

Félévi NagyHF specifikáció

Üzleti Intelligencia Labor

Félév: 2025 tavasz

S&P 500 Reporting

Dancsó Marcell - (AZTVS7)

dancsomarci@gmail.com

Bemutató

A választott feladat az S&P 500 index elemzése és report készítése „megalapozott” befektetési döntések meghozatalához. A historikus adatok több nézetben, cég- illetve szektor szerint szűrhetők. A teljes indexre előrejelzés készítése.

Főbb funkciók

- Adatforrások:
 - o <https://datahub.io/core/s-and-p-500-companies> (Cégek részletes adatai/szektorok szerinti megoszlása)
 - o <https://pypi.org/project/yfinance/> (Yahoo finances api custom adatok elérésre)
 - o <https://www.kaggle.com/datasets/camnugent/sandp500> (minta adathalmaz az index alakulásáról)
- Adattárolás:
 - o Napi szintű adatok betöltése, szűrése, validálása -> raw
 - o Végül csillag sémába kerül át az adat:
 - Tény: napi adatok
 - Dimenzió1: cégek adatai
 - Dimenzió2: előrejelzéshez szükséges adatok, predikció
 - Ha valamilyen nehezen aggregálható adatra lenne szükség, a struktúra kiegészíthető további aggregált tény táblákkal.
- Fő KPI-k:
 - o Cég szintű árfolyam, különböző granularitással
 - o Megtérülés
 - o Volatilitás
 - o Trend/Mozgó Átlag
 - o P/E
 - o Etc.
- Reportok jellemzői:
 - o Időszakra/szektorra/cégre szűrés/lefűrés
 - o Index alakulása
- Adatelemzése/data science funkciók.
 - o Index előrejelzése
 - o Feature engineering: lags, mozgó átlagok, volatilitás, RSI
 - o Több modell kipróbálása pl.: SARIMA, Prophet, DL

Választott technológiák

- BI Stack:
 - o MSSQL
 - o SSIS
 - o PowerBI
- Adatelemzés, előrejelzés: python (numpy, tensorflow, prophet, pandas, matplotlib, yfinance)