

# Capítulo 2: Manipulación y Visualización

- [Práctica 4: Datos que Cuentan Historias \(Gráficas\).](#)
- [Práctica 5: Detectives de Datos \(Limpieza de datos\).](#)

# Práctica 4: Datos que Cuentan Historias (Gráficas).

Los seres humanos somos visuales. Si te muestro una tabla con 800 filas de números, te dolerá la cabeza. Pero si te muestro un gráfico colorido, tu cerebro detectará patrones en segundos.

Hoy convertiremos los datos aburridos en historias visuales. Usaremos dos librerías artísticas: **Matplotlib** (la base) y **Seaborn** (el maquillaje para que se vea genial).

---

## ? Objetivo de la Práctica

- Entender la importancia de la visualización de datos.
  - Crear **Histogramas** para ver distribuciones (¿Qué tan frecuentes son los datos?).
  - Crear **Diagramas de Dispersión** (Scatter Plots) para ver relaciones entre dos variables.
  - Interpretar gráficos para sacar conclusiones.
- 

## ? Conceptos Clave

### ? Matplotlib vs Seaborn

- **Matplotlib:** Es el abuelo de las gráficas en Python. Es muy poderoso pero a veces requiere mucho código para que se vea bonito.
- **Seaborn:** Es una librería construida sobre Matplotlib. Hace gráficos estadísticos increíbles con una sola línea de código. ¡Usaremos ambos!

### ? Tipos de Gráficos Básicos

1. **Histograma:** Barras que nos dicen "cuántos hay de cada cosa".
  2. **Scatter Plot:** Puntos en un plano X-Y. Sirve para ver si una variable afecta a la otra (ej. ¿Más altura = Más peso?).
- 

## ?? Instrucciones Paso a Paso

*Nota: Como es una nueva sesión de Colab, necesitamos cargar los datos de nuevo.*

## Paso 1: Importar Librerías y Datos

En tu primera celda, preparamos todo. Fíjate que ahora importamos dos cosas nuevas.

```
import pandas as pd
import matplotlib.pyplot as plt # La librería gráfica base
import seaborn as sns # La librería gráfica bonita

# Cargamos el dataset de Pokémon nuevamente
url =
"[https://raw.githubusercontent.com/KeithGalli/pandas/master/pokemon_data.csv](https://raw.githubusercontent.com/KeithGalli/pandas/master/pokemon_data.csv)"
df = pd.read_csv(url)

print("Librerías gráficas y datos listos.")
```

## Paso 2: El Histograma (Distribución del Ataque)

Queremos saber: ¿La mayoría de los Pokémon son débiles, fuertes o promedios? Para esto, analizaremos la columna Attack.

Copia y ejecuta:

```
# Configuramos el tamaño de la figura (Ancho, Alto)
plt.figure(figsize=(10, 6))

# Creamos el histograma
sns.histplot(df['Attack'], bins=20, color='skyblue')

# Ponemos títulos (¡Gráfico sin título no sirve!)
plt.title('Distribución del Ataque de los Pokémon')
plt.xlabel('Puntos de Ataque')
plt.ylabel('Cantidad de Pokémon')
plt.show() # Mostramos el dibujo
```

Analiza el resultado: Verás una "montaña". La parte más alta te dice dónde se encuentra la mayoría de los Pokémon (probablemente entre 50 y 80 de ataque).

## Paso 3: Diagrama de Dispersión (Ataque vs. Defensa)

Ahora una pregunta más compleja: ¿Los Pokémon que pegan fuerte también aguantan mucho? (¿Hay relación entre Ataque y Defensa?).

Usaremos un Scatter Plot (nube de puntos).

```
plt.figure(figsize=(10, 6))

# x = Eje horizontal, y = Eje vertical
sns.scatterplot(data=df, x='Attack', y='Defense')

plt.title('Relación Ataque vs Defensa')
plt.show()
```

**Analiza:** Verás una nube dispersa. Si los puntos formaran una línea diagonal hacia arriba, significaría que "más ataque = más defensa". Aquí parece que no hay una regla clara, están por todos lados.

## Paso 4: Agregando una tercera dimensión (Color)

El gráfico anterior es un poco aburrido. Vamos a ver si el Tipo de Pokémon (Fuego, Agua, Roca, etc.) influye. Usaremos el argumento hue (matiz/color) para colorear los puntos según su Type 1.

```
plt.figure(figsize=(12, 8)) # La hacemos más grande

# Agregamos hue='Type 1' para colorear por tipo
sns.scatterplot(data=df, x='Attack', y='Defense', hue='Type 1')

plt.title('Ataque vs Defensa por Tipo')
plt.show()
```

**Analiza:** ¡Ahora es arte moderno!

- Busca los puntos de color Roca (Rock) o Acero (Steel). ¿Notas que suelen estar más arriba en el eje Y (Defensa)?
- Busca los puntos Normal o Bicho (Bug). ¿Dónde suelen estar?

# ? Reto Final (Entregable)

Ahora te toca a ti ser el analista de datos.

- Crea un nuevo Scatter Plot que compare la Velocidad (Speed) en el eje X contra la Vida (HP) en el eje Y.
- Usa hue para colorearlos según si son Legendarios o no (la columna se llama Legendary).
- Pregunta de reflexión: Observando tu gráfico, ¿Los Pokémon legendarios (puntos de un color distinto) tienden a ser más rápidos y tener más vida que los normales? ¿O están mezclados?

**Sube tu Notebook con los gráficos generados.**

# Práctica 5: Detectives de Datos (Limpieza de datos).