

# **Coyuntura Demográfica de Empresas (CODEM)**

## **Metodología y líneas operativas**

Instituto Nacional de Estadística

Enero de 2021



## **Perspectiva general; Objetivos y elementos de partida**

La estadística experimental de *Coyuntura Demográfica de Empresas* nace de la necesidad de información de mayor frecuencia, especialmente durante la crisis generada por la pandemia COVID-19, sobre los fenómenos vinculados al dinamismo empresarial, así como su caracterización mediante variables de clasificación, especialmente las vinculadas a los expedientes de regulación temporal de empleo (ERTES), actividad económica, tamaño o características sociodemográficas para los trabajadores autónomos.

Esta operación complementa el papel del sistema DIRCE, como elemento generador de información sobre el tejido empresarial en España. Hasta ahora, los principales análisis estadísticos en este ámbito han tenido un carácter estructural, considerando tanto la perspectiva estática (punto fijo en el tiempo) como dinámica (flujos anuales). Estas nuevas iniciativas se centran en la producción de indicadores a corto plazo, de periodicidad trimestral.

El ciclo de producción del DIRCE continuará desarrollándose con periodicidad anual, generándose el sistema de marcos interconectados de unidades estadísticas. Con este modelo de trabajo, se garantiza el cumplimiento de los compromisos Reglamentarios a nivel europeo y las necesidades de los usuarios nacionales sobre marcos de muestreo. El DIRCE seguirá siendo la base para la generación de todas las operaciones vinculadas al dominio *Business Demography*, de acuerdo al actual Reglamento SBS y al futuro EBS.

Ahora se trata de aprovechar la experiencia adquirida, para dar respuesta a las demandas crecientes sobre dinamismo empresarial a corto plazo. Aunque no exista consistencia plena con los indicadores tradicionalmente obtenidos sobre Demografía Armonizada de Empresas, el propósito es que contribuyan a aportar valor a la base estadística actualmente disponible, especialmente en el contexto que estamos viviendo.

La operación se centra en dos poblaciones de interés: una dedicada a empresas empleadoras y otra dedicada a autónomos, incluyendo diversas series de datos sobre los stocks en cada trimestre y los flujos ocurridos entre trimestres sucesivos. La cobertura de actividades económicas es idéntica a la utilizada para la Explotación Estadística del DIRCE. Quedan por tanto excluidas las actividades agrícolas, ganaderas y pesqueras, los servicios administrativos de la Administración Central, Autonómica y Local, las actividades de los hogares que emplean personal doméstico y los Organismos Extraterritoriales.

# CODEM\_1.- Demografía trimestral de empresas empleadoras

## 1.1 Características

Esta iniciativa aborda una aproximación a la metodología internacional sobre *Demografía de Empleadores*, pero incorporando la componente del corto plazo. Se centra en la población de Unidades Legales con empleo desarrollando un rango de actividades consistente con la cobertura publicada en la Explotación Estadística del DIRCE.

Dado que a nivel internacional, hay establecido un plazo de 18 meses para la producción nacional de los indicadores referidos a la anterior metodología, se aplicará obviamente una versión simplificada de dicha metodología para la producción estos indicadores.

## 1.2 Fuentes básicas de información

Se utilizan principalmente dos fuentes:

- **Registro de Cuentas de Cotización de la Seguridad Social.** Se considera el conjunto de Cuentas con al menos un trabajador de alta a una fecha de referencia determinada. Incluye información relativa a los siguientes Regímenes de la SS:

- ✓ Régimen General
- ✓ Régimen especial del mar
- ✓ Régimen especial de minería del carbón
- ✓ Conciertos asistencia sanitaria

El INE recibe habitualmente esta fuente en diferentes puntos del tiempo a lo largo de un año natural. Para este proyecto, se han utilizado cuatro bases de datos mensuales conteniendo información de unidades con trabajadores a 31 Diciembre, 31 de marzo, 30 de junio y 30 de septiembre. Cada Cuenta de Cotización contiene datos identificativos de la Unidad Legal responsable.

- **Unidades Legales empleadoras con asalariados acogidos a ERTES.** Dada la situación actual, esta fuente aporta información relevante para caracterizar los flujos demográficos que serán objeto de estudio. El fichero básico recibido del SEPE se usa para múltiples operaciones estadísticas con diferentes propósitos y contiene información personal a nivel de asalariado. Para este proyecto, la población objetivo estará conformada por *el conjunto de unidades legales que, tomando las mismas fechas de la fuente anterior, tenían al menos un asalariado acogido a ERTE.*

## 1.3 Tratamientos y outputs

Ninguna de las fuentes anteriores se usa de forma directa. En línea con la experiencia adquirida para el ciclo de producción del DIRCE, se ha diseñado un programa específico de tratamientos y mejora de la calidad del dato primario.

### 1.3.1 Registro de Cuentas de Cotización

Los principales tratamientos se centran en un análisis macro, mejoras micro de variables nucleares y transformación de la unidad elemental.

Análisis macro.- Eliminación de Cuentas con cero trabajadores, ajuste de la cobertura por actividad y generación de agregados por número de trabajadores. Se realiza un análisis comparativo con periodos precedentes.

Mejoras micro de variables nucleares.- Especialmente dirigidas al identificador NIF (control de validez, incorporación de identificadores válidos con base a información DIRCE, normalización, cambios de forma legal, etc). Contraste de códigos CNAE contra diccionario e incorporación de códigos válidos con base a información DIRCE. Contraste de códigos geográficos contra Nomenclátor INE e incorporación de códigos válidos mediante apoyo de variables complementarias (código postal o incluso literal de dirección postal).

Transformación.- Ejecutados los procesos anteriores, se genera una base de datos, conteniendo información sobre Unidades Legales Empleadoras (relaciones n:1). Las variables asociadas a esta nueva unidad se generan por procedimientos ya experimentados en el DIRCE. Para cada referencia temporal considerada, se produce *información individualizada sobre Unidades Legales con trabajadores al final del periodo*.

### 1.3.2 Unidades Legales con trabajadores acogidos a ERTES

Los tratamientos se centran en mejoras para el identificador NIF, replicando los procesos de la fuente anterior e incluyendo un control de duplicados.

### 1.3.3 Output: Base dinámica sobre Unidades Legales empleadoras

Se genera una base de datos sobre unidades empleadoras con información longitudinal relativa a los diferentes trimestres observados. Se mantiene por integración de la información retrospectiva disponible y los nuevos ficheros recibidos, una vez finaliza su programa de mejora. Es la referencia para la extracción de los indicadores y agregados identificados en el plan de difusión.

## 1.4 Eventos demográficos

Se ha tomado como punto de arranque el 1 de Enero de 2020. A esa fecha de referencia se genera la población de Unidades Legales empleadoras. Con el paso de los sucesivos trimestres se identifican las siguientes categorías:

- **Altas de unidades legales empleadoras en el trimestre.-** Conjunto de Unidades Legales que no tenían empleo al comienzo del trimestre (no figuraban en la población de Unidades Legales empleadoras), han contratado trabajadores y siguen con empleo al final del trimestre.
- **Permanencias de unidades legales empleadoras en el trimestre.-** Unidades Legales con empleo al comienzo y final del trimestre.
- **Reactivaciones de unidades legales empleadoras en el trimestre.-** Unidades Legales con asalariados al comienzo del trimestre anterior, pero sin asalariados al comienzo del trimestre actual. A lo largo del mismo, han contratado trabajado-

res y siguen con empleo al final del trimestre. Este evento se comienza a identificar a partir del segundo trimestre de 2020.

- **Bajas de unidades legales empleadoras en el trimestre.-** Unidades Legales empleadoras al comienzo del trimestre, pero que han perdido su empleo a largo del mismo.
- **Cohorte de unidades legales empleadoras.-** Conjunto de unidades con empleo a una fecha determinada. Será utilizada como población de partida para su seguimiento temporal y detección de supervivencias. Por ejemplo, la cohorte a 1 Julio 2020 es el conjunto de unidades con empleo al final del segundo trimestre y con independencia de su estatus demográfico (Altas, Permanencias o Reactivaciones).
- **Supervivencias de unidades legales empleadoras.-** Cada cohorte se monitoriza en el tiempo. Las unidades legales que persisten con empleo al final de cada sucesivo trimestre se consideran supervivencias en ese trimestre.

Los sucesos de Alta y Baja no significan que necesariamente la unidad legal se haya constituido / disuelto o haya estado involucrada en algún proceso implicando una redistribución de factores de producción (fusión, absorción, etc). El factor *empleo* es el determinante en la identificación del flujo demográfico.

## 1.5 Variables de Clasificación

Las poblaciones que conforman los diferentes eventos demográficos se desglosarán de acuerdo a las siguientes variables

- **ERTE.-** Variable dicotómica asignada a nivel microdato con base a la segunda fuente del proyecto. La modalidad "Con ERTE" se asigna si la unidad tiene al menos un empleado en esa situación. "Sin ERTE" se asigna a la unidad en caso contrario.
- **Actividad principal.-** De acuerdo a la Clasificación Nacional de Actividades Económicas (CNAE 2009). En el proceso de transformación a *Unidad Legal*, se implementa una jerarquización de códigos CNAE, basada en el empleo registrado en sus Cuentas de cotización. La actividad principal de la Unidad Legal es la que registra mayor número de empleados.
  - **Estrato de Asalariados.-** En el mismo proceso de transformación, se agrupa el empleo registrado en todas las Cuentas de cotización de cada Unidad Legal. Se han adoptado las siguientes modalidades: 1 a 5, 6 a 9, 10 a 99, 100 a 249, 250 y más.
- **Códigos geográficos.-** Para cada Unidad Legal, se toma la ubicación de la Cuenta de cotización con mayor número de trabajadores. El plan de difusión contempla diversos grados de detalle, incluido el nivel municipal.

## 1.6 Plan de difusión

Los resultados generados para cada trimestre se estructuran en cuatro series

- **Serie 1: Cuadros comparativos de eventos demográficos.** Permite analizar la evolución de las categorías demográficas de un trimestre base a lo largo del siguiente trimestre. Para cada año natural se generan cuatro tablas, desglosadas a nivel nacional y por comunidades autónomas.

- **Serie 2: Supervivencias de cohortes de unidades empleadoras.** Para cada cohorte identificada, se efectúa un seguimiento en los siguientes cuatro trimestres, identificándose las supervivencias (agregado y tasa) en cada periodo. Alternativamente se puede analizar la rapidez de la pérdida de efectivos con el paso del tiempo y los factores que contribuyen a este fenómeno. Las supervivencias se desglosan territorialmente a nivel nacional, por comunidades autónomas y provincias. A su vez, para cada ámbito geográfico los agregados se clasifican según dos criterios, combinando las variables de clasificación:

- ERTE y actividad principal (División CNAE)
- ERTE y estrato de asalariados

- **Serie 3: Dinamismo de las actividades económicas.** Se toma como punto de partida el conjunto total de unidades al inicio del año distribuido por actividad. Para cada División CNAE se calculan los flujos (Altas + Reactivaciones, Permanencias y Bajas) ocurridos en los sucesivos trimestres del año. Esta información se desglosa a nivel nacional y por provincias.

- **Serie 4: Indicadores municipales.** Esta información se proporciona con soporte cartográfico. Se han generado las siguientes magnitudes.

- **Tasa neta de crecimiento.** Cociente donde el numerador es *Altas + Reactivaciones – Bajas* y el denominador es el Stock total de unidades, *Altas + Reactivaciones + Permanencias + Bajas*
- **Tasa de Altas.** Cociente donde el numerador es *Altas* y el denominador es el *Stock*
- **Tasa de Bajas.** Cociente donde el numerador es *Bajas* y el denominador es el *Stock*
- **Tasa de unidades con ERTE.** Cociente donde el numerador es *Total de unidades con ERTE* y el denominador se restringe al conjunto de unidades activas *Altas + Reactivaciones + Permanencias*

## CODEM\_2.- Demografía trimestral de trabajadores autónomos

### 2.1 Características

Esta iniciativa aborda el análisis dinámico trimestral de las personas físicas registradas en el Registro de Trabajadores Autónomos. La población objetivo es el conjunto total de Afiliados a los distintos regímenes por cuenta propia de la Seguridad Social: Régimen Especial de Trabajadores Autónomos (RETA) y Mar.

La población considerada se entiende en sentido amplio y abarca todas las personas físicas que legalmente están afiliados en los citados regímenes y desarrollan un rango

de actividades consistente con la cobertura publicada en la Explotación Estadística del DIRCE.

## 2.2 Fuentes básicas de información

Se utiliza el **Registro de Trabajadores Autónomos de la Seguridad Social (RETA)**, que el INE recibe de forma habitual en diferentes puntos del tiempo a lo largo de un año natural. Para este proyecto, se han utilizado cuatro bases de datos mensuales conteniendo información sobre personas físicas de alta a 31 Diciembre, 31 de marzo, 30 de junio y 30 de septiembre. Desde el punto de vista temporal, esta fuente es totalmente consistente con la utilizada en el primer proyecto.

Otra fuente utilizada es el **Padrón de habitantes**, registro administrativo donde constan los vecinos de cada municipio con datos identificativos. Esta base de datos está disponible en el INE y es la fuente de la que se obtiene la información sociodemográfica del trabajador autónomo.

## 2.3 Tratamientos y outputs

En línea con la experiencia adquirida para el ciclo de producción del DIRCE, los principales tratamientos diseñados para el registro RETA se centran en análisis macro y mejoras micro de variables nucleares.

Análisis macro.- Eliminación de registros que no estaban dados de alta al final del periodo de referencia, ajuste de la cobertura por actividad, generación de agregados por actividad y análisis comparativo con periodos precedentes.

Mejoras micro de variables nucleares.- Especialmente dirigidos al identificador NIF - DNI (normalización y control de validez). Contraste de códigos CNAE contra diccionario e incorporación de códigos válidos con base a información DIRCE. Contraste de códigos geográficos contra Nomenclátor INE e incorporación de códigos válidos mediante apoyo de variables complementarias (código postal o incluso literal de dirección postal).

Para cada fichero mensual tratado en los cuatro puntos del año, se produce *información individualizada sobre trabajadores autónomos de alta al final del periodo*.

El output generado tras aplicar los tratamientos anteriores es una base de datos dinámica sobre trabajadores autónomos con información longitudinal relativa a los diferentes trimestres observados. Se mantiene por integración de la información retrospectiva disponible y los nuevos ficheros recibidos, una vez finaliza su programa de mejora. Es la referencia para la extracción de los indicadores y agregados identificados en el plan de difusión.

## 2.4 Eventos demográficos

Se establece como punto de arranque de este proyecto el 1 de Enero de 2020. A esa fecha de referencia se genera la población de trabajadores autónomos dados de Alta. Con el paso de los sucesivos trimestres se identifican las siguientes categorías:

- **Altas de autónomos en el trimestre.-** Personas físicas no detectadas al comienzo del trimestre y dados de Alta en RETA al final del trimestre.
- **Permanencias de autónomos en el trimestre.-** Personas físicas detectadas en RETA al comienzo y final del trimestre.
- **Reactivaciones de autónomos en el trimestre.-** Personas físicas detectadas en RETA al comienzo del trimestre anterior, no detectadas al comienzo del trimestre actual pero nuevamente detectadas al final del trimestre. Este evento se comienza a identificar a partir del segundo trimestre de 2020.
- **Bajas de autónomos en el trimestre.-** Personas físicas detectadas en RETA al comienzo del trimestre, pero que se dan de baja a largo del mismo.
- **Cohorte de autónomos.-** Conjunto de personas físicas dadas de Alta en RETA, utilizado como población de partida para su seguimiento temporal y detección de supervivencias. Por ejemplo, la cohorte a 1 Julio 2020 es el conjunto de personas de Alta al final del primer trimestre y con independencia de su estatus demográfico (Altas, Permanencias o Reactivaciones).
- **Supervivencias de autónomos.-** Cada cohorte se monitoriza en el tiempo. Las personas que persisten de Alta en RETA al final de cada sucesivo trimestre se consideran supervivencias en ese trimestre.

## 2.5 Variables de clasificación

Las poblaciones que conforman los diferentes eventos demográficos se desglosarán de acuerdo a las siguientes variables

- **Actividad principal.-** De acuerdo a la Clasificación Nacional de Actividades Económicas (CNAE 2009).
- **Sexo.-** Incorporada a partir de la información de Padrón.
- **Intervalo de edad.-** Idéntico procedimiento, considerando las siguientes modalidades: Menor de 30 años, 30 a 39, 40 a 49, 50 a 59, 60 y más.
- **Códigos geográficos.-** Tomados del domicilio que consta en el propio registro RETA. El plan de difusión contempla diversos grados de detalle, incluido el nivel municipal.

## 2.6 Plan de difusión

Los resultados generados para cada trimestre se estructuran en cuatro series

- **Serie 1: Cuadros comparativos de eventos demográficos.** Permite analizar la evolución de las categorías demográficas de un trimestre base a lo largo del siguiente trimestre. Para cada año natural se generan cuatro tablas, desglosadas a nivel nacional y por comunidades autónomas.
- **Serie 2: Supervivencias de cohortes de autónomos.** Para cada cohorte identificada en un trimestre base, se efectúa un seguimiento en los siguientes cuatro trimestres, identificándose las supervivencias (agregado y tasa) en cada periodo. Alternativamente



se puede analizar la rapidez de la pérdida de efectivos con el paso del tiempo y los factores que contribuyen a este fenómeno. Las supervivencias se desglosan territorialmente a nivel nacional, por comunidades autónomas y provincias. A su vez, para cada ámbito geográfico los agregados se clasifican según dos criterios, combinando las variables de clasificación:

- Actividad principal (División CNAE) y sexo
- Intervalo de edad y sexo

• **Serie 3: Dinamismo de las actividades económicas.** Se toma como punto de partida el conjunto total de autónomos al inicio del año distribuido por actividad. Para cada División CNAE se calculan los flujos (Altas + Reactivaciones, Permanencias y Bajas) ocurridos en los sucesivos trimestres del año. Esta información se desglosa a nivel nacional y por provincias.

• **Serie 4: Indicadores municipales.** Esta información se proporciona con soporte cartográfico. Se han generado las siguientes magnitudes.

- **Tasa neta de crecimiento.** Cociente donde el numerador es *Altas + Reactivaciones – Bajas* y el denominador es el Stock total de autónomos, *Altas + Reactivaciones + Permanencias + Bajas*
- **Tasa de Altas.** Cociente donde el numerador es *Altas* y el denominador es el *Stock*
- **Tasa de Bajas.** Cociente donde el numerador es *Bajas* y el denominador es el *Stock*