



INSTITUTO TECNOLÓGICO DE DURANGO

Analítica de datos en las organizaciones

UNIDAD 1

Prácticas de Laboratorio 1,2 & 3

Litzy Yulissa Nevarez García
19040372

07 de septiembre de 2022

Práctica de laboratorio: Estudiar fuentes de datos sin restricciones

Objetivos

Estudie los diferentes sitios web que proporcionan datos disponibles sin restricciones.

Parte 1: Open Government Partnership

Parte 2: Sitios web de datos sin restricciones

Parte 3: Visualización de los datos

Información básica/situación

A la fecha, se han generado y almacenado cantidades inconmensurables de datos. Los datos son útiles solamente si se los puede emplear para generar más conocimientos. Cuando los datos se ponen a libre disposición de otras personas, sin restricciones relacionadas con la propiedad intelectual, se los considera datos sin restricciones (o abiertos). Es posible que los datos sin restricciones requieran que se mencione la fuente original. Los productos derivados del trabajo original también se pueden designar como datos sin restricciones.

En esta práctica de laboratorio estudiará varias fuentes de datos sin restricciones. También estudiará la capacidad de visualización de datos disponible en estos sitios web para generar conclusiones basadas en hechos.

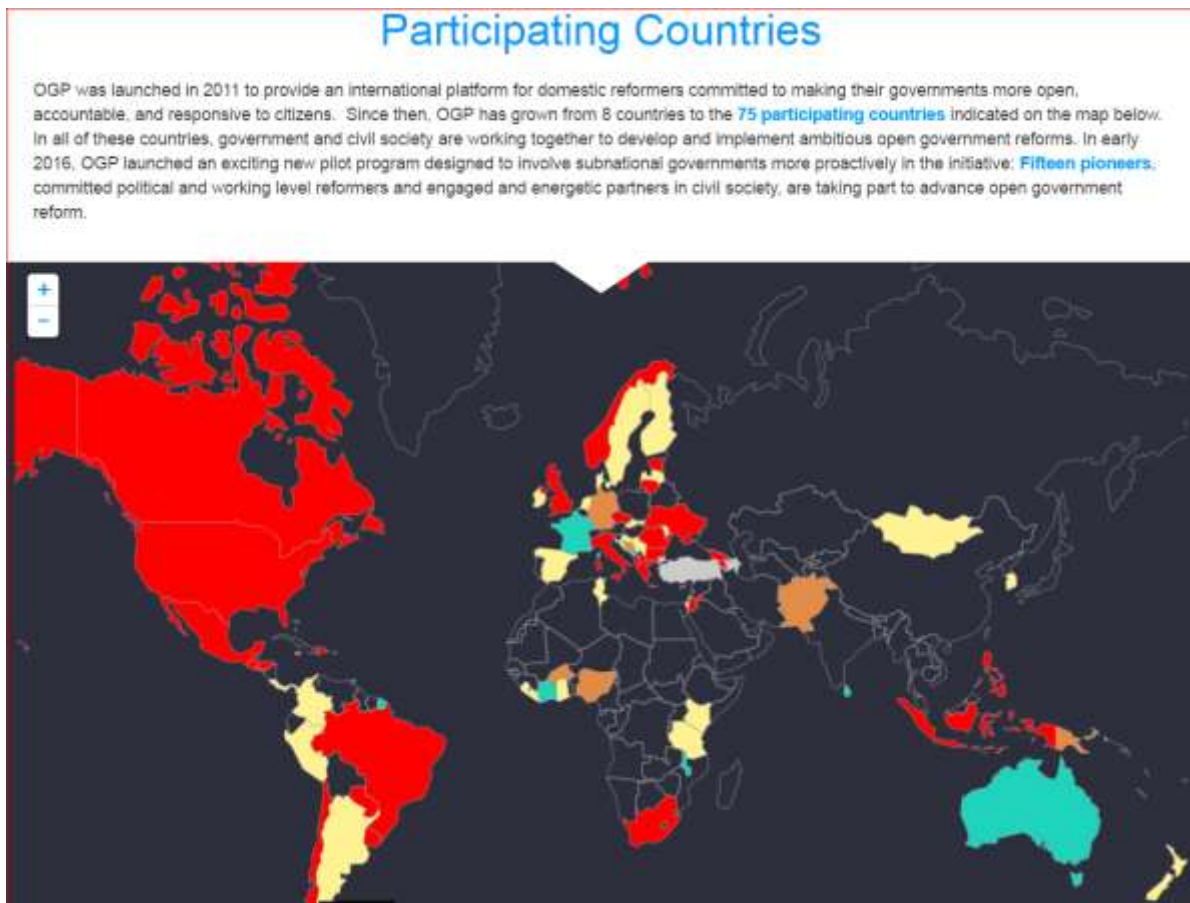
Recursos necesarios

- Computadora o dispositivo móvil con acceso a Internet

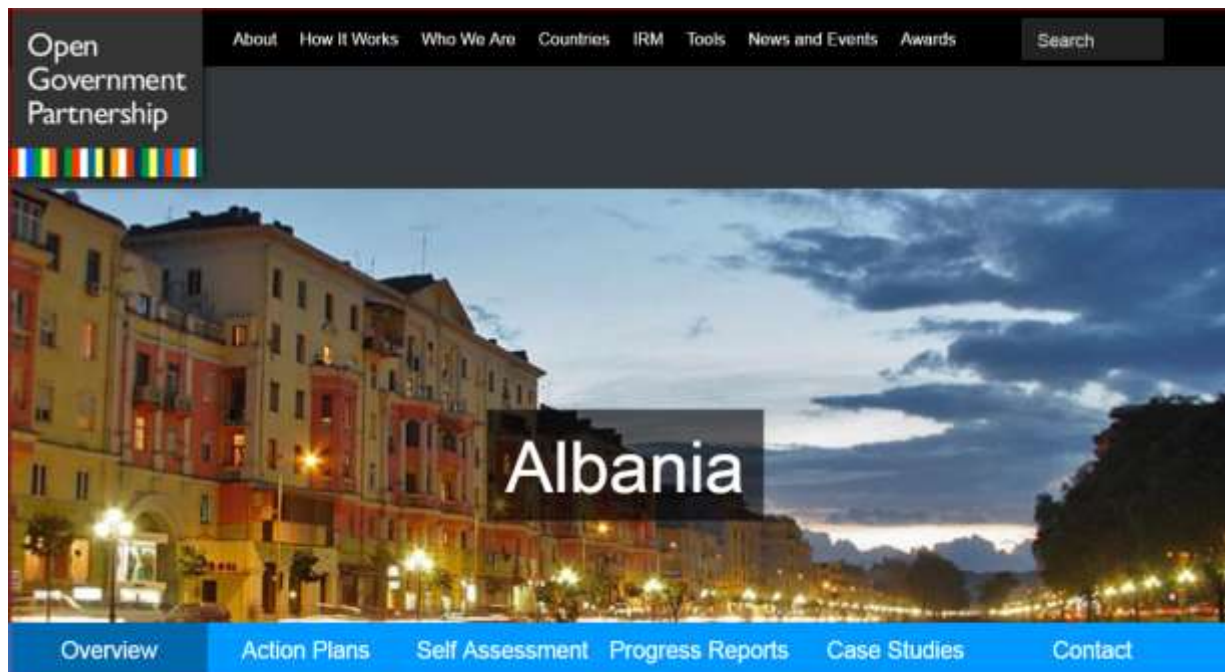
Parte 1: Open Government Partnership

Open Government Partnership (OGP) se fundó en 2011. La meta de la sociedad es crear una plataforma internacional con el fin de que los gobiernos se muestren más abiertos, sean más responsables y respondan mejor a sus ciudadanos.

- Diríjase a www.opengovpartnership.org/ para encontrar más información sobre OGP.
- Haga clic en **Countries** (Países) en el centro superior de la pantalla.



- c. Seleccione un país en el mapa o en la lista de la parte inferior de la página para ver información sobre la participación de dicho país en los datos sin restricciones. Con Albania como ejemplo, vemos que aparece una página con la descripción general de los pasos que siguió para unirse a OGP. También vemos gran cantidad de enlaces a información relacionada con la transición de ese país hacia un gobierno más abierto.



- d. En la página de cada país hay un anuncio de color azul con enlaces a información relacionada con su compromiso por lograr un gobierno abierto. Haga clic en **Case Studies (Casos de estudio)** para leer más acerca de los pasos seguidos.
- e. Haga clic en **Tools (Herramientas)** en la parte superior de la pantalla. En la página de herramientas podrá desarrollar su plan de acción, aprender de sus pares, seguir el progreso en OGP y transmitir su efecto.



- f. Por ejemplo: en la categoría Learn from Your Peers (Aprender de sus pares), haga clic en **Case Study** (Caso práctico). Aquí podrá buscar por temas (Topics), tipo de recurso (Resource Type) y país (Country). En Topics, seleccione Open Data (Datos sin restricciones). En Resource Type, seleccione Case Study. En Country, seleccione cualquiera de los países de esta lista. En la lista de casos prácticos verá ejemplos del efecto de los datos sin restricciones sobre las operaciones de los gobiernos y la participación de los ciudadanos en el gobierno. Busque un caso práctico que le interese y responda la siguiente pregunta.

En su carácter de ciudadano, ¿qué beneficios puede obtener si su gobierno es más abierto?

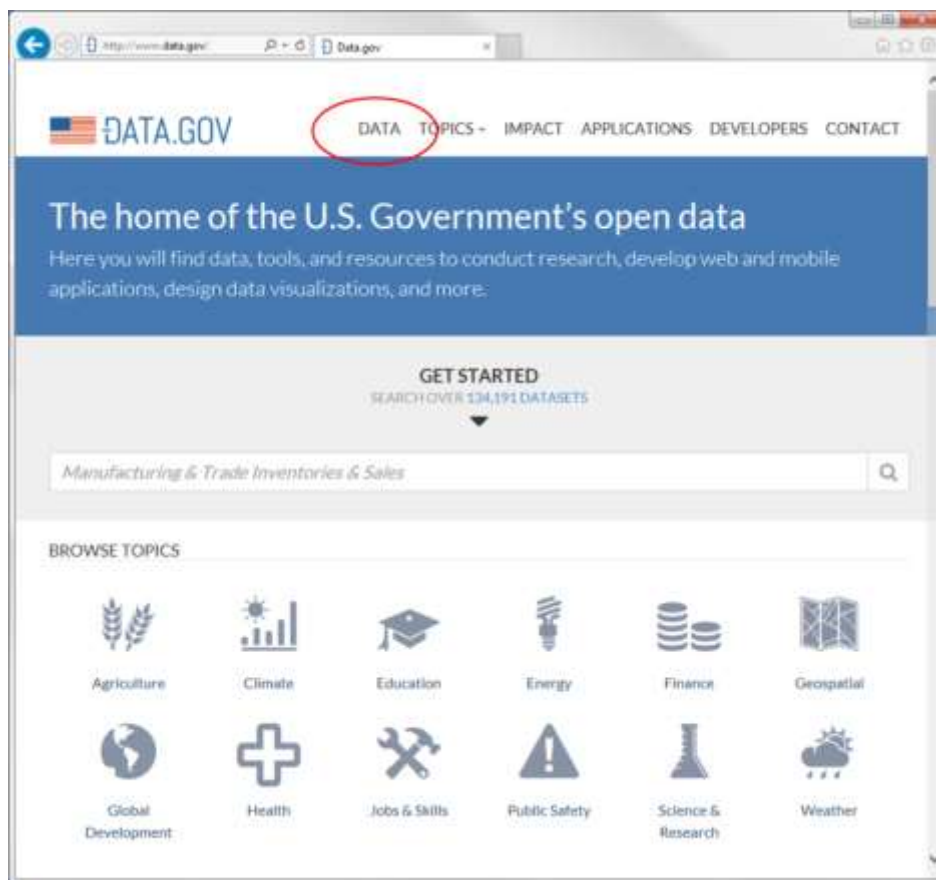
Abra mas oportunidades para todos porque no habrá restricciones de ciertos datos y los y las profesionistas pueden investigar mas y ser mas abiertos

Parte 2: Sitios web de datos sin restricciones

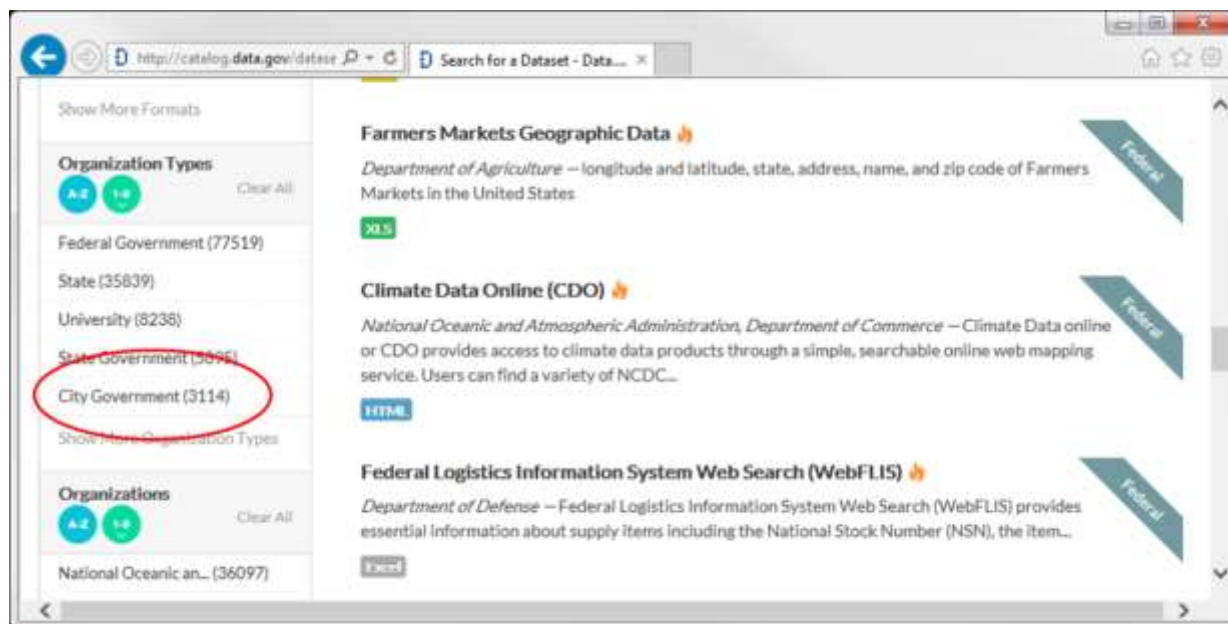
Muchas organizaciones (gobiernos, universidades, empresas y organizaciones sin fines de lucro) publican sus datos sin restricciones. ¿Le interesan los datos sobre el clima, sobre viviendas o sobre el delito?

- a. Navegue hasta <http://www.data.gov>. Este sitio web es la página central de los datos sin restricciones del Gobierno de EE. UU. Aquí podrá encontrar datos y recursos sobre investigaciones, desarrollo de software y otros proyectos relacionados con los datos. Desde este sitio web también podrá dirigirse a los sitios web originales que proporcionaron la información publicada.

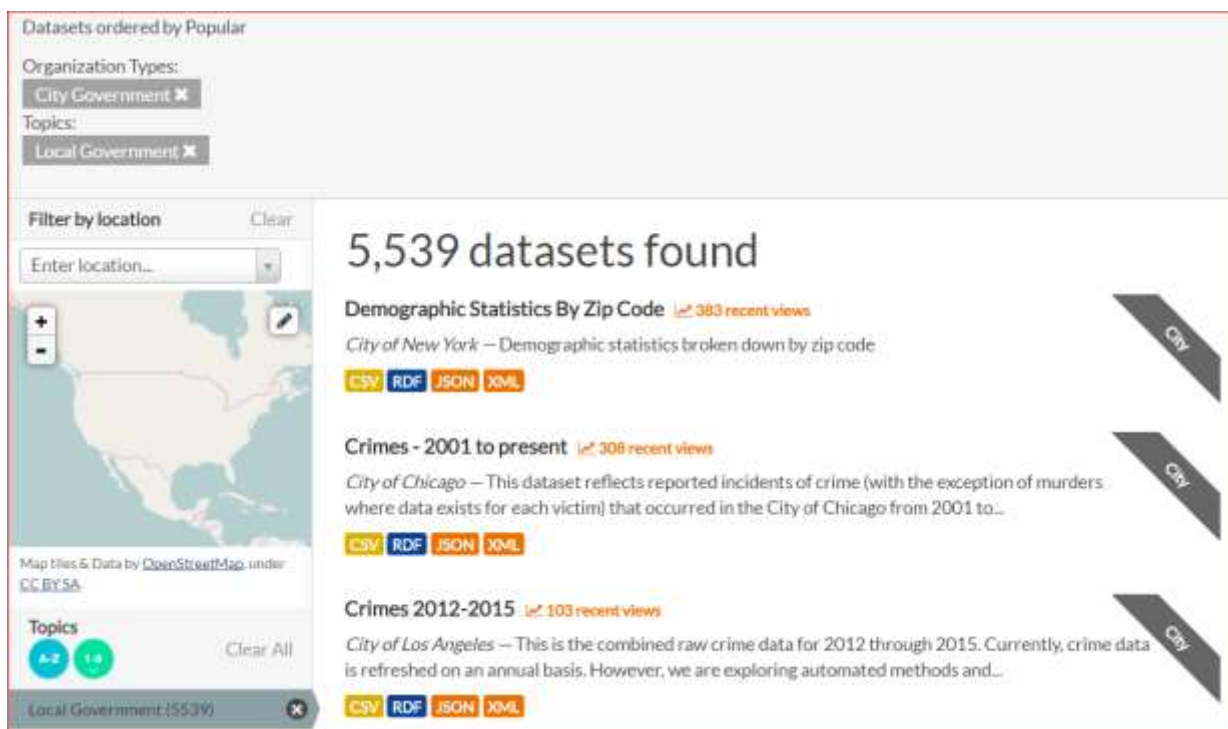
- b. Haga clic en **Data** (Datos) en la parte superior de la pantalla.



- c. ¿Cuántos conjuntos de datos hay disponibles? 317,801
- d. Estudie un conjunto de datos interesante u otra categoría temática. Por ejemplo: puede filtrar por diferentes categorías. Si solo está interesado en los datos provistos por el gobierno de una ciudad, haga clic en **Data** en la parte superior de la pantalla. Haga clic en **Local Government** (Gobierno local). Desplácese por la lista hasta ubicar **Organizational Types** (Tipos de organizaciones) en el panel izquierdo. Haga clic en **City Government** (Gobierno de la ciudad).




- e. Haga clic en **Crimes – 2001 to present** (Delitos: 2001 al presente). En este conjunto de datos se ofrecen las estadísticas delictivas correspondientes a la ciudad de Chicago del año 2001 al presente.



- f. Puede descargarlo si se dirige a la sección de descargas y recursos (Downloads and Resources).
¿Cuáles son los formatos de archivo para descargar disponibles?


ArcGis Hub Dataset, ArcGis GeoService, GeoJSON, CSV, KML, Shapefile

Downloads & Resources




Comma Separated Values File 11607 views

Download




RDF File 791 views

Download




JSON File 1739 views

Download



XML File 3153 views

Download



Landing Page

Visit page

Haga clic en el enlace de la **página de bienvenida** en la sección de descargas y recursos para dirigirse al portal de datos de la ciudad de Chicago correspondiente a la página de Delitos: 2001 al presente.

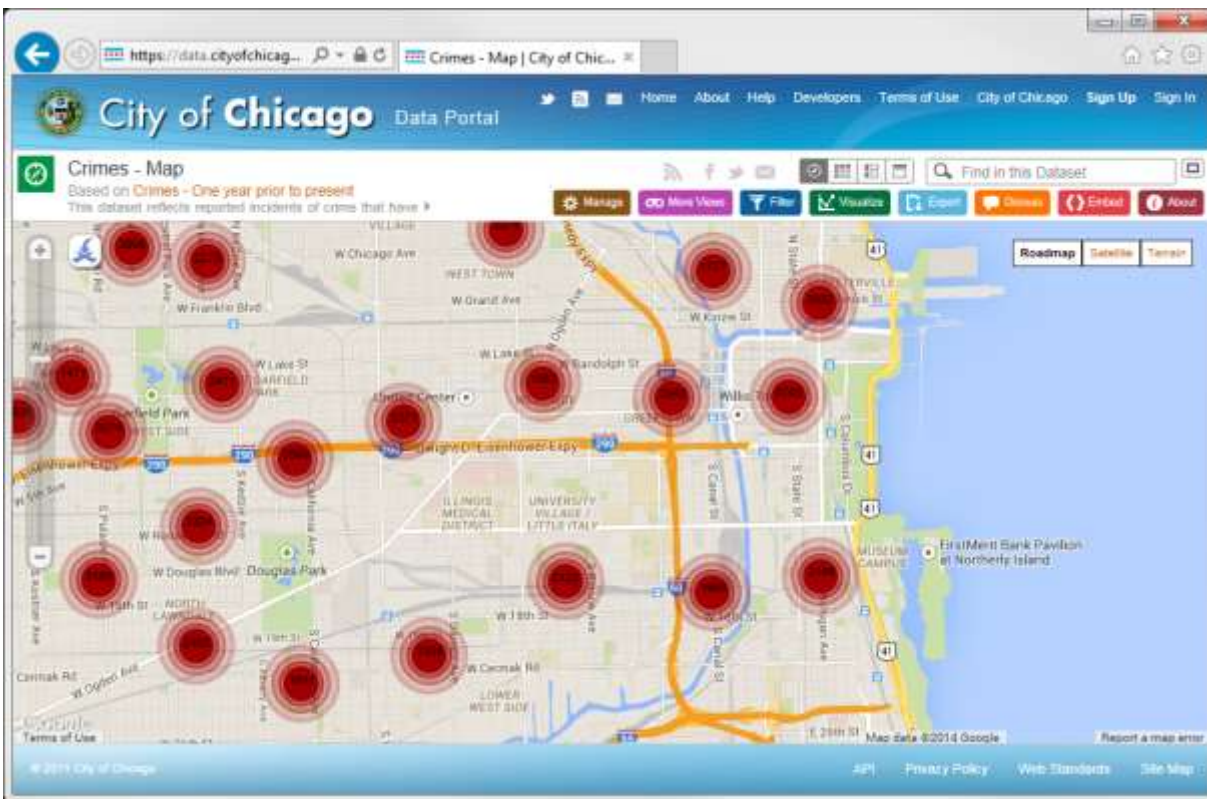


ID	Case Number	Date	Block	Side	Primary Type	Description	Location Description
1	9800786	JUL10/03	013002017 11 04 00 PM	0800	BURGLARY	HOUSE INVASION	APARTMENT
2	9800726	JUL10/03	013002017 11 30 00 PM	0400	CRIMINAL DAMAGE	TO VEHICLE	STREET
3	9800860	JUL10/03	013002017 11 30 00 PM	0800	BATTERY	DOMESTIC BATTERY SIMPLE	SIDEWALK
4	9800716	JUL10/03	013002017 11 40 00 PM	0800	BATTERY	AGE HANDSPRINTS/FEET SERIOUS BATTERY	STREET
5	9800707	JUL10/03	013002017 11 40 00 PM	0800	CRIMINAL DAMAGE	TO PROPERTY	APARTMENT
6	9800301	JUL10/03	013002017 11 40 00 PM	0800	BURGLARY	FORCIBLE ENTRY	BARBERSHOP
7	9800867	JUL10/03	013002017 11 30 00 PM	0800	ASSAULT	SIMPLE	APARTMENT
8	9800797	JUL10/03	013002017 11 30 00 PM	0400	BATTERY	DOMESTIC BATTERY SIMPLE	APARTMENT

- g. En esta página podrá ver los datos sobre delitos correspondientes a la ciudad de Chicago del año 2001 al presente sin descargarlos. También podrá ver una parte de estos datos en un mapa. En la parte superior de la página, haga clic en **City of Chicago** (Ciudad de Chicago) para dirigirse a la lista de conjuntos de datos disponibles del portal de datos de la ciudad de Chicago. En View Types (Tipos de vista), haga clic en **Maps** (Mapas).

- h. En la sección **View Types** del panel izquierdo, haga clic en **Maps** para ver una lista de mapas. Haga clic en el conjunto de datos de título **Crimes – Map** (Delitos: Mapa). Aparecerá un mapa interactivo del área de Chicago con los lugares puntuales en los que se cometieron delitos el año pasado.

- i. Haga clic en el conjunto de datos de título **Crimes – Map**. Aparecerá un mapa interactivo del área de Chicago con los lugares puntuales en los que se cometieron delitos el año pasado. Puede ir a cada uno de los lugares en los que se cometieron delitos si hace clic en los puntos de color rojo o si acerca la vista con la barra de desplazamiento de la izquierda.



- j. Si está interesado en otros datos de organización, puede regresar a <http://www.data.gov> y buscar otros temas. Si busca en Internet, podrá encontrar sitios web similares correspondientes a otros países. Por ejemplo:

Australia: <http://data.gov.au/>

Reino Unido: <http://data.gov.uk>

Canadá: <http://data.gc.ca>

Parte 3: Visualización de datos

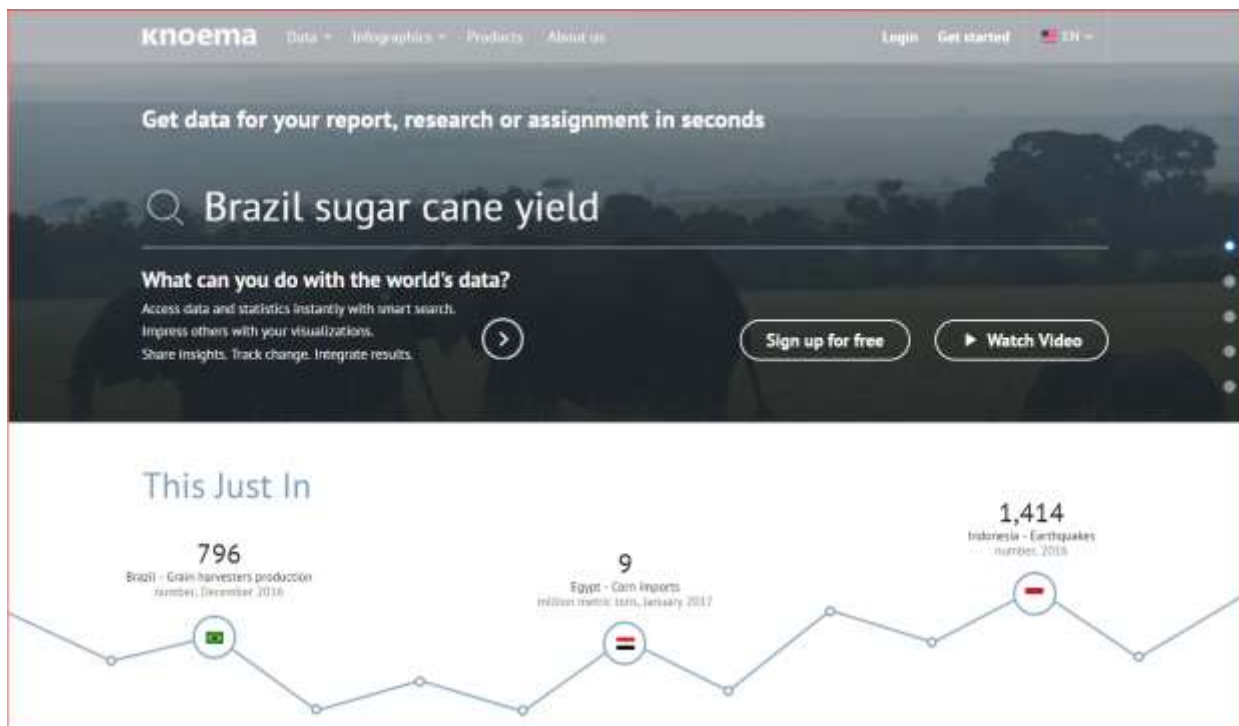
La visualización de datos puede ayudarle a interpretar listas de cifras y a exponerlas de un modo significativo. La visualización puede ayudarle a “ver” tendencias que podrían pasar desapercibidas de otro modo.

Diversos sitios web le proporcionan visualización de datos. Estudiará la visualización de datos con los gráficos que se proporcionan en estos dos sitios web: www.knoema.com y www.gapminder.org.

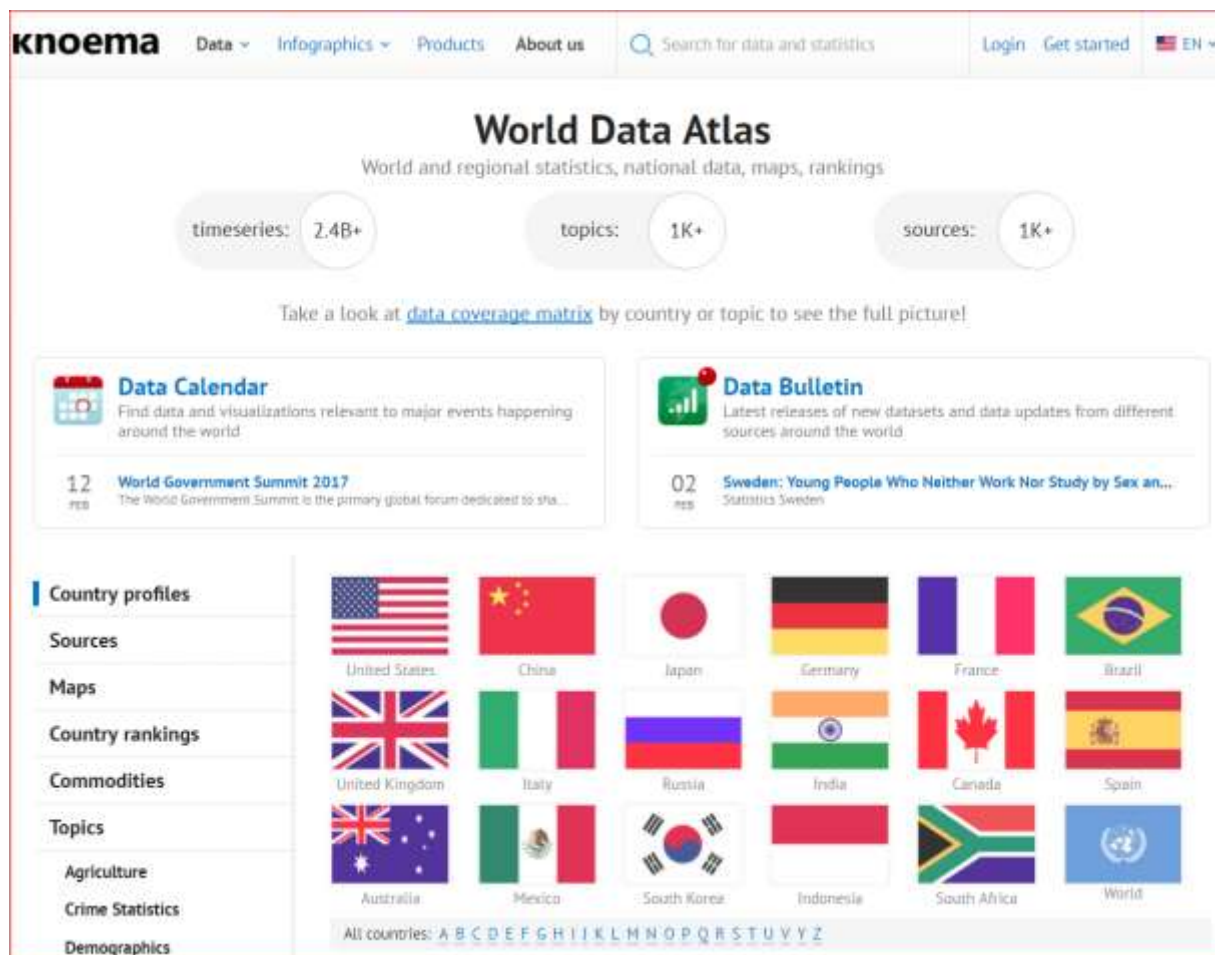
Paso 1: Navegar en knoema

En www.knoema.com puede encontrar muchas formas de representar datos estadísticos gráficamente.

- a. Desplácese hacia abajo en la página de inicio. Verá una lista de distintos temas, como This Just In (Lo más reciente) y Viz of the day (Visualización del día). Desplácese hacia abajo por la página para ver muchos otros temas que puede consultar.



- b. ¿Cuál es el tema de la Visualización del día?
Aluminum price forecast: 2021,2022 and long term to 2035
- c. Haga clic en el título y estudie los temas y los diferentes gráficos que se le presentan.
Asígnele nombres a las fuentes de datos.
Muestra el precio anual del aluminio en USD al igual de un pronostico
- d. En la parte superior de la página, haga clic en la **flecha hacia abajo** adyacente a **Data** (Datos). Haga clic en **World Data Atlas** (Atlas de datos mundial). Desde aquí (o desde cualquier categoría), puede estudiar la amplia variedad de datos disponibles.

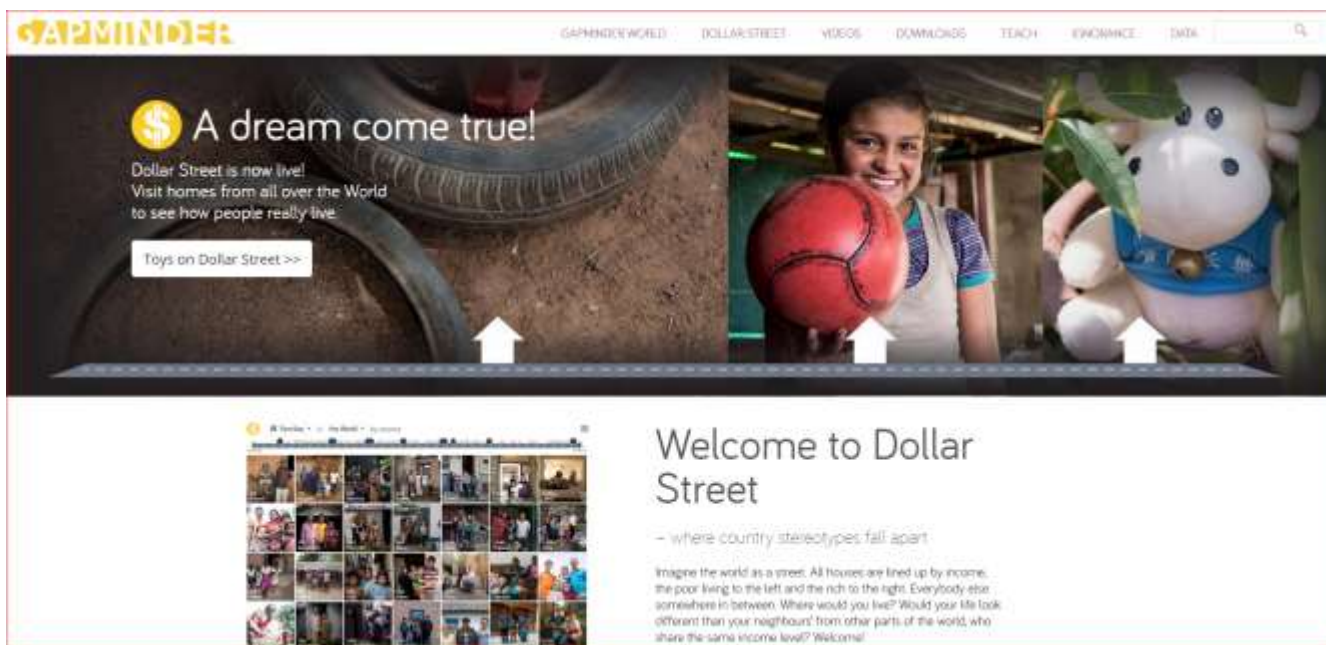


Paso 2: Navegar en Gapminder

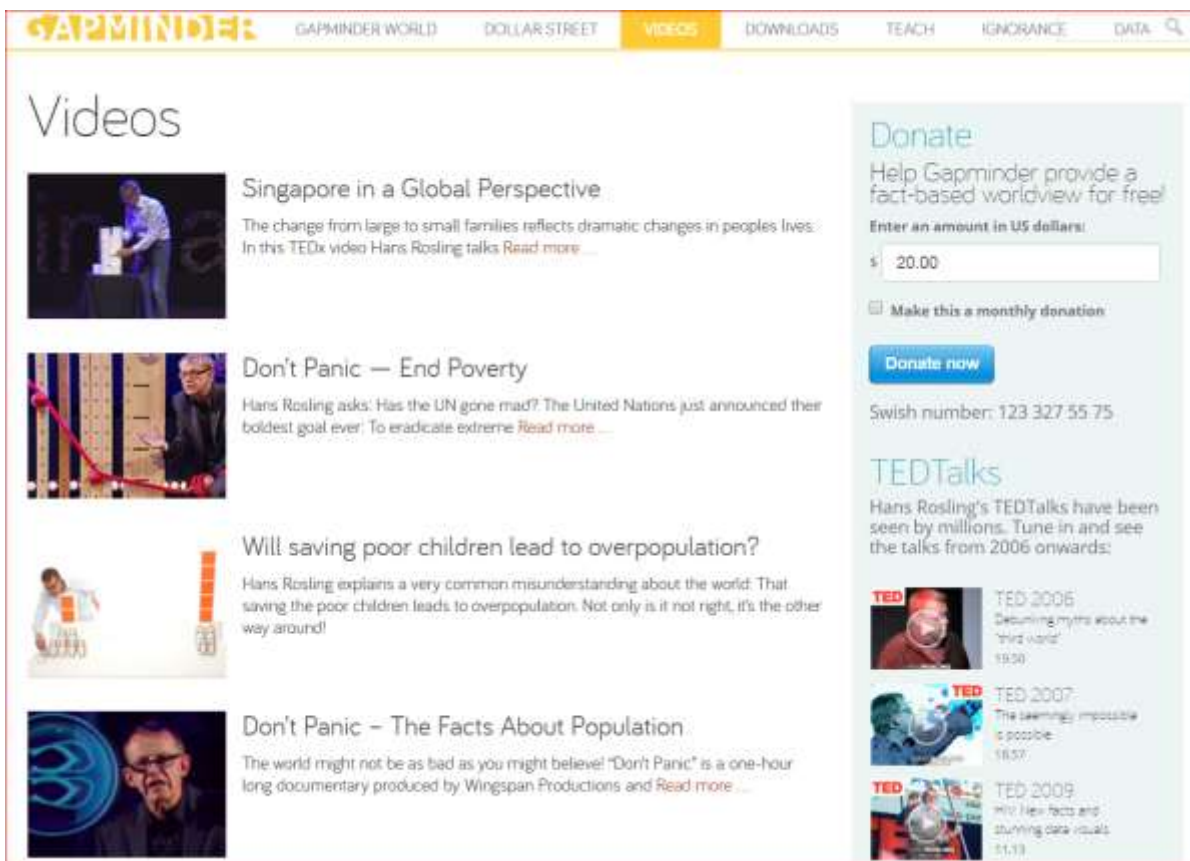
La misión de la Fundación Gapminder es combatir el desconocimiento por medio de datos estadísticos recopilados de diversas fuentes. Como se generan visualizaciones de datos de excelente calidad y presentaciones animadas de datos, las estadísticas pasan a ser interesantes.

En este paso estudiará los materiales gratuitos que se ofrecen en www.gapminder.org.

- a. Diríjase a www.gapminder.org.



- b. Haga clic en **Videos** (Vídeos) para ver presentaciones en las que se corporizan datos estadísticos y se realza su relevancia.



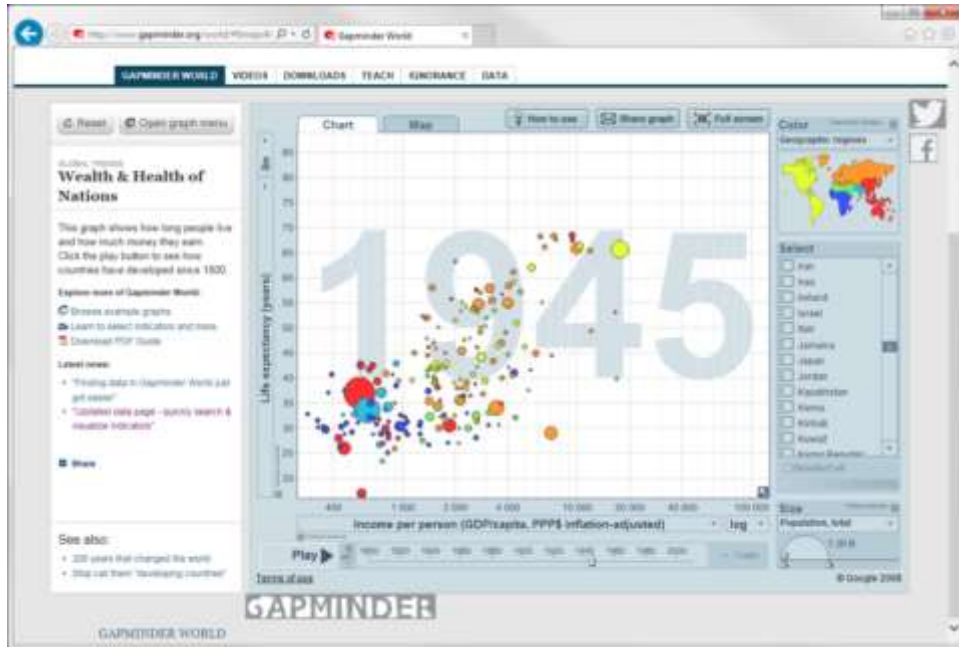
Busque vídeos que le interesen. Después de ver algunos de los vídeos, ¿de qué manera se modificó su forma de ver el mundo?

En la mayoría de los vídeos que vi, están los temas de estadística por ejemplo el ébola y de como se mostraron los datos de las personas que tienen esa enfermedad y como con el paso del tiempo se puede quitar o eliminar

- c. Para interactuar con los gráficos de burbuja que se presentaron en los vídeos, haga clic en **Gapminder World** (Mundo de Gapminder)



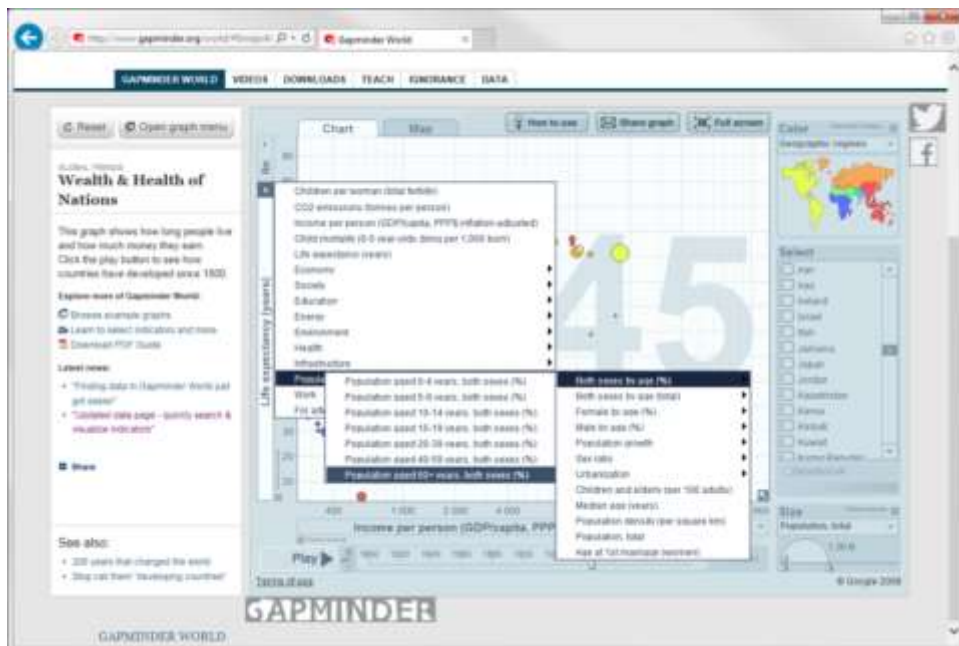
- d. Haga clic en **Play** (Reproducir). Observe los cambios del gráfico de dispersión correspondientes a la tendencia entre Life expectancy (years) (Expectativa de vida [años]) e Income per person (Ingreso por persona). Cada una de estas burbujas representa el progreso de un país.



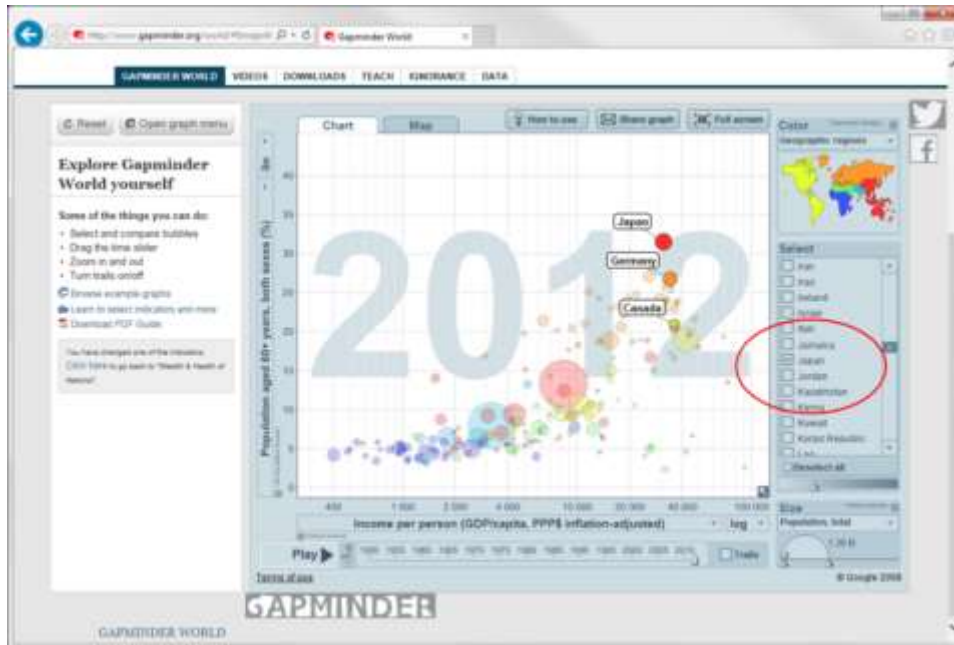
¿Qué relaciones observó?

Que la tendencia de vida en el 1945 era mas baja que en el 2013, ya es mas larga y puedes vivir mas de 60 años

- e. Cambie el eje y **Life Expectancy (years)** a **Population aged 60+ years, both sexes (%)** (Población de más de 60 años, ambos sexos [%]) para que en el cuadro se compare el porcentaje de la población mayor de 60 años con el ingreso por persona.



- f. Ahora seleccione algunos países para resaltar su progreso desde 1950 hasta 2012. En este ejemplo se han seleccionado Canadá, Alemania y Japón. Japón tendrá el porcentaje más elevado de su población mayor de 60 años en el año 2030, de acuerdo con la visualización de datos de <http://knoema.com/awppbnd/how-well-ageing-population-is-faring>.



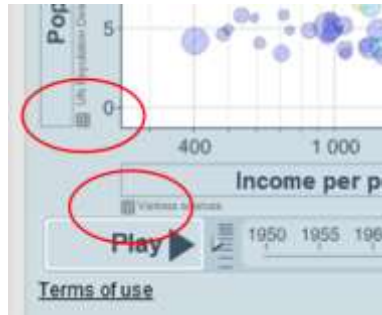
- g. Haga clic en **Play** para observar las líneas correspondientes a estos tres países.

¿Qué relaciones ve?

Que Japón es el país que con tendencia a vivir mas



- h. Puede consultar las fuentes de datos que se utilizaron en estos gráficos si hace clic en el símbolo de cuadrícula adyacente a los ejes x o y del gráfico. Haga clic en el símbolo de cuadrícula para ver o descargar los datos sin procesar.



Fuente: Material gratuito de www.gapminder.org

Reflexión

Acabamos de presentarle brevemente la visualización de datos ofrecida por diferentes sitios web. Elija el sitio web de visualización de datos que más le agrada y ejecute el proceso.

Describa brevemente la visualización que realizó.

La página que mas me gusto fue la knoema porque muestra las graficas de manera clara y facil de comprender y las fuentes que tienen me parece que son verídicas

¿Qué aprendió de la visualización?

Que como con solo ver graficas y visualizaciones puedes ver un progreso en los datos que manejes o también puedes ver como van bajando con el paso del tiempo y eso nos puede ayudar hacer un análisis para mejorar cada año

Práctica de laboratorio: ¿dónde están mis datos?

Objetivos

Analice la propiedad de los datos cuando estos no se almacenen en un sistema local.

Parte 1: Analizar la política de Condiciones del servicio

Parte 2: ¿Sabe a qué se suscribió?

Aspectos básicos/situación

Las redes sociales y el almacenamiento en línea se han convertido en una parte integral de la vida de muchas personas. Los archivos, las fotos y los videos se comparten entre amigos y familiares. La colaboración y las reuniones en línea se realizan en el lugar de trabajo con personas a muchos kilómetros de distancia entre sí. El almacenamiento de datos ya no se limita solo a los dispositivos a los que tiene acceso localmente. La ubicación geográfica de los dispositivos de almacenamiento ya no es un factor limitante para almacenar o generar respaldos en ubicaciones remotas.

En esta práctica de laboratorio, explorará los acuerdos legales requeridos para usar diversos medios sociales en línea y servicios de almacenamiento. También explorará algunas de las formas en las que puede proteger sus datos.

Recursos necesarios

- Computadora o dispositivo móvil con acceso a Internet

Parte 1: Explore los acuerdos de Condiciones del servicio

Si utiliza servicios en línea para almacenar datos o comunicarse con sus amigos o familiares, probablemente hizo un acuerdo con el proveedor. Las Condiciones del servicio, también conocidas como Condiciones de uso o Términos y condiciones, es un contrato legalmente vinculante que rige las reglas de la relación entre usted, su proveedor y otras personas que utilizan el servicio.

Diríjase al sitio web de un servicio en línea que utilice y busque el acuerdo de los Condiciones del servicio. A continuación se muestra una lista de muchas redes sociales y servicios de almacenamiento en línea conocidos.

Redes sociales

Facebook: <https://www.facebook.com/policies>

Instagram: <http://instagram.com/legal/terms/>

Twitter: <https://twitter.com/tos>

Pinterest: <https://about.pinterest.com/en/terms-service>

Almacenamiento en línea

iCloud: <https://www.apple.com/legal/internet-services/icloud/en/terms.html>

Dropbox: <https://www.dropbox.com/terms2014>

OneDrive: <http://windows.microsoft.com/en-us/windows/microsoft-services-agreement>

Revise los términos y responda las siguientes preguntas.

- ¿Tiene una cuenta con un proveedor de medios sociales en línea y/o de almacenamiento? Si es así, ¿ha leído el acuerdo de Condiciones del servicio?

Si, si cuento con algunas. No he leído el acuerdo de términos y condiciones, solo lo he leído así de manera muy rápida pero no todo completo solo muy pocas partes

- b. ¿Cuál es la política de uso de datos?

Por ejemplo Facebook usa mis datos para personalizar funciones y el contenido de sus apps como sugerir grupos, amigos, publicidad que me podría interesar.

- c. ¿Cuáles son las configuraciones de privacidad?

Nos permite administrar todo el contenido que podemos compartir de manera publica para poder habilitar o no que información pueden ver

- d. ¿Cuál es la política de seguridad?

Señala que es de exclusiva aplicación para ellos y no para aplicaciones o sitios web que utilizan su plataforma, también establece que no se crean cuentas a menores de 13 años en caso de que si se bloquean

- e. ¿Cuáles son sus derechos en relación con sus datos? ¿Puede solicitar una copia de sus datos?

Tengo el control de configurar quien puede ver mis datos públicos y si se puede pedir una copia de mis datos o hacer que otra persona los pueda tener

- f. ¿Qué puede hacer el proveedor con los datos que usted carga?

Los recopila para así poder darme mas anuncios o publicaciones que sean de mi interes, mas que nada para vender cosas

- g. ¿Qué sucede con sus datos cuando cierra su cuenta?

Todo se elimina no se podrá recuperar nada

Parte 2: ¿Sabe a qué se suscribió?

Luego de haber creado una cuenta y aceptado las Condiciones del servicio, ¿sabe realmente a qué se suscribió?

En la parte 2, explorará cómo los proveedores pueden interpretar y utilizar las Condiciones del servicio.

Utilice Internet para buscar información sobre cómo los acuerdos de Condiciones del servicio pueden interpretarse.

A continuación se encuentran algunos artículos de ejemplo para comenzar.

- Facebook:

<http://www.telegraph.co.uk/technology/social-media/9780565/Facebook-terms-and-conditions-why-you-dont-own-your-online-life.html>

- icloud:

http://www.americanbar.org/publications/law_practice_today_home/law_practice_today_archive/april12/have-attorneys-read-the-icloud-terms-and-conditions.html

- Dropbox:

<http://www.legalgenealogist.com/blog/2014/02/24/terms-of-use-change-dropbox/>

Revise los artículos y responda las siguientes preguntas.

- a. ¿Qué puede hacer para protegerse?

No dar mucha información a la plataforma y estar enterada de cada condicion que tiene la plataforma

b. ¿Qué puede hacer para proteger su cuenta y proteger sus datos?

No publicar información muy personal como por ejemplo, dirección de la casa, teléfonos, cuentas de bancos y administrar quien puedo o no ver mis datos, también tener contraseñas seguras y difíciles de adivinar

Práctica de laboratorio: Demostrar las limitaciones de una hoja de cálculo

Objetivos

Observe las limitaciones de una hoja de cálculo cuando se la utiliza para el análisis de datos.

Parte 1: Estudiar y descargar los datos

Parte 2: Determinar el puente raíz

Aspectos básicos/situación

En Internet dispone de enormes cantidades de datos sin restricciones que puede descargar y utilizar para realizar análisis. En esta práctica de laboratorio estudiará las limitaciones de las hojas de cálculo en el análisis de datos con un conjunto de datos sin restricciones proveniente de la ciudad de Seattle.

Recursos necesarios

- Dispositivo PC con acceso a Internet de alta velocidad
- Programa de hojas de cálculo; por ejemplo: Microsoft Excel, Google Sheet, LibreOffice Calc, Apple Numbers u OpenOffice Calc

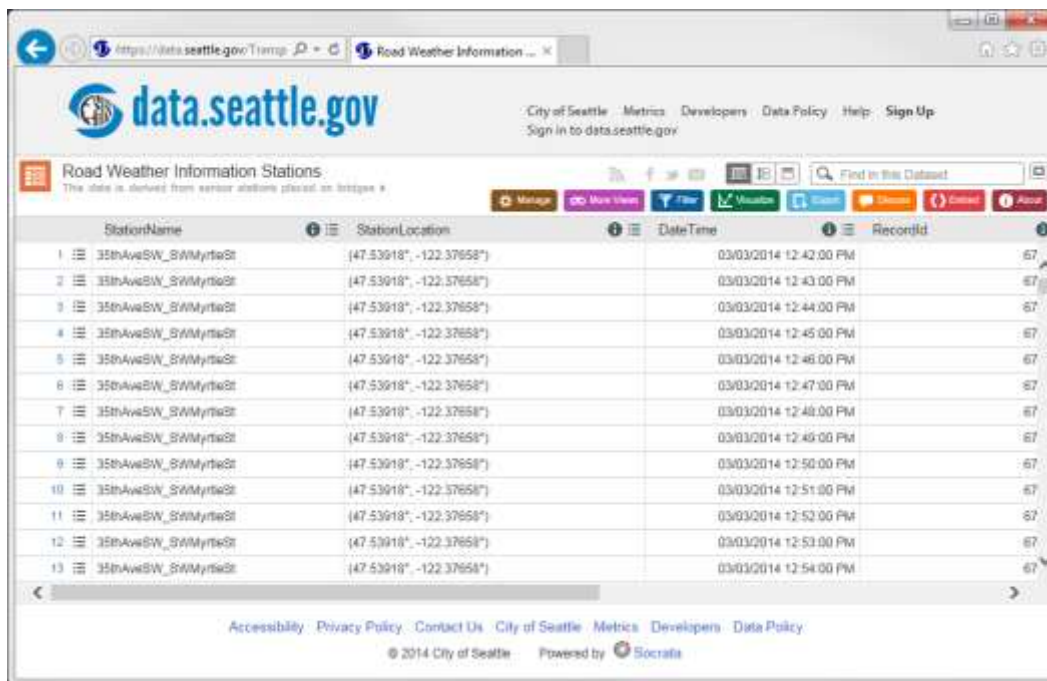
Parte 1: Estudiar y descargar los datos

En esta parte descargará el conjunto de datos de nombre *Road Weather Information Stations* (Estaciones de información climatológica vial) de <https://data.seattle.gov>. También estudiará algunas de las características de este conjunto de datos.

- Diríjase a <https://data.seattle.gov/Transportation/Road-Weather-Information-Stations/egc4-d24i?>.
- Haga clic en **View Data** (Ver datos).

Práctica de laboratorio: Demostrar las limitaciones de una hoja de cálculo en el análisis de datos

En este conjunto de datos se indican las temperaturas de los caminos y del aire obtenidas por sensores incrustados en puentes y en calles locales pertenecientes a la ciudad de Seattle. Los datos se actualizan cada quince minutos.



	StationName	StationLocation	DateTime	RecordId	
1	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:42:00 PM	67	
2	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:43:00 PM	67	
3	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:44:00 PM	67	
4	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:45:00 PM	67	
5	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:46:00 PM	67	
6	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:47:00 PM	67	
7	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:48:00 PM	67	
8	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:49:00 PM	67	
9	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:50:00 PM	67	
10	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:51:00 PM	67	
11	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:52:00 PM	67	
12	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:53:00 PM	67	
13	35thAveSW_SWMYrteSt	(47.53918°, -122.37658°)	03/03/2014 12:54:00 PM	67	

¿Cuántos campos (columnas) hay en el conjunto de datos? ¿Cuántos registros (filas)?

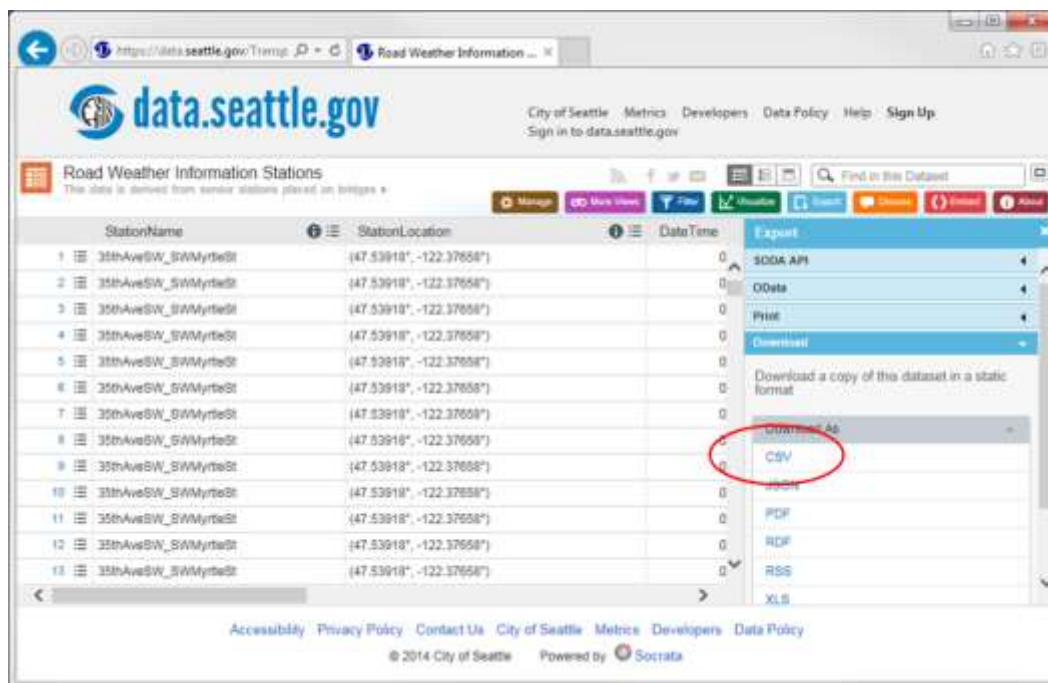
Son 6 columnas con 100 filas de 1,250,000

¿Cómo se llaman los campos?

Nombre de estación, ubicación de la estación, Fecha y hora, Id de registro, Carretera Superior Temperatura y Temperatura

- c. Haga clic en **Exportar**. Haga clic en **CSV** para descargar este conjunto de datos en formato CSV. La descarga demorará algunos minutos. Prosiga con la parte siguiente si no puede descargar los datos.

Nota: Si no pudo terminar de descargar el archivo *Road_Weather_Information_Stations.csv*, busque *Road_Weather_Information_Stations.csv.part* en su sistema. Lo más probable es que esté en su carpeta de Descargas. Cámbiele el nombre a *Road_Weather_Information_Stations.csv*.



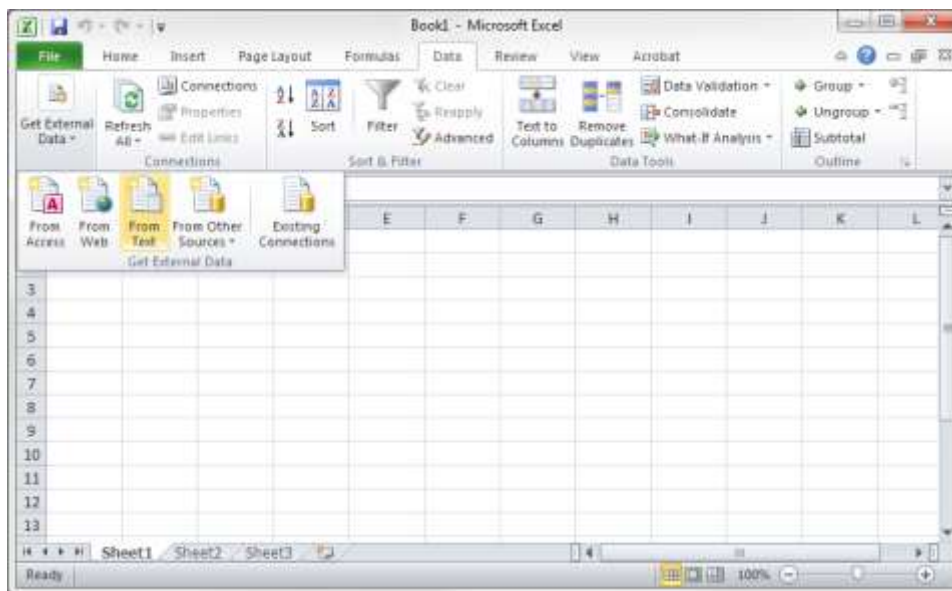
¿Qué tamaño tiene el archivo que se descargó? 110,048 KG

Parte 2: Estudiar las limitaciones de una hoja de cálculo

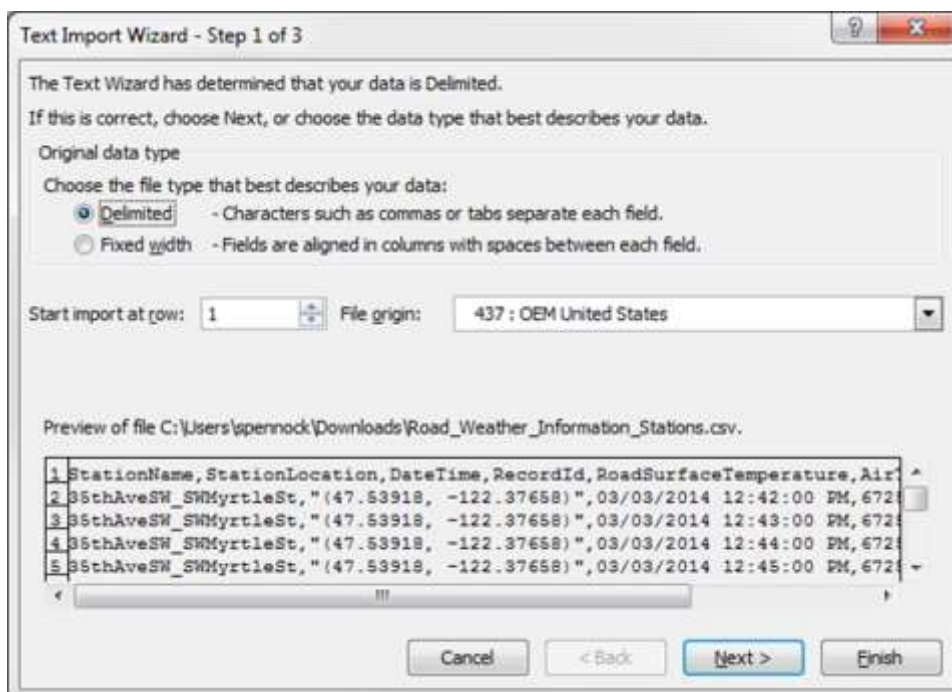
Intentará abrir el conjunto de datos que descargó en el programa de hojas de cálculo que prefiera. Si no pudo descargar el archivo, puede revisar los pasos y estudiar las limitaciones de las hojas de cálculo sin abrir el archivo.

- Abra su programa de hojas de cálculo. En este ejemplo se utiliza Microsoft Excel.

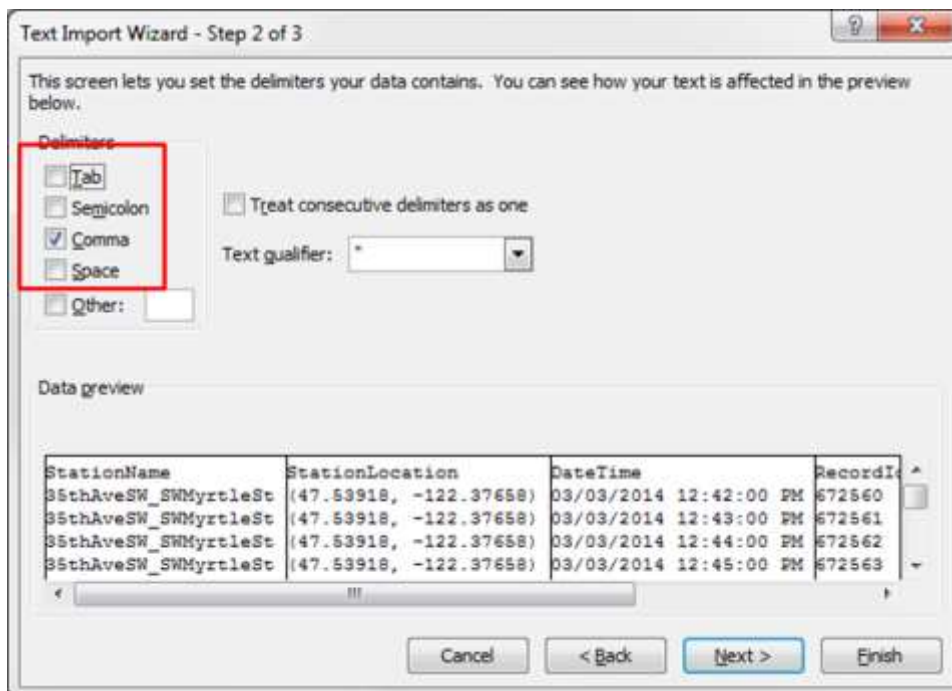
- b. Importe el archivo csv que descargó: *Road_Weather_Information_Stations.csv*. Haga clic en **Get External Data** (Obtener datos externos). Haga clic en **From Text** (Desde texto) y elija *Road_Weather_Information_Stations.csv*.



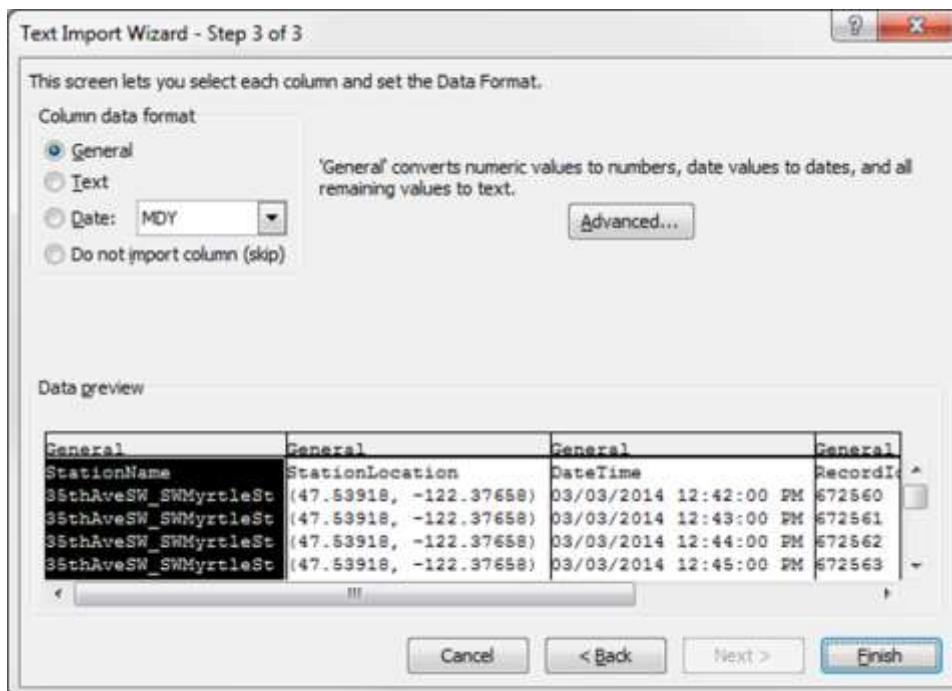
- c. Aparecerá el asistente para importar texto. En el paso 1 de 3, haga clic en **Next** (Siguiente) para continuar.



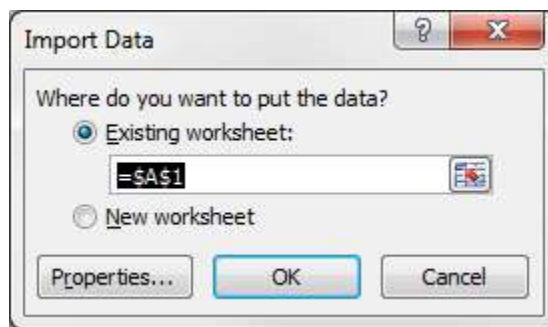
- d. En el paso 2 de 3, seleccione **Comma** (Coma) como el delimitador y desmarque **Tab** (Tabulación). Haga clic en **Siguiente** para continuar.



- e. En el paso 3 de 3, haga clic en **Finish** (Finalizar) para continuar.



- f. Haga clic en **OK** (Aceptar) para importar los datos.



¿Qué mensaje de advertencia ve en la pantalla?

El resultado de la consulta es demasiado grande para cargarse en la ubicación especificada en la hoja de calculo



¿De qué manera esta limitación le impide analizar los datos?

Que no me muestra todos los datos completos

- g. Abra una hoja de cálculo nueva. Trate de alcanzar la cantidad máxima de filas de la hoja de cálculo.
¿Cuál es la cantidad máxima de filas en la hoja de cálculo que eligió?

1.048.576

Compare la cantidad máxima de filas de la hoja de cálculo con la cantidad de registros presentes en el conjunto de datos Road Weather Information Stations de <https://data.seattle.gov>. En función de lo que observe, explique la limitación de la hoja de cálculo.

Se pasa por 200 filas mas para que Excel las pueda cargar

Reflexión

Si los programas de hojas de cálculo más usados no pueden manejar conjuntos de datos más grandes, ¿qué herramientas de análisis de datos pueden utilizarse? Busque algunas herramientas posibles en Internet.

Podría ser Power BI que es un software muy potente para analizar muchos datos de manera mas fácil
