



Faculdade Senac Porto Alegre

Trabalho Avaliativo 01

Unidade Curricular	Análise de Dados e Inteligência Empresarial	Avaliação	Individual
Professor	Luis Vaz	Ano/Sem	2022/02

Competências avaliadas:

- Saber interpretar e seguir as instruções solicitadas;
- Saber apresentar os conceitos sobre os tópicos solicitados;
- Saber utilizar os comandos da Linguagem Python;
- Saber apresentar uma resposta viável para os problemas propostos;

Regras Gerais:

- O presente trabalho **DEVE** ser realizado individualmente;
- Quando identificado o plágio nas respostas para as questões a nota dos envolvidos na respectiva questão serão **ZERADAS**. Será considerado plágio a cópia entre colegas ou cópia de material da internet;

Instruções do Trabalho:

- A data limite para a entrega do trabalho é o dia **17/10/2022 às 19:00:00**;
- A entrega do trabalho DEVE ser realizada pela plataforma **Blackboard**.
- Não será aceito o envio total ou parcial do trabalho por e-mail;
- Características deste trabalho:
 - O presente trabalho pode conter questões dissertativas, objetivas e práticas (construção de scripts);
 - As respostas para todas as questões devem constar de um **ÚNICO ARQUIVO**.
 - Para as questões que exigem respostas dissertativas ou objetivas estas devem ser apresentada como um comentário ou markdown conforme o arquivo utilizado para a entrega;
 - Todas as respostas devem estar devidamente identificadas em relação as instruções aqui definidas;
 - Foram fornecidos na pasta de Conteúdo do Blackboard dois modelos para a entrega do trabalho:
 - modelo script python (extensão “.py”)
 - modelo script jupyter notebook (extensão “.ipynb”);
 - O nome do aluno **DEVE CONSTAR** dentro do arquivo de script;

Importante: as questões que envolvem o desenvolvimento de scripts serão submetidas ao interpretador python para a sua validação.



Questões

Todas as questões a seguir devem utilizar o arquivo de dados “**nba.csv**” o qual apresenta informações de diversos atletas da liga americana de basquete (NBA).

Estrutura simplificada do *dataset*:

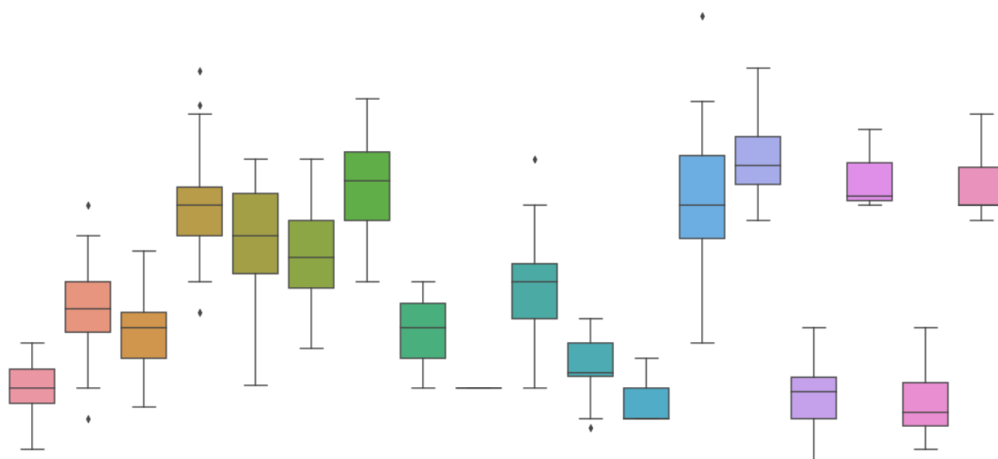
Campo	Descrição
Name	Nome do atleta
Team	Equipe do atleta
Number	Número utilizado pelo atleta na equipe
Position	Posição do atleta: <ul style="list-style-type: none">• PG = Point Guard (armador)• SF = Small Forward (ala)• SG = Shooting Guard (ala-armador)• PF = Power Forward (ala-pivô)• C = Center (pivô)
Age	Idade do atleta
Height	Altura do atleta (unidade: pés e polegadas)
Weight	Peso do atleta (unidade: libras)
College	Universidade de origem do atleta
Salary	Salário anual do atleta

- Q1.** Como você avalia os dados deste *dataset*. Existem valores NULOS? Se sim, quantos e quais. Demonstre.
- Q2.** Com relação a posição dos atletas, apresente a distribuição de frequência deste parâmetro (absoluta e percentual). Apresente também o histograma com a linha de distribuição de frequência para estes dados. Apresente uma análise sobre o resultado do diagrama gerado.
- Q3.** Faça uma segmentação da base de dados com base na idade dos atletas. Para esta atividade deve ser realizada uma classificação fixa onde: o menor valor é 18 o maior valor é 45 e a amplitude das classes é 3. Apresentar a informação de idade utilizando os gráficos de boxplot e histograma (com linha de distribuição de frequência). Realizar uma breve análise dos resultados obtidos.
- Q4.** Apresentar a média do salário mensal dos atletas agrupados por equipe e posição. Após apresentar um gráfico de histograma com os valores médios dos salários.

- Q5.** Faça uma análise e prepare uma tabela que apresente a frequência dos atletas relacionando as informações de Altura, Peso e Posição. Após construa um gráfico boxplot com as informações de peso e altura. Com estas duas informações apresente uma análise relacionada a estes dois atributos.

DICA: utilizar a função de agregação 'count'.

Observação: o gráfico gerado nesta questão deve ter um aspecto conforme a imagem apresentada a seguir:



Bom Trabalho!
Prof. Luis Vaz