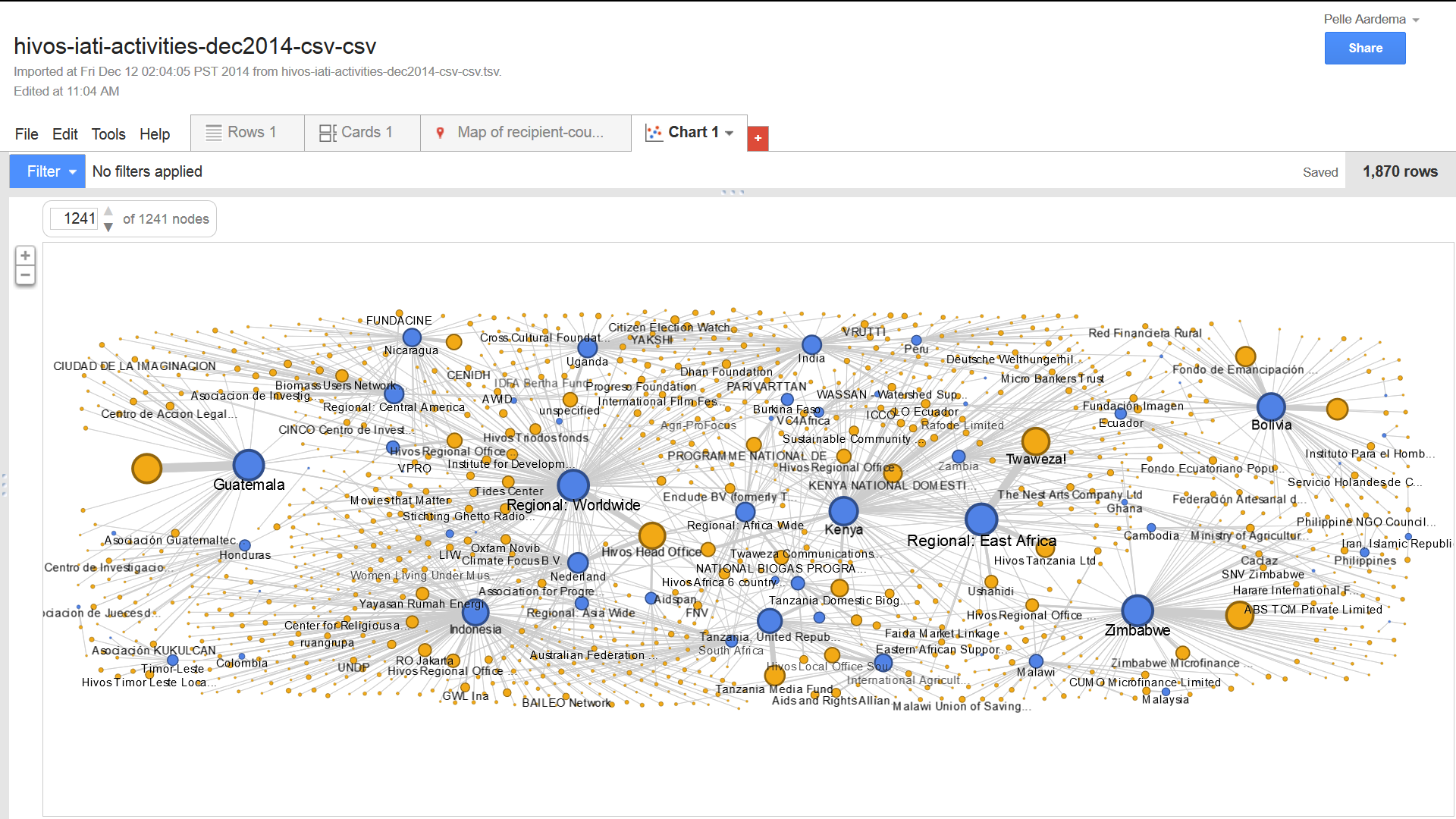
Recipe: Spaghetti à la Hivos

Recipe: Spaghetti à la Hivos

Showing Hivos’ partner network using IATI data

Showing Hivos’ partner network using IATI data

# Begin with the end in mind

Think of what you’d like to show, what kind of data do you need to make that possible and check if you have the right data.

In this case we want to show the network of partners Hivos works with, and get an impression of the budgets they work with.

# Get your ingredients (the data)

We need to check if Hivos includes implementing partners in its IATI data set.

1. Go to the IATI Registry and download the latest Hivos IATI file.

*It’s always a good idea to open the file in a text editor (for instance Notepad++), to see what’s in it.* [*http://notepad-plus-plus.org/*](http://notepad-plus-plus.org/)

We look at the data to see if there are elements like below – which is the case.

<participating-org role="Implementing">...</participating-org>

For a further analysis of the data, we’ll transform it to a table format (CSV) first.

1. Download EditiX free edition (<http://www.freexmleditorsite.com/download.html>) and install it.
2. Download XSLT transformations here: <https://github.com/IATI/IATI-XSLT> and unzip the library to a place where you can easily find it again.
3. Transform the file to CSV using XSLT:
   * Open EditiX
   * Open the Hivos .xml file
   * Choose XSLT  Transform using XSLT
   * In the box ‘XSLT document’ browse to the library you’ve just unzipped, and on to templates  csv. Now choose iati-activities-xml-to-csv.xsl
   * Choose a location for the resulting file. Make sure to set the File type to ‘All files’ and end your filename with “.csv”
   * Click “OK”

And then we’ll open the data set in OpenRefine (*a power tool for working with messy data*)

1. Download OpenRefine from <http://openrefine.org/download.html> and install it.
2. When you run Open Refine, it will open in your browser, and let you create a project by importing data from your computer. Open the .csv file you’ve just created in step 4 and click next.
   * Make sure you set the field "Character encoding" to UTF-8
   * Choose “Columns are separated by Commas (CSV)”
   * You will see a preview of the data. In the top-right corner, click “create project” to finish importing the data.
3. Now let’s analyse the data Hivos has included in their IATI file. You can easily do this using facets and filters.
   * At the top of each column you’ll find an arrow pointing downward that offers you a range of powerful options to analyse and edit the data in that column.
   * For instance: go to the column with recipient-countries
   * Choose Facet  Text facet
   * On the left hand side, you’ll now see a box showing all the different values in this column

Notice there are 59 different values. There is only 1 blank cell.

And there are different ‘Regional: …’ values included in this column.

1. Looking through the data set we see a couple of things:
   * All the recipient-country entries contain a single value
   * There are only implementing partners in the data set (no funding or extending partner organizations)
   * All activities have a budget (except one)
   * Budgets are given in different currencies: EUR (1599), USD (204), GBP (65) and KES (1)
   * 185 activities don’t contain any transaction data.
   * 1401 activities contain sector information (classified using DAC-3), 469 don’t contain sector information.
2. In this visualization, we want to show network of organisations Hivos works with. Given our analysis we’ll be able to show this network per country, or we can link organizations to the projects they work on. But we’ll make sure to include more fields so that we keep some more options open.
   * Nodes should have a weight (which determines their size in the eventual visualization), for instance based on project budget.
   * We'll save fields such as 'recipient-country' and 'sector' so that we can show the visualization for a specific country or sector.
   * We'll also save some other additional information per partner-project relation, for other possible visualizations.

# Shaping & cleaning your data

Now we're going to shape the data, using our “power tool” OpenRefine.

1. First we’re going to reorder the columns in the data set. Go to the first column (All), click on the arrow and choose ‘Edit columns’  ‘Reorder / remove columns’
   * Reorder or remove a few columns and click “OK”
   * Notice how in the top-left corner the Undo / Redo tab now shows (1)
   * When you click the tab, you’ll see all the ‘transformations’ you’ve made so far. Here you can easily undo and redo your steps.
   * Also notice the buttons “Extract” and “Apply”. Here you have access to the code that OpenRefine produced to execute your transformation. You can copy this code and use it in future projects.
2. Go back to how the original data was by clicking on “0. create project”. The columns you removed are back again. Now click the “Apply...” button, and copy and paste the code below to reorder the columns in the data set, keeping the columns we want to use & save as extra information.

[

{

"op": "core/column-reorder",

"description": "Select columns we need",

"columnNames": [

"iati-identifier",

"titles",

"participating-orgs\_implementing",

"recipient-countries",

"sectors",

"sector\_vocabularies",

"sector\_codes",

"budget\_values",

"budget\_value\_currencies"

]

},

1. Make sure there are no empty values

Why is this important? From the Schoolofdata.org course on “*Cleaning data with spreadsheets*”: (<http://schoolofdata.org/handbook/recipes/cleaning-data-with-spreadsheets>):

*“In many spreadsheets you come across there will be empty (“blank”) cells. They may have been left blank intentionally, or in error. Either way, they are missing data, and it’s useful to be able to find, quantify, display and correct them if needed.*

*You may want to actively specify that the data is missing, rather than leaving a blank cell. - Filling blank cells isn’t always useful and it’s important to choose the right term to denote a missing value.”*

Not all visualization tools know how to handle empty values, or empty values may lead to misrepresentations, so we’ll replace all the empty values in the columns we are going to use.

* + For instance: select the column sectors
  + Open a Text Facet
  + Scroll down to the (blank) value, and click ‘edit’ (appears when you mouse over the value)
  + Change the value to ‘unspecified’.

The code below fills all the empty values in the columns we need for our visualization:

{

"op": "core/mass-edit",

"description": "Replace blanks in column participating-orgs\_implementing",

"engineConfig": {

"facets": [],

"mode": "row-based"

},

"columnName": "participating-orgs\_implementing",

"expression": "value",

"edits": [

{

"fromBlank": true,

"fromError": false,

"from": [],

"to": "unspecified"

}

]

},

{

"op": "core/mass-edit",

"description": "Replace blanks in column recipient-countries",

"engineConfig": {

"facets": [],

"mode": "row-based"

},

"columnName": "recipient-countries",

"expression": "value",

"edits": [

{

"fromBlank": true,

"fromError": false,

"from": [],

"to": "unspecified"

}

]

},

{

"op": "core/mass-edit",

"description": "Replace blanks in column sectors",

"engineConfig": {

"facets": [],

"mode": "row-based"

},

"columnName": "sectors",

"expression": "value",

"edits": [

{

"fromBlank": true,

"fromError": false,

"from": [],

"to": "unspecified"

}

]

},

{

"op": "core/mass-edit",

"description": "Replace blanks in column budget\_values",

"engineConfig": {

"facets": [],

"mode": "row-based"

},

"columnName": "budget\_values",

"expression": "value",

"edits": [

{

"fromBlank": true,

"fromError": false,

"from": [],

"to": "unspecified"

}

]

},

1. Because our CSV-transformation ‘flattened’ the dataset – and may have put multiple implementing partners in the same cell (separated by ‘;’) – we now want to split those again. Again, OpenRefine comes to the rescue.

* Go to the column ‘participating-orgs\_implementing’
* Click on the arrow
* Select ‘Edit cells’  ‘Split multi-valued cells’
* Multiple values are currently separated by ‘;’.
* Click ‘OK’.
* Did you notice that the number of rows hasn’t changed? So (if you were using the Hivos IATI data set) no rows have been split.

1. The code below will split multi-valued cells in all the columns containing partner organizations.

{

"op": "core/multivalued-cell-split",

"description": "Split multi-valued cells in column participating-orgs\_implementing",

"columnName": "participating-orgs\_implementing",

"keyColumnName": "iati-identifier",

"separator": ";",

"mode": "plain"

},

In **step 8** we already concluded that Hivos included only implementing partners in their IATI data. Also, there were no cells that contained multiple values. So our data set should be in good shape now.

1. Export your data from OpenRefine. Choose ‘Export’  ‘Tab separated value’.

**On to the visualization!**

In this recipe we’ll be using Google Fusion Tables to visualize the network.   
(You need a Google account in order to use this tool. If you haven’t got one, set it up now)

1. Open the resulting file in Google Fusion tables. In Google Drive, choose ‘New’  ‘More’  ‘Google Fusion Tables’. Now select the file you just exported, and follow the wizard.   
   *NOTE: If this is the first time you use Google Fusion tables, click on ‘Connect more apps’ and search for Google Fusion.*
2. You'll see a few tabs (including a “Map of recipient-countries”: if you click on this, Fusion Tables will take some time to look up all countries). We're going to do something different first. Click on the red + sign, choose 'Add chart' and select 'Network Graph' at the bottom of the column on the left.
3. Now set the values that should be shown:
   * **Show link between**: 'recipient-countries' & 'participating-orgs-implementing'
   * **Appearance**: select 'Color by columns' – this will give the countries and organizations a different color.
   * **Weight by**: 'budget-values' – you will now see something like “0 of 1241 nodes” in the top-left corner of the graph. If you click the up arrow, you will see a preview with 10 nodes.

Click “Done” in the top-right corner to finish this step.

Now you have a network graph showing all the partners Hivos works with in different countries. All countries are shown in blue, the organizations are shown in orange. The size of the circles (and the weight of the links) is determined by budget size.

*NOTE: Make sure you show all the nodes using the box in the top-left corner.*

1. Click on the blue 'Filter' button, and add filters for 'recipient-countries'
2. Drag the nodes around until the network ‘looks nice’
3. You can embed this visualization on any website, or you can invite others to look at it too. Choose ‘Tools’  ‘Publish’ and change the privacy settings. You’ll now find a link you can share and a code to embed the chart.

Congratulations!

You have just made your first (?) interactive network chart!

