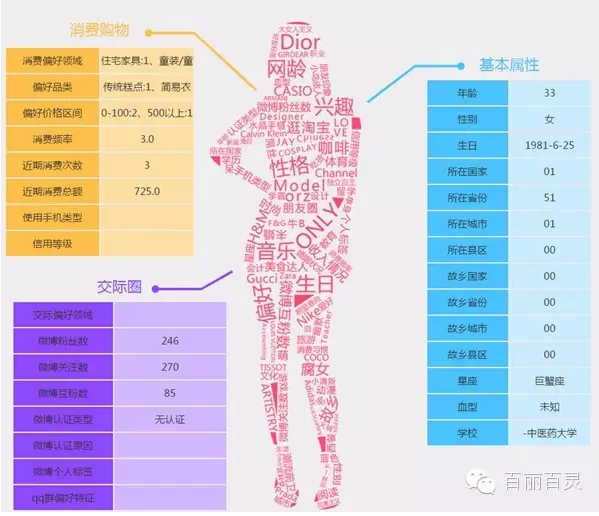
# 1、用户画像是什么

用户画像（User Profile），作为大数据的根基，它完美地抽象出一个用户的信息全貌，为进一步精准、快速地分析用户行为习惯、消费习惯等重要信息，提供了足够的数据基础，奠定了大数据时代的基石。

男，31岁，已婚，收入1万以上，爱美食，团购达人，喜欢红酒配香烟。这样一串描述即为用户画像的典型案例。如果用一句话来描述，即：用户信息标签化。



用户画像，即用户信息标签化，就是企业通过收集与分析消费者社会属性、生活习惯、消费行为等主要信息的数据之后，完美地抽象出一个用户的商业全貌作是企业应用大数据技术的基本方式。用户画像为企业提供了足够的信息基础，能够帮助企业快速找到精准用户群体以及用户需求等更为广泛的反馈信息。

**用户画像有很多的的标签组成**，每个标签都规定了**观察、认识、描述**用户的角度。标签根据企业业务的发展情况，或多或少，对外而言都是一个整体，这个整体称之为用户画像。

# 2、为什么需要用户画像

用户画像的核心工作是为用户打标签，**打标签的重要目的之一是为了让人能够理解并且方便计算机处理**，如，可以做分类统计：喜欢红酒的用户有多少？喜欢红酒的人群中，男、女比例是多少？

也可以做数据挖掘工作：利用关联规则计算，喜欢红酒的人通常喜欢什么运动品牌？利用聚类算法分析，喜欢红酒的人年龄段分布情况？

大数据处理，离不开计算机的运算，标签提供了一种便捷的方式，使得计算机能够程序化处理与人相关的信息，甚至通过算法、模型能够“理解” 人。**当计算机具备这样的能力后，无论是搜索引擎、推荐引擎、广告投放等各种应用领域，都将能进一步提升精准度，提高信息获取的效率。**

****

**用户画像本质就是从业务角度出发对用户进行分析**，了解用户需求，寻找目标客户。另外一个方面就是，金融企业利用统计的信息，开发出适合目标客户的产品。

提到用户画像，很多厂商都会提到**360度用户画像**，其实经常360度客户画像是一个广告宣传用语，根本不存数据可以全面描述客户，透彻了解客户。人是非常复杂的动物，信息纬度非常复杂，仅仅依靠外部信息来刻画客户内心需要根本不可能。

用户画像一词具有很重的场景因素，不同企业对于用户画像有着不同对理解和需求。举个例子，金融行业和汽车行业对于用户画像需求的信息完全不一样，信息纬度也不同，对画像结果要求也不同。每个行业都有一套适合自己行业的用户画像方法，但是其核心都是为客户服务，为业务场景服务。

# 3、用户画像怎么设计

一个标签通常是人为规定的**高度精炼的特征标识**。

如年龄段标签：25~35岁，地域标签：北京，标签呈现出两个重要特征：语义化，人能很方便地理解每个标签含义。这也使得用户画像模型具备实际意义。能够较好的满足业务需求。

如，判断用户偏好。短文本，每个标签通常只表示一种含义，标签本身无需再做过多文本分析等预处理工作，这为利用机器提取标准化信息提供了便利。

人制定标签规则，并能够通过标签快速读出其中的信息，机器方便做标签提取、聚合分析。所以，用户画像，即：用户标签，向我们展示了一种朴素、简洁的方法用于描述用户信息。

## 3.1 数据源分析

构建用户画像是为了还原用户信息，因此数据来源于：所有用户相关的数据。

对于用户相关数据的分类，引入一种重要的分类思想：封闭性的分类方式。如，世界上分为两种人，一种是学英语的人，一种是不学英语的人；客户分三类，高价值客户，中价值客户，低价值客户；产品生命周期分为，投入期、成长期、成熟期、衰退期…所有的子分类将构成了类目空间的全部集合。

这样的分类方式，有助于后续不断枚举并迭代补充遗漏的信息维度。不必担心架构上对每一层分类没有考虑完整，造成维度遗漏留下扩展性隐患。另外，不同的分类方式根据应用场景，业务需求的不同，也许各有道理，按需划分即可。

本文将用户数据划分为静态信息数据、动态信息数据两大类。

### 3.1.1、静态信息数据

用户相对稳定的信息，如图所示，主要包括人口属性、商业属性等方面数据。这类信息，自成标签，如果企业有真实信息则无需过多建模预测，更多的是数据清洗工作，因此这方面信息的数据建模不是本篇文章重点。

### 3.2.2、动态信息数据

用户不断变化的行为信息，如果存在上帝，每一个人的行为都在时刻被上帝那双无形的眼睛监控着，广义上讲，一个用户打开网页，买了一个杯子；与该用户傍晚溜了趟狗，白天取了一次钱，打了一个哈欠等等一样都是上帝眼中的用户行为。当行为集中到互联网，乃至电商，用户行为就会聚焦很多，如上图所示：浏览凡客首页、浏览休闲鞋单品页、搜索帆布鞋、发表关于鞋品质的微博、赞“双十一大促给力”的微博消息。等等均可看作互联网用户行为。

本篇文章以互联网电商用户，为主要分析对象，暂不考虑线下用户行为数据（分析方法雷同，只是数据获取途径，用户识别方式有些差异）。

在互联网上，用户行为，可以看作用户动态信息的唯一数据来源。如何对用户行为数据构建数据模型，分析出用户标签，将是本文着重介绍的内容。

## 3.2 目标分析

用户画像的目标是通过分析用户行为，最终为每个用户打上标签，以及该标签的权重。

如，红酒 0.8、李宁 0.6。

标签，表征了内容，用户对该内容有兴趣、偏好、需求等等。

权重，表征了指数，用户的兴趣、偏好指数，也可能表征用户的需求度，可以简单的理解为可信度，概率。

## 3.3 数据建模方法

下面内容将详细介绍，如何根据用户行为，构建模型产出标签、权重。一个事件模型包括：时间、地点、人物三个要素。每一次用户行为本质上是一次随机事件，可以详细描述为：**什么用户，在什么时间，什么地点，做了什么事**。

### 3.3.1、什么用户

**关键在于对用户的标识，用户标识的目的是为了区分用户、单点定位。**

以上列举了互联网主要的用户标识方法，获取方式由易到难。视企业的用户粘性，可以获取的标识信息有所差异。

### 3.3.2、什么时间

时间包括两个重要信息，时间戳+时间长度。时间戳，为了标识用户行为的时间点，如，1395121950（精度到秒），1395121950.083612（精度到微秒），通常采用精度到秒的时间戳即可。因为微秒的时间戳精度并不可靠。浏览器时间精度，准确度最多也只能到毫秒。时间长度，为了标识用户在某一页面的停留时间。

### 3.3.3、什么地点

用户接触点，Touch Point。对于每个用户接触点。潜在包含了两层信息：网址 + 内容。网址：每一个url链接（页面/屏幕），即定位了一个互联网页面地址，或者某个产品的特定页面。可以是PC上某电商网站的页面url，也可以是手机上的微博，微信等应用某个功能页面，某款产品应用的特定画面。如，长城红酒单品页，微信订阅号页面，某游戏的过关页。

### 3.3.4、什么内容

每个url网址（页面/屏幕）中的内容。可以是单品的相关信息：类别、品牌、描述、属性、网站信息等等。如，红酒，长城，干红，对于每个互联网接触点，其中网址决定了权重；内容决定了标签。

注：接触点可以是网址，也可以是某个产品的特定功能界面。如，同样一瓶矿泉水，超市卖1元，火车上卖3元，景区卖5元。商品的售卖价值，不在于成本，更在于售卖地点。标签均是矿泉水，但接触点的不同体现出了权重差异。这里的权重可以理解为用户对于矿泉水的需求程度不同。即，愿意支付的价值不同。

标签 权重

矿泉水 1 // 超市

矿泉水 3 // 火车

矿泉水 5 // 景区

案例：你是我的优乐美，优乐美用户 促销

类似的，用户在京东商城浏览红酒信息，与在品尚红酒网浏览红酒信息，表现出对红酒喜好度也是有差异的。这里的关注点是不同的网址，存在权重差异，权重模型的构建，需要根据各自的业务需求构建。

所以，网址本身表征了用户的标签偏好权重。网址对应的内容体现了标签信息。

### 3.3.5、什么事

用户行为类型，对于电商有如下典型行为：浏览、添加购物车、搜索、评论、购买、点击赞、收藏 等等。

不同的行为类型，对于接触点的内容产生的标签信息，具有不同的权重。如，购买权重计为5，浏览计为1

红酒 1 // 浏览红酒

红酒 5 // 购买红酒

综合上述分析，用户画像的数据模型，可以概括为下面的公式：用户标识 + 时间 + 行为类型 + 接触点（网址+内容），某用户因为在什么时间、地点、做了什么事。所以会打上\*\*标签。

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×网址子权重

如：用户A，昨天在品尚红酒网浏览一瓶价值238元的长城干红葡萄酒信息。

标签：红酒，长城

时间：因为是昨天的行为，假设衰减因子为：r=0.95

行为类型：浏览行为记为权重1

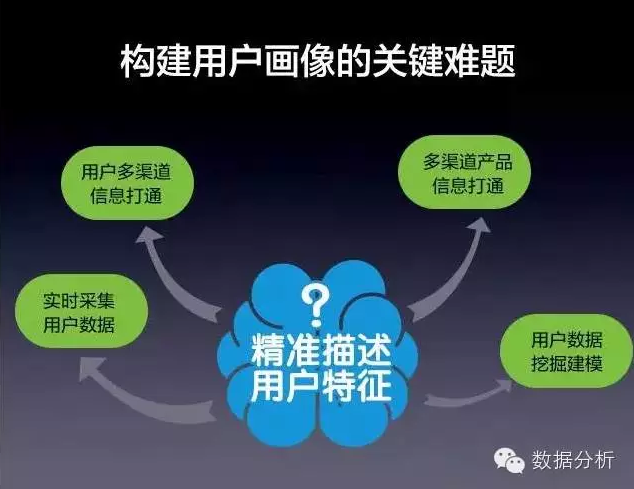
地点：品尚红酒单品页的网址子权重记为 0.9（相比京东红酒单品页的0.7）

假设用户对红酒出于真的喜欢，才会去专业的红酒网选购，而不再综合商城选购。

则用户偏好标签是：红酒，权重是0.95\*0.7 \* 1=0.665，即，用户Ａ：红酒 0.665、长城 0.665。

上述模型权重值的选取只是举例参考，具体的权重值需要根据业务需求二次建模，这里强调的是如何从整体思考，去构建用户画像模型，进而能够逐步细化模型。

# 4、用户画像怎么开发



人口属性、资产特征、营销特性、兴趣爱好、购物爱好、需求特征

# 5、用户画像工作坚持的原则(金融企业)

市场上用户画像的方法很多，许多企业也提供用户画像服务，将用户画像提升到很有逼格一件事。金融企业是最早开始用户画像的行业，由于拥有丰富的数据，金融企业在进行用户画像时，对众多纬度的数据无从下手，总是认为用户画像数据纬度越多越好，画像数据越丰富越好，某些输入的数据还设定了权重甚至建立了模型，搞的用户画像是一个巨大而复杂的工程。但是费力很大力气进行了画像之后，却发现只剩下了用户画像，和业务相聚甚远，没有办法直接支持业务运营，投入精力巨大但是回报微小，可以说是得不偿失，无法向领导交代。

事实上，用户画像涉及数据的纬度需要业务场景结合，既要简单干练又要和业务强相关，既要筛选便捷又要方便进一步操作。用户画像需要坚持三个原则，分别是人口属性和信用信息为主，强相关信息为主，定性数据为主。下面就分别展开进行解释和分析。

## 5.1、信用信息和人口属性为主

描述一个用户的信息很多，信用信息是用户画像中重要的信息，信用信息是描述一个人在社会中的消费能力信息。任何企业进行用户画像的目的是寻找目标客户，其必须是具有潜在消费能力的用户。信用信息可以直接证明客户的消费能力，是用户画像中最重要和基础的信息。一句戏言，所有的信息都是信用信息就是这个道理。其包含消费者工作、收入、学历、财产等信息。

定位完目标客户之后，金融企业需要触达客户，人口属性信息就是起到触达客户的作用，人口属性信息包含姓名、性别，电话号码，邮件地址，家庭住址等信息。这些信息可以帮助金融企业联系客户，将产品和服务推销给客户。

## 5.2、采用强相关信息，忽略弱相关信息

我们需要介绍一下强相关信息和弱相关信息。强相关信息就是同场景需求直接相关的信息，其可以是因果信息，也可以是相关程度很高的信息。

如果定义采用0到1作为相关系数取值范围的化，0.6以上的相关系数就应该定义为强相关信息。例如在其他条件相同的前提下，35岁左右人的平均工资高于平均年龄为30岁的人，计算机专业毕业的学生平均工资高于哲学专业学生，从事金融行业工作的平均工资高于从事纺织行业的平均工资，上海的平均工资超过海南省平均工资。从这些信息可以看出来人的年龄、学历、职业、地点对收入的影响较大，同收入高低是强相关关系。简单的将，对信用信息影响较大的信息就是强相关信息，反之则是弱相关信息。

用户其他的信息，例如用户的身高、体重、姓名、星座等信息，很难从概率上分析出其对消费能力的影响，这些弱相关信息，这些信息就不应该放到用户画像中进行分析，对用户的信用消费能力影响很小，不具有较大的商业价值。

用户画像和用户分析时，需要考虑强相关信息，不要考虑弱相关信息，这是用户画像的一个原则。

## 5.3将定量的信息归类为定性的信息

用户画像的目的是为产品筛选出目标客户，定量的信息不利于对客户进行筛选，需要将定量信息转化为定性信息，通过信息类别来筛选人群。

例如可以将年龄段对客户进行划分，18岁-25岁定义为年轻人，25岁-35岁定义为中青年，36-45定义为中年人等。可以参考个人收入信息，将人群定义为高收入人群，中等收入人群，低收入人群。参考资产信息也可以将客户定义为高、中、低级别。定性信息的类别和方式方法，金融可以从自身业务出发，没有固定的模式。

将金融企业各类定量信息，集中在一起，对定性信息进行分类，并进行定性化，有利与对用户进行筛选，快速定位目标客户，是用户画像的另外一个原则。

## 5.4、用户画像的方法介绍，不要太复杂

金融企业需要结合业务需求进行用户画像，从实用角度出发，我们可以将用户画像信息分成五类信息**。分别是人口属性，信用属性，消费特征，兴趣爱好，社交属性。**它们基本覆盖了业务需求所需要的强相关信息，结合外部场景数据将会产生巨大的商业价值。我们先了解下用户画像的五大类信息的作用，以及涉及的强相关信息。特别复杂的用户画像纬度例如八个纬度，十个纬度信息都不利于商业应用，不建议金融企业进行采用，其他具有价值的信息，基本上都可以归纳到这五个纬度。金融企业达到其商业需求，从这五个纬度信息进行应用就可以了，不需要过于复杂用户画像这个工作，同时商业意义也不太大。

### 5.4.1、人口属性

用于描述一个人基本特征的信息，主要作用是帮助金融企业知道客户是谁，如何触达用户。姓名，性别，年龄，电话号码，邮箱，家庭住址都属于人口属性信息。

### 5.4.2、信用属性

用于描述用户收入潜力和收入情况，支付能力。帮助企业了解客户资产情况和信用情况，有利于定位目标客户。客户职业、收入、资产、负债、学历、信用评分等都属于信用信息。

### 5.4.3、消费特征

用于描述客户主要消费习惯和消费偏好，用于寻找高频和高价值客户。帮助企业依据客户消费特点推荐相关金融产品和服务，转化率将非常高。为了便于筛选客户，可以参考客户的消费记录将客户直接定性为某些消费特征人群，例如差旅人群，境外游人群，旅游人群，餐饮用户，汽车用户，母婴用户，理财人群等。

### 5.4.4、兴趣爱好

用于描述客户具有哪方面的兴趣爱好，在这些兴趣方面可能消费偏好比较高。帮助企业了解客户兴趣和消费倾向，定向进行活动营销。兴趣爱好的信息可能会和消费特征中部分信息有重复，区别在于数据来源不同。消费特征来源于已有的消费记录，但是购买的物品和服务不一定是自己享用，但是兴趣爱好代表本人的真实兴趣。例如户外运动爱好者，旅游爱好者，电影爱好者，科技发烧友，健身爱好者，奢侈品爱好者等。兴趣爱好的信息可能来源于社交信息和客户位置信息。

## 5.5、金融企业用户画像的基本步骤

参考金融企业的数据类型和业务需求，可以将金融企业用户画像工作进行细化。基本上从数据集中到数据处理，从强相关数据到定性分类数据，从引入外部数据到依据业务场景进行筛选目标用户。

### 5.5.1、画像相关数据的整理和集中

金融企业内部的信息分布在不同的系统中，一般情况下，人口属性信息主要集中在客户关系管理系统，信用信息主要集中在交易系统和产品系统之中，也集中在客户关系管理系统中，消费特征主要集中在渠道和产品系统中。

兴趣爱好和社交信息需要从外部引入，例如客户的行为轨迹可以代表其兴趣爱好和品牌爱好，移动设备到位置信息可以提供较为准确的兴趣爱好信息。社交信息，可以借助于金融行业自身的文本挖掘能力进行采集和分析，也是可以借助于厂商的技术能力在社交网站上直接获得。社交信息往往是实时信息，商业价值较高，转化率也较高，是大数据预测方面的主要信息来源。例如用户在社交网站上提出罗马哪里好玩的问题，就代表用户未来可能有出国旅游的需求；如果客户在对比两款汽车的优良，客户购买汽车的可能性就较大。金融企业可以及时介入，为客户提供金融服务。

客户画像数据主要分为五类，人口属性、信用信息、消费特征、兴趣爱好、社交信息。这些数据都分布在不同的信息系统，金融企业都上线了数据仓库（DW），所有画像相关的强相关信息都可以从数据仓库里面整理和集中，并且依据画像商业需求，利用跑批作业，加工数据，生成用户画像的原始数据。

**数据仓库成为用户画像数据的主要处理工具**，依据业务场景和画像需求将原始数据进行分类、筛选、归纳、加工等，生成用户画像需要的原始数据。

用户画像的纬度信息不是越多越好，只需要找到这五大类画像信息强相关信息，同业务场景强相关信息，同产品和目标客户强相关信息即可。根本不存在360度的用户画像信息，也不存在丰富的信息可以完全了解客户，另外数据的实效性也要重点考虑。

### 5.5.2、找到同业务场景强相关数据

依据用户画像的原则，所有画像信息应该是五大分类的强相关信息。强相关信息是指同业务场景强相关信息，可以帮助金融行业定位目标客户，了解客户潜在需求，开发需求产品。

只有强相关信息才能帮助金融企业有效结合业务需求，创造商业价值。例如姓名、手机号、家庭地址就是能够触达客户的强人口属性信息，收入、学历、职业、资产就是客户信用信息的强相关信息。差旅人群、境外游人群、汽车用户、旅游人群、母婴人群就是消费特征的强相关信息。摄影爱好者、游戏爱好者、健身爱好者、电影人群、户外爱好者就是客户兴趣爱好的强相关信息。社交媒体上发表的旅游需求，旅游攻略，理财咨询，汽车需求，房产需求等信息代表了用户的内心需求，是社交信息场景应用的强相关信息。

金融企业内部信息较多，在用户画像阶段不需要对所有信息都采用，只需要采用同业务场景和目标客户强相关的信息即可，这样有助于提高产品转化率，降低投资回报率（ROI），有利于简单找到业务应用场景，在数据变现过程中也容易实现。

千万不要将用户画像工作搞的过于复杂，同业务场景关系不大，这样就让很多金融企业特别是领导失去用户画像的兴趣，看不到用户画像的商业，不愿意在大数据领域投资。为企业带来商业价值才是用户画像工作的主要动力和主要目的。

### 5.4.1、对数据进行分类和标签化（定量to定性）

金融企业集中了所有信息之后，依据业务需求，对信息进行加工整理，需要对定量的信息进行定性，方便信息分类和筛选。这部分工作建议在数据仓库进行，不建议在大数据管理平台（DMP）里进行加工。

定性信息进行定量分类是用户画像的一个重要工作环节，具有较高的业务场景要求，考验用户画像商业需求的转化。其主要目的是帮助企业将复杂数据简单化，将交易数据定性进行归类，并且融入商业分析的要求，对数据进行商业加工。例如可以将客户按照年龄区间分为学生，青年，中青年，中年，中老年，老年等人生阶段。源于各人生阶段的金融服务需求不同，在寻找目标客户时，可以通过人生阶段进行目标客户定位。企业可以利用客户的收入、学历、资产等情况将客户分为低、中、高端客户，并依据其金融服务需求，提供不同的金融服务。可以参考其金融消费记录和资产信息，以及交易产品，购买的产品，将客户消费特征进行定性描述，区分出电商客户，理财客户，保险客户，稳健投资客户，激进投资客户，餐饮客户，旅游客户，高端客户，公务员客户等。利用外部的数据可以将定性客户的兴趣爱好，例如户外爱好者，奢侈品爱好者，科技产品发烧友，摄影爱好者，高端汽车需求者等信息。

将定量信息归纳为定性信息，并依据业务需求进行标签化，有助于金融企业找到目标客户，并且了解客户的潜在需求，为金融行业的产品找到目标客户，进行精准营销，降低营销成本，提高产品转化率。另外金融企业还可以依据客户的消费特征、兴趣爱好、社交信息及时为客户推荐产品，设计产品，优化产品流程。提高产品销售的活跃率，帮助金融企业更好地为客户设计产品。

### 5.4.1、依据业务需求引入外部数据

利用数据进行画像目的主要是为业务场景提供数据支持，包括寻找到产品的目标客户和触达客户。金融企业自身的数据不足以了解客户的消费特征、兴趣爱好、社交信息。

金融企业可以引入外部信息来丰富客户画像信息，例如引入银联和电商的信息来丰富消费特征信息，引入移动大数据的位置信息来丰富客户的兴趣爱好信息，引入外部厂商的数据来丰富社交信息等。

外部信息的纬度较多，内容也很丰富，但是如何引入外部信息是一项具有挑战的工作。外部信息在引入时需要考虑几个问题，分别是外部数据的覆盖率，如何和内部数据打通，和内部信息的匹配率，以及信息的相关程度，还有数据的鲜活度，这些都是引入外部信息的主要考虑纬度。外部数据鱼龙混杂，数据的合规性也是金融企业在引入外部数据时的一个重要考虑，敏感的信息例如手机号、家庭住址、身份证号在引入或匹配时都应该注意隐私问题，基本的原则是不进行数据交换，可以进行数据匹配和验证。

外部数据不会集中在某一家，需要金融企业花费大量时间进行寻找。外部数据和内部数据的打通是个很复杂的问题，手机号／设备号／身份证号的MD5数值匹配是一种好的方法，不涉及隐私数据的交换，可以进行唯一匹配。依据行业内部的经验，没有一家企业外部数据可以满足企业要求，外部数据的引入需要多方面数据。一般情况下，数据覆盖率达到70%以上，就是一个非常高的覆盖率。覆盖率达到20%以上就可以进行商业应用了。

金融行业外部数据源较好合作方有银联、芝麻信用、运营商、中航信、腾云天下、腾讯、微博、前海征信，各大电商平台等。市场上数据提供商已经很多，并且数据质量都不错，需要金融行业一家一家去挖掘，或者委托一个厂商代理引入也可以。独立第三方帮助金融行业引入外部数据可以降低数据交易成本，同时也可以降低数据合规风险，是一个不错的尝试。另外各大城市和区域的大数据交易平台，也是一个较好的外部数据引入方式。

### 5.4.1、按照业务需求进行筛选客户（DMP的作用）

用户画像主要目的是让金融企业挖掘已有的数据价值，利用数据画像技术寻找到目标客户和客户的潜在需求，进行产品推销和设计改良产品。

用户画像从业务场景出发，实现数据商业变现重要方式。用户画像是数据思维运营过程中的一个重要闭环，帮助金融企业利用数据进行精细化运营和市场营销，以及产品设计。用户画像就是一切以数据商业化运营为中心，以商业场景为主，帮助金融企业深度分析客户，找到目标客户。

DMP（大数据管理平台）在整个用户画像过程中起到了一个数据变现的作用。从技术角度来讲，DMP将画像数据进行标签化，利用机器学习算法来找到相似人群，同业务场景深度结合，筛选出具有价值的数据和客户，定位目标客户，触达客户，对营销效果进行记录和反馈。大数据管理平台DMP过去主要应用在广告行业，在金融行业应用不多，未来会成为数据商业应用的主要平台。

DMP可以帮助信用卡公司筛选出未来一个月可能进行分期付款的客户，电子产品重度购买客户，筛选出金融理财客户，筛选出高端客户（在本行资产很少，但是在他行资产很多），筛选出保障险种，寿险，教育险，车险等客户，筛选出稳健投资人，激进投资人，财富管理等方面等客户，并且可以触达这些客户，提高产品转化率，利用数据进行价值变现。DMP还可以了解客户的消费习惯、兴趣爱好、以及近期需求，为客户定制金融产品和服务，进行跨界营销。利用客户的消费偏好，提高产品转化率，提高用户黏度。

DMP还作为引入外部数据的平台，将外部具有价值的数据引入到金融企业内部，补充用户画像数据，创建不同业务应用场景和商业需求，特别是移动大数据、电商数据、社交数据的应用，可以帮助金融企业来进行数据价值变现，让用户画像离商业应用更加近一些，体现用户画像的商业价值。

用户画像的关键不是360度分析客户，而是为企业带来商业价值，离开了商业价值谈用户画像就是耍流氓。金融企业用户画像项目出发点一定要从业务需求出发，从强相关数据出发，从业务场景应用出发。用户画像的本质就是深度分析客户，掌握具有价值数据，找到目标客户，按照客户需求来定制产品，利用数据实现价值变现。

# 6、银行用户画像实践介绍

银行具有丰富的交易数据、个人属性数据、消费数据、信用数据和客户数据，用户画像的需求较大。但是缺少社交信息和兴趣爱好信息。

到银行网点来办业务的人年纪偏大，未来消费者主要在网上进行业务办理。银行接触不到客户，无法了解客户需求，缺少触达客户的手段。分析客户、了解客户、找到目标客户、为客户设计其需要的产品，成了银行进行用户画像的主要目的。银行的主要业务需求集中在消费金融、财富管理、融资服务，用户画像要从这几个角度出发，寻找目标客户。

银行的客户数据很丰富，数据类型和总量较多，系统也很多。可以严格遵循用户画像的五大步骤。先利用数据仓库进行数据集中，筛选出强相关信息，对定量信息定性化，生成DMP需要的数据。利用DMP进行基础标签和应用定制，结合业务场景需求，进行目标客户筛选或对用户进行深度分析。同时利用DMP引入外部数据，完善数据场景设计，提高目标客户精准度。找到触达客户的方式，对客户进行营销，并对营销效果进行反馈，衡量数据产品的商业价值。利用反馈数据来修正营销活动和提高ROI。形成市场营销的闭环，实现数据商业价值变现的闭环。另外DMP还可以深度分析客户，依据客户的消费特征、兴趣爱好、社交需求、信用信息来开发设计产品，为金融企业的产品开发提供数据支撑，并为产品销售方式提供场景数据。

简单介绍一些DMP可以做到的数据场景变现。

A 寻找分期客户

**利用发卡机构数据＋自身数据＋信用卡数据，发现信用卡消费超过其月收入的用户，推荐其进行消费分期。**

B 寻找高端资产客户

**利用发卡机构数据＋移动位置数据（别墅／高档小区）＋物业费代扣数据＋银行自身数据＋汽车型号数据，发现在银行资产较少，在其他行资产较多的用户，为其提供高端资产管理服务**。ookilllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllll

C 寻找理财客户

利用自身数据（交易＋工资）＋移动端理财客户端／电商活跃数据。发现客户将工资／资产转到外部，但是电商消费不活跃客户，其互联网理财可能性较大，可以为其提供理财服务，将资金留在本行。

D 寻找境外游客户

利用自身卡消费数据＋移动设备位置信息＋社交好境外强相关数据（攻略，航线，景点，费用），寻找境外游客户为其提供金融服务。

E 寻找贷款客户

利用自身数据（人口属性＋信用信息）＋移动设备位置信息＋社交购房／消费强相关信息，寻找即将购车／购房的目标客户，为其提供金融服务（抵押贷款／消费贷款）。

# 7、保险行业用户画像实践

保险行业的产品是一个长周期产品，保险客户再次购买保险产品的转化率很高，经营好老客户是保险公司一项重要任务。保险公司内部的交易系统不多，交易方式不是很复杂，数据主要集中在产品系统和交易系统之中，客户关系管理系统中也包含丰富了信息，但是数据集中在很多保险公司还没有完成，数据仓库建设可能需要在用户画像建设前完成。

保险公司主要数据有人口属性信息，信用信息，产品销售信息，客户家人信息。缺少兴趣爱好、消费特征、社交信息等信息。保险产品主要有寿险，车险，保障，财产险，意外险，养老险，旅游险。

保险行业DMP用户画像的业务场景都是围绕保险产品进行的，简单的应用场景可以是。

A

依据自身数据（个人属性）＋外部养车App活跃情况，为保险公司找到车险客户。

B

依据自身数据（个人属性）＋移动设备位置信息，为保险企业找到商旅人群，推销意外险和保障险。

C

依据自身数据（家人数据）＋人生阶段信息，为用户推荐理财保险，寿险，保障保险，养老险，教育险。

D

依据自身数据＋外部数据，为高端人士提供财产险和寿险。

6.3