# 用户行为分析项目拆解

## 自我介绍

面试官您好，我是xxxxx

本科毕业于XXXX，计算机科学与技术专业，本科期间绩点3.7，排名年级前十；

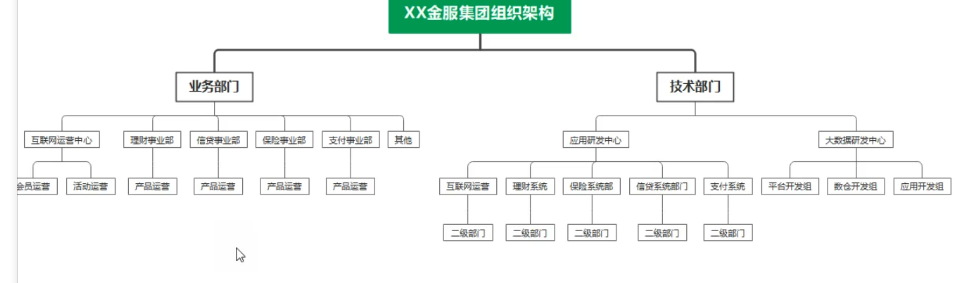
硕士就读于xxxxx学，软件工程专业，目前研二

研一期间在校参与了实验室与华为合作的Mindspore项目组件开发，协助师兄开发神经网络模型Rexnet;以及与中电20所的合作的跑冒滴漏智能检测项目；并于研一的暑假在杭州的网易严选实习了4个月，主要担任数据开发实习生，进组之后主要做的就是网易严选电商（神相—用户行为分析平台，专门做用户行为分析的一个数据产品）用户行为分析板块的相关工作，数据源主要来自三端的埋点数据app,wap,pc端。通过数据集成工具和数据开发过程实现数据分层，满足业务的各类用户行为数据应用，主要对接的是用户行为分析平台，比如事件分析，漏斗分析，留存分析，网页热力分析，用户群像分析等；

个人负责的内容主要包括：

* 负责对接网易严选电商平台用户行为分析日常漏斗模型的配置，这块主要是协助业务配一配漏斗；
* 完成渠道引流和用户行为相关的报表的开发；
* ODS层数据采集工作，完成相关数据的接入,以及规整化清洗等工作；
* 协助业务进行日常全链路行为表的数据核对工作，业务的临时取数工作，SQL编写工作等；
* 参与部分统计类标签和规则类标签的开发、用户行为统计标签的开发，协助开发了规则类标签；
* 协助优化全链路表中曝光表的百亿曝光、漏斗交易数据存在的性能与倾斜问题；

## **实习介绍：用户行为分析项目**

1. 你们组多少人？公司架构什么样子？  
     
    我们是大数据研发中心部门，专门搞平台的，搞应用开发的，我们是数仓开发组，负责用户行为分析这条线的，加上用户画像标签开发的人员，一共十多个人；主要是数仓开发+运维+产品经理+数据治理；其他的业务线有营销，采购（供应商选品），仓储，物流；
2. 进公司干了什么？

进去以后有1对1带教，导师会给我制定一个新人学习计划，第一周没有什么开发的需求，从熟悉平台开始，主要是看一些学习文档，初期跟我导师参加需求评审学习一下企业开发流程，然后导师会给排一些简单的需求，然后就是渐渐开始进行开发工作（组内氛围很好，前辈都很照顾新人，都会耐心回答一些开发上的问题），主要是开发一些报表，以及日常的漏斗模型配置，后面还带着参与了用户画像标签的建设，然后大促活动会协助导师进行一些调优和数据探查工作。

1. 一些开发常识

开发平台：网易易数一站式开发平台（数据传输，离线开发，任务调度，运维，数据治理集成），

主要使用spark sql进行开发；

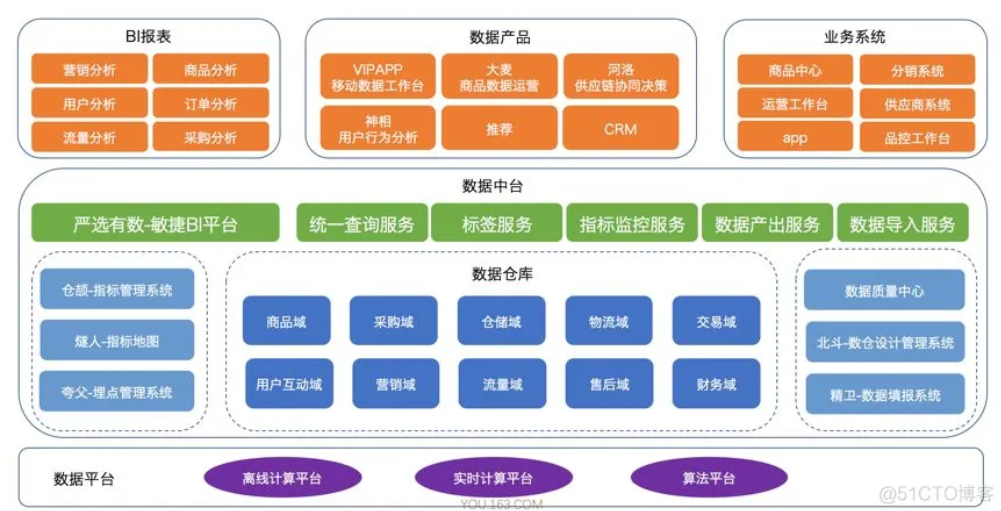
开发流程：任务上线 => 新建任务 => 代码编写 => 任务提交 => 导师review => 冒烟测试 => 发布上线

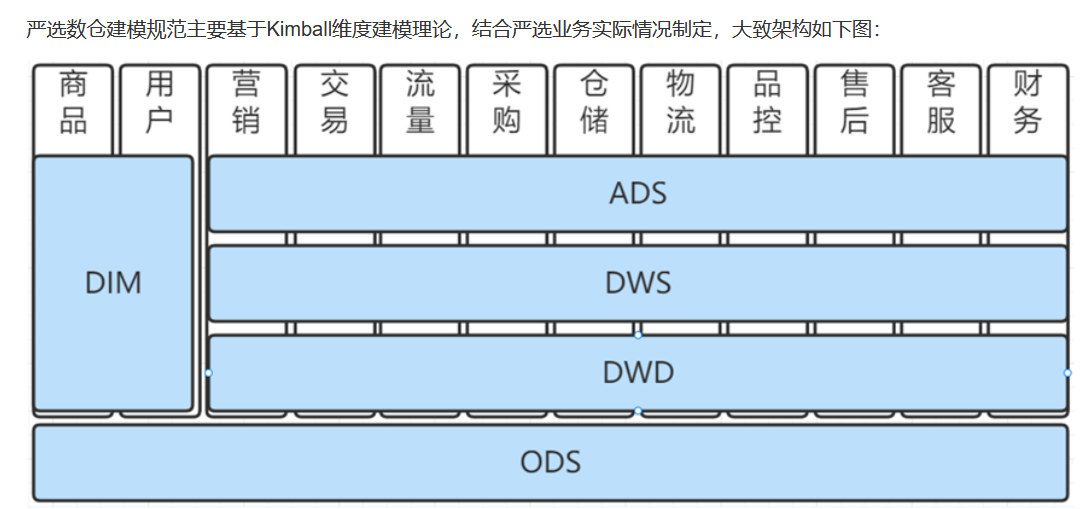
1. 公司大致信息？
2. 开发规范
3. 具体开发流程

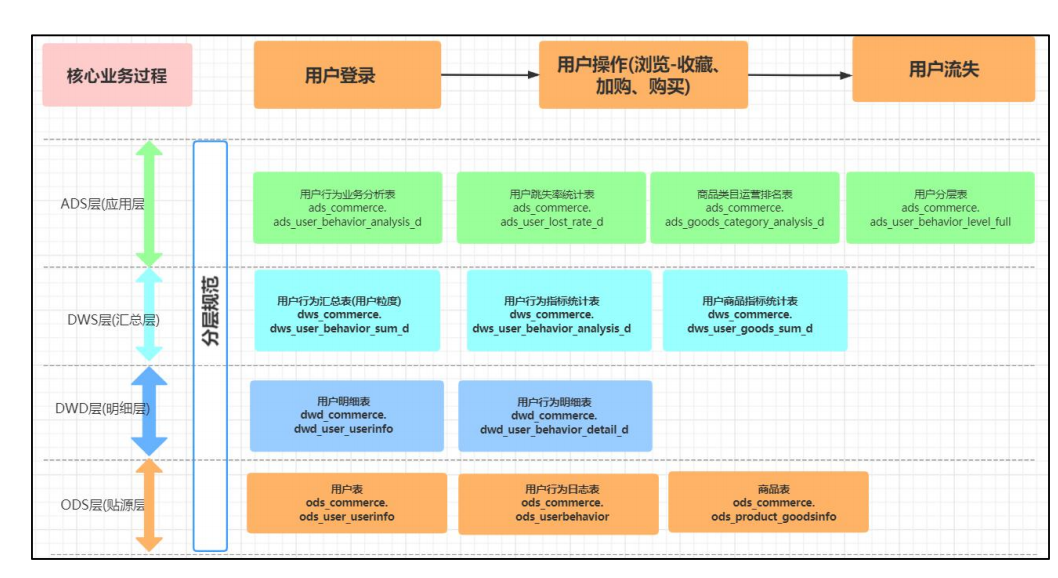
具体开发是优先参考以前方案，因为好的模型或者链路，它是有很强的参考性和复用性。

具体的的操作是 查询以前的类似的报表，找指标口径文档，然后去资产管理平台找到这个报表对应的节点ID，去运维平台看报表整体生产链路，参考开发代码。根据技术方法产出的数据模型进行开发。然后是测试上线，配置DQC和SLA。

## 业务场景，项目架构，数仓架构，开发管理规范

1. 业务场景  
    用户行为分析场景建立在严选电商场景之上，主要是支撑神相用户行为分析平台；包括一些件分析，漏斗分析，留存分析，网页热力分析，用户群像分析等；用于辅助业务做一系列精细化的运营。
2. 项目架构  
   
3. 数仓架构

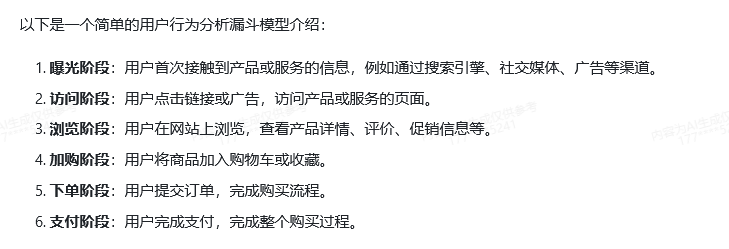




## 4.漏斗模型配置和优化维护（价值：直接价值+间接价值（报表）用户行为分析的业务流程）

### 4.1 漏斗模型介绍

主要用于分析一个多步骤过程中每一步的转化与流失情况。

  
我在实习过程中主要做了一些用户行为相关的日常漏斗的一个配置，

### 4.2 实习中漏斗模型配置案例以及对应需求承接流程

主要是配合业务去使用这个漏斗模型这块，实际开发不是我来做的，我主要负责漏斗的配置，比如业务方提出需求可能要哪些漏斗，我就在平台漏斗配置模块，去创建漏斗，  
包括漏斗拿出来数据和业务配不上的时候会协助进行一些数据探查；

**比如说当时业务方想做一个运营位的优化以及，板块下推荐商品和该品类热销商品的一个匹配情况（可能发现热销产品，并不是通过板块推荐渠道来的，而是用户自行搜索来的；那么是不是可以考虑，将这些产品所在坑位进行一个顺序优化，放在更首位的位置）；**

**我就要配严选养生板块的一个各个坑位的一个用户转化情况；**

当时配置了从主页浏览到 养生板块下各个板块下，完成订单的一个转换情况；

还有对某几个业务方给的特定商品的一个转换情况，我记得什么燕窝和什么益生菌；

就需要创建个漏斗，配置一个时间的范围：也就是漏斗第一个步骤生成的时间

以及窗口期：用户完成漏斗的时间限制，即只有在这个时间范围内，用户从第一个步骤，行进到最后一个步骤，才能被视为一次成功的转化。

然后就是配置漏斗步骤了，去根据业务需求，  
去定义每一层的内容和一个筛选条件；

* 第一层是主页浏览情况
* 第二层是一个严选养生馆模块的一个点击情况；
* 第三层是配置养生馆内各个模块的一个点击情况
* 第四层就是配了一下商详浏览情况
* 第五层完成订单的情况  
    
  有时候他要看具体比较火的某个商品的卖的转换情况，就比如燕窝和益生菌类的商品，就需要配置一些具体商品id，层间进行关联，精确具体商品的一个转化情况；

包括设置漏斗任意几步的属性进行关联，假设需要精确了解用户浏览了某个商品，并完成此商品购买的情况，创建的漏斗为 浏览商品详情页 -> 提交订单 -> 支付订单，并且在每个事件中设置了商品 ID 的属性，此时就可以将该属性作为关联 ID，以保证用户浏览商品详细页到支付订单的商品都是同一个。如果不设置商品 ID 为关联属性，则用户浏览商品详细页与支付订单的商品不是同一个，也会被算作转化成功。  
  
漏斗完成后，可以对每个阶段用户详情进行查看，以及对应用户画像查看，判断一个什么类型的用户流失； **漏斗配置完成后怎么 提交任务，要给谁看？**

漏斗维护相关

* 做了一些漏斗漏斗模型问题的收集，比如  
  精确具体商品的一个转化情况；

包括设置漏斗任意几步的属性进行关联，假设需要精确了解用户浏览了某个商品，并完成此商品购买的情况，创建的漏斗为 浏览商品详情页 -> 提交订单 -> 支付订单，并且在每个事件中设置了商品 ID 的属性，此时就可以将该属性作为关联 ID，以保证用户浏览商品详细页到支付订单的商品都是同一个。如果不设置商品 ID 为关联属性，则用户浏览商品详细页与支付订单的商品不是同一个，也会被算作转化成功。

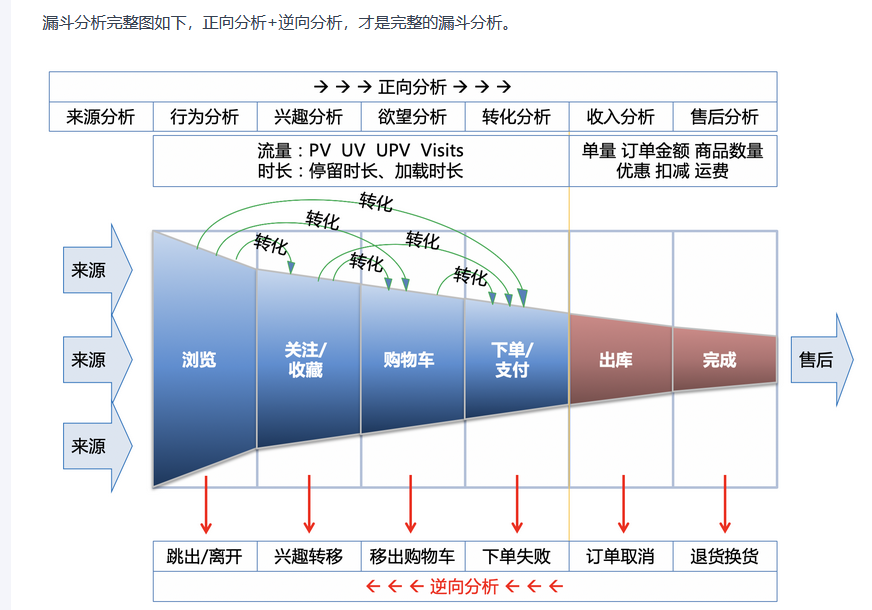
如果发现漏斗产出数据实时跑数，有时跑不出数；数不太对；我协助进行做一些漏斗的（底层大宽表）做一些（拆分）数据探查；对某个阶段的数据进行一个查看；

优化，

大促活动（618）期间，各个活动之间的漏斗转换出现数据倾斜问题；

补充：

AARRR模型是指Acquisition、Activation、Retention、Revenue、Referral，即用户获取、用户激活、用户留存、用户收益以及用户传播



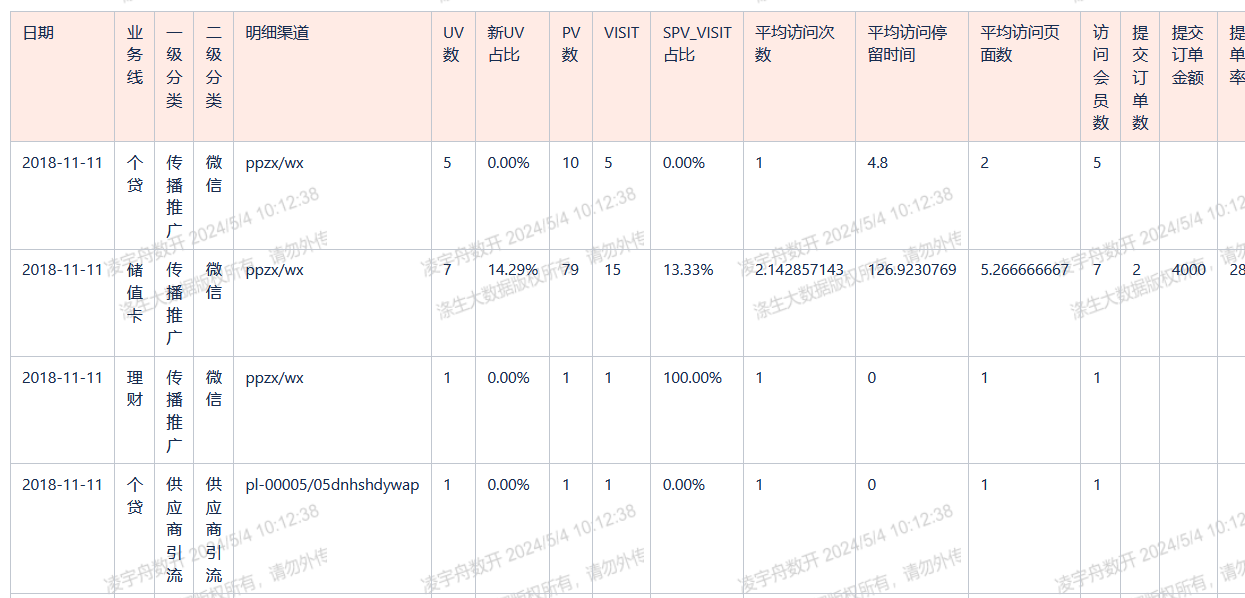
添加至购物车环节为例，这里可能需要结合其他的数据指标来进行综合分析，比如页面停留时长、页面跳出率、上级流量来源、历史数据、同行业平均数据等，从而确定是商品本身存在问题，不具有吸引力，还是说添加至购物车的操作步骤存在问题、还是商品详情页与列表页出入过大导致跳出率很高、或者是推荐给了错误的用户群体等等

## 渠道引流和用户行为相关报表开发

这块主要是一个是为了**优化营销策略**：通过分析不同渠道的引流效果，企业可以评估渠道的投入产出比，从而优化营销资源配置，提高营销效率。  
一个是**用户行为洞察**：用户行为分析可以帮助企业了解用户的真实需求和行为习惯，为产品设计、营销活动提供依据，以更好地满足用户需求。  
  
我这工作一般是 产品经理提了需求后，我导师再对我进行一个需求下发和需求排期，会将具体报表需求告知我（这个过程主要通过我们公司需求管理平台（类似jira））,我可以知道具体报表结构，要哪些维度，哪些指标，以及对应一个指标口径；包括我要去哪些表中取数；基本上都会做一个整体概述，我的任务就是把这张ads层表开发出来；然后我实习期间主要是开发了渠道引流和用户行为分析相关一些报表；

我记得主要当时排期比较长的，就是这张渠道引流报表；这张表数据主要来源于我们的一张全链路表，一张大宽表；  
结合全链路表，准备报表所需要的数据，例如用户访问日志、交易记录、商品信息等。需要确保数据质量，并进行必要的清洗和转换。

### **渠道引流报表开发 （粒度！！！）**



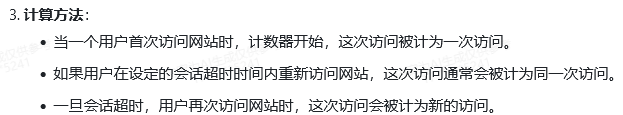
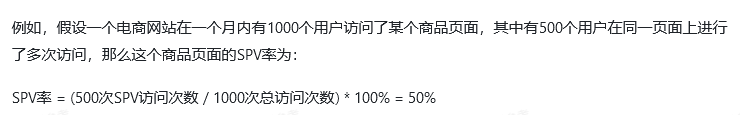
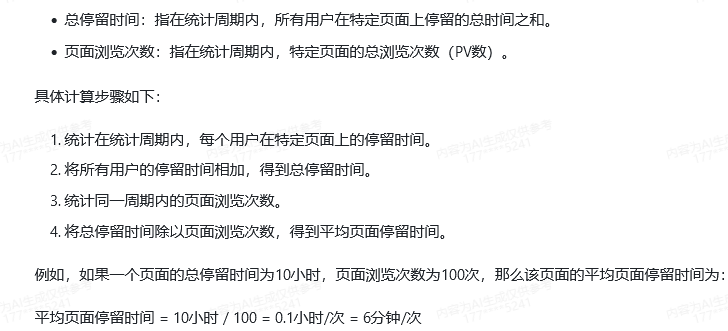
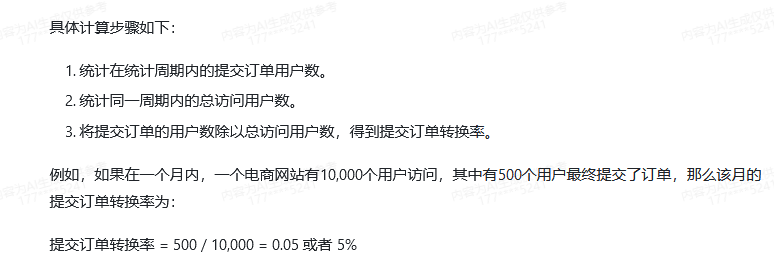
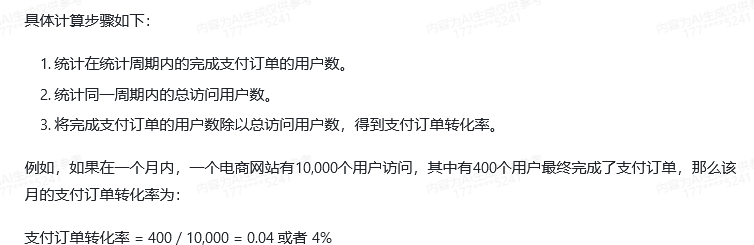




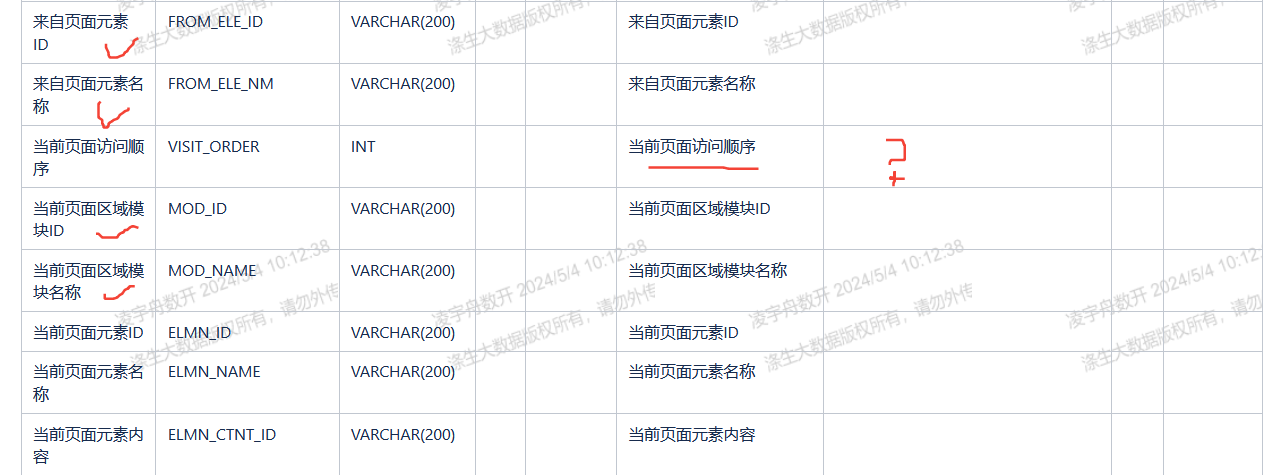
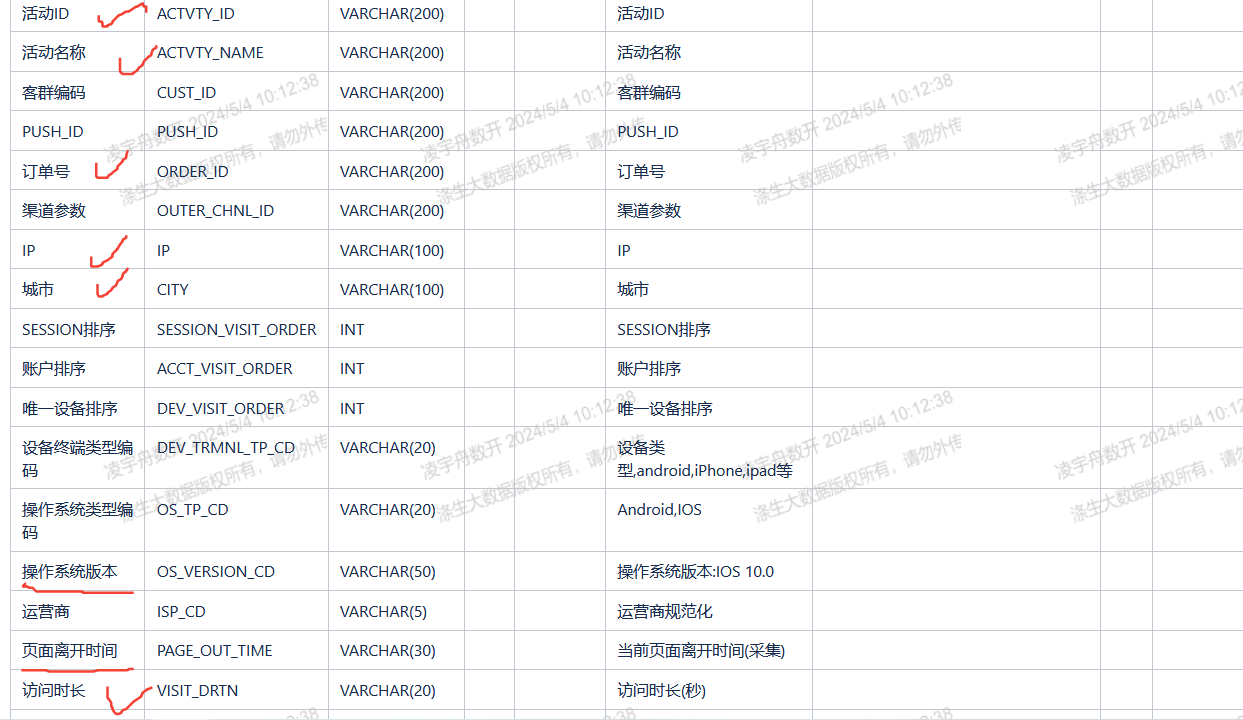
维度：日期，;统计周期（1/7/30）

粒度：渠道

指标：

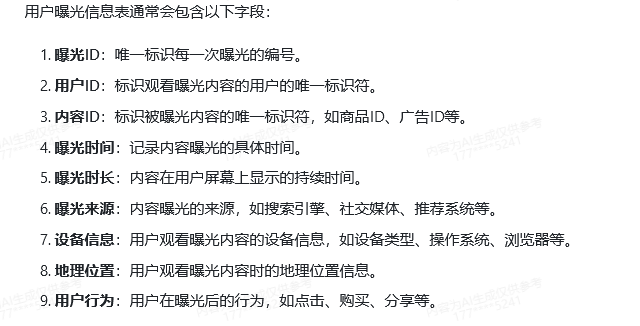
* UV数：统计某个日期，各个渠道（app,pc）带来的独立访客数，以及一个新访客数占比  
  涉及：全链路表
* PV数：每次页面访问都会产生一个PV  
  涉及：全链路表
* Visit：访问次数，跟踪session\_id;来源全链路表  
  
* SPV数占比：同一页面上的多次访问的用户占总用户的比例；  
    
  例如，用户可能在查看一个商品页面时，多次返回该页面以查看商品详情、评论或其他相关信息。
* 平均访问次数：总访问次数 / 访问用户数
* 跳出率，访问一个页面的总数/总访问数
* 平均访问页面数： 总页面浏览数（PV数） / 总访问次数
* 平均访问停留时间：  
  
* 访问的会员数：会员计数，涉及 会员基本信息表
* 提交订单数：提交订单总数，涉及订单基本信息表
* 提交订单金额：金额汇总
* 提交订单转化率：  
  
* 支付订单数：订单支付数
* 支付订单金额：订单支付总金额
* 支付订单转化率： 涉及订单表  
  

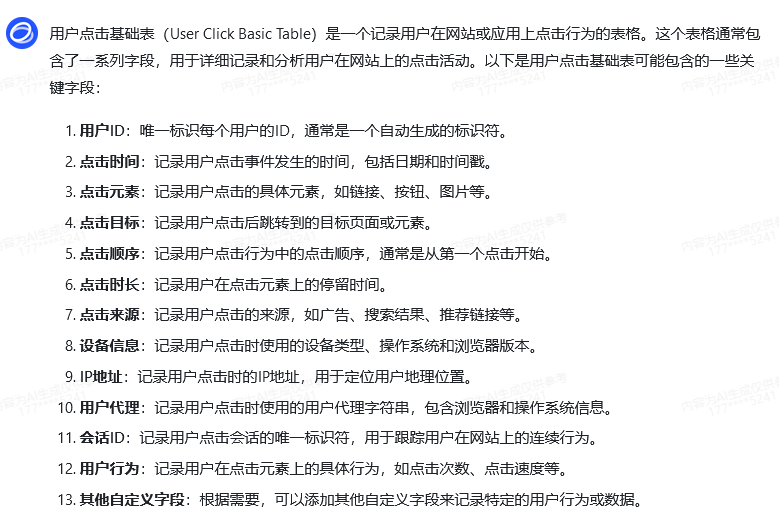
### **全链路表介绍**

全链路表主要就是集成了一些用户行为分析的核心表，用户访问基础表啊，用户点击基础表，用户搜索信息表，用户曝光信息表；总之包含了用户的各种行为数据，如页面浏览、搜索、点击、加购、以及这些行为发生的时间、用户身份、设备信息的一张大宽表（300多个字段）；  
  
  
  


* 分区：天
* 更新周期：日增表
* 表一天的数据量：几十亿

用户曝光信息表



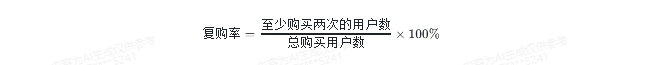
用户访问信息基础表  
  
用户点击基础表  


### **用户行为分析相关报表**

**热门活动页面浏览报表：（**显示用户在访问当前页面之前和之后访问了多少个页面**）**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **页面id，页面名称** | **uv** | **pv** | **平均停留时间** | **跳出率** | **页面浏览深度**： |  |
|  |  |  |  |  |  |  |

**某类目商品浏览，购买行为报表：（通过分析用户浏览商品的行为，了解用户对商品的兴趣和喜好）**



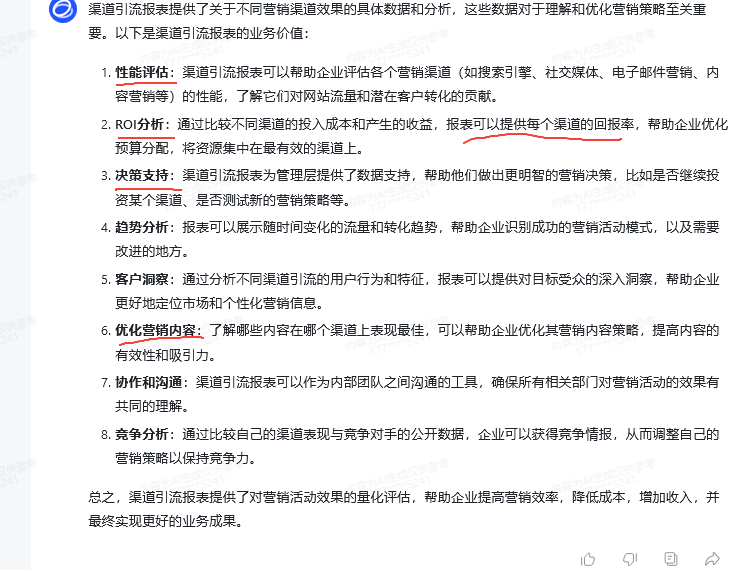
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **商品id** | **搜索量/商详页uv/浏览排名** | **商详页Pv添加购物车点击 UV；**  **添加购物车点击 PV；**  **立即购买点击 UV；**  **立即购买点击 PV。** | **平均停留** | **最终下单数量** | **二次下单复购** |
|  |  |  |  |  |  |

### **报表的业务价值**

为公司的用户增长团队提供了基础的分析数据，原本他们每周需要

花 2 天时间来做数据统计、图表分析，现在只需要 0.5 天即可获取到他们的数据，

极大地提升了他们日常取数的效率。



## ODS数据采集工作，包括用户行为，交易，商品等相关信息的接入；

用户行为相关数据，主要是前端一些用户行为埋点数据，

交易相关数据，

商品相关数据，

## 协助业务进行日常全链路行为表的数据核对工作，业务的临时取数工作，SQL编写

因为这张表数据量大，其实生成逻辑还挺复杂，有时候可能数据对不上，一些数据倾斜问题，我主要是协助导师做一些数据探查；具体的维护任务还是我的导师做；  
数据核对：

获取全链路行为表的最新数据。

与业务团队一起核对数据，确保数据的准确性和完整性。

发现数据相关问题，比如  
**业务团队发现，在某个时间段内，部分用户在访问了商品页面后，没有记录到加入购物车的行为数据。这意味着这些用户在商品页面上的行为路径可能不完整，无法准确评估商品页面的转化效果。**

* **转换路径缺失**：业务团队发现某些用户的转化路径数据不完整，例如，他们可能没有完成预期的转化步骤，如从商品页跳转到购买页。
* **时间戳数据缺失**：某些用户的行为数据的时间戳可能缺失或不准确，导致无法准确追踪用户的行为路径和转化时间。
* **异常值的缺失**：业务团队发现某些用户的异常值数据缺失，例如，可能没有记录异常行为，如异常高的访问频率或异常低的页面停留时间。

主要解决方案：

* 检查埋点代码  
  首先，需要检查商品页面上的埋点代码是否正确配置，确保能够捕获加入购物车的行为事件。如果埋点代码存在问题，需要及时修复或更新。
* 数据修复  
  对于已经发生的缺失数据，可以尝试通过其他途径获取，例如，查看前端日志或后端日志，以查找相关的操作记录。如果可以找到缺失的数据，可以手动将其补充到全链路行为表中。
* 系统优化：如果数据缺失问题经常发生，可能需要对系统进行优化，以提高数据的捕获和记录准确性。例如，可以增加埋点代码的检查频率

临时取数：  
业务人员可能有些临时的需求，要我们做一下临时的某些指标的取数；  
一般我都会和业务人员去交涉一下取数口径已经要从哪些表去取数。

比如618，做一些活动时候，电商业务团队需要分析即将到来的促销活动前几周内，用户在不同渠道的转化路径和购买行为。  
涉及用户行为全链路表、以及订单表等。

理解业务需求，确定需要提取的数据和字段，

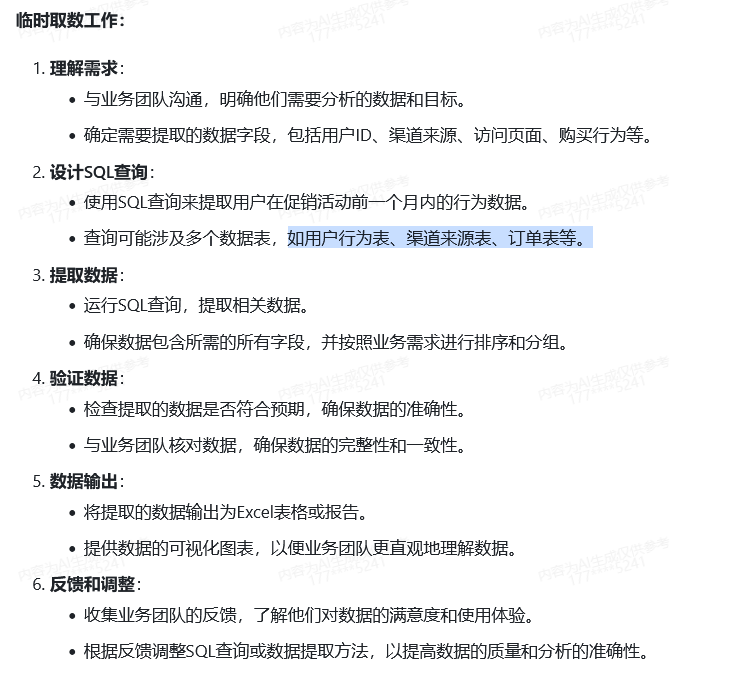
例如大促前

**用户在某些活动页上停留的平均时长。较长的平均停留时长往往表示用户对网站内容**

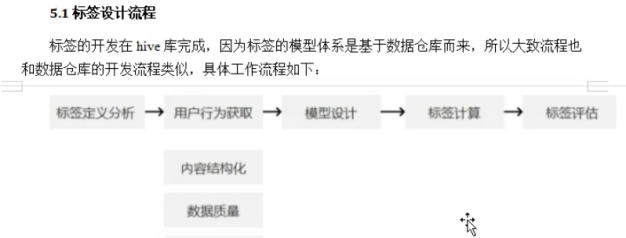
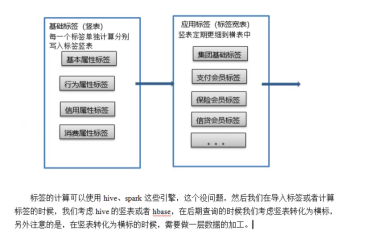
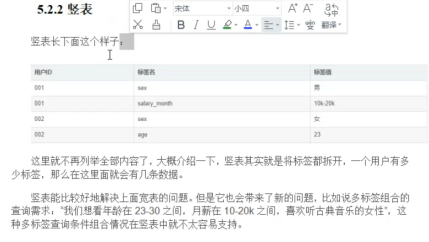
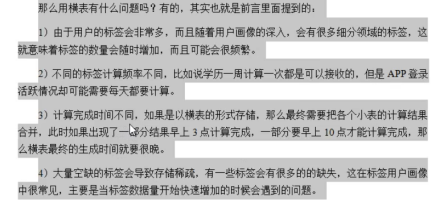
**的关注度高。**

**用户对某类热门商品的加购情况**

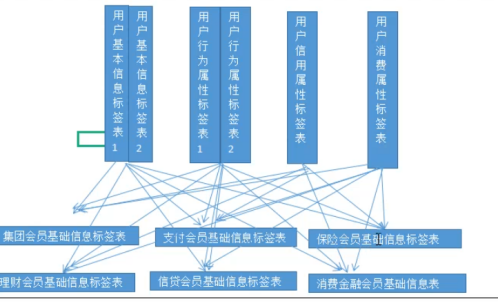
**活动下发一周，各个渠道引流情况**



## 负责部分统计类标签和规则类标签的开发如用户基本属性标签、用户行为统计标签、用户RFM模型等

1. 用户画像基本知识  
    标签？：标签是对业务实体（用户，商品，店铺）某个特征的刻画和描述，是一种面向业务的数据组织形式。比如 基本性别，年龄段，有车族，消费水平高低这些；例如，我们在大众点评上看到某家店是“必吃店”，企业的运营主要围绕业务对象和业务活动，所谓的精细化运营；  
    画像：多个标签的组合，为的是更好的去刻画一个业务实体，描述他经可能详细的特征  
    画像作为标签的集合，可以综合提供多维度的正交信息，帮助我们更准确、形象地洞察实体。
2. 严选画像架构  
    目前行业中的标签和画像主要都是用户标签和用户画像，其应用场景主要是精准营销和消费者洞察。严选的业务场景较多，如消费者洞察、供应商寻源、爆品打造等，涉及到的核心业务实体比较多，如用户、商品、供应商、渠道等等。  
   
3. 标签开发流程  
    产品经理把标签开发逻辑口径理好，中高级建模，根据便签模型和口径，对标签进行开发； 需求承接，标签评审（开发什么标签），告诉我后，我从标签管理平台创建对应标签（录入元数据信息）；然后再进行标签开发（写好代码后审批后才能上调度系统跑批，还是说调度后有一部分结果了再审批），开发好后进行测试（测试怎么测，谁测？）；标签报警是谁配置？
4. 标签的下线，暂停
5. 标签规则的变更：修改的是元数据，但如果标签计算逻辑修改呢，
6. 标签变更后，会记录其变更数据，记录标签的新建，审批，编辑，启停操作人（可回滚）
7. 用户画像标签表结构设计
8. 标签开发  
   
9. 横表和竖表  
   

前期开发一些竖表；后期开发一些横表（用于查询分析，圈群，直接对接画像产品相应模块；可以以宽表为单位进行，权限隔离）



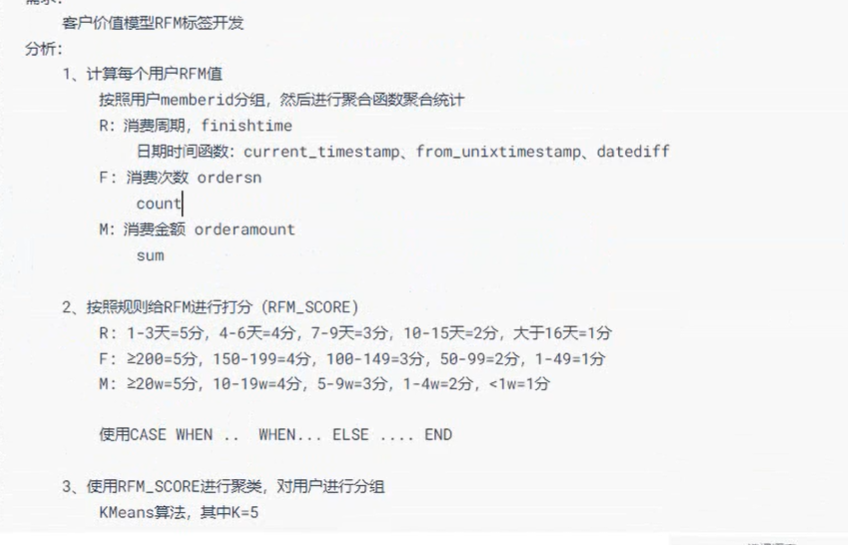
设备类型标签，

关联操作系统版本标签，

付款页面停留时间超过40s，

付款尝试次数超过2次的人（是不是没找到合适的支付方式，导致付款失败）

问题：

1. 基础标签开发哪些比较合适？（没开发过，这块体系已经较完善了,性别，城市，省份，常用电话，邮箱啥的，有不有小孩，已婚，注册日期）关联操作系统版本标签，关联操作系统版本标签，
2. 用户行为统计标签：（购买业务偏好：闪购，包邮，促销；最近一次商品收藏类目；站内广告偏好：首页，弹窗，悬浮；活跃时间段：白天中午晚上等等）;  
     
   近一月会员访问深度\_平均页面停留时间(（星选（做会员活动价的））页面)；近一月会员访问深度\_跳出率(（星选（做会员活动价的））页面)；计算会员近三个月访问（星选）页面的次数，开会员付款页面停留时间超过1mins；
3. RFM（最近一个客单价）模型：（衡量客户价值，给用户就行一个定级）属性（多少个标签值）：  
     
     
    --dwd.dwd\_act\_acct\_info 用户主表

--dwd.dwd\_pty\_pay\_svc\_user 用户注册信息表 --dwd.ssa\_epp\_login\_ftis 用户登录表 ；  
--订单信息表



1. 生命周期：帮着做了一些前置逻辑的实现，不是首先得吧一年内有过交易的，交易次数这些数据拿出；  
   
2. 怎么接的标签开发需求，需求包括哪些内容，为什么提这个需求？  
   标签的活是我导师给我安排的，需求评审是他参与的，具体是业务人员那边想迭代一批用户画像标签，他们就去讨论标签可行性以及之后的一个开发安排；我这边收到的需求，开发什么标签以及一些取数口径都是写好了的，我些在开发平台上写sql就好了；我具体开发了几个用户行为统计标签，巴拉巴拉（将逻辑）；主要是涉及（全链路表巴拉巴拉）； 其次我还帮我导师做了一些规则类标签，rfm模型标签的一些测试（那个迭代代码有点复杂，我也有过研究 巴拉巴拉；我主要是怎么测试）
3. 讲一下你们用户画像开发的总体流程以及模型架构？

这一块我涉及的不深，我大概说下，就是我们标签，是通过竖表+宽表的形式进行一个组织的；表签管理平台，创建好要开发的标签后，我们会对标签进行一个开发，也就是逻辑计算，像用户基本信息标签表，用户行为属性标签表，用户消费属性标签表我们都是竖表进行一个组织的，标签id,标签名称，用户id,标签类型，标签值这些的；然后标签开发完后，会将他们做一个行转列，会汇总到宽表上（面向用户画像系统一个查询，筛选用户id,可以读取其对应的画像信息）；然后这边还会将部分标签数据，按场景推数到mysql（用户标签的元数据，结果集的校验；（标签每天计算跑批任务，表量级的一个监控，以及同步的HBASE这种数据校验，以及波动检验相关的数据）；同步到业务系统，比如客服系统），HBASE（实时数据的查询，比如说渠道运营想做一个新用户的促进下单；可能会进行一个app的首页弹窗，运营就会通过组合标签（未注册，安装距今 小与多少多少天，筛选用户分群然后做一个push，推送广告（通过广告系统））），上；（因为跑用户相关的作业计算量非常大，相关 作业的执行一般使用 mp 或 spark 作为执行引擎，将跑出来的结果写入到 HDFS 上，这个作业过程不可能在 MySQL、hbase等其他数据库中实现，但是当我们进行业务处理 展现的时候需要用到其他数据库，所以需要将对应的数据推送到对应数据库中。）

1. 存储系统介绍

元数据管理；结果集校验；



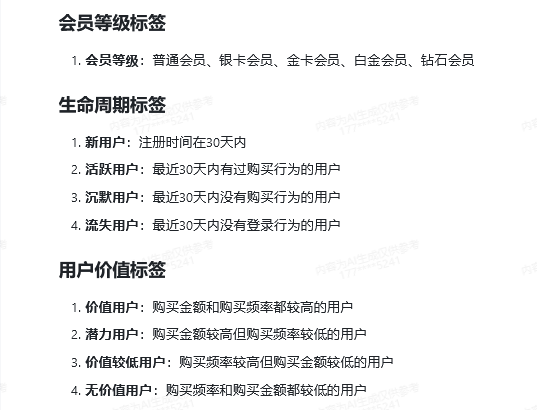
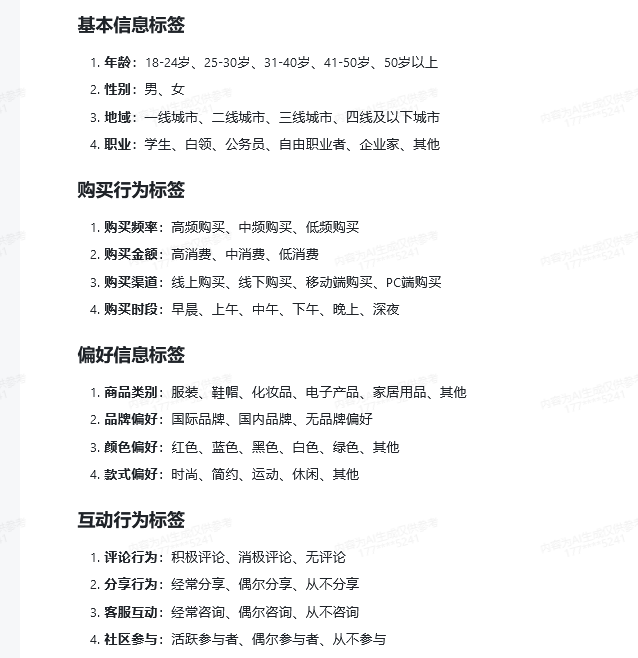
HBASE ：这样每天画像系统的 ETL 调度完成后对应人群数据就被推送到 HBase 数据库进行存储。满足条件的新用户来访 App 时，由在线接口读取 HBase 数据库，在查询到该用户时为其推送该弹窗。

1. 当你们用了宽表，竖表，数据量很大你们是怎么优化的：建立分区表（日期，某个标签值用户量很大时(也可以用标签值进行分区)，主要目的是优化查询效率）；标签生成脚本调优，hive sql调优那种了；
2. 介绍一下你们的标签体系：用户行为维度，用户属性维度；

w我了解到的，标签体系，主要是用户属性标签，用户行为标签，用户消费标签：

其中用户属性：基本属性标签，年龄，性别，地域，职业，设备类型；包括一些会员等级，价值内标签（RFM评级），是否首单用户啊等等；用户行为标签：首单营销方式，进30天行为标签，登录次数，高频活跃时间段（早？），停留某页面时间，跳失率；用户消费标签，购买品类（电子产品，日用啊）这些；大概就这些；

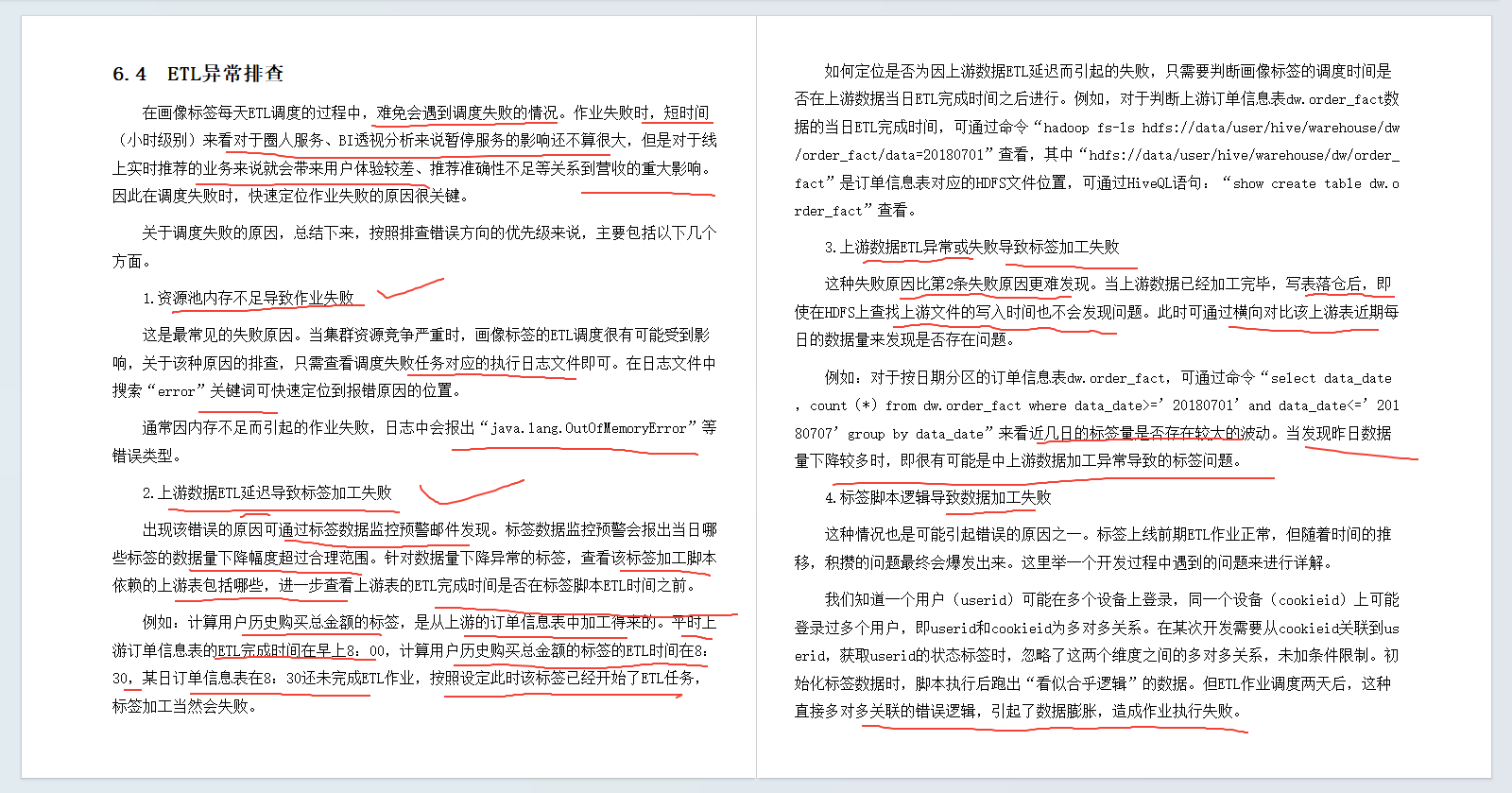
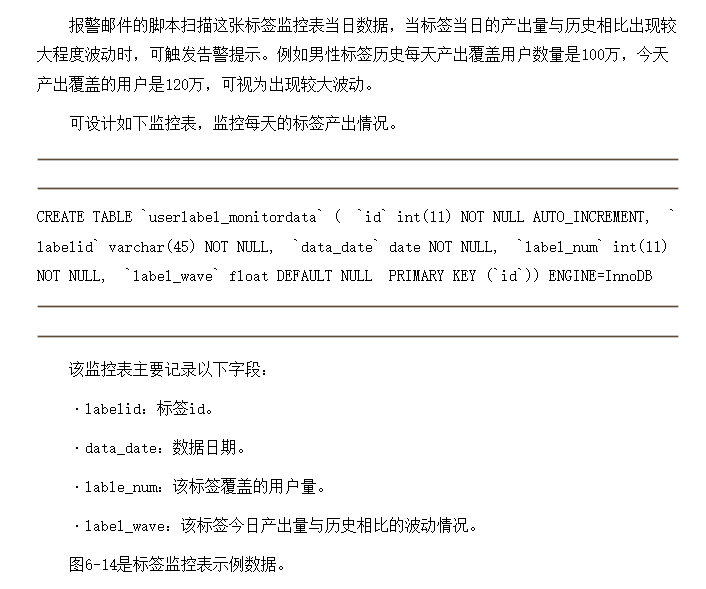
|  |
| --- |
| 重要价值用户 |
| 重要发展用户 |
| 重要保持用户 |
| 重要挽留用户 |
| 一般价值用户 |
| 一般发展用户 |
| 一般保持用户 |
| 一般挽留用户 |



1. 标签怎么进行的调度，部署，上线（标签测试怎么做的）
2. 用户画像的竖表和横表存在哪层，dw汇总层吗，那dm层存的是啥
3. 标签开发好后的，元数据管理，以及数据监控是谁来做，是开发人员来做吗

相比Hive，由于MySQL等关系型数据库对小量的数据读写速度较快，所以开发时考虑将数据的监控相关表维护在MySQL中。

数据监控预警整体来看涵盖下面几个主要模块。

**标签监控预警：用于监控每个标签当日的ETL是否产生问题，当数据量超出正常范围时，发出报警邮件。  
服务层数据监控预警：数据从数据仓库走出提供到服务层时，该过程中是否正常进行，一般通过对比数据仓库（Hive）中各业务线的数量和各业务系统（如MySQL、Hbase、csv文件等）中对应的业务线的数量进行监控。**

## 协助优化全链路表中曝光表的百亿曝光、漏斗交易数据存在的性能与倾斜问题

做618核心活动的时候，用全链路表做一些页面核心坑位曝光分析的时候，可能是会出现一些数据倾斜问题；

基于渠道做一些数据分析的时候可能会出现数据倾斜，比如判断哪些是热点坑位，数据激增就会发生数据倾斜；

漏斗分析的时候，卖一个东西，它从哪些渠道来，有些渠道核心渠道，数据量很大，就会产生数据倾斜；

业务人员观看自己投的广告做的效果怎么样；首页来的90%；  
怎么发现的，看任务跑得慢，然后去看yarn 日志，具体task，的一个时间，或者处理数据量，

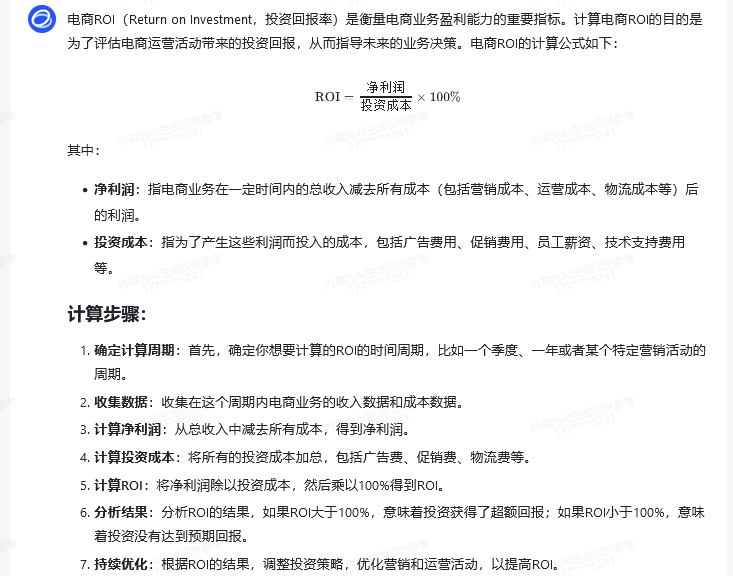
618 这种临时加资源，加大节点内存；核心任务监控，盯着；数据倾斜5个g,6个g,

漏斗倾斜脏数据，打散；

## 编写实习生学习文档，从新人视角介绍业务理解、数仓总揽、开发思路和相关建议等；

实习态度还可以，领导让我从新手视角从用户行为分析业务线，写学习文档，方便新的同学来后快速上手实习；公司平台？业务线要熟悉哪些业务架构？新手哪几天干哪些事；

## 11.常规问题与扩展问题

* 你当时是怎么上手工作的？
* 你说一下你们报表的开发流程？
* 说一下你们这个项目的数据源有哪些，埋点数据和业务系统电商相关商铺，商品，订单，用户信息
* 除了spark sql,有没有用过Scala等语言编写spark脚本  
  没用过，但会用，答流程
* Ods-dws层都做了什么操作
* Dqc你们是怎么配置的，怎么监控的
* 你们公司怎么处理缓慢变换维的
* 你们公司采用的什么模型居多
* 你们公司怎么做的数据治理，你了解多少，说一说
* 如果面试官问你这个组的业务是啥，一会用户画像，一会行为分析；
* 你们整个有多少张表？
* 说说你们公司的指标体系  
   我们公司指标主要分为，用户，流量，商品，订单，营销，这几类指标（电商那一套）；用户，基本属性啊，用户生命周期的指标啊，注册用户数，新用户，老用户，复购用户（订单来定位）；流量，就是渠道引流相关的指标，一些page 坑位，访问点击曝光，主要是uv,pv,；包括一些渠道roi指标；商品：商品一些基本信息指标，库存，上新，销量，倾销率啥的；订单：下单数量，成交额，成交客单价，成交转化率，关单数量，售后啊；GMV；营销：活动ROI,优惠券的一些使用指标；  
  ；