## **Desafio Final**

**Entrega** 20 nov em 23:59 **Pontos** 100 **Perguntas** 15 **Disponível** até 20 nov em 23:59 **Limite de tempo** Nenhum

# Instruções

### O Desafio Final está disponível!

### 1. Instruções para realizar o desafio

Consulte a data de entrega no teste e em seu calendário.

Reserve um tempo para realizar a atividade, leia as orientações e enunciados com atenção. Em caso de dúvidas utilize o "Fórum de dúvidas do Desafio Final".

Para iniciá-lo clique em "Fazer teste". Você tem somente **uma** tentativa e não há limite de tempo definido para realizá-lo. Caso precise interromper a atividade, apenas deixe a página e, ao retornar, clique em "Retomar teste".

Clique em "Enviar teste" **somente** quando você concluí-lo. Antes de enviar confira todas as questões.

Caso o teste seja iniciado e não enviado até o final do prazo de entrega, a plataforma enviará a tentativa não finalizada automaticamente, independente do progresso no teste. Fique atento ao seu teste e ao prazo final, pois novas tentativas só serão concedidas em casos de questões médicas.

O gabarito será disponibilizado partir de sexta, 20/11/2020, às 23h59.

Bons estudos!

### 2. O arquivo abaixo contém o enunciado do desafio

Enunciado do Desafio Final - Cientista de Dados.pdf

### Histórico de tentativas

	Tentativa	Tempo	Pontuação
MAIS RECENTE	Tentativa 1	475 minutos	100 de 100

(!) As respostas corretas estarão disponíveis em 20 nov em 23:59.

Pontuação deste teste: 100 de 100

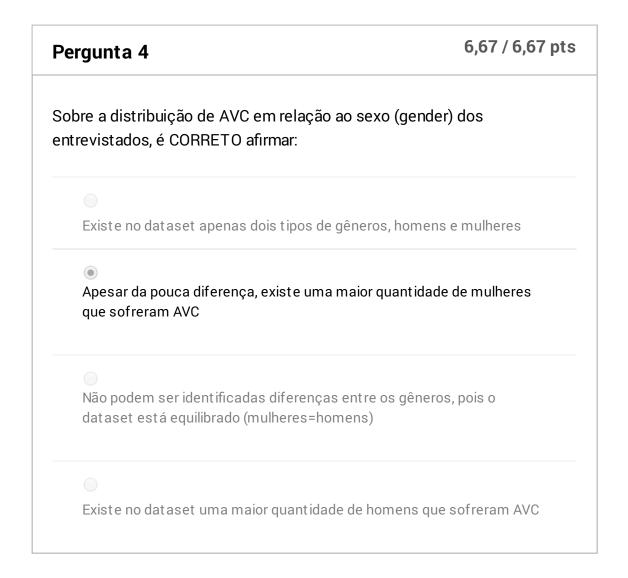
Esta tentativa levou 475 minutos.

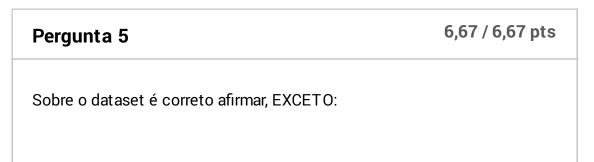
Pergunta 1	6,67 / 6,67 pts
Quantas instâncias e características existem, respectivo no dataset?	amente,
(5110, 12)	
(5000, 7)	
(12, 5110)	
<b>(7, 5000)</b>	

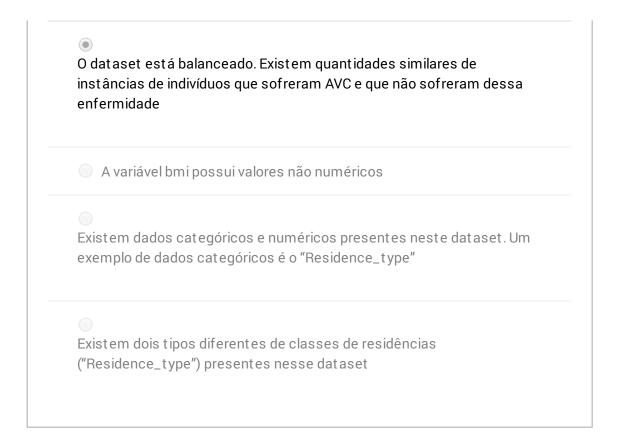
Pergunta 2	6,67 / 6,67 pts
Quantas variáveis do tipo "string" estão presentes no o	dataset?
<ul><li>6</li></ul>	
O 2	
<b>4</b>	
<b>3</b>	

Pergunta 3	6,67 / 6,67 pts
Qual é a idade (age) média dos entrevistados?	

<ul><li>43,22 anos</li></ul>	
45,28 anos	
22,61 anos	
55,12 anos	

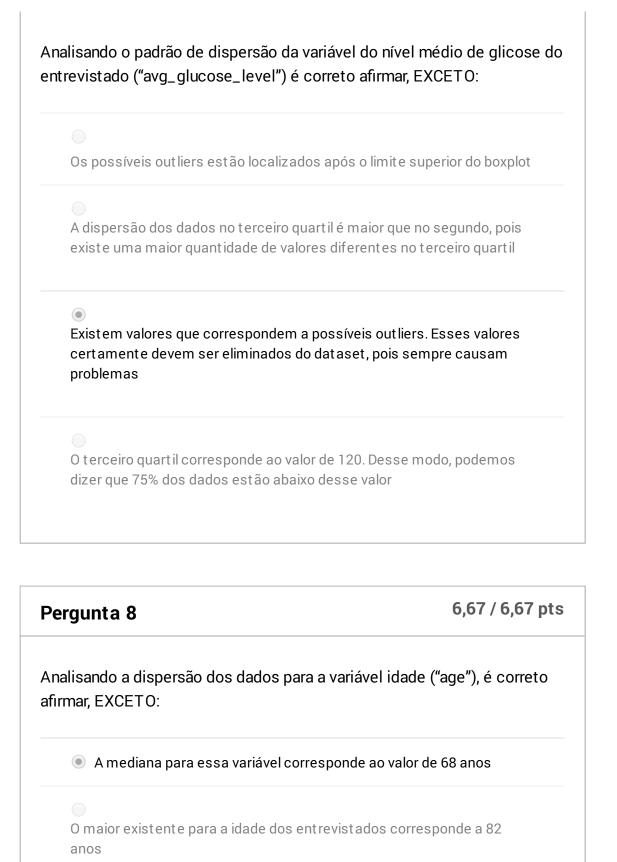






Pergunta 6	6,67 / 6,67 pts
Qual é o valor da mediana para a variável do nível médio entrevistado ("avg_glucose_level")?	de glicose do
O 120	
O 271,74	
<ul><li>95</li></ul>	
O 78	

Pergunta 7 6,67 / 6,67 pts



Pelo Boxplot n\u00e3o \u00e9 poss\u00edvel identificar poss\u00edvels outliers

O primeiro quartil indica que 25% dos dados estão abaixo de 30 anos

Pergunta 9	6,67 / 6,67 pts
Quantas classes diferentes para a variável "work_type no dataset?	e" existem
<ul><li>2</li></ul>	
<b>6</b>	
<ul><li>5</li></ul>	
<b>4</b>	

Pergunta 10	6,67 / 6,67 pts
Dentre as classes de tipos de trabalhos existentes (wo aquela que possui uma maior quantidade de instâncias?	
Govt_job	
Private	
Self-employed	
Never_worked	

Pergunta 11

6,67 / 6,67 pts

# Qual foi, respectivamente, o percentual de dados utilizados para o treinamento e teste do modelo? (30%, 70%) (80%, 20%) (70%, 30%) (20%, 80%)

# Analisando as variáveis "bmi" e "smoking\_status", é CORRETO afirmar: A variável "bmi" possui apenas valores numéricos Ambas são variáveis numéricas Existem oito classes distintas de "smoking\_status" Ambas possuem instâncias com valores desconhecidos

### Pergunta 13

6,67 / 6,67 pts

Após o agrupamento dos dados de 'smoking\_status' e 'stroke', é CORRETO afirmar que:

<ul><li>Existem seis classes difer</li></ul>		
	ntes de "smoki	ing_status"
Não é possível realizar o agrup dimensões diferentes	mento, pois os	dados possuem
Dentre os entrevistados que s quantidade de indivíduos da cla		

# Pergunta 14

6,67 / 6,67 pts

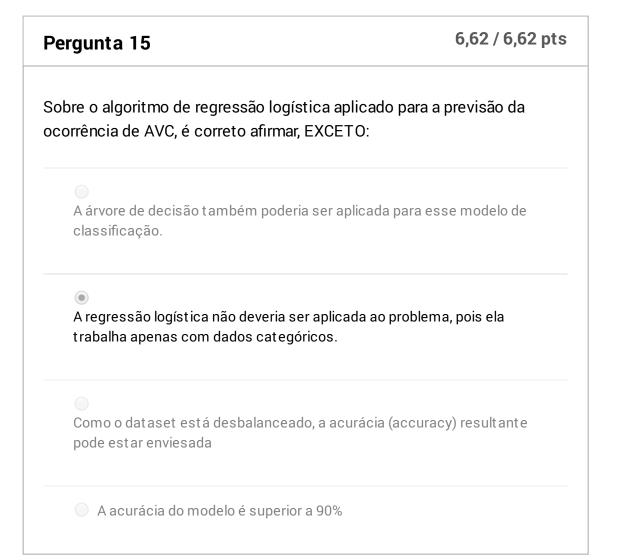
Sobre a relação entre a hipertensão (hypertension) e o AVC (stroke) presente neste dataset, é CORRETO afirmar:

Os dados mostram que este dataset está balanceado.

A proporção de incidência de AVC é maior nos indivíduos que sofrem de hipertensão.

Existe uma maior quantidade de dados de indivíduos hipertensos.

A proporção entre indivíduos hipertensos e não hipertensos no dataset é a mesma.



Pontuação do teste: 100 de 100