

Desafio do Módulo 4

Entrega 29 set em 21:00**Pontos** 40**Perguntas** 15**Disponível** até 29 set 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, **01/10/2021**, às 23h59.

Bons estudos!

2. O arquivo abaixo contém o enunciado do desafio

[Enunciado do Desafio - Módulo 4 - Bootcamp Engenheiro\(a\) de Dados.pdf](#)

Histórico de tentativas

| | Tentativa | Tempo | Pontuação |
|---------------------|-----------------------------|---------------|-----------|
| MAIS RECENTE | Tentativa 1 | 2.329 minutos | 40 de 40 |

⚠ As respostas corretas estarão disponíveis em 1 out em 23:59.

Pontuação deste teste: **40** de 40

Enviado 29 set em 9:46

Esta tentativa levou 2.329 minutos.

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

O desenho arquitetural da solução implementada corresponde a uma solução:

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

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

Com relação à etapa de extração, ela deve vir:

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

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

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

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

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

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:

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

Pergunta 5

2,67 / 2,67 pts

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

- ☐ Utilizar um token pessoal de login.
- ☒ Utilizar secrets para autenticação em serviços externos.
- ☐ Utilizar um runner token para execução do flow.
- ☐ 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 autônoma, independente de intervenção humana.
- ☐ Seja escalável.

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

Pipelines de Airflow e Prefect são definidas com:

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

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

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

- ☐ Kubernetes.

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

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

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

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

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

A API Stream do Twitter retorna que tipo de dados?

- ☐ XML.
- ☐ HTML.
- ☐ CSV.
- ☒ JSON.

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.

☐ SQL.

☐ Cache.

☒ NoSQL.

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__

☐ Os métodos I e III.

☒ Os métodos I e II.

☐ Os métodos II e IV.

☐ Os métodos III e IV.

Pergunta 13

2,67 / 2,67 pts

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

☐ 5.

☒ 4.

☐ 2.

☐ 3.

Pergunta 14

2,67 / 2,67 pts

As etapas do processo de ETL devem:

☐

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.

☒

Estar o mais desacopladas quanto possível, possibilitando paralelismos.

Pergunta 15

2,62 / 2,62 pts

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



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



Sua implementação usando linguagens de programação de baixo nível.



A utilização de um SGBD para controle de estados.



Sua capacidade de ser facilmente implantável em ambiente de nuvem.

Pontuação do teste: **40** de 40