* **Base Paper**

**T**otal Attribute: 42

Total Dataset: 4681

**N**ame Of Attributes:

1. Avg\_min\_between\_sent\_tnx

2. Avg\_min\_between\_received\_tnx

3. Time\_Diff\_between\_first\_and\_last (Mins)

4. Sent\_tnx

5. Received\_tnx

6. Number\_of\_Created\_Contracts

7. Unique\_Received\_From\_Addresses

8. Unique\_Sent\_To\_Addresses

9. Min\_Value\_Received

10. Max\_Value\_Received

11. Avg\_Value\_Received

12. Min\_Val\_Sent

13. Max\_Val\_Sent

14. Avg\_Val\_Sent

15. Min\_Value\_Sent\_To\_Contract

16. Max\_Value\_Sent\_To\_Contract

17. Avg\_Value\_Sent\_To\_Contract

18. Total\_Transactions (Including\_Tnx\_to\_Create\_Contract)

19. Total\_Ether\_Sent

20. Total\_Ether\_Received

21. Total\_Ether\_Sent\_Contracts

22. Total\_Ether\_Balance

23. Total\_ERC20\_Tnxs

24. ERC20\_Total\_Ether\_Received

25. ERC20\_Total\_Ether\_Sent

26. ERC20\_Total\_Ether\_Sent\_Contract

27. ERC20\_Uniq\_Sent\_Addr

28. ERC20\_Uniq\_Rec\_Addr

29. ERC20\_Uniq\_Rec\_Contract\_Addr

30. ERC20\_Avg\_Time\_Between\_Sent\_Tnx

31. ERC20\_Avg\_Time\_Between\_Rec\_Tnx

32. ERC20\_Avg\_Time\_Between\_Contract\_Tnx

33. ERC20\_Min\_Val\_Rec

34. ERC20\_Max\_Val\_Rec

35. ERC20\_Avg\_Val\_Rec

36. ERC20\_Min\_Val\_Sent

37. ERC20\_Max\_Val\_Sent

38. ERC20\_Avg\_Val\_Sent

39. ERC20\_Uniq\_Sent\_Token\_Name

40. ERC20\_Uniq\_Rec\_Token\_Name

41. ERC20\_Most\_Sent\_Token\_Type

42. ERC20\_Most\_Rec\_Token\_Type

* **Omitted Attributes From Base Paper (Not In Generated Dataset):**

1. Number of Created Contracts

2.Min Value Sent to Contract

3.Max Value Sent to Contract

4.Avg Value Sent to Contract

5.Total Ether Send Contract

6. ERC20\_total\_Ether\_sent\_contract

7. ERC20\_Avg\_Time\_Between\_Contract\_Tnx

8. ERC20\_Uniq\_Rec\_Contract\_Addr

**NOTE**: As it is End User Addresses Not the smart Contract address.

* **Generated Database**

Description: The dataset was generated through web scraping from Etherscan.io API and also from the base research paper dataset. Due to the presence of 54 attributes, the data extraction process was divided into three parts. The first part focuses on transaction-related attributes such as average, minimum, and maximum values etc. The second part covers ERC20 token-related attributes, while the third part includes newly added gas-related attributes.

**T**otal Attributes: 54

Total Dataset: 2806

Illicit: 1177 (Addresses Taken From The Base Paper Dataset)

Non-Illicit: 1628 (New Active Accounts , Not Tagged - Addresses Taken By Web Scraping Using API Of Etherscan.io)

Newly Added Attributes: 9 new attributes

1. **Avg Gas Price**

* Gas price refers to the amount of Ether a user is willing to pay for each unit of gas to execute a transaction.
* This attribute represents the average gas price paid across all transactions made by the address.
* Higher gas prices lead to faster transaction processing, while lower gas prices may result in delays.

✅ **Formula:**

Avg Gas Price = sum(gas\_prices) / len(gas\_prices)

1. **Avg Gas Limit**

* Gas limit defines the maximum amount of gas a user is willing to spend on a transaction.
* This attribute represents the average gas limit set by the address for all transactions.
* If the gas limit is too low, the transaction may fail.

✅ **Formula:**

Avg Gas Limit=sum(gas\_limits) / len(gas\_limits)

1. **Avg Gas Used**

* Every Ethereum transaction requires gas to process.
* Some transactions use more gas, some use less.
* "Avg Gas Used" tells us, on average, how much gas is consumed per transaction

**✅ Formula:**

Avg Gas Used**=**sum(gas\_used) / len(gas\_used)

1. **Transaction Frequency in min.**

* Represents how frequently the address initiates transactions.
* A lower value means the account is more active, while a higher value suggests infrequent activity.

✅ **Formula:**Transaction Frequency= Time Difference Between First and Last Transaction/(Total Transaction-1)

1. **Gas Price Volatility**

* Measures how much the gas price fluctuates for an address over multiple transactions.
* A higher value indicates that the user paid inconsistent gas fees, possibly due to network congestion or urgent transactions.
* A lower value means stable gas prices.

✅ **Formula:**

Gas Price Volatility=Max Gas Price - Min Gas Price

1. **Transaction Success-Failure Ratio**

* Measures the reliability of an address in executing transactions.
* A higher ratio means the user has successfully executed most transactions.
* A lower ratio suggests frequent failures, which could be due to insufficient gas fees or contract execution errors.

✅ **Formula:**

Transaction Success-Failure Ratio=success\_count / failure\_count

1. **Base Fee (Retrived 0 – Calculated Using Formula In Excel )**

* In the Ethereum blockchain, the base fee is the minimum amount you must pay for a transaction.
* It automatically changes based on how busy the network is. If a lot of people are making transactions, the base fee goes up; if fewer people are using the network, the base fee goes down.

✅ **Formula:**

Base Fee=Avg Gas Price−(Gas Price Volatility/2)

1. **Tip (Retrived 0 – Calculated Using Formula In Excel )**

* Additional Fee Paid to Miners to Prioritize Transactions
* A user can **add a tip** (also called **priority fee**) to incentivize miners/validators to process their transactions faster.

✅ **Formula:**

Tip=Gas Price Volatility/2

* **The gas-related features (Gas Used, Gas Limit, Gas Price, etc.) in Ethereum transactions are typically associated only with outgoing transactions (sent transactions), not incoming transactions.**