

Protocolo de Finalização Nº 000083545308

O aluno RAPHAEL FERNANDO REIS DA SILVA com RA **25123537-4** finalizou a atividade **PROVA - PYTHON AVANÇADO - 41_2025** em **18/01/2025 00:36:39**

PROVA - PYTHON AVANÇADO - 41_2025

Período:02/12/2024 08:00 a 18/01/2025 23:59 (Horário de Brasília)

Status:ABERTO

Nota máxima:6,00

Nota obtida:

1ª QUESTÃO

Uma Series é um objeto do tipo _____ contendo uma sequência de valores (de tipos semelhantes aos tipos do _____) e um array associado de rótulos (labels) de dados, chamado de _____. A Series mais simples é composta de apenas um _____ de dados

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Diante do exposto, na sequência assinale a alternativa que preencha corretamente as lacunas.

RESPOSTA: NumPy, array unidimensional, índice, array.

2ª QUESTÃO

Os dados contínuos geralmente são discretizados ou separados em caixas para análise. O Objeto pandas pode ser tratado como um array de strings indicando o nome do compartimento; internamente, ele contém uma array de categorias especificando os nomes das categorias distintas, juntamente com uma rotulagem para os dados.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

O principal método utilizado para contagem de compartimentos em Pandas é:

Assinale a alternativa correta:

RESPOSTA: value_counts(): Retorna uma série contendo contagens de linhas exclusivas no DataFrame.

3ª QUESTÃO

Os objetos Pandas são equipados com um conjunto de métodos matemáticos e estatísticos comuns. A maioria deles se enquadra na categoria de reduções ou estatísticas de resumo, métodos que extraem um único valor (como a soma ou a média) de uma série ou uma série de valores das linhas ou colunas de um DataFrame. Em comparação com os métodos semelhantes encontrados em matrizes NumPy, eles têm tratamento interno para dados ausentes.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Sobre as funcionalidades estatísticas do Pandas, assinale a alternativa correta sobre a descrição do método describe().

RESPOSTA: Este método imprime dados que resumem a tendência central, dispersão e forma da distribuição de um conjunto de dados, excluindo os valores de não nulos.

4ª QUESTÃO

Se você escrever em recursividade infinita por engano, confira se a sua função tem um caso-base que não faz uma chamada recursiva.

DOWNEY, Allen B.; EM PYTHON, Pense como um cientista da computação. São Paulo: Novatec, 2016.

Assim, com base nesta informação, analise as asserções abaixo:

I. Aplicar um caso-base no início de uma função recursiva faz com que essa função nunca seja infinita.

PORQUE

II. Dado que em algum momento o caso base será atingido antes de executar a próxima iteração da recursão, dessa forma desempilhando a chamada recursiva.

A respeito dessas asserções, assinale a opção correta.

RESPOSTA: As asserções I e II são proposições verdadeiras e a II é uma justificativa correta da I.

5ª QUESTÃO

A mecânica fundamental de interação com os dados contidos em uma Série ou DataFrame é apresentada pelas funcionalidades essenciais. Estas funcionalidades apresentam os recursos mais básicos da biblioteca.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Sobre as funcionalidades básicas, assinale a alternativa correta.

I. reindexação: aplicando o método drop ele devolverá um novo objeto como valor ou os valores indicados apagados de um eixo.

II. descartar entradas de um eixo: implica em criar um novo objeto com os dados de acordo com um novo índice.

III. seleção com iloc e loc, permitem selecionar um subconjunto de linhas e colunas de um DataFrame com uma notação semelhante a do NumPy.

As afirmações I, II e III são, respectivamente:

RESPOSTA: V, F, V.

6ª QUESTÃO

Cada índice tem vários métodos e propriedades para lógica de conjunto, que respondem a outras perguntas comuns sobre os dados que ele contém.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Avalie as afirmações a seguir como (V) para verdadeiras e (F) para falsas, sobre as propriedades dos índices de uma Dataframe .

I. difference: Calcular a diferença do conjunto como um índice.

II. isin: Compute boolean array indicando se cada valor está contido na coleção passada.

III. drop: Calcula novo índice com elemento no índice que foi excluído.

IV. unique: Calcula o novo índice inserindo o elemento no índice i.

As afirmações I, II, III e IV são, respectivamente:

RESPOSTA: V, V, F, F.**7ª QUESTÃO**

A biblioteca Matplotlib tem como objetivo gerar visualizações informativas, chamadas de plotagens. Este passo é uma das tarefas mais importantes em análise de dados. Essa tarefa faz parte do processo exploratório, para ajudar a identificar valores discrepantes (outliers) ou transformações necessárias nos dados, ou como uma forma de gerar ideias modelos.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Sobre a plotagem de gráficos com pandas e Matplotlib, assinale a alternativa correta.

RESPOSTA: por padrão da o método plot() do DataFrame e Series invoca a biblioteca Matplotlib exibindo gráfico de linhas.**8ª QUESTÃO**

Iterável: qualquer objeto do qual a função interna do iter pode obter um iterador. Objetos que implementam um método `__iter__` retornando um iterador são iteráveis. As sequências são sempre iteráveis; como são objetos que implementam um método `__getitem__` que usa índices baseados em 0.

RAMALHO, Luciano. Fluent Python 2nd ed. 2015.

Sobre funções objetos iteráveis em Python, avalie as afirmações abaixo:

I. Dados que não cabem na memória, precisam ser buscados em modo lazy (um de cada vez e sob demanda), para isso é usado os objetos iteradores

II. Uma string em Python é um objeto iterável.

III. A iteração é fundamental para o processamento de dados.

É correto o que se afirma em:

RESPOSTA: I, II, III.**9ª QUESTÃO**

Às vezes, a maneira como os dados são armazenados em arquivos ou bancos de dados não está no formato correto para uma determinada tarefa. Muitos pesquisadores optam por fazer o processamento ad hoc de dados de um formulário para outro usando uma linguagem de programação de uso geral, como Python, Perl, R ou Java, ou ferramentas de processamento de texto Unix como sed ou awk.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Sobre as funções e métodos do Pandas, assinale a alternativa correta sobre a descrição do método `head()`.

RESPOSTA: Esta função retorna as primeiras n linhas do objeto com base na posição. É útil para testar rapidamente se o seu objeto contém o tipo certo de dados.**10ª QUESTÃO**

Durante a análise e modelagem de dados, uma quantidade significativa de tempo é gasta na preparação de dados: carregamento, limpeza, transformação e reorganização. Essas tarefas costumam ocupar 80% ou mais do tempo de um analista, para isso o Pandas oferece muitas funções para compreensão dos DataFrames.

MCKINNEY, Wes. Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2019.

Sobre as funções e métodos do Pandas, assinale a alternativa correta sobre a descrição do método `info()`.

RESPOSTA:

Este método imprime informações sobre um DataFrame, incluindo o índice `dtype` e colunas, valores não nulos e uso de memória.