

Anexo 2

METODOLOGÍA

Encuesta Europea de Violencia de Género 2022



Índice

ANEXO 2. METODOLOGÍA	1
1. Introducción	1
2. Objetivo de la encuesta	1
3. Ámbitos de la investigación	2
3.1. Ámbito geográfico	2
3.2. Ámbito poblacional	2
3.3. Ámbito temporal	2
4. Diseño muestral	2
4.1. Tipo de muestreo	2
4.2. Marco de selección	2
4.3. Estratos	3
4.4. Tamaño muestra	3
4.5. Estimadores	3
4.6. Cálculo de errores de muestreo	4
5. Diseño del cuestionario	5
5.1. Idiomas del cuestionario	5
6. Recogida de la información	5
6.1. Métodos de recogida	5
6.2. Tratamiento de la información	6
6.2.1. Validación de datos	6
6.2.2. Compilación de datos	7
7. Confidencialidad	7
8. Difusión	8

ANEXO 2. METODOLOGÍA

1. Introducción

La Encuesta Europea de Violencia de Género (EEVG), en terminología de EUROSTAT “European Survey on Gender Based Violence”, ha sido promovida por la Comisión Europea, específicamente por la DG JUST, debido a la necesidad de disponer de datos sobre violencia contra la mujer armonizados en Europa. Esta encuesta da respuesta a los requerimientos de información recogidos en el Convenio del Consejo de Europa sobre prevención y lucha contra la violencia contra la mujer y la violencia doméstica (Convenio de Estambul)¹ ratificado en España en el año 2014 y cuyos resultados ayudarán en la toma de decisiones políticas en este ámbito en Europa.

Antecedentes

La Agencia de Derechos Fundamentales (FRA) realizó una encuesta sobre violencia contra la mujer en 2012: “Violence against women: an EU Wide Survey”. La Encuesta Europea de violencia de género de EUROSTAT se diferencia de la encuesta llevada a cabo por la FRA en que se ha implementado por los propios países, siguiendo en la medida de lo posible las mismas directrices de Eurostat (utilizando el cuestionario y la metodología) y dentro del Sistema Estadístico Europeo² (SEE, integrado por Eurostat, los Institutos de Estadística nacionales y otras autoridades nacionales responsables en cada estado miembro del desarrollo, producción y difusión de las estadísticas europeas).

2. Objetivo de la encuesta

El objetivo de la encuesta es estimar la prevalencia de la violencia contra la mujer en España, mediante un cuestionario común que permita obtener datos armonizados en Europa. Se recaba información sobre la violencia contra la mujer que se produce en las relaciones de pareja, y la violencia fuera de este ámbito. Se obtiene información sobre la prevalencia de distintos tipos de violencia contra la mujer: psicológica, física, sexual en el caso de la violencia en la pareja, y violencia física y sexual en el caso de la violencia ejercida por otras personas. También se investiga la prevalencia del acoso sexual en el trabajo, el acoso reiterado y la violencia en la infancia.

Se estudian principalmente las siguientes características de la violencia: repetición y frecuencia. La encuesta permite calcular la prevalencia en toda la vida, del último año y los últimos 5 años.

Las variables de clasificación primordiales son: edad, nivel de formación, país de nacimiento, discapacidad (proxy) y grado de urbanización. Sobre su frecuencia, duración, gravedad (miedo, lesiones), consecuencias (psicológicas y físicas) y otras características. También se recoge información sobre la búsqueda de ayuda: denuncias, contacto con otros servicios de atención a víctimas o de apoyo (sanitarios, sociales, asociaciones, etc.).

¹ https://www.boe.es/diario_boe/txt.php?id=BOE-A-2014-5947

² <https://www.ine.es/dyngs/INE/index.htm?cid=542>

3. Ámbitos de la investigación

3.1. Ámbito geográfico

El ámbito geográfico de la investigación lo constituye todo el territorio español.

3.2. Ámbito poblacional

La encuesta en España proporciona información de las mujeres entre 16 y 74 años que residen en viviendas familiares en España.

3.3. Ámbito temporal

La encuesta se realiza en el año 2022.

Periodo de referencia de los resultados de la encuesta:

La recogida de datos se realiza en el año 2022.

Periodo de referencia de la información:

Se utilizan principalmente tres periodos temporales: toda la vida, los últimos 5 años, y los últimos 12 meses.

4. Diseño muestral

El Instituto Nacional de Estadística (INE) colabora con la DGVG en esta encuesta aportando el diseño de la muestra, y extrayéndola para su uso en la recogida de datos, gracias a un convenio de colaboración³ entre ambos organismos. También realiza el cálculo de los errores de muestreo y apoya con otros aspectos metodológicos generales de la encuesta. El INE diseña la muestra probabilística representativa de las mujeres residentes en España entre 16 y 74 años en hogares privados y proporciona a la DGVG la muestra inicial teórica para poder realizar las entrevistas y obtener una muestra efectiva de mujeres encuestadas.

Las características básicas del diseño de la muestra son las siguientes:

4.1. Tipo de muestreo

Se utiliza un muestreo bietápico estratificado, en el que las unidades de primera etapa son secciones censales, y las unidades de segunda etapa las mujeres de 16 a 74 años que residen en viviendas familiares principales. Se selecciona una muestra independiente en cada comunidad autónoma.

4.2. Marco de selección

Como marco para la selección de la muestra se utiliza el Marco de Direcciones Georreferenciadas (MDG) con año de referencia 2021. El Marco de Direcciones

³ Convenio de colaboración INE-DGVG: <https://www.boe.es/boe/dias/2019/04/25/pdfs/BOE-A-2019-6176.pdf>

Georreferenciadas (MDG) es un sistema de Información estructurado y jerarquizado compuesto por todas las entidades que forman parte del modelo territorial. El MDG aprovecha la información de viviendas que proviene del fichero precensal generado en el Censo de 2011. Esta información se complementa con la información sobre las personas que residen en cada vivienda extraída del fichero padronal. El MDG se actualiza trimestralmente, a partir de las variaciones padronales tanto de la población como del territorio.

4.3. Estratos

Las secciones censales se agrupan en estratos dentro de cada comunidad autónoma en función del grado de urbanización del municipio. Los estratos son los siguientes:

- Ciudades: Zonas densamente pobladas
- Pueblos y barrios periféricos: Zonas de densidad intermedia
- Áreas rurales: Zonas escasamente pobladas

4.4. Tamaño muestra

El INE calcula el tamaño necesario de la muestra teórica para que sea suficiente para satisfacer las necesidades de información y construir los indicadores apropiados.

Usando la encuesta piloto y otras encuestas similares para estimar la tasa de respuesta, se considera una muestra teórica de 14.760 mujeres repartidas en 984 secciones censales a lo largo de todo el territorio nacional. Este reparto corresponde a seleccionar 15 mujeres por sección censal.

La muestra de secciones se distribuye por estratos mediante una afijación de compromiso entre uniforme y proporcional.

Selección de la muestra

Unidades de primera etapa: se seleccionan en cada estrato con probabilidad proporcional al tamaño, medido por el número de habitantes.

Unidades de segunda etapa: se seleccionan dentro de cada unidad primaria con probabilidades iguales.

4.5. Estimadores

Se utilizan estimadores con factores de elevación deducidos de las probabilidades de selección del diseño. Estos factores se ajustan mediante técnicas de reponderación y calibrado para reducir el sesgo asociado a la falta de respuesta y recuperar los totales poblacionales a nivel nacional.

El estimador previo a la fase de calibrado es el siguiente:

$$\hat{Y} = \sum_{k \in S} d_k Y_k$$

donde:

$$d_k = \frac{P_h}{p_h}$$

Siendo P_h la población total del estrato h-ésimo y p_h es la población efectiva del estrato h-ésimo.

El estimador final viene dado por la expresión:

$$\hat{Y}_w = \sum_{k \in S} w_k Y_k$$

Donde los pesos w_k son los pesos calibrados.

El proceso de calibrado elegido incorpora las siguientes variables:

- Grupos de edad: [16, 24], [25, 34], [35,44], [45, 54], [55,64], [65, 69], [70, 74] en cada CAU.
- Total extranjeras en cada CAU.
- Total mujeres en Ceuta y total de mujeres en Melilla.

En los resultados de la encuesta se tiene en cuenta el número de observaciones de cada celda y dato a estimar para considerar si el tamaño muestral es adecuado para proporcionar datos de calidad. Teniendo en cuenta el diseño muestral, el porcentaje a estimar y el número de observaciones de cada celda, siguiendo criterios estadísticos y metodología habitual⁴, se marcan los datos en los que el tamaño muestral resulta escaso y no permite proporcionar una estimación totalmente fiable. Asimismo, en el caso en el que el tamaño muestral resulta insuficiente para proporcionar el dato, éste se elimina de las tablas publicadas.

4.6. Cálculo de errores de muestreo

Para el cálculo de los errores de muestreo de las principales características investigadas se utiliza el método indirecto Jackknife.

Este método se basa en la formación de submuestras en las que cada una de ellas se obtiene eliminando una unidad primaria de la muestra total. La estimación de la varianza del estimador tiene la expresión:

$$\hat{V}(\hat{X}) = \sum_h \frac{(n_h - 1)}{n_h} \sum_{j \in h} (\hat{X}_{(hj)} - \hat{X})^2$$

donde:

$\hat{X}_{(hj)}$ es la estimación de la característica X cuando se quita de la muestra la unidad primaria j del estrato h .

\hat{X} es la estimación de X obtenida con toda la muestra.

n_h es el número de unidades primarias en el estrato h .

⁴<https://ec.europa.eu/eurostat/documents/3859598/5927001/KS-RA-13-029-EN.PDF.pdf/a3155d11-4bf0-48d2-943d-2b1e9d096442?t=1414782791000>

5. Diseño del cuestionario

El cuestionario que se utiliza en esta encuesta está basado en el diseñado y propuesto por Eurostat. Es el resultado de grupos de trabajo⁵ coordinados por Eurostat y dedicados al desarrollo de la metodología y cuestionario de la encuesta, así como del análisis de los pre-test del cuestionario y de las encuestas piloto realizadas en muchos países.

En España se sigue fielmente la estructura de este cuestionario europeo⁶, eliminando algunas de las preguntas propuesta como opcionales debido a la longitud del cuestionario e incluyendo puntualmente alguna pregunta necesaria para los objetivos nacionales.

5.1. Idiomas del cuestionario

El cuestionario se prepara y traduce para su utilización en las encuestas en las distintas comunidades autónomas al español y a las lenguas cooficiales (Catalán, Euskera, Gallego y Valenciano). También está disponible en inglés en el caso de necesitar encuestar a mujeres extranjeras que no entiendan los idiomas oficiales en España.

6. Recogida de la información

El trabajo de campo se lleva a cabo en todo el territorio español, obteniendo muestra efectiva en todas las comunidades autónomas, y con una representación adecuada de las zonas de distinto grado de urbanización.

El período de recogida de información abarca los meses de marzo a junio de 2022. La muestra efectiva recoge información sobre 6465 mujeres.

6.1. Métodos de recogida

Se utilizaron dos métodos de recogida:

- CAWI (Computer Assisted Web Interview): los datos que proceden de este canal se recogen a través de un cuestionario electrónico disponible en Internet, por medio de una aplicación de recogida de datos.

Este canal se ofrece a todas las mujeres de la muestra antes de ser visitadas. Se proporciona el apoyo telefónico puntual (clarificación de conceptos en alguna pregunta, apoyo técnico) a la cumplimentación de la entrevista.

- CAPI (Computer Assisted Personal Interview): entrevista personal con ordenador portátil o tableta: las entrevistadoras disponen de dispositivos portátiles (tabletas y ordenadores portátiles) en los que se instala una aplicación de recogida en la se encuentra el cuestionario de la encuesta. Este se cumplimenta electrónicamente.

En primer lugar, se ofrece la posibilidad a la mujer de colaborar por CAWI, de no hacerse en los plazos establecidos se procede a intentar recoger la información vía CAPI.

⁵ “Eurostat task force on the development of a survey on gender-based violence” y “Task force for implementing the EU survey on gender-based violence against women and other forms of interpersonal violence”

⁶ Dicho cuestionario se encuentra incluido en el [manual metodológico de Eurostat](#).

También se consideró durante la recogida de datos la posibilidad de realizar entrevistas CATI (Computer Assisted Telephone Interview), debido a la alta incidencia del COVID en 2022 y como última opción. En la encuesta de 2022 no se utiliza.

6.2. Tratamiento de la información

La información se trata teniendo en cuenta los estándares habituales de calidad. La Encuesta Europea de Violencia de Género cuenta además con un marco de garantía de calidad para las estadísticas basado en el ESSCoP, el Código de Buenas prácticas de las Estadísticas Europeas de Eurostat. Se implementan una serie de medidas que contribuyen a garantizar la calidad del proceso y de los resultados: recogida de datos por ordenador, inspecciones, formación de calidad a los equipos de recogida de datos, revisión de los microdatos, tanto de recogida en campo (incidencias, etc.) como de cuestionarios (controles de inconsistencias, flujos, valores válidos, etc.).

6.2.1. Validación de datos

El proceso de supervisión de la información para garantizar su calidad se realiza en dos fases:

1. Durante el desarrollo de la entrevista CAPI/CAWI se incorporan en el cuestionario electrónico reglas de validación de los errores e inconsistencias que deben ser depurados en el mismo momento de la entrevista. Asimismo, se incorporan avisos sobre inconsistencias leves para ser cotejadas directamente con la entrevistada.

Los errores graves o valores inválidos, en caso de existir, se detectan en las primeras semanas del periodo de recogida de información y permiten perfeccionar el cuestionario electrónico, corrigiendo los errores en la programación.

Los errores que puedan deberse a una mala interpretación de preguntas, se transmiten a las entrevistadoras para que lo tengan en cuenta durante el desarrollo de su trabajo.

2. Trabajos de oficina: Una vez recibidos los datos se contrasta:

a) al recibir además de los microdatos del cuestionario, datos acerca de los trabajos de campo se comprueba que no exista una mala interpretación en la clasificación de las incidencias en campo, y que la recogida se esté realizando según las normas previstas (sustituciones, tiempos, número visitas, etc.).

b) sobre los microdatos del cuestionario: segunda revisión del flujo del cuestionario (ya se habrá comprobado en campo, pero en la primera semana de datos se pueden detectar aún errores), corrección de inconsistencias, valores extremos, chequeo de distribuciones no esperadas en alguna variable, errores de codificación, o problemas con variables auxiliares.

De forma complementaria se programan cruces o se extraen subconjuntos de datos para analizar variables concretas.

6.2.2. Compilación de datos

Entre los procesos aplicados a los datos iniciales hasta la obtención de los datos agregados que para difundir se encuentran:

1. Depuración - validación de los datos.
2. Construcción, a partir del fichero de microdatos de los cuestionarios, del fichero de microdatos de las variables solicitadas por EUROSTAT (transformaciones de algunas variables, construcción de ficheros según formato requerido).
3. Cálculo de estimadores para la publicación

La encuesta ha sido sometida, también por haber sido financiada por EUROSTAT, a un control exhaustivo, tanto de los procesos como de sus resultados.

7. Confidencialidad

La Ley 12/1989, de 9 de mayo, de la Función Estadística Pública obliga al Ministerio de Igualdad no difundir en ningún caso los datos personales cualquiera que sea su origen. Se entiende que son datos personales los referentes a personas físicas o jurídicas que o bien permitan la identificación inmediata de los interesados, o bien conduzcan por su estructura, contenido o grado de desagregación a la identificación indirecta de los mismos.

Tratamiento de datos confidenciales

El Ministerio de Igualdad adopta las medidas lógicas, físicas y administrativas necesarias para que la protección de los datos confidenciales sea efectiva, desde la recogida de datos hasta su publicación.

En los cuestionarios de las encuestas se incluye una cláusula legal que informa de la protección que ampara a los datos recogidos. En las fases de tratamiento de la información los datos que permiten la identificación directa solo se conservan mientras son estrictamente necesarios para garantizar la calidad de los procesos.

En la publicación de las tablas de resultados se analiza el detalle de la información para evitar que puedan deducirse datos confidenciales de las unidades estadísticas y respetar así el secreto estadístico.

Los ficheros de microdatos siempre están anonimizados, es decir se distribuyen sin información que pudiera permitir identificar ni localizar a las unidades estadísticas.

La recogida de datos de la encuesta la realizará una empresa, en su contratación se tiene en cuenta que respete, la política de confidencialidad del Ministerio, teniendo en cuenta lo establecido en la Ley 12/1989, de 9 de mayo, de la Función Estadística Pública, y lo dispuesto en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, respecto de los límites en el uso de la informática y otras técnicas y medios de tratamiento automatizado de los datos de carácter personal, para garantizar el honor, la intimidad personal y familiar de las personas físicas y el pleno ejercicio de sus derechos.

8. Difusión

La Delegación del Gobierno contra la Violencia de Género tiene previsto difundir los datos y metadatos de la Encuesta Europea de Violencia de Género en su página web. Además de este estudio que se pone a disposición del público en dicha web, se difundirán también los resultados más destacables en gráficos y tablas estadísticas acompañadas de breves notas descriptivas. También se harán públicos los resultados más importantes en nota de prensa.