

Lab2. Red Wine Quality Data Analysis using NumPy Part-II

Suriya S

225229140

Import necessary modules

In [4]:

```
#1
import numpy as np
```

In [5]:

```
wines = np.genfromtxt("winequality-red.csv", delimiter=";", skip_header=1)
```

In [6]:

```
#2
x=wines[:,3]
sum(x)
```

Out[6]:

```
4059.550000000003
```

In [7]:

```
#3
sum(wines)
```

Out[7]:

```
array([[13303.1    ,  843.985   ,  433.29    ,  4059.55    ,  139.859    ,
        25384.    ,  74302.    , 1593.79794,  5294.47    ,  1052.38    ,
        16666.35   ,  9012.     ]])
```

In [8]:

```
#4
wines[:,:].sum(axis=1)
```

Out[8]:

```
array([ 74.5438 , 123.0548 ,  99.699   , ..., 100.48174, 105.21547,
        92.49249])
```

In [9]:

```
#5  
wines.shape[0]
```

Out[9]:

1599

In [10]:

```
#6  
y=wines[:,3]  
y=y.astype('int32')  
y
```

Out[10]:

array([1, 2, 2, ..., 2, 2, 3])

In [11]:

```
#7  
max(y)
```

Out[11]:

15

In [12]:

```
#8  
min(y)
```

Out[12]:

0

In [13]:

```
#9  
np.mean(x)
```

Out[13]:

2.53880550343965

In [14]:

```
#10  
np.percentile(x,25)
```

Out[14]:

1.9

In [15]:

```
#11
np.percentile(x,75)
```

Out[15]:

2.6

In [16]:

```
#12
n=wines[:,:]
n
n.mean(axis=0)
```

Out[16]:

```
array([ 8.31963727,  0.52782051,  0.27097561,  2.5388055 ,  0.08746654,
        15.87492183, 46.46779237,  0.99674668,  3.3111132 ,  0.65814884,
        10.42298311,  5.63602251])
```

NumPy Array Comparison

In [17]:

```
#1
wines[:,11]>5
```

Out[17]:

```
array([False, False, False, ...,  True, False,  True])
```

In [18]:

```
#2
a=wines[:,11]>7
a
```

Out[18]:

```
array([False, False, False, ..., False, False, False])
```

In [19]:

```
#3
True in a
```

Out[19]:

True

In [20]:

```
34
hq=wines[:,11]>7
hq
```

Out[20]:

```
array([False, False, False, ..., False, False, False])
```

In [21]:

```
#5
wines[hq][0:3]
```

Out[21]:

```
array([[7.900e+00, 3.500e-01, 4.600e-01, 3.600e+00, 7.800e-02, 1.500e+01,
        3.700e+01, 9.973e-01, 3.350e+00, 8.600e-01, 1.280e+01, 8.000e+00],
       [1.030e+01, 3.200e-01, 4.500e-01, 6.400e+00, 7.300e-02, 5.000e+00,
        1.300e+01, 9.976e-01, 3.230e+00, 8.200e-01, 1.260e+01, 8.000e+00],
       [5.600e+00, 8.500e-01, 5.000e-02, 1.400e+00, 4.500e-02, 1.200e+01,
        8.800e+01, 9.924e-01, 3.560e+00, 8.200e-01, 1.290e+01, 8.000e+00]])
```

In [22]:

```
#6
al=wines[:,10]>10
al
```

Out[22]:

```
array([False, False, False, ..., True, True, True])
```

In [23]:

```
#7
w=al & a
w
```

Out[23]:

```
array([False, False, False, ..., False, False, False])
```

In [24]:

```
#8
wines[:,10:]
```

Out[24]:

```
array([[ 9.4,  5. ],
       [ 9.8,  5. ],
       [ 9.8,  5. ],
       ...,
       [11. ,  6. ],
       [10.2,  5. ],
       [11. ,  6. ]])
```

Combining NumPy Arrays

In [25]:

```
#1
w_wines = np.genfromtxt("winequality-white.csv", delimiter=";", skip_header=1)
```

In [26]:

```
#2
w_wines.shape
```

Out[26]:

```
(4898, 12)
```

In [28]:

```
#3
all_wines=np.vstack((wines,w_wines))
```

In [29]:

```
all_wines.shape
```

Out[29]:

```
(6497, 12)
```

In [31]:

```
#4
all_wines1=np.concatenate((wines,w_wines),axis=0)
```

Matrix Operations and Reshapes

In [32]:

```
#1
all_wines1.shape
```

Out[32]:

```
(6497, 12)
```

In [33]:

```
tran=wines.T
tran.shape
```

Out[33]:

```
(12, 1599)
```

In [34]:

```
#2
wines.ravel()
```

Out[34]:

```
array([ 7.4 ,  0.7 ,  0.   , ...,  0.66, 11.   ,  6.   ])
```

In [35]:

```
#3
wines.ravel().shape
```

Out[35]:

```
(19188,)
```

In [36]:

```
#4
wines[1].reshape((2,6))
```

Out[36]:

```
array([[ 7.8   ,  0.88   ,  0.    ,  2.6   ,  0.098 , 25.    ],
       [67.    ,  0.9968,  3.2   ,  0.68   ,  9.8   ,  5.    ]])
```

In [37]:

```
#5
sorted_alcohol=np.sort(wines[:,2])
```

In [38]:

```
sorted_alcohol
```

Out[38]:

```
array([ 8.4,  8.4,  8.5, ..., 14. , 14. , 14.9])
```

In [39]:

```
#6
wines[:,2].sort()
```

In [40]:

```
#7
wines[:,2]
```

Out[40]:

```
array([ 8.4,  8.4,  8.5, ..., 14. , 14. , 14.9])
```

In [41]:

```
#8  
sorted_alcohol_desc=np.sort(wines[:,10])[:, -1]
```

In [42]:

```
sorted_alcohol_desc
```

Out[42]:

```
array([14.9, 14. , 14. , ...,  8.5,  8.4,  8.4])
```

In [43]:

```
#9  
wines[:, -2]
```

Out[43]:

```
array([ 8.4,  8.4,  8.5, ..., 14. , 14. , 14.9])
```

In []: