

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN



Extracción de Conocimiento de Base de Datos

Reporte de Investigación de Técnicas de Visualización

IDGS91N

Alumno:

Erick Fabian Terrazas Hernandez

Docente:

Enrique Mascote

Chihuahua, Chih., 30 de noviembre de 2025

Introducción

La visualización de datos es uno de los elementos más importantes dentro del análisis moderno de información. En un mundo donde las organizaciones generan millones de datos cada segundo, la capacidad de presentarlos de manera clara, comprensible y visualmente efectiva se ha convertido en una habilidad esencial. Las visualizaciones facilitan la identificación de patrones, tendencias, anomalías y relaciones que no serían evidentes en su formato tabular.

La importancia de la visualización radica en su capacidad para transformar datos complejos en conocimiento accionable. Un gráfico bien diseñado permite comunicar hallazgos clave de forma precisa y rápida, apoyando la toma de decisiones estratégicas, el análisis exploratorio y la comunicación ejecutiva.

El objetivo del presente reporte es documentar las técnicas, fundamentos, herramientas y bibliotecas más utilizadas en la visualización de datos, así como explicar el proceso del *Data Storytelling*, un enfoque que combina narrativa con análisis visual para transmitir mensajes claros y convincentes.

El alcance del trabajo incluye conceptos básicos, tipos de gráficas, mejores prácticas, herramientas de Business Intelligence, bibliotecas de programación, ejemplos de uso, conclusiones y referencias académicas.

Fundamentos de Visualización de Datos

3.1 Conceptos básicos

Sistemas de coordenadas

Los sistemas de coordenadas permiten ubicar datos dentro de un espacio visual. Los principales son:

- **Coordenadas cartesianas:** el sistema más común, utiliza ejes perpendiculares X y Y. Se emplea en gráficos de barras, líneas y dispersión.
- **Coordenadas polares:** basadas en radios y ángulos. Se utilizan para gráficos circulares, radiales y de área polar.

- **Coordenadas geoespaciales:** se basan en latitud y longitud. Son esenciales para mapas, análisis geográficos y visualización de patrones espaciales.

Una elección incorrecta del sistema de coordenadas puede distorsionar la percepción de los datos.

Tipos de ejes

Los ejes permiten contextualizar valores y facilitar la lectura del gráfico.

- **Ejes lineales:** aumentan de forma constante. Adecuados para datos homogéneos.
- **Ejes logarítmicos:** aumentan en potencias de 10. Útiles cuando los valores varían drásticamente.
- **Ejes categóricos:** representan grupos o etiquetas, comunes en barras.
- **Ejes temporales:** se utilizan para series de tiempo y permiten analizar tendencias cronológicas.

La elección correcta del eje es clave para interpretar correctamente la información.

Esquemas de colores y su importancia psicológica

El color transmite emociones y guía la atención del usuario. Su uso debe ser estratégico:

- **Azul:** confianza, estabilidad.
- **Rojo:** alerta, advertencia, intensidad.
- **Verde:** crecimiento, estabilidad.
- **Amarillo:** advertencia suave, atención.

Además, existen paletas especializadas:

- **Secuenciales:** muestran gradientes de menor a mayor.
- **Divergentes:** contrastan dos extremos con un punto neutro.
- **Categóricas:** asignan colores distintos a grupos sin relación numérica.

Un mal uso del color puede confundir al lector o transmitir un mensaje equivocado.

Principios de diseño visual

Los gráficos deben seguir reglas de diseño para ser claros y eficaces:

- Priorizar claridad sobre estética.
- Evitar elementos innecesarios (*chartjunk*).
- Usar etiquetas y leyendas claras.
- Mantener consistencia en colores, formatos y escalas.
- Maximizar la densidad de información sin saturar.

Tipos de representación gráfica

A. Visualización de cantidad

1. Gráficos de barras y columnas

Representan valores comparativos entre categorías; son ideales para resaltar diferencias claras entre grupos.

2. Pictogramas

Usan íconos para representar cantidades. Son atractivos para público general, educación y presentaciones informales.

B. Visualización de distribución

1. Histogramas

Dividen los valores en intervalos y muestran la frecuencia. Permiten identificar dispersiones, sesgos y patrones.

2. Box plots

Muestran mediana, cuartiles e outliers. Son esenciales para comparar distribuciones.

3. Gráficos de violín

Integran la forma del boxplot con una representación de la densidad. Ideales para muestras medianas y grandes.

C. Visualización de proporción

1. Gráficos de pastel y donut

Representan porcentajes de un total. Funcionan mejor con pocas categorías.

2. Treemaps

Muestran proporciones mediante rectángulos. Son eficientes con grandes cantidades de categorías.

D. Visualización de relación XY

1. Diagrama de dispersión

Permite observar correlaciones entre dos variables numéricas.

2. Gráficos de burbujas

Incorporan una tercera variable mediante el tamaño del punto.

E. Visualización de datos geoespaciales

1. Mapas de calor

Representan concentraciones mediante colores.

2. Mapas coropléticos

Colorean regiones según un indicador (ejemplo: población, ingresos).

3. Cartogramas

Deforman regiones según una variable cuantitativa. Útiles para análisis poblacionales.

F. Visualización de incertidumbre

1. Barras de error

Indican margen de error, desviación estándar o variabilidad.

2. Intervalos de confianza

Muestran rangos de estimación estadística confiable.

Proceso de Storytelling con Datos

Definición y componentes

El *Data Storytelling* combina datos, visualizaciones y narrativa para comunicar hallazgos de manera efectiva.

Sus componentes son:

- **Contexto:** define el propósito y la audiencia.
- **Narrativa:** estructura lógica que guía al lector.
- **Visualización:** soporta el mensaje principal y crea impacto.

Etapas del proceso

1. Comprender a la audiencia

Identificar su nivel técnico, necesidades y objetivos.

2. Definir el mensaje clave

Cada gráfico debe respaldar este mensaje.

3. Seleccionar datos relevantes

Evitar incluir datos irrelevantes o redundantes.

4. Diseñar la narrativa visual

Ordenar los gráficos de forma que cuenten una historia lógica.

5. Presentación efectiva

Usar claridad, precisión y diseño profesional.

Mejores prácticas

- Utilizar gráficos adecuados para el tipo de dato.
- Evitar saturar de colores o texto.
- Respetar jerarquía visual.
- Guiar al lector con anotaciones y resaltados.

Errores comunes incluyen:

- Gráficos irrelevantes, escalas incorrectas, colores exagerados y exceso de información.

Herramientas de Visualización

Tableau

Herramienta líder en análisis visual interactivo.

- **Ventajas:** intuitiva, potente, excelente para storytelling.
- **Desventajas:** costo elevado.

Power BI

Herramienta de Microsoft para análisis empresarial.

- **Ventajas:** económica, integración con Office.
- **Desventajas:** menor flexibilidad gráfica que Tableau.

Google Data Studio / Looker Studio

Herramienta gratuita orientada a dashboards web.

- **Ventajas:** gratuita, ideal para marketing.
- **Desventajas:** limitada para análisis avanzados.

Bibliotecas de Visualización

Matplotlib (Python)

La biblioteca base para gráficos científicos.

```
import matplotlib.pyplot as plt
plt.plot([1,2,3], [2,4,6])
plt.show()
```


Seaborn (Python)

Incluye estilos avanzados y funciones estadísticas.

```
import seaborn as sns
sns.boxplot(data=[1,2,3,4,5])
```

Plotly (Python/JS)

Permite gráficos interactivos con zoom y hover.

```
import plotly.express as px
px.scatter(x=[1,2], y=[3,4])
```

D3.js (JavaScript)

Una de las librerías más avanzadas y personalizables para web.

Conclusiones

Las herramientas de Business Intelligence permiten crear dashboards interactivos sin necesidad de programar, mientras que las bibliotecas de visualización ofrecen mayor control y flexibilidad técnica.

El aprendizaje obtenido resalta la importancia del diseño visual, el manejo del color y la narrativa como elementos clave para construir visualizaciones claras, efectivas y alineadas con los objetivos del análisis.

Referencias (APA)

Knaflitz, C. N. (2015). *Storytelling with Data*. Wiley.

Few, S. (2012). *Show Me the Numbers*. Analytics Press.

Tableau Software. (2024). *Official Documentation*.

Microsoft. (2024). *Power BI Documentation*.

Plotly Technologies. (2024). *Plotly Express User Guide*.