Python网络爬虫+网站（一个项目文件中），基于Windows平台开发，开发工具可以用VS、Pycharm等，Python使用3.5及3.5以上的版本

Python爬虫：使用python对目标网站进行数据爬取（目标网站：慕课网、CSDN学院、网易云课堂、腾讯课堂中任意三个或者是一些有较多学习课程资源的网站）网站中免费的课程，注意只要免费的课程。

主要爬取的课程类有四类：JAVA、C/C++/C#、Python、网络安全（信息安全）

目标数据：课程封面、课程名、主讲人（老师）、累计参与课程人数、评分（或者星级）、课程的超链接（点击可以跳转至此课程）、此课程下的评论数、评论数每门课程中点赞数最多的二十个评论

爬取的时候尽量遵守robots协议（robots.txt）或者爬取数据不能过快导致对目标网站服务器形成压力，防止目标网站的反爬技术封禁爬虫，爬取的数据存入数据库中。爬虫每周一次自动爬取并且更新数据库中的数据。

网站：

制作一个网站（用python）：

大致模样



网站用于浏览爬取到的数据（上面提到的数据库），

用户登录模块功能需求：注册、登录、修改密码，输入验证码才能登录

用户登录后可以查看自己收藏的资源

Python 数据清洗与可视化：

从爬取到的数据中进行分析（能整合放在网页另一个页面上最好）

1. 四类（Python、Java、网络安全、C/C++/C#）免费课程的数量

举例：

1. 四类（Python、Java、网络安全、C/C++/C#）免费课程的总参与人数

举例：

1. 网络课程的时长分布与此课程参与人数（散点图）

举例：

1. 此课程参与人数分布与此课程下的评论数对比（散点图）

要有讲解或者是开发文档，程序代码有相应的注释，比如这个类/模块是干嘛的，怎么实现的，这个变量是什么意思，网页结构解析原理等

开发周期：半个月到一个月 2020.04.17前给我

对上面有啥问题可以和我说