题目： 豆瓣网图书爬虫

简介：基于豆瓣网爬取评分较高的图书，并根据tag分析那些图书更容易受到人们的喜爱。并且分析人的年龄，职业，和社会阶级对不同图书的喜好的不同

技术：webMagic java 数据挖掘 K-means聚类分析等各种机器学习算法