实验报告

成绩：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 班级学号 | 202015116 | 姓名 | 梁祺若 | 实验日期 | 2022-9-28 | 仪器编号 |  |
| 实验名称 | 实验2-3 R的数据可视化，R的近邻分析 | | | | | | |

**1．实验题目**

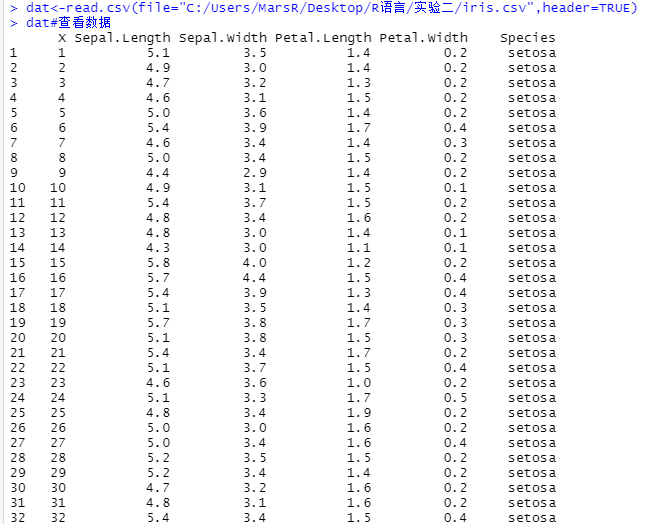
R的数据可视化，R的近邻分析

**2．实验步骤**

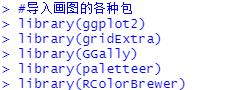
本题目由两部分组成：（1）利用R软件做数据可视化；（2）利用R软件对鸢尾花数据集数据做近邻分析。

1. 利用R软件做数据可视化

* 数据来源：https://cloud.tencent.com/developer/article/1869024
* 导入鸢尾花数据集

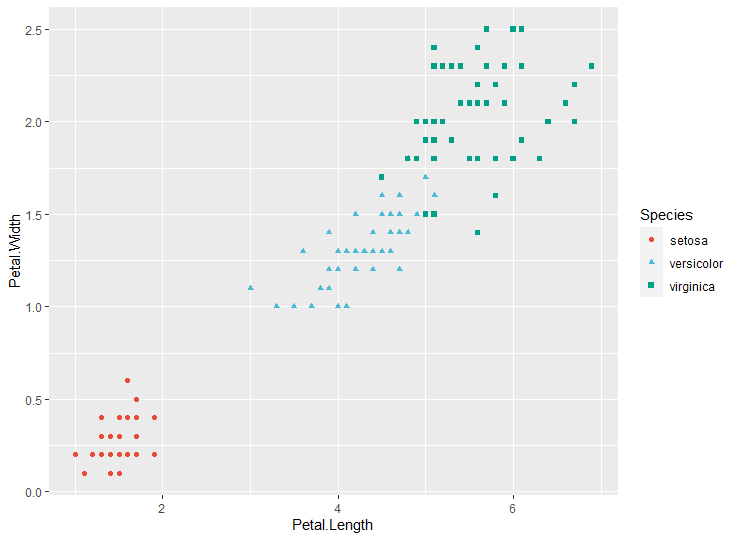
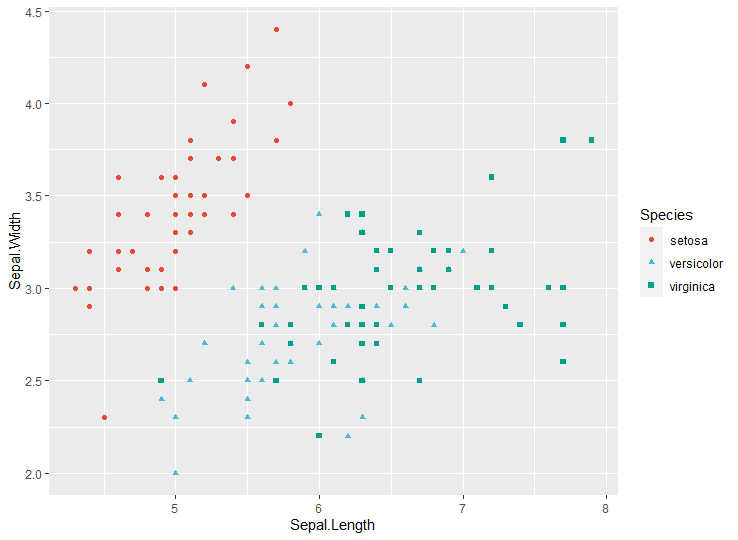


* 导入画图的各种包

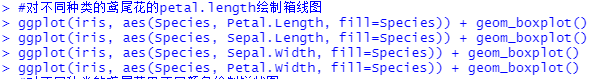
****

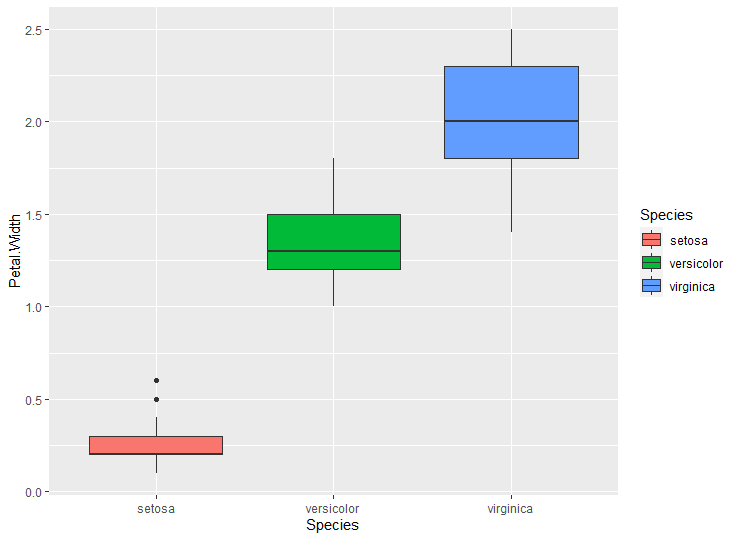
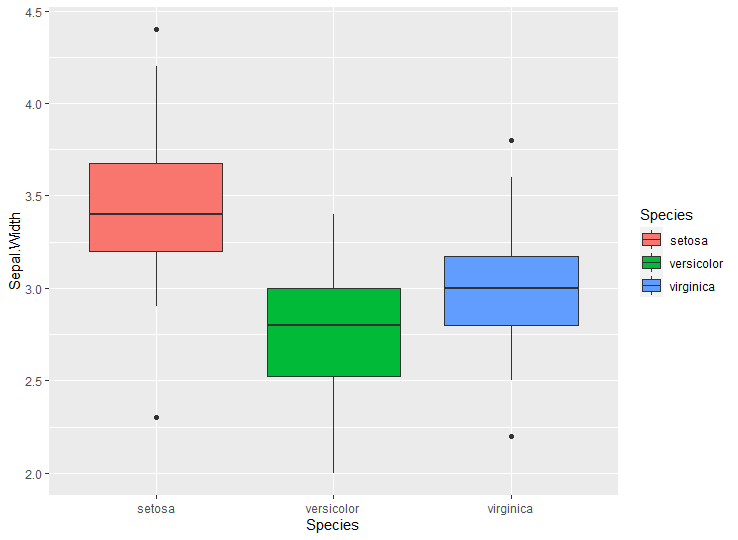
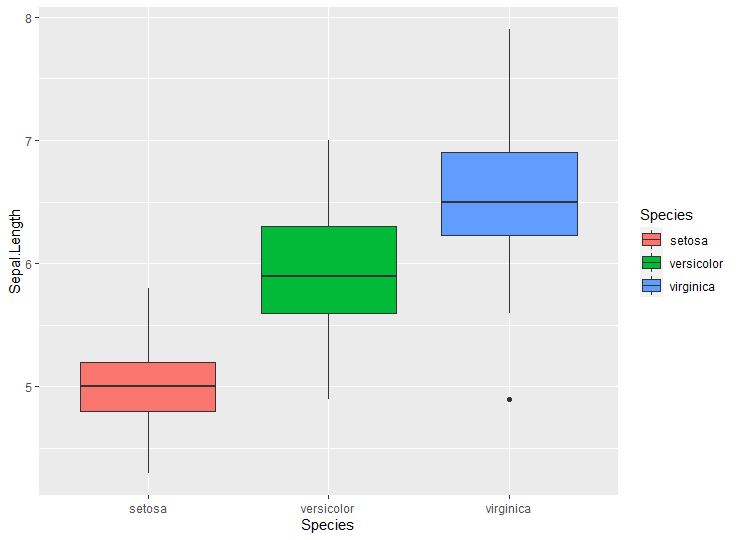
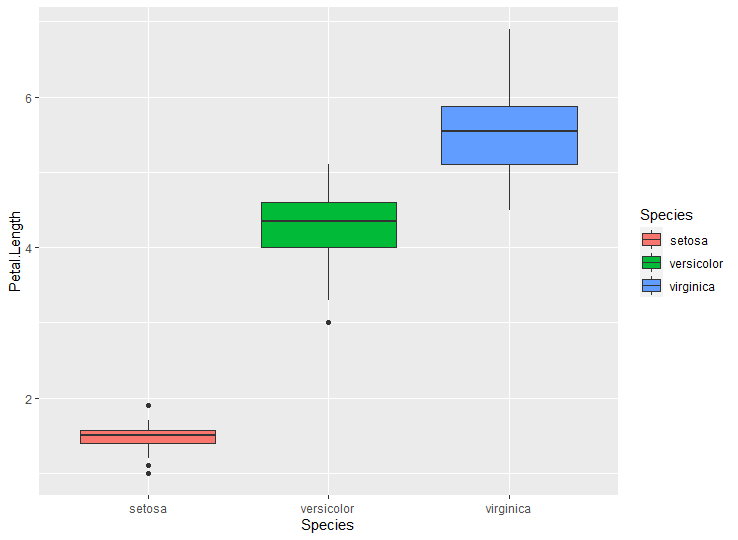
* 对不同种类的鸢尾花用不同颜色绘制散点图





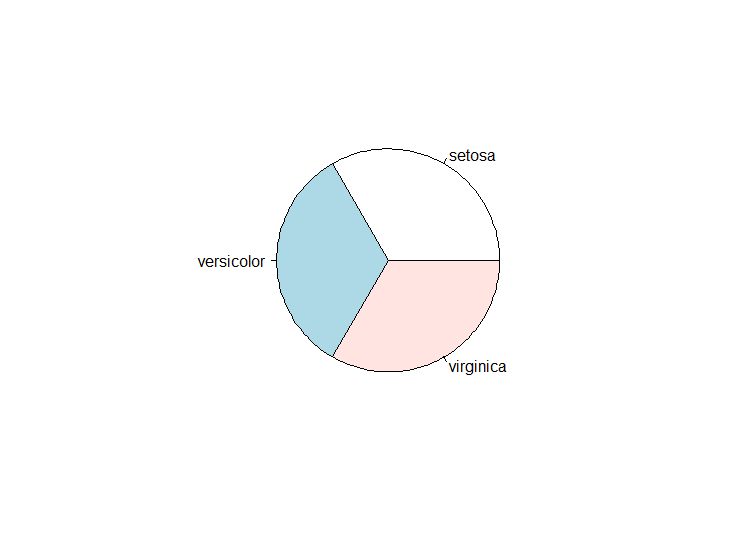
* 对不同种类的鸢尾花绘制箱线图



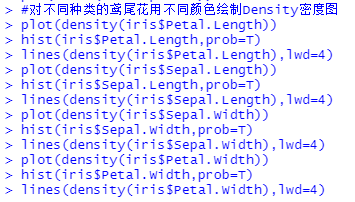


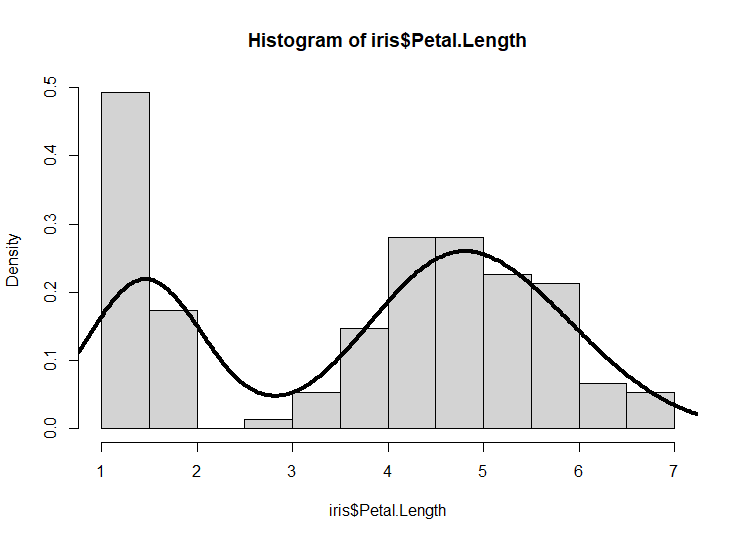
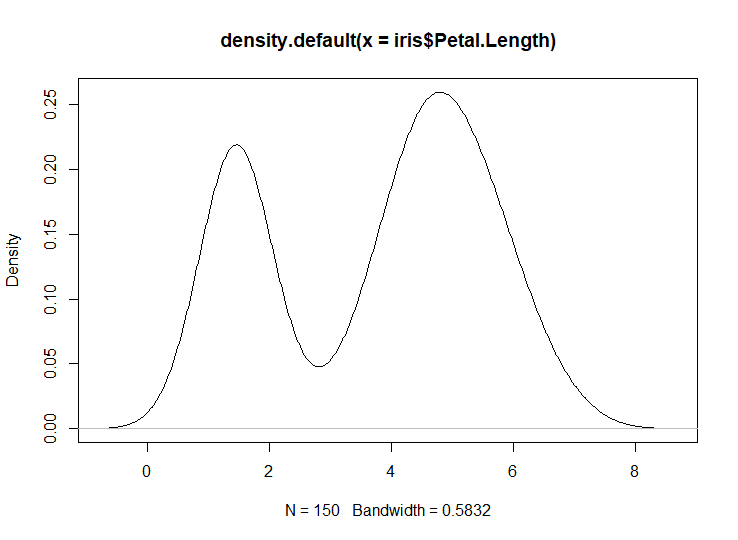
* 对不同种类的鸢尾花用不同颜色绘制饼状图

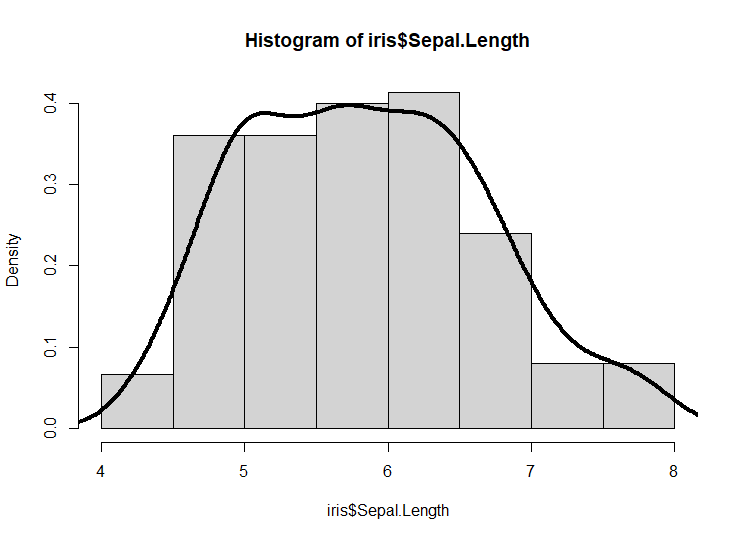
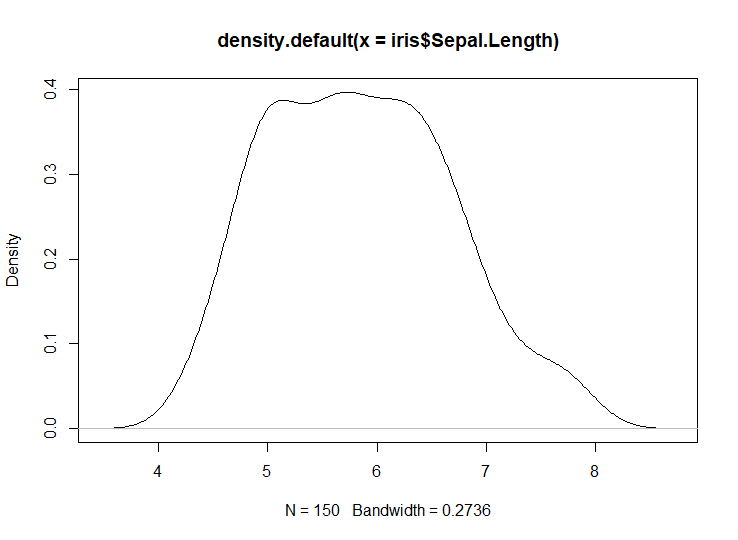


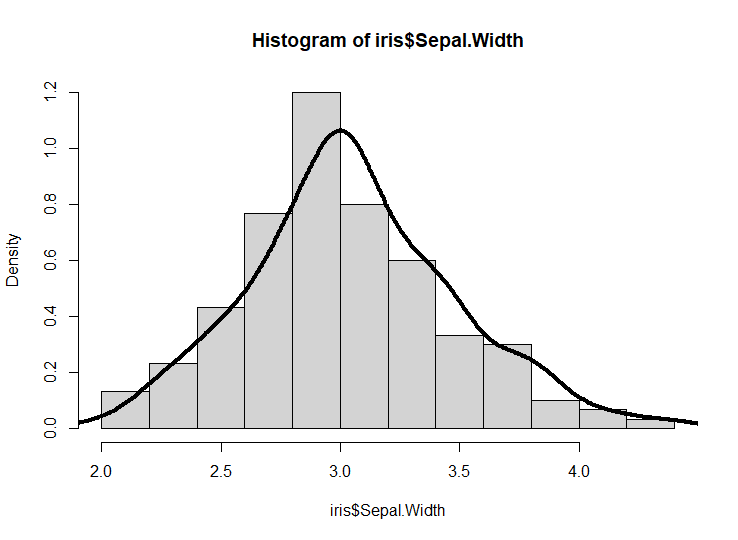
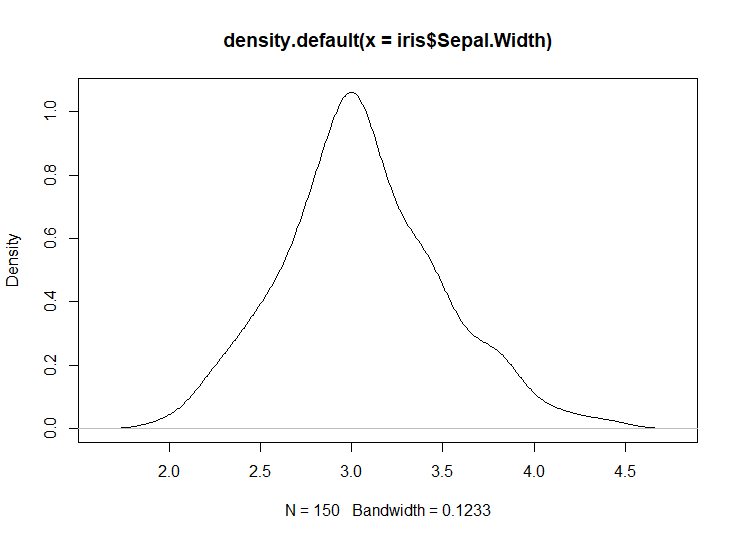


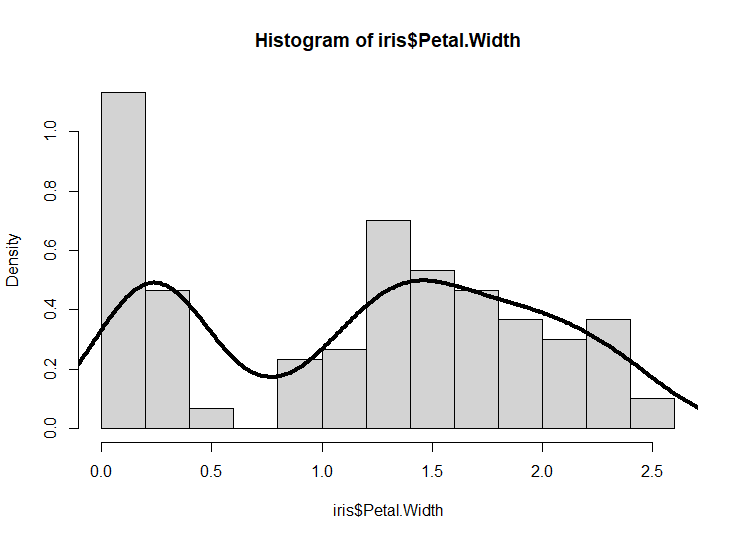
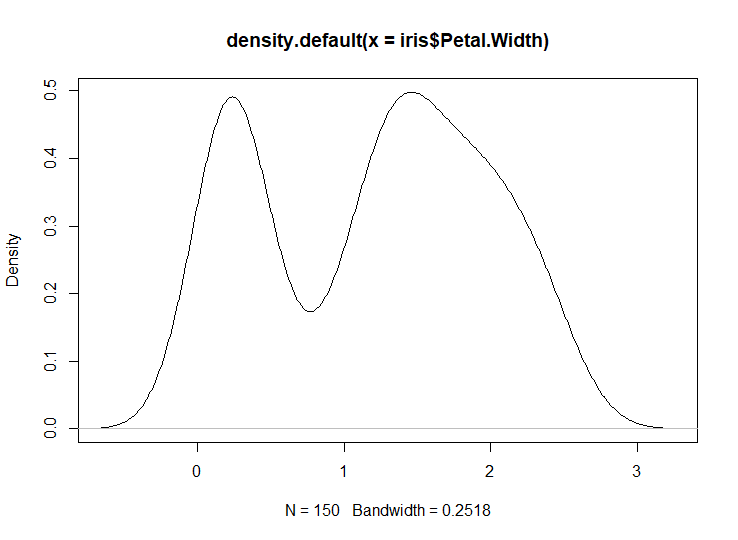
* 对不同种类的鸢尾花用不同颜色绘制Density密度图



****

****

****

****

（2）利用R软件对鸢尾花数据集数据做近邻分析。

* 可视化样本之间相关性



* 导入包



* 构建测试集



* 构建训练集

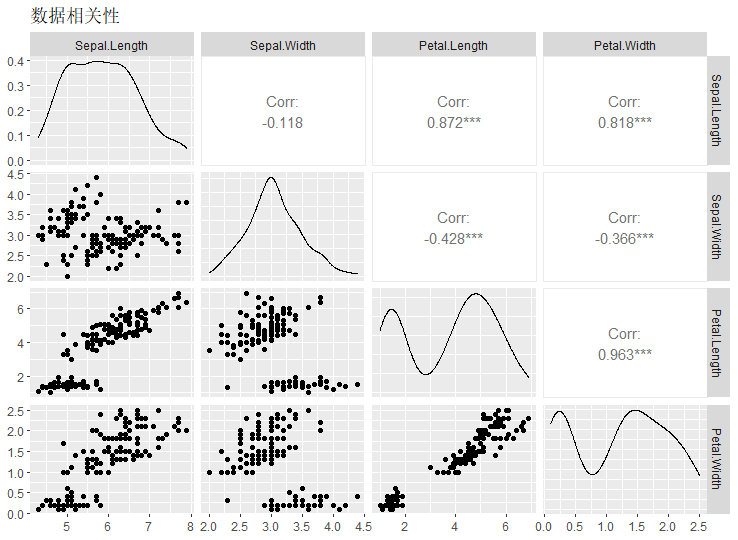


* 利用knn进行训练



* 查看多分类问题的混淆矩阵





**3．实验结果**

导入鸢尾花数据集，并对该数据集进行可视化处理，得到散点图，箱线图饼状图和Density密度图。最后对鸢尾花数据进行近邻分析后得到混淆矩阵，通过混淆矩阵可知预测的准确度较高。