

Module 7) DA - Analyzing Data with Python

NUMPY

Q.1 Convert a 1D array to a 2D array with 2 rows

Q.2 Get the common items between **a** and **b**

Input:

```
a = np.array([1,2,3,2,3,4,3,4,5,6])
```

```
b = np.array([7,2,10,2,7,4,9,4,9,8])
```

Desired Output:

```
array([2, 4])
```

Q.3 Get all items between 5 and 10 from **a**.

Input:

```
a = np.array([2, 6, 1, 9, 10, 3, 27])
```

Desired Output:

```
(array([6, 9, 10]),)
```

Q.4 Limit the number of items printed in python NumPy array **a** to a maximum of 6 elements.

Input:

```
array([ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14])
```

Desired Output:

```
array([ 0, 1, 2, ..., 12, 13, 14])
```

PANDAS

1. Compute the minimum, 25th percentile, median, 75th, and maximum of ser.

2. Creating A Pandas Data Frame and Using Sample Data Sets

3. Using NumPy, create a Pandas Data Frame with five rows and three columns.

4. For a Pandas Data Frame created from a NumPy array, what is the default behavior for the labels for the columns? For the rows?

5. take csv file contains at least 10,000 rows and 12 columns which numerical and text values according to that continue following steps.

6. Write the code to show the number of rows and columns in data frame.

7. How might you show the first few rows of data frame?

8. If you select a single column from the diamonds Data Frame, what will be the type of the return value?

9. Create a line plot using the x and y values provided below. Label the y-axis as “Y” and label the x-axis as “X”.

x = [3, 4, 5, 6]

y = [1.5, 2, 2.5, 3]

10. Create an array of numbers between 0 and 6 with increments of 0.3 and name it “x”. Then on the same plot, plot x , x^2 , x^3 , and x^4 . For consistency, use the following style lines respectively, “ro”, “bs”, “g”, and “:”. Lastly, make sure that the x-axis covers 0 to 6, while the y-axis spans from 0 to 125. Do not worry if you are not familiar with the style lines — you will recognize them as soon as you see the plot.

11. Heights and initials of a group of individuals are provided below. Create a bar plot titled “Height Comparison” to compare the heights among this group.

height = [179, 155, 191, 152, 188, 177]

names = ['QA', 'WB', 'EC', 'RD', 'TE', 'YF']

12. Plot a histogram of x , where x consists of 100,000 randomly-selected points with a normal distribution (hint: you can use `numpy.random.randn()` to generate the random points). The histogram should have 10 bins. Look at how the histogram changes when we try 20 and 50 bins.

13. What are the features of TensorFlow?

14. List a few limitations of TensorFlow.

15. What do you know about supervised and unsupervised machine learning?