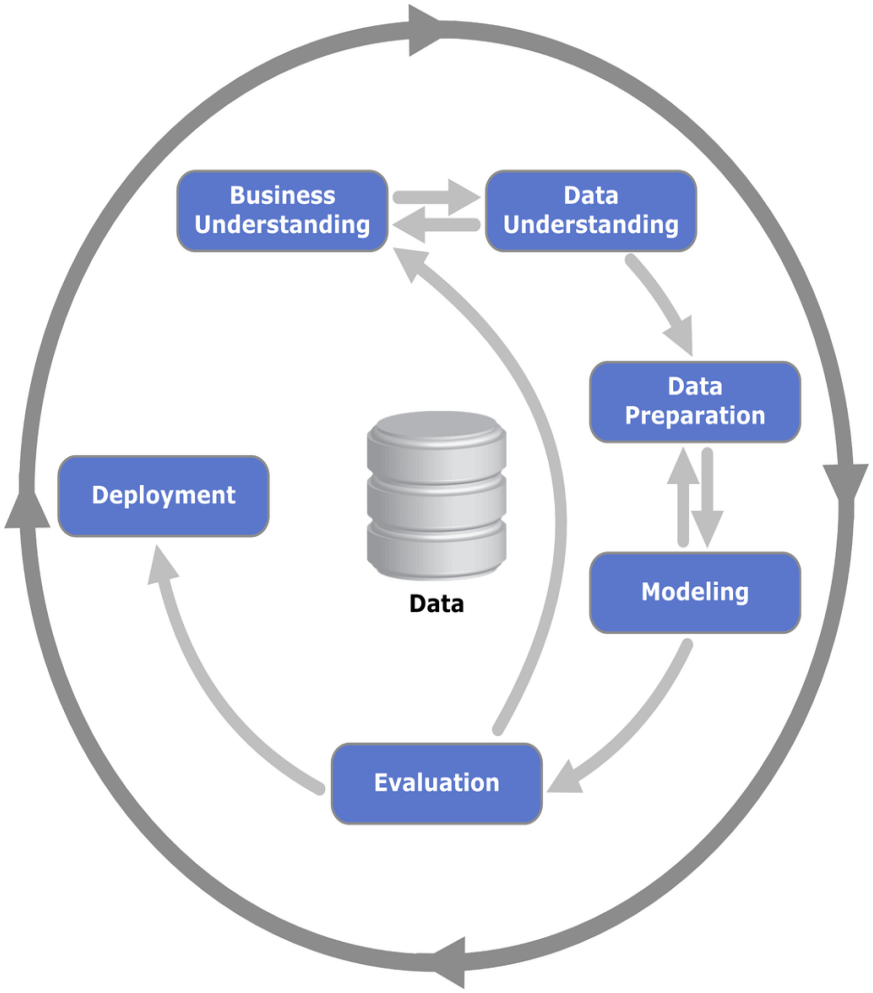
# Quantitative Analysis
# Vietnam Stock Exchange

Binh Kieu Nguyen
Madrid, Dec 24, 2021

Springboard

# Agenda




1. **Problem Statements**
2. **EDA & Hypothesis Definition**
3. **Features selection & Modeling**
4. **Conclusions & Recommendations**

# Problem Statement

Over 10 years historical dataset following quant-trading approach is used in integration of market risks with 03 well-known alpha factors: Momentum, Mean reversion & Seasonality. This is the massive trading algorithm that allows investors quantitatively working with an enormous number of stocks at a time. It is considered as a promising approach for investors, which could minimize bad behaviors during trading, beating the benchmark in Vietnam market.

## 1 Context

Vietnam stock exchange is one of the most potential raising up players in frontier markets. The country which has a dynamic, young population structure & a highly open economic regime with more than 20 years stock exchange life-long is explicitly a promising opportunity for domestic & foreign investors making profits. The market is in an early developing phase & lacking massive & advanced financial tools such as algo-trading (quant-trading), an alternative of fundamental & technical analysis.

## 2 Criteria for success

Vietnam stock exchange in a nutshell

Recommending a betas system & potential alpha factors

Build a hybrid model blending market risks & alpha factors in order to form a set of optimum weight schemes that then are used to design trading strategies to defeat the benchmark (VNindex/VN30).

## 3 Scope of solution space

10 years of historical stock data of Vietnam market.
Project only targets to build a multi-factor model, the combination of risk & alpha model. Multi-factor model outcomes are optimum alpha weights for each stock & portfolio net factor exposures for each principal component. Strategy design & backtest are out of scope.
Deliverables include: code, documentation & slice deck

## 4 Constraints within solution space

Solution is just considered as a reference, not financial advice.
Solution is time bounded & subject to type of security

## 5 Stakeholders to provide key insight

Data from online open-source platforms

Solution could be advised from: stock traders, quantitative engineers/analysts, portfolio managers

## 6 Key data sources

https://www.vndirect.com.vn/portal/thong-ke-thi-truong-chung-khoan/lich-su-gia.shtml
https://s.cafef.vn/Lich-su-giao-dich-VNINDEX-1.chn#data

H D E I P

# Project Quantitative Analysis – Vietnam Stock Exchange

## Data Wrangling

**Dataset:**

1. Historical stock data: OHLC Volume Dividend of 621 tickers, from 2006-2021
2. Sector: 621 stocks' sector, from 2006-2021

**Data cleansing:**

Remove duplicate & unnecessary columns
Extract states from locations, correct states 'name convention
Convert data type from object to numeric one

**Raw features:**

'open', 'high', 'low', 'close', 'volume', 'dividend', 'sector''

**Features definition:**
*Key feature: close (stock closing price)*

**Features transforming:**
Create daily simple returns & log returns features

**Feature extraction:**
Create Bollinger band feature set
Create breakout feature set
Create Momentum feature set for 1 week, 1 month, 1 quarter, 1 half & 1 year

**Features for visualization:**
close, simple_returns, log_returns, Bollinger band feature set, breakout feature set

# Project Quantitative Analysis – Vietnam Stock Exchange

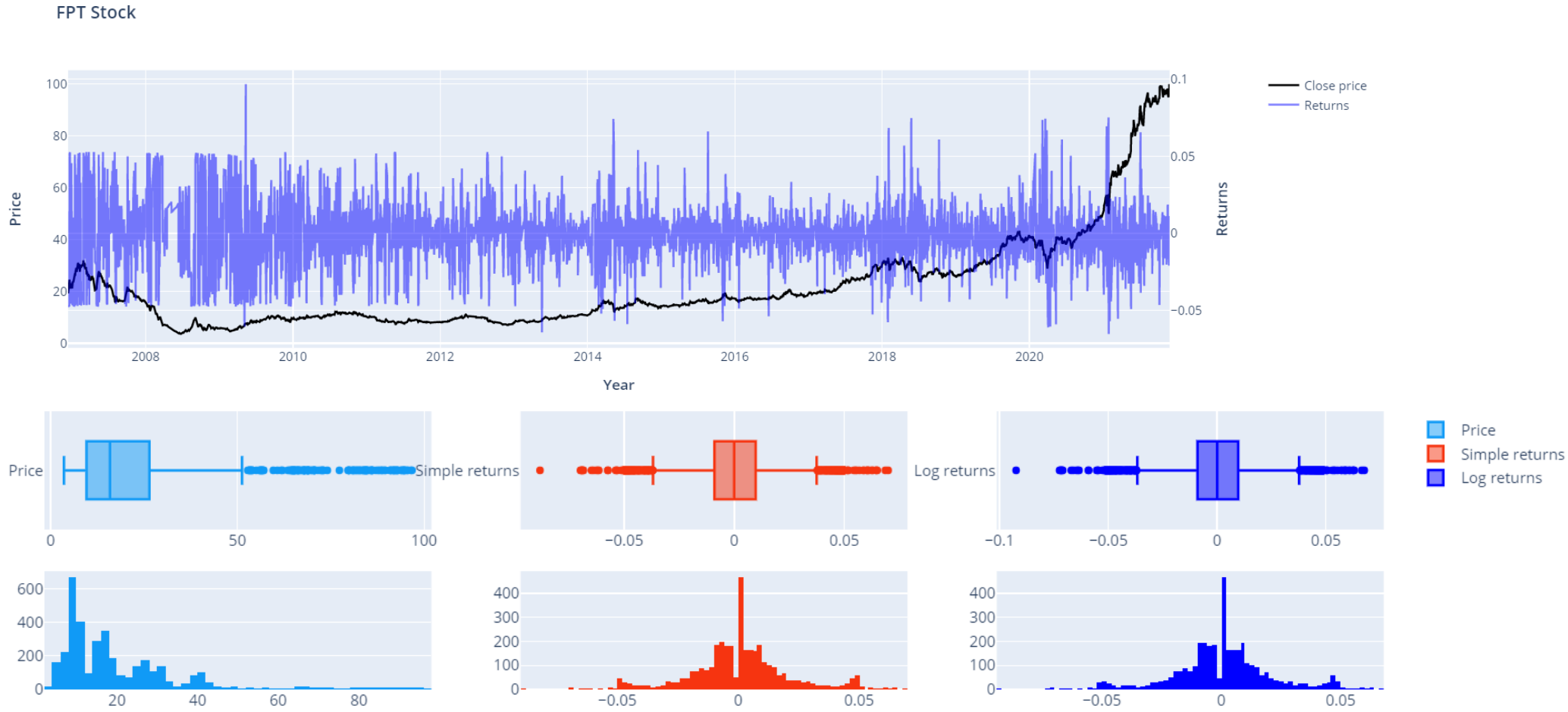
## Exploration Data Analysis – Data Overview

**FPT stock:**

Price is non-stationary data, not normal distribution with mean & variance change by time

Returns data is more stationary, normal distribution with mean=0 & variance/standard deviation stable by time

Return data will be the input for model construction later on

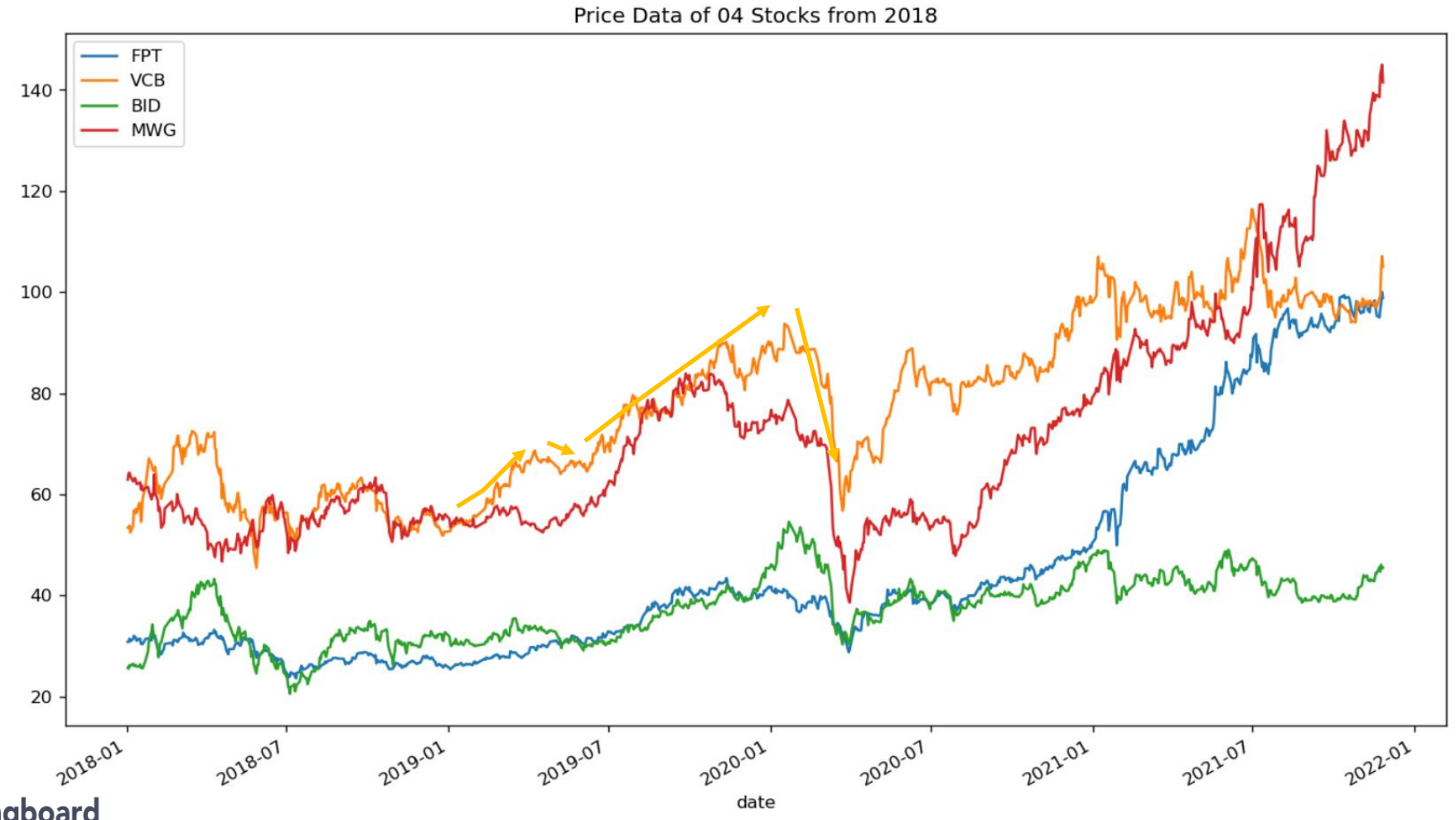


# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Visualization Stock Price

**Top 4 stocks of the market:**

- Both momentum & mean reversion effect could be seen

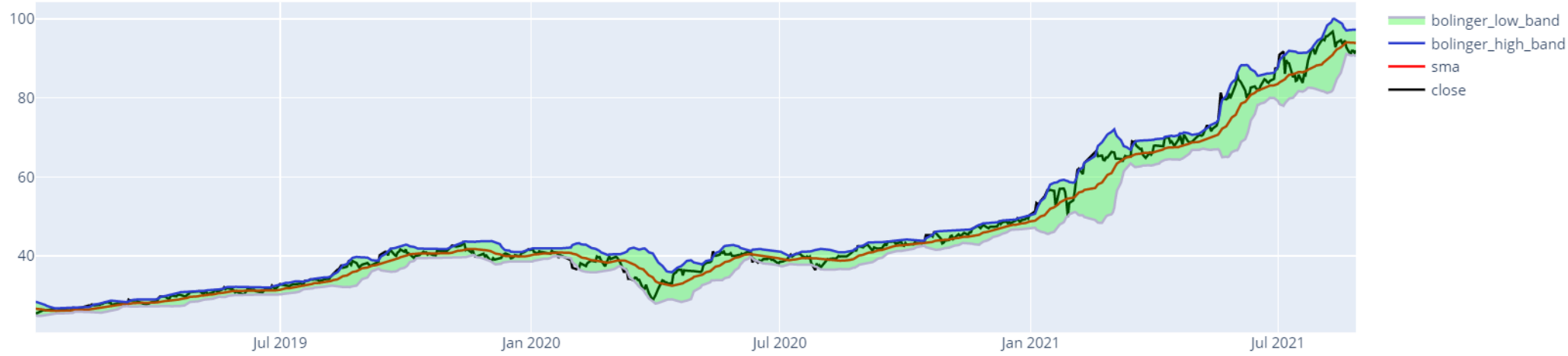


# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Bollinger Band Effect

**FPT stock Bollinger band effect from 2018-2021:**

- 1 month simple moving average curve (sma), upper & lower band cut the ticker price curve which could be long & short signal

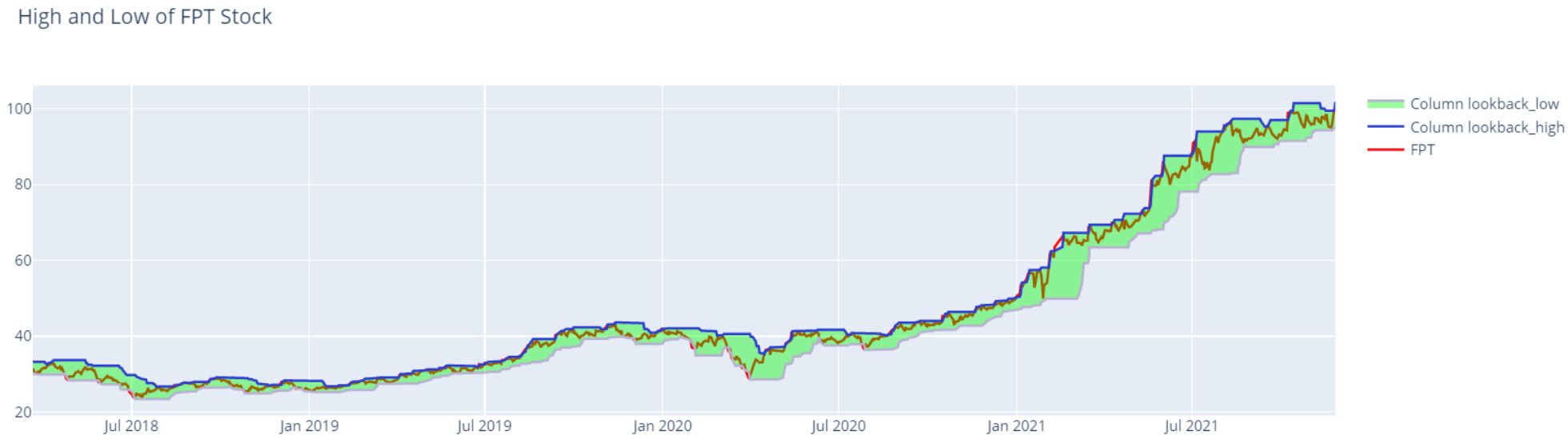


# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Breakout Effect

**Features visualization by time: 2009-2018:**

- 1 month lookback low & high curves cut the FPT price one creating potential trading signals

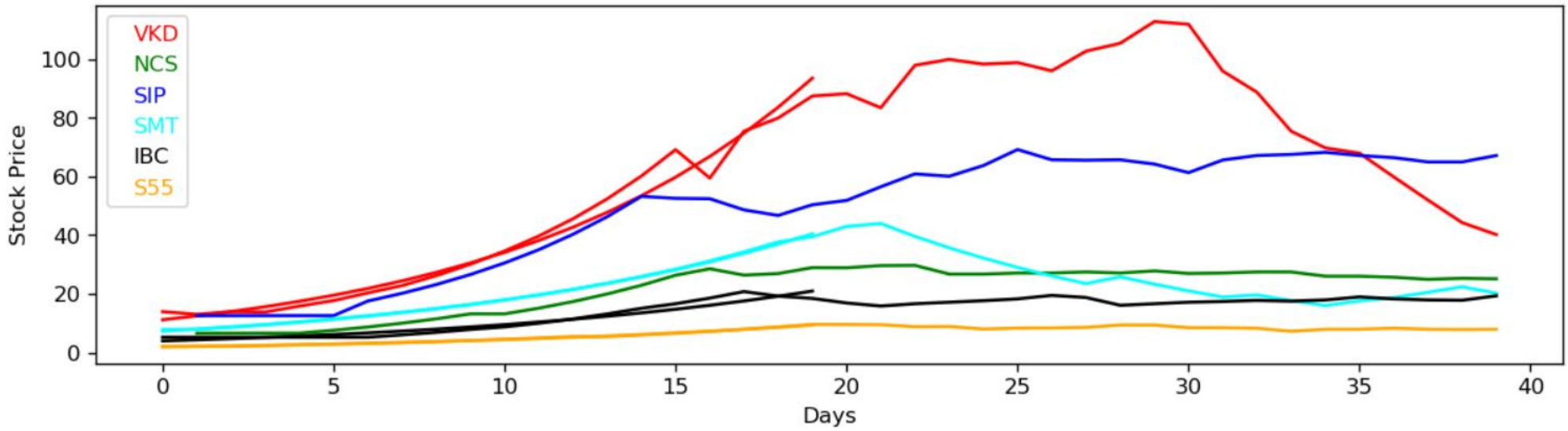


# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Momentum Effect

**Time series analysis: 2011-2021:**

- Top 6 stocks have highest 1month momentum effect
- The first 3 stocks (VKD, SIP, SMT) seems having a significant momentum intensity
- 1month momentum effect for the rest are predicted  from fair to weak, even negative.

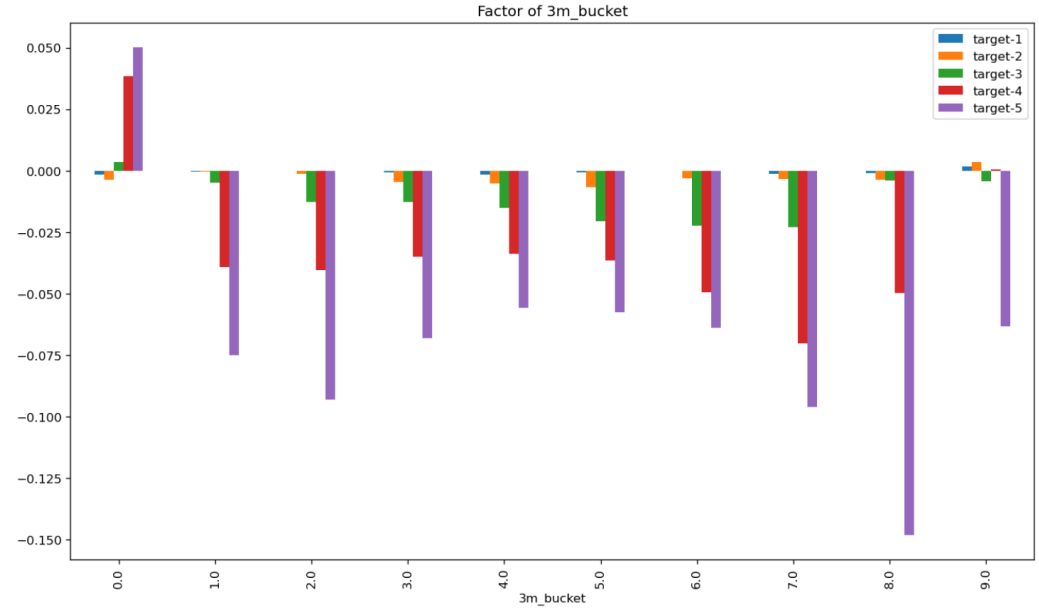


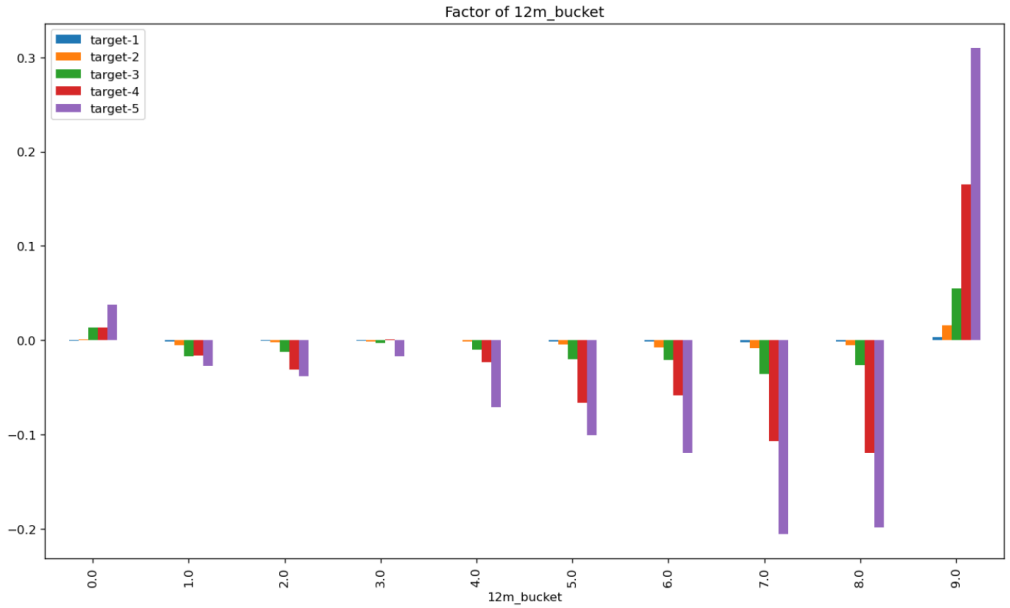
# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Momentum Effect

**Cross sectional momentum analysis: 04 stocks basket (FPT, VCB, BID, MWG)**

- A year momentum works better than quarter one, quantile 10th also perform the best



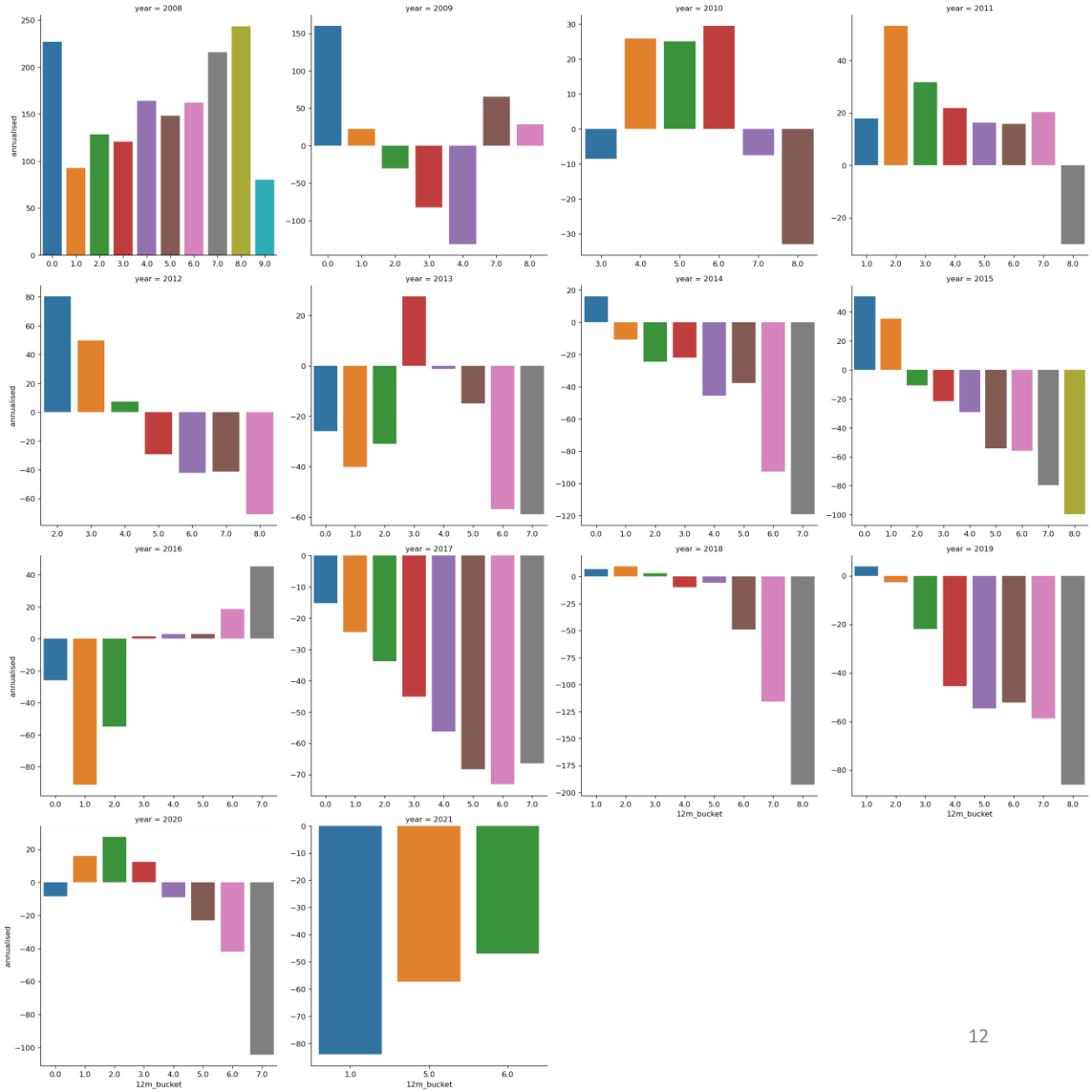
# Project Quantitative Analysis – Vietnam Stock Exchange

## Exploration Data Analysis – Momentum Effect

**Cross sectional momentum analysis:**

**04 stocks basket (FPT, VCB, BID, MWG) 2008-2021**

- 12months momentum displays by year
- The first half of quantiles seem more resilient than the last part



# Project Quantitative Analysis – Vietnam Stock Exchange

## Hypotheses Statements

**Momentum**

Weekly, monthly, quarterly, halfly & yearly momentum indicators have been observed both in time series & cross sectional analysis, they are considered as the short, mid & long-term alpha factors.

- Yearly momentum hypothesis: it is "Higher past 12-months (252 days) returns are proportional to future return."
- Halfly momentum hypothesis: it is "Higher past 6-months (120 days) returns are proportional to future return."
- Quarterly momentum hypothesis: it is "Higher past 3-months (60 days) returns are proportional to future return."
- Monthly momentum hypothesis: it is "Higher past 1-month (20 days) returns are proportional to future return."
- Weekly momentum hypothesis: it is "Higher past 1-week (5 days) returns are proportional to future return."

The other indicators such as mean reversion & overnight sentiment - seasonality, both are short-time factors, are also taken into account in the model.

**Mean reversion**

Mean reversion theory states that security prices and economic indicators such as interest rates will tend to revert to the historical mean prices.

- Mean_reversion_5day "Short-term outperformers(underperformers) compared to their sector will revert."
- Mean_reversion_5days_smoothed is a smoothed version of Mean_reversion_5day factor
- Mean_reversion_20day "Short-term outperformers(underperformers) compared to their sector will revert.

**Overnight sentiment**

The US stocks have significantly higher return during the night session compared to the daily session.

- Overnight Sentiment
- Overnight Sentiment Smoothed

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction

**Candidate selection**

This is a pretty high level research, it covers the whole VN stock exchange. 621 exchange-traded stocks, from 2011-09-01 to 2021-09-01 will be used as input for modeling process.

Poor candidate selection could lead to scattering/divergent in analysis & low performance of the models.

$$R_p(t) = \sum(B_{p,k} * R_m(t) + \alpha_p(t))$$

Risk factors are well-known and significant contributors to the variance of asset returns. **Risk model** actually is **Portfolio Covariance Matrix** will be built on stock returns data under CAPM, EMH & Modern Portfolio Theory-Effective Frontier theories to compute model factors: factor_betas, factor_returns, factor_cov_matrix, idiosyncratic_var_matrix & idiosyncratic_var_vector to estimate the portfolio risk. It will then be used as the constrains for alpha operation in the weight optimization process of multi-factor model

Factors that are proprietary and significant in describing the mean of asset returns can be candidates for alpha factors. **Alpha model** is built under Fundament law. Momentum, mean-reversion and seasonality effects are the key resilient factors for alpha model.

Finally, the two models are combined as a **Multi-factor model** in order to derive optimum weight schemes that neutralize market risks and maximize alpha returns. From that, investors/trader/portfolio managers can quantitatively design their own strategies & backtests that allows them quantitatively make up their further decision.

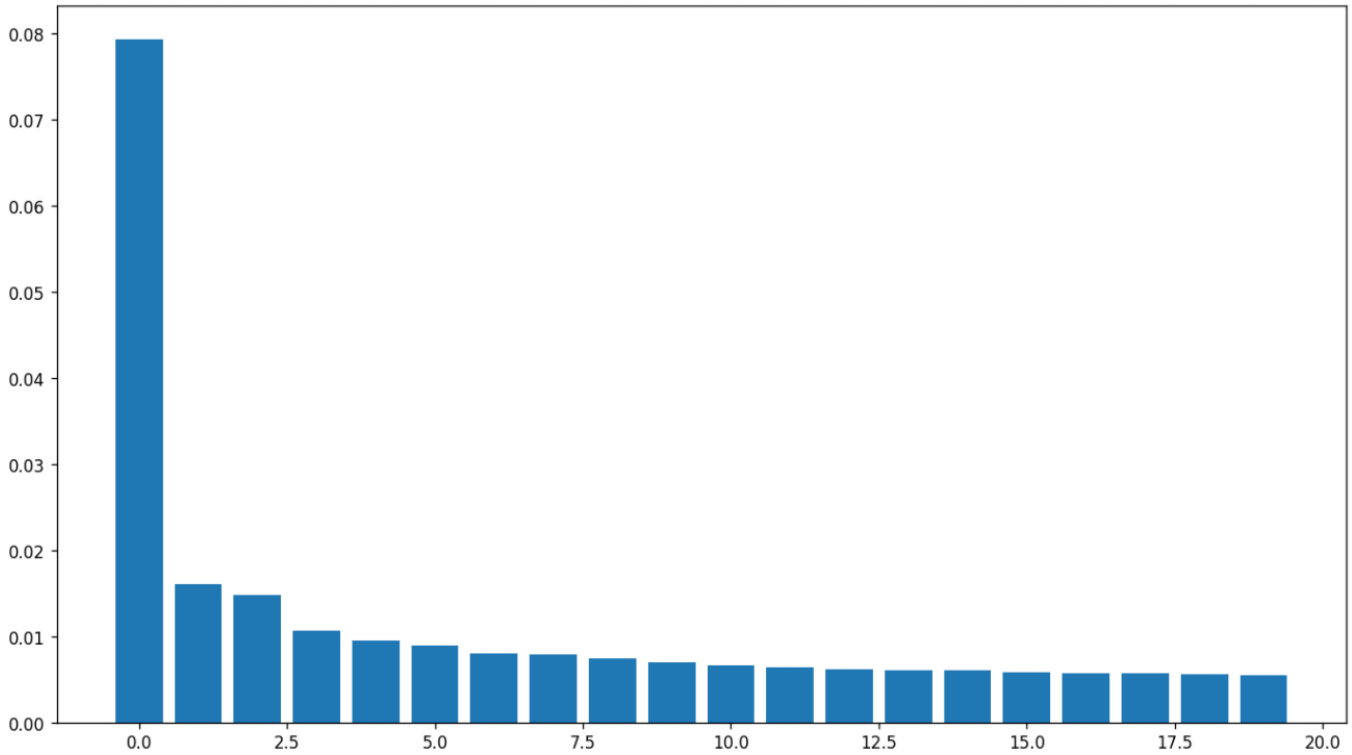
# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Risk Model

PCA squeezes stock dataset from (2517x621) matrix to (621x20) components.

Total variance of data is explained by 20 components (factor_betas) is about 23% which is low. It because of missing values from at least 30% of the universe has less than 5 year life long.

Top 5 stocks contribute the most in the variance of the first principal component
Equity(572 [VIG]) 9.6%, Equity(237 [HUT]) 10.3%, Equity(420 [SCR]) 9.8%, Equity(443 [SHS]) 11.7%, Equity(355 [PFL]) 10.2%
Sum square of stocks' contribution in each component approximates 1

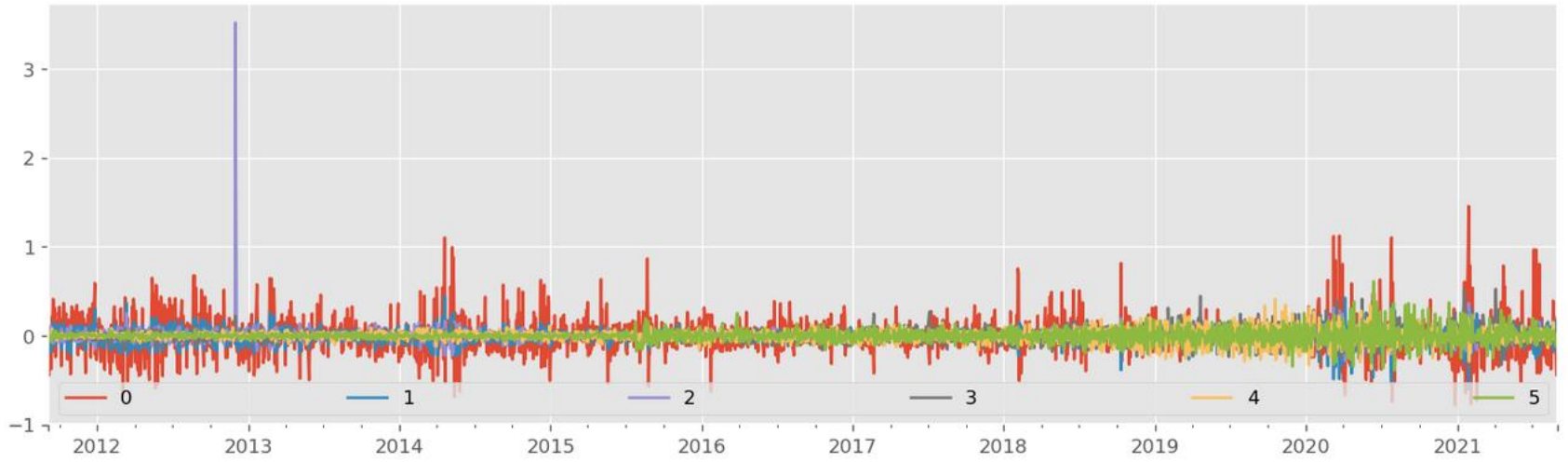



Springboard

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Risk Model

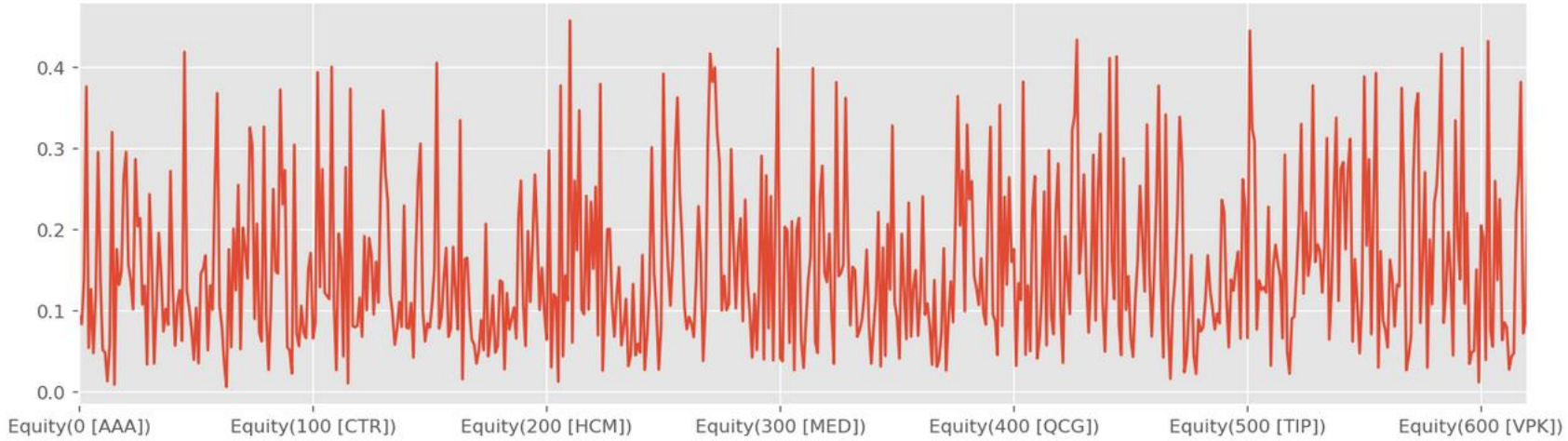
**factor_returns** is the returns time series for some kind of latent or unknown driver of return variance.
The first 06 components fator_returns




# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Risk Model

**idiosyncratic variance vector** over 621 stocks, they are alpha risks that are related to each individual stocks, not the market.




**Portfolio risk** prediction for equal weights **SQRT($X^T(BFB^T + S)X$) = 10.52%**

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model

Alpha factors and risk factors can be interchangeable by time and markets. Once alpha becomes publicly known and available, its profitable power will depreciate and turn to be a risk factor and vice versa. Alpha model is built & assessed using **Quantopian** API.

Raw alpha factors are created by **Zipline** module then processed for sector neutralizing, ranking & zscore

Processed alpha factors are modeled & assess the performances by **Alphalens** module

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model – Alpha Factors Processing

Raw alpha factors are processed for sector neutralization, ranking & zscore

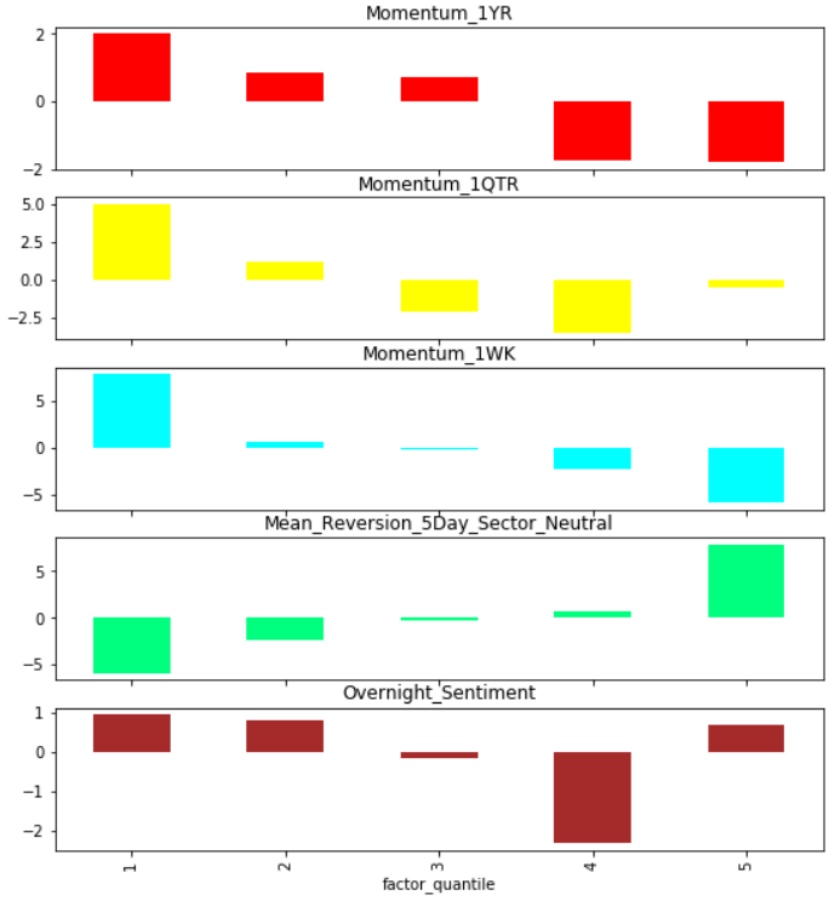
10 alpha factors are computed for a 5 year period:
- Momentum 1 Year, 1 Half Year, 1 Quarter, 1 Month, 1 Week Factors
- Mean Reversion 5Days, 20Days Sector Neutral Factors
- Mean Reversion 5 Day Sector Neutral Smoothed Factor
- Overnight Sentiment Factor
- Overnight Sentiment Smoothed Factor

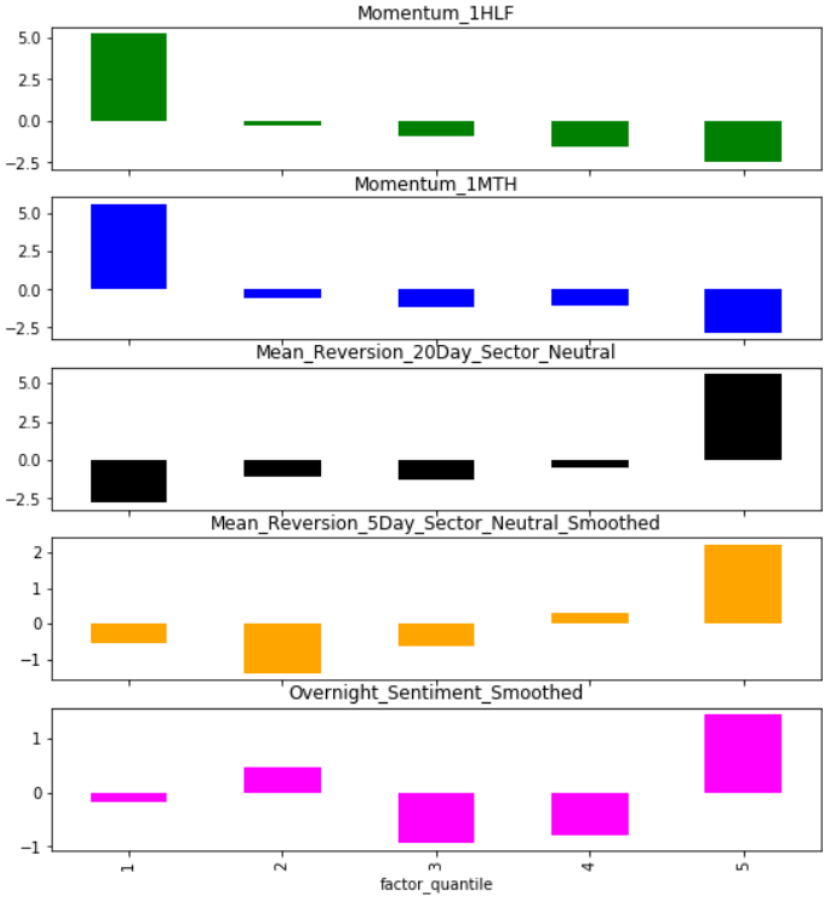
# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model – Factor Operation

Processed alpha factors & pricing data are used to compute 1 day forward returns & classify to 5 quantiles:

- Momentum returns performances seems working well in the 1st quantile
- Mean reversion & sentiment returns performances are good at the 5th quantile



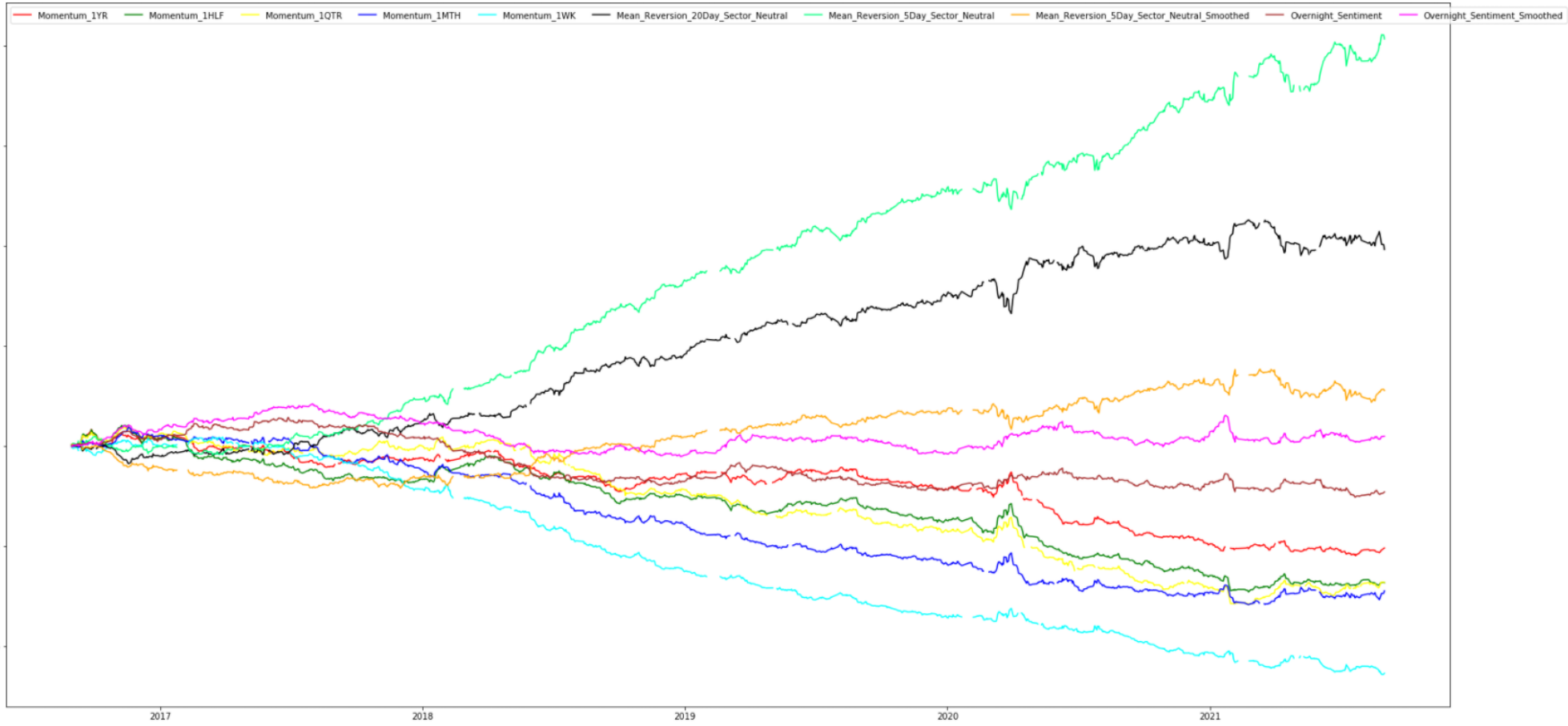



# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model – Factor Performance

Alpha factors cumulative returns for all quantiles.

- Momentum factors perform from weak to negative
- Sentiment factors show a fair to week performance
- Mean reversion factors perform the best



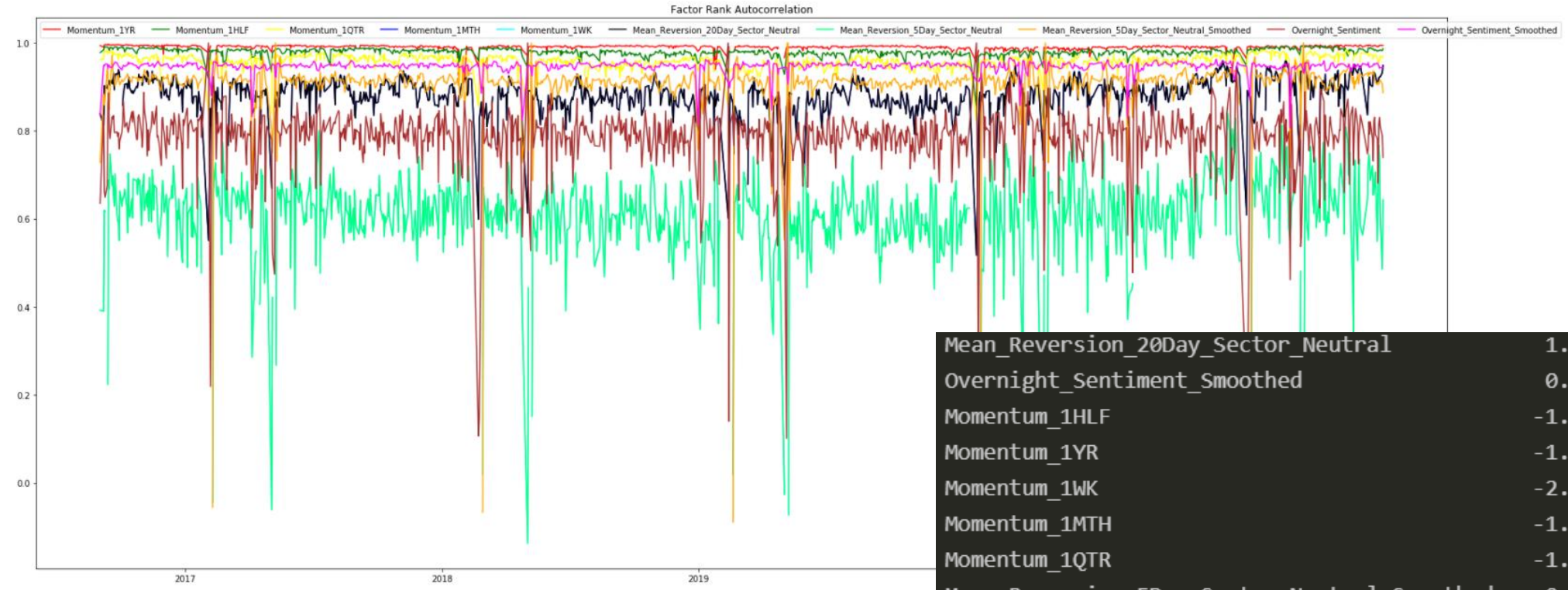
# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model – Factor Assessment

Turnover & Sharpe Ratio are two key indicators  to assess alpha factors performances:
Turnover or Factor Rank Autocorrelation expresses how persistent the factors are for future prediction. The closer to 1 the more consistent performance of a factor is & less trading

- Momentum & smoothed factors are showing a stable status
- Mean reversion & sentiment factors are unstable

Sharp Ratio shows how well the factor making profit. Mean reversion factors are good candidates




| | |
|---|---|
| Mean_Reversion_20Day_Sector_Neutral | 1.44 |
| Overnight_Sentiment_Smoothed | 0.20 |
| Momentum_1HLF | -1.61 |
| Momentum_1YR | -1.02 |
| Momentum_1WK | -2.77 |
| Momentum_1MTH | -1.44 |
| Momentum_1QTR | -1.48 |
| Mean_Reversion_5Day_Sector_Neutral_Smoothed | 0.35 |
| Overnight_Sentiment | -0.47 |
| Mean_Reversion_5Day_Sector_Neutral | 2.77 |

Springboard

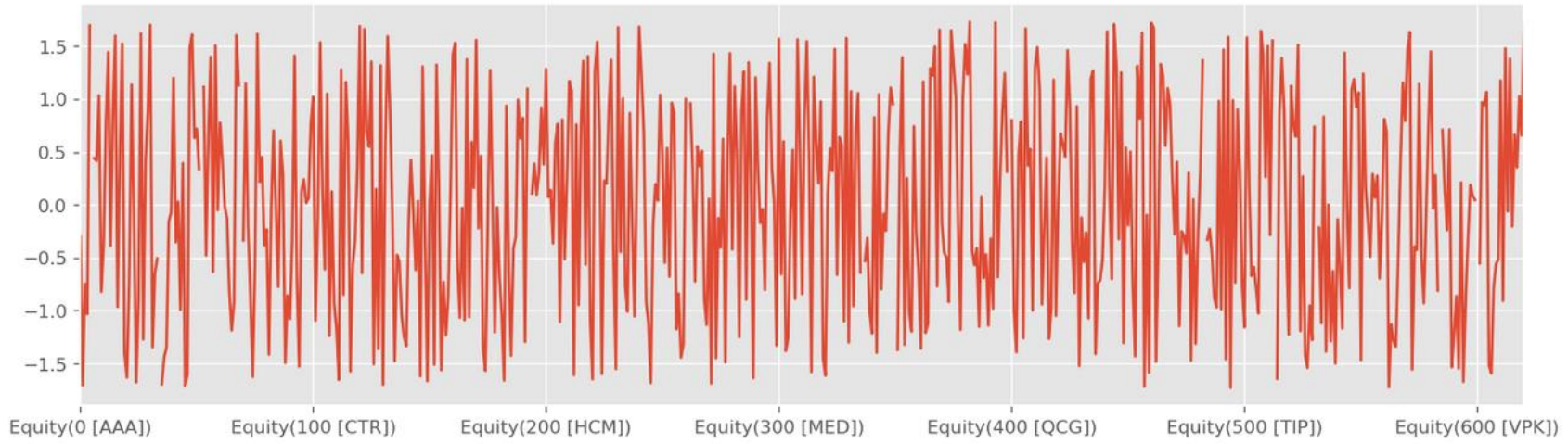
# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Alpha Model – Factor Assessment

Regardless of low FRA, Mean_Reversion_5Day_Sector_Neutral & Mean_Reversion_20Day_Sector_Neutral Factors are still selected as input for Multi-Factor Model.

They are simply blended by averaging to the alphas. Below is the alpha vector of the last day 2021-08-31 of all 621 stocks that we think will be proportional to the next day returns

The higher alpha, the better arbitrage opportunities . Multi factor model need to preserve these alpha from market & idiosyncratic risks by weight optimizing process.



# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Multi Factor Model

Alpha model and risk model are available. Let's find a portfolio that trades as close as possible to the alpha model but limiting the risks as measured by the risk model by optimizing the weights under risks constrain.

**Multi factor model comprises objective function & its constraints**

Where**:**

- Objective function is maximizing (**$\alpha^T * x$**) with **x** is the portfolio weights and **α** is the alpha vector.
- Constraints are risk model as following:

r is the portfolio risk, $risk^2_{cap}$ is market (vnindex/ETF…) volatility

B is the factor betas, $factor_{max,min}$ are max & min portfolio factor exposures

sum of weights = 0 (market neutral)

leverage (borrow cash/stocks) constraint <= 1

minimum and maximum limits on individual holdings, kind of insurance.

$$r \leq risk^2_{\text{cap}}$$
$$B^T * x \preceq factor_{\max}$$
$$B^T * x \succeq factor_{\min}$$
$$x^T 1 = 0$$
$$\|x\|_1 \leq 1$$
$$x \succeq weights_{\min}$$
$$x \preceq weights_{\max}$$

Output is optimum weights scheme for a period of time (1 day forward for this case). Negative weight is short position, positive is long position

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Multi Factor Model

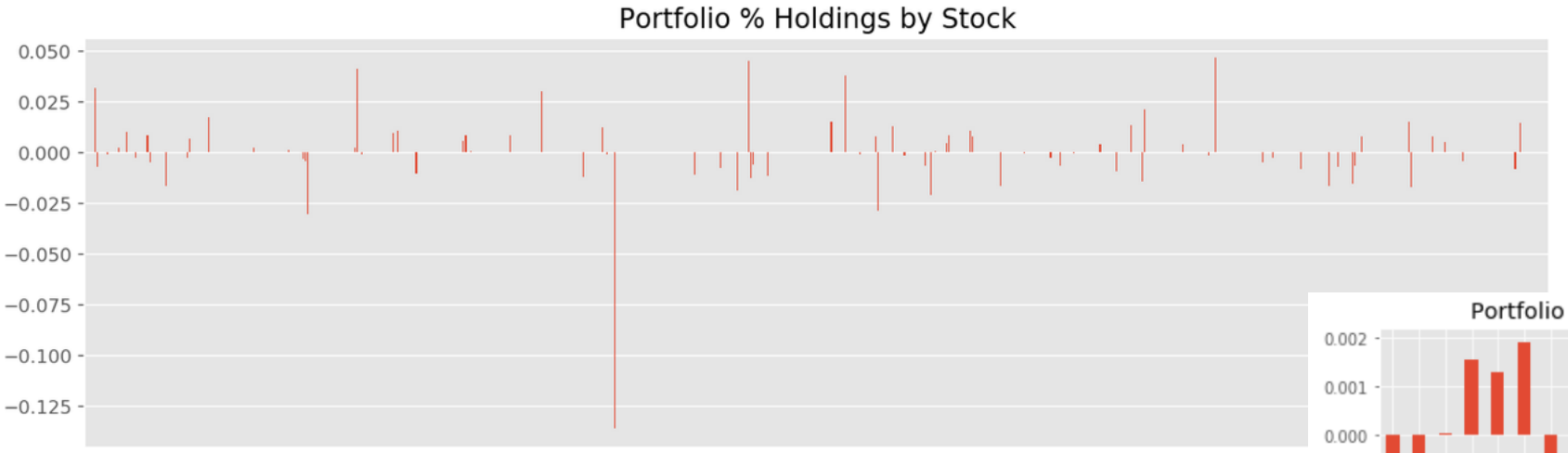
**Optimizing object function ($\alpha^T * x$)**

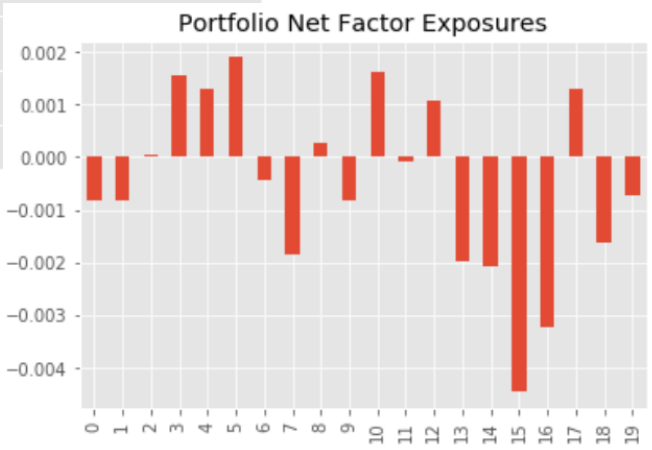
Default portfolio constraints: risk_cap=0.05, factor_max=10, factor_min=-10, weights_max=0.55, weights_min=-0.55

- Weights allocation seems not diversity, not optimum for lowering risks
- Weights allocation for the case of portfolio contains 20 components

**Portfolio risk** prediction for optimum_weights is **$SQRT(X^T(BFB^T + S)X)$ = 5.0%**

**Transfer coefficient = 20.4%,** optimum_weights is limited match with alpha_vector




Springboard

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Multi Factor Model

**Optimizing object function** $(\alpha^T * x) + \lambda||x||_2$

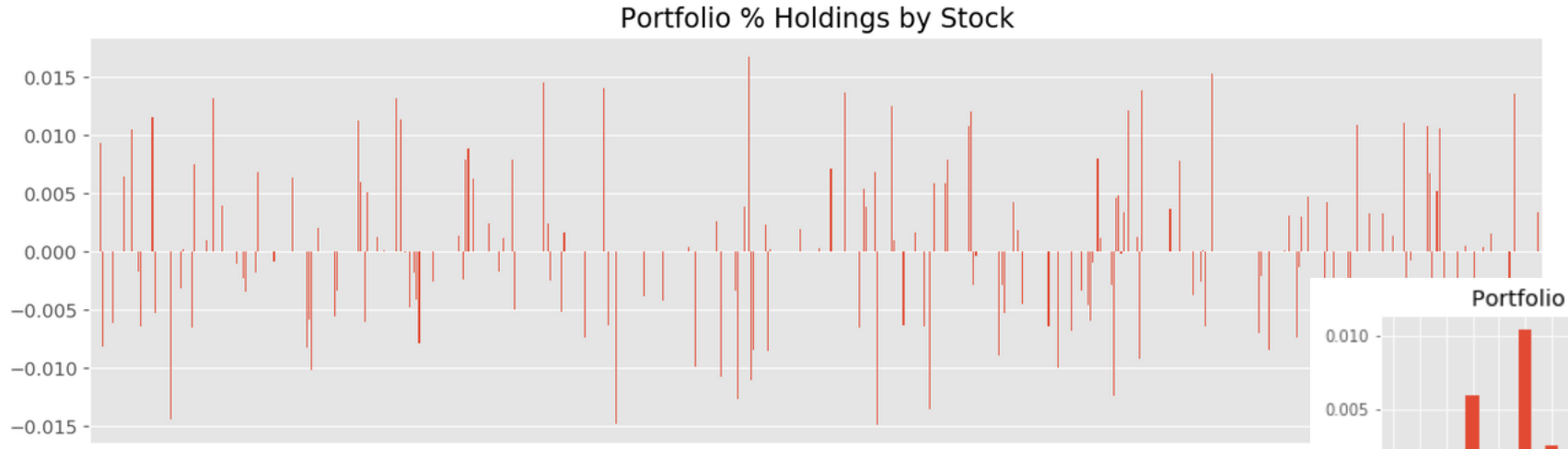
Regularization parameter penalize the extreme weights & re-attribute to the whole portfolio

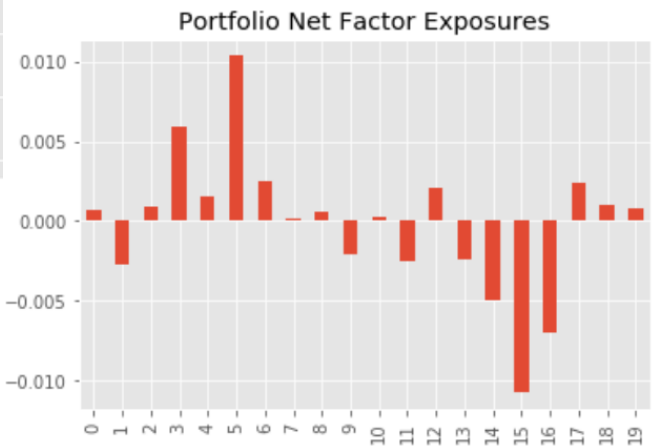
Default portfolio constraints: **λ** =0.5, risk_cap=0.05, factor_max=10.0, factor_min=-10.0, weights_max=0.55, weights_min=-0.55

- Weights allocation is now pretty diversified
- Weights allocation for the case of portfolio contains 20 components

**Portfolio risk** prediction for optimum_weights_1 is **SQRT($X^T(BFB^T + S)X$)** = 4.5%

**Transfer coefficient = 35.5%,** optimum_weights_1 fairly matched with alpha_vector




Springboard

# Project Quantitative Analysis – Vietnam Stock Exchange

## Model Construction – Multi Factor Model

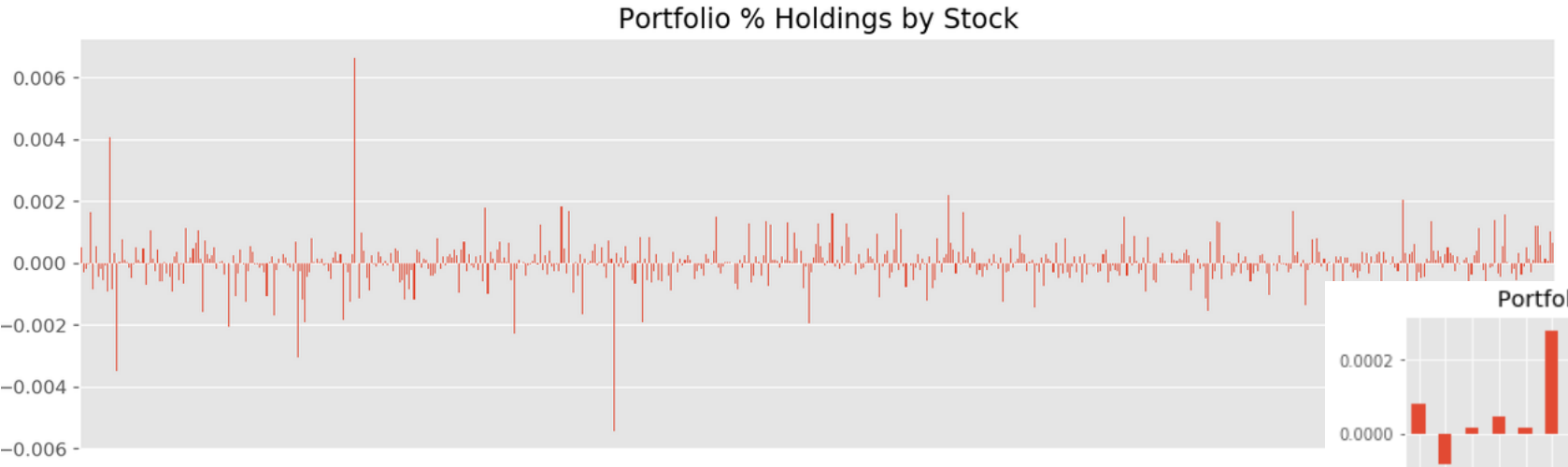
**Optimizing object function** $-(\alpha^T * x) + ||x - x^*||_2$

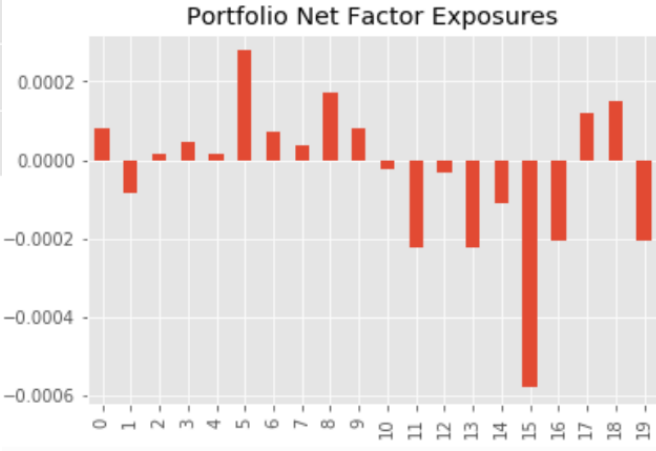
Specific portfolio constraints: weights_max=0.02, weights_min=-0.02, risk_cap=0.0015, factor_max=0.015, factor_min=-0.015

- Portfolio weights scheme is much more diversified
- Weights allocation for the case of portfolio contains 20 components

**Portfolio risk** prediction for optimum_weights_2 is **SQRT($X^T(BFB^T + S)X$)** = 0**.4%,** much lower than equal weight scheme (10.5%)

**Transfer coefficient = 31.3%,** optimum_weights_2 fairly matched with alpha_vector




Springboard

# Project Quantitative Analysis – Vietnam Stock Exchange

## Conclusions

**Project achievements:**

- Successfully ingest Vietnam stock data to Quantopian API format
- Conventional quant effect analysis: Bollinger band, breakout & momentum
- Fully create risk model for universe of 621 stocks
- Create 10 alpha factors model, operating & performance assessment: **Mean reversion factors outperform** for the universe of 621 stocks working for 1 day looking forward
- 03 cases of weight optimization under 02 risks constraints cases. The 3rd weights scheme reflects the best real trading condition.

**Discussion:**

- Predicted portfolio risk for non-managed case is 10.52% and fully managed case is 0.4% which is a significant market risk neutral
- Selected alpha factors have low turnover indicator(FRA) which is not good for future prediction
- Full-managed case's weights allocation is well diversified, but fair transfer coefficient (31.4%) which is a not good indicator of well-preserved alpha.

# Project Quantitative Analysis – Vietnam Stock Exchange

## Recommendations

**for Retail investors, Fund managers, Portfolio managers:**

Stock universe (candidate selection) should be carefully selected ex: filtered by

- size (volume),
- value (price),
- other fundamental factors : dividend, P/E…
- or even following ETF indexes: vnindex, vn30index…

Recommended further steps:

- create more alphas,
- apply weight scheme for alphas selected by sharpe ratio or other alpha performance indicators,
- back-testing under real trading condition in order to fully assess how effective, persistent & resilient the factor models are