**《编程和应用程序R语言》期末考核答案**

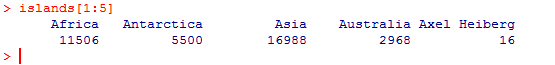
根据学习的R语言编程技术完成给定案例项目的数据处理与基本分析任务，回答相关数据分析问题。要求：

1、综合应用题，提交**R Console界面或者Rstudio操作界面**截图，操作代码与结果截图，截图务必与题目要求、分析要求相关，且**清晰**；

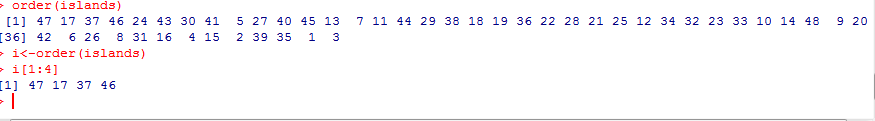
2、**截图**包括操作的命令语句、操作结果。

**一、综合应用题（共6小题，100分）**

1、使用系统内建数据集islands（该数据集含有全球48个岛屿的名称及面积），列出：（10分）

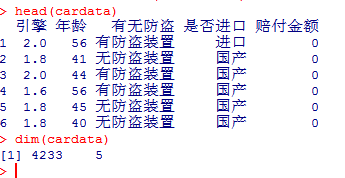
（1）取出前5个岛屿的面积。

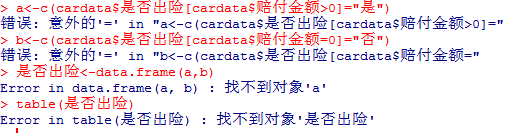
（2）对islands进行升序排序，并显示排序后的前4个岛屿与面积。

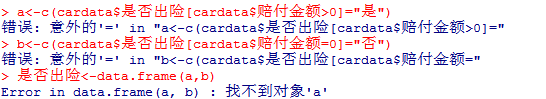


2、将车险数据auto.csv导入R，并进行以下操作：（20分）

（1）导入后命名为cardata。

（2）显示数据前6行，并计算导入后的数据维度（行列数）。



（3）新生成一个字段起名为“是否出险”，其中字段“赔付金额”中大于0的在新生成的字段“是否出险”中记为“是”， 字段“赔付金额”中等于0的在新生成的字段“是否出险”中记为“否”。

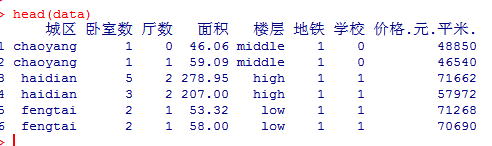
（4）计算上一问中“是否出险”中的“是”和“否”分别有多少个？

3、对二手房数据“二手房.csv”进行以下操作：（20分）

（1）导入“二手房.csv”数据，命名为data。



（2）查看该数据集data的前六行数据。



（3）找出价格最高的房子的全部信息。

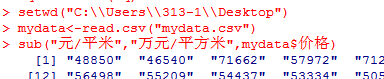


（4）取出面积变量在（30,300）之间的观测值，将取出的数据集命名为data0，并计算其行列数。

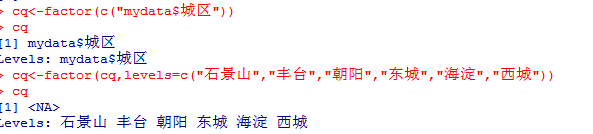


4、对数据集mydata.csv，进行以下操作：（30分）

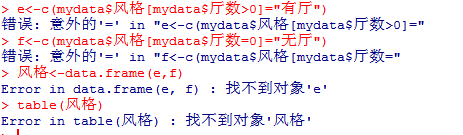
（1）导入数据集命名为mydata,并将价格的单位转换成万元/平方米；



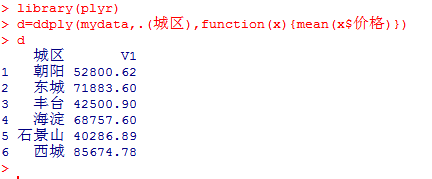
（2）调整城区的因子水平顺序，调整成的顺序为"石景山","丰台","朝阳","东城","海淀","西城"；



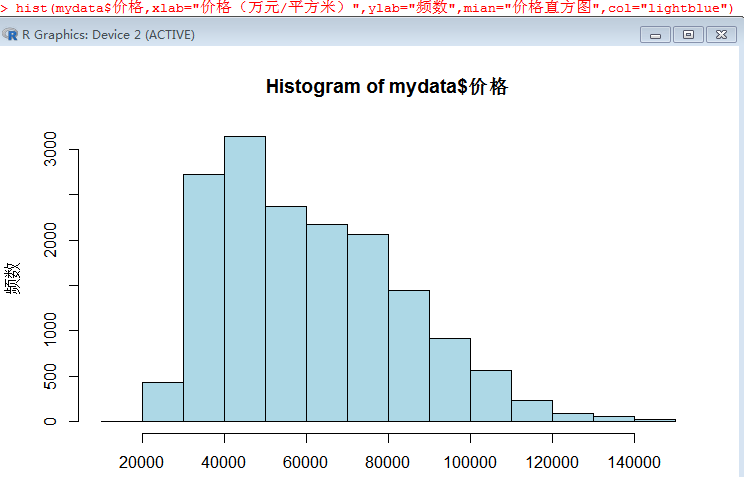
（3）新生成一列字段‘风格’，将字段厅数中等于0的在字段“风格”中记为“无厅”，大于0的记为“有厅”；计算此时有厅和无厅的个数分别是多少。



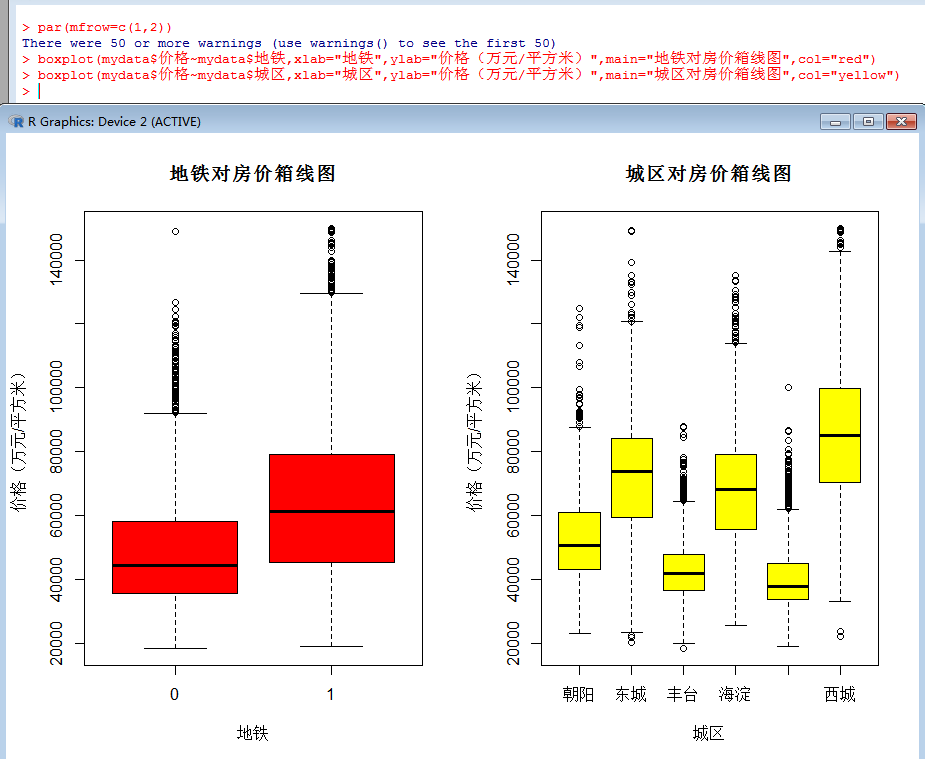
（4）计算不同城区的房价均值；



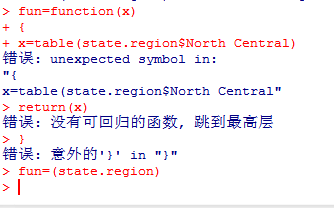
（5）绘制价格分布的直方图，需要设置颜色为“lightblue”、横轴为“价格（万元/平方米）”、纵轴为“频数”，整个图的名字为“价格直方图”；



（6）设置1\*2的画布，绘制地铁、城区对房价的分组箱线图。其中，地铁对房价的分组箱线图需要：设置颜色为“red”、横轴为“地铁”、纵轴的名字为“价格（万元/平方米）”，整个图的名字为“地铁对房价箱线图”。城区对房价的分组箱线图需要：设置颜色为“yellow”、横轴为“城区”、纵轴的名字为“价格（万元/平方米）”，整个图的名字为“城区对房价箱线图”。城区的图需要放大，让每个城区的名字都显示出来。

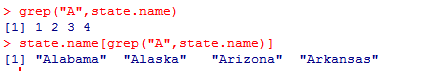


5、编写函数fun，其功能是计算系统内建数据向量state.region属于‘North Central’的有多少个州。（10分）



6、字符串操作（10分）

（1）请搜索state.name数据集中，州名含‘A’的州。



（2）将state.name数据集中，州名中‘N’的替换成‘n’，并显示此时的全部州名。

