**数据建设总结**

**高意向人群区分**

**需求：**业务反馈高价楼盘中存在很多非高消费人群的标签，使得报告结果与实际情况相差较大，需调整

**过程：**取部分高价楼盘的用户标签进行校验，发现很多低消费标签用户有低频次点击高价盘的行为，故做了算法优化，保留原始真实点击数据的前提下，增加一版【高意向】用户的报告，根据楼盘均价所处城市均值的水平，划分为高、中、低三挡楼盘，并过滤低点击频次用户，再结合用户居住小区价格，线下消费场景划分消费能力高、中、低三挡与楼盘对应，匹配核心高意向用户人群。

**结果：**高意向版报告与业务认知情况相似度提高到80%，个别标签的认知出入可能源于线上浏览与实际线下案场的人群不同，不排除这部分人是目标楼盘潜客，故保留了部分出入结果。

**出行方式标签优化**

**需求：**业务反馈高价楼盘中的用户驾车出行比例过低，骑行与地铁比例过高

**过程：**核查驾车出行标签判断逻辑，发现判断用户有车的逻辑主要取决于个别APP是否安装，结果判定过于片面，故做了判断算法优化，添加居住地与工作地的距离+用户年龄+消费能力+线下消费场景出行频次等标签综合判断，不在局限个别APP是否安装。

**结果：**校验多个楼盘报告中的用户驾车、地铁、骑行的出行方式比例符合业务认知