**资源管家原子任务插件**

**开发文档**

版本:<0.1>

日期:2019-5-24

版权所有（C）中国移动通信集团广东有限公司，保留所有权利。

Copyright© China Mobile Group GuangDong Co., Ltd. All rights reserved

修订历史记录

| **日期** | **版本** | **说明** | **作者** |
| --- | --- | --- | --- |
| 2019-5-24 | 0.1 | 文档初始化 | 王云茜 |

目录

[修订历史记录 2](#_Toc9853399)

[1. 引言 3](#_Toc9853400)

[1.1. 文档目的 3](#_Toc9853401)

[2. 原子任务插件总体说明 3](#_Toc9853402)

[2.1. 原子任务插件的定义 3](#_Toc9853403)

[2.2. 原子任务类别 3](#_Toc9853404)

[3. 原子任务插件编写规范 4](#_Toc9853405)

[3.1. 原子任务SQL 4](#_Toc9853406)

[3.1.1. 通用规范 4](#_Toc9853407)

[3.1.2. 数据质量-比率 4](#_Toc9853408)

[3.1.3. 数据质量-问题数 6](#_Toc9853409)

[3.1.4. 数量统计 7](#_Toc9853410)

[3.2. 原子任务Python 7](#_Toc9853411)

[3.3. 原子任务JAVA 7](#_Toc9853412)

# 引言

## 文档目的

此文档供资源管家研发人员了解原子任务插件形态并规范原子任务插件开发。

数据库账号密码：

RMPDB

pdb\_2019\_#$

# 原子任务插件总体说明

## 原子任务插件的定义

原子任务插件是在资源管家系统中配置与运行的程序代码。为保证原子任务插件的正常运行，该插件的程序代码需要遵循固定的编写规范和输出方式，同时，根据原子任务类别不同，编写规范和输出方式也不同。目前，原子任务插件可支持的编程语言包括SQL、JAVA、Python。

## 原子任务类别

* 原子任务大类

分为：数据质量、数量统计、存储过程。

不同大类的原子任务解决不同的业务问题，数据质量反映不合规数据占比和问题数据；数量统计反映资源数量于不同维度的分类；存储过程用于复杂算法中中间表数据生成的任务。

* 原子任务小类

分为：比率、问题数。

比率小类的原子任务需提供三组代码：问题数据分子、总数据数量、问题数据明细。

问题数小类的原子任务需要提供两组代码：问题数据分子、问题数据明细。

* 目前已有的各类原子任务的支持语言及支撑场景列在下表中。

|  |  |  |  |
| --- | --- | --- | --- |
| **任务大类** | **任务小类** | **支持语言** | **支撑场景** |
| 数据质量 | 比率 | SQL | 数据质量百分比指标 |
| 问题数 | SQL | 数据质量已较高，只需关注问题数据数量 |
| 比率 | JAVA | 算法复杂，需要通过JAVA程序实现 |
| 比率 | Phython | 算法复杂，需要通过Phython语句实现 |
| 数量统计 |  | SQL | 资源数量分地市、多维度分组统计 |
| 存储过程 |  | SQL | 对于复杂算法需要建立中间表时使用 |

# 原子任务插件编写规范

## 原子任务SQL

### 通用规范

1. 原子任务的名称中不可包含下列任一字符：: \ / ? \* [ ]。
2. 每个SQL程序段结尾均不可带分号等任何符号。
3. SQL程序段中不可出现注释。
4. SQL标识符的关键字包括CITY、NUM。
5. SQL标识符可用来做枚举值翻译，标识符格式：“对象名（中文）$字段名（中文）”或者“数据库表名（英文）$表字段（英文）”。按此格式进行标识，输出明细、数量统计图表呈现等即可进行枚举KEY值的翻译。
6. 在调试SQL程序时，要留意输出明细的数据量级及完整运行的耗时。（因为大多数编译工具，如DataGrip，在显示结果时采取分页的形式，造成耗时的误评估）

### 数据质量-比率

关键字：CITY, NUM

**1、问题数据分子SQL格式**

该段SQL执行完输出数据是每个地市质量比率的分子数值。

SQL执行完成输出表的表头为CITY、NUM，顺序不分先后。

|  |  |
| --- | --- |
| CITY | NUM |
|  |  |
|  |  |

示例：

**SELECT** (**select** D.**LABEL\_CN  
 from** RESDB.DISTRICT D  
 **where** D.**CUID** = *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 26)) CITY,  
 *COUNT*(1) NUM  
**FROM** RESDB.ROOM T  
**WHERE** *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 20) =  
 **'DISTRICT-00001-00009'  
 AND** T.**LABEL\_CN IS NULL  
GROUP BY** *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 26)

**2、总数据数量SQL格式**

该段SQL执行完输出数据是每个地市质量比率的分母数值。

SQL执行完成输出表的表头为CITY、NUM，顺序不分先后。

|  |  |
| --- | --- |
| CITY | NUM |
|  |  |
|  |  |

示例：

**SELECT** (**select** D.**LABEL\_CN  
 from** RESDB.DISTRICT D  
 **where** D.**CUID** = *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 26)) CITY,  
 *COUNT*(1) NUM  
**FROM** RESDB.ROOM T  
**WHERE** *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 20) =  
 **'DISTRICT-00001-00009'  
GROUP BY** *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 26)

**3、问题数据明细SQL格式**

该段SQL执行完输出数据是全省的分子数据的明细。

SQL执行完成输出表的表头为CITY及其他需要呈现的字段，顺序不分先后。

|  |  |  |  |
| --- | --- | --- | --- |
| CITY | 字段1 | 字段2 | …… |
|  |  |  |  |
|  |  |  |  |

示例：

**SELECT** T.**CUID**,  
 T.**LABEL\_CN** 机房名称,  
 (**select** D.**LABEL\_CN  
 from** RESDB.DISTRICT D  
 **where** D.**CUID** = *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 26)) CITY,  
 (**SELECT** S.LABEL\_CN  
 **FROM** RESDB.SITE S  
 **WHERE** S.CUID = T.**RELATED\_SITE\_CUID**) 所属站点,  
 T.**STATUS** 机房$生命周期状态  
**FROM** RESDB.ROOM T  
**WHERE** *SUBSTR*(T.**RELATED\_SPACE\_CUID**, 0, 20) = **'DISTRICT-00001-00009'  
 AND** T.**LABEL\_CN IS NULL**

### 数据质量-问题数

**1、问题数据分子SQL格式**

该段SQL执行完输出数据是每个地市质量的问题数据数值。

SQL执行完成输出表的表头为CITY、NUM，顺序不分先后。

|  |  |
| --- | --- |
| CITY | NUM |
|  |  |
|  |  |

示例：

**select** (**select** D.**LABEL\_CN  
 from** RESDB.DISTRICT D  
 **where** D.**CUID** = *SUBSTR*(T.RELATED\_SPACE\_CUID, 0, 26)) CITY,  
 *count*(1) NUM  
**from** RESDB.SITE T  
**where** (T.LOCATION **is null or** T.LOCATION = **''**)  
**GROUP BY** *SUBSTR*(T.RELATED\_SPACE\_CUID, 0, 26)

**2、问题数据明细SQL格式**

该段SQL执行完输出数据是全省分子部分数据的明细；SQL执行完成输出表的表头为CITY及其他需要呈现的字段。

|  |  |  |  |
| --- | --- | --- | --- |
| CITY | 字段1 | 字段2 | …… |
|  |  |  |  |
|  |  |  |  |

示例：

**select** T.CUID **id**,  
 T.LABEL\_CN **name**,  
 T.STATUS status,  
 (**select** D.**LABEL\_CN  
 from** RESDB.DISTRICT D  
 **where** D.**CUID** = *SUBSTR*(T.RELATED\_SPACE\_CUID, 0, 26)) CITY  
**from** RESDB.SITE T  
**where** (b.LOCATION **is null or** b.LOCATION = **''**)

### 数量统计

数量统计只需要一段SQL程序，SQL执行完成输出表的表头为CITY、字段1（统计维度1）……字段n（统计维度n）、NUM，顺序不分先后。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CITY | 字段1（统计维度1） | …… | 字段n（统计维度n） | NUM |
|  |  |  |  |  |
|  |  |  |  |  |

统计维度建议不超过两个（包括地市）；如按地市（21个）、设备厂家（10个）、设备类别（6个）对传输设备进行数量分组统计，则统计表共有1260条分组数量，每个地市也有60个分组数量，数据展示的效果较差。若对于统计维度的枚举值数量较少的情况，也可以做多维度分组统计。

统计维度字段如果为枚举字段，在SQL中请使用枚举翻译的标识符格式进行标识，以便于统计图表的展示。

示例：

**SELECT** (**SELECT** district.**LABEL\_CN  
 FROM** RESDB.DISTRICT district  
 **WHERE** *SUBSTR*(trans.RELATED\_DISTRICT\_CUID, 0, 26) =  
 district.**CUID**) CITY,  
 vendor.LABEL\_CN 设备厂家,  
 trans.SIGNAL\_TYPE TRANS\_ELEMENT$SIGNAL\_TYPE,  
 *COUNT*(*\**) NUM  
**FROM** RESDB.TRANS\_ELEMENT trans  
 **JOIN** RESDB.DEVICE\_VENDOR vendor **ON** trans.RELATED\_VENDOR\_CUID = vendor.CUID  
**GROUP BY** *SUBSTR*(trans.RELATED\_DISTRICT\_CUID, 0, 26),  
 vendor.LABEL\_CN,  
 trans.SIGNAL\_TYPE

## 原子任务Python

编写规范待补充。

## 原子任务JAVA

编写规范待补充。