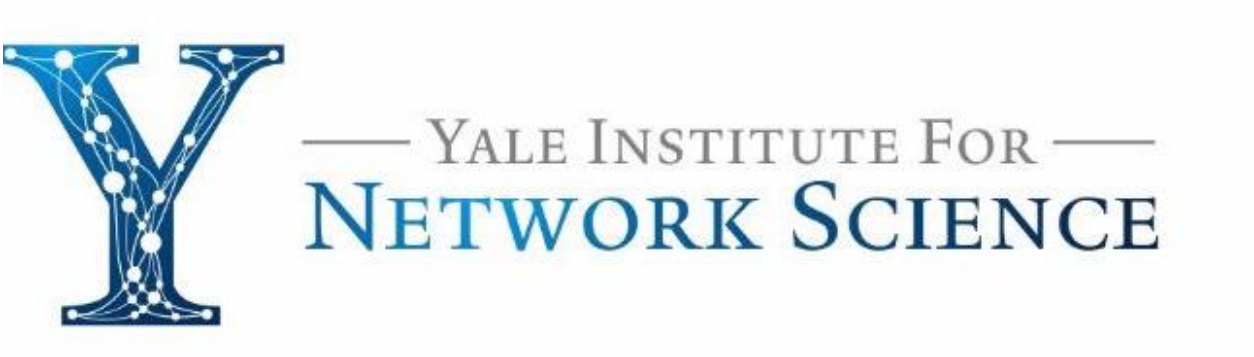


DATAHACK: INDENTIFYING GROUP OF TRADERS THAT MANIPULATE THE FINANCIAL MARKETS

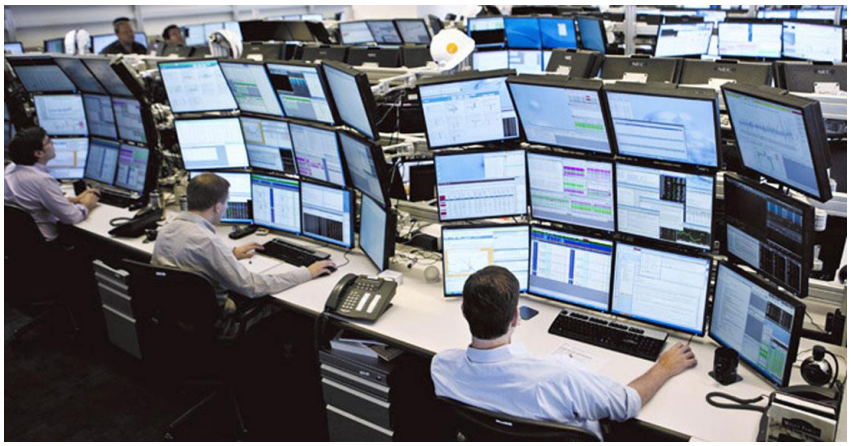
It is a common practice for the Compliance Departments within big financial organizations to set the appropriate controls and procedures to ensure that employees comply with applicable rules and regulations. For example, "Spoofing" is a practice in which traders attempt to give an artificial impression of market conditions by entering and quickly canceling large buy or sell orders onto an exchange, in an attempt to manipulate prices. The 2010 Dodd-Frank Act specifically forbids spoofing. Monitoring and identifying spoofing at an individual trader level is straightforward; however, finding a group of traders that manipulate the market is more involved. One of the mandates of a contemporary compliance officer is to identify potential group spoofing activities, which we define as follows:

Definition: [Potential Group Spoofing Activity (PGSA)]: We say that there is a Potential Group Spoofing Activity when two traders trade the same financial instrument (e.g., a stock) at some timestamp *t* and they communicate (for example, via email or phone) at the same timestamp *t*.

In this datahack, we are going to explore methodologies that identify Potential Group Spoofing Activities within big financial organizations.

Dataset: We are given *trading data* and *communication dat*a - all corresponding to a single day.

1. Trading data: 500 traders trade (buy/sell positions) 1,000 stocks in 100 different timestamps.
2. Communication data: 500 traders communicate with each other in the same 100 timestamps.



Questions:

1. Determine if there is at least one PGSA at each timestamp.
2. Find the timestamp where the fewest PGSAs occur.
3. Find all the PGSAs in this dataset.
4. Find the “riskiest” PGSA and explain your reasoning in detail.