Quiz 2 2024 May

1. 选择题

Q1: 建 NumPy 数组的正确语法是什么？

A.np.creatArray([1,2,3,4,5])

B.np.array([1,2,3,4,5])

C.np.object([1,2,3,4,5])

Q2: 检查数组中维数的正确语法是什么？

A.np.ndim()

B.np.dim

C.np.dim()

D.np.ndim

Q3: 哪一句是打印数组第一项？

A.print(myArr,1)

B.print(myArr[1])

C.print(myArr[0])

Q4: 从下面数组中打印数字8的正确语句是哪一句？

Arr=np.arry([[1,2,3,4,5],[6,7,8,9,10]])

A.print(Arr[3,0])

B.print(Arr[1,2])

C.print(Arr[7,2])

Q5: 创建浮点类型数组的正确语法是什么？

A.arr=np.float([1,2,3,4])

B.arr=np.array([1,2,3,4],dtype=’f)

C.arr=np.float([1,2,3,4]).tpFloat()

Q6: 在数组中搜索某个值的正确方法是？

A.search()

B.find()

C.where()

Q7: 从数组中返回所有值为 4 的项目的索引的正确语法是什么

A.np.where(arr==4)

B.np.where(arr=4)

C.arr.search(4)

Q8: 使用 NumPy 随机模块时，如何返回 0 到 100 的随机整数？

A.random.rand()

B.random.rand(100)

C.random.randint(100)

Q9:将标签“x”、“y”和“z”添加到 Pandas 系列的正确语法是什么？

A.pd.Series(mylist, index=[“x”,“y”,“z”])

B.pd.Series(mylist, lables=[“x”,“y”,“z”])

C.pd.Series(mylist.names=[“x”,“y”,“z”])

Q10:读取 Pandas DataFrame 中第一行的正确语法是什么

A.df.get(0)

B.df.loc[0]

C.df[0]

Q11: P将 CSV 文件加载到 DataFrame 中的正确 Pandas 函数是什么

A.read\_file()

B.read\_csv()

C.ReadCSV()

D.readFile()

Q12:对于 Pandas 的 head() 方法，如果不指定，默认返回多少行

A.5

B.10

C.20

Q13:什么是删除包含空单元格的行的正确 Pandas 方法？

A.remove\_null()

B.dropna()

C.delete\_null

Q14:Which of the following is true about user- defined functions in Python?

A.Every function should have return statement at the end

B.The name of the functions defined by the user should follow some guidelines

C.The body of the function is indented and is enclosed by the brackets

D.None of the above

Q15: How can you calculate the mean along a specific axis of a 2D NumPy array arr?

1. mean\_values = np.mean(arr, axis=1)
2. mean\_values = arr.mean(axis=0)
3. mean\_values = np.average(arr, axis=1)
4. mean\_values = arr.mean(axis=2)

Q16: Which of the following is code gives an error?  A. a = np.array([(1,2,3),(4,5,6)]) a[(0,1)] B. a = np.array([(1,2,3),(4,5,6)]) a.reshape(2,4) C. a = np.array([(1,2,3),(4,5,6)]) a[np.arange(1), :] D. All the above

Q17: What is a correct syntax to create a NumPy array?

1. np.array([1, 2, 3, 4, 5])

B. np.createArray([1, 2, 3, 4, 5])C. np.object([1, 2, 3, 4, 5])

Q18: Which of the following arrays is a two dimensional (2-D) array?

1. np.array([1, 2, 3, 4, 5])
2. np.array( [[1, 2, 3], [4, 5, 6]])
3. np.array(42)

Q19: What is a correct syntax to check the number of dimensions in an array?

1. arr.dim()
2. arr.ndim
3. arr.dim
4. arr.ndim()

Q20: What is a correct syntax to print the first item of an array? C

print(myArr,1)print(myArr[1])print(myArr[0])

Q21: What is a correct syntax to print the number 8 from the array below: arr = np.array([[1,2,3,4,5], [6,7,8,9,10]]) A

1. print(arr[1, 2])
2. print(arr[3, 0])
3. print(arr[7, 2])

Q22: What is a correct syntax to print the numbers [3, 4, 5] from the array below: arr = np.array([1,2,3,4,5,6,7])C

1. print(arr[2:4])
2. print(arr[2:6])
3. print(arr[2:5])
4. print(arr[3:6])

Q23: What is the correct Pandas function for loading CSV files into a DataFrame?

1. pandas.ReadFile()
2. pandas.read\_file()
3. pandas.ReadCSV()
4. pandas.read\_csv()

Q24: How can you select a specific column from a DataFrame?

df.select\_column("ColumnName")df.get\_column("ColumnName")df["ColumnName"]df.column("ColumnName")

Q25: What does the groupby() function in Pandas allow you to do?

1. Group DataFrame columns
2. Group DataFrame rows based on a condition
3. Group data based on one or more columns
4. Group data based on the index

Q26: How can you handle missing values in a Pandas DataFrame?

1. Use df.fillna(value) to replace missing values
2. Use df.dropna() to remove rows with missing values
3. Both A and B
4. Neither A nor B

Q27. Given a dataset named ‘data’ containing the 5 columns and 10 rows, find the output of the below code?print(len(data.columns))

1. 5
2. 10
3. 15
4. 50

Q28: In Pandas \_\_\_\_\_\_\_\_\_ is used to store data in multiple columns.A. SeriesB. DataFrameC. Both of the aboveD. None of the above

Q29: Assume D1 is a dataframe, D1[ : ] = 77 will set \_\_\_\_\_\_\_\_ values of a Data Frame ‘D1’ to 77.A. Only First RowB. Only First ColumnC. AllD. None of the above

Q30:  In given code dataframe ‘D1’ has \_\_\_\_\_ rows and \_\_\_\_\_ columns.import pandas as pdS1=pd.Series([1, 2, 3, 4], index = ['a', 'b','c','d'])S2=pd.Series([11, 22, 33, 44], index = ['a', 'bb','c','dd'])D1 = pd.DataFrame([S1,S2])

1. 2, 4

B. 4, 6C. 4, 4D. 2, 6

Q31:In DataFrame, axis 0 is for \_\_\_\_\_

A. RowsB. ColumnsC. Rows and Columns BothD. None of these

Q32: Suppose a data frame Df1 contains information about student having columns rollno, name, class and section. Write the code for the following:Add one more column as fee: A. Df1[‘fee’]=([100,200,300])B. Df1.add[‘fee']= ([100,200,300])C. Df1.AddNewCol[‘fee’]= [100,200,300])D. Df1.addnewcol[‘fee']= ([100,200,300])

Q33: Which of the following is the correct syntax to select or access columns from the dataframe using column names? A. df(col1,col2,…,coln)B. df[[col1,col2,…,coln]]C. df[col1,col2,…,coln]D df{col1:col2:…,:coln}

Q34: Replace the column label from 2016 to 2020 in DataFrame ‘DF’ A. DF.rename({2016 : 2020}, axis = ‘columns’)B. DF.rename({2016 : 2020}, axis = ‘index’)C. DF.rename({2016 : 2020}, axis = ‘column’)D. DF.rename({2016 : 2020}, axis = columns)

Q35: Which display first two columns of dataframe ‘DF’. A. DF[DF.columns[ 0 : 2 ] ]B. DF.columns[ 0 : 2 ]C. Both of the aboveD. None of the above

Q36: Write a python statement to print the details of employees having Sal more than 50000

A. df1.Sal>=50000B. df1[df1.Sal>=50000]C. df1[df1.'Sal'>=50000]D. df1.iloc[df1.Sal>=50000]

Q37: Display the marks of Harry in Maths Subject. A. print(DF.loc[‘Maths’, ‘Harry’])B. print(DF.Loc[‘Maths’, ‘Harry’])C. print(DF.loc(‘Maths’, ‘Harry’))D. None of the above

Q38: Display data of 1st to 3rd rows A. data[1:4]B. data.iloc[1:4]C. data.loc[:, 1:3]D. None of the above

Q39: What is the correct syntax to return both the first row and the second row in a Pandas DataFrame df? A. df.loc[[0,1]]B. df.[[0,1]]C. df.loc[[0-1]]D. df.[[0-1]]

Q40: Which of the following is the correct statement to access index 3rd and 5th values using positional index for series s? A. s[3,5]B. s[[3,5]]C. s[(3,5)]D. s([3,5])

Q41: Which of the following is correct way of assigning a labelled index to series? A. s=pd.Series(index=range(5,10),[22,33,44,56,78])B. s=pd.Series(index=range(5,10),dt=[22,33,44,56,78])C. s=pd.Series({22,33,44,56,78},index=range(5,10))D. s=pd.Series([22,33,44,56,78],index=range(5,10))

Q42: What will be the output of following code? import pandas as pds=pd.Series([10,20,30,40,50],index={'a','b'','c','d','e'})print(s['d'])A. dB. 30C. 40C. 4

Q43: Write the output of the following : import pandas as pdS1=pd.Series([1,2,3,4])S2=pd.Series([7,8])print((S1+S2).count())A. 6

1. 4
2. 2
3. 0
4. 代码填空题

Q1: 补充完整下列代码，使得实现函数get\_max，函数最终返回列表lst的最大值，不要使用max函数

def get\_max(lst):

if not lst:

return None

max\_value = .

for item in lst:

if item > :

max\_value = item

return max\_value

lst = [4, 2, 1, 6, 7, 9]

max\_value = get\_max(lst)

print(max\_value)

Q2: 补充完整下列代码，使得实现函数get\_max\_socre， 执行程序，最终输出： 英语 98

def get\_max\_socre(score\_dic):

"""

返回学生考试成绩的最高分的科目和分数

:param score\_dic:

:return:

"""

max\_score = 0

max\_score\_course = ''

for course, score in :

if score > max\_score:

max\_score = .

max\_score\_course = .

return , .

dic = {

'语文': 90,

'数学': 97,

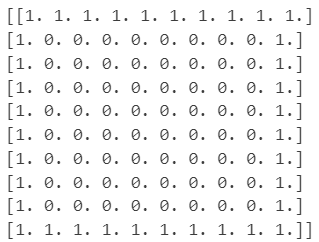
'英语': 98

}

course, score = get\_max\_socre(dic)

print(course, score)

Q3: 补充完整下列代码, 使得创建一个10 x 10的零数组，然后用一个边框“框”起来，如下图所示：



import numpy as np

arr = .

arr[ : , : ] = 0

print(f"The final matrix is\n {arr}")

Q4: 补充完整下列代码, 使得创建一个5×5数组，其（i，j）项等于i+j

import numpy as np

arr\_size = .

A = np.arange(arr\_size)

B = .

C = .

print(f"The generated array is\n {C}")

Q5: 补充完整下列代码, 使得通过使用Pandas过滤分数，打印未通过考试（分数＜60）的男生的姓名

import pandas as pd

records = {

'Name': ['Mary', 'Maria', 'Anna', 'John', 'Jake', 'Joe'],

'Gender': ['Female', 'Female', 'Female', 'Male', 'Male', 'Male'],

'Score': [88, 92, 56, 70, 55, 48]

}

# 创建DataFrame

df = .

# 使用条件过滤

failed\_male\_students = df[( ) & ( )]

print(failed\_male\_students['Name'])

Q6: 补充完整下列代码, 使得将screening.csv读取到DataFrame，并生成列高血压（ Hypertension），诊断标准：如果SBP>=140,或者 DBP>=90, 则为高血压

import pandas as pd

# 读取CSV文件

df = ("screening.csv")

# 生成列 高血压

df['Hypertension'] = (( ) | ( )).astype(int)

Q7: 补充完整下列代码, 使得将test\_virus.csv作为DataFrame读取。使用日期作为索引、病毒作为列以及每3天的计数作为值来透视表。

import pandas as pd

df = (r'test\_vrius.csv', encoding='gb2312')

df3['date'] = (df3['date'])

df\_counts = df3.groupby(['date', 'virus']).size().reset\_index(name='counts')

pivot\_df = df\_counts.pivot\_table(index=pd.Grouper(key=' ', freq='3D'),

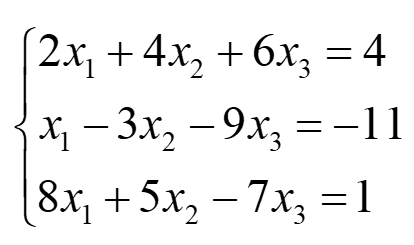
columns=' ',

values=' ',

aggfunc='sum')

print(pivot\_df)

Q8: 补充完整下列代码, 使得使用sciPy/numpy求解以下方程



import numpy as np

import scipy

A = np.array([[ ], [ ], [ ]])

b = np.array([ ])

x = scipy.linalg.solve( , )

print(x)

Q9: 补充完整下列代码, 使得将两个 DataFrame合并到一个 DataFrame 中，并添加一个新的列来标识每个数据来自哪个地区

import pandas as pd

# 创建两个示例 DataFrame，分别表示不同地区的销售数据

data1 = {'Date': ['2024-05-01', '2024-05-02', '2024-05-03'],

'Sales': [1000, 1500, 1200]}

df1 = pd.DataFrame(data1)

data2 = {'Date': ['2024-05-01', '2024-05-02', '2024-05-03'],

'Sales': [900, 1200, 1100]}

df2 = pd.DataFrame(data2)

df\_concatenated = pd. ( , keys=['Region1', 'Region2'])

print(df\_concatenated)

Q10: 补充完整下列代码, 使得计算DataFrame中每个学生的平均成绩和每门课程的平均成绩

import pandas as pd

# 示例数据

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

'Course': ['Math', 'Math', 'English', 'English', 'Math', 'English'],

'Score': [88, 90, 85, 78, 92, 85]}

df = pd.DataFrame(data)

# 计算每个学生的平均成绩

student\_avg\_scores = 。

# 计算每门课程的平均成绩

course\_avg\_scores = 。

Q11：补充完整下列代码, 求一个3\*3矩阵中对角线上元素之和

lst = [

[3,5,6],

[4,7,8],

[2,4,9]

]

sum = 0

for i in range(3):

sum += lst[ ][ ]

print sum

Q12: 补充完整下列代码, 编写一个可以接受可变数量的值并计算其乘积的函数

def product( ):

p = .

for arg in args:

p \*= arg

return p

print(product(2,4))

print(product(2,4,6))