

数据分析系统控件

数据分析可视化

- 购买时间分布
- 购买时间和得分...
- 交互特征（点赞...

情感分类和情感关键词挖掘

- 情感分类-Snow...
- 情感分类后关键...

关键信息提取

- TextRank关键词...
- LDA主题分析

共现语义网络

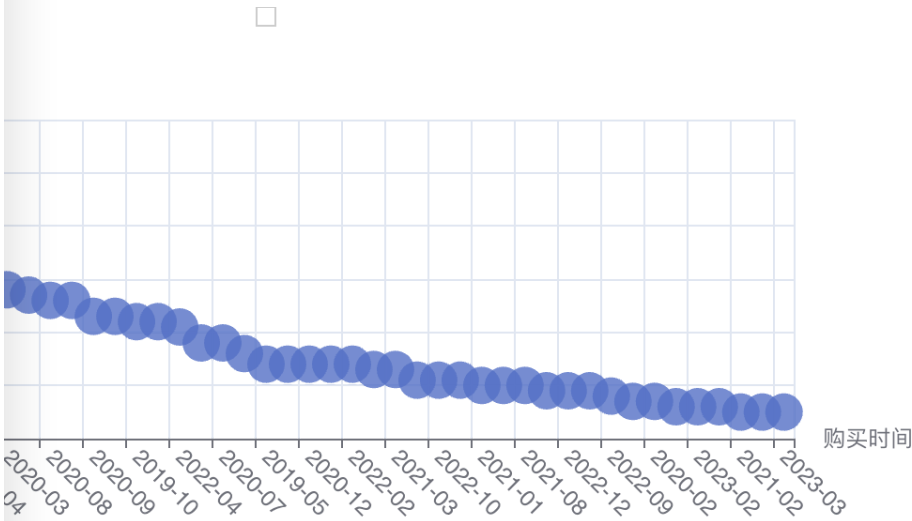
- 语义网络图(有向)

总结和其他分析角度

- 其他潜在分析角度
- 总结

power) 商品评论分析系统

式挖掘京平台下“电小二 (Dxpowers) 电小二 (Dxpowers) 户外电源  
量电脑充电宝露营应急备用储能电源 (型号: 户外电源500W) ”的商品



期间, 即11月和6月

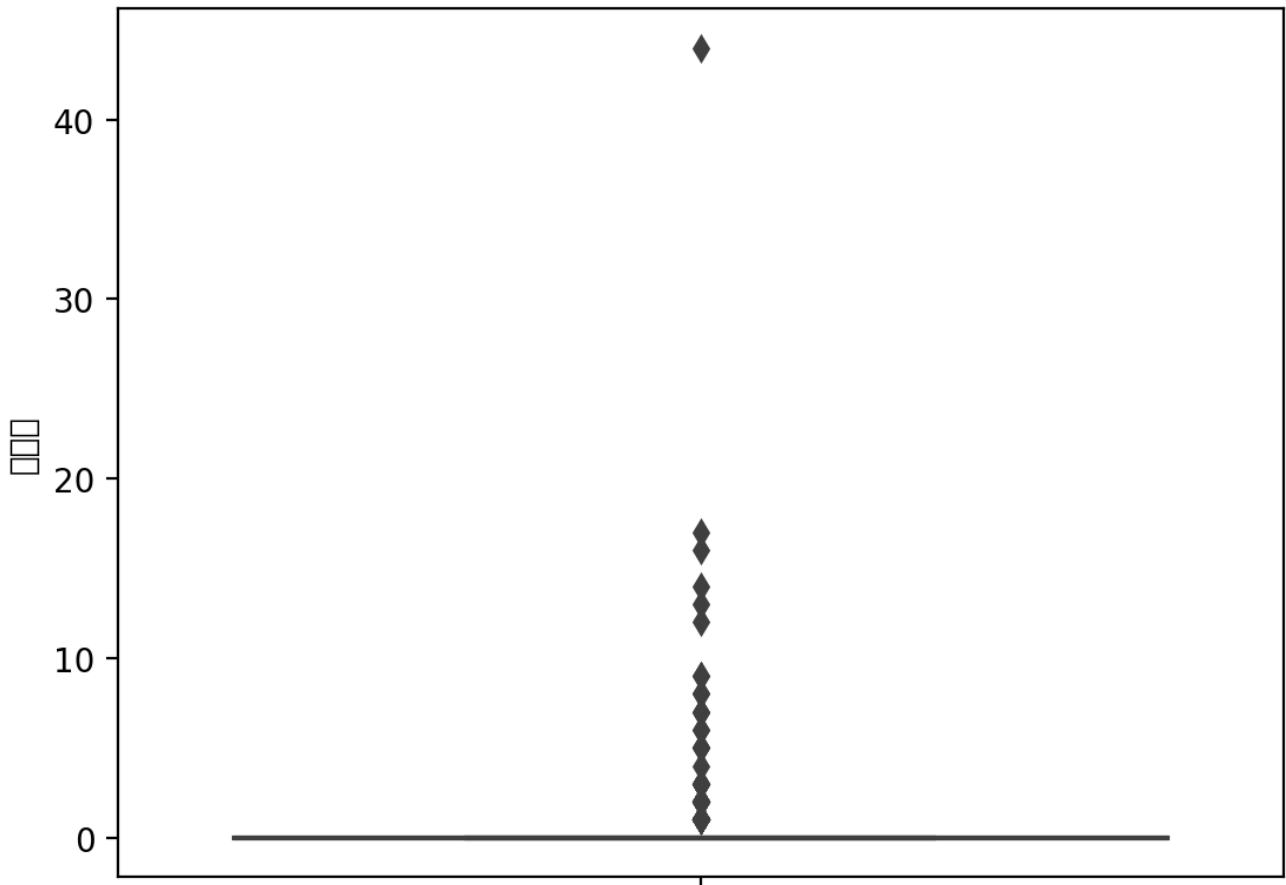
的关系

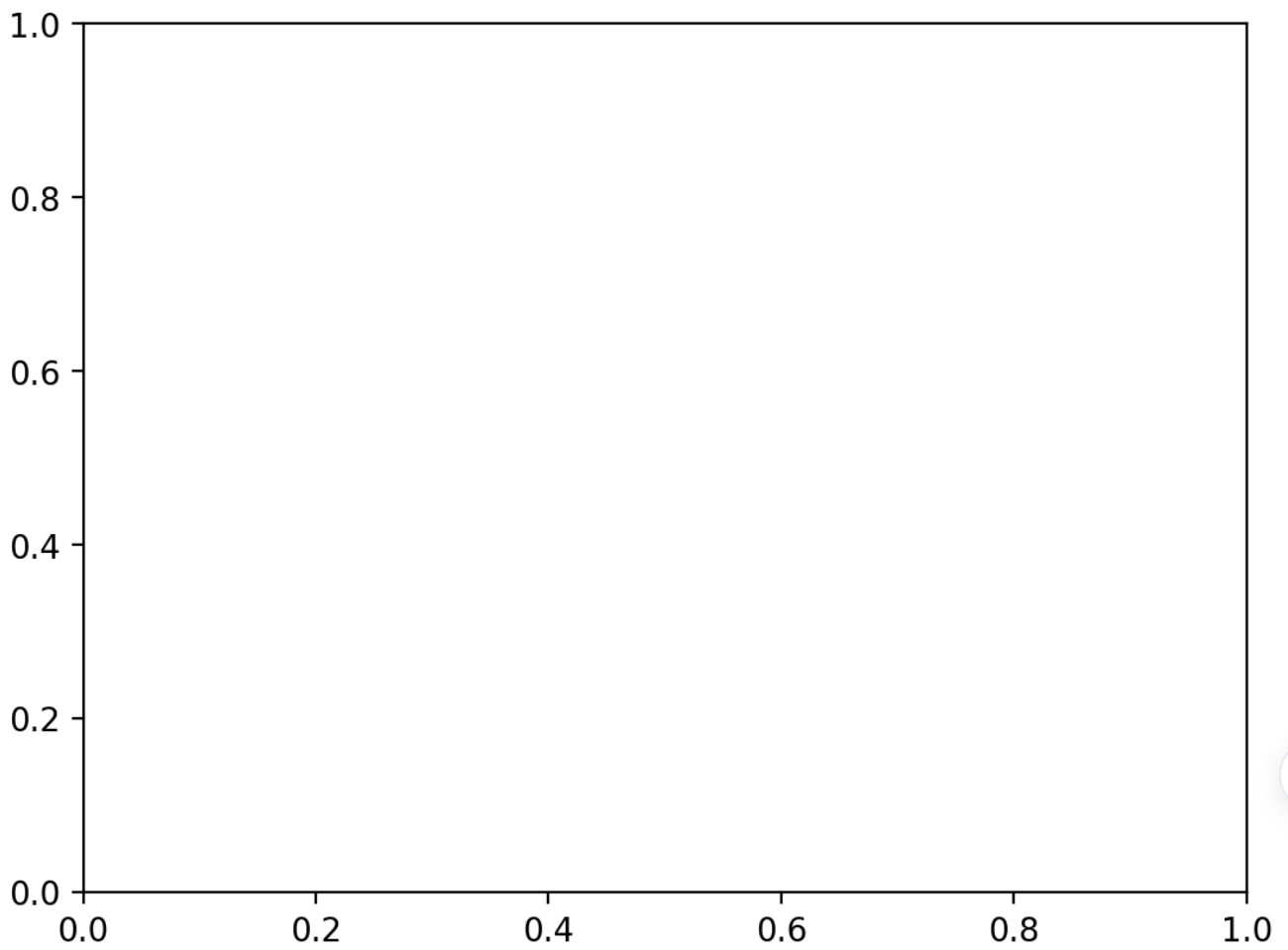
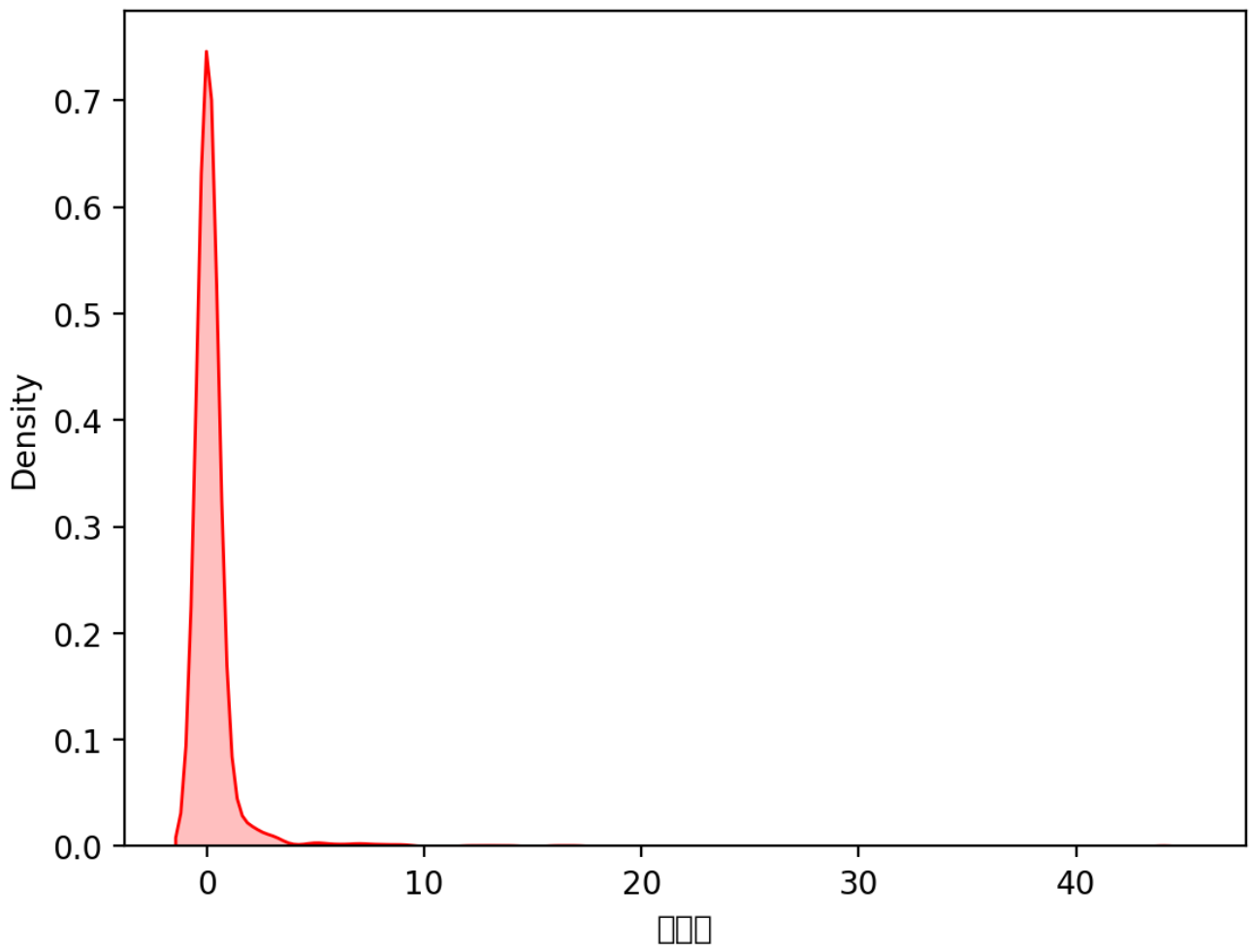


因为评价得分大多为4或5分，对得分进行归一化，即得分取值为[0,1]

在通过对得分进行归一化之后，发现用户的评价得分商品在大促期间（如6，10，11月）下降明显。结合后文负面情绪关键词挖掘的结果，暗示在大促期间，由于商品发货量大，客服售后应对不及时，难免出现中差评，这会影响卖家DSR动态评分的高低

## 交互特征（点赞数）的数量分布





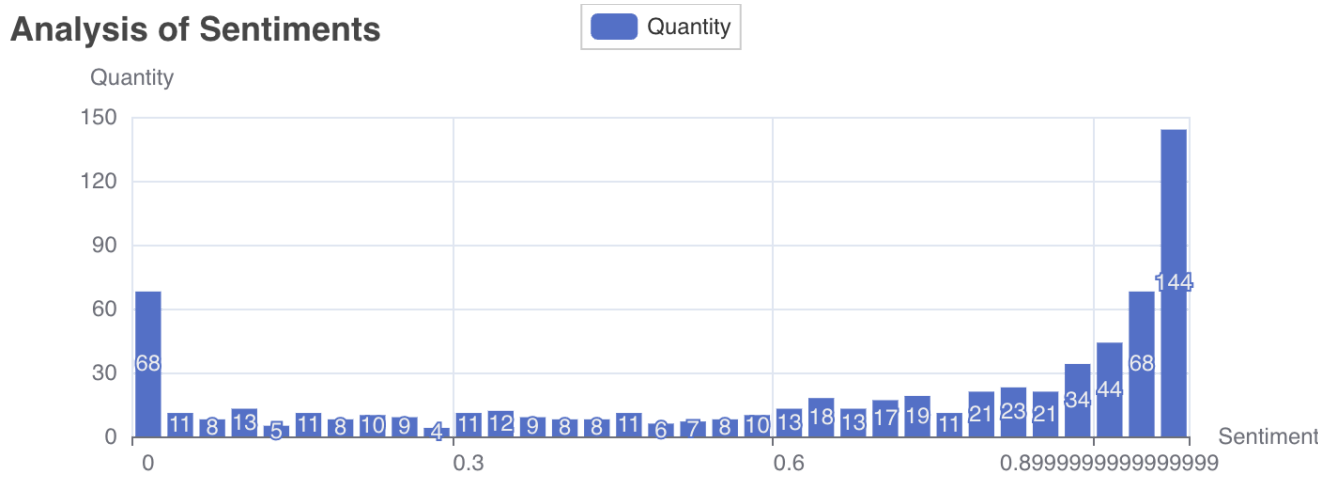
最大点赞数: 44

平均点赞量: 0.31451612903225806

发现评论的点赞数主要集中在0-4之间，最大点赞数为44，平均点赞量为0.314

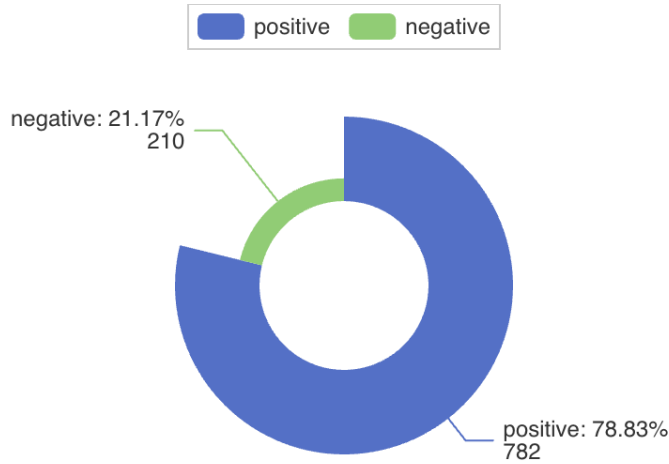
## 情感分类-SnowNLP

经过人工判别，发现3分得分以上的评论多为称赞，持鼓励态度。判定3分为情感极性分界线（即>3分为“积极”），提取正向评论的特征 'pos'



SnowNLP的情感得分呈现两头分布，主要集中在0分和1分附近，中间分布较少，1分附近最多

### 情绪分布玫瑰图



根据情感分数的分布情况和分数数量，判断阈值为0.5，将情感分数进行二分类，得到积极情感评论782条，消极情感评论210条

以0.5为阈值，对情感得分进行0，1截断。查看混淆矩阵，分析情感得分的分类准确率

precision socres:

0.8044354838709677

	0	1
0	18	192
1	2	780

None

SnowNLP准确率：0.8044354838709677

SnowNLP准确率为：0.804。SnowNlp的结果中，正向情感得分有782个，而负向情感得分仅有210个，样本极不均衡。类别1（积极）判对780个，判错了2个，类别0（消极）判对了18个，判错了192个，总的来看类别1（积极）的准确率较高

## 情感分类后关键词挖掘-词云图

### 全体评价词云图



### 积极评价词云图



# 好评



## 消极评价词云图



## 小结：

正面/积极评价的关注点在于'做工质感', '便携性能', '蓄电容量', '充电速度', '抗摔能力'

负面/消极评价的关注点在于'服务'

## 潜在问题：

根据消极情绪的词云图和混淆举证的得分反映出SnowNLP对长句或复合句的情感极性判读的准确率(80.4%)有待提升，可以通过以下方法解决：

- 建立自定义电商评论情感词词典
- 用更多相关语料来定制化训练模型

## TextRank关键词提取

### Top 20 关键词

充电 1.0000

做工 0.8218

电源 0.5627

使用 0.4650

质感 0.4056

速度 0.3905

没有 0.3146

电小二 0.3045

这个 0.3000

容量 0.2854

质量 0.2577

问题 0.2483

收到 0.2481

小时 0.2394

蓄电 0.2348

电量 0.2294



外观 0.2170

客服 0.2113

值得 0.2021

满意 0.1973

Top20关键词分别是 {充电, 做工, 电源, 使用, 质感, 速度, 没有, 电小二, 这个, 容量, 质量, 问题, 收到, 小时, 蓄电, 电量, 外观, 客服, 值得, 满意}

频词汇中除了词云图中出现的词外, “外观”和“质量”也进入了前排, 说明除了质量这个硬指标, 外观颜值也是消费者够买电小二的重要关注点

“满意”一词的高频程度, 说明不少消费者对电小二的总体满意度较高

## LDA主题分析

### 正面主题分析

主题0

0.044\*的" + 0.032\*很" + 0.026\*好" + 0.021\*了" + 0.017\*不错" + 0.012\*非常" + 0.011\*也" + 0.010\*用" + 0.009\*电源" + 0.009\*小二"

主题1

0.026\*的" + 0.018\*了" + 0.011\*不错" + 0.009\*用" + 0.006\*电" + 0.006\*时间" + 0.006\*很" + 0.006\*做工" + 0.006\*还" + 0.005\*充电"

主题2

0.024\*做工" + 0.020\*的" + 0.019\*好" + 0.017\*了" + 0.017\*很" + 0.016\*不错" + 0.015\*质感" + 0.015\*充电" + 0.013\*工质" + 0.013\*便携"

### 负面主题分析

主题0

0.039\*了" + 0.025\*的" + 0.011\*用" + 0.008\*充电" + 0.008\*户外" + 0.008\*就" + 0.007\*好" + 0.007\*是" + 0.006\*收到" + 0.006\*很"

主题1

0.035\*的" + 0.014\*也" + 0.014\*很" + 0.014\*好" + 0.013\*了" + 0.009\*用" + 0.008\*给" + 0.006\*户外" + 0.006\*客服" + 0.006\*非常"

主题2





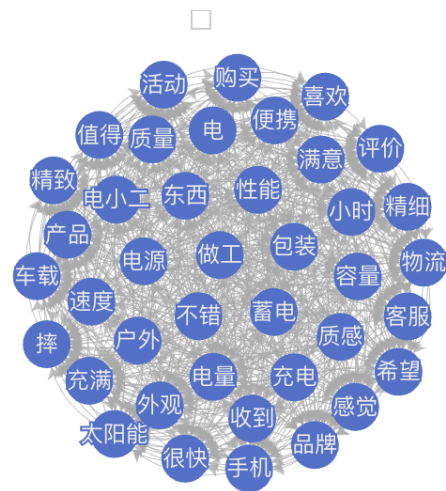
0.028\*的"+ 0.020\*充电"+ 0.017\*用"+ 0.017\*了"+ 0.011\*可以"+ 0.011\*电源"+ 0.009\*还"+ 0.008\*好"+ 0.007\*很"+ 0.006\*移动"

LDA主题模型结果没能挖掘电小二评价文本中的潜在主题，由于正面评论和负面评论的三个主题之间交集较多，说明主题分离的效果不好，无法进一步洞悉有效的主题。

## 语义网络图(有向)

复杂语义网络还可以用Rost.CM6、Gephi或者Neo4j软件自动化构建

电小二共现语义网络（有向）



## 小结：

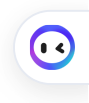
由于关键词之间的连线越多，说明共现的次数越多，发现

- 比如“外观”这个积极情感词主要指向了“精致”，“精细”和“品牌”，说明外观可以作为消费者对电小二品牌忠诚度的考量因素
- 比如“车载”主要指向了“太阳能”和“活动”，暗示消费者在户外活动时由于车载设备的大量充电需求，有购买太阳能配件的倾向性

## 总结

根据对京东平台上电小二（Dxpowers）户外电源600W大功率220V移动便携大容量电脑充电宝露营应急备用储能电源（型号：户外电源500W）的消费者评价情况进行多维度的分析，得出以下4点结论：

- 1- 消费者对电小二在做工质感、便携程度、蓄电容量、充电速度量、抗摔能力等方面满意度高
- 2- 电小二在售前和售后服务质量上仍有改善空间。可以通过提升智能客服机器人的回复精准度，对差评用户进行人工回访，以提升品牌口碑
- 3- 电小二可以进一步提升在外观设计上的研发，以进一步提升其品牌忠诚度
- 4- 电小二可以进一步探索不同储能设备和太阳能配件捆绑销售的组合营销策略



# 其他潜在分析角度

## 1- 话题聚类

- 通过kmeans聚类，尝试将评论信息归纳为若干话题

## 2- 用户画像/用户分层（RFM模型）

- 在拥有消费者的\*\*基本人口统计信息（如性别、年龄、教育水平），地理位置信息，兴趣爱好信息，消费行为信息（包括消费者的购买历史、购买渠道、购买频率、购买金额），社交媒体信息等\*\*的前提下，来建立消费者的用户画像，了解他们的兴趣、偏好和行为模式，以帮助电小二品牌更好地了解他们的目标市场，并针对性地进行市场推广和产品开发

- RFM用户分层模型通过用户历史消费数据，以三维坐标系（最近消费时间（R）、消费频次（F）、消费金额（M））来进行用户价值分析

## 3- 关联分析

- 通过分析商品型号与用户评价之间的关系，可以发现不同型号的商品在用户心目中的位置，帮助电小二了解市场需求并根据需求调整产品的研发方向

## 4- 竞品分析

- 对竞争对手的评论数据来对电小二在服务、价格、销售策略等方面的优势和劣势



