

Distribución geográfica de los establecimientos productivos, versión exploratoria: metodología

Diciembre 2023



Ministerio de Economía
Argentina

Secretaría de Industria
y Desarrollo Productivo

CEPXXI
CENTRO DE ESTUDIOS
PARA LA PRODUCCIÓN



INTRODUCCIÓN

El Centro de Estudios para la Producción (CEP XXI) 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 está compuesta por la ubicación geográfica aproximada a tres dígitos de más de los establecimientos productivos registrados, la cantidad de trabajadores asignados a ese establecimiento productivo –agrupada en rangos–, el sector de actividad al que pertenecen, la proporción de trabajadoras mujeres en la fina, el nivel de exportaciones –agrupada en quintiles– y su jurisdicción. La información proviene de los registros administrativos del Sistema Integrado Previsional Argentino (SIPA) y de registros de AFIP, excepto en el caso de las exportaciones que provienen del sistema registral de la Aduana. La base presenta un promedio de la información reportada durante 2021 y 2022, siendo la información referida al nivel de exportaciones, la proporción de mujeres y el tamaño dependiente del año en cuestión.

Frente a la versión publicada en 2022, la actual cuenta con una serie de mejoras: indica el carácter exportador de la empresa; permite analizar la cantidad de sucursales por empresa, ya que la anonimización se realizó a nivel de CUIT y no de establecimiento; permite la comparación interanual de los datos, de forma tal de contrastar 2021 y 2022.

Dados los cambios detallados en el quinto punto de este documento, la versión actual del Mapa no es estrictamente comparable con la versión publicada en 2022 (datos pertenecientes a 2021).

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



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 y 2022, 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 es incluido en la información provista.

Los establecimientos productivos presentados forman parte del sector privado y de empresas con participación estatal mayoritaria (como por ejemplo ocurre en ramas como energía, minería, transporte, algunas industrias, entre otras). Se excluyó del proceso a los empleadores públicos que no sean 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, 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. En caso de que una empresa utilizase el clasificador de actividades previo, se le imputó una actividad nueva sobre la base de la [tabla de conversión utilizada por INDEC](#). De este modo, la información sectorial provista refleja la rama de actividad de la empresa en la que se inscribe la relación laboral.



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.¹

¹ En la georreferenciación se utilizaron principalmente los sistemas de Open Street Map y ArcGIS que cuentan con APIs públicas y gratuitas, así como también se complementa con información provista por el sistema de georreferenciación de Google.



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 donde 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 larga antigüedad) 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, se realizaron tareas de imputación, en función de la distribución geográfica de las relaciones laborales, del domicilio de la persona trabajadora y de las características de la empresa empleadora, de forma tal de completar la base de datos.

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.

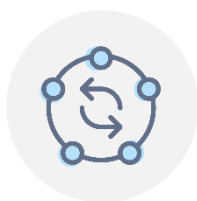
Información provista

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 y 2022. Para ello se cuenta con las siguientes variables:

- cuit: anonimización del CUIT de la empresa a la que pertenece el establecimiento;
- sucursal: indicador único por sucursal de cada cuit. El par cuit-sucursal conforma un establecimiento;
- anio: año al que refiere la información del establecimiento en cuestión;
- lat/lon: coordenadas geográficas (latitud y longitud);
- 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);
- in_departamentos: código de departamento del Instituto Geográfico Nacional;
- 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;
- quintil: quintil de exportaciones de bienes en el que se ubica la empresa según el nivel de exportaciones del año en cuestión. Consta de seis valores: 0 para aquellas empresas que no hayan tenido exportaciones, los valores del 1 al 5 indican el nivel de exportaciones de la empresa a la que pertenece el establecimiento en cuestión. 1 refiere al menor nivel de exportaciones (según el valor exportado) y 5 el mayor nivel.
- 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);
- proporcion_mujeres: indica la cantidad de mujeres que trabajaron en el período analizado en esa sucursal sobre la totalidad de empleados

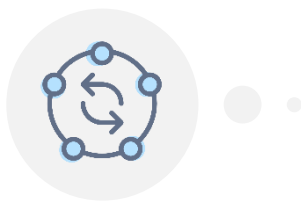


DIFERENCIAS METODOLÓGICAS CON SU VERSIÓN ANTERIOR

La presente versión del Mapa Productivo Laboral no resulta comparable con la versión de 2022, dada una serie de modificaciones metodológicas tendientes a mejorar la información provista.

Dado que el registro de sucursales es una herramienta relativamente nueva en ciertos casos las empresas incurrir en errores de registro, aduciendo la existencia de más de una sucursal laboral para el mismo espacio de trabajo. Con el objetivo de evitar la duplicación de sucursales se redondearon las coordenadas geográficas a tres dígitos, de forma tal de unificar todos los casos en los que una firma tuviera más de una sucursal en un radio menor a 111 metros. Esto se debe a que los procesos de georreferenciación no brindan la misma coordenada en caso de haber mínimas diferencias en la información cargada, por lo cual si una misma sucursal estuviera cargada más de una vez con un pequeño cambio en su dirección, la coordenada alcanzada no sería la misma y no podrían unificarse -excepto mediante el redondeo de la coordenada-.

La actual versión del mapa no contempla las sucursales que registraron menos de un trabajador en seis meses, de forma tal de evitar el cómputo de establecimientos temporales.



En este sentido, si bien gran parte de la información publicada es consistente con su versión anterior, deben evitarse comparaciones, ya que los procesos de georreferenciación en algunos casos cambiaron, así como también se modificaron algunas sucursales que se detectaron en la primera versión como incorrectamente georreferenciadas.



COMPARABILIDAD INTERANUAL

A fin de garantizar la comparabilidad de los datos entre 2021 y 2022, en esta nueva edición del Mapa Productivo-Laboral se volvieron a georreferenciar los datos de la versión anterior siguiendo las modificaciones metodológicas utilizadas para la versión actual. De esta manera, se garantiza que los métodos, imputaciones y aproximaciones utilizadas en los dos años tienen el mismo grado de error y se garantiza la comparabilidad entre los dos años disponibles.

A fin de posibilitar la comparación interanual de los datos se subió a Datos Abiertos el dataset “Datos de establecimientos por departamento, actividad y género” donde se pueden observar la cantidad de establecimientos, puestos de trabajo y empresas exportadoras con una apertura por género, sector de actividad (clae6) y año.

Este dataset incluye las siguientes variables:

- **anio:** año al que refiere la información del establecimiento en cuestión;
- **in_departamentos:** código de departamento del Instituto Geográfico Nacional;
- **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;
- **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);
- **genero:** desagregación según el sexo biológico de los trabajadores;

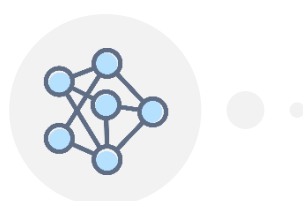


- Empleo: cantidad de trabajadores en el departamento para el nivel de desagregación y el sexo biológico mencionado;
- Establecimientos: cantidad de establecimientos en el departamento para el nivel de desagregación y el sexo biológico mencionado;
- empresas_exportadoras: cantidad de empresas exportadoras en el departamento para el nivel de desagregación y el sexo biológico mencionado



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**.



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.



Ministerio de Economía
Argentina

Secretaría de Industria
y Desarrollo Productivo

CEPXXI CENTRO DE ESTUDIOS
PARA LA PRODUCCIÓN