

《淘宝展示广告点击率数据集分析》

- 分析目的：作为一个完整的数据分析实践，旨在系统性掌握从业务理解、数据处理、多维度分析到商业洞察的全流程能力。
- 数据来源：阿里天池
- 分析工具：Python (Pandas, Matplotlib)
- 报告日期：2025. 12. 14~12. 20
- 目录：
 1. 数据集信息
 - (1) 数据集来源
 - (2) 数据集简介
 2. 数据处理结果
 - (1) 曝光量
 - (2) 点击率：
 - ① 总点击率
 - ② 日点击率
 - ③ 各时段点击率
 - ④ 价格区间点击率
 - (3) 用户画像
根据特定标签分析
 - ① 性别
 - ② 是否为大学生
 - ③ 购物深度
 - ④ 年龄层次
 3. 数据分析结论

一、数据集信息

1. 数据集来源

阿里天池的淘宝展示广告点击率预估数据集，是阿里巴巴提供的一个淘宝展示广告点击率预估数据集。

2. 数据集简介

数据名称	说明	属性
raw_sample	原始的样本骨架	用户 ID, 广告 ID, 时间, 资源位, 是否点击
ad_feature	广告的基本信息	广告 ID, 广告计划 ID, 类目 ID, 品牌 ID
user_profile	用户的基本信息	用户 ID, 年龄层, 性别等

二、数据处理结果

1. 曝光量

26557961 次

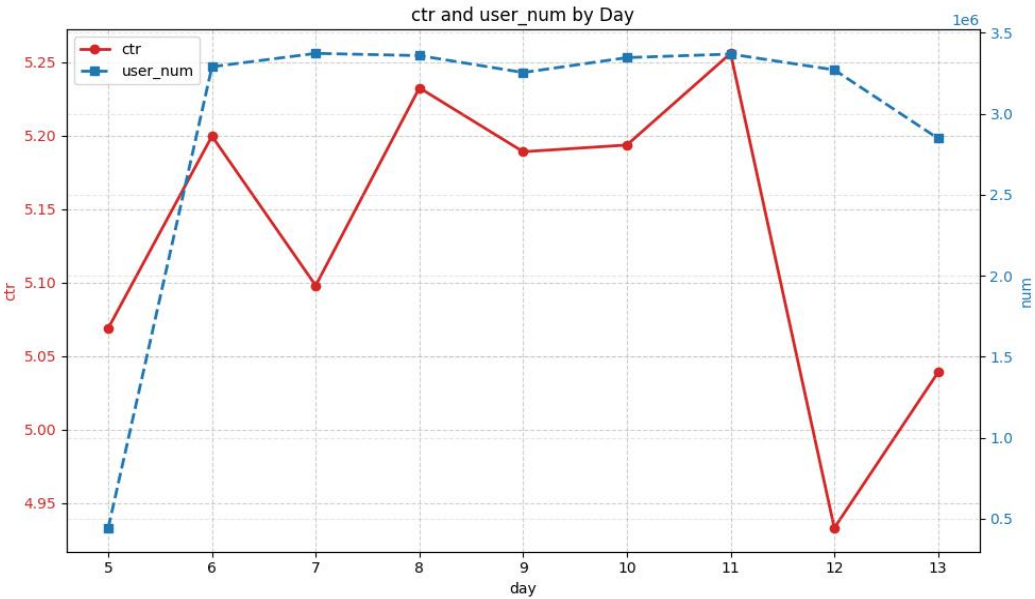
2. 点击率

(1) 总点击率

总点击数: 1366056

总点击率: 5.143678%

(2) 日点击率



日期	5 号	6 号	7 号	8 号	9 号
浏览量 (次)	443443	3290746	3373322	3359443	3255134
点击率 (%)	5.068972	5.199581	5.097972	5.232385	5.188972
星期	五	六	七	一	二
日期	10 号	11 号	12 号	13 号	
浏览量 (次)	3347045	3368931	3271392	2848505	
点击率 (%)	5.193536	5.255851	4.933191	5.038959	
星期	三	四	五	六	

● 浏览量：

最大值： 3373322 次（7 号） 最小值 443443 次（5 号） 极差 2929879 次

● 点击率：

最大值： 5.2559% 最小值 4.9332% 极差 0.3227%

● 分析小结

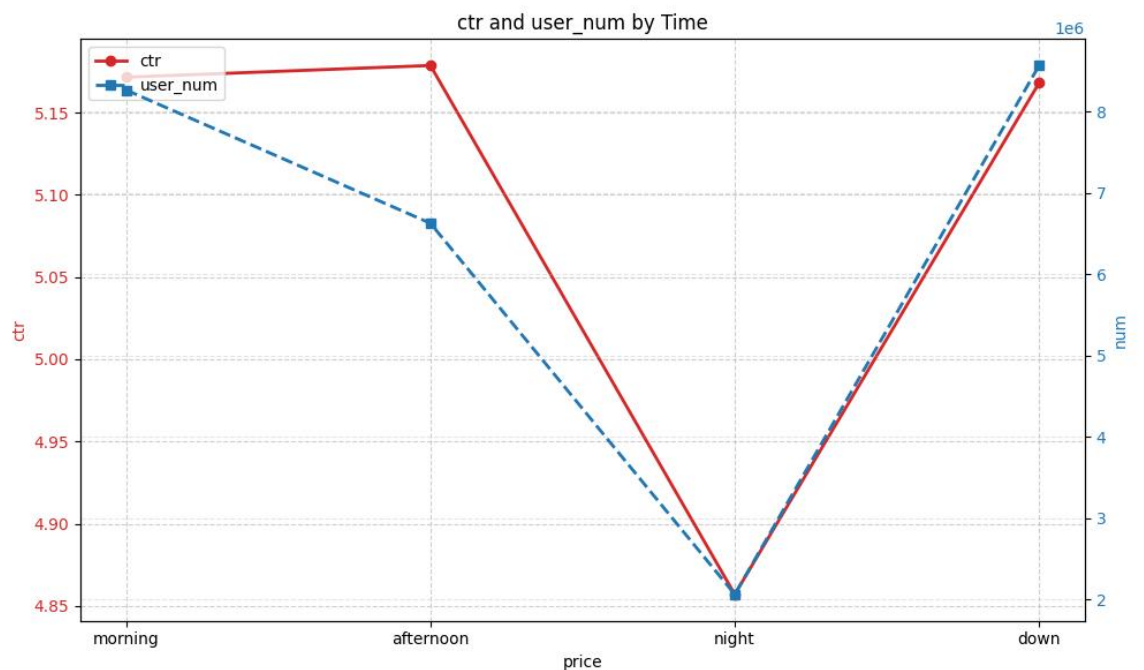
在去掉记录不全的五号和十三号后进行分析，发现浏览量波动较小，同时周中的点击率相对周五六七要高一些，可以在周中投入更多预算同时根据实时数据调整周末的预算安排。

● 其他信息

5 号记录的时间范围：18~23

13 号记录的时间范围：0~15

(3) 各时段点击率



时间	早上	中午	晚上	凌晨
浏览量（次）	8264815	6625314	2060127	8566243
点击率（%）	5.1715	5.1786	4.8571	5.1680
范围	(6:12]	(12:18]	(18:24]	(0:6]

● 浏览量

最大值： 8566243 次（0~6） 最小值 2060127 次（18~24）

极差 6506116

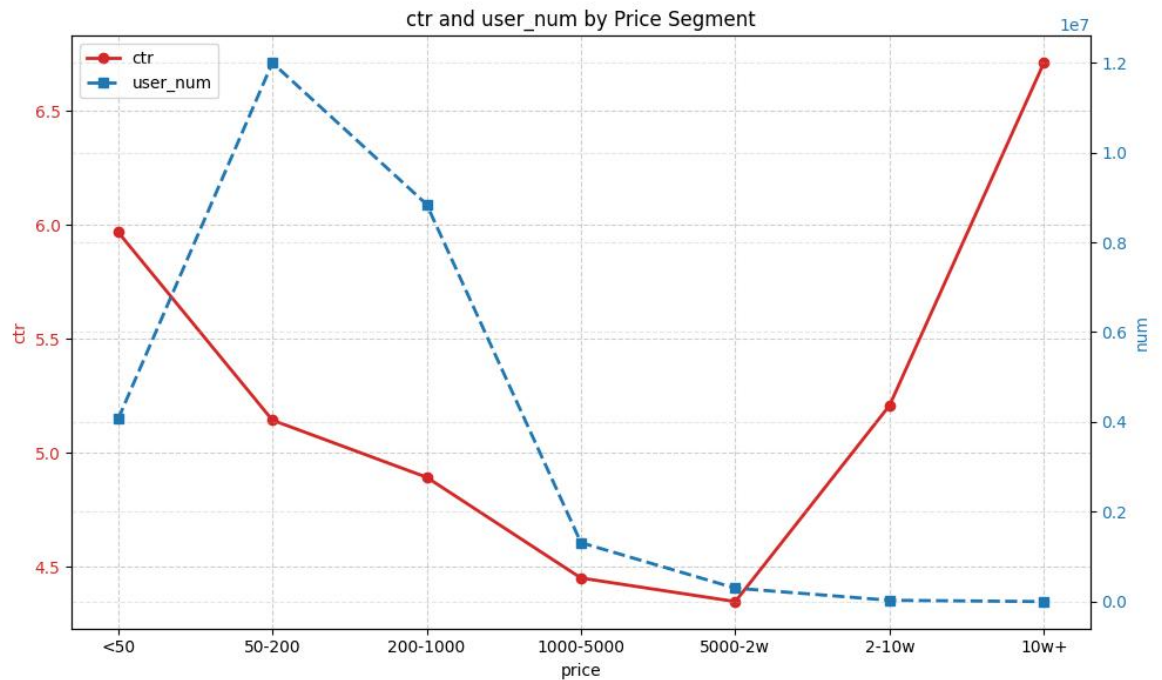
● 点击率

最大值： 5.1786% 最小值 4.8571% 极差 0.3215%

● 分析小结

对数据进行分析可以发现，浏览量和点击率趋势基本一致。早上和凌晨的浏览次数最多，同时凌晨的浏览量高于早上，但是点击率低于早上，而中午是点击率最高的时间，但是浏览量稍低。可以对晚上投放的广告进行创意优化以提升点击率。

(4) 价格区间点击率



价格	<50	50~200	200~1000	1000~5000
浏览量（次）	4065189	11994477	8842789	1314144
点击率（%）	5.9681	5.1437	4.8932	4.4516
价格	5000~2w	2w~10w	10w+	
浏览量（次）	302184	32736	3383	
点击率（%）	4.3483	5.2083	6.7100	

● 浏览量

最大值： 11994477（50~200） 最小值 3383（10w+） 极差 11991094

● 点击率

最大值： 6.7100% 最小值 4.3483% 极差 2.3617%

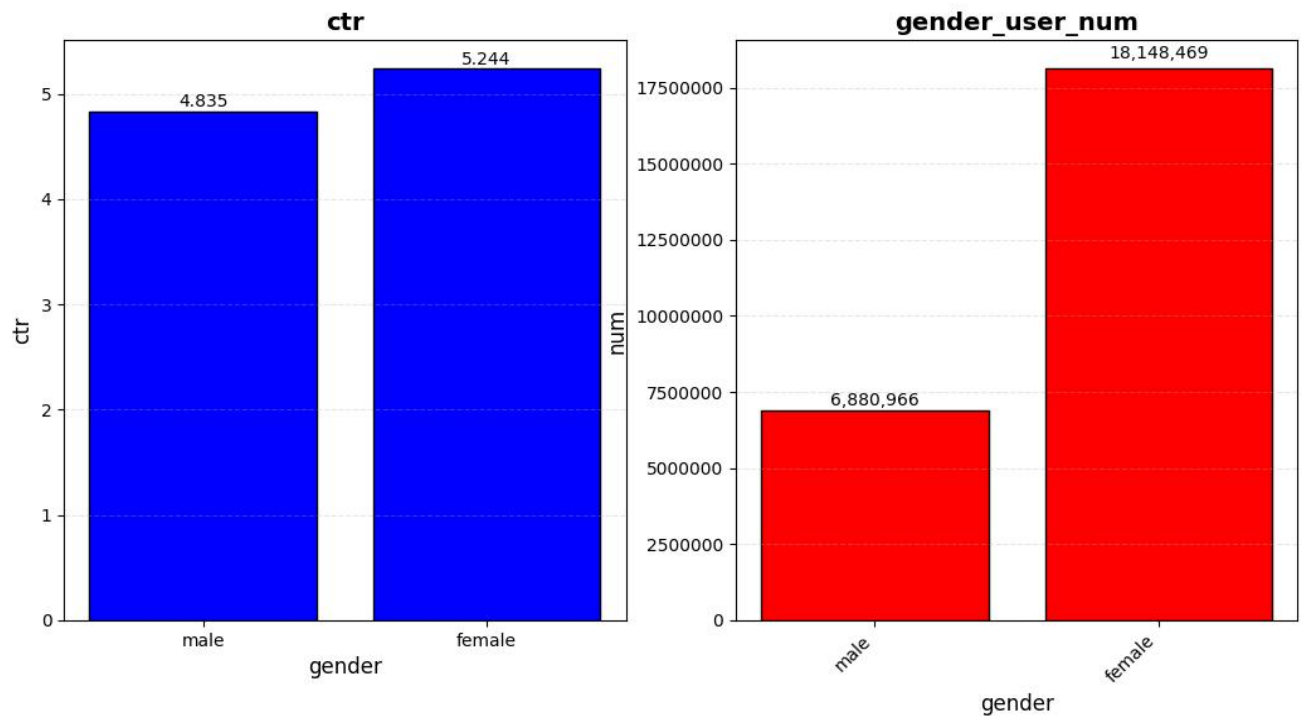
● 分析小结

点击率呈现两头高中间低的两极分化趋势。用户对性价比高的商品和高价奢侈品的有强烈的兴趣，而对价格处于两种层次之间的兴趣平平。可以对浏览量最高的 50~200 区间的广告进行优化从而获得更多点击率。

3. 用户画像

(1) 根据性别分析点击率

性别	男	女
点击率 (%)	4.8354	5.2445
浏览量 (次)	6880966	18148469



- 点击率

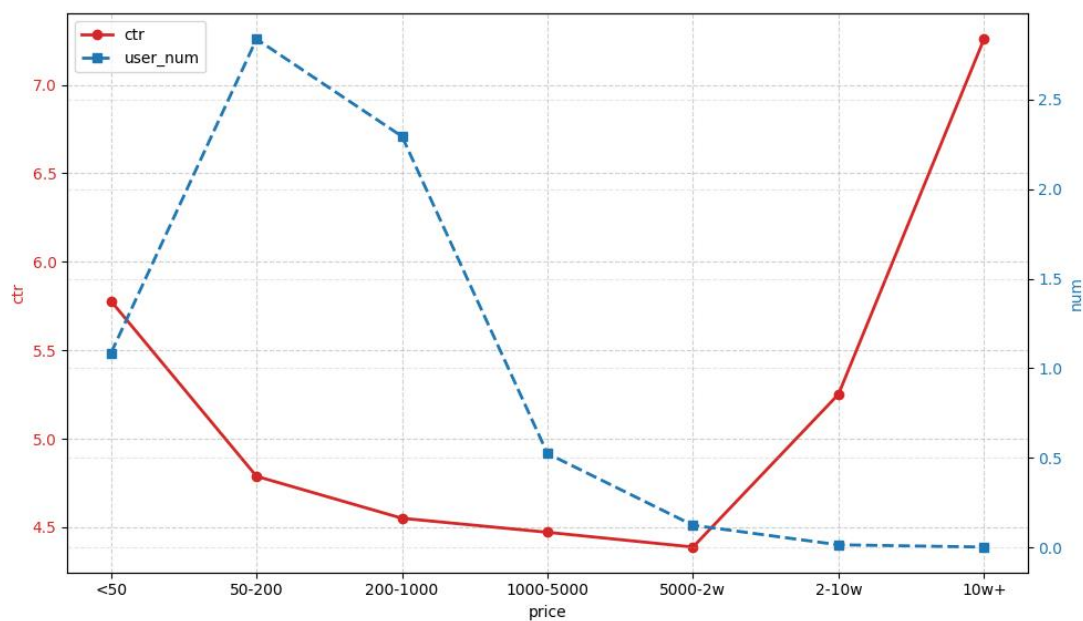
最大值： 5.2444%最小值 4.8354% 极差 0.4090%

- 浏览量

最大值： 18148469 次最小值 6880966 次 极差 11267503 次

(2) 根据性别分析点击的商品价格分布

◆ 男性点击价格分布：



价格	<50	50~200	200~1000	1000~5000
浏览量 (次)	1084719	2837933	2294467	523544
点击率 (%)	5. 7781	4. 7879	4. 5506	4. 4713
价格	5000~2w	2w~10w	10w+	
浏览量 (次)	123204	14015	1474	
点击率 (%)	4. 3878	5. 2515	7. 2591	

- 点击率

最大值： 7. 2591%最小值 4. 3878% 极差 2. 8713%

- 浏览量

最大值： 2837933 次最小值 1474 次 极差 2836459 次

- 男性浏览的商品价格分布

总数： 6879376 均值： 677. 0769 标准差： 6997. 1483

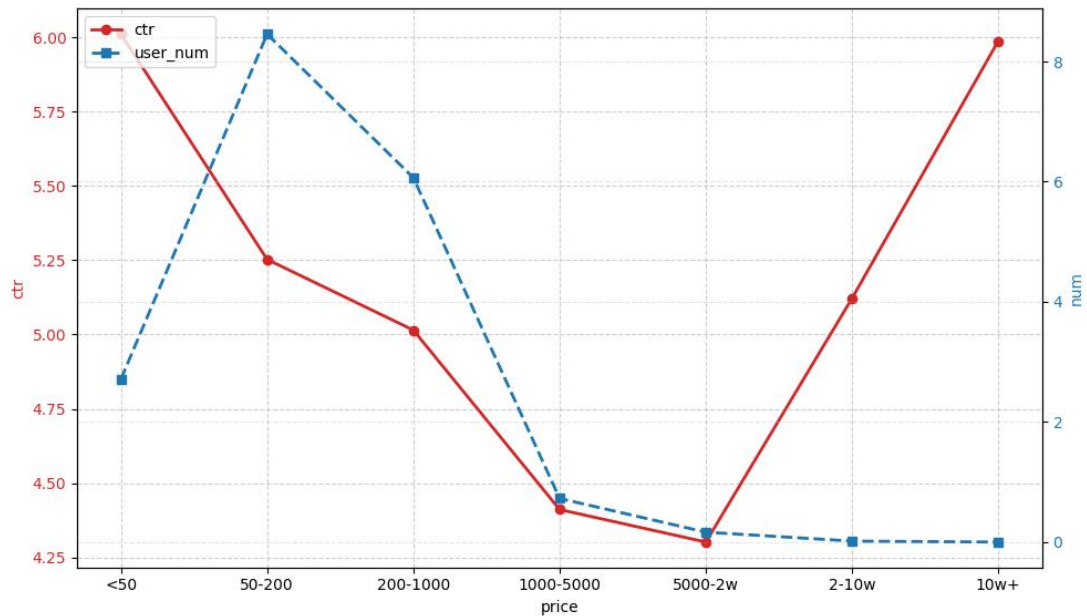
25% ： 87. 0000 50%： 178. 0000 75%： 390. 0000

- 男性点击的商品价格分布

总数： 332625 均值： 689. 1428 标准差： 8233. 2060

25% ： 70. 0000 50%： 168. 0000 75%： 368. 0000

- ◆ 女性点击价格分布：



价格	<50	50~200	200~1000	1000~5000
浏览量（次）	2712580	8456870	6065585	727111
点击率（%）	6.0117	5.2528	5.0149	4.4107
价格	5000~2w	2w~10w	10w+	
浏览量（次）	166042	17259	1720	
点击率（%）	4.3013	5.1220	5.9884	

- 点击率

最大值： 6.0117% 最小值 4.3013% 极差 1.7104%

- 浏览量

最大值： 8456870 次 最小值 1720 次 极差 8455150 次

- 女性浏览的商品价格分布

总数：18147179 均值：435.8595 标准差：4595.9059

25% : 87.0000 50%: 159.8000 75%: 302.0000

- 女性点击的商品价格分布

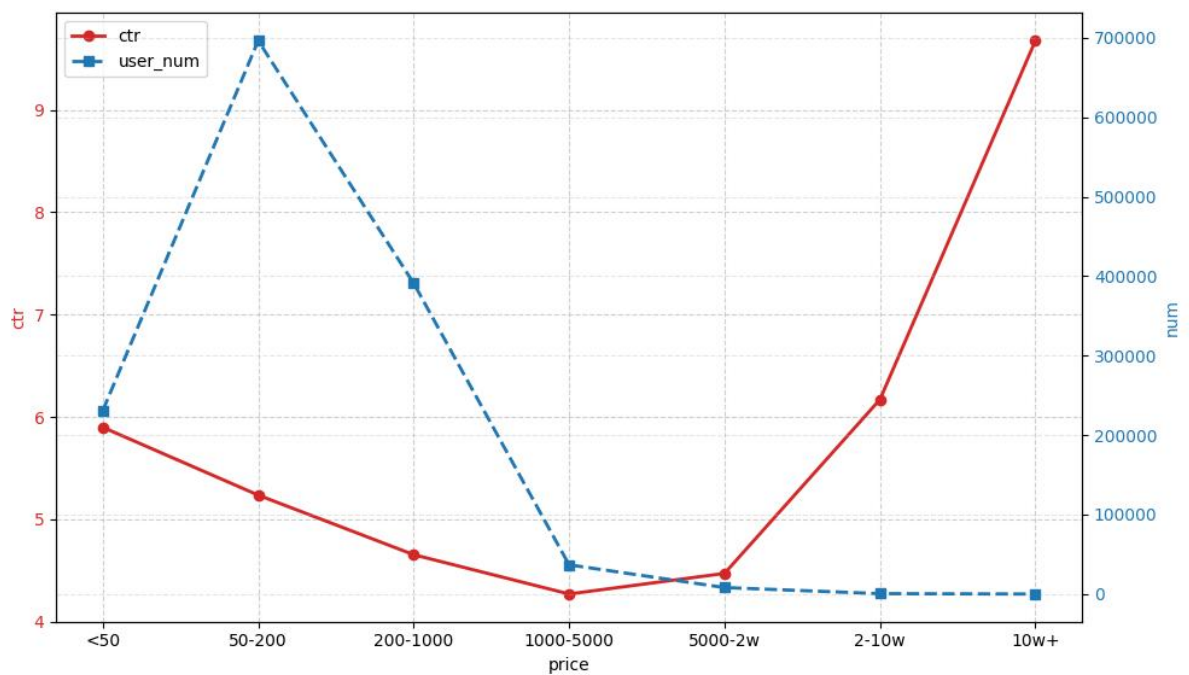
总数：951681 均值：409.9873 标准差：5248.6868

25% : 78.0000 50%: 158.0000 75%: 298.0000

- 针对性别的分析小结

用户与点击率的女性用户贡献了近三分之二的浏览量且有更高的点击率，而男性用户在高价位商品的点击率突出。说明女性用户是主要面向群体，而男性是高价产品的不可忽视的潜在客户。

(3) 大学生和商品价格分布



总点击率：5.1516%

价格	<50	50~200	200~1000	1000~5000
浏览量（次）	230907	696215	391068	36816
点击率（%）	5.8989	5.2365	4.6542	4.2699
价格	5000~2w	2w~10w	10w+	
浏览量（次）	8208	551	93	
点击率（%）	4.4712	6.1706	9.6774	

● 点击率

最大值： 9.6774% 最小值 4.2699% 极差 5.4075%

● 浏览量

最大值： 696215 次 最小值 93 次 极差 696122 次

● 大学生浏览的商品价格分布

总数：1363859 均值：336.4758 标准差：4453.6789

25%：78.0000 50%：138.0000 75%：268.0000

● 大学生点击的商品价格分布

总数：70261 均值：348.6856 标准差：7061.1777

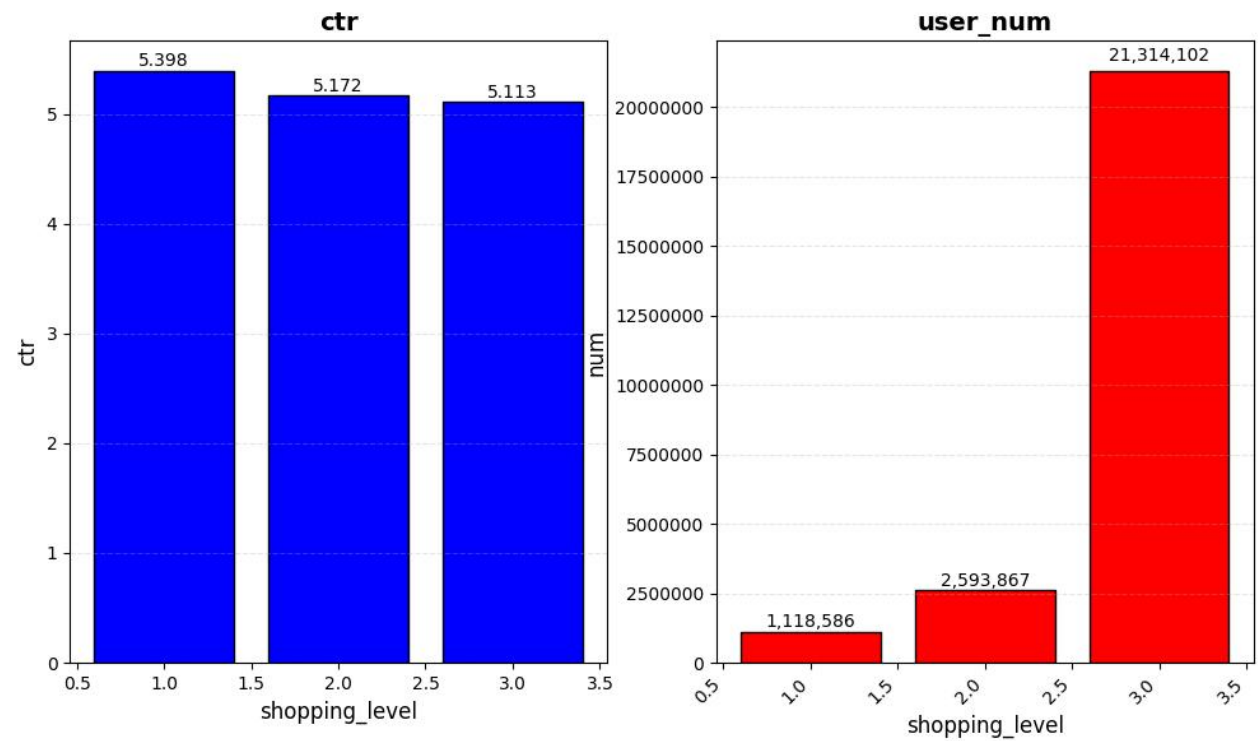
25%：69.0000 50%：133.0000 75%：238.0000

● 针对大学生的分析小结

大学生群体对奢侈品广告抱有极高的好奇心和探索欲，虽然浏览量不高但是点击率奇高，属于较高的层次，但其消费能力与兴趣点存在明显落差。针对此群体应该注重广告的创意和潮流。

(4) 根据用户购物深度分析

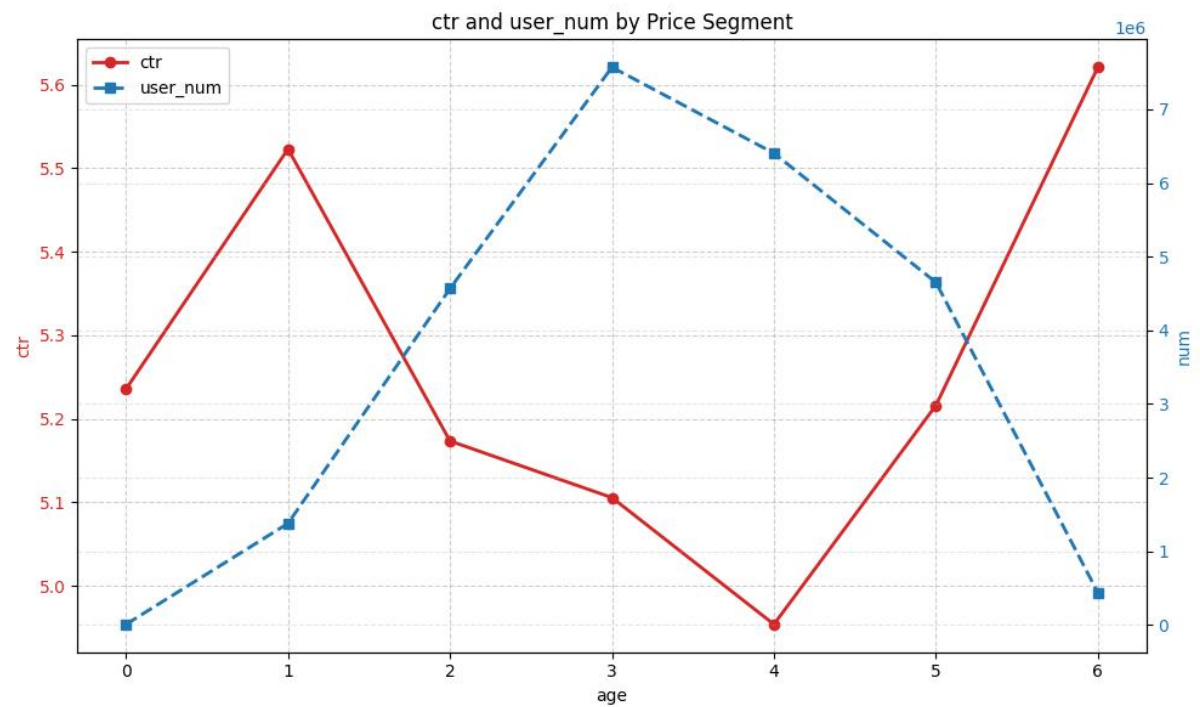
购物深度，1:浅层用户,2:中度用户,3:深度用户



购物深度	浅层（1）	中度（2）	深度（3）
浏览量	1118586	2593867	21314102
点击率	5.3976	5.1723	5.1129

- 点击率
最大值： 5.3976%最小值 5.1129% 极差 0.2847%
- 浏览量
最大值： 21314102 次最小值 1118586 次 极差 20195516 次
- 针对购物深度的分析小结
浅层用户点击率最高，深度用户贡献了绝大部分流量但点击率略低。
新用户更容易点击广告，但是数量还是以深层用户为主。

(5) 根据用户年龄分析



年龄层次	0	1	2	3
浏览量（次）	9187	1379587	4567485	7572879
点击率（%）	5. 2357	5. 5230	5. 1734	5. 1056
年龄层次	4	5	6	
浏览量（次）	6405237	4651896	440284	
点击率（%）	4. 9539	5. 2159	5. 6209	

● 点击率

最大值： 5. 6209%最小值 4. 9539% 极差 0. 6670%

● 浏览量

最大值： 7572879 次最小值 9187 次 极差 7563692 次

● 针对年龄层次的分析小结

年轻群体（年龄层 1、2）点击率较高，中年群体（年龄层 4）流量大但点击率出现低谷。说明年轻用户对广告的接受度和互动性更高，而中年用户虽有较高的浏览量但是点击率偏低，可能是因为对广告内容审美疲劳，

需要更多内容充实的广告。

三、数据分析结论

● 核心发现总述：

本分析揭示了影响淘宝展示广告点击率的三个关键维度：时间、价格与人群差异。广告效果并非由单一因素决定，而是用户意图（何时看）、商品属性（是什么，价多少）与用户身份（谁在看）共同作用的结果。

● 战略性建议：

时段策略：集中资源优化午间时段 的投放效率，并重构晚间广告创意以提升用户兴趣。

商品与定价策略：优化中间价位商品的展示形式，并利用低价与奢侈品的“眼球效应” 为店铺引流。

● 人群精细化运营：

面向女性：作为基本盘，持续优化整体体验。

面向男性：挖掘其高端品类的精准价值。

面向年轻人与大学生：侧重品牌内容与兴趣激发。

面向深度/中年用户：提供更具决策信息与实用价值的广告内容。

● 项目价值与展望：

本项目完成了一次从数据清洗、多维度分析到商业解读的完整实践。分析结论可为广告平台的流量分配、广告主的选品定价与创意优化提供数据支持。后续可进一步结合“行为日志”数据，深入挖掘点击后的转化路径。