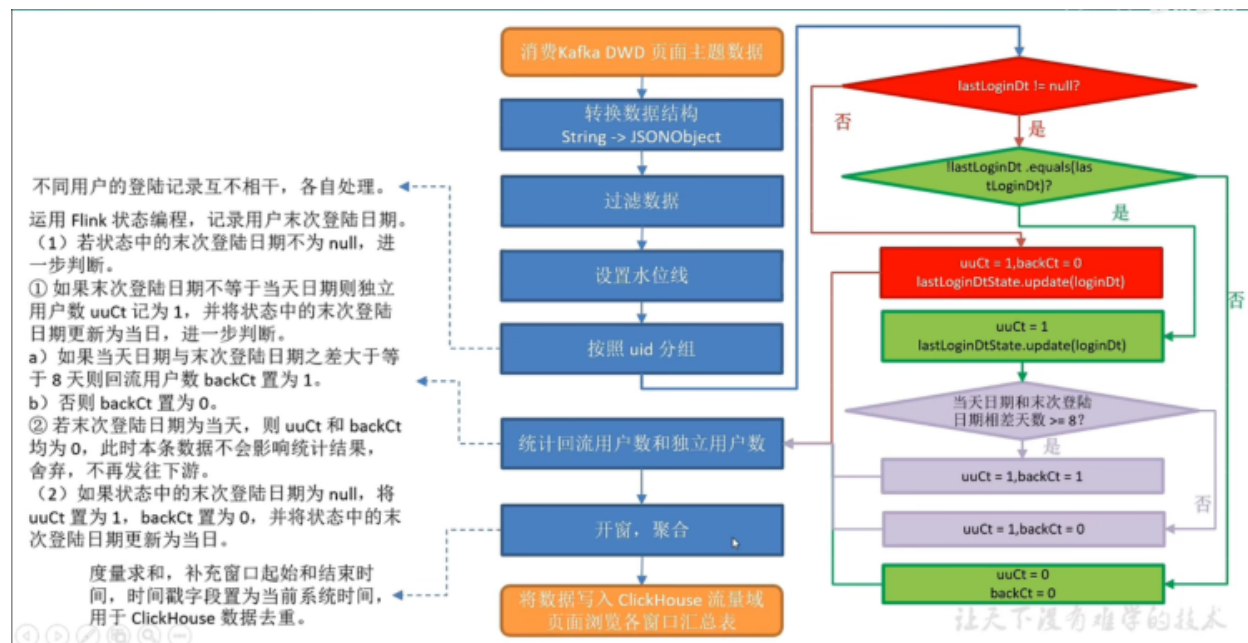
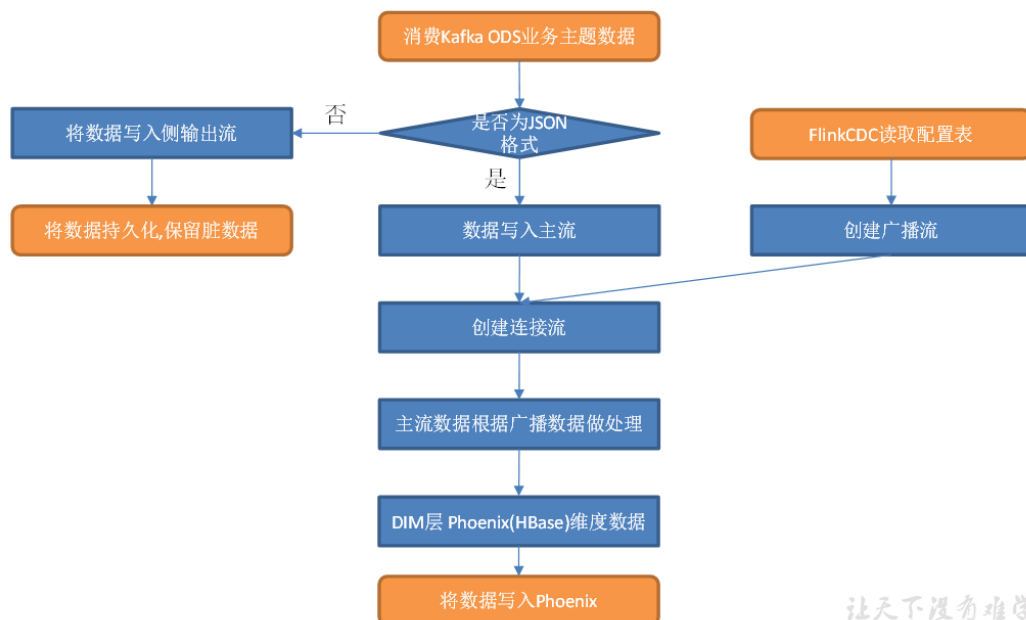


# 业务流程拆解



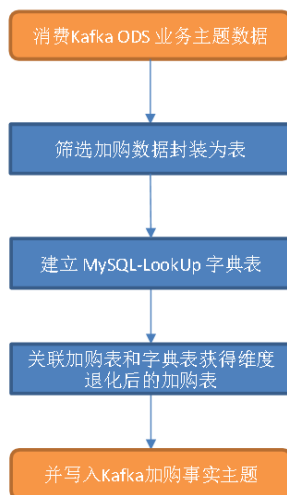
## 一、DIM层数据流程



## 二、ODS-DWD

## 交易域加购事实事物表

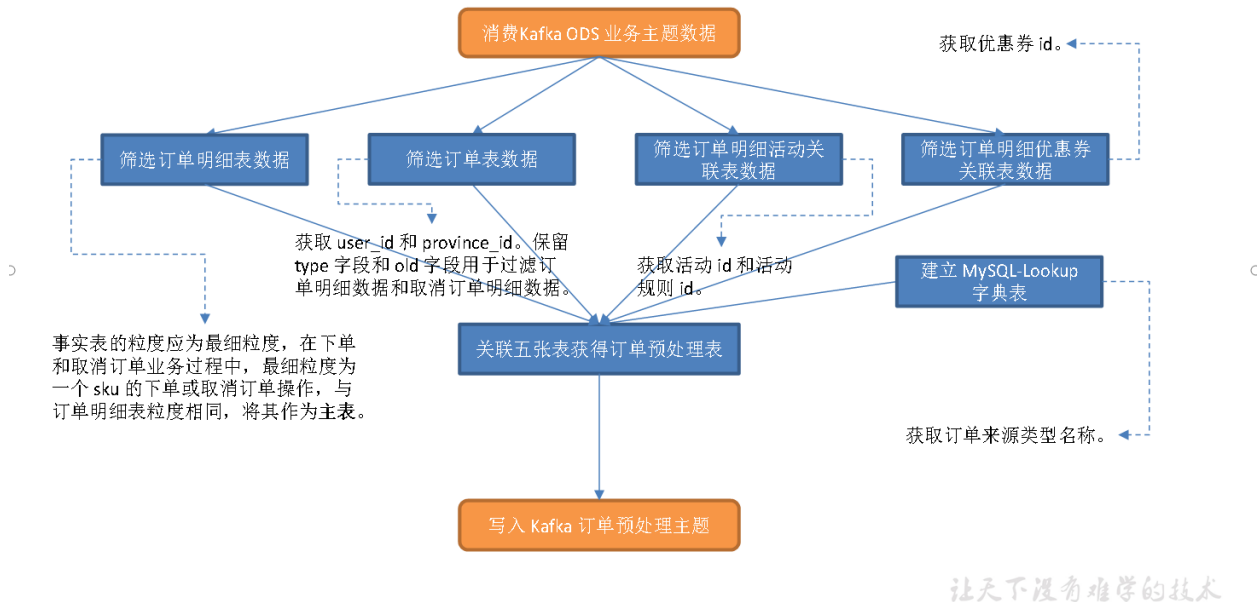
提取加购操作生成加购表，并将字典表中的相关维度退化到加购表中，写出到Kafka 对应主题。



让天下没有难学的技术

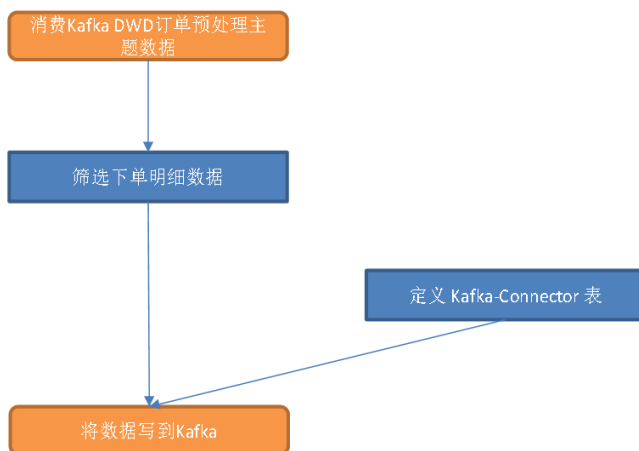
## 交易域订单预处理表

关联订单明细表、订单表、订单明细活动关联表、订单明细优惠券关联表四张事实业务表和字典表（维度业务表）形成订单预处理表，写入Kafka 对应主题。



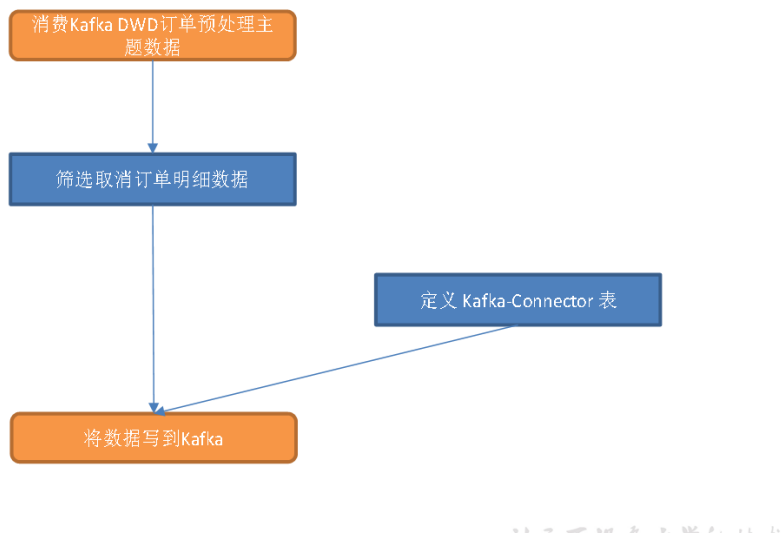
## 交易域下单事务事实表

从Kafka 读取订单预处理表数据，筛选下单明细数据，写入Kafka 对应主题



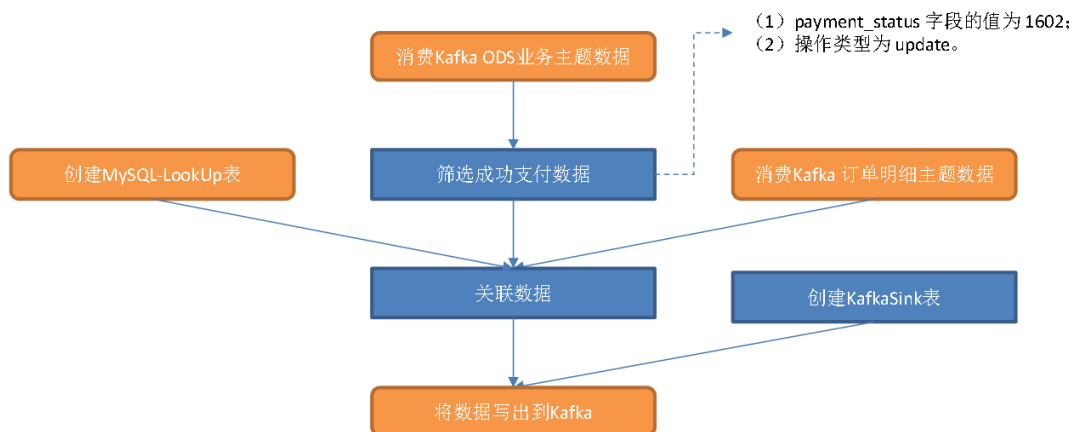
## 交易域取消订单事务事实表

从Kafka 读取订单预处理表数据，筛选取消订单明细数据，写入Kafka 对应主题。



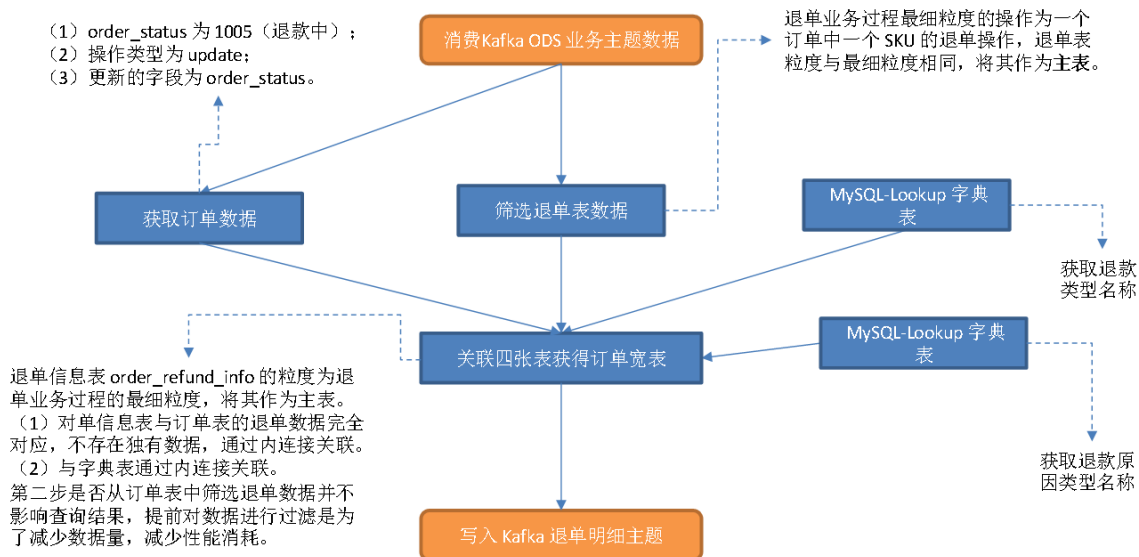
## 交易域支付成功事务事实表

从Kafka topic\_db主题筛选支付成功数据、从dwd\_trade\_order\_detail主题中读取订单事实数据、MySQL-LookUp字典表，关联三张表形成支付成功宽表，写入 Kafka 支付成功主题。（支付成功需求）



## 交易域退单事务事实表

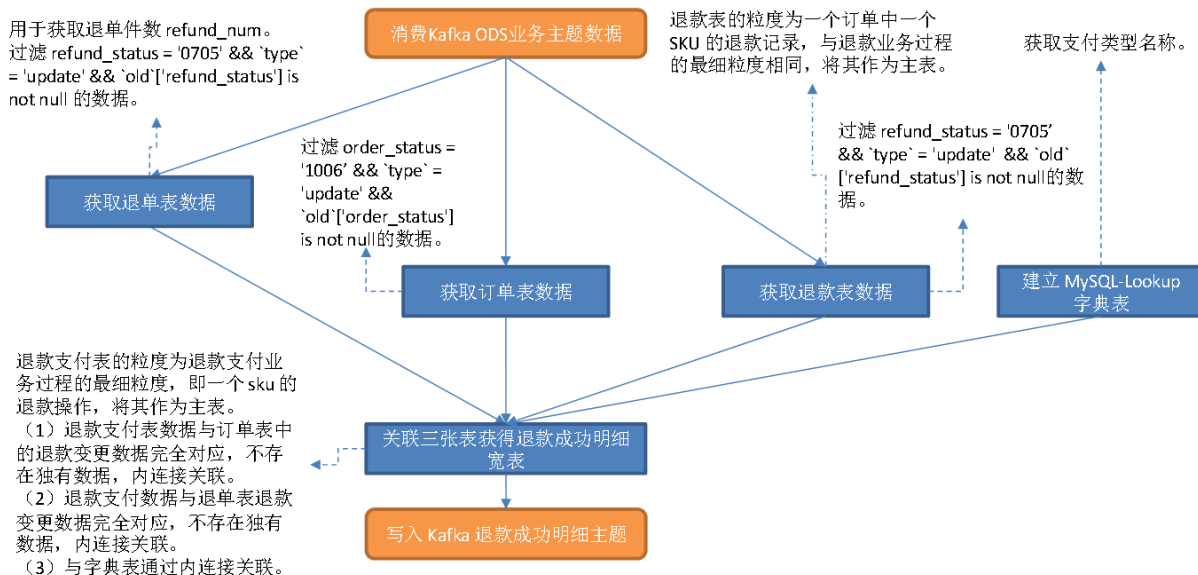
从Kafka 读取业务数据，筛选退单表数据，筛选满足条件的订单表数据，建立MySQL-Lookup 字典表，关联三张表获得退单明细宽表。（退单需求）



让天下没有难学的技术

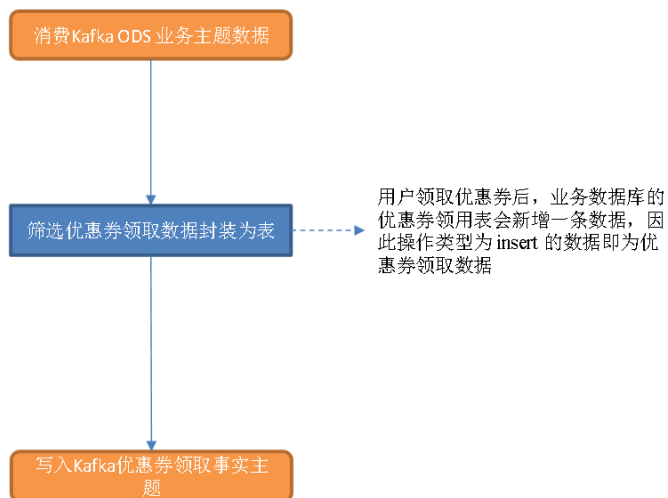
## 交易域退款成功事务事实表

从退款表中提取退款成功数据，并将字典表的dic\_name 维度退化到表中，从订单表中提取退款成功订单数据，从退单表中提取退款成功的明细数据



## 优惠券相关需求

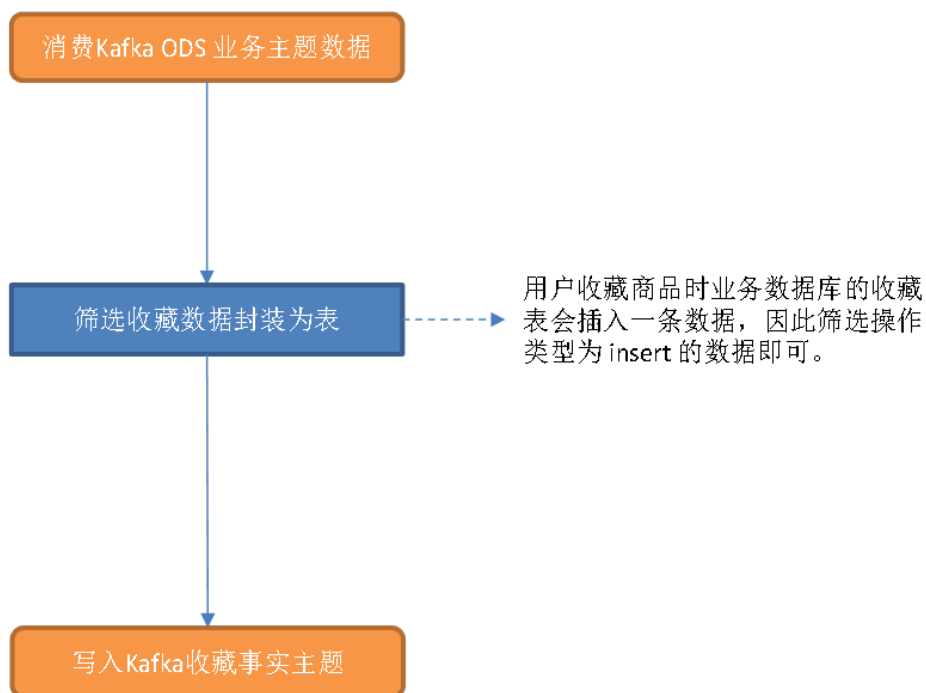
读取优惠券领用数据，写入Kafka 优惠券领用主题



让天下没有难学的技术

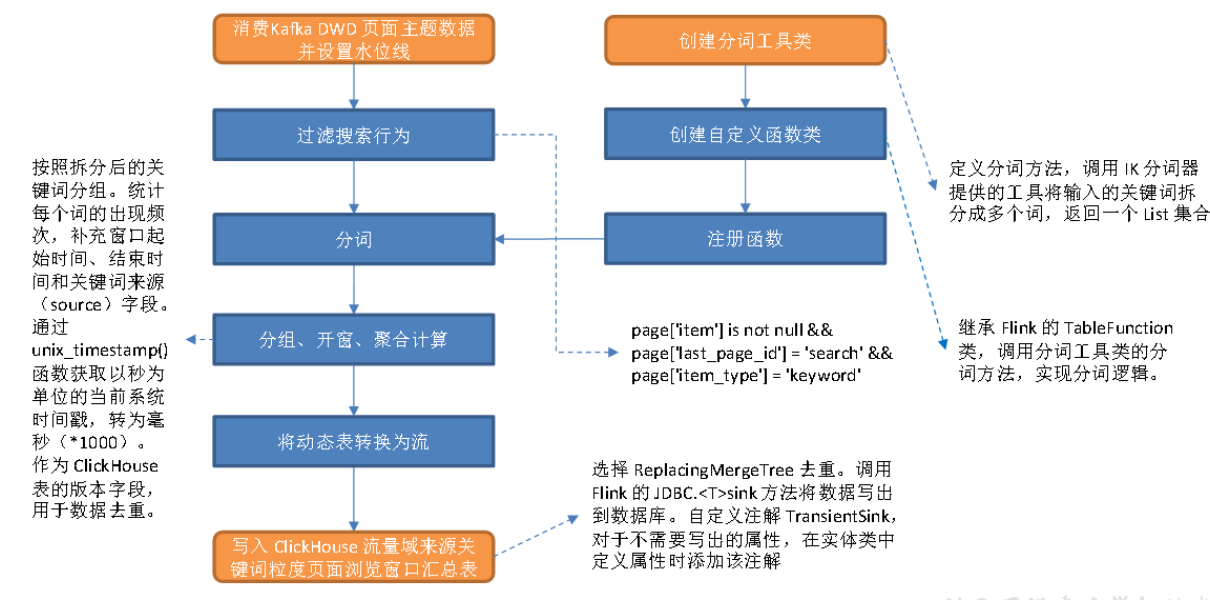
## 收藏商品，评价，用户注册需求

读取收藏数据，写入Kafka 收藏主题



流量域来源关键词粒度页面浏览各窗口汇总表

从Kafka 页面浏览明细主题读取数据，过滤搜索行为，使用自定义UDTF（一进多出）函数对搜索内容分词。统计各窗口各关键词出现频次，（查看用户搜索的频次哪个最高）写入 ClickHouse。

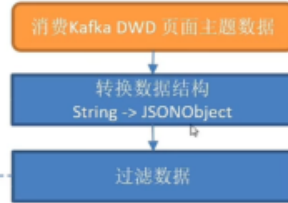


三、DWS层

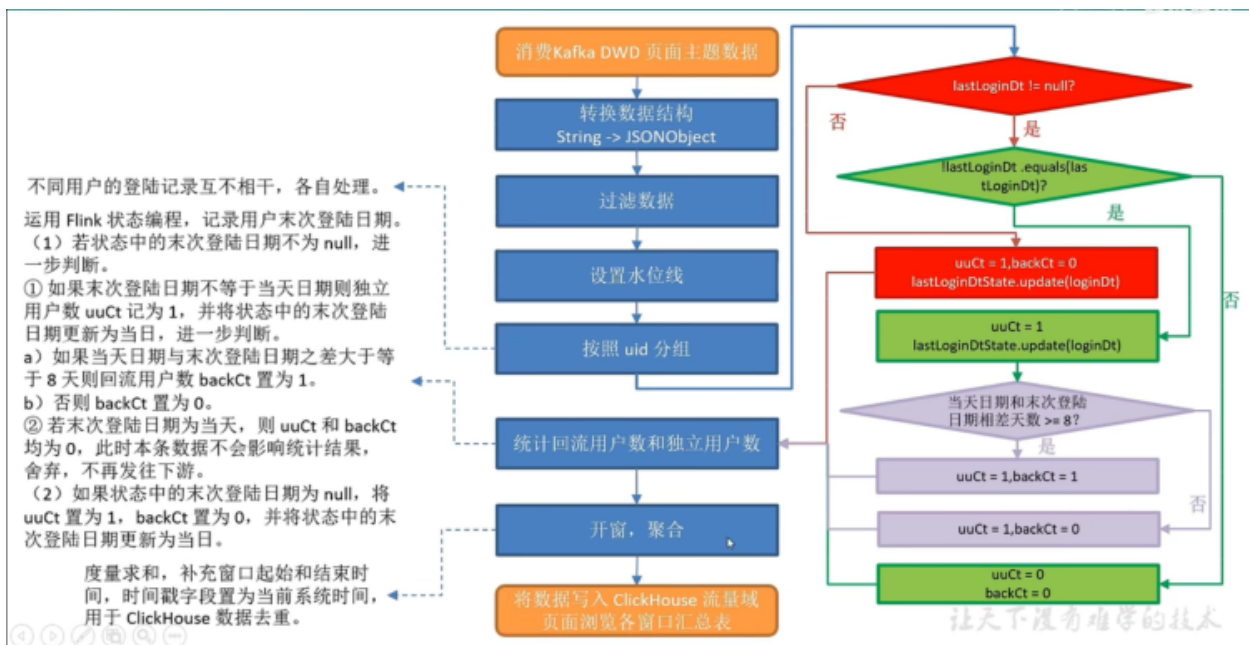
用户域用户登陆各窗口汇总表

从 Kafka 页面日志主题读取数据，统计七日回流用户和当日独立用户数。

统计的指标与用户有关，uid 不为 null 的数据才是有用的。此外，登陆分为两种情况：（1）用户打开应用后自动登录；（2）用户打开应用后没有登陆，浏览部分页面后跳转到登录页面，中途登陆。对于情况（1），登录操作发生在会话首页，所以保留首页即可；对于情况（2），登录操作发生在 login 页面，login 页面之后必然会跳转到其它页面，保留 login 之后的页面即可记录情况（2）的登陆操作。  
 综上，我们应保留 uid 不为 null 且 last\_page\_id 为 null 或 last\_page\_id 为 login 的浏览记录。



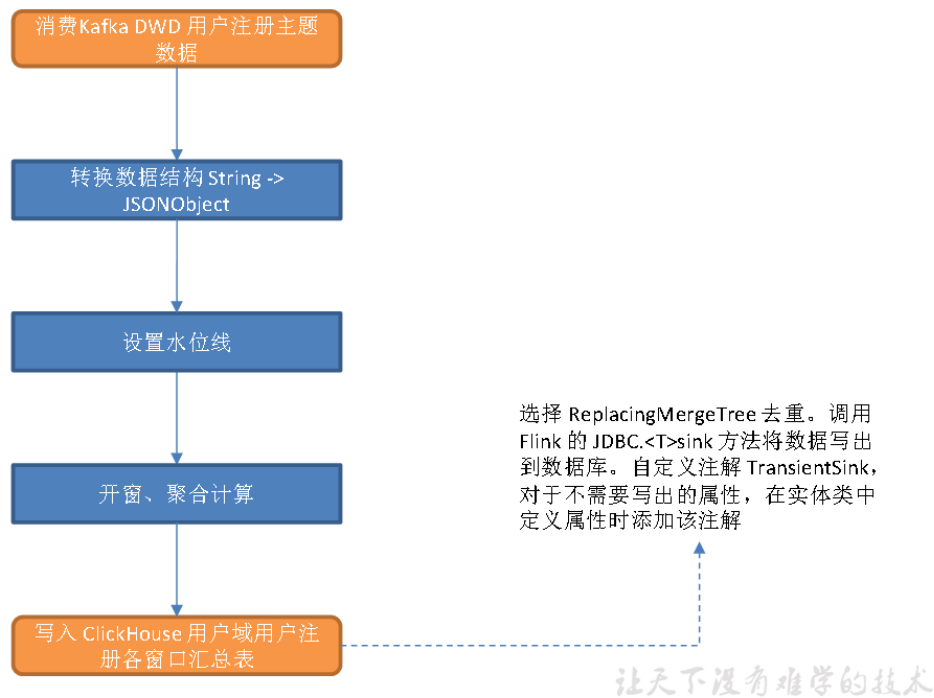
让天下没有难学的技术



## 用户域用户注册各窗口汇总表

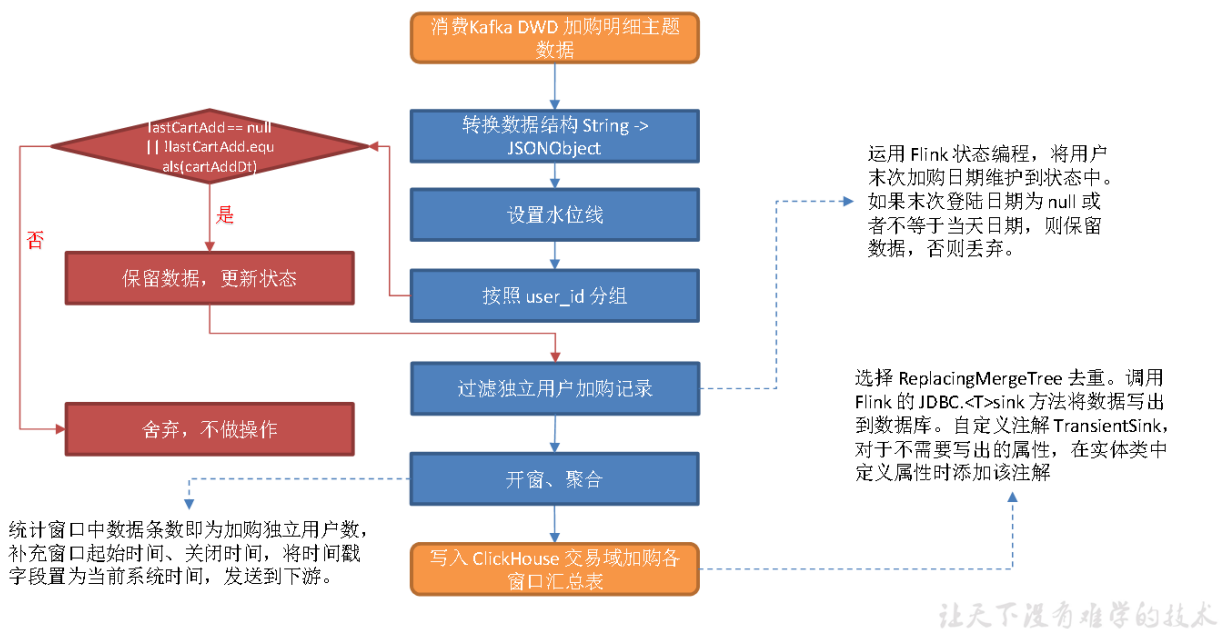
DWD 层用户注册表中读取数据，统计各窗口注册用户数，写入ClickHouse。





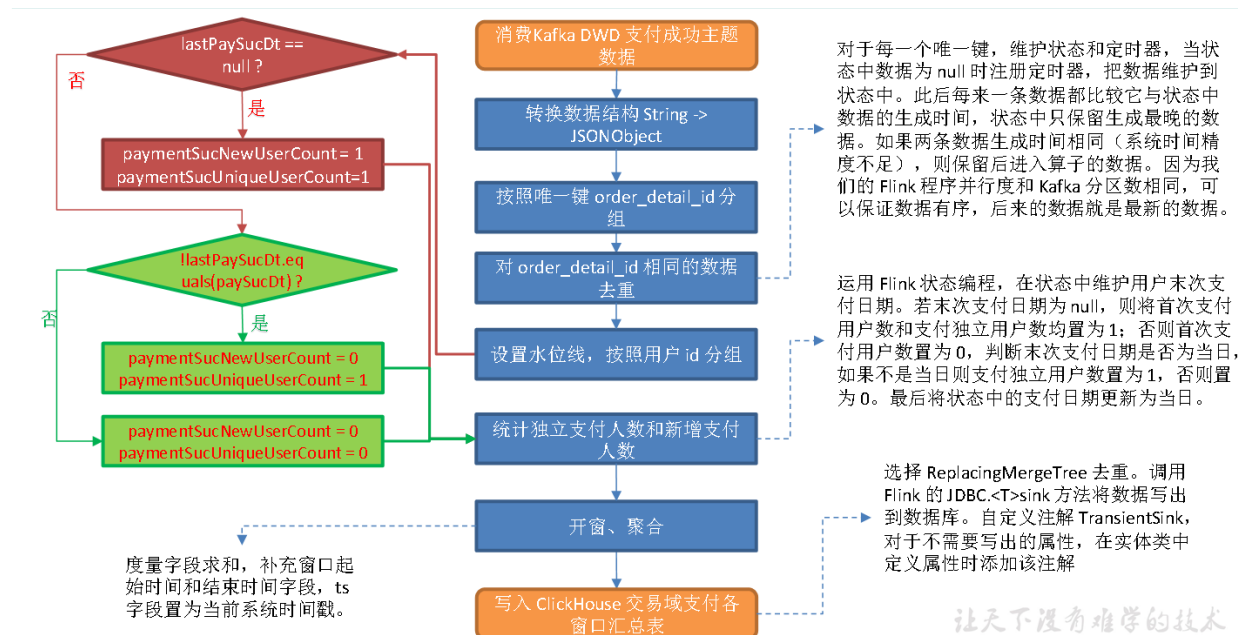
## 交易域加购各窗口汇总表

从Kafka 读取用户加购明细数据，统计每日各窗口加购独立用户数，写入ClickHouse。



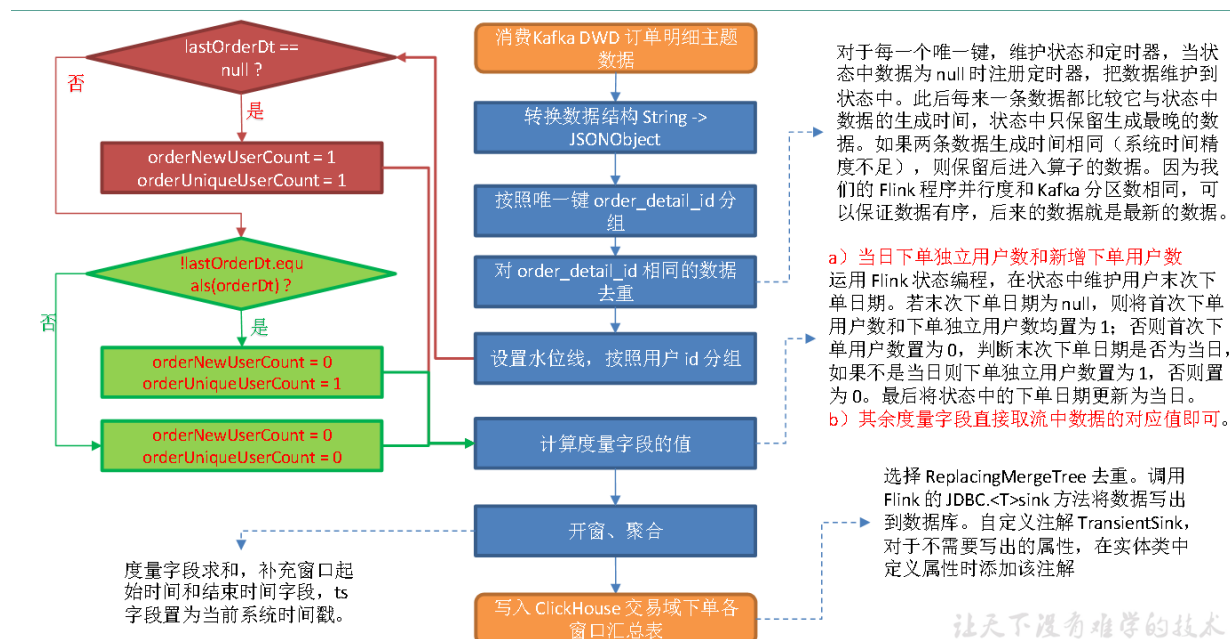
## 交易域支付各窗口汇总表

从Kafka 读取交易域支付成功主题数据，统计支付成功独立用户数和首次支付成功用户数。



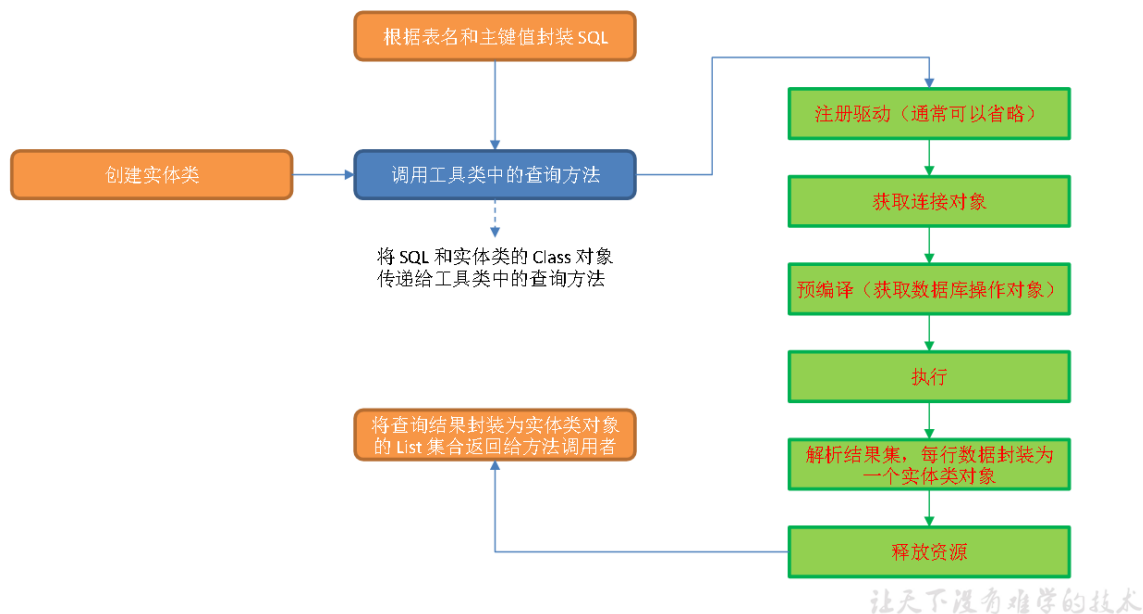
## 交易域下单各窗口汇总表

从 Kafka 订单明细主题读取数据，对数据去重，统计当日下单独立用户数和新增下单用户数，封装为实体类，写入ClickHouse。



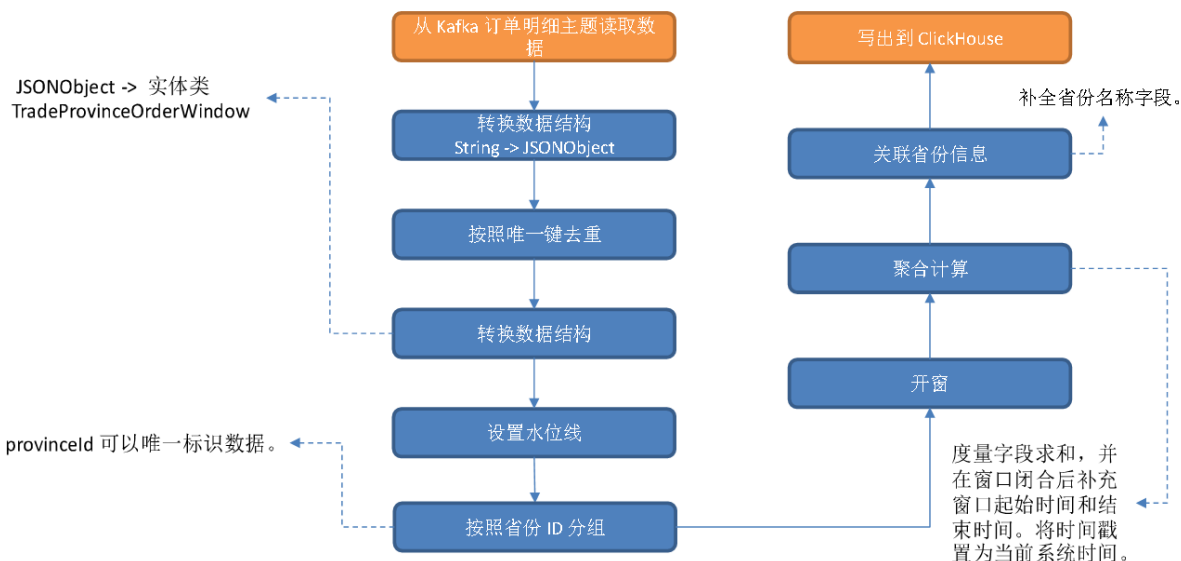
## 交易域用户-SPU粒度下单各窗口汇总表

从Kafka 订单明细主题读取数据，过滤null 数据并按照唯一键对数据去重，关联维度信息，按照维度分组，统计各维度各窗口的订单数和订单金额，将数据写入ClickHouse 交易域品牌-品类-用户-SPU粒度下单各窗口汇总表。



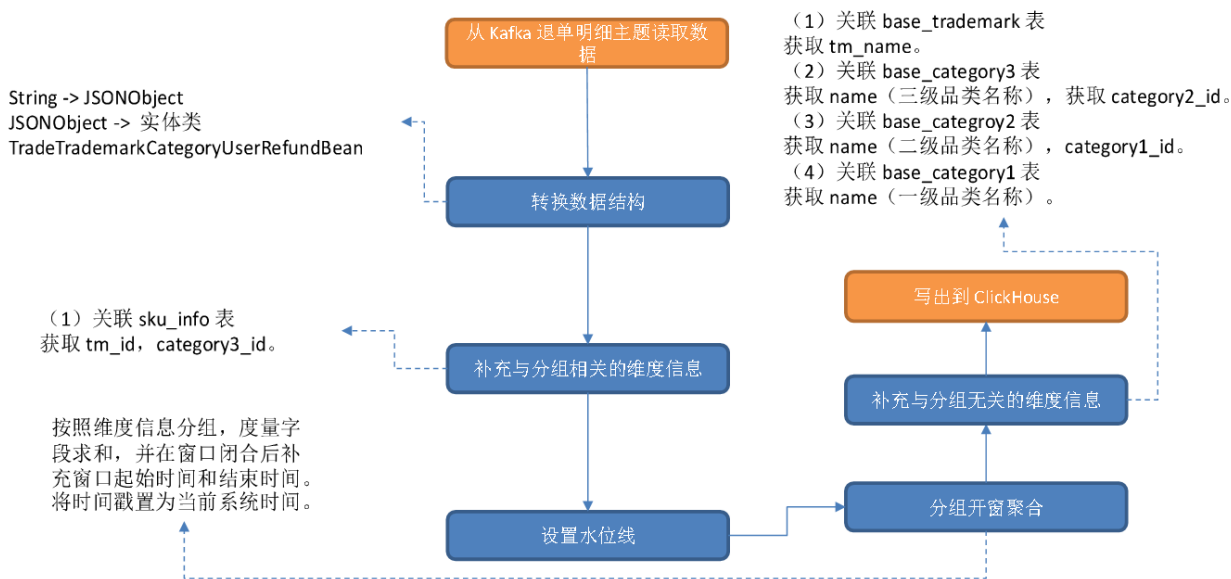
## 交易域省份粒度下单各窗口汇总表

从Kafka 读取订单明细数据，过滤 null 数据并按照唯一键对数据去重，统计各省份各窗口订单数和订单金额，将数据写入ClickHouse 交易域省份粒度下单各窗口汇总表。



## 交易域品牌-品类-用户粒度退单各窗口汇总表

从Kafka 读取退单明细数据，过滤null 数据并按照唯一键对数据去重，关联维度信息，按照维度分组，统计各维度各窗口的订单数和订单金额，将数据写入ClickHouse 交易域品牌-品类-用户粒度退单各窗口汇总表。



让天下没有难学的技术