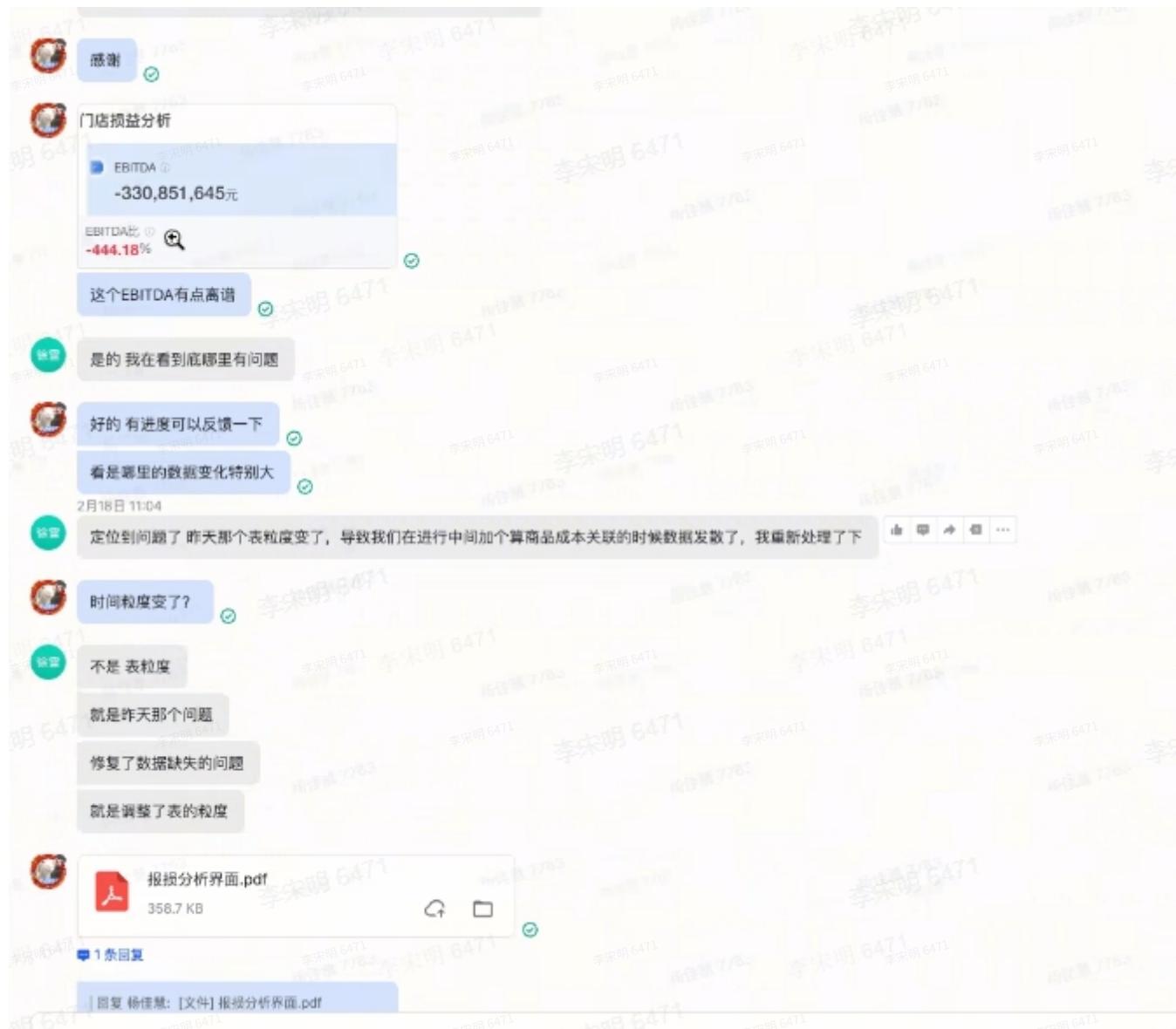


对茶姬驭事台（智能运营助手）重点数据（EBITDA等数据项）做数据异常检测

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 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:

④ 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：

1. 数仓在指定的时间之前没有完成数据汇总

2. 数仓中的数据与数据侧中的数据的条数不一致

3. 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	预警时效	指标负责人	统计粒度(区域)	统计粒度(时间)	预警定义	取值字段	数据源表(by日)	原开发	数据下发SR	数
PBP		是	1个月	门店	月				ads_store_piece_rate_salary_di	@王秘书		
FRTDA		是	1个月	门店	月				ads_franchise_material_cost_all_di	@王秘书		
投入产出		是	1个月	门店	月				ads_inventory_produce_frbloss_di	@王秘书		
现金流利润率		是	1个月	门店	月				ads_store_channel_goods_summary_di	@王秘书		
Gmv		是	1个月	门店	月				dwd_finance_ue_detail_flat_avg_di	@王秘书		
货补(使用)			1个月	门店	日				ads_store_channel_goods_summary_di	@王秘书		
货补(新增)					日							
人售比		是	1个月	门店	月				ads_store_piece_rate_salary_di	@王秘书		
人力成本		是	1个月	门店	月				ads_store_piece_rate_salary_di	@王达		
人均薪资支出			1个月						ads_store_piece_rate_salary_di	@王达		
薪资			1个月		日							

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



门店损益分析

EBITDA	12,012,920元	GMV	74,974,720元
EBITDA比	16.02%	实收比	77.80%
人售比	9.43%	租售比	8.54%
实收金额	58,331,142元	货售比	34.43%
		其他费率	3.95%

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/

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/

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/

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/

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/

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/
渠道门店清单表 chagee_dwd_prod.dwd_doc_channel_store_list_df

④ 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 徐雷 他们计算的结果是否正确。

⑤ UE分析树指标清单

3.2 沿用驾驶舱数据异常检测算法, 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