

Guía de OpenSpending (Gasto Abierto)

Por el equipo de OpenSpending Data

Anders Pedersen, Rufus Pollock, Michael Bauer, Neil Ashton, Lisa Evans, Tony Hirst, Pierre Chrzanowski, Félix Ontañón, Oluseun Onigbinde

[Introducción](#)

[¿Qué es OpenSpending?](#)

[¿Qué tipos de datos financieros pueden entrar en OpenSpending?](#)

[Datos de gastos transaccionales](#)

[Datos presupuestarios](#)

[Cómo presenta los datos OpenSpending?](#)

[Conjuntos de datos](#)

[Modelos](#)

[Agregar datos a OpenSpending](#)

[Resumen](#)

[Recogiendo datos](#)

[Dándole formato a los data](#)

[CSV](#)

[El formato de OpenSpending](#)

[Publicando datos en la web](#)

[Google Drive](#)

[Gist](#)

[Creando un conjunto de datos en OpenSpending](#)

[Creando un nuevo conjunto de datos](#)

[Agregando una nueva fuente de datos](#)

[Modelando tus datos en OpenSpending](#)

[Dimensiones obligatorias: cantidad y tiempo](#)

[Las dimensiones claves \(key\) y compuestas \(compound\)](#)

[Medidas y otras dimensiones](#)

[Para terminar: guardar y cargar](#)

[Visualizaciones](#)

[Crear una Visualización](#)

[BubbleTree](#)

Introducción

¿Qué es OpenSpending?

OpenSpending es una comunidad que comparte datos y aplicaciones web con el fin de hacer un seguimiento a cada transacción financiera de gobiernos y empresas alrededor del mundo, y presentar los datos de forma útil y atractiva. OpenSpending es un proyecto abierto mantenido por una comunidad de colaboradores. Cualquier persona interesada en datos de gastos de cualquier tipo está invitada a contribuir a la base de datos de OpenSpending, crear visualizaciones utilizando software de OpenSpending y utilizar la IPA (Interfaz de Programación de Aplicaciones) de OpenSpending.

Este capítulo introduce a nuevos colaboradores de Open Spending a los conceptos centrales del sistema. Describe el tipo de datos financieros que OpenSpending soporta, y explica como OpenSpending representa los datos.

¿Qué tipos de datos financieros pueden entrar en OpenSpending?

OpenSpending es muy flexible en el tipo de datos financieros que soporta. Aunque el proyecto de OpenSpending tiene un fuerte enfoque en finanzas de gobiernos, esto no es un limitante técnico. OpenSpending soporta cualquier conjunto de datos que consista en un conjunto de transacciones, cada una asociada a una cantidad de dinero y tiempo.

La mayoría de los datos actualmente hospedados en OpenSpending pueden ser categorizados ya sea por transacción o datos presupuestarios. La principal diferencia entre estas categorías es el nivel de detalle. Datos transaccionales llevan el historial de transacciones individuales, mientras que los datos presupuestarios agrega las transacciones en categorías.

Datos de gasto transaccional

Datos transaccionales, o simplemente "datos de gastos", lleva el rastro de las transacciones financieras individuales. Cada pago de una entidad a otra en una fecha determinada y con un propósito específico (por ejemplo, un proyecto o servicio) aparece de forma individual. Datos de gastos transaccionales incluye varios tipos de registros, incluida la información sobre las subvenciones del gobierno, compromisos y los gastos reales.

La información en conjunto (por ejemplo, subtotales) no debe incluirse en los datos de transacciones. Los datos que han sido parcial o totalmente agregados exigen un modo diferente de análisis y deben ser tratados como datos presupuestarios en lugar de datos

transaccionales. Esto no significa, sin embargo, que varios pagos "físicos" que equivalen a una sola transacción "lógica" no pueden ser representados por una única transacción en los datos transaccionales.

Datos transaccionales en OpenSpending incluye:

- [Vendedores y Contratistas del Distrito de Columbia](#)
- [Agencia de Desarrollo Austríaca](#)

Otro tipo de datos relacionados tienen que ver con procedimientos de contratación pública. Los datos de Contratación Pública son los datos sobre licitaciones públicas: qué se licitó, por cuánto dinero y quién ganó la licitación. Puede considerarse un subconjunto de los datos transaccionales.

Datos de licitaciones en OpenSpending incluye:

- [Marchés publics au Sénégal](#)
- [Marchés publics France 2011](#)

Datos Presupuestarios

En los datos presupuestarios, los gastos y los ingresos se agrupan en categorías. El objetivo de esta formación es ayudar al lector a comprender el presupuesto, que suele ser un documento de política que se utiliza para ofrecer a los lectores una visión general de las opciones financieras más importantes del gobierno. La asignación suele estar estructurada por un esquema de clasificación y no por los beneficiarios reales de los fondos.

Los datos presupuestarios a menudo presentan conjuntamente datos sobre los resultados y créditos pasados para un período futuro. En esta presentación, los importes gastados en años anteriores en un sector en particular se utilizan para informar cuánto se debe asignar durante el período presupuestario que viene. La información presupuestaria se basa a menudo en datos agregados y en las estimaciones estadísticas.

Las diferentes regiones hacen poner a disponibilidad diferentes tipos de información presupuestaria, incluyendo: Declaraciones Pre-Presupuesto; Propuestas Presupuestarias Ejecutivas; Presupuesto Estatal Promulgado; y presupuestos de los Ciudadanos (versiones simplificadas del presupuesto para el beneficio de los ciudadanos).

Los datos presupuestarios sobre OpenSpending incluye:

- [Presupuesto de Berlin](#)
- [Gasto presupuestario de Sevilla](#)

¿Cómo representa los datos OpenSpending?

OpenSpending mantiene una colección de conjuntos de datos, cada uno de los cuales representa un conjunto de datos de una fuente separada. Dentro de cada conjunto de

datos, las transacciones individuales están representados por un conjunto de entradas. Cada conjunto de datos tiene su propio modelo que correlaciona la estructura de los datos. El modelo codifica las propiedades de cada conjunto de datos de entrada en términos de *dimensiones*.

Conjuntos de Datos

La unidad básica en el sistema OpenSpending es el conjunto de datos. Las transacciones financieras que comparten un tema común (por ejemplo, el gasto de una ciudad en particular de un presupuesto para un año en particular) se agrupan y almacenan como un conjunto de datos. Un conjunto de datos es una colección de "entradas", y cada entrada representa una única transacción asociada con una cantidad de dinero y tiempo.

Los conjuntos de datos también incluyen metadatos para caracterizar su contenido. Los metadatos incluyen una descripción del conjunto de datos, la información sobre el origen de los datos, y otra información que ayuda a los usuarios a encontrar el conjunto de datos e interpretar su contenido.

Modelos

La estructura de cada conjunto de datos depende completamente del creador del conjunto de datos. Esta estructura se crea mediante la especificación de un modelo, que proporciona las dimensiones a lo largo de las entradas que pueden diferir unos de otros.

Un modelo consiste en un conjunto de dimensiones. Una dimensión es potencialmente una propiedad que diferencia a una entrada de otra. Si usted se imagina un conjunto de datos como una hoja de cálculo, cada dimensión puede considerarse como una columna. Sin embargo, las dimensiones pueden tener más estructura que una columna ordinaria hoja de cálculo.

Las dimensiones vienen en varios tipos. El más importante es el tipo de medida. Las medidas son dimensiones que puede contener un único valor numérico. Otro tipo de dimensión importante es el tipo de tiempo, que representan fechas y horas. Todos los datos tienen por lo menos una de las dimensiones de cada medida y el tiempo, lo que representa, respectivamente, la cantidad de dinero representado por la transacción y el momento en que tuvo lugar.

Los tipos de dimensiones restantes se utilizan para representar otras propiedades que las entradas pueden tener, por ejemplo, números de transacción, las etiquetas de un sistema de clasificación, o los nombres de las personas o empresas involucradas. Estas dimensiones incluyen *atributos*, que puede contener un único valor, y las *dimensiones compuestas*, que puede contener un conjunto jerarquizado de valores. Las dimensiones compuestas son útiles cuando la propiedad incluye varios sub-propiedades que podrían ser utilizados para cada conjunto de datos.

Agregando Datos a OpenSpending

Resumen

Una de las contribuciones más valiosas que puede realizar en el proyecto OpenSpending es añadir un nuevo conjunto de datos. Esta sección de la guía le guiará por el proceso de agregar nuevos datos.

Un flujo de trabajo típico para la importación de un conjunto de datos en OpenSpending implica los siguientes pasos:

- Recopilar datos legibles por máquina a partir de una fuente de confianza.
- Convertir los datos a un archivo CSV en el formato esperado por OpenSpending, limpiándolo para eliminar inconsistencias y errores.
- Publicar los datos en la web.
- Crear un conjunto de datos, agregar los datos publicados como una nueva fuente de datos.
- Modelar el conjunto de datos para asignar un papel lógico a cada columna de la tabla de origen.
- Cargar los datos, o refinar los datos basados en la retroalimentación dada por la plataforma sobre la consistencia de los datos.

Cada uno de estos pasos serán explicados en detalle en las siguientes secciones.

Recopilando datos

Para agregar un conjunto de datos a OpenSpending, primero debe tener datos. Si ya los tiene, puede proceder. Sino, necesita conseguirlos.

Comience su búsqueda de datos mediante la consulta a recursos digitales tales como la Escuela de Datos ([School of Data](#)) y el Manual de Periodismo de Datos ([Data Journalism Handbook](#)). También puede obtener ideas sobre cómo ir sobre la búsqueda visitando el grupo OpenSpending ([OpenSpending group](#)) en datahub.io, y usted puede hacer preguntas en el canal de IRC # openspending en Freenode.

La información que usted encuentre en el mejor de los casos estará en un formato "legible" por máquina, por ejemplo en forma de una hoja de cálculo Excel o un archivo CSV. Si usted encuentra los datos en un formato como PDF o un documento de Word, será muy difícil de trabajar, y usted podría considerar simplemente utilizar diferentes datos.

Formateando los datos

OpenSpending espera que todos los datos estén en un formato sencillo.

CSV

OpenSpending acepta datos en un solo formato de archivo, el archivo de Valores Separados por Comas (CSV, por sus siglas en inglés). Un CSV es un archivo de texto simple que representa los datos en forma de tabla, similar a una hoja de cálculo. En una tabla, cada punto de datos está representado por una fila, y las propiedades de cada punto de datos están representados por una columna. Los archivos CSV codifican las tablas dando a cada fila una línea en el archivo de texto y, separando las columnas con comas.

CSV aceptadas por OpenSpending están *sin normalizar*, lo que significa que no ahorran espacio mediante la eliminación de los valores redundantes. Los CSV de OpenSpending también son rectangulares, teniendo exactamente el mismo número de columnas en cada fila.

El formato de OpenSpending

Los CSV para OpenSpending deben tener las siguientes propiedades:

1. Una fila de encabezado. La primera fila del archivo CSV debe contener los nombres de las columnas, separados por comas. Todas las demás filas se tratan como filas de datos.
2. Al menos tres columnas. El mínimo de las columnas son: una cantidad, una fecha (que podría ser sólo un año), y un inversionista o un recipiente (que sólo podría ser el nombre de una cuenta).
3. Columnas consistentes. Cada columna debe representar consistentemente un solo tipo de valor para todas las filas. (No puede haber ninguna fila sub-cabecera, por ejemplo.)
4. Las filas son los puntos de datos. Las filas deben contener un solo tipo de información: una transacción o una línea. Una fila debe representar un máximo de un período de tiempo. (Las filas no pueden representar múltiples transacciones.)
5. No debe haber filas o celdas en blanco. Cada fila de un archivo de datos importado debe contener toda la información necesaria para construir el elemento resultante.
6. No debe haber totales pre-agregados (por ejemplo, subtotales o auto sumas). OpenSpending hará las matemáticas y calculará los totales automáticamente.
7. Un identificador único. Debe haber una columna (o una combinación de columnas) cuyos valores identifican de forma exclusiva cada fila. La manera más fácil de crear dicho identificador es agregar una columna ficticia para el conjunto de datos en la que se pone un número que se incrementa en cada fila. Usted puede hacer esto en Excel, escriba los números en las dos primeras filas, la selección de ambas células y arrastrando la esquina inferior derecha de la celda para extender la serie.
8. Fechas en el formato correcto. Las fechas deben estar en el formato "aaaa-mm-dd".
9. Cifras en el formato correcto. Las cifras deben contener únicamente dígitos y un periodo adicional —sin comas! (Cifras legibles como "12,345.67" deben convertirse a

cifras como "12345.67".)

La comunidad OpenSpending ha recopilado algunos ejemplos de hojas de cálculo ([example spreadsheets](#)) con el fin de ilustrar "buenos" y "malos" datos tabulares. Estos son algunos ejemplos de las hojas de cálculo con formato incorrecto:

- [Muchas celdas en blanco](#) (probablemente se omitió información redundante)
- [Múltiples transacciones en una fila](#) (varios años en una fila)
- [Cifras equivocadas](#) (las cifras tienen comas para ser legibles por máquina)

Este es un ejemplo de una buena hoja de cálculo:

- [Washington, DC](#)

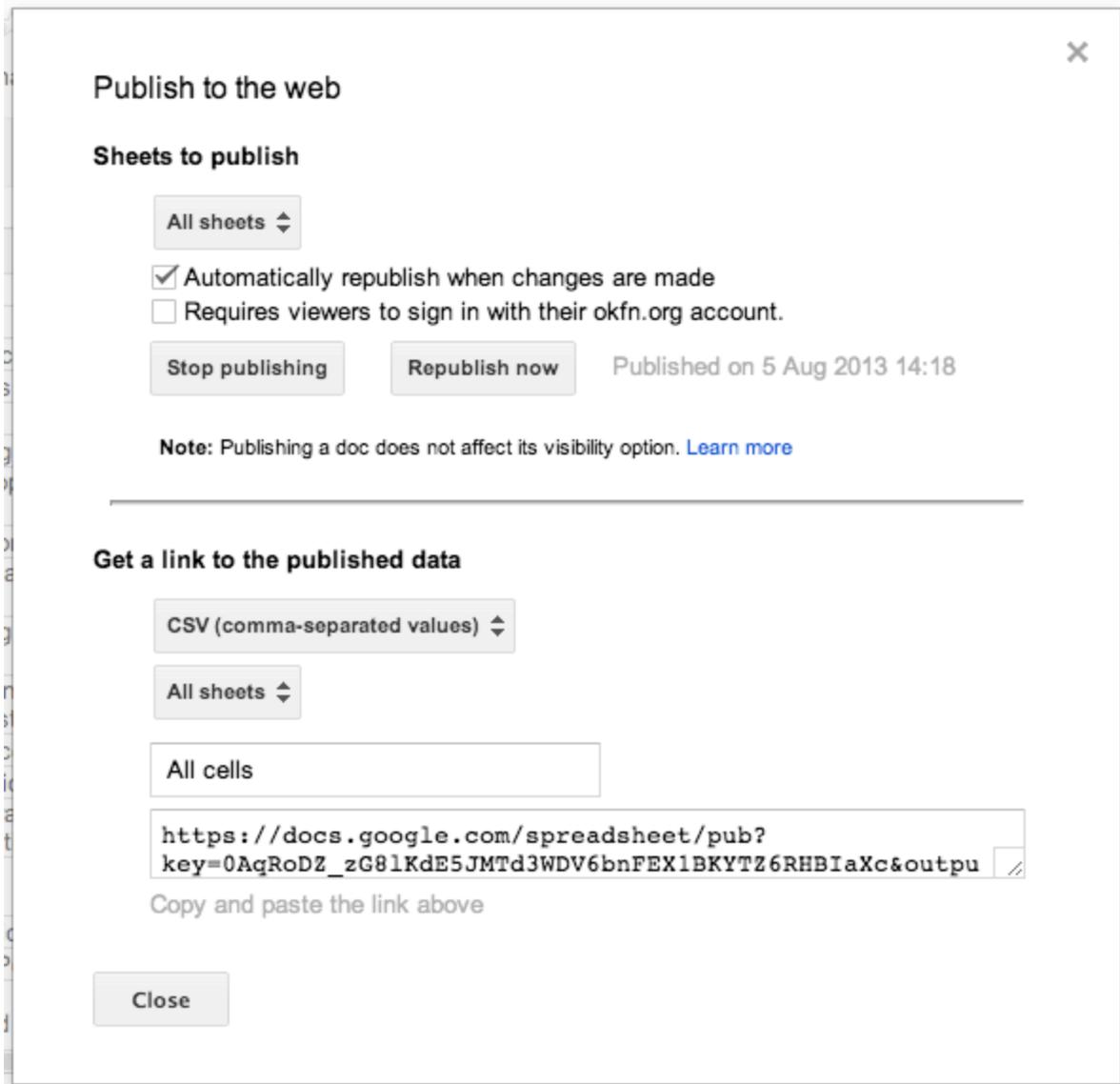
Publicando datos en la web

Los datos no pueden (aún) subirse directamente a OpenSpending. Con el fin de ser añadido a la base de datos OpenSpending, los datos primero deben ser accesibles desde la web. Esta sección presenta dos formas convenientes para publicar conjuntos de datos en línea.

Google Drive

Usted puede hacer sus datos accesibles en la web, convirtiéndola en una hoja de cálculo de Google Drive.

1. Importe sus datos. Crear una nueva hoja de cálculo de Google Drive, a continuación, seleccione *Importar ...* en el menú *Archivo*. Seleccione *Reemplazar hoja de cálculo*, haga clic en **Seleccionar archivo** y navegue hasta su archivo CSV.
2. Asegúrese de que Google Docs no cambie el formato de fechas de sus datos. Seleccione la columna que contiene fechas. Haga clic en el menú *Formato* y seleccione *Número -> Más formatos -> 2008-09-26*. Sus fechas deben aparecer en el formato **aaaa-mm-dd** prescrito.
3. Haga clic en el menú *Archivo* y seleccione *Publicar en la web* En el cuadro que aparece, haga clic en **Iniciar publicación**. Debajo *Obtener un enlace a los datos publicados*, seleccione **CSV (valores separados por comas)**.



La URL en la parte inferior de la caja ahora apunta a sus datos.

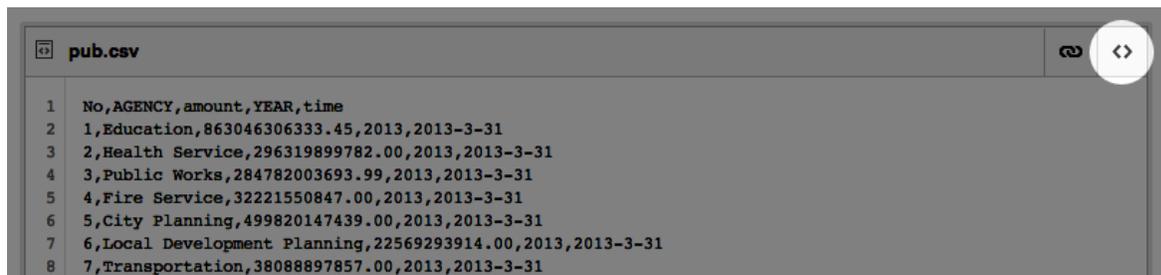
Gist

GitHub Gist es una manera conveniente de organizar pequeñas cantidades de texto, como archivos CSV.

GitHub Gist is a convenient way to host small quantities of text, including CSV files.

1. Ingresar a GitHub (o regístrate si aún no lo ha hecho), a continuación, vaya a gist.github.com.

2. Haga clic y arrastre el archivo CSV desde el administrador de archivos de su sistema operativo en la página Gist GitHub de su navegador. Aparecerá el nombre y el contenido del archivo.
3. Haga clic en **Crear Gist Público** para ser llevado a la página de inicio de su nuevo "gist". La dirección URL original de sus datos será accesible a través del botón de "escuadras" en la esquina superior derecha del archivo.



```
pub.csv
1 No, AGENCY, amount, YEAR, time
2 1, Education, 863046306333.45, 2013, 2013-3-31
3 2, Health Service, 296319899782.00, 2013, 2013-3-31
4 3, Public Works, 284782003693.99, 2013, 2013-3-31
5 4, Fire Service, 32221550847.00, 2013, 2013-3-31
6 5, City Planning, 499820147439.00, 2013, 2013-3-31
7 6, Local Development Planning, 22569293914.00, 2013, 2013-3-31
8 7, Transportation, 38088897857.00, 2013, 2013-3-31
```

Creando un conjunto de datos en OpenSpending

Para empezar a compartir datos en la plataforma de OpenSpending, regístrese en OpenSpending y cree un nuevo conjunto de datos OpenSpending. Para crear un conjunto de datos, simplemente llene algunos de los metadatos que caracterizan sus datos y proporcione el URL donde sus datos están hospedados.

Creando un nuevo conjunto de datos

Ingresa a OpenSpending.org con su información de usuario, o regístrese si todavía no lo ha hecho. La primera pantalla a la que llegará es el Tablero (Dashboard), donde verá un botón azul con el nombre **Importar un Conjunto de Datos (Import a Dataset)**. Haga clic para crear un nuevo conjunto de datos en OpenSpending.

La siguiente pantalla le pide que proporcione metadatos que caracteriza sus datos. Esto incluye los siguientes campos:

- *Título*: un nombre descriptivo y significativo para el conjunto de datos. Puede ser cualquier frase.
- *Identificador*: un título más corto, que se utiliza como parte de la dirección URL del conjunto de datos. Sólo puede contener caracteres alfanuméricos, guiones y guiones bajos - sin espacios en blanco o puntuación.
- *Categoría*: una de "Presupuesto", "Gastos", y "Otros". Vea la sección de guía sobre los tipos de datos financieros para obtener más información sobre estas categorías.
- *Moneda*: la moneda en la que el gasto descrita por el conjunto de datos se lleva a cabo.
- *Países*: una lista de los países mencionados en el conjunto de datos. La elección de los países se ve limitado por una lista de países válidos.
- *Idiomas*: una lista de los idiomas utilizados en el conjunto de datos. Elección de

idiomas está limitado por una lista de idiomas válidos.

- *Descripción*: una caracterización del conjunto de datos en prosa sencilla. Puede ser cualquier frase.

Rellene todos los campos. Asegúrese de incluir una descripción que explica el origen de su conjunto de datos y reconoce los cambios que ha introducido (por ejemplo, la limpieza que han hecho).

Una vez que todos los metadatos ha sido rellenado, pulse **Siguiente Paso** para continuar.

Agregando una nueva fuente de datos

Al hacer clic a través del siguiente paso crea el nuevo conjunto de datos OpenSpending y usted llegará a la página *Administrar (Manage)*. La página de administración se utiliza para añadir fuentes de datos. También se utiliza para proporcionar información de esquema que permite OpenSpending para interpretar los datos, un proceso llamado "modelado", que se tratarán en la siguiente sección de la guía.

Para agregar un origen de datos a un conjunto de datos, haga clic en **Agregar una fuente**. Aparecerá un mensaje que le solicita una URL. Proporcione la dirección URL del archivo CSV que ha publicado en la web en el apartado anterior de la guía y haga clic en **Crear**. Verá un cuadro de texto en azul indica que OpenSpending está pensando en sus datos.

The screenshot shows a light blue notification box at the top with the text: "Let me see... The data is now being checked to see if it can be imported. This should not take very long." Below this text is a "Refresh" button. Underneath the notification is a section titled "Source files" which contains a table with two columns: "Link" and "Operations". The table has one row with the link "https://docs.go .. gle=true&gid=1&output=csv" and two buttons: "Test a sample" and "Load". Below the table, there is an "Analysis:" section with a blue button that says "STILL ANALYSING" and the text "Try refreshing in a moment."

Haga clic en **Actualizar** o simplemente utilizar el botón de actualización del navegador. Si en OpenSpending lograron analizar sus datos, debería ver un cuadro de texto verde que indica que los datos están listos. También debe ver una lista correcta de las columnas de la CSV.

Good stuff! Your data is now ready. Next, create a **data model** to structure the dataset.

Create a model

Source files

Link	Operations
OK https://docs.go .. gle=true&gid=1&output=csv 12 columns	Test a sample Load
Analysis:	We found these columns in the CSV file: <ul style="list-style-type: none">• Head-account• Head-Account Description• Sub-account• Sub-account Description• Year• Reporting Type• Amount• Revenue/Expenditure• Recurrent/Investment• Council Name• Expense Type• Identifier

Tenga en cuenta que si usted proporciona incorrectamente a OpenSpending con un archivo HTML en lugar de un archivo CSV válido, no reclamará, sino que simplemente intentará analizar el código HTML como si fuera un archivo CSV. El resultado tiene el siguiente aspecto.

Good stuff! Your data is now ready. Next, create a **data model** to structure the dataset.

Create a model

Source files

Link	Operations
OK http://openspending.org/help/data-loading.html 1 columns	Test a sample Load
Analysis:	We found these columns in the CSV file: <ul style="list-style-type: none">• <!DOCTYPE html>

Si agregó una fuente de datos mal, no se preocupe. Usted no tiene que utilizar la fuente de

su último conjunto de datos: OpenSpending obliga a hacer más trabajo en un origen de datos antes de que se pueda publicar. Sólo tiene que añadir una nueva fuente correcta y olvidarse de la mala.

Modelando sus datos en OpenSpending

Para cargar datos en OpenSpending, debe construir un *modelo* de sus datos. Un modelo especifica cómo sus datos se traducen en términos que OpenSpending entiende. OpenSpending representa las propiedades de datos en términos de *dimensiones*. El modelado de datos consiste en enumerar las dimensiones que le gustaría que el conjunto de datos tenga en OpenSpending y especifique cómo se relacionan con las columnas de los datos de origen.

Dimensiones obligatorias: cantidad y tiempo

Cada modelo tiene que tener al menos dos dimensiones: una cantidad y un tiempo. Estas especifican el tamaño de la operación y el momento en que se efectuó la transacción. La cantidad y el tiempo están asociados con tipos especiales de dimensiones. Una cantidad está representada por una dimensión de *medida*, y un tiempo está representado por una *fecha*. Las dimensiones genéricas no pueden representar estos valores especiales.

Al modelar los datos, no es una mala idea comenzar con las dimensiones obligatorias. Para empezar, haga clic en la pestaña **Dimensiones y Medidas** en el conjunto de datos de la página **Administración de datos**.

Cameroon Local Council Budget / Manage the dataset

Cameroon Local Council Budget

Dashboard Dimensions & Measures Views Team Meta Templates

Source files

Link	Operations
OK https://docs.go..gle=true&gid=1&output=csv 12 columns	<input type="button" value="Test a sample"/> <input type="button" value="Load"/>

A continuación, haga clic en **Agregar Dimensión** para abrir el panel *Agregar nueva dimensión*. Haga clic en el botón de selección de *Fecha*. Verá el cuadro *Nombre* llenado automáticamente con "tiempo", como se muestra a continuación. Haga clic en el botón verde **Agregar**.

Add new dimension

Name:

The name you choose should only contain letters, numbers and dashes (i.e., no spaces)

Where the dimension represents the ultimate source or destination of the funds, call it **from** or **to**.

Type:

- Dimension
A defined set, like suppliers or classifications, with multiple attributes (e.g. supplier name and VAT number).
- Attribute
A simple value, e.g. descriptions or transaction IDs.
- Date
Temporal value, e.g. start of project, date of disbursement.
- Measure
Monetary information, e.g. amount disbursed or allocated.

Unmapped source data columns

- Amount
- Council Name
- Expense Type
- Head-Account Description
- Head-account
- Identifier
- Recurrent/Investment
- Reporting Type
- Revenue/Expenditure
- Sub-account
- Sub-account Description
- Year

La siguiente pantalla que aparece le dará algo de información sobre el significado del tiempo. En el cuadro desplegable junto a *Columna:*, seleccione la columna de los datos que representa el valor del tiempo.

time (date)

The time dimension represents the time or period over which the spending occurred. Please choose the column of your dataset which contains an ISO8601 formatted date (YYYY, YYYY-MM, YYYY-MM-DD, etc.).

Label:

A human-readable title for this dimension.

Include in unique key

Make this dimension part of the set of uniquely identifying values for each column.

Column:

Source column in the source CSV file.

Una vez que identifica la columna de tiempo, haga clic en **Agregar Dimensión** de nuevo para añadir la cantidad. Esta vez, seleccione el botón etiquetado *Medida*, que se completará automáticamente el nombre de "cantidad", y haga clic en **Agregar**. Seleccione la columna que representa el valor de la transacción en el cuadro desplegable junto a *Columna*.

Las dimensiones claves (key) y compuestas

Sólo una dimensión adicional es necesaria para hacer el modelo suficiente: la dimensión (o conjunto de dimensiones) cuyo valor identifica de manera única cada punto de datos, la *clave*.

Un punto de datos no necesita ser identificado por el valor de una sola columna. Puede ser identificado por la combinación de varias en una *dimensión compuesta*. Dado que las claves *pueden* ser compuestas, el tipo de dimensión compuesta *debe* ser utilizada para que los represente, aunque su clave particular, no sea compuesta.

Para añadir la dimensión clave, haga clic en **Agregar Dimensión** y seleccione el botón de **Dimensión**. Introduzca un nombre para la clave, como "clave", en el cuadro *Nombre*. Haga clic en **Agregar**. Marque la casilla *Incluir en la clave única* para identificar esta dimensión como parte de la clave.

A continuación, eche un vistazo a la lista de **Campos**, que contiene dos filas etiquetados *nombre* y *etiqueta*. Una dimensión compuesta puede contener un número arbitrario de *campos*, cada uno de los cuales tiene un nombre y un tipo y cada uno de los cuales puede ser asociado con una columna en sus datos. Este es el sentido en que estas dimensiones son "compuestas": se agrupan varias columnas de los datos de origen en una única propiedad del conjunto de datos de destino.

key (compound)

Label:

A human-readable title for this dimension.

Include in unique key

Make this dimension part of the set of uniquely identifying values for each column.

Description:

Use as facet in entries browser

Select a few dimensions which are useful to slice the dataset by.

Field	Column	Type	Default
name	<input type="text" value="Identifier"/>	<input type="text" value="id"/>	<input type="text" value="Default Value"/>
label	<input type="text" value="Identifier"/>	<input type="text" value="float"/>	<input type="text" value="Default Value"/>

Una dimensión compuesta requiere al menos dos campos: *nombre* y *etiqueta*, que debe ser, respectivamente, de tipo *identificación* y *cadena*. El nombre de la dimensión se utiliza

para dotarla de una URL, y la etiqueta se utiliza para presentarlo en la interfaz de usuario.

Para crear una mínima dimensión compuesta, simplemente asocie la misma columna del origen de datos con el *nombre* y *etiqueta*. Seleccione la columna apropiada para cada uno y deje sin cambios los tipos predeterminados.

Medidas y otras dimensiones

Con una cantidad, tiempo, y clave, el modelo es suficientemente rico. Un modelo muy completo, sin embargo, incluirá las dimensiones de cada propiedad significativa de los datos de origen. Después de ciertas convenciones hace que esto sea más conveniente.

Un patrón común en las fuentes de datos es difundir la información que identifica a las entidades - grupos, cuentas, etc - a través de múltiples columnas. Información acerca de una cuenta asociada con una transacción se puede dividir en una columna de "cuenta" con un número de identificación y una columna "Descripción de cuenta" con una descripción verbal, por ejemplo. La imagen de abajo muestra este patrón en "Cuenta Principal" y "Sub-cuenta".

	A	B	C	D
1	Head-account <input type="checkbox"/>	Head-Account Description <input type="checkbox"/>	Sub-account <input type="checkbox"/>	Sub-account Description <input type="checkbox"/>
2	110.xxx	Uncategorized	110.101	Reserves for overheads
3	110.xxx	Uncategorized	110.101	Reserves for overheads
4	110.xxx	Uncategorized	110.101	Reserves for overheads
5	110.xxx	Uncategorized	110.101	Reserves for overheads
6	710.xxx	Fiscal revenue (1)	710.1	Proceeds from global taxes
7	710.xxx	Fiscal revenue (1)	710.1	Proceeds from global taxes
8	710.xxx	Fiscal revenue (1)	710.1	Proceeds from global taxes
9	710.xxx	Fiscal revenue (1)	710.1	Proceeds from global taxes
10	710.xxx	Fiscal revenue (1)	710.101	Proceeds from business
11	710.xxx	Fiscal revenue (1)	710.101	Proceeds from business
12	710.xxx	Fiscal revenue (1)	710.101	Proceeds from business

Las dimensiones compuestas de OpenSpending están diseñadas para modelar este tipo de información dispersa. Para ello, agregue una nueva dimensión compuesta y asocie cada

columna a uno de los campos de la dimensión. Intente emparejar una columna legible a *etiqueta* y una columna más concisa a *nombre*. En la imagen de abajo, "Cuenta Principal" se corresponde con *nombre* y la "Descripción de la cuenta principal" a *etiqueta*.

head-account (compound)

Label:

A human-readable title for this dimension.

Include in unique key

Make this dimension part of the set of uniquely identifying values for each column.

Description:

Use as facet in entries browser

Select a few dimensions which are useful to slice the dataset by.

Field	Column	Type	Default
name	<input type="text" value="Head-account"/>	<input type="text" value="id"/>	<input type="text" value="Default Value"/>
label	<input type="text" value="Head-Account Description"/>	<input type="text" value="string"/>	<input type="text" value="Default Value"/>

Algunas columnas de sus datos son más autónomos, representando atributos particulares de cada punto de datos. Una columna que organiza cada tipo de transacción en alguna categoría, por ejemplo, es de este tipo. En la imagen de abajo, las columnas Tipo de informes, Ingresos/Gastos, Recurrente/Inversión son así.

E	F	G	H	I
Year	Reporting Type	Amount	Revenue/Expenditu	Recurrent/Investme
2008	Budget	2000000	REVENUE	RECURRENT
2008	Actual	0	REVENUE	RECURRENT
2009	Budget	0	REVENUE	RECURRENT
2009	Actual	175000	REVENUE	RECURRENT
2008	Budget	6847200	REVENUE	RECURRENT
2008	Actual	3127600	REVENUE	RECURRENT
2009	Budget	6847200	REVENUE	RECURRENT
2009	Actual	6293198	REVENUE	RECURRENT
2008	Budget	4166100	REVENUE	RECURRENT
2008	Actual	1528655	REVENUE	RECURRENT

Columnas autónomas que especifican atributos o categorías se modelan mejor con unas dimensiones de *atributos*. Un atributo es esencialmente una dimensión con un solo campo, que puede tener cualquier tipo. Para crear un atributo, sólo tiene que seleccionar el botón *Atributo* al añadir una dimensión.

reporting-type (attribute)

Label:

A human-readable title for this dimension.

Include in unique key

Make this dimension part of the set of uniquely identifying values for each column.

Column:

Source column in the source CSV file.

Default value:

Empty cells will be padded with this value.

Data type:

Description:

Use as facet in entries browser

Select a few dimensions which are useful to slice the dataset by.

Para terminar: salvar y cargar

Cuando haya especificado todas las dimensiones y vinculado a las columnas de los datos de origen, haga clic en **Guardar Dimensiones** para guardar el modelo. Si algo está mal con el modelo, aparecerá un mensaje de error, solicitándole corregir sus parámetros. De lo contrario, aparecerá un mensaje que invita a volver al tablero, donde se puede proceder a cargar sus datos.

Una vez que los datos se han cargado, el modelo que ha creado se fija y se inhabilitará la edición. Por lo tanto, puede que desee probar el modelo antes de cargarlo. Para ello, haga clic en **Probar una muestra** en la fila del origen de datos en el tablero. Espere unos segundos y vuelva a cargar la página. Si aparece un mensaje que dice COMPLETA con un fondo verde, el modelo está listo para ir. Si ve ERRORES, se necesitan reparaciones.

Cameroon Local Council Budget

Dashboard Dimensions & Measures Views Team Meta Templates

Source files

Link	Operations
OK https://docs.go .. gle=true&gid=1&output=csv 12 columns	<input type="button" value="Test a sample"/> <input type="button" value="Load"/>
Reports:	<ul style="list-style-type: none">• 01.08.2013, 20:15 COMPLETE (0 Errors)• 01.08.2013, 20:13 ERRORS (1 Errors)• 01.08.2013, 20:13 COMPLETE (0 Errors)• 01.08.2013, 19:14 COMPLETE (0 Errors)

Si el modelo está libre de errores, haga clic en **Cargar** para cargar el conjunto de datos de origen y aplicar el modelo. A continuación, puede volver a la página principal de la base de datos haciendo clic en su nombre en la parte superior de la pantalla, donde puede proceder a la construcción de las visualizaciones y jugar con sus datos de otra manera .

Visualizaciones

Crear una Visualización

La plataforma de OpenSpending hace que sea fácil para crear e incrustar visualizaciones de datos. Hay tres tipos de visualizaciones compatibles: BubbleTree, TreeMap, y la Tabla de los agregados.

Todas las visualizaciones de OpenSpending permiten elegir una serie de medidas a lo largo de la cual se agregan los datos, profundizando en detalles más y más finos. Cada visualización se crea de la misma manera: por la elección de las dimensiones para agregar y el orden en el que se pueda averiguar.

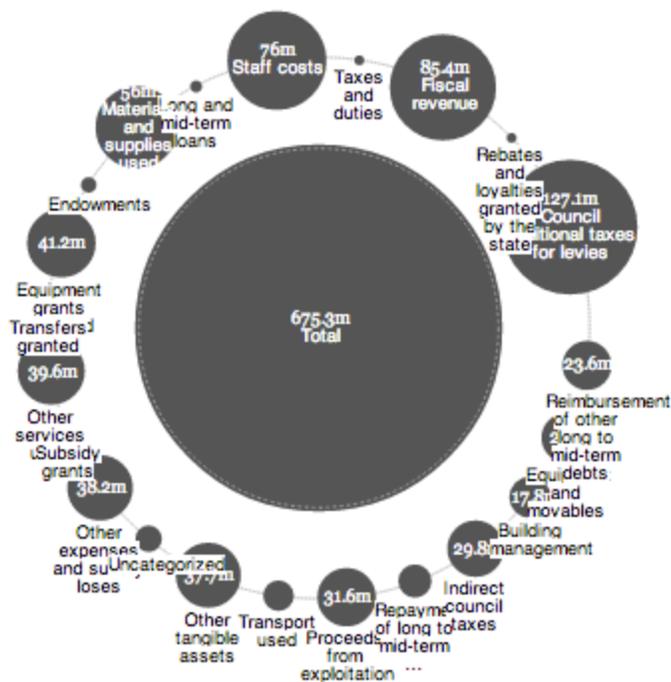
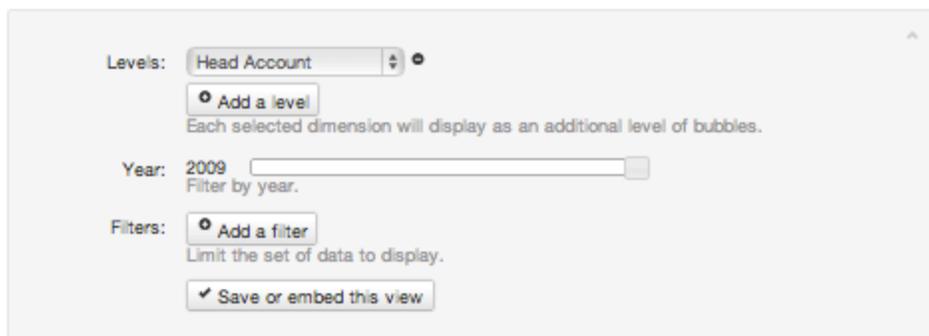
Para empezar a crear una visualización, vaya a la página principal de un conjunto de datos y seleccione **Crear una visualización** del menú *Visualizaciones*.

BubbleTree

El BubbleTree es una visualización interactiva que presenta los datos agregados de gasto como un círculo de burbujas. Cada burbuja representa un (sub) total agregado. La burbuja central representa una suma agregada y sus burbujas que rodean representan las otras

sumas de las que se compone. Al hacer clic en cualquier burbuja, se muestra al usuario cómo la suma se divide en más sub-totales.

Para crear un BubbleTree, elija las dimensiones para agregar y el orden para agregarlos. Elija la dimensión principal en el menú desplegable *Nivel*. Verá el total agregado de esa dimensión como la burbuja central, con los valores de la dimensión que la rodea con sus propios totales.



Para añadir un segundo nivel, haga clic en **Agregar un nivel** y elegir una nueva dimensión. Los usuarios podrán hacer clic en las burbujas para "profundizar" y ver cómo los valores del primer nivel se dividen en valores en el segundo nivel.

Levels:

Each selected dimension will display as an additional level of bubbles.

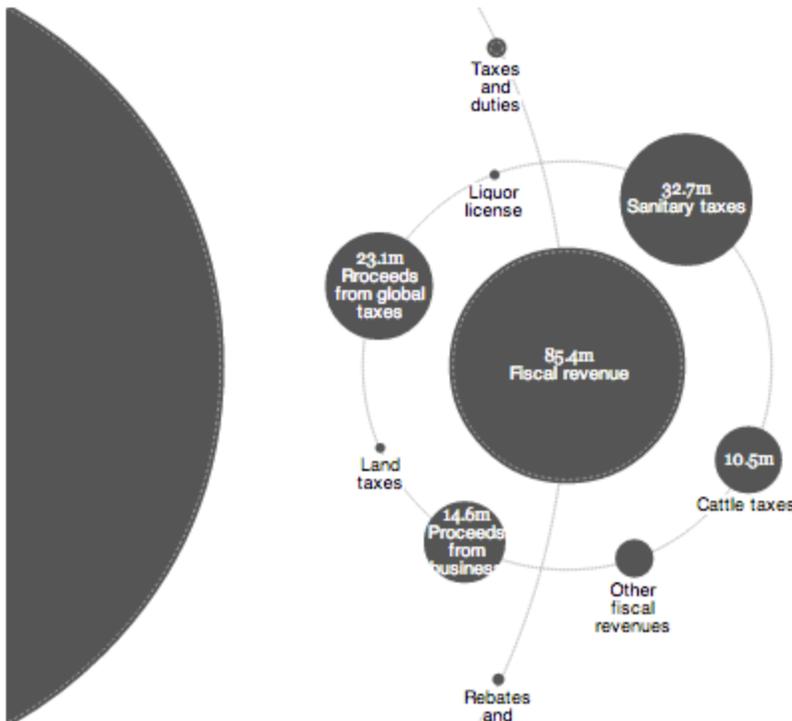
Year: 2009

Filter by year.

Filters:

Limit the set of data to display.

Save or embed this view



TreeMap

El TreeMap presenta datos agregados de gasto como un rectángulo interactivo de mosaicos de colores. Cada mosaico representa valores agregados para una dimensión particular de los datos. Al hacer clic en el mosaico, se acerca para mostrar cómo se desglosa a lo largo de las dimensiones más globales.

Para crear un TreeMap, introduzca las dimensiones para agregar y su orden. Seleccione la dimensión principal en el menú *Tile*. Verá un TreeMap mostrando cómo el gasto total se descompone en esa dimensión.

Tiles: Each selected dimension will display as an additional level of tiles for the treemap.

Year: 2009 Filter by year.

Filters: Limit the set of data to display.

Save or embed this view



La visualización no tiene interactividad todavía. Al añadir más niveles de azulejo le permite profundizar para ver cómo los valores agregados se descomponen en agregados más pequeños. Para añadir un segundo nivel de azulejos, haga clic en **Agregar un nivel** y elija una nueva dimensión. Ahora los usuarios pueden hacer clic en las fichas para ver cómo sus totales se descomponen.

Tiles:

Each selected dimension will display as an additional level of tiles for the treemap.

Year: 2009

Filter by year.

Filters:

Limit the set of data to display.

Save or embed this view

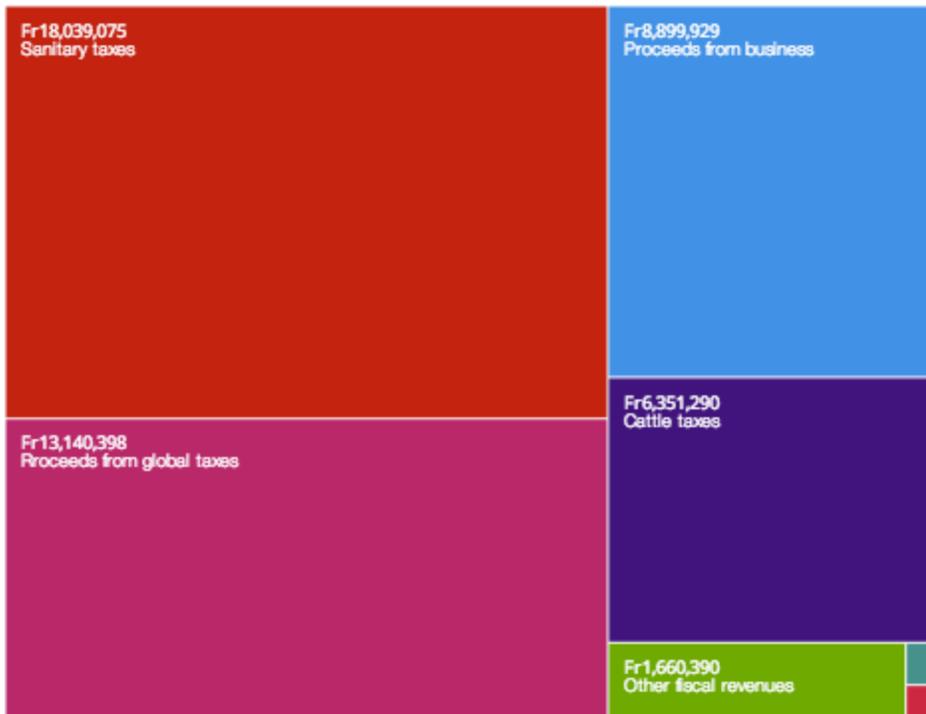


Table of Aggregates

La tabla de Agregados es una vista de tabla simple de un conjunto de datos que agrega totales a través de dimensiones seleccionadas. Una tabla de Agregados se especifica eligiendo dimensiones para sus columnas.

La elección de una dimensión primaria a través del menú *Columnas* mostrará los datos en forma de tabla, con cantidades y porcentajes acumulados del total general. De forma predeterminada, las filas se ordenarán en función del porcentaje.

Columns:

Each selected dimension will display as a column, and values will be drilled down by it.

Year: 2009

Filters:

Limit the set of data to display.

Head-account	Amount (Fr)	%
Council additional taxes for levies	62,341,392	17.35%
Fiscal revenue (1)	48,272,332	13.44%
Staff costs (19)	39,335,512	10.95%
Equipment grants received (12)	34,164,475	9.51%
Materials and supplies used (16)	29,845,704	8.31%
Other tangible assets (28)	22,708,188	6.32%
Other services used (18)	20,768,408	5.78%
Other expenses and sundry losses (24)	18,908,318	5.26%
Indirect council taxes (3)	16,040,184	4.46%
Proceeds from exploitation of council properties and services (4)	12,874,000	3.58%
Building management (29)	12,433,106	3.46%
Reimbursement of other long to mid-term debts (26)	11,244,134	3.13%
Equipment and movables (30)	9,653,865	2.69%
Transport used (17)	7,654,000	2.13%
Repayment of long to mid-term loans resulting from commitment made by the council	5,984,456	1.67%
Total	359,255,890	100.00%

Showing 1 to 15 of 22 entries

← Previous 1 2 Next →

Al agregar otra columna, haciendo clic en **Agregar un nivel**, desagregará cada subtotal de la primera columna de las cantidades de la nueva columna. Tenga en cuenta que esto generalmente cambia los valores de porcentaje y por lo tanto reorganiza las filas.

Columns:

Each selected dimension will display as a column, and values will be drilled down by it.

Year: 2009

Filter by year.

Filters:

Limit the set of data to display.

Head-account	Sub-account	Amount (Fr)	%
Council additional taxes for levies	Council additional taxes for levies	62,341,392	17.35%
Staff costs (19)	Basic wages and gross salary	24,268,589	6.76%
Equipment grants received (12)	Equipment grants from FEICOM	20,164,000	5.61%
Fiscal revenue (1)	Sanitary taxes	18,039,075	5.02%
Equipment grants received (12)	Equipment grants from urban council	14,000,475	3.90%
Materials and supplies used (16)	Fuel and lubricants	13,960,905	3.89%
Fiscal revenue (1)	Proceeds from global taxes	13,140,398	3.66%
Other tangible assets (28)	Construction of road networks (roads, ferries, dams, drains, sewers, electricity, water, public lighting, road signs)	13,110,000	3.65%
Reimbursement of other long to mid-term debts (26)	Short-term SONEL, SNEC and P & T loans included in the accounts consolidated as long to mid term loans	11,244,134	3.13%
Other services used (18)	Total Sub-Head 6.12	10,384,204	2.89%
Fiscal revenue (1)	Proceeds from business	8,899,929	2.48%
Indirect council taxes (3)	Market tolls	8,209,150	2.29%
Proceeds from exploitation of council properties and services (4)	Rents and hires of council properties	8,120,000	2.26%
Building management (29)	Layout and construction of building	6,912,908	1.92%
Fiscal revenue (1)	Cattle taxes	6,351,290	1.77%
Total	Total	359,255,890	100.00%

Showing 1 to 15 of 121 entries

← Previous 1 2 3 4 5 Next →

Insertar una visualización en su sitio web

Usted puede fácilmente integrar cualquiera de las visualizaciones creadas en OpenSpending en su propio sitio web. Esto significa que usted puede tener las pantallas interactivas completas en su sitio.

Digamos que usted ha elegido una visualización en la plataforma Open Spending. Observe que hay un botón **Insertar** en la parte inferior derecha de la página. Haga clic en este botón y se le presentará con el código para incrustar la visualización de su sitio web y algunas opciones para el tamaño (en píxeles) del interactivo. El resto es sólo cortar y pegar este código en su sitio. Si no está seguro de cómo pegar el código correctamente, póngase en contacto con el administrador del sitio.

La razón por la que es posible incrustar códigos se reduce a los *widgets*. En términos muy

simplificados, un widget es una pieza de código que puede agregar a su página web, y extrae los datos - en este caso, de la base de datos de Open Spending Open - por lo que no es necesario que almacene bases de datos usted mismo.

Sitios web Satellites

INESC - Orçamento ao seu Alcance (Presupuesto a su alcance)

Esta fue una colaboración entre Brasil y OKF INESC (Instituto de Estudios Socio-Económicos), una organización no gubernamental brasileña. El objetivo era hacer más fácil para el público en general ver el presupuesto federal brasileño, y cómo se reparte entre las diferentes entidades públicas, con especial atención a exceso de gastos o infrautilización de fondos.

Los datos provienen de SIGA Brasil, un agregador de los muchos sistemas utilizados por el gobierno para organizar el presupuesto. Nos permite elegir las columnas que queremos, como organismo público, categoría, subcategoría, presupuesto, gastos, etc, y la exportación como un archivo CSV. Hay datos desde 2001 hasta la actualidad, actualizada diariamente. Aparte de algunos problemas, como filas con mes "00", no tuvimos que limpiar mucho para cargarlo en OpenSpending.

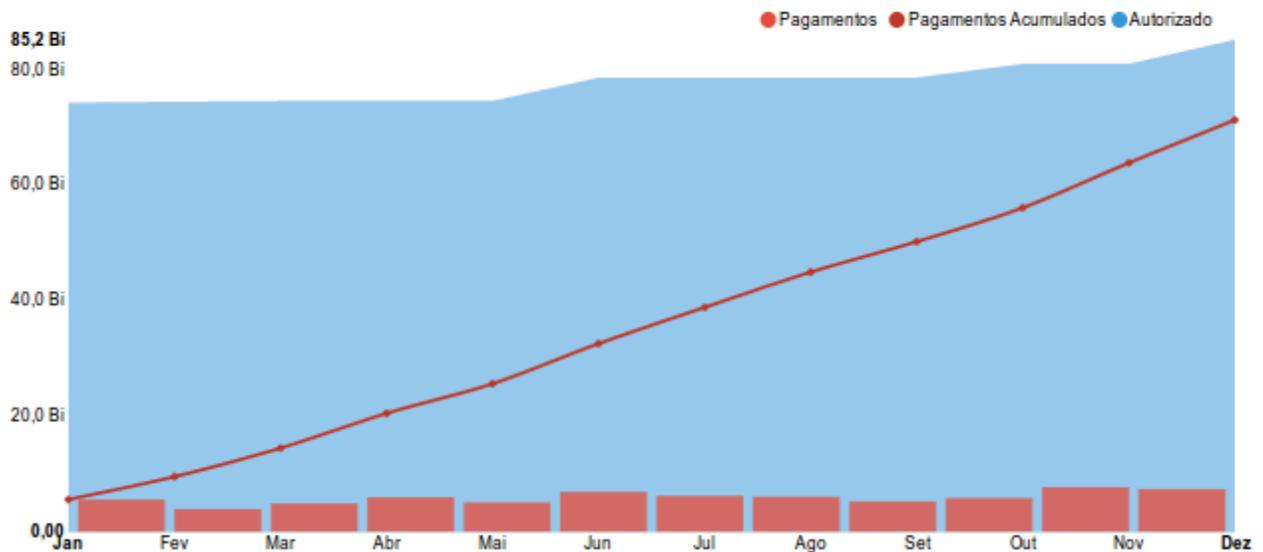
Construyendo el sitio

Sabíamos que queríamos centrarnos en infrautilización, lo que ocurre en diversos grados en todos los organismos públicos. En 2012, por ejemplo, el Ministerio de Educación no gastó el 16,3% de su presupuesto (alrededor de US\$ 6,100,000,000). OpenSpending no tenían un gráfico para este tipo de datos, así que lo diseñamos.

Después de unas horas dibujando, decidimos hacer un análisis con gráficos de área, barras y líneas. El gráfico de la figura X muestra datos del Ministerio de Educación de 2012. El área azul representa el presupuesto (tenga en cuenta que cambia a lo largo del año). Las barras muestran cuánto se pagó en ese mes en particular, y la línea muestra el gasto total hasta ese momento. Usted puede ver, por la distancia entre la línea roja en diciembre hasta la punta (o la zona azul), que se infrautilizó bastante.

Pagamentos por mês para 2012 ▼

Posicione o cursor sobre as áreas para visualizar detalhes.



Para construir este gráfico utilizamos NVD3, uma biblioteca de JavaScript com uma coleção de tabelas reutilizáveis hechas encima de la D3. La información proviene de OpenSpending, utilizando la API Agregada. Es muy bueno, y después de que hace una agregación, almacena en caché el resultado y se vuelve muy rápido. Pero hay una limitación que nos dio algunos problemas: sólo se puede utilizar una medida a la vez.

En este gráfico, tenemos 2 medidas: el presupuesto y los gastos. Sin embargo, a nivel interno, los gastos se dividen en dos partes: lo que se pagó por el año en curso, y las deudas de años anteriores que fueron pagadas. Así, nos encontramos con 3 medidas. A medida que la API Agregada sólo permite una, tuvimos que hacer tres solicitudes para construir este gráfico.

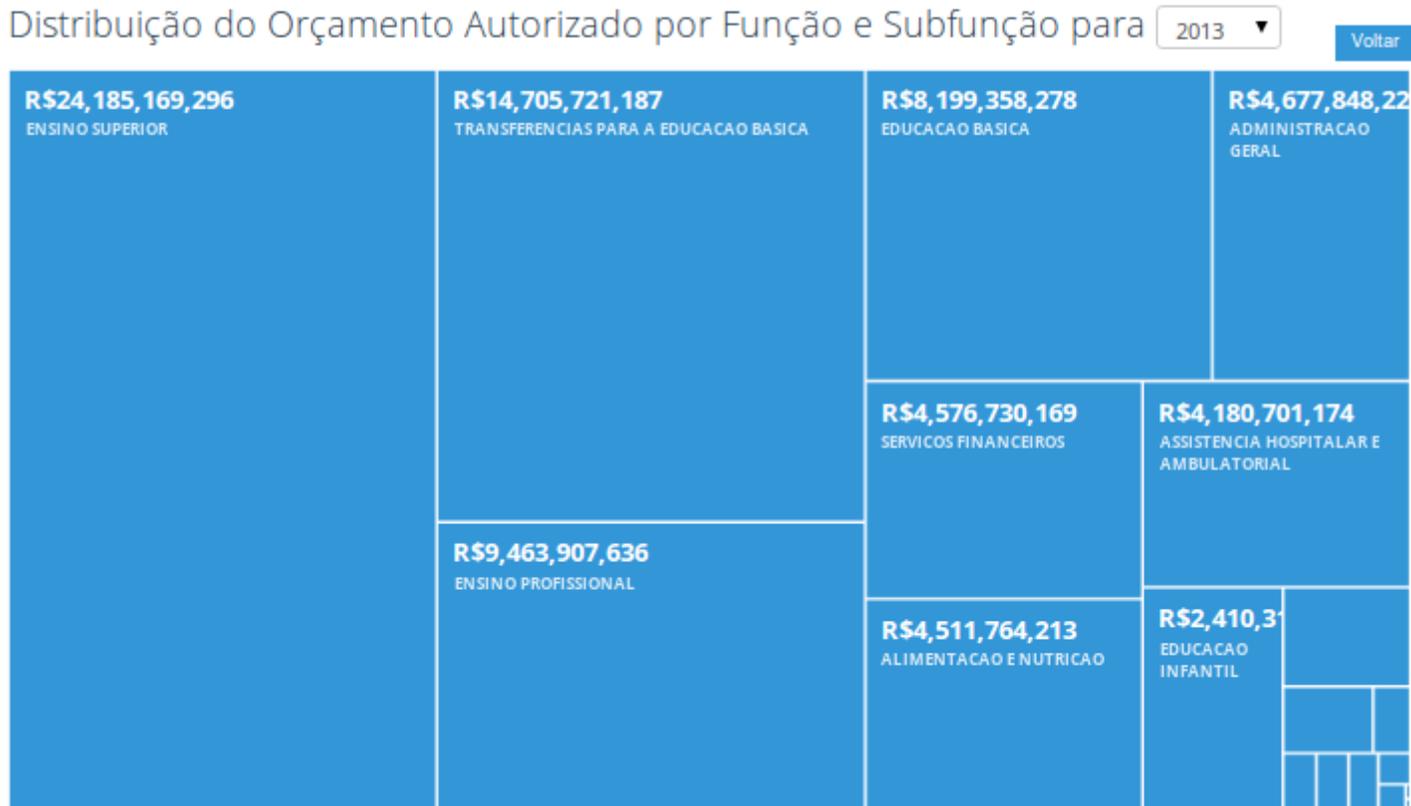
Esto, obviamente, genera un problema de rendimiento, tanto para nuestro proyecto y para OpenSpending en sí. Pero, como las solicitudes se almacenan en caché después del primer intento, funciona muy bien. Ya hay planes para soportar múltiples medidas en el API, por lo que el problema se resolverá pronto.

Usando el Treemap

En la página de índice, hemos querido mostrar una visión amplia del presupuesto a través de todos los organismos públicos. Más que eso, hemos querido mostrar la cantidad de dinero que se utiliza en cada función y subfunción, como la educación y la educación básica. Para mostrar esto, elegimos el Treemap.

Incrustar esto como un widget era sencillo: simplemente créelo en OpenSpending, copie el código y péguelo en el sitio. Sin embargo, hay algunas limitaciones.

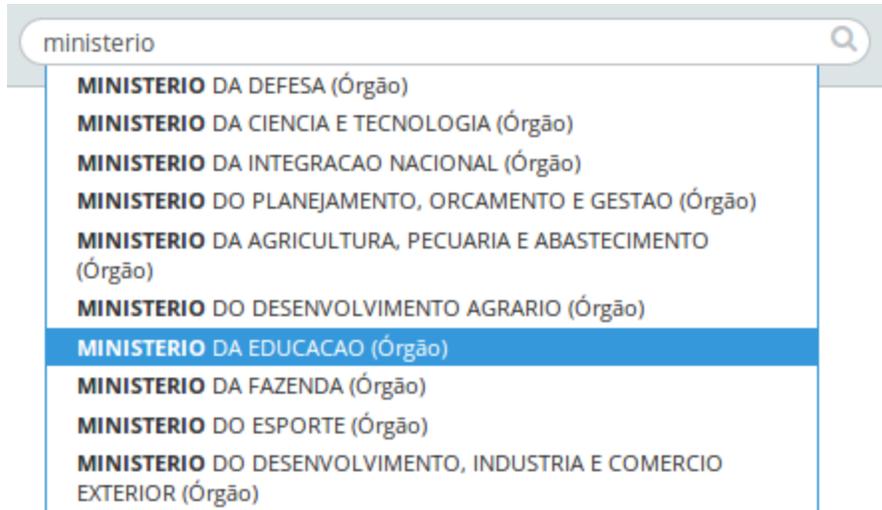
Los widgets están hechos cuando usted quiere poner el cuadro en un blog o en un artículo de periódico. No se puede personalizar. Esto no funcionó para nosotros. Tuvimos que cambiar sus fuentes y colores, para que se ajustara en el diseño del resto de la página. Como se trata de un *iframe*, no hay manera de cambiarlo usando sólo CSS. Pero hay una solución fácil: copiar el código *iframe* del widget en su página.



No se necesita mucho para construir un Treemap, sólo unas pocas líneas de código de inicialización y funciona. Cuando está en su página, puede utilizar CSS. Desafortunadamente, eso no funciona para todo: los colores no se pueden cambiar de esa forma. Pero es fácil de configurar otro esquema de colores, sólo tiene que cambiarlo en el código de inicialización. También hemos añadido un botón "Volver", para que pueda navegar entre función y subfunción fácilmente.

Búsqueda

Para ayudar al usuario a encontrar los organismos públicos, se implementó una caja de búsqueda de auto-completar, usando el plugin de *typeahead* de Twitter Bootstrap. Como no hay muchas entidades (alrededor de 500), decidimos cargarlas todas cuando el usuario entra por primera vez a la página, por lo que la búsqueda será instantánea.



Para hacerlo, necesitábamos una lista con los organismos públicos, y sus identificadores. El API Agregada, de nuevo, nos ayudó. Con un simple desglose por organismo público, hemos sido capaces de obtener todos los datos que necesitábamos, y construir la búsqueda. Si tuviéramos más entidades, tendríamos que utilizar la API de Búsqueda.

Conclusion

<OpenSpending helped a lot>

<We don't even have a database, our page is 100% cacheable. We could have done everything with just HTML+JS>

<INESC is able to update the data, with no dependency on us>

<Everyone's happy!>