



Series largas de algunos agregados económicos y demográficos regionales: Actualización de *RegData* hasta 2023

(*RegData* y *RegData Dem* versión 7.0-2023)

ÁNGEL DE LA FUENTE

Estudios sobre la Economía Española 2025/01

Enero 2025

fedea

Las opiniones recogidas en este documento son las de sus autores y no coinciden necesariamente con las de Fedea.

Series largas de algunos agregados económicos y demográficos regionales: Actualización de *RegData* hasta 2023 (*RegData* y *RegData Dem* versión 7.0-2023)

Angel de la Fuente*
FEDEA e Instituto de Análisis Económico (CSIC)

Enero de 2025

Palabras clave: renta, empleo y población regional, series homogéneas
Clasificación JEL / JEL Classification: E01, R1

Key words: regional income, employment and population, homogeneous series, Spain

1. Introducción

En esta nota se describe brevemente la última actualización de *RegData*, una base de datos que recoge los principales agregados económicos y demográficos de las regiones españolas durante las últimas siete décadas. En su mayoría, las series comienzan en 1950 o 1955 y se extienden hasta 2023. En años recientes, la información que ofrece *RegData* coincide con las últimas series oficiales que publica el INE. Estas series se enlazan con series oficiales anteriores y con otras fuentes con el fin de obtener series enlazadas homogéneas que se extienden hacia atrás tanto como ha sido posible.

La actualización de este año es algo más complicada de lo habitual pues se realiza tras la revisión estadística de 2024, que vendría a ser un cambio de base en la terminología tradicional. Por tanto, las nuevas series de Contabilidad Regional de España (CRE) del INE, que comienzan en 2021 (de aquí en adelante la CRE21 para abreviar) han de enlazarse con las series correspondientes de la versión anterior de REGDATA (versión 6.3). Esta complicación no afecta al módulo demográfico de REGDATA, que se extienden un año más por el procedimiento habitual de incorporar los nuevos datos del INE sin modificación alguna.

2. La revisión estadística de 2024

A finales de 2024, el INE publicó las primeras series de CRE que incorporan los resultados de la revisión estadística de 2024.¹ Como los antiguos cambios de base, esta revisión incluye una actualización de las fuentes estadísticas y métodos de estimación empleados en la elaboración de la Contabilidad Nacional de España (CNE) y la CRE e incorpora diversas mejoras metodológicas.² Como resultado de todo ello, las nuevas series de la CRE21 difieren de las anteriores (la CRE16) en el año de referencia (2021 en este

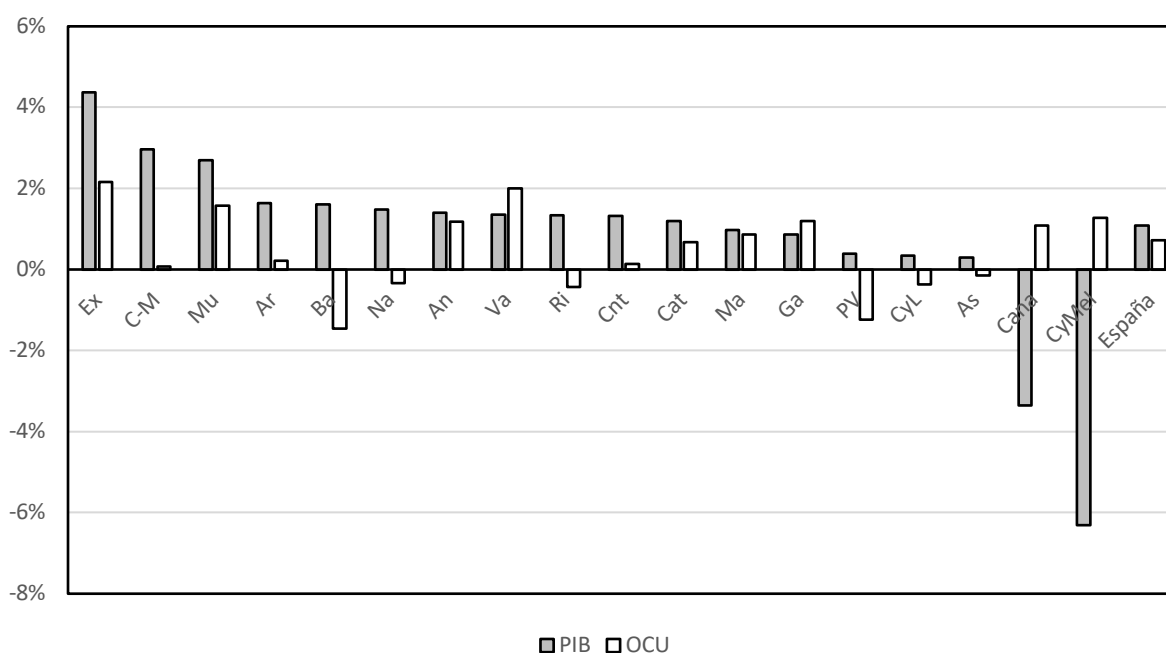
* Este trabajo forma parte de un proyecto de investigación cofinanciado por BBVA Research.

¹ Véase https://ine.es/prensa/nota_RE_2024_CN_09.pdf

² Entre otras novedades, se incorporan los datos de los nuevos censos de Población y Viviendas 2021, lo que entre otras cosas afecta a los factores de elevación de varias encuestas utilizadas en la CRE, así como la nueva información disponible tras la mejora de la cobertura de la Encuesta Estructural de Empresas por ramas de actividad. También se incorporan diversas mejoras técnicas en métodos y procedimientos de estimación, entre otras cosas para incorporar las recomendaciones de los grupos de trabajo del Sistema Estadístico Europeo. Para más detalles, véase INE (2024b).

caso) y es necesario realizar un enlace entre ellas. La revisión estadística de 2024 supone una leve revisión al alza de los grandes agregados nacionales de output y empleo, pero tiene efectos muy significativos a nivel regional en algunos casos. Como se aprecia en el Gráfico 1, las nuevas estimaciones agregadas de PIB y ocupados nacionales en 2021 están en torno a un punto porcentual por encima de las estimaciones anteriores de la CRE16. A nivel regional, sin embargo, las diferencias son mayores, superando los dos puntos porcentuales de PIB en cinco territorios, destacando los casos de Extremadura, con una revisión al alza de 4,4 puntos, y de Ceuta y Melilla, con una reducción de 6,3 puntos. Los ajustes a la baja del PIB en Canarias y Ceuta y Melilla se deben sobre todo a la utilización de un nuevo procedimiento de regionalización de los impuestos sobre los productos que tiene en cuenta la menor presión fiscal en este ámbito que se registra en estos territorios (INE, 2024c).

Gráfico 1: Serie nueva CRE21 – serie CRE16, diferencia porcentual entre las dos últimas series de PIB y empleo de la CRE



- *Clave:* An: Andalucía; Ar: Aragón; As: Asturias; Ba: Baleares; Cana: Canarias; CyL: Castilla y León; C-M: Castilla-la Mancha; Cat: Cataluña; CyMel: Ceuta y Melilla; Cnt: Cantabria; Va: Comunidad Valenciana; Ex: Extremadura; Ga: Galicia; Ma: Madrid; Mu: Murcia; Na: Navarra; Ri: Rioja; PV: País Vasco.

3. Procedimiento de enlace

El INE enlaza las nuevas series de la CRE21 con sus antecesoras para construir una “serie homogénea” 2000-23. El Instituto utiliza el procedimiento conocido como *retropolación*, consistente en extender hacia atrás las nuevas series utilizando las tasas de crecimiento de las antiguas. Manteniendo la metodología habitual en la construcción de *REGDATA*, en este trabajo el enlace entre la CRE21 y nuestras series anteriores se realizará utilizando el llamado *método mixto* desarrollado en de la Fuente (2014).

Esta metodología proporciona una vía intermedia entre los dos procedimientos más habituales en la literatura: la *interpolación* y la *retropolación*. El nombre del método hace referencia al hecho de que éste reparte a lo largo del tiempo la discrepancia entre las dos series que aflora en el momento del enlace de una forma que afecta tanto a las tasas de crecimiento de la serie más antigua como a su valor en el año

base. Con este método, la senda de la corrección a la serie más antigua que ha de introducirse para obtener la serie enlazada se obtiene a partir del supuesto de que el "error de medición" que presenta la primera de estas series es proporcional a un componente "mal medido" del producto o del empleo total cuya tasa bruta de crecimiento es un múltiplo constante de la tasa bruta de crecimiento del resto del agregado correspondiente.

Para describir con precisión el procedimiento, resulta necesario introducir algo de notación. Supongamos que en el momento T se produce alguna incidencia, tal como un cambio de base, que genera una discontinuidad en alguna variable de interés. Llamemos x_t a la serie original (no corregida) del agregado medida en logaritmos, y_t al logaritmo de la serie enlazada o corregida por la incidencia que queremos construir para $0 \leq t \leq T$ y d_t a la diferencia logarítmica entre ambas series. Interpretaremos d_t como el error de medición contenido en la serie original antes de la introducción del cambio metodológico e intentaremos adivinar su evolución en los años anteriores a T a partir del único valor conocido de esta magnitud -- el correspondiente al punto de enlace, d_T que se obtiene como la diferencia logarítmica entre las dos series a enlazar en ese punto. Con el método mixto, y_t se estima como

$$(1) \hat{y}_t = x_t + \hat{d}_t \quad \text{para } 0 \leq t \leq T \quad \text{con } \hat{d}_t = d_T \left(\frac{1}{2} \right)^{\frac{T-t}{H}}$$

donde H es la *vida media* del error de medición aflorado con el cambio metodológico, esto es, el tiempo necesario (yendo hacia atrás) para que el error observado en la serie más antigua en el momento T se reduzca a la mitad.

Para realizar el enlace, es necesario hacer algún supuesto sobre el valor del parámetro de vida media, H . Puesto que la vida media real no se conoce, el supuesto será inevitablemente arbitrario, pero seguramente más razonable que los supuestos implícitos en los métodos tradicionales (que suponen que el error ha existido siempre hacia atrás en el caso de la retropolación o que se origina en el año base de la serie antigua en el de la interpolación). Como se ha hecho habitualmente en la construcción de REGDATA cuando no existe información adicional que pueda ayudar a fijar ese parámetro, en lo que sigue se supondrá una vida media de 10 años, lo que implica un grado de persistencia de tal error que cabría considerar moderado.

Cuando x e y son magnitudes monetarias, se plantea también la cuestión de cómo enlazar sus valores "reales" (en volumen o a precios constantes del año base). La opción más natural consiste en expresar ambas magnitudes en precios de un mismo año y aplicar después la ecuación (1) para construir el enlace. Un procedimiento equivalente que en la práctica resulta algo más sencillo consiste en enlazar por retropolación los correspondientes índices de precios y, seguidamente, deflactar la serie nominal previamente enlazada utilizando esta serie de precios.

4. Construcción de algunas series auxiliares

Antes de proceder al enlace entre las dos fuentes, es necesario construir algunas variables que la CRE21 no proporciona directamente. Estas son, en particular, el número de puestos de trabajo y puestos de trabajo equivalentes a jornada completa (totales y asalariados) y las rentas totales del trabajo.

El número de puestos de trabajo equivalentes a jornada completa (totales y asalariados, $PTEJC$ Y $PTASEJC$) en cada región se aproxima dividiendo las horas totales trabajadas (por el total de los ocupados y por los asalariados respectivamente) en la región de acuerdo con la CRE21 por la duración media estimada de la correspondiente jornada completa a nivel nacional. Esta variable se obtiene como el cociente entre el número de horas trabajadas y el número de puestos de trabajo equivalentes (totales y asalariados) en el conjunto de España utilizando datos de las nuevas series de la CNE y la Contabilidad Trimestral (CNTR) con base 2021 que, a diferencia de la CRE21, sí proporcionan información sobre estas variables. El cálculo es sólo aproximado porque supone que la duración media de la jornada completa es igual en todas las regiones, lo que no tiene por qué ser cierto.

La serie de puestos de trabajo (PT) se construye a partir de la serie de ocupados (OCU) de la nueva base de la CRE, de las series nacionales de ambas variables que ofrecen la CNE y la CNTR y de la información regional que ofrece la EPA sobre el porcentaje de los ocupados que declaran una segunda actividad además de su ocupación principal ($\%SEGEMP$), elaborada a partir de los microdatos de la encuesta.

Calculamos en primer lugar el cociente entre el número de puestos de trabajo y el número de ocupados a nivel nacional con datos de la CNE21, obteniendo así un factor de conversión entre las dos variables a nivel agregado:

$$(2) RPTOC_t^{CNE} = \frac{PT_t^{CNE}}{OCU_t^{CNE}}$$

Seguidamente, dado que puestos de trabajo y ocupados cumplen la siguiente relación

$$(3) PT = (1 + \%SEGEMP) * OCU,$$

la información de la EPA sobre la fracción de los trabajadores que declaran una segunda ocupación, $\%SEGEMP$, puede utilizarse para modular al alza o a la baja en cada región el factor de conversión nacional que nos da la CNE21. En particular, estimamos el número de puestos de trabajo de la región r en el año t como el producto

$$(4) \widehat{PT}_{rt}^{CRE} = OCU_{rt}^{CRE} * RPTOC_t^{CNE} * \frac{1 + \%SEGEMP_{rt}^{EPA}}{1 + \%SEGEMP_t^{EPA}}$$

donde la ausencia del subíndice regional, r , indica que estamos trabajando con el promedio nacional de la variable de interés.³

Para el caso del empleo asalariado, se procede exactamente de la misma forma y utilizando las mismas fuentes, incluyendo la información regional de la EPA sobre el porcentaje de los asalariados que declaran una segunda actividad además de su ocupación principal ($\%SEGEMPAS$).

³ El dato de ocupados correspondiente a la extra-regio se multiplica sólo por $RPTOC$ puesto que la EPA no proporciona información sobre este territorio ficticio que permita afinar la estimación.

Finalmente, las rentas totales del trabajo (*RTL*) se construyen como sigue utilizando información de la CRE21. En primer lugar, se calcula el salario medio para el sector *s* en la región *r* dividiendo la remuneración total de los asalariados en el sector-región (*RAS*) por el número total de personas asalariadas (*AS*),

$$(5) w_{sr} = \frac{RAS_{sr}}{AS_{sr}}$$

Seguidamente, el salario medio de cada sector-región se multiplica por el número total de ocupados (asalariados y no asalariados) correspondiente (*OCU*) y los resultados se agregan sobre sectores para obtener una estimación de las *rentas totales del trabajo (RTL)* en la región

$$(6) RTL_r = \sum_s w_{sr} OCU_{sr}$$

5. Algunos detalles del enlace

Para cada una de las variables de interés, el enlace entre las dos fuentes se realiza en primer lugar a nivel nacional por el procedimiento descrito en la sección 3. Seguidamente, se enlazan las correspondientes series regionales por el mismo procedimiento, tratando a la *extra regio* como un territorio más. Las series regionales preliminares obtenidas de esta forma se ajustan todas en la misma proporción de forma que su suma coincida exactamente con el correspondiente agregado nacional. Esta corrección se aplica a las variables de empleo, a la *RAS* y *RTL*, al *VAB* y al *PIB* a precios corrientes y constantes. Finalmente, las series definitivas de deflatores del *VAB* y el *PIB* se obtienen como el cociente entre las series definitivas de *VAB (PIB)* nominal y real. El procedimiento supone dar prioridad a las series nacionales construidas directamente sobre las que se obtendrían a partir de la suma de los enlaces de las variables regionales. Se ha optado por hacer así las cosas porque en algunos casos se dispone de mejor información a nivel nacional que regional.

Como se ha dicho más arriba, los deflatores del *VAB* y el *PIB* se enlazan por retroproyección y las series enlazadas de deflatores se aplican a las correspondientes series enlazadas de producción a precios corrientes para obtener series a precios constantes (aproximados) de 2021. Puesto que la CRE ya no publica deflatores, estos se aproximan a partir de los correspondientes índices de volumen (que se renormalizan a 100 en 2021 antes de realizar los cálculos pertinentes).⁴ Para ello, se construyen en primer lugar series aproximadas de *PIB* "a precios constantes" de 2021 multiplicando los índices de volumen del *PIB* por el *PIB* del año base medido a precios corrientes. Los deflatores aproximados se calculan seguidamente como el cociente entre el *PIB* a precios corrientes y el *PIB* aproximado a precios constantes del año base. La serie de *PIB* a precios constantes obtenida a partir de esta serie preliminar de deflatores se ajusta en la misma proporción en todas las regiones de forma que su suma sobre comunidades autónomas (y la extra-regio) coincida con la serie nacional construida directamente por el mismo procedimiento. Finalmente, la serie final de deflatores se obtiene dividiendo la serie final de *PIB* a precios corrientes por la serie final de *PIB* a precios constantes. El procedimiento es el mismo en el caso del *VAB*.

⁴ Rompiendo con la práctica histórica habitual, los índices de volumen de la CRE se normalizan a 100 no en el año base de cada revisión sino en el año anterior a éste.

6. Variables disponibles

Los Cuadros 1 y 2 describen respectivamente las variables disponibles en el módulo central de *RegData* y en su módulo demográfico, *RegData Dem*. Los datos correspondientes se adjuntan en dos hojas de Excel. Una descripción detallada de la metodología y los datos originales de base que se han utilizado para construir las series de REGDATA pueden encontrarse en los trabajos citados en las referencias, que están disponibles en las webs de FEDEA (<http://www.fedea.net>) y de BBVA Research (<https://www.bbva.com/>).

El módulo central de *RegData* contiene datos de renta, precios, empleo y paro y rentas del trabajo. Se incluyen datos de PIB y VAB a precios corrientes y constantes de 2021 y a precios medios del período 1955-2023. El empleo total y asalariado se mide en personas, puestos de trabajo, horas y puestos de trabajo equivalentes a jornada completa. Las dos últimas variables comienzan en 1977 y 1981 respectivamente por falta de la información necesaria para construirlas en años anteriores. Finalmente, se ofrecen datos de remuneración de asalariados y salarios medios, así como una serie de rentas totales del trabajo, que se aproximan imputando a cada trabajador no asalariado unas rentas salariales iguales a la remuneración media percibida por los asalariados de su misma región y sector.

En el módulo demográfico se ofrecen datos de población residente y su desglose por grupos de edad, nacimientos, defunciones y saldos migratorios netos. También se construye una serie de población residente corregida por migraciones no declaradas que intenta aproximar la población que realmente reside de forma habitual en cada territorio.

Cuadro 1: Contenido del módulo principal de *RegData*

VAB nom:	VAB nominal = valor añadido bruto a precios básicos, medido a precios corrientes (miles de euros corrientes)	1955-2023
VAB real:	VAB real = valor añadido bruto a precios básicos, medido a precios constantes (miles de euros de 2021)	1955-2023
Pvab21:	deflactor del VAB, año 2021 = 1	1955-2023
Pvabmed:	deflactor del VAB a precios medios del período 1955-2023	1955-2023
VAB pmed	VAB real valorado a precios medios del período 1955-2023	1955-2023
PIB nom:	PIB nominal = producto interior bruto a precios de mercado, medido a precios corrientes (miles de euros corrientes)	1955-2023
PIB real:	PIB real = producto interior bruto a precios de mercado, medido a precios constantes (miles de euros de 2021)	1955-2023
Ppib10:	deflactor del PIB, año 2021 = 1	1955-2023
Ppibmed:	deflactor del PIB a precios medios del período 1955-2023	1955-2023
PIB pmed	PIB real valorado a precios medios	1955-2023
POBd:	población residente o de derecho	1955-2024
PARADOS:	parados, miles de personas, promedio de los cuatro trimestres	1955-2023
OCU:	ocupados, miles de personas	1955-2023
AS:	asalariados, miles de personas	1955-2023
PT:	puestos de trabajo, en miles	1955-2023

Cuadro 1: Contenido del módulo principal de RegData-- continuación

<i>PTAS</i> : puestos de trabajo asalariados, en miles	1955-2023
<i>H</i> : horas efectivas anuales trabajadas por los ocupados, en millones	1977-2023
<i>HAS</i> : horas efectivas anuales trabajadas por los asalariados, millones	1977-2023
<i>PTEJC</i> : puestos de trabajo equivalentes a jornada completa, en miles	1981-2023
<i>PTASEJC</i> : puestos de trabajo asalariados equivalentes a jornada completa, en miles	1981-2023
<i>RAS</i> remuneración de asalariados, miles de euros corrientes	1955-2023
<i>RTL</i> rentas totales del trabajo, miles de euros corrientes	1955-2023
<i>PARTL</i> : participación corregida del trabajo en el VAB = RTL/VAB	1955-2023
<i>w</i> : salario medio = $RAS/PTAS$	1955-2023
<i>rml</i> : remuneración media del trabajo = RTL/PT	1955-2023

Cuadro 2: Contenido del módulo demográfico, RegData Dem

<i>NAC(t)</i> = nacimientos durante el año que comienza el 1 de julio de t	1950-2023
<i>DEF(t)</i> = defunciones durante el año que comienza el 1 de julio de t	1950-2023
<i>SVEG(t)</i> = saldo vegetativo del año que comienza el 1 de julio de $t = NAC(t) - DEF(t)$.	1950-2023
<i>POBd(t)</i> = población de derecho o residente a 1 de julio de t.	1950-2024
<i>SMIG1(t)</i> = saldo (in-)migratorio neto consistente con la serie de población de derecho = $POBd(t+1) - POBd(t) - SVEG(t)$	1950-2023
<i>CORR_MIGND</i> = Ajuste a la población de derecho por migraciones no declaradas	1950-2023
<i>POBRES_CORR</i> = Población residente corregida por migraciones no declaradas = $POBd + CORR_MIGND$	1950-2023
<i>SMIG2</i> = saldo (in-)migratorio neto consistente con la serie de población residente corregida = $POBRES_CORR(t+1) - POBRES_CORR(t) - SVEG(t)$	1950-2023
<i>wPOBd_0-14</i> = peso de la población de 0 a 14 años en la población total de derecho	1950-2024
<i>wPOBd_65+</i> = peso de la población de 65 o más años en la población total de derecho	1950-2024
<i>wPOBd_0-15</i> = peso de la población de 0 a 15 años en la población total de derecho	1950-2024
<i>POBd_16+</i> = población de derecho de 16 o más años de edad	1950-2024
<i>PET</i> = Población en edad de trabajar, 16-64	1950-2024

(*) La fuente de datos de población residente pasa a ser la Estadística Continua de Población (INE, 2023).

Referencias

- de la Fuente, A. (2016a). "Series enlazadas de Contabilidad Regional para España, 1980-2014. (RegData_80-14, v4.3)." FEDEA, Estudios sobre economía española no. 2016-35, Madrid.
<http://www.fedea.net/datos-economia-regional-y-urbana/>
- de la Fuente, A. (2016b). "Series enlazadas de PIB y otros agregados de Contabilidad Nacional para España, 1955-2014. (RegDat_Nac versión 4.2)." FEDEA, Estudios sobre Economía Española no. 2016-02, Madrid.
<http://www.fedea.net/datos-economia-regional-y-urbana/>
- de la Fuente, A. (2016c). "Series largas de algunos agregados demográficos regionales, 1950-2015. RegData-Dem versión 5.0." FEDEA, Estudios de Economía Española no 2016-14, Madrid.
<http://www.fedea.net/datos-economia-regional-y-urbana/>
- de la Fuente, A. (2017). "Series enlazadas de algunos agregados económicos regionales, 1955-2014. Parte I: Metodología, VAB, PIB y puestos de trabajo (RegData_55-14, Versión 5.0-parte I)." FEDEA, Estudios sobre Economía Española no. 2017-08, Madrid.
<http://www.fedea.net/datos-economia-regional-y-urbana/>
- de la Fuente, A. (2017). "Series enlazadas de algunos agregados económicos regionales, 1955-2014. Parte II: Otras variables de empleo, rentas del trabajo y paro. RegDat Versión 5.0." FEDEA, Estudios sobre la Economía Española no. 2017-
<http://www.fedea.net/datos-economia-regional-y-urbana/>
- de la Fuente, A. (2024). "Series largas de algunos agregados económicos y demográficos regionales: Actualización de RegData hasta 2022. (RegData y RegData Dem versión 6.3-2022)." FEDEA, Estudios sobre la Economía Española no. 2024-06. Madrid.
<https://documentos.fedea.net/pubs/eee/2024/eee2024-06.pdf>
- Instituto Nacional de Estadística (INE, 2024a). Revisión Estadística 2024 de las Cuentas Nacionales: estimación definitiva de impacto sobre el PIB para el año de referencia 2021. Nota informativa.
https://ine.es/prensa/nota_RE_2024_CN_09.pdf
- Instituto Nacional de Estadística (INE, 2024b). Revisión Estadística 2024 de las operaciones de Contabilidad Nacional del INE. Proyecto técnico.
https://ine.es/proyectos_mejora/registros_eee/proyecto_tecnico_CNRE.pdf
- Instituto Nacional de Estadística (2024c). Nota de prensa. Contabilidad Regional de España Producto Interior Bruto regional. Serie 2000-2023. <https://ine.es/dyngs/Prensa/es/CRE2023.htm>
- Instituto Nacional de Estadística (INE, 2024d). Contabilidad Regional de España. Resultados. Serie contable 2021-23.
https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736167628&menu=resultados&idp=1254735576581#
- Instituto Nacional de Estadística (INE, 2024e). Contabilidad Nacional Anual de España. Principales agregados.
https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177057&menu=ultiDatos&idp=1254735576581
- Instituto Nacional de Estadística (INE, 2024f). Contabilidad Nacional Trimestral de España. Principales agregados.
https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736164439&menu=resultados&idp=1254735576581
- Instituto Nacional de Estadística (INE, 2024g). Encuesta de población activa.
https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=resultados&idp=1254735976595#!tabs-1254736195129

INE (2024h). Estadística continua de población.

https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177095&menu=resultados&idp=1254735572981

Véase también:

<https://www.bbvaresearch.com/category/tipos-de-documento/documento-de-trabajo/>