# Auquan

*Capstone Project*
*Prepared by Team Illuminati*
*Members- Vipul Bajaj, Saumya Shah*

## OVERVIEW & PURPOSE

The aim is to implement mean reversion trading strategy for most profitable 20 stocks handpicked from NSE using various data science techniques .

## DATA SCRAPING

1. We scraped data of 1 year(2017) for 110 stocks listed at the NSE  from their respective log files.
2. The features that we extracted include StockVWAP. FutureVWAP, Ask Price, Bid Price, Total traded volume,Total traded size
3. [Stock data](#)

## IDENTIFY STOCKS (PROFITS)

1. We checked the time frames over which stocks were mean-reverting. In particular, we checked for 2hr, 1 day, 5 days and 2 weeks.
2. It was accomplished by calculating their Hurst value over different time frames.
3. We consider a particular stock to be mean-reverting over a certain time frame if its Hurst exponent was roughly below 0.5.
4. Now for each time frame we calculate the mean value(StockVWAP) over 3 times the respective time frames for each stock assuming that to be the mean in the future.
5. Next, we find the deviations of every stock from this mean. Deviation was taken to be StockVWAP-Brokerage(20p for buy+sell)-spread(Top Ask Price-Top Bid Price)-respective_mean.
6. Now we take mean over all the positive deviations and divide it by number of positive deviations(number of trades) to get avg profit per trade. Total profit is the

sum of positive deviations.

7. Now we consider every stock one by one manually looking at their Hurst exponent for respective time frames, avg profit, total profit, num_trades and select 20 which are expected to give good RoC when traded with our strategy.

8. In the above process, we ignored some stocks with abnormally high profits by carefully scrutinizing their distributions and found the data having outliers.

## ARIMA STRATEGY AND THRESHOLDS

1. The ARIMA method was used to predict the future values of the stocks identified above.

2. If the stock was mean reverting in, say, time T, then the future values of the stock were predicted for times in the range 2T - 3T.

3. If the value of the stock was greater than (mean + threshold), then we sell that stock. If it is less than (mean - threshold), then we buy the stock.

4. The mean was calculated as the mean of all the future predicted prices. The threshold was determined by considering different features. In the end, the standard deviation was found to be the most suitable feature.

## ACKNOWLEDGEMENT

We would like to thank Auquan and Chandini Jain Ma'am for providing us with this great learning opportunity by giving  useful and informative resources during the training weeks. Also, we are grateful to them for providing support in every step of the project, through video calls and via emails. And we would like to work with Auquan in the coming future if feasible.