

# **PLAN DE VALIDACIÓN DE LAS ESTADÍSTICAS DE NACIMIENTO**

**Caracas, Enero de 2009**

## I. INTRODUCCIÓN.

Desde noviembre del año 2001, fue creado por ley el Instituto Nacional de Estadística (INE), quien según su reglamento interno, designa a la Gerencia General de Estadísticas Demográficas, coordinar en conjunto a la Oficina de Tecnología de la Información, el procesamiento de los datos de las estadísticas vitales registradas; considerando para ello, los procesos de control estadístico de la recolección y procesamiento de los registros vitales con base al registro civil, a fin de garantizar la calidad estadística de los hechos vitales, así como la comparabilidad nacional e internacional de sus resultados.

En virtud de cumplir con este objetivo, la Gerencia General de Estadísticas Demográficas tiene adscrito dos gerencias de línea, la Gerencia de Censo de Población y Vivienda y la Gerencia de Análisis Demográfico, esta última tiene a su cargo la Coordinación de Investigación y Desarrollo y la Coordinación de Estadísticas Vitales. Estas coordinaciones tienen como función establecer, conjuntamente con la Gerencia de Diseño Estadístico y Control de Calidad y la Gerencia de Normalización Estadística, los parámetros de calidad a ser aplicados a la recolección y procesamiento de las estadísticas de nacimientos.

Considerando lo hasta ahora expuesto y reconociendo que desde la creación del **INE**, uno de sus principios fundamentales es producir información estadísticas con características de transparencia, comparabilidad, confiabilidad y neutralidad, se ha llevado a cabo simultáneamente al diligenciamiento de los instrumentos estadísticos referidos a los nacimientos, un conjunto de actividades de revisión y mejoramiento de los procesos de control, validación y depuración de los datos que permiten detectar y depurar los errores de distinta naturaleza. Sin embargo, como se verá a continuación, la validación, depuración e imputación no se deberá entender como procesos independientes sino como acciones interrelacionadas que forman un todo centrado en la obtención de información de calidad.

Es por ello que el presente documento, recoge y sintetiza las experiencias identificadas en los talleres de mejoramiento de las estadísticas vitales, así como las validaciones básicas existentes en el procesamiento de las estadísticas vitales que habían sido elaborados con anterioridad, con la intención de consolidar el plan de validación de las estadísticas de nacimientos.

## **II. FINALIDAD**

El propósito de este documento es sintetizar las consideraciones generales y específicas del plan de validación del proceso de captura de los instrumentos estadísticos referidos a las estadísticas defunción mediante la captura en línea de los datos, específicamente por la plataforma tecnológica institucional reconocida como SISTINE. En síntesis, este documento representa un plan escrito que indica cómo la validación será conducida en el proceso de captura de los datos contenidos en los instrumentos estadísticos, identificando para cada ítem, los valores posibles, omitidos y no aplica; así como las condiciones y restricciones a la cual se someten estos valores, considerando para ello, la dependencia de los ítems con sus predecesores.

La información presentada en este plan, podrá ser utilizada o consultada para realizar los criterios de validación del proceso de captura de los ítems de los instrumentos estadísticos de nacimientos, los cuales deberán ser utilizados como insumo necesario para la actualización y/o desarrollo del sistema de captura de los datos referidos a los nacimientos vivos registros.

Es por ello que este plan de validaciones tiene una relevancia muy importante para garantizar que los datos capturados por el sistema automatizado, puedan considerarse lo suficientemente válidos, ya que de no ser delimitado perfectamente al marco de la validación aquí presentado, va a ser imposible garantizar la transparencia y confiabilidad de las estadísticas de defunción. En otras palabras, la validación busca proporcionar un alto grado de confianza de que un instrumento de recolección, un proceso o un sistema específico producirán en forma consistente un resultado que cumpla con sus especificaciones predefinidas.

## **III. PLAN DE VALIDACIÓN**

Para acometer en el detalle del plan de validación, previamente es menester conocer que una validación es una verificación documentada que proporciona un alto grado de confianza de que un sistema integral o un proceso funciona de la manera prevista, en el ambiente de operación normal, lo cual aplica de manera similar al caso de los instrumentos de recolección, ya que por validación, también se entiende como el proceso de comprobar la precisión de los datos, a partir de un conjunto de reglas que se pueden aplicar como control, para especificar el tipo y el intervalo de valores y datos que son válidos en el proceso de captura y/o transcripción, los cuales pasan por un proceso de verificación y cotejo que certifica la calidad, congruencia e integridad de la información que se pretende caracterizar.

Las validaciones o procedimientos de control que se establecerán en este plan, serán el insumo para monitorear la transcripción y validar la data transcrita a partir de los instrumentos estadísticos, a fin de disminuir los errores que puedan causar inconsistencia en la información de los instrumentos estadísticos.

#### **IV. VALIDACIONES GENERALES**

Las validaciones generales se refieren las consideraciones universales para convertir en número la información recolectada por el instrumento estadístico de defunción. En este sentido se debe identificar primero los tipos de ítems o variables que conformarán la estructura de datos, a saber:

- Para los ítems que presentan opciones cerradas y pre-codificadas como lo son: sexo, sitio del hecho, tipo de embarazo, tipo de parto, persona que atendió el parto, situación conyugal, tipo de matrimonio o unión y alfabetismo; tienen asignados los valores numéricos, impresos generalmente al lado izquierdo de las opciones de cada pregunta, los cuales deberán ser considerados como los valores válidos que deberá tener en la data al momento de almacenarlos en un registro informático.
- Para el caso de los ítems o preguntas abiertas, vale decir, aquellas que no tienen alternativas ó opciones de respuesta preestablecidas, en lugar de ello, exigen una respuesta textual ó numérica, como se detalla a continuación:
  - Si son referidas a los casos de identificación de las entidades, municipios, parroquias y localidades que identifiquen los lugares de ocurrencia, residencia o registro, se deberán codificar utilizando para ello, los clasificadores de códigos elaborados para este fin. Estos son: División Político Territorial (DPT) y el clasificador uniforme de ocupación.
  - En el caso de que el ítem solamente acepte valores numéricos, como el asunto referido al número de hijos y la duración del embarazo; pueden presentarse los casos de que coloquen abreviaturas o textos, por ejemplo: (uno, dos, tres, guión "-", etc.), para ello se deberá colocar el número correspondiente a fin de almacenar valores numéricos en lugar de un texto.
  - En el caso de que el ítem solamente acepte valores numéricos, como el tema referido a las fechas de registro, ocurrencia y nacimiento; pueden presentarse los casos de que coloquen abreviaturas o textos, por ejemplo: (ene, feb, mar, etc.), para ello se deberá colocar el número correspondiente a fin de almacenar valores numéricos en lugar de un texto y a su vez, deberá de validarse que sean en conjunto el día, mes y año, una valor válido como fecha calendario.

- En el caso de que el ítem no posea un clasificador de códigos, que permita almacenar un valor numérico en lugar de un texto, se deberá considerar el almacenamiento del texto: cédula de identidad, nombres y apellidos del niño, la madre y el padre y la dirección de residencia de estos.
- Por último están los ítems que presentan una combinación de las preguntas pre-codificadas y abiertas, es decir, tienen opciones preimpresas con su respectiva codificación y espacios para completar la descripción de las opciones. Este tipo de ítems al igual que las preguntas pre-codificadas Y abiertas, deben ser transformados a códigos, como se detalla a continuación:
  - Si son referidas a los casos de identificación del lugar de nacimiento del fallecido, se deberá identificar si fue en Venezuela o en el extranjero como lo indica su pre-codificación, sin embargo, el detalle de ello, deberá codificarse utilizando para ello, los clasificadores de códigos elaborados para este fin. Estos son: División Político Territorial (DPT) para identificar la entidad federal, en caso de que naciera en el país, de lo contrario, se deberá utilizar el clasificador de países para identificar el país.
  - En el caso de que se refiera a las características educativas de los progenitores de el niño o la niña, vale decir, el nivel educativo y el último grado o año aprobado, se deberá identificar primero el nivel educativo como lo indica su pre-codificación, sin embargo, el detalle del último grado o año aprobado, se deberá cargar según el valor numérico declarado en el espacio correspondiente.
- Para todos los casos, es necesario registrar la totalidad de los ítems o variables del instrumento estadístico, independientemente que exista omisión en los datos de los progenitores del recién nacido.
- Todos los datos registrados en el sistema, deberán ser consistentes con los valores de los ítems o variables del instrumento estadístico de defunción, a fin de garantizar que el registro transcrito será lo más semejante posible a los valores pre-codificados.
- No deben existir campos nulos en el registro de defunción, con excepción de las variables abiertas que acepten caracteres alfanuméricos; como es el caso de los nombres y apellidos del fallecido.
- **IV. VALIDACIONES DE UNICIDAD DE REGISTRO**

Las validaciones referidas a la identificación de un registro único, se refieren a las consideraciones de identificación como registro unitario en la base de datos, que en el caso de defunciones se refiere a la concatenación de cuatro variables: tipo de defunción, el número interno del INE (asignado por las gerencias de

estadísticas estatales en los procesos previos a la transcripción), el código de la entidad de registro y la fecha de registro (considerando para ello el formato válido según la estructura de día/mes/año).

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Tip. de def.	Tipo de Defunción	3	1-2-3	-
NUM_BOLETA	NÚMERO DE BOLETA	7	1-99.998	ENTERO
COD_ENTIDAD_REG	CÓDIGO DE LA ENTIDAD DE REGISTRO	2	1-24	ENTERO
FECHA_REG	FECHA DE REGISTRO CON ESTRUCTURA "DD/MM/YYYY"	10	FECHA CALENDARIO	FECHA O CARACTER

- Para el caso de estos tres campos, todos deben de aceptar solamente valores numéricos.
- Para el tema de entidades federales, deberán coincidir con la división político territorio utilizada por el INE.
- La fecha de registro deberá contemplar para el caso de los días valores válidos de uno a treinta y uno; de uno a doce para el caso del mes de registro; y el valor del año preferiblemente predefinido por el sistema.
- Se recomienda utilizar controles que permitan identificar la fecha de registro, de lo contrario se deberá garantizar una fecha calendario válida.
- Previo a la búsqueda en el registro de nacimientos, de la data que se intenta transcribir se deberá validar que todos los valores sean numéricos y válidos.
- En el caso de identificar en la data transcrita la información de identificación igual a la que se intenta transcribir, se deberá de evitar la continuidad de transcripción, y mostrar un mensaje de advertencia donde se notifique que ya existe un registro con las características de identificación señaladas.

#### IV. VALIDACIONES PARTICULARES

##### Tipo de Defunción:

Es tipo de dato es numérico y la longitud debe ser de 3 posiciones al cual corresponda el tipo de defunción: 1 corresponde a una muerte fetal, 2 muerte de menor de 1 año y 3 muerte mayores de 1 año.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Tip. de def.	Tipo de Defunción	3	1-2-3	-

##### Sexo:

El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones. Debe ser 1 femenino (F) y 2 masculino (M).

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Sex.	Sexo	2	1-2	9

##### Fecha de Nacimiento

El tipo de dato es fecha y la longitud del campo debe ser de 6 posiciones. El formato de la fecha es DD/MM/AAAA. Día (2 dígitos), mes (2 dígitos), año (4 dígitos) para un total de 10 dígitos.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Fec. Nac	Fecha de Nacimiento	6		99/99/9999

##### Fecha de Defunción

El tipo de dato es fecha y la longitud del campo debe ser de 6 posiciones. El formato de la fecha es DD/MM/AAAA: día (2 dígitos), mes (2 dígitos), año (4 dígitos) para un total de 10 dígitos.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Fec. Nac	Fecha de Nacimiento	6		99/99/9999

##### Lugar de nacimiento:

El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones. Debe ser 1 Venezuela y 2 Exterior. La entidad federal para país se adjudica de acuerdo a la División Política territorial.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Lug.Nac	Lugar de Nacimiento	2	1-2(códigos de entidades y países)	99

#### Sitio donde ocurrió la muerte:

El tipo de dato es numérico y la longitud de la variable debe ser de 6 posiciones. Debe ser 1 Hospital, 2 clínicas, 3 casa, 4 vía pública y 6 ignorado.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Sit. Ocu. Mue	Sitio donde ocurrió la muerte	6	1-2-3-4-5-6	9

#### Fecha de Registro:

El tipo de dato es fecha y la longitud del campo debe ser de 6 posiciones. El formato de la fecha es DD/MM/AAAA: día 1-31 (2 dígitos), mes 1-12 (2 dígitos), año (4 dígitos) para un total de 10 dígitos. Compruebe el siguiente filtro: la fecha de registro DD/MM/AAAA deber igual o mayor a la fecha de defunción.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
FECHA_REG	FECHA DE REGISTRO CON ESTRUCTURA "DD/MM/YYYY"	10	FECHA CALENDARIO	99/99/9999

#### Entidad federal:

El tipo de dato es numérico y la longitud de la variable debe ser 2 dígitos.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Ent.fed	Entidad federal:	2	1-24	99

#### Duración del embarazo (semana de gestación):

Es tipo de dato es numérico y la longitud debe ser de 2 posiciones al cual corresponda a la semana de gestación: 1 corresponde menor a 28 semanas y 2 se adjudicar a 28 semana y más. Para valores omitidos 99.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Dur. Emba	Duración del Embarazo	2	1-2	9

#### Tipo de embarazo:

El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones. Debe ser 1 único y 2 múltiple.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Tip. Emba	Tipo de Embarazo	2	1-2	9

#### Asistente del parto:

El tipo de dato es carácter y la longitud de la variable debe ser de 3 posiciones. Debe ser 1 médico, 2 comadrona y 3 otros.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Asi. Part	Asistente del parto	3	1-2-3	9

#### Edad de la madre:

El tipo de dato es numérico y la longitud de la variable debe ser de 3 posiciones (10-100).



NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Eda.mad	Edad de la Madre	2	12-49	99

#### Número de hijos tenidos:

- **Número de hijos nacidos:** El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones 0-30.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Núm.hij.Nac	Número de hijos Nacidos	2	0-30	99

- Establecer la siguiente condición: el número de hijos nacidos vivos, sea mayor o igual al número de hijos actualmente vivo y al número de hijos nacidos que han fallecidos.

- **Número de hijos actualmente vivo:** El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones 0-30.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Núm.hij. act.vivo	Número de hijos actualmente vivo	2	0-30	99

Establecer la siguiente condición:

- Número de hijos nacidos vivos – número de hijos actualmente vivos= número de hijos fallecidos.

- Edad de la madre – Número de hijo actualmente vivo(NHNV) >11

- **Número de muertes fetales:** El tipo de dato es numérico y la longitud de la variable debe ser de 2 posiciones 0-30.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Núm. Mue.Fet	Número de muertes fetales	2	0-30	99

#### Situación conyugal:

El tipo de dato es numérico y de carácter, la longitud de la variable debe ser de 6 posiciones: 1 soltera, 2 casada, 3 divorciada, 4 viuda, 5 unidad y 6 separada.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Sit.Conj	Situación conyugal	6	1-2-3-4-5-6	9

#### Sabe leer y escribir:

El tipo de dato es numérico y de carácter, la longitud de la variable debe ser de 1-2 posiciones: SI 1 y NO 2.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Sab.lee.escri	Saber leer y escribir	2	1-2	9

#### Nivel Educativo y último Grado o año aprobado:

El tipo de dato es numérico y la longitud de la variable debe ser de 1-6 posición. Se debe registrar la opción que corresponde al último título obtenido en el nivel educativo aprobado, si la madre no posee ningún título debe registrar la opción 1 Ninguno, 2 preescolar, 3 básica, 4 media, diversificada y profesional, 5 superior y 6 ignorado.

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Nív. Edu.ult.año aprob	Nivel Educativo y último Grado o año aprobado	6	1-2-3-4-5-6	99

--	--	--	--	--

Debe establecerse lo siguiente filtro entre la edad de la madre y su último grado o año aprobado

Edad	10	11	12	13	14	15	16	17 y más
Año o grado aprobado	5	6	7	8	9	1 media	2 media	Superior

#### Profesión Habitual:

El tipo de dato es numérico y la longitud de la variable debe ser de 3 posición de 0-999, según el Clasificador Uniforme de Ocupaciones (CIU).

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Prof.Hab	Profesión Habitual	3	CIU-58	999

#### Ocupación Habitual:

El tipo de dato es numérico y la longitud de la variable debe ser de 3 posición de 0-999, según el Clasificador Uniforme de Ocupaciones (CIU).

NOMBRE DEL CAMPO	DESCRIPCIÓN	TAMAÑO	VALORES VÁLIDOS	OMITIDOS
Ocu.Hab	Ocupación habitual	3	CIU-58	999