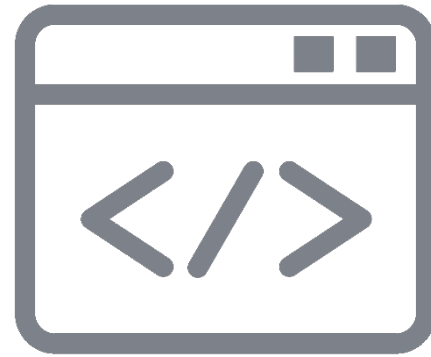


# Reporte de indicadores de precisión y confiabilidad en los metadatos DDI



Tercera sesión 2018

Comité de Aseguramiento de la Calidad

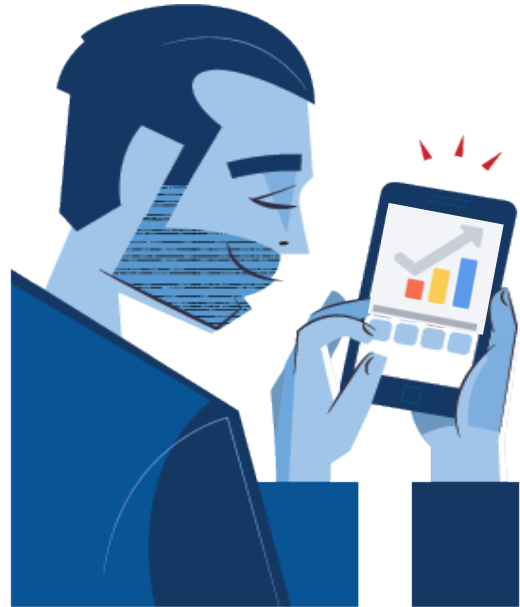
Agosto 14, 2018



# 1. Motivación



# ¿Para qué?



- El fin último de la información producida por las ONE es la **satisfacción de las necesidades de los usuarios**, que está relacionada con la facilidad en el **uso e interpretación de los datos**.
- Los usuarios deben poder **evaluar** hasta qué punto **la precisión de los datos** y otros atributos de calidad son consistentes con el uso previsto de la información.
- La publicación de metadatos muestra **apertura y transparencia, y promueve la confianza**.
- Es importante apegarse a los principios para su documentación y **mantener los metadatos actualizados**.

[Metadata an Integral Part of Statistics Canada's Data Quality Framework](#)



# ¿Por qué?



- Cumplimiento normativo eficiente
  - Norma de Aseguramiento de la Calidad: Reportes de calidad de cada programa
  - Norma para la Elaboración de Metadatos: Evaluación estadística de la calidad de los datos
- Meta estratégica:
  - **PESNIEG**: 95% de los **programas de IIN incluye en sus metadatos indicadores de calidad** de los productos basados en los Principios y Buenas Prácticas
    - **Avance 2018**: Todos los proyectos con muestreo **publicados a partir del segundo trimestre de 2018** reportan de manera estandarizada los indicadores de calidad aprobados por el CAC en sus metadatos.

# Compromisos PAEG



- **INEGI-025-DGIAI:** *Dar seguimiento al reporte de los indicadores de calidad aprobados por el Comité de Aseguramiento de la Calidad del INEGI.*
  - En el informe de resultados 2018 del Comité de Aseguramiento de la Calidad del INEGI se analizará el cumplimiento del reporte de los indicadores de calidad aprobados por el Comité.
- **INEGI-023-DGEE:** *Contar con Indicadores de precisión y confiabilidad para evaluar la calidad de la información estadística.*
  - Calcular los Indicadores de precisión y confiabilidad para evaluar la calidad de la información estadística de proyectos en unidades económicas según aplique:
    - Coeficiente de variación, error estándar, intervalo de confianza y tasa de no respuesta antes de imputación a nivel de unidad de observación: EMEC, EMS, EMIM, ENEC, EMOE, EAC, EAEC, EASPNF, EAT, EAIM y ENA.
    - Cobertura de la variable de diseño: EMEC, EMIM, EMS, EAC, EAIM, EASPNF y EAT.

## 2. Reporte de indicadores de calidad en metadatos





# European Statistical System Metadata Structures



- **Euro SDMX Metadata Standards (ESMS)**

- [ESMS. Production in industry. Czech Statistical Office](#)
- [ESMS. Production in construction. Statistics Netherlands](#)

- **Single Integrated Metadata Structure (SIMS)**

- **ESS Standard for Quality Report Structure (ESQRS)**

- [Quality Report of Structure of Earnings Survey in the Netherlands](#)
- [Quality Report for the MIDSS \(Monthly Inquiry into the Distribution and Services Sector\) Survey. Office for National Statistics](#)



EUROPEAN  
STATISTICAL  
SYSTEM



# Especificación del estándar DDI versión 2.0



Facilitan la posterior búsqueda, recuperación y explotación de la información



# Usos de los metadatos DDI por tipo de usuario



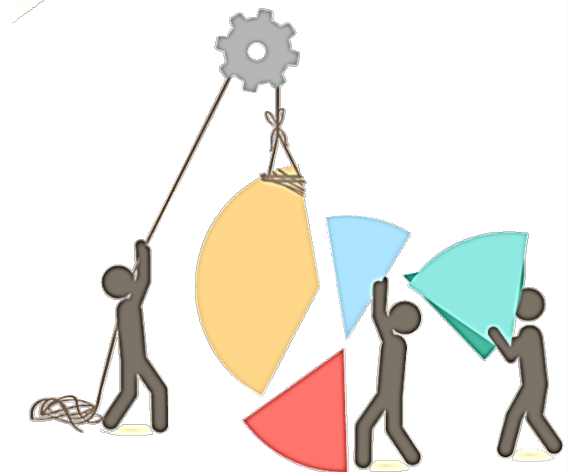
## Usuarios:

- Buscar y consultar programas estadísticos relevantes para sus necesidades de información
- Interpretar, comprender y analizar los datos
- Replicar o calcular información a partir de los procesos y métodos descritos en los metadatos

## METADATOS DE REFERENCIA Y ESTRUCTURALES

### Productores

- Evaluar la calidad en general de los programas estadísticos
- Comparar las prácticas nacionales en la compilación de estadísticas
- Validar la consistencia de los conjuntos de datos



# Plantilla de metadatos DDI INEGI\_2.0



## Evaluación estadística de calidad de datos

- Descripción de estimadores de precisión de los datos del proyecto estadístico, tales como: errores de muestreo; coeficientes de variación, intervalos de confianza, etc.
- Para el caso de los errores de muestreo se deberá proporcionar la siguiente información:
  - Las estimaciones para las cuales fueron calculados los errores muestrales.
  - Referencia a reportes u otro tipo de documentos donde se pueden encontrar los resultados.
  - Detalles referentes al software usado para calcular el error muestral y una referencia a los programas usados para realizar los cálculos.

### 3. Cumplimiento de acuerdos del CAC

---



# Acuerdos del CAC



- **Acta de la tercera sesión 2017 ( 30-10-2017)**

Como resultado de la discusión, por unanimidad se aprobaron los 12 indicadores propuestos, conforme a las especificaciones de las fichas técnicas mediante el acuerdo CAC-004/03/2017;

- **Acta de la cuarta sesión 2017 (12-12-2017)**

Asimismo, mediante el acuerdo CAC-006/04/2017, el Comité aprobó que el valor del CV sea publicado, junto con el error estándar y el intervalo de confianza, conforme a las especificaciones de las fichas técnicas aprobadas, para permitir a los usuarios especializados realizar su propia interpretación.

Por su parte, Enrique de Alba Guerra opinó que el cálculo de los indicadores de tasa de no respuesta debe ser de uso interno. Como resultado de la discusión, mediante el acuerdo CAC-008/04/2017 se aprobó el cálculo para uso interno de la tasa de no respuesta antes de imputación a nivel unidad de observación para los proyectos con muestreo, conforme a las especificaciones de la nueva ficha técnica.



# Implementación de los acuerdos del CAC



- Fichas técnicas aprobadas

3.5 Medios de difusión	
3.5.1 Indique los medios de difusión externa y/o interna del indicador, según sea el caso, así como el lugar donde podrá consultarse:	
Medio de difusión <i>Ejemplo: Apartado de indicadores de calidad de la ficha de metadatos</i>	Lugar de consulta <i>Ejemplo: Página Web del INEGI (incluir URL)</i>
Difusión externa: Página Web del INEGI, según proyecto	<a href="http://www.inegi.org.mx/default.aspx">http://www.inegi.org.mx/default.aspx</a>
Difusión Externa: Apartado de Evaluación estadística de la calidad de los datos en la Red Nacional de Metadatos para cada proyecto	<a href="http://www3.inegi.org.mx/rnm/index.php/catalog/central">http://www3.inegi.org.mx/rnm/index.php/catalog/central</a>



- Formatos para el reporte de los indicadores en los metadatos

- Ejemplos de indicadores publicados

- Encuesta Nacional de Seguridad Pública Urbana 2018, Datos para el primer y segundo trimestre del año.
- Encuesta Nacional sobre Confianza del Consumidor 2018, Datos al mes julio



# Reporte de indicadores en metadatos



- Enfocado en usuarios especializados
- Uso de formato homologado para estandarizar el reporte
- Entrega de indicadores 30 días naturales después de la fecha de publicación de resultados comprometida en el calendario
- Acceso restringido a **los indicadores de uso interno.**



# Programas/ciclos con indicadores entregados y publicados

Programa	Fecha de publicación de resultados	¿Se entregaron los indicadores para metadatos?	¿Se publicaron los metadatos?
<b>DGEGSPyJ</b>			
ENSU, Primer trimestre 2018	18 de abril	SI 17 de abril	SI 18 de abril
ENSU, Segundo trimestre 2018	18 de julio	SI 16 de julio	SI 18 de julio
<b>DGES</b>			
ENCO, mayo 2018	3 de agosto	SI 1° de agosto	SI 3 de agosto
ENCO, mayo 2018	5 de junio	SI 4 de junio	SI 5 de junio
ENCO, junio 2018	5 de julio	SI 4 de julio	SI 5 de julio
Modulo de Trabajo Infantil 2017	12 de junio	SI* 4 de julio	Falta definir fecha de entrega
<b>DGGMA</b>			
Módulo de Hogares y Medio Ambiente (MOHOMA) 2017	5 de junio	SI 29 de junio	En proceso de revisión

\*Pendiente Tasa de no respuesta antes de imputación a nivel de unidad de observación (uso interno).



¡ Muchas Gracias !



# Anexo I: Visualización de la Red Nacional de Metadatos



# Red Nacional de Metadatos



## Encuesta Nacional de Seguridad Pública Urbana 2018, Primer Trimestre Enero - Marzo 2018

Información del proyecto estadístico

Diccionario de Datos

Material de Referencia

Información general

Muestreo

Recolección de Datos

Procesamiento de  
Datos

**Evaluación de  
datos**

Política de Acceso a  
los microdatos

### Indicadores de la Calidad/Evaluación de la Calidad

#### Evaluación estadística de la calidad de los datos

A través de los siguientes vínculos se pueden consultar los indicadores de calidad sobre precisión y confiabilidad estadística aprobados por el Comité de Aseguramiento de la Calidad del INEGI.

Los valores de los indicadores de calidad se presentan para las encuestas publicadas a partir del segundo trimestre del 2018.

- Indicadores para proyectos con muestreo probabilístico:

#### Coeficiente de variación, error estándar e intervalo de confianza

[ENSU 2018. Primer trimestre](#)

[Metadatos de las tablas de indicadores](#)

¿Quieres abrir o guardar **IndicadoresPrecisionEstadistica\_CV-EE-IC\_ENSU-2018-I.xls** (109 KB) desde [www3.inegi.org.mx](http://www3.inegi.org.mx)?

Abrir

Guardar

Cancelar



# Indicadores disponibles como Materiales de Referencia Externos



Información del proyecto estadístico

Diccionario de Datos

Material de Referencia

## - Otros Materiales

Indicadores de precisión y confiabilidad estadística



106.48 KB

Metadatos de las tablas de indicadores



136.8 KB

[http://www3.inegi.org.mx/rnm/index.php/catalog/327/related\\_materials?idPro=](http://www3.inegi.org.mx/rnm/index.php/catalog/327/related_materials?idPro=)

# Resguardo de los indicadores de uso interno



<http://www3.inegi.org.mx/rnm/index.php/auth/login/?destination=admin>

**Ingresar**

Dirección de correo electrónico:

Contraseña:

**Iniciar sesión**

[Registrar nuevo usuario](#) | [¿Olvidó su](#)

[contraseña?](#)

Administrar Archivos 1

Recursos Externos 9

Citas 0

Notas 0

[Cargar archivos](#)

<input type="checkbox"/>	Nombre	Tamaño	Permisos	Modificado	Acciones
<input type="checkbox"/>	IndicadoresPrecisionEstadistica_CV-EE-IC_ENSU-2018-I.xls	109.5 KB	-rw-rw-rw-	06/07/2018: 22:25:11	
<input type="checkbox"/>	IndicadoresPrecisionEstadistica_TNR- AntesImputacion_ENSU_2018_1ER_TRIM.xlsx	14.71 KB	-rw-rw-rw-	06/07/2018: 22:25:13	
<input type="checkbox"/>	Metadatos_Indicadores_Encuestas.pdf	141.09 KB	-rw-rw-rw-	06/08/2018: 17:38:38	
<input type="checkbox"/>	MEX-INEGI-ENSU-2018.xml	461.06 KB	-rw-rw-rw-	06/08/2018: 17:39:18	

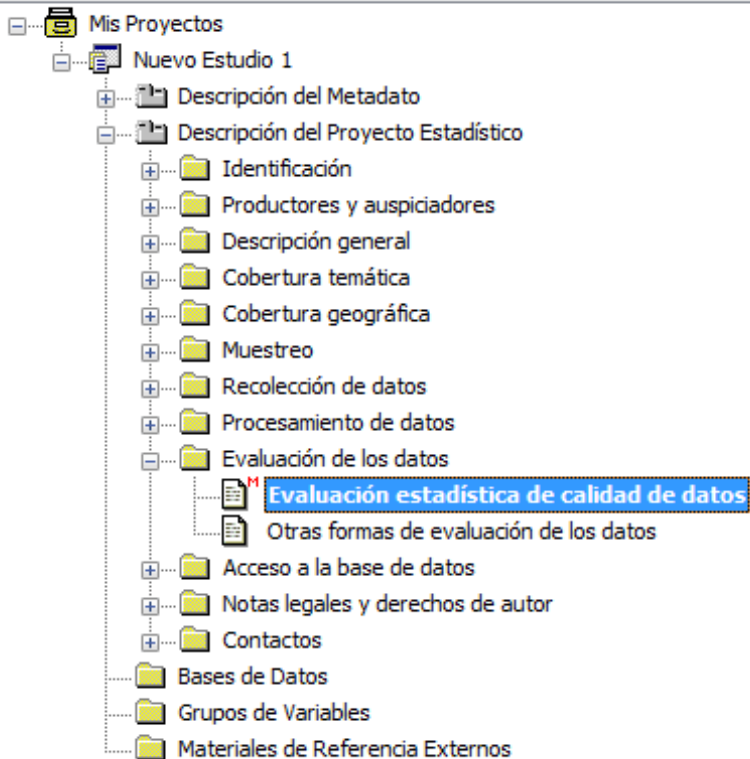
No vinculado Archivo de Datos Otros recursos

Total de archivos: 4

# Captura en el editor de metadatos antes de enviar a DASC



Proyectos:



## Evaluación estadística de calidad de datos

- Las UA responsables deben dejar este elemento en blanco
- El DASC incorpora el texto fijo acordado

### Descripción del elemento:

Descripción de estimadores de precisión de los datos del proyecto estadístico, tales como: errores de muestreo; coeficientes de variación, intervalos de confianza, etc.

Para el caso de los errores de muestreo se deberá proporcionar la siguiente información:

- Una lista de razones/indicadores para los cuales fueron calculados los errores debidos a la muestra.
- Detalles referentes al software usado para calcular el error muestral y una referencia a los programas usados para realizar los cálculos.

Referencia a reportes u otro tipo de documentos donde se pueden encontrar los resultados.



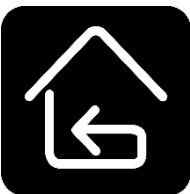
# Anexo II: Bibliografía experiencia internacional



# Referencias internacionales sobre el reporte de los principios de calidad en los metadatos



- **Metadata: an Integral Part of Statistics Canada's Data Quality Framework**
  - <http://www.stats.gov.cn/english/icas/papers/p020071114297538592238.pdf>
- **Metadata management at Statistics Canada**
  - [https://www.census.gov/fedcas/cfc2007/ppt/wp4\\_born.pdf](https://www.census.gov/fedcas/cfc2007/ppt/wp4_born.pdf)
- **The Role of Metadata in Statistics**
  - <https://www.bls.gov/ore/pdf/st000040.pdf>
- **Statistical Metadata: a Unified Approach to Management and Dissemination**
  - <https://content.sciendo.com/view/journals/jos/31/2/article-p325.xml>
- **Quality and Metadata at Statistics Denmark**
  - <https://www.dst.dk/ext/335729722/0/jordan/3-7-PPT-Metadata-at-Statistics-Denmark--pdf>
- **Metadata as a tool to improve quality of the statistical production process**
  - <https://www.stat.fi/isi99/proceedings/arkisto/varasto/froe0738.pdf>
- **UN Statistical Commission international recommendations for industrial and distributive trade statistics. Chapter IX. Data quality and metadata**
  - <https://unstats.un.org/oslogroup/meetings/og-04/docs/oslo-group-meeting-04--escm-ch09-draft1.pdf>
- **Metadata, the DNA of statistical data. Institute of Statistics, INSTAT**
  - [http://www.instat.gov.al/media/1497/metadata\\_the\\_dna\\_of\\_statistical\\_data.pdf](http://www.instat.gov.al/media/1497/metadata_the_dna_of_statistical_data.pdf)
- **ESS. Reference metadata in business statistics**
  - <http://ec.europa.eu/eurostat/documents/54610/7779382/Reference-metadata-in-business-statistics.pdf>
- **Role of Metadata in Quality Assurance of Multi-Country Statistical Data In the case of UNIDO Industrial Statistics**
  - <https://unstats.un.org/unsd/ccsa/cdqio-2004/3i-unido.pdf>





# European Statistical System Metadata Structures

ESMS (SDMX)	SIMS	ESQRS (Reportes de calidad)
11. Quality management <ul style="list-style-type: none"> <li>• 11.1 Quality Assurance</li> <li>• 11.2 Quality assessment</li> </ul>	S.11. Quality management <ul style="list-style-type: none"> <li>• S.11.1 Quality Assurance</li> <li>• S.11.2 Quality assessment</li> </ul>	4. Quality management <ul style="list-style-type: none"> <li>• 4.1 Quality Assurance</li> <li>• 4.2 Quality assessment</li> </ul>
<b>13 Accuracy and reliability</b>	<b>S.13 Accuracy and reliability</b>	<b>6 Accuracy and reliability</b>
13.1 Overall accuracy	S.13.1 Overall accuracy	6.1 Overall accuracy
13.2 Sampling error Quality indicators: A1. The coefficient of variation (or relative standard error)	S.13.2 Sampling error <ul style="list-style-type: none"> <li>• A1. Sampling errors - indicators</li> </ul>	6.2 Sampling error
13.3 Non-sampling error Quality indicators: <ul style="list-style-type: none"> <li>• A2. Rate of overcoverage</li> <li>• A3. Edit failure rate</li> <li>• A4. Unit response rate</li> <li>• A5. Item response rate</li> <li>• A6. Imputation rate</li> <li>• A7. Number of mistakes made, by type</li> </ul> <a href="http://ec.europa.eu/eurostat/documents/64157/4373903/SI-MS-2-0-Revised-standards-November-2015-ESSC-final.pdf/47c0b80d-0e19-4777-8f9e-28f89f82ce18">http://ec.europa.eu/eurostat/documents/64157/4373903/SI-MS-2-0-Revised-standards-November-2015-ESSC-final.pdf/47c0b80d-0e19-4777-8f9e-28f89f82ce18</a>	S13.3 Non-sampling error <ul style="list-style-type: none"> <li>• A2. Over-coverage – rate</li> <li>• A3. Common units - proportion</li> <li>• A4. Unit non-response - rate</li> <li>• A5. Item non-response - rate</li> <li>• Coverage error</li> <li>• Measurement error</li> <li>• Non response error</li> <li>• Processing error</li> <li>• Model assumption error</li> </ul>	6.3 Non-sampling error <ul style="list-style-type: none"> <li>• Coverage error</li> <li>• Over-coverage – rate</li> <li>• Common units - proportion</li> <li>• Measurement error</li> <li>• Non response error</li> <li>• Unit non-response - rate</li> <li>• Item non-response - rate</li> <li>• Processing error</li> <li>• Imputation - rate</li> <li>• Model assumption error</li> <li>• Seasonal adjustment</li> </ul>



# Anexo III: Mantenimiento de la vigencia de los metadatos



# Consideraciones sobre la actualización de los metadatos



- La **accesibilidad** de la información implica que los programas de información estadística y geográfica estén **acompañados de metadatos estandarizados y vigentes**.
- Independientemente de la publicación de indicadores de precisión y confiabilidad, todos los programas con periodicidad menor a la anual requieren **actualizar sus metadatos en cada levantamiento**.
- Se recomienda realizar **un solo documento de metadatos** para el programa estadístico, en el cual, **se deberán actualizar los elementos que así lo requieran derivado del cambio de estatus de los datos** (preliminares a definitivos).

# Programas con metadatos actualizados de conformidad a los ciclos publicados



## DGEE

- [Balanza Comercial de Mercancías de México 2018, información revisada al mes de mayo.](#)
- [Estadística de Sacrificio de Ganado en Rastros Municipales 2018, Datos al mes de mayo](#)
- [Estadística del Programa de la Industria Manufacturera, Maquiladora y de Servicios de Exportación 2018, Datos al mes de mayo](#)
- [Estadística de la Industria Minerometalúrgica 2018, Datos al mes de mayo](#)
- [Directorio Estadístico Nacional de Unidades Económicas 2018, Información para la actualización e incorporación de unidades económicas al DENUE; datos a marzo de 2018](#)

## DGES

- [Encuesta Nacional sobre Confianza del Consumidor 2018, Datos al mes junio.](#)

## DGEGSPJ

- [Encuesta Nacional de Seguridad Pública Urbana 2018, Datos para el primer y segundo trimestre del año.](#)

# Elementos susceptibles de actualización permanentemente



## 1. Título y Subtítulo

- Proyecto EMX 2018
  - Datos al mes de marzo
  - Primer trimestre

## 2. Tamaño de la muestra\*

## 3. Desviaciones del diseño muestral\*

## 4. Tasa de respuesta\*

Ejemplo: ENCO, junio 2018

<http://www3.inegi.org.mx/rnm/index.php/catalog/336#page=sampling&tab=study-desc>

*\* La actualización de estos elementos no aplica para Censos y Registros administrativos*



# Elementos susceptibles de actualización permanentemente



## 5. Periodo de ejecución del proyecto estadístico

### Periodo de ejecución del proyecto estadístico

Inicio	Fin	Período
2018-05-02	2018-06-05	Mayo
2018-05-02	2018-05-18	Levantamiento
2018-05-04	2018-05-22	Crítica y supervisión
2018-05-02	2018-05-23	Descarga, validación y envío de archivos a Aguascalientes Semana 1
2018-05-24	2018-05-24	Revisión de archivos
2018-05-25	2018-05-25	Entrega de base de datos a diseño estadístico
2018-05-28	2018-05-29	Recepción de factores de expansión
2018-05-30	2018-05-30	Entrega de base de datos y documentos
2018-06-05	2018-06-05	Difusión del índice de confianza del consumidor
2018-06-01	2018-07-05	Junio
2018-06-01	2018-06-19	Levantamiento
2018-06-05	2018-06-21	Crítica y supervisión
2018-06-01	2018-06-21	Descarga, validación y envío de archivos a Aguascalientes Semana 1
2018-06-22	2018-06-22	Revisión de archivos
2018-06-25	2018-06-25	Entrega de base de datos a diseño estadístico
2018-06-26	2018-06-27	Recepción de factores de expansión
2018-06-29	2018-06-29	Entrega de base de datos y documentos
2018-07-05	2018-07-05	Difusión del índice de confianza del consumidor

## 6. Periodo de referencia

### Periodo de referencia

Inicio	Fin	Período
2018-05-01	2018-05-31	Mayo
2018-06-01	2018-06-30	Junio



# Elementos susceptibles de actualización permanentemente



## 7. Evaluación estadística de la calidad de datos\*

- El **Comité de Aseguramiento de la Calidad** acordó la **publicación de los indicadores de calidad estadística aprobados hasta el momento** (EE, CV, IC y cobertura de la variable de diseño) **como parte de los metadatos DDI** de las encuestas con muestreo probabilístico y no probabilístico a través de esta apartado.



# Elementos susceptibles de actualización permanentemente



## Descripción de los archivos de datos

### 8. Nombre

- EMX\_2018\_enero
- EMX\_2018\_febrero

### 9 . Contenido

Información sobre... captada por la EMX al mes de enero de 2018.

- Variables: 43
- **Casos validos: 75,546**

Información sobre... captada por la EMX al mes de febrero de 2018.

- Variables: 43
- **Casos validos: 76,876**



# Elementos susceptibles de actualización permanentemente



## 10. Descripción de las variables

- Se debe contar con documentación de todas las variables de cada archivo de microdatos de los ciclos de programa que se vayan incorporando.
- **Los metadatos** a nivel de variable **pueden “reciclarse”** a partir de la documentación del primer ciclo del programa o incluso de una documentación de un periodo anterior, **reemplazando la matriz de datos** para que los metadatos sean consistentes con los valores de la base del mes, trimestre o semestre que corresponda.



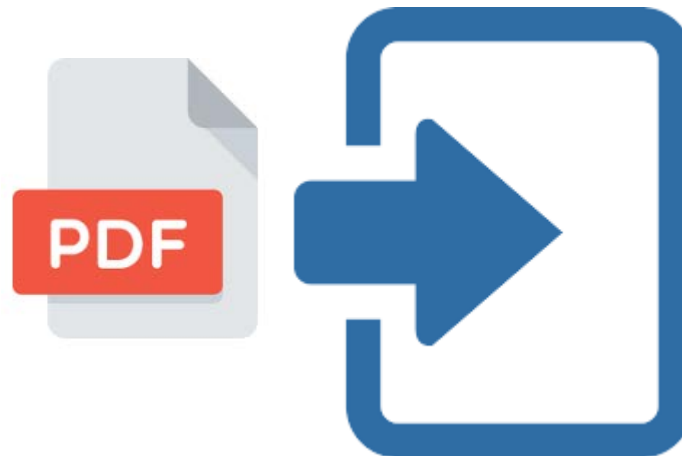


# Elementos susceptibles de actualización permanentemente



## 11. Descripción de los Materiales de Referencia Externos

- De ser necesario, deberán añadirse y documentarse los materiales de referencia (cuestionarios, instructivos, clasificadores, manuales), derivado de la incorporación archivos de microdatos de los periodos de referencia que se vayan actualizando.





# Elementos susceptibles de actualización permanentemente



**12. Fecha de conclusión de la documentación**

**13. Versión**

**14. Número de identificación del documento de metadatos**

Estos campos deberán actualizarse cada vez que se modifique o actualice la documentación del programa estadístico, la base de datos o el material de referencia externo.

