

Trabalho Prático do Módulo 3

Entrega 13 jan em 21:00 **Pontos** 25 **Perguntas** 15
Disponível até 13 jan em 21:00 **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 de sexta-feira, **15/01/2021**, às 23h59.

Bons estudos!

2. O arquivo abaixo contém o enunciado do trabalho prático

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

Histórico de tentativas

| | Tentativa | Tempo | Pontuação |
|--------------|-----------------------------|-------------|-----------|
| MAIS RECENTE | Tentativa 1 | 443 minutos | 25 de 25 |

⚠ As respostas corretas estarão disponíveis em 15 jan em 23:59.

Pontuação deste teste: **25** de 25

Enviado 11 jan em 16:50

Esta tentativa levou 443 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?

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

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 ano 1989?

- ☐ 1989 Baseball|12.93
- ☐ 1989 Streets of Rage|2.41
- ☐ 1989 Streets of Rage|3.99
- ☒ 1989 Baseball|0.93

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 ano 2019?

- ☐ 2019 Gears of War 3|5.64
- ☐ 2019 Madden NFL 17|4.37
- ☒ Não existem dados para o ano de 2019.
- ☐ 2019 Madden NFL 17|1.42

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 ano em que houve o maior faturamento?

- ☐ 2014 com 6,01.
- ☐ 1980 com 4,26.
- ☐ 2002 com 4,22.
- ☒ 2011 com 5,64.

Pergunta 5

1,67 / 1,67 pts

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

- ☐ O mesmo que o número de linhas do arquivo ExemploIGTI.java.

- ☐ Uma vez para cada chamada ao método `copyFromLocalFile`.
- ☒ O mesmo que o número de linhas que o arquivo de entrada possui.
- ☐ 64 vezes.

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

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

- ☒ Uma vez para cada chave.
- ☐ Uma vez para cada execução da função `Map`.
- ☐ 7008 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?

- ☐ `/usr/local/hadoop/bin/hdfs dfs -ls /usr/igti/PastaSaida/part-00000`
- ☐ `/usr/local/hadoop/bin/hdfs dfs -rmr /usr/igti/PastaSaida/part-00000`
- ☒ `/usr/local/hadoop/bin/hdfs dfs -cat /usr/igti/PastaSaida/part-00000`
- ☐ `/usr/local/hadoop/bin/hdfs /usr/igti/PastaSaida/part-00000`

Pergunta 8**1,67 / 1,67 pts**

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

☐

1. Leitura de dados do HDFS, na pasta rotulada como *Dados*.
2. Processamento da função *Map*.
3. Processamento da função *Reduce* com os dados de saída da função *Map*.
4. Gravação dos dados no HDFS na pasta rotulada como *PastaEntrada*.

☐

1. Leitura de dados do HDFS, na pasta rotulada como *Entrada*.
2. Processamento da função *Map*.
3. Processamento da função *Combine* com os dados de saída da função *Map*.
4. Gravação dos dados no HDFS na pasta rotulada como *Saida*.

☒

1. Leitura de dados do HDFS, na pasta rotulada como *DADOS_GAME*.
2. Processamento da função *Map*.
3. Processamento da função *Reduce* com os dados de saída da função *Map*.
4. Gravação dos dados no HDFS na pasta rotulada como *PastaSaida*.

☐

1. Leitura de dados do sistema de arquivos do sistema operacional.
2. Processamento da função *Reduce*.
3. Processamento da função *Map* com os dados de saída da função *Reduce*.
4. 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 `dadosGame` no método `map` da classe `MapIGTI`. Quantas posições existem nesse array e o que é armazenado nas posições 0, 2, 3 e 4?

São 11 posições.

Posição 0: Nome do Jogo.

Posição 2: Ano de lançamento.

Posição 3: Fabricante.

☐ Posição 4: Vendas realizadas.

São 4 posições.

Posição 0: Identificador.

Posição 2: Ano de lançamento.

Posição 3: Dados de controle.

☐ Posição 4: Quantidade de itens vendidos.

São 8 posições.

Posição 0: Nome do Jogo.

Posição 2: Data de Lançamento.

Posição 3: Estilo do jogo.

☐ Posição 6: Fabricante.

São 11 posições.

Posição 0: Identificador.

Posição 2: Nome do Jogo.

Posição 3: Ano de Fabricação.

☒ Posição 4: Estilo/classificação do jogo.

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 uma barra vertical. A chave foi o ano de lançamento do jogo e o valor é representado por uma concatenação de dados, sendo: Nome do Jogo e vendas em todo o mundo.

☒

O programa gravou pares chave/valor, separados por um TAB. A chave foi o ano de lançamento do jogo e o valor é representado por uma concatenação de dados, sendo: Nome do Jogo e a soma das vendas do jogo realizadas nos continentes europeu e americano.

☐

O programa gravou pares chave/valor, separados por um arroba. A chave foi o nome do jogo e o valor é representado por uma concatenação de dados, sendo: Nome do Jogo e a soma das vendas do jogo realizadas no mundo todo.

☐

O programa gravou pares chave/valor, separados por um TAB. A chave foi ano de lançamento do jogo e o valor pelo fabricante do jogo.

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 ano 1997?

☐ 1997 Sonic the Hedgehog 3|1.49

☒ 1997 Half-Life|4.03

☐ 1997 SimCity 2000|2.04

☐ 1997 Tetris 2|0.75

Pergunta 12

1,67 / 1,67 pts

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

☐ O Fabricante do jogo.

☐ A linha completa do arquivo.

☒ O Ano de Lançamento do jogo.

☐ A quantidade de vendas.

Pergunta 13

1,67 / 1,67 pts

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

☒ FileSystem.

☐ JobClient.

☐ JobConf.

☐ MapIGTI.

Pergunta 14

1,67 / 1,67 pts

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

☐ setNameJob.

☐ FileInputFormat.

☒ setJobName.

☐ setName.

Pergunta 15

1,62 / 1,62 pts

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

☐ runJob.

☒ mkdirs.

☐ setInputPaths.

☐ setOutputPaths.

Pontuação do teste: **25** de 25