

1. 导入 House Price.csv 数据
2. 输出数据集的大小、所有列名以及目标变量的信息
3. 删除 “id” 列，输出删除后的数据大小
4. 将 “GrLivArea” 和 “SalePrice” 的关系进行可视化(散点图), 并找到离群值
5. 删除 4 中的离群值，并将删除后的数据进行可视化
6. 计算所有特征值每两个之间的相关系数，并作图表示
7. 作 SalePrice 的直方图，观察是否呈正态分布，若不是，将其转换为正态分布，并直方图显示
8. 找到训练集和测试集的缺失值并打印
9. 对有缺失值的列都进行填充（注意填充的格式），并输出填充后的数据集（无缺失值）
10. 将 “MSSubClass”， “OverallCond”， “YrSold”， “MoSold” 列的数据类型修改为字符串类型
11. 将 “TotalBsmtSF”， “1stFlrSF”， “2ndFlrSF” 数据进行整合