

数据库插入使用说明

该版本是使用批量造数据然后一次性提交当天txt文件实现的

该脚本为单线程版本，即造完当天数据后再插入。后期可改完双进程，一个进程造数据，一个进程插入数据库，这样脚本执行时间会大大缩减

该脚本SQL为10.12之前上线版本SQL，新需求SQL变更后需要自行修改 234、245、255行代码。注意字符串拼接方式，可不填写字段使用 '+' '\t\t' 方式拼接

```
234 dataListSQL1 = '''+increaseID1() + '\t\t' + uid + '\t\t' + robot_flow + '\t\t' + caller + '\t\t' + create_phone() + '\t\t' + call_type + '\t\t' + formatTime + '\t\t' + \
235 formatTime + '\t\t' + duration + '\t\t' + call_status() + '\t\t' + talk_times + '\t\t' + labelValues + '\t\t' + \
236 slots_info + '\t\t' + get_gender() + '\t\t' + province + '\t\t' + flow_path + '\t\t' + department + '\t\t' + \
237 disconnect_reason() + '\t\t' + current_count + '\t\t' + "0" + '\t\t' + get_success() + '\t\t' + cmos_modify_time + ''
238 #print(dataListSQL1)

245 dataListSQL2 = '''+increaseID2() + '\t\t' + uid + '\t\t' + alternate() + '\t\t' + interaction1 + '\t\t' + formatTime + '\t\t' + labelValues + '\t\t' + uid + '\t\t' + musicURL + '\t\t' + \
246 '1.6分_取删_询问不再易原因' + '\t\t' + cmos_modify_time + ''
247 #print(dataListSQL2)
248 writeTxt(formatTime1, 'dataListSQL2', dataListSQL2)

253 for value in labelValues:
254     # dataListSQL3 = (uid, value, '2022-06-21 11:23:23.467')
255     dataListSQL3 = '''+ uid + '\t\t' + value + '\t\t' + increaseID3() + '\t\t' + '2022-06-21 11:23:23.467' + ''
256     #print(dataListSQL3)
257     writeTxt(formatTime1, 'dataListSQL3', dataListSQL3)
258
```

数据库权限查看与配置

查看 secure_file_priv 路径

```
SHOW VARIABLES LIKE "secure_file_priv";
```

192.168.128.48 x86_27_all_bj 运行 停止 解释

1 SHOW VARIABLES LIKE "secure_file_priv";

信息	结果 1	剖析	状态
	Variable_name	Value	
	secure_file_priv	/var/lib/mysql-files/	

可以看出，mysql 安全文件路径为 /var/lib/mysql-files/data。

如果脚本中的 baseFile 字段路径不为该安全路径，请更改 196 行

```
191
192 if __name__ == '__main__':
193     # 程序使用CPU开始时间
194     start=datetime.now()
195
196     baseFile = '/var/lib/mysql-files/data'
197     #baseFile = r"C:\Users\William\Desktop\data"
198
```

将 local_infile 改为 ON

在mysql 8.0.22 运行load data local 从本地文本导入数据时, 报错: "ERROR 3948(42000): Loading local data is disabled - this must be enabled on both the client and server sides".[\[https://blog.csdn.net/young_kp/article/details/109523153\]](https://blog.csdn.net/young_kp/article/details/109523153)

首先, 检查一个全局系统变量 'local_infile' 的状态:

```
show global variables like 'local_infile';
```

192.168.128.48 x86_27_all_bj 运行 停止 解释

```
1 -- SHOW VARIABLES LIKE "secure_file_priv";
2 show global variables like 'local_infile';
```

信息	结果 1	剖析	状态
	Variable_name	Value	
	local_infile	ON	

如果不是ON, 则需要执行以下代码

```
set global local_infile=1;
```

保存 查询创建工具 美化 SQL () 代码段 文本 导出结果

192.168.128.48 x86_27_all_bj 运行 停止 解释

```
1 -- SHOW VARIABLES LIKE "secure_file_priv";
2 set global local_infile=1;
3 show global variables like 'local_infile';
4
```

信息	结果 1	剖析	状态
	Variable_name	Value	
	local_infile	ON	

注: 每次重启 mysql 后需要重新设置, 包括最大连接数

好像可以不用设置

在 docker 配置的 mysql 文件中 /docker/mysql8/conf/my.cnf 最后写入该代码 local_infile = 1

```
!includedir /etc/mysql/conf.d/
max_connections=1000
max_allowed_packet=10g
init-connect="SET collation_connection=utf8mb4_0900_ai_ci"
init connect="SET NAMES utf8mb4"
local_infile = 1
```

脚本配置

设置执行时间

脚本第 207 行为设置从哪一天开始插入，即三张表的开始时间和结束时间都为该时间（有需求可以自行改代码）

```
206     dataListSQL3 = ''
207     start_date = datetime.strptime("2023-09-11 16:53:04", "%Y-%m-%d %H:%M:%S")
208     end_date = start_date + timedelta(days=7) # 30天后
```

208 行为要造数据的天数，**造数据天数不要超过造数的当月时间。**

例如：我要从9月25号造数据，207行设置为：start_date = datetime.strptime("2023-09-25 16:53:04", "%Y-%m-%d %H:%M:%S")

那 days 就要小于6天，end_date = start_date + timedelta(days=5)

这里208 行days 设置成5天，不超过本月。（因为262行、265行、268行为当月数据表，详见设置提交表）

设置每天插入数量

在 226 行可以设置每天插入数量，默认是 50w

```
225     # 每天生成50w数据
226     for i in range(0, 500000):
227         uid = str(create_uid())
228         # 随机拿5个call_label
229         labelValues = label()
```

设置提交表

```
260     txtPath = baseFile + '/' + formatTime1 + '/' + 'dataListSQL1' + '.txt'
261     #print(txtPath)
262     commitDB(txtPath, 'al_job_202309')
263     # 提交数据库
264     txtPath = baseFile + '/' + formatTime1 + '/' + 'dataListSQL2' + '.txt'
265     commitDB(txtPath, 'al_mat_202309')
266     # 提交数据库
267     txtPath = baseFile + '/' + formatTime1 + '/' + 'dataListSQL3' + '.txt'
268     commitDB(txtPath, 'al_tag_202309')
269
```

代码中提交表是固定的，如果你要插入10月的数据，就需要更改表名称了。

注意插入表的开始时间，以及插入表的结束时间，是否超过了提交表的名称时间

数据清理

导入后的原始数据及时清理，很占空间！！

默认原始数据路径：/var/lib/mysql-files/data

智能语音机器人

当前用户组：河北

AI 流程

统计分析

通话明细

menu_callDataOverview

短信

系统设置

● 通话明细

查询

重置

生成精简报表

生成全量报表

呼叫方向

全音

呼叫状态

全部

主叫号码

对话起止次数

机器人流程

全部

CALL_ID

通话起止时长

被叫号码

呼叫起止时间

2023-09-11 00:00:00 ~ 2023-09-11 23:59:59

标签

请选择

呼叫方向	被叫号码	主叫号码	呼叫时间	呼叫状态	意向标签	操作
呼出	13580077792	10086	2023-09-11 16:53:04	对方拒接	反对,肯定,SMS_测试,差评,支持	<div></div> <div></div>
呼出	17757052176	10086	2023-09-11 16:53:04	正常挂机	支持,完成调研未打分,电话,机主本人,反对	<div></div> <div></div>
呼出	15397094950	10086	2023-09-11 16:53:04	强制挂机	反对,举报,电话,差评,同意	<div></div> <div></div>
呼出	14952819501	10086	2023-09-11 16:53:04	识别错误挂断	赞同,宽带,肯定,支持,完成调研	<div></div> <div></div>

共50003条

20条/页

首页

上一页

下一页

尾页

Version 3.1.3