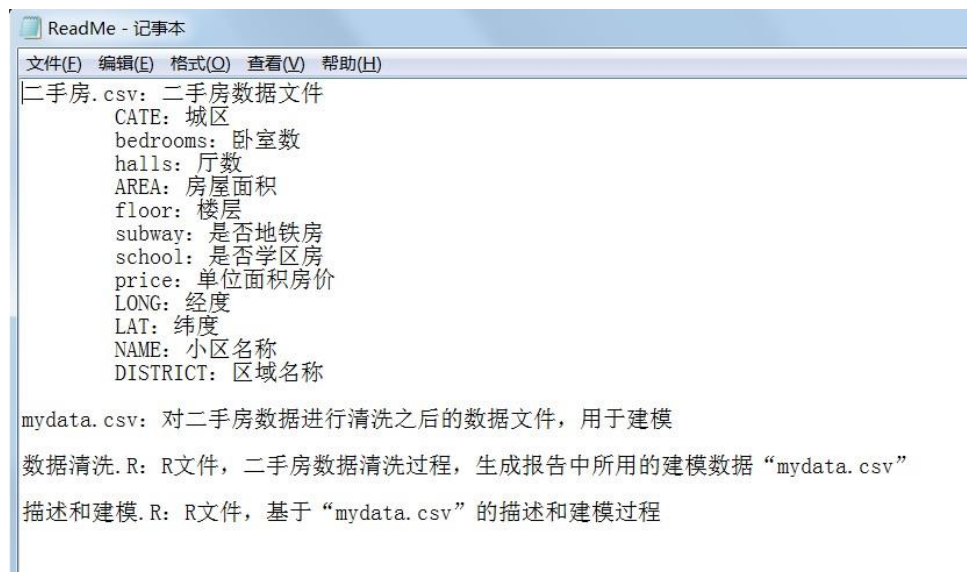


# 作业规范

## 一、 主要构成（基于 R 语言版本）

1. ppt（必须）
2. R 语言代码（必须）
3. README 文档（有多个数据文件或 R 代码，必须），如图示：



图示 README 文档

4. 案例报告（必须）

## 二、ppt 制作规范

1. 不强制 ppt 模板，可自选；
2. ppt 的模板配色与图表颜色保持高度一致；
3. ppt 字体保持一致（强烈推荐微软雅黑）；
4. ppt 内容至少包含以下几个部分（可以有侧重点）：
  - 4.1. 背景介绍
  - 4.2. 描述分析
  - 4.3. 统计模型
  - 4.4. 结论与建议
5. ppt 要有目录页，每一部分标题明确，切忌胡乱堆砌；
6. ppt 要有封面和封底页，封面选图与 ppt 内容一致；
7. 如果没有案例报告，ppt 中对于结果解读的文字尽量详尽。避免一个统计图，几个 keyword；或者一个回归结果表，没有任何解读。
8. 请回忆二手房案例，感受 ppt 制作强迫症患者的作品，并以此为奋斗目标。

## 三、R 代码规范和注意事项

1. 最终编码为“UTF-8”；
2. 代码以简洁规范为主，能用简单语句实现的分析避免复杂化；
3. 自定义函数和加载 R 包需在代码第一行注释清楚；
4. 保持 R 代码的美观，以 block 的样子呈现出来（见图示）；
5. 必须逐行写注释，注释要清楚详细；
6. R 代码的书写顺序与 ppt 保持一致；
7. 一段 R 代码示例（来自二手房案例）：

```

###厅数做因子化处理，变成二分变量，使得建模有更好的解读
style=rep("其他",n)
style[which(dat0$hall==0)]= "无厅"
style[which(dat0$hall>0)]= "有厅"
style=factor(style,levels=c("无厅","有厅"))

###将是否有厅bind到已有数据集
dat0=cbind(dat0,style)

###线性回归模型
lm1=lm(price~CATE+school+subway+style+floor+bedrooms+AREA,data=dat0)
summary(lm1)           #回归结果展示
par(mfrow=c(2,2))      #画2*2的图
plot(lm1,which=c(1:4))  #模型诊断图，存在异方差现象，对因变量取对数

###对数线性模型
lm2=lm(log(price)~CATE+school+subway+style+floor+bedrooms+AREA,data=dat0)
summary(lm2)           #回归结果展示
par(mfrow=c(2,2))      #画2*2的图
plot(lm2,which=c(1:4)) #模型诊断图，异方差现象得到消除

```

图示 一段规范的 R 代码

8. 示范代码说明：在写 R 代码注释的时候，鼓励将代码运行之后展示的结果注上，比如图示中最后一行注释“异方差现象得到消除”这种评论。这样读者运行完代码，能够同时体会到注释中的简单解读。