

数据编码

目录

SPSSAU 操作..... 1

数据编码是指将数字进行重新赋值，比如将数字 1 替换成数字 5，又或者比如年龄介于 18~30 岁之间编码为数字 1，又或者将数据按照平均值大小分为两组（低于平均值编码成数字 1 高于平均值编码成数字 2），SPSSAU 中支持下述功能：

- ✓ 单个或者批量标题进行数字编码；
- ✓ 单个或者批量标题进行范围编码；
- ✓ 单个或者批量标题进行自动分组；
- ✓ 数据编码时，支持覆盖原始数据（或者新生成新的标题）；
- ✓ 进行数据编码时，标题标签可自动进行同步。

SPSSAU 操作

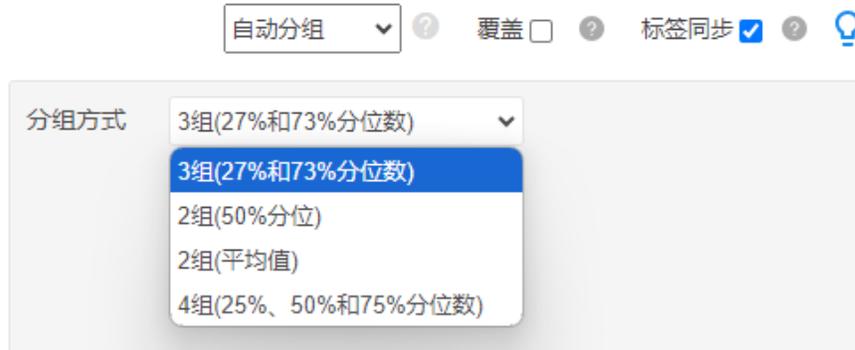


- ✓ 选中‘标题’，可配合 **ctrl** 或者 **shift** 键同时选择多个标题，批量数据编码；
- ✓ 数字编码时：将数字进行替换成自己需要的数字；
- ✓ 范围编码时：比如 18 岁以下设置为数字 1，18~30 岁设置为数字 2，30 岁以上设置为数字 3，操作如下图：



提示：范围编码时为闭区间，比如 0 到 18 时包括 0 和 18，并且以第 1 次出现为准，比如图中 18 出现 2 次，但 18 编码为数字 1。

✓ 自动分组时，如下图：



- ✧ SPSSAU 提供四种方式，分别是按 27%分位数和 73%分位数分成三组（分别是小于 27%分位数、27%分位数~73%分位数和大于 73%分位数），分别编码数字为 1、2 和 3；
- ✧ 按 50%分位数为界分成两组，分别编码数字为 1 和 2；
- ✧ 按平均值为界分成两组，分别编码数字为 1 和 2；
- ✧ 按 25%、50%和 75%分位数为界分成四组（分别是小于 25%分位数、25%分位数~50%分位数、50%分位数~75%分位数和大于 75%分位数），分别编码数字为 1、2 和 3；
- ✧ 自动分组仅方便用于快速分组处理，其可以完全被范围编码替换；自动分组会新生成标题进行标识。

✓ 针对两个参数即‘覆盖’和‘标签同步’，其说明如下：

参数	说明
覆盖	✓ 选中该参数后，数字编码和范围编码后，直接覆盖原始数据。如果不选中该参数，则新生成标题，标题名称类似为：“New_****”。
标签同步	✓ 选中该参数后，编码标题有标签信息，自动同步到新标题中，相当于比如原来数据中 1 表示男性，现在 1 编码为 2，此时 2 也表示男性。