

# PRA1 - Tipologia i cicle de vida de les dades

- Lluís Bernardez Casaus
- Aina Garcia Espriu

## Respostes:

### 1. Context.

DatosMacro.com és un web que depen del diari econòmic expansion, diari líder a Espanya en aquest àmbit fundat a Madrid l'any 1986 , el portal és una base de dades de caràcter financer/econòmic tals com l'index de preus al consumidor (IPC), valors sobre materies primes, salaris, etc on a més podem obtenir aquests valors per països i veure l'evolució al llarg del temps. Així hem escollit del reguitzell de dades, l'històric dels preus dels derivats del petroli per països i l'evolució del preu dels barrils de brent de la OPEP.

### 2. Títol.

Evolució del preu del petroli i els seus derivats

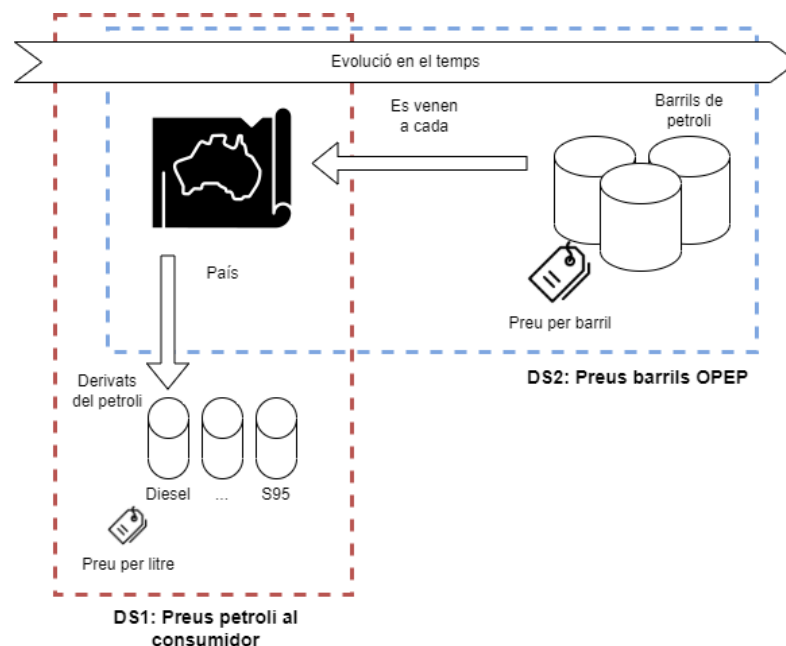
### 3. Descripció del dataset.

Evolució del preu del petroli tant dels barrils OPEP com el preu final al consumidor, per país. El dataset conté dues taules:

La primera recull l'evolució del preu dels barrils de brent per l'OPEP en dolars i euros.

La segona recull l'evolució del preu de la gasolina súper 95, el diesel i el diesel domèstic amb i sense impostos per països europeus i els EEUU en euros

### 4. Representació gràfica.



## 5. Contingut.

Mostrem la informació en la següent taula:

Taula OPEP		
Camps	Descripció	Període
Fecha	Data del valor en format mes en espanyol i any numero amb 4 dígit	Gener 2003-Actualitat
Precio dolar	Preu dels barrils de Brent en dolars americans	Gener 2003-Actualitat
Precio euro	Preu dels barrils de Brent en euros	Gener 2003-Actualitat
Taula derivats petroli		
Camps**	Descripció	Període*
Fecha	Data del valor en format AAAA-MM-DD	Gener 2005-Acualitat
Super 95	Import mitjà de la gasolina 95	Gener 2005-Acualitat
Super 95 (Sin imp.)	Import mitjà de la gasolina 95 sense impostos	Gener 2005-Acualitat
Diesel	Import mitjà del gasoil	Gener 2005-Acualitat
Diesel (Sin imp.)	Import mitjà del gasoil sense impostos	Gener 2005-Acualitat
Diesel Cal.	Import mitjà del gasoil domèstic (per la calefacció)	Gener 2005-Acualitat
Diesel Cal. (Sin imp.)	Import mitja del gasoil domèstic (per la calefacció) sense impostos	Gener 2005-Acualitat
País	País d'on provenen els imports dels derivats, països EU mes EEUU	Gener 2005-Acualitat

\*La gran majoria de països de la UE, EEUU conté serie desde 1999 i d'altres com per exemple romania contenen series més petites en aquest cas desde 2008.

\*\*EEUU només conté dades de Super 95 i Diesel amb impostos

## 6. Propietari.

El propietari de [www.datosmacro.com](http://www.datosmacro.com) és Alldatanow, S.L., amb C.I.F.: B-87328910 i Registre Mercantil de Madrid Tomo 33.655, Secció 8ª, Foli 60, Fulla M-605792, e-mail: [info@datosmacro.com](mailto:info@datosmacro.com)

Copiem la descripció que fa el portal:

*“Datosmacro tiene como objetivo ofrecer las principales variables económicas y socio demográficas de cientos de países, para ofrecer una visión global de la situación económica en cada momento y en cada país.”*

Bàsicament extreuen les dades dels instituts nacionals d'estadística i els banc centrals de cada país i organismes internacionals d'àmbit socio-econòmic com Banc Mundial. Les dades son per tant de tipus públic.

Ens hem assegurat que la web no tingui cap llicència privativa. A més, hem mirat l'arxiu robots.txt de la web per tal d'assegurar-nos de que estigui permès fer web-scraping. Les URL que hem scrapejat no estaven prohibides en aquest arxiu. (<https://datosmacro.expansion.com/robots.txt>)

```
Sitemap: https://datosmacro.expansion.com/sitemap.xml
Sitemap: https://datosmacro.expansion.com/sitemap.xml?set=cmp
User-agent: *
# CSS, JS, Images
Allow: /core/*.css$
Allow: /core/*.css?
Allow: /core/*.js$
Allow: /core/*.js?
Allow: /core/*.gif
Allow: /core/*.jpg
Allow: /core/*.jpeg
Allow: /core/*.png
Allow: /core/*.svg
Allow: /profiles/*.css$
Allow: /profiles/*.css?
Allow: /profiles/*.js$
Allow: /profiles/*.js?
Allow: /profiles/*.gif
Allow: /profiles/*.jpg
Allow: /profiles/*.jpeg
Allow: /profiles/*.png
Allow: /profiles/*.svg
```

## 7. Inspiració.

Hem escollit aquest conjunt de dades per ser un tema d'actualitat, la crisi energètica degut al peakoil i la recessió derivada d'aquesta es poden analitzar desde el preu dels barrils i els seus derivats.

Esperem que el dataset serveixi per respondre i analitza diferents punts:

1. Diferències dels preus dels derivats per països.
2. Diferències de càrregues impositives en el derivats per països.
3. Comparar les dues taules per observar períodes amb més especulació per les suministradores de derivats.
4. Diferència de preus dels derivats per països.
5. Diferència de preus dels derivats i dels barrils per series temporals.

## 8. Llicència.

Escollim la llicència CCO ja que l'origen dels datasets és d'àmbit públic, també al ser un tema d'actualitat i que ens hauria de preocupar a tota la societat, compartim el projecte per si pot ser d'utilitat.

## 9. Codi.

El codi té dues funcions principals que fan scrapping sobre diferents indicadors: preu dels derivats del petroli al consumidor i preu dels barrils OPEP.

Cada una d'aquestes funcions navega dins de la subpàgina a través d'extreure la URL del mateix HTML de la pàgina principal. De fet, tots els canvis de pàgina que es fan durant l'execució del codi són a través de links extrets del HTML de la pàgina actual. Per exemple, quan volem iterar per països, cada pàgina específica per país, s'hi accedeix per la URL proporcionada per la web. El mateix passa quan iterem per anys, tot s'accedeix desde navegació interna.

Això ens ha portat un problema en els dos tipus d'indicadors ja que les URLs que fan servir són circulars en el temps. Es a dir, si anem accedint als links dels anys anteriors i arribem al últim any amb dades, aquest ens porta al any actual. Això ho hem resolt a base de parar la iteració quan llegim dades d'un any major al actual (detecció de circularitat)

A part d'això, també hem hagut de tractar algunes dades com ara els noms dels països que estaven del format Nom [+] i hem eliminat els caràcters sobrants. Finalment, una funció escriu el dataframe resultant de cada crida en un csv indicant-li el nom.

## 10. Dataset.

<https://doi.org/10.5281/zenodo.7322701>

## 11. Vídeo.

<https://drive.google.com/drive/folders/1-gl37DyA2h2ccqagBhQbEFiXLuBHvhQJ?usp=sharing>

<b>Contribucions</b>	<b>Signatura</b>
Investigació prèvia	Lluís, Aina
Redacció de les respostes	Lluís, Aina
Desenvolupament del codi	Lluís, Aina
Participació al vídeo	Lluís, Aina