**ANÁLISE EXPLORATÓRIA FUTEBOL**

Elisa de Oliveira Soares, Gabrielly Esteves Pinheiro Chácara,   
João Vitor Tomaz Alves Ferreira, Luiz Eduardo Bravin, Nícolas Mateus Spaniol

**RESUMO**

*Neste artigo analisamos dados do futebol europeu, incluindo informações detalhadas sobre partidas, clubes, jogadores e eventos ocorridos em campo extraídos do site Transfermarkt e disponibilizados publicamente no Kaggle. Procuramos responder, usando Python e bibliotecas de visualização e manipulação de dados, à hipóteses desenvolvidas pelo grupo acerca do dataset. Exploramos correlações entre o desempenho de jogadores e seu custo, entre a performance dos clubes em partidas e o fato destes estarem jogando dentro/fora do país e as distribuições dos jogadores por mês de nascimento e de cartões recebidos por posição em campo. Com relação ao mês de nascimento, vimos que a quantidade de jogadores no dataset cai linearmente ao longo do ano, com dezembro tendo cerca de metade dos jogadores de janeiro. A distribuição dos cartões, similarmente, se concentrou nas posições defensivas. Descobrimos que, em média, clubes que vendem e depois compram um mesmo jogador ficam com saldo positivo, e tem o intervalo entre as transações de três anos. Quanto ao desempenho dos jogadores e clubes, ….*

**Palavras-chave:** Futebol, Python, visualização de dados, gráfico, estatística, análise de dados

**1 INTRODUÇÃO**

Por meio do uso da linguagem Python, foram analisados diferentes aspectos dos times e jogadores, com base em alguns fatores que podem ou não influenciar o desempenho dos mesmos.

**2 DESENVOLVIMENTO**

**2.1** **Há uma diferença na performance dos times quando estes jogam em casa, como visitantes e fora do país?**

Ddjfhte

**2.2** **Quais foram as compras de jogadores com melhores e piores custo-benefício registradas?**

Ggreryd

**2.3 A posição dos jogadores em campo influencia na quantidade de cartões que estes recebem?**

ESdfbdfrhyerd

**2.4** **Vendas e compras posteriores de um jogador por um mesmo time costumam gerar lucro para o time?**

DFHTEyedvcx

**2.5** **Jogadores com preço fora do comum tem o desempenho proporcional?**

Para realizar a análise proposta, primeiro definimos os conceitos que utilizamos, que são: preços fora do comum e desempenho do jogador.

Com as colunas “*date*” e “*market\_value\_in\_eur*” da tabela “*player\_valuations.csv*”*,* criamos o conceito de preços fora do comum. Primeiramente, corrigimos os valores de “*market\_value\_in\_eur*” de acordo com a inflação decorrida de *“date”* até os dias atuais. Em seguida, calculamos a média ponderada dos preços dos jogadores ao longo de suas carreiras, considerando os pesos iguais ao tempo, em dias, que o jogador permaneceu com aquela cotação. Após isto, calculamos o limite superior – segundo a fórmula LS = Q3 + 1,5DI, onde DI = Q3 – Q1 e Q1 e Q3 são, respectivamente, os primeiro e terceiro quartis – da coluna *“mean\_price”* da tabela *“performance.csv”* onde os dados foram inseridos. Finalmente, os jogadores considerados com preço fora do comum são aqueles cujo preço é maior ou igual a LS.

Para definir o desempenho do jogador, usamos as colunas “*yellow\_cards*”, “*red\_cards*”, “*goals*”e “*assists*” da tabela “*appearances.csv*”

**3 CONSIDERAÇÕES FINAIS**

REGRrygxcbgszdegsr

**REFERÊNCIAS**

**GLADWELL, Malcolm.** Outliers: The Story of Success. 1. ed. Nova York: Little, Brown and Company, 2008.

**CARIBOO, David.** Player Scores. Kaggle, 2023. Disponível em: <https://www.kaggle.com/datasets/davidcariboo/player-scores>. Acesso em: 02 out. 2024.