

Zadání projektu do předmětu Business Intelligence a datové sklady

Členové: David Michalica, Adam Pech

Téma: Analýza finančních trhů

Zdroje dat:

<https://algotrading101.com/learn/yahoo-finance-api-guide/>

Data jsou načítána pomocí knihovny yfinance (Python), která poskytuje data ve formát viz dokumentace:

- Date – datum (a čas) specifické pro danou cenu a aktivum (akcii, měnový pár, ...), (např. 1/1/2020 18:01)
- Open – cena při daném čase (začátek dne, minuty, hodiny,...) - dle zvolené granularity (např. 508.50)
- High – nejvyšší cena v rámci časové jednotky (např. 525.45)
- Low – nejnižší cena v rámci časové jednotky (např. 502.40)
- Close – cena při závěru časové jednotky (dne, minuty,...) (např. 516.00)
- Adj Close – cena při závěru časové jednotky s ohledem na např. "stock split"
- Volume – objem obchodovaných akcií v ks (prodej a nákup za časovou jednotku) (např. 105000)
- Ticker – název akcie, ve formátu pro burzovní operace (např. MSFT)

<https://www.kaggle.com/datasets/gpreda/bbc-news/discussion>

Kolekce dat, týkající se světových zpráv (od 10. 7. 2021)

- title – název zprávy (např. Microsoft presented new laptops)
- pubDate – datum a čas publikace zprávy (např. Mon, 07 Mar 2022 08:01:56 GMT)
- guid – URL odkaz typu (např. <https://www.bbc.co.uk/news/world-europe-60638042>)
- link – odkaz RSS (např. https://www.bbc.co.uk/news/world-europe-60638042?at_medium=RSS&at_campaign=KARANGA)
- description – detailní popis zprávy (např. New laptops from Microsoft has been introduced with higher price,...)

<https://www.kaggle.com/datasets/therohk/million-headlines>

Kolekce dat, týkající se světových zpráv (od 19. 2. 2003 do 31. 12. 2021)

- publish_date – datum publikování ve formátu yyyyMMdd (např. 20030219)
- headline_text – nadpis zprávy (např. aba decides against community broadcasting licence)

Charakteristika a cíl:

Analýza popsaných dat se bude zaměřovat na to jak se vyvíjí akciový trh, měnový trh, komodity, ... v závislosti na čase a událostech. Jsou zkoumány vzájemné korelace mezi jednotlivými daty a jejich reakce na události. Takový přehled by mohl sloužit pro predikci budoucího vývoje trhů.

Klíčové indikátory:

- Cena aktiva v závislosti na čase, vztah jeho ceny vůči ostatním aktivům. Může sloužit ke sledování vývoje cen aktiv v závislosti na jiných aktivech.

Zadání projektu do předmětu Business Intelligence a datové sklady

- Vývoj jednotlivých trhů (americký, japonský, atd.), což může sloužit ke sledování vývoje ekonomik daných zemí a ekonomických sektorů.
- Závislost cen aktiv na událostech ve světě indikující změnu pohledu obchodníků s akcemi a společností v závislosti na dané události.
- Vztah otevíracích cen pro den oproti uzavíracím cenám na konci předchozího dne, lze pozorovat obchodování mimo obchodní hodiny burz a reakce akcí na případné “price gaps”
- Objem obchodovaných aktiv v závislosti na počtu zpráv o daném aktivu v určitém období (lze vypořadovat strach investorů)