

2021 中级计量经济学作业 4

1. 课本 574 页习题 28.2
2. (选做)housing.csv 是美国加州 1990 年的房价数据。请分别用线性回归和决策树模型（也可以用任意你觉得合理、有效的其他模型，或者 Ensembling）对房价的中值（median_house_value）做出预测。注意以下几个问题：
 1. 设定模型好坏判断标准 (evaluation metrics)
 2. 数据划分 training, validation 以及 test(train-validation-test split)
 3. 防止模型过拟合 (overfitting)
 4. 可自行创建新的自变量 (feature engineering)