# 目录

[目录 1](#_Toc7225)

[安装python 2](#_Toc9541)

[安装pip 4](#_Toc12297)

[安装pycharm 7](#_Toc19893)

[使用pycharm 10](#_Toc27578)

[设置pycharm字体 12](#_Toc22975)

[通过ASIN获取reviews内容 13](#_Toc28646)

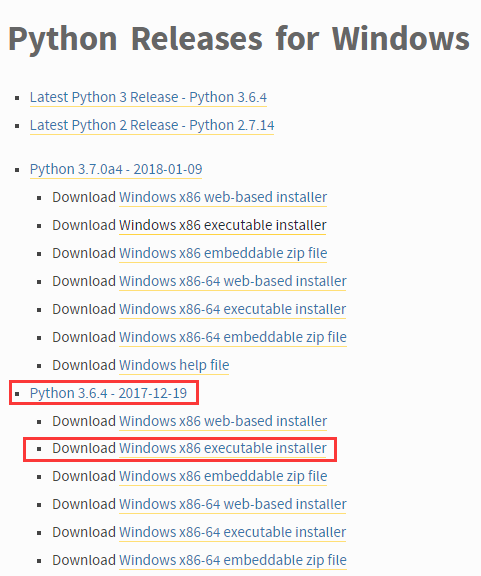
[通过keyword, ASIN, SKU获取搜索页排名 17](#_Toc9086)

[通过keyword获取前2页listing的具体信息 20](#_Toc8827)

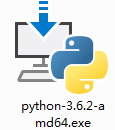
## 安装python

去官网下载python3.6程序

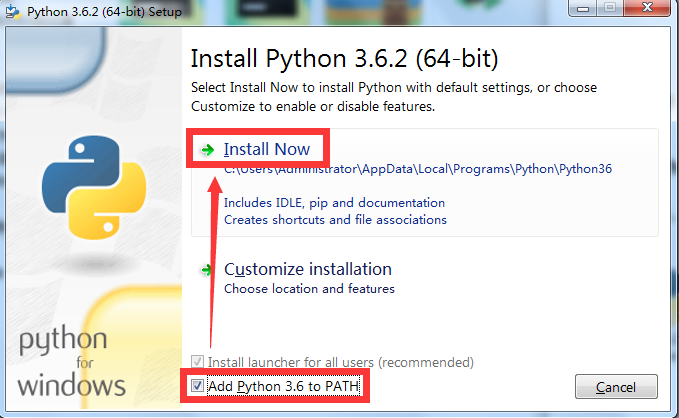
https://www.python.org/downloads/windows/



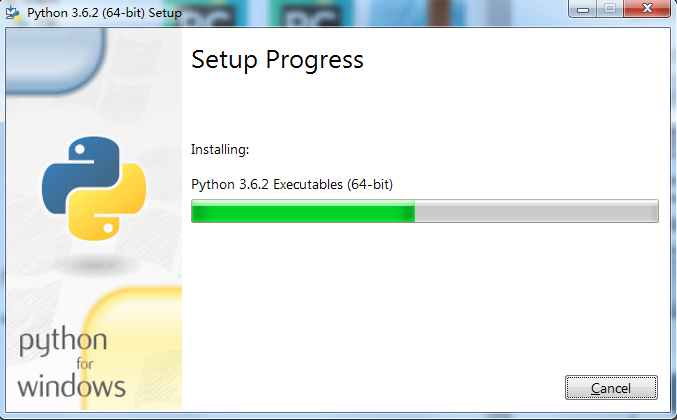
双击如下图标；



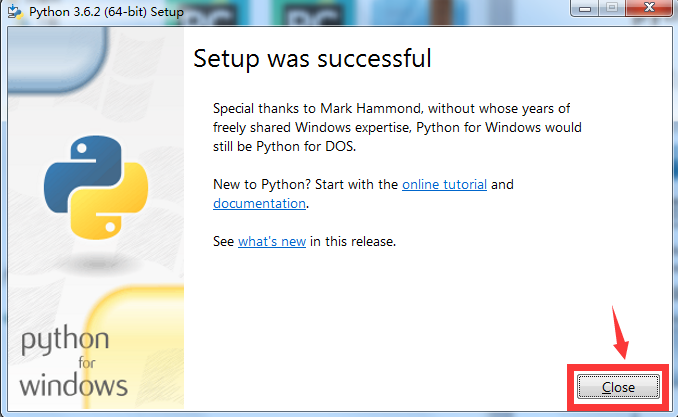
勾选底部复选框！！！！！！！ 点Install Now；



安装进行中，等一分钟左右；



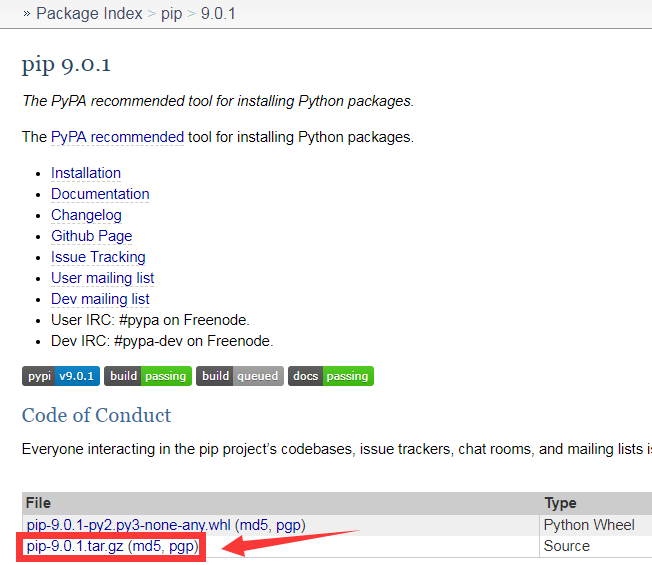
python安装完成，点close；



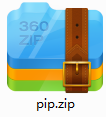
## 安装pip

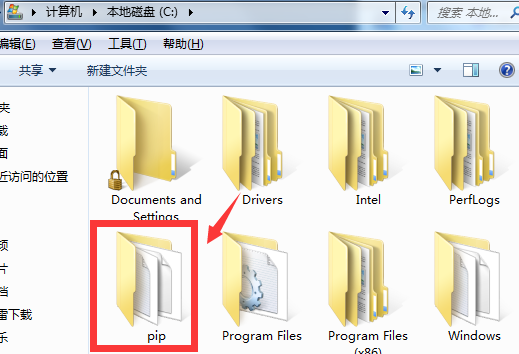
下载pip压缩包，解压缩到C盘， 文件夹重命名为pip；

https://pypi.python.org/pypi/pip

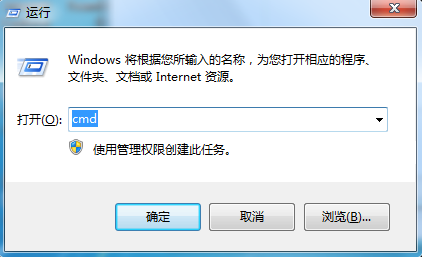


或者到共享下载，把pip.zip压缩包复制到C盘，解压缩到C盘；





按下键盘左下角Windows按键+r；弹出运行窗口，填如cmd，点确定；



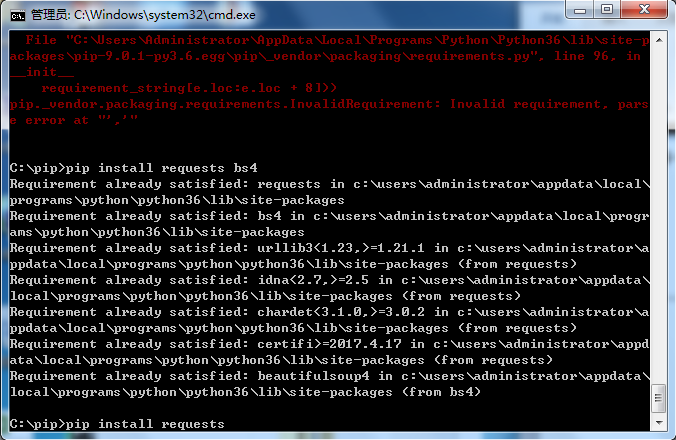
弹出黑色命令行窗口，输入 cd C:\pip ，敲回车；



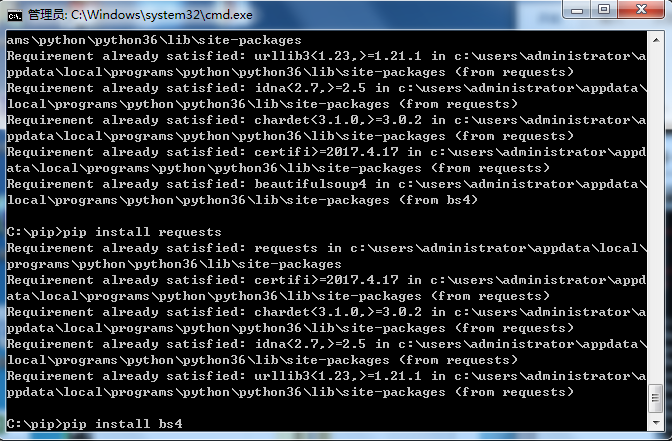
输入 python setup.py install ，敲回车；



等一长串字符跑完以后，输入 pip install requests ，敲回车；



等一长串字符跑完以后，输入 pip install bs4 ，敲回车；

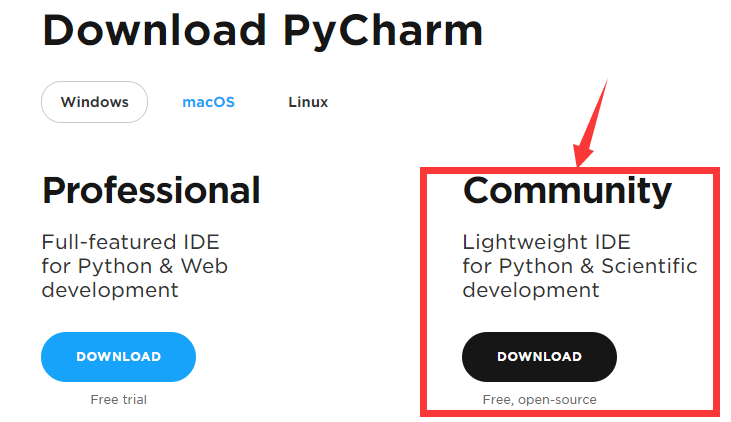


一长串字符跑完以后，pip安装完成；

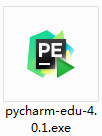
## 安装pycharm

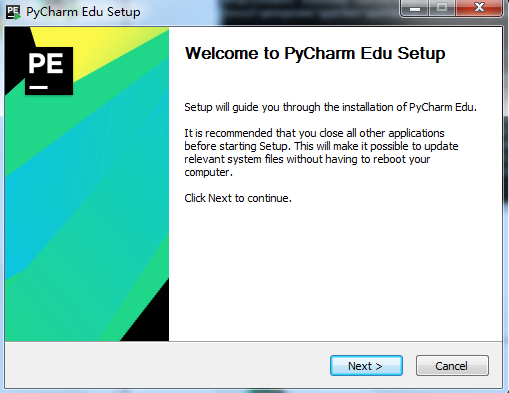
下载软件

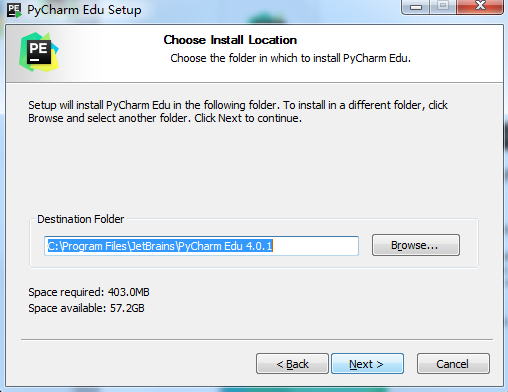
https://www.jetbrains.com/pycharm/download/



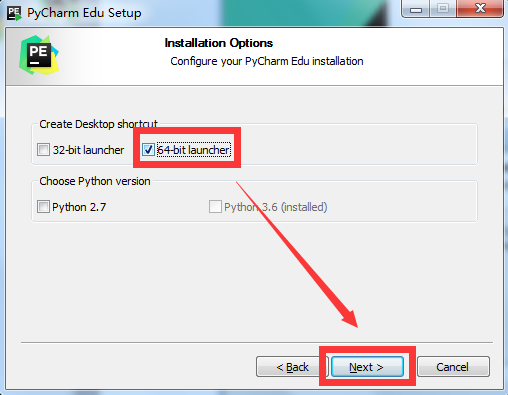
双击如下程序，一路next；



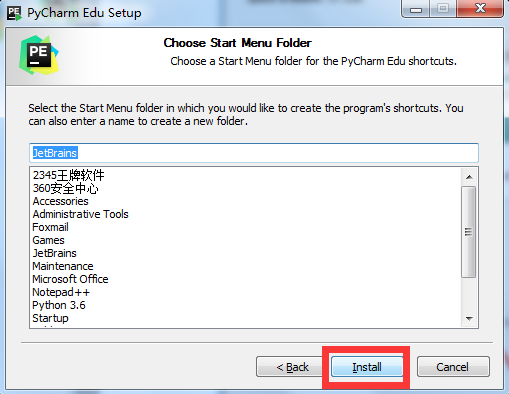




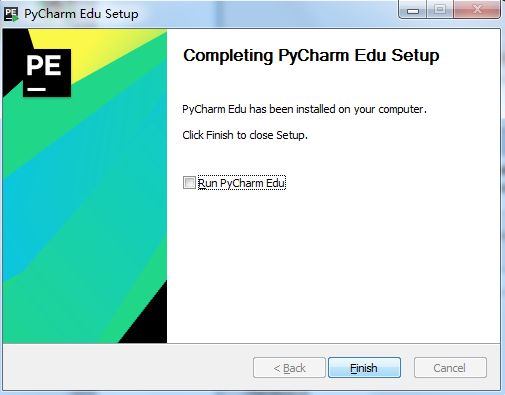
勾选64-bit launcher，Next；（如果你的电脑是32位系统，就勾选32-bit launcher）



Install



安装完成，Finish；

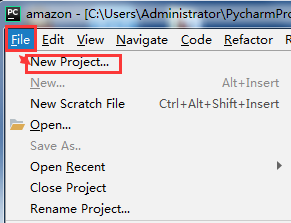


## 使用pycharm

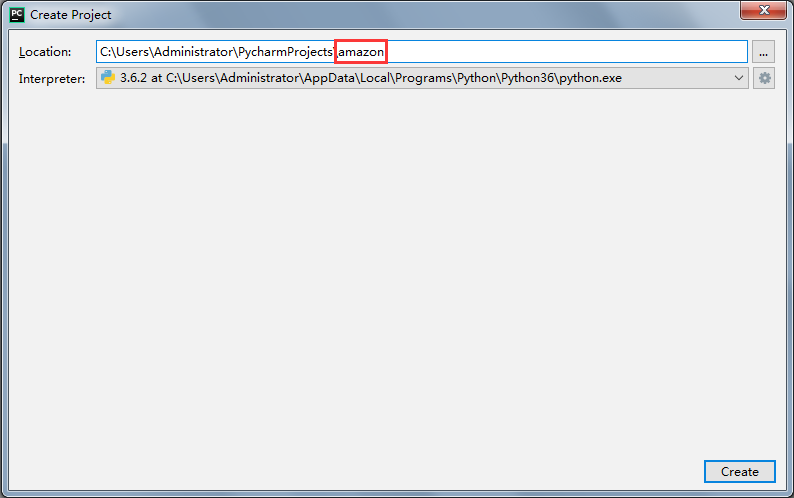
双击桌面如下图标；



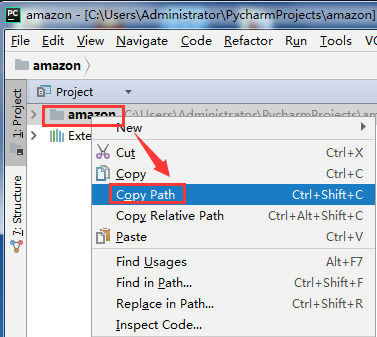
左上角File--New Project；

M

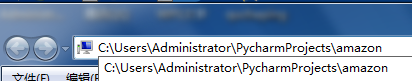
改成你想要的名字，比如amazon，点右下角Create；



右击左侧amazon文件夹，选Copy Path；



打开我的电脑，粘贴到地址栏，敲回车；

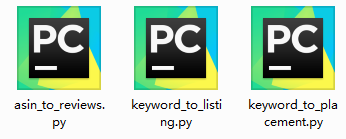


会打开文件夹，把下面如图所示的3个文件复制到这个amazon文件夹；

asin\_to\_reviews.py 会根据asin获取该产品最近20页的，所有星级的reviews，并保存到此文件下page rank子文件夹里；

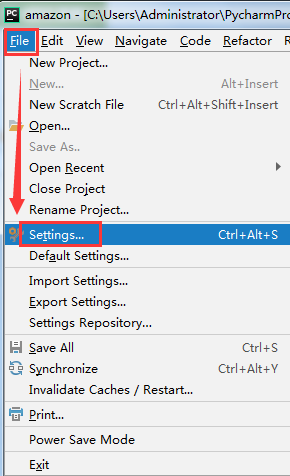
keyword\_to\_placement.py 会根据你设置的asin，关键词，sku来抓取前10页自然排名和广告位（不包含headline/右侧广告位）；

Keyword\_to\_listing.py 会根据你设置的asin，来抓取前2页所有listing的信息（如首图，价格，标题，五点，大类目排名等等）

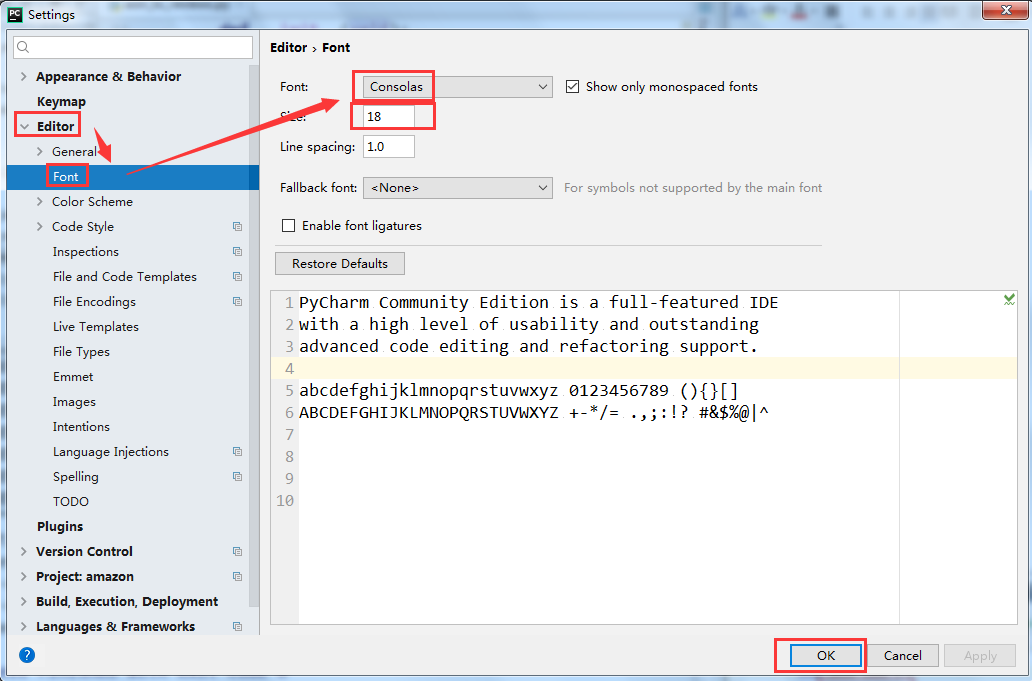


## 设置pycharm字体

File--Settings...；



Editor--Font，设置字体，大小，点右下角OK；



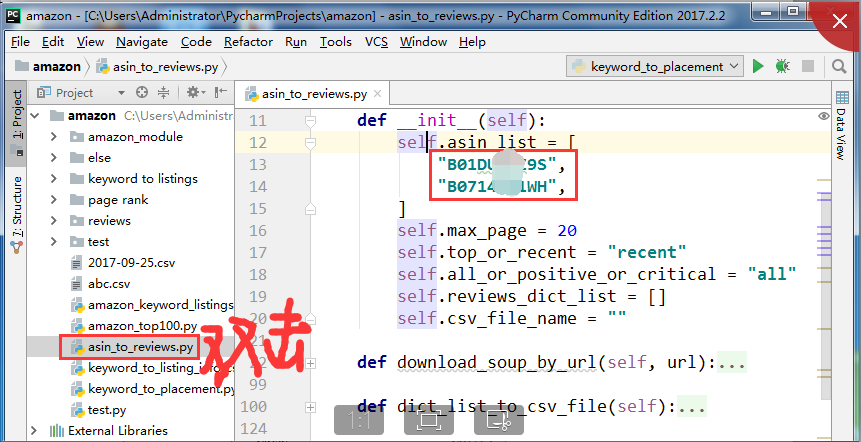
## 通过ASIN获取reviews内容

回到pycharm软件，双击左侧asin\_to\_reviews.py，右侧会出现代码；

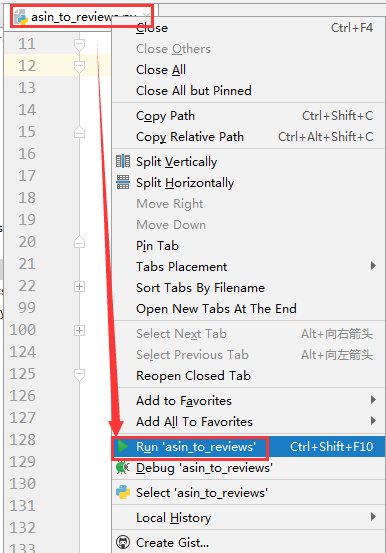
把ASIN修改成你想爬去的ASIN，可以一次写多个ASIN；

注意每行的格式为：英文双引号+ASIN+英文双引号+英文逗号；

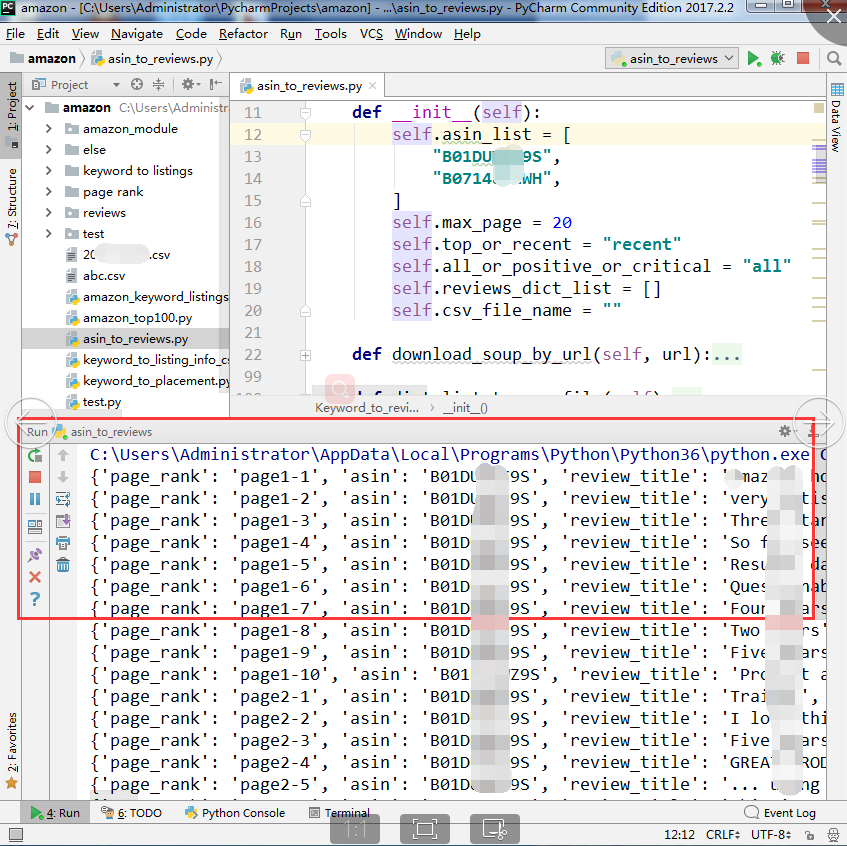
每行必须左对齐；



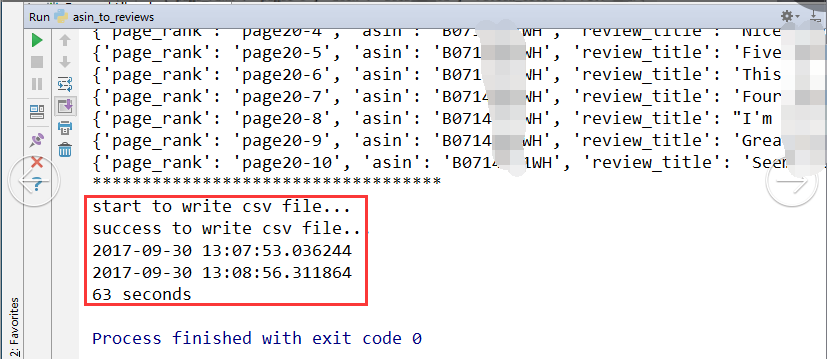
右击标题，选Run；



程序开始运行；

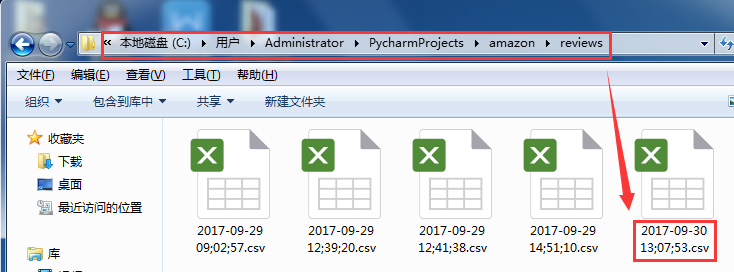


爬取一个ASIN大概需要30秒，二个大概60秒；



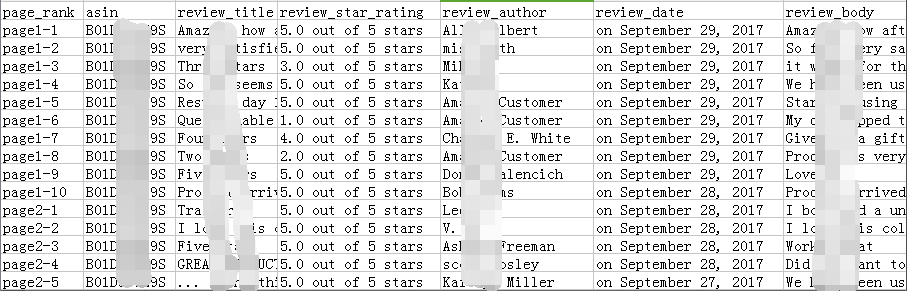
爬取好的数据在reviews文件夹里，路径如下；

文件名为开始爬取的日期，如2017-09-30 13;07;53代表2017-09-30 13:07:53；（因为windows文件名不能出现冒号，所以用分号代替）



文件中内容如下：

review的页码位置，asin，review标题，星级，留评人，日期，review内容；



## 通过keyword, ASIN, SKU获取搜索页排名

双击左侧keyword\_to\_placement.py，修改右侧内容；

把ASIN修改成你想爬去的ASIN，可以一次写多个ASIN；

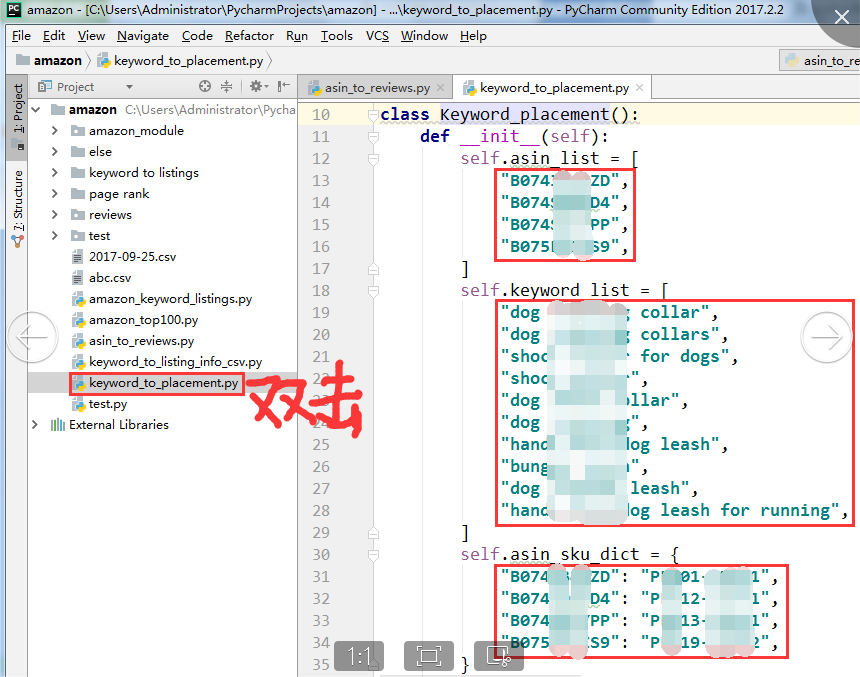
把keyword修改成你想爬去的keyword，可以一次写多个keyword；

做一个ASIN和SKU的对应表；（如果会用excel的vlookup，可以不用在这里设置）

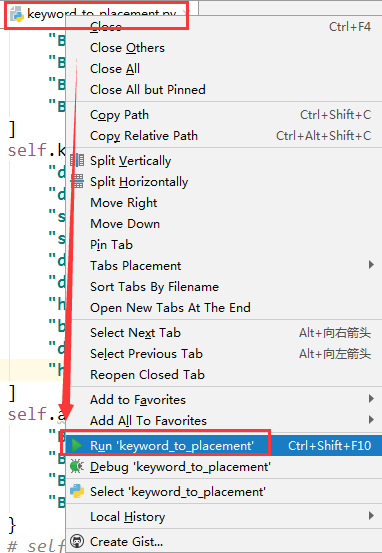
注意ASIN, keyword的格式为：英文的双引号+ASIN+英文的双引号+英文逗号；

注意ASIN和SKU的对应表的格式为：英文双引号+ASIN+英文双引号+英文冒号+英文双引号+SKU+英文双引号+英文逗号；

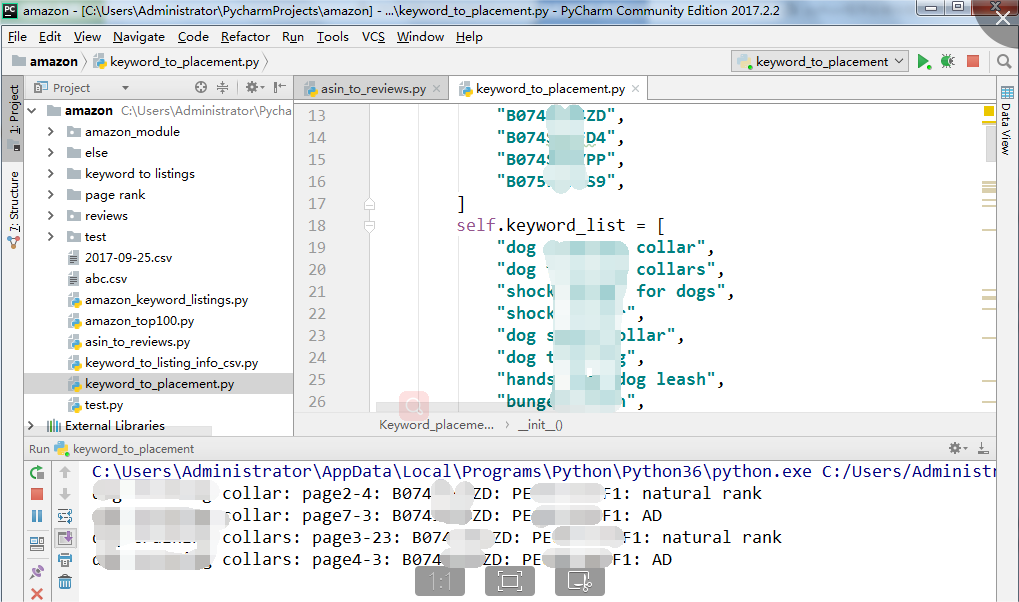
每行必须左对齐；



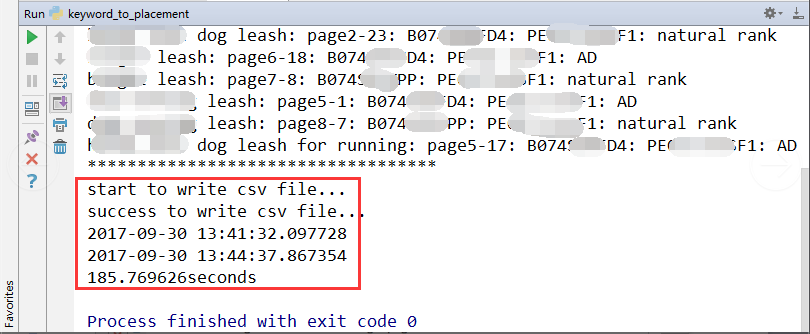
右击标题，选Run；



程序运行中；

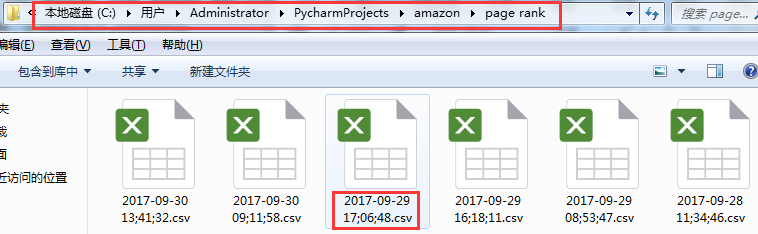


运行完成；



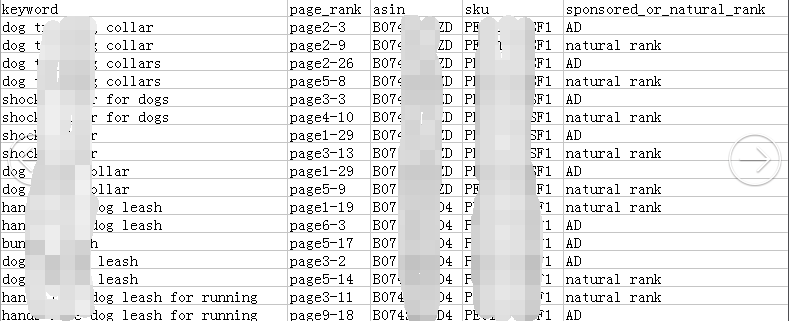
爬取好的数据在page rank文件夹里，路径如下；

文件名为开始爬取的日期，如2017-09-30 13;07;53代表2017-09-30 13:07:53；（因为windows文件名不能出现冒号，所以用分号代替）



文件中内容如下：

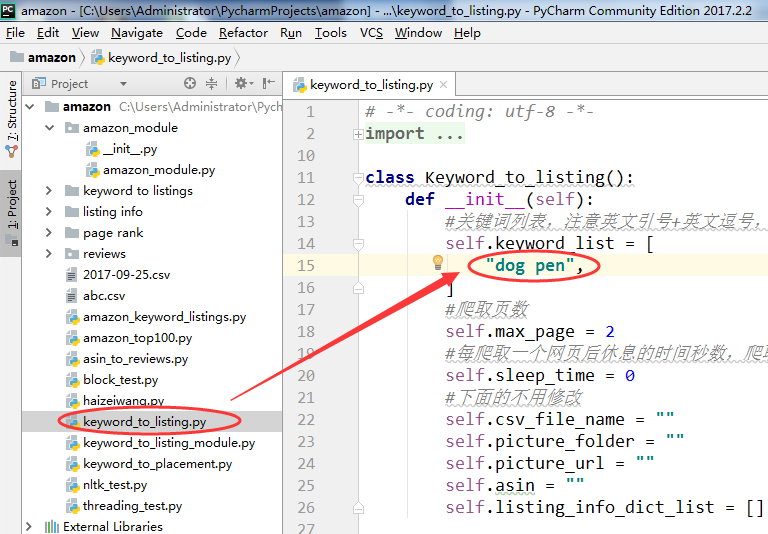
Keyword，页码位置，asin，sku，自然排名或广告位；



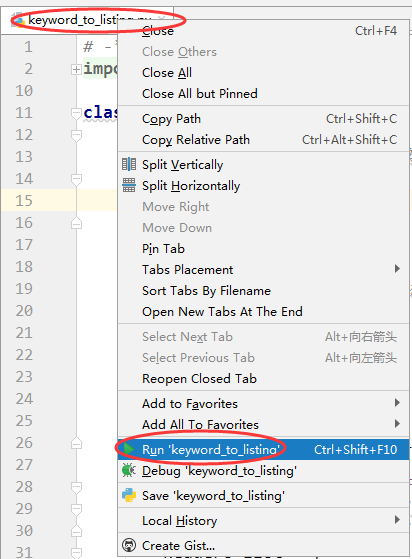
## 通过keyword获取前2页listing的具体信息

双击左侧keyword\_to\_listing.py，修改右侧内容；

把ASIN修改成你想爬去的ASIN，可以一次写多个ASIN；



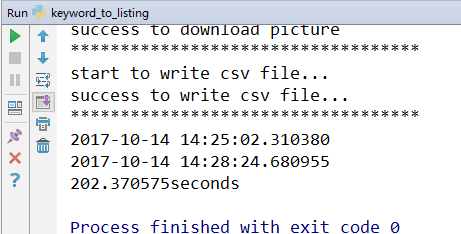
右击标题，点运行；



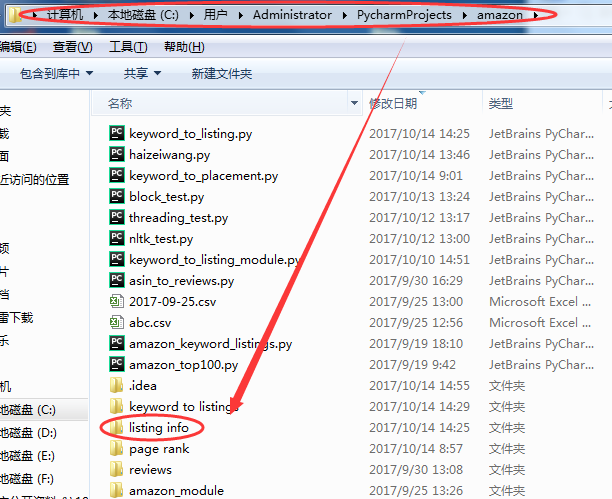
软件底部会实时显示爬取情况；



爬取完成，会显示运行时间；



默认下载到amazon文件夹下的listing info文件夹里；



Excel表以关键词命名；

