

In [3]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv("Z:/College/3.2/ML LAB/Final lab/datasets/salary_data.csv")
x=df.iloc[:, :-1].values
y=df.iloc[:, -1].values
mean=np.mean(x,axis=0)
std=np.std(x,axis=0)
x=(x-mean)/std
def compute_cost(x,y,theta):
    m=len(y)
    h=x.dot(theta)
    j=(1/2*m)*np.sum((h-y)**2)
    return j
def batch_gradient_descent(x,y,alpha,num_iters):
    m=len(y)
    n=x.shape[1]
    theta=np.zeros(n)
    j_history=[]
    for i in range(num_iters):
        h=x.dot(theta)
        error=h-y
        gradient=(1/m)*x.T.dot(error)
        theta=theta-alpha*gradient
        j=compute_cost(x,y,theta)
        j_history.append(j)
    return theta,j_history
alpha=0.002
num_iters=10
theta,j_history=batch_gradient_descent(x,y,alpha,num_iters)
print("optimal parameters values:",theta)
```

optimal parameters values: [522.62278014]

In [ ]: