

用户流失原因 数据分析实战

安思聪

数数科技数据分析师



如何发现用户流失的原因



- ◆ 首先需要明确对流失用户的定义
- ◆ 在不同的阶段，用户流失的原因并不相同，确定流失用户之后，还需按照流失用户所处的生命周期，进一步对流失用户进行分类
- ◆ 对比活跃用户和流失用户的差异，进行解释驱动的数据分析
- ◆ 改变策略，对活跃用户的流失情况进行验证，调优
- ◆ 观察活跃用户情况，对老用户进行召回

01 对流失用户的定义





什么是流失用户



如何界定流失用户

按用户活动

注册后流失、活跃流失、付费流失

按用户类型

新/老用户流失、核心用户流失、付费用户流失

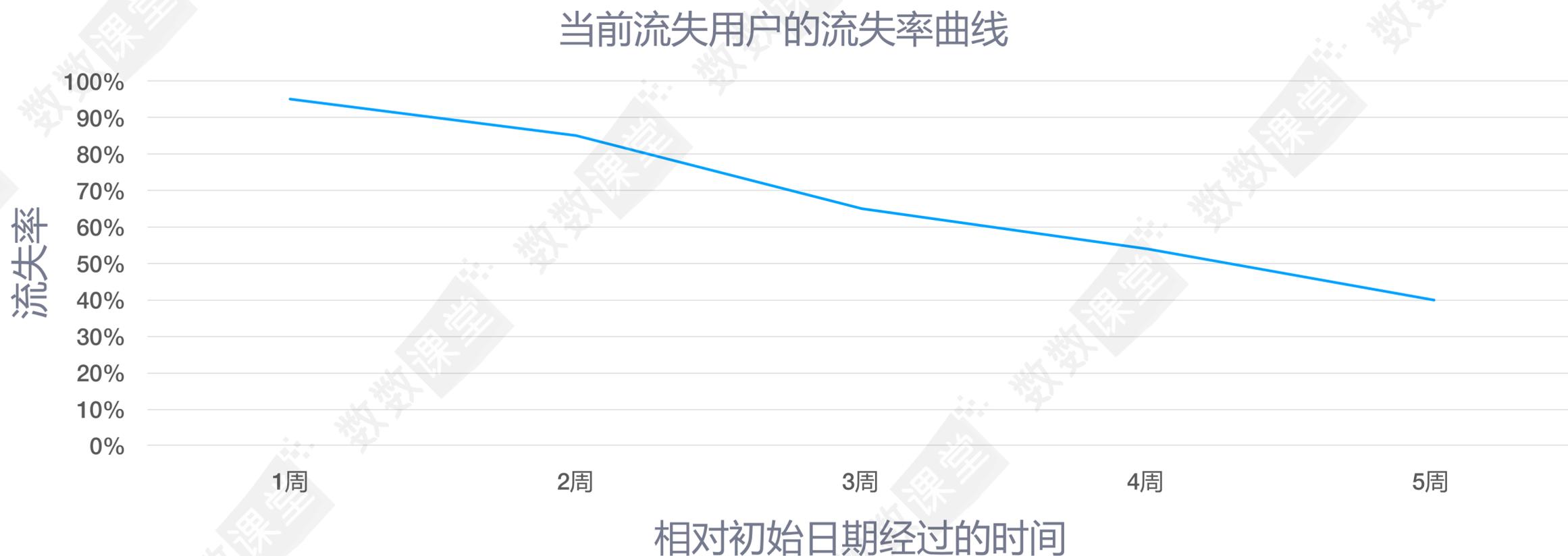
按流失状态

当前流失、历史流失、流失但回流



当前流失用户的流失率曲线

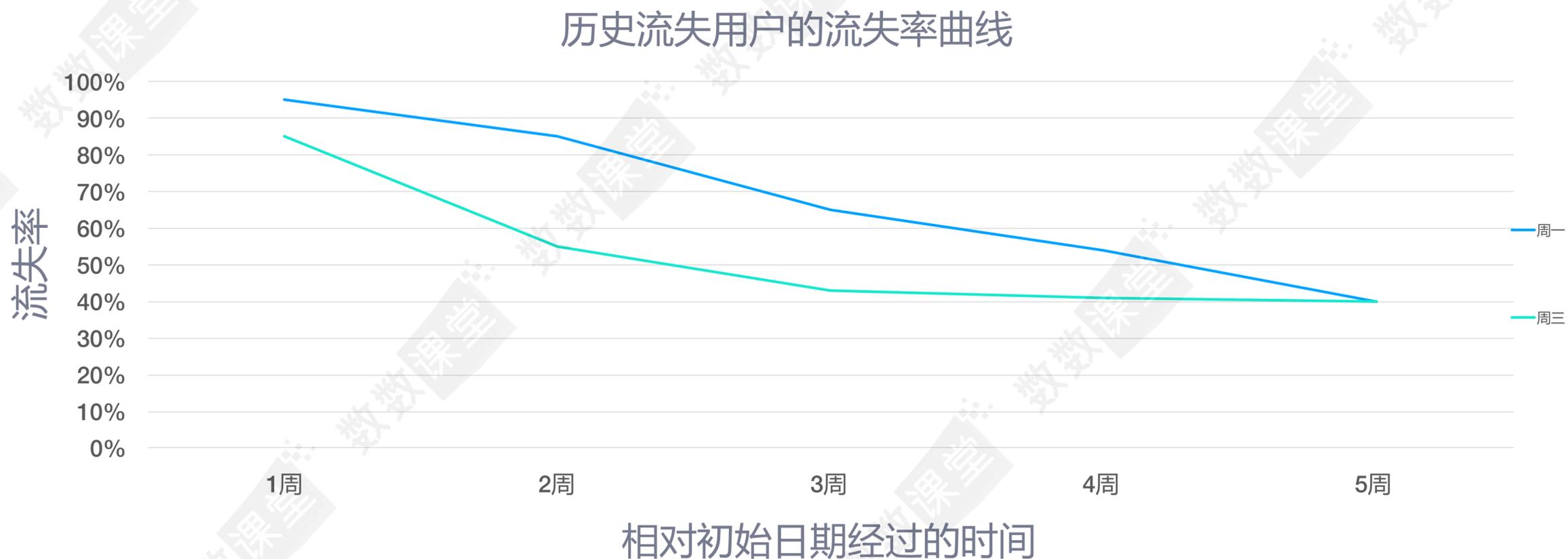
- 用户在某个确定的时间范围内**到当前的**未回访用户的比例
- ◆ 流失率=在该段时间内未登录的用户/初始选定的用户数*100%
- ◆ 以用户在该时间点之后时间为横轴，流失率为纵轴作图，得到流失率曲线
- ◆ 曲线单调递减是因为，一旦用户产生了行为，就不算是流失用户了





历史流失用户的流失率曲线

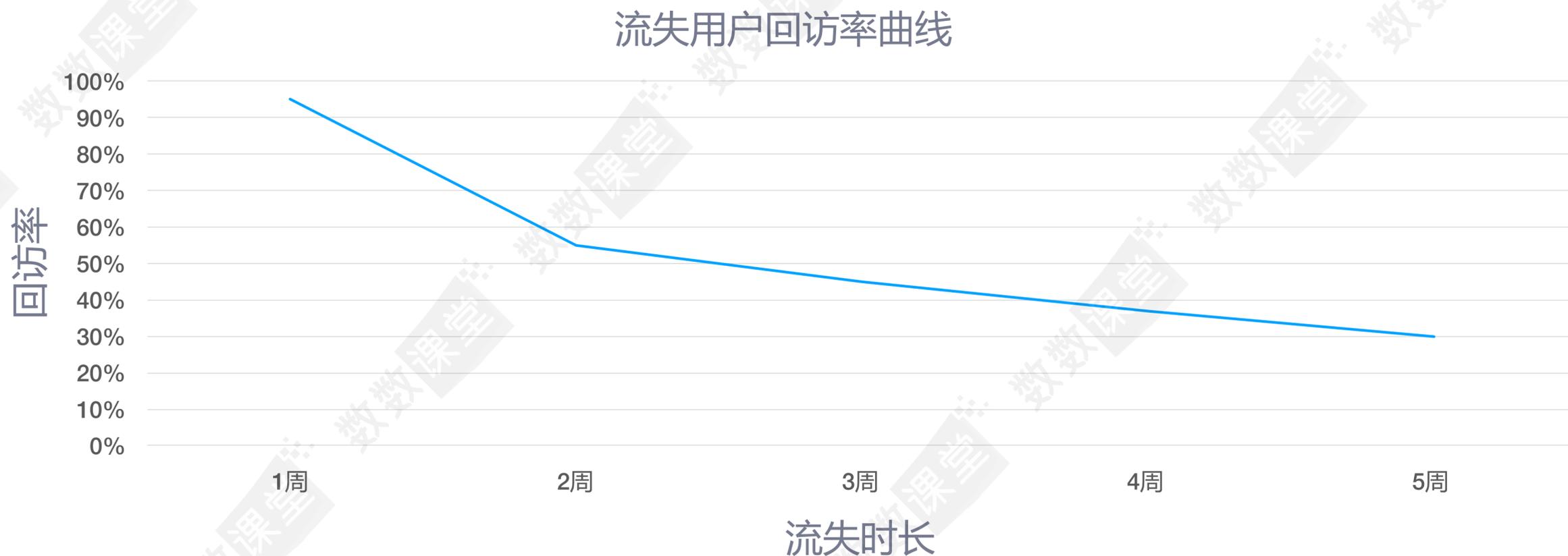
- 用户在某个确定的时间范围内到**之后一段时间内**的未回访用户的比例
- ◆ 流失率=在该段时间内未登录的用户/初始选定的用户数*100%
- ◆ 以用户在该时间点之后时间为横轴，流失率为纵轴作图，得到流失率曲线
- ◆ 曲线单调递减是因为，一旦用户产生了行为，就不算是流失用户了





流失但回流用户的回访率曲线

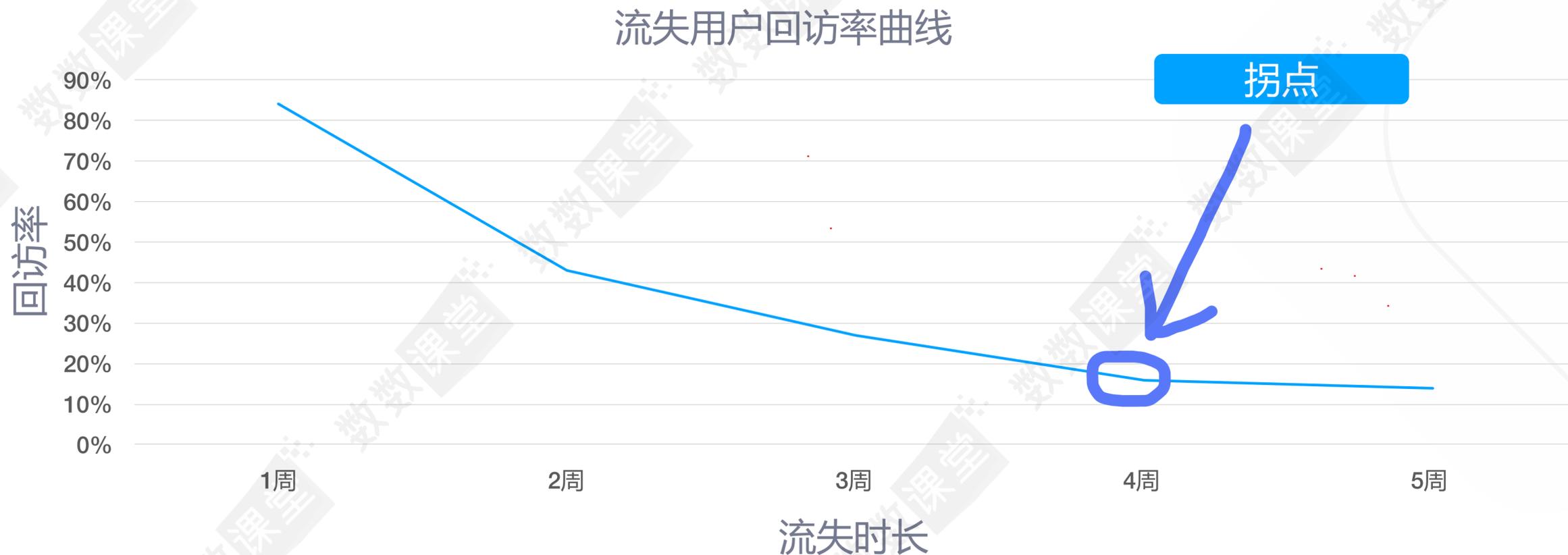
- 用户在流失后还可能回访，回访用户**实际没有流失**
- ◆ 用户的流失时间越长，回访率越低
- ◆ 以用户的流失时长（两次活跃的间隔时长）为横轴，回访率为纵轴作图，得到回访率曲线，可以衡量不同流失时长用户的回访潜力





回访率曲线计算方法

- ◆ 确定横轴的时间粒度，日/周/月
- ◆ 计算每日/周/月，活跃用户在经过多久后，才首次再次回访的人数
- ◆ $\text{回访率} = \frac{\text{在上次登录后的第N日/周/月才再次登录的用户数}}{\text{上次登录的用户数}} * 100\%$
- ◆ 将回访率数据按照初始日/周/月的活跃用户数进行加权平均，得到总回访率曲线





不同流失计算方式的特点

流失状态	当前流失	历史流失	流失但回流
计算方式	查看某一时间点 到当前 的用户未回访情况	查看某一时间点到 其之后 的 第N日/周/月 的用户未回访情况	在用户上次登录后的 第N日/周/月 再次登录的情况
举例	上周登录的用户，有多少人在这周末登录	上周注册的用户，注册后的7日内流失情况	半年前注册的用户，还有多少用户留在游戏中
体现的信息	现阶段多少用户未回流	历史上用户在登录后的N日内未回流的情况	历史上用户在登录后的第N日出现了回流
优势	计算方便，反应当前用户粘性与流失情况	计算较复杂，反应历史上一定周期内用户的粘性	计算较复杂，反应历史用户的实际回流情况
不足	用户上次活跃距离当前越近之后的回访概率越大，这部分用户实际是非流失用户	未统计用户的回访情况，实际的流失率会小于算出的流失率	该数据存在一定滞后性并且模糊了不同阶段用户回流的差异

数数课堂

02 用户生命周期的划分



用户生命周期

用户在游戏阶段的阶段可分为：新手期，探索期，成长期，佛系期
不同阶段用户流失的原因不同

生命周期	新手期	探索期	成长期	佛系期
特点	对游戏核心玩法和核心内容不了解不熟悉	进入新手流程，并开始体验前期快速成长，建立了对游戏内容的初步认识	登录频率和游戏时长比较稳定，在游戏中有坚定目标	对游戏的核心玩法熟悉，不再追逐更好的收益
触发标志	用户首次进入游戏	完成新手引导	等级，关卡进度，获得的物品到达关键节点	等级，关卡进度，获得的物品到达关键节点
流失原因	内容/操作/设定/画面/玩法	流程过长/成长不明显/内容单一	成长难度过大，道具爆率过低，用户氛围不好	游戏内容深度广度不足，可探索内容过少
需要关注的内容	买量渠道，引导流程，bug闪退	道具获得情况，任务完成情况	等级驻留，关卡驻留，行为频率，充值情况	用户当前的进度，等级，持有道具，排名

数数课堂

03 如何进行数据分析



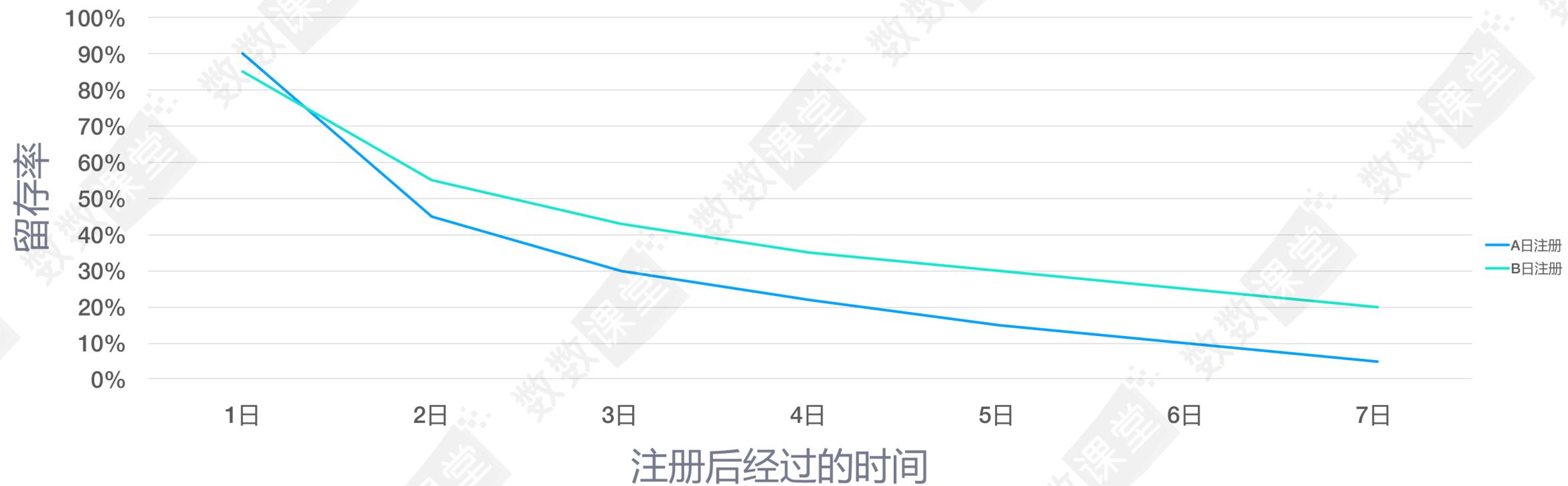


案例



某两日新增用户的留存率出现了明显的差异，该如何定位原因

新用户的7日留存曲线





案例 第一步：提出问题

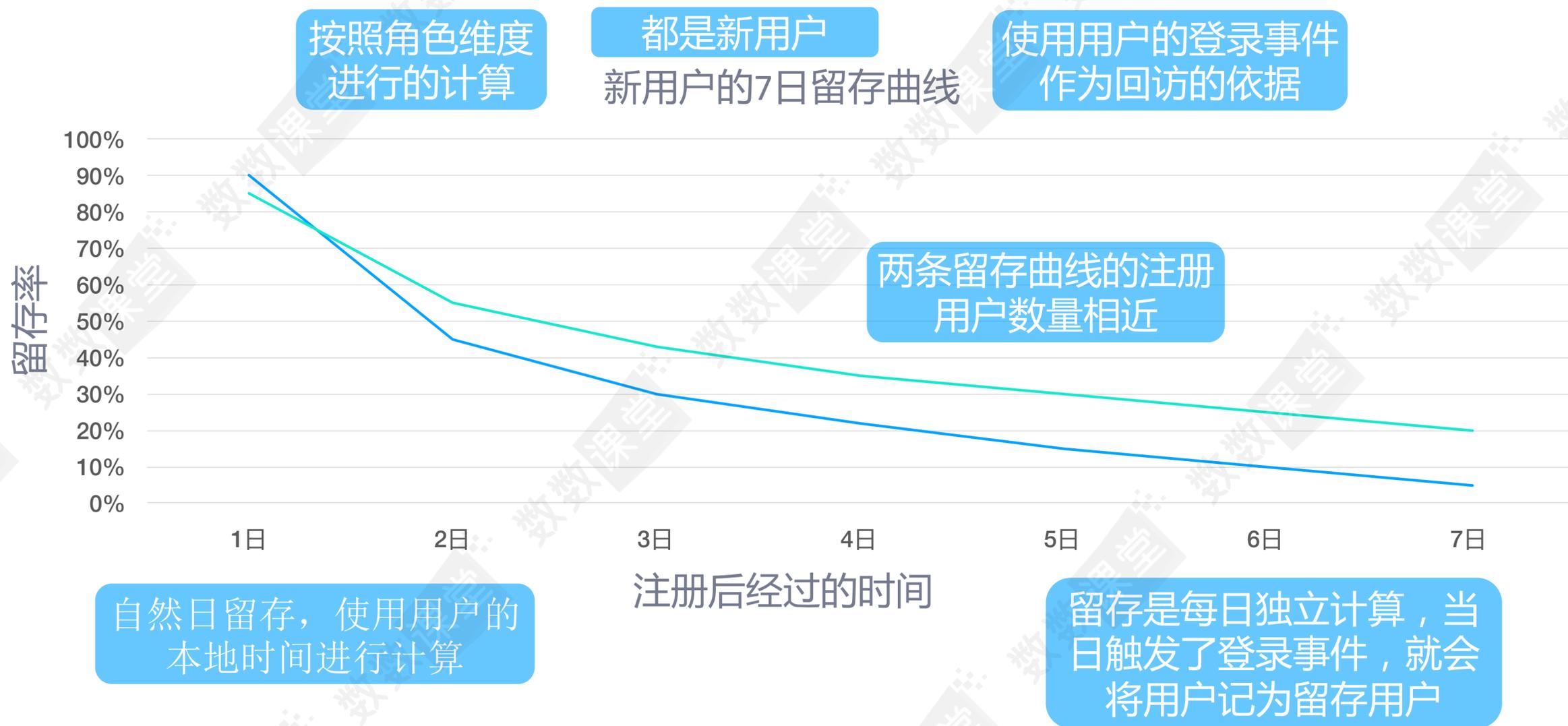


- ◆ 留存的数据是什么样的计算逻辑？
- ◆ 在这两天里，新注册用户的人数差异是什么样的？
- ◆ 横轴上的日期，是自然日还是非自然日？
- ◆ 行为数据是按用户本地时间计算的，还是将所有数据偏移到某个确定的时区下计算的？
- ◆ 计算的用户维度是角色？ 账户？ 设备？
- ◆ 该部分用户包含多少种不同的生命周期？



案例

某两日新增用户的留存率出现了明显的差异，该如何定位原因





案例 第二步：确定分析范围

- ◆ 新用户的留存，不是老用户的留存，用户的生命周期集中分布在新手至成长期之间
- ◆ 确定用户的历史实际流失情况，用历史流失曲线去分析
- ◆ 准备分析的维度

留存分析 UTC+01:00 TA用户ID 报表

分析时段 过去30天 7日 流失

图表设置 时间详情 显示设置 导出

先登录，后登录流失情况

日期	登录用户数	1日	2日	3日	4日	5日	6日	7日
阶段均值	-	43.64%	36.41%	32.91%	30.19%	28.13%	26.51%	25.28%
2020-06-23(二)	469	168 35.82%	138 29.42%	116 24.73%	99 21.11%	93 19.83%	90 19.19%	89 18.98%
2020-06-24(三)	421	134 31.83%	105 24.94%	89 21.14%	85 20.19%	83 19.71%	78 18.53%	74 17.58%
2020-06-25(四)	464	169 36.42%	133 28.66%	118 25.43%	112 24.14%	104 22.41%	100 21.55%	94 20.26%
2020-06-26(五)	513	212 41.33%	177 34.50%	151 29.43%	142 27.68%	135 26.32%	130 25.34%	120 23.39%
2020-06-27(六)	643	290 45.10%	262 40.75%	244 37.95%	232 36.08%	227 35.30%	217 33.75%	200 31.10%
2020-06-28(日)	613	318	284	259	251	228	202	184



数据对比维度



- ◆ 用户渠道
- ◆ 用户区服
- ◆ 用户生命周期
- ◆ 单次在线时长

- ◆ 累计付费
- ◆ 流失等级
- ◆ 关卡停留
- ◆ 玩法偏好



案例 第三步：进行数据分析与对比

- ◆ 需要确定流失用户的id，圈定这批要分析的流失用户
- ◆ 两组数据之间进行对比，与对等的留存活跃用户进行对比

The screenshot shows a data analysis interface with the following components:

- Header:** 事件分析 (Event Analysis) | UTC+01:00 | 近似计算 (Approximate Calculation) OFF | 刷新 (Refresh) | 下载 (Download) | SQL | 报表 (Report)
- Analysis Object 1 (①):** 充值 (Recharge) | 触发用户数 (Triggered User Count) | 流失 (Lost) | 属于分群 (Belongs to Group)
- Analysis Object 2 (②):** 充值 (Recharge) | 触发用户数 (Triggered User Count) | 低等级人群 (Low-level group) | 属于分群 (Belongs to Group)
- Filters:** 全事件符合 (All events match) | 属性筛选 (Attribute Selection)
- Grouping:** 分组项 (Grouping Item) | 总体 (Overall)
- Buttons:** 开始计算 (Start Calculation) | 保存报表 (Save Report)
- Analysis Period:** 分析时段 (Analysis Period) | 过去30天 (Last 30 days) | 合计 (Total)
- Chart Settings:** 图表设置 (Chart Settings) | 指标详情 (Indicator Details) | 柱状图 (Bar Chart) | 折线图 (Line Chart) | 饼图 (Pie Chart) | 仪表盘 (Dashboard)



案例



发现：**A**天注册的用户之所以留存率高，是因为付费用户整体留存率更高。而流失的用户很多是未付费用户。而**B**天注册的用户，付费用户占比少，导致留存率整体偏低。平均比A天注册的用户低**5%**

下一步做什么？

案例

第四步：进行解释驱动的数据分析

◆ 流失率高5%，不是原因，而是现象，需要发掘原因

5W法探究原因

When
Where
Who
Why
What

解释驱动的用户调查

行为观察
用户访谈
用户问卷



案例 第五步：验证猜想，设置优化策略



A/B测试

04 如何看待用户流失分析



流失数据的特点

流失数据是一个侧面

流失数据存在滞后性



流失用户分析中可能碰到的误区



- ◆ 将相同的流失定义，应用到所有的流失分析
- ◆ 试图明确所有用户的流失原因
- ◆ 以为看到的数据确实反应了用户的实际流失原因
- ◆ 不去关注现有的活跃用户
- ◆ 忽略了用户流失的时间点



数数科技创立于2015年,是国内专业的游戏数据分析服务商。公司总部位于上海,在北京、广州、深圳和成都设有分部。

数数科技自主研发的Thinking Analytics(简称“TA系统”),提供“数据采集-存储建模-实时计算-分析展示”的一体化解决方案。平台支持全端数据采集、多维交叉分析和私有化部署,可帮助企业应对复杂的游戏业务场景,实现精细化运营和业务增长。

目前,数数科技已经服务了300+家游戏企业,其中包括Funplus、昆仑游戏、中手游、电魂网络、英雄互娱等知名游戏公司。TA系统接入的游戏多达2000+款,《弓箭传说》、《最强蜗牛》、《江南百景图》、《闪耀暖暖》、《古剑奇谭》等爆款都在使用。2021年3月,数数科技完成1亿元B轮融资。

