Python网络爬虫学习：

## 爬虫基本库：

### urllib库：

是一个python自带的库，用来模拟浏览器的行为，向服务器发送一个请求，并可以保存服务器返回的数据。

#### Urlopen()函数：

创建一个表示远程url的类文件对象，然后像本地文件一样操作这个类文件对象来获取远程数据。

url:请求的Url

data:请求的data,如果设置了这个值，将变成Post请求。

返回值：返回一个http.client.HTTPResponse对象，是一个类文件对象。有read(size),readline(),readlines()以及getcode()等方法。

示例：from urllib import request

Res = request.urlopen()

Printf(res.read())

#### Urlretrieve()函数：

将网页上的一个文件保存到本地。

示例：from urllib import request

Request.retrieve(“<http://www.sougou.com>”,”sougou.html”)

#### Urlencode()函数：