# 案例背景

假设公司产品（一款本地服务类app），近期发现日新增用户数下滑明显。老板要求你分析：数据异动的原因是什么？

最开始你只有“问题”sheet中的数据。其他数据是随着分析深入，和业务部门沟通获取的。



# 项目思路

用数据分析万能三步走展开分析：

1. 依次进行：明确问题、分析原因、提出建议这三个步骤。

2. 需要确定：每个步骤包含哪些内容？运用什么分析方法？

3. 根据分析结果给出业务建议。

# 参考答案：

## 明确问题



1. **明确数据来源和准确性**

时间为：4.1- 4.13日，其中10-13日数据异常；

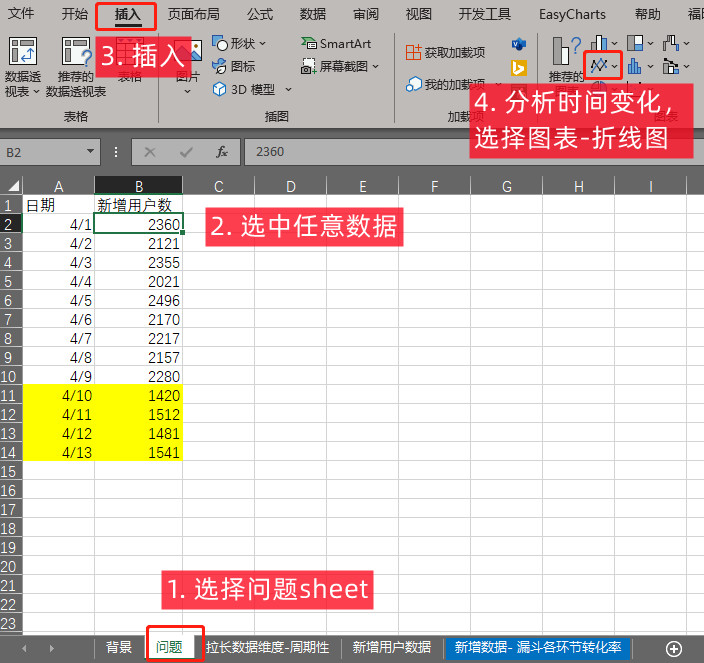
地点和数据来源：通过产品后台数据确认，并核实准确性。

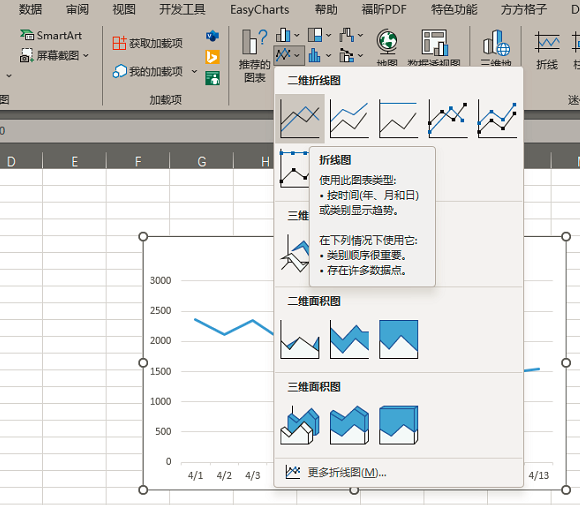


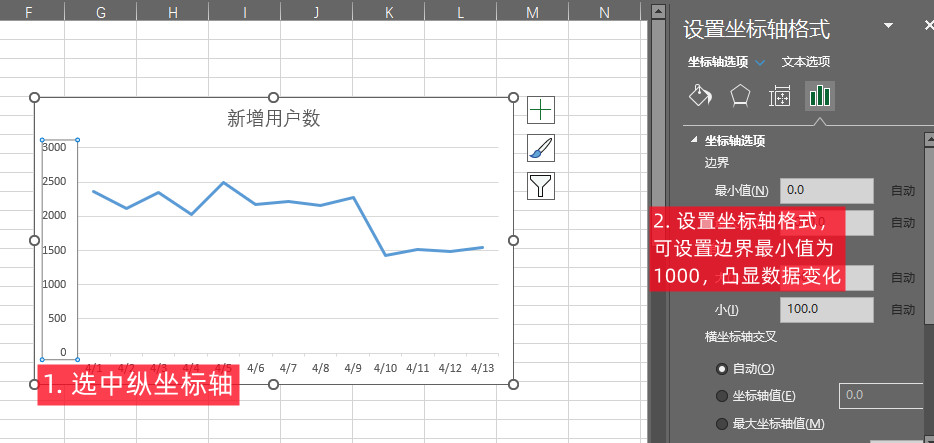
**2）指标理解**

**指标理解1：和谁比下降了？（自身环比数据 + 同比数据）**

**环比数据（制作折线图）**——



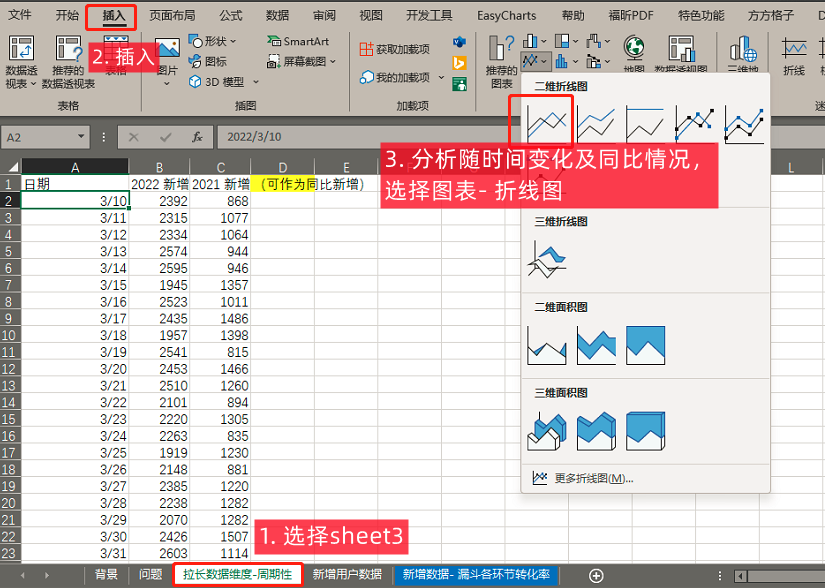




**同比数据（制作折线图）**——

为了确定数据是否为周期性波动，需要从产品部门获取更多日期的的新增用户数据：包含4月1日之前的数据；以及去年同时期数据。

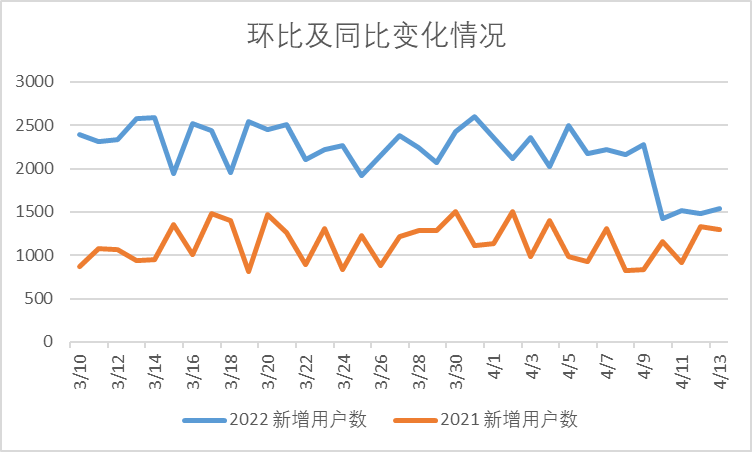




**结论：**

从折线图看出，2022年（蓝色折线）3/10- 4/9新增用户数平稳，4/10开始才明显下滑；并且2021年曲线整体平稳，没有周期性下降。

因此确定：2022年4/10- 4/13的数据存在异常，需要进一步分析原因。

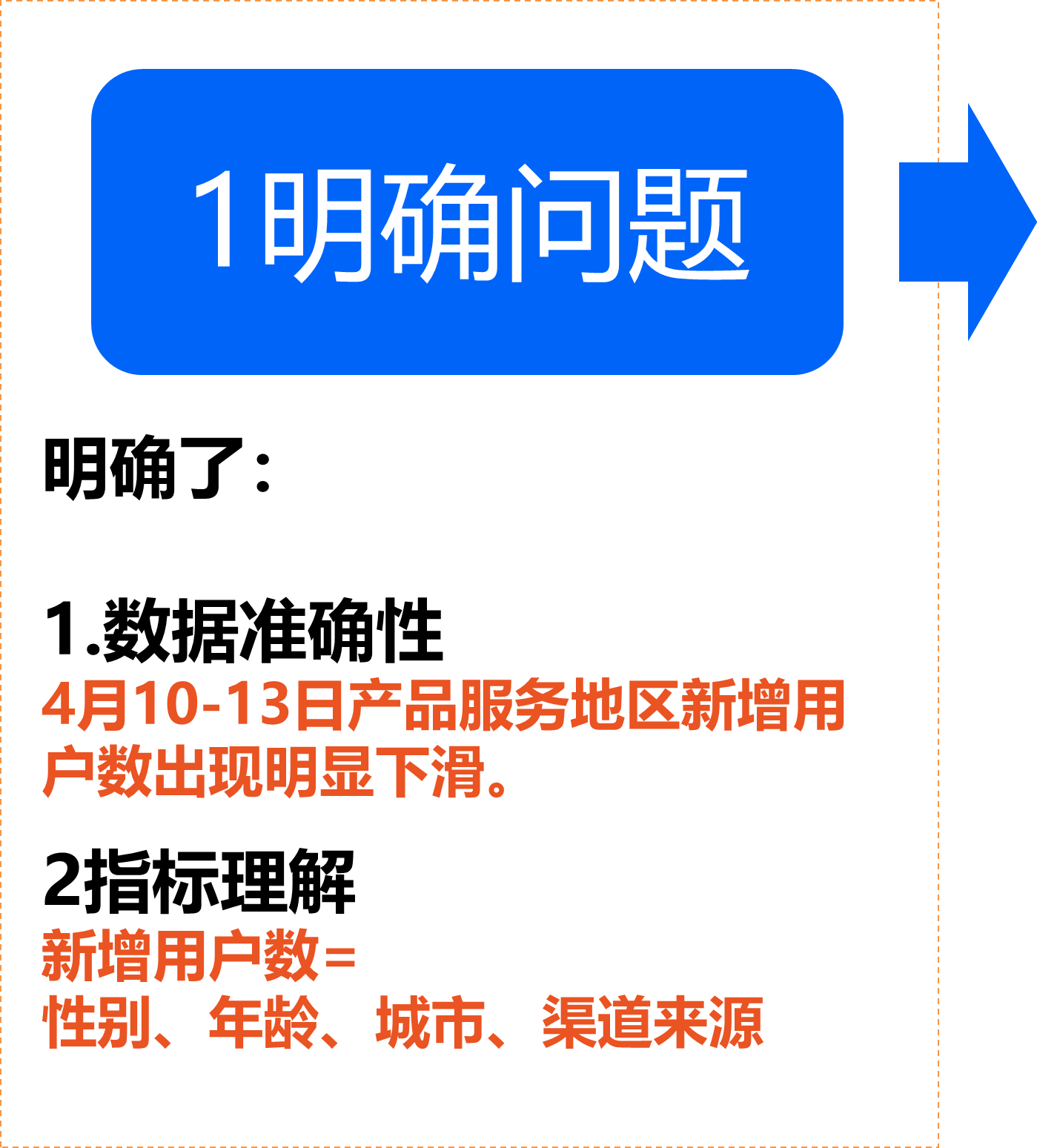


**指标理解2：指标含义**

和业务人员沟通确定，“新增用户数”指标的定义。



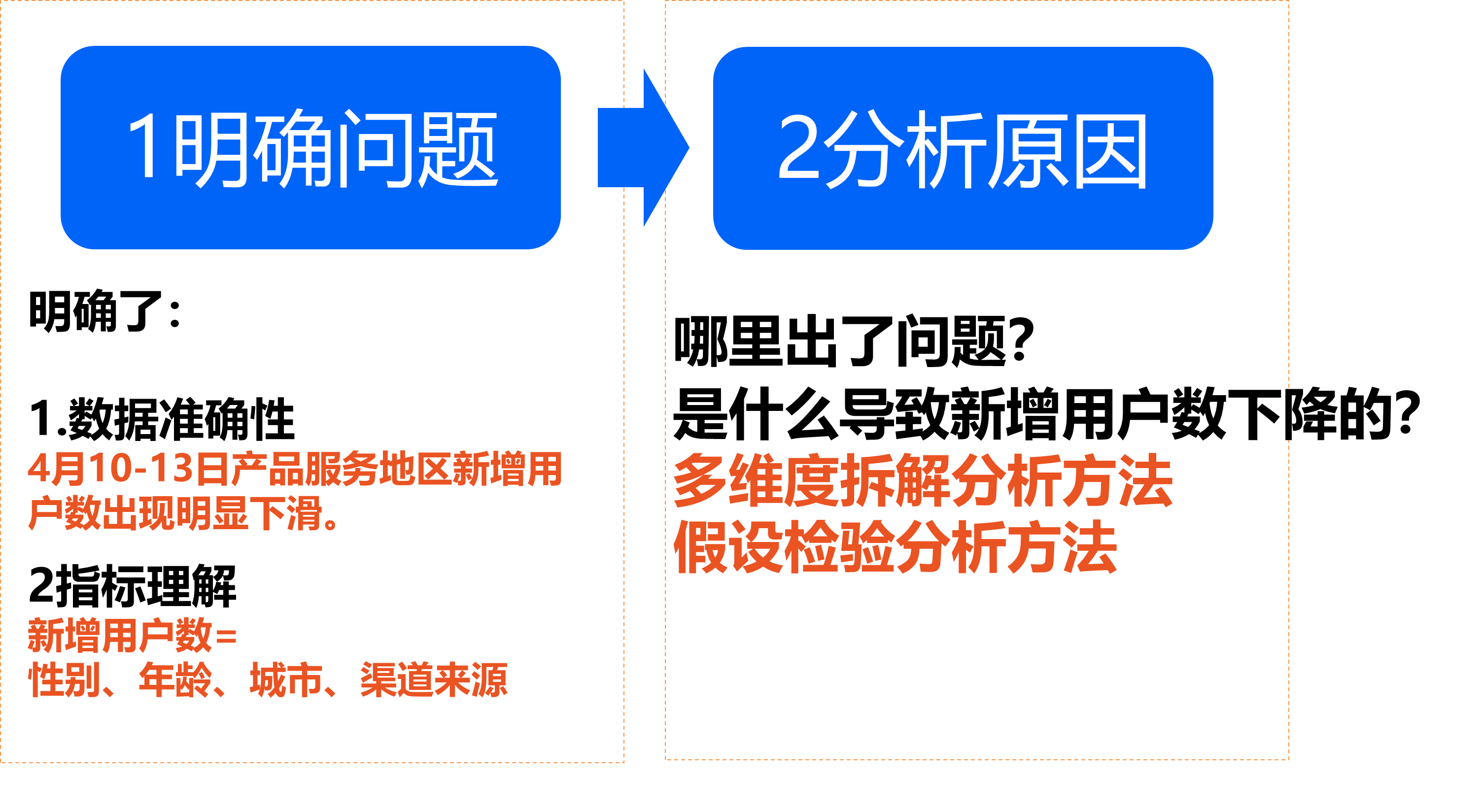
**将以上分析结论，填入步骤“明确问题”中：**



## 分析原因

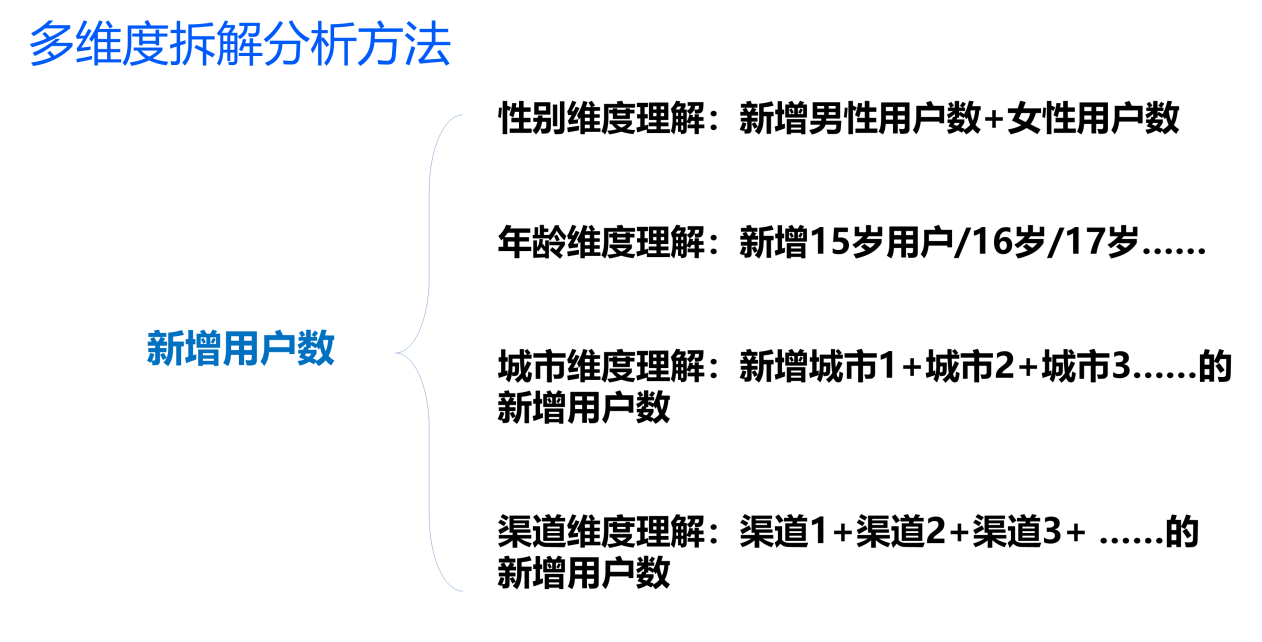
万能三步走的第二步，分析原因。哪里出了问题？是什么导致新增用户数下降的？

分析原因阶段，常见的分析方法包括：多维度拆解分析方法、假设检验分析方法。（使用“多维度拆解分析方法"对问题进行拆解，将一个复杂问题细化成各个子问题；再对拆解的每个部分，使用“假设检验”找到哪里出了问题。）

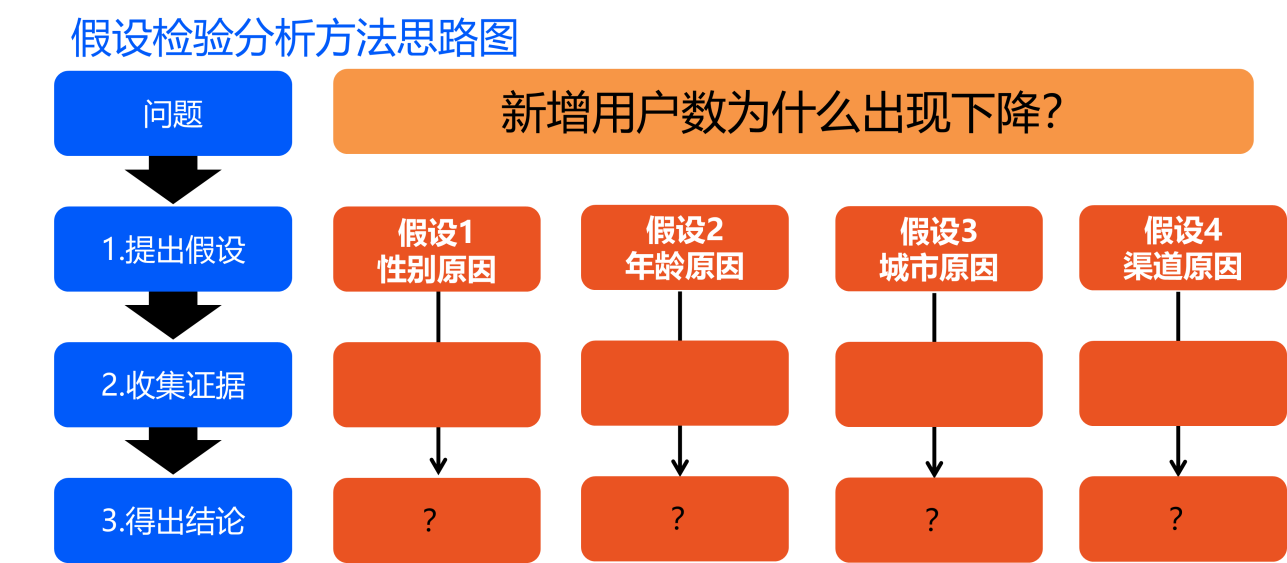


1. **分析思路**

前面已经拆解了新增用户数指标，从性别、年龄、城市和渠道维度拆解。

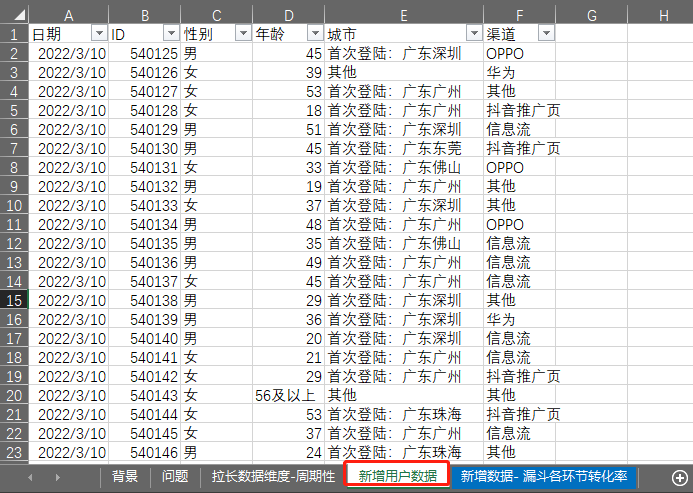


再使用假设检验分析方法，分别假设：是性别原因、年龄原因、城市原因和渠道原因，导致新增用户数下降。

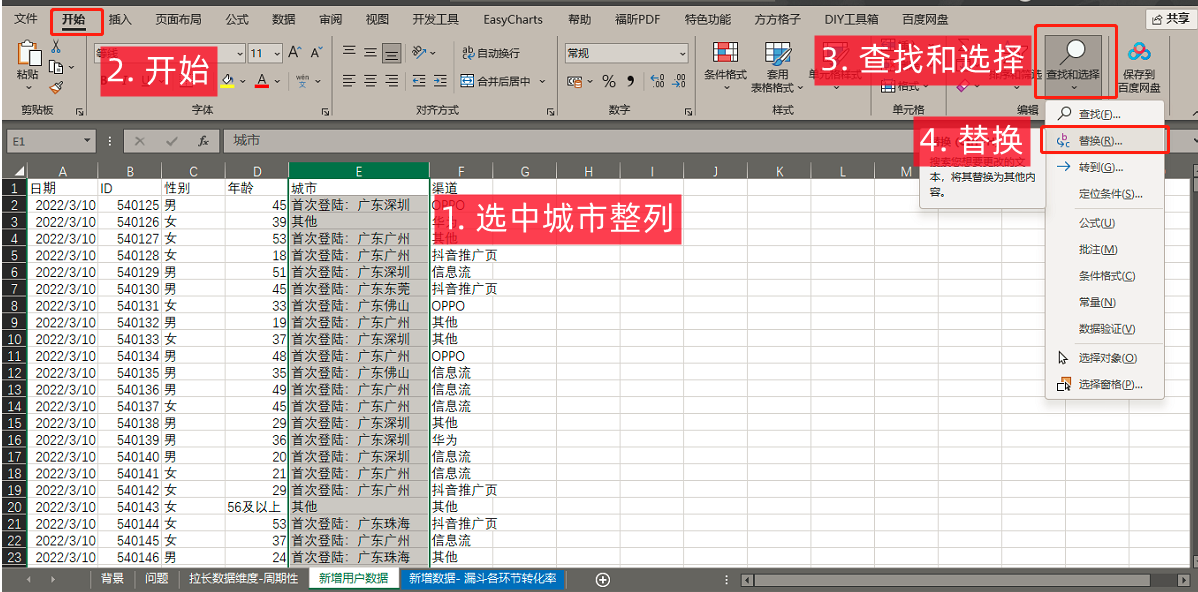


1. **补充数据和清洗**

由于以上假设维度的分析需要数据支持，因此和业务部门沟通需求后，补充了具体的新增用户数据，包含日期、ID、性别、年龄、城市和渠道等字段。详见sheet4：新增用户数据。



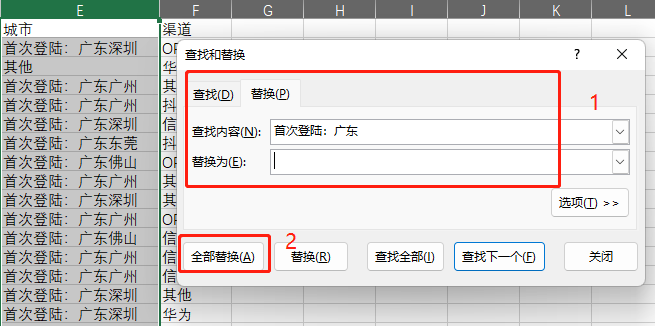
城市列数据太过繁杂，通过数据清洗，只需保留具体城市名称：



查找内容填写- 首次登陆：广东

替换为- 保持空值

点击“全部替换”。



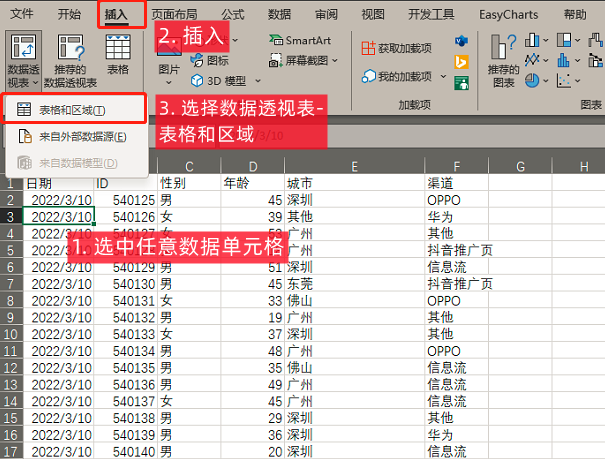
1. **展开具体分析**

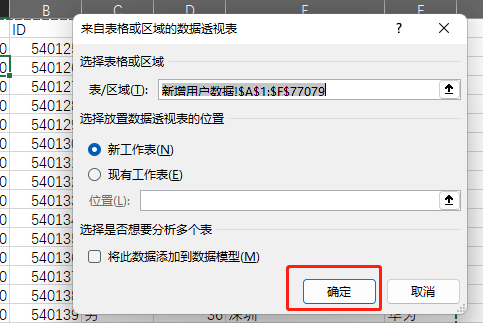
**假设是性别原因导致新增用户数下降：**

需要收集数据下降前后（4月10之前和之后）的性别比例数据，进行论证。

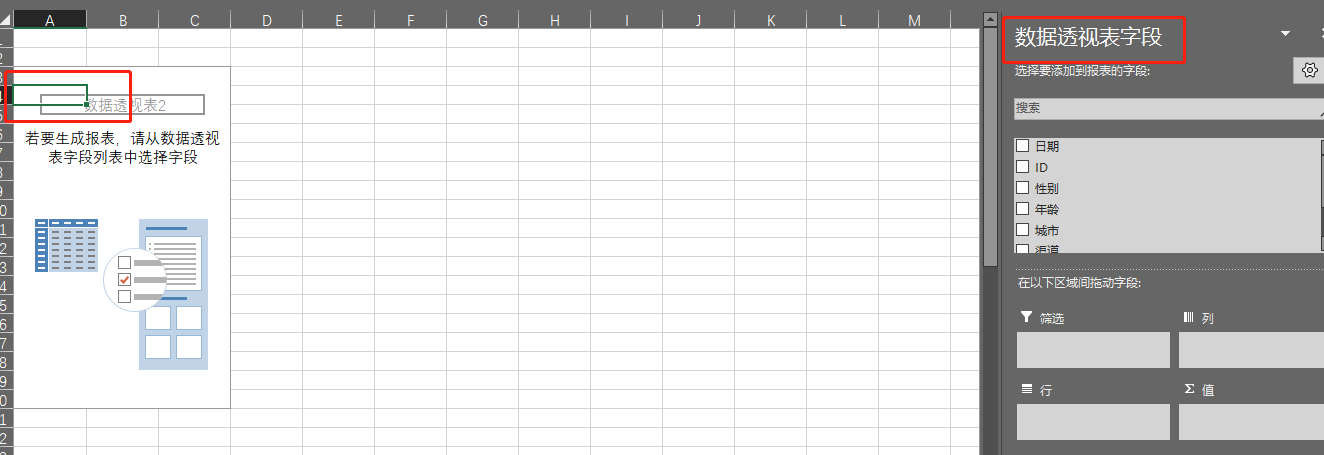


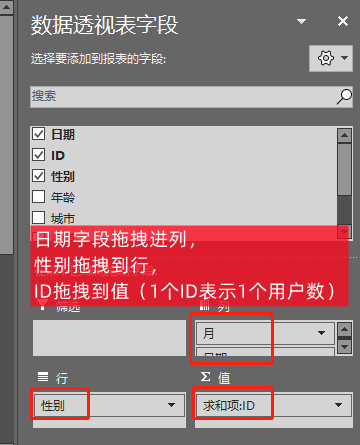
由于数据较多，插入数据透视表方便数据的汇总和分析。



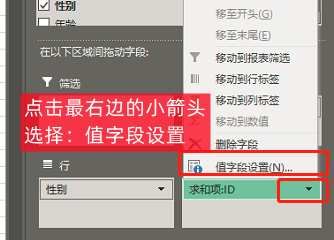


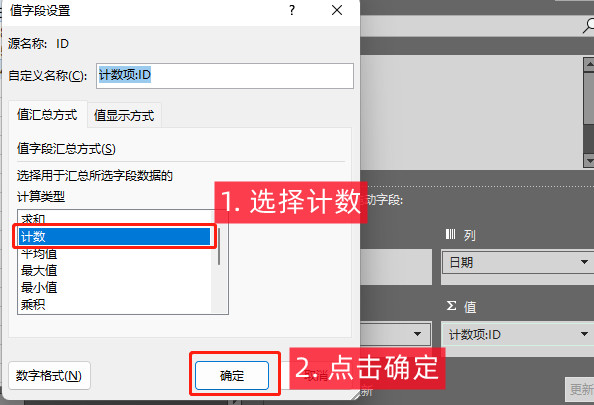
点击数据透视表任意位置，出现数据透视表字段







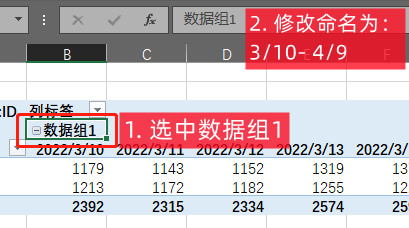


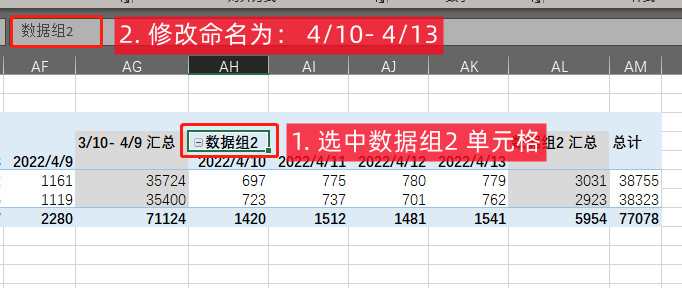


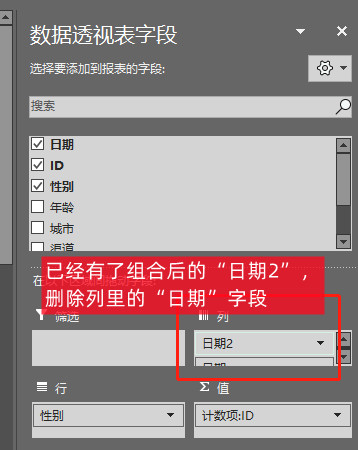


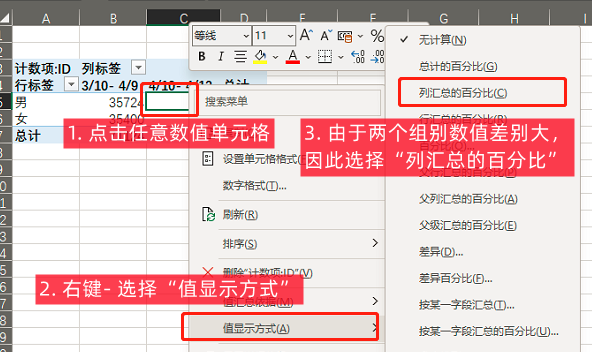


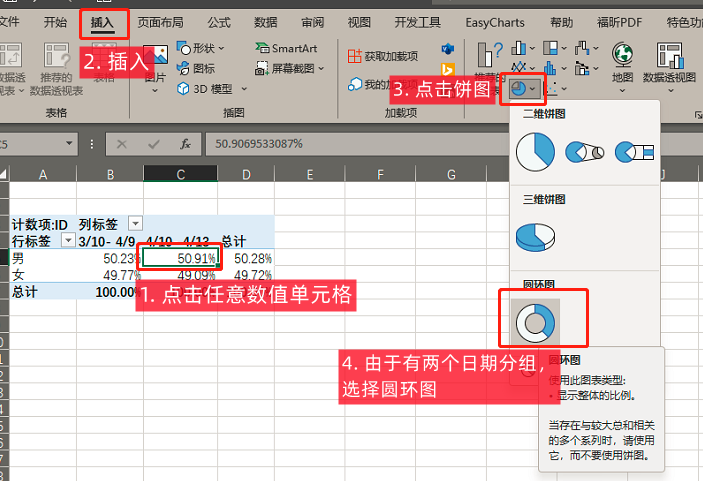


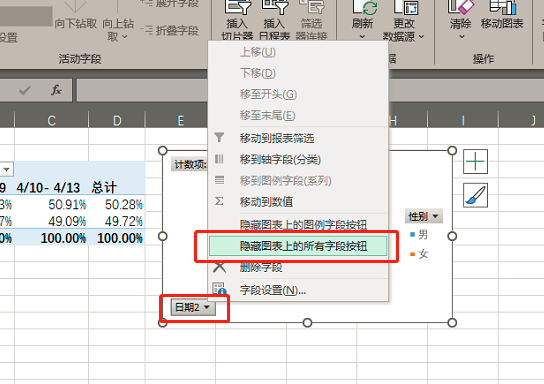


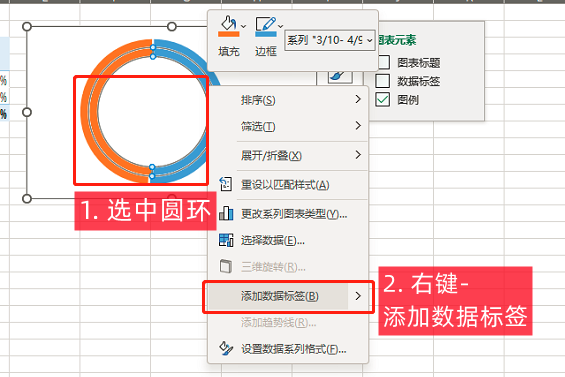


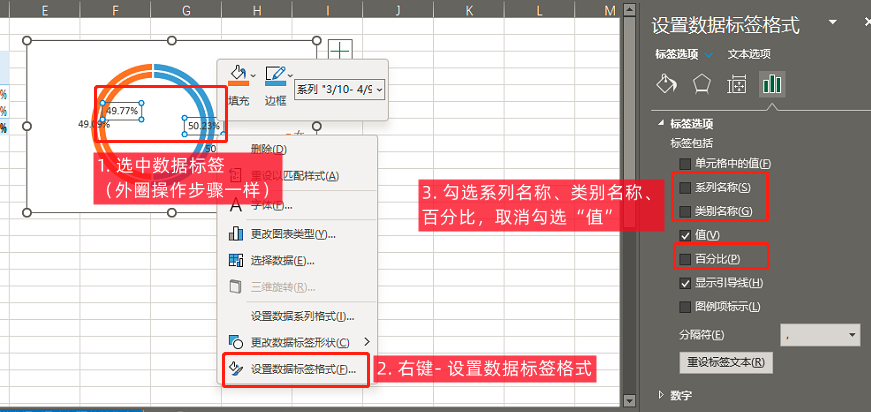


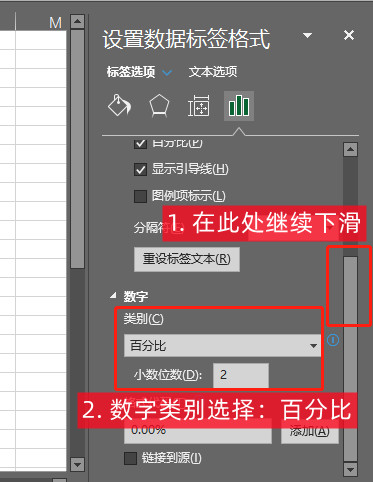






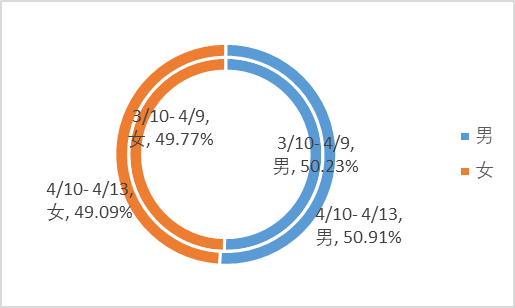






结论：

从圆环图中看出，两个日期阶段男女比例没有明显变化，都是50%左右。因此假设1不成立。



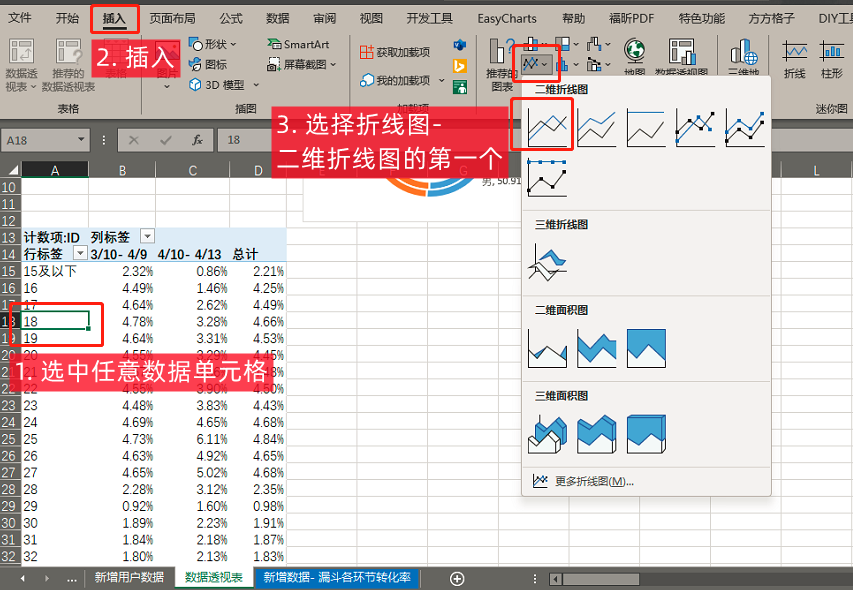
**假设是年龄原因导致新增用户数下降：**

需要收集数据下降前后（4月10之前和之后）的年龄比例数据，进行论证。



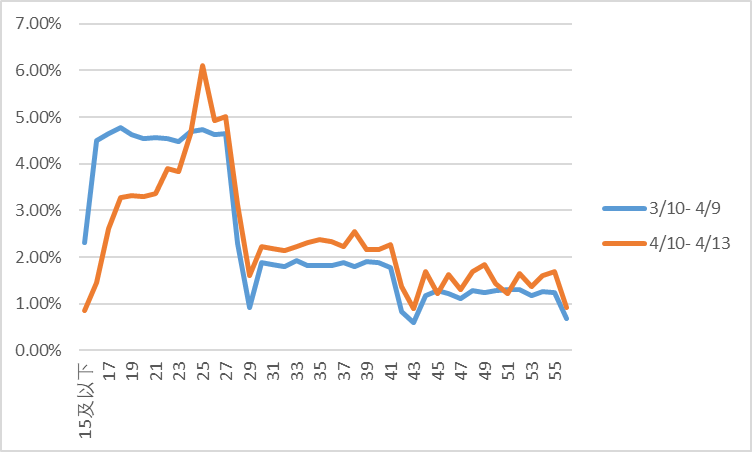






结论：

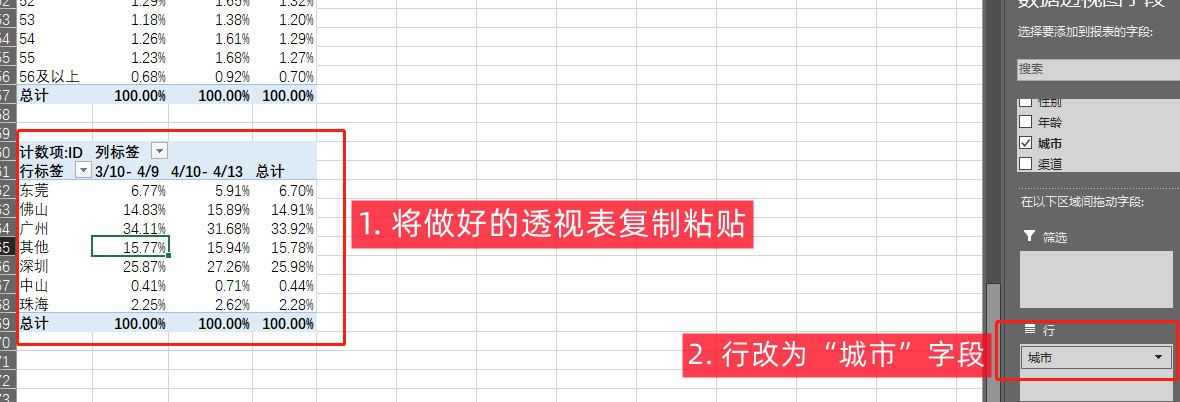
从折线图对比看出，4/10- 4/13的组别中（橙色折线），23及以下的新增用户占比明显更少。因此假设2成立，23及以下年龄段出现数据异常。

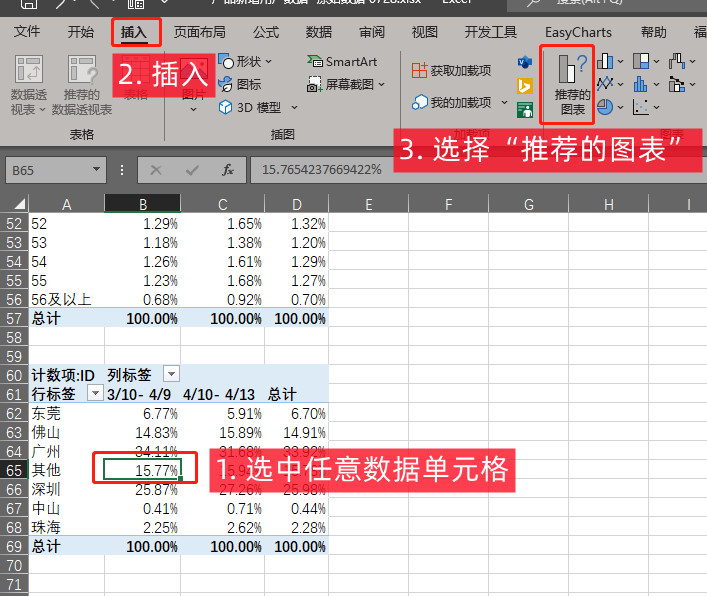


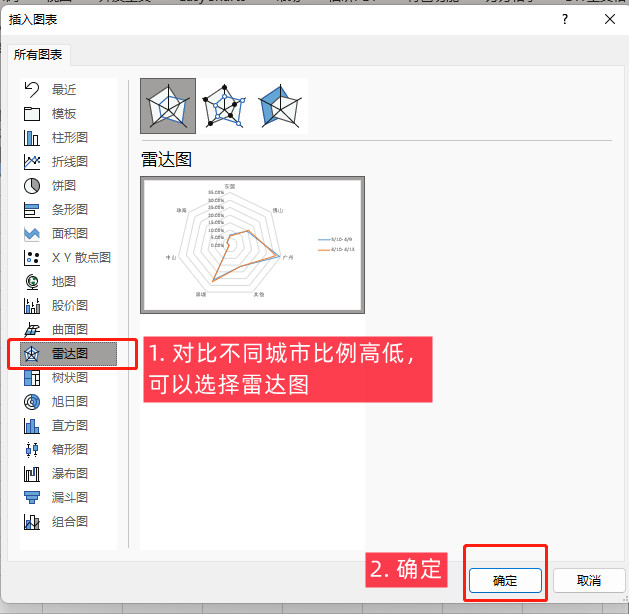
**假设是城市原因导致新增用户数下降：**

需要收集数据下降前后（4月10之前和之后）的城市比例数据，进行论证。



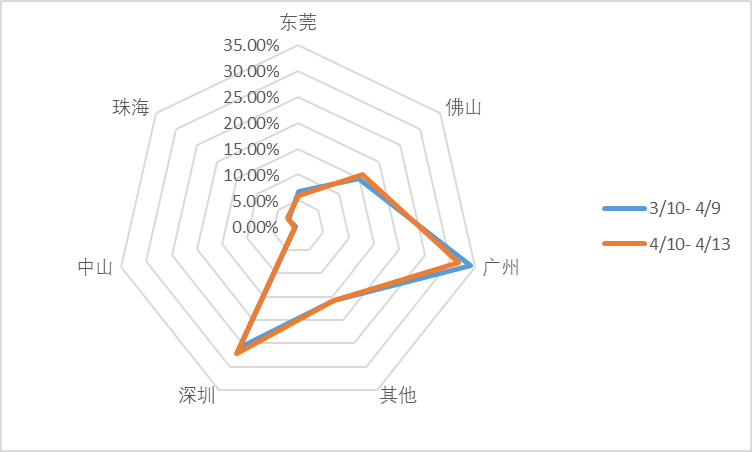






结论：

从雷达图对比看出，两个日期阶段，各城市新增用户比例基本相同。因此假设3不成立。

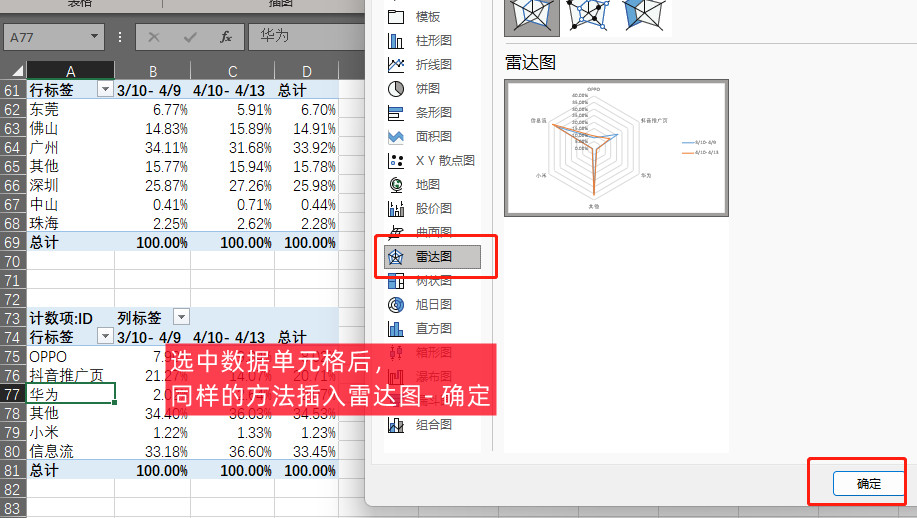


**假设是渠道原因导致新增用户数下降：**

需要收集数据下降前后（4月10之前和之后）的渠道比例数据，进行论证。

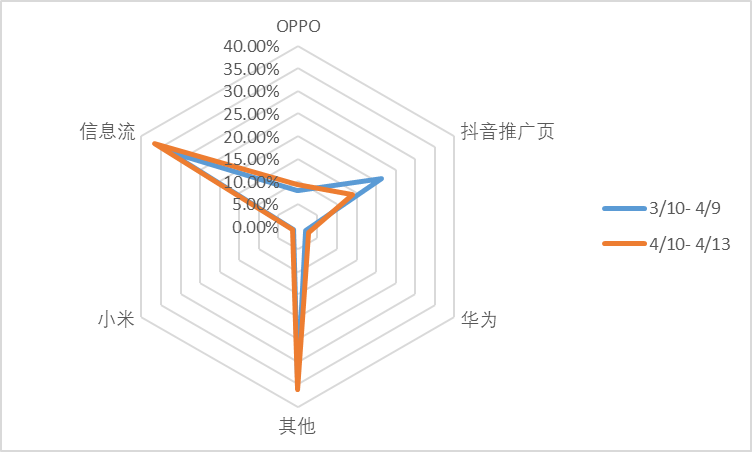






结论：

从折线图对比看出，4/10- 4/13的组别中（橙色线条），抖音推广页的新增用户占比明显更少。因此假设4成立，抖音推广页出现了问题。



**总结四个假设分析的结论**：

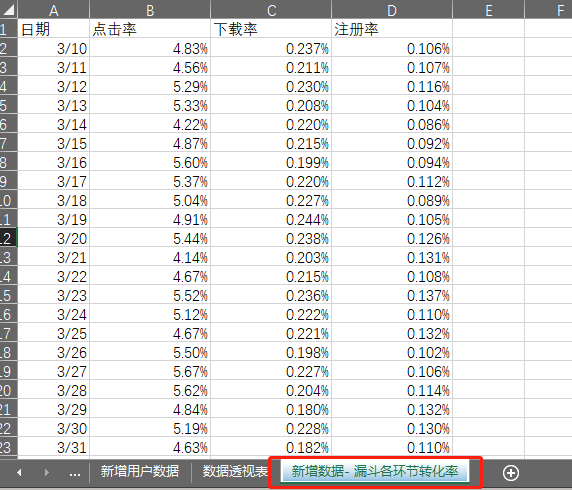


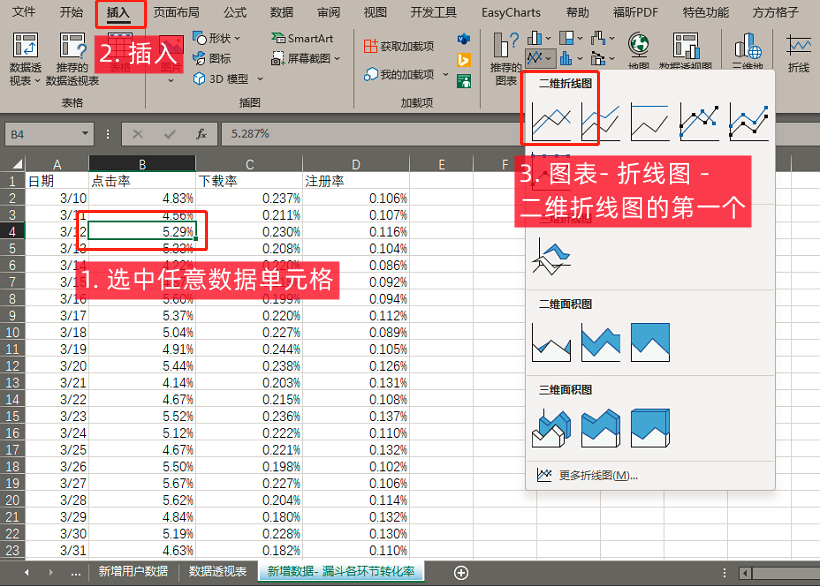
**补充假设- 漏斗分析方法：从业务流程出发防止遗漏问题**

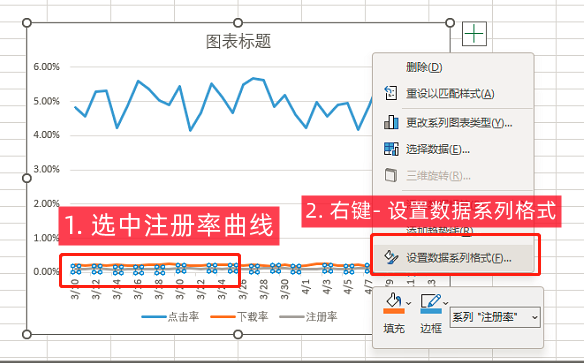
用户成为新增用户前，要经历渠道曝光- 点击- 下载- 注册的漏斗阶段。因此需要补充分析，是否某个漏斗阶段出现问题，导致新增用户减少。

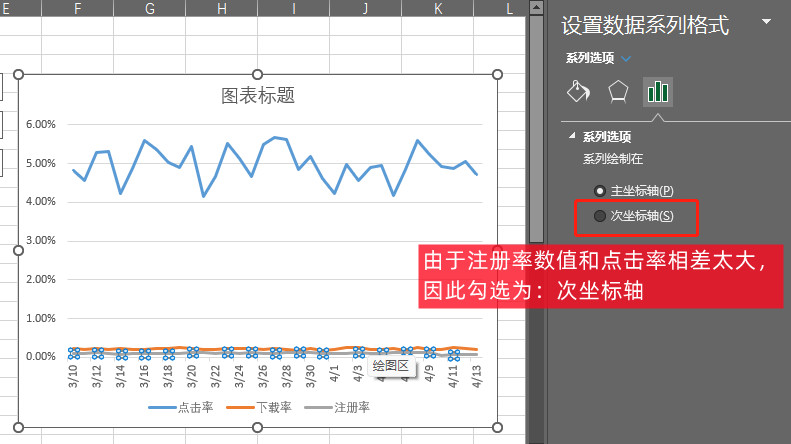


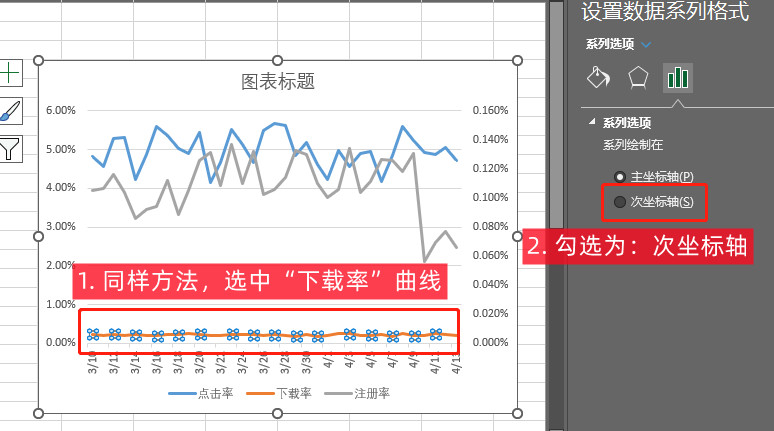
和业务部门沟通获取3/10- 4/13的用户点击率、下载率和注册率数据。





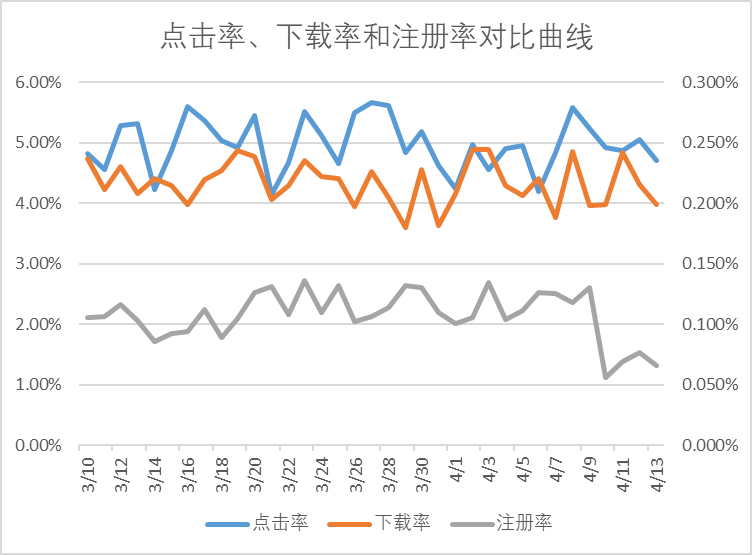






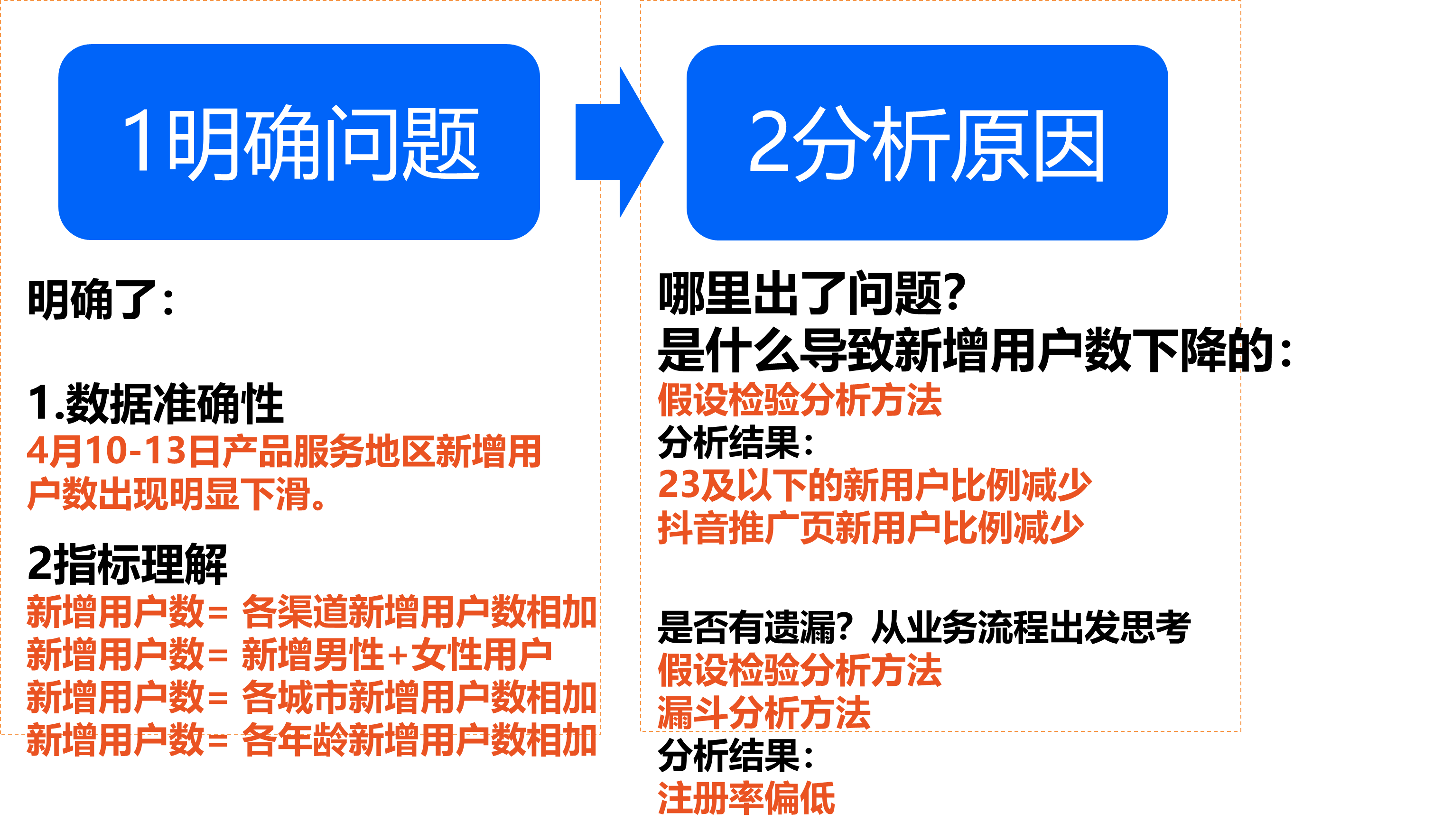
结论：

从折线图对比看出，4/10- 4/13只有注册率曲线出现明显低谷。因此补充假设成立，注册率明显下滑。



1. 总结

将以上5个假设检验的分析结果，填入“分析原因”步骤中。



## 提出建议

针对排查出的问题，分别提出业务建议和后续解决方向。

