

2024 年全国大学生数据分析大赛

A 题：直播带货与电商产品的大数据分析

直播带货，是直播娱乐行业在直播的同时带货，由主播在直播间里推介，也称为好物推荐官。2023 年 11 月，中国消费者协会对外发布的《2023 年“双 11”消费维权舆情分析报告》显示，消费者维权问题主要集中在直播带货乱象等方面。对于直播带货平台和主播来说，通过顾客的交易行为、交易金额、交易频率、交易时间等信息，分析顾客的行为和电商产品的销售情况，可促进二次营销，精准营销，促活等。请根据附件中提供的 541910 条数据，进行数据的分析与建模。

一、数据说明：

字段	注释
EventID	订单的唯一标识
ProductID	商品的唯一标识
Product	购买商品的内容
Quantity	购买商品的数量
Date	购买商品的日期
Price	购买商品的价格，单位元
CustomerID	用户的唯一标识

Country	顾客所在国家
---------	--------

二、任务：

- **问题 1.** 对附件中直播带货商品进行统计分析并可视化。最畅销的商品前 10 名分别是哪些商品？购买直播带货商品总数最多的顾客前 10 名分别是哪些顾客？购买直播带货商品总数最多的国家前 10 名分别是哪些国家？购买直播带货商品的时间段主要集中在一天的什么范围？直播带货商品的价格主要集中在什么范围？（15 分）
- **问题 2.** 顾客画像是指根据顾客的属性、顾客偏好、顾客行为等信息而抽象出来的标签化用户模型。通俗说就是给顾客打标签，而标签是通过对顾客信息分析而来的高度精炼的特征标识。通过打标签可以利用一些高度概括、容易理解的特征来描述顾客，可以让人更容易理解顾客，请构建顾客属性、消费水平和消费偏好等标签，并绘制顾客画像。（20 分）
- **问题 3.** 直播带货平台拟推出不同价格的代金券，用于发放给顾客，请根据顾客的交易行为完成特征构造，并进行客户价值分析，推动此活动（25 分）
- **问题 4.** 请任选一位顾客，根据该顾客的交易行为，为该顾客推荐代金券，最后评估模型的性能。（25 分）
- **问题 5.** 请设计一份优惠券的投放策略，需要考虑优惠券的数量、优惠券的金额、投放时间段和投放商品种类等因素。（15 分）