### 分析流程 数据源： 附件.xlsx 算法配置： 算法： Yates校正卡方检验 分析结果： 暂无数据

### 分析步骤 1. 根据列联表的数据情况，分析Yates校正卡方检验是否呈现显著性（P<0.05，呈现显著性，拒绝原假设，则说明分类变量X与分类变量Y之间存在显著性差异）。 2. 若Yates校正卡方检验呈现显著性，可接着根据效应指标对差异进行深入量化分析。

### 详细结论

**输出结果1：Yates校正卡方检验结果**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 题目 | 名称 | 颜色 | | | | | | | | 合计 | X² | P |
| 浅绿 | 浅蓝 | 深绿 | 深蓝 | 紫 | 绿 | 蓝绿 | 黑 |
| 表面风化 | 无风化 | 2 | 8 | 3 | 2 | 2 | 1 | 6 | 0 | 24 | 6.287 | 0.507 |
| 风化 | 1 | 12 | 4 | 0 | 2 | 0 | 9 | 2 | 30 |
| 合计 | | 3 | 20 | 7 | 2 | 4 | 1 | 15 | 2 | 54 |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | | | | | | | |

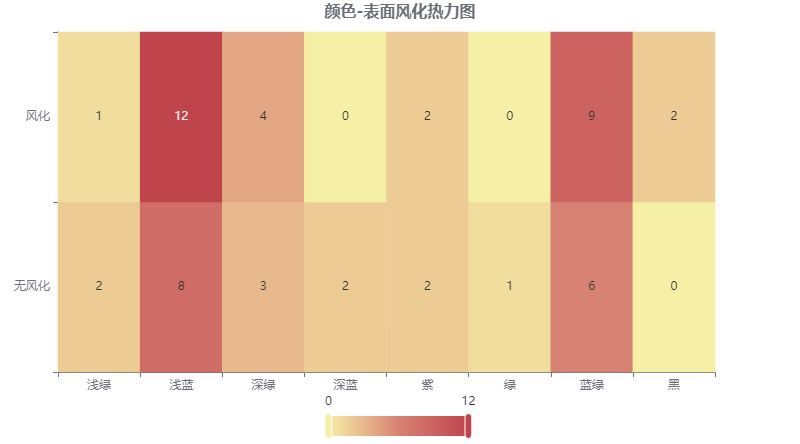
**图表说明：**

上表展示了Yates校正卡方检验的结果，包括数据的频数、卡方值、显著性P值。  
● 若P<0.05，呈现显著性，拒绝原假设，则说明分类变量X与分类变量Y之间存在显著性差异。  
● 若P>0.05，不呈现显著性，不拒绝原假设，不存在显著性差异。

**智能分析：**

Yates校正卡方检验分析的结果显示，显著性P值为0.507，水平上不呈现显著性，不能拒绝原假设，因此颜色和表面风化数据不存在显著性差异。

**输出结果2：交叉列联表热力图**



**图表说明：**

上图展示了热力图的形式展示了交叉列联表的值，主要通过颜色深浅去表示值的大小。

**输出结果3：效应量化分析**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名/分析项 | Phi | Crammer‘s V | 列联系数 | lambda |
| 表面风化-颜色 | 0.341 | 0.341 | 0.323 | 0.167 |

**图表说明：**

上表展示了效应量化分析的结果，包括phi、Crammer's V、列联系数、lambda ，用于分析样本的相关程度。  
1. 当呈现出显著性差异（前提），结合分析效应量指标对差异性进行量化分析。  
2. 效应量化指标反映的是变量之间的相关程度。  
3. 根据交叉类型的不同，可以选用不同的效应量指标（交叉类型表示：交叉表横向格子数×纵向格子数）。  
4. phi系数：phi相关系数的大小，表示两样本之间的关联程度。当phi系数小于0.3时，表示相关较弱；当phi系数大于0.6时，表示相关较强（用于2×2交叉类型表）。  
5. Cramer's V：与phi系数作用相似，但Cramer's V系数的作用范围较广。当两个变量相互独立时，V=0，当数据中只有2个二分类变量时，Cramer's V系数的结果与phi相同（若m≠n，建议使用Cramer's V ）。  
6. 列联系数：简称C系数，用于3×3或4×4交叉表，但其受行列数的影响，随着R和C 的增大而增大。因此根据不同的行列和计算的列联系数不便于比较，除非两个列联表中行数和列数一致。  
7. lambda：用于反应自变量对因变量的预测效果，一般情况下，其值为1时表示自变量预测因变量效果较好，为0时表明自变量预测因变量较差（X或Y有定序数据时，建议使用lambda）。

**智能分析：**

效应量化分析的结果显示，分析项：表面风化Cramer's V值为0.341，因此颜色和表面风化的差异程度为中等程度差异。

### 参考文献 [1] Scientific Platform Serving for Statistics Professional 2021. SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com. [2] 杨倩. 基于信息恒等式的属性数据分析[D]. 厦门大学, 2012.