命名规范

版本变更历史：

|  |  |  |  |
| --- | --- | --- | --- |
| 时间 | 版本 | 作者/修改人 | 变更原因 |
| 2017-06-20 | V1.0 | 喻健民 | 初版 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

[1. 简介 4](#_Toc485829398)

[1.1目标 4](#_Toc485829399)

[1.2范围 4](#_Toc485829400)

[2. 命名规则 4](#_Toc485829401)

[2.1通用规则 4](#_Toc485829402)

[2.2 表分类和命名规则 4](#_Toc485829403)

[2.2.1数据缓冲层 4](#_Toc485829404)

[2.2.2基础数据层 5](#_Toc485829405)

[2.2.3数据应用层 6](#_Toc485829406)

[2.3 字段命名规则 7](#_Toc485829407)

[2.3.1 时间类 7](#_Toc485829408)

[2.3.2 统计类 7](#_Toc485829409)

[2.3.3 编码类 8](#_Toc485829410)

[2.3.4 判别类 9](#_Toc485829411)

[2.3.5衍生类 9](#_Toc485829412)

[2.3.6 其他类 9](#_Toc485829413)

[2.4 任务名，Kylin模型名，Cube名 10](#_Toc485829414)

[2.4.1任务名称 10](#_Toc485829415)

[2.4.2 Kylin模型名称 10](#_Toc485829416)

[2.4.3 Cube名称 10](#_Toc485829417)

# 简介

在数据仓库建设以及使用过程中由于参与方众多，不同的命名方式给大家带来理解的不便，甚至给使用者带来困惑，因此有必要对命名做统一的规范。

## 1.1目标

* 规范表分类、表名、字段名、任务名
* 提高表名、字段名、任务名的辨识度

## 1.2范围

适用于对数据仓库、元数据管理平台表、字段、任务的命名。

# 命名规则

## 2.1通用规则

* 命名采用英文单词小写和下划线来命名。

## 2.2 表分类和命名规则

### 2.2.1数据缓冲层

STG层，表结构跟生产数据源保持一致，作为数据清洗的缓冲。命名规则是：

数据层名称\_源库名\_源表名\_抽取频率和方式

其中：

数据层名称为：STG

抽取频率为： MM/H/D/W/M 【分钟/小时/天/周/月】

抽取方式： I/A 【增量/全量】

例子：stg\_cust\_customer\_label\_da 表示来自客源库【cust】的客源标签表【customer\_label】,抽取频率是按天抽取【d】，抽取方式是全量抽取【a】

### 2.2.2基础数据层

基础数据层包括：

* ODS层

操作数据存储层，结构贴近STG，对敏感字段（能通过该字段定位到客户或者房源的字段）脱敏，空值处理等，详见ODS处理规范。命名规则如下：

数据层名称\_源库名\_源表名\_更新频率和方式

中文名称：源库名-源表名

中文描述：描述这个表是在哪个业务场景中产生的。

其中：

数据层名称为：ODS

抽取频率为： MM/H/D/W/M 【分钟/小时/天/周/月】

抽取方式： I/A 【增量/全量】

例子： ods\_house\_sh\_housedel\_role\_da 表示来自客源库【house】的房源角色表【sh\_housedel\_role】,抽取频率是按天抽取【d】，抽取方式是全量抽取【a】

* DW层

数据仓库层，保存按照主题划分的、集成的、反应历史变化的数据，详见DW层处理规范。命名规则如下：

数据层名称\_主题域\_表内容描述[\_宽表标识]\_更新频率和方式

* 中文名称：表内容的描述，【对实体的动作】+【实体名】不多于20个文字，比如【预约】【带看】表

中文描述：描述这个表是在哪个业务场景中产生的。

其中：

数据层名称为：DW

表描述： 用名词描述表的内容，不超过3个单词，如有多个单词中间用下划线分开

宽表标识： 仅适用于宽表，内容为all\_info

抽取频率为： MM/H/D/W/M 【分钟/小时/天/周/月】

抽取方式： I/A 【增量/全量】

主题域：

|  |  |
| --- | --- |
| 主题【中文名】 | 主题【英文缩写】 |
| 潜客 | Cust |
| 客源 | Cdel |
| 房屋 | House |
| 房源 | Hdel |
| 事件 | Evt |
| 协议 | Agr |
| 交易 | Tra |
| 财务 | Fin |
| 组织机构 | Org |
| 日志 | Log |

例子： dw\_hdel\_hdel\_da 表示来自房源主题域【hdel】的房源委托表【hdel】,抽取频率是按天抽取【d】，抽取方式是全量抽取【a】

### 2.2.3数据应用层

应用层包括以下几个逻辑层：

* OLAP
* DIM
* RPT
* API

由于应用层是面向业务方的，因此会在命名中体现业务方，命名规则如下：

数据层名称\_业务方编码\_用途\_指标标识\_更新频率和方式

中文名称：【业务方描述】+【用途】，不多于20个文字。

中文描述：描述这个表的使用场景。

其中：

数据层名称为：OLAP/DIM/RPT/API 【数据集市/维表/固定格式报表/对外数据输出接口】

用途： 不超过3个单词描述用途，如有多个单词中间用下划线分开

指标标识： indicator，仅用于后面直接产生指标的表

更新频率为： MM/H/D/W/M 【分钟/小时/天/周/月】

更新方式： I/A 【增量/全量】

业务方编码为：

|  |  |
| --- | --- |
| 业务方名称 | 业务方编码 |
| 市场 | MKT |
| 财务 | FIN |
| 业务运营 | BOC |
| 策略 | STRA |
| 租赁 | RENT |
| 二手房 | SHH |
| 新房 | NH |
| 产品 | PROD |
| 人力资源 | HR |
| 品控 | QC |
| 研究院 | RI |
| 房产生态 | ECOSYS |
| 金融 | FFIN |
| 内容 | MEDIA |
| 总部 | HQ |
| 楼盘中心 | REDS |
| 平台产品运营中心 | POC |
| Center of Excellence | BOE |
| Business Partner | BP |
| 端口支持部 | PORT |

## 2.3 字段命名规则

### 2.3.1 时间类

* 日期

日期修饰语\_date

其中：

日期修饰语： 代表该日期的业务含义，不超过两个单词，比如：contract\_sign

* 时间

时间修饰语\_time

其中：

时间修饰语： 代表该时间的业务含义，不超过两个单词，比如：contract\_draft

### 2.3.2 统计类

* 数量

业务修饰语\_统计对象\_cnt

其中：

业务修饰语： 代表该统计字段的业务含义，不超过两个单词

统计对象： 一个单词描述统计对象，比如客【cust】,房【house】，点击【click】

例子：showing\_cust\_cnt【带看的客户数量】

* 次数

业务修饰语\_统计对象\_times

其中：

业务修饰语： 代表该统计字段的业务含义，不超过两个单词

统计对象： 一个单词描述统计对象，比如涨价【price\_rise】

例子：price\_rise\_times【涨价次数】

* 金额

业务修饰语\_统计对象\_amt

其中：

业务修饰语： 代表该统计字段的业务含义，不超过两个单词

统计对象： 一个单词描述统计对象，比如佣金【commission】

例子：rent\_commission\_amt【租赁佣金金额】

* PV

业务修饰语\_统计对象\_pv

其中：

业务修饰语： 代表该统计字段的业务含义，不超过两个单词

统计对象： 一个单词描述统计对象，主要指页面或者频道，比如首页【home】

例子：app\_home\_pv【APP首页PV】

* UV

业务修饰语\_统计对象\_uv

其中：

业务修饰语： 代表该统计字段的业务含义，不超过两个单词

统计对象： 一个单词描述统计对象，主要指页面或者频道，比如首页【detail】

例子：app\_shh\_detail\_uv【APP二手房详情页UV】

### 2.3.3 编码类

* 无业务含义的编码

生产系统中由表实体序列产生的自增id

业务修饰语\_id

其中：

业务修饰语： 主要为表实体描述，不超过两个单词，比如客源【custdel】

例子：custdel\_id 【客源ID】

* 业务操作流水号

生产系统中由某某业务操作产生的流水号，比如合同号，理房通缴费流水号

业务修饰语\_no

其中：

业务修饰语： 主要为业务操作描述，不超过两个单词，比如合同【contract】

例子：contract\_no 【合同流水号】

* 有业务含义的编码

主要指维度编码，比如业务类型编码

业务修饰语\_code

其中：

业务修饰语： 主要为维度描述，不超过两个单词，比如城市【city】，业务类型【biz\_type】

例子：city\_code 【城市编码】

* 编码对应的名称

业务修饰语\_name

其中：

业务修饰语： 主要为维度描述，不超过两个单词，比如城市【city】，业务类型【biz\_type】

例子：biz\_type\_name 【业务类型名称】

### 2.3.4 判别类

* 判别类

Is\_判别对象

其中：

判别对象： 需要判断是否的对象，不超过两个字段，比如是否有效【valid】

例子： is\_valid【是否有效】

### 2.3.5衍生类

原有字段名\_衍生标识

其中：

衍生标识： MD5【MD5加密】

例子：mobile\_md5【MD5加密移动电话号码】

### 2.3.6 其他类

业务修饰\_字段表述对象

其中：

业务修饰： 对描述对象的限定

字段表述对象： 描述对象

例子： home\_phone【家庭电话】

## 2.4 任务名，Kylin模型名，Cube名

### 2.4.1任务名称

任务类型标识\_对象名称

其中：

任务类型标识： Kylin/mail/[空] 【Kylin的任务名称/邮件任务/Hive表的更新任务此项为空】

对象名称： 表、Kylin模型、Cube、邮件主体

对Hive任务来讲一个任务只能对应一个目标表。

### 2.4.2 Kylin模型名称

对应OLAP表名\_MODEL\_Vx

其中：

Vx中x为版本号用阿拉伯数字命名如：1,2,3,4……

### 2.4.3 Cube名称

对应OLAP表名\_CUBE\_Vx

其中：

Vx中x为版本号用阿拉伯数字命名如：1,2,3,4……