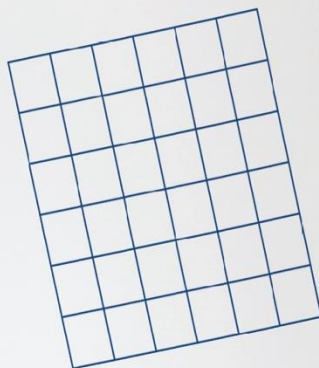


# ENCUESTA ESTRUCTURAL EMPRESARIAL 2018

Metodología de  
Diseño Muestral

Abril- 2020



Dirección de Infraestructura Estadística y Muestreo

# Metodología del Diseño Muestral de la Encuesta Estructural Empresarial

Año 2018

Quito - Ecuador, 2020

**Dirección**

Dirección de Infraestructura Estadística y Muestreo

**Unidad**

Gestión de Diseño Muestral

**Elaborado por:**

Pilar Carrillo

**Revisado por:**

Christian Garcés

Javier Núñez

**Aprobado por:**

Sebastián Carvajal

## Tabla de contenido

<b>INTRODUCCIÓN</b> .....	<b>6</b>
<b>ANTECEDENTES</b> .....	<b>7</b>
<b>1 PLANIFICACIÓN</b> .....	<b>8</b>
<b>1.1 OBJETIVOS Y DELIMITACIÓN DE LA OPERACIÓN ESTADÍSTICA</b> .....	<b>8</b>
<b>1.1.1 OBJETIVOS</b> .....	<b>8</b>
<i>Objetivo General</i> .....	8
<i>Objetivos Específicos</i> .....	8
<b>1.1.2 DELIMITACIÓN</b> .....	<b>8</b>
<i>Tipo de operación estadística</i> .....	8
<i>Cobertura temática</i> .....	8
<i>Universo de estudio</i> .....	9
<i>Unidad de observación y unidad de análisis</i> .....	9
<i>Cobertura geográfica</i> .....	9
<i>Población objetivo</i> .....	9
<i>Periodicidad y continuidad</i> .....	10
<b>2 DISEÑO Y CONSTRUCCIÓN</b> .....	<b>10</b>
<b>2.1 DISEÑAR EL MARCO Y LA MUESTRA</b> .....	<b>10</b>
<b>2.1.1 TIPO Y ETAPAS DE DISEÑO MUESTRAL</b> .....	<b>10</b>
<b>2.1.2 DOMINIOS DE ESTUDIO</b> .....	<b>10</b>
<b>2.1.3 MARCO MUESTRAL</b> .....	<b>10</b>
<b>2.1.4 VARIABLES DE DISEÑO</b> .....	<b>11</b>
<b>2.3.5 TAMAÑO MUESTRAL</b> .....	<b>11</b>
<b>2.3.6 SELECCIÓN DE LA MUESTRA</b> .....	<b>12</b>
<b>2.3.7 CONTROL DE COBERTURA DE CAMPO Y MUESTRAL</b> .....	<b>12</b>
<b>2.3.8 FACTORES DE EXPANSIÓN</b> .....	<b>14</b>
<b>2.3.9. POSTESTRATIFICACIÓN</b> .....	<b>14</b>
<b>2.4.0 ESTIMACIÓN DE CARACTERÍSTICAS</b> .....	<b>16</b>
<b>2.4.1 ERROR DE MUESTREO</b> .....	<b>16</b>
<b>2.4.2 DECLARACIÓN DEL PLAN DE MUESTREO</b> .....	<b>16</b>
<b>BIBLIOGRAFÍA</b> .....	<b>18</b>

## Lista de tablas

Tabla 1. Tamaño de empresas según la estratificación de la CAN .....	9
Tabla 2. Distribución de la muestra según el tamaño de empresa.....	12
Tabla 3. Estructura de condición de levantamiento de información .....	12
Tabla 4. Seguimiento de cobertura por dominio de estudio.....	13
Tabla 5. Cambio de tamaño de las empresas de la muestra.....	14
Tabla 6. Cambio de sector económico de las empresas pertenecientes a la muestra .....	15

# INTRODUCCIÓN

El Instituto Nacional de Estadísticas y Censos (INEC), organismo coordinador de la Producción Estadística en el Ecuador, realiza anualmente encuestas económicas dirigidas a empresas a nivel nacional.

En este sentido, la Dirección de Estadísticas Económicas solicita el apoyo de la Dirección de Infraestructura Estadística y Muestreo en la creación del diseño muestral para la "Encuesta Estructural Empresarial 2018". Esta encuesta se realiza a nivel nacional para las distintas secciones económicas contempladas por el Clasificador Industrial Internacional Uniforme de Actividades Económicas CIIU-4 para empresas grandes y medianas, en esta encuesta en particular se utiliza el CIIU-4 a un dígito.

El proceso que se realiza para la generación del diseño muestral consta de tres etapas:

En la primera etapa se toma en cuenta:

- Determinación del marco de muestreo
- El universo de la investigación
- Las unidades de muestreo y análisis
- Cobertura geográfica
- Cobertura temática

En la segunda etapa:

- Metodología para el diseño muestral
  - Determinación de los dominios de estudio
- Determinación del tamaño de la muestra
- Selección de la muestra
- Validación de la cobertura obtenida en campo

Por último, se realiza:

- El cálculo de los factores de expansión
- Postestratificación
- Estimación de características
- Error de muestreo
- Declaración del plan de muestreo

Cabe mencionar, que debido a la dinámica de las empresas, se utiliza el método conocido como postestratificación, con la finalidad de ajustar los cambios producidos en las unidades de observación por actualización del Directorio de Empresas y Establecimientos.

## ANTECEDENTES

En el año de 1955, por disposición del Departamento de Investigaciones Económicas del Banco Central, se realizó el primer Censo Económico del sector industrial de Ecuador. Y a partir de 1962, se empezó a realizar la Encuesta de Manufactura y Minería bajo el nombre de "Encuestas Industriales", la cual se ejecutó hasta 1964.

En 1965 se llevó a cabo el segundo censo de Manufactura y Minería y el primer censo de Comercio y Servicios. Más adelante, desde 1966 hasta 1979, se investigó formalmente la Encuesta de Manufactura y Minería, Encuesta de Comercio y Encuesta de Hoteles Restaurantes y Servicios, utilizando por primera vez la Clasificación Industrial Internacional Uniforme (CIIU) de las Naciones Unidas.

A partir de 1981, y en base al Censo Económico de 1980, comenzó a ejecutarse la encuesta en dos partes: Tomo I y Tomo II. Cabe mencionar que en 1980 no se realizaron las Encuestas Industriales debido al Censo Económico, puesto que éste incluyó todos los sectores económicos. En el marco de la mejora continua, y de satisfacer las necesidades de los usuarios, a partir del 2008 se cambió el modo de recolección, al pasar de formularios físicos a electrónicos. Así también, se incrementaron, eliminaron o desagregaron algunas variables y capítulos.

En el 2010 el INEC realizó el Censo Nacional Económico (CENEC), el cual fue insumo principal para la actualización del marco muestral, ya que fue realizado en dos etapas. La primera consistió en el empadronamiento a fin de ubicar geográficamente a los establecimientos del universo y contar con un directorio actualizado; la segunda consistió en el levantamiento de la Encuesta Exhaustiva con año de referencia 2011, periodo para el cual no se realizaron las Encuestas Industriales. Es así que, a partir del periodo 2012 se retomó su producción, pero bajo un esquema de marcos muestrales, diseño y estructura de formularios que no la hacía comparable ni con la serie de las Encuestas Industriales de 2010 hacia atrás, ni con la Encuesta Exhaustiva de 2011.

Frente a esta situación, y con el fin de dar continuidad a las series de las principales variables macroeconómicas calculadas en las Encuestas Industriales del 2010 hacia atrás, se procedió a realizar el ejercicio de empalme de encuestas industriales 2010 y 2012 con la Encuesta Exhaustiva 2011, que consistió en analizar la composición de los sectores económicos por cada rama de actividad, con la finalidad de dar continuidad al cálculo de los principales agregados económicos para que sean comparables a través del tiempo. Para la construcción de la serie, se tomó como línea base la información de las Encuestas Industriales (2010). Así, se parte del análisis de la Encuesta Exhaustiva (2011) y se retoman las Encuestas Industriales (2012 -2015).

Para el periodo de referencia 2016, surge la necesidad de implementar mejoras en las estadísticas empresariales, es por esta razón que se realiza una revisión metodológica de la Encuesta Estructural Empresarial. Esta situación derivó en la implementación de un nuevo diseño muestral, homologándose las Encuestas Industriales (Manufactura, Comercio y Servicios) en un formulario denominado Encuesta Estructural Empresarial, con el fin de reducir la carga de respuesta de los informantes, como una aproximación al uso de registros administrativos, así como, la mejora en los módulos de Información Ambiental, y de uso de Tecnologías de la Información y la Comunicación (TIC).

# 1 PLANIFICACIÓN

## 1.1 Objetivos y delimitación de la operación estadística

### 1.1.1 Objetivos

#### Objetivo General

Elaborar el Diseño Muestral de la “Encuesta Estructural Empresarial 2018”, utilizando técnicas de muestreo probabilístico que permitan emitir resultados confiables de las principales variables que se investigará en dicha encuesta.

#### Objetivos Específicos

- Garantizar la representatividad de la encuesta tanto a nivel geográfico como de los dominios de estudio.
- Calcular factores de expansión y errores asociados al muestreo

### 1.1.2 Delimitación

#### Tipo de operación estadística

Esta operación estadística corresponde a una encuesta dirigida a empresas.

#### Cobertura temática

A nivel de secciones económicas de la CIIU-4, las empresas a investigar pertenecen a las siguientes ramas:

- B.** Explotación de minas y canteras.
- C.** Industrias manufactureras.
- D.** Suministro de electricidad, gas, vapor y aire acondicionado.
- E.** Distribución de agua; alcantarillado, gestión de desechos y actividades de saneamiento
- F.** Construcción.
- G.** Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas.
- H.** Transporte y almacenamiento.
- I.** Actividades de alojamiento y de servicio de comidas.
- J.** Información y comunicación.
- K.** Actividades financieras y de seguros (No incluye K64 y K66).
- L.** Actividades Inmobiliarias
- M.** Actividades profesionales científicas y técnicas.
- N.** Actividades de servicios administrativos y de apoyo.
- P.** Enseñanza.
- Q.** Actividades de atención de la salud humana y de asistencia social (No incluye Q88).
- R.** Artes, entretenimiento y recreación.
- S.** Otras actividades de servicios (No incluye sólo la división S94).



Se excluyen las secciones:

- A.** Agricultura, ganadería, silvicultura y pesca.
- O.** Administración pública y defensa; planes de seguridad social de afiliación obligatoria.
- T.** Actividades de los hogares como empleadores; actividades no diferenciadas de los hogares como productores de bienes y servicios para uso propio.
- U.** Actividades de organizaciones y órganos extraterritoriales.

En el caso de la sección **A**, se debe a que la forma de investigar dicha sección consta en una encuesta específica de estadísticas agropecuarias. Por otro lado, las secciones **O**, **T** y **U** no hacen referencia a la población objetivo que son las empresas.

Como parte de la cobertura temática se tiene los tamaños de empresa:

- Mediana empresa "A",
- Mediana empresa "B", y,
- Grande empresa.

A los cuales se los han definido en base a la estratificación de la Comunidad Andina de Naciones (CAN), establecida de la siguiente manera:

**Tabla 1. Tamaño de empresas según la estratificación de la CAN**

Clasificación de la empresa	Personal Ocupado	Ventas anuales (USD\$)
Pequeña empresa	10 – 49	100.001 - 1.000.000
Mediana empresa Tipo A	50 – 99	1.000.001 - 2.000.000
Mediana empresa Tipo B	100 – 199	2.000.001 - 5.000.000
Grande empresa	200 y más	5.000.001 y más

Fuente: CAN, Decisión 702, Artículo 3.

### **Universo de estudio**

Las empresas de tamaño grande, mediana A y mediana B que cumplen con los criterios de clasificación de cobertura temática para el año 2018.

### **Unidad de observación y unidad de análisis**

Empresas

### **Cobertura geográfica**

La cobertura geográfica está definida por las empresas que se encuentren ubicadas dentro del territorio ecuatoriano incluyendo la región Insular.

### **Población objetivo**

Una consideración importante a la hora de estructurar un marco maestro de muestreo, es la relación entre la población objetivo de las encuestas y las unidades de selección. Las unidades de selección deberán estar determinadas en el marco de muestreo, así como también la probabilidad de selección de cada una de las unidades de selección en las diferentes etapas del muestreo. Las unidades del marco se constituyen en un

elemento básico para el cálculo, asignación y selección de la muestra de unidades poblacionales.

### **Periodicidad y continuidad**

Anual

## **2 DISEÑO Y CONSTRUCCIÓN**

### **2.1 Diseñar el marco y la muestra**

#### **2.1.1 Tipo y etapas de diseño muestral**

Muestreo probabilístico de elementos con selección aleatoria.

#### **2.1.2 Dominios de estudio**

Los dominios de estudio de la encuesta son los sectores económicos por tamaño de empresa a nivel nacional, detallados en la sección de Cobertura temática. Dentro de cada dominio de estudio se considera como grupo de inclusión forzosa a las empresas con tamaño "Grande empresa", es decir, estas empresas entran con probabilidad uno (1) a formar parte de la muestra. Cabe recalcar que se incluirá a las empresas que en 2017 eran de tamaño "Grande empresa" y en el 2018 son "Medianas B", con el fin de llevar el seguimiento de estas empresas en sus características.

#### **2.1.3 Marco muestral**

El Marco de Muestreo para la "Encuesta Estructural Empresarial 2018" está conformado por las empresas que constan en el Directorio de Empresas y Establecimientos Económicos 2017 (DIEE-2017), que contiene principalmente variables de identificación; variables de ubicación; variables de estratificación; y, variables de diseño y control de las empresas; indispensables para el diseño muestral.

Para la construcción del Marco Muestral, se toma en cuenta los sectores económicos a los que se enfoca la encuesta que están determinados por la cobertura temática y son los siguientes:

- Manufactura
- Minería
- Construcción
- Comercio
- Servicios

Luego de revisar la base de datos del directorio Empresas y Establecimientos del año 2017 (DIEE-2017) conformado por 888.946 empresas. Para la construcción del Marco Muestral se seleccionó empresas que presenten las siguientes características, respondiendo a las necesidades de información planificadas para la encuesta:

- Empresas de tamaño: Grande Empresa, Mediana Empresa Tipo B y Mediana Empresa Tipo A.
- Se excluye empresas cuya actividad Económica Principal de acuerdo a la sección de Clasificación Industrial Internacional Uniforme de Actividades Económicas

Revisión (CIIU-4) corresponde a Agricultura (A), Administración Pública (O), Actividades de hogares (T) y Actividades de organizaciones (U).

- Se excluye de la selección las siguientes divisiones del CIIU
  - Actividades de servicio financiero, excepto de seguros (K64)
  - Actividades auxiliares de servicio financiero (K66)
  - Actividades de Asistencia Social sin alojamiento (Q88)
  - Actividades de Asociaciones (S94)
- Se excluyen las empresas con forma institucional "Institución Pública".
- Se excluye empresas reportadas como no ubicadas de acuerdo a la variable "empresas\_noubicadas" del DICE.
- Se excluye las empresas Grandes y Medianas "B" con forma institucional "Personas naturales no obligadas a llevar contabilidad".

#### 2.1.4 Variables de diseño

El tamaño de la muestra se calcula en función a su variable de diseño "Ventas totales 2017"

#### 2.3.5 Tamaño muestral

Para la determinación del tamaño de la muestra se requiere establecer la característica o características a estimar, el nivel de confianza y la precisión requeridas, de tal manera que los resultados obtenidos no sean demasiado costosos y/o imprecisos.

En este sentido, la variable de diseño, en base a la cual se calcula el tamaño de la muestra es "Ventas totales 2017". Los valores poblacionales (marco muestral) de esta variable se toman de la base de datos del Directorio de Empresas y Establecimientos del año 2017 (DIEE 2017).

La fórmula para la determinación del tamaño de la muestra corresponde a un muestreo probabilístico para el estimador del total.

$$n_i \geq \frac{(N_i \cdot S_i)^2}{\left(\frac{N_i - 1}{N_i}\right) \cdot \left(\frac{e}{z}\right)^2 + (N_i \cdot S_i^2)} \cdot (1 + SOB)$$

Dónde:

- $n_i$  = Tamaño de la muestra por dominio
- $N_i$  = Tamaño del dominio i
- $S_i$  = Cuasivarianza del dominio i
- $e$  = Error relativo máximo admisible
- $z$  = Coeficiente que representa el nivel de seguridad o confianza
- SOB = Sobremuestreo

La cuasivarianza representa la variabilidad del dominio a partir de la información de la variable de diseño.

A partir del marco de muestreo constituido por 13.695 empresas, se ha diseñado una muestra de tamaño igual a 4.088 empresas, con el 10% de error relativo y un nivel de confianza del 90% para la variable de diseño: Ventas totales 2017, con un sobremuestreo del 35% obteniéndose una fracción de muestreo global de 0,2985.

La muestra se distribuye de la siguiente manera:

**Tabla 2. Distribución de la muestra según el tamaño de empresa**

Tamaño de empresa	Número de empresas
Mediana empresa "A"	392
Mediana empresa "B"	633
Grande empresa	3.063
<b>Total</b>	<b>4.088</b>

Fuente: Directorio de Empresas y Establecimientos Económicos 2017 (DIEE 2017)

### 2.3.6 Selección de la muestra

La selección de las empresas que forman parte de la muestra, de acuerdo al tamaño establecido, se realiza independientemente en cada uno de los dominios de manera aleatoria, asignando a cada empresa igual probabilidad.

### 2.3.7 Control de cobertura de campo y muestral

La cobertura hace referencia a la información obtenida en campo y proporcionada por los informantes de las empresas, puede estar constituida de dos formas:

**Tabla 3. Estructura de condición de levantamiento de información**

Condición	Novedad
Efectiva	Cambio de rama de actividad
	Cambio de sector económico
	Cambio de razón social
	Cambio de RUC
	Cambio de Provincia de la empresa
	Empresas que producen para terceros
	Fusionadas
	Absorbidas
	Desintegración
	Escisión
No efectiva	No ubicadas
	Rechazos
	Liquidadas
	Sin características
	Inactivas
	Sin contabilidad en el año de referencia

Fuente: Encuesta Estructural Empresarial (ENESEM 2017)

Se debe considerar que la cobertura está en función de cada uno de los dominios y se toma en cuenta variables como:

- universo,
- muestra mínima requerida (muestra sin sobremuestreo); y,
- muestra final.

Posteriormente, se realizan seguimientos de cobertura con el fin de verificar el total de empresas efectivas, ya que de este dependerán las posteriores estimaciones en los indicadores.

Como se puede observar en la tabla 4, del total de la muestra de la Encuesta Estructural Empresarial 2018 (4.088 casos) se obtuvo una cobertura de 3.723 casos (efectivas), es decir, una efectividad del 91%.

En cuanto a la cobertura final, se hace referencia a las efectivas sobre la muestra final, lo que permite conocer la cobertura por dominio de estudio como se observa en la siguiente tabla:

**Tabla 4. Seguimiento de cobertura por dominio de estudio**

Dominio	Universo	Muestra sin sobremuestreo	Muestra con sobremuestreo	Efectivas	Cobertura final (%)
3B	51	28	38	25	65,8%
3C	662	13	18	11	61,1%
3D	10	4	5	4	80,0%
3E	22	16	22	18	81,8%
3F	418	27	36	12	33,3%
3G	3.306	13	18	11	61,1%
3H	424	14	19	13	68,4%
3I	183	16	21	17	81,0%
3J	105	13	17	14	82,4%
3K	6	4	6	4	66,7%
3L	136	14	19	8	42,1%
3M	445	20	27	15	55,6%
3N	216	24	33	20	60,6%
3P	155	24	33	32	97,0%
3Q	117	24	32	26	81,3%
3R	40	27	37	24	64,9%
3S	32	8	11	9	81,8%
4B	39	24	33	27	81,8%
4C	549	51	69	54	78,3%
4D	10	7	10	8	80,0%
4E	17	13	17	17	100,0%
4F	255	53	71	49	69,0%
4G	2.289	85	115	96	83,5%
4H	264	30	41	36	87,8%
4I	90	19	26	25	96,2%
4J	69	19	25	23	92,0%
4K	4	3	4	3	75,0%
4L	83	23	31	24	77,4%
4M	217	34	46	36	78,3%
4N	172	33	44	41	93,2%
4P	101	24	33	33	100,0%
4Q	114	27	37	36	97,3%
4R	16	12	16	16	100,0%
4S	15	11	15	14	93,3%
5B	74	74	74	69	93,2%
5C	679	679	679	651	95,9%
5D	19	19	19	18	94,7%
5E	24	24	24	23	95,8%
5F	169	169	169	148	87,6%
5G	1.436	1.436	1.436	1.379	96,0%
5H	146	146	146	142	97,3%
5I	51	51	51	51	100,0%
5J	60	60	60	58	96,7%
5K	45	45	45	42	93,3%
5L	36	36	36	29	80,6%
5M	90	90	90	84	93,3%
5N	78	78	78	75	96,2%
5P	67	67	67	64	95,5%
5Q	65	65	65	65	100,0%
5R	12	12	12	12	100,0%
5S	12	12	12	12	100,0%
<b>Totales</b>	<b>13.695</b>	<b>3.822</b>	<b>4.088</b>	<b>3.723</b>	<b>91,07%</b>

**Fuente:** Diseño Muestral, INEC 2020

**Nota:** La variable "Dominio" está constituido por un número y una letra, el número describe el tamaño de la empresa, es decir, 3: Mediana tipo "A", 4: Mediana tipo "B" y 5: Grande. Por otro lado, la letra adjunta representa la rama de actividad a la que pertenece y está anclada con el CIIU-4, el cual está especificado en el apartado "Cobertura temática".

### 2.3.8 Factores de expansión

El objetivo de una encuesta por muestreo es hacer inferencias acerca de la población a partir de la información contenida en la muestra. En este sentido se utilizan factores de expansión para llevar la información de la muestra a niveles poblacionales. Se define como el inverso de la probabilidad de escoger una empresa en un determinado estrato, considerando el informe de cobertura que registra las novedades de la empresa.

El factor de expansión para cada dominio de estudio, se determina en base al número de empresas tanto del marco de muestreo como de la muestra, así:

$$a_{k_i} = \frac{N_k}{n_k}$$

Dónde:

$a_{k_i}$  = Factor de expansión de la empresa  $i$  perteneciente al dominio  $k$ .

$N_k$  = Tamaño de la población del dominio  $k$ .

$n_k$  = Tamaño de la muestra del dominio  $k$ .

A partir de ahora, se notará por  $a_k$  al factor de expansión de la empresa  $i$  perteneciente al dominio  $k$   $a_{k_i}$ .

### 2.3.9. Postestratificación

Considerando que la muestra fue seleccionada a partir del DIEE 2017 y el periodo de la información levantada en la encuesta corresponde al año 2018, existen empresas que por el dinamismo económico cambiaron de tamaño y/o sector económico. En consecuencia, los dominios de estudio originales sufrieron cambios.

En la tabla 5, se evidencia el cambio del tamaño de las empresas en la muestra, donde se registran 725 variaciones, lo cual corresponde al 17,7% del total de la muestra.

**Tabla 5. Cambio de tamaño de las empresas de la muestra**

	DIEE 2018							
	Tamaño de empresa	Sin tamaño	Micro empresa	Pequeña empresa	Mediana empresa tipo "A"	Mediana empresa tipo "B"	Grande Empresa	Totales
DIEE 2017	Mediana empresa tipo "A"	8	12	92	234	41	5	158
	Mediana empresa tipo "B"	8	20	47	75	391	92	242
	Grandes Empresa	24	19	32	26	224	2.738	325
	<b>Totales</b>	<b>40</b>	<b>51</b>	<b>171</b>	<b>101</b>	<b>265</b>	<b>97</b>	<b>725</b>

Fuente: Diseño Muestral, INEC 2020

Adicional, para el 2018 se dan de baja 262 empresas de la población objetivo, ya que se distribuyen entre sin tamaño (40), micro empresa (51) y pequeñas empresas (171).

De igual forma existen cambios en las secciones económicas, estos se presentan en 142 empresas, es decir, el 3,47% del total de la muestra. Por otro lado, es necesario destacar que 40 empresas (0,98%) pasaron a no tener ninguna rama de actividad.

**Tabla 6. Cambio de sector económico de las empresas pertenecientes a la muestra**

	DIEE 2018																				
	Sección Económica	Sin rama	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P	Q	R	S	Total
DIEE 2017	B	0	0	141	1	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	4
	C	4	0	0	751	0	0	2	6	0	0	1	0	0	2	0	0	0	0	0	15
	D	0	0	0	0	33	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	E	0	0	0	0	0	63	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	9	0	1	2	0	0	260	1	1	0	0	0	0	2	0	0	0	0	0	16
	G	16	3	1	6	0	0	1	1.537	1	0	0	0	1	2	0	0	0	0	1	32
	H	2	0	0	1	0	0	0	7	195	0	0	0	0	0	1	0	0	0	0	11
	I	0	1	0	0	0	0	0	4	0	93	0	0	0	0	0	0	0	0	0	5
	J	0	0	0	0	0	0	1	1	0	0	99	0	0	0	1	0	0	0	0	3
	K	2	0	0	0	0	0	0	0	0	0	0	53	0	0	0	0	0	0	0	2
	L	2	0	0	0	0	0	6	6	0	0	0	0	71	1	0	0	0	0	0	15
	M	3	0	0	2	0	1	5	3	1	0	2	1	1	141	3	0	0	0	0	22
	N	1	0	0	0	0	0	1	1	1	0	0	0	1	0	150	0	0	0	0	5
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	133	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	133	0	0	1
	R	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	59	1	6
	S	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	34	4
<b>Total</b>	<b>40</b>	<b>4</b>	<b>2</b>	<b>13</b>	<b>0</b>	<b>2</b>	<b>17</b>	<b>34</b>	<b>5</b>	<b>0</b>	<b>3</b>	<b>1</b>	<b>3</b>	<b>7</b>	<b>8</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>142</b>	

Fuente: Diseño Muestral, INEC 2020

Es importante considerar que así como existen estos cambios de dominio en las empresas seleccionadas en la muestra, existen cambios en todo el marco muestral. Con el fin de solventar estos inconvenientes se utiliza un método denominado postestratificación.

El método de postestratificación, permite realizar un ajuste a los factores de expansión mediante la incorporación de información auxiliar. En este caso, la información auxiliar utilizada proviene del DIEE 2018, desde el cual se realiza un seguimiento a las empresas de la población objetivo. Esta información se convierte en la base para realizar actualizaciones al universo de investigación y asignar el dominio de estudio al que pertenece cada una de las empresas investigadas.

De esta manera, un postestrato está conformado por las empresas que tienen un mismo tamaño de empresa y sección económica de acuerdo al DIEE 2018, fragmentando los dominios originales de la muestra.

La fórmula de cálculo del factor de expansión por empresa es:

$$fexp_{kp} = a_k \times \frac{N_p}{\hat{N}_p}$$

Dónde:

$fexp_{kp}$  = Factor de expansión de la empresa  $k$  perteneciente al postestrato  $p$ .

$\hat{N}_p$  = Tamaño estimado de la población del postestrato  $p$ .

$N_p$  = Tamaño de la población del postestrato  $p$ .

Es necesario mencionar que las empresas que tienen factor 1 corresponden a:

- Empresas grandes de acuerdo al dominio de selección de la muestra original (DIEE 2017).
- Empresas que se convirtieron en grandes de acuerdo al DIEE 2018.

### 2.4.0 Estimación de características

Debido a la postestratificación que se realiza, la estimación del Total Poblacional para determinada característica o variable  $Y$ , se obtiene a partir de la siguiente fórmula:

$$\hat{Y}_{POST} = \sum_{p=1}^P \sum_{sp} fexp_{k_p} y_k$$

Dónde:

- $\hat{Y}_{POST}$  = Estimador postestratificado de la característica  $y$ .
- $fexp_{k_p}$  = Factor de expansión de la empresa  $k$  perteneciente al postestrato  $p$ .
- $y_k$  = Característica  $y$  de la empresa  $k$ .
- $s_p$  = Empresas que pertenecen al postestrato  $p$ .

### 2.4.1 Error de muestreo

El error de muestreo cometido para la estimación de características, podrá ser calculado a partir de la estimación de la varianza del estimador del total postestratificado  $\hat{Y}_{POST}$ , cuya fórmula es la siguiente:

$$CV(\hat{Y}_{POST}) = \frac{\sqrt{\widehat{Var}(\hat{Y}_{POST})}}{\hat{Y}_{POST}}$$

Con:

$$\widehat{Var}(\hat{Y}_{POST}) = \sum_s \sum_{kl} \frac{\Delta_{kl} g_k e_k g_l e_l}{\pi_{kl} \pi_k \pi_l}$$

$$\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$$

$$\pi_k = \frac{1}{ak}$$

$$e_k = y_k - \tilde{y}_{sp}$$

$$\tilde{y}_{sp} = \frac{1}{N_p} \sum_{sp} fexp_{k_p} y_k$$

$$g_k = \frac{N_p}{\hat{N}_p}$$

Dónde:

- $CV(\hat{Y}_{POST})$  = Coeficiente de variación del estimador de total postestratificado.
- $\widehat{Var}(\hat{Y}_{POST})$  = Estimador de la varianza del estimador de total postestratificado.
- $\pi_{kl}$  = Probabilidad de inclusión conjunta de las empresas  $k$  y  $l$ .
- $\tilde{y}_{sp}$  = Estimador de la media en el postestrato  $p$ .

### 2.4.2 Declaración del Plan de muestreo

Para el cálculo de los estimadores y sus errores, varianzas, coeficientes de variación, intervalos de confianza y efecto de diseño considerando el diseño muestral, se puede utilizar



el software estadístico SPSS, en el que se construye un plan de muestreo para muestras complejas. Aunque el paquete no tiene las facilidades para realizar un diseño con postestratificación, se puede estimar con un diseño aproximado, para el cual la sintaxis del plan de muestreo es la siguiente:

```
CSPLAN ANALYSIS  
  
/PLAN FILE='C:\Users\Estructurales\plan_muestreo_final.csaplan'  
  
/PLANVARS ANALYSISWEIGHT=f_exp  
  
/PRINT PLAN  
  
/DESIGN STRATA=Pst  
  
/ESTIMATOR TYPE=EQUAL_WOR  
  
/POPSIZE VARIABLE=N_fin.
```

Dónde las variables corresponden a:

- “**Pst**”: postestrato,
- “**fexp**”: factor de expansión; y,
- “**N\_fin**”: total de la población por postestrato.

De igual manera, también se puede calcular utilizando el software Stata con las siguientes líneas de programación:

```
svyset [pweight=f_exp], strata(Pst) fpc(N_fin) vce(linearized)  
  
svyset, singleunit(certainty) noclear
```

Por último, en el software R se utiliza la siguiente sintaxis:

```
plan_muestreo <- factores_fin_V1_1 %>%  
  as_survey_design(ids = "id_empresa",  
                  strata="Pst",  
                  fpc= "N_fin",  
                  weights = f_exp,  
                  nest = T)  
options(survey.lonely.psu = "certainty")
```


## **BIBLIOGRAFÍA**

Hidiroglu, M., Särndal, C., & Binder, D. (1995). Weighting and Estimation in Business Surveys. En Cox, Binder, Chinnappa, Christianson, Colledge, & Kott, *Business Survey Methods* (págs. 477-502). Jhon Wiley & Sons.

Kish, L. (1972). *Muestreo de Encuestas*. México: Trillas.

Särndal, C., Swensson, B., & Wretman, J. (2003). *Model Assisted Survey Sampling*, Nueva York, USA. New York: Springer.


**CADA  
HECHO  
DE TU  
VIDA**  
*Cuenta*

 @ecuadorencifras

 INEC/Ecuador

 @InecEcuador

 INECEcuador

 t.me/ecuadorencifras

 INEC Ecuador