

Le Guide OpenSpending

Par l'équipe OpenSpending : Anders Pedersen, Rufus Pollock, Michael Bauer, Neil Ashton, Lisa Evans, Tony Hirst, Pierre Chrzanowski, Félix Ontañón, Oluseun Onigbinde, Vítor Baptista

Traduction française : Pierre Chrzanowski, Elisabetta Lombardo, Claire-Marie Foulquier-Gazagnes

Introduction

[Qu'est ce que OpenSpending ?](#)

[Quels types de données financières peut-on ajouter dans OpenSpending ?](#)

[Les données de transactions financières](#)

[Les données budgétaires](#)

[Représentation des données dans OpenSpending](#)

[Jeu de données](#)

[Modèle](#)

Ajouter des données dans OpenSpending

[Introduction](#)

[Rassembler les données](#)

[Formater les données](#)

[Fichier CSV](#)

[Le format OpenSpending](#)

[Publier les données sur le web](#)

[Google Drive](#)

[Gist](#)

[Créer un jeu de données sur OpenSpending](#)

[Créer un nouveau jeu de données](#)

[Ajouter une nouvelle source de données](#)

[Modéliser vos données dans OpenSpending](#)

[Dimensions obligatoires : montant et date](#)

[Autre dimension indispensable : l'identifiant](#)

[Mesures et autres dimensions](#)

[Pour finir : sauvegarder et charger](#)

[Visualisation](#)

[Créer une visualisation](#)

[BubbleTree](#)

[TreeMap](#)

[Tableau d'agrégats](#)

[Inclure une visualisation sur votre site web](#)

Introduction

Qu'est ce que OpenSpending ?

OpenSpending est une application web et une communauté de contributeurs dont l'objectif est de répertorier toutes les transactions financières des entreprises et gouvernements à travers le monde et présenter ces données de manière utile et engageante. OpenSpending est un projet ouvert maintenu par la communauté. Toutes les personnes intéressées par les données financières de toutes sortes sont invitées à contribuer à la base de données OpenSpending, à créer des visualisations en utilisant l'application OpenSpending, et à utiliser l'API OpenSpending.

Ce chapitre introduit les concepts principaux du système aux nouveaux contributeurs d'OpenSpending. Il décrit les différents types de données financières que OpenSpending supporte, et explique comment l'application représente les données.

Quels types de données financières peut-on ajouter dans OpenSpending ?

OpenSpending est très flexible et admet de nombreux type de données financières. Et si le projet OpenSpending s'intéresse en particulier aux données financières des gouvernements, cela n'est pas une limitation technique. OpenSpending supporte ainsi n'importe quel jeu de données contenant une liste de transactions associées à une quantité monétaire et à une date.

La plupart des données actuellement hébergées sur OpenSpending sont des données budgétaires ou bien des données transactionnelles. La principale différence entre ces deux types de données est leur granularité. Les données transactionnelles portent sur des transactions uniques réalisées, alors que les données budgétaires sont des données de dépenses agrégées par catégories.

Les données de transactions financières

Les données transactionnelles, ou plus simplement les "données de dépense", renseignent sur les transactions financières. Chaque paiement d'une entité à une autre entité pour une date donnée et avec un objet particulier (un projet, un service, etc.) est alors renseigné.

Les informations agrégées (ex: un sous-total de transactions) ne devraient pas être incluses dans les données de transactions. Les données qui ont été partiellement ou complètement agrégées ne peuvent pas être analysées de la même manière et devraient donc être traitées comme des données budgétaires plutôt que comme des données transactionnelles. En revanche, si plusieurs paiements ont été effectués à différentes dates pour régler une même dépense, ceux-ci devraient pouvoir être agrégés pour représenter une transaction unique.

Exemples de données transactionnelles sur OpenSpending :

- [Contrats publics Washington D.C.](#)
- [Agence de Développement Autrichienne](#)

Un autre type de données concerne les marchés publics. Les données sur les marchés publics informent sur l'objet du marché, le montant du marché, et qui a remporté le marché. Les données des marchés publics peuvent être considérées comme un sous-type de données transactionnelles (dans le cas où il s'agit d'un marché réalisé).

Exemples de données sur les marchés publics sur OpenSpending :

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

Les données budgétaires

Pour les données budgétaires, les dépenses et les recettes sont agrégées par catégories. L'objectif de ces agrégations est d'aider le lecteur à comprendre le budget, qui est habituellement un document de politique publique utilisé pour donner au lecteur un aperçu des choix financiers les plus importants du gouvernement. Les fonds alloués sont habituellement structurés en fonction de secteurs plutôt qu'en fonction des actuels bénéficiaires.

Les données budgétaires présentent généralement à la fois les recettes et dépenses passées et le budget prévisionnel. Ainsi, les montants des dépenses pour les années précédentes dans un secteur particulier donnent une indication sur les fonds qui devrait être attribués pour la période suivante. Les données budgétaires sont en général composées de données agrégées et d'estimations statistiques.

Les pays publient différents types d'informations budgétaires, incluant : le pré-budget, la proposition de budget (projet de loi de Finance), le budget ratifié, et le budget citoyen (qui est une version simplifiée du budget pour une meilleure compréhension par les citoyens).

Des exemples de données budgétaires sur OpenSpending :

- [Budget de la ville de Berlin](#)
- [Budget de la ville de Séville](#)

Représentation des données dans OpenSpending

OpenSpending contient une collection de jeux de données, dont chacun provient d'une source différente. Dans un jeu de données, chaque transaction est représentée par une liste d'entrées (ou attributs). Chaque jeu de données a son propre modèle de structure de données. Un modèle décrit les propriétés d'un jeu de données en termes de dimensions.

Jeu de données

L'unité de base dans OpenSpending est le jeu de données. Les données financières partageant un sujet commun (un budget pour une ville, le budget d'une année) sont regroupées ensemble et enregistrées dans un jeu de données. Un jeu de données est une collection d'«entrées», et chaque entrée représente une transaction unique associée à un montant et à une date.

Un jeu de données inclut également des métadonnées pour décrire son contenu. Les métadonnées contiennent une description du jeu de données, des informations sur la source des données, et d'autres informations qui peuvent aider l'utilisateur à trouver et interpréter le contenu de ce jeu de données.

Modèle

Le créateur d'un jeu de données a la maîtrise complète sur la structure du jeu de données. Cette structure est créée en définissant un modèle qui est composé de dimensions.

Une dimension est une propriété qui définit une entrée du jeu de données. En considérant qu'un jeu de données est un tableau, les dimensions peuvent alors être assimilées aux colonnes. Cependant, une dimension peut avoir une structure plus complexe qu'une seule colonne.

Les dimensions sont de différentes natures. La plus importante étant celle de type **measure** (montant). Une *measure* est une dimension qui contient une valeur numérique. Une autre dimension importante est celle de type **time** (temps), qui représente la date et l'heure. Chaque entrée nécessite au moins une dimension *measure* et une dimension *time*, représentant respectivement le montant et la date d'une transaction.

Les autres dimensions sont utilisées pour représenter d'autres propriétés du jeu de données, telles que le numéro de la transaction, le type de la transaction, le nom de l'administration et le nom de la société ou l'individu impliqué. Ces dimensions incluent des **attributes** (attributs), qui contiennent une valeur unique, et les **compound dimensions** (dimensions composées), qui peuvent contenir plusieurs valeurs. Les dimensions composées peuvent être utiles lorsqu'une propriété est définie par plusieurs autres propriétés.

Ajouter des données dans OpenSpending

Introduction

L'une des contributions les plus utiles que vous pouvez faire sur OpenSpending est d'ajouter un nouveau jeu de données. Ce paragraphe vous explique comment faire.

Une démarche classique pour importer un jeu de données sur OpenSpending suit les étapes suivantes :

- Rassembler les données lisibles par une machine depuis une source officielle.
- Convertir les données dans le format CSV supporté par OpenSpending. Nettoyer le jeu de données pour supprimer les erreurs et les éventuelles incohérences.
- Publier le jeu de données sur le web.
- Créer un jeu de données sur OpenSpending en indiquant la source (URL) du jeu de données.
- Créer un modèle en définissant le rôle de chaque colonne du jeu de données.
- Charger les données, ou corriger le jeu de données en vous aidant des

recommandations fournies par OpenSpending sur la cohérence des données.

Chacune de ces étapes est détaillée dans les sections suivantes.

Rassembler les données

Pour ajouter un jeu de données sur OpenSpending, vous devez déjà avoir des données à disposition. Si ce n'est pas le cas, vous devrez chercher une source de données sur le web (ou en ajouter une).

Commencez votre recherche en consultant les ressources telles que [School of Data](#) (en anglais) ou le [Guide du Datajournalisme](#). Consultez également les sources officielles de données telles que [data.gouv.fr](#). Vous pouvez également trouver de l'aide et des idées en vous rendant sur le groupe [OpenSpending](#) sur datahub.io, ou en posant des questions sur le channel IRC #openspending sur Freenode.

Si les données trouvées ne sont pas dans un format lisible par une machine, comme un fichier Excel ou CSV, mais dans un format du type PDF ou Word, il sera très difficile de travailler avec. Vous devriez alors peut être penser à essayer avec d'autres données.

Formater les données

OpenSpending suppose que les données soient fournies dans un seul format simple.

Fichier CSV

OpenSpending accepte les données dans un seul type de format de fichier, le format CSV (Comma-Separated Values). Un fichier CSV est un fichier texte qui représente les données sous la forme d'un tableau, similaire à un format tableau. Dans un tableau, chaque entrée est représentée par une ligne, et chaque propriété de l'entrée est représentée par une colonne. Dans un fichier CSV, les lignes sont subdivisées en colonnes à l'aide de virgules.

Les fichiers CSV acceptés par OpenSpending sont "dénormalisés": cela signifie qu'ils n'économisent pas d'espace en supprimant des valeurs redondantes. Ils doivent également être "rectangulaires", c'est-à-dire qu'ils ont exactement le même nombre de colonnes pour chacune des lignes.

Le format OpenSpending

Un fichier CSV doit avoir les caractéristiques suivantes pour OpenSpending.

1. Une ligne de titre. La première ligne du fichier CSV devrait contenir les noms des colonnes, séparés par des virgules.
2. Au moins 3 colonnes : un montant, une date (qui peut être juste une année), et le payeur ou le bénéficiaire (qui peut juste être le nom d'un compte).

3. Des colonnes homogènes. Chaque colonne doit toujours représenter un seul type de valeur pour toutes les lignes (Il ne peut y avoir, par exemple, de sous-division).
4. Les lignes sont des points de données. Chaque ligne ne doit contenir qu'un type de données : une transaction ou une ligne budgétaire (elles ne peuvent représenter plusieurs transactions)
5. Aucune ligne ou cellule vide. Chaque ligne d'un fichier devrait contenir toutes les informations requises afin de pouvoir construire l'objet résultant.
6. Pas de résultats agrégés (exemple : sous total ou calculs à l'aide de formules). Ils seront calculés automatiquement par OpenSpending.
7. Un identifiant unique. Il doit y avoir une colonne ou une combinaison de colonnes permettant d'identifier chaque ligne. La méthode la plus simple pour créer un tel identifiant est de créer une colonne supplémentaire dans laquelle vous ajoutez un chiffre qui s'incrémente à chaque ligne. Vous pouvez le faire dans Excel en ajoutant les deux premiers chiffres de la série, sélectionnez les deux cellules, et faire glisser le coin droit de la seconde cellule tout en bas du jeu de données pour étendre automatiquement la série. Si le jeu de données a trop de colonnes pour un logiciel comme Excel, vous pouvez utiliser [MySQL](#).
8. Les dates dans le bon format. Les dates doivent toutes être dans le format "aaaa-mm-jj"
9. Les nombres dans le bon format. Les nombres doivent contenir uniquement des chiffres et éventuellement un point (pas une virgule) pour séparer les décimales (un nombre écrit "12,345.67" devrait être converti en "12345.67").

La communauté OpenSpending a rassemblé des [exemples](#) afin d'illustrer les «bons» et les «mauvais» exemples de formatage de jeux de données.

Voici des exemples de jeux de données **mal formatés** :

- [Plusieurs transactions, une ligne](#) (plusieurs années sur une seule entrée)
- [Format de chiffre incorrect](#) (les chiffres contiennent des virgules pour la lisibilité)

Voici des exemples de **bons** jeux de données :

- [Washington, DC](#) (données de dépenses)
- [Minsk, Biélorussie](#) (données budgétaires)

Publier les données sur le web

Les données ne peuvent pas encore être téléchargées directement sur OpenSpending. Pour publier des données, celles-ci doivent être mises à disposition sur le web.

Google Drive

Vous pouvez mettre vos données à disposition sur le web en utilisant Google Drive spreadsheet.

1. Importez vos données dans Google Drive. Créez une nouvelle feuille de calcul sur

Google Drive, sélectionnez ensuite importer depuis le menu Fichier. Sélectionnez Remplacer la feuille de calcul, cliquez Choose File, et choisissez votre fichier CSV. Ensuite, cliquez sur Importer.

2. Assurez-vous que Google Docs formate correctement votre jeu de données. Sélectionnez la colonne qui contient les dates. Cliquez sur le menu Format et sélectionnez Nombre -> Autre formats -> 2008-09-26. Vos dates devraient apparaître dans le format aaaa-mm-jj utilisé par OpenSpending.
3. Cliquez sur le menu Fichier et sélectionnez Publier sur le web ... Dans la fenêtre qui apparaît, cliquez sur Démarrer la publication. En dessous, dans la section Insérer un lien vers les données publiées, sélectionnez CSV (valeurs séparées par des virgules). L'URL en bas est celle de votre jeu de données.

Publish to the web

Sheets to publish

All sheets ▾

Automatically republish when changes are made
 Requires viewers to sign in with their okfn.org account.

Stop publishing Republish now Published on 5 Aug 2013 14:18

Note: Publishing a doc does not affect its visibility option. [Learn more](#)

Get a link to the published data

CSV (comma-separated values) ▾

All sheets ▾

All cells

`https://docs.google.com/spreadsheet/pub?key=0AqRoDZ_zG8lKdE5JMTd3WDV6bnFEX1BKYTZ6RHBIaXc&outpu`

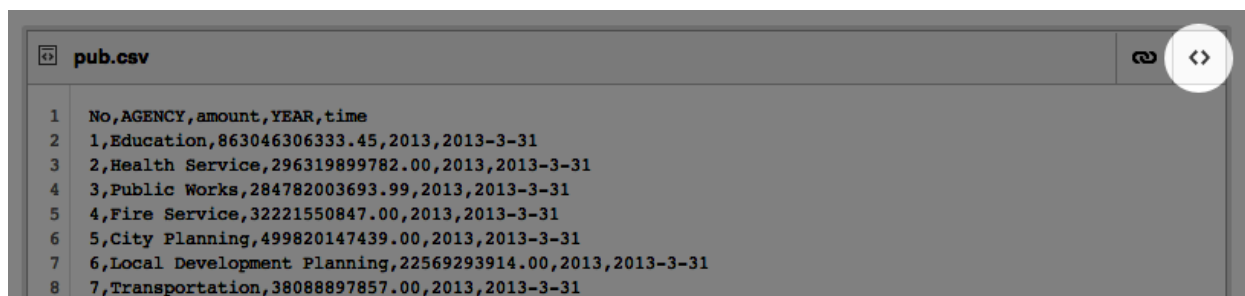
Copy and paste the link above

Close

Gist

GitHub Gist est un moyen très pratique pour héberger des fichiers textes de petite taille, notamment des fichiers CSV.

1. Connectez-vous à GitHub (ou enregistrez-vous si vous ne l'avez pas déjà fait), allez ensuite sur gist.github.com.
2. Cliquez et déplacez votre fichier CSV depuis le répertoire sur votre ordinateur vers la page GitHub Gist dans votre navigateur. Le nom du fichier et son contenu apparaîtront.
3. Cliquez sur **Create Public Gist** pour être redirigé sur la page d'accueil de votre nouveau gist. L'URL de votre jeu de données est accessible via le bouton "< >" dans le coin en haut à droite du fichier.



```
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
```

Créer un jeu de données sur OpenSpending

Pour pouvoir partager vos données sur OpenSpending, vous devez d'abord vous enregistrer sur le site [OpenSpending.org](https://www.openspending.org). Pour créer un jeu de données, il vous suffira ensuite d'indiquer ses métadonnées et son URL.

Créer un nouveau jeu de données

Connectez-vous sur OpenSpending avec vos informations d'utilisateur. Vous serez ensuite redirigé vers le *Dashboard*. Cliquez sur le bouton *Import a Dataset* pour créer un nouveau jeu de données.

L'écran suivant vous invite à fournir les métadonnées qui définissent votre jeu de données. Voici les informations demandées :

- Title (Titre) : une description et un nom explicite pour le jeu de données.
- Identifier (Identifiant) : un titre court, utilisé dans l'URL du jeu de données. L'identifiant peut contenir uniquement des caractères alphanumériques, des tirets, et des traits bas – ni espace blanc ni signe de ponctuation.
- Category (Catégorie) : pouvant être "Budget" (Budget), "Expenditure" (Dépense) ou "Other" (Autres). Référez-vous aux types de données financières pour des détails sur ces catégories.

- Currency (devise) : la devise utilisée pour les montants dans le jeu de données.
- Countries (Pays) : une liste de pays référencés dans le jeu de données. Le choix des pays est contraint à une liste de pays valides.
- Languages (Langue) : une liste des langues utilisées dans le jeu de données. Le choix des langues est contraint par une liste de langues valides.
- Description (Description) : une définition du jeu de données en langage simple.

Remplissez ces champs. Assurez-vous d'inclure une description qui explique l'origine de votre jeu de données et les modifications apportées (par exemple que le jeu de données a été nettoyé).

Dès que toutes les métadonnées ont été renseignées, cliquez sur **Next Step** pour continuer.

Ajouter une nouvelle source de données

L'étape suivante vous emmène sur la page *Manage* (Gestion). La page *Manage* est utilisée pour ajouter la source du jeu de données. Elle est aussi utilisée pour créer le modèle de schéma de données qui permet à OpenSpending d'interpréter les informations. Cette étape, dénommée *modelling* (modélisation) est expliquée en détail dans la section suivante du guide.

Pour ajouter une source de données à un jeu de données, cliquez sur **Add a source**. Une fenêtre apparaît vous demandant d'entrer une URL. Indiquez l'URL du fichier CSV que vous avez précédemment publié sur le web comme indiqué dans la section précédente du guide et cliquez sur **Create**. Vous verrez alors une fenêtre bleue indiquant qu'OpenSpending est en train d'examiner les données.

Let me see... The data is now being checked to see if it can be imported. This should not take very long.

Source files

Link	Operations
https://docs.go .. gle=true&gid=1&output=csv	<input type="button" value="Test a sample"/> <input style="background-color: #4caf50; color: white;" type="button" value="Load"/>

Analysis: STILL ANALYSING Try refreshing in a moment.

Cliquez sur **Refresh** ou utilisez le bouton de rafraîchissement de votre navigateur. Si OpenSpending parvient à analyser vos données, vous verrez alors un message vert vous

disant que les données sont prêtes. Vous devriez également voir une liste correcte des colonnes de votre 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

Notez que si vous fournissez à OpenSpending un fichier HTML à la place d'un fichier CSV, l'application analysera quand même le HTML comme si c'était un CSV. Le résultat ressemblera à cela :

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>

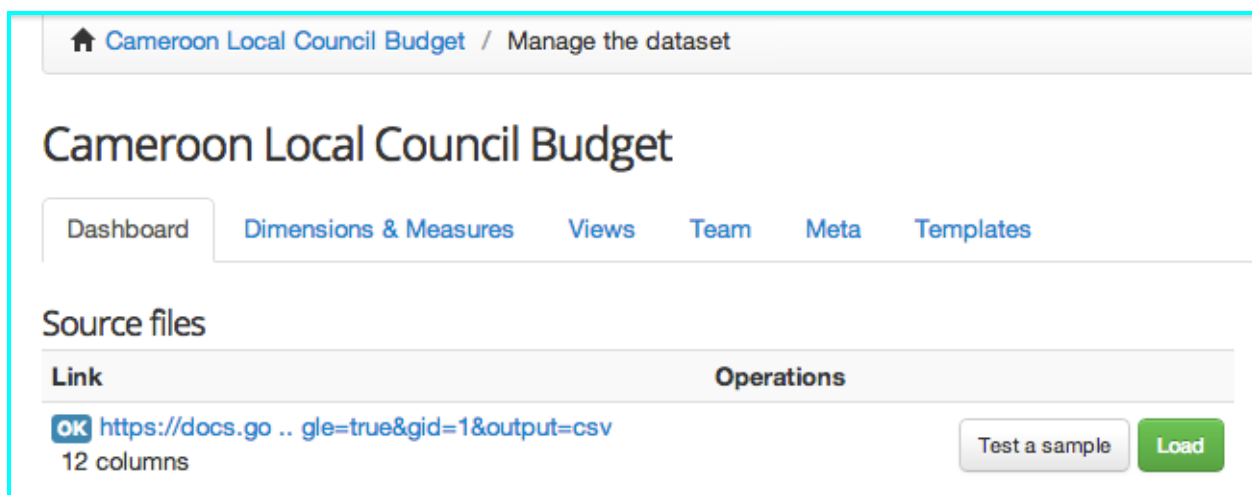
Si vous avez ajouté une mauvaise source, ce n'est pas grave. Vous n'avez pas à utiliser cette source pour votre jeu de données final. OpenSpending demande un travail supplémentaire sur les données avant de pouvoir les publier. Ajoutez alors simplement la source correcte, et oubliez l'autre source.

Modéliser vos données dans OpenSpending

Pour charger vos données dans OpenSpending, vous devez construire un modèle de vos données. Un modèle définit comment les données seront comprises par OpenSpending. L'application représente les propriétés des données en termes de *dimensions*. Modéliser les données consiste à lister les dimensions que vous souhaitez inclure dans le jeu de données OpenSpending et définir pour chaque dimension la ou les colonnes correspondantes dans le jeu de données d'origine.

Dimensions obligatoires : montant et date

Chaque modèle nécessite au moins deux dimensions : un montant et une date. Elles spécifient la taille de la transaction et le moment de sa réalisation. Le montant et la date sont associés à des indicateurs spécifiques. Le montant est représenté par un indicateur de mesure et le temps est représenté par une date. Des dimensions génériques ne peuvent pas représenter ces valeurs particulières.



The screenshot shows the 'Cameroon Local Council Budget' dataset management page. At the top, there is a breadcrumb trail: 'Cameroon Local Council Budget / Manage the dataset'. Below this, the title 'Cameroon Local Council Budget' is displayed. A navigation menu includes 'Dashboard', 'Dimensions & Measures', 'Views', 'Team', 'Meta', and 'Templates'. Under the 'Source files' section, a table lists a source file with the following details:

Link	Operations
https://docs.google.com/spreadsheets/d/1.../edit?usp=sharing 12 columns	Test a sample Load

Lorsque vous modélisez vos données, ce n'est pas une mauvaise idée de commencer par les dimensions obligatoires. Tout d'abord, cliquez sur le menu **Dimensions & Measures** au sein de la page **Manage the dataset** de votre jeu de données.

Après, cliquez sur **Add Dimension** pour faire apparaître le panneau *Add new dimension*. Cliquez sur la case correspondant à *Date*. Vous verrez la fenêtre *Name* automatiquement complétée avec la date, comme ci-dessous. Cliquez sur le bouton vert **Add**.

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 or .

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

Le prochain écran que vous verrez vous donnera des informations sur la signification du temps. Dans la liste déroulante à côté de *Column*, sélectionnez la colonne de vos données qui représentera la valeur du temps.

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.

Quand vous avez identifié la colonne du temps, cliquez sur **Add Dimension** encore une fois

pour ajouter le montant. Cette fois-ci, sélectionnez la case correspondant à *Measure*, qui sera automatiquement complétée sous le nom "montant" et cliquez sur **Add**. Choisissez la colonne représentant la valeur de la transaction dans la liste déroulante à côté de *Column*.

Autre dimension indispensable : l'identifiant

Il y a une autre dimension nécessaire au fonctionnement du modèle: la dimension (ou groupe de dimensions) dont la valeur identifie de manière unique chaque entrée de données, la *clé*.

Une entrée peut être identifiée aussi par une combinaison de plusieurs colonnes, une *dimension composée*. Étant donné que les clés peuvent être composées, le type "dimension composée" doit être utilisé pour les représenter, même si votre clé en l'occurrence n'est pas composée.

La dimension clé peut être assignée en cliquant sur **Add Dimension** et ensuite en sélectionnant le bouton radio *Dimension*.

Ajoutez le nom de votre clé, par exemple "clé", dans la case appropriée. Cliquez sur **Add**. Cochez la case *include in unique* afin d'identifier cette dimension en tant que partie de votre clé.

On passe maintenant à la liste de **Fields**, qui contient deux rangées nommées *name* (nom) et *label* (étiquette). Une dimension composée peut contenir un nombre arbitraire de champs (*fields*), chacun ayant un nom et un type et pouvant être associé à une colonne dans vos données. Cela nous permet d'expliquer le sens du mot "composé" dans ce contexte: les dimensions sont "composées" car elles regroupent plusieurs colonnes de données dans une seule propriété du jeu de données.

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="Identifiant"/>	<input type="text" value="id"/>	<input type="text" value="Default Value"/>
label	<input type="text" value="Identifiant"/>	<input type="text" value="float"/>	<input type="text" value="Default Value"/>

Une dimension composée nécessite au moins deux champs, *name* et *label*. Ces derniers doivent être de type *id* et *string*. Le nom de la dimension est utilisé afin de fournir à cette dernière une URL opérationnelle et l'étiquette (*label*) est utilisée pour présenter la dimension au sein de l'interface.

Pour créer une dimension composée minimale, il suffit d'associer la même colonne de vos données d'origine avec le nom (*name*) et l'étiquette (*label*). Choisissez la colonne appropriée pour chacun et laissez les options sous "types" inchangées.

Mesures et autres dimensions

Avec un montant, une date et un identifiant, votre modèle est suffisamment riche. Toutefois, un modèle parfaitement abouti devrait inclure une dimension pour chaque caractéristique importante du jeu de données d'origine. Suivre certaines conventions est alors utile.

	A	B	C	D
1	Head-account ▾	Head-Account Description ▾	Sub-account ▾	Sub-account Description ▾
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	Roceeds from global taxes
7	710.xxx	Fiscal revenue (1)	710.1	Roceeds from global taxes
8	710.xxx	Fiscal revenue (1)	710.1	Roceeds from global taxes
9	710.xxx	Fiscal revenue (1)	710.1	Roceeds 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

Une pratique courante dans la présentation des jeux de données d'origine est la segmentation de l'information qui caractérise chaque donnée sur de multiples colonnes. L'information concernant un compte associé à une transaction peut être réparti entre une colonne "Compte" qui identifie le montant chiffré et une colonne "Description du compte" qui donne une description écrite. L'image ci-dessous illustre cette pratique avec les colonnes "Head-account" et "Sub-account".

Les dimensions composées d'OpenSpending sont conçues pour modéliser ce type d'informations éparpillées. Pour ce faire, ajoutez une nouvelle dimension composée et associez chaque colonne à l'un des champs de la dimension. Essayez de faire correspondre une colonne aux commentaires très détaillés à *label* et une colonne plus concise à *name*. Dans l'image ci-dessous "Head-account" correspond à *name* et "Head-account description" à *label*.

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"/>

Certaines colonnes de votre base de données sont plus indépendantes, représentant certaines propriétés particulières de chaque donnée. Par exemple, une colonne qui attribue à chaque transaction une catégorie rentre dans ce cadre. Dans l'image ci-dessous, les colonnes *Reporting Type*, *Revenue/Expenditure* et *Recurrent/Investment* sont de ce type.

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

Les colonnes indépendantes qui spécifient des propriétés ou des catégories sont mieux modélisées avec des dimensions d'attribut. Un attribut est essentiellement une dimension qui ne connaît qu'un seul champ, quel que soit son type. Pour créer un attribut, sélectionnez simplement la case *Attribute* quand vous ajoutez une dimension.

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.

Pour finir : sauvegarder et charger

Lorsque chaque dimension a été spécifiée et reliée aux colonnes dans les données source, cliquez sur **Save Dimensions** pour sauvegarder votre modèle. En cas d'erreurs, un message apparaîtra, vous demandant de corriger certains paramètres. S'il n'y a pas d'erreurs, vous serez redirigé vers le Dashboard, où vous pourrez charger vos données.

Une fois que les données ont été téléchargées, le modèle que vous avez créé sera figé et vous n'aurez plus la possibilité de l'éditer. C'est pourquoi il est préférable de tester le modèle avant de télécharger les données. Pour ce faire, cliquez sur **Test a sample** dans votre tableau de sources de données. Attendez quelques secondes et rechargez la page. Si vous voyez un message indiquant COMPLETE sur un fond vert, alors votre modèle est prêt. Si vous voyez un message indiquant ERRORS, des corrections sont nécessaires.

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 votre modèle ne contient plus d'erreurs, cliquez sur **Load** pour charger le jeu de données source et lui appliquer le modèle. Vous pouvez ensuite retourner sur la page d'accueil du jeu de données en cliquant sur son nom en haut de l'écran et créer des visualisations ou explorer le jeu de données.

Visualisation

Créer une visualisation

La plateforme OpenSpending permet de créer facilement des visualisations de données. Trois visualisations sont proposées : le BubbleTree, la TreeMap et la Table des agrégats.

Toutes les modélisations d'OpenSpending vous permettent de choisir une série de dimensions grâce auxquelles vous pouvez agréger vos données et les segmenter selon des propriétés de plus en plus précises. Chaque modélisation est créée de la même manière : en choisissant les dimensions à agréger et l'ordre dans lequel les segmenter.

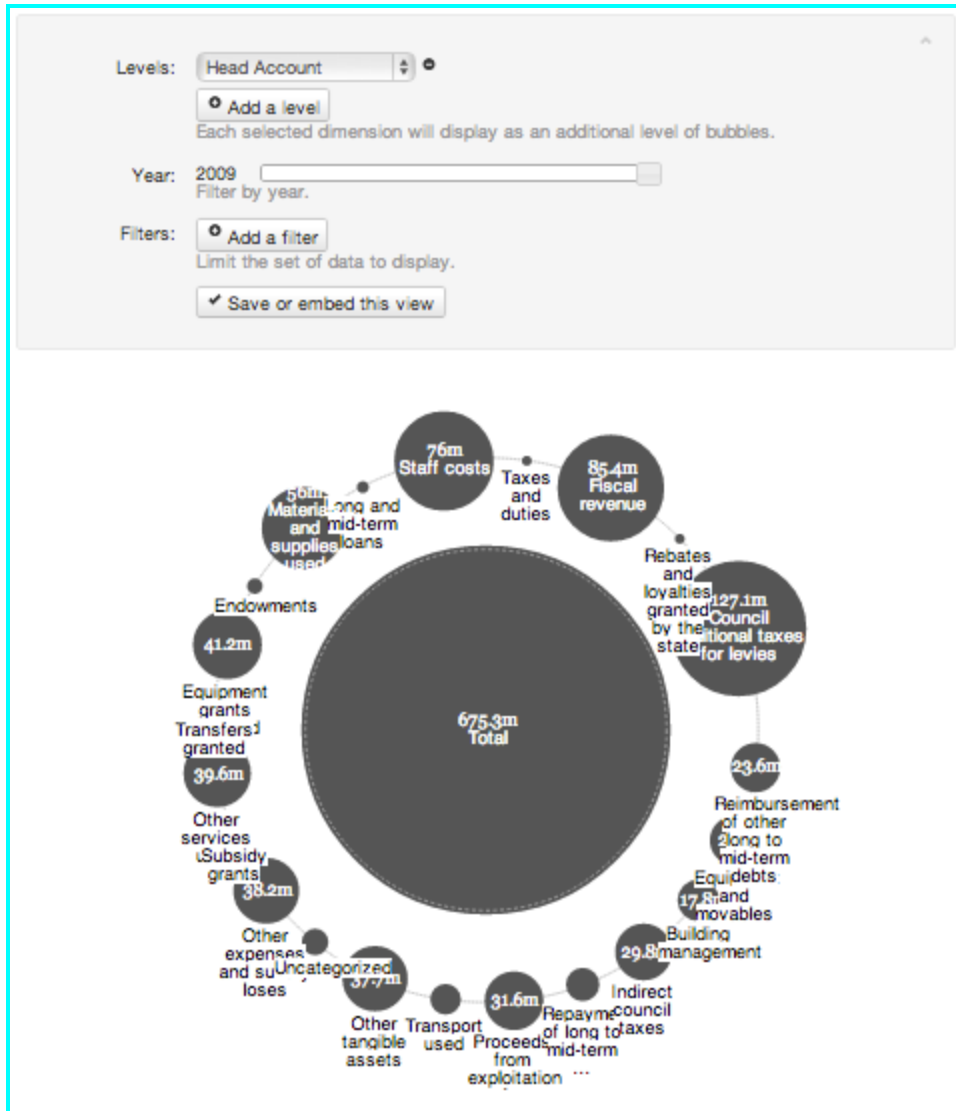
Pour commencer à créer une modélisation, allez dans la page d'accueil du jeu de données et sélectionnez **Create a visualisation** dans le menu *Visualization*.

BubbleTree

Le BubbleTree est une visualisation interactive qui présente les données de dépense agrégée sous forme de bulles. Chaque bulle représente un (sous-)total agrégé. La bulle centrale représente la somme agrégée et les bulles qui l'entourent représentent les autres sommes qui la composent. En cliquant sur n'importe quelle bulle, l'utilisation voit la

manière dont la somme se divise en d'autres sous-totaux.

Pour créer un BubbleTree, choisissez les dimension d'agrégation et l'ordre dans lequel vous voulez les agréger. Choisissez la première dimension dans la liste déroulante *Level*. Vous verrez alors le total agrégé de cette dimension sous la forme de la bulle centrale, avec les valeurs des dimensions qui l'entourent et leurs propres totaux.



Pour ajouter un deuxième niveau, cliquez sur **Add a level** et choisissez une nouvelle dimension. Les utilisateurs seront alors capables de cliquer sur les bulles pour les segmenter et voir comment les valeur du premier niveau se divisent dans les valeurs du second niveau.

Levels:

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

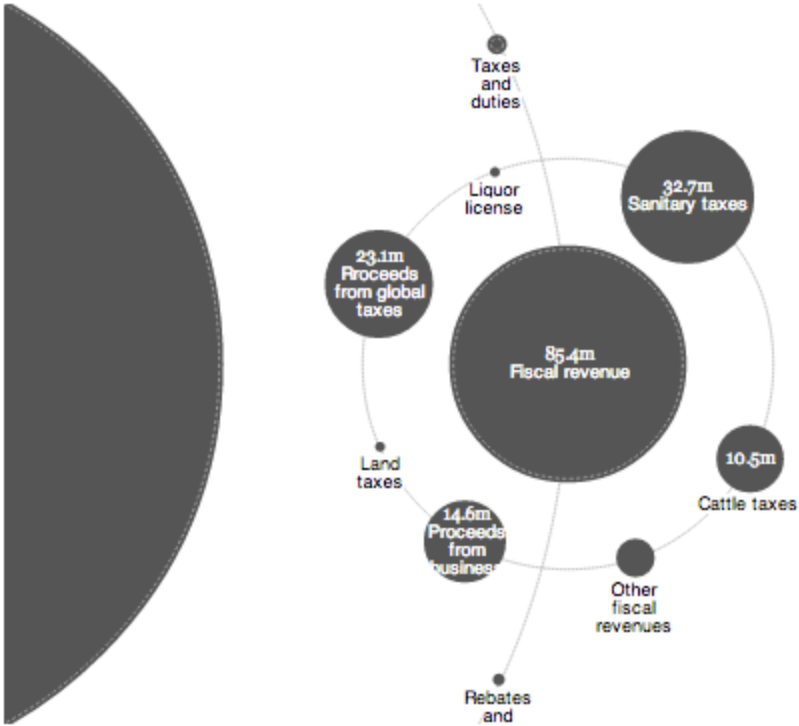
Year: 2009

Filter by year.

Filters:

Limit the set of data to display.

Save or embed this view



TreeMap

Le TreeMap représente les données de dépense agrégées sous forme de rectangle de couleur. Chaque rectangle représente un montant agrégé selon une dimension du jeu de données. Cliquer sur un rectangle permet d’explorer la répartition des sous-totaux suivant une autre dimension.

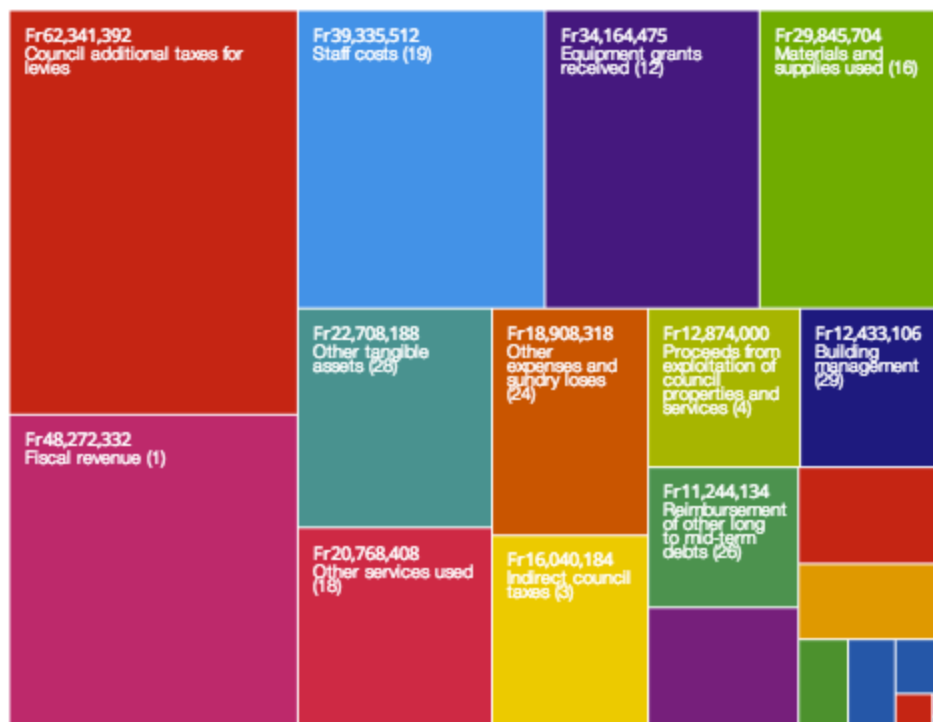
Pour créer un TreeMap, choisissez simplement les dimensions à agréger et leur niveau. En sélectionnant la dimension de premier niveau vous verrez apparaître la répartition des montants dans les rectangles suivant la dimension choisie.

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



Pour rendre la visualisation plus interactive, ajoutez des niveaux supplémentaires qui vous permettront d'explorer comment chaque montant se divise en sous-montants. Pour ajouter un second niveau, cliquez sur **Add a level** et choisissez une nouvelle dimension. Les utilisateurs peuvent désormais cliquer pour voir comment les montants se divisent.

Tableau d'agrégats

Le tableau d'agrégats est un simple tableau qui offre une vue des données numériques agrégées (additionnées) suivant la dimension choisie.

Sélectionnez une dimensions dans le menu *Columns*. Vous verrez alors apparaître un tableau avec les montants et le pourcentage du montant total qu'ils représentent. Par défaut, les lignes sont triées par pourcentage décroissant.

Columns:

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

Year:

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 loses (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 →

En ajoutant une autre colonne en cliquant sur **Add a level**, le montant de la première colonne sera divisé en sous-totaux dans la seconde suivant la dimension sélectionnée. Notez que cette opération modifie les valeurs des pourcentages et donc l'ordre des lignes.

Inclure une visualisation sur votre site web

Il est très simple d'insérer dans votre site Web des visualisations créées grâce à OpenSpending. Votre site bénéficiera alors des visualisations interactives complètes.

Imaginons que vous ayez choisi une visualisation particulière sur la plateforme OpenSpending. Remarquez le bouton **Embed** en bas à droite de la page. En cliquant sur ce bouton, vous accédez au code nécessaire à l'inclusion de la visualisation sur votre site Web, ainsi qu'à quelques options concernant la taille (en pixels) de la visualisation interactive. Il ne vous reste qu'à copier et coller le code dans votre site. Si vous n'êtes pas sûr de savoir comment coller le code correctement, contactez l'administrateur de votre site.

L'inclusion de visualisations est rendue possible par les *widgets*. Ces morceaux de code

que vous pouvez inclure dans une page Web récupèrent à distance des données — dans notre cas, depuis la base de données OpenSpending — sans que vous ayez à stocker vous-même ces données.