

就你话多

——今日头条用户评论数据报告（下）



今日头条算数中心



研究目的

- › 本报告分析今日头条过去一年的用户阅读大数据，研究在移动阅读环境下，什么样的资讯更易引发评论，为创作者提供参考。本报告主要通过评论率这一指标来衡量资讯引发评论的难易程度。



样本说明

- › 本报告选取的样本是2016.7.15-2017.7.15期间，今日头条平台上分发的共计66,492,398篇资讯。



核心发现

篇幅

- › 中长篇（1000字以上）资讯比短资讯（1-500字）更易引发评论。

体裁

- › 文章（包括含有插图的文章）比组图和视频更易引发评论。

题材

- › 政务、媒体、移民、游戏、本地、数码、房地产、科技、财经、社会等资讯更易引发评论。

写作方式

- › 严肃+倾向性的资讯评论率最高；严肃+蹭热点的资讯评论率最高。

结尾方式

- › 评论率高的资讯，多以祈使句结尾，或是以问句结尾，或是结尾设置投票。

目录

CONTENTS

01

篇幅

02

体裁

03

题材

04

写作方式

05

结尾方式

01.

PART ONE

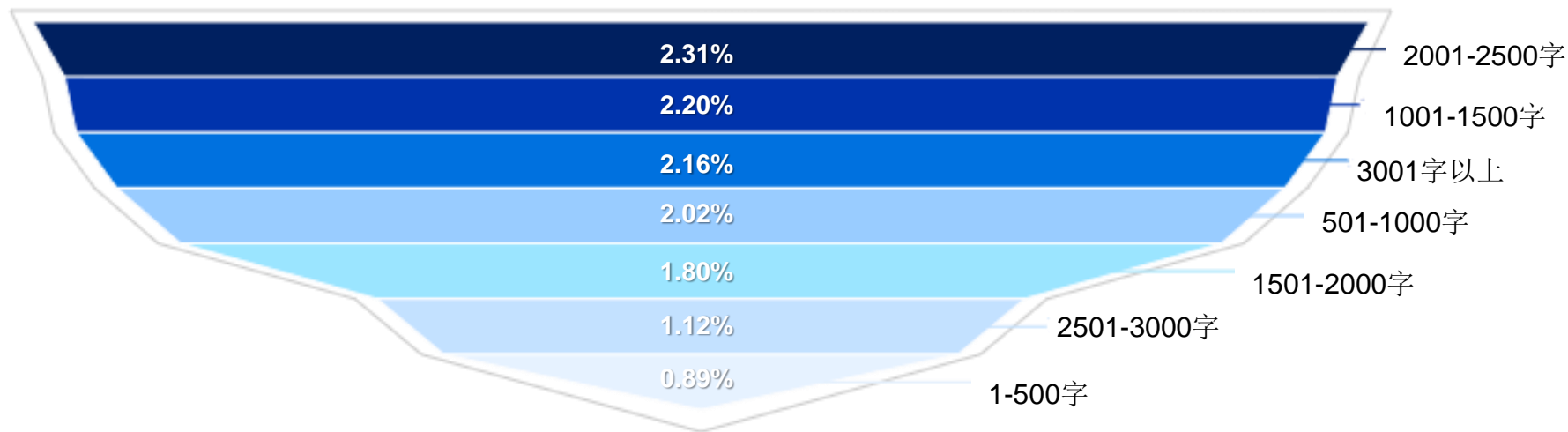
篇幅 Length

我们将移动资讯按照字数进行划分，
每500字划分一档，共分为七档。

表1呈现了在剔除无效样本后，各档
的样本量和总样本量。

篇幅（字）	样本量
1-500	31,290,680
501-1000	11,373,373
1001-1500	4,775,300
1501-2000	2,310,461
2001-2500	1,109,613
2501-3000	733,473
3001以上	927,043
总计	52,519,943

表1



2001-2500字 1001-1500字 3001字以上 501-1000字 1501-2000字 2501-3000字 1-500字

中长篇资讯（1000字以上）比短资讯（1-500字）更易引发评论。

数据来源：头条指数

注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

© 2017.07 今日头条算数中心



“ 即便在移动互联网上，某些规律依然成立——复杂的有价值的事情需要更多的字数描述，中长篇的内容丰富度更高，容纳矛盾的可能性更大，可阐释性更高，更易获得评论。”

今日头条副总编辑 徐一龙

02.

PART TWO

体裁 Genre

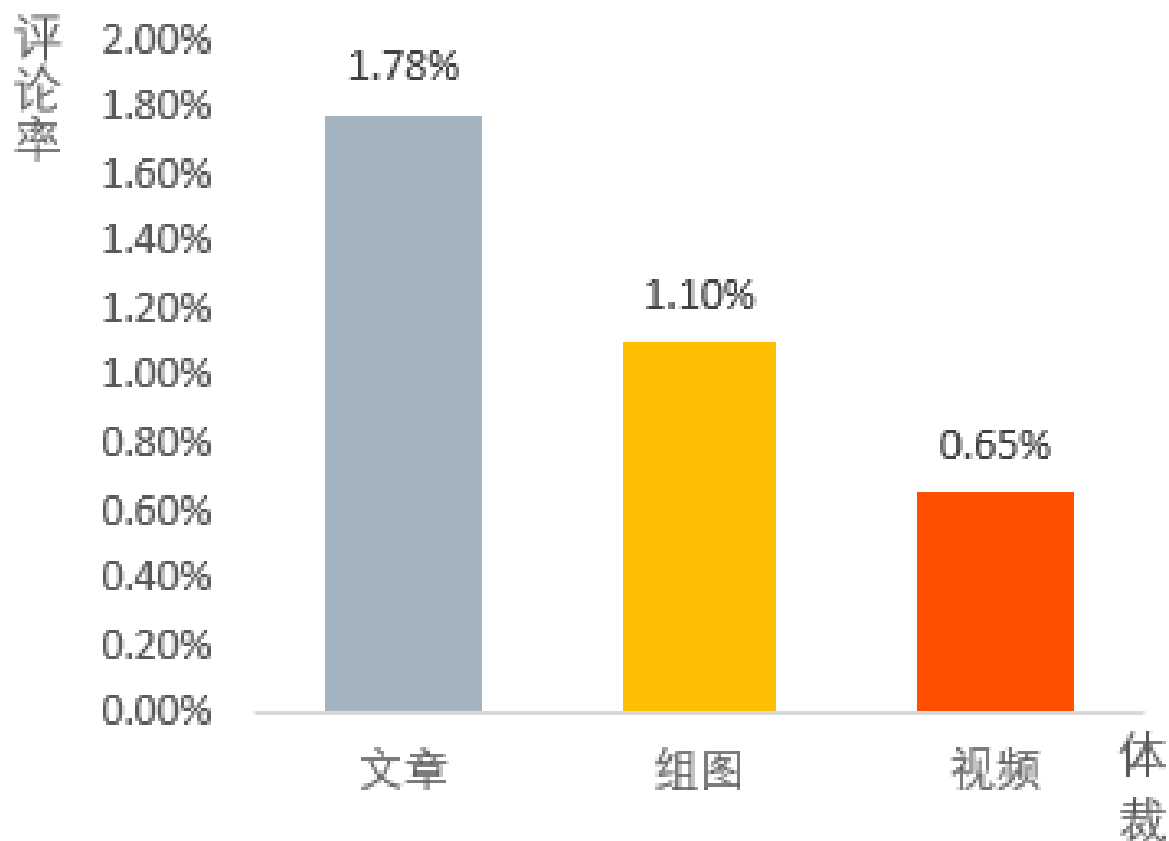
我们将移动资讯按照体裁分为三类：

文章（包括含有插图的文章）、组图和视频。

表2呈现了在剔除无效样本后，三类的样本量和总样本量。

体裁	样本量
文章	48402296
组图	3246246
视频	14843856
总计	66492398

表2



文章（包括含有插图的文章）比组图和视频更易引发评论。

数据来源：头条指数

注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

© 2017.07 今日头条算数中心

03.

PART THREE

题材 Theme

样本说明



我们将移动资讯按照内容题材进行分类，剔除无效样本后，对各类样本进行分析。

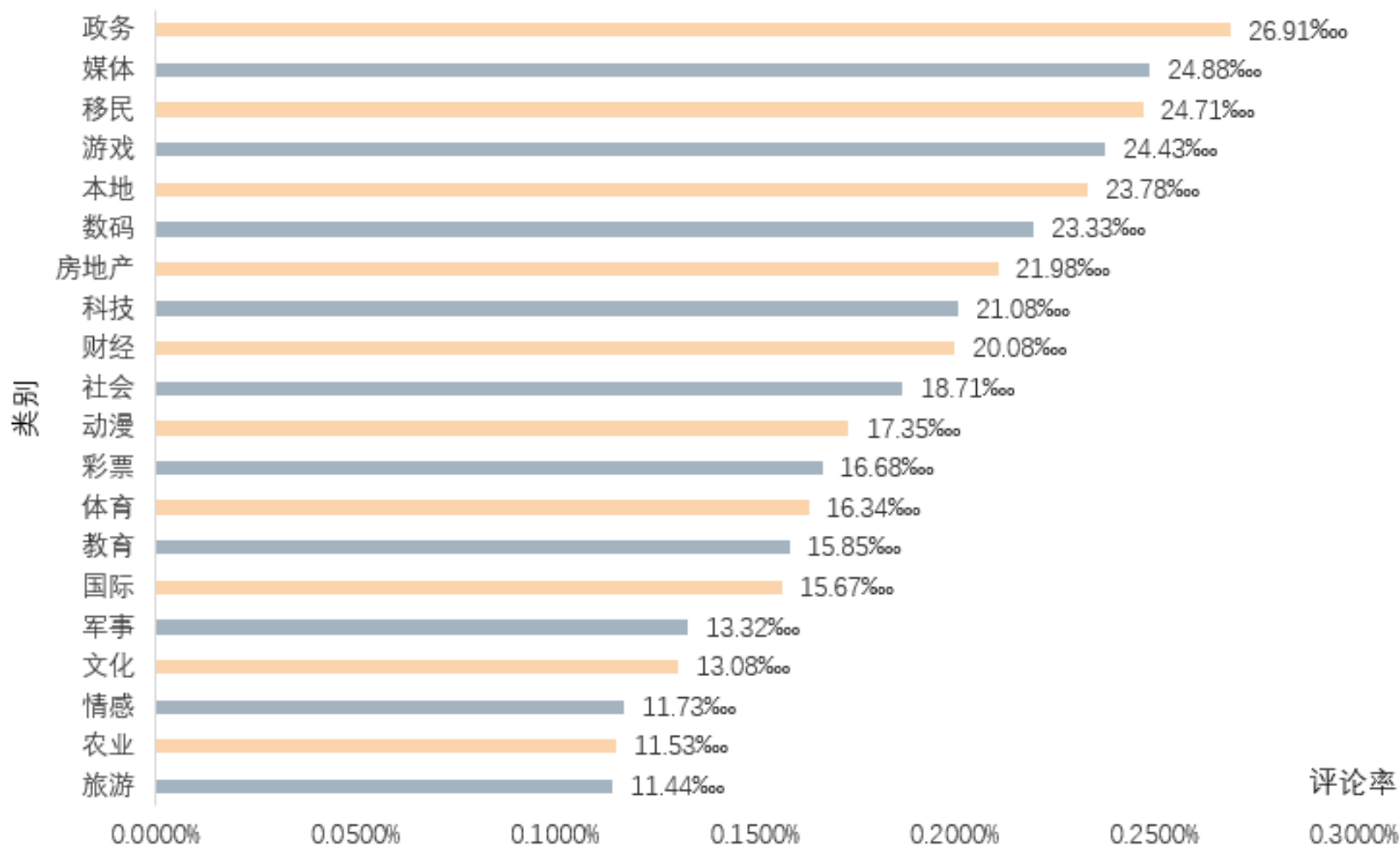
表3仅展示部分类别的样本量。

题材	样本量
农业	18327890
社会	8318385
财经	4249367
体育	2390962
科技	2336720
教育	1982724
旅游	1880833
文化	1633823
情感	1580555
数码	1541048
游戏	1293588
国际	768964
房地产	723718
军事	551284
政务	435635
动漫	432065
本地	396347
彩票	198820
媒体	100787
移民	39781

表
3



不同类别的文章评论率（取Top20）



数据来源：头条指数

注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

此处评论率以‰（万分号）为单位，表示每10000次阅读中产生评论的数量。

© 2017.07 今日头条算数中心

“

丰富多维的内容为今日头条引来了大量明星、名人、达人用户，为头条增加了鲜活、有趣、有个人特色的短内容，引发大量网友评论。



2017年6月，头条用户用「叫错名」的玩法在评论区玩得不亦乐乎。明星名人在评论区也非常活跃，他们不仅彼此互动，还与网友互动，诞生了一位准现象级网红「色厨」李奇斌。”

今日头条UGC运营总监 王一波

04.
PART FOUR

写作方式
Style of Writing

4-1

我们在今日头条平台上，选择了科技类、娱乐类、体育类、财经类、社会类五个细分类别，在剔除无效样本后，按照评论数选择Top300的资讯作为研究对象。

在这部分中，依照资讯特点，我们从严肃/活泼、倾向性/非倾向性两个维度进行写作方式的分析和标注，探讨不同写作方式的资讯评论率的高低。

严肃是指信源真实可靠，用词严谨，在陈述事实时不用或少用夸张、比喻、拟人等修辞手法，不用或少用网络流行语，图片原始真实，没有使用表情包类图片。

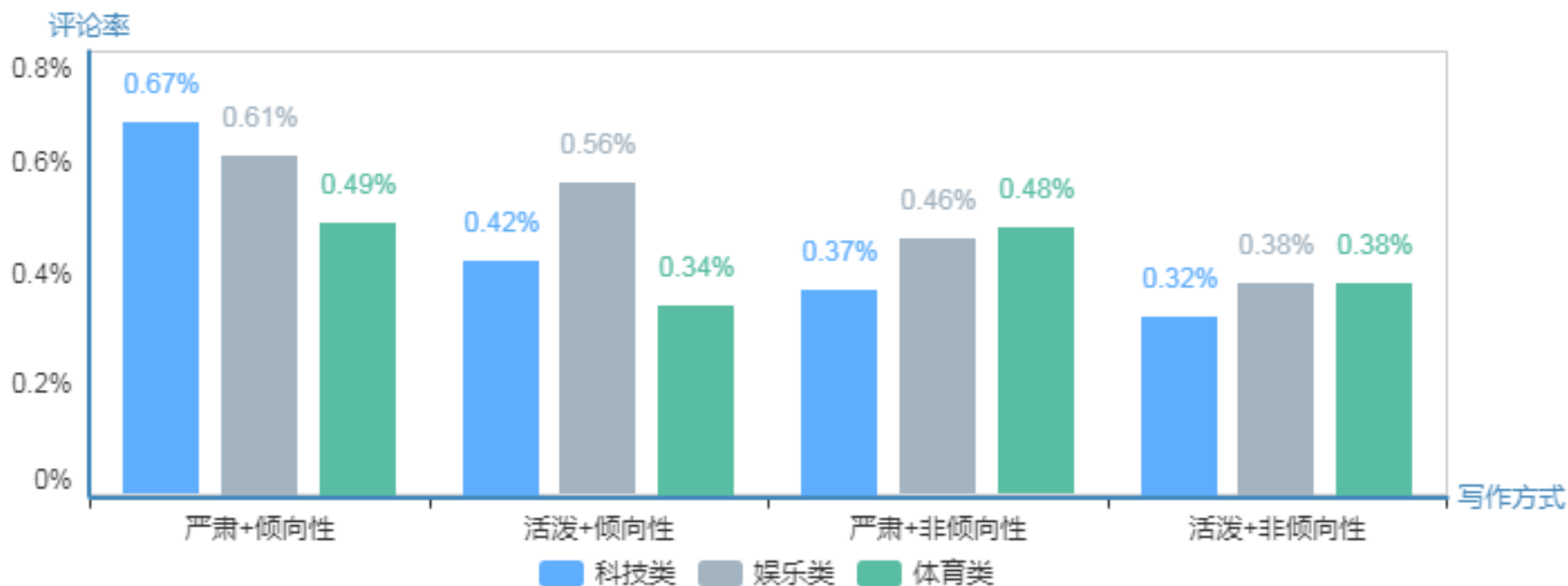
倾向性是指在陈述完事件后对该事件作出了褒贬或赞同反对等主观性评论。

表4是五大类别不同写作方式的样本量。

	严肃+倾向性	活泼+倾向性	严肃+非倾向性	活泼+非倾向性	合计
科技类	69	100	93	38	300
娱乐类	34	112	94	60	300
体育类	92	34	163	11	300
财经类	76	65	150	9	300
社会类	25	14	251	10	300

表4

严肃+倾向性 的资讯评论率最高，如科技类、娱乐类和体育类。



数据来源：头条指数

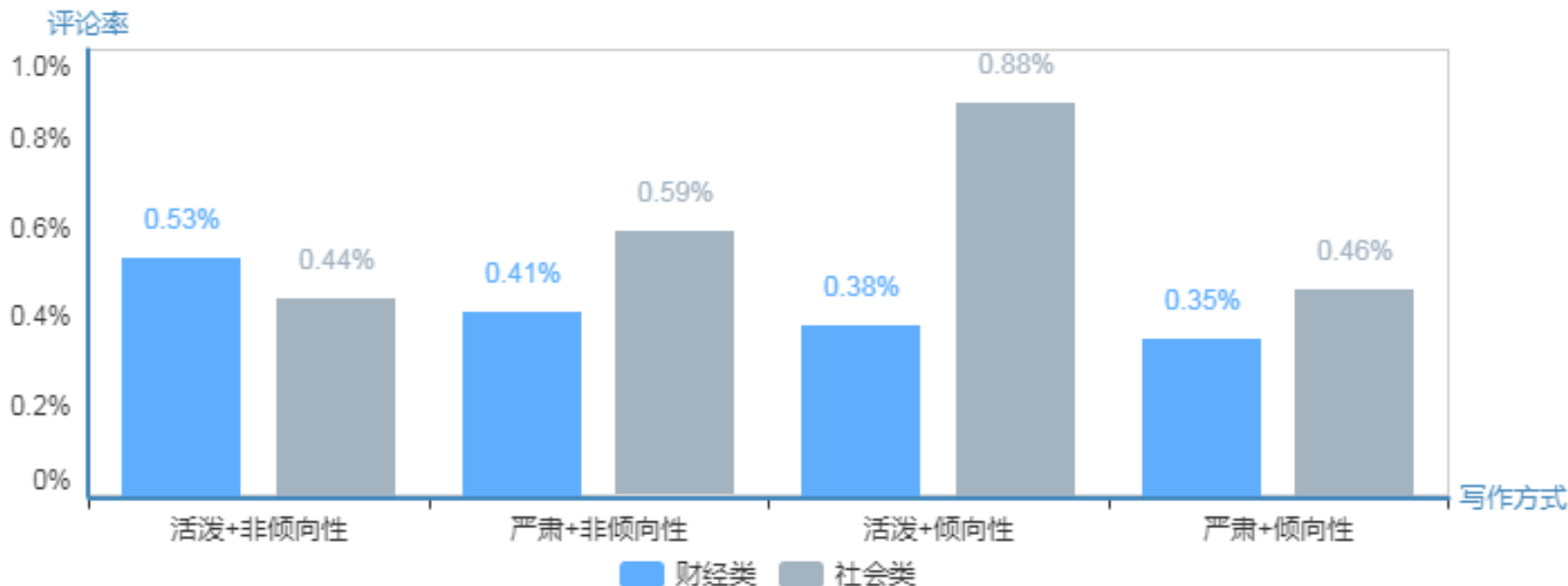
注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

© 2017.07 今日头条算数中心

但是，也存在一些例外的情况。

对于 **财经类** 而言，**活泼+非倾向性** 的资讯评论率最高。

对于 **社会类** 而言，**活泼+倾向性** 的资讯评论率最高。



数据来源：头条指数

注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

© 2017.07 今日头条算数中心

4-2

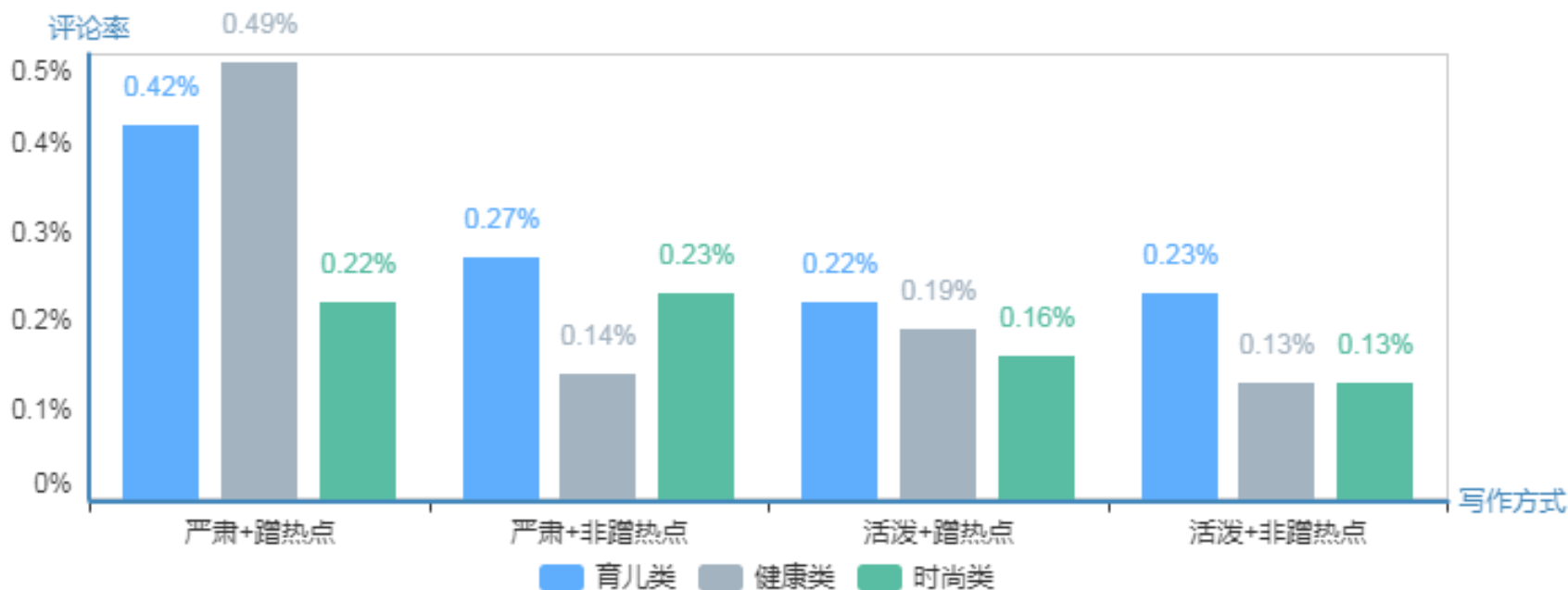
- 我们在今日头条平台上选择了育儿类、健康类、时尚类三个细分类别，在剔除无效样本后，按照评论数选择Top300的资讯作为研究对象。
- 在这一部分，依照资讯特性，我们从严肃/活泼、蹭热点/非蹭热点两个维度进行写作方式的分析和标注，从而探讨不同写作方式的资讯评论率的高低。
- 蹭热点是指使自己的内容与当前热点产生关联，或围绕当前热点创作内容。

表5是三大类别不同写作方式的样本量。

	严肃+蹭热点	活泼+蹭热点	严肃+非蹭热点	活泼+非蹭热点	合计
育儿类	15	6	204	75	300
健康类	21	14	174	91	300
时尚类	106	60	27	107	300

表5

总体而言，**严肃+蹭热点** 的资讯评论率最高，如 **育儿类和健康类**。
但是对于 **时尚类** 而言，**严肃+非蹭热点** 的资讯，评论率最高。



数据来源：头条指数

注：评论率反映某类文章引发评论的难易程度，评论率=评论数/阅读数。

© 2017.07 今日头条算数中心

“今日头条绝不「唯流量是举」、「唯点击率是举」，我们相信用户愿意接受各种好的东西，平台也会继续在推荐机制设计上给严肃的、优质的内容更多资源扶持和倾斜。”

今日头条高级运营总监 吴达



05.
PART FIVE

结尾方式
Ways of Ending

○ 以祈使句结尾



在健康类资讯中，大约每**8**个样本中就有1个以祈使句结尾（**13.00%**）。

文章《这5种症状都在表明：你体内可能有大量蛔虫寄生》的结尾为：如果大家有什么经验和方法，请务必和我们分享，让我们能帮助到更多的人远离蛔虫病！下期见！

对于蛔虫病，主要还是以预防为主。比如干净饮食，不在外面喝生水、吃直接用生水洗的食物，勤洗手等。在家，大家也可以多喝一点醋或者吃葱白、藿香正气水等，都可以一定程度上抑制蛔虫。

如果大家有什么经验和方法，请务必和我们分享，让我们能帮助到更多的人远离蛔虫病！下期见！

11,970 条评论



写下您的评论...

评论

以问句结尾



在育儿类资讯中，大约每**7**个样本中就有1个以问句结尾（14.33%）。

文章《你站哪边？2岁小孩被4岁熊娃从滑梯推落 孩爸怒踹熊娃》的结尾为：这些事儿，你怎么看？

草頭酥餅：父亲踢小孩肯定不对，再怎样也不能对4岁孩子动手；但这位母亲有反省自责过自家孩子做错事儿了吗？双方监护人都有问题吧，餐厅员工挺好的，首先护住孩子。

这些事儿，你怎么看？

13,050 条评论



写下您的评论...

评论



结尾设置投票



在科技类资讯中，大约每**37**个样本中就有1个结尾设置投票（2.67%）。

文章《华为荣耀系统停更，请愿无效！花粉反目》的结尾为：

“荣耀6系统停更你怎么看？(单选)”

- 支持！配置跟不上就不更
- 不支持！老用户也需要关爱
- 打酱油

对于此事你怎么看？

荣耀6系统停更你怎么看？(单选)

- ☐ 支持！配置跟不上就不更
- ☐ 不支持！老用户也需要关爱
- ☐ 打酱油

投票

3,810 条评论



写下您的评论...

评论

结语

Conclusion

人们更多地参与到创作与交流中

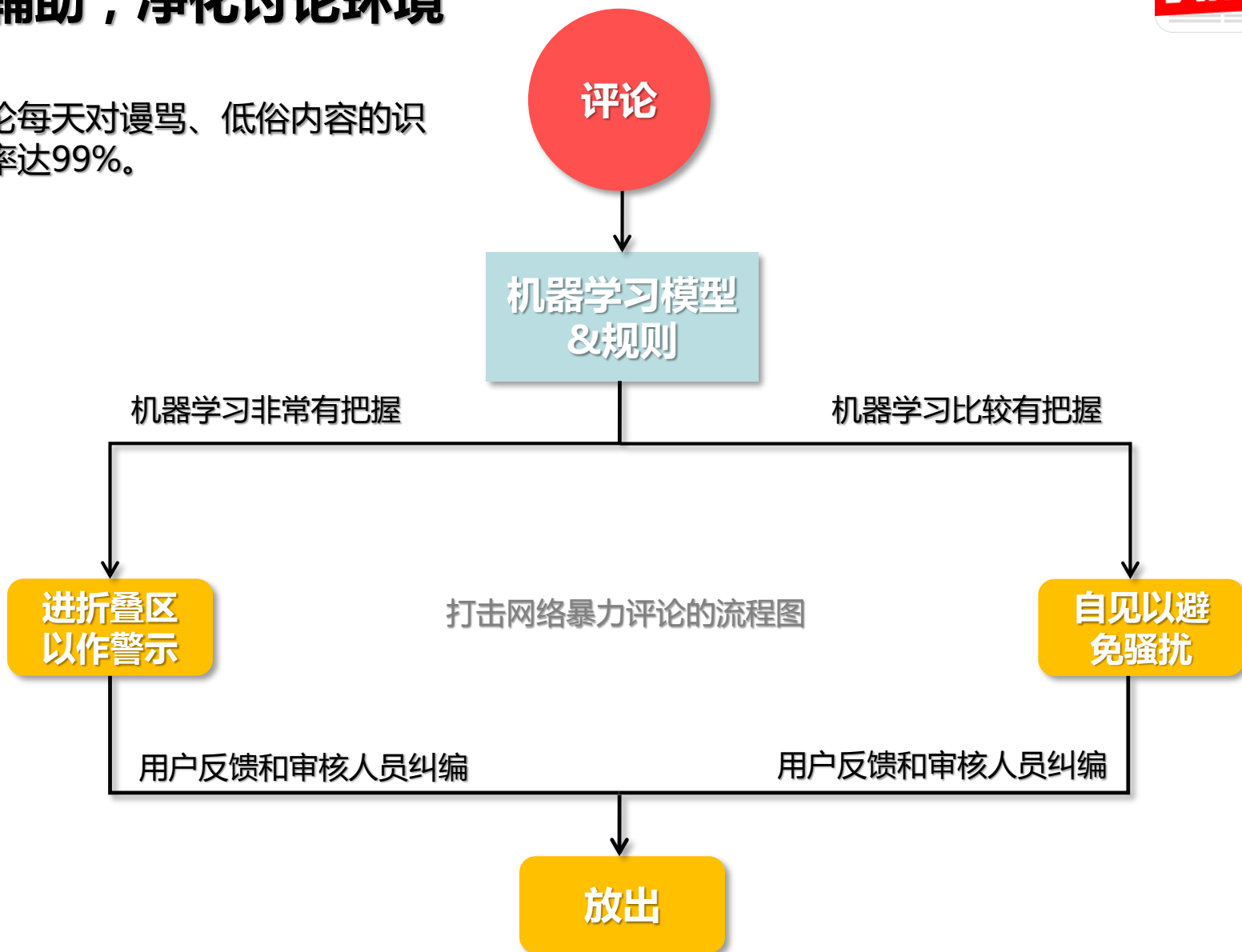
我们监测了今日头条平台2017年第15周至第30周的平均每百万人评论量，增长率为49.76%。



○ 算法辅助，净化讨论环境



头条评论每天对谩骂、低俗内容的识别准确率达99%。



就你话多

——今日头条用户评论数据报告（下）



今日头条算数中心



重要声明

Declaration

本文《就你话多——今日头条用户评论数据报告（下）》（以下简称“本报告”）由今日头条算数中心（以下简称“算数中心”）制作。报告中文字、数据等受中国知识产权法律法规保护。除注明的引用第三方数据及公开信息，本报告所有权归今日头条算数中心所有。

算数中心致力于更及时、更全面、更深度的阅读数据挖掘和数据应用创新。未经允许，不得对本报告进行加工或改造。如有转载或引用，需及时与我们联系并注明出处“**今日头条算数中心**”，同时不得删减或改写报告内容。

特此声明。



关于我们 About Us

今日头条算数中心依靠今日头条海量用户每天数以百亿次级的阅读行为产生的数据，测量每一则资讯的传播力，描绘每一个用户更精准的画像，洞悉用户阅读行为下更深层的动机，致力于更广更深的阅读数据挖掘，用数据服务各个行业。

下载更多原版报告，请访问**今日头条媒体实验室** (mlab.toutiao.com) 报告下载区。

电邮：datadance@bytedance.com

地址：北京海淀区知春路63号中国卫星通信大厦F17

研究人员：刘志毅&裴俊良&吴珏



报告下载区
二维码



公众号二维码