

致敬大师 致敬未来的你



PRODUCT INTROD

# 图灵Python手册

图灵Python学院出品

我应该学  
Python么？

怎么做好  
编程准备？

零基础  
如何入门？

跳转到第3页

跳转到第11

跳转到第25

# 第一章

# 为什么要学习Python



# Why Python?

那些最好的程序员不是为了得到更高的薪水或者得到公众的仰慕而编程，他们只是觉得这是-件有趣的事情。

——Linux 之父 Linus Torvalds

作为一个实用主义的学习者，最关心的问题-定是「我为什么要选择学 Python，学会之后我可以用来做什么2(

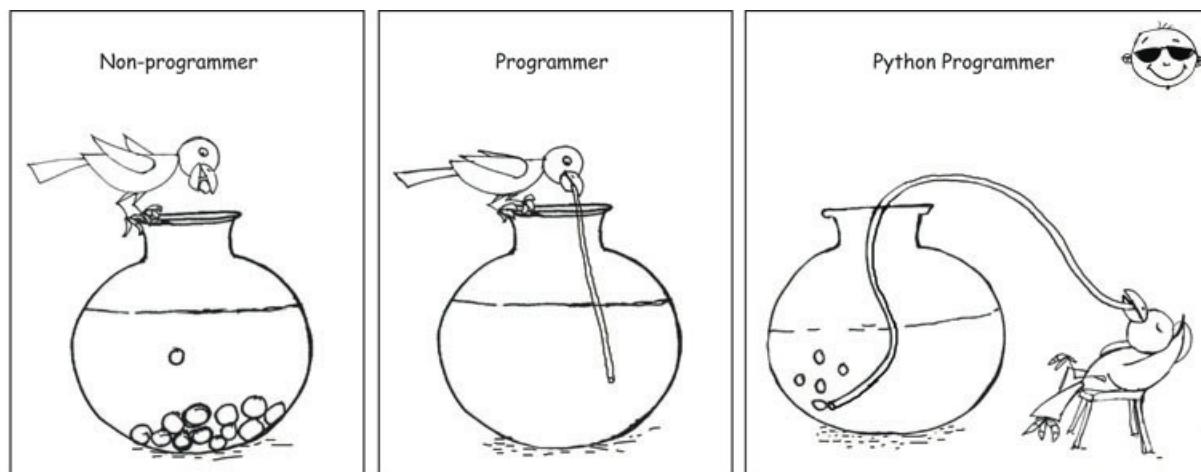
首先，对于初学者来说，比起其他编程语言，Python 更容易上手。

Python 的设计哲学是优雅、明确、简单。在官方的 The Zen of Python( Python 之禅) 中，有这样-句话，

There should be one-- and preferably onIy one --obvious way to do it.

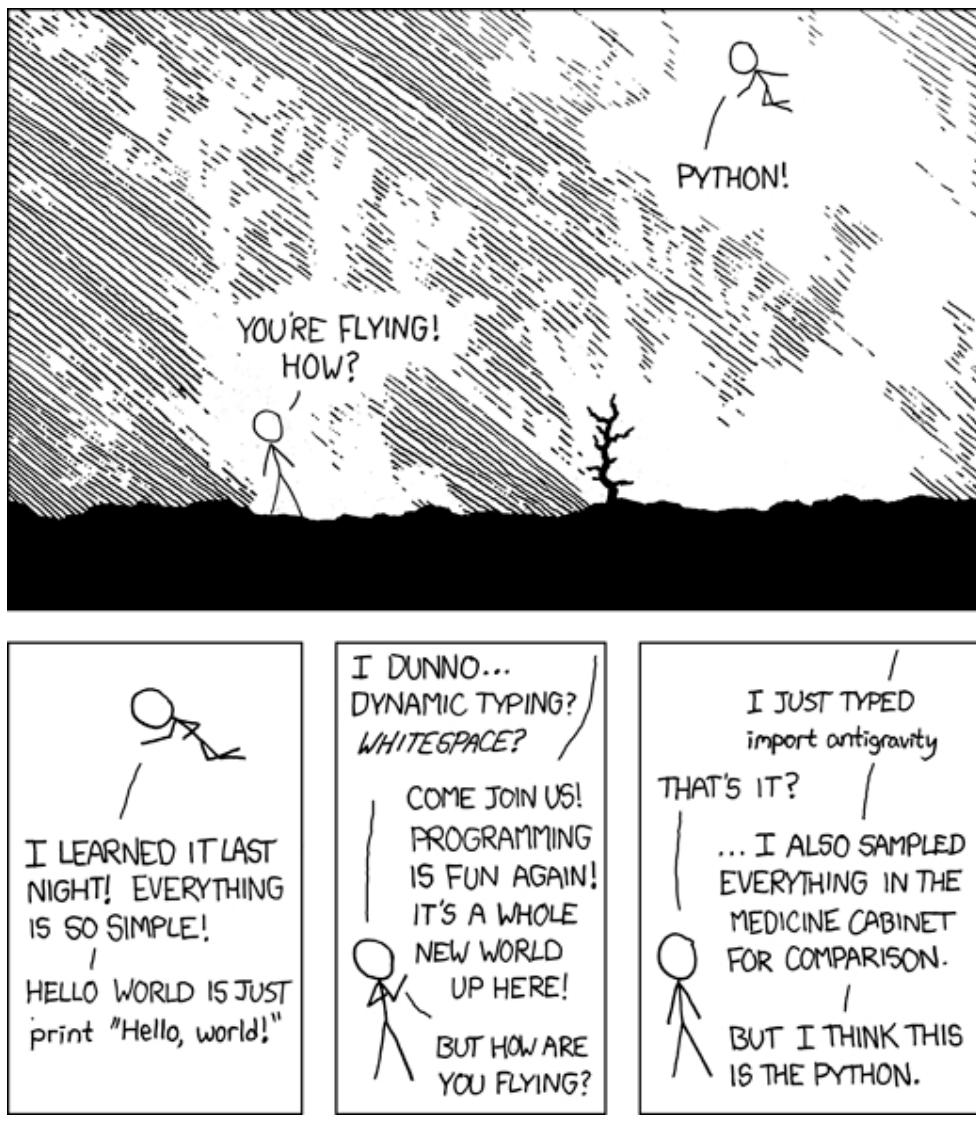
Python 追求的是找到最好的解决方案。相比之下，其他语言追求的是多种解决方案。

如果你试着读-段写的不错的 Python 代码，会发现像是在读英语-样。这也是 Python 的最大优点，它使你能够专注于解决问题而不是去搞明白语言本身。



THE TH1RSTY Python PROGRAMMER, from Pycot

其次，Python 功能强大，很多你本来应该操心的事情，Python 都替你考虑到了。当你用 Python 语言编写程序的时候，你不需要考虑如何管理你的程序使用的内存之类的底层细节。并且，Python 有很丰富的库，其中有官方的，也有第三方开发的，你想做的功能模块很有可能已经有人写好了，你只需要调用，不需要重新发明轮子。这就像是拥有了智能手机，可以任意安装需要的 app。



Python, from xkcd

这幅漫画形容了 Python 的库有多强大，导入一个反重力库就可以飞起来了。

第三, Python 能做的事情有许多。

在职场中, 使用 Python 工作的主要就是这样几类人:

- 网站后端程序员:使用 Python 搭建网站、后台服务会比较容易维护, 当需要增加新功能, 用 Python 可以比较容易的实现。但如果使用 php , 往往需要重写代码。不少知名网站都使用了 Python 开发, 比如:



Gmail



Youtube



Reddit



Spotify



知乎



豆瓣

- 自动化运维:越来越多的运维开始倾向于自动化, 批量处理大量的运维任务。Python 在系统管理上的优势在于强大的开发能力和完整的工具链。
- 数据分析师:Python 能快速开发的特性可以让你迅速验证你的想法, 而不是把时间浪费在程序本身上, 并且有丰富的第三方库的支持, 也能帮你节省时间。
- 游戏开发者:一般是作为游戏脚本内嵌在游戏中, 这样做的好处是即可以利用游戏引擎的高性能, 又可以受益于脚本化开发的优点。只需要修改脚本内容就可以调整游戏内容, 不需要重新编译游戏, 特别方便。
- 自动化测试:对于测试来说, 要掌握 Script 的特性, 会在设计脚本中, 有更好的效果。Python 是目前比较流行的 Script。

如果你是一名业余开发者，只是想在资源少的情况下快速做出自己想要的东西、自动化的解决生活中的问题，那么 Python 可以帮你做到这几类事情

### ● 网站的开发

借助功能丰富的框架 django, flask, 丰富的设计模板 bootstrap，你可以快速搭建自己的网站，还可以做到移动端自适应。



私教计划的竞品分析器项目

动态效果可以查看视频:<http://v.qq.com/boke/page/k/5/g/k0171oqo95g.html>

- 用爬虫实现数据挖掘、批量处理

爬虫的本质是模仿人去获取网页数据，当你需要获取大批量数据或是不停地获取的时候，Python 可以快速做到，节省你的重复劳动时间。比如：微博私信机器人、批量下载美剧、运行投资策略、刷便宜机票、爬合适房源、系统管理员的脚本任务等等。



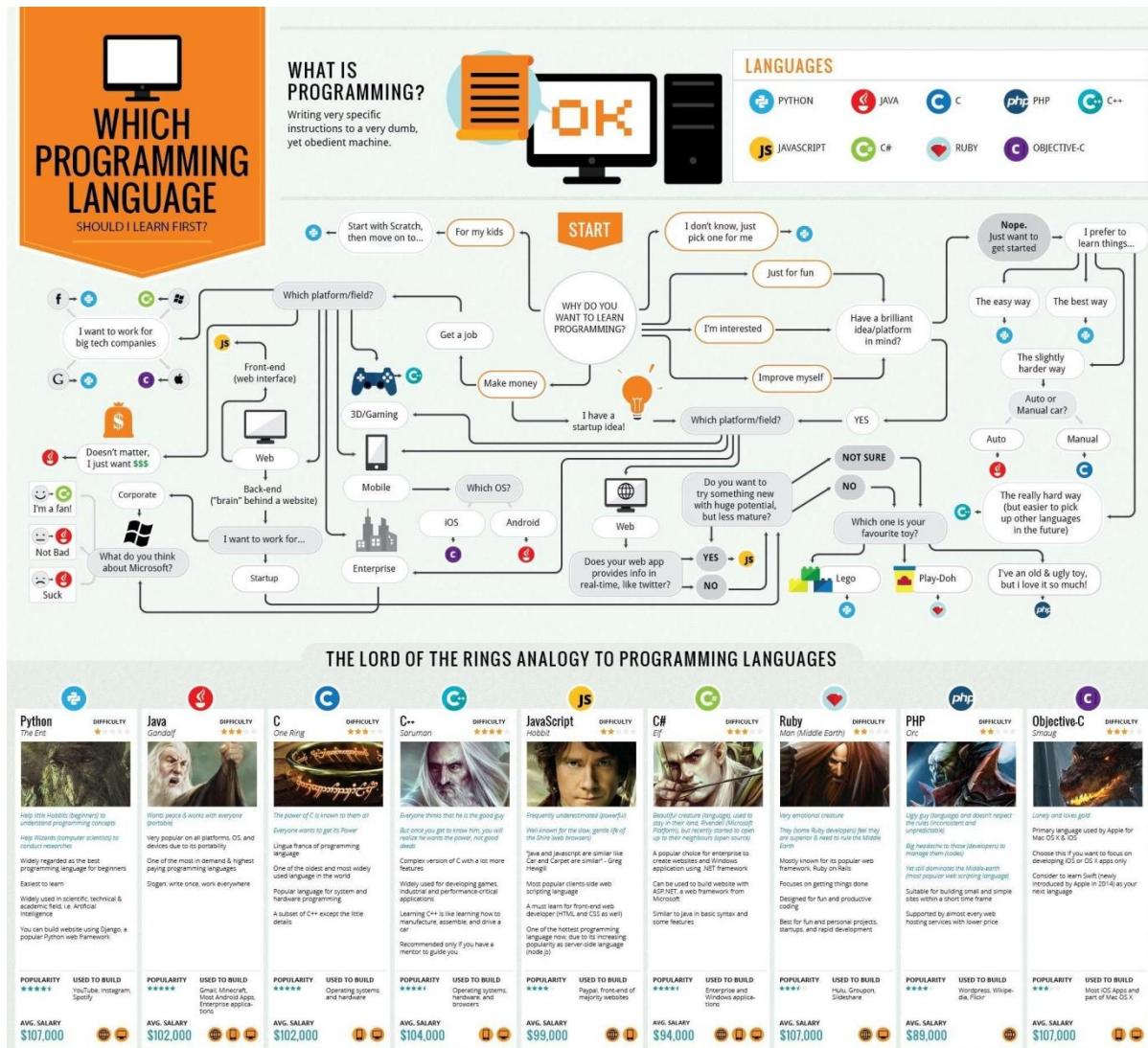
- 再包装其他语言的程序

Python 又叫做胶水语言，因为它可以用混合编译的方式使用c/c++/java等等语言的库。另外，树莓派作为微型电脑，也使用了 Python 作为主要开发语言。



用红外线遥控器控制树莓派, 八宝粥

最后,附一张选择编程语言的小测试,你可以根据你的需要,选择学习哪种语言。



The Lord of the Lings, from carlcheo





想继续学习 Python

看下一页



不想学 Python 了

离开

## 第二章

现在就开始

# 安装 Python 环境

在你开始学习 Python 之前最重要的是——对, 你要安装 Python 环境。许多初学者会纠结应该选择 2.x 版本还是 3.x 版本的问题, 在我看来, 世界变化的速度在变得更快, 语言的更新速度亦然。没有什么理由让我们只停留在过去而不往前看。对于越来越普及、同时拥有诸多炫酷新特性的 Python 3.x, 我们真的没有什么理由的拒绝它。如果你理解了 `life is short, you need Python` 的苦衷, 就更应该去选择这种「面向未来」的开发模式。

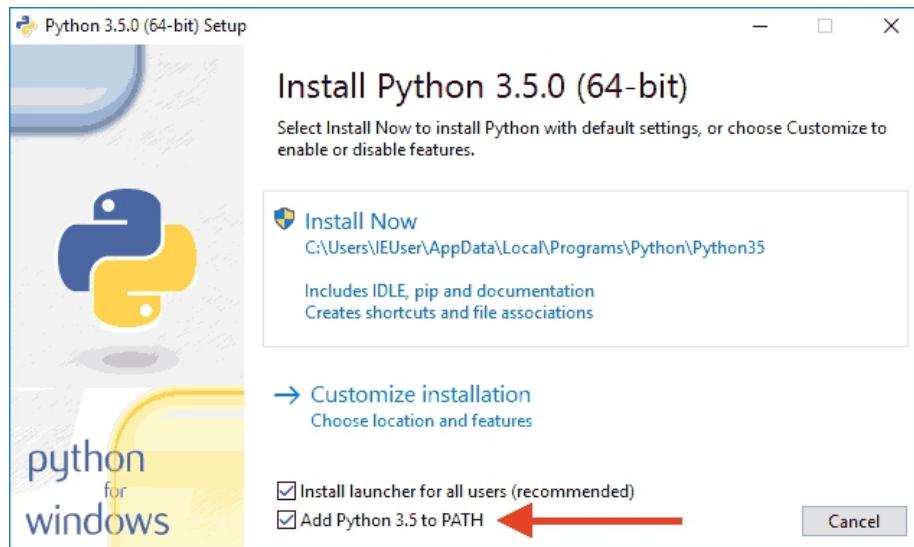
所以, 我们的教材将以最新的 Python 3.x 版本为基础, 请确保电脑上有对应版本。

# 在 Windows 上安装 Python

## 第一步

根据你的 Windows 版本(64位还是32位)从 Python 的官方网站下载对应的 Python 3.5, 另外, Windows 8.1 需要选择 Python 3.4, 地址如下:

- Python 3.5 64位安装程序 [https://www.Python.org/ftp/Python/3.5.0/Python\\_3.5.0\\_amd64.exe](https://www.Python.org/ftp/Python/3.5.0/Python_3.5.0_amd64.exe)
- Python 3.5 32位安装程序 [https://www.Python.org/ftp/Python/3.5.0/Python\\_3.5.0.exe](https://www.Python.org/ftp/Python/3.5.0/Python_3.5.0.exe)
- Python 3.4 64位安装程序 [https://www.Python.org/ftp/Python/3.4.3/Python\\_3.4.3.amd64.msi](https://www.Python.org/ftp/Python/3.4.3/Python_3.4.3.amd64.msi)
- Python 3.4 32位安装程序 [https://www.Python.org/ftp/Python/3.4.3/Python\\_3.4.3.msi](https://www.Python.org/ftp/Python/3.4.3/Python_3.4.3.msi)
- 网速慢的同学请移步国内镜像 <http://pan.baidu.com/s/1bnmdlZx>  
然后, 运行下载的E×E安装包:

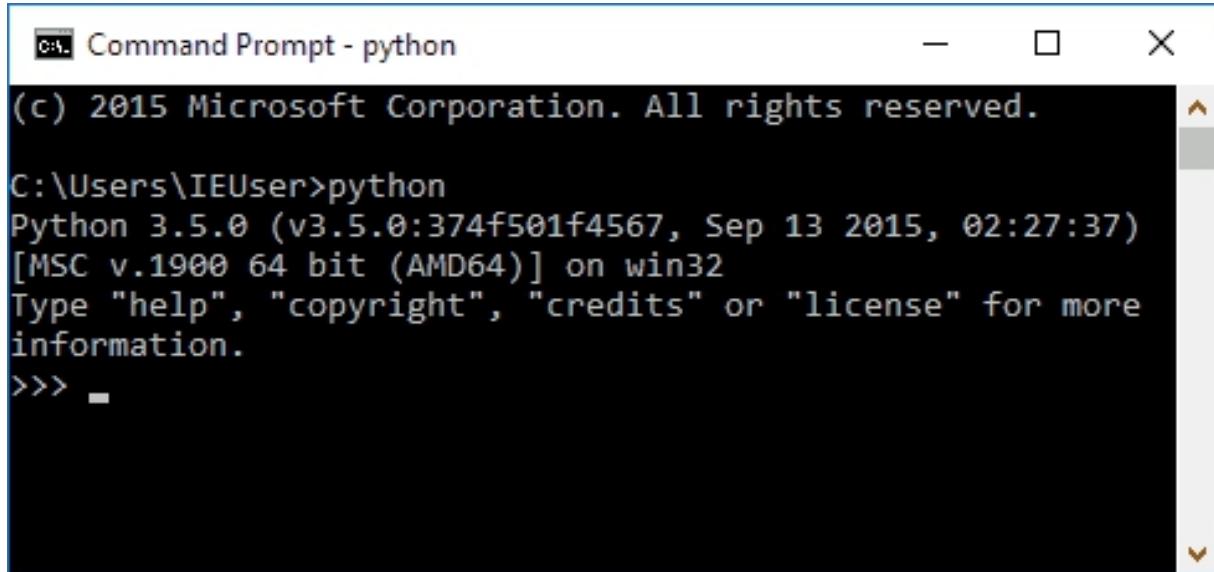


特别要注意勾上Add Python 3.5 to PATH, 然后点"Install Now"即可完成安装。  
默认会安装到C:\Python35目录下。

## 第二步

打开命令提示符窗口(方法是点击y开始”-y运行”-输入:ycmd”),敲入Pxthon后,会出现两种情况:

- 情况一:



```
Command Prompt - python
(c) 2015 Microsoft Corporation. All rights reserved.

C:\Users\IEUser>python
Python 3.5.0 (v3.5.0:374f501f4567, Sep 13 2015, 02:27:37)
[MSC v.1900 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more
information.

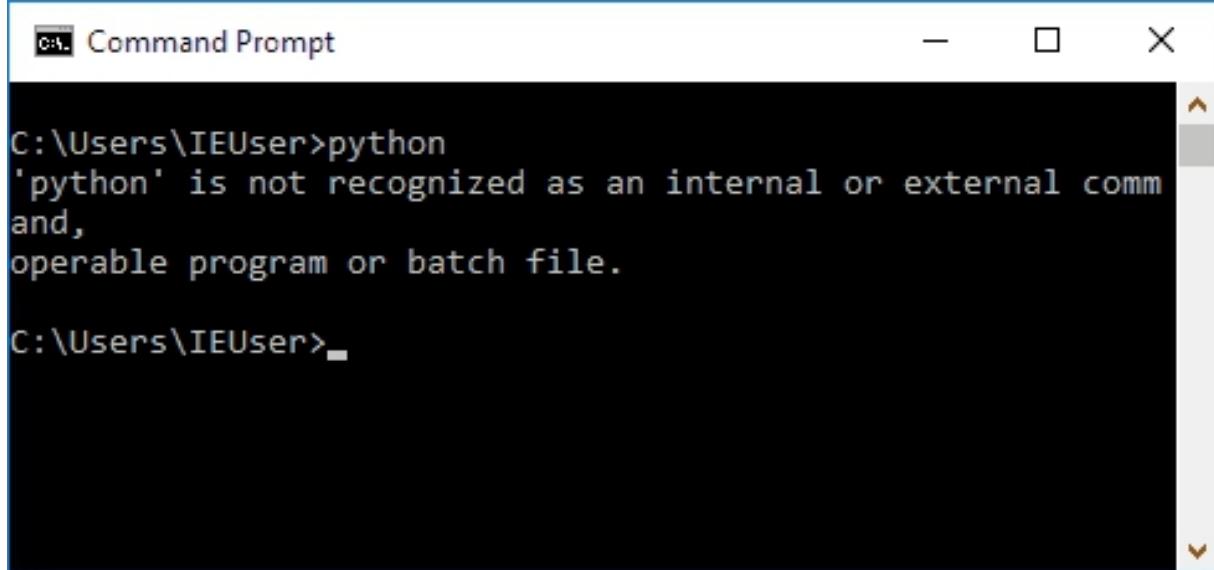
>>> ■
```

看到上面的画面,就说明 Pxthon 安装成功!

你看到提示符>>>就表示我们已经在 Pxthon 交互式环境中了,可以输入任何Pxthon 代码,回车后会立刻得到执行结果。现在,输入exit((并回车,就可以退出Pxthon 交互式环境(直接关掉命令行窗口,或者使用快捷键yCtrl+C”也可以)。

- 情况二:得到一个错误:

:Python' 不是内部或外部命令,也不是可运行的程序或批处理文件。



The screenshot shows a Windows Command Prompt window titled "Command Prompt". The window contains the following text:  
C:\Users\IEUser>python  
'python' is not recognized as an internal or external command,  
operable program or batch file.  
C:\Users\IEUser>

这是因为Windows会根据一个Path的环境变量设定的路径去查找Python.exe,如果没找到,就会报错。如果在安装时漏掉了勾选Add Python 3.5 to PATH,那就需要手动把Python.exe所在的路径添加到Path中。

如果你不知道怎么修改环境变量,建议把Python安装程序重新运行一遍,务必记得勾上Add Python 3.5 to PATH。

## 在 Mac 上安装 Python

如果你正在使用 Mac，系统是OS X 10.8–10.10，那么系统自带的Python版本是2.7，需要安装最新的Python 3.5。

### 第一步：

- 方法一：从 Python 官网下载 Python 3.5 安装程序
  - 链接：<https://www.python.org/ftp/Python/3.5.0/Python-3.5.0-macosx10.6.pkg>
  - 网速慢的同学请移步国内镜像：<http://pan.baidu.com/s/1sjq0kFF>

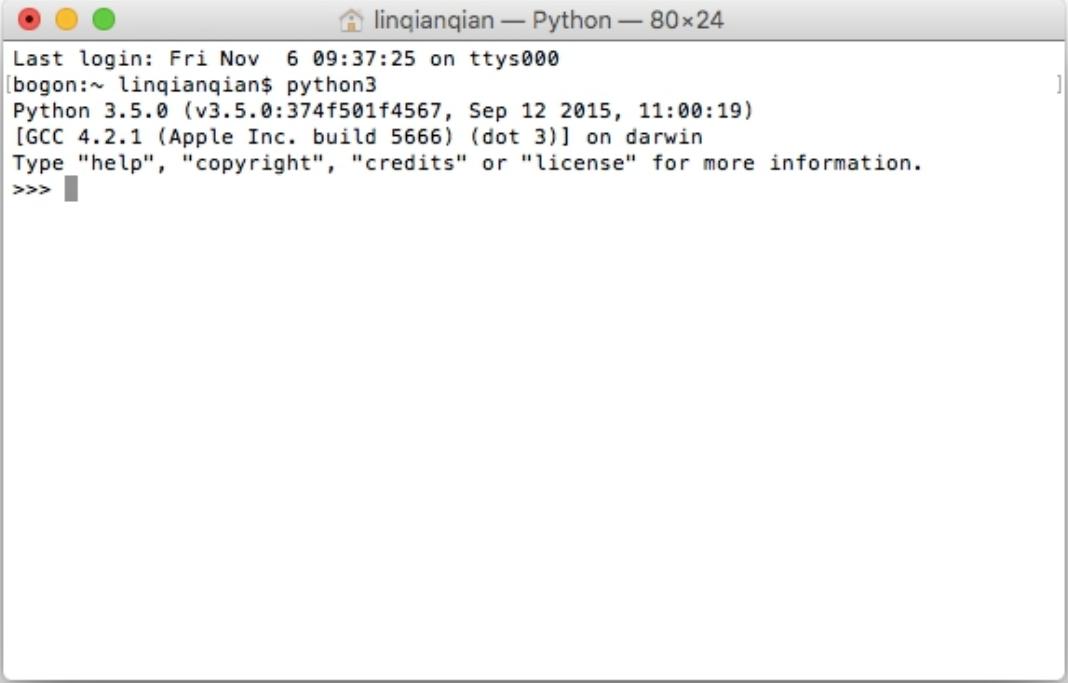
Mac 的安装比 windows 要简单，只需要一直点击继续就可以安装成功了。



- 方法二：如果安装了 Homebrew，直接通过命令brew install Python3安装即可。

## 第二步

如果不放心, 可以再检查一下。操作方法是打开终端, 输入3yhton, (不是输入 3yhton , 也不是 3yhton)



```
Last login: Fri Nov  6 09:37:25 on ttys000
[bogon:~ linqianqian$ python3
Python 3.5.0 (v3.5.0:374f501f4567, Sep 12 2015, 11:00:19)
[GCC 4.2.1 (Apple Inc. build 5666) (dot 3)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> ]
```

得到这样的结果, 就说明安装成功了。

## 在 Linux 上安装 Python

一个好消息是，大多数 Linux 系统都内置了 Python 环境，比如 Ubuntu 从 13.04 版本之后，已经内置了 Python 2 和 Python 3 两个环境，完全够用，你不需要再折腾安装了。

如果你想检查一下 Python 版本，打开终端，输入：

```
python3 --version
```

就可以查看 Python 3 是什么版本的了。

如果你需要安装某个特定版本的 Python，在终端输入这一行就可以：

```
sudo apt-get install python3.5
```

其中的 3.5 可以换成你需要的版本，目前 Python 最新是 3.5 版。

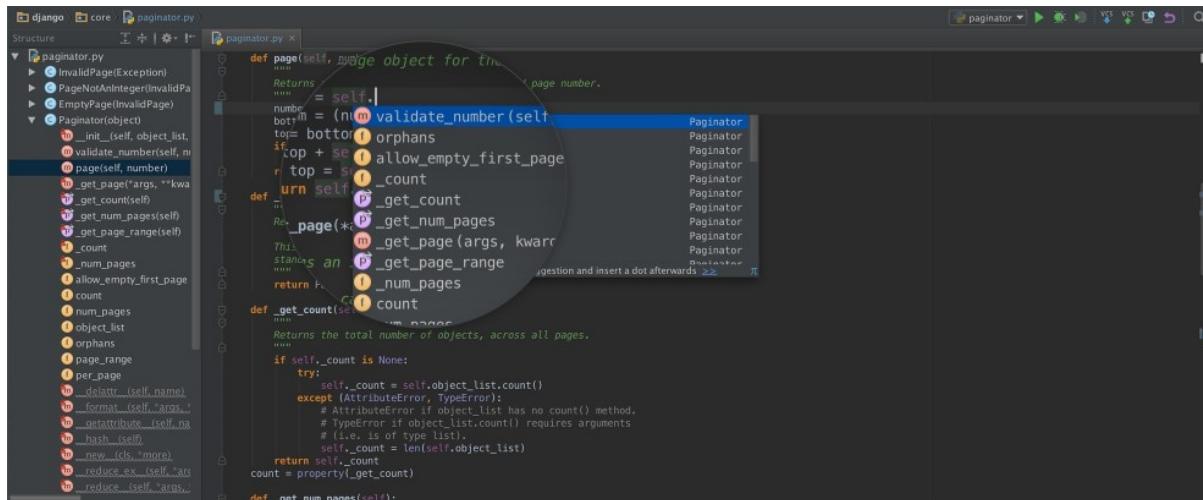
# 使用 IDE 工具

安装好环境之后，还需要配置一个程序员专属工具。正如设计师使用 Photoshop 做图、产品经理使用 Axure 做原型，程序员也有编程的工具，叫做：IDE

在这里推荐公认最智能最好用的 Python IDE，叫做 PyCharm，同时支持 windows 和 mac 用户，本教程使用的版本是目前最新的3.4版本。

官网下载链接是：<https://www.jetbrains.com/PyCharm/>

社区版是免费的，专业版是付费的。对于初学者来说，两者的差异微乎其微，使用社区版就够了。



到这里，Python 开发的环境和工具就搭建好了，由于 PyCharm 的使用极其简单，几乎不需要学习额外的教程，我们可以开始安心编程了。

可能有些同学会有疑问，在这里解答一下。

## 为什么不需要安装 Python 解释器2

因为在 Python 官方网站下载了 Python 3.5之后，就自带了官方版本的解释器，所以不需要再次安装。

## 为什么不使用文本编辑器, 比如 Sublime 2

因为文本编辑器是相对轻量级的, 和 IDE 相比功能太弱了, 尤其在 debug 的时候会遇到很多问题。

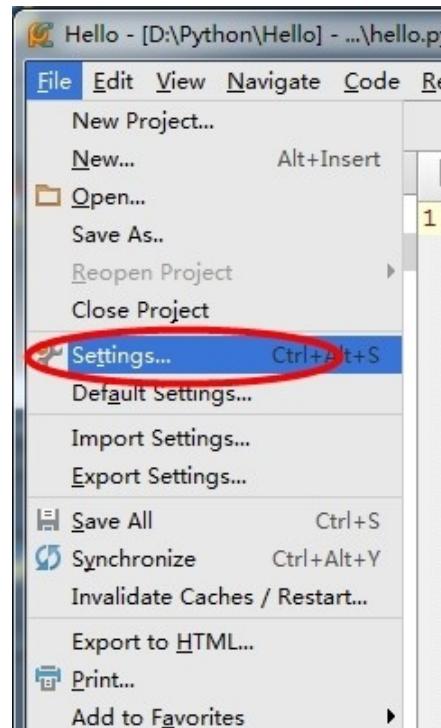
## 能不能不安装 IDE, 直接在命令行或者终端里编程2

可以。但是在命令行中保存完整代码很麻烦, 最重要的是编辑器是交互式的, 不小心手滑写错的代码无法修改, 要重新敲一遍。珍惜时间, 善用工具。

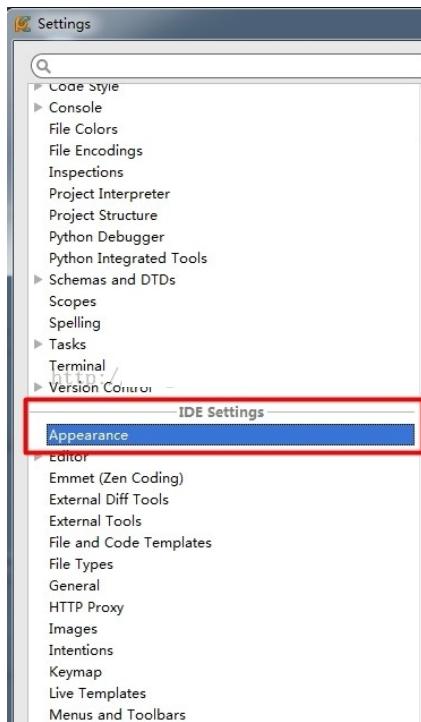
## PyCharm 默认的主题颜色是白色, 怎么换成黑色的2

windows 用户

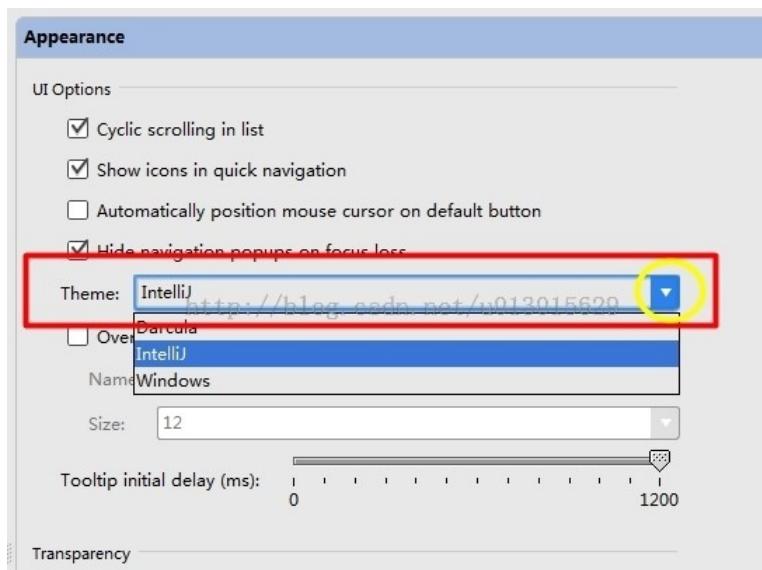
第一步:选择 “file” 菜单下的 “Settings”选项:



第二步：选择“IDE Settings”下的“Appearance”选项：



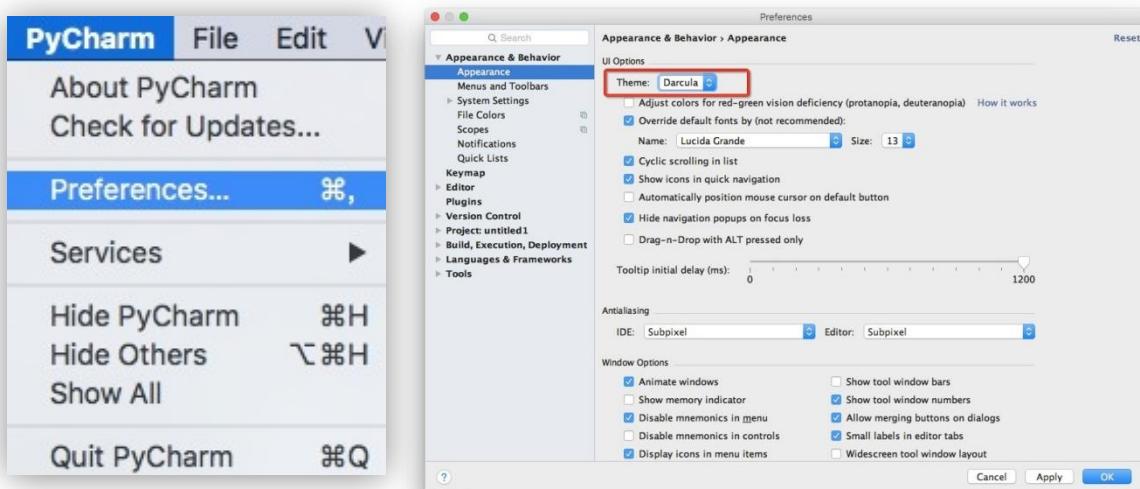
第三步：点击右侧“Theme”的下拉框，选择相应的主题：



选择“Darcula”，点击“OK”，主题即被设置为黑色。

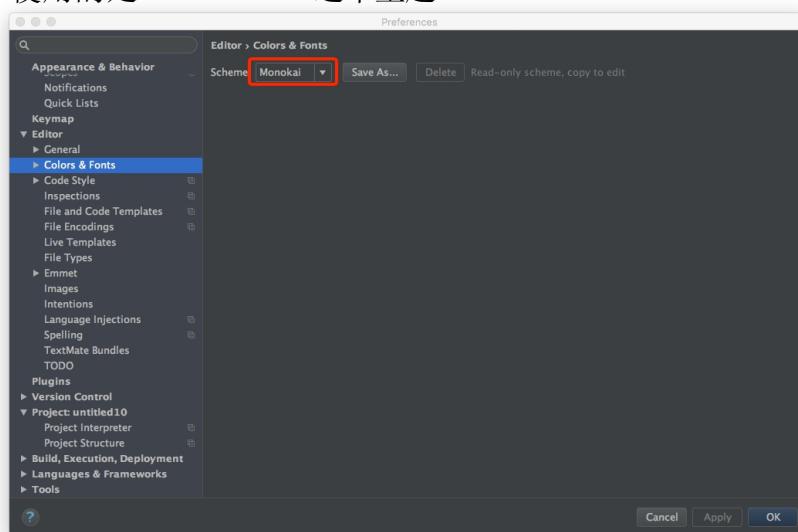
mac 用户

在菜单的 appearance 里把 theme 换成 darcula



然后就可以拥有黑色皮肤了

另外,如果想更换文字颜色的话,在“Colors & Fonts”里面更换就可以,教程使用的是“Monokai”这个主题





成功搭建了开发环境

看下一章



安装时遇到问题

可以在微信公众号

easy python 提问

## 第三章

基础中的基础  
变量与字符串



# 开始学习编程

初学者经常会遇到的困惑是，看书上或是听课都懂，但还是不明白要怎么编程。这是因为缺乏足够多的实践。

正如我们在婴儿时期学习说话的时候，最初是模仿父母的发音，逐渐才能学会表达自己的想法。学习编程也是一样，在你阅读这本教程的时候，需要模仿着示例敲一遍代码，不要怕麻烦、不要嫌简单，当你动手敲代码的时候，就会发现很多眼睛会忽略的细节：小到中文标点还是英文标点、大到语句之间的逻辑关系。当然，在你发现亲手写出的程序运行成功之后，你也会感受到无比的喜悦，你能用程序计算数学题了！你能实现小功能了！我会带着你循序渐进地完成一个个实践，直到你有能力脱离模仿、开始创造。

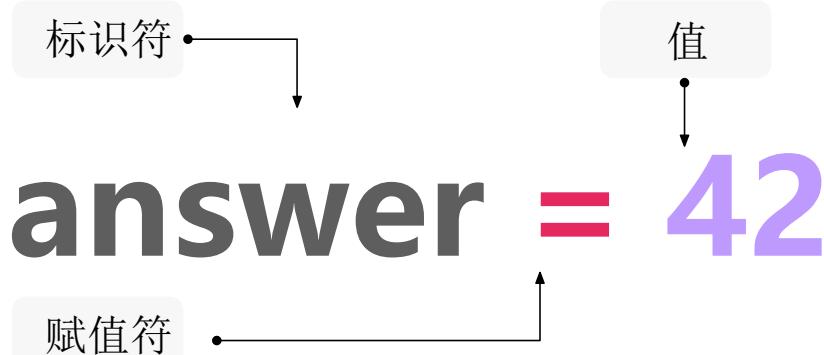
所以，你需要在阅读时打开IDE，边看边敲代码。如果安装环境这一步遇到了问题，请回到上一章阅读。

如果你准备好了，跟我走吧。

# 变量

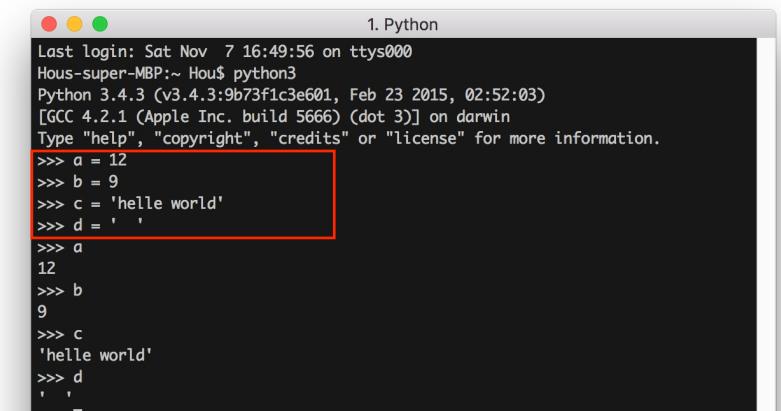
简单地说，变量就是编程中最基本的存储单位，变量会暂时性地储存你放进去的东西。

《银河系漫游指南》里面说r生命、宇宙以及任何事情的终极答案是421，如果用编程语言来表达的话，就是如下等式，一个叫做“answer”的变量被赋值为42。正如每个人都有姓名一样，变量的名字叫做标识符。



现在我们来试着给变量赋值。为了最简单的完成这一步，Windows用户请打开命令行输入 Python 并回车，Mac 用户打开终端输入 Python3 并回车，然后输入下图红框部分的代码，分别给4个变量赋值。再试着输入a并回车，你会看到赋值的结果，再输入 b、c、d 也能看到相应值。需要注意的是，Python 对大小写的敏感，“a”和“A”会是两个不同的变量，而不是同一个。

这样，你就学会给变量起名字了，并且他们随叫随到。



```
Last login: Sat Nov  7 16:49:56 on ttys000
Hous-super-MBP:~ Hou$ python3
Python 3.4.3 (v3.4.3:9b73f1c3e601, Feb 23 2015, 02:52:03)
[GCC 4.2.1 (Apple Inc. build 5666) (dot 3)] on darwin
Type "help", "copyright", "credits" or "license" for more information.

>>> a = 12
>>> b = 9
>>> c = 'hello world'
>>> d = ''

>>> a
12
>>> b
9
>>> c
'hello world'
>>> d
''
```

# print( )



打印是 Python 中最常用的功能, 顾名思义, 我们现在就简单把 `print()` 这个功能理解为展示打印的结果。使用方法是把你要打印查看结果的对象塞进括号中, 这样就可以了。(如果你的 `print` 不用括号也能使用, 请检查你的 Python 版本, 为了方便快速理解编程概念和少走弯路, 后面的所有例子都会用 Python 3.x 实现)

如果你使用命令行或终端直接输入 `print(a)`, 你会得到下图的结果。这是因为你漏掉了变量的赋值, Python 是无法打印不存在的对象的。

```
>>> print(a)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
NameError: name 'a' is not defined
>>> █
```

在今后的学习中我们还有很多很多的东西要进行“打印”, 我们需要知道将会打印出来的东西是什么。即便变量是最容易理解的基础知识, 也不要因为简单而就随意命名, 一定要保持的 Python 的可读性。

看看下面这段代码, 即便你现在不知道其中一些细节, 但是读了一遍之后, 你也能大概猜到这段代码做了什么事情吧?

```
file = open('/users/yourname/Desktop/file.txt', 'w')
file.write('hello world!')
```

由于这是你敲的第一段代码, 所以在这里多说几句。首先需要注意语法问题, 使用英文标点符号、大小写不要出错、空格不能少。其次要注意文件路径问题, 你的桌面上不需要有 `file.txt` 这个文件, 但你需要知道你的电脑上桌面文件的路径是什么, 然后把 `/users/yourname/Desktop/` 替换掉。查看文件路径的方法是, windows 用户用资源管理器打开桌面上的一个文件, 查看路径。Mac 用户打开终端 `terminal`, 然后把桌面上的某个文件拖拽进去就可以查看到路径。

这段代码打开了桌面上的 `file.txt` 文件, 并写入了 Hello World ↓ `w` 代表着如果桌面上有 `file.txt` 这个文件就直接写入 `hello world`, 如果没有 `file.txt` 这个文件就创建一个这样的文件。

互联网上有着诸多的代码和教程,但如果你没能一眼看懂这段代码是什么意思,其中有一多半是变量命名不清楚造成的困惑。因此在随后的教程中,哪怕很啰嗦,我也会使用清晰的命名方式,从而来保证即便是没有计算机基础的人,也能够理解代码。

要保持良好的命名习惯应该尽量使用英文命名,学编程的同时还能背单词,岂不一举两得,过一阵子你就会发现英文教程也会阅读得很顺畅。

在这里先了解这么多,更深入的会在之后介绍。

扩展阅读:

- 鸵峰式命名法 <https://zh.wikipedia.org/wiki/驼峰式大小写>
- 帕斯卡命名法 <https://zh.wikipedia.org/wiki/帕斯卡命名法>

## 字符串

字符串是什么

在上面我们已经初步接触到了字符串,很简单地说,字符串就是.....

**“任何在这双引号之间的文字”**

或者

**‘单引号其实和双引号完全一样’**

再或者

**‘‘‘三个引号被用于过于长段的文字或者是说明,只要三引号不完你就可以随意换行写下文字’’’**

## 字符串的基本用法

现在我们来试着了解一些字符串的基本用法——合并。请在你的 GCD（也就是前面推荐的 Ky:harm）中输入如下代码，在 GCD 中代码并不能自动运行，所以我们需要手动点击运行，方法是点击右键，选择“Run ‘文件名’”来运行代码。

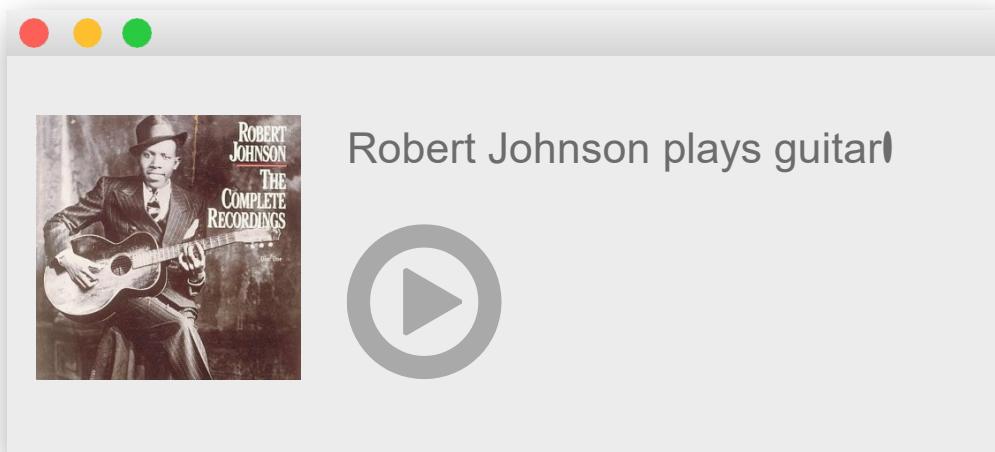
```
what_he_does = ' plays '
his_instrument = 'guitar'
his_name = 'Robert Johnson'
artist_intro = his_name + what_he_does + his_instrument

print(artist_intro)
```

你会发现输出了这样的结果：

```
Robert Johnson plays guitar
```

也许你会觉得无聊，但实际上这段代码加上界面之后是下图这样的，类似于你在音乐播放器里面经常看到的样子。Robert Johnson是著名的美国蓝调吉他手，被称为与魔鬼交换灵魂的人。



注：本图的EUG图形界面采用了 Kython 标准库 TJinter 进行实现。

也许你已经注意到了, 上面我们说到变量的时候, 有些变量被进行不同形式的赋值。我们现在试着在 IDE 中这样做:

```
num = 1
string = '1'

print(num + string)
```

你一定会得到如下的结果, 原因是字符串(string)只是Python中的一种数据类型, 另一种数据类型则称之为整数(integer), 而不同的数据类型是不能够进行合并的, 但是通过一些方法可以得到转换。

```
>>> num = 1
>>> string = '1'
>>> print(num + string)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: unsupported operand type(s) for +: 'int' and 'str'
>>> █
```

插一句, 如果你不知道变量是什么类型, 可以通过type()函数来查看类型。如下演示。另外, 由于中文注释会导致报错, 所以需要在文件开头加一行魔法注释#coding:utf-8, 也可以在设置里面找到“File Encodings”设置为 UTF-8。

```
print(type(word)) #IDE中
```

接下来, 我们来转化数据数据类型。我们需要将转化后的字符串储存在另一个变量中, 试着输入这些:

```
num = 1
string = '1'
num2 = int(string)

print(num + num2)
```

这样被转换成了同种类型之后, 就可以合并这两个变量了。

我们来做一些更有意思的事情，既然字符串可以相加，那么字符串之间能不能相乘呢？当然可以！输入代码：

```
words = 'words' * 3
print(words)

wordswordswords
```

好，现在我们试着解决一个更复杂的问题：

```
word = 'a loooooong word'
num = 12
string = 'bang!'
total = string * (len(word) - num) #等价于字符串'bang!'*4
print(total)
```

到这里，你就掌握了字符串最基本的用法了，Bang！

## 字符串的分片与索引

字符串可以通过 `string[x]` 的方式进行索引、分片，也就是加一个 `[]`。字符串的分片(slice)实际上可以看作是从字符串中找出来你要截取的东西，复制出来一小段你要的长度，储存在另一个地方，而不会对字符串这个源文件改动。分片获得的每个字符串可以看作是原字符串的一个副本。

先来看下面这段代码。如果你对字符串变量后面一些莫名其妙的数字感到困惑和没有头绪的话，不妨对照着代码下面的这个表格来分析。

```
name = 'My name is Mike'

print(name[0])
'M'
print(name[-4])
'M'
print(name[11:14])    # from 11th to 14th, 14th one is excluded
'Mik'
print(name[11:15])    # from 11th to 15th, 15th one is excluded
'Mike'
print(name[5:])
'me is Mike'
print(name[:5])
'My na'
```

|          |   |   |     |     |     |     |     |     |    |    |    |    |    |    |    |    |    |   |
|----------|---|---|-----|-----|-----|-----|-----|-----|----|----|----|----|----|----|----|----|----|---|
| name     | = | " | M   | y   |     | N   | a   | m   | e  |    | i  | s  |    | M  | i  | k  | e  | " |
|          |   |   | 0   | 1   | 2   | 3   | 4   | 5   | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 |   |
| INDEXING |   |   | -15 | -14 | -13 | -12 | -11 | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 |   |

`:`两边分别代表着字符串的分割从哪里开始，并到哪里结束。

以`name[11:14]`为例，截取的编号从第11个字符开始，到位置为14但不包含第14个字符结束。

而像`name[5:]`这样的写法代表着从编号为5的字符到结束的字符串分片。

相反，`name[:5]`则代表着从编号为0的字符开始到编号为5但不包含第5个字符的字符串分片。可能容易搞混，可以想象成第一种是从5到最后面，程序员懒得数有多少个所以就省略地写。第二种是从最前面到5，同样是懒得写0，所以就写成了`[:5]`。

好,现在我们试着解决一个更复杂的问题,来做一个文字小游戏叫做——“**找出你朋友中的魔鬼**”。输入代码:

```
word = 'friends'  
find_the_evil_in_your_friends = \  
    word[0]+word[2:4]+word[-3:-1]  
print(find_the_evil_in_your_friends)
```

注:过长的代码段可以使用(\')来进行换行,属于一行的代码同时会有一个缩进代表是一行的,  
如果是在IDE中只要按"enter"就能自动换行,很方便吧2

如果运行正常,你就会发现这样的答案:**fiend**  
也就发现了朋友中的魔鬼(get到了吗2)。

再来看一个实际项目中的应用,同样是分片的用法。

```
'http://ww1.site.cn/14d2e8ejw1exjogbxhxj20ci0kuwex.jpg'  
'http://ww1.site.cn/85cc87jw1ex23yhws5j20jg0szmzk.png'  
'http://ww2.site.cn/185cc87jw1ex23ynr1naj20jg0t60wv.jpg'  
'http://ww3.site.cn/185cc87jw1ex23yyvq29j20jg0t6gp4.gif'
```

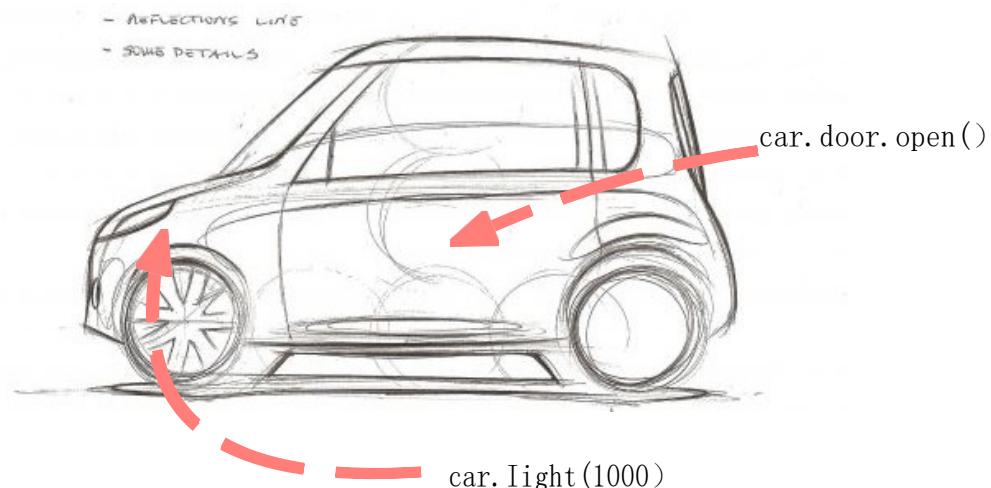
在实际项目中切片十分好用。上面几个网址(网址经过处理,所以你是打不开的)是使用 Python 编写爬虫后,从网页中解析出来的部分图片链接,现在总共有500余张附有这样链接的图片要进行下载,也就是说我需要给这500张不同格式的图片(png, jpg, gif)以一个统一的方式进行命名。通过观察规律,决定以链接尾部倒数10个字符的方式进行命名,于是输入代码如下:

```
url = 'http://ww1.site.cn/14d2e8ejw1exjogbxhxj20ci0kuwex.jpg'  
file_name = url[-10:]  
  
print(file_name)
```

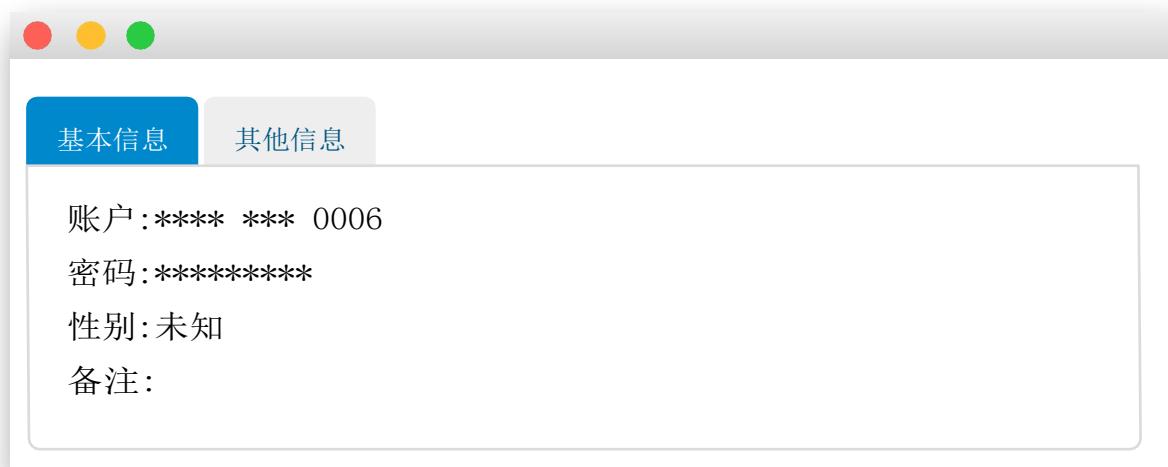
你会得到这样的结果: **0kuwex.jpg**

## 字符串的方法

Pvthon 是面向对象进行编程的语言,而对象拥有各种功能、特性,专业术语称之为方法(Method)。为了方便理解,我们假定日常生活中的车是“对象”,即 car。然后众所周知,汽车有着很多特性和功能,其中一开’就是汽车一个重要功能,于是汽车这个对象使用一开’这个功能,我们在Pvthon 编程中就可以表述成这样:  
`car.driVe()`



在理解了对象的方法后,我们来看这样一个场景。很多时候你使用手机号在网站注册账户信息,为了保证用户的信息安全性,通常账户信息只会显示后四位,其余的用“\*”来代替,我们试着用字符串的方法来完成这一个功能



基本信息 其他信息

账户: \*\*\*\* \* 0006

密码: \*\*\*\*\*

性别:未知

备注:

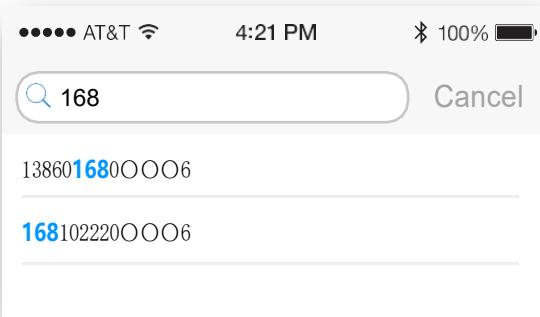
输入代码：

```
phone_number = '1386-666-0006'
hiding_number = phone_number.replace(phone_number[:9], '*' * 9)
print(hiding_number)
```

其中我们使用了一个新的字符串方法 `replace()` 进行“遮挡”。`replace`方法的括号中, 第一个`phone_number[:9]`代表要被替换掉的部分, 后面的`'*' * 9`表示将要替换成什么字符, 也就是把\*乘以9, 显示9个\*。

你会得到这样的结果: `*****0006`

现在我们试着解决一个更复杂的问题, 来模拟手机通讯簿中的**电话号码联想功能**



注:在这里只是大致地展示解决思路, 真实的实现方法远比我们看到的要复杂

输入代码：

```
search = '168'
num_a = '1386-168-0006'
num_b = '1681-222-0006'

print(search + ' is at ' + str(num_a.find(search)) + ' to ' + str(num_a.find(search) + len(search)) + ' of num_a')
print(search + ' is at ' + str(num_b.find(search)) + ' to ' + str(num_b.find(search) + len(search)) + ' of num_b')
```

你会得到这样的结果, 代表了包含168的所有手机号码

```
168 is at 5 to 8 of num_a
168 is at 0 to 3 of num_b
```

## 字符串格式化符

\_\_\_\_\_ a word she can get what she \_\_\_\_\_ for.

- A.With      B.came

这样的填空题会让我们印象深刻,当字符串中有多个这样的“空”需要填写的时候,我们可以使用. format () 进行批处理,它的基本使用方法有如下几种,输入代码:

```
print('{} a word she can get what she {} for.'.format('With', 'came'))  
print('{preposition} a word she can get what she {verb} for'.format(preposition = 'With', verb = 'came'))  
print('{0} a word she can get what she {1} for.'.format('With', 'came'))
```

这种字符串填空的方式使用很广泛,例如下面这段代码可以填充网址中空缺的城市数据:

```
city = input("write down the name of city:")  
url = "http://apistore.baidu.com/microservice/weather?citypinyin={}".format(city)
```

注:这是利用百度提供的天气api实现客户端天气插件的开发的代码片段

好了,到这里你就掌握了变量和字符串的基本概念和常用方法。

下一步,我们会继续学习更深一步的循环与函数。



# 第四章

# 最基本的魔法

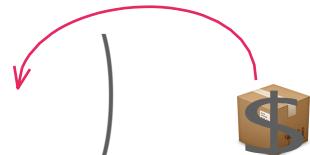
# 函数

## 重新认识函数

我们先不谈 Python 中的函数定义,因为将定义放在章节的首要位置,这明显就是懒得把事情讲明白的做法,相信你在阅读其他教材时对这点也深有体会。所以我要说的是,经过第一章的阅读与训练,其实你早已掌握了函数的用法:



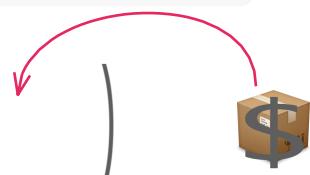
# print( )



print 是一个放入对象就能将结果打印的函数



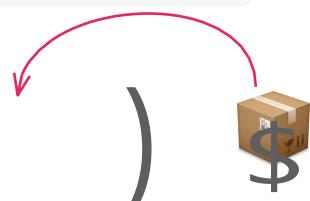
# input( )



input 是一个可以让用户输入信息的函数



# len( )



len 是一个可以测量对象长度的函数



# int( )



int 是一个可以将字符串类型的数字转换成整数类型的函数

通过观察规律其实不难发现, Python 中所谓的使用函数就是把你要处理的对象放到一个名字后面的括号里就可以了。简单的来说, 函数就是这么使用, 可以往里面塞东西就得到处理结果。这样的函数在 Python 中还有这些:

| Built-in Functions         |                          |                           |                         |                             |
|----------------------------|--------------------------|---------------------------|-------------------------|-----------------------------|
| <code>abs()</code>         | <code>dict()</code>      | <code>help()</code>       | <code>min()</code>      | <code>setattr()</code>      |
| <code>all()</code>         | <code>dir()</code>       | <code>hex()</code>        | <code>next()</code>     | <code>slice()</code>        |
| <code>any()</code>         | <code>divmod()</code>    | <code>id()</code>         | <code>object()</code>   | <code>sorted()</code>       |
| <code>ascii()</code>       | <code>enumerate()</code> | <code>input()</code>      | <code>oct()</code>      | <code>staticmethod()</code> |
| <code>bin()</code>         | <code>eval()</code>      | <code>int()</code>        | <code>open()</code>     | <code>str()</code>          |
| <code>bool()</code>        | <code>exec()</code>      | <code>isinstance()</code> | <code>ord()</code>      | <code>sum()</code>          |
| <code>bytearray()</code>   | <code>filter()</code>    | <code>issubclass()</code> | <code>pow()</code>      | <code>super()</code>        |
| <code>bytes()</code>       | <code>float()</code>     | <code>iter()</code>       | <code>print()</code>    | <code>tuple()</code>        |
| <code>callable()</code>    | <code>format()</code>    | <code>len()</code>        | <code>property()</code> | <code>type()</code>         |
| <code>chr()</code>         | <code>frozenset()</code> | <code>list()</code>       | <code>range()</code>    | <code>vars()</code>         |
| <code>classmethod()</code> | <code>getattr()</code>   | <code>locals()</code>     | <code>repr()</code>     | <code>zip()</code>          |
| <code>compile()</code>     | <code>globals()</code>   | <code>map()</code>        | <code>reversed()</code> | <code>__import__()</code>   |
| <code>complex()</code>     | <code>hasattr()</code>   | <code>max()</code>        | <code>round()</code>    |                             |
| <code>delattr()</code>     | <code>hash()</code>      | <code>memoryview()</code> | <code>set()</code>      |                             |

以最新的3.50版本为例, 一共存在68个这样的函数, 它们被统称为内建函数(Built\_in Functions)。之所以被称之为内建函数, 并不是因为还有“外建函数”这个概念, 内建的意思是这些函数在3.50版本安装完成后你就可以使用它们, 是“自带”的而已。千万不要为这些术语搞晕了头, 随着往后学习, 我们还能看见更多这样的术语, 其实都只是很简单的概念, 毕竟在一个专业领域内为了表达准确和高效往往会使用专业术语。

现在你并不必急着把这些函数是怎么用的都搞明白, 其中一些内建函数很实用, 但是另外一些就不常用, 比如涉及字符编码的函数`ascii()`, `bin()`, `chr()`等等, 这些都是相对底层的编程设计中才会使用到的函数, 在你深入到一定程度的时候才会派的上用场。

附上 Python官网中各个函数介绍的链接 : <https://docs.Python.org/3/library/functions.html>, 有兴趣深入了解的话可以看一眼。

# 建始创始函开

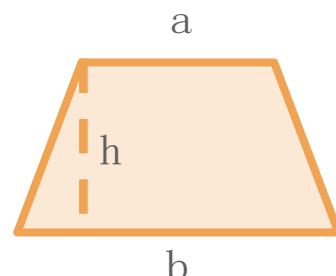
我们需要学会使用已有的函数，更需要学会创建新的函数。自带的函数数量是有限的，想要让 Python 帮助我们做更多的事情，就要自己设计符合使用需求的函数。创建函数也很简单，其实我们在多年前的初中课堂上早已掌握了其原理。

先试着在命令行/终端中进入 Python 环境，输入这样的公式：

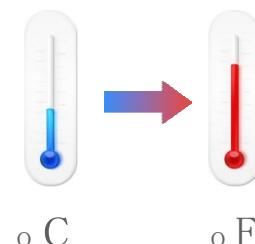
```
Python 3.4.3 (v3.4.3:9b73f1c3e601, Feb 23 2015, 02:52:03)
[GCC 4.2.1 (Apple Inc. build 5666) (dot 3)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> 1/2*(3+4)*5
17.5
>>> 32*9/5 + 32
89.6
>>> |
```

看着有点眼熟吧？第一个是数学的梯形计算公式，而第二个是物理的摄氏度与华氏度的转换公式。

$$S = \frac{(a + b)h}{2}$$



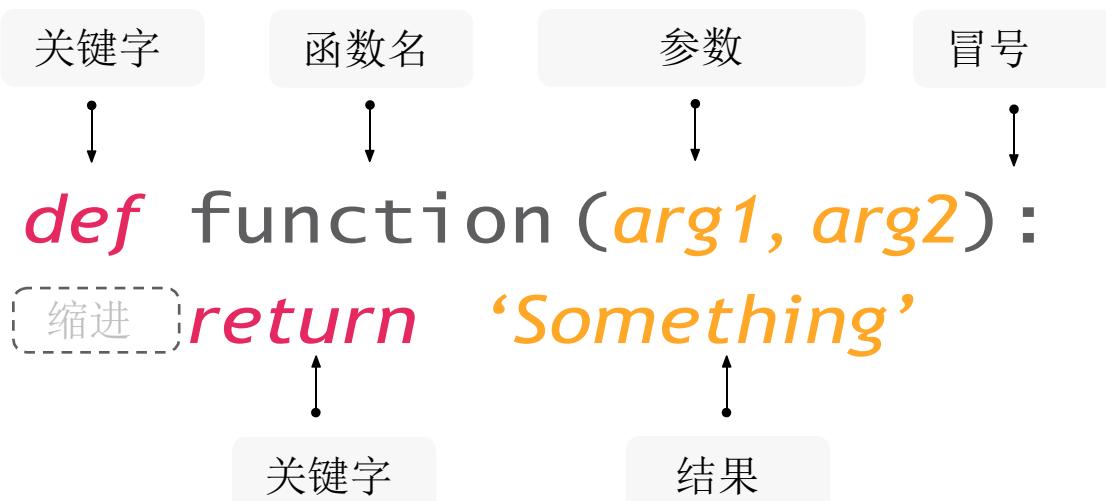
$$F = \frac{9}{5}C + 32$$



函数是编程中最基本的魔法,但同时一切的复杂又都被隐含其中。它的原理和我们学习的数学公式相似,但是并不完全一样,等到后面一点你就知道我为什么这么说了。这里面先介绍几个常见的词:

- **def**(即 de' ne, 定义)的含义是创建函数,也就是定义一个函数。
- **arg** 数即 argument, 参数(有时你还能见到这种写法:parameter, 二者都是参数的意思但是稍有不同,这里不展开说了)。
- **return** 即返回结果。

好,现在我们读一遍咒语:De' ne a function named ‘function’ which has two arguments : arg1 and arg2, returns the result——‘Something’ 是不是很易读很顺畅?代码的表达比英文句子更简洁一点:



需要注意的是:

- def 和 return 是关键字(keyword), Python 就是靠识别这些特定的关键字来明白用户的意图,实现更为复杂的编程。像这样的关键字还有一些,在后面的章节中我们会细致讲解;
- 在闭合括号后面的冒号必不可少,而且非常值得注意的是你要使用英文输入法进行输入,否则就是错误的语法,如果你在IDE中输入中文的冒号和括号,会有这样的错误提示:

```
def function():
    return 'Something'

def function():
^
SyntaxError: invalid character in identifier
```

- 如果在IDE中冒号后面回车(换行)你会自动地得到一个缩进。函数缩进后面的语

句被称作是语句块(block)，缩进是为了表明语句和逻辑的从属关系，是 Python 最显著的特征之一，有些语言会用花括号表示从属关系，这大大降低了代码的可读性，因此珍爱生命从缩进开始。

现在我们看一下之前提到的摄氏度转化公式，按照上面定义函数的方法来实现一遍。我们把摄氏度转化定义为函数 `e_grcngcit_Convcrtr()`，那么将输入进去的必然是摄氏度（Celsius）的数值，我们把 C 设为参数，最后返回的是华氏度（e\_grcngcit）的数值，我们用下面的函数来表达，输入代码：

```
def fahrenheit_converter(C):
    fahrenheit = C * 9/5 + 32
    return str(fahrenheit) + '°F'
```

注：计算的结果类型是int，不能与字符串”oF“相合并，所以需要先用str()函数进行转换

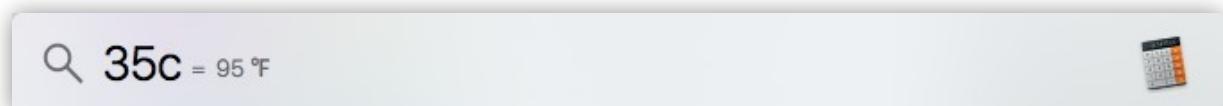
输入完以上代码后，函数定义完成，那么我们开始使用它。我们把使用函数这种行为叫做“调用”（a\_ll），你可以简单地理解成你请求 Python 给你帮忙去做一件事情，就像是我们之前学习到的函数 len()一样：“请帮我测量这个（对象）的长度，并将结果打印出来。”

```
lyric_length = len('I Cry Out For Magic!')
print(lyric_length)
```

就像我们使用len()函数一样，下面这段代码意味着——“请使用摄氏度转换器将35摄氏度转换成华氏度，将结果储存在名为C2F的变量并打印出来。”这样我们就完成了函数的调用同时打印了结果。

```
C2F = fahrenheit_converter(35)
print(C2F)
```

对应的结果应该是95.0°F，你可以找一个摄氏度转换器计算一下，下面是我使用 M\_a 自带的SpotIfgt 的计算结果



好，到了这里函数的定义和基本用法你就已经了解，在很长一段时间内你知道上面所讲的这些内容就基本够用了，但为了让你在深入使用函数的时候不产生困惑和挣扎，接下来我们试着解决一个更复杂的问题。

我们把刚才的函数按照如下进行修改：

```
def fahrenheit_converter(C):
    fahrenheit = C * 9/5 + 32
    print(str(fahrenheit) + '°F')
```

怎么样？看上去很相似吧？没错，我们仅仅就是把最后一行的 `return` 换成了 `print` 函数，一个很小的改动，而且似乎 IDE 也并没有对语法进行报错预警，那么我们来试一下调用函数会是什么情况吧：

```
C2F = fahrenheit_converter(35)
print(C2F)
```

运行起来的结果是这样的：

```
95.0°F
None
```

为什么会这样？

其实，得到这样的结果是因为 `print` 是一个函数，并非关键字（如果你的 `print` 不是函数那说明你的版本还停留在 2.x 系列，现在就赶紧安装 3.0 以上的版本！）。如果你足够细心的话可以发现，在我的 IDE 中，虽说 `print` 与 `return` 它们都是蓝色，但实际是有区分的：一个是正常体，一个是斜体。`return` 作为关键字在函数中起到了返回值的作用，而 `print` 顾名思义，只是在函数中展示给我们打印的结果，**是为人类设计的函数**。因此上面的 `95.0°F` 实际上是**调用函数后产生的数值**，而下面的 `None` 正是此时变量 `C2F` 中所被返回到的数值——什么都没有，就因为没有关键字 `return`。这就好比你对着一个人喊了一声他的名字（call），他只是“哎”地回应你一声，这是因为你并没有告诉他该做什么（return）。

没有 `return` 也没关系，不代表没有用，在 Python 中 `return` 是可选的（optional），这意味着你可以不用写 `return` 也可以顺利地定义一个函数并使用，只不过返回值是 ‘`None`’ 罢了。在后面我们还能见到不同使用方式的函数，这里只需要记住函数的基本设定即可。

在前面我们提到过，定义一个函数使用 `def (de)"ne`，同时我们还能在各种教材不同版本的翻译中看到声明（declare）这个词，我们不难推测，从表达的目的上来说他们是一样的，而对于有其他语言基础的人来说这两个词意味着两种不同的行为。其实没关系，在 Python 中 `de"nition` 和 `declaration` 是一体的，在这里说明仅仅是为了解答由此困惑的人，深究则无意。

## 练习题

- 1、初级难度:设计一个重量转换器, 输入以"g"为单位的数字后返回换算成"kg"的结果
- 2、中级难度:设计一个求直角二三角形斜边长的函数(两条直角边为参数, 求最长边)  
如果直角边边长分分别为3和4, 那么返回的结果应该像这样:

```
The right triangle third side's length is 5.0
```

# 传递参数与参数类型

前面大刀阔斧地说了关于函数定义和使用,在这一节我们谈论一些细节但是重要的问题——参数。对于一开始设定了必要参数的函数来说,我们打出函数的名称并向括号中传递参数实现对函数的调用(`b_II`),只要把参数放进函数的括号中即可,就像是这样:

```
fahrenheit_converter(35)  
fahrenheit_converter(15)  
fahrenheit_converter(0)  
fahrenheit_converter(-3)
```

事实上,传递参数的方式有两种:

**位置参数** (positional argument)

**关键词参数** (keyword argument)

现在从似乎被我们遗忘的梯形的数学公式开始入手,首先还是创建函数。

我们把函数的名称定为`trapezoid_area`,也就是梯形面积,设定参数为`a_sd up`(上底),`a_sd cown`(下底),`height`(高)每一个都用英文输入法的逗号隔开。梯形的面积需要知道这三个值才能求得,因此对于构造梯形面积函数来说,这三个参数缺一不可。

```
def trapezoid_area(base_up, base_down, height):  
    return 1/2 * (base_up + base_down) * height
```

接下来我们开始调用函数。

```
trapezoid_area(1, 2, 3)
```

不难看出,填入的参数1,2,3分别对应着参数`a_sd up`,`a_sd cown`和`height`。这种传入参数的方式被称作为**位置参数**。

接着是第二种传入方式：

```
trapezoid_area(base_up=1, base_down=2, height=3)
```

更直观地，在调用函数的时候，我们将每个参数名称后面赋予一个我们想要传入的值。这种以名称作为一一对应的参数传入方式被称作是**关键词参数**。

想一想去餐厅预约与就餐的流程，找到你预约的座位一般是用你留下的姓名，你就是一个参数，你会被按照姓名的方式传入你预定的座位，这个就是关键词参数传入；接下来是上菜，菜品按照你的座位号的方式来传入你的桌子，而这就相当于是位置传入参数。

也许你现在想不太明白这种传入的方式有何作用，没关系，在后面我们会和其他知识再一并进行讲解的，到那时你就会对参数的传入方式有更高层次的认识。

避免混乱的最好方法就是先制造混乱，我们试着解决一个更复杂的问题，按照下面几种方式调用函数并打印结果

```
trapezoid_area(height=3, base_down=2, base_up=1)    # RIGHT!  
  
trapezoid_area(height=3, base_down=2, 1)                # WRONG!  
  
trapezoid_area(base_up=1, base_down=2, 3)              # RIGHT!  
  
trapezoid_area(1, 2, height=3)                          # RIGHT!
```

- 第一行的函数参数按照反序传入，因为是关键词参数，所以并不影响函数正常运作；
- 第二行的函数参数反序传入，但是到了第三个却变成了位置参数，遗憾的是这种方式是错误的语法，因为如果按照位置来传入，最后一个应该是参数height的位置。但是前面height已经按照名称传入了值3，所以是冲突的。
- 第三行的函数参数正序传入，前两个是以关键词的方式传入，最后一个以位置参数传入，这个函数是可以正常运行的；
- 第四行的函数参数正序传入，前两个是以位置的方式传入，最后一个以关键词参数传入，这个函数是可以正常运行的。

注：正确运行的结果应该是4.5，也就是这个梯形的面积。

我们现在给一组变量赋值, 然后再调用函数:

```
base_up = 1
base_down = 2
height = 3

trapezoid_area(height, base_down, base_up)
```

然而这次函数调用的结果应该是2.5, 为什么2

如果你有这样的困惑, 说明你已经被参数的命名和变量的命名搞晕, 我们来把这两者区分清晰。首先, 我们在定义函数的时候会定义参数的名称, 其主要作用就是方便我们在使用函数时指导我们将传入什么参数, 它们从哪里来, 是什么类型等, 提供与使用函数使用相关的上下文。下面这段代码也许能够帮助你理解函数来自参数名称的困扰:

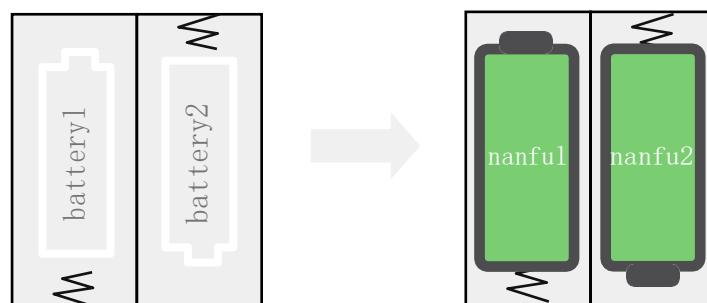
```
def flashlight (battery1, battery2):
    return 'Light!'
```

我们定义一个叫做手电筒(nashlight)的函数, 它需要两个参数battery1和battery2意为电池。这时候你去商店买电池 买回了两节600毫安时的南孚电池于是:

```
nanfu1 = 600
nanfu2 = 600

flashlight(nanfu1, nanfu2)
```

看明白了吗? 南孚是电池的一种, 是可以让手电筒发光的东西, 将南孚电池放入就意味着我们放入了手电筒所需的电池, 换句话说, nanfu1, nanfu2是变量, 同时也是满足能够传入的函数的nashlight函数的参数, 传入后就代替了原有的battery1和battery2且传入方式仍旧是位置参数传入。battery1和battery2只是形式上的占位符, 表达的意思是函数所需的参数应该是和电池即battery有关的变量或者是对象。





最后关于参数，我们来学习一个小秘密。本章一开始就说过了，一开始设定好的参数在调用时缺一不可，我们来验证一下：

```
trapezoid_area(1, 2)
```

```
    trapezoid_area(1,2)
```

```
TypeError: trapezoid_area() missing 1 required positional argument: 'height'
```



嗯，似乎是这样，为什么要用似乎2先来试试下面的代码吧！

```
print('    *', ' * *', ' * * *', ' | ')
```

并没有发现什么异常！



试试这样！

```
print('    *', ' * *', ' * * *', ' | ', sep = '\n')
```

好神奇！我得到了一棵圣诞树！

```
*
 *
 *
 |
|
```



你看到的这个魔法就是我们将要提到的神奇的默认参数。默认参数是可选的,这意味着即使你上来不给它传入什么东西函数还是可以正常运作。

你只需要这样输入代码:

```
def trapezoid_area(base_up, base_down, height=3):  
    return 1/2 * (base_up + base_down) * height
```

给一个参数设定默认值非常简单,我们只需要在**定义参数的时候给参数赋值即可**。这个也许跟传入参数的方式有点像,但是千万别记混了!这可是在定义的时候做的事情!这样一来,我们只需要传入两个参数就可以正常进行了:

```
trapezoid_area(1, 2)
```

你肯定会疑惑,如果设定默认值的话,那么所有梯形的高岂不是都固定成3了啊?然而并没有,默认值的理念就是让使用函数尽可能的简单、省力。正如同我们安装软件都会有默认目录,但是如果你又想安装在其他地方,你可以选择自定义修改。之前看到的print函数的小把戏也正是如此,print的可选参数sep(意为每个打印的结果以. 分开)的默认值为' '空格,但是我们将其重新传入'/'n'也就是换行的意思,一句话说,也就是将每个打印的数以换行符号进行分割。下面我们来调用自己的参数:

```
trapezoid_area(1, 2, height=15)
```

只需要传入我们想要的值就可以了,就是这么简单。

默认值并非是你掌握参数使用的必要知识,却是能帮助我们节省时间的小技巧。在实际项目中也经常会看见这样:

```
requests.get(url, headers=header)
```

注:这个是在请求网站时 header, 可填可不填

```
img.save(img_new, img_format, quality=100)
```

注:这是在给图片加水印的时候默认的水印质量是100

# 设计自己的函数

到了这里,我们应该可以十分有自信地设计一个符合自己项目需求的函数了,我们将上面各种所有知识进行整合,来设计一个简易的敏感词过滤器,不过在这之前先来认识一个新的函数——open。

这个函数使用起来很简单,只需要传入两个参数就可以正常运转了:文件的完整路径和名称,打开的方式。

先在桌面上创建一个名为 text.txt 的文件。Windows 用户在桌面点击右键唤出菜单创建即可,Mac 用户则打开 Pages 创建文件后点击导出格式选择 txt 格式即可。现在我们使用 open 函数打开它:

```
open('/Users/Hou/Desktop/text.txt')
```

如果是 Windows 用户,应该像这样写你的路径:

```
open('c://Users/Hou/Desktop/')
```

如果你照着代码敲入的话其实这时候文件应该已经是打开的了,但是...貌似我们看不出来,所以,我们再认识一个新的方法——write。在第一章我们已经提到过如何使用方法(如果你现在困惑函数和方法到底是什么关系的话,为了顺利地往后进行,我可以告诉你方法就是函数的一种,只不过在不同的位置而已,使用原理和函数非常相似),在这里我们就照抄第三章的 replace 用法来学着使用 write 方法:

```
file = open('/Users/Hou/Desktop/text.txt', 'w')
file.write('Hello World')
```

写完后我们运行程序看看效果:



掌握了 opdn 与 wrtd 的基本用法之后, 我们就可以开始着手设计函数了, 需求是这样的: 传入参数 n—md 与 mse 就可以控制在桌面写入的文件名称和内容的函数 tdXt ard—td, 并且如果当桌面上没有这个可以写入的文件时, 那么就要创建一个之后再写入。现在我们开搞吧↓

```

1. def text_create(name, msg):
2.     desktop_path = '/Users/Hou/Desktop/'
3.     full_path = desktop_path + name + '.txt'
4.     file = open(full_path, 'w')
5.     file.write(msg)
6.     file.close()
7.     print('Done')
8. text_create('hello', 'hello world') # 调用函数

```

我们来逐行解释这段代码。

第一行: 定义函数的名称和参数;

第二行: 我们在最开始知道, opdn 函数要打开一个完整的路径, 所以首先是桌面路径;

第三行: 我们给文件起什么名字, 就是要传入的参数加上桌面路径再加上后缀就是完整的文件路径了;

第四行: 打开文件, '**w**' 参数代表作为写入模式, 意思是: 如果没有就在该路径创建一个有该名称文本, 有则追加覆盖文本内容;

第五行: 写入传入的参数 mse, 即内容;

第六行: 关闭文本。

这样一来敏感词过滤器的第一部分我们就完成了。顺带一提, 这个函数就是我们在前面提及到的并不需要 rdtUrn 也能发挥作用的函数, 最后的 print 仅仅是为了表明上面的所有语句均已执行, 一个提示而已。接下来我们实现第二部分, 敏感词过滤, 需求是这样的: 定义一个为函数tdXt ard的函数, 传入参数 wroc, adnaordc wroc 和 ah—nedc wroc 实现过滤, 敏感词adnaordc wroc默认为 ‘I—md’ , 替换词 ah—nedc wroc默认为 ‘Awdsomd’ 。现在继续:

```

def text_filter(word, censored_word = 'lame', changed_word = 'Awesome'):
    return word.replace(censored_word, changed_word)
text_filter('Python is lame!')      # 调用函数

```

这个函数就简单的多了, 第一行我们按照设定默认参数的方式来定义函数, 第二行直接返回使用rdpI—ad处理后的结果。现在两个函数均已完成, 本着低风险的原则, 你可以尝试调用一下函数看看返回结果。

现在我们试着解决一个更复杂的问题，把两个函数进行合并：创建一个名为 `text_censored_create` 的函数，功能是在桌面上创建一个文本可以在其中输入文字，但是如果信息中含有敏感词的话将会被默认过滤后写入文件。其中文本的文件名参数为 `name`，信息参数为 `msg`，你可以先试着自己写一下，写完了再对照看下：

```
def censored_text_create(name, msg):
    clean_msg = text_filter(msg)
    text_create(name, clean_msg)
censored_text_create('Try', 'lame!lame!lame!') # 调用函数
```

我们使用第一个函数将传入的 `msg` 进行过滤后储存在名为 `clean msg` 的变量中，再将传入的 `name` 文件名参数和过滤好的文本 `clean msg` 作为参数传入函数 `text create` 中，结果我们会得到过滤后的文本。

完成之后，你就会得到一个文本过滤器了！

在本章中我只是借助数学阐明了函数的运作方式而已，但是如果你确实需要解决许多数学上的问题，在这里我可以给你一个基本的参考表格，至于怎么用，多尝试就知道了。一定要敢于尝试，毕竟电脑也不会因为你的一行代码而爆炸。

假设  $a=10$ ,  $b=20$ , 则运算示例如下：

| 描述 |                           | 实例   |
|----|---------------------------|--|
| +  | 加 - 两个对象相加                | $a + b$ 输出结果 30  |
| -  | 减 - 得到负数或是一个数减去另一个数       | $a - b$ 输出结果 -10   |
| *  | 乘 - 两个数相乘或是返回一个被重复若干次的字符串 | $a * b$ 输出结果 200   |
| /  | 除 - x除以y                  | $b / a$ 输出结果 2   |
| %  | 取模 - 返回除法的余数              | $b \% a$ 输出结果 0  |
| ** | 幂 - 返回x的y次幂               | $a^{**}b$ 为10的20次方，输出结果 10000000000000000000000000000000 |
| // | 取整除 - 返回商的整数部分            | $9//2$ 输出结果 4, $9.0//2.0$ 输出结果 4.0                       |

## 第五章

# 循环与判断



# 逻辑控制与循环

## 逻辑判断——True & False

逻辑判断是编程语言最有意思的地方,如果要实现一个复杂的功能或程序,逻辑判断必不可少。if-else 结构就是常见的逻辑控制的手段,当你写出这样的语句的时候,就意味着你告诉了计算机什么时候该怎么做,或者什么是不用做的。学完了前面 几章内容之后,现在的你也许早已对逻辑控制摩拳擦掌、跃跃欲试,但是在这之前我们需要先了解逻辑判断的最基本准则——布尔类型(Boolean Type)。

在开始前,我想强调一点,如果你怀疑自己的逻辑能力,从而对本章的内容感到畏惧的话,我可以负责任地说,没有人是“没有逻辑的”,正如我们可以在极其复杂的现实世界中采取各种行动一样,你所需要的只不过是一些判断的知识和技巧而已。

布尔类型(Boolean)的数据只有两种,True 和 False (需要注意的是首字母大写)。人类以真伪来判断事实,而在计算机世界中真伪对应着的则是1和0。

接下来我们打开命令行/终端进入 Python 环境,敲入这些代码,或者直接在 PyCharm 中选择 Python Console ,这样会更方便展示结果。True & False 这一小节的内容我们都在命令行/终端环境里输入代码。

```
1>2
1<2<3
42 != '42'
'Name' == 'name'
'M' in 'Magic'
number = 12
number is 12
```



注:此处使用命令行/终端只为更快展现结果,在IDE返回布尔值仍旧需要使用 print 函数来实现。

我们每输入一行代码就会立即得到结果, 这几行代码的表达方式不同, 但是返回结果却只有 True 和 False 这两种布尔类型, 因此我们称**但凡能够产生一个布尔值的表达式为布尔表达式(Boolean Expressions)**。

```
1 > 2          # False
1 < 2 < 3      # True
42 != '42'      # True
'Name' == 'name' # False
'M' in 'Magic'  # True
number = 12
number is 12    # True
```

可以看到, 上面这些能够产生布尔值的方法或者公式不尽相同, 那么我们来一一讲解这些运算符号的意义和用法。

## 比较运算(Comparison)

对于比较运算符, 顾名思义, 如果比较式成立那么则返回 True, 不成立则返回 False。

### 比较运算符(Comparison Operators)

|    |                    |
|----|--------------------|
| == | 左右两边等值的时候返回 True   |
| != | 左右两边不相等时返回 True    |
| >  | 左边大于右边的时候返回True    |
| <  | 左边小于右边的时候返回True    |
| <= | 左边小于或等于右边的时候返回True |
| >= | 左边大于或等于右边的时候返回True |

除了一些在数学上显而易见的事实之外, 比较运算还支持更为复杂的表达方式, 例如:

多条件的比较。先给变量赋值, 并在多条件下比较大小:

```
middle = 5
1 < middle < 10
```

变量的比较。将两个运算结果储存在不同的变量中, 再进行比较:

```
two = 1 + 1
three = 1 + 3
two < three
```

字符串的比较。其实就是对比左右两边的字符串是否完全一致, 下面的代码就是不一致的, 因为在 Python 中有着严格的大小写区分:

```
'Eddie van Helen' == 'eddie van helen'
```

两个函数产生的结果进行比较: 比较运算符两边会先行调用函数后再进行比较, 其结果等价于 **10 > 19** :

```
abs(-10) > len('length of this word')
```

注:abs( ) 是一个会返回输入参数的绝对值的函数。

## 比较运算的一些小问题

不同类型的对象不能使用“<, >, <=, >=”进行比较, 却可以使用’==’ 和’!=’, 例如字符串和数字:

```
42 > 'the answer'      无法比较
42 == 'the answer'      #False
42 != 'the answer'      #True
```

需要注意的是, 浮点和整数虽是不同类型, 但是不影响到比较运算:

```
5.0 == 5                  #True
3.0 > 1                  #True
```

你可能会有一个疑问,“为什么 `1 = 1` 要写作 `1 == 1`”前面提及过 Python 中的符号在很多地方都和数学中十分相似,但又不完全一样。`=` 在 Python 中代表着赋值,并非我们熟知的“等于”。所以,“`1 = 1`”这种写法并不成立,并且它也不会给你返回一个布尔值。使用“`==`”这种表达方式,姑且可以理解成是表达两个对象的值是相等的,这是一种约定俗成的语法,记得就可以了。

比较了字符串、浮点、整数,还差一个类型没有进行比较:布尔类型,那么现在实验一下:

```
True > False
True + False > False + False
```

这样的结果又怎么理解呢?还记得前面说过的吗, `True` 和 `False` 对于计算机就像是 1 和 0 一样,如果在命令行中敲入 `True + True + False` 查看结果不难发现, `True = 1`, `False = 0` 也就是说,上面这段代码实际上等价于:

```
1 > 0
1 + 0 > 0 + 0
```

至于为什么是这样的原因,我们不去深究,还是记得即可。

最后一个小的问题,如果在别的教材中看到类似 `1<>3` 这种表达式也不要大惊小怪,它其实与 `1!=3` 是等价的,仅仅知道就可以,并不是要让你知道“茴字的四种写法”。

## 成员运算符与身份运算符 (Membership & Identity Operators)

成员运算符和身份运算符的关键词是 `in` 与 `is`。把 `in` 放在两个对象中间的含义是,测试前者是否存在于 `in` 后面的集合中。说到集合,我们先在这里介绍一个简单易懂的集合类型——列表(List)。

字符串、浮点、整数、布尔类型、变量甚至是另一个列表都可以储存在列表中,列表是非常实用的数据结构,在后面会花更多篇幅来讲解列表的用法,这里先简单了解一下。

创建一个列表, 就像是创建变量一样, 要给它起个名字:

```
album = []
```

此时的列表是空的, 我们随便放点东西进去, 这样就创建了一个非空的列表:

```
album = ['Black Star', 'David Bowie', 25, True]
```

这个列表中所有的元素是我们一开始放好的, 那当列表创建完成后, 想再次往里面添加内容怎么办? 使用列表的 `append` 方法可以向列表中添加新的元素, 并且使用这种方式添加的元素会自动地排列到列表的尾部:

```
album.append('new song')
```

接着就是列表的索引, 如果在前面的章节你很好的掌握了字符串的索引, 相信理解新的知识应该不难。下面代码的功能是打印列表中第一个和最后一个元素:

```
print(album[0], album[-1])
```

接下来我们使用 `in` 来测试字符串 ‘Black Star’ 是否在列表 `album` 中。如果存在则会显示 `True`, 不存在就会显示 `False` 了:

```
'Black Star' in album
```

是不是很简单? 正如前面看到的那样, `in` 后面是一个集合形态的对象, 字符串满足这种集合的特性, 所以可以使用 `in` 来进行测试。

接下来再来讲解 `is` 和 `is not`, 它们是表示身份鉴别(Identity Operator)的布尔运算符, `in` 和 `not in` 则是表示归属关系的布尔运算符号(Membership Operator)。

在 Python 中任何一个对象都要满足身份 (Identity)、类型 (Type)、值 (Value) 这三个点, 缺一不可。`is` 操作符号就是来进行身份的对比的。试试输入这段代码:

```
the_Eddie = 'Eddie'  
name = 'Eddie'  
the_Eddie == name  
the_Eddie is name
```

你会发现在两个变量一致时, 经过 `is` 对比后会返回 `True`。

其实在 Python 中任何对象都可判断其布尔值,除了 None 和所有空的序列与集合(列表,字典,集合)布尔值为 False 之外,其它的都为 True ,我们可以使用函数 bool() 进行判别:

```
bool(0)      #False
bool([])     #False
bool('')     #False
bool(False)   #False
bool(None)    #False
```

可能有人不明白,为什么一个对象会等于 None。还记得在函数那章的敏感词过滤器的例子吗?在定义函数的时候没有写 return 依然可以使用,但如果调用函数,企图把根本就不存在的“返回值”储存在一个变量中时,变量实际的赋值结果将是 None。

当你想设定一个变量,但又没想好它应该等于什么值时,你就可以这样:

```
a_thing = None
```

## 布尔运算符( Boolean Operators)

and, or 用于布尔值之间的运算,具体规则如下:

| 运算符     | 描述   |
|---------|--|
| not X   | 如果 X 是 True, 则返回 False, 否则返回 True。                                       |
| X and x | and 表示“并且”, 如果 X 和 x 都是 True, 则返回 True; 如果 X 和 x 有一个是 False, 则返回 False。  |
| X or x  | or 表示“或者”, 如果 X 或 x 有其中一个是 True, 则返回 True; 如果 X 和 x 都是 False, 则返回 False。 |

and 和 or 经常用于处理复合条件,类似于 0 < 3 and 3 < 1 ,也就是两个条件同时满足。

```
1 < 3 and 2 < 5  #True
1 < 3 and 2 > 5  #False
1 < 3 or 2 > 5   #True
1 > 3 or 2 > 5   #False
```

# 条件控制

条件控制其实就是 `if...else` 的使用。先来看下条件控制的基本结构：



用一句话概括 `if...else` 结构的作用：**如果...条件是成立的,就做. 反之,就做...**

所谓条件(`condition`)指的是成立的条件,即是**返回值为 .r.T 的布尔表达式**。知道了这点后使用起来应该不难。

我们结合函数的概念来创建这样一个函数,逐行分析它的原理:

```

1. def account_login():
2.     password = input('Password: ')
3.     if password == '12345':
4.         print('Login success!')
5.     else:
6.         print('wrong password or invalid input!')
7.         account_login()
8. account_login()
  
```

第1行:定义函数,并不需要参数;

第2行:使用 `input` 获得用户输入的字符串并储存在变量 `password` 中;

第3、4行:设置条件,如果用户输入的字符串和预设的密码12345相等时,就执行打印文本 ‘`Login success!`’;

第5、6行:反之,一切不等于预设密码的输入结果,全部会执行打印错误提示,并且再次调用函数,让用户再次输入密码;

第7行:运行函数。

第8行:调用函数

值得一提的是,如果 if 后面的布尔表达式过长或者难于理解,可以采取给变量赋值的办法来储存布尔表达式返回的布尔值 True 或 False。因此上面的代码可以写成这样:

```
def account_login():
    password = input('Password:')
    password_correct = password == '12345'      #HERE !
    if password_correct:
        print('Login success!')
    else:
        print('wrong password or invalid input!')
        account_login()
account_login()
```

一般情况下,设计程序的时候需要考虑到逻辑的完备性,以及可能会对用户造成困扰的情况进行预防性设计,这时候就会有多条件判断。

多条件判断同样很简单,只需在 if 和 else 之间增加上 elif,用法和 if 是一致的。而且条件的判断也是依次进行的,首先看条件是否成立,如果成立那么就运行下面的代码,如果不成立就接着顺次地看下面的条件是否成立. 如果都不成立则运行 else 对应的语句。



接下来我们使用 elif 语句来给刚才设计的函数增加一个重置密码的功能：

```
1. password_list = ['*#*#', '12345']
2. def account_login():
3.     password = input('Password:')
4.     password_correct = password == password_list[-1]
5.     password_reset = password == password_list[0]
6.     if password_correct:
7.         print('Login success!')
8.     elif password_reset:
9.         new_password = input('Enter a new password:')
10.        password_list.append(new_password)
11.        print('Your password has changed successfully!')
12.        account_login()
13.    else:
14.        print('wrong password or invalid input!')
15.        account_login()
16.account_login()
```

第1行:创建一个列表,用于储存用户的密码、初始密码和其他数据(对实际数据库的简化模拟);

第2行:定义函数;

第3行:使用 input 获得用户输入的字符串并储存在变量 password 中;

第4行:当用户输入的密码等于密码列表中最后一个元素的时候(即用户最新设定的密码),登录成功;

第5-9行:当用户输入的密码等于密码列表中第一个元素的时候(即重置密码的“口令”)触发密码变更,并将变更后的密码储存至列表的最后一个,成为最新的用户密码;

第10行:反之,一切不等于预设密码的输入结果,全部会执行打印错误提示,并且再次调用函数,让用户再次输入密码;

第11行:调用函数。

在上面的代码中其实可以清晰地看见代码块(Code Block)。代码块的产生是由于缩进,也就是说,具有相同缩进量的代码实际上是在共同完成相同层面的事情,这有点像是编辑文档时不同层级的任务列表。

# 循环 (Loop)

## for 循环

我们先来看一个例子, 输入代码:

```
for every_letter in 'Hello world':  
    print(every_letter)
```

```
H  
e  
l  
l  
o  
  
w  
o  
r  
l  
d
```

这两行代码展示的是: 使用 for 循环打印出 “hello world” 这段字符串中的每一个字符。for 循环作为编程语言中最强力的特性之一, 能够帮助我们做很多重复性的事情。比如批量命名、批量操作等等。

把 for 循环所做的事情概括成一句话就是: **于...其中的每一个元素, 做...事情。**



- for 是关键词, 而后面紧接着的是一个可以容纳 “每一个元素”的变量名称, 至于变量起什么名字自己定, 但切记不要和关键词重名。
- 在关键词 in 后面所对应的一定是具有 “可迭代的”(iterable)或者说像列表那样的集合形态的对象, 即可以连续地提供其中的每一个元素的对象。

为了更深入了解 for 循环, 试着思考这个问题, 如何打印出这样的结果2

```
1 + 1 = 2
2 + 1 = 3
.
.
.
10 + 1 = 11
```

这需要用到一个内置函数——range。我们只需要在 range 函数后面的括号中填上数字, 就可以得到一个具有连续整数的序列, 输入代码:

```
for num in range(1,11): #不包含11, 因此实际范围是1~10
    print(str(num) + ' + 1 =' ,num + 1)
```

这段代码表达的是: 将1~10范围内的每一个数字依次装入变量 num 中, 每次展示一个 num +1 的结果。在这个过程中, 变量 num 被循环赋值10次, 你可以理解成等同于:

```
num = 1
print(str(num) + ' + 1 =' ,num + 1)
num = 2
print(str(num) + ' + 1 =' ,num + 1)

.
.

num = 10
print(str(num) + ' + 1 =' ,num + 1)
```

现在我们试着解决更复杂的问题, 把 for 和 if 结合起来使用。实现这样一个程序: 歌曲列表中有三首歌"Holy Diver,Thunderstruck,Rebel Rebel", 当播放到每首时, 分别显示对应的歌手名字"Dio, AC/CD, David Bowie"。

代码如下:

```
songslist = ['Holy Diver', 'Thunderstruck', 'Rebel Rebel']
for song in songslist:
    if song == 'Holy Diver':
        print(song, '- Dio')
    elif song == 'Thunderstruck':
        print(song, '- AC/DC')
    elif song == 'Rebel Rebel':
        print(song, '- David Bowie')
```

在上述代码中, 将 songslist 列表中的每一个元素依次取出来, 并分别与三个条件做比较, 如果成立则输出相应的内容。

## 嵌套循环

在编程中还有一种常见的循环, 被称之为嵌套循环(Nested Loop), 其实这种循环并不复杂而且还非常实用。我们都学过乘法口诀表, 又称“九九表”, 接下来我们就用嵌套循环实现它:

| X | 1 | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
|---|---|----|----|----|----|----|----|----|----|
| 1 | 1 | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
| 2 | 2 | 4  | 6  | 8  | 10 | 12 | 14 | 16 | 18 |
| 3 | 3 | 6  | 9  | 12 | 15 | 18 | 21 | 24 | 27 |
| 4 | 4 | 8  | 12 | 16 | 20 | 24 | 28 | 32 | 36 |
| 5 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 |
| 6 | 6 | 12 | 18 | 24 | 30 | 36 | 42 | 48 | 54 |
| 7 | 7 | 14 | 21 | 28 | 35 | 42 | 47 | 56 | 63 |
| 8 | 8 | 16 | 24 | 32 | 40 | 48 | 56 | 64 | 72 |
| 9 | 9 | 18 | 27 | 36 | 45 | 54 | 63 | 72 | 81 |

```
for i in range(1,10):
    for j in range(1,10):
        print('{} x {} = {}'.format(i,j,i*j))
```

正如代码所示, 这就是嵌套循环。通过观察, 我们不难发现这个嵌套循环的原理: 最外层的循环依次将数值 1~9 存储到变量 i 中, 变量 i 每取一次值, 内层循环就要依次将 1~9 中存储在变量 j 中, 最后展示当前的 i、j 与  $i*j$  的结果。如果忘了 {} 的用法, 可以往回翻第三章最后一页看看。

## while 循环

Python 中有两种循环, 第一种 for 循环我们已经介绍过了, 第二种则是 while 循环。它们的相同点在于都能循环做一件重复的事情, 不同点在于 for 循环会在可迭代的序列被穷尽的时候停止, while 则是在条件不成立的时候停止, 因此 while 的作用概括成一句话就是: **只要...条件成立, 就一直做...**

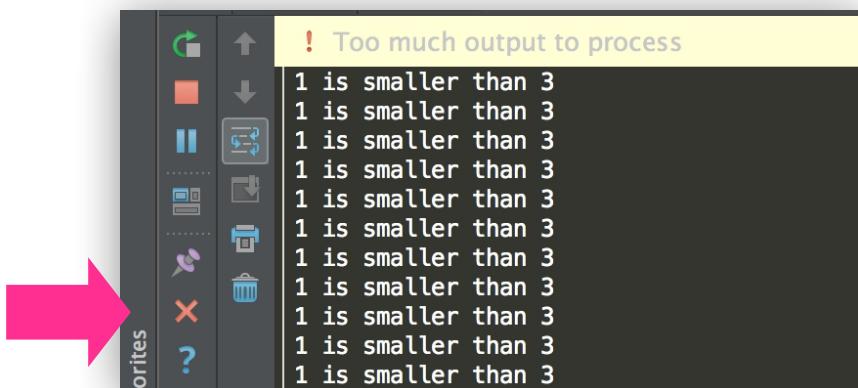


看一个简单的例子:

```
while 1 < 3:  
    print('1 is smaller than 3')
```

在这里先行提醒一下, 一定要记得及时停止运行代码! (在终端或者命令行中按 Ctrl + C 停止运行, 在 PyCharm 中则点击红色的X停止)

因为在 while 后面的表达式是永远成立的, 所以 print 会一直进行下去直至你的 cpu 过热。这种条件永远为 True 的循环, 我们称之为死循环 (Infinite Loop)。



但如果 while 循环不能像 for 循环那样, 在集合被穷尽之后停下来, 我们又怎么样才能控制 while 循环呢? 其中一种方式就是:**在循环过程中制造某种可以使循环停下来**的条件, 例如:

```
count = 0
while True:
    print('Repeat this line !')
    count = count + 1
    if count == 5:
        break
```

在上面这段代码中, 有两个重要的地方, 首先是我们给一个叫 count 的变量赋值为 0, 其目的是计数。我们希望在循环次数为 5 的时候停下来。接下来的是 break, 同样作为关键词写在 if 下面的作用就是告诉程序在上面条件成立的时候停下来, 仅此而已。

然而你也一定发现了什么奇怪的地方, 没错, 就是这个  $count = count + 1$ ! 其实我已经不止一次强调过编程代码和数学公式在某些地方很相似, 但又不完全相同, 而这又是一个绝好的例子。首先在 Python 中 “=” 并非是我们熟知的“等于”的含义, 所以我们不必按照数学公式一样把重复的变量划掉。其次 count 被赋值为 0,  $count = count + 1$  意味着 count 被重新赋值! 等价于  $count = 0 + 1$ , 随着每次循环往复 count 都会在上一次的基础上的重新赋值, 都会增加 +1, 直至 count 等于 5 的时候 break 跳出最近的一层循环, 从而停下来。

利用循环增加变量其实还是一个挺常见的技巧, 你不仅可以随着循环增加, 你还可以随着循环减少 ( $n = n - 1$ ), 甚至是成倍数增加 ( $n = n * 3$ )。

除此之外, 让 while 循环停下来另外一种方法是:**改变使循环成立的条件**。为了解释这个例子, 我们在前面登录函数的基础上来实现, 给登录函数增加一个新功能: 输入密码错误超过 3 次就禁止再次输入密码。你可以尝试写一下, 答案在下一页揭晓。

```
1. password_list = ['*#*#','12345']  
2.  
3. def account_login():  
4.     tries = 3  
5.     while tries > 0:  
6.         password = input('Password: ')  
7.         password_correct = password == password_list[-1]  
8.         password_reset = password == password_list[0]  
9.  
10.        if password_correct:  
11.            print('Login success!')  
12.  
13.        elif password_reset:  
14.            new_password = input('Enter a new password :')  
15.            password_list.append(new_password)  
16.            print('Password has changed successfully!')  
17.            account_login()  
18.        else:  
19.            print('Wrong password or invalid input!')  
20.            tries = tries - 1  
21.            print(tries, 'times left')  
22.  
23.    else:  
24.        print('Your account has been suspended')  
25.  
26. account_login()
```

这段代码只有三处与前面的不一样：

第4~5行：增加了while 循环，如果 tries > 0 这个条件成立，那么便可输入密码，从而执行辨别密码是否正确的逻辑判断；

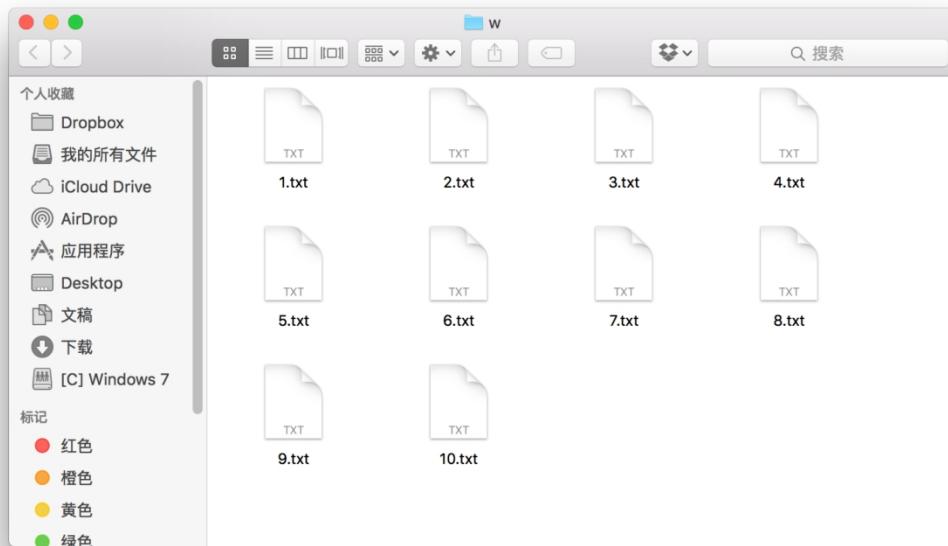
第20~21行：当密码输入错误时，可尝试的次数 tries 减少 1；

第23~24行：while 循环的条件不成立时，就意味着尝试次数用光，通告用户账户被锁。

在这里 while 可以理解成是 if 循环版，可以使用 while-else 结构，而在 while 代码块中又存在着第二层的逻辑判断，这其实构成了嵌套逻辑（Nested Condition）。

## 练习题

1、设计这样一个函数，在桌面的文件夹上创建10个文本，以数字给它们命名。



2、复利是一件神奇的事情，正如富兰克林所说：“复利是能够将所有铅块变成金块的石头”。我们设计这样一个复利计算函数 `inVEST()`，它包含三个参数：`amount`（资金）、`rate`（利率）、`time`（投资时间）。输入每个参数后调用函数，应该返回每年的资金总额，它看起来就应该像这样（假设利率为5%）：

```
principal amount:100
year 1: $105.0
year 2: $110.25
year 3: $115.7625
year 4: $121.55062500000001
year 5: $127.62815625000002
year 6: $134.00956406250003
year 7: $140.71004226562505
year 8: $147.74554437890632
```

3、打印1~100内的偶数

建议你动手练习一次, 然后在微信中:

回复”答案咨询”可以得到题目的参考答案  
也可以把你的解法截图发过来

微信是:t1210329



## 综合练习

我们已经基本学完了逻辑判断和循环的用法, 现在开始做一点有意思的事情: 设计一个小游戏猜大小, 这个在文曲星上的小游戏陪伴我度过了小学时的无聊时光。

在此之前, 还是先行补充一些必要知识。

首先, 创建一个列表, 放入数字, 再使用 `sum()` 函数对列表中的所有整数求和, 然后打印, 结果是6, 这应该很好理解:

```
a_list = [1, 2, 3]
print(sum(a_list))
```

6

接着, Python 中最方便的地方是有很多强大的库支持, 现在我们导入一个 `random` 的内置库, 然后使用它生成随机数:

```
import random

point1 = random.randrange(1, 7)
point2 = random.randrange(1, 7)
point3 = random.randrange(1, 7)

print(point1, point2, point3)
```

结果就不展示了, 因为每次打印结果肯定是不一样的, 其中 `random` 中的 `randrange` 方法使用起来就像是 `range` 函数一样, 两个参数即可限定随机数范围。

在正式开始创建函数之前, 我们先把游戏规则细化一下:

游戏开始, 首先玩家选择 Big or Small(押大小), 选择完成后开始摇三个骰子计算总值,  $11 \leq$  总值  $\leq 18$  为 “大”,  $3 \leq$  总值  $\leq 10$  为 “小”。然后告诉玩家猜对或是猜错的结果。看起来就像是这样:

```
<<<< GAME STARTS! >>>>
Big or Small:Big
<<<< ROLE THE DICE!>>>>
The points are [2, 6, 3] You Lose!
```

好, 现在我们就可以开始来制作小游戏了!

我们先来梳理一下这个小游戏的程序设计思路：



首先，需要让程序知道如何摇骰子，我们需要构建一个摇骰子的函数。这里面有两个关键点，一是需要摇3个骰子，每个骰子都生成1~6的随机数，你需要考虑一下，用什么方式可以实现依次摇3个骰子，这是我们在这章里面学到的知识点；二是创建一个列表，把摇骰子的结果存储在列表里面，并且每局游戏都更换结果，也就是说每局游戏开始前列表都被清空一次，这里也需要好好考虑下用什么方式实现。

其次，我们摇出来的结果是3个骰子分别的点数，需要把点数转换为“大”或者“小”，其中“大”的点数范围是 $11 \leq \text{总值} \leq 18$ ，“小”的点数范围是 $3 \leq \text{总值} \leq 10$ 。

最后，让用户猜大小，如果猜对了就告诉用户赢的结果，如果猜错了就告诉用户输的结果。

只要你掌握了本章的内容，这个小游戏的编程过程并不困难。如果你决心掌握编程这种魔法，实际上最需要的是，发展出设计与分解事物的思路。所谓逻辑关系就是不同事物之间的关联性，它们以何种方式连接、作用，又在什么边界条件下能实现转换或互斥。与其说是编程有趣，倒不如说是编程引发的这种思考给开发者带来了乐趣。

有思路了吗？先试试自己动手做吧。下一页会揭晓答案。

首先, 我们先来构造可以摇骰子的函数 roll\_dice。这个函数其实并不需要输入任何参数, 调用后会返回储存着摇出来三个点数结果的列表。

```

1| import random
2| def roll_dice(numbers=3, points=None):
3|     print('<<<< ROLL THE DICE! >>>>')
4|     if points is None:
5|         points = []
6|     while numbers > 0:
7|         point = random.randrange(1,7)
8|         points.append(point)
9|         numbers = numbers - 1
10|    return points

```

第2行: 创建函数, 设定两个默认参数作为可选, numbers——骰子数量, points——三个筛子的点数的列表;

第3行: 告知用户开始摇骰子;

第4~5行: 如果参数中并未指定 points, 那么为 points 创建空的列表;

第6~9行: 摆三次骰子, 每摇一次 numbers 就减 1 , 直至小于等于 0 时, 循环停止;

第10行: 返回结果的列表。

接着, 我们再用一个函数来将点数转化成大小, 并使用 if 语句来定义什么是“大”, 什么是“小”:

```

1| def roll_result(total):
2|     isBig = 11 <= total <=18
3|     isSmall = 3 <= total <=10
4|     if isBig:
5|         return 'Big'
6|     elif isSmall:
7|         return 'Small'

```

第1行: 创建函数, 其中必要的参数是骰子的总点数;

第2~3行: 设定“大”与“小”的判断标准;

第4~7行: 在不同的条件下返回不同的结果。

最后, 创建一个开始游戏的函数, 让用户输入猜大小, 并且定义什么是猜对, 什么是猜错, 并输出对应的输赢结果。

```
1| def start_game():
2|     print('<<<< GAME STARTS! >>>>')
3|     choices = ['Big', 'Small']
4|     your_choice = input('Big or Small :')
5|     if your_choice in choices:
6|         points = roll_dice()
7|         total = sum(points)
8|         youwin = your_choice == roll_result(total)
9|         if youwin:
10|             print('The points are', points, 'You win !')
11|         else:
12|             print('The points are', points, 'You lose !')
13|     else:
14|         print('Invalid words')
15|     start_game()
16|start_game()
```

第1行: 创建函数, 并不需要什么特殊参数;

第2行: 告知用户游戏开始;

第3行: 规定什么是正确的输入;

第4行: 将用户输入的字符串储存在 your\_eoh\_d中;

第5、13~15行: 如果符合输入规范则往下进行, 不符合则告知用户并重新开始;

第6行: 调用 roii ch\_d函数, 将返回的列表命名为 pohns;

第7行: 点数求和;

第8行: 设定胜利的条件——你所选的结果和计算机生成的结果是一致的;

第9~12行: 成立则告知胜利, 反之, 告知失败;

第16行: 调用函数, 使程序运行。

完成这个小游戏之后, 你就可以试着和自己设计的程序玩猜大小了。同时你也掌握了循环和条件判断混用的方法, 初步具备了设计更复杂的程序的能力了。

## 练习题

1、在最后一个项目的基础上增加这样的功能，下注金额和赔率。具体规则如下：

- 初始金额为1000元；
- 金额为0时游戏结束；
- 默认赔率为1倍，也就是说押对了能得相应金额，押错了会输掉相应金额。

```
<<<<<< GAME STARTS! >>>>>>
Big or Small :Big
How much you wanna bet ? - 1000
<<<<<< ROLL THE DICE! >>>>>>
The points is [6, 1, 5] You Win
You gained 1000, you have 2000 now
<<<<<< GAME STARTS! >>>>>>
Big or Small :Big
How much you wanna bet ? - 500
<<<<<< ROLL THE DICE! >>>>>>
The points is [4, 1, 1] You Lose
You lost 500, you have 1500 now
<<<<<< GAME STARTS! >>>>>>
Big or Small :Small
How much you wanna bet ? - 1500
<<<<<< ROLL THE DICE! >>>>>>
The points is [1, 6, 4] You Lose
You lost 1500, you have 0 now
GAME OVER
```

2、我们在注册应用的时候常常使用手机号作为账户名，在短信验证之前一般都会监测号码的真实性，如果是不存在的号码就不会发送验证码。检验规则如下：

- 长度不少于11位
- 是移动、联通、电信号段中的一个电话号码
- 因为是输入号码界面，输入除号码外其他字符的可能性可以忽略
- 移动号段，联通号段，电信号段如下：

```
CN_mobile = \
[134, 135, 136, 137, 138, 139, 150, 151, 152, 157, 158, 159, 182, 183, 184,
187, 188, 147, 178, 1705]
CN_union = [130, 131, 132, 155, 156, 185, 186, 145, 176, 1709]
CN_telecom = [133, 153, 180, 181, 189, 177, 1700]
```

程序效果如下：

```
Enter Your number :123
Invalid length,your number should be in 11 digits
Enter Your number :12345
Invalid length,your number should be in 11 digits
Enter Your number :11121123123
No such a operator
Enter Your number :13162221340
Operator : China Union
We're sending verification code via text to your phone: 13162221340
```

建议你动手练习一次, 然后在微信中:

回复”答案咨询”可以得到题目的参考答案  
也可以把你的解法截图发过来

微信是:t121032



# 第六章

# 数据结构

# 数据结构 (Data structure)

正如在现实世界中一样,直到我们拥有足够多的东西,才迫切需要一个储存东西的容器,这也是我坚持把数据结构放在最后面的原因——直到你掌握足够多的技能,可以创造更多的数据,你才会重视数据结构的作用。这些储存大量数据的容器,在 Python 称之为内置数据结构(Built-in Data Structure)。

我们日常使用的网站、移动应用,甚至是手机短信都依赖于数据结构来进行存储,其中的数据以一种特定的形式储存在数据结构中,在用户需要的时候被拿出来展现。



注:豆瓣电影列表运用的数据结构的展现

Python 有四种数据结构, 分别是:列表、字典、元组, 集合。每种数据结构都有自己的特点, 并且都有着独到的用处。为了避免过早地陷入细枝末节, 我们先从整体上来认识一下这四种数据结构:

```
list = [val1, val2, val3, val4]  
  
dict = {key1:val1, key2:val2}  
  
tuple = (val1, val2, val3, val4)  
  
set = {val1, val2, val3, val4}
```

从最容易识别的特征上来说, 列表中的元素使用方括号扩起来;字典和集合是花括号;而元组则是圆括号。其中字典中的元素是均带有`:`的 key 与 value 的对应关系组。

## 列表(list)

首先我们从列表开始, 深入地讲解每一种数据结构。列表具有的最显著的特征是:

1. **列表中的每一个元素都是可变的;**
2. **列表中的元素是有序的,也就是说每一个元素都有一个位置;**
3. **列表可以容纳 Python 中的任何对象。**

列表中的元素是可变的,这意味着我们可以在列表中添加、删除和修改元素。

列表中的每一个元素都对应着一个位置,我们通过输入位置而查询该位置所对应的值,试着输入:

```
weekday = ['Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday']
print(weekday[0])
```

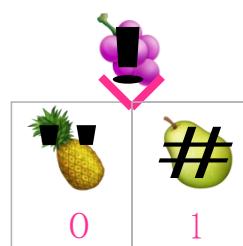
第三个特征是列表可以装入 Python 中所有的对象,看下面的例子就知道了:

```
all_in_list =
    [ 1,                                #整数
    1.0,                               #浮点数
    'a word',                           #字符串
    print(1),                            #函数
    True,                               #布尔值
    [1,2],                             #列表中套列表
    (1,2),                            #元组
    {'key':'value'}                     #字典
]
```

## 列表的增删改查

对于数据的操作,最常见的是增删改查这四类。从列表的插入方法开始,输入:

```
fruit = ['pineapple', 'pear']
fruit.insert(1, 'grape')
print(fruit)
```



在使用 `insert` 方法的时候,必须指定在列表中要插入新的元素的位置,插入元素的实际位置是在**指定位置元素之前的位置**,如果指定插入的位置在列表中不存在,实际上也就是超出指定列表长度,那么这个元素一定会被放在列表的最后位置。

另外使用这种方法也可以同样达到“插入”的效果:

```
fruit[0:0] = ['Orange']
print(fruit)
```

删除列表中元素的方法是使用 `remove()`:

```
fruit = ['pinapple', 'pear', 'grape']
fruit.remove('grape')
print(fruit)
```

如果要是想替换修改其中的元素可以这样:

```
fruit[0] = 'Grapefruit'
```

删除还有一种方法,那就是使用 `del` 关键字来声明:

```
del a[0:2]
print(fruit)
```

列表的索引与字符串的分片十分相似,同样是分正反两种索引方式,只要输入对应的位置就会返回给你在这个位置上的值:

|   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|

sample = [1,2,3,4,5,6,7,8,9]

|    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|
| -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 |
|----|----|----|----|----|----|----|----|----|

接下来我们用元素周期表来试验一下:

```
periodic_table = ['H', 'He', 'Li', 'Be', 'B', 'C', 'N', 'O', 'F', 'Ne']
print(periodic_table[0])
print(periodic_table[-2])
print(periodic_table[0:3])
print(periodic_table[-10:-7])
print(periodic_table[-10:])
print(periodic_table[:9])
```

你会发现列表的索引和字符串是一样的,十分简单对吧?但是如果要是反过来,想要查看某个具体的值所在的位置,就需要用别的方法了,否则就会报错:

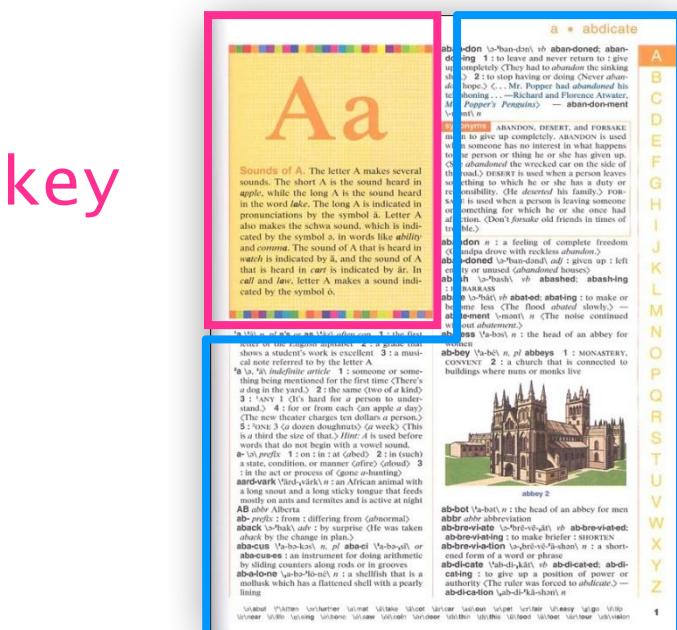
```
print(periodic_table['H'])
```

```
Traceback (most recent call last):
  File "/Users/lingqianqian/untitled8/test.py", line 3, in <module>
    print(periodic_table['H'])
TypeError: list indices must be integers or slices, not str
```

报错是因为列表只接受用位置进行索引,但如果数据量很大的话,肯定会记不住什么元素在什么位置,那么有没有一种数据类型可以用人类的方式来进行索引呢?其实这就是字典,我们一起来继续学习。

## 字典 (Dictionary)

编程世界中其实有很多概念都基于现实生活的原型，字典这种数据结构的特征也正如现实世界中的字典一样，使用名称-内容进行数据的构建，在 Pyteol 中分别对应着键(hay)-值(vPku)，习惯上称之为键值对。



# value

字典的特征总结如下：

1. 字典中数据必须是以键值对的形式出现的；
2. 逻辑上讲，键是不能重复的，而值可以重复；
3. 字典中的键(key)是不可变的，也就是无法修改的；而值(value)是可变的，可修改的，可以是任何对象。

用下面这个例子来看一下，这是字典的书写方式：键与值并不能脱离对方而存在，如果你这样写会引发一个语法错误

```
NASDAQ_code =  
    { 'BIDU': 'Baidu'  
    ,  
    'SINA': 'Sina',  
    'YOKU': 'Youku'
```

一个字典中键与值并不能脱离对方而存在，如果你写成{'BIDU':}会引发一个语法错误：

```
NASDAQ_code = {'BIDU':}  
                           ^  
SyntaxError: invalid syntax
```

我们再试着将一个可变(mutable)的元素作为 key 来构建字典，比如列表：

```
key_test = {[]: 'a Test'}  
print(key_test)
```

```
key_test = {[]: 'a Test'}  
TypeError: unhashable type: 'list'
```

想必一次次的报错会让你深深的记住这两个特征：key 和 value 是一一对应的，key 是不可变的。

同时字典中的键值不会有重复，即便你这么做，相同的键值也只能出现一次：

```
a = {'key': 123, 'key': 123}  
print(a)
```

## 字典的增删改查

首先我们按照映射关系创建一个字典, 继续使用前面的例子:

```
NASDAQ_code = {'BIDU': 'Baidu', 'SINA': 'Sina'}
```

与列表不同的是, 字典并没有一个可以往里面添加单一元素的“方法”, 但是我们可以通过这种方式进行添加:

```
NASDAQ_code['YOKU'] = 'Youku'  
print(NASDAQ_code)
```

列表中有用来添加多个元素的方法 `extend()`, 在字典中也有对应的添加多个元素的方法 `update()`:

```
NASDAQ_code.update({'FB': 'Facebook', 'TSLA': 'Tesla'})
```

删除字典中的元素则使用 `del` 方法:

```
del NASDAQ_code['FB']
```

需要注意的是, 虽说字典是使用的花括号, 在索引内容的时候仍旧使用的是和列表一样的方括号进行索引, 只不过在括号中放入的一定是, 字典中的键, 也就是说需要通过键来索引值:

```
NASDAQ_code['TSLA']
```

同时, 字典是不能够切片的, 也就是说下面这样的写法应用在字典上是错误的:

```
chart[1:4] # WRONG!
```



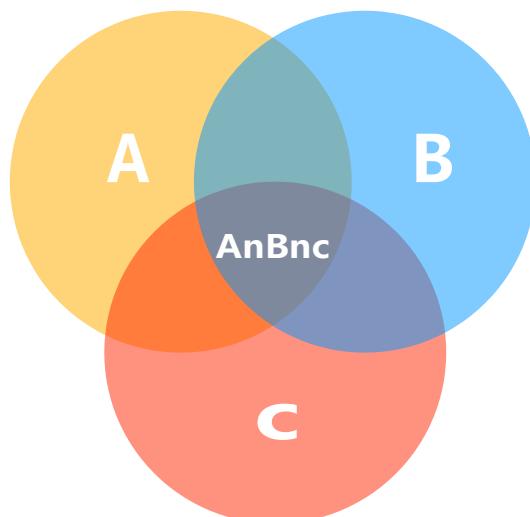
## 元组 (Tuple)

元组其实可以理解成一个稳固版的列表, 因为元组是不可修改的, 因此在列表中的存在的方法均不可以使用在元组上, 但是元组是可以被查看索引的, 方式就和列表一样:

```
letters = ('a', 'b', 'c', 'd', 'e', 'f', 'g')  
letter[0]
```

## 集合 (Set)

集合则更接近数学上集合的概念。每一个集合中的元素是无序的、不重复的任意对象, 我们可以通过集合去判断数据的从属关系, 有时还可以通过集合把数据结构中重复的元素减掉。



集合不能被切片也不能被索引, 除了做集合运算之外, 集合元素可以被添加还有删除:

```
a_set = {1, 2, 3, 4}  
a_set.add(5)  
a_set.discard(5)
```

# 数据结构的一些技巧

## 多重循环

有很多函数的用法和数据结构的使用都是息息相关的。前面我们学习了列表的基本用法,而在实际操作中往往回遇到更多的问题。比如,在整理表格或者文件的时候会按照字母或者日期进行排序,在 Python 中也存在类似的功能:

```
num_list = [6, 2, 7, 4, 1, 3, 5]
print(sorted(num_list))
```

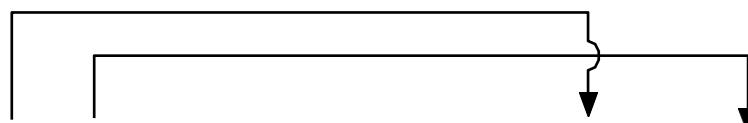
怎么样,很神奇吧?sorted 函数按照长短、大小、英文字母的顺序给每个列表中的元素进行排序。这个函数会经常在数据的展示中使用,其中有一个非常重要的地方,sorted 函数并不会改变列表本身,你可以把它理解成是先将列表进行复制,然后在进行顺序的整理。

在使用默认参数 reverse 后列表可以被按照逆序整理:

```
sorted(num_list, reverse=True)
```

在整理列表的过程中,如果同时需要两个列表应该怎么办?这时候就可以用到zip 函数,比如:

```
for a, b in zip(num, str):
    print(b, 'is', a)
```



*for a, b in zip(num, str):*

## 推导式

现在我们来看数据结构中的推导式(List comprehension),也许你还看到过它的另一种名称叫做列表的解析式,在这里你只需要知道这两个说的其实是一个东西就可以了。

现在我有10个元素要装进列表中,普通的写法是这样的:

```
a = []
for i in range(1,11):
    a.append(i)
```

下面换成列表解析的方式来写:

```
b = [i for i in range(1,11)]
```

列表解析式不仅非常方便,并且在执行效率上要远远胜过前者,我们把两种不同的列表操作方式所耗费的时间进行对比,就不难发现其效率的巨大差异:

```
import time

a = []
t0 = time.clock()
for i in range(1,20000):
    a.append(i)
print(time.clock() - t0, "seconds process time")

t0 = time.clock()
b = [i for i in range(1,20000)]
print(time.clock() - t0, "seconds process time")
```

```
8.99999999998592e-06 seconds process time
0.001232000000000005 seconds process time
```

列表推导式的用法也很好理解,可以简单地看成两部分。红色虚线后面的是我们熟悉的 for 循环的表达式,而虚线前面的可以认为是我们想要放在列表中的元素,在这个例子中放在列表中的元素即是后面循环的元素本身。

```
list = [item for item in iterable]
```

为了更好地理解这句话, 我们继续看几个例子:

```
a = [i**2 for i in range(1,10)]
c = [j+1 for j in range(1,10)]
k = [n for n in range(1,10) if n % 2 ==0]
z = [letter.lower() for letter in 'ABCDEFGHIJKLMN']
```

字典推导式的方式略有不同, 主要是因为创建字典必须满足键-值的两个条件才能达成:

```
d = {i:i+1 for i in range(4)}
```

```
g = {i:j for i,j in zip(range(1,6), 'abcde')}
g = {i:j.upper() for i,j in zip(range(1,6), 'abcde')}
```

## 循环列表时获取元素的索引

现在我们有一个字母表, 如何能在索引的时候像图中一样得到每个元素的具体位置的展示呢?

```
letters = ['a', 'b', 'c', 'd', 'e', 'f', 'g']
```

```
a is 1
b is 2
c is 3
d is 4
e is 5
f is 6
g is 7
```

前面提到过, 列表是有序的, 这时候我们可以使用 Python 中独有的函数 `enumerate` 来进行:

```
letters = ['a', 'b', 'c', 'd', 'e', 'f', 'g']
for num, letter in enumerate(letters):
    print(letter, 'is', num + 1)
```

## 综合项目

为了深入理解列表的使用方法, 在本章的最后我们来进行做一个词频统计。需要瓦尔登湖的文本, 可以在这里下载:<http://pan.baidu.com/s/1o75GKZ4>, 下载后用 PyCharm 打开文本重新保存一次, 这是为了避免编码的问题。

之前还是提前做一些准备, 学习一些必要的知识。

```
lyric = 'The night begin to shine, the night begin to shine'
words = lyric.split()
```

现在我们使用 `split` 方法将字符串中的每个单词分开, 得到独立的单词:

```
['The', 'night', 'begin', 'to', 'shine']
```

接下来是词频统计, 我们使用 `count` 方法来统计重复出现的单词:

```
path = '/Users/Hou/Desktop/walden.txt'
with open(path, 'r') as text:
    words = text.read().split()
    print(words)
    for word in words:
        print('{} - {} times'.format(word, words.count(word)))
```

结果出来了, 但是总感觉有一些奇怪, 仔细观察得出结论:

1. 有一些带标点符号的单词被单独统计了次数;
2. 有些单词不止一次地展示了出现的次数;
3. 由于 Python 对大小写敏感, 开头大写的单词被单独统计了。

现在我们根据这些点调整一下我们的统计方法, 对单词做一些预处理:

```
1. import string
2. path = '/Users/Hou/Desktop/walden.txt'
3. with open(path, 'r') as text:
4.     words = [raw_word.strip(string.punctuation).lower() for raw_word in text.read().split()]
5.     words_index = set(words)
6.     counts_dict = {index:words.count(index) for index in words_index}
7. for word in sorted(counts_dict, key=lambda x: counts_dict[x], reverse=True):
8.     print('{} -- {} times'.format(word, counts_dict[word]))
```

1. 引入了一个新的模块 strgne 。其实这个模块的用法很简单, 我们可以试着把 strgne.punbtu、tgon 打印出来, 其实这里面也仅仅是包含了所有的标点符号——!"#\$%&'() \*+, -./:; <=>2@[\]A \_{I}]^~。
4. 在文字的首位去掉了连在一起的标点符号, 并把首字母大写的单词转化成小写;
5. 将列表用 sdt 函数转换成集合, 自动去除掉了其中所有重复的元素;
6. 创建了一个以单词为键(kdy)出现频率为值(v、lud)的字典;
- 7~8. 打印整理后的函数, 其中 `key=lambda x: counts_dict[x]` 叫做 1、mac、 表达式, 可以暂且理解为以字典中的值为排序的参数。



# 第七章

# 开始使用第三方库

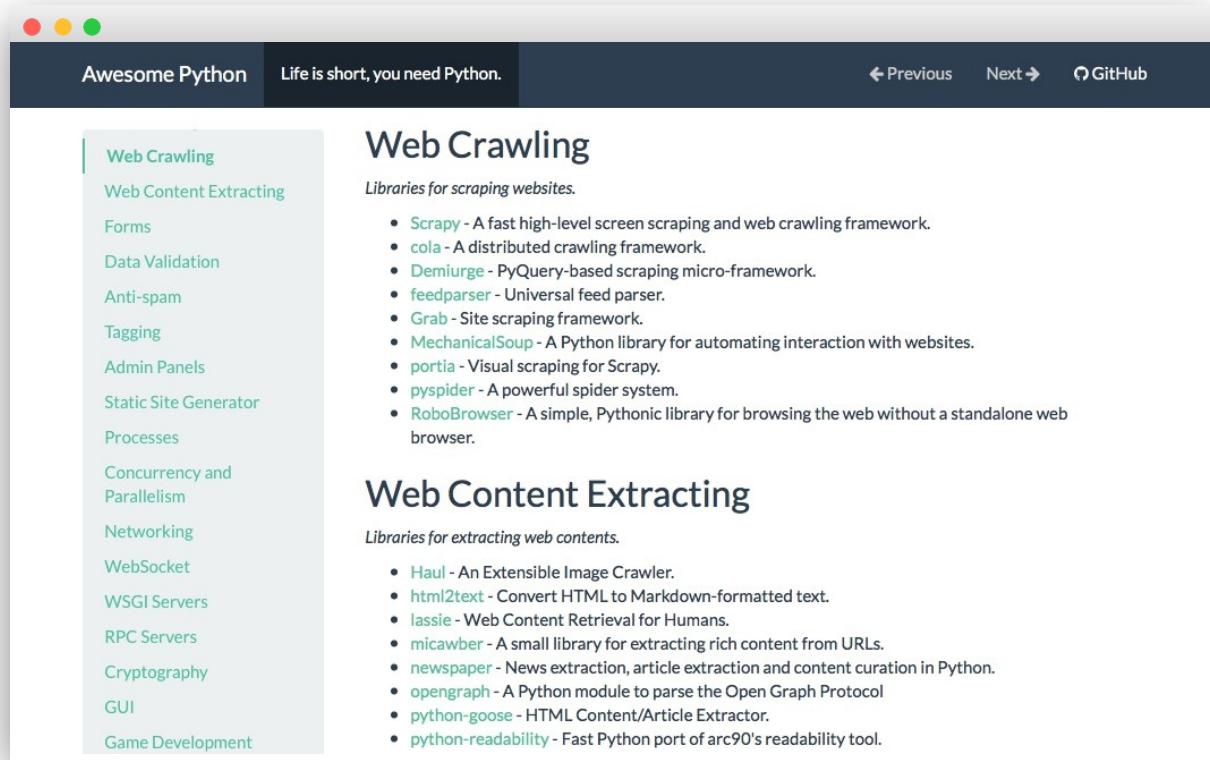
# 令人惊叹的第三方库

如果用手机来比喻编程语言,那么 Python 是一款智能机。正如海量的手机应用出现在 iOS、Android 平台上,同样有各种各样的第三方库为 Python 开发者提供了极大的便利。

当你想搭建网站时,可以选择功能全面的 Django、轻量的 Flask 等 web 框架;当你想写一个小游戏的时候,可以使用 PyGame 框架;当你想做一个爬虫时,可以使用 scrapy 框架;当你想做数据统计分析时,可以使用 Pandas 数据框架……这么多丰富的资源可以帮助我们高效快捷地做到想做的事,就不需要再重新造轮子了。

那么,如何根据自己的需求找到相应的库呢?

可以在 [awesome-python.com](http://awesome-python.com) 这个网站上按照分类去寻找,上面收录了比较全面的第三方库。比如当我们想找爬虫方面的库时,查看 Web Crawling 这个分类,就能看到相应的第三方库的网站与简介:



The screenshot shows the 'Awesome Python' website interface. At the top, there's a dark header bar with three circular icons (red, yellow, green), the text 'Awesome Python', 'Life is short, you need Python.', and navigation links for 'Previous', 'Next', and 'GitHub'. Below the header is a sidebar on the left containing a list of categories: 'Web Crawling', 'Web Content Extracting', 'Forms', 'Data Validation', 'Anti-spam', 'Tagging', 'Admin Panels', 'Static Site Generator', 'Processes', 'Concurrency and Parallelism', 'Networking', 'WebSocket', 'WSGI Servers', 'RPC Servers', 'Cryptography', 'GUI', and 'Game Development'. The main content area has a title 'Web Crawling' and a subtitle 'Libraries for scraping websites.' It lists several libraries: Scrapy, cola, Demiurge, feedparser, Grab, MechanicalSoup, portia, pypspider, and RoboBrowser. Below this is another section titled 'Web Content Extracting' with a subtitle 'Libraries for extracting web contents.' It lists Haul, html2text, lassie, micawber, newspaper, opengraph, python-goose, and python-readability.

可以进入库的网站查看更详细的介绍,并确认这个库支持的是 python 2 还是 python 3,不过绝大多数常用库已经都支持了这两者。

另外,还可以直接通过搜索引擎寻找,比如:



Google 搜索结果展示了约 202,000 条关于 Python 爬虫库的结果。以下是一些搜索结果摘要：

- 用Python写爬虫, 用什么方式、框架比较好? - Python ... - 知乎**  
<https://www.zhihu.com/question/19899608>  
 前只写过很简单的Python爬虫, 直接用内置库实现, 有没有谁用Python爬过规模较大的数据, 用的是什么方法? 还有, 采用现有的Python爬虫框架, 相比与直接使用 ...
- 开源python网络爬虫框架Scrapy - CORE - 51CTO技术博客**  
[lujiesky.blog.51cto.com/332319/1328176](http://lujiesky.blog.51cto.com/332319/1328176)  
 2013年11月19日 - 开源python网络爬虫框架Scrapy ... 标签: Scrapy 网络爬虫框架  
 Python ... Scrapy 使用Twisted这个异步网络库来处理网络通讯, 架构清晰, 并且包含 ...
- Python开发爬虫常用库 - 豆瓣**  
[www.douban.com/group/topic/39919487/](http://www.douban.com/group/topic/39919487/)  
 2013年6月1日 - Python开发爬虫常用库 ..... google中这样搜"python spider github", 可以找到很多优秀的源代码, 用来 ... 初学者, 开发一段小爬虫代码, 增加自信心。
- Python爬虫入门三之Urllib库的使用| 静觅**  
[cuiqingcai.com/947.html](http://cuiqingcai.com/947.html)  
 那么接下来, 小伙伴们就一起和我真正迈向我们的爬虫之路吧。1. 分分钟扒一个网页下来怎样扒网页呢? 其实就是根据URL来获取它的网页信息, 虽然我们在浏览器中 ...

如果你能尝试用英文搜索, 会发现更大的世界, 比如 stackoverflow 上的优质讨论。



Google 搜索结果展示了约 229,000 条关于 Python 网络爬虫库的结果。以下是一些搜索结果摘要：

- Scrapy | A Fast and Powerful Scraping and Web Crawling ...**  
[scrapy.org/](http://scrapy.org/) 翻译此页  
 An open source and collaborative framework for extracting the data you need ... PyPI  
 Version PyPI monthly downloads Wheel Status Python 3 Porting Status ...  
 Documentation - Download - Companies - Frequently Asked Questions
- Anyone know of a good Python based web crawler that I ...**  
[stack overflow.com/.../anyone-know-of-a-good-python-based-...](http://stackoverflow.com/.../anyone-know-of-a-good-python-based-...) 翻译此页  
 2009年1月7日 - Hound is a simple web crawler developed in python. - mfontanini Sep 29 '11 ... It is a twisted-based web crawler framework. Still under heavy ...
- Build a Python Web Crawler with Scrapy - DevX**  
[www.devx.com/webdev/build-a-python-web-crawler-with-scr...](http://www.devx.com/webdev/build-a-python-web-crawler-with-scr...) 翻译此页  
 2015年2月12日 - ... different sources is required. Learn to create a web scraper with Python combined with an open source web crawler framework called Scrapy.
- 11. Introduction to web-crawling in Python — python\_for\_ss ...**  
[www-rohan.sdsu.edu/.../python\\_for\\_ss/.../web/web\\_intro.html](http://www-rohan.sdsu.edu/.../python_for_ss/.../web/web_intro.html) 翻译此页  
 This section contains an introduction to some basic python web crawling tools. We will focus on programs that request web services from service providers and ...

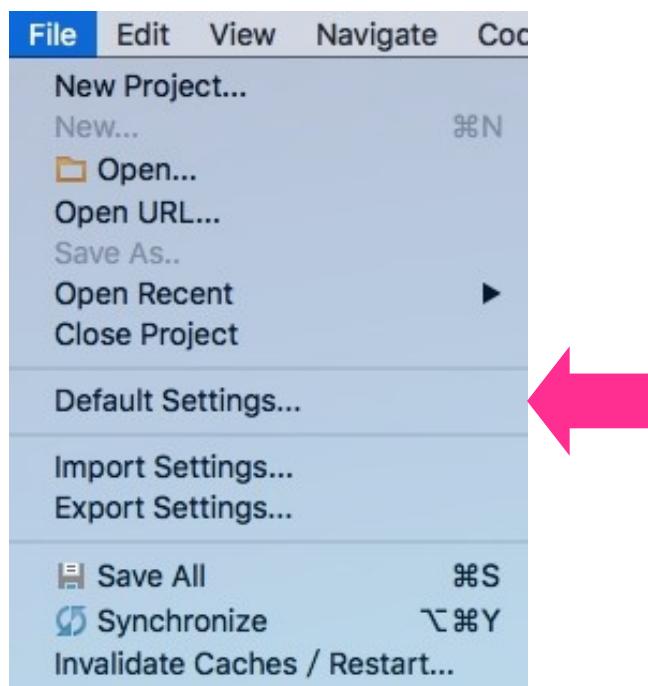
# 安装第三方库

无论你想安装哪一种库，方法基本都是通用的。下面开始介绍安装第三方库的方法。

## 最简单的方式：在 PyCharm 中安装

推荐大家使用 PyCharm，就是因为它贴心地考虑了开发者的使用体验，在 PyCharm 中可以方便快捷地安装和管理库。

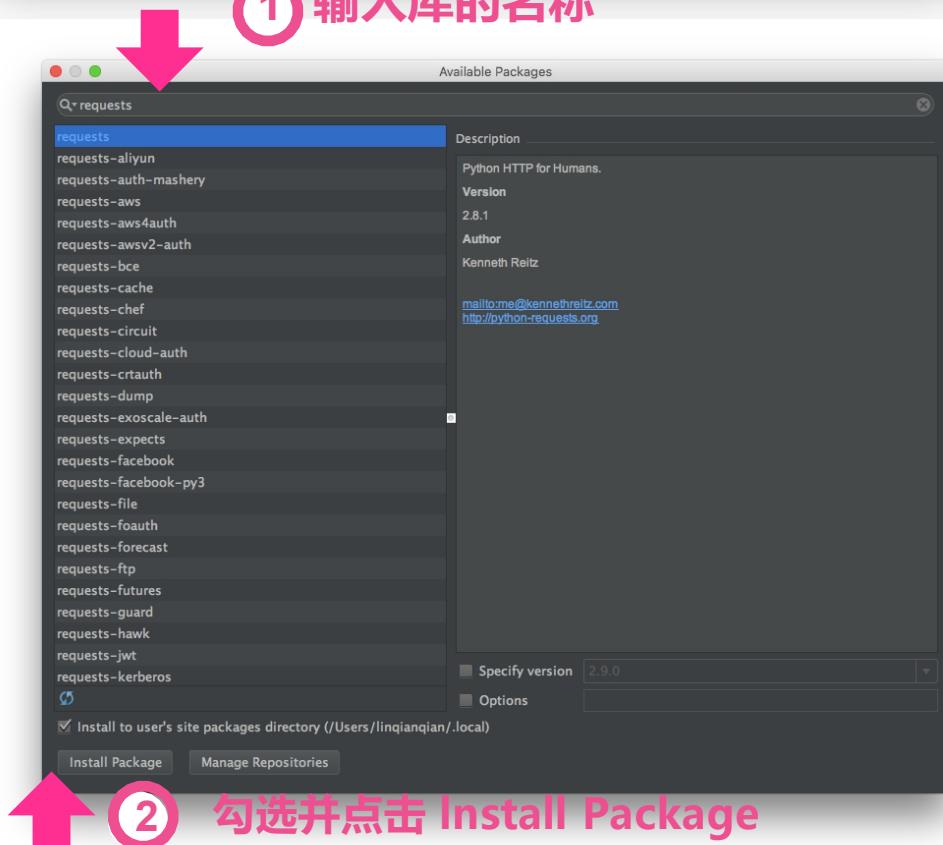
- 第一步：在 PyCharm 的菜单中选择：File > Default Settings



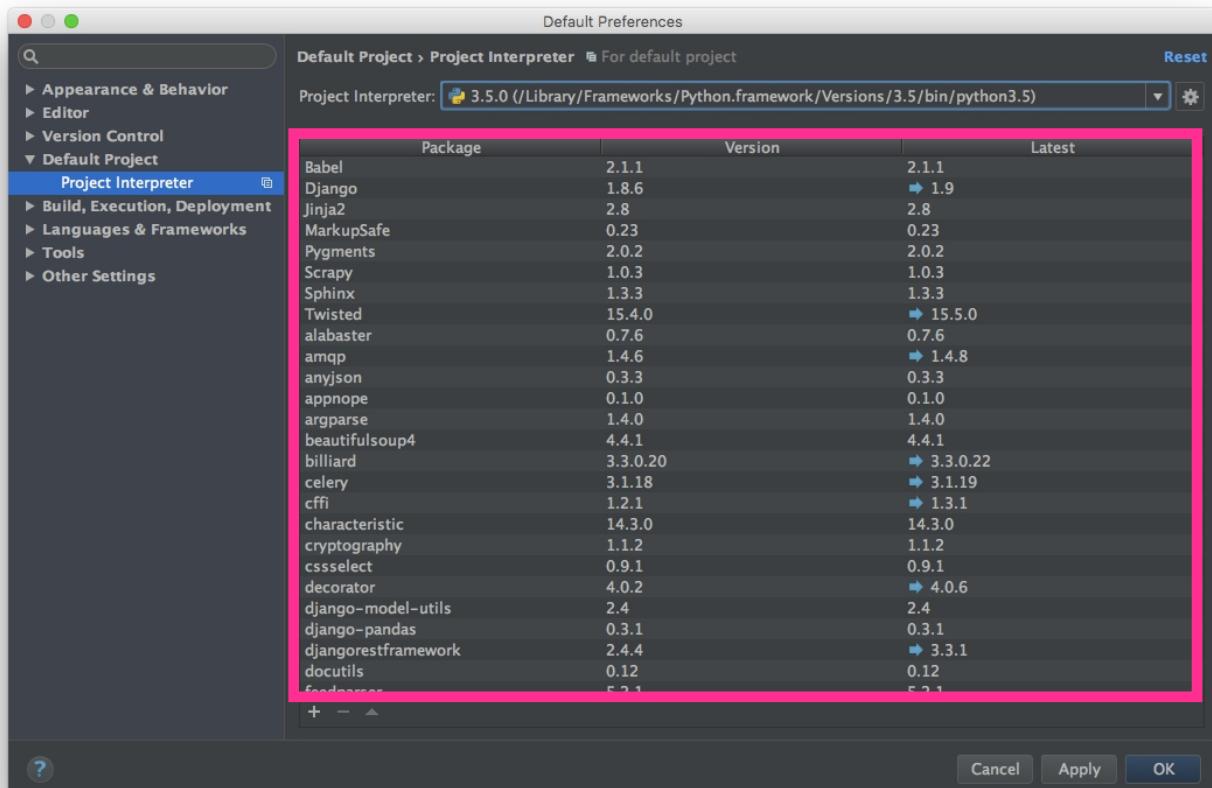
- 第二步：



- 第三步：



在安装成功后, PyCharm 会有成功提示。你也可以在 project interpreter 这个界面中查看安装了哪些库, 点`_`号就可以卸载不再需要的库。



## 最直接的方式:在终端/命令行中安装

### 安装 pip

在 Python 3.4 之后, 安装好 Python 环境就可以直接支持 pip, 你可以在终端/命令行里输入这句检查一下:

```
pip --version
```

如果显示了 pip 的版本, 就说明 pip 已经成功安装了。

如果发现没有安装 pip的话, 各系统安装的方法不同:

- Windows 用户请查看:<https://taiZilongxu.gitbooks.io/stackoverzow-about-python/content/8/README.html>
- Mac 用户请查看:<https://www.mobibrw.com/p=1274>
- Linux 用户请查看:<http://pip-cn.readthedocs.org/en/latest/installing.html>

### 使用 pip 安装库

在安装好 pip 之后, 以后安装库, 只需要在命令行里面输入:  
(如果你想安装到 python 2中, 需要把 pip3换成 pip )

```
pip3 install PackageName
```

如果你安装了 python 2和3两种版本, 可能会遇到安装目录的问题, 可以换成:  
(如果你想安装到 python 2中, 需要把 python3换成 python )

```
python3 -m pip install PackageName
```

如果遇到权限问题, 可以输入:

```
sudo pip install PackageName
```

安装成功后会提示:

```
Successfully installed PackageName
```

再介绍几个pip 的常用指令：

```
pip install --upgrade pip          #升级 pip
pip uninstall flask                #卸载库
pip list                          #查看已安装库
```

异常情况：安装某些库的时候，可能会遇到所依赖的另一个库还没安装，导致无法安装成功的情况，这时候的处理原则就是：缺啥装啥，举个例子：

```
danbao$ scrapy version -v
:0: UserWarning: You do not have a working installation of the
service_identity module: 'No module named service_identity'.
Please install it from <https://pypi.python.org/pypi/
service_identity> and make sure all of its dependencies are
satisfied. without the service_identity module and a recent
enough pyOpenSSL to support it, Twisted can perform only
rudimentary TLS client hostname verification. Many valid
certificate/hostname mappings may be rejected.
Scrapy : 0.24.6
lxml   : 3.4.4.0
libxml2 : 2.9.0
Twisted : 15.2.1
Python  : 2.7.9 (default, May 27 2015, 22:47:13) - [GCC 4.2.1
Compatible Apple LLVM 6.1.0 (clang-602.0.53)]
```

这时候的解决方法是：

```
pip install service_identity
```

## 最原始的方式:手动安装

为了应对异常情况,再提供一种最原始的方法:手动安装。往往是 Windows 用户需要用到这种方法。

**进入pypi.python.org,搜索你要安装的库的名字,这时候有3种可能,**

- 第一种是 exe 文件,这种最方便,下载满足你的电脑系统和 python 环境的对应的exe,再一路点击 next 就可以安装。
- 第二种是 .whl 类文件,好处在于可以自动安装依赖的包。
- 第三种是源码,大概都是 zip 、 tar.zip、 tar.bz2 格式的压缩包,这个方法要求用户已经安装了这个包所依赖的其他包。例如pandas依赖于numpy,你如果不安装numpy,这个方法是无法成功安装pandas的。如果没有前两种类型的文件,那只能用这个了。

### I. 如果你选择了下载.whl 类文件,下面是安装方法:

1) 到命令行输入:

```
pip3 install wheel
```

等待执行完成,不能报错。(python2 要换成 pip)

2) 从资源管理器中确认你下载的.whl 类文件的路径,然后在命令行继续输入:

```
cd C:\download
```

此处需要改为你的路径,路径的含义是文件所在的文件夹,不包含这个文件名字本身

然后再在命令行继续输入:

```
pip3 install xxx.whl
```

xxx.whl 是你下载的文件的完整文件名。

## 2、如果你选择了下载源码压缩包,下面是安装方法:

1) 解压包,进入解压好的文件夹,通常会看见一个 `setup.py` 的文件。从资源管理器中确认你下载的文件的路径,打开命令行(cmd),输入:

```
cd C:\download
```

此处需要改为你的路径,路径的含义是文件所在的文件夹,不包含这个文件名字本身

2) 然后在命令行中继续输入:

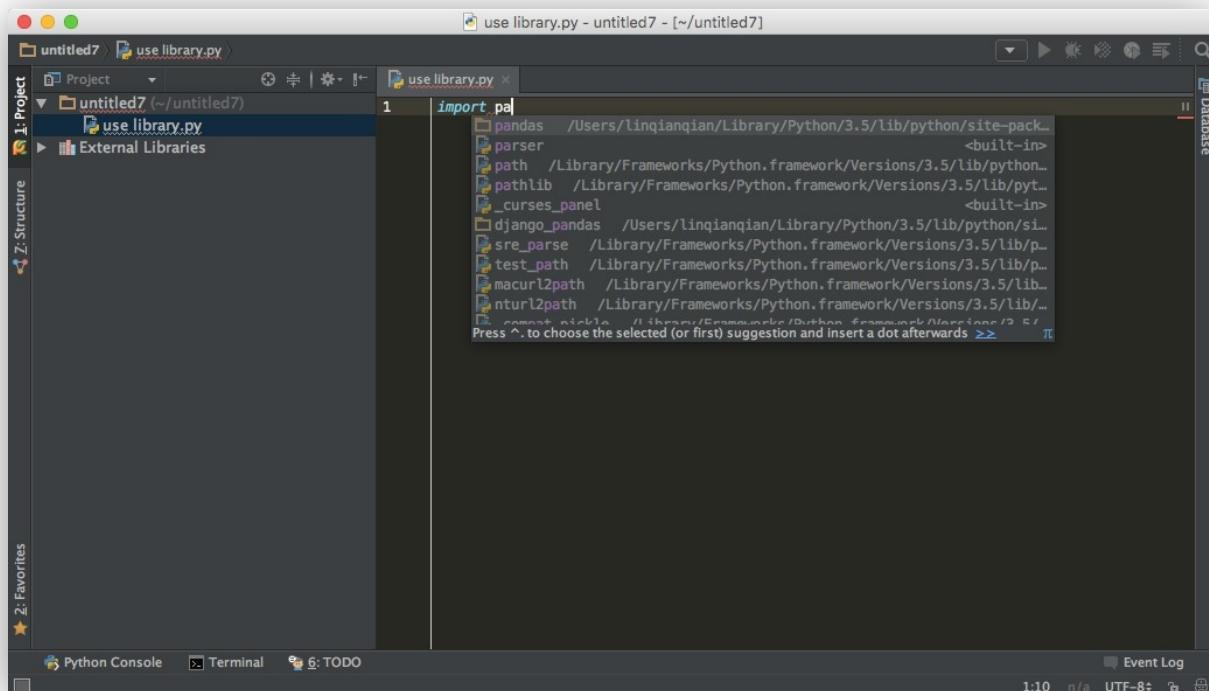
```
python3 setup.py install
```

这个命令,就能把这个第三库安装到系统里,也就是你的 Python 路径, windows 大概是在 `C:\Python3.5(或2.7)\Lib\site-packages`。

想卸载库的时候,找到 python 路径,进入 `site-packages` 文件夹,在里面删掉库文件就可以了。

# 使用第三方库

在 PyCharm 中输入库的名字, 就会自动提示补全了:



输入之后你会发现是灰色的状态`import pandas`, 这是因为还没有在程序中使用这个库, 而不是因为库还没安装(想检查库是否安装的话, 可以使用前面提到的 PyCharm 或者 pip 的方式来确认)。

关于库的使用指南到这里就结束了。

我是教程的作者夏洛,我在学习编程时发现,市面上没有一本教程适合非程序员这个人群的初学者,于是开始用生动易懂的方式来系统地整理编程知识。



感谢创作教程过程中这些朋友给予的宝贵意见



诸葛:图灵学院联合创始人,前唯品会资深架构师。

顾安:金提科技python资深技术专家,天眼查项目主导者之一

周瑜:美团点评高级架构师,美团外卖分布式架构设计者

monkey:国美技术委员会成员,开源框架Flasher作者