作业第三问：

训练集划分：

from sklearn.model\_selection import train\_test\_split

使用train\_test\_split()函数划分训练集和测试集，

x\_train:划分的训练集数据

x\_test:划分的测试集数据

y\_train:划分的训练集标签

y\_test:划分的测试集标

test\_size=0.2:分割比例

random\_state=30:随机数种子，之后划分结果都相同

x\_train**,** x\_test**,** y\_train**,** y\_test = train\_test\_split(x**,** y**,** test\_size=**0.2,** random\_state=**30**)

使用模型：最小二乘法多元线性回归和简单线性回归模型

数据集验证：x\_test、y\_test，测试集验证

验证公式：MAE、MSE、RMSE

MAE: 3.352956659680632

MSE: 20.469440010390198

RMSE: 4.524316524116123

效果：我觉得，效果一般

运行代码的结果：

多元最小二乘线性回归：

w:[-1.97782083e-02 4.06929747e-02 1.55552334e-02 3.09545222e+00

-1.39901130e+01 4.33621883e+00 2.74196225e-02 -1.21731126e+00

2.74309634e-01 -1.12067681e-02 -8.86228044e-01 1.31679553e-02

-6.68710053e-01]

b:27.142292978826703

MAE: 3.352956659680632

MSE: 20.469440010390198

RMSE: 4.524316524116123

最小二乘线性回归模型可视化：

w:9.613204275451118

b:-37.07467942279478

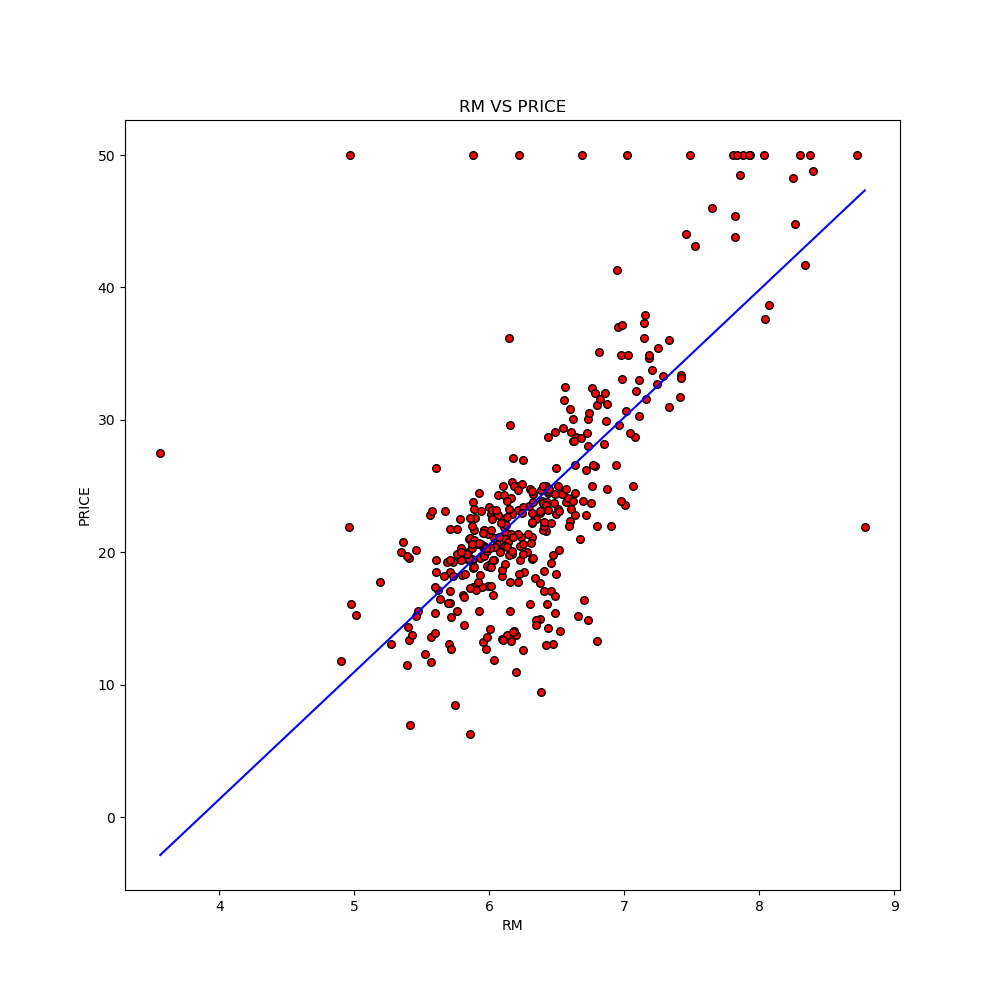
由于MSE值较高,使用梯度下降算法再建立一次模型，得到w、b数值近似，最小二乘模型没有大问题，问题可能出现在异常数据中

梯度下降线性回归：

w:9.777870607934407

b:-38.275847475092675

最小二乘线性回归模型可视化：



波士顿数据集指标解释：

文本

中度可信度描述已自动生成