

Portales de Empleo - SPRINT I

Parte 1 - Extracción de Datos

- Navegar la web de indeed.com y entender su estructura.
- Desarrollar un script que extraiga la información de una o más ofertas de empleo.
- Desarrollar un script que extraiga la información de la sección “Buscar sueldos”.
- Ejecutar los scripts con una muestra relativamente pequeña de acciones y verificar que todo funcione correctamente.

Parte 2 - Limpieza y EDA inicial

- Hacer limpieza general de datos (si aplica)
- Modelar los datos para trabajar cómodamente para poder cargarlos en una base de datos. La base de datos que van a utilizar queda a vuestro criterio.
- En el modelo de datos, cada registro debe tener un timestamp de extracción y un identificador para que se pueda manejar el sistema de actualización.
- Ejecutar los scripts de recopilación de datos, aplicándoles la limpieza y transformaciones de modelado correspondientes.
- Estudiar cómo se relacionan las diferentes variables entre sí.
- Hacer una segmentación del mercado y buscar grupos de ofertas de empleo usando técnicas de clustering.
- Preparar un dashboard usando **plotly** exponiendo los resultados del análisis de segmentación.

Parte 3 - Base de Datos

- Definir una estructura de bases de datos.
- Realizar un script para popular la base de datos.
- Realizar un script para extraer datos mediante queries.

Parte 4 - Streamlit

- Preparar la estructura del proyecto (carpetas, entorno virtual, dependencias, etc.)
- Preparar un archivo `.env` o un `st.secrets` para almacenar información sensible como usuario y contraseña de la base de datos.
- Preparar la configuración general de Streamlit.
- Definir la paginación y las secciones principales
 - Las vista principal (landing page).
 - La vista de presentación de datos (filtros, mapa, dashboard, etc.)
 - La vista detallada donde se puede buscar una o varias ofertas de empleo.
- Montar placeholders para las visualizaciones, media y el texto que se vaya a incluir en cada una de las vistas.