

Trayecto Datos

Clase 4: Datos abiertos



Proyecto de Análisis de Datos

- Definición del problema
- Búsqueda de datos
- Análisis exploratorio
- Limpieza
- Visualizaciones
- Presentación de resultados



Fuente: <https://www.actionsdata.com/blog/tendencias-en-analisis-de-datos>

Definición del problema o tema

- ¿Cuál es el tema que queremos tratar?
- ¿Con qué datos contamos para hacerlo?
- ¿Cuántos datos tenemos?
- ¿El tema es de interés?
- ¿Cómo lo queremos abordar?



<https://www.thinkingondata.com/tag/svn/>

Datos Abiertos

Como analistas de datos, nuestra materia prima para trabajar son los datos. Si trabajamos en una empresa probablemente necesitemos crear y usar datos generados por la misma, pero, para proyectos educativos o de investigación, se suelen trabajar con datos abiertos.

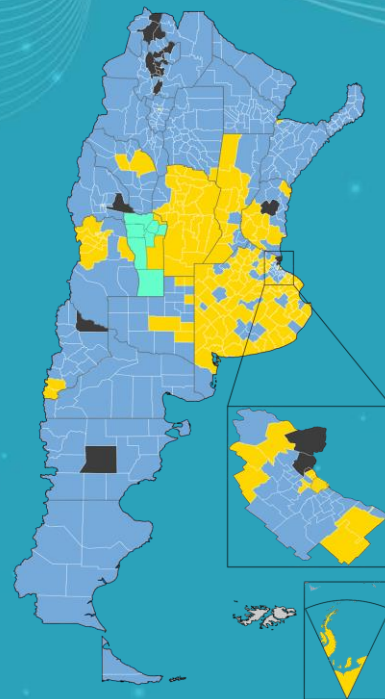
Una de las grandes fuentes de datos abiertos son los gobiernos, quienes suelen colocar información gratuita sobre diferentes aspectos de las naciones.



Datos Abiertos en Argentina

La Ley de Acceso a la Información Pública establece que la información en poder del Estado debe ser accesible para todas las personas y estar disponible en formatos electrónicos abiertos para facilitar su circulación y redistribución. En el año 2016, mediante Decreto 117/2016 el gobierno nacional impulsó el “Plan de Apertura de Datos”.

Datos Abiertos es una iniciativa global, ligada a las políticas de Gobierno Abierto. Se trata de un medio que posibilita un mejor conocimiento del funcionamiento del gobierno, el fortalecimiento del rendimiento de cuentas y la mejora de la vida en ciudadanía.



datos.gob.ar

Datos Argentina

Datos Argentina (datos.gob.ar) es el sitio web oficial donde se publican los datos abiertos de la nación.

Los invitamos a recorrer el sitio.

The screenshot displays the homepage of the Datos Argentina website. At the top, the logo 'datos.gob.ar' is on the left, and navigation links for 'Datasets', 'Series', 'Organizaciones', 'APIs', and 'Acerca' are on the right. The main header features the title 'Datos Argentina' and a descriptive paragraph: 'Ponemos a tu alcance datos públicos en formatos abiertos para que puedas usarlos, modificarlos y compartirlos. Estos datos son tuyos. Podés crear visualizaciones, aplicaciones y grandes herramientas con ellos.' To the right of the header, two large statistics are shown: '1019 DATASETS' and '033 ORGANIZACIONES CON DATOS'. Below the header is a search bar with the placeholder text '¿Qué dataset buscás?' and a magnifying glass icon. The main content area consists of a grid of category tiles, each with an icon and a label: 'Agroganadería, pesca y forestación', 'Asuntos internacionales', 'Ciencia y tecnología', 'Economía y finanzas', 'Educación, cultura y deportes', 'Energía', 'Gobierno y sector público', 'Justicia, seguridad y legales', 'Medio ambiente', 'Población y sociedad', 'Regiones y ciudades', and 'Salud'. A 'Transporte' tile is located at the bottom right of the grid.

datos.gob.ar

Datasets Series Organizaciones APIs Acerca ▾

Datos Argentina

Ponemos a tu alcance datos públicos en formatos abiertos para que puedas usarlos, modificarlos y compartirlos. Estos datos son tuyos. Podés crear visualizaciones, aplicaciones y grandes herramientas con ellos.

1019
DATASETS

033
ORGANIZACIONES
CON DATOS

¿Qué dataset buscás?

Agroganadería, pesca y forestación

Asuntos internacionales

Ciencia y tecnología

Economía y finanzas

Educación, cultura y deportes

Energía

Gobierno y sector público

Justicia, seguridad y legales

Medio ambiente

Población y sociedad

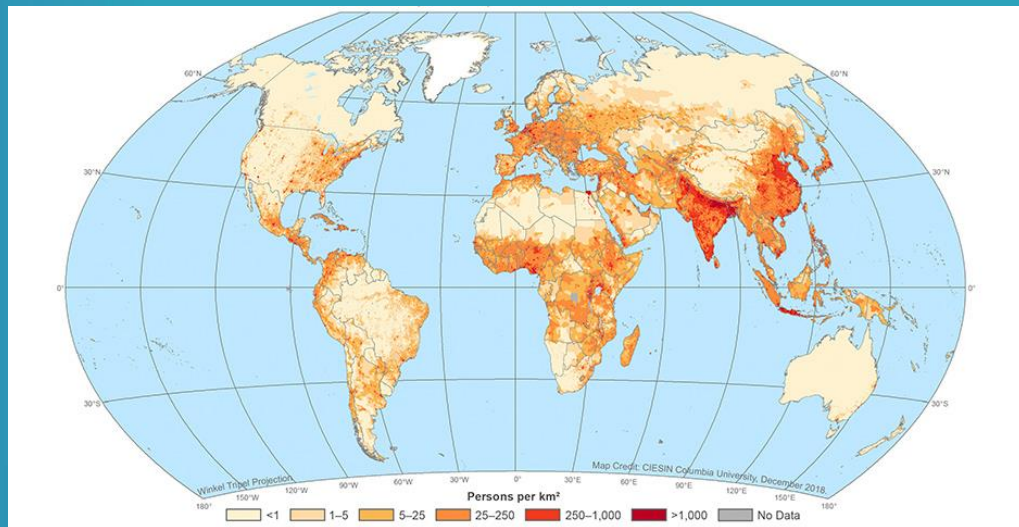
Regiones y ciudades

Salud

Transporte

Datos del mundo

Una fuente de datos mundiales en temas de economía, salud, educación, etc. Es el banco mundial de datos.



Redes sociales

Otra fuente de datos abiertos son las API de redes sociales, para usar las mismas generalmente solo debes create una cuenta de desarrollador en la red social



Datos de todo tipo

Existen muchas otras fuentes de datos abiertas, en las que podemos encontrar datos de cualquier índole, una de las más populares en el mundo de la ciencia de datos es la página Kaggle, se debe tomar en cuenta que la mayoría de estos datos son subidos por la comunidad, por lo que es recomendable verificar la fuente para poder confiar en su veracidad.

The Kaggle logo, consisting of the word "kaggle" in a lowercase, blue, sans-serif font, is centered within a white rectangular box.

Este es un Documento de Google con sitios donde pueden buscar datasets, hay algunos dataset propuestos y sitios donde pueden buscar: [Documento](#)

¡MUCHAS GRACIAS!

