

# Visualización de datos

*Big data e inteligencia artificial*



tech

# CONTENIDO

## 1. Objetivo

---

## 2. La importancia de hacer una visualización de datos

---

## 3. Herramientas para llevar a cabo visualizaciones de datos

---

*Tableau*

D3

*Python (Matplotlib)*

*R(Shiny)*

Otros

## 4. Ejemplo de *Dashboard*

---

Análisis del turismo y transporte (*Tableau*)

Venta al por mayor (*Tableau*)

Educación (*Tableau*)

Análisis de la atención médica (*Tableau*)

Seguros (*Tableau*)

## 5. Bibliografía

---

## OBJETIVO

Conocer la importancia de realizar visualizaciones de datos, así como herramientas para su correcto desarrollo.

## LA IMPORTANCIA DE HACER UNA VISUALIZACIÓN DE DATOS

La visualización de datos se ha convertido en un punto clave para comunicar, impactando en las decisiones de las empresas. Los datos son una información, en ocasiones en cantidades ingentes y difíciles de entender, que circulan de forma física y virtual por el mundo. Entender e interpretar estos datos brinda herramientas para mejorar el trabajo o la vida; ayuda a tomar decisiones.

En todo proceso de decisión, visualizar correctamente no solo los datos, sino su correcto significado, cambiará la experiencia como usuarios, clientes o consultores. Leer nunca será lo mismo que interactuar con la información.

Los gráficos e infografías ofrecen una comprensión difícil de igualar. Visualizar un mapa, una tabla o una lista de ideas ordenadas, mejora el entendimiento. Por eso, en la era de la información, visualizar correctamente la información es vital para todos.

Lo más importante de obtener ingentes cantidades de datos, información masiva, estadísticas corporativas, etc., es que esta información sea posteriormente comprensible, a pesar de contener millones de cifras.

Para conseguir este objetivo, se requiere de *software de análisis de datos*. Este tipo de herramientas ayudan a construir contenido visual o multimedia a partir de tablas de datos.

## HERRAMIENTAS PARA LLEVAR A CABO VISUALIZACIONES DE DATOS

### TABLEAU

*Tableau Software* es una empresa de *software* con su sede principal en Seattle, Estados Unidos, la cual desarrolla productos de visualización de datos interactivos que se enfocan en inteligencia empresarial.

- ¿Qué es *Tableau*?

*Tableau* es una plataforma de análisis visual que transforma la manera en que se usan los datos para resolver problemas. Además, permite a las personas y las organizaciones sacar el máximo partido de los datos

Únicamente requiere la inversión inicial en la licencia que más se ajuste a las necesidades. El nivel de complejidad de los análisis que permite es ilimitado y muy intuitivo. Técnicamente, *Tableau* se integra con casi todas las fuentes de datos existentes.

Hace que sea más fácil explorar y administrar los datos. Asimismo, permite descubrir y compartir información más rápidamente a fin de generar grandes cambios en los negocios y en el mundo.

### D3

D3.js es una biblioteca de *JavaScript* para manipular documentos basados en datos. D3 le ayuda a dar vida a los datos mediante HTML, SVG y CSS. El énfasis de D3 en los estándares *web* le brinda todas las capacidades de los navegadores modernos sin atarse a un marco propietario, combinando poderosos componentes de visualización y un enfoque basado en datos para la manipulación de DOM.



Como todas las herramientas sofisticadas, no es fácil de aprender, especialmente si no se tienen conocimientos de SVG, pero se puede destacar la inmensa comunidad que está detrás de este proyecto, apoyándolo, publicando abundante documentación: tutoriales, ejemplos, etc. Aun así, se hace hincapié: tener una base de *JavaScript*, *HTML5*, *DOM* y *CSS* y *SVG* mejorará la curva de aprendizaje.

- **Conceptos básicos de SVG utilizados en la visualización de D3.js**

- **Sistema coordinado:** en un sistema de coordenadas matemático normal, el punto  $x = 0$ ,  $y = 0$  está en la esquina inferior izquierda de la gráfica. Pero en el sistema de coordenadas SVG, este punto (0,0) está en la esquina superior izquierda del 'lienzo', es similar a CSS cuando se especifica la posición en absoluto / corregir y se usa la parte superior e izquierda para controlar el Punto exacto del elemento.

- **losElemento:** `<rect>` representa un rectángulo, aparte de las propiedades estéticas como trazo y relleno, el rectángulo se definirá por ubicación y tamaño.

En cuanto a la ubicación, está determinada por los atributos  $x$  y  $y$ . La ubicación es relativa a la matriz del rectángulo. Y si no especifica el atributo  $x$  o  $y$ , el valor predeterminado será 0 en relación con el elemento padre.

Después de especificar la ubicación, o más bien el 'punto de inicio' del rect, lo siguiente es especificar el tamaño, que es esencial si realmente desea dibujar algo en el lienzo, es decir, si no lo hace especifique los atributos de tamaño o el valor se establece en 0, no verá nada en el lienzo.

- **los:** `<svg>` elemento `<svg>` es el elemento raíz, o el lienzo, ya que se dibujan gráficos en él.

Los elementos SVG se pueden anidar uno dentro del otro y, de esta manera, las formas SVG se pueden agrupar, mientras tanto, todas las formas anidadas dentro de un elemento se posicionarán con relación a su elemento de encierro.

Algo que podría ser necesario mencionar es que no se puede anidar dentro de otra, no funcionará.

## PYTHON (MATPLOTLIB)

*Matplotlib* es una biblioteca completa para crear visualizaciones estáticas, animadas e interactivas en *Python*. Genera gráficos a partir de datos contenidos en listas o arrays y su extensión matemática *NumPy*.

- **Gráficos interesantes de muestra en Matplotlib**

- **Imágenes:** *Matplotlib* puede mostrar imágenes (asumiendo dimensiones horizontales igualmente espaciadas) usando la función `imshow()`.

- **Rutas:** puede agregar rutas arbitrarias en *Matplotlib* usando el módulo `matplotlib.path`.



- **Trazado tridimensional:** el *kit* de herramientas *mplot3d* (consulte introducción y trazado en 3D) admite gráficos simples en 3D, incluidos gráficos de superficie, de estructura alámbrica, de dispersión y de barras.
- **Streamplot:** la función *streamplot()* traza las líneas de corriente de un campo vectorial. Además de simplemente trazar las líneas de corriente, le permite asignar los colores y / o anchos de línea de las líneas de corriente a un parámetro separado, como la velocidad o la intensidad local del campo vectorial.
- **Elipses:** en apoyo de la misión *Phoenix* a Marte (que usó *Matplotlib* para mostrar el seguimiento terrestre de la nave espacial), Michael Droettboom se basó en el trabajo de Charlie Moad para proporcionar una aproximación extremadamente precisa de ocho *splines* a los arcos elípticos, que son insensibles al nivel de zoom.
- **Barras:** utilice la función *bar()* para crear gráficos de barras, que incluye personalizaciones como barras de error.
- **Notación TeX para objetos de texto:** a continuación, se expone una muestra de las muchas expresiones *TeX* que ahora son compatibles con el motor de texto matemático interno de *Matplotlib*. El módulo de texto matemático proporciona expresiones matemáticas de estilo *TeX* utilizando *FreeType* y las fuentes *DejaVu*, *BaKoMa computer modern* o *STIX*.

## R(SHINY)

*Shiny* es un paquete de R que facilita la creación de aplicaciones web interactivas directamente desde R. Puede alojar aplicaciones independientes en una página web o incrustarlas en documentos de *R Markdown* o crear paneles de control. También puede ampliar sus aplicaciones *Shiny* con temas CSS, *htmlwidgets* y acciones de *JavaScript*.

*Shiny* combina el poder computacional de R con la interactividad de la web moderna.

Las aplicaciones *Shiny* son fáciles de escribir. No se requieren habilidades de desarrollo web.

- **Hospedaje e implementación**

- **Implementar en la nube**

- » **Shinyapps.io:** aloje sus aplicaciones *Shiny* en la web en minutos con *Shinyapps.io*. Es fácil de usar, seguro y escalable. No se requiere hardware, instalación o contrato de compra anual. Opciones gratuitas y de pago disponibles.

- **Implementar localmente (código abierto)**

- » **Shiny server:** implemente sus aplicaciones *Shiny* y documentos interactivos en las instalaciones con *Shiny Server* de código abierto, que ofrece características como múltiples aplicaciones en un solo servidor e implementación de aplicaciones detrás de *firewalls*.

- **Implementar localmente (comercial)**

- » **RStudio Connect:** es la plataforma de publicación insignia para el trabajo que crean sus equipos en R. Con *RStudio Connect*, puede compartir aplicaciones *Shiny*, informes de *R Markdown*, cuadros de mando, gráficos y más en un lugar conveniente con la publicación de botones desde *RStudio* IDE. Las características incluyen la ejecución programada de informes y políticas de seguridad flexibles para llevar el poder de la ciencia de datos a toda su empresa.

## OTROS

- **Power BI (Microsoft)**

*Power BI* es un programa que se parece a *Excel* en términos de interfaz. Se integra a la perfección en una infraestructura existente de *Microsoft*. Requiere la descarga del *software* que se ajuste al sistema operativo y cuenta con una versión gratuita que evidentemente puede ser ampliado con paquetes de pago personalizados. Como *Microsoft 365* incluye *Power BI*, se puede realizar un primer proyecto de *business intelligence* con muy poco capital. Es intuitiva y potente, pero no tan sencilla de utilizar como otras.

Existen tres tipos de usuarios de *Power BI*:

- **Creadores de contenido:** los creadores de contenido crean visualizaciones de datos y elaboran informes. Usan la herramienta *Power BI Desktop* gratuita para preparar, modelar y crear visualizaciones de datos.
- **Editores de contenido:** los editores de contenido precisan colaboración, modelado de datos, creación de contenido, uso compartido de paneles, análisis *ad hoc* y publicación de informes. Estos usuarios necesitan una licencia de *Power BI Pro*.
- **Consumidores de contenido:** los consumidores de contenido usan paneles e informes publicados previamente. Existen dos opciones para proporcionar a los consumidores acceso a contenido: conceder una licencia a todos los usuarios de *Power BI Pro* o bien conceder una licencia de *Power BI Premium* para bases de usuarios a gran escala.

- **Data Studio (Google)**

Herramienta gratuita y sencilla de utilizar. Una de sus grandes virtudes es que se conecta con otras herramienta como *Google Analytics* o *Adwords* y, mediante pago, también puede hacerlo con otras como Facebook. La integración/conexión con los datos es un poco limitada a productos Google, de momento, solo la conexión con *MySql*, *PostgreSQL* o subidas de ficheros CSV son la excepción.

Se accede a través del navegador sin necesidad de instalar *software* adicional.

## EJEMPLO DE DASHBOARD

### ANÁLISIS DEL TURISMO Y TRANSPORTE (TABLEAU)

En los sectores del turismo y el transporte siempre abundaron los datos, ya que se recopilan en cada punto de interacción. Tanto si se trata de pasajeros como de cargas, los datos ayudan a equilibrar la oferta y la demanda, con el fin de optimizar los ingresos y la rentabilidad del inventario y las capacidades.

El negocio de los viajes y la hospitalidad incluye aerolíneas, hoteles, compañías de alquiler de autos y empresas de cruceros. Todos ellos capturan comportamientos y experiencias de los clientes que contienen información valiosa para el negocio. Pero los métodos de análisis del sector de turismo y transporte suelen ser lentos y no siempre son intuitivos.

Al mismo tiempo, las empresas de logística y transporte, como transportistas de camiones, trenes de carga, entregas aéreas, marítimas y proveedores de servicios logísticos, intentan vincular fuentes de datos diversas para administrar sus redes y planificación de capacidades, además de optimizar las rutas para una mayor rentabilidad.

Con *Tableau*, puede combinar y vincular con rapidez datos similares a fin de obtener la información necesaria para mejorar la administración y el rendimiento de los ingresos. Esto se consigue gracias a estrategias de precios basadas en los datos y la visibilidad del inventario en tiempo real. Además, todo ello ayuda a generar fidelidad en los clientes. Descubra en esta sección cómo *Tableau* ayuda a mejorar los análisis de turismo y transporte.

### VENTA AL POR MAYOR (TABLEAU)

Debido a los márgenes ajustados, los mercados volátiles y las exigencias cambiantes de los clientes, los comerciantes minoristas y mayoristas no pueden confiar solo en la intuición para generar crecimiento y adaptarse a los cambios. Necesitan información basada en datos.

Los comercios minoristas eligen *Tableau* con el propósito de obtener la información de análisis visual necesaria para responder a las demandas del sector. Más del 80 % de los 100 principales comercios minoristas y más de 7000 empresas de productos minoristas y de consumo de todo el mundo confían en *Tableau* para entender sus datos y tomar medidas en consecuencia.

### EDUCACIÓN (TABLEAU)

Las escuelas y las instituciones necesitan los datos para mejorar. Desde la educación temprana, hasta la matriculación y las becas, las fuentes de datos del mundo de la educación son enormes.

Cumplir con los requisitos de análisis y generación de informes en todos los departamentos y campus puede suponer todo un reto. Llevar a cabo el análisis de la educación con *Tableau* puede ser la solución a estos problemas. *Tableau* facilita esta tarea, ya que permite que cualquier persona realice un análisis de educación complejo y comparta sus descubrimientos.

Explore cómo los planes de estudio se relacionan con el rendimiento de los estudiantes. Descubra cómo aumentar la participación de los niños en riesgo de abandono escolar. Encuentre la manera de hacer más eficiente el proceso de inscripción. *Tableau* también puede ser una herramienta eficaz en el aula, ya que permite a los estudiantes y los profesores ver y comprender sus datos. En un mundo en el que el análisis de datos es un activo cada vez más valioso, resulta fundamental que todos los miembros del sistema educativo se sientan a gusto con los datos. Por esta razón, *Tableau* se compromete a que todos en el ámbito de la educación puedan acceder a herramientas de análisis de primer nivel. Los estudiantes pueden disfrutar de *Tableau* gratis, los profesores tienen disponible una gran variedad de recursos y los administradores pueden obtener descuentos especiales.



## ANÁLISIS DE LA ATENCIÓN MÉDICA (TABLEAU)

Los datos de atención médica pueden llegar a reducir costos, aumentar la calidad del servicio y mejorar la experiencia del paciente. El desafío es descubrir cómo pasar de los datos a la información y, después, a la acción. *Tableau* motiva a todos en las organizaciones de atención médica a encontrar el camino para avanzar.

En un sector en el que cada segundo cuenta, *Tableau* permite conectarse a los datos, visualizarlos, analizarlos y encontrar las respuestas mucho más rápido.

## SEGUROS (TABLEAU)

Pólizas. Primas. Reclamaciones. Indemnizaciones. En el mundo de los seguros, cada transacción es un punto de datos listo para el análisis y la acción. *Tableau* le ofrece un análisis de seguros visual para todos los miembros de su aseguradora.

Los complejos conjuntos de datos sobre seguros no representan ningún problema para *Tableau*, que permite explorarlos en tiempo real y reaccionar con mayor velocidad ante los cambios.

## BIBLIOGRAFÍA

- [1] H. Jones, *Data analytics: an essential beginner's guide to data mining, data collection, big data analytics for business, and business intelligence concepts*.
- [2] C. Nussbaumer Knaflic, *Telling with data: a data visualization guide for business professionals*.
- [3] J. Berengueres, *Introduction to data visualization & storytelling: a guide for the data scientist (English edition)*.