Relatório Projeto Integrador 3 - Análise de transferência de jogadores

Instituto de Educação Superior de Brasília

Filipe Carvalho de Jesus Marcos Sergio Jordão de Sousa Willian Marreiros Santos

Orientador: Prof. Jose Reginaldo de Sousa Mendes Junior

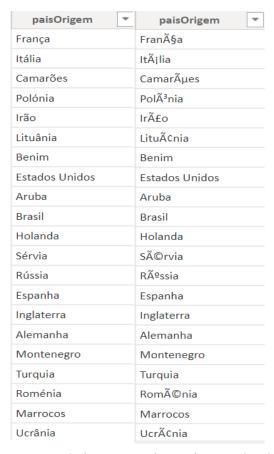
Resumo

A ideia principal do grupo foi implementar um ETL via python, por meio de uma técnica chamada Web Scraping, aplicada então no site Transfermarkt. É um site dedicado ao futebol e à avaliação de valores de mercado de vários jogadores. Em seguida, foi feita a carga de dados, transformação, implementação de gráficos e tabelas sumarizadas baseadas nos dados coletados, no qual, a ferramenta orquestradora foi o software PowerBI.

Desafios

O principal desafio durante o desenvolvimento do projeto foi a elaboração do código para a raspagem de dados do site utilizado pelo grupo. Como nenhum dos integrantes nunca havia feito algo do tipo, foi gasto um tempo considerável para a automatização do processo.

Durante a inserção de atualização dos dados no PowerBI, foi identificado um erro de encoding que o grupo não conseguiu encontrar a solução para obter o resultado esperado, então foi decidido que seria mantida a base atualizada mas com o problema encontrado, exemplo:



A tabela à esquerda refere-se ao período anterior de atualização dos dados. Já a tabela à direita representa os dados após tal atualização.

Conclusão

O grupo conseguiu compreender, aprender e aplicar técnicas da engenharia e ciência de dados, envolvidas durante o desenvolvimento do projeto. Onde o grupo, desenvolveu técnicas de modelagem de dados, como o Star Schema (Modelo Dimensional) e a análise e exploração dos dados coletados.