y el valor real.
- Error de proceso: Se producen en la fase de procesamiento tras la recogida de los datos, errores en la depuración, codificación, grabación, tabulación, etc. de los resultados.
- Error de cobertura: Se define el marco de una encuesta como «todo el material y toda la información que identifique, distinga y permita acceder a los elementos que conforman la población objetivo, y que permite enlazar e identificar los elementos de la población con las unidades de muestreo bases del diseño de la encuesta». El marco perfecto es aquel en el que es posible asociar a cada elemento de la población una única unidad de muestreo de las delimitadas en el marco y, recíprocamente, es posible asociar cada unidad de muestreo con un único elemento de la población. En el caso de no disponer de un marco perfecto se introducirán imperfecciones que a su vez producirán en las estimaciones los llamados errores de cobertura. El error de cobertura surge si no existe una correspondencia uno a uno entre las unidades del marco y las de la población objeto de estudio. Existen varias situaciones en las que se pueden producir estos errores: a) algunos elementos de la población objeto de estudio no se encuentran en el marco, produciéndose una infraestimación de la población b). algunas unidades del marco no están en la población objeto de estudio, dando lugar a una sobreestimación de la población c) varias unidades del marco se corresponden con una unidad en la población objeto de estudio (uno a muchos) y d) varias unidades de la población objeto de estudio se corresponden con un único elemento del marco (muchos a uno).
- El error de no respuesta es aquel que se produce al fracasar en la recolección de la información de las unidades muestrales. Se distinguen dos tipos de no respuesta: No respuesta total: no se consigue realizar una entrevista con la persona seleccionada y no respuesta parcial, que se solventa en algún caso con imputación.

En los siguientes apartados se mostrarán los datos relativos al impacto de la no respuesta total en el campo.

3.1. Encuestas realizadas según canal

Antes de analizar la no respuesta, se muestra la distribución de las encuestas a través del canal que se ha empleado para su recogida. Dentro de las encuestas telefónicas se señalan como una categoría específica aquellas que se han recuperado, es decir, encuestas que en un primer momento no se pudieron realizar pero que finalmente se realizaron como fruto de trabajos adicionales de recuperación. Por ejemplo, personas que en un primer momento se negaron a realizar la encuesta, pero que después de una nueva llamada días después por otro equipo distinto de encuestación, cambiaron de opinión y accedieron a realizarla. A continuación se muestran los datos por canal de recogida.

Encuestas realizadas por canal de encuesta según variables de residencia, edad y sexo. Absolutos.

	Vía Telefónica			Vía Web	Total
	Vía Telefónica	Vía Telefónica (recuperación)	Total Vía Telefónica		
Total	2369	185	2554	451	3005
Hombre	1072	88	1160	227	1387
Mujer	1297	97	1394	224	1618
16-24	230	25	255	86	341
25-34	306	35	341	87	428
35-44	426	49	475	85	560
45-54	529	37	566	81	647
55-64	398	21	419	62	481
65 y más	480	18	498	50	548
Almería	173	13	186	26	212
Cádiz	365	25	390	62	452
Córdoba	240	19	259	49	308
Granada	281	28	309	41	350
Huelva	141	9	150	26	176
Jaén	188	13	201	41	242
Málaga	425	34	459	89	548
Sevilla	556	44	600	117	717
Ciudades	1137	100	1237	239	1476
Zonas de densidad intermedia	921	65	986	161	1147
Zonas rurales	311	20	331	51	382

**Encuestas realizadas por canal de encuesta
según variables de residencia, edad y sexo. Porcentajes.**

	Vía Telefónica			Vía Web	
	Vía Telefónica	Vía Telefónica (recuperación)	Total Vía Telefónica		
Total	78,8%	6,2%	85,0%	15,0%	100,0%
Hombre	77,3%	6,3%	83,6%	16,4%	100,0%
Mujer	80,2%	6,0%	86,2%	13,8%	100,0%
16-24	67,4%	7,3%	74,8%	25,2%	100,0%
25-34	71,5%	8,2%	79,7%	20,3%	100,0%
35-44	76,1%	8,8%	84,8%	15,2%	100,0%
45-54	81,8%	5,7%	87,5%	12,5%	100,0%
55-64	82,7%	4,4%	87,1%	12,9%	100,0%
65 y más	87,6%	3,3%	90,9%	9,1%	100,0%
Almería	81,6%	6,1%	87,7%	12,3%	100,0%
Cádiz	80,8%	5,5%	86,3%	13,7%	100,0%
Córdoba	77,9%	6,2%	84,1%	15,9%	100,0%
Granada	80,3%	8,0%	88,3%	11,7%	100,0%
Huelva	80,1%	5,1%	85,2%	14,8%	100,0%
Jaén	77,7%	5,4%	83,1%	16,9%	100,0%
Málaga	77,6%	6,2%	83,8%	16,2%	100,0%
Sevilla	77,5%	6,1%	83,7%	16,3%	100,0%
Ciudades	77,0%	6,8%	83,8%	16,2%	100,0%
Zonas de densidad intermedia	80,3%	5,7%	86,0%	14,0%	100,0%
Zonas rurales	81,4%	5,2%	86,6%	13,4%	100,0%

3.2. Encuestas no realizadas: tasas de respuesta y análisis de las incidencias finales

Durante la realización de las encuestas telefónicas surgen distintas situaciones que pueden tener como consecuencia que la encuesta no se realice. Es lo que se denomina **incidencias** en campo.

De estas incidencias, las más importantes son las **incidencias finales**: aquellas que después de varios intentos tienen como resultado final que esa encuesta no puede realizarse. En la siguiente tabla se muestran las tasas de respuesta según variables sociodemográficas y territoriales. Las tasas de respuesta hacen referencia al porcentaje de encuestas realizadas o muestra efectiva (3.005 encuestas) respecto al total de la muestra teórica (5.000 unidades muestrales).

**Encuestas realizadas en proporción a la muestra teórica (tasa de respuesta)
 según variables de residencia, edad y sexo.**

	Total encuestas realizadas	Total encuestas no realizadas	Total muestra	Tasa de respuesta
Total	3005	1995	5000	60,1%
Hombre	1387	1024	2411	57,5%
Mujer	1618	971	2589	62,5%
16-24	341	227	568	60,0%
25-34	428	342	770	55,6%
35-44	560	365	925	60,5%
45-54	647	363	1010	64,1%
55-64	481	240	721	66,7%
65 y más	548	458	1006	54,5%
Almería	212	201	413	51,3%
Cádiz	452	287	739	61,2%
Córdoba	308	159	467	66,0%
Granada	350	201	551	63,5%
Huelva	176	136	312	56,4%
Jaén	242	142	384	63,0%
Málaga	548	444	992	55,2%
Sevilla	717	425	1142	62,8%
Ciudades	1476	944	2420	61,0%
Zonas de densidad intermedia	1147	795	1942	59,1%
Zonas rurales	382	256	638	59,9%

Más allá de las tasas de respuesta, es importante analizar de forma específica los motivos o incidencias finales por los que las encuestas no se realizan. Estas incidencias finales se pueden agrupar en dos tipos:

- **Incidencias de marco:** agrupan los motivos que imposibilitan la realización de la encuesta por problemas en el marco muestral. Bien porque no se pudo localizar un teléfono con el que poder establecer contacto con la persona en muestra, o bien porque el marco de viviendas no estaba lo suficientemente actualizado.
- **Incidencias de relación-situación:** agrupan los motivos que imposibilitan la realización de la encuesta por situaciones de diversa índole que tienen que ver con las personas encuestadas: no se las puede localizar o no quieren realizar la encuesta o hay algún elemento que lo imposibilita.

A continuación se muestran las incidencias concretas de cada uno de estos tipos.

A) INCIDENCIAS DE MARCO

Teléfono no existe, FAX (código: TNE).

Esta incidencia incluye situaciones en las que el marco telefónico proporciona información de números que son erróneos. Por ejemplo cuando las operadoras telefónicas indicaban que el teléfono marcado no existía, cuando saltaba un fax o se escuchaba una locución en la que se informaba que ese teléfono tenía las llamadas restringidas.

No contactable (código: NC)

Estos casos se producen cuando la información del marco está desactualizada. Este problema puede afectar tanto al marco de viviendas como al marco telefónico. Las incidencias que se agrupan en esta categoría se producen principalmente porque la persona seleccionada vive en un municipio distinto al que aparece en la muestra o porque el marco telefónico no tenía teléfono en algunas unidades muestrales. En casos minoritarios se produce porque la persona no es contactable en ese teléfono/vivienda por circunstancias tales como fallecimiento, divorcio/separación, etc.)

B) INCIDENCIAS DE RELACIÓN-SITUACIÓN

Sin contacto (código: SC).

Tras un número de intentos, que se explican en el Cuadro 1, no se logra contactar con la vivienda: nadie atiende el teléfono o salta el contestador.

Ausencia (código: AUS)

Se produce cuando no se consigue contactar con la persona seleccionada en los días y horarios de duración de los trabajos de campo. Antes de sustituir la vivienda por esta incidencia, los encuestadores tenían que realizar un número concreto de intentos (Cuadro 1).

Imposibilidad para contestar (código: INC)

Se produce esta incidencia la persona seleccionada no podía realizar la encuesta por incapacidad para responder a la misma a causa de alguna discapacidad, edad, enfermedad o cualquier otra circunstancia. En caso de no poder comunicarse en español con la persona con la que se contactaba, la encuesta se realizaba en inglés: el cuestionario estaba traducido y había encuestadores con capacidad para realizarlo en este idioma.

Negativas (código: NEG)

Esta incidencia se produce cuando la persona seleccionada o la persona que atiende el teléfono se niega a realizar la encuesta. También se considera negativa cuando un encuestado se niega a continuar después de haber empezado la encuesta.

Algunas de estas incidencias provocaban que la encuesta se elimine de la muestra. En otras por el contrario, se ponía en marcha un protocolo de intentos por parte de los encuestadores.

	Incidencia	Protocolo Sustitución de muestra
Incidencias de marco	Teléfono no existe (código: TNE)	Eliminación directa
	No contactable (código: NC)	Eliminación directa
Incidencias de relación	Sin contacto (código: SC)	Solo se elimina si no hay contacto tras intentarlo 15 veces: 5 días distintos, en 3 intentos diarios distribuidos en franjas horarias de mañana, tarde y noche.
	Ausencia (código: AUS)	Solo hay sustitución si no hay contacto tras intentarlo 15 veces: 5 días distintos, en 3 intentos diarios distribuidos en franjas horarias de mañana, tarde y noche.
	Imposibilidad para contestar (código: INC)	Eliminación directa
	Negativa (código: NEG)	Tres intentos más por parte de un equipo especialista en recuperación

Una vez presentados los tipos de incidencias, a continuación se muestra la presencia de cada una de ellas en los trabajos de campo telefónicos de la “Encuesta Social 2018. Hogares y Medio Ambiente en Andalucía”.

**Encuestas no realizadas. Incidencias finales en campo
según variables de residencia, edad y sexo. Absolutos.**

	Incidencias de Marco		Incidencias de relación-situación				
	TNE	NC	SC	AUS	NEG	INC	Total
Total	388	413	723	16	346	109	1995
Hombre	215	220	381	7	153	48	1024
Mujer	173	193	342	9	193	61	971
16-24	40	74	89	1	23	0	227
25-34	75	85	131	4	45	2	342
35-44	76	66	157	3	52	11	365
45-54	63	78	150	3	63	6	363
55-64	50	37	91	2	49	11	240
65 y más	84	73	105	3	114	79	458
Almería	35	51	86	0	15	14	201
Cádiz	57	52	108	1	52	17	287
Córdoba	29	35	57	4	25	9	159
Granada	32	40	65	3	50	11	201
Huelva	25	29	56	2	22	2	136
Jaén	30	29	59	0	21	3	142
Málaga	101	94	148	1	72	28	444
Sevilla	79	83	144	5	89	25	425
Ciudades	190	199	329	9	170	47	944
Zonas de densidad intermedia	136	160	311	5	140	43	795
Zonas rurales	62	54	83	2	36	19	256

**Encuestas no realizadas. Incidencias finales en campo
 según variables de residencia, edad y sexo. Porcentajes.**

	Incidencias de Marco		Incidencias de relación-situación				
	TNE	NC	SC	AUS	NEG	INC	Total
Total	19,4%	20,7%	36,2%	0,8%	17,3%	5,5%	100,0%
Hombre	21,0%	21,5%	37,2%	0,7%	14,9%	4,7%	100,0%
Mujer	17,8%	19,9%	35,2%	0,9%	19,9%	6,3%	100,0%
16-24	17,6%	32,6%	39,2%	0,4%	10,1%	0,0%	100,0%
25-34	21,9%	24,9%	38,3%	1,2%	13,2%	0,6%	100,0%
35-44	20,8%	18,1%	43,0%	0,8%	14,2%	3,0%	100,0%
45-54	17,4%	21,5%	41,3%	0,8%	17,4%	1,7%	100,0%
55-64	20,8%	15,4%	37,9%	0,8%	20,4%	4,6%	100,0%
65 y más	18,3%	15,9%	22,9%	0,7%	24,9%	17,2%	100,0%
Almería	17,4%	25,4%	42,8%	0,0%	7,5%	7,0%	100,0%
Cádiz	19,9%	18,1%	37,6%	0,3%	18,1%	5,9%	100,0%
Córdoba	18,2%	22,0%	35,8%	2,5%	15,7%	5,7%	100,0%
Granada	15,9%	19,9%	32,3%	1,5%	24,9%	5,5%	100,0%
Huelva	18,4%	21,3%	41,2%	1,5%	16,2%	1,5%	100,0%
Jaén	21,1%	20,4%	41,5%	0,0%	14,8%	2,1%	100,0%
Málaga	22,7%	21,2%	33,3%	0,2%	16,2%	6,3%	100,0%
Sevilla	18,6%	19,5%	33,9%	1,2%	20,9%	5,9%	100,0%
Ciudades	20,1%	21,1%	34,9%	1,0%	18,0%	5,0%	100,0%
Zonas de densidad intermedia	17,1%	20,1%	39,1%	0,6%	17,6%	5,4%	100,0%
Zonas rurales	24,2%	21,1%	32,4%	0,8%	14,1%	7,4%	100,0%

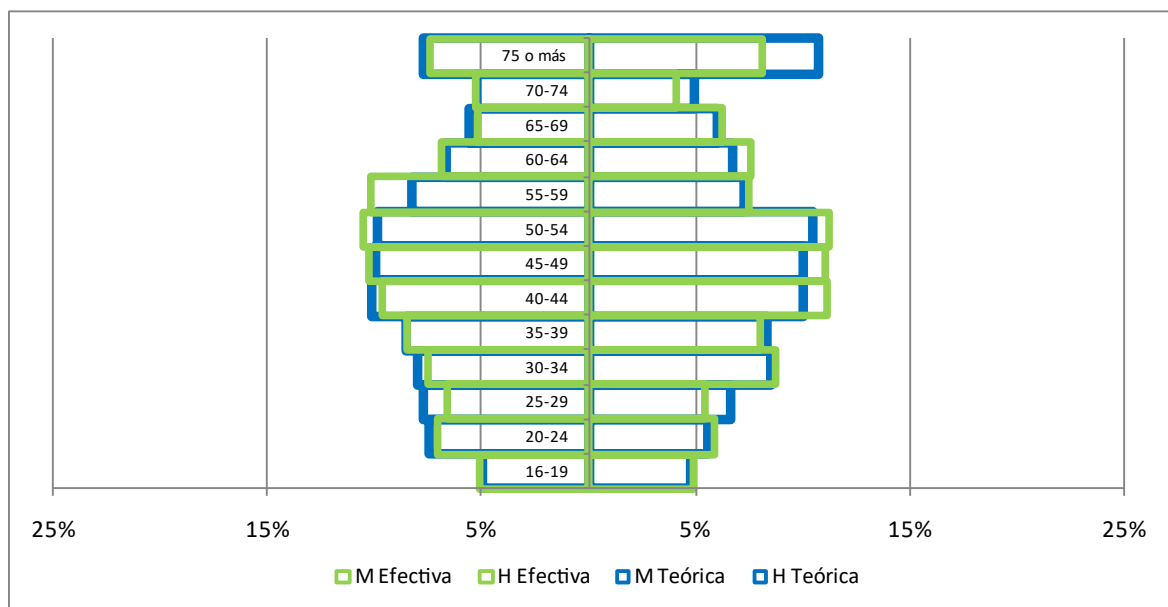
3.3. Muestra teórica y efectiva

Como efecto de estas incidencias, se producen desviaciones entre la muestra teórica (la resultante del diseño original) y la muestra efectiva (la resultante después de los trabajos de campo). Para profundizar más en los sectores de la población en los que se han producido mayores desviaciones, utilizamos la edad y el sexo. Estas variables están disponibles en el marco original, por lo que tenemos información de ellas tanto de las personas encuestadas como de las que no.

Desviaciones entre la muestra teórica y efectiva por sexo y edad. Absolutos y porcentajes

	Muestra teórica				Muestra efectiva			
	Hombre	Mujer	Hombre %	Mujer %	Hombre	Mujer	Hombre %	Mujer %
16-19	120	124	5,0%	4,8%	70	79	5,1%	4,9%
20-24	180	144	7,5%	5,6%	97	95	7,0%	5,9%
25-29	186	172	7,7%	6,6%	92	88	6,6%	5,4%
30-34	192	220	8,0%	8,5%	104	142	7,5%	8,8%
35-39	206	216	8,5%	8,3%	117	130	8,4%	8,0%
40-44	243	260	10,1%	10,0%	133	181	9,6%	11,2%
45-49	240	260	10,0%	10,0%	142	179	10,3%	11,0%
50-54	238	272	9,9%	10,5%	146	182	10,5%	11,2%
55-59	199	188	8,3%	7,3%	141	122	10,2%	7,5%
60-64	160	174	6,6%	6,7%	95	123	6,9%	7,6%
65-69	135	155	5,6%	6,0%	72	101	5,2%	6,2%
70-74	125	127	5,2%	4,9%	73	67	5,3%	4,1%
75 o más	187	277	7,8%	10,7%	103	131	7,4%	8,1%
Total	2411	2589	100,0%	100,0%	1385	1620	100,0%	100,0%

Pirámide de desviaciones entre la muestra teórica y efectiva por sexo y edad (grupos quinquenales)



4.- Análisis de la eficacia y eficiencia de los trabajos de campo

Además de la calidad de la muestra, hay otros factores que son interesantes para evaluar cómo se han desarrollado los trabajos de campo en una operación de encuesta. Uno de ellos es conocer cómo se ha funcionado en términos de eficacia y eficiencia. La forma más directa de medir este particular es calcular el número de intentos o llamadas que se han tenido que realizar para poder hacer cada encuesta. Este tipo de información también es muy útil para poder diseñar estrategias que optimicen los intentos y aumenten los niveles de muestra en próximas operaciones.

En las siguientes tablas se pueden observar estos datos segmentados por las mismas variables que en las tablas anteriores.

Número de intentos telefónicos para realizar las encuestas según variables de residencia, edad y sexo. Absolutos

	1 intento	2-3 intentos	4-6 intentos	7-10 intentos	Más de 10 intentos	Total
Total	518	717	473	285	261	2254
Hombre	207	336	211	143	118	1015
Mujer	311	381	262	142	143	1239
16-24	29	64	53	29	40	215
25-34	51	70	77	48	37	283
35-44	85	139	88	40	57	409
45-54	110	168	104	65	59	506
55-64	109	135	69	47	28	388
65 y más	134	141	82	56	40	453
Almería	31	50	35	24	23	163
Cádiz	85	103	72	41	43	344
Córdoba	62	54	68	30	18	232
Granada	60	80	56	39	36	271
Huelva	27	53	25	19	16	140
Jaén	46	55	35	20	20	176
Málaga	88	144	70	45	52	399
Sevilla	119	178	112	67	53	529
Ciudades	224	366	232	136	114	1072
Zonas de densidad intermedia	223	252	180	113	116	884
Zonas rurales	71	99	61	36	31	298

Número de intentos telefónicos para realizar las encuestas según variables de residencia, edad y sexo. Porcentajes

	1 intento	2-3 intentos	4-6 intentos	7-10 intentos	Más de 10 intentos	Total
Total	23,0%	31,8%	21,0%	12,6%	11,6%	100,0%
Hombre	20,4%	33,1%	20,8%	14,1%	11,6%	100,0%
Mujer	25,1%	30,8%	21,1%	11,5%	11,5%	100,0%
16-24	13,5%	29,8%	24,7%	13,5%	18,6%	100,0%
25-34	18,0%	24,7%	27,2%	17,0%	13,1%	100,0%
35-44	20,8%	34,0%	21,5%	9,8%	13,9%	100,0%
45-54	21,7%	33,2%	20,6%	12,8%	11,7%	100,0%
55-64	28,1%	34,8%	17,8%	12,1%	7,2%	100,0%
65 y más	29,6%	31,1%	18,1%	12,4%	8,8%	100,0%
Almería	19,0%	30,7%	21,5%	14,7%	14,1%	100,0%
Cádiz	24,7%	29,9%	20,9%	11,9%	12,5%	100,0%
Córdoba	26,7%	23,3%	29,3%	12,9%	7,8%	100,0%
Granada	22,1%	29,5%	20,7%	14,4%	13,3%	100,0%
Huelva	19,3%	37,9%	17,9%	13,6%	11,4%	100,0%
Jaén	26,1%	31,3%	19,9%	11,4%	11,4%	100,0%
Málaga	22,1%	36,1%	17,5%	11,3%	13,0%	100,0%
Sevilla	22,5%	33,6%	21,2%	12,7%	10,0%	100,0%
Ciudades	20,9%	34,1%	21,6%	12,7%	10,6%	100,0%
Zonas de densidad intermedia	25,2%	28,5%	20,4%	12,8%	13,1%	100,0%
Zonas rurales	23,8%	33,2%	20,5%	12,1%	10,4%	100,0%