## **简介**

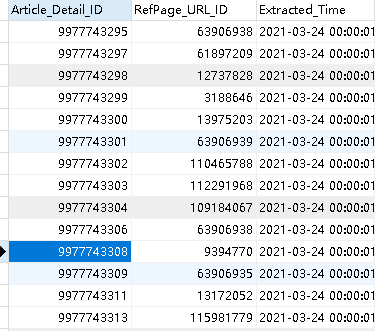
本程序是对KWM的网站进行调度调整的功能，针对listpage\_url很多的分成多个website的情况，例如 搜狐号API列表2，url非常大的时候，全部同样的调度会导致某些热门的url没办法很及时采集。本程序根据url所采集的数据量进行调整，把多个website\_no的情况分成5个website\_no，分别对应A、B、C、D、E等级，每个等级占比分别为10%、15%、20%、25%、30%。

## **流程分析**

#### **1.从主数据库116导出detail到118**

a\_output\_detail\_to\_datasource

导出最近7天的detail数据到118，结果表为listpage\_url\_by\_detail。



#### **2.从主数据库116导出listpage\_url到118**

b\_output\_listpage\_to\_datasource

#### 

#### **3.计算每个listpage\_url的数据量**

c\_update\_extracted\_data\_count

#### 

#### **4.对某个网站进行等级划分**

d\_rank\_listpage\_url

需要输入要调整的网站编号列表和调整后每个等级对应的网站编号，例如： 搜狐号： 目前在KWM里面所建的网站编号 website\_no\_input = ('S15347', 'S18604', 'S18603', 'S17969', 'S18309', 'S17968', 'S15326', 'S17967') 调整后A、B、C、D、E五个等级的编号 website\_no\_output = ('S17967', 'S15347', 'S15326', 'S17968', 'S17969')

#### 

#### **5.把最终接口更新回主数据库116**

e\_reset\_listpage\_url\_in\_center

## 

## **运行**

打开 run.py ，编辑 need\_init，website\_no\_input，website\_no\_output 三个参数， need\_init 是指是否需要初始化，即重新导detail和listpage\_url到118，一般一周运行一次即可。 website\_no\_input 是需要调整的网站编号列表，如 搜狐号API列表1 有8个一样的网站 website\_no\_output 是调整后的5个等级的网站编号列表。