

# Trabalho Prático do Módulo 4

---

<b>Entrega</b> 6 nov em 19:00	<b>Pontos</b> 25	<b>Perguntas</b> 15
<b>Disponível</b> até 6 nov em 19:00	<b>Limite de tempo</b> Nenhum	

---

## Instruções

O Trabalho Prático do Módulo 4 está disponível!

### 1. Instruções para realizar o trabalho prático

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 Trabalho Prático do Módulo 4".

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-feira, **06/11/2020**, às 23h59.

Bons estudos!

### 2. O arquivo abaixo contém o enunciado do trabalho prático

[Enunciado do Trabalho Prático - Módulo 4 - Bootcamp Analista de Dados.pdf](#) 

## Histórico de tentativas

	Tentativa	Tempo	Pontuação
MAIS RECENTE	<a href="#">Tentativa 1</a>	605 minutos	25 de 25

---

⚠ As respostas corretas estarão disponíveis em 6 nov em 23:59.

Pontuação deste teste: **25** de 25

Enviado 1 nov em 7:25

Esta tentativa levou 605 minutos.

### Pergunta 1

1,66 / 1,66 pts

Qual é a biblioteca em Python utilizada para criação de dataframes?

☐ Matplotlib.

☒ Pandas.

☐ Numpy.

☐ Seaborn.

### Pergunta 2

1,66 / 1,66 pts

Quantas linhas (instâncias) e colunas (características) existem, respectivamente, no dataset?

☐ (7, 125).

☒ (440, 8).

☐ (125, 7).

☐ (8, 440).

### Pergunta 3

1,66 / 1,66 pts

Para o dataset presente no trabalho, quantos valores nulos existem para a coluna de gastos anuais com leite (Milk)?

☐ 440.

☒ 0.

☐ 8.

☐ 14.

#### Pergunta 4

1,66 / 1,66 pts

Para o dataset presente no trabalho, qual é o valor médio gasto anualmente com compra de derivados do leite (Milk)?

☒ 5796,2.

☐ 7380,3.

☐ 55,0.

☐ 7951,2.

#### Pergunta 5

1,66 / 1,66 pts

Qual é o maior valor anual gasto com produtos congelados (Frozen)?

☐ 40827,0.

☐ 5796,2.

☐ 3071,9.

☒ 60869,0.

#### Pergunta 6

1,66 / 1,66 pts

Analizando o boxplot das colunas "Fresh" e "Grocery" podemos dizer, EXCETO:



A mediana dos dados da coluna "Fresh" é superior ao dos dados da coluna "Grocery".



Para a coluna "Fresh" existem valores superiores a 100000.



A dispersão dos dados da coluna "Fresh" é maior que os da coluna "Grocery".



Não existem possíveis outliers nos dados.

### Pergunta 7

1,66 / 1,66 pts

Analizando o gráfico de distribuição dos consumidores quanto à região, é CORRETO dizer:



Existem 3 vezes mais consumidores entrevistados na região 2 que na região 3.



A menor quantidade de consumidores está presente na região 1.



Existem mais consumidores na região 3.



Existem 5 vezes mais consumidores entrevistados na região 1 que na região 2.

### Pergunta 8

1,66 / 1,66 pts

Analizando a distribuição dos gastos de consumidores com produtos derivados do leite (Milk) é CORRETO afirmar:



A maior concentração de consumidores encontra-se com gastos de até 10000 euros com produtos derivados do leite.



Existem consumidores que gastam mais de 300000 euros por ano com produtos derivados do leite.



Não existem consumidores que gastem mais de 50000 euros por ano com produtos derivados do leite.



Mais de 50% dos consumidores gastem mais que 20000 euros com produtos derivados do leite.

### Pergunta 9

1,66 / 1,66 pts

Analizando os gastos dos consumidores por tipo de canal de compra (1 e 2), é CORRETO afirmar que:



Através do canal 1 existe a maior quantidade de clientes que gastam mais de 40000 euros por ano com produtos derivados do leite.



Os consumidores realizam compras de derivados do leite apenas através do canal 2.



A maioria dos consumidores gastam menos de 10000 euros com produtos derivados do leite através do canal 1.



Os consumidores gastam mais através do canal 1.

### Pergunta 10

1,66 / 1,66 pts

Analisando o histograma do gasto anual de consumidores com produtos de supermercado, é correto afirmar, EXCETO:



Não existem consumidores que gastem mais de 60000 euros anuais com produtos de supermercado.



A maioria dos consumidores gastam até 20000 euros anuais com supermercado.



À medida que os gastos com supermercado aumentam, existe uma redução do número de consumidores para cada faixa de valor gasto (bins).



Uma pequena parcela dos consumidores gasta entre 40000 e 60000 euros com supermercado ao ano.

### Pergunta 11

1,66 / 1,66 pts

Analisando os gastos com produtos de supermercado por região, é CORRETO afirmar:



Não existem consumidores da região 2 que gastam mais de 20000 euros por ano com supermercado.



Os consumidores da região 3 tendem a gastar mais com produtos de supermercado.



Os consumidores da região 2 são os que mais gastam com supermercado, pois possuem maior poder aquisitivo.



A região 1 possui a maior quantidade de consumidores entrevistados.

### Pergunta 12

1,66 / 1,66 pts

Analisando a correlação de Pearson entre os gastos com produtos de supermercado (Grocery) e produtos de higiene (Detergents\_Paper), qual é o valor de coeficiente?



-0,35.



-0,92.



0,92.



0,35.

### Pergunta 13

1,66 / 1,66 pts

Após a aplicação do modelo de regressão linear para a previsão de gastos com produtos de higiene baseado nos gastos com produtos de

supermercado, qual é o valor previsto para os gastos com produtos de higiene, dado que o cliente gastou 30000 euros com supermercado?

☐ 35234.

☒ 12739.

☐ 9546.

☐ 42716.

#### Pergunta 14

1,66 / 1,66 pts

Dada a segmentação dos consumidores com base nos gastos anuais com produtos de higiene e compra de supermercados, é CORRETO afirmar:



O grupo de consumidores que gastam muito com produtos de supermercado e muito com produtos de higiene é o menor existente.



Todos os clusters (grupos de clientes) gastam a mesma quantidade com supermercado.



O grupo de consumidores que gastam pouco com supermercado e pouco com produtos de higiene é o menor existente.



O maior grupo de clientes corresponde aqueles que gastam muito com supermercado e muito com produtos de higiene

#### Pergunta 15

1,76 / 1,76 pts



Sobre o coeficiente de determinação encontrado para a reta do trabalho prático, é CORRETO afirmar:



Possui um valor muito próximo de 1, indicando que ocorreu um bom "fit" para os dados.



Nada pode ser dito, pois apresenta um valor próximo a zero.



Não pode ser calculado.



Apresenta valor negativo.

Pontuação do teste: **25** de 25