



Criterios para la estructura de bases de datos abiertos

Formatos CSV, JSON, GPKG, GeoTIFF

Este documento contiene los criterios técnicos determinados por la Dirección de Datos Abiertos (DDA) de la Agencia de Transformación Digital y Telecomunicaciones (ATDT) para publicar datos abiertos en la Plataforma Nacional de Datos Abiertos.

Última actualización 21-08-2026.

Contenido

CSV

1. Conceptos importantes

// Apoyos visuales que identifican los conceptos utilizados a lo largo del listado de criterios

2. Criterios generales del archivo (criterios 1 a 7)

// características de codificación y errores a evitar en la estructura general

3. Criterios para los datos

// criterios categorizados en criterios para columnas, filas y tipo de dato explicados con ejemplos visuales

- **Columnas (criterios 8 a 20)**

- estructura (8 a 12)

- nombres (13 a 20)

- **Filas (criterios 21 a 24)**

- **Por tipo de dato (criterios 25 a 56)**

- estructura (25 a 28)

- número (29 a 36)

- texto (37 a 46)

- fecha (47 a 53)

- coordenada (54 a 56)





JSON

1. Criterios para archivo (criterios 1 a 4)
2. Criterios para nombres (criterios 5 a 7)
3. Criterios para elementos (criterios 8 a 13)

GPKG

1. Criterios para archivo (criterios 1 a 2)

GeoTIFF

1. *Open Geospatial Consortium (OGC)*





CSV

1. Conceptos importantes

Las partes básicas que componen el archivo tabular (CSV) son:

COLUMNA		A	B	C	D	E
FILA	1					
	2					
	3					
	4					
	5					
	6					
	7					
	8					
	9					
	10					
	11					

Diferencias entre criterios de texto en las bases de datos CSV

NOMBRE DEL ARCHIVO (CRITERIO 4)

- Sin caracteres especiales.
- Palabras clave y temporalidad.

NOMBRE DE COLUMNAS (CRITERIOS 14, 15)

- Sin caracteres especiales.
- Minúsculas separadas con guión bajo.

inmuebles_INDEP_1er_trim_2026.csv				
	A	B	C	D
1	nombre_inmueble	municipio	entidad_federativa	superficie_terreno
2	Promotora Francisco Villa	Comondú	Baja California Sur	1302.86
3	Sucursal de Correos Tlahualilo	Tlahualilo	Durango	350.127
4	Tienda No. 103 Manzanillo	Manzanillo	Colima	3770.12
5	Lote 16 Segunda Sección	León	Guanajuato	180
6	Patio Fiscal	Suchiate	Chiapas	1784.63
7	Distrito de Riego 4 Don Martín	Anáhuac	Nuevo León	20028.84
8	Tienda No. 228 Chilpancingo	Chilpancingo	Guerrero	3370.3
9	Terreno Paseo del Ángel	Tecámac	Estado de México	635.68
10	Concentrador Valle	Benito Juárez	Ciudad de México	250
11	Predio en Reserva Tizayuca	Tizayuca	Hidalgo	4253.34
12	Lote Miguel Lerdo	Veracruz	Veracruz	135.1
13	Productora Nacional de Semillas	Mexicali	Baja California	43164.898
14	Río Amazonas 93 y 93 A	Cuauhtémoc	Ciudad de México	355.3

DATOS (CRITERIOS 37, 38)

- Escritos en correcto español.
- Se permiten caracteres especiales.
- No separar con guión bajo.

miró



2026
año de
Margarita Maza



2. Criterios generales del archivo

Criterios de archivo CSV (1-7)

1 El formato del archivo de datos tabulares debe ser CSV.

El formato de archivo CSV tiene un nivel alto de apertura. **Evitar usar archivos formato XLSX** (desarrollado por Microsoft y cuyo nivel de apertura es medio).

2 La codificación del archivo debe ser UTF8.

Es importante **no mezclar diferentes codificaciones en la misma base de datos** ya que esto provocará que algunos datos se lean bien y otros mal. Este es un error común al copiar datos de diferentes archivos y unificarlos en una sola base. Tanto Excel como Libreoffice tienen opciones para guardar en codificación UTF8.

1	A	B	1	A	B
2	clave	nombre_sustancia	2	clave	nombre_sustancia
3	100-00-5	Acido propanoico, éster etílico	3	100-00-5	Acido propanoico, éster etílico
4	100-01-6	Bencenamina, 4-nitro -	4	100-01-6	Bencenamina, 4-nitro -
5	100-02-7	Fenol, 4-nitro -	5	100-02-7	Fenol, 4-nitro -
6	100-06-1	Etanona, 1 - (4-metoxifenil) -	6	100-06-1	Etanona, 1 - (4-metoxifenil) -
7	100-09-4	Acido benzoico, 4-metoxi-	7	100-09-4	Acido benzoico, 4-metoxi-
8	100-10-7	Benzaldehído, 4 - (dimetilamino) -	8	100-10-7	Benzaldehído, 4 - (dimetilamino) -
9	100-16-3	Hidracina, (4-nitrofenil) -	9	100-16-3	Hidracina, (4-nitrofenil) -
10	100-17-4	Benceno, 1-metoxi-4-nitro -	10	100-17-4	Benceno, 1-metoxi-4-nitro -
11	100-20-9	1,4-Dicloruro de benceno dicarbonilo	11	100-20-9	1,4-Dicloruro de bencenodicarbonilo
12	100-21-0	Acido 1,4-bencendicarboxílico	12	100-21-0	Acido 1,4-bencendicarboxílico
13	100-25-4	Acido nítrico, pentil éster	13	100-25-4	Acido nítrico, pentil éster
14	100-36-7	Acido nítrico, sal de bario (2:1)	14	100-36-7	Acido nítrico, sal de bario (2:1)
15	100-37-8	Óxido de nitrógeno (N2O)	15	100-37-8	Óxido de nitrógeno (N2O)
16	100-38-9	Cloruro de platino (PtCl2)	16	100-38-9	Cloruro de platino (PtCl2)
17	100-39-0	Cloruro de azufre (S2Cl2)	17	100-39-0	Cloruro de azufre (S2Cl2)
18	100-41-4	Acido silícico (H4SiO4), sal de circo	18	100-41-4	Acido silícico (H4SiO4), sal de circo

3 En caso de que las bases de datos sean generadas en un editor de texto, las columnas deben separarse con comas (,) y las filas con salto de línea (enter).

Este criterio no aplica si las bases de datos se generan en Excel o Libreoffice.

4 Los archivos deben nombrarse con palabras clave del recurso de datos al que pertenecen.

Los nombres de los archivos deben:

1. usar únicamente letras y números [ASCII \(a-z, 0-9\)](#),
2. excluir acentos y caracteres especiales,
3. escribirse en minúsculas (salvo en el uso de siglas),
4. no usar artículos o preposiciones,
5. las palabras deben separarse con guión bajo.





Errores comunes a evitar

- 5 La base de datos solo debe contener los datos estructurados. Evitar incluir encabezados con información institucional, notas al pie, comentarios, entre otros.**

Un error común es añadir en las primeras filas información institucional como: logotipo, nombre de la institución, fecha. Tampoco es correcto agregar notas al pie, comentarios, imágenes o algún otro elemento que no forme parte de la estructura de los datos.

- 6 No se deben aplicar criterios gráficos a los datos: combinar celdas, usar colores, dar calidad a líneas, aplicar color y tamaño a textos.**

No se deben aplicar criterios gráficos con el objetivo de codificar información. De ser el caso, evaluar la estructura de la base o hacer dos o más bases de datos.

- 7 Cada base de datos debe estructurarse en una sola hoja. No se deben incluir múltiples hojas con información en el mismo archivo.**

Evitar colocar metadatos, diccionarios de datos o datos relacionados dentro de hojas en el mismo archivo. Cada archivo debe tener una sola base de datos.

Errores comunes al estructurar datos en formato CSV

5

6

7

	A	B	C	D	E	F	G	H
1		GOBIERNO DE MEXICO						
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Datos 1 | Metadatos | Diccionario





3. Criterios para los datos

Criterios para columnas: estructura (8-12)

8 Cada columna debe referirse a un solo atributo y no a información separable (a excepción de fechas o claves compuestas).

Los datos deben estar lo más desagregados posible y cada atributo debe estar en una columna, de ser necesario, se pueden generar varias columnas.

Incorrecto (X)

	A	B	C
1	año	municipio_entidad	tipo_numero_solicitudes
2	2015	León, Guanajuato	Poda y arbolado, 25
3	2015	Manzanillo, Colima	Recolección de residuos, 50
4	2015	Tizayuca, Hidalgo	Recolección de residuos, 36
5	2015	Cajeme, Sonora	Poda y arbolado, 13

Correcto (✓)

	A	B	C	D	E
1	año	municipio	entidad	tipo_solicitudes	numero_solicitudes
2	2015	León	Guanajuato	Poda y arbolado	25
3	2015	Manzanillo	Colima	Recolección de residuos	50
4	2015	Tizayuca	Hidalgo	Recolección de residuos	36
5	2015	Cajeme	Sonora	Poda y arbolado	13

9 De preferencia, los datos deben estructurarse de forma vertical en lugar de horizontal.

Los datos deben desplegarse hacia abajo. Una base de datos estructurada de forma vertical facilita el análisis y tratamiento de los datos.

Incorrecto (X)

	A	B	C	D	E
1	año	municipio	entidad	numero_solicitudes_poda_arbolado	numero_solicitudes_recoleccion_residuos
2	2015	León	Guanajuato	25	0
3	2015	Manzanillo	Colima	0	50
4	2015	Tizayuca	Hidalgo	0	36
5	2015	Cajeme	Sonora	13	0

Correcto (✓)

	A	B	C	D	E
1	año	municipio	entidad	tipo_solicitudes	numero_solicitudes
2	2015	León	Guanajuato	Poda y arbolado	25
3	2015	Manzanillo	Colima	Recolección de residuos	50
4	2015	Tizayuca	Hidalgo	Recolección de residuos	36
5	2015	Cajeme	Sonora	Poda y arbolado	13

10 Cada columna debe contener solo un tipo de dato: textos, números, fechas o coordenadas.

Correcto (✓)

	A	B	C	D
1	fecha	entidad	tipo_solicitud	numero_solicitudes
2	2024-01-03	Guanajuato	Poda y arbolado	125
3	2024-05-01	Colima	Recolección de residuos	150
4	2024-10-03	Hidalgo	Recolección de residuos	136
5	2024-11-12	Sonora	Poda y arbolado	113

FECHA TEXTO TEXTO NÚMERO

11 La base de datos no debe tener columnas ni filas duplicadas.





- 12** Las bases de datos de actualización deben respetar la estructura de los datos que actualizan: nombres de columnas, categorías, despliegue de datos.

Criterios para columnas: nombres (13-20)

- 13** Todas las columnas deben estar nombradas únicamente en la primera fila de la base de datos.

No se deben combinar celdas para agrupar nombres de columnas (ver criterio 6).

14 Los nombres de columnas

DEBEN:

1. estar en español,
2. ser lo más explícitos y descriptivos posible,
3. estar escritos en minúsculas (a excepción de siglas: "UNAM", "SRE", "IMSS"),
4. separar las palabras con guión bajo "_", en vez de usar espacios o palabras juntas,
5. tener máximo 5 palabras (para nombres más largos, definir claves (ver criterio 16).

15 Los nombres de columnas NO

DEBEN:

1. tener caracteres especiales: ñ, tildes, diéresis, circunflejos, tabuladores, puntos, /, etc.,
2. tener saltos de línea ni espacios,
3. usar artículos o preposiciones.
4. empezar con un número*,
5. contener en su nombre un "_ 1" (de ser necesario se puede usar "_ 01" en su lugar)*.

**Por criterios de accesibilidad a los datos.*

Un error común es confundir las características de los nombres de las columnas con los datos (o valores) los cuales, sí deben estar escritos en correcto español (ver criterio 37).

✗

	A	B	C	D	E
1	ID	YEAR	NOMBRE DEL PROGRAMA	MONTO EN \$ DE BECA	PERÍODO DE BECA
2	18-000-456	2014	LICENCIATURA EN MATEMÁTICAS	18000	01-03-2014 A 01-03-2018
3	18-000-457	2014	LICENCIATURA EN ESTADÍSTICA	24000	01-03-2014 A 01-03-2018
4	18-000-458	2014	LICENCIATURA EN PROBABILIDAD	19200	01-03-2014 A 01-03-2018

✓

	A	B	C	D	E	F	G
1	programa_id	año	nombre_programa	monto_beca	denominacion_moneda	inicio_beca	conclusion_beca
2	18-000-456	2014	Licenciatura en matemáticas	18000	MXN	2014-01-30	2018-01-30
3	18-000-457	2014	Licenciatura en estadística	24000	MXN	2015-01-30	2019-01-30
4	18-000-458	2014	Licenciatura en probabilidad	19200	MXN	2016-01-30	2020-01-30





16 En caso de tener un nombre de columna con más de 5 palabras, se debe estructurar una clave.

Los datos se refieren a:

Hombres mayores de 80 años sin enfermedades crónico-degenerativas

El nombre "por defecto" de la columna sería:

hombres_mayores_80_anios_sin_enfermedades_cronico_degenerativas

+ 5 palabras

Las claves incorrectas serían:

✗ HMA80ASCD H_MA_80A_S_CD

Un ejemplo de clave correcta sería:

✓ h_ma80_se_crode

Y en el diccionario de datos se vería así:

	A	B	C	D
1	nombre_columna	tipo_dato	etiqueta_columna	descripcion
2	h_ma80_se_crode	numérico	Hombres 80+ sin enfermedades	Hombres mayores de 80 años sin enfermedades crónico-degenerativas

*Las claves deben respetar los criterios 14 y 15 en su estructura.

**Considerar que otros elementos (diccionario, metadatos y nombres de columnas) dan contexto a las claves.

17 Ninguna columna puede llamarse únicamente "id" o "ID".

En caso de necesitar utilizar el sufijo "id" se debe describir a qué se refiere.

✓

	A	B
1	produccion_id	nombre_campo
2	A-0001	Abkatun
3	A-0003	Agave
4	A-0008	Akal

18 Los nombres de las columnas no se pueden repetir en la misma base de datos.

19 Todos los nombres de columnas, incluidas las claves, deben describirse en el diccionario de datos.

20 Los diccionarios de datos deben especificar el tipo de dato de cada columna de la base: número, texto, fecha, coordenada





Criterios para filas (21-24)

21 La primera fila de la base de datos siempre debe tener el nombre de las columnas. A partir de la segunda fila sólo debe haber datos, nunca un nombre de columna.

22 Cada fila debe ser una unidad o registro.

23 Las bases de datos no deben tener filas vacías con el objetivo de separar los datos. En caso de que sea necesario, es preferible hacer varias bases de datos.

24 Todas las filas deben tener el mismo número de columnas.

Un error común que provoca que este criterio no se cumpla es tener datos de texto que incluyen comas (,) y éstos no se contienen entre comillas dobles (" "). También se debe indicar la existencia de valores nulos dependiendo del tipo de dato (ver cómo indicar datos nulos dependiendo del tipo de dato en los criterios 36, 46, 53, 56 de este listado).

✓

	A	B	C	D	E
21	1	año	municipio	entidad	tipo_solicitud
	2	2024	León	Guanajuato	Poda y arbolado
	3	2024	Manzanillo	Colima	Recolección de residuos
	4	2024	Tizayuca	Hidalgo	Recolección de residuos
	5	2024	Veracruz	Veracruz	Poda y arbolado
	6	2025	Mexicali	Baja California	Recolección de residuos
22	7	2025	Cuauhtémoc	Ciudad de México	Recolección de residuos
	8	2025	Cajeme	Sonora	Recolección de residuos
	9	2025	Compostela	Nayarit	Poda y arbolado
	10	2026	Matías Romer	Oaxaca	Recolección de residuos
	11	2026	Guanajuato	Guanajuato	Recolección de residuos
	12	2026	Mexicali	Baja California	Recolección de residuos

✗

	A	B	C	D	E
	1	año	municipio	entidad	tipo_solicitud
	2	2024	León	Guanajuato	Poda y arbolado
	3	2024	Manzanillo	Colima	Recolección de residuos
	4	2024	Tizayuca	Hidalgo	Recolección de residuos
	5	2024	Veracruz	Veracruz	Poda y arbolado
23	6				113
	7	2025	Mexicali	Baja California	Recolección de residuos
	8	2025	Cuauhtémoc	Ciudad de México	Recolección de residuos
	9	2025	Cajeme	Sonora	Recolección de residuos
24	10	2025	Compostela	Nayarit	Poda y arbolado
	11				340
	12	2026	Matías Romer	Oaxaca	Recolección de residuos
	13	2026	Guanajuato	Guanajuato	Recolección de residuos
	14	2026	Mexicali	Baja California	Recolección de residuos





Criterios para datos: estructura (25-28)

- 25** Todos los datos deben estar en una columna y una fila definida.
- 26** Ningún tipo de dato (número, texto, fecha, coordenada) debe tener espacios al principio o al final ni el carácter apóstrofe (').
- 27** Pueden existir celdas sin datos sin embargo, depende del tipo de dato (número, texto, fecha, coordenada) es la forma en la que se debe indicar el valor nulo.
(ver cómo indicar datos nulos dependiendo del tipo de dato en los criterios 36, 46, 53, 56 de este listado).
- 28** Ninguna base de datos debe tener datos que vulneren información clasificada, reservada o confidencial¹, a menos que por ley la institución tenga la facultad de hacer pública la información.

Criterios para datos: números (29-36)

- 29** Los datos numéricos no deben mezclar unidades (por ejemplo, metros y kilómetros) en la misma columna.
- 30** Para agregaciones de datos (promedios, sumas, totales, etc.):
 1. Evitar tener celdas aisladas con promedios, sumas parciales, totales generales, etc.
 2. Se puede tener promedios, sumas parciales, totales generales, etc. de datos de filas, pero no de columnas.

Incorrecto (X)

	A	B	C	D	E
1	año	nombre_programa	egresadas_mujeres	egresados_hombres	
2	2024	Licenciatura en matemáticas	55	40	
3	2024	Licenciatura en estadística	60	55	
4	2024	Licenciatura en computación	45	39	
5	2024	Licenciatura en probabilidad	53	53	
6	2025	Licenciatura en matemáticas	59	55	
7	2025	Licenciatura en estadística	45	39	
8	2025	Licenciatura en computación	53	53	
9	2025	Licenciatura en probabilidad	60	55	
10			Total mujeres: 430	Total hombres: 389	Total general: 819

Correcto (✓)

	A	B	C	D	E
1	año	nombre_programa	egresadas_mujeres	egresados_hombres	total
2	2024	Licenciatura en matemáticas	55	40	95
3	2024	Licenciatura en estadística	60	55	115
4	2024	Licenciatura en computación	45	39	84
5	2024	Licenciatura en probabilidad	53	53	106
6	2025	Licenciatura en matemáticas	59	55	114
7	2025	Licenciatura en estadística	45	39	84
8	2025	Licenciatura en computación	53	53	106
9	2025	Licenciatura en probabilidad	60	55	115





- 31** Los datos numéricos nunca deben tener coma (,) ni siquiera para separar miles. Los decimales se deben separar con un punto (.)

- 32** Los datos numéricos no deben tener ningún símbolo de unidad (kg, m, cm).

De ser necesario especificar el tipo de unidad, se debe incluir en una columna extra como dato de texto.

- 33** Los datos numéricos referentes a moneda no deben incluir símbolos ni códigos monetarios (¢, \$, USD, MXN).

De ser necesario especificar el tipo de unidad, se debe incluir en una columna extra como dato de texto.

X

	A	B	C
1	año	nombre_programa	monto_beca
2	2025	Licenciatura en matemáticas	\$18,000.00 31
3	2025	Licenciatura en estadística	\$20,000.00
4	2025	Licenciatura en computación	\$19,200.00 32
5	2025	Licenciatura en probabilidad	\$19,000.00 33

✓

	A	B	C	D
1	año	nombre_programa	monto_beca	denominacion_moneda
2	2025	Licenciatura en matemáticas	18000	MXN
3	2025	Licenciatura en estadística	20000	MXN
4	2025	Licenciatura en computación	19200	MXN
5	2025	Licenciatura en probabilidad	19000	MXN

- 34** En datos numéricos normalmente se utilizan dos decimales. Si una columna solo contiene valores enteros no se debe incluir separador decimal.

Es importante no confundir este criterio con datos de coordenadas las cuales, mientras más decimales tengan mayor será la precisión (ver criterio 55).

- 35** Los datos numéricos no deben tener ningún símbolo especial, por ejemplo, porcentaje (%).

De ser necesario especificar, puede hacerse en el diccionario de datos.





36 Representación de datos numéricos nulos: evitar usar texto u otro símbolo para representar valores numéricos nulos, se debe dejar la celda en blanco.

Evitar colocar guiones (-), "NA", "sin dato" u otro elemento que indique que no se tiene un registro, ya que esto ocasionará que no se pueda asignar el tipo de dato numérico.

Incorrecto (X)

	A	B	C	D	E	F
1	caseta	tipo	motos	autos	autobus	camion
2	Compostela	Camino	19	38	-	NA
3	Zacatecas	Camino	NA	38	54	54
4	Culiacán	Puente	11	-	40	40
5	Sinaloa	Puente	11	-	40	40
6	Tampico	Puente	10	20	38	NA

Correcto (✓)

	A	B	C	D	E	F
1	caseta	tipo	motos	autos	autobus	camion
2	Compostela	Camino	19	38		
3	Zacatecas	Camino		38	54	54
4	Culiacán	Puente	11		40	40
5	Sinaloa	Puente	11		40	40
6	Tampico	Puente	10	20	38	

Criterios para datos: texto (37-46)

37 Los datos de texto deben estar escritos en correcto español. Es decir, si pueden contener caracteres especiales (acentos, mayúsculas, minúsculas, etc.)

38 Los datos de texto no deben estar escritos todos en mayúsculas, ni separados por guiones bajos.

(ver criterios 14, 15).

Correcto (✓)

	A	B	C	D
1	año	entidad	tipo_solicitud	numero_solicitudes
2	2024	Guanajuato	Poda y arbolado	230
3	2024	Colima	Recolección de residuos	150
4	2024	Hidalgo	Recolección de residuos	120
5	2024	Veracruz	Poda y arbolado	113
6	2025	Baja California	Recolección de residuos	200
7	2025	Ciudad de México	Recolección de residuos	76
8	2025	Sonora	Recolección de residuos	89
9	2025	Nayarit	Poda y arbolado	340

Incorrecto (X)

	A	B	C	D
1	ANIO	ENTIDAD	TIPO_SOLICITUD	NUMERO_SOLICITUDES
2	2024	GUANAJUATO	PODA_Y_ARBOLADO	230
3	2024	COLIMA	RECOLECCION_DE_RESIDUOS	150
4	2024	HIDALGO	RECOLECCION_DE_RESIDUOS	120
5	2024	VERACRUZ	PODA_Y_ARBOLADO	113
6	2025	BAJA CALIFORNIA	RECOLECCION_DE_RESIDUOS	200
7	2025	CIUDAD_DE_MEXIC	RECOLECCION_DE_RESIDUOS	76
8	2025	SONORA	RECOLECCION_DE_RESIDUOS	89
9	2025	NAYARIT	PODA_Y_ARBOLADO	340





39 Si la base de datos contiene información estatal:

- se deben usar los nombres cortos para los siguientes estados: Coahuila, Michoacán, Veracruz,
- de considerar necesario especificar el nombre oficial, se puede agregar en una columna extra,
- no se deben usar las abreviaturas de los estados.

Aunque en estos estados el nombre oficial es el siguiente:

Coahuila de Zaragoza

Michoacán de Ocampo

Veracruz de Ignacio de la Llave

En las bases de datos se debe optar por su "nombre corto":

Coahuila

Michoacán

Veracruz

X	A	B	✓	A	B
1	entidad	numero_solicitudes	1	entidad	numero_solicitudes
2	AGS	230	2	Aguascalientes	230
3	CDMX	150	3	Ciudad de México	150
4	COAH	120	4	Coahuila	120
5	EDOMEX	113	5	Estado de México	113
6	MICH	200	6	Michoacán	200
7	SLP	145	7	San Luis Potosí	145
8	VER	150	8	Veracruz	150

40 Si la base de datos contiene columnas de estado, municipio o localidad, se debe agregar su clave geoestadística en una columna.

Es importante agregar las claves geoestadísticas para la interoperabilidad de las bases de datos que contienen esta información. Las claves de estados, municipios y localidades se pueden obtener del [Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades](#).

✓	A	B	C	D	E	F
1	entidad	cve_ent	municipio	cve_mun	localidad	cve_loc
2	Aguascalientes	01	Asientos	01002	Las Adjuntas	010020002
3	Chiapas	07	Altamirano	07004	Delicias Pachán	070040014
4	Guanajuato	11	Irapuato	11017	El Carmen	110170067
5	Guerrero	12	Cocula	12017	Cocula	120170001
6	Puebla	21	Acajete	21001	Buenavista	210010018
7	Yucatán	31	Tekax	31079	Canek	310790004

41 El uso de "Distrito Federal" en lugar de "Ciudad de México" debe ser únicamente para datos generados durante ese periodo histórico.





- 42** En los datos de texto que describan categorías, debe evitarse el uso de múltiples formas de nombrarla misma categoría, por ejemplo:

	A	B
✗ 1	entidad	sexo
2	AGS	F
3	Aguascalientes	FEMENINO
4	aguascalientes	femenino
5	AGUASCALIENTES	mujer

	A	B
✓ 1	entidad	sexo
2	Aguascalientes	femenino
3	Aguascalientes	femenino
4	Aguascalientes	femenino
5	Aguascalientes	femenino

- 43** No se deben dejar celdas vacías para indicar un “agrupamiento conceptual”.

✗	A	B	C	✓	A	B	C
1	año	nombre_programa	monto_beca	1	año	nombre_programa	monto_beca
2	2025	Licenciatura en matemáticas	18000	2	2025	Licenciatura en matemáticas	18000
3	2026		20000	3	2026	Licenciatura en matemáticas	20000
4	2025	Licenciatura en estadística	19500	4	2025	Licenciatura en estadística	19500
5	2026		21500	5	2026	Licenciatura en estadística	21500
6	2025	Licenciatura en computación	22000	6	2025	Licenciatura en computación	22000
7	2026		24000	7	2026	Licenciatura en computación	24000

- 44** Los datos pueden incluir hipervínculos, siempre y cuando se especifique que son datos de texto.

- 45** Si los datos de texto incluyen comas internas (,) el valor completo debe estar contenido en comillas dobles (“”). Se debe evitar el uso de comillas dobles si no es para esta función.

- 46** Representación de datos de texto nulos: la base de datos debe ser explícita en indicar que el dato no existe. Idealmente se debe marcar “sin dato”.

✗

	A	B	C	D
1	unidad_programa	tematica_programa	grado	beca
2	Guanajuato	Ciencias con Orientación en Ciencias de la Computación	Doctorado	
3		Ciencias con Orientación en Matemáticas Aplicadas		
4	Aguascalientes	Métodos Estadísticos en Línea	Especialidad	
5	Guanajuato	Probabilidad y Estadística		
6		Ciencias con Especialidad en Computación y Matemáticas Industriales	Maestría	Sí
7	Zacatecas	Ingeniería de Software	Maestría	Sí
8		Análisis Estadístico y Computación		
9		Cómputo Estadístico		
10	Aguascalientes	Modelación y Optimización de Procesos	Maestría	Sí

✓

	A	B	C	D
1	unidad_programa	tematica_programa	grado	beca
2	Guanajuato	Ciencias con Orientación en Ciencias de la Computación	Doctorado	sin dato
3	sin dato	Ciencias con Orientación en Matemáticas Aplicadas	sin dato	sin dato
4	Aguascalientes	Métodos Estadísticos en Línea	Especialidad	sin dato
5	Guanajuato	Probabilidad y Estadística	sin dato	sin dato
6	sin dato	Ciencias con Especialidad en Computación y Matemáticas Industriales	Maestría	Sí
7	Zacatecas	Ingeniería de Software	Maestría	Sí
8	sin dato	Análisis Estadístico y Computación	sin dato	sin dato
9	sin dato	Cómputo Estadístico	sin dato	sin dato
10	Aguascalientes	Modelación y Optimización de Procesos	Maestría	Sí





Criterios para datos: fechas (47-53)

- 47** El formato para todas las fechas en las bases de datos debe seguir el estándar [ISO 8601](#):

(YYYY-MM-DDTHH:MM:SS[.mmmmmm] [+HH:MM])

Año:

YYYY (2026)

Año y mes:

YYYY-MM (2026-06)

Año, mes y día*:

YYYY-MM-DD (2026-06-28)

T: Separador entre fecha y hora

Año, mes, día, horas y minutos:

YYYY-MM-DDTHH:MM:SS (2026-06-28T08:00)

Año, mes, día, horas, minutos y segundos:

YYYY-MM-DDTHH:MM:SS (2026-06-28T08:00:30)

[.mmmmmm] [+HH:MM]: información opcional que indica:

[.mmmmmm]: milisegundos en decimal (hasta 6 cifras)

[+HH:MM]: usos horarios respecto a Greenwich

guión medio

2026-06-28

4 dígitos 2 2

*La especificación de mes y día siempre debe incluir dos dígitos.

**El separador de fecha siempre debe ser un guión medio "-".

***No se permite la anotación am o pm.



2026
año de
Margarita
Maza



48 Todas las bases de datos deben incluir una columna de fecha que permita identificar el periodo de la información.

49 Se debe buscar generar una columna que contenga la fecha completa, aunque la base solo incluya año y mes. Para este caso se puede tomar el último día del periodo al que se refiere.

50 Si la base de datos contiene fechas de inicio y fin de un periodo se deben indicar en columnas separadas.

X

	A	B	C
1	nombre_programa	monto_beca	periodo_beca
2	Licenciatura en matemáticas	18000	2014-01-30/2018-01-30
3	Licenciatura en estadística	24000	2015-01-30/2019-01-30
4	Licenciatura en probabilidad	19200	2016-01-30/2020-01-30

✓

	A	B	C	D
1	nombre_programa	monto_beca	inicio_beca	conclusion_beca
2	Licenciatura en matemáticas	18000	2014-01-30	2018-01-30
3	Licenciatura en estadística	24000	2015-01-30	2019-01-30
4	Licenciatura en probabilidad	19200	2016-01-30	2020-01-30

51 No se deben mezclar formatos de fechas.

X

	A	B	C	D
1	nombre_programa	monto_beca	inicio_beca	conclusion_beca
2	Licenciatura en matemáticas	18000	2014-01-30	30-01-2018 X
3	Licenciatura en estadística	24000	X 30-01-2015	2019-01-30
4	Licenciatura en probabilidad	19200	2016-01-30	30-01-2020 X

52 Con base en los criterios de nombres de columnas, aquellas que refieran al año deberán nombrarse “anio”.

53 Representación de datos de fechas nulos: se debe dejar la celda en blanco. Evitar usar texto u otro símbolo para representar valores de fecha nulos,

No colocar guiones (-), “SF”, “sin dato” u otro elemento que indique que no se tiene un registro de fecha.

X

	A	B	C	D
1	nombre_programa	monto_beca	inicio_beca	conclusion_beca
2	Licenciatura en matemáticas	18000	2014-01-30	sin fecha
3	Licenciatura en estadística	24000	-	2019-01-30
4	Licenciatura en probabilidad	19200	2016-01-30	sin fecha

✓

	A	B	C	D
1	nombre_programa	monto_beca	inicio_beca	conclusion_beca
2	Licenciatura en matemáticas	18000	2014-01-30	
3	Licenciatura en estadística	24000		2019-01-30
4	Licenciatura en probabilidad	19200	2016-01-30	





Criterios para datos: coordenadas (54-56)

- 54** Las coordenadas deben estructurarse en columnas separadas por latitud y longitud.

X

	A	B	C
1	nombre_museo	tematica	longitud_latitud
2	Museo de Arte Contemporáneo	Arte	-102.29489, 21.88283
3	Museo Ferrocarrilero	Historia	-102.28093, 21.88477
4	Museo de la Fauna	Ciencia	-102.29124, 21.85177

✓

	A	B	C	D
1	nombre_museo	tematica	longitud	latitud
2	Museo de Arte Contemporáneo	Arte	-102.29489	21.88283
3	Museo Ferrocarrilero	Historia	-102.28093	21.88477
4	Museo de la Fauna	Ciencia	-102.29124	21.85177

- 55** En coordenadas no se deben redondear los valores. Es importante utilizar números decimales ya que éstos influyen en la precisión de la ubicación.

X

	A	B
1	latitud	longitud
2	34° 11' 30"	58° 43' 20"

✓

	A	B
1	latitud	longitud
2	-34.6043222	-58.4134862

- 56** Representación de datos de coordenadas nulos: se debe dejar la celda en blanco. Evitar usar texto u otro símbolo para representar valores de coordenadas nulos.

Evitar colocar guiones (-), "NA", "sin dato" u otro elemento que indique que no se tiene un registro de coordenadas.

X

	A	B	C	D
1	nombre_museo	tematica	longitud	latitud
2	Museo de Arte Contemporáneo	Arte	-	-
3	Museo Ferrocarrilero	Historia	-102.28093	21.88477
4	Museo de la Fauna	Ciencia	sin dato	sin dato

✓

	A	B	C	D
1	nombre_museo	tematica	longitud	latitud
2	Museo de Arte Contemporáneo	Arte		
3	Museo Ferrocarrilero	Historia	-102.28093	21.88477
4	Museo de la Fauna	Ciencia		





JSON

1. Criterios para archivo (criterios 1 a 4)

-
- 1 Sólo se debe hacer un archivo JSON si los datos son multinivel o tienen jerarquías semánticas.**

Si los datos se pueden estructurar en una tabla, se debe optar por un formato CSV.

- 2 La codificación del archivo debe ser UTF8.**
-

- 3 El archivo es de texto y debe tener la extensión .json**
-

- 4 Los archivos JSON deben seguir la especificación [RFC 7159](#).**
-

2. Criterios para nombres (criterios 5 a 7)

-
- 5 Los nombres sólo pueden ser cadenas de texto y deben estar contenidos en comillas dobles (" ").**
-

- 6 Los nombres deben estar separados por guión bajo (_) y en minúsculas. Ningún nombre debe contener espacios.**
-

- 7 Cada nombre y su valor deben estar separados por dos puntos (:) ej. nombre: valor.**
-

3. Criterios para elementos (criterios 8 a 13)

-
- 8 El archivo JSON debe estar conformado por colecciones no ordenadas de parejas de elementos.**

Cada pareja tiene dos partes: un nombre y el valor correspondiente a ese nombre. Los valores pueden ser:

- cadenas de texto,
 - valores numéricos,
-





- elementos nulos que siempre se deben identificar con la palabra "null",
- listas de elementos (secuencia ordenada de valores),
- objetos (subcolecciones de parejas de elementos con sus propias llaves { } y valor).

9 Los objetos se deben representar por una pareja de llaves { } que contienen al conjunto de parejas (nombre: valor) separados por el carácter (.). Los nombres usados dentro de un objeto deben ser únicos.

10 Las listas se representan por una pareja de corchetes [] que contienen al conjunto de valores y en la cual los elementos están separados por el carácter (,)

La especificación del formato JSON no menciona que los valores dentro de una lista deban ser el mismo tipo de valor: texto, numérico, coordenadas o fechas, sin embargo, se sugiere que así sea.

11 Los números deben seguir las mismas especificaciones que las mencionadas en los datos numéricos del formato CSV.

Ver criterios para datos numéricos del formato CSV (29 a 36).

12 Las cadenas de texto deben estar siempre contenidas entre comillas dobles (" ") y seguir las mismas especificaciones que las mencionadas en los datos de texto del formato CSV.

Ver criterios para datos de texto del formato CSV (37 a 46).

13 Si el archivo contiene listas de objetos que en su interior tengan los mismos nombres, deben tener coherencia entre ellos, es decir, todos deben tener el mismo conjunto de nombres.





GPKG

1. Criterios para archivo (criterios 1 y 2)

1 El archivo debe tener extensión .gpkg

2 La tabla de datos debe satisfacer los mismos criterios especificados para formato CSV, en medida de lo posible.

Este criterio debe cumplirse a menos que sea imposible por limitaciones del formato GPKG.

Nota. Se recomienda revisar a profundidad el estándar internacional en el siguiente enlace: [OGC GeoPackage Encoding Standard](#).

GeoTIFF

El archivo GeoTIFF está compuesto por dos archivos: (1) **.tiff** es el archivo principal de la imagen ráster que contiene los datos visuales; (2) **.tfw** es el archivo de los datos georreferenciados es decir, define las coordenadas (latitud y longitud).

Por la complejidad de sus criterios, se recomienda revisar el estándar internacional del *Open Geospatial Consortium (OGC)* en el siguiente enlace: [OGC GeoTIFF standard](#).