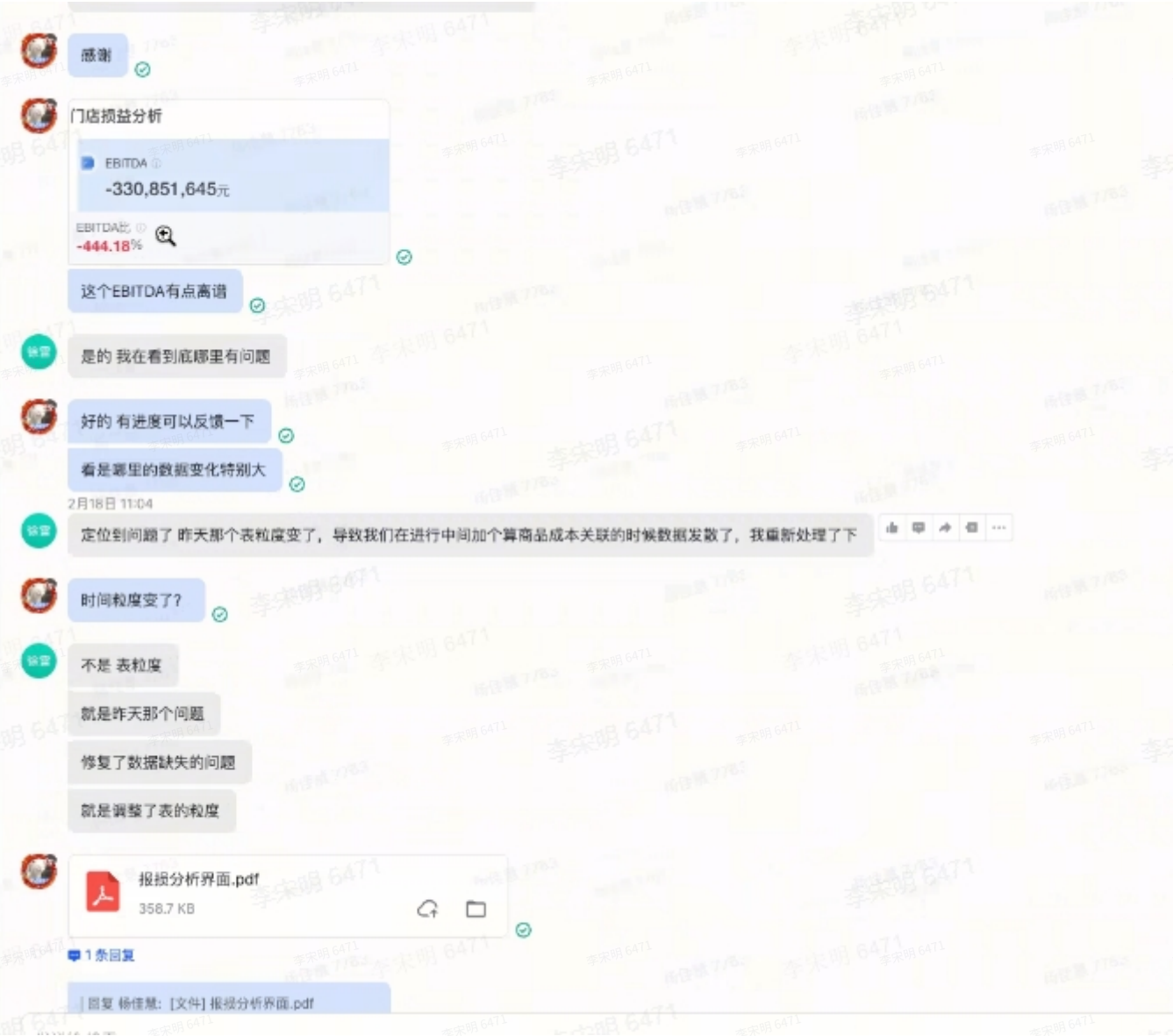


# 对茶姬驭事台（智能运营助手）重点数据(EBITDA等数据项)做数据异常检测

## 一、业务需求

### 1. 需求目标

- 对经营输出指标（如GMV、转化率等）等数据做异常检测，及时发现异常的数据。



### 2. 需求结果

像目前的驾驶舱数据检测一样，每天对全国，各个子公司，各个门店的经营输出指标做一次检测（等经营输出指标计算完后），如果有异常，将异常信息发送到经营输出指标告警群中。

## 驾驶舱数据新版

报警时间

2025-02-20 11:10:09

报警频率

1小时1次

干系人

处理状态

已忽略

## 1. 数据对比

周期	指标项	指标值	下限	上限	异常
当月	新增门店	11.00	42.60	181.40	1

明细数据

重新处理

## 二、需求分析

## 1. 异常原因分析:

## 国驭事台告警规则

和杨佳慧讨论了一下，经营输出指标出现异常大概有这几种原因：

- 数仓在指定的时间之前没有完成数据汇总
- 数仓中的数据与数据侧中的数据的条数不一致
- 数仓中的数据存在一些空值

由于这些原因的存在，可能造成经营输出指标产生异常数据。

## 2. 目前的解决办法:

指标名称	备注说明	是否有数据支持	报警时段	报警负责人	统计粒度 (区域)	统计粒度 (时间)	报警定义	取值字段	数据源表 (by 日)	数据源表	报表开发	数据下发SR
PBP		是	1个月	门店	月						@王松博	
EBITDA		是	1个月	门店	月				ads_store_piece_rate_salary_di		@王松博	
收入合计		是	1个月	门店	月				ads_franchise_material_cost_all_di		@王松博	
现金净利润率		是	1个月	门店	月				ads_inventory_produce_frmloss_di		@王松博	
GMV		是	1个月	门店	月				ads_store_channel_goods_summary_di		@王松博	
餐补 (使用)			1个月	门店	日				dwd_finance_us_detail_flat_avg_di		@王松博	
餐补 (新增)					日				ads_store_channel_goods_summary_di		@王松博	
人餐比		是	1个月	门店	月				ads_store_piece_rate_salary_di		@王松博	
人力成本		是	1个月	门店	月				ads_store_piece_rate_salary_di		@王松博	
人均薪资支出			1个月						ads_store_piece_rate_salary_di		@王松博	
薪资			1个月		日						@王松博	

由于经营输出指标都是从数仓中不同的表中计算得出，同时会将结果存入数仓中的对应的数据表中（经营输出指标和数据源存在不同的表中）。所以每天定时读取经营输出指标表，获取前30天的数据

(具体天数待定)，使用的驾驶舱数据检测算法，来检测当天的经营输出指标值是否存在异常，无论有没有异常，都将结论发送到"茶姬驭事台数据监控"报警群中。

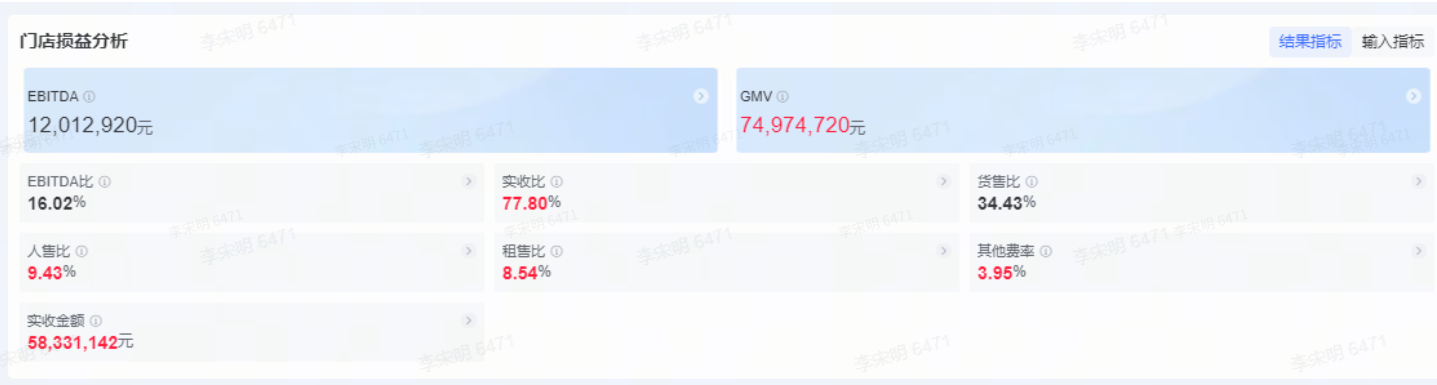
### 三、实现详解

#### 1. 创建"茶姬驭事台数据监控"报警群，完成人员添加

报警群的创建以及数据卡片的建立已经跑通

#### 2. 经营输出指标的获取

找 王崧骅 和 徐雷 确定经营输出指标存在哪些的表中，编写代码每天定时读取表中的数据



以下是 徐雷 给的目前他所用到的经营输出指标所在的表名，我先研究一下里面的内容

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_ads\\_prod.ads\\_franchise\\_material\\_cost\\_di/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_ads_prod.ads_franchise_material_cost_di/)

原材料成本 ads\_franchise\_material\_cost\_di

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_uat.ads\\_store\\_piece\\_rate\\_salary\\_di/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_uat.ads_store_piece_rate_salary_di/)

人力成本 ads\_store\_piece\_rate\_salary\_di

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_uat.ads\\_inventory\\_produce\\_frmloss\\_di/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_uat.ads_inventory_produce_frmloss_di/)

当日损耗 ads\_inventory\_produce\_frmloss\_di

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_ads\\_prod.ads\\_store\\_channel\\_goods\\_summary\\_di/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_ads_prod.ads_store_channel_goods_summary_di/)

GMV 智能运营\_渠道门店商品日汇总表 ads\_store\_channel\_goods\_summary\_di

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_dwd\\_prod.dwd\\_finance\\_ue\\_detail\\_flat\\_avg\\_di/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_dwd_prod.dwd_finance_ue_detail_flat_avg_di/)

租金成本 dwd\_finance\_ue\_detail\_flat\_avg\_di

[https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee\\_dwd\\_prod.dwd\\_doc\\_channel\\_store\\_list\\_df/](https://dataworks.data.aliyun.com/cn-shanghai/dmc/odps-table/odps.chagee_dwd_prod.dwd_doc_channel_store_list_df/)  
渠道门店清单表 chagee\_dwd\_prod.dwd\_doc\_channel\_store\_list\_df

📖 python操作MaxCompute指南

### 3. 普通公式检测以及异常检测算法的使用

3.1 使用现有的公式来检测EBITDA, EBITDA比， 实收比等检查  徐雷 他们计算的结果是否正确。

📖 UE分析树指标清单

3.2 沿用驾驶室数据异常检测算法， 启动定时任务每天9点遍历数仓中的经营输出指标， 和之前30天的数据做类比， 将结果写入"茶姬驭事台数据监控"报警群中。

### 4. 将检测结果写入"茶姬驭事台数据监控"报警群

写入数据的流程已经跑通

茶姬驭事台数据预警

报警时间	报警频率	干系人	处理状态
2025-02-20 17:52:12	1小时1次		正常

明细数据

茶姬驭事台数据预警

报警时间	报警频率	干系人	处理状态
2025-02-20 17:52:30	1小时1次	111	未处理

数据对比

周期	指标项	指标值	下限	上限
当月	新增门店	100	0	100
当月	新增门店	200	0	200
当月	新增门店	300	0	300
当月	新增门店	400	0	400
当月	新增门店	500	0	500

明细数据

UE分析树指标清单