



Etapa 1: Idea y planificación

- Brainstorming de ideas, análisis de viabilidad e interés del equipo.
- Definición del **alcance y los objetivos del proyecto**.



Etapa 2: Obtención de Datos (Scraping) 32.000 vinos de vivino.es

Implementación del proceso en varias etapas :

- Recopilación de todas las URLs.
- Recopilación de la información dentro de cada página: uso de BeautifulSoup y Selenium.
- Adaptación de formato y unión de los datasets.

TIEMPO : La extracción tomó unas 3 semanas.



Etapa 3 : EDA

EDA: análisis y comparación entre vinos, bodegas, con datos sobre situación geográfica, precio, valoración, frecuencia, tipo de uva, notas de sabor. Los gráficos son interactivos usando altair, seaborn dentro del entorno de Streamlit.



Etapa 4 : Perfect match

Bodega perfecta y Vino perfecto : mejor combinación por país (análisis multivariado entre ubicación, tipo de uva, valoración, precio).

Elige tu vino: desarollo



Etapa 7: Implementación sobre Streamlit



Etapa 6: Valorify

- **Predicción de la valoración** de un vino nuevo
- **Reconocimiento de foto de la etiqueta por IA** (Gemini) para facilitar al usuario el relleno de los campos necesarios.

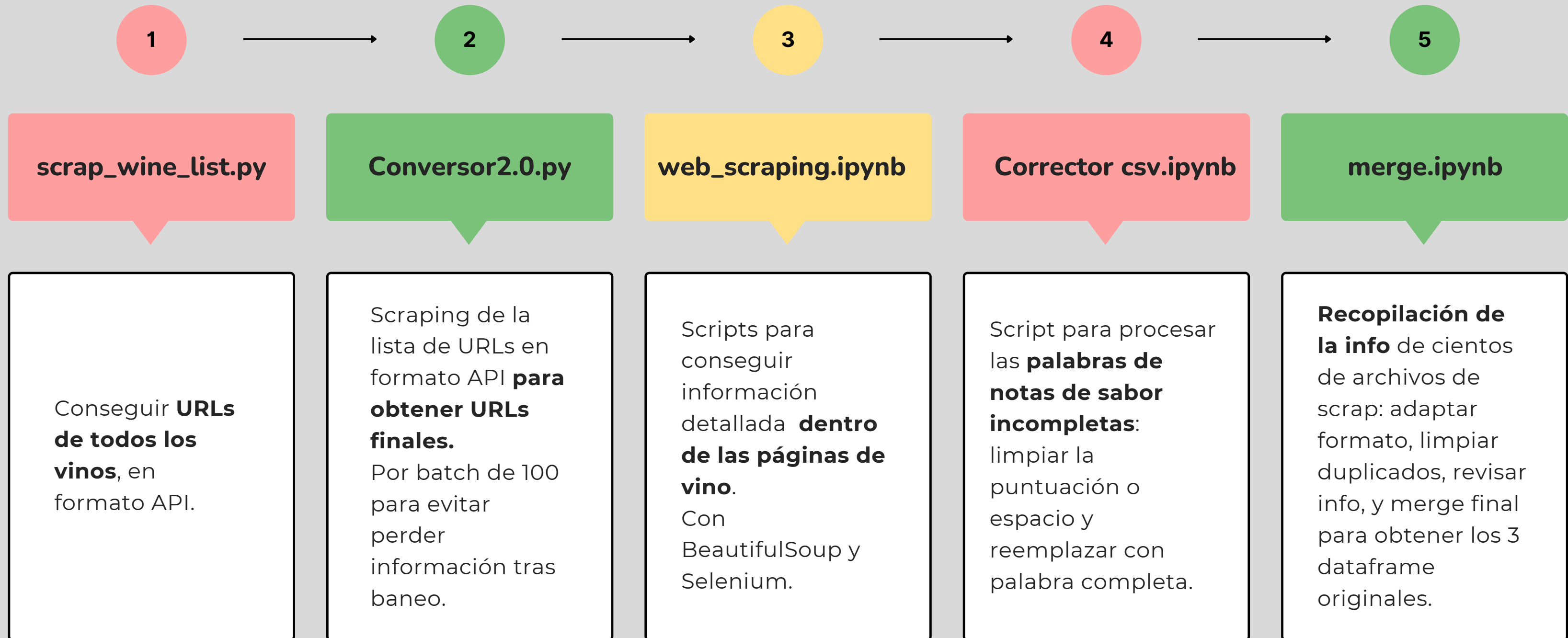


Etapa 5: Recomendador de vino

Recomendación de un vino similar al elegido por el usuario, dentro de un rango de precio y con la mejor valoración.

Web Scraping - Proceso y scripts

32.000 vinos screapeados entre tintos, blancos y espumosos



Dificultades encontradas :

- Tiempo : horas y horas de script en ejecución
- Baneo de parte de la web de vivino.es: scripts con batchs a lanzar regularmente
- Formato : Ciertas informaciones más complicadas que otras a ir a buscar, varios scripts fueron necesarios.

Integración de script para EDA, modelos predictivos y streamlit

