# FS DataScience

Energy Analytics Optimization

- **We are a Research & Software Development firm**
- **We help companies find and create value with their data using mathematics and computer science techniques**


**Jeff Fong**
Co-Founder, Data Scientist
Chief Product Architect

**Michael Simantov**
Co-Founder, Data Scientist
Chief Quantitative Researcher

# Outline

I. Why are Big Data & Analytic concepts important in Energy Markets?

II. What is Big Data & Analytics?

III. A “small” Big Data & Analytics example

IV. Big Data, Analytics and Energy Markets

V. Getting started if you haven’t already


FS DATASCIENCE

Energy Analytics Optimization

# Why are Big Data & Analytic concepts important in Energy Markets?

# Big Data & Analytics

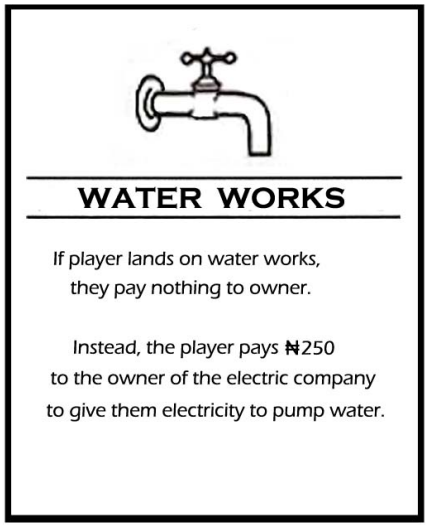
- The Past:  A competitive advantage was often obtained with an asset.
  - Railroads, pipelines, refineries

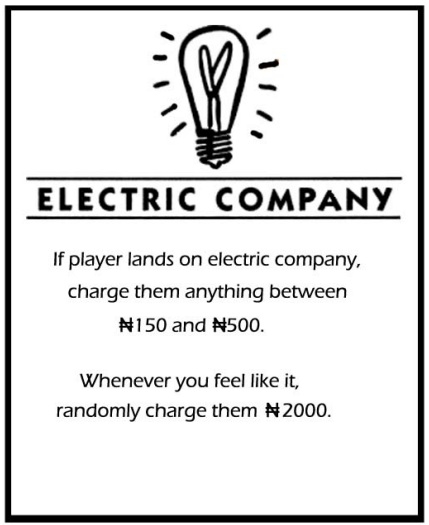
# Big Data & Analytics

- The Past:  A competitive advantage was often obtained with an asset.
  - Railroads, pipelines, refineries

# Big Data & Analytics

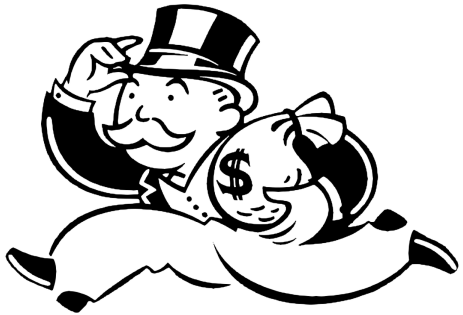
- The Past:  A competitive advantage was often obtained with an asset.
  - Railroads, pipelines, refineries,
  - Water utilities
  - Power generation



# Big Data & Analytics

- The Past:  A competitive advantage was often obtained with an asset.
  - Railroads, pipelines, refineries,
  - Water utilities
  - Power generation
  - Asset based competitive advantage
  - **Monopolies**



# Big Data & Analytics

- What we see today is a trend where analytics revolutionize industries
- Lower capital intensive industries were first
- We have already seen many industries revolutionized by companies that specialize in data and analytics (Google, Amazon)
  - Advertising and Sales
  - Logistics
  - Travel and Airlines

# Big Data & Analytics

- What we see today is a trend where analytics revolutionize industries
- Lower capital intensive industries were first
- We have already seen many industries revolutionized by companies that specialize in data and analytics (Google, Amazon)
  - Advertising and Sales
  - Logistics
  - Travel and Airlines
- **Energy markets are next!**

I. **The data is being stored**
II. **Sufficient processing power is now available**
III. **Markets are becoming integrated**

# Big Data & Analytics

- As energy markets continue to become more global, the competitive advantage from analytics will continue to grow
  - increased integration of global energy markets
  - new pipelines, transmission, energy technologies
  - growing amounts of data
  - increased competition and demand response
  - integration with financial markets

- ***The future of strategic discussions will include include data analytics and algorithms as being competitive advantages***


FS DataScience

Energy Analytics Optimization

# What is Big Data & Analytics?

# What is “Big Data & Analytics”

- “Big” is relative
  - data starts to become big when you can’t make use of the raw data without summarizing it
- Three types of analytics
  1. Descriptive
  2. Predictive
  3. Prescriptive

# Analytics: Descriptive

- Most data in its raw form is not suitable for human consumption
- Data must be condensed into useful pieces of information
- The purpose of descriptive analytics is to summarize what has happened
- The focus is on learning about your data (data discovery)
- 80% of business analytics is descriptive (think about your BI reports)

# Analytics: Descriptive

- Averages, variance, correlations, principle components, distributions and summary statistics are all descriptive
- They help you learn about your data, but are not in themselves, predictive or prescriptive

- Example: How do interest rates, spot prices and volatility move together?

# Analytics: Predictive

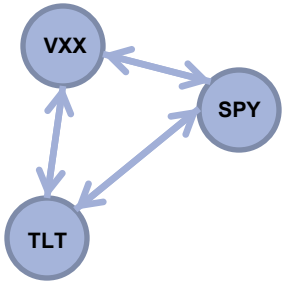
- Predictive analytics do not tell you will happen; it tells you what ***might*** happen.  (probabilistic)
- In other words, using data you do have; fill in data you don't have
  - predict possible future data
  - fill in possible values for missing data

# Analytics: Predictive (=modeling)

- Predictive analytics do not tell you what will happen; it tells you what ***might*** happen.  (probabilistic)
- In other words, using data you do have; fill in data you don't have
  - predict possible future data
  - fill in possible values for missing data
- Model building is a big part of predictive analytics
  - Statistical models, data mining, machine learning, time-series, monte-carlo simulation, forecasting, neural nets are all predictive
  - They tell you what might happen, but not what to do about it

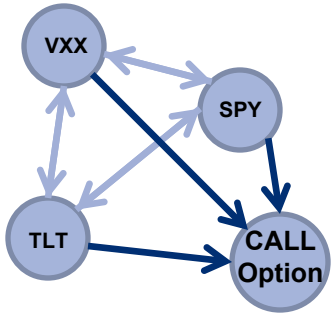
# Analytics: Predictive

- Example: ‘Predict’ or model the value of an OTC call option
  - We use data we have (*interest rates, spot prices and volatility*) to predict data we don’t have (*the call option value*)



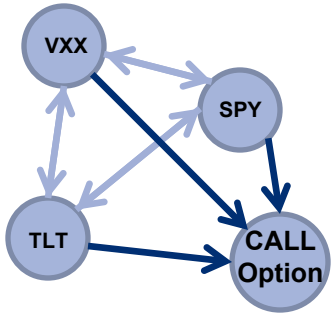
# Analytics: Predictive

- Example: ‘Predict’ or model the value of an OTC call option
  - We use data we have (*interest rates, spot prices and volatility*) to predict data we don’t have
    - *Possible future interest rates, spot prices and volatility*
    - *And possible future option values*



# Analytics: Prescriptive

- Using our data, descriptive analytics and predictive models
  - We can prescribe an action that best achieves an objective
- Prescriptive analytics needs “actionable data”
  - What actions can be taken that affect outcomes?
  - **What are the objectives?**



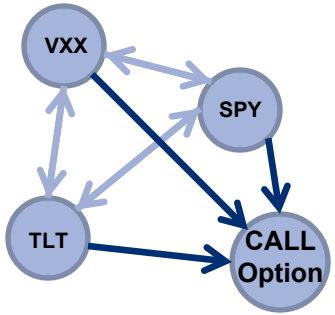

***Minimize risk?***
***Maximize reward?***
***Maximize Sharpe ratio?***

# Analytics: Prescriptive

- Example:  Actions to hedge a call option
  - We have the data for interest rates, spot prices & volatility
  - We have a predictive model for an option price
  - Objective:  Maximize return per unit of risk

***Dynamic hedging is a prescriptive strategy***

- ***delta / gamma hedge***
- ***hedge greek exposures***



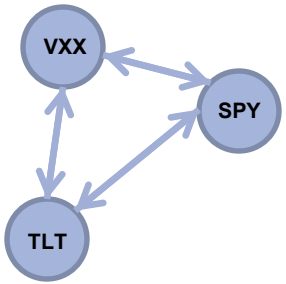

FS DATASCIENCE

Energy Analytics Optimization

# A “small” Big Data & Analytics example

# A Portfolio Trading example

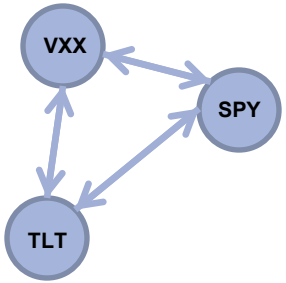
1. SPY - S&P500 index
2. TLT - 20+ year Treasury bonds ETF
3. VXX – S&P500 Volatility ETF



# A Portfolio Trading example

Mini universe of 9 stocks and ETFs

1. SPY - S&P500 index
2. TLT - 20+ year Treasury bonds ETF
3. VXX – S&P500 Volatility ETF
4. JNK - High yield corporate bond ETF
5. UGA - Gasoline futures ETF
6. VLO - Valero Energy Corp (refinery)
7. USO - Oil (WTI) futures ETF
8. JO - Coffee ETF
9. SBUX - Starbucks

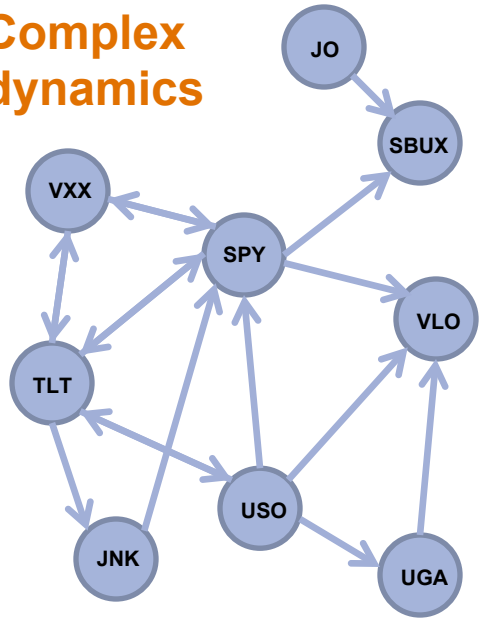


# A Portfolio Trading example

Mini universe of 9 stocks and ETFs

1. SPY - S&P500 index
2. TLT - 20+ year Treasury bonds ETF
3. VXX – S&P500 Volatility ETF
4. JNK - High yield corporate bond ETF
5. UGA - Gasoline futures ETF
6. VLO - Valero Energy Corp (refinery)
7. USO - Oil (WTI) futures ETF
8. JO - Coffee ETF
9. SBUX - Starbucks

**Complex dynamics**




***How can we methodically construct a trading strategy?***

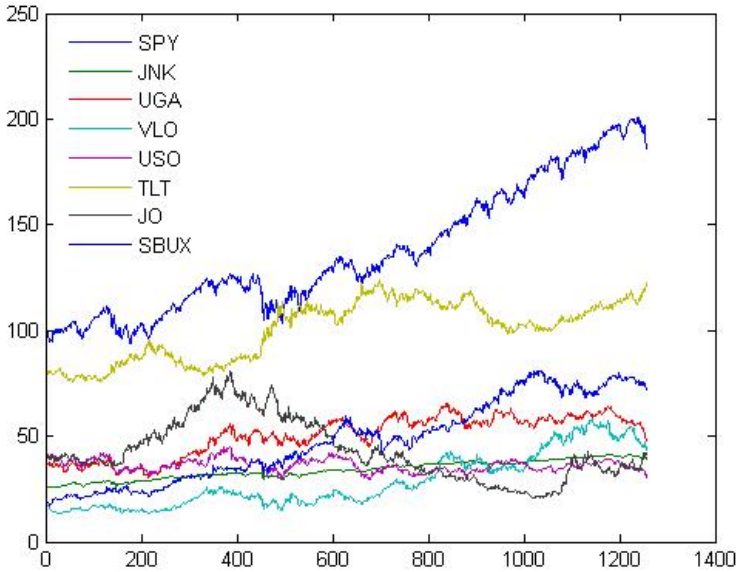
# Descriptive

Mini universe of 8 stocks and ETFs

Calculate:

- means, variances
- correlations, autocorrelations
- principle components

**Discover features about this dataset**



# Predictive

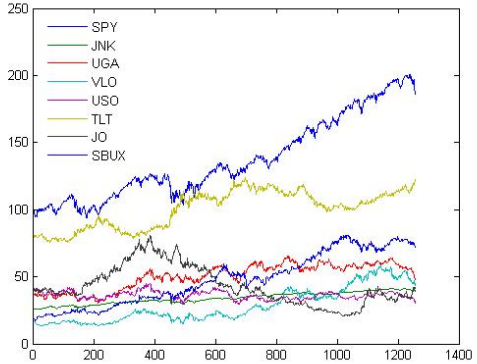
- This “small” example isn’t really that small.  There are 2^8=256 possible combinations of underliers to include, each with endless possibilities of share combinations
- With some data mining *(We’ll make the code available for you to try yourself)*, we can build a (predictive) model of a cointegrated portfolio from our mini universe of stocks and ETFs

# Predictive

- This “small” example isn’t really that small.  There are 2^8=256 possible combinations of underliers to include, each with endless possibilities of share combinations
- With some data mining *(We’ll make the code available for you to try yourself)*, we can build a (predictive) model of a cointegrated portfolio from our mini universe of stocks and ETFs

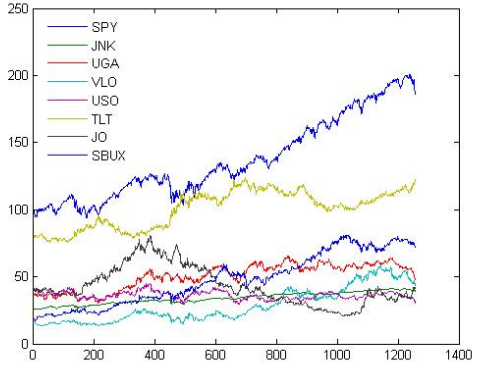
- In otherwords, **we can build a portfolio that mean reverts** using predictive (modeling) analytics

# Predictive

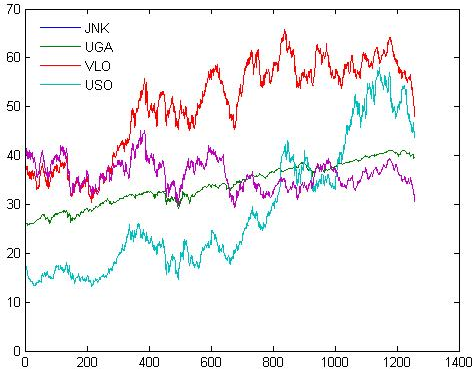



Mini universe

# Predictive

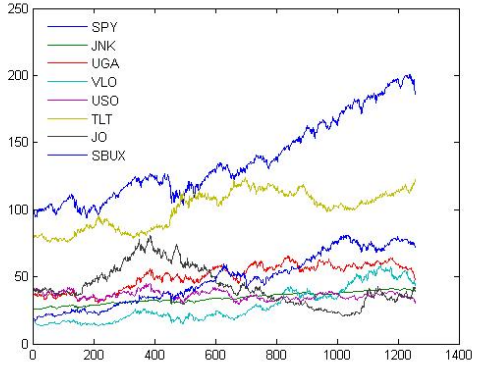



Mini universe

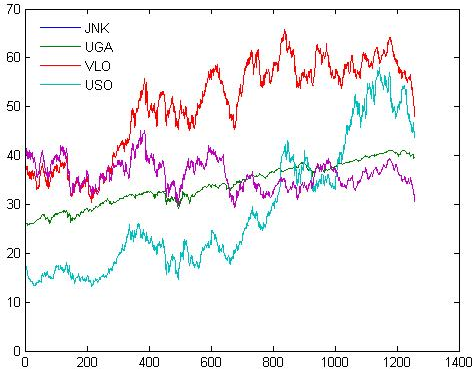



Data-Mined subset

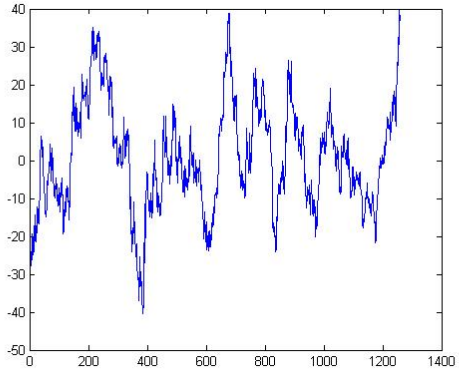
# Predictive




Mini universe



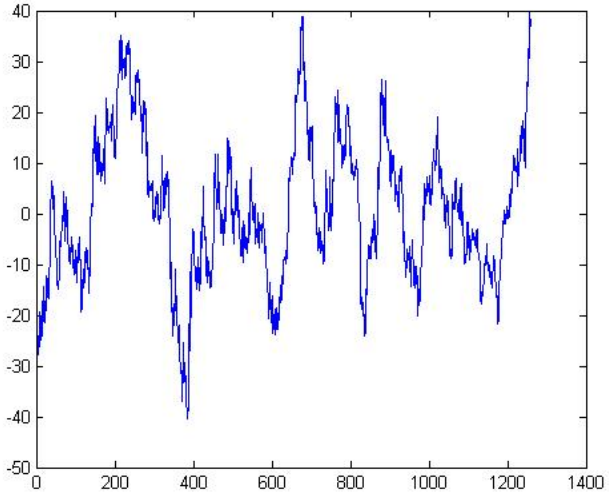

Data-Mined subset




Cointegrated portfolio allocation

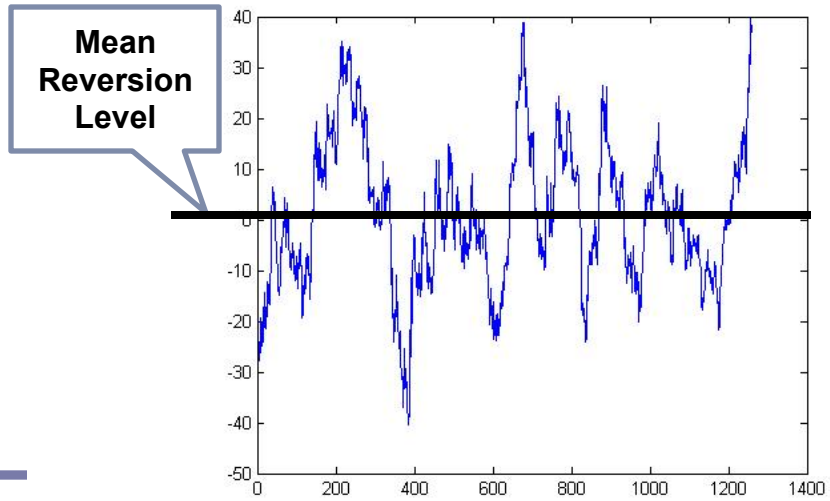
1. long 7.0136 shares of JNK @ $40.31 = $282.72
2. short 2.0775 shares of UGA @ $48.87 = ($101.53)
3. short 1.2010 shares of VLO @ $48.36 = ($58.08)
4. short 2.7351 shares of USO @ $30.64 = ($83.80)

# Predictive

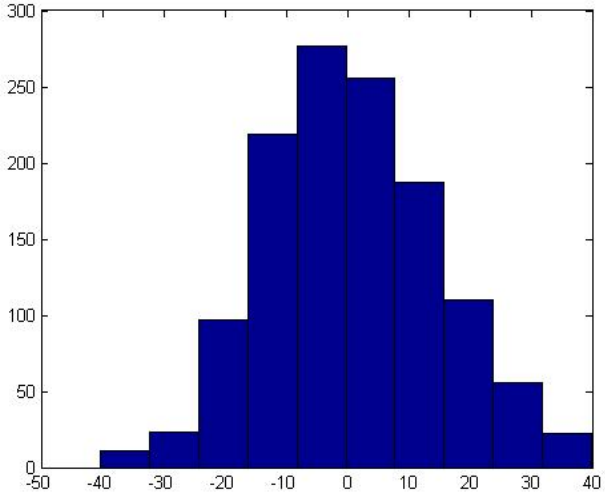



5 year history of the net value of the model cointegrated portfolio

# Predictive

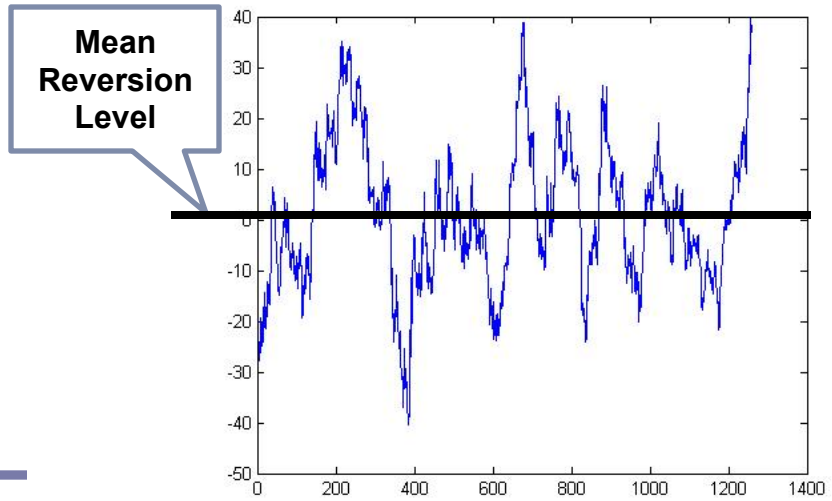


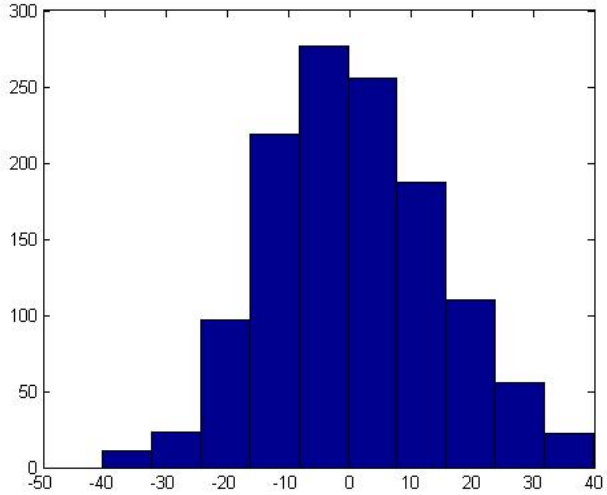

5 year history of the net value of the model cointegrated portfolio




Distribution of the net portfolio value over the 5 years

# Predictive

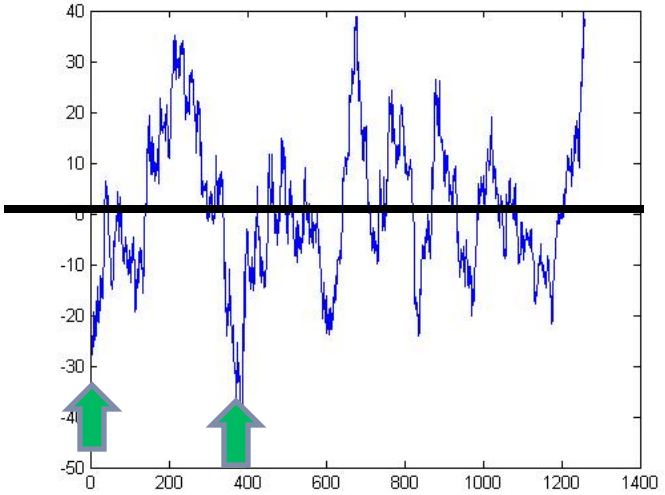


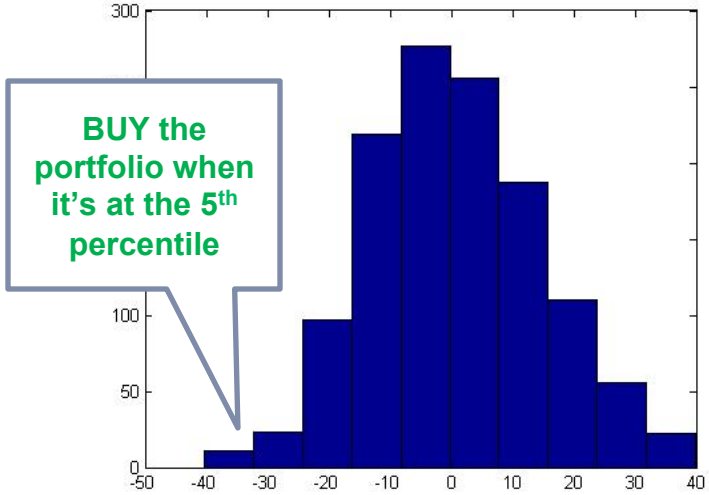




Let's create a simple (prescriptive) strategy with the objective being to generate trading profits!

***This is not a recommendation! Do your own due diligence!***

# Prescriptive

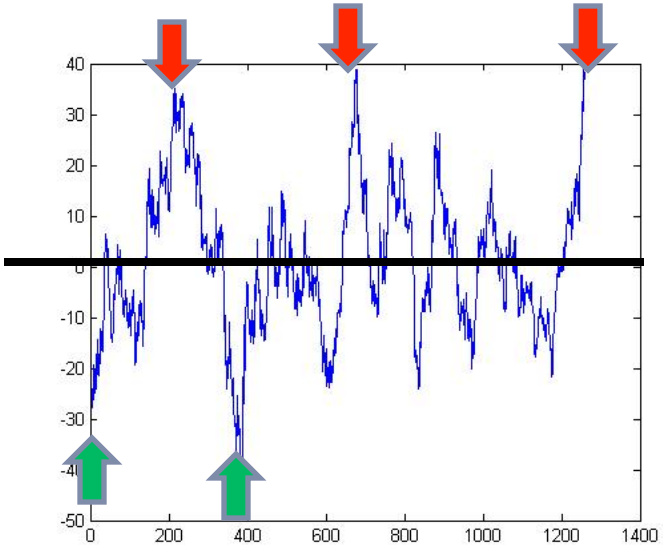


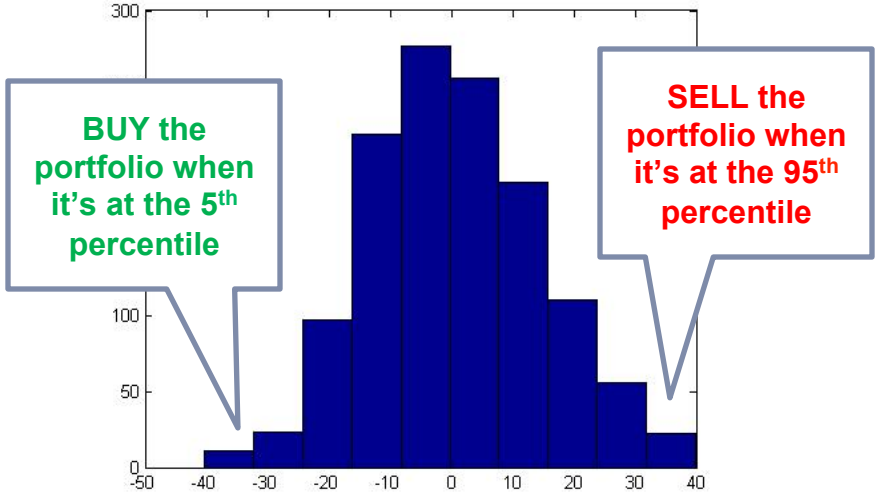




Let's create a simple (prescriptive) strategy with the objective being to generate trading profits!

***This is not a recommendation! Do your own due diligence!***

# Prescriptive

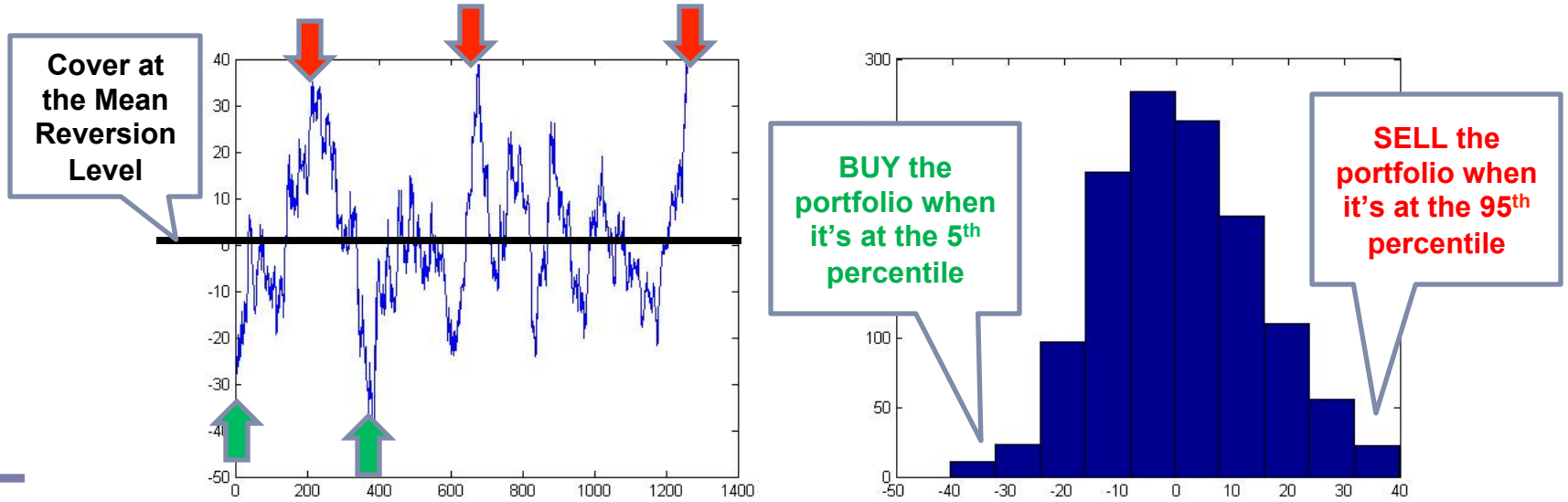



Let's create a simple (prescriptive) strategy with the objective being to generate trading profits!

***This is not a recommendation! Do your own due diligence!***

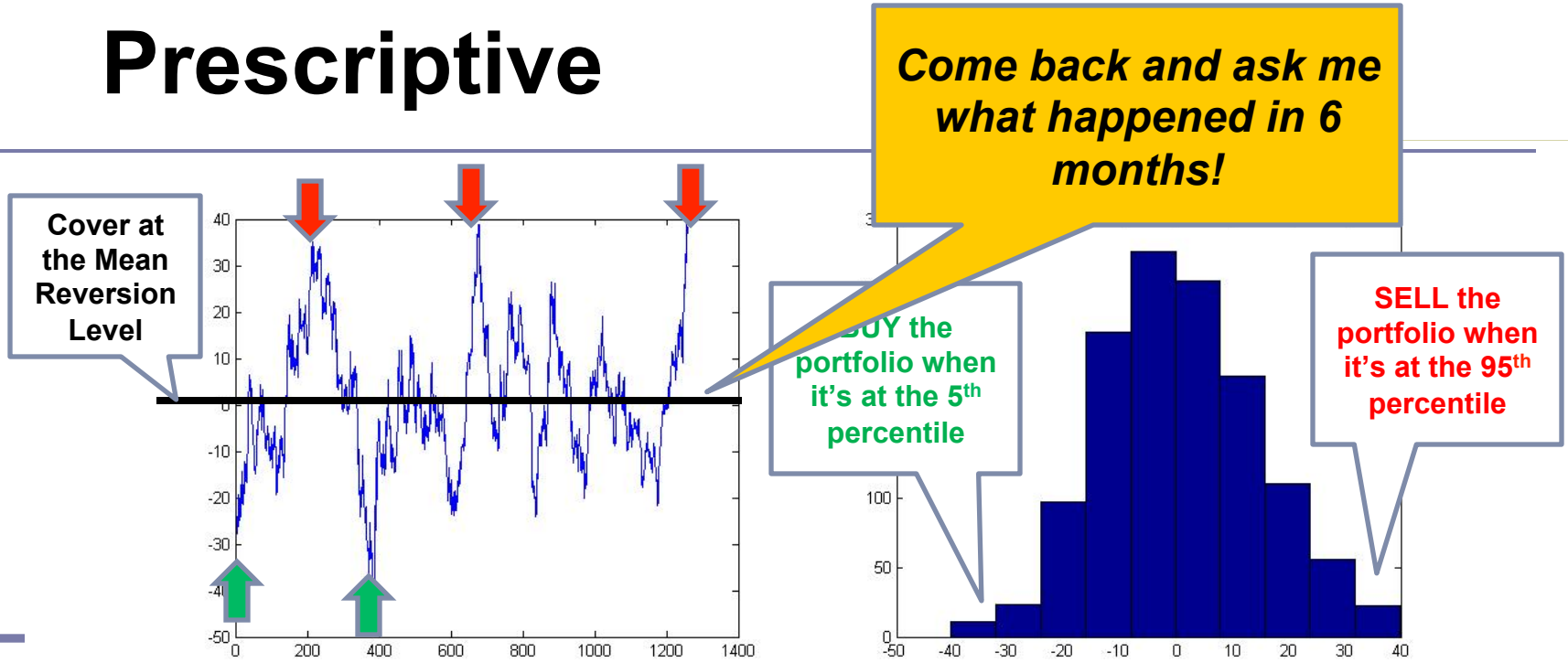
# Prescriptive




Let's create a simple (prescriptive) strategy with the objective being to generate trading profits!

***This is not a recommendation! Do your own due diligence!***

# Prescriptive




Let's create a simple (prescriptive) strategy with the objective being to generate trading profits!

***This is not a recommendation! Do your own due diligence!***


FS DataScience

Energy Analytics Optimization

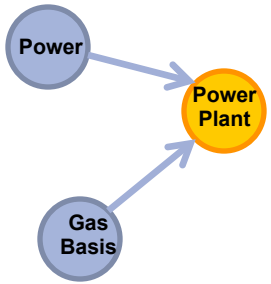
# Big Data, Analytics and Energy Markets

# Energy Market Big Data Analytics

- Modeling energy markets is a difficult and complex problem
  - complex distributions
  - large constrained optimization problems
  - market feedback loops
- Problems in energy markets, get very big, very fast!

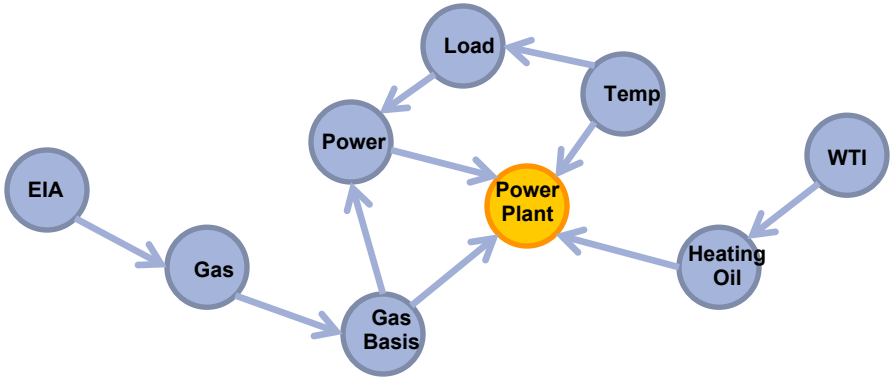
# Energy Market Big Data Analytics

- Digging into the details is like opening pandora's box
- If we look at a power plant, at its surface it seems simple



# Energy Market Big Data Analytics

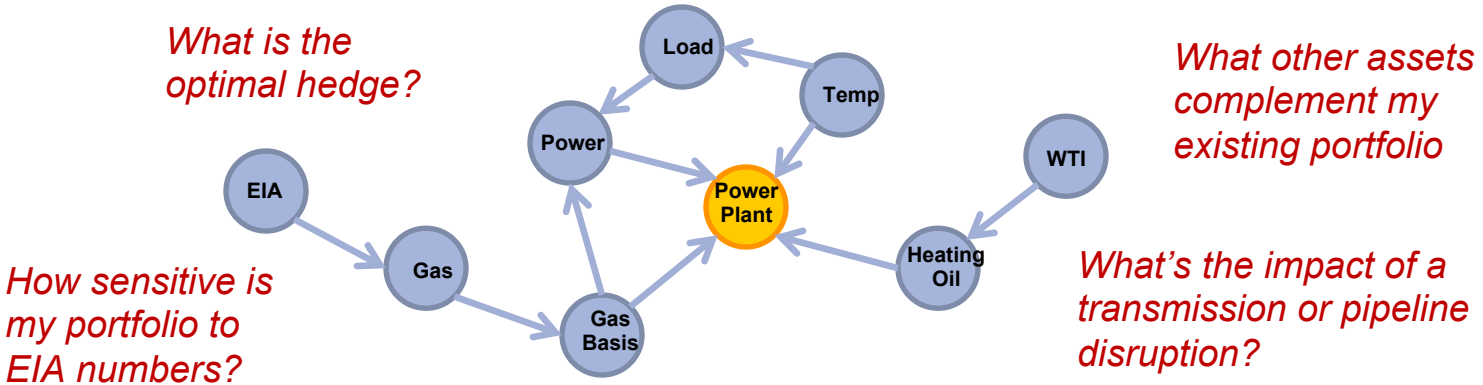
- Digging into the details is like opening pandora's box
- If we look at a power plant, at its surface it seems simple




- Once we peel back the surface, it gets complicated very fast

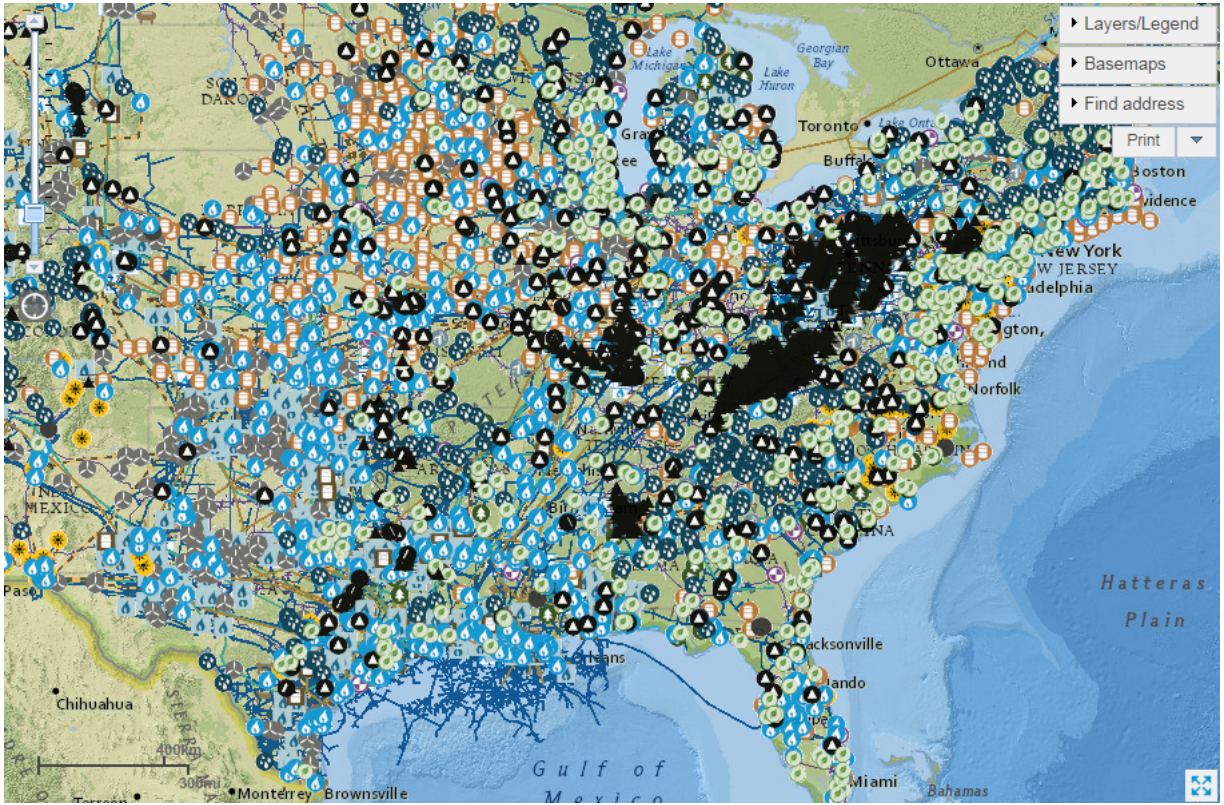
# Energy Market Big Data Analytics

- Digging into the details is like opening pandora's box
- If we look at a power plant, at its surface it seems simple

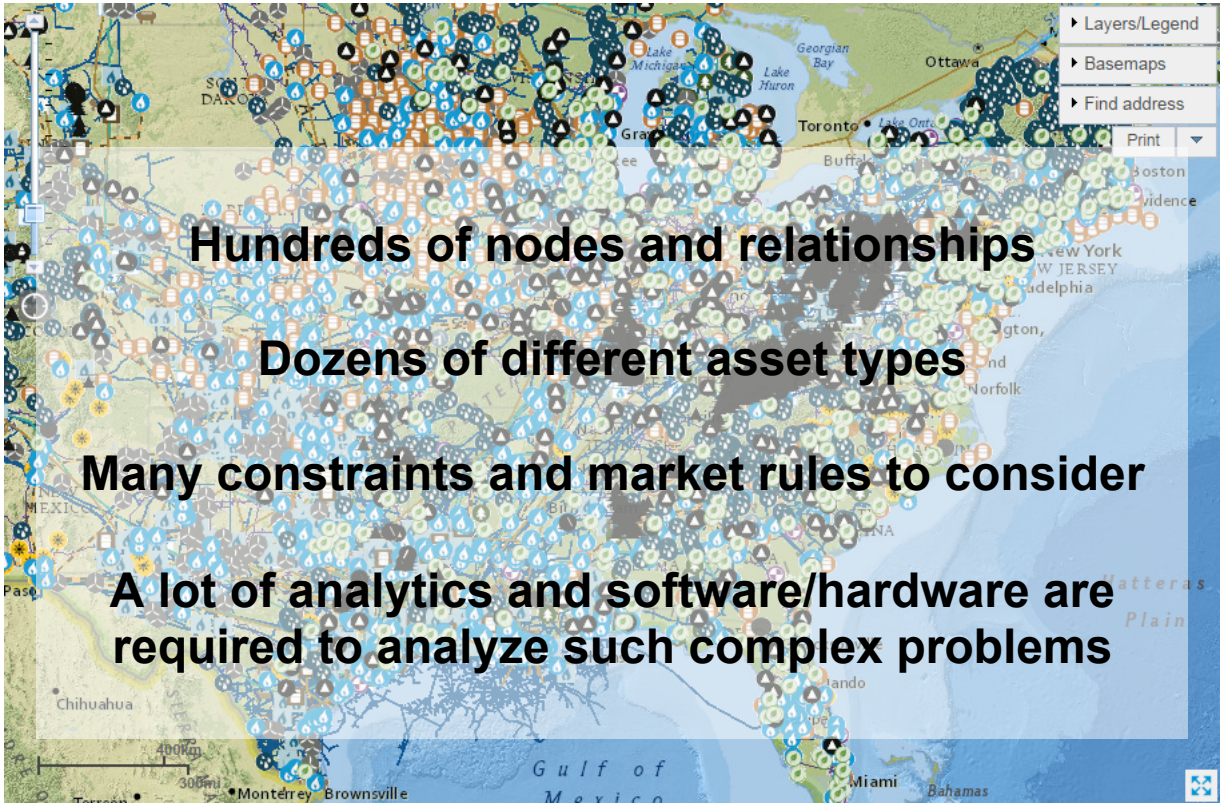



- Once we peel back the surface, it gets complicated very fast

# The “Real” World



# The “Real” World



# Energy Market Big Data Analytics

- Even the largest problems can still be broken down
  - Use an analytics based approach to simplify the problems
  - Include top down quantitative methods with bottoms up fundamental based modeling
  - Leverage “Big Data” technology once its needed
    - Hadoop (reading / writing)
    - Distributed and cloud computing

      (large scale computations)

***Hadoop and distributed / cloud technologies can be extremely powerful in helping you solve large problems***


FS DataScience

Energy Analytics Optimization

# Getting started if you haven't already

# Learn as much as you can!

- Learn to work with and analyze large data sets
  - SQL (any will do)
  - Any good programming or statistical language such as R

# Learn more!

- Learn to work with and analyze large data sets
  - SQL (any will do)
  - Any good programming or statistical language such as R
- Start by solving small optimization problems to gain experience
- Free online universities and excellent courses
  - udacity, coursera, etc
- Learn more about other fields where you have interests

# Learn more!

- Learn to work with and analyze large data sets
  - SQL (any will do)
  - Any good programming or statistical language such as R
- Start by solving small optimization problems to gain experience
- Free online universities and excellent courses
  - udacity, coursera, etc
- Learn more about other fields where you have interests
- **Go write the ERP exam!**

# Gaining Adoption

- Gaining adoption can be difficult at first
- Start with small, easy to handle projects
  - Try descriptive analytics first
- Focus on solving specific problems first to show ROI
- Share results often and invite feedback
  - Incorporate feed back and use rapid prototyping
- Teach & **collaborate**!

# Resources

- Free online courses
  - coursera.org
  - udacity.com
- Open source software options
  - www.mysql.com
  - hadoop.apache.org
  - www.r-project.org
  - http://d3js.org
  - Javascript libraries
- Commercial  software options
  - www.tableausoftware.com
  - Matlab
  - MS SQL / Oracle

FS DataScience

Energy Analytics Optimization

# Questions?