

# Distribución geográfica de los establecimientos productivos

## Versión exploratoria

### Metodología

#### 1. Introducción

El Centro de Estudios para la Producción (CEP XXI) –dependiente de la Secretaría de Industria y Desarrollo Productivo del Ministerio de Economía– y la Subsecretaría de Planificación, Estudios y Estadísticas del Ministerio de Trabajo, Empleo y Seguridad Social ponen a disposición una base de datos de establecimientos productivos a lo largo del país. La base, que presenta un promedio de la información reportada durante 2021 según los registros administrativos del Sistema Integrado Previsional Argentino (SIPA) y de la AFIP, contiene datos referidos a: la ubicación geográfica -exacta o aproximada- de 706.498 establecimientos productivos registrados, la cantidad de trabajadores asignados a ese establecimiento productivo -agrupada en rangos-, el sector de actividad al que pertenecen, y su jurisdicción.

El proyecto es exploratorio y, por lo tanto, la información considerada debe tomarse como provisoria y sujeta a futuras modificaciones.

#### 2. Consideraciones generales

Los datos recabados del SIPA provienen de la declaración jurada que cada empleador realiza, mediante el formulario N° 931 de la AFIP, sobre la nómina de personal en relación de dependencia. Dado que dichas declaraciones pueden ocurrir con cierto rezago, la información presentada puede variar levemente ante actualizaciones de los datos.

La información presentada refleja la situación del empleo asalariado registrado en las empresas entre enero y diciembre de 2021, por lo que no forman parte de la base las relaciones laborales no registradas y las diversas formas de trabajo por cuenta propia, sea registrado o no. A su vez, dado que los trabajadores de casas particulares cuentan con un régimen especial, este tampoco se encuentra incluido en la información provista.

Los establecimientos productivos presentados forman parte del sector privado y de empresas con participación estatal mayoritaria (como ocurre, por ejemplo, entre otras, en las ramas de energía, minería, transporte, y en algunas industrias). Se excluyó del proceso a los trabajadores públicos que no trabajan en empresas públicas.

La apertura sectorial que se utiliza es la actividad principal declarada por cada empresa empleadora en el formulario N°883, según la última versión del Clasificador de Actividades Económicas (CLAE) de AFIP. Se incluyen todas las ramas de actividad de la economía en tres desagregaciones: a nivel de seis dígitos, de dos dígitos y a nivel de letra.

Si bien la mayoría de las empresas utiliza el CLAE 2010 (asociado al formulario N°883), algunas continúan informando su actividad según versiones anteriores del clasificador. En estos casos, se optó por trasladar las clasificaciones a su última versión y se suprimieron de la base aquellos CLAE que no pudieron ser actualizados. Asimismo, a cada empresa se le asignó la última actividad declarada como principal en AFIP.

### 3. Georreferenciación de la información

El registro de empleados en la AFIP se acompaña por un formulario de Alta Temprana –Simplificación Registral– en el que se detalla el domicilio de explotación laboral en la que se enmarca cada una de las relaciones laborales. A partir de la dirección provista por el sistema se utilizaron diferentes herramientas de georreferenciación, a fin de establecer las coordenadas geográficas de los establecimientos productivos.<sup>1</sup>

Dado que la georreferenciación consiste en un proceso de normalización y comprensión de la información provista en los registros, aquella información incompleta o mal cargada fue georreferenciada de forma parcial, siendo una aproximación del domicilio real. A su vez, ante los casos en los cuales la información del domicilio estaba ausente o se indicaba la pertenencia a una zona rural, se realizó una estimación respetando las localidades y departamentos informados.

Asimismo, en algunos casos (sobre todo en relaciones laborales de más antiguas) la información provista por Simplificación Registral (un sistema más reciente en el tiempo) no representa la totalidad de las relaciones laborales existentes en el período de análisis. En esos casos, y a fin de completar la base de datos, se realizaron tareas de imputación, en función de la distribución geográfica de las relaciones laborales, las características de la empresa empleadora y el domicilio de la persona trabajadora.

Finalmente, hay ramas de actividad que por definición se encuentran parcialmente deslocalizadas, por ejemplo los servicios de seguridad y de transporte. En estos casos se optó por respetar la información provista por las empresas empleadoras, referida a oficinas centrales.

### 4. Información provista

La base de datos consiste en una serie de variables que permiten realizar una radiografía de la distribución geográfica del empleo asalariado formal y los establecimientos productivos durante el período enero-diciembre 2021. Para ello se cuenta con las siguientes variables:

- ID: indicador único por establecimiento productivo;
- lat/lon: coordenadas geográficas (latitud y longitud);
- in\_departamentos: código de departamento del Instituto Geográfico Nacional;

---

<sup>1</sup> En la georreferenciación se utilizó principalmente el sistema de georreferenciación de Google y se complementó con información provista por los sistemas de Open Street Map (OSM) y ArcGIS que cuentan con APIs públicas y gratuitas.

- departamento: nombre del departamento en el que se radica el establecimiento productivo;
- provincia\_id: código de la provincia del Instituto Geográfico Nacional;
- provincia: nombre de la provincia en la que se radica la sucursal;
- empleo: promedio de trabajadores que realizaron actividades en el establecimiento productivo en cuestión agrupados en rangos (1-9, 10-49, 50-199, 200-499, 500+ trabajadores);
- clae6: sector de actividad a seis dígitos en base al Clasificador Nacional de Actividades Económicas (CLAE);
- clae2: sector de actividad a dos dígitos en base al Clasificador Nacional de Actividades Económicas (CLAE);
- letra: sector de actividad a nivel de letra en base al Clasificador Nacional de Actividades Económicas (CLAE);
- Tipo\_coordenada: indica si la georreferenciación provista es exacta o una aproximación realizada

Las coordenadas geográficas provistas están redondeadas a dos o a tres dígitos en función de la cantidad de establecimientos productivos del departamento. Para los primeros cien departamentos la información georreferenciada está a tres decimales (lo que equivale a un margen de error de alrededor de 100 metros), mientras que para el resto lo está a dos (lo que equivale a alrededor de 1km de margen de error).

## 5. Base auxiliar

A fin de reducir el tamaño de la base de datos y facilitar la experiencia del usuario, las variables vinculadas a las ramas de actividad se presentan únicamente en forma de código. Para añadir la descripción a los procesamientos realizados se deberá recurrir al archivo auxiliar **actividades.csv**.

## 6. Coordenadas geográficas

Las coordenadas corresponden al Sistema Geodésico Mundial 1984 (WGS84 y EPSG 4326). Por lo tanto, en caso de realizar una representación geográfica se recomienda mantener el mismo sistema, de forma de evitar inconvenientes en la asignación de los puntos.