



MEDICIONES 2014-2016

Notas a las bases de datos

RESUMEN

Como parte de los esfuerzos tanto de la Red PAL como del Proyecto MIA por transparentar y facilitar el uso público de los datos, se ha aplicado un proceso exhaustivo de verificación y limpieza de los conjuntos de datos generados en MIA a partir de las distintas rondas de medición llevadas a cabo. El resultado de dichos procesos da lugar a las notas aquí presentadas y en las que se exponen los procedimientos empleados para generar los conjuntos de datos de las mediciones de 2014, 2015 y 2016 que acompañan.

DAVID LOYO PÉREZ | MARIANA ORTIZ CARMONA
Departamento de Análisis y Calidad de Datos

CONTENIDO

| | |
|--|----|
| GLOSARIO | 2 |
| PRESENTACIÓN | 5 |
| NOTAS A LAS DIFERENTES RONDAS DE EVALUACIÓN | 6 |
| MEDICIÓN 2014 | 9 |
| Limitación | 9 |
| RazónAbandono | 10 |
| EconomíaHogar | 12 |
| MEDICIÓN 2015 | 14 |
| Lengua | 14 |
| Participación9B | 16 |
| ViveOtro | 17 |
| UtilidadEscuela | 19 |
| MEDICIÓN 2016 | 21 |
| OcupaciónM y OcupaciónP | 21 |
| LecturasAdicionales | 23 |
| Reporte de violencia | 25 |
| MedioCalle y MiedoEscuela | 26 |

GLOSARIO

Categoría. Cada uno de los subconjuntos en los que se ordenan las respuestas dadas a una variable de respuesta abierta de tipo cadena en un conjunto ordenado de datos.

Categorización. En la generación de bases o conjuntos de datos, es el subproceso mediante el cual las respuestas de una variable de respuesta abierta son agrupadas dentro de categorías con la intención de reducir la cantidad de datos y facilitar su análisis.

Codificación. Valores asignados a las respuestas previamente categorizadas de una variable. Proceso mediante el cual se asignan valores numéricos a respuestas abiertas de tipo cadena durante el proceso de categorización.

Código [también entendido como valor]. En las bases o conjuntos ordenados de datos, es cada uno de los números que se asignan a las posibles condiciones en las que se puede hallar la respuesta a una variable. Es decir, los códigos representan tanto a las posibles respuestas a dicha variable (siempre que sean categorías), como a los posibles valores perdidos que puede asumir. Si la variable es de respuesta abierta de tipo numérico, entonces los códigos se limitarán a sus posibles valores perdidos.

Cómputo. Durante la generación de las bases o conjuntos de datos, es la operación aplicada para la transformación de los valores de las variables que permite ajustarlas entre las distintas versiones de las mediciones sin modificar la magnitud o sentido original de las mismas.

Criterios de clasificación. Dentro del subproceso de categorización, son los juicios según los cuales una respuesta dada puede ser incluida o no dentro de una categoría.

Edición. En términos generales, durante el proceso de generación de bases o conjuntos de datos, se refiere a la transformación de los valores mediante la aplicación de cómputos que hacen coincidir las variables encuestadas en las diferentes rondas de medición. También se refiere a los procesos de

categorización de los datos cadena contenidos en variables de respuesta abierta y a la generación de nuevas variables a partir de la información disponible.

Libro de códigos. Para cada conjunto de datos, es una descripción detallada de cada variable incluida, con información sobre su naturaleza, posibles valores asumidos y significado de sus categorías (si aplica), así como los códigos empleados para señalar, entre otros, valores perdidos, variables no evaluadas u ocasiones en las que algunas variables pierden vigencia en función de otras variables filtro.

Magnitud original. Durante el proceso de captura y la generación de las bases o conjuntos de datos, cada uno de los valores que asumen las diferentes variables según las respuestas registradas de cada uno de los encuestados en los cuestionarios e instrumentos de evaluación.

Nivel distrital. Dentro de las base o conjuntos de datos, es el nivel máximo de desagregación al que puede accederse para la generación de información. Este nivel hace referencia a la división geopolítica-electoral del país.

Registro. Asociado al proceso de captura, acción de asignar en cada una de las variables, los datos proporcionados por los encuestados.

Revisión. En términos generales, proceso que compara los valores registrados en un conjunto de datos con los contenidos en los cuestionarios e instrumentos de evaluación empleados para el levantamiento de la información en cada ronda de medición.

Tipo cadena. Hace referencia a un tipo de dato que contiene esencialmente texto.

Valor perdido. En un conjunto ordenado de datos, en los procesos de medición e intervención y en los libros de códigos, es un valor numérico que se asigna a una variable cuando durante el levantamiento de la información –sea por error del encuestador o por omisión del entrevistado– no se obtuvo respuesta a alguna de las preguntas del cuestionario aplicado.

Variables categorizadas. Es cada una de las respuestas que recibe una pregunta de respuesta abierta y que ha sido transformada tras incluirse dentro de una

categoría a la que se le ha asignado un valor numérico para representarla dentro de un conjunto ordenado de datos y en su respectivo libro de códigos.

Variable de respuesta abierta. En los cuestionarios y en los conjuntos ordenados de datos, se refiere a preguntas que esperan una respuesta del encuestado sin restricciones preestablecidas o categorías a las que ésta deba ajustarse.

Variable condicionante. En los conjuntos de datos, son variables que, tal como su nombre indica, condiciona las respuestas que otras pueden asumir.

Variable condicionada. En los conjuntos de datos y en los libros de códigos, son las variables cuyas respuestas dependen de la información contenida en las variables condicionantes. Ante determinadas respuestas de las variables condicionantes, las condicionadas pueden asumir sólo ciertas respuestas o, en algunos casos, perder vigencia.

Versión anonimizada. Conjunto de datos en el cual se encuentra suprimida toda información que permita la identificación de un municipio, localidad u hogar encuestado, dejando sólo la posibilidad de extraer información al nivel de desagregación distrital.

Versión Post Edición. Conjunto de datos resultante del proceso de edición.

PRESENTACIÓN

El proyecto Medición Independiente de Aprendizajes (MIA), busca incrementar la participación ciudadana a través de la inclusión de la sociedad en el ámbito de la educación, mediante la investigación colaborativa, para mejorar los aprendizajes básicos de niñas, niños y jóvenes; ampliar la rendición de cuentas y construir ciudadanía en México.

Consiste en evaluaciones periódicas a nivel estatal de aprendizajes básicos en lectura y matemáticas a través de entrevistas orales uno a uno en el hogar, haciendo uso de instrumentos simples que permiten seguir el proceso fácilmente tanto a los encuestadores como a los adultos en el hogar visitado.

Este método fue desarrollado en la India hace más de 10 años, y ha sido replicado en varios países de Asia y África. Todos ellos –incluido México– constituyen la Red de Acción Ciudadana por el Aprendizaje (PAL, por sus siglas en inglés).

Uno de los principales compromisos de la Red PAL es poner a disposición del público los conjuntos de datos generados por los diferentes países con el objetivo de que ciudadanos, investigadores y tomadores de decisiones del ámbito educativo regional, nacional e incluso internacional, puedan hacer uso de ellos para mejorar los procesos educativos.

En México, tras haber conducido 3 grandes rondas de evaluación del año 2014 al 2016, MIA ganó presencia en la región Sureste del país. Los estados evaluados han sido Veracruz en 2014; Puebla, Quintana Roo y Yucatán, en 2015; y Campeche, Tabasco y Veracruz (por segunda ocasión), en 2016. Esta nota y las bases de datos que acompaña, son parte del compromiso por generar datos de calidad abiertos al público.

NOTAS A LAS DIFERENTES RONDAS DE EVALUACIÓN

Si bien no es objeto de la presente nota describir la revisión y edición de las bases de datos, es necesario comentar que ambos son procesos llevados a cabo por el Departamento de Análisis y Calidad de Datos con la intención de garantizar que los valores registrados sean lo más fiables posible. Esto involucra, aunque no se limita a, el contraste de la información entre cuadernillos y bases de datos, su corrección (siempre que sea necesario), y posterior edición para que los valores coincidan con criterios predefinidos para facilitar su análisis.

Asimismo, es importante garantizar que, si bien las ediciones pueden y de hecho modifican los valores de algunas variables, esto es sólo parte de un cómputo que bajo ninguna situación altera la magnitud original. Esto quiere decir, que los valores pueden no coincidir entre las diferentes versiones de la misma base de datos, pero la categoría representada por ese valor es la misma en todo momento.

Son tres las versiones que se generan durante los procesos de revisión y edición: Post revisión, Post edición y Anonimizada. Ésta última es la versión disponible al público. En ella se suprimen todas las variables que permiten la identificación de participantes, escuelas o localidades ya que, además de cumplir con los lineamientos de la Red PAL, se atiende al llamado ético de garantizar la confidencialidad de los involucrados. Por ello, el nivel más bajo de desagregación del que se puede obtener información es el distrital.

Asimismo, desde la versión Post edición, las variables que originalmente registran respuestas de tipo cadena son analizadas con el objetivo de generar categorías que permitan realizar los análisis pertinentes. En otras palabras, las variables que se enlistan en la pestaña Variables categorizadas de los documentos de soporte (como se explica más adelante) pertenecen a la versión Post revisión de la base de datos en cuestión.

Especialmente ahora que la edición se ha llevado a cabo sobre las tres rondas de medición, algunos de los valores originales de 2014 y 2015 han sido transformados para hacerlos coincidir con los de la última medición, sobre todo en

aquellas variables que se han evaluado en cada ronda. O, en su defecto, algunos valores de la ronda de medición 2016 han sido categorizados a partir de los códigos de la medición 2015.

Por ejemplo, en la Tabla 1 se muestran los valores y las respectivas categorías que asumió la variable parentesco durante las diferentes rondas de evaluación, así como la codificación que le ha sido asignada para efectos de hacerla coincidir a través de las distintas bases de datos.

Tabla 1. Valores y categorías de la variable Parentesco.

| Valores | Categorías originales | | | Categorización actual |
|---------|-----------------------|--|-------------|----------------------------|
| | 2014 | 2015 | 2016 | |
| 1 | Mamá | Soy su mamá | Soy su mamá | Mamá |
| 2 | Papá | Soy su papá | Soy su papá | Papá |
| 3 | Abuela(o) | Soy su abuelo(a) | Otro | Abuelo(a) |
| 4 | Tía(o) | Soy su tío(a) | | Tío(a) |
| 5 | Otro | Soy su hermano(a) mayor | | Hermano(a) |
| 6 | | Soy su cuidador, nana o empleado(a) doméstico(a) | | Otro familiar |
| 7 | | Otro | | Otra persona (no familiar) |

Por otro lado, los documentos de soporte se presentan en formato .xls, y contienen por lo menos 3 pestañas: Libro de códigos, Variables categorizadas y Variables condicionantes.

La primera de ellas, como su nombre indica, contiene el libro de códigos. Esto es un condensado de la información requerida para entender la naturaleza de las variables. Inicia exponiendo la fuente de la información que se registra, la descripción, nombre y etiqueta perteneciente a cada variable. A continuación, se enlistan los códigos y categorías de todos los casos en los que aplica. Posteriormente, se describen los posibles valores perdidos que podrían encontrarse en los registros y la interpretación que debe hacerse de los mismos. Para finalizar con información sobre la escala de medición, tipo y extensión de cada variable.

La segunda pestaña corresponde a las variables que han sido categorizadas durante la edición del conjunto de datos. El número varía de una ronda de medición a otra, pero en todos los casos se enlistan las respuestas abiertas de

tipo cadena originales de cada uno de los encuestados, así como las categorías a las que han sido asignadas, de tal suerte que para los usuarios sea posible entender más a fondo las categorías de una variable en cuestión cuando éstas sean demasiado complejas o, en su defecto, elaborar nuevas categorías a partir de las respuestas originales.

Por último, la tercera pestaña contiene un listado de lo que llamamos variables condicionantes. Éstas reciben ese nombre porque, tal como puede suponerse, condicionan los valores que otras pueden asumir. Un ejemplo sencillo de este tipo es la variable *AsisteEscuela*, donde si la respuesta es negativa, otras variables como *Turno* o *Sostenimiento* no deberían recibir ningún valor. Por supuesto, el número de variables también varía de una medición a otra, por ello se incluye el listado de variables condicionantes asociadas a sus respectivas variables condicionadas, más los posibles valores que las segundas pueden asumir en función de los datos contenidos en las primeras.

MEDICIÓN 2014

El conjunto de datos de la medición 2014 llevada a cabo en el estado de Veracruz contiene un total de 242,658 valores (omitiendo FN), provenientes de 78 variables y 3111 encuestados. Y fue, contrario a lo que podría pensarse, el último en atravesar por los procesos de revisión y edición. Esto tiene implicaciones importantes puesto que para entonces se habían tomado decisiones importantes con respecto a la forma en la que habrían de tratarse algunas variables desde el momento de la revisión.

Retomando el ejemplo de la variable Parentesco de la Tabla 1: tal como puede verse en la columna correspondiente a la medición 2014, originalmente hay sólo 5 categorías posibles para dicha variable. Éstas se corresponden con las opciones posibles de respuesta del cuestionario empleado en esa ronda de medición. Sin embargo, los valores fueron transformados para hacerlos coincidir con la categorización actual (columna 5 de la Tabla 1) desde el momento de la revisión y así facilitar el posterior proceso de edición. Algunas otras decisiones de esta naturaleza también fueron tomadas por lo que es posible encontrar valores que no corresponden perfectamente con los que están planteados originalmente en las codificaciones de los cuestionarios.

Por otro lado, las variables que fueron categorizadas en esta ronda de evaluación son Parentesco, Limitación, RazónAbandono y EconomíaHogar. La primera es relativamente transparente y se ha abordado con suficiencia a lo largo del documento, por lo que no se le dará mayor atención.

Para las tres variables restantes, se explica a continuación sus respectivos procesos de categorización.

Limitación

Se preguntó a los responsables del niño(a) encuestado si éste sufría de alguna limitación de cualquier tipo que afectara su aprendizaje. Del análisis de las respuestas afirmativas se crearon cuatro categorías:

1. Limitaciones físicas o motoras (LFyM)

2. Limitaciones del desarrollo psicológico y del aprendizaje (LDPyA)
3. Limitaciones por enfermedades con tratamiento médico accesible (LETMA)
4. Limitaciones por enfermedades graves o crónicas (LEGC)

LFyM hace referencia a aquellas condiciones del aparato motor y de los órganos de los sentidos que sin caer en la cronicidad afectan el desarrollo normal del niño(a). En esta categoría se encuentran, por ejemplo, las limitaciones visuales y auditivas, pero también puede incluir, por mencionar algunas, pérdida de un ojo, displasia de cadera y soplo en el corazón.

LDPyA engloba los aspectos que, si bien pueden no tener una manifestación física, sí generan alteraciones en el comportamiento que, a la larga, tiene un impacto sobre el desempeño de los menores tanto en la escuela como en la vida diaria. Aquí se incluyen los problemas de hiperactividad, atención, lento aprendizaje, dislexia, entre otras.

LETMA es la categoría en la que se inscriben aquellas limitaciones a las que, sin restar importancia, se les consideró de fácil acceso a tratamiento médico.¹ Incluye, entre otras, asma, problemas de la tiroides y migraña.

Por último, *LEGC* refiere a aquellas afectaciones de salud cuya presencia puede suponer graves alteraciones al bienestar general del menor. Se incluyen en esta categoría, por mencionar algunas, leucemia, hidrocefalia y enfisema.

Razón Abandono

En esta variable se registran los motivos por los cuales algunos de los niños encuestados no están asistiendo actualmente o no asisten más a la escuela.

Originalmente, la variable contiene sólo las siguientes categorías:

1. Nunca estuve inscrito
2. Falta de dinero
3. Matrimonio o embarazo

¹ El proceso de categorización no contó con asesoramiento médico por lo que los expertos podrían no estar del todo de acuerdo tanto con las categorías como con los criterios de clasificación. Se insiste en que los valores originales están incluidos en los documentos de soporte con la intención de que puedan ser manipulados bajo otras perspectivas.

4. No me gusta la escuela
5. Tuve que trabajar
6. Discapacidad
7. No tengo documentos
8. Cambio de domicilio
9. Tuve que cuidar a mis hermanos
10. No hay escuela
11. Otro

El proceso implicó asignar valores adicionales a través de la categorización de las respuestas abiertas de tipo cadena de la opción de respuesta 11. Otro². El resultado son 5 nuevas categorías que se enlistan a continuación:

11. Deficiencias o problemas de salud
12. Problemas personales o familiares
13. Problemas de o con la escuela
14. Desempeño escolar insuficiente
15. Estudio otra cosa

La categoría *Deficiencias o problemas de salud* enlista las razones de abandono vinculadas a la salud del menor e incluye respuestas como las siguientes: “por problemas de salud”, “porque es hiperactivo” y “porque sufre de desmayos”.

La categoría *Problemas personales o familiares* incluye las dificultades asociadas a los ámbitos personal y familiar que son motivo de abandono escolar, y hace referencia a respuestas del tipo: “porque no tengo padres”, “porque no me dejaron seguir estudiando” y “porque no compraron mis libros”.

Problemas de o con la escuela se refiere a aquellos motivos asociados a problemas de funcionamiento de la escuela que impiden que el menor siga asistiendo o, en su defecto, a problemas del menor dentro del centro escolar. En esta categoría se incluyen los siguientes tipos de respuesta: “por bullying”, “por problemas con el maestro”, “porque no alcancé lugar” y “porque me expulsaron”.

² Estas categorías se aplicaron de forma idéntica para las rondas de medición 2015 y 2016.

La siguiente categoría, *Desempeño escolar insuficiente*, se vuelca sobre los motivos que están directamente vinculados con el desempeño del menor dentro del centro escolar con respecto a lo esperado para el grado escolar que cursa o cursaba. Por ejemplo: “porque me cuesta aprender”, “porque tengo lento aprendizaje”, y “porque no pasé de grado escolar”.

Por último, *Estudio otra cosa*, tal como sugiere, hace referencia a los motivos asociados al abandono de la escuela con objeto de atender otra clase de educación como pueden ser los cursos de capacitación para el trabajo.

12

Economía Hogar

Esta variable recaba información brindada por los menores sobre las actividades que, si aplica, realizan para apoyar económicamente a sus padres. Las categorías elaboradas se enlistan a continuación³:

1. Trabajos de limpieza a terceros
2. Trabajos de agricultura o ganadería
3. Trabajos de comercio o servicios
4. Labores que requieren esfuerzo físico
5. Labores domésticas
6. Actividades resultantes del esfuerzo intelectual, deportivo o artístico
7. Aportando dinero o disminuyendo el gasto familiar

La primera categoría, *Trabajos de limpieza a terceros*, se refiere a cualquier tipo de labor doméstica (generalmente de limpieza) que se lleva a cabo para personas no familiares del menor encuestado. Puede incluir, aunque no se limita a: limpiar casas, lavar ajeno, etcétera.

La segunda, *Trabajos de agricultura o ganadería*, engloba cualquier tipo de trabajo vinculado con la producción agropecuaria. Pueden encontrarse respuestas como las que se enlistan a continuación: abonar el campo, cultivar la milpa, cosechar alimentos, cuidar animales, etcétera.

³ Las categorías aquí enlistadas aplican de igual forma para la medición 2016 en la variable ActividadRemuneradaCat.

La tercera, *Trabajos de comercio o servicios*, incluye todas aquellas actividades vinculadas con el proceso de compra-venta de bienes o servicios. Ejemplos de actividades en esta categoría son: atender un negocio, vender comida, meserear, empacar mercancía, rentar mesas en la playa, etcétera.

La cuarta, *Labores que requieren esfuerzo físico*, está conformada por aquellas actividades para las que, aun cuando por su naturaleza podrían incluirse en otra categoría, la demanda de esfuerzo físico es esencial para llevarlas a cabo. Se encuentran incluidas respuestas como las enlistadas a continuación: trabajar de albañil, cortar y acarrear leña, realizar mudanzas, desmontar terrenos, ser cargador, entre muchas otras⁴.

La quinta categoría, *Labores domésticas*, se circunscribe a las actividades propias del hogar, de su mantenimiento y organización. En este caso se encuentran respuestas como, por mencionar algunas, barrer la casa, lavar trastes, realizar mandados, cuidar niños, haciendo el quehacer.

La sexta categoría, *Actividades resultantes del esfuerzo intelectual, deportivo o artístico*, tal como el nombre sugiere, enlista respuestas vinculadas con la posesión de habilidades que permiten al menor comerciar con éstas. Implica respuestas como vender tareas, dar clases o asesorías, ser músico, jugar fútbol, estar becado, ser payaso en el circo, entre otras.

La séptima y última, *Aportando dinero o disminuyendo el gasto familiar*, se refiere a las respuestas que, sin necesariamente especificar cómo se consigue el dinero, reportan aportar a la economía familiar ya sea solventando gastos o disminuyéndolos por medio del ahorro. Las respuestas de este tipo son similares a ahorrar, pagar servicios, y aportar dinero.

⁴ También se incluyen en esta categoría aquellos casos que sólo respondieron “trabajar”.

MEDICIÓN 2015

El de la medición 2015 es un conjunto de datos con un total de 540,015 valores (omitiendo FolioHogarGeneral y FolioNiñoGeneral), provenientes de 105 variables y 5143 encuestados de los estados de Puebla, Quintana Roo y Yucatán.

Las variables que fueron categorizadas en esta ronda de evaluación son Lengua, Participación9B, ViveOtro y UtilidadEscuela. Parentesco y RazónAbandono también forman parte de este conjunto de datos, pero recibieron un tratamiento idéntico al expuesto en apartados anteriores, por lo que la exposición se concentrará en las primeras cuatro.

14

Lengua

La codificación de esta variable inició con la asignación arbitraria de valores según el tipo de lengua adicional al español que el adulto reporta que se habla en el hogar. Éstas se dividieron en *Extranjeras* y *Nacionales*. Para las primeras no hubo más que asignar un valor conforme aparecieron en las respuestas de los encuestados. Sin embargo, codificar las lenguas nacionales es mucho más complejo puesto que existen 11 familias lingüísticas de las que se derivan 68 lenguas, y de éstas, a su vez, se derivan 364 variantes lingüísticas.

Por si fuera poco, algunos hogares hablan más de una lengua adicional al español ya sean nacionales o extranjeras e, incluso, en algunos casos se hablan lenguas de ambos tipos. Para dar solución a esta situación se elaboró una codificación de cinco dígitos que se construye de la siguiente manera:

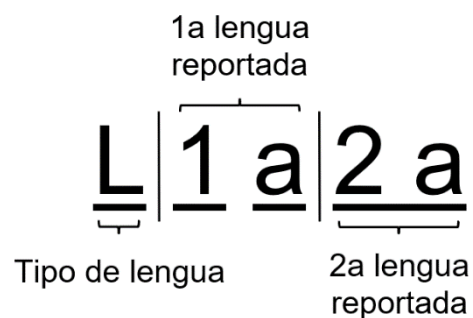


Figura 1. Construcción de la codificación de la variable Lengua.

El primer dígito, entonces, hace referencia a si la o las lenguas adicionales reportadas son nacionales, extranjeras o una combinación de ambas:

1. Lengua extranjera
2. Lengua nacional
3. Tanto lengua extranjera como nacional

Inmediatamente después se agrega uno de los valores contenidos en la Tabla 2 dependiendo de la respuesta dada. En todos los casos donde se habla más de una lengua adicional, el valor que se asigna a la primera posición es el que aparece primero en la lista. Por ejemplo, si una familia habla portugués y francés, el primer valor será 02. Francés y el segundo será 07. Portugués⁵. Si hablase Zapoteco y Mixteco, aparecería primero el valor 05. Mixteco y después 10. Zapoteco.

Tabla 2. Lista actualizada de las lenguas reportadas en las rondas de medición 2015 y 2016.

| Valores | Lenguas extranjeras | Lenguas nacionales |
|---------|---------------------|-------------------------|
| 01 | Alemán | Chinanteco |
| 02 | Francés | Cuicateco |
| 03 | Inglés | Mazateco |
| 04 | Italiano | Mixe/Popoluca/Zoque |
| 05 | Latín | Mixteco |
| 06 | Japonés | Náhuatl/Mexicano/Mexica |
| 07 | Portugués | Papanteco |
| 08 | Árabe | Otomí |
| 09 | | Totonaco/Tepehua |
| 10 | | Zapoteco |
| 11 | | Maya/Huasteco |
| 12 | | Purépecha |
| 13 | | Chol |
| 14 | | Tzeltal/Tseltal/Tzotzil |
| 15 | | Kanjobal |
| 16 | | Mam |
| 17 | | Mazahua |
| 18 | | Chontal |
| 19 | | Yokot'an |

⁵ La asignación de estos valores es completamente arbitraria y obedece al orden en el que fueron apareciendo en las respuestas de los encuestados.

Si, en cambio, se trata de una familia que habla tanto lenguas nacionales como extranjeras, luego del valor correspondiente a esta situación, se asigna el valor de la lengua extranjera a la primera posición y el valor de la lengua nacional a la segunda posición.

Por supuesto, una codificación de esta naturaleza limita la cantidad posible de valores que puede asumir a dos. Debe reconocerse que hay familias en las que se hablan hasta cuatro lenguas diferentes. En esos casos, los valores que son asignados al código son los que corresponden a las lenguas con menor prevalencia entre los encuestados. De este modo, si una familia hablase inglés, japonés y árabe, el código resultante sería 10608. 1 = Lengua extranjera, 06 = Japonés, 08 = Árabe.

Se entiende que los expertos pueden no estar de acuerdo con esta clasificación. Y se reconoce que aún hay trabajo por hacer para ofrecer una mejor versión de la misma.

Participación9B

Esta variable es parte de la escala de involucramiento parental contenida en el cuestionario empleado en la ronda de medición de 2015. El propósito de esta escala es indagar sobre cómo los padres se involucran en las actividades que a nivel escolar tienen injerencia en la educación de sus hijos(as). Para esta variable se les solicitó especificar de qué forma participaban en la toma de decisiones de sus respectivas escuelas, según el caso. Para ello, se crearon las siguientes categorías:

1. Dando opiniones/Aportando ideas/Aconsejando
2. En votaciones
3. Apoyando económicamente
4. Aportando mano de obra
5. Asistiendo a juntas
6. Hablando/Cuestionando/Discutiendo con los profesores
7. Hablando u organizándose con los demás padres o tutores

8. Dando seguimiento a las decisiones tomadas
9. Gestionando recursos

Todas ellas son lo suficiente claras como para requerir una explicación adicional, sin embargo, se aclara que en la categoría *Aportando mano de obra*, se incluyen todas las respuestas que se limitan a decir “Apoyando” o “Participando”. De igual manera, la categoría *Hablando u organizándose con los demás padres o tutores* incluye las respuestas donde los padres reportan ser parte de la mesa directiva o de la asociación de padres de familia.

17

ViveOtro

La variable *ViveOtro* también forma parte de un conjunto de variables que exploran con qué personas vive el niño en el hogar. En la ronda de evaluación 2015, se pregunta a los niños si vive con ellos el padre, la madre, los hermanos, abuelos, tíos y primos, o alguna otra persona.

A partir de las respuestas dadas, se crearon seis categorías enlistadas a continuación:

1. Familiar nuclear consanguíneo
2. Familiar no nuclear consanguíneo
3. Familiar nuclear político
4. Familiar no nuclear político
5. Miembros de su propia familia
6. Persona no familiar

De manera adicional, esta variable fue empleada para calcular otra a la que se ha llamado *TipoFamilia*. Esta hace referencia a una clasificación de los diferentes tipos de familia a partir de la cantidad y miembros que la conforman. Los diferentes valores y categorías son los siguientes:

1. Familia nuclear consanguínea
2. Familia extendida consanguínea
3. Familia nuclear política

4. Familia extendida política
5. Tiene su propia familia
6. Familia constituida por personas sin lazos sanguíneos o políticos

Las familias en la primera categoría *Familia nuclear consanguínea* son aquellas conformadas exclusivamente por los abuelos, padres y hermanos del niño(a) encuestado(a).

La segunda *Familia extendida consanguínea* sin omitir a cualquiera de los miembros de la primera categoría, incluye a los siguientes familiares con respecto al niño(a) encuestado(a): bisabuelos, cónyuges de los hermanos (cuñados), hijos de los hermanos (sobrinos), y hermanos de los padres (tíos).

La tercera categoría *Familia nuclear política*, a diferencia de la primera con la que comparte los mismos miembros de la familia, incluye a los cónyuges de los padres (padrastros o madrastras), y cónyuges de los abuelos (abuelastros o abuelastras).

La cuarta categoría *Familia extendida política* incluye a todos los familiares anteriores, más los enlistados a continuación con respecto al niño(a) encuestado(a): primos, cónyuges de tíos y cónyuges de primos.

La quinta *Tiene su propia familia*, como su nombre indica, hace referencia a todos los miembros de una familia de la que el menor encuestado(a) forma parte como cónyuge, en otras palabras, se trata de la familia constituida por el o la menor en concubinato, matrimonio o cualquier otra forma de asociación con su pareja. Por lo tanto, además de los cónyuges del menor encuestado(a), esta categoría puede también incluir a los hijos del menor, a sus suegros y, ocasionalmente, también a sus cuñados.

Por último, la sexta categoría *Familia constituida por personas sin lazos sanguíneos o políticos* está conformada, además de cualquier familiar, y tal como el nombre sugiere, por personas sin parentesco con el o la menor encuestada. Entre éstos pueden incluirse a las personas de servicio del hogar, amigos de los padres o personas a las que por alguna razón se les da alojamiento.

Todas estas categorías no son restrictivas hacia atrás, es decir que, aun cuando alguna de las familias pueda no contar con uno o más de los miembros que se enlistan en cada una, sigue siendo parte del mismo tipo de familia. Por ejemplo, familias nucleares sin padre o madre, o sin abuelos, o de hijos únicos.

En cambio, sí son restrictivas hacia adelante. En este sentido, un hogar conformado por padres e hijos más personas sin lazos sanguíneos se vuelve, en automático, parte de la categoría *Familia constituida por personas sin lazos sanguíneos o políticos*.

Ambas variables y sus respectivas categorizaciones fueron empleadas de manera idéntica para la ronda de medición de 2016.

UtilidadEscuela

El proceso de categorización de la variable UtilidadEscuela dio por resultado la generación de ocho categorías diferentes entre sí⁶:

0. No sirve.
1. Sirve ahora
2. Sirve para el futuro
3. Sirve para ser mejor
4. Sirve para ser exitoso
5. Sirve para conocerme
6. Sirve para convivir
7. No sé para qué sirve.

La primera categoría *No sirve*, y que es totalmente transparente, es asignada a las respuestas de los menores que consideran que la escuela no tiene utilidad.

La segunda *Sirve ahora*, está asociada a respuestas que reflejan un sentido de inmediatez sobre la utilidad que los menores ven en la escuela. Algunos ejemplos son: “Para aprender cosas”, “Para estudiar”, “Para tener bonita letra”, etcétera.

⁶ Las categorías aquí enlistadas aplican de igual forma para la medición 2016.

La tercera, *Sirve para el futuro*, a diferencia de la anterior, se refiere a las respuestas que reflejan una suerte de ganancia posterior tras la asistencia a la escuela, es decir, que la utilidad que ven en la escuela está proyectada hacia el futuro a mediano o a largo plazo. Puede contener respuestas como las siguientes: “Para conseguir trabajo”, “Para aprender a lo que te vas a dedicar”, “Para ser alguien”, etcétera.

La cuarta, *Sirve para ser mejor*, se asigna a respuestas que denotan una relación estrecha entre la utilidad que los menores ven en la escuela y la mejora de sí mismos a través de la educación. Algunos ejemplos son: “Para ser mejor persona”, “Para ayudar a otras personas”, “Para pensar mejor”, etcétera.

La quinta categoría, *Sirve para ser exitoso*, engloba respuestas que se refieren a la búsqueda de crecimiento personal a partir de la asistencia a la escuela. Algunos ejemplos de este tipo de respuesta son: “Para conseguir tus sueños”, “Para resolver problemas”, “Para no depender de nadie”, etcétera.

La sexta, *Sirve para conocerme*, expresa una función de la escuela que los menores asocian al descubrimiento de las propias potencialidades y limitaciones, generalmente a partir de poner a prueba sus propias capacidades. En esta categoría se encuentran respuestas como las siguientes: “Para saber en qué soy bueno”, “Para establecer metas”, “Para ver en qué soy hábil”, etcétera.

La séptima, *Sirve para convivir*, engloba respuestas que denotan casi exclusivamente la función social de la institución. Por ejemplo: “Para hacer amigos”, “Para convivir”, “Porque te premian/castigan”, etcétera.

Por último, la octava categoría, *No sé para qué sirve*, tal como el nombre sugiere se asigna a quienes dicen no saber para qué sirve la institución.

MEDICIÓN 2016

El conjunto de datos de la medición 2016 tiene un total de 820,512 valores (omitiendo FolioHogarGeneral y FolioNiñoGeneral), provenientes de 96 variables y 8547 encuestados de los estados de Veracruz, Tabasco y Campeche.

Las variables que fueron categorizadas en esta ronda de evaluación son OcupaciónM, OcupaciónP, LecturasAdicionales, RoboEscuela, RoboCalle, AgresiónEscuela, AgresiónCalle, AmenazasEscuela, AmenazasCalle, BurlaEscuela, BurlaCalle, MiedoEscuela y MiedoCalle. Parentesco, Lengua, ViveOtro, ActividadRemunerada, RazónAbandono y UtilidadEscuela también forman parte de este conjunto de datos, pero recibieron un tratamiento idéntico al expuesto en apartados anteriores, por lo que la exposición se concentrará en las enunciadas inicialmente.

21

OcupaciónM y OcupaciónP

La codificación de estas variables se basó, aunque no fielmente, en el Sistema Nacional de Clasificación de Ocupaciones (SINCO) del Instituto Nacional de Estadística y Geografía (INEGI).

Tales códigos fueron empleados durante la ronda de medición de 2015, y fueron replicados para categorizar las respuestas a las variables OcupaciónM y OcupaciónP de la ronda de medición de 2016. Para tal objetivo, previamente se incluyeron categorías adicionales con el fin de agrupar las respuestas abiertas de tipo cadena de la opción de respuesta 15. Otra, incluida en la versión de 2015. Asimismo, se modificaron algunos de los valores originales dando por resultado la siguiente lista de valores y categorías:

0. Desempleado(a).
1. Labores del hogar.
2. Empleado(a) doméstico(a).
3. Empleado(a).
4. Ocupaciones elementales.
5. Operadores(as) de instalaciones y maquinaria pesada.

6. Obreros(as), operarios(as) y artesanos(as).
7. Agricultores(as) y trabajadores(as) calificados(as) agropecuarios(as).
8. Trabajadores(as) de servicios y vendedores(as) en comercios y mercados.
9. Personal de apoyo administrativo.
10. Técnico(a) profesional nivel medio.
11. Profesional, científico(a) e intelectual.
12. Director(a) o gerente.
13. Jubilado(a) o pensionado(a).
14. Militar, policía.

Originalmente, el valor 3 correspondía a la categoría *Estudiante*, sin embargo, esta categoría se incluyó dentro del valor 0 con el objetivo de contener dentro del valor 3 a todos los encuestados que respondieron llanamente “Empleado(a)” y que generaba un problema en tanto que la categorización dentro de las opciones disponibles se volvía inviable. A este valor también se le sumaron los Trabajadores independientes y Trabajadores de empresas y compañías, siempre que no especificaran qué tipo de tarea realizan.

Por otro lado, el valor 0 incluye entonces tanto a Desempleados(as) como estudiantes y emigrantes. Al valor 7 se le agregó la ocupación Pescadores. Al valor 9 los Agentes municipales, Diligenciaros y Servidores públicos. Al valor 12 los Supervisores. Y, por último, al valor 14 los Marinos.

Asimismo, la cuarta categoría, *Ocupaciones elementales*, incluye ocupaciones como las siguientes: Peones, Jornaleros, Ambulantes, Albañiles, Ayudantes de oficios, Taxistas, Bodegueros, Carniceros, Electricistas, Soldadores, Veladores, Fumigadores, Recamaristas, entre otras.

La sexta categoría, *Obreros(as), operarios(as) y artesanos(as)*, también incluye las siguientes ocupaciones: Empleados de PEMEX, Herreros, Verificadores de embarque, Zapateros, Trabajadores de fábricas, Empleados operarios de seguridad pública, entre otras.

La octava categoría, *Trabajadores(as) de servicios y vendedores(as) de comercios y mercados*, también incluye las ocupaciones de la lista siguiente: Panaderos, Mecánicos, Arrendadores, Contratistas, Estilistas, Modistos, Diseñadores gráficos, Seguridad privada, Volqueteros, Agentes de seguros, Meseros, Pastores, Fotógrafos, Técnicos sin educación superior, Músicos sin educación superior, Bomberos, entre otras.

Por último, en la décima categoría, *Técnico(a) profesional nivel medio* también se incluye la siguiente lista de ocupaciones: Enfermeros, Buzos soldadores, Técnicos en sistemas, Asistentes educativos, Asistentes de guarderías, Trabajadores sociales sin formación superior, Periodistas sin formación superior, entre otras.

23

Lecturas Adicionales

Esta variable se desprende de preguntarle a los(as) participantes si realizan lecturas adicionales a las de los libros de texto escolares. Sin ser necesariamente exhaustivas, a partir de las respuestas sobre el tipo de lecturas que realizan, se crearon las siguientes categorías:

1. Libros interactivos
2. Cuentos y fábulas
3. Leyenda y mitología
4. Material gráfico
5. Novela infantil
6. Novela
7. Enciclopedias y conocimiento
8. Otro

La distinción entre una y otra categoría no necesariamente obedece a criterios predefinidos y se basan principalmente en formato, extensión y contenido. Así, por ejemplo, los *Libros interactivos* se presentan en un formato que de una u otra forma interactúa con el lector, con una extensión relativamente breve y contenido variable. Este tipo de texto puede perseguir, de forma didáctica, enseñar algo a los

niños o, simplemente, incitarlos a realizar alguna actividad. Pero también pueden ser lúdicos como los libros de adivinanzas o de chistes.

Por otro lado, la distinción entre *Cuentos y fábulas*, y *Leyendas y mitología*, pese a que sus características son similares en términos de formato y extensión, es la dimensión histórica del contenido de las segundas ya que mientras que los cuentos y las fábulas tienen un objetivo esencialmente lúdico y didáctico, las leyendas y la mitología generalmente forman parte de la cosmovisión de los pueblos y están basados en sus mitos fundacionales, por lo tanto, al acceder a ellos, más o menos de forma consciente, los lectores también se están vinculando con diferentes culturas, aprendiendo de ellas, y transformando su propia manera de ver el mundo.

Los textos integrados dentro de la categoría *Material gráfico* son aquellos que, si bien pueden contener narrativa breve y ser generalmente de temática accesible, su formato es predominantemente visual. Algunos ejemplos son las revistas, los cómics, entre otros.

Caso aparte han sido las categorías que distinguen la *Novela infantil* de la *Novela*. Esto debido a que en vista de los involucrados cierta cantidad de textos que históricamente han sido considerados parte de la literatura infantil y juvenil como Alicia en el País de las Maravillas, El principito o Harry Potter, exceden las limitaciones de contenido o extensión que las clasificarían como novelas infantiles. En cambio, las actualmente consideradas novelas como Crepúsculo, Los juegos del hambre o Maze Runner, si bien por extensión entrarían en la categoría novela, el contenido las hace parecer dirigidas más bien a un público juvenil.

Para salvar esta dificultad, se suavizó el criterio de contenido y se matizó el de extensión. Sin embargo, la categoría *Novela* también incluyó, pese a su extensión, todas aquellas obras que son consideradas como literatura clásica. Entonces, algunos ejemplos de novela infantil y juvenil son: A golpe de calcetín, Buscando a Dory, Hasta el domingo, Cicatrices, entre otros títulos. Y algunos ejemplos de novela son: ¿Y si quedamos como amigos?, Alicia en el país de las maravillas, Bajo la misma estrella, El principito, Drácula, entre muchos otros.

En otro tema, la categoría *Enciclopedias* y conocimiento contiene aquellos textos de corte académico, de divulgación del conocimiento, de consulta o investigación científica. Algunos ejemplos mencionados por los encuestados son: Aguas marinas, Enciclopedia de animales, Asquerosología y Cuerpo y salud.

Por último, se incluyó la categoría *Otro* como un recurso para incluir cualquier texto no clasificado por omisión o incapacidad en alguna de las anteriores. Entre éstos se han destacado los textos de poesía y de catecismo.

25

Reporte de Violencia

Por último, el reporte de violencia es una variable incluida en esta ronda de medición con la intención de explorar si los niños(as) y adolescentes encuestados han sufrido o sufren algún tipo de violencia al interior de sus escuelas, o en los espacios públicos, específicamente en la calle. Sobre este aspecto se les preguntó si habían sido víctimas de robo de alguna de sus pertenencias, agresión física, amenazas o burlas constantes. Para este conjunto de variables, se agruparon los agresores de la siguiente forma:

1. Compañeros de grado
2. Amigos
3. Compañeros de grados superiores
4. Familiares
5. Otros niños/jóvenes
6. Vecinos
7. Adultos
8. No sé⁷

Una aclaración que debe hacerse sobre la categoría *Amigos* es que ésta incluyó también, además de todas las respuestas donde explícitamente los encuestados dijeron haber sufrido alguna de las agresiones exploradas por alguno de sus amigos, todas aquellas respuestas que contenían un nombre personal. Evidentemente, en términos de análisis esto no dice más que el encuestado

⁷ Esta categoría sólo aplica para las variables RoboEscuelaCat y RoboCalleCat.

conoce al agresor, y bajo esta mirada se decidió colocar todas esas respuestas dentro de dicha categoría. Si lo desea, el usuario final puede modificar estas categorías tanto como a sus objetivos convenga.

MiedoCalle y MiedoEscuela

Como parte adicional del Reporte de violencia, se indagó si acaso los niños(as) y adolescentes encuestados sienten algún tipo de temor a estar en la calle o asistir a sus escuelas. Las respuestas resultantes de valores cadena dieron lugar a las siguientes categorías:

0. No reporta sentir miedo
1. Por causas internas atribuibles al niño
2. Por causas internas no atribuibles al niño
3. Por causas externas atribuibles a eventos concretos
4. Por causas externas no atribuibles a eventos concretos
5. No sabe por qué

Fuera de la categoría en el valor 0 y que es completamente transparente, la primera, *Por causas internas atribuibles al niño*, hace referencia a inseguridades o temores cuya raíz son eventos concretos y que dependen de la actuación del menor. Ejemplos de respuestas incluidas en esta categoría son: “porque olvido mis útiles”, “porque no me gustan las matemáticas” y “porque no puedo hacer las cosas que me piden”.

La segunda, *Por causas internas no atribuibles al niño*, se refiere a temores por situaciones que son imaginadas o atribuidas por el infante a agentes externos. Por ejemplo: “porque siento que algo malo me va a pasar”, “porque no me gusta que mi mamá se vaya” y “porque siento que me pueden robar”.

La tercera, *Por causas externas atribuibles a eventos concretos*, corresponde a temores justificados sobre situaciones más o menos concretas que pueden o no estar en el entorno. Algunos ejemplos pueden ser: “porque vengo solo(a) y el camino es muy solitario”, “porque los coches pasan muy rápido” y “porque me pegan”.

La cuarta, *Por causas externas no atribuibles a eventos concretos*, agrupa los temores originados por situaciones abstractas que el niño atribuye a partir de la información que obtiene del medio. Ejemplos de esta categoría son: “porque me pueden robar”, “por la inseguridad” y “porque pasan noticias de secuestros”.

También se incluye una quinta categoría que engloba aquellas respuestas donde los encuestados dan una respuesta afirmativa a sentir temor en alguna o ambas situaciones exploradas, pero no saber referir la causa.



La revisión de este documento se concluyó el 24 de septiembre de 2018.

El contenido aquí vertido es responsabilidad de los miembros del
Departamento de Análisis y Calidad de Datos.