

题目

摘要

摘要

关键字：关 键 字

一、问题重述与问题分析

1.1 问题重述

玻璃是早期丝绸之路的重要商品，从西亚和埃及地区传入我国后，古人因地制宜、就地取材制作出外观相似但化学成分不同的玻璃制品。玻璃炼制时需要助熔剂，不同的助熔剂会导致玻璃的化学成分不同，例如铅钡玻璃和钾玻璃分别以铅矿石和草木灰作为助熔剂。同时，古代玻璃易与外界环境进行元素交换而导致化学成分比例发生改变。

现有提供高钾玻璃和铅钡玻璃的相关数据，完成下列问题：

1. 分析玻璃文物表面风化和玻璃类型、纹饰和颜色的关系；对不同的玻璃类型分析表面风化与否和玻璃化学成分含量之间的统计规律；根据风化点的检测数据预测风化前的化学成分含量；
2. 根据附件数据分析铅钡玻璃和高钾玻璃的分类规律；对不同的玻璃类型，选择适当的化学成分划分亚类，并给出具体的方法和结果，并对分类结果进行合理性和敏感性分析；
3. 针对附件表单 3 的化学成分，鉴别其所属的类型，同样需要对结果进行敏感性分析；
4. 分析不同类别玻璃文物的化学成分之间的关系以及比较不同类别化学成分关系的差异性。

1.2 问题分析

1.2.1 数据预处理

1.2.2 问题一的分析

问题一要求分析附件表单 1 中文物表面风化与玻璃类型、颜色和纹饰之间的关系，随后分析表单 2 数据，得出风化与化学成分含量的统计规律，最后依据表单 2 风化点的检测数据，预测风化前的化学成分含量。本文首先对附件表单 1 的数据进行可视化，作出以玻璃类型、纹饰和颜色为自变量，是否风化为因变量的三维气泡图，再对高钾玻璃和铅钡玻璃分别作风化与纹饰和颜色的二维散点图得到四者之间的定性关系。未完待续

1.2.3 问题二的分析

问题二要求分析两类玻璃的分类规律，之后选取合适的化学成分划分亚类并给出方法和具体结果，并对亚类划分结果做合理性和敏感性分析。本文首先作出不同玻璃文物的不同化学成分含量的散点图，得出初步的高钾玻璃和铅钡玻璃的分类规律。未完待续

1.2.4 问题三的分析

问题三要求预测附件表单 3 中文物的类型，并对结果做敏感度分析。未完待续

1.2.5 问题四的分析

问题四要求分析同类玻璃化学成分的联系以及不同类别化学成分联系的差异性。未完待续

二、模型假设与符号说明

2.1 符号说明

三、数据预处理

3.1 数据清洗

3.1.1 处理缺失值

对于表单 1 的数据，我们首先找出缺失值，缺失值数据如表1所示

表 1 表单 1 缺失值数据

文物编号	纹饰	类型	颜色	表面风化
19	A	铅钡	NaN	风化
40	C	铅钡	NaN	风化
48	A	铅钡	NaN	风化
58	C	铅钡	NaN	风化

可以观察到，表单 1 的数据仅存在颜色属性的缺失，缺失数据量为 4，缺失量较少，可以忽略。

3.1.2 处理异常值

3.2 数据规约

由于表单 2 中不同的化学成分含量相差较大，故考虑首先对各列化学成分数据进行标准化处理，使处理后的数据更加能体现该化学成分含量的相对大小，使数据更加直观且具有可比性。此处使用了最小——最大规范化方法 (minmax-scale)，方法公式如下：

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}}$$

参考文献

- [1] 参考文献

附录 A 附录

```
import math

def hix(x):
    if x > 0:
        return 1+1=2
    else:
        return 0
```