



DIRECTORIO NACIONAL DE EMPRESAS INFORME ANUAL 2007

Santiago, Junio 2009

Instituto Nacional de Estadísticas
DIRECTORIO ANUAL INFORME 2007

Período de la información: 2007

Fecha última de publicación: Junio de 2009

Subdirección de Operaciones

Proyecto Directorio Nacional de Empresas

Profesionales responsables:

Oriana Villegas Pérez, oriana.villegas@ine.cl

Paseo Bulnes 209, piso 2 Of. 21, Santiago Centro, Chile.

Fono (56-2) 7962451 / (56-2) 3667777

Santiago-Chile.

RESUMEN EJECUTIVO

El Directorio corresponde a un listado de empresas constituidas formalmente con representación geográfica, actividad económica y tamaño por ventas al que pertenecen, atributos válidos para la obtención continua de muestras utilizadas en la producción estadística económica oficial del INE. Este documento describe la elaboración del Directorio Anual 2007 (DIR2007), proporciona los antecedentes generales de la vigencia del Proyecto, los cambios metodológicos de mejora entre las versiones disponibles (2004, 2005, 2006), así como las principales características de la versión 2007. Se identifican las fuentes de información utilizadas y el procesamiento aplicado a éstas, para luego, incorporar como variables del Directorio los atributos que hacen posible la diferencia entre un registro administrativo (RRAA) y un Directorio de Empresas.

ÍNDICE

INTRODUCCIÓN	4
ANTECEDENTES GENERALES	5
JUSTIFICACIÓN	7
DEFINICIÓN Y OBJETIVOS	8
METODOLOGÍA	8
DIRECTORIO FICTICIO Y DIRECTORIO REAL	8
4.1 FUENTES DE INFORMACIÓN	9
4.1.1 FUENTE PRIMARIA	9
4.1.2 FUENTES SECUNDARIAS	10
4.2 PROCESAMIENTO DE LAS FUENTES DE INFORMACIÓN	11
VENTAS	11
ACTIVIDAD ECONÓMICA	15
UBICACIÓN GEOGRÁFICA	17
RESULTADOS ESTADÍSTICOS GENERALES	19
VENTAS TOTALES ANUALES	19
DISTRIBUCIÓN DE LA SUMA DE LAS VENTAS ANUALES POR SECCIÓN	19
DISTRIBUCIÓN DE LA SUMA DE LAS VENTAS ANUALES POR REGIÓN	20
DISTRIBUCIÓN DE EMPRESAS POR TAMAÑO SEGÚN ESTRATO DE VENTAS	20
DISTRIBUCIÓN DE LAS EMPRESAS POR TIPO DE VENTA	21
ANEXOS	22
ANEXO A	23
ANEXO B	24
DESCRIPCIÓN CAMPOS ÁMBITO REAL DIR2007	24
DESCRIPCIÓN CAMPOS ÁMBITO FICTICIO DIR2007	25

INTRODUCCIÓN

El presente documento describe la elaboración del Directorio Anual 2007 (DIR2007¹), producto desarrollado por el Proyecto Directorio Nacional de Empresas, del Instituto Nacional de Estadísticas (INE). El Directorio corresponde a un listado de empresas constituidas formalmente con representación geográfica, actividad económica y tamaño por ventas al que pertenecen, atributos válidos para la obtención continua de muestras utilizadas en la producción estadística económica oficial del INE.

En este documento se proporcionan antecedentes generales de la vigencia del Proyecto, cambios metodológicos de mejora que explican el comportamiento de los Directorios de empresa disponibles (versiones 2004, 2005, 2006), así como las principales características de la versión 2007 aquí detallada. Se identifican las fuentes de información utilizadas y el procesamiento aplicado a éstas, para luego, incorporar como variables del Directorio los atributos que hacen posible la diferencia entre un registro administrativo (RRAA) y un Directorio de Empresas, haciendo de éste la base única de agentes productivos con perfil de uso estadístico económico.

Finalmente, se entrega el resultado del procesamiento de las fuentes de información. Esto permite representar la realidad de la organización industrial formal vigente durante el año 2007, con características compatibles a las prácticas implementadas por instituciones homónimas a nivel mundial facilitando la interpretación de los datos aún cuando no dispone de filtros por organización jurídica, deciles, o tipo de empresa, etc.

Respecto de los clientes, se exponen las demandas específicas de información, así como el medio a través del cual el Proyecto Directorio pone a disposición el material de apoyo y la base de datos resultante para su utilización.

El Proyecto Directorio, en particular las personas responsables de la elaboración del DIR2007, agradecen al organismo proveedor de registros, Servicio de Impuestos Internos (SII) así como a quienes colaboraron con asesoría técnica para la toma de decisiones (Comité Técnico y jefes de productos).

¹ Respecto de las versiones anteriores del Directorio de empresas cambia la abreviatura MMA (Marco Muestral Anual) por DIR (Directorio).

ANTECEDENTES GENERALES

El Proyecto Directorio, dependiente de la Subdirección de Operaciones, se orienta al apoyo de la producción estadística económica del INE: en lo coyuntural mediante procesos continuos de actualización de muestras y, en lo estructural, a través de la construcción del Directorio de Empresas.

A diferencia de otros países como Brasil y Francia, Chile no cuenta con una legislación que le permita acceder con facilidad a la información de empresas y/o establecimientos del país. Esto sólo es posible mediante el cumplimiento de la Ley N° 17.374 INE del 15 de octubre de 1970 que, entre otras cosas, señala:

Artículo 2°, letra L: Corresponderá al INE confeccionar un registro de las personas naturales o jurídicas que constituyan "Fuente de Información Estadística".

Artículo 20: "Todas las personas naturales o jurídicas chilenas y las residentes o transeúntes están obligadas a suministrar datos, antecedentes o información de carácter estadístico que el Instituto Nacional de Estadísticas les solicite".

Artículo 22: "Las personas a que se refiere el artículo 20, de esta ley, que se negaren a suministrar los datos estadísticos que le fueren solicitados, o que falsearen o alteraren, sufrirán una multa. La aplicación y pago de la multa no exime al infractor de cumplir la obligación correspondiente".

Artículo 29: "El Instituto Nacional de Estadísticas y cada uno de sus funcionarios no podrán divulgar los hechos que se refieren a personas o entidades. El estricto mantenimiento de esta reserva constituye el "Secreto Estadístico".

Respecto de las fuentes de información, el modelo utilizado por países, como España, por ejemplo, promueve la utilización de datos tipo administrativo, mientras que en otros, son los censos económicos. Inicialmente en Chile, las investigaciones sobre qué fuente utilizar detectaron la existencia de directorios de empresas con representación nacional limitada y/o con antecedentes que no facilitaban el muestreo.

Chile cuenta, y está trabajando con registros administrativos provenientes del SII principalmente. Éstos no proporcionan antecedentes relevantes nominados² por contribuyente, velando por el secreto tributario (lo que ha obligado históricamente a construir en paralelo dos versiones del Directorio de empresas, uno real y otro ficticio, asociados a RUT real y ficticios, respectivamente). Adicionalmente, estos datos provienen de declaraciones tributarias anuales y mensuales, que provocan un desfase entre la elaboración

² Datos asociados a RUT real.

del Directorio y el año al que hace referencia (nueve meses aproximadamente), además de diferencias entre la declaración de actividad económica declarada y aquella ejercida. A pesar de estos aspectos, el Directorio, constituye un registro con propiedades potenciales de ser explotadas en la producción estadística mediante muestreo.

Por lo anteriormente expuesto, es preciso señalar que la información de la fuente es necesaria, no suficiente, para dar a la agrupación de empresas el carácter de Directorio. Los diversos procesos que se desarrollan en el Proyecto permiten incorporar mejoras en el ámbito real que, a través de conversiones³, se reflejan en el ámbito ficticio. Estos procesos son adicionales a la elaboración del Directorio de empresas y se realizan continuamente entre la construcción de uno y otro permitiendo contar con:

Un Directorio de empresas ficticio (actividad, ventas, región, salario y trabajadores) susceptible de ser utilizado en la obtención de las muestras para la producción estadística económica del INE.

Un Directorio de empresas real utilizado específicamente para entrevistas telefónicas con el fin de actualizar datos de ubicación geográfica y actividad económica especialmente, como también para capturar establecimientos asociados a empresas.

A la fecha, el INE cuenta con un Directorio de empresas para los años 2004, 2005, 2006 y 2007. Sólo el año 2004 se recurre al concepto y atributo ventas distorsionadas, utilizadas alternativamente por no contar en ese momento con el cálculo de ventas oficiales.

Respecto de la estructura, cada nueva versión, entre los años 2004 – 2007, posee características similares pero particulares en su composición, producto de los criterios con que se definen las mejoras continuas entre años. Éstas se traducen, por ejemplo, en: cambio de fórmula y mejora de la misma para el cálculo de ventas, método de asignación para actividad económica, depuración de la información que proporciona la fuente principal.

El resultado se visualiza en la tabla siguiente:

Total Registros	VERSIÓN DIRECTORIO DE EMPRESAS			
	2004	2005	2006	2007
Registros declarantes	902.202	902.116	916.818	941.411
Directorio de empresas	690.475	773.455	915.108	931.511
Con clasificación	99,30%	100,0%	98,87%	99,70%
Con DPA	99,10%	98,28%	97,88%	99,98%

³ Proceso informático realizado por el SII. A través de éste, facilita al INE la relación real – ficticia – real sin ser uno a uno (1:1)

JUSTIFICACIÓN

Los siguientes puntos contienen los argumentos fundamentales que sustentan la construcción de un Directorio de empresas, enfatizando aquellos de carácter técnico principalmente.

1. Existía desconocimiento del número de empresas y establecimientos en funcionamiento, formales e informales, públicos y privados, a nivel nacional, regional, provincial, comunal y las características de éstas(os).
2. No existen organismos y/o instituciones que confeccionen directorios de empresas, mas bien poseen RRAA que permiten obtener algún grado de contabilidad y caracterización de éstas, pero su realización no forma parte de la misión de ninguno de ellos.
3. De acuerdo con lo indicado previamente, es competencia del rol rector del INE en la producción de estadísticas oficiales, tal como lo señala la Ley 17.374 en su artículo 2, letra L, Confeccionar un registro de las personas naturales o jurídicas que constituyan "Fuente de información Estadística".
4. Alcanzar los estándares de instituciones equivalentes a nivel mundial, de acuerdo a lo establecido por la Organización para la Cooperación y el Desarrollo Económico (OCDE).
5. Satisfacer las demandas de los productores de estadísticas económicas que, entre otros, requieren:
 - a. Agentes económicos productivos legalmente constituidos, personas naturales o jurídicas, con atributos que permitan muestreo o censos de encuestas económicas.
 - b. Información de agentes "empresas" como características de tamaño por ventas, ubicación geográfica y actividad económica principal desarrollada.
6. Potenciar el mejoramiento continuo de los procesos orientados a satisfacer la demanda directa e indirecta, interna y externa al INE, respectivamente (Depto. Investigación & Desarrollo, productores de estadísticas económicas y Banco Central de Chile).

DEFINICIÓN Y OBJETIVOS

De acuerdo a los insumos disponibles y a los aprendizajes de experiencias en seminarios y/o pasantías internacionales, el INE de Chile define el Directorio de empresas como el conjunto organizado de registros cuya unidad de estudio es la empresa, existente en soporte magnético⁴, con atributos de identificación, dirección, distribución geográfica, clasificación por tamaño (ventas y número de trabajadores) y actividad económica. Representa el universo de empresas determinado a través del procesamiento de fuentes primarias. Los atributos para éstas se consiguen de fuentes secundarias, internas y externas.

El Directorio de empresas del INE es el principal resultado de las actividades realizadas por el Proyecto Directorio, teniendo en cuenta las restricciones de las fuentes de información, disponibilidad, oportunidad, compatibilidad y costo en el acceso. Éste es un consolidado único y formal que tiene como propósito ser el insumo útil para el desarrollo de estudios económicos y facilitar la obtención de sub poblaciones (muestras) para encuestas, por sí solo o para complementar otras fuentes oficiales (directorios provenientes de Subsecretarías y/o Ministerios).

METODOLOGÍA

Para confeccionar un Directorio de empresas, el Proyecto estableció la estructura a partir de fuentes de información primaria y secundaria de que dispone, en armonía con las variables de muestreo que buscan satisfacer las demandas de la producción estadística. Además, dado las fuentes administrativas recepcionadas y utilizadas por el INE, el listado de agentes económicos productivos (empresas) se obtiene de la declaración tributaria realizada por los contribuyentes a través del Formulario 29 (F29), es decir, conformada por todos los agentes que pagan Impuesto al valor agregado (IVA). Estos antecedentes permiten el desarrollo de dos ámbitos, ficticio y real.

Directorio ficticio y Directorio real

Como se señaló anteriormente, por restricciones de la fuente primaria de información (secreto tributario), cada Directorio de empresas requiere de la construcción de dos versiones. Éstas se comportan como "espejo" en la información agregada que proporcionan y no en la relación uno a uno, impidiendo conocer el valor estimado de las ventas mensuales y/o anuales por

⁴ Documentación y Archivos. Formato MS Access, Directorio Anual versiones Real y Ficticia. **Servidor Lingue link \\Maule\unidad directorio**

empresa. A partir de esto, las variables de muestreo⁵ que el INE necesita, se encuentran principalmente en el ámbito ficticio. La distribución de las mismas es la siguiente:

Dimensión de la variable de muestreo	Directorio de empresas 2007	
	Ficticio	Real
1. Identificación	Rut ficticio	Rut real, dv Razón social
2. Ventas	Mensuales y anuales en pesos chilenos. Anual además en UF.	Estrato por ventas
3. Actividad Económica	CIIU.CL	CIIU.CL
	CIIU.3	CIIU.3
4. Ubicación geográfica	Región	Región
	Provincia	Provincia
	Comuna	Comuna

El valor para cada variable de muestreo se obtiene desde fuentes de datos específicas que a continuación se describen.

4.1 Fuentes de Información

Las fuentes que el INE dispone para la elaboración del Directorio son de dos tipos: primaria y secundaria. La primera permite determinar el número de empresas y la estimación de las ventas mensuales y anuales realizadas en un año en particular. La segunda, es utilizada para asignar atributos de identificación, actividad económica, localización y empleo a las empresas.

Debido a que el objetivo es generar información, a partir del dato que proporciona el registro administrativo, el Proyecto tiene la misión de adaptar las fuentes mediante procesos de mejoramiento continuo como tratamientos de depuración, selección y agrupación de casos aislados para tratamiento especial con el fin de obtener estructuras válidas a las que puedan incorporarse los demás atributos útiles para la producción estadística, permitiendo año a año aumentar la calidad del mismo.

4.1.1 Fuente Primaria

Formulario 29 (F29) de Impuesto al Valor Agregado (IVA)⁶. Esta fuente proviene del SII y corresponde a las declaraciones tributarias mensuales, del impuesto al valor agregado (IVA) que legalmente son retenidas por la misma

⁵ En lo que se refiere a las variables componentes del Directorio (cantidad y nombres) han cambiado entre años. La descripción para las correspondientes al año 2007 se encuentran en el **ANEXO B** de este documento.

⁶ Además de utilizarse para la declaración de IVA, el SII lo establece para declaración de otros impuestos y retenciones mensuales por la aplicación del Impuesto a la Renta.

institución para formar parte de las arcas fiscales. En el Directorio, cumple la función de determinar las empresas que pertenecerán al universo, constituidas formalmente en Chile al año de referencia y estimar ventas a partir de ciertos códigos declarados en ellas.

Las empresas declarantes consideradas se presumen en movimiento durante el año de referencia de los datos, no importando la frecuencia de sus declaraciones mensuales. Cabe señalar que el año en que se hace efectiva la declaración no corresponde al año de la tributación, por el contrario, un proceso de declaración tributaria 2008 podría representar alguna corrección del ejercicio tributario 2007.

4.1.2 Fuentes Secundarias

Las fuentes secundarias, a diferencia de la primaria, no determinan la estructura o número de empresas que componen el Directorio sino a RRAA que contienen atributos asociables a un porcentaje del total, enriqueciéndolos. Estas fuentes son internas y externas:

FUENTES SECUNDARIAS	
Externas	Internas
Catastro de Contribuyentes. Esta fuente proviene del SII y, de ella se obtiene información complementaria relacionada con dirección, teléfonos, fax, email. En suma, medios de contacto con la empresa.	Consolidado de Clasificación de Actividad Económica (CCAЕ). Corresponde a la agrupación de registros únicos de empresa, en los que la actividad económica proviene de procesos continuos de verificación, desarrollo de bases sectoriales y bases de datos departamentales provenientes de productores de estadísticas económicas estructurales y coyunturales del INE. Para su utilización, los registros son depurados con el propósito de identificar una única actividad económica por empresa en la medida que sea posible. Esta tarea es previa a la elaboración de cada Directorio y, en su totalidad, representan un porcentaje del mismo.

Externas	Internas
<p>Síndico de quiebras. Información facilitada mensualmente por la Superintendencia de Quiebras. Contiene las empresas que se encuentran en proceso legal de quiebra declarada y permite identificar empresas con éste estado legal (no necesariamente operacional, podría ser en trámite)</p>	
<p>Formulario 22 (F22) de la Declaración sobre Impuestos a la Renta. Esta fuente proviene del SII y proporciona información especialmente de la actividad económica y localización para gran parte de las empresas pertenecientes al Directorio. Afecta a personas naturales y jurídicas, útil para sociedades anónimas e inversionistas extranjeros que realizan actividad económica en Chile, empresas constructoras exportadores, entre otros. A diferencia de la versión 2006 del Directorio, la versión 2007 no cuenta con actualizaciones desde esta fuente debido a que sólo se recepcionó en agosto 2008 y sin actividad económica.</p>	

4.2 PROCESAMIENTO DE LAS FUENTES DE INFORMACIÓN

De la fuente primaria se obtiene la cantidad de empresas que conforman el Directorio y la estimación de las ventas (mensuales – anuales). Las fuentes secundarias, en cambio, enriquecen la información asociada a las empresas permitiendo su identificación. El procesamiento de éstas se realiza a partir de la variable de muestreo de interés que corresponden a:

- Ventas.
- Actividad económica.
- Ubicación geográfica.

Ventas

Recuérdese que la fuente primaria a partir de la cual se obtiene el número de empresas pertenecientes al Directorio es el F29 del SII. Éstas no son declaraciones de ventas, sino de impuestos (19%) y, por lo tanto, la información que recibe el INE es de todos los códigos que el Formulario posee (con valor o no en la medida que cada agente declare o le corresponda declarar), asociados a cada contribuyente, sobre este concepto.

Considerando el contenido de declaración, se generó el cálculo de ventas, imprescindible en la construcción del Directorio ya que determina el universo de las empresas vigentes en el año desarrollado. Este ejercicio permite además obtener el tamaño de las mismas por estrato considerando que:

El SII envía periódicamente, con RUT ficticio, las declaraciones primitivas de los formularios F29 y F50⁷ del mes indicado (enero – diciembre, o enero – febrero del año siguiente, según corresponda). Debido a que las declaraciones son recibidas por el INE con posterioridad al período en que las ventas se llevan a cabo, registra el ingreso de la información con tiempo de desfase cercano a un mes o más, y los antecedentes mensuales se remiten acumulados. Recepcionados los datos, son cargados y procesados por el Depto. de Informática INE (aplica a F29 y no a F50). En paralelo, Directorio realiza el ejercicio para identificar eventuales inconsistencias.

Las bases de datos provenientes del F29 contienen antecedentes del IVA pagado por contribuyentes declarantes del periodo en desarrollo y anteriores. Debido a que esta versión del Directorio hace referencia al año 2007, para el cálculo de ventas sólo se consideraron registros de este periodo.

Dado que el F29 contiene información sobre IVA pagado, fue necesario evaluar los campos declarados en el mismo, con el fin de obtener una fórmula de cálculo de ventas que incorpore incluso a contribuyentes que son declarantes en formularios antiguos (códigos 108 y 112).

Asociadas a este formulario existen las declaraciones rectificatorias⁸. Éstas se reciben periódicamente e incorporaron al procesamiento de acuerdo a lo indicado por el SII, es decir, la última disponible (puede haber más de una por contribuyente) reemplaza a la declaración original. En consideración a que una declaración rectificatoria podría ingresar al SII hasta con cinco años de desfase (en el extremo), para la elaboración del DIR2007 se utilizó como criterio incorporar las declaraciones provenientes por este concepto con hasta tres meses terminado el año tributario (declaraciones mensuales del 2007 realizadas durante enero, febrero y marzo de 2008)

En síntesis, la fórmula de cálculo de ventas se aplica a declaraciones juradas de impuestos F29 provenientes del Servicio de Impuestos Internos con

⁷ Formulario proveniente del SII, representa a los contribuyentes que declaran remesas al exterior. Recepcionado por la unidad de Informática – INE.

⁸ Primitiva es la primera declaración para el período tributario. Las rectificatorias se realizan sobre las primitivas reemplazando su información por completo. Éstas son sustitutivas, siendo válida la que tiene el número identificador más alto. El F50 (remesas al exterior) registra más de una declaración por contribuyente para el mismo período y se pueden rectificar individualmente, no obstante, son de baja frecuencia. Actualmente el SII no ha resuelto como asociar rectificatorias con primitivas.

periodicidad mensual. De este modo, se estiman ventas mensuales y anuales por contribuyente.

Como antecedente de mejora, la versión del DIR2007 se procesó con la fórmula de cálculo 2006 revisada por el INE y el BCCH⁹ en la que el principal cambio es la exclusión del código 054 (Monto por el 0,5% de las ventas efectuadas por contribuyentes calificados como suplementeros).

La estructura de la fórmula y la descripción de los campos contenidos se presenta a continuación:

$$\sum_i^{12} [C20_i + C142_i + C587_i + (C108_i + C502_i + C111_i + C112_i + C513_i + C510_i + C709 + (C517_i * 19 / 10) + C501_i) / 0,19 + C152_i * 10 + C529_i]$$

Código	Descripción
20	Monto por exportaciones realizadas que no generan débito.
108	Monto por el 19% de ventas y servicios facturados afectos del mes. F29 - 2003.
111	Monto por el 19% de ventas efectuadas con boletas a otro comerciante no industrial (débito)
112	Monto por el 19% de ventas efectuadas sin boletas por valores inferiores al mínimo.
142	Monto por ventas y/o servicios prestados. Internos - exentos o no gravados.
152	Monto por el 10% de renta percibida por contribuyentes que efectúan prestación de servicios.
501	Monto por el 19% de comisión (si es liquidación de factura recibida) o 19% por la venta (si es liquidación de factura emitida)
502	Monto por el 19% de ventas efectuadas con facturas a otro comerciante industrial (débito)
510	Monto por el 19% de disminuciones de venta consignado en Notas de crédito emitidas por facturas.
513	Monto por el 19% de aumentos de venta consignado en Notas de débito emitidas.
517	Monto por el 10% de ventas con retención parcial.
529	Monto de las ventas del período para contribuyentes afectos a tributación simplificada (Art. 29 DL 827/74)
587	Monto por facturas de compra recibidas con retención total (contribuyentes retenidos) La retención es realizada por un tercero.
709	Monto por el 19% de disminuciones de venta consignado en Notas de crédito emitidas por vales de máquina autorizadas por SII.

Una vez realizado el procesamiento del cálculo se ejecutaron validaciones de primer y segundo nivel, con el fin de detectar casos erróneos y/u omisiones y

⁹ Banco Central de Chile.

proponer depuraciones y/o imputaciones. Los casos irregulares fueron tratados en paralelo, con el apoyo del SII y se asocian a:

Coherencia mensual a nivel de códigos declarados en el F29. Se realiza una vez obtenido el cálculo de ventas a fin de imputar valores por declaraciones mal completadas y sin rectificatorias.

Resultado irregular del procesamiento de ventas¹⁰:

Detalle	N° registros únicos	Detalle	N° registros únicos
Ventas con problemas	141.827	Ventas cero	11
		Ventas negativas	50
		Ventas nulas	141.766
Ventas normales	789.684		
Total DIR2007	931.511		

Ventas anuales cero; corresponde a registros que declaran códigos de la fórmula distintos a cero, en los que la suma aritmética de éstos da como resultado ventas anuales iguales a cero. Son 11 registros asociados a 89 declaraciones.

Ventas anuales negativas; corresponde a registros que declaran códigos de la fórmula distintos a cero, en los que la suma aritmética de éstos da como resultado ventas anuales negativas. Son 50 registros asociados a 421 declaraciones.

Ventas anuales nulas; corresponde a registros que declaran códigos de la fórmula con valor cero o sin valor provocando indeterminación de la fórmula. Son 141.766 registros asociados a 1.137.968 declaraciones.

Depuración de registros con declaraciones con repeticiones y rectificatorias F29¹¹. En una situación ideal, la estimación de ventas mensuales por registro es posible si un contribuyente proporciona al SII una declaración por mes, sin correcciones posteriores. Sin embargo, en caso de que la declaración proporcionada por el contribuyente sea incorrecta, éste tiene la posibilidad de generar rectificatorias posteriores con un plazo de hasta 5 años.

El procesamiento del DIR2007 da como resultado 15.225 contribuyentes con declaraciones que tenían repeticiones mensuales. Para éstas se realizó un proceso de depuración en base a criterios contables y normativa tributaria que permitieron conseguir la relación un mes - una declaración para cada empresa

¹⁰ Para mayor información Documento Técnico N° 1 "*Validación e imputación de contribuyentes con ventas anuales cero, negativas y nulas dir2007*" del 170908.

¹¹ Para mayor información Documento Técnico N° 3 "*Descripción del proceso de depuración de registros asociados a declaraciones con repeticiones F29 para la construcción del DIR2007*" del 170908.

y, finalmente imputar un valor de venta mensual para éstos registros. La distribución es la siguiente:

Tipo de contribuyente	Rut Únicos	Nº declaraciones
Unideclarantes	14.960	17.507
Multideclarantes	265	1.736
Total	15.225	19.243

Actividad Económica

En relación a la asignación del código de actividad económica, ésta es usualmente proporcionada por el SII asociado al ámbito ficticio y la obtiene a través de las declaraciones juradas que los contribuyentes realizan ante el SII en el F22 (código 14). La versión del Directorio 2007 no cuenta con esta actualización en consideración a que, como se mencionó anteriormente, el RRAA no fue recepcionado oportunamente por el INE, éste ingresó en agosto 2008 sin actividad económica.

Lo anterior significó en esta versión del DIR2007 que, a los RUT que se mantienen desde año 2006 se les asignara la actividad del MMA2006 y, para los nuevos RUT se solicitó al SII enviar la información declarada en algún formulario (por ejemplo, el Formulario de Inicio de actividades).

El clasificador de actividad económica proporcionado por el SII es el Clasificador Internacional Industrial Uniforme – Adaptación Chilena (CIIU.cl, versión 2007). Sin embargo, y a pesar de que el INE lidera las mesas de trabajo sobre actualización de la CIIU.cl, por razones prácticas, de continuidad y comparabilidad internacional, se utiliza además el Clasificador Internacional Industrial Uniforme – Revisión 3 (CIIU3) a cuatro dígitos. La versión del DIR2007 contiene ambos clasificadores cuando a nivel de clase son iguales a partir de la comparación de la actividad de origen del SII con la actividad mejorada en CIIU3 del MMA06. Esto representa un 80% del total de los registros aproximadamente. Para aquellos casos en los que no existía relación uno a uno, se realizó comparación a nivel de grupo. Este ejercicio permitió conseguir actividad económica con CIIU.cl y CIIU3 para el 99,70% y 97,85% respectivamente.

Finalmente, cuando no se encontró código de actividad económica asignable en ninguno de los dos clasificadores ó para los casos en que no existía CIIU.cl pero sí CIIU3 (del año anterior 2006 conseguido a través de equivalencias con CASII) se indagó en el Formulario de Inicio de Actividades a través del SII. Esto permitió conseguir información para 8,9% de los casos en CIIU.cl. El resumen de completitud es el siguiente:

Descripción	Total		
Registros antiguos	779.181		
Registros nuevos	152.350		
Total DIR2007	931.511	Registros en CIU.ci	911.557
		Registros en CIU_3	928.572

Teniendo en consideración que una empresa puede tener más de una actividad en el MMA2006, que la información facilitada por el SII para el DIR2007 sólo correspondió a declarantes nuevos y que el Directorio debe contener una actividad como principal, se tiene lo siguiente:

Registros del DIR2007 coincidentes con el MMA2006: aproximadamente 767.000 registros a los que, inicialmente se les imputó la actividad 2006.

Registros con códigos múltiples: Esta situación se evidenció en 39.000 registros y la depuración para obtener una actividad se efectuó a través del atributo razón social y estrato de ventas (haciendo énfasis en la revisión de estratos¹² altos).

Empresas consideradas declarantes en moneda extranjera: Estos son 39 registros y se determinó mantenerlos sin atributos de actividad económica y ubicación geográfica. Estos se encuentran marcados en el ámbito ficticio.

Registros con códigos únicos con clasificación no válida: Este problema se evidenció sólo para 11 registros. Se procedió a obtener datos desde la fuente Catastro de Contribuyentes del SII, despejando dudas con llamados telefónicos. El resumen es el siguiente:

Registros Nuevos	Total
Registros con actividades únicas	112.846
Registros con códigos no válidos	11
Registros con códigos múltiples	39.493
TOTAL	152.350

El CCAE agrupa procesos continuos de verificación en el año de referencia, desarrollo de bases sectoriales, registros proporcionados por las estadísticas departamentales INE¹³ (económicas y sociales asociadas a empresas), entre otros.

Una vez obtenidas las actividades base, se procedió a imputar la actividad mejorada aplicando el árbol de decisiones elaborado por el Proyecto

¹² Comparables a los utilizados por la Corporación de Fomento de la Producción CORFO.
¹³ Para mayor información Documento Técnico N° 4 "*Proceso de validación a bases de datos departamentales*", julio 2008.

Directorio y revisado por el Comité Técnico del INE. Ésta se almacena y utiliza desde el CCAE, proveniente a su vez, de una matriz de posibilidades de actividad¹⁴. El análisis de la matriz es el siguiente:

Descripción	Total
Registros con clasificación equivalente (única)	66.324
Registros con mas de un código asociado	8.373
Registros con código de 3 y 5 dígitos	1.000
Registros sin código solo glosa	2.354
TOTAL REGISTROS MATRIZ	78.051

Como antecedente se utiliza la razón social, la clasificación del MMA06 y se identifica los registros que el BCCH considera como más importantes respecto del total.

La representación gráfica del árbol de decisiones se encuentra en el Anexo A de este documento.

Ubicación Geográfica¹⁵

La información que alimenta este atributo es proporcionada por el SII. Éste, a su vez, obtiene los antecedentes mayoritariamente de las declaraciones que los contribuyentes realizan a través de Internet. Además, se incorpora y completa este atributo con la información rescatada del consolidado de Directorios de Productos INE y Verificaciones del periodo.

La asignación general de la comuna proviene primero de lo recepcionado desde el SII a través del Catastro de Contribuyentes 2007. Para los registros que no se logró información, se imputó el dato desde el MMA06 en la medida que los registros coinciden. Finalmente se asigna la información rescatada desde la matriz de direcciones que corresponde a las posibilidades de dato provenientes del consolidado de Directorios Departamentales INE, Verificación del periodo y bases de datos sectoriales. En este último proceso se identificó y asignó la oferta única de comuna y se verificaron los registros que presentaban más de una oferta o no contaban con dato para este atributo, identificando en este proceso la ubicación de la casa matriz, para luego asignarlas a los registros respectivos en el Directorio.

¹⁴ La información para registros que resultaron con más de una actividad diferente se encuentra disponible en la tabla denominada "base_final_registros_con_mas_de_una_actividad" de la entrega final. Sobre estos casos se realizó un trabajo de escritorio utilizando el árbol de decisiones y realizando consultas directas a los jefes de productos.

¹⁵ En pro del secreto estadístico y en beneficio de la selección sólo la versión disponible para I&D - INE tiene a disposición el archivo electrónico del DIRECTORIO ficticio con el campo comuna.

Por último y para los registros que no se logró completar información, se buscó en la fuente Patentes Municipales encontrándose un número menor de registros con este atributo.

El SII no utiliza el codificador oficial de la División Político Administrativa, por lo tanto, el procesamiento requirió generar equivalencias entre la codificación de la fuente y la División Político – Administrativa (DPA), oficial utilizada por INE.

Cabe señalar que la ubicación geográfica asignada dependerá de si el declarante informó la comuna donde tributa ó el lugar donde está ubicado el establecimiento ó genera la producción. Esto explica en ocasiones los cambios de la distribución de las ventas por región del país.

En el DIR2007, 158 empresas no cuentan con el atributo y el contenido del campo para estos es 99999, representando una mejora respecto del año anterior de 95,8%. Cabe señalar que estos registros pertenecen a estratos de ventas pequeños e indeterminados.

El resumen de la distribución procesada para este atributo es la siguiente¹⁶:

Año	Total registros	Observación
Registros 2006	779.609	Existe coincidencia con 76.116 registros, pertenecientes a la matriz de direcciones.
Nuevos 2007	151.744	
Nulos	158	
TOTAL DIR2007	931.511	

¹⁶ La matriz de direcciones corresponde a las posibilidades existentes para el dato a partir de fuentes (bases de datos) del tipo departamentales, bases sectoriales y provenientes de verificación.

RESULTADOS ESTADÍSTICOS GENERALES

VENTAS TOTALES ANUALES

N° empresas	Total ventas en pesos	Total ventas en UF
931.511	195.324.191.182.380	10.396.934.385

DISTRIBUCIÓN DE LA SUMA DE LAS VENTAS ANUALES POR SECCIÓN

SECCION	Total ventas en pesos	Total ventas en UF	N° de empresas
A	5.260.310.734.013	280.001.700	91.952
B	755.917.754.891	40.236.835	3.266
C	3.012.168.426.530	160.335.068	3.699
D	64.560.208.128.943	3.436.482.925	63.363
E	7.313.902.561.324	389.312.581	2.241
F	10.194.323.747.314	542.634.860	48.565
G	51.221.843.557.199	2.726.493.546	304.638
H	2.258.761.172.655	120.231.865	35.609
I	14.426.896.130.273	767.930.954	83.997
J	18.439.868.297.668	981.537.922	18.592
K	11.769.389.030.031	626.474.195	110.645
L	276.854.919.839	14.736.743	1.121
M	2.366.472.942.444	125.965.267	9.988
N	1.766.597.807.996	94.034.443	44.673
O	1.630.693.045.988	86.800.352	105.818
P	2.206.929.511	117.473	370
Q	75.181.114	4.002	35
SIN ACTIVIDAD	67.700.814.667	3.603.655	2.939
TOTALES	195.324.191.182.399	10.396.934.385	931.511

DISTRIBUCIÓN DE LA SUMA DE LAS VENTAS ANUALES POR REGIÓN

Región	Total ventas en pesos	Total ventas en UF	Nº de empresas
01	2.341.424.130.629	124.631.941	29.164
02	3.808.249.676.811	202.709.771	26.676
03	1.438.000.906.891	76.543.520	14.544
04	2.253.668.141.984	119.960.767	35.545
05	7.339.833.686.588	390.692.872	96.949
06	5.106.806.377.758	271.830.798	47.721
07	3.327.361.189.279	177.112.501	63.953
08	12.726.379.868.699	677.413.973	99.166
09	2.392.617.138.205	127.356.900	47.505
10	6.577.627.071.611	350.121.286	63.972
11	481.214.088.584	25.614.601	6.744
12	858.748.843.726	45.710.443	10.911
13	146.671.992.618.828	7.807.220.776	388.503
99 Sin región	267.442.806	14.236	158
TOTALES	195.324.191.182.399	10.396.934.385	931.511

DISTRIBUCIÓN DE EMPRESAS POR TAMAÑO SEGÚN ESTRATO DE VENTAS

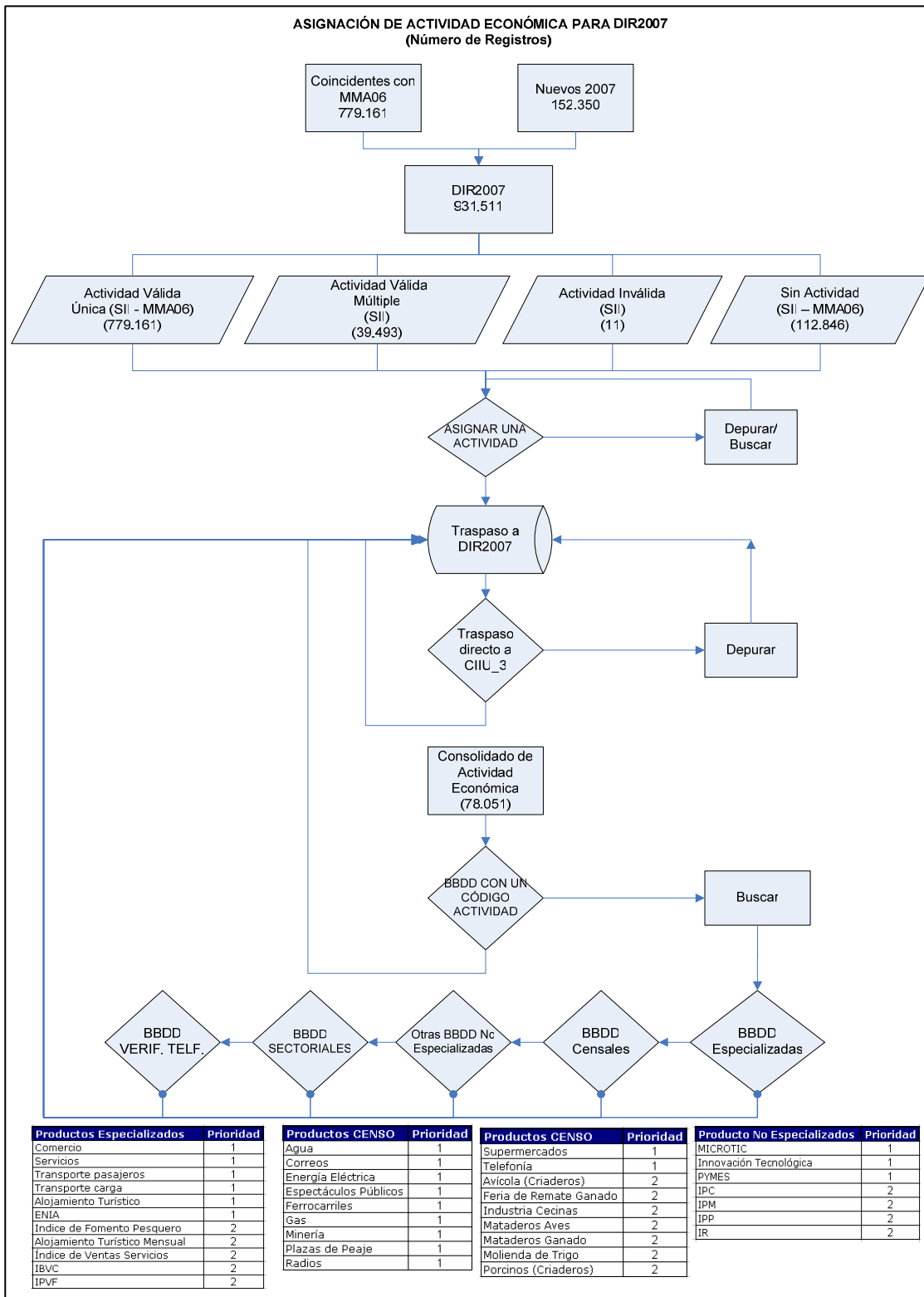
Estrato	Descripción	Total
0	Sin ventas	11
1	Micro, (0 UF; 200 UF]	332.177
2	Micro, (200; 600 UF]	162.940
3	Micro, (600; 2400 UF]	163.235
4	Pequeña, (2400; 5000 UF]	56.151
5	Pequeña, (5000; 10.000 UF]	34.772
6	Pequeña, (10.000; 25.000 UF]	25.948
7	Mediana, (25.000; 50.000 UF]	10.397
8	Mediana, (50.000; 100.000 UF]	6.253
9	Grande, (100.000; 200.000 UF]	3.655
10	Grande, (200.000; 600.000 UF]	2.851
11	Grande, (600.000; 1.000.000 UF]	648
12	Grande, (1.000.000 y más UF]	1.141
20	Venta indeterminada	131.282
21	Negativa	50
	931.511	

DISTRIBUCIÓN DE LAS EMPRESAS POR TIPO DE VENTA

Marca tipo_vta	Descripción	Nº de empresas
1	Venta indeterminada imputada en cero	131.243
2	Venta cero (el calculo de la venta da cero)	11
3	Venta positiva	790.187
4	Venta negativa imputada en cero	50
5	Declarante en moneda extranjera	39
6	Venta indeterminada imputada con venta distinta de cero	9.977
7	Imputación por códigos abultados	4
		931.511

ANEXOS

ANEXO A



ANEXO B

DESCRIPCIÓN CAMPOS ÁMBITO REAL DIR2007

Nombre campo	Descripción
RUT	Rol único tributario
DV_RUT	Dígito verificador del RUT
RAZON_SOCIAL	Nombre de la empresa o persona natural
REGION	Código de la región
PROVINCIA	Código de la provincia
COMUNA	Código de la comuna
SECCION	Código que identifica el primer nivel de clasificación económica CIIU3 (1 letra)
DIVISION	Código que identifica el segundo nivel de clasificación económica CIIU3 (2 dígitos)
GRUPO	Código que identifica el tercer nivel de clasificación económica CIIU3 (3 dígitos)
CIIU3_4	Código que identifica el cuarto nivel de clasificación económica CIIU3 (4 dígitos)
CIIUCL	Código que identifica la adaptación nacional del CIIU3 (6 dígitos)
QUIEBRA	Marca (X) que indica las empresas que se encuentran en quiebra, según Síndico de quiebras.
ESTRATO_VTA	Número que identifica un rango de ventas, según un tramo determinado.
TAMAÑO	Número que identifica un rango de ventas, según un grupo de estratos de ventas determinados.
MATRIZ	Marca (X) que indica las empresas que se encuentran en el consolidado de mejoras de actividades económicas.
NO_PERT_SII	Marca (X) que indica las empresas que según la página del SII, son "Rut no existe" o "Rut consultado sin inicio de actividades", pero que presentan declaración.

DESCRIPCIÓN CAMPOS ÁMBITO FICTICIO DIR2007

Nombre campo	Descripción
RUT_FICTICIO	Número que permite innominar (ocultar) el RUT real
SECCION	Código que identifica el primer nivel de clasificación económica CIIU3 (1 letra)
DIVISION	Código que identifica el segundo nivel de clasificación económica CIIU3 (2 dígitos)
GRUPO	Código que identifica el tercer nivel de clasificación económica CIIU3 (3 dígitos)
CIIU3_4	Código que identifica el cuarto nivel de clasificación económica CIIU3 (4 dígitos)
CIIUCL	Código que identifica la adaptación nacional del CIIU3 (6 dígitos)
REGION	Código de la región
PROVINCIA	Código de la provincia
COMUNA	Código de la comuna
VTA_MES_ENERO (\$)	Valor mensual de ventas estimadas en pesos enero
VTA_MES_FEBRERO (\$)	Valor mensual de ventas estimadas en pesos febrero
VTA_MES_MARZO (\$)	Valor mensual de ventas estimadas en pesos marzo
VTA_MES_ABRIL (\$)	Valor mensual de ventas estimadas en pesos abril
VTA_MES_MAYO (\$)	Valor mensual de ventas estimadas en pesos mayo
VTA_MES_JUNIO (\$)	Valor mensual de ventas estimadas en pesos junio
VTA_MES_JULIO (\$)	Valor mensual de ventas estimadas en pesos julio
VTA_MES_AGO (\$)	Valor mensual de ventas estimadas en pesos agosto
VTA_MES_SEPTIEMBRE (\$)	Valor mensual de ventas estimadas en pesos septiembre
VTA_MES_OCTUBRE (\$)	Valor mensual de ventas estimadas en pesos octubre
VTA_MES_NOVIEMBRE (\$)	Valor mensual de ventas estimadas en pesos noviembre
VTA_MES_DICIEMBRE (\$)	Valor mensual de ventas estimadas en pesos diciembre
VTA_ANUAL (\$)	Valor Anual de ventas estimadas en pesos
VTA_ANUAL UF (\$) ¹⁷	Valor Anual de ventas estimadas en unidades de fomento (promedio 2007 18.786,71)

¹⁷ Cabe señalar que el procedimiento utilizado por Directorio para el cálculo de la U.F promedio (y que se aplica en la estratificación) consiste en el promedio de los promedios por mes. También existe la posibilidad de obtener este valor a través de la sumatoria del monto de la U.F. diaria dividida por los días hábiles. El resultado de ambos procesos en grandes volúmenes de datos no es el mismo.

MARCA_TIPO_V TA	Número que identifica los diferentes tipos de ventas según marca determinada.
MATRIZ	Marca (X) que indica las empresas que se encuentran en el CCAE.