# 第五次实践-聚类算法

**按组评分（撰写作业报告，每组提交一份）提交作业名称：”组号+组内同学姓名.pdf/.docx/.doc”**

**组内同学自行分工，在报告中务必写明具体分工！！！**

**按照作业报告模版撰写（包括：数据介绍、模型/方法介绍、模型训练、模型预测、实验结果与分析、实验总结等部分）**

#### 数据集（实验数据集可以自行选择，选定数据集后要在报告对数据集进行介绍）

(1) 鸢尾花数据集：这是一个非常经典的数据集，常被用于测试聚类算法。它包含了 150 个样本，每个样本有 4 个特征，代表了鸢尾花的不同测量维度。

https://archive.ics.uci.edu/dataset/53/iris

(2) Wine 数据集：这个数据集包含了 178 个样本，每个样本有 13 个特征，代表了不同种类的葡萄酒的化学成分。https://archive.ics.uci.edu/dataset/109/wine

(3) Wholesale customers 数据集：这个数据集包含了 440 个样本，每个样本有 8 个特征，记录了客户的年度消费习惯。

https://archive.ics.uci.edu/dataset/292/wholesale+customers

#### 实验要求

理解K-means聚类算法的基本思想和工作原理

掌握使用Python的sklearn库实现K-Means算法的方法

对实验数据进行预处理

使用K-Means算法对数据进行聚类

对聚类结果进行可视化展示

分析和讨论实验结果

#### 实验说明

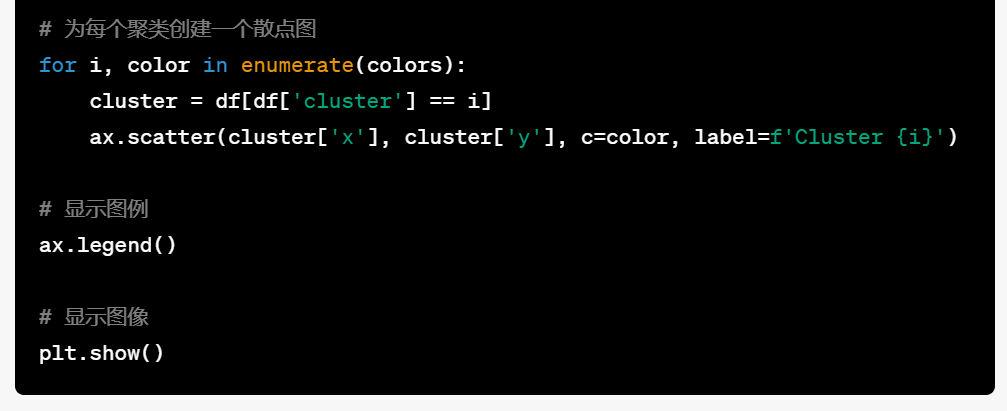
本实验主要是通过Python的sklearn库实现K-Means算法，并对一组数据进行聚类。通过对聚类结果的分析和讨论，深入理解K-Means聚类算法的应用。

#### 实验步骤









#### 结果分析与讨论

在实验报告的这部分，你需要分析并讨论你的结果。你可以根据你的结果来提出观察结果，例如，某些聚类似乎有共享特征。你也可以讨论K-Means的限制，例如，需要预先知道聚类的数量，以及对初始值的敏感性。

#### 结论

在这个实验中，我们使用了K-Means聚类算法对一组数据进行了分类，并将结果进行了可视化。从结果图像中，我们可以清晰地看到不同的聚类。这个实验使我们更好地理解了K-Means聚类算法的工作原理和应用，以及其在实际问题中的应用和限制。