**Coding question:**

Find the Bitcoin transactions from year 2011 at  \url{https://utdallas.app.box.com/s/bu30643q4l0a79b4907c2a51tx31s16a} where transaction inputs and transaction outputs are given in two separate files.

(Our full dataset and more description can be found at <https://www.kaggle.com/sapere0/bitcoin-dataset>)

Data is zipped together and contains 12 input edge files and 12 output edge files of transactions that were mined in the blocks of that year/month.

Each line in the input edge file is tab separated with the format:

1) Unix time of transaction $T\_3$, 2) hash of transaction $T\_3$, 3) hash of transaction $T\_1$ that created the first input, 4) index of output from $T\_1$, 5) (if exists) hash of transaction $T\_2$ that created the second input 6) (if exists) index of output from $T\_2$, 7) (additional inputs, if exist)

Each line in the output edge file is tab separated with the format:

1) Unix time of transaction $T\_3$, 2) first output address $a\_1$, 3) satoshi amount sent to $a\_1$, 4) (if exists) second output address $a\_2$, 5) (if exists) satoshi amount sent to $a\_2$, (additional outputs, if exist)

By using any graph library (e.g., Jung at Java, GraphTools, NetworkX at Python and iGraph at R), create an address graph from transactions of 2011 and find 1) average node degree, 2) the histogram of time an output remains unspent (use block time as a proxy for the transaction time).  In the histogram, x axis is time in seconds. Below are some cases to consider when coding.

\begin{enumerate}

    \item Some transaction inputs have been created in earlier transactions, but the dataset does not contain them. You can omit these inputs from histogram computations in question 2.

    \item Some outputs remain unspent. You can omit these from the histogram as well.

    \item Coinbase transactions are not included in the dataset.

    \item Some outputs are created and spent in the same block. You may take their unspent time as 0 seconds.

\end{enumerate}