

多方视角下的外卖大战舆情演化分析

关于市场动态与消费者情绪的研究计划

贺颖，刘益弟，唐可姝，陈意舒

南京大学
2025-11-27

选题背景：即时零售的价格战争

- **参与平台：**由京东率先挑起，美团、饿了么（后并入淘宝闪购）随后跟进。
- **核心策略：**“百亿补贴”与激烈价格战，旨在争夺即时零售的流量入口。
- **消费者态度：**从最初对0元购补贴的兴奋转变为审慎。担忧随之上升：
 - ✓ 商家利润受到挤压。
 - ✓ 食品质量与安全保障下降的可能性。
 - ✓ 顾客对于薅羊毛策略的疲劳。



主要研究问题

市场动态

- 消费者兴趣如何随关键事件和补贴波动而变化？
- 2月至7月是否存在明显的参与度阶段？

核心话题

- 讨论的中心议题是什么？
- 随着价格战的持续，这些议题如何演变？

情绪驱动因素

- 什么因素驱动了整体情绪的变化？
- 负面情绪是否与质量问题或服务问题强相关？

区域差异

- 不同省份的消费者对于平台战的态度或关注点是否存在显著差异？

关注事件时间线



数据来源



百度指数

- **目的：** 分析整体趋势以衡量公众关注度。
- **关键数据：** “外卖大战”及相关关键词的搜索量。
- **主要作用：** 建立时间框架并识别兴趣峰值。



小红书

- **目的：** 深入挖掘消费者内容与情绪。
- **关键数据：** 关键节点的高赞笔记、用户评论、话题标签、用户ip地址。
- **主要作用：** 定性分析与情绪建模的数据源。

数据处理过程



1. 数据获取

爬取小红书高互动笔记，并提取百度指数趋势数据。



2. 数据清洗

文本预处理：去噪、去除停用词，并统一日期格式以进行时间序列分析。



3. 数据分析

描述性统计、情绪分类建模、TF-IDF关键词提取及地理聚类。

1. 数据清理与去噪处理

- 去除推广贴、种草广告、品牌营销号内容
- 删除重复文本、疑似机器人账号内容

2. 文本结构化处理

- 提取关键词（价格、配送时间、补贴幅度、骑手权益等）
- 标签化关键特征（含品牌名、事件提及、情绪词等）

3. 地域信息识别

- 从用户主页/标签中提取省市

4. 情绪分析准备

- 对贴子标题与正文进行清洗分句
- 建立情感词典样本

5. 时间序列对齐

- 对百度指数与小红书贴子时间统一格式
- 按天聚合，补全缺失日期

6. 可视化准备

数据形式

名称	修改日期	类型	主页url:
逛逛超市_650e837f0000000001e02384d	2023/10/5 22:15	文件夹	https://www.douyin.com/u
好期待秋天O.o_650ae4690000000001e0227a4	2023/10/5 22:15	文件夹	用户名: 鱼鱼不吃鱼
南宁好GING, 我爱南宁~轨道地产惊喜多多! _...	2023/10/5 22:15	文件夹	介绍: 还没有简介
去看余佳运咯~_65112fdf0000000001e02f9bb	2023/10/5 22:15	文件夹	关注数量: 3
有没有好心人帮我把旁边的男生p掉_64c3f1e8...	2023/10/5 22:15	文件夹	粉丝数量: 335
抓住夏末的最后一缕阳光🌞_65081b83000000...	2023/10/5 22:15	文件夹	作品被赞和收藏数量: 2282
avatar.jpg	2023/10/5 21:18	JPG 图片文件	ip地址: 广西
detail.txt	2023/10/5 21:18	文本文档	性别: 女
			标签: []

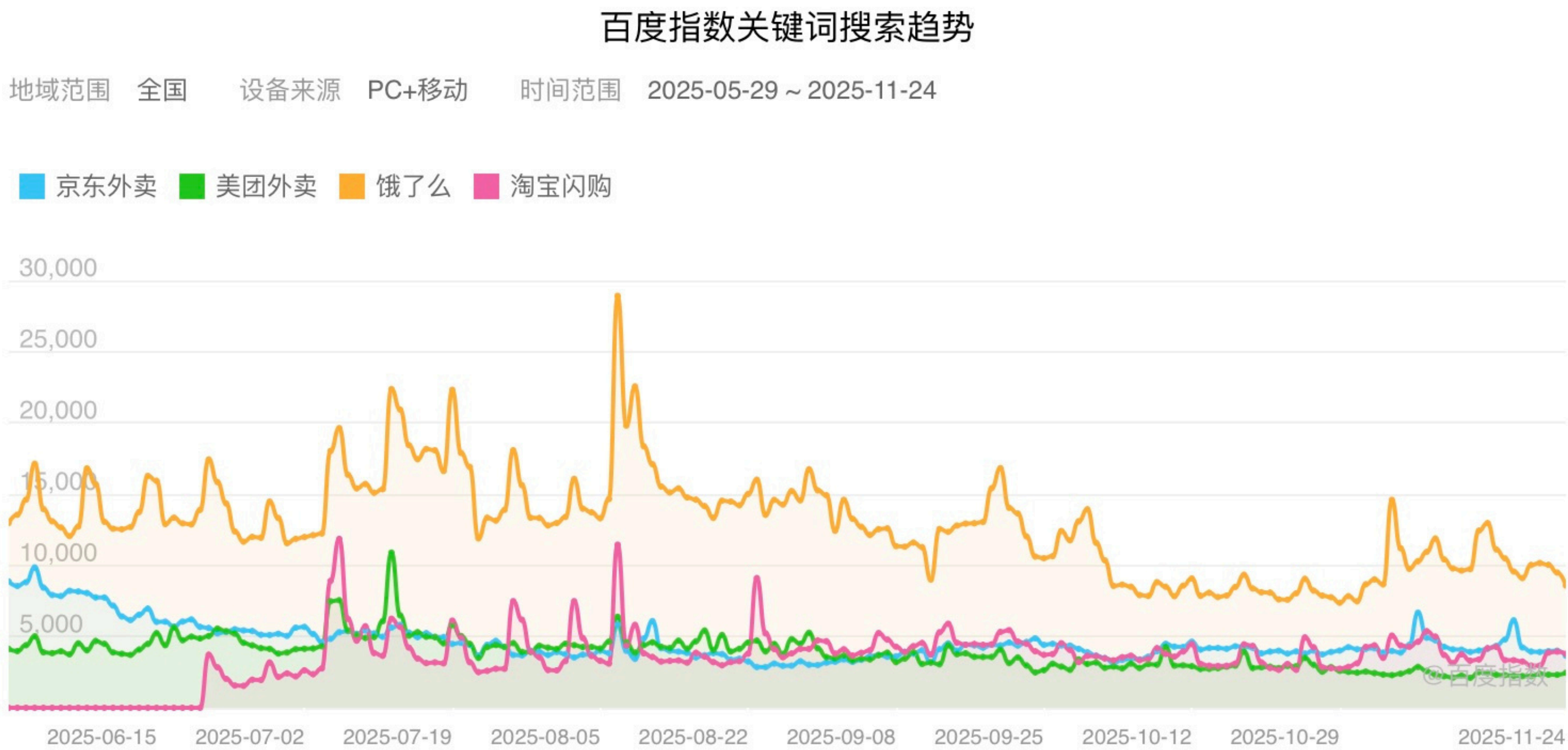
笔记url:
https://www.xiaohongshu.com/explore/65081b830000000001e02069e
笔记类型: normal
笔记标题: 抓住夏末的最后一缕阳光🌞
笔记描述: 🌞夏日库存-1

笔记点赞数量: 4
笔记收藏数量: 1
笔记评论数量: 4
笔记分享数量: 0
笔记上传时间: 2023-09-18 17:42:27
笔记标签: ['森林公园', '浪漫生活的记录者', 'I人', '秋天第1张照片']
笔记ip归属地: 广西

(Based on the open-source repository cv-cat/Spider_XHS)

计划：时间趋势洞察

以百度指数绘制外卖大战及相关关键词的搜索趋势时序图；统计小红书每日相关帖子发布量，与百度指数趋势进行对比验证。

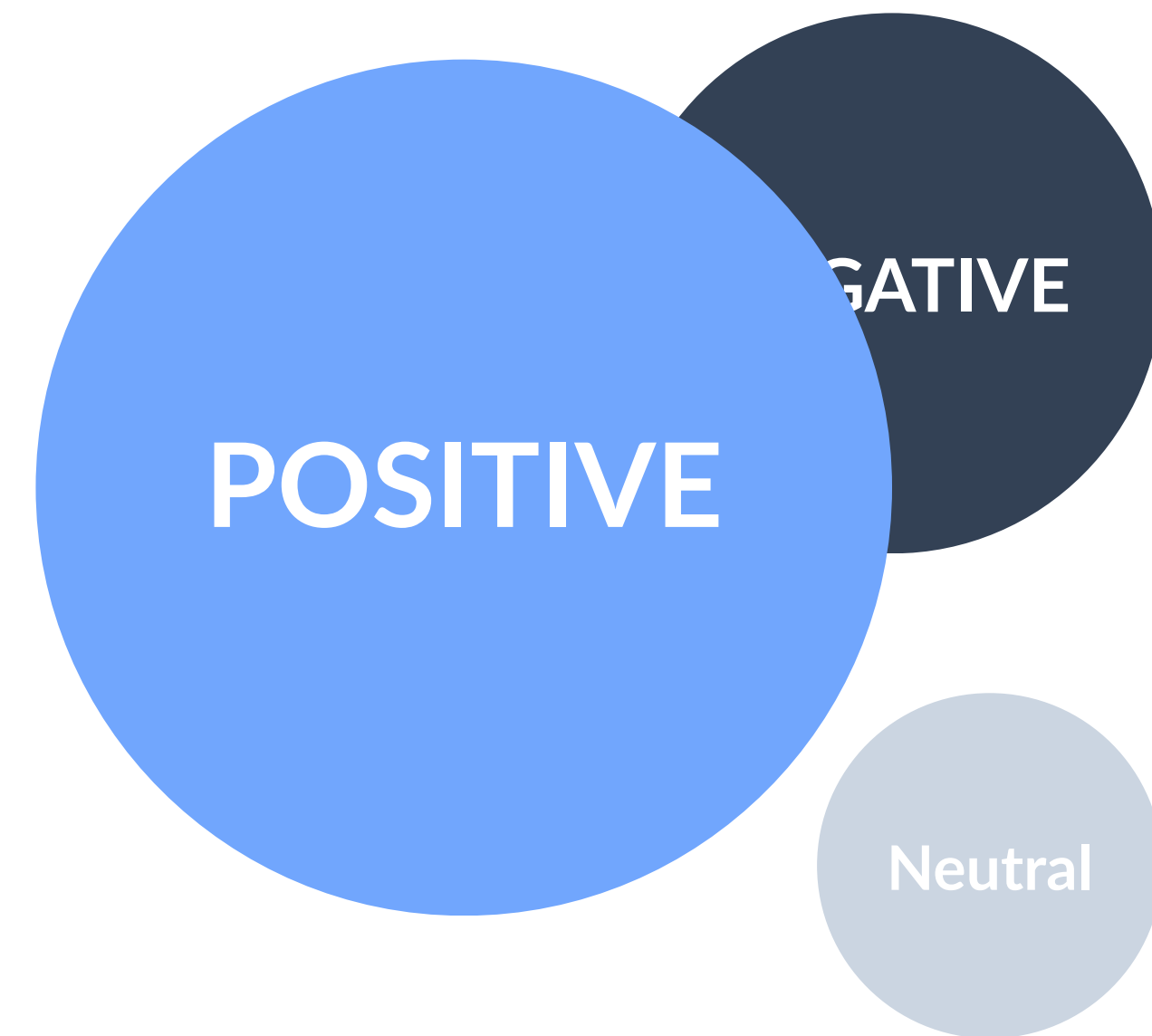


计划：用户情感分析

情感分类评估

- 训练情绪分类模型，对用户帖子内容中展现的情感类型进行评分。（使用Chinese-ELECTRA进行情绪分类）
- **可视化计划：**使用堆叠面积图展示不同情感（正面/负面/中性）随时间的比例变化。

验证猜测：用户情感是否随时间推移存在规律性变化？是否存在地区差异性？



计划：地域分析

地区热力图

- 提取小红书数据中用户的ip地址信息，按省份对进行讨论和热度分级。
- **可视化计划：** 使用地理热力图直观呈现热度区域及区域情绪差异，讨论外卖大战及相关不同话题热点的地区差异性。

验证猜测：一线城市对服务质量的敏感度是否高于对补贴价格的敏感度？



预期结果

1. 舆情热度的结构性波动

- 关键事件（补贴加码、监管介入、平台策略调整）将呈现显著峰值。
- 小红书讨论与百度指数热度可能存在提前或滞后的差异。

2. 议题结构随阶段更迭而变化

- 前期：以补贴力度、平台优惠、配送速度为核心。
- 中后期：讨论重心转向食品安全、品质或骑手压力；可能出现“从消费狂欢到品质焦虑”的议题转折节点。

3. 消费者情绪的非对称演化

- 正面情绪在大战初期占比高；随补贴衰减与负面体验上升，负向情绪占比扩大。
- 情绪波动可能与议题强关联

4. 地域差异：一线城市与其他地区的分化

- 一线城市更关注体验、稳定与服务品质；新一线或二线更敏感于补贴价格。
- 若存在价格战疲态，可能在新一线城市先出现负面拐点。

Q & A

谢谢！