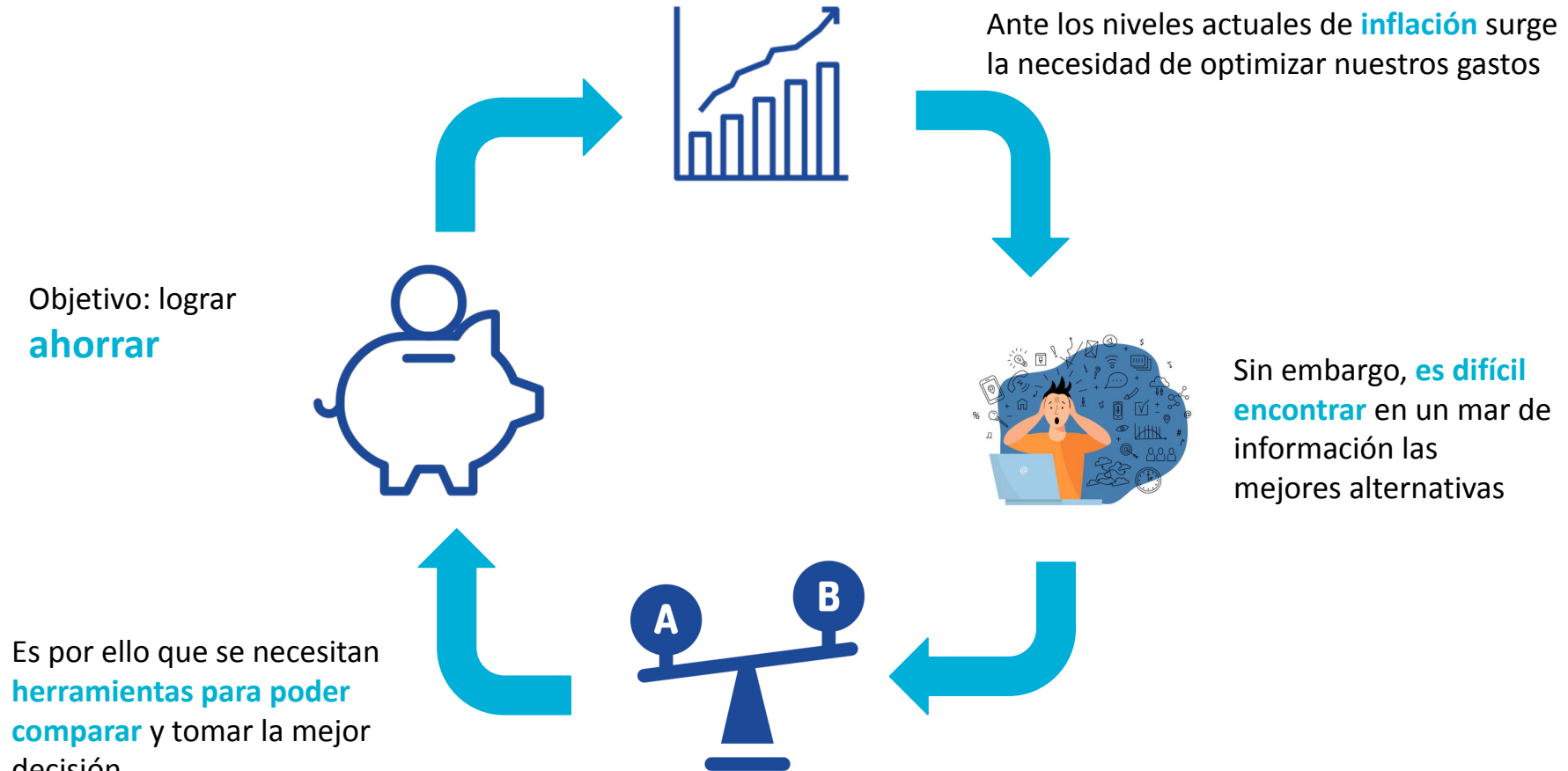


Web Scraping

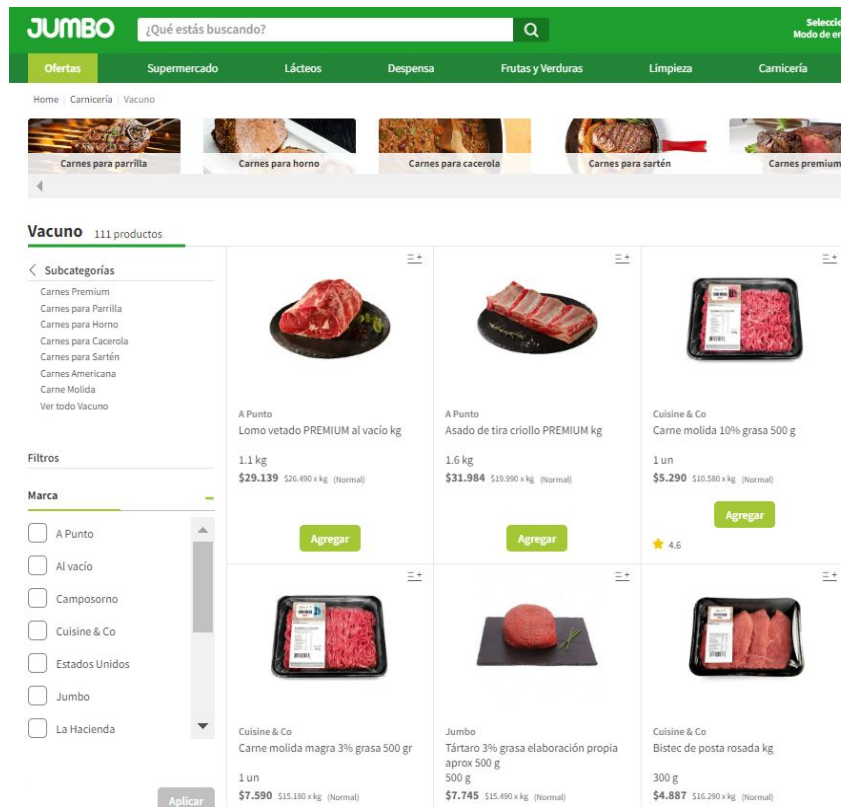
Tracker Precios Supermercados

Florencia Sutter | Francisco Figuera | Zhan Gao

1. Contexto

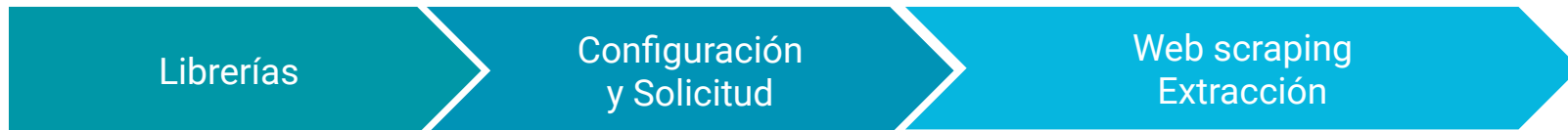


2. Pasos a seguir



- 1 **Escoger** páginas web:
<https://www.jumbo.cl/>
<https://www.lider.cl/>
- 2 **Definir un objetivo** e información a ser capturada:
 Comparar y analizar evolución de precios (carnes 18 sep)
- 3 **Desarrollar un scraper** para extraer y almacenar información.
- 4 **Seleccionar formato** de salida de la información.
- 5 **Prueba y depuración** de errores y correcciones.
- 6 **Automatizar** scraper según definición de la actualización periódica
- 7 **Visualización** de los resultados

3. Pseudocódigo



- Selenium
- BeautifulSoup
- Pandas
- Numpy
- Time
- Datetime

webdriver.ChromeOptions()
requests.get

Sitios Web:
URL1 = <https://www.lider.cl/>
URL2 = <https://www.jumbo.cl/>

Por cada página:

Por categoría:

Lista de producto:

Recorre por: marca, nombre, precio

-Extraer el dato

-Agregarlo al DF

4. Dataframe

El **DataFrame** resultante contará posiblemente con las siguientes columnas:

N°	Nombre	SKU	Precio Normal	Precio Promocional	Ahorro	Unidad

N°= Número de evento

Nombre = Nombre del producto

SKU = Código identificador del producto en la cadena

Precio Normal = Valor regular del producto

Precio Promocional = Valor con descuento del producto

Ahorro = Diferencia entre precio normal vs Precio Regular

Unidad = Unidad métrica del producto