

成本分析-存储分析

成本分析下的存储分析分为四个功能模块，分别为：Hive表详情、推荐下线表、生命周期诊断和目录分析。

Hive表详情

在当前项目下，针对表总数、表总存储、净新增表数量、净新增表存储量和修改表数量、访问表数量均进行了数量统计，可以一目了然的对项目下的资产进行了解。

在**表总数**页签下，下面同时展示近30天的表趋势图和表所属主题域的Top5，点击**查看详情**，还可以查看详细的库表的访问次数。



在**表总存储**页签下，可以查看近30日表存储量的趋势图，右侧也会有表存储的Top5，点击**查看详情**，可以查看每张表的存储量。



在**净新增表数量**页签下，会展示净新增表趋势、新增表趋势和删除表趋势的近30日表数量的折线图，且对新增表所属主题域Top也进行了汇总，[点击查看详情](#)，可查看新增表所属的主题域和表分层的具体内容。



在**净新增表存储量**页签下，展示净新增表趋势、新增表趋势和删除表趋势的近30日表存储的折线图，对新增表存储Top也进行了汇总，[点击查看详情](#)，支持弹窗查看表归属的主题域和表分层以及存储量等信息。



在**修改表数量**页签下，展示近30日表数量和修改表次数的折线图，右侧也会展示表修改次数Top5，[点击查看详情](#)，可了解每张表的修改次数。



在**访问表数量**页签下，展示近30日表数量和表访问次数的折线图，右侧汇总展示表访问次数Top5，点击“**查看详情**”，可查看每张表的访问次数。



在Hive表详情功能模块下方，有Hive表详情列表，支持根据库名、表类型、负责人等快速筛选定位，列表展示表新增存储量、表生命周期、分区生命周期、分区数、新增文件数等详细内容，在列表中，支持批量选择，可修改表负责人(只有当前项目管理员/负责人和表负责人才可修改，其他角色功能键置灰)。

批量修改表负责人 [↓ csv txt](#)

库名: 全部 | 表类型: 全部 | 负责人: 全部 | 状态: 全部 | 主题域: 全部 | 表分层: 全部 | 表文件是否符合规范: 全部 | 重置

库名	表名	分区生命周期	分区数	存储成本/日(元)	直接计算成本/日(元)	总成本/日(元)	文件数	新增文件数
jinyx_hivedb	parent_tb08	永久	0	0.00	0.00	0.00	48	0
jinyx_hivedb	parent_tb02	永久	0	0.00	0.00	0.00	37	0
jinyx_hivedb	parent_tb03	永久	0	0.00	0.00	0.00	36	0
jinyx_hivedb	parent_tb04	永久	0	0.00	0.00	0.00	17	0
jinyx_hivedb	parent_tb05	永久	0	0.00	0.00	0.00	17	0
jinyx_hivedb	parent_tb06	永久	0	0.00	0.00	0.00	17	0

同时，在Hive表详情列表中，支持快速筛选出表文件是否符合规范的表，对于不符合规范的表，不支持在平台进行下线操作或设置生命周期。规范的目录文件需满足以下条件：1、分区路径处于表路径下；2、表路径未被其他表引用。

批量修改表负责人 [↓ csv txt](#)

库名: 全部 | 表类型: 全部 | 负责人: 全部 | 状态: 全部 | 主题域: 全部 | 表分层: 全部 | 表文件是否符合规范: 否 | 重置

库名	表名	分区生命周期	分区数	存储成本/日(元)	直接计算成本/日(元)	总成本/日(元)
dev	dwd_fir02_fir02...	永久	1	0.00	0.00	0.00
dev	aaa_testkudu1_c	永久	0	0.00	0.00	0.00
dev	ed_test248	永久	0	0.00	0.00	0.00

表路径被其他表引用，生命周期到期后不可删除目录文件，建议治理

推荐下线表

下线表趋势图展示最近30天确认下线表和推荐下线表数量变化趋势图。



推荐下线表占比展示最新一次统计的推荐下线表占全部表的比例。

推荐下线表占比



- 推荐下线表存储量及占比 99.68%
- 推荐下线表数量及占比 93.01%

在**推荐下线表**页面下方提供推荐下线列表，可在列表中查看推荐下线的表，并支持对推荐下线表进行下线、加入白名单等操作。

批量确认下线 批量加入白名单 下载 csv txt 白名单列表(0) 下线失败表清单(0)

负责人: 全部 库名: 全部 下线状态: 全部 推荐下线强度: 全部 重置 表名 请输入表名称

<input type="checkbox"/>	库名	表名	表类型	所属项目	状态	推荐下线强度	表负责人	表存储量 TB	文件数	近30日文件open次数	操作
<input type="checkbox"/>	demo	ods_sys_usertab...	内部表	poc	建议下线	强	poc	0.00	47057	0	确认下线 加入白名单
<input type="checkbox"/>	demo	t_ods_acct_accou...	内部表	poc	建议下线	强	poc	0.00	6	0	确认下线 加入白名单
<input type="checkbox"/>	demo	ods_acct_accoun...	内部表	poc	建议下线	强	poc	0.00	4	0	确认下线 加入白名单
<input type="checkbox"/>	demo	ods_acct_acc_tr...	内部表	poc	建议下线	强	poc	0.00	207	0	确认下线 加入白名单
<input type="checkbox"/>	demo	testvw	内部表	poc	建议下线	强	poc	0.00	1	0	确认下线 加入白名单
<input type="checkbox"/>	cdm	dws_trd_sample...	内部表	poc	建议下线	强	poc	0.00	4	0	确认下线 加入白名单
<input type="checkbox"/>	demo	stage_acct_accou...	内部表	poc	建议下线	强	poc	0.00	213	0	确认下线 加入白名单

参数信息	说明
状态	建议下线。建议下线遵循表下线规则：扫描周期内的文件open次数(仅读取，不含写入)=0，表访问次数=0，日均表引用数=0。
推荐下线强度	强推荐表示各项指标均符合下线规则，弱推荐则有可能该表存在写操作，建议均二次确认后，再执行下线。
确认下线	点击确认下线后，后台会对该表数据进行删除操作，默认会先移入灰色空间，7天之后正式清空。
加入白名单	若表虽满足推荐下线，但仍需要保存，可加入白名单，加入白名单的表不会再被推荐。

说明：

- 1.仅当前项目管理员、负责人和表负责人有操作权限；
- 2.若申请加入白名单的工单处于审批状态，则不允许操作；
- 3.加入白名单后，将不作为**推荐下线表**显示，可在**白名单列表**中移出；
- 4.申请加入白名单，默认需要当前用户的上级领导审批，如未对接OA系统则由项目负责人进行审批。

白名单资源

在表分析的推荐下线表模块，若表虽然属于推荐下线表，但数据比较重要，不做下线操作，可将表加入白名单，加入白名单的表将不再被推荐，加入白名单的表会在**白名单列表**中呈现，项目管理员/负责人和表负责人有权利将表移出白名单。

下线失败表清单

在推荐下线表中，若操作确认下线，对于下线失败的表，提供列表查看，展示下线失败的具体的表和失败原因等内容。

生命周期诊断

对于表设置生命周期，项目管理员/负责人对当前项目下所有表均有权限，其他角色，只有表负责人对自己的表有权限设置。

下线存储量趋势：当前项目下，因生命周期到期而下线的表和分区存储量的趋势图。



生命周期时间分布：当前项目下，表和分区的生命周期时间分布图，若均设置生命周期，则周期=Min(表生命周期，分区生命周期)。



生命周期诊断页面下方提供表生命周期列表，支持表的生命周期设置以及批量设置生命周期，对于想永久保存的表，可点击**永久保存**按钮进行保存。**永久保存**功能和白名单功能比较相似，加入到其中的表/分区的使用寿命将变成永久且不会进行资产健康的评分。

批量设置生命周期 批量永久保存 下载 csv txt 永久保存列表(0)

负责人: 全部 | 表类型: 全部 | 状态: 全部 | 是否分区表: 全部 | 主题域: 全部 | 表分层: 全部 | 表/分区是否设置生命周期: 全部

生命周期时间分布(天): 全部 | 重置

库名	表名	表类型	所属项目	状态	是否为分区表	表/分区是否设置生命周期	分区数	主题域	表分层	表负责人	操作
xf	ndi_performance...	内部表	poc	建议下线	否	否	0	-	-	-	设置生命周期 永久保存
xf	ods_jira_all_df	内部表	poc	状态正常	否	否	0	-	-	-	设置生命周期 永久保存
xf	ads_ui3	内部表	poc	建议下线	是	否	9	-	-	-	设置生命周期 永久保存
xf	ods_jira_update...	内部表	poc	状态正常	是	否	1035	-	-	-	设置生命周期 永久保存
xf	mid01_ods_jira_...	内部表	poc	建议下线	否	否	0	-	-	-	设置生命周期 永久保存

说明：

- 1.仅未设置生命周期的表/分区支持添加至永久保存列表；
- 2.加入永久保存列表后，资产健康分将不再扣除此类未设置生命周期表的分数，可在**永久保存列表**中移出；
- 3.申请加入白名单，默认需要当前用户的上级领导审批，如未对接OA系统则由项目负责人进行审批。

设置生命周期

修改表生命周期时需要注意：

1. 对于分区表，可修改表和分区生命周期，对于非分区表，仅可修改表生命周期；
2. 生命周期到期后，对于内部表，系统将自动删除表元数据和目录文件；
3. 对于外部表，系统将根据所选策略进行删除；
4. 外部表的目录文件，需满足以下条件才会删除：(1)分区路径处于表路径下；(2)表路径未被其他表引用。

目录分析

在目录分析功能模块，支持查看存储总量和日净增存储量相关内容。

在**存储总量**页签下，左侧展示最近30天，当前项目下数据存储量变化趋势，右侧展示当前项目下，最新一次统计的数据存储量排名前5的目录，点击[查看详情](#)，可查看目录具体的存储内容。



在**日净增存储量**(当前项目下，最新一次统计的净增数据存储量)页签下，左侧展示最近30天，当前项目下净增数据存储量变化趋势，右侧展示当前项目下，最新一次统计的近30天净增数据存储量排名前5的目录，点击[查看详情](#)，可查看目录具体的存储内容。

