

Desafio do Módulo 4

Entrega 20 jul em 21:00 **Pontos** 40 **Perguntas** 15
Disponível até 20 jul em 21:00 **Limite de tempo** Nenhum

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".

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, **22/07/2022**, às 23h59.

Bons estudos!

2. O arquivo abaixo contém o enunciado do desafio

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

Este teste foi indisponível 20 jul em 21:00.

Histórico de tentativas

	Tentativa	Tempo	Pontuação
MAIS RECENTE	<u>Tentativa 1</u>	6.418 minutos	40 de 40

Pontuação deste teste: **40** de 40

Enviado 18 jul em 21:13

Esta tentativa levou 6.418 minutos.

Pergunta 1**2,67 / 2,67 pts**

O desenho arquitetural da solução implementada corresponde a uma solução:

Correto!

- ☒ Desacoplada.
- ☐ Dependente de intervenção humana.
- ☐ Monolítica.
- ☐ Nenhuma das alternativas está correta.

Pergunta 2**2,67 / 2,67 pts**

Com relação à etapa de extração, ela deve vir:

Correto!

- ☐ Após o primeiro passo da transformação.
- ☒ Antes do primeiro passo de transformação.
- ☐ Ao mesmo tempo que o primeiro passo de transformação.
- ☐ Todas as alternativas contêm soluções possíveis.

Pergunta 3**2,67 / 2,67 pts**

Com relação às etapas de transformação, elas devem ser programadas:

- ☐ Sequenciais.
- ☐ Com o máximo de integração possível visando consistência.

Correto!

- ☐ Nenhuma das alternativas está correta.
- ☒ Com o máximo de desacoplamento possível.

Pergunta 4**2,67 / 2,67 pts**

Caso você tenha escolhido implementar o seu scheduler com Airflow, o modo de definir o scheduler das soluções consiste em:

Correto!

- ☐ Definir intervalos com “timedelta”.
- ☐ Programar expressões cron.
- ☒ Todas as alternativas estão corretas.
- ☐ Palavras- chave estabelecendo a regularidade, como “@once”, “@hourly”, “@daily”.

Pergunta 5**2,67 / 2,67 pts**

Para executar uma pipeline utilizando a nuvem do Prefect, NÃO é necessário:

Correto!

- ☐ Utilizar um runner token para execução do flow.
- ☒ Utilizar secrets para autenticação em serviços externos.
- ☐ Utilizar um token pessoal de login.
- ☐ Executar um agente para a query dos flows a serem executados pela nuvem do Prefect.

Pergunta 6**2,67 / 2,67 pts**

Pensando na melhor prática de soluções de orquestração de Jobs, é bom que a solução:

- ☐ Possa se recuperar contra falhas.
- ☒ Todas as alternativas estão corretas.
- ☐ Seja escalável.
- ☐ Seja autônoma, independente de intervenção humana.

Correto!**Pergunta 7****2,67 / 2,67 pts**

Pipelines de Airflow e Prefect são definidas com:

- ☐ Yaml.
- ☐ Json.
- ☐ SQL.
- ☒ Python.

Correto!**Pergunta 8****2,67 / 2,67 pts**

Qual destas NÃO é uma opção para execução de pipelines com Airflow e Prefect?

- ☐ Servidor virtualizado.

Correto!

- ☐ Kubernetes.
- ☐ Deploy local em minha máquina pessoal.
- ☒ API rest.

Pergunta 9**2,67 / 2,67 pts**

Com relação ao licenciamento para utilização do Airflow e do Prefect:

Correto!

- ☐ Prefect é open source, Airflow é proprietário.
- ☐ Airflow é proprietário, Prefect é open source.
- ☒ Ambos são open source.
- ☐ Ambos são proprietários pagos.

Pergunta 10**2,67 / 2,67 pts**

A API Stream do Twitter retorna que tipo de dados?

Correto!

- ☒ JSON.
- ☐ CSV.
- ☐ XML.
- ☐ HTML.

Pergunta 11**2,67 / 2,67 pts**

A estrutura do dado retornado pela API Stream do Twitter corresponde a qual sistema de bancos de dados?

☐ Graph.

Correto!

☒ NoSQL.

☐ SQL.

☐ Cache.

Pergunta 12

2,67 / 2,67 pts

A classe do “Listener” que implementamos para acessar a API do Twitter (pacote tweepy) precisa ter quais métodos sobrescritos? Assinale a alternativa CORRETA.

I. on_data

II. on_error

III. __init__

IV. __main__

Correto!

☒ Os métodos I e II.

☐ Os métodos I e III.

☐ Os métodos III e IV.

☐ Os métodos II e IV.

Pergunta 13

2,67 / 2,67 pts

Quantas chaves são necessárias para autenticar na API Stream do Twitter?

Correto!

☒ 4.

☐ 3.

☐ 5.

☐ 2.

Pergunta 14

2,67 / 2,67 pts

As etapas do processo de ETL devem:

Correto!

☒

Estar o mais desacopladas quanto possível, possibilitando paralelismos.

☐

Ser interdependentes, garantindo que etapas subsequentes tenham acesso a dados processados em etapas anteriores.

☐

Necessariamente trabalhar em containers, pois esta é a única possibilidade de instância de computação disponível.

☐

Ser tão acopladas quanto possível, no intuito de garantir a consistência do fluxo.

Pergunta 15

2,62 / 2,62 pts

A capacidade das soluções de orquestração de pipelines de escalar vem, sobretudo, de:

Correto!



Sua implementação de um sistema de fila para gerenciar workers de processamento.



A utilização de um SGBD para controle de estados.



Sua implementação usando linguagens de programação de baixo nível.



Sua capacidade de ser facilmente implantável em ambiente de nuvem.

Pontuação do teste: **40** de 40