**实验二 认识数据**

**一、实验名称**：

认识数据

**二、实验目的：**

掌握数据属性类型，掌握数据的基本统计描述以及数据的邻近性度量

**三、实验要求**

1、掌握数据分析手段和方法

2、实现数据的可视化

**四、实验平台**

计算机、Anaconda

**五、实验数据**

Iris.csv

**六、通过实验,要求同学在实验中解决以下问题，并递交完整的实验报告**

1、使用Numpy，分析Iris.csv数据集中每个属性的中心趋势以及散布(包括均值、众数、中位数、极差、四分数、四分位极差，合图)。

2、 通过直方图分析Iris各个属性的分布情况，选择其中的两个或多个属性，分析这两个属性之间的关系。

3、选择20个样本，计算Iris数据集中的不同样本之间的距离。

4、将3中的20个样本规格化后在计算不同样本间的距离，对比规格化前后距离的变化情况