

## 总结

### 流失率项目

- 交易流水
  - 交易使用红包金额
  - 同一店铺交易金额、交易次数、不成功交易次数
  - 某段时间店铺平均交易金额、交易次数、交易金额的标准差（交易波动性）、交易日期分布的熵（交易日期是否集中）
  - 不同支付方式使用金额占比（支付方式的多样性）
- 抛出店铺
  - 抛出次数（每个商户关联的店铺被抛出的次数）

### 信贷项目

- 贷款申请
  - 同一贷款人 id 贷款次数、逾期 30 天还款现象的贷款交易数、未逾期还贷频率、逾期率（逾期的次数与总贷款次数的比值）、是否出现最后还款期限内未还完的情况（通过理论结清日和实际结清日判断）
  - 实际还款和理论还款之间的天数差异
  - 历史逾期率：商户历史逾期的次数占总贷款次数的比率
  - 商户对应的有逾期 30 天还款现象的次数（要通过计算理论结清日与实际结清日进行重新计算）、商户对应的没有出现逾期 30 天还款现象的次数（要通过计算理论结清日与实际结清日进行重新计算）
  - 商户逾期率与平均贷款期数的比值
  - 商户历史贷款总金额、历史平均贷款期数、历史 30 天逾期次数
  - 每期应还款数（lend/amout）
  - 还款金额统计特征：平均值、总额、最大值、最小值等
- 交易流水
  - 商户平均每日消费者支付金额、平均每日红包金额
  - 商户储蓄卡、微信余额、支付宝余额、支付宝余额宝的平均每日消费者支付金额，商户信用卡、花呗、集分宝的平均每日消费者支付金额
  - 商户交易活动度（某段时间的交易次数、交易总金额、交易金额的标准差）
  - 信用卡支付金额与总支付金额的比例、花呗支付金额与总支付金额的比例
  - 支付方式统计（是否只用单一方式交易）、支付方式比例（不同支付方式的交易金额占比）
  - 交易频率、交易金额统计特征（平均值、总额、最大值、最小值）

## 示例

### ➤ 商户信息特征

- **连锁门店数量**: 指该商户旗下所有门店的总数量, 反映了商户规模
- **系统预警次数**: 指商户在  $t-29$  至  $t_0$  历史时段内被系统预警的总次数

### ➤ 交易流水特征

- **无交易占比**: 指  $t-29$  至  $t_0$  历史时段内商户不存在交易的天数与总天数的比值, 可以反映商户的经营状况. 给定历史时段商户存在交易的天数为  $T$ , 则无交易占比表示为  $(30-T)/30$
- **休息日**: 指系统预警时间点  $t_2$  是否为节假日或双休日, 反映了商户的经营模式
- **平均交易笔数**: 指  $t-29$  至  $t_0$  历史时段内商户每天交易笔数的平均值
- **平均交易金额**: 指  $t-29$  至  $t_0$  历史时段内商户每天交易金额的平均值
- **笔均交易金额**: 指  $t-29$  至  $t_0$  历史时段内商户每笔交易金额的平均值
- **交易失败占比**: 指  $t-29$  至  $t_0$  历史时段内商户交易失败笔数与总交易笔数的比值
- **三天交易笔数**: 指商户在  $t-2$  至  $t_0$  时段内的总交易笔数
- **三天交易金额**: 指商户在  $t-2$  至  $t_0$  时段内的总交易金额
- **平均交易笔数比较**: 指三天平均交易笔数与历史平均交易笔数的比值取对数
- **平均交易金额比较**: 指三天平均交易金额与历史平均交易金额的比值取对数
- **金额标准差**: 指  $t-29$  至  $t_0$  历史时段内商户每天交易金额的标准差

### ➤ 风控特征

- **信用卡支付笔数占比**: 指  $t-29$  至  $t_0$  历史时段内, 商户的所有收款记录中使用信用卡支付的笔数与总交易笔数的比值
- **信用卡支付金额占比**: 指  $t-29$  至  $t_0$  历史时段内, 商户的所有收款记录中使用信用卡支付的金额与总交易金额的比值
- **整 10 金额交易笔数**: 指  $t-29$  至  $t_0$  历史时段内满足交易金额是整十数的总交易笔数, 整 10 金额数量越多, 则越可能存在严重的刷单行为
- **高额信用天数**: 指  $t-29$  至  $t_0$  历史时段内满足信用卡支付大于 2000 元的总天数
- **高额信用笔数**: 指  $t-29$  至  $t_0$  历史时段内满足信用卡支付大于 2000 元的总笔数