实验报告

成绩：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 班级学号 | 202015116 | 姓名 | 梁祺若 | 实验日期 | 2022-9-14 | 仪器编号 |  |
| 实验名称 | 实验1 R的数据组织和整理 | | | | | | |

**1．实验题目**

R的数据组织和整理

**2．实验步骤**

本题目由两部分组成：（1）利用R软件做数据对象的创建和访问；（2）利用R软件自定义函数，对\*\*数据做\*\*\*分析。

1. 利用R软件做数据对象的创建和访问

* 读取数据

ConsumptData<-read.table("ConsumptData.txt",header=TRUE,sep="\t")

str(ConsumptData)

names(ConsumptData)

head(ConsumptData)

* 保存数据

write.table(ConsumptData,file="out.txt",sep="\t",row.names=FALSE,col.names=TRUE,quote=FALSE,append=FALSE,na="NA" )

* 数据框的创建与访问

ConsumptData<-read.table("ConsumptData.txt",header=TRUE, sep="\t")

str(ConsumptData)

ConsumptData$price

ConsumptData[["ID"]]

* 数组的创建与访问

a<-(1:120)

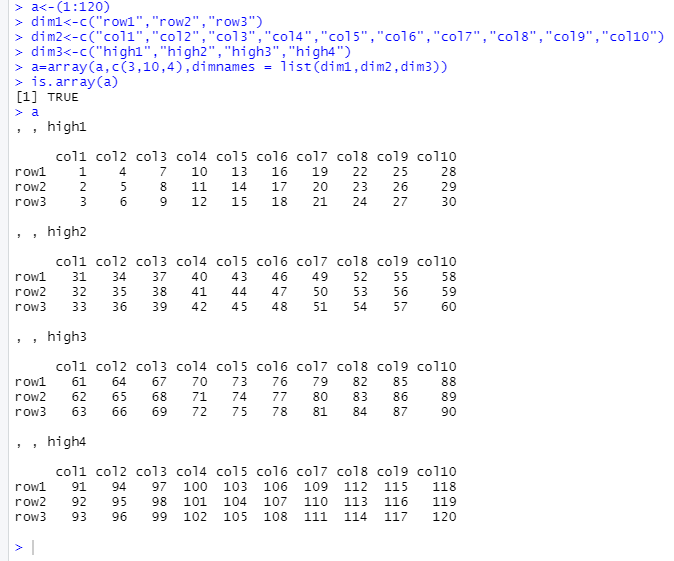
dim1<-c("row1","row2","row3")

dim2<-c("col1","col2","col3","col4","col5","col6","col7","col8","col9","col10")

dim3<-c("high1","high2","high3","high4")

a=array(a,c(3,10,4),dimnames = list(dim1,dim2,dim3))

is.array(a)



* 矩阵的创建和访问

MatA<-(1:42)

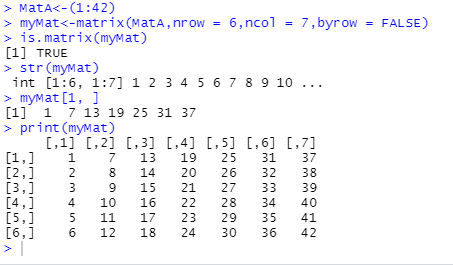
myMat<-matrix(MatA,nrow = 6,ncol = 7,byrow = FALSE)

is.matrix(myMat)

str(myMat)

myMat[1, ]

print(myMat)



* 列表的创建和访问

myMatb<-matrix(nrow=5,ncol=2)

lista<-c(1:10)

listb<-myMatb

listb[ ,1]=seq(from=6,to=30,by=6)

listb[ ,2]=seq(from=20,to=0,by=-5)

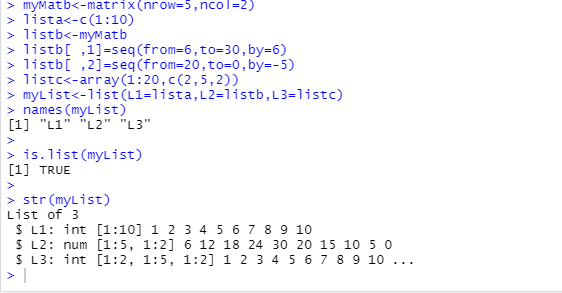
listc<-array(1:20,c(2,5,2))

myList<-list(L1=lista,L2=listb,L3=listc)

names(myList)

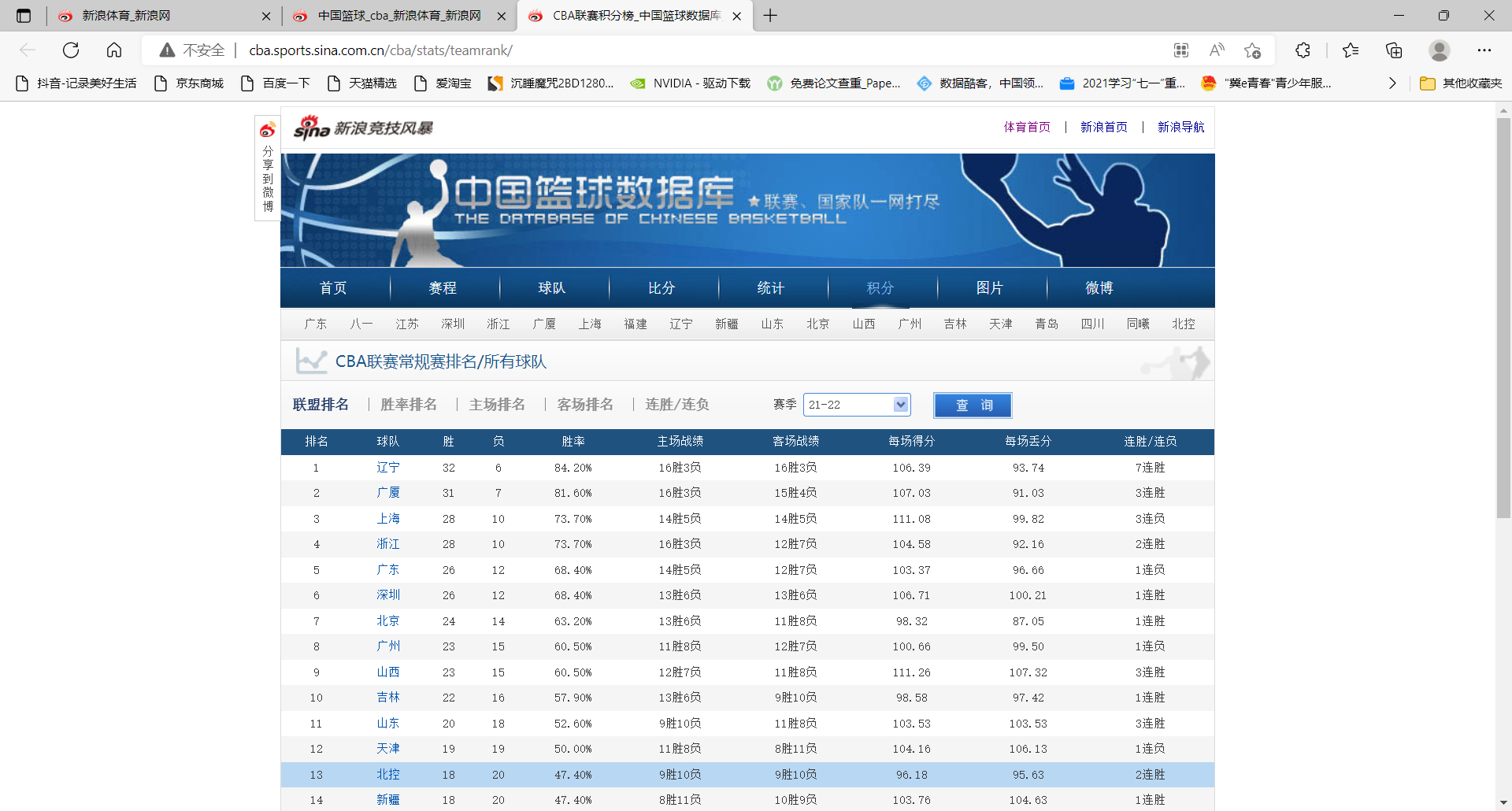
is.list(myList)

str(myList)



（2）利用R软件自定义\*\*函数，对\*\*数据做\*\*\*分析。

* 数据来源：[CBA联赛积分榜\_中国篮球数据库\_篮球-CBA\_新浪体育\_新浪网 (sina.com.cn)](http://cba.sports.sina.com.cn/cba/stats/teamrank/)



* 从中国篮球数据库抓取数据。

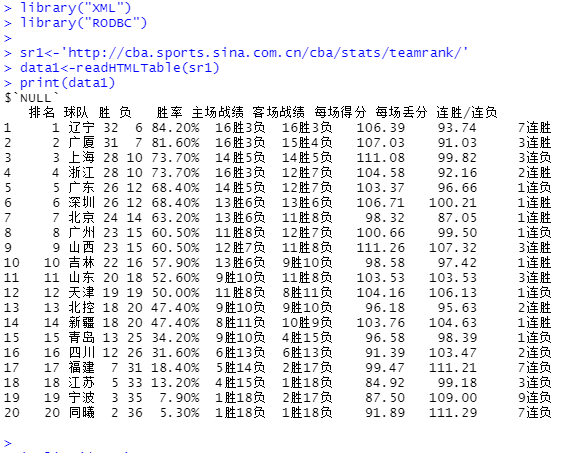
library("XML")

library("RODBC")

sr1<-'http://cba.sports.sina.com.cn/cba/stats/teamrank/'

data1<-readHTMLTable(sr1)

print(data1)



* 由于原先抓取的数据有表头，且原先的表头是字符型数据，故我们需要先将原先的表头“胜”删除，并将数据转化为整数型。最后进行排序。

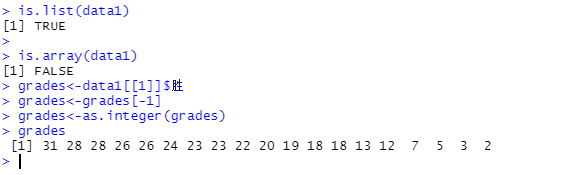
is.list(data1)

is.array(data1)

grades<-data1[[1]]$胜

grades<-grades[-1]

grades<-as.integer(grades)



* 自定义冒泡排序函数

bubblesort=function(grades){

n=length(grades)

for(i in 1:(n-1)){

for(j in (i+1):n){

if(grades[i]<=grades[j]){

t<-grades[i]

grades[i]<-grades[j]

grades[j]<-t

}

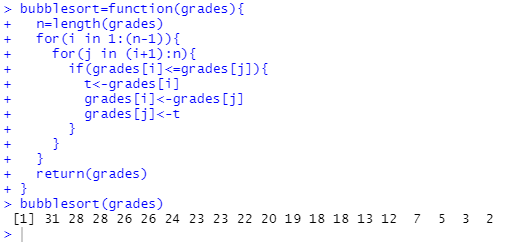
}

}

return(grades)

}

bubblesort(grades)



**3．实验结果**

* 在实验的第一部分，我们对数据进行了读取和保存。然后我们分别创建并访问了我们自行设计的数据框、数组、矩阵和列表。
* 在实验的第二部分，我们从网页中抓取中国篮球数据库中“2021-2022赛季CBA联赛常规赛排名/所有球队”中的数据，通过自定义冒泡排序函数对数据进行降序排序。