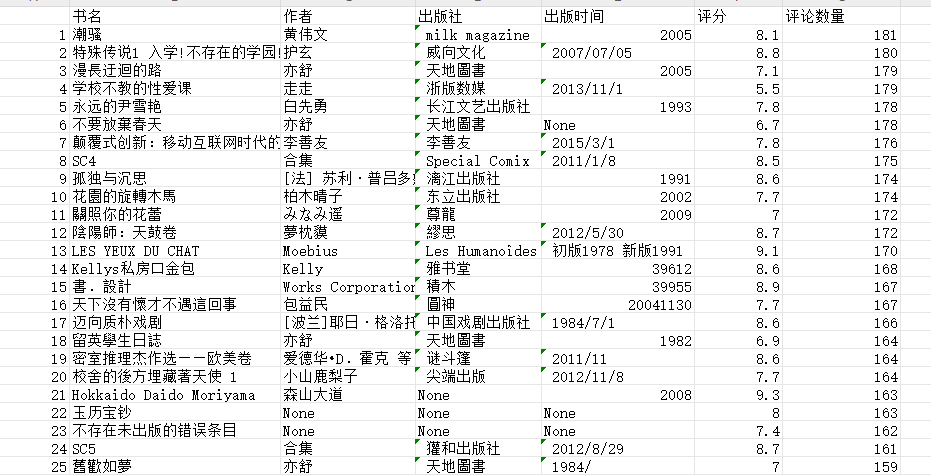
1. 将每条书籍评分记录存储至Mongdb中，并清理NULL值。从数据集中可以发现出版时间的数据格式多样，有1999,2012/12,1923-4,2019年六月，因此需要提取出其年份。最后把文件下载到本地（25分）



2. 由于相同的书籍可能出于多个出版社，为避免书籍评分重复导入，实现BF过滤器，每当有一部电影被存储后将其加入BF过滤器，并能够使用BF过滤器查询上述电影是否已经被存储。如果已经被存储则以列表的形式更新出版社的属性值，同时对于其评分和评论数量也需更新（20分）



**加分项：实现CF过滤器完成上述功能（15分）**

3. 为了判断近些年哪些书值得一读，根据书籍评分记录中的“出版时间”和“评分数量”构建-Tree索引。根据-Tree索引实现书籍出版时间和评论数量的范围查询，例如将出版时间在1990和2020之间以及评论数量大于50000的书籍提取出来，并按照评分降序排序，输出前十，结果如下：（25分）



**加分项：**选择书籍评分记录中的“出版时间”和“评分数量”任意一个属性值来构建基于ALEX的**学习索引。（15分）**