

**Control de calidad  
de la BDD del  
Directorio de  
Empresas y  
Establecimientos  
DIEE-2019**

—  
Noviembre, 2020



Buenas cifras,  
**mejores vidas**

<b>Contenido</b>	
<b>Introducción</b> .....	3
<b>Objetivo</b> .....	3
<b>Objetivos específicos</b> .....	3
<b>1. Consideraciones preliminares para el cálculo de los Indicadores</b> .....	3
1.1. <b>Primera Fase</b> .....	3
1.2. <b>Segunda Fase</b> .....	4
1.3. <b>Cálculo y selección de la muestra</b> .....	5
1.3.1. <b>Verificación de la información en Call Center</b> .....	6
<b>2. Resultados</b> .....	6
2.1. <b>Indicadores de la primera fase:</b> .....	6
2.1.1. <b>Indicadores por cobertura de identificación</b> .....	6
2.1.2. <b>Indicadores por cobertura de Estratificación</b> .....	9
2.1.3. <b>Indicadores por variables demográficas</b> .....	12
2.1.4. <b>Indicadores por cobertura de imputación</b> .....	12
2.2. <b>Indicadores de la segunda fase:</b> .....	14
2.2.1. <b>Indicadores por cobertura de identificación</b> .....	14
2.2.2. <b>Indicadores por cobertura de Estratificación</b> .....	15

## Introducción

El Directorio de Empresas y Establecimientos (DIEE) es un producto estadístico importante del Instituto Nacional de Estadística y Censos (INEC), ya que, en su función de proveedor oficial de las estadísticas económicas para Ecuador, pone a disposición esta herramienta clave, que proporciona información de las empresas con una cobertura casi completa del sector formal, para obtener estadísticas para la toma de decisiones de política económica.

El DIEE una operación estadística basada en registros administrativos cuyas principales fuentes de información proviene del Instituto Ecuatoriano de Seguridad Social (IESS), del Servicio detentas Internas (SRI) y de diferentes encuestas a empresas que el INEC realiza. Con ello se produce información estadística actualizada, para satisfacer requerimientos sobre marcos muestrales para la mayoría de las encuestas dirigidas a empresas; directorio que se compone con unidades económicas públicas o privadas con personería jurídica propia o sin ella, que se encuentran dentro del territorio nacional y que generaron alguna actividad económica en el periodo de estudio.

Con este antecedente y dado los resultados que se obtuvieron para la publicación del DIEE-2019, el presente documento va enfocado a obtener indicadores de calidad que permitan medir su cobertura y calidad.

## Objetivo

Calcular indicadores que permitan medir la cobertura del Directorio de Empresas y Establecimientos que publica la unidad de Gestión de Estadísticas Económicas en Base a Registros Administrativos.

## Objetivos específicos

- Calcular indicadores por cobertura de identificación.
- Calcular indicadores por cobertura de estratificación.
- Calcular indicadores por cobertura de imputación.
- Calcular indicadores por variables demográficas.

## 1. Consideraciones preliminares para el cálculo de los Indicadores

Para el cálculo de los indicadores antes mencionados se utilizará la base de datos completa final (datos para publicación) DIEE-2019, la obtención de los indicadores se ejecutará en dos fases:

### 1.1. Primera Fase

En esta fase se procesa la **BDD completa** y se calcula los siguientes indicadores:

- Porcentaje de empresas del DIEE que disponen detalle de dirección completa.
- Porcentaje de empresas que se actualizó detalle de dirección.
- Porcentaje de empresas que disponen información de actividad económica a 4 y 6 dígitos de la CIU.

- Porcentaje de actividad económica a 4 y 6 dígitos de la CIIU que fueron actualizadas.
- Porcentaje de empresas que disponen de información de ubicación geográfica.
- Porcentaje de empresas que se actualizó información de ubicación geográfica.
- Porcentaje de empresas falsas activas.
- Porcentaje de empresas imputadas para definir tamaño por ventas.
- Porcentaje de empresas imputadas para definir tamaño por plazas de empleo registrado y remuneraciones.

Las variables de la BDD esenciales que intervienen en esta fase son: Dirección, Código de actividad económica a 4 y 6 dígitos de la CIIU. Ubicación de geografía, ventas, empleo y remuneraciones.

## 1.2. Segunda Fase

En esta fase se selecciona de la BDD DIEE2019 las empresas definidas en los tamaños: **grandes, medianas A y B**; se realiza esta selección dado que no se dispone de tiempo y recursos para analizar todos los tamaños de empresas. De los tamaños de empresas seleccionados se calcula una muestra, la cual permitirá obtener los siguientes indicadores:

- Porcentaje de empresas que presentaron error de dirección.
- Porcentaje de empresas que presentaron error de codificación.
- Porcentaje de empresas que presentaron error de ubicación geográfica.

Para gestionar la muestra seleccionada de los tamaños de empresas antes descritas, se utilizará el sistema Call Center, en el cual se verificará el siguiente listado de variables:

- Identificador de empresa, RUC, razón social y nombre comercial (esenciales para llevar un control de las empresas a revisar).
- Código CIIU.
- Descripción de la Actividad Económica.
- Ubicación Geográfica.
- Variables de Dirección.
- Variables de contacto.

Para ello se establece las siguientes directrices:

- Investigar si la empresa estuvo activa en el año 2019.
- Investigar la actividad económica de la empresa en el año 2019.
- Investigar la ubicación geográfica a nivel de provincia, cantón, parroquia y dirección en el año 2019.

Esta actividad será realizada por el equipo verificador de GEERA y lo hará directamente con la empresa a través de llamada telefónica para investigar la información requerida. Únicamente en caso de que no se logre ubicar a la empresa por la vía antes descrita, deberá buscar la información mediante páginas WEB, de preferencia en la página de la empresa si es que esta la dispone.

### 1.3. Cálculo y selección de la muestra

La muestra se calcula del universo de empresas obtenidas de la BDD total final del DIEE-2019 de los tamaños: Grande, medianas A y B; por lo tanto esta será la población objetivo. La muestra a calcular debe reproducir las características del universo o población; por ello, la muestra debe ser representativa de esta, al considerar un tamaño de muestra suficiente, un nivel de confianza seguro y un máximo de error permitido; con estas características definidas se puede argumentar que la muestra obtenida es representativa de la población objetivo.

Para aplicar las características descritas, se utiliza la fórmula para muestreo aleatorio simple:

$$n = \frac{N * Z_{\alpha}^2 * \rho * q}{(N - 1) * e^2 + Z_{\alpha}^2 * \rho * q}$$

Donde:

n = Tamaño de la muestra

N = Total poblacional (sumatoria de los tres estratos)

$Z_{\alpha}$  = Variables de estandarización dado para el nivel de confianza

(Valor medio para lograr el nivel de confianza deseado, distribución de Gauss)

$\rho$  = Proporción esperada en el estudio (éxito)

q = Proporción no esperada en el estudio (fracaso)

e = Margen de error máximo admitido en el estudio

Para el estimador “p” se considera el porcentaje de 0.5, y para la probabilidad de fracaso de “q” de 0.5; con un nivel de confianza del 95% (según tablas estadísticas)  $Z_{\alpha}=1.96$ , y con el fin de construir indicadores precisos el error “e” será igual a 3,17%.

Bajo lo expuesto se obtuvo una muestra de 752 empresas. Dado que se cuenta con tamaños de empresa, la muestra obtenida se reparte para cada estrato; este tipo de muestreo es una extensión del muestreo aleatorio simple, debido que en primer lugar se calcula el tamaño total de la muestra y posterior a ello y función al número de elementos con el que está compuesto cada estrato se redistribuye la muestra; esto último se lo conoce como afijación proporcional tal como se muestra en la tabla 2. Para ello se utiliza la siguiente formulación:

$$n_i = \left( \frac{N_i}{N} * n \right) * 100$$

**Tabla 1** Tamaño de la muestra

Tamaño Empresas	Total, Universo	Muestra	Porcentaje frente al total	Porcentaje acumulado muestra
Mediana empresa "A"	8.544	345	45,80%	45,80%
Mediana empresa "B"	5.798	234	31,08%	76,88%
Grande empresa	4.312	174	23,12%	100,00%
<b>Total, muestra</b>	<b>18.654</b>	<b>752</b>	<b>100%</b>	

Fuente: BDD del DIEE-2019.

Elaboración: GEERA.

### 1.3.1. Verificación de la información en Call Center

El tiempo que utilizará el verificador para realizar esta actividad, se calculó con una piloto inicial, en la cual se determinó que el verificador debe emplear 15 minutos por empresa; tiempo en el que se consideró varios factores negativos como: no todas las empresas cuentan con números de teléfonos fijos correctos; tampoco todas las empresas disponen de página WEB; aun cuando se verifica que si existe información en las fuentes el SRI, SUPERCAS y diferentes páginas WEB, estas difieren de la información que se registra en la base de datos del DIEE; adicional a ello está la búsqueda de códigos CIU para actividades económicas muy específicas.

Si se emplea el 100% de tiempo a la actividad de verificar y actualizar la información en el sistema Call Center, en promedio el verificador gestionará 32 registros diarios. A continuación, se describe el tiempo a ser utilizado:

**Tabla 2** Distribución del tiempo

Tamaño Empresas	Muestra	Carga diaria* 90%	Tiempo en días al 90%	% de avance 90%	Carga* diaria 70%	Tiempo en días al 70%	% de avance 70%
Mediana empresa "A"	345	29	5,97	22,87%	22	7,68	22,88%
Mediana empresa "B"	234	29	4,06	38,41%	22	5,22	38,42%
Grande empresa	174	29	3,02	49,97%	22	3,88	49,97%
<b>Total muestra</b>	<b>752</b>		13,05			16,78	

Fuente: BDD del DIEE-2019

Elaboración: GEERA.

Nota\*: Tiempo asignado en función de las actividades de cada revisor.

Para realizar la actividad de verificación de la muestra obtenida, se debe utilizar un total de 17 días considerando el 70% de tiempo asignado al segundo verificador.

## 2. Resultados

El cálculo de los indicadores del control de calidad de la BDD del DIEE-2019, se realizó bajo cuatro categorías de variables: identificación, estratificación, demografía y otras variables como las imputaciones; a continuación, el detalle de los indicadores obtenidos:

### 2.1. Indicadores de la primera fase:

#### 2.1.1. Indicadores por cobertura de identificación

##### ✓ Porcentaje de empresas que disponen detalle de dirección completa.

Para calcular este indicador se considera las variables que intervienen en la estructura de identificación, tal como se lista en la siguiente tabla:

**Tabla 3** Estructura para definir registros con dirección completa.

No. de Bloque	Estructura de Dirección completa	Bloque selección	Condición por bloque	Condición general
1	Calle principal	1	>=1	>=3
	Intersección	1		
2	Numero	2,3,4, 5 y 6	>=1	
	Kilometro			
	Urbanización			
	Nombre edificio			
	Número piso			
3	Número oficina	2,3 y 4	>=1	
	Ciudadela			
	Barrio			
	Referencia			
	Manzana			

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

Aquellos registros que no cumplieron la condición general se determinaron como direcciones no válidas, los bloques de selección se establecen en función de:

- Si no existe ningún dato en los campos de dirección (calle principal y secundaria).
- Si solo existe 1 dato en cualquiera de los campos: número, kilómetro, urbanización, nombre edificio, número piso y número oficina.
- Si solo existe un dato en los campos: ciudadela, barrio, referencia y manzana

Si no se cumple la condición en cada uno de los bloques se determinó como dirección no válida. Para calcular cada uno de los bloques se utilizó el programa SPSS con la función recodificar variables en nuevas variables con ello se calculará las nuevas variables. Posterior a ello se empleó la siguiente formulación, la cual se aplicó tanto al universo de la población como a los estratos por tamaño de empresas:

$$PEDC_t = \frac{TESDC_t}{TEA_t} * 100$$

Donde:

**PEDC<sub>t</sub>**= Porcentaje de empresas con dirección completa en el año de referencia.

**TESDC<sub>t</sub>**= Total empresas sin dirección completa en el año de referencia.

**TEA<sub>t</sub>**= Total empresas activas en el año de referencia.

### Resultados:

$$PEDC_t = \frac{736.410}{882.766} * 100$$

$$PEDC_t = 83,42\%$$

Se puede evidenciar que el 83,42% de empresas disponen de dirección para su ubicación tal como se muestra en la siguiente tabla.

**Tabla 4** total empresas que disponen o no de información de dirección.

Detalle de dirección completa	# de empresas	%
Dirección vacía	5.621	0,64%
Dirección en al menos dos de las variables	140.735	15,94%
Dirección completa	736.410	83,42%
<b>Total</b>	<b>882.766</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

#### ✓ Porcentaje de empresas que se actualizó detalle de dirección.

Para el cálculo de este indicador se realizó un conteo de todas las empresas que se actualizaron dirección completa (fecha\_registro\_direccion), en función de las diferentes fuentes como: IND=Industriales, CALL=Cal Center, SRI= Servicio de Rentas Internas, entre otras y que correspondan al año de referencia. Se aplicó la siguiente fórmula tanto para el universo como para los estratos de empresas.

$$PEAD = \frac{TESAD}{TEA} * 100$$

Donde:

PEAD= Porcentaje de empresas que se actualizaron dirección completa en el año de referencia.

TESAD= Total empresas sin actualizar dirección completa en el año de referencia.

TEA= Total empresas activas en el año de referencia.

#### Resultados:

$$PEAD = \frac{154.909}{882.766} * 100$$

$$PEAD = 17,55\%$$

La base del DIEE 2019 presenta una actualización de la variable dirección del 17,55% del directorio total. De este total general se analiza a las empresas grandes, medianas A y B, con lo cual se puede observar que: este grupo de empresas se actualizó en un 19,85%, de ello, el mayor porcentaje de actualización de detalle de dirección corresponde a las empresas Mediana A con el 21,43%, tal como se muestra en la siguiente tabla:

**Tabla 5** Actualización de detalle de dirección, según tamaño de empresa.

Actualización de detalle de dirección	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%	Total	%
SI	835	<b>19,36%</b>	1.036	<b>17,87%</b>	1.831	<b>21,43%</b>	<b>3.702</b>	<b>19,85%</b>
NO	3.477	<b>80,64%</b>	4.762	<b>82,13%</b>	6.713	<b>78,57%</b>	<b>14.952</b>	<b>80,15%</b>
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>	<b>18.654</b>	<b>100,00%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

## 2.1.2. Indicadores por cobertura de Estratificación

- ✓ **Porcentaje de empresas que disponen información de actividad económica a 4 y 6 dígitos de la CIIU Rev. 4.0.**

$$PAE_i = \frac{TESAE_t}{TEA_t} * 100$$

**Donde:**

**PAE<sub>i</sub>**= Porcentaje de empresas que disponen código de actividad económica de la CIIU 4.0 a nivel de 6 dígitos.

**TESAE<sub>i</sub>**= Total empresas sin actividad económica en el año de estudio.

**TEA<sub>i</sub>**= Total empresas activas en el año de estudio.

### Resultados:

El 100% de las actividades registradas en la BDD del DIEE-2019 dispone de códigos a nivel de 4 dígitos de la CIIU Rev. 4.0.

Para los códigos a nivel de 6 dígitos de la CIIU Rev. 4.0 se indica que:

$$PAE = \frac{547}{882.766} * 100$$

$$PAE = 0,06\%$$

El 0,06% de empresas no disponen de código de actividad a 6 dígitos de la CIIU Rev. 4.0 (incluye todos los tamaños de empresas). Si separamos las empresas que no dispone de código de actividad económica en: grupo 1 empresas Grandes y medianas A y B y grupo 2 las demás (Pequeña empresa y microempresa), se puede verificar que el primer grupo aporta apenas con el 0,37% y las demás con el 99,63%.

**Tabla 6** Empresas por tamaño que no dispone de actividad económica a 6 dígitos de CIIU Rev. 4.0

Tamaño de empresas	# de empresas	%
Grupo 1 (Empresas grandes, mediana A y B)	2	0,37%
Grupo 2 (Pequeña empresa y microempresa)	545	99,63%
Total	547	100%

**Fuente:** Base de datos DIEE 2019.

**Elaboración:** GEERA.

Ahora bien, de este porcentaje (0.37%) que no disponen código de actividad, lo representa las empresas de tamaño "Empresas Mediana A", tal como se muestra en la siguiente tabla:

**Tabla 7** Distribución de empresas que no disponen de código de actividad económica a 6 dígitos de la CIU Rev. 4.0 por tamaño de empresa.

Dispone de actividad económica	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%	Total	%
SI	4312	100,00%	5798	100,00%	8542	99,98%	18652	99,99%
NO	0	0,00%	0	0,00%	2	0,02%	2	0,01%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

✓ **Porcentaje de empresas a nivel de 4 y 6 dígitos de la CIU Rev. 4.0 que fueron actualizadas.**

Para calcular este indicador se toma la variable "fecha\_registro\_actividad", cuyos registros tengan año de actualización 2019.

$$PAE_t = \frac{TEAE_t}{TEA_t} * 100$$

Donde:

**PAE<sub>t</sub>**= Porcentaje de empresas que se actualizó actividad económica a nivel de 4 y 6 dígitos de la CIU 4.0.

**TEAE<sub>t</sub>**= Total empresas que se actualizó actividad económica a nivel de 4 y 6 dígitos de la CIU 4.0.

**TEA<sub>t</sub>**= Total empresas activas en el año de estudio.

**Resultados:**

$$PAE = \frac{146.111}{882.766} * 100$$

$$PAE = 16,55\%$$

Para el año de referencia el Directorio de Empresas y Establecimientos actualizó la variable actividad económica en un 16,55%. De este porcentaje a continuación, se selecciona a las empresas según su tamaño: grandes, mediana A y B, se puede observar que el mayor porcentaje de actualización lo presentan las empresas "mediana A" con un 10,70%.

**Tabla 8** Distribución de empresas por tamaño que se actualizó código de actividad económica a nivel de 6 dígitos de la CIU 4.0.

Se actualizó Actividad Económica	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%	Total	%
SI	176	4,08%	313	5,40%	914	10,70%	1403	7,52%
No	4.136	95,92%	5.485	94,60%	7.630	89,30%	17.251	92,48%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

✓ **Porcentaje de empresas que dispone información de ubicación geográfica.**

Para este indicador se toma las variables "codigo\_parroquia", con ello se puede

verificar que todos los registros dispongan de geografía y se aplica la siguiente formulación:

$$PEUG_t = \frac{TESUG_t}{TEA_t} * 100$$

**Donde:**

PEUG= Porcentaje de empresas que disponen ubicación geográfica en el periodo de referencia.

TESUG= Total empresas sin ubicación geográfica en el periodo de referencia.

TEA= Total empresas activas en el periodo de referencia.

**Resultados:**

$$PEUG = \frac{882.766}{882.766} * 100$$

$$PEUG = 100\%$$

Se puede verificar que el 100% de los registros disponen de ubicación de geografía. Debido a ello los estratos de empresas grande, mediana A y B también disponen de ubicación de geografía al 100%.

✓ Porcentaje de empresas que se actualizó información de ubicación geográfica.

$$PEAUG_t = \frac{TEAUG_t}{TEA_t} * 100$$

**Donde:**

PEAUG<sub>t</sub>= Porcentaje de empresas que se actualizó ubicación geográfica en el periodo de referencia.

TEAUG<sub>t</sub>= Total empresas que se actualizó ubicación geográfica en el periodo de referencia.

TEA<sub>t</sub>= Total empresas activas en el periodo de referencia.

**Resultados:**

$$PEAUG = \frac{94.845}{882.766} * 100$$

$$PEAUG = 10,74\%$$

El 10,74% de registros de empresas se actualizó información de ubicación geográfica; este resultado se obtiene a partir de la variable "registro\_fecha\_geografia".

Seleccionando sólo las empresas grandes, medianas A y B, se verifica que en la base de datos de empresas del DIEE 2019, se actualizó información de ubicación

geográfica en un 18,35% y el mayor porcentaje lo representa las empresas grandes 19,13%.

**Tabla 9.** Empresas por tamaño que se actualizó o no información de ubicación geográfica.

Actualización Ubicación Geográfica	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%	Total	%
Si	825	19,13%	986	17,01%	1612	18,87%	3423	18,35%
No	3.487	80,87%	4.812	82,99%	6.932	81,13%	15.231	81,65%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

### 2.1.3. Indicadores por variables demográficas

#### ✓ Porcentaje de empresas falsas activas.

Las empresas falsas activas hacen referencia a empresas que no registran ventas, empleo o pertenecen al RISE y en el año de referencia se encuentran activas.

Se puede verificar que a nivel de estratos no se presentan registros que contengan con las categorías antes descritas, a continuación, tabla de desglose:

**Tabla 10** Porcentaje de empresas falsas activas por tamaño.

Situación 2019 (falsas activas)	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%	Total	%
SI	0	0,00%	0	0,00%	0	0,00%	0	0,00%
NO	4.312	100,00%	5.798	100,00%	8.544	100,00%	18.654	100,00%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

### 2.1.4. Indicadores por cobertura de imputación

Las imputaciones son procedimientos que permiten asignar información de variables que por algún motivo no se dispone en el período de análisis. Para el caso del DIEE, al ser una operación estadística basada en registros administrativos, se realiza imputaciones a aquellas empresas que no disponen datos de ventas para el año de publicación, debido a ello se considera datos de años anteriores, se tomó información del INAR, considerando el valor mensual y obteniendo la suma total para el año. Adicional a ello también se utiliza la información de la encuesta industriales, del SIPRO y las que deriven del análisis del equipo técnico de la DIEE, en ello están las siguientes empresas: "EMPRESA PUBLICA DE EXPLORACION Y EXPLOTACION DE HIDROCARBUROS PETROAMAZONAS EP" y "EMPRESA PUBLICA DE HIDROCARBUROS DEL ECUADOR EP PETROECUADOR".

#### ✓ Porcentaje de empresas imputadas para definir tamaño por ventas

$$PIM_t = \frac{TRIM_t * 100}{TEA_t}$$

**Donde:**

**PIM<sub>i</sub>**= Porcentaje de registros imputados en el periodo de referencia.

**TRIM<sub>i</sub>**= Total de registros imputados por ventas en el periodo de referencia.

**TEA<sub>i</sub>**= Total registros activos en el periodo de referencia.

**Resultados:**

$$PIM = \frac{2}{882.766} * 100$$

$$PIM = 0.0002\%$$

Para el DIEE\_2019 se imputo un minúsculo número de empresas para definir tamaño de empresas por ventas

✓ **Porcentaje de empresas imputadas para definir tamaño por empleo y remuneración**

Para el caso de las variables empleo y remuneraciones se imputan aquellas empresas que corresponden a Educación y Administración Pública.

Para ello se considera a cada una de las Direcciones Distritales con las Instituciones que se encuentran adscritas a cada una de ellas, procediendo a agruparlas para identificar el empleo mensual que presentaron las instituciones educativas hasta el año 2014 para relacionarlo con el empleo mensual que muestra la Dirección Distrital que la contiene. A partir del año 2015 la cantidad de empleo que muestra la Dirección Distrital se distribuye para cada una de las instituciones que se encuentran a su cargo. Para ello se coloca como valor de empleo, el último que declaró cada institución en el 2014 y se resta la cantidad de empleados que tiene la Dirección Distrital.

Por otro lado, la imputación de remuneraciones se sigue la misma metodología que la de empleo, que para este caso es el valor total de salarios que está asignado a la Dirección Distrital y se distribuye para las instituciones educativas adscritas a la misma.

En función a ello se obtiene los registros de empresas en los cuales se cumple las imputaciones descritas y se aplicará la siguiente formulación:

$$PIM = \frac{TRIM}{TEA} * 100$$

Donde:

PIM= Porcentaje de registros imputados en el periodo de referencia.

TRIM= Total de registros imputados por plazas y/o empleo registrado y remuneraciones en el periodo de referencia.

TEA= Total registros activos en el periodo de referencia.

## Resultados:

$$PIM = \frac{13.250}{882.766} * 100$$

$$PIM = 1.50\%$$

El porcentaje de empresas imputadas para definir tamaño de las empresas por plazas de empleo registrado, empleo registrado y remuneraciones es del 1.50%, que representa un total de 13.250 empresas.

## 2.2. Indicadores de la segunda fase:

Resultados de la verificación en el sistema CALL CENTER.

De la muestra (752) obtenida para verificar en el sistema Call Center se diligenció toda la muestra. De ello 99,07% de registros fueron terminados; lo restante 0,93% corresponde a registros en estado gestionadas, que corresponde a información que no se pudo corroborar alguna de las variables como: ubicación o actividad económica válida, por ninguno de los medios establecidos en las directrices; con estos resultados se extrapoló al universo de los estratos de estudio definidos al inicio de este documento y que corresponde al DIEE2019.

**Tabla 11** Gestión de la muestra

Tamaño Empresas	Total Universo	Muestra	Verificadas (terminado)	Porcentaje frente al total
Mediana empresa "A"	8.544	336	333	99,11%
Mediana empresa "B"	5.798	242	238	98,35%
Grande empresa	4.312	174	174	100,00%
<b>Total muestra</b>	<b>18.654</b>	<b>752</b>	<b>745</b>	<b>99,07%</b>

Fuente: Muestra Call Center DIEE 2019.

Elaboración: GEERA.

### 2.2.1. Indicadores por cobertura de identificación

#### ✓ Qué porcentaje de empresas presentan error de dirección\*.

#### Resultados:

Aproximadamente 1.901 empresas presentaron error en la dirección, tal como se observa a continuación:

**Tabla 12** Porcentaje de empresas que presentaron error de dirección.

Error dirección	Total empresa	%
SI	1.901	10,19%
NO	16.753	89,81%
<b>Total</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

De las empresas verificadas el 90,87% se ubica en la misma dirección, mientras que el 9,13% de las empresas registra una ubicación diferente al de la base de datos DIEE-2019.

Ahora bien, analizando el indicador por tamaño de empresa resulta que el porcentaje obtenido se distribuye en función del peso que presenta cada tamaño de empresa, en función a ello el mayor porcentaje de error en la dirección es de 9,68% en las empresas Mediana A , tal como se observa en la siguiente tabla:

**Tabla 13** Distribución de empresas que presentaron error de dirección por tamaño de empresa.

Error dirección	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%
SI	347	8,05%	529	9,13%	827	9,68%
NO	3.965	91,95%	5.269	90,87%	7.717	90,32%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

## 2.2.2. Indicadores por cobertura de Estratificación

El cálculo se lo realiza tanto para la población como para los estratos de estudio.

### ✓ Porcentaje de empresas que presentaron error de codificación

Para calcular este indicador se toma la variable cambio de sector económico de los registros verificados en el sistema Call Center.

Según la verificación realizada, se tiene que 301 empresas presentaron error en la codificación de la actividad económica, que representa el 1,61%, tal como se muestra a continuación:

**Tabla 14** Porcentaje de empresas que presentaron error de codificación.

Error codificación	Total empresa	%
SI	301	1,61%
NO	18.353	98,39%
<b>Total</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

A continuación, se desglosa el error de codificación en función del tamaño de las empresas:

**Tabla 15** Porcentaje de empresas que presentaron error de codificación Actividad Económica.

Error codificación	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%
SI	25	0,57%	76	1,30%	200	2,35%
NO	4.287	99,43%	5.722	98,70%	8.344	97,65%
<b>Total</b>	<b>4.312</b>	<b>100%</b>	<b>5.798</b>	<b>100%</b>	<b>8.544</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

Se puede advertir que las empresas con mayor error de codificación, son las empresas

definidas como: empresas "Mediana A" con el 2,35%.

✓ **Porcentaje de empresas presentan error en la ubicación geográfica**

Luego de la verificación en el CALL CENTER, se tiene que: El 0,67% de empresas presentaron error en la ubicación geográfica lo que representa 125 registros.

**Tabla 16.** Empresas por tamaño que presentaron error de ubicación geográfica.

Error geografía	Total empresa	%
SI	125	0,67%
NO	18.529	99,33%
<b>Total</b>	<b>18.654</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

De las empresas verificadas aproximadamente el mayor porcentaje de error lo representa las empresas Mediana A con el 0,86, tal como se muestra en la siguiente tabla:

**Tabla 17** Porcentaje de empresas por tamaño que presentaron error en ubicación geográfica.

Error geografía	Empresas Grandes	%	Empresas Mediana B	%	Empresas Mediana A	%
SI	50	0,58%	25	0,59%	50	0,86%
NO	8.494	99,42%	4.262	99,41%	5.773	99,14%
<b>Total</b>	<b>8.543</b>	<b>100%</b>	<b>4.288</b>	<b>100%</b>	<b>5.823</b>	<b>100%</b>

Fuente: Base de datos DIEE 2019.

Elaboración: GEERA.

<b>Elaborado por:</b>	Rocío Andino	Firma:
<b>Revisado y aprobado por:</b>	Libertad Trujillo	Firma:
<b>Fecha de aprobación:</b>	30/11/2020	