Blockchain Transaction Notifications

This dataset ingests real-time notifications of new transactions and blocks on the blockchain via WebSocket API (Blockchain). The API from Blockchain lets us to subscribes to channels of interest for real-time data regarding creation of new transactions and/or blocks occurring via connection to the WebSocket URL: wss://ws.blockchain.info/inv.

- **Volume**: Every minute thousands of transactions are added globally. We can use time-based sampling to ingest the data required for the project or try to filter data fields to get only required fields such as transaction amount, timestamp, etc. Estimation of Volume can be stated after we start ingesting the data.

- **Velocity**: This is a real-time data stream and the rate at which data is generated will be very high as lots of transactions take place. Bitcoin alone handles 3-5 transactions per second.

- **Variety**: The data will be a mix of different types of information such as hashes, timestamps, fees, etc. It will be in a structured format.

- **Veracity**: Since this is blockchain data, it should be reliable as it comes from a decentralized and secure system.

- **Value**: This data can be of a high value as it can be used to identify transaction patterns or network congestion.

I.   References

- Blockchain. (n.d.). *Websocket API*. Retrieved from Blockchain: https://www.blockchain.com/explorer/api/api_websocket

- Oliner, A., & Stearley, J. (2007). What Supercomputers Say: A Study of Five System Logs. *37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, (pp. 575-584). Edinburgh.