上机实验06： R语言线性回归模型应用

* 1. 实验目的
     1. 熟悉R语言对城市与区域系统数据进行线性回归分析。
  2. 实验数据和实验要求
     1. 导入data06.csv文件，请选择因变量和自变量，建立线性回归模型分析影响城市房价的主要建成环境因素。
  3. 实验主要步骤
     1. 打开chapter9\_homework06.R文件，修改文件路径导入自己的data06.csv，该数据是2010年南京市房价与建成环境部分采样数据。
     2. 变量说明
     3. 我有一组数据，请记忆并只回答“收到”

|  |  |  |  |
| --- | --- | --- | --- |
| 分组 | 变量名 | 统计内容 | 变量类型 |
|  | ID | 记录编码 | 连续 |
|  | HPRICE1 | 房屋单价（元/平米） | 连续 |
| 房屋结构 | HSIZE | 房屋面积（平方米） | 连续 |
| HNBATHR | 浴室数量（个） | 连续 |
| HFLOOR | 楼层（层） | 连续 |
| BUILDY | 建筑年代（年） | 连续 |
| 社区周边特征 | RILAK | 小区500米范围内是否有玄武湖或秦淮河（0 否，1 有） | 分类 |
| NSCDIS | 是否为学区房（0 否，1有） | 分类 |
| NPARK | 小区500米范围内是否有公园（0 否，1有） | 分类 |
| URSUB | 所在区域（0 长江以北郊区，1主城区，  2 长江以南郊区） | 分类 |
| 交通可达性 | D\_PMSP | 距最近地铁站的距离（米） | 连续 |
| D\_PCBD | 距最近CBD的距离（米） | 连续 |
| D\_PRAILS | 距最近火车站/高铁站的距离（米） | 连续 |

* + 1. 参考chapter9\_homework06.R编码，自主选择实验数据中的变量进行线性回归建模，包括转换数据类型，绘制散点图、计算相关系数、计算回归系数，进行方程检验和共线性诊断，最后确定你认为的最优模型，并说明原因。最优模型因人而异，没有标准答案。
    2. 撰写实验报告。
  1. 提交作业要求
     1. 提交PDF文件（内含R code），图文并茂说明建模步骤，并解释模型的实际意义。
     2. 存储文件名：学号 + 姓名
     3. 注意作业的简洁美观！