

**PROPUESTA DE IMPLEMENTACIÓN**

**“Estandarización de los nombres de las**

**variables del Marco Maestro de Empresas”**

**Versión: julio año 2023**

**Entrega N°1**

**INSTITUTO NACIONAL DE ESTADÍSTICAS**

**Julio/ 2023**

**DEPARTAMENTO DE METODOLOGÍAS E INNOVACIÓN ESTADÍSTICA**

**SUBDEPARTAMENTO DE DISEÑO DE MARCOS Y MUESTRAS**

**UNIDAD DE MUESTRAS ECONÓMICAS**

Estandarización de los nombres de las variables del Marco Maestro de Empresas.

Versión 2023

Entrega N°1.

Instituto Nacional de Estadísticas, Chile.

31 de julio de 2023.

ÍNDICE

[INTRODUCCIÓN 5](#_Toc141701072)

[I. OBJETIVOS 6](#_Toc141701073)

[I.1. Objetivo general 6](#_Toc141701074)

[I.2. Objetivos específicos 6](#_Toc141701075)

[II. ANTECEDENTES 7](#_Toc141701076)

[III. DESCRIPCIÓN DE LOS MARCOS MAESTROS DE EMPRESAS 8](#_Toc141701077)

[III.1. Marco maestro de empresas 8](#_Toc141701078)

[III.1.1. Descripción general 8](#_Toc141701079)

[IV. ESTANDARIZACIÓN DE VARIABLES 13](#_Toc141701080)

[V. PROBLEMÁTICA 15](#_Toc141701081)

[VI. PROPUESTA 16](#_Toc141701082)

DICCIONARIO DE ACRONIMOS

Para una mejor comprensión acerca de los términos mencionados en el documento, se presenta las definiciones para los acrónimos y siglas utilizadas.

**Acrónimo Descripción**

INE : Instituto Nacional de Estadísticas

MM : Marcos Maestros

MME : Marco Maestro de Empresas

SEN : Sistema Estadístico Nacional

SII : Servicio de Impuestos Internos

SDMM : Subdepartamento de Diseño de Marcos y Muestras

UL : Unidad legal(es)

RRAA : Registros Administrativos

# INTRODUCCIÓN

El Instituto Nacional de Estadísticas (INE) es el organismo estatal encargado de producir y difundir las estadísticas oficiales de Chile. Su misión es proporcionar información confiable y accesible a los usuarios para respaldar la toma de decisiones. Además, el INE se encarga de coordinar el Sistema Estadístico Nacional (SEN), asegurando que el país cuente con información pertinente, precisa, confiable, oportuna, accesible, clara, coherente y comparable a nivel nacional e internacional.

El Marco Maestro de Empresas (MME) es un listado de unidades legales que realizan actividades económicas en el país. Estas unidades son caracterizadas por diversos atributos, como su identificación, ubicación geográfica, estratificación por ventas anuales, número de trabajadores y medios de contacto.

La administración de la base de datos del MME es responsabilidad del Subdepartamento de Diseño de Marcos y Muestras (SDMM). La construcción y actualización del MME son operaciones estadísticas fundamentales para comprender las unidades legales formalmente constituidas que desarrollan actividades económicas en el territorio nacional. Para esto, se utilizan diversas fuentes de datos consideradas insumos, proporcionadas por otras instituciones en forma de registros administrativos.

En la elaboración del MME, es importante evidenciar la procedencia de los datos. Actualmente, algunas variables contienen información sobre la composición del nombre y el año del insumo al que pertenece el microdato, y esta procedencia puede variar. Sin embargo, las variables del MME tienen asociado el nombre a un año específico del insumo que actualiza el microdato, esto es incorrecto, por lo cual, se muestra la necesidad de normalizar el producto final (MME).

El objetivo de este documento es presentar una propuesta de estandarización de nombres de variables del MME para facilitar la creación de una automatización. Es por esto por lo que, se utilizará un marco maestro de empresas genérico, con año de entrega AAAA y en ámbito real y ficticio.

# OBJETIVOS

## Objetivo general

Realizar un proceso de estandarización del nombre de las variables en el Marco Maestro de Empresas (MME), con el fin de mejorar la difusión del producto estadístico y la utilización de este en cualquier software.

## Objetivos específicos

Estos objetivos específicos permiten abordar de manera clara y concisa cada uno de los aspectos que se quieren lograr en el estudio de estandarización del nombre de las variables en el Marco Maestro de Empresas (MME).

1. Presentar las variables incluidas en el MME.
2. Identificar y exponer los inconvenientes en la estandarización de nombre de las variables en el MME.
3. Desarrollar una propuesta de estandarización de nombres para las variables del MME, con el objetivo de mejorar y facilitar el uso del producto estadístico en cualquier software.

# ANTECEDENTES

El Marco Maestro de Empresas (MME) es una base de datos que se genera para difundirse como un registro estadístico que contiene información sobre las unidades legales (UL) que realizan actividades económicas dentro del territorio nacional durante un período de tiempo específico. Estas unidades son identificadas mediante un Rol Único Tributario (RUT), obtenido de registros administrativos y encuestas del INE. Luego de depurar los datos, se asignan éstos a las variables de cada una de las dimensiones del MME.

Algunos nombres de variables de ambos ámbitos poseen un año asociado al cual pertenece el insumo, siendo este uno de los grandes impedimentos para realizar una automatización óptima para el marco maestro de empresas.

En el contexto de la propuesta de estandarizar el nombre de las variables para el marco maestro de empresas (MME), es que este documento sugiere realizar este proceso en el producto final (MME) corrigiendo esa situación.

Es común que la información en una base de datos cambie y se actualice periódicamente, pero eso se registra en un documento o reporte de la base de datos al momento de difundir el producto final en lugar de incluirlo en los nombres de las variables. Es importante asegurar una estructura de base de datos adecuada para optimizar la automatización del producto final, y lograr un análisis más prolijo y exhaustivo.

Por lo tanto, en general, los nombres de variables en una base de datos deben ser descriptivos y representar la información que contienen, sin necesidad de incluir un año asociado en su nombre.

A continuación, se expone el estado actual de las variables del MME en todos sus ámbitos.

# DESCRIPCIÓN DE LOS MARCOS MAESTROS DE EMPRESAS

A continuación, este capítulo mostrará una descripción general del MME, las variables contenidas y descripción en ámbito real y ficticio (diccionario de variables actual del marco maestro de empresas).

## Descripción general

El marco maestro de empresas en su versión contable del año AAAA en ámbito real y ficticio contiene las siguientes variables y descripciones:

Tabla 1. Descriptor de campos del Marco Maestro de Empresas, ámbito real.

|  |  |
| --- | --- |
| ***Nombre de la variable*** | ***Descripción*** |
| RUT\_REAL | Rol único tributario otorgado por la administración fiscal, que identifica a la unidad legal en forma oficial. |
| DV\_RUT | Dígito verificador del RUT |
| ID\_GRUPO\_EMPRESAS\_INE | Identificador(es) de cada Grupo según Grupos de Empresas INE. |
| ID\_GRUPO\_EMPRESAS\_CMF | Identificador de cada grupo según la Comisión del Mercado Financiero (CMF). |
| NOM\_GRUPO\_EMPRESAS\_INE | Nombre del grupo de empresas identificado por INE. |
| RAZON\_SOCIAL | Nombre o denominación de la unidad legal que desarrolla una actividad económica, sea persona natural o jurídica. |
| NOM\_REPRESENTANTE\_LEGAL | Nombre del representante legal de la empresa. |
| NOMBRE\_FANTASIA | Nombre o denominación comercial de la unidad legal que desarrolla una actividad económica, sea persona natural o jurídica. |
| REGION\_CUT\_2018 | Código de 2 dígitos correspondiente a la Región, de acuerdo con la división político-administrativa oficial (RPC), vigente desde el 06 de septiembre 2018. |
| PROVINCIA\_CUT\_2018 | Código de 3 dígitos correspondiente a la Provincia, de acuerdo con la división político-administrativa oficial (RPC), vigente desde el 06 de septiembre 2018. |
| COMUNA\_CUT\_2018 | Código de 5 dígitos correspondiente a la Comuna, de acuerdo con la división político-administrativa (RPC), vigente desde el 06 de septiembre 2018. |
| DIRECCION | Ubicación Geográfica asociada a la unidad local principal. |
| INICIO\_DE\_ACTIVIDADES | Fecha de inicio de actividades según el SII. |
| SINDICO\_DE\_QUIEBRAS | Marca de la nómina de síndico de quiebras. |
| SECCION\_CIIU4CL | Primer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (letra). En caso de no tener actividad asignable, toman el valor 0. |
| DIVISION\_CIIU4CL | Segundo nivel de clasificación económica CIIU Rev.4. Adaptación nacional (2 dígitos). En caso de no tener actividad asignable, toman el valor 00. |
| GRUPO\_CIIU4CL | Tercer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (3 dígitos). En caso de no tener actividad asignable, toman el valor 000. |
| CLASE\_CIIU4CL | Cuarto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (4 dígitos). En caso de no tener actividad asignable, toman el valor 0000. |
| SUBCLASE\_CIIU4CL | Quinto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (5 dígitos). En caso de no tener actividad asignable, toman el valor 00000. |
| ESTRATOS\_F29 | Estrato calculado según ventas anuales en UF obtenidas de la selección de algunos códigos del F29. |
| FONO | Teléfono principal de la empresa. |
| CORREO\_ELECTRONICO | Correo electrónico principal de la Unidad Legal. |
| PAGINA\_WEB | Página WEB principal de la Unidad Legal. |
| DIRECTORIO\_PRIORIZADO | Marca que indica 1 si cumple con ser Registros de tamaño grande (estrato 9 al 12) o registros coincidentes con la producción estadística o Grupos de empresas INE, y 0 para los no coincidentes. |
| GRANDES\_CONTRIBUYENTES | Marca de grandes contribuyentes, 1 para los grandes contribuyentes y 0 para los que no lo son. |
| REGIMEN\_TRIBUTARIO | Marca de la nómina de Régimen Tributario. |
| CATEGORIA\_TRIBUTARIA | Marca que indica con un 1 si la unidad legal es de primera categoría tributaria, 2 si es de segunda categoría, 3 si presenta ambas categorías y 0 si no tiene categoría asociada. |
| TIPO\_CONTRIBUYENTE | Tipo de sociedad constituida por el contribuyente según el SII. |
| SUBTIPO\_CONTRIBUYENTE | Mayor desagregación del tipo de sociedad declarada por el contribuyente según SII. |
| COMERCIO\_EXTERIOR | Marca de las unidades legales presentes en BBDD de Aduana. Solo exporta (1), solo importa (2) o ambas (3). |
| MONEDA\_EXTRANJERA | Marca de las unidades legales que declaran formularios del SII en moneda extranjera, 1 para los coincidentes y 0 para los no coincidentes. |
| AFECTO\_A\_IVA | Empresas tiene la obligación de para el Impuesto al Valor Agregado, 1 para afecto a IVA y 0 para los que no. |
| PERS\_NAT\_CEPAL | Marca de las unidades legales según CEPAL,2013. "COMPENDIO DE RECOMENDACIONES TÉCNICAS Y BUENAS PRÁCTICAS PARA EL DESARROLLO DE LOS DIRECTORIOS DE EMPRESAS Y ESTABLECIMIENTOS ", pág.6. |
| TAMAÑO\_OCUPADOS\_AAAA | Tamaño de empresa según ocupados, regidos por la Ley N°20.416 del Estatuto Pyme. |
| TIPO\_DE\_REGISTRO | Característica del registro de acuerdo con su aparición en la ejecución del Directorio anterior como registro continuo (0) o registro nuevo (1). |
| PERS\_NAT\_SII | Marca de las unidades legales según criterio de corte de Rut menor a cincuenta millones (SII). |
| TERMINO\_GIRO\_AAAA | Marca de la nómina de registros que posee término de giro ante el SII. |
| ID | Identificador correlativo de uso interno. |
| NORMALIZACION | Marca que indica con un SI la dirección geográfica fue normalizada, NO si no lo fue y NA si no presenta dirección. |
| PEE\_AAAA | Marca que indica con un 1 si el registro fue levantado en la producción estadística estructural del año AAAA y con un 0 si no lo fue. |
| PEC\_AAAA | Marca que indica con un 1 si el registro fue levantado en la producción estadística coyuntural del año AAAA y con un 0 si no lo fue. |
| PRODUCCION\_ESTADISTICA | Marca que indica con un 1 si pertenece a la producción estadística estructural, 2 si pertenece a la producción coyuntural, 3 si pertenecen a ambas producciones estadísticas INE y, 0 si no pertenecen a los casos anteriores. |

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Tabla 2. Descriptor de campos del Marco Maestro de Empresas, ámbito ficticio.

|  |  |
| --- | --- |
| ***Nombre de la variable*** | ***Descripción*** |
| RUT\_FICTICIO | RUT ficticio asociado a una empresa. |
| REGION\_CUT\_2018 | Código de 2 dígitos correspondiente a la Región, de acuerdo con la división político-administrativa oficial (RPC), vigente desde el 06 de septiembre 2018. |
| PROVINCIA\_CUT\_2018 | Código de 3 dígitos correspondiente a la Provincia, de acuerdo con la división político-administrativa oficial (RPC), vigente desde el 06 de septiembre 2018. |
| COMUNA\_CUT\_2018 | Código de 5 dígitos correspondiente a la Comuna, de acuerdo con la división político-administrativa (RPC), vigente desde el 06 de septiembre 2018. |
| SECCION\_CIIU4CL | Primer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (letra). En caso de no tener actividad asignable, toman el valor 0. |
| DIVISION\_CIIU4CL | Segundo nivel de clasificación económica CIIU Rev.4. Adaptación nacional (2 dígitos). En caso de no tener actividad asignable, toman el valor 00. |
| GRUPO\_CIIU4CL | Tercer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (3 dígitos). En caso de no tener actividad asignable, toman el valor 000. |
| CLASE\_CIIU4CL | Cuarto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (4 dígitos). En caso de no tener actividad asignable, toman el valor 0000. |
| SUBCLASE\_CIIU4CL | Quinto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (5 dígitos). En caso de no tener actividad asignable, toman el valor 00000. |
| VTA\_$\_ENERO | Ventas del mes - 1 según F29. |
| VTA\_$\_FEBRERO | Ventas del mes - 2 según F29. |
| VTA\_$\_MARZO | Ventas del mes - 3 según F29. |
| VTA\_$\_ABRIL | Ventas del mes - 4 según F29. |
| VTA\_$\_MAYO | Ventas del mes - 5 según F29. |
| VTA\_$\_JUNIO | Ventas del mes - 6 según F29. |
| VTA\_$\_JULIO | Ventas del mes - 7 según F29. |
| VTA\_$\_AGOSTO | Ventas del mes - 8 según F29. |
| VTA\_$\_SEPTIEMBRE | Ventas del mes - 9 según F29. |
| VTA\_$\_OCTUBRE | Ventas del mes - 10 según F29. |
| VTA\_$\_NOVIEMBRE | Ventas del mes - 11 según F29. |
| VTA\_$\_DICIEMBRE | Ventas del mes - 12 según F29. |
| VTA\_ANUAL\_$\_F29\_AAAA | Suma de ventas mensuales según F29. |
| VTA\_ANUAL\_UF\_F29\_AAAA | Ventas anuales en UF según F29. |
| ESTRATOS\_F29 | Estrato calculado según ventas anuales en UF obtenidas de la selección de algunos códigos del F29. |
| TAMAÑO\_F29 | Tamaño de la unidad legal obtenido según el monto de la venta anual calculado desde el Formulario 29 en UF (micro, pequeña, mediana, grande). |
| CUENTA\_OCUPADOS\_AAAA | Total de trabajadores según F1887. |
| TAMAÑO\_OCUPADOS\_AAAA | Tamaño de empresa según ocupados, regidos por la Ley N°20.416 del Estatuto Pyme. |
| PROM\_TRAB\_OCUPADOS\_AAAA | Promedio anual de trabajadores según F1887. |
| PROM\_TRAB\_HONORARIO\_AAAA | Promedio anual de trabajadores a honorarios según F1879. |
| SINDICO\_DE\_QUIEBRAS | Marca de síndico de quiebras, 1 para los coincidentes y 0 para los que no lo son. |
| TERMINO\_GIRO\_AAAA | Marca de la nómina de registros que posee término de giro ante el SII. |
| GRUPO\_EMPRESAS\_INE | Marca de las unidades que pertenecen a algún grupo de empresas según levantamientos de encuestas INE. 1 para los coincidentes y 0 para los que no lo son. |
| MONEDA\_EXTRANJERA | Marca de moneda extranjera, 1 para los coincidentes y 0 para los que no lo son. |
| GRANDES\_CONTRIBUYENTES | Marca de grandes contribuyentes, 1 para los coincidentes y 0 para los que no lo son. |
| GRUPO\_EMPRESAS\_CMF | Marca de las empresas que pertenecen a un grupo empresarial según la CMF, 1 para los coincidentes y 0 para los que no lo son. |
| PRODUCCION\_ESTADISTICA | Marca que indica con un 1 si pertenece a la producción estadística estructural, 2 si pertenece a la producción coyuntural, 3 si pertenecen a ambas producciones estadísticas INE y, 0 si no pertenecen a los casos anteriores. |
| CATEGORIA\_TRIBUTARIA | Marca que indica con un 1 si la unidad legal es de primera categoría tributaria, 2 si es de segunda categoría, 3 si presenta ambas categorías y 0 si no tiene categoría asociada. |
| DIRECTORIO\_PRIORIZADO | Marca de directorio priorizado, 1 para los coincidentes y 0 para los que no lo son. |
| REGIMEN\_TRIBUTARIO | Marca de la nómina de Régimen Tributario. |
| COMERCIO\_EXTERIOR | Marca de las unidades legales presentes en BBDD de Aduana. Solo exporta (1), solo importa (2) o ambas (3). |
| AFECTO\_A\_IVA | Empresas tiene la obligación de para el Impuesto al Valor Agregado, 1 para afecto a IVA y 0 para los que no. |
| PERS\_NAT\_SII | Marca de las unidades legales según criterio de corte de Rut menor a cincuenta millones (SII). |
| PERS\_NAT\_CEPAL | Marca de las unidades legales según CEPAL,2013. "COMPENDIO DE RECOMENDACIONES TÉCNICAS Y BUENAS PRÁCTICAS PARA EL DESARROLLO DE LOS DIRECTORIOS DE EMPRESAS Y ESTABLECIMIENTOS ", pág.6. |
| ID | Identificador correlativo de uso interno. |
| TIPO\_DE\_REGISTRO | Indicador de si el registro es continuo o nuevo. |
| PEE\_AAAA | Marca que indica con un 1 si el registro fue levantado en la producción estadística estructural del año AAAA y con un 0 si no lo fue. |
| PEC\_AAAA | Marca que indica con un 1 si el registro fue levantado en la producción estadística coyuntural del año AAAA y con un 0 si no lo fue. |
| NOMINA\_F22\_AAAA | Marca de registros que están contenidos en el Formulario de declaración de Renta (F22). |

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

A continuación, se muestra la propuesta de estandarización de nombre de variables, ventajas y desventajas, con el fin de cumplir con el objetivo propuesto en este documento.

# ESTANDARIZACIÓN DE NOMBRE DE VARIABLES

En el ámbito del análisis de datos, la estandarización del nombre de las variables es un proceso fundamental que busca mejorar los distintos procesos o etapas de la construcción o actualización del MME. La estandarización es una práctica ampliamente utilizada en la ciencia de datos y estadística, ya que proporciona múltiples ventajas y desventajas que impactan directamente en la calidad y confiabilidad del producto final.

* **Ventajas**:
* Claridad y comprensión: Los nombres estandarizados hacen que sea más fácil entender la información contenida en una base de datos, tanto para los usuarios actuales como para los futuros. Los nombres descriptivos y consistentes reducen la ambigüedad y mejoran la comprensión de los datos.
* Facilita la búsqueda y recuperación de datos[[1]](#footnote-1): Los nombres estandarizados permiten a los usuarios encontrar rápidamente los datos que necesitan, ya que pueden buscar por nombres coherentes y significativos en lugar de intentar descifrar nombres confusos o ambiguos.
* Interoperabilidad[[2]](#footnote-2): Los datos con nombres estandarizados son más fáciles de compartir y combinar con otras bases de datos. Esto es especialmente importante en entornos donde se requiere intercambio de datos entre diferentes sistemas y organizaciones.
* Consistencia y calidad de los datos[[3]](#footnote-3): Al usar un enfoque estandarizado, se reduce la posibilidad de errores y duplicaciones en los nombres de variables. Esto mejora la calidad general de los datos y evita confusiones.
* Análisis más eficiente: Los nombres estandarizados agilizan el proceso de análisis, ya que los analistas y científicos de datos pueden identificar rápidamente las variables relevantes sin necesidad de adivinaciones o verificaciones adicionales.
* **Desventajas:**
* Pérdida de información contextual: En algunos casos, al estandarizar los nombres de variables, es posible que se pierda cierta información contextual que podría ser útil para comprender completamente los datos.
* Adaptación a estándares cambiantes: Los estándares de nomenclatura pueden cambiar con el tiempo debido a cambios en la industria o en las regulaciones. Esto puede requerir modificaciones frecuentes en la base de datos para mantenerse al día con los estándares actuales.
* Necesidad de capacitación: Implementar nombres estandarizados puede requerir una curva de aprendizaje para los usuarios y administradores de la base de datos, lo que podría generar resistencia inicial o confusión.
* Costo y tiempo: La estandarización de nombres puede requerir inversión de tiempo y recursos para modificar nombres existentes y asegurarse de que los nuevos nombres se adhieran al estándar establecido.
* Requisitos de adaptabilidad: En algunos casos, puede haber necesidades específicas en una organización que no se ajusten completamente a un estándar de nomenclatura, lo que podría requerir excepciones o adaptaciones.

# PROBLEMÁTICA

En esta sección, se muestra la identificación de algunas problemáticas relacionadas con la estandarización de nombres de las variables en el Marco Maestro de Empresas. Una de las problemáticas observadas es la presencia de números asociados a los insumos que se utilizan para asignar el dato en los distintos nombres de variables del MME, por ejemplo, se tiene el MME ámbito real y se quiere asignar información sobre un insumo de producción estadística coyuntural del año 2020, entonces, para la situación actual del nombre de la variable en el MME es denominada “PE\_2020”, esto provoca un problema el cual consiste en que cuando se actualizan los datos con un insumo más actualizado (año 2021), entonces la variable vuelve a cambiar de nombre dentro MME y trae consecuencias al momento de automatizar el proceso de actualización del MME, además de complejizar el proceso de análisis longitudinal del dato, ya que al llamarse de distintos nombres de variables, éstos no son posible realizar análisis de manera fluida.

Lo anterior, es una de las principales problemáticas correspondiente a la estandarización del nombre de las variables del MME.

Además, el MME posee problemas con caracteres especiales, uso de la ñ, tildes y nombres de variables compuesta de una o más espacios entre medio, siendo etas dos últimas las no concurridas por el MME actual.

Las problemáticas expuestas en el ámbito real, también se replican en el ámbito ficticio del MME.

# PROPUESTA

A continuación, se muestra el nombre de la variable original, el nuevo nombre propuesto para la misma variable con problemas y su justificación correspondiente, usando las tablas 4 y 5.

Cabe destacar que, se propone de antemano, que el nombre de las variables compuestas de dos o más monosílabas deben ir unidas por un guion bajo, el cambio de nombre de las variables originales que se proponen modificar es para aplicar la estandarización del nombre de las variables del marco maestro de empresas (ambos ámbitos, respectivamente).

Es importante destacar que todas las variables deben cumplir con ciertas normas de estandarización para garantizar la consistencia y facilidad en el análisis. Por lo tanto, se sugiere eliminar el uso de la letra “ñ”, los acentos, signos ASCII como “$”, y escribir todas las variables en minúsculas, al seguir estas pautas de estandarización, se logrará una base de datos más homogénea y coherente, lo que facilitará el procesamiento y análisis de la información, así como la comparación entre variables y la obtención de resultados más confiables y significativos.

Tabla 4. Nombre de variables actuales en el marco maestro de empresas ámbito real, su propuesta de estandarización de variables y justificación.

|  |  |  |
| --- | --- | --- |
| ***Nombre de la variable original*** | ***Nombre de la variable propuesta*** | ***Justificación*** |
| REGION\_CUT\_2018 | region | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| PROVINCIA\_CUT\_2018 | Provincia | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| COMUNA\_CUT\_2018 | Comuna | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| SECCION\_CIIU4CL | seccion\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| DIVISION\_CIIU4CL | division\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| GRUPO\_CIIU4CL | grupo\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| CLASE\_CIIU4CL | clase\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| SUBCLASE\_CIIU4CL | subclase\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| ESTRATOS\_F29 | estrato\_venta | Se elimina la extensión del nombre del insumo del cual se asigna el dato |
| FONO | telefono | Se incorpora la palabra completa |
| TAMAÑO\_OCUPADOS\_AAAA | tamanio\_ocupados | Se reemplaza la letra ñ por la sílaba ni, y se elimina el año de referencia del insumo |
| TERMINO\_GIRO\_AAAA | termino\_de\_giro | Se elimina la extensión del año de referencia del insumo que se asigna el dato |
|  | anio | Se incorpora esta variable para homogeneizar la base de datos |
|  | periodo | Se incorpora esta variable para homogeneizar la base de datos |

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Tabla 5. Nombre de variables contenidas en el marco maestro de empresas ámbito ficticio, su propuesta de estandarización de variables y justificación.

|  |  |  |
| --- | --- | --- |
| ***Nombre de la variable original*** | ***Nombre de la variable propuesta*** | ***Justificación*** |
| REGION\_CUT\_2018 | region | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| PROVINCIA\_CUT\_2018 | provincia | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| COMUNA\_CUT\_2018 | comuna | Se elimina la extensión con el nombre del código único territorial, año 2018 |
| SECCION\_CIIU4CL | seccion\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| DIVISION\_CIIU4CL | division\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| GRUPO\_CIIU4CL | grupo\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| CLASE\_CIIU4CL | clase\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| SUBCLASE\_CIIU4CL | subclase\_acteco | Se elimina la extensión del clasificador chileno de actividad económica |
| VTA\_$\_ENERO | venta\_enero | Se elimina el signo "$" |
| VTA\_$\_FEBRERO | venta\_febrero | Se elimina el signo "$" |
| VTA\_$\_MARZO | venta\_marzo | Se elimina el signo "$" |
| VTA\_$\_ABRIL | venta\_abril | Se elimina el signo "$" |
| VTA\_$\_MAYO | venta\_mayo | Se elimina el signo "$" |
| VTA\_$\_JUNIO | venta\_junio | Se elimina el signo "$" |
| VTA\_$\_JULIO | venta\_julio | Se elimina el signo "$" |
| VTA\_$\_AGOSTO | venta\_agosto | Se elimina el signo "$" |
| VTA\_$\_SEPTIEMBRE | venta\_septiembre | Se elimina el signo "$" |
| VTA\_$\_OCTUBRE | venta\_octubre | Se elimina el signo "$" |
| VTA\_$\_NOVIEMBRE | venta\_noviembre | Se elimina el signo "$" |
| VTA\_$\_DICIEMBRE | venta\_diciembre | Se elimina el signo "$" |
| VTA\_ANUAL\_$\_F29\_AAAA | venta\_anual | Se elimina el signo "$" y la extensión del insumo junto con el año de procedencia de este |
| VTA\_ANUAL\_UF\_F29\_AAAA | venta\_anual\_uf | Se elimina la extensión del insumo junto con el año de procedencia de este |
| ESTRATOS\_F29 | estrato\_ventas | Se elimina la extensión del nombre del insumo del cual se asigna el dato |
| TAMAÑO\_F29 | tamanio\_ventas | Se elimina la extensión del nombre del insumo del cual se asigna el dato |
| CUENTA\_OCUPADOS\_AAAA | cuenta\_ocupados | Se elimina el año de referencia del insumo |
| TAMAÑO\_OCUPADOS\_AAAA | tamanio\_ocupados | Se elimina el año de referencia del insumo; se reemplaza la letra ñ por la sílaba ni. |
| TERMINO\_GIRO\_AAAA | termino\_de\_giro | Se elimina el año de referencia del insumo |
| NOMINA\_F29\_AAAA | nomina\_ventas\_iva | Se elimina el nombre y año de referencia del insumo |
| NOMINA\_F22\_AAAA | nomina\_ventas\_renta | Se elimina el nombre y año de referencia del insumo |
| PROM\_TRAB\_HONORARIO\_AAAA | promedio\_trabajadores\_honorario | Se elimina el año de referencia del insumo |
| PROM\_TRAB\_OCUPADOS\_AAAA | promedio\_trabajadores\_ocupados | Se elimina el año de referencia del insumo |
| PEE\_AAAA | produccion\_estadistica\_estructural | Se elimina el año de referencia del insumo |
| PEC\_AAAA | produccion\_estadistica\_coyuntural | Se elimina el año de referencia del insumo |
|  | anio | Se incorpora esta variable para homogeneizar la base de datos |
|  | periodo | Se incorpora esta variable para homogeneizar la base de datos |

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

## Difusión

A continuación, este capítulo mostrará una propuesta de difusión del nombre de las variables y su definición, si se cuenta con aceptación de la propuesta planteada. Es decir, el diccionario de variables propuesto para la base de datos del marco maestro de empresas, ámbito real y ficticio quedaría de la siguiente manera (ver tabla 6 y 7).

Tabla 6. Nombre de variables propuestas para difusión del marco maestro de empresas ámbito real y su definición.

|  |  |
| --- | --- |
| ***Nombre del Campo Propuesto*** | ***Definición del Campo Propuesto*** |
| rut\_real | Rol único tributario otorgado por la administración fiscal, que identifica a la unidad legal en forma oficial. |
| dv\_rut | Dígito verificador del RUT |
| id\_grupo\_empresas\_ine | Identificador(es) de cada Grupo según Grupos de Empresas INE. |
| id\_grupo\_empresas\_cmf | Identificador de cada grupo según la Comisión del Mercado Financiero (CMF). |
| nom\_grupo\_empresas\_ine | Nombre del grupo de empresas identificado por INE. |
| razon\_social | Nombre o denominación de la unidad legal que desarrolla una actividad económica, sea persona natural o jurídica. |
| representante\_legal | Nombre del representante legal de la empresa. |
| nombre\_fantasia | Nombre o denominación comercial de la unidad legal que desarrolla una actividad económica, sea persona natural o jurídica. |
| region | Código de 2 dígitos correspondiente a la región, de acuerdo con la división político-administrativa oficial del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| provincia | Código de 3 dígitos correspondiente a la Provincia, de acuerdo con la división político-administrativa oficial del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| comuna | Código de 5 dígitos correspondiente a la Comuna, de acuerdo con la división político-administrativa del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| direccion | Ubicación Geográfica asociada a la unidad local principal. |
| inicio\_de\_actividades | Fecha de inicio de actividades según el SII. |
| sindico\_de\_quiebras | Marca de la nómina de síndico de quiebras. |
| seccion\_acteco | Primer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (letra) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 0. |
| division\_acteco | Segundo nivel de clasificación económica CIIU Rev.4. Adaptación nacional (2 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 00. |
| grupo\_acteco | Tercer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (3 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 000. |
| clase\_acteco | Cuarto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (4 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 0000. |
| subclase\_acteco | Quinto nivel de clasificación económica CIIU Rev.4. Adaptación nacional (5 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 00000. |
| estrato\_ventas | Estrato calculado según ventas anuales en UF obtenidas de la selección de algunos códigos del F29. |
| fono | Teléfono principal de la empresa. |
| correo\_electronico | Correo electrónico principal de la Unidad Legal. |
| pagina\_web | Página WEB principal de la Unidad Legal. |
| directorio\_priorizado\* | Marca que indica 1 si cumple con ser registro clasificado como empresa con venta de tamaño grande (estrato 9 al 12), registros coincidentes con la producción estadística o grupos empresariales INE, y 0 para los no coincidentes. |
| grandes\_contribuyentes | Marca de grandes contribuyentes, 1 para los grandes contribuyentes y 0 para los que no lo son. |
| regimen\_tributario | Marca de la nómina de Régimen Tributario. |
| categoria\_tributaria | Marca que indica con un 1 si la unidad legal es de primera categoría tributaria, 2 si es de segunda categoría, 3 si presenta ambas categorías y 0 si no tiene categoría asociada. |
| tipo\_contribuyente | Tipo de sociedad constituida por el contribuyente según el SII. |
| subtipo\_contribuyente | Mayor desagregación del tipo de sociedad declarada por el contribuyente según SII. |
| comercio\_exterior | Marca de las unidades legales presentes en comercio exterior del Servicio Nacional de Aduanas, que indica como 1, si sólo exporta, 2 si solo importa o 3 si realiza ambas transacciones. |
| moneda\_extranjera | Marca de las unidades legales que están autorizados a declarar formularios asociados a compras y/o ventas en moneda extranjera del SII, donde 1 es para declarantes de moneda extranjera y 0 para los no declarantes de moneda extranjera. |
| afecto\_a\_iva | Marca de empresas que tienen obligación de pagar el Impuesto al Valor Agregado (IVA), donde 1 es para declarantes afectos a IVA y 0 para los que no. |
| pers\_nat\_cepal\* | Marca de las unidades legales de acuerdo con criterio CEPAL del año 2013, el cual indica 1 si el contribuyente tiene cinco o más ocupados en la empresa y 0 si tiene cuatro o menos ocupados. |
| tamanio\_ocupados | Tamaño de la empresa según ocupados declarados en la DJ1887 en el año AAAA, y clasificados de acuerdo con la Ley N°20.416 del Estatuto Pyme. |
| tipo\_de\_registro | Estado del registro en función a su aparición en la ejecución del marco maestro de empresas anterior como registro continuo (0) o registro nuevo (1). |
| pers\_nat\_sii\* | Clasificación de las unidades legales mediante el criterio de umbral de RUT inferior a cincuenta millones (SII) |
| termino\_de\_giro | Marca de la nómina de registros que contenga término de algún giro que posea el contribuyente ante el SII. |
| id\* | Identificador correlativo de uso interno. |
| normalizacion\* | Marca que indica con un SI la dirección geográfica fue normalizada, NO si no lo fue y NA si no presenta dirección. |
| produccion\_estadistica\_estructural\* | Marca que indica con un 1 si el registro fue levantado en la producción estadística estructural del año AAAA y con un 0 si no lo fue. |
| produccion\_estadistica\_coyuntural\* | Marca que indica con un 1 si el registro fue levantado en la producción estadística coyuntural del año AAAA y con un 0 si no lo fue. |
| produccion\_estadistica | Marca que indica con un 1 si pertenece a la producción estadística estructural, 2 si pertenece a la producción coyuntural, 3 si pertenecen a ambas producciones estadísticas INE y, 0 si no pertenecen a los casos anteriores. |
| anio | Año (AAAA) contable que define el marco maestro de empresas. |
| periodo | fecha (MM/AAAA) correspondiente al mes y año calendario de entrega del marco maestro de empresas. |

/\* Se propone eliminar estas variables ya que son de uso interno del Subdepartamento de marcos y muestras.

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Tabla 7. Nombre de variables propuestas para difusión del marco maestro de empresas ámbito ficticio y su definición.

|  |  |
| --- | --- |
| ***Nombre del Campo Original*** | ***Nombre del Campo Propuesto*** |
| rut\_fictcio | Rut real encriptado asociado a una empresa, denominado "rut ficticio". |
| tipo\_de\_registro | Estado del registro en función a su aparición en la ejecución del marco maestro de empresas anterior como registro continuo (0) o registro nuevo (1). |
| moneda\_extranjera | Marca de las unidades legales que están autorizados a declarar formularios asociados a compras y/o ventas en moneda extranjera del SII, donde 1 es para declarantes de moneda extranjera y 0 para los no declarantes de moneda extranjera. |
| grandes\_contribuyentes | Marca de grandes contribuyentes, 1 para los grandes contribuyentes y 0 para los que no lo son. |
| grupo\_empresas\_cmf | Identificador de cada grupo según la Comisión del Mercado Financiero (CMF). |
| sindico\_de\_quiebras | Marca de la nómina de síndico de quiebras. |
| estructurales\* | Marca que indica con un 1 si el registro fue levantado en la producción estadística estructural del año AAAA y con un 0 si no lo fue. |
| segunda\_categoria | Marca de las unidades que pertenecen a algún grupo empresarial según levantamientos de encuestas INE, donde 1 es para los coincidentes y 0 para los que no lo son. |
| grupo\_empresas\_ine | Marca correspondiente a la coincidencia en unidad legal con respecto al insumo de grupo de empresas INE, marcado como 1, y 0 en otro caso. |
| region | Código de 2 dígitos correspondiente a la región, de acuerdo con la división político-administrativa oficial del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| provincia | Código de 3 dígitos correspondiente a la Provincia, de acuerdo con la división político-administrativa oficial del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| comuna | Código de 5 dígitos correspondiente a la Comuna, de acuerdo con la división político-administrativa del Código Único Territorial (CUT), vigente desde el 06 de septiembre 2018. |
| seccion\_acteco | Primer nivel de clasificación económica CIIU Rev.4. Adaptación nacional (letra) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 0. |
| division\_acteco | Segundo nivel de clasificación económica CIIU Rev.4. Adaptación nacional  (2 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 00. |
| grupo\_acteco | Tercer nivel de clasificación económica CIIU Rev.4. Adaptación nacional  (3 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 000. |
| clase\_acteco | Cuarto nivel de clasificación económica CIIU Rev.4. Adaptación nacional  (4 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 0000. |
| subclase\_acteco | Quinto nivel de clasificación económica CIIU Rev.4. Adaptación nacional  (5 dígitos) vigente a partir del año 2012. En caso de no tener actividad asignable, toman el valor 00000. |
| venta\_enero | Ventas del mes - 1 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_febrero | Ventas del mes - 2 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_marzo | Ventas del mes - 3 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_abril | Ventas del mes - 4 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_mayo | Ventas del mes - 5 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_junio | Ventas del mes - 6 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_julio | Ventas del mes - 7 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_agosto | Ventas del mes - 8 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_septiembre | Ventas del mes - 9 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_octubre | Ventas del mes - 10 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_noviembre | Ventas del mes - 11 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_diciembre | Ventas del mes - 12 informado por el contribuyente en la declaración mensual de IVA - F29 del año tributario AAAA. |
| venta\_anual | Ventas acumuladas en pesos declarada los último 12 meses por los contribuyentes en la declaración mensual de IVA - F29, año tributario AAAA. |
| venta\_anual\_uf | Ventas acumuladas en UF declarada los último 12 meses por los contribuyentes en la declaración mensual de IVA - F29, año tributario AAAA. |
| produccion\_estadistica | Marca que indica con un 1 si pertenece a la producción estadística estructural, 2 si pertenece a la producción coyuntural, 3 si pertenecen a ambas producciones estadísticas INE y, 0 si no pertenecen a los casos anteriores. |
| comercio\_exterior | Marca de las unidades legales presentes en comercio exterior del Servicio Nacional de Aduanas, que indica como 1, si sólo exporta, 2 si solo importa o 3 si realiza ambas transacciones. |
| regimen\_tributario | Marca de la nómina de Régimen Tributario. |
| directoro\_priorizado | Marca que indica 1 si cumple con ser registro clasificado como empresa con venta de tamaño grande (estrato 9 al 12), registros coincidentes con la producción estadística o grupos empresariales INE, y 0 para los no coincidentes. |
| afecto\_a\_iva | Marca de empresas que tienen obligación de pagar el Impuesto al Valor Agregado (IVA), donde 1 es para declarantes afectos a IVA y 0 para los que no. |
| estrato\_ventas | Estrato calculado según ventas anuales en UF obtenidas de la selección de algunos códigos del F29, año tributario AAAA. |
| tamanio\_ventas | Tamaño de la empresa de acuerdo con el monto de la venta anual en UF, calculado desde el Formulario 29 año tributario AAAA. |
| cuenta\_oupados | Cantidad total de trabajadores declarados por los contribuyentes en el F1887 anual, año tributario AAAA. |
| tamanio\_oupados | Tamaño de la empresa según ocupados declarados en la DJ1887 en el año AAAA, y clasificados de acuerdo con la Ley N°20.416 del Estatuto Pyme. |
| pers\_nat\_cepal\* | Marca de las unidades legales de acuerdo con criterio CEPAL del año 2013, el cual indica 1 si el contribuyente tiene cinco o más ocupados en la empresa y 0 si tiene cuatro o menos ocupados. |
| pers\_nat\_sii\* | Clasificación de las unidades legales mediante el criterio de umbral de RUT inferior a cincuenta millones (SII) |
| categoria\_tributaria | Marca que indica con un 1 si la unidad legal es de primera categoría tributaria, 2 si es de segunda categoría, 3 si presenta ambas categorías y 0 si no tiene categoría asociada. |
| termino\_de\_giro | Marca de la nómina de registros que posee término de giro ante el SII. |
| id\* | Identificador correlativo de uso interno. |
| nomina\_ventas\_iva\* | Marca de registros que están contenidos en el Formulario de declaración Mensual de IVA (F29), AT AAAA. |
| nomina\_ventas\_renta\* | Marca de registros que están contenidos en el Formulario de declaración Anual de Renta (F22), AT AAAA. |
| prom\_trabajadores\_honorario | Promedio anual de trabajadores declarados por los contribuyentes en el Formulario de trabajadores - F1879, AT AAAA. |
| prom\_trabajadores\_ocupados | Promedio anual de trabajadores declarados por los contribuyentes en el Formulario de trabajadores - F1887, AT AAAA. |
| produccion\_estadistica\_estructural\* | Marca que indica con un 1 si el registro fue levantado en la producción estadística estructural del año AAAA y con un 0 si no lo fue. |
| produccion\_estadistica\_coyuntural\* | Marca que indica con un 1 si el registro fue levantado en la producción estadística coyuntural del año AAAA y con un 0 si no lo fue. |
| anio | Año (AAAA) contable que define el marco maestro de empresas |
| periodo | fecha (MM/AAAA) correspondiente al mes y año calendario de entrega del marco maestro de empresas. |

/\* Se propone eliminar estas variables ya que son de uso interno del Subdepartamento de marcos y muestras.

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

1. • Inmon, W. H. (2005). Building the Data Warehouse. John Wiley & Sons. [↑](#footnote-ref-1)
2. • Redman, T. C. (2001). Data Quality: The Field Guide. Digital Press. [↑](#footnote-ref-2)
3. • Batini, C., & Scannapieco, M. (2016). Data and Information Quality: Dimensions, Principles and Techniques. Springer. [↑](#footnote-ref-3)