| Medium |
| --- |
| In this code fragment, what will the values of c and d be equivalent to?  import numpy as np  a = np.array([1,2,3])  b = np.array([4,5,6])  c = a\*b  d = np.dot(a,b)  print(d)  1. c = [ a[1] \* b[1], a[2] \* b[2], a[3] \* b[3] ] d = sum(c)  2. c = a[0] \* b[0], a[1] \* b[1], a[2] \* b[2] d = [ a[0] \* b[0], a[1] \* b[1], a[2] \* b[2] ]  3. c = [ a[0] \* b[0], a[1] \* b[1], a[2] \* b[2] ] d = sum(a) + sum(b)  4. c = [ a[0] \* b[0], a[1] \* b[1], a[2] \* b[2] ] d = sum(c) |
| You have a DataFrame df with columns 'Name', 'Age', and 'Salary'. You want to select only the rows where the age is greater than 30 and the salary is above $50,000. Which of the following code snippets achieves this?  1. df[df.Age > 30, df.Salary > 50000]  2. df[df.Age > 30 and df.Salary > 50000]  3. df[(df['Age'] > 30) & (df['Salary'] > 50000)]  4. df[(df['Age']>30) &&(df['Salary'] > 50000)] |
| Suppose you have a DataFrame "df" with a column named "Marks" that contains numerical values. You want to normalize the values in the "Marks" column such that they range between 0 and 1. Which of the following code snippets correctly performs the normalization?  1. df['Marks'] = (df['Marks'] - df['Marks'].mean()) / (df['Marks'].max() - df['Marks'].min())  2. df['Marks'] = (df['Marks'] - df['Marks'].mean()) / df['Marks'].std()  3. df['Marks'] = (df['Marks'] - df['Marks'].min()) / (df['Marks'].max() - df['Marks'].min())  4. df['Marks'] = (df['Marks'] - df['Marks'].min()) / df['Marks'].std() |
| What is the output of this code?  import numpy as np  table = np.array([[1,3],[2,4]])  print(table.max(axis=1))  1. [2, 4]  2. [3, 4]  3. [4]  4. [1,2] |
| What will be output for the following code?  import numpy as np  dt = np.dtype([('age',np.int8)])  a = np.array([(10,),(20,),(30,)], dtype = dt)  print(a)  1. [[10 20 30]]  2. [(10,)( 20,) (30,)]  3. [10]  4. ['age'] |
| What will be output for the following code?  import numpy as np  a = np.array([1, 2, 3,4,5], ndmin = 2)  print(a)  1. [[1, 2, 3, 4, 5]]  2. [1, 2, 3, 4, 5]  3. [(1,2,3,4,5)]  4. [{1,2,3,4,5}] |
| What will be syntax for pandas dataframe?  1. pandas.DataFrame( data, index, dtype, copy)  2. pandas.DataFrame( data, index, rows, dtype, copy)  3. pandas\_DataFrame( data, index, columns, dtype, copy)  4. pandas.DataFrame( data, index, columns, dtype, copy) |
| Fill in the blank to get output as 13 import pandas as pd a = [1, 7, 2,8,13,11] myvar = pd.Series(a, index = 'u','v','w',"x", "y", "z"]) print(myvar[\_\_\_\_\_\_\_\_\_])  1. y  2. 5  3. 'y'  4. 13 |
| What will be output for the following code? import pandas as pd data = { "calories": [420, 380, 390], "duration": [50, 40, 45]} df = pd.DataFrame(data) print(df.loc[0].values)  1. [420]  2. [50]  3. [420 50]  4. Error |
| Which library is commonly used for web scraping in Python?  1. NumPy  2. Pandas  3. Matplotlib  4. BeautifulSoup |
| What is the primary purpose of NumPy in data analysis with Python?  1. Data visualization  2. Data transformation  3. Data storage  4. Data manipulation |
| In NumPy, what does "vectorization" refer to?  1. Converting data to a vector format  2. Performing element-wise operations without explicit loops  3. Plotting data points on a vector space  4. Storing data as vectors |
| What is the primary function of a Pandas DataFrame?  1. Data manipulation  2. Data visualization  3. Data storage  4. Both A and C |
| Which library is commonly used for creating data visualizations in Python?  1. NumPy  2. Pandas  3. Matplotlib  4. Sklearn |
| In data visualization, what does a "box plot" show?  1. Distribution of a single continuous variable  2. Relationship between two continuous variables  3. Relationship between a continuous and a categorical variable  4. Visualizing the central tendency, spread, and presence of outliers in a dataset |
| What is the purpose of the requests library in web scraping?  1. To create regular expressions  2. To send HTTP requests and retrieve web pages  3. To parse XML data  4. To create graphical user interfaces |
| Which of the following is NOT a property of a NumPy array?  1. Homogeneous data type  2. Multidimensional  3. Variable size  4. Efficient element-wise operations |
| In Pandas, what is a "Series"?  1. A Python list  2. A one-dimensional labeled array  3. A two-dimensional table  4. A data visualization library |
| How can you filter data in a Pandas DataFrame based on a specific condition?  1. Using the filter method  2. Using the sort\_values method  3. Using the query method  4. Using the pivot method |

| Medium |
| --- |
| What is NumPy broadcasting?  1. A live streaming feature for arrays  2. A method for distributing arrays over a network  3. A mechanism that allows NumPy to work with arrays of different shapes  4. A protocol for sharing arrays across multiple processes |
| NumPy is often used along with packages like?  1. Node.js  2. SciPy  3. Matplotlib  4. Both B and C |
| What will be output for the following code? import numpy as np dt = np.dtype([('age',np.int8)]) a = np.array([(10,),(20,),(30,)], dtype = dt) print (a['age'])  1. [10]  2. [10 20 30]  3. [[10 20 30]]  4. Error |
| For what purpose a Pandas is used?  1. To create a GUI programming  2. data analysis  3. data cleaning  4. both b and c |
| In Pandas python, the process of changing the structure of the DataFrame is known as \_\_\_\_\_\_\_\_.  1. Importing  2. Exporting  3. Reshaping  4. None of the above |
| What is expected output for the given code? import pandas as pd df=pd.DataFrame({'A':[12,4,5,44,1],'B':[5,2,54,3,2], 'C':[20,16,7,3,8],'D':[14,3,17,2,6]}) df.max(axis=0)  1. A 44 B 54 C 20 D 17  2. A 1 B 2 C 3 D 2  3. A 12 B 5 C 20 D 14  4. A NAN B 2 C NAN D NAN |
| Which argument keyword can be used to change the style of the plotted line?  1. line  2. lines  3. linestyles  4. linestyle |
| Which parameter(s) is/are used to define the color of the points in a scatter plot?  1. color  2. c  3. Both A and B  4. None of the above |
| Which is used for plotting a horizontal line?  1. hline()  2. ahline()  3. xhline()  4. axhline() |
| What function displays row data in a column or column data in a row?  1. Hyperlink  2. Index  3. Transpose  4. None of these |
| You have two columns of numerical data and want to create a visual to help determine if there is a relationship between them. What kind of chart is designed to do this?  1. bar chart  2. bubble chart  3. line chart  4. scatter chart |
| Observe the following code and identify what will be the outcome? import numpy as np x = np.array([[0, 1], [2, 3]]) np.transpose(x)  1. array([[0, 2], [1, 3]])  2. array([[0, 1], [2, 3]])  3. array([[2, 3],[0, 1]])  4. None of the mentioned above |
| Observe the following code and identify what will be the output when we run following code - import pandas as pd df = pd.DataFrame() print (df)  1. Empty DataFrame Columns: [] Index: []  2. Empty Series Columns: [5] Index: [0]  3. Empty DataFrame Columns: [2] Index: [3]  4. None of the mentioned above |
| What will be the output of the following code? import numpy as np array1=np.array([100,200,300,400,500,600,700]) print(array1[1:5:2])  1. [200 300]  2. [200 700]  3. [200 400]  4. [200 600] |
| Python pandas was developed by?  1. Guido van Rossum  2. Travis Oliphant  3. Wes McKinney  4. Brendan Eich |
| How do you set a title for the axis using matplotlib library?  1. ax.set\_title("Axis Title")  2. ax.title("Axis Title")  3. ax.axis\_title("Axis Title")  4. ax.setTitle("Axis Title") |
| Which line of code would create a line plot in the form of red squares with x = x and y = y?  1. plt.plot(x, y, 'red', marker = 'square')  2. plt.plot(x, y, 'R-', marker = 'sqr')  3. plt.plot(x, y, 'r', 's')  4. plt.plot(x, y, 'r-', marker = 's') |
| To display a histogram with a well defined edge we can write?  1. sns.histplot(df["horsepower"], edge="red")  2. sns.histplot(df["horsepower"], edgecolor="red")  3. sns.histplot(df["horsepower"], line="red")  4. sns.histplot(df["horsepower"], linecolor ="red") |
| Which function is used to set a label for the x-axis in pyplot?  1. xlabel()  2. set\_x()  3. x-axis()  4. xaxi() |

| Low |
| --- |
| Which of the following codes is used to extract two columns and three rows of a dataframe ?  1. df.iloc[0:2, 0:2]  2. df.iloc[1:3, 0:2]  3. df.iloc[0:3, [0, 1]]  4. df.iloc[0:3, 0:3] |
| The following code will create a dataframe with \_\_\_\_\_ rows and \_\_\_\_\_ columns Lt = [{'a':10, 'b':20}, {'a':5, 'b':10, 'c':20}] D1 = pd.DataFrame(Lt)  1. 5, 3  2. 3, 2  3. 2, 3  4. 3, 5 |
| The \_\_\_\_\_\_\_\_ function returns its argument with a modified shape, whereas the \_\_\_\_\_\_\_\_ method modifies the array itself.  1. reshape, resize  2. resize, reshape  3. reshape2, resize  4. all of the mentioned |
| What type of error is returned by the following code ? import pandas as pd S1 = pd.Series(data=(31,2,-6),index=[7,9,2,3]) Print(S1)  1. Syntax Error  2. Index Error  3. Value Error  4. None of the above |
| Observe the following code and identify what will be the outcome? import numpy as np x = np.array([ [0, 1] , [2, 3] ]) np.transpose(x)  1. Array ( [ [ 0, 3 ], [ 1, 2 ] ] )  2. Array ( [ [ 3, 2 ], [ 1, 0 ] ] )  3. Array ( [ [ 2, 3 ], [ 0, 1 ] ] )  4. Array ( [ [ 0, 2 ], [ 1, 3 ] ] ) |
| Which of the following options describes a pie chart?  1. Comparison in different categories of data  2. Relationship between two sets of numeric data  3. Contribution of individual values to a total value  4. Data trend over a period of time |
| By default, Plot() function plots :  1. Bar Chart  2. Pie Chart  3. Line Chart  4. Horizontal Bar Chart |
| How can you drop a column named 'column\_name' from a Pandas DataFrame, df?  1. df.remove\_column('column\_name')  2. df.drop('column\_name', axis=1,inplace=True)  3. df.drop\_column('column\_name')  4. df.delete('column\_name') |
| What is the primary data structure in Pandas for handling 1D labeled data?  1. DataFrame  2. Series  3. Array  4. List |
| Which of the following is true about a Pandas DataFrame?  1. It can have different data types for each column.  2. It cannot have missing values.  3. It can only have integers as column names.  4. It is suitable only for small datasets. |
| What is the purpose of the np.zeros() function in NumPy?  1. It creates an array with random values.  2. It returns the length of an array.  3. It creates an array with ones.  4. It creates an array with all elements initialized to zero. |
| What is the purpose of the NumPy function np.linspace(0, 1, 5)?  1. It creates an array with values from 0 to 5.  2. It generates an array with 5 equally spaced values between 0 and 1.  3. It returns an array with 5 random values.  4. It creates an array with values from 0 to 1. |
| What is the purpose of the NumPy function np.linspace(0, 1, 5)?  1. It creates an array with values from 0 to 5.  2. It generates an array with 5 equally spaced values between 0 and 1.  3. It returns an array with 5 random values.  4. It creates an array with values from 0 to 1. |
| Series is a one-dimensional labeled array capable of holding any data type.  1. True  2. False  3. -  4. - |
| Point out the wrong combination with regards to the kind keyword for graph plotting.  1. 'scatter' for scatter plots  2. kde' for hexagonal bin plots  3. 'pie' for pie plots  4. None of the mentioned |
| The \_\_\_\_\_\_\_\_ function returns its argument with a modified shape, whereas the \_\_\_\_\_\_\_\_ method modifies the array itself.  1. reshape, resize  2. resize, reshape  3. reshape2, resize  4. all of the mentioned |
| What does the sns.heatmap() function in Seaborn visualize?  1. Scatter plot  2. Pair plot  3. Correlation matrix  4. 3D plot |
| Which of the following object you get after reading CSV file?  1. DataFrame  2. Character Vector  3. Panel  4. All of the mentioned |
| Which function is used to draw multiple figures in one plot?  1. subplot()  2. subplots()  3. pyplots()  4. pysubplots() |