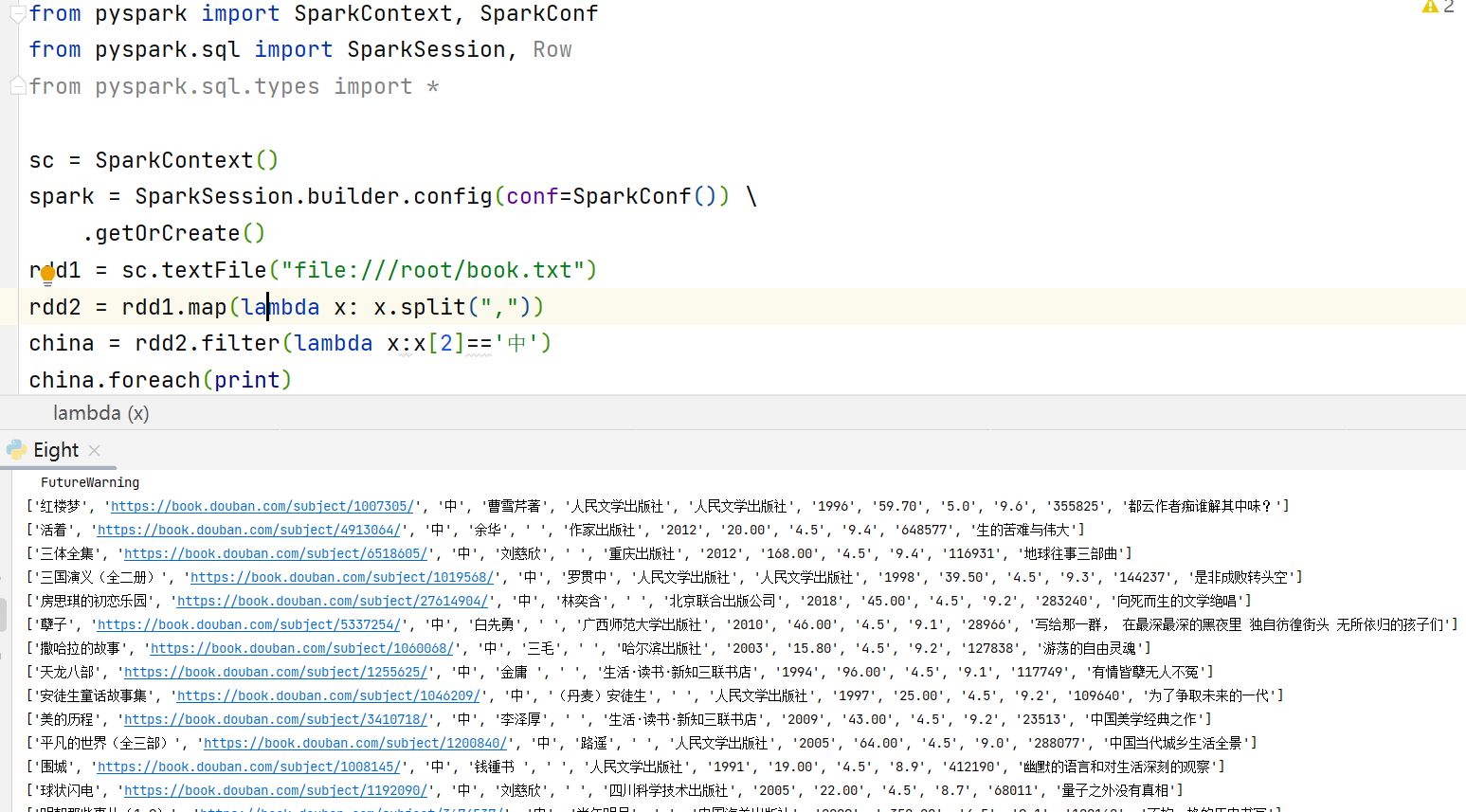
**实验八**

**【实验目的】**

**复习所学Rdd和DataFrame的内容**

**【实验内容】**

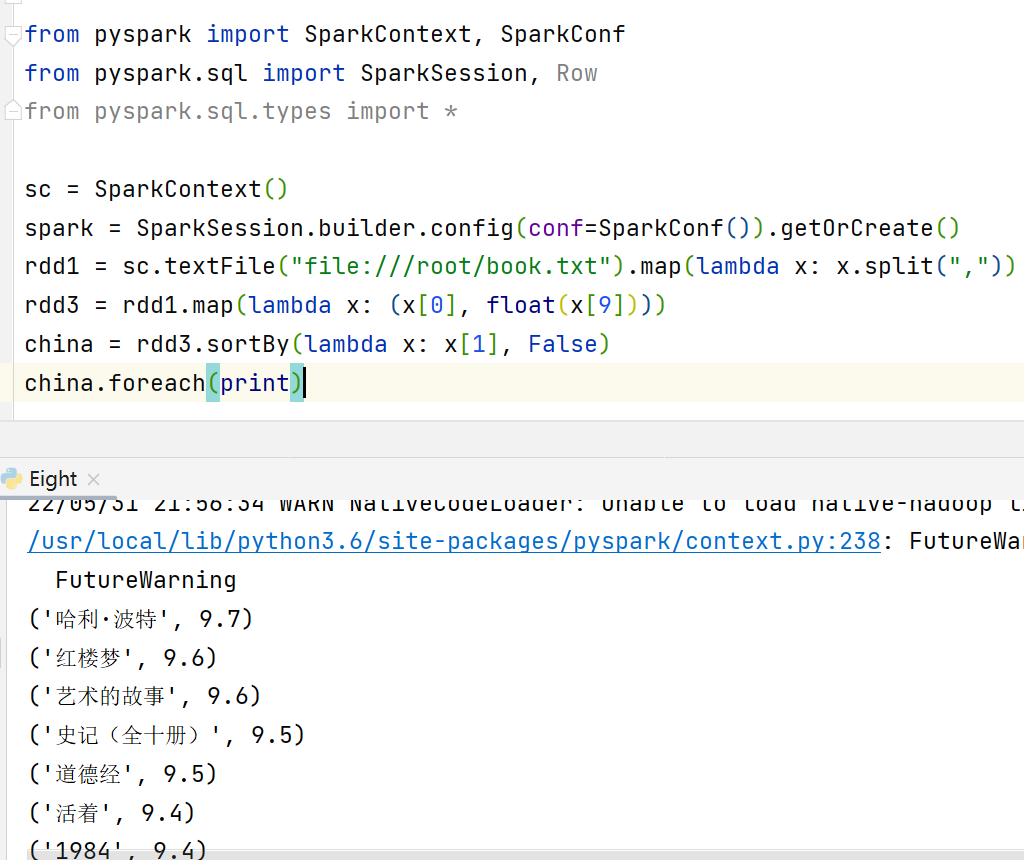
1. **现有book.txt，包含内容有（书名,链接,国家,作者,翻译者,出版社,出版时间,价格,星级,评分,评价人数,简介），请将book.txt转换成rdd，使用rdd各种转换操作完成下面题目：**
2. **筛选出中国的书籍；**



1. **筛选出2010年之后出版的书籍；**



1. **按照评分高低将书籍排序；**



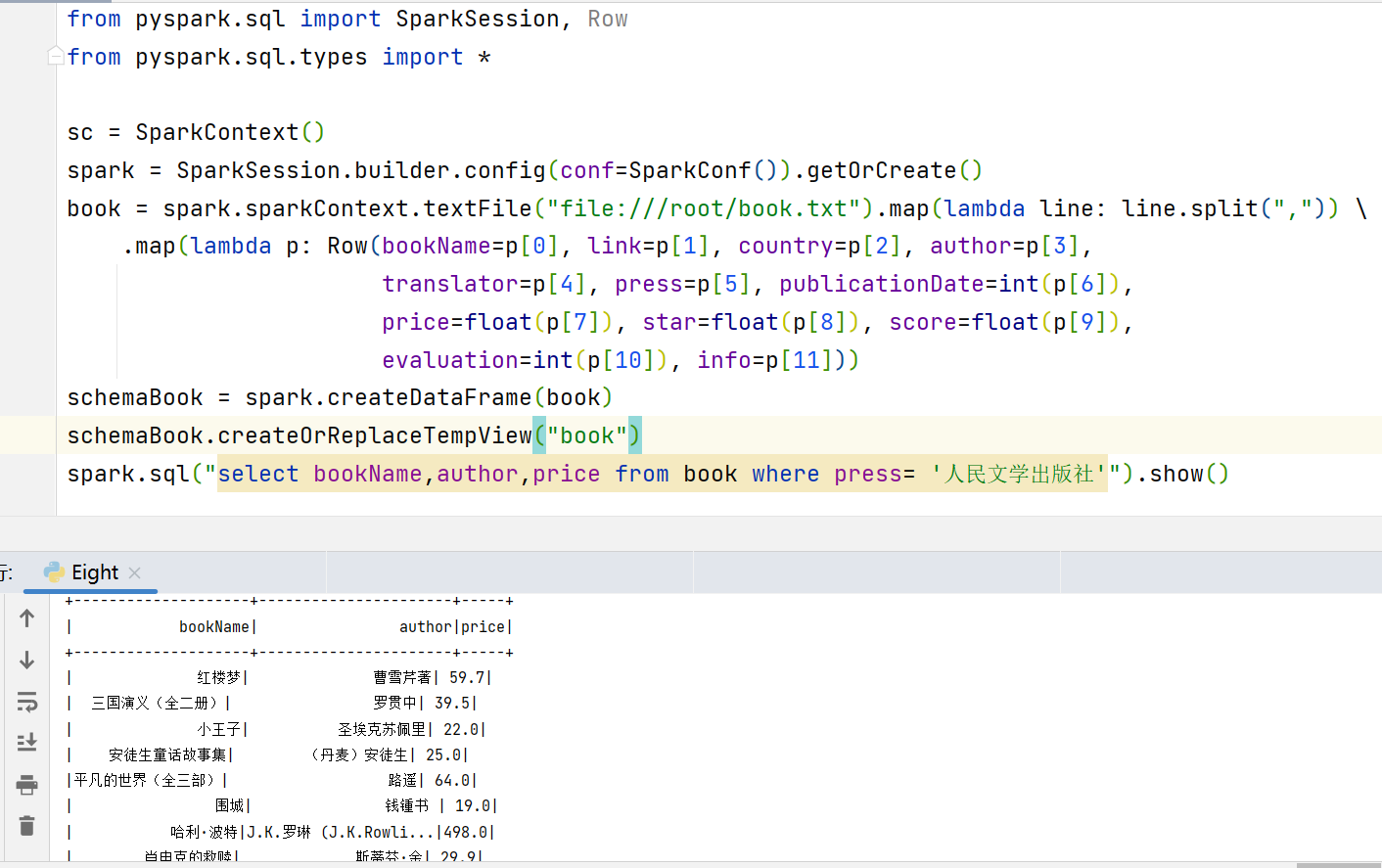
1. **按照出版时间从低到高排序书籍；**



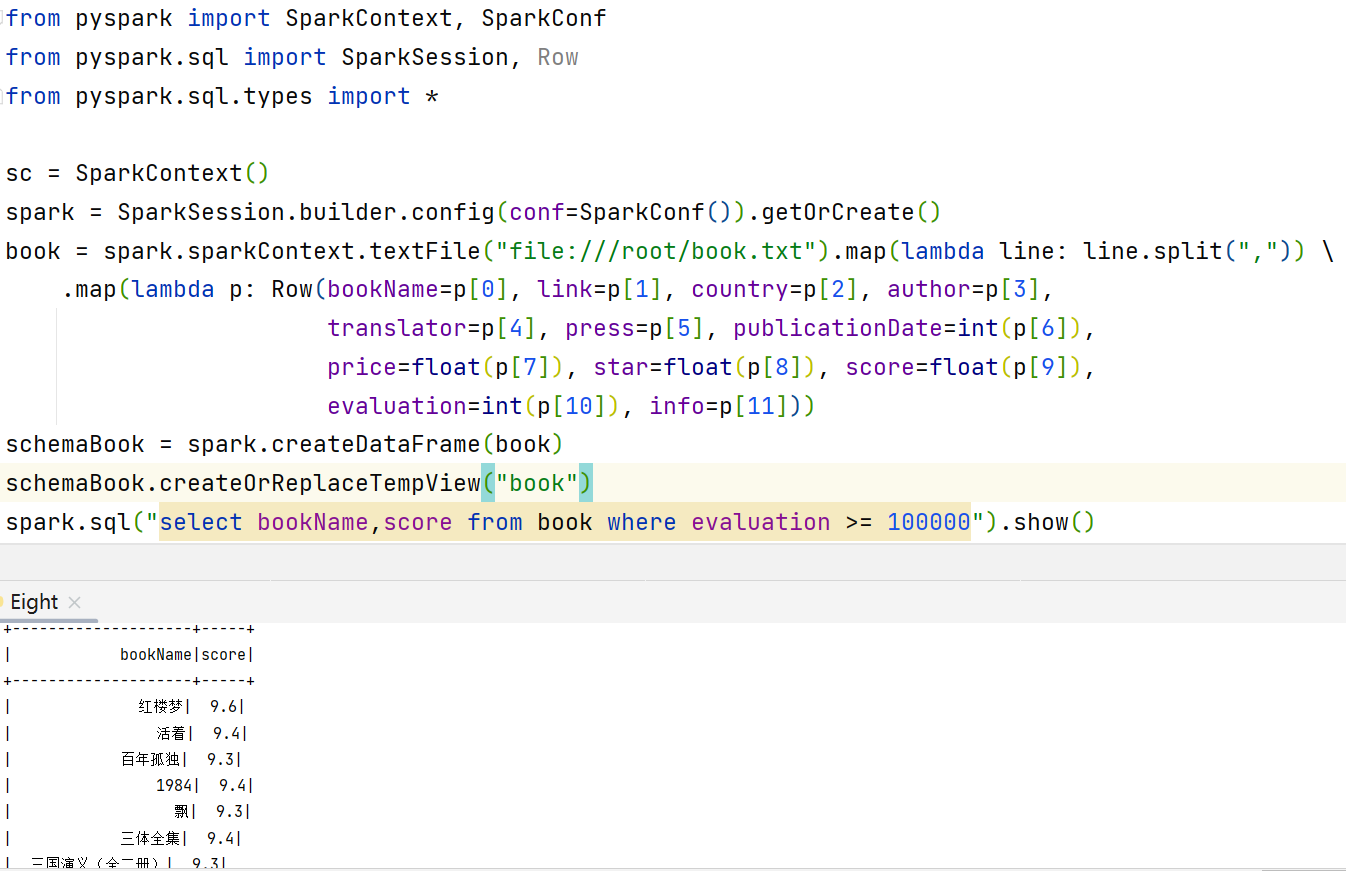
1. **数出不同国家书籍的数量；**



1. **将上题生成的rdd转换成DataFrame后并生成临时表进行查询完成下面题目：**
2. **查询“人民文学出版社”出版的书籍的书名，作者和价格；**



1. **查询评价人数超过100000的书籍的书名和评分；**



1. **查询不同星级书籍的数量。**

