大数据菁英班第一次作业

**班级**：营销14 **学号**：2014311323 **姓名**：李婉荣

1. 简单阐述自己学习这门课过程中存在的问题(限100字内)
2. 知识储备少，学习过程中要多花时间
3. 在用help的过程中发现术语类的英文理解能力太差，即使是用help也要好长时间。
4. 说明两个以上使用R进行数据处理的优势?
5. 数据处理比较灵活，可以有选择性的输出自己想要的数据
6. 脚本和数据处理结果分开，便于观察
7. 作图形及图表比较方便。
8. 参考RStudio中的help帮助，查看numeric的介绍，定义一个变量Num，它由4个numeric的元素组成。

**num=numeric(length = 4)**

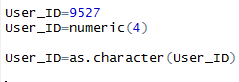


4.定义变量User-ID=9527，是一个数字类型，查看character的帮助说明，将变量User-ID转换为字符串格式(提示:as. character )

**User\_ID=9527**

**User\_ID=numeric(4)**

**User\_ID=as.character(User\_ID)**

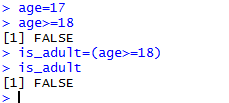


5已知表示小明年龄的变量为age，赋值为17(age=17)，根据条件判断符>或者<，判断小明是否成年，并将结果保存到变量is\_adult中。

**age=17**

**age>=18**

**is\_adult=（age>=18）**



6.以下是小学四年级应用题，请将题中的数字保存到变量中，并通过加减乘除等运算符得到结果，保存到一个变量中(提示:计算过程中可以定义其他变量，所有的变量名由自己定义，取名尽力有意义，便于理解意思)

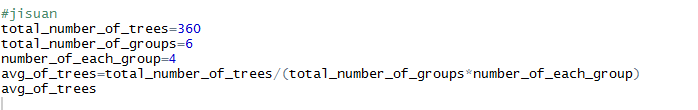
a.红星小学分成6个小组去浇树，每组有4人，一共浇树360棵，平均每人浇树多少棵?

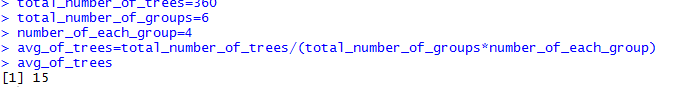
**total\_number\_of\_trees=360**

**total\_number\_of\_groups=6**

**number\_of\_each\_group=4**

**avg\_of\_trees=total\_number\_of\_trees/(total\_number\_of\_groups\*number\_of\_each\_group)**

**avg\_of\_trees** 



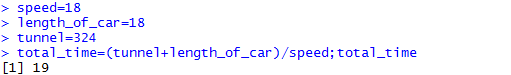
b.一辆汽车每秒行18米，车的长度是18米，隧道长324米，这辆汽车全部通过隧道要用多长时间?

**speed=18**

**length\_of\_car=18**

**tunnel=324**

**total\_time=(tunnel+length\_of\_car)/speed;total\_time**



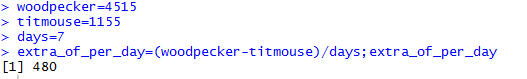
c.一只啄木鸟7天能吃4515只害虫，一只山雀一周能吃1155只害虫，啄木鸟平均每天比山雀多吃多少只害虫?

**woodpecker=4515**

**titmouse=1155**

**days=7**

**extra\_of\_per\_day=(woodpecker-titmouse)/days;extra\_of\_per\_day**



d.一个长方形长12米，宽8米，若长和宽各增加2米，面积增加了多少平方米?

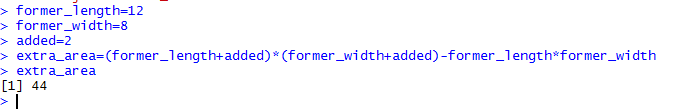
**former\_length=12**

**former\_width=8**

**added=2**

**extra\_area=(former\_length+added)\*(former\_width+added)-former\_length\*former\_width**

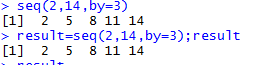
**extra\_area**



7.利用seq()函数，生成一列数字[2,5,8,11,14]，并将结果保存到变量result中。

**seq(2,14,by=3)**

**result=seq(2,14,by=3);result**



8.我们常用符号c()来得到向量，一个xiao’zu小组中有三个人，请将三个人zhangsan,lisi,wangwu的名字保存到向量类型的变量names中，假设三个人的身高分别是165, 175, 170厘米，存储到变量height中。

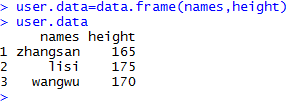
**names=c("zhangsan","lisi","wangwu")**

**height=c(165,175,170)**



9.根据第8题，构造数据框(dataframe)变量。user.data，由两列构成，第一列是names，第二列是height.

**user.data=data.frame(names,height)**

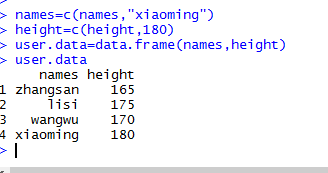


10.现在有一名新同学xiaoming加入了这个组，他身高是180，请利用符号c()将他的信息加入到变量names和height中，然后更新数据框变量。user.data中。

**names=c(names,"xiaoming")**

**height=c(height,180)**

**user.data=data.frame(names,height)**



(老师，这里更新数据表的时候不知道是不是还有别的简单的方法，我只是感觉再输一遍data.frame不会错^-^)

11.请结合for和if命令，判断哪些同学的身高不低于170厘米，并用print输出他们的名字。

**for(i in 1:4)**

**{**

**print(height[i])**

**}**

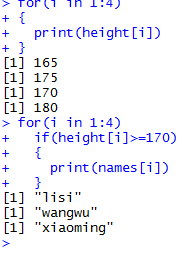
**for(i in 1:4)**

**if(height[i]>=170)**

**{**

**print(names[i])**

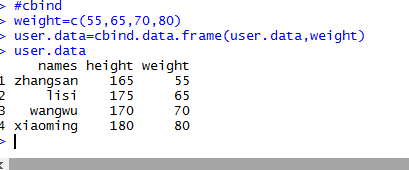
**}**



12.为了锻炼使用帮助的能力，自己通过help帮助，查看函数cbind()，实现如下功能:现在得到上面四个人的体重为(55,65,70,80)公斤，保存到变量weight中。然后使用cbind函数，将该列数据插入到问题10中得到的数据框。user.data。中，得到一个四行三列的数据框。

**weight=c(55,65,70,80)**

**user.data=cbind.data.frame(user.data,weight)**



(这道题也不知道做的对不对)