ipsos KPI使用文档

1. 项目目标

客户黏度的营销管理

单车消费的营销提升

1. 数据检查(已实现自动化)

1.分割符检查

数据格式：

在原始数据中数据是英文逗号分割的，每个字段的数据是使用双引号包裹的

数据检查原因：

为了保证数据在进行数据格式转换的时候保持数据的字段数保持一致，使数据字段不会发生缺失现象

2.字段数检查

字段数量：

1. doss的原始数据一共是17个字段
2. customer的原始数据一共是48个字段
3. claim的原始数据一共是20个字段
4. part 的原始数据一共是27个字段
5. order的原始数据一共是37个字段

字段数量检查原因：

字段数量不足会造成数据缺失导致数据在计算出的数据会有偏大或者偏小或为空的现象。

3.字段对应顺序检查

字段对应顺序检查需要与五张表表头字段表进行比较

检查原因：

字段顺序不对应可能会造成数据错位，在进行数据关联的时候可能出现错误。

4.原始数据条数检查

检查原因：

需要与后面数据条数检查进行对比

1. 数据导入

注意事项：

原始数据导入路径需要注意检查

1. MR计算(分区：传入分区需要规范化)

1.分隔符转换

(1) 需要执行auto\_mr.sh format\_delimiter 对五张表的原始数据进行分隔符转换的操作

(2) 检查输出条数与原始数据条数做对比

2.检查新增信息(自动化)

从order表（品牌车系车系）和part表（零件）以及label\_order表(车型)、customer表(经销商\_asc\_code)对比现有的维度表，查找缺失的信息

3.检查辅助文件以及辅助表

每次需要检查的表有asc\_mapping 、mon\_mapping、mapping

文件有：mark\_doss,doss\_asc,mark,filter,engineoil,CHE

注意：

(1)如果是新加数据则一定套检查asc\_mapping和mon\_mapping中的日期是否已经达到数据所给的日期

(2)如果是除八大区、全国以及单店以外新加区域性计算需要更改asc\_mapping表

(3)如果是除年、季度以及月份以外的时间计算方式则需要更改mon\_mapping表

(4)如果零件需要进行版本控制的时候需要修改mapping表

(5)从order表中提取出所有的品牌车系车型与conf文件夹中的CHE文件进行对比

(6)从part里面机油与机滤编号与filter,engineoil文件进行对比

(7)从order表里面拿出品牌车系车型字段与mark文件进行对比

(8)从doss表中拿出品牌车系车型字段与mark\_doss文件进行对比

(9)如果有新增的经销商则需要与doss\_asc进行对比，有新增则需要我们自己修改

4.数据去重

功能是对4s店的数据进行去重，需要对doss、customer、order表进行去重操作

规则是含有a，b，c三个字段都相同的若干数据为重复

1. doss表需要根据VIN和DEALER\_NUMBER去重，在数据中是第八列和第十三列(下标从0开始)
2. customer表需要根据VIN,LNUMBER,CH\_CODE,OWNER\_CODE去重，在数据中是第一列、第二列、第十一列、地四十七(下标从0开始)
3. order表需要根据ORDER\_NUMBER,ASC\_CODE去重，在数据中是第四列和第二十八列(下标从0开始)
4. 打标签
5. 删除标签

1）给order表打 “删除” 或者 “” 标签，会在order表的最后一列加上标签  
注意跑程序的时候需要添加品牌到 CHE 文件

2）删除标签规则：

1. 如果4s店的code不存在则 状态为删除
2. 删除非雪佛兰品牌的车型
3. 将“品牌”列为类似于QT的工单删除
4. 只保留售后部分的工单：将“维修类别”为"延保”"空值”和出厂里程小于等于200的 “装潢”删除，如果维修类为 SQWX 也打删除标签

e.假单处理：同一个4s店工单“结算金额”满足当天有10张及以上重复金额，且金额大于0小于等于50元的工单，判定为

3）注意：order表已经在最后一列打入一列“删除”标签，字段数 +1,这一步结束后会在order中再加入两列，第一列是否首保，第二列是否索陪

4）执行流程

1.检查数据流程

2.打删除标签的流程

sh auto\_mr.sh delete\_useless <参数：时间>

1. order表标签(claimAndFirstMaintenance)  
    1) 首保与索赔

注意：order表已经在最后一列打入一列“删除”标签，字段数 +1,这一步结束后会在order中再加入两列，第一列是否首保，第二列是否索赔  
 需要关联claim表,关联条件是索赔单号与asc\_code  
 2) 维修类别1(maintType)  
 需要关联part表,关联条件是asc\_code,工单号，结算日期，vin，  
 注：打完标签后order表会有40个字段  
 3) other标签  
 标签种类：

出库日期,车龄月数,车龄分类,一级分类,二级分类,保内,保外,车龄分类1(others)  
 需要关联doss，customer，关联条件vin  
 (3) doss表标签  
 标签种类：一级分类,二级分类

(4) 执行流程

执行如下命令：

sh auto\_mr.sh label\_fm <参数：时间>

sh auto\_mr.sh label\_fix <参数：时间>

sh auto\_mr.sh label\_other <参数：时间>

1. 添加分区(传参规范化)

为计算之前的五张原始数据表添加分区，命令入下：

sh auto\_mr.sh add\_part <参数：时间>

1. 检查数据条数

检查打完标签以后的数量与原始数据是否在同一数量级上

1. 计算KPI

执行kpi的时候需要注意传入的参数，无特殊说明是按照以下的顺序执行程序以及参数

1. 全国月度 whole\_country everymonth
2. 单店月度asc everymonth
3. 区域月度8areas everymonth
4. 全国全年whole\_country year
5. 单店全年asc year
6. 区域全年8areas year
7. 全国季度whole\_country quarter
8. 单店季度 asc quarter
9. 区域季度 8areas quarter

注意：

1. 计算的时候需要先计算直接相加类和占比类，然后在计算非基盘和基盘类
2. 计算多次的时候直接相加类和占比类只计算一次。

3.在传时间参数时：如果是算全年数据传入的时间应当是整年，计算季度的时候传入的季度时间也应该是整季度的值