

Introducción

En este análisis se simuló el flujo de clics de usuarios utilizando **Python, PySpark y Pandas** para observar cómo evoluciona la interacción en una plataforma fitness.

Código utilizado

```
# analisis.py (fragmento)
import pandas as pd
from pyspark.sql import SparkSession
import matplotlib.pyplot as plt
import random, time
```

Ejecución y resultados

Durante la ejecución, se generó el dataset `data_simulada.csv` y un gráfico de clics acumulados por usuario:

Figura 1. Clics acumulados por usuario tras el procesamiento del flujo simulado.

Conclusiones

PySpark permite procesar datos de flujo simulados de forma eficiente.

El análisis en tiempo real facilita la detección de patrones de comportamiento.

El gráfico muestra el crecimiento progresivo del número de clics por usuario.

Despliegue del blog y arquitectura

El blog fue desarrollado con Jekyll y desplegado en GitHub Pages.

La estructura general del proyecto es:

```
|— _posts/
|   |— 2025-11-03-analisis-flujo-streaming.md
|— index.html
|— assets/
|   |— grafico_clicks.png
```