HUDSON
AND THAMES

# Hackathon

## Data Driven Investments

# Overview

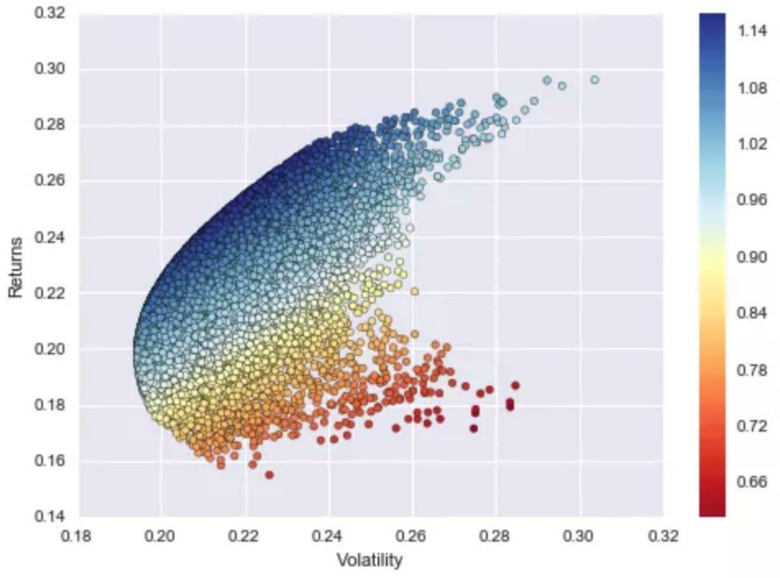
Hudson & Thames is in a strong position to leverage its talent and infrastructure to create a data driven investment product. The purpose of this presentation is to introduce advanced portfolio optimisation algorithms, that work out of sample (de Prado 2018).

This presentation was developed for the Capitec Bank Hackathon at which we competed.

# Financial Machine Learning

Machine Learning is only beginning to transform Finance:
– 2016: Studies show that ML methods (like Hierarchical Risk Parity) deliver portfolios that systematically outperform Markowitz optimization out-of-sample.
– 2016: The GIS-Liquid Strategies group manages $13 billion with 12 people.
– 2017: Four funds of Man/AHL manage $12.3 billion using AI.
– 2018: KPMG's report argues that hedge funds must embrace technology or face 'treadmill to oblivion'.
– 2018: First graduate-level textbook on ML, specifically applied to Finance (Advances in Financial Machine Learning)

# Portfolio Optimization: Mean Variance




Searching for the efficient frontier (Python for Finance, 2017)

- Mathematical framework for assembling a portfolio of assets such that the expected return is maximized for a given level of risk.
- Maximise returns
- Reduce variance of returns
- Harry Markowitz (1952)

HUDSON AND THAMES

# New Optimization: Hierarchical Risk Parity

## Building Diversified Portfolios that Outperform Out of Sample

Marcos López de Prado

Marcos López de Prado
is a senior managing director at Guggenheim Partners in New York, NY, and a research fellow at the Lawrence Berkeley National Laboratory in Berkeley, CA.
lopezdeprado@lbl.gov

Portfolio construction is perhaps the most recurrent financial problem. On a daily basis, investment managers must build portfolios that incorporate their views and forecasts on risks and returns. Before Markowitz earned his Ph.D. in 1954, he left academia to work for the RAND Corporation, where he developed the Critical Line Algorithm (CLA), a quadratic optimization procedure specifically designed for inequality-constrained portfolio optimization problems. This algorithm is notable in that it guarantees finding the exact solution after a known number of iterations—and it ingeniously circumvents the Karush-Kuhn-Tucker conditions (Kuhn and Tucker [1952]). A description and open-source implementation of this algorithm can be found in Bailey and López de Prado [2013]. Surprisingly, most financial practitioners still seem unaware of CLA, as they often rely on generic-purpose quadratic programming methods that do not guarantee the correct solution or a stopping time.

Despite of the brilliance of Markowitz's theory, CLA solutions are somewhat unreliable because of a number of practical problems. A major caveat is that small deviations in the forecasted returns cause CLA to produce very different portfolios (Michaud [1998]). Given that returns can rarely be forecasted with sufficient accuracy, many authors have opted to drop them altogether and focus on the covariance matrix. This has led to risk-based asset allocation approaches, of which "risk parity" is a prominent example (Jurczenko [2015]). Dropping the forecasts on returns improves the instability issues; however, it does not prevent them, because quadratic programming methods require the inversion of a positive-definite covariance matrix (all eigenvalues must be positive). This inversion is prone to large errors when the covariance matrix is numerically ill-conditioned—that is, it has a high condition number (Bailey and López de Prado [2012]).

### MARKOWITZ'S CURSE

The condition number of a covariance, correlation (or normal, thus diagonalizable) matrix is the absolute value of the ratio between its maximal and minimal (by moduli) eigenvalues. Exhibit 1 plots the sorted eigenvalues of several correlation matrices, where the condition number is the ratio between the first and last values of each line. This number is lowest for a diagonal correlation matrix, which is its own inverse. As we add correlated (multicollinear) investments, the condition number grows. At some point, the condition number is so high that numerical errors make the inverse matrix too unstable: A small change on any entry will lead to a very different inverse. This is *Markowitz's*

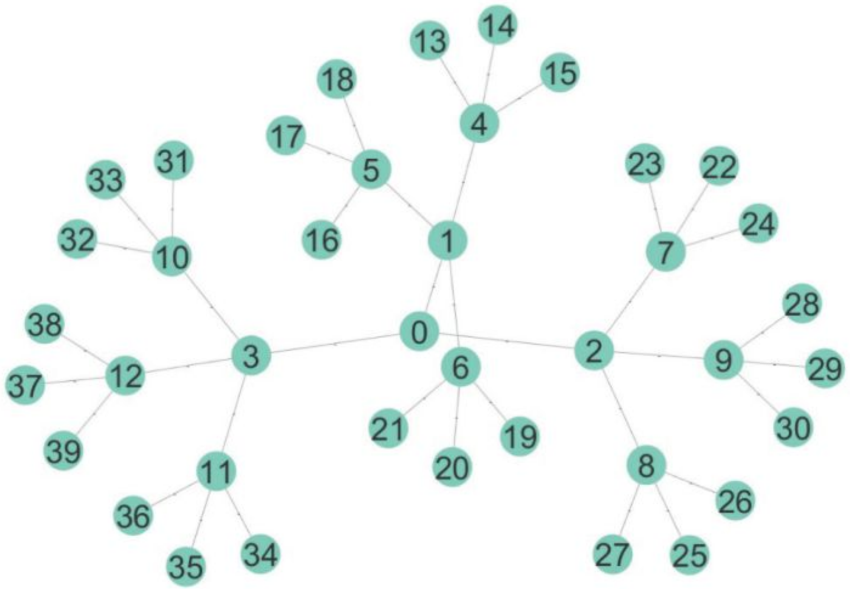


- HRP does not require the invertibility of the covariance matrix.
- In fact, HRP can compute a portfolio on an ill-degenerated or even a singular covariance matrix, an impossible feat for quadratic optimizers.
- Monte Carlo experiments show that HRP delivers lower out-of-sample variance than CLA, even though minimum-variance is CLA's optimization objective.
- HRP also produces less risky portfolios out-of-sample compared to traditional risk parity methods.
- Marcos Lopez de Prado

# New Optimization: Hierarchical Risk Parity

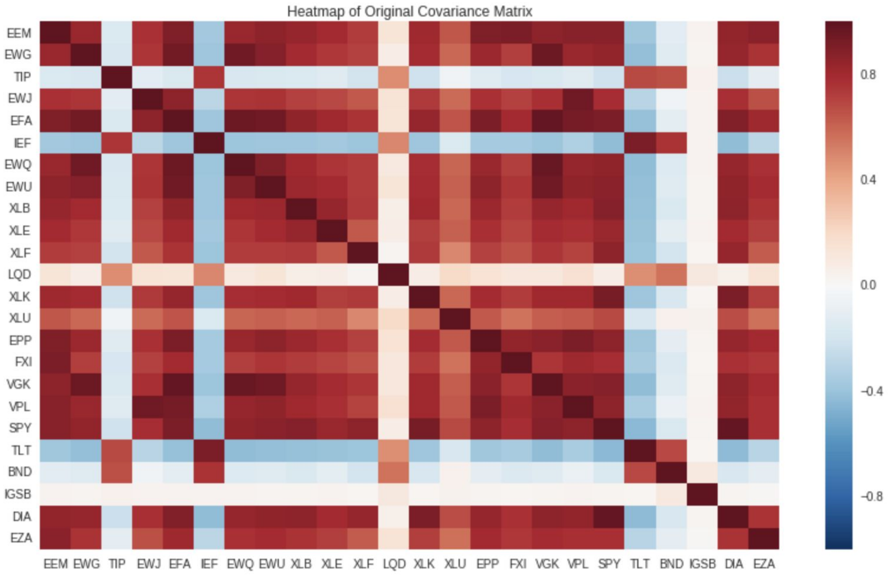
**Trade optimality for stability in unsure future.**

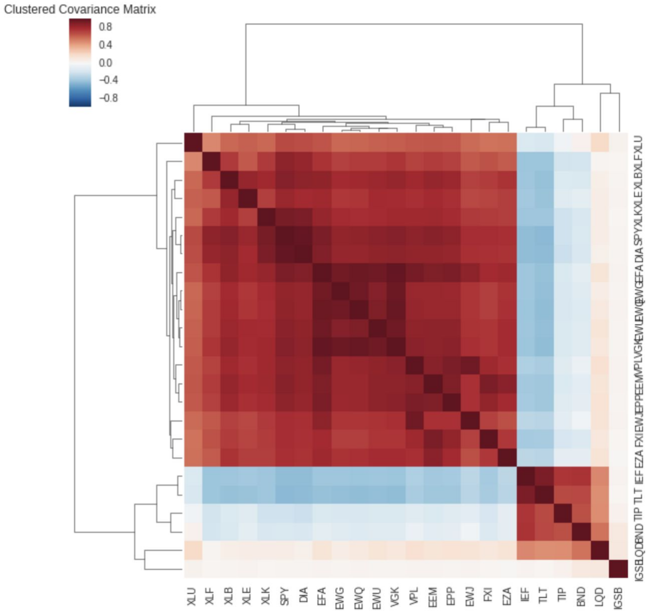
1. **Tree Clustering**
2. **Quasi-Diagonalization (Order by Clustering)**
3. **Recursive Bisection (Allocate by Inverse variance on clusters)**



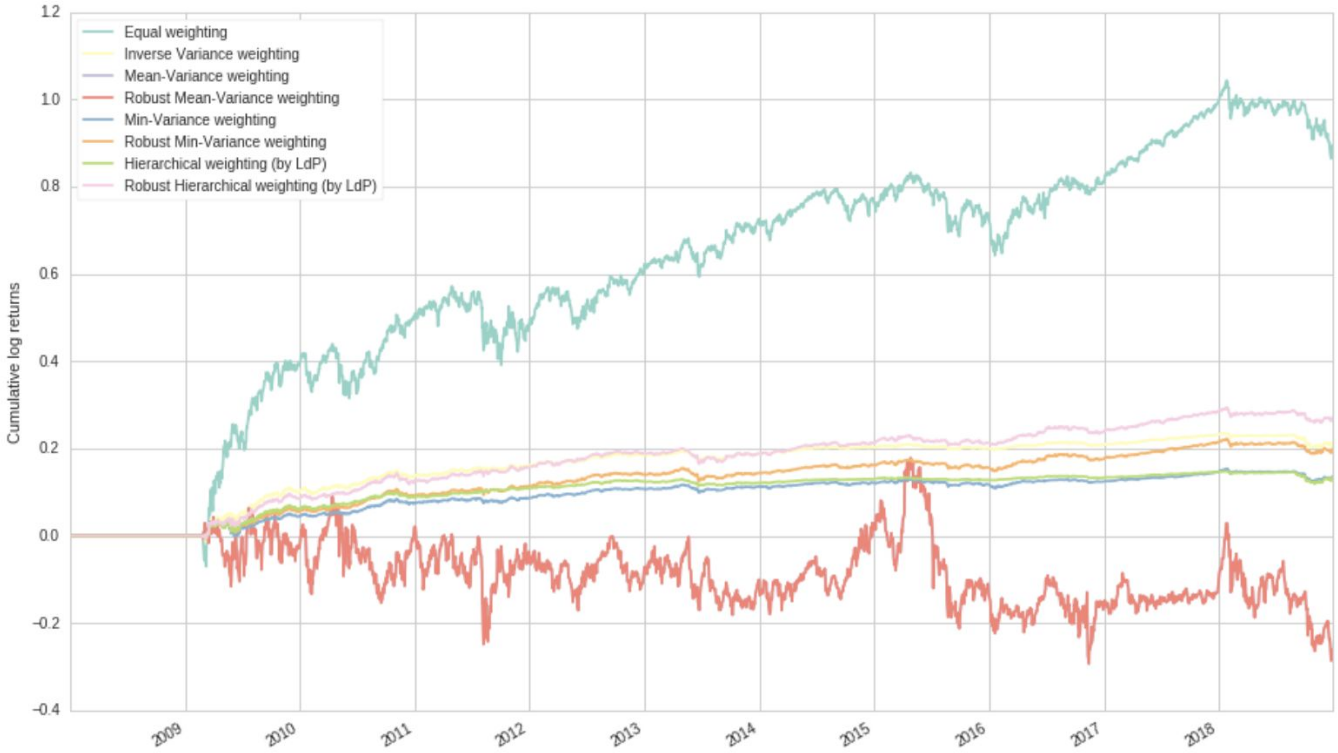
HUDSON AND THAMES

# Quasi-diagonalize (Lots of stable structure)

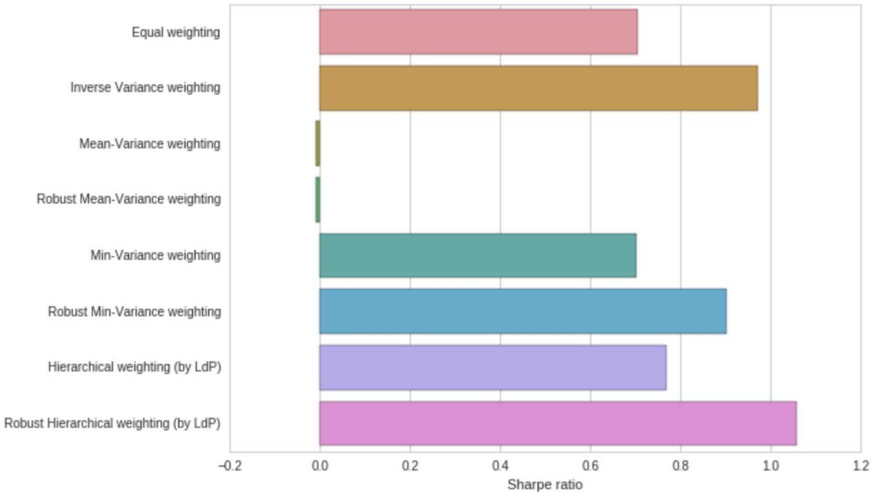


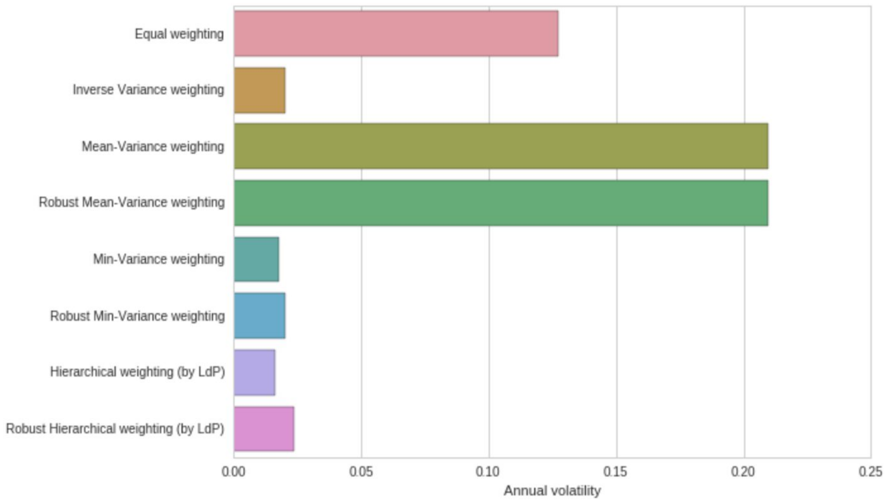



# Performance



# What Really Matters

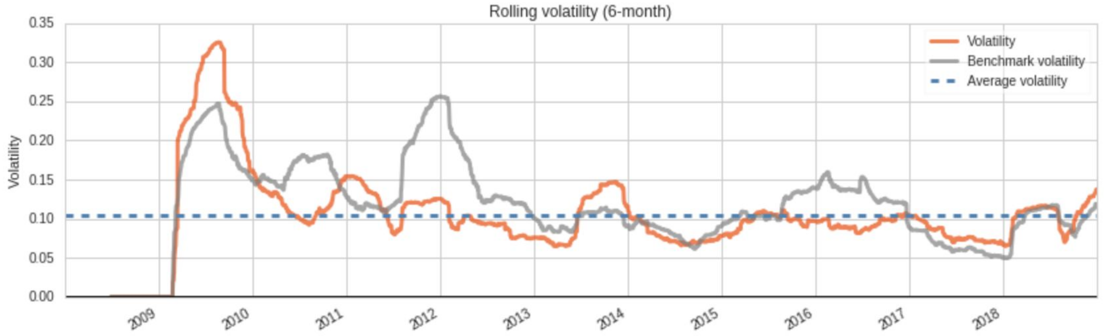


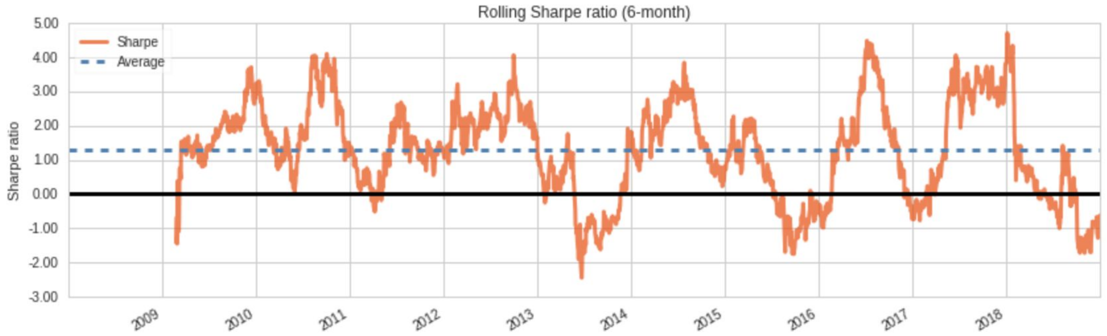



# Scaled Portfolio

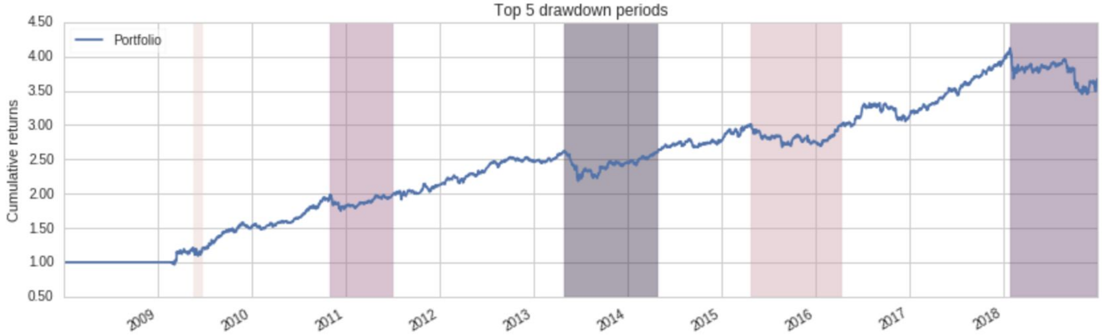
| | |
|---|---|
| Start date | 2008-01-03 |
| End date | 2018-12-31 |
| Total months | 131 |
| | Backtest |
| Annual return | 12.6% |
| Cumulative returns | 266.9% |
| Annual volatility | 11.9% |
| Sharpe ratio | 1.06 |
| Calmar ratio | 0.75 |
| Stability | 0.92 |
| Max drawdown | -16.9% |
| Omega ratio | 1.23 |
| Sortino ratio | 1.63 |
| Skew | 1.71 |
| Kurtosis | 34.35 |
| Tail ratio | 1.10 |
| Daily value at risk | -1.4% |
| Alpha | 0.09 |
| Beta | 0.41 |

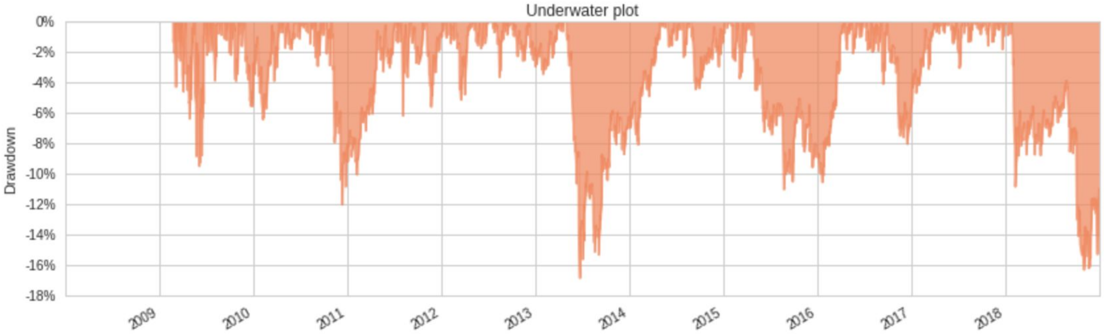
# Scaled Portfolio



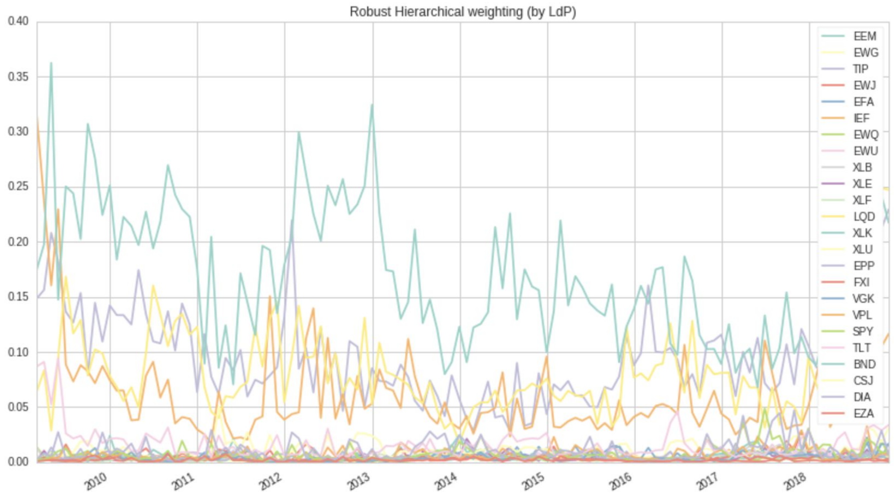



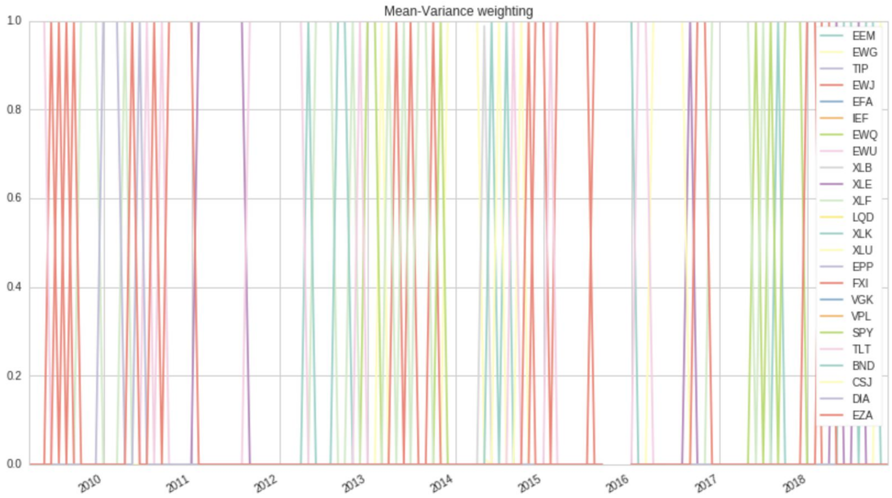
# Scaled Portfolio

# Rebalancing



# Conclusion

Machine learning is opening up new opportunities in financial markets. Linear models are prime for disruption and it is time that we all update our toolkit. This presentation introduced a new technique for portfolio optimisation, that works out-of-sample.

We highly recommend the new text book: Advances in Financial Machine Learning by Lopez de Prado.

A special thank you to Adriaan Janse Van Vuuren, I such fun working with you on this project and will remember it always.


HUDSON
AND THAMES