

JD Notebook Analysis 使用手册

V1.0 2013.12.15

成员：范英明、赵仕荣、王贵存、李洪宇、牛童、马笑、李轶

目录

1	配置.....	2
2	使用说明.....	2
2.1	抓取数据.....	2
2.2	分析数据.....	2
2.3	显示分析结果.....	3
2.3.1	各品牌销量、平均售价、好评率.....	3
2.3.2	各品牌好评率及评论关键字.....	4
2.3.3	市场份额.....	5
2.3.4	CPU 分析.....	6
2.3.5	屏幕尺寸分析.....	6
2.3.6	价格区间分析.....	7
3	致谢.....	8

1 配置

测试支持的服务器：apache-tomcat-7.0.47。

数据库：MySQL；

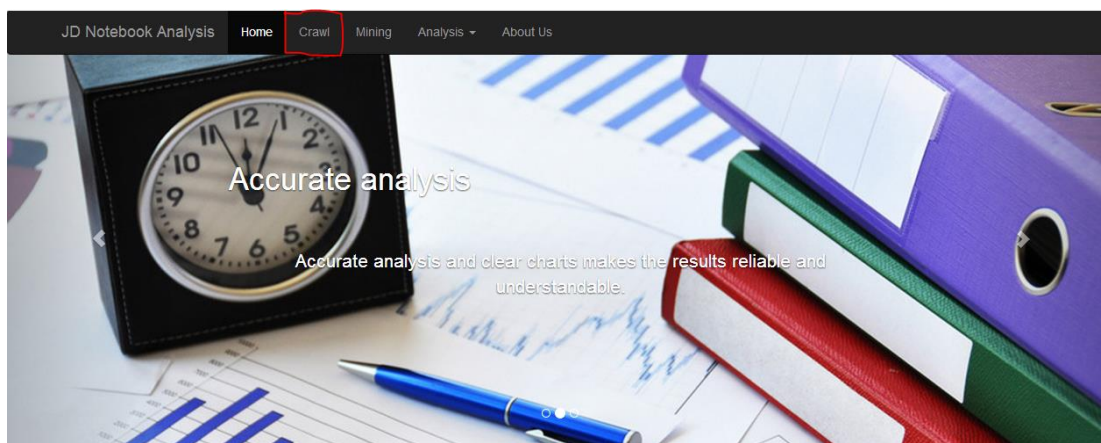
说明：将 tomcat 服务器配置好并开启，且保证安装完 MySQL 数据库，然后在 GlobalData 类里设置 MySQL 的用户名和密码；爬取数据速度较慢，建议从附带的 jdcrawler.sql 导入数据。

项目地址：<https://bitbucket.org/YingmingFan/datamining>

2 使用说明

2.1 抓取数据

点击如图 2-1 中的 Crawl 按钮，系统将开始抓取 JD 网的笔记本电脑数据；系统采用 4 线程同时抓取，抓取过程中将实时显示目前抓取到的数据条数，当抓取结束后，将提示“抓取完毕”，用户也可以手动点击“停止抓取”按钮进行停止。

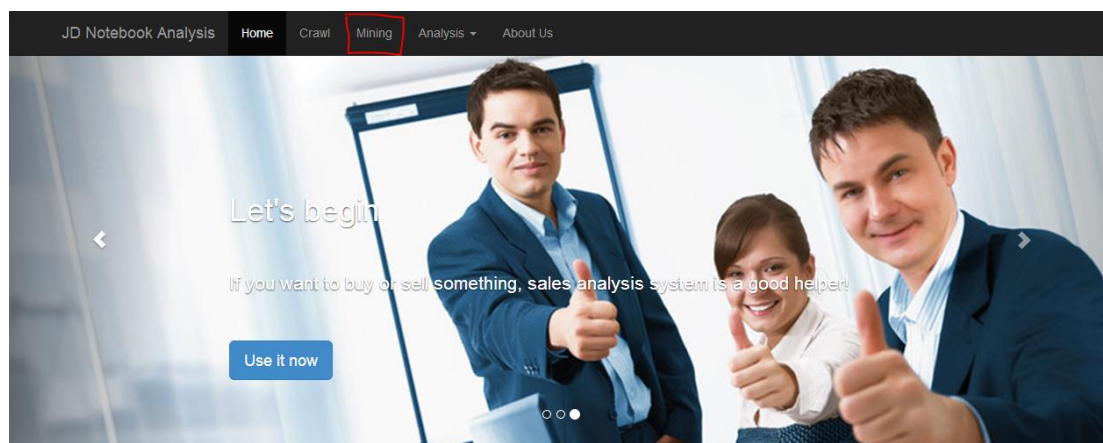


Welcome To JD Notebook Analysis!

图 2-1 抓取数据页面

2.2 分析数据

点击如图 2-2 中的 Mining 按钮，系统将会分析抓取到的数据，并将结果存到项目 data 文件夹下，以供用户通过可视化的工具进行查看；分析数据需要的时间根据数据量决定，当分析完数据后，会提示用户，如图 2-3 所示。



Welcome To JD Notebook Analysis!

图 2-2 分析数据按钮

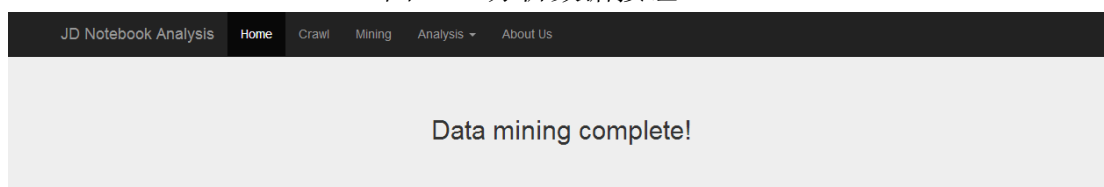


图 2-3 分析完数据的提示

2.3 显示分析结果

2.3.1 各品牌销量、平均售价、好评率

点击“All Brands Computer Sales Analysis Chart”按钮，如图 2-4，将查看各个品牌的销售量、平均价格以及好评率，如图 2-5。

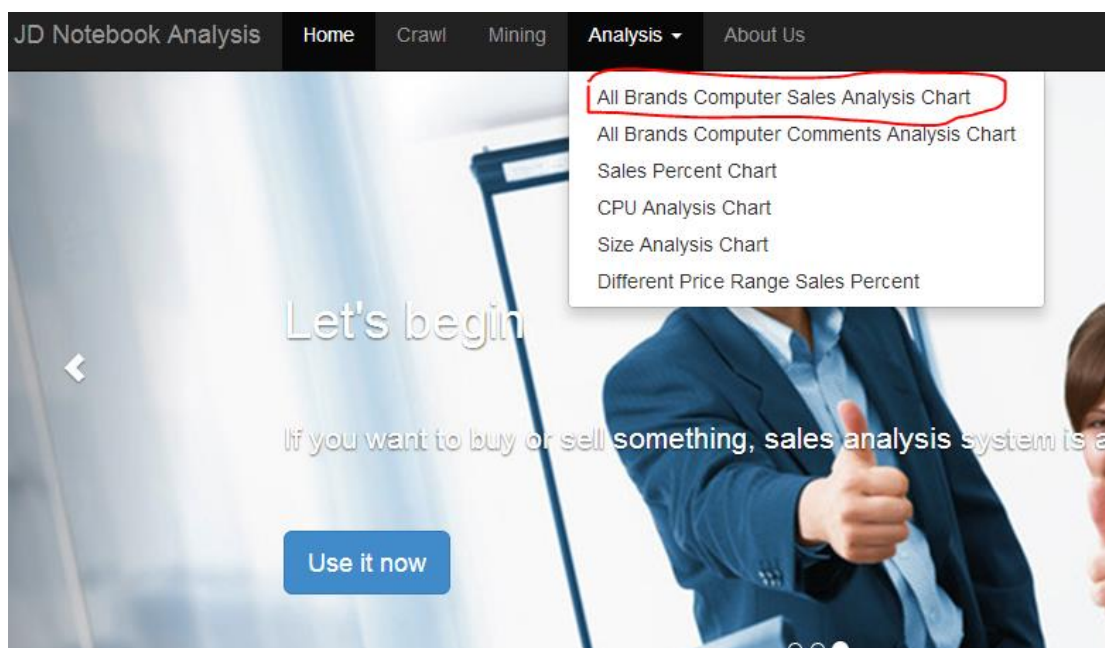


图 2-4

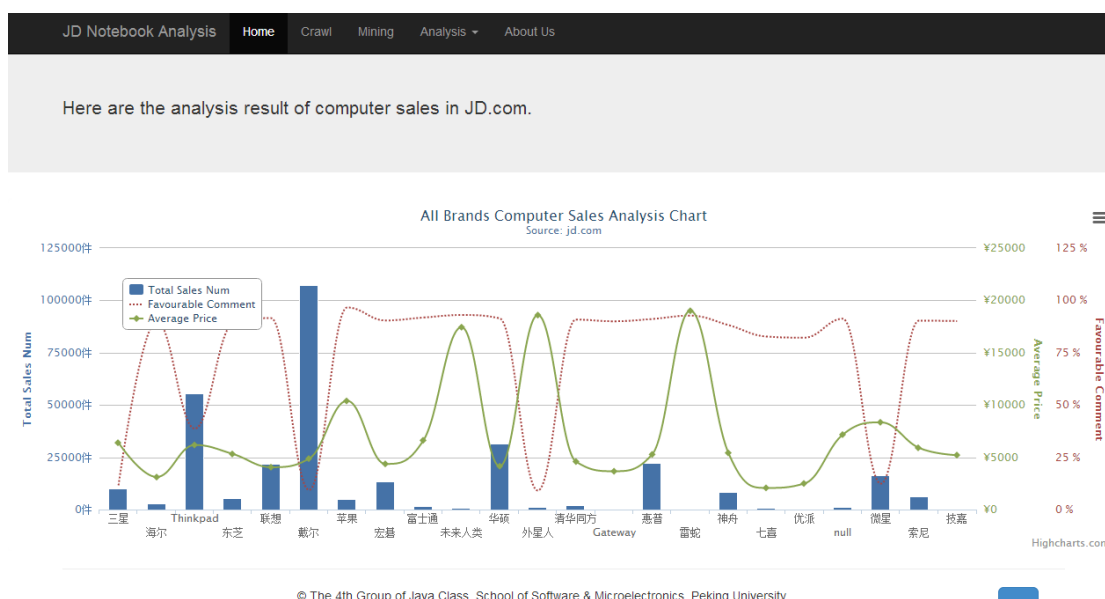


图 2-5 各品牌的销量、平均售价、好评率

2.3.2 各品牌好评率及评论关键字

点击“All Brands Computer Comments Analysis Chart”按钮，将进入如图 2-6 的页面，显示各个品牌的好评率；点击某个品牌的柱状图，将进入如图 2-7 的页面，将显示当前品牌所有评价的关键字中出现频率最多的词。

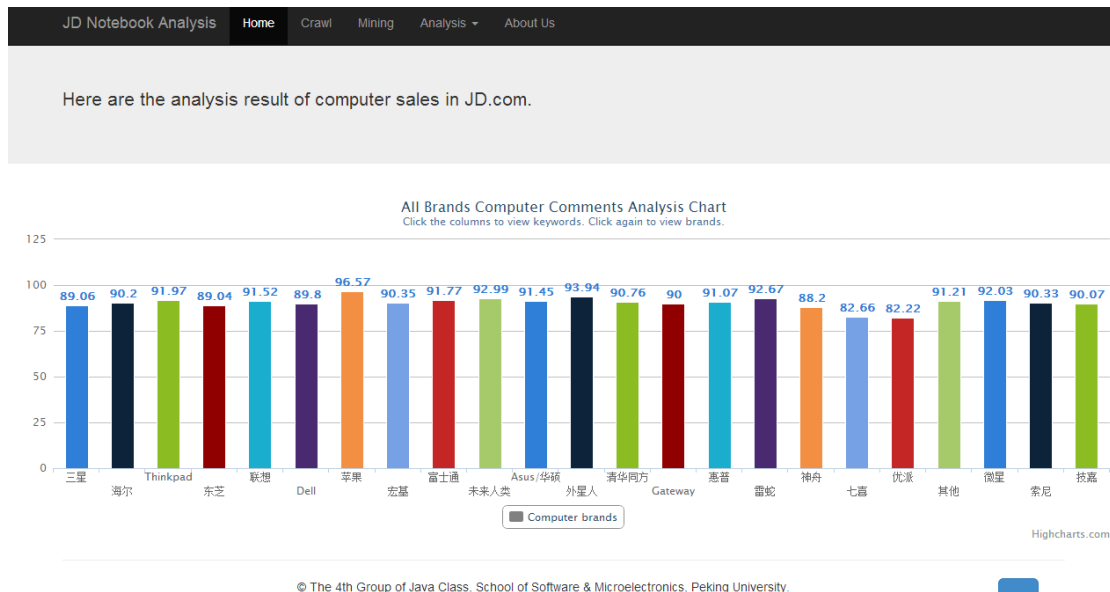


图 2-6 各品牌好评率

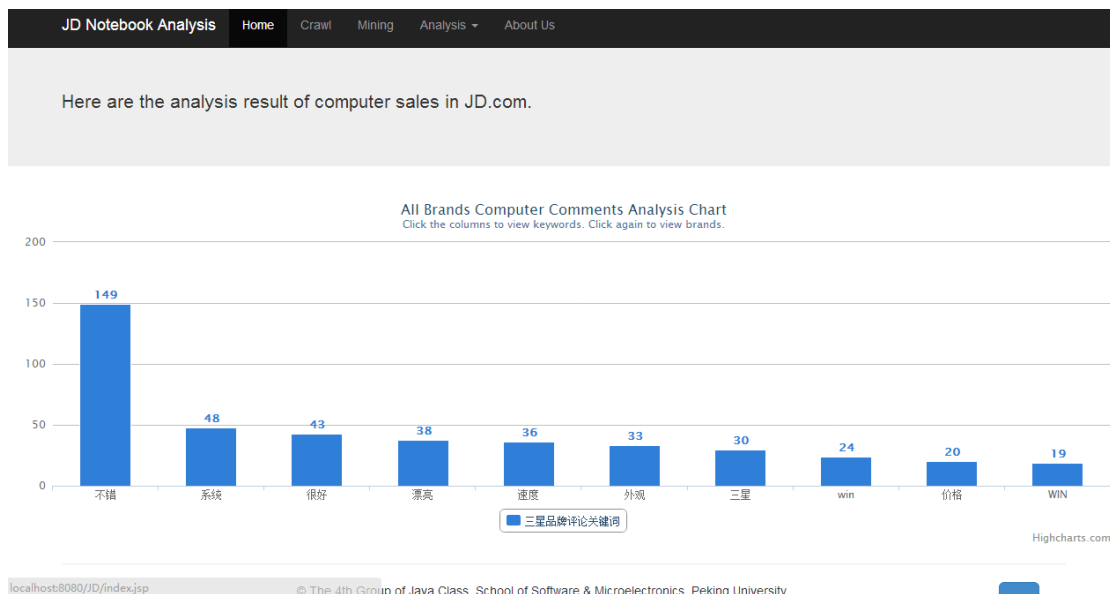


图 2-7 当前品牌下评论中关键字出现的个数

2.3.3 市场份额

点击“Sales Percent Chart”按钮，将进入如图 2-8 的页面，饼状图显示每个品牌在 JD 销售笔记本中所占的比例，点击某个品牌，这个品牌的饼会突出显示。

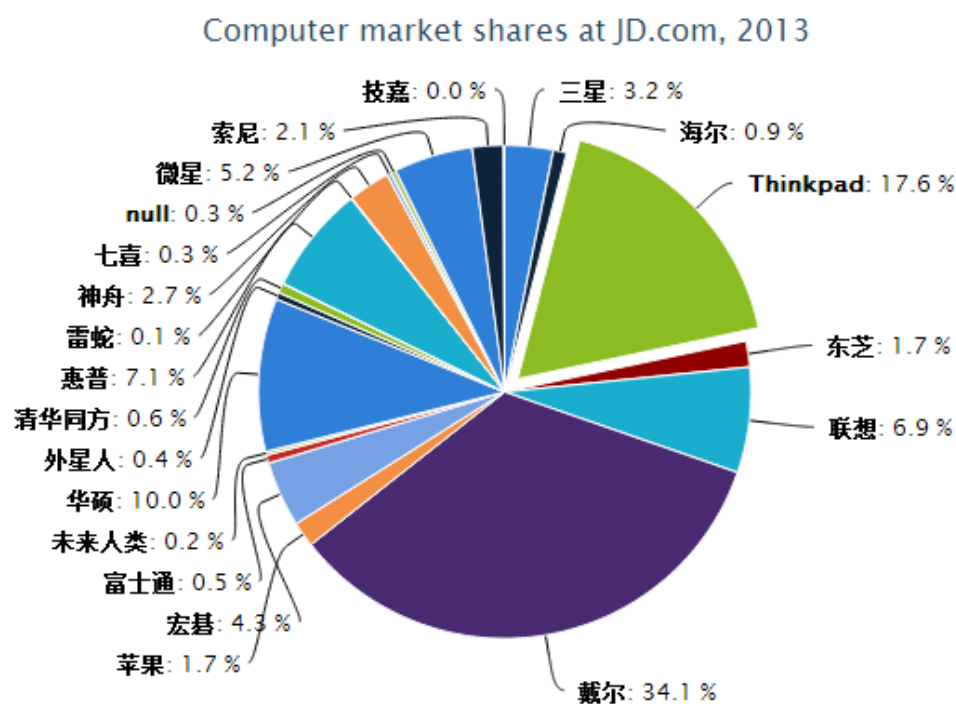


图 2-8 各品牌在 JD 中的市场份额

2.3.4 CPU 分析

点击“CPU Analysis Chart”按钮，将显示如图 2-9 的页面，页面中显示各个 CPU 的价格、销售量，左上角还显示每个 CPU 的型号个数。

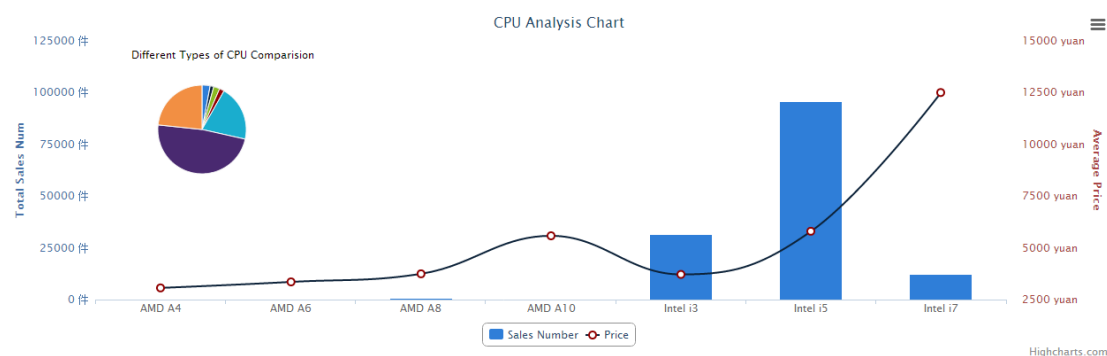


图 2-9 各 CPU 的价格和销售量以及型号个数

2.3.5 屏幕尺寸分析

点击“Size Analysis Chart”按钮，将显示如图 2-10 的页面，页面展示了各个屏幕尺寸的笔记本的销量、平均价格以及各个尺寸的笔记本型号数。

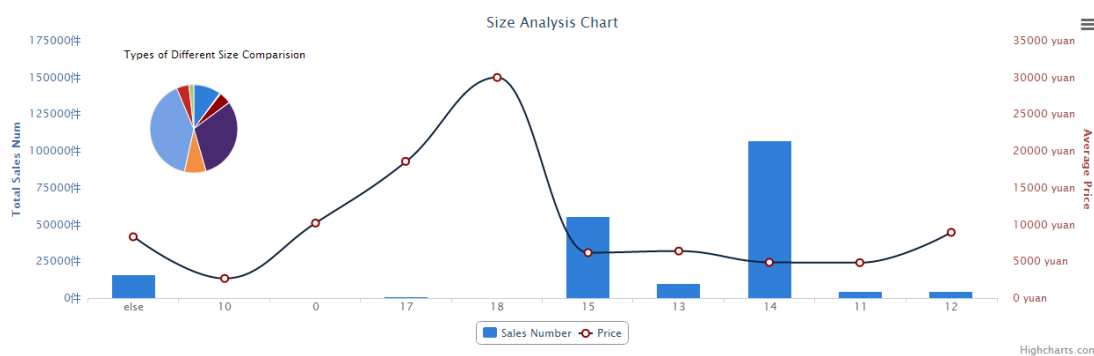


图 2-10 各屏幕尺寸的价格、销量、型号数

2.3.6 价格区间分析

点击“Different Price Range Sales Percent”按钮，将显示如图 2-11 的页面，页面展示了不同价格区间下，各个品牌的销售份额。

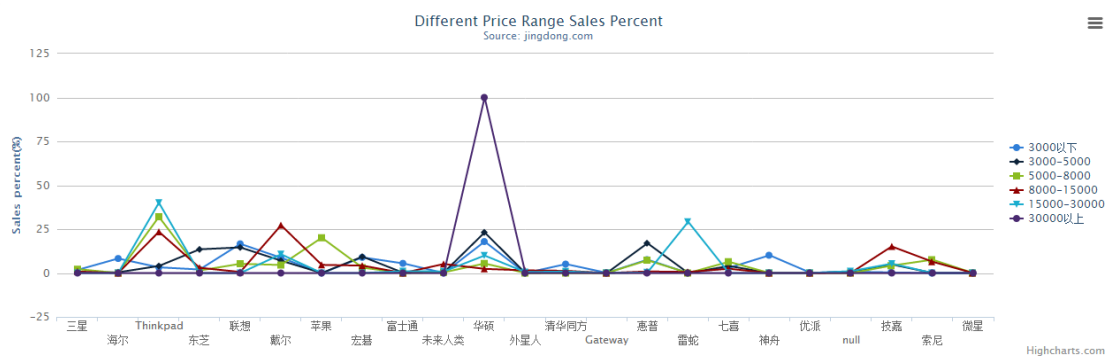


图 2-11 各品牌在不同价格区间下的销售量

3 致谢

感谢小组成员的努力付出，感谢 Highcharts 项目的开源贡献。



Members Introduction

We are a TEAM!