



DATOS ABIERTOS

¿POR DÓNDE
EMPEZAR?



#DatosAbiertos

Elaborado en mayo de 2024,
publicado en abril de 2025.

 Info CDMX | Estado Abierto

 @InfoCdMex | @Estado_Abierto



Pleno del INFO CDMX

LAURA LIZETTE ENRÍQUEZ RODRÍGUEZ Comisionada Presidenta.

MARÍA DEL CARMEN NAVA POLINA Comisionada Ciudadana.

JULIO CÉSAR BONILLA GUTIÉRREZ Comisionado Ciudadano.



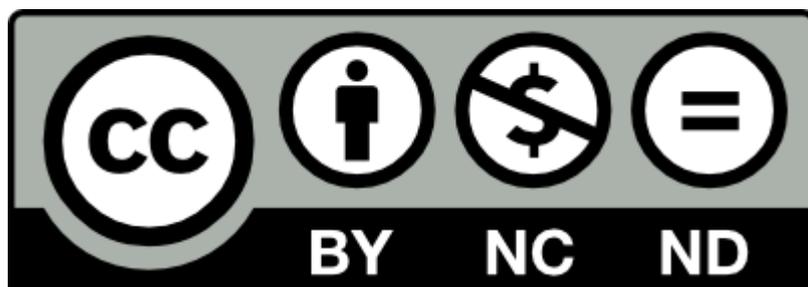
Autor: Equipo de Estado Abierto.

Entidad: Instituto de Transparencia, Acceso a la Información, Protección de Datos Personales y Rendición de Cuentas de la Ciudad de México (INFO CDMX).

La Morena, 865, Col. Narvarte Poniente, Alcaldía Benito Juárez, C.P. 03020, Ciudad de México.

Lugar y año de publicación: Ciudad de México, abril 2025. **Idioma:** Español.

Tipo de licencia: Creative Commons Atribución No Comercial 4.0 Internacional.



Cómo citar este documento:

Equipo de Estado Abierto, "Datos abiertos: ¿por dónde empezar?". Ciudad de México: INFO CDMX, 2025.

Este material de socialización deriva del Plan Estratégico para el desarrollo de la agenda de Estado Abierto desde el INFO CDMX de la gestión 2019-2025 de la Comisionada Ciudadana María del Carmen Nava Polina, su ponencia y el área que coordina, la Dirección de Estado Abierto, Estudios y Evaluación. Agradecimientos especiales a Juan Manuel Casanueva de Social Tic y a Andrea Barenque de Datamorfosis.

Ciudad de México, elaborado en mayo de 2024 y publicado en abril de 2025.



Esta edición de **Datos abiertos: ¿por dónde empezar?** es un producto cocreado en el marco de la Red Ciudad en Datos Abiertos, en la cual participaron:

Info CDMX

Coordinación: Comisionada Ciudadana
María del Carmen Nava Polina

Redacción:

María Soledad Rodrigo
Evelyn Salazar Flores
Daniel Saavedra Lladó

SocialTIC, A.C.

Juan Manuel Casanueva, Director

Datamorfofis

Andrea Barenque, Fundadora y
Directora

Diseño editorial:

Sarai Zulema Oviedo Hernández
Melisa Citlali Romero Castillo
Lourdes García Miranda

Cita sugerida: Equipo de Estado Abierto, “Datos abiertos: ¿por dónde empezar?”. Ciudad de México: INFO CDMX, 2025.



Índice

Prólogo	6
Introducción	9
Objetivo	9
1. ¿Qué son los datos abiertos?	10
¿Qué no es un dato abierto?	14
¿Cómo tiene que ser un dato abierto?	15
Ciclo de vida de un dato abierto	15
2. ¿Para qué sirven los datos abiertos?	16
3. ¿Cómo publicar datos abiertos?	22
4. ¿Dónde puedo aprender más?	27
Conclusión	28
Referencias	29
Otros ejemplos	30



Prólogo

Desde el equipo de Estado Abierto y con base en el Modelo Apertura por Diseño, vemos la apertura institucional como un diálogo permanente entre la sociedad con las instituciones y actores públicos para detectar necesidades y problemáticas públicas así, como para proponer y cocrear soluciones a dichos problemas.

La apertura institucional debe siempre contar con los cinco elementos: transparencia y rendición de cuentas; participación y cocreación; integridad y prevención de conflicto de intereses; lenguaje sencillo, formatos abiertos y accesibles; y memoria y archivo. Los datos abiertos los encuadramos en el cuarto elemento por considerarlos parte de los formatos abiertos. Es así que impulsamos la Red Ciudad en Datos Abiertos, para que en colaborativo con organizaciones de la sociedad civil, instituciones y actores públicos y personas expertas, avancemos en la apertura de datos en la Ciudad de México.

Los datos abiertos, liberados y publicados alcanzan un mayor impacto, cuando son aprovechados y reutilizados. Por estas razones, impulsamos este material que desarrollamos en colaborativo con SocialTIC, el documento *Datos abiertos: ¿por dónde empezar?*, para contribuir a que las personas que laboran en, las instituciones y actores públicos, así como la población interesada en datos abiertos, tengan a su alcance un documento escrito en lenguaje claro, sencillo y con ejemplos, que les permitan no sólo conocer qué son los datos abiertos y cuáles son sus características, sino cómo pueden iniciar con su implementación y cuáles recursos que pueden utilizar.

La apertura de datos es un primer paso para publicar información que luego las personas puedan utilizar y reutilizar para generar conocimiento y mejorar la toma de decisiones por parte de la sociedad como para las instituciones y actores públicos.

María del Carmen Nava Polina

Comisionada Ciudadana del Instituto de Transparencia, Acceso a la Información Pública, Protección de Datos Personales y Rendición de Cuentas de la Ciudad de México



Hace poco más de una década, la agenda de datos abiertos surgió en América Latina como una extensión natural de los esfuerzos por promover la transparencia. Sin embargo, desde su origen tuvo una ambición distinta: no se trataba solo de abrir información para que la ciudadanía pudiera fiscalizar a sus gobiernos, sino de crear las condiciones para que cualquier persona —desde una periodista hasta una desarrolladora o una organización de la sociedad civil— pudiera transformar los datos en algo nuevo. Nuevas empresas, nuevos mapas, nuevos reportajes, nuevas visualizaciones. Nuevas formas de entender lo público. Hoy, esa agenda enfrenta un nuevo punto de inflexión.

Así como la tecnología y la sociedad han cambiado, también lo ha hecho la forma en la que producimos, compartimos y usamos los datos. Es momento de actualizar nuestras preguntas, nuestras prácticas y nuestras herramientas. No basta con abrir archivos: hay que abrir conversaciones, conectar capacidades y pensar en cómo convertir la información en conocimiento útil, en decisiones mejor informadas y en políticas públicas más justas.

Hablar de datos es hablar de la historia misma de la humanidad. Desde los primeros censos hasta los registros digitales que hoy generan nuestros teléfonos, los datos son una forma de mirar el mundo y tratar de entenderlo. También son una palanca para transformarlo. Por eso, este documento busca ofrecer un punto de partida para gobiernos locales, equipos técnicos y personas interesadas en comenzar —o renovar— su camino hacia una política de datos abiertos. Porque abrir datos no es el fin, sino el medio para construir un presente más conectado y un futuro más inteligente.

A lo largo de estas páginas se abordan preguntas clave para cualquier institución que quiera iniciar una política de apertura: ¿qué son y qué no son los datos abiertos?, ¿cómo elegir qué conjuntos de datos vale la pena compartir?, ¿cuáles son los formatos adecuados?, ¿qué buenas prácticas existen para estructurar los datos dentro de una tabla? Entender estos aspectos técnicos es tan importante como tener claridad política. Elegir bien qué publicar, cómo nombrar las columnas, cómo organizar los registros y cómo facilitar su comprensión puede marcar la



diferencia entre una base de datos que transforma y una que simplemente ocupa espacio.

Este documento busca justamente ser una brújula en ese proceso. Nada de esto es posible sin liderazgo. La apertura de datos no ocurre por accidente: requiere visión, coordinación y una estrategia clara. El rol del Chief Data Officer —o su equivalente en los equipos de gobierno— es clave para acercar los datos correctos a las personas correctas en el momento correcto. Pero también es fundamental reconocer a todas las personas que intervienen en el ciclo de vida de los datos: quienes los capturan, los mantienen, los documentan, los visualizan y los reutilizan. Una política de datos abiertos debe ser, sobre todo, una política de colaboración.

Como parte de la comunidad de datos en México, desde Datamorfosis, trabajamos por un mundo en el que una persona nunca sea el impedimento para acceder a un dato y poder usarlo para mejorar la vida de otra. Este documento es una pequeña contribución a ese propósito.

Andrea Barenque Hernández
Fundadora y Directora General de Datamorfosis



Introducción

El documento Datos abiertos: ¿por dónde empezar?, es un trabajo realizado desde la Red Ciudad en Datos Abiertos, entre el Equipo de Estado Abierto del InfoCDMX y la organización SocialTIC y Datamorfosis, para poner al alcance de las personas que laboran en las instituciones y actores públicos, así como de la población interesada en datos abiertos, de una manera clara y sencilla, la información acerca de qué son los datos abiertos y cuáles estrategias se pueden implementar para la apertura de datos.

Este material de socialización busca impulsar la cultura de la apertura de datos en la Ciudad de México, así como abonar a la acción estratégica de la Política Nacional de Datos Abiertos de fortalecer los procesos de planeación, publicación y uso de los datos en México, con base en los conceptos, principios, características y criterios definidos en la propia Política Nacional.

Los datos abiertos son una manera de acercar proactivamente la información a las personas, por lo que es importante que quienes generen los datos, puedan identificar las necesidades de información de la sociedad para que la liberación de las bases de datos atiendan a dichas necesidades, bajo los estándares y características que hacen abiertos los datos.

Es así que los datos abiertos permiten aprovechar la información pública contenida en bases de datos, mediante el uso herramientas técnicas, de procesamiento y visualización de datos que faciliten su análisis, para la generación de propuestas, mejorar políticas públicas y facilitar soluciones a problemas sociales.

Objetivo

El objetivo de este documento sobre Datos Abiertos consiste en que las personas que laboran en las instituciones y actores públicos, así como la población interesada en datos abiertos, comprendan y conozcan de una manera sencilla, las acciones que pueden realizar para publicar información en formatos de datos abiertos, la puedan leer, visualizar y aprovechar, así como encontrar plataformas, aplicaciones digitales, visualizadores y otros tipos de recursos desarrollados con base en los datos abiertos que pudieran resultarles de interés.



1. ¿Qué son los datos abiertos?

Para empezar, veamos qué es un *dato*.

Los **datos** se pueden definir como características que describen un objeto, acontecimiento, lugar, situación o hecho. Algunos ejemplos de datos son:

- la **magnitud** de un sismo.
- la **velocidad** con la que se mueve un trolebús.
- la **fecha** en que hubo contingencia ambiental en la ciudad.
- la **ubicación** de un hospital.
- la **cantidad** de personas registradas en algún programa social.

A partir de estos ejemplos, vemos que un dato se puede representar con un número o con algunas cuantas palabras.

NOTA: Otro concepto importante es el de **información**. La información se puede describir como un conjunto de datos, organizados y estructurados, a partir de los cuales se puede generar conocimiento.

Día con día, se genera una gran cantidad de datos, lo cual puede volver un poco difícil trabajar con conjuntos o bases de datos tan grandes. Afortunadamente tenemos a las computadoras, que son una herramienta muy poderosa que nos puede ayudar no sólo a manejar todos esos datos, sino que además nos permite generar visualizaciones y análisis estadísticos.

Entre los datos que se generan, están los datos abiertos que son datos digitales de carácter público accesibles en línea, y que pueden ser usados, reutilizados y redistribuidos por cualquier persona interesada. Los datos de carácter público son aquellos que el sector público genera y/o administra para el cumplimiento de sus funciones.

El artículo 3, fracción VIII de la Ley General de Transparencia y Acceso a la Información Pública (LGTAIP), y el artículo 6, fracción XI de la Ley de Transparencia, Acceso a la Información Pública y Rendición de Cuentas de la Ciudad de México, establecen las características que deben tener los datos abiertos, las cuales se señalan a continuación:

De acuerdo con estos artículos, los **datos abiertos**

- **Accesibles** - Los datos están disponibles para la gama más amplia de usuarios, para cualquier propósito.
- **Integrales** - Contienen el tema que describen a detalle y con los metadatos necesarios.
- **Gratuitos** - Se obtienen sin entregar a cambio contraprestación alguna.



- **No discriminatorios** - Los datos están disponibles para cualquier persona, sin necesidad de registro.
- **Oportunos** - Son actualizados, periódicamente, conforme se generen.
- **Permanentes** - Se conservan en el tiempo, para lo cual, las versiones históricas relevantes para uso público se mantendrán disponibles con identificadores adecuados al efecto.
- **Primarios** - Proviene de la fuente de origen con el máximo nivel de desagregación posible.
- **Legibles por máquinas** - Deberán estar estructurados, total o parcialmente, para ser procesados e interpretados por equipos electrónicos de manera automática.
- **En formatos abiertos** - Los datos estarán disponibles con el conjunto de características técnicas y de presentación que corresponden a la estructura lógica usada para almacenar datos en un archivo digital, cuyas especificaciones técnicas están disponibles públicamente, que no suponen una dificultad de acceso y que su aplicación y reproducción no estén condicionadas a contraprestación alguna.
- **De libre uso** - Citan la fuente de origen como único requerimiento para ser utilizados libremente. (Los datos deben ser gratuitos bajo una licencia abierta, como por ejemplo las desarrolladas por Creative Commons.)

Si los datos abiertos cumplen con estas características, tendrán un mayor potencial para su uso y aprovechamiento.

NOTA: Los **metadatos** son datos que describen datos. En otras palabras, los metadatos son los detalles relevantes que describen un conjunto de datos, por ejemplo, cómo se generaron los datos, en qué formato se encuentran, fecha en que se crearon o fecha de la última actualización, versión del conjunto de datos, etc.

Por lo tanto, **si un dato puede ser utilizado, modificado y compartido libremente de manera digital por cualquier persona en cualquier momento para cualquier propósito, entonces es un dato abierto.**

Para que un dato pueda ser utilizado, modificado y compartido libremente, tiene que estar en un **formato abierto** que las computadoras puedan reconocer de manera automática y que fácilmente pueda ser procesado. Algunos formatos abiertos disponibles son:

- **CSV.** *Comma Separated Values.* Es un formato donde las **columnas se separan por comas** y las filas por saltos de línea.
- **JSON.** *JavaScript Object Notation.* Es un formato donde los datos se organizan en una **colección de pares nombre-valor.**
- **SHP.** *Shapefile.* Es un formato que se utiliza para almacenar la ubicación geométrica y la información de **atributos de las entidades geográficas.** Las



entidades geográficas de un shapefile se pueden representar **por medio de puntos, líneas o polígonos (áreas)**.

- **KML**. *Keyhole Markup Language*. Es un formato basado en XML que se utiliza para visualizar **información en un contexto geográfico**. Se utiliza para mostrar datos geográficos en un navegador terrestre, como Google Earth y Google Maps.
- **GEOJSON**. *Geographic JSON*. Es un formato diseñado para **representar elementos geográficos**.
- **GEOPACKAGE**. Es un formato de **datos geoespaciales** implementado como un contenedor de base de datos SQLite.

NOTA: El formato **PDF** es útil para la publicación de documentos, informes, libros, minutas y otro tipo de información escrita, pero **no para la publicación de datos abiertos**. Los datos en archivos PDF no son accesibles porque no se pueden procesar automáticamente de una manera fácil y confiable, lo que obliga a la persona usuaria a realizar el procesamiento adicional para incorporar los datos en un formato que le sea más fácil de utilizar.

Para trabajar con datos en formatos abiertos podemos utilizar programas o aplicaciones que permiten visualizar estos datos de una manera clara y fácil de interpretar. Por ejemplo, para trabajar con archivos .csv podemos utilizar programas como Excel o Google Sheets, que transforman los datos en una tabla donde por cada registro u observación hay una **fila** y por cada característica de estos registros hay una **columna**. A los datos que se pueden acomodar de esta manera se les conoce como **datos tabulares**.

(1) Archivo CSV

```
fecha,mes,año,tipo_pago,afluencia
01/01/2022,Enero,2022,Prepago," 14,251 "
01/01/2022,Enero,2022,Gratuidad," 1,806 "
02/01/2022,Enero,2022,Prepago," 21,875 "
02/01/2022,Enero,2022,Gratuidad," 2,259 "
03/01/2022,Enero,2022,Prepago," 49,290 "
03/01/2022,Enero,2022,Gratuidad," 5,774 "
04/01/2022,Enero,2022,Prepago," 52,538 "
04/01/2022,Enero,2022,Gratuidad," 5,467 "
```



(2) Archivo CSV en Google Sheets

	A	B	C	D	E
1	fecha	mes	anio	tipo_pago	afluencia
2	01/01/2022	Enero	2022	Prepago	14251
3	01/01/2022	Enero	2022	Gratuidad	1806
4	02/01/2022	Enero	2022	Prepago	21875
5	02/01/2022	Enero	2022	Gratuidad	2259
6	03/01/2022	Enero	2022	Prepago	49290
7	03/01/2022	Enero	2022	Gratuidad	5774
8	04/01/2022	Enero	2022	Prepago	52538
9	04/01/2022	Enero	2022	Gratuidad	5467
10	05/01/2022	Enero	2022	Prepago	52288
11	05/01/2022	Enero	2022	Gratuidad	5389
12	06/01/2022	Enero	2022	Prepago	45249



Plataforma: Portal de Datos Abiertos

Datos: Afluencia Diaria Tren Ligero (desglosada)

Archivo: *afluencia_desglosada_tl_02_2024.csv*

Institución: Secretaría de Movilidad

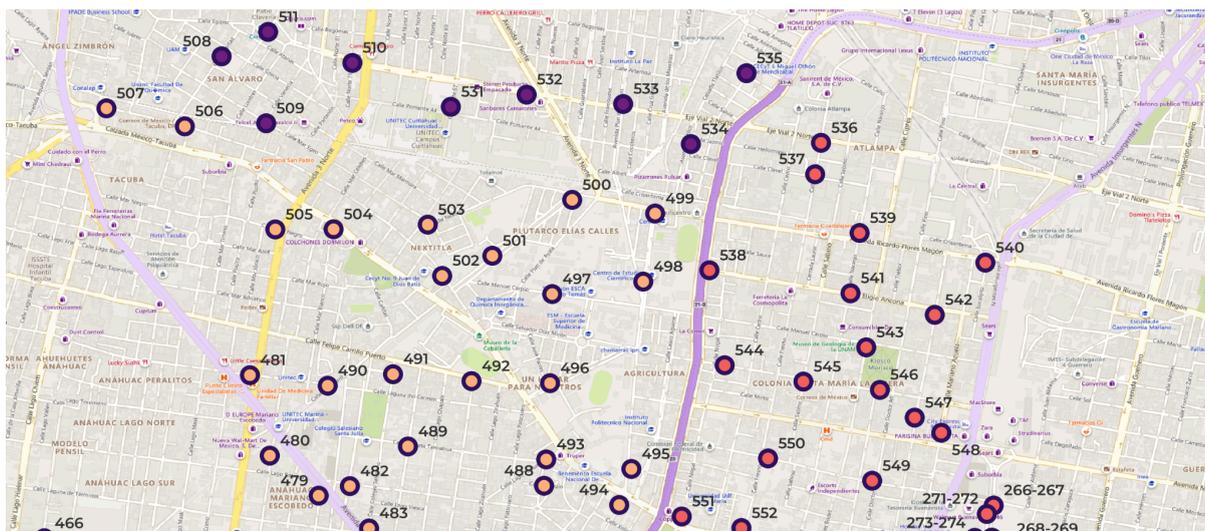
URL:

<https://datos.cdmx.gob.mx/dataset/afluencia-diaria-servicio-de-transportes-electricos>

Por otro lado, para trabajar con **datos georreferenciados** en formatos SHP, KML, GEOJSON y GEOPACKAGE existen programas de código abierto como *QGIS*, que permite visualizar los datos de dos maneras que se complementan mutuamente: como una **tabla** y como **objetos geométricos** ubicados en un mapa.



(1) Archivo SHP en QGIS



sistema	num_cicloes	calle_prin	calle_secu	colonia	alcaldia	latitud	longitud	sitio_de_e	estatus
Ecobici	503	Nexttita	Mar Egeo	Nexttita	Miguel Hidalgo	19.455531000000001	-99.175222000000005	Arroyo	Instalada
Ecobici	504	Mar Blanco	Calzada Mexico Tacuba	Popotla	Miguel Hidalgo	19.455342999999999	-99.179056000000003	Arroyo	Instalada
Ecobici	505	Mar Negro	Eje 4 Poniente	Popotla	Miguel Hidalgo	19.455334799999999	-99.181427799999994	Arroyo	Instalada
Ecobici	506	Mar Mediterraneo	Mar Arabigo	Tacuba	Miguel Hidalgo	19.459553000000000	-99.185111000000006	Arroyo	Instalada
Ecobici	507	Mar Mediterraneo	Avenida Azcapotzalco	Tacuba	Miguel Hidalgo	19.460291999999999	-99.188303000000005	Arroyo	Instalada
Ecobici	508	Libertad	José Sánchez Trujillo	San Alvaro	Azcapotzalco	19.462420999999999	-99.183614000000006	Banqueta	Instalada
Ecobici	509	Benit Juárez	Vanecia	San Alvaro	Azcapotzalco	19.459679999999999	-99.181799999999996	Arroyo	Instalada
Ecobici	510	Egipto	Calle Norte 79 A	Claveria	Azcapotzalco	19.462152000000000	-99.178291000000002	Arroyo	Instalada
Ecobici	511	Irapuato	Egipto	Claveria	Azcapotzalco	19.463422000000001	-99.181720999999996	Arroyo	Pendiente
Ecobici	512	Avenida Claveria	Floresta	Claveria	Azcapotzalco	19.465636000000000	-99.186341999999996	Arroyo	Instalada

Plataforma: Portal de Datos Abiertos

Datos: Cicloestaciones Ecobici

Archivo: *cicloestaciones_ecobici.shp*

Institución: Secretaría de Movilidad

URL: <https://datos.cdmx.gob.mx/dataset/cicloestaciones-ecobici-nuevo-sistema>

El color de los puntos está definido por la variable “alcaldia”.

El número sobre los puntos está definido por la variable “num_cicloes”.

¿Qué no es un dato abierto?

La información que publican los sujetos obligados a través de los formatos de obligaciones de transparencia, tanto en la Plataforma Nacional de Transparencia como en los portales institucionales de transparencia, **no son datos abiertos por no reunir las características definidas** en la ley, a pesar de encontrarse publicados de forma tabular y poder ser leídos en un programa como Excel.

La principal diferencia de los datos abiertos respecto al resto de la información pública a la cual es posible tener acceso, radica en sus **características jurídicas y técnicas para su disponibilidad y explotabilidad** por sistemas y herramientas así como por cualquier persona u organización con las capacidades necesarias para



hacerlo, y para transformar dichos datos en reportes, visualizaciones, aplicativos o cualquier otro producto que pueden dar una nueva relevancia y utilidad a la información que, de otra forma, posiblemente ya habría cumplido su ciclo útil al interior de las instituciones gubernamentales.¹

Otro formato frecuentemente utilizado por su facilidad de ser exportado y en el que se comparten datos tabulares es el conocido como .pdf; este formato no es un formato abierto ya que dificulta la lectura y disponibilidad de los datos que presenta, o se trata de información que fue generada en otro formato para ser impresa y después digitalizada, por lo que se trata de una imagen que no es legible por máquina.

¿Cómo tiene que ser un dato abierto?

Para que los datos abiertos sean de utilidad para las personas y los puedan aprovechar, estos deben considerar las siguientes características:

- **Estandarizados:** se refiere a si los datos siguen un estándar en particular, esto facilita el uso e intercambio de datos entre diferentes áreas o entre instituciones. Un ejemplo de estándar es el Estándar de Datos para las Contrataciones Abiertas (EDCA) (OCDS, por sus siglas en inglés *Open Contracting Data Standard*) que permite la divulgación de datos y documentos en todas las etapas de un proceso de contratación gracias a que define un modelo de datos común.
- **Limpios y procesados:** es decir, los datos pasaron por un procesamiento y ordenamiento de manera que son más fáciles de comprender, analizar y utilizar.
- **Documentados:** se refiere a acompañar los datos con algún documento de registro como manuales de usuario, guías, bitácoras, diccionarios, donde se explique cómo se obtuvieron los datos y a qué procesos se sometieron.
- **Íntegros:** es decir, que la información contenida en los datos abiertos no haya sido modificada.
- **Actualizados:** los datos abiertos publicados deberán corresponder a la última versión generada, y contar con la fecha en que fueron publicados.
- **Libres:** los datos abiertos deben contar con licencias y permisos de uso para que las personas puedan aprovecharlos.

Ciclo de vida de un dato abierto

El ciclo de vida de un dato abierto es una serie de pasos ordenados que sigue el proceso de apertura de datos, el cual inicia con la planeación del proceso para

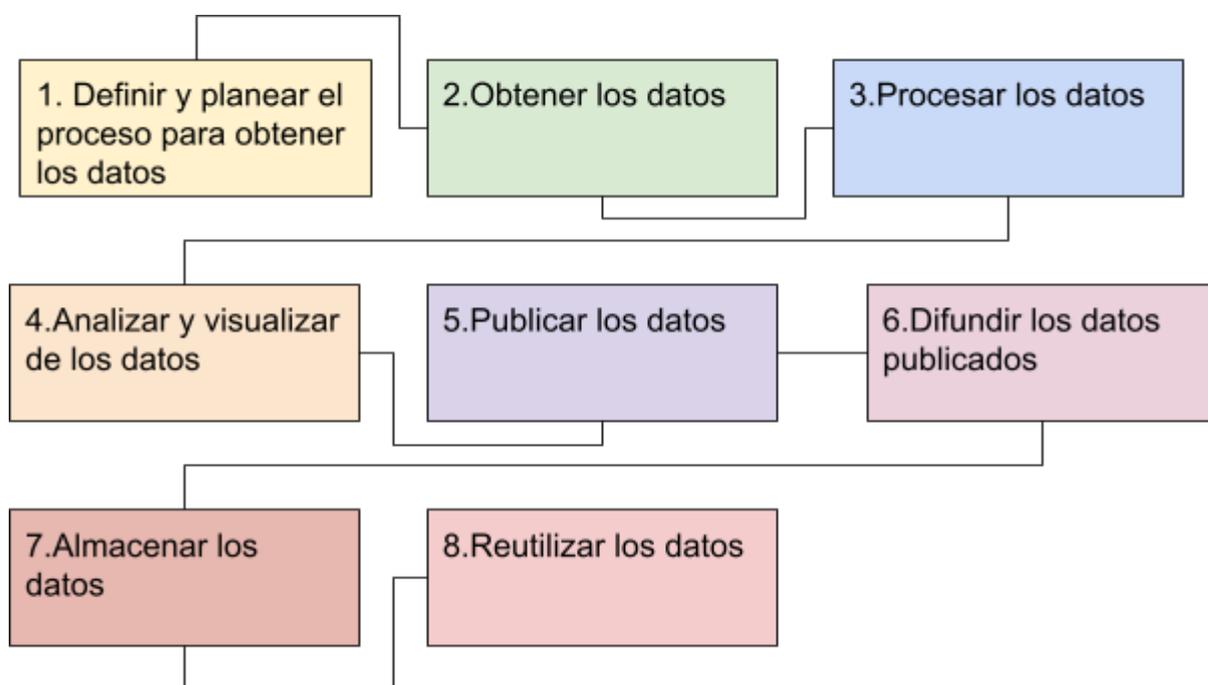
¹ México, A. (2023). Política Nacional de Datos Abiertos. *Abramos México*. Tomado de <https://abramosmexico.pubpub.org/pub/politica-nacional-de-datos-abiertos>



obtener o generar los datos que se van a abrir, hasta su reutilización una vez que son públicos.

Como se presenta en el esquema, este ciclo inicia con la definición y planeación del proceso para la obtención de los datos y continúan con la obtención y procesamiento de los mismos para seguir con el análisis y visualización de los datos previo a su publicación.

Una vez que los datos abiertos están publicados, estos se pueden difundir para que las personas los conozcan y se encuentren disponibles en algún repositorio de acceso público que permita su descarga y reutilización de los datos para fines distintos a los que fueron creados.



Fuente: Elaboración propia

2. ¿Para qué sirven los datos abiertos?

Los datos abiertos tienen el potencial de beneficiar a cualquier persona sin importar su género, edad, nivel educativo, ocupación o religión, así como a cualquier organización del sector privado, sociedad civil, grupo de académicos, emprendedores, periodistas, o incluso a la misma institución que genera los datos, y pueden tener un impacto en la solución de problemáticas en áreas como la salud, educación, desarrollo social, movilidad y transporte, comercio y seguridad pública, entre muchas otras.



De acuerdo con los resultados del Barómetro Global de Datos, el cual tiene una cobertura de 109 países, los grupos más propensos a ser identificados como usuarios de datos son los medios de comunicación (36.4% de los casos revisados) y las organizaciones de la sociedad civil (37.6%), seguidos por el gobierno (30.3%) y la academia (25.7%).

Algunas de las posibles aplicaciones de los datos abiertos son:

- Entender y resolver problemas en cualquier ámbito.
- Tomar decisiones de manera informada.
- Realizar estudios e investigaciones para entender algún fenómeno o suceso.
- Diseñar políticas públicas y medir el impacto de las políticas existentes.
- Hacer análisis e identificar áreas de mejora de políticas públicas, programas y servicios.
- Crear nuevos servicios y aplicaciones.

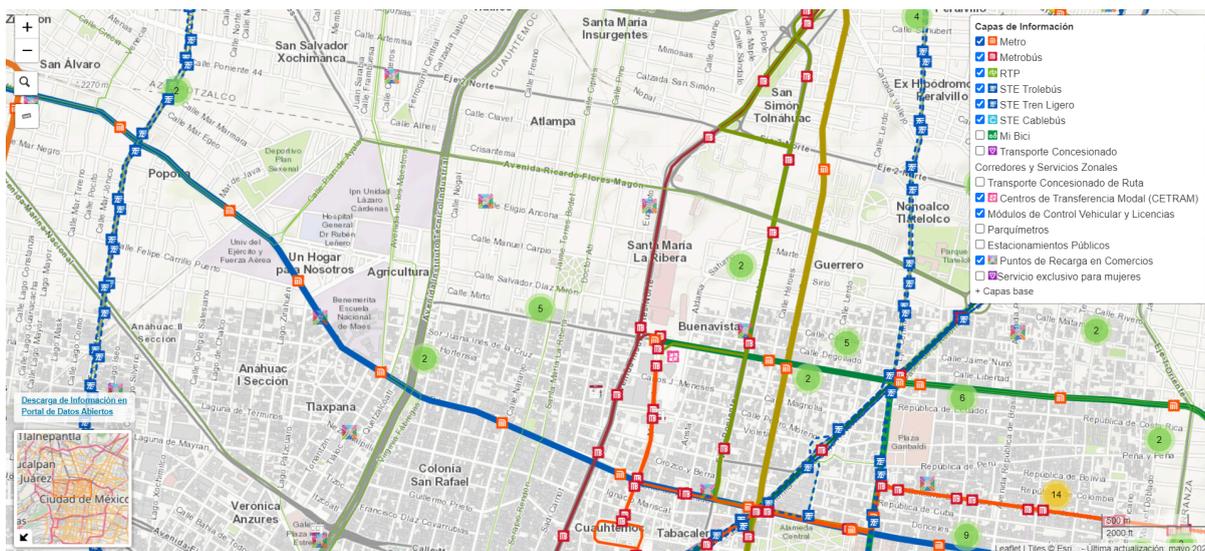
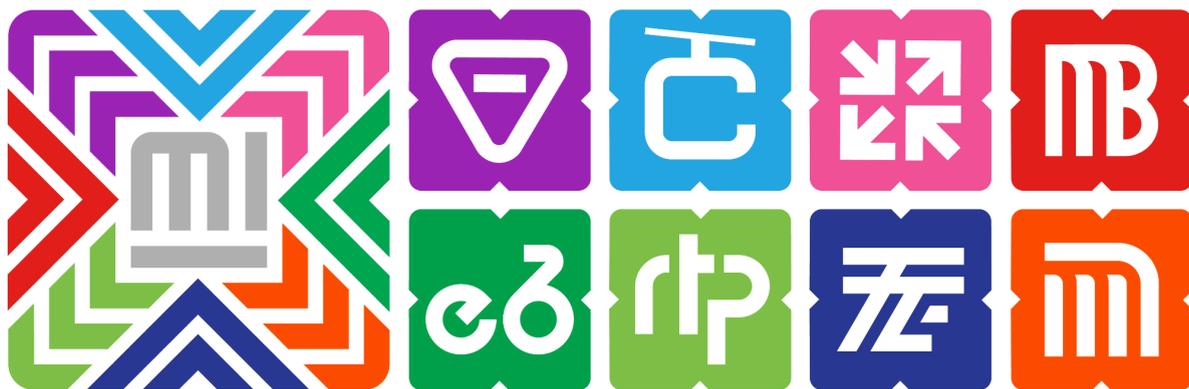
Un caso que demuestra la contribución de los datos abiertos al diseño de **políticas públicas** es el de la Política Nacional Anticorrupción en México, donde gracias a la sistematización de datos se detectaron y definieron asuntos prioritarios basados en evidencia².

Sobre la aplicación de datos abiertos para el análisis e identificación de mejoras de **programas sociales**, en México se han usado datos georreferenciados que permiten obtener más y mejor información sobre las comunidades receptoras, así como los criterios y reglas de operación de dichos programas. El uso de datos abiertos puede ayudar a conocer mejor a las personas beneficiarias reales del gasto social del país³.

Son numerosos los ejemplos en los que una institución puede sacar provecho de los mismos datos abiertos que publica, tal como es el caso del **Visor de Movilidad Integrada** de SEMOVI, en el cual se puede visualizar la ubicación de las estaciones y rutas de cada sistema de movilidad, además de la posición de los Centros de Transferencia Modal (CETRAM), de los módulos de control vehicular y otros puntos clave.

² México, A. (2023). Política Nacional de Datos Abiertos. *Abramos México*. Tomado de <https://abramosmexico.pubpub.org/pub/politica-nacional-de-datos-abiertos>

³ México, A. (2023). Política Nacional de Datos Abiertos. *Abramos México*. Tomado de <https://abramosmexico.pubpub.org/pub/politica-nacional-de-datos-abiertos>



Plataforma: Visor de Movilidad Integrada

URL: https://app.semovi.cdmx.gob.mx/visor_sitp/

Otro ejemplo en este sentido es el del **Mapa Ecobici**, el cual muestra de una manera interactiva la ubicación de las estaciones de Ecobici, así como el número de bicicletas y anclajes disponibles.



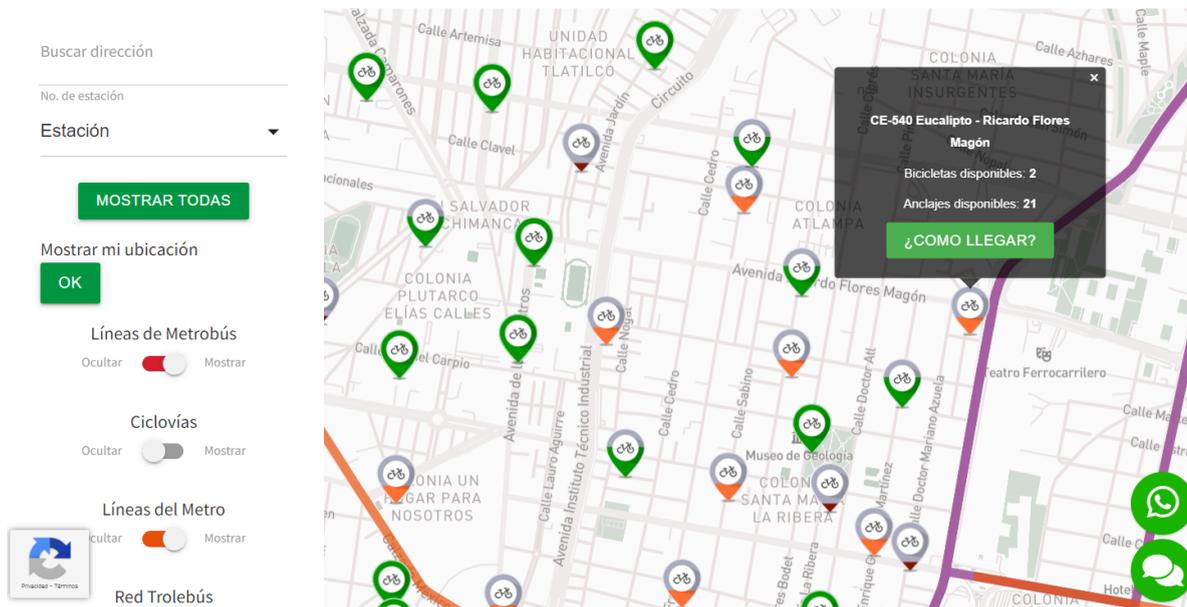
GOBIERNO DE LA CIUDAD DE MÉXICO



ECOBICI



Encuentra tu Ecobici



Plataforma: Mapa - Ecobici

URL: <https://ecobici.cdmx.gob.mx/mapa/>

Por otro lado, un ejemplo de cómo una institución puede utilizar y aplicar los datos abiertos creados por otra institución es el caso del **Sistema de Indicadores de Género (SINGE)** de la Secretaría de las Mujeres, quienes recopilaron, analizaron y difundieron información estadística relacionada con la situación y posición de las mujeres en la Ciudad de México a través de una plataforma de fácil acceso y que contiene información, visualizaciones, metadatos y las fuentes de datos de cada indicador.

Plataforma: Sistema de Indicadores de Género (SINGE)

URL: <https://indicadoresdegenero.semujeres.cdmx.gob.mx/>

Indicador: Alcaldesas y alcaldes de la Ciudad de México por sexo

Fuente de datos: Instituto Electoral de la Ciudad de México

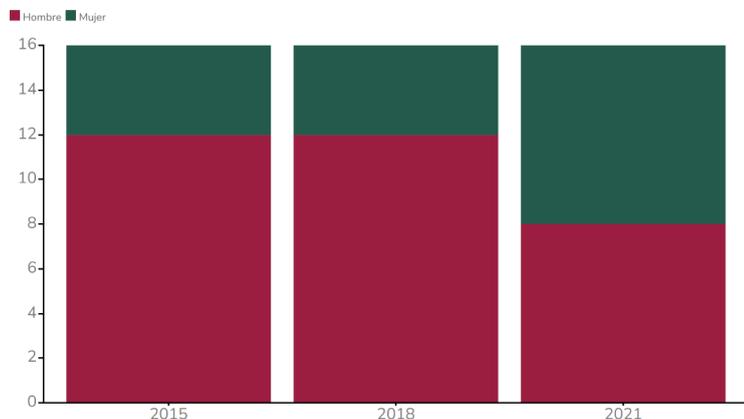


INMUJERES
INSTITUTO NACIONAL DE LAS MUJERES



GOBIERNO DE LA
CIUDAD DE MÉXICO

SECRETARÍA
DE LAS MUJERES



	A	B	C	D
1	Año	Sexo	Total	Porcentaje
2	2015	Hombre	12	75
3	2015	Mujer	4	25
4	2018	Hombre	12	75
5	2018	Mujer	4	25
6	2021	Hombre	8	50
7	2021	Mujer	8	50

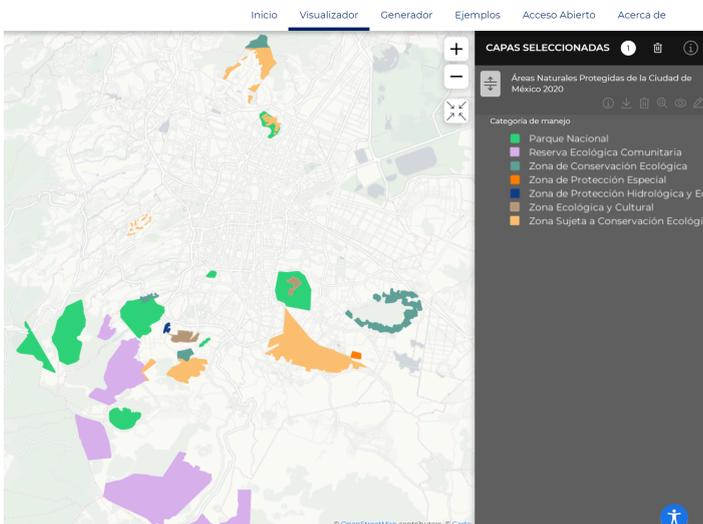
Otro ejemplo de este tipo es **Gema**, un gestor de mapas del Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) que permite que las personas usuarias puedan visualizar y generar cartografías del territorio nacional a partir de bases de datos de varias instituciones tanto federales como estatales, tales como el INEGI, INPI, SEP, Conagua, SEDEMA, entre otros. *Gema* toma su nombre de la combinación de las palabras “gestor” y “mapas”, haciendo alusión a una piedra preciosa y al proceso paulatino para convertirse en una gema, mismo que se puede comparar con la transformación de datos en información y conocimiento.

Plataforma: Gema

URL: <https://gema.conahcyt.mx/>

Mapa: Áreas Naturales Protegidas de la Ciudad de México 2020

Fuente de datos: Secretaría del Medio Ambiente



	A	B	C	D	E
	g_id	nombre	categoria	area_real	perimetro
1					
2	1	Bosque de Tlalpan	Zona Ecológica y Cultural	2528596.9	103.507.422
3	2	Bosques de Las Lomas	Zona Sujeta a Conservación Ecológica	3.087.642.612	138.990.907
4	3	Ecoguardas	Zona de Conservación Ecológica	1.326.358.126	51.086.034
5	4	Ejidos de Xochimilco y San Gregorio	Zona Sujeta a Conservación Ecológica	25224420.37	383.772.655
6	5	La Armella	Zona de Conservación Ecológica	1.932.508.953	87.090.604
7	6	La Loma	Zona de Conservación Ecológica	7.572.621.596	83.844.907
8	7	Los Encinos	Zona de Protección Hidrológica y Ecológica	2.501.015.543	44.890.221
9	8	Parque Ecológico de la CDMX	Zona Sujeta a Conservación Ecológica	7.276.189.543	223.936.348
10	9	San Bernabé Ocoatepec	Reserva Ecológica Comunitaria	2.403.808.332	90.830.543
11	10	San Miguel Ajusco	Reserva Ecológica Comunitaria	11759808.5	144.078.776
12	11	San Miguel Topilejo	Reserva Ecológica Comunitaria	60003293.7	579.431.152
13	12	San Nicolás Totolapan	Reserva Ecológica Comunitaria	19895004.16	281.393.484
14	13	Sierra de Santa Catarina	Zona de Conservación Ecológica	7.497.141.267	56.172.487
15	14	Cerro de La Estrella	Parque Nacional	1.0615470.64	213.476.277
16	15	Cumbres del Ajusco	Parque Nacional	5.013.857.525	101.181.334
17	16	Desierto de los Leones	Parque Nacional	15239366.58	178.416.436
18	17	El Tepeyac	Parque Nacional	1.648.690.188	193.861.691
19	18	Fuentes Brotantes de Tlalpan	Parque Nacional	2.202.754.378	32.984.996
20	19	Lomas de Padierna	Parque Nacional	11590912.74	153.272.257
21	20	Histórico Coyoacán	Parque Nacional	3.976.904.906	3.976.904.906
22	21	Cerro de La Estrella	Zona Ecológica y Cultural	1.217.801.842	80.196.613
23	22	Sierra de Cuadalupe	Zona Sujeta a Conservación Ecológica	6.335.327.944	468.862.455
24	23	Insurgente Miguel Hidalgo y Costilla	Parque Nacional	1.263.335.992	22.751.625
25	24	Insurgente Miguel Hidalgo y Costilla	Parque Nacional	5.124.146.557	136.990.735
26	25	Templulli	Zona de Protección Especial	4.775.511.711	29.944.096

Los datos abiertos que generan y publican los actores y las instituciones públicas sobre sus actividades y resultados abonan a la rendición de cuentas porque posibilitan la realización de estudios e investigaciones por parte de personas académicas, periodistas e investigadoras de la sociedad civil.

De la misma manera, es una buena práctica a impulsar desde la sociedad civil y la academia, publicar las bases de datos abiertas utilizadas y resultantes de sus investigaciones, con el fin de que los análisis sean replicables.



3. ¿Cómo publicar datos abiertos?

Si a una institución le interesa publicar datos abiertos, el primer paso es identificar cuáles son los datos que potencialmente pueden convertirse en datos abiertos, ya sean datos existentes o datos que en un futuro se planea generar. Una vez hecho este levantamiento, de acuerdo con la Política Nacional de Datos Abiertos se debe reunir toda esta información en un **inventario de datos**, el cual debe actualizarse periódicamente.

Una vez elaborado el inventario institucional de datos, se deberán **analizar y priorizar** los conjuntos de datos para su publicación. El análisis y priorización de los conjuntos de datos permitirá seleccionar y organizar el proceso para la apertura y posterior publicación. De acuerdo con la Política Nacional de Datos Abiertos, la priorización se hace de la siguiente manera:

Mayor prioridad	1 Conjuntos de datos que se encuentren asociados a la planeación institucional, así como a las atribuciones y facultades del sujeto obligado.
↓	2 Conjuntos de datos que resulten estratégicos o relevantes por contribuir al propósito y objetivos de política pública institucional o del sector (salud, educación, seguridad, etc.).
	3 Conjuntos de datos que respondan a necesidades concretas de información de audiencias estratégicas o grupos poblacionales.
Menor prioridad	4 Conjuntos de datos que hayan sido definidos como mínimos o prioritarios derivado de la construcción o implementación de la Política Nacional.

Al instrumento que marcará la ruta de trabajo que se llevará a cabo para aperturar y poner a disposición los conjuntos de datos se le llama **Plan de Apertura** y cada sujeto obligado debe contar con el suyo.

El siguiente paso es la **preparación y transformación** de los datos. Cuando se abren datos, éstos deben ser accesibles para cualquier persona usuaria, independientemente de su conocimiento y experiencia en el manejo y uso de datos. Por esta razón es importante que al preparar los datos para su publicación se sigan ciertas reglas o estándares (también conocidas como “mejores prácticas”) que tengan como finalidad facilitar el uso de datos. Otra razón por la que es importante que se sigan los estándares establecidos es que existe la posibilidad de que una persona usuaria requiera usar más de una base de datos. Por ejemplo, si hay una base de datos que se actualiza anualmente puede que haya una usuaria o usuario que necesite los datos de los últimos tres años; o también puede ser el caso de una persona que esté realizando un estudio o investigación y requiere utilizar datos de



múltiples fuentes. Por lo tanto, si todos los datos abiertos se publican siguiendo los mismos estándares entonces será mucho más sencillo combinar y manejar datos de varias fuentes, consiguiendo así obtener un mayor aprovechamiento de toda la información disponible. Aquí te mencionaremos algunas de estas reglas.

FORMATOS ABIERTOS. La primera recomendación es una que ya se mencionó anteriormente y es convertir los datos en formatos abiertos (.csv, .shp, .kml, etc.), los cuales son altamente accesibles para el público general pues existen programas y aplicaciones gratuitas y de código abierto con las cuales se pueden manejar y procesar los datos fácilmente.

NOMBRES DE COLUMNAS. Es sumamente recomendable que los nombres de las columnas en una base de datos tabular sean claros (sin ser demasiado extensos) para que cualquier persona usuaria pueda entender rápidamente y sin dificultad sobre qué son los datos. En general, se recomienda que el nombre de las columnas se escriba con letras minúsculas, sin acentos y que los espacios se reemplacen por un guión bajo “_”, por ejemplo: **primer_apellido, codigo_postal, alcaldia**. Por otra parte, se aconseja evitar poner encabezados en las bases de datos abiertos. Es común encontrarse con archivos donde en la parte superior se colocan logotipos o información referente a la institución que generó la base de datos, sin embargo es recomendable que al publicar datos abiertos se evite colocar este tipo de encabezados para que el primer renglón del archivo corresponda al nombre de las columnas de la base de datos. Además, si vas a publicar datos tabulares de manera periódica (por ejemplo cada trimestre o cada año) es recomendable que los nombres de las columnas se mantengan constantes. Por ejemplo, si publicaste una tabla de datos abiertos con la columna **fecha_envio** entonces se sugiere que cuando actualices los datos mantengas ese nombre y no otras variantes como **FECHA** o **fecha_de_envio**.

ESCRITURA CONSISTENTE. Es de gran importancia que el texto tenga buena ortografía y gramática. En las ocasiones en las que un término se repite varias veces en una misma columna se sugiere mantener el mismo formato para ese término (acentuación, puntuación, uso de letras mayúsculas y minúsculas, uso de espacios, etc.) en toda la columna.



ciudad
Ciudad De México
Ciudad de México
CDMX
México
Ciudad de Mexico
Ciudad de México.



ciudad
Ciudad de México

Otro ejemplo de este tipo es el de las columnas con fechas. En este caso lo que se sugiere es que se mantenga constante el formato en que se escribe la fecha en toda la columna, ya sea **día/mes/año**, **día-mes-año** o cualquier otro formato de tu elección.



fecha
8 de abril de 2024
12/01/2024
23-09-22
7/julio/23



fecha
08/04/2024
12/01/2024
23/09/2022
07/07/2023

NOMBRE DE LOS ARCHIVOS. Otra recomendación es que cuando se genere un archivo se le asigne un nombre fácil de leer y que deje claro cuál es el contenido del archivo, así como la fecha de creación o cualquier otra descripción que pueda ser relevante. Por ejemplo:

Datos: Sitios de servicios de atención a la violencia de género durante contingencia COVID19

Archivo: *servicios-atencion-violencia-genero-covid19.csv*

NOTA: En la medida de lo posible, asegúrate de que los datos abiertos que vas a publicar estén verificados y sean confiables.



NOTA: Es importante que antes de publicar cualquier tipo de dato, te asegures de que no es información sensible o que pueda poner en riesgo la seguridad de una persona.

Es necesario asegurar que los conjuntos de datos incluyan **metadatos**, pues facilita el entendimiento del origen, preparación y significado de los datos.

Nombre	Etiqueta	Tipo	Descripción
num_ciclo	Número de cicloestacion	Numérico	Número de cicloestacion
calle_prin	Calle principal	Texto	Calle principal
calle_secu	Calle secundaria	Texto	Calle secundaria
colonia	Colonia	Texto	Colonia
alcaldia	Alcaldía	Texto	Alcaldía
latitud	Latitud	Numérico	Latitud
longitud	Longitud	Numérico	Longitud
sitio_de_e	Sitio de instalación	Texto	Sitio de instalación
sistema	Sistema que la opera	Texto	Sistema que la opera el ciclistacionamiento
estatus	Estatus de instalación	Texto	Estatus de instalación, si ya está instalada o está pendiente de instalación.

Plataforma: Portal de Datos Abiertos

Datos: Metadatos de Cicloestaciones Ecobici

Archivo: *diccionario_cicloestaciones_ecobici.xlsx*

Institución: Secretaría de Movilidad

URL:

<https://datos.cdmx.gob.mx/dataset/afluencia-diaria-servicio-de-transportes-electricos>

Además, para asegurar que los datos sean de uso libre se tienen que publicar con alguna de las **licencias** diseñadas para datos abiertos, provistas por la Red de la Fundación del Conocimiento Abierto (Open Knowledge Foundation Network), específicamente:

- *Open Data Commons versión 4.0, Open Data Commons Public Domain Dedication and License (PDDL)* que permite difundir, reutilizar o adaptar los datos sin restricción alguna
- *Open Data Commons Attribution License* al igual que la licencia de Creative Commons Attribution License, se exige la referencia a la autoría o fuente de los datos para la reutilización de la información
- *Open Data Commons Open Database License (ODbL)* que permite la reutilización de los datos siempre que se reconozca la autoría de la información original; se mantenga la misma licencia en las obras derivadas las cuales pueden restringir su uso y si, además, se distribuye una versión sin dichas restricciones de uso.

Ya que se tiene una base de datos preparada y con sus respectivos metadatos y licencia, el siguiente paso es **publicarla** en una plataforma donde se puedan



almacenar, compartir, conectar y visualizar los conjuntos de datos y que sea accesible para el público en general a través de internet. Dos de las plataformas a nivel nacional que cumplen con estas características son el Portal nacional de Datos Abiertos de México (<https://datos.gob.mx/>) y la Plataforma Nacional de Transparencia. Para el caso particular de la Ciudad de México se cuenta con el Portal de Datos Abiertos (<https://datos.cdmx.gob.mx/>), el cual al momento de esta publicación cuenta 476 conjuntos de datos abiertos provenientes de 33 instituciones, clasificados en 21 categorías.



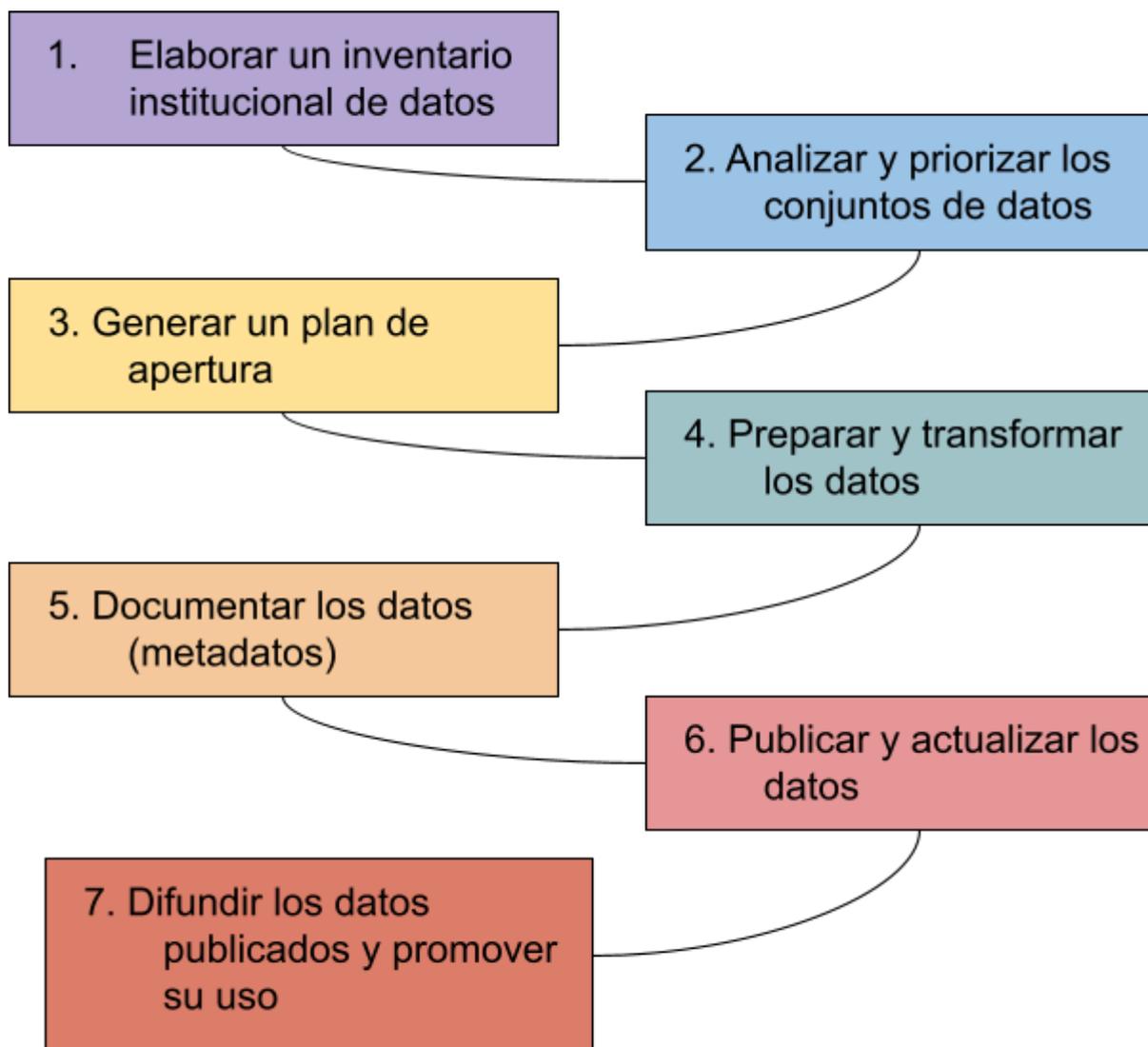
Si te interesa publicar tus datos abiertos en el Portal de Datos Abiertos de la Ciudad de México, puedes entrar al siguiente enlace y seguir los pasos ahí descritos:

https://politicadedatos.cdmx.gob.mx/Politica_datos

Una vez publicados los conjuntos de datos, éstos deberán ser **actualizados** de manera permanente con la intención de mantener vigente los datos.

La socialización o **difusión de los datos abiertos** consiste en el desarrollo de las actividades correspondientes para el acercamiento de los datos publicados en audiencias estratégicas y en la población en general, con el objetivo de potenciar las ventajas y los beneficios que pudieran resultar del aprovechamiento de los datos.

En resumen, las etapas a seguir para la apertura de datos son:



Fuente: Elaboración propia

4. ¿Dónde puedo aprender más?

Todas las personas con conocimientos en un nivel básico en el uso de un programa de hojas de cálculo como Microsoft Excel, Google Sheets, OpenOffice Calc o Apple Numbers pueden hacer uso de datos abiertos que estén disponibles en formato .csv

Para otro tipo de formatos, se requiere un conocimiento básico de lenguajes de programación de código abierto y gratuitos como R o Python, que permiten leer todos los formatos de datos abiertos.

Si te interesa saber más sobre los datos abiertos, cómo se producen y visualizan y cómo utilizarlos y aprovecharlos, el ecosistema de opciones en línea ha crecido mucho en los últimos años y están disponibles muchos cursos y tutoriales. Por mencionar algunos:



Cursos virtuales y presenciales	Actualmente están disponibles en línea diversas plataformas de cursos como Escuela de Datos de SocialTIC o Datacrunchers, así como tutoriales y grupos de personas usuarias de datos abiertos y programación en redes sociales.
Licenciaturas, posgrados y especializaciones	La oferta académica actual en México tiene programas que van desde Ingeniería en Computación hasta Ciencia de Datos, en niveles de licenciatura, posgrado y diplomados.
Open Data Day	El primer sábado de marzo de cada año, la comunidad dataera se reúne para celebrar los datos abiertos y compartir proyectos y aprendizajes.
Hackatones y Datatones	Cada año, diversas instituciones públicas y organizaciones sociales crean espacios de colaboración para generar retos y productos que utilicen datos abiertos, con una temática focalizada y enfoque social.

El equipo de Estado Abierto del InfoCDMX ha realizado diferentes actividades de asesoría en el uso y aprovechamiento de los datos abiertos.

- Segunda Jornada de Asesorías Técnicas Especializadas 2022: Apertura Institucional <https://www.youtube.com/watch?v=n0B46G-BYS8&t=4556s>
- Jornada por la Inclusión y los Datos Abiertos https://www.youtube.com/watch?v=SHKIS_rzcYk
- Repositorio de información de la por la Inclusión y los Datos Abiertos <https://infocdmx.org.mx/images/biblioteca/2023/Documento-Jornada-por-la-inclusion-y-los-datos-abiertos.pdf>
- Portal de Datos Abiertos de Transparencia: <https://datosabiertos.infocdmx.org.mx/>
- Estadísticas de solicitudes de información pública en la Plataforma Nacional de Transparencia: https://infocdmx.org.mx/images/biblioteca/Estadisticas_de_Solicitudes_de_Informacion_en_la_PNT.pdf

Conclusión

El mundo está pasando por una era digital en la que hay un avance acelerado en cuanto a tecnología e información, por lo que es de gran importancia identificar las áreas de oportunidad y aprovechar estos avances para lograr transformar los modelos de gobierno y alcanzar uno que sea mejor para toda la sociedad. Una de estas grandes oportunidades está en los datos abiertos, los cuales son una



herramienta poderosa al alcance de cualquier persona y que tiene un potencial incalculable para encontrar y crear soluciones a problemas en distintos sectores, como medio ambiente y seguridad. Por esta razón es que vale la pena que como sociedad se haga un esfuerzo tanto para generar datos abiertos como para aprovecharlos y utilizarlos como base de soluciones innovadoras e incluyentes. En la Ciudad de México existe actualmente un avance considerable en datos abiertos, sin embargo aún hay trabajo por hacer para alcanzar el máximo potencial de los datos.

Referencias

[1] México, A. (2023). Política Nacional de Datos Abiertos. *Abramos México*. Tomado de <https://abramosmexico.pubpub.org/pub/politica-nacional-de-datos-abiertos>

[2] Ley General de Transparencia y Acceso a la Información Pública (LGTAIP). <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGTAIP.pdf>

[3] Guía para la apertura de datos en organismos de la Administración Pública Nacional. Paquete de Apertura de Datos de la República Argentina. <https://datosgobar.github.io/paquete-apertura-datos/>

[4] Zins, Chaim. *Conceptual Approaches for Defining Data, Information, and Knowledge*. Journal Of The American Society For Information Science And Technology, 2007

[5] Open Data Handbook. <https://opendatahandbook.org/guide/es/how-to-open-up-data/>

[6] La diferencia entre Datos Abiertos y Gobierno Abierto. <https://datos.gob.mx/blog/la-diferencia-entre-datos-abiertos-y-gobierno-abierto>

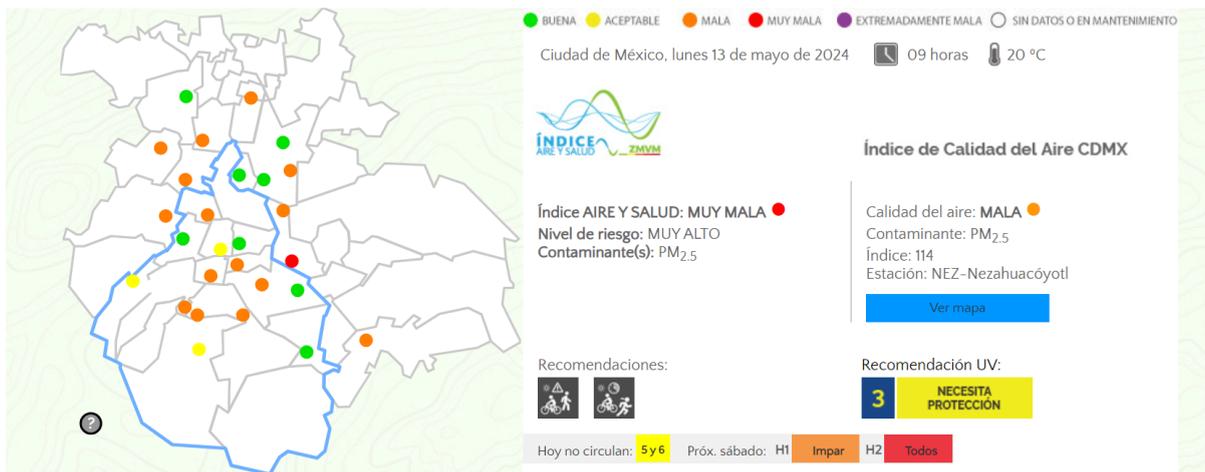
[7] Portal de la Política de Datos de la Ciudad de México. <https://politicadedatos.cdmx.gob.mx/>



Otros ejemplos



Inicio | Mapa del sitio | Preguntas frecuentes | Contacto | CDMX | SEDEMA |

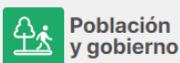


CALIDAD DEL AIRE CDMX

<http://www.aire.cdmx.gob.mx/default.php>

MONITOR ESTADÍSTICO DE LA CIUDAD DE MÉXICO

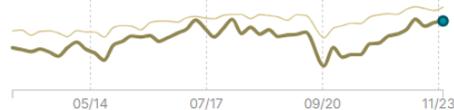
Alcaldía:



96% Tasa de Ocupación

Porcentaje de personas que cuentan con un empleo respecto a la población económicamente activa

De acuerdo con la última actualización de la Encuesta Nacional de Ocupación y Empleo (ENOE), la tasa de ocupación cerró en 2023 con un valor de 96.09%. Manteniendo una tendencia a la alza, con relación a los dos trimestres anteriores de ese año.

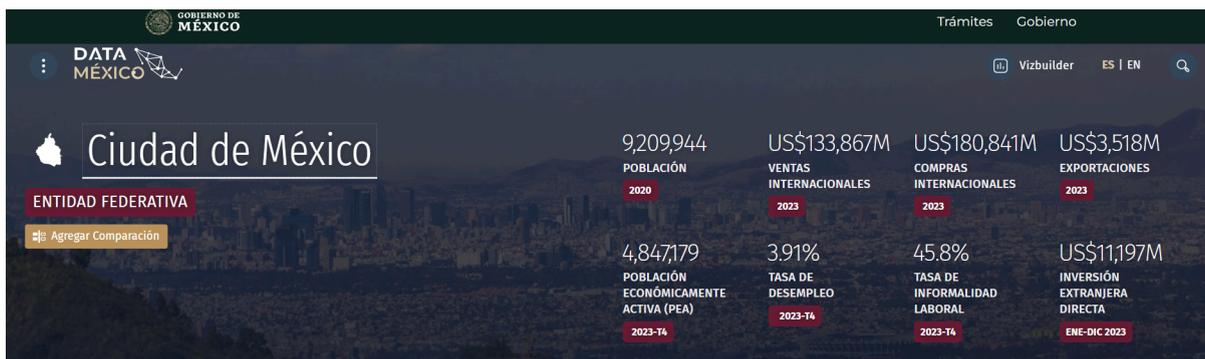


Datos



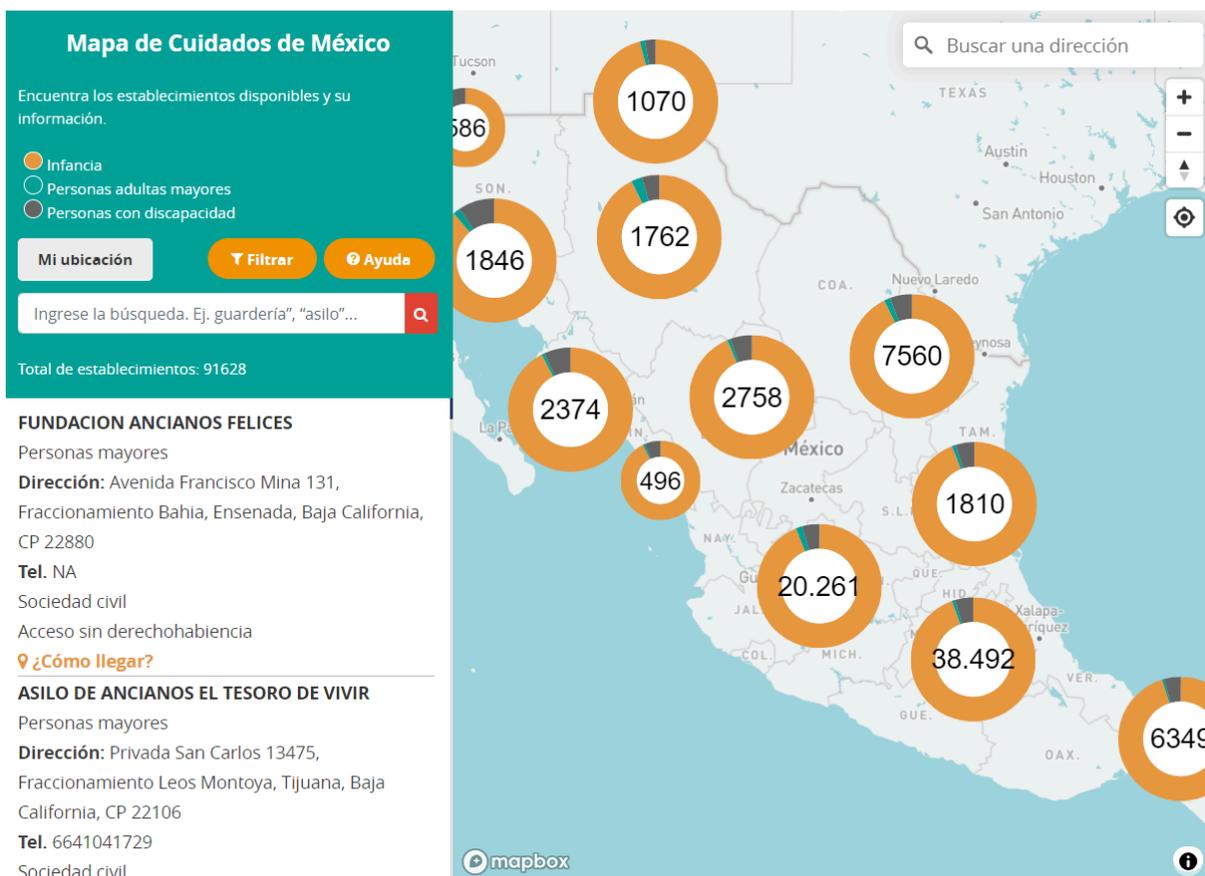
MONITOR ESTADÍSTICO DE LA CIUDAD DE MÉXICO

<https://visorestadistico.cdmx.gob.mx/>



Ciudad de México: Economía, empleo, equidad, calidad de vida, educación, salud y seguridad pública | Data México

<https://www.economia.gob.mx/datamexico/es/profile/geo/ciudad-de-mexico-cx>



Mapa de Cuidados de México

<https://mapadecuidados.inmujeres.gob.mx/buscador/>