

PRÀCTICA 1

1. Context:

La inversió financera és una part important de l'economia i una font d'ingressos per a moltes persones i empreses. Malauradament, és un sector on es treballa amb molta informació i incertesa i la majoria de gent no té coneixements ni eines suficients a l'abast per a escollir quines són les millors accions on invertir.

La finalitat d'aquest projecte és fer una proposta de rànquing de les millors accions.

Les principals webs que hem fet servir per recopilar la informació necessària són **Morningstar** (<https://www.morningstar.es/es/>) i **Yahoo Finance** (<https://finance.yahoo.com/>). Ambdues són webs de referència pel que fa a informació financera. Proporcionen informació fiable i completa sobre la gran majoria de fons i accions disponibles. Aquesta és la raó per la qual hem decidit escollir-les.

D'altra banda, també hem fet servir **Wikipedia** per a aconseguir els tickers de les accions. Es tracta d'una font amb un objectiu clar de compartir informació, que cada cop millora més els seus criteris d'edició i que, pel que fa a les dades que hem extret, també considerem fiable.

2. Títol:

El títol que considerem que descriu millor el dataset és: **Best Stocks**

El nom de l'arxiu generat és **best_stocks.csv**

3. Descripció del dataset:

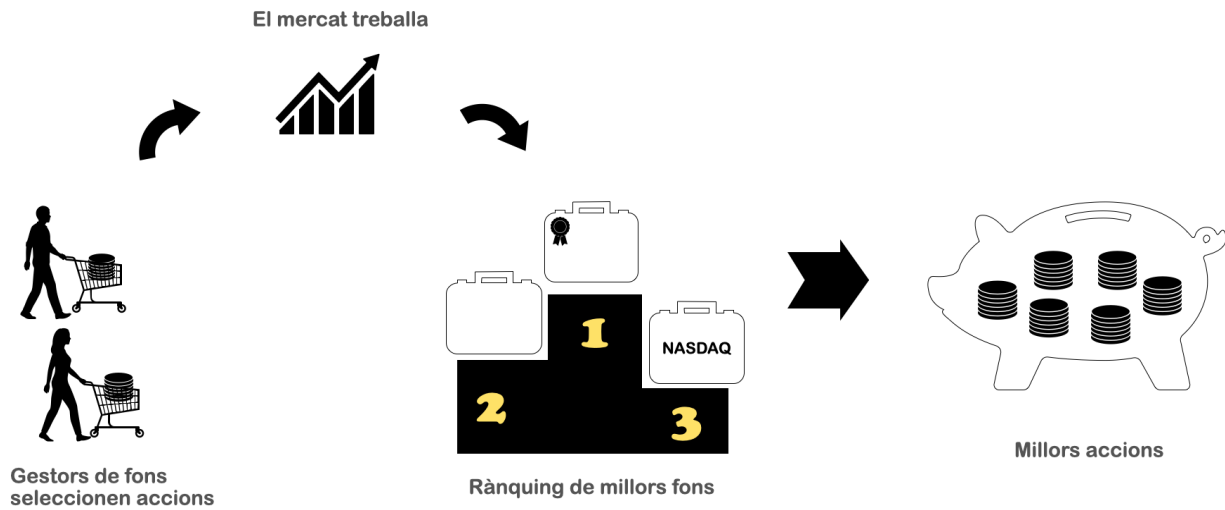
El dataset és un llistat de les millors accions segons el criteri dels millors gestors de fons.

El procés que hem seguit per a valorar quins són els millors fons i les millors accions és el següent: al món de les finances l'objectiu principal és batre al mercat, que vol dir obtenir més rendiment que el mercat de referència o benchmark. Podríem dir que el mercat està representat pels principals índexs i l'índex que ha obtingut més rendiment durant els darrers 10 anys ha sigut el NASDAQ, així que el NASDAQ serà el nostre mercat de referència. Hem extret de la web Morningstar el rendiment dels fons i hem seleccionat els que han aconseguit un rendiment superior al del NASDAQ durant els darrers 10 anys. Considerem que els gestors d'aquests fons són els millors perquè han batut el mercat, en aquest cas el NASDAQ que és el nostre índex de referència.

Després, hem extret les principals accions on inverteixen aquests fons i les hem ordenat per nombre d'aparicions al nostre dataset: quantes més aparicions entre els millors fons, millor és l'acció segons els millors gestors.

Finalment, hem fet servir Yahoo Finance per a recopilar informació sobre aquestes accions.

4. Representació gràfica:



5. Contingut:

El dataset final **best_stocks.csv** inclou els següents camps:

- **stock**: nom de l'acció
- **appearances**: aparicions entre els millors fons
- **ticker**: identificador de l'acció
- **previous close**: preu de l'acció al tancament de la darrera sessió
- **fifty-two weeks range**: rang de preus de les darreres 52 setmanes
- **volume**: volum de negociació actual
- **avg. volume**: volum de negociació promig dels darrers dies
- **beta**: mesura de volatilitat de l'acció
- **per**: rati del preu de l'acció entre el benefici de l'empresa per acció
- **1y target**: estimació del preu a un any vista de Yahoo Finance

Les accions es mostren ordenades per número d'aparicions entre els millors fons, que és la principal informació del dataset.

La resta d'informació és auxiliar però podem destacar la *beta* per tenir una referència de la volatilitat de l'acció en comparació amb el mercat, el *per* per tenir una referència de com de barata o de cara està l'acció comparant amb el *per* d'altres accions del mateix

sector i el *1y target* que és una estimació del preu de l'acció a un any vista que ens permet calcular un rendiment esperat de l'acció segons l'anàlisi de Yahoo Finance.

Les dades recollides són en temps real. És té en compte el rendiment a 10 anys, però no s'han recollit dades històriques.

La informació sobre els fons (índexs, dades de rendiment i llistat de les principals accions) procedeixen de la web Morningstar (<https://www.morningstar.es/es/>). Primer hem fet servir un arxiu json per a extreure el codi que Morningstar assigna a cada fons. Fent servir aquest codi, podem iterar les URL específiques de cada fons. Hem fet servir la llibreria BeautifulSoup per a recopilar la informació sobre cada fons.

De les pàgines de Wikipèdia <https://en.wikipedia.org/wiki/Nasdaq-100> i https://en.wikipedia.org/wiki/List_of_S%26P_500_companies hem extret els tickers de les diferents accions fent servir la funció `read_html` de pandas.

Amb el paquet `process` de la llibreria `fuzzywuzzy` hem trobat les coincidències adequades entre els noms de les accions que fa servir Morningstar i els tickers que identifiquen a cada acció a Wikipèdia.

Finalment, hem extret tota la informació referent a les accions de la web de Yahoo Finance <https://finance.yahoo.com/>. Hem fet servir el llistat de tickers per iterar sobre les URLs de cada acció i amb BeautifulSoup hem extret la informació.

6. Agraïments:

Agraïm a les webs esmentades anteriorment per compartir la informació amb el públic de forma oberta:

- **Morningstar** (<https://www.morningstar.es/es/>)
- **Yahoo Finance** (<https://finance.yahoo.com/>)
- **Wikipedia** (<https://www.wikipedia.org/>)

En una primera versió del codi hem fet servir aquest dataset de Kaggle per a aconseguir els ISINs (International Securities Identification Number) dels fons, com a punt de partida per a extreure la informació de Morningstar:

<https://www.kaggle.com/stefanoleone992/european-funds-dataset-from-morningstar>

En la versió final del codi ja no el fem servir, ja que les dades no estan actualitzades i Morningstar fa servir codis propis per a indexar les URL de cada fons, així que vam decidir extreure aquests codis de la mateixa Morningstar.

Altres webs consultades són:

<https://medium.com/ironhack/data-nurturing-con-pandas-web-scrapping-5068d7596bfe>

La informació obtinguda està disponible a internet de forma oberta. A més, hem comprovat que compleix amb els termes d'ús de les webs utilitzades.

Link als termes d'ús de Morningstar:

<https://www.morningstar.es/es/Disclaimer/Disclaimer.aspx?id=TermsofUse>

Link als termes d'ús de Yahoo:

<https://legal.yahoo.com/us/en/yahoo/terms/otos/index.html>

7. Inspiració

Les webs com Morningstar, Yahoo Finance i moltes altres recopilen tota la informació disponible sobre tots els productes financers de forma indiscriminada.

La finalitat d'aquest projecte és jerarquitzar tota la informació disponible per prendre millors decisions d'inversió en base al criteri dels gestors de fons que obtenen més rendiment.

Contestem a la pregunta: quines són les millors accions en temps real segons els millors gestors?

8. Llicència

Seleccionem la llicència "Released Under CC BY-NC-SA 4.0 License (Attribution-NonCommercial-ShareAlike)"



L'escollim ja que és una llicència que permet l'ús i la distribució del material, a més de l'adaptació, modificació o ampliació del mateix. D'altra banda, estableix algunes condicions que en semblen adequades:

- **Attribution:** l'usuari del material ha de donar crèdit als autors, proporcionar un link a la llicència i indicar les modificacions fetes.
- **NonCommercial:** no es pot fer servir el material amb finalitats comercials.
- **ShareAlike:** l'usuari que modifiqui o amplii el material, ha de distribuir les seves contribucions amb la mateixa llicència.

Es pot trobar més informació a aquesta web:

https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en_US

I el text legal aquí:

<https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

9. Codi

El codi que hem desenvolupat per generar el dataset es troba disponible al següent repositori de **GitHub**:

https://github.com/BlueDunes/PRA1_best_stocks

10. Dataset

El dataset obtingut es troba disponible en format CSV a **Zenodo** amb una breu descripció:

<https://doi.org/10.5281/zenodo.5648614>

Contribucions	Signatura
Investigació prèvia	MDG, JAP
Redacció de les respostes	MDG, JAP
Desenvolupament del codi	MDG, JAP