

Taller 2

1. Big Data, Ciencia de Datos y Visualización

La visualización de datos juega un papel fundamental dentro del Big Data, la Ciencia de Datos, el Análisis de Datos y la Inteligencia de Datos. En escenarios de Big Data, caracterizados por grandes volúmenes de información, alta velocidad de generación y variedad de formatos, la visualización permite transformar datos complejos en representaciones comprensibles para los usuarios.

En la Ciencia de Datos, la visualización es clave durante el análisis exploratorio, ya que ayuda identificar patrones, tendencias y anomalías. Además, facilita la comunicación de resultados a tomadores de decisiones.

2. Niveles de abstracción de la información

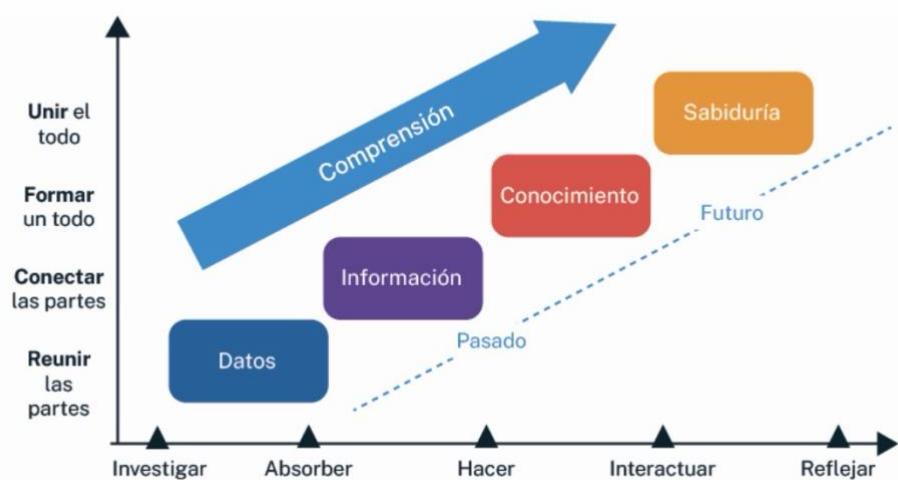
La jerarquía DIKW describe la transformación progresiva del dato hasta convertirse en sabiduría.

Dato: valores crudos sin contexto.

Información: datos procesados y organizados.

Conocimiento: interpretación de patrones.

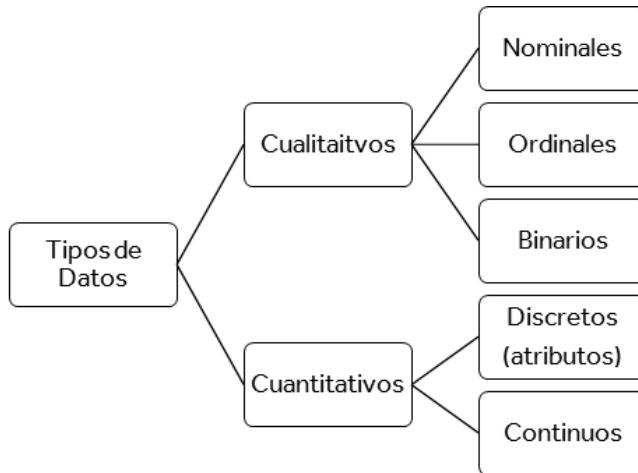
Sabiduría: toma de decisiones informadas.



3. Clasificación de los datos

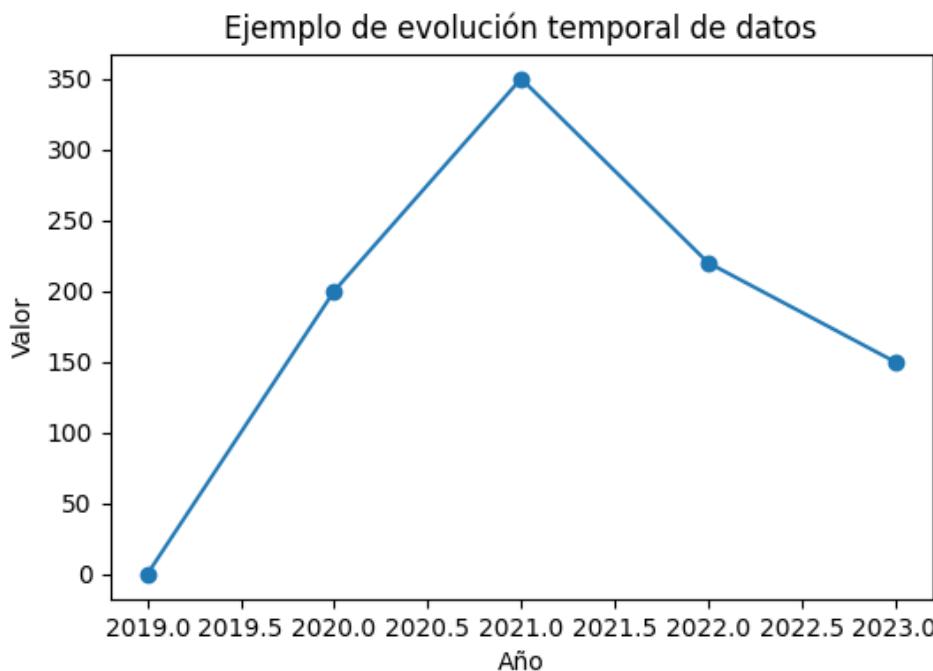
Los datos del mundo real pueden clasificarse según su naturaleza y forma de medición.

Datos cuantitativos (discretos y continuos), datos cualitativos (nominales y ordinales), datos temporales y datos geográficos.



4. Representación gráfica de la información

Ejemplo de visualización temporal que permite observar tendencias.



5. Análisis de una visualización interactiva

Se seleccionó una visualización interactiva de la plataforma Our World in Data sobre COVID-19.

Enlace: <https://ourworldindata.org/coronavirus>

WHAT

Eje X: tiempo.

Eje Y: número de casos.

Color: países o regiones.

Interacción: filtros y selección.

El análisis del WHAT se realiza . La visualización analizada utiliza un dataset estructurado de tipo tabla, donde cada fila representa un ítem (por ejemplo, un país en una fecha específica) y cada columna corresponde a un atributo del dataset.

Tipo de dataset: Tables.

Tipo de datos y dataset: combinación de Tables y Fields (continuos), ya que los valores numéricos evolucionan en el tiempo.

Tipos de atributos: categóricos (país, región), ordenados u ordinales (fecha) y cuantitativos (número de casos).

Dirección de orden: secuencial, dado que el tiempo avanza progresivamente.

Disponibilidad del dataset: dinámica, ya que los datos se actualizan continuamente.

WHY

La posición permite analizar la evolución temporal.

El color facilita la comparación.

La interactividad permite explorar los datos.

El objetivo principal de la visualización corresponde a la acción de alto nivel ANALYZE. La visualización permite al usuario analizar grandes volúmenes de datos para comprender su comportamiento.

Acciones principales:

- Analyze → Consume → Discover y Present, permitiendo identificar patrones y comunicar información.
- Analyze → Produce → Record y Derive, facilitando la obtención de conclusiones.
- Query → Compare y Summarize, para contrastar países, períodos y valores.

Targets del análisis:

- All Data: identificación de tendencias, valores atípicos y características

relevantes.

- Attributes (Many): análisis de dependencias y correlaciones entre múltiples atributos.

6. Conclusiones

En conclusión, la visualización está diseñada para apoyar el análisis exploratorio de datos, permitiendo descubrir patrones, comparar atributos y derivar conocimiento útil para la toma de decisiones, cumpliendo correctamente con los objetivos definidos en el framework.