

**INSTITUTO NACIONAL DE ESTADÍSTICA,
GEOGRAFÍA E INFORMÁTICA**

**Principios básicos de las
clasificaciones estadísticas en el
ámbito sociodemográfico**

Marzo 2005

Objetivo general

El presente documento tiene como objetivo plantear algunos conceptos básicos sobre las clasificaciones estadísticas y los principios generales que las rigen para su elaboración o revisión; con ello se busca llegar a términos y definiciones comunes que guíen el trabajo de homologación de las clasificaciones sociodemográficas que actualmente se está llevando a cabo en el INEGI.

Cabe aclarar que este documento forma parte de un trabajo más amplio que da cuenta de las clasificaciones usadas en el ámbito sociodemográfico que recaba información estadística a partir de censos de población y vivienda (conteos intercensales), encuestas en hogares y registros administrativos.

Índice

Introducción	5
Consideraciones generales sobre las clasificaciones	6
¿Qué es una clasificación?	6
Términos semejantes	7
Principios de la clasificación	9
Las clasificaciones estadísticas	9
Estructura de una clasificación	10
Principios básicos de la clasificación estadística	11
Comentarios finales	13
Bibliografía	17

Introducción

En México, los censos, las encuestas y los registros administrativos hasta hace muy poco tiempo se habían desarrollado de manera autónoma unos de otros, si acaso sólo con algunos contactos tangenciales. Esto dio lugar a que existiera una disparidad de criterios y resultados en aspectos que son comunes y que por lo tanto deberían asemejarse. Con el fin de unificar la producción de las estadísticas generadas por estas fuentes, el INEGI ha emprendido un esfuerzo por estandarizar los procesos, que implica identificar los aspectos comunes de los proyectos, para encontrar coincidencias entre ellos y adoptar decisiones que los unifiquen. Esta revisión actualmente se está llevando a cabo para las estadísticas sociodemográficas.

El trabajo de homologación que se ha emprendido busca asegurar la complementariedad de las fuentes y la integración del Sistema Nacional de Información Estadística. Homologación significa uniformar las fuentes de información en los siguientes aspectos:

- En los conceptos, para que un mismo término designe el mismo fenómeno y no se aplique a distintos aspectos. Esto significa que si el concepto es abordado por varias fuentes, éste debe ser definido en los mismos términos para asegurar que la medición es la misma y por ende los datos resultantes son comparables; esto independientemente de diferencias temporales, geográficas o grado de profundidad de cada fuente.
- En las preguntas y opciones de respuesta, para conservar la misma redacción y evitar cambios de sentido o malas interpretaciones por parte del informante.
- En la validación, es decir, para aplicar los mismos criterios de revisión y consistencia y evitar resultados distintos.
- En las clasificaciones, y que cada fuente determine el nivel de desglose en el cual la va a utilizar.
- En el tratamiento de la información, esto es, que se sigan las mismas normas en la captura y validación.
- En criterios de agrupamiento, cálculo de indicadores y presentación de resultados estadísticos, para que la generación de cifras sea homogénea.

Actualmente se está trabajando la homologación de los marcos conceptuales, las preguntas de los cuestionarios, las opciones de respuesta, los criterios de llenado y las clasificaciones utilizadas en las variables, abiertas o cerradas, de las distintas fuentes o proyectos de

información. Precisamente sobre este último punto, el presente documento busca plantear un marco común de conceptos y definiciones que guíen el trabajo de homologar las clasificaciones.

Consideraciones generales sobre las clasificaciones

Las clasificaciones han sido una herramienta básica del desarrollo de las ciencias. De hecho, la operación básica y a la vez más sencilla de toda ciencia es la de delimitar, ordenar, clasificar. Las clasificaciones nacen de la necesidad de entender, identificar y simplificar el manejo de la multiplicidad del mundo fenoménico. El primer esfuerzo real para desarrollar un sistema de clasificación empezó con los antiguos griegos; hacia el año 350 a.C., Aristóteles dividió a los organismos en dos grupos: reino animal y reino vegetal. Posteriormente, con el desarrollo de las ciencias modernas se extendió el uso de los sistemas de clasificación a las disciplinas particulares y se ha generalizado en la actualidad como una práctica de todas las actividades humanas.

Aquí se abordarán algunos principios básicos de las clasificaciones estadísticas. En este ámbito, las clasificaciones son elementos fundamentales de la infraestructura estadística, pues se utilizan para la generación de la información y la presentación de resultados, además de que aseguran un tratamiento uniforme de los datos estadísticos y permiten la comparabilidad con estadísticas producidas anteriormente o por diversos proyectos de la misma institución o de otras instituciones nacionales o internacionales.

¿Qué es una clasificación?

De acuerdo con el *Diccionario de la Lengua Española*¹ clasificación es la “Acción y efecto de clasificar” y clasificar es “Ordenar o disponer por clases”. Por su parte, *The American Heritage Dictionary*² define clasificación como: “El acto, proceso o resultado de clasificar” y clasificar como: “Ordenar u organizar de acuerdo con clases o categorías.”

Al consultar obras más especializadas, encontramos que la clasificación constituye un campo de la lógica. Algunas referencias que se hacen en torno a las clasificaciones son las siguientes:

¹ *Diccionario de la Lengua Española*, <http://www.rae.es/>

² *The American Heritage Dictionary of the English Language*, fourth edition, Houghton Mifflin Company, 2000.

“La clasificación consiste en agrupar los objetos por sus semejanzas y en separarlos por sus diferencias.”³

Una clasificación es “...una serie ordenada de divisiones y subdivisiones de cosas en grupos distintos pero relacionados.”⁴

Términos semejantes

Sin embargo, antes de llegar a una definición de clasificación más precisa para nuestro campo de interés, es importante aclarar otros términos semejantes que también tienen como objetivo ordenar sucesos; estos términos son los de ‘taxonomía’, ‘nomenclatura’ y ‘catálogo’.

Nuevamente, de acuerdo con el *Diccionario de la Real Academia Española*: ‘taxonomía’ es la “Ciencia que trata de los principios, métodos y fines de la clasificación. Se aplica en particular, dentro de la biología, para la ordenación jerarquizada y sistemática, con sus nombres, de los grupos de animales y de vegetales.”⁵

Por su parte, *The American Heritage Dictionary* dice que ‘taxonomía’ es: “1. La clasificación de los organismos en un sistema ordenado que indica relaciones naturales. 2. La ciencia, leyes o principios de las clasificaciones; sistemática. 3. División en grupos ordenados de categorías.”⁶ Sistemática aquí se refiere, según este mismo diccionario, a la ciencia de la clasificación sistemática, es decir, es sinónimo de taxonomía.

En cuanto al término ‘nomenclatura’, los diccionarios ya citados dicen que se refiere a una lista de nombres de personas o cosas o al procedimiento de asignar nombres a los tipos y grupos de organismos listados en una clasificación taxonómica.

Es importante aclarar que en el ámbito de las clasificaciones internacionales por lo general se utilizan indistintamente los términos clasificación y nomenclatura. Tal parece que en Europa ha predominado el segundo término,⁷ pero de acuerdo con la ONU, la definición de nomenclatura es más estrecha que la de clasificación;

3 José Rubén Sanabria, *Lógica*, editorial Porrúa, décimosegunda edición, México, 1980, p. 91.

4 Francisco Montes de Oca, *Lógica*, editorial Porrúa, México, 1971.

5 *Diccionario de la Lengua Española*, op. cit.

6 *The American Heritage Dictionary*, op.cit.

7 Un ejemplo lo constituye la NACE, que significa Nomenclatura de Actividades Económicas de la Comunidad Europea.

nomenclatura es el “Nombramiento sistemático de cosas o un sistema de nombres o de términos para las cosas. En clasificaciones, la nomenclatura implica un nombramiento sistemático de categorías o elementos (ítems)...Una nomenclatura es esencialmente una convención para describir observaciones, mientras que una clasificación estructura y codifica las observaciones también.”⁸ Por lo tanto, ‘nomenclatura’ designa la existencia de una lista de nombres o términos (o categorías) con los que se designan los elementos observados, pero cuando estos nombres o términos se estructuran y codifican hablamos de una clasificación, por ello podemos decir que una clasificación es una nomenclatura estructurada.

La palabra ‘catálogo’ está asociada con nomenclatura, pues también se refiere a una lista o enumeración; este significado se deriva de su etimología: del latín *catalogus*, y éste del gr. *κατάλογος*, lista, registro. El *Diccionario de la Real Academia Española* lo define así: “Relación ordenada en la que se incluyen o describen de forma individual libros, documentos, personas, objetos, etc., que están relacionados entre sí.”⁹ Por su parte, las Naciones Unidas no tienen considerado este término en su glosario de clasificaciones, ni se ha encontrado en otros documentos de este organismo.

Recapitulando lo expuesto, tenemos que ‘taxonomía’ es la ciencia de la clasificación, y el término tiene un uso muy importante en la biología; ‘nomenclatura’ y ‘catálogo’ se refieren a una lista o enumeración, y clasificación conlleva la estructuración de sus elementos y su codificación.

Las características de una clasificación son:

- Agrupa los objetos por sus semejanzas y los separa por sus diferencias.
- Ordena las cosas por divisiones y subdivisiones que constituyen grupos distintos pero relacionados.
- Designa los elementos de la clasificación por medio de *categorías*. También se utiliza el término ‘clase’ para designar cualquier agrupación; de ahí clasificación como formación de clases.

⁸ *UN Glossary of classification terms*, ONU, disponible en Internet en: http://unstats.un.org/unsd/class/family/glossary_short.htm

⁹ *Diccionario de la Lengua Española*, *op. cit.*

- Agrupa, separa y jerarquiza de acuerdo con criterios específicos, conocidos como *principios de clasificación*, los cuales se explicitan.

Principios de la clasificación¹⁰

Las principales reglas que rigen la elaboración de una clasificación son las siguientes:

1. Todos los elementos de la clasificación han de determinarse con arreglo al mismo criterio de clasificación.
2. La clasificación ha de ser exhaustiva, es decir, contener todas las especies o miembros del concepto genérico.
3. Las partes deben excluirse entre sí.
4. La división debe ser ordenada y gradual.

Las clasificaciones estadísticas

Diversos organismos internacionales han desarrollado investigaciones para normar el desarrollo de las clasificaciones estadísticas; sin duda la Organización de las Naciones Unidas es la que ha desarrollado un trabajo sistemático para organizar, normar y estandarizar los sistemas de clasificación, por ello la mayor parte de las referencias que se hacen aquí son a este organismo.

En primer lugar, para la ONU clasificación “Es un conjunto de observaciones discretas, exhaustivas y mutuamente excluyentes que se pueden asignar a una o más variables que se medirán en la recolección y/o la presentación de datos.”¹¹

Y en específico “Una clasificación estadística es una clasificación que tiene un conjunto de categorías discretas, que se pueden asignar a una variable específica registrada en un estudio estadístico o en un archivo administrativo, y es usada en la producción y presentación de estadísticas.”¹²

En el ámbito de las clasificaciones estadísticas, la categoría es el término genérico para referirse a los elementos (ítems) que componen

¹⁰ Raúl Gutiérrez Sáenz, *Introducción a la lógica*, Esfinge, 1992 y Francisco Larroyo, *Lecciones de lógica, ética y estética*, editorial Porrúa, 1975.

¹¹ UN *Glossary of classification terms*, op. cit.

¹² Eivind Hoffmann y Maria Chamie, *Standard Statistical Classifications: Basic Principles*, <http://unstats.un.org/unsd/class/family/default.htm>

una clasificación, en cualquiera de sus niveles en que se subdividen. “Las categorías de la clasificación son identificadas generalmente por los códigos (alfabéticos o numéricos) que proporcionan un identificador único para cada categoría y denotan su lugar dentro de la jerarquía. Contienen los elementos que son subconjuntos de la clasificación a la cual pertenecen, por ejemplo actividades, productos, tipos de ocupaciones, tipos de educación, etc.”¹³

Vale la pena decir que en el INEGI la Dirección General de Estadística ha buscado desarrollar documentos técnicos y normativos sobre la generación de estadísticas básicas, y en este contexto uno de los primeros trabajos que emprendió fue la elaboración de un documento de carácter metodológico sobre el diseño conceptual en proyectos de estadística básica, que cubriera los aspectos comunes a todo proyecto, independientemente del método de captación. En este documento, se dice que las clasificaciones: “Son delimitaciones conceptuales de subconjuntos mutuamente excluyentes con referencia a una variable. Cada subconjunto se denomina clase y en determinados casos pueden también desglosarse en subclases...En el formato de captación de datos la clasificación se relaciona con las opciones de respuesta a una pregunta, mientras que en la presentación de resultados estadísticos la clasificación se aplica para conocer la distribución del total de elementos de una categoría...”¹⁴

Estructura de una clasificación

Se refiere a cómo las categorías se ordenan, se agrupan y se subdividen; puede ser una estructura jerárquica o plana. En esta última las categorías se arreglan en un solo nivel, por ejemplo la clasificación de sexo. Las clasificaciones jerárquicas tienen varios niveles, que corresponden a los diversos grados en que se especifica la variable que es observada; estos niveles pueden ser: secciones, subsecciones, divisiones, subdivisiones, grupos, subgrupos, clases y subclases.

¹³ UN Glossary of classification terms, op. cit.

¹⁴ Metodología para el Diseño Conceptual en Proyectos de estadística Básica, versión preliminar, diciembre 2003, disponible en la página de Intranet de la Dirección General de Estadística, Dirección de Normatividad.

Principios básicos de la clasificación estadística

Existen principios generales que sirven de guía para la elaboración de una clasificación. Algunos de ellos, son los siguientes:¹⁵

1. Una clasificación es un método de generalización.
2. Debe estar conformado por un número limitado de categorías mutuamente excluyentes, capaces de abarcar el amplio rango de situaciones.
3. Las categorías deben estar estructuradas de tal forma que permitan facilitar el estudio estadístico del fenómeno.
4. Cada caso¹⁶ que tenga particular importancia debe tener su propia categoría.
5. Cada caso debe tener un lugar bien definido en la lista de categorías.
6. Debe haber categorías residuales para casos que no pueden ser ubicados en las categorías más específicas.
7. Las categorías residuales deben ser tan pocas como sea posible.

Por su parte, las Naciones Unidas, en su glosario¹⁷ enumera las características de una buena clasificación:

1. Las categorías son exhaustivas y mutuamente excluyentes (es decir, cada miembro de una población puede ser asignado solamente a una categoría sin duplicación u omisión).
2. La clasificación es comparable a otras clasificaciones estándares (nacionales o internacionales) afines.
3. Las categorías son estables, es decir, no se cambian con demasiada frecuencia, o sin una revisión, justificación y documentación apropiadas.
4. Las categorías se describen con un título en un formato estándar y son respaldadas por las notas explicativas, índices codificados, codificadores y tablas de correspondencia para relacionar las clasificaciones (que incluyen versiones anteriores de la misma clasificación).

¹⁵ *Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud*, décima revisión, vol. 2, Organización Panamericana de la Salud, Oficina Sanitaria Panamericana, Oficina Regional de la Organización Mundial de la Salud.

¹⁶ En el documento original de la OMS se emplea la palabra ‘enfermedades’ y no ‘casos’; aquí se ha cambiado el término para generalizar los principios a cualquier clasificador.

¹⁷ *UN Glossary of classification terms*, op. cit.

5. Las categorías están equilibradas dentro del sistema de límites por los principios de la clasificación (es decir, ni demasiadas categorías ni demasiado pocas). Esto se establece generalmente aplicando los criterios de la significación (e.g. límites de tamaño en variables tales como empleo, volumen de ventas, etc.).
6. Las categorías reflejan las realidades del campo (e.g. la sociedad o la economía) con el cual se relacionan (e.g. en una clasificación de la industria, las categorías deben reflejar el cuadro total de las actividades industriales del país); y
7. La clasificación está respaldada por la disponibilidad de instrucciones, manuales, índices de codificación y el entrenamiento.

Como puede verse, algunos puntos planteados por la Organización Mundial de la Salud coinciden con los de la ONU y otros no; los que no coinciden, son complementarios. Así, tomando como base estos planteamientos y añadiendo otros, se hace un intento de sistematizar lo que podríamos considerar algunos principios básicos de una clasificación estadística. Estos principios son:

- El número de categorías debe ser limitado (“ni demasiadas categorías ni demasiado pocas”).
- Las categorías deben ser exhaustivas, es decir, deben contener todos los elementos de la población considerada.
- Las categorías deben ser mutuamente excluyentes, esto significa que cada unidad del universo considerado debe corresponder a un solo agrupamiento o categoría.
- El conjunto de categorías debe estar bien estructurado, esto es, ordenado y en división gradual; si se trata de una estructura jerárquica, las categorías deben estar organizadas de tal forma que “representen los diferentes niveles de detalle de medición de la variable”¹⁸. Y este nivel de detalle depende de “las necesidades descriptivas y analíticas” de la clasificación, que a su vez están determinadas por los usuarios.
- Cada caso que tenga particular importancia debe tener su propia categoría.
- “Se requieren definiciones descriptivas o listas exhaustivas de los contenidos de las categorías definidas.”¹⁹
- Los nombres o títulos de las categorías deben responder al requisito de simplicidad y precisión.

¹⁸ Eivind Hoffmann y Maria Chamie, op. cit.

¹⁹ Ibid.

- “Las categorías de la clasificación son identificadas generalmente por los códigos (alfabéticos o numéricos) que proporcionan un identificador único para cada categoría y denotan su lugar dentro de la jerarquía.”²⁰
- Debe haber categorías residuales para casos que no pueden ser ubicados en las categorías más específicas.
- Las categorías residuales deben ser tan pocas como sea posible.
- Es necesario considerar un rango libre en los códigos para que sean utilizados en una futura revisión o expansión de la clasificación.
- Las categorías son estables, es decir, no se cambian con demasiada frecuencia, o sin una revisión, justificación y documentación apropiadas.

Comentarios finales

De acuerdo con los planteamientos anteriores y para establecer un lenguaje común que facilite la comunicación en el área sociodemográfica de la Dirección General de Estadística, se propone utilizar el término ‘clasificación’ como genérico para designar tanto las clasificaciones como los catálogos que en la actualidad se usan. Esta propuesta se basa en:

- *La conveniencia de apegarnos a una recomendación y convención emitida por la ONU.*
- *La definición que la ONU emite sobre ellas, corresponde o se ajusta al tipo de documento que en la actualidad se denomina indistintamente catálogo o clasificación.*
- *El término es suficientemente comprensivo como para dar cabida a estructuras de clasificación plana o jerárquicas y alude a la existencia de claves únicas para codificar las categorías correspondientes a variables abiertas*

De la misma manera, para manejar un lenguaje común con los planteamientos internacionales es conveniente adoptar las definiciones sobre ‘clasificación’ y ‘clasificación estadística’ elaboradas por la ONU, esto es:

Clasificación. *“Es un conjunto de observaciones discretas, exhaustivas y mutuamente excluyentes que se pueden asignar a una o más*

²⁰ UN Glossary of classification terms, op. cit.

*variables que se medirán en la recolección y/o la presentación de datos.*²¹

*“Una clasificación estadística es una clasificación que tiene un conjunto de categorías discretas, que se pueden asignar a una variable específica registrada en un estudio estadístico o en un archivo administrativo, y es usada en la producción y presentación de estadísticas.”*²²

Respecto de los principios de las clasificaciones estadísticas se propone adoptar los siguientes.

- *El número de categorías debe ser limitado.*
- *Las categorías deben ser exhaustivas.*
- *Las categorías deben ser mutuamente excluyentes.*
- *El conjunto de categorías debe estar bien estructurado, ordenado y en división gradual.*
- *Casos con particular importancia deben tener su propia categoría.*
- *Debe existir definiciones descriptivas o listas exhaustivas de los contenidos de las categorías definidas.*
- *Los nombres o títulos de las categorías deben responder al requisito de simplicidad y precisión.*
- *Las categorías de la clasificación deben ser identificadas por códigos (alfabéticos o numéricos) o identificador único.*
- *Debe existir categorías residuales para casos no ubicados en categorías específicas.*
- *Las categorías residuales deben ser tan pocas como sea posible.*
- *Es necesario considerar un rango libre en los códigos, para una futura revisión o expansión de la clasificación.*
- *Las categorías deben ser estables, es decir, no se cambian sin previa revisión y justificación.*

En las clasificaciones estadísticas sociodemográficas del INEGI estos principios y definiciones han sido rectores y han tenido una aplicación práctica; por ello en el ámbito de estadísticas sociodemográficas existen tanto clasificaciones sencillas y planas como complejas y jerárquicas. A continuación a modo de ejemplo, esto es sin la pretensión de exhaustividad, se muestra como los

²¹ UN Glossary of classification terms, op. cit.

²² Eivind Hoffmann y Maria Chamie, *Standard Statistical Classifications: Basic Principles*, <http://unstats.un.org/unsd/class/family/default.htm>

principios que rigen una clasificación se han puesto en práctica en clasificaciones como son la Mexicana de Ocupaciones (CMO), el Sistema de clasificación de actividades económicas (SCIAN-México); la clasificación de Carreras profesionales y técnicas y la correspondiente a Lenguas indígenas.

Éstas se caracterizan por:

- Estructurarse en 4 o 6 niveles, según sea el caso. La CMO, por ejemplo se organiza desde grupo principal a ocupaciones individuales, en cambio el SCIAN se estructura de sector a clase.
- Asignar una clave única a cada conjunto de categorías. Tratándose de la CMO, por ejemplo, su estructura da lugar a una clave de cuatro dígitos al menor nivel de la jerarquía, que corresponde a las ocupaciones individuales. En el caso de las actividades económicas (SCIAN México) se contemplan seis dígitos, que corresponden al nivel de clase, esto es el máximo detalle o desglose de la clasificación.
- Considerar, en cada nivel de la jerarquía, claves para elementos insuficientemente especificados; de tal forma de “ir filtrando” conforme el grado de indefinición de la respuesta. En el caso por ejemplo de la clasificación de ocupaciones esto significa que se crearon claves para información insuficientemente especificada desde el nivel de grupo principal hasta grupo unitario.
- Contemplar una denominación para cada agrupamiento de categorías; de tal forma que de manera sucinta se expresa su contenido; sin embargo también se detallan los elementos incluidos y los excluidos, garantizando tanto la integridad (todas las categorías sujetas a clasificación tienen un lugar) como la exclusión (es decir cada categoría tiene un único lugar – por ende una clave – en la clasificación).
- Tomar como base o referencia las clasificaciones internacionales disponibles o recomendadas por oficinas de estadística u organismos que constituyen una autoridad

en la materia; con el propósito de garantizar la comparabilidad entre países.

Estas clasificaciones son las que se utilizan para recolectar, procesar y publicar la información sociodemográfica en el instituto. De hecho, a partir de la experiencia que se deriva de esta práctica, ha sido posible sistematizar los principios de clasificación propuestos para llevar a cabo el actual trabajo de homologación.

Bibliografía

1. Guibert, Bernard, Laganier, Jean y Volle, Michel. *An essay on industrial classification*, http://europa.eu.int/comm/eurostat/ramon/other_documents/index.cfm?TargetUrl=DSP_READINGS.
2. Hoffmann, Eivind y Chamie, Maria. *Standard Statistical Classifications: Basic Principles*, <http://unstats.un.org/unsd/class/family/default.htm>
3. Montes de Oca, Francisco. *Lógica*, editorial Porrúa, México, 1971.
4. Rainer, Norbert. *El sistema revisado de nomenclaturas y clasificaciones internacionales*, http://europa.eu.int/comm/eurostat/ramon/other_documents/index.cfm?TargetUrl=DSP_READINGS
5. Sanabria, José Rubén. *Lógica*, editorial Porrúa, décimosegunda edición, México, 1980, p. 91.
6. Real Academia Española. *Diccionario de la Lengua Española*, <http://www.rae.es/>
7. *The American Heritage Dictionary of the English Language*, fourth edition, Houghton Mifflin Company, 2000.
8. Organización de las Naciones Unidas. *UN Glossary of classification terms*, http://unstats.un.org/unsd/class/family/glossary_short.htm
9. Organización Mundial de la Salud. *Clasificación Estadística internacional de Enfermedades y Problemas Relacionados con la Salud*, décima revisión, vol. 2, Organización Panamericana de la Salud, Oficina Sanitaria Panamericana, Oficina Regional de la Organización Mundial de la Salud.