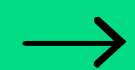


OpenBuild

Web3 数据分析师 系列公开课 入门基础课程



openbuild.xy

第二讲

链上数据结构及分析基础

区块链的核心数据结构，常用数据指标及分析工具概述

◆ 讲师简介

Ye @wyeeeh

- 清华大学-美国南加州大学 传媒&数据科学双硕士
- Exp. 链上数据分析2年经验，DeFi及MEV策略研究
- Prev. 清华大学学生区块链协会研究部长，EigenPhi数据分析师，Optimism NumbaNerd
- X | Telegram | Dune | GitHub: @wyeeeh





目录

Contents

01/ 链上数据结构

链上数据基本类型及结构示例

02/ 常见链上数据指标

总锁仓量(TVL)、活跃地址数、交易量、Gas 费用

03/ 链上数据分析流程及工具

数据ETL, Etherscan, Nansen, Dune

01/ 链上数据结构

链上数据基本类型、交易、区块及合约数据示例

◆ 链上数据基本类型

区块数据 2

区块自身信息数据，如时间戳、验证者、包含的交易及交易排序等；

1 交易数据

包含交易信息的数据，如转账资金、收发地址、手续费等；

3 合约数据

区块链上部署的非用户交互产生的代码数据，即预先定义好的智能合约代码。

回顾：链上交易

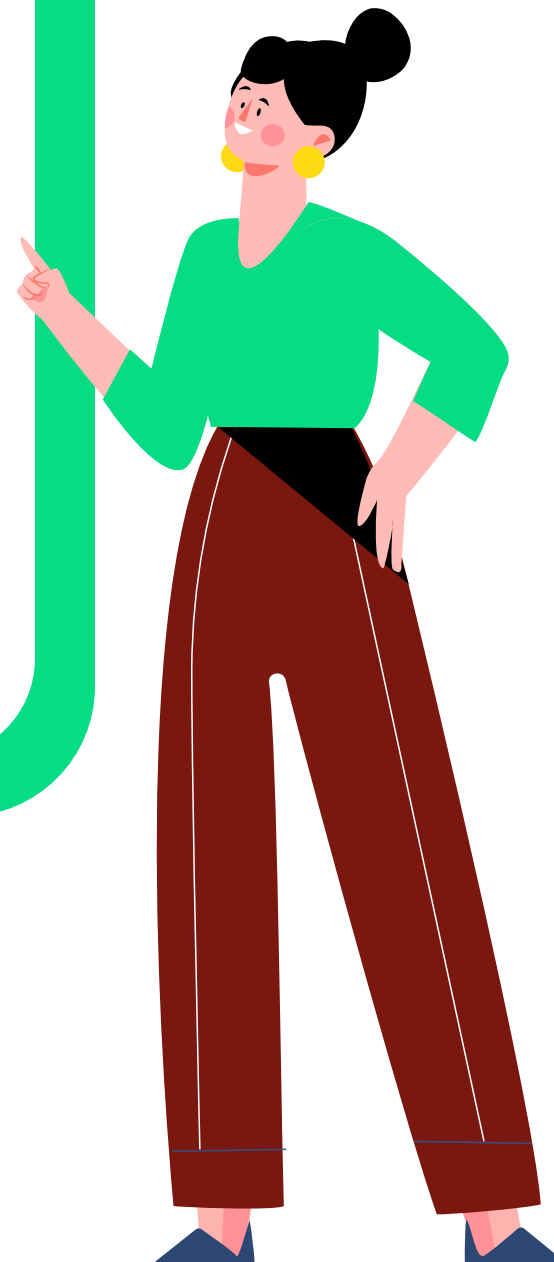
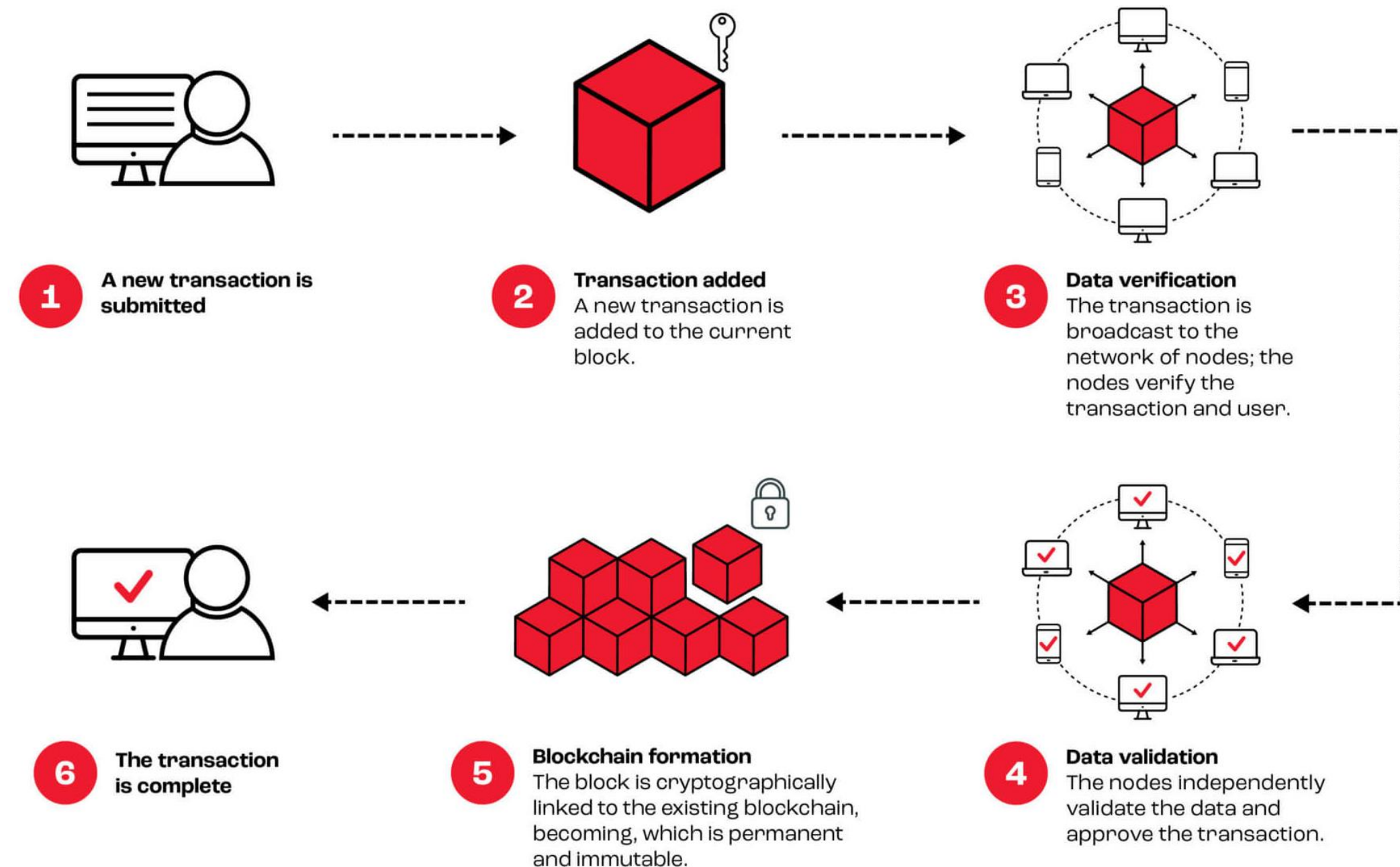


链上交易是区块链状态改变的主要方式，每笔交易的基本生命周期包括创建、广播、打包、确认。

交易创建：用户通过钱包或合约交互生成交易，指定发送方、接收方和金额。

交易广播：交易被发送到区块链网络，所有节点收到交易并验证合法性（如余额是否充足）。

交易确认：矿工或验证节点将交易打包到区块中，添加到区块链上。



✦ 交易数据

Transaction Details

Buy

Exchange

Earn

Gaming

Overview

State

Comments

Transaction Hash:

0x3af096859a880d9c33718eda59cb96e1504db7390d0e086c7260d91e87139eab

Status:

Success

Block:

15537394

545143 Block Confirmations

Timestamp:

76 days 4 hrs ago (Sep-15-2022 06:42:59 AM +UTC) | Confirmed within 30 secs

Sponsored:

From:

0xeb9e17781b43517999827d2b256ee0ebdd00b287

To:

0xf99c02cae906ff0c3b3b277cb67ab5f4f39d7796

Value:

0.031391199870524086 Ether

(\$39.78)

Transaction Fee:

0.001781047686495 Ether

(\$2.26)

Gas Price:



0.000000084811794595 Ether (84.811794595 Gwei)

Ether Price:

\$1,472.51 / ETH

- 交易哈希
- 交易状态
- 区块
- 时间戳
- 收发双方地址
- 该笔交易转账的ETH数量
- 交易费
- Gas费
- Nonce：交易发起者账户的交易计数器，用于防止重复交易。

区块数据

Block Height:	15537394 < >
Status:	Finalized
Timestamp:	796 days ago (Sep-15-2022 06:42:59 AM +UTC)
Proposed On:	Block proposed on slot 4700013 , epoch 146875
Transactions:	80 transactions and 1 contract internal transaction in this block
Fee Recipient:	Fee Recipient: 0xeE...4C8 in 17 secs
Block Reward:	45.031378244766393234 ETH (0 + 46.494902574979045804 - 1.46352433021265257)
Total Difficulty:	58,750,003,716,598,352,816,469
Size:	18,559 bytes
Gas Used:	29,983,006(99.94%)  +100% Gas Target
Gas Limit:	30,000,000
Base Fee Per Gas:	0.000000048811794595 ETH (48.811794595 Gwei)
Burnt Fees:	 1.46352433021265257 ETH
Extra Data:	0x (Hex:Null)
Ether Price:	\$1,472.51 / ETH

- 区块高度
- 区块哈希
- 时间戳
- 验证者/矿工地址
- 区块中的交易列表
- Gas费

Slot:	4700013
Epoch:	146875
Proposer Index:	347963
Slot Root Hash:	0x810a00400a80cdfc11ffdcf17ac404ac4dba215b95221955a9dfddf163d0b0d
Parent Root Hash:	0x60e751f7d2cf0ae24b195bda37e9add56a7d8c4b75469c018c0f912518c3bae8
Beacon Chain Deposit Count:	435603
Slot Graffiti:	0x (Hex:Null)
Block Randomness:	0xa86c2e601b6c44eb4848f7d23d9df3113fbcac42041c49cbcd5000cb4f118777
Randao Reveal:	0x91bbcabf3543508a785153162ce0920e6c25d1dc75832c31d5a14bfaa0eb151409a2215652138e2dc6ababc08361d238106c4e5df7d8ee8e5e67619cd89aceff328fa3758c87354d08544808c4efa0fce6a06827ff0d0852eaf3b647a9826d95

◆ 智能合约数据

✔ Contract Source Code Verified (Exact Match)



Contract Name: TetherToken

Optimization Enabled: No with 0 runs

Compiler Version v0.4.18+commit.9cf6e910

Other Settings: default evmVersion, Audited

📄 Contract Source Code (Solidity) [Audit Report](#)

b IDE Outline More Options



```
1  /**
2   *Submitted for verification at Etherscan.io on 2017-11-28
3   */
4
5   pragma solidity ^0.4.17;
6
7   /**
8    * @title SafeMath
9    * @dev Math operations with safety checks that throw on error
10   */
11  library SafeMath {
12      function mul(uint256 a, uint256 b) internal pure returns (uint256) {
13          if (a == 0) {
14              return 0;
15          }
16          uint256 c = a * b;
17          assert(c / a == b);
18          return c;
19      }
20
21      function div(uint256 a, uint256 b) internal pure returns (uint256) {
22          // assert(b > 0); // Solidity automatically throws when dividing by 0
23          uint256 c = a / b;
24          // assert(a == b * c + a % b); // There is no case in which this doesn't hold
25          return c;
26      }
27  }
```

📄 Contract Security Audit

- Callisto Network - July 10th, 2019 - [Security Audit Report](#)

☰ Contract ABI

Export ABI



```
[{"constant":true,"inputs":[{"name":"name","outputs":[{"name":"","type":"string"}],"payable":false,"stateMutability":"view","type":"function"}],{"constant":false,"inputs":[]}]
```

- 合约地址
- 合约定义逻辑（函数）
- 合约定义事件
- 合约定义变量

02/ 常见链上数据指标

总锁仓量(TVL)、活跃地址数、交易数及交易量、Gas 费用

◆ 总锁仓量 (TVL)



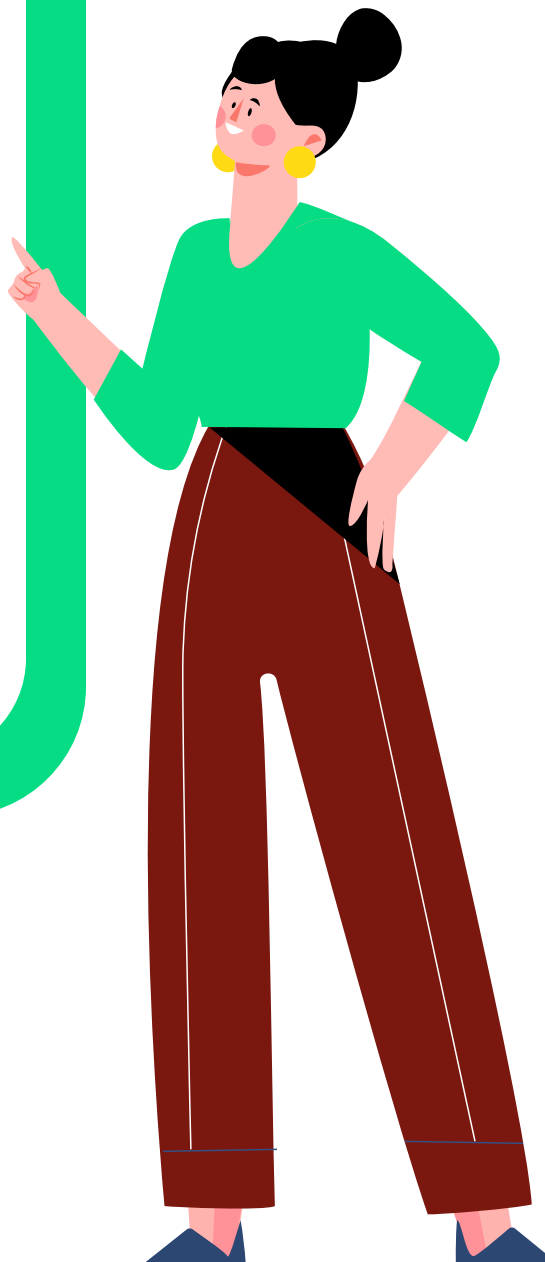
总锁仓量 (Total Value Locked, TVL) 表示存入 DeFi 协议中的总资产价值，包括流动性池、抵押贷款和收益聚合中的总资产价值。

● 应用

TVL是衡量 DeFi 协议市场占有率和吸引力的重要指标，TVL的增加通常表示协议的增长和吸引力。

● 计算

TVL = 所有存入协议中的资产价值的总额
(包含币种价值，常用USD或ETH计价)。
在某个质押协议上，Alice 存入 10 ETH，
Bob 存入 20 ETH，则该协议的 TVL 为 30 ETH。



◆ 活跃地址数



活跃地址数是指在一定时间段内至少执行过一次交易的钱包地址数量。与Web2中的活跃用户 (Active User) 类似，链上的“活跃地址”表示的是有实际行为发生的用户地址。

● 应用

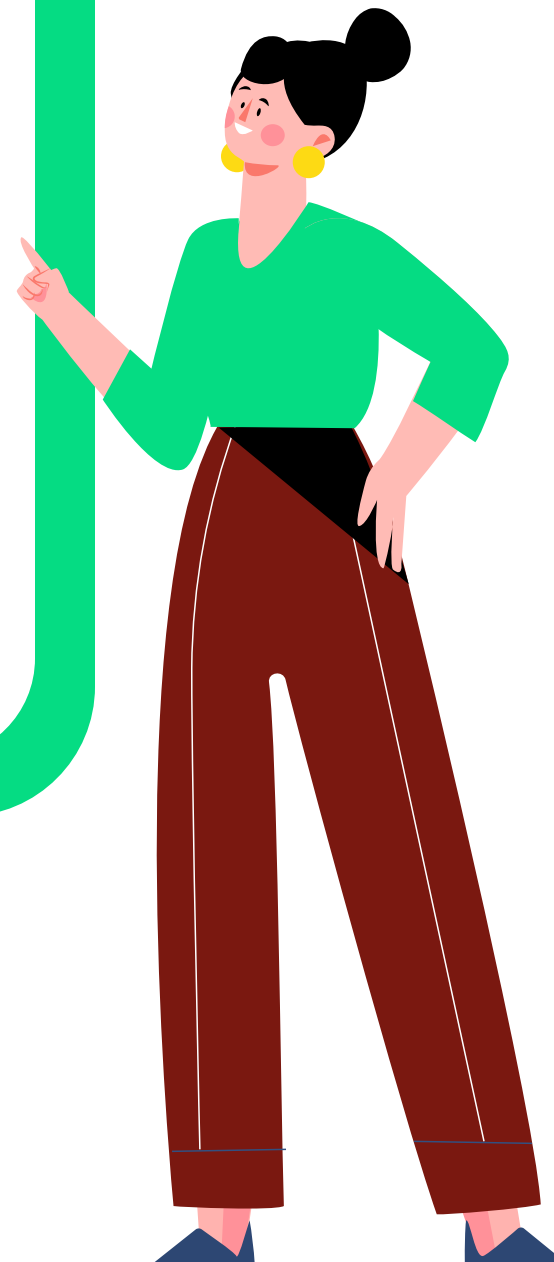
在区块链网络中，活跃地址的数量能够反映区块链及应用的用户基础及增长趋势。

对于DeFi协议，活跃地址可能是参与质押、交易、借贷等活动的用户地址。

● 计算

活跃地址数 = 在指定时间段内进行过交易的唯一地址数量

在某个去中心化交易所的活动中，Alice、Bob和Charlie各自进行了交易，因此该时间段内有三个活跃地址。



◆ 交易数及交易量



交易数 (Transaction Count) 是在一定时间段内链上发生的所有交易的数量，交易量 (Trading Volume) 是这些交易涉及资产价值的总额。

● 应用

交易量通常衡量区块链网络的活跃度，它反映了有多少资产被转移或交换。对于DeFi协议，交易量可能包括通过协议进行的代币交换或质押操作。

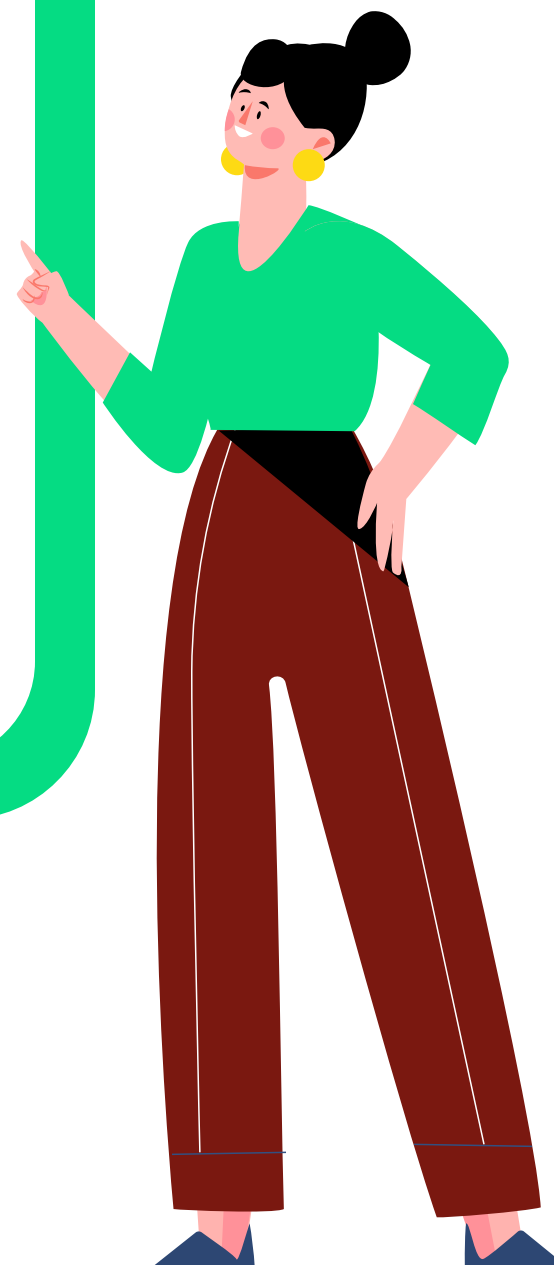
交易量是评估区块链网络健康、交易活动以及市场情绪的重要指标。高交易量通常意味着网络的需求增加。

● 计算

交易数 = 特定时间段内所有交易哈希数量

交易量 = 特定时间段内所有交易的代币数量价值总和（常用USD或ETH计价）。

假设Alice和Bob在某个区块链上进行了三笔交易，分别转账2ETH、3ETH和1ETH，那么该时间段的交易量就是6ETH（无论收发地址是谁）。



◆ 交易费用/Gas费用



交易费用 (Transaction Fee) 是用户在区块链网络中进行交易时支付给矿工或验证者用打包上链的费用，是用户交易的成本，通常又叫Gas费用。

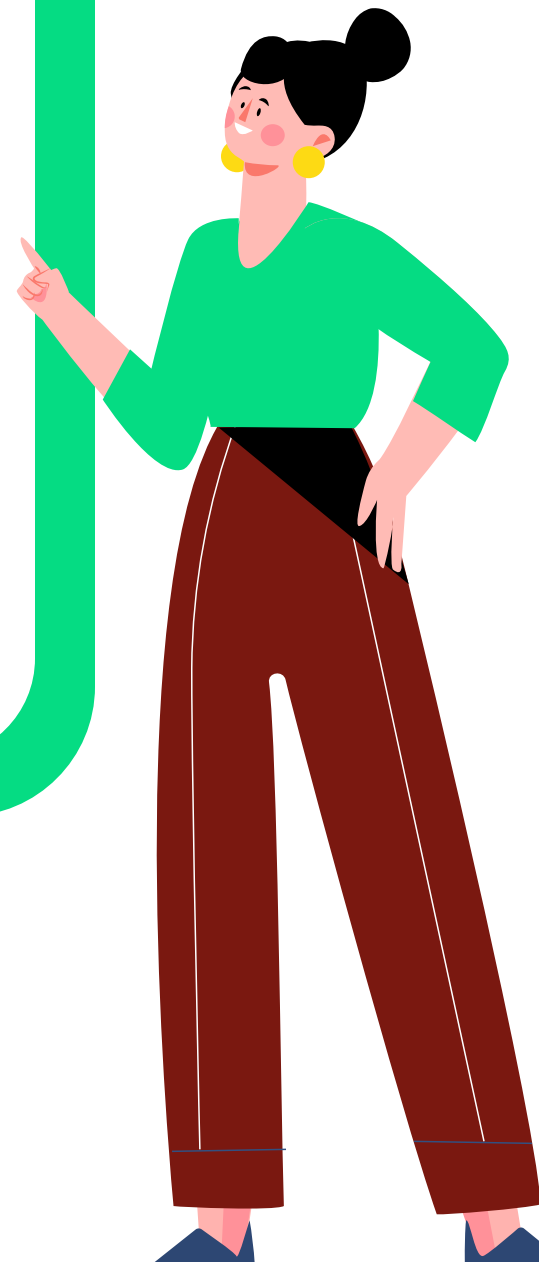
● 应用

交易费反映了区块链网络的拥堵程度和交易成本。在动态交易费用的背景下，高交易费通常表示网络较为拥堵，用户交易成本高。

● 计算

交易费 = 交易使用的Gas数量 × Gas价格。

Transaction Fee:	0.001781047686495 ETH \$5.94	= 21000 * Gas Price
Gas Price:	84.811794595 Gwei (0.000000084811794595 ETH)	
Gas Price = Base + Priority = (48.81 + 36) = 84.81 Gwei		
Ether Price:	\$1,472.51 / ETH	
Gas Limit & Usage by Txn:	21,000 21,000 (100%)	
Gas Fees:	Base: 48.811794595 Gwei Max: 133.62358919 Gwei Max Priority: 36 Gwei	
Burnt & Txn Savings Fees:	Burnt: 0.001025047686495 ETH (\$3.42)	Txn Savings: 0.001025047686495 ETH (\$3.42)



03/ 链上数据分析流程及工具

数据ETL, 主流工具 (Etherscan, Dune)

◆ 数据ETL



和Web2数据分析类似，一个完整的链上数据分析流程需要经过以下几个步骤：数据抽取 (Extract)→清洗转换 (Transform) →加载 (Load)到数据仓库→依据不同业务线及观测指标分析数据→输出量化结果。

● 数据抽取 (Extract)

数据抽取从区块链RPC节点或区块链浏览器 (如Etherscan)中获取原始数据。

区块链数据通常以非结构化的日志形式存储，且包含大量的交易数据、状态信息和事件日志 (Event, Log, Traces)。

● 清洗转换 (Transform)

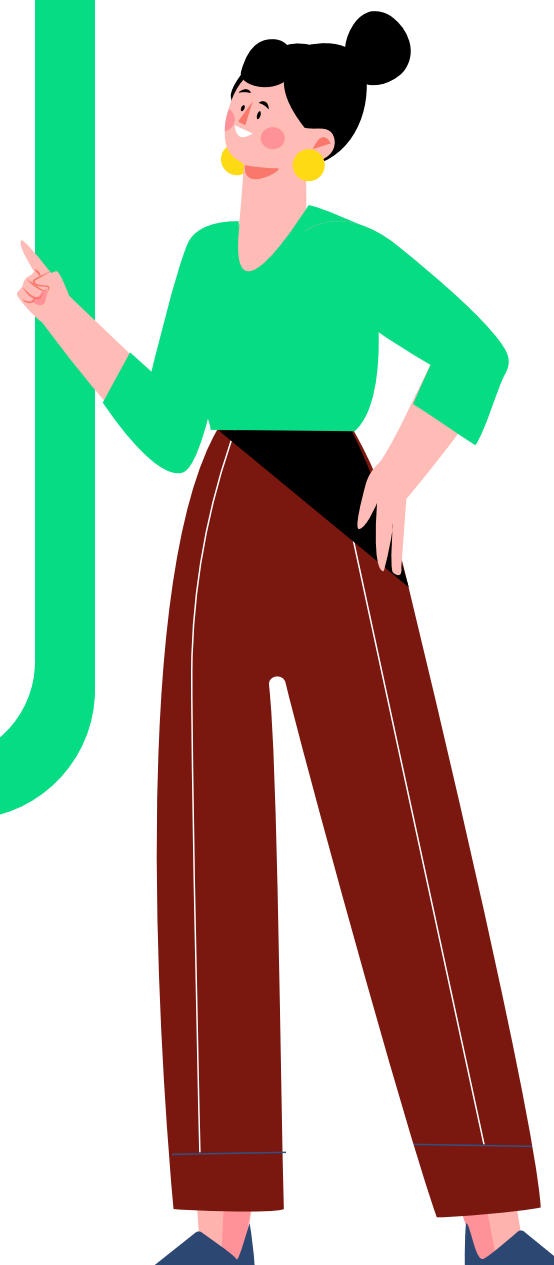
智能合约被调用执行时的交易输入输出和日志数据通常以十六进制HEX格式存储，无法直接分析。

合约解析需要通过智能合约ABI（应用二进制接口）将这些HEX格式的输入数据解码，恢复成原始函数参数等信息，使其能够以结构化的数据形式呈现（如数字、地址、字符串等）。

● 数据加载及分析 (Load)

根据分析目标进行计算和查询。在区块链数据分析中，加载的数据通常是合约相关的交易记录和事件日志，并依据特定的指标进行筛选和聚合分析。

例如，针对某一特定合约的TVL，首先需要确定合约地址，再对该合约触发的deposit事件交易记录进行筛选和汇总计算。



◆ 链上数据产品生态



基于链上数据分析的ETL流程，区块链数据产品可以分为数据源、数据开发工具（Data Dev）和数据应用（Data App）三类。

● 数据源

数据源提供对原始区块链数据的索引，它们通过节点或API访问不同的区块链网络，提取链上的交易、事件日志、合约信息等原始数据。

典型应用：The Graph, Alchemy

● 数据开发工具

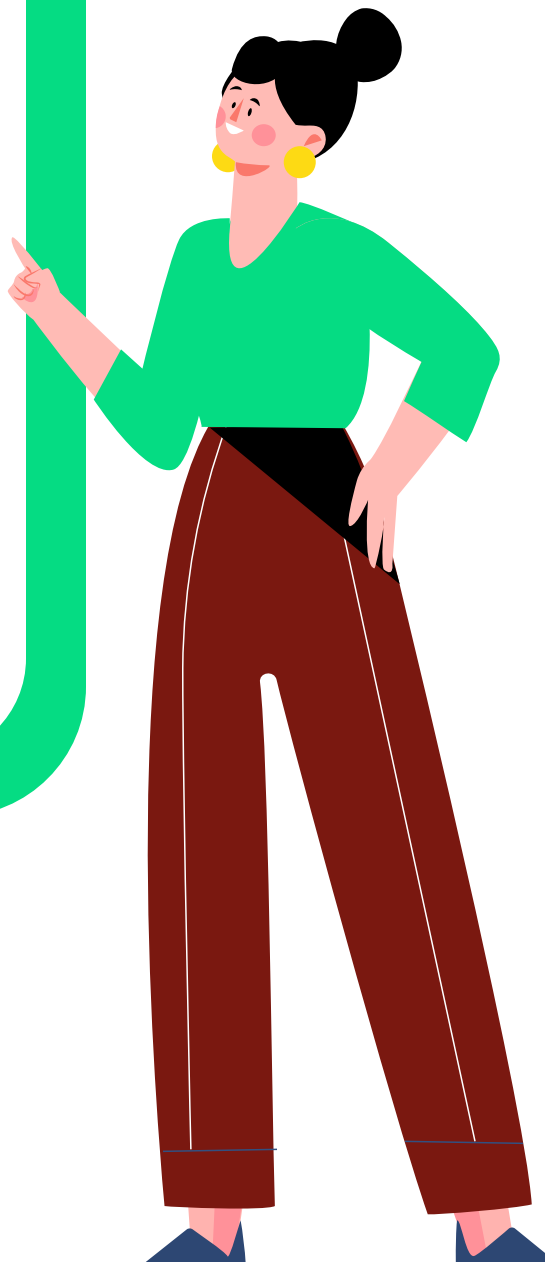
面向数据分析开发人员，可以完成ETL的工作，将交易、状态和事件日志数据解析成可以用SQL或GraphQL等传统语言查询的格式并存储在数据库中，供后续查询。自由定制程度高，不直接提供数据分析结果。

典型应用：Dune, Flipside

● 数据应用

面向普通用户，提供可直接访问的数据内容，供投资者参考决策。一般会以数据可视化看板等提供结果，易读性强。自由定制程度低，用户只能阅读预先分析好的数据结果，较难按需求自行定制输出。

典型应用：Nansen, Debank, DefiLlama

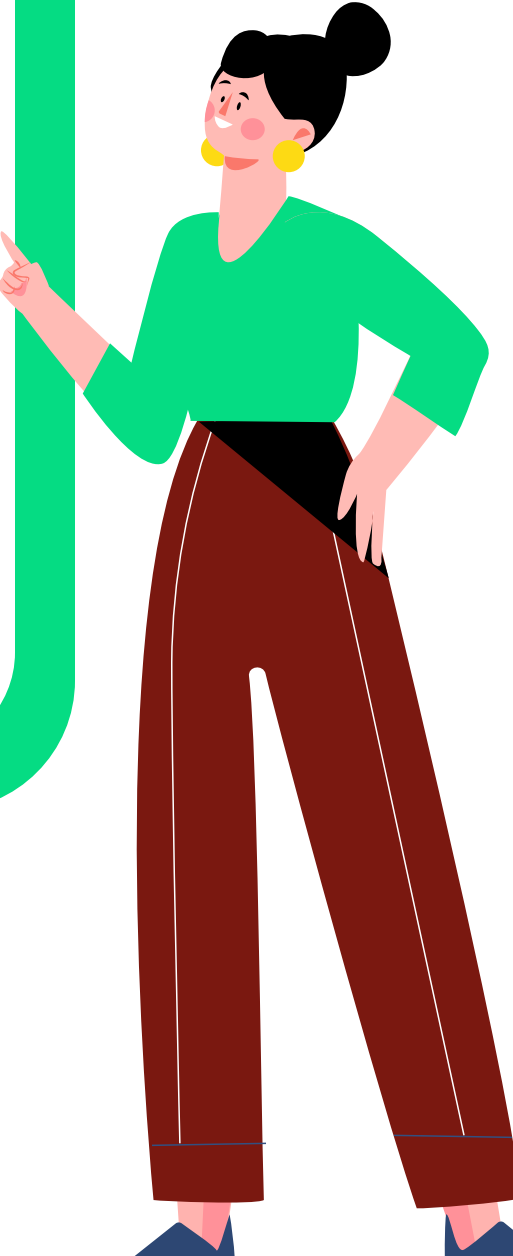


链上数据产品生态

目前链上数据分析赛道生态位已经有诸多玩家，根据面向的不同用户类型、不同赛道都有对应的数据产品布局。



加密数据产品堆栈 @zk7hao



常用链上数据分析工具：区块链浏览器 Etherscan

Etherscan

HomeBlockchainTokensNFTsResourcesDevelopersMore

Transaction Details

BuyExchangePlay

Sponsored: CryptoSlots - 1,000,000 USD Jackpot & Provably Fair Casino Games

OverviewState

Transaction Hash:

0x3af096859a880d9c33718eda59cb96e1504db7390d0e086c7260d91e87139eab

Status:

Success

Block:

155373945710485 Block Confirmations

Timestamp:

799 days ago (Sep-15-2022 06:42:59 AM UTC)

Transaction Action:

Transfer 0.031391199870524086 (\$104.99) ETH To 0xF99c02CaE906ff0c3b3B277Cb67ab5f4F39D7796

Sponsored:

BC.GAME

TOP CRYPTO GAMES AND BETS IN CRYPTO

REGISTER NOW

18+ GambleAware

From:

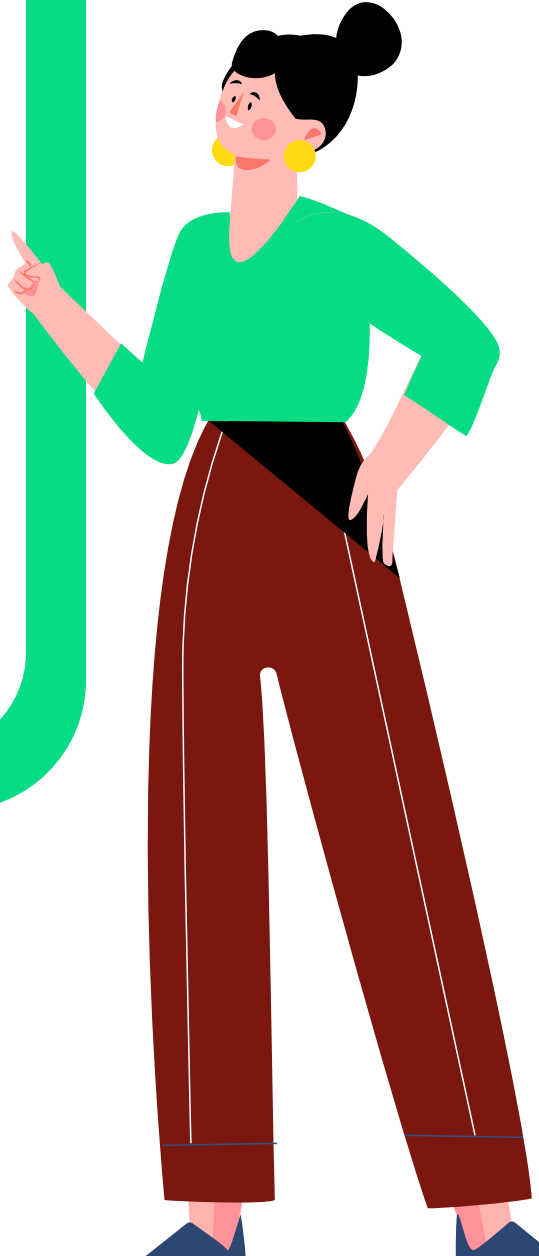
0xeB9E17781b43517999827D2B256ee0EBDd00B287

To:

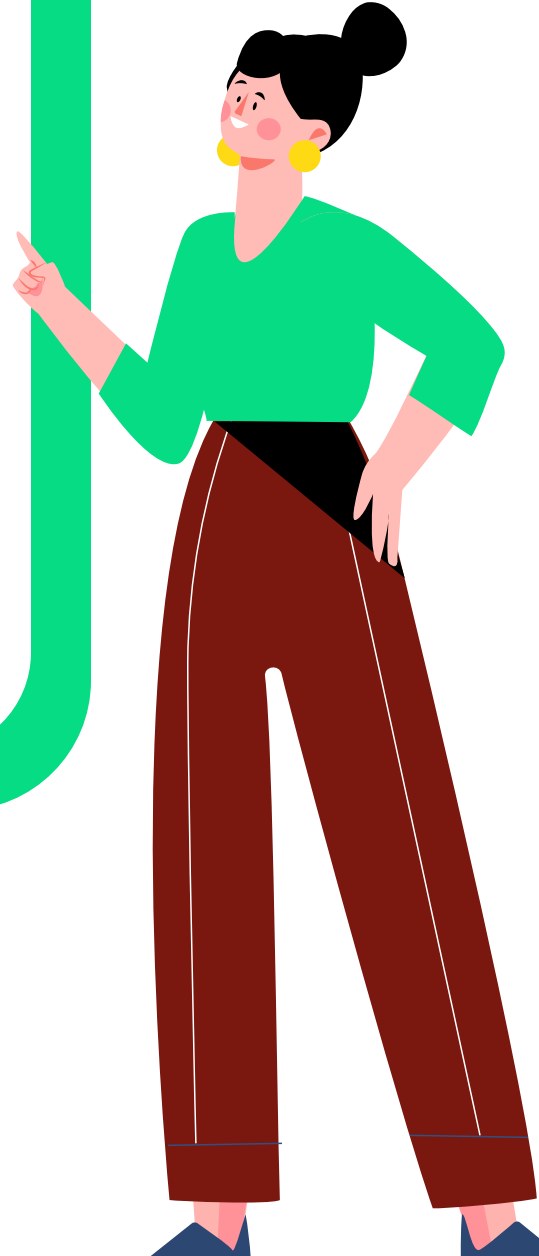
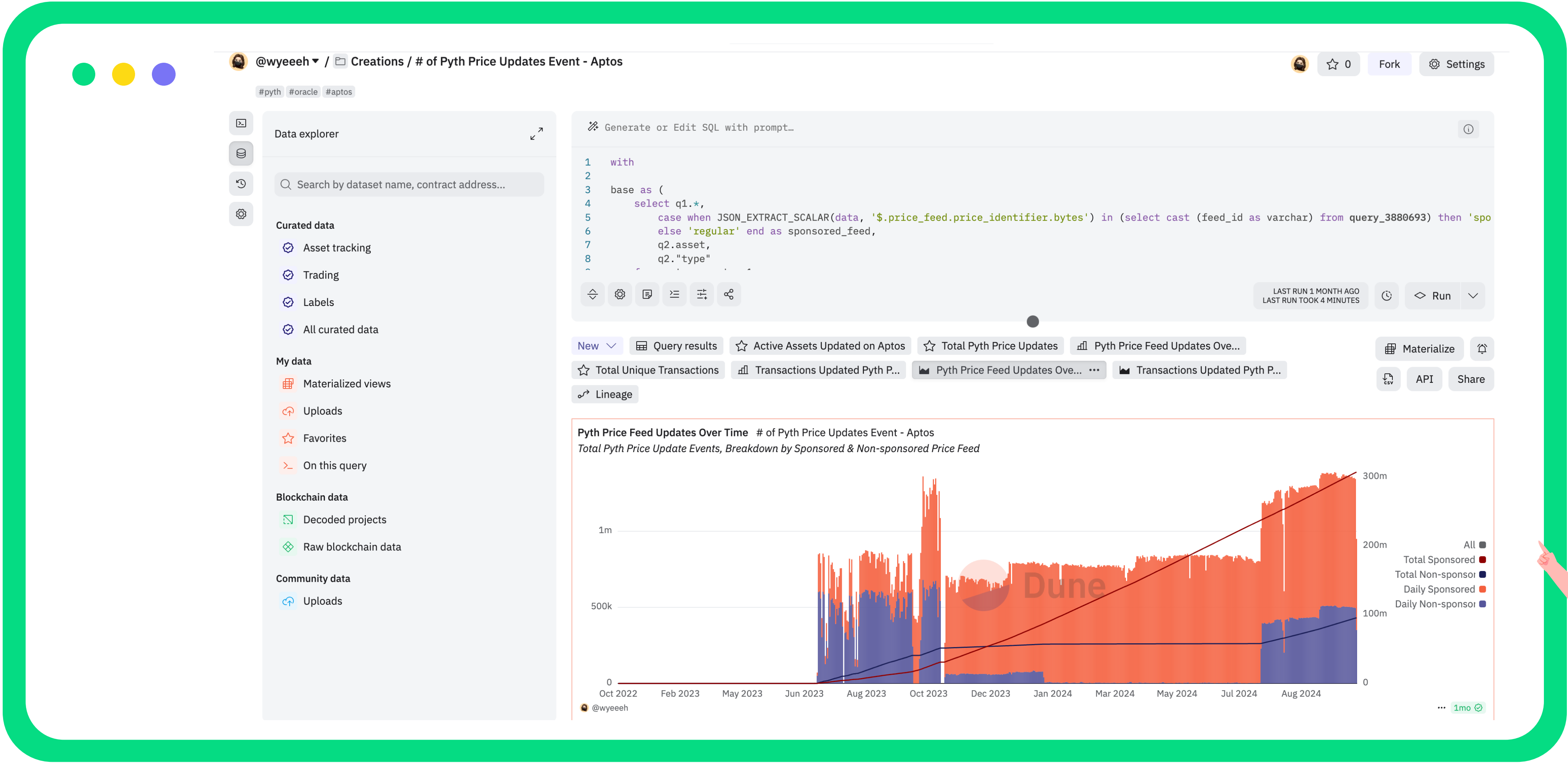
0xF99c02CaE906ff0c3b3B277Cb67ab5f4F39D7796

Value:

0.031391199870524086 ETH\$104.99



常用链上数据分析工具：Dune



◆ 参考资料

- 链上数据分析：从野蛮生长的DeFi世界洞察价值
<https://mp.weixin.qq.com/s/UzeL6y8iKN4am6xgPlj3wA>
- 当我们谈 Web3 数据，我们在谈些什么？ | ZONFF Research
<https://mp.weixin.qq.com/s/TnVlj93CYchY0wnaazLRTg>
- IOSG Weekly Brief | 链上数据分析平台现状与展望 #132
https://mp.weixin.qq.com/s/o1pO7unj3cUS9sWt4q_gBw
- 成为链上数据分析师
<https://sixdegreelab.gitbook.io/mastering-chain-analytics/>



Thanks for watching!



WeChat



Twitter