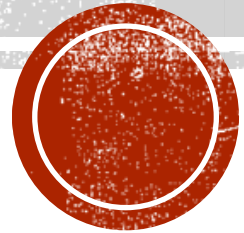
VIX Predicting Modeling



Jingyi Yang
Carnegie Mellon University
The Data Incubator Project

AGENDA

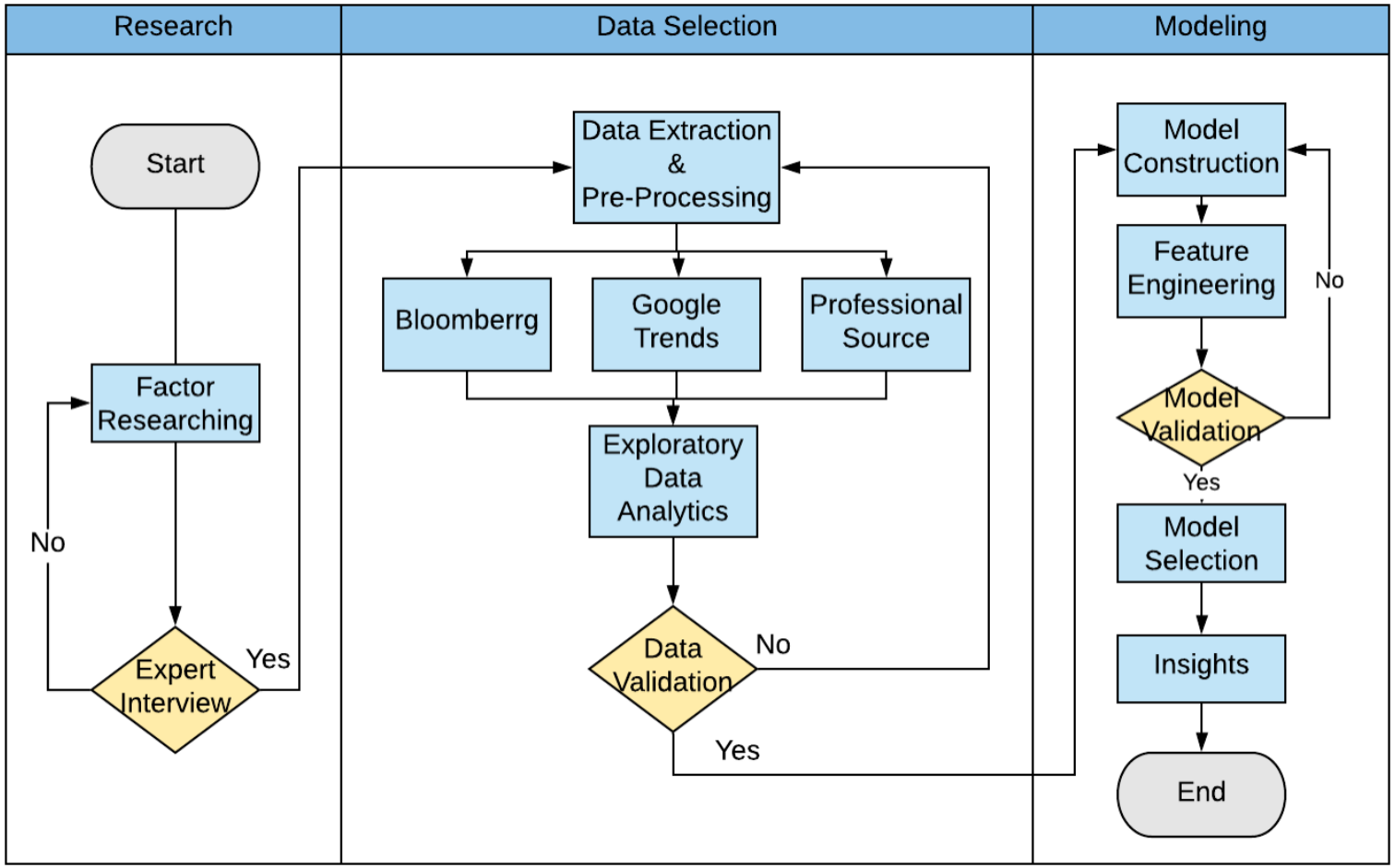
- Problem Statement
- Process Flow Diagram
  - Data Extraction & Pre-processing
  - Exploratory Data Analysis
- Prediction Result
  - Regression Analysis
  - Stepwise Selection Feature
- Key Insights & Conclusion
- Q & A

PROBLEM STATEMENT

- VIX: CBOE Volatility Index, a real-time market index representing the market's expectations for volatility over the coming 30 days
- Investors use the VIX to **measure the level of risk, fear, or stress** in the market when *making investment decisions*
- Project Goal:
  - Use predictive modeling to estimate the change in VIX
  - Research additional relevant factors that will affect VIX changes
- **Dependent Variable:** the VIX "close" difference from day *t-1* to day *t*
- **Predictive Formula:**

$$y = \Delta VIX = VIX_{day1} - VIX_{day0}$$
$$= f(\Delta x_1, \Delta x_2, \dots, \Delta x_n)$$

PROCESS FLOW DIAGRAM



DATA EXTRACTION & PRE-PROCESSING

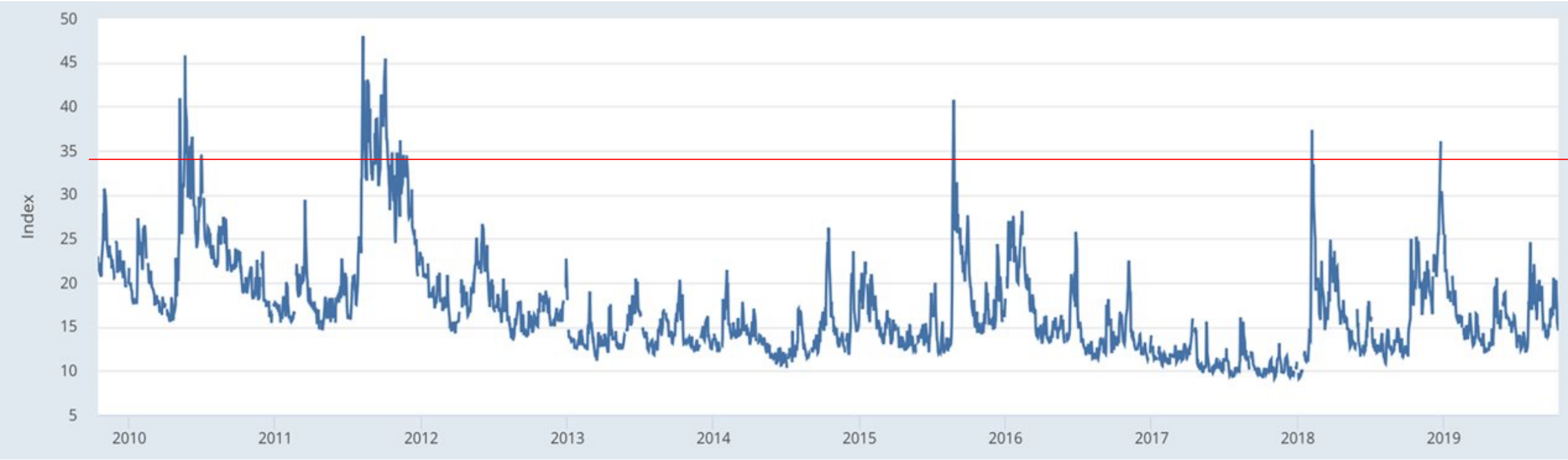
1 **Data Extraction:**

- Collected data from multiple sources: Bloomberg, FRED, CBOE, QUANTIL and Google etc.
- The dataset includes **1546** pieces of training data, **387** pieces of testing data over a **10-year** span
- The model incorporates **20** different factors

2 **Data Pre-processing:**

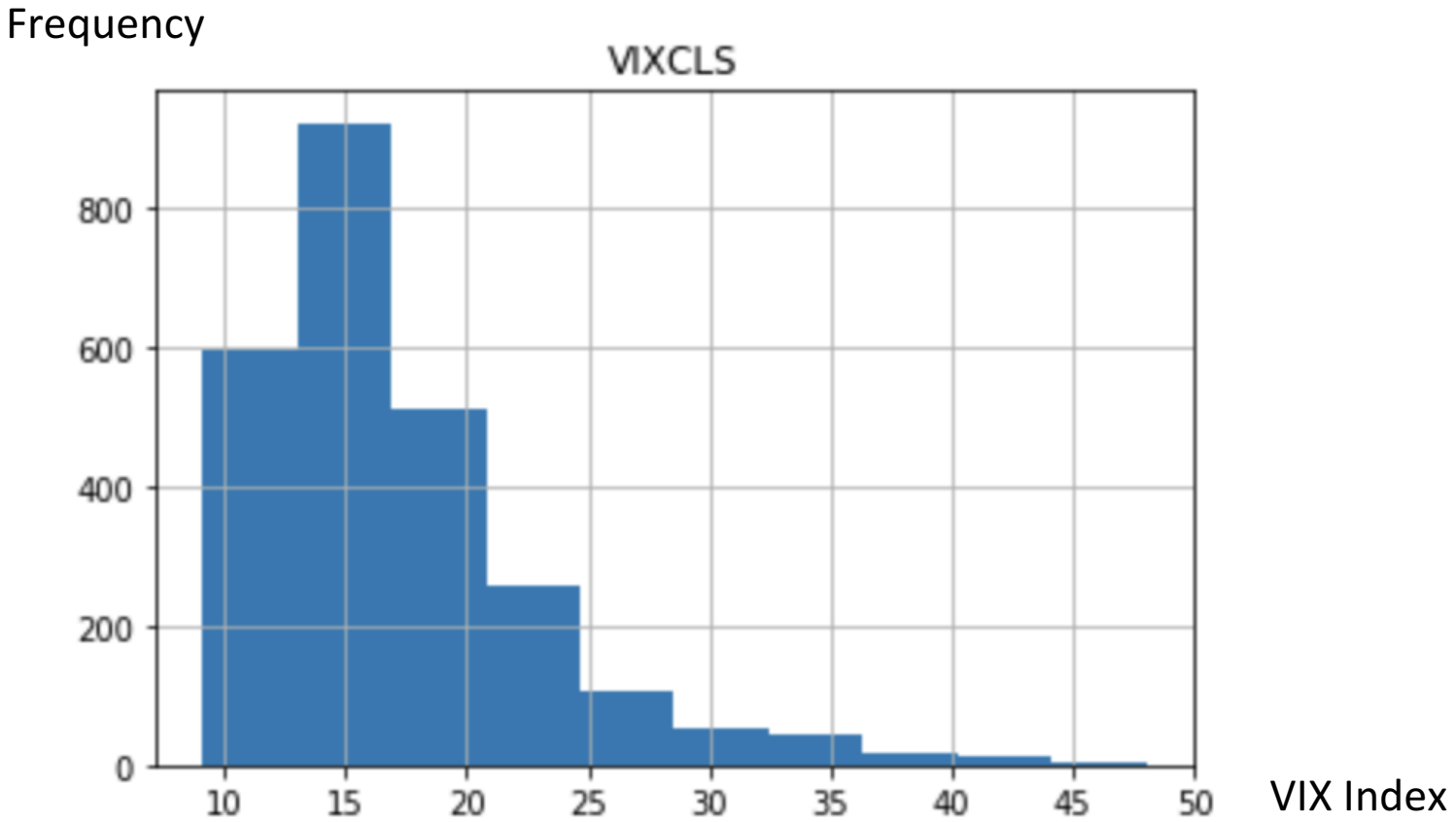
- Standardize the tables into a single format
- Merge datasets into one data frame using Python (pandas)

EXPLORATORY DATA ANALYSIS - TIME SERIES ANALYSIS FOR VIX

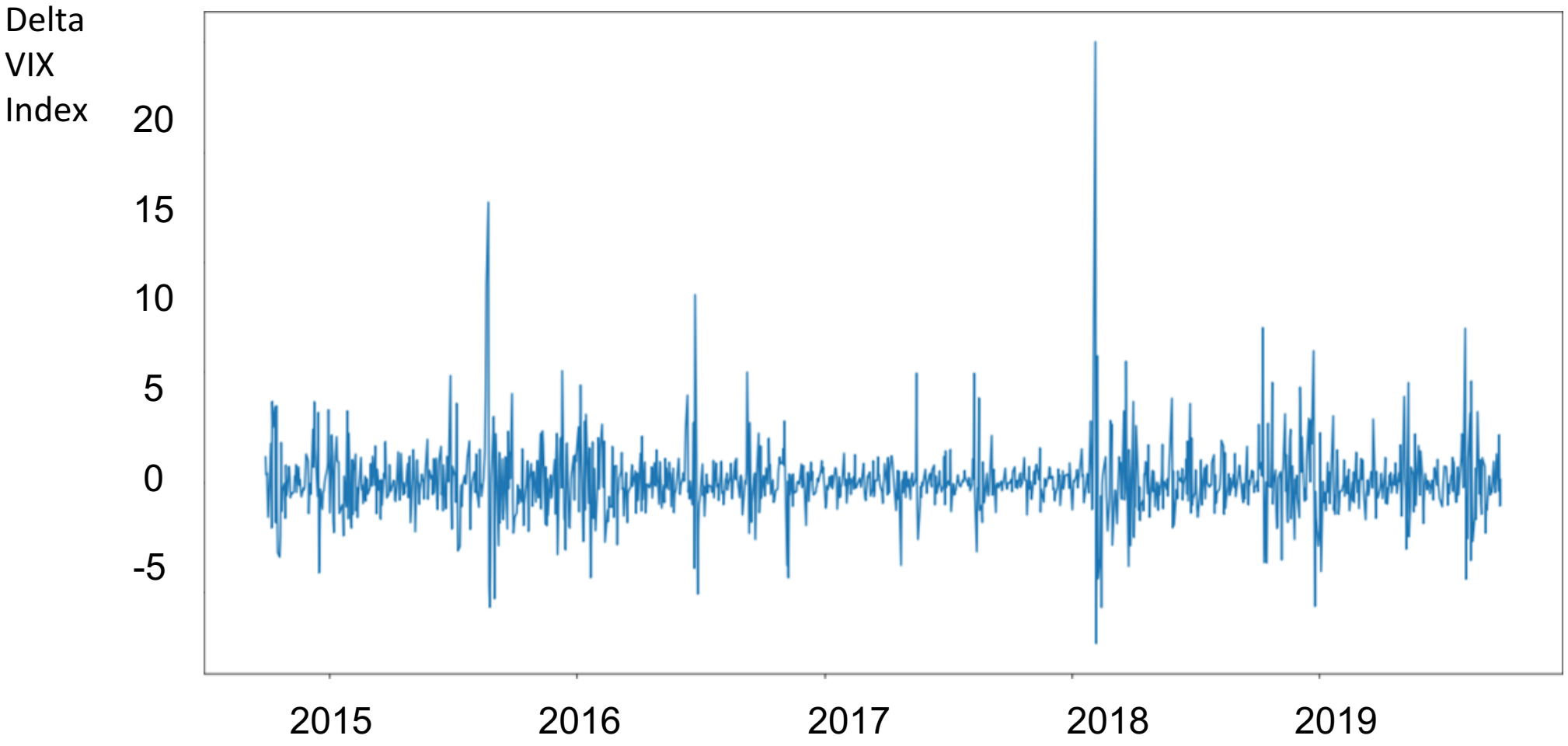


Over the past 10 years, VIX mean: 17.08 standard deviation: 5.69
Using 3 standard deviation, we define the cutoff for outlier at 34.15

EXPLORATORY DATA ANALYSIS – VIX DISTRIBUTION HISTOGRAM



EXPLORATORY DATA ANALYSIS - DELTA VIX



RELEVANT FACTORS THAT AFFECT VIX CHANGES

Exchange Rate
- AUD vs JPY
- EUR vs USD
- USD vs CNY

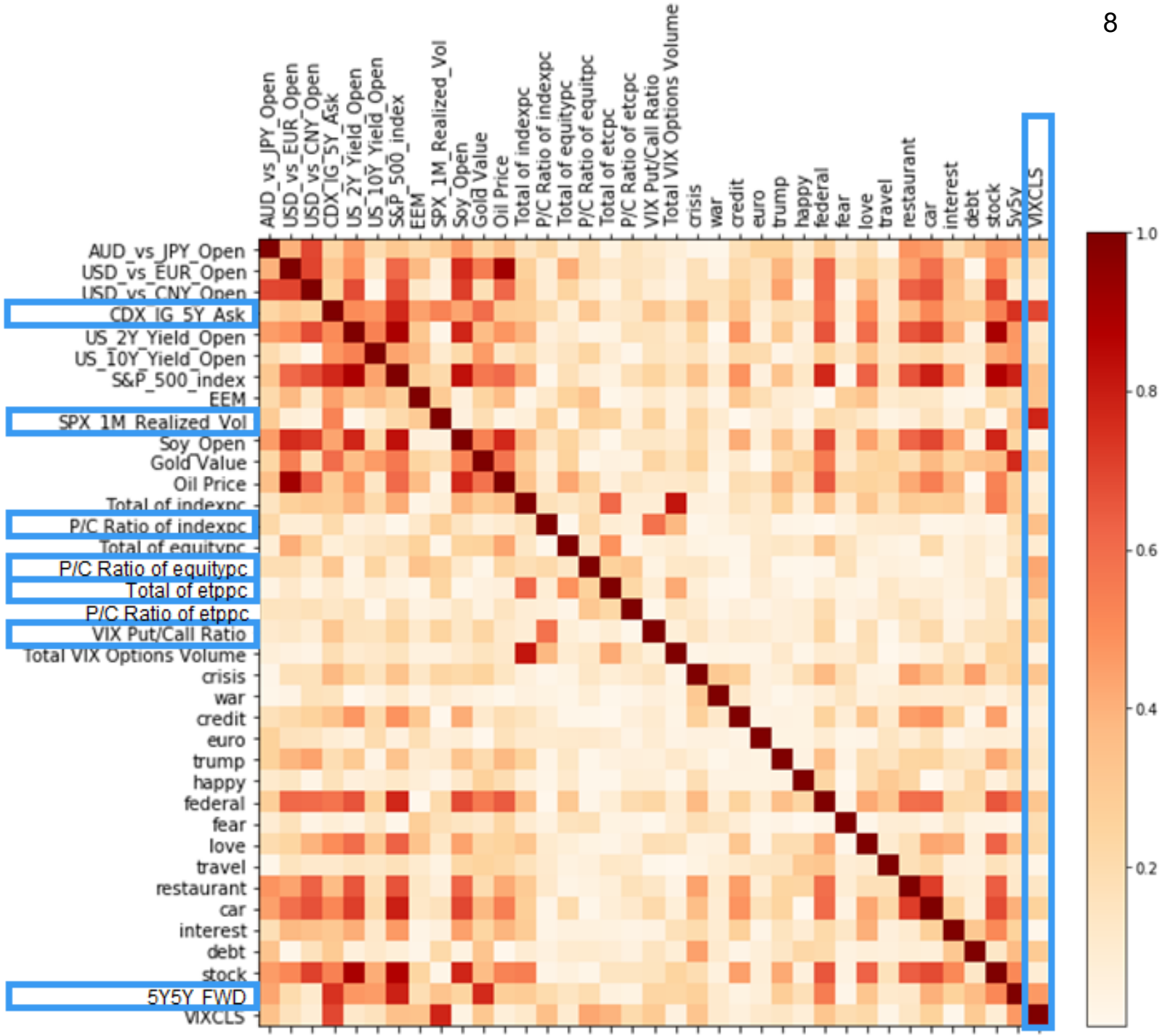
Interest Rate
- US 2Y Yield
- US 10Y Yield

Financial Index
- S&P 500
- CDX IG 5Y
- EEM
- SSE Composite Index
- P/C Options Ratios
- SPX 1M Realized Vol
- VIX Option Volume
- 5y5y FWD
- Dow Jones Index

Commodity Price
- Oil Price
- Gold Price
- Soybean price

Social Influence
- Google Search Volume Index (SVI)

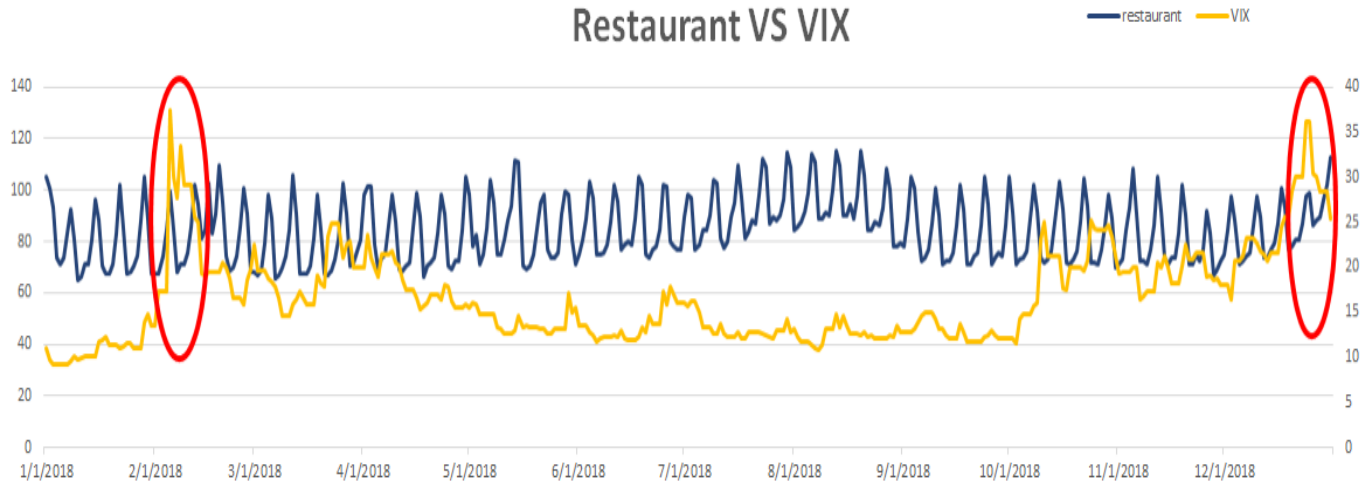
RELEVANT FACTORS' CORRELATIONS



GOOGLE SEARCH VOLUME INDEX

Google Search Volume Index (SVI)		
Category	Keyword	Assumed Correlation
Economic	debt, interest, stock, salary, credit, crisis	↑
Emotion	happy, fear, depress, love, angry, sad	↑ ↓
Consumer	flight, car, restaurant, travel, weather	↓
Political	trump, war, pollice, euro, federal	↑

Tobias Preis, Helen Susannah Moat, & H. Eugene Stanley. (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. Scientific Reports, 3(1), 1684. https://doi.org/10.1038/srep01684



TEST METHOD

- accuracy of 0.5:

  count(abs(predict - real) <= 0.5)/total

- accuracy of 1:

  count(abs(predict - real) <= 1)/total

Stepwise

- Feature Selection method

COMPARISON ACROSS MODELS

	Accuracy of 0.5	Accuracy of 1	Factors
# 22 data factors	47.80%	72.35%	AUD vs JPY, USD vs EUR, USD vs CNY, CDX IG 5Y Ask, US 2Y Yield, US 10Y Yield, S&P 500 index, EEM, SPX 1M Realized Volume, Soy Bean, Gold Value, Oil Price, Total of indexpc, P/C Ratio of indexpc, Total of equitypc, P/C Ratio of equitpc, Total of etcpc, P/C Ratio of etcpc, VIX Put/Call Ratio, Total VIX Options Volume, 5y5y, Dow Jones
# 22 data factors + # 15 google search words with stepwise selection	44.96%	72.61%	Dow Jones, CDX IG 5Y Ask, *Fear,* SPX 1M Realized Volume, *Restaurant*, Total etcpc, *Trump*, *Stock, Euro*, SPX 1M Realized Volume, USD vs CNY