

The background features a light gray grid of cubes. In the top left corner, there is a cluster of faint, overlapping numbers and symbols. The main title is positioned in the upper right quadrant, and the subtitle is below it. The INEC logo is at the bottom center.

# **GUÍA DE USO DE LAS BASES DE DATOS**

**CIAEGP 2015**

**Guía para el manejo de las bases de datos de la Operación Estadística “Censo de Información Ambiental Económica en GAD Provinciales” 2015**

**Dirección responsable de la operación estadística:**

**DIRECCIÓN DE ESTADÍSTICAS AGROPECUARIAS Y AMBIENTALES**

**Responsable**

**Carlos Pilataxi**

## Contenido

1. ANTECEDENTES .....	4
2. ESTRUCTURA DE LAS BASES DE DATOS DE LA OPERACIÓN ESTADÍSTICA .....	4
3. DESARROLLO DEL ANÁLISIS DE PERTINENCIA DE LAS NECESIDADES DE INFORMACIÓN.....	8

## 1. ANTECEDENTES

---

La Dirección de Estadísticas Agropecuarias y Ambientales presenta anualmente el Censo de Información Ambiental Económica en GAD Provinciales, el cual investiga varias temáticas: gestión ambiental, personal empleado, fomento y desarrollo productivo, riego y drenaje, ingresos y gastos, en la unidad de observación (consejos provinciales) con el fin de proporcionar al país (autoridades, investigadores, y público en general), una herramienta real y actualizada para la toma de decisiones, el planteamiento y la ejecución de proyectos en el área ambiental.

Para el usuario que requiera un acceso a las fuentes primarias de información (bases de datos), se requiere de ciertas destrezas en el procesamiento de bases de datos estadísticas, particularmente si éstas vienen en el formato IBM SPSS® Statistics, versión 18 o posterior. Adicional a ello, también es importante contar con una guía que permita manejar las bases de datos de esta operación estadística en forma fácil, rápida y precisa.

Este documento presenta algunas sugerencias para lograr ese objetivo. Ciertamente, esta guía no tiene el alcance de un manual de programación de bases de datos estadísticas. Sin embargo, cubre los aspectos básicos e indispensables a tomar en cuenta antes de procesar las bases de datos del Censo de Información Ambiental Económica en GAD Provinciales 2015.

## 2. ESTRUCTURA DE LAS BASES DE DATOS DE LA OPERACIÓN ESTADÍSTICA

---

La presente operación estadística cuenta con un conjunto de 10 bases de datos en formato SPSS®, las cuales son:

- **Gobierno:** Es la base más extensa de la operación estadística. Consta de 256 variables por 24 registros o casos (uno por cada GAD provincial), de las cuales 6 son de cabecera o variables de identificación de cada GAD provincial. Esta base contiene todas las variables que en el formulario físico aparecen como variables simples –como las variables de Sí / No o variables sueltas que registran opciones múltiples o valores numéricos abiertos–. En cambio, no se encuentran aquí las variables que en el formulario aparecen como tablas cruzadas –una variable para filas y dos o más en las columnas (ejemplo, la Pregunta 2.2)–, tablas anidadas –listados de valores de variables dentro de filas o columnas de tablas cruzadas (ejemplo, la Pregunta 2.10, variables 2.10.3 y 2.10.4). Este tipo de

- variables aparecen justamente en bases de datos separadas cuyos nombres siempre harán referencia a la pregunta o familia de preguntas<sup>1</sup> que contiene las respectivas tablas cruzadas y/o tablas anidadas.
- **2.2\_3.2\_4.2\_Personal:** Es la base de datos que contiene las variables de las preguntas 2.2, 3.2 y 4.2. Consta de 9 variables por 360 registros o casos (15 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. Las variables tienen un nombre que comienza con el prefijo **vX2**, dependiendo de si el contenido de la variable será para la pregunta 2.2, 3.2 o 4.2. De hecho, la variable de segmentación (o selectora) de la tabla es la variable **vX2**, en donde X es un índice que vale 1 (pregunta 2.2), 2 (pregunta 3.2) o 3 (pregunta 4.2). Luego, la variable **vX2\_cargo** tiene 5 grupos de categorías, donde cada grupo representa en realidad la misma instancia de cargo / perfil del personal que aparece en cualquiera de las tablas 2.2, 3.2 o 4.2. Así, las categorías {11, 12, 13} representan a las respuestas de la categoría de cargo / perfil “a. Director, jefe o su equivalente”, en tanto que las categorías {21, 22, 23} representan a la categoría de cargo / perfil “b. Personal Administrativo / Financiero”, y así sucesivamente. Si se desea obtener resultados únicamente para una de las tablas 2.2, 3.2 o 4.4, se debe segmentar la base a partir de la variable **vX2**. Las demás variables tienen sufijos que representan al número de la columna etiquetada en el formulario. Por ejemplo, la variable **vX26** representa a la variable 2.2.6, 3.2.6 o 4.2.6, dependiendo del valor de la variable selectora **vX2**.
  - **2.10\_ViverosDetalles:** Es la base de datos que contiene las variables de las preguntas 2.2, 3.2 y 4.2. Consta de 4 variables por 284 registros o casos, de las cuales una es variable de cabecera o de identificación del GAD provincial. Esta base contiene las variables de las preguntas 2.10.3 (Especies existentes) y 2.10.4 (Número de plantas / año). El resto de las variables de la Pregunta 2.10 vienen incorporadas en la base Gobierno, porque son variables simples que no requieren de multiplicidad de registros por GAD provincial. La variable selectora de tipo de vivero es **v210\_tip\_esp**, la cual tiene los valores a (Viveros con especies nativas), b (Viveros con especies introducidas / adaptadas), c (Viveros con ambas especies nativas e introducidas – Mixtos).
  - **2.13\_Programas:** Es la base de datos que contiene las variables de la pregunta 2.13. Consta de 8 variables por 168 registros o casos (7 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. La base tiene la variable **v213\_recurso** desglosa los recursos naturales para los cuales se han desarrollado planes, programas y/o proyectos. Las demás variables representan a las columnas de los tipos de planes, programas y/o proyectos, como son


<sup>1</sup> Existen bases de datos que contienen variables que son idénticas en significado y contenido en varias tablas cruzadas, debido a que básicamente la misma tabla cruzada se repite en varios capítulos. El mejor ejemplo es el de la base 2.2\_3.2\_4.2\_Personal, la cual contiene las variables de las preguntas 2.2 (personal para la gestión ambiental), 3.2 (personal para fomento y desarrollo productivo) y 4.2 (personal para manejo de riego y drenaje). Más adelante se detallará la estructura de esta base de datos.

capacitación y manejo (**v213\_capac**), conservación y protección (**v213\_conser**), investigación (**v213\_invest**), fomento (**v213\_fomen**), industrialización (**v213\_indust**) y administración (**v213\_admin**).

- **3.3\_SectorProductivo:** Es la base de datos que contiene las variables de la pregunta 3.3. Consta de 6 variables por 264 registros o casos (11 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. La variable **v33\_sec\_pro** es una variable selectora de los sectores productivos en donde el GAD invirtió en el 2015. Sus valores son 1 (“a. Agricultura”), 2 (“b. Pesca”), 3 (“c. Ganadería”), 4 (“d. Acuicultura”),..., 10 (“j. Economía solidaria”), 11 (“k. Otro...especifique”). Las demás variables corresponden a las 4 columnas de la Pregunta 3.3.
- **3.4\_3.5:** Es la base de datos que contiene las variables de las preguntas 3.4 y 3.5. Consta de 12 variables por 360 registros o casos (15 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. Cabe señalar que las Preguntas 3.4 y 3.5 son tablas cruzadas que comparten la misma estructura de variables de columna. Para segmentar la base y obtener resultados asociados a c/u de las Preguntas 3.4 y 3.5, se requiere usar la variable **v34\_35**, que es igual a 1 (Pregunta 3.4) o 2 (Pregunta 3.5). Las demás variables se corresponden con las columnas de ambas tablas.
- **4.7\_Riego:** Es la base de datos que contiene las variables de la pregunta 4.7. Consta de 10 variables por 96 registros o casos (4 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. La base contiene todas las columnas asociadas al total de número de sistemas (**v47\_total**), número de sistemas de riego superficial (**v471**), hectáreas cubiertas por de riego superficial (**v472**), hectáreas efectivamente regadas por de riego superficial (**v473**), número de sistemas de riego presurizado (**v474**), hectáreas cubiertas por de riego presurizado (**v475**), hectáreas efectivamente regadas por de riego presurizado (**v475**). Existe una variable selectora del modelo de gestión del riesgo, **v47\_modelo**, la cual elige de entre los 4 posibles modelos listados en la Pregunta 4.7
- **4.9\_4.10\_4.11\_RiegoProyectos:** Es la base de datos que contiene las variables de la pregunta 3.3. Consta de 9 variables por 288 registros o casos (12 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. Esta base ha sido construida bajo la misma lógica que la base **2.2\_3.2\_4.2\_Personal**. Existe una variable selectora de la pregunta, **v4X**, la cual vale 1 (Pregunta 4.9), 2 (Pregunta 4.10) y 3 (Pregunta 4.11). Existe una variable desagregadora de tipo de proyectos (de riego, de drenaje o mixtos), **v4X\_tipo**, la cual tiene 4 grupos de categorías, donde cada grupo representa en realidad la misma instancia de tipo de proyectos que aparece en cualquiera de las tablas 4.9, 4.10 o 4.11. Así, las categorías {11, 12, 13} representan a las respuestas

de la categoría de tipo de proyecto “a. Estudios de factibilidad de Sistemas de Riego”, en tanto que las categorías {21, 22, 23} representan a la categoría de cargo / perfil “b. Nuevos sistemas de riego”, y así sucesivamente. Si se desea obtener resultados únicamente para una de las tablas 4.9, 4.10 o 4.11, se debe segmentar la base a partir de la variable **v4X**. Las demás variables tienen sufijos que representan al número de la columna etiquetada en el formulario. Por ejemplo, la variable **v4X1** representa a las variables 4.9.1, 4.10.1 o 4.11.1, dependiendo del valor de la variable selectora **v4X**.

- **4.12\_MecanismosArticulacion:** Es la base de datos que contiene las variables de la pregunta 3.3. Consta de 6 variables por 244 registros o casos, de las cuales una es variable de cabecera o de identificación del GAD provincial. La variable **v412\_inst** es una variable de desagregación que abarca las entidades y/o instituciones que impulsaron o generaron mecanismos de articulación con los GADs provinciales. Sus valores son: 1 (“a. MAGAP”), 2 (“b. SENAGUA”), 3 (“c. CNC”), 4 (“d. SENPLADES”), ..., 10 (“j. Asociaciones productivas”) y 11 (“k. Otro ... Especifique”). Las demás variables tienen sufijos que representan al número de la columna etiquetada en el formulario.
- **6.3\_Gastos:** Es la base de datos que contiene las variables de la pregunta 6.3. Consta de 4 variables por 336 registros o casos (14 por cada GAD provincial), de las cuales una es variable de cabecera o de identificación del GAD provincial. La variable **v63\_item** es una variable de desglose que abarca a todas las posibles categorías de ítems de gasto para gestión y protección ambiental. Sus valores son: 1 (“a. Proteger el aire, clima y la capa de ozono, incluyendo la implementación de medidas de mitigación y adaptación al cambio climático”), 2 (“b. Prevenir, controlar y mitigar la contaminación de suelos, aguas subterráneas y superficiales”),..., 12 (“l. Promover y fortalecer la institucionalidad de políticas públicas para ambiente”), 13 (“m. Gasto en personal ocupado en actividades ambientales”), 14 (“Total (a+b+c+d+e+f+g+h+i+j+k+l+m)”). Las demás variables representan a las columnas de valor programado (USD) y valor devengado (USD) del gasto en gestión ambiental.

 <p>INEC instituto nacional de estadística y censos</p>	<p>GUÍA DE USO DE LAS BASES DE DATOS – CIEAGP 2015</p>	<p>Código:</p>
---	--	----------------

### 3. RECOMENDACIONES GENERALES

---

Finalmente, se sugiere tomar en cuenta las siguientes recomendaciones generales:

- Antes de procesar las bases de datos, se sugiere ordenar las filas o registros, tanto por la clave identificadora de GAD provincial (variable **gob\_cod**), como por la(s) variable(s) que forman el desglose de las tablas cruzadas o tablas anidadas que conformen las diferentes variables de las bases de datos.
- Se sugiere que se preste atención al formulario, junto con las variables, haciendo una inspección visual previa a los cálculos de tabulados y otras transformaciones de las variables de las bases de datos,.



