

Visualización de datos

Definición, tecnologías y herramientas

Iniciativa
aporta



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA, ENERGÍA,
Y TURISMO

MINISTERIO
DE HACIENDA
Y ADMINISTRACIONES PÚBLICAS

red.es

ÍNDICE

1. Visualización de datos

1.1. Visualización: explotación de datos

1.1.1. Tipos de visualización de datos

1.2. Visualización de datos y Open Data

2. Software de visualización de datos

2.1. Estándares web

2.1.1. Estándar HTML5 y Canvas

2.1.2. Estándar SVG

2.1.3. Estándar WebGL

2.2. Bibliotecas y framework para la visualización

2.2.1. Bibliotecas para gráficos 2D

2.2.2. Bibliotecas para gráficos 3D

2.2.3. Bibliotecas para mapas

2.2.4. Otro tipo de visualizaciones

3. Estado del arte en herramientas y plataformas de visualización de datos

3.1. Descripción general

3.2. Ejemplos de herramientas

3.2.1. Tableau Software

3.2.2. Tabulae

3.2.3. QlikView y QlikSense

3.2.4. CartoDB

3.2.5. Socrata

3.2.6. CKAN (Módulos de visualización)

4. Ejemplos de visualización de datos

- 4.1. Datos.gob.es
- 4.2. Banco mundial
- 4.3. Portal de transparencia del Principado de Asturias
- 4.4. Plataforma openREDBAG
- 4.5. Google Public Data
- 4.6. Eurostat
- 4.7. Fundación Civio
- 4.8 Otros ejemplos de visualización
 - 4.8.1. Evolución de la Web
 - 4.8.2. Calentamiento global.
 - 4.8.3. Retirada de Kobe Bryant
 - 4.8.4. Historia de la música

5. Conclusiones

- 5.1. Tendencias futuras

1. VISUALIZACIÓN DE DATOS

Vivimos en un mundo rodeados de datos que nuestro cerebro procesa de manera continua para construir la realidad, entender el entorno que nos rodea y tomar decisiones sobre nuestro futuro. En la actualidad nuestro consumo de información se ha multiplicado de manera exponencial debido a dos factores: **cada vez se produce más información** (redes sociales, dispositivos, etc.) y cada vez **tenemos más capacidad de acceso a dicha información**, especialmente a través de Internet y de la Web. La capacidad de sacar partido y entender la información bruta está íntimamente ligada a nuestra capacidad para explotarla y transformarla en algo más que puro dato: **los datos adquieren significado**.

No obstante, los datos, entendidos como registros aislados, no aportan un significado concreto. Sólo cuando nos acercamos a ellos y le aplicamos una interpretación, estos cobran sentido y se transforman en conocimiento. En el ámbito de la tecnología, la explotación de datos ha evolucionado en las últimas décadas para diseñar mecanismos de interpretación cada vez más robustos y asequibles. Y, entre estos mecanismos de explotación, el más importante es la **visualización de datos**.

1.1. Visualización: explotación de datos

La visualización de datos es la presentación gráfica de información con dos propósitos. Por un lado, la interpretación y construcción de significado a partir de los datos (es decir, el análisis); y por otro lado, la comunicación.

La visualización es una herramienta muy potente para descubrir y comprender la lógica que se encuentra detrás de un conjunto de datos, así como para compartir esta interpretación con otras personas desde un punto de vista objetivo. Como se suele decir: *“una imagen vale más que mil palabras”*, especialmente cuando el significado que se quiere comunicar se representa mejor gráfica que verbalmente, y esta representación gráfica está diseñada de acuerdo a los principios formales que rigen la visualización de datos.

Podemos estar durante horas ante una tabla de números y no visualizar lo que es inmediatamente obvio cuando esta información se presenta gráficamente. Por ejemplo, imaginemos la siguiente serie de datos histórica de ventas de hipotecas en España (Fuente INE:

serie 2003-2011), en la que se desglosa el volumen de venta de hipotecas por año y entidad financiera:

Periodo	Entidad	Nº Hipotecas
2010	Otras entidades	103,055
2008	Otras entidades	122,332
2009	Banca	367,298
2011	Banca	258,667
2008	Banca	408,712
2010	Banca	337,825
2003	Cajas de ahorro	685,464
2011	Cajas de ahorro	279,461
2006	Banca	576,777
2006	Cajas de ahorro	1,016,728
2004	Banca	583,213
2009	Otras entidades	111,066
2005	Banca	594,477
2003	Banca	486,579
2004	Cajas de ahorro	791,349
2003	Otras entidades	119,617
2005	Cajas de ahorro	924,882
2007	Banca	518,638
2008	Cajas de ahorro	683,546
2007	Cajas de ahorro	1,017,548
2011	Otras entidades	86,420
2007	Otras entidades	154,809
2009	Cajas de ahorro	549,558
2010	Cajas de ahorro	475,536
2005	Otras entidades	184,773
2004	Otras entidades	147,843
2006	Otras entidades	202,889

A pesar de que comprendemos los datos que se muestran en la tabla, para el usuario final será mucho más sencillo de interpretar dicha información si se crea una gráfica con la evolución del mercado inmobiliario en España y el efecto burbuja, tal y como se muestra en la siguiente figura.

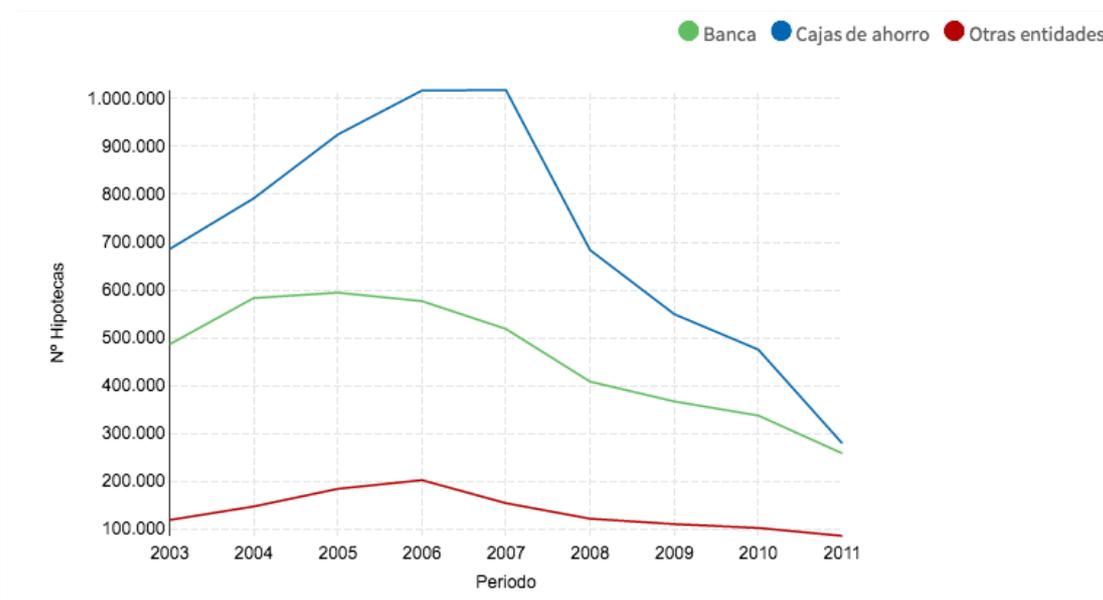


Figura 1. Serie histórica venta de hipotecas en España: 2003-2011 (Fuente: INE)

Lo que los números no pueden comunicar cuando son presentados en una tabla, se vuelve visible e inteligible cuando se comunica visualmente. **Este es el “poder” de la visualización de datos.**

Es importante tener en cuenta que, aunque la visualización de datos se utiliza para representar generalmente variables cuantitativas y relaciones entre ellas, también puede utilizarse para representar relaciones entre entidades de naturaleza cualitativa. Por ejemplo, las relaciones entre las personas de una determinada red social, que pueden ser además “tipadas” de acuerdo a la naturaleza de esta relación: amistad, familia, trabajo, etc. Estas visualizaciones de representación de entidades y propiedades de tipo relacional se sustentan sobre la tipología de la estructura que se quiere representar y utilizan gráficos basados en grafos (nodos y arcos).

Históricamente la visualización ha existido de manera consustancial a los propios datos. Especialmente en el ámbito de la cartografía. Sin embargo, es a **finales del siglo XVIII y principios del siglo XIX cuando aparecen los primeros estudios y aplicaciones de la visualización de datos** para construir narrativas y entender fenómenos reales: desde indicadores económicos hasta eventos históricos. En este sentido, hay que destacar el trabajo pionero del economista escocés William Playfair y su libro: *The Commercial and Political Atlas and Statistical Breviary*.

Playfair se le considera el fundador de aplicación de técnicas gráficas para el análisis estadístico, inventando gráficos como los de líneas, áreas (como en la Figura 2), barras y de tarta.

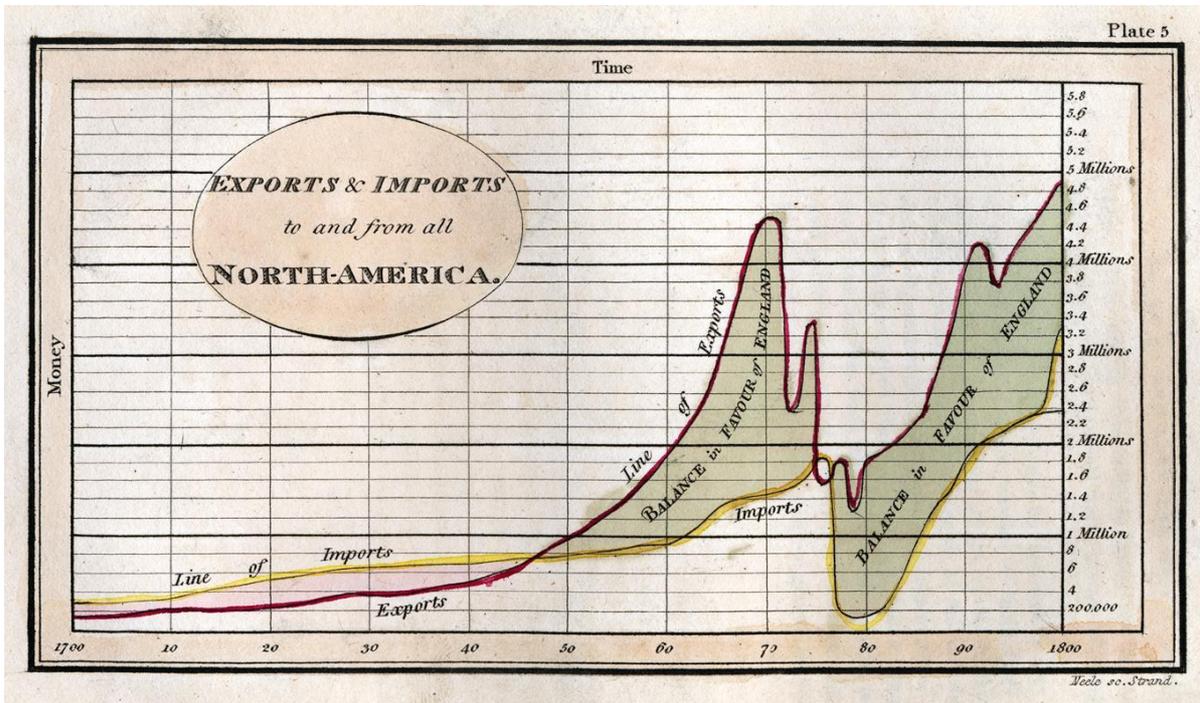


Figura 2. Serie histórica (1700 - 1800): balanza comercial entre Inglaterra y Norteamérica

El trabajo innovador de Playfair se vio acompañado por el trabajo de otros ingenieros y economistas de la época, como es el caso de Charles Minard y su famosa narrativa gráfica sobre la marcha de Napoleón sobre Rusia en 1812, M.H.P.R. Sankey y los diagramas y mapas de flujo, Jon Snow y Charles Dupin y el uso de mapas temáticos o de coropletas para representar variables cuantitativas en una cartografía (véase Figura 3), y así un largo etcétera.

Ya en el siglo XX, podemos destacar a dos autores de referencia: Tukey¹ y Tufte² y sus respectivas obras, que establecieron principios formales la visualización. Tukey estableció lo que se conoce como EDA (Exploratory Data Analysis) y explicó cómo usar métodos visuales para entender los datos y formular hipótesis sin necesidad de utilizar modelos estadísticos.

¹ Tukey, John W (1977). *Exploratory Data Analysis*. Addison-Wesley.

² Tufte, Edward R (2001) [1983], *The Visual Display of Quantitative Information* (2nd ed.), Cheshire, CT: Graphics Press

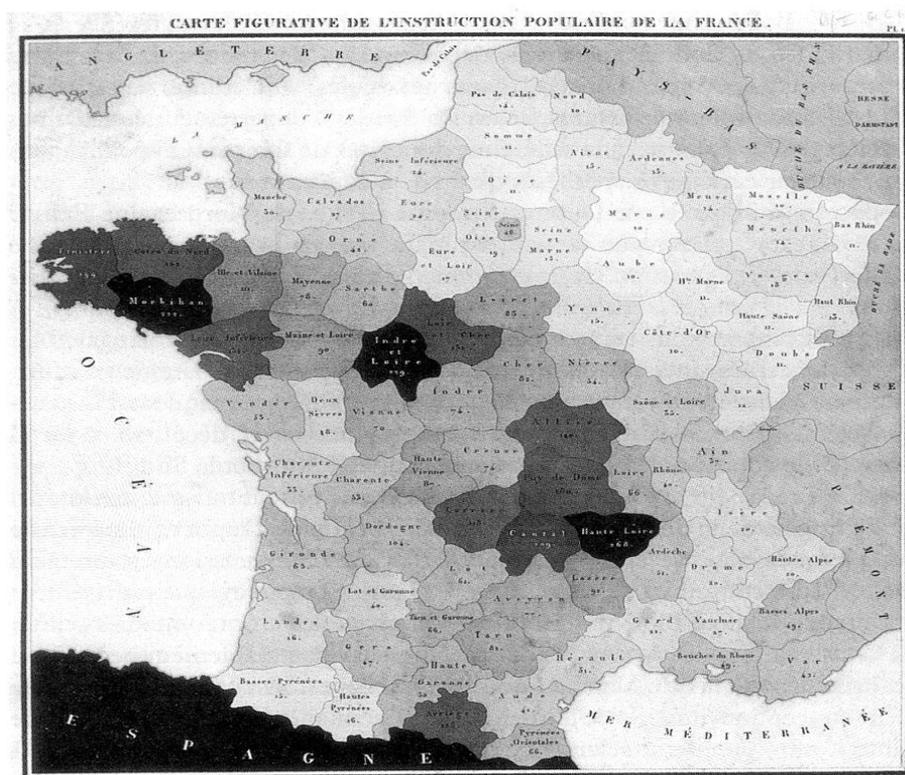


Figura 3. Representación de Dupin de la distribución de población en Francia en 1826

Hoy en día, tal y como veremos en las siguientes secciones del presente informe, **existen multitud de herramientas para la generación de gráficos y mapas** de maneras asistida, que permiten construir valor sobre los datos. Aunque la visualización puede seguir siendo un proceso artesanal basado en el diseño y la elaboración manual, como son algunos ejemplos de infografías, en la mayor parte de aplicaciones se utiliza software específico para la construcción de estas representaciones gráficas.

1.1.1. Tipos de visualización de datos

Existen multitud de técnicas y aproximaciones para la visualización según sea la naturaleza del dato de la información. Desde el punto de vista de los datos, especialmente datos estructurados (o semi-estructurados) y su explotación visual, podemos establecer *grosso modo* la siguiente clasificación³ de **tipos de visualización según complejidad y elaboración de la información**.

³ Esta clasificación no obedece a criterios académicos. Sólo tiene un propósito ilustrativo.

1. Elementos básicos de representación de datos.

Es el caso más sencillo. Un elemento básico de visualización puede ser un gráfico, un mapa, un KPI, tablas de datos, un grafo etc. A continuación se señalan algunos tipos de visualizaciones básicas.

- **Gráficas:** barras, líneas, columnas, puntos, “tree maps”, tarta, semi-tarta, etc.
- **Mapas:** burbujas, coropletas (o mapa temático), mapa de calor, de agregación (o análisis de drill-down)
- **Tablas:** con anidación, dinámicas, de drill-down, de transiciones, etc.

2. Cuadros de mando.

Un cuadro de mando es una composición compleja de visualizaciones individuales que guardan una coherencia y una relación temática entre ellas. Son ampliamente utilizados en las organizaciones para análisis de conjuntos de variables y toma de decisiones.



Figura 4 – Ejemplo de cuadro de mandos empresarial

3. Infografías

Una infografía también es una composición elaborada de visualizaciones que construye, a partir de diferentes elementos, un significado complejo para el usuario. Sin embargo, las infografías no están tan destinadas al análisis de variables sino a la **construcción de narrativas a partir de los datos**; es decir, las infografías se utilizan para contar “historias”.

1.2. Visualización de datos y open data

El movimiento de apertura de datos dentro de las administraciones se ha focalizado en los últimos años en los aspectos de su publicación. Por un lado, se han definido políticas y normativas para empujar a las administraciones a abrir sus datos a la sociedad, tanto por compromisos de transparencia como en la búsqueda de un retorno en la reutilización de la información a través de otras organizaciones y de desarrollo de un tejido productivo a su alrededor.

Por otro lado, se han definido [formatos de publicación y buenas prácticas](#) para que las administraciones supiesen no sólo qué publicar, sino cómo y cuándo, de manera que los procesos de búsqueda, intercambio y acceso de la información obedezcan a principios de diseño estándares que faciliten la **automatización de la reutilización del open data**.

En este sentido, podemos destacar la catalogación de los diferentes niveles de publicación de datos y la definición de buenas prácticas en cuanto a reutilización (por ejemplo, la distinción 5 estrellas de Tim Berners-Lee⁴)

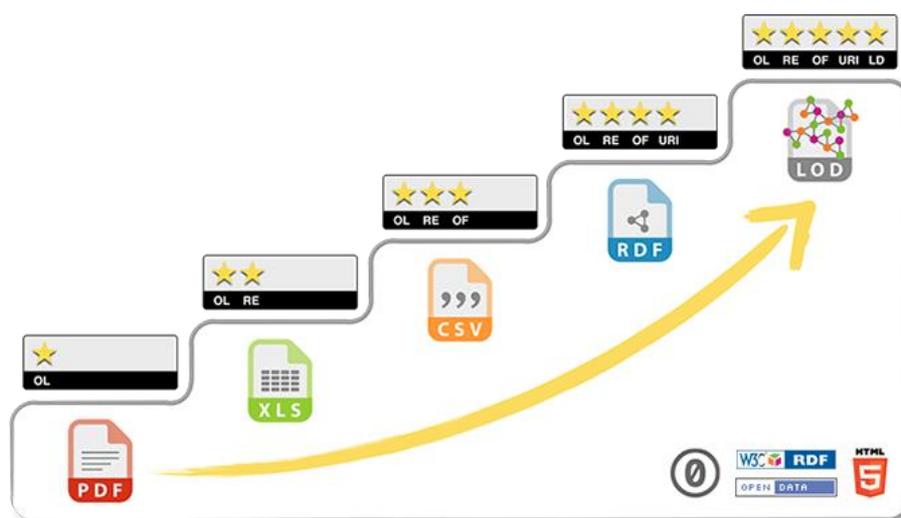


Figure 4. Categorías de publicación de datos abiertos (imagen de 5stardata.info)

Organizaciones como el W3C, la Open Knowledge Foundation (OKFN) o la World Wide Web Foundation han resultado determinantes para la evangelización sobre la necesidad de abrir los datos públicos y empujar el trabajo técnico sobre la publicación de los mismos. Tanto en España

⁴ Tim Berners-Lee (2006): *Linked Data - Design Issues*. <https://www.w3.org/DesignIssues/LinkedData.html>

en particular, como en Europa en general, esto ha desembocado en numerosas iniciativas y proyectos, así como la aparición de legislación específica al respecto. La Unión Europea, de manera muy sensible, está promoviendo de manera activa la armonización de estas dinámicas a través de [propuestas tecnológicas concretas como DCAT-AP](#) y fomentando el impacto empresarial a través de programas de [financiación como el H2020](#).

Sin embargo una de las críticas⁵ más habituales en torno al open data es que el esfuerzo actual se concentra en la publicación de datos y no en su **usabilidad**; es decir, en cómo esos datos se van a consumir por el usuario final. Muchos proyectos open data se han concentrado excesivamente en estas cuestiones técnicas: formatos, *endpoints*, etc., ajenas al usuario potencial, y no han prestado atención en cómo esos datos pueden utilizarse y se les puede extraer valor. Esto ha provocado que muchos portales se hayan convertido en meros repositorios de datos, con un tráfico marginal y con un impacto social muy bajo.

Así, es fundamental entender que es natural que los primeros pasos en el movimiento open data se hayan concentrado precisamente en la definición de qué es un dato abierto y cuál es su procedimiento de publicación. Ahora mismo, sin embargo, tanto la tecnología disponible y la madurez del movimiento como la demanda ciudadana, obligan a dar un paso más allá y evolucionar el concepto open data a una dimensión más amplia en la que no sólo se tengan en cuenta cuestiones referentes a la publicación y la reutilización desde el punto de vista técnico.

El siguiente hito es **proporcionar herramientas para que los usuarios puedan consumir y explotar los datos de manera autónoma**, haciendo realidad el objetivo inicial de que la información de las administraciones tenga un impacto positivo y real en nuestra sociedad. En este punto, la visualización, como hemos comentado antes, es la herramienta más potente para acercar los datos a cualquier usuario.

Actualmente la visualización forma parte de cualquier proceso estándar de gestión y análisis de datos en el mundo empresarial, especialmente cuando hablamos de cuestiones relacionadas con el análisis estadístico y la “Inteligencia de negocio” (*Business Intelligence*, en adelante BI), aunque

⁵ Helbig, N., Cresswell, A.M., Burke, G.B. and Luna-Reyes, L. (2012) *The Dynamics of Opening Government Data: A White Paper*. Centre for Technology in Government, State University of New York, Albany.
<http://www.ctg.albany.edu/publications/reports/opendata/opendata.pdf>

no de manera exclusiva. En cualquier caso, podemos definir estos procesos como el **conjunto de técnicas y herramientas para la adquisición, tratamiento y transformación de datos brutos** en información y conocimiento útiles para un propósito de análisis de determinado.

En el ámbito de la empresa, el propósito suele estar ligado a procesos y aspectos de negocio, pero la aplicación de las técnicas y herramientas es genérico respecto de la intención del análisis. En estos casos, la visualización forma parte del ciclo de vida de los datos, en los que después de su preparación y transformación, la información se explota por parte del usuario final a través de representaciones gráficas que le permiten interpretar el sentido de los datos.

El movimiento open data puede reutilizar tanto la tecnología disponible actualmente en el mercado, como de las lecciones aprendidas en este ámbito en los últimos quince años. No se trata de empezar todo desde cero, sino de aplicar el conocimiento adquirido en estos años para dar un salto cualitativo y hacer que los datos adquieran el uso esperado por parte de la ciudadanía.

2. SOFTWARE DE VISUALIZACIÓN DE DATOS

A lo largo de esta sección, se describen los estándares web que se han ido desarrollando en los últimos años para la evolución de las aplicaciones web, base fundamental para creación de visualizaciones web basadas en datos. A su vez, también se describen diferentes bibliotecas JavaScript que utilizan estos estándares para la creación de las partes fundamentales de la visualización.

2.1. Estándares web

En los últimos años, el desarrollo de contenidos y aplicaciones web ha sufrido una revolución motivada por los nuevos estándares web. Desde que en 2014 el W3C publicó la versión definitiva del estándar HTML5⁶ **se han ido sumando nuevos estándares** que permiten a los desarrolladores crear, no solo visualizaciones estáticas, sino potentes aplicaciones de visualización que incluyen un alto grado de dinamismo y personalización estética. Estándares como HTML5 o la nueva versión del estándar de hojas de estilo en cascada (CSS3⁷) junto con la evolución de los navegadores modernos permiten a los desarrolladores crear aplicaciones web multidispositivo sin tener que preocuparse, como solía ocurrir, de qué navegador o dispositivo se use para ver la visualización.

Estándar	Última Versión	Función
HTML5	v5	Canvas: elemento HTML para dibujar gráficos 2D
CSS3	v3	Permite diferenciar el contenido de las páginas web de la presentación de este contenido
SCV	v2	Utilizado para crear gráficos 2D
WebGL	v1	Gráficos 3D haciendo uso de Canvas

⁶ <https://www.w3.org/TR/html5/>

⁷ <https://www.w3.org/Style/CSS/>

A continuación, se explican detalladamente cada uno de ellos:

2.1.1. Estándar HTML5 y Canvas

El W3C publicaba en 2014 la versión 5 del estándar más importante para la web: **HTML5**, terminando así con 6 años de desarrollo en colaboración con las principales empresas tecnológicas. **HTML es el lenguaje básico sobre el que se construyen todos los contenidos**. Esta nueva versión supuso un salto de calidad en este estándar incluyendo nuevas API para la interacción con el navegador y el dispositivo, y lo más importante, obligando a todos los navegadores web a actualizarse para soportar esta nueva versión.

Esto trajo como consecuencia que las compañías responsables de cada uno de los navegadores modernos estuviesen obligadas a mejorar la compatibilidad de sus productos con esta nueva versión. Se facilitó la **homogeneización de los contenidos web** y se terminó, al menos parcialmente, con los “hacks” de código necesarios para visualizar los contenidos correctamente en cada uno de ellos.

Dentro de esta nueva versión HTML5 se definió un elemento importante para la visualización de contenidos, el **Canvas**. Este elemento HTML es usado para dibujar gráficos, normalmente 2D aunque también puede ser utilizado junto con WebGL para visualizar gráficos 3D, usando scripts, normalmente escritos en JavaScript. Canvas puede ser utilizado para **dibujar gráficos, realizar composición de fotos y animaciones**. Existen múltiples bibliotecas JavaScript que facilitan la creación de gráficos para este elemento.

2.1.2 Estándar CSS3

Las hojas de estilo en cascada o CSS (Cascading Style Sheets) es un lenguaje que permite definir la presentación de un documento estructurado como HTML o CSS. Este lenguaje es un estándar del consorcio W3C que permite diferenciar el contenido de las páginas web de la presentación de este contenido. Tras muchos años de desarrollo, en 2011 se publicó la versión 3 de este lenguaje que supone un gran avance en cuanto a potencia y funcionalidad.

Debido a su definición modular, no todos los elementos del lenguaje CSS3 tienen el mismo nivel de madurez tecnológica. Respecto a la visualización de datos, los módulos más relevantes son los siguientes:

- **Fondos y colores:** Añadida la posibilidad de añadir múltiples fondos de pantalla, así como nuevos elementos que facilitan controlar mucho mejor los tamaños y posiciones de estos.
- **Animaciones:** Evolución de las animaciones que se pueden describir para los diferentes elementos del lenguaje.
- **Posicionamiento:** Este módulo es una evolución del posicionamiento clásico que permite distribuir los elementos de forma sencilla y mucho más flexible.

2.1.3. Estándar SVG⁸

Los Gráficos Vectoriales Redimensionables o SVG son una especificación estándar para describir gráficos vectoriales bidimensionales, tanto estáticos como animados (estos últimos con ayuda de SMIL), en formato XML. El W3C lo convirtió en recomendación en 2001 pasando a ser uno de **los estándares más utilizados a la hora de crear gráficos en 2D**. Este estándar nos permite definir conceptos básicos como punto, línea, polígono para después combinarlos y formar el gráfico vectorial complejo. Resulta especialmente útil para la creación de componentes de gráficos estadísticos que permiten la visualización y el análisis de series de datos.

Actualmente la versión 2 de SVG está en desarrollo y promete añadir un amplio conjunto de características que faciliten y amplíen el uso de las nuevas características de los navegadores y estándares web modernos.

2.1.4. Estándar WebGL⁹

WebGL es un estándar que nace de la mano de Mozilla pero que actualmente incluye a muchas de las grandes empresas tecnológicas. Este estándar permite definir gráficos y animaciones 3D haciendo uso del componente de HTML5, Canvas. WebGL está basado en OpenGL ES 2.0 y

⁸ <https://www.w3.org/Graphics/SVG/>

⁹ <https://www.khronos.org/registry/webgl/specs/1.0/>

proporciona una API para gráficos 3D. Actualmente está ampliamente soportado por los navegadores pasando a ser el lenguaje más utilizado para las visualizaciones web en 3D.

Como WebGL es una tecnología diseñada para trabajar directamente con el procesador gráfico del dispositivo resulta bastante difícil de codificar en comparación con otros estándares web más accesibles. Para resolver esto han ido surgiendo múltiples bibliotecas JavaScript que facilitan el uso de esta especificación.

2.2. Bibliotecas y framework para la visualización

2.2.1. Bibliotecas para gráficos 2D

Existen múltiples bibliotecas que facilitan la creación de gráficos 2D para visualizaciones de datos. Dentro de este grupo podemos diferenciar las que están basadas en Canvas, y las que están basadas en SVG.

	Estándar	Paleta de gráfico	Acabado final	Facilidad de uso	Extensible	Modalidad de uso
D3.js	SVG	Altamente extensible	Bueno	Difícil	Si	Gratuita
Google Charts	SVG	Muy amplia	Bueno	Sencillo	No	Gratuita
Chart.JS	Canvas	Líneas, Barras, Radar y Tarta	Muy bueno	Sencillo	No	Gratuita
HighchartsJS	Canvas	Muy amplia	Muy bueno	Sencillo	No	Pago/Gratuita
JavaScript InfoVis Toolkit	WebGL	Bastante reducida	Regular	Difícil	No	Gratuita

2.2.1.1. D3.js¹⁰

D3.js es una biblioteca JavaScript para la manipulación de documentos basados en datos. D3 ayuda a dar vida a los datos usando HTML, SVG y CSS. El punto clave de D3 es la utilización de estos estándares web para que el editor/desarrollador pueda explotar todas las nuevas capacidades de los navegadores web modernos, facilitando la manipulación del DOM para la creación de poderosas visualizaciones basadas en datos.

¹⁰ <https://d3js.org/>

D3 nos da una capa de abstracción por encima de SVG ofreciendo una serie de utilidades para la manipulación de datos y su posterior visualización. De este modo, D3, nos ofrece **la creación de gráficos interactivos que pueden ser personalizados utilizando estándares con CSS3**.

Por todo esto, D3 se ha convertido en **una de las bibliotecas más utilizadas** no solo para la creación de visualizaciones sino como base de un gran conjunto de bibliotecas más específicas basadas en esta biblioteca.

Algunos ejemplos son n3-charts, NVD3¹¹, etc. Este tipo de envoltorios minimiza el principal problema que encontramos en D3, la genericidad. Al tratarse de una librería genérica para gráficos 2D, resulta complicada su utilización. Por ello las bibliotecas como NVD3 nos ofrecen clases Javascript para los principales tipos de gráficos, como por ejemplo: gráfico de barras, de líneas, de área, de burbujas, etc., facilitando la creación de este tipo de visualizaciones.

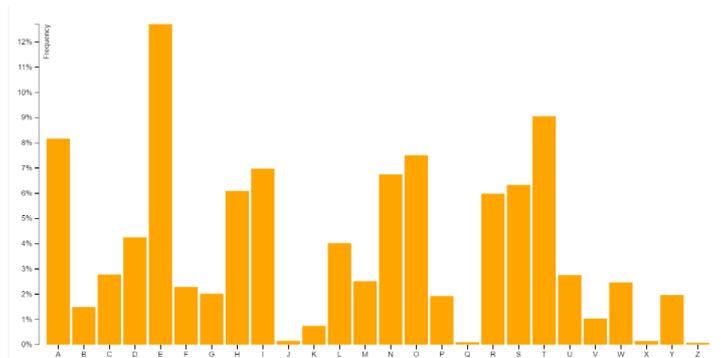


Figura 5 - Ejemplo de gráfica de barras de D3

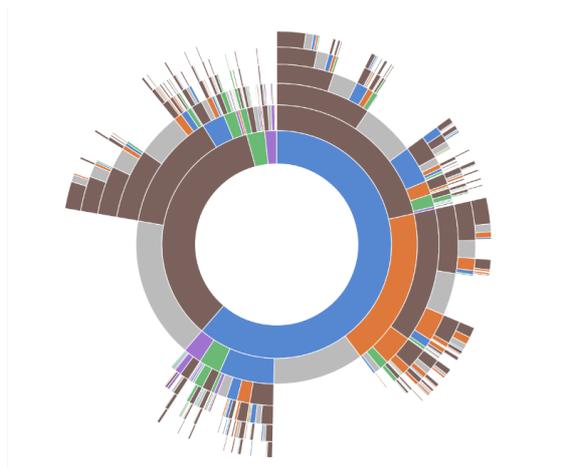


Figura 6 - Ejemplo de una visualización más compleja de D3

¹¹ <http://nvd3.org/>

2.2.1.2. Google Charts¹²

Google ha apostado fuerte por la visualización de datos y ha creado su propia biblioteca para la creación de visualizaciones. Esta biblioteca también está basada en SVG y dispone de un amplio número de tipos de gráficos y mapas. Dispone también de lo que han denominado datos dinámicos, que nos permiten hacer consultas SQL sobre datos almacenados en Google Spreadsheets, Google Fusion Tables o Salesforce y visualizarlos. Google ofrece también, aunque de forma limitada, algunos componentes para creación de dashboards interactivos.

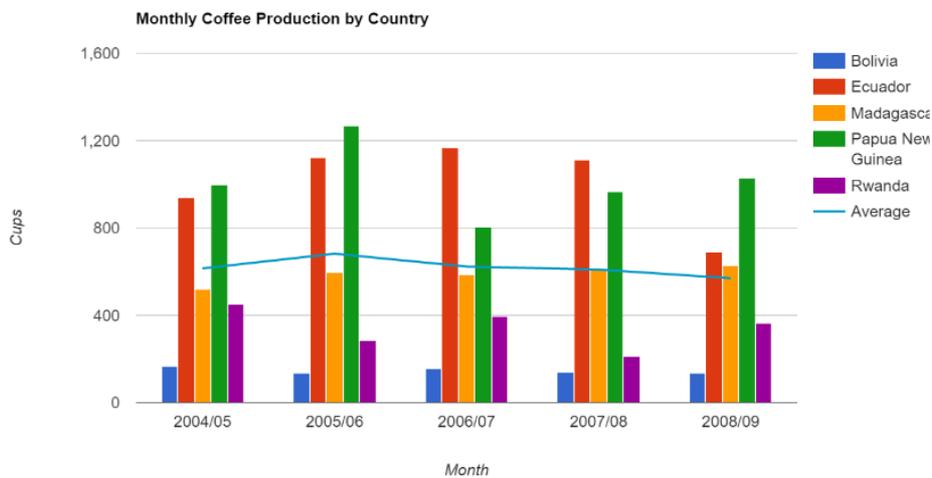


Figura 7- Ejemplo de gráfica de Google Charts

2.2.1.3. Chart.JS¹³

Chart.JS es el primero de los ejemplos incluidos que utiliza Canvas en lugar de SVG. Este tipo de bibliotecas nos ofrecen clases JavaScript para dibujar el gráfico dentro del nuevo elemento de HTML5 Canvas. El uso de Canvas en lugar de SVG limita la interoperabilidad de los gráficos creados con estándares como CSS3 pero tiene la ventaja de que necesitan menos recursos computacionales para su para que el navegador web los represente:

¹² <https://developers.google.com/chart/>

¹³ <http://www.chartjs.org/>

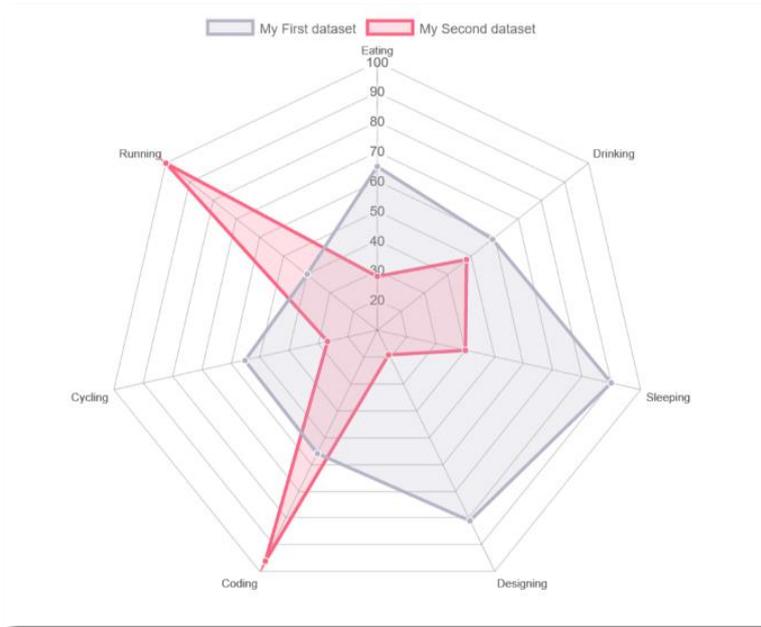


Figura 8 - Ejemplo de grafico creado con Chart.JS

2.2.1.4. HighchartsJS¹⁴

HighchartsJS es una biblioteca comercial, gratuita para uso no comercial, que cuenta un gran número de seguidores. Esta biblioteca también utiliza el estándar SVG para sus gráficos. Como principal punto a su favor, **dispone de un gran número de componentes y una documentación muy extensa y cuidada**. La principal desventaja es la licencia, que hay que adquirir si se quiere usar para fines comerciales.

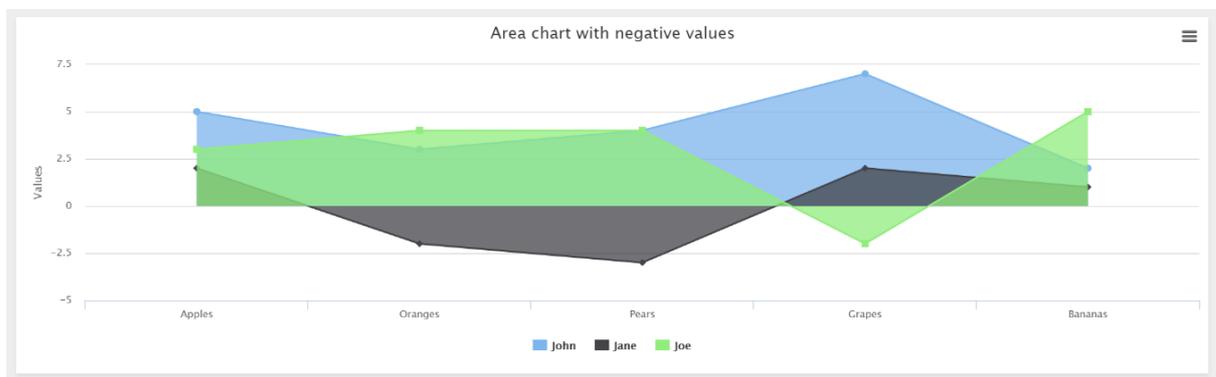


Figura 9 - Ejemplo de grafico creado con HightchartsJS

¹⁴ <http://www.highcharts.com/>

2.2.1.5. JavaScript InfoVis Toolkit¹⁵

Ejemplo de biblioteca de gráficos que utiliza estándar WebGL. Esta biblioteca está desarrollada por Nicolás García Belmonte bajo copyright de Sencha Company y ofrece la incorporación de ciertos aspectos 3D a los tipos de gráficos más utilizados. Este tipo de características permite crear mapas interactivos ofreciendo espectaculares animaciones 3D.

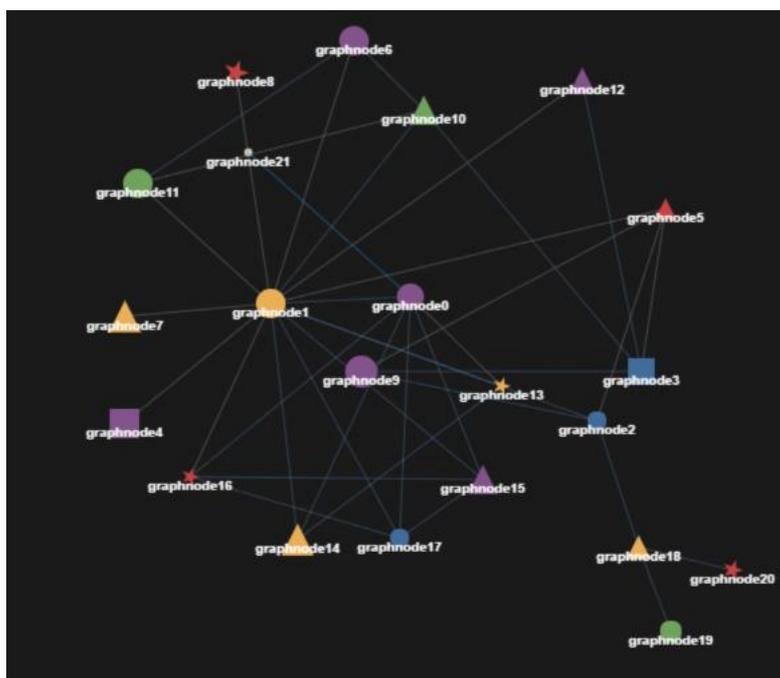


Figura 10 - Ejemplo de grafo creado con JavaScript InfoVis Toolkit

¹⁵ <http://philogb.github.io/jit/>

2.2.2. Bibliotecas para gráficos 3D

Existen múltiples bibliotecas que utiliza WebGL para la creación de gráficos en 3D todas ellas están orientadas a la creación de animaciones y de juegos HTML5 pero son muy útiles para la creación de visualizaciones con un alto índice de impacto.

	Propósito	Estándar	Gráficas y Animaciones 3D	Dependencias
Three.js	General	WebGL	Sí	No
BabylonJS	Animaciones, juegos	HTML5, WebGL	Sí	Hand.js

2.2.2.1. Three.js¹⁶

Three.js es una librería para construir escenas 3D con WebGL. Podemos verlo con un D3 para WebGL, es decir, una biblioteca que elimina la complejidad de escribir una visualización en WebGL. Ofrece todas las clases necesarias para aspectos comunes de las escenas 3D como las luces, cámaras, materiales y objetos, etc.



Figura 11 - Ejemplo de animación creada por Google usando Three.js

2.2.2.2. BabylonJS¹⁷

¹⁶ <http://threejs.org/>

¹⁷ <http://www.babylonjs.com/>

Framework para la creación de juegos 3D con HTML5 y WebGL que también puede ser utilizado para crear espectaculares visualizaciones. Como ocurre con Three.js ofrece las clases necesarias de Cámara, Luz y Figura, de modo que nos abstraer de toda la complejidad de WebGL.

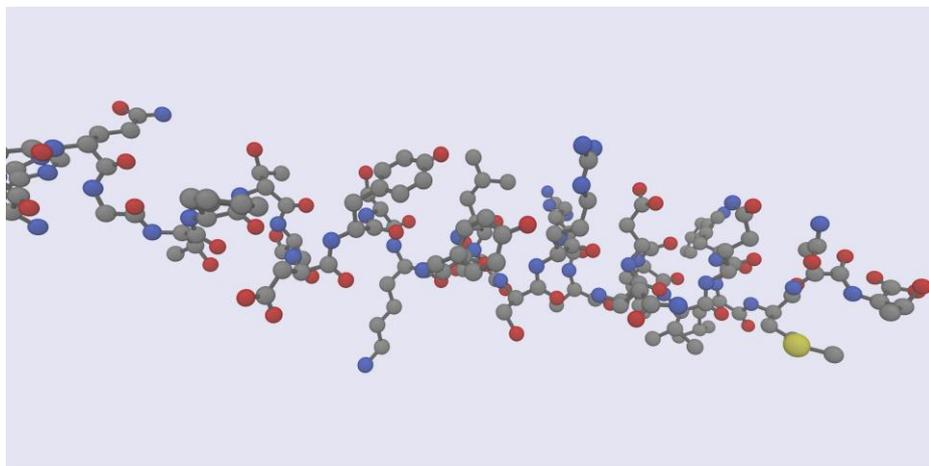


Figura 12 - Ejemplo de visualización 3D con BabylonJS

2.2.3. Bibliotecas para mapas

Con el auge de los servicios de mapas como Google Maps, OpenStreetMap y los sistemas GIS (Sistemas de Información Geográfica), han ido apareciendo múltiples bibliotecas para la visualización y el análisis de datos georreferenciados. Muchos de estas bibliotecas y clientes se basan en los estándares definidos por el Open Geospatial Consortium (OGC)¹⁸, como los servicios WMS (Web Map Service) o WFS (Web Feature Service).

Cabe destacar el uso del vocabulario GeoJSON¹⁹ como estándar de facto para la descripción de capas y objetos georreferenciados en el entorno web. Este vocabulario define los conceptos básicos de punto, ruta y polígono y permitiendo su combinación para la descripción de capas de visualización

¹⁸ <http://www.opengeospatial.org/>

¹⁹ <http://geojson.org/>

Existen múltiples bibliotecas para la visualización de datos georreferenciados, destacamos: polymaps.js y leaflet como dos de las más utilizadas.

	Elementos vectoriales	Capas GeoJson	Conexión a capas WMS	Conexión a capas WFS	Extensible (plugins)
Polymaps.js	Sí	Sí	No	No	No
Leaflet	Sí	Sí	Sí	Sí	Sí

2.2.3.1. Polymaps.js²⁰

Polymaps es una biblioteca JavaScript para la creación de mapas dinámicos e interactivos para la web. Proporciona mecanismos para visualizar capas vectoriales de datos sobre las cartografías más utilizadas, como OpenStreetMap, CloudMade, Bing, etc.

Polymaps utiliza el estándar SVG para la creación de estas capas a partir de, por ejemplo, datos definidos con GeoJSON, permitiendo también su posterior personalización a través de la definición de reglas de visualización escritas en CSS.

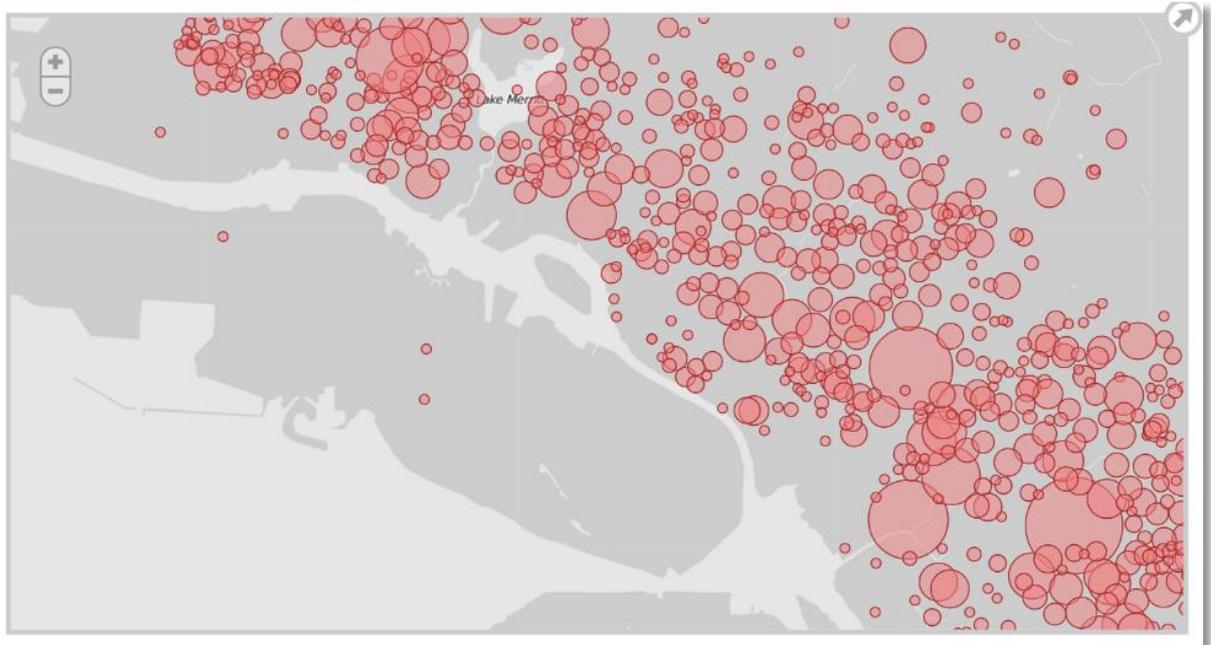


Figura 13 - Ejemplo de un mapa de burbujas creado con Polymaps

²⁰ <http://polymaps.org/>

2.2.3.2. Leaflet²¹

Leaflet es una de las bibliotecas de visualización y creación de mapas interactivos más usadas en la actualidad. Permite la creación de capas animadas a partir de múltiples formatos de datos, entre los que se encuentran GeoJSON ó WMS. Además ofrece un gran conjunto de clases JavaScript para la creación y modificación de dichas capas así como para la visualización e interacción con ellas.

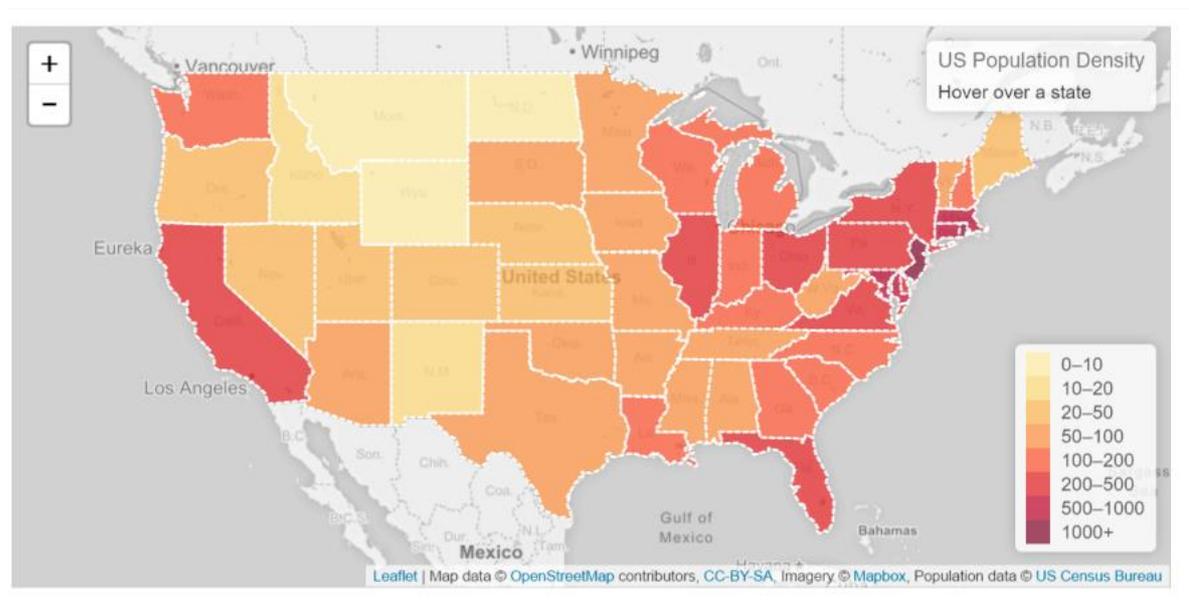


Figura 14- Ejemplo de un mapa de Leaflet

²¹ <http://leafletjs.com/>

2.2.4. Otro tipo de visualizaciones

Aunque muchas de las visualizaciones se basan en gráficos y mapas existen otros tipos de visualización que, dependiendo del problema, pueden resultar muy útiles para interactuar con los datos. Algunos ejemplos de estas visualizaciones son los grafos, líneas de tiempos o visores de ontologías.

2.2.4.1. Sigma.js²²

Sigma.js es una biblioteca JavaScript para la visualización de grafos para su utilización en entornos web. Esta biblioteca utiliza el elemento HTML5 Canvas para la visualización y ofrece todo un conjunto de utilidades para la visualización y el análisis de los elementos de grafo.

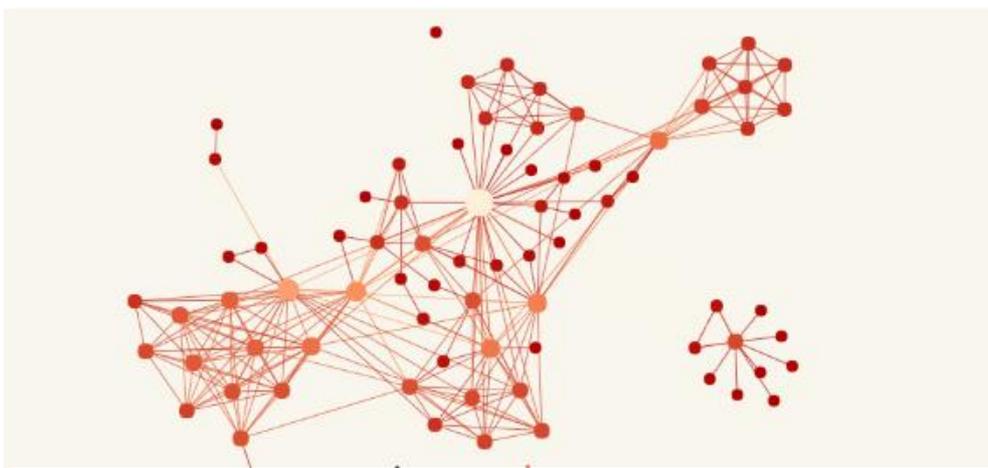


Figura 15- Ejemplo de un grafo creado Sigma

2.2.4.2. TimelineJS²³

Uno de los componentes recurrentes en las visualizaciones son las líneas de tiempo, es decir, en la localización de un conjunto de evento dentro de una línea temporal. TimelineJS ofrece de manera sencilla la creación de un componente de línea de tiempo que nos permite navegar entre los eventos representados.

²² <http://sigmaj.js.org/>

²³ <https://timeline.knightlab.com/>

Nelson Mandela's Extraordinary Life: An Interactive Timeline

By TIME Staff | Dec. 05, 2013 | Add a Comment

Share Like 2.3K Tweet G+ 16 LinkedIn Share 22 Pin it Read Later



Jurgen Schadeberg / Getty Images

Mandela in the office of Mandela and Tambo

12:00 AM
August 1, 1952

Law Firm

Mandela and Oliver Tambo found South Africa's first black-run law firm.

12:00 AM
Arrest



Figura 16 - Ejemplo de visualización de línea temporal

3. ESTADO DEL ARTE EN HERRAMIENTAS Y PLATAFORMAS DE VISUALIZACIÓN DE DATOS

En la actualidad el número de herramientas y plataformas de visualización de datos al que se puede acceder es muy extenso. En función de las necesidades del usuario se puede elegir entre un buen número de opciones de vanguardia. Dada la magnitud y variedad existente, es complicado hacer una clasificación rigurosa. Sin embargo se pueden distinguir varios grupos de herramientas bien diferenciados en función del **manejo de los datos** y del **tipo de resultado final que se desea**:

- **Ofimática** - Se permite mover/copiar contenidos entre diferentes aplicaciones
- **Web** - Orientación a crear contenidos web, no archivos
- **Analítica** - Se pueden realizar cálculos (avanzados) con los datos
- **Programación** - El usuario tiene control total mediante lenguaje de programación
- **Open data** - Orientación a usar datos abiertos (disponibles en la web)
- **Mapas** - Creación nativa de mapas (datos sobre una capa base)

	Ofimática	Web	Analítica	Programación	Opendata	Mapas
Excel	✓	✓		~		
Openoffice	✓			~		
Google Sheets	✓	✓			✓	
SAP BI		~	✓			~
SAS		~	✓	~		~
MS Power BI		~	✓			~
IBM Gognos		~	✓			~
Pentaho		~	✓	~		~
Qlik		✓	✓			✓
Tableau Sof.		✓	✓			✓
R			✓	✓		~
Weka			✓	✓		~
Matlab			✓	✓		~

Mathematica			✓	✓		~
Maple			✓	✓		~
Tabulae		✓	~			✓
CartoDB		✓				✓
Socrata		✓			✓	✓
CKAN		✓			✓	✓

✓=100%

~=50%

3.1. Descripción general

Como inicio de la descripción o estado del arte de las herramientas de visualización de datos en primer lugar hay que presentar la herramienta históricamente más popular de visualización: la hoja de cálculo -**MS Excel, OpenOffice Calc o Google Sheets**- entre otras son usadas a diario por millones de personas para todo tipo de representación visual basada en datos. Su facilidad de uso y el hecho de que en muchas ocasiones forman parte de *suits* de ofimática las convierte en la opción ideal para un uso sencillo y rápido. Además, hay versiones web con soporte *cloud* que las equiparan con otras herramientas más modernas, por lo que se resisten a abandonar su lugar de preferencia.

Cuando las hojas de cálculo se quedan “cortas” (gran volumen de datos, operaciones avanzadas, opciones de diseño o estéticas, etc.), aparecen herramientas y plataformas más avanzadas. En general, en este segundo grupo se engloban las herramientas asociadas al *Business Intelligence*, pensadas para ser la versión “profesional” de las hojas de cálculo.

Estas permiten una gestión completa de datos no como archivos individuales sino como sistemas de información con (posiblemente) cierta complejidad, y también una explotación avanzada de los datos a nivel analítico, que finaliza con las opciones de visualización. SAP, SAS, Microsoft, IBM, Pentaho, Qlik o Tableau entre otras, están presentes en miles de organizaciones que hacen uso extensivo de ellas.

Otra opción, con un gran número de seguidores entusiastas es el uso de lenguajes de programación con foco en la representación de datos. R, Weka, Matlab, Mathematica o Maple entre otros, representan la versión “técnica” de las hojas de cálculo y de las herramientas de BI,

en el sentido de que **precisan de unos conocimientos técnicos imprescindibles**, mientras que las segundas son accesibles para usuarios sin conocimientos especiales. Por contrapartida, las posibilidades para la creación de visualizaciones son prácticamente infinitas.

Hasta aquí el repaso de las herramientas “tradicionales”, porque con la mejora de la tecnología web durante los últimos años, **la visualización de datos se desplaza hacia la web**. Ciertamente los tres grupos anteriores tienen presencia web, hoy en día es un requisito indispensable. Sin embargo hay un grupo verdaderamente diferenciado, que consiste en las **aplicaciones web nativas**. Estas herramientas entienden la visualización de datos como parte natural de una página web (bien como un contenido embebido o directamente como una página propiamente), se adaptan a las necesidades actuales de contenidos dinámicos y responsivos, y a parte simplifican el proceso de creación de la visualización, tanto a nivel puramente de gestión de los datos como de la parte estética o visual. Un ejemplo en esta línea es la [plataforma Tabulae](#).

Finalmente, y destacando desde el punto de vista de la gestión open data, hay varios portales web que nacen para cubrir la necesidad de hacer una gestión de datos online. Socrata o CKAN son representantes de una herramienta menos enfocada a empresa que las herramientas de BI, y que entiende de forma nativa en la web la gestión de datos especialmente procedentes de fuentes públicas, su manipulación y su posterior visualización.

3.2. Ejemplos de herramientas

Algunas de estas herramientas están más cercanas al ámbito de análisis y de la visualización pura de datos (en la web). Otras, aunque nacen con el propósito inicial de dar soporte a la publicación de datos, posteriormente añaden a su funcionalidad capacidades de representación gráfica de datos para su consumo. ¿Cuáles son las herramientas principales?

3.2.1. Tableau Software²⁴

Tableau Software es una empresa dedicada íntegramente al análisis y visualización de datos. A través de su producto Tableau Desktop y Tableau Server ofrece a los usuarios una potente herramienta de análisis y enriquecimiento de datos. Además, su potente motor gráfico permite generar impactantes visualizaciones sobre grandes volúmenes de datos. Tableau Software es

²⁴ <http://www.tableau.com/es-es>

utilizado ampliamente en el sector de la comunicación (New York Times, O'Reilly Media), farmacéutico (Bayer CropScience) y en el de la educación (universidades americanas principalmente).

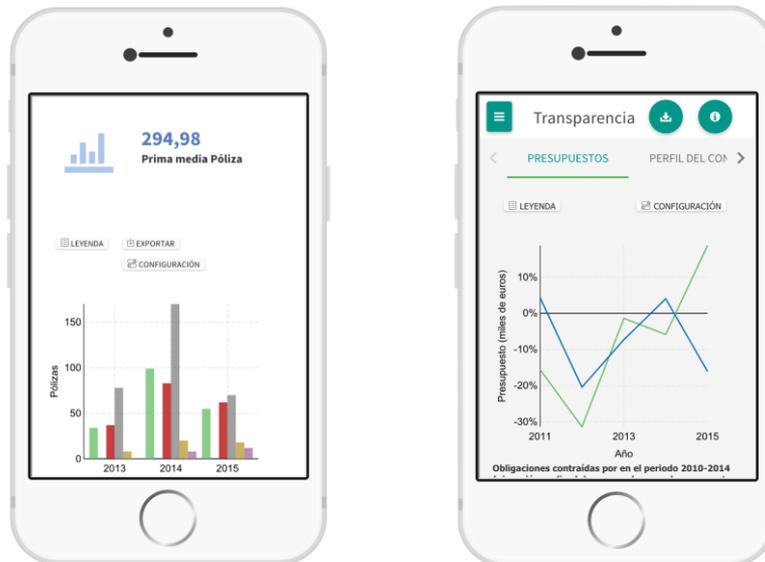
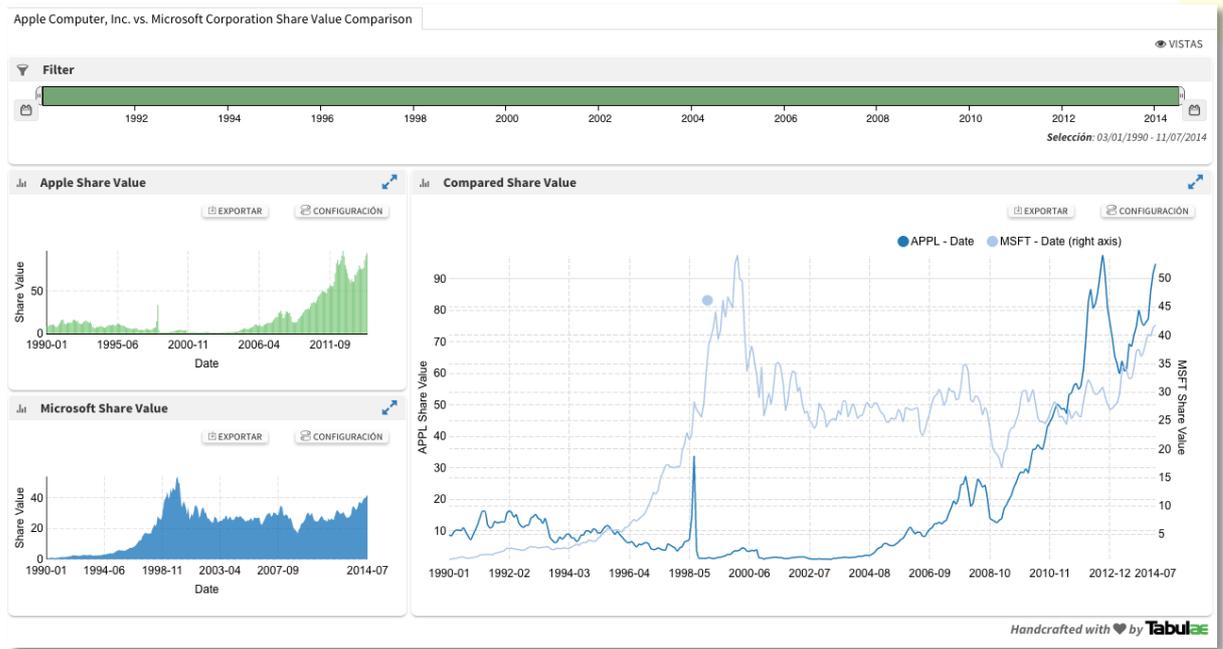


3.2.2. Tabulae²⁵

Tabulae es una plataforma web, basada en estándares W3C, de explotación y análisis visual de datos. Es una herramienta ágil, que destaca por su sencillez de uso y su capacidad de personalizar completamente la presentación de la información a través de cuadros de mando e informes interactivos.

Tabulae proporciona mecanismos de explotación dinámica, enriqueciendo la experiencia de usuario y su capacidad de interpretación de los datos. Sin conocimientos de programación, se pueden transformar cualquier conjunto de datos en una aplicación web interactiva y accesible desde cualquier dispositivo.

²⁵ <http://tabulaeapp.com/>



3.2.3. QlikView y QlikSense²⁶

QlikTech ofrece dos productos principales QlikView y QlikSense. Su producto estrella, QlikView, es un software de Business Intelligence que destaca por su sencillez de uso y por ser muy visual permitiendo a los usuarios de negocio la toma de decisiones basadas en datos. QlikView permite

²⁶ <http://www.qlik.com/>

la consolidación de datos procedentes de múltiples fuentes en una sola aplicación, lo que facilita la exploración de las asociaciones entre los datos. Por otro lado QlikSense es una versión más ligera y sencilla orientada más a la visualización que a la analítica. Ambas soluciones permiten la construcción de cuadros de mando e informes de forma sencilla. Su web dispone de un catálogo de demos muy variado, que nos permite descubrir rápidamente su funcionalidad y apariencia visual.

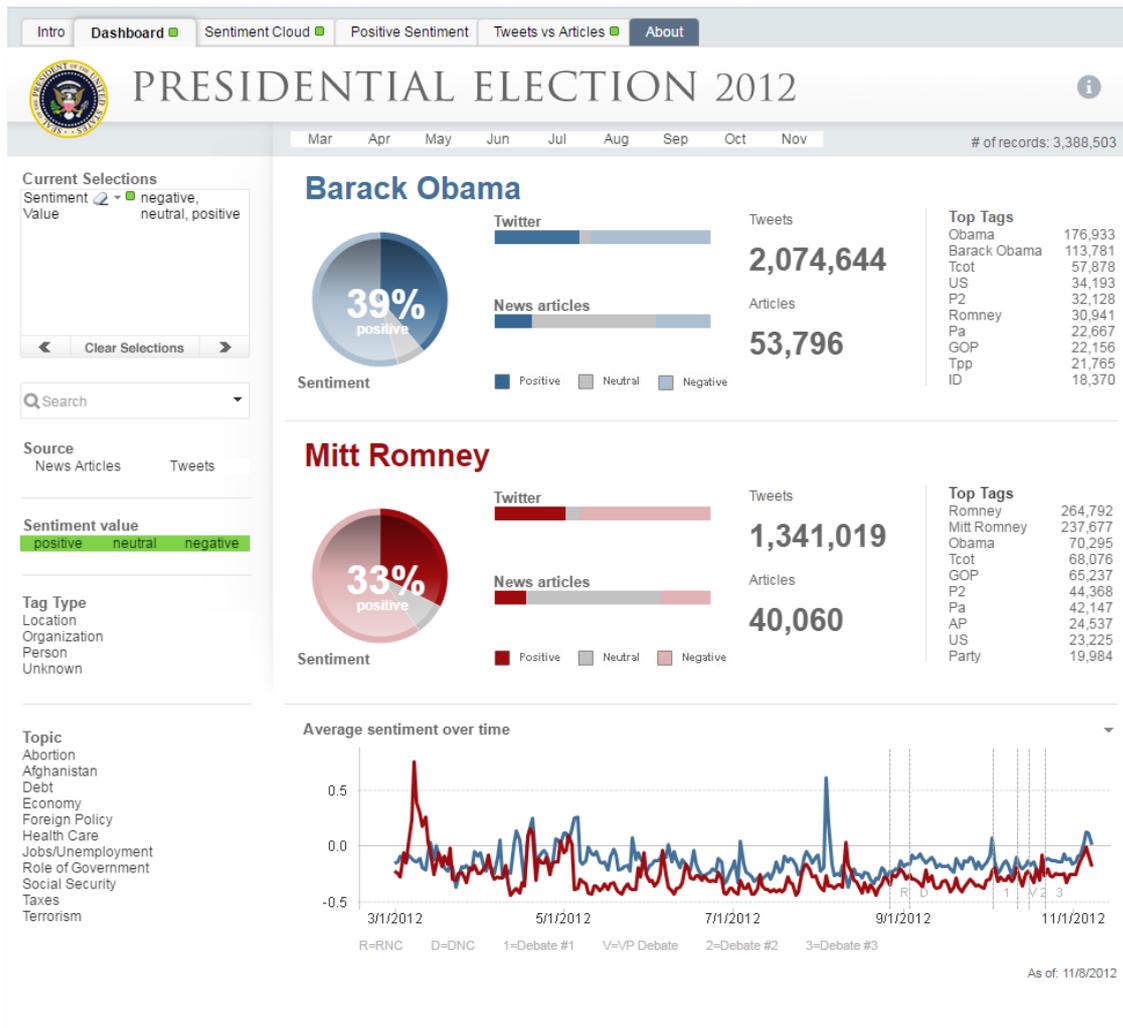


Figura 17 - Ejemplo de cuadro de mandos sobre las elecciones a la presidencia de EEUU de 2012 desarrollado con QlikView

3.2.4. Carto²⁷

Carto es un servicio en la nube que ofrece funcionalidades de GIS y herramientas para la construcción de mapas accesibles mediante un navegador web. Carto es un servicio ampliamente usado para la creación de **mapas temáticos** debido a su sencillez de uso y a unos resultados visualmente muy atractivos, por lo que podemos encontrar mapas creados con esta herramienta en multitud de noticias de periódicos digitales, blogs, etc.

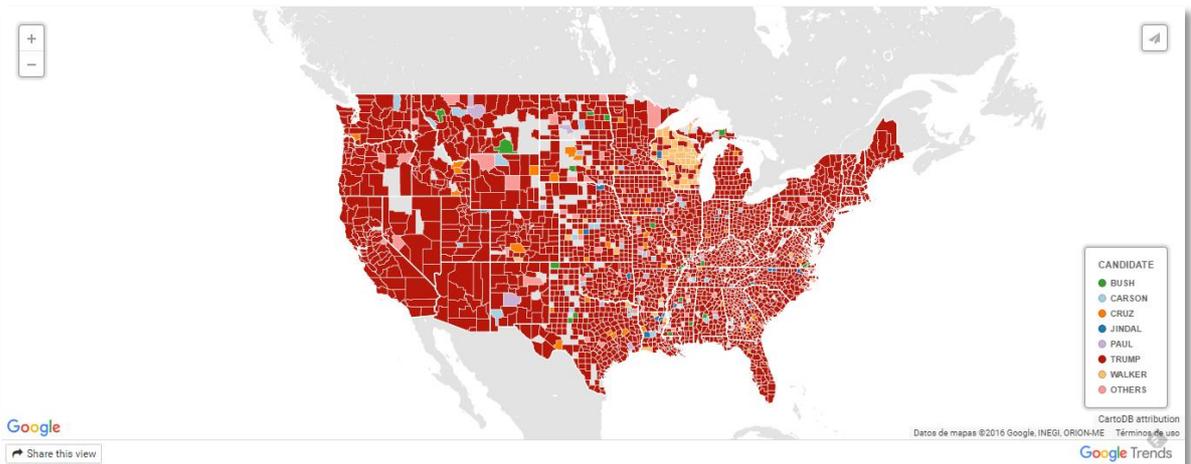


Figura 18 - Ejemplo de mapa temático desarrollado con Carto sobre las búsquedas en google de los candidatos republicanos a la presidencia de los EE.UU.

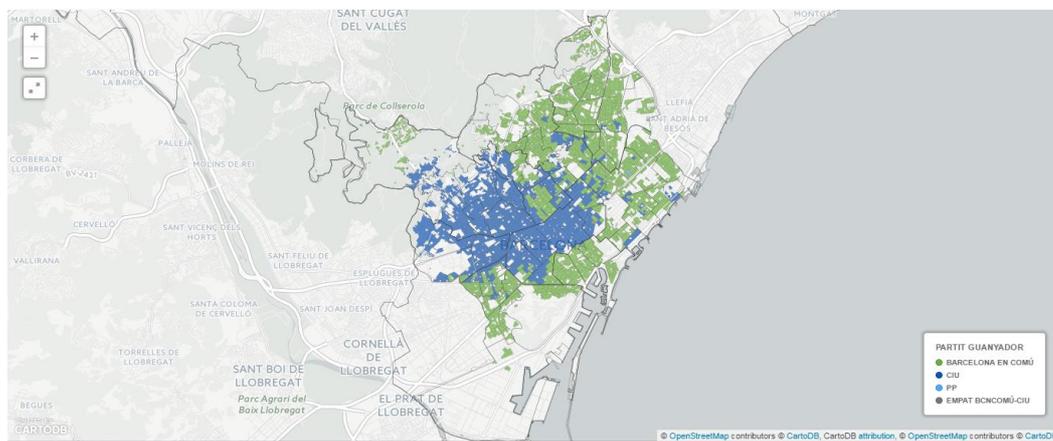


Figura 19 – Ejemplo de visualización de los resultados de las elecciones municipales de Barcelona de Mayo 2015

²⁷ <https://cartodb.com/>

3.2.5. Socrata

Socrata es una compañía que ofrece llevar a los usuarios no técnicos que quieran compartir información la posibilidad de crear fácilmente visualizaciones de estos datos. Fue fundada en 2007 y se especializa en el apoyo al sector público ofreciendo un servicio integral que va desde el apoyo para la **recolección de datos** hasta su **conversión en aplicaciones**. Actualmente está presente en varios portales de datos abierto como es el caso del portal de transparencia de datos del Ayuntamiento de Gijón (Figura 5).

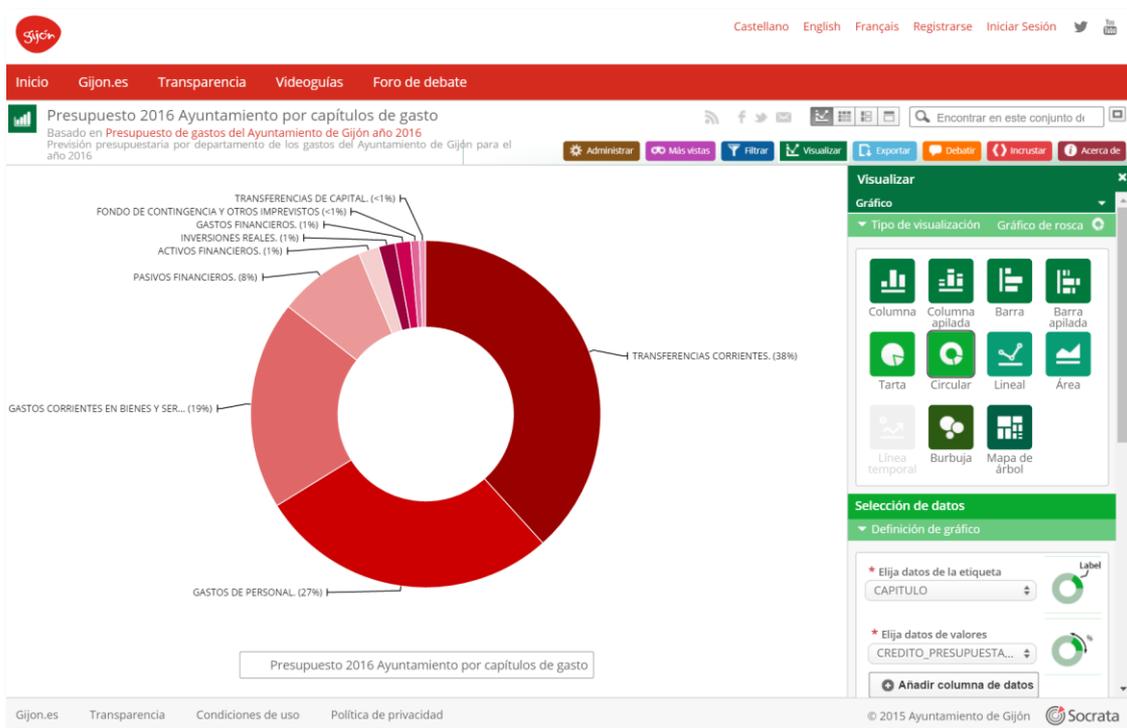


Figura 20 - Visualización de presupuesto del Ayto. de Gijón: <https://transparencia.gijon.es/>.

3.2.6. CKAN (Módulos de visualización)

La ampliamente conocida plataforma para la publicación de catálogos de datos, CKAN, ofrece un conjunto de módulos de visualización que permiten generar diferentes representaciones gráficas.

La **Table view** nos permite disponer de una versión tabular, en formato tabla, de los datos de un dataset en particular, ofreciendo **herramientas de listado** y **vista facetada**. También se dispone de un módulo para la visualización de datos en forma de gráfico, permitiendo la comparación de variables a través de series dentro de un mismo gráfico. De igual modo, si los datos contienen

datos georreferenciados se puede utilizar el módulo de mapas para crear visualizaciones interactivas.

Estos módulos de visualización dentro de la plataforma CKAN son Data Explorer, DataStore Grid, DataStore Graph, DataStore Map, Text view, Image view y Web page view. Y son instalados junto con la plataforma. Además existen numerosas extensiones que permite completar esta funcionalidad básica, como por ejemplo *ckanext-basiccharts*, *ckanext-dashboard*, *ckanext-map*, *ckanext-mapsearch*, *ckanext-mapviews*, etc.

4. EJEMPLOS DE VISUALIZACIÓN DE DATOS

En esta sección se describen ejemplos de referencia de visualización, tanto a nivel nacional como internacional, pioneras en el ámbito de datos abiertos. El objetivo es, por tanto, ilustrar a través de aplicaciones concretas la potencia que tiene la explotación gráfica para proporcionar valor a los datos desde el punto de vista del usuario (bien sea por el ciudadano o empleados de las propias administraciones). Estas visualizaciones se han construido utilizando librerías y herramientas presentadas en las secciones anteriores del documento.

4.1. Datos.gov.es

La iniciativa nacional que organiza y gestiona el **Catálogo Nacional de datos abiertos**, es el punto único de acceso a los conjuntos de datos que las Administraciones Públicas ponen a disposición para su reutilización en España. Este catálogo ofrece un conjunto de datos elaborados o custodiados por organismos públicos disponibles en formatos electrónicos y libres de derechos. El Catálogo Nacional ofrece multitud de datasets de una temática variada: transporte, industria, sanidad, educación, finanzas, medio ambiente, etc. La publicación de estos datos ofrece no sólo la posibilidad de acceso de los ciudadanos a un gran conjunto de información relevante, sino que posibilita también la creación de productos y servicios basados en dichos datos.

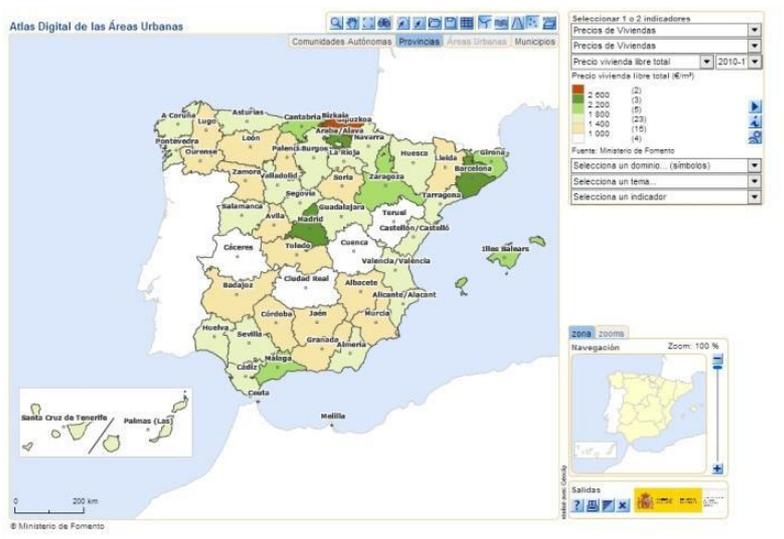


Figura 21 - Ejemplo de visualización de información urbanística de cada región española a partir de los datos del Ministerio de Fomento

4.2. Banco mundial²⁸

El Banco mundial es uno de los principales referentes en cuanto a la publicación de datos abiertos. Ofrecen un amplio catálogo de datos y la posibilidad de visualizar esos datos mediante gráficas y mapas. Uno de sus objetivos es combatir la pobreza en el mundo por lo que ofrecen gran cantidad de datos sobre el desarrollo de todos los países del mundo. De esta forma disponemos de una gran cantidad de datos sobre agricultura, economía, educación, salud, etc.

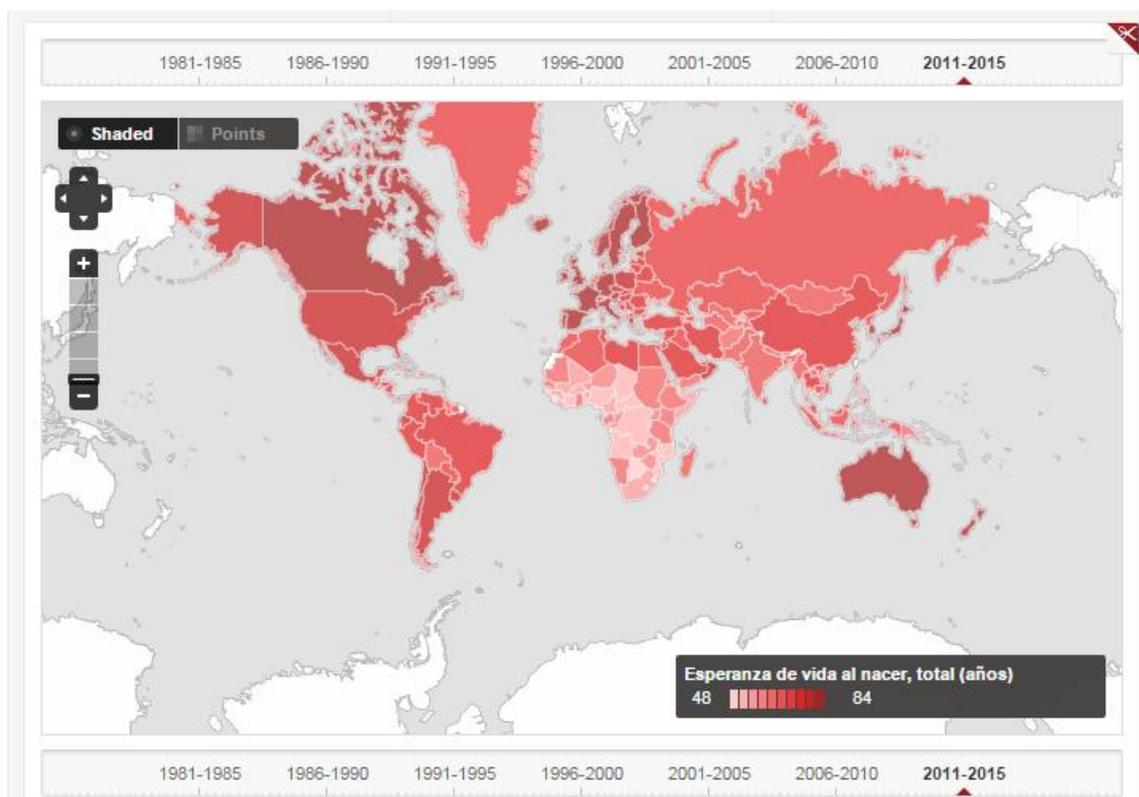


Figura 22 - Ejemplo de visualización del Banco Mundial. Esperanza de vida por países

4.3. Portal de transparencia del Principado de Asturias²⁹

El Gobierno del Principado de Asturias ofrece a sus ciudadanos un portal de transparencia con una gran variedad de datos y visualizaciones dinámicas. Estas visualizaciones, desarrolladas con

²⁸ <http://www.bancomundial.org/>

²⁹ <http://asturias.transparenciaendatos.es/>

la herramienta Tabulae, permiten al usuario explorar de forma sencilla información sobre asuntos sociales, economía y finanzas, empleo, medio ambiente y salud.

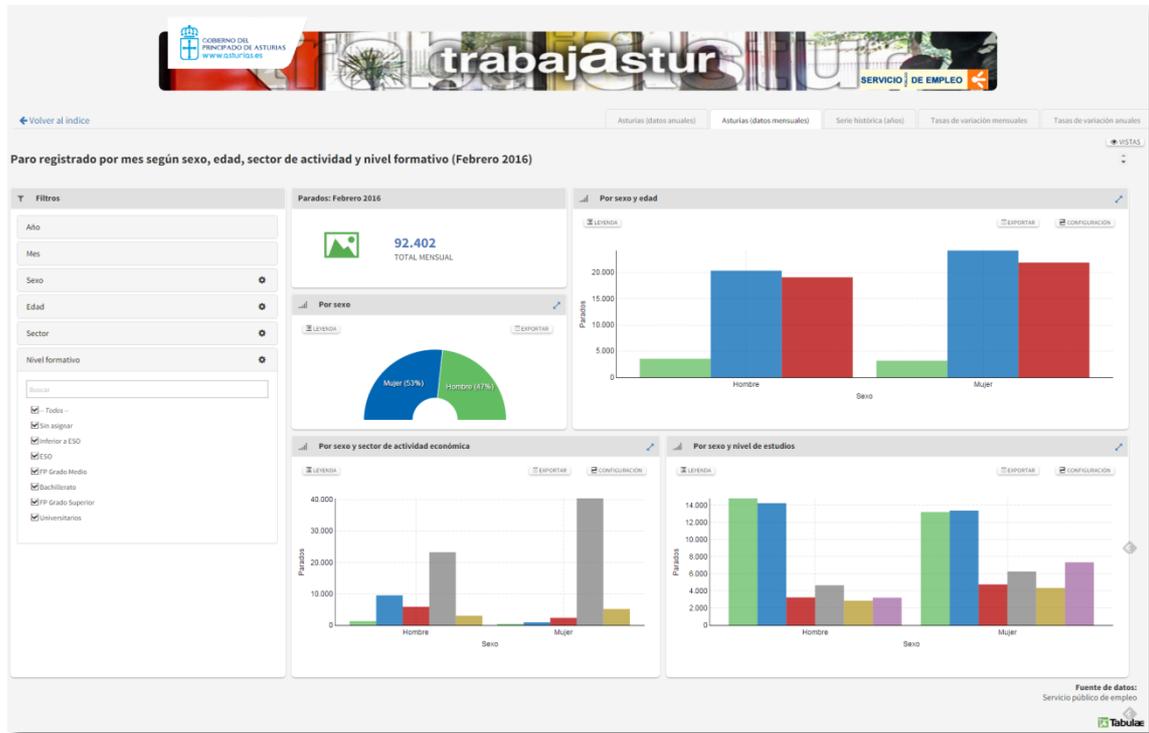


Figura 23 - Ejemplo de cuadro de mandos interactivo sobre datos de parados en el Principado de Asturias (Tabulae)

4.4. Plataforma openREDBAG³⁰

openREDBAG es un proyecto de la Asociación Ibero-Macaronésica de Jardines Botánicos (AIMJB), el MAGRAMA y la Fundación Biodiversidad para construir una **plataforma de acceso a datos sobre recursos fitogenéticos** silvestres conservados en bancos de germoplasma. Esta plataforma ofrece a los usuarios una aplicación que a través de los datos de la AIMJB y de GBIF (Global Biodiversity Information Facility, la base de datos abierta de biodiversidad a nivel mundial), los enriquece y permite visualizarlos mediante mapas y tablas información sobre la conservación de especies, catálogos de protección, especies amenazadas, etc. del territorio nacional.

³⁰ <http://www.redbag.es/openredbag/index.php>

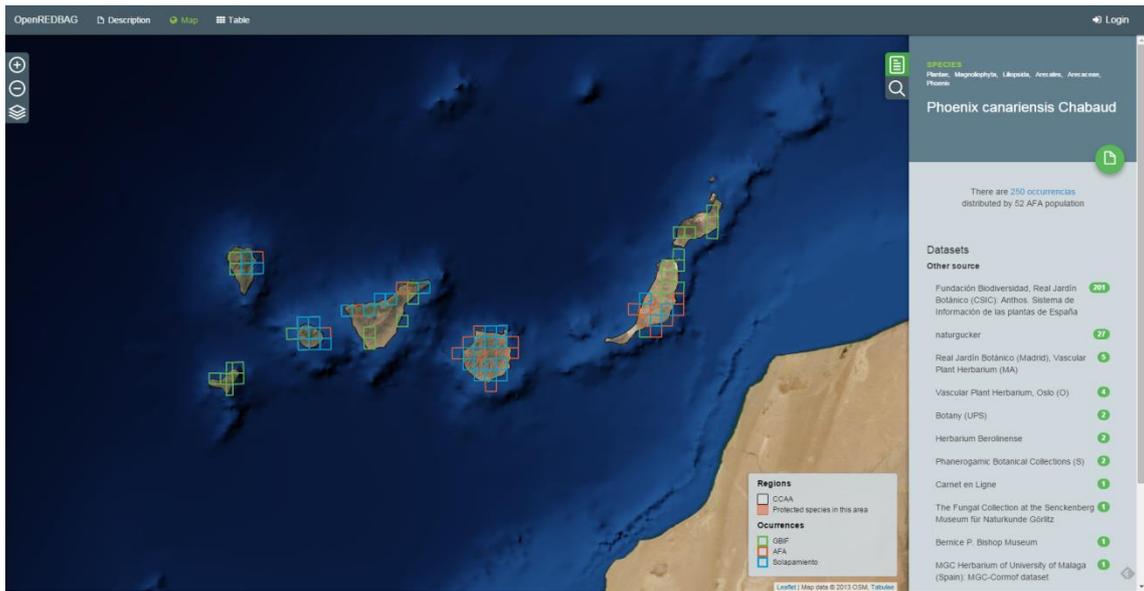
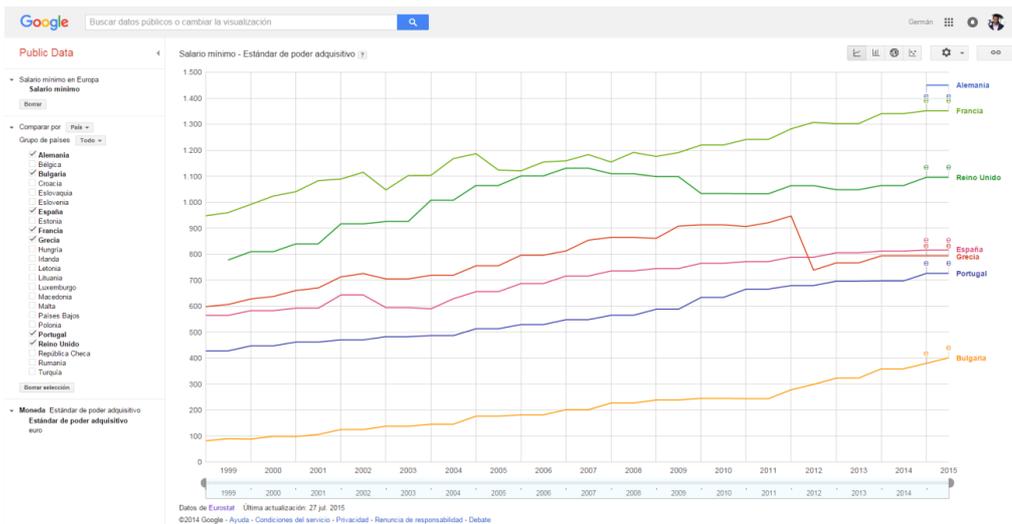


Figura 24 - Ejemplo de visualización de especies conservadas en bancos de germoplasma de la REDBAG

4.5. Google Public Data³¹

Google Public Data es un servicio de Google que permite explorar de forma visual gran variedad de conjuntos de datos de múltiples fuentes. Ofrece la posibilidad de visualizar tanto gráficas como mapas e incluso animar las visualizaciones a partir de la dimensión temporal.



³¹ <https://www.google.com/publicdata/directory>

Figura 25 - Ejemplo de visualización del salario mínimo por país a partir de los datos de Eurostat

4.6. Eurostat

Eurostat (Oficina Europea de Estadística), es la oficina estadística de la Comisión Europea. Su función principal es la de ofrecer información estadística sobre las instituciones de la Unión Europea y la armonización de los métodos estadísticos de los estados miembros. En la Web de Eurostat podemos encontrar multitud de datasets y visualizaciones sobre datos económicos, sociales, de industria, etc. tanto a nivel general como regional.

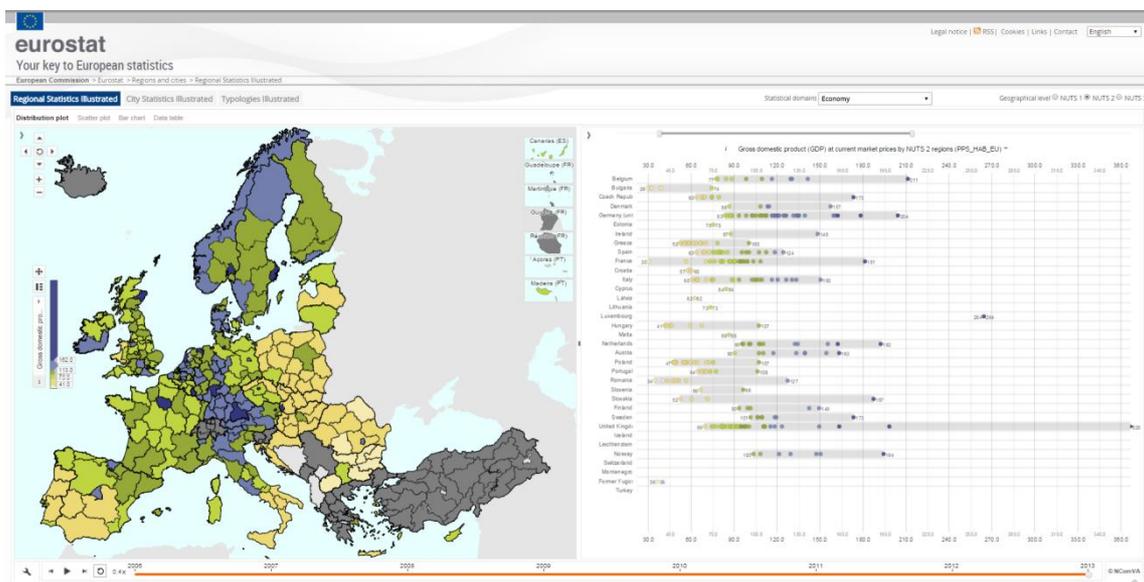


Figura 26 - Ejemplo de visualización de Eurostat sobre el producto interior bruto

4.7. Fundación Civio³²

La fundación Civio es una organización sin ánimo de lucro que pretende lograr una transparencia real y un acceso libre a los datos públicos tanto para los ciudadanos como para las organizaciones. Para ello desarrollan herramientas y generan informes que permiten sacar valor a los datos públicos y promover la transparencia. Civio utiliza **fuentes de datos públicas** como el INE o el BOE e investiga para generar información relevante sobre la gestión pública, el objetivo final es

³² <http://www.civio.es/>

la rendición de cuentas por parte de las instituciones para mejorar la democracia. Algunos de los proyectos más destacados de la Fundación Civio son:

- **¿Dónde van mis impuestos?** Permite a los ciudadanos conocer la distribución de los presupuestos regionales durante los años 2006 al 2015. El ciudadano puede ver la distribución de presupuestos por comunidades autónomas sobre el presupuesto total o su distribución por habitante, así como ver esa distribución por cada una de las áreas funcionales (sanidad, educación, cultura, etc.)

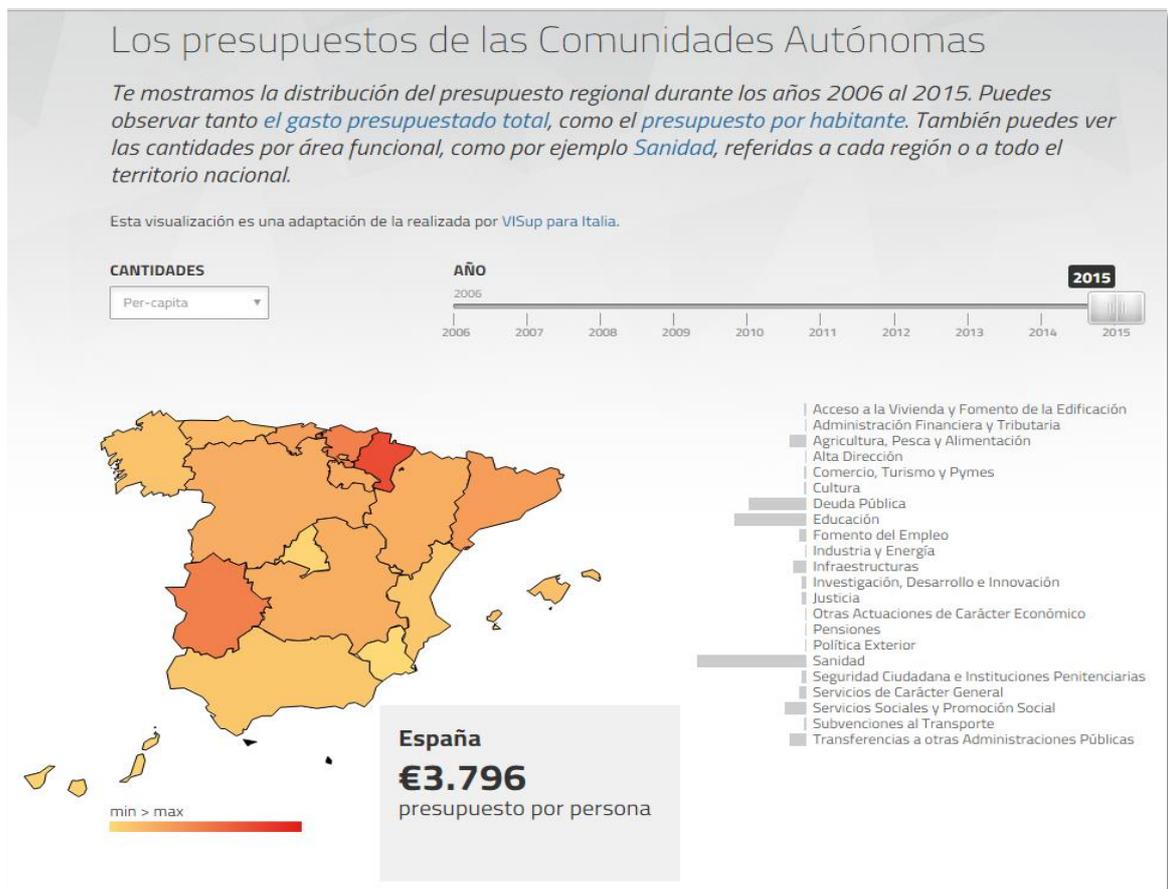


Figura 27 - ¿Dónde van mis impuestos? Ejemplo de visualización de la distribución de presupuestos creada a partir de los datos de Presupuestos autonómicos del Ministerio de Hacienda

España en llamas: La herramienta de España en llamas permite visualizar los datos de todos los incendios forestales ocurridos en España entre 2001 y 2013. España en llamas ofrece multitud de

datos relacionados con los incendios, como pueden ser la causa, número de hectáreas quemadas, si son intencionados o no, si han causado muertes, etc.

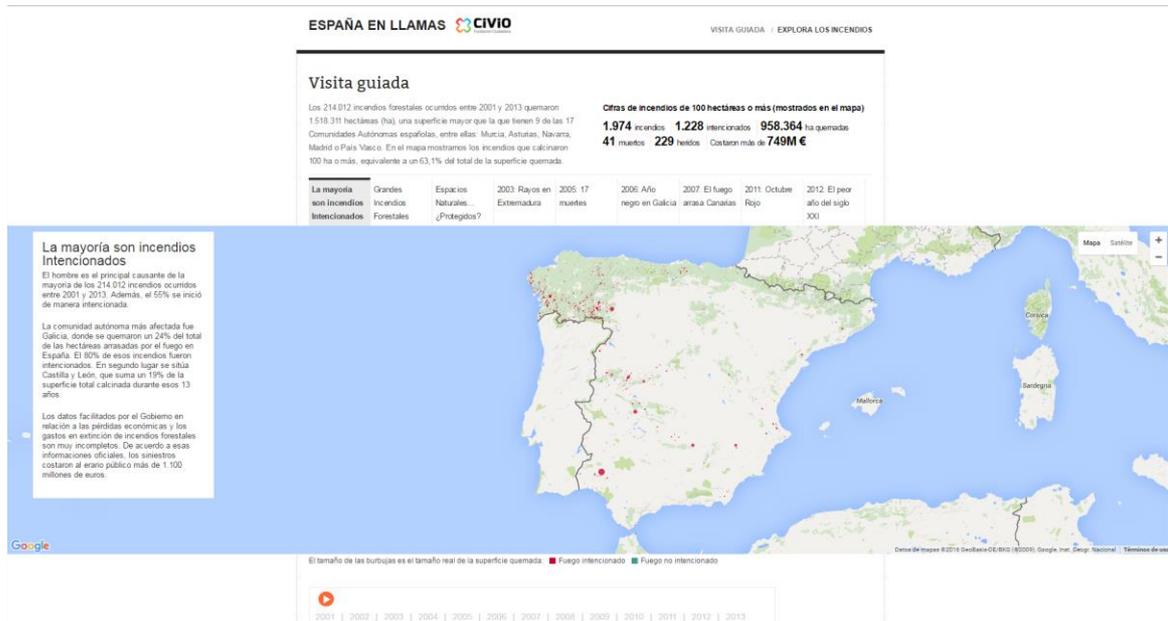


Figura 28 - España en llamas. Ejemplo de visualización de incendios forestales en España entre 2001 y 2013

- **El indultómetro:** El indultómetro recopila y clasifica toda la información publicada en el BOE sobre los indultos concedidos en España desde el año 1996.



Figura 29 - Ejemplo de visualización sobre una línea temporal de los indultos más destacados

4.8 Otros ejemplos de visualización

En los últimos años las visualizaciones interactivas han revolucionado la manera de comunicar en Internet. Además de instituciones e iniciativas son muchos los medios de comunicación, empresas e incluso usuarios que utilizan visualizaciones, infografías o cuadros de mando para comunicar. En muchos casos esas visualizaciones no están desarrolladas con librerías de componentes gráficos convencionales sino que están implementadas de una forma más *ad-hoc* y haciendo un especial hincapié en **buscar el impacto visual**. A continuación se muestran algunos ejemplos:

4.8.1. Evolución de la Web³³

Algunos de los miembros del equipo de desarrollo del navegador Google Chrome desarrollaron una atractiva visualización sobre la evolución de la Web.

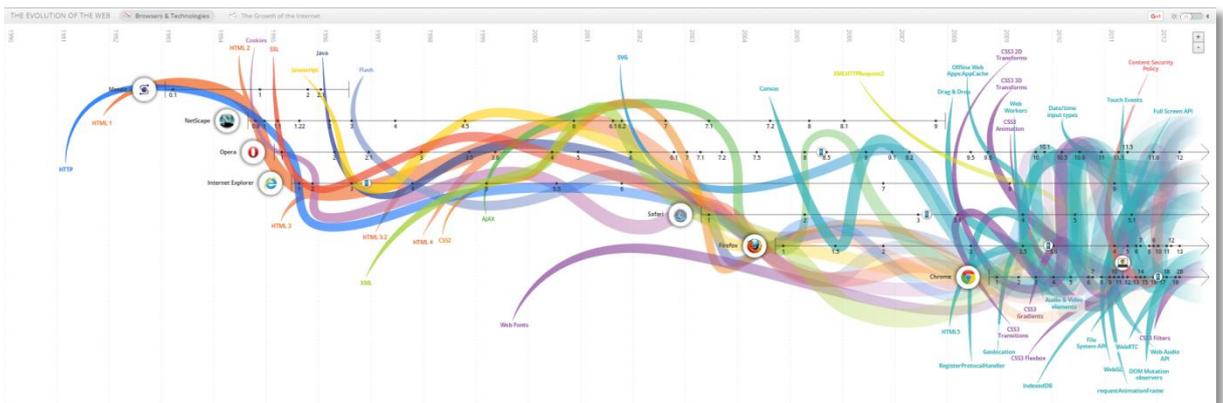


Figura 30 - Visualización sobre una línea temporal de la evolución de la Web

Esta visualización muestra la evolución de la web atendiendo a los diferentes estándares, navegadores y tecnologías que han ido apareciendo a lo largo de los años. El usuario puede interactuar con la visualización para resaltar cada uno de los elementos visualizados y acceder a imágenes y capturas de pantalla de las diferentes versiones de cada uno.

³³ <http://evolutionofweb.appspot.com/>

4.8.2. Calentamiento global³⁴

Ed Hawkins, científico y profesor del departamento de meteorología de la Universidad de Reading, publicó recientemente una impactante visualización sobre el aumento global de la temperatura desde 1850 hasta la actualidad.

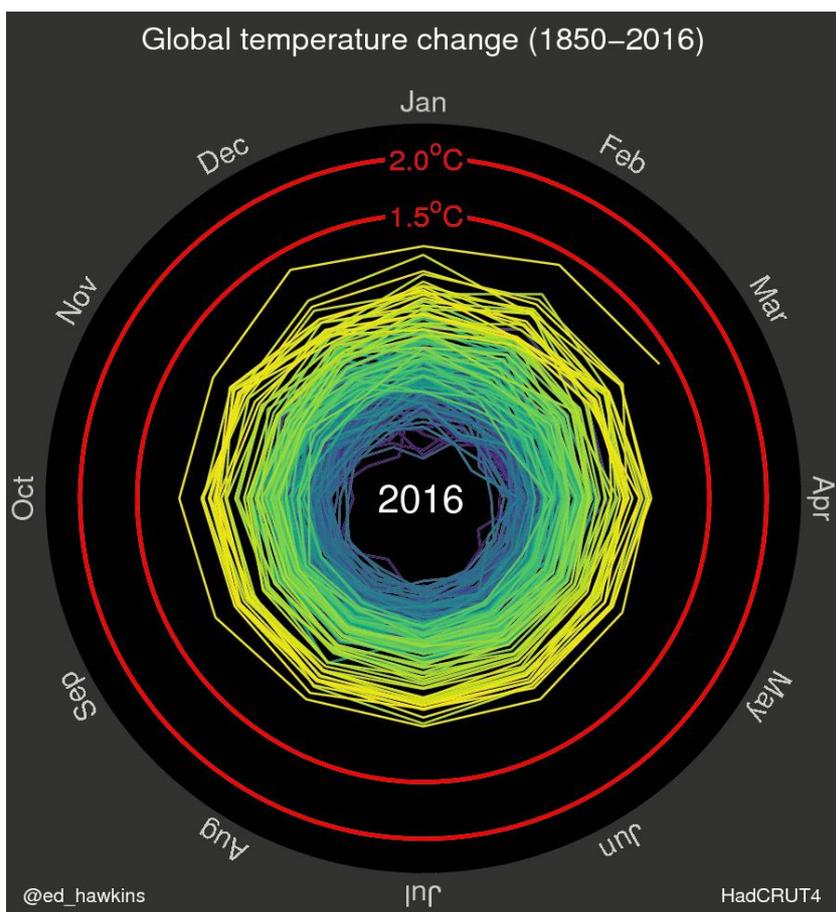


Figura 31 - Visualización sobre el crecimiento global de la temperatura desde 1850 hasta la actualidad

La visualización anterior muestra mediante una espiral la variación en la temperatura global del planeta. Se trata de una visualización animada en la que se simula el paso de los años y que permite apreciar cómo en algunos periodos, especialmente en los primeros años, la espiral muestra líneas muy juntas indicando variaciones pequeñas e incluso se contrae indicando un enfriamiento. Sin embargo con el paso del tiempo y especialmente en las últimas décadas se

³⁴ http://www.climate-lab-book.ac.uk/files/2016/05/spiral_optimized.gif

aprecia una mayor separación entre las líneas y cómo se van alejando cada vez más del centro, mostrando claramente que el calentamiento del planeta se acelera.

4.8.3. Retirada de Kobe Bryant³⁵

Con motivo de la retirada de Kobe Bryant, uno de los mejores jugadores de la historia de la NBA, el diario LA Times creó una fantástica visualización interactiva que muestra los más de 30.000 tiros a canasta que realizó durante su carrera deportiva.

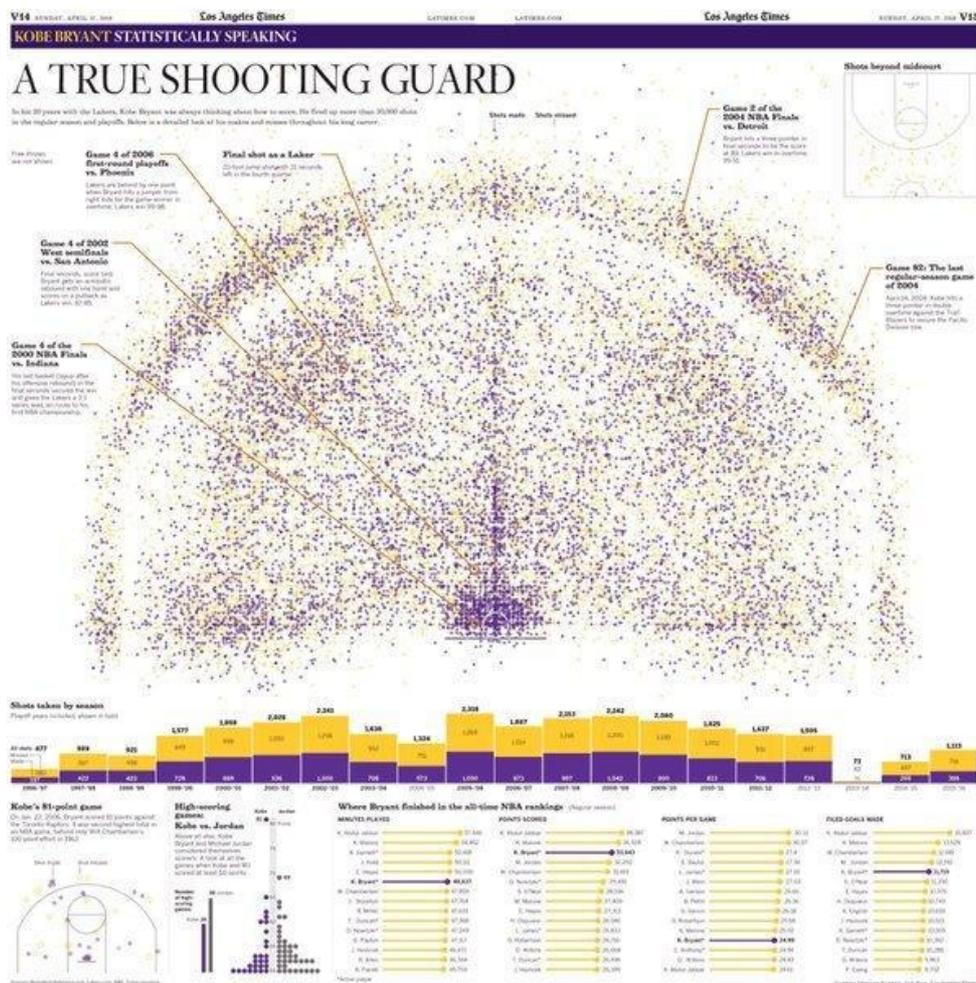


Figura 32 - Visualización que muestra todos los tiros realizados por Kobe Bryant durante su carrera deportiva

³⁵ <http://graphics.latimes.com/kobe-every-shot-ever/>

Esta visualización muestra sobre una pista de baloncesto la posición de cada tiro realizado por Kobe Bryan y permite a los lectores situarse sobre cada uno de esos puntos para ver su detalle (tipo de tiro, fecha, distancia al aro y rival). El color del punto indica si el tiro se convirtió en canasta o fue fallido.

4.8.4. Historia de la música³⁶

Google ha desarrollado una impresionante visualización que muestra la popularidad de cada género musical a lo largo de cada década desde 1950.

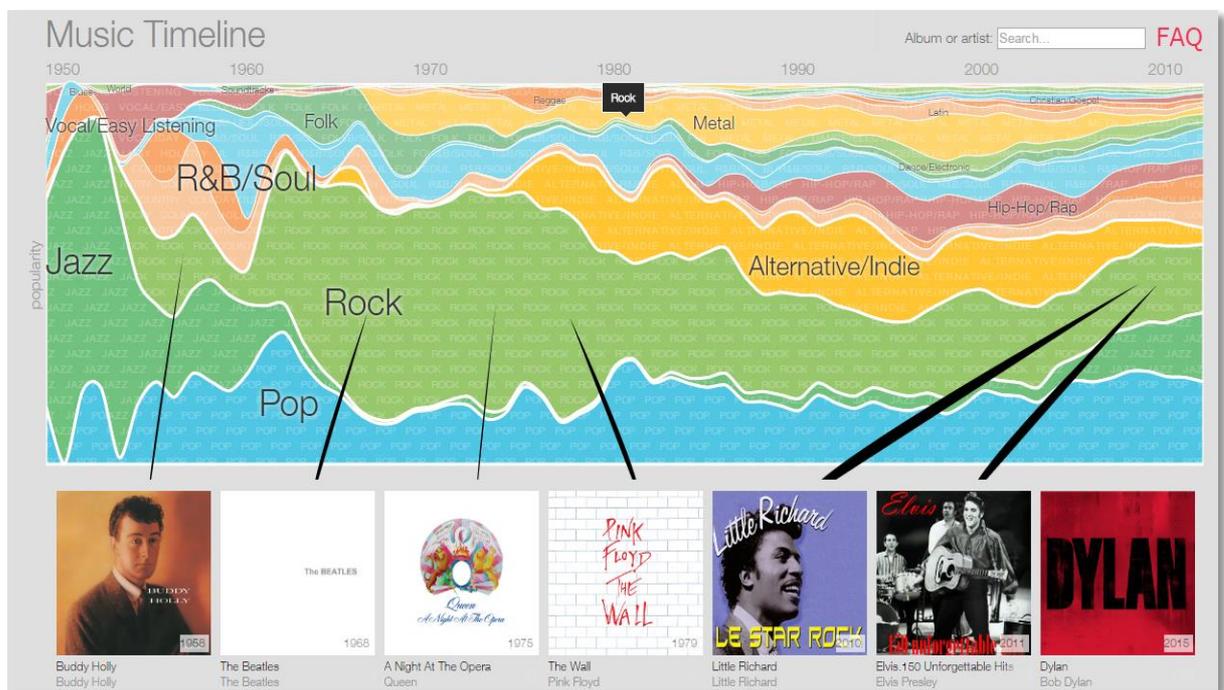


Figura 33 - Visualización que muestra los géneros musicales y su popularidad en el tiempo

Este ejemplo permite además a los usuarios interactuar para desglosar cada género musical en subgéneros, muestra también artistas y discos para cada una de las categorías y permite reproducirlos mediante el servicio musical de Google, comportándose como un original y atractivo buscador.

³⁶ <http://research.google.com/bigpicture/music/>

5. CONCLUSIONES

Una vez se ha avanzado de manera significativa en la publicación de datos abiertos: legislación, formatos, buenas prácticas, definición de vocabularios estándar, etc., el siguiente paso es facilitar la usabilidad y el acceso a los datos por parte de cualquier usuario. **La visualización de datos es uno de los mecanismos más potentes para explotar y analizar de manera autónoma el significado implícito en los datos**, independientemente del grado de conocimiento tecnológico del usuario. La visualización nos permite construir significado sobre los datos y la creación de narrativas basadas en la representación gráfica.

La visualización de datos, desde un punto de vista formal y científico, es una disciplina madura. Además su grado de implantación en la industria es muy amplio, desde el ámbito industrial y de proceso, hasta aspectos más ligados al *reporting* financiero y la gestión de negocio. Su aplicación es transversal a cualquier proceso en el que existan datos, como es el caso del movimiento open data.

Esto supone una gran ventaja desde el punto de vista de introducción de herramientas de explotación de datos basados en visualización. Se conocen perfectamente los principios metodológicos y de usabilidad. Además existe actualmente una gran variedad de software disponible para el desarrollo de proyectos de visualización dentro de los portales open data.

En este informe hemos centrado nuestro análisis en dos bloques claramente diferenciados:

1. **Tecnología de visualización.** Frameworks y bibliotecas de programación que permiten la construcción de aplicaciones y servicios basados en datos. El informe se centra exclusivamente en tecnología web de última generación y especialmente en aquella que se construye sobre **estándares**, como es el caso de HTML5 (Canvas), SVG y WebGL.
2. **Plataformas de visualización.** Aplicaciones que permiten la construcción de cuadros de mando y visualizaciones interactivas completas (como son el caso de los mapas de CartoDB). Se presentan más en detalle aquellas plataformas más orientadas al mundo web, que se pueden dividir en dos grandes bloques. Por un lado, las que vienen del **mundo del BI y análisis de datos**, y que se pueden aplicar a la explotación de datos

abiertos: Tableau Soft., Qlik y Tabulae. Por otro, las herramientas de publicación open data que incorporan ciertas funcionalidades de visualización: CKAN y Socrata.

5.1. Tendencias futuras

Las tecnologías y técnicas relacionadas con la visualización de datos siguen evolucionando de manera significativa. Este informe es un estado del arte que refleja qué alternativas más relevantes se encuentran en el mercado para trabajar en esta disciplina. Desde el punto de vista de la innovación y líneas futuras, existen varios caminos que se están siguiendo tanto en la industria del software como desde la parte más académica, entre las que destacan:

1. Visualización y grandes volúmenes de información

Un problema relacionado con la visualización es cuando se trabaja con grandes volúmenes de datos. Este problema no existe sólo a nivel de back-end, donde está más vinculado al análisis y tratamiento de grandes volúmenes de información estructurada y no estructurada (ie., Big Data), si no con la capacidad del navegador para soportar grandes conjuntos de datos (por ejemplo, cientos de miles de registros). Las restricciones de computación de los clientes en los que se ejecuta la visualización pueden ser una barrera para la construcción de ciertas visualizaciones. Este es un campo de investigación en el que se está trabajando actualmente.

2. Visualizaciones que se puedan construir por el usuario experto en datos, e incluso por el usuario final (en el caso de open data, por el propio ciudadano).

Muchas herramientas no son accesibles más que para perfiles técnicos, ya sean a nivel informático o de tratamiento y análisis de datos. Esta barrera de la tecnología es uno de los grandes desafíos actuales: conseguir que el usuario que entiende la información tenga la autonomía suficiente para construir a través de asistentes intuitivos (WYIWYG) sus propias visualizaciones y explotar la información de manera dinámica.

3. Visualizaciones 3D

En los últimos años han aparecido, como se ha identificado en el informe, tecnología de aceleración de gráficos 3D para el entorno web. La aplicación de esta tecnología en la visualización de datos todavía está en un estado incipiente. La mayor parte de las herramientas siguen trabajando con perspectiva 2D. Es de esperar que según la tecnología alcance un estado

de madurez más sólido, aparezcan nuevas técnicas de visualización integradas con las herramientas de usuario.

4. Visualización de datos y realidad aumentada

Un ámbito todavía poco explorado en la actualidad, pero que seguro supondrá una ruptura en la forma en cómo se consumen los datos, es la tecnología de realidad aumentada, que a través de dispositivos combina información real y física en un entorno interactivo para el usuario. Actualmente el estado del arte de la visualización se concentra en un consumo clásico y multidispositivo de los datos: ya sean aplicaciones web o aplicaciones móviles de propósito específico. La tecnología de realidad aumentada, con sus posibilidades de representación 3D y en tiempo real, es un territorio desconocido en este momento sobre sus potenciales aplicaciones e impacto en la disciplina de la visualización. Sin embargo, abre un vasto espacio para la imaginación y para diseñar nuevas soluciones en un entorno hasta hace poco “futurista”.