**一、Python的安装**

就在这安装就好了

<https://www.python.org/getit/>

​

一定要注意选择自己电脑的型号，其他的倒也是没啥

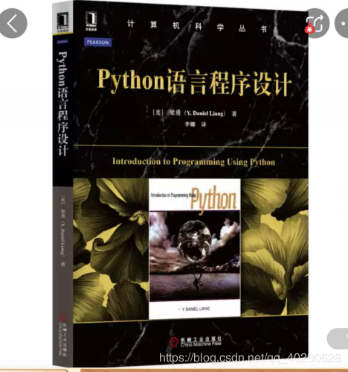
这里我推荐的编译器是PyCharm，别问什么，我要我觉得，不要你觉得（明学♂其实Jupyter据说也不错，读者可以试试）

下载网站：<https://www.jetbrains.com/>（其实他家的clion也不错，就是最近clion不知道抽了什么风）

**二、python的入门使用**

就是简单的简绍一下python的基本语法等等，我这里写了比较草率了。

如果你真的想学学python基础的话，我建议买计科类的《python语言程序设计》

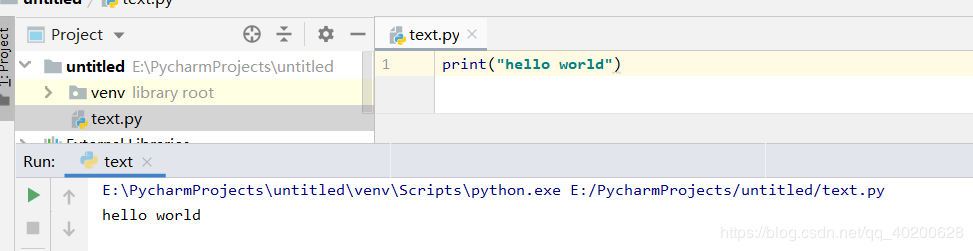
​

这里就不发链接了，免得有人说我打广告（滑稽）

可以配套这个<https://github.com/zhuwenzhang/Python-base>一遍自学一遍练习。

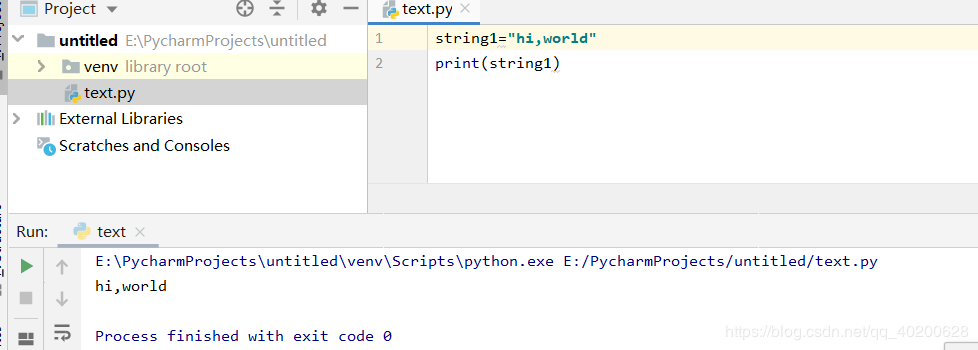
基本指令：

这里讲一个最简单的指令print，就是和c语言的printf函数是一样的。这里最简单的程序了

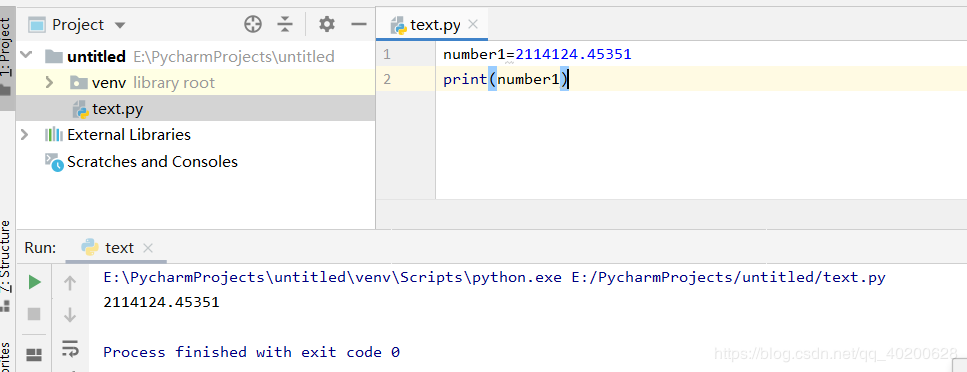
​

数据类型：

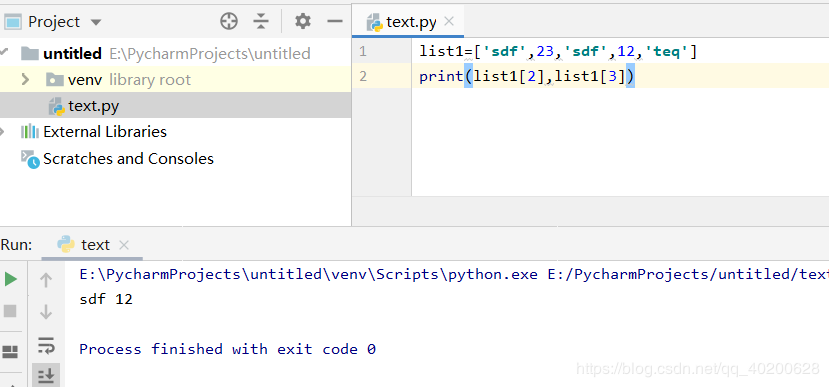
1.字符串（String）

​

2.数字（number）

​

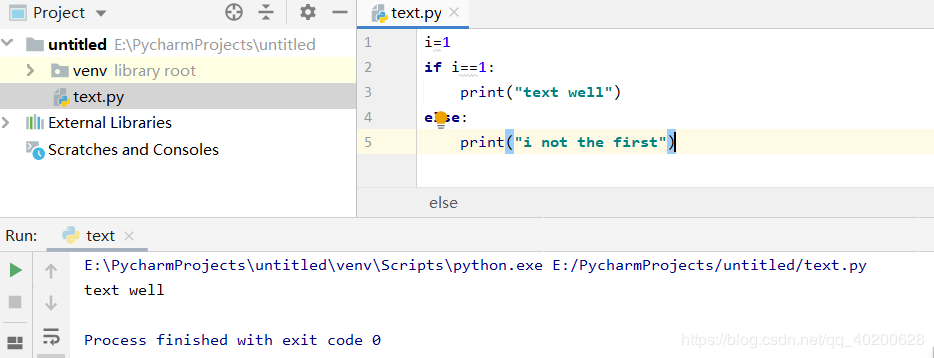
3.列表（list）

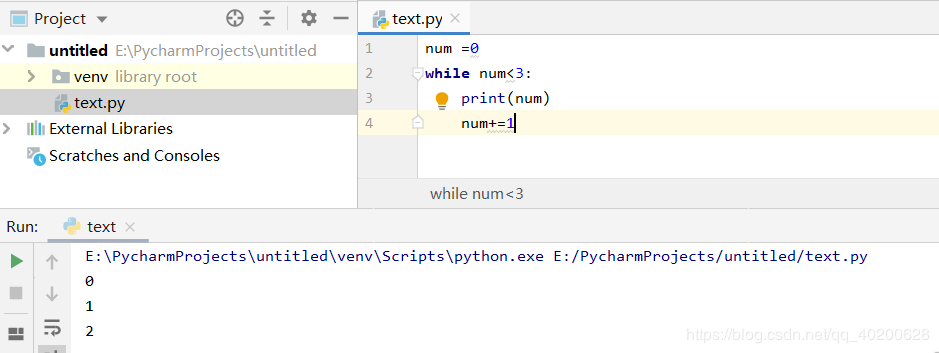
​

4.字典（dictionaries）

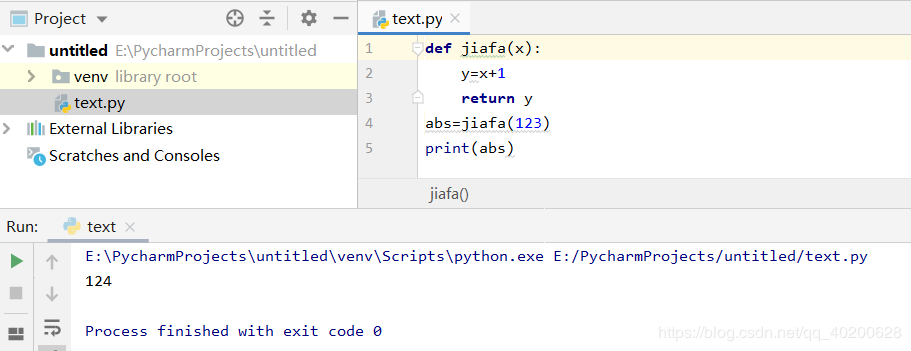
​

条件语句和循序语句：

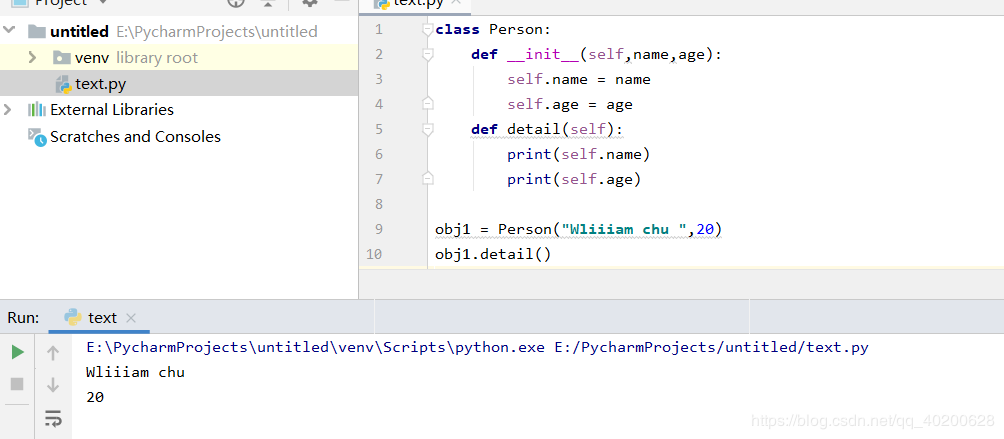
​

​

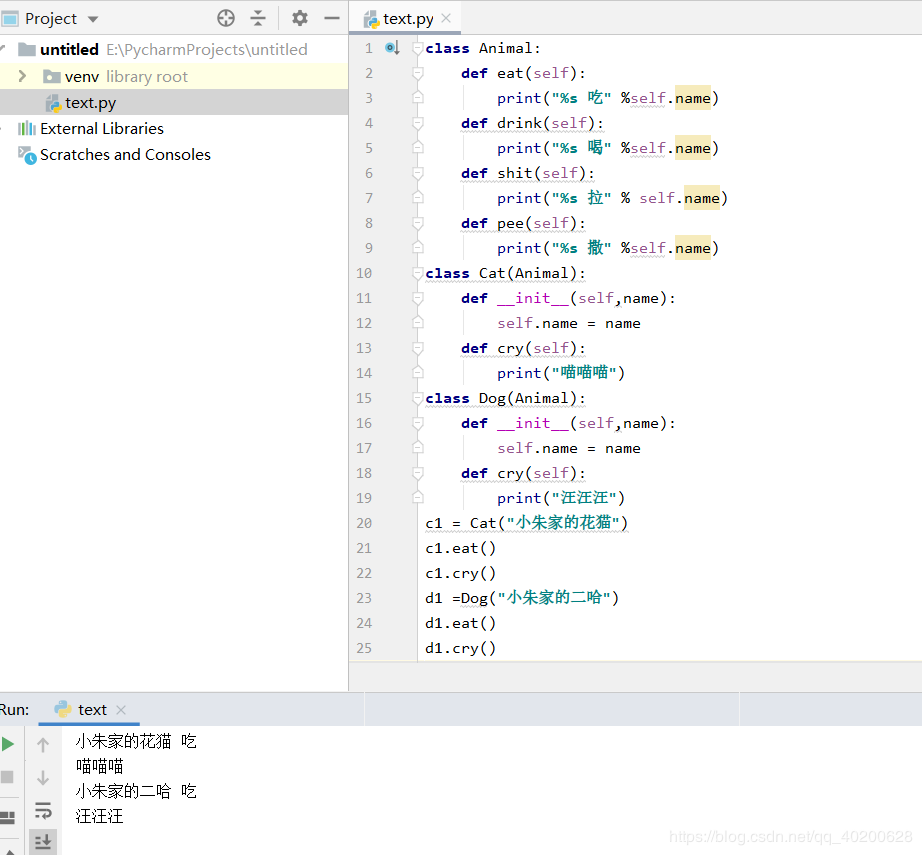
函数：

​

面向对象编程：

​

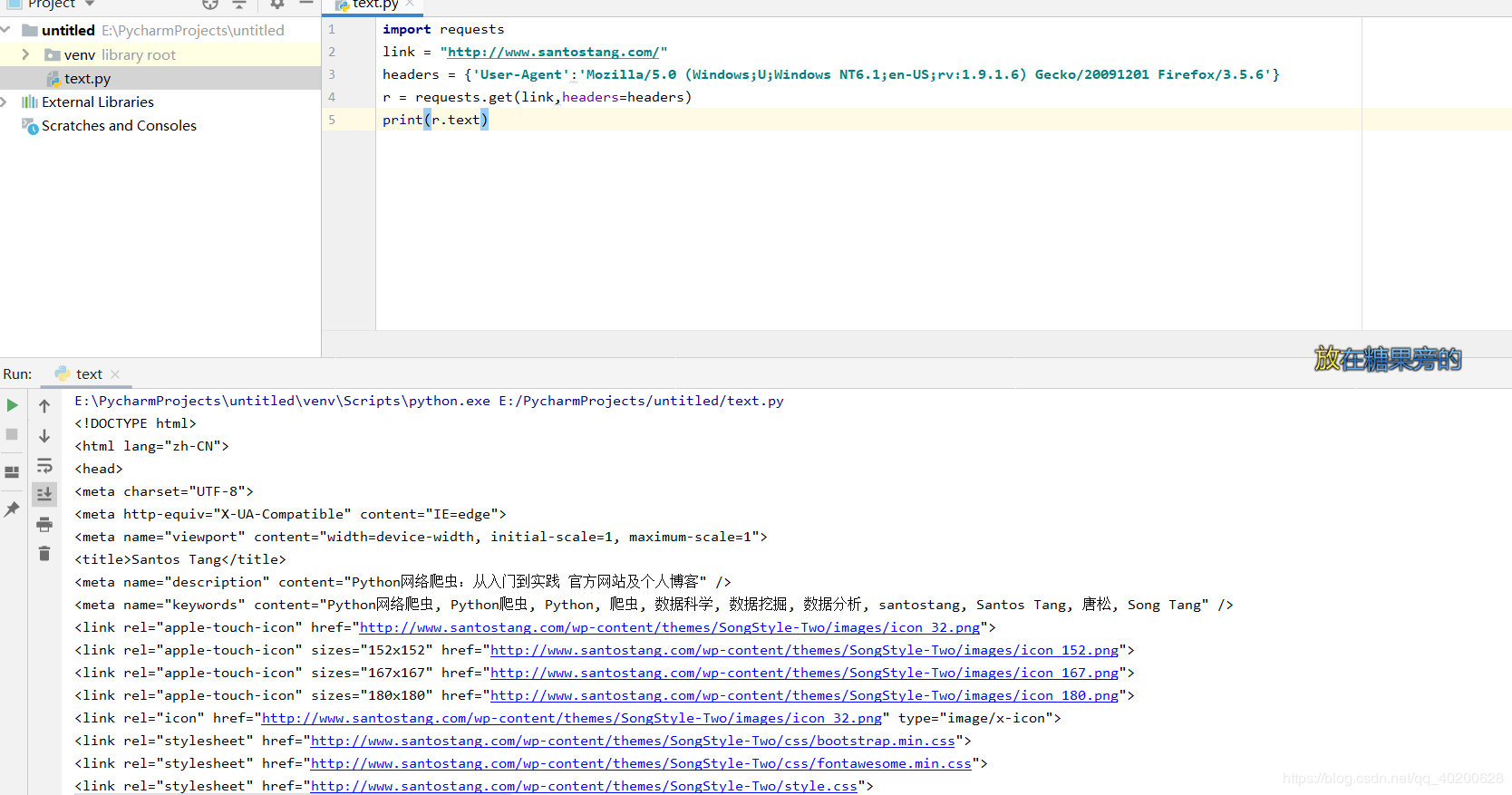
继承：

​

**三、编写第一个简单的爬虫**

激不激动，嘿嘿，这里以大佬唐松的博客为例：

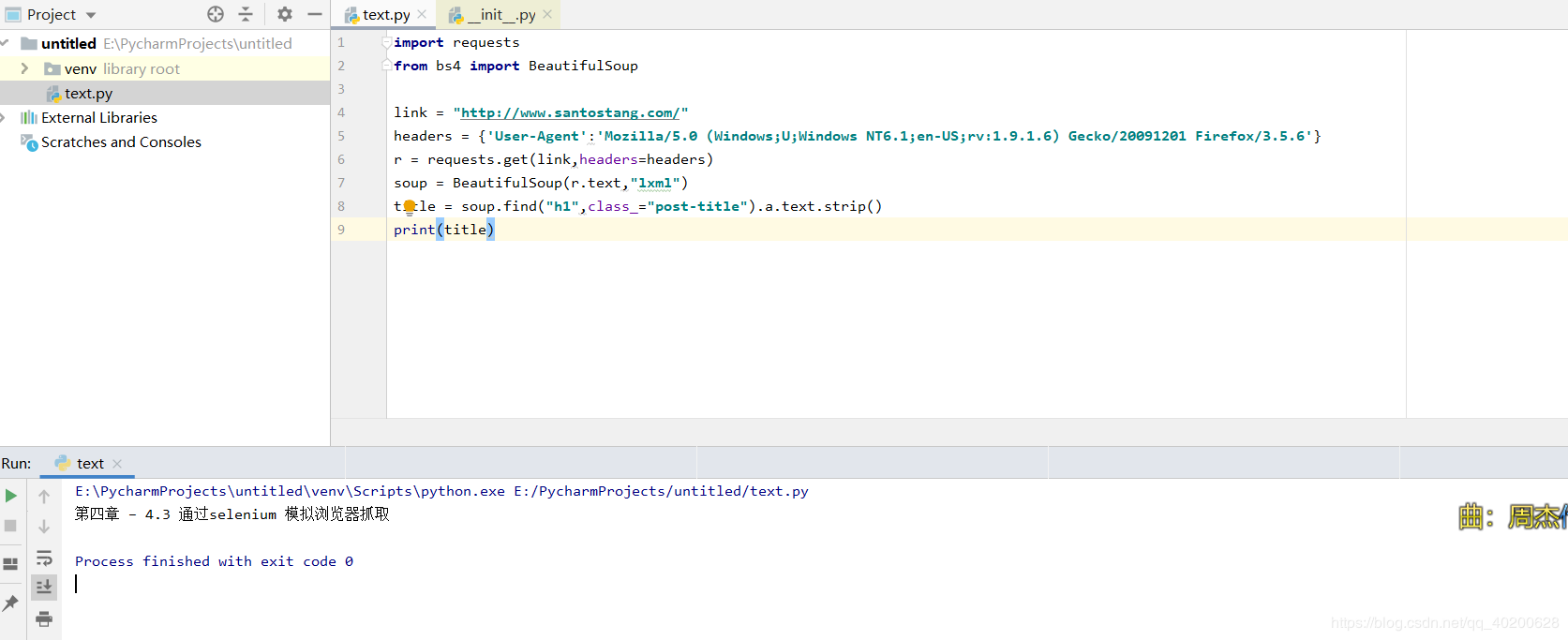
第一步：获取页面

​

这里的import requests，使用requests.get（link，headers=headers）获得网页。值得注意的是：

r这里是回复对象，我们从中可以获得想要的信息。r.text是获得的网页内容的代码

第二步：提取需要的数据

​

这里主要要重视的是要寻找出你要爬取的内容在哪里，这个你可以检测一下源代码进行查看，在就是系统的学一下bs4，小白的话就是把其中的find函数知道怎么用就可以了。

第三步：存储数据

保存数据到本地十分的简单

就加三行代码就行了

​

这里我们就可以看到title.txt了，你打开就是你提取的文字了。

**四、特别注意**

1.练习是最有效提高你编程能力的方式了

2.百度和google都是你最好的老师，遇到问题看到Stack Overflow基本上就解决了

3.不要复制粘贴要自己练习，复制粘贴一点用的没有，如果你是要提高自己的能力的话