一,环境准备

Python 的安装<http://pianshen.com/article/9328297421/>

二,scrapy环境的安装

参考文章 <http://pianshen.com/article/2293331286/>

三,scrapy框架的目录结构

1. scrapy.cfg: 项目的配置文件，现在可以先忽略
2. settings.py: 项目的设置文件
3. pipelines.py: 项目中的pipelines文件。

  Scrapy提供了pipeline模块来执行保存数据的操作。在创建的 Scrapy 项目中自动创建了一个 pipeline.py 文件，同时创建了一个默认的 Pipeline 类。比如我们要把item提取的数据可以保存到mysql数据库。

1. middlewares.py 中间件
2. item.py   Item 是保存爬取到的数据的容器；
3. \_init\_.py 暂时可以忽略