

2021 年全国大学生数据统计与分析竞赛

题目 B：用户消费行为价值分析

在互联网不断发展的今天，各领域公司都在拓展互联网获客渠道，为公司产品引入新鲜活跃的用户，提高用户购买产品的欲望，提升公司的品牌影响力。但如何判别高质量的用户和渠道，优化营销成本一直都是各公司的痛点；这需要对用户的行为数据进行分析，判别用户的价值，进而对用户制定专门的营销策略，实现小成本促销，达到提高用户转化率的目标。

现有用户信息表 (user_info.csv)，用户登录情况表 (login_day.csv)，用户访问统计表 (visit_info.csv)，用户下单表 (result.csv)，利用数据对用户行为进行数据统计与分析，并判断用户是否会下单购买。

任务 1：获取数据并进行预处理，提高数据质量；

任务 2：对用户的各城市分布情况、登录情况进行分析，并分别将结果进行多种形式的可视化展现；

任务 3：构建模型判断用户最终是否会下单购买或下单购买的概率，并将模型结果输出为 csv 文件（参照结果输出样例 sample_output.csv）。要求模型的效果达到 85%以上；

任务 4：通过用户消费行为价值分析，给企业提出合理的建议。