



Con el apoyo de:



## Herramientas para hacer Periodismo de datos

Elaborado por: Hassel Fallas @HasselFallas

Con este manual doy respuesta a una pregunta recurrente entre periodistas y alumnos de periodismo:

- ¿Necesito un programador y un ingeniero para hacer Periodismo basado en análisis de datos?
- No siempre.

Veamos porqué en este paso a paso por una serie de herramientas que te recomiendo aprender.

Usaremos como ejemplo esta publicación:

[¿Qué buscan los medios a la hora de contratar a un periodista de datos?](#)

<http://hasselfallas.com/2016/11/20/esto-buscan-los-medios-a-la-hora-de-contratar-a-un-periodista-de-datos/>

Los datos que debemos de raspar (scrap) están en estos vínculos:

- <http://datajournalismjobs.com/>

### Scrapping básico

- Descargar Outwit, la version gratuita de Outwit hub (<http://www.outwit.com/>)



↓ Download Application (free)

- Abrir Outwit hub

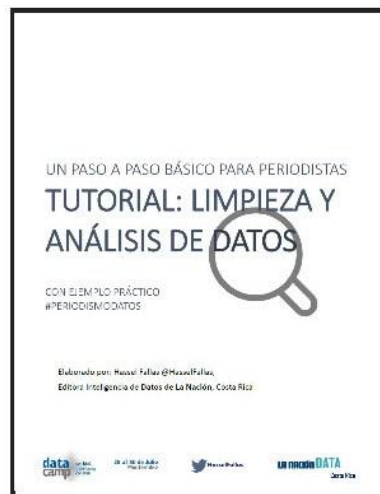
- Introducir la dirección: <http://datajournalismjobs.com/>
- <https://web.archive.org/web/20161003021052/http://datajournalismjobs.com/>
- Editar
- Exportar como EXCEL
- Abrir el archive y limpiarlo

### Scrappear imágenes

- Descargar Outwit images, desde <http://www.outwit.com/>
- Abrir Outwit images
- Abrir el link <http://icom.museum/recursos/banco-de-datos-de-las-listas-rojas/categoria/figuras-y-vasijas-jama-coaque/L/1/>
- Arrastrar las imágenes
- Salvar

### Excel

- ¿Cómo trabajar con Excel? Descárgate esta guía <http://hasselfallas.com/wp-content/uploads/2015/07/tutorial-limpieza-y-anc3a1lisis-de-datos-bc3a1sico-hassel-fallas1.pdf>



### Access

[Aquí hay un curso básico y útil](#) para hacer preguntas y cruzar tablas con datos

Símbolo	Palabras
>	mayor que
<	menor que
≥	mayor o igual que
≤	menor o igual que

## Simple criteria for all data types:

	Criteria Name	Write it like...	Function
1	Equals	"x"	Searches for values equal to x
2	Does Not Equal	Not in ("x")	Searches for all values except those equal to x
3	Null	Is Null	Searches for empty fields
4	Not Null	Is Not Null	Searches for non-empty fields

## Simple criteria for text:

	Criteria Name	Write it like...	Function
1	Contains	Like "*x*"	Searches for all values that contain x
2	Does Not Contain	Not like "*x*"	Searches for all values except those that contain x
3	Begins With	Like "x*"	Searches for all values beginning with x
4	Ends With	Like "*x"	Searches for all values ending with x
5	Comes After	>= "x"	Searches for all values that come after x in alphabetical order
6	Comes Before	<= "x"	Searches for all values that come before x in alphabetical order

## Simple criteria for numbers:

	Criteria Name	Write it like...	Function
1	Between	Between "x" and "y"	Searches for values in the range between x and y
2	Less Than	< x	Searches for all values smaller than x
3	Less Than or Equal To	<= x	Searches for all values smaller than or equal to x
4	Greater Than	> x	Searches for all values larger than x
5	Greater Than or Equal To	>= x	Searches for all values larger than or equal to x

## Simple criteria for dates:

	Criteria Name	Write it like...	Function
1	Between	Between "#mm/dd/yyyy#" and "#mm/dd/yyyy#"	Searches for dates that fall between two dates
2	Before	<#mm/dd/yyyy#	Searches for dates before a certain date
3	After	>#mm/dd/yyyy#	Searches for dates after a certain date
4	Today	=Date()	Searches for all records containing today's date
5	Days Before Today	<=Date()-x	Searches for all records containing dates x or more days in the past

### R (software para análisis estadístico)

- Descárgalo desde:
- <https://www.r-project.org/>
- Eliges la opción PAQUETES
- Descargar paquetes
- Eliges CRAN Mirror cercano a tu país de origen
- Luego eliges la librería: Rcmdr

Una vez instalado digitas en la consola: library (Rcmdr)

Desde la opción fichero debes de establecer el Directorio donde se encuentra el archivo CSV (en mi caso es el escritorio)

Desde el menú datos/ importar / desde archivo de datos o portapapeles

Eliges la opción valores separados por **;** (punto y coma)

En decimales digitas **,** (coma)

Cargas el archivo

\*(Es muy importante que el archivo que importes esté en formato CSV y que en lugar de espacios vacíos aparezca la advertencia NA)

- Estadísticos/Conjunto de datos activos
- Estadísticos/ resúmenes/distribución de frecuencias/sexo
- Estadísticos/tablas de contingencia/tabla de doble entrada

## Tableau Public

Descárgalo aquí: <https://public.tableau.com/en-us/s/download>

¿Quieres practicar con la base de datos de habilidades en Periodismo de datos? Bájala aquí ([https://docs.google.com/spreadsheets/d/1S8W68Tsud4i1ynvi\\_fAaY4-92Q0bLWNUhVWUeyfDtbQ/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1S8W68Tsud4i1ynvi_fAaY4-92Q0bLWNUhVWUeyfDtbQ/edit?usp=sharing))

Una vez que la cargues en Tableau practica acomodar las variables siguiendo este orden:

- Habilidad en filas
- Puesto en columnas
- Empresa en filas
- Número de registros
- Arrastras el campo habilidad a color
- Eliges Forma cuadrado
- Quita la suma de números