

④ Linear Regression

```
* import csv
import numpy as np
import matplotlib.pyplot as plt
```

```
def estimate_coef(x, y):
```

```
    n = np.size(x)
```

```
    m_x = np.mean(x) #mean
```

```
    m_y = np.mean(y)
```

```
    ss_xy = np.sum(y*x) - n*m_y*m_x
```

```
    ss_xx = np.sum(x*x) - n*m_x*m_x
```

```
    # calculating regression coefficients
```

```
    b_1 = ss_xy / ss_xx
```

```
    b_0 = m_y - b_1*m_x
```

```
    return (b_0, b_1)
```

```
def plot_regression(x, y, b):
```

```
    plt.scatter(x, y, color="m", marker="o", s=30)
```

```
    y_pred = b[0] + b[1]*x
```

```
    plt.plot(x, y_pred, color="g")
```

```
    plt.xlabel('x')
```

```
    plt.ylabel('y')
```

```
    plt.show()
```

(4)

```
def main():
```

```
    x1, y1 = [], []
```

```
    with open('..\\input\\sampledata\\sampledata.csv', 'r') as csvfile:
```

```
        next(csvfile)
```

```
        for row in csv.reader(csvfile):
```

```
            i, j = row[0], row[1]
```

```
            x1.append(i)
```

```
            y1.append(j)
```

```
    x = np.array(x1).astype(np.float)
```

```
    y = np.array(y1).astype(np.float)
```

```
    b = estimate_coef(x, y)
```

```
    Print("Estimated coefficients: ")
```

```
    Print("b_0 =", b[0])
```

```
    Print("b_1 =", b[1])
```

```
    # graph
```

```
    plot = regression(x, y, b)
```

```
main()
```


(u)

Dataset used

x	y
43	95
21	65
25	79
42	75
57	87
59	81
55	?