

⚠ Este teste foi reavaliado; sua nova pontuação reflete 3 perguntas que foram afetadas.

# Trabalho Prático do Módulo 1

**Entrega** 20 mai em 17:00

**Pontos** 25

**Perguntas** 12

**Disponível** até 20 mai em 17:00

**Limite de tempo** Nenhum

## Instruções

### Informações sobre o Trabalho Prático

[TP.py](#) 

O Trabalho Prático do Módulo 1 está disponível! 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".

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.

O gabarito será disponibilizado partir de sexta-feira, 22/05/2020, às 21h.

Bom trabalho!

### Objetivos

Exercitar os seguintes conceitos trabalhados no Módulo 1:

- Uso do Google Colab
- Uso do Power BI
- Algoritmo K-Means

### Enunciado

Para as questões de 1 a 7 trabalharemos com python e o algoritmo K-Means. Serão fornecidos alguns trechos de código para auxiliar na execução. Considere um dataset gerado pelo seguinte comando:

```
make_blobs(n_samples=500, centers=20, random_state=999)
```

Ao executar o algoritmo K-Means, considere os parâmetros:

- `init = k-means++`
- `max_iter = 300`

- `n_init = 10`

Para as questões de 8 a 12, instale o [Power BI Desktop](https://powerbi.microsoft.com/pt-br/downloads/) e baixe o dataset [Novel Corona Virus 2019](https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=covid_19_data.csv) e carregue os dados no Power BI. Atenção aos tipos de dados de cada coluna. Crie uma nova medida chamada “Active” para calcular o número de casos ativos, que é no caso o valor da coluna “Confirmed” menos a soma das colunas “Recovered” e “Deaths”. Crie gráficos para facilitar a sua visualização, de acordo com a questão.

Este teste foi indisponível 20 mai em 17:00.

## Histórico de tentativas

	Tentativa	Tempo	Pontuação	Reavaliado
MAIS RECENTE	<a href="#">Tentativa 1</a>	1.102 minutos	22,92 de 25	25 de 25

Pontuação deste teste: **25** de 25

Enviado 20 mai em 15:10

Esta tentativa levou 1.102 minutos.

### Pergunta 1

2,08 / 2,08 pts

Considere o dataset gerado a partir das instruções do enunciado. Caso o algoritmo seja executado buscando 5 clusters, qual será a soma dos quadrados intra-clusters (wcsc) ao final da execução?

- ☐ 1853.7783832527882
- ☐ 9464.847250570496
- ☐ 6005.101775346179
- ☒ 4717.404091931214

Correto!

### Pergunta 2

2,08 / 2,08 pts

Considere o dataset gerado a partir das instruções do enunciado. Caso o algoritmo K-Means seja executado buscando apenas 1 cluster, qual será a coordenada do centroide ao final da execução?

- ☐ [1.9050166 4.91200285]
- ☐ [ 2.40603415 6.29715718]
- ☒ [ 0.02618523 -1.30384304]
- ☐ [1.03796387 -5.49096674]

**Correto!**

### Question 3

Pontuação original: 0 / 2,08 pts **Pontuação reavaliada: 2,08 / 2,08 pts**

⚠ Esta pergunta foi reavaliada.

Considere o dataset gerado a partir das instruções do enunciado. Qual o número ideal de clusters?

**Você respondeu**

- ☒ 5
- ☐ 7
- ☐ 3
- ☐ 4

**Resposta correta**

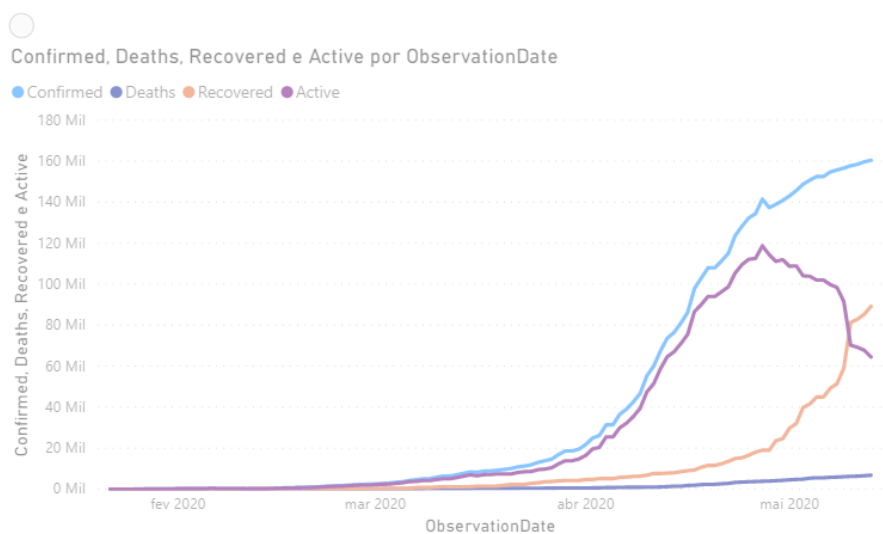
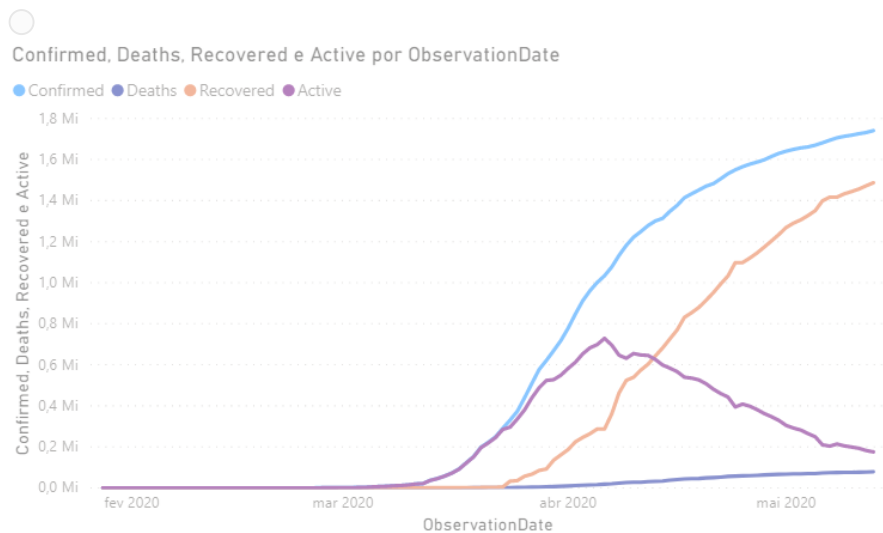
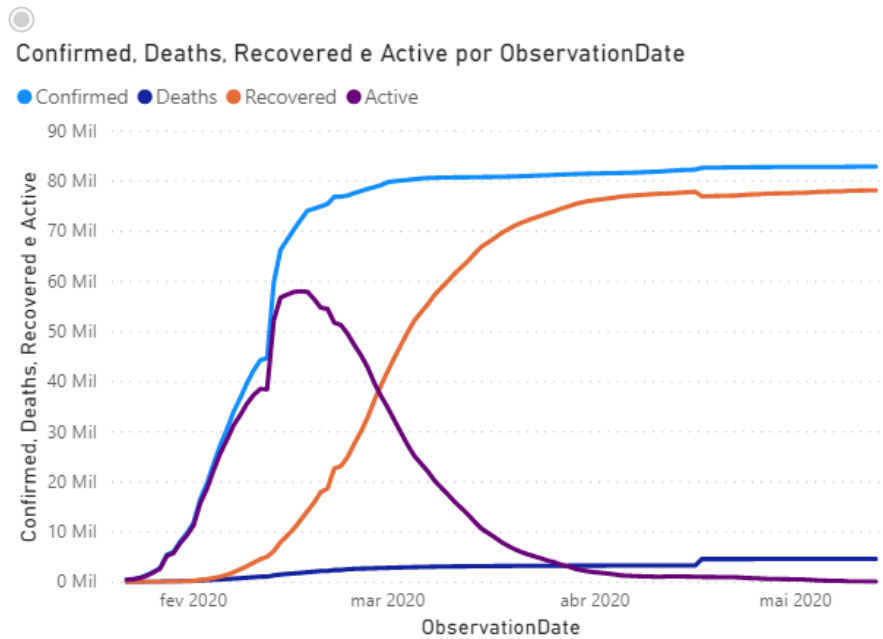
### Pergunta 4

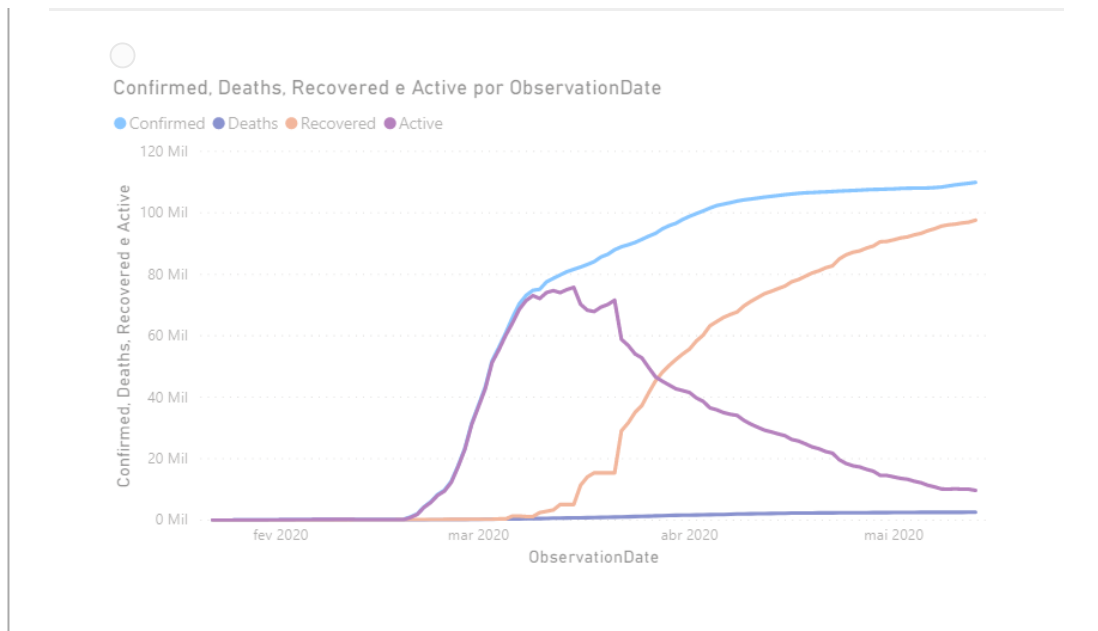
**2,08 / 2,08 pts**

Utilize o Power BI e o dataset das instruções do enunciado. Ao criar um gráfico com as medidas “Confirmed”, “Deaths”, “Recovered” e

“Active” apenas para a China (Mainland China), como fica o gráfico?

Correto!



**Pergunta 5****2,08 / 2,08 pts**

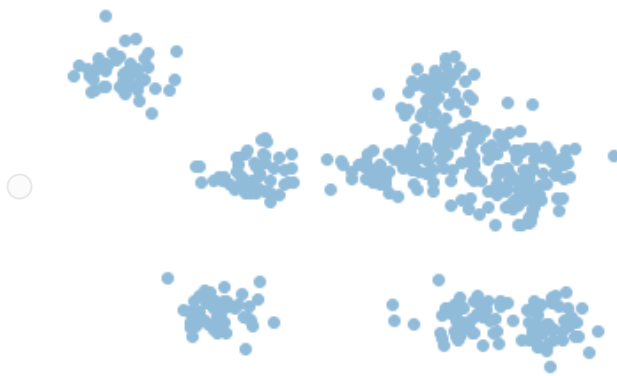
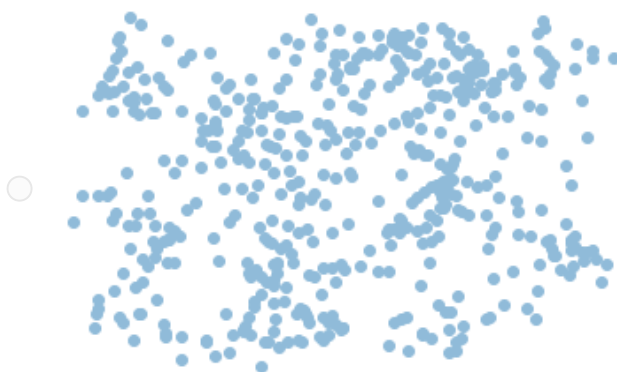
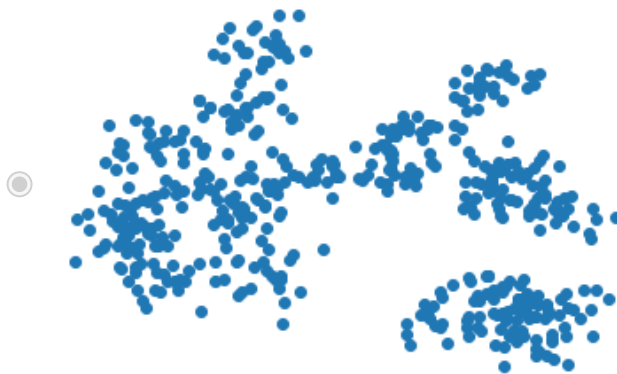
Qual o objetivo do algoritmo K-Means?

**Correto!**

- ☒ Clusterização
- ☐ Classificação
- ☐ Correlação
- ☐ Seleção

**Pergunta 6****2,08 / 2,08 pts**

Considerando o dataset gerado a partir das instruções do enunciado, qual das imagens representa o gráfico correto?

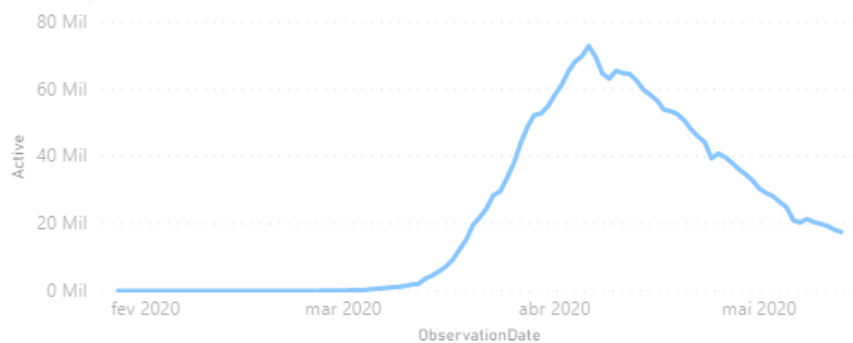
**Correto!****Pergunta 7****2,08 / 2,08 pts**

Utilize o Power BI e o dataset das instruções do enunciado. Ao criar um gráfico com a medida “Active”, apenas para a Itália (Italy), como

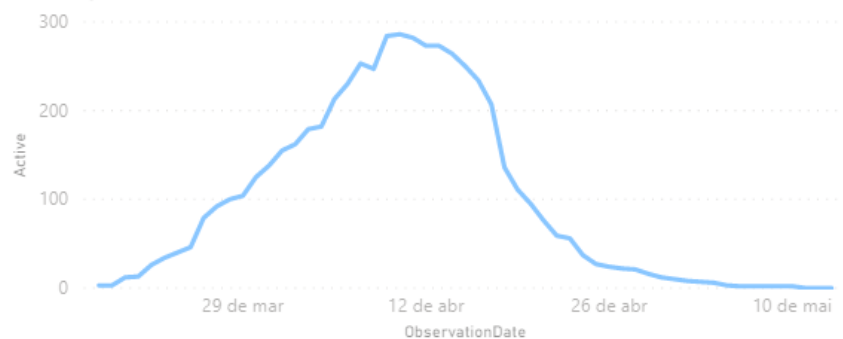
fica o gráfico?



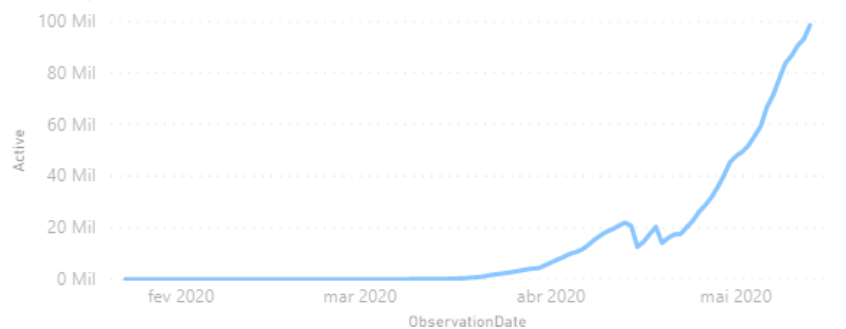
Active por ObservationDate



Active por ObservationDate



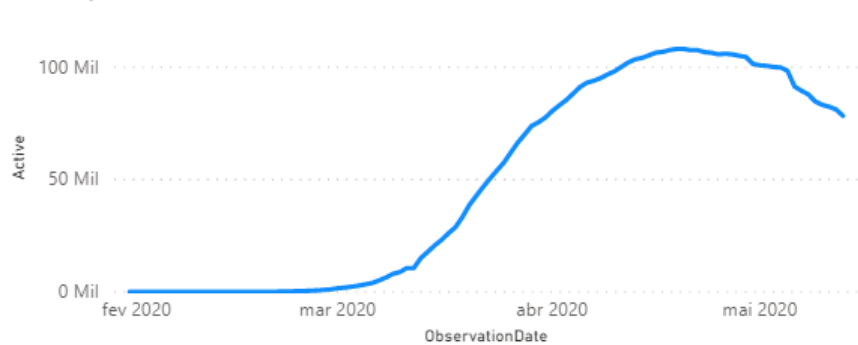
Active por ObservationDate



Correto!



Active por ObservationDate



**Question 8**

Pontuação original: 2,08 / 2,08 pts **Pontuação reavaliada: 2,08 / 2,08 pts**

⚠ Esta pergunta foi reavaliada.

Utilize o Power BI e o dataset das instruções do enunciado. Filtre por país, selecionando a Itália (Italy). Analise as informações do dataset e responda: em que mês a Itália passou a ter mais casos recuperados do que casos ativos?

Correto!

- ☒ Maio
- ☐ Março
- ☐ Abril
- ☐ Fevereiro

**Pergunta 9**

2,08 / 2,08 pts

O algoritmo K-Means é baseado em qual tipo de aprendizagem?

Correto!

- ☐ Supervisionada
- ☒ Não supervisionada
- ☐ Por reforço
- ☐ Aleatória

**Pergunta 10**

2,08 / 2,08 pts



Considere o dataset gerado a partir das instruções do enunciado. Caso o algoritmo seja executado buscando apenas 1 cluster, qual será a soma dos quadrados intra-clusters (wcss) ao final da execução?

**Correto!**

- ☒ 31547.372471390485
- ☐ 1068.2049059143994
- ☐ 34737.56779526727
- ☐ 61186.525771621105

### Question 11

Pontuação original: 2,08 / 2,08 pts **Pontuação reavaliada: 2,08 / 2,08 pts**

⚠ Esta pergunta foi reavaliada.

Utilize o Power BI e o dataset das instruções do enunciado. Filtre por país, selecionando a Coreia do Sul (South Korea). Em que mês esse país atingiu o pico de casos ativos?

**Correto!**

- ☒ Março
- ☐ Fevereiro
- ☐ Abril
- ☐ Janeiro

### Pergunta 12

2,12 / 2,12 pts

Utilize o Power BI e o dataset das instruções do enunciado. Filtre por país, selecionando a Venezuela, analise as informações, perceba que

algumas vezes o número de casos recuperados foi maior que o de casos ativos. Em que mês isso ocorreu pela primeira vez?

☐ Fevereiro

☐ Maio

Correto!

☒ Abril

☐ Março

Pontuação do teste: **25** de 25