

Especificações do Mini-Projeto

Prazo final: 09/12 — 23:59

Os alunos que tiverem interesse em participar da 2ª etapa do projeto deverão submeter esse mini-projeto até a data limite. De todos, apenas dez alunos serão selecionados para continuar na 2ª etapa do PET, então faça o seu melhor!

Objetivo do Mini-Projeto

Neste projeto, o objetivo é gerar um relatório com base nos CSVs disponibilizados no arquivo “FIFA - World Cup (2006 - 2018).zip” anexado à atividade. O projeto será dividido em três partes e em cada uma delas deve conter uma breve descrição do que foi realizado e a justificativa das decisões que foram tomadas (bibliotecas, abordagens...).

Formato da entrega: deve ser feita usando o Google Colab e o .ipynb deve ser anexado na atividade.

Arquivos CSV: FIFA - 2006.csv, FIFA - 2010.csv, FIFA - 2014.csv e FIFA - 2018.csv.

Esses CSV são dados das seleções nas copas do mundo de 2006, 2010, 2014 e 2018.

Referência: FIFA - Football World Cup Dataset | Kaggle

Variáveis de interesse: Goals For e Goals Against.

Tratamento dos dados

Objetivo: Realizar a abertura dos arquivos e ler apenas as variáveis de interesse dos **quinze primeiros** colocados das copas de cada arquivo. A lista de valores de cada variável devem ser a junção de todos os valores contidos daquela variável nos 3 CSVs.

Exemplo:

FIFA - 2006.csv, Goals For = 6, 4, 1, 2, 8.

FIFA - 2010.csv, Goals For = 1, 3, 4, 5, 6.

FIFA - 2014.csv, Goals For = 5, 4, 2, 3, 6.

FIFA - 2018.csv, Goals For = 9, 8, 4, 5, 7.

Lista de valores da variável “Goals For”: 6, 4, 1, 2, 8, 1, 3, 4, 5, 6, 5, 4, 2, 3, 6, 9, 8, 4, 5, 7.

Análise estatística

Objetivo: Geração de média, desvio padrão e limites superiores e inferiores de cada variável de interesse. Nesse projeto, o limite superior é dado a partir da soma de 3 desvios padrão à média e o limite inferior é dado a partir da subtração de 3 desvios padrão da média.

Exemplo:

Média: 6 / **Desvio padrão:** 0,5

Limite superior: $6 + 3 \cdot (0,5) = 7,5$ / **Limite inferior:** $6 - 3 \cdot (0,5) = 4,5$

Plotagem de gráficos

Objetivo: Gerar um gráfico de linha para cada variável escolhida. Esses gráficos devem ter, além de uma linha para os valores da variável, uma linha para o limite inferior e uma para o limite superior. Ou seja, cada gráfico terá 3 linhas.