黑客派

大数据对外服务专题分析标准化案例

作者：relyn

原文链接：https://hacpai.com/article/1472544995655

来源网站：黑客派

许可协议：署名-相同方式共享 4.0 国际 (CC BY-SA 4.0)

<p># 前言
 大数据对外服务是支撑中心大数据中心的一项业务，也是大数据变现的一项重要手。虽然我们掌握了许多数据资源，熟悉多种分析手段，但由于各类行业都有其独特性，每次大数据对服务需求都应该充分调研，了解行业客户真正的需求，必要的时候还要到实地去考察，与客户面对面通，以达到最好的分析效果。
 # 交付成果
 大数据对外服务一般提供给客户一份专题分析告 PPT，PPT 中仅提供脱敏的统计级数据，或者趋势分析数据。不可包含任何用户清单级的敏感数据
 # 分析流程
 大数据对外服务流程为 152 流程，是由县市公司政企部客户经理发起，经过属领导、法务部、信息安全等部门审批，通过后方可进行。客户经理应尽可能详细地写清楚客户需求并告知客户最终的可交付成果。商务部分也需要提供，例如是否通过其他项目来置换大数据对外服务用等。
 ## 需求调研
 为了满足客户需求，避免分析结果出现问题，第一次接到大数据对外务流程时，可与项目发起人一同前往客户所在地，当面沟通，进一步明确客户的具体需求。
 ## 可行性分析
 与客户沟通完毕并明确了客户需求之后，应对该需求进行可行性分析，考虑客户需是否为当前系统中可提取的数据。此外，考虑交付成果是否为脱敏的统计级数据。
 ## 数据提取br> 数据提取过程一般通过客户标签库、客户轨迹表、常用基站表等维度进行提取。
 ###1.区确认
 根据商铺的位置，在地图上进行坐标拾取，通常情况下，我们先选取商铺所在位置多边形四个点，分别提取这四个点的 GPS 坐标（注意：只有谷歌地球的坐标才是真实的 GPS 坐标，其他坐都是经过转换的“火星坐标”）
 <img src="https://static.hacpai.com/images/img-loading.sg" alt="6666.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/c14c426e07384feb807257662f4d4ce.jpg">
 通过获取这 4 个坐标的最大值和最小值，得出实际的坐标取值范围如下图色方框所示。
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="777.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/51fef9fdcaa341c38a17f679cc04b830.jpg"
 从数据库中取得落在该区域内的基站清单。
 ###2.基站确认
 获得这个矩形区域中的站，如下图所示，可见，所得基站均处于商铺范围内。
 <img src="https://static.hacpai.com/mages/img-loading.svg" alt="base.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/0c2c1e287fa4417893d58d09f6d935a.jpg">
 ###3.用户分布热力图
 根据基站所覆盖的用数绘制出用户分布密度热力图。
 <img src="https://static.hacpai.com/images/img-loading.sg" alt="rlt.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/9f78db73b42d450e8f453bb2dc0061f.jpg">
 ###4.用户提取
 根据基站清单，提取用户信息，并剔除在该区域 7 天及 7 天以上的常驻用户。如下图所示，根据每天的访客情况可见，该区域的日均访客数为 23301.9 人，大值为 29008 人，最小值为 14818 人。
 <img src="https://static.hacpai.com/images/img-oading.svg" alt="444.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/cfda1a1db1fb4d5ac1496029a180c08.jpg">
 根据用户清单，进一步挖掘用户的详细信息。
 ## 数据挖掘br> 对于已获取的用户清单，可直接进行用户标签的关联，对于潜在目标客户的分析，可让客户提供效的真实会员资料，通过该清单进行建模，以得到潜在目标客户模型。
 ###1.基本信息
 （1）用户年龄分布
 如下图所示，该区域的用户年龄集中分布在 20-35 之间其中年龄为 26 岁的用户数最多，达到 1.8 万，随着年龄的增长呈递减趋势。
 <img src="https:/static.hacpai.com/images/img-loading.svg" alt="5555.jpg" data-src="http://7xw87f.com2.z0.lb.qiniucdn.com/3139b2e391844df88cea6b4bf88eb70f.jpg">
 （2）用户性别比
 如下图所示，该区域的用户为男性居多，占比高达 63%。
 <img src="https:/static.hacpai.com/images/img-loading.svg" alt="666666.jpg" data-src="http://7xw87f.com2.z.glb.qiniucdn.com/46d83766708c403a922af7cf65a0fdbb.jpg">
 （3）用户归属
 如下图所示，在一个月的监控样本中，用户归属来自丰泽的最多，达到 13 万人，出第二名两倍以上，晋江归属的用户达 5.78 万，鲤城归属的用户达 5.35 万，惠安归属的用户达 4.2。
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="77777.jpg" data-sc="http://7xw87f.com2.z0.glb.qiniucdn.com/d37841d798864f12b1ef3a57ca98ffba.jpg">
 （4）用户职业特征
 如下图所示，该区域近一个月来访的用户分布情况如，其中商旅人士最多，占比达 62%，学生占比 21%，公务人员占比 7%。
 <img src="https://satic.hacpai.com/images/img-loading.svg" alt="zyzy.jpg" data-src="http://7xw87f.com2.z0.glbqiniucdn.com/41265f976b3741419febf9343948ac0a.jpg">
 （5）用户星座分布
 该区域用户的星座分布如下图所示，其中天秤座用户占比最高，达到 10.54%，其次天蝎座用户，占比 9.95%，处女座用户占比第三，达到 8.92%。
 <img src="https://static.hacai.com/images/img-loading.svg" alt="star.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.om/4ddbea0336784b42b6a3af39ee0a0de8.jpg">
 ###2.消费能力
 （1）月 ARPU 值
 该区域用户平均 ARPU 值分布情况如下图所示，可见，绝大多数用户的 APU 值集中在 100-200 之间。
 <img src="https://static.hacpai.com/images/img-loading.sv

" alt="ARPU.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/4fa213fd68d04dcc8312003b1a4577c.jpg">
 （2）用户余额分布
 该区域用户余额分布情况下图所示，可见余额小于 50 元的用户数最多。
 <img src="https://static.hacpai.com/images/mg-loading.svg" alt="余额.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/5dff91f7e2046488b861c924582246a.jpg">
 （3）月均流量使用
 该区域的户流量使用情况如下图所示，其中，月均流量处于 500MB-1GB 的用户数最多，达到 8.7 万户。
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="ll.jpg" data-src="http://7x87f.com2.z0.glb.qiniucdn.com/cb7a34acc5b34cf88ca9638e367d0906.jpg">
 （）终端品牌分布
 下图所示为该区域用户所使用终端品牌的占比情况，其中苹果终端比最高达到 42.92%，华为、欧珀、小米并列占比 11%，步步高、三星品牌占比分别为 8% 和 6%。<r> <img src="https://static.hacpai.com/images/img-loading.svg" alt="zdzb.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/e78510ac62974e3fa29f40ba511fed92.jpg">
 <strog>（5）拥有汽车的用户占比
 如下图所示，该区域中，拥有汽车的用户占比为 26%
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="iscar.jpg" data-src=http://7xw87f.com2.z0.glb.qiniucdn.com/83e7287f2ac5425da6d6ef5b099a2e4a.jpg">
 （6）汽车品牌占比情况
 该区域用户的汽车品牌占比情况如下图所示，其中田汽车占比最高，达到 29.76%，丰田汽车占比仅次于本田，达 27.60%，大众汽车占比 23.73%，标汽车占比 3.38%，同时也存在宾利、保时捷、宝马、奔驰、雷克萨斯等高端汽车品牌。
 <img sr="https://static.hacpai.com/images/img-loading.svg" alt="carband.jpg" data-src="http://7xw7f.com2.z0.glb.qiniucdn.com/c8526007fa0449c6bb0e3a150b20f64f.jpg">
 ###3.偏好特征
 （1）上网内容偏好
 该区域用户上网行为偏好如下图所示，其中搜索偏好最高，达到 2.86 万次，IT 类偏好达 1.92 万次，游戏和购物类偏好分别为 1.39 和 1.33 万次，交生活类偏好为 1.06 万次。
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="swxw.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/6b2b756c27a8490f96b88e054fb97ad.jpg">
 （2）APP 使用偏好
 该区域中，社交通信类 APP 使用的频次最高，其次为 IT 类 APP 和系统工具类 APP。
 <img src="https://static.hacpai.co/images/img-loading.svg" alt="appuse.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.cm/e45a6b86a6f3471e8d9d301b7858e765.jpg">
 ###4.通信特征
 （1）用通话次数分布
 下图为该区域用户的月通话次数，其中月通话在 100-200 次的用户最，达到 91925 户，通话 500 次以上的为 75168 户。
 <img src="https://static.hacpai.com/iages/img-loading.svg" alt="thcs1.jpg" data-src="http://7xw87f.com2.z0.glb.qiniucdn.com/9a098fe83df437186f8c272eef18a6a.jpg">
 （2）用户影响力
 根据户通话特征评估模型对该区域的用户进行分析，得出该区域用户有 45% 的用户具有较高的社会影响。
 <img src="https://static.hacpai.com/images/img-loading.svg" alt="yxl.jpg" data-src=http://7xw87f.com2.z0.glb.qiniucdn.com/9dba402025584cd3b4c47399156cf286.jpg">
 # 分析报告制作
 分析报告一般提供 PPT 或 WORD 形式，尽量避免使用专业术语，应使用可供客阅读的简单描述或举例。
 ## 数据审核
 数据审核一般由 B 角、、信息安全部门、法务部财务部依次进行，仅当审核完成之后才可将最终的分析报告提交给客户。
 ## 分析报告提交
 分析报告通过流程进行流转，最终由需求发起人将该报告直接发送给客户。</p>

<script async src="https://pagead2.googlesyndication.com/pagead/js/adsbygoogle.js"></script>
<!-- 黑客派PC帖子内嵌-展示 -->
<ins class="adsbygoogle" style="display:block" data-ad-client="ca-pub-5357405790190342" data-ad-slot="8316640078" data-ad-format="auto" data-full-width-responsive="true"></in>
<script>
    (adsbygoogle = window.adsbygoogle || []).push({});
</script>