# **Projeto Final – MovieFlix Analytics**

Vocês foram contratados como equipe de tecnologia da **MovieFlix**, uma startup fictícia de streaming de filmes.

A empresa deseja lançar uma **plataforma simples de cadastro e avaliação de filmes**, mas também precisa **analisar os dados gerados pelos usuários** para entender preferências, tendências e apoiar decisões de negócio.

O desafio da equipe é construir:

1. Uma **aplicação web com docker** e publicada via **pipeline de CI/CD**.
2. Um **fluxo de dados** que simula **Data Lake, Data Warehouse e Data Mart**, com consultas que geram insights para a empresa.

## **Objetivos do Projeto**

* Compreender como funciona o ciclo de vida de uma aplicação em containers.
* Aplicar conceitos de **servidores, proxies e DNS**.
* Implementar um **pipeline de CI/CD** para build, teste e publicação de imagens Docker.
* Simular um **ecossistema de dados** (Data Lake, DW e Data Mart).
* Realizar consultas analíticas para apoiar decisões de negócio.

## **Aplicação Web**

* Criar uma aplicação simples (pode ser em qualquer linguagem) que permita **cadastrar filmes e avaliar com notas**.
* A aplicação deve rodar no **Docker**.
* Configurar o **Nginx como proxy reverso** para a aplicação.
* Opcional: utilizar um domínio ou subdomínio (mesmo gratuito, ex: DuckDNS).

## **Pipeline (CI/CD)**

Configurar no **GitHub Actions** um workflow com as etapas:

1. **Build da imagem Docker**.
2. **Teste simples** (ex.: verificar se a aplicação sobe e responde na porta correta).
3. **Push da imagem para o Docker Hub**.

**Dados (DW, Data Lake e Data Mart)**

Vocês vão simular o fluxo de dados da MovieFlix, recomendo utilizar uma API para buscar os dados (ex:.omdbapi.com)

### **Data Lake (dados brutos)**

Armazenar arquivos CSV com dados brutos, exemplo:

* movies.csv (informações sobre os filmes)
* users.csv (informações sobre usuários)
* ratings.csv (avaliações dos usuários nos filmes)

Esses arquivos podem ficar em um diretório.

**Data Warehouse (dados tratados)**

Criar tabelas no PostgreSQL (ou outro banco) e carregar os CSVs do Data Lake nessas tabelas.

### **Data Mart (visões de negócio)**

Criar visões/tabelas resumidas a partir do DW, por exemplo:

* **Top 10 filmes mais bem avaliados por gênero**
* **Nota média por faixa etária dos usuários**
* **Número de avaliações por país**

### **Consultas Analíticas**

Escrever queries SQL que respondam perguntas como:

* Quais os 5 filmes mais populares?
* Qual gênero tem a melhor avaliação média?
* Qual país assiste mais filmes?

## **Entregáveis**

1. **Repositório GitHub** com:  
   * Código da aplicação
   * Dockerfile
   * Workflow do GitHub Actions
   * Scripts de carga de dados (ETL simples CSV → Postgres (ou outro banco)
   * README com explicação da arquitetura
2. **Demonstração prática**:  
   * Pipeline rodando no GitHub Actions.
   * Imagem publicada no Docker Hub.
   * Aplicação acessível via Nginx/DNS.
   * Consultas SQL ou dashboards mostrando resultados do Data Mart.

**Prazo:**

Os projetos serão aceitos até o dia 01 de outubro de 2025 às 23:59.

O envio precisa ser feito via e-mail: thayse.frankenberger@gmail.com

Título: Seu nome completo + Projeto final