|  |
| --- |
| **成绩** |
|  |



中国农业大学

课程练习

（2021 -2022 学年 秋季学期）

**练习题目：** 期中练习

**课程名称：** SPSS统计分析实践

**任课教师：** 王 昊

**班　　级：**

**学　　号：**

**姓　　名：**

**作业排版格式**

全文：1.5倍行距

正文：宋体、五号。段落首行缩进2字符。

\* SPSS统计分析实践期中练习

\* 姓名：

\* 学号：

/\* 答题说明

请统一在题目后作答，形式如下列例题：

\*/

/\*

例题：

请打开系统数据auto，并以mpg为因变量，weight、weight的平方（命名为wtsq2）

和foreign为自变量，做线性回归分析。

回答问题：

(1) 根据模型结果，列出因变量与自变量拟合关系的方程式

\*/

/\*回答示例\*/

sysuse auto, clear

gen wtsq=weight^2

reg mpg weight wtsq foreign

/\*答：

(1) 回归方程式为

mpg= -0.0165729 weight + 1.59 \* 10^(-6)^ wtsq + 2.2035 foreign + 56.53884

\*/

/\* 提示：

（1）具体的小数点可酌情省略到两位或三位；格式自主安排，无需严格对齐；

（2）文字性的回答请用“注释”的格式，即前面加星号“\*”，或者首尾加"/\*"、"\*/"，

以免do文件报错、log文件不完整。

（3）文字或命令的长度不要超过右侧直线（位置Col: 51），以免影响打印。

\*/

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*练习开始\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

/\*

0.在桌面新建一个文件夹叫"exercise"。使用CD命令, 将该文件夹设定为工作文件夹。

使用log，将log文件名设定为自己的姓名（拼音）

\*/

\*1.打开系统自带数据auto（提示：使用命令sysuse)

/\*

2.使用命令，浏览、观察数据的基本情况（提示：使用命令describe)

回答问题：

(1) 数据是哪一年的automobile data？

(2) 样本量（又称“观测值”）是多少？

(3) 变量有多少个？

(4) 变量rep78的变量标签是什么？

(5) 数据里哪个变量有设定取值标签？

\*/

/\*

3.使用命令，查看数据的标签（提示：使用命令note)

回答问题：

(1) 显示的内容是什么？

\*/

/\*

4.仅使用一个命令，输出所有变量的基本描述统计，包括样本量、平均数、

标准差、最小值、最大值等（提示：使用命令summarize)

回答问题：

(1) make有多少个样本？为什么是这样？

(2) rep78有多少个样本？为什么比别的变量少？

(3) weight的最小值、最大值是什么？

\*/

/\*

5.使用codebook命令，分析变量make, rep78

回答问题：

(1) 从中能获取什么信息？

(2) 这些信息对你回答问题4(1)、4(2)是否有用？

\*/

/\*

6.使用list命令，列出变量rep78取值为缺失值的个案(提示：if missing(rep78))

\*/

/\*

7.使用summarize命令，输出price的详细描述统计（提示：,detail)

回答问题：

(1) price的中位值、平均值分别是多少？

(2) price的最小值、最大值是多少？

\*/

/\*

8.对foreign进行描述分析（提示：使用命令tab)

回答问题：

(1) 进口车、国产车的比例分别是多少？

(2) 取值标签Foreign对应的数字是多少？（提示：在逗号后加,nol，取消标签显示）

\*/

/\*

9.比较国产车、进口车的维修记录

回答问题：

(1) 在rep78等于3的样本当中，国产车、进口车的比例分别是多少？(提示：if rep78==3)

\*/

/\*

10.比较国产车、进口车的里程数(变量mpg)

回答问题：

(1) 国产车、进口车的平均里程数分别为多少？

(2) 国产车、进口车的平均里程数，是否存在显著差异？

（提示：使用ttest命令）

\*/

/\*

11.输出mpg weight的相关矩阵（提示：使用correlate命令）

回答问题：

(1) 两者的相关系数是多少？

(2) 这表明两者是什么关系？(提示：正相关，还是负相关？）

\*/

/\*

12.对国产车、进口车，分别输出mpg weight的相关矩阵

（提示：使用correlate命令；前面使用前缀bysort）

回答问题：

(1) mpg与weight的相关系数在国产车、进口车分别为多少？

(2) mpg与weight的相关关系，在国产车中，还是在进口车中更突出？

（提示：相关系数的绝对值更大）

\*/

/\*

13.对mpg与weight的分布情况做散点图（提示：scatter）

回答问题：

(1) 从图中可以看出什么信息？（提示：两者是正相关，还是负相关？

两者的关系是线性的，还是非线性的？等等）

\*/

/\*

14.对国产车、进口车，分别对mpg与weight的分布情况做散点图

(提示：使用scatter命令；在options当中，使用by）

回答问题：

(1) 从图中可以看出什么信息？（提示：如上）

\*/

/\*

15. 生成变量gp100m，其计算方式是100/mpg。给变量gp100m打上变量标签，内容为

"Gallons per 100 miles"。并对国产车、进口车，分别对gp100m与weight的分

布情况做散点图。

回答问题：

(1) 从图中可以看出什么信息？（提示：如上）

\*/

/\*

16.将rep78的缺失值，替换为99（提示，使用replace)。

\*/

/\*

17.利用系统变量\_n，生成变量id，作为样本编号。

\*/

/\*

18.使用命令，打开数据Chapter6.dta。对girl与enroll做两个变量的交互表。要求：

①提供两个变量关系的卡方；

②提供列变量的百分比；

③提供行变量的百分比；

④提供缺失变量的比例

回答问题：

(1) 从表中可以看出什么信息？分性别的入学情况有何差异？

试简要进行解释数据的结果。（几句就可以，字数不需要很多）

\*/

/\*

19.执行以下命令，并回答问题

\*/

use "Chapter7.dta", clear

sort age

by age:egen weightm=mean(weight)

scatter height age || line weightm age,yaxis( 1 2) xaxis(1 2) /\*

\*/ title(This is the Title) subtitle(This is the Subtitle) /\*

\*/ caption(This is the caption)note(This is the Note) /\*

\*/ ytitle(Height of Children, axis(2))ytitle(Average Weight of Children) /\*

\*/ xtitle(Age of Children) xtitle(Child Age, axis(2))ysize(4)xsize(5)

/\*

回答问题:

(1) 通过图片可以发现，数据当中的一个很大的问题是什么？

(2) 上述作图命令实际做了多少个图？分别是什么类型的图？哪个命令或符号，

是拼接这几个图的关键？

(3) 原始的命令很长，上述命令是如何实现“换行”的（提示：用什么符号？）

\*/

/\*

20.执行以下命令，并回答问题

\*/

gen sib0=sibs==0

gen sib1=sibs==1

gen sib2=sibs==2

graph bar (mean) sib0 sib1 sib2, over(urban, relabel(1 Rural 2 Urban)) ///

stack legend(row(1)) blabel(bar, position(base)format(%4.2f)) percentages ///

ysize(3) xsize(4) ytitle(Percent)over(girl, relabel(1 Boy 2 Girl))

/\*

回答问题:

(1) 从图中可以看出什么信息？试简要进行解释数据的结果。

\*/

/\*

21.执行以下命令

\*/

sysuse auto,clear

list if foreign==1

/\*

回答问题：

(1)这个命令的作用是什么？

(2)为什么"foreign==1"中要使用两个等号"=="？

\*/

/\*

22.继续上面的分析，使用系统数据auto，

使用命令，进行如下操作：

(1) 将变量make的变量名改为producer

(2) 将变量price的变量标签改为“价格”

(3) 将变量foreign的取值重新编码为变量名为type的新变量，其中Domestic取值为1、

Foreign取值为2，并将取值标签分别改为中文“国产”、“进口”

(4) 生成变量x，x的取值为变量price取值的平方根

\*/

/\*

23.试执行以下命令

\*/

log using "newlog.log",replace

/\*

回答问题：

(1)是否报错？如果报错，为什么？

(2)试用命令暂停log记录

(3)试用命令重新开始log记录

/\*

24.使用命令关闭log文件。保存do文件（提示：可以用鼠标点击）。将本do文件的全部内容，

复制保存到word文件中，并附上课程要求的封面。提交到腾讯课堂作业“期中练习”中。

\*/

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*练习结束\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*