数据集说明

1. 数据集构成
   1. cust\_dataset，客户基本属性文件， 数量1108827行7列，主键**cust\_no（客户号）**birth\_ym（出生年月）、loc\_cd（区域编码）、 gender（性别）、init\_dt（开户日期）、edu\_bg（学历编码）、marriage\_situ\_cd（婚姻状况编码）
   2. event\_dataset，2024年度业务事件文件，数量338939行 9列，关联字段cust\_no（客户号）、prod\_id（产品号）、event\_id（事件号）、event\_type（事件类型）、event\_level（事件级别）、event\_date（事件发生日期）、event\_term（事件期限，如有则有值）、event\_rate（事件利率分段，如有则有值）、event\_amt（事件额度分段，如有则有值）
   3. productLabels\_multiSpreadsheets，产品属性文件，多页文件，D代表储蓄类产品、C为信贷类产品、A为财富类产品、N为渠道类产品、P为委托支付类产品，每张子表的prod\_id（产品号），可与event\_dataset进行关联
2. 标记为成功事件的条件为：event\_type（事件类型）为A或B的，（字段event\_type：码值清单：A开立、B开通、D关闭）
3. 新产品为信贷产品，识别方式为：new\_flag列为1

注：客户为全量简单抽样，各项业务样本客户与全量客户有业务客户占比一致，所有数据集不含模拟数据。

附录：

|  |  |
| --- | --- |
| **事件（业务属性）** | **event\_id** |
| 信用卡开立 | E0001 |
| 信用卡激活 | E0002 |
| 信用卡关闭 | E0003 |
| 借款发生 | E0004 |
| 借款到期 | E0005 |
| 借款结清 | E0006 |
| 存款业务开立 | E0007 |
| 存款业务到期 | E0008 |
| 存款账户开通 | E0009 |
| 渠道类签约 | E0010 |
| 支付渠道签约 | E0011 |
| 代收付业务签约 | E0012 |
| 收单开立 | E0013 |
| 收单交易 | E0014 |
| 理财购入 | E0015 |