



**MANUAL DE USUARIOS DE BASES DE DATOS
VIII ENCUESTA DE PRESUPUESTOS FAMILIARES
(STATA)**

DEPARTAMENTO DE PRESUPUESTOS FAMILIARES (DPF)

INSTITUTO NACIONAL DE ESTADÍSTICAS

Junio / 2018

INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Presupuestos Familiares

VIII ENCUESTA DE PRESUPUESTOS FAMILIARES

Período de Información: julio 2016 - junio 2017

Manual de Usuarios de Bases de Datos – VIII Encuesta de Presupuestos Familiares (Stata)

Junio de 2018

El uso de un lenguaje que no discrimine ni marque diferencias entre hombres y mujeres ha sido una preocupación en la elaboración de este documento. Sin embargo, y con el fin de evitar la sobrecarga gráfica que supondría utilizar en castellano "o/a" para marcar la existencia de ambos sexos, se ha optado por utilizar -en la mayor parte de los casos- el masculino genérico, en el entendido de que todas las menciones en tal género representan siempre a hombres y mujeres, abarcando claramente ambos sexos.

Dirección: Av. Presidente Bulnes 418, Santiago de Chile

Casilla de correo: 498 – Correo 3

Teléfono: 562 2892 4000

Sitio web: www.ine.cl

Correo electrónico: ine@ine.cl

Facebook: @ChileINE

Twitter: @INE_Chile

ÍNDICE

1 INTRODUCCIÓN	5
2 ACCESO Y SOLICITUD DE LOS DATOS.....	7
3 ESTRUCTURA DE LOS DATOS (BASES).....	8
3.1 Captura de la información	8
3.2 Base de personas VIII EPF	9
3.3 Base de gastos VIII EPF	10
3.4 Base de Clasificación del Consumo Individual por Finalidades (CCIF) VIII EPF	12
4 INDICACIONES PARA EL USO DE LAS BASES DE DATOS	14
4.1 Especificación del diseño muestral complejo	14
4.1.1 Descripción del diseño muestral.....	14
4.1.2 Variables que identifican el diseño muestral complejo.....	15
4.1.3 Plan de muestreo	16
4.2 Utilización de factores de expansión	16
5 INDICADORES Y TABULADOS.....	18
5.1 Códigos especiales de la encuesta	19
5.2 Hogares y personas encuestados	20
5.2.1 Cantidad de hogares	20
5.2.2 Número personas integrantes del hogar	21
5.2.3 Edad de las personas integrantes del hogar	22
5.2.4 Sexo.....	22
5.2.5 Educación.....	22
5.3 Vivienda	24
5.3.1 Tenencia de vivienda principal.....	24
5.3.2 Tenencia de la vivienda secundaria	25
5.4 Ocupación	25
5.4.1 Población según su condición de actividad económica.....	25
5.4.2 Categoría en la ocupación por sexo	26
5.5 Ingresos y gastos	26
5.5.1 Ingresos.....	26
5.5.2 Gasto	48

5.5.3 Gasto e ingreso según características del sustentador o sustentadora principal del hogar	53
6 QUINTILIZACIÓN POR INGRESO DE LAS BASES DE DATOS	56
7 PRECISIÓN DE LOS PRINCIPALES INDICADORES	64
7.1 Grados de libertad	65
7.2 Interpretación	72
8 BIBLIOGRAFÍA.....	74
9 ANEXOS.....	75
Anexo 1. Variables contenidas en la base de datos de personas de la VIII EPF	75
Anexo 2. Componentes del ingreso	93
Anexo 3. Comandos útiles	95

1 INTRODUCCIÓN

El presente manual está dirigido a los usuarios de las bases de datos de la VIII Encuesta de Presupuestos Familiares (EPF), cuyo trabajo de campo se desarrolló entre julio de 2016 y junio de 2017.

Su objetivo es orientar en el uso de las bases de datos y en la construcción de los principales indicadores de la VIII EPF para que, con este insumo, investigadores y usuarios en general puedan replicar algunos de los tabulados oficiales, generar gráficos y estudiar la precisión de sus estimaciones. Además, busca entregar herramientas para que las personas puedan adaptar las programaciones aquí ofrecidas y alcanzar sus propios objetivos de investigación.

La asistencia se proporciona mediante el uso del paquete estadístico Stata¹ (en su versión 13), que permite revisar desde la configuración inicial para trabajar con un diseño muestral complejo hasta la construcción de variables y el desarrollo de gráficos. No obstante, además del formato Stata (.dta), las bases de datos de la VIII EPF se encuentran disponibles para su descarga y análisis en formato SPSS (.sav) en el micrositio de la encuesta².

Los contenidos del presente documento están estructurados en ocho capítulos, incluida la introducción (capítulo 1). En el capítulo 2 se instruye a los usuarios sobre cómo acceder a las bases de datos de la encuesta y cómo solicitarlas. A continuación se presentan y describen las 3 bases de datos oficiales de la VIII EPF (capítulo 3), señalando los tipos de información contenida en ellas, para luego presentar, en el capítulo 4, las indicaciones de uso necesarias, en atención al diseño muestral de la encuesta y sus factores de expansión.

A partir del capítulo 5 se ofrecen programaciones para construir variables, tabular información y generar gráficos de interés, de acuerdo con temáticas de análisis general con las bases (hogares y personas, educación, vivienda, ingresos, gastos, etc.). En el capítulo 6 se aborda la quintilización de las bases de datos que se realiza para generar los tabulados oficiales y, finalmente, en el capítulo 7 se presentan las rutinas para calcular los indicadores de precisión y confiabilidad de los principales indicadores de gasto e ingresos y se señalan precauciones que es necesario tener en cuenta, desde un punto de vista estadístico, al momento de realizar interpretaciones.

¹ Stata es un paquete estadístico creado por StataCorp. Para más información respecto a este *software*, visitar www.stata.com

² Si bien la encuesta se pone a disposición de los usuarios en formato Stata y SPSS, estos pueden transformar las bases de datos para utilizar los *software* que más les acomoden (R, SAS, Matlab, EViews, entre otros).

El Anexo N°2, al final de este manual, incluye una tabla que describe los componentes del ingreso, esquematizando los conceptos utilizados en el apartado 5.5.1 y, a continuación, el Anexo N°3 ofrece a los usuarios una compilación de comandos útiles empleados a lo largo del documento.

Si bien esta publicación desarrolla el cálculo de algunos de los principales resultados de la encuesta, es importante comentar que no busca presentar todos los tabulados de la VIII EPF ni explotar todos los usos de la encuesta, sino que entregar guías respecto al uso de los datos con el fin de que las personas puedan explorar las bases de datos de la EPF y utilizar la información entregada por la encuesta de la mejor forma posible. Además, es necesario considerar que las programaciones propuestas entregan una guía para trabajar la encuesta, dado que no existe una única forma de trabajar y obtener los resultados presentados.

2 ACCESO Y SOLICITUD DE LOS DATOS

Las bases de datos de la VIII EPF se encuentran publicadas en el portal web del Instituto Nacional de Estadísticas (INE), específicamente en el micrositio de la encuesta (www.ine.cl/epf), junto con otros documentos metodológicos asociados al estudio³. Para acceder a los datos, los usuarios deben ir al portal y descargar las bases de datos directamente desde el sitio web mencionados, las que están disponibles en formato Stata y SPSS. También pueden requerir la información directamente en las oficinas de la institución, ubicadas en la avda. Presidente Bulnes 418, Santiago de Chile.

Quienes deseen información detallada o tengan consultas sobre el uso de la base de datos, pueden realizar las solicitudes y preguntas a través del portal del Sistema Integral de Información y Atención Ciudadana (SIAC) del instituto (www.ine.cl/SIAC).

La información adicional requerida por los usuarios puede contemplar diversos módulos de la encuesta y niveles de desagregación. Sin embargo, la entrega de esta información se encuentra sujeta a los siguientes criterios:

- Disponibilidad de la información solicitada.
- Protección de la confidencialidad de la información, de acuerdo con el *secreto estadístico* establecido en la Ley 17.374.

Al momento de utilizar la información de la encuesta es importante tener en cuenta que los datos de ingreso y gasto solo tienen representatividad estadística con temporalidad anual a nivel de áreas de estimación de la muestra⁴. De este modo, si bien el trabajo de campo de la encuesta se realiza a un nivel más desagregado en términos geográficos (manzana, comuna, región), los estimadores no son estadísticamente significativos a estos niveles.

³ Se recomienda a los usuarios acompañar el análisis de las bases de datos con la lectura de la Metodología VIII EPF, el documento de Principales Resultados VIII EPF, las libretas utilizadas, los manuales de trabajo de campo (encuestadores y encargados de grupo), manual de supervisores y del equipo de codificación, los metadatos de la encuesta y la revisión de los tabulados oficiales.

⁴ La representatividad geográfica de la muestra comprende Gran Santiago, Resto de capitales regionales (en su conjunto) y Total de capitales regionales (Gran Santiago + Resto de capitales regionales). Para una explicación más detallada, revisar la Metodología de la VIII EPF (Eurostat, 2014), disponible en el micrositio de la encuesta.

3 ESTRUCTURA DE LOS DATOS (BASES)

La información de la VIII EPF se encuentra publicada en tres bases de datos. La **“BASE PERSONAS VIIIIEPF”** (en adelante, *Base de personas*) contiene la caracterización sociodemográfica y económica de los hogares encuestados y sus integrantes. La **“BASE GASTOS VIIIIEPF”** (en adelante, *Base de gastos*) contiene información desagregada del gasto por producto de los hogares. La tercera base de datos, **“CCIF VIIIIEPF”** (en adelante, *Base CCIF*), es la Base de Datos de la Clasificación del Consumo Individual por Finalidades (CCIF), que corresponde a la adaptación del clasificador de Naciones Unidas que fue empleado para clasificar los bienes y servicios capturados por la encuesta. Este clasificador se utiliza hasta el quinto nivel de desagregación⁵; los primeros tres niveles (que describen el nivel de división, grupo y las clases de producto) se adoptaron según el clasificador internacional, mientras que los siguientes dos (que respectivamente describen las subclases y los productos) fueron adaptados a la realidad nacional.

En tanto, para facilitar la accesibilidad y comprensión de las variables pertenecientes a cada base de datos, en el micrositio de la encuesta se encuentran a disposición los metadatos de la VIII EPF, que contienen el nombre de cada variable, su etiqueta en base de datos, su definición y la fuente de referencia (libreta o libretas) utilizada para la construcción de cada variable⁶.

3.1 Captura de la información

La VIII EPF se compone de seis libretas que capturan la información de los hogares. Las libretas se pueden clasificar según las siguientes tres temáticas:

- Características sociodemográficas del hogar y sus integrantes, y de la vivienda principal y secundaria, capturadas en la libreta Registro de Personas en el Hogar.
- Gasto del hogar, que se captura en cinco libretas: Libreta de Gastos Individuales, Libreta de Gastos del Hogar, Libreta de Gastos del Recuerdo, Libreta de Ingresos⁷ y Libreta de Gastos en Servicios Financieros y Seguros.
- Ingresos del hogar, capturados a través de la Libreta de Ingresos.

⁵ La clasificación CCIF cuenta con cinco niveles de desagregación: división (primer nivel), grupo (segundo nivel), clase (tercer nivel), subclase (cuarto nivel) y producto (quinto nivel).

⁶ Revisar los metadatos de la encuesta en el micrositio www.ine.cl/epf.

⁷ Si bien la Libreta de Ingresos captura principalmente las partidas asociadas a los ingresos con los que cuenta el hogar, también recoge una pequeña proporción del gasto de los hogares, asociada a gastos en transferencias emitidas, comisiones en administradoras de pensiones e ingresos en especie.

El contenido detallado de las libretas que componen la encuesta se puede consultar en la Metodología de la VIII EPF (INE Chile, 2018a), específicamente en el apartado 5.3 *sobre Instrumentos de Captura de Información*, y en el Manual de Trabajo de Campo para Encuestadores, Tomos I y II (INE Chile, 2018c).

3.2 Base de personas VIII EPF

La Base de personas contiene 247 variables, mientras que el número de observaciones totales de la base es de 48.308 (personas), las que integran los 15.239 hogares entrevistados correspondientes a 14.999 viviendas entrevistadas⁸. La Base de personas incluye información desagregada de las características sociodemográficas a nivel del hogar y de sus integrantes. Las variables de esta base se construyen⁹ a partir de la combinación de información de los distintos cuestionarios de la encuesta. A continuación se detalla, según temáticas, la información de las variables de esta base de datos.

- **Variables sociodemográficas:** provienen de la información obtenida mediante el Registro de Personas en el Hogar. Esta sección cuenta con variables de identificación del hogar, un identificador genérico de las personas que integran del hogar (FOLIO¹⁰), las variables que caracterizan sus relaciones de parentesco, el estado civil y la identificación de quien cumple el rol de sustentador principal y de la o las personas que cumplen el rol de administrador de gastos en cada hogar, además de las variables de educación y escolaridad de quienes integran el hogar.
- **Variables sobre la condición de actividad económica, ingresos y situación sobre la afiliación de las personas en los sistemas de salud y previsión:** provienen de la información obtenida en el Registro de Personas en el Hogar y en la Libreta de Ingresos. Estas variables permiten identificar la población ocupada, el tipo de ocupación de la actividad principal y la situación de afiliación y cotización en los sistemas de salud y/o previsional. Además, la base presenta las variables de los ingresos declarados por cada persona y las variables calculadas que cuantifican los diferentes tipos o partidas de

⁸ Una vivienda puede contener más de un hogar. La EPF busca entrevistar a todos los hogares que habitan al interior de una vivienda. Para más detalle respecto a estos conceptos, revisar la Metodología de la VIII EPF (INE Chile, 2018a).

⁹ La “**BASE PERSONAS VIII EPF**” contiene la suma del gasto e ingreso total del hogar con y sin arriendo imputado, por lo que posee información de todos los cuestionarios de la encuesta.

¹⁰ La variable FOLIO sirve de enlace entre la base de personas y la de gasto. En la primera se repite para cada integrante de cada hogar, mientras que en la segunda, para cada gasto que realice el hogar a nivel de producto.

ingresos percibidos, construidas a partir de la información reportada en la Libreta de Ingresos.

- **VARIABLES DE GASTO:** provienen de las distintas libretas en las que se mide el gasto. Para generar las variables de gasto total del hogar, ya sea con arriendo imputado o sin arriendo imputado (*per cápita* o a nivel hogar), se considera la información de todas las libretas. Además, se agregan variables de gasto en educación, que son construidas a partir del módulo de Educación (ED) del Registro de Personas en el Hogar y de la información reportada en la Libreta de Gastos del Hogar. Se consideran los gastos en matrículas, aranceles, centros de padres y cuotas de curso.
- **VARIABLES SOBRE LA VIVIENDA PRINCIPAL Y SECUNDARIA:** provienen del Registro de Personas en el Hogar. Este set de variables refieren al tipo de vivienda principal, al tipo de tenencia de la vivienda principal y secundaria, además de la comuna donde se encuentra la vivienda secundaria.
- **VARIABLES QUE IDENTIFICAN SI EL HOGAR REALIZA GASTOS ASOCIADOS A LA VIVIENDA PRINCIPAL Y SECUNDARIA:** provienen de información obtenida en el Registro de Personas en el Hogar y la Libreta de Gastos del Hogar. En esta sección se encuentran variables categóricas que denotan la existencia o no de diferentes tipos de gastos asociados a la vivienda principal y secundaria¹¹, como por ejemplo, el pago de servicios, agua, electricidad, gastos comunes, entre otros.

Para más información respecto a las variables y su construcción, consultar los metadatos de la encuesta, disponibles en el micrositio www.ine.cl/epf. Además, en el Anexo N°1 se presenta una descripción de las variables incorporadas en la base de datos de personas de la encuesta.

3.3 Base de gastos VIII EPF

Los datos presentados en esta base contemplan la información de los gastos realizados por el hogar a nivel de bienes y servicios, de acuerdo con la Clasificación del Consumo Individual por Finalidades (CCIF) utilizada por la VIII EPF.

La base incluye 13 variables y 1.064.239 observaciones. El elevado número de observaciones se explica por su estructura: una fila en la Base de gastos corresponde al gasto mensualizado que realizó un hogar en un tipo de producto. Entonces, la base tendrá tantas filas como tipos de

¹¹ Estas variables son depuradas y analizadas con la información de la Libreta de Gastos del Hogar.

productos en los que el hogar declaró haber realizado gastos en el período de referencia, situación que se repetirá para cada hogar (FOLIO) presente en la Base de personas.

La Base de gastos contiene información compilada de todas las libretas de la encuesta. Es decir, los registros de esta base se construyen al combinar información de las seis libretas¹², integrando en una misma fila los gastos provenientes de todas ellas.

Un hogar puede haber gastado en un mismo producto (pan, por ejemplo) más de una vez durante la quincena en la cual fue entrevistado. Que el gasto se encuentre integrando en una misma fila significa que el gasto del hogar contiene la suma del producto declarado. Por lo tanto, en la base de gastos, el producto asociado al gasto en pan, por ejemplo, contendrá la suma de todos los gastos declarados por los integrantes del hogar en dicho producto.

Esta base incluye las siguientes variables¹³:

- **ZONA:** área de estimación de la encuesta.
- **FOLIO:** identificador único del hogar que permite vincular la Base de personas con la Base de gastos.
- **FE:** factor de expansión.
- **VARSTRAT** y **VARUNIT:** estrato y conglomerado de varianza (VARSTRAT y VARUNIT, respectivamente, utilizados para especificar el diseño muestral complejo).
- **CCIF:** código del clasificador de gastos en bienes y servicios adaptado para la VIII versión de la encuesta. Se obtiene al concatenar los distintos niveles del clasificador (D.G.C.SC.P).
- **D, G, C, SC** y **P:** distintos niveles de clasificación de la CCIF que corresponden a División (D), Grupo (G), Clase (C), Subclase (SC) y Producto (P). Permiten la compilación de la Base de gastos en los diferentes niveles y la combinación con la estructura de la Base CCIF, que entrega la glosa teórica de cada nivel. En la base de gastos estas variables están desplegadas a nivel de productos. En el apartado 5.5.2 se entrega una guía para combinar ambas bases de datos.
- **GASTO:** gasto total del hogar por producto.
- **GLOSA:** glosa teórica del clasificador CCIF adaptado para la VIII EPF.

¹² En el Registro de Personas en el Hogar solo se registra el arriendo imputado, lo que se analiza con lo declarado en la Libreta de Gastos del Hogar. En la Libreta de Ingresos se registran las transferencias emitidas en especies.

¹³ Muchas de las variables de esta base de datos son de tipo texto para mantener el largo de cada una y poder replicar los códigos CCIF.

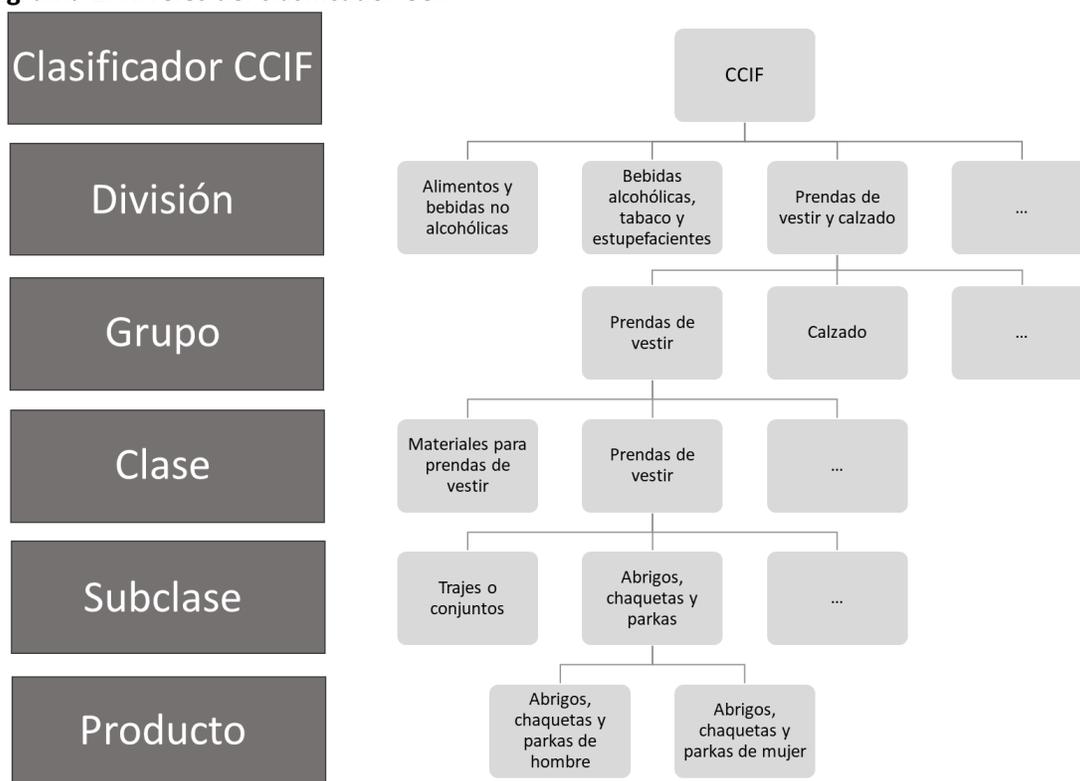
3.4 Base de Clasificación del Consumo Individual por Finalidades (CCIF) VIII EPF

Esta base presenta las glosas teóricas de la Clasificación del Consumo Individual por Finalidades (CCIF), según la cual se ordena la información en la Base de gastos.

La CCIF es una clasificación funcional del Sistema de Cuentas Nacionales (SCN) que estandariza los bienes y servicios capturados por la encuesta. Se utiliza desagregado hasta nivel de producto (quinto nivel de desagregación)¹⁴. Los primeros tres niveles del clasificador (hasta nivel de clase) se adoptaron según el clasificador internacional, mientras que los siguientes dos niveles fueron adaptados a la realidad nacional (subclase y producto). La adaptación fue realizada por el INE con el apoyo de instituciones externas¹⁵.

El siguiente diagrama muestra un ejemplo del funcionamiento del clasificador CCIF con sus diferentes niveles.

Diagrama 1: Niveles del clasificador CCIF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

¹⁴ Los cinco niveles se encuentran jerarquizados de la siguiente forma: división (primer nivel), grupo (segundo nivel), clase (tercer nivel), subclase (cuarto nivel) y producto (quinto nivel).

¹⁵ La adaptación de la CCIF contó con el apoyo de Cepal, del Ministerio de Desarrollo Social y del Ministerio de Salud.

El clasificador de la VIII EPF tiene 12 divisiones, 59 grupos, 126 clases, 285 subclases y 1.186 productos. En coherencia con esta estructura, la base CCIF puesta a disposición cuenta con siete variables y un total de 1.668 observaciones.

La Base de datos de la CCIF contiene una observación según el nivel de la clasificación propuesto. Siguiendo el ejemplo planteado en el diagrama, los códigos asociados al ejemplo se presentan a continuación:

Tabla 1. Ejemplo de los niveles de clasificación de la CCIF

Nivel de desagregación	Código CCIF	Glosa CCIF
División (D)	03.0.0.00.00	Prendas de vestir y calzado
Grupo (G)	03.1.0.00.00	Prendas de vestir
Clase (C)	03.1.2.00.00	Prendas de vestir
Subclase (SC)	03.1.2.02.00	Abrigos, chaquetas y parkas
Producto (P)	03.1.2.02.02	Abrigos, chaquetas y parkas de mujer

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La base de datos de la CCIF incluye las siguientes variables:

- **D, G, C, SC y P:** son los distintos niveles de clasificación de la CCIF y corresponden a División (D), Grupo (G), Clase (C), Subclase (SC) y Producto (P). Permiten la compilación de la Base de gastos en los diferentes niveles y la combinación con la estructura de la Base CCIF, que entrega la glosa teórica de cada nivel. En la base de CCIF estas variables están presentes según los cinco niveles de desagregación del clasificador. En el apartado 5.5.2 se entrega una guía para combinar ambas bases de datos.
- **CCIF:** códigos del clasificador de gastos en bienes y servicios adaptado para la VIII versión de la encuesta. Se obtiene al concatenar los distintos niveles del clasificador (D, G, C, SC y P).
- **GLOSA:** glosa teórica del clasificador CCIF adaptado para la VIII EPF.

4 INDICACIONES PARA EL USO DE LAS BASES DE DATOS

En el presente apartado se provee al usuario de elementos para un uso adecuado de las bases de datos de la VIII EPF. En esta versión de la encuesta se incluyen las variables Estrato de varianza y Conglomerado de varianza al que pertenece cada hogar, tanto en la Base de personas como en la Base de gastos, lo que permite realizar estimaciones considerando el muestreo complejo utilizado en la encuesta y, de este modo, tener la posibilidad de conocer la precisión de las estimaciones para el nivel en el que se esté trabajando.

A continuación se presenta (1) la especificación del diseño muestral complejo, incluyendo una descripción de este, la definición de las variables de estrato y conglomerado de varianza, además de su aplicación a través del plan de muestreo complejo en Stata, y (2) la definición y utilización de los factores de expansión de la encuesta.

4.1 Especificación del diseño muestral complejo

Cuando en una encuesta las unidades de observación son seleccionadas en una etapa, mediante un muestreo aleatorio simple, cada unidad tiene la misma probabilidad de ser seleccionada y, por tanto, cada una representa a una misma proporción de la población que busca representar. Cuando el diseño muestral es complejo (estratificado y multietápico), la probabilidad de incluir a cada unidad de observación es diferente según su procedencia, por lo que representará a un número distinto de unidades en la muestra (Observatorio Social, 2012).

La Encuesta de Presupuestos Familiares cuenta con un diseño muestral complejo, lo que implica ciertos antecedentes que deben ser considerados al momento de calcular la precisión de sus estimadores y al interpretar las ponderaciones realizadas mediante su factor de expansión.

El diseño muestral del estudio se describe en extenso en la metodología publicada de la encuesta (INE Chile, 2018a), por lo que en este documento solo se enunciarán sus características principales.

4.1.1 Descripción del diseño muestral

El diseño muestral de la encuesta es:

- **Probabilístico:** cada unidad en la muestra tiene una probabilidad de selección conocida y distinta de cero.
- **Estratificado:** las unidades de observación se distribuyen en estratos que buscan minimizar la variabilidad de las características de interés en su interior y maximizarla en relación con otros estratos. En la VIII EPF el estrato se define por la interacción entre la comuna, la clasificación socioeconómica de la manzana y el grupo de tamaño (respecto a la cantidad de viviendas que contiene) al que pertenece cada manzana.

- **Bietápico:** al interior de cada estrato se seleccionan conglomerados que corresponden a la unidad primaria de muestreo (manzanas), y al interior de cada conglomerado se seleccionan las unidades secundarias de muestreo, que son las viviendas.

Considerando lo anterior, para establecer el diseño muestral complejo se requiere introducir tres variables relevantes: conglomerados de varianza, estratos de varianza y factor de expansión (VARUNIT, VARSTRAT, FE- según se identifican en las Bases de datos). En las secciones siguientes se explica la importancia de considerarlas entre los insumos que se requieran para generar los análisis de interés del usuario.

4.1.2 Variables que identifican el diseño muestral complejo

Para que en un muestreo estratificado la estimación de la varianza sea factible, se deben seleccionar como mínimo dos conglomerados por estrato. Adicionalmente, para evitar estimaciones de varianza volátiles, se recomienda seleccionar un número mínimo de unidades muestrales (es decir, viviendas) por conglomerado.

Dado que es frecuente que los estratos y conglomerados no contengan el mínimo de unidades requeridas para la estimación de la varianza, ya sea por razones de diseño o por falta de respuesta, para la VIII EPF se reagruparon algunos de los conglomerados y estratos definidos en el diseño muestral¹⁶. El objetivo de esta reagrupación fue estabilizar el cálculo de la varianza para la estimación de los indicadores principales del estudio y de sus indicadores de confiabilidad, considerando también el diseño complejo.

En los párrafos siguientes se explica qué son los conglomerados y estratos de varianza y cuáles fueron los criterios utilizados para su construcción:

Conglomerados de varianza, pseudoconglomerados o VARUNIT: corresponden a una reagrupación de los conglomerados originales¹⁷, para cumplir con el requisito de cantidad de unidades en su interior, el que se estableció en un rango entre 39 y 60 viviendas¹⁸. Además, cada conglomerado de varianza debe estar contenido en solo un estrato de varianza.

¹⁶ El detalle se encuentra en la Metodología, en el capítulo 6 sobre el *Diseño Estadístico de la Encuesta* (INE Chile, 2018a).

¹⁷ Los conglomerados originales son combinaciones de las variables comuna, clasificación socioeconómica y grupo de tamaño de las manzanas que se utilizan para el diseño muestral.

¹⁸ Las unidades de muestreo en segunda etapa son las viviendas, por ello, el tamaño del pseudoconglomerado se establece con el número mínimo y máximo de viviendas. A nivel hogar, el rango total se encuentra entre 39 a 65 hogares.

Estratos de varianza, pseudoestratos o Varstrat: son una reagrupación generada a partir de los estratos de muestreo y los conglomerados de varianza. La reagrupación se realiza con el objetivo de cumplir con el estándar definido para asegurar un mínimo de variabilidad al interior de cada estrato, que en este caso es de al menos 2 conglomerados en su interior.

Cada uno de los hogares contenidos en la base de datos corresponde a uno de los 304 conglomerados de varianza, los que, a su vez, están contenidos en alguno de los 56 estratos de varianza. Las variables que contienen esta información tanto en la Base de Personas como en la Base de Gastos de la VIII EPF son las siguientes:

Etiqueta	Nombre
Conglomerado varianza	VARUNIT
Estrato varianza	VARSTRAT

4.1.3 Plan de muestreo

Utilizando las variables que identifican el diseño muestral complejo (VARUNIT y VARSTRAT) y el factor de expansión (FE, que será descrito en el apartado siguiente), es posible especificar el plan de muestreo complejo. De no realizarse este paso, el programa entenderá que se trata de un muestreo aleatorio simple.

Programación 1: identificación del diseño muestral complejo

```
svyset VARUNIT [pw=FE], strata(VARSTRAT) singleunit(certainty)
```

En Stata, el plan de muestreo (Csplan) se declara antes de realizar las estimaciones, como se revisará en el apartado 7.

Cada vez que se solicite una estimación, si se requiere conocer intervalos de confianza, niveles de error, coeficientes de variación, etc., debe utilizarse el comando `svy`, como se revisará en el capítulo 7.

4.2 Utilización de factores de expansión

El factor de expansión se interpreta como la cantidad de unidades en la población a las que representa una unidad de la muestra. De acuerdo con el diseño muestral de la encuesta, el factor de expansión se calcula en función del total de manzanas y viviendas seleccionadas en cada estrato, así como también del número de manzanas totales en el estrato (según información reportada por el Marco Muestral de Viviendas) y del número de viviendas particulares enumeradas.

La VIII EPF, al igual que sus versiones anteriores, cuenta con un factor de expansión, a nivel de hogar, el que debe ser idéntico para quienes residen en una misma vivienda.

El factor de expansión de esta versión de la encuesta cuenta con innovaciones respecto al de la VII EPF. Destacan, por ejemplo, el ajuste por no respuesta mediante un modelo predictivo y una calibración por población realizada con el método *raking*. Este método permite ajustar el factor de expansión a nivel vivienda, considerando las variables demográficas sexo y edad en tres tramos para ajustar las estimaciones de la encuesta a *stocks* poblacionales más precisos. Para conocer más detalles sobre la construcción de los factores de expansión de la VIII EPF, se recomienda consultar el documento metodológico de la encuesta (INE Chile, 2018a).

La variable que contiene el factor de expansión, presente tanto en la Base de Personas como en la Base de Gastos de la VIII EPF, es la siguiente:

Etiqueta	Nombre
Factor de expansión	FE

El factor de expansión debe estar especificado en cada análisis que se le solicite al *software*, con el fin de que este pondere los registros y se obtengan resultados a nivel poblacional de la encuesta. Si no se incorpora el factor de expansión en los análisis, se trabajaría la encuesta considerando los resultados solo de la muestra. Para ejemplificar esto, en la Programación 2 se muestra la especificación del factor de expansión, a través de un tabulado para la variable SEXO. Si no se agrega la indicación entre corchetes, el programa no pondera los casos por el factor.

Programación 2: uso de factor de expansión

```
// Uso del factor de expansión // FE
table SEXO [pweight=FE], row f(%10,0fc)
```

5 INDICADORES Y TABULADOS

A continuación se ofrecen alternativas para la construcción de algunos tabulados que se pueden obtener con las bases de datos de la VIII EPF. Tal como se señaló en la introducción, es importante recalcar que el presente documento trata de ejemplificar mediante algunos gráficos y tabulados el tipo de información que se pone a disposición de la ciudadanía, pero en ningún caso el objetivo es mostrar todo lo que se puede hacer o todo lo que se publica a partir de esta encuesta. Además, las programaciones provistas no son la única vía para obtener las mismas estimaciones o resultados. En este sentido, este manual busca guiar a los usuarios en el manejo y tratamiento de la información, con el fin de reducir las complejidades de la base de datos, sin pretender ser un catálogo exhaustivo de todos los comandos o procedimientos con los que es posible realizar las estimaciones ejemplificadas.

Se advierte que si bien gran parte de las programaciones que se ofrecen a lo largo del documento pueden ejecutarse a través de la consola de Stata (la interfaz principal del programa), están programadas en un *do file*, por lo que se recomienda ejecutarlas a través de dicho ambiente de programación¹⁹.

Otro aspecto técnico a considerar es la interacción entre este documento (formato PDF) y el ambiente de programación Stata. Para una buena compatibilidad en la acción de copiar las programaciones del documento y pegarlas en un *do file*, este debe ser abierto mediante una versión estándar de PDF. De esta manera, en el *do file* el texto se desplegará de la misma forma en que estaba estructurado en el *do file* original en que fue construido. De lo contrario, de abrirse este archivo a través de plataformas de edición de PDF, como Nitro PDF Reader o PDF Complete, el ordenamiento de la programación respetará la estructura en que se visualiza en PDF, que en muchos casos impedirá que esta se ejecute correctamente en el *software* estadístico, teniendo que ajustar las líneas de programación propuestas.

Ya en el programa, el primer paso para comenzar a trabajar en forma directa en las bases es definir una ruta dentro del equipo en donde estas se encuentran alojadas. Idealmente, para efectos del estudio de este manual, se recomienda que estén alojadas en una misma carpeta, para así optimizar el acceso a cada una cada vez que sea necesario. Para esto se utiliza el comando *cd*, y luego para llamar a cada base, el comando *use*. Por ejemplo:

¹⁹ Hay elementos presentes en las programaciones como los comentarios señalados con “//”, que el *do file* reconoce pero la consola no, generándose un error.

Programación 3: Especificación de ruta de acceso

```
// Ruta de acceso a la carpeta
cd "C:\VIII EPF\BBDD" //Especificar la dirección de acceso a los archivos
// Para abrir una base de datos
use "BASE_PERSONAS_VIII_EPF.dta", clear
```

El comando *clear* generalmente se agrega para limpiar los datos que están abiertos. Si ya se tenía una base abierta en forma previa, solo empleando este comando se podrá abrir la base. Una vez realizado el paso indicado en la Programación 3, ya estará abierta la Base de personas. Para abrir la Base de gastos o la Base CCIF, solo debe reemplazarse el nombre de la base tras el comando *use*, indicándolo entre comillas (“ ”).

5.1 Códigos especiales de la encuesta

Antes de comenzar a trabajar con las bases de datos de la encuesta, es importante tener en consideración que existen los siguientes códigos especiales en el registro de las variables:

- -99: código que corresponde a la categoría “No sabe”
- -88: código que corresponde a la categoría “No responde”
- -77: código que corresponde a la categoría “No aplica”

La siguiente programación presenta el tabulado de la pregunta AE01, de la base de datos de personas que contiene los tres códigos especiales:

Programación 4: Ejemplo de uso de códigos especiales

```
// Categorías de respuesta a la pregunta AE01 "Durante la semana pasada, ¿tuvo
un trabajo, de al menos una hora, por el cual recibió o recobirá pago en dinero
o en especies?"
table AE01, f(%9,0fc) //dato de la muestra
table AE01 [pweight=FE], f(%9,0fc) //dato expandido al total poblacional
```

Al ejecutar la programación anterior es posible observar que, a nivel muestral, al responder la pregunta AE01, 2 casos contestaron “No sabe”, 42 casos plantearon “No responde” y para 9.811 casos la pregunta no aplicaba (código -77), ya que los menores de 15 años no responden esta pregunta.

Los códigos -88 y -99 se utilizan para identificar la no respuesta parcial de la encuesta en estas distintas categorías. Dicha no respuesta puede ser por desconocimiento de la información por parte del informante²⁰ o bien porque no se quería declarar algún dato en particular²¹.

²⁰ Muchas veces se encuesta a un informante idóneo que entrega información por otras personas del hogar, por lo que es habitual que no conozca toda la información consultada. Si bien se hacen esfuerzos desde los

Al momento de realizar el análisis, es importante considerar que las distintas variables pueden contener estos valores y, por lo tanto, podrían distorsionar eventuales interpretaciones de resultados si no se tiene en consideración la existencia de estos códigos especiales.

5.2 Hogares y personas encuestados

5.2.1 Cantidad de hogares

El número de hogares expandidos en la Base de personas de la VIII EPF se puede calcular a partir de la variable JHOGAR (jefe/a del hogar). Esta variable permite identificar a la persona reconocida como jefe o jefa de hogar (valor 1; "Sí" de esta variable). De acuerdo con la metodología de la encuesta, solo existe un integrante con esta condición para cada hogar incluido en la base de datos.

Alternativamente, para la identificación de la cantidad de hogares en la base de datos también pueden emplearse las variables N_LINEA, PERSONA y PARENTESCO, las que, al igual que JHOGAR, le atribuyen el valor 1 exclusivamente al jefe o jefa de hogar²². Esto se debe a que el registro de los integrantes de un hogar en el Registro de Personas en el Hogar se inició una vez que se identificaba a la persona jefa del hogar, cuyos datos siempre se registraban en la primera línea de los módulos con preguntas a nivel persona. Así, en la Base de personas todas las variables antes mencionadas tienen un único valor igual a 1, correspondiente a la jefatura del hogar. Para los ejemplos expuestos a continuación, cualquier tipo de cálculo que considere solo los hogares se demostrará utilizando como condición el valor "1" de la variable JHOGAR.

Programación 5: Cantidad de hogares estimados a partir de la muestra

```
// Cantidad de hogares estimados a partir de la muestra
table JHOGAR if JHOGAR==1 [pweight=FE], f(%9,0fc)
```

La variable FOLIO es la que identifica a todos los integrantes de un mismo hogar. En consecuencia, también se puede calcular el número de hogares en la Base de Personas a partir de ella. Para esto,

equipos de campos por recuperar esta información, es habitual que no pueda recuperarse la totalidad de los datos faltantes. Aunque parte de la no respuesta parcial es imputada, algunos registros no lo son (en esta versión de la encuesta no se imputaron variables cualitativas). Para más información, se sugiere revisar la metodología de la encuesta (INE Chile, 2018a).

²¹ Algunos informantes no desean entregar información respecto a algún ítem en particular, no obstante, participan de la encuesta contestando el resto de la información y módulos. Para conocer los mínimos de calidad exigidos por la encuesta, se puede revisar el apartado de Grilla de Mínimos de Calidad de la metodología de la VIII EPF (INE Chile, 2018a).

²² Por lo tanto, para filtrar una observación por hogar se puede sustituir JHOGAR==1 por PERSONA==1 o N_LINEA==1 o PARENTESCO==1 y el resultado será el mismo en los diferentes ejemplos.

se construye una variable que contabilice el número de observaciones por hogar, identificando la primera observación que posea el valor, lo que permite tabular una observación por hogar. En la Base de gastos se calcula de la misma manera que en la Base de personas, pero tomando en consideración que existen productos que no presentaron gasto en ningún hogar para esta versión de la encuesta. Estos gastos se presentan en la Base de gastos con la variable FOLIO vacía, por lo que es necesario indicar en la programación que no cuente estos folios, de modo que no se agreguen hogares adicionales al conteo.

Programación 6: Número de hogares en la muestra

```
// Número de hogares en la muestra //
use "BASE_PERSONAS_VIIIIEPF.dta", clear //Utilizar base de datos de persona
bysort FOLIO (PERSONA): gen hogar=1 if _n==1 //Se etiqueta a una única
observación por hogar
tab hogar //La frecuencia mostrará la cantidad de hogares

use "BASE_GASTOS_VIIIIEPF.dta", clear //Se abre la base de gastos
bysort FOLIO (CCIF): gen hogar=1 if _n==1 & FOLIO!="" //Se etiqueta a una única
observación por hogar
tab hogar //La frecuencia mostrará la cantidad de hogares
drop hogar //se borra las variable creada
```

5.2.2 Número personas integrantes del hogar

El número de personas del hogar (NPERSONA) es una variable que se encuentra contenida en la Base de personas de la VIII EPF. No obstante, en caso de requerirlo, se puede recalcular a través de la generación de una variable auxiliar que enumere las personas del hogar, para luego calcular el valor máximo de aquella variable auxiliar. Además de este cálculo, en la Programación 7 es posible observar cómo se tabula la cantidad de hogares según tamaño del mismo (cantidad de personas que componen el hogar), así como el tamaño promedio de los hogares por área de estimación (ZONA), considerando el total de hogares expandidos para ambos casos.

Programación 7: Número personas integrantes del hogar

```
// Cantidad de integrantes por hogar //numerop
use "BASE_PERSONAS_VIIIIEPF.dta", clear //Utilizar base de datos de persona
bysort FOLIO (PERSONA): gen contarp=_n //Se cuentan las personas al interior de
cada hogar
bysort FOLIO : egen numerop=max(contarp) //Se genera el total de personas
bysort FOLIO (PERSONA): gen numerop2=_N //Alternativamente se puede realizar en
un paso
count if numerop== numerop2 & numerop==NPERSONA //se verifican los cálculos,
debe ser 48.308
table numerop if JHOGAR==1 [pweight=FE], row f(%9,0f) //Se tabula la cantidad
de hogares por tamaño
tabstat numerop [aweight=FE], stat(mean) f(%8,3fc) by(ZONA) //Se tabula el
tamaño promedio por área de estimación
drop contarp numerop numerop2 // se borran las variables creadas
```

5.2.3 Edad de las personas integrantes del hogar

La variable EDAD es continua y tiene un rango amplio de datos, ya que llega a un máximo de 108 años. Es por eso que el usuario podría requerir construir una variable que indique la edad en tramos (que en la programación se denominará EDAD_T) para simplificar sus tabulados. En la variable calculada en la Programación 8 se agrupa la edad en tramos de diez años. Cabe señalar que existen códigos -88 (no responde) y -99 (no sabe) en esta variable, los que se agrupan en una categoría "No sabe/No responde".

Programación 8: Construcción de tramos etarios

```
// Edad en tramos // EDAD_T
gen EDAD_T=1 if EDAD>=0 & EDAD<=9
replace EDAD_T=2 if EDAD>=10 & EDAD<=19
replace EDAD_T=3 if EDAD>=20 & EDAD<=29
replace EDAD_T=4 if EDAD>=30 & EDAD<=39
replace EDAD_T=5 if EDAD>=40 & EDAD<=49
replace EDAD_T=6 if EDAD>=50 & EDAD<=59
replace EDAD_T=7 if EDAD>=60 & EDAD<=69
replace EDAD_T=8 if EDAD>=70 & EDAD<=79
replace EDAD_T=9 if EDAD>=80
replace EDAD_T=10 if EDAD== -88 | EDAD== -99
la var EDAD_T "Tramos de edad" // etiquetar variable
la de EDAD_T 1 "0 - 9" 2 "10 - 19" 3 "20 - 29" 4 "30 - 39" 5 "40 - 49" 6 "50 - 59" 7 "60 - 69" 8 "70 - 79" 9 "80 y más" 10 "No sabe/No responde", modify //
etiquetar categorías
la val EDAD_T EDAD_T
// Para tabular la edad en tramos
tabstat FE, stat(sum) f(%10,0fc) by(EDAD_T)
```

5.2.4 Sexo

La variable SEXO se preguntó en el cuestionario Registro de Personas en el Hogar, específicamente en la pregunta MH04. A continuación se muestran algunos tabulados simples por sexo y utilizando los tramos de edad generados en la Programación 8.

Programación 9: Tabulados con la variable sexo

```
table SEXO [pweight=FE], row f(%10,0fc) //Población total por sexo
tabstat FE if SEXO==1, by(EDAD_T) stat(sum) f(%9,0fc) miss //Población total.
Hombre por Tramos Etarios
tabstat FE if SEXO==2, by(EDAD_T) stat(sum) f(%9,0fc) miss //Población total.
Mujer por Tramos Etarios
```

5.2.5 Educación

En este apartado se mostrará la construcción y tabulación de las variables escolaridad (EDUE), nivel educacional por tramos (EDU_T) y nivel educacional de acuerdo con el estándar de la Clasificación Internacional Normalizada de la Educación (CINE). Si bien las variables EDUE y CINE están disponibles en la base de datos, de todos modos se explicará cómo fueron construidas.

5.2.5.1 Escolaridad

La escolaridad es presentada en la Base de personas de la VIII EPF bajo el nombre de EDUE. A través de la Programación 10 es posible conocer las decisiones tomadas en la construcción de esta variable, por ejemplo, qué hacer en aquellos casos en que una o más variables de las utilizadas como base para el cálculo contienen no respuesta.

Programación 10: Escolaridad

```
//Escolaridad //EDUE
recode EDUNIVEL (7/8=0) (9/10=8) (11/12=6) (13/14=12) (15/17=17),
gen(nivel_años)
gen EDUE2=0
//Si nivel y curso son conocidos
replace EDUE2=nivel_años + EDUCURSO if EDUCURSO!=-88 & EDUNIVEL!=-88 &
EDUCURSO!=-99 & EDUNIVEL!=-99
//Si nivel es conocido y el curso ignorado
recode EDUNIVEL (7=8) (8=6) (9/10=12) (11/12=12) (13=15) (14=17) (15=18)
(16=19) (17=21), gen(curso_max) // años de escolaridad máximos por nivel cuando
este está completo
replace EDUE2= curso_max if (EDUCURSO== -88 | EDUCURSO== -99) & EDUTERMINA==1
recode EDUNIVEL (7=7) (8=5) (9/10=11) (11=11) (12=9) (13=13) (14=15) (15=18)
(16=18) (17=19), gen(curso_aux) // auxiliar de curso cuando nivel es incompleto
o se ignora su completitud
replace EDUE2= curso_aux if (EDUCURSO== -88 | EDUCURSO== -99) & (EDUTERMINA==2 |
EDUTERMINA== -88 | EDUTERMINA== -99 | EDUTERMINA== -66)
// Si nivel esta entre 1 y 6
replace EDUE2=0 if EDUNIVEL>=1 & EDUNIVEL<=6
//Si nivel y curso son ignorados
replace EDUE2=-88 if EDUNIVEL<0 & EDUCURSO<0
//pasos finales de etiquetado
lab var EDUE2 "Años de escolaridad de la persona"
lab def EDUE2 -88 "No responde", modify
lab val EDUE2 EDUE2
//Tabla para media de escolaridad por sexo
table SEXO [pweight=FE], cont(mean EDUE2) row f(%5,2f) //La variable EDUE2
contiene la misma información que la variable EDUE
//Comprobación de cálculos
count if EDUE!=EDUE2 // Resultado debe ser cero 0
count if EDUE==EDUE2 // Resultado debe ser cero 48.308 (total de personas)
drop nivel_años curso_max curso_aux EDUE2 // se borran las variables creadas
```

5.2.5.2 Nivel educacional

Se muestra a continuación la construcción de una variable de nivel educacional denominada EDU_T, que corresponde a una reagrupación de las categorías de la variable que indica el nivel educativo más alto alcanzado (EDUNIVEL, presente en la Base de personas) y que se utiliza para los tabulados oficiales de la encuesta y aquellos presentes en el documento de Principales Resultados VIII EPF (INE Chile, 2018d).

Programación 11: Nivel educacional

```

//Nivel educacional //EDU_T
gen EDU_T=15
replace EDU_T=1 if EDUNIVEL==1 | EDUNIVEL==2 //Incluye está empezando a
estudiar
replace EDU_T=2 if EDUNIVEL==3
replace EDU_T=3 if EDUNIVEL==4 | EDUNIVEL==5 | EDUNIVEL==6 //Incluye niveles de
Sala cuna, nivel medio y transición
replace EDU_T=4 if EDUNIVEL==7
replace EDU_T=5 if EDUNIVEL==8
replace EDU_T=6 if EDUNIVEL==9
replace EDU_T=7 if EDUNIVEL==10
replace EDU_T=8 if EDUNIVEL==11
replace EDU_T=9 if EDUNIVEL==12
replace EDU_T=10 if EDUNIVEL==13
replace EDU_T=11 if EDUNIVEL==14
replace EDU_T=12 if EDUNIVEL==15
replace EDU_T=13 if EDUNIVEL==16
replace EDU_T=14 if EDUNIVEL==17

la var EDU_T "NIVEL EDUCACIONAL"
la de EDU_T 1 "NUNCA ASISTIÓ" 2 "DIFERENCIAL" 3 "PREESCOLAR" 4 "BÁSICA" 5
"PRIMARIA (SISTEMA ANTIGUO)" /*
*/ 6 "MEDIA CIENTÍFICO HUMANISTA" 7 "MEDIA TÉCNICO PROFESIONAL" 8 "HUMANIDADES
(SISTEMA ANTIGUO)" 9 /*
*/ "TÉCNICA COMERCIAL, INDUSTRIAL O NORMALISTA (SISTEMA ANTIGUO)" 10 "EDUCACIÓN
SUPERIOR CARRERAS HASTA 4 AÑOS" /*
*/ 11 "EDUCACIÓN SUPERIOR CARRERAS SOBRE 4 AÑOS" 12 "POSTÍTULO" 13 "MAESTRÍA"
14 "DOCTORADO" 15 "NIVEL IGNORADO"
la val EDU_T EDU
//Para tabular el nivel educativo según sexo
table EDU_T SEXO [pweight=FE], row f(%10,0fc) col

```

5.3 Vivienda

La información de la vivienda principal y secundaria se pregunta a nivel hogar en la encuesta, por lo que en la Base de personas la información se replica para todos los miembros del hogar.

5.3.1 Tenencia de vivienda principal

Para obtener información sobre el tipo de tenencia (TVP) y valor del arriendo imputado de la vivienda principal se debe utilizar la Base de personas de la VIII EPF. El cálculo se realiza filtrando por la variable JHOGAR, para tener solo una observación por folio.

Programación 12: Tenencia de la vivienda principal

```

//Tipo de vivienda principal // TVP
table TVP if JHOGAR==1 [pweight=FE], row f(%9,0fc)
//Valor promedio del arriendo imputado por el tipo de vivienda principal
ocupada
table TVP if JHOGAR==1 & AI_VP!=. & (TVP!=3 & TVP!=4 & TVP!=5 & TVP!=10)
[pweight=FE], cont(mean AI_VP) row f(%9,0fc) //existen dos casos en que se les
imputa arriendo imputado, no obstante la tenencia la vivienda es "no sabe" y
"no responde"

//Alternativamente
tabstat AI_VP [w=FE] if JHOGAR==1 & (TVP!=3 & TVP!=4 & TVP!=5 & TVP!=10),
stat(mean) f(%9,0fc) miss by(TVP)

```

El arriendo imputado de la vivienda principal no se calcula para los hogares que arriendan, con o sin contrato, ni para quienes tienen viviendas cedidas por trabajo o servicio. En estos casos AI_VP tiene valor -77.

5.3.2 Tenencia de la vivienda secundaria

La VIII EPF considera que un hogar cuenta con vivienda secundaria y, por tanto se incluye el gasto realizado por el hogar a causa de ella, si se cumple con cuatro condiciones: (1) debe ser una vivienda propia totalmente pagada, pagándose o una herencia o sucesión compartida; (2) el hogar debió disponer de ella durante el mes anterior a la aplicación de la encuesta; (3) la vivienda debe ser usada principalmente para el descanso, recreación y/o vacaciones, y (4) el hogar debió disponer de ella al menos tres meses durante el último año. Para resumir el cumplimiento de estas condiciones se construyó la variable VS en la Base de personas. Esta indica con valor “1” la tenencia de una vivienda secundaria y “0” cuando el hogar no cuenta con vivienda secundaria para efectos de la encuesta.

La variable VS está construida a partir de los valores de las variables VS01, VS02, VS03 y VS04, provenientes del Registro de Personas en el Hogar. Para obtener el detalle de las opciones de cada pregunta que se utilizó en la construcción, ver el Manual de Trabajo de Campo para Encuestadores (INE Chile, 2018c) o el Registro de Personas en el Hogar.

Programación 13: Tenencia de la vivienda secundaria

```
//Generación la variable tenencia de vivienda secundaria //VS
gen VS2=0 //No tiene vivienda secundaria
replace VS2=1 if VS01==1 & (VS02==1 |VS02==2 |VS02==9) & VS03==1 & VS04>=3 &
VS04<=12 // Las variables son las preguntas que se realiza en el último
módulo del Registro de Personas en el Hogar. Tiene vivienda secundaria.
//Tenencia de vivienda secundaria
table VS2 if JHOGAR==1 [pweight=FE], row f(%9,0fc)
//Valor promedio y rango intercuartil del arriendo imputado por el tipo de
vivienda ocupada
tabstat AI_VS if JHOGAR==1 & VS==1 [aweight=FE], stat(mean p25 p50 p75)
f(%9,0fc)
```

5.4 Ocupación

5.4.1 Población según su condición de actividad económica

La variable CAEG presente en la Base de personas indica la condición de actividad económica de los entrevistados, permitiendo conocer quiénes estaban ocupados durante la semana de referencia, quiénes estaban desocupados o inactivos, y quiénes son menores de 15 años y no les corresponde responder. Se construye en función de las variables AE01, AE02, AE03 y la EDAD. En la Programación 14 se muestra cómo recalcularla y tabular la población de ocupados según sexo. En el ejemplo, además, se considera la variable AEPRL, que colabora en identificar los casos en que la actividad principal declarada (semana pasada) coincide con el período de referencia para consultar ingresos (mes anterior).

Programación 14: Población según su condición de actividad económica por sexo

```
//Condición de actividad económica //CAEG
gen CAEG2=1 if (EDAD>=15 | EDAD==99 | EDAD==88) & (AE01==1 | AE02==1 |
AE03==1)
replace CAEG2=2 if (EDAD>=15 | EDAD==99 | EDAD==88) & (AE01==2 & AE02==2 &
AE03==2)
replace CAEG2=-88 if (EDAD>=15 | EDAD==99 | EDAD==88) & ((AE01==88 |
AE01==99) | (AE02==88 | AE02==99) | (AE03==88 | AE03==99)) & CAEG2!=1
replace CAEG2=-77 if EDAD< 15 & EDAD!=99 & EDAD!=88
label def CAEG2 1 "Trabaja durante la semana de referencia" 2 "Desocupado o
Inactivo durante la semana de referencia" -77 "Menores de 15 años" -88 "No
responde", modify
label values CAEG2 CAEG2
//Población de ocupados // CAEG2
table SEXO if CAEG2==1 & AEPRL==1 [pweight=FE], row f(%9,0fc)
```

5.4.2 Categoría en la ocupación por sexo

Se muestra a continuación una alternativa para tabular a la población ocupada según sexo y categoría en la Clasificación Internacional de la Situación en el Empleo (CISE). La variable específica que se usa para tabular la situación en el empleo es AECISE, correspondiente a la variable utilizada para segmentar la población en el documento Principales Resultados VIII EPF (INE Chile, 2018d).

Programación 15: Categoría en la ocupación por sexo

```
// categoría en la ocupación por sexo // CISEG
table AECISE SEXO if CAEG==1 & AEPRL==1 [pweight=FE], row f(%9,0fc) col
```

5.5 Ingresos y gastos

En este apartado se ofrece a los usuarios una alternativa para construir variables agregadas de ingreso y gasto, así como también algunos ejemplos de tabulados y gráficos a partir de algunas variables seleccionadas.

Es importante mencionar que las variables de ingresos y gastos presentadas están expresadas en pesos chilenos corrientes.

5.5.1 Ingresos

Con el objetivo de poner la mayor cantidad de información posible en manos de los usuarios, en pos de la calidad y transparencia, se incluyeron en la Base de personas las preguntas de la Libreta de Ingresos, a partir de las cuales se calculan los agregados por fuente. Al encontrarse estas variables a disposición de los usuarios, es posible presentar la forma en que se construyeron estos

agregados. Además, se entregan ya construidas las variables observadas e imputadas de ingreso de la misma forma que se hizo en la VII EPF.

En primer lugar (en el apartado 5.5.1.1) se aborda la construcción de los agregados de ingresos del trabajo, es decir, los ingresos del trabajo dependiente, del trabajo independiente y de otros trabajos²³.

Posteriormente (apartado 5.5.1.2), se aborda el cálculo de los agregados de ingresos de la propiedad, dentro de los que se incluyen los ingresos por arriendos de propiedades y por la tenencia de instrumentos financieros.

Luego, el apartado 5.5.1.3 comienza revisando el cálculo del ingreso de la producción²⁴, el que considera los ingresos del trabajo (asalariados e independientes), además del ingreso por arriendo imputado²⁵, para luego revisar la construcción del ingreso primario, posible de construir a partir de la suma de los ingresos de la producción y los de la propiedad. Cabe destacar que si bien los ingresos del trabajo y asociados a las rentas de la propiedad están a nivel de personas, por razones analíticas los ingresos de la producción y primarios se presentan y calculan a nivel hogar.

Tras esto, en el apartado 5.5.1.4 se presenta el cálculo de las transferencias corrientes recibidas, que está compuesta por los ingresos de jubilaciones y otras transferencias recibidas en dinero desde el Estado e Instituciones Sin Fines de Lucro al Servicio de los Hogares (ISFLSH) y aquellas recibidas en dinero o especie desde otros hogares.

²³ Estas actividades pueden corresponder a segundas actividades o aquellas actividades económicas que se terminaron antes de la semana de referencia, pero que de igual manera generaron ingresos en el mes de referencia para los ingresos.

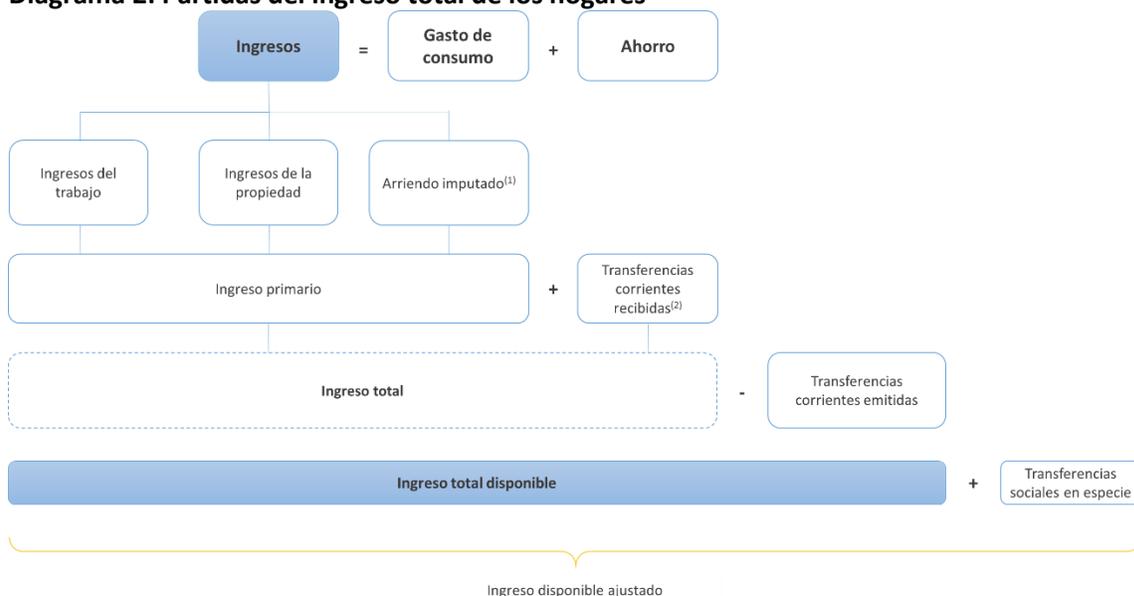
²⁴ Como se puede ver más en detalle en la Metodología VIII Encuesta de Presupuestos Familiares (INE Chile, 2018a), los ingresos de la producción incluyen los ingresos del trabajo y los de la autoprovisión de servicios producidos por el hogar. Estos últimos incluyen el arriendo imputado, el valor de los servicios domésticos no remunerados y el valor de los servicios aportados por el consumo de bienes durables. Los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables son excluidos de esta encuesta, por lo que los ingresos de la producción equivalen a los del trabajo cuando no se incluye el arriendo imputado. La Base de personas, para fines analíticos, incluye los ingresos de la producción con arriendo imputado y sin arriendo imputado, en variables separadas.

²⁵ El ingreso por arriendo imputado corresponde a un concepto que mide el valor de la prestación del servicio de alojamiento en las viviendas ocupadas por sus propietarios (arriendo imputado). El arriendo imputado es parte de un concepto más amplio conocido como **ingresos de autoprovisión de servicios producidos por el hogar**, el que, además, incluye el valor de los servicios domésticos no remunerados y el valor de los servicios aportados por el consumo de bienes durables. No obstante, los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables son excluidos de esta encuesta y han sido tradicionalmente excluidos de la medición de encuestas y de Cuentas Nacionales (UN et al., 2016).

Los últimos dos apartados de esta sección (5.5.1.5 y 5.5.1.6) entregan el cálculo de los ingresos totales y disponibles²⁶ por hogar y *per cápita*.

El Diagrama 2 ilustra cómo se complementan las fuentes de ingreso, para componer los diferentes agregados a los cuales se hace mención en los párrafos anteriores. Como es posible observar, al agregar al ingreso primario las transferencias corrientes recibidas (jubilaciones, pensiones, transferencias en dinero o especie desde otros hogares, etc.), se obtiene el ingreso total, y al restar de ese ingreso las transferencias corrientes emitidas, se consigue el ingreso disponible de los hogares. Para mayor profundidad en el marco conceptual de la medición de ingresos, ver Metodología VIII EPF (INE Chile, 2018a).

Diagrama 2: Partidas del ingreso total de los hogares



⁽¹⁾ El esquema de Cuentas Nacionales para los ingresos incorporan en esta partida los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables, los cuales son excluidos de la medición de esta encuesta, considerándose solo el arriendo imputado.

⁽²⁾ En el caso chileno, las jubilaciones poseen componentes asociados a la seguridad social, componentes asociados a capitalización individual o una combinación de ambos. En la VIII EPF, no es posible diferenciar entre ellos, por lo que se incluyen todos en las transferencias corrientes recibidas.

Fuente: Elaboración propia en base a Eurostat (2003) y OECD (2013). En el caso de la EPF, los ingresos por jubilaciones y/o pensiones de vejez se encuentran agregadas a las transferencias corrientes recibidas.

²⁶ Para el cálculo de los ingresos disponibles, en la misma sección se explicará cómo construir la partida de transferencias corrientes emitidas.

Las variables de ingresos susceptibles de imputación en la VIII EPF son las de ingresos provenientes del trabajo asalariado e independiente y de las jubilaciones, así como del arriendo imputado para aquellos hogares que hacen usufructo de la vivienda donde residen²⁷. Las variables de ingreso total y disponible (del hogar y *per cápita*) que provee la Base de personas VIII EPF se construyen a partir de las variables imputadas para las fuentes de ingreso antes mencionadas. Se muestra la construcción de estas variables en dos variantes, sin considerar el arriendo imputado como ingreso y también sumándolo al ingreso del hogar²⁸.

Es importante tener en cuenta que siempre se respeta el dato observado en las variables originales de ingresos, por lo que las imputaciones y depuraciones se realizan en variables construidas para presentar el dato corregido. Esto se traduce en que, por ejemplo, si una persona reporta recibir jubilación, pero no declara el monto, entonces en JU02 indicará que no sabe o no responde, su ingreso bruto en jubilaciones observado (INGJ) será igual a cero y su ingreso bruto en jubilaciones imputado por HD (INGJ_HD) tendrá el valor donado mediante la imputación. Por otro lado, cuando se corrigen las cotizaciones previsionales o de salud (por falta de dato o por ser anómalas en relación con el ingreso bruto o neto), igualmente se respetan los valores observados en las variables originales (TA04 y TI03, en salud; TA06 y TI05, en previsión; TA08, en impuesto a la renta) y se presentan corregidos en variables las construidas SALUD, PREVISION e IMPUESTO.

Respecto al uso de la base de datos, antes de replicar las programaciones de ingreso, es importante que los usuarios introduzcan el comando ***set type double***. Este código instruirá al programa para que cada vez que genere una variable numérica lo haga con *doble precisión*. Dado que el programa trabaja utilizando memoria RAM, los manuales de Stata recomiendan programar orientándose a la parsimonia, para así lograr tiempos de procesamiento más eficientes. Esto significa que si tenemos una variable categórica o la edad simple, que trabaja con números enteros, se está desperdiciando memoria al almacenarla con doble precisión (Stata Corp, 2018). Al contrario, para la construcción de las variables de ingreso debe utilizarse la máxima precisión posible, lo que permitirá, además, que las variables calculadas coincidan con las variables publicadas.

²⁷ La estimación del gasto en arriendo corresponde al gasto que debería realizar un hogar si no fuera propietario de su vivienda, o bien si no tuviera un beneficio gratuito de ella. Las categorías de tenencia de vivienda que se consideran para el cálculo del arriendo imputado son: propia, totalmente pagada; propia, pagándose; cedida por un pariente o amigo; ocupada de hecho; propiedad en litigio; herencia o sucesión compartida. No se calcula arriendo imputado para aquellos hogares que pagan arriendo (con o sin contrato) o si la vivienda fue cedida por el trabajo.

²⁸ Dado que el ingreso de la producción considera la agregación de los ingresos del trabajo y el arriendo imputado, al excluir el arriendo imputado, el ingreso de la producción coincide con los ingresos del trabajo.

Para ilustrar al usuario sobre el cálculo de los agregados de ingreso bruto observado, en las siguientes programaciones, se calcula una variable espejo con el mismo nombre y el sufijo "2". Por ejemplo, para recalculer INGJ de ingresos por jubilaciones, se genera una variable INGJ2. Si se generara una variable con el mismo nombre que la original, el programa no funcionaría correctamente, arrojando un mensaje de error que indica que la variable ya existe.

A continuación se revisa la construcción de las variables de los ingresos del trabajo.

5.5.1.1 Ingresos del trabajo

a) Ingresos del trabajo dependiente, sueldos y salarios

Los ingresos brutos del trabajo dependiente corresponden al total haberes o al ingreso total derivado de una actividad asalariada dependiente. Se presenta el cálculo a partir de las variables de ingresos observados y luego con las variables imputadas por *hot-deck*²⁹.

Se advierte la existencia de 21 casos de ocupados que muestran sueldo o salario bruto observado igual a cero (TA02), pero que presentan montos mayores a cero en cotizaciones previsionales (TA06) y/o de salud (TA04), o en impuesto único a las remuneraciones (TA08), durante el mes anterior a la aplicación de la encuesta. Estos casos corresponden a ocupados que declararon haber recibido ingresos por su actividad principal asalariada, no informaron el monto, pero sí lo hicieron en relación con las cotizaciones o impuestos antes mencionados. Esta situación se solventa mediante la imputación de ingresos, la que se ve reflejada en la variable de Ingresos brutos del trabajo dependiente imputado por HD (INGTD_HD).

Programación 16: Ingresos brutos del trabajo dependiente observados

```
//Ingresos brutos del trabajo dependiente observado //INGTD
foreach v in TA02 TA25 TA29 {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan los valores en blanco para
  poder sumarlas
}
gen double INGTD2=TA02+TA25+TA29 if AECISE<=7 & AECISE>=3 //Se suman los
ingresos monetarios y en especies
replace INGTD2=0 if INGTD2==.
count if INGTD!=INGTD2 //debe ser cero
```

Programación 17: Ingresos brutos del trabajo dependiente imputados por HD

```
//Ingresos brutos del trabajo dependiente imputado por HD // INGTD_HD
gen double INGTD_HD2=INGA_HD if AECISE<=7 & AECISE>=3 & AE08!=1 //Se suman los
ingresos de asalariados
replace INGTD_HD2=INGH_HD if AECISE<=7 & AECISE>=3 & AE08==1 //Se suman los
```

²⁹ Se utiliza la abreviación HD para indicar *hot-deck* en los cuadros de programación.

```

ingresos de asalariados que entregan boleta de honorarios
replace INGTD_HD2=0 if INGTD_HD2==.
count if INGTD_HD!=INGTD_HD2 //debe ser cero

```

A continuación se presentan los cálculos realizados para los asalariados no honorarios y honorarios. Esta distinción utilizada por la encuesta se realiza solo para efectos de imputación, ya que para los cálculos de los ingresos dependientes se consideran ambas modalidades de asalariados.

- Ingreso de asalariados no honorarios

El ingreso del empleo asalariado corresponde a todos los haberes o ingresos derivados del trabajo dependiente, tales como sueldos y salarios, comisiones, propinas, bonificaciones o bonos en efectivos, indemnización por despido y contribuciones de seguridad social. Dentro de los ingresos del empleo asalariado se contemplan todos los pagos: monetarios y de prestaciones en especie o servicios. Estos deben ser devengados en el mes anterior a la aplicación de la encuesta. Los ingresos del trabajo asalariado son brutos cuando no se realizan los descuentos legales (previsión, salud e impuesto de segunda categoría).

La base de datos entrega los ingresos brutos y disponibles del trabajo asalariado, observados e imputados por *hot-deck*. La comparación del ingreso bruto de asalariados observado e imputado por *hot-deck* se puede desarrollar graficando la información como se muestra en la Programación 18.

Programación 18: Ingreso bruto de asalariados no honorarios imputados versus observados

```

//Ingreso bruto de asalariados imputados versus observados // INGA_HD INGA
tabstat INGA_HD INGA if INGA_HD>0, stat(n mean min p1 p25 p50 p75 p99 max)
col(stat) f(%13,0fc) //Se tabulan los principales descriptivos de distribución
de los ingresos para poder determinar cómo especificar el gráfico de tal manera
de visualizar la distribución
//Se especifica un histograma truncado en 5 millones, se establecen las
etiquetas de los ejes, los tamaños de letra y los colores
histogram INGA if (INGA>0 & INGA<5000000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)25, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Observado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcolor(gs7) blcolor(gs7)
graph save uno, replace //Se guarda en formato Stata

histogram INGA_HD if (INGA_HD>0 & INGA_HD<5000000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)25, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Imputado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcolor(gs4) blcolor(gs4)
graph save dos, replace //Se guarda en formato Stata
graph combine uno.gph dos.gph, title("Comparación de los ingresos brutos

```

```
observados v/s imputados" "de asalariados (excluye honorarios)", span
size(large)) note("Fuente: VIII EPF- INE", size(vsmall)) scheme(s1color)
//Se combinan los gráficos, se pone título y fuente
graph export "Asalariados_HD.png", replace //Se exporta en formato png
```

El ingreso disponible del empleo asalariado corresponde al ingreso bruto del empleo asalariado restándole los descuentos legales. Se presenta a continuación una de las formas de calcular el ingreso disponible de asalariados, observado y con imputación *hot-deck*.

Programación 19: Ingreso disponible de asalariados no honorarios imputados versus observados

```
foreach v in INGA TA02 TA04 TA06 TA08 TA25 TA29 SALUD PREVISION IMPUESTO {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingreso disponible de asalariados no honorarios imputado por HD //INGDA_HD
gen double INGDA_HD2=INGA_HD-SALUD-PREVISION-IMPUESTO if AECISE<=7 & AECISE>=3
& AE08!=1 //A los ingresos brutos imputados se les resta los descuentos legales
replace INGDA_HD2=0 if INGDA_HD2==.
  count if INGDA_HD!=INGDA_HD2 //debe ser cero
//Ingreso disponible de asalariados no honorarios observado //INGDA
gen double INGDA2=TA02+TA25+TA29-TA04-TA06-TA08 if AECISE<=7 & AECISE>=3 &
AE08!=1 & (TA02>0|TA25>0|TA29>0) //A los ingresos brutos observados se les
resta los descuentos legales
replace INGDA2=0 if INGDA2==.
  count if INGDA!=INGDA2 //debe ser cero
```

- Ingreso de asalariados honorarios

El ingreso de asalariados honorarios corresponde al total de haberes o al ingreso total por trabajo asalariado percibido en el mes anterior a la aplicación de la encuesta, por personas que declaran explícitamente entregar boleta honorarios.

El ingreso de honorarios es bruto cuando no se descuenta la retención del impuesto (10%), ni las cotizaciones voluntarias de previsión y/o de salud. La comparación del ingreso bruto de honorarios observado versus el imputado por *hot-deck* puede hacerse de la siguiente manera:

Programación 20: Ingreso bruto de asalariados honorarios imputados versus observados

```
//Ingreso bruto de asalariados honorarios imputados versus observados //
INGH_HD INGH
tabstat INGH_HD INGH if INGH_HD>0, stat(n mean min p1 p25 p50 p75 p99 max)
col(stat) f(%13,2f) //Se tabulan los principales descriptivos de distribución
de los ingresos para poder determinar cómo especificar el gráfico de tal manera
de visualizar la distribución
//Se especifica un histograma truncado en 5 millones, se establecen las
etiquetas de los ejes, los tamaños de letra y los colores
histogram INGH if (INGH<5000000 & INGH>0), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)25, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Observado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcOLOR(gs7) blcolor(gs7)
```

```

graph save uno.gph, replace //Se guarda en formato Stata

histogram INGH_HD if (INGH_HD<5000000 & INGH_HD>0), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)25, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Imputado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(slcolor) bfcolor(gs4) blcolor(gs4)
graph save dos.gph, replace //Se guarda en formato Stata
graph combine uno.gph dos.gph, title("Comparación de los ingresos brutos
observados v/s imputados" "asalariados que entregan boleta de honorarios", span
size(large)) note("Fuente: VIII EPF- INE", size(vsmall)) scheme(slcolor)
//Se combinan los gráficos, se pone título y fuente
graph export "Honorarios_HD.png", replace // Se exporta en formato png

```

El ingreso disponible de honorarios corresponde al ingreso bruto de honorarios restándole la retención del impuesto (10%) y las cotizaciones voluntarias de previsión y/o de salud. El ingreso disponible de honorarios observados y con imputación *hot-deck* se puede calcular como lo indica la Programación 21:

Programación 21: Ingreso disponible de asalariados honorarios imputados versus observados

```

foreach v in INGA TA02 TA04 TA06 TA08 TA25 TA29 SALUD PREVISION IMPUESTO {
replace `v'=0 if `v'==. | `v'<0 // Se reemplazan para poder sumarlas
}
//Ingreso disponible de honorarios imputado por HD //INGDH_HD
gen double INGDH_HD2=INGH_HD-SALUD-PREVISION-IMPUESTO if AECISE<=7 & AECISE>=3
& AE08==1 //A los ingresos brutos imputados se les resta los descuentos legales
replace INGDH_HD2=0 if INGDH_HD2==.
count if INGDH_HD!=INGDH_HD2 //debe ser cero
//Ingreso disponible de honorarios observado //INGDH
gen double INGDH2=TA02+TA25+TA29-TA04-TA06-TA08 if AECISE<=7 & AECISE>=3 &
AE08==1 & (TA02>0|TA25>0|TA29>0) //A los ingresos brutos observados se les
resta los descuentos legales
replace INGDH2=0 if INGDH2==.
count if INGDH!=INGDH2 //debe ser cero

```

b) Ingresos netos del trabajo independiente

Se identifican como ingresos netos del trabajo independiente los ingresos generados por la realización de una actividad o negocio por cuenta propia descontando los costos asociados a su realización: impuestos, salarios pagados o bienes intermedios que deben ser comprados para la producción de la actividad. Se presenta el cálculo a partir de las variables de ingresos observados y luego con las variables imputadas por *hot-deck*.

En relación con los ingresos del trabajo independiente, también se ve la existencia de trabajadores independientes que muestran un ingreso neto observado de su actividad igual a cero (TI01), pero que presentan montos mayores a cero en cotizaciones previsionales (TI05) y/o de salud (TI03), durante el mes anterior a la aplicación de la encuesta. Son 140 registros los que presentan esta condición. Estos casos corresponden a ocupados que declararon haber recibido ingresos por su

actividad principal independiente, no informaron el monto, pero sí lo hicieron en relación con las cotizaciones antes mencionadas. Esta situación se solventa mediante la imputación de ingresos, la que se ve reflejada en la variable de ingresos netos del trabajo independiente imputado por HD (INGTI_HD).

Programación 22: Ingresos brutos del trabajo independiente observados

```
//Ingresos brutos del trabajo independiente observado //INGTI
foreach v in TI01 TI12 {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGTI2=TI01+TI12 if AECISE<=2 & AECISE>=1 //Se suman los ingresos
monetarios y en especies
replace INGTI2=0 if INGTI2==.
count if INGTI!=INGTI2 //debe ser cero
```

Programación 23: Ingresos brutos del trabajo independiente imputados por HD

```
//Ingresos brutos del trabajo independiente imputado por HD //INGTI_HD
gen double INGTI_HD2= INGNP_HD if AECISE<=2 & AECISE>=1 & TI10==1 //Se suman
los ingresos de quienes tienen negocios por cuenta propia
replace INGTI_HD2= INGPI_HD if AECISE<=2 & AECISE>=1 & TI10==2 //Se suman los
ingresos de quienes tienen una actividad profesional independiente
replace INGTI_HD2=INGNP_HD if AECISE<=2 & AECISE>=1 & TI10==88 //Se suman los
ingresos de las 41 personas que no contestaron TI10
replace INGTI_HD2=0 if INGTI_HD2==.
count if INGTI_HD!=INGTI_HD2 //debe ser cero
```

Los ingresos del trabajo independiente se pueden desagregar según la situación en el empleo de quienes recibieron aquellos ingresos durante el mes de referencia: (1) patrón o empleador y (2) trabajador independiente o por cuenta propia. Para distinguir a estos tipos de trabajador independiente se utiliza la pregunta AECISE, que corresponde a una autoclasificación, que consulta si la persona tuvo empleados a su cargo durante el período de referencia.

- Ingreso de empleadores

El ingreso neto de los empleadores corresponde a los ingresos generados por la realización de una actividad o negocio por cuenta propia, descontando los costos asociados a su realización, siempre y cuando la persona que declara estos ingresos en su ocupación principal se autoclasifique como patrón o empleador (da empleo a otras personas) durante el mes de referencia (la AECISE toma el valor "1"). El cálculo de este ingreso se puede realizar de la siguiente manera:

Programación 24: Ingreso neto de los empleadores imputado versus observado

```
//Ingreso neto de empleadores imputado por HD //INGEMP_HD
gen double INGEMP_HD2= INGTI_HD if AECISE ==1 //Se suman los ingresos
monetarios y en especies
replace INGEMP_HD2=0 if INGEMP_HD2==.
count if INGEMP_HD2!= INGEMP_HD2
//Ingreso neto de empleadores observado //INGEMP
foreach v in TI01 TI12 {
```

```

replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGEMP2=TI01+TI12 if AECISE ==1 //Se suman los ingresos monetarios y
en especies
replace INGEMP2=0 if INGEMP2==.
count if INGEMP!= INGEMP2 //debe ser cero

```

El ingreso disponible de los empleadores se obtiene descontando las cotizaciones voluntarias que haya realizado en previsión y/o salud.

Programación 25: Ingreso disponible de los empleadores imputado versus observado

```

//Ingreso disponible de empleadores imputado por HD //INGDEMP_HD
foreach v in INGEMP_HD TI01 TI03 TI05 TI12 SALUD PREVISION {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingreso disponible de empleadores imputado por HD //INGDA_HD
gen double INGDEMP_HD2= INGEMP_HD-SALUD-PREVISION if AECISE==1 //A los ingresos
brutos imputados se les resta los descuentos legales
replace INGDEMP_HD2=0 if INGDEMP_HD2==.
count if INGDEMP_HD!= INGDEMP_HD2 //debe ser cero
//Ingreso disponible de empleadores observado //INGDEMP
gen double INGDEMP2 = TI01 +TI12-TI03-TI05 if AECISE==1 & (TI01>0|TI12>0) //A
los ingresos brutos observados se les resta los descuentos legales
replace INGDEMP2=0 if INGDEMP2==.
count if INGDEMP!= INGDEMP2 //debe ser cero

```

- Ingreso de trabajadores por cuenta propia

En contrapartida a lo anterior, los ingresos de los trabajadores por cuenta propia son aquellos recibidos por trabajadores independientes que se autoclasifican como trabajador independiente o por cuenta propia (la AECISE toma el valor “2”) durante el mes anterior a la aplicación de la encuesta. Se pueden obtener ejecutando los pasos de la Programación 26.

Programación 26: Ingreso neto de trabajadores por cuenta propia imputado versus observado

```

//Ingreso neto de trabajadores por cuenta propia imputado por HD //INGTCP_HD
gen double INGTCP_HD2= INGTI_HD if AECISE ==2 //Se suman los ingresos
monetarios y en especies
replace INGTCP_HD2=0 if INGTCP_HD2==.
count if INGTCP_HD2!= INGTCP_HD2 //debe ser cero
//Ingreso neto de trabajadores por cuenta propia observado //INGTCP
foreach v in TI01 TI12 {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGTCP2=TI01+TI12 if AECISE ==2 //Se suman los ingresos monetarios
y en especies
replace INGTCP2=0 if INGTCP2==.
count if INGTCP!= INGTCP2

```

Al igual que en el caso anterior, el ingreso disponible de estos trabajadores se obtiene restando del neto las cotizaciones voluntarias en previsión y/o salud.

Programación 27: Ingreso disponible de trabajadores por cuenta propia imputado versus observado

```

foreach v in INGTCP_HD TI01 TI03 TI05 TI12 SALUD PREVISION {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingreso disponible de trabajadores por cuenta propia imputado por HD
//INGDTCP_HD
gen double INGDTCP_HD2= INGTCP_HD -SALUD-PREVISION if AECISE==2 //A los
ingresos brutos imputados se les resta los descuentos legales
replace INGDTCP_HD2=0 if INGDTCP_HD2==.
  count if INGDTCP_HD!= INGDTCP_HD2 //debe ser cero
//Ingreso neto de trabajadores por cuenta propia observado //INGDTCP
gen double INGDTCP2 = TI01+TI12-TI03-TI05 if AECISE==2 & (TI01>0|TI12>0) //A
los ingresos brutos observados se les resta los descuentos legales
replace INGDTCP2=0 if INGDTCP2==.
  count if INGDTCP!= INGDTCP2 //debe ser cero

```

A continuación se presentan los ingresos de trabajadores independientes según la clasificación alternativa referente a la variable TI10, que indica una autoclasificación del informante respecto a si en su actividad principal lo realizó para su propio negocio, empresa o actividad por cuenta propia o como actividad profesional independiente (clasificación también utilizada en la VII EPF). De esta respuesta se determina si el trabajador podrá o no tener ingresos en especies, ya que se excluyen los servicios.

- Ingreso de quienes tienen negocios por cuenta propia

Este tipo de ingreso corresponde al que proviene de la realización de una actividad o negocio por cuenta propia. El ingreso de una actividad o negocio por cuenta propia es neto cuando se descuentan los costos asociados a su realización (impuestos, salarios que deban ser pagados a empleados u otros y bienes intermedios que deban ser comprados para la producción de la actividad). También debe sumarse a los ingresos lo que retiró del negocio o empresa para el consumo personal o del hogar, valorado a precio de mercado. El ingreso neto de quienes tienen negocios por cuenta propia observado e imputado por *hot-deck* se puede graficar de la siguiente manera a modo de comparación:

Programación 28: Ingreso neto de quienes tienen negocio por cuenta propia imputado versus observado

```

//Ingreso disponible de quienes tienen negocio por cuenta propia imputado
versus observados //INGNP_HD INGNP
tabstat INGNP_HD INGNP if INGNP_HD>0, stat(n mean min p1 p25 p50 p75 p99 max)
col(stat) f(%13,0fc) //Se tabulan los principales descriptivos de distribución
de los ingresos para poder determinar cómo especificar el gráfico de tal manera
de visualizar la distribución
//Se especifica un histograma truncado en 5 millones, se establecen las
etiquetas de los ejes, los tamaños de letra y los colores
histogram INGNP if (INGNP>0 & INGNP<5000000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Observado (en miles de

```

```

pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcOLOR(gs7) blcolor(gs7)
graph save uno.gph, replace //Se guarda en formato Stata
histogram INGNP_HD if (INGNP_HD>0 & INGNP_HD<5000000), leg(off) percent normal
normopts(lwidth(thick))/*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Imputado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcOLOR(gs4) blcolor(gs4)
graph save dos.gph, replace // Se guarda en formato Stata
graph combine uno.gph dos.gph, title("Comparación de los ingresos neto
observados v/s imputados" "de quienes tienen negocio por cuenta propia", span
size(large)) note("Fuente: VIII EPF- INE", size(vsmall)) scheme(s1color)
//Se combinan los gráficos, se pone título y fuente
graph export "CuentaPropia_HD.png", replace //Se exporta en formato png

```

El ingreso disponible de quienes tienen negocio por cuenta propia corresponde al ingreso neto ya definido, al que se le descuentan las cotizaciones voluntarias en previsión y/o salud. Se presenta a continuación una alternativa para el cálculo del ingreso disponible de quienes tienen negocio por cuenta propia, observado e imputado por *hot-deck*:

Programación 29: Ingreso disponible de quienes tienen negocios por cuenta propia imputados versus observados

```

foreach v in INGTI_HD TI01 TI03 TI05 TI12 SALUD PREVISION {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingreso disponible de quienes tienen negocio por cuenta propia imputado por
HD //INGDNP_HD
gen double ING_DNP_HD2= INGTI_HD-SALUD-PREVISION if AECISE<=2 & AECISE>=1 &
(TI10==1 | TI10== -88) //A los ingresos brutos imputados se les resta los
descuentos legales. La categoría no responde de TI10 se considera ingreso por
negocio por cuenta propia
replace ING_DNP_HD2=0 if ING_DNP_HD2==.
count if ING_DNP_HD!= ING_DNP_HD2 //debe ser cero

//Ingreso disponible de quienes tienen negocio por cuenta propia observado
//INGDNP
gen double ING_DNP2= TI01+TI12-TI03-TI05 if AECISE<=2 & AECISE>=1 &
(TI10==1|TI10== -88) & (TI01>0|TI12>0) //A los ingresos brutos observados se les
resta los descuentos legales. La categoría no responde de TI10 se considera
ingreso por negocio por cuenta propia
replace ING_DNP2=0 if ING_DNP2==.
count if ING_DNP!= ING_DNP2 //debe ser cero

```

- Ingreso de la actividad profesional independiente

El ingreso de la actividad profesional independiente corresponde al ingreso generado por las personas que se declaran independientes, pero que se distinguen de quien tiene algún negocio o empresa. Su actividad se relaciona más con la prestación de servicios y no realizan, por ese motivo, retiros en especie desde su negocio.

El ingreso de la actividad profesional independiente es neto cuando se restan los costos asociados a su realización (similar al caso anterior, ya que ambas categorías son consideradas como independientes). El ingreso neto de los profesionales independientes observado e imputado por *hot-deck* puede ser calculado del modo en que se muestra en la Programación 30.

Programación 30: Ingreso neto de la actividad profesional independiente imputados versus observados

```
//Ingreso neto de la actividad profesional independiente imputado por HD
//INGPI_HD INGPI
tabstat INGPI_HD INGPI if INGPI_HD>0, stat(n mean min p1 p25 p50 p75 p99 max)
col(stat) f(%13,0fc) //Se tabulan los principales descriptivos de distribución
de los ingresos para poder determinar cómo especificar el gráfico de tal manera
de visualizar la distribución
//Se especifica un histograma truncado en 5 millones, se establecen las
etiquetas de los ejes, los tamaños de letra y los colores
histogram INGPI if (INGPI>0 & INGPI <5000000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Observado (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcolor(gs7) blcolor(gs7)
graph save uno.gph, replace //Se guarda en formato Stata
histogram INGPI_HD if (INGPI_HD>0 & INGPI_HD<5000000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 1000000 "1.000" 2000000 "2.000" 3000000 "3.000" 4000000 "4.000"
5000000 "5.000", axis(1) gmax labsize(small)) xtitle("Imputados (en miles de
pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcolor(gs4) blcolor(gs4)
graph save dos.gph, replace //Se guarda en formato Stata
//Se combinan los gráficos, se pone título y fuente
graph combine uno.gph dos.gph, title("Comparación de los ingresos netos
observados v/s imputados" "de la actividad profesional independiente", span
size(large)) note("Fuente: VIII EPF- INE", size(vsmall)) scheme(s1color)
graph export "ProfIndep_HD.png", replace //Se exporta en formato png
```

El ingreso disponible de quienes reciben ingresos por este tipo de actividades se obtiene descontando al neto las cotizaciones voluntarias en previsión y/o salud. El ingreso disponible de la actividad profesional independiente observado e imputado por *hot-deck* puede ser calculado del modo expuesto en la Programación 31.

Programación 31: Ingreso disponible de la actividad profesional independiente imputados versus observados

```
foreach v in INGTI_HD TI01 TI03 TI05 TI12 SALUD PREVISION {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingreso disponible de la actividad profesional independiente imputado por HD
//INGDPI_HD
gen double INGDPI_HD2= INGTI_HD -SALUD-PREVISION if AECISE<=2 & AECISE>=1 &
TI10==2 //A los ingresos brutos imputados se les resta los descuentos legales
replace INGDPI_HD2=0 if INGDPI_HD2==.
count if INGDPI_HD!= INGDPI_HD2 //debe ser cero
//Ingreso disponible de la actividad profesional independiente observado
//INGDPI
```

```

gen double INGDPI2 = TI01+TI12-TI03-TI05 if AECISE<=2 & AECISE>=1 & TI10==2 &
(TI01>0 | TI12>0) //A los ingresos brutos observados se les resta los
descuentos legales
replace INGDPI2 =0 if INGDPI2==.
count if INGDPI!= INGDPI2 //debe ser cero

```

c) Ingresos por otros trabajos asalariados e independientes

El Ingreso por otros trabajos asalariados e independientes es aquel que proviene de actividades secundarias y de actividades principales que, habiendo generado ingresos durante el mes pasado (período de referencia de los ingresos), finalizaron antes de la semana de referencia para la ocupación. A continuación se presenta el cálculo del ingreso bruto y disponible (restando descuentos legales y no legales) derivados de otros trabajos asalariados.

Programación 32: Ingreso bruto y disponible por otros trabajos asalariados

```

//Ingreso bruto otros trabajos dependiente //INGOTA
gen double INGOTA2=OT03 if OT02==1 & OT03>0
replace INGOTA2=0 if INGOTA2==.
count if INGOTA!=INGOTA2 //debe ser cero
//Ingreso disponible otros trabajos dependiente
gen double INGDOTA2=OT07 if OT02==1 & OT07>0 //Alternativa 1 de cálculo
foreach v in OT03 OT05 OT06 {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGDOTA3=OT03-OT05-OT06 if OT02==1 & OT03>0 //Alternativa 2 de
cálculo
replace INGDOTA2=0 if INGDOTA2==.
replace INGDOTA3=0 if INGDOTA3==.
count if INGDOTA!=INGDOTA2 & OT07>0 //debe ser cero
count if INGDOTA!=INGDOTA3 & OT07>0 //debe ser cero

```

En el caso de los ingresos de otros trabajos independientes no se entrega el disponible, dado que no se pregunta sobre cotizaciones previsionales ni de salud en relación con esos ingresos. En la Programación 33 se muestra un ejemplo de cómo calcular el ingreso bruto por otros trabajos independientes.

Programación 33: Ingreso disponible por otros trabajos independientes INGOTI

```

//Ingreso disponible por otros trabajos independientes //INGOTI
gen double INGDOTI2=OT09 if OT08==1 & OT09>0
replace INGDOTI2=0 if INGDOTI2==.
count if INGDOTI!=INGDOTI2 //debe ser cero

```

d) Ingresos del trabajo según persona

En los apartados anteriores se revisó la construcción de los distintos componentes del ingreso de las personas. A continuación se presenta la programación de las sumas para obtener los ingresos provenientes del trabajo a nivel de personas perceptoras de estos ingresos, considerando los ingresos de la ocupación principal, otros ingresos y el agregado de ambos componentes.

Programación 34: Agregados brutos y disponibles de los ingresos del trabajo

```

foreach v in INGA INGA_HD INGDA_HD INGH INGH_HD INGDH_HD INGNP INGNP_HD
INGDNP_HD INGPI INGPI_HD INGDP_HD INGDP_HD INGDP_HD INGDP_HD INGDP_HD {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Ingresos del trabajo de la ocupación principal (bruto observado)
g double INGTRABOP=INGA+INGH+INGNP+INGPI
//Ingresos del trabajo de la ocupación principal (bruto imputado)
g double INGTRABOP_HD=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD
//Ingresos del trabajo de la ocupación principal (disponible imputado)
g double INGDTABOP_HD=INGDA_HD+INGDH_HD+INGDNP_HD+INGDPI_HD
//Ingresos del trabajo de la ocupación secundaria y de ocupación del mes
anterior que no coincide con la ocupación declarada como principal
g double INGTRABOS=INGOTA+INGDOTI
//Ingresos disponibles del trabajo de la ocupación secundaria y de ocupación
del mes anterior que no coincide con la ocupación declarada como principal
g double INGDTABOS=INGDOTA+INGDOTI
//Total por persona Ingresos del trabajo con imputación
g double INGTRAB_HD=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI
//Total por persona Ingresos disponibles del trabajo con imputación
g double INGDTAB_HD=INGDA_HD+INGDH_HD+INGDNP_HD+INGDPI_HD+INGDOTA+INGDOTI

```

e) Ingresos medios por la ocupación principal según sexo

A continuación se muestra una alternativa de programación que tabula la media de los ingresos de la ocupación principal según sexo. Para que la programación funcione, se debe generar la variable que hace referencia a los ingresos del trabajo de la ocupación principal de la Programación 34 (INGDTRABOP_HD).

Programación 35: Ingresos medios por ocupación principal según sexo

```

//Valor promedio de los ingresos de la ocupación principal según sexo
table SEXO if INGDTABOP_HD>0 [pweight=FE], cont(mean INGDTABOP_HD) row
f(%9,0fc) //calculada para quienes tienen un ingreso de la ocupación mayor a
cero

//Alternativamente
tabstat INGDTABOP_HD [w=FE] if INGDTABOP_HD>0, stat(mean) f(%9,0fc) miss
by(SEXO)

```

5.5.1.2 Ingreso por rentas de la propiedad e instrumentos financieros*a) Ingresos de propiedades*

Los ingresos de la propiedad son aquellos percibidos durante el mes anterior a la aplicación de la encuesta que provienen del arriendo de bienes inmuebles (casas, departamentos, oficinas, locales comerciales, estacionamientos, tierras u otras propiedades) o de bienes muebles (vehículos, maquinarias de todo tipo y animales). Se pueden calcular como se muestra en la Programación 36.

Programación 36: Ingresos de propiedades

```
//Ingresos de propiedades //INGP
foreach v in IP02 IP04 {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGP2=IP02+IP04
count if INGP!=INGP2 //debe ser cero
```

b) Ingresos por tenencia de instrumentos financieros

Los ingresos por instrumentos financieros corresponden a las ganancias percibidas durante el mes anterior a la aplicación de la encuesta, derivadas de la tenencia de este tipo de instrumentos. Incluyen dividendos por acciones, utilidades por la participación en sociedades de personas, intereses de cuentas de ahorro e intereses por otros instrumentos financieros, como depósitos a plazo, fondos mutuos, etc. Se pueden calcular del modo en que se expone en la Programación 37.

Programación 37: Ingresos por tenencia de instrumentos financieros

```
//Ingresos por instrumentos financieros //INGF
foreach v in IP06 IP08 IP10_GANANCIA IP12_GANANCIA {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double INGF2=IP06+IP08+IP10_GANANCIA+IP12_GANANCIA
count if INGF!=INGF2 //debe ser cero
```

5.5.1.3 Ingreso de la producción e ingreso primario

En el presente apartado se desarrolla el cálculo del ingreso de la producción y del ingreso primario de los hogares. A diferencia de las programaciones revisadas de forma previa, estos agregados son construidos a nivel de ingresos del hogar.

a) Ingresos de la producción

Los ingresos de la producción corresponden a aquellos ingresos del trabajo derivados de la actividad principal o secundaria, ya sea por una actividad asalariada o actividad independiente. Además, contiene aquellos ingresos derivados de la producción de los hogares para el autoconsumo, los que en la encuesta solo contempla el arriendo imputado.

Es preciso indicar que hay componentes del ingreso de la producción que no son capturados por la VIII EPF y que han sido tradicionalmente excluidos de la medición Cuentas Nacionales (UN, 2001). Estos son los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables. Aquellos ingresos de la producción que sí son capturados en la VIII EPF son los ingresos del trabajo y también el arriendo imputado, que es considerado como un tipo de autoprovisión de servicios producidos por el hogar. Por motivos metodológicos, se presenta la variable ingresos de la producción (ING_PRODUC) con arriendo imputado y sin arriendo imputado (lo que sería equivalente a los ingresos del trabajo).

Los ingresos de la producción observados e imputados por *hot-deck* agregados a nivel hogar, sin y con arriendo imputado, pueden ser calculados de la siguiente manera³⁰:

Programación 38: Ingresos imputados de la producción a nivel hogar sin arriendo imputado

```
//Ingreso de la producción imputado a nivel hogar//ING_PRODUC_HOG_HD
foreach v in INGA_HD INGH_HD INGNP_HD INGPI_HD INGOTA INGDOTI {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double ING_PRODUC_IND=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI
bys FOLIO: egen double ING_PRODUC_HOG_HD2= sum(ING_PRODUC_IND)
replace ING_PRODUC_HOG_HD2=0 if ING_PRODUC_HOG_HD2==.

count if ING_PRODUC_HOG_HD!=ING_PRODUC_HOG_HD2 //debe ser cero
```

Programación 39: Ingresos imputados de la producción a nivel hogar con arriendo imputado

```
//Ingreso de la producción imputado con arriendo imputado
//ING_PRODUC_HOG_HD_AI
foreach v in INGA_HD INGH_HD INGNP_HD INGPI_HD INGOTA INGDOTI AI_VP AI_VS {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double ING_PRODUC_IND_AI=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI
if JHOGAR!=1
replace
ING_PRODUC_IND_AI=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI+AI_VP+AI_VS
if JHOGAR==1 //el arriendo imputado se suma solo para el jefe o jefa de hogar
bys FOLIO: egen double ING_PRODUC_HOG_HD_AI2= sum(ING_PRODUC_IND_AI)
replace ING_PRODUC_HOG_HD_AI2=0 if ING_PRODUC_HOG_HD_AI2==.

count if ING_PRODUC_HOG_HD_AI!=ING_PRODUC_HOG_HD_AI2 //debe ser cero
```

b) Ingreso primario

El ingreso primario corresponde a la suma de los ingresos de la producción y los ingresos de la propiedad. Los ingresos de la propiedad incluyen ingresos derivados de arriendos de activos (inmuebles y muebles) y los ingresos derivados de la tenencia de instrumentos financieros.

El ingreso primario a nivel hogar, imputado por *hot-deck* con y sin arriendo imputado, puede ser calculado de la siguiente manera:

³⁰ Para el cálculo del ingreso de la producción se asume que el ingreso bruto por otros trabajos independientes coincide con el ingreso disponible asociado a otros trabajos independientes (segundas ocupaciones independientes o bien ocupaciones independientes del mes pasado que no coinciden con la ocupación declarada en la semana de referencia).

Programación 40: Ingresos imputados primarios a nivel hogar con y sin arriendo imputado

```

//Ingreso primario con ingresos imputados por HD (sin arriendo imputado)
//ING_PRIMARIO_HOG_HD
foreach v in INGA_HD INGH_HD INGNP_HD INGPI_HD INGOTA INGDOTI INGP INGF AI_VP
AI_VS {
  replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double ING_PRIMARIO_IND=INGA_HD+ INGH_HD+INGNP_HD+ INGPI_HD+INGOTA+INGDOTI+
INGP+INGF
bys FOLIO: egen double ING_PRIMARIO_HOG_HD2= sum(ING_PRIMARIO_IND)
replace ING_PRIMARIO_HOG_HD2=0 if ING_PRIMARIO_HOG_HD2==.
  count if ING_PRIMARIO_HOG_HD!=ING_PRIMARIO_HOG_HD2 //debe ser cero
//Ingreso primario con ingresos imputados por HD (con arriendo imputado)//
ING_PRIMARIO_HOG_HD_AI
gen double ING_PRIMARIO_IND_AI=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+ INGOTA+
INGDOTI+INGP+INGF if JHOGAR!=1
replace ING_PRIMARIO_IND_AI=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+ INGOTA+INGDOTI+
INGP+INGF+AI_VP+AI_VS if JHOGAR==1 //el arriendo imputado se suma solo para el
Jefe o jefa de hogar
bys FOLIO: egen double ING_PRIMARIO_HOG_HD_AI2= sum(ING_PRIMARIO_IND_AI)
replace ING_PRIMARIO_HOG_HD_AI2=0 if ING_PRIMARIO_HOG_HD_AI2==.
  count if ING_PRIMARIO_HOG_HD_AI!=ING_PRIMARIO_HOG_HD_AI2 //debe ser cero

```

5.5.1.4 Transferencias corrientes*a) Ingreso por jubilaciones*

El ingreso por jubilaciones corresponde a la suma de los ingresos por jubilaciones y/o pensiones de vejez pagados por el Instituto de Previsión Social (IPS), por alguna Administradora de Fondos de Pensiones (AFP) o por algún sistema de las FF.AA.

El ingreso por jubilaciones es bruto cuando no se han descontado las cotizaciones legales en salud. El ingreso bruto en jubilaciones observado corresponde a la pregunta JU03. Al descontar las cotizaciones legales en salud, se obtiene el ingreso disponible en jubilaciones, que equivale a JU02.

Programación 41: Ingreso bruto y disponible en jubilaciones observado

```

//Ingreso bruto observado de Jubilaciones y/o pensiones de vejez //INGJ
gen double INGJ2=JU02 if JU01==1 & (JU02!=-99 & JU02!=-88)
replace INGJ2=0 if INGJ2==.
  count if INGJ!=INGJ2 //debe ser cero
//Ingreso disponible observado de Jubilaciones y/o pensiones de vejez //INGDJ
gen double INGDJ2=JU03 if JU01==1 & (JU03!=-99 & JU03!=-88)
replace INGDJ2=0 if INGDJ2==.
  count if INGDJ!=INGDJ2 //debe ser cero

```

La imputación de los ingresos por jubilaciones y/o pensiones de vejez se realiza a partir del valor bruto, al cual se descuenta el 7% de salud según el valor de las variables SP01 y SP02 (afiliación e identificación de sistema de salud al que pertenece). La comparación del ingreso bruto de jubilaciones y/o pensiones de vejez observado versus el imputado por *hot-deck* se puede generar según la programación que a continuación se presenta:

Programación 42: Ingreso bruto en jubilaciones imputado y observado

```

//Ingreso bruto en Jubilaciones y/o pensiones de vejez observado vs imputado
por HD //INGJ INGJ_HD
tabstat INGJ_HD INGJ if INGJ_HD>0, stat(n mean min p1 p25 p50 p75 p99 max)
col(stat) f(%13,0fc) //Se tabulan los principales descriptivos de distribución
de los ingresos para poder determinar cómo especificar el gráfico de tal manera
de visualizar la distribución
//Se especifica un histograma truncado en 1.5 millones, se establecen las
etiquetas de los ejes, los tamaños de letra y los colores
histogram INGJ if (INGJ>0 & INGJ <1500000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 5e5 "500" 1e6 "1.000" 1.5e6 "1.500", axis(1) gmax
labsize(small)) xtitle("Observado (en miles de pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcOLOR(gs7) blcolor(gs7)
graph save uno.gph, replace // Se guarda en formato Stata
histogram INGJ_HD if (INGJ_HD>0 & INGJ_HD<1500000), leg(off) percent normal
normopts(lwidth(thick)) /*
*/ ylabel(0(5)30, ang(h) grid labsize(small)) ytitle("Porcentaje (%)", axis(1)
size(small)) /*
*/ xlabel(0 "0" 5e5 "500" 1e6 "1.000" 1.5e6 "1.500", axis(1) gmax
labsize(small)) xtitle("Imputado (en miles de pesos)", axis(1) size(small)) /*
*/ scheme(s1color) bfcOLOR(gs4) blcolor(gs4)
graph save dos.gph, replace //Se guarda en formato Stata
graph combine uno.gph dos.gph, title("Comparación de los ingresos brutos
observados v/s imputados" "Jubilaciones y/o pensiones de vejez", span
size(large)) note("Fuente: VIII EPF- INE", size(vsmall)) scheme(s1color) //Se
combinan los gráficos, se pone título y fuente
graph export "Jubilación_HD.png", replace //Se exporta en formato png

```

b) Ingresos por otras transferencias (corrientes)

La encuesta mide las transferencias corrientes (en dinero) que los hogares reciben desde distintos agentes.

El ingreso por transferencias recibidas es la suma de las transferencias monetarias corrientes recibidas desde otros hogares, el Estado o Instituciones Sin Fines de Lucro al Servicio de los Hogares (ISFLSH) y las transferencias en especies recibidas desde otros hogares. Dado que el objetivo de la encuesta es medir el ingreso disponible de los hogares y no el ajustado, se excluyen del cálculo las transferencias recibidas en especie desde el Estado y desde las ISFLSH³¹. Los ingresos por transferencias recibidas se pueden calcular de la siguiente manera:

³¹ Si bien la pregunta TR14 consulta sobre transferencias en especie recibidas desde el Estado y desde las ISFLSH, su medición no es exhaustiva y no es posible reconstruir las transferencias sociales en especie a partir de dicha medición (por ejemplo, no se incluye la valoración de educación, asistencia sanitaria, entre otros componentes).

Programación 43: Ingresos por transferencias recibidas

```
//Ingresos por transferencias recibidas //TRANSFERENCIAS_RECIBIDAS
foreach v in TR02 TR04 TR07 TR09 TR10 TR14 {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Transferencias corrientes recibidas de otros hogares (monetarias y en
especies)
gen double ING_TRANS1=TR04+TR07 if TR13!=1
replace ING_TRANS1=TR04+TR07+TR14 if TR13==1
//Transferencias corrientes recibidas del Estado e ISFLH
gen double ING_TRANS2=TR02+TR09
replace ING_TRANS2=TR02+TR10 if TR10>0 & TR09==0

gen double TRANSFERENCIAS_RECIBIDAS2=ING_TRANS1+ING_TRANS2
count if TRANSFERENCIAS_RECIBIDAS!= TRANSFERENCIAS_RECIBIDAS2 //debe ser cero
```

5.5.1.5 Ingreso total del hogar imputado por hot-deck

El ingreso total del hogar corresponde a la sumatoria de todas las fuentes de ingreso que fueron percibidos el mes anterior a la aplicación de la encuesta. Incluye los ingresos del trabajo (ya sea generados por la actividad principal o por otras actividades laborales), ingresos de la propiedad (rentas de arriendos y ganancias asociadas a instrumentos financieros), alquiler imputado, jubilaciones y transferencias corrientes (en dinero) recibidas desde otros hogares, el Estado o ISFLSH. Como se explicó al comienzo del apartado de ingresos, los ingresos totales por hogar presentados siempre son los imputados mediante *hot-deck*.

Programación 44: Ingreso total del hogar (imputado por hot-deck) sin arriendo imputado

```
//Ingreso total del hogar (imputado por hot-deck) sin arriendo imputado
//ING_TOTAL_HOG_HD
foreach v in ING_PRIMARIO_HOG_HD TRANSFERENCIAS_RECIBIDAS INGJ_HD {
replace `v'=0 if `v'==.
}
bys FOLIO: egen double
TRANSFERENCIAS_RECIBIDAS_HOG=sum(TRANSFERENCIAS_RECIBIDAS)
bys FOLIO: egen double INGJ_HD_HOG=sum(INGJ_HD)
gen double
ING_TOTAL_HOG_HD2=ING_PRIMARIO_HOG_HD+TRANSFERENCIAS_RECIBIDAS_HOG+INGJ_HD_HOG
replace ING_TOTAL_HOG_HD2=0 if ING_TOTAL_HOG_HD2==.
count if ING_TOTAL_HOG_HD!=ING_TOTAL_HOG_HD2 //debe ser cero
```

Para calcular el ingreso total del hogar imputado con *hot-deck* incluyendo el arriendo imputado, se debe agregar el arriendo imputado por la vivienda principal (AI_VP) y secundaria (AI_VS) a la variable calculada mediante la Programación 44 (ingreso total del hogar imputado). La Programación 45 utiliza elementos calculados en la Programación 44.

Programación 45: Ingreso total del hogar imputado con arriendo imputado

```
use "BASE_PERSONAS_VIII_EPF.dta", clear
//Ingreso total del hogar (imputado por hot-deck) con arriendo imputado
//ING_TOTAL_HOG_HD_AI
foreach v in ING_PRIMARIO_HOG_HD TRANSFERENCIAS_RECIBIDAS INGJ_HD AI_VP AI_VS {
replace `v'=0 if `v'==. | `v'<0
}
bys FOLIO: egen double
```

```

TRANSFERENCIAS_RECIBIDAS_HOG=sum(TRANSFERENCIAS_RECIBIDAS)
bys FOLIO: egen double INGJ_HD_HOG=sum(INGJ_HD)
gen double
ING_TOTAL_HOG_HD_AI2=ING_PRIMARIO_HOG_HD+TRANSFERENCIAS_RECIBIDAS_HOG+INGJ_HD_HOG+AI_VP+AI_VS
replace ING_TOTAL_HOG_HD_AI2=0 if ING_TOTAL_HOG_HD_AI2==.
count if ING_TOTAL_HOG_HD_AI!=ING_TOTAL_HOG_HD_AI2 //debe ser cero

```

Para conseguir el ingreso total del hogar imputado *per cápita* se divide el ingreso total del hogar imputado por *hot-deck*, o su símil que incluye el arriendo imputado, por la variable NPERSONA, que cuenta la cantidad de miembros al interior del hogar.

Programación 46: Ingreso total *per cápita* del hogar imputado con y sin arriendo imputado

```

//Ingreso total per cápita del hogar imputado sin arriendo imputado
//ING_TOTAL_HOG_HD_PC
gen double ING_TOTAL_HOG_HD_PC2=ING_TOTAL_HOG_HD/NPERSONA
count if ING_TOTAL_HOG_HD_PC!= ING_TOTAL_HOG_HD_PC2 //debe ser cero
//Ingreso total per cápita del hogar imputado con arriendo imputado
//ING_TOTAL_HOG_HD_PC_AI
gen ING_TOTAL_HOG_HD_PC_AI2=ING_TOTAL_HOG_HD_AI/NPERSONA
count if ING_TOTAL_HOG_HD_PC_AI!= ING_TOTAL_HOG_HD_PC_AI2 //debe ser cero

```

5.5.1.6 Ingreso disponible del hogar imputado por *hot-deck*

El ingreso disponible del hogar corresponde a la sumatoria de las fuentes de ingreso que el hogar percibió durante el mes anterior a la aplicación de la encuesta (ingresos totales) menos las transferencias corrientes emitidas.

a) *Transferencias corrientes emitidas*

Las transferencias corrientes emitidas corresponden a la suma de las transferencias monetarias corrientes emitidas a otros hogares o a Instituciones Sin Fines de Lucro al Servicio de los Hogares, así como también las contribuciones a la seguridad social (ya sean obligatorias o voluntarias). Se excluyen de las transferencias corrientes emitidas las transferencias en especies a otros hogares, pues se consideran gasto y se codifican en la Base de gastos como productos o servicios según corresponda. En la base de datos, la variable TRANSFERENCIAS_EMITIDAS hace referencia solo a las transferencias corrientes emitidas a otros hogares y a Instituciones sin Fines de Lucro que Sirven a los Hogares. Las transferencias emitidas se pueden calcular del modo que se muestra en la Programación 47:

Programación 47: Transferencias corrientes emitidas

```

//Transferencias corrientes emitidas //TRANSFERENCIAS_EMITIDAS
foreach v in TE02 TE05 TE06 TE08 SALUD PREVISION IMPUESTO OT05 OT06 INGDJ_HD
INGJ_HD {
replace `v' =0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
//Transferencias corrientes emitidas a otros hogares (monetarias)

```

```

gen double TRANS_EM1=TE02+TE05
replace TRANS_EM1=TE02+TE06 if TE05==0 & TE06>0
//Transferencias corrientes emitidas a ISFLH (monetarias)
gen double TRANS_EM2=TE08
//Contribuciones a la seguridad social (ya sean obligatorias como voluntarias).
No computan en la suma de la variable TRANSFERENCIAS_EMITIDAS
gen double TRANS_EM3=SALUD+PREVISION+IMPUESTO+OT05+OT06 +(INGJ_HD-INGDJ_HD)
//Se suman todas las transferencias emitidas
gen double TRANSFERENCIAS_EMITIDAS2=TRANS_EM1+TRANS_EM2
count if TRANSFERENCIAS_EMITIDAS!=TRANSFERENCIAS_EMITIDAS2 //debe ser cero

```

El ingreso neto por transferencias es el resultado de la suma de las transferencias corrientes recibidas por el hogar, menos las transferencias emitidas. Este ingreso puede resultar en un valor negativo. Se puede calcular de la siguiente forma:

Programación 48: Ingresos por transferencias netas

```

//Ingresos por transferencias netas //INGT_NETA
foreach v in TRANSFERENCIAS_RECIBIDAS TRANSFERENCIAS_EMITIDAS {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder restarlas
}
gen double INGT_NETA2=TRANSFERENCIAS_RECIBIDAS-TRANSFERENCIAS_EMITIDAS
replace INGT_NETA2=0 if INGT_NETA2==.
count if INGT_NETA!=INGT_NETA2 //debe ser cero

```

b) Agregado del Ingreso disponible del hogar

El ingreso disponible del hogar imputado, con y sin arriendo imputado, se puede calcular de la siguiente manera:

Programación 49: Ingreso disponible del hogar imputado con y sin arriendo imputado

```

//Ingreso disponible del hogar (sin arriendo imputado) //ING_DISP_HOG_HD
foreach v in INGA_HD INGH_HD INGNP_HD INGPI_HD INGOTA INGDOTI INGP INGF
INGDJ_HD TRANSFERENCIAS_RECIBIDAS SALUD PREVISION IMPUESTO OT05 OT06
TRANSFERENCIAS_EMITIDAS AI_VP AI_VS {
replace `v'=0 if `v'==. | `v'<0 //Se reemplazan para poder sumarlas
}
gen double ING_DISP_PERSONA_HD2=INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+
INGDOTI+INGP+INGF+INGDJ_HD+TRANSFERENCIAS_RECIBIDAS-SALUD-PREVISION-IMPUESTO-
OT05-OT06-TRANSFERENCIAS_EMITIDAS
bys FOLIO: egen double ING_DISP_HOG_HD2=total(ING_DISP_PERSONA_HD2)
count if ING_DISP_HOG_HD!=ING_DISP_HOG_HD2 //debe ser cero

//Ingreso disponible del hogar imputado (con arriendo imputado)
//ING_DISP_HOG_HD_AI
g double ING_DISP_HOG_HD_AI2=ING_DISP_HOG_HD+AI_VP+AI_VS
count if ING_DISP_HOG_HD_AI!=ING_DISP_HOG_HD_AI2 //debe ser cero

```

Para calcular el ingreso disponible del hogar imputado *per cápita*, se divide el ingreso disponible del hogar imputado por *hot-deck* por la variable NPERSONA, que cuenta la cantidad de miembros al interior del hogar. A continuación la versión de esta variable con y sin arriendo imputado:

Programación 50: Ingreso disponible *per cápita* del hogar imputado con y sin arriendo imputado

```

//Ingreso disponible per cápita del hogar imputado sin arriendo imputado
//ING_DISP_HOG_HD_PC
gen double ING_DISP_HOG_HD_PC2=ING_DISP_HOG_HD/NPERSONA
la var ING_DISP_HOG_HD_PC2 "Ingreso disponible per cápita del hogar imputado
(sin arriendo imputado)"
count if ING_DISP_HOG_HD_PC!= ING_DISP_HOG_HD_PC2 //debe ser cero
//Ingreso disponible per cápita del hogar imputado con arriendo imputado
//ING_DISP_HOG_HD_PC_AI
gen double ING_DISP_HOG_HD_PC_AI2=ING_DISP_HOG_HD_AI/NPERSONA
la var ING_DISP_HOG_HD_PC_AI2 "Ingreso disponible per cápita del hogar imputado
(con arriendo imputado)"
count if ING_DISP_HOG_HD_PC_AI!= ING_DISP_HOG_HD_PC_AI2 //debe ser cero

```

5.5.2 Gasto

El gasto de consumo final de los hogares, que es el concepto de gasto medido en la EPF, cubre una amplia gama de bienes y servicios, los que se pueden desagregar según la Clasificación del Consumo Individual por Finalidades (variable CCIF en la base). Mientras que a nivel de división, clase y subclase se adopta la clasificación internacional, para cada versión de la encuesta esta clasificación es adaptada a nivel de subclase y producto para dar cuenta de mejor manera de los patrones de consumo en Chile (ver apartado 3.4). El gasto de consumo final de los hogares se entrega detallado a nivel de producto a los usuarios a través de la Base de gastos y, como una variable agregada, en la Base de personas.

En la primera sección de este apartado se explica al usuario cómo obtener los agregados principales del gasto por hogar. Luego se aborda la estructura del gasto total por hogar (a niveles funcionales de la CCIF).

5.5.2.1 Gasto total por hogar

Para obtener el gasto total por hogar imputado por *hot-deck* (HD), se puede utilizar tanto la Base de gastos como la Base de personas. Cabe mencionar que el cálculo de este tabulado se obtiene de una forma más directa en la Base de personas.

Cuando se quiere obtener agregados de gasto sin arriendo imputado a partir de la Base de gastos, es preciso tener en cuenta que se deben eliminar los códigos CCIF que reflejan este concepto de gasto, que son:

- 04.2.1.01.01: Arriendo imputado de la vivienda principal solo para propietarios.
- 04.2.2.01.01: Arriendo imputado de la vivienda principal para hogares que no pagan arriendo y no son propietarios.
- 04.2.2.01.02: Arriendo imputado de la segunda vivienda (solo para propietarios).

Al comienzo de este manual, mediante la Programación 3, se indicó cómo definir las rutas de acceso que debe emplear Stata para utilizar las bases de datos. Si la ruta donde se encuentra la Base de gasto fue especificada previamente, mediante el comando `cd`, y corresponde a la misma

ruta señalada para acceder a la Base de personas, bastará solo con abrir la base como se indica al comienzo de las programaciones siguientes.

Programación 51: Gasto total por hogar imputado por HD sin arriendo imputado

```
//Gasto total por hogar imputado por HD sin arriendo imputado a partir de la
base de gastos//GASTOT_HD
use "BASE_GASTOS_VIII_EPF.dta", clear //Se abre la base de gastos
drop if (CCIF=="04.2.1.01.01" | CCIF=="04.2.2.01.01" | CCIF=="04.2.2.01.02")
//Se quitan los códigos asociados al arriendo imputado
bys FOLIO: egen double GASTOT_HD=total(GASTO) //Se genera el gasto total por
hogar
bys FOLIO (CCIF): gen aux=1 if _n==1 & FOLIO!=" " //Se etiqueta una observación
por hogar
tabstat GASTOT_HD [aweight=FE] if aux==1, stat(mean) f(%9,0fc) by(ZONA)
//Tabulado

//Para lograr el mismo tabulado mediante la base personas
use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar base de datos de persona
tabstat GASTOT_HD [aweight=FE] if JHOGAR==1, stat(mean) f(%9,0fc) by(ZONA)
//Tabulado
```

Programación 52: Gasto total por hogar imputado por HD con arriendo imputado

```
//Gasto total por hogar imputado por HD con arriendo imputado a partir de la
base de gastos//GASTOT_HD_AI
use "BASE_GASTOS_VIII_EPF.dta", clear //Se abre la base de gastos
bys FOLIO: egen double GASTOT_HD_AI=total(GASTO) //Se genera el gasto total por
hogar
bys FOLIO (CCIF): gen aux=1 if _n==1 & FOLIO!=" " //Se etiqueta una observación
por hogar
tabstat GASTOT_HD_AI [aweight=FE] if aux==1, stat(mean) f(%9,0fc) by(ZONA)
//Tabulado

//Para lograr el mismo tabulado mediante la base personas
use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar base de datos de persona
tabstat GASTOT_HD_AI [aweight=FE] if JHOGAR==1, stat(mean) f(%9,0fc) by(ZONA)
```

En la Base de personas vienen construidas las variables gasto total *per cápita* por hogar imputado HD sin arriendo imputado (GASTOT_HD_PC) y gasto total *per cápita* por hogar imputado HD con arriendo imputado (GASTOT_HD_PC_AI). Ambas variables se obtuvieron dividiendo el gasto total del hogar imputado por *hot-deck* (sin o con arriendo imputado, respectivamente) por el número de persona en el hogar NPERSONA.

Programación 53: Gasto total *per cápita* por hogar imputado HD con y sin arriendo imputado

```
use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar base de datos de persona
//Gasto per cápita del hogar imputado sin arriendo imputado //GASTOT_HD_PC
gen double GASTOT_HD_PC2=GASTOT_HD/NPERSONA
count if GASTOT_HD_PC!=GASTOT_HD_PC2 //debe ser cero
la var GASTOT_HD_PC "Gasto per cápita del hogar imputado por hot-deck, sin
arriendo imputado"

//Gasto per cápita del hogar imputado con arriendo imputado //GASTOT_HD_PC_AI
gen double GASTOT_HD_PC_AI2=GASTOT_HD_AI/NPERSONA
count if GASTOT_HD_PC_AI!=GASTOT_HD_PC_AI2 //debe ser cero
```

```
la var GASTOT_HD_PC_AI "Gasto per cápita del hogar imputado por hot-deck, con arriendo imputado"
```

5.5.2.2 Estructura del gasto de los hogares imputado por hot-deck

Para calcular el gasto en sus distintas desagregaciones (y no solo el gasto a nivel del hogar), se debe trabajar de manera conjunta la Base de gastos y la Base CCIF VIII EPF. Para fusionar la Base CCIF con la Base de gastos sin modificar la Programación 54, ambas deben estar guardadas en la misma ubicación que se haya especificado previamente mediante el comando cd.

A continuación se detallan los pasos que deben seguirse para transformar la base de datos y obtener resultados a nivel de gasto promedio por hogar según división para el total de capitales regionales.

Programación 54: Estructura del gasto promedio de los hogares según división

```
//Gasto promedio por hogar a nivel de divisiones para el total de capitales regionales (Sin Arriendo Imputado)
use "BASE_GASTOS_VIII_EPF.dta", clear
drop if (CCIF=="04.2.1.01.01" | CCIF=="04.2.2.01.01" | CCIF=="04.2.2.01.02")
//Se quitan los códigos asociados al arriendo imputado
gen double GASTO_EXP=GASTO*FE //Se genera el gasto expandido por hogar
bysort FOLIO: gen AUX=_n==1 //Se genera el número total de hogares
gen double AUX2=AUX*FE //Se genera el número de hogares que representa cada hogar encuestado
egen double Total_hogares_VIII=sum(AUX2) //Número total de hogares encuestados (expandido)
collapse (sum) GASTO_EXP, by(D /*G C SC P*/ Total_hogares_VIII)
gen double GASTO_D= GASTO_EXP/ Total_hogares_VIII //Variable con el gasto promedio por hogar por división
g G="0"
g C="0"
g SC="00"
g P="00"
merge 1:m D G C SC P using "CCIF_VIII_EPF.dta" //unión con la Base CCIF para obtener las glosas al nivel especificado
keep if _merge==3 //Elimina los bienes y servicios del resto de las categorías (Grupo-Clase-Subclase-Producto) y aquellos que no tuvieron gasto
drop _merge Total_hogares_VIII GASTO_EXP /*D*/ G C SC P //CCIF
order CCIF GLOSA GASTO_D
format GASTO_D %15,0fc
list, clean noobs
```

En la anterior programación, las variables G, C, SC y P son variables auxiliares que sirven para unir la Base de gastos con la Base CCIF, donde se encuentran las glosas de la estructura completa del

clasificador³². Para obtener el tabulado incluyendo el arriendo imputado en el gasto de los hogares, se debe eliminar la instrucción que borra los códigos de arriendo imputado (códigos "04.2.1.01.01", "04.2.2.01.01" y "04.2.2.01.02").

La siguiente programación muestra cómo convertir la información obtenida en la programación anterior en un gráfico de barras para la estructura del gasto (en monto y porcentaje).

Programación 55: estructura del gasto de los hogares

```
//Gráfico de barra para el gasto promedio por División (monto sin arriendo
imputado)
graph bar GASTO_D, over(D) bar(1,color(gs12) fintensity(inten60) lwidth(medium))
///
yttitle("Gasto por división (en miles)") yla(0 "0" 5e4 "50" 1e5 "100" 1.5e5 "150"
2e5 "200" 2.5e5 "250", ang(h) grid labsize(small)) ///
blabel(total, format(%15,0fc)) bltitle("(Divisiones del CCIF)") scheme(s1color)
title("Gasto promedio por División", span size(large)) subtitle("(sin arriendo
imputado)", span size(medium))
graph export "Gasto promedio por D.png", replace //Se exporta en formato png
//Gráfico de barra para el gasto promedio por División (porcentaje sin arriendo
imputado)
egen double GASTO_T=sum(GASTO_D)
g double PORC_D=(GASTO_D/GASTO_T)*100
graph bar PORC_D, over(D) bar(1,color(gs12) fintensity(inten60) lwidth(medium))
///
yttitle("Gasto por división (en % respecto al gasto total)") yla(0 "0" 5 "5" 10
"10" 15 "15" 20 "20", ang(h) grid labsize(small)) ///
blabel(total, format(%15,2fc)) bltitle("(Divisiones del CCIF)") scheme(s1color)
title("Estructura del gasto promedio por División", span size(large))
subtitle("(sin arriendo imputado)", span size(medium))
graph export "Estructura del gasto promedio por D.png", replace //Se exporta en
formato png
```

Si se desea graficar el gasto incorporando el arriendo imputado, ya sea para el monto de gasto o para el porcentaje de gasto por división, no se deben eliminar los códigos de gasto asociados al arriendo imputado. Por el contrario, se recomienda ajustar los valores de las etiquetas del eje de las ordenadas, tal como se muestra en la siguiente programación.

Programación 56: Estructura del gasto promedio de los hogares según grupo

```
//Con arriendo imputado:
use "BASE_GASTOS_VIII_EPF.dta", clear
gen double GASTO_EXP=GASTO*FE //Se genera el gasto expandido por hogar
bysort FOLIO: gen AUX=_n==1 //Se genera el número total de hogares
gen double AUX2=AUX*FE //Se genera el número de hogares que representa cada
hogar encuestado
egen double Total_hogares_VIII=sum(AUX2) //Número total de hogares encuestados
```

³² La Base de Gastos incorpora las glosas a nivel de producto, por lo tanto, si se requiere las glosas a otros niveles de desagregación, la base de datos debe ser combinada con la Base CCIF VIII EPF.

```

(expandido)
collapse (sum) GASTO_EXP, by(D /*G C SC P*/ Total_hogares_VIII)
gen double GASTO_D_AI= GASTO_EXP/ Total_hogares_VIII //Variable con el gasto
promedio por hogar por división

graph bar GASTO_D_AI, over(D) bar(1,color(gs12) fintensity(inten60)
lwidth(medium)) ///
yttitle("Gasto por división (en miles)") yla(0 "0" 5e4 "50" 1e5 "100" 1.5e5
"150" 2e5 "200" 2.5e5 "250" 3e5 "300" 3.5e5 "350" 4e5 "400", ang(h) grid
labsize(small)) ///
lblabel(total, format(%15,0fc)) b1title("(Divisiones del CCIF)") scheme(s1color)
/*
*/ title("Gasto promedio por División", span size(large)) subtitle("(con
arriendo imputado)", span size(medium))
graph export "Gasto promedio por D con arriendo imputado.png", replace //Se
exporta en formato png
//Gráfico de barra para el gasto promedio por División (porcentaje sin arriendo
imputado)
egen double GASTO_T_AI=sum(GASTO_D_AI)
g double PORC_D_AI=(GASTO_D_AI/GASTO_T_AI)*100
graph bar PORC_D_AI, over(D) bar(1,color(gs12) fintensity(inten60) lwidth(medium))
///
yttitle("Gasto por división (en % respecto al gasto total)") yla(0 "0" 5 "5" 10
"10" 15 "15" 20 "20" 25 "25" 30 "30", ang(h) grid labsize(small)) ///
lblabel(total, format(%15,2fc)) b1title("(Divisiones del CCIF)") scheme(s1color)
/*
*/ title("Estructura del gasto promedio por División", span size(large))
subtitle("(con arriendo imputado)", span size(medium))
graph export "Estructura del gasto promedio por D con arriendo imputado.png",
replace //Se exporta en formato png

```

Si se deseara obtener el gasto a un nivel menos desagregado, por ejemplo a nivel de grupo de gasto de la CCIF, se debería volver a ejecutar la Programación 54 con ajustes en los parámetros, tal como se ejemplifica a continuación:

Programación 57: Estructura del gasto promedio de los hogares según grupo

```

//Gasto promedio por hogar a nivel de divisiones para el total de capitales
regionales (Sin Arriendo Imputado)
use "BASE_GASTOS_VIII_EPF.dta", clear
drop if (CCIF=="04.2.1.01.01" | CCIF=="04.2.2.01.01" | CCIF=="04.2.2.01.02")
//Se quitan los códigos asociados al arriendo imputado
gen double GASTO_EXP=GASTO*FE //Se genera el gasto expandido por hogar
bysort FOLIO: gen AUX=_n==1 //Se genera el número total de hogares
gen double AUX2=AUX*FE //Se genera el número de hogares que representa cada
hogar encuestado
egen double Total_hogares_VIII=sum(AUX2) //Número total de hogares encuestados
(expandido)
collapse (sum) GASTO_EXP, by(D G /*C SC P*/Total_hogares_VIII) //Colapsando
según las distintos variables censuradas se consiguen distintos niveles de
desagregación
gen double GASTO_D= GASTO_EXP/ Total_hogares_VIII //Variable con el gasto
promedio por hogar por división
//g G="0" //Se censura esta eliminación, ya que se quiere obtener información
a nivel de grupo de gasto
g C="0"
g SC="00"
g P="00"
merge 1:m D G C SC P using "CCIF_VIII_EPF.dta" //unión con la Base CCIF para
obtener las glosas al nivel especificado

```

```

keep if _merge==3 //Elimina los bienes y servicios del resto de las categorías
(Grupo-Clase-Subclase-Producto) y aquellos que no tuvieron gasto
drop _merge Total_hogares_VIII GASTO_EXP D G C SC P
order CCIF GLOSA GASTO_D
format GASTO_D %15,0fc
list, clean noobs

```

Además de los procedimientos antes explicados, para el usuario podría ser de interés combinar la información del gasto desagregado con las características del hogar y de sus integrantes (Base de personas). En la siguiente sección se ejemplifica cómo realizar combinaciones entre las diferentes bases para generar tabulados y gráficos de interés.

5.5.3 Gasto e ingreso según características del sustentador o sustentadora principal del hogar

Con la Base de personas de la VIII EPF se pueden realizar diferentes tabulados según las características del hogar y sus integrantes. Por ejemplo, tabulados de gastos según ciertos atributos del sustentador o sustentadora principal³³ del hogar (SPRINCIPAL =1), obtener el gasto e ingreso promedio según sexo del sustentador o sustentadora principal del hogar por zona (Gran Santiago y/o Resto de capitales regionales), entre otros. En la Programación 58 se muestra, por ejemplo, cómo tabular gasto e ingreso promedio según el sustentador o sustentadora principal del hogar.

Programación 58: Tabulados de gasto e ingreso por sexo del sustentador o sustentadora principal del hogar

```

use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar base de datos de persona
//Gasto promedio por hogar para el total de capitales regionales por sexo del
sustentador o sustentadora principal del hogar (sin arriendo imputado)
tabstat GASTOT_HD [w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss by(SEXO)

//Ingreso promedio por hogar para el total de capitales regionales por sexo del
sustentador o sustentadora principal del hogar
tabstat ING_DISP_HOG_HD [w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss
by(SEXO)

```

³³ El sustentador principal corresponde al miembro del hogar que más aporta al presupuesto familiar, evaluando los últimos seis meses. No es, necesariamente, el integrante que recibe más ingresos o el designado como jefe o jefa de hogar. Para más información respecto al uso del sustentador principal para los tabulados de caracterización de los hogares, consultar el documento de principales resultados de la encuesta (INE Chile, 2018d).

A partir de este tabulado se pueden realizar diferentes combinaciones, tanto en la variable de interés como en la característica del sustentador o sustentadora principal del hogar que se quiera utilizar como variable de clasificación, reemplazando las variables en la programación ofrecida.

Con la ayuda del comando *foreach* se puede automatizar la ejecución de una función a cada variable listada. En la Programación 59, por ejemplo, se muestra cómo tabular los tipos de ingreso y gasto promedio (con y sin arriendo imputado), según sexo del sustentador o sustentadora principal del hogar, obteniendo como resultado tabulados por separado o un tabulado conjunto para cada variable de la lista (GASTOT_HD; GASTOT_HD_AI; ING_DISP_HOG_HD; ING_DISP_HOG_HD_AI).

Programación 59: Tabulados de gasto e ingreso por sexo del sustentador o sustentadora principal del hogar y área de estimación

```
//Se pueden tabular por separado
foreach x in GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI {
  bys ZONA: tabstat `x' [w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss
  by(SEXO)
}
//También se pueden tabular en conjunto
bys ZONA: tabstat GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI
[w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss by(SEXO)
//o bien:
table SEXO if SPRINCIPAL==1 [pweight=FE], by(ZONA) cont(mean GASTOT_HD mean
GASTOT_HD_AI mean ING_DISP_HOG_HD mean ING_DISP_HOG_HD_AI) row f(%9,0fc)
```

El siguiente ejemplo muestra cómo realizar tabulados individuales de ingresos y gastos con la información publicada, según ciertas características del sustentador o sustentadora principal del hogar utilizando tramos etarios (ver Programación 8).

Programación 60: Tabulados de ingresos y gastos según tramo etario del sustentador o sustentadora principal del hogar

```
//Luego de generar los Tramos Etarios (utilizando la programación 8)
//Tabular el gasto e ingreso promedio por hogar para el total capitales
regionales y por tramos etarios del sustentador o sustentadora principal del
hogar
foreach x in GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI {
  tabstat `x' [w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss by(EDAD_T)
}
//En un solo comando
tabstat GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI [w=FE] if
SPRINCIPAL==1, stat(mean) f(%9,0fc) miss by(EDAD_T)
//o bien
table EDAD_T if SPRINCIPAL==1 [pweight=FE], cont(mean GASTOT_HD mean
GASTOT_HD_AI mean ING_DISP_HOG_HD mean ING_DISP_HOG_HD_AI) row f(%9,0fc)
```

Para tabular el gasto e ingreso promedio del hogar para el total de capitales regionales según el sexo y el nivel educativo del sustentador o sustentadora principal del hogar, se utiliza la variable

EDU_T, que es la que se emplea para los tabulados oficiales de la encuesta y que fue construida mediante la Programación 11.

Programación 61: Tabulados de ingresos y gastos según sexo y nivel educacional del sustentador o sustentadora principal del hogar

```
//Luego de generar los Niveles Educativos (utilizando la programación 11)
//Gasto e ingreso total con y sin arriendo imputado según sexo y nivel
educativo del sustentador o sustentadora principal del hogar
foreach x in GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI {
  bysort SEXO: tabstat `x' [w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss
  by(EDU_T)
}
//En un solo comando
bys SEXO: tabstat GASTOT_HD GASTOT_HD_AI ING_DISP_HOG_HD ING_DISP_HOG_HD_AI
[w=FE] if SPRINCIPAL==1, stat(mean) f(%9,0fc) miss by(EDU_T)
//o bien
bys SEXO: table EDU_T if SPRINCIPAL==1 [pweight=FE], cont(mean GASTOT_HD mean
GASTOT_HD_AI mean ING_DISP_HOG_HD mean ING_DISP_HOG_HD_AI) row f(%9,0fc)
```

Los gráficos permiten visualizar una o más variables de interés, mostrando cómo se distribuyen estas estimaciones según las características que se seleccionen del sustentador o sustentadora principal del hogar, del jefe o jefa de hogar o de alguna persona de referencia. Se pueden seleccionar características sociodemográficas y económicas presentes en la Base de personas, por ejemplo, sexo, edad, educación, condición de actividad económica, área de estimación, entre otros.

El siguiente gráfico muestra cómo realizar un gráfico de barras para el gasto promedio según área de estimación, utilizando además la desagregación por sexo del sustentador o sustentadora principal del hogar. En esta programación se puede incluir más variables, por ejemplo, el ingreso disponible.

Programación 62: Gráfico del gasto promedio

```
//El orden en cómo se establecen el comando over determina el agrupamiento de
las barras
graph hbar (mean) GASTOT_HD if SPRINCIPAL==1 [pweight=FE], over(ZONA,
label(labsize(vsmall))) over(SEXO, label(labsize(small))) ///
bar(1,color(gs12) fintensity(inten60) lwidth(small)) ytitle("Gasto total del
hogar (en miles de pesos)", size(vsmall)) scheme(s1color) ///
ylabel(0 "0" 5e5 "500" 1e6 "1.000" 1.5e6 "1.500",labsize(small))
format(%10,0fc) blabel(total, format(%15,0fc)) ///
title("Gasto total del hogar según sexo del sustentador principal del hogar por
zona", span size(med) color (gs0)) subtitle ("(excluye arriendo imputado)",
span size(medium))
graph export "GASTOT_HD.png", replace //Se exporta en formato png
```

6 QUINTILIZACIÓN POR INGRESO DE LAS BASES DE DATOS

En las bases que se encuentran publicadas no se entrega la variable quintil. Sin embargo, en esta sección se explica la manera de replicar las quintilizaciones utilizadas para la tabular los resultados publicados. Los quintiles corresponden a una variable de orden que indica en qué 20% de la distribución de los ingresos (del hogar o *per cápita*) se encuentra el hogar encuestado. El orden se genera de forma ascendente en relación con el ingreso respecto al que se construyen. Cabe aclarar que no corresponde a una variable de estratificación socioeconómica, ya que para lograr una estratificación se deben incorporar al análisis otras dimensiones. El primer quintil de ingresos representa al 20% de hogares de menores ingresos, mientras que el quinto quintil representa al 20% de los hogares de mayores ingresos.

Para la tabulación de resultados de la encuesta se utilizan cuatro definiciones de ingresos por los cuales se realiza el ordenamiento de los hogares: (1) ingreso disponible total mensual por hogar sin arriendo imputado, (2) ingreso disponible total mensual por hogar con arriendo imputado, (3) ingreso disponible *per cápita* mensual por hogar sin arriendo imputado y, finalmente, (4) ingreso disponible *per cápita* mensual por hogar con arriendo imputado. Cada una de estas definiciones se aplicaron para tres ámbitos geográficos, que corresponden a las áreas de estimación de la encuesta: Gran Santiago, Resto de capitales regionales (en su conjunto) y Total de capitales regionales (Gran Santiago + Resto de capitales regionales).

Como resultado, los tabulados presentados en el micrositio de la encuesta fueron realizados según doce quintilizaciones diferentes, que corresponden a las combinaciones entre las cuatro definiciones de la variable principal de ordenamiento y las tres desagregaciones geográficas. Para la quintilización correcta de la base a nivel personas, es necesario considerar que al ordenar la base se identificarán hogares que están en el límite del corte entre un quintil y el siguiente. Dado lo anterior, para obtener el 20% exacto de hogares por quintil³⁴, es necesario particionar el factor de expansión de cuatro hogares. Esto se realiza sobre la variable del factor de expansión (FE), ya que las demás características se mantienen sin variación. Como se desprende de lo anterior, los hogares que quedan en los límites de dos quintiles quedan en ambos, pero repartiendo el peso de su factor de expansión.

³⁴ El proceso de quintilización obtiene cinco grupos con la misma cantidad de hogares cada uno, representando cada uno de estos grupos un 20% de los hogares respecto al total nacional. Estos grupos son obtenidos a partir de una determinada ordenación realizada a través de los ingresos, según se explicó con anterioridad. El primero de estos grupos contiene al 20% de los hogares de menos ingresos, mientras que el quinto quintil aglutina el 20% de los hogares más ricos según la variable de ordenación.

A continuación se ofrece una programación para replicar la quintilización de ingresos a nivel nacional de capitales regionales, utilizando como variable de ordenamiento el ingreso disponible total por hogar sin arriendo imputado:

Programación 63: Quintilización a nivel hogar por ingreso disponible total mensual por hogar sin arriendo imputado

```
//Ejemplo: Ámbito: Nacional
//Variable de ordenamiento: Ingreso disponible total por hogar sin arriendo
imputado
use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar Base de personas
set type double //Importante para no tener diferencias en los cálculos
keep if JHOGAR==1 //Nivel hogar

*1.-DISPONIBLE HOGAR SIN ARRIENDO IMPUTADO ** NSAI GSSAI RRS AI
gen double vector_orden=ING_DISP_HOG_HD
replace vector_orden=GASTOT_HD if ID_MISSING==1
//Ordenamiento de los hogares
sort vector_orden ING_DISP_HOG_HD_PC GASTOT_HD FOLIO //Se ordena según el orden
que se generará la quintilización
g double s=sum(FE)
g double c=(s)/(s[_N])
g double tag_QNSAI=0
g double w2=.
foreach 1 in 20 40 60 80 {
  replace tag_QNSAI=1 if c>=`1'/100 & c[_n-1]<`1'/100
  replace w2=s-(`1'/100*s[_N]) if c>=`1'/100 & c[_n-1]<`1'/100
}
g double w1=FE-w2
g double id=_n
expand 2 if w2<.
g double FW_QNSAI=FE
bys id: replace FW_QNSAI=w1 if _n==1 & w1<.
by id: replace FW_QNSAI=w2 if _n==2
g double sum=sum(FW_QNSAI)
g double cum=sum/sum[_N]
g double q_NSAI=1 if cum<=0.20000001
replace q_NSAI=2 if cum<=0.40000001 & cum>0.20000001
replace q_NSAI=3 if cum<=0.60000001 & cum>0.40000001
replace q_NSAI=4 if cum<=0.80000001 & cum>0.60000001
replace q_NSAI=5 if cum<=1 & cum>0.80000001
list FOLIO FE FW_QNSAI q_NSAI ING_DISP_HOG_HD ING_DISP_HOG_HD_PC GASTOT_HD cum
if tag_QNSAI==1
tabstat ING_DISP_HOG_HD_PC, stat(mean sd min max) by(q_NSAI)
keep FOLIO PERSONA FE FW_QNSAI q_NSAI tag_QNSAI ID_MISSING
sort FOLIO PERSONA
tab FOLIO if tag_QNSAI==1
//Se guarda una base auxiliar a nivel hogar
bysort FOLIO: gen aux=_n-1
label var q_NSAI "Quintil Nacional sin AI"
save "BASE_QNSAI_HOGAR.dta", replace

use "BASE_PERSONAS_VIII_EPF.dta", clear
merge m:m FOLIO PERSONA using "BASE_QNSAI_HOGAR.dta"
replace _merge=. if _merge!=3
by FOLIO, sort: egen sum_merge=total(_merge)
expand 2 if sum_merge==6 & PERSONA!=1
bys FOLIO PERSONA: gen id=_n
bys FOLIO id: egen double FW_QNSAIH=total(FW_QNSAI)
bys FOLIO id: egen double q_NSAIH=total(q_NSAI)
bys FOLIO id: egen double tag_QNSAIH=total(tag_QNSAI)
```

```

drop FW_QNSAI q_NSAI _merge sum_merge id tag_QNSAI
ren FW_QNSAIH FW_QNSAI
ren q_NSAIH q_NSAI
ren tag_QNSAIH tag_QNSAI
tabstat FW_QNSAI if PERSONA==1, stat(sum) by(q_NSAI) format(%15,4fc) //Deben
ser todas iguales
keep FOLIO PERSONA FW_QNSAI q_NSAI
save "BASE_QNSAI_PERSONAS.dta", replace

//El comando para ordenar a los hogares para cada tipo de quintilización
restante sigue la siguiente lógica.
//Estos comandos pueden ser reemplazados al comienzo de la quintilización para
obtener las distintas quintilizaciones:
*1.-DISPONIBLE HOGAR SIN ARRIENDO IMPUTADO ** NSAI GSSAI RRS AI
gen double vector_orden=ING_DISP_HOG_HD
replace vector_orden=GASTOT_HD if ID_MISSING==1
//Ordenamiento de los hogares
sort vector_orden ING_DISP_HOG_HD_PC GASTOT_HD FOLIO
*2.-DISPONIBLE PER CÁPITA SIN ARRIENDO IMPUTADO ** NSAI_PC GSSAI_PC RRS AI_PC
gen double vector_orden=ING_DISP_HOG_HD_PC
replace vector_orden=GASTOT_HD_PC if ID_MISSING==1
//Ordenamiento de los hogares
sort vector_orden ING_DISP_HOG_HD_PC GASTOT_HD_PC FOLIO
*3.-DISPONIBLE HOGAR CON ARRIENDO IMPUTADO ** NCAI GSCAI RRCAI
gen double vector_orden=ING_DISP_HOG_HD_AI
replace vector_orden=GASTOT_HD_AI if ID_MISSING==1
//Ordenamiento de los hogares
sort vector_orden ING_DISP_HOG_HD_AI GASTOT_HD_AI FOLIO
*4.-DISPONIBLE PER CÁPITA CON ARRIENDO IMPUTADO ** NCAI_PC GSCAI_PC RRCAI_PC
gen double vector_orden=ING_DISP_HOG_HD_PC_AI
replace vector_orden=GASTOT_HD_PC_AI if ID_MISSING==1
//Ordenamiento de los hogares
sort vector_orden ING_DISP_HOG_HD_PC_AI GASTOT_HD_PC_AI FOLIO

```

En la Programación 64 se da un ejemplo de un tabulado que utiliza la quintilización realizada por la rutina anterior. En él se tabula la composición del ingreso por fuentes de ingreso para los quintiles de hogares. El ordenamiento escogido es el ingreso del hogar según fuentes de ingreso, total nacional (excluye arriendo imputado). Es importante que el usuario note que en caso de utilizar la base quintilizada, debe emplear el factor de expansión calculado en la Programación 63.

Programación 64: Composición del ingreso de los quintiles de hogares ordenados por ingreso del hogar según fuentes de ingreso, total nacional (excluyendo arriendo imputado)

```

use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar Base de personas
set type double //importante para no tener diferencias en los cálculos
merge 1:m FOLIO PERSONA using "BASE_QNSAI_PERSONAS.dta", keepusing(FW_QNSAI
q_NSAI) generate(MERGE)
//Se cambia el uso del factor al nuevo creado
tabstat FW_QNSAI if JHOGAR ==1, stat(sum) by(q_NSAI) format(%15,4fc)
foreach N in INGDA_HD INGDH_HD INGDNP_HD INGDPI_HD INGOTA INGDOTI INGP INGF
INGJ_HD INGT_NETA AI_VP AI_VS {
quietly replace `N'=0 if `N'==.
}
quietly gen double
INGDTRAB_HD=INGDA_HD+INGDH_HD+INGDNP_HD+INGDPI_HD+INGDOTA+INGDOTI //Ingresos de
la ocupación
quietly gen double INGDTD_HD2=INGDA_HD+INGDH_HD //esta variable ya existe en la
BD

```

```

count if INGDTD_HD!=INGDTD_HD2 //debe ser cero
quietly gen double INGDTI_HD2=INGDNP_HD + INGDPH_HD //esta variable ya existe
en la BD
count if INGDTI_HD!=INGDTI_HD2 //debe ser cero
quietly gen double INGDEMP_HD2=INGDTI_HD if AECISE==1 //esta variable ya existe
en la BD
replace INGDEMP_HD2=0 if INGDEMP_HD2==.
count if INGDEMP_HD!=INGDEMP_HD2 //debe ser cero
quietly gen double INGDTCP_HD2=INGDTI_HD if AECISE==2 //esta variable ya existe
en la BD
replace INGDTCP_HD2=0 if INGDTCP_HD2==.
count if INGDTCP_HD!=INGDTCP_HD2 //debe ser cero
quietly gen double INGDTRABOS=INGDOTA+INGDOTI //Ingresos disponibles del
trabajo de la ocupación secundaria y de ocupación del mes anterior que no
coincide con la ocupación declarada como principal
quietly gen double RENTAS_PROPIEDAD=INGP+INGF
quietly gen double ARRIENDO_IMPUTADO=AI_VP+AI_VS if (AI_VP>=0 & AI_VS>=0)
quietly replace ARRIENDO_IMPUTADO=AI_VP if (AI_VP>=0 & AI_VS<0)
quietly replace ARRIENDO_IMPUTADO=AI_VS if (AI_VP<0 & AI_VS>=0)
quietly replace ARRIENDO_IMPUTADO=0 if PERSONA!=1 | ARRIENDO_IMPUTADO==. //se
debe una obserbación por persona para trabajar los ingresos a nivel hogar

//Primera forma de presentar los ingresos del hogar
//(A+B) Ingreso total del hogar imputado por hot-deck (sin arriendo imputado)
bys FOLIO q_NSAI: egen double
ING_TOTAL_HOG_HD2=sum(INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI+INGJ_HD+
RENTAS_PROPIEDAD+TRANSFERENCIAS_RECIBIDAS)
//(A=1+2+3+4) Ingreso de la ocupación del hogar imputado por hot-deck
bys FOLIO q_NSAI: egen double
INGTRAB_HOG_HD=sum(INGA_HD+INGH_HD+INGNP_HD+INGPI_HD+INGOTA+INGDOTI)
//(1) Sueldos y Salarios del hogar
bys FOLIO q_NSAI: egen double INGTD_HOG_HD=sum(INGTD_HD)
//(2) Ing de Patrones y empleadores del hogar
bys FOLIO q_NSAI: egen double INGEMP_HOG_HD=sum(INGEMP_HD)
//(3) Ing de Cuenta propia del hogar
bys FOLIO q_NSAI: egen double INGTCR_HOG_HD=sum(INGTCP_HD)
//(4) Ingresos de otros trabajos del hogar
bys FOLIO q_NSAI: egen double INGTRABOS_HOG=sum(INGOTA+INGDOTI)

//(B=4+5+6) Ingreso de otras fuentes del hogar
bys FOLIO q_NSAI: egen double
ING_OTF_HOG_HD=sum(INGJ_HD+RENTAS_PROPIEDAD+TRANSFERENCIAS_RECIBIDAS)
//(4) Jubilaciones o pensiones de vejez del hogar
bys FOLIO q_NSAI: egen double INGJ_HD_HOG=sum(INGJ_HD)
//(5) Rentas de la propiedad del hogar
bys FOLIO q_NSAI: egen double RENTAS_PROPIEDAD_HOG=sum(RENTAS_PROPIEDAD)
//(6) Transferencias corrientes recibidas del hogar
bys FOLIO q_NSAI: egen double
TRANSFERENCIAS_RECIBIDAS_HOG=sum(TRANSFERENCIAS_RECIBIDAS)

//(C=7+8+9) Transferencias corrientes emitidas del hogar
bys FOLIO q_NSAI: egen double
TRANSFERENCIAS_EMITIDAS_HOG=sum(IMPUESTO+CONTRIBUCIONES+TRANSFERENCIAS_EMITIDAS
)
//(7) Impuesto a la renta del hogar
bys FOLIO q_NSAI: egen double IMPUESTO_HOG=sum(IMPUESTO)
//(8) Contribuciones del hogar
bys FOLIO q_NSAI: egen double CONTRIBUCIONES_HOG=sum(CONTRIBUCIONES)
//(9) Transferencias emitidas del hogar a Hog e ISFLSH
bys FOLIO q_NSAI: egen double TRANSFERENCIAS_HOG=sum(TRANSFERENCIAS_EMITIDAS)

//(D=A+B-C) Ingreso total Disponible del hogar imputado por hot-deck (sin
arriendo imputado)

```

```

gen double ING_DISP_HOG_HD2=ING_TOTAL_HOG_HD2-TRANSFERENCIAS_EMITIDAS_HOG

//Se agregan los labels a lo que se tabulará
label var ING_TOTAL_HOG_HD2 "Ing total del hogar (A+B)"
label var INGTRAB_HOG_HD "Ing de la ocupación del hogar(A=1+2+3+4)"
label var INGTD_HOG_HD "Sueldos y Salarios del hogar (1)"
label var INGEMP_HOG_HD "Ing de Patrones y empleadores del hogar (2)"
label var INGTCP_HOG_HD "Ing de Cuenta Propia del hogar (3)"
label var INGTRABOS_HOG "Ing de otros trabajos del hogar (4)"
label var ING_OTF_HOG_HD "Ing de otras fuentes del hogar (B=4+5+6)"
label var INGJ_HD_HOG "Ing jubilaciones o pensiones de vejez del hogar (4)"
label var RENTAS_PROPIEDAD_HOG "Ing rentas de la propiedad del hogar (5)"
label var TRANSFERENCIAS_RECIBIDAS_HOG "Transferencias corrientes recibidas del hogar (6)"
label var TRANSFERENCIAS_EMITIDAS_HOG "Transferencias corrientes emitidas del hogar (C=7+8+9)"
label var IMPUESTO_HOG "Impuesto a la renta del hogar (7)"
label var CONTRIBUCIONES_HOG "Contribuciones del hogar (8)"
label var TRANSFERENCIAS_HOG "Transferencias emitidas del hogar a Hog e ISFLSH (9)"
label var ING_DISP_HOG_HD2 "Ingreso Disponible del hogar (A+B-C)"

tabstat ING_TOTAL_HOG_HD2 INGTRAB_HOG_HD INGTD_HOG_HD INGEMP_HOG_HD
INGTCP_HOG_HD INGTRABOS_HOG_HD ING_OTF_HOG_HD INGJ_HD_HOG RENTAS_PROPIEDAD_HOG ///
TRANSFERENCIAS_RECIBIDAS_HOG TRANSFERENCIAS_EMITIDAS_HOG IMPUESTO_HOG
CONTRIBUCIONES_HOG TRANSFERENCIAS_HOG ING_DISP_HOG_HD2 [w=FW_QNSAI] if
JHOGAR==1, s(mean) format(%9,0fc) col(stat)
//Para generar el mismo tabulado pero considerando los quintiles
bys q_NSAI: tabstat ING_TOTAL_HOG_HD2 INGTRAB_HOG_HD INGTD_HOG_HD INGEMP_HOG_HD
INGTCP_HOG_HD INGTRABOS_HOG_HD ING_OTF_HOG_HD INGJ_HD_HOG RENTAS_PROPIEDAD_HOG ///
TRANSFERENCIAS_RECIBIDAS_HOG TRANSFERENCIAS_EMITIDAS_HOG IMPUESTO_HOG
CONTRIBUCIONES_HOG TRANSFERENCIAS_HOG ING_DISP_HOG_HD2 [w=FW_QNSAI] if
JHOGAR==1, s(mean) format(%9,0fc) col(stat)

ssc install fsum //esta programación instalará el comando fsum necesario para
ejecutar el próximo tabulado
fsum ING_TOTAL_HOG_HD2 INGTRAB_HOG_HD INGTD_HOG_HD INGEMP_HOG_HD INGTCP_HOG_HD
INGTRABOS_HOG_HD ING_OTF_HOG_HD INGJ_HD_HOG RENTAS_PROPIEDAD_HOG ///
TRANSFERENCIAS_RECIBIDAS_HOG TRANSFERENCIAS_EMITIDAS_HOG IMPUESTO_HOG
CONTRIBUCIONES_HOG TRANSFERENCIAS_HOG ING_DISP_HOG_HD2 [aw=FW_QNSAI] if
JHOGAR==1, s(mean) f(%9.0f) uselabel

//Para generar el mismo tabulado pero considerando los quintiles
bys q_NSAI: fsum ING_TOTAL_HOG_HD2 INGTRAB_HOG_HD INGTD_HOG_HD INGEMP_HOG_HD
INGTCP_HOG_HD INGTRABOS_HOG_HD ING_OTF_HOG_HD INGJ_HD_HOG RENTAS_PROPIEDAD_HOG ///
TRANSFERENCIAS_RECIBIDAS_HOG TRANSFERENCIAS_EMITIDAS_HOG IMPUESTO_HOG
CONTRIBUCIONES_HOG TRANSFERENCIAS_HOG ING_DISP_HOG_HD2 [aw=FW_QNSAI] if
JHOGAR==1, s(mean) f(%9.0f) uselabel

//Segunda forma de presentar los ingresos del hogar
//(A=1+2+3+4)Ingreso disponible de la ocupación del hogar imputado por hot-deck
bys FOLIO q_NSAI: egen double
INGDTRAB_HOG_HD=sum(INGDA_HD+INGDH_HD+INGDNP_HD+INGDPI_HD+INGDOTA+INGDOTI)
//(1)Sueldos y Salarios disponible del hogar
bys FOLIO q_NSAI: egen double INGDTD_HOG_HD=sum(INGDTD_HD)
//(2)Ing disponible de Patrones y empleadores del hogar
bys FOLIO q_NSAI: egen double INGDEMP_HOG_HD=sum(INGDEMP_HD)
//(3)Ing disponible de Cuenta propia del hogar
bys FOLIO q_NSAI: egen double INGDTCP_HOG_HD=sum(INGDTCP_HD)
//(4)Ing disponible de otros trabajos del hogar
bys FOLIO q_NSAI: egen double INGDTABOS_HOG_HD=sum(INGDOTA+INGDOTI)

//(B=4+5+6)Ingreso disponible de otras fuentes del hogar

```

```

bys FOLIO q_NSAI: egen double INGD_OTF_HOG_HD=sum(INGDJ_HD+ RENTAS_PROPIEDAD
+INGT_NETA)
//(4)Jubilaciones o pensiones de vejez disponible del hogar
bys FOLIO q_NSAI: egen double INGDJ_HD_HOG=sum(INGDJ_HD)
//(5)Rentas de la propiedad del hogar
bys FOLIO q_NSAI: egen double INGP_HOG=sum(INGP) //Ing por arriendo de
propiedades
bys FOLIO q_NSAI: egen double INGF_HOG=sum(INGF) //Ing por tenencia de
instrumentos financieros
//(6)Transferencias netas del hogar (recibidas-emitidas)
bys FOLIO q_NSAI: egen double TRANSFERENCIAS_NETAS_HOG=sum(INGT_NETA)

label var ING_DISP_HOG_HD2 "Ingreso Disponible del hogar (A+B)"
label var INGDTRAB_HOG_HD "Ingreso disponible de la ocupación del hogar
imputado por hot-deck (A=1+2+3+4)"
label var INGDTD_HOG_HD "Ing por sueldos y salarios disponibles del hogar (1)"
label var INGDEMP_HOG_HD "Ing disponible de Patrones y empleadores del hogar
(2)"
label var INGDTCP_HOG_HD "Ing disponible de Cuenta propia del hogar (3)"
label var INGDTRABOS_HOG "Ing disponible de otros trabajos del hogar (4)"
label var INGD_OTF_HOG_HD "Ingreso disponible de otras fuentes del hogar
(B=4+5a+5b+6)"
label var INGDJ_HD_HOG "Jubilaciones o pensiones de vejez disponible del hogar
(4)"
label var INGP_HOG "Ing por arriendo de propiedades (5a)"
label var INGF_HOG "Ing por tenencia de instrumentos financieros (5b)"
label var TRANSFERENCIAS_NETAS_HOG "Transferencias netas del hogar (recibidas-
emitidas) (6)"

//Para tabular la información generada
fsum ING_DISP_HOG_HD2 INGDTRAB_HOG_HD INGDTD_HOG_HD INGDEMP_HOG_HD
INGDTCP_HOG_HD INGDTRABOS_HOG INGD_OTF_HOG_HD INGDJ_HD_HOG INGP_HOG INGF_HOG
TRANSFERENCIAS_NETAS_HOG [aw=FW_QNSAI] if JHOGAR==1, s(mean) f(%9.0f) uselabel

//Para generar el mismo tabulado pero considerando los quintiles
bys q_NSAI: fsum ING_DISP_HOG_HD2 INGDTRAB_HOG_HD INGDTD_HOG_HD INGDEMP_HOG_HD
INGDTCP_HOG_HD INGDTRABOS_HOG INGD_OTF_HOG_HD INGDJ_HD_HOG INGP_HOG INGF_HOG
TRANSFERENCIAS_NETAS_HOG [aw=FW_QNSAI] if JHOGAR==1, s(N mean) f(%9.0f)
uselabel

```

A continuación se entrega una rutina sugerida a través de otro ejemplo, utilizando en esta oportunidad la Base de gastos y considerando que la Programación 63 fue ejecutada. En ella, se calcula el gasto total promedio por hogar por división, de acuerdo con el quintil al que se asocia el hogar.

Programación 65: Gasto total promedio por hogar por división y quintil

```

use "BASE_GASTOS_VIII_EPF.dta", clear //Utilizar Base de gastos
set type double //importante para no tener diferencias en los cálculos
expand 2 if FOLIO=="02344-1"|FOLIO=="09954-1"|FOLIO=="11678-1"|FOLIO=="14723-
1", generate(aux) //folios que se deben expandir
//La expansión de los hogares de corte, dependerá del tipo de método de
quintilización utilizado
drop if GASTO==0 & FOLIO==" " //Se eliminan los códigos sin gasto
//Se quita el arriendo imputado
drop if (CCIF=="04.2.1.01.01" | CCIF=="04.2.2.01.01" | CCIF=="04.2.2.01.02")
merge m:m FOLIO aux using "BASE_QNSAI_HOGAR.dta" //SOLO DEBE EXISTIR _MERGE==3
drop _merge
gen double gasto_exp = GASTO*FW_QNSAI

```

```

replace gasto_exp=0 if gasto_exp==.
by FOLIO q_NSAI, sort: gen AUX=_n==1
gen double n_hogares=AUX* FW_QNSAI
bysort q_NSAI: egen double n_hogares_totales=sum(n_hogares) //Número total de
hogares encuestados por quintil
bysort q_NSAI: egen double gasto_total_hogares=sum(gasto_exp) //Gasto total de
todos los hogares por quintil
bysort q_NSAI: gen double G_Total_HOG=gasto_total_hogares/n_hogares_totales
collapse (sum) gasto_exp, by(CCIF GLOSA G_Total_HOG q_NSAI n_hogares_totales)
gen double Gasto_hog_prod= gasto_exp/ n_hogares_totales
gen double Porcentaje_hog_prod= (Gasto_hog_prod/ G_Total_HOG)*100
//Nota: Hasta aquí ya está la base a nivel de gasto promedio por hogar en cada
producto para cada quintil
//Para guardar la base, se pueden aplicar los siguientes comandos:
preserve
drop n_hogares_totales G_Total_HOG gasto_exp
drop if CCIF=="-77"
reshape wide Gasto_hog_prod Porcentaje_hog_prod, i( CCIF GLOSA) j(q_NSAI)
foreach x in Gasto_hog_prod1 Porcentaje_hog_prod1 Gasto_hog_prod2
Porcentaje_hog_prod2 Gasto_hog_prod3 Porcentaje_hog_prod3 Gasto_hog_prod4
Porcentaje_hog_prod4 Gasto_hog_prod5 Porcentaje_hog_prod5 {
replace `x'=0 if `x'==.
}
save "BASE_q_Productos.dta", replace //La base de datos guardada está a nivel
de productos por quintil
restore
*****A NIVEL DE DIVISIONES *****
gen D=substr(CCIF,1,2)
collapse(sum) Gasto_hog_prod, by(D q_NSAI G_Total_HOG)
gen double Porcentaje_hog_prod = (Gasto_hog_prod/ G_Total_HOG)*100
sort q_NSAI D
order q_NSAI D Gasto_hog_prod Porcentaje_hog_prod G_Total_HOG
gen G="0"
gen C="0"
gen SC="00"
gen P="00"
merge m:m D G C SC P using "CCIF_VIII_EPF.dta" //No debe existir merge==1
drop if _merge==2
order q_NSAI CCIF GLOSA Gasto_hog_prod Porcentaje_hog_prod
keep q_NSAI CCIF GLOSA Gasto_hog_prod Porcentaje_hog_prod
drop if CCIF==" "
reshape wide Gasto_hog_prod Porcentaje_hog_prod, i(CCIF GLOSA) j(q_NSAI)

```

Puede ser útil para los usuarios graficar la estructura del gasto por división obtenida en la Programación 65. Para ello, la siguiente programación ofrece un ejemplo de cómo generar un gráfico de la estructura del gasto por división según los distintos quintiles de ingreso disponible del hogar.

Programación 66: Estructura del gasto promedio mensual del hogar según quintil de ingreso disponible total del hogar

```
//Se debe primero ejecutar la programación 65
g D=substr(CCIF,1,2) //genera la variable división
//Graficar la estructura del gasto promedio mensual del hogar según quintil de
ingreso disponible total del hogar
graph bar (mean) Porcentaje_hog_prod*, over(D, label(labsize(vsmall)))
bar(1,color(gs2)) bar(2,color(gs5)) bar(3,color(gs8)) bar(4,color(gs11))
bar(5,color(gs14)) ///
yttitle("Porcentaje del gasto(%)", size(small)) scheme(s1color)
bltitle("(Divisiones del CCIF)") ///
ylabel(0 "0" 10 "10" 20 "20" 30 "30",labsize(small) ang(h) format(%10,0fc)) ///
title("Estructura del gasto según quintil de ingreso disponible del hogar",
size(medium)) subtitle("(excluye arriendo imputado)", size(small)) ///
legend(size(small) label(1 "Quintil hogares 1") label(2 "Quintil hogares 2")
label(3 "Quintil hogares 3") label(4 "Quintil hogares 4") label(5 "Quintil
hogares 5") symxsize(4))
graph export "Estructura_gasto_q.png", replace //Se exporta en formato png
```

7 PRECISIÓN DE LOS PRINCIPALES INDICADORES

Para una mejor comprensión de las medidas de precisión de las principales estimaciones o variables de interés de la VIII EPF, es necesario introducir el concepto de error estándar o desviación estándar. En particular, en este apartado se presentan las medidas de precisión asociadas al error estándar de la media del gasto e ingreso.

El coeficiente de variación es una medida de variabilidad relativa de las variables de interés (gasto e ingreso), que se calcula dividiendo el error estándar o desviación estándar de una estimación de interés (por ejemplo, el promedio del gasto) por la media del mismo.

El error absoluto corresponde a la máxima diferencia entre la estimación de la variable de interés y su valor real. No obstante, dada la imposibilidad de encuestar a la población completa, el valor “verdadero” no puede conocerse y se utiliza una aproximación de este, dada por el error estándar y un valor de la tabla t-Student para un nivel de confianza dado (usualmente, 95%) y que requiere el cálculo de los grados de libertad.

El error relativo corresponde al error absoluto dividido por la estimación de la variable de interés y se expresa en porcentaje.

El intervalo de confianza es un rango de valores determinados a partir de los valores observados en la muestra. Sus límites, inferior y superior, se determinan de forma tal que la probabilidad de que el parámetro poblacional (de la variable de interés) se ubique entre ellos sea de $(1 - \alpha)\%$. La probabilidad asociada al intervalo es también denominada como “nivel de confianza” del intervalo $(1 - \alpha)\%$. Usualmente se utiliza el 95% de nivel de confianza.

Para el cálculo de los indicadores, es necesario considerar un valor t-Student asociado al nivel de confianza escogido y los grados de libertad que permitan considerar los datos de la encuesta. Para las estimaciones principales de la encuesta (gasto promedio e ingreso disponible del hogar), no se identifican pseudoestratos o pseudoconglomerados de varianzas vacíos para las áreas de estimación³⁵. Dado lo anterior, las medidas de precisión se pueden obtener rápidamente sin hacer ajustes al cálculo de grados de libertad para el error absoluto, relativo y los límites del intervalo de confianza (inferior y superior).

³⁵ Dado que en la estructura del gasto se les imputa un valor cero a quienes no declaran realizar el gasto o tener el ingreso, siempre se trabaja con los grados de libertad máximos permitidos por la encuesta (248 a nivel nacional y 127 para Gran Santiago y 121 para el Resto de las capitales regionales).

Programación 67: Cálculo de coeficientes de variación para el gasto promedio del hogar

```

use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar Base de personas
set type double //importante para no tener diferencias en los cálculos
//Variables de interés Gasto total o Ingreso Disponible, del hogar o per cápita
con y sin arriendo imputado
//GASTOT_HD GASTOT_HD_AI GASTOT_HD_PC GASTOT_HD_PC_AI
//ING_DISP_HOG_HD ING_DISP_HOG_HD_AI ING_DISP_HOG_HD_PC ING_DISP_HOG_HD_PC_AI
svyset VARUNIT [pw=FE], strata(VARSTRAT) singleunit(certainty)
svy: mean GASTOT_HD if JHOGAR==1
estat cv
mat GP = (r(b)', r(se)', r(cv)')
foreach Z in 1 2 {
gen double stock_`Z'=1 if ZONA==`Z' //de esta forma se controla que los grados
de libertad sean correctos para cada área de estimación
svy: mean GASTOT_HD if JHOGAR==1, subpop(stock_`Z') //Aquí se puede reemplazar
la variable de interés
estat cv
mat GP`Z' = (r(b)', r(se)', r(cv)')
}
mat T = (GP\ GP1 \ GP2) //El orden coincide en cómo se crea posteriormente la
variable AREA
clear
svmat T
ren (T1 T2 T3) (Promedio_Gasto Error_Estandar CV) //El orden coincide en cómo
se guardaron las matrices GP y GPZ
gen AREA = "Total Capitales Regionales" if _n==1
replace AREA = "Gran Santiago" if _n==2
replace AREA = "Resto Capitales Regionales" if _n==3
order AREA Promedio Error_Estandar CV
list AREA Promedio Error_Estandar CV, noobs

```

En las siguientes secciones se revisa el cálculo de los cuatro indicadores de confiabilidad presentados, considerando el diseño complejo y la corrección de los grados de libertad. De la misma manera, se entrega una recomendación respecto a la interpretación de los datos haciendo referencia a los coeficientes de variación calculados y la cantidad de datos.

7.1 Grados de libertad

Es importante destacar que para el cálculo de grados de libertad se realiza un ajuste, de modo que estos sean calculados como la diferencia entre los conglomerados no vacíos y los estratos no vacíos, es decir, donde existe la variable de interés respecto a la cual se quieren calcular medidas de precisión (INE Chile, 2018b).

Los ajustes se realizan para las desagregaciones que dividen a los hogares de la encuesta según distintas características sociodemográficas. Estas desagregaciones podrían generar conglomerados que no tuviesen observaciones, por lo tanto, se hace necesario ajustar los grados de libertad para calcular las medidas de precisión de las estimaciones generadas.

Esto implica realizar cálculos individuales para cada una de las características o cruces que se realizan, ajustando los grados de libertad para cada uno de estos cruces o desagregaciones de la encuesta³⁶. En la siguiente programación se indica cómo efectuar estos cálculos en el programa.

Programación 68: Cálculo de las medidas de precisión del gasto considerando el ajuste a los grados de libertad (promedio y total)

```

qui mat drop _all //Se borran todas las matrices creadas hasta ahora, por lo
que los totales nacionales se calculan nuevamente
use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar Base de personas
set type double //importante para no tener diferencias en los cálculos
//Medidas para el gasto promedio
//Total de Capitales Regionales
svyset VARUNIT [pw=FE], strata(VARSTRAT) singleunit(certainty)
svy: mean GASTOT_HD ING_DISP_HOG_HD if JHOGAR==1 //Gasto promedio sin arriendo
imputado

mat datos = r(table) '
mat GP= datos[1..2,1]
mat EE= datos[1..2,2]
mat LI= datos[1..2,5]
mat LS= datos[1..2,6]

estat cv //Coeficiente de variación
qui mat CV = r(cv) '

//Por área de estimación (Gran Santiago y Resto de Capitales Regionales)
foreach Z in 1 2 {
gen double stock_`Z'=1 if ZONA==`Z' //de esta forma se controla que los grados
de libertad sean correctos para cada área de estimación
svy: mean GASTOT_HD ING_DISP_HOG_HD if JHOGAR==1, subpop(stock_`Z') //Gasto
promedio sin arriendo imputado para cada área de estimación
mat datos2 = r(table) '

mat GP`Z'= datos2[1..2,1]
mat EE`Z'= datos2[1..2,2]
mat LI`Z'= datos2[1..2,5]
mat LS`Z'= datos2[1..2,6]

estat cv //Coeficiente de variación
qui mat CV`Z' = r(cv) '
}
mat GP_1= (GP\GP1\GP2)
mat EE_1= (EE\EE1\EE2)
mat CV_1= (CV\CV1\CV2)
mat LI_1= (LI\LI1\LI2)
mat LS_1= (LS\LS1\LS2)

clear
mat TABLA= (GP_1, EE_1, CV_1, LI_1, LS_1)
svmat TABLA //Para la construcción del Tabulado

```

³⁶ Si no se realiza un ajuste de los grados de libertad, Stata considerará el total de conglomerados y estratos disponibles, sin restar aquellos en los cuales no hay información de cierta característica a estudiar.

```

ren (TABLA1 TABLA2 TABLA3 TABLA4 TABLA5) (Promedio Error_Estándar CV Lím_inf
  Lím_sup)
gen Etiqueta="Total Capitales Regionales" if _n==1 | _n==2
replace Etiqueta="Gran Santiago" if _n==3 | _n==5
replace Etiqueta="Resto Capitales Regionales" if _n==4 | _n==6

gen GRUPO=1 if _n==1 | _n==3 | _n==5 //GASTO
replace GRUPO=2 if _n==2 | _n==4 | _n==6 //INGRESO

preserve
keep if _n==1 | _n==3 | _n==5
gen Z=_n
two (bar Promedio Z, lwidth(thin) color(gs12)) (rcap Lím_inf Lím_sup Z,
lwidth(medthick) lcolor(navy)) if GRUPO==1, /*
*/ subtitle("Gasto Promedio", span size(medium)) scheme(s1color) xlabel( 1
"Total Capitales Regionales" 2 "Gran Santiago" 3 "Resto Capitales Regionales",
angle(45) lsize(small)) /*
*/ legend(order(1 "Promedio estimado" 2 "95% int. conf.)) xtitle("")
ytitle("Pesos", size(small)) /*
*/ ylabel( 0 "0" 200000 "200" 400000 "400" 600000 "600" 800000 "800"
1000000 "1.000" 1200000 "1.200" 1400000 "1.400" , lsize(small)
angle(h)) legend(size(vsmall)) /*
*/ yscale(range(0 1400000))
graph save uno, replace //Se guarda en formato Stata
restore
preserve
keep if _n==2 | _n==4 | _n==6
gen Z=_n
two (bar Promedio Z, lwidth(thin) color(gs7)) (rcap Lím_inf Lím_sup Z,
lwidth(medthick) lcolor(navy)) if GRUPO==2, /*
*/ subtitle("Ingreso Disponible", span size(medium)) scheme(s1color) xlabel( 1
"Total Capitales Regionales" 2 "Gran Santiago" 3 "Resto Capitales Regionales",
angle(45) lsize(small)) /*
*/ legend(order(1 "Promedio estimado" 2 "95% int. conf.)) xtitle("")
ytitle("Pesos", size(small)) /*
*/ ylabel(0 "0" 200000 "200" 400000 "400" 600000 "600" 800000 "800" 1000000
"1.000" 1200000 "1.200" 1400000 "1.400", lsize(small) angle(h))
legend(size(vsmall)) /*
*/ yscale(range(0 1400000))
graph save dos, replace //Se guarda en formato Stata
graph combine uno.gph dos.gph, title("Gasto Promedio e Ingreso Disponible" "del
hogar según área de estimación", span size(medium)) note("Fuente: VIII EPF-
INE", size(vsmall)) scheme(s1color)
//Se combinan los gráficos, se pone título y fuente
graph export "Gasto_Ingreso_xÁrea.png", replace //Se exporta en formato png
restore

```

La estimación de las medidas de precisión de las variables de ingreso (ingreso total del hogar sin arriendo imputado, ingreso total del hogar con arriendo imputado y sus respectivos promedios) se realiza de una manera similar al cálculo de las variables de gasto, utilizando las variables correspondientes: ING_TOTAL_HOG_HD e ING_TOTAL_HOG_HD_AI.

No obstante, la estimación del gasto a nivel de división no genera estratos o conglomerados de varianza vacíos, por ende, las medidas de precisión no variarán entre la Programación 67 y una ajustada, como se indica en la Programación 68 o en la Programación 69.

Programación 69: Cálculo de las medidas de precisión del gasto por división (promedio y total)

```

//Medidas de precisión del gasto por división
use "BASE_GASTOS_VIII_EPF.dta", clear //Se abre la Base de gastos
set type double //importante para no tener diferencias en los cálculos
replace GASTO=0 if (CCIF=="04.2.1.01.01" | CCIF=="04.2.2.01.01" |
CCIF=="04.2.2.01.02") //Se excluye el arriendo imputado
collapse (sum) GASTO, by(D FOLIO FE VARSTRAT VARUNIT) //Se colapsa la base a
nivel de división y hogar manteniendo la suma del gasto en esos niveles
destring D, replace //Se transforma la variable "División" de string a numérica
  foreach v in 1 2 3 4 5 6 7 8 9 10 11 12 {
    gen double GASTO_`v'=GASTO if D==`v'
    replace GASTO_`v'=0 if GASTO_`v'==. //Se reemplaza el "no gasto" por gasto 0
  }
collapse (sum) GASTO GASTO_1 GASTO_2 GASTO_3 GASTO_4 GASTO_5 GASTO_6 GASTO_7
GASTO_8 GASTO_9 GASTO_10 GASTO_11 GASTO_12, by(FOLIO FE VARSTRAT VARUNIT) //Se
colapsa la base a nivel del total de gasto por división
drop if FOLIO==" "
//Medidas de precisión para el gasto promedio por división
svyset VARUNIT [pw=FE], strata(VARSTRAT) singleunit(certainty)
svy: mean GASTO_1 GASTO_2 GASTO_3 GASTO_4 GASTO_5 GASTO_6 GASTO_7 GASTO_8
GASTO_9 GASTO_10 GASTO_11 GASTO_12
mat datos = r(table)'
gen D=_n
label define D 1 "01" 2 "02" 3 "03" 4 "04" 5 "05" 6 "06" 7 "07" 8 "08" 9 "09"
10 "10" 11 "11" 12 "12", modify
label values D D
mat GP= datos[1..12,1]
mat EE= datos[1..12,2]
mat LI= datos[1..12,5]
mat LS = datos[1..12,6]
estat cv //Coeficientes de variación
mat CV=r(cv)'
clear
mat TABLA= (GP, EE, CV, LI, LS)
svmat TABLA //Para la construcción del Tabulado
ren (TABLA1 TABLA2 TABLA3 TABLA4 TABLA5) (Promedio Error_Estándar CV Lím_inf
Lím_sup)
gen D=_n
label define D 1 "01" 2 "02" 3 "03" 4 "04" 5 "05" 6 "06" 7 "07" 8 "08" 9 "09"
10 "10" 11 "11" 12 "12", modify
label values D D

two (bar Promedio D, lwidth(thin) color(gs12)) (rcap Lím_inf Lím_sup D,
lwidth(medthick) lcolor(navy)), ///
title("Gasto Promedio del hogar", span size(large)) subtitle("Con intervalo de
confianza", span size(medium)) scheme(s1color) xlabel(1(1)12, valuelabel
angle(h) labsize(small)) ///
legend(order(1 "Promedio estimado" 2 "95% int. conf.)) xtitle("Divisiones
del CCIF", size(medsmall)) ytitle("Gasto por división (en miles)",
size(medsmall)) ///
ylabel( 0 "0" 20000 "20" 40000 "40" 60000 "60" 80000 "80" 100000 "100" 120000
"120" 140000 "140" 160000 "160" 180000 "180" 200000 "200" 220000 "220",
labsize(small) angle(h)) ///
yscale(range(0 220000)) note("Fuente: VIII EPF", size(small))
graph export "Gasto promedio por D_c_intervalo.png", replace

```

De la misma manera, se pueden calcular las medidas de precisión del ingreso desagregado a nivel de fuentes de ingreso: de la ocupación (sueldos y salarios, empleadores, trabajadores por cuenta propia, otros trabajos) y de otras fuentes (jubilaciones y pensiones de vejez, rentas de la

propiedad y transferencias netas). Para ello se debe ejecutar la programación detallada a continuación y utilizar las variables de ingreso creadas.

Programación 70: Cálculo de las medidas de precisión del ingreso

```

use "BASE_PERSONAS_VIII_EPF.dta", clear //Utilizar Base de personas
set type double //importante para no tener diferencias en los cálculos
foreach N in INGDA_HD INGDH_HD INGDNP_HD INGDPH_HD INGOTA INGDOTI INGP INGJ
INGJ_HD INGT_NETA AI_VP AI_VS {
quietly replace `N'=0 if `N'==.
}
quietly gen double
INGDTRAB_HD=INGDA_HD+INGDH_HD+INGDNP_HD+INGDPH_HD+INGDOTA+INGDOTI //Ingresos de
la ocupación
quietly gen double INGDTD_HD2=INGDA_HD+INGDH_HD //esta variable ya existe en la
BD
count if INGDTD_HD!=INGDTD_HD2 //debe ser cero
quietly gen double INGDTI_HD2=INGDNP_HD + INGDPH_HD //esta variable ya existe
en la BD
count if INGDTI_HD!=INGDTI_HD2 //debe ser cero
quietly gen double INGDEMP_HD2=INGDTI_HD if AECISE==1 //esta variable ya existe
en la BD
replace INGDEMP_HD2=0 if INGDEMP_HD2==.
count if INGDEMP_HD!=INGDEMP_HD2 //debe ser cero
quietly gen double INGDTCP_HD2=INGDTI_HD if AECISE==2 //esta variable ya existe
en la BD
replace INGDTCP_HD2=0 if INGDTCP_HD2==.
count if INGDTCP_HD!=INGDTCP_HD2 //debe ser cero
quietly gen double INGDTRABOS=INGDOTA+INGDOTI //Ingresos disponibles del
trabajo de la ocupación secundaria y de ocupación del mes anterior que no
coincide con la ocupación declarada como principal
quietly gen double RENTAS_PROPIEDAD=INGP+INGF
quietly gen double ARRIENDO_IMPUTADO=AI_VP+AI_VS if (AI_VP>=0 & AI_VS>=0)
quietly replace ARRIENDO_IMPUTADO=AI_VP if (AI_VP>=0 & AI_VS<0)
quietly replace ARRIENDO_IMPUTADO=AI_VS if (AI_VP<0 & AI_VS>=0)
quietly replace ARRIENDO_IMPUTADO=0 if PERSONA!=1 | ARRIENDO_IMPUTADO==. //se
debe una obserbación por persona para trabajar los ingresos a nivel hogar

//Se generan los agregados de los ingresos disponibles (segunda opción de
programación 63)
//(A=1+2+3+4)Ingreso disponible de la ocupación del hogar imputado por hot-deck
bys FOLIO: egen double
INGDTRAB_HOG_HD=sum(INGDA_HD+INGDH_HD+INGDNP_HD+INGDPH_HD+INGDOTA+INGDOTI)
//(1)Sueldos y Salarios disponible del hogar
bys FOLIO: egen double INGDTD_HOG_HD=sum(INGDTD_HD)
//(2)Ing disponible de Patrones y empleadores del hogar
bys FOLIO: egen double INGDEMP_HOG_HD=sum(INGDEMP_HD)
//(3)Ing disponible de Cuenta propia del hogar
bys FOLIO: egen double INGDTCP_HOG_HD=sum(INGDTCP_HD)
//(4)Ing disponible de otros trabajos del hogar
bys FOLIO: egen double INGDTRABOS_HOG_HD=sum(INGDOTA+INGDOTI)
//(B=4+5+6)Ingreso disponible de otras fuentes del hogar
bys FOLIO: egen double INGDOTF_HOG_HD=sum(INGDJ_HD+RENTAS_PROPIEDAD+INGT_NETA)
//(4)Jubilaciones o pensiones de vejez disponible del hogar
bys FOLIO: egen double INGDJ_HOG_HD=sum(INGDJ_HD)
//(5)Rentas de la propiedad del hogar
bys FOLIO: egen double RENTAS_PROPIEDAD_HOG_HD=sum(RENTAS_PROPIEDAD)
//(6)Transferencias netas del hogar (recibidas-emitidas)
bys FOLIO: egen double TRANSFERENCIAS_NETAS_HOG_HD=sum(INGT_NETA)
//(D=A+B)Ingreso total Disponible del hogar imputado por hot-deck (sin arriendo
imputado)
gen double ING_DISP_HOG_HD2=INGDTRAB_HOG_HD+INGDOTF_HOG_HD

```

```

label var ING_DISP_HOG_HD2 "Ingreso Disponible del hogar (A+B)"
label var INGDTTAB_HOG_HD "Ingreso disponible de la ocupación del hogar
imputado por hot-deck (A=1+2+3+4)"
label var INGDTD_HOG_HD "Ing por sueldos y salarios disponibles del hogar (1)"
label var INGDEMP_HOG_HD "Ing disponible de Patrones y empleadores del hogar
(2)"
label var INGDTCP_HOG_HD "Ing disponible de Cuenta propia del hogar (3)"
label var INGDTTABOS_HOG "Ing disponible de otros trabajos del hogar (4)"
label var INGDOTF_HOG_HD "Ingreso disponible de otras fuentes del hogar
(B=4+5a+5b+6)"
label var INGDIJ_HD_HOG "Jubilaciones o pensiones de vejez disponible del hogar
(4)"
label var RENTAS_PROPIEDAD_HOG "Ing rentas de la propiedad del hogar (5)"
label var TRANSFERENCIAS_NETAS_HOG "Transferencias netas del hogar (recibidas-
emitidas) (6)"
//Se establece el plan de muestreo
svyset [pweight=FE], strata(VARSTRAT) psu(VARUNIT) singleunit(certainty)
//Se renombran las variables para poder transponer fácilmente la base
ren (ING_DISP_HOG_HD2 INGDTTAB_HOG_HD INGDTD_HOG_HD INGDEMP_HOG_HD
INGDTCP_HOG_HD INGDTTABOS_HOG INGDOTF_HOG_HD INGDIJ_HD_HOG RENTAS_PROPIEDAD_HOG
TRANSFERENCIAS_NETAS_HOG) (ING_0 ING_1 ING_2 ING_3 ING_4 ING_5 ING_6 ING_7
ING_8 ING_9)
foreach r in 0 1 2 3 4 5 6 7 8 9 {
gen stocki_`r'=1 if JHOGAR==1
//se generan variables auxiliares para poder estimar la población por cada
fuente de ingresos
}
forvalues r = 0(1)9 { // esto permite guardar el total y por fuente de ingresos
svy: mean ING_`r', subpop(stocki_`r')
estat cv
mat est_`r' = r(b)' //Guarda el promedio
if est_`r'[1,1] == 0 {
mat se_`r' = 0 //Guarda el error estándar
mat cv_`r' = 0 //Guarda el coeficiente de variación
}
if est_`r'[1,1] != 0 {
mat se_`r' = r(se) //Guarda el error estándar
mat cv_`r' = r(cv) //Guarda el coeficiente de variación
}
}
//Cálculo de los grados de libertad a través del conteo de pseudoconglomerados
y pseudoestratos
by VARUNIT, sort: gen n_conglo = _n
sort VARUNIT, stable
forvalues r = 0(1)9 { //genera una matriz que cuenta los pseudoconglomerados y
pseudoestratos del ingreso disponible y para cada una de las fuentes
if `r' == 0 {
egen subp_conglo_T_`r' = max(stocki_`r'), by(VARUNIT)
gen conglo_T_`r' = 1 if(subp_conglo_T_`r')==n_conglo)
}
else if `r' > 0{
egen subp_conglo_T_`r' = max(stocki_`r') if (stocki_`r' == 1), by(VARUNIT)
gen conglo_T_`r' = 1 if(subp_conglo_T_`r')==n_conglo)
}
sum conglo_T_`r'
mat conglo_T_`r' = r(N)
}
by VARSTRAT, sort: gen n_VARSTRAT = _n
sort VARSTRAT, stable
forvalues r = 0(1)9 {
if `r' == 0 {
egen subp VARSTRAT T_`r' = max(stocki_`r'), by(VARSTRAT)

```

```

gen VARSTRAT_T_`r' = 1 if(subp_VARSTRAT_T_`r'==n_VARSTRAT)
}
else if `r' > 0 {
egen subp_VARSTRAT_T_`r' = max(stocki_`r') if (stocki_`r' == 1), by(VARSTRAT)
gen VARSTRAT_T_`r' = 1 if(subp_VARSTRAT_T_`r'==n_VARSTRAT)
}
sum VARSTRAT_T_`r'
mat VARSTRAT_T_`r' = r(N)
}
forvalues r = 0(1)9 {
if (conгло_T_`r'[1,1] - VARSTRAT_T_`r'[1,1] <= 2) {
mat df_T_`r' = 2
}
else {
mat df_T_`r' = conгло_T_`r' - VARSTRAT_T_`r'
}
//Cálculo del valor t-Student que corresponde al 95% y con los grados de
libertad calculados
mat crit_T_`r' = invt(df_T_`r'[1,1],0.975)
mat eabs_T_`r' = se_`r'[1,1]*crit_T_`r'
if est_`r'[1,1] <= eabs_T_`r'[1,1]{
mat ll_T_`r' = 0
}
else if est_`r'[1,1] > eabs_T_`r'[1,1]{
mat ll_T_`r' = est_`r'[1,1]-eabs_T_`r'
}
mat ul_T_`r' = est_`r'[1,1]+eabs_T_`r'
}
forvalues s = 1/9 { //Compilación de matrices total ingresos y por fuente
mat est_0 = est_0\est_`s'
mat se_0 = se_0\se_`s'
mat cv_0 = cv_0\cv_`s'
mat df_T_0 = df_T_0\df_T_`s'
mat crit_T_0 = crit_T_0\crit_T_`s'
mat eabs_T_0 = eabs_T_0\eabs_T_`s'
mat ll_T_0 = ll_T_0\ll_T_`s'
mat ul_T_0 = ul_T_0\ul_T_`s'
}
mat T= (est_0, se_0, cv_0, df_T_0, crit_T_0, eabs_T_0, ll_T_0, ul_T_0)
clear
svmat T //Para la construcción del Tabulado
ren (T1 T2 T3 T4 T5 T6 T7 T8) (Promedio Error_Estándar CV GradosLibertad ValorT
Error_Absoluto Lim_inf Lim_sup)
gen FUENTE="Ingreso Disp del hogar (A+B)" if _n==1
replace FUENTE="Ing disp de la ocupación(A=1+2+3+4)" if _n==2
replace FUENTE="Ing por sueldos y salarios disp(1)" if _n==3
replace FUENTE="Ing disp de Patrones y empleadores(2)" if _n==4
replace FUENTE="Ing disp de Cuenta propia(3)" if _n==5
replace FUENTE="Ing disp de otros trabajos(4)" if _n==6
replace FUENTE="Ing disp de otras fuentes (B=4+5a+5b+6)" if _n==7
replace FUENTE="Jubilaciones o pensiones de vejez disp(4)" if _n==8
replace FUENTE="Ing rentas de la propiedad(5)" if _n== 9
replace FUENTE="Transferencias netas(6)" if _n==10
drop GradosLibertad ValorT
format Promedio Error_Estándar Error_Absoluto Lim_inf Lim_sup %15,0fc
format CV %15,3fc
order FUENTE Promedio Error_Estándar Error_Absoluto CV Lim_inf Lim_sup //ordena
la presentación de las fuentes
gen FUENTE_n= n
labmask FUENTE_n, val(FUENTE) //este comando usa la variable string como
etiqueta de la variable numérica
two (bar Promedio FUENTE_n, lwidth(thin) color(gs12)) (rcap Lim_inf Lim_sup
FUENTE_n, lwidth(medthick) lcolor(navy) ) ///

```

```

if_n<=15, title("Ingreso Disponible del hogar", span size(large))
subtitle("Con intervalo de confianza", span size(medium)) ///
scheme(s1color) xlabel(1(1)10, value label angle(45) labsize(vsmall)) ///
legend(order(1 "Promedio estimado" 2 "95% int. conf.)) xtitle("Fuentes de
ingresos", size(vsmall)) legend(size(vsmall)) ///
yttitle("Promedio de ingresos (en miles)", size(vsmall)) legend(size(vsmall))
///
ylabel(0 "0" 500000 "500" 1000000 "1000" 1500000 "1500", labsize(small)
angle(horizontal)) ///
yscale(range(0 1500000)) ///
note("Fuente: VIII EPF- INE" "Nota: todas las fuentes de ingresos se encuentran
agregadas a nivel de hogar", size(vsmall))
graph export "Ingreso_estructura.png", replace

```

7.2 Interpretación

Dado que las bases de datos de la VIII EPF contienen diversas variables con las cuales se pueden realizar múltiples cruces de información y desagregaciones, es necesario incorporar consideraciones metodológicas respecto al uso de los datos y desagregaciones obtenidas.

Dados los tamaños muestrales de la VIII EPF (considerando la unidad de análisis hogares), el que una estimación sea o no aceptable dependerá del coeficiente de variación obtenido y de la desagregación del nivel de estimación, ya que al analizar un subgrupo de la muestra³⁷ se debe considerar la menor cantidad de observaciones disponibles. Por tanto, es relevante considerar la cantidad de unidades con las que se busca realizar alguna inferencia estadística.

Es importante, luego de realizar una estimación, evaluar su precisión. Para ello se recomienda analizar el coeficiente de variación (CV) como medida del error estándar del estimador. De forma complementaria, se deberá considerar el tamaño de la subpoblación analizada en el cruce de información, la que también afectará los grados de libertad ajustados, como se mostró en la sección anterior.

Para interpretar el error muestral, usualmente se selecciona el coeficiente de variación asociado a una estimación, para lo cual se recomienda considerar, a modo de referencia, los siguientes rangos:

- hasta un 5% se considera que la estimación es óptima en términos de precisión.
- sobre un 5% y hasta un 15% se considera una buena estimación en términos de precisión.
- sobre un 15% la estimación se considera deficiente en términos de precisión y, por ende, no es recomendable realizar análisis o inferencias, sino solo para fines descriptivos.

³⁷ Por ejemplo, niveles funcionales de la clasificación de productos o cruces de características de los integrantes del hogar.

Aun cuando el coeficiente de variación presente rangos adecuados en términos de precisión, las estimaciones podrían derivar en errores de interpretación por parte de los usuarios de las bases de datos. Esto, debido a la disparidad de los tamaños muestrales (número de hogares) que sustenta el cálculo de la estimación desagregada. Lo anterior podría implicar detectar estimaciones de gastos en productos con coeficientes de variación aceptables, pero calculado en base a un número reducido de hogares (observaciones), lo que no permite dirimir si el valor obtenido en este estimador es o no confiable.

En cuanto al número de observaciones se recomienda que las estimaciones evaluadas consideren al menos seis hogares con información, en el contexto de la VIII EPF. En efecto, las variables de interés con una cantidad muy baja de observaciones presentan en principio una alta dispersión en la estimación de las medidas de error, lo que dificulta su interpretación como indicador de calidad.

Por consiguiente, para definir el uso que se le puede dar a las estimaciones, se debe considerar un análisis conjunto del número de conglomerados con información de la variable de interés y del coeficiente de variación efectivo, al nivel de desagregación evaluado:

- **Estimador confiable:** el estimador se puede utilizar según los objetivos propuestos.
- **Uso descriptivo:** se recomienda que el estimador se utilice para estadística descriptiva. Si se desea utilizar para otro tipo de análisis, se recomienda agruparlo a un nivel superior de la clasificación de consumo. En todos estos casos se recomienda incluir una nota técnica advirtiendo sobre las restricciones de la estimación que mencione el valor del coeficiente de variación.
- **Estimador poco confiable:** no se recomienda utilizar la estimación del gasto obtenida. Se debe reevaluar la estimación de consumo a un nivel superior de la clasificación de consumo, para allí evaluar su calidad y determinar su aptitud para un uso inferencial o solamente descriptivo.

8 BIBLIOGRAFÍA

- Eurostat. (2003). *Household Budget Surveys in the EU, Methodology and recommendations for harmonization*. Luxembourg: Eurostat. Obtenido de http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-BF-03-003-__-N-EN.pdf
- Eurostat. (2014). *European Statistical System Handbook for Quality Reports*. Obtenido de <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf>
- INE Chile. (2018a). *Metodología VIII Encuesta de Presupuestos Familiares*. Santiago de Chile: Instituto Nacional de Estadísticas.
- INE Chile. (2018b). *Manual Conceptual y Metodológico Diseño Muestral Encuesta Nacional de Empleo (ENE)*. Instituto Nacional de Estadísticas, Santiago de Chile.
- INE Chile. (2018c). *Manual de trabajo de campo para encuestadores, TOMO 1 y 2*. Instituto Nacional de Estadísticas, Santiago de Chile.
- INE Chile. (2018d). *Principales Resultados VIII EPF*. Santiago de Chile: Instituto Nacional de Estadísticas.
- Observatorio Social. (2012). *Casen 2011: Metodología del Diseño Muestral y Factores de Expansión Encuesta de Caracterización Socioeconómica Nacional*. Ministerio de Desarrollo Social, Observatorio Social, Santiago de Chile. Obtenido de http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/Informe%20Diseno%20Muestral_Revision_13sep12.pdf
- OECD. (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. París: OECD. Obtenido de http://www.oecd-ilibrary.org/economics/framework-for-statistics-on-the-distribution-of-household-income-consumption-and-wealth_9789264194830-en
- Stata Corp. (22 de Marzo de 2018). <https://www.stata.com>. Obtenido de <https://www.stata.com>: <https://www.stata.com/manuals13/datatypes.pdf>
- UN. (2001). *Ingresos y gastos de consumo de los hogares en el marco del SCN y en encuestas de hogares*. Comisión Económica para América Latina y el Caribe, División de Estadística y Proyecciones Económicas. Santiago de Chile: CEPAL. Obtenido de https://repositorio.cepal.org/bitstream/handle/11362/4718/1/S01010054_es.pdf
- UN et al. (2016). *Sistema de Cuentas Nacionales 2008*. Nueva York: Naciones Unidas, Comisión Europea, OCDE, FMI, Banco Mundial. Obtenido de <https://unstats.un.org/unsd/nationalaccount/docs/SNA2008Spanish.pdf>

9 ANEXOS

Anexo N°1. Variables contenidas en la base de datos de personas de la VIII EPF

A continuación se presentan las variables contenidas en la Base de Datos de Personas de la VIII EPF. Para una mayor explicación respecto a la construcción de estas variables, se sugiere consultar los metadatos de la encuesta disponibles en el micrositio de la encuesta (www.ine.cl/epf).

- **ZONA:** identificador de la zona geográfica a la que pertenece el hogar, considerando Gran Santiago y Resto de capitales regionales.
- **FOLIO_V:** número identificador de la vivienda encuestada. Se encuentra compuesto de cinco dígitos que generan un valor único para cada vivienda.
- **FOLIO:** número identificador del hogar encuestado. Se encuentra compuesto de cinco dígitos que generan un valor único para la vivienda y un dígito posterior al guion, el cual indica el número del hogar dentro de la misma. El número de folio es único y corresponde a una variable clave (llave) que permite unir la información de cada hogar contenida en la "Base de gastos", con la información de la "Base de personas" de la encuesta.
- **FE:** el Factor de expansión corresponde a la cantidad de unidades en la población que representa una unidad de la muestra. Esta variable corresponde al ponderador de la encuesta.
- **VARSTRAT:** agrupación de estratos unitarios junto con otros estratos para el cálculo de la varianza. Un estrato se contruye a través del cruce de las variables comuna, clasificación socioeconómica y grupo de tamaño. Cada estrato de varianza debe estar conformado por un mínimo de 2 conglomerados (manzanas).
- **VARUNIT:** agrupación de conglomerados que no tengan el mínimo de unidades requeridas para la estimación de la varianza. Un conglomerado se identifica por medio de la variable manzana. Cada conglomerado de varianza debe contener al menos 30 viviendas.
- **PERSONA:** número correlativo que identifica a los miembros del hogar, definiéndose como quién reside habitualmente en el hogar y comparte un presupuesto común para alimentación.
- **N_LINEA:** número de línea de la persona según el Registro de Personas en el Hogar (RPH).
- **NPERSONAS:** corresponde al número total de personas que compone cada hogar encuestado y se construye a partir de la variable PERSONA o N_LINEA. Se define a los miembros del hogar como aquellas personas que cumplen con los criterios de residencia habitual en la vivienda y comparten presupuestos en alimentación. En la encuesta se excluyen como miembros del hogar al personal de servicio doméstico, parientes de este y los pensionistas.
- **JHOGAR:** identifica a la persona mayor de 15 años a la que todos los miembros del hogar reconocen como jefe/a de hogar. No se definen criterios objetivos en su elección, sin embargo, generalmente coincide con la persona de mayor edad, el propietario de la vivienda o la persona que toma las decisiones importantes.

- **PARENTESCO:** identifica para cada uno de los miembros del hogar el tipo de parentesco respecto al jefe/a de hogar. El parentesco puede ser por consanguinidad, adopción y/o afinidad.
- **SEXO:** señala el sexo de cada uno de los miembros del hogar, hombre o mujer.
- **EDAD:** señala para cada uno de los miembros del hogar la edad en años efectivamente cumplidos al momento de aplicar la encuesta.
- **ECIVIL:** identifica el estado civil o conyugal de los miembros del hogar. Los menores de 15 años se consideran solteros a menos que informen lo contrario en la entrevista de aplicación.
- **SPRINCIPAL:** identifica al sustentador principal que corresponde al miembro del hogar que más aporta al presupuesto familiar, evaluando los últimos seis meses antes de la aplicación de la encuesta. El sustentador principal no corresponde, necesariamente, al miembro que recibe más ingresos o al designado como jefe/a de hogar.
- **ECOMPRAS:** identifica al administrador de gastos del hogar, que corresponde a la persona que habitualmente realiza las compras para el hogar, evaluando los últimos seis meses antes de la aplicación de la encuesta. Es posible señalar a más de una persona por hogar.
- **EDUE:** identifica, para cada uno de los miembros del hogar, el número de años de escolaridad formal (se excluye del cálculo los cursos de párvulo). Los años de escolaridad varían dependiendo de los niveles educacionales alcanzados; se espera que una persona con estudios de enseñanza media completa alcance un total de 12 años de escolaridad.
- **EDUACTUAL:** identifica, para cada miembro del hogar, la situación de estudios formales o complementarios considerando el período académico al momento de la aplicar la encuesta.
- **EDUCURSO:** identifica, para cada uno de los miembros del hogar, el último curso aprobado dentro del nivel de educación más alto alcanzado.
- **EDUNIVEL:** identifica el nivel educacional más alto alcanzado para cada miembro del hogar. Se agrupan las personas que terminaron el nivel, así como las personas que no lo terminaron.
- **EDUTERMINA:** identifica, para cada miembro del hogar, el cumplimiento de todos los requisitos para obtener la certificación del nivel educacional más alto alcanzado.
- **EDUDEPENDENCIA:** identifica, para cada miembro del hogar que declara estudiar en el momento de aplicación de la encuesta, el tipo de dependencia administrativa del establecimiento educativo.
- **EF01:** pregunta que permite diferenciar entre aquellas personas que actualmente se encuentra estudiando en educación parvularia, básica o media (EF01=1) y los estudiantes de educación superior (EF01=2).
- **MATRICULA:** corresponde a una variable que identifica el monto pagado por aquellas personas que se encontraban estudiando al momento de aplicar la encuesta, por el derecho a asistir a un establecimiento educacional. La variable se construye a partir de la información recogida por el cuestionario de gastos del hogar, específicamente las preguntas EF02, EF03, EF04 y EF05. Este pago se encuentra mensualizado y se expresa en pesos chilenos.

- **ARANCEL:** corresponde a una variable que identifica el monto efectivamente pagado por aquellas personas que se encontraban estudiando al momento de aplicar la encuesta, por concepto de colegiatura o arancel educativo. La variable se construye a partir de la información recogida por el cuestionario de gastos del hogar, específicamente las preguntas EF6, EF07, EF08, EF09, EF10, EF11, EF12, EF13, EF14, EF15, EF16, EF17, EF18 y EF19. Este pago se encuentra mensualizado y se expresa en pesos chilenos.
- **EF10:** pregunta dicotómica que identifica a los estudiantes de educación parvularia, básica, media y superior que poseen financiamiento compartido o alguna beca que descuenta parte del arancel, colegiatura, mensualidad. También incluye a los estudiantes que se encuentran adscritos a gratuidad en la educación superior.
- **EF14:** pregunta dicotómica que identifica a los estudiantes que poseen algún crédito para costear sus estudios de educación superior.
- **ARANCEL_SUB:** corresponde a una variable dicotómica que permite identificar una subdeclaración en el monto efectivamente pagado por concepto de colegiatura o arancel educativo, para aquellas personas que se encontraban estudiando al momento de aplicar la encuesta. Esto se puede explicar, ya sea por una ausencia en la declaración del crédito educativo, una inconsistencia en el monto del crédito declarado o bien debido a una subdeclaración en el monto efectivamente pagado por el estudiante.
- **CCPP:** corresponde a una variable que identifica el monto pagado por concepto de Centro de padres y apoderados del establecimiento educativo. Para aquellas personas que se encontraban estudiando en educación preescolar, básica y media, al momento de aplicar la encuesta. Este pago se encuentra mensualizado y se expresa en pesos chilenos.
- **CUOTA_CURSO:** corresponde a una variable que identifica el monto pagado por concepto de cuotas de curso, para aquellas personas que se encontraban estudiando en educación preescolar, básica y media. Este pago se encuentra mensualizado y se expresa en pesos chilenos.
- **ADMISION:** corresponde a una variable que identifica el monto pagado por aquellas personas que se encontraban estudiando, por concepto de cuotas de incorporación y/o examen de admisión. Este pago se encuentra mensualizado y se expresa en pesos chilenos.
- **OTROS_EDU:** corresponde a una variable que identifica el monto pagado por aquellas personas que se encontraban estudiando en educación, por concepto de otros derechos como: certificados varios (notas, títulos, etc.), examen de grado, entre otros. Este pago se encuentra mensualizado y se expresa en pesos chilenos.
- **PETG:** variable que identifica a las personas que cumplen con los requisitos legales de edad para optar a un trabajo remunerado, que se extiende a las personas de 15 años o más.
- **CAE:** variable que clasifica a todos los miembros del hogar según su condición de actividad económica en la semana de referencia, considerando el momento de aplicación de la encuesta. La clasificación subdivide a la población en los siguientes grupos: menores de 15 años, ocupados, desocupados o inactivos.

- **CAEG:** variable que identifica a las personas en edad de trabajar que se encuentran activas en la semana de referencia al momento de aplicar la encuesta (cumplen los requisitos para ser incluidos como ocupados o desocupados).
- **AE01:** pregunta que identifica a las personas de 15 años o más, que durante la semana anterior a la aplicación de la encuesta realizaron algún trabajo, de al menos una hora, y por el cual recibió o recibirá pago en dinero o en especies. Las personas que responden de forma afirmativa son consideradas parte de la Población Económicamente Activa (PEA).
- **AE02:** pregunta que identifica aquellas personas que son parte de la Población Económicamente Activa (PEA), pero no fueron capturadas en la pregunta AE01.
- **AE03:** pregunta que identifica a las personas que son parte de la Población Económicamente Activa (PEA), pero no fueron capturadas en la pregunta AE01 y AE02, por encontrarse temporalmente ausentes.
- **AE04:** pregunta que se aplica a las personas cuya respuesta ha sido negativa en AE01, AE02 y AE03. Identifica a las personas que en las últimas cuatro semanas han estado en búsqueda de trabajo o intentando establecer su propio negocio. Las personas que responden de manera negativa son consideradas inactivas y aquellas que responden de manera afirmativa pasan a responder AE05.
- **AE05:** pregunta que se aplica a las personas cuya respuesta ha sido negativa en AE01, AE02 y AE03 y, además, han respondido afirmativamente AE04. Esta variable identifica a las personas que estarían disponibles para trabajar en los próximos 15 días. Las personas que responden de forma afirmativa son consideradas parte de la Población Económicamente Activa (PEA). Por el contrario, las personas que responden de forma negativa son consideradas Inactivas.
- **AEPR1:** la variable período de referencia laboral identifica si los ingresos del mes de referencia (mes anterior) corresponden a la actividad principal declarada en el RPH al momento de aplicar la encuesta (actividad realizada la semana anterior). Asigna 1 cuando la ocupación corresponde al mismo empleo del mes anterior y 2 cuando no corresponde al mismo empleo del mes anterior.
- **AE06A:** pregunta que identifica el mes de inicio en el empleo, actividad o negocio, declarado previamente por las personas de 15 años o más.
- **AE06B:** pregunta que identifica el año de inicio en el empleo, actividad o negocio, declarado previamente por las personas de 15 años o más.
- **AECISE:** adaptación de la Clasificación Internacional de la Situación en el Empleo (CISE-93), la que clasifica a la población económicamente activa de 15 años o más, según la existencia de subordinación o autonomía en la relación laboral, incluyendo también desagregaciones más específicas para trabajadores asalariados y para trabajadores del servicio doméstico.
- **CAT_IMP:** corresponde a una variable de clasificación de la ocupación principal construida para el proceso de imputación de ingresos del trabajo en la encuesta, la cual permite clasificar a la población económicamente activa de 15 años o más bajo las categorías de asalariados, honorarios, negocios por cuenta propia y profesionales independientes. La variable permite clasificar únicamente aquellas personas que durante el período de

referencia registraron ingresos por la actividad principal. En la VII versión de la encuesta esta variable se denominó CISE_4.

- **CISEG**: corresponde a una variable resumen construida para la encuesta, la que permite clasificar a la población activa según el tipo de ingresos que reciben: ingresos derivados del trabajo dependiente (condición de subordinación en la relación laboral) o ingresos derivados del trabajo independiente (existencia de autonomía en la relación laboral).
- **AE08**: pregunta que identifica la forma en que reciben sus ingresos aquellas personas de 15 años o más, que declaran al momento de ser encuestados en la semana de referencia una ocupación como asalariados o personal del servicio doméstico.
- **AECIUO88**: la Clasificación Internacional de Ocupaciones (CIUO-88) es un sistema clasificador que agrupa las ocupaciones tomando como base la similitud de las competencias requeridas para cumplir las tareas y funciones del empleo. Cada empleo es clasificado a partir del nivel de competencias (en función de la variedad y complejidad de las tareas) y la especialización de las competencias (el tipo de conocimientos aplicados, la naturaleza de los bienes y servicios producidos, etc.). La estructura original propuesta por la OIT se compone de cuatro niveles de desagregación, subdivididos en 10 grandes grupos, 28 subgrupos principales, 116 subgrupos y 390 grupos primarios. La variable se encuentra agregada en 9 grandes grupos y un grupo residual.
- **AECIUO88_2**: la Clasificación Internacional de Ocupaciones (CIUO-88) a dos dígitos corresponde a la variable ocupación desagregada a nivel de 28 subgrupos principales.
- **AECIUO08**: la Clasificación Internacional de Ocupaciones adaptada a Chile (CIUO-08 CL) es un sistema clasificador que agrupa las ocupaciones tomando como base la similitud de las competencias requeridas para cumplir las tareas y funciones del empleo. Cada empleo es clasificado a partir del nivel de competencias (en función de la variedad y complejidad de las tareas) y la especialización de las competencias (el tipo de conocimientos aplicados, la naturaleza de los bienes y servicios producidos, etc.). La estructura original propuesta por la OIT se compone de cuatro niveles de desagregación, de los cual se utilizan solo 2 niveles, subdivididos en 10 grandes grupos y 42 subgrupos principales. En la encuesta la variable se encuentra agregada en 9 grandes grupos y un grupo residual.
- **AECIUO08_2**: la Clasificación Internacional de Ocupaciones adaptada a Chile (CIUO-08 CL) a dos dígitos corresponde a la variable ocupación desagregada a nivel de 42 subgrupos principales.
- **AE10A**: corresponde a una variable dicotómica que permite identificar la fuente de ingresos por concepto de trabajo asalariado, para las personas de 15 años o más.
- **AE10B**: corresponde a una variable dicotómica que permite identificar la fuente de ingresos por concepto de trabajo independiente, para las personas de 15 años o más.
- **AE10C**: corresponde a una variable dicotómica que permite identificar la fuente de ingresos por concepto de jubilaciones o pensión de vejez, para las personas de 15 años o más.
- **AE10D**: corresponde a una variable dicotómica que permite identificar otro tipo de ingresos percibidos por las personas de 15 años o más.

- **SP01:** variable que determina, para cada uno de los miembros del hogar de 18 años o más, la pertenencia a un sistema de salud (Fonasa, Isapre u otro), permitiendo conocer la condición de afiliación de las personas, independiente de si tienen contrato laboral.
- **SP02:** el tipo de sistema de salud corresponde a la aseguradora a la cual se declara afiliado cada uno de los miembros del hogar de 18 años o más. La encuesta identifica los tres sistemas de salud que existen en nuestro país y que están definidos por ley: sistema público (Fonasa), sistema privado (Isapre), el sistema de salud de las FFAA y de Orden y el sistema Prais que es gratuito. Además, incluye la categoría otro sistema de salud, la que corresponde a seguros de salud extranjeros, servicios médicos cubiertos por otras instituciones, entre otros.
- **SP03:** variable que determina, para cada uno de los miembros del hogar de 18 años o más, la pertenencia a un sistema de previsión (AFP, Capredena, Dipreca, etc.). Específicamente, es posible conocer la condición de afiliación de cada persona.
- **SP04:** variable que identifica el tipo de sistema al cual se declara afiliado cada uno de los miembros del hogar. La encuesta especifica tres sistemas previsionales: sistema de Administradoras de Fondo de Pensiones (AFP), que constituye la forma de actual de acumulación de recursos para la vejez; sistema público o antiguo sistema a cargo del Instituto de Previsión Social (IPS, ex INP), y sistema de las FFAA (Capredena o Dipreca). Incluye, además, la categoría otro sistema previsional para los depósitos realizados en otras cajas de previsión.
- **SP04A:** variable que especifica el nombre de la AFP a la cual se encuentran afiliados los miembros del hogar de 18 años o más.
- **VP:** identifica mediante la observación el tipo de vivienda que habitan los miembros del hogar.
- **TVP:** identifica el tipo de tenencia de la vivienda principal que habitan los miembros del hogar. El tipo de tenencia puede estar constituida por una relación jurídica o de palabra, lo que permite a las personas ocupar una vivienda y disfrutar de los atributos que entrega. Se distinguen nueve formas de tenencia de la vivienda.
- **VP02:** corresponde a una variable dicotómica que identifica, mediante la observación, la existencia de un negocio adosado a la vivienda que habitan los miembros del hogar.
- **VP05A:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en arriendo por el uso de la vivienda que habitan los miembros del hogar.
- **VP05B:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en dividendo por el uso de la vivienda que habitan los miembros del hogar.
- **VP05C:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en agua por el uso de la vivienda que habitan los miembros del hogar.
- **VP05D:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en electricidad (luz) por el uso de la vivienda que habitan los miembros del hogar.
- **VP05E:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en gas de cañería por el uso de la vivienda que habitan los miembros del hogar.

- **VP05F:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en servicios de seguridad (monitoreo, alarma, etc.) por el uso de la vivienda que habitan los miembros del hogar.
- **VP05G:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en servicio doméstico por el uso de la vivienda que habitan los miembros del hogar.
- **VP05H:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en gastos comunes por el uso de la vivienda que habitan los miembros del hogar.
- **VP05I:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en servicios de telecomunicaciones por el uso de la vivienda que habitan los miembros del hogar. Por ejemplo: telefonía fija, internet, televisión pagada, *pack* de telecomunicaciones, etc.
- **VP05J:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en extracción de basura por el uso de la vivienda que habitan los miembros del hogar.
- **VS:** corresponde a una variable dicotómica que indica la posesión de una vivienda secundaria considerando los últimos doce meses antes de la aplicación de la encuesta. En el caso de la encuesta, se consideran viviendas secundarias aquellas viviendas propias que estuvieron a disposición el mes anterior y al menos tres meses dentro del último año, cuyo uso está destinado al descanso, recreación y/o vacaciones.
- **VS01:** corresponde a una variable dicotómica que indica la posesión de una vivienda secundaria adicional a la vivienda principal que habitan los miembros del hogar.
- **VS02:** identifica el tipo de tenencia de la vivienda secundaria. Se distinguen nueve formas de tenencia de la vivienda.
- **VS03:** indica el uso principal de la vivienda secundaria declarada por el hogar. Distingue si la vivienda está destinada al descanso, arriendo u otro uso.
- **VS04:** indica la cantidad de meses en que el hogar ha tenido a su disposición la vivienda secundaria.
- **VS05:** indica la comuna donde se ubica la vivienda secundaria declarada por el hogar.
- **VS06A:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en agua por el uso de la vivienda secundaria.
- **VS06B:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en electricidad (luz) por el uso de la vivienda secundaria.
- **VS06C:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en gastos comunes por el uso de la vivienda secundaria.
- **VS06D:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en servicios de telecomunicaciones por el uso de la vivienda secundaria.
- **VS06E:** corresponde a una variable dicotómica que permite identificar la existencia de gastos en extracción de basura por el uso de la vivienda secundaria.
- **GASTOT_HD:** es la suma de gastos realizados por el hogar, sin considerar el arriendo imputado. Estos gastos corresponden al valor por concepto de determinados bienes y servicios adquiridos por cada hogar encuestado, el que se construye a partir de la

información obtenida en los diferentes cuestionarios de gastos y del cuestionario autocompletado por los miembros del hogar. La variable se genera mediante la suma del gasto a nivel de producto; luego, son mensualizados e imputados por *hot-deck*.

- **GASTOT_HD_AI**: corresponde al gasto total por hogar imputado por *hot-deck*, incluyendo el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella, así como el arriendo imputado estimado por la vivienda secundaria (para los hogares que poseen vivienda secundaria). Esta variable se puede construir utilizando la variable de gasto total GASTOT_HD, además de las variables de arriendo imputado AI_VP y AI_VS.
- **GASTOT_HD_PC**: corresponde al gasto total por hogar imputado por *hot-deck*, sin considerar el arriendo imputado y subdividido por el número de miembros del hogar.
- **GASTOT_HD_PC_AI**: es el gasto total por hogar imputado por *hot-deck*, incluyendo el arriendo imputado y subdividido por el número de miembros del hogar.
- **ING_TOTAL_HOG_HD**: corresponde a la sumatoria de todas las fuentes habituales de ingreso del hogar que fueron percibidos el mes anterior a la aplicación de la encuesta. Incluye los ingresos por la actividad principal, otras actividades laborales, rentas de la propiedad, ingresos por instrumentos financieros, jubilaciones, transferencias en efectivo y en especie recibidas desde otros hogares, así como las transferencias en efectivo recibidas desde el Gobierno y/o instituciones sin fines de lucro al servicio de los hogares (ISFLSH).
- **ING_TOTAL_HOG_HD_AI**: corresponde a la sumatoria de todas las fuentes habituales de ingreso del hogar percibidos en el mes anterior a la aplicación de la encuesta. Incluye el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella, así como el arriendo imputado estimado por la vivienda secundaria (para los hogares que poseen vivienda secundaria).
- **ING_TOTAL_HOG_HD_PC**: corresponde al ingreso total del hogar sin considerar el arriendo imputado y subdividido por el número de miembros del hogar.
- **ING_TOTAL_HOG_HD_PC_AI**: corresponde al ingreso total del hogar incluyendo el arriendo imputado y subdividido por el número de miembros del hogar.
- **ING_DISP_HOG_HD**: corresponde a la sumatoria de las fuentes habituales de ingreso del hogar percibidos en el mes anterior a la aplicación de la encuesta. Se calcula como la suma entre el ingreso primario compuesto por remuneración de los asalariados, independientes, ingresos de la propiedad, ingresos financieros y autosuministro de bienes producidos por el hogar, más las transferencias corrientes recibidas en efectivo. De la sumatoria se deben descontar las transferencias corrientes pagadas, tales como: transferencias emitidas y donaciones realizadas a otros hogares o instituciones sin fines de lucro al servicio de los hogares (ISFLSH); pago de impuestos, contribuciones, tasas y multas, y pago de contribuciones sociales como salud y previsión, entre otros.
- **ING_DISP_HOG_HD_AI**: corresponde al ingreso disponible del hogar incluido el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella, así como el arriendo imputado estimado por la vivienda secundaria (para los hogares que poseen vivienda secundaria).

- **ING_DISP_HOG_HD_PC:** corresponde al ingreso disponible del hogar sin considerar el arriendo imputado y subdividido por el número de miembros del hogar.
- **ING_DISP_HOG_HD_PC_AI:** corresponde al ingreso disponible del hogar incluyendo el arriendo imputado y subdividido por el número de miembros del hogar.
- **AI_VP:** corresponde al valor estimado por el hogar respecto de lo que debería pagar si tomara la vivienda principal en arriendo.
- **AI_VS:** corresponde al valor estimado por el hogar respecto de lo que debería pagar si tomara la vivienda secundaria en arriendo.
- **ING_PRODUC_HOG_HD:** corresponde a los ingresos del trabajo observado y los ingresos imputados por *hot-deck* derivados de una actividad asalariada y/o independiente, así como, una actividad principal y/o secundaria declarada por los miembros del hogar. También contempla aquellos ingresos derivados de la auto-provisión de bienes producidos por el hogar, el cual se encuentra recogido en la encuesta como retiro de productos de la empresa o negocio por cuenta propia.
- **ING_PRODUC_HOG_HD_AI:** corresponde a los ingresos de la producción, incluyendo los ingresos imputados por *hot-deck* y el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella, así como el arriendo imputado estimado por la vivienda secundaria (para los hogares que poseen vivienda secundaria).
- **ING_PRIMARIO_HOG_HD:** corresponden a la suma de los ingresos derivados de la producción, los ingresos de la propiedad y los ingresos por instrumentos financieros. Además, incluye los ingresos imputados por *hot-deck*.
- **ING_PRIMARIO_HOG_HD_AI:** corresponde al ingreso primario, incluyendo los ingresos imputados por *hot-deck* y el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella, así como el arriendo imputado estimado por la vivienda secundaria (para los hogares que poseen vivienda secundaria).
- **INGTD:** corresponde al total haberes o al ingreso total derivado de una actividad asalariada dependiente. La clasificación de actividades asalariadas dependientes deriva de la adaptación realizada de la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGTD_HD:** corresponde al ingreso bruto derivado del trabajo dependiente observado, más los ingresos imputados por *hot-deck*. La clasificación de actividades asalariadas dependientes deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDTD:** corresponde al ingreso de los trabajadores dependientes observado luego de los descuentos legales (previsión, salud e impuesto a las remuneraciones). La clasificación de actividades asalariadas dependientes deriva de la adaptación realizada de la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDTD_HD:** corresponde al ingreso de los trabajadores dependientes luego de los descuentos legales (previsión, salud e impuesto a las remuneraciones), incluyendo los ingresos imputados por *hot-deck*. La clasificación de actividades asalariadas dependientes deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).

- **INGTI**: corresponde los ingresos generados o ganados por la realización de la actividad o negocio por cuenta propia, menos los costos asociados a su realización. Los costos mencionados anteriormente consideran impuestos, salarios que deben ser pagados a empleados u otros y bienes intermedios que deben ser comprados para la producción de la actividad.
- **INGTI_HD**: corresponde al ingreso neto del trabajo independiente observado, más los ingresos imputados por *hot-deck*.
- **INGDTI**: corresponde al ingreso neto generado por los trabajadores independientes observado, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGDTI_HD**: corresponde al ingreso neto generado por los trabajadores independientes descontando las cotizaciones voluntarias en previsión y/o salud. Incluye los ingresos imputados por *hot-deck*.
- **INGA**: Corresponde al total de haberes o ingreso bruto percibido por trabajadores asalariados que no entregan boleta de honorarios. Estos ingresos incluyen los descuentos legales (previsión, salud e impuesto de segunda categoría) y los descuentos no legales. Los ingresos relacionados con el empleo asalariado se componen de todos los pagos, prestaciones en efectivo, especie o servicios devengados en el mes anterior a la aplicación de la encuesta.
- **INGA_HD**: corresponde a los ingresos brutos percibido por trabajadores asalariados que no entregan boleta a honorarios, más los ingresos imputados por *hot-deck*.
- **INGDA**: corresponde al ingreso de los trabajadores asalariados que no entregan boleta de honorarios luego de los descuentos legales (previsión, salud e impuesto a las remuneraciones). Los ingresos relacionados con el empleo asalariado se componen de todos los pagos, prestaciones en moneda, especie o servicios percibidos en el mes anterior a la aplicación de la encuesta.
- **INGDA_HD**: corresponde a los ingresos disponibles percibidos por trabajadores asalariados que no entregan boleta a honorarios, más los ingresos imputados por *hot-deck*.
- **INGH**: corresponde al total haberes o el ingreso bruto por trabajo asalariado percibido por trabajadores asalariados en el mes anterior a la aplicación de la encuesta, correspondiente a personas que declaran explícitamente entregar boleta honorarios. Este ingreso considera el salario a recibir por una persona sin descontar la retención del impuesto (10%), así como tampoco las cotizaciones voluntarias de previsión y/o de salud.
- **INGH_HD**: corresponde a los ingresos brutos percibidos por trabajadores asalariados que declaran entregar boleta de honorarios, más los ingresos imputados por *hot-deck*.
- **INGDH**: corresponde al ingreso percibido por trabajadores asalariados que declaran explícitamente entregar boleta honorarios en el mes anterior a la aplicación de la encuesta. Los ingresos se componen de todos los pagos en efectivo, especies o servicios, menos el descuento por retención del impuesto (10%), así como las cotizaciones voluntarias de previsión y/o de salud.
- **INGDH_HD**: corresponde al ingreso disponible percibido por trabajadores asalariados que declaran entregar boleta de honorarios, más los ingresos imputados por *hot-deck*.

- **INGEMP**: corresponde al ingreso neto percibido por las personas clasificadas como patrones o empleadores según la variable AECISE. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGEMP_HD**: corresponde al ingreso neto percibido por las personas clasificadas como patrones o empleadores en la variable AECISE y los ingresos imputados por *hot-deck*. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDEMP**: corresponde al ingreso neto percibido por las personas clasificadas como patrones o empleadores en la variable AECISE, descontando las cotizaciones voluntarias en previsión y/o salud. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDEMP_HD**: corresponde al ingreso neto percibido por las personas clasificadas como patrones o empleadores en la variable AECISE y los ingresos imputados por *hot-deck*, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGTCP**: corresponde al ingreso neto percibido por las personas clasificadas como trabajador por cuenta propia según la variable AECISE. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGTCP_HD**: corresponde al ingreso neto percibido por las personas clasificadas como trabajador por cuenta propia en la variable AECISE y los ingresos imputados por *hot-deck*. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDTCP**: corresponde al ingreso neto percibido por las personas clasificadas como trabajador por cuenta propia en la variable AECISE, descontando las cotizaciones voluntarias en previsión y/o salud. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGDTCP_HD**: corresponde al ingreso neto percibido por las personas clasificadas como trabajador por cuenta propia en la variable AECISE y los ingresos imputados por *hot-deck*, descontando las cotizaciones voluntarias en previsión y/o salud. Esta clasificación de actividades deriva de la adaptación realizada a la Clasificación Internacional de Situación de Empleo (CISE 93).
- **INGNP**: corresponde al ingreso generado por las personas que realizan una actividad o negocio por cuenta propia, descontando los costos asociados a su realización (impuestos, salarios que deban ser pagados a empleados u otros, bienes intermedios que deban ser comprados para la producción de la actividad). También se descuentan los retiros de productos del negocio o empresa para el consumo personal o del hogar, considerando los precios de mercado.
- **INGNP_HD** : el ingreso neto de las personas que realizan una actividad o negocio por cuenta propia, más los ingresos imputados por *hot-deck*.

- **INGDNP**: corresponde al ingreso neto generado por las personas que realizan una actividad o negocio por cuenta propia, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGDNP_HD**: el ingreso disponible de las personas que poseen negocios por cuenta propia y los ingresos imputados por *hot-deck*, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGPI**: corresponde al ingreso generado por las personas que realizan una actividad como profesionales independientes, descontando los impuestos legales.
- **INGPI_HD**: corresponde al ingreso neto de las personas que se declaran profesionales independientes y los ingresos imputados por *hot-deck*, descontando los impuestos legales.
- **INGDPI**: corresponde al ingreso neto generado por las personas que realizan una actividad como profesionales independientes, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGDPI_HD**: el ingreso disponible de las personas que se declaran profesionales independientes y los ingresos imputados por *hot-deck*, descontando las cotizaciones voluntarias en previsión y/o salud.
- **INGJ**: corresponde a la suma de los ingresos por jubilaciones y/o pensiones de vejez pagados por el Instituto de Previsión Social (IPS), Administradora de Fondos de Pensiones (AFP), así como de los sistemas de las FFAA.
- **INGJ_HD**: los ingresos brutos por jubilaciones y/o pensiones de vejez, más los ingresos imputados por *hot-deck*.
- **INGDJ**: suma de los ingresos brutos por jubilaciones y/o pensiones de vejez, descontando las cotizaciones legales en salud. Estos ingresos son pagados por el Instituto de Previsión Social (IPS), Administradora de Fondos de Pensiones (AFP) y el sistema de las FFAA.
- **INGDJ_HD**: ingresos disponibles por jubilaciones observado más los ingresos imputados por *hot-deck*.
- **INGOTA**: ingreso que proviene de actividades secundarias y actividades principales que han finalizado antes de la semana de referencia para la ocupación. Esta partida considera solo ingresos del trabajo por actividades asalariadas dependientes.
- **INGDOTA**: ingreso disponible que proviene de actividades secundarias y actividades principales que han finalizado antes de la semana de referencia para la ocupación. Esta partida considera ingresos del trabajo por actividades asalariadas dependientes, menos los descuentos legales (previsión, salud e impuesto a las remuneraciones).
- **INGDOTI**: ingreso disponible que proviene de actividades secundarias y actividades principales que han finalizado antes de la semana de referencia para la ocupación. Esta partida considera solo ingresos del trabajo por actividades independientes, menos los costos asociados a su realización, descontando además las cotizaciones voluntarias en previsión y/o salud. Los costos mencionados anteriormente consideran impuestos, salarios que deben ser pagados a empleados u otros, bienes intermedios que deben ser comprados para la producción de la actividad, etc.

- **INGP:** ingresos que provienen del arriendo mensual de bienes inmuebles como: casas, departamentos, oficinas, locales comerciales, estacionamientos, tierras u otras propiedades, o de bienes muebles como: vehículos, maquinarias de todo tipo y animales. Considera los ingresos percibidos el mes anterior a la aplicación de la encuesta.
- **INGF:** ingresos derivados de la tenencia de instrumentos financieros como dividendos por acciones, utilidades por la participación en sociedades de personas, intereses de cuentas de ahorro, intereses por otros instrumentos financieros como depósitos a plazo, fondos mutuos, etc. Considera los ingresos percibidos el mes anterior a la aplicación de la encuesta.
- **TRANSFERENCIAS_RECIBIDAS:** corresponde a la suma de las transferencias corrientes en efectivo recibidas desde otros hogares, el Gobierno y/o instituciones sin fines de lucro al servicio de los hogares (ISFLSH), más las transferencias en especies recibidas únicamente desde otros hogares.
- **TRANSFERENCIAS_EMITIDAS:** corresponde a la suma de las transferencias corrientes en efectivo pagadas a otros hogares y/o instituciones sin fines de lucro al servicio de los hogares (ISFLSH).
- **INGT_NETA:** es el resultado de la suma de las transferencias corrientes recibidas por el hogar, menos las transferencias o donaciones emitidas. Las transferencias provienen de diferentes fuentes, como otros hogares, el Gobierno y/o instituciones sin fines de lucro al servicio de los hogares (ISFLH).
- **SALUD:** es el monto por concepto de cotizaciones de salud de todos los trabajadores (tanto dependientes como independientes). El cálculo de los descuentos se encuentra ajustado según las tasas de referencia manejadas por el sistema de salud y lo declarado previamente por los informantes.
- **PREVISION:** es el monto por concepto de cotizaciones previsionales en el sistema de AFP de todos los trabajadores (tanto dependientes como independientes). Los descuentos se encuentran ajustados para los asalariados según la tasa de referencia que utilizan las administradoras de pensiones y lo declarado previamente por los informantes.
- **IMPUESTO:** es el monto de descuento por concepto de impuestos legales a las remuneraciones, declarado por los trabajadores asalariados y honorarios dependientes.
- **CONTRIBUCIONES:** corresponde a la suma de los montos por concepto de descuentos en cotizaciones de salud y previsión de las personas que declararon durante la semana de referencia alguna actividad principal, actividad secundaria y/o jubilación.
- **ID_MISSING:** corresponde a la identificación de los hogares que presentan ingresos negativos o menores a \$80.000 pesos. Los ingresos negativos en el hogar podrían explicar una subdeclaración de ingresos por los informantes o también indicar que las transferencias emitidas son superiores a los ingresos declarados en su conjunto. El ingreso disponible de los hogares que presentan esta característica es considerado como dato *missing* (perdido) para el cálculo de los quintiles.
- **TA01:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo asalariado, ya sea en efectivo y/o especies.

- **TA02:** identifica el monto bruto recibido por aquellas personas que declararon ingresos por trabajo asalariado el mes anterior a la aplicación de la encuesta.
- **TA03:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo asalariado y declararon descuentos por cotizaciones en salud.
- **TA04:** identifica el monto descontado por cotizaciones en salud.
- **TA05:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo asalariado y declararon descuentos por cotizaciones previsionales.
- **TA06:** identifica el monto descontado por cotizaciones previsionales.
- **TA07:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo asalariado y declararon descuentos por impuesto único a las remuneraciones.
- **TA08:** identifica el monto descontado por impuesto único a las remuneraciones.
- **TA09:** identifica el monto líquido recibido por aquellas personas que declararon ingresos por trabajo asalariado el mes anterior a la aplicación de la encuesta.
- **TA10:** identifica aquellas personas que durante el mes pasado percibieron ingresos variables por la actividad asalariada declarada anteriormente.
- **TA11:** identifica el monto líquido promedio mensual de los últimos 12 meses, recibido por aquellas personas que declararon ingresos por trabajo asalariado el mes anterior a la aplicación de la encuesta.
- **TA12:** identifica aquellas personas que durante el mes pasado percibieron ingresos por horas extras en la actividad asalariada declarada anteriormente.
- **TA13:** identifica el monto recibido por concepto de horas extras.
- **TA14:** identifica aquellas personas que durante el mes pasado percibieron ingresos por aguinaldos en la actividad asalariada declarada anteriormente.
- **TA15:** identifica el monto recibido por concepto de aguinaldos.
- **TA16:** identifica aquellas personas que durante el mes pasado percibieron ingresos por asignaciones en la actividad asalariada declarada anteriormente.
- **TA17:** identifica el monto recibido por concepto de asignaciones.
- **TA18:** identifica aquellas personas que durante el mes pasado percibieron ingresos por incentivos y premios en la actividad asalariada declarada anteriormente.
- **TA19:** identifica el monto recibido por concepto de incentivos y premios.
- **TA20:** identifica la frecuencia de pago del ingreso percibido por concepto de incentivos y premios.
- **TA21:** identifica aquellas personas que durante el mes pasado percibieron ingresos por otros tipos de bonos y formas de ingresos variables en la actividad asalariada declarada anteriormente.
- **TA22:** identifica el monto recibido por concepto de otros tipos de bonos y formas de ingresos variables.
- **TA23:** identifica la frecuencia de pago del ingreso percibido por concepto de otros tipos de bonos y formas de ingresos variables.

- **TA24:** identifica aquellas personas que durante el mes pasado percibieron ingresos por alimentos y bebidas (tales como vales de colación, cheque restaurant, canasta familiar, etc.) en la actividad asalariada declarada anteriormente.
- **TA25:** identifica el monto recibido por concepto de alimentos y bebidas (tales como vales de colación, cheque restaurant, canasta familiar, etc.)
- **TA26:** identifica la frecuencia de pago del ingreso percibido por concepto alimentos y bebidas.
- **TA27:** identifica aquellas personas que durante el mes pasado percibieron ingresos por otras especies en la actividad asalariada declarada anteriormente. Por ejemplo: ingresos por vestuario y/o calzado, transporte (vehículo y bencina), vivienda, servicio de telefonía, entre otros.
- **TA29:** identifica el monto promedio mensual de lo recibido en especies valorizado a precios de mercado.
- **TI01:** identifica el monto recibido por aquellas personas que declararon ingresos por trabajo independiente. Corresponde al ingreso neto percibido el mes anterior a la aplicación de la encuesta, excluyendo los gastos en insumos, remuneraciones e impuestos.
- **TI02:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo independiente y declararon pagar de forma voluntaria cotizaciones en salud.
- **TI03:** identifica el monto pagado por cotizaciones en salud.
- **TI04:** identifica aquellas personas que durante el mes pasado percibieron ingresos por trabajo independiente y declararon pagar de forma voluntaria cotizaciones previsionales.
- **TI05:** identifica el monto pagado por cotizaciones previsionales.
- **TI06:** identifica el monto recibido por ingresos del trabajo independiente. Corresponde al ingreso disponible, excluyendo los pagos de cotizaciones en salud y previsión, percibido el mes anterior a la aplicación de la encuesta.
- **TI07:** identifica aquellas personas que durante el mes pasado percibieron ingresos variables por la actividad independiente declarada con anterioridad.
- **TI08:** identifica el monto disponible promedio mensual de los últimos 12 meses, recibido por ingresos del trabajo independiente el mes anterior a la aplicación de la encuesta.
- **TI09:** identifica aquellas personas que el mes anterior percibieron ingresos independientes y declararon tener personas a su cargo, excluyendo a familiares no remunerados.
- **TI10:** pregunta que permite diferenciar entre aquellas personas que percibieron ingresos derivados de su actividad por cuenta propia (negocio, empresa, etc.) o por su actividad profesional independiente.
- **TI11:** identifica aquellas personas que el mes anterior percibieron ingresos independientes y retiraron productos de su empresa, negocio o actividad por cuenta propia, para su consumo personal o del hogar.
- **TI12:** identifica el monto de los productos retirados en su empresa, negocio o actividad por cuenta propia, valorizado a precios de mercado.

- **OT01:** identifica aquellas personas que el mes anterior percibieron otros ingresos del trabajo, ya sea por segundas actividades realizadas o ingresos generados el mes anterior por actividades distintas a la declarada en el módulo AE de RPH (como ocupación principal).
- **OT02:** pregunta que permite diferenciar aquellas personas que el mes anterior percibieron otros ingresos de carácter asalariado.
- **OT03:** identifica el monto bruto recibido por otros ingresos del trabajo que corresponden a una actividad asalariada.
- **OT04:** identifica aquellas personas que durante el mes pasado percibieron otros ingresos del trabajo por una actividad asalariada y declaran descuentos por cotizaciones en salud y/o previsión.
- **OT05:** identifica el monto descontado por cotizaciones en salud.
- **OT06:** identifica el monto descontado por cotizaciones previsionales.
- **OT07:** identifica el monto líquido recibido por aquellas personas que declararon ingresos por trabajo asalariado el mes anterior a la aplicación de la encuesta.
- **OT08:** pregunta que permite diferenciar aquellas personas que el mes anterior percibieron otros ingresos por una actividad por cuenta propia o una actividad profesional independiente.
- **OT09:** identifica el monto disponible recibido el mes anterior por esa actividad independiente o bien el monto promedio recibido de los últimos 12 meses.
- **JU01:** identifica aquellas personas que durante el mes pasado percibieron ingresos por jubilación y/o pensiones de vejez.
- **JU02:** identifica el monto bruto recibido por ingresos de jubilación y/o pensiones de vejez, el mes anterior a la aplicación de la encuesta.
- **JU03:** identifica el monto líquido recibido por ingresos de jubilación y/o pensiones de vejez, el mes anterior a la aplicación de la encuesta.
- **TR01:** identifica aquellas personas que durante el mes pasado percibieron ingresos provenientes de otras pensiones. Por ejemplo: pensiones de invalidez, discapacidad, asistenciales, montepío, etc.
- **TR02:** identifica el monto líquido recibido ingresos provenientes de otras pensiones.
- **TR03:** identifica aquellas personas que durante el mes pasado percibieron ingresos provenientes de pensiones alimenticias.
- **TR04:** identifica el monto recibido ingresos provenientes de pensiones alimenticias.
- **TR05:** identifica aquellas personas que durante el mes pasado percibieron ingresos provenientes de otras transferencias en efectivo o donaciones desde algún otro hogar, ya sea dentro o fuera del país.
- **TR06:** identifica el origen del ingreso percibido por concepto de transferencias en efectivo o donaciones.
- **TR07:** identifica el monto recibido por ingresos provenientes de otras transferencias en efectivo o donaciones desde algún otro hogar, ya sea dentro o fuera del país.
- **TR08:** identifica aquellas personas que durante el mes pasado percibieron ingresos provenientes desde el Gobierno y/o alguna institución sin fines de lucro. Por ejemplo:

subsidio único de cesantía, subsidio único familiar, becas de estudio u otros subsidios del Estado.

- **TR09:** identifica el monto recibido por ingresos provenientes desde el Gobierno y/o alguna institución sin fines de lucro.
- **TR10:** identifica el monto promedio mensual de los últimos 12 meses, recibido por ingresos derivados de transferencias desde el Gobierno y/o alguna institución sin fines de lucro.
- **TR11:** identifica aquellas personas que durante el mes pasado percibieron ingresos por transferencias en especies.
- **TR13:** permite identificar el origen de los productos recibidos como transferencias en especies.
- **TR14:** identifica el monto de los productos recibidos como transferencias en especies, valorizado a precios de mercado.
- **NH01:** identifica aquellas personas que durante el mes pasado percibieron ingresos de carácter no habitual. Por ejemplo: ingresos por herencias, loterías y juegos de azar, ventas de vehículos del hogar, indemnizaciones u otros ingresos, etc.
- **NH02A:** identifica los ingresos no habituales percibidos por herencias.
- **NH02B:** identifica los ingresos no habituales percibidos por loterías y juegos de azar.
- **NH02C:** identifica los ingresos no habituales percibidos por ventas de vehículos del hogar.
- **NH02D:** identifica los ingresos no habituales percibidos por indemnizaciones.
- **NH02E:** identifica los otros ingresos no habituales percibidos por una persona y que no se encuentran especificados.
- **NH04:** identifica el monto recibido por ingresos de carácter no habitual.
- **IP01:** identifica aquellas personas que durante el mes pasado percibieron ingresos por arriendo de casas, departamentos, oficinas o locales comerciales, estacionamientos, tierras u otros.
- **IP02:** identifica el monto recibido por ingresos en arriendo de casas, departamentos, oficinas o locales comerciales, estacionamientos, tierras u otros.
- **IP03:** identifica aquellas personas que durante el mes pasado percibieron ingresos por arriendo de otro tipo de bienes, por ejemplo: maquinaria, taxis, vehículos colectivos, animales e implementos, etc.
- **IP04:** identifica el monto recibido por ingresos en arriendo de otro tipo de bienes.
- **IP05:** identifica aquellas personas que participan de alguna sociedad de personas.
- **IP06:** identifica el monto promedio mensual de los ingresos por retiro de utilidades en alguna sociedad de personas.
- **IP06B:** identifica el tramo de ingresos por retiro de utilidades en alguna sociedad de personas.
- **IP07:** identifica aquellas personas que durante el mes pasado poseían inversiones en algún tipo de acciones.
- **IP08:** identifica el monto promedio mensual de los ingresos percibidos por dividendos en acciones.
- **IP08B:** identifica el tramo de ingresos por dividendos de acciones.

- **IP09:** identifica aquellas personas que declaran poseer cuentas de ahorro o depósitos a plazo.
- **IP10_GANANCIA:** corresponde a una variable que identifica la ganancia promedio mensual que genera la tenencia de cuentas de ahorro o depósitos a plazo.
- **IP11:** identifica aquellas personas que declaran poseer otro tipo de instrumentos financieros no especificados, como fondos mutuos, APV, etc.
- **IP12_GANANCIA:** corresponde a una variable que identifica la ganancia promedio mensual que genera la tenencia de otros instrumentos financieros no especificados anteriormente, como fondos mutuos, APV, etc.
- **TE01:** identifica aquellas personas que durante el mes pasado realizaron transferencias en efectivo por concepto de pensiones alimenticias.
- **TE02:** identifica el monto entregado por concepto de pensiones alimenticias.
- **TE03:** identifica aquellas personas que durante el mes pasado realizaron transferencias en efectivo a otros hogares, ya sea dentro del país y/o fuera del país.
- **TE04:** pregunta que permite identificar el destino de las transferencias emitidas a otros hogares, ya sea dentro y/o fuera del país.
- **TE05:** identifica el monto entregado por transferencias en efectivo a otros hogares, ya sea dentro y/o fuera del país.
- **TE06:** identifica el monto promedio mensual entregado por transferencias en efectivo a otros hogares, ya sea dentro y/o fuera del país.
- **TE07:** identifica aquellas personas que durante el mes pasado realizaron transferencias en efectivo a instituciones sin fines de lucro al servicio de los hogares (ISFLSH). Por ejemplo: iglesias, Hogar de Cristo, etc.
- **TE08:** identifica el monto entregado por transferencias en efectivo a instituciones sin fines de lucro al servicio de los hogares (ISFLSH). Por ejemplo: iglesias, Hogar de Cristo, etc.
- **TE09:** identifica aquellas personas que durante el mes pasado realizaron transferencias en especies a otros hogares.
- **TE10_HOG:** identifica el monto asociado a los productos entregados por el hogar como transferencias en especies, valorizado a precios de mercado.

Anexo N°2. Componentes del ingreso

Componentes		Variables Observadas (nombres en cuestionario)	Variables Agregadas
I1. Ingresos del trabajo (producción)			ING_PRODUC_HD
I1.1.	Ingresos del empleo dependiente	TA02, TA25, TA29, OT03	INGTD_HD, INGOTA⁽¹⁾
I1.1.1.	Sueldos y salarios en efectivo	TA02, OT03	
	Horas adicionales en efectivo	TA13	
	Aguinaldo o asignaciones en efectivo	TA15, TA17	
	Incentivos y premios en efectivo	TA19	
	Otros bonos en efectivo	TA22	
I1.1.2.	Bienes y servicios gratuitos o subsidiados por el empleador ⁽¹⁾	TA25, TA29	
I1.2.	Ingresos del empleo independiente	TI01, TI12, OT09	INGTI_HD, INGDOTI⁽¹⁾
I1.2.1.	Ingresos de la empresa, negocio u actividad (excluyendo los gastos de producción)	TI01, OT09	
I1.2.2.	Bienes y servicios producidos para su propio consumo (excluye los costos de los insumos) ⁽¹⁾	TI12	
I2. Ingresos de la propiedad e intereses por instrumentos financieros			
I2.1.	Rentas de la propiedad	IP02, IP04	INGP
I2.2.	Intereses o beneficios derivados de activos financieros	IP06, IP08, IP10_GANANCIA, IP12_GANANCIA	INGF
I3. Transferencias corrientes recibidas (excluye transferencias sociales en especies desde el Gobierno e instituciones sin fines de lucro)			
I3.1.	Otro tipo de pensiones y beneficios sociales	TR02	TRANSFERENCIAS_RECIBIDAS
I3.2.	Jubilaciones y/o pensiones de vejez	JU02	INGJ_HD
I3.3.	Beneficios de asistencia social en efectivo derivados del Gobierno	TR09	TRANSFERENCIAS_RECIBIDAS
I3.5.	Transferencias en efectivo recibidas desde otros hogares	TR04, TR07	TRANSFERENCIAS_RECIBIDAS
I3.6.	Transferencias o donaciones en efectivo recibidas desde instituciones sin fines de lucro	TR09	TRANSFERENCIAS_RECIBIDAS
I3.7.	Otras transferencias recibidas en especie desde otros hogares	TR14	TRANSFERENCIAS_RECIBIDAS
	Ingreso primario = (I1 + I2)		ING_PRIMARIO_HD
IT	Ingreso total sin arriendo imputado = (I1 + I2 + I3)		ING_TOTAL_HOG_HD

(Continúa en la siguiente página)

Componentes		Variables Observadas (nombres en cuestionario)	Variables Agregadas
E2. Transferencias corrientes emitidas			
E2.1.	Impuesto a la renta	TA08	IMPUESTO
E2.2.	Contribuciones por seguridad social (descuento en salud y previsión) ⁽²⁾	TA04, TI03, OT05, TA06, TI05, OT06,	CONTRIBUCIONES
E2.3.	Transferencias en efectivo pagadas/emitidas a otros hogares	TE02, TE05	TRANSFERENCIAS_EMITIDAS
E2.4.	Transferencias en efectivo pagadas/emitidas a instituciones sin fines de lucro	TE08	TRANSFERENCIAS_EMITIDAS
ID	<i>Ingreso total disponible = (IT - E2)</i>		ING_DISP_HOG_HD ⁽²⁾
I4	<i>Arriendo imputado</i>		AI_VP, AI_VS
IT_AI	<i>Ingreso total con arriendo imputado = (IT + I4)</i>		ING_TOTAL_HOG_HD_AI ⁽³⁾
ID_AI	<i>Ingreso total disponible con arriendo imputado = (ID + I4)</i>		ING_DISP_HOG_HD_AI ⁽³⁾

Nota: - Existen dos grandes diferencias que se observan entre la clasificación de los componentes de los ingresos planteada por la OCDE y la utilizada en la VIII EPF. La primera, en la VIII EPF se excluyó los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables (en OCDE incorporadas en **I4**) y la segunda, las *Jubilaciones y/o pensiones de vejez* se encuentran separadas del agregado *Transferencias corrientes recibidas* (en OCDE incorporadas en **I3**).

- Para el cálculo de los ingresos del hogar y *per cápita*, totales y disponibles, se utilizan las variables imputadas por *hot-deck* (HD).

⁽¹⁾ Ingresos en especies no se corrigen por no respuesta al ítem y son sumadas a las partidas agregadas según corresponda la fuente de ingreso (INGTD_HD o INGTI_HD). Además, se debe considerar que INGOTA es igual a OT03, de la misma forma que, INGDOTI es igual a OT09, ya que no se corrigen por no respuesta al ítem.

⁽²⁾ La construcción de la variable Contribuciones incluye el pago de cotizaciones en salud de las personas que perciben ingresos por jubilación y/o pensiones de vejez. El monto se calcula a partir de la diferencia entre el ingreso bruto (INGJ_HD) y el líquido o disponible (INGDJ_HD) para aquellas personas que cotizan, ya que no se registra una pregunta específica asociada en la libreta.

⁽³⁾ En la Base de Personas se encuentran calculadas a nivel de hogar (Ingreso total e Ingreso total disponible).

Fuente: VIII EPF-INE en base a Eurostat (2003) y OECD (2013).

Anexo N°3. Comandos útiles

A continuación se presenta una breve descripción de algunos comandos útiles para trabajar las bases de datos de la encuesta.

1. Tipos de ponderación del factor de expansión:

Comando	Interpretación
<pre>weight command ... [weight=exp] ...</pre>	<p>Al ponderar con <i>weight</i> sin especificar el tipo de ponderación, Stata asume una forma por defecto para ponderar, de acuerdo al comando utilizado.</p> <p>Cada comando señala en su descripción en el <i>help</i>, qué tipo de ponderaciones permite.</p>
<pre>fweight command ... [fweight=exp] ...</pre>	<p>O <i>frequency weight</i>, duplica observaciones hasta alcanzar el peso que tiene cada caso en el factor de expansión. Por ejemplo, si FE==10, entonces <i>fweight</i> operará como si realmente hubieran 10 casos con información idéntica en la base de datos.</p> <p>Funciona cuando el factor de expansión es siempre un número entero.</p>
<pre>pweight command ... [pweight=exp] ...</pre>	<p>O <i>probability weight</i> denota el inverso de la probabilidad de que una observación sea incluida en la muestra a causa del diseño muestral.</p> <p>La escala en la que está el ponderador no es relevante en términos de parámetros estimados y errores estándar, excepto cuando se estiman totales y se calculan correcciones de población finita con comandos <i>svy</i>.</p>
<pre>aweight command ... [aweight=exp] ...</pre>	<p>O <i>analytic weight</i> es inversamente proporcional a la varianza de una observación.</p> <p>Típicamente, las observaciones representan promedios, y los pesos corresponden al número de observaciones que dieron origen a la media.</p>
<pre>iweight command ... [iweight=exp] ...</pre>	<p>O <i>importance weight</i> indica la importancia relativa de la observación y no tiene una definición estadística formal.</p> <p>Cada comando que lo permite define cómo tratará los datos. Es un comando destinado más bien al uso por parte de programadores que desean utilizar algún cálculo en específico.</p>

2. Configuraciones:

Comando	Interpretación
<code>set type double</code>	Se indica al programa usar su capacidad máxima, en uso y guardado de decimales.
<code>set dp comma / period</code>	Se indica al programa el separador de decimales preferido coma o punto. Si se desea que Stata recuerde la preferencia especificada permanentemente, es necesario agregar la opción <i>permanently</i> (ej: <code>set dp comma, permanently</code>)
<code>cd " "</code>	Se determina el directorio desde donde se accede directamente a la información a utilizar y a guardar.
<code>format var1 %20,2f</code>	Se define un formato a la variable, en el caso de las numéricas la f al final instituye una forma fija para presentarla, mientras que el número indicará la cantidad total de caracteres que se muestran y el número después de la coma la cantidad de decimales. Si se agrega la letra c después de la f, se agregarán los separadores de miles (ejemplo: <code>format var1 %20,2fc</code>)
<code>preserve</code>	Se guarda provisionalmente la base de datos.
<code>restore</code>	Vuelve a la versión guardada provisionalmente.
<code>quietly</code>	No se muestran los resultados en la ventana de resultados para la sentencia especificada.
<code>log using "logfile.log"</code>	Se guarda en block de notas el historial de los comandos ejecutados y los resultados, puede reemplazarse o ir añadiendo contenido.
<code>log close</code>	Se cierra el archivo log y se termina de grabar el historial.

3. Generación de variables simples y complejas:

Comando	Interpretación
<code>gen</code>	Para crear variables
<code>egen</code>	Para generar variables de forma compleja, es una extensión del comando anterior.
<code>gen var2=var1</code>	Se genera una variable nueva igual a la anterior, si bien copia la capacidad de almacenamiento es indiferente al formato de la variable- texto, numérica.
<code>gen var=_n</code>	Se genera una nueva variable que enumera las observaciones según su orden dentro de un conjunto de observaciones.
<code>gen var=_N</code>	Se genera una variable cuyo valor siempre es el mismo y ese valor es la cantidad total dentro de un conjunto de observaciones.
<code>gen var=.</code>	Se genera una variable numérica sin datos.
<code>gen var=""</code>	Se genera una variable de texto sin datos.

Comando	Interpretación
<code>gen var2=var1^2</code>	Se genera una variable que es el cuadrado de otra variable
<code>egen var3=rowtotal(var1 var2)</code>	Se genera una variable que es la suma horizontal de las variables seleccionadas, considerando los vacíos o missing como cero.
<code>label var var1 "Etiqueta"</code>	Se establece una etiqueta para la variable, si esta tiene una, se reemplaza sin advertencia.
<code>rename var1 var2</code>	Se cambia el nombre de la variable, var1 ahora se llama var2.
<code>replace var1=0 if var1==2</code>	Se cambia el valor de la variable de 2 a 0.
<code>drop var1</code>	Se elimina la variable var1.
<code>keep if var1 >=6</code> <code>drop if var1<6</code>	Se conservan solo las observaciones que tienen un valor menor de 6 en la variable (5 o menos). Es equivalente a borrar todas las observaciones que tienen un valor de 5 o menos en la variable var1.

4. Para estadística descriptiva:

Comando	Interpretación																																												
<pre> tabstat EJEMPLO use "BASE_PERSONAS_VIII_ EPF.dta", clear tabstat EDAD [aw=FE] if EDAD>=0, stat(mean sd p50 min max) f(%5,2f) </pre>	<p>El comando permite calcular los estadísticos de una variable o una lista, considerando filtros (aplicados al conjunto de análisis) y factores de expansión.</p> <p>Entre los estadísticos que se pueden calcular:</p> <table border="1"> <tbody> <tr> <td>mean</td> <td><i>Promedio</i></td> <td>p1</td> <td><i>Percentil 1</i></td> </tr> <tr> <td>count o n</td> <td><i>Cantidad (unidades muestrales)</i></td> <td>p5</td> <td><i>Percentil 5</i></td> </tr> <tr> <td>sum</td> <td><i>Suma total de la variable</i></td> <td>p10</td> <td><i>Percentil 10</i></td> </tr> <tr> <td>max</td> <td><i>Valor máximo</i></td> <td>p25</td> <td><i>Percentil 25</i></td> </tr> <tr> <td>min</td> <td><i>Valor mínimo</i></td> <td>median o p50</td> <td><i>Percentil 50 (mediana)</i></td> </tr> <tr> <td>range</td> <td><i>Rango (resta = max - min)</i></td> <td>p75</td> <td><i>Percentil 75</i></td> </tr> <tr> <td>sd</td> <td><i>Error Estándar</i></td> <td>p90</td> <td><i>Percentil 90</i></td> </tr> <tr> <td>variance</td> <td><i>Varianza</i></td> <td>p95</td> <td><i>Percentil 95</i></td> </tr> <tr> <td>cv</td> <td><i>Coficiente de variación (sd/mean)</i></td> <td>p99</td> <td><i>Percentil 99</i></td> </tr> <tr> <td>skewness</td> <td><i>Oblicuidad</i></td> <td>iqr</td> <td><i>Rango intercuartil (p75-p25)</i></td> </tr> <tr> <td>kurtosis</td> <td><i>Curtosis</i></td> <td>q</td> <td><i>Equivalente a P25 p50 p75</i></td> </tr> </tbody> </table> <p>En el ejemplo se estima el promedio, el error estándar, la mediana, el mínimo y el máximo de la EDAD (sin contar los tres casos que no cuentan esta información).</p>	mean	<i>Promedio</i>	p1	<i>Percentil 1</i>	count o n	<i>Cantidad (unidades muestrales)</i>	p5	<i>Percentil 5</i>	sum	<i>Suma total de la variable</i>	p10	<i>Percentil 10</i>	max	<i>Valor máximo</i>	p25	<i>Percentil 25</i>	min	<i>Valor mínimo</i>	median o p50	<i>Percentil 50 (mediana)</i>	range	<i>Rango (resta = max - min)</i>	p75	<i>Percentil 75</i>	sd	<i>Error Estándar</i>	p90	<i>Percentil 90</i>	variance	<i>Varianza</i>	p95	<i>Percentil 95</i>	cv	<i>Coficiente de variación (sd/mean)</i>	p99	<i>Percentil 99</i>	skewness	<i>Oblicuidad</i>	iqr	<i>Rango intercuartil (p75-p25)</i>	kurtosis	<i>Curtosis</i>	q	<i>Equivalente a P25 p50 p75</i>
mean	<i>Promedio</i>	p1	<i>Percentil 1</i>																																										
count o n	<i>Cantidad (unidades muestrales)</i>	p5	<i>Percentil 5</i>																																										
sum	<i>Suma total de la variable</i>	p10	<i>Percentil 10</i>																																										
max	<i>Valor máximo</i>	p25	<i>Percentil 25</i>																																										
min	<i>Valor mínimo</i>	median o p50	<i>Percentil 50 (mediana)</i>																																										
range	<i>Rango (resta = max - min)</i>	p75	<i>Percentil 75</i>																																										
sd	<i>Error Estándar</i>	p90	<i>Percentil 90</i>																																										
variance	<i>Varianza</i>	p95	<i>Percentil 95</i>																																										
cv	<i>Coficiente de variación (sd/mean)</i>	p99	<i>Percentil 99</i>																																										
skewness	<i>Oblicuidad</i>	iqr	<i>Rango intercuartil (p75-p25)</i>																																										
kurtosis	<i>Curtosis</i>	q	<i>Equivalente a P25 p50 p75</i>																																										

5. Exportación de tabulados y gráficos

Comando	Interpretación
<pre>putexcel ssc install putexcel EJEMPLO use "BASE_PERSONAS_VIII_ EPF.dta", clear svyset VARUNIT [pw=FE], strata(VARSTRAT) singleunit(certainty) svy: mean GASTOT_HD if JHOGAR==1 estat cv qui mat GT = (r(b)', r(se)', r(cv)') qui putexcel C3=matrix(GT) C2=("Promedio") D2=("Error Estándar") E2=("CV") using "Tabla ejemplo.xlsx", sheet("Hoja1") modify</pre>	<p>Se debe instalar primero.</p> <p>Es un comando que sirve para la exportación de expresiones, matrices, y resultados almacenados.</p> <p>En el ejemplo, se guardan los resultados de la estimación del promedio, error estándar y coeficiente de variación del gasto total. Los títulos especificados en la fila 2 y la información en la fila 3.</p> <p>El archivo de destino debe ser creado antes de exportar algún dato.</p> <p>Se debe considerar que en versiones 14 y superiores se podrá mantener el formato de las celdas de destino (<code>keepcellformat</code>).</p>
<pre>graph export EJEMPLO Programación 18</pre>	<p>Se debe ejecutar la realización de un gráfico primero, luego se puede guardar en formato ps, eps, wmf, emf, pdf, png, o tif.</p>

6. Programación de loops (foreach y forvalues)

Comando	Interpretación
<pre>foreach var in var1 var2 { comando sobre `var' }</pre>	<p>Es un comando que permite repetir una rutina para cada variable que se enliste en la misma línea antes de la llave "{". Este procedimiento puede ser cualquiera enlistado en el punto 2, 3 o 4 del presente anexo. Los procedimientos de la rutina (puede ser más de uno) o comandos sobre la lista de variables se finalizan con la llave de cierre "}". Desde este punto empezará a iterar hasta cumplir con la ejecución de la lista.</p>
<pre>forvalues val=n1(n2)n3 { comando con `val' }</pre>	<p>Es un comando es similar al anterior. Se ejecutará la rutina de tomando los valores que se establezcan en el rango n1 a n3, para la cada n1, n1+n2, n1+n2+n2..., hasta n3.</p>