Python Homework NumPy and Visualization Exercises

WEEK 2

Data Science Academy

You are asked to write a Python program according to the following questions.

1. Create a vector with values ranging from 15 to 55 and print all values except the first and last.

2. Create a 3X4 array using np.full().

3. Create a 3x3 matrix filled with values from 10 to 18. Use np.arange() and np.reshape().

4. Create a 5x5 zero matrix with elements on the main diagonal equal to 1, 2, 3, 4, 5 using np.diag().

5. Create a null vector of size 10 using np.zeros() and update sixth value to 11.

6. Convert an array to a float type using np.asfarray().

7. Swap columns in a given array. Such as:

Before swap:	After swap:

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

8. Capitalize the first letter, lowercase, uppercase, swapcase, title-case of all the elements of a given array. Use np.char.capitalize(), np.char.lower(), np.char.upper(), np.char.swapcase(), np.char.title().

9. Get the dates of yesterday, today and tomorrow using np.datetime64() and np.timedelta64().

10. Append values to the end of an array using np.append().

   Original array: [10, 20, 30]
   Expected Output: [10 20 30 40 50 60 70 80 90]

Data Science Academy

Disclaimer: You have been given the 'cars.xlsx' excel dataset. This dataset contains new released car models between 1999 and 2008. Please run the code given below to use it:

import pandas as pd
cars = pd.read_excel('cars.xlsx')
(If necessary please change the directory)

Variable dictionary is listed below:

Manufacturer: manufacturer name
Model: model name
Displ: engine displacement, in litres
Year: year of manufacture
Cyl: number of cylinders
Trans: type of transmission
Drv: the type of drive train, where f = front-wheel drive, r = rear wheel drive, 4 = 4wd
Cty: city miles per gallon
Hwy: highway miles per gallon
Fl: fuel type
Class: "type" of car

Please, answer following questions:
1. Show relationship between **highway** and **city** miles per gallon.
   a. Use scatter in matplotlib.
   b. Interpret how are variables correlated according to scatter.
2. Show distributions and scatters between all variables. While *hue* equals to **type of drive train** which variables have the lowest and the highest correlation? Find according to scatterplot.
3. Which **type of car** is most frequent in dataset? Show by using countplot.
4. Display number of **cylinders** for each **drive train** in bar chart. Which **drive train** is the most frequent?
5. Visualize **engine displacement** by each **class** using boxplot. Do the same thing in violinplot.

Data Science
Academy