

大数据智能体工作台 DataBuddy

数据分析



腾讯云

【 版权声明 】

©2013–2026 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

数据分析

数据探索概述

数据分析

数据探索概述

最近更新时间：2026-09-11 20:57:31

什么是数据探索

数据探索（Data Exploration）是 DataBuddy 中面向数据分析师的轻量化分析入口。您在这里通过 Spark SQL 直接查询统一 Catalog 中的库表，把查询结果立即转成图表，并将常用查询沉淀为收藏，无需进入 Workflow 编排或 Notebook 开发流程。

举一个典型例子：业务同学想看「最近 7 天某个产品线的下单趋势」，只需要在数据探索中新建一段 SQL，选择 Catalog → Schema → 运行，下方表格立即出现结果，再点一下 **图表** 就可以得到一张折线图，分享链接给同事即可。

与数据工程中 SQL 的关系

数据探索与 开发工作台 SQL 共享同一套底层能力：相同的 Spark SQL 引擎、相同的 Serverless 数据分析型计算资源组、相同的统一 Catalog、相同的查询执行链路。两者的区别在于产品入口与定位：

维度	数据探索（5.1）	Studio SQL（3.1.3）
目标读者	数据分析师、业务分析人员	数据工程师、ETL 开发人员
产品入口	左侧导航 数据探索	左侧导航 开发工作台
主要场景	即席查询、可视化探查、临时分析	数据加工脚本开发、参数化复用、调度准备
协作语境	独立 SQL 脚本文件 + 全局运行历史 + 可视化	Notebook / SQL / Python 多文件协作 + Git
与 Workflow 的关系	通常不调度，验证想法用	可一键作为 SQL 任务 加入工作流

提示

两个入口的 SQL 文件在 Workspace 中互通：数据探索中保存的 `.sql` 文件可以在 Studio 中打开继续编辑，反之亦然；都遵循同一套 **文件权限** 规则。这意味着分析师沉淀的查询，工程师可以直接复用并加入 Workflow 调度。

核心能力

能力	说明
----	----

即席查询	在统一编辑器中编写 Spark SQL，秒级提交到 Serverless 资源组运行，结果以表格展示在编辑器下方。
查询收藏	将常用查询沉淀为收藏列表，下次打开后一键加载到编辑器。
结果可视化	一键将查询结果生成柱状图、折线图、饼图等可视化图表，无需切换到仪表盘。
结果导出	将查询结果下载为 CSV / Excel / TXT，用于离线汇报或对外交付。
AI 辅助	内置 Buddy「数据分析」模式，支持自然语言生成 SQL、查询解读、报告生成、归因分析等。
联邦查询	通过统一 Catalog 接入外部数据源（MySQL、Doris、ClickHouse 等），用同一套 Spark SQL 语法跨源查询。

关键概念

理解数据探索，您需要熟悉以下核心概念：

- **数据探索空间**：Workspace 中所有数据探索 SQL 文件的集合，按多级文件夹组织。文件级 ACL 与 Studio SQL 一致。等同于 Databricks 的 *SQL Editor / Queries*、阿里 DataWorks 的 *数据分析*。
- **Spark SQL 方言**：底层执行引擎统一使用 Spark SQL；所有跨数据源查询都走统一 Catalog，使用同一套语法。详见 SQL IDE 基础操作。
- **交互式分析型 Serverless 资源**：数据探索使用的执行资源，由平台动态扩缩容，分析师无需感知具体实例。详见 [计算资源概述](#)。
- **统一 Catalog**：基于 Unity Catalog 的三段命名空间 `catalog.schema.table`，承载内部引擎与外部数据源的元数据。详见 [三级命名空间](#)。
- **查询历史、查询收藏**：分别对应「过去运行过什么」「想再次复用什么」两类工作流；支持按文件 ACL 过滤个人可见范围。

完整术语请参考 [术语表](#)。

适用场景

场景	适用方式
临时业务问数	写一段 SQL，看一次结果，不需要保存
重复性日常查询	收藏常用 SQL，每次打开秒级加载
趋势 / 占比分析	在结果区直接生成图表，无需进入仪表盘
离线汇报 / 数据交付	将结果下载为 CSV / Excel
跨源联邦查询	通过 External Catalog 接入外部数据源

分析结果固化为报表	将探索成果沉淀到仪表盘，形成可分享视图
探索成果转生产任务	将 SQL 文件直接绑定为 Workflow SQL 任务

相关文档

- [SQL IDE 基础操作](#)
- [SQL 脚本运行](#)
- [SQL 任务](#)
- [仪表盘概述](#)
- [三级命名空间](#)
- [计算资源概述](#)