

SAS Base Programming Using SAS 9.4 (改版后)

Haoxi Ma

2020-11-20

今天刚考完 based 考试，分享下经验，和女票都是这几天考的，因为女票先考，然后给我说了她考试的内容，我自己考的时候很多题都没怎么看就做了，基本都是一样的，最后考了 959hhhhh。下面凭记忆，分点说明：

Lab 板块：一共 8 个 lab, 6 个让你自己 code, 2 个改 code 输出结果即可

Lab1: 第一问，导入 xlsx 并查看 sheet name: 用 libname 导入 xlsx 表，就可以在 library 看到名字了（我们两都是一个 income sheet 一个 people sheet）。第二问，问第几个 obs 的某个变量值：在 proc print 用 (firstobs=n,obs=n) 即可

Lab2: 问的是某个数据集的 encoding 和 label: 用 proc contents 即可（记得是 encoding=window, label=employee）

Lab3: major 变量中有一部分使用 and 连接两个 major (computer science and statistics)，分成 major 和 second major，鉴于很难表述清楚不如直接上代码：data output; set input; index=find(major,"and"); if index then do; secondmajor=substr(major,index+4); major=substr(major,1,index-2); end; run;

然后让你求 major 和 secondmajor 中频数最大的那个，用 proc freq 就可以轻松解决（记得是 major=chemistry, secondmajor=economics）

Lab4: 要求把每个 postcode 中最高的 income 输出到一个数据集中：只需要 proc sort by postcode descending income 然后在 data 步里用 by postcode; 和 if first.postcode; 即可

Lab5: 要求 one-to-one merge 两个 datasets (要记得先 sort 再 merge) 然后创建个新变量 (记得是 merge by product 然后创建 $total=price*qty$)

Lab6: 要求按 age 分组然后输出到 young, mid, notasyoung 三个 dataset 里, 这里用 if not missing(age) and age LE 39 then output=young;else if 同理)

Lab7 and Lab8 是两个改错题, 可以先跑一下, 从 log 中就可以知道大部分错误, 但有些错误不显示 (如缺失值运算), 这里记不太清只能简单写写记得的错误: dataset 名称 younger 改成 young, proc means 中 total 改成 sum, if, then 语句的作用范围, 还有一个要你求当 $bonus > 7k$ 时 $total = sum(income, bonus)$ 这里尤其要注意有些 total 运行出来是缺失值, 是因为 $bonus < 7000$ 。

选择填空选择题: 70 题完全搞懂就够用了, 基本可以拿满分

填空题: 1. 宏的用法: 定义宏用 %let amount=5000, 引用宏用 &amount 即可

2. transpose 的用法: 给你一个原始数据集和输出结果问你 by, var, id 和 idlabel 都填什么 (强烈建议完全搞透 transpose, 我们选择填空各遇到了一题)

3. proc import 语句中让你填 dbms= 什么, 这个很简单, 看前面的数据类型, .xlsx 就写 dbms=xlsx, .xls 就写 dbms=excel

4. 是 70 题中的原题, 变成了填空 (记得是问数据集中有几个 obs, 因为 code 里没有 output, 所以就 $X=1$)

Conclusion 我们都有一些 SAS 基础, 就复习了 2 天, 但感觉没有 sas 基础, 1 周也完全拿的下来, 多看看 proc means, proc contents, proc freq 的用法, 再结合机经, 问题不大。最后, 推荐一本书给不是完全为了拿证的小伙伴, the little SAS book。Finally, good luck and stay safe.