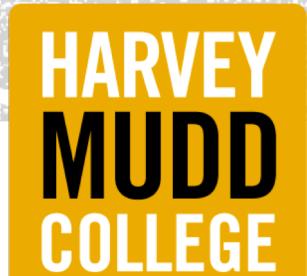
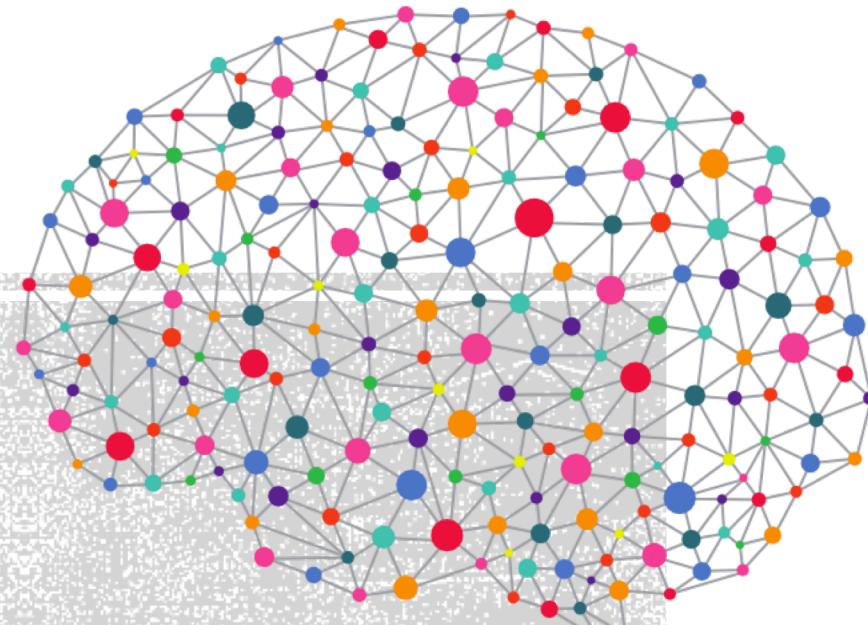


# DNN'S FOR TIMESERIES

Predicting Brexit and the 2016 Presidential Election

By: Avi Thaker

A Lecture for Professor Gary Evans Markets and Modeling Class



# ABOUT ME

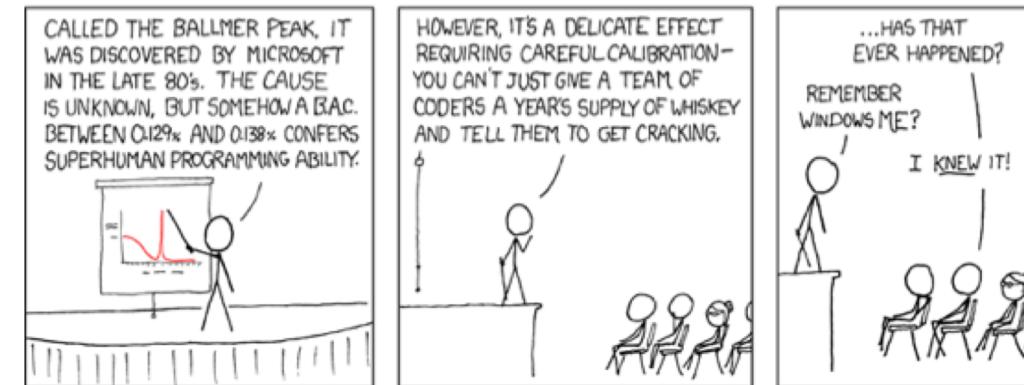
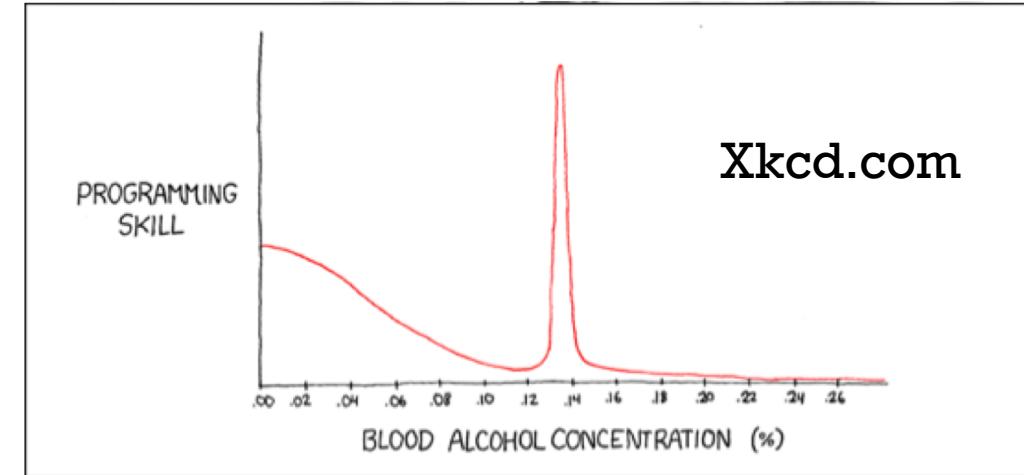
- Data Scientist - Microsoft
  - Deep learning models for customers
  - Implemented a semantic knowledge graph for Ads
- Trader
  - Fully automated trading systems
    - Crypto and Equity Markets
  - Research Driven Process
  - Deep learning models for Audio/Text/Timeseries



avi.thaker@gmail.com



Microsoft



# LECTURE HIGHLIGHTS

- Algorithmic Trading at a Glance
- Introduction to Neural Networks
- Predicting the Trump Election
- Careers in Data Science / AI
- Building your own prop-shop (time permitting)



# ALGORITHMIC TRADING AT A GLANCE

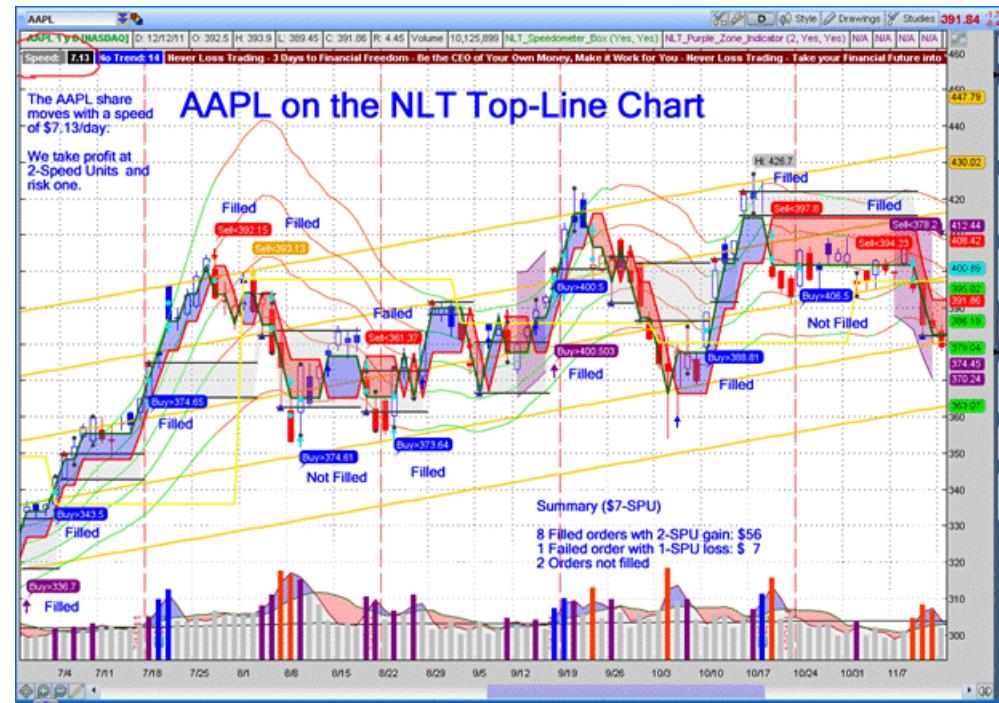
avi.thaker@gmail.com

# THE EFFICIENT MARKET HYPOTHESIS

- Market efficiency causes prices to incorporate and reflect all relevant information
  - Impossible to “beat the market”
- Warren Buffett? Citadel? Virtu?
  - *When filing for its IPO in March 2014, it was disclosed that during five years Virtu Financial made profit 1,277 out of 1,278 days, losing money just one day.*

# WHAT IS ALGORITHMIC TRADING

- The use of electronic platforms for entering trading orders with an algorithm which executes pre-programmed trading instructions whose variables may include timing, price, or quantity of the order.



<http://neverlosstrading.com/Algorithmic%20Trading.html>

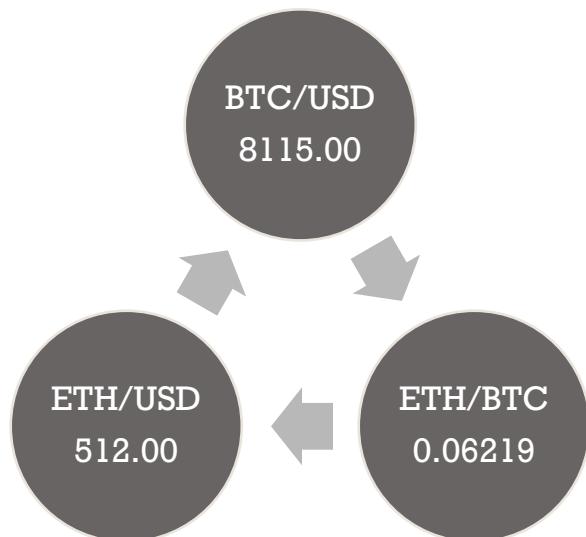
# KINDS OF STRATEGIES

- Arbitrage
- Alpha Taking – Predicting the market
  - Mean reverting
  - Momentum
- Market Making
- Index/Pairs Trading
  - Can be part of any of the above

\*More info can be found in the appendix for all of these kinds of trading strategies

# ARBITRAGE (ARB)

## Inter Exchange Arb (Triangle)



8115 USD -> 1 BTC  
1 BTC -> 16.079 ETH  
16.079 ETH -> 8232.83 USD  
**Net: 117.83 USD**

## Cross Exchange Arb

### Console output Crypto

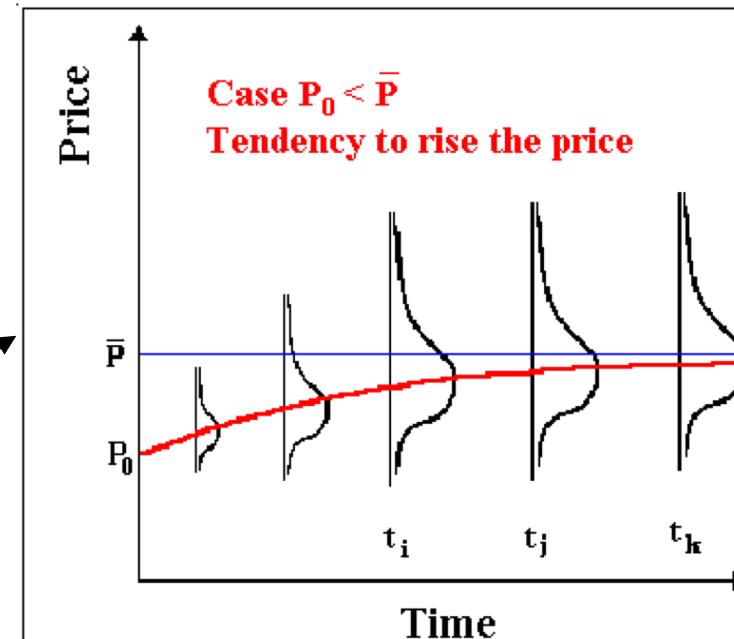
```
[0.0404%, 'BTC/USD', <ccxt.gdax.gdax>, <ccxt.bitstamp.bitstamp>]  
[0.0332%, 'ETH/USD', <ccxt.gdax.gdax>, <ccxt.bitstamp.bitstamp>]
```



Buy BTC on GDAX transfer it to Bitstamp  
Sell BTC for USD on Bitstamp  
**Net: 3.28 USD**

# ALPHA TAKING

- Believes the product will move in one direction
- Mean Reverting – price reverts back to a “average” price, Bollinger Bands
  - Stochastic Math
- Trend following / Momentum
  - Price will move in one direction
  - Impulse response filters



<http://marcoagd.usuarios.rdc.puc-rio.br/revers.html>



Stockcharts.com

# MARKET MAKING

- Increase liquidity – rewarded with the spread
- Can be dangerous in times of volatility

Bids	Price	Asks
	100.03	2,1
	100.02	3,7,8
	100.01	5,2,15
	100	1,2,5
1,2	99.99	
2,5,8	99.98	
3,8,1,5,3	99.97	
2,3	99.85	

- Lots of machine learning, think Bayesian networks, and neural networks
- The focus is on market microstructure

# TECHNICAL INDICATORS

- Technical indicators are mostly **useless** on their own
  - Must identify a market phenomenon and statistically validate it
  - Simplicity and understandability usually are better
- ▶ Accumulation/Distribution (AD)
  - ▶ Average Directional Movement (ADX)
  - ▶ Absolute Price Oscillator (APO)
  - ▶ Aroon Oscillator (ARO)
  - ▶ Average True Range (ATR)
  - ▶ Bollinger Band (BBANDS)
  - ▶ Directional Movement Indicators (DMI)
  - ▶ Exponential (EMA)
  - ▶ Fill Indicator (FILL)
  - ▶ Ichimoku (ICH)
  - ▶ Keltner Channel (KC)
  - ▶ Linear Regression (LR)
  - ▶ Moving Average Convergence Divergence (MACD)
  - ▶ On Balance Volume (OBV)
  - ▶ Relative Strength Indicator (RSI)
  - ▶ Parabolic Sar (SAR)
  - ▶ Simple Moving Average (SMA)
  - ▶ Triple Exponential Moving Average (TEMA)
  - ▶ Triangular Moving Average (TRIMA)
  - ▶ Triple Exponential Moving Average Oscillator (TRIX)
  - ▶ Time Series Forecast (TSF)
  - ▶ TT Cumulative Vol Delta (TT CVD)
  - ▶ Ultimate Oscillator (ULTOSC)
  - ▶ Volume Delta (Vol Δ)
  - ▶ Volume Weighted Average Price (VWAP)
  - ▶ ...
  - ▶ ...



# INTRODUCTION TO NEURAL NETWORKS

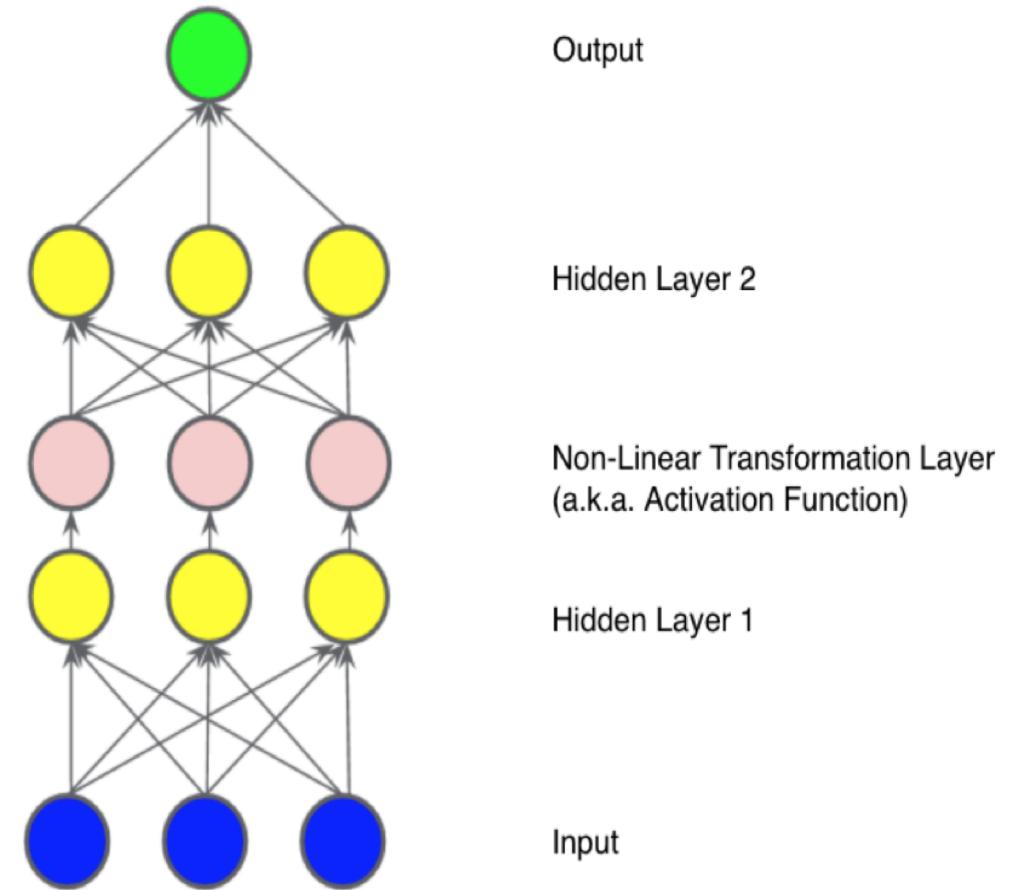
avi.thaker@gmail.com

avi.thaker@gmail.com

<https://www.youtube.com/watch?v=bxe>

# ANATOMY OF A NEURAL NETWORK

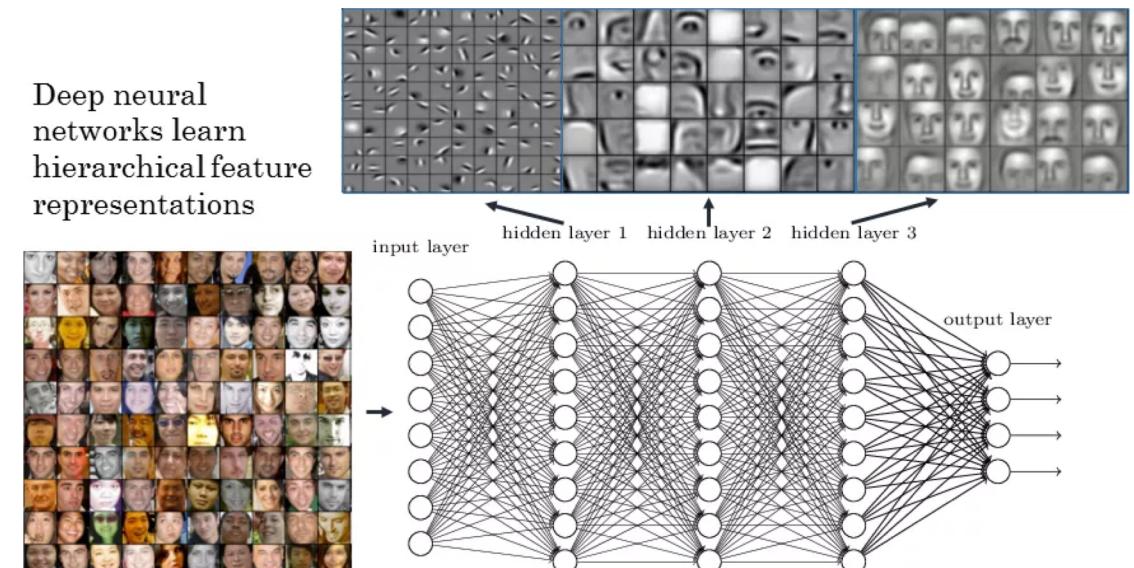
- Set of nodes organized in layers
- Set of weights representing the connections between each layer
- Set of biases one for each node
- Activation function that transforms the output of each node in a layer
- Can be thought of as a black box mapping inputs to outputs



# CLASSICAL ML VS DEEP LEARNING

- Classical ML
  - Code up features to some model
- Deep Learning
  - Model learns the useful features/feature combinations by looking at raw input
- Garbage in -> garbage out
  - Models are highly dependent on data quality no matter the quality of the network architecture

Deep neural networks learn hierarchical feature representations



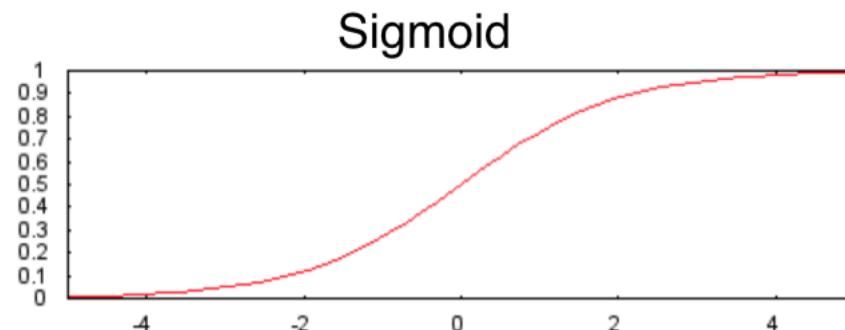
<https://www.strong.io/blog/deep-neural-networks-go-to-the-movies>

# COMMON ACTIVATION FUNCTIONS

Sigmoid

Weighted sum to a value between 0 and 1

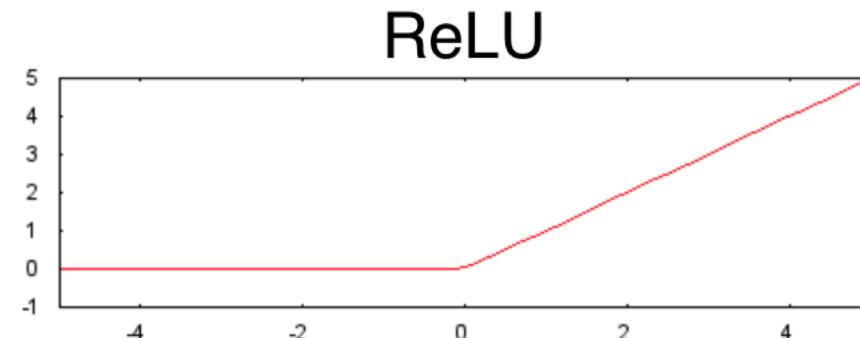
$$F(x) = \frac{1}{1 + e^{-x}}$$



ReLU

Rectified linear unit - much easier to compute

$$F(x) = \max(0, x)$$



# APPLICATIONS OF NEURAL NETWORKS

- Image/Video
  - Classification, Detection, Segmentation
- Text
  - Search, classification, summarization
- Speech
  - Text to speech, speech to text
- Timeseries
  - What will the future value of the stock be?
- Many More
  - Risk analysis, medicine, etc.



<https://www.youtube.com/watch?v=xW2hpkoalIM>



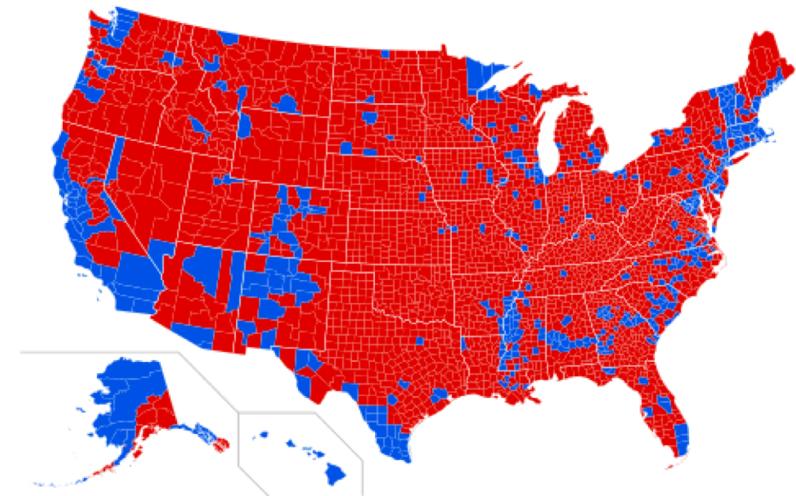
# PREDICTING THE TRUMP ELECTION

This will provide the basics and principles can be applied to predict other events like Brexit. In practice I built a more general example version of big events.

# OVERVIEW

- Use a DNN and features to predict the Trump election and the expected market reaction
- Python utilizing Tensorflow
- [https://github.com/athaker/econ\\_136](https://github.com/athaker/econ_136)
- Identify the sources of data and the resulting outputs
- What features do you think are most important to the model?

Candidate	Party	Electoral Votes	Popular Votes
Donald J. Trump	Republican	304	62,980,160
Hillary R. Clinton	Democratic	227	65,845,063
Gary Johnson	Libertarian	0	4,488,931
Jill Stein	Green	0	1,457,050
Evan McMullin	Independent	0	728,830

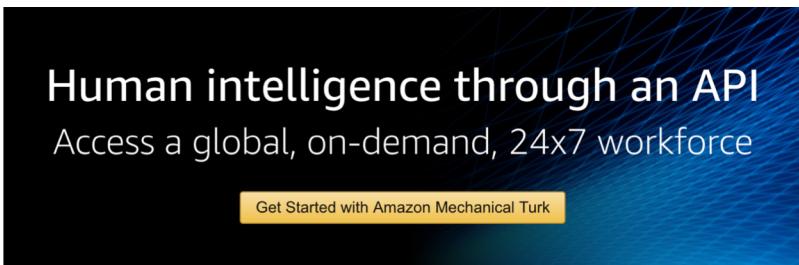


# DATA ENGINEERING

- Arguably the most important thing in Deep/Machine Learning
- High quality labels and features drive model performance
  - Perfect features would be forward looking
- Personally, I look for model insensitivity\*
- Look for model biases
  - Forward looking bias
  - Curve fitting bias
  - Survivorship bias
  - Psychological bias

\*I do not have a formal rigorous approach to this, but I strive for models to get similar results with similar features, i.e. features mean something

# OBTAINING DATA



**Mechanical Turk: Build an API  
for humans to generate your  
labels**

- Better and more data will give you better models

## Google Trends

- League of Legends  
Online game

⋮

- Fortnite  
Survival game

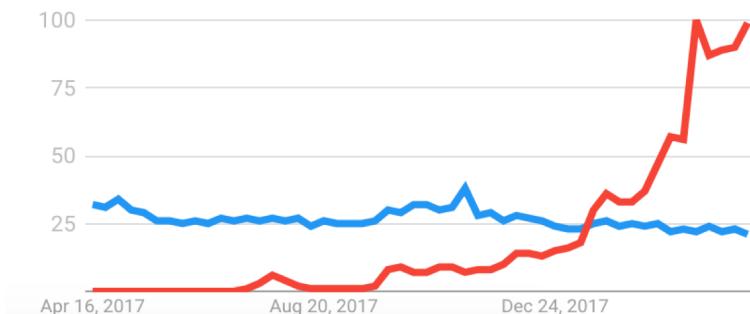
⋮

+ ADD COMPARISON

— Worldwide, Past 12 months

Interest over time

⋮

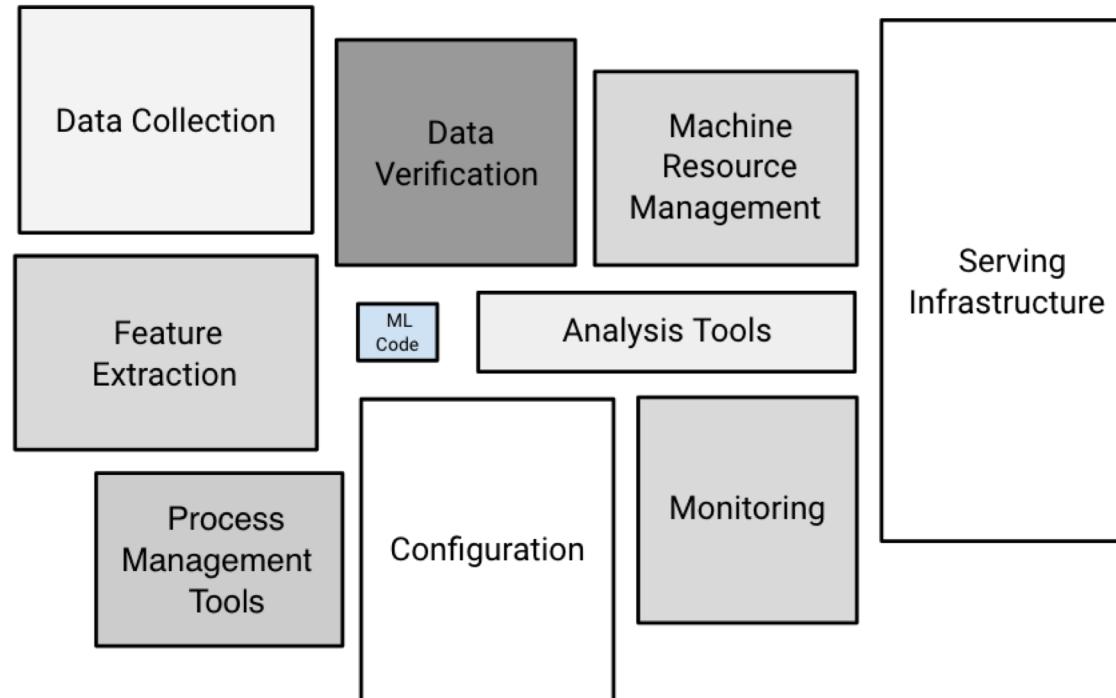




# TENSORFLOW: DNNREGRESSOR

- Tensorflow: DNNRegressor
  - [https://www.tensorflow.org/api\\_docs/python/tf/estimator/DNNRegressor](https://www.tensorflow.org/api_docs/python/tf/estimator/DNNRegressor)
- Model Features:
  - Political Party
  - Previously held an office
  - Previous party to win
  - Was VP
  - Historical market prices (Dow Jones Industrial Average)
- Predicting
  - Political Party
  - Market price (Dow Jones Industrial Average)

# ML/DL IS ONLY A SMALL FRACTION OF THE WORK



<https://developers.google.com/machine-learning/crash-course>



# CAREERS IN DATA SCIENCE / AI

What can I do with this new found knowledge?

# HOT TOPICS

- Finance hires “quants”
- This is the hot topic right now, tech, finance, ... all sectors hiring like crazy
- Smart interaction with the world
  - Smart Cars, Search
- 3rd Party Data
  - Satellite imagery parsing for hedge funds, quant focused approach
- Big name companies say it is their focus
  - Google, Amazon, Facebook, Microsoft
  - Microsoft “AI First” company now
- Think about and be careful what you build



A breakthrough in machine learning would be worth ten Microsofts.

— Bill Gates —

AZ QUOTES



Google will fulfill its mission only when its search engine is AI-complete. You guys know what that means? That's artificial intelligence.

— Larry Page —

AZ QUOTES



“WITH ARTIFICIAL INTELLIGENCE WE ARE SUMMONING THE DEMON.”  
-ELON MUSK

HUFF POST



MOODY COLLEGE

# IMPORTANT TO KNOW

- Math
  - Linear Algebra
  - Statistics
- Programming
  - Python - Most used and developed in language
  - Tensorflow
  - Keras
  - Scikit Learn
  - C/C++
  - CUDA
- If path in Finance
  - Start trading and get a track record!
  - Math/Computer Science majors tend to be valuable over Economics/Business
- Read Papers and Blogs!

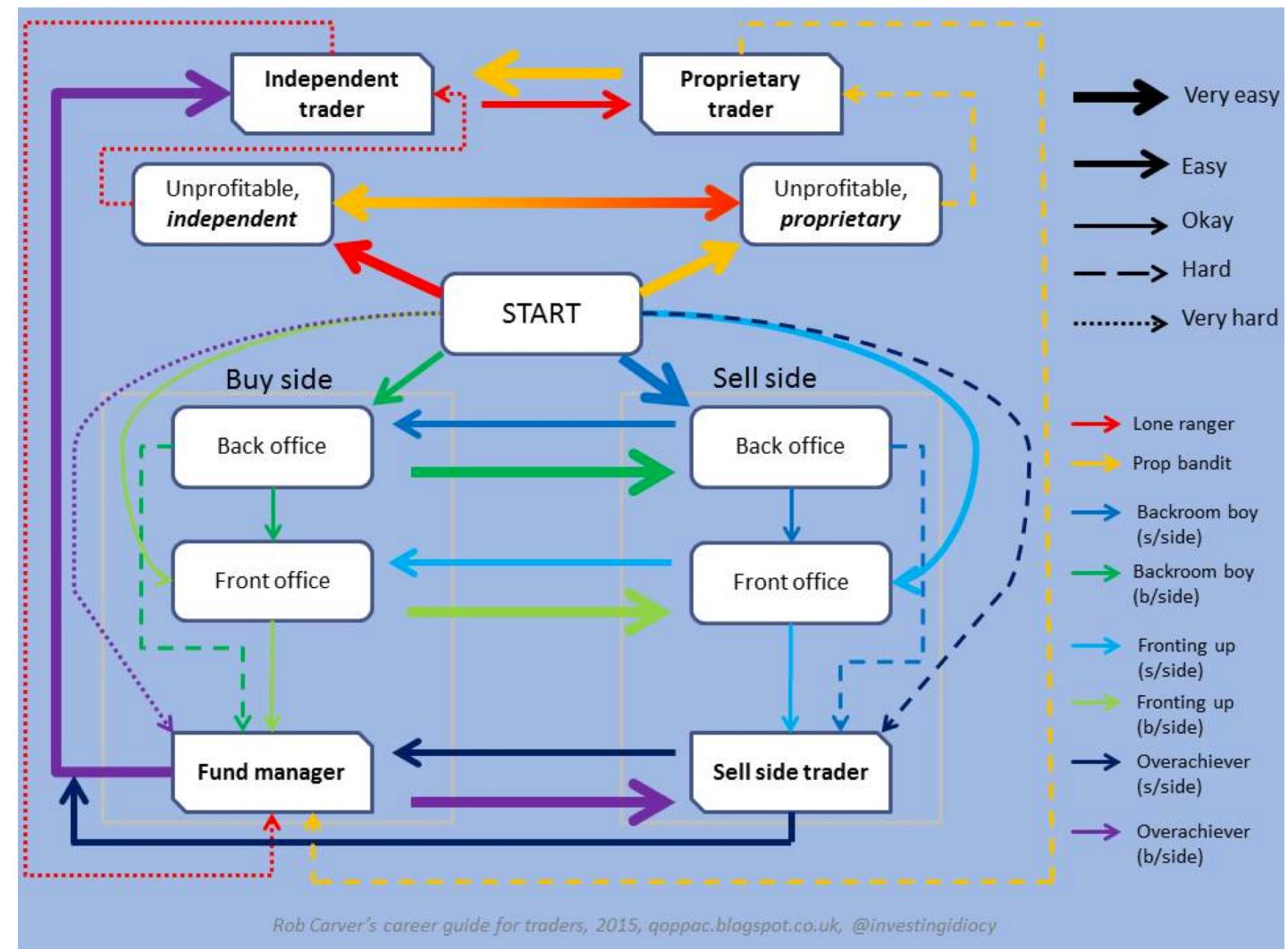


# GETTING STARTED: PROP SHOP / FUND

avi.thaker@gmail.com

# A VISUAL GUIDE

- Get started trading!
  - Start small
- Prove you can make money
- Higher Sharpe Ratio is more valuable
- Quantitative skills right now are incredibly important



# ONLINE RESOURCES: DATA SCIENCE

- **Golden Standard - Andrew NG (Stanford)**
  - <https://www.coursera.org/learn/machine-learning>
  - <https://www.coursera.org/specializations/deep-learning>
- **Father of Deep Learning**
  - <https://www.coursera.org/learn/neural-networks> (hardest)
- **Google**
  - <https://developers.google.com/machine-learning/crash-course/> (easiest)
  - <https://www.udacity.com/course/deep-learning--ud730>
- **Quantstart**
  - <https://www.quantstart.com/>
- **Wikipedia**
  - [https://en.wikipedia.org/wiki/Deep\\_learning](https://en.wikipedia.org/wiki/Deep_learning)

# ONLINE RESOURCES: TRADING

- Golden Standard
  - <http://www.quantstart.com>
- Execution Environment/Backtester/Community
  - <https://www.quantopian.com>
- Suggested intro trading platform / broker
  - [www.interactivebrokers.com](http://www.interactivebrokers.com)
- Data:
  - Yahoo Finance, Google Finance, Quandl
- Other:
  - <https://qoppac.blogspot.com/>
  - <http://epchan.blogspot.com/>



# QUESTIONS?

Ask me about:

Trading? DNNs? Crypto Currency? Crypto Mining? Life at Microsoft?



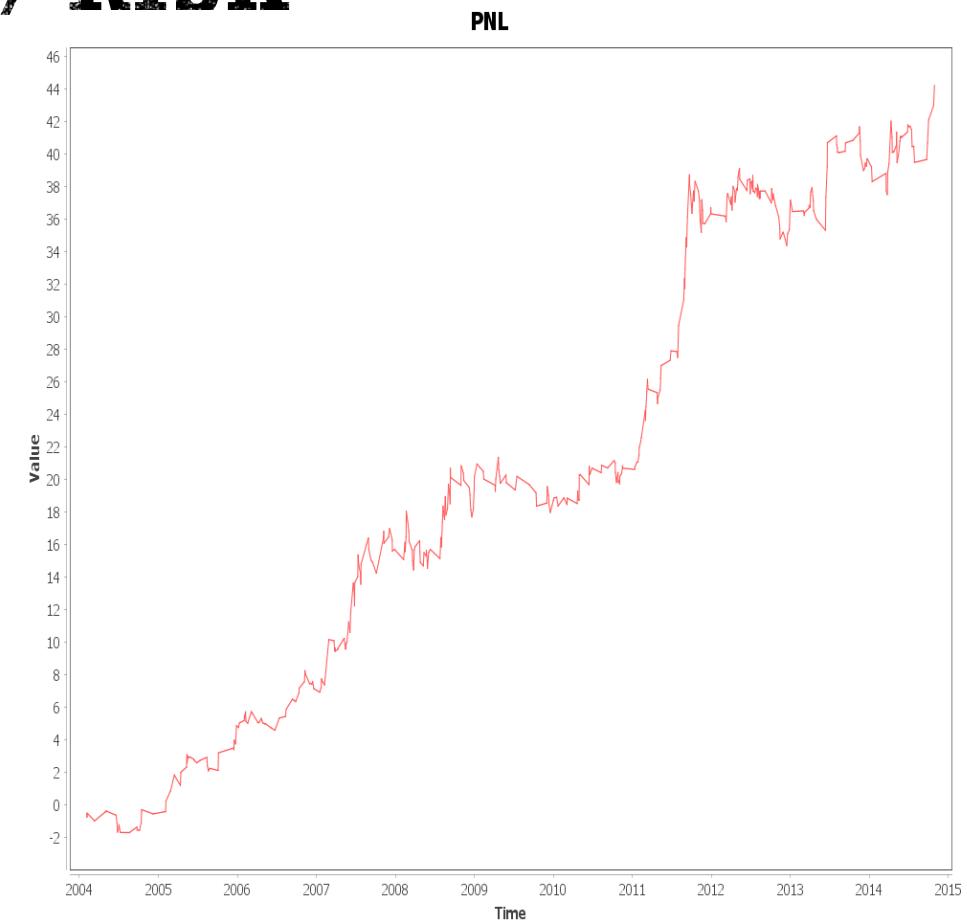
# APPENDIX: PERSONAL LEARNINGS

# BACKTESTING A STRATEGY/RISK

- Provide evidence of profitability
  - Curve fitting/ optimization bias
  - In-sample vs out-of-sample
  - Forward looking bias
- Risk tolerance

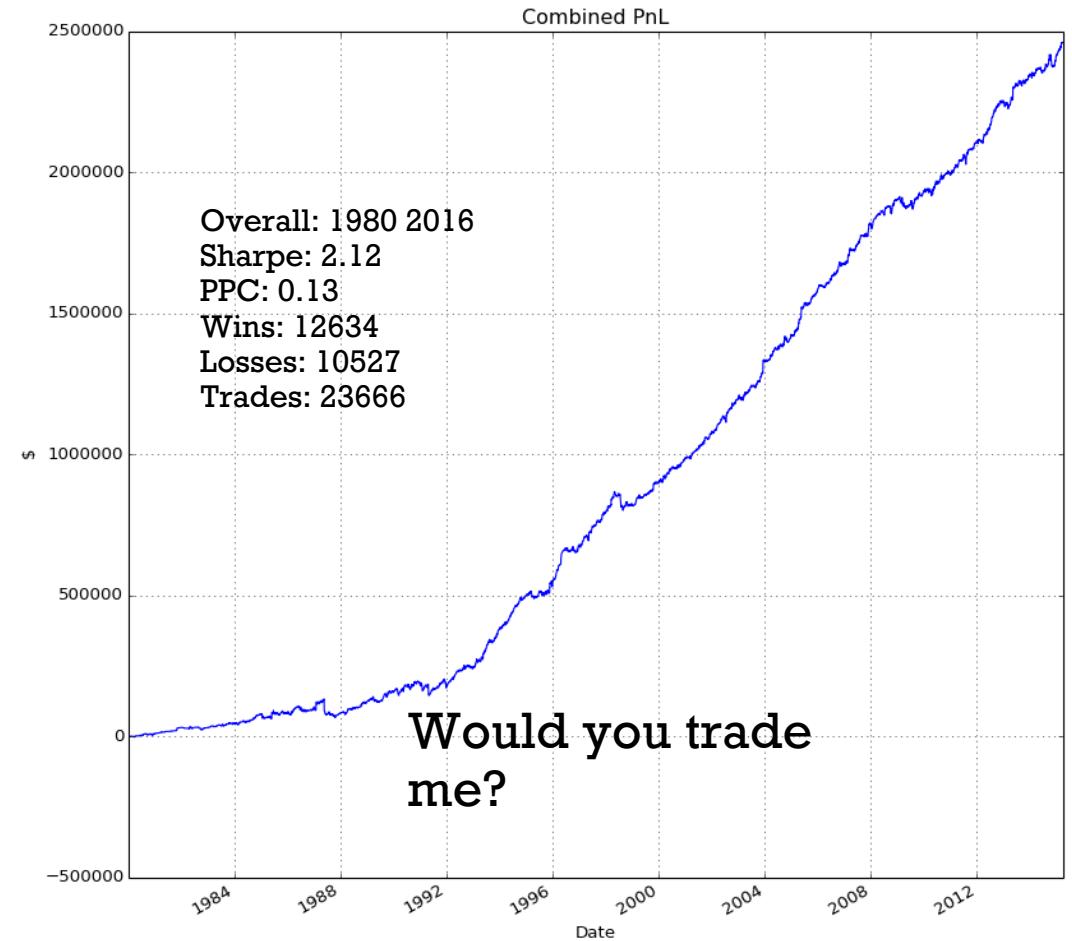
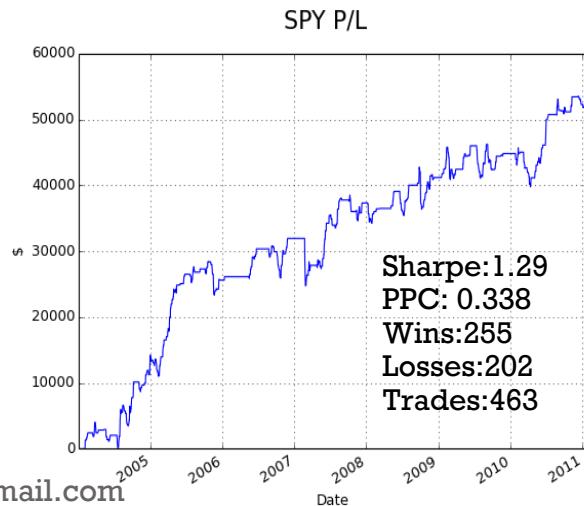
## Key Statistics

Average wins	:: 0.637 USD
Average loss	:: -0.438 USD
# Wins	:: 214
# Losses	:: 210
# Neutrals	:: 3
Win Rate	:: 0.501
PPC	:: 0.104USD
# Traded	:: 427.0
Ann. Sharpe	:: 2.335



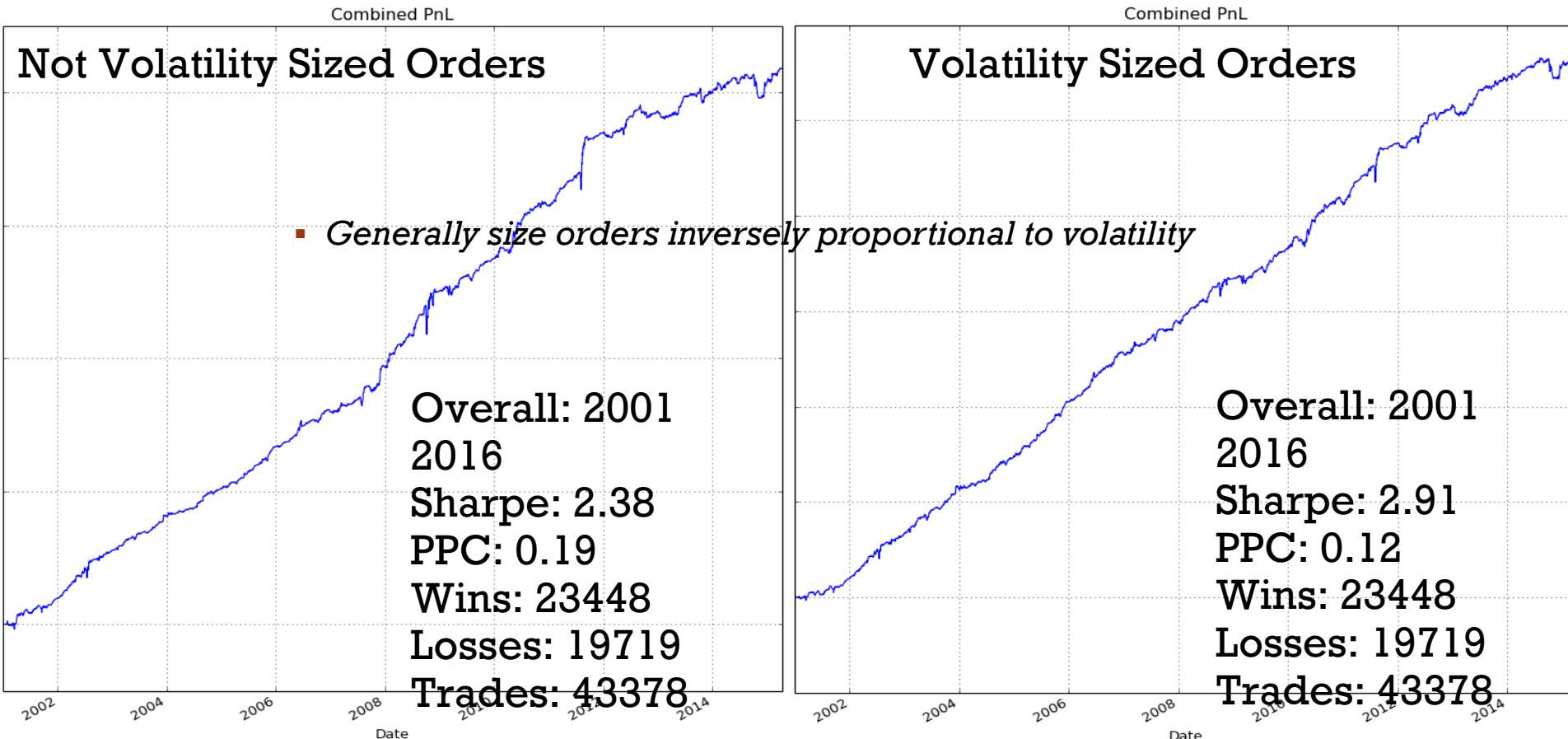
# BACKTESTING

- Does the strategy work across many assets?
- How many years does it work for?
- Does it escape the bid-ask bounce?
- Risk Tolerance?
  - Maximum Drawdown?
- Fees? Trading frequency?



In Sample: SPY 2004-2010  
Out of Sample: Assets Randomly Selected:  
ADBE XLNX BBBY CFN EMC ADP AFL DE T SPLS DG ADS ALL MET CL PX  
WYN

# ORDER SIZING



# GENERAL TIPS

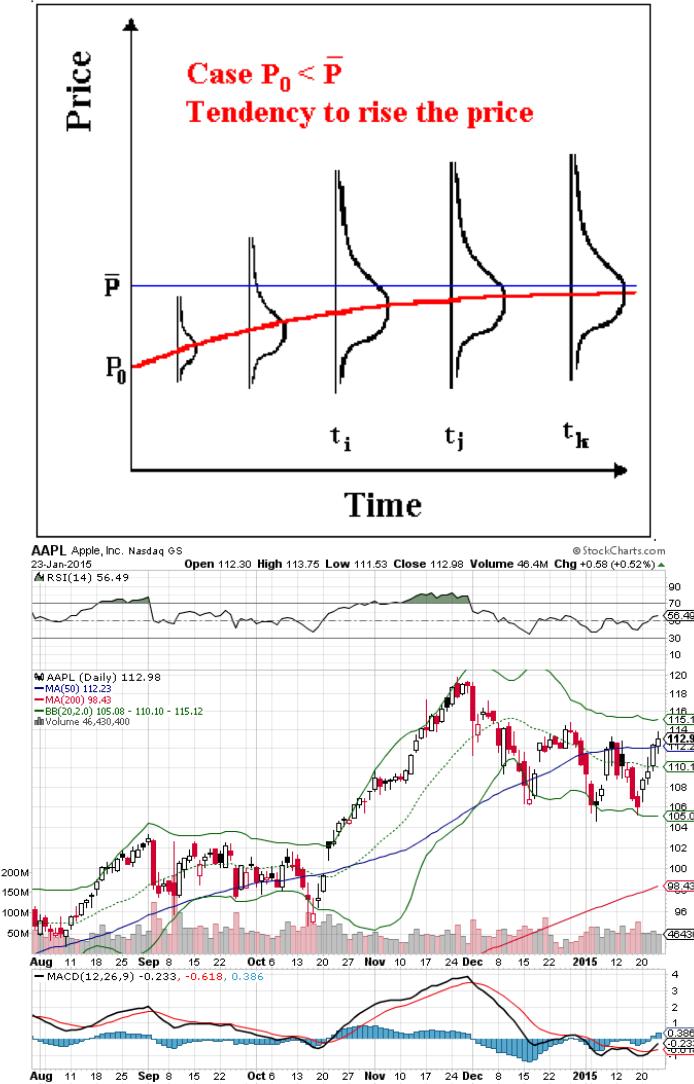
- This is not a get rich quick scheme
- Finding alpha is hard, do not get discouraged
- Drawdown are painful, be careful with leverage
- Trust your alpha (if you have some), strategies are usually simple
- Performance
  - Out of sample performance is generally  $\frac{1}{2}$  of in sample performance
  - Live trading performance is generally  $\frac{1}{4}$  of in sample performance
  - Due to curve fitting, unexpected slippage, etc.
- Make sure you account for ***transaction fees*** and ***slippage*** and ***order sizes***
- Fun and exciting way to learn not only the markets but also computer science and math
- Data is your friend
- Build your own backtester/execution environment

# APPENDIX: MEAN REVERSION AND MOMENTUM

This section covers the details about statistical testing and a detailed look into how to test classically for a mean reverting or trending (momentum) timeseries.

# STATISTICAL TESTING

- Mean Reversion
  - A process that refers to a time series that displays a tendency to revert to its historical mean
  - More specifically: if the prices within the series move away from their initial value faster than that of Geometric Brownian Motion
  - Ornstein-Uhlenbeck process (a random walk has no memory)
- Momentum
  - The exact opposite of mean reversion
  - Movement away from the initial value faster than that of random walk
- Mean reversion and momentum go hand in hand, in identifying one you may identify the other
- Will cover two methods: Augmented Dickey-Fuller test, and the Hurst Exponent



Pictures From:

<http://marcoagd.usuarios.rdc.puc-rio.br/revers.html>

<http://www.stockcharts.com>

Avi.thaker@gmail.com

# STATISTICAL TESTING: TERMS

- Orenstein-Uhlenbeck SDE
- Change in price series in next time period is proportional to the difference between the mean price and the current price with Gaussian noise
- Motivates Augmented Dickey-Fuller (ADF) Test

$$dx_t = \theta(\mu - x_t)dt + \sigma dW_t$$

$\theta$  = rate of reversion to mean  
 $\mu$  = mean value of process  
 $\sigma$  = variance of the process  
 $W_t$  = Wiener Process or Brownian Motion

# AUGMENTED Dickey-Fuller (ADF) TEST

- Identify presence of a unit root in autoregressive time series
- Relies on the fact that if a price series has a mean reversion then the next price will be proportional to the current price
- Linear Model of Order  $p$

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

$\alpha = \text{constant}$

$\beta = \text{coefficient of time trend (long term drift)}$

$$\Delta y_t = y(t) - y(t - 1)$$

- Testing null hypothesis:  $\gamma = 0$ 
  - Indicates that process is a random walk ( $\alpha = \beta = 0$ )

# AUGMENTED DICKEY-FULLER (ADF) TEST

- Test statistic: sample proportionality / standard error of sample proportionality

$$DF_{\tau} = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

- Negative number, and must be less than critical values to be significant
- Code `adf_test.py`

Calculated Test Statistic: -2.1900105031287529

P-Value: 0.2098910250427564

# Datapoints: 2106  
10%: -2.5675011176676956  
5%: -2.8629133710702983  
1%: -3.4334588739173006

Cannot reject null hypothesis,  
and unlikely to have found a  
mean reverting time series

# HURST EXPONENT

- A stochastic process is strongly stationary if its joint probability distribution is invariant under translations in time or space
  - Mean and variance of process do not change over time and do not follow a trend
- Hurst Exponent helps to characterize the stationarity of a time series
  - Reverting, trending, or neither
- Variance of a log price series to identify rate of diffusive behavior
$$Var(\tau) = \langle |\log(t + \tau) - \log(t)|^2 \rangle$$
- Since large  $\tau$ , variance is proportional to  $\tau$  for Geometric Brownian Motion
$$\tau \sim \langle |\log(t + \tau) - \log(t)|^2 \rangle$$
- If autocorrelations exist the relationship is not valid, but can be modified to include  $2H$  with the Hurst Exponent value  $H$ 
$$\tau^{2H} \sim \langle |\log(t + \tau) - \log(t)|^2 \rangle$$

# HURST EXPONENT: MEANING

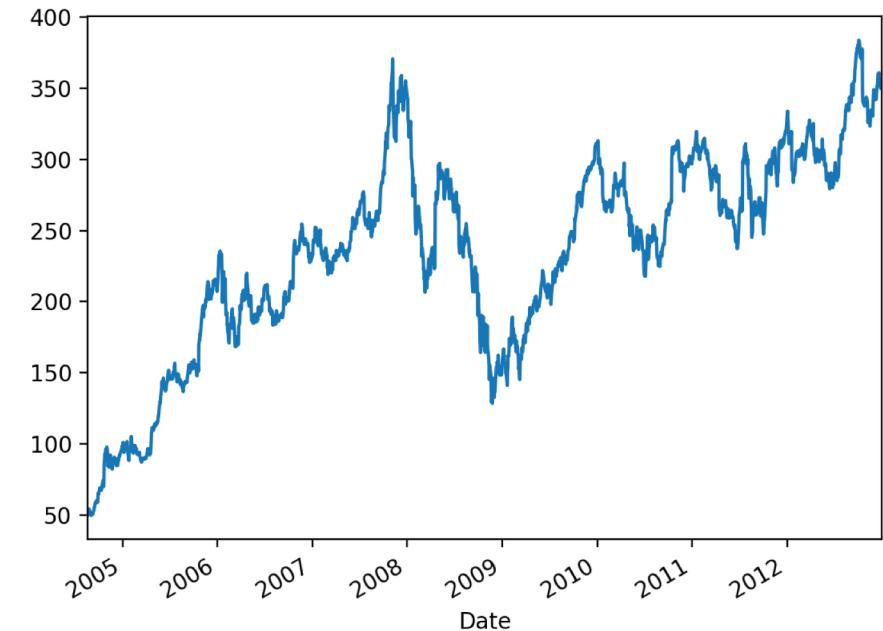
$H < 0.5$  mean reverting process

$H == 0.5$  GBM

$H > 0.5$  trending process

- Characterizes extent
  - Closer to 0 more mean reverting
  - Closer to 1 more trending
- Try different time periods, different stocks

Avi.thaker@gmail.com



Hurst(GBM): 0.498349157279

Hurst(MR): -

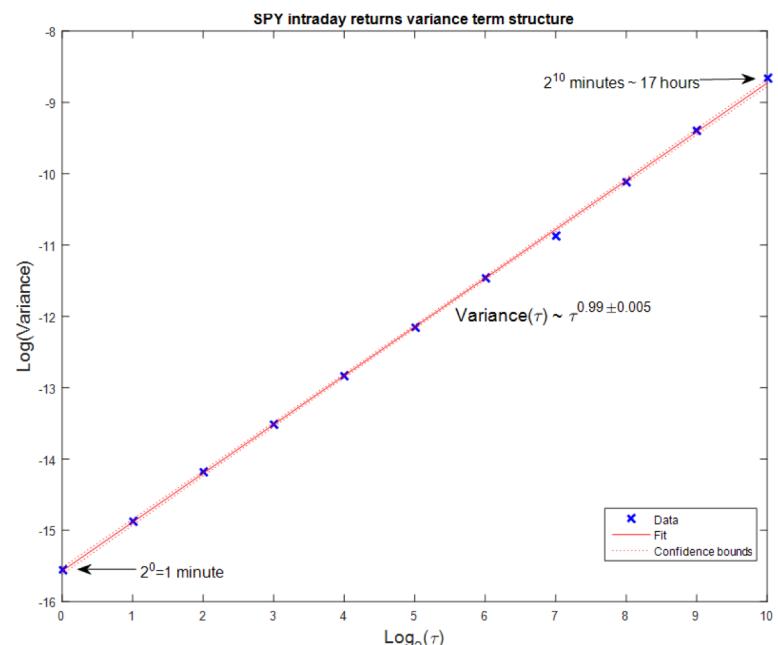
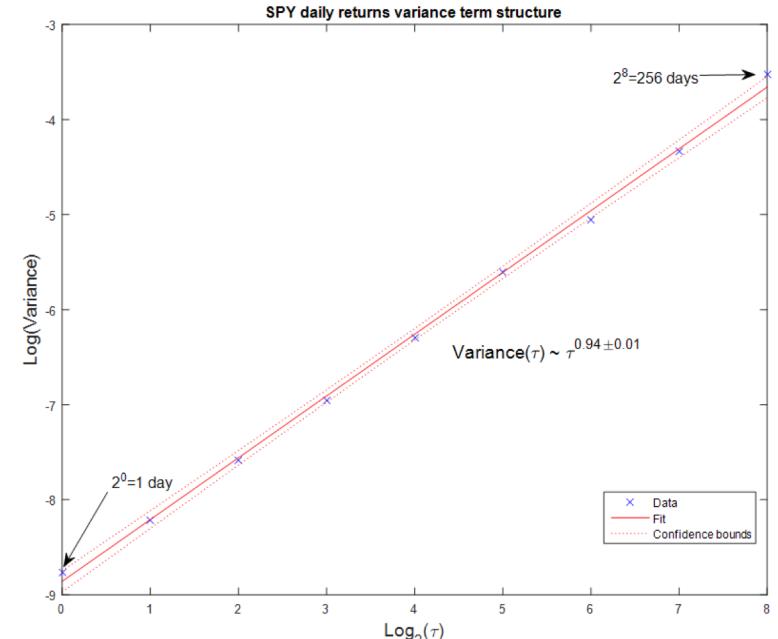
6.26637088795e-05

Hurst(TR): 0.95964231812

Hurst(GOOG): 0.50788012279

# VARIANCE AND TERM STRUCTURE

- Not included in code
- Plot of  $\log(Var(\tau))$  vs  $\log(\tau)$  for SPY
  - Slope/2 is the Hurst exponent
  - Intraday
    - Returns of mid-prices from 1 minute to  $2^{10}$  minutes
    - $H = 0.494 \pm 0.003$ ; slightly mean reverting
  - Daily
    - Returns from 1 day to  $2^8$  days
    - $H = 0.469 \pm 0.007$ ; strongly mean reverting
- Mean reversion strategies should work better than intraday strategies on SPY

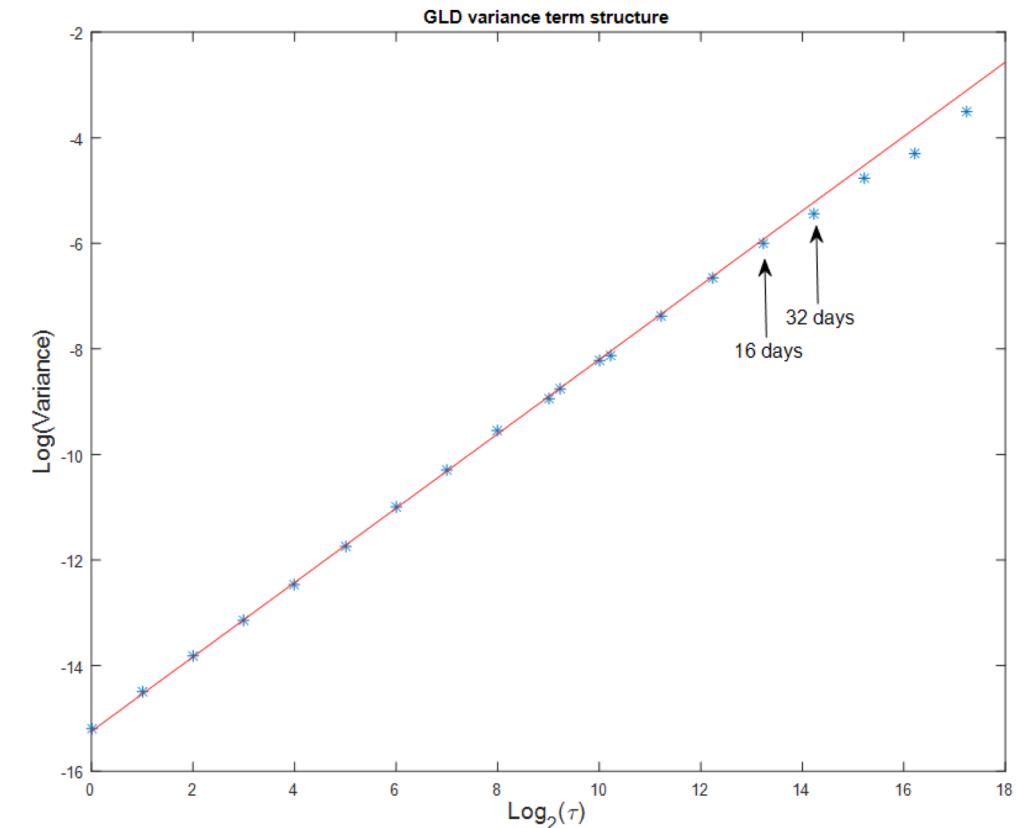


# VARIANCE AND TERM STRUCTURE: GOLD

- Intraday:  $H = 0.505 \pm 0.002$
- Daily:  $H = 0.469 \pm 0.007$
- 16-32 days volatilities drift from the regression
  - ***This is where we should switch from momentum to mean reversion strategies***

A Trending Example: USO  
Intraday  $H = 0.515 \pm 0.001$   
Daily  $H = 0.560 \pm 0.020$

*Momentum strategies should work well here*



<http://epchan.blogspot.com/2016/04/mean-reversion-momentum-and-volatility.html>



# APPENDIX: KINDS OF STRATEGIES

A Brief Introduction

# HFT VS NON HFT

- HFT or High Frequency Trading, relies on quick algorithmic trading to capture small spreads or small alpha (less than 1 minute holds)
  - Difficult for the retail investor to get into
  - Extremely competitive
    - Fastest usually takes all profits
    - Makes up 70%+ of all trading volume
- Non HFT (trades of longer periods)
  - Easier for the retail investor to get into
  - Not as competitive
    - Smarter people make profits
  - This is our focus

# STRATEGY: PAIR TRADING

- Assumes two contracts move together
- Revert moves when one moves against the other
- Can be highly risky
  - 2007 Crash?
  - Spreads blow up
- Commonly used by Hedgefunds
- High usage of Kalman Filters, CADF (Co-integrated Augmented Dickey Fuller) tests, autocorrelation, etc.



<http://www.investopedia.com/university/guide-pairs-trading/pairs-trade-example.asp>

# STRATEGY DETAIL: COMMON SIGNALS

## SIMPLE MOVING AVERAGE: SMA

- Mean of previous  $n$  periods
- Smooths data, easier to identify trends
- Common signal

$$SMA = \frac{p_M + p_{M-1} + \dots + p_{M-(n-1)}}{n}$$

$$SMA_{current} = SMA_{previous} - \frac{p_{M-n}}{n} + \frac{p_M}{n}$$

$p_i$  = price at given time



# STRATEGY DETAIL: COMMON SIGNALS

## EXPONENTIAL MOVING AVERAGE: EMA

- Infinite impulse response filter
- Less lag than SMA
- Commonly used signal

$$\alpha = \frac{2}{n + 1}$$

$$\begin{aligned} EMA_{current} &= \frac{p_1 + (1 - \alpha)p_2 + (1 - \alpha)^2 + \dots}{1 + (1 - \alpha) + (1 - \alpha)^2 + \dots} \\ &= EMA_{previous} \\ &\quad + \alpha(p_{current} - EMA_{previous}) \end{aligned}$$



# TRADING THE EMA



Enter Long: Close > EMA & Prev\_Close > EMA\_Prev

Enter Short: Close < EMA & Prev\_Close < EMA\_Prev

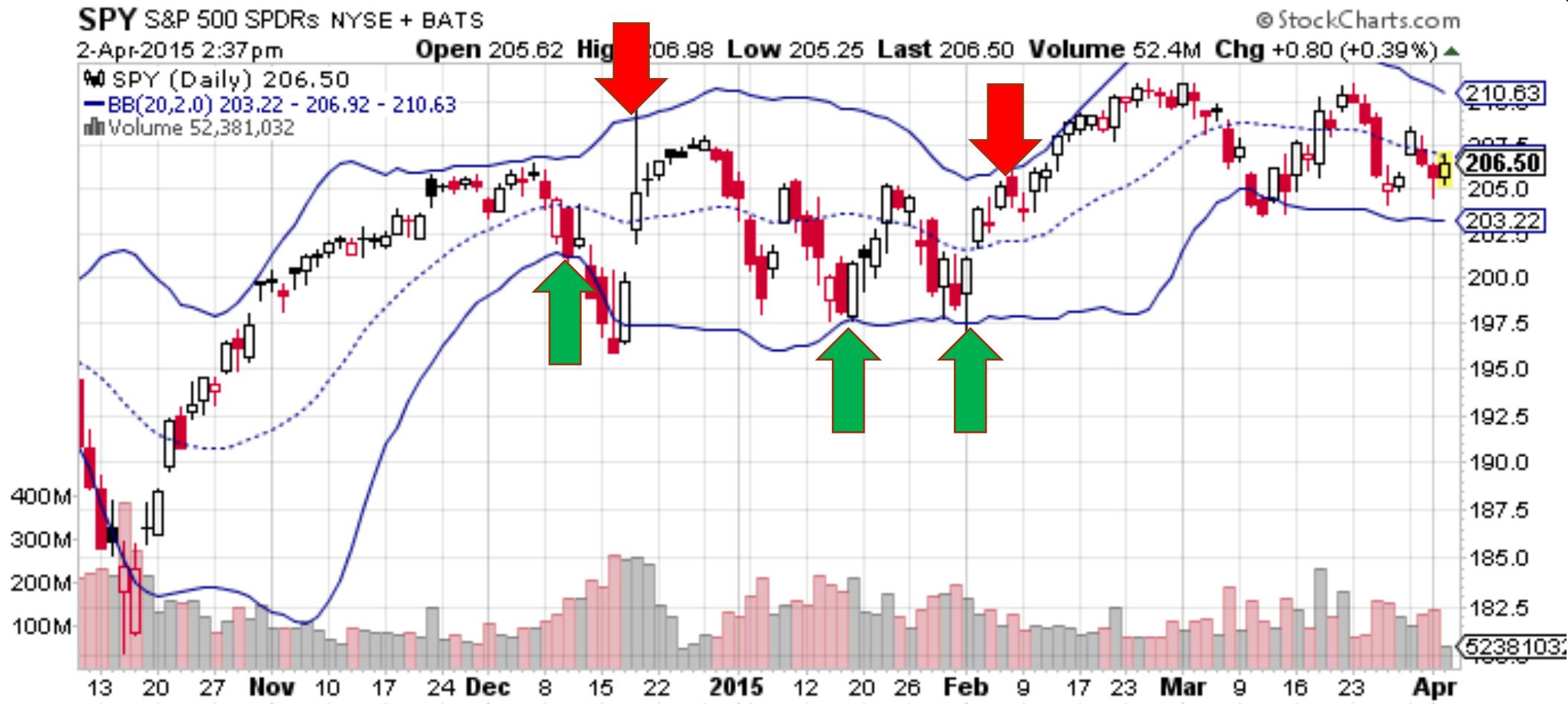
# STRATEGY DETAIL: COMMON SIGNALS

## BOLLINGER BANDS

- Volatility Bands
  - Based upon standard deviation
  - Identifies points of reversion
- Middle Band = 50-Day SMA
- Upper Band = 50-Day SMA + 50-Day SD of Price
- Lower Band = 50-Day SMA - 50-Day SD of Price



# EXAMPLE: TRADING THE BOLLINGER BANDS



# ETF CREATION/REDEMPTION

- Authorized participant (market maker, institutional investor, specialist) borrows stock shares and places them in a trust to form ETF **creation units** – bundles of stock units
- Trust provides shares to the AP, and shares sold to public on open market
- Redeeming ETF
  - Sell shares on open market
  - Form a creation unit and exchange for underlying security
    - Tax efficient

Creation Unit	Last Trade Bid	Ask	Size Net	Percentage
AMD	13.7	13.69	13.7 100	1370 16.31%
INTC	35.16	35.16	35.17 100	3516 41.85%
AAPL	140.64	140.63	140.64 25	3516 41.85%
<b>Creation Unit</b>	<b>84.02</b>		<b>100 8402</b>	<b>100.00%</b>

# ETF/INDEX ARBITRAGE

Unit	Last Trade	Bid	Ask	Size	Net	Percentage
AMD	13.7	13.69	13.7	100	1370	16.31%
INTC	35.16	35.16	35.17	100	3516	41.85%
AAPL	140.64	140.63	140.64	25	3516	41.85%
Creation Unit	84.02	84.01	84.03	100	8402	100.00%

- This is the most common strategy employed by most quant firms and banks
  - Jane Street, AQR, Jump Trading
- You need to be **fast** or have **flow**
  - Some of you may have done this in the game
- Arbitrage happens when ETF trades at a discount or premium to the NAV
  - **Institutional:** When ETF price > NAV, the AP will sell shares it received during creation and make a spread between the cost of the assets it bought for the ETF issuer and the selling price from the ETF shares. AP can also buy the underlying shares that compose the ETF directly at lower prices, sell ETF shares on the open market at the higher price, capturing the spread.
  - **Individuals:** When the ETF is selling at a premium (or discount), individuals can buy (short) the underlying securities in the same proportions and short (or buy) the ETF. Limited by liquidity and spread
    - If inside the spread need to know if the ETF goes to share price or share price goes to ETF price
- Do this at an international level with ADR's

Avi.thaker@gmail.com

Example	Price
Calculated Ask	84.03
Calculated Bid	84.0075
ETF Bid	84.04
ETF Ask	84.05

What is the potential profit of this trade?

# HFT AND FLOW

- Pure Arbitrage
  - Fastest always wins
- Deal Flow
  - Orders executed on behalf of another client
  - E\*TRADE – guaranteed 2 second execution market
  - Smart Orders
- Example Companies
  - Citadel, market makers, Goldman, JP Morgan, etc.



# DEAL FLOW: THE RUSSELL REBALANCE

- Bank will trade all of the positions on behalf of FTSE Russell (moves ~20 Billion in a few hours) for a single client
  - Buys up in anticipation of the trade and sells their own shares to the client
  - Massive market moves
- Legalized insider trading and market manipulation due to sheer size of orders
- Goldman actually pays Russell (and the like) for their order flow!

## Winners and losers since 2015 Russell rebalance

Market capitalization and stock price percentage changes from June 26, 2015 to June 17, 2016

### TOP MARKET CAP INCREASES



### TOP MARKET CAP DECLINES



Source: Thomson Reuters

Staff, 19/06/2016

