**华为帐号历史开发手册**

1. **数据字典**

**1.1表级**

英文名：dwd\_pty\_up\_ds\_his

中文名：华为帐号历史

表类型：拉链表

ETL调度频率：日表

**1.2字段级**

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名** | **字段类型** | **字段中文名** | **备注** |
| up\_id | VARCHAR(128) | 华为帐号编号 |  |
| start\_date | VARCHAR(8) | 开始日期 |  |
| up\_name | VARCHAR(256) | 华为账号名称 |  |
| language\_cd | VARCHAR(16) | 语言代码 |  |
| last\_name | VARCHAR(256) | 姓氏 |  |
| name | VARCHAR(256) | 名称 |  |
| gender\_cd | VARCHAR(16) | 性别代码 |  |
| birth\_date | VARCHAR(10) | 出生日期 |  |
| career | VARCHAR(256) | 职业 |  |
| service\_open\_situ | VARCHAR(256) | 业务开通情况 |  |
| invite\_up\_id | VARCHAR(128) | 邀请华为帐号编号 |  |
| invite\_user | VARCHAR(256) | 邀请用户 |  |
| register\_time | VARCHAR(30) | 注册时间 |  |
| account\_cancel\_time | VARCHAR(30) | 销户时间 |  |
| client\_type\_cd | VARCHAR(16) | 客户端类型代码 |  |
| register\_ip\_addr | VARCHAR(256) | 注册IP地址 |  |
| register\_src | VARCHAR(256) | 注册来源 |  |
| register\_channel\_id | VARCHAR(128) | 注册渠道编号 |  |
| up\_rank\_cd | VARCHAR(16) | 华为会员等级代码 |  |
| email\_login\_acct | VARCHAR(128) | 邮箱登录帐号 |  |
| mobile\_login\_acct | VARCHAR(128) | 手机登录帐号 |  |
| safe\_email | VARCHAR(128) | 安全邮箱 |  |
| safe\_mobile | VARCHAR(128) | 安全手机 |  |
| up\_service\_status\_cd | VARCHAR(16) | 华为帐号业务状态代码 |  |
| up\_type\_cd | VARCHAR(16) | 华为帐号类型代码 |  |
| etl\_time | VARCHAR(30) | ETL时间 |  |
|  |  |  |  |

1. **数据来源**

**2.1．来源于4个表:**

dwd\_up\_user\_info\_ds (用户信息表)

dwd\_up\_user\_device\_info\_ds (华为账号与IMEI绑定关系表)

dwd\_pty\_up\_login\_identity\_ds\_his (华为帐号登录身份历史)

dwd\_evt\_up\_oper\_log\_dm (华为帐号操作日志)

**2.2．其他依赖的表:**

dwd\_pty\_up\_addr\_ds\_his (华为帐号地址历史):用于获取华为帐号的注册IP地址

ods\_up\_t\_up\_memberright\_dm (华为帐号会员权益):用于获取华为帐号的华为会员等级代码

1. **数据逻辑**

**3.1．主体逻辑**

1，汇总4个来源表的所有帐号，获取用户信息

2，关联华为帐号地址历史，获取注册IP地址

3，关联华为帐号会员权益，获取会员等级代码

4，关联华为帐号操作日志中有销户操作的用户，获取销户用户的销户时间

5，用当日和昨日的数据对比，生成拉链

**3.2．字段逻辑**

| 字段名 | 处理逻辑 |
| --- | --- |
| up\_id | 主键。SHA256加密 |
| start\_date | 拉链表逻辑字段，标识整条记录状态的开始时间 |
| up\_name | 来源于用户信息表 |
| language\_cd | 来源于用户信息表 |
| last\_name | 来源于用户信息表 |
| name | 来源于用户信息表 |
| gender\_cd | 来源于用户信息表 |
| birth\_date | 来源于用户信息表 |
| career | 来源于用户信息表 |
| service\_open\_situ | 来源于用户信息表 |
| invite\_up\_id | 来源于用户信息表 |
| invite\_user | 来源于用户信息表 |
| register\_time | 用户信息表：register\_time  华为账号与IMEI绑定关系表：first\_report\_time  华为帐号登录身份历史：first\_report\_time  华为帐号操作日志: 操作为注册的oper\_time |
| account\_cancel\_time | 从华为帐号操作日志取的操作为销户的oper\_time |
| client\_type\_cd | 来源于用户信息表 |
| register\_ip\_addr | 来源于华为帐号地址历史 |
| register\_src | 来源于用户信息表 |
| register\_channel\_id | 取用户信息表和华为帐号操作日志里注册时间最早的渠道 |
| up\_rank\_cd | 来源于华为帐号会员权益表 |
| email\_login\_acct | 来源于华为帐号登录身份历史，帐号类型为1 |
| mobile\_login\_acct | 来源于华为帐号登录身份历史，帐号类型为2 |
| safe\_email | 来源于华为帐号登录身份历史，帐号类型为5 |
| safe\_mobile | 来源于华为帐号登录身份历史，帐号类型为6 |
| up\_service\_status\_cd | 来源于用户信息表 |
| up\_type\_cd | 来源于华为帐号登录身份历史，取最在注册的帐号类型 |
| etl\_time | ETL处理时间 |
| end\_date | 拉链表逻辑字段同时也是分区字段，标识整条记录状态的结束时间。 |

**3.3．ETL逻辑**

参考拉链表处理逻辑。

先在dwd\_eqp\_device\_ds\_src表中累积每个业务每天的全量数据，分区字段pt\_d:天分区 pt\_service:业务分区。然后取出终端设备历史表的前一天数据(start\_date<=’$last\_date’ AND end\_date > ‘$last\_date’)与dwd\_eqp\_device\_ds\_src当天全量数据(pt\_d = ‘$date’)做比较(主键关联)，若有字段有变动。则将前一天的数据的end\_date修改为$date，并且新插入一条记录(start\_date=$date, end\_date为99991231)。若当天的数据有昨天的数据没有则属于新增记录，在终端设备历史表中新增一条记录(start\_date为$date, end\_date为99991231)。

1. **其他**

**4.1．销户**

针对华为帐号操作日志表里有销户操作的用户，会给该用户打上销户时间（account\_cancel\_time），并且把处理up\_id和主键之外的所有字段置为空（’’）

**4.2．拉链表几种简单的取数用法**

**4.2.1 取某天数据**

WHERE start\_date <=’$date’ AND end\_date>’$date’

**4.2.2 取当前最新的数据**

WHERE end\_date = ’99991231‘

**4.2.3取某天新增的记录(主键新增)**

**4.2.3.1方法1：**

正常方法取当前日期和前一天日期，两天的数据 做对比

**4.2.3.2方法2：**

SELECT

imei

FROM dwd\_pty\_up\_ds\_his

GROUP BY imei

HAVING MIN(start\_date) = '$date';

**4.2.3.3 方法3：**

SELECT

imei

FROM dwd\_pty\_up\_ds\_his t1

WHERE end\_date > '$date' AND start\_date = '$date'

AND NOT EXISTS (

SELECT 1 FROM dwd\_pty\_up\_ds\_his t2

WHERE end\_date = '$date' AND t1.imei = t2.imei

);