

Desafio do Módulo 1

Iniciado: 23 nov em 15:43

Instruções do teste

O Desafio do Módulo 1 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 1".

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, **04/12/2020**, às 23h59.

Bons estudos!

2. O arquivo abaixo contém o enunciado do desafio

[Enunciado do Desafio - Módulo 1 - Bootcamp Engenheiro de dados.pdf](#)

Pergunta 1

2,67 pts

O desenho arquitetural da solução implementada corresponde a uma solução:

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

Pergunta 2

2,67 pts

Com relação à etapa de extração, ela deve vir:

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

Pergunta 3

2,67 pts

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

- ☐ Com o máximo de integração possível visando consistência.
- ☐ Nenhuma das alternativas está correta.
- ☐ Com o máximo de desacoplamento possível.
- ☐ Sequenciais.

Pergunta 4

2,67 pts

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

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

Pergunta 5

2,67 pts

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

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

Pergunta 6

2,67 pts

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

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

Pergunta 7

2,67 pts

Pipelines de Airflow e Prefect são definidas com:

- ☐ Python.
- ☐ SQL.

☐ Json.

☐ Yaml.

Pergunta 8

2,67 pts

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

☐ API rest.

☐ Kubernetes.

☐ Servidor virtualizado.

☐ Deploy local em minha máquina pessoal.

Pergunta 9

2,67 pts

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

☐ Airflow é proprietário, Prefect é open source.

☐ Ambos são proprietários pagos.

☐ Ambos são open source.

☐ Prefect é open source, Airflow é proprietário.

Pergunta 10

2,67 pts

A API Stream do twitter retorna que tipo de dados?

☐ XML.

- ☐ CSV.
- ☐ JSON.
- ☐ HTML.

Pergunta 11

2,67 pts

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

- ☐ Cache.
- ☐ Graph.
- ☐ NoSQL.
- ☐ SQL.

Pergunta 12

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 III e IV.
- ☐ Os métodos I e III.
- ☐ Os métodos II e IV.
- ☐ Os métodos I e II.

Pergunta 13**2,67 pts**

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

☐ 2.

☐ 5.

☐ 3.

☐ 4.

Pergunta 14**2,67 pts**

As etapas do processo de ETL devem:

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

☐ 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.

Pergunta 15**2,62 pts**

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

☐ A utilização de um SGBD para controle de estados.

☐ Sua capacidade de ser facilmente implantável em ambiente de nuvem.

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

☐ Sua implementação usando linguagens de programação de baixo nível.

Salvo em 15:50

Enviar teste