



Olá, aluno(a)!

Seja bem-vindo(a) à aula interativa!

Você entrará na reunião com a câmera e o microfone desligados.

Sua presença será computada através da enquete.
Fique atento(a) e não deixe de respondê-la!



Aprenda com quem faz

> Módulo II - Coleta e Armazenamento de Dados em Renda Variável

Primeira Aula Interativa

Prof. Victor A Mancini



Planejamento para hoje

- Introdução ao módulo -> 10 min
- Um pouco sobre contexto e mercado -> 30 min
- Exemplo prático -> 30 min
- Apresentar o trabalho prático -> 10 min
- Dúvidas -> 10 min



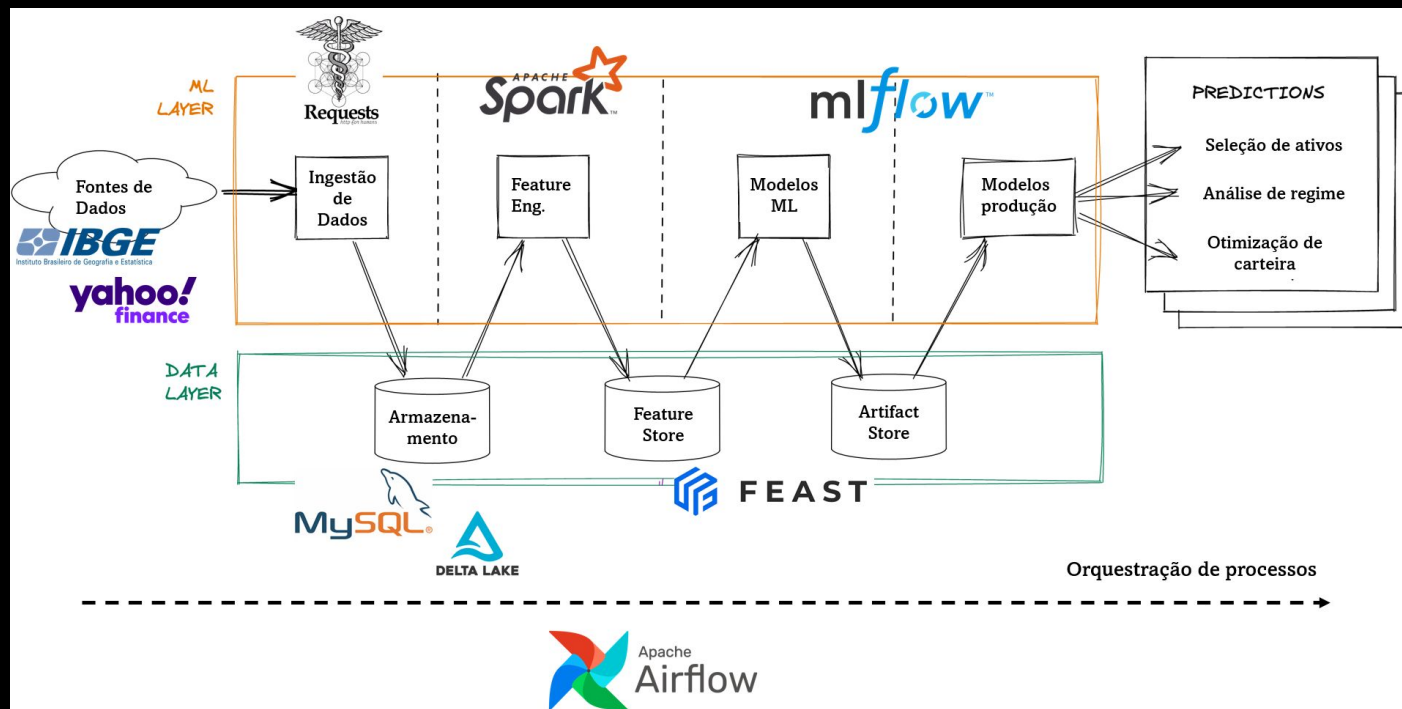


- Pouco sobre mim

- Cientista dos materiais e matemático computacional, mestre em tecnologia nuclear
- +5 anos trabalhando com pesquisa e desenvolvimento
- +5 anos de experiência em ciência de dados no mercado financeiro
 - Hedge fund (~4 anos)
 - Tesouraria (~1 ano)
- Atualmente sou Lead Data Scientist na Fourth Sail Capital



● Introdução ao módulo



- Pouco sobre mercado
 - [Fundos de investimento quantitativo no Brasil](#)
 - [Estratégias quantitativas](#)
 - [Vaga para Engenheiro de Dados](#)
 - [Outra vaga para Engenheiro de Dados](#)
 - [Startup de dados alternativos](#)
 - [Startup de dados de mercado é adquirida pelo TC](#)



- Exemplo prático: coleta e armazenamento de dados de fundos de investimento
1. Coletar histórico de cotas de fundos da CVM;
 2. Armazenar os dados brutos no data lake;
 3. Processar os dados brutos e armazenar os dados de interesse no banco de dados MySQL;
 4. Ler os dados processados em Python;
 5. Rodar uma [análise de risco de fatores](#).



- Trabalho prático - Enunciado

Suponha que trabalhamos em uma gestora de fundos de investimentos e o gestor de portfólio de ações nos solicitou apresentar dados que possam ajudá-lo a decidir investir ou não em Ambev (ABEV3).

Em particular, ele gostaria de olhar dados de mercado, dados fundamentalistas, dados macroeconômicos e dados alternativos em um relatório bastante minucioso.

Inicialmente ele nos solicitou apenas os dados de mercado e dados macroeconômicos para iniciar nossa análise.



● Trabalho prático - Questões

- Colete a série histórica dos dados de inflação de bebidas do IBGE usando a API sidra;
- Colete a série histórica dos dados diários de preços e volume (OHLCV) para a ação ABEV3 do provedor de sua preferência (yahoo finance, marketstack, eodhistoricaldata, finnhub, alphavantage, etc.);
- Crie e armazene esses dados em um banco de dados relacional SQL local ou em nuvem;
- Esboce como seria a implementação em Python de uma DAG Airflow que executaria as coletas implementadas anteriormente. O código não precisa rodar com sucesso, é apenas um esboço.





Muito obrigado!

