

Curso de Especialização em *Data Science* e Estatística Aplicada

Módulo I - Estatística descritiva para *Data Science* Atividade Avaliativa - Parte I

Profa. Dra. Amanda Buosi Gazon Milani

2024-07-20

Instruções

- O desenvolvimento desta atividade deve ser realizada de forma individual.
- Deve-se completar o arquivo Rmd enviado na atividade.
- É necessário devolver o arquivo em Rmd e em pdf.
- Valor da atividade: 10 pontos.

Data disponível: 19/07/2024 (Parte 1) e 03/08/2024 (Parte 2).

Data máxima para entrega: 30/08/2024

Atividade - Parte I

Conjunto de dados - Notificações de Síndrome Gripal 2024

O Ministério da Saúde implementou a Vigilância da Síndrome Gripal (SG) de casos leves e moderados suspeitos de Covid-19 (doença causada pela infecção pelo vírus SARS-Cov-2). Para a atividade avaliativa o conjunto de dados que será analisado será o de notificações de síndrome gripal de casos leves e moderados suspeitos de Covid-19 do Estado de Goiás do ano de 2024. Os dados foram obtidos no site do OpenDataSUS <https://opendatasus.saude.gov.br/dataset/notificacoes-de-sindrome-gripal-leve-2024> (acesso e download em 16/07/2024). O conjunto de dados e seu respectivo dicionário foram postados no MOODLE Ipê.

A Parte I da Atividade Avaliativa consiste nos itens a seguir.

Itens - Parte I

Desenvolva e/ou responda os itens a abaixo.

Item 1 - Faça uma tabela de frequências e responda à seguinte pergunta: Qual a porcentagem dos profissionais da saúde com notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás? Utilize duas casas decimais.

Item 2 - Faça uma tabela de frequências e responda à seguinte pergunta: Qual é o sexo que apresenta o maior número de notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás?

Item 3 - Faça uma tabela de frequências e responda à seguinte pergunta: Qual a raça que apresenta a maior proporção de notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás? Utilize quatro casas decimais.

Item 4 - Faça um gráfico em setores para a variável “recebeu vacina” (`codigoRecebeuVacina`) das notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás. Utilize a frequência relativa em porcentagem. Além disso, responda: Qual é a categoria mais frequente nas notificações? Para o gráfico, adicione os nomes das categorias e as porcentagens como rótulos. Utilize uma casa decimal.

Item 5 - Faça um gráfico em barras para a variável “classificação final” (`classificacaoFinal`) das notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás. Utilize a frequência absoluta (em centenas de notificações). Além disso, analisando o gráfico, responda: Qual é a classificação mais frequente?

Item 6 - Faça um histograma para a idade (em anos) das pessoas que tiveram notificações de suspeita de Covid-19 no ano de 2024 no Estado de Goiás. Utilize a densidade de frequência e classes de amplitude 10 (com a primeira iniciando na idade 0). Além disso, analisando o gráfico, responda: Qual é a faixa etária mais frequente?

Observações:

- (i) Antes de responder aos itens, é preciso manipular o conjunto de dados;
- (ii) A exclusão dos *NA*'s deve ser realizada considerando apenas dados faltantes nas variáveis de interesse;
- (iii) A entrega da atividade deve ser realizada em Rmd e um arquivo no formato pdf, gerado a partir do RMarkdown, juntamente com a resolução da Parte II (entrega única).