

(1) Table schema

```
CREATE TABLE "map*" (  
    "countID" INTEGER NOT NULL,  
    "parkID" TEXT NOT NULL,  
    "stid" TEXT NOT NULL,  
    "opid" TEXT NOT NULL,  
    "feature" TEXT,    # 是否有停车信息, 1 为有 ; 0 为无  
    "distance" NUMERIC,  
    "parkName" TEXT,  
    "stationName" TEXT,  
    "parkAddress" TEXT,  
    "stationAddress" TEXT,  
    "parkCompanyName" TEXT,  
    "stationConstruct" TEXT,  
    "lat_park" TEXT,  
    "lng_park" TEXT,  
    "lat_station" TEXT,  
    "lng_station" TEXT,  
    PRIMARY KEY("parkID","stid","opid")  
);
```

(2) Table info:

Table Name	Info
Map2	准确匹配
Map1	不确定匹配
Map0	错误匹配

(3) Mapping method

STEP 1: 对所有停车场寻找 threshold=300 米以内的充电站，结果记为 M1

STEP 2: 对 M1 中的每条记录，使用正则表达式将停车场、充电站地址分解为 **市、**区、**街、**路、**弄、**号等几个单位，连同名字，使用 KMP 方法进行字符串匹配。

STEP 3: 根据匹配结果，将 map 结果插入到数据库

$Sim1 = 0.1 * \text{区} + 0.4 * (\text{道} + \text{街}) + 0.1 * \text{弄} + 0.4 * \text{号}$

$Sim2 = 40 / \text{distance} + 0.6 * \text{Name}$

IF sim1 > 0.8 THEN insert into map2

ELSE IF sim2 > 0.8 THEN insert into map1

ELSE insert into map0

(4) 结果分析：

map0 中共有 376 条记录，人工筛选一遍后，共有 2 条不确定匹配数据，分别为 :parkID=zb31010800238 和 parkID = pt31010700042, 其它均是错误匹配。

map2 中共有 131 条数据，人工检查后，配准结果均准确，

map1 表经过人工筛选后，有 10 条确定配准数据（放到了 insert Query 中，运行该 sql 可以将这 10 条数据插入道 map2 中，暂未插入），map1 还剩余 41 条

不确认纪录（例如名字相同地址不同等）

（5）下一步：需要对 map1 中的 41 条记录进行处理，筛选出正确匹配