

```
In [14]: from numpy import genfromtxt
import matplotlib.pyplot as plt
import chardet
import numpy as np
```

```
In [23]: def simlin_coef(x, y):
n = np.size(x)
m_x = np.mean(x)
m_y = np.mean(y)
AA_xy = np.sum(y*x)-n*m_y*m_x
AA_xx = np.sum(x*x)-n*m_x*m_x
b_1 = AA_xy / AA_xx
b_0 = m_y - b_1*m_x
return(b_0, b_1)
```

```
In [30]: def simlin_plot(x, y, b):
plt.scatter(x, y, color = "b", s = 30)
y_pred = b[0] + b[1]*x
plt.plot(x, y_pred, color = "r")
plt.xlabel('Experience')
plt.ylabel('Salary')
plt.show()
```

```
In [31]: def main():
#READ THE DATA SET
with open('team_big.csv', 'rb') as f:
dataset = chardet.detect(f.read())
dataset = genfromtxt('team_big.csv', dtype=None, delimiter=',', names=True, encoding=dataset['encoding'])
#I added encoding because it gave to me encoding error
x = dataset['Experience']
y = dataset['Salary']
b = simlin_coef(x, y)
simlin_plot(x, y, b)
```

```
In [32]: if __name__ == "__main__":
main()
```

