

PostgreSQL example

建立資料表

請將下列程式碼存成.sql並匯入PostgreSQL，或直接對PostgreSQL運行下列SQL語法

FB

```
-- -----  
-- Table structure for fb  
-- -----  
DROP TABLE IF EXISTS "public"."fb";  
CREATE TABLE "public"."fb" (  
  "_rid" text NOT NULL PRIMARY KEY,  
  "post_id" text ,  
  "post_time" timestamptz(6),  
  "source" text ,  
  "data_category" text,  
  "page_category" text,  
  "post_status" text,  
  "post_identity" text,  
  "from_id" text,  
  "from_name" text,  
  "page_id" text,  
  "page_name" text,  
  "body" text,  
  "title" text,  
  "caption" text,  
  "description" text,  
  "hashtag" jsonb,  
  "url" text,  
  "related_link" text,  
  "image_links" jsonb,  
  "image_alts" jsonb,  
  "attachment_type" text,  
  "attachment_url" text,  
  "post_tags" jsonb,  
  "comments" jsonb,  
  "image_md5" text,  
  "cs" numeric,  
  "cc" numeric,  
  "cl" numeric,  
  "ca" numeric,  
  "cd" numeric,  
  "ch" numeric,  
  "co" numeric,  
  "cw" numeric,  
  "ct" numeric,  
  "cu" numeric,  
  "cp" numeric,  
  "interval" numeric,  
  "diff" numeric,  
  "diff_comment" numeric,  
  "diff_share" numeric,  
  "track_time" timestamptz(6)  
)  
;
```

PTT

```
-- -----  
-- Table structure for ptt  
-- -----  
DROP TABLE IF EXISTS "public"."ptt";  
CREATE TABLE "public"."ptt" (  
  "_rid" text NOT NULL PRIMARY KEY,  
  "source" text ,  
  "post_id" text ,  
  "post_time" timestamptz(6),  
  "from_board" text ,  
  "from_id" text ,  
  "from_name" text ,  
  "from_ip" text ,  
  "post_status" text ,  
  "url" text ,  
  "title" text ,  
  "body" text ,  
  "cr" int4,  
  "cl" int4,  
  "cn" int4,  
  "cb" int4,  
  "co" int4,  
  "cc" int4,  
  "comments" jsonb,  
  "author_replies" jsonb,  
  "image_links" jsonb,  
  "related_links" jsonb,  
  "edit_infos" jsonb,  
  "istrans" bool,  
  "trans_to_infos" jsonb,  
  "trans_from_infos" jsonb,  
  "trans_from_boards_infos" jsonb,  
  "comment_image_links" jsonb,  
  "track_time" timestamptz(6)  
)  
;
```

新聞媒體

```
-- -----  
-- Table structure for rss  
-- -----  
DROP TABLE IF EXISTS "public"."rss";  
CREATE TABLE "public"."rss" (  
  "_rid" text NOT NULL PRIMARY KEY ,  
  "author" text,  
  "url" text ,  
  "post_time" timestamptz(6),  
  "title" text,  
  "description" text,  
  "body" text,  
  "size" int8,  
  "keyword" text,  
  "image_links" text,  
  "lang" text,  
  "source" text,  
  "category" text,  
  "fetch_time" timestamptz(6)  
)  
;
```

匯入資料

以透過Python匯入FB資料集為例

1. 打開insert_fb_example.py
2. 修改資料集存放路徑
3. 修改欲輸入的資料表名稱

```
for root, dirs, files in os.walk(".\\example_data"): #修改資料集(.gz)存放路徑
    for file in files:
        if file.find(".gz") != -1:
            total = 0
            total=total+insert(os.path.join(root, file), 'fb') #修改欲輸入的資料表名稱
print("total:"+ str(total))
```