

# Trabalho Prático do Módulo 3

**Entrega** 25 out em 23:59**Pontos** 25**Perguntas** 15**Disponível** até 25 out em 23:59**Limite de tempo** Nenhum

## Instruções

O Trabalho Prático do Módulo 3 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 3".

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 do domingo, **25/10/2020**, às 23h59.

Bons estudos!

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

[Enunciado do Trabalho Prático - Módulo 3 - Bootcamp Cientista de Dados.pdf](#)

[Clique aqui](https://drive.google.com/drive/u/1/folders/17Kv5wwkpVnJ0MPNPYn7uEckWmp2cRE-Y) [\(https://drive.google.com/drive/u/1/folders/17Kv5wwkpVnJ0MPNPYn7uEckWmp2cRE-Y\)](https://drive.google.com/drive/u/1/folders/17Kv5wwkpVnJ0MPNPYn7uEckWmp2cRE-Y) para fazer download da máquina virtual

## Histórico de tentativas

	Tentativa	Tempo	Pontuação
MAIS RECENTE	<a href="#">Tentativa 1</a>	46 minutos	25 de 25

⚠ As respostas corretas estarão disponíveis em 25 out em 23:59.

Pontuação deste teste: **25** de 25

Enviado 20 out em 19:12

Esta tentativa levou 46 minutos.

### Pergunta 1

1,67 / 1,67 pts

Em nossos exemplos, utilizamos nas funções Map e Reduce o método collector do objeto output (output.collector). Qual a função desse método, que está presente na interface OutputCollector?

- ☐ Encerrar a execução de todo o programa.
- ☐ Atribuir logs ao objeto output.
- ☒ Adicionar pares chave/valor (key/value) ao objeto output.
- ☐ Atribuir pares chave/valor (key/value) ao objeto jobconf.

### Pergunta 2

1,67 / 1,67 pts

Sabemos que no programa que você desenvolveu, os resultados foram armazenados em pares chave/valor no diretório PastaSaida do HDFS. Qual foi o valor (value) do dia (key) 27/03/2020?

- ☐ 313|Brazil|518|Equador.
- ☒ 8578|Spain|660|Italy.
- ☐ 649|China|97|China.
- ☐ 517|China|71|China.

### Pergunta 3

1,67 / 1,67 pts

Sabemos que no programa que você desenvolveu, os resultados foram armazenados em pares chave/valor no diretório PastaSaida do HDFS. Qual foi o valor (value) do dia (key) 26/01/2020?

☒ 688|China|15|China.

☐ 7324|Germany|625|Italy.

☐ 585|Germany|90|Italy.

☐ 4207|Italy|473|Italy.

#### Pergunta 4

1,67 / 1,67 pts

Sabemos que no programa que você desenvolveu, os resultados foram armazenados em pares chave/valor no diretório PastaSaida do HDFS. Qual foi o valor (value) do dia (key) 24/02/2020?

☒ 231|Republic of Korea|150|China.

☐ 23104|United States of America|1443|France.

☐ 30018|France|5006|United States of America.

☐ 802|Brazil|9087|France.

#### Pergunta 5

1,67 / 1,67 pts

No programa que você desenvolveu, quantas vezes a função Map foi executada?

☐ 64 vezes.

- ☐ O mesmo que o número de linhas do arquivo ExemploIGTI.java.
- ☒ O mesmo que o número de linhas que o arquivo covidData.txt possui.
- ☐ Uma vez para cada chamada ao método copyFromLocalFile.

**Pergunta 6****1,67 / 1,67 pts**

No programa que você desenvolveu, quantas vezes a função Reduce foi executada?

- ☐ Uma vez para cada execução da função Map.
- ☒ Uma vez para cada chave.
- ☐ 708 vezes.
- ☐ Uma vez para cada valor maior que zero.

**Pergunta 7****1,67 / 1,67 pts**

Ao utilizar o HDFS utilizando linhas de comando, como você poderia “imprimir na tela” os dados referentes ao resultado final do seu programa?

- ☒ 7. Ao utilizar o HDFS utilizando linhas de comando, como você poderia “imprimir na tela” os dados referentes ao resultado final do seu programa?
- ☐ `/usr/local/hadoop/bin/hdfs dfs -rmr /usr/igti/PastaSaida/part-00000`
- ☐ `/usr/local/hadoop/bin/hdfs /usr/igti/PastaSaida/part-00000`

☐ /usr/local/hadoop/bin/hdfs dfs -ls /usr/igti/PastaSaida/part-00000

## Pergunta 8

1,67 / 1,67 pts

No programa que você desenvolveu, qual seria a sequência CORRETA de eventos?

☐

- i) Leitura de dados do sistema de arquivos do sistema operacional.
- ii) Processamento da função Reduce.
- iii) Processamento da função Map com os dados de saída da função Reduce.
- iv) Gravação dos dados no HDFS na pasta rotulada como PastaSaida.

☐

- i) Leitura de dados do HDFS, na pasta rotulada como Dados.
- ii) Processamento da função Map.
- iii) Processamento da função Reduce com os dados de saída da função Map.
- iv) Gravação dos dados no HDFS na pasta rotulada como PastaEntrada.

☐

- i) Leitura de dados do HDFS, na pasta rotulada como Entrada.
- ii) Processamento da função Map.
- iii) Processamento da função Combine com os dados de saída da função Map.
- iv) Gravação dos dados no HDFS na pasta rotulada como Saida.



- i) Leitura de dados do HDFS, na pasta rotulada como PastaEntrada.
- ii) Processamento da função Map.
- iii) Processamento da função Reduce com os dados de saída da função Map.
- iv) Gravação dos dados no HDFS na pasta rotulada como PastaSaida.

## Pergunta 9

1,67 / 1,67 pts

No programa que você desenvolveu, existe um array chamado dadosCovid no método map da classe MapIGTI. Quantas posições existem nesse array e o que é armazenado nas posições 0, 2, 4 e 6?

São 4 posições.

Posição 0: Data do evento.

Posição 2: País do evento.

Posição 4: Quantidade de novos casos informado.

☐ Posição 6: Quantidade de novos óbitos informado.

São 8 posições.

Posição 1: Data do evento.

Posição 2: País do evento.

Posição 4: Quantidade de novos casos informado.

☐ Posição 6: Quantidade de novos óbitos informado.

São 8 posições.

Posição 0: Data do evento.

Posição 2: País do evento.

Posição 4: Quantidade de novos casos informado.

☒ Posição 6: Quantidade de novos óbitos informado.

São 4 posições.

Posição 1: Data do evento.

Posição 2: País do evento.

Posição 4: Quantidade de novos casos informado.

☐ Posição 6: Quantidade de novos óbitos informado.

## Pergunta 10

1,67 / 1,67 pts

Para responder esta questão, remova a execução do método `reduce` do seu programa, comentando o método `setReducerClass(...)`. Em seguida, execute o programa. Qual é o resultado dos dados gravados no HDFS?



O programa gravou pares chave/valor, separados por um TAB. A chave foi a data do evento e o valor é representado por uma concatenação de dados, sendo: país do evento, quantidade de novos casos e quantidade de novos óbitos.



O programa gravou pares chave/valor, separados por um TAB. A chave foi a data do evento e o valor é representado por uma concatenação de dados, sendo: quantidade de novos casos e quantidade de novos óbitos.



O programa gravou pares chave/valor, separados por uma barra vertical. A chave foi o país do evento e o valor é representado por uma concatenação de dados, sendo: data do evento, quantidade de novos casos e quantidade de novos óbitos.



O programa gravou pares chave/valor, separados por um arroba. A chave foi o país do evento e o valor é representado por uma concatenação de dados, sendo: data do evento, quantidade de novos casos e quantidade de novos óbitos.

**Pergunta 11****1,67 / 1,67 pts**

Sabemos que no programa que você desenvolveu, os resultados foram armazenados em pares chave/valor no diretório PastaSaida do HDFS. Qual foi o valor (value) do dia (key) 05/03/2020?

- ☐ 20803|Spain|1088|United States of America.
- ☒ 769|Italy|41|Italy.
- ☐ 15829|France|1785|United States of America.
- ☐ 11122|Spain|188|Brazil.

**Pergunta 12****1,67 / 1,67 pts**

Qual o conteúdo do argumento key (LongWritable) do método *reduce* da classe MapIGTI?

- ☐ O país do evento.
- ☐ A linha completa do arquivo.
- ☐ A quantidade de óbitos.
- ☒ A data do evento.

**Pergunta 13****1,67 / 1,67 pts**

No programa que você desenvolveu, qual foi a classe responsável por criar o diretório PastaSaida no HDFS?



☐ JobClient.☒ FileSystem.☐ JobConf.☐ MapIGTI.**Pergunta 14****1,67 / 1,67 pts**

Qual o método responsável por atribuir um nome para o job que vai ser executado pelo Hadoop/MapReduce?

☐ FileInputFormat.☐ setName.☒ setJobName.☐ setNameJob.**Pergunta 15****1,62 / 1,62 pts**

Qual o método responsável por criar o diretório PastaEntrada no HDFS?

☐ runJob.☐ setInputPaths.☐ setOutputPaths.☒ mkdirs.

Pontuação do teste: **25** de 25