“**HomesForSale**”数据变量含义：

*State* : 美国的州（NY, NJ, PA， CA）

*Price* : 房价（in thousands）

*Size* : 房子的大小 （in thousands of square feet）

*Beds* : 房子的房间个数

*Baths* : 房子的卫生间个数

1. 做数据各个变量，及相关变量间的描述性分析。
2. NY与NJ州的房子价格是否相同？（这里先用t-test，但主要是为了强调第一点中观测到价格分布quite skewed，t-test不一定适合。）
3. 有3个以上卫生间的房子的比例在每个州是多少？这个比例在NY与NJ之间是否显著不相同？
4. 研究不同州的房价是否一样；房子的房间个数是否影响了房价（这里可能需要把提出一个新的变量，以房间个数分类。）；卫生间个数是否影响了房价
5. 单看CA的房子，卫生间的个数能预测房子的大小吗？还有什么能预测？
6. 对房价建模，找出能预测房价的变量，并做模型检验。
7. 对房价聚类