

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/) [\(https://powerbi.microsoft.com/pt-br/downloads/\)](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) [\(https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=covid_19_data.csv\)](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 é 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.

Fazer o teste