

HW2: Exploring Bitcoin transactions

Yihui Wang UFID# 8316-4355

Part 1: Transactions analysis

1. What is the number of transactions and addresses in the dataset?

The number of transactions is 10000055, the number of addresses is 8385065.

2. What is the Bitcoin address that is holding the greatest amount of bitcoins? How much is that exactly? Note that the address here must be a valid Bitcoin address string. To answer this, you need to calculate the balance of each address. The balance here is the total amount of bitcoins in the UTXOs of an address.

The Bitcoin address is 1933phfhK3ZgFQNLGSDXvqCn32k2buXY8a.
The greatest amount is 11111100000000.

3. What is the average balance per address?

The average balance per address is 125990554.98689432

4. What is the average number of input and output transactions per address? What is the average number of transactions per address (including both inputs and outputs)? An output transaction of an address is the transaction that is originated from that address. Likewise, an input transaction of an address is the transaction that sends bitcoins to that address.

The average number of input transactions per address: 2.405294890379502
The average number of output transactions per address: 2.7747914893921513
The average number of transactions per address: $2.405294890379502 + 2.7747914893921513 = 5.18008637977$

5. What is the transaction that has the greatest number of inputs? How many inputs exactly? Show the hash of that transaction. If there are multiple transactions that have the same greatest number of inputs, show all of them.

The transaction that has the greatest number of inputs is
9621b3c67f9bddd3de65fafc488087b8f2b40b638e3a06209a904c66c0b32982.
The number of inputs is 1312.

6. What is the average transaction value? Transaction value is the sum of all outputs' value.

The average transaction value is 12315588064.035421

7. How many coinbase transactions are there in the dataset?

The number of coinbase transactions is 212576

8. What is the average number of transactions per block?

The average number of transactions per block is 47.04225782778865

Part 2: Address de-anonymization

1. How many users are there in the dataset?

There are 4241424 users in the dataset

2. Answer questions 2, 3, and 4 in part 1 by replacing "address" with "user". Note that each user is identified by the addresses that are owned by him/her. Thus, in answering question 2 (i.e., the user who is holding the greatest amount of bitcoins), you need to list all the user's addresses.

2.1 The Bitcoin addresses that is holding the greatest amount of bitcoins are 0, 316, 470, 586, 628..... (There are 2578820 addresses more)
The largest balance is 86562187546646

2.2 The average balance per user is 188007724

2.3 The average number of input transactions per user: 4.705303

The average number of output transactions per user: 5.428123

The average number of transactions per user: $4.705303 + 5.428123 = 10.13343$

3. Give the hash of the transaction sending the greatest number of bitcoins to the user who is holding the greatest balance.

Hash of the transaction sending the greatest number of bitcoins to the user who is holding the greatest balance:

9621b3c67f9bddd3de65fafc488087b8f2b40b638e3a06209a904c66c0b32982.

Instructions on running the code:

All the code is written by Python.

To run the code, first, put the code in the same folder with the datasets needed.

And then, open the Terminal and enter *python <filename>.py* to run the code.