

Python Word Cloud Analysis of Consumers' Attitudes towards Different Brands of Skin Care Products

Xinhong Lin, Enchi Liu, Wendan Liu, Huqing Yan

Xiamen National Accounting Institute, Xiamen Fujian
Email: 870527069@qq.com

Received: Mar. 21st, 2020; accepted: Apr. 2nd, 2020; published: Apr. 9th, 2020

Abstract

In this paper, the BeautifulSoup web crawler tool, Wordcloud word cloud analysis tool and Jieba Chinese character processing tool provided by Python language are used to collect consumer evaluation component data from Tmall's flagship stores under Taobao's official website, and six famous skin care brands are selected. After preprocessing the collected data, high-frequency words are counted and word cloud maps are made. This paper analyzes consumers' views on skin care brands in different countries and explores the future development path of skin care brands in China.

Keywords

Skincare, Brand, Consumer Market

消费者对不同品牌护肤品态度的Python词云分析

林鑫洪, 刘恩赤, 刘文丹, 阎虎勤

厦门国家会计学院, 福建 厦门
Email: 870527069@qq.com

收稿日期: 2020年3月21日; 录用日期: 2020年4月2日; 发布日期: 2020年4月9日

摘要

本文采用Python语言所提供的BeautifulSoup网页爬虫工具、Wordcloud词云分析工具、以及Jieba汉字

处理工具,从淘宝官网下的天猫各大品牌旗舰店网站中采集消费者评价组成数据,选取了六大知名护肤品品牌,对采集的数据通过预处理,统计出高频词汇,制作词云图,据以分析消费者对不同国家护肤品品牌的看法,探索中国护肤品品牌的未来发展道路。

关键词

护肤品, 品牌, 消费市场

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

改革开放以来,我国逐渐进入一个稳步、可持续发展的社会主义新时代,各行各业发展迅速,与各行各业的发展类似,我国化妆品行业也进入高速成长阶段,经过 20 年的发展,我国的化妆品市场已跃居世界第二,排名仅次于美国。日益庞大的中国化妆品消费市场,吸引着全球化妆品行业的眼光,国际知名品牌纷踏而至,给我国化妆品行业带来巨大的冲击。从《中国化妆品》杂志社和赛迪顾问股份有限公司联合发布的《2019 年第三季度中国化妆品市场盘点数据》(图 1)可以看出,护肤品占据化妆品行业的第一大品类,高达 54% 占据着最大的市场份额[1]。

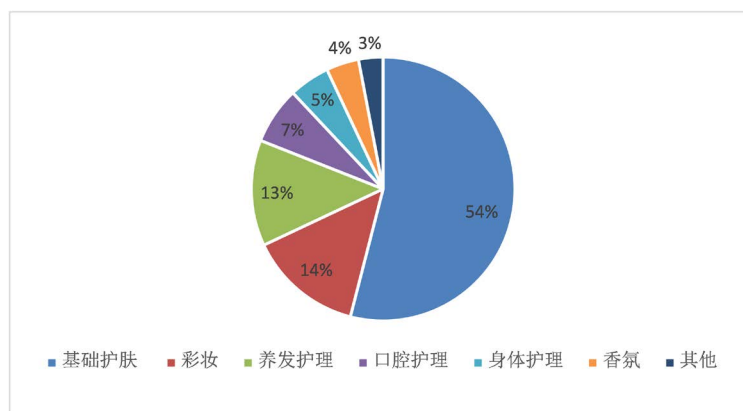


Figure 1. China cosmetics segment market structure in the third quarter of 2019
图 1. 2019 年第三季度中国化妆品细分市场结构

2018 年将我国护肤品行业市场份额中前 10 位企业的行业市场占有率加起来,尚不及护肤品行业总占有率的 50%。市场前十名中,国产护肤品企业仅有三家[2]。同样 2019 年天猫双十一护肤品行业 TOP10 销售额排行榜中,国产品牌百雀羚、自然堂、薇诺娜占据了三席。针对这些数据,不少认为中国国产护肤品品牌在与外资品牌的竞争中节节败退,难以有美好的未来。也有人认为中国民族护肤品品牌正是在跟外资护肤品品牌的交锋中不断成长和壮大[3]。但唯一可以肯定是,国货美妆品牌仍需继续成长。

同时随着全球化进程的加快和产品技术差异的缩小,导致同类商品间的差异越来越小,品牌竞争也随之加剧。在这样的基础上,企业不断通过各种沟通手段试图达到预期的品牌状态,使得品牌在市场中反映出独特的定位和主张。品牌识别同时关注了产品属性和品牌的象征意义以及品牌的价值主张[4]。

本文分析不同国家的护肤品品牌在我国现状，利用 Python 语言支持下的 BeautifulSoup 网页爬虫、Wordcloud 词云分析、Jieba 汉字处理等工具，分析了消费者对不同国家护肤品品牌的看法，探索中国护肤品品牌在未来的发展道路。

2. 数据获取及数据预处理

从淘宝官网下的天猫各大品牌旗舰店网站中采集所需的消费者评价词条组成数据，选取了六大知名护肤品品牌，调查对象分别为国产品牌(表 1): 自然堂、麦吉丽; 日韩品牌(表 2): 珂润、Sk2; 欧美品牌(表 3): 娇兰、兰蔻。对采集的数据用 anaconda 中的 matplotlib 进行预处理分析, 统计出词频, 运用 wordcloud 依据消费者使用评价的词频制作词云图, 让统计结果更加直观, 便于寻找主要属性特征。本文一共采集了 286,783 条统计数据, 其中, 珂润品牌选取了 19,739 条累计评价, Sk2 品牌选取了 121,072 条累计评价, 自然堂品牌选取了 97,136 条累计评论, 麦吉丽品牌选取了 12,195 条累计评论。娇兰选取了 6115 条累计评价, 兰蔻选取了 30,526 条累计评价

Table 1. Chinese lotion and cream evaluation words frequency

表 1. 国产水乳评价词频

词序	国产			
	自然堂		麦吉丽	
	词	频数	词	频数
1	保湿滋润	852	包装挺好	212
2	控油	247	保湿滋润	116
3	味道好闻	245	正品	66
4	吸收效果不错	202	吸收效果不错	40
5	是正品	196	清洁度强	28
6	清洁度强	193	提亮肤色	27
7	味道不好闻	67	不适合敏感肌	13

Table 2. Japanese and Korean lotion and cream evaluation word frequency

表 2. 日韩水乳评价词频

词序	日韩			
	珂润		Sk2	
	词	频数	词	频数
1	保湿滋润	778	包装挺好	2313
2	控油	210	保湿滋润	558
3	吸收效果不错	197	吸收效果不错	352
4	物流快	165	是正品	330
5	清洁度强	161	控油	230
6	温和不刺激	152	清洁度强	181
7	控油效果一般	69	味道不好闻	216

Table 3. European and American lotion and cream evaluation word frequency
表 3. 欧美水乳评价词频

词序	欧美			
	娇兰		兰蔻	
	词	频数	词	频数
1	保湿滋润	1570	物流快	11,595
2	包装不错	1460	保湿滋润	10,022
3	味道好闻	1045	吸收效果不错	3981
4	吸收效果不错	976	味道好闻	2660
5	控油	744	控油	1775
6	清洁强度高	243	不保湿滋润	493
7	控油效果一般	77		

3. 总结热点词及情感分析

3.1. 国产品牌水乳

3.1.1. 自然堂(平价)

基于词频统计结果,分析词频表(图 2),其中“保湿滋润”、“控油”、“吸收效果不错”、“清洁度强”都代表了产品功能和实用性,合计占到总词频数的 74.6%,可见在选择国产平价产品时,产品功效仍然是消费者最为关心的因素。其次,产品的味道也在很大程度上影响了消费者的评价,说明在产品消费者对于产品使用的外在体验也有一定的要求。



Figure 2. Chando's lotion and cream evaluation word cloud illustration
图 2. 自然堂水乳评价词云图

3.1.2. 麦吉丽水乳(高端)

基于词频统计结果,分析词频表(图 3),其中排名第一的“包装挺好”的词频数占到总词频的 42.2%，“保湿滋润”、“吸收效果不错”、“清洁度强”、“提亮肤色”、“不适合敏感肌”等代表产品功能的次频数占到 44.6%,可见消费者对高端品牌产品外在的直观感受几乎与产品功能的关注度持平,这是与国产平价品牌效用优先最为不同之处。在平价和高端国货品牌的关键词的筛选中,“正品”都占到了一定的比例,说明消费者对于护肤品安全质量问题仍然非常关心,尤其是在麦吉丽所代表的高端品牌中,“正品”的频数更是排到第三,这可能与高端产品的售价高、盈利空间大,容易被仿制有关。



Figure 3. Mageline's lotion and cream evaluation word cloud illustration
图 3. 麦吉丽水乳评价词云图

3.2. 日韩品牌水乳

3.2.1. 珂润水乳(平价)

基于词频统计结果，分析词频表(图 4)，排名前三的关键词为此护肤品的实际效用，其中，“保湿滋润”的词频出现频率最高，为 778 条，占据总词频数的 44.92%。从中可以看出，消费者对日韩品牌中较为平价的护肤用品，最为重视的因素是产品的实际性能，即一分产品一分货，在不需要过多考虑的情况下更为重视产品的性价比，产品使用效果会导致消费者对这产品的评价有一定的负面情绪。其次，以商家服务质量为代表的词频“物流快”，也极大地影响了消费者的对该产品的评价，在产品的实际效用能够达到消费者预期的情况下，商家服务的质量也会影响消费者的满意度。



Figure 4. Curel's lotion and cream evaluation word cloud illustration
图 4. 珂润水乳评价词云图

3.2.2. Sk2 水乳(高端)

基于词频统计结果，分析词频表(图 5)，排名第一的关键词为“包装挺好”，出现频率为 2313 条，占据总词频数的 55.33%，其次才是产品的实际使用效果，与上文较为平价的护肤产品对比可以看出，消费者对日韩品牌中较为高端的护肤用品，关注的不仅有产品效用，还有产品外表所带来的直观感受。由此可以看出，高端品牌所带来的品牌效应，高端护肤品的品牌形象极大影响了消费者的品牌态度，也许这种影响力量能够决定消费者是不是喜欢某一个品牌，是不是会对这一品牌生出购买意向[5]。观察词频统计我们仍然可以发现，排名第四的词频“正品”也与前文分析的平价产品有所区别，在护肤品市场中，由于高端产品的售价高、盈利空间大，容易被不良供应商仿制，需要正品商家在积累产品口碑的同时，严厉提防假冒伪劣商品给消费者带来的损失。在该产品的负面评价中，词频数最高的评价是体验感上的

“味道不好闻”，对比平价护肤品牌负面评价最多的是产品的实际效用，可以看出高端护肤品带给高端消费者的用户体验是评价的最主要影响因素。



Figure 5. Sk2's lotion and cream evaluation word cloud illustration
图 5. Sk2 水乳评价词云图

3.3. 欧美品牌

3.3.1. 娇兰

基于词频统计结果，分析词频表(图 6)，排名第一的关键词为“保湿滋润”，出现频率为 1570 条，占据总词频数的 25.67%，其次是产品的包装不错出现频率为 1460，占了总词频的 23.88%。娇兰的词频体现了欧美护肤品牌的一大特征—产品具有较好的效果，欧美国家本身对时尚有着强烈的愿望，雄厚的经济实力及发达的科学技术，加上化工和现代工业的辅助，各种护肤品层出不穷，以大众为发展基础，加之纯天然配方，良好的使用效果使之很快席卷全球。同时，欧美品牌也善于宣传，在这个追求创意的时代，质感上乘，特点显著，富有创造性的包装是吸引顾客的法宝[6]，从娇兰的词频可以看出该产品的精美包装也为其吸引了一大批顾客。

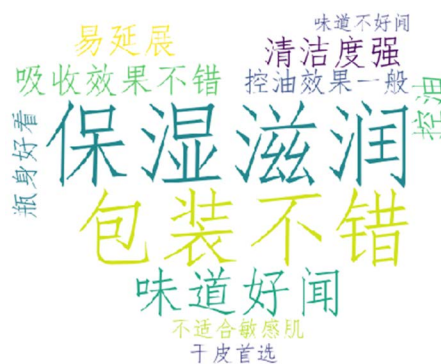


Figure 6. Guerlain's lotion and cream evaluation word cloud illustration
图 6. 娇兰水乳评价词云图

3.3.2. 兰蔻

基于词频统计结果，分析词频表(图 7)，排名第一的关键词为“物流快”，出现频率为 11,595 条，占据总词频数的 38%，其次是“保湿滋润”出现频率为 10,022，占了总词频的 32%，从中可以发现，兰蔻不仅具有欧美品牌的一般特征，还与我国市场完美结合，利用我国发达的运输服务，给消费者带来更好的体验。

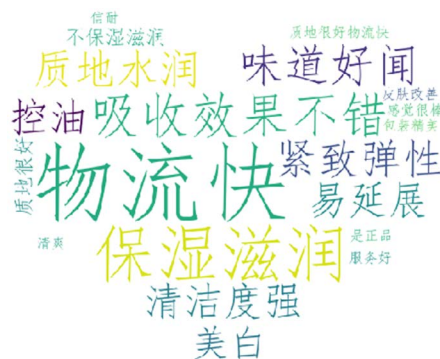


Figure 7. Lancome's lotion and cream evaluation word cloud illustration

图 7. 兰蔻水乳评价词云图

3.4. 分析结果

综合分析上述 6 个不同品牌的护肤品, 我们不难发现, 由于不同价位的护肤品的消费群体有所不同, 消费者对产品关注的侧重点也有所差异。消费者在使用平价产品时更为关注产品的实际效用, 在使用高端产品时, 消费者则把侧重点向产品外在包装带来的直观感受转移。针对消费者的评价, 不同品牌的商家可以从中获得产品改进的参考。平价护肤品应更加注重产品实际效用的提升和服务质量的提升, 适当利用广告及价格策略不断为产品积累口碑; 高端护肤品应提升品牌形象, 谨慎防范假冒伪劣产品, 注重消费者体验感的提升。

4. 总结

本文利用 Python 语言支持下的 BeautifulSoup 网页爬虫、Wordcloud 词云分析、Jieba 汉字处理等工具, 分析了消费者对不同国家护肤品品牌的看法。

基于对国货平价和高端品牌的分析, 消费者在选择平价产品时最为关注产品性能, 其次是外在的使用体验感。在选择高端产品时, 消费者同样关注产品性能和外在包装的直观感受。由于我国美妆品牌中走高端路线品牌较少, 因此在分析消费者对于国货品牌的选择因素时, 平价品牌可能会更具有代表性。据《中国国货化妆品消费者行为影响因素分析》文中所列的问卷调查, 结果显示有 69.5%、46.32% 和 53% 的调查人群认为最重要的因素为护肤品的质量、成分和效果[7]。腾讯发布的 2019 年《国货美妆洞察报告》显示, 在选择美妆国货主要原因中, 84% 消费者认为产品功能是最为重要的[8]。《我国护肤品行业现状与消费者需求分析》中间卷调查同样表明, 在选择护肤品时, 90.86% 的消费者认为产品功效使他们所需要考虑的因素[9]。这些结论也与本文通过词频分析得到的结论基本一致。

基于日韩护肤品牌的分析, 在价格、形象、实际效用不同的情况下, 消费者的品牌态度有所不同。平价产品的实际效用和商家服务态度, 影响了消费者对这一品牌产品的评分, 消费者选择平价护肤品, 关注度最高的是其实际效用, 其次是物流的快慢; 高端产品由于已经积累了一定程度的口碑, 产品的形象、产品是否为正品对消费者对品牌的感知产生重大的影响。

基于欧美护肤品牌的分析, 我们不难发现, 随着欧美文化近几年风靡全国, instagram、facebook、youtube 等平台上的明星博主已成为当下的时尚风向。越来越多的中国女性开始追捧欧美时尚。目前越来越多的欧美护肤品品牌入驻中国, 大多数欧美护肤品品牌都经过时间和历史积淀而成, 其独特的个性和风格吸引世界各地的人们共同向往, 共同消费, 同时其较好的产品使用效果也为其带来了大批的忠诚消费者。

我国护肤品行业正处于快速发展的黄金时期, 国民对护肤品的采购数量和消费额的不断增加, 明确

地表明了国民对护肤品的强烈需求和愿望。但国产护肤品企业仍需继续成长。结合本文所做的词频分析，提出以下几点建议：

增加产品研发投入，提升国产护肤品品质。无论是对平价还是高端品牌，产品功能与实用性，都是消费者极为关心的考虑因素。国产护肤品品牌应以满足顾客所需作为企业战略选择的方向，通过研究消费者需求的变化来更好地研发适销对路的产品。优质的产品就是企业的核心竞争力，若是能打造出爆品，可以更好带动品牌的整体发展。

优化产品设计包装，提高包装的营销作用。从词频分析中，我们得知很大部分消费者都很注重产品包装所带来的外在体验感。实际上，包装除了保护功能作用之外，还具有促进商品销售的功能。通过关注包装的结构设计和颜色使用等方面与客户之间的紧密联系，包装也可以表达产品的内部文化，触及顾客的内心深处，并利用客户的情感心理来提升品牌价值[10]。

重视人才培养或引进，壮大企业人才队伍。虽然国产护肤品行业有了一定的实力和基础，但是在产品研发、市场营销、供应链等方面仍缺乏优秀的人才。因此企业要根据自身的需求培养或引进优秀的人才去弥补自身的短板。同时要为人才提供的良好企业的工作环境和氛围，以便其愿意长期留在企业跟品牌共同发展。

基金项目

本论文得到了厦门国家会计学院 2019 年“云顶课题：python 财务数据分析项目的支持”。

参考文献

- [1] 中国护肤. 2019 年第三季度中国护肤品市场盘点[J]. 中国护肤品, 2020(1): 64-71.
- [2] 陈燮达. 我国护肤品行业现状分析及战略思路[J]. 中外企业家, 2020(6): 111-112.
- [3] 龚述辉. 中国民族护肤品品牌的发展策略浅析[J]. 日用化学品科学, 2020(1): 26-27.
- [4] 王红侠. 消费者护肤品品牌识别实证研究[D]: [硕士学位论文]. 西安: 西安理工大学, 2007.
- [5] 王玥. 品牌来源国效应下广告诉求方式及价格因素对护肤品消费者品牌态度的影响分析[D]: [硕士学位论文]. 上海: 华东师范大学, 2015.
- [6] 蒋帆. 中国护肤品市场的贸易分析——以欧美护肤品为例[J]. 商场现代化, 2018(15): 97-99.
- [7] 杨蕾. 中国国产护肤品消费者行为影响因素分析[D]: [硕士学位论文]. 青岛: 中国海洋大学, 2015.
- [8] 腾讯, 凯度. 2019 年国货美妆洞察报告[R]. 2019.
- [9] 杨雨鑫. 我国护肤品行业现状与消费者需求分析[J]. 价值工程, 2019(17): 258-260, 10.
- [10] 陈金晶, 李艳花. 浅析中国传统设计元素在现代包装中的应用[J]. 福建茶叶, 2020(2): 157-158.