消费者数据模拟

为了精准的描述用户特征，从用户微观画像的建立→用户画像的标签建模→用户画像的数据架构，我们由微观到宏观，逐层分析

第一分类：人口属性、资产特征、营销特性、兴趣爱好、购物爱好、需求特征

第二分类…

第三分类…

这里仅对第一分类的数据进行模拟

用户数据包含人口属性和需求特征

人口属性，年龄性别，地域，工作类型，家庭背景，资产特征

指标与维度是数据分析中最常用到的术语，它们是非常基础的，但是又很重要，经常有朋友没有搞清楚它们之间的关系，只有掌握理解了，我们的数据分析工作开展就就容易多了。现在就来说说指标与维度的那些事。

1、指标

指标，用于衡量事物发展程度的单位或方法，它还有个IT上常用的名字，也就是度量。例如：人口数、GDP、收入、用户数、利润率、留存率、覆盖率等。很多公司都有自己的KPI指标体系，就是通过几个关键指标来衡量公司业务运营情况的好坏。

指标需要经过加和、平均等汇总计算方式得到，并且是需要在一定的前提条件进行汇总计算，如时间、地点、范围，也就是我们常说的统计口径与范围。

指标可以分为绝对数指标和相对数指标，绝对数指标反映的是规模大小的指标，如人口数、GDP、收入、用户数，而相对数指标主要用来反映质量好坏的指标，如利润率、留存率、覆盖率等。我们分析一个事物发展程度就可以从数量跟质量两个角度入手分析，以全面衡量事物发展程度。

刚才说过，指标用于衡量事物发展程度，那这个程度是好还是坏，这就需要通过不同维度来对比，才能知道是好还是坏。

2、维度

维度：是事物或现象的某种特征，如性别、地区、时间等都是维度。其中时间是一种常用、特殊的维度，通过时间前后的对比，就可以知道事物的发展是好了还是坏了，如用户数环比上月增长10%、同比去年同期增长20%，这就是时间上的对比，也称为纵比;

另一个比较就是横比，如不同国家人口数、GDP的比较，不同省份收入、用户数的比较、不同公司、不同部门之间的比较，这些都是同级单位之间的比较，简称横比;

维度可以分为定性维度跟定量维度，也就是根据数据类型来划分，数据类型为字符型(文本型)数据，就是定性维度，如地区、性别都是定性维度;数据类型 为数值型数据的，就为定量维度，如收入、年龄、消费等，一般我们对定量维度需要做数值分组处理，也就是数值型数据离散化，这样做的目的是为了使规律更加明 显，因为分组越细，规律就越不明显，最后细到成最原始的流水数据，那就无规律可循。

最后强调一点，只有通过事物发展的数量、质量两大方面，从横比、纵比角度进行全方位的比较，我们才能够全面的了解事物发展的好坏。

Google Analytics（分析）中的每个报告都由维度和指标组成。

“维度”是指数据的属性。举例来说，“城市”维度表示的是发起会话的城市，例如“巴黎”或“纽约”。“网页”维度表示的是用户浏览过的网页的网址。

“指标”是量化衡量标准。“会话数”指标是指总会话次数。“每次会话浏览页数”指标是指每次会话的平均网页浏览量。

大多数 Google Analytics（分析）报告中的表格会逐行显示维度值，逐列显示指标值。 例如，下表显示的是一个维度（“城市”）和两个指标（“会话数”和“每次会话浏览页数”）

参考链接：

你确定你真的懂用户画像？

<http://www.woshipm.com/it/250043.html>

数据分析必须想清楚的两个概念：指标和维度

<https://www.cnblogs.com/xitingxie/p/5718572.html>

维度和指标：数据分析入门必须了解的2个基础概念

<https://www.jianshu.com/p/4551bcf98a93>