

实习 3：SAS 基本过程的应用

一、对应课堂讲授内容：

第 07 讲 过程步语句和全程语句

第 08 讲 SAS 服务过程

第 09 讲 定量资料的统计描述

二、实习用文本文件描述

这里提供两份原始数据资料。其名称和内容如下：

1. 儿童成绩.txt：包含编号、姓名、班级、绘画、音乐、劳动共 6 个变量，采用“,”分隔，第一行是变量名。
2. 儿童身高和体重.txt：包含编号、性别、年龄、身高、体重共 5 个变量，采用空格分割，第一行是变量名。

注意：原始数据资料，只能采用 INFILE 读入，不得人为修改或者复制粘贴。

二、实习内容与步骤：

1. 合并儿童成绩和儿童身高和体重数据。三项要求：
 - ①最终建立永久数据集，数据集名称 sasuser.child1；
 - ②各个变量指定中文标签；
 - ③在 OUTPUT 中显示合并结果。
2. 利用 sasuser.child1 建立一个永久数据集 sasuser.child2。要求：
 - ①按姓名排序观测值；
 - ②不要显示观察值序号；
 - ③去除冗余的观察值；
 - ④按顺序显示下面几个变量：姓名、性别、年龄、身高、体重；

3. 利用 `sasuser.child1` 建立一个永久数据集 `sasuser.child3`。要求：

①先按班级排序，班级内按三门课程总分从高到低的顺序排列；

（用函数实现总分的计算；舍弃冗余观察值）

②需要保留的变量清单：班级、姓名、总分。

4. 利用 `sasuser.child2` 建立一个永久数据集 `sasuser.child4`。要求：

①年龄输出为属相；

②性别输出为“男”“女”；

③身高输出为“高个”“中等”“低个”；

④需要显示的变量清单：姓名、性别、年龄、身高。

5. 利用 `sasuser.child2` 数据集，采用 MEANS 过程，按性别分组计算身高、体重的基本统计量。两种方法各作一次，体会其中的异同：

①用 BY 分类作一次。

②用 CLASS 分组作一次。

6. 利用 `sasuser.child3` 数据集，采用 MEANS 过程，计算总分的基本统计量，要求把结果输出到新数据集中。

7. 采用 UNIVARIATE 分析学生总分的分布特征？绘制茎叶图和盒子图，并作出正态性检验。

請同學們獨立完成作業，可以商量借鑒但不得抄襲。