**店品销量band设计**

1. **背景**

目前基于分类维度统计分析dms偏差度，使用简单粗暴的方式：

SUM(dms)-SUM(sale)/SUM(sale)

这种累加求偏差度的方式，数据会有很大偏差，只能简单总体反映出dms与实际销量差值。而系统目前是基于店品维度进行销量预测的，那么需要基于店品维度减少dms与实际销量的偏差。基于现有数据，如何赛选出部分店品进行数据分析呢。

1. **Band 店品等级方案**
2. 随机抽样法：比如选取门店最近1个月门店数据，随机收取部分sku进行多维度分析(随机性，数据干扰大)
3. Top20：比如选取门店最近1个月门店数据，选取销量最高的前20%sku，进行多维度分析(数据倾斜严重，容易造成数据孤岛)
4. dms 偏差度 等级法：根据偏差度范围进行店品打标(目的明确，找出需要提升的店品重点分析)
5. **DMS偏差度等级法+ 实际销量大的店品**

|  |  |
| --- | --- |
| **偏差度：P 预测值：dms 实际销量：sale** | |
| band正向等级 | |
| 等级 | 说明 |
| A | (dms>dms中位数>=0 & sale=0 ) 或p> 100% |
| B | 50%<=p<100% |
| C | 30%< p< 50% |
| D | P < = 30% |
| band负向等级 | |
| 等级 | 说明 |
| A | (dms=0 & sale>销量中位数>=0) p< -100% |
| B | -100%<P<=-50% |
| C | -50%<p<-30% |
| D | -30%<=p<0% |
|  | |
| X | sales == 0 and dms == 0 或者 p=0 |

**四.店品band数据结构**

|  |  |
| --- | --- |
| 字段 | 说明 |
| mart | 商家id |
| Shop\_code | 门店id |
| sku | 店品id |
| band\_type | 0:正向 1:负向 |
| band\_level | 等级：A B C D |
| Weight\_p | 权重p |
| Weight\_q | 权重q |
| factor | 因子 |
| band\_ext | 扩展字段 |
| dt | 分区字段 |