



- Risk is an important determinant of investment in stock market.
- Volatility is the measure of how wildly the stock return swings.
- However, volatility is not constant over time.
- Isn't it great to know the dynamics of volatility itself?

Two behaviors commonly found:
1) volatility clustering
2) Asymmetric responses to shocks

- Approaches
1. Develop the proxy of variance (LR, RF, RNN)
 2. Infer the variance from residuals (ARCH, GARCH)

Trade-offs

clustering

$$\sigma_t^2 = f(\sigma_{t-1}^2, \varepsilon_{t-1}^2, \dots)$$

asymmetry

- NOT observed
- Measure the deviation from return mean
- Trade-off in choosing est period (At least 500)

Problem: Our data are not ready to be used in modelling yet. (80% of time used in cleaning)

Algorithm & Data Structure comes into rescue

sector: Banking	
Price	SCB(P) KTB(P)
Volume	SCB(VO) KTB(VO)

Cleaning Tasks (Pandas-Python):

1. Combine: merge price + vo (regex & database)
2. Split: from sector \rightarrow stock (Hash & filter)
3. Combine: from stock \rightarrow group (Sort, apply & reduce)

Result: 1. Volatility dynamics exist in individual level. \rightarrow study dynamics at ind level helps.
2. Relationship between inv period and vol dynamics

