



回复：签到+ERP



大数据专家认证

# 大数据专家认证培训之产品篇

--数据知识&数据集成开发

孙爱林  
sunailin@jd.com

[www.jd.com](http://www.jd.com)

Tuesday, September 08, 2015

- 产品功能概述
- 数据知识功能简介与使用
- 数据开发功能简介与使用

# 大数据平台首页

地址：[bdp.jd.com](http://bdp.jd.com)

大数据专家认证

产品使用



# 产品功能概述

大数据专家认证

## 业务导向

满足京东各体系的业务需求为首要职责，推动数据产品化进程，发挥数据价值。

## 技术驱动

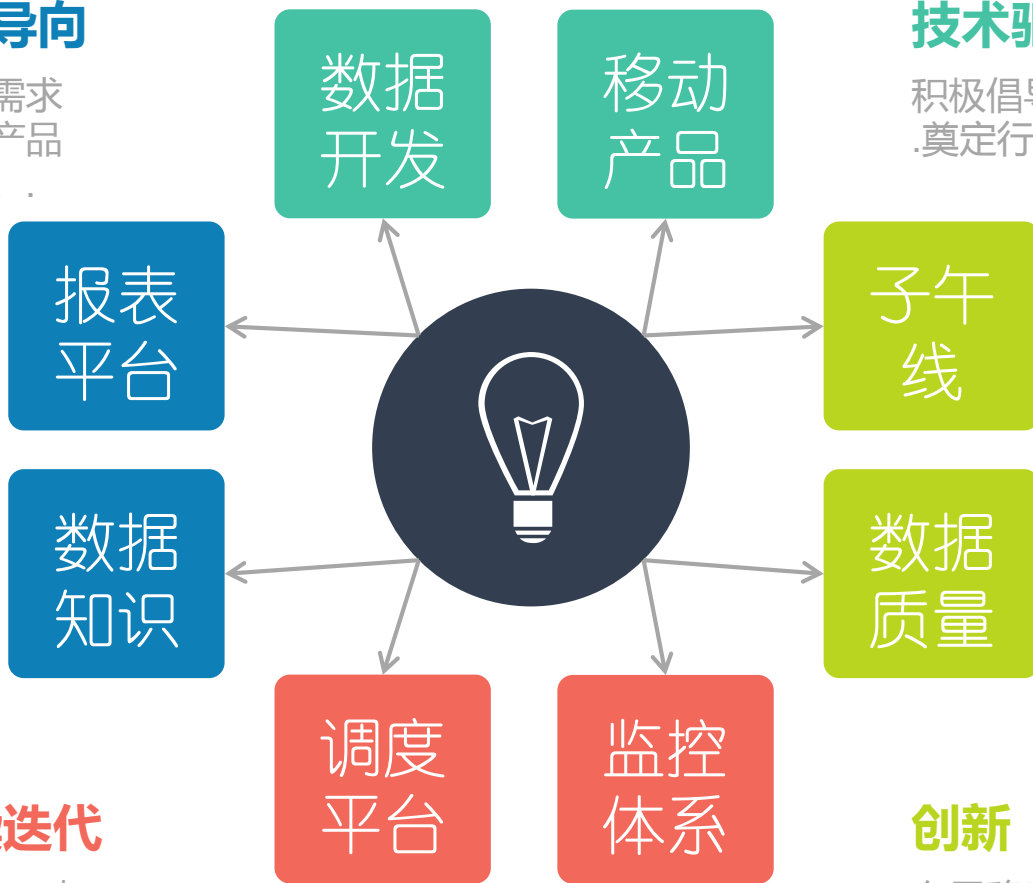
积极倡导自主研发技术应用，奠定行业领先地位。

## 持续迭代

把控产品功能和易用性，建立反馈及快速响应机制，推动产品优化工作。

## 创新

布局移动端数据产品，以创新性的思维鼓励团队开发新产品，创造新价值。



产品使用

# 产品功能概述

- **数据知识基本功能**

查询元数据（业务信息存储的数据库、数据表、字段信息）

沟通交流，询问、评论相关表及字段信息，问题反馈

收藏常用表

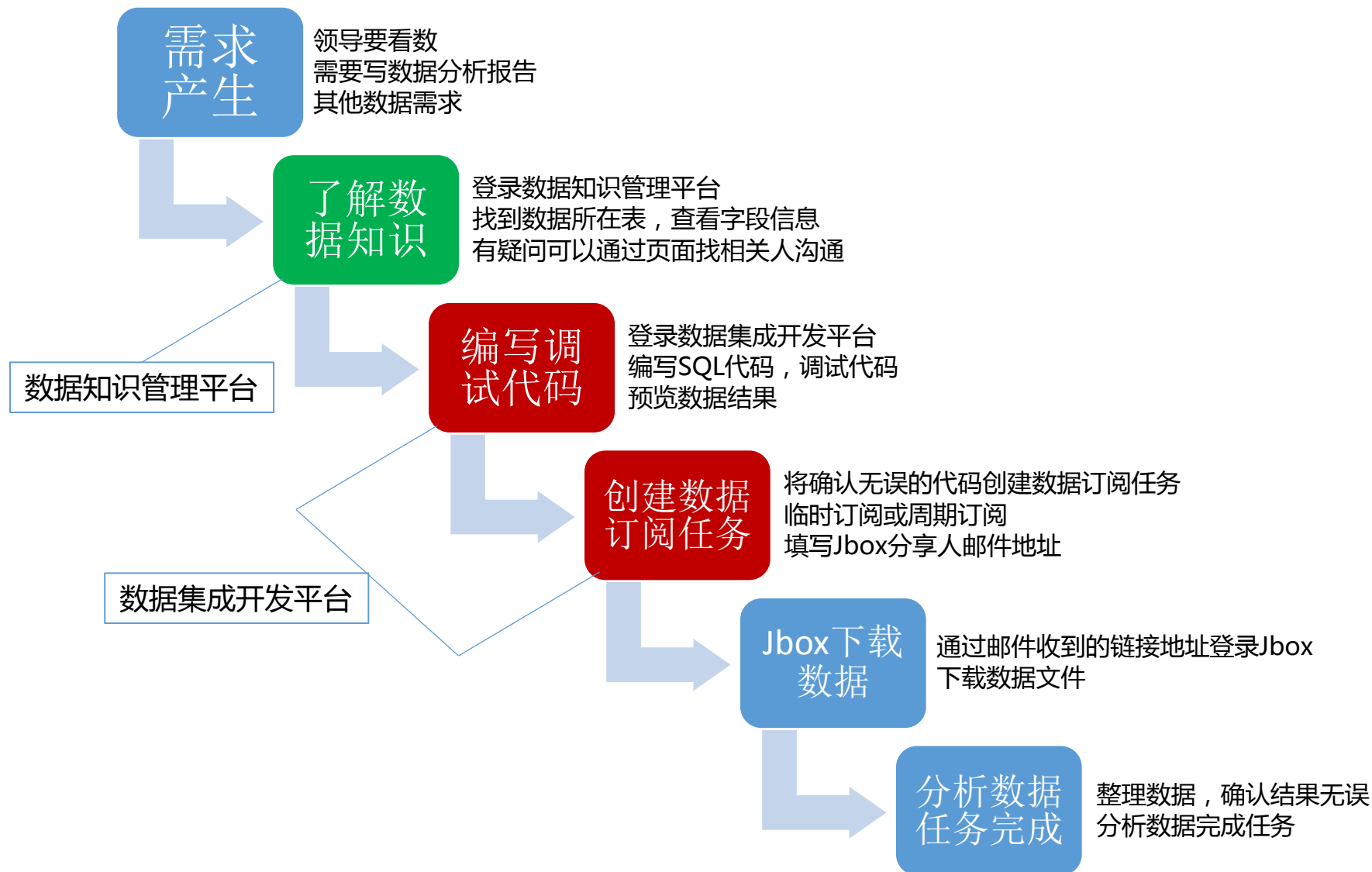
- **数据开发基本功能**

即时编写、调试SQL查询代码，预览数据结果

获取数据结果

配置任务定期自动获取数据

# 产品功能概述



# 数据知识管理平台功能简介与使用

## 菜单路径

京东大数据平台

平台首页

产品分类

数据知识

王威

数据知识

表

表中/英文名称、第一负责人姓名/ERP

搜索

展开筛选条件

今日更新表

6

今日新增表

1709

今日下线表

1603

我收藏的表(7)

我负责的表(4)

数据资料库

<div>gdm_m04_ord_sum</div> <div>订单宽表</div> <div>★★★★★</div> <div>使用中</div> <div>数据仓库</div> <div>取消收藏</div>	<div>FDM模型的使用说明</div> <div>2015-07-30 11:51:56</div>
<div>gdm_m04_ord_det_o2o</div> <div>O2O订单明细模型表</div> <div>★★★★★</div> <div>使用中</div> <div>数据仓库</div> <div>取消收藏</div>	<div>最近访问的表</div>
<div>fdm_jd_hotel_hotel_order_chain</div> <div>酒店订单主表</div> <div>★★★★★</div> <div></div> <div>数据仓库</div> <div>取消收藏</div>	<div>fdm_product_book_pr...</div> <div>营销集市</div>
<div>app_bda_mkt_user_value_q2_sum_...</div> <div>用户利润价值汇总表</div> <div>★★★★★</div> <div></div> <div>数据仓库</div> <div>取消收藏</div>	<div>fdm_product_self_prod...</div> <div>营销集市</div>
<div>gdm_m04_gp_sum</div> <div></div> <div>★★★★★</div> <div>将下线</div> <div>数据仓库</div> <div>取消收藏</div>	<div>fdm_pop_vender_vend...</div> <div>数据仓库</div>



# 数据知识管理平台功能简介与使用

数据知识

表

表中/英文名称、第一负责人姓名/ERP

搜索

收起筛选条件 ^

所在集市:	数据仓库	财务集市	营销集市	运营集市	建模集市	移动集市	审计	审计new集市				收起 ^	
	O2O集市	微信手Q	易迅	京东拍拍	风控	推荐	比价集市	广告	NSNG集市				
	POP集市	销量预测	网站移动研发集市		交易平台集市		hr集市	供应商集市	成都集市	职能集市			
	云平台集市												
集团公司:	京东商城	京东到家	京东智能	京东集团	金融集团	拍拍							
业务主题:	用户	供应商结算	支付	P2P	众筹	供应链金融	店铺	服务评价	商品	销售订单			收起 ^
	订单结算	供应链	发票	促销	JSHOP	POP	EDM	仓储	SEM	配送	通讯		
	大数据	售后	客服	履约	O2O	微信手Q	易迅	广告	流量	闪购	团购		
	云平台	搜索推荐	组织架构		用户账户	EBS							
业务系统:	用户系统	问卷系统	Think	CRM2.0系统		青龙系统	赔付系统	维修中心	备件库	亚一系统	更多 v		
数据库:	mysql	oracle	sqlserver	hive	hbase	Jes	temp						

# 数据知识管理平台功能简介与使用

所在集市:	数据仓库	财务集市	营销集市	运营集市	建模集市	移动集市	审计	审计new集市	更多 ▾			
所在层级:	adm	app	gdm	fdm	dim							
集团公司:	京东商城	京东到家	京东智能	京东集团	金融集团	拍拍						
业务主题:	用户	商品	销售订单	供应链	促销	POP	仓储	配送	售后	客服	履约	更多 ▾
业务系统:	用户系统	问卷系统	Think	CRM2.0系统	青龙系统	赔付系统	维修中心	备件库	亚一系统	更多 ▾		
数据库:	mysql	oracle	sqlserver	hive	hbase	Jes	temp					
全部清空												

共搜索到6条结果

**gdm\_m04\_ord\_det\_sum**

★★★★☆

订单明细模型

创建: 2015-01-21 00:00:00 更新: 2015-01-21 16:10:28

1. 分区字段: dt, dp 2. dt每日生产系统订单归档日期, dp分为active和history, active为生产系统未归档的数据, history为生产系统按照一定规则归档的数据 3. 用法示例: 取某日订单数据: select \* from gdm\_m04\_ord\_det\_sum where dt>=yyyy-mm-dd; 4. 建议: 使用时在3中建议加上相应的业务时间限制

所在集市: 运营集市 供应商集市 营销集市 POP集市 O2O集市 网站移动研发集市 NSNG集市

微信手Q 数据仓库 风控 云平台集市 财务集市 建模集市 销量预测 广告 比价集市

推荐 交易平台集市 移动集市 成都集市

标签: 京东商城 销售订单 hive gdm

# 数据知识管理平台功能简介与使用

大数据专家认证

产品使用

gdm\_m04\_ord\_det\_sum 使用中

编辑 取消关注 评论

基本信息 字段信息 分区信息 版本历史 人工干预历史

查看各类元数据

中文名: 订单明细模型 标签:

表编号: 436214033 GDM EDW 订单 EDW 2.0 AD RECSYS

数据库: gdm CFO CMO COO BDA Mobile

表拥有者: dd\_edw

类型: 基础表

安全级别:

创建时间: 2013-03-29

最后访问时间:

第一负责人: 宫敬财

第一负责人email: gongjingcai@jd.com

第二负责人:

第二负责人email:

表引擎:

数据行数:

表同义词编号:

用途描述: 1. 分区字段: dt, dp 2, dt每日生产系统订单归档日期, dp分为active和history, active为生产系统未归档的数据, history为生产系统按照一定规则归档的数据 3. 用法示例: 取某日订单数据: select \* from gdm\_m04\_ord\_det\_sum where dt>=yyyy-mm-dd; 4. 建议: 使用时在3中建议加上相应的业务时间限制

备注: 因订单宽表gdm\_m04\_ord\_det\_sum 中的 长、宽、高、体积、型号, 这几个字段数据源 (fdm\_prd\_wareqita\_chain和fdm\_prd\_wareseo\_chain) 12月31日下线, POP商家编号 POP店铺编号, 这两个字段数据源 (pop\_ware\_sku) 可用性差, 考虑均由新数据源 odm m03 item sku da 替换上洋字段。历史数据不再重跑, 2013-12-31号及以后的数据将由新

加关注、评论

需要沟通时的联系人

可添加评论

评论

评论

李凡博 bj17hachal 发布于2014-08-05 16:04:03

因数据源对bj17hachal: 发现商品二条分拆编号有空的, 往回看, 这个商品好像是新出的, 订单明细应该是空的, 其他好多字段也是空的...

宫敬财 bj17hachal 发布于2014-07-10 15:12:26

因数据源bj17hachal: 电话明细是无效的

宫敬财 bj17hachal 发布于2014-07-10 15:15:53

因数据源bj17hachal: ordermemberid

宫敬财 bj17hachal 发布于2014-07-10 15:11:13

因数据源bj17hachal: 无效的。

李平国 bj17hachal 发布于2014-07-04 09:22:35

user\_id 是cookie里的userid吗?id中的第二个字段?

# 数据知识管理平台功能简介与使用

大数据专家认证

产品使用

gdm\_m04\_ord\_sum 使用中

编辑

关注

评论

基本信息 字段信息 分区信息 版本历史 人工编辑历史

字段名称	字段类型	字段中文名	备注
sale_ord_id	string	销售订单编号	俗称订单号，订单编号等。

基本信息 字段信息 分区信息 版本历史 人工编辑历史

分区路径	创建分区时间	创建时间
dp=HISTORY/dt=2014-04-29	2014-04-30 01:09:11	2014-04-30 07:59:55








基本信息 字段信息 分区信息 版本历史 人工编辑历史


更新属性	更新类型	原值	新值	更新时间	修改人
item_original_amount	新增		item_original_amount	2014-03-25 03:05:17	dd.edw

基本信息 字段信息 分区信息 版本历史 人工编辑历史

中文名	第一负责人	第一负责人email	第二负责人	第二负责人email	安全级别	使用状态	用途描述	备注
订单模型	宫敬财	gongjingcai@jd.com				1002		1.针对大家经常迷惑的移动端订单的提问，我整理了如下4张表供大家参考，如果有疑问可以提问：fdm_mobile_app_m_orderstat M版订单 fdm_mob_m_virtualorderstat 移动端虚拟订单，client='m' M版，其他为客户端 fdm_mobile_app_app_tuanorderstat_chain 移动端团购订单 client='mtuan' M版，其他为客户端 fdm_mob_app_orderstat 所有客户端

# 数据集成开发平台功能简介与使用

 京东大数据平台
  平台首页
  产品分类
  数据计算
 

 王威



数据订阅

任务管理

任务回收站

任务审批状态

任务执行记录

收件人组管理

数据查询

数据查询

代码执行记录

数据分享记录




数据加载记录

任务管理

数据仓库

ID/任务名/负责人/描述

更多条件



 新手指南
  任务执行记录
  集市状态

全部

+ 新建任务

数据仓库

恢复上次编辑信息

ID

任务名

负责人

描述

186161 PO

app

186089 PO

app

185071 PO

表

表名 / 描述

字段 / 描述

184987 PO

全部

最近使用

收藏

字段 / 描述

类型

184980 PO

app\_item\_third\_cate\_cd\_price\_ban...

item\_third\_cate\_cd

string

184980 PO

app\_jq\_data\_dif

item\_third\_cate\_...

string

184958 PO

app\_jq\_data\_total

price\_band

string

184944 PO

app\_jq\_data\_total

sale\_qtty

bigint

184080 ID

dual

181999 PO

app\_zwx\_region\_user\_target\_app\_d

子午线APP端地域用户指标

app\_s14\_app\_os\_ver\_region\_d

APP端操作系统按客户端版...

app\_zwx\_region\_impression\_app\_d

子午线APP端地域效果分析

app\_zwx\_visitor\_remain\_app\_d

运行

执行计划

优化格式

导入数据

打开

保存

订阅配置

打开自动提示

日志

结果

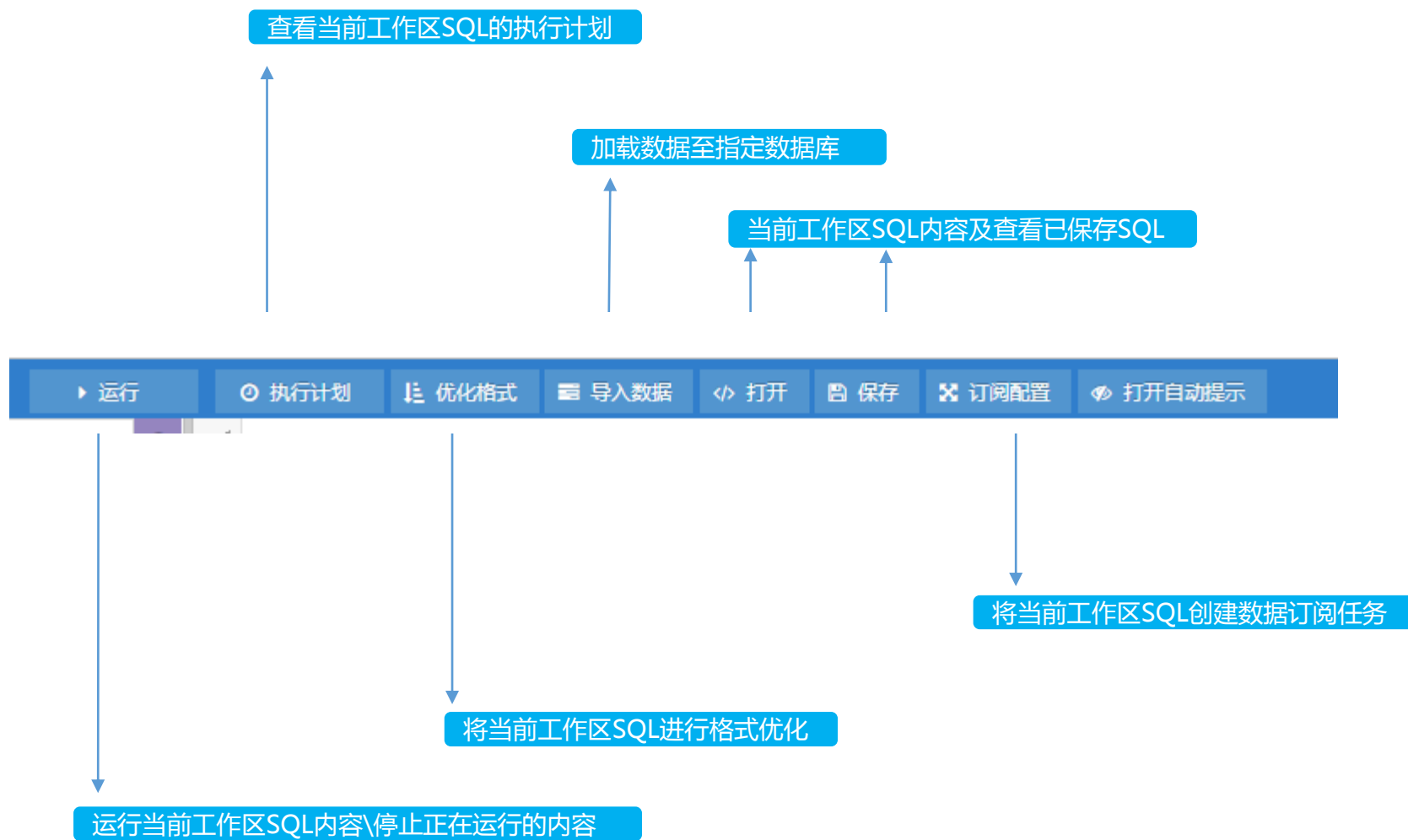
# 数据集成开发平台功能简介与使用

## 数据开发界面布局



# 数据集成开发平台功能简介与使用

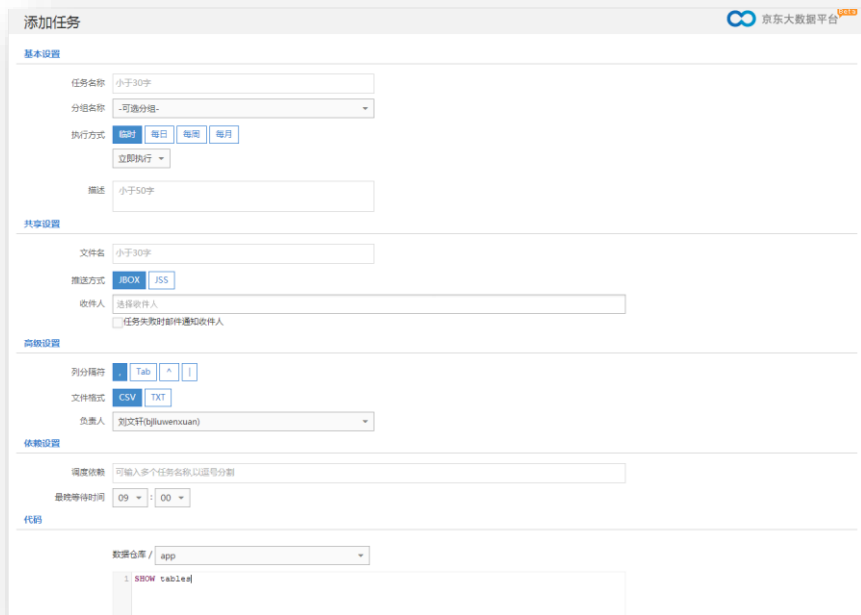
## 数据开发功能区



# 数据集成开发平台功能简介与使用

## 订阅任务面板

组名	数据订阅任务分组管理
名称	数据订阅任务名称
期望文件名	数据订阅任务生成的数据文件名称
列间隔符	数据文件列分隔符，下拉可选
文件类型	数据订阅任务生成的数据文件类型
负责人	数据订阅任务负责人，默认当前登录账户
数据源	SQL执行所在数据库
优先级	多个任务同时运行时的优先顺序
推送方式	生成数据文件的推送方式，选择JBOX即可
用户列表	生成数据文件存储在JBOX之后共享对象，邮箱前缀即可
类型	数据订阅任务的类型，临时或周期，年月日及时间等条件随之调整设置即可
下次运行时间	设定当前任务下一次的运行时间，可不填
失败后通知用户	当勾选此项时，若任务运行失败，会发邮件告知用户列表中全部用户
启用调度依赖	开发者功能，可不选
描述	对SQL内容及数据订阅任务可追加描述，用以说明任务用途信息等





# 数据集成开发平台功能简介与使用

任务管理界面即可查看提交的任务状态

任务管理

数据仓库

ID/任务名/描述

更多条件

新手指南 任务审批情况 任务执行记录 回收站 集市状态

全部

ide运营周报

测试分组测试分组测试5

分组设置

+ 新建任务

移动到

启用

禁用

删除

批量导出

ID	任务名	负责人	分组	状态	执行方式	执行时间	描述	有效期	操作
61132	test1	刘文轩	ide运营周报	成功	临时	2014-07-08 09:00:00		-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
61566	表名字数统计	刘文轩	ide运营周报	等待	每天	2014-08-08 09:00:00		74天	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
34457	ide八周用户数	刘文轩	ide运营周报	等待	每周	2014-08-10 12:00:00	ide到当前位置共八周的每周用户数统计, ...	35天	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
34333	数据集成开发平台上周五到...	刘文轩	ide运营周报	成功	临时	2014-07-22 21:06:18		-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
34329	数据集成开发平台上周五至...	刘文轩	ide运营周报	成功	临时	2014-07-05 11:39:33	数据集成开发平台上周五至本周四pv	-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
34360	ide上周五到本周四任务运行...	刘文轩	ide运营周报	成功	临时	2014-07-05 11:38:43		-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
34392	ide八周任务数统计	刘文轩	ide运营周报	成功	临时	2014-07-11 16:08:18	统计时间点为每周周四, 统计当前周到前...	-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>
57797	重跑任务	刘文轩	ide运营周报	成功	临时	2014-07-05 11:37:55	ide到当前位置共八周的每周用户数统计, ...	-	<a href="#">▶</a> <a href="#">≡</a> <a href="#">✎</a> <a href="#">🔗</a>

1 共 1 页

20

1 - 8 共 8 条

# 数据集成开发平台功能简介与使用

## 任务运行日志及jbox邮件通知

执行明细 ide八周用户数										
进入队列时间	开始时间	结束时间	耗时	状态	生成文件名	格式	推送方式	大小	行数	操作
2014-08-03 12:00:04	2014-08-03 12:00:08	2014-08-03 12:01:59	1分51秒	成功	ide_weekly_account_20140803120006.csv	JBOX		65.0B	10	<a href="#">≡</a> <a href="#">&lt;/&gt;</a>
2014-07-27 12:00:04	2014-07-27 12:00:09	2014-07-27 12:02:12	2分3秒	成功	ide_weekly_account_20140727120007.csv	JBOX		65.0B	10	<a href="#">≡</a> <a href="#">&lt;/&gt;</a>
2014-07-22 21:05:32	2014-07-22 21:05:38	2014-07-22 21:06:26	48秒	成功	ide_weekly_account_20140722210535.csv	JBOX		65.0B	10	<a href="#">≡</a> <a href="#">&lt;/&gt;</a>
2014-07-22 21:03:40	2014-07-22 21:03:43	2014-07-22 21:04:42	59秒	成功	ide_weekly_account_20140722210340.csv	JBOX		65.0B	10	<a href="#">≡</a> <a href="#">&lt;/&gt;</a>
2014-07-20 12:00:04	2014-07-20 12:00:07	2014-07-20 12:01:50	1分43秒	成功	ide_weekly_account_20140720120006.csv	JBOX		65.0B	10	<a href="#">≡</a> <a href="#">&lt;/&gt;</a>
2014-07-13 12:00:04	2014-07-13 12:00:12	2014-07-13 12:01:17	1分5秒							
2014-07-11 16:15:35	2014-07-11 16:15:37	2014-07-11 16:16:36	59秒							
2014-07-06 12:00:04	2014-07-06 12:00:12	2014-07-06 12:01:33	1分21秒							



2014/4/30 (周三) 16:11  
cdyjj-jbox@jd.com  
提数结果\_临时测试任务\_20140430

收件人 ☐ Undisclosed recipients:



京东 多·快·好·省

[京东网盘首页](#) [帮助](#)

亲爱的京东网盘用户，您好！

[用户数据部@jbox.com](#) 分享了文件——“临时测试任务\_20140430161035.csv” 给您。  
链接地址: <http://jd.com/f5dYvE>

以下是 [数据部@jbox.com](#) 给您的留言：

“任务名称：临时测试任务 任务负责人：王威”

本邮件由JBOX系统自动发出，请勿直接回复！

如果您有任何疑问或建议，请联系我们：[cdyjj-jbox@jd.com](mailto:cdyjj-jbox@jd.com)

谢谢！  
Thank you!

# 大数据专家认证

京东·大数据平台

---

北京市朝阳区北辰西路8号北辰世纪中心A座16层  
16F Building A, North-Star Century Center, 8 Beichen West Street,  
Chaoyang District, Beijing 100101