Python 数据科学 *速查表*

Python 基础

变量与数据类型

变量赋值

```
>>> x=5
>>> x
 5
```

变量计算

```
>>> x+2
 7
>>> x-2
 3
>>> x*2
 10
>>> x**2
 25
>>> x%2
 1
>>> x/float(2)
 2.5
```

加 减 乘 幂 取余 除

类型与类型转换

str()	'5', '3.45', 'True'	转为字符串
int()	5, 3, 1	转为整数
float()	5.0, 1.0	转为浮点数
bool()	True, True, True	转为布尔值

调用帮助

```
>>> help(str)
```

列表 参阅 Numpy 数组

```
>>> a = 'is'
>>> b = 'nice'
>>> my_list = ['my', 'list', a, b]
>>> my_list2 = [[4,5,6,7], [3,4,5,6]]
```

选择列表元素 索引始于0

子集

```
>>> my_list[1]
>>> my_list[-3]
```

选择索引1对应的值
选择倒数第3个索引对应的值

切片

```
>>> my_list[1:3]
>>> my_list[1:]
>>> my_list[:3]
>>> my_list[:]
```

选取索引1和2对应的值
选取索引0之后对应的值
选取索引3之前对应的值
复制列表

子集列表的列表

```
>>> my_list2[1][0]
>>> my_list2[1][:2]
```

my_list[list][itemOfList]

列表操作

```
>>> my_list + my_list
['my', 'list', 'is', 'nice', 'my', 'list', 'is', 'nice']
>>> my_list * 2
['my', 'list', 'is', 'nice', 'my', 'list', 'is', 'nice']
>>> my_list2 > 4
True
```

列表方法

>>> my_list.index(a)	获取某值的索引
>>> my_list.count(a)	统计某值出现的次数
>>> my_list.append('!')	追加某值
>>> my_list.remove('!')	移除某值
>>> del(my_list[0:1])	反转列表
>>> my_list.reverse()	添加某值
>>> my_list.extend('!')	移除某值
>>> my_list.pop(-1)	插入某值
>>> my_list.insert(0,'!')	列表排序
>>> my_list.sort()	

字符串

```
>>> my_string = 'thisStringIsAwesome'
>>> my_string
'thisStringIsAwesome'
```

字符串运算

```
>>> my_string * 2
 'thisStringIsAwesomethisStringIsAwesome'
>>> my_string + 'Innit'
 'thisStringIsAwesomeInnit'
>>> 'm' in my_string
 True
```

字符串操作 索引始于0

```
>>> my_string[3]
>>> my_string[4:9]
```

字符串方法

>>> my_string.upper()	设为大写字符
>>> my_string.lower()	设为小写字符
>>> my_string.count('w')	统计某字符出现的次数
>>> my_string.replace('e', 'i')	替换字符
>>> my_string.strip()	清除空格

Python库

导入库

```
>>> import numpy
>>> import numpy as np
```

导入指定功能

```
>>> from math import pi
```

pandas 数据分析 | learn 机器学习 | NumPy 科学计算 | matplotlib 二维视图

安装 Python

ANACONDA. Python 首选开源数据科学平台

spyder Anaconda 内置的免费IDE

jupyter 创建包含代码、可视图与文本的文档

Numpy 数组 参阅 列表

```
>>> my_list = [1, 2, 3, 4]
>>> my_array = np.array(my_list)
>>> my_2darray = np.array([[1,2,3],[4,5,6]])
```

选取 Numpy 数组的值 索引始于0

子集

```
>>> my_array[1]
  2
```

选择索引1对应的值

切片

```
>>> my_array[0:2]
  array([1, 2])
```

选择索引0和1对应的值

二维 Numpy 数组的子集

```
>>> my_2darray[:,0]
  array([1, 4])
```

my_2darray[rows, columns]

Numpy 数组运算

```
>>> my_array > 3
 array([False, False, False,  True], dtype=bool)
>>> my_array * 2
 array([2, 4, 6, 8])
>>> my_array + np.array([5, 6, 7, 8])
 array([6, 8, 10, 12])
```

Numpy 数组函数

>>> my_array.shape	获取数组形状
>>> np.append(other_array)	追加数据
>>> np.insert(my_array, 1, 5)	插入数据
>>> np.delete(my_array,[1])	删除数据
>>> np.mean(my_array)	平均值
>>> np.median(my_array)	中位数
>>> my_array.corrcoef()	相关系数
>>> np.std(my_array)	标准差

原文作者 DataCamp
Learn Python for Data Science Interactively