

Uso de los microdatos de la Encuesta de Percepción Social sobre el apoyo a las energías renovables en España 2022 [edición 1.0]

Junio de 2023

Esta nota contiene información sobre los microdatos de la Encuesta de Percepción Social sobre el apoyo a las energías renovables en España de 2022 impulsada por la Fundación Española de la Ciencia y la Tecnología (FECYT). Los microdatos se hacen públicos con el fin de aportar transparencia a la investigación, así como proporcionar nuevas fuentes de datos a la comunidad científica.

Citar los datos:

Spanish Foundation for Science and Technology, Díaz Catalán, C., & Cabrera Álvarez, P. (2023). Encuesta de Percepción Social sobre el apoyo a las energías renovables en España. FECYT [Data set]. Fundación Española para la Ciencia y la Tecnología (FECYT). <https://doi.org/10.58121/7SCT-4010>

1. Contenido de la carpeta

TABLA 1. CONTENIDO DE LA CARPETA CON LOS DATOS DE LA ENCUESTA DE DESINFORMACIÓN

Archivo	Descripción
cuestionario_ev-2022_fecyt_2023.pdf	Documentación. Cuestionario completo de la encuesta en formato pdf.
nota_ev-2022_fecyt_2023.pdf	Documentación. Guía para el uso y análisis de los datos.
microdatos-epsct-2022_fecyt_2023.sav	Microdatos de la encuesta en formato SPSS.
microdatos-epsct-2022_fecyt_2023.dta	Microdatos de la encuesta en formato Stata.

2. Resumen del diseño de la encuesta

La encuesta, dirigida a la población general residente en España de 18 o más años, fue administrada mediante entrevista telefónica asistida por ordenador a una muestra final de 1.240 entrevistados. La muestra estuvo compuesta por una combinación de números de teléfonos de líneas fijas (30%) y móviles (70%) generados aleatoriamente. El trabajo de campo fue llevado a cabo por la empresa Celeste-Tel entre los días 15 de septiembre y 29 de septiembre de 2022, siendo la duración media de la entrevista de 15 minutos.

La selección de la muestra, estratificada por comunidad autónoma y tamaño de hábitat, se realizó en dos etapas. En la primera se generaron números de teléfono fijo y móviles de forma aleatoria; en el caso de los fijos, fueron estratificados por comunidad y tamaño de hábitat. En la segunda etapa, una vez que se había establecido contacto con un número de teléfono fijo correspondiente a un hogar, se procedió a seleccionar a la persona entrevistada empleando cuotas de sexo y edad. En el caso de las llamadas a teléfono móviles, la selección de la persona se realizó con cuotas de sexo y edad.

Una vez concluido el trabajo de campo, la muestra fue calibrada con el fin de generar un coeficiente de ponderación para corregir posibles desviaciones del perfil de la muestra derivados de la falta de cobertura de la población o de respuesta a la encuesta. La calibración es una técnica de ajuste que consiste en generar un coeficiente de ponderación que hace coincidir el perfil de la muestra con el de la población con respecto a una serie de variables que están presentes en los datos poblacionales y en la encuesta. Las variables empleadas en la calibración fueron una combinación de las variables sexo, edad (seis grupos: 15-24, 25-34, 35-44, 45-54, 55-64, 65+) y nivel educativo (cuatro grupos: primarios o menos, secundarios de primer ciclo, secundarios de segundo ciclo, superiores) y, por otro lado, comunidad autónoma. Los totales poblacionales fueron tomados de la EPA, una encuesta que por su diseño y tamaño muestral puede actuar como referente para la encuesta.

En la presente edición del estudio se ofrecen una serie de parados que aportan información complementaria sobre la recogida de los datos. Estos parados son la fecha y hora de inicio ($z1_1$ y $z2_2$) y final de las entrevistas ($z2_1$ y $z2_2$) y la duración de la entrevista en segundos ($z3$).

3. Recomendaciones para el análisis de los datos de la encuesta

El diseño de la encuesta incluye la estratificación de la muestra por comunidad autónoma y tamaño de hábitat y el uso de cuotas de sexo y edad para seleccionar a la persona entrevistada. Además, tras un análisis preliminar del perfil sociodemográfico de la muestra se detectó la sobrerrepresentación de los grupos con un mayor nivel educativo. Para corregir esta desviación se generó un coeficiente de ponderación que ajusta el perfil de la muestra con respecto a las variables sexo, grupos de edad, nivel educativo y comunidad autónoma.

A la hora de analizar los datos, estos elementos del diseño deben ser tenidos en cuenta ya que afectan tanto a las estimaciones (porcentajes, medias, coeficientes...) como a la precisión con la que se puede inferir de la muestra a la población (errores típicos). Para ello, en el fichero de datos se pueden encontrar la variable `peso`, que corresponde al coeficiente de ponderación, y la variable `estrato`, que refleja la estratificación de la muestra.

Por un lado, usar la ponderación (`peso`) en el análisis es necesario para mitigar el sesgo que puede introducir en la estimación la falta de respuesta a la encuesta y la selección introducida por el sistema de cuotas. Ignorar la ponderación en el análisis de los datos implica asumir que la muestra se diseñó usando un muestreo aleatorio simple y que se obtuvo una tasa de respuesta del 100%, dos supuestos que no se cumplen en este estudio. Como resultado de utilizar la ponderación, es muy

probable que la media, porcentajes o coeficientes del análisis sean diferentes a los que se obtienen al ejecutar el análisis sin ponderar. Además, el uso de la ponderación resta precisión a la estimación, lo que resulta en unos errores típicos más altos, que reflejan la incertidumbre que genera la falta de respuesta a la encuesta y la selección que produce el uso de las cuotas.

Por el otro, la inclusión de la estratificación (*estrato*) en el análisis de los datos afecta únicamente a los errores típicos. La estratificación, puede tener un efecto positivo, reduciendo los errores típicos, si las variables por las que se estratificó la muestra están relacionadas con la variable que se pretende estimar.

Una duda recurrente es la conveniencia de utilizar la ponderación cuando se analizan subpoblaciones, por ejemplo, las mujeres o aquellos que están más interesados temas científicos. En estos casos se debe utilizar la ponderación de la misma forma que se emplea en los análisis de la muestra completa. Por ejemplo, en el caso del análisis de las mujeres, la ponderación ajusta la composición de la muestra en términos de edad, nivel educativo y comunidad de residencia.

Los principales programas de análisis estadísticos incorporan comandos para tener en cuenta el diseño muestral a la hora de analizar los datos. Un amplio resumen de estas herramientas se puede encontrar en el artículo de West *et al.* (2018).

Referencias

West, B., Sakshaug, J. y Aurelien, G. (2018). Accounting for Complex Sampling in Survey Estimation: A Review of Current Software Tools. *Journal of Official Statistics*, 34(3) 721-752. <https://doi.org/10.2478/jos-2018-0034>