

NUMPY_PRACTICE

June 20, 2021

```
[1]: import numpy as np
```

```
[79]: data = np.loadtxt("Salary_Data.csv", delimiter=",", dtype = float, skiprows=1)
      data.dtype
```

```
[79]: dtype('float64')
```

```
[80]: data
```

```
[80]: array([[1.10000e+00, 3.93430e+04],
            [1.30000e+00, 4.62050e+04],
            [1.50000e+00, 3.77310e+04],
            [2.00000e+00, 4.35250e+04],
            [2.20000e+00, 3.98910e+04],
            [2.90000e+00, 5.66420e+04],
            [3.00000e+00, 6.01500e+04],
            [3.20000e+00, 5.44450e+04],
            [3.20000e+00, 6.44450e+04],
            [3.70000e+00, 5.71890e+04],
            [3.90000e+00, 6.32180e+04],
            [4.00000e+00, 5.57940e+04],
            [4.00000e+00, 5.69570e+04],
            [4.10000e+00, 5.70810e+04],
            [4.50000e+00, 6.11110e+04],
            [4.90000e+00, 6.79380e+04],
            [5.10000e+00, 6.60290e+04],
            [5.30000e+00, 8.30880e+04],
            [5.90000e+00, 8.13630e+04],
            [6.00000e+00, 9.39400e+04],
            [6.80000e+00, 9.17380e+04],
            [7.10000e+00, 9.82730e+04],
            [7.90000e+00, 1.01302e+05],
            [8.20000e+00, 1.13812e+05],
            [8.70000e+00, 1.09431e+05],
            [9.00000e+00, 1.05582e+05],
            [9.50000e+00, 1.16969e+05],
            [9.60000e+00, 1.12635e+05],
            [1.03000e+01, 1.22391e+05],
```

```
[1.05000e+01, 1.21872e+05]])
```

```
[33]: data.ndim
```

```
[33]: 2
```

```
[35]: data.shape
```

```
[35]: (30, 2)
```

```
[81]: for i,n in np.ndenumerate(data):  
      print(f"el indice es {i} y el los valores son {n}")
```

```
el indice es (0, 0) y el los valores son 1.1  
el indice es (0, 1) y el los valores son 39343.0  
el indice es (1, 0) y el los valores son 1.3  
el indice es (1, 1) y el los valores son 46205.0  
el indice es (2, 0) y el los valores son 1.5  
el indice es (2, 1) y el los valores son 37731.0  
el indice es (3, 0) y el los valores son 2.0  
el indice es (3, 1) y el los valores son 43525.0  
el indice es (4, 0) y el los valores son 2.2  
el indice es (4, 1) y el los valores son 39891.0  
el indice es (5, 0) y el los valores son 2.9  
el indice es (5, 1) y el los valores son 56642.0  
el indice es (6, 0) y el los valores son 3.0  
el indice es (6, 1) y el los valores son 60150.0  
el indice es (7, 0) y el los valores son 3.2  
el indice es (7, 1) y el los valores son 54445.0  
el indice es (8, 0) y el los valores son 3.2  
el indice es (8, 1) y el los valores son 64445.0  
el indice es (9, 0) y el los valores son 3.7  
el indice es (9, 1) y el los valores son 57189.0  
el indice es (10, 0) y el los valores son 3.9  
el indice es (10, 1) y el los valores son 63218.0  
el indice es (11, 0) y el los valores son 4.0  
el indice es (11, 1) y el los valores son 55794.0  
el indice es (12, 0) y el los valores son 4.0  
el indice es (12, 1) y el los valores son 56957.0  
el indice es (13, 0) y el los valores son 4.1  
el indice es (13, 1) y el los valores son 57081.0  
el indice es (14, 0) y el los valores son 4.5  
el indice es (14, 1) y el los valores son 61111.0  
el indice es (15, 0) y el los valores son 4.9  
el indice es (15, 1) y el los valores son 67938.0  
el indice es (16, 0) y el los valores son 5.1  
el indice es (16, 1) y el los valores son 66029.0  
el indice es (17, 0) y el los valores son 5.3
```

```

el indice es (17, 1) y el los valores son 83088.0
el indice es (18, 0) y el los valores son 5.9
el indice es (18, 1) y el los valores son 81363.0
el indice es (19, 0) y el los valores son 6.0
el indice es (19, 1) y el los valores son 93940.0
el indice es (20, 0) y el los valores son 6.8
el indice es (20, 1) y el los valores son 91738.0
el indice es (21, 0) y el los valores son 7.1
el indice es (21, 1) y el los valores son 98273.0
el indice es (22, 0) y el los valores son 7.9
el indice es (22, 1) y el los valores son 101302.0
el indice es (23, 0) y el los valores son 8.2
el indice es (23, 1) y el los valores son 113812.0
el indice es (24, 0) y el los valores son 8.7
el indice es (24, 1) y el los valores son 109431.0
el indice es (25, 0) y el los valores son 9.0
el indice es (25, 1) y el los valores son 105582.0
el indice es (26, 0) y el los valores son 9.5
el indice es (26, 1) y el los valores son 116969.0
el indice es (27, 0) y el los valores son 9.6
el indice es (27, 1) y el los valores son 112635.0
el indice es (28, 0) y el los valores son 10.3
el indice es (28, 1) y el los valores son 122391.0
el indice es (29, 0) y el los valores son 10.5
el indice es (29, 1) y el los valores son 121872.0

```

```
[38]: len(data)
```

```
[38]: 30
```

```
[83]: años_trabajo = data[:,0]
```

```
[84]: salario = data[:,1]
```

```
[85]: orden = np.arange(1,31)
orden
```

```
[85]: array([ 1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16, 17,
        18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30])
```

```
[86]: #años de trabajo promedio
promedio_simple = np.sum(años_trabajo)/len(años_trabajo)
promedio
```

```
[86]: 5.3133333333333335
```

```
[95]: #promedio ponderado de trabajo
promedio_ponderado = np.sum(salario*años_trabajo)/sum(años_trabajo)
```

```
promedio_ponderado
```

```
[95]: 89849.19071518193
```

```
[115]: np.mean(años_trabajo)
```

```
[115]: numpy.float64
```

```
[109]: np.min(años_trabajo)
```

```
[109]: 1.1
```

```
[111]: np.max(años_trabajo)
```

```
[111]: 10.5
```

```
[98]: np.median(años_trabajo)
```

```
[98]: 4.7
```

```
[100]: np.var(años_trabajo)
```

```
[100]: 7.785155555555556
```

```
[99]: np.std(años_trabajo)
```

```
[99]: 2.790189161249745
```

```
[101]: np.corrcoef(años_trabajo)
```

```
[101]: 1.0
```

```
[122]: def estadisticos(data):  
        print("la media ", np.mean(data))  
        print("el minimo valor es ", np.min(data), "\nel maximo valor es", np.  
        ↪max(data))  
        print("la mediana es ", np.median(data))  
        print("la varianza es ", np.var(data))  
        print("la desviacion estandar es ", np.std(data))  
        print("el coeficiente de correlacion es ", np.corrcoef(data))  
  
        estadisticos(salario)
```

```
la media 76003.0  
el minimo valor es 37731.0  
el maximo valor es 122391.0  
la mediana es 65237.0  
la varianza es 726499261.7333333
```

la desviacion estandar es 26953.65024877583
el coeficiente de correlacion es 1.0

```
[131]: import matplotlib.pyplot as plt

plt.scatter(años_trabajo,salario)
plt.xlabel("años de trabajo")
plt.ylabel("salario")
plt.xlim(0,12)
plt.ylim(30000,130000 )
plt.title("distribucion del salio y años de la empresa S.A. ")
plt.show()
```

