IRIS鸢尾花数据集

参考链接：<https://www.kesci.com/mw/dataset/58a942bc7159a710d916af11/content>

**背景介绍**

Iris也称鸢尾花卉数据集，是一类多重变量分析的数据集。数据集包含150个数据集，分为3类，每类50个数据，每个数据包含4个属性。可通过花萼长度，花萼宽度，花瓣长度，花瓣宽度4个属性预测鸢尾花卉属于（Setosa，Versicolour，Virginica）三个种类中的哪一类。

如何在线使用数据集

创建项目后：  
Python用户，输入ls ../input/iris/ 查看数据路径  
R用户，输入list.files("../input/iris/") 查看数据路径  
使用相关包读取数据

**数据描述**

**数据字典**

| **变量名** | **sepal\_length** | **sepal\_width** | **petal\_length** | **petal\_width** | **species** |
| --- | --- | --- | --- | --- | --- |
| **变量解释** | 花萼长度（单位cm） | 花萼宽度（单位cm） | 花瓣长度（单位cm） | 花瓣宽度（单位cm） | 种类 |
| **数据类型** | numeric | numeric | numeric | numeric | categorical |

**数据预览**

iris.csv 数据集大小：4k

| **sepal\_length** | **sepal\_width** | **petal\_length** | **petal\_width** | **species** |
| --- | --- | --- | --- | --- |
| 5.1 | 3.5 | 1.4 | 0.2 | setosa |
| 4.9 | 3 | 1.4 | 0.2 | setosa |
| 4.7 | 3.2 | 1.3 | 0.2 | setosa |
| 4.6 | 3.1 | 1.5 | 0.2 | setosa |
| 5 | 3.6 | 1.4 | 0.2 | setosa |
| 5.4 | 3.9 | 1.7 | 0.4 | setosa |
| 4.6 | 3.4 | 1.4 | 0.3 | setosa |
| 5 | 3.4 | 1.5 | 0.2 | setosa |
| 4.4 | 2.9 | 1.4 | 0.2 | setosa |
| 4.9 | 3.1 | 1.5 | 0.1 | setosa |
| … | … | … | … | … |

**10** of of150 rows, 5 columns