

2020 DataFunTalk 精选

用户画像技术及方法论



姚凯飞

用户画像技术及方法论

作者：姚凯飞

导读：用户画像也是近几年比较热的一个词，不过很多小伙伴对于画像的认知还只是标签化的层面，或者是只是利用其做一些简单的分群分析；如何全面地认知并做系统性地尝试，背后有非常多的点需要我们深思挖掘。今天就根据自己的一些浅见进行分享，因为与商品画像的联系，中间也会掺杂一些商品画像的知识。如果不足和错误之处，还望大家批评指正，enjoy。

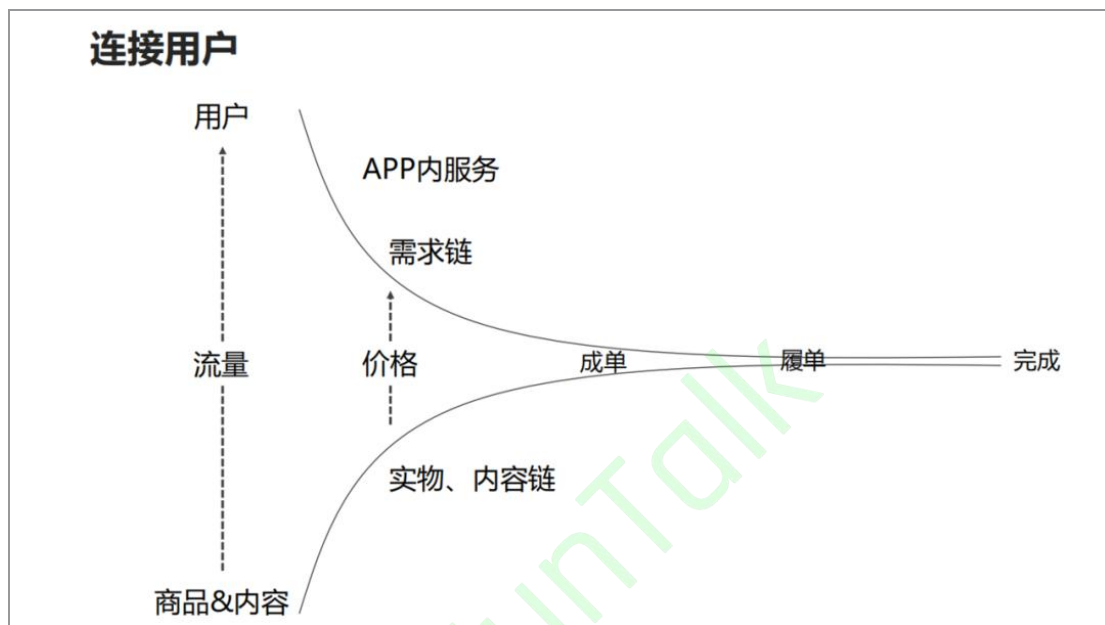
今天的分享希望通过介绍如何通过用户及商品画像来构建数字化体系，前面是整个内容的概述，然后从浅到深我们去挖掘和讲解其中的各个模块，由于时间的原因，我们会主要通过用户画像展开，对于重要或者有意思的商品画像相关的点会略做辅助性的介绍。

互联网公司的核心

用户体验(商品,供应链,物流,营销,活动...)

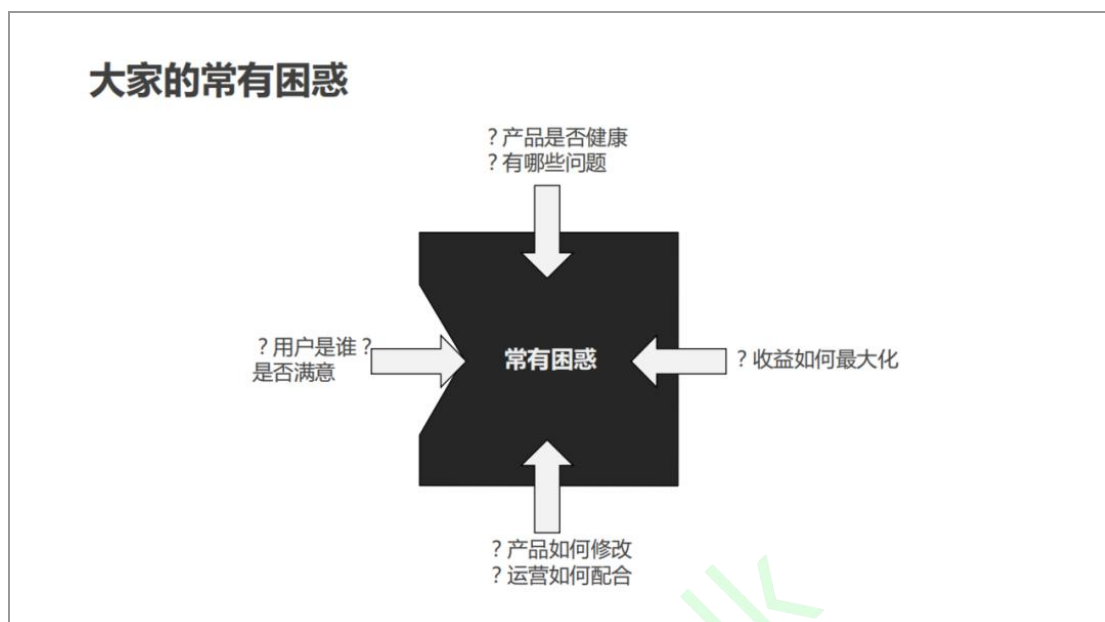


首先我们来看一下什么是互联网公司的核心。鄙人的浅见是围绕商品、供应链、物流、营销活动等的整个购物旅程背后的用户体验的。



既然用户体验非常重要，那如何去度量和优化整个流程呢，那就是站在用户角度收集其在各个模块的数据，并利用统计、概率思维建模分析；在产品运营、增长过程中找到雪球效应的撬动点施以影响，最终建立起良性的闭环。





作为互联网从业人员,其实我们常有一些困惑,产品的用户是谁?他们是否满意?产品是否健康?有哪些问题?如何才能服务好我们的用户,并且让彼此的收益最大化?

如何去构建起更多用户使用的产品,如何创作一个更多用户喜欢的课程?就像小破站(B站)的成长中一直在解决问题,如何突破二次元的定义,如何将用户圈层拉开,如何让尽量多的人能在B站找到自己喜欢的内容,如何做到存量用户保有及增量用户拉取。当然肯定不存在每个人都满意的产品和课程,但是如何在现有体系下去最大化满意度呢?这个我觉得是大家希望得到答案的根本。



其中一种手段

以增量用户拉取，存量用户数量 and 价值的保有为目标，利用大数据技术精准**用户画像**，并结合**商品画像**，**解构**用户及商品，**精细化**耕作，最终在重点目标客户群体上形成突破。

那其中的一个手段就是以增量用户拉取，存量用户数量 and 价值的保有为目标，利用大数据技术精准用户画像，并结合商品画像，解构用户及商品，精细化耕作，最终在重点目标客户群体上形成突破。比如电商猜你喜欢中的个性化推荐技术，万能的淘宝上有海量的商品几乎可以满足每个人的采购商品需求，但是如何让一个用户可以精准快速地找到是大问题，个性化是其根本，用户画像和塑造则是个性化的根本，希望可以用户感受到产品为 ta 而造的满足感。

如何解构用户呢，举一个例子就是说当你很了解某一个人的时候，你跟 ta 一起去吃饭，你如果在之前历次跟 ta 的进餐中收集了 ta 喜欢及不喜欢吃什么(用户动态数据)，然后 ta 的性格是怎么样的(比较犹豫)(用户数据)，是不是喜欢吃辣(用户数据)，然后你再结合有哪些菜、是不是辣的(商品静态数据)，两端做匹配就可以提供更好的体验了。所以不是单单用户画像，商品画像或其它维度也是需要，最重要的是匹配。



目录

- 什么是商品&用户画像(what)
- 为什么需要商品&用户画像/精细化运营 (why)
- 如何构建商品&用户画像(how)
- 举例(方法论)
- 总结

然后我们来看看目录，第一部分是说 what，就是什么是用户画像；第二部分是说 why，为什么我们需要用户画像以及精细化运营；然后第三部分是 how，我们如何构建商品和用户画像；第四部分是举例，是一些方法论的介绍；最后，我们做一个整体的总结。

课程目标

- 了解商品&用户画像及构建方法
- 熟悉常见商品&用户画像应用
- 尝试结合工作做一些实践



课程目标是希望通过分享，首先让大家去了解用户画像和商品画像，整体概念框架以及构建的一些方法；熟悉常见的商品和用户画像的一些应用；尝试结合自己的工作做一些实践。出于对于过往公司数据及技术的保密，无法直接分享给大家特别细粒度且对应的数据。所以分享过程中我会结合业界的一些案例，以及其他前辈公开在网上的分享内容和数据，帮助大家去理解整个应用。

01 什么是商品/用户画像

什么是用户画像

- A user profile is a visual display of personal data associated with a specific user, or a customized desktop environment
- 用户画像就是与该用户相关联的数据的可视化的展现；**一句话**来总结就是：**用户信息标签化**
- 用户画像可以简单理解成是海量数据的标签，根据用户的目标、行为和观点的差异，将他们区分为不同的类型，然后每种类型中抽取出典型特征，赋予名字、照片、一些人口统计学要素、场景等描述，形成了一个人物原型 (personas)。

首先我们来看一下什么是用户画像，根据维基百科的定义，用户画像就是与该用户相关联的数据的可视化的展现，然后一句话来总结的话就是用户信息标签化。从用户海量的信息里面去找到一些标签，为用户去贴上这些标签，当然这些标签的来源就是一些用户的行为。



什么是用户画像



举个范冰冰的例子，客户这个实体通过客户信息的收集，映射到一个客户的画像，最终通过这些画像来建立认知，比如说范冰冰是一个演员，是一个女性，有参加过哪些电影节等。当然你会问那上面的这些标签有什么用呢？这会根据具体的业务场景来定，有些标签在 X 场景下它是没用的，在 Y 场景下则可能非常有用，当然这也提前说明了画像构建的过程肯定是需要与业务结合的展开。



常见用户画像

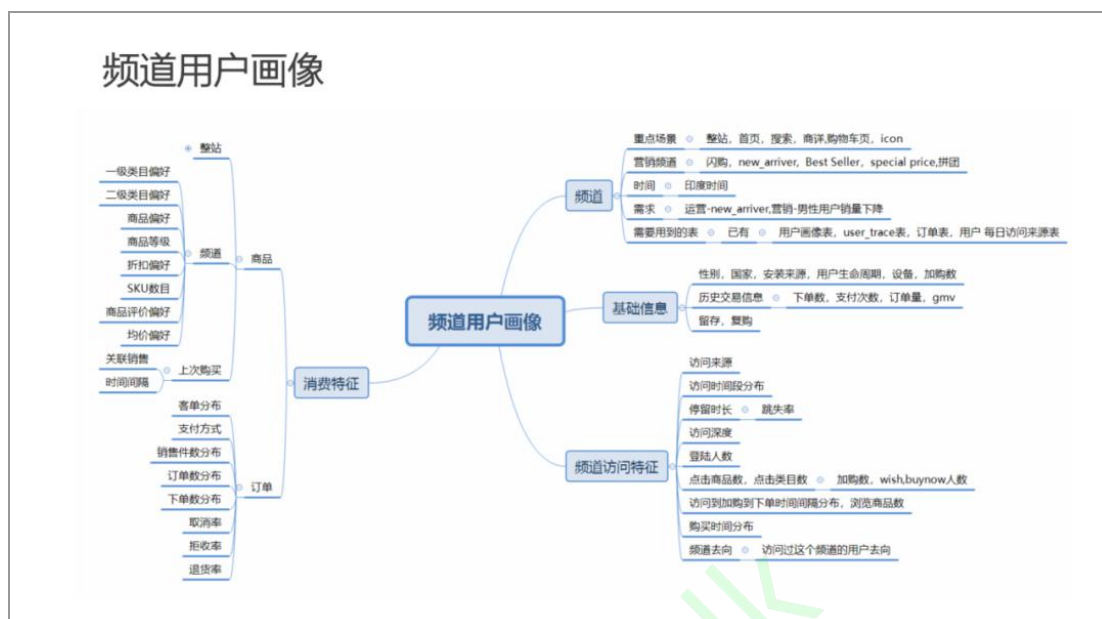


常见用户画像



比较常见的画像是我们需要建立平台客户群体化认知时，我们需要借助可视化的标签，我们看一下京东的一个食品用户画像。对于食品这个业务场景，平台上购买的用户他们是长成什么样的？比如性别比例，蓝领与白领占比，评价敏感的人的占比等，通过拉平垂类偏好用户画像与全站的对比，来寻找差异点和优化点。





举个例子来说，你是抢购秒杀频道的运营负责人，你通过类似的上述分析，找到了秒杀频道和全站的用户画像差异点，那有什么用呢？比如我们发现秒杀频道女性用户或者大龄女性用户，相对于全站占比高很多，然后通过全站分析出这类用户喜欢购买的品类是 A、B、C；但是频道内我们发现几乎没有 A、B、C 这些品类的商品，或者 A、B、C 这些类目的价格段高了很多，甚至是大量的男性商品。那我们马上就找到了一个优化点，是否可以在该场景做一些品类和商品的调整进行一些测试，可能就会在测试中找到明显的提升。

再比如食品品类下的商品运营如何去拓宽自己的流量池，通过类似上述的分析，比如在一个某个频道发现有非常多的 18-30 岁的男性，并他们在平台有非常多的电子类产品的购买，结合你的专家知识，你的碳酸饮料是否可以考虑进入这个频道呢。



上面的是较直观的应用，进一步其实我们可以借助算法挖掘群体偏好（没有算法基础的可以略过），对频道内商品做进一步改造。比如我们通过算法可以对近 N 天频道内的用户，通过算法进行商品召回和排序生成全站他们购买的商品，再通过用户在频道的活跃度和频道 LTV 对这些商品通过 embedding 后的向量进行归约，形成最终的 300 或者 500 个商品池，这些商品是你典型用户的一个集合，可以与你现有的商品集合进行融合测试，会有不错的收益。用户消费购物非常复杂，需要根据具体的业务场景去展开的，上述方法不一定适用所有的业务，不同的业务场景需要做调整和适配。

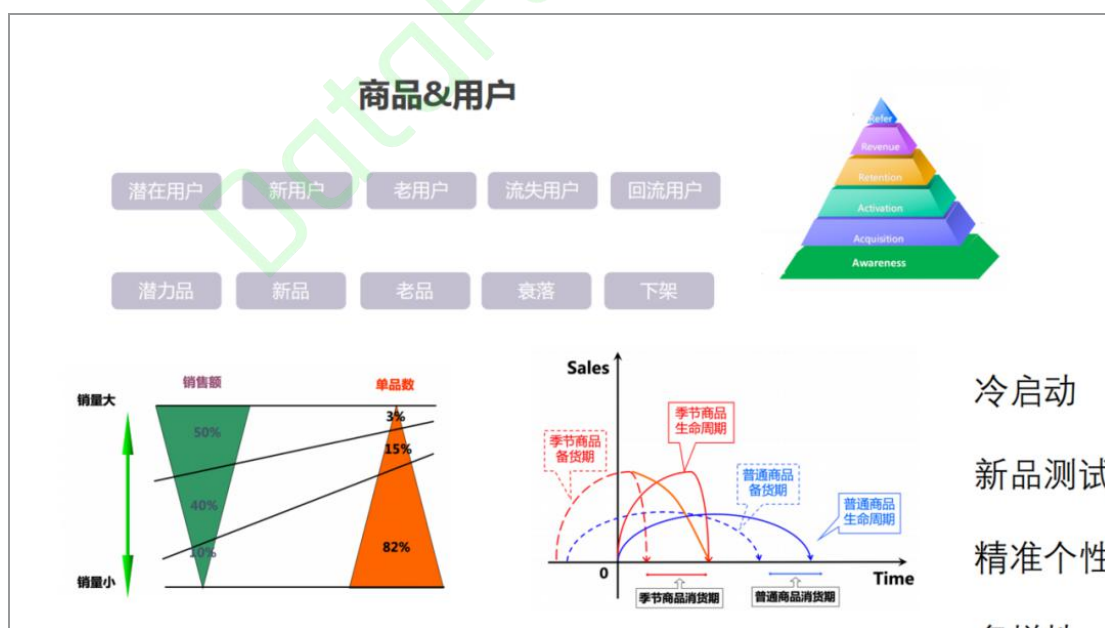
什么是商品画像

- 商品画像如同用户画像一样，可以简单理解成是商品海量数据的标签，根据商品的特征、设计、功能、口味、波次、价位段、流行度、销售状况、促销力度、销售渠道差异，将他们区分为不同的类型，然后每种类型中抽取出典型特征，赋予名字、特点、场景、统计学要素等描述。

顺便介绍一下什么是商品画像，可以简单的理解为商品画像如同用户画像一样，可以简单理解成是商品海量数据的标签，根据商品的特征、设计、功能、口味、波次、价位段、流行度、销售状况、促销力度、销售渠道差异，将他们区分为不



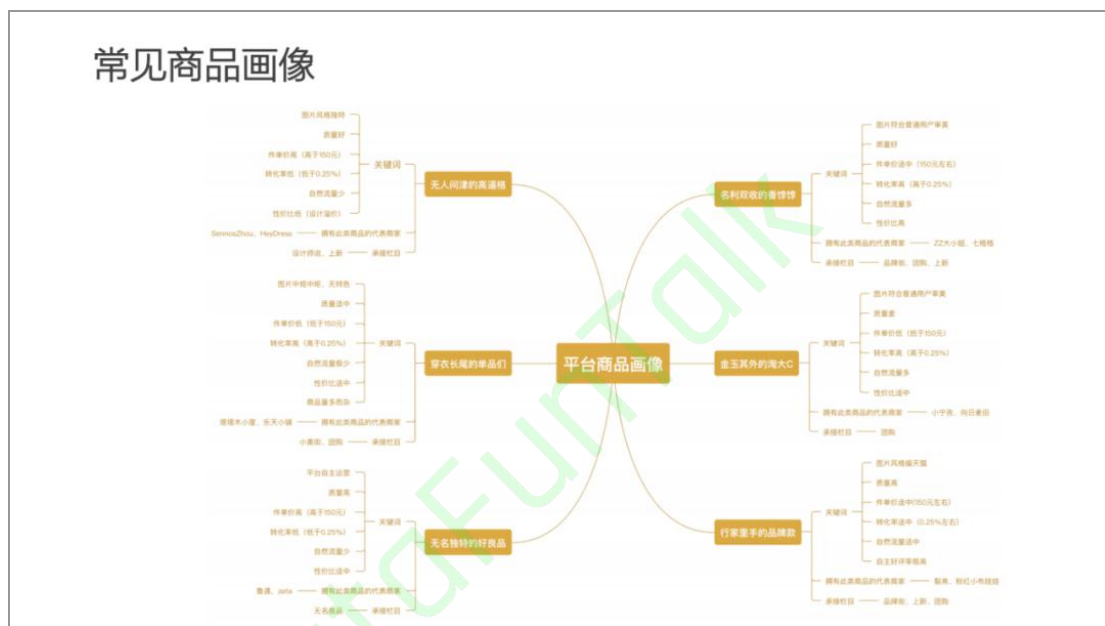
同的类型，然后每种类型中抽取出典型特征，赋予名字、特点、场景、统计学要素等描述。



其实商品和用户画像有很多交集和互相交互的，比如商品有个标签是适用人群，比如女性或者男性，可能就是有很多女性用户经常购买，或者说用户经常购买某



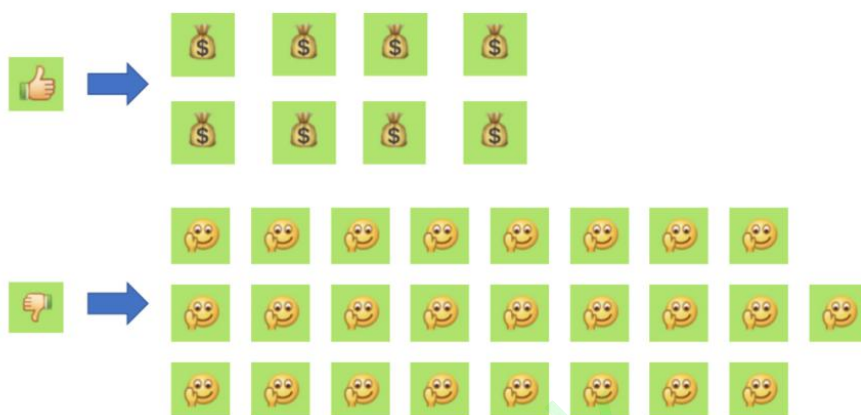
个品类或者某种颜色的商品，那用户就可以贴上类似的标签。商品有了标签，运营在做活动的时候，就可以根据活动首页和分会场人群定位，去选品、氛围塑造，比如时尚会场的商品池选取，比如对于履约不好(拒收、退货、低评分)的、高点击&低转化商品的控制。



02 为什么需要商品/用户画像



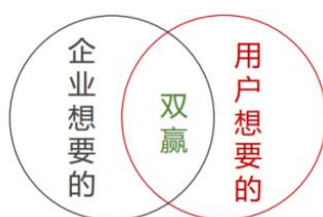
为什么需要商品&用户画像 (why)



为什么需要用户画像呢？其实在用户使用产品的过程中，如果我们做了对的事情让用户的体验非常棒，然后其实我们就可以有不错的收益。但如果我们做了用户体验较差的事情，损失会非常巨大。

为什么需要商品&用户画像 (why)

做用户满意的产品



数据驱动决策



如何在企业商业化过程中与用户达成双赢是需要逐步来探索的，一种途径是通过数据去驱动决策，做到利益最大化，也就是这里分享的。通过用户和商品画像找到并更好地服务为你产品买单买单的那一群人。

为谁服务

- 1.为谁创造价值: 建立对**客户**的尽量全面的**理解**
 - 2.把握: 学会**识别**用户的解决方案和用户在场景下的**诉求**。
- 永远不要幻想你的用户能告诉你，他们真正需要的是什么。
- 持续验证**你的假设

产品十定论

- 人** 用户通过内容形成粉丝，**精准引导及匹配消费**
- 货** 商品走向柔性供应链，数据驱动**产销配一体化**
- 场** 渠道**匹配**品牌和运营能力的节奏布局(含场景)



通过初期形成的产品，形成了我们平台的一些忠实的客户，对于这部分群体的解读，就可以驱动产销配一体化。需求（用户）与供给（商品）两端的匹配 GAP 的存在，可以推动精准的供给体系构建，比如 C2M、C2B 的可能尝试。

用户画像的作用

- 微观
 - 勾画用户与联系用户需求与产品设计，便于**解构场景**内用户
 - 各类数据应用：作为**推荐/搜索/风控**等重要一环而存在
 - 定性和量化方法很好结合在一起的载体，**数据化运营和用户分析**
 - **精准营销和定向投放**，使产品的服务对象更加**聚焦**，更加的专注。
- 宏观
 - 构建**具象认知**以及**战略、战术**方向
 - 探索用户足迹，**市场细分与用户分群**，并以其为导向(市场)

用户画像的作用

应用类型	说明	应用业务效果	应用业务范围
运营	消息推送类精细化运营	★★★★★	★★★★★
商业分析	用户分析决策支持	★★★	★★★★★
搜索	搜索个性化排序模型特征	★★★	★★★★
广告	CTR等指标优化	★★	★
推荐	推荐个性化排序模型特征	★★	★★
展现	千人千面展现	★★	★★★★
反作弊	刷单等作弊识别,风险控制	★	★



宏观和微观上我们看看画像的作用，微观上主要是精细化，比如推荐、搜索、精准营销、定向投放、风控、定量和定性的分析、数据化的运营用户分析。宏观上来说就是具象化对于用户的认知，在探索用户足迹的过程中，建立市场细分和用户分群。

商品画像的作用

- 对商品进行精准的定位，让不同的商品迅速匹配到处在不同地域、时间、偏好、阶层中的消费者，进而去更加优化用户的体验，同时商品画像给商品所贴上的各种“标签”可以驱动后端供应链的各种的行为，如预测、补货、促销、库存、采购、生产、物流等等，都是要和这些标签相匹配，标签不同，模型不同，流程不同，管理的模式也有所不同，而且一切都是动态的。

这里顺便讲一下商品画像的作用，商品画像的前台作用更多是在流量分发上与用户画像的 match，关联后的一些推荐、精准推送等。比如用户标签为 a 类目的价格段偏好，然后商品画像也有价格段的标签，比如说用户偏好的是 30-50 人民币的连衣裙，然后这个商品是连衣裙并且正好是在 30-50 人民币的，所以可能相对于其它价格段的商品而言，在这一维上更合理；接着再结合其它维度的商品和用户标签的 match 就可以构建更多的匹配维度。



而偏后端一些的维度，商品画像也可以带来非常多的帮助，可以去驱动后端的供应链。通过商品画像的历史情况(近期曝光、点击、订单、转化等)，可以去做销量预估，进行备货、采购、越库、生产物流等等环节的优化。

统一建模预测分析

- 人口属性细分
明确“是谁”，“购买了什么”，“为什么”
- 购买行为细分
提供**市场机会**、市场规模等关键信息
- 产品需求细分
提供更具**差异化**竞争力的产品规格和业务价值
- 兴趣态度细分
提供人群类别画像: **渠道策略，定价策略，产品策略，品牌策略**

这里讲一下统一建模预测分析，首先如果有了人口属性细分，比如说一些性别，年龄段等等，就可以去明确用户是谁，他在平台上面买了什么，为什么买；有了购买行为细分以后，可以去明确市场机会，市场规模等关键信息；再次是产品需求的细分，提供根据差异化的竞争的产品规格和业务价值，可以去做更加差异化的竞争，不做海量的库存，而是更精准、成本更可控；最后是兴趣态度的细分，通过用户画像可以去做不同的渠道的策略，定价的策略，产品的策略，品牌策略；比如渠道策略，我们可以通过分渠道归因并捞取对应用户群体，并观察各个渠道群体的差异，比如类目偏好、价格段偏好等的差异，在投放过程中进行商品池和投放策略的优化。



精细化用户分群与品类管理

行为特点/回访留存/群体画像/漏斗转化/商品层级

- 根据用户行为的特征将其按需拆分成不同属性的用户群
- 例如：做过A事件的人拆分成一个用户群，做过B事件的拆分成另一个群，看群体用户画像有什么区别，看他的留存和回访有什么区别。

Why-回顾

商品&用户画像很重要！！！！

用户画像是产品的根基，但是“商品画像”是产品的外延生命力的象征。

两者相辅相成！

总结一下就是商品和用户画像很重要，两者是相辅相成的，在整个用户体验过程中非常重要。



03 如何构建商品/用户画像

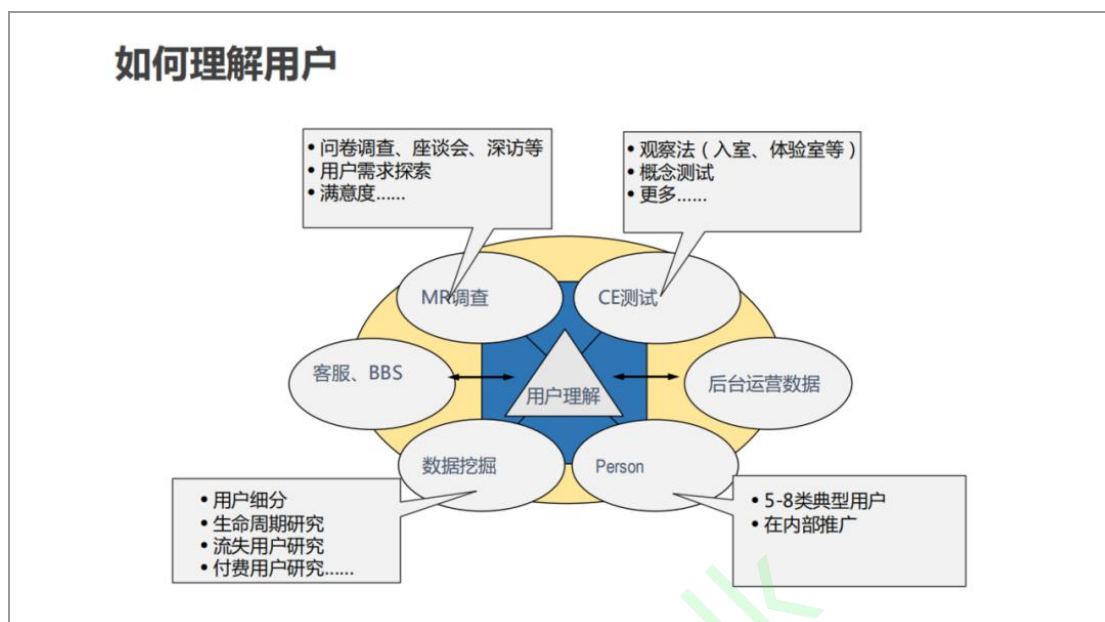
如何构建用户画像

- 侧写师



那到底如何构建用户和商品画像呢，这里可以提一个概念叫侧写师，电影中经常有一些概念就是侧写师通过观察和收集犯罪现场的信息（数据收集），去想象作案人的心理和案发时的动作（预测），并最终为破案带来机会(实施)，构建用户画像的方法论与其类似。





第一步是收集数据，图中方法比较多，各有优缺点，不过目前主流的还是通过用户静态数据结合海量的平台用户行为数据的方式。通过埋点，我们可以看到用户在我们平台上看了什么商品，点了什么商品买了什么商品，将这些数据串起来，可以作为入口去理解用户，更进一步，通过一些统计、概率知识进行建模，比如说用户细分、生命周期划分、流失用户研究等，做深层次的挖掘。接下来我们来解构用户画像。

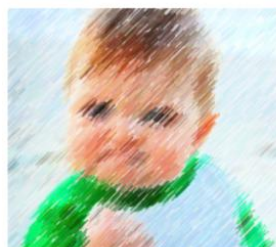
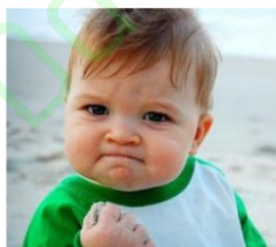


用户画像-维度

- 人口属性——用户是谁（性别、年龄、职业等个人基本信息）
- 消费需求——消费习惯和消费偏好
- 购买能力——收入及购买力、购买频次和渠道
- 兴趣爱好——品牌偏好、个人兴趣
- 社交属性——用户活跃场景（社交媒体等）

主要是人口统计学属性、消费需求、购买能力、兴趣爱好、社交属性等

用户画像-粒度



用户画像的粒度简单来讲，比如年龄标签是 20-30 岁和 21 岁，这两个就是明显不同粒度的标签



用户画像-层级



用户画像-标签层级



再次就是从层级的角度来看，用户基本属性和行为标签为浅层用户画像，这些标签画像主要是收集就行，比如年龄，性别。第二层是稍作处理的汇总层，比如手机端大概什么时间活跃，经常购买的品类等。第三层是基于前2层可以做到



营销的敏感度，社交关系，上网时间的预测；最后一个更是更深层的，需要结合业务知识去做定制。比如保险领域的高中低价值、贷款需求程度，风险高低等。

用户画像-标签层级					
行为属性	上网习惯	活跃情况		模型标签	
		活跃情况-老用户	账号开通以来，发生以下之一的业务： 1. 发生a行为至少1次。 2. 发生b行为至少1次。 3. 发生c行为至少1次。		
用户分类	人群属性	活跃情况-流失用户	属于老用户，但不符合以下条件之一： 1. 过去30天时间里，发生a行为1次。 2. 过去30天时间里，发生b行为1次。		
		活跃情况-微信48小时活跃粉丝	符合微信活跃条件，48小时进行以下操作： 1. 新关注 2. 点击自定义菜单 3. 发送消息 4. 扫描二维码 5. 支付成功 6. 用户授权	事实标签	
商业属性		年龄阶段	年龄阶段-80后 出生时间：1980-1989 年龄阶段-90后 出生时间：1990-1999	事实标签	
		地区分布	地区分布-xx 选择城市	事实标签	
		电商业务	购买额度-高频用户 过去12月内，累计订单数超过24 购买额度-中频用户 过去12月内，累计订单数5-24 购买额度-低频用户 过去12月内，累计订单数小于5 购买额度-新用户 至今，累计订单数为0	模型标签	
		金融支付	支付额度-高频用户 过去30日内，累计支付笔数大于150 支付额度-中频用户 过去30日内，累计支付笔数在20-150 支付额度-低频用户 过去30日内，累计支付笔数小于20 支付额度-新用户 至今，支付笔数为0	模型标签	
		消费订单比例-消费狂	消费订单比例高于60%或过去30日内，超过30件	模型标签	
		消费订单比例-消费达人	消费订单比例达到在20-60%或过去30日内，在10-30件之间	模型标签	
		消费订单比例-普通者	消费订单比例达到低于10%或过去30日内低于10件	模型标签	
		充值-充值新用户	至今，未充值过	模型标签	

网上找了一个金融领域的例子，比如根据活跃情况，可以区分活跃用户和流失用户，再根据时间窗口维度可以拆为48小时活跃和不活跃用户。简单的统计就可以为用户贴上各类的标签，比如过去12个月内累计订单超过20次为界限，可以划分为高低频用户等。不过多展开，可以自行阅读上图。



常见用户画像标签

基本特征	社会身份	顾客用户生命周期	类目偏好	购物属性	风险控制	其它偏好	族群
<ul style="list-style-type: none"> 性别 母婴年龄预测 顾客消费层级 顾客年龄 地域气候 	<ul style="list-style-type: none"> 家庭用户 学生 公司白领 中老年人 顾客职业的行业 	<ul style="list-style-type: none"> 注册用户转新客 PC转移动 类目半新客转化 流失得分 	<ul style="list-style-type: none"> 果粉 吃货 高品质生活 家庭日用品 手机数码达人 礼物礼券 	<ul style="list-style-type: none"> 跨区域购买用户 日用品周期购买 顾客价值得分 促销敏感 辣妈、丽人 	<ul style="list-style-type: none"> 黄牛小号判别得分 注册异常用户判别得分 积分获取异常用户得分 	<ul style="list-style-type: none"> 购买力 关键词 品牌 单品 店铺 销量 B/C 	<ul style="list-style-type: none"> 用户基于语义空间的聚类 用户活跃度得分 达人 微群

类目标签（主题推荐）

女装	饼干/糕点	茶叶	流行首饰	身体护理	公共
<ul style="list-style-type: none"> 甜美文艺 职业通勤 个性街头 妩媚性感 气质名媛 	<ul style="list-style-type: none"> 三高人群 瘦身减脂 独爱花香 香甜 鲜咸 	<ul style="list-style-type: none"> 清热解暑 补血益气 清肝明目 呵护女性 健胃消食 	<ul style="list-style-type: none"> 恋恋深情 卡通图案 平安 乔迁 金饰 	<ul style="list-style-type: none"> 抗敏感 滋润型 中草药 清香型 防晒隔离 	<ul style="list-style-type: none"> 儿时回忆 懒人必备 便携旅游 送礼必备 宴会待客

用户画像-微观画像事例

标签基于银行业务需求、场景进行构建，划分主题、颗粒度等



当然用户画像依赖的商品基础标签是必不可少的，比如用户服装偏好下有颜色偏好(白色)，这个是通过用户历史的点击、加购、收藏等行为关联的商品所带标签的聚合，那其实底层商品的颜色标签是必不可少的。



接下来介绍的是一个构建方法论，但不是唯一方法论。

如何构建用户画像

- 用户静态/动态画像
 - 静态画像-评估价值:用户**相对稳定**的信息
 - 动态画像-循迹: 用户**不断变化**的行为信息，根据用户行为将物品的结构化结果传递给用户
- 统计/预测类画像
 - 每天看新闻类型占比
 - 用户的性别，生理性别与行为性别

先从大维度来说，划分为动态和静态两类。比如说人口统计学的性别、年龄、地域、收入、生活习惯，是相对静态的，性别一般都不会变，年龄的话一年就涨一岁，然后地域的话也很少有大的变动，收入职业所处行业也不会有突然的变化，这些都可以划分到静态标签。而像用户访问设备、用户的 48 小时是否活跃、内容&商品消费偏好等属于时常在发生变动的，这些可以划入动态特征。



如何构建用户画像

- 静态画像：**实时性弱，覆盖广，粒度粗**
- 动态画像：
 - 一般建立兴趣模型，更加实时并且动态
 - 用户行为可以解构为**5w**
 - 时效性非常敏感
 - 在空间上，不同应用领域侧重点不同，**营销领域更侧重消费习惯，推荐领域更侧重喜好**

静态和动态的划分，其实是根据某个维度来定的，或者是在某个时间窗口内的。

静态画像一般是实时性弱更新慢；动态画像变化相对更快一些，可以再时效性、空间上有差异。

如何构建用户画像

- 不同行业、不同业务场景对用户画像需求大不同。
- 目标相关的
 - 产品/品牌相关
 - 商业目标相关
 - 项目目标相关



如何构建用户画像



不同的行业，不同的业务，也会有不一样的构建目标和方式。在这个过程中，通过业务应用的要素分析，去驱动标签维度的扩展，然后通过标签维度的扩展。然后我又可以去驱动业务。

如何构建用户画像

- 偏好画像数据建模方法:
 - 标签=用户标识+时间+行为类型+接触点（网址+内容）的聚合
 - 事件模型: 主要通过收集用户行为，并结合上下文构建事件模型，主要为5w
 - 整体思考建模: 用户标签的权重可能随时间的增加而衰减，因此定义时间为衰减因子 r ，行为类型、网址决定了权重，内容决定了标签，进一步转换为公式：**标签权重 = 衰减因子 × 行为权重**



用户画像-偏好类标签

- 需求量最大
- 通常只反映用户一段时间内的兴趣点，并有可能**随着时间的推移发生改变**，比如衣服尺码
- 无法精确的用0和1表示
- 一般与**业务强相关**，哪种行为权重更大，时间窗口多长比较合适
- 无复杂逻辑，加工过程和流程非常相似：加权汇总一段时间内的几种行为后，归一化到[0,1]
- 如彩妆功能偏好，美妆肤质偏好，洗发护发功效偏好，品牌偏好，类目偏好

标签系统

•事实标签:

用户的基本属性、行为频次等基于事实及简单行为统计的标签,如用户的基本属性、行为标签都属于这类.

•模型标签:

利用模型(业务规则/算法模型),根据业务需求构造出一些标签,如大部分的偏好标签采用的加权模型:

$$\sum_{i=1}^n w_i \epsilon_i$$
$$\epsilon_i = e^{-d_i/\alpha}$$

w : 行为类型权重, ϵ : 时间衰减, d : 日期差, α : 可调参数

难点: 权重怎么定? -> 目前主要还是根据业务经验逻辑来定.

比如在类目偏好中, 各行为权重降序: 下单 > buynow > 加购 > 收藏 > 点击.



用户画像-实时风格偏好预测

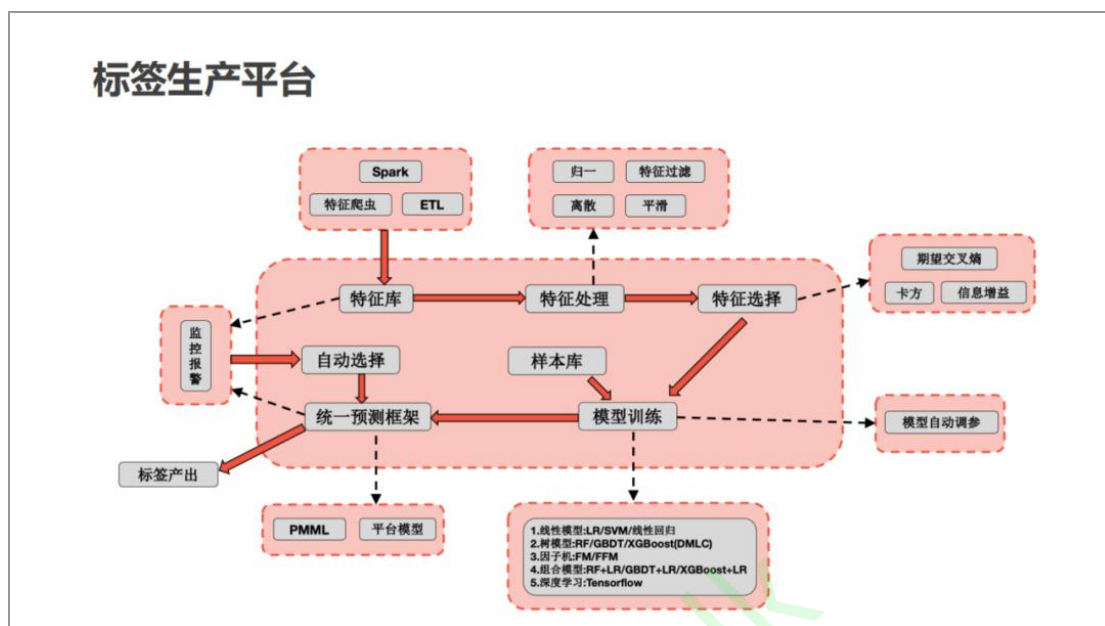
- 实时风格偏好预测
 - 除了性别、年龄、购买力等，一些抽象的兴趣点也非常有帮助
 - 用户**行为和兴趣的多变性**，长期偏好抽取的方式无法满足用户的需求，需实时的抽取用户的风格偏好

用户画像-基于标签聚类的人群生成（圈人）

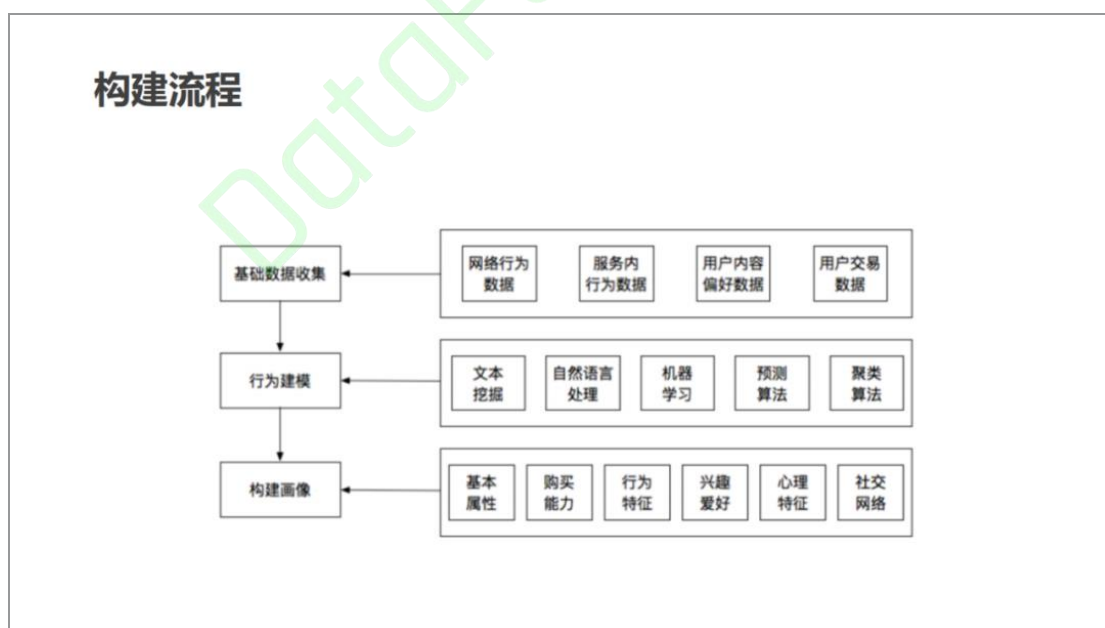
- 实现方案
 - 按维度（n天内）加权汇总某类主题和实体（买家）在某种对象（品牌名）上的相关行为（点击，收藏，加购，交易），然后归一化到[0,1]之间，取TopN或全部输出。
 - 可以调整的为维度、主体、对象、行为、权重等
 - 并结合时间衰减，使越久远的历史数据影响越小

上述图片介绍了常用的偏好标签建模的思路，不做过多展开，有兴趣的小伙伴可以私信交流。





那系统化的标签生产如何来做呢，可以参考上图，上图元素和参考了很多美团的做法，由于篇幅限制，这里不过多展开，后续的分可以专门来将构建的细节。



如何构建用户画像



重点是结合业务，构建闭环，用数据来提升业务理解，通过业务理解来快速构建画像相关的数据体系。

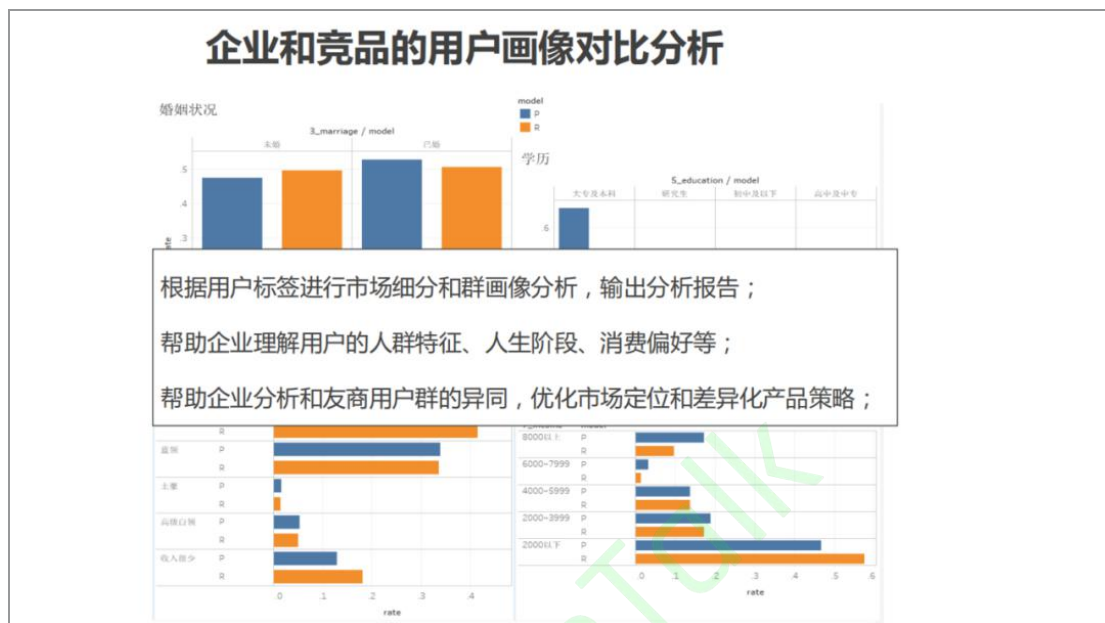
How-回顾

- 维度/粒度
- 动态/静态
- 常见画像及层级
- 构建方法举例
- 闭环/实践出真知思维

04 商品/用户画像的应用



第四部分我们来介绍一些画像的应用。



第一个我们介绍一下通过用户画像来做与竞品对比分析，在大维度战略调整和选择的时候，可以有差异化市场定位。根据用户标签进行市场细分和群画像分析，输出分析报告；帮助企业理解用户的人群特征、人生阶段、消费偏好等；帮助企业分析和友商用户群的异同，优化市场定位和差异化产品策略。这部分数据有自己内部的数据，竞品数据可以通过第三方等间接获取。

比如今天 shein 作为跨境快时尚平台做得不错，如果同类型的平台是否可以考虑在年龄段标签上做差异化，比如他们切的是 18-35 岁的年轻女性，那是否可以考虑 35-50 岁的中年女性作为用户群，在运营过程中做调整，反而在利基点上有不错的收益。



另一个方向的思考是，如果我们在部分用户标签维度差异较大，如果我们就是希望拿下这部分用户，那就说明我们可能有做的不对的地方，通过数据去排查问题，比如是我们投放的过程中渠道选择带来这样的差异？货是不是有问题，价格过高？品类深度不够？认识到了这种差异，可以进一步深挖分析或者用户访谈，来对我们的产品给用户的心智营造做调整。

用户画像-应用

- **竞品**

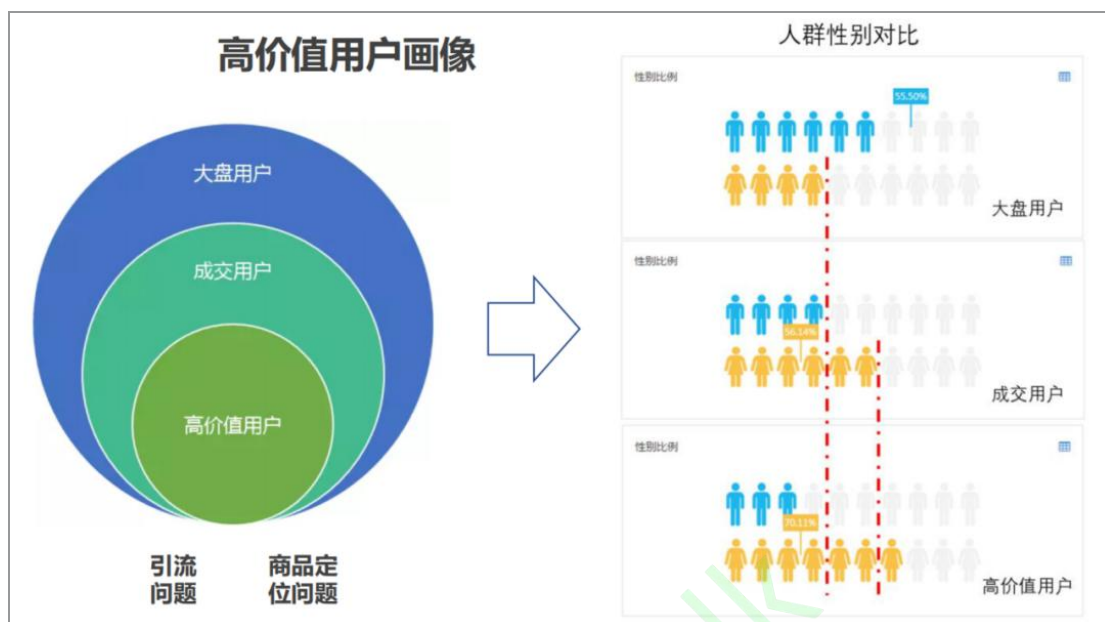
- 人群分析后，你知道了你真正的竞品是谁

- **品牌刚刚建立**

需要快速提升知名度，可以按照不同媒体目标人群覆盖率的高低进行预算分配

通过画像数据的对比，可以去确定一件事情，就是说我们的竞品到底是谁。真正的竞品突然在你眼前，画像跟你高度重合的那个，而不是我们假想的那个。如果你发现原来被认为是竞品的典型用户群体是在 30 到 50 岁之间，而你们的群体在 20 岁到 35 岁了，那至少在当前情况下，你们是弱竞争关系。

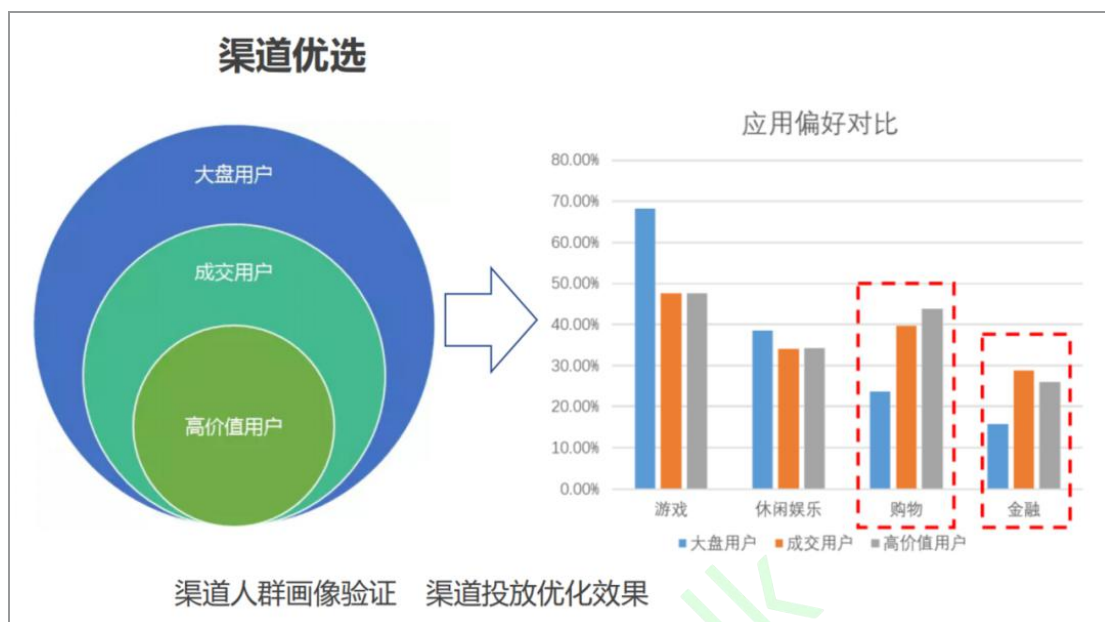




定位了不同的人群，并结合数据做一定划分后，我们也可以根据情况针对性做投放和媒体选取。在广告营销推广的时候，我们结合画像对用户进行划分，区分为大盘->成交->高价值用户，然后通过某些标签进行数据分析，如上图中的性别维度，我们对比中发现其实女性用户比例在高价值用户中与大盘用户中刚好反过来。

那接下来就有几个动作，可以对投放渠道、商品进行分析，是引流的问题？还是货品上对于男性用户天然不高？商品运营有没有问题？如果不做调整，是否后期的投放渠道可以选取女性标签会更精准？





我们希望根据不同的媒体目标覆盖率高低进行一个预算的分配，在平台上人群通过标签进行拆分。上面的问题，我们拆分了性别，是否有其它的维度可以进一步拆分，比如通过用户上报的 APP 安装列表，在平台上高价值用户有多种购物类 APP 和金融类 APP 的安装，那在投放过程中可以根据渠道的差异进行筛选，或者在投放过程中选取对应的一些标签词进行尝试。



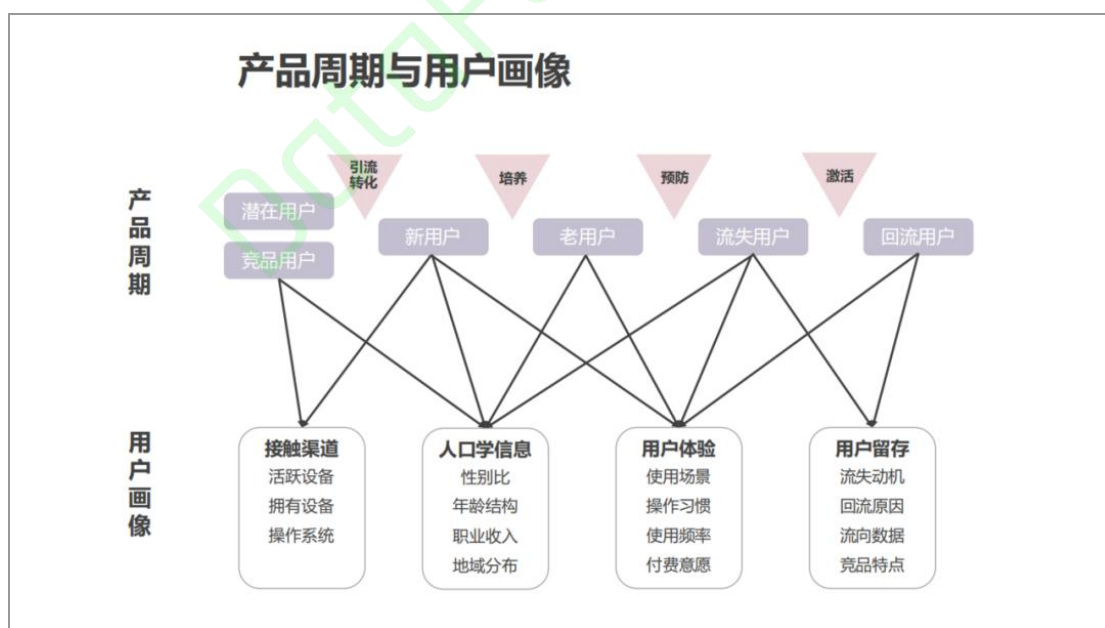
渠道优选

引流渠道进行优化

- 使用上述分析结论，进行策略制定
- 对各个渠道的量进行重新分配，并实验

商品定位调整

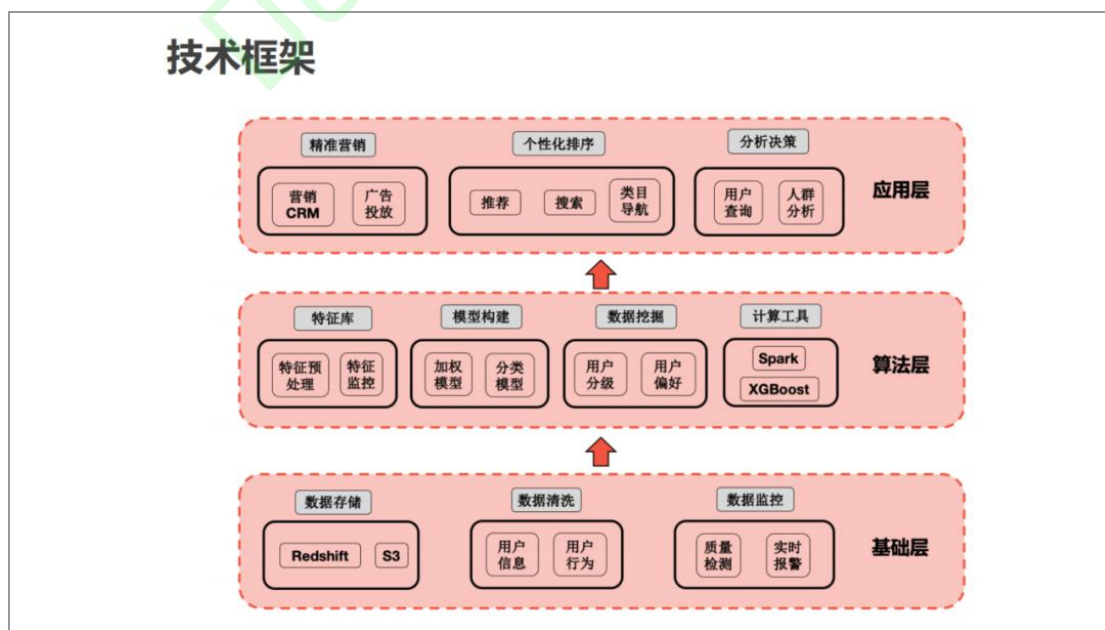
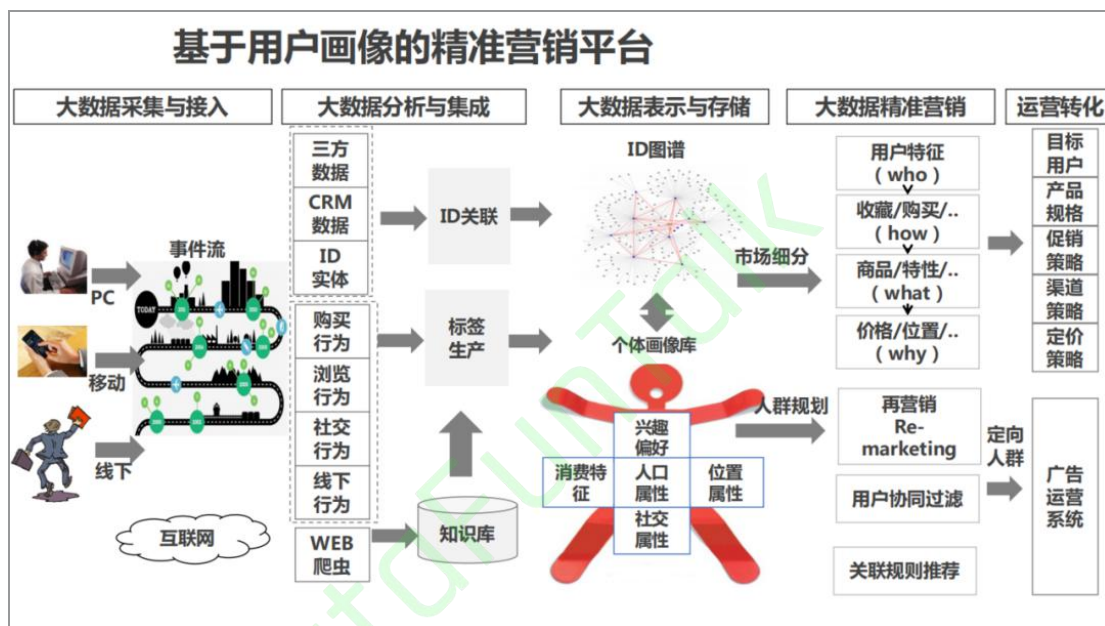
上面的例子是结合用户画像做不同渠道的选择，逐步测试和优化。而对于商品的定位调整，可以结合商品生命的周期，尝试做一些品类深度与宽度的运营。



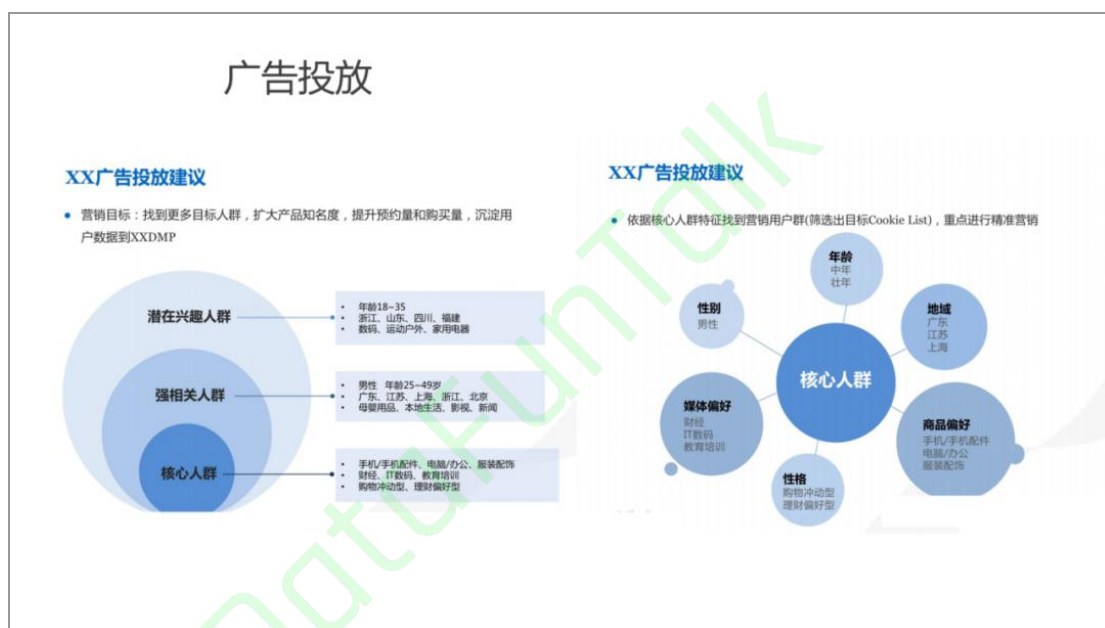
对于不同生命周期内的用户，其实数据和画像的丰富度也会有差异，在精细化运营过程中可以根据情况做差异化的运营。这里再讲一个其他案例，在用户分析维



度，初期会通过用户调研访谈进行初期的产品定位，在投放扩量过程中，结合用户站内的情况作出效率分析，新用户有什么特征？核心的用户属性是否有变化？前后一年内获客可能会有很大差异，这种差异时常去关注，可能会了解到产品的变化是否符合预期。



在精准营销领域的的一些尝试，可以参考上图。分析历史某个时间窗口内的用户，哪些方面是有特性的，比如年龄段，性别，地域，收入，偏好等，有了这样的分析，可以选择相应的人群投放。输出到媒体广告，做预算看效果，然后将效果数据回流，再去迭代，以提高进一步的准确率。通过某个特性找人，再去分析这些潜在用户，找到可以触达他们的渠道。



广告投放建议

XX广告投放建议

强相关人群特征：

- 性别：男性
- 年龄：中年、壮年（25~49岁）
- 地域：重点投放广东、江苏、上海、浙江、北京等地
- 商品偏好：母婴用品、本地生活
- 媒体偏好：母婴育儿、影视、新闻

此类用户最有可能转化成XX核心人群，建议加大营销力度。

XX广告投放建议

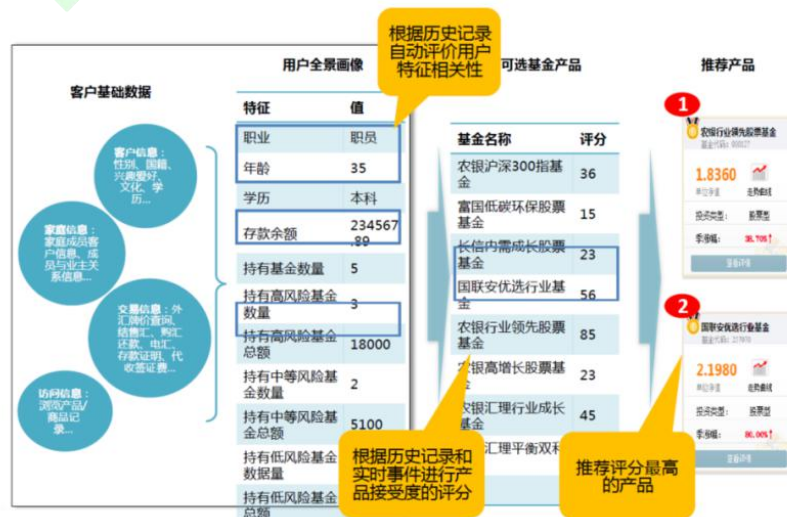
潜在兴趣人群特征：

- 年龄：XX在青壮年用户群覆盖度不够，明显低于小米和苹果，建议重点对青壮年人群进行营销，同时可以考虑与耐克、阿迪达斯一起合作吸引青年用户群
- 地域：重点浙江、山东、四川、福建等地，同时扩展到全国地区
- 商品偏好：数码、运动户外、家用电器
- 媒体偏好：时尚、历史、医疗健康

扩大目标用户范围，通过营销数据沉淀出XX核心用户群和强相关用户群。

上图也跟前面介绍的有点类似，作用类似，就是通过筛选已有用户中的核心用户，结合这群用户的标签，进行投放标签的初步筛选。在成本需要控制的情况下，可以通过核心人群标签投放->强相关人群标签投放->一般用户标签投放，这里也需要结合不同人群经常交互的商品进行商品池选取。

基于画像的基金推荐流程



在看一个金融领域的应用案例，根据职业、年龄、存款等进行筛选，并结合历史理财的偏好等进行基金推荐。

用户画像应用

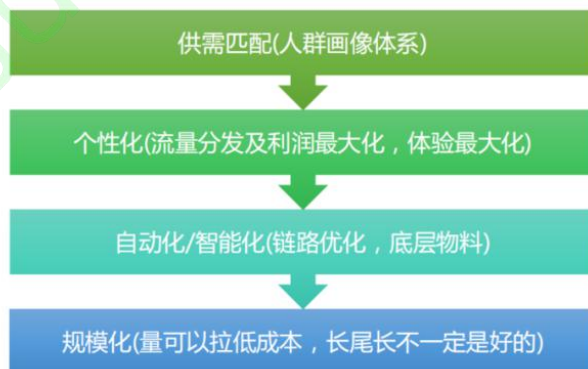
- 精准营销：

从**粗放式到精细化**，将用户群体**切割成更细的粒度**，辅以短信、推送、邮件、活动等手段，驱以**关怀、挽回、激励**等策略。

- 数据应用：

推荐/搜索/广告系统。操作过各大广告投放系统的同学想必都清楚，广告投放基于一系列人口统计相关的标签，性别、年龄、学历、兴趣偏好、手机等等。

优化链



推送

条件选择：☒ 实时+离线 ☐ 实时

实时查询条件：

用户基本信息 性别、用户长期所在省份、用户...	用户近期行为	用户天气感受	用户公司、校园
用户小区类型			

离线查询条件：

用户行为维度	用户行为指标	用户会员信息
--------	--------	--------

任务名称：

当前查询条件：

性别：男 <input checked="" type="checkbox"/>	用户长期所在省份：河北、重庆 <input checked="" type="checkbox"/>	用户长期所在城市：石家庄市、衡水市 <input checked="" type="checkbox"/>
--	--	---

人群定向

人群定向

时间维度：

用户行为：

类目维度：

行为频次： 次

性别：☐ 男 ☐ 女

促销关注：☐ 敏感 ☐ 不敏感

人群一：
近7天搜索，点击，浏览过食品类目X次男性

人群二：
近7天搜索，点击，浏览过食品类目X次男性

人群三：
近7天搜索，点击，浏览过食品类目X次男性

人群四：
近7天搜索，点击，浏览过食品类目X次男性

人群五：
近7天搜索，点击，浏览过食品类目X次男性



CRM常见问题

- 现状1: CRM库里有千万的会员资料，但不知道怎么用-会员行为标签少
 - 没有统一的标签采集维度
 - 缺少分析模型
 - 数据杂乱，未作画像匹配现状
- 现状2: 花了很多钱做推广营销，流量抓到了，用户留存始终很低
 - 缺少用户生命周期的系统规划
 - 用户培育运营成本较高，企业没有动力投入
 - 缺少用户画像，难以系统地进行用户培育现状
- 现状3: 花了很多钱做粉丝福利，可是粉丝转化为忠诚会员的比例还是很低
 - 缺少用户生命周期的系统规划
 - 抓不到用户痛点，促销机制不能满足核心用户需求
 - 缺少用户画像，难以系统化地进行会员忠诚度管理

再讲一个概念就是 CRM，业界有很多探讨，其中也有一些问题点，如上图，后面分享我们对 CRM 可以专门展开。

CRM整体框架



用户的构成
用户是谁
怎么影响用户
所做的事情效果如何



注意点

- 现有用户是不是等于目标用户？
- 猜用户是男是女，哪里人，工资多少，有没有谈恋爱，喜欢什么，准备剁手购物吗？探讨这些是没有**意义**的。是男是女**如何影响消费决策**，工资多少**影响消费能力**，有没有谈恋爱会否带来**新的营销场景**，剁手购物怎么**精准推荐**，这些才是用户画像**背后的逻辑**。

有几个注意点也需要大家去注意的，其实我们有时候现有的用户不等于目标用户。在产品运营过程中，会发现平台上用户画像是这样的，但是结合战略以及竞品的一些情况，更大的去做差异化的竞争。重合的那部分是高价值用户，所以应该投入更多人力去做？结合用户画像去做更细粒度扩人群，是 A 人群还是 B 人群还是 C 人群？

然后第二点，构建什么样的用户画像，比如说用户是男性还是女性，她工资多少，喜欢什么样的商品。这些标签可能是没有意义的，如果说有意义的地方是这些标签如何去影响决策，如何为用户带来更好的体验，带来更高的留存、复购。

05 回顾/总结



实践出真知，我们需要结合具体业务场景，做尝试。

画像-总结

- **结合业务场景**去分析维度
- 不要简单看画像，一定要做**对比**。单纯看分布是没有太多信息含量的，不对比看不出差异
- 环比的对比，哪些人触达了，哪些人到app落地了，哪些人注册了，哪些人真正浏览了，哪些是留存，哪些是付费，漏斗的**每个环节**都需要分析
- 直觉未必靠谱，一定要**通过反馈来检测**，要有**数据闭环**，
- 活动和用户场景可以做**专属**的标签

我们要结合业务场景去分析，然后去不要单看画像，而是要做一些对比，通过前后对比，跟竞品的对比，跟频道内与大盘的对比等手段去分析，发现不足和优势，做纠正和调整。通过数据反馈形成数据的闭环，最终在产品的迭代过程中拿到更好的业务结果。总结一下，做画像要有目标要有数据，也不拘泥于技术细节，大胆的尝试，然后先粗粒度，后细粒度。

今天的分享就到这里，谢谢大家。



感谢收藏！

