

# 中心主题

## ▼ 数据

- books.dbf
- labels.dbf
- mark.dbf
- users.dbf
- book\_labels.dbf
- bx-book-ratings.csv  
用户评分原始数据
- books.sql  
各个数据表的建立格式

## ▼ 数据挖掘与处理

- scrapy\_books.py  
根据标签爬取图书简单信息
- books\_info.py  
根据id爬取图书详细信息
- multi\_thread.py  
改进的多线程图书数据爬虫
- tags.py  
图书标签爬取
- crawl\_picture.py  
爬取图书封面
- process.py  
将用户评分的原始数据转换成该系统评分数据。

## ▼ 推荐模块

- redis\_star.py  
计算图书的平均分并将其存入redis中
- favoriate.py  
计算评价分数最高的图书列表，并将其存入redis中。
- recommend.py  
计算用户的推荐列表并将其存入redis中

## ▪ 书城