

# 淘宝用户行为分析

项目背景：

为进行用户精细化管理，针对用户从浏览到下单的流程进行拆解与分析，找到关键点制定适合的用户运营策略。

分析流程

1. 提出问题
2. 数据理解
3. 数据清洗
4. 构建模型
5. 分析并总结建议

## 一、提出问题

销售额=用户\*产品价格， 所以从用户和产品两个角度进行分析。

- 1.用户从浏览到下单的流失情况，流失最严重的环节是什么？策略是什么？
- 2.用户使用时间模式
- 3.核心付费用户群及其购买最多的产品
- 4.销量最好的产品？

## 二、数据理解：

针对以上四个问题涉及到的指标有：

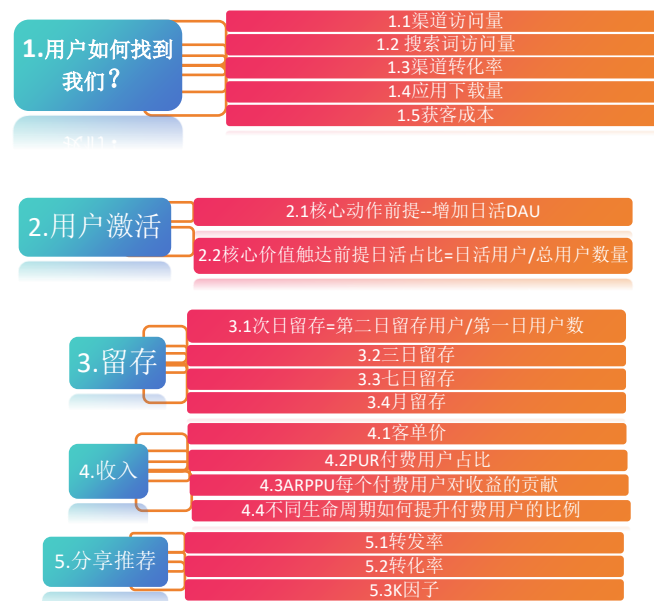
流量指标（uv、pv、pv/uv、跳失率）

用户行为转化指标（点击、加购收藏、下单、下单转化率）

用户时间指标（每日活跃点击量、每时段活跃点击量）

产品销售指标（产品销售排名、产品种类排名、产品累计销售、价值用户（复购率、复购率最高的用户购买的产品））

## AARRR 模型：



数据选取自阿里云天池。本数据集包含了 2017 年 11 月 25 日至 2017 年 12 月 3 日之间，有行为的约一百万随机用户的所有行为（行为包括点击、购买、加购、喜欢）。数据集的每一行表示一条用户行为，由用户 ID、商品 ID、商品类目 ID、行为类型和时间戳组成，并以逗号分隔。数据包括：

| 列名称    | 说明                                      |
|--------|---|
| 用户ID   | 整数类型，序列化后的用户ID                          |
| 商品ID   | 整数类型，序列化后的商品ID                          |
| 商品类目ID | 整数类型，序列化后的商品所属类目ID                      |
| 行为类型   | 字符串，枚举类型，包括('pv', 'buy', 'cart', 'fav') |
| 时间戳    | 行为发生的时间戳                                |

行为分类与解释：

| 行为类型 | 说明            |
|------|---------------|
| pv   | 商品详情页pv，等价于点击 |
| buy  | 商品购买          |
| cart | 将商品加入购物车      |
| fav  | 收藏商品          |

每个维度数据的大小：

| 维度     | 数量          |
|--------|-------------|
| 用户数量   | 987,994     |
| 商品数量   | 4,162,024   |
| 用户数量   | 987,994     |
| 商品类目数量 | 9,439       |
| 所有行为数量 | 100,150,807 |

### 三、数据清洗

由于数据量一亿过大，为了方便分析，选取第 500 万行到 7999999 行数据进行分析。

```
create table new_userbehavior as(SELECT
*
FROM
`userbehavior`
LIMIT 3000000 OFFSET 4999999)
```

时间一致化处理：原数据时间戳使用的是 epoch&unix timestamp 格式，需要转换为标准可读的日期时间形式。在原数据表增加 3 个新字段 datetime、date、hour，把转换好的日期时间放进去。

```
ALTER TABLE new_userbehavior ADD COLUMN datetime TIMESTAMP(0) NULL;
UPDATE new_userbehavior SET datetime=FROM_UNIXTIME(timestamp);
ALTER TABLE new_userbehavior ADD COLUMN date CHAR(10) NULL;
UPDATE new_userbehavior SET date=SUBSTRING(datetime FROM 1 FOR 10);
ALTER TABLE new_userbehavior ADD COLUMN hour CHAR(2) NULL;
UPDATE new_userbehavior SET hour=SUBSTRING(datetime FROM 12 FOR 2);
```

使用 SQL、R、Python 完成数据清洗及预处理：

缺失值：

```
select count(*) AS 缺失值数量
from userbehavior500_800
where user_id is null
or item_id is null
or category_id is null
or `timestamp` is null
;
```

缺失值数量

0

重复值检查：

```
SELECT
    user_id,
    item_id,
    behavior_type,
    `timestamp`,

    count(*) AS 重复值数量

FROM
    userbehavior500_800
GROUP BY
    1,
    2,
    3,
    4
HAVING
    count(*) > 1;
```

| 数据      |         | 信息            |           |       |
|---------|---------|---------------|-----------|-------|
| user_id | item_id | behavior_type | timestamp | 重复值数量 |
| (N/A)   | (N/A)   | (N/A)         | (N/A)     | (N/A) |

异常值检查：

检查日期是否在规定范围内（2017 年 11 月 25 日至 2017 年 12 月 3 日），将不符合规定的删除。

| MAX(date)  | MIN(date)  |
|------------|------------|
| 2020-10-19 | 2017-11-14 |

```
DELETE
FROM
    new_userbehavior
WHERE
    DATE < "2017-11-25"
    OR DATE > "2017-12-03";
```

共删除 136 行

再次验证日期时间的准确性：

```
Select  
MAX(DATE),  
MIN(DATE)  
FROM new_userbehavior;
```

| MAX(DATE)  | MIN(DATE)  |
|------------|------------|
| 2017-12-03 | 2017-11-25 |

都符合要求，数据处理完毕。

## 四、建模分析

### 4.1 流量分析：

#### 4.1.1 客流量 UV，浏览量 PV，用户平均浏览量 PV/UV，跳失率：

```
WITH click_only_user AS (  
    SELECT DISTINCT  
        u1.user_id  
    FROM  
        new_userbehavior u1  
    WHERE  
        NOT EXISTS (  
            SELECT 1 FROM userbehavior500_800 u2 WHERE u1.user_id = u2.user_id AND  
            u2.behavior_type IN ('fav', 'cart', 'buy'))  
    ) SELECT  
        COUNT(DISTINCT u1.user_id) AS UV,  
        COUNT(CASE WHEN u1.behavior_type = 'pv' THEN 1 END) AS PV,  
        CASE  
            WHEN COUNT(DISTINCT u1.user_id) = 0 THEN  
                0  
            ELSE
```

```

COUNT(CASE WHEN u1.behavior_type = 'pv' THEN 1 END) / COUNT(DISTINCT
u1.user_id)
END AS 'PV/UV',

COUNT(DISTINCT c.user_id) AS '仅点击用户数',

CASE
WHEN COUNT(DISTINCT u1.user_id) = 0 THEN
0
ELSE
COUNT(DISTINCT c.user_id) / COUNT(DISTINCT u1.user_id)

END AS '跳失率'

FROM
new_userbehavior u1
LEFT JOIN click_only_user c ON u1.user_id = c.user_id;

```

| UV    | PV      | PV/UV   | 仅点击用户数 | 跳失率    |
|-------|---------|---------|--------|--------|
| 29233 | 2683659 | 91.8024 | 1628   | 0.0557 |

用户访问量是 29233 人，平均每个用户浏览 92 次网页，仅点击网页的用户是 1628 个，跳失率为 5.57%，大多数用户浏览产品页面后会进行加购收藏购买等操作。问题出在用户渠道上，可以通过社交媒体平台引流与广告投入，扩大用户量。

#### 4.1.独立用户漏斗分析

```

WITH c AS (SELECT COUNT(DISTINCT user_id) AS total_user FROM new_userbehavior)
SELECT
CASE
WHEN
behavior_type IN ('cart', 'fav') THEN

'意向行为'

ELSE
behavior_type

END AS '行为类型',

COUNT(DISTINCT user_id) AS 用户数,

```

```
CONCAT(ROUND(COUNT(DISTINCT user_id) * 100.0 / c.total_user, 2),
```

```
"%") AS 每种行为用户占比
```

```
FROM
```

```
new_userbehavior,
```

```
c
```

```
GROUP BY
```

```
CASE
```

```
WHEN behavior_type IN ('cart', 'fav') THEN
```

```
'意向行为'
```

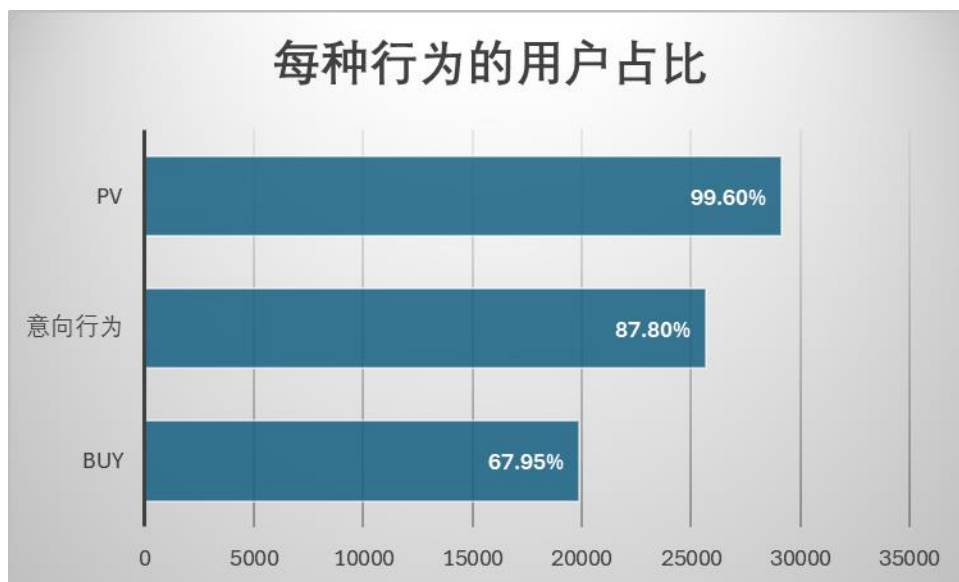
```
ELSE
```

```
behavior_type
```

```
END,
```

```
c.total_user;
```

| 行为类型 | 用户数   | 每种行为用户占比 |
|------|-------|----------|
| buy  | 19865 | 67.95%   |
| pv   | 29116 | 99.60%   |
| 意向行为 | 25666 | 87.80%   |



可以看到大多数用户点进网页后会浏览商品，且大多数用户浏览完商品后会加购或者收藏，比例大约 88%，但是从浏览到购买的转化率仅 68%，77%的用户从加购到购买进行转化。应该加强加购到购买这一环节的优化。

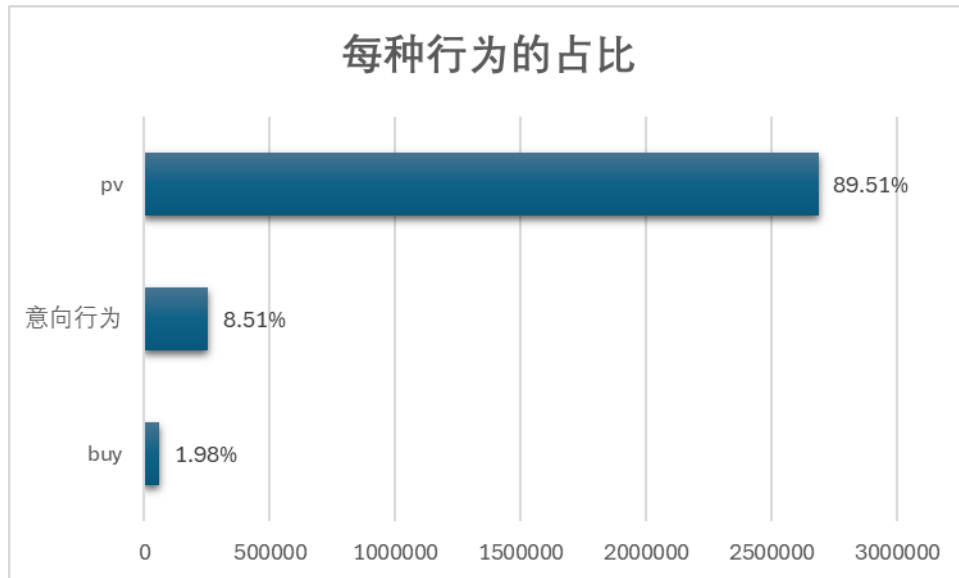


### 4.3 用户行为漏斗分析

接下来看一下用户行为的漏斗转化情况观察用户在浏览-加购&收藏-购买这一流程中的卡点。

```
WITH c AS (SELECT COUNT(*) AS total_user FROM new_userbehavior) SELECT
CASE
  WHEN
    behavior_type IN ('cart', 'fav') THEN
    '意向行为'
  ELSE
    behavior_type
END AS '行为类型',
COUNT(*) AS 行为数,
CONCAT(ROUND(COUNT(*) * 100.0 / c.total_user, 2), "%") AS '每种行为用户占比'
FROM
  new_userbehavior,
  c
GROUP BY
CASE
  WHEN behavior_type IN ('cart', 'fav') THEN
    '意向行为'
  ELSE
    behavior_type
END,
c.total_user;
```

| 行为类型 | 行为数     | 每种行为用户占比 |
|------|---------|----------|
| pv   | 2685348 | 89.51%   |
| buy  | 59329   | 1.98%    |
| 意向行为 | 255323  | 8.51%    |



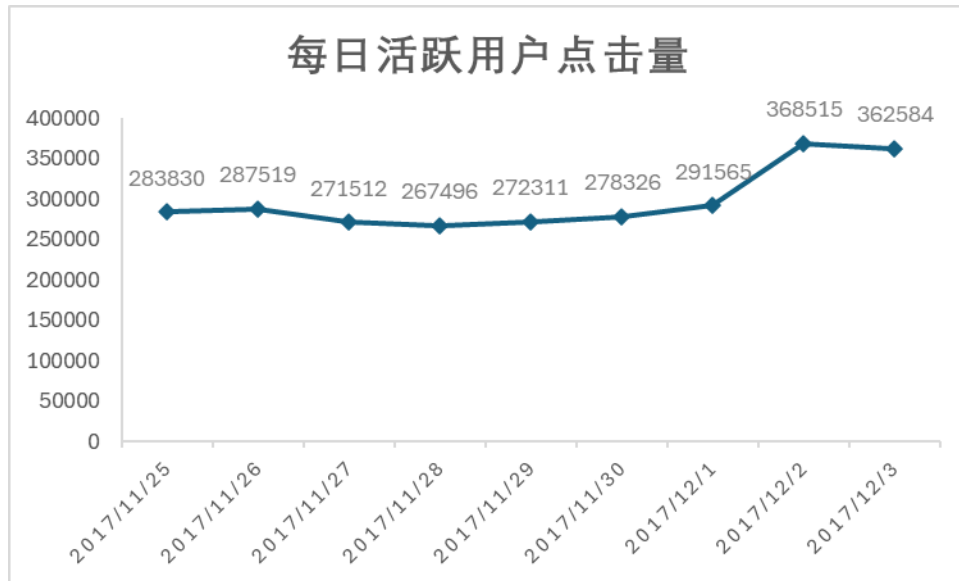
可以发现从行为漏斗上观察浏览到购买行为的转换率很低仅 2.2%，从浏览到加购的转化约 9.5%。所以关键环节在加购和收藏，加强浏览到收藏加购的补贴激励，可以通过数据埋点进一步分析商品详情面的浏览时长，是否因为商品页面不够简洁，吸引人而退出页面。通过 A/B 测试分析是否是加购按键图表过于隐蔽导致。

#### 4.4 用户行为模式分析

解决问题：在研究的时间段里找出用户最活跃的日子以及每天活跃的时间段，了解用户的行为时间模式。

##### 4.4.1 每日活跃点击量

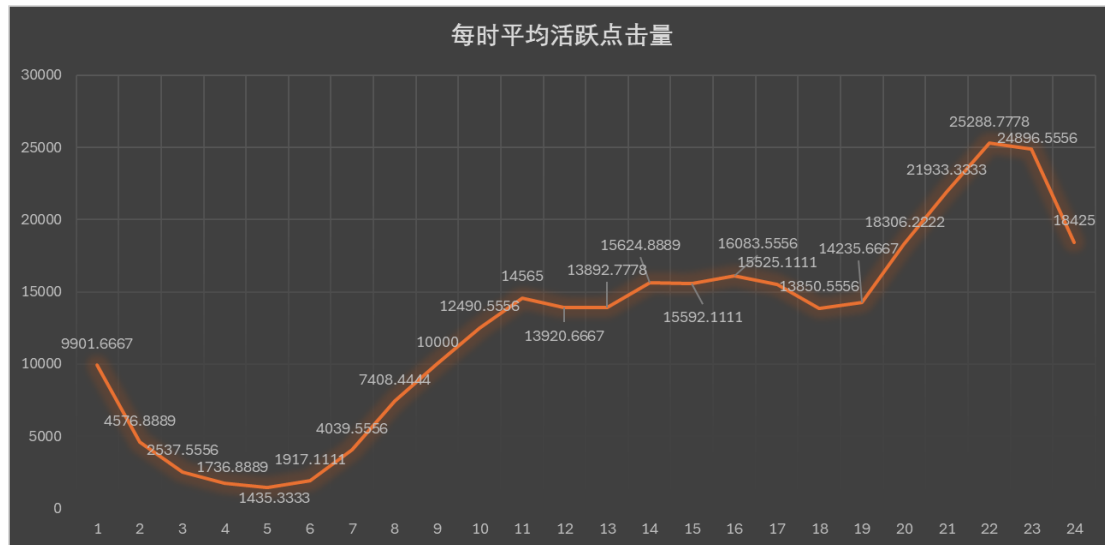
```
SELECT date,COUNT(*) as pv
FROM new_userbehavior
WHERE behavior_type ='pv'
GROUP BY date
ORDER BY date;
```



可以发现 25 号到 1 号的点击量都属于比较平稳的状态, 在 12 月 2 日的点击量突然增加约 7 万人次。推测是上班族因工作逛淘宝的时间少, 而(12 月 2 日-12 月 3 日)是周末, 有充足的精力和有较多空闲时间访问淘宝。因此运营可以将活动集中在周末进行。

#### 4.4.2 每时段平均活跃用户量:

```
SELECT
    `hour`,
    COUNT(*)/9
FROM new_userbehavior
WHERE behavior_type = 'pv'
GROUP BY `hour`
ORDER BY `hour`;
```



每天早上点击量不断增长,从6点到11点时段是一个高峰到11点达到峰值后下降在下午4点又达到小高峰,晚上7点到9点是个高峰,10点左右达到峰值。晚上是比较集中浏览网页的时段,某些品类促销和推荐活动可以在这个时间开展。

#### 4.4.3 次日留存、三日留存率、七日留存率

```

SELECT
    a.date AS '基准日' ,

    COUNT(DISTINCT a.user_id) AS 基准日活跃用户数,

    COUNT(DISTINCT b.user_id) AS 次日留存用户数,

    ROUND(
        COUNT(DISTINCT b.user_id) * 100.0 / COUNT(DISTINCT a.user_id),
        2
    ) AS '次日留存率(%)'

FROM
    (SELECT DISTINCT user_id, date FROM new_userbehavior) AS a
LEFT JOIN
    (SELECT DISTINCT user_id, date FROM new_userbehavior) AS b
    ON a.user_id = b.user_id
    AND b.date = DATE_ADD(a.date, INTERVAL 1 DAY)
GROUP BY
    a.date
ORDER BY
    a.date;

```

| 基准日        | 基准日活跃用户数 | 次日留存用户数 | 次日留存率(%) |
|------------|----------|---------|----------|
| Null       | 1        | 0       | 0.00     |
| 2017-11-25 | 21025    | 16643   | 79.16    |
| 2017-11-26 | 21313    | 16662   | 78.18    |
| 2017-11-27 | 21083    | 16617   | 78.82    |
| 2017-11-28 | 21109    | 16742   | 79.31    |
| 2017-11-29 | 21315    | 16992   | 79.72    |
| 2017-11-30 | 21630    | 17244   | 79.72    |
| 2017-12-01 | 21939    | 21593   | 98.42    |

从留存率来看用户整体的留存情况较好,说明周一浏览的用户在接下来一周会持续对该产品关注可能是加购物车或者是下单,用户粘性较好。

## 4.5 产品销量分析

问题:

- 1: 什么产品以及产品类目的购买率最高,找出最受欢迎的产品,优化产品销售。
- 2: 哪些用户购买次数最多,找出最核心的付费用户群,并且统计出这些用户购买的产品以及类目,针对这些用户的购买偏好推送个性化的产品销售方案。

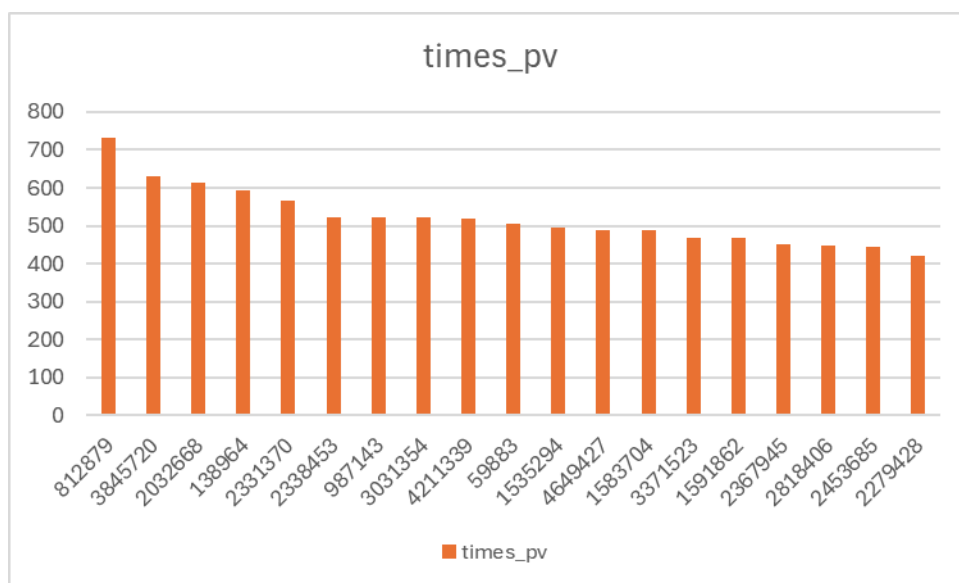
### 4.5.1 产品排名

```
SELECT
    item_id,
    count(user_id) AS times_pv
FROM
    new_userbehavior
WHERE
    behavior_type = 'pv'
GROUP BY
    item_id
ORDER BY
```

```
times_pv DESC
LIMIT 20;
```

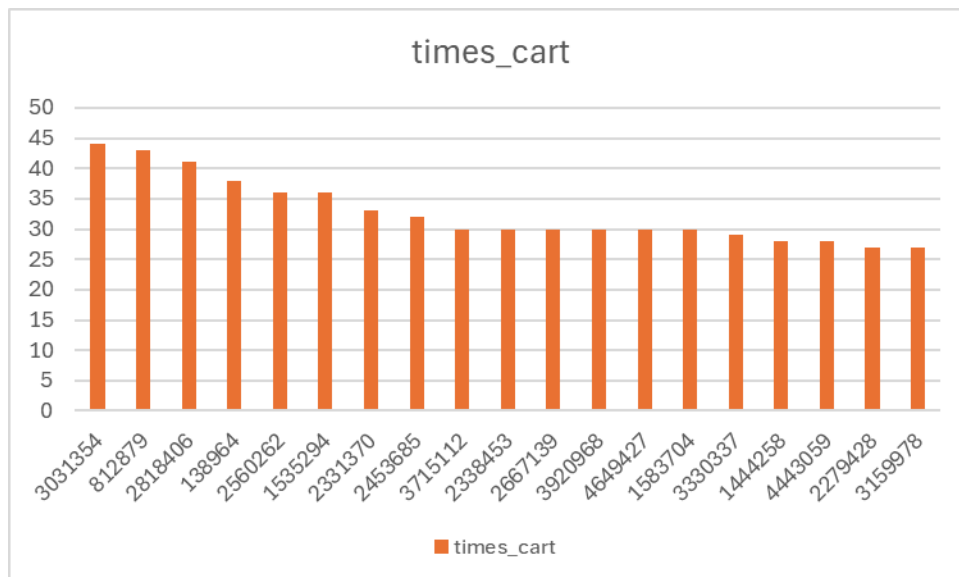
| item_id | times_pv |
|---------|----------|
| 812879  | 914      |
| 3845720 | 733      |
| 2032668 | 632      |
| 138964  | 615      |
| 2331370 | 593      |
| 2338453 | 568      |
| 987143  | 524      |
| 3031354 | 524      |
| 4211339 | 523      |
| 59883   | 520      |
| 1535294 | 505      |
| 4649427 | 494      |
| 1583704 | 490      |
| 3371523 | 489      |
| 1591862 | 469      |
| 2367945 | 469      |
| 2818406 | 451      |
| 2453685 | 448      |
| 2279428 | 446      |
| 4443059 | 420      |

浏览次数前 20 的产品 id:



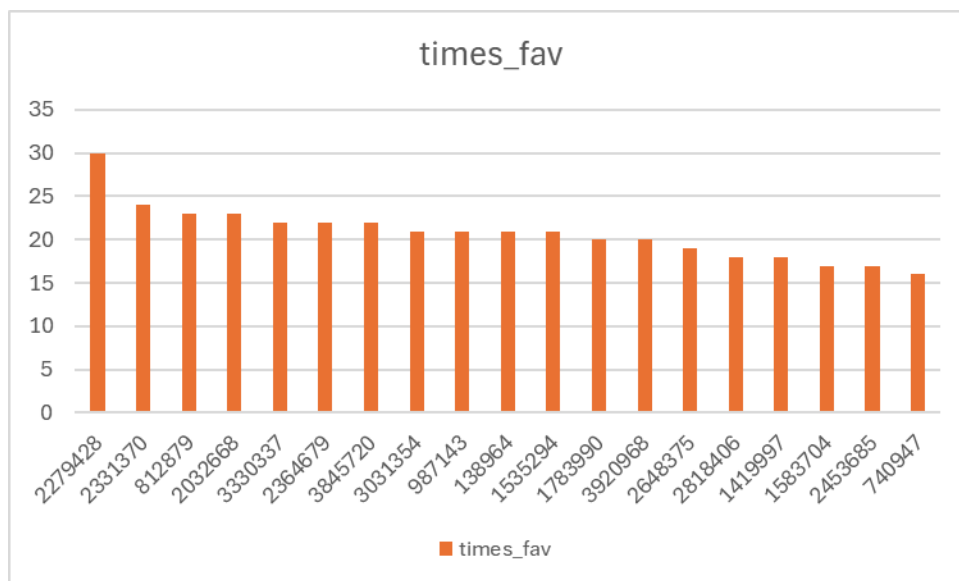
加入购物车前 20 的产品 id:

| item_id | times_car |
|---------|-----------|
| 3031354 | 46        |
| 812879  | 44        |
| 2818406 | 43        |
| 138964  | 41        |
| 2560262 | 38        |
| 1535294 | 36        |
| 2331370 | 36        |
| 2453685 | 33        |
| 3715112 | 32        |
| 2338453 | 30        |
| 2667139 | 30        |
| 3920968 | 30        |
| 4649427 | 30        |
| 1583704 | 30        |
| 3330337 | 30        |
| 1444258 | 29        |
| 4443059 | 28        |
| 2279428 | 28        |
| 3159978 | 27        |
| 4211339 | 27        |



收藏的前 20 名产品 id:

| item_id | times_cart |
|---------|------------|
| 2279428 | 34         |
| 2331370 | 30         |
| 812879  | 24         |
| 2032668 | 23         |
| 3330337 | 23         |
| 2364679 | 22         |
| 3845720 | 22         |
| 3031354 | 22         |
| 987143  | 21         |
| 138964  | 21         |
| 1535294 | 21         |
| 1783990 | 21         |
| 3920968 | 20         |
| 2648375 | 20         |
| 2818406 | 19         |
| 1419997 | 18         |
| 1583704 | 18         |
| 2453685 | 17         |
| 740947  | 17         |
| 1343119 | 16         |

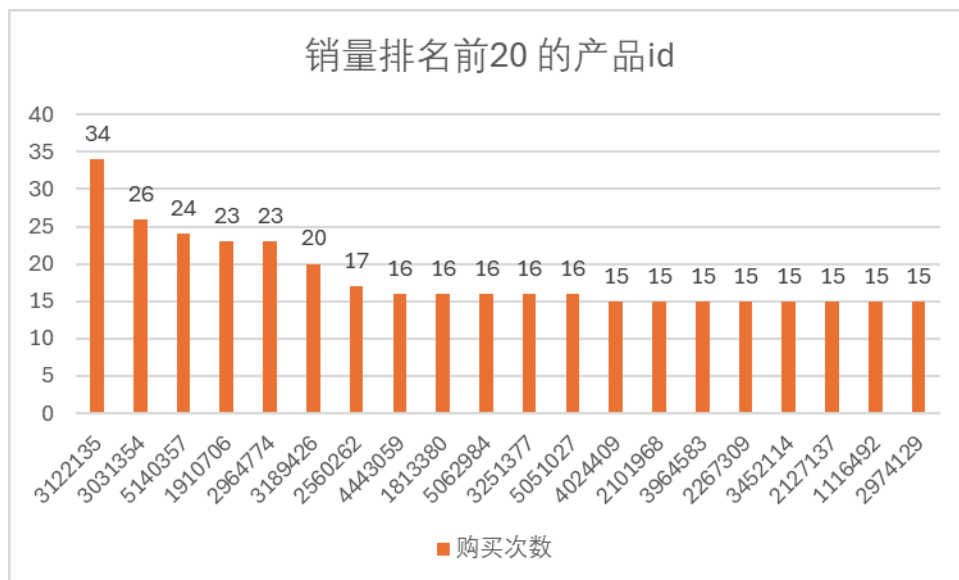


购买量前 20 的产品 id:

| item_id | 购买次数 |
|---------|------|
| 3122135 | 34   |
| 3031354 | 26   |
| 5140357 | 24   |
| 1910706 | 23   |
| 2964774 | 23   |



|         |    |
|---------|----|
| 3189426 | 20 |
| 2560262 | 17 |
| 4443059 | 16 |
| 1813380 | 16 |
| 5062984 | 16 |
| 3251377 | 16 |
| 5051027 | 16 |
| 4024409 | 15 |
| 2101968 | 15 |
| 3964583 | 15 |
| 2267309 | 15 |
| 3452114 | 15 |
| 2127137 | 15 |
| 1116492 | 15 |
| 2974129 | 15 |

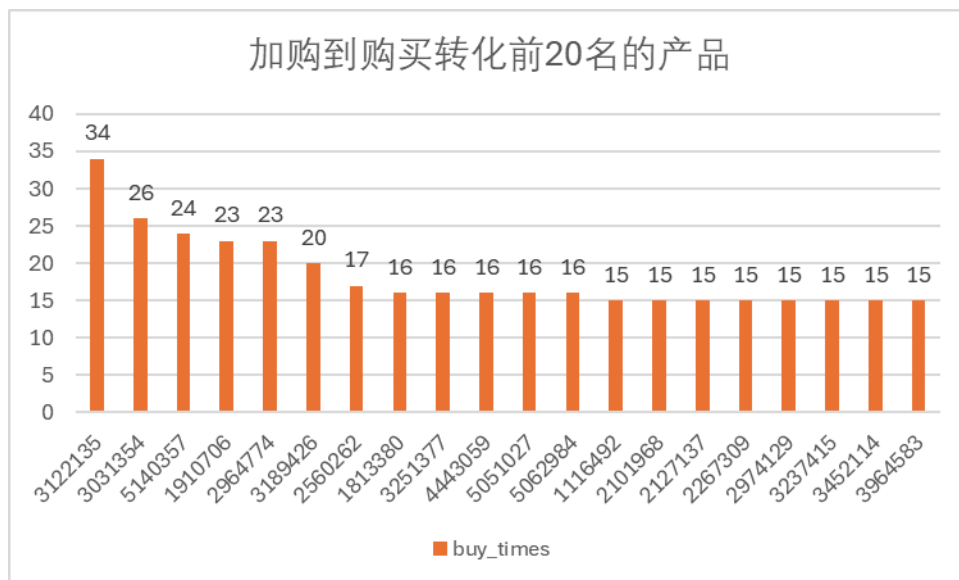


如果我们将所有加购并最后购买的产品导出，提取销量前 20 名的产品，就能得到下面这张

表：

| item_id | buy_times |
|---------|-----------|
| 3122135 | 34        |
| 3031354 | 26        |
| 5140357 | 24        |
| 2964774 | 23        |
| 1910706 | 23        |
| 3189426 | 20        |

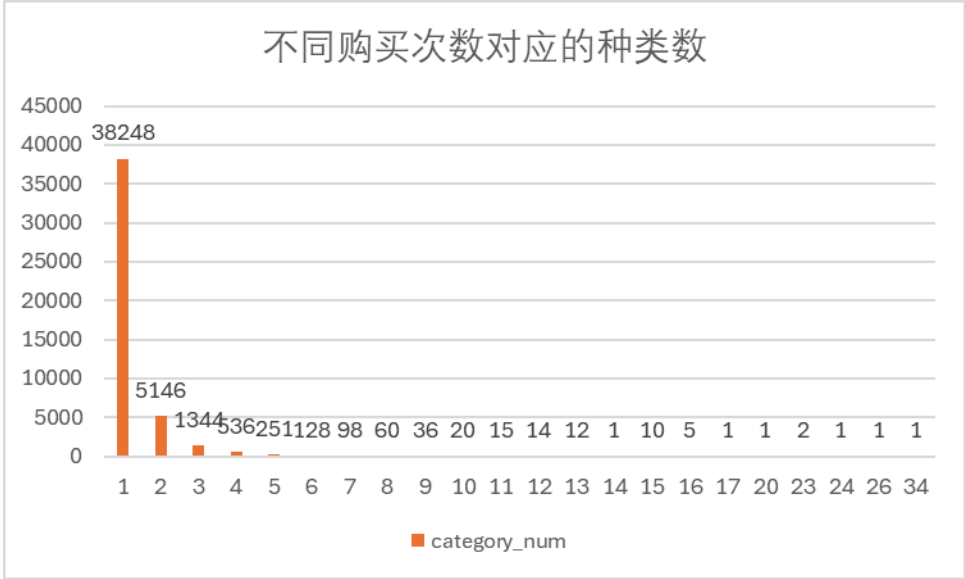
|         |    |
|---------|----|
| 2560262 | 17 |
| 1813380 | 16 |
| 4443059 | 16 |
| 5051027 | 16 |
| 5062984 | 16 |
| 3251377 | 16 |
| 2974129 | 15 |
| 3237415 | 15 |
| 3452114 | 15 |
| 2101968 | 15 |
| 2127137 | 15 |
| 2267309 | 15 |
| 1116492 | 15 |
| 3964583 | 15 |



3122135 没有出现在加购的前 20 名中但在转化为成交额的产品中排名第一。说明很多用户直接购买了这款产品省去了加入购物车的环节,可以进一步分析促销活动或者是产品详情页的影响。2560262 被大多数用户加入购物车但是仅排在购买量第二,可以对比 3122135 的产品页及促销活动的区别。大多数被用户加入购物车的产品最后都购买了。浏览量排名前三的产品中仅 812879 是出现在购物车前 3 的区间。为了进一步研究电商销售产品的规律,对不同用户购买次数对应的产品种类数进行分析。

4.5.2 不同购买次数对应的产品种类数：

| buy_times | category_num |
|-----------|--------------|
| 34        | 1            |
| 26        | 1            |
| 24        | 1            |
| 23        | 2            |
| 20        | 1            |
| 17        | 1            |
| 16        | 5            |
| 15        | 10           |
| 14        | 1            |
| 13        | 12           |
| 12        | 14           |
| 11        | 15           |
| 10        | 20           |
| 9         | 36           |
| 8         | 60           |
| 7         | 98           |
| 6         | 128          |
| 5         | 251          |
| 4         | 536          |
| 3         | 1344         |
| 2         | 5146         |
| 1         | 38248        |
| 0         | 759899       |



可以发现 有 38248 种产品被购买一次，产生销售额的产品种类是 45931 种，83% 的产品种类

都对总销售额有贡献，说明能够贡献销售额的产品并非集中于爆款，而是来源于产品种类的积累。

### 3) 整体复购率

不同购买次数下的用户数，计算购买两次及以上的用户数量占整体的比例。

```
SELECT
    a.buy_times,
    COUNT(user_id) AS '人数'
FROM
    (SELECT user_id, COUNT(behavior_type) AS buy_times FROM new_userbehavior
    WHERE behavior_type = 'buy' GROUP BY user_id) AS a
GROUP BY
    a.buy_times
ORDER BY
    a.buy_times;
```

| buy_times | 人数   |
|-----------|------|
| 1         | 6787 |
| 2         | 4586 |
| 3         | 3004 |
| 4         | 1857 |
| 5         | 1204 |
| 6         | 785  |
| 7         | 495  |
| 8         | 327  |
| 9         | 219  |
| 10        | 174  |
| 11        | 108  |
| 12        | 84   |
| 13        | 46   |
| 14        | 40   |
| 15        | 27   |
| 16        | 18   |
| 17        | 25   |
| 18        | 16   |
| 19        | 8    |
| 20        | 8    |
| 21        | 4    |
| 22        | 5    |
| 23        | 9    |

|    |   |
|----|---|
| 24 | 6 |
| 25 | 1 |
| 26 | 4 |
| 27 | 2 |
| 28 | 3 |
| 30 | 2 |
| 32 | 2 |
| 33 | 1 |
| 35 | 1 |
| 36 | 2 |
| 39 | 1 |
| 40 | 1 |
| 43 | 1 |
| 45 | 1 |
| 93 | 1 |



平台整体的复购率是 88%，大部分用户会再次使用淘宝平台下单，用户的忠诚度较高，接下来需要分析什么类型的产品会让用户重复购买,因此我们对平台复购率最高的用户所购买的产品进一步分析。

4.5.3 平台复购率最高的用户喜爱的前 20 个产品：

| user_id | item_id | 购买次数 |
|---------|---------|------|
| 337305  | 4675518 | 3    |
| 337305  | 4403610 | 3    |
| 337305  | 2288434 | 3    |
| 337305  | 2930515 | 3    |

|        |         |   |
|--------|---------|---|
| 337305 | 1105585 | 2 |
| 337305 | 1911827 | 2 |
| 337305 | 3156760 | 2 |
| 337305 | 4361112 | 2 |
| 337305 | 2405542 | 2 |
| 337305 | 2957350 | 2 |
| 337305 | 4844191 | 2 |
| 337305 | 674807  | 2 |
| 337305 | 2053331 | 2 |
| 337305 | 2481271 | 2 |
| 337305 | 5094401 | 2 |
| 337305 | 3836655 | 2 |
| 337305 | 451490  | 2 |
| 337305 | 3679545 | 2 |
| 337305 | 987361  | 2 |
| 337305 | 2049903 | 2 |

由于本数据集没有产品具体信息表，无法进一步分析热门产品特征。但这些商品没有出现在销量前 20 的产品列表中，也能说平台复购不是以爆品为导向的。

#### 4.5.4 复购率最高的产品

```
SELECT
    item_id,
    user_id,
    COUNT(*) AS buy_times
FROM
    new_userbehavior
WHERE
    behavior_type = 'buy'
GROUP BY
    item_id,
    user_id
HAVING
    buy_times >= 2
ORDER BY
    buy_times DESC
LIMIT 20;
```

| item_id | user_id | buy_times |
|---------|---------|-----------|
| 616792  | 381745  | 10        |
| 1905432 | 340119  | 7         |
| 4015112 | 3735    | 7         |
| 5140357 | 426328  | 7         |
| 5073358 | 339658  | 6         |
| 3643327 | 340119  | 6         |
| 2195809 | 357565  | 6         |
| 4587613 | 381745  | 6         |

重复购买次数最高的产品是 616792，重复购买率最高的用户是 381745，可以在后续针对产品进行分析，提出产品详情页等改进建议。

重复购买率最高的用户，购买的其他产品如下，可以结合产品信息进一步分析原因，提升产品的复购率。

```
SELECT
    item_id,
    user_id,
    count(*) AS buy_times
FROM
    new_userbehavior
WHERE
    behavior_type = 'buy'
    AND user_id = 381745
GROUP BY
    item_id,
    user_id
ORDER BY
    buy_times DESC LIMIT 20;
```

| item_id | user_id | buy_times |
|---------|---------|-----------|
| 616792  | 381745  | 10        |
| 4587613 | 381745  | 6         |
| 2094821 | 381745  | 3         |
| 3458317 | 381745  | 3         |
| 4956769 | 381745  | 2         |
| 2757020 | 381745  | 2         |
| 4113691 | 381745  | 1         |
| 227510  | 381745  | 1         |

## 五、总结与建议

### 5.1 获取用户

从前面的分析可以发现, 用户进入淘宝的商品页面后从浏览到加购物车再到下单的转化总体是顺畅的, 所以获取用户, 和用户增量是关键, 可以通过加强海外社媒引流的方式增加新用户。

### 5.2 激活用户

从用户行为的漏斗分析中可以发现浏览到加入购物车的转化率不高, 这一环节可以成为激活用户的关键环节, 加强浏览到收藏加购的补贴激励, 可以通过数据埋点进一步分析商品详情面的浏览时长, 是否因为商品页面不够简洁, 或者是功能按键不够明显导致, 可以设计 A/B 测试进一步分析。

从用户行为模式上分析可以从在周五或者周六晚间 7-9 点投放活动津贴, 周五投放量大于周六, 周三至周六的中午 11-14 时段进行活动宣传, 例如直播活动抢购物券等方式激活用户的加购操作。并计算每周用户从加入购物车到使用购物券完成消费的人数占总人数的比例。了



解完成产品核心价值的用户占比。分析这一占比与平台复购率以及产品复购率的变化是否有关联，为后续进一步提升产品的复购率打基础； 计算用户消耗的津贴数占总发放津贴的比例了解津贴使用的情况，对后续补贴活动进行改进。

### 5.3 留存用户

根据次日留存率、七日留存率可以看到平台整体的用户留存情况较好,继续保持的策略如下:

- 1.按照使用频率和购买次数积攒积分，每天上线点击量达到某个数值即可自动领取积分，到月末换取购物礼券。
- 2.对于年购买次数和金额达到规定量的客户推出 VIP 服务，购买次数同比上升之后相应福利也上升，利用这种方法提高高价值用户的留存率和对平台的忠诚度。

### 5.4 增加收入

从用户行为分析漏洞可以发现从加购到购买的转化率是 23%，独立用户 68%会从浏览到下单说明现存用户对平台总体满意。83%的产品在创造销售额说明产品种类对销售量影响很大。复购率达到 88%，复购次数是 93 次，我们可以通过复购率、购买金额（本次数据集没有提供）等来确定价值用户，通过分析找出价值用户的购买偏好，产品和类目等，给价值用户制定个性化的产品推荐，从而提高用户体验和电商平台销售情况。

综上，需要不断丰富产品种类外加精准推荐产品并优化产品的详情信息。对于优化产品种类可以从几个角度出发：

#### 1.用户分层与需求挖掘

对现有用户按消费频率、偏好、场景等分层（如“高频刚需用户”“尝鲜型用户”“价格敏感型用户”），分析不同群体的潜在需求。

收集用户反馈：通过问卷、客服咨询、评论区留言等，提炼“用户提到但未被满足”的需求。

## 2. 场景化延伸

针对用户使用产品的不同场景，开发适配场景的细分品类。

比如咖啡品牌现有“家庭装咖啡豆”，可延伸出“办公室挂耳咖啡”（便捷场景）、“旅行迷你装”（外出场景）、“低糖冷萃液”（控糖场景）。

## 3. 基于产品线协同：构建“互补型产品矩阵”

丰富种类时需考虑与现有产品的协同性，形成“1+1>2”的效果，降低用户决策成本，提升复购。

横向扩展：同一品类的差异化延伸

在现有核心品类中，通过“功能、设计、规格、价格”等维度细分，覆盖更多用户偏好。

例：护肤品品牌核心产品是“保湿面霜”，横向扩展可包括：

功能差异：添加抗衰成分的“保湿抗衰面霜”、针对敏感肌的“无添加保湿霜”；

规格差异：50g 家庭装、15g 便携装、100g 大容量装；

价格差异：高端线（主打成分）、平价线（性价比）。

纵向扩展：上下游或关联品类延伸

围绕用户“购买 - 使用 - 复购”的全链路，补充相关联的产品，形成消费闭环。

例：运动品牌核心产品是“跑鞋”，纵向扩展可包括：

上游：运动袜、鞋垫（提升跑鞋使用体验）；

下游：运动背包、运动水壶（满足运动场景配套需求）；

关联：运动后修复霜、肌肉贴（解决运动后需求）。

## 4. 基于市场趋势：捕捉“新兴需求与行业风口”

结合行业动态、技术进步或社会趋势，提前布局潜力品类，抢占市场先机。

跟踪行业数据与竞品动态

通过电商平台（如京东）的“热销榜”“新品榜”，观察用户近期偏好（如“环保材质”“智能化”“国潮设计”）；

分析竞品的产品结构，寻找其未覆盖的细分市场（如竞品主打“成人服饰”，可切入“亲子装”细分赛道）。

按照趋势丰富品类：如“健康化”趋势下，食品品牌可新增“低糖、低脂、无添加”产品线；“懒人经济”趋势下，家电品牌可开发“自动清洁、一键操作”的智能小家电。如新材料（石墨烯、可降解材料）、新技术（AI 算法、物联网）的应用，催生新产品（如石墨烯保暖内衣、AI 智能测温水杯）。

## 5.5 推荐

由于数据集没有用户的推荐数据。在此提出通用方案：

- 1.产品在购买的时候提供拼团服务，让用户主动推荐给其他人；
- 2.每当推出新功能，比如前面提到的一键下单，让体验过的用户转发和分享领取优惠券，快速实现新功能推广；
- 3.当用户使用优惠券购物或者通过某种行为积分购物之后提供朋友圈打卡功能，分享给好友，实现传播功能。

在实行以上方案之后需要关注转发率、转化率、通过用户分享链接点击购买的用户比例以及 K 因子来检测提出方案的有效性。