

基于微博言论的 年轻人情感倾向 可视分析

■ 浅浅组个队

■ 成员：何畅2210620
鄢佳宇2213026
卢艺晗2213583

目录

01

组员分工



02

选题介绍



03

数据获取



04

作品展示

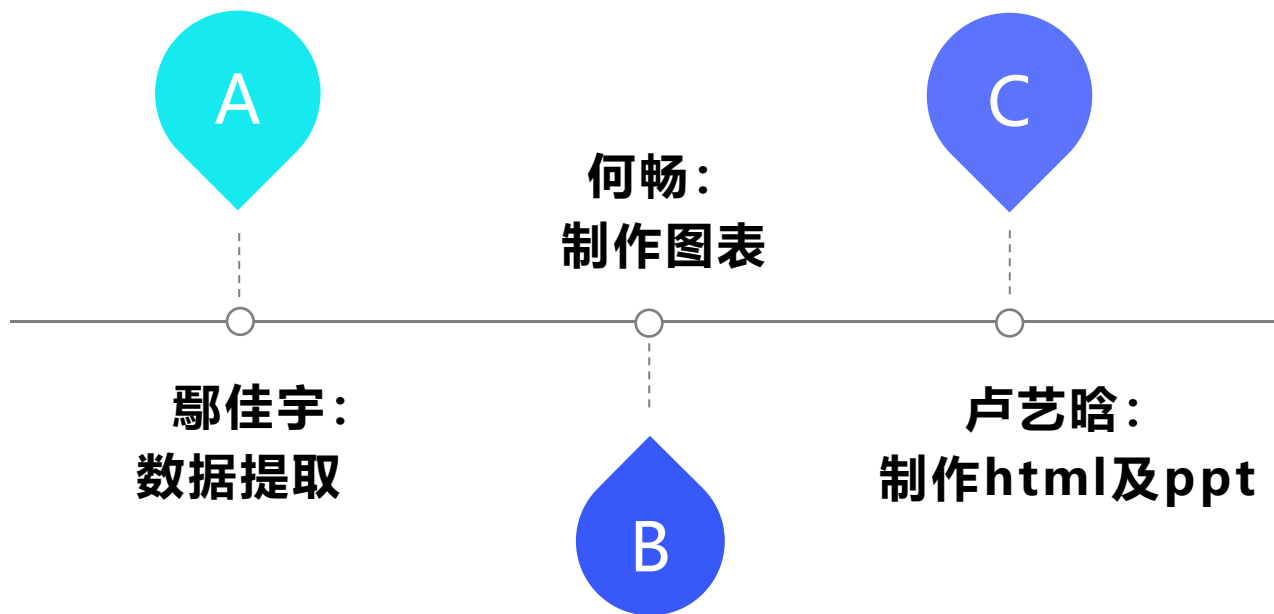


05

作品发现
及案例分
析



组员分工



选题介绍

选题意义：

把握年轻人的主要心态走向，深入了解当下年轻人的情感倾向及因素。

分析目标：

- (1) 对微博上年轻人言论进行情感分类（使用深度学习模型预测）；
- (2) 分析地区、时代、性别等因素对年轻人情感倾向的影响（使用图表展示）；
- (3) 找出影响年轻人心态的主要因素并进行可视展示（使用图表展示）。

数据获取

1、数据描述：基于2010~2023年微博用户的id信息及动态，提取地理位置、发文时间、性别及文本内容等信息；

2、数据来源：微博

3、数据处理：（历时三周）

（1）利用python中的request和lxml库来爬取和处理网页数据；总共爬取数据约28万条，筛选得到有效数据161962条；

（2）使用正则表达式等进行文本数据清洗，剔除与主题无关的信息；

（3）调用大语言模型api进行关键词提取；

（4）使用text RNN+LSTM模型进行情感分析：使用2023年COAE比赛26万数据训练模型200轮，用于微博文本情感预测。



作品展示

html框架

1、首页内容：

- (1) 介绍选题意义、数据资料及结论分析；
- (2) 各种心情发文的微博网页截图，10张图片轮播，真实再现用户的多元情感；
- (3) 数据提取过程中的词云展示；

2、功能设计：

- (1) 采用按钮实现页面跳转；
- (2) 背景采用底图+乳白色半透明叠加；
- (3) 标题使用Canvas绘制艺术字；
- (4) 插入图片，动态轮播；
- (5) 采用表格介绍数据信息，展示网页的可靠程度；

欢迎来到年轻人的心态空间站！

首页	地区差异	时间差异	性别差异	影响因素
----	------	------	------	------

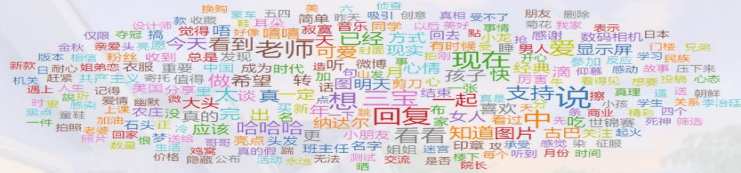
选题意义

欢迎来到年轻人的心态空间站！

在这里，我们基于微博数据，深度挖掘并揭示当下年轻人的心态情感倾向。通过地区差异、时代差异、性别因素等维度的多方位解析，我们得以感知到当今年轻人的心态走向，看到年轻人内心世界的多元化和复杂性。同时展示引发其心态的主要因素，为深入理解年轻人提供一把钥匙。

数据来源

数据来源：	微博
有效数据总量：	161962条
数据处理：	爬虫：数据清洗：大模型预测：
数据处理工具：	python的request和lxml库-爬虫； 大语言模型API-关键词提取； text RNN+LSTM模型-情感预测；
数据覆盖范围：	中国各省份；2010~2023年；



作品发现和分析总结

- 本作品统计了2010~2023年微博言论中体现出的年轻人情感倾向（非常积极、积极、消极、非常消极），综合四个图表得出以下分析发现：
- (1) 在微博言论中，年轻人主流情感倾向为积极，但疫情前后消极情感占比显著升高；
 - (2) 在上海、江苏等经济高度发达、生活压力大的地区，年轻人更容易产生消极情绪；而在贵州、天津等相对宽松的生活中环境中，年轻人更容易产生积极情绪；
 - (3) 由于传统社会角色期待和压力应对策略的性别差异等原因，男性比女性更不易在公开言论中表达负面情绪；



图表A： 中国地图 (支持交互)

- 1、主题：从空间维度分析年轻人情感差异并进行可视展示；
- 2、效果：鼠标悬浮地图上的省份区域，该区域高亮并显示相应的统计信息
- 3、图表设计：
 - (1) 省份颜色：展现该省份年轻人心态的积极/消极程度（越红表示越积极，整个作品系统颜色映射逻辑相同）；
 - (2) 地图上的资料卡：统计该省份的数据总量、积极/消极心态占比
 - (3) 底部配有图表分析发现；

2010~2023年年轻人情感倾向地区差异及数据概览

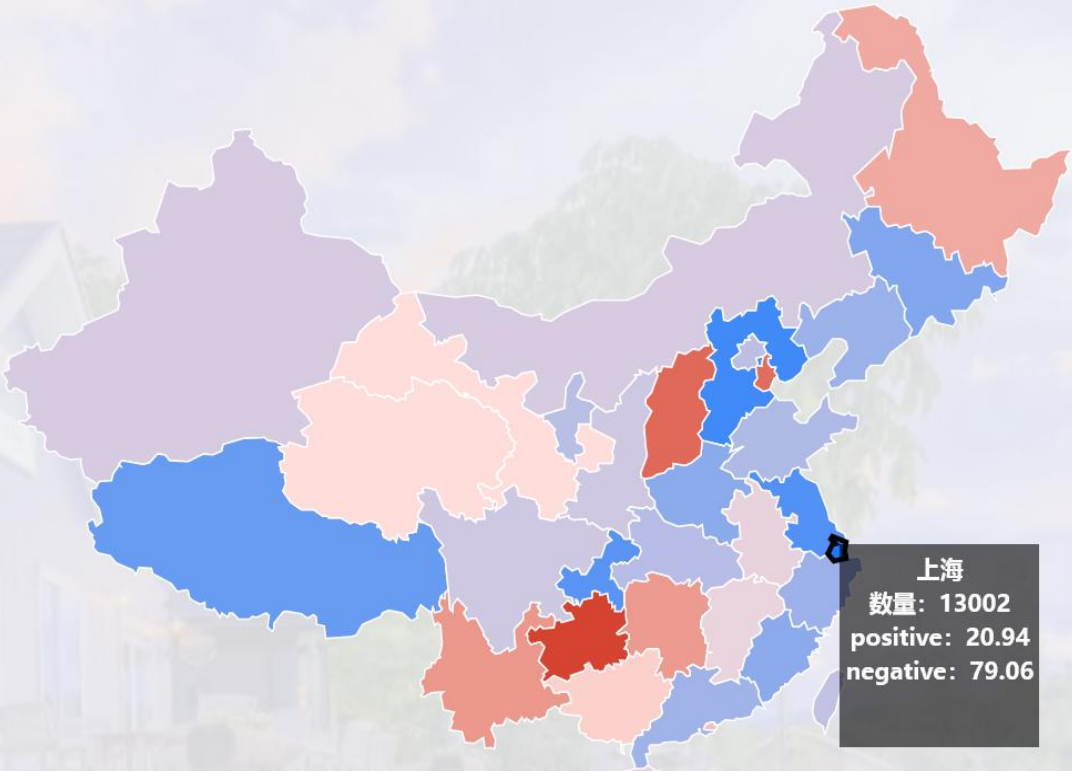
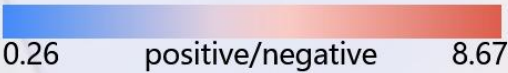
首页

地区差异

时间差异

性别差异

影响因素



发现:

情感倾向积极排名前三的省份: 贵州、天津、陕西;

情感倾向消极排名前三的省份: 上海、江苏、河北;

图表B

柱状时间追及图

(动态+交互)

- 主题：

从时间维度分析年轻情感差异
并进行可视展示

- 图表设计：

展现从2010年到2023年
年轻人主要心态的种类及
变化；

可以点击按钮进行重置；

底部配有图表分析发现；

2010~2023年年轻人情感倾向时间柱状追及图



2014

重新播放

发现：

微博上年轻人言论以积极为主，2020年前后情感倾向中消极占比最大，其余均为非常积极占比最大；
占比最低的始终都是非常消极；

图表C

旭日图

(静态)

01

主题：

从性别角度分析年轻人情感差异并进行可视展示

02

图表设计：底层展示情感倾向的比例，外层展示各时间段内各种情感倾向的男女占比

2010~2023年年轻人情感倾向性别差异旭日图

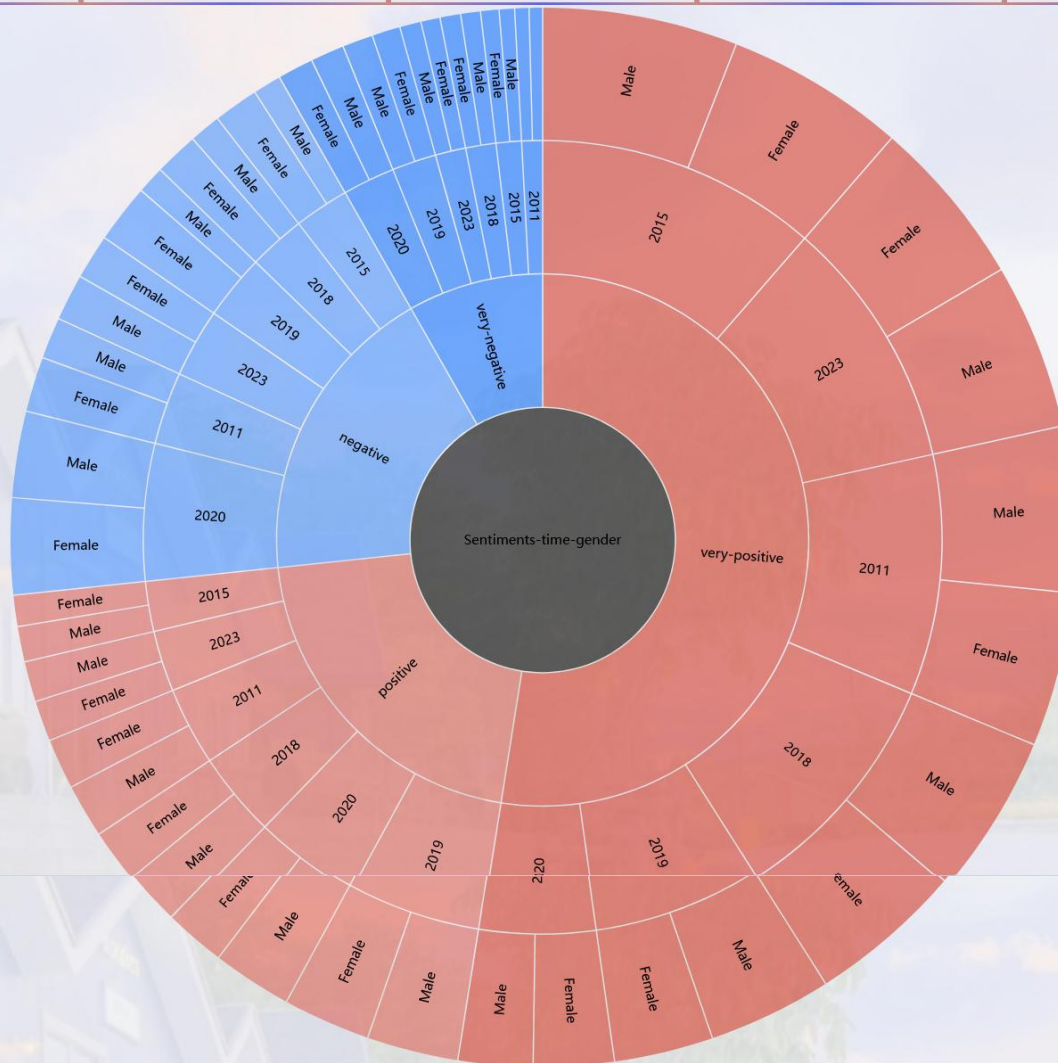
首页

地区差异

时间差异

性别差异

影响因素



发现:

2011~2023年, 情感倾向以非常积极占比最大, 非常消极占比最小;

各类情感倾向的男女比例无显著差异, 但积极和非常积极中, 多数年份男性占比偏高;

图表D

可缩放圆形图

(交互)

- 主题：

展现当下（2023年）的年轻人情感倾向占比及影响因素。

- 图表设计：

大圆为情感倾向，半径表示占比；
点击放大可在小圆内浮现出相关的影响因素关键词，同时展现占比。

2023年度年轻人情感倾向主要影响因素关键词

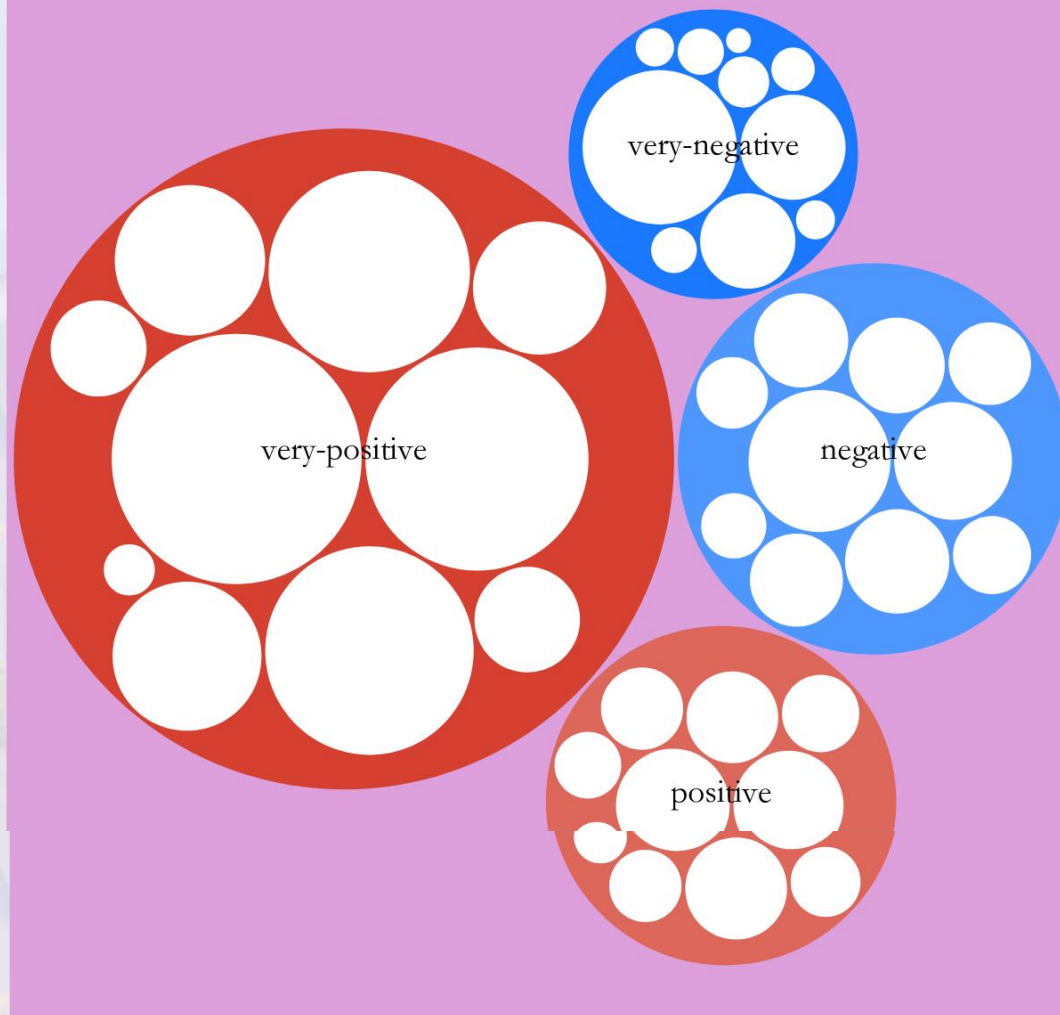
首页

地区差异

时间差异

性别差异

影响因素



发现:

2023年度情感倾向，仍以非常积极占比最大

积极情感中，主要相关关键词有学业成就、吃、追星、旅游等；

作品发现 及 案例分析

(案例截图已在前面展示，并伴有相应分析，在此综合以上做出总结和分析)

本作品统计了2010~2023年微博言论中体现出的年轻人情感倾向(非常积极、积极、消极、非常消极)，综合四个图表得出以下分析发现：

(1) 在微博言论中，年轻人主流情感倾向为积极，但疫情前后消极情感占比显著升高；

(2) 在上海、江苏等经济高度发达、生活压力大的地区，年轻人更容易产生消极情绪；而在贵州、天津等相对宽松的生活环境中，年轻人更容易产生积极情绪；

(3) 由于传统社会角色期待和压力应对策略的性别差异等原因，男性比女性更不易在公开言论中表达负面情绪；

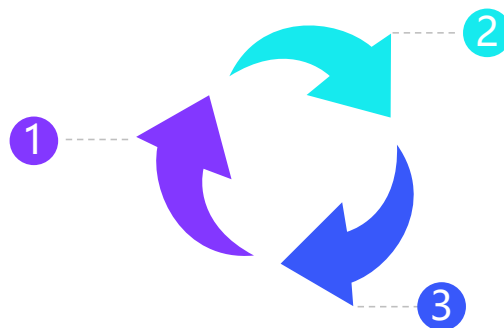
(4) 近些年在内卷的大环境下，“内卷”“加班”等事件深刻影响着年轻人的情绪，而“吃”和“回家”等缓解压力的事件更容易为年轻人带来幸福感。

综合以上分析发现，地区的经济发展状况、生活氛围、时代背景、传统观念等因素始终深刻影响着年轻人的情感倾向，每个个体的情感与这个时代息息相关。

但即使当下生存压力增大，年轻人在公开言论中的主流情感依然为积极，可见年轻人在面对困难和挑战时，依然保持了相对积极的心态。

总结

特色：
自行收集数据并进行分析，提供了分析年轻人情感倾向的新视角；



困难回顾：
数据获取和处理（耗时长、技术要求高）

作品局限性：

(1) 微博用户并不代表所有年轻人，这份报告的结果可能并不能全面反映年轻人的情感倾向。

(2) 作品主要关注了微博上的公开表达，可能无法捕捉到年轻人更深层次、私人的情感状态。

(3) 情感倾向的复杂性和多元性可能超出了单一的关键词和地区、时间、性别的分类所能描述的范围。