

Tipologia i cicle de vida de les dades

Pràctica 1: Web scraping

Memòria

Nota: Les respostes a les diferents preguntes estan escrites amb **blau** per tal de ser fàcilment identificables

Components de l'equip (usuaris GitHub): Andrés Laverde Marín ([alaverma](#))
Josep M^a Espasa Verdés ([ilergeta](#))

Enllaç GitHub: <https://github.com/alaverma/web-scraping-uoc>

1. Context

Explicar en quin context s'ha recol·lectat la informació. Explicar per què el lloc web triat proporciona aquesta informació.

L'objectiu d'aquesta pràctica ha estat recol·lectar informació de la cotització borsària històrica de les empreses que cotitzen a la Borsa de Madrid, principal mercat de valors d'Espanya, integrada en la Sociedad de Bolsas y Mercados Españoles (BME).

El lloc web www.bolsamadrid.es proporciona aquesta informació, ja que és el lloc web de la societat rectora de la borsa de valors de Madrid, propietària de d'aquestes dades.

2. Títol del dataset

Definir un títol pel dataset. Triar un títol que sigui descriptiu.

Les dades recol·lectades es recullen en un dataset emmagatzemat en un fitxer de format CSV amb el nom corresponent al *ticker* de l'empresa cotitzada de la que s'està obtenint les dades històriques de cotització. Recordar que el *ticker* no deixa de ser un identificador de les empreses que cotitzen en un mercat borsari i que es correspon a una abreviació que pot contenir lletres i/o nombres.

Vàries empreses / fitxers??

3. Descripció del dataset

Desenvolupar una descripció breu del conjunt de dades que s'ha extret (és necessari que aquesta descripció tingui sentit amb el títol triat).

TODO

4. Representació gràfica

Presentar una imatge o esquema que identifiqui el dataset visualment

TODO

5. Contingut

Explicar els camps que inclou el dataset, el període de temps de les dades i com s'ha recollit.

TODO

6. Agraïments

Presentar el propietari del conjunt de dades. És necessari incloure cites de recerca o anàlisis anteriors (si n'hi ha).

El propietari de les dades és: **Sociedad Rectora de la Bolsa de Valores de Madrid, S.A. Sociedad Unipersonal**

D'altra banda, per la realització d'aquesta pràctica, ja sigui, pròpiament, aquest document de memòria o el codi font, s'ha consultat la següent bibliografia:

Col·laboradors de Viquipèdia. 'Borsa de Madrid', *Viquipèdia, l'Enciclopèdia Lliure*. [Data de consulta: 10 de novembre de 2018].
<https://ca.wikipedia.org/wiki/Borsa_de_Madrid>

Col·laboradors de Viquipèdia. 'Ticker symbol', *Viquipèdia, l'Enciclopèdia Lliure*. [Data de consulta: 10 de novembre de 2019].
<https://en.wikipedia.org/wiki/Ticker_symbol>

Lawson, R. (2015). *Web Scraping with Python* (1a ed.). Birmingham: Packt Publishing Ltd.

7. Inspiració

Explicar per què és interessant aquest conjunt de dades i quines preguntes es pretenen respondre.

TODO

8. Llicència

Seleccionar una d'aquestes llicències pel dataset resultant i explicar el motiu de la seva selecció.

TODO

9. Codi

Adjuntar el codi amb el qual s'ha generat el dataset, preferiblement en Python o, alternativament, en R.

TODO

10. Dataset

Presentar el dataset en format CSV.

TODO