**停车数据处理流程**

**1合并原始数据（线下）**

每个停车场的原始数据在多个文件不同表单中，需先将数据以停车场为单位合并到同一个文件，同一张表单中。

|  |  |  |
| --- | --- | --- |
| 日期 | 停车场 | 停车记录（条） |
| 2019/8 – 2020/1 | 土星 | 522320 |
| 2019/8 – 2020/3 | 渝兴广场 | 878137 |
| 2019/8 – 2020/1 | 两江星界 | 22190 |

**2提取、导入、处理原始数据**

1）渝兴物业与高科物业使用了不同的停车场管理系统，分别提取如下字段

高科：车牌号、进场VIP类型、进场时间、出场时间

渝兴：修改车牌号、车牌类型、车主名称、进出时间、进出地点、车牌颜色

2）将数据存入数据库并剔除无效数据与临时车

渝兴：剔除无车牌数据、剔除临时车、剔除车牌不全的数据（tx：229284 | yxgc：199616）

高科：剔除临时车（21077）

|  |  |  |  |
| --- | --- | --- | --- |
| 停车场 | 临时车进出记录（条） | 无车牌记录（条） | 车牌不全（条） |
| 土星 | 292299 | 286 | 737 |
| 渝兴广场 | 678104 | 1474 | 332 |
| 两江星界 | 1113 | 0 | 0 |

~~3）将车牌中不包含中文的数据补全~~

|  |  |
| --- | --- |
| ~~停车场~~ | ~~数量（条）~~ |
| ~~土星~~ | ~~737条记录/136辆车~~ |
| ~~渝兴广场~~ | ~~332条记录/134辆车~~ |
| ~~两江星界~~ | ~~0~~ |

~~/\* 查找哪些车牌可以更新 \*/~~

~~CREATE TEMPORARY TABLE tmp\_1 (~~

~~SELECT DISTINCT~~

~~plateno,~~

~~substr(plateno, 2) AS `key`,~~

~~substr(plateno, 1, 1) AS `value`~~

~~FROM~~

~~tx\_org~~

~~WHERE~~

~~substr(plateno, 2) IN (~~

~~SELECT DISTINCT~~

~~(plateno)~~

~~FROM~~

~~tx\_org~~

~~WHERE~~

~~char\_length(plateno) = length(plateno)~~

~~)~~

~~);~~

**3原始数据整理**

土星与渝兴广场的车辆进出记录是分开的，因此需要将车辆进出合并为同一条记录。（通过进出地点判断记录是进场还是出场、将数据根据车牌号以及进出时间进行升序排列）。

剔除无进场时间或无出场时间的数据。

|  |  |
| --- | --- |
| 停车场 | 无进场或出场时间记录 |
| Ljxj | 0 |
| Tx | 2192 |
| Yxgc | 1344 |

**4合并整理后的数据**

将各个停车场整理后的数据合并到同一个文件表单中，其中字段包含plateno（车牌号）、owner（车主）、entertime（入场时间）、exittime（出场时间）、parking（停车场）。

CREATE TABLE ljxj (SELECT plateno,cartype,’’ as `owner`,entertime,exittime,’ljxj’ as parking FROM ljxj\_org WHERE cartype != ‘临时车’);

/\* 将三张表合并为一张 \*/

create table total (

select \* from

(

select plateno,`owner`,entertime,exittime,parking from ljxj

UNION ALL

select plateno,`owner`,entertime,exittime,parking from tx where (entertime is not null and exittime is not null)

union ALL

select plateno,`owner`,entertime,exittime,parking from yxgc where (entertime is not null and exittime is not null)

) a

);

|  |  |
| --- | --- |
| 停车场 | 非临时车数量 |
| tx | 1506 |
| yxgc | 1194 |
| ljxj | 292 |

**5将数据按照天为单位整理**

跨天出场的记录需要按天进行拆分；

剔除周末入场的数据。（周六 DAYOFWEEK(date) = 7、周日 DAYOFWEEK(date) = 1）

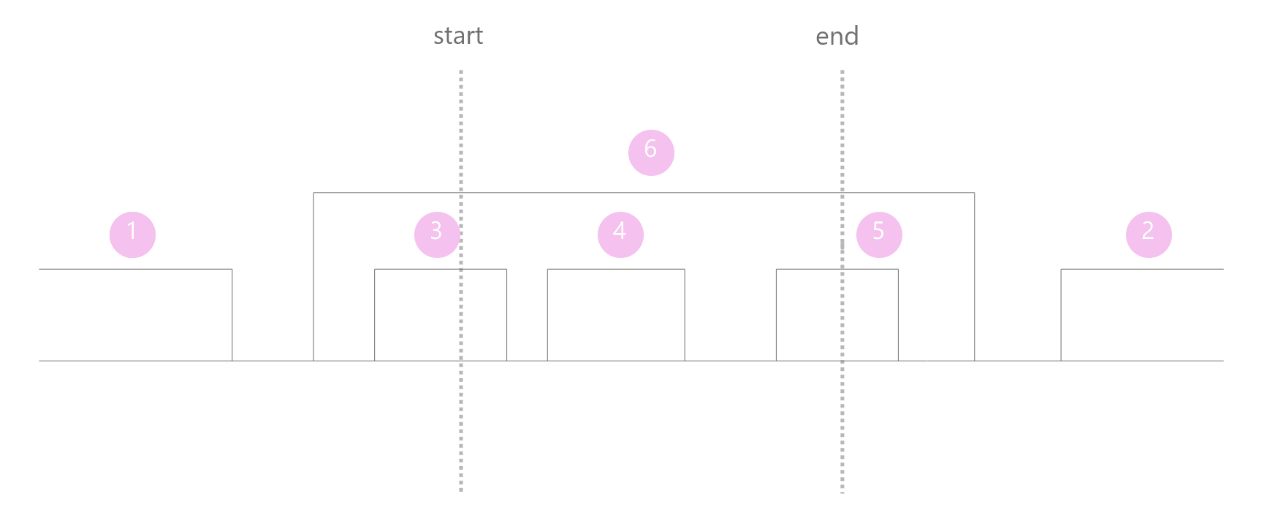
|  |  |
| --- | --- |
| 停车场 | 拆分并排除周末后数据量 |
| tx | 133987 |
| yxgc | 115021 |
| ljxj | 24810 |

**统计方案一**

**6工作时间停车时长统计**

输入起止日期，获取时段内的所有数据；

获取起止时间，并计算相应的停车时长；



1. Exit <= start **0**
2. Enter >= end **0**
3. Enter < start & start < exit < end **exit - star**t
4. Enter > start & exit < end **exit - enter**
5. Start < enter < end & exit > end **end - enter**
6. Enter < start & end < exit **end - start**

**7同一车辆一天中多次进出数据去重处理**

若其中一条记录的duration = end – start（符合情况6）则只保留这条记录；

除此之外，生成一条新纪录，duration为所有记录的duration叠加（1 + 2 + 3 + 4 + 50）。

**8车位空闲时间统计**

略

**统计方案二**

**9参数选择**

停车场（parking）：三选一

起止日期（start, end）：先选日期，再选时间段

//查询语句

select count(distinct plateno)

from total\_split

where parking = 'tx'

and DATE(entertime) = '2019-08-17'

and (TIME(entertime) <= '15:00:00' and TIME(exittime) >= '15:05:00');