说明：该站的反爬虫通过添加headers请求头和每个页面sleep 10秒的方式解决

2.（1）爬取<https://book.douban.com/top250?icn=index-book250-all>网站的图书信息，主要包括图书名（只需要爬取中文名称即可）、作者、出版社、出版日期、价格、评分及评价人数，将所爬取的输入保存为文本文件。

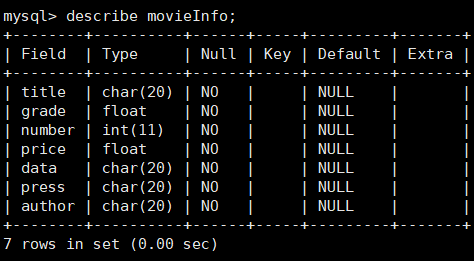
代码见 test.py

文件见data.data

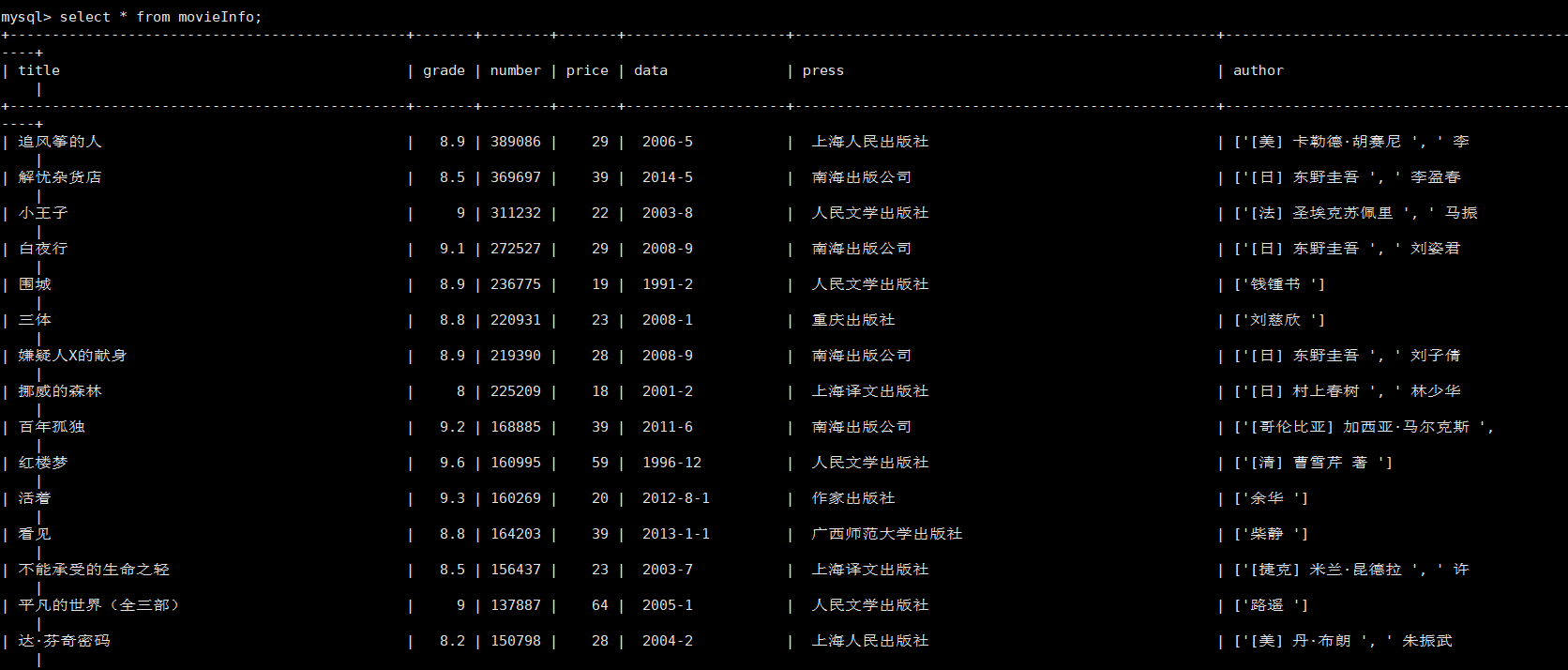
（2）将（1）中爬取结果保存到mysql数据库中。

该题自定义了操作数据库的类MySQLDB，在test.py文件中调用

数据库格式设置：



保存结果：



（3）对（1）中的数据进行处理，求出样本网页的图书平均分，结果保留两位小数。

该题使用re模块匹配文件中评分数据，将数据放入列表，最后用sum(列表)/len(列表)的方式求出