

# 京东消费行为分析

## 报 告 书

指导老师: 何老师

著作人: 李尧

# 目 录

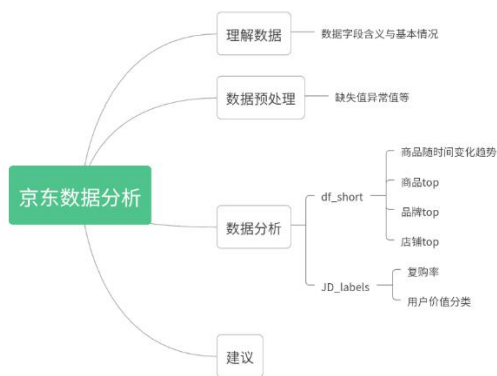
一、背景.....	3
二、 研究问题.....	3
三、 数据分析.....	4
1、数据字段介绍.....	4
1.1、JD_labels.....	4
1.2、df_short.csv.....	5
2.用户行为分析.....	5
2.1 用户从 2018 年 2 月 1 日到 2018 年 4 月 15 日各个行为变化趋势? .....	5
2.2 有购买行为的用户人均购买多少次? .....	7
2.3.用户复购情况如何? .....	7
2.4.用户下单时间大多在什么时候? .....	9
2.5.点击就购买的用户和加购就购买的用户占比? .....	9
2.6.用户 7 天与 30 天活跃次数分析? .....	10
2.7.根据 kmeans 对用户进行分类.....	10
3.商品与品牌，店铺分析.....	12
3.1 最受用户欢迎的品牌? .....	12
3.2 最受用户欢迎的物品种类? .....	13
3.3 最受用户欢迎的商店种类? .....	14
3.4 两个半月以来各商店种类用户点击的变化趋势? .....	15
四、 建议.....	15

# 一、背景

京东 JD.COM-专业综合网上购物商城，销售超数万品牌，4020 万种商品，囊括家电、手机、电脑、服装、居家、健康、母婴、美妆、个护、食品、旅游等品类。数据时间是从 2018 年 2 月 1 日到 4 月 15 日两个半月的销售数据。本文主要通过 python 对用户数据进行分析，探讨用户行为体现出的购买趋势，为商家提供一些运营方面的建议与策略。

## 二、研究问题

1. 从 2018 年 2 月 1 日到 2 月 15 日商品的浏览，收藏，购买，加购物车的变化趋势如何？为什么会有这种趋势？
2. 为什么此类商品会被用户购买甚至复购？有何特征？
3. 哪种商品销量最好？ 哪些品牌最受用户欢迎？哪些店铺排名靠前？
4. 用户浏览，购买，加购，收藏商品主要时间在一天中的什么时候？为什么在这个时间段？
5. 哪些用户是重点用户？哪些用户即将流失？



## 三、数据分析

### 1、数据字段介绍

#### 1.1、JD\_labels

JD\_labels.csv

- Unnamed: 0: 未知
- customer\_id: 用户id
- time\_browse: 浏览时间
- time\_Order: 订购时间
- cate\_most\_browse: 浏览最多的物品种类
- cate\_most\_Follow: 收藏最多的物品种类
- cate\_most\_SavedCart: 加购最多的物品种类
- cate\_most\_Order: 购买最多的物品种类
- month\_buy: 近30天购买次数
- month\_Cart: 近30天加购物车次数
- month\_active: 近30天活跃次数
- week\_buy: 近7天购买次数
- week\_Cart: 近7天加购次数
- week\_active: 近7天活跃次数
- last\_browse: 上一次浏览距离今天天数
- last\_SavedCart: 上一次加购距离今天天数
- last\_Order: 上一次购买距离今天天数
- interval\_buy: 最近两次购买间隔天数
- browse\_not\_buy: 是否有商品浏览未下单
- cart\_not\_buy: 是否有商品加购未下单
- buy\_again: 是否复购

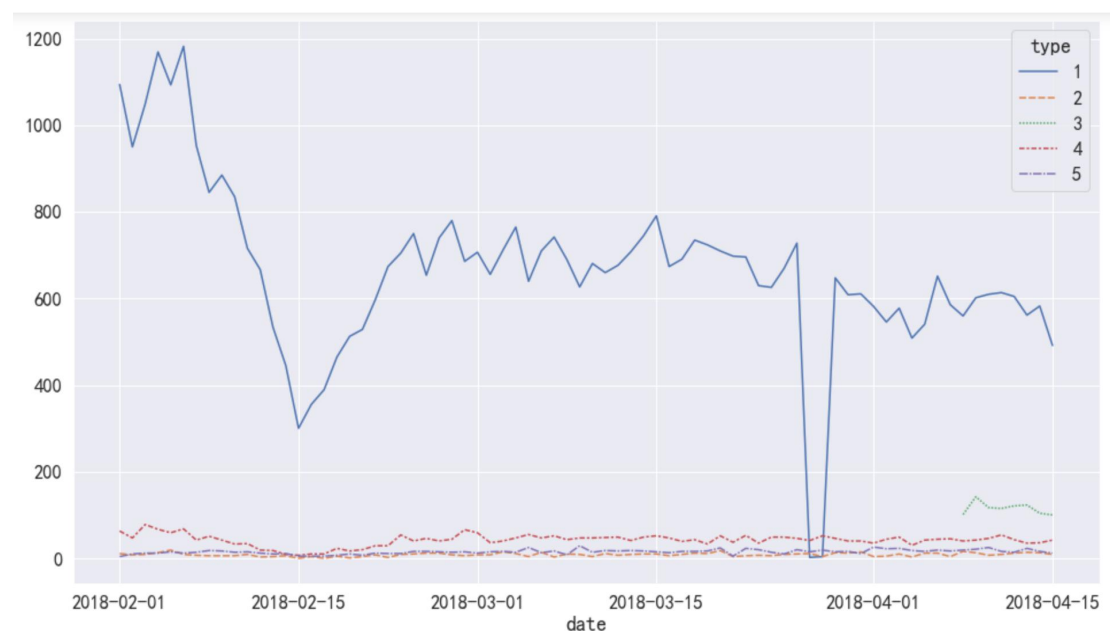
## 1.2、df\_short.csv

df\_short.csv

- Unnamed: 0: 未知
- customer\_id: 用户id
- product\_id: 商品id
- type: 类别 1: 浏览, 2: 收藏, 3: 购买, 4: 加入购物车, 5: 评论
- brand: 品牌
- category: 物品种类
- shop\_category: 商店种类
- date: 日期

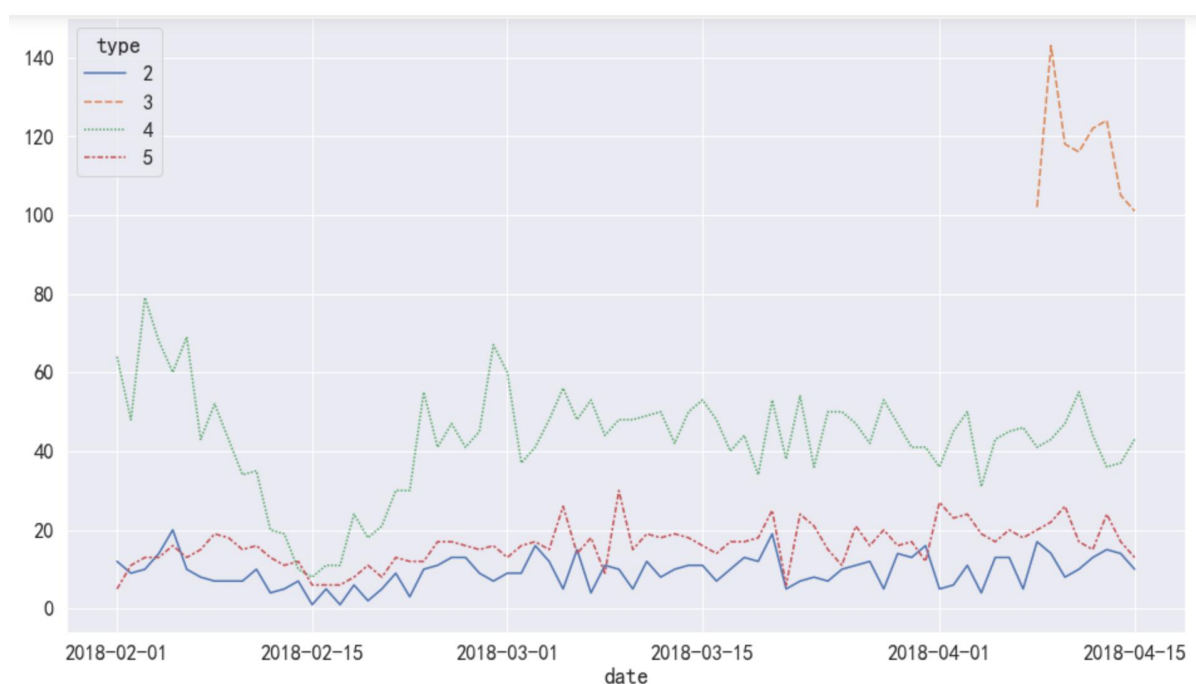
## 2.用户行为分析

### 2.1 用户从 2018 年 2 月 1 日到 2018 年 4 月 15 日各个行为变化趋势？



分析：

1. 可看出在 2 月 15 日前后，浏览量收藏量加购量都减为 0 了然后又都逐渐上升而 2018 年 2 月 15 日是除夕，2 月 16 日是春节，猜测这种变化是受春节影响。
2. 可看出浏览量很多，且浏览量有随着时间逐渐减少的趋势，而加购物车，购买，评论，收藏都相对较少，可以把浏览量去掉分析收藏，加购，购买这三种行为的变化趋势。
3. 在 2018 年 3 月 25 日左右浏览量急速降为 0，而后又急剧上升，推测原因：  
①春节消费过度，用户购买欲望下降②3 月下旬是春季旅游高峰期，上网人数可能会减少

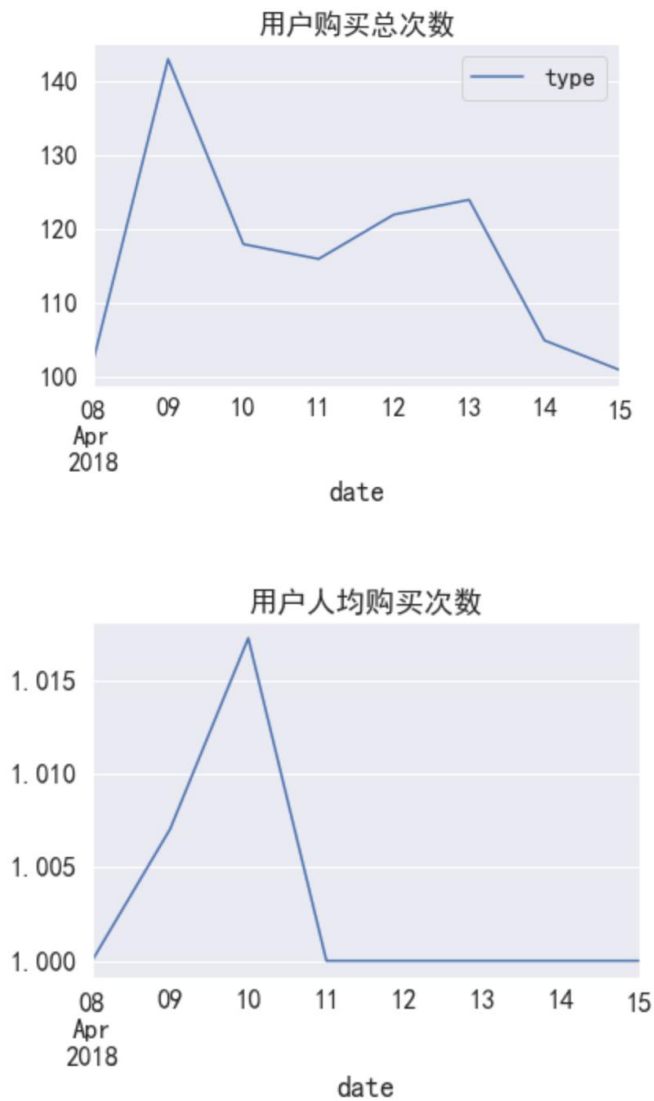


分析：

type: 类别 1: 浏览, 2: 收藏, 3: 购买, 4: 加入购物车, 5: 评论。

1. 可看出加入购物车随时间变化相对于其他的都较多，加入购物车用户量超过了收藏评论的用户量。购买量从一开始到 4 月 7 日没有，从 4 月 8 日突然迅速增多然后又稍微下降，猜测：①是商品有活动了②节假日前后（2018 年 4 月的 5, 6, 7 日为清明放假时间）③该商品具有季节性（清明后天气回暖）所以大家纷纷购买

## 2.2 有购买行为的用户人均购买多少次？



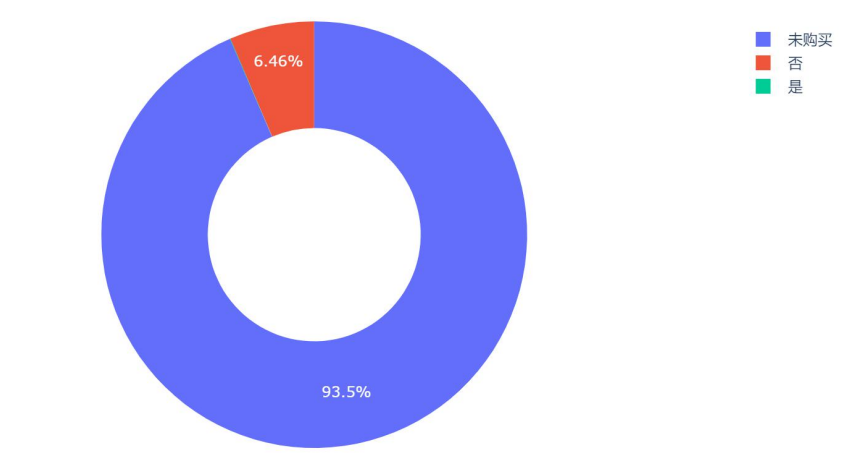
分析：

4月8日到4月15日期间，日均消费110次左右，人均消费1次左右，商品复购率不高。可提升商品质量或加大广告力度。

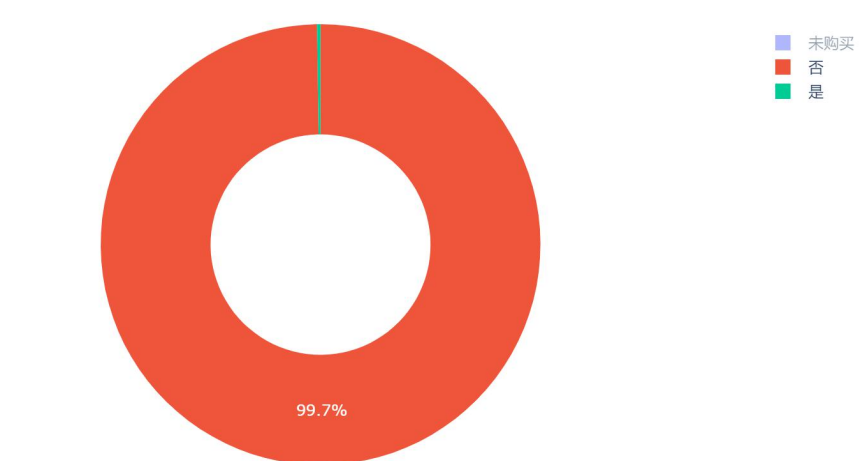
## 2.3.用户复购情况如何？

```
未购买    45831
否         3164
是           8
Name: buy_again, dtype: int64
```

复购情况



复购情况



分析：  
大多数用户 ‘只看不买’ ， 购买并复购的用户仅占 0. 3%

customer_id	time_Order	cate_most_Order	last_Order	interval_buy
1240926	上午,晚上	Basketball	2.0	33.0
118758	上午,晚上	Coat,Tea	5.0	11.0
502169	中午,下午	Face Cream,Sunscreen Cream	1.0	5.0
895001	上午	Face Cream	46.0	NaN
1496979	上午,中午	Necklace	52.0	19.0
558952	下午	Tea	18.0	19.0
121982	上午,晚上	Facial Essence	9.0	62.0
675346	上午,中午	Cookie,Milk	37.0	14.0

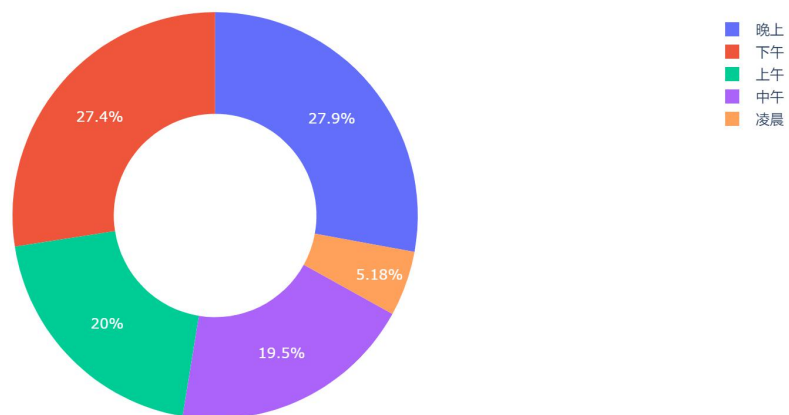
分析：



可能是数据量不大的原因，复购用户复购的多为日常生活用品，复购时间在 1-50 天内均有分布，复购时间在 10 天内客户占了复购用户的一半，商家应注重一周内对客户的宣传。

## 2.4.用户下单时间大多在什么时候？

下单时间情况



分析：

用户下单最多的时间是下午和晚上，凌晨消费最少，所以活动可以选择在下午和晚上进行推送。

## 2.5.点击就购买的用户和加购就购买的用户占比？

```
labels['browse_not_buy'].value_counts()
```

```
是    43465
否    5538
Name: browse_not_buy, dtype: int64
```

```
print('有 {} 点击行为转为了购买行为'.format(5538/49003))
```

有0.11301348897006305点击行为转为了购买行为

```
labels['cart_not_buy'].value_counts()
```

```
是    43467
否    5536
Name: cart_not_buy, dtype: int64
```

```
print('有 {} 点击加购转为了购买行为'.format(5536/49003))
```

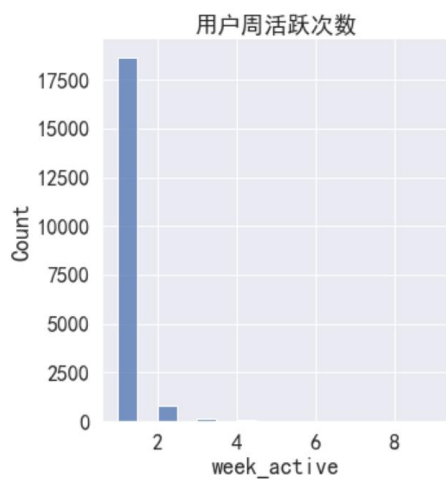
有0.11297267514233822点击加购转为了购买行为

分析：

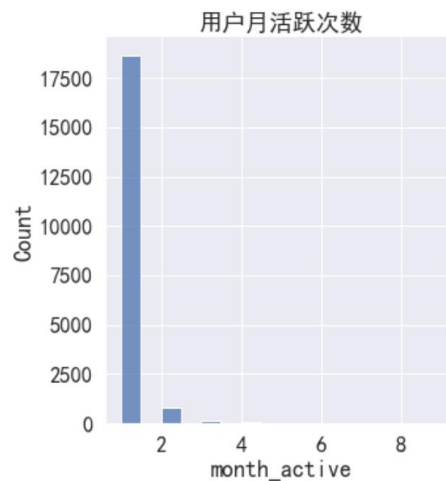
加入购物车的商品和浏览的商品转为被购买商品几率差不多大，都为 11.2% 左右，用户加入购物车却不买，用户是有购买欲望的，可能是商品的价格等不合适，建议办一些活动来促销，有 0.11297267514233822 点击行为转为了购买行为，可以增强网站的推荐机制，或许可以提高此比率

## 2.6.用户 7 天与 30 天活跃次数分析？

```
1.0    18617
2.0     818
3.0     161
4.0      38
5.0      16
6.0       6
9.0       2
7.0       1
Name: week_active, dtype: int64
```



```
1.0    18617
2.0     818
3.0     161
4.0      38
5.0      16
6.0       6
9.0       2
7.0       1
Name: month_active, dtype: int64
```



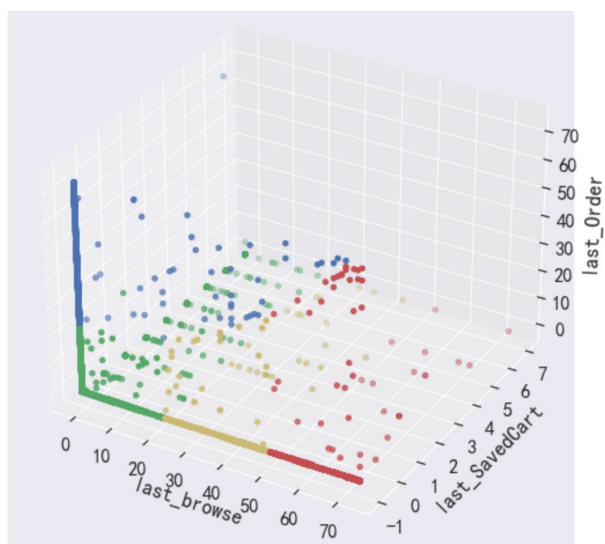
分析：

用户月活跃次数与用户周活跃次数完全一样。

推测：可能数据里 7 天内的新用户占比多，然后新用户近 1 个月活跃次数和 7 天内基本一样，所以数据看起来差别不大。

## 2.7.根据 kmeans 对用户进行分类

kmeans 根据用户上一次行为距离今天天数聚类



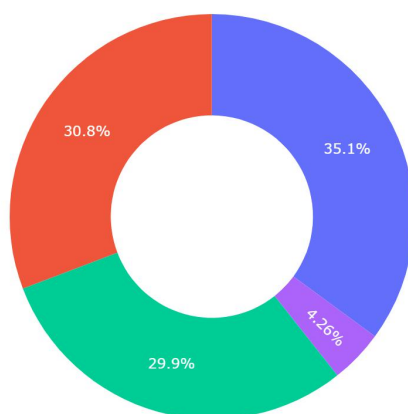
Red:0

Green:1

Blue:2

Yello:3

各类用户占比

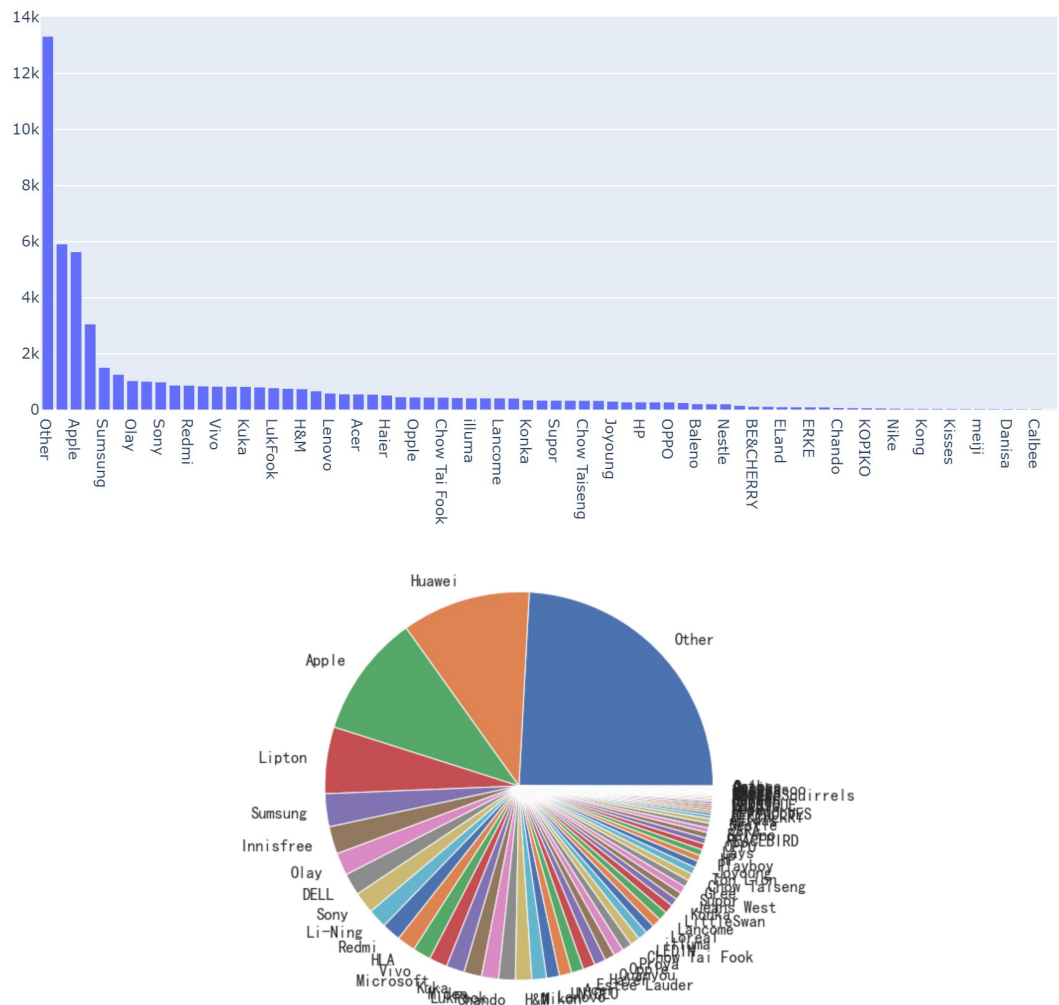


- 标签 0: 一般客户
- 标签 1: 重要深耕客户
- 标签 2: 高价值客户
- 标签 3: 一般客户

### 3.商品与品牌， 店铺分析

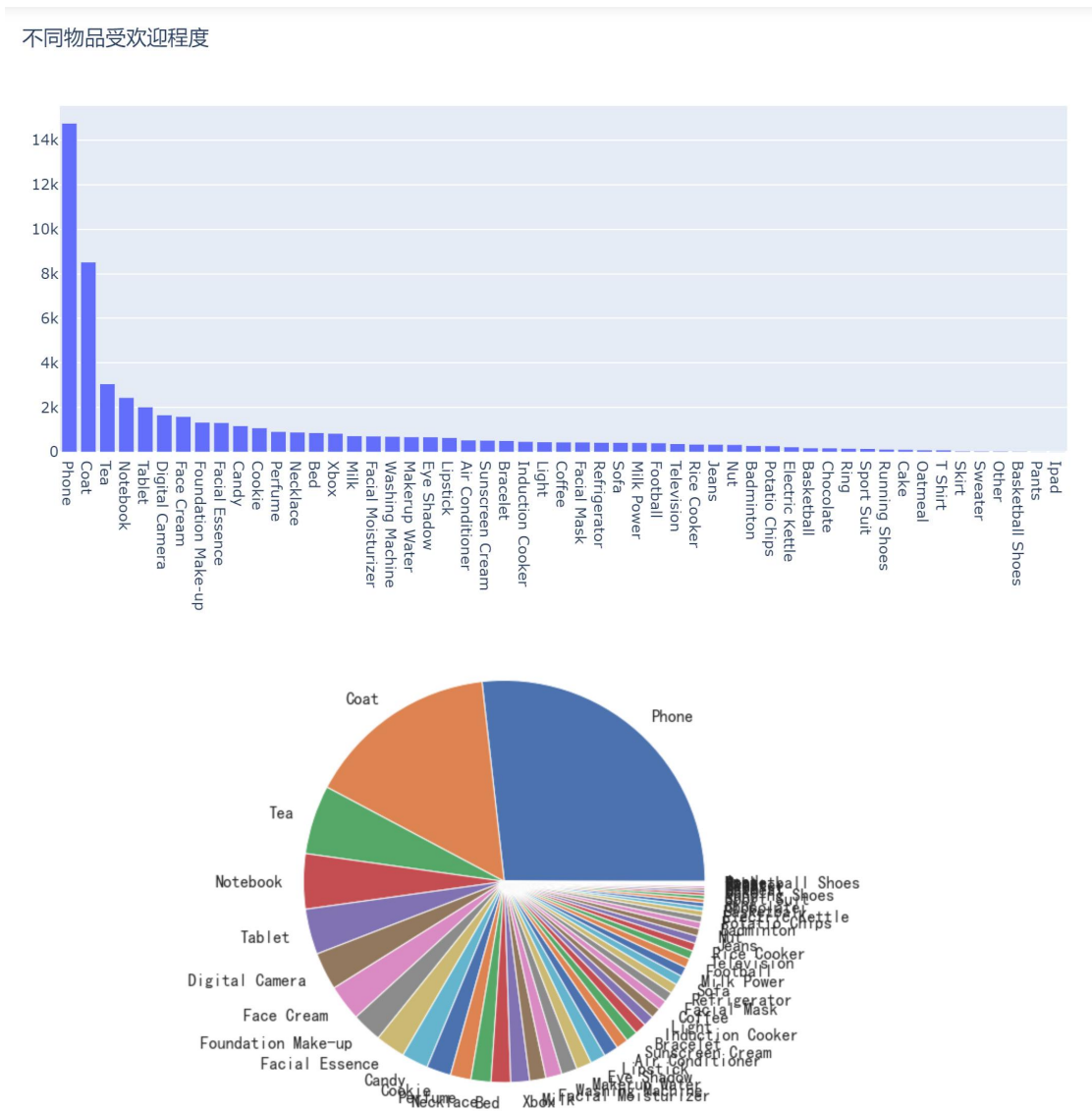
#### 3.1 最受用户欢迎的品牌？

不同品牌受欢迎程度



Others 可能为一些其他的品牌，从图中可看出 others 为第一，其次就是各大知名品牌了，所以建议，在营销时除了知名品牌的宣传外，也要注重一些小品牌的宣传。

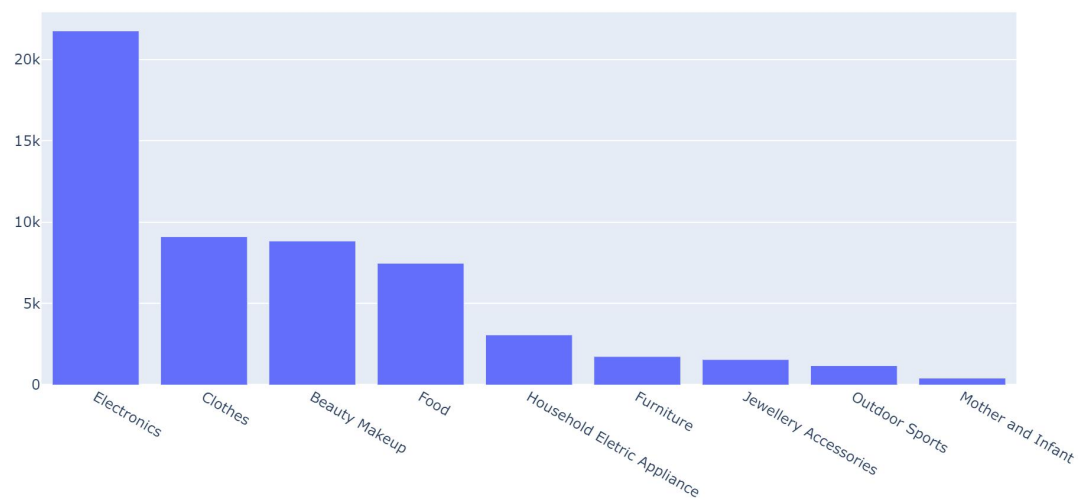
### 3.2 最受用户欢迎的物品种类？



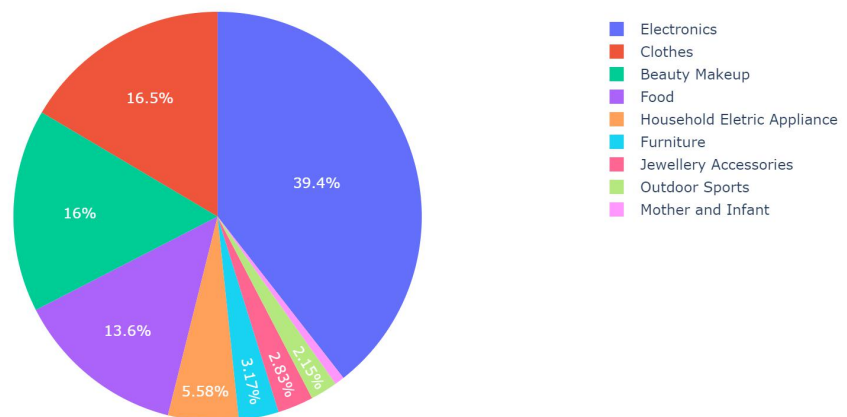
分析：

最受用户欢迎的物品种类前 7 分别是：手机，外衣，茶，笔记本/笔记本电脑，平板电脑，数码相机，洗面奶。

### 3.3 最受用户欢迎的商店种类？

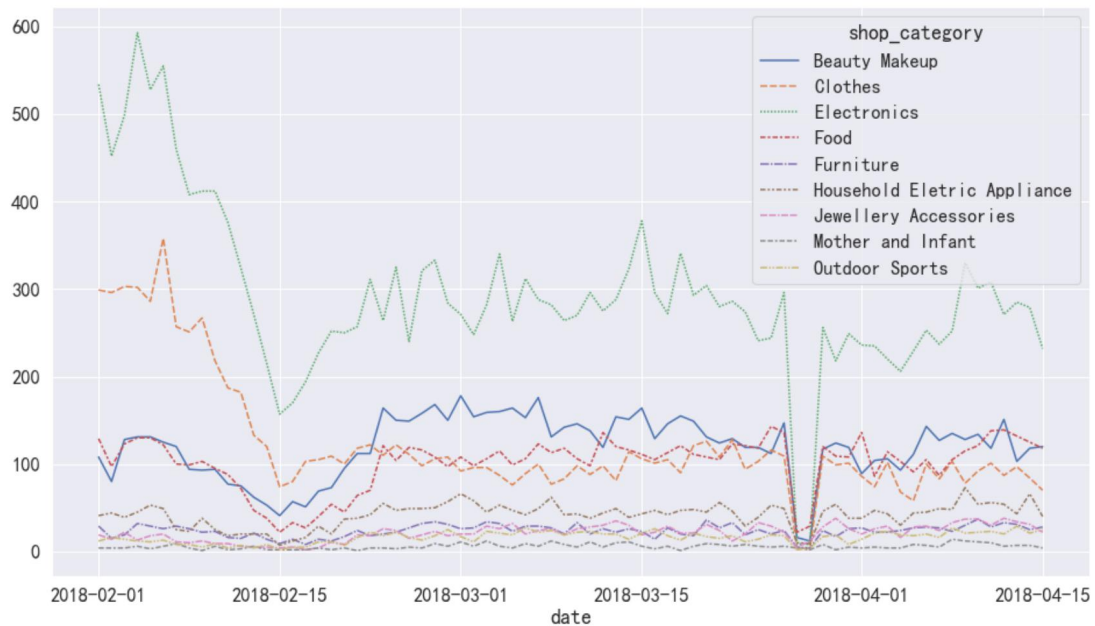


用户访问商店种类



两个半月以来用户访问商店种类前 4 的是电子产品，衣服，美妆，食物，可在这四个方面做好售后服务，打造好口碑

### 3.4 两个半月以来各商店种类用户点击的变化趋势？



分析：

2月15日和3月底的大变化前面分析了，这里不再赘述。

从图中还可看出，电子类商品一直以来热度很高，其次是衣服，美妆，食物也很靠前。可以加大这方面的营销

## 四、建议

- 1、促销活动时间：下午，晚上
- 2、积极完善用户推荐机制，多做活动降低商品价格提升商品质量
- 3、有商品复购行为的用户少，应积极完善商品排名等
- 4、应积极关注电子产品，衣物，美妆，食物这四大类商品，因为这是用户最关注的