# 概述

## 聚类的方法

聚类分析本身有很多办法，常用的有层次聚类、基于划分的聚类、基于密度的聚类：

(1) 层次聚类

(2) 基于划分的聚类(k-means)

(3) 基于密度的聚类

## 聚类的逻辑

不同聚类方法在算法方面差异很大，但是逻辑是相同的，即把类似的样本归为一组，以层次聚类方法为例。

层次聚类分析的基本逻辑为：

(1) 生成初始聚类数据，即特征数据集。

(2) 根据特征，计算两两观测之间的距离，生成距离矩阵。

(3) 将距离较近的观测点聚为一类，最终达到组间的距离最大化，组内的距离最小化。