用户画像 - 描述统计学

应用案例

*<2019年1月27日更新 V1.0>*

修订记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **修改人** | **修改原因** |
| 2019年 1月27日 | V1.0 <初稿> | 李旭龙, Kylie | * 创建“用户画像”文档v1.0版本。 |

# 互动：

那么，大家在自己的工作场景中，有没有用到用户画像呢？这些用户画像是如何帮到我们工作的呢？大家可以谈谈自己的经历~

# 一. 用户画像基本概念

**1 什么是用户画像**

用户画像是根据用户属性、生活习惯和消费行为等信息而抽象出的一个标签化的用户模型。

**2 用户属性**

用户属性又可分为静态属性和动态属性。

静态属性：一般指一些较稳定不会频繁变化的属性，静态属性是用户画像建立的基础，最基本的用户信息记录。如性别、年龄、学历、角色、收入、地域、婚姻状态。

动态属性：动态属性常常指具有可变性的属性，如用户的兴趣爱好、在互联网上的活动行为特征，比如网页浏览行为、使用行为、交易行为等。

**获取手段：**用户画像的数据获取手段有很多种，我们一般会采用调查问卷的方式获取用户的基本数据，也可以通过产品长时间的运营数据以及活动数据来收集。

**用户属性举例：**

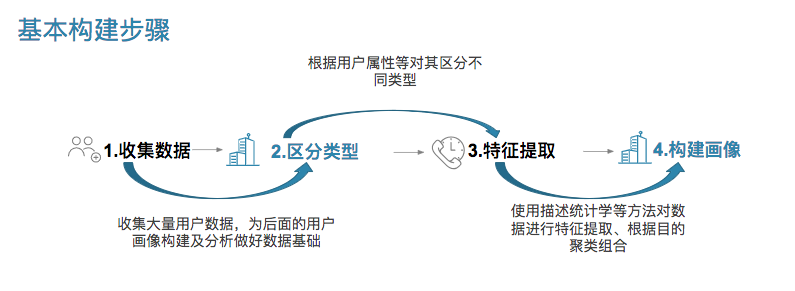
1. 你在网站填写的个人资料，比如性别，年龄等个人属性。

2. 如果经常购买玩具，电商网站可能根据购买玩具的情况给你打上有孩子的标签，甚至推算出你孩子的大概年龄，贴上有5-10岁孩子的更为具体的标签，更有可能推算出你30-35岁的年龄

**3 用户画像的构建步骤**

首先我们可以收集到大量关于自己用户的数据，这些数据可能从调查问卷或者运营历史数据中获得；随后根据他们的用户属性、生活习惯和消费行为等，将他们区分为不同的类型。接着我们会在每种类型中去抽取典型用户特征，也就是特征提取，赋予群类、年龄、一些人口统计要素、场景等描述，确定用户的属性，最后根据这些特征组合形成自己的用户画像，然后对这些画像进行更为深入的分析和利用等。

当然不同的岗位关注的用户画像属性，以及最终构建出的用户画像也是有很大不同的。我们会在后面举例子。



# 二. 用户画像的作用

我们将首先明解工作中为何需要使用用户画像，以及用户画像会为我们带来的好处，要明白一个优质的用户画像会起到意想不到的结果。

**原因一：准确地了解现有用户**

使用用户画像能够提炼和挖掘核心用户的特征。可以使得产品或者运营的人员可以抛开个人喜好，将焦点关注在核心用户的某些共性动机以及行为上进行产品以及运营方案的设计。

## **原因二：获取类似画像的新用户**

使用用户画像还能帮我们做什么呢，对于营销部门，我们需要通过获取用户画像，来获取和找到类似画像的新用户，这样可以实现精准营销，提高在市场投放上的投资回报率（ROI）。

除了这两点以外，还可以对销售运营的数据进行深度挖掘，来对我们的用户进行智能商品的推荐；

当然我们还有很多地方促使我们使用用户画像，完善产品运营，进行私人化定制，甚至进行竞品分析等等，这些领域中都有用户画像广阔的使用空间。

用户画像虽然好用，但是我们还是有一些地方是**需要注意**的，不然可能会适得其反：

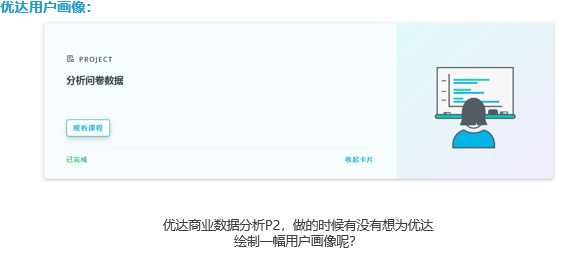
1. 首先要注意需要控制数量，也就是说我们**不能同时为三个以上**的用户画像涉及产品，如果一开始就目标过多，容易顾此失彼；
2. 随后也需要考虑清楚不同用户画像的**优先级**，比如主推3C产品等，优先级就在男性画像，如果主推美妆、购物，那目标应该指向女性画像；
3. 还有就是不同的职位需要根据自己的情况选取分类，确定画像，**不能共用别人的画像**，很可能对自己的情况完全不相符，后面我们会就这个详细展开一个例子。

# 三. 实例分析

**1 简介**

优达的学员到底是什么样的人群呢，针对优达的目标用户我们可以如何分析呢，下面实例采用虚拟数据，帮助大家切实了解一下不同职位绘制用户画像的关注点，以及最终的可能的数据结果。

大家也可以采用项目2：分析问卷中的真实优达学城数据进行分析。



这个项目大家都做了，不知道大家在做项目的过程中，看到优达的调查问卷数据有什么拓展想法呢，在完成项目的同时有没有想用这些数据绘制一幅优达的用户画像呢，那绘制出来的画像可能会呈现什么样子呢，接下来我们来细看。

**2 数据概览**

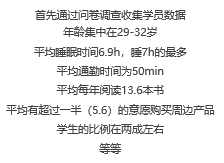
首先不管什么职位，拿到调查问卷数据后，我们都可以先大致了解一下数据里面有什么维度，根据这些维度可以大致感知一下可以探究或者分析的方向，需要筛出哪些数据帮助我们深入探究、

下面我们就简单看两个例子，对于不同的维度我们用什么描述性统计、可视化来探究比较合适（这在之前的课程应该学过）：

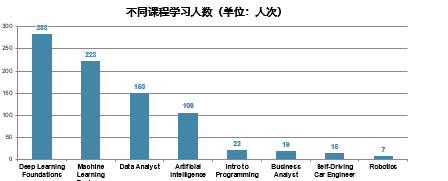
学员年龄：对于年龄这个连续数据变量，我们可以使用均值来探究，但是因为数据本身有很多异常的数据，用均值展现出来并不客观，于是可以考虑将不同年龄分段聚类，以柱形图、饼图形式展现，取众数，也就是占比最多的数据作为人物画像的一个特征；

对于学员所在城市同样采用方法，可以来出大多来自北京上海这两个城市，这里我们采用了柱形图。

了解了如何结合数据维度和描述性统计之后，我们可以对自己的数据进行快速查看，大致了解用户的特征，为后面分行业的挖掘寻找方向，例：



**3 市场人员应该如何分析**  
作为一名市场的同学，我们的目标就是找到目标人群，定向投放，卖出更多的课程，提高ROI（投资回报率），根据自己的目标，可以先定位不同课程的人员画像，最后再根据此画像定位目标人群，实现转化。



于是输出了不同课程的学习人数柱状图，得出了学习人数最多的三个课程：深度学习、机器学习和数据分析，将目标用户确定在这三个课程之内，去分别寻找符合的用户画像，就得到了下面这三组特征：



我们能看到，对于不同的课程，人群特征多少有些差异，这正是我们分类输出用户画像的用意，可以更精准的定位到不同人群。我们可以根据这些特征，必要时加上定性访谈作为辅助，生成更加具象化的用户画像。

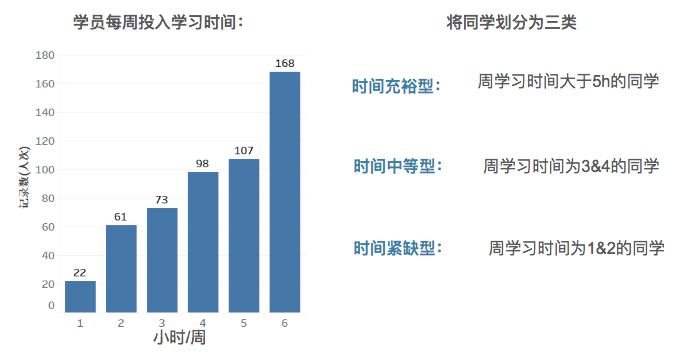
首先深度学习和机器学习较为类似，多属于互联网科技行业，职位也多为软件工程师，基本都是理工科男生居多，不过深度学习的同学年级偏大，在31-35岁之间，学历多为学士和硕士学位，学习深度学习是兴趣使然，基本没有太多转行需求，平均每周学习4.6小时，是一个成熟的理工中产形象，工作有成，但仍谋求成长；而机器学习的同学年级稍微年轻：26-300岁之间，学历大多是硕士学位，对机器学习不仅是兴趣使然还有一部分转行的诉求，每周学习4.5小时，所以基本呈现一个年轻有为，成长潜力不错还想要在不同行业也有更大作为的年轻形象

再看数据分析的画像，有比较大的不同了，年龄横跨20-35岁，行业不仅有互联网，也多了很多金融行业的朋友，职业多为商业智能或者商业工程师，学历多为学士和硕士学位，平均每周学习4.4小时，用于转行的诉求较大，总体呈现一个求变的女生形象。

得到了上述三个画像之后，市场人员可以根据这些画像，在人群之中寻找更符合这行画像的用户，精准营销，更有可能提升转化率。

**4 运营人员的用户画像**

同样的，作为一名运营同学，关注的角度自然和市场有所不同，运营一般针对不同的人群可以采用定向运营的策略，以提升课程满意度和课程毕业率，基于这个痛点，我针对学员每周投入的学习时间，绘制出投入时间柱状图，将投入时间不同的学员划分成三大类：时间充裕型（投入时间 >5-6小时）、时间中等型（3-4小时）和时间紧缺型（1-2小时）：



接下来的运营同学的步骤和之前一样，我们需要根据自己区分的学员类型，对其特征进行分别提取，并具象化成不同的人物画像：



首先能看出来，时间充裕型的同学，大部分是谋求转行，通勤时间再40分钟左右，呈中等态势，每年会看非常多的书，达到了18本之多，学历为本科及以上，每周投入大量时间运用知识，完成一个项目需要投入48个小时之久，综合来看时间充裕型的同学是跨行业学习且自驱能力很强，乐意花费大量时间再项目学习和运用的学霸形象；

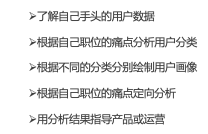
再来看时间中等型同学，学习多是凭借各人兴趣，谋求自己的成长、通勤时间最长，多达57分钟，每年看书13本，多是硕士学历，美中运用14小时学习知识，完成一个项目需要41 小时，可以看出，按时完成的同学，有着自己的节奏，凭借自己的兴趣支撑，安排好各方面的工作学习时间的自主型同学；

而时间紧缺型同学，虽然也是凭借兴趣学习，但是投入时间比较少，且运用知识的时间也不多，只有欧4小时每周，每天通勤时间最少，只有30是分钟，看来省下来的通勤时间也没有投入学习中去，学历在本科，每年看7本书，完成一个项目的时间最少，12小时，由于最后这个数据有点反常是，于是深入探究了一下数据，发现有一部分这类同学，完成项目的时间集中在项目完结前，说明很大程度上是项目马上到期，赶鸭子上架，有追进度的嫌疑，看来虽然项目时间是一个方面，但是质量却是另一方面了。

总结了上面三个画像，相信运营同学的策略和资源倾斜不言而喻，定向运营，提高满意度和毕业率可以说是目标清晰，那结果一定比没有策略来的好。

**5 小结**

详细看了上面两个例子之后，相信大家对用户画像步骤也比较清晰了，总结下来就是以下5点：



我们这里学习用户画像的目的是要灵活运用描述性统计学，运用这些知识，结合我们手头的数据，用数据的思维方式指导工作和生活，这样会有意想不到的结果。

# 四. 拓展内容

在本次分享结束之前，还是要说明的一点就是，用户画像目前并没有形成统一的方法论，本分享以及拓展内容的部分大家取精去伪，自己多尝试多总结才会达到自己想要的目的，下面这两个网站同学可以补充学习，祝大家在未来的工作中摸索出自己的一套优质方法。