Broadview®
www.broadview.com.cn

赠送本书教学课件PPT
涵盖13个电商平台的数据工具箱
高校电子商务数据实验室必备



第2版

基于Excel和SmartMining玩转电商数据分析

淘宝、天猫
电商数据分析与挖掘实战

零一 著

数十所院校共同选择的电商数据分析教材
电商运营与新零售接轨的必修课

中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

需要全本请访问 www.ertongbook.com

第2版

淘宝、天猫

电商数据分析与挖掘实战

零一 著

学院图书馆
藏书专用章

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书主要针对电商从业者（运营和店长）和数据分析入门者，以电商业务实战为主线，介绍数据分析相关的知识。本书的上半部分主要介绍淘宝的操作方法，以及探讨未来的电商布局之路。本书的下半部分以实战为主，主要介绍淘宝卖家如何应用 Excel 和数据来做决策。数据从来都离不开业务层，数据分析师必不可少的 4 个要素是思维、业务、工具和数据，而前两者才是最重要的。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

淘宝、天猫电商数据分析与挖掘实战 / 零一著. —2 版. —北京：电子工业出版社，2017.10
ISBN 978-7-121-31950-1

Ⅰ. ①淘… Ⅱ. ①零… Ⅲ. ①电子商务－数据管理 Ⅳ. ①F713.36

中国版本图书馆 CIP 数据核字(2017)第 139691 号

策划编辑：张慧敏
责任编辑：王　静
印　　刷：北京富诚彩色印刷有限公司
装　　订：北京富诚彩色印刷有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编：100036
开　　本：787×980　1/16　印张：17.5　字数：346 千字
版　　次：2016 年 3 月第 1 版
　　2017 年 10 月第 2 版
印　　次：2018 年 1 月第 2 次印刷
定　　价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

第 2 版序言

自阿里提出“数据赋能”以来，越来越多的电商企业和运营开始关注数据的应用，2013 年淘宝运营技巧的爆发，其实就是在享受数据带来的红利。业内人说阿里唯一不变的就是变化，互联网的节奏是“快”“狠”“准”，适者生存，优胜劣汰，只要慢上一两步，时机将一去不复返。数据产品也在变化的过程中更迭，本书第 1 版中涉及的淘宝指数、数据魔方都下线了，同时新版的阿里指数、生意参谋上线了。Excel 从 2013 版本更新到了 2016 版本，SQL Server 从 2012 版本更新到了 2016 版本，SQL Server 数据挖掘套件停止了更新，据悉，Excel 将会集成更多的分析方法。因为有太多的变化，很多读者写邮件给我希望可以更新本书内容。

回顾本书第 1 版，虽然获得了很多读者的认可，甚至有不少高校将其改编成了教程，但本书依旧有许多不足之处，比如，对于运营模块内容较少而且不够深入，特别是数据挖掘套件的安装和部署难倒了许多读者。

本书第 2 版为了跟上时代的变化和解决软件部署烦琐的问题，首先将线下的数据平台改成新平台，或是用其他平台替代；其次是对于软件方面的升级，将 SQL Server 的挖掘套件换成了 SmartMining 个人版，部分章节使用了 Excel 2016 版本，用 Power Pivot 实现数据建模，用 Power Query 提升数据清洗的能力。

第 2 版的内容更加充实饱满，新增的实例分别在 3.2.6 节和 3.4.3 节，给读者参考以提升分析能力，4.2 节简单介绍了阿里的生意参谋数据平台，4.4.2 节介绍了运营数据建模的方法，5.5 节介绍了降维的方法。

本书可作为电商行业数据分析的高级教程。

零　一

2017.5.26

前言

数据有毒，我也会被数据误导，因为平台提供的数据可能并不精准，甚至误差极大。我在操作单品时发现过生意参谋里面统计的关键词数据有错误，例如我通过实时访客掌握了成交词为 A 词，并确认无误，但第二天的单品统计数据告诉我昨天的成交词是 B 词，如果我没有掌握真实的信息，那么我就会着重去优化 B 词，甚至在后面干脆就把 A 词删掉了，因为数据告诉我 A 词没有任何效果。

数据原本的作用是帮助我们做决策，一旦数据产生错误，就可能导致我们做出错误的决策。但没有数据，我们又只能是一个瞎子，毫无方向感。

所以，使用数据要慎重、胆大、心细，同时也要看我们利用数据来做什么。对商家而言，数据有两种用途：第一种，用来做预测，通过参考数据从而决定卖什么货；第二种，用来指导下一步的运营操作。对于第一种用途，数据仅作为参考，因为我们对市场的判断，除数据外，还有自己的行业经验。对于第二种用途，则建议多验证数据，例如当我看到 B 词有数据时，通过搜索 B 词发现我的商品排名很靠后，不用想都知道排名这么靠后，买家怎么可能会找到我的商品并下单呢?

本书的上半部分主要介绍淘宝的操作方法，以及探讨未来的电商布局之路，这部分内容有时效性。

本书的下半部分主要介绍淘宝卖家如何应用 Excel、SmartMining 和数据来做决策。数据从来都离不开业务层，数据分析师必不可少的 4 个要素是思维、业务、工具和数据，前两者才是最重要的，希望读者在阅读本书时着重研究业务层面的内容。

需要全本请访问 www.ertongbook.com

说到业务层面，在电商人的眼里，一切皆流量，流量是电商赖以生存的食粮。平台流量和社交流量是可以相结合的。在电商运营中有一个 6 字真言——引流、转化、留存，这 6 个字概括了电商运营的精髓，字字珠玑，都不简单。无数的运营者夜以继日地研究的是前 4 个字，而最后两个字“留存”却鲜有运营者钻研。我们曾经统计过获取一个新客户的成本是维护老客户的 100 倍。“引流、转化”是让电商暂时得以生存，“留存”则是决定了电商能存活多久。未来，流量会越来越贵，如果我们做不出能够粘着用户的产品，那么就用人去粘着人。

用人去粘着人的前提是消费者在我们这里有过很棒的购物体验，做体验其实并不难，关键是卖家要有做体验的意识。我的淘宝集市店发货用的是顺丰快递，甚至退换货也用顺丰快递，我的换货流程是：顺丰快递员到我们这里拿一个新品送到消费者手里的同时，把旧的送回给我们。所有产生换货的消费者都会给我们好评，因为他们想不到一家集市店的购物体验居然能做得和京东一样，甚至超过京东（我是以京东的购物体验作为标杆的）。消费者会因为物流的时效、赠品或者产品本身而感到震撼，他们在我这里爽过了之后，我的个人微信号的加粉率接近 100%，因为好的购物体验让消费者对素未谋面的淘宝集市卖家产生了信任。

数据要落到实处，必然是要落到业务上，落到人上，技术可以不精湛，但思维要跟上，数据时代已经到来！

谁适合看这本书？

如果对下面的所有问题都能肯定地回答“是”，那么本书适合你。

① 你是不是淘宝店铺的运营人员或者店长，目前还不会做数据分析，渴望提升自己？

② 你是不是打算在淘宝开店，目前尚在学习中？

③ 你是不是对数据分析感兴趣，尚在入门阶段？

④ 你是不是更愿意亲自动手实操，在实践中应用所学，而不是看我吹嘘？

谁可能不适合看这本书？

如果满足下面任何一种情况，那么本书不适合你。

① 你是不是已经擅长数据分析？

② 你是不是看到数据就犯困，对数据毫无兴趣？

③ 你是不是只愿意花时间抱怨，做不好就找借口把责任推给他人？

如何联系作者？

作者的联系方式：

邮箱：lingyi187@qq.com

微信公众号：start_data

本书可作为电商行业数据分析的高级教程，已有数十所院校选用此书作为电子商务专业数据分析课程专用教程。

零　一

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- **提交勘误**：您对书中内容的修改意见可在 提交勘误 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动**：在页面下方 读者评论 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：http://www.broadview.com.cn/31950



目录

上篇 做数据分析之前要先学会做淘宝

业务能力是数据分析师成长最大的瓶颈之一，做淘宝数据分析当然要先学会做淘宝，只有了解淘宝是如何操作的，才能更好地理解淘宝数据。

第 1 章 什么是淘商？未来如何布局 1

1.1 什么是淘商 2
1.2 未来如何布局 2
1.3 电商媒体化和媒体人格化 4

第 2 章 让店铺活下来 5

2.1 小卖家的生存之道 6
2.2 中大卖家的生存之道 15
2.3 引火线——流量之战 17

下 篇　跟着实战学做数据分析

> 数据分析师必须实战，只有把数据应用到业务上，然后通过市场反馈来验证自己的分析结果，才能积累属于自己的经验。

第 3 章　市场分析——告诉你这个市场能不能做，能做多大 19

3.1　数据分析的五大思维方式 20
3.2　如何评判一个行业市场的规模 24
3.3　进一步研究市场的方法 91
3.4　通过竞争对手分析了解行业态势 127

第 4 章　运营分析——万变归宗 148

4.1　了解运营数据 149
4.2　生意参谋数据的解读 155
4.3　店铺数据规划 166
4.4　分析实例 175

第 5 章　常见的数据挖掘方法 201

5.1　预测 202
5.2　异常检测 225
5.3　探索关系 237
5.4　聚类分析 249
5.5　降维 257

第 6 章　数据分析报告撰写思路 263

后记 272

第1章

什么是淘商？未来如何布局

淘宝电商已经遇到了瓶颈，再也回不到七八年前淘宝流量红利的辉煌时期，并且还要面对日益壮大的社交电商，看着竞争对手们不断地在圈人，淘宝电商们虽然焦急，却束手无策。

01

1.1 什么是淘商

自从有了实体经济，销售的本质就从来都没有变过。淘宝网相当于线下的一栋百货大楼，里面有形形色色的商品，消费者因为有需求，才会到淘宝上购买商品，因此淘宝也只是销售渠道之一。在淘宝开店的商家，被称为淘商或平台电商。对淘商而言，最重要的莫过于流量，因为有了流量才能有销量，而流量的产生是由消费者的需求而产生的。淘商之战，本质上讲就是流量之战，比的是谁的流量多，谁的转化率高。

淘商的玩法有很多种，大商家玩资源，中小商家玩操作技术（SEO，直通车），也有商家玩价格战，还有瞎玩的、不会玩的。中小商家玩的操作技术是根据淘宝网的一些规则和商家自己通过观测得出的一些规律，形成的一套操作方案。在价格战中，虽然大家都说自己的价格是最低的，但不是人人都可以玩得起的，价格战是资本游戏。

很多新手会在不清楚状况的情况下贸然进入淘宝，并且会犯同一个致命的错误——忽略消费者的需求。在消费者不需要这个商品时去销售这个商品，完全忽略了商的本质。商的本质是供需关系，即一方产生需求，一方供给需求。

淘商和其他平台一样并没有本质上的区别，淘商群体是电商群体中具有代表性的一个群体。

1.2 未来如何布局

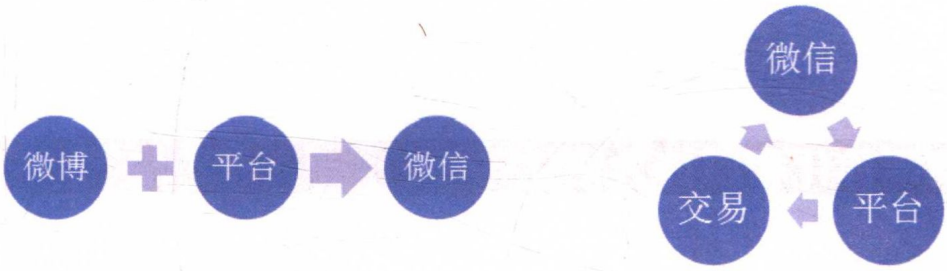
今天的淘商在淘宝网上浴血奋战，他们很痛，非常痛。大部分淘商没有流量，而有流量的淘商也遇到了天花板。这个现象一方面和淘宝内部环境竞争加剧有关，在淘宝内部分行业的商家已经饱和，而需求量就那么多，这就导致僧多粥少；另一方面也和淘宝在电商行业的优势变化有关，淘宝也面临着多方的压力，竞争对手的崛起等都让淘宝在行业内的地位发生了微妙的变化。

淘宝最大的竞争对手不是京东、唯品会等电商平台，因为平台无法干掉平台，而是社交电商。从支付宝仿照微信加入了社交元素的事件可以看出，淘宝内部是急躁的，因为社交电商一旦形成气候，淘宝就会陷入被动。

平台电商和社交电商是有本质区别的，平台电商的运营思路是运营货品，社交电商的运营思路是运营人群和内容。这两种运营思路看似矛盾，但不冲突。社交电商是未来的趋势，因为社交比平台更具有信任基础，而且富有人情味。社交就是内容，我们看微商们就可以知道，他

需要全本请访问 www.ertongbook.com

们拼命地在朋友圈中生产内容。2017 年淘宝也在淘宝内开展了轰轰烈烈的内容运营改革，从手机淘宝 APP 的改版到商家教育运动，都标志着电商已经进入以内容为导向的时代。所以，淘商未来的布局应该是如下图所示的样子。



1. 品牌窗口：微博

微博的定位是一个窗口，只要用户高兴就可以来这个窗口看一看你，不高兴就可以不理你，这就是社交里面的弱关系。在微博中，可以通过举行抽奖活动，找“大 V”转发微博等各种手段去吸引大量的粉丝，平时要多发段子维持微博的人气，随后再通过营销把喜欢你的用户进行转化（通过平台或者非平台成交）。

2. 交易承载页：淘宝、微店等具有交易功能的平台

如果你暂时没有粉丝，那么可以通过淘宝找到你的粉丝。或者说哪里有流量，哪里就可能有你的粉丝。如今，淘宝中的流量虽然没有以前大了，但它还是目前国内最大的电商平台，有流量不用，天理不容！所以淘商们都会想尽办法在淘宝中获取流量，但很多人都忽略了一个问题：当消费者在你这里下单后，你们之间就建立了初步的联系，如果在交易后消费者体验不好，那么他就不会再信任你，这等于让刚建立的关系破裂了。相反，如果在交易后消费者体验很爽，那么他就会信任你，你们彼此之间的关系就变成强关系了。好的体验，必须建立在好的产品基础上。现在电商平台如雨后春笋，别看它们小，但淘宝要面对这么多元素丰富的平台还是很累的。

3. CRM 端口：个人微信

微信最开始的定位就是强关系，因此微信特别适合做 CRM（客户关系管理）。微信也是整个系统的终点，所有的流量入口都是为了微信这个端口，最终的形态是用一部手机和一个微信号，便可赚钱。从长远来看，要让每个顾客都会在你这里重复消费不同品类的产品，前提就是每次交易你都能让他感到爽。

我不是危言耸听，以后的电子商务绝对是靠信任关系才可以出单的，平台电商会生存下来，但是内部必定经过改革和洗牌。我已经看不到中小淘商在淘宝的未来，甚至大商家的未来也是阴晴不定的。哪怕是现在如日中天的淘宝红人，也必须要能掌握客户，客户在哪里，商家就要到哪里。所以，淘商们把微信做起来吧，现在的微商只是时代进步的中间产物，未来，我们都要游走在电商和微信之间，直至有另一个产品替代了微信。

1.3 电商媒体化和媒体人格化

1.2 节提到电商的未来布局，是内容的布局，电商媒体化和媒体人格化就是内容的一种体现形式。电商媒体化的目的是把品牌（自己的）价值传递给消费者。

我们先了解一下媒体是什么？媒体是传递信息的实体，因此每个人都能成为媒体。电商其实也是一种媒体，但不一样的是，电商以商店或者品牌的身份作为媒体，对消费者来讲，其和电商媒体的心理距离比较远。所以玩电商媒体化的商家就无须再建立一个如此遥远的媒体，这个时候需要建立的是一个和消费者心理距离很近的媒体，把媒体人格化，具象化，如下图所示。



品牌即是媒体，媒体即是人，一个活生生的人。通过提炼品牌的特征，可以虚拟化出来一个和品牌特征相符的个人，用这个虚拟出来的人作为传播媒体。例如杜蕾斯的媒体，网友亲昵地称之为“杜杜”。记住了杜杜，就是记住了杜蕾斯这个品牌。这个过程就是以内容为承载，而内容可以投放在任何地方。

第2章

让店铺活下来

做淘宝就是做数据！活下来才是硬道理，2015年电商的关键词就是“生存”，存者为王。在淘宝上要生存有很多种方法，本章介绍作者了解的一些让店铺活下来的有效玩法，读者了解即可。

2.1 小卖家的生存之道

02

一切皆流量。我们在电商领域里做的任何一个动作，背后都是为了流量，所有的连接点也都是流量。

淘宝的搜索流量有一套完整的排序体系，就像我们念书时的试卷一样，例如满分是 100 分，里面会有几十道题目，而且每道题目的分值不同，单道题目答对得满分，答对一半也会得到相应的分值。这里面每道题目就像搜索排序的一个因子，而每道题目的分值就像因子的权重。权重分值计算的概念和答题是一样的。之前经常有人问我，直通车推广费“烧”得越多是不是权重就越大，其实直通车就是试卷里面的一道题目而已，只要你答对了就能得到满分，不会因为你多写几个字，就额外给你分值。

小卖家缺的就是思维和执行力。有思维和执行力的卖家不会是小卖家。这里不是要长篇大论和读者论证这么一个事实，因为这已经是一个事实。

小卖家的玩法可以分成下面几种，因为这些玩法具有时效性，读者在阅读的时候可能这些方法已经失效。玩法终究是技巧性的东西，**小卖家找到合适的产品，自行拓展流量渠道，同时兼顾平台流量，才是最佳的状态。**

（1）上新玩法

玩法定位：这个玩法适合新手，属于纯粹练手的玩法，店铺可能会因为这次练手而被废掉。

适合类目：服饰鞋包等非标品类目。

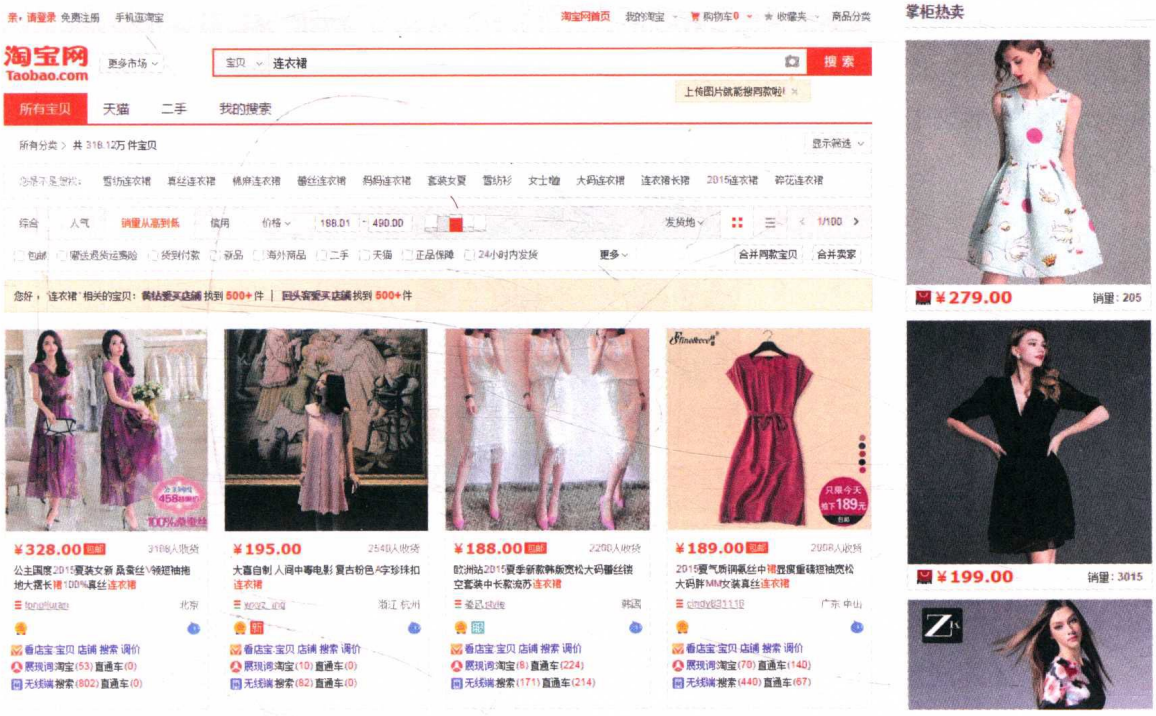
玩法原理：① 利用市场证明过的热销产品；② 利用上新率、动销率权重和新品的测试流量；③ 利用非标品的定价差；④ 配合重复铺货同个产品铺多个链接。

操作流程：

第一步，在淘宝中通过搜索、直通车[1]或者淘宝排行榜[2]找到销量高，并且价格在中档偏上的产品，如下图所示。

1 直通车是淘宝网的付费推广方式之一，可以直接通过关键词竞价出现在搜索页面的右侧

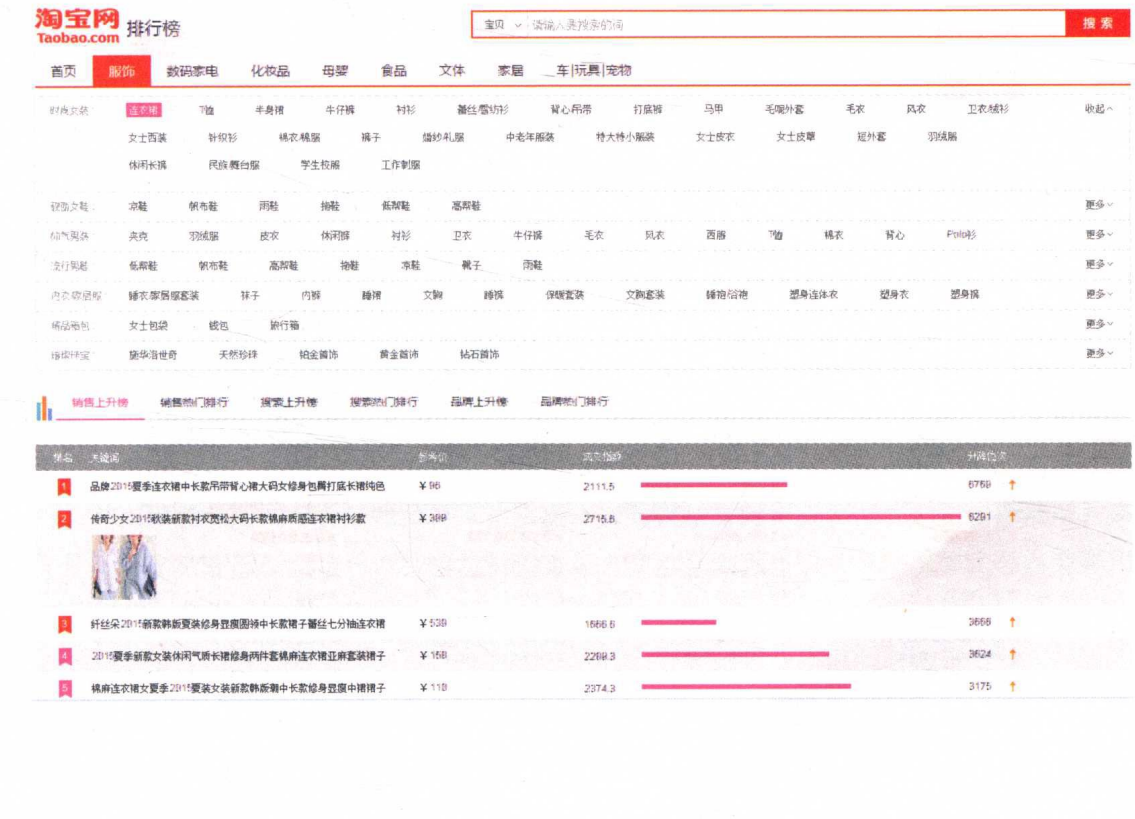
2 淘宝排行榜是淘宝网的热销和暴涨商品的排行榜，网址为 http://top.taobao.com


淘宝网
Taobao.com
搜索
连衣裙
所有宝贝
天猫
二手
我的搜索
¥328.00
¥195.00
¥188.00
¥189.00

掌柜热卖

¥279.00
销量：205

¥199.00
销量：3015


淘宝网 排行榜
Taobao.com
搜索
首页
服饰
数码家电
化妆品
母婴
食品
文体
家居
车|玩具|宝物
销售上升榜
销售热门排行
搜索上升榜
搜索热门排行
品牌上升榜
品牌热门排行

02

第二步，在淘宝搜索同款产品，找到产品价格低，并确认有货的卖家，这一步是为了之后我们卖出产品时，可以在淘宝的其他店铺中拍下产品，让其他卖家代发货，如下图所示。





需要全本请访问 www.ertongbook.com