

《新媒体数据运营与分析》

Excel (8): 统计分析-方差分析

教师：林志良

邮箱：linzhl@nfu.edu.cn

个人网站：www.zhilianglin.com

目录

- 简介
- 单因素方差分析
- 无重复双因素方差分析
- 可重复双因素方差分析

简介

- **方差分析作用：**
 - 比较多组（一般三组及以上）样本的均值 **（对比：t检验）**
 - **换种说法：**类别变量是否影响数值型变量
- **适用条件：**自变量-多类别变量；因变量-数值型变量
- **分类：**
 - 单因素方差分析
 - 无重复双因素方差分析
 - 可重复双因素方差分析
- **因素——**可理解为自变量。
 - **单因素**就是只有一个自变量；**双因素**就是两个自变量
- **无重复/可重复：**无重复是指两个自变量无关联；**可重复**是指两个自变量有关联。

单因素方差分析

- 例子：
 - 文传各**专业**（汉语言/新闻/网新）同学**社交媒体使用时间**差异
 - 专业是否影响社交媒体使用时间？
 - 不同**生源地**（珠三角/非珠三角省内/省外）同学**英语口语能力**差异
 - 生源地是否影响英语口语能力？

单因素方差分析

假设

- H_0 : 所有组的平均值都相等。/自变量不能影响因变量。
- H_1 : 不是所有组的平均值都相等的。/存在两组平均值不相等/
自变量影响因变量。

- 方差分析不仅考虑不同组别之间的平均值差异，也会考虑小组内部的标准差（方差）。

某品牌饮料的主要销售渠道是各种类型的超市。为了研究**不同区域**的超市是否对**销量**产生影响，从居民区、学校附近、写字楼各随机抽取10家，统计每日销售量如下图。

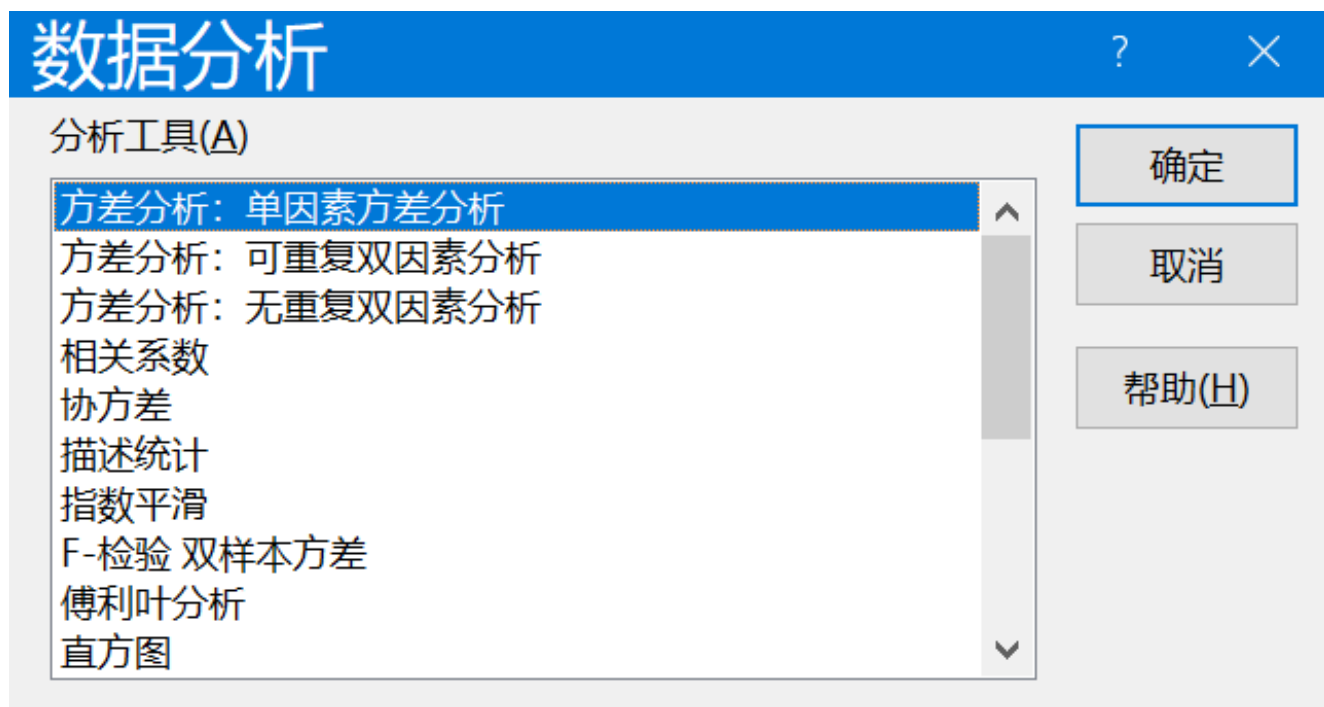
居民区	学校附近	写字楼
77	85	90
76	83	92
66	97	97
46	83	88
68	77	90
88	79	91
67	80	90
56	76	97
65	83	97
57	89	94

单因素方差分析

Excel操作



【工具】>【数据分析】>【方差分析：单因素方差分析】>选择数据范围



无重复双因素方差分析

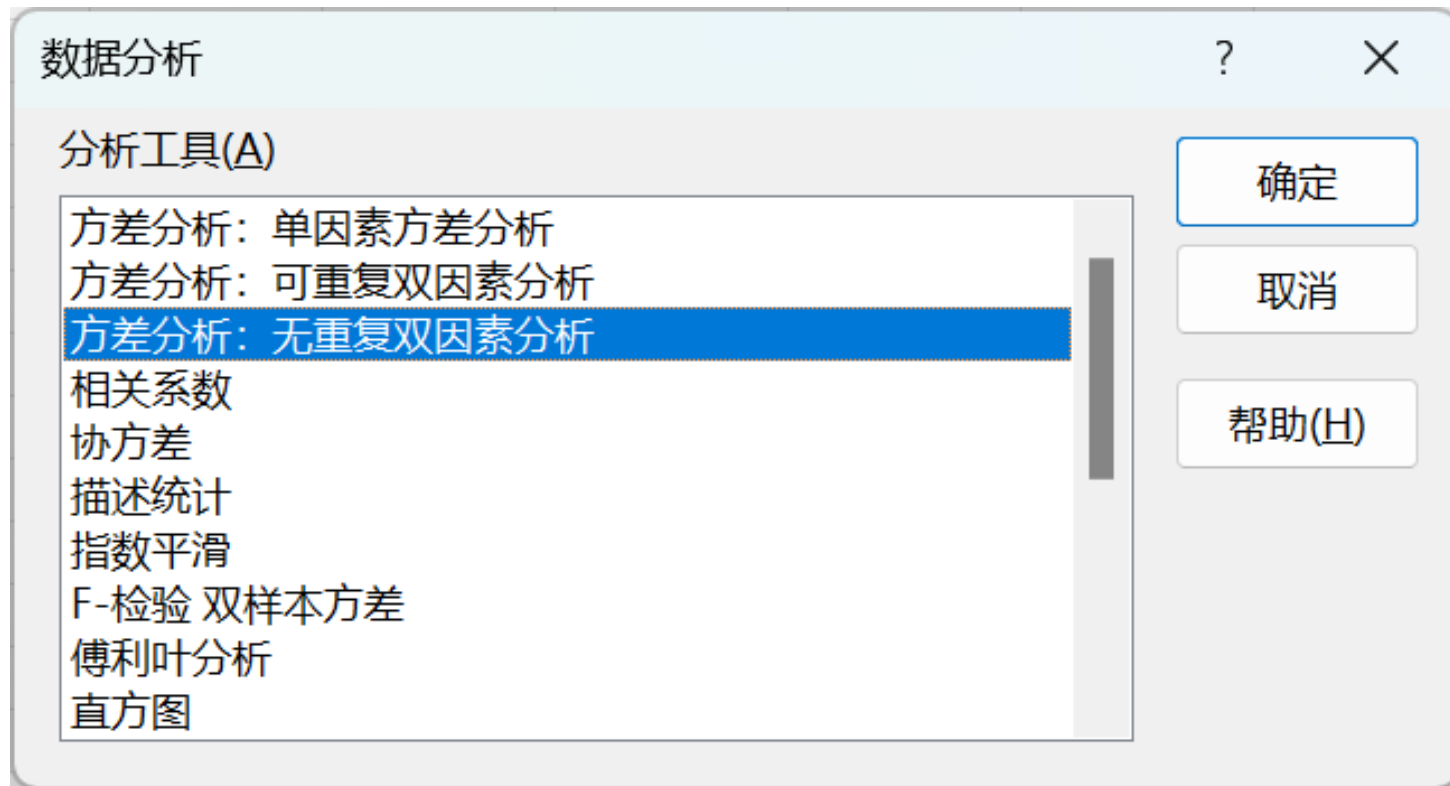
该品牌饮料一共有铁罐、玻璃瓶、塑料瓶三种不同包装。请问**包装**、**地段**对**销量**有没有影响？

	居民区	学校附近	写字楼
铁罐	289	352	485
玻璃瓶	310	401	457
塑料瓶	297	425	406

无重复双因素方差分析



【工具】>【数据分析】>【方差分析：无重复双因素方差分析】>选择数据范围



可重复双因素方差分析

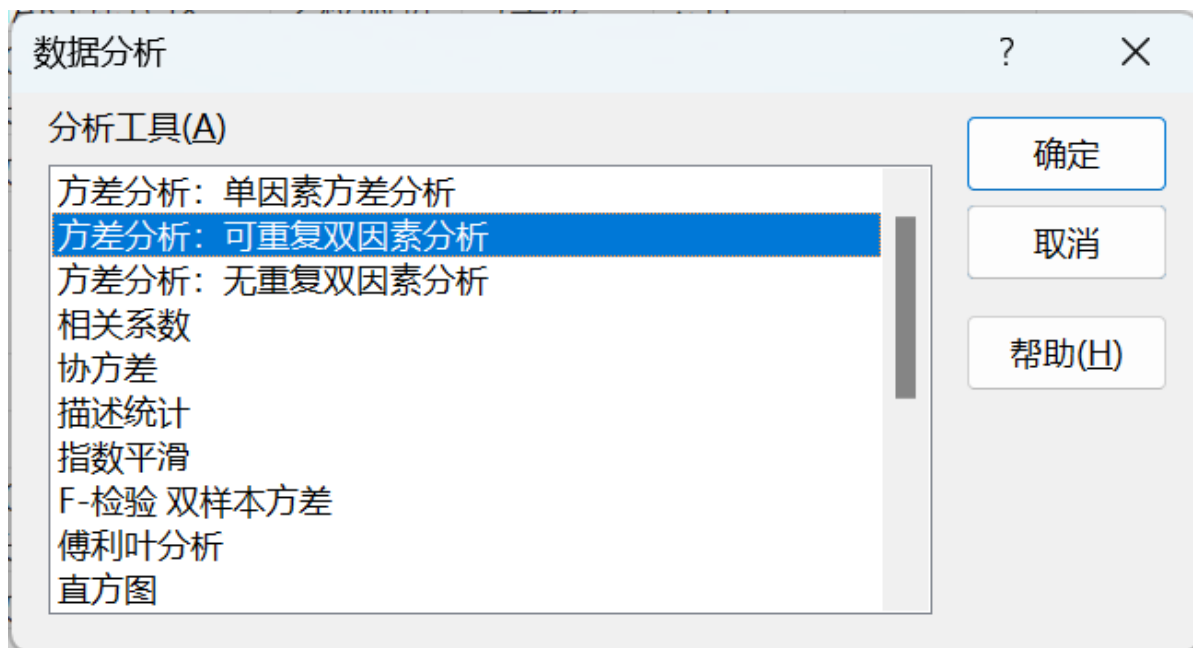
这些超市面积一共有80-120平方米和200-300米两种类型。现统计不同超市面积、区域该品牌饮料的销量。请问超市面积、区域能否影响销量？

面积	居民区	学校附近	写字楼
80-120平方米	65	95	110
	64	93	112
	54	107	117
	34	93	108
	56	87	110
200-300平方米	98	71	91
	77	72	90
	66	68	97
	75	75	97
	67	81	94

可重复双因素方差分析



【工具】>【数据分析】>【方差分析：可重复双因素方差分析】>选择数据范围



可重复双因素方差分析

行变量的观测值数量，这里两种不同面积的超市在不同区域都是有5个观测值

面积	居民区	学校附近	写字楼		方差分析：可重复双因素分析				
80-120平方米	65	95	110						
	64	93	112						
	54	107	117						
	34	93	108						
	56	87	110						
200-300平方米	98	71	91						
	77	72	90						
	66	68	97						
	75	75	97						
	67	81	94						

方差分析：可重复双因素分析

输入

输入区域(I):

每一样本的行数(R):

α (A):

输出选项

☒ 输出区域(O):

☐ 新工作表组(P):

☐ 新工作簿(W)

确定

取消

帮助(H)

参考资料

- [Lizongzhang的个人空间-合集 · Excel 数据分析实战](#)



谢谢！