

# Introducción a la visualización de datos

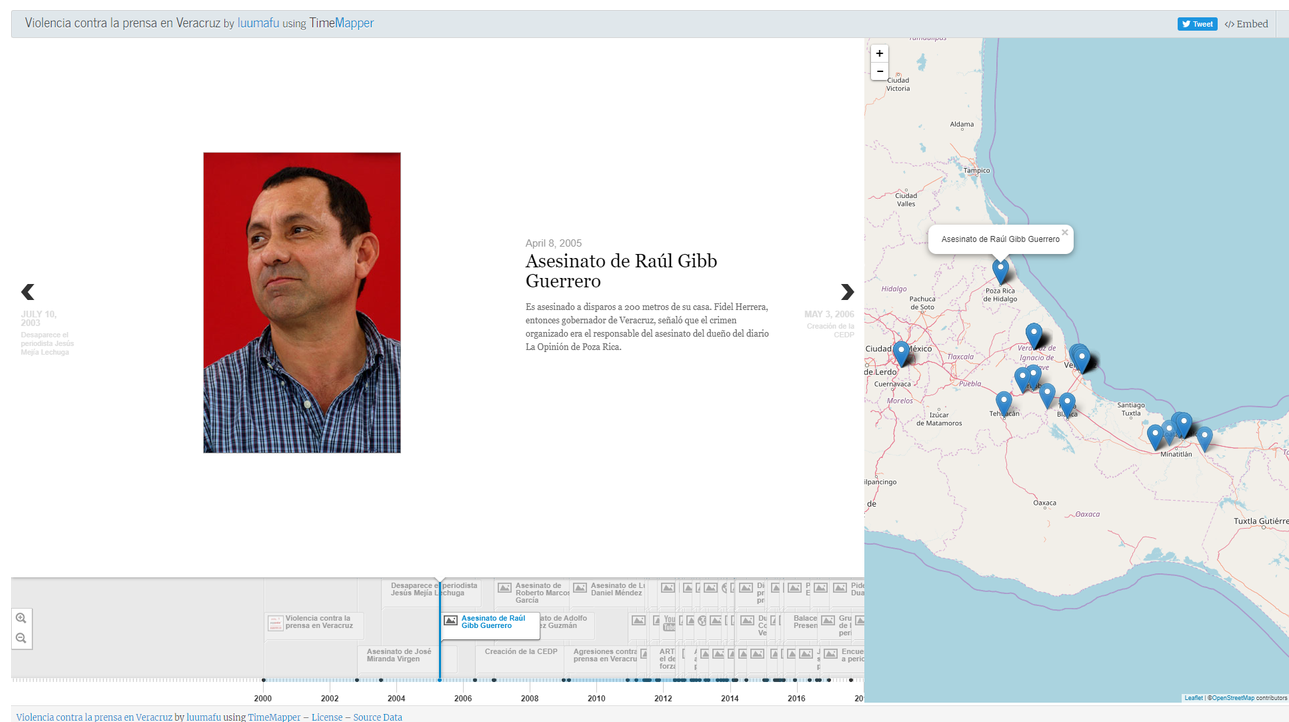
Elaborado por: Silvia Gutiérrez De la Torre (Licencia: CC-BY-4.0)

Contacto: segutierrez@colmex.mx

## Herramienta 1. TimeMapper

### Parte 1. Conociendo la herramienta

1. En esta sección aprenderemos a usar una herramienta de crono-mapa desarrollado por la Open Knowledge Foundation (una ONG que aboga por los datos y contenidos abiertos)
2. Nosotros usaremos los datos preparados por Lucía Vergara, Coordinadora multimedia en Mexicanos contra la Corrupción: <http://bit.do/timemapperdatos>
3. Estos datos son los que están detrás de este mapa: <http://timemapper.okfnlabs.org/luumafu/violencia-prensa-veracruz>



Antes de poner manos a la obra, observa:

- a. ¿Qué podemos aprender de una visualización así?
- b. ¿Qué tipo de datos están siendo visualizados?

## Parte 2. Aprender modificando

1. Entrar a la página de TimeMapper: <http://timemapper.okfnlabs.org/create>
2. Opcional: crear una cuenta (así se guardan todos tus mapas)
3. Normalmente se toma la tabla muestra de TimeMapper ([Click here to use a pre-prepared example](#) ») y se quitan los datos pre-cargados pero usaremos los datos de Vergara.
4. Las columnas más importantes son: título, descripción, año de inicio o “start” (puedes poner uno ficticio y arrastrar con control más mouse hacia abajo), lugar y “location”
5. Analiza qué está contenido en cada una de estas columnas.

### Ejercicio:

1. Buscaremos repetir el ejercicio de Vergara, visibilizando ahora los asesinatos de periodistas en todo el territorio mexicano durante 2017.
2. Para esto tomaremos como base la lista publicada por El Universal: <http://www.eluniversal.com.mx/nacion/sociedad/los-11-periodistas-asesinados-en-mexico-en-2017>
3. Llenaremos todos los campos con información de fuentes confiables (cada quien llenará la información sobre un periodista) en la siguiente tabla colaborativa: <http://bit.do/violenciaprensa2017>
4. ⚠ Antes de empezar, noten que la columna “location” tiene una fórmula:

```
=JOIN(", ", ImportXML(CONCATENATE("http://nominatim.openstreetmap.org/search/?format=xml&q=", F2), "//place[1]/@lat | //place[1]/@lon"))
```

Además de las funciones básicas de una hoja de cálculo (JOIN, CONCATENATE), ésta fórmula utiliza una de las múltiples funciones especiales de GoogleSheets: [ImportXML](#), la cual importa varios tipos de datos [semi]estructurados (XML, HTML, CSV, etc).

Su sintaxis es:

```
IMPORTXML(url; consulta_xpath)
```

En este caso, de la respuesta a una consulta a OpenStreetMap (&q=?) que se llena con lo que escribamos en F2 y luego una extracción de una porción de los datos usando XPath:

```
//place[1]/@lat  
//place[1]/@long
```

Observen que he llenado la primera fila con la información “Guerrero, México”. Lo que hace la fórmula en es:

- a) Hacer la consulta <http://nominatim.openstreetmap.org/search/?format=xml&q=Guerrero%20Mexico>
- b) Tomar al primer hijo “place” (place[1]) y a su atributo lat y long

Ahora que han entendido cómo funciona el “geocoding” empiecen a llenar cada uno de los rubros.

<http://timemapper.okfnlabs.org/espejolento/violencia-prensa-2017>

### Parte 3. Compartir y divulgar

1. Para visualizar la tabla en TimeMapper tenemos que, antes, hacerla pública.
2. Para esto se va a File >> Publish to web >> Publish
3. Después, en “compartir documento” se copia el hipervínculo en la sección 2 de la página de crear mapas de TimeMapper (ver paso Parte 2.1)

## 2. Connect and Customize

ALERT: you are **not signed in** so your timemap will be created '**anonymously**'.

If want to 'own' your timemap you should [sign in \(or sign-up\)](#) now »  
(Sign-up takes a few seconds with your [twitter account](#) »)

Find out more on anonymous vs logged in – [read FAQ below](#) »

Data Source



Select from Your Google Drive

(If nothing happens check you are not blocking popups ...)

Or paste the Google Spreadsheet URL directly

**Important:** you must "publish" your Google Spreadsheet: go to File Menu in your spreadsheet, then 'Publish to the Web', then click 'Start Publishing'. See [the FAQ below](#) for more details

4. Se rellenan los parámetros según la visualización deseada
5. Se pulsa sobre "Publish" ¡y listo!

## Herramienta 2. Google Fusion Tables

1. Entra a la página

<https://fusiontables.google.com/DataSource?dsrclid=implicit&pli=1>

2. Descargar datos para la sesión: Proyectos de Obra Pública Abierta 2017

<https://datos.gob.mx/busca/dataset/obra-publica-abierta>

3. Primero, a explorar:

- a. Observa la página de transparencia presupuestaria:

[http://www.transparenciapresupuestaria.gob.mx/en/PTP/Obra\\_Publica\\_Abierta](http://www.transparenciapresupuestaria.gob.mx/en/PTP/Obra_Publica_Abierta)

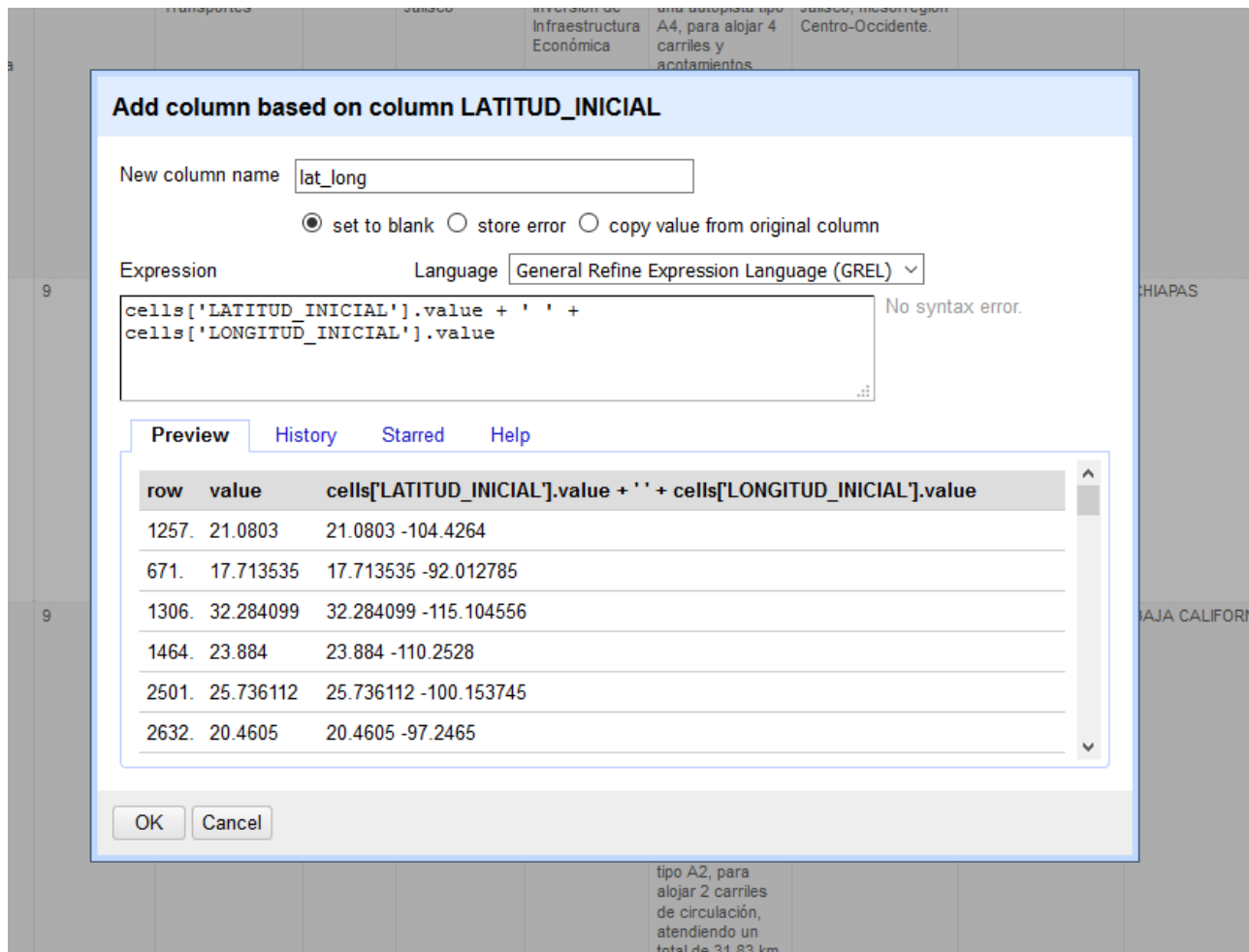
- b. Ahora compara lo que puedes aprender ahí con lo que descubres al abrir los datos en GoogleRefine:

- i. Tipos de Programas y Proyectos de Inversión (PPI)

- ii. Entidades Federativas

- iii. Administradores (¿qué rubro nos permitiría identificarlos mejor?)


- iv. Montos (¿qué podríamos hacer para comprender mejor estos datos?)
- 4. Después, a transformar:
  - a. Más adelante usaremos las coordenadas, para esto hay que





- 5. Exportar nueva tabla desde GoogleRefine
- 6. Abrir datos en fusion tables

## Import new table



 From this computer

 Google Spreadsheets

 Create empty table


Seleccionar archivo proyectos\_opa.csv

Separator character ☒ Comma ☐ Tab ☐ Colon ☐ Other

Character encoding UTF-8 ▼

You can upload spreadsheets, delimited text files (.csv, .tsv, or .txt), and Keyhole Markup Language files (.kml) [Learn more](#)

Or search public data tables



### New to Fusion Tables?

Take a peek! [Play with a data set](#) or [try a tutorial](#).

Cancel

« Back

Next »

## 6. Añadir metadatos

Import new table

Table name

proyectos\_opa-csv

Allow export

☒ ?

Attribute data to

SHCP ?

Attribution page link

<https://datos.gob.mx/busca/d>

Description

Contiene información de los Programas y Proyectos de Inversión de la Obra Pública Abierta para el año fiscal 2017, además brinda datos para poder revisar su seguimiento y el detalle de cada Proyecto o Programa de Inversión.

For example, what would you like to remember about this table in a year?

New to Fusion Tables?

Cancel

« Back

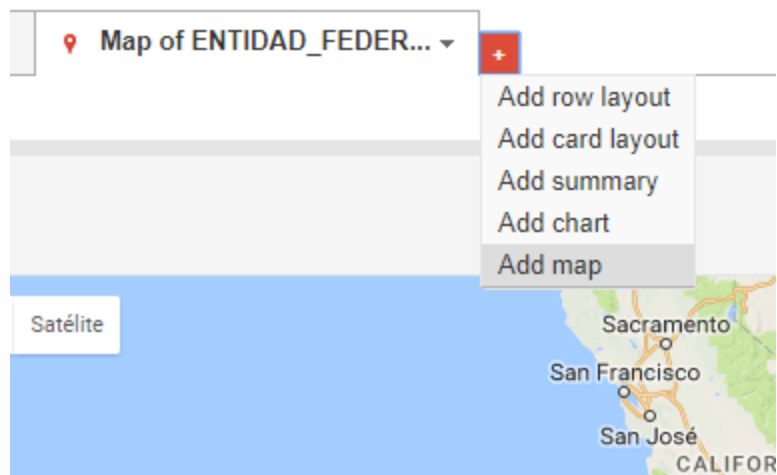
Finish

Take a peek!

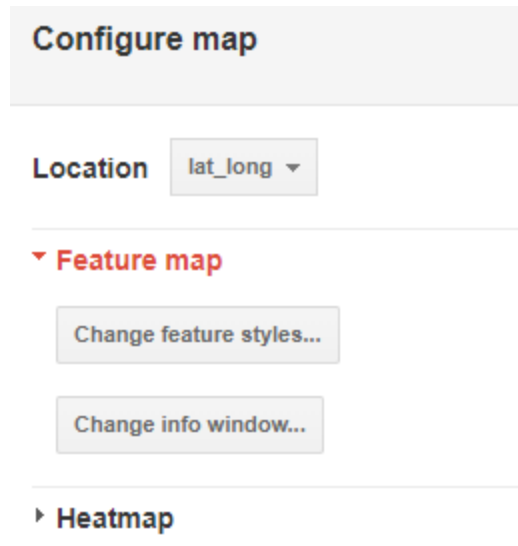
[Play with a data set](#) or [try a tutorial](#).

## 6. Manos a la obra I: Mapear los datos

- Añadir un nuevo mapa → <https://www.google.com/fusiontables/DataSource?docid=1kxNuDXTvHqqCsf79vdAMpRLcWfCdJDpRDd3UawVz>



b. Crearlo basado en nuestra columna de lat\_long

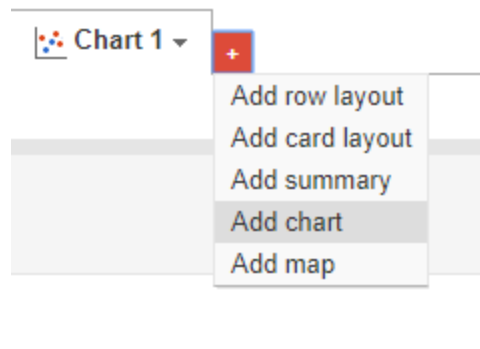


c. Cambiar visión a HeatMap

- i. observar las opciones “radius” y “opacity”
- ii. cambiar el “weight” a “Monto\_Total\_Inversion”, después a “Total\_Gasto\_No\_Considerado”

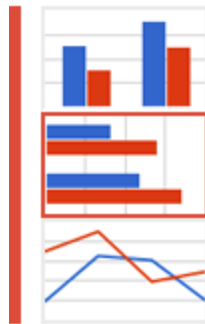
## 8. Manos a la obra II: Graficar los datos

a. Añadan ahora una “chart”



b. Seleccionen la opción de gráficos de barras





c. Seleccionen los siguientes parámetros:

**Category**

ENTIDAD\_FEDERATIVA ▾

**Summarize data?** ☒

**Values**

ENTIDAD\_FEDERATIVA  
☒ count

TOTAL\_GASTO\_OPERACION\_HE  
☐ min ☐ max ☐ avg ☐ sum

CVE\_PPI  
☐ min ☐ max ☐ avg ☐ sum

Choose... ▾

**Sort by** [reverse](#)

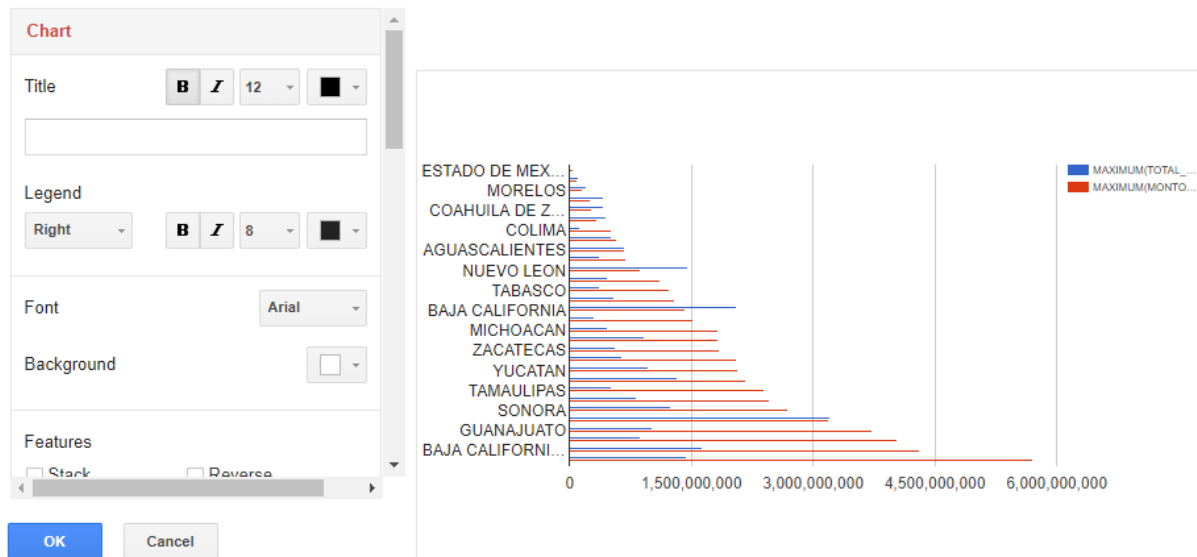
Count ▾

**Maximum categories**

30

- i. ¿A qué creen que se refiere la opción “summarize data?”
- ii. ¿Por qué Entidad\_Federativa tendrá la opción de “count”?
- iii. ¿Para qué sirve cambiar el número de “Maximum categories”?
- iv. Cambien el formato de la gráfica en el editor

## Chart Editor



v. Interpreten la gráfica resultante

vi. Cambien la categoría a Tipo PPI y exploren

## 9. Manos a la obra III: Graficar en red los datos

a. Den clic en la última opción que es un gráfico de redes

b. Seleccionen los siguientes parámetros:

### Configure network graph

Show link between

ENTIDAD\_FEDERATIVA ▼

NOMBRE ▼

Appearance

☐ Link is directional

☒ Color by columns

Weight by

MONTO\_TOTAL\_INVERSION ▼

- i. ¿qué creen que significa “link is directional”?
- ii. ¿a qué creen que se hace referencia con “weight by”?