

Data Analysis

Velen Kong

2019 年 8 月 9 日

摘要

基于Python的数据分析入门笔记，希望能自学掌握基本的数据分析能力，毕竟是统计学的基础之一。

内容安排主要参考

(美) Wes McKinney 著. Python for Data Analysis:2nd Edition [M]
USA:O'Reilly Media 2017

该书作者同时也是pandas库的作者，同时借用其在GitHub上的数据资料。

1 第三方库与Python入门

1.1 IPython & Jupyter

用到的库：ipython, jupyter。

jupyter确实好用，主要扩展了`Tab`, `?`, `*`的功能。

还有一些快捷键仅限IPython使用。

另外就是一些常用的Magic Command，比如`%run`, `%load`,
`%debug`, `%matplotlib inline`

1.2 Python基础

基本的条件和循环。

#注释。

所有变量都是对象(object)。

基本的函数调用，利用函数批量处理。

变量赋值类似引用。

动态类型，利用`type()`进行类型检查。

2 Numpy库