**Questões de Múltipla Escolha**

1. Um cientista de dados está trabalhando em um projeto de análise de sentimentos de avaliações de clientes. Qual das seguintes tarefas **NÃO** se beneficiaria diretamente do uso da biblioteca Pandas?
   * a) Carregar e limpar um grande conjunto de dados de avaliações de um arquivo CSV.
   * b) Calcular a frequência de palavras específicas nas avaliações.
   * c) Criar visualizações complexas de dados, como mapas de calor de correlação entre palavras.
   * d) Converter as avaliações de texto para representações numéricas usando técnicas de processamento de linguagem natural.
   * e) Filtrar avaliações com base em critérios específicos, como classificação do cliente ou data da avaliação.
2. Em um projeto de Aprendizado de Máquina para prever preços de casas, qual das seguintes etapas envolveria o uso de Pandas para **manipulação e preparação dos dados**?
   * a) Escolher o algoritmo de Aprendizado de Máquina a ser utilizado.
   * b) Dividir o conjunto de dados em conjuntos de treinamento e teste.
   * c) Tratar valores ausentes nas colunas de área, número de quartos e número de banheiros.
   * d) Treinar o modelo de regressão linear.
   * e) Avaliar o desempenho do modelo usando métricas como erro quadrático médio.
3. Considere o seguinte código Python que utiliza Pandas:

import pandas as pd

df = pd.DataFrame({'A': [1, 2, 2, 3, 3, 3],

'B': ['a', 'b', 'b', 'c', 'c', 'c']})

df\_result = df.drop\_duplicates()

print(df\_result)

Qual será a saída do código?

* + a)
  + A B
  + 0 1 a
  + 1 2 b
  + 2 2 b
  + 3 3 c
  + 4 3 c
  + 5 3 c
  + b)
  + A B
  + 0 1 a
  + 1 2 b
  + 3 3 c
  + c)
  + A B
  + 0 1 a
  + 1 2 b
  + 3 3 c
  + 4 3 c
  + d)
  + A B
  + 0 1 a
  + 2 2 b
  + 5 3 c
  + e) O código resultará em um erro.

1. Um desenvolvedor está trabalhando em um sistema de recomendação de filmes. Qual das seguintes tarefas **envolveria conceitos de Sets em Python**?
   * a) Armazenar a lista de todos os filmes disponíveis.
   * b) Armazenar os gêneros de cada filme.
   * c) Encontrar a interseção de filmes assistidos por dois usuários para identificar recomendações em comum.
   * d) Ordenar os filmes por classificação.
   * e) Calcular a média das classificações dos filmes.
2. Um analista de dados precisa ler um arquivo JSON que contém informações sobre vendas de produtos, incluindo detalhes do produto, quantidade vendida e data da venda. Qual método do Pandas é mais adequado para essa tarefa e como ele deve ser usado para acessar as quantidades vendidas?
   * a) read\_csv(); Acessar as quantidades usando df['Quantidade Vendida'].
   * b) read\_excel(); Acessar as quantidades usando df.Quantidade Vendida.
   * c) read\_json(); Acessar as quantidades usando df['Quantidade']['VendidoEm'].
   * d) read\_json(); Acessar as quantidades usando df['Quantidade Vendida'].
   * e) read\_table(); Acessar as quantidades usando df.loc['Quantidade Vendida'].
3. Em um script Python, você precisa abrir um arquivo para adicionar logs de eventos. O arquivo já existe . Qual modo de abertura de arquivo você utilizaria?
   * a) "r"
   * b) "w"
   * c) "a"
   * d) "x"
   * e) "r+"
4. Qual das seguintes expressões lógicas em Python avalia para True se a variável x for maior que 10 e menor ou igual a 20?
   * a) x > 10 or x <= 20
   * b) x > 10 and x < 20
   * c) x >= 10 and x <= 20
   * d) x > 10 and x <= 20
   * e) x in range(11, 21)
5. Considere o seguinte código Python:

lista = [10, 20, 30, 40, 50]

for i in range(len(lista)):

if lista[i] > 25:

print(lista[i])

Qual será a saída do código?

* + a) 10 20 30 40 50
  + b) 30 40 50
  + c) 10 20
  + d) 20 30 40
  + e) 40 50

1. Um programador precisa iterar sobre uma string chamada mensagem e realizar uma ação específica para cada caractere que não seja um espaço em branco. Qual das seguintes abordagens é mais eficiente e Pythonica?
   * a)

for i in range(len(mensagem)):

if mensagem[i] != ' ':

# Ação

* + b)

i = 0

while i < len(mensagem):

if mensagem[i] != ' ':

# Ação

i += 1

* + c)

for caractere in mensagem:

if caractere != ' ':

# Ação

* + d)

for i, caractere in enumerate(mensagem):

if caractere != ' ':

# Ação

* + e)

for caractere in mensagem.split():

# Ação

1. Considere o seguinte código Python:

frase = "Python é uma linguagem poderosa"

palavras = frase.split()

nova\_frase = " ".join(palavras[::-1])

print(nova\_frase)

Qual será a saída do código?

* + a) Python é uma linguagem poderosa
  + b) poderosa linguagem uma é Python
  + c) poderosa linguagem uma é
  + d) linguagem poderosa uma é Python
  + e) onhtyP é amu megaugnil asodrep

1. Em um projeto de Aprendizado de Máquina, um engenheiro de dados precisa converter uma coluna de um DataFrame Pandas que contém datas em formato de string para o tipo datetime. Qual função do Pandas é mais adequada para essa conversão?
   * a) to\_string()
   * b) astype()
   * c) to\_datetime()
   * d) convert\_dates()
   * e) parse\_dates()
2. Complete o código :(UTILIZE APENAS 1 dígito no seu trecho do CODIGO)

minha\_string = "Python é poderoso"

Acessar um único caractere primeiro\_caractere = print(minha\_string[**<CODIGO01>**]) # Saída: 'P' ultimo\_caractere = print(minha\_string[**<CODIGO02>**]) # Saída: 'o'

# Acessar uma substring

print(minha\_string[**<CODIGO03>**]) # Saída: 'é poderoso'

print(minha\_string[**<CODIGO04>**]) # Saída: 'Python'

1. Qual é a saída do código?

import pandas as pd

data = {'Nome': ['Alice', 'Bob', 'Charlie'],

'Idade': [25, 30, 28],

'Cidade': ['São Paulo', 'Rio de Janeiro', 'Belo Horizonte']}

df = pd.DataFrame(data)

print(df)

1. Qual é a saída do código?

import pandas as pd

# Criando um arquivo CSV de exemplo (você pode substituir pelo seu arquivo)

data = {'Nome': ['Alice', 'Bob', 'Charlie', 'David', 'Emily', 'Fernando'],

'Idade': [25, 30, 28, 22, 35, 29],

'Cidade': ['São Paulo', 'Rio de Janeiro', 'Belo Horizonte', 'Salvador', 'Brasília', 'Curitiba']}

df = pd.DataFrame(data)

df.to\_csv('clientes.csv', index=False)

# Ler o arquivo CSV para um DataFrame

df\_clientes = pd.read\_csv('clientes.csv')

# Imprimir as 5 primeiras linhas

print(df\_clientes.head())

# Calcular a média da idade

media\_idade = df\_clientes['Idade'].mean()

print(f"\nMédia da idade dos clientes: {media\_idade}")

1. Qual é a saída do código?

import pandas as pd

data = {'A': [1, 2, 3], 'B': [4, 5, 6]}

df = pd.DataFrame(data)

print(df.loc[1]['B'])

1. Qual é a saída do código?

strFrase = "Python eh top"

print(strFrase[1:4:2])

1. O que o seguinte trecho de código faz (Descreva com suas palavras):

import pandas as pd

df = pd.read\_csv('dados3.csv')

df.dropna(inplace=True)

print(len(df))

1. Considerando que o arquivo dados3.csv tem, ao menos, 10 registros (células) sem valores, o que podemos afirmar sobre os dois prints?

import pandas as pd

df = pd.read\_csv('dados3.csv')

print(len(df))

df.dropna(inplace=True)

print(len(df))

1. Qual é uma das saídas possíveis do código abaixo?

lista = [1, 2, 2, 3, 4, 4, 5]

print(set(lista))

1. Qual é a saída do seguinte código:

import pandas as pd

df = pd.DataFrame({'col1': [1, None, 3], 'col2': [None, 5, 6]}) df['col1'].fillna(0, inplace=True)

print(df['col1'].tolist())