DVA_LAB_01

NAME : RETHINAGIRI.G

ROLL_NO : 225229130

Red Wine Quality Data Analytics using NumPy Part-I

In [1]:
```
import numpy as num
```

In [2]:
```
#Read  the Data

w=num.genfromtxt("winequality-red.csv",delimiter=";",skip_header=1)
print(w)
```

```
[[ 7.4    0.7    0.    ...  0.56   9.4    5.   ]
 [ 7.8    0.88   0.    ...  0.68   9.8    5.   ]
 [ 7.8    0.76   0.04  ...  0.65   9.8    5.   ]
 ...
 [ 6.3    0.51   0.13  ...  0.75  11.     6.   ]
 [ 5.9    0.645  0.12  ...  0.71  10.2    5.   ]
 [ 6.     0.31   0.47  ...  0.66  11.     6.   ]]
```

In [3]:
```
#size of the data

si=w.size
si
```

Out[3]: 19188

In [4]:
```
#shape of the data

print(w.shape)
print("rows=",w.shape[0])
```

```
(1599, 12)
rows= 1599
```

In [5]:
```
#Dimention

w.ndim
```

Out[5]: 2

In [6]:
```
#Type Of the Data

print(type(w))
```

```
<class 'numpy.ndarray'>
```

In [7]:
```
#Datatype

w.dtype
```

Out[7]: dtype('float64')

In [8]:
```
#top 5 rows

top=w[:5,:]
print(top)
```

```
[[7.400e+00 7.000e-01 0.000e+00 1.900e+00 7.600e-02 1.100e+01 3.400e+01
  9.978e-01 3.510e+00 5.600e-01 9.400e+00 5.000e+00]
 [7.800e+00 8.800e-01 0.000e+00 2.600e+00 9.800e-02 2.500e+01 6.700e+01
  9.968e-01 3.200e+00 6.800e-01 9.800e+00 5.000e+00]
 [7.800e+00 7.600e-01 4.000e-02 2.300e+00 9.200e-02 1.500e+01 5.400e+01
  9.970e-01 3.260e+00 6.500e-01 9.800e+00 5.000e+00]
 [1.120e+01 2.800e-01 5.600e-01 1.900e+00 7.500e-02 1.700e+01 6.000e+01
  9.980e-01 3.160e+00 5.800e-01 9.800e+00 6.000e+00]
 [7.400e+00 7.000e-01 0.000e+00 1.900e+00 7.600e-02 1.100e+01 3.400e+01
  9.978e-01 3.510e+00 5.600e-01 9.400e+00 5.000e+00]]
```

In [9]:
```
# 3rd row,4th column of Data

w[2,3]
```

Out[9]:
```
2.3
```

In [10]:
```
#First 3item in 4th column

w[0:3,3]
```

Out[10]:
```
array([1.9, 2.6, 2.3])
```

In [12]:
```
#1st column

w[0:,0]
```

Out[12]:
```
array([7.4, 7.8, 7.8, ..., 6.3, 5.9, 6. ])
```

In [13]:
```
#2nd row

w[1,0:]
```

Out[13]:
```
array([ 7.8   ,  0.88  ,  0.    ,  2.6   ,  0.098 , 25.    , 67.    ,
        0.9968,  3.2   ,  0.68  ,  9.8   ,  5.    ])
```

In [14]:
```
# rows 1 to 3 and 5th column

w[1:4,4]
```

Out[14]:
```
array([0.098, 0.092, 0.075])
```

In [15]:
```
#Entire Data

w
```

Out[15]:
```
array([[ 7.4  ,  0.7  ,  0.   , ...,  0.56 ,  9.4  ,  5.   ],
       [ 7.8  ,  0.88 ,  0.   , ...,  0.68 ,  9.8  ,  5.   ],
       [ 7.8  ,  0.76 ,  0.04 , ...,  0.65 ,  9.8  ,  5.   ],
       ...,
       [ 6.3  ,  0.51 ,  0.13 , ...,  0.75 , 11.   ,  6.   ],
       [ 5.9  ,  0.645,  0.12 , ...,  0.71 , 10.2  ,  5.   ],
       [ 6.   ,  0.31 ,  0.47 , ...,  0.66 , 11.   ,  6.   ]])
```

In [16]:
```
#chance 1st Value to 100

print(w[0,0])
r=w[0,0]=100
print(r)
```

```
7.4
100
```

In [17]:
```
#Change it back to 7.4

r=7.4
r
```

Out[17]:
```
7.4
```

1-Dimentional Numpy Arrays

In [18]:
```
#select 4th row

fr=w[3,]
print(fr)
```

```
[11.2    0.28   0.56   1.9    0.075 17.    60.     0.998  3.16   0.58
  9.8    6.   ]
```

In [19]:
```
#show 2nd Value

fr[1]
```

Out[19]:
```
0.28
```

In [20]:
```
#Convert Data to Integervalue

w.astype(int)
```

Out[20]:
```
array([[100,   0,   0, ...,   0,   9,   5],
       [  7,   0,   0, ...,   0,   9,   5],
       [  7,   0,   0, ...,   0,   9,   5],
       ...,
       [  6,   0,   0, ...,   0,  11,   6],
       [  5,   0,   0, ...,   0,  10,   5],
       [  6,   0,   0, ...,   0,  11,   6]])
```

2-Vectorization Opearation

In [21]:
```
# Increase wine Quality Score by 10

v=w[:,11]
s=v.astype(int)
s
```

Out[21]:
```
array([5, 5, 5, ..., 6, 5, 6])
```

In [22]:
```
s+=10
```

In [23]:
```
print([s])
```

```
[array([15, 15, 15, ..., 16, 15, 16])]
```

In [24]:
```
# Multiply data by 3 times

m=w[:,10]*3
m
```

Out[24]:
```
array([28.2, 29.4, 29.4, ..., 33. , 30.6, 33. ])
```

In [25]:
```
#add Quality by it self

add=v+v
add
```

Out[25]:
```
array([10., 10., 10., ..., 12., 10., 12.])
```

In [26]:
```
# Multiply 11th and 12th colums

mul=s*m
mul
```

Out[26]:
```
array([423., 441., 441., ..., 528., 459., 528.])
```

3-Broadcasting

In [27]:
```
# Add every row of winne Data with a random array of Values

ran=num.random.rand(12)
ran
```

Out[27]:
```
array([0.44243305, 0.79434499, 0.34734571, 0.49321355, 0.97687396,
       0.4054912 , 0.68387849, 0.13918862, 0.06135908, 0.38906348,
       0.85517212, 0.28752534])
```

In [28]:
```
# Add wines and random array

w+ran
```

Out[28]:
```
array([[100.44243305,   1.49434499,   0.34734571, ...,   0.94906348,
          10.25517212,   5.28752534],
       [  8.24243305,   1.67434499,   0.34734571, ...,   1.06906348,
          10.65517212,   5.28752534],
       [  8.24243305,   1.55434499,   0.38734571, ...,   1.03906348,
          10.65517212,   5.28752534],
       ...,
       [  6.74243305,   1.30434499,   0.47734571, ...,   1.13906348,
          11.85517212,   6.28752534],
       [  6.34243305,   1.43934499,   0.46734571, ...,   1.09906348,
          11.05517212,   5.28752534],
       [  6.44243305,   1.10434499,   0.81734571, ...,   1.04906348,
          11.85517212,   6.28752534]])
```