

---

# 短视频数据分析报告

---

唐艺璇

1

分析背景及目的

2

数据可视化分析

3

结论及建议

**01**

## 分析背景及目的



公司主要业务为创作短视频，发布于各个平台（今日头条，B站，youtube）。为了提高平台的视频播放量，在了解平台推荐机制的基础上进行分析（今日头条为例）：

- **关键词总结**，包括热点关键词及沉积关键词
- 根据往期数据，总结影响推荐量的 **相关因子**
- 建立**预测**推荐量的模型

# 02

## 数据可视化分析

关键词分析

相关因子分析

预测模型建立

[illegible]

排名前十**高频词汇**包括：  
农村，男子，美女，女子，结婚，揭秘，秘密，手机，竟然，老人

# 相关因子分析

## 数据准备:

- 缺失值处理
- 删除重复值
- 构建特征值

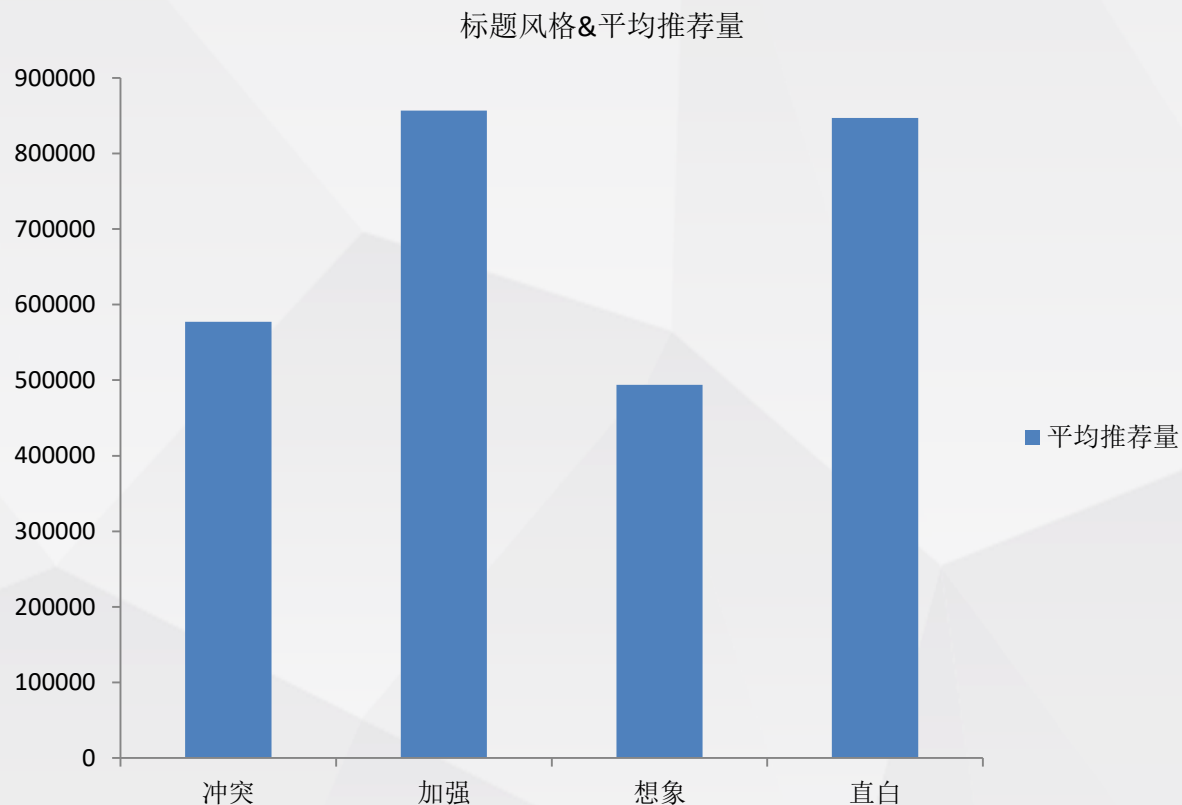
## 分析方向:

标题风格&推荐量

影片分类&推荐量

用户行为&推荐量

## 2.1 标题风格



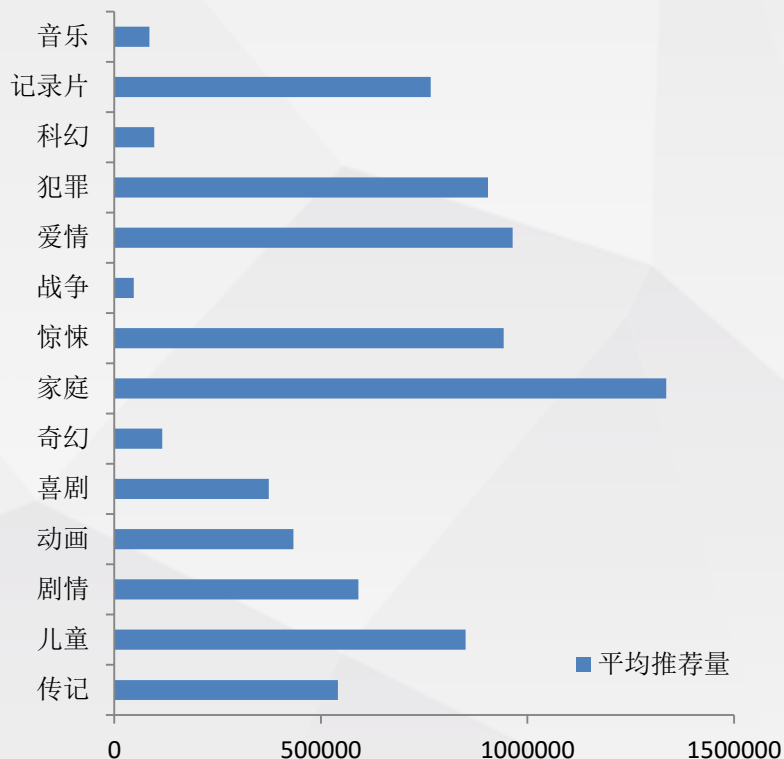
标题风格包括  
冲突型，加强型，  
想象型，直白型。

其中，加强型  
及直白型语言风格  
标题平均推荐量更  
高，冲突型随后。



## 2.2 影片分类

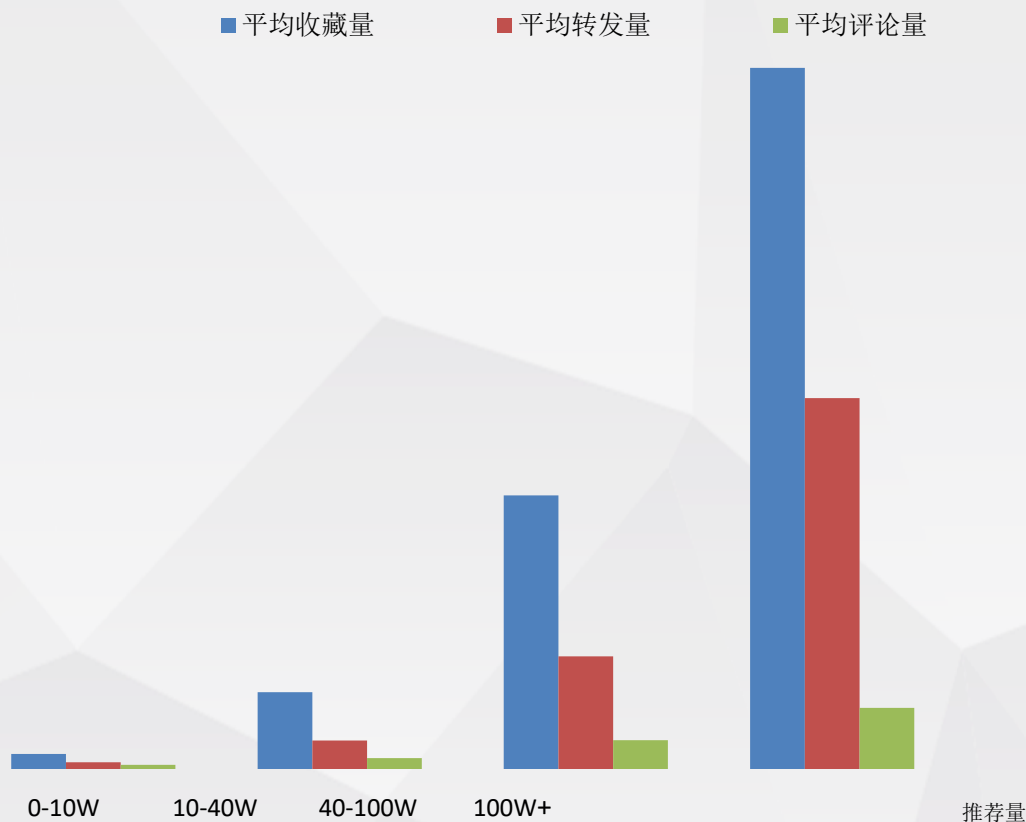
影片分类&平均推荐量



将视频加上分类标签，分析不同分类与推荐量的关系。

总体来看，**家庭类**影片平均推荐量最高，**爱情，犯罪，惊悚**类型随后，**科幻，战争**类推荐量最低。

## 2.3 用户行为分析



用户行为代表着来自用户方对视频的反馈，这也是影响推荐量的重要因子，包括收藏，转发，评论等。

可以看出，推荐量越高的视频，其平均收藏量，转发量，评论量均更高。说明**用户行为对推荐量有着积极的作用**。

# 预测模型建立

根据转发量，收藏量，评论量，标题类型，影片分类等特征清理数据，构建新特征，训练模型用来预测推荐量

Model = ridge(alpha=35)

Rmse\_cv 为0.96

# 03

## 结论及建议

1. 平台系统最初会根据关键词分类进行首次推荐，根据词云分析，可在标题种加入合适的关键词
2. 标题类型，影片分类和推荐量有一定的关系，用户行为对推荐量有着积极的作用
3. 构建预测模型，用以通过视频投放一段时间的数据预测最终的表现

---

**THANKS**

---