Q1. Import numpy as `np` and print the version number.

Q2. Create a 1D array of numbers from 0 to 9

Q3. Extract all odd numbers from numpy array

Q4. Replace all odd numbers in array with -1

Q5. Convert a 1D array to a 2D array with 2 rows

Q6. Stack the arrays a and b horizontally.

Q7. Get the common items between a and b

Q8. From array a remove all items present in array b

Q9. Get the positions where elements of a and b match

Q10. Get all items between 5 and 10 from a.

Q11. Swap columns 1 and 2 in the array arr.

Q12. Swap rows 1 and 2 in the array arr:

Q13. Reverse the rows of a 2D array arr.

# Input

arr = np.arange(9).reshape(3,3)

**19. How to reverse the columns of a 2D array?**

Difficulty Level: L2

Q. Reverse the columns of a 2D array arr.

# Input

arr = np.arange(9).reshape(3,3)

**20. How to create a 2D array containing random floats between 5 and 10?**

Difficulty Level: L2

Q. Create a 2D array of shape 5x3 to contain random decimal numbers between 5 and 10.

**21. How to print only 3 decimal places in python numpy array?**

Difficulty Level: L1

Q. Print or show only 3 decimal places of the numpy array rand\_arr.

Input:

rand\_arr = np.random.random((5,3))

**22. How to pretty print a numpy array by suppressing the scientific notation (like 1e10)?**

Difficulty Level: L1

Q. Pretty print rand\_arr by suppressing the scientific notation (like 1e10)

Input:

# Create the random array

np.random.seed(100)

rand\_arr = np.random.random([3,3])/1e3

rand\_arr

#> array([[ 5.434049e-04, 2.783694e-04, 4.245176e-04],

#> [ 8.447761e-04, 4.718856e-06, 1.215691e-04],

#> [ 6.707491e-04, 8.258528e-04, 1.367066e-04]])

Desired Output:

#> array([[ 0.000543, 0.000278, 0.000425],

#> [ 0.000845, 0.000005, 0.000122],

#> [ 0.000671, 0.000826, 0.000137]])

**23. How to limit the number of items printed in output of numpy array?**

Difficulty Level: L1

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

Input:

a = np.arange(15)

#> 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])

**24. How to print the full numpy array without truncating**

Difficulty Level: L1

Q. Print the full numpy array a without truncating.

Input:

np.set\_printoptions(threshold=6)

a = np.arange(15)

a

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

Desired Output:

a

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

**25. How to import a dataset with numbers and texts keeping the text intact in python numpy?**

Difficulty Level: L2

Q. Import the [iris dataset](https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data) keeping the text intact.

**26. How to extract a particular column from 1D array of tuples?**

Difficulty Level: L2

Q. Extract the text column species from the 1D iris imported in previous question.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_1d = np.genfromtxt(url, delimiter=',', dtype=None)

**27. How to convert a 1d array of tuples to a 2d numpy array?**

Difficulty Level: L2

Q. Convert the 1D iris to 2D array iris\_2d by omitting the speciestext field.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_1d = np.genfromtxt(url, delimiter=',', dtype=None)

**28. How to compute the mean, median, standard deviation of a numpy array?**

Difficulty: L1

Q. Find the mean, median, standard deviation of iris's sepallength (1st column)

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

**29. How to normalize an array so the values range exactly between 0 and 1?**

Difficulty: L2

Q. Create a normalized form of iris's sepallength whose values range exactly between 0 and 1 so that the minimum has value 0 and maximum has value 1.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

sepallength = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0])

**30. How to compute the softmax score?**

Difficulty Level: L3

Q. Compute the softmax score of sepallength.

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

sepallength = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0])

**31. How to find the percentile scores of a numpy array?**

Difficulty Level: L1

Q. Find the 5th and 95th percentile of iris's sepallength

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

sepallength = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0])

**32. How to insert values at random positions in an array?**

Difficulty Level: L2

Q. Insert np.nan values at 20 random positions in iris\_2d dataset

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='object')

**33. How to find the position of missing values in numpy array?**

Difficulty Level: L2

Q. Find the number and position of missing values in iris\_2d's sepallength (1st column)

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float')

iris\_2d[np.random.randint(150, size=20), np.random.randint(4, size=20)] = np.nan

**34. How to filter a numpy array based on two or more conditions?**

Difficulty Level: L3

Q. Filter the rows of iris\_2d that has petallength (3rd column) > 1.5and sepallength (1st column) < 5.0

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0,1,2,3])

**35. How to drop rows that contain a missing value from a numpy array?**

Difficulty Level: L3:

Q. Select the rows of iris\_2d that does not have any nan value.

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0,1,2,3])

**36. How to find the correlation between two columns of a numpy array?**

Difficulty Level: L2

Q. Find the correlation between SepalLength(1st column) and PetalLength(3rd column) in iris\_2d

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0,1,2,3])

**37. How to find if a given array has any null values?**

Difficulty Level: L2

Q. Find out if iris\_2d has any missing values.

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0,1,2,3])

**38. How to replace all missing values with 0 in a numpy array?**

Difficulty Level: L2

Q. Replace all ccurrences of nan with 0 in numpy array

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='float', usecols=[0,1,2,3])

iris\_2d[np.random.randint(150, size=20), np.random.randint(4, size=20)] = np.nan

**39. How to find the count of unique values in a numpy array?**

Difficulty Level: L2

Q. Find the unique values and the count of unique values in iris's species

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**40. How to convert a numeric to a categorical (text) array?**

Difficulty Level: L2

Q. Bin the petal length (3rd) column of iris\_2d to form a text array, such that if petal length is:

* Less than 3 --> 'small'
* 3-5 --> 'medium'
* '>=5 --> 'large'

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**41. How to create a new column from existing columns of a numpy array?**

Difficulty Level: L2

Q. Create a new column for volume in iris\_2d, where volume is (pi x petallength x sepal\_length^2)/3

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris\_2d = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**42. How to do probabilistic sampling in numpy?**

Difficulty Level: L3

Q. Randomly sample iris's species such that setose is twice the number of versicolor and virginica

# Import iris keeping the text column intact

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

**43. How to get the second largest value of an array when grouped by another array?**

Difficulty Level: L2

Q. What is the value of second longest petallength of species setosa

# Input

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**44. How to sort a 2D array by a column**

Difficulty Level: L2

Q. Sort the iris dataset based on sepallength column.

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**45. How to find the most frequent value in a numpy array?**

Difficulty Level: L1

Q. Find the most frequent value of petal length (3rd column) in iris dataset.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

**46. How to find the position of the first occurrence of a value greater than a given value?**

Difficulty Level: L2

Q. Find the position of the first occurrence of a value greater than 1.0 in petalwidth 4th column of iris dataset.

# Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

**47. How to replace all values greater than a given value to a given cutoff?**

Difficulty Level: L2

Q. From the array a, replace all values greater than 30 to 30 and less than 10 to 10.

Input:

np.random.seed(100)

a = np.random.uniform(1,50, 20)

**48. How to get the positions of top n values from a numpy array?**

Difficulty Level: L2

Q. Get the positions of top 5 maximum values in a given array a.

np.random.seed(100)

a = np.random.uniform(1,50, 20)

**49. How to compute the row wise counts of all possible values in an array?**

Difficulty Level: L4

Q. Compute the counts of unique values row-wise.

Input:

np.random.seed(100)

arr = np.random.randint(1,11,size=(6, 10))

arr

> array([[ 9, 9, 4, 8, 8, 1, 5, 3, 6, 3],

> [ 3, 3, 2, 1, 9, 5, 1, 10, 7, 3],

> [ 5, 2, 6, 4, 5, 5, 4, 8, 2, 2],

> [ 8, 8, 1, 3, 10, 10, 4, 3, 6, 9],

> [ 2, 1, 8, 7, 3, 1, 9, 3, 6, 2],

> [ 9, 2, 6, 5, 3, 9, 4, 6, 1, 10]])

Desired Output:

> [[1, 0, 2, 1, 1, 1, 0, 2, 2, 0],

> [2, 1, 3, 0, 1, 0, 1, 0, 1, 1],

> [0, 3, 0, 2, 3, 1, 0, 1, 0, 0],

> [1, 0, 2, 1, 0, 1, 0, 2, 1, 2],

> [2, 2, 2, 0, 0, 1, 1, 1, 1, 0],

> [1, 1, 1, 1, 1, 2, 0, 0, 2, 1]]

Output contains 10 columns representing numbers from 1 to 10. The values are the counts of the numbers in the respective rows.  
For example, Cell(0,2) has the value 2, which means, the number 3 occurs exactly 2 times in the 1st row.

**50. How to convert an array of arrays into a flat 1d array?**

Difficulty Level: 2

Q. Convert array\_of\_arrays into a flat linear 1d array.

Input:

# Input:

arr1 = np.arange(3)

arr2 = np.arange(3,7)

arr3 = np.arange(7,10)

array\_of\_arrays = np.array([arr1, arr2, arr3])

array\_of\_arrays

#> array([array([0, 1, 2]), array([3, 4, 5, 6]), array([7, 8, 9])], dtype=object)

Desired Output:

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

**51. How to generate one-hot encodings for an array in numpy?**

Difficulty Level L4

Q. Compute the one-hot encodings (dummy binary variables for each unique value in the array)

Input:

np.random.seed(101)

arr = np.random.randint(1,4, size=6)

arr

#> array([2, 3, 2, 2, 2, 1])

Output:

#> array([[ 0., 1., 0.],

#> [ 0., 0., 1.],

#> [ 0., 1., 0.],

#> [ 0., 1., 0.],

#> [ 0., 1., 0.],

#> [ 1., 0., 0.]])

**52. How to create row numbers grouped by a categorical variable?**

Difficulty Level: L3

Q. Create row numbers grouped by a categorical variable. Use the following sample from iris species as input.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

species = np.genfromtxt(url, delimiter=',', dtype='str', usecols=4)

species\_small = np.sort(np.random.choice(species, size=20))

species\_small

#> array(['Iris-setosa', 'Iris-setosa', 'Iris-setosa', 'Iris-setosa',

#> 'Iris-setosa', 'Iris-setosa', 'Iris-versicolor', 'Iris-versicolor',

#> 'Iris-versicolor', 'Iris-versicolor', 'Iris-versicolor',

#> 'Iris-versicolor', 'Iris-virginica', 'Iris-virginica',

#> 'Iris-virginica', 'Iris-virginica', 'Iris-virginica',

#> 'Iris-virginica', 'Iris-virginica', 'Iris-virginica'],

#> dtype='<U15')

Desired Output:

#> [0, 1, 2, 3, 4, 5, 0, 1, 2, 3, 4, 5, 0, 1, 2, 3, 4, 5, 6, 7]

**53. How to create groud ids based on a given categorical variable?**

Difficulty Level: L4

Q. Create group ids based on a given categorical variable. Use the following sample from iris species as input.

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

species = np.genfromtxt(url, delimiter=',', dtype='str', usecols=4)

species\_small = np.sort(np.random.choice(species, size=20))

species\_small

#> array(['Iris-setosa', 'Iris-setosa', 'Iris-setosa', 'Iris-setosa',

#> 'Iris-setosa', 'Iris-setosa', 'Iris-versicolor', 'Iris-versicolor',

#> 'Iris-versicolor', 'Iris-versicolor', 'Iris-versicolor',

#> 'Iris-versicolor', 'Iris-virginica', 'Iris-virginica',

#> 'Iris-virginica', 'Iris-virginica', 'Iris-virginica',

#> 'Iris-virginica', 'Iris-virginica', 'Iris-virginica'],

#> dtype='<U15')

Desired Output:

#> [0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2]

**54. How to rank items in an array using numpy?**

Difficulty Level: L2

Q. Create the ranks for the given numeric array a.

Input:

np.random.seed(10)

a = np.random.randint(20, size=10)

print(a)

#> [ 9 4 15 0 17 16 17 8 9 0]

Desired output:

[4 2 6 0 8 7 9 3 5 1]

**55. How to rank items in a multidimensional array using numpy?**

Difficulty Level: L3

Q. Create a rank array of the same shape as a given numeric array a.

Input:

np.random.seed(10)

a = np.random.randint(20, size=[2,5])

print(a)

#> [[ 9 4 15 0 17]

#> [16 17 8 9 0]]

Desired output:

#> [[4 2 6 0 8]

#> [7 9 3 5 1]]

**56. How to find the maximum value in each row of a numpy array 2d?**

DifficultyLevel: L2

Q. Compute the maximum for each row in the given array.

np.random.seed(100)

a = np.random.randint(1,10, [5,3])

a

#> array([[9, 9, 4],

#> [8, 8, 1],

#> [5, 3, 6],

#> [3, 3, 3],

#> [2, 1, 9]])

**57. How to compute the min-by-max for each row for a numpy array 2d?**

DifficultyLevel: L3

Q. Compute the min-by-max for each row for given 2d numpy array.

np.random.seed(100)

a = np.random.randint(1,10, [5,3])

a

#> array([[9, 9, 4],

#> [8, 8, 1],

#> [5, 3, 6],

#> [3, 3, 3],

#> [2, 1, 9]])

**58. How to find the duplicate records in a numpy array?**

Difficulty Level: L3

Q. Find the duplicate entries (2nd occurrence onwards) in the given numpy array and mark them as True. First time occurrences should be False.

# Input

np.random.seed(100)

a = np.random.randint(0, 5, 10)

print('Array: ', a)

#> Array: [0 0 3 0 2 4 2 2 2 2]

Desired Output:

#> [False True False True False False True True True True]

**59. How to find the grouped mean in numpy?**

Difficulty Level L3

Q. Find the mean of a numeric column grouped by a categorical column in a 2D numpy array

Input:

url = '<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>'

iris = np.genfromtxt(url, delimiter=',', dtype='object')

names = ('sepallength', 'sepalwidth', 'petallength', 'petalwidth', 'species')

Desired Solution:

#> [[b'Iris-setosa', 3.418],

#> [b'Iris-versicolor', 2.770],

#> [b'Iris-virginica', 2.974]]

**60. How to convert a PIL image to numpy array?**

Difficulty Level: L3

Q. Import the image from the following URL and convert it to a numpy array.

URL = '<https://upload.wikimedia.org/wikipedia/commons/8/8b/Denali_Mt_McKinley.jpg>'

**61. How to drop all missing values from a numpy array?**

Difficulty Level: L2

Q. Drop all nan values from a 1D numpy array

Input:

np.array([1,2,3,np.nan,5,6,7,np.nan])

Desired Output:

array([ 1., 2., 3., 5., 6., 7.])

**62. How to compute the euclidean distance between two arrays?**

Difficulty Level: L3

Q. Compute the euclidean distance between two arrays a and b.

Input:

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

b = np.array([4,5,6,7,8])

**63. How to find all the local maxima (or peaks) in a 1d array?**

Difficulty Level: L4

Q. Find all the peaks in a 1D numpy array a. Peaks are points surrounded by smaller values on both sides.

Input:

a = np.array([1, 3, 7, 1, 2, 6, 0, 1])

Desired Output:

#> array([2, 5])

where, 2 and 5 are the positions of peak values 7 and 6.

**64. How to subtract a 1d array from a 2d array, where each item of 1d array subtracts from respective row?**

Difficulty Level: L2

Q. Subtract the 1d array b\_1d from the 2d array a\_2d, such that each item of b\_1d subtracts from respective row of a\_2d.

a\_2d = np.array([[3,3,3],[4,4,4],[5,5,5]])

b\_1d = np.array([1,1,1]

Desired Output:

#> [[2 2 2]

#> [2 2 2]

#> [2 2 2]]

**65. How to find the index of n'th repetition of an item in an array**

Difficulty Level L2

Q. Find the index of 5th repetition of number 1 in x.

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

**66. How to convert numpy's datetime64 object to datetime's datetime object?**

Difficulty Level: L2

Q. Convert numpy's datetime64 object to datetime's datetime object

# Input: a numpy datetime64 object

dt64 = np.datetime64('2018-02-25 22:10:10')

**67. How to compute the moving average of a numpy array?**

Difficulty Level: L3

Q. Compute the moving average of window size 3, for the given 1D array.

Input:

np.random.seed(100)

Z = np.random.randint(10, size=10)

**68. How to create a numpy array sequence given only the starting point, length and the step?**

Difficulty Level: L2

Q. Create a numpy array of length 10, starting from 5 and has a step of 3 between consecutive numbers

**69. How to fill in missing dates in an irregular series of numpy dates?**

Difficulty Level: L3

Q. Given an array of a non-continuous sequence of dates. Make it a continuous sequence of dates, by filling in the missing dates.

Input:

# Input

dates = np.arange(np.datetime64('2018-02-01'), np.datetime64('2018-02-25'), 2)

print(dates)

#> ['2018-02-01' '2018-02-03' '2018-02-05' '2018-02-07' '2018-02-09'

#> '2018-02-11' '2018-02-13' '2018-02-15' '2018-02-17' '2018-02-19'

#> '2018-02-21' '2018-02-23']

**70. How to create strides from a given 1D array?**

Difficulty Level: L4

Q. From the given 1d array arr, generate a 2d matrix using strides, with a window length of 4 and strides of 2, like [[0,1,2,3], [2,3,4,5], [4,5,6,7]..]

Input:

arr = np.arange(15)

arr

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

Desired Output:

#> [[ 0 1 2 3]

#> [ 2 3 4 5]

#> [ 4 5 6 7]

#> [ 6 7 8 9]

#> [ 8 9 10 11]

#> [10 11 12 13]]