Artificial Intelligence in Fintech Quiz (1)

Shaif Chowdhury

February 2, 2022

1 Question 1

Summarize singular value decomposition (SVD) and its possible applications in Fintech (10 points)

SVD is a matrix factorization technique that would represent a matrix as a product of two unitary matrix (U,V*) and a rectangular diagonal matrix ($\sum$), $i.e : X = U \sum V *$.

SVD has application across different range of problems like least square, pseudo inverse and data types like image, remote sensing, financial data, time series etc.

Here U is formed of orthonormal eigenvectors of XX^T,

V* is transpose of that i.e : X$^T X$

All matrix have a SVD, which makes it a good option compared to other ways of decomposition.

Application of SVD in Fintech :

1. SVD can help reduce dimensionality of financial data using decomposition.
2. SVD can help us measure the spread/concentration of variance in the data.
3. SVD can be used to do principle Component Analysis, which helps in visualization, analysis and prediction of financial data.
3. SVD [1]is useful in multi-variate time series data, ie : daily stock data.

2 Question 1

APPL, BAC, WMT, AEO, EU-Option, Option-OTM Visualize data with any methods you like

1. Data Visualization : AAPL

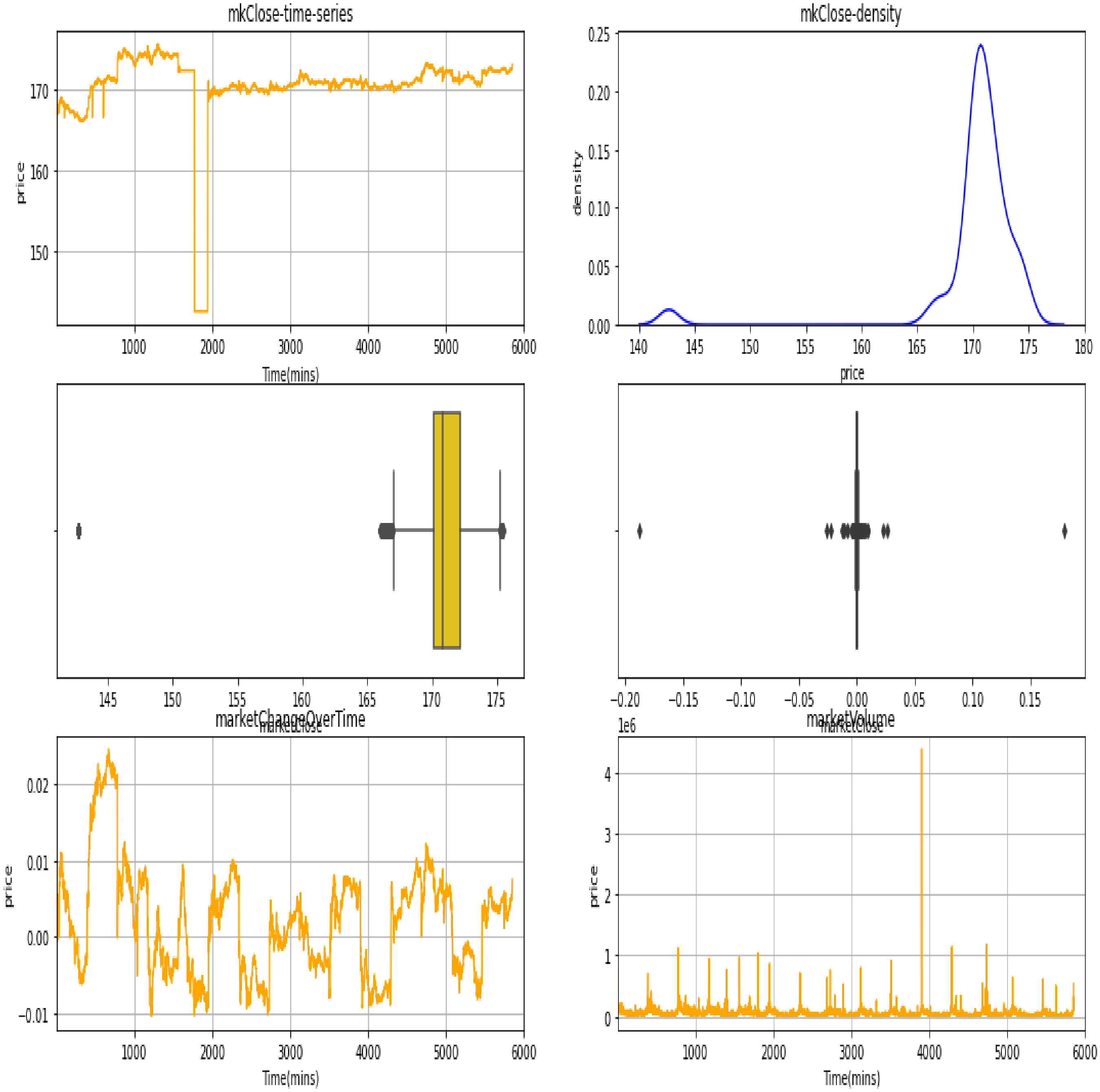


Figure 1: This is a plot for market Close, Change over time and volume.

2. Data Visualization : BAC

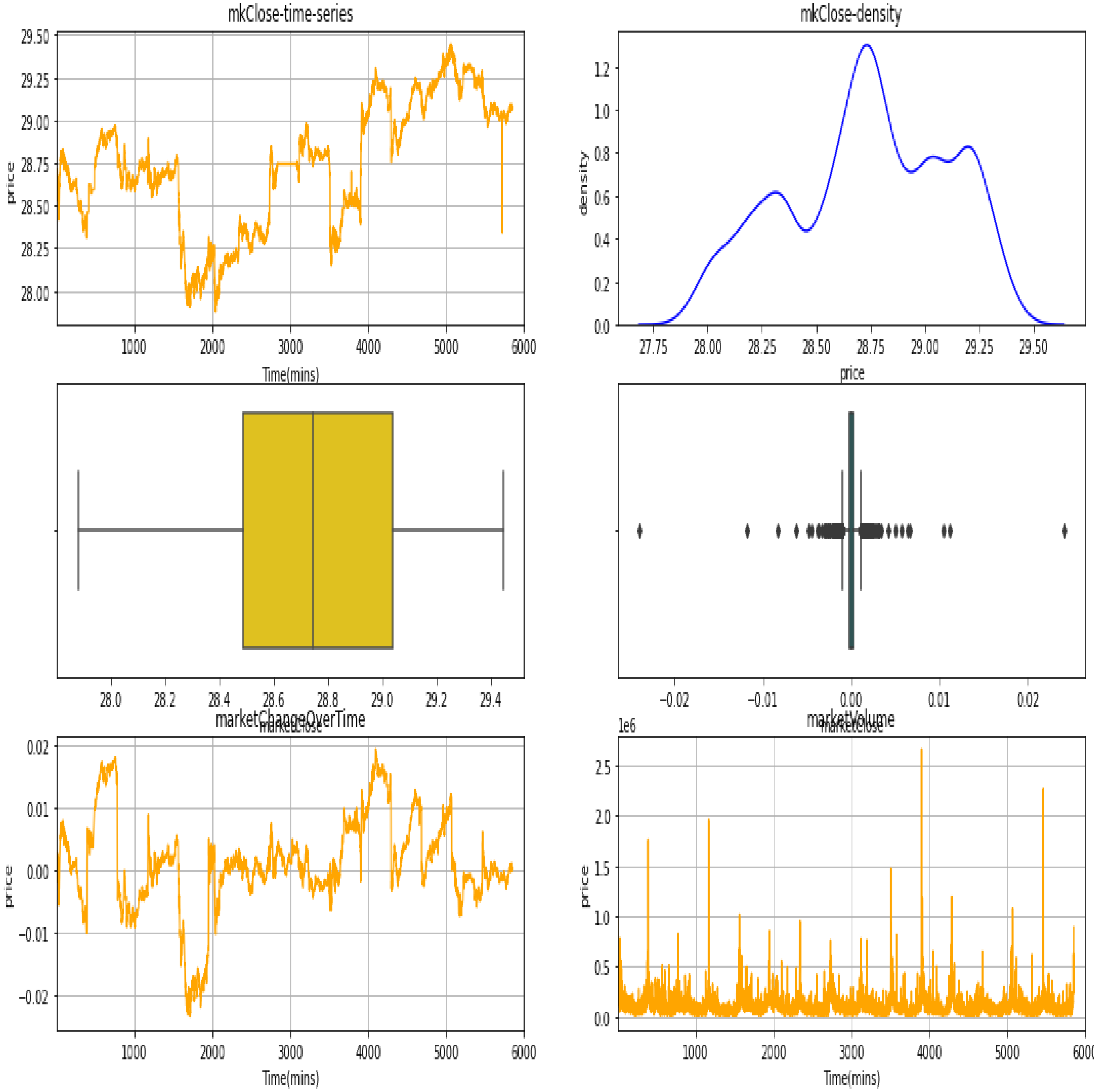


Figure 2: This is a plot for market Close, Change over time and volume.

3. Data Visualization : WMT

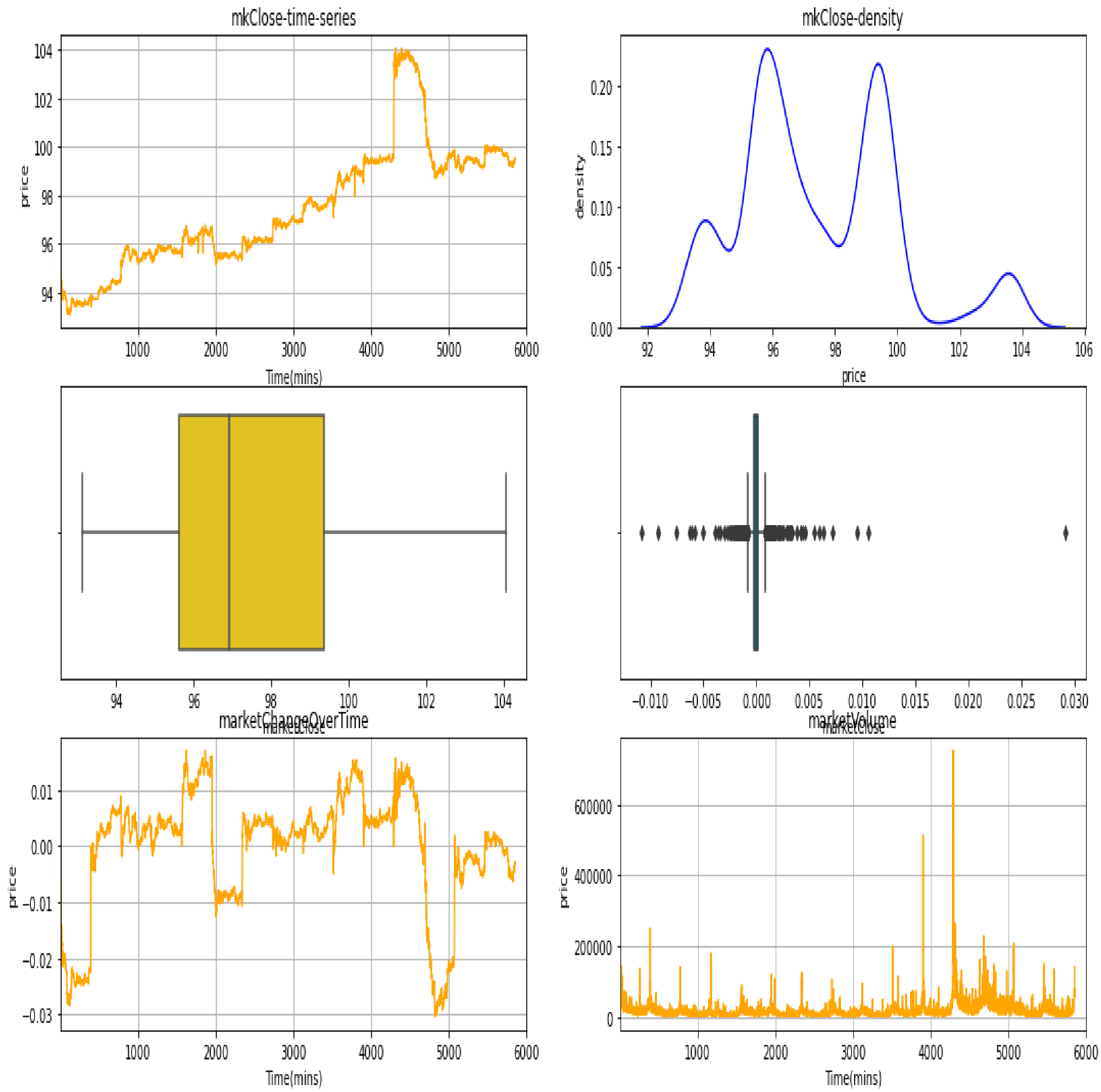


Figure 3: This is a plot for market Close, Change over time and volume.

4. Data Visualization : AEO

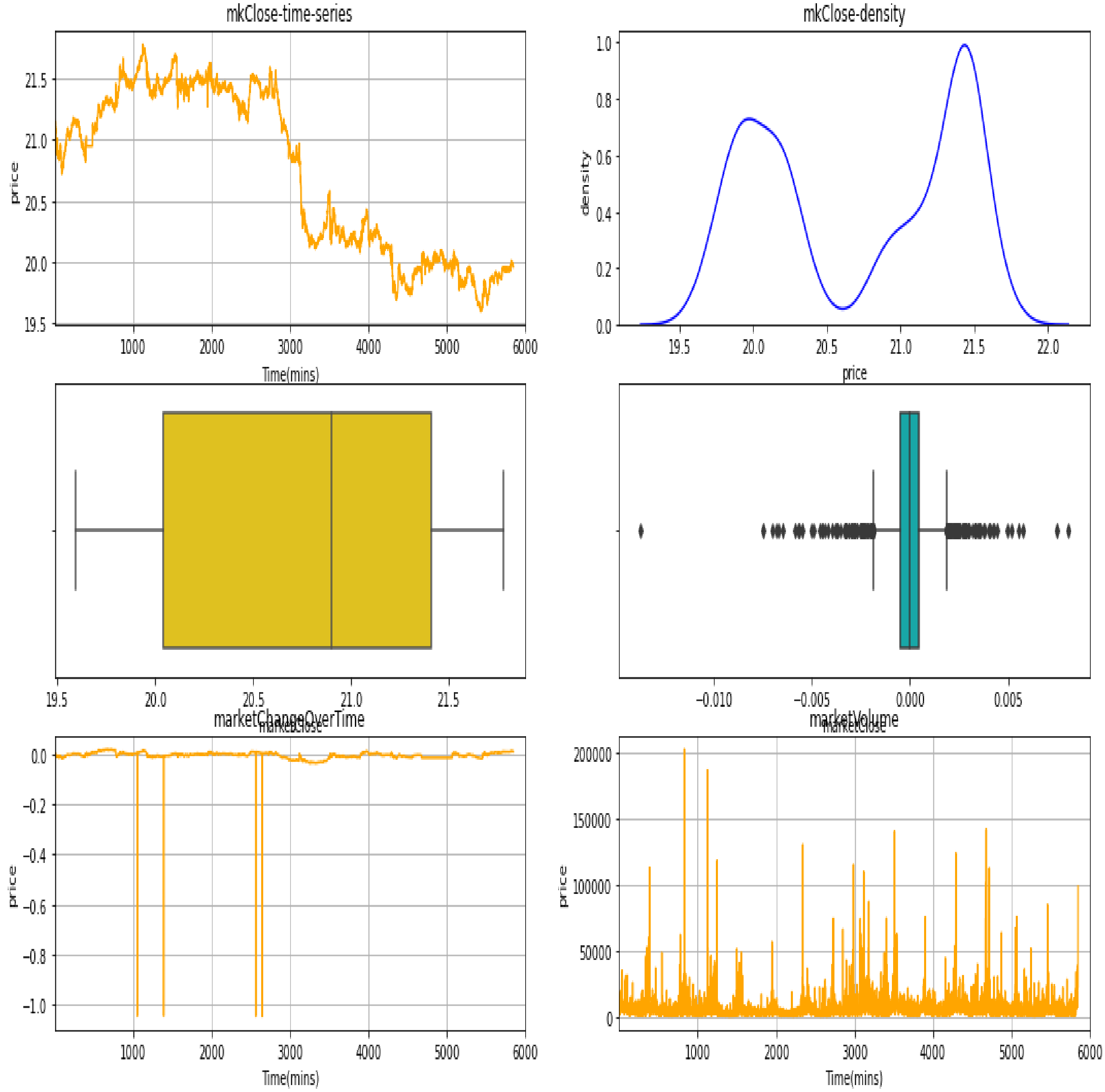


Figure 4: This is a plot for market Close, Change over time and volume.

3 Question 2,3

2. Compute the variance concentration ratios of data and visualize it. What can you find?

3. Compare the variance concentration ratios of the normalized data of the data by using different normalization methods

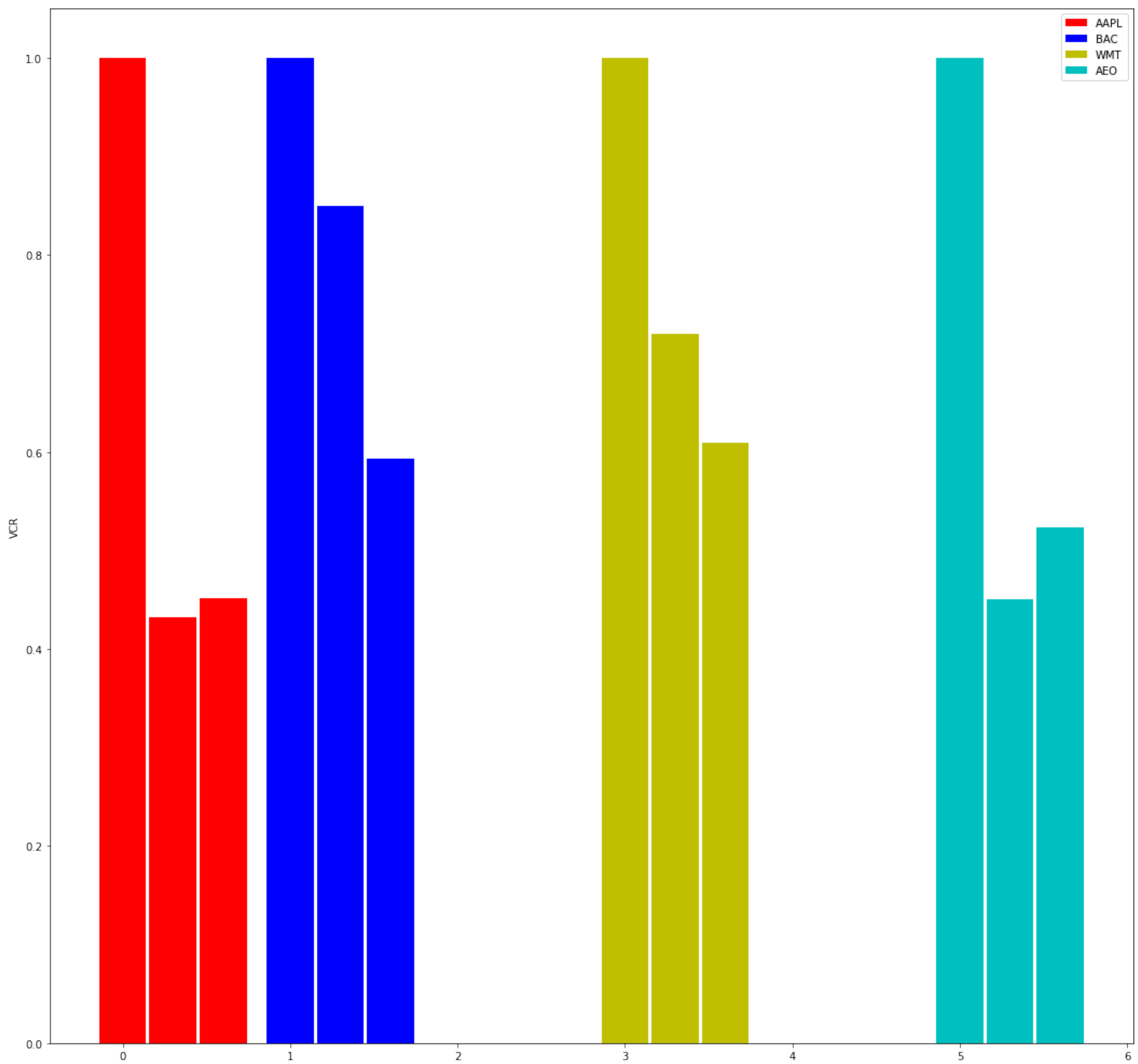


Figure 5: The first bar is for VCR, second for MinMax VCR and third for StandardScaler VCR.

4 Question 4

2. Apply 1,2,3 to at least your own two datasets in finance

1. Data Visualization : GOOG

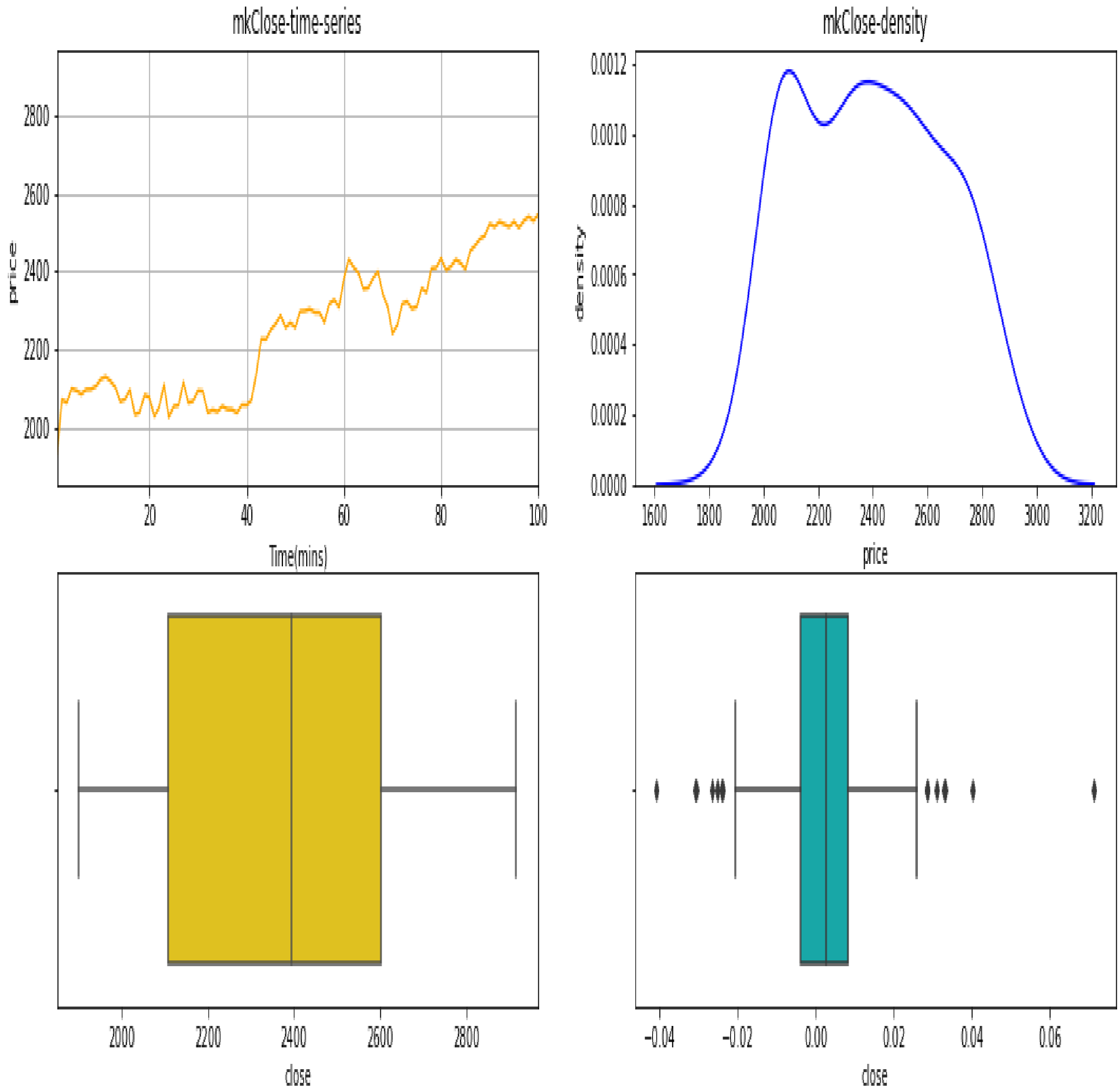


Figure 6: This is a plot for market Close, Change over time and volume.

2. Data Visualization : HSBC

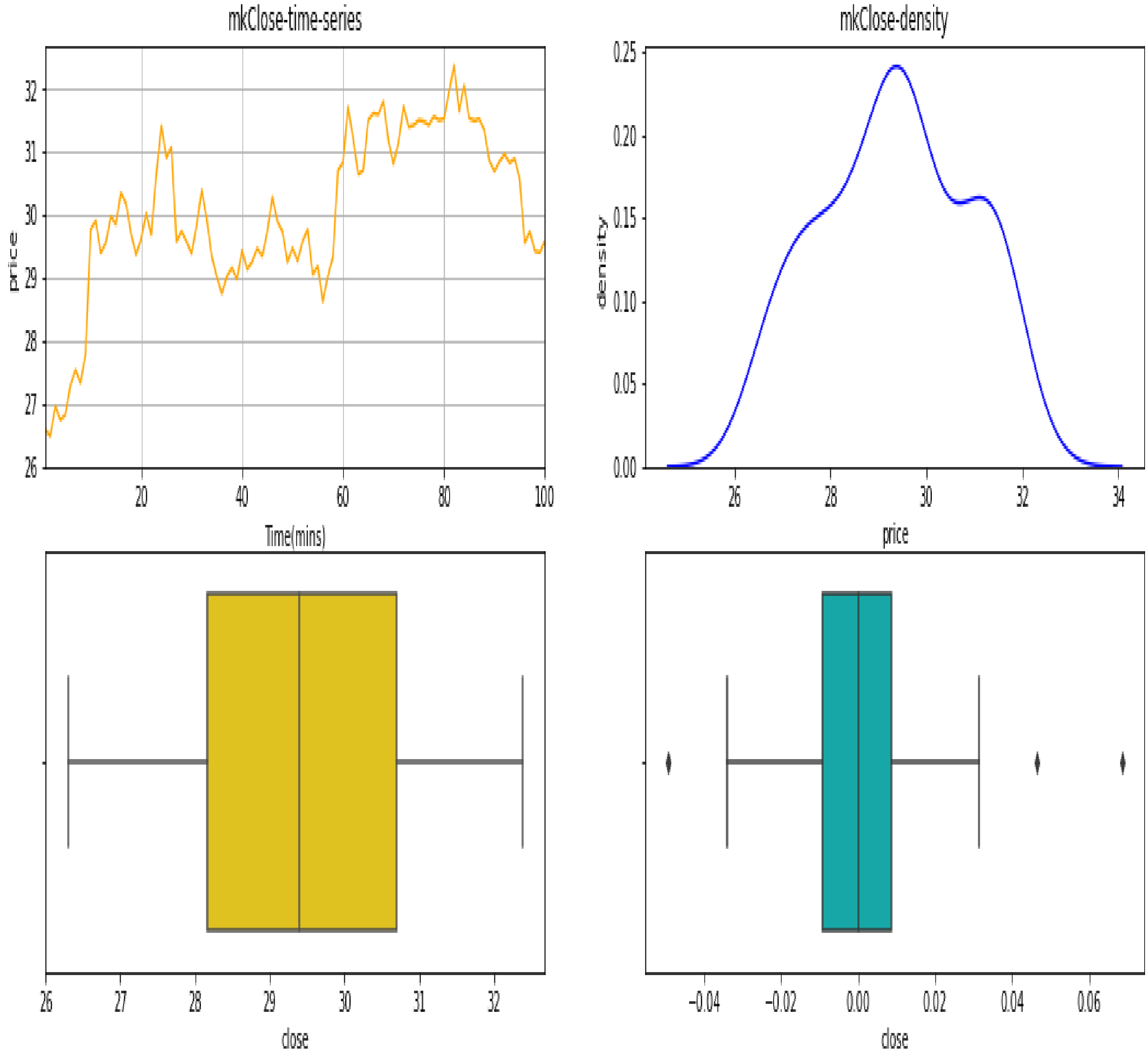


Figure 7: This is a plot for market Close, Change over time and volume.

VCR Plot for GOOG and HSBC

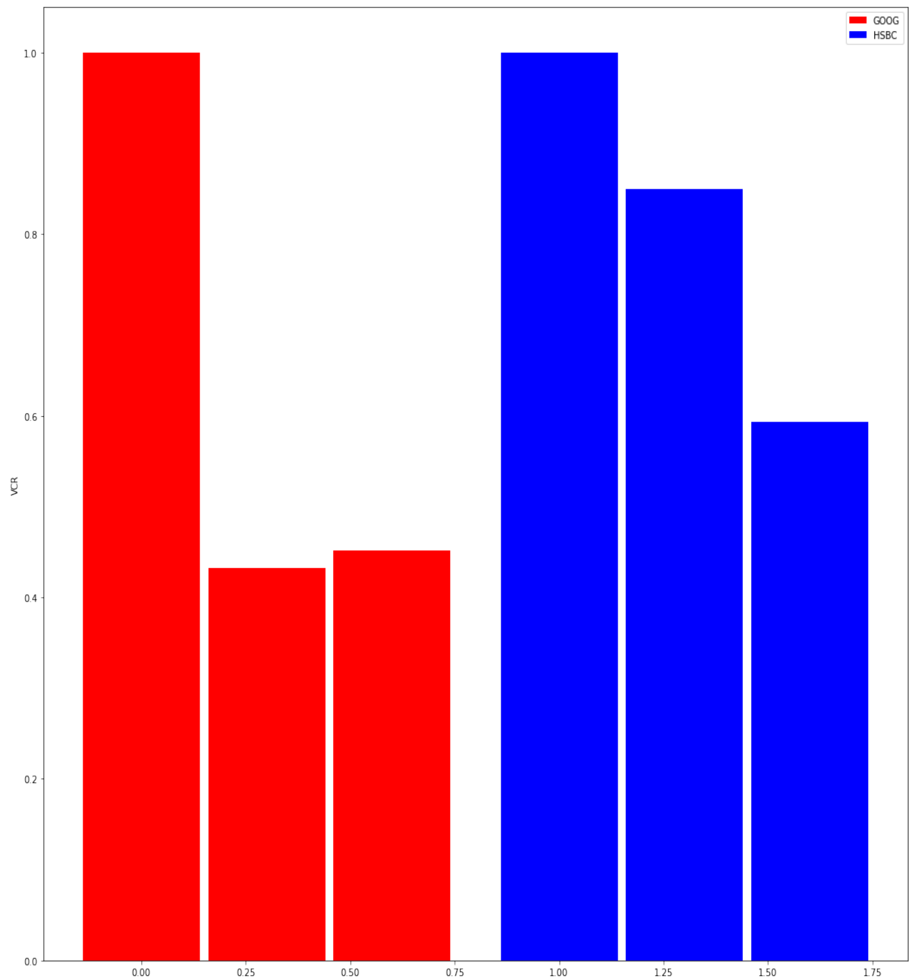


Figure 8: The first bar is for VCR, second for MinMax VCR and third for StandardScaler VCR.

5 Question 5

What are your conclusions? why?

Variance concentration ratios are very similar across different stocks for raw data. There is more difference in MinMax and Standard Scalar normalized data. So, normalization is really crucial to analyze variance concentration , which might give insights into volatility in stock price.

6 Question 1

1. Write software to retrieve recent 1 week HFT data for the following stocks

```
def retrieveStockH(ticker_name,  start_time, end_time, API):

  TI='af3fb5397b00b669904c3e1212800ed9b586c084'

  stock_data=web.get_data_tiingo(ticker_name,  start_time, end_time, api_key=TI)

  return stock_data

goog = retrieveStockH('AMZN','2/1/2021', '9/1/2021', api_key)
```

2. Compare their variance concentration ratios under different normalization methods, what can you find?

IT: GOOG, AAPL, MSFT, AMZN, FB

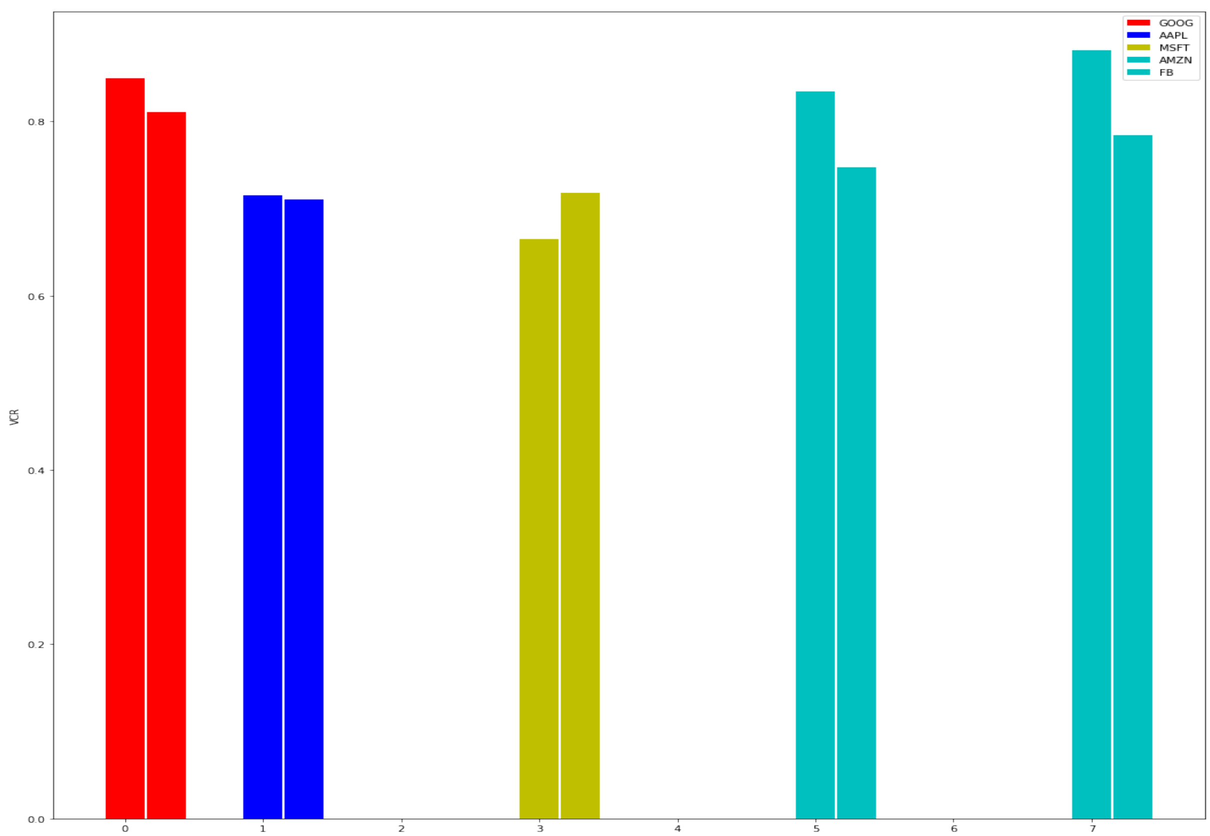


Figure 9: The first bar is for MinMax VCR and second for StandardScaler VCR.

Bank: 'JPM', 'BAC', 'C' , 'GS, 'HSBC'

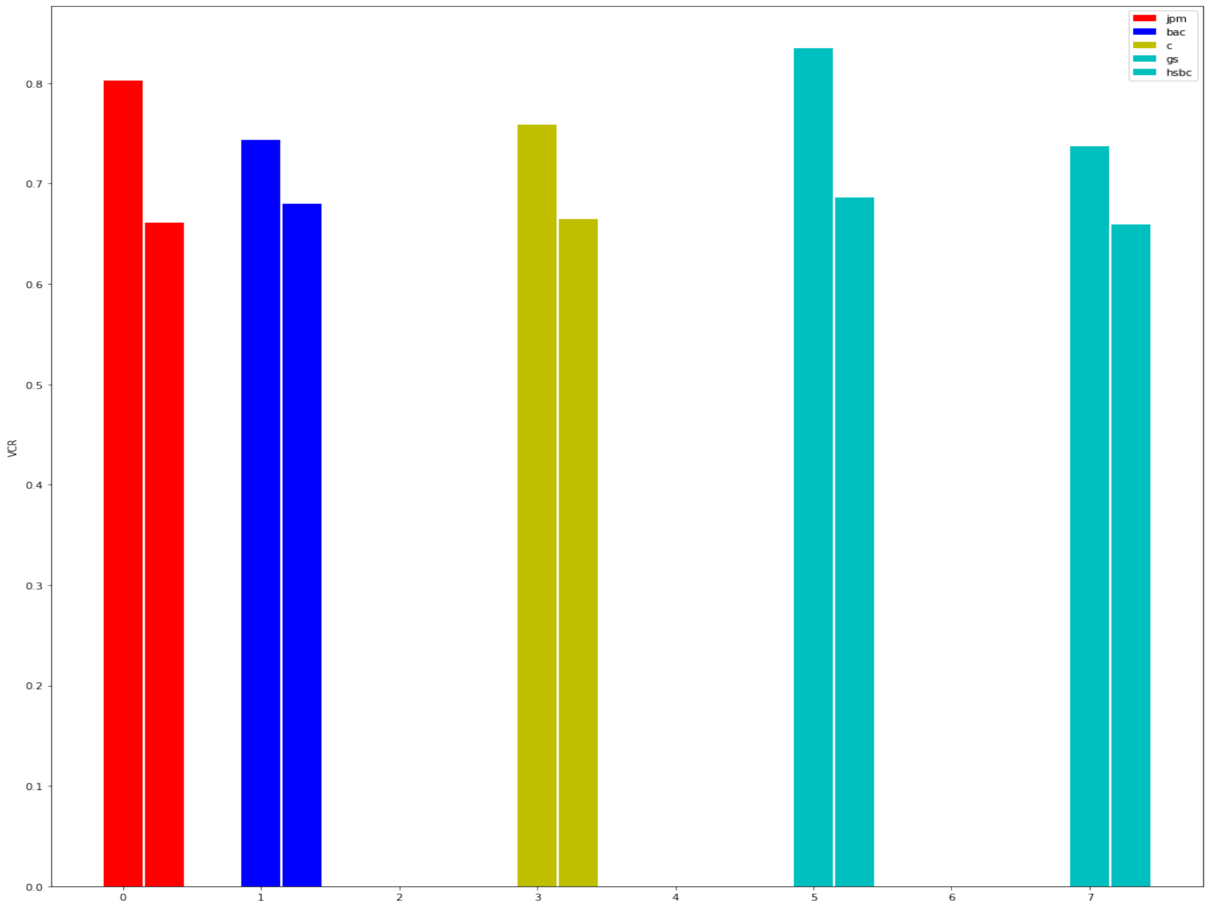


Figure 10: The first bar is for MinMax VCR and second for StandardScaler VCR.

Compare their variance concentration ratios under different normalization methods, what can you find?

There is more difference in MinMax normalized data. So, minmax normalization is probably better than scaling the data. A comparison with centering can be interesting too.

1. Calculate the stock log returns in each time interval for each stock and plot it

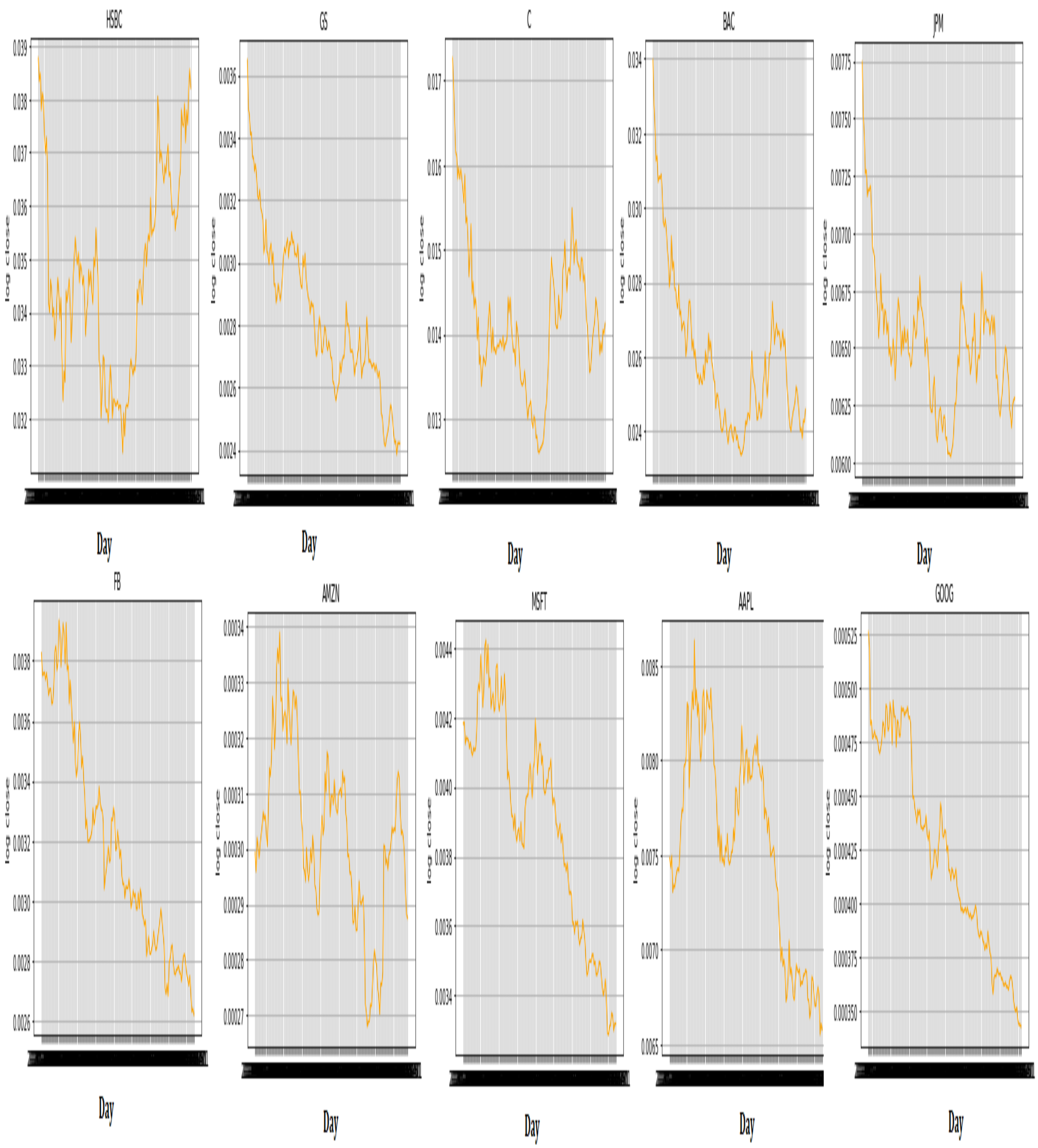
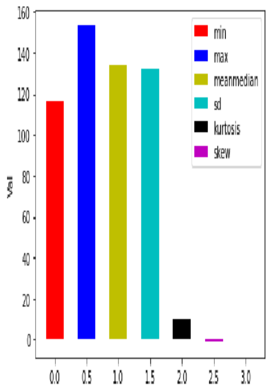


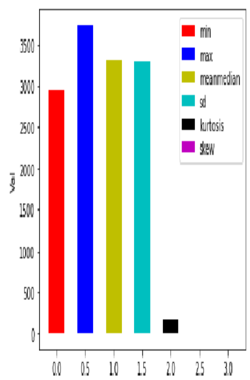
Figure 11: Log return for daily closing price.

2. Compute the max, min, mean, median, standard deviation, skewness, kurtosis for close price and volume for each data set (need plots).

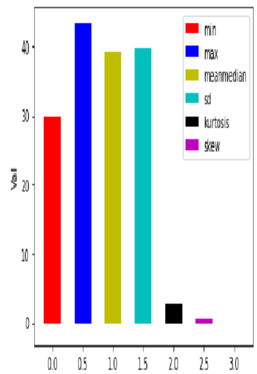
Close Price : Comparison by Stock.



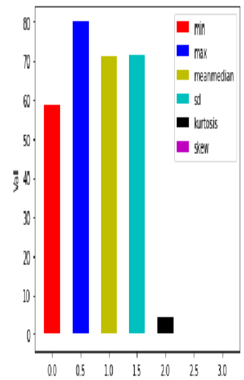
AAPL



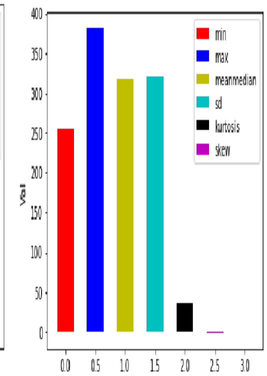
AMZN



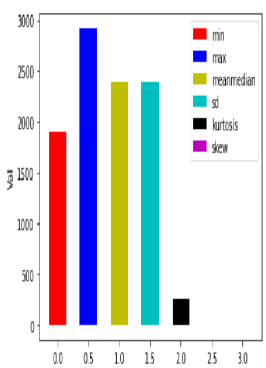
BAC



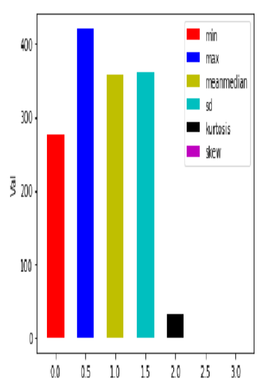
C



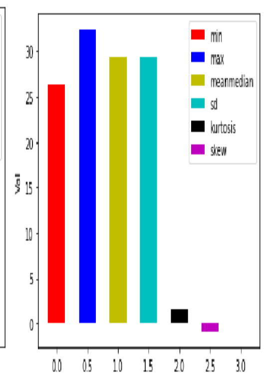
FB



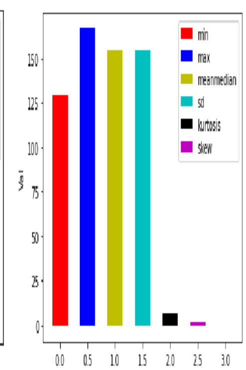
GOOG



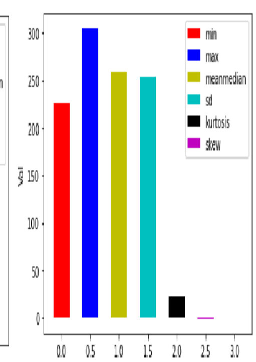
GS



HSBC



JPM



MSFT

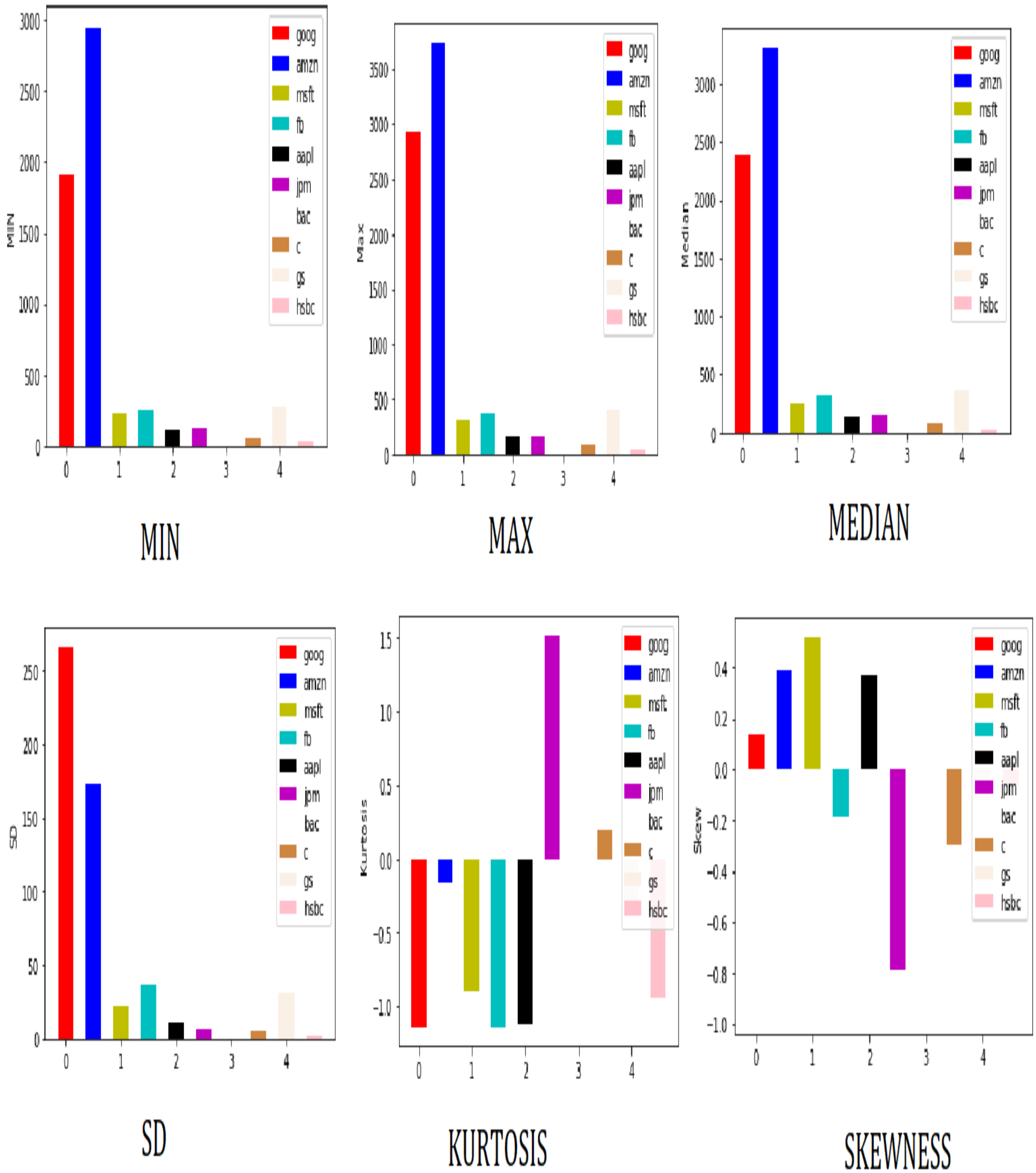


Figure 12: Comparison by Attribute.

Volume

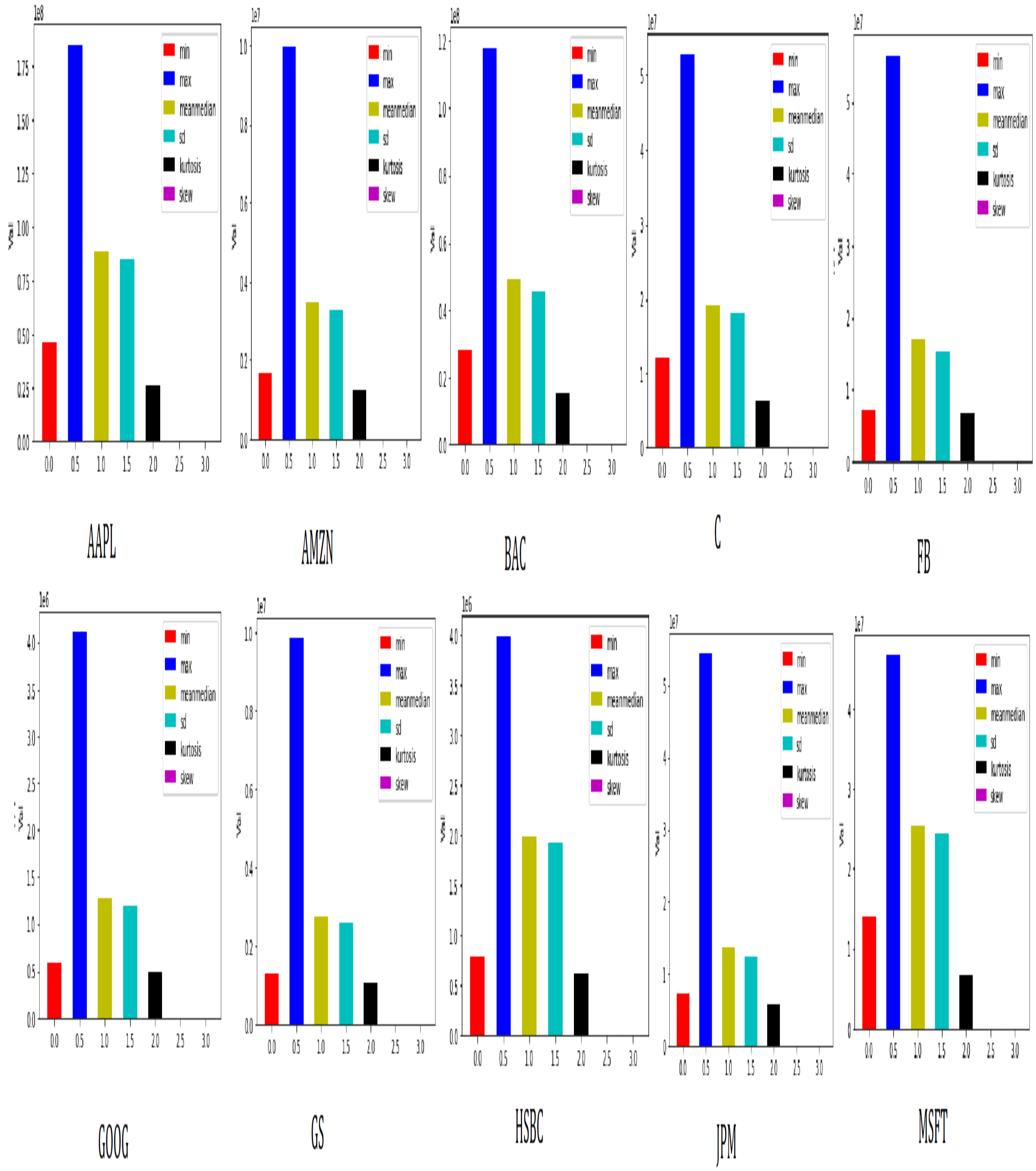


Figure 13: Comparison by Stock.

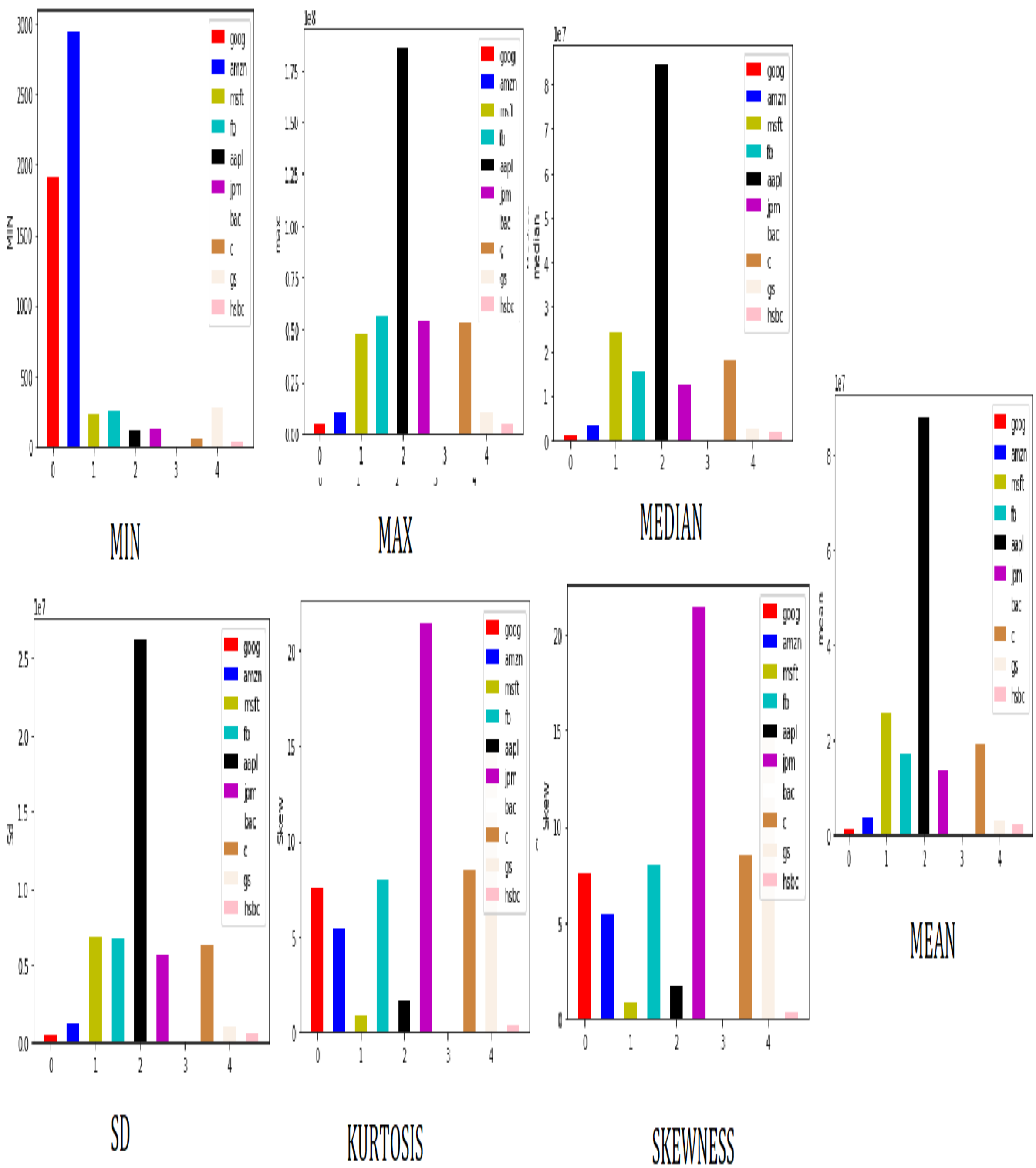


Figure 14: Comparison by Attribute.

1. Retrieve daily stock data for the following types of companies from 02/01/2015 to 02/01/2021and write it in csv – IT: GOOG, AAPL, MSFT, AMZN, FB – Bank: 'JPM', 'BAC', 'C' , 'GS, 'HSBC'

1. **Compare the stock price patterns of these companies during the years**

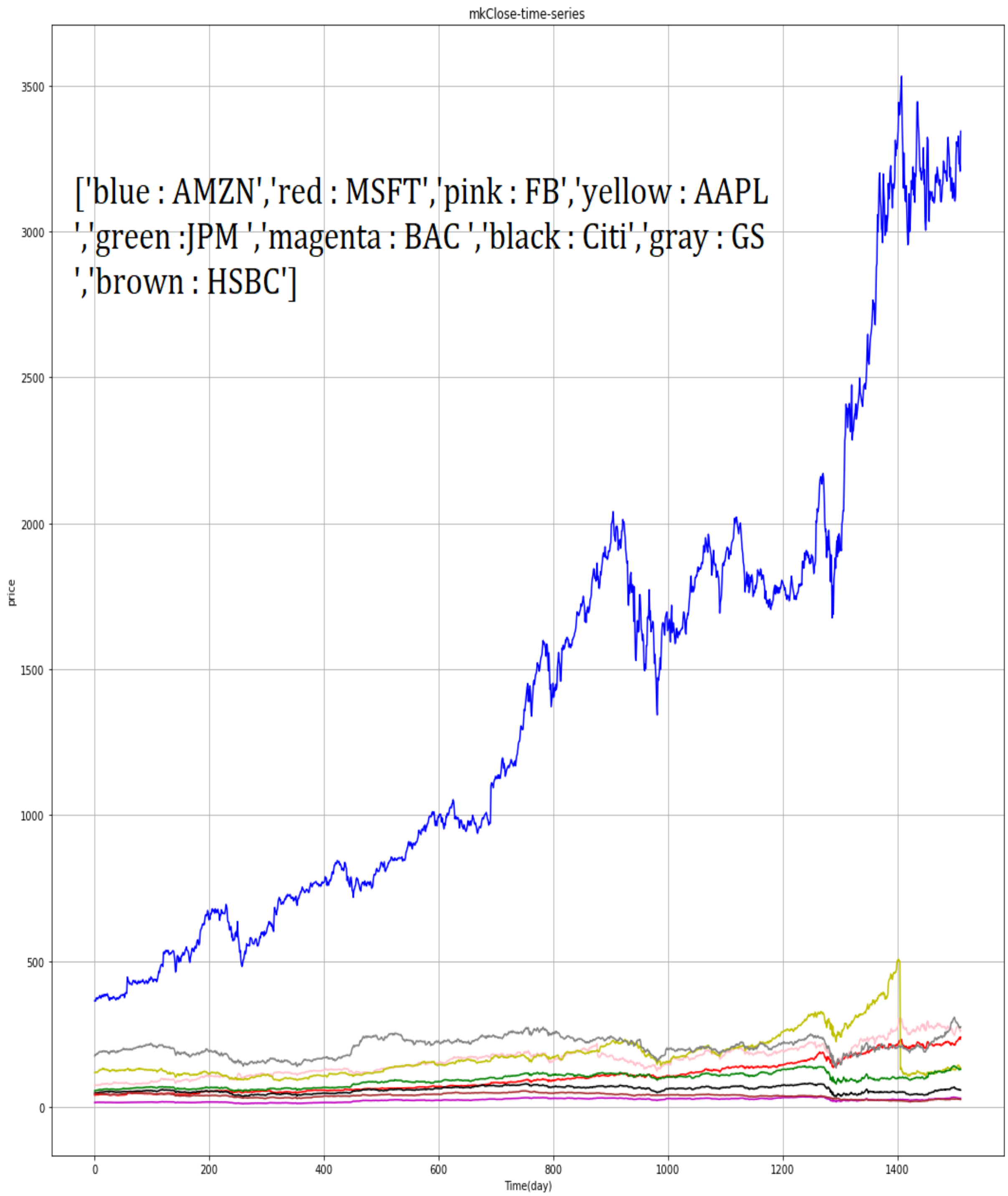


Figure 15: ['blue : AMZN','red : MSFT','pink : FB','yellow : AAPL ','green :JPM ','magenta : BAC ','black : Citi','gray : GS ','brown : HSBC']

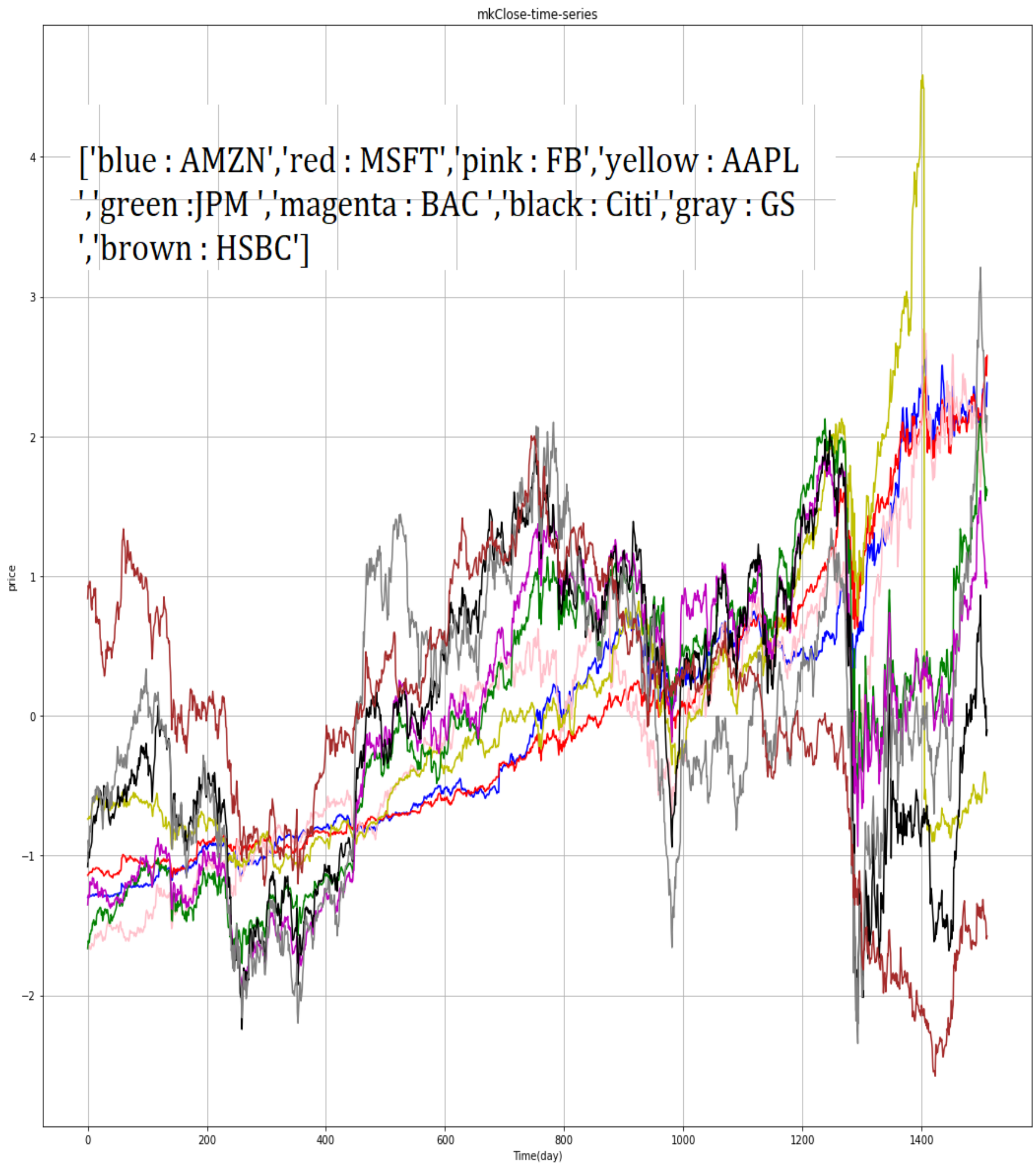


Figure 16: [Standard Sclaler Norm Plot

2. Calculate the days of up and down for each stock in each year.

Number of Days up and Down based on open and close price

GOOG : 88
AMZN : 753
MSFT : 696
FB : 730
AAPL : 696
JPM : 723
BAC : 736
C : 723
GS : 754
HSBC : 742

References

[1] Khoshrou2017SVDbasedVA, title=SVD-based visualisation and approximation for time series data in smart energy systems, author=Abdolrahman Khoshrou and Andre B. Dorsman and Eric J. Pauwels, journal=2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), year=2017, pages=1-6