## 模型设计中心:组织和刻画数据,规范化建设数仓

模型设计中心主要提供组织数据、刻画数据、规范约束和效果衡量的能力,用于实现数仓的规范化建设,同时指引数仓建设的方向,提供主题域、表分层、字典集以及表设计规则管理等功能。

连锁便利店的数据已从业务库集成至平台中,也通过数据测试中心进行了数据表的形态探查。但想要获取销量榜单 以及不同地域下的商品门店销售诊断报告等内容,首先需要通过模型设计中心产品,对主题域、表分层以及表设计 规则进行组织,对数据进行刻画和规范约束,实现数仓的规范化建模,基于数仓建设,以提供规范化业务数据。

对于当前案例想要达到的目标,在数仓规范化建设过程中,对于DIM—DWD—DWS—ADS各个分层所要创建的规范化表,如下所示(根据不同的案例场景,首先需要确定好各个分层下的表字段内容,结合模型设计中心的产品功能,完成表的规范化设计),如下图为数仓不同分层下的表字段设计:



在本案例中,结合模型设计中心的产品功能,最终可实现不同分层下,多张表的规范化设计,最终可构建如下表内容:

- dim层表——-dim\_product\_region\_cookbook(连锁便利店地域维表)
- dwd层表———dwd\_product\_sales\_cookbook\_month(连锁便利店商品供货表)
- dwd\_product\_sales\_cookbook2\_month(连锁便利店商品销售情况纪录表)
- dwd\_product\_sales\_cookbook3\_month(连锁便利店分布及供货情况纪录表)
- dwd\_product\_sales\_cookbook4\_month(连锁便利店用户评价体系纪录表)
- dws层表——-dws\_product\_sales\_cookbook1\_quarter(连锁便利店商品销量和供货情况宽表)

- dws\_product\_sales\_cookbook2\_quarter(连锁便利店不同地域下商品销量和用户评价体系宽表)
- ads层表——-ads\_product\_sales\_cookbook1(连锁便利店商品销量数据统计表)
- ads\_product\_sales\_cookbook2(连锁便利店不同地域下的商品诊断统计分析表)

## 具体实现

下面讲解如何通过模型设计的功能实现上述的规范化表设计目录。

点击左上角的logo,打开导航栏,点击数据治理模块的模型设计中心,进入产品界面(用户需为项目下**数仓开发**角色,才可以访问模型设计中心产品,对于申请开通的试用账号,管理员已默认将用户加入该角色,均可访问模型设计中心产品)。

<b>数据集成</b> ④ 数据传输	<b>数据服务及应用</b> ④ 流程协作与通知中心				图表
数据开发	问 问答中心	dws ads ods	TestLaver		
■ 实时计算					
<ul> <li>製 新加速</li> <li>製 新加速</li> <li>製 新加速</li> <li>製 新加速</li> </ul>	111 数据服务	未设置主题域表数量 ⑦	) 一级下游表数量 ⑦	昨日下游使用任务数 ②	
● 自助分析	数据治理 ◎ 模型设计中心	0	0	0	
任务运维	页 指标系统				
🙆 任务运维中心(旧)	◎ 数据管理				/•
🖗 任务运维中心 (新)	③ 数据质量中心				
	2 数据资产中心				
	<ol> <li>数据地图</li> </ol>	14 03-15 03-16 03-17 03-18 03-19 03-20 0	3-21 03-22 03-23 03-24 03-25 03-26 03-27 03-28 0	3-29 03-30 03-31 04-01 04-02 04-03 04-04 04-05 0	4-06 04-07 04-0
			■ dim - 已设置主题域表数量		

## 最初需求:

- 制作连锁便利店近30天内新品Top10销量榜单报表,并且可以清楚的了解近30天内Top10中每件商品的具体销售量;
- 制作不同地域下的商品门店销售诊断报告,可以根据地域、城市等进行筛选查找,了解不同地域下的销量、铺 货率、售罄率、达成率等内容。

维度:地域(东北区、华北区、华东区、华南区、西北区和西南区)。

主题域:商品域—销量。

在进行表规则设计之前,需要将所需主题域、分层等内容提前创建好,点击产品界面左下角的**数仓配置**(仅当前项目 组负责人或**数据团队管理者**角色有权限查看该功能按钮),进入数仓配置界面。

《 模型设计中心	*AP#1	回幸 キャ
😮 数仓建设概览		<b>國</b> 來   次  日
■ 主题域	dim map dwd dws ads ods TestLayer	
▲ 维度		
※ 度量	已设置主题城表数量     未设置主题城表数量 ⑦     一级下游表数量 ⑦     昨日下游使用任务数 ⑦	
▲ 基础字典	<b>1</b> 0 0 0	
■ 表设计工单管理 ^	压由纯热	
与我相关		-••
全部	0.75	
	0.5	
	0.25	
	0	04-07 04-08
◎ 数仓配置	■ dim - 已设置主题城表数量	

在**主题域配置**下,新增一级主题域,弹框中填写主题域名称、英文标识,指定域负责人和域审批人,域负责人有权 限在该一级域下创建二级,且对该域有编辑权限,域审批人用来对归属于该域下的工单进行审批。

《 模型设计中心	+ 新建主题域				
数仓配置	主题域名称	新増主题域	×		操作
● 分层配置	订单主题	* 主题城名称 商品城		2021-03-25 15:32:1 9	編輯 删除 上移 下移
<ul> <li>字典集配置</li> <li>最设计规则管理</li> </ul>	乘客主题	* 英文缩写 product		2021-03-25 15:52:0 9	编辑 删除 上移 下移
	保险	* 城负责人 1 ×		2021-03-26 13:38:0 2	编辑 删除 上移 下移
	法务主题	* 城亩批人		2021-04-06 09:55:5 5	编辑 删除 上移 下移
	未分组(系统)	<b>确定</b> 取消		-	-

在主题域列表中,可以对创建好的主题域进行编辑、删除、上移和下移操作,上下移动操作会影响主题域在模型设 计中心的展示顺序。

23	模型设计中心	+ 新建主题域		✓ 上移	✓ 上移成功						
		主题域名称	英文缩写	域责任人	域审批人	添加人	添加时间	操作			
*	分层配置	商品域	product	V0003			2021-04-09 16:17:4 1	编辑 删除 上移 下移			
	字典集配置 表设计规则管理	订单主题	order		C1 10		2021-03-25 15:32:1 9	编辑 删除 上移 下移			
		乘客主题	passager		and the set		2021-03-25 15:52:0 9	编辑 删除 上移 下移			
		保险	insurance		Jummgoul		2021-03-26 13:38:0 2	编辑 删除 上移 下移			
		法务主题	FW				2021-04-06 09:55:5 5	编辑 删除 上移 下傘			

在**分层配置**下,系统提供预置分层,包括dim、map、dwd、dws、ads和ods,系统预置的分层可以选择启用或者 不启用,也可以选择暴露或者不暴露(是否暴露并不影响模型设计中心的产品功能,仅针对数据地图产品,是否决定 将当前分层对所有用户开放的设置项),也可以开启自动审批,开启后,归属于该分层下的工单将自动通过。

模型设计中心	◇ 系统预置分层										
数仓配置		分层名称	英文标识	描述	启用状态	显示控制⑦	表抓取规则	允许建表的库	自动审批	操作	
<ul> <li>主题域配置</li> <li>分层配置</li> </ul>		维表	dim	维度数据	⊘启用	⊘暴露	-	-		编辑 禁用 不暴露	
<ul> <li>字典集配置</li> <li>事心计规则等理</li> </ul>	■ 申集配置 ■设计规则管理	桥接层	map	用于存放桥接维度 的关系表	⊘启用	⊘暴露		-		编辑 禁用 不暴露	
		明细层	dwd	从ods层经过ETL 得到的明细数据, 表示具体的事实, 主要由维度和度量 等构成	⊘启用	⊘暴露	-	-		编辑 禁用 不暴露	
		汇总层	dws	由明细数据经过汇 总得到的数据,主 要由统计维度和指 标构成	⊘启用	⊘暴露	-	-		编辑 禁用 不暴露	
_=		应用层	ads	由明细或汇总层加 工的出来面向报 表、服务、应用等 使用的数据	⊘启用	⊘暴露	-	-		编辑 禁用 不暴露	

若系统提供的预置分层不满足需求,也可以自定义分层。自定义分层时,可以设置抓取规则,对于没有设置分层的 表,若符合抓取规则,则可以抓取归属到自定义的分层中,比如自定义分层为adm,抓取规则为bdms.adm\_%,表 示bdms库下以adm开头的表均归属于adm层,创建完成之后,可以在对应的产品界面查看抓取到的表信息(**模型设** 

**计中心—主题域—未分组列表**,点击**立即抓取**,会呈现对应符合抓取规则的表分层下的表信息)。

	ルーベンテス	uws	四993-90006至22/2 总得到的数据,主 要由统计维度和指 标构成	UPU1	し登場	-	-		海祖 示□ 不暴露
<b>罪</b> 主题域配置	应用层	ads	由明细或汇总层加工的出来面白铝	⊘启用	⊘暴露	-	-		编辑 禁用
♦ 分层配置			上的山来回问报 表、服务、应用等 使用的粉提						小次的
🔄 字典集配置			PROSECUTION						
田 表设计规则管理	原数据层	ods	由业务系统同步到 数据仓库的原始数 据,一般不经过加 工	⊘启用	⊗不暴露	-	-		编辑 禁用 暴露
	∨ 自定义分层								
	+ 新建自定义	分层							
	分层名称	英文标识	描述	启用状态	显示控制②	表抓取规则	允许建表的库	自动审批	操作
	测试层	TestLayer	test	⊘启用	⊘暴露	命中规则: db1.at _ <sup>%</sup>	-		编辑 禁用 不暴露 删除
ē									

在字典集Tab下,可以新建字典集,比如对日期的后缀统一标准,day表示每日,hour表示每小时等,点击**+新建字 典集**,填写名称和描述等内容。字典集创建完成之后,在模型设计中心主页面的基础字典中可以针对该字典集添加 字典内容。

🕱 模型设计中心	+ 新建	<b>性字典集</b>				
数仓配置	#	字典集类型	新增字典集		×	( 操作
■ 主题域配置	1	简单字典集				编辑 删除 上移 下移
● 分层配置	2	复杂字曲集	字典集类型	简单字典集	$\vee$	编辑 删除 上线 下移
	3	复办之中集	* 字典集名称	日期后缀		
H AUTANEL		每九今曲 <b>佳</b>	描述	对日期的后缀进行统一定义和管理		
		at the				
					15/256	
			确定取测	Ä		
					_	

在**表设计规则管理**下,可以针对各个分层下表的设计规则进行设置,包括表的命名、表的分区生命周期、默认的存储格式等内容,若开启强制分区表,则归属于该分层下的表必须为分区表。

☆ 構用に合きます。										
教令配置	表分层	表名规则	表属性默认项配置	强制分区表						
<ul> <li>計 主题域配置</li> <li>⑦层配置</li> <li>字典集配置</li> <li>素设计规则管理</li> </ul>	维表 - dim	无设置	Impala同步: 是 表述型: 内部表 分区生命周期: 永久 表生命周期: 永久 存储格式: Parquet <mark>编辑</mark>							
	桥接层 - map	无设置	Impala同步: <b>是</b> 表送型:内部表 分区生命周期: 永久 表生命周期: 永久 存储格式: Parquet 編辑							
	明细层 - dwd	无设置	Impala同步: 是 表述型: 内部表 分区生命周期: 永久 表生命周期: 永久 存储指式: Parquet 编辑							
Ē	汇总层 - dws		Impala同共:							

点击表名规则下的**设置**,可以针对当前分层下的表命名规则进行设置,包括设置表的前缀和后缀,均可以选择固定 字符串、主题域缩写、二级主题域缩写、字典集或是正则表达式,对于常见的配置,系统也给出参考**配置建议**,可 以灵活调整各个分层下的表命名规则。在本示例中,设置ads应用层的表,前缀为固定字符串ads,后缀为二级主题 域的缩写,系统会根据用户选择的二级主题域自动填充表名后缀。

	编辑表名	规则 - 应用层	×	-
数仓配置	<ul> <li>1)表</li> </ul>	名拼接规则为: f1_f2_f3_f4_f5 表名主体_s1_s2_s3。		
計 主题域配置 ★ 分层配置	2) 表 3) 字 应用层 - ads 4) 常	名(文字)/马,对于正则表达式清按/马设置。 典集(文字持选择"简单字典集", 见的配置可参考 配置建议		
字典集配置	前綴			
■ 表设计规则管理		n 回定子付串 ∨ aos X + 添加		
	原数据层 - ods 后缀	后裔 控制逻辑		
		s1 二级主题缩写 ∨ ズロエは気がいった(HD) ストンビロの) 目別 X + 添加		
	测试层 - TestLaye 建表说明	用于展示给建表人员的说明性文字,比如未采用基础字典配置的后缀枚 举值说明、表名主体的命名建议等,从而指导建表人员设置合理的表名		
c				

以上,以项目组负责人(或数据团队管理者)的身份创建完了一级主题域、分层、字典集以及各个分层下的表设计规则,下面以数仓开发的角色进入模型设计中心主页面。

在基础字典下,针对已经创建的日期后缀字典集添加字典,填写字典的名称和描述,点击添加,即在当前选中的字 典集下快速创建字典,添加完成后,可以对字典进行编辑和删除等操作(当前只有项目组管理员或数据团队管理者角 色或项目负责人或项目管理员可以添加字典内容,其他角色可查看并使用已创建好的字典集)。

《 模型设计中心	字典集	快捷添加:	日期后缀 >	字典名称, 仅支持英文、数字		字典描述,不可超过256个字	
数 仓 建 设 概 じ	全部	添加					
■ 主题域	保险字典集	_					
▲ 维度	取消原因字典	批重删除			C	<b>、</b> 请输入字典名称或描述	С
<b>☆</b> 度量	字曲1		字典名称	◆ 描述	所属字典集	操作	
- 基础字典			day	每日	日期后缀	编辑 删除	
■ 表设计工单管理 ^	法劳子典		hour	每小时	日期后缀	编辑 删除	
与我相关	日期后缀		minute	每分钟	日期后缀	编辑 删除	
全部			month	每月	日期后缀	编辑删除	
			non	从不	日期后缀	编辑删除	
			quarter	每季	日期后缀	编辑删除	
◎ 数仓配置			week	每周	日期后缀	编辑 删除	
ē			year	每年	日期后缀	编辑 删除	

在**维度**下,针对本案例中的分析维度**地域**(东北区、华北区、华东区、华南区、西北区和西南区)进行创建。维度创建 完成后,可以对维度进行筛选、编辑和删除等,若维度被表所引用,也可以在维度列表中查询当前维度的引用详 情。

∃						新建维度			Х
28	模型设计中心	+ 新建维度				* 维度名称	地域		
	主题域	<b>负责人:</b> 请选择负	责人		(778) (4)	* 英文标识	region 如果有维表。	建议和维夷丰富设置为相同	
*	维度度量	所属王题域: 王部 维度名称	(5) 订单主题(1) 英文标识 ⑦	乘客主题 (2) 所属主题域	保险(1) 描述	* 所属主题域 ②	商品域	v	
8	基础字典	🛕 下单城市	city	订单主题	-	关联维表	维表属性 主维表	表 「请选择表,支持表名搜索 ↓	
	与我相关	▲ 渠道	channel	乘客主题	-		次维表	可多选,支持表名搜索	
	全部	🛕 票种维度	ticket_dim	法务主题	-	描述	请输入维	度描述	
		🛕 test	123	乘客主题	-			0/2	56
تا ا	数仓配置	🛕 管理机构	dept	保险	-	确定取消	i 确:	定后继续新建另一个	
C									

在**主题域**下,针对已经创建的一级主题域**商品域**创建二级主题域**销量**(只有一级主题域的域负责人才有权限在该域下 创建子主题,填写主题名称和英文标识,可以选择主题负责人为同父级或者自定义选择其他成员为子主题的负责 .

人。

2022/7/30

产品功能详述 - 模型设计中心:组织和刻画数据,规范化建设数仓 - 《有数数据开发及管理平台用户手册 V6.5.0》

ጰ 模型设计中心	主题域 (5)	<b>1</b>	部主题域		Q 请输入表				0
🚳 数仓建设概览	全部主题域 (系统)	新增子主题			х				
<b>罪</b> 主题域	商品域								
▲ 维度	▶ 订单主题	主题归属	商品域						
<b>次</b> 度量	<ul> <li>乘客主题</li> </ul>	* 主题名称	销量						
• 基础字典	▶ 保险	******	sales						
■ 表设计工单管理 ^	* 法方土题 主公组 (至広)	* 央文编与	Sures				主题归属	操作	
与我相关	不刀珇 (赤元)	主题负责人	● 同父级主题负责人 ○ 自定义			dwd	乘客主题/顺风车	查看核心属性	
全部						dwd	乘客主题/顺风车	查看核心属性	
						dim	法务主题	查看核心属性	
				取消	新增				

子主题域创建完成之后,可对主题域进行编辑、删除,也可以再创建二级主题域下的子主题域,最多可以创建至四级主题域。

ጰ 模型设计中心	主题域 (5)	✿ 销量 - sales	+ 添加关联表		Q、请输入表名称	Q、请输入表名称				
🚳 数仓建设概览	全部主题域 (系统)		主题介绍数位表							
₩ 主题域	▼ 商品域	主题介绍数包表								
▲ 维度	销量 + ∠ ×	表分层: 全部 dim	map dwd dws	ads ods Testl	ayer					
2 度量	▶ 订单主题	# 定 表 2 农	志伝書↓	表描述	主键 表分回	士駒旧尾	<b>揭作</b>			
■ 基础字典	▶ 乘客主题	·/ /+-///1///	AXAAA	733022	王隆    秋月按		19KT P			
■ 表设计工单管理 ^	▶ 保险			无数	据					

以上有关主题域、维度、字典等内容创建完成后,就可以规范化设计表了。在**表设计工单管理的与我相关**下,可以 创建建表工单,来规范化设计各个分层和主题域下的表。点击**+创建建表工单**,选择表分层、库和主题归属,根据在 **数仓配置**的表设计规则管理中设计的表名规则选择或填写完善表名称。

<b>爹</b> 模	型设计中心	← 创建建表工单								
<b>6</b> ž	收仓建设概览		1 表名 2 基:	出属性及字段						
₩ ±	主题域									
▲ 绡	住度	* 表分层	dwd - 明细层 V	查看分层和选库说明						
12 度	建									
<b>e</b> i	基础字典	* 库	dwd 🗸							
■表	長设计工单管理 ^	* 主题归属	商品域/ 销量 × ~ ~							
与	我相关									
全	注部	∨ 表名配置								
		您选择的表分层下设置了表设计规则,请按照规则设置表名及其他信息								
		表名拼接规则	dwd_一级主题域缩写_二级主题域缩写_表名主体							
愈 数	位配置	建表说明	无							
e III		下一步								

基本信息配置完后,在下一步可以设计表的字段、分区、业务以及基础属性等内容。下图为本示例设计的dwd层表的基本属性的部分字段,其中, area1\_name字段表示地域内容,和维度相关,在**标准化标签**一栏,可以选择维度,然后将该字段与之前创建好的"地域"维度相关联起来。创建完成后,点击提交,需要该表所归属的**商品域**的域审批人进行工单审批(若在数仓配置中,分层配置中开启了自动审批,则归属于该分层下的工单会自动通过)。

模型设计中心	← 创建建表工单										
🚳 数仓建设概览		<ul><li>✓ त्र</li></ul>	長名	2 基础属性及字段	]						
■ 主题域											
▲ 维度	✓ 字段配置										
26 度量	+ 添加字段										
■ 基础字典	# 字段名称 ⑦	类型	描述	主键 标	隹化标签	操作					
■ 表设计工单管理 ^	1 sku id	STRING V	商品id	维	度 度量 请先选择标准化标签	删除 上移 下移					
与我相关											
全部	2 area1_name	STRING ~	地域名称	维	度度量地域、	删除 上移 下移					
	3 city	STRING V	城市	维	渡 度量 请先选择标准化标签	删除 上移 下移					
송 **-스피포	4 area2_name	STRING V	地域别名	✔ 维	度 度量 请先选择标准化标签	删除 上移 下移					
	上————————————————————————————————————	3 伊存									

在**表设计工单管理**的**与我相关**页签下,可查看工单状态(在**数仓配置**页面中**主题域**配置下,指定域审批人,负责审批 归属于该域下的工单),支持快速查看表的DDL语句,也可以在**更多**下快速复制DDL语句。

	模型设计中心	未提交 0	待审批 0	被驳回 0 待我审批	比0 已通过 1	+ 创建	修改表工单 + 创建建表工单	⑦ 不合理命名表清单
8	数仓建设概览				L	]	表名称 > 清輸入工単編号、表名和	ф <b>С</b>
17	主题域	工单编号 ¢	工单类型	工单操作对象	对象配置模式	工单状态	工单信息	操作
A	维度	9	新建Hive表	dwd.dwd_product_sale	表单模式	<ul> <li>已通过</li> </ul>	创建人:) 审批摄	查看DDL 更多
X	度量			s_cookbook_month			提交备注: cookbook的dwd层表 审批意见: 自动审批通过	
4	基础字典	4						Þ
	表设计工单管理 ^							
[	与我相关							
	全部							

在模型设计中心的建表工单通过之后,系统会自动在当前项目下,对应的库下创建表,比如本示例中在dwd库下申 请创建的dwd\_product\_sales\_cookbook\_month表,工单通过之后,可以在**主题域**下,查看已经创建成功的表。

✗ 模型设计中心	主题域 (5)	・         ・         ・	Q、请输入表名称	0
<ul> <li>数仓建设概览</li> <li>主题域</li> </ul>	全部主题域(系统) 商品域	主题域介绍 数仓表		
▲ 维度	<ul> <li>订单主题</li> <li>乘客主题</li> </ul>	表分层: 全部 dim map dwd dws ads ods Te 二级主题: 全部 销量	stLayer	
<ul><li>✓ 皮里</li><li>▲ 基础字典</li></ul>		# 库、表名称 表负责人 表描述	主鍵 表分层	主题归属 操作
田 表设计工单管理 ^     与我相关	• 法务主题 未分组 (系统)	1 I dwd.dwd_product_sales_coo cookbook示例 kbook2_month wd层的表2	中d sku_id dwd	商品域/销量 查看核心属性 变更主题归属 取消当前域关联
全部		2 dwd.dwd_product_sales_coo kbook_month 表的分析	wd层 area2_name dwd store_id	商品域/销量 查看核心属性 变更主题归属 取消当前域关联
⑦ 数合配置		3   dwd.dwd_product_sales_coo cookbook示例 kbook3_month wd层的表3	Fd city_id dwd	商品域/销量 查看核心属性 变更主题归属 取消当前域关联
C		4 Ⅲ dwd.dwd_product_sales_coo cookbook示例 kbook4_month wd层的表4	中d sku_id dwd	商品域/销量 查看核心属性 变更主题归属 取消当前候关联

在本示例中,以上面同样的方式,设计并创建以下各个分层的表:

- dim层表——-dim\_product\_region\_cookbook;
- dwd层表———dwd\_product\_sales\_cookbook2\_month、dwd\_product\_sales\_cookbook3\_month、dwd \_product\_sales\_cookbook4\_month;
- dws层的表——-dws\_product\_sales\_cookbook1\_quarter和表 dws\_product\_sales\_cookbook2\_quarter;
- ads层的表——-ads\_product\_sales\_cookbook1、ads\_product\_sales\_cookbook2。

其中,在dwd层,可以将表字段与度量关联,在dws和ads层,对于表字段可以关联指标。本示例中在表ads\_product\_sales\_cookbook1的ord\_prod\_qty、ord\_prod\_amt等字段中,关联了在指标系统已经创建的指标 总进货数量、实际商品销售量 等(对于已经创建完成的表,若也想将字段与标签,即指标、度量和维度关联,可以在数据地图

的表详情页面,针对字段进行编辑,关联标签化标签,前提是当前表已归属于某个主题域和表分层下),如下图所示,创建表工单的第二步,选择指标,下拉框中选择对应的指标即可。

着型设计中心	← 创建建表工单										
	4	area2_name		STRING V				维度	指标	请先选择标准化标签	删除 上移 下移
■ 主题域	E	etoro num		CTDINIC				維密	t≌t∓	2年14-34-12+二2年11/+二大六	

以上,完成了在模型设计中心对于各个分层的规范性表设计,基于系统已经创建成功的各个分层下的表,接下来就可以在离线开发中心构建数仓任务,对数据进行ETL加工和处理,完成任务的开发和提交上线等流程。