

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

Desarrollo y Gestión de Software



**Extracción de Conocimiento
en Bases de Datos**

Reporte de investigación de técnicas de visualización

IDGS 91N

PRESENTA:

Chaparro Estrada Hugo Uriel

DOCENTE:

Enrique Mascote

Chihuahua, Chih., 2 dic 2025

Introducción.....	3
1. Importancia de la Visualización de Datos.....	4
2. Tipos de Visualización según el Tipo de Datos.....	4
2.1. Datos Univariados.....	4
2.2. Datos Bivariados.....	5
3. Técnicas Avanzadas de Visualización.....	6
3.1. Visualización Dimensional Reducida.....	6
3.2. Dashboards Interactivos.....	6
3.3. Diagramas de Red.....	7
• Representan relaciones entre entidades (nodos) y sus conexiones (aristas). • Usados en análisis de redes sociales, comunicaciones, grafos semánticos.....	7
4. Visualización y Extracción de Conocimiento.....	7
5. Herramientas y Lenguajes de Visualización Actualmente, existen múltiples plataformas para la visualización de datos: • Lenguajes de programación:.....	7
6. Buenas Prácticas en Visualización.....	8
Conclusión.....	10
Referencias.....	11

Introducción

En la era del Big Data y la inteligencia artificial, el volumen de información generado por diversas fuentes ha alcanzado proporciones masivas. Frente a este escenario, la visualización de datos se ha convertido en una herramienta esencial para comprender, analizar y extraer conocimiento a partir de grandes volúmenes de información. La representación gráfica permite transformar datos complejos en formas comprensibles, facilitando la toma de decisiones y permitiendo una exploración más intuitiva. Este reporte tiene como objetivo explorar las principales técnicas de visualización y representación de información, así como su relación con el proceso de descubrimiento de conocimiento en bases de datos (KDD, por sus siglas en inglés).

1. Importancia de la Visualización de Datos

La visualización de datos es el proceso de representar gráficamente la información con el fin de comunicar patrones, tendencias y relaciones que podrían no ser evidentes en los datos en bruto. No solo ayuda a los analistas a identificar características importantes, sino que también facilita la interpretación por parte de usuarios no técnicos, como ejecutivos, tomadores de decisiones o el público general.

La visualización cumple múltiples funciones:

- **Exploración de datos:** permite detectar anomalías, agrupaciones o tendencias.
- **Comunicación de hallazgos:** presenta resultados de manera clara.
- **Interacción con datos:** posibilita la manipulación visual para un análisis más profundo.

2. Tipos de Visualización según el Tipo de Datos

Dependiendo de la naturaleza de los datos (categóricos, numéricos, temporales, espaciales), se emplean diferentes técnicas de visualización.

2.1. Datos Univariados

- **Histogramas:** muestran la distribución de una variable numérica.
- **Gráficos de barras:** ideales para variables categóricas.
- **Boxplots (diagramas de caja):** resumen la distribución estadística de una

variable.

2.2. Datos Bivariados

- **Diagramas de dispersión (scatter plots):** muestran relaciones entre dos variables numéricas.
- **Diagramas de burbujas:** extensión del scatter plot que incorpora una tercera variable.

2.3. Datos Multivariados

- **Matriz de correlación:** representa relaciones entre múltiples variables numéricas.
- **Gráficos paralelos:** permiten visualizar múltiples variables para diferentes observaciones.
- **Mapas de calor (heatmaps):** representan valores en una matriz de colores.

2.4. Datos Temporales

- **Series de tiempo:** líneas que muestran cómo varía una variable a lo largo del tiempo.
- **Gráficos de área apilada:** muestran la evolución acumulativa de distintas variables.

2.5. Datos Espaciales

- **Mapas de coropletas:** representan valores por regiones geográficas.

- **Mapas de calor geoespaciales:** muestran concentración o intensidad en una ubicación geográfica.

3. Técnicas Avanzadas de Visualización

Con el crecimiento de la complejidad de los datos, se han desarrollado técnicas avanzadas que combinan la estadística, la minería de datos y el diseño interactivo:

3.1. Visualización Dimensional Reducida

- **t-SNE (t-distributed stochastic neighbor embedding):** útil para representar datos de alta dimensión en 2D o 3D.
- **PCA (Análisis de Componentes Principales):** transforma datos multidimensionales en componentes visualizables.

3.2. Dashboards Interactivos

- Integran múltiples visualizaciones y permiten a los usuarios filtrar y explorar datos dinámicamente.
- Herramientas como Tableau, Power BI o D3.js permiten construir tableros complejos con alto impacto visual.

3.3. Diagramas de Red

- Representan relaciones entre entidades (nodos) y sus conexiones (aristas).
- Usados en análisis de redes sociales, comunicaciones, grafos semánticos.

4. Visualización y Extracción de Conocimiento

La visualización no solo ayuda a mostrar datos, sino que es un paso crucial en el proceso de **extracción de conocimiento**:

- **Preprocesamiento**: permite identificar datos faltantes, errores o outliers.
- **Minería de datos**: se apoya en visualizaciones para mostrar resultados de algoritmos (por ejemplo, clusters en un scatter plot).
- **Interpretación**: una vez descubierto un patrón, la representación visual ayuda a comprenderlo y comunicarlo.

El ciclo KDD se complementa con herramientas visuales en cada etapa, desde la selección de atributos hasta la interpretación de resultados.

5. Herramientas y Lenguajes de Visualización

Actualmente, existen múltiples plataformas para la visualización de datos:

- **Lenguajes de programación**:

- *Python*: matplotlib, seaborn, plotly.

- *R*: ggplot2, lattice.
- *JavaScript*: D3.js, Chart.js.

- **Herramientas especializadas:**

- Tableau
- Power BI
- QlikView
- Google Data Studio

Estas herramientas ofrecen flexibilidad para representar desde datos simples hasta estructuras complejas con interactividad en tiempo real.

6. Buenas Prácticas en Visualización

Al diseñar visualizaciones efectivas, se deben seguir principios como:

- **Claridad**: evitar la sobrecarga visual.
- **Coherencia**: mantener un diseño uniforme.
- **Relevancia**: mostrar solo lo necesario para responder una pregunta.
- **Contexto**: incluir leyendas, escalas y títulos descriptivos.

La mala visualización puede conducir a interpretaciones erróneas o decisiones

equivocadas, por lo que es vital una correcta implementación.

Conclusión

La visualización de datos es mucho más que una herramienta estética; es una disciplina fundamental en el análisis de información moderna. Desde histogramas simples hasta visualizaciones interactivas en 3D, cada técnica tiene un propósito específico en el descubrimiento de patrones, la comunicación de hallazgos y la toma de decisiones basada en datos. Comprender las diferentes formas de representar datos, sus aplicaciones y sus limitaciones permite aprovechar al máximo el potencial de la información en el proceso de extracción de conocimiento.

Referencias

- Few, S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Analytics Press.
- Ware, C. (2019). *Information Visualization: Perception for Design* (4th ed.). Morgan Kaufmann.
- McKinney, W. (2018). *Python for Data Analysis* (2nd ed.). O'Reilly Media.
- Tableau Software. (s.f.). *Visual Analytics Best Practices*. <https://www.tableau.com/>
- Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2016). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (6th ed.). Pearson.