

# Desafio do Módulo 4

**Entrega** 23 out em 23:59

**Pontos** 40

**Perguntas** 15

**Disponível** até 23 out em 23:59

**Limite de tempo** Nenhum

**Tentativas permitidas** 2

## Instruções

O Desafio do Módulo 4 está disponível!

### 1. Instruções para realizar o desafio

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 Desafio do Módulo 2".

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 a partir de sexta-feira, **23/10/2022**, às 23h59.

Bons estudos!

### 2. O arquivo abaixo contém o enunciado do desafio

**Enunciado do Desafio - Módulo 4 - Engenheiro(a) de Dados Cloud.pdf**

Este teste foi indisponível 23 out em 23:59.

## Histórico de tentativas

	Tentativa	Tempo	Pontuação
MANTIDO	<u>Tentativa 2</u>	7 minutos	40 de 40
MAIS RECENTE	<u>Tentativa 2</u>	7 minutos	40 de 40
	<u>Tentativa 1</u>	3.855 minutos	26,7 de 40

Pontuação desta tentativa: **40** de 40

Enviado 13 out em 13:13

Esta tentativa levou 7 minutos.

### Pergunta 1

2,66 / 2,66 pts

São tecnologias utilizadas para Event Streaming, EXCETO:

☐ Amazon Kinesis Streams

☐ Azure EventHub

☒ Google Dataproc

☐ Amazon Kinesis Firehose

Correto!

### Pergunta 2

2,66 / 2,66 pts

Qual das ferramentas abaixo podemos utilizar para testar o Kubernetes em ambiente local?

☐ Amazon ECS

☒ Minikube

☐ Google Cloud PubSub

☐ Amazon Kinesis Firehose

Correto!

### Pergunta 3

2,66 / 2,66 pts

A *Confluent* é a maior impulsionadora de desenvolvimento de qual ferramenta?

**Correto!**

- ☐ Apache Airflow
- ☐ Apache Atlas
- ☒ Apache Kafka
- ☐ Apache Spark

**Pergunta 4****2,66 / 2,66 pts**

O que o Apache Pinot NÃO é?

**Correto!**

- ☐ Uma base de dados analítica para camada OLAP em Real Time
- ☒ Uma database relacional
- ☐ Uma database não relacional
- ☐ Um Data Warehouse

**Pergunta 5****2,66 / 2,66 pts**

Qual serviço Kafka pode ser utilizado para se conectar a uma base de dados?

**Correto!**

- ☐ Zookeeper
- ☐ Kafka Topic
- ☐ Kafka Sink Connector
- ☒ Kafka Connect

**Pergunta 6****2,66 / 2,66 pts**

Até quanto tempo o *log retention* do Kafka pode ser aumentado?

- ☐ 7 dias
- ☒ Infinito, dependendo apenas do Hardware disponível
- ☐ 7 anos
- ☐ 7 meses

**Correto!****Pergunta 7****2,66 / 2,66 pts**

Qual das frases abaixo melhor define o Spark Structured Streaming?

- ☐ Nenhuma das alternativas.
- ☒ Uma extensão do core do Apache Spark, para ler, processar e disponibilizar eventos em tempo real.
- ☐ Capacidade de ler e disponibilizar eventos em tempo real.
- ☐ Capacidade de criar tópicos no Kafka.

**Correto!****Pergunta 8****2,66 / 2,66 pts**

Qual das frases abaixo melhor define um *streaming* de dados?

**Correto!**

- ☐ Fluxos de ingestão de dados agendados para execução em horários específicos.
- ☐ Fluxos de dados extraídos de bases relacionais.
- ☒ Fluxo de dados contínuo e sem fim.
- ☐ Fluxos criados em ferramentas de orquestração, como o Apache Airflow.

**Pergunta 9****2,66 / 2,66 pts**

Qual desses métodos NÃO é viável para fazer o deploy de um cluster Kafka num ambiente Kubernetes?

**Correto!**

- ☐ Nenhuma das alternativas.
- ☒ docker-compose
- ☐ Manifestos .yaml
- ☐ Helm Chart

**Pergunta 10****2,66 / 2,66 pts**

Dentre as opções, qual delas NÃO é uma possibilidade de processamento em real time?

- ☐ Processar eventos com KSQLDB lendo de tópicos e escrevendo dados processados em outros tópicos.

**Correto!**☐

Processar eventos com Spark Structured Streaming lendo de tópicos Kafka e escrevendo dados processados no Amazon S3.

☐

Processar eventos com Spark Structured Streaming lendo arquivos da camada landing do S3 e escrevendo em outra camada posterior.

☒

Processar eventos com KSQLDB lendo de tópicos e escrevendo dados processados no S3 direto do KSQLDB.

**Pergunta 11****2,66 / 2,66 pts**

Qual é o método de processamento utilizado pelo Spark Structured Streaming?

☐

Streaming e processamento de cada evento isoladamente

☐

Micro leituras de tópicos Kafka

☒

Micro Batch

☐

Nenhuma das alternativas.

**Correto!****Pergunta 12****2,66 / 2,66 pts**

Qual é a ferramenta responsável por ajudar o KSQLDB entender os schemas das tabelas ingeridas por Kafka Connect?

☐

Topic Partitions

☐

Source Connector

**Correto!**☒ Schema Registry☐ Zookeeper**Pergunta 13****2,66 / 2,66 pts**

Com relação à escrita de dados do Spark Structured Streaming, qual das opções abaixo pode ser considerada INCORRETA?

☐ Nenhuma das alternativas.☒ Escrita de streaming com agregação e outputMode “append”☐ Escrita de streaming com agregação e outputMode “complete”☐ Escrita de streaming sem agregação e outputMode “append”**Correto!****Pergunta 14****2,66 / 2,66 pts**

Qual das opções abaixo é a MENOS indicada para camada de consumo de dados em Real Time?

☐ Pinot☐ DynamoDB☐ Elastic Search☒ PostgreSQL**Correto!****Pergunta 15****2,76 / 2,76 pts**

Qual das opções abaixo melhor define o método “JBOD” de deployment do Kafka no Kubernetes?



Consiste na reserva de memória em cluster para melhorar a performance do processamento do Kafka



Consiste na criação de PVCs para persistir os dados do Kafka em disco.



Nenhuma das alternativas.



Consiste na criação de LoadBalancers para distribuir a entrada de mensagens nas distintas partições dos tópicos Kafka

**Correto!**

Pontuação do teste: **40** de 40