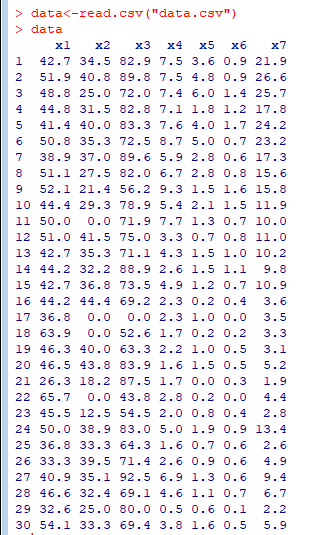
**利用聚类和判别分析法**

**对NBA篮球运动员做技术鉴定**

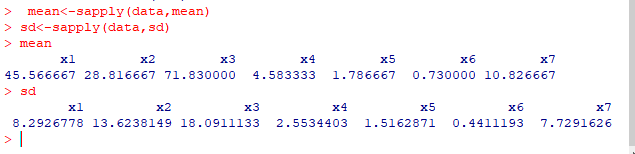
1.聚类分析

在对运动员数据集进行判别分析之前，先对数据集的整体进行聚类分析，用read.table()函数加载数据集，观察数据集的结构，代码如下：

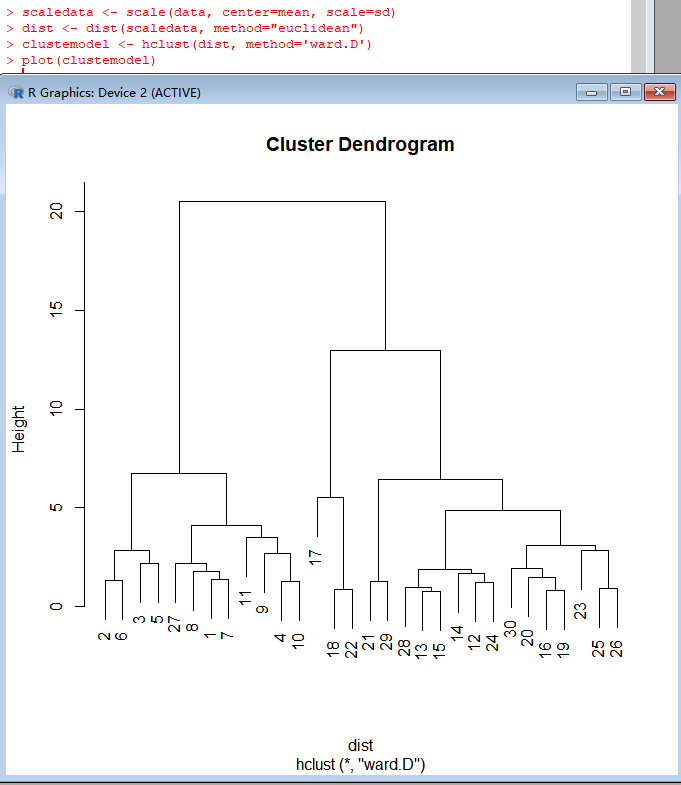


包含30个样本数据，每条样本数据包含7个数据，分别代表NBA各球投篮、三分、罚篮、篮板、助攻、抢断和场均得分。

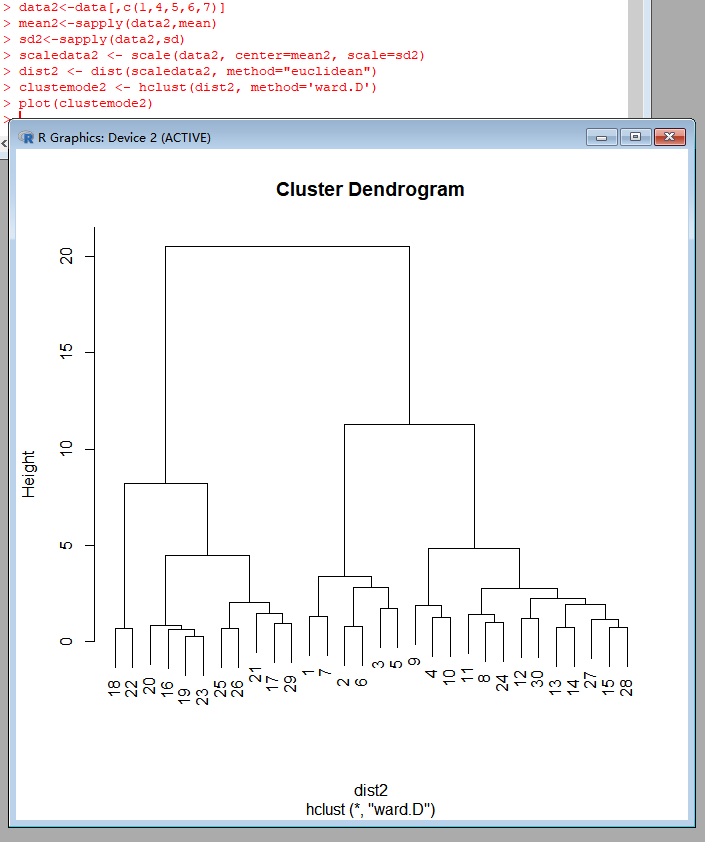
计算其均值和标准差



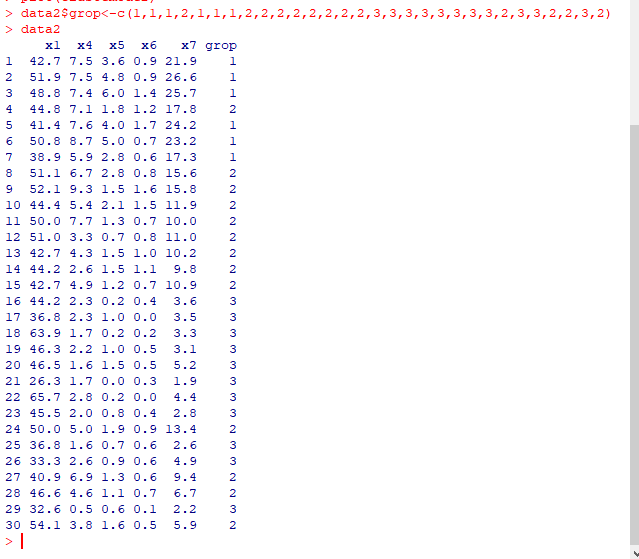
使用均值和方差对数据进行标准化，绘制层次聚类图



三类的观察值数目差距悬殊，所以因该排除一些聚类因子，以达 到更好的聚类效果。在此，考虑到前锋的特性，我们选择投篮命中率、篮板、 助攻、抢断和场均得分作为主要因子，再次进行聚类分析

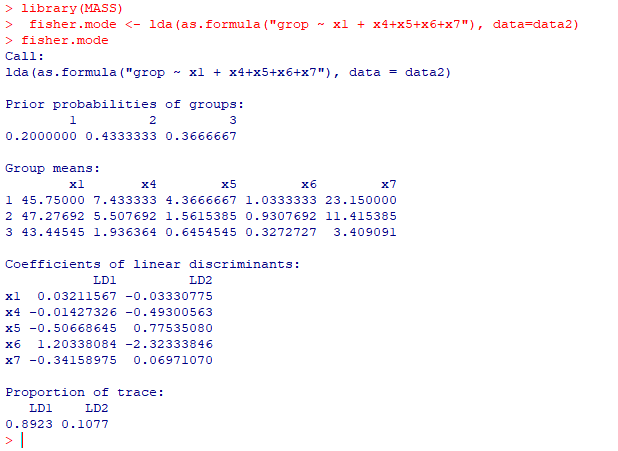


聚类结果较好，根据聚类将其标注为三类

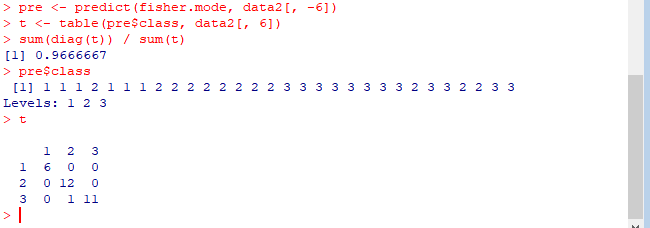


2.Fisher 判别法

调用 lda() 函数建立统计模型，其代码如下：



使用 Fisher 模型进行分类标签的预测和验证，其代码如下：



综上所述，使用fisher判别分析方法，对球星聚类的预测分类结果的正确率在 95% 以上。其中误判的样本主要分布在第二类。

由聚类分析的数据也可以得出，将三分和罚篮除掉之后，整体聚类效果更好。