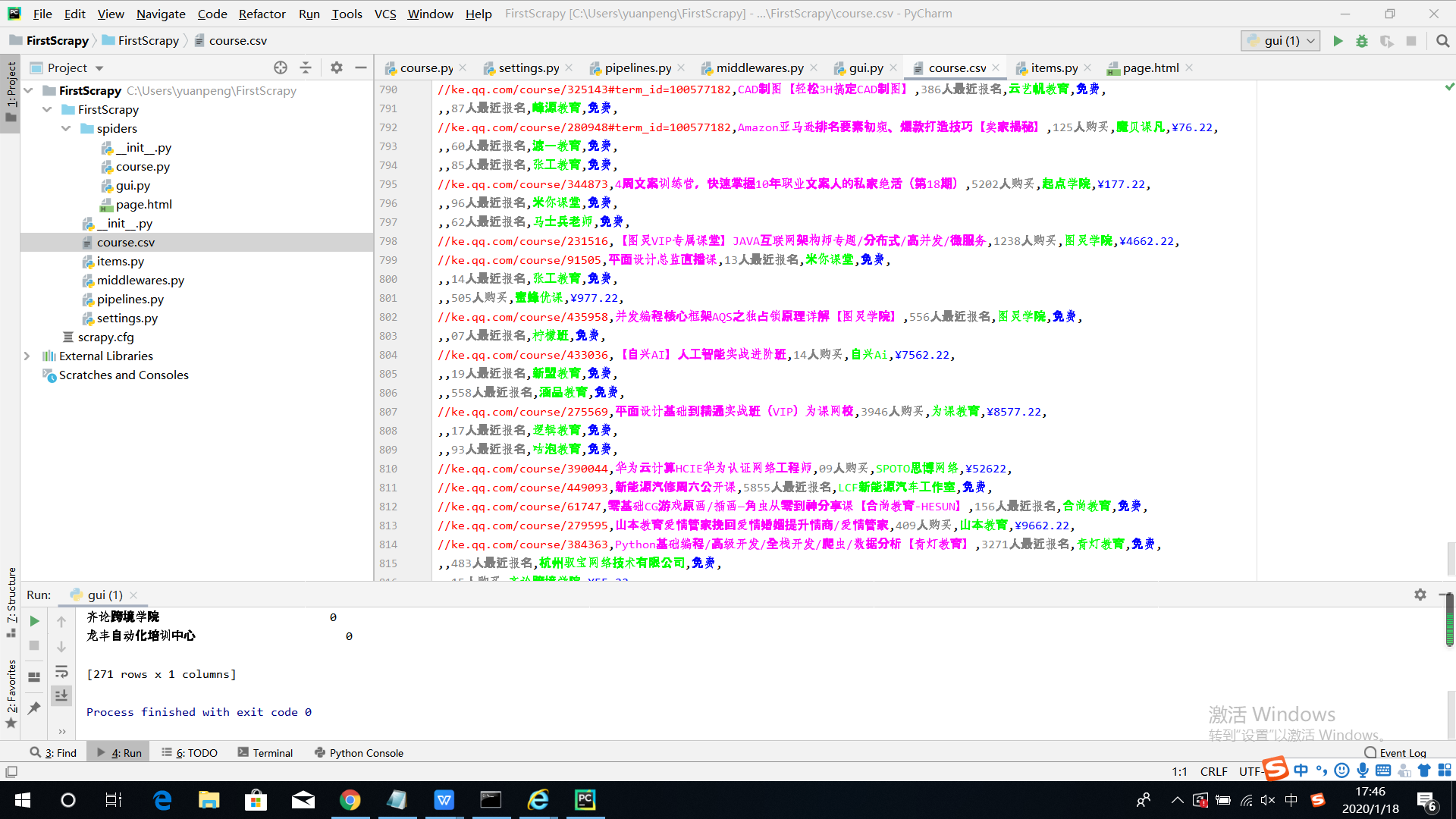
主要爬取腾讯课堂所有开课信息，包括课程名，培训机构，课程链接，课程状态，课程价格等信息，使用数据项，缓存爬取的数据字段，使用管道保存数据，爬取总共34页内容数据，使用分页结束，对url进行分析，使用正则表达式进行动态匹配，爬取内容如下所示，保存为csv格式。



然后使用pandas，numpy数据包进行数据分析，使用matplotlib进行绘图，对所有课程培训机构按组group by计数，看哪些培训机构在腾讯课堂上开的课程最多，有可能是广告做的好，或者课程推荐排行高。注意这里出现中文，需要解决中文乱码问题。

