

RELATÓRIO DE ANÁLISE EXPLORATÓRIA DE DADOS – JOGOS OLÍMPICOS (1896–2016)

1 INTRODUÇÃO

O presente relatório tem como objetivo apresentar uma análise exploratória de dados referente à história dos Jogos Olímpicos, com base no conjunto de dados “120 years of Olympic history: athletes and results”, disponível na plataforma Kaggle. A análise abrange os dados de atletas, modalidades, países, resultados e outras variáveis, considerando os eventos realizados entre os anos de 1896 e 2016. O estudo contempla as etapas fundamentais de um processo profissional de análise de dados, com foco na exploração, limpeza, preparação e visualização de informações.

2 METODOLOGIA

A base de dados utilizada consiste em duas tabelas principais:

- athlete_events.csv: contém informações sobre atletas, países (NOC), esportes, eventos, altura, peso, idade e medalhas;
- noc_regions.csv: mapeia os códigos dos comitês olímpicos nacionais (NOC) para os respectivos países e regiões.

O tratamento dos dados foi realizado com a linguagem de programação Python, utilizando bibliotecas como Pandas, Seaborn e Plotly. Foram tratadas inconsistências, valores ausentes e realizadas manipulações para criação de novas variáveis derivadas. A seguir, apresentam-se os principais achados.

3 PRINCIPAIS INSIGHTS

3.1 Evolução da Participação Olímpica

Verificou-se um crescimento expressivo no número de atletas ao longo dos anos, sobretudo após a Segunda Guerra Mundial. A partir da década de 1980, observa-se uma aceleração ainda mais significativa, refletindo o aumento da abrangência e importância dos Jogos Olímpicos no cenário global.

3.2 Países com Maior Número de Medalhas

Os Estados Unidos lideram historicamente o quadro de medalhas, seguidos por países como a Rússia (e antiga URSS), Alemanha, Grã-Bretanha e China. Tal domínio está relacionado a fatores como investimento em esporte, estrutura organizacional e tradição olímpica.

3.3 Crescimento da Participação Feminina

Houve um notável crescimento na participação de atletas do sexo feminino ao longo das edições dos Jogos. Em 1896, não havia mulheres competindo, e atualmente, a proporção entre os gêneros é praticamente equilibrada, o que reflete avanços em políticas de igualdade de gênero no esporte.

3.4 Perfil Físico dos Atletas por Esporte

A análise de altura e peso dos atletas demonstra perfis distintos conforme a modalidade. Esportes como basquete e remo concentram atletas mais altos e pesados, enquanto ginástica artística e corrida de longa distância apresentam padrões mais leves e baixos.

3.5 Faixas Etárias e IMC

A criação de uma coluna de faixa etária (“grupo etário”) e a derivação do Índice de Massa Corporal (IMC) permitiram uma visão mais aprofundada sobre os perfis dos atletas. Verificou-se que modalidades como ginástica e natação concentram jovens, enquanto hipismo e tiro esportivo envolvem atletas com idade média mais elevada.

4 VISUALIZAÇÕES

Foram criadas cinco visualizações principais, com o intuito de facilitar a compreensão dos achados:

1. Gráfico de linha com a evolução do número de atletas por edição;
2. Gráfico de barras com os 10 países mais premiados;
3. Gráfico de pizza com a proporção por sexo;

4. Gráfico de boxplot com a distribuição do IMC por grupo etário;
5. Mapa coroplético com a distribuição geográfica de medalhas.

As visualizações foram desenvolvidas com a biblioteca Plotly, permitindo interatividade e melhor interpretação dos dados.

5 CONSIDERAÇÕES FINAIS

A análise histórica dos Jogos Olímpicos permite observar transformações sociais, esportivas e políticas ocorridas ao longo de mais de um século. A crescente inclusão de mulheres, a diversificação geográfica da participação e o aumento de atletas refletem mudanças relevantes no cenário esportivo mundial. Os dados analisados oferecem subsídios valiosos para pesquisadores, profissionais do esporte e formuladores de políticas públicas.

6 REFERÊNCIAS

KAGGLE. 120 years of Olympic history: athletes and results. Disponível em: <<https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-results>>. Acesso em: 15 abr. 2025.