



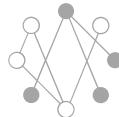
Bloque 1 • Datos por todas partes

Unidad 2 • Píldora 1

LA QUINTA UVE DEL BIG DATA



LA QUINTA UVE DEL BIG DATA



Tenemos una gran cantidad de datos a nuestro alcance y, con su análisis, podemos obtener un conocimiento muy valioso. Sin embargo, cuando manejamos un gran volumen de datos masivos, son indispensables herramientas que nos faciliten este análisis. Es aquí donde entra en juego la quinta uve del *Big Data*: la visualización.

VISUALIZAR PARA COMPRENDER

La visualización de datos es una técnica para poder leer, analizar y comparar datos a partir de su representación gráfica. Se utiliza en ámbitos de todo tipo como la educación, la ciencia o los negocios y resulta indispensable para obtener conclusiones válidas.

Piensa que la información se almacena en bases de datos. A simple vista estas bases de datos tienen el aspecto de hojas de cálculo interminables, llenas de textos y cifras. Mirando solo filas y columnas es casi imposible observar relaciones significativas entre unos datos y otros, o determinar qué elementos predominan. Por ello es necesario traducir estos datos a modelos gráficos que sí nos permitan hacer una lectura útil de la información y extraer conclusiones que muchas veces nos pasarían desapercibidas.

La visualización es una herramienta eficaz que puede aplicarse a cualquier tipo de estudio, desde identificar las relaciones entre participantes en una red social como Facebook hasta entender cómo se propagan enfermedades y epidemias en un territorio, o cómo ha evolucionado la esperanza de vida a lo largo del tiempo.



fuente: pexels

DE LAS HOJAS DE CÁLCULO A LOS SERVICIOS ON-LINE DE VISUALIZACIÓN

Siempre han existido datos y siempre se ha tratado de visualizarlos. Muchos programas de hojas de cálculo (como Microsoft Excel) incorporan desde hace tiempo herramientas para crear gráficos a partir de los datos de nuestras tablas, para poder cotejarlos de una manera visual.



Unidad 2º Píldora 1

• LA QUINTA UVE DEL BIG DATA

Actualmente contamos con muchas más formas de generar y acceder a un volumen enorme de datos y, al mismo tiempo, tenemos dispositivos más potentes que nos permiten procesarlos de manera más fácil y rápida. Por lo tanto, es lógico que hayan surgido nuevos tipos de software y lenguajes de programación para poder visualizar mejor la información con la que contamos.

Tenemos, por ejemplo, plataformas *on-line*, como [Infogr.am](#) o [Datawrapper](#), con las que podemos crear gráficos y visualizaciones. Algunas de ellas utilizan información disponible en bases de datos abiertas, como [Google Public Data](#), y otras, además, permiten incorporar información de bases de datos propias, como [Analytics de IBM](#) o [Tableau](#).

EL MEJOR GRÁFICO PARA NUESTROS DATOS

A la hora de generar visualizaciones con estos servicios es importante saber qué tipo de gráfico nos ayuda mejor a explicar la información que tenemos. Por ejemplo, si los datos contienen un elemento geográfico asociado, puede que lo ideal sea mostrar nuestra información sobre un mapa, o, si los datos explican la evolución temporal de un fenómeno, la mejor visualización puede ser una línea de tiempo.

Existen algunos recursos útiles para identificar qué tipo de gráfico o visualización es la idónea en función de los datos a mostrar. [Vizipedia](#) es una buena recopilación *on-line* de visualizaciones organizadas según la tipología de datos que representan.

