

Práctica 3: El Petróleo del Siglo XXI (Carga de datos con Pandas).

Dicen que "los datos son el nuevo petróleo". Pero el petróleo crudo no sirve de mucho si no se refina para convertirlo en gasolina.

En el mundo de Python, nuestra "refinería" se llama **Pandas**. Es una librería increíblemente potente que nos permite tomar archivos gigantes de Excel o CSV y manipularlos en milisegundos. Hoy dejaremos de ver datos como simples números y empezaremos a ver información.

? Objetivo de la Práctica

- Entender qué es un **DataFrame**.
 - Aprender a cargar datos externos desde internet (`.read_csv()`).
 - Aprender a explorar datos desconocidos (`.head()`, `.tail()`, `.describe()`).
 - Realizar ordenamientos básicos para responder preguntas.
-

? Conceptos Clave

? Pandas

No es el oso 🐼. Es la librería de **PA**nel **DA**ta. Es la herramienta estándar en la industria para análisis de datos.

? DataFrame (df)

Es el objeto principal de Pandas. Imagínalo simplemente como una **Tabla de Excel superpoderosa** que vive en la memoria de tu código. Tiene filas (índices) y columnas (variables).

? CSV (Comma Separated Values)

Es el formato de archivo más común para datos. Es texto plano donde cada columna está separada por una coma. Es ligero y cualquier programa lo puede leer.

?? Instrucciones Paso a Paso

Para esta práctica, usaremos una base de datos con las estadísticas de combate de los **Pokémon**.

Paso 1: Importar la librería

En tu primera celda de código, necesitamos llamar a nuestra herramienta.

```
import pandas as pd

# El "as pd" significa que de ahora en adelante,
# en lugar de escribir "pandas", solo escribiremos "pd" para ahorrar tiempo.
```

Paso 2: Cargar los datos (El oleoducto)

Vamos a traer los datos directamente desde un repositorio en Internet. Copia esta URL con cuidado.

```
# Guardamos la dirección del archivo en una variable
url =
"[https://raw.githubusercontent.com/KeithGalli/pandas/master/pokemon_data.csv](https://raw.githubusercontent.com/KeithGalli/pandas/master/pokemon_data.csv)"

# Leemos el archivo CSV y lo convertimos en un DataFrame llamado "df"
df = pd.read_csv(url)

print("¡Datos cargados correctamente!")
```

Paso 3: Exploración Inicial (Mirando los datos)

Nunca confíes en datos que no has visto. Vamos a ver las primeras filas.

```
# .head(n) muestra las primeras n filas
df.head(5)
```

Observa las columnas: Name, Type 1, HP (Vida), Attack, Defense, etc.

También podemos ver el final de la tabla:

```
# .tail(n) muestra las últimas n filas
df.tail(3)
```

Paso 4: Estadísticas Rápidas

¿Cuál es el promedio de ataque de todos los Pokémon? ¿Cuál es la defensa máxima? Pandas te lo dice en una línea.

```
# .describe() nos da un resumen estadístico automático
df.describe()
```

Busca la fila "mean" (promedio), "min" (mínimo) y "max" (máximo).

Paso 5: Ordenando Datos (Respondiendo preguntas)

Vamos a responder: ¿Cuáles son los Pokémon más débiles? Para eso, ordenamos la tabla por "Total" (la suma de sus poderes) de menor a mayor.

```
# sort_values ordena los valores.
# ascending=True significa "ascendente" (de menor a mayor)
df.sort_values("Total", ascending=True).head(5)
```

Ahora al revés: ¿Cuáles son los más fuertes?

```
# ascending=False significa "descendente" (de mayor a menor)
df.sort_values("Total", ascending=False).head(5)
```

? Reto Final (Entregable)

En tu Notebook, escribe el código necesario para responder las siguientes preguntas utilizando lo que aprendiste hoy.

- Muestra los 5 Pokémon con mayor Defensa (Defense).
- Muestra los 5 Pokémon con mayor Velocidad (Speed).
- Pregunta de reflexión (escríbela en una celda de texto): Si tuvieras que elegir un Pokémon basándote solo en estos datos para ganar una carrera, ¿cuál elegirías y por qué?
- Sube el enlace de tu Colab a la plataforma.

Revision #6

Created 15 December 2025 04:20:46 by Admin

Updated 15 December 2025 05:07:10 by Admin