GROUP WORK PROJECT # _2__ MScFE 622: Stochastic Modeling
GROUP NUMBER: 7372

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Alpha K. Kamara	USA	alpha.learn@outlook.com	
Eric Komla Anku	Ghana	ankueric1@gmail.com	
GABRIEL OMONDI ODERO	KENYA	gabriel.odero@yahoo.com	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an "X" above).	
Team member 1	**Alpha K. Kamara**
Team member 2	**Eric Komla Anku**
Team member 3	**Gabriel Omondi Odero**

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed. **Note:** You may be required to provide proof of your outreach to non-contributing members upon request.

Introduction

The current project involved modeling regime-switching time series to reflect the new normals in financial markets, focusing on the pre-COVID year in 2019 and 2 years after the pandemic (2020). In that respect, we can classify three phases: pre-covid, pandemic, and post-pandemic. Financial time series from different asset classes (equities, cryptocurrencies, rates, credit, among others) spanning 2019 to 2022 were collected. The financial time series were visualized to identify regime change, and then a decision was taken to model a particular time series based on the outcome of the visualizations. The Markov-regime switching model was done under different assumptions, including different numbers of states, different expectations with constant variance, different variance with constant expectation, and different expectations and variances. Model performance was assessed using standard information criteria, which were also used to rank the models. In addition, models were estimated assuming that the time series selected for this project was an autoregressive process with state-dependent autoregressive coefficients and variance.

Step 1: Data Collection

Asset Classes Chossen

For this project, we each picked some financial time series from stocks, bonds, and crypto, the - types of assets traded to make a profit or stack value.

We tracked prices daily starting 2019 into 2022, capturing major financial and economic events like the COVID-19 pandemic.

Asset Classes Chosen

Stocks - Big tech companies investors bet on

- GOOGL (Alphabet Inc.)
- NVDA (NVIDIA Corporation)

Bonds - Loans in a bundle.

- IEF (iShares 7-10 Year Treasury Bond ETF)
- TLT (iShares 20+ Year Treasury Bond ETF)

Crypto - Digital currencies that can bubble or bust quickly.

- BTC-USD (Bitcoin)
- ETH-USD (Ethereum)

Blending these assets gives a complete view. We can explain how they move during panics and chill times to understand market behavior. Tracing the flows reveals human nature - fear, greed, hope, and doubt all tied up in finance.

Table 1: Descriptive statistics were calculated for each asset: mean, median, variance, skewness, and kurtosis, as well as the results of normality tests.

	Mean	Median	Variance	Standard Deviation	Skewness	Kurtosis	Shapiro-Wilk p-value
BTC-USD	**0.002106**	**0.001164**	**0.001657**	**0.040701**	**-0.80524**	**11.17057**	**2.57E-20**
ETH-USD	**0.002037**	**0.000882**	**0.00257**	**0.050696**	**-0.94192**	**8.336324**	**1.63E-17**
GOOGL	**0.001109**	**0.00109**	**0.00038**	**0.019505**	**0.277129**	**3.611346**	**5.46E-16**
IEF	**8.60E-05**	**0.000262**	**1.90E-05**	**0.004339**	**-0.07272**	**4.687699**	**8.47E-15**
NVDA	**0.002038**	**0.002123**	**0.001011**	**0.031795**	**0.31454**	**1.866091**	**3.57E-08**
TLT	**0.000255**	**0.000784**	**0.000113**	**0.010625**	**-0.09427**	**7.383436**	**1.74E-17**

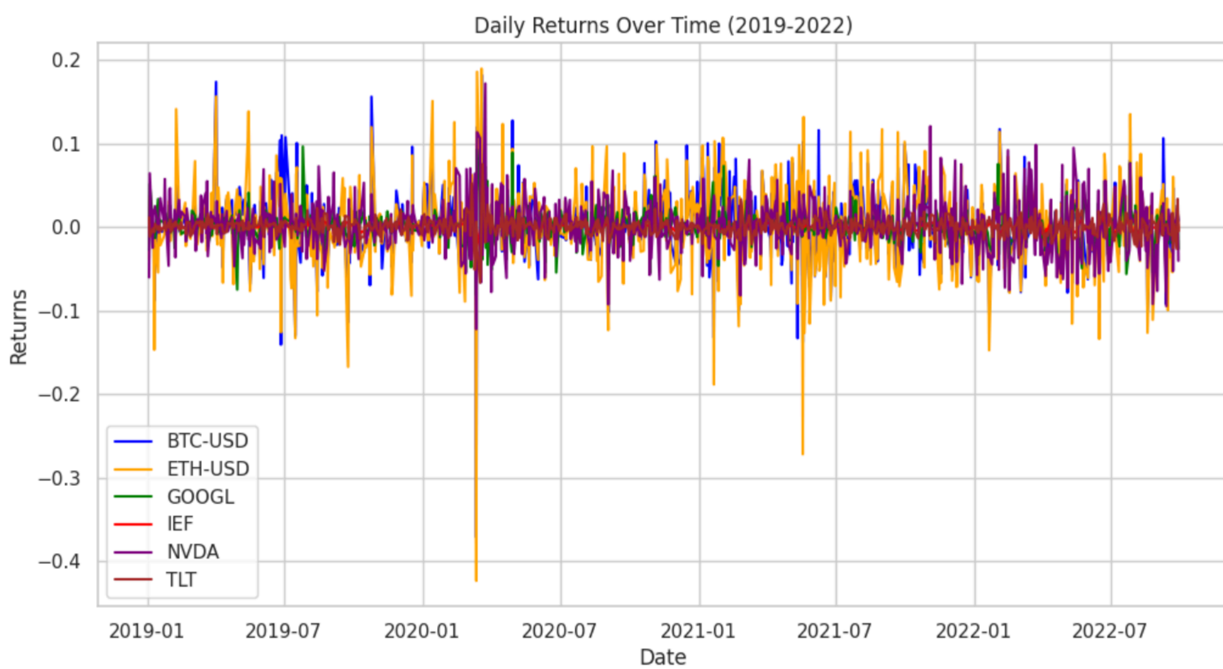


Fig. 1: Returns of selected assets

The analysis from Figure 1 indicates higher volatility from the different asset classes. More specifically, cryptocurrencies had the highest volatility. This indicated that BTC-USD and ETH-USD had higher volatilities than the other assets. The trend indicated that each of the assets had differences in volatilities, with the treasury bonds (IEF and TLT) used in the analysis having the least volatility.

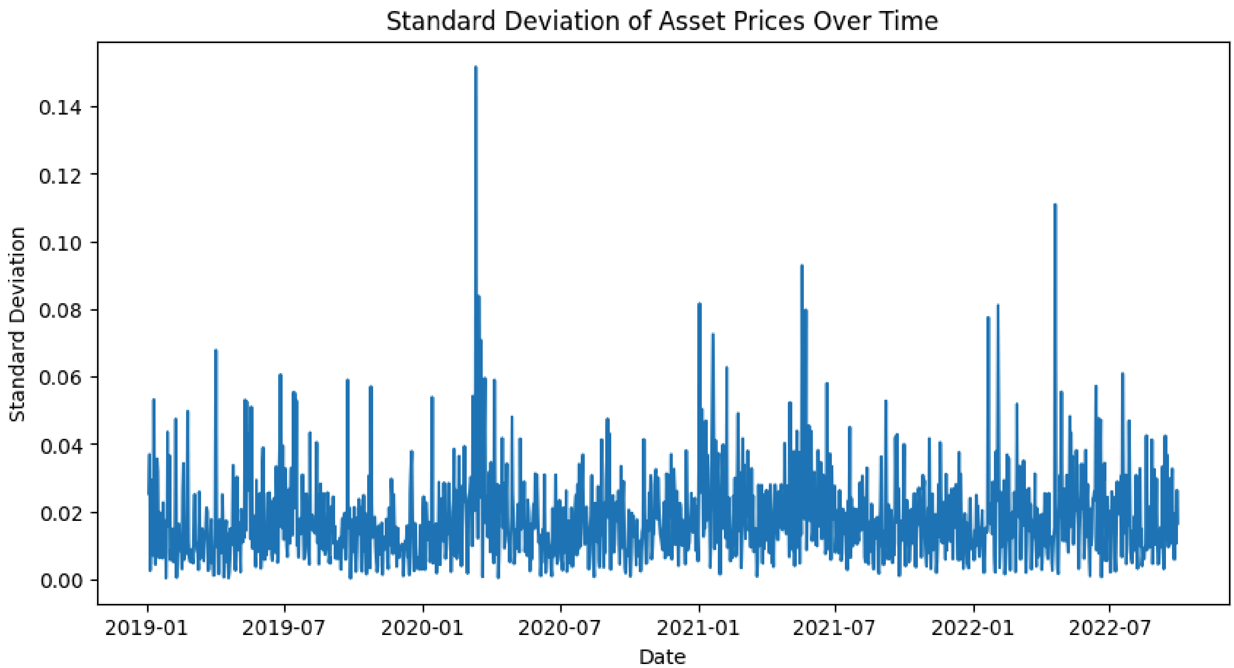


Fig 2: Standard deviations of asset returns over time

The standard deviations associated with the joint assets' returns indicated that there were instances of high and low volatilities. The period from 2020-01-2020-07 had the highest volatilities, followed by the period from 2022-01 to 2022-07. The least volatility was experienced in 2019-01 to 2019-07. The trend reports instances of regime changes in terms of the reported returns. The trend informed the need for a **univariate time series** to explain regime changes across the period from the various assets.

	GOOGL	META	NFLX	NVDA	IEF	TLT	BND	SHY	BTC-USD	ETH-USD
GOOGL	1.000000	0.637505	0.479718	0.599668	-0.066798	-0.102834	0.120263	-0.008966	0.138473	0.157481
META	0.637505	1.000000	0.486877	0.526563	-0.049913	-0.072592	0.099928	-0.009208	0.119046	0.126946
NFLX	0.479718	0.486877	1.000000	0.485524	-0.003875	-0.034543	0.111712	0.033538	0.117281	0.118609
NVDA	0.599668	0.526563	0.485524	1.000000	-0.063126	-0.070321	0.106927	-0.038125	0.158726	0.171722
IEF	-0.066798	-0.049913	-0.003875	-0.063126	1.000000	0.913258	0.837550	0.812771	-0.004950	-0.013037
TLT	-0.102834	-0.072592	-0.034543	-0.070321	0.913258	1.000000	0.779779	0.602636	-0.016240	-0.026235
BND	0.120263	0.099928	0.111712	0.106927	0.837550	0.779779	1.000000	0.707461	0.083854	0.077027
SHY	-0.008966	-0.009208	0.033538	-0.038125	0.812771	0.602636	0.707461	1.000000	0.026732	0.016726
BTC-USD	0.138473	0.119046	0.117281	0.158726	-0.004950	-0.016240	0.083854	0.026732	1.000000	0.811146
ETH-USD	0.157481	0.126946	0.118609	0.171722	-0.013037	-0.026235	0.077027	0.016726	0.811146	1.000000

Fig 3: Correlation analysis of the asset returns

The findings reported strong positive correlations between the cryptocurrencies (BTC-USD and ETH-USD) and a low relationship with the other assets. The trend is in line with symmetric trends of cryptos, owing to the evolution in returns based on sentiments from the greater market. This translates to a higher correlation between the cryptos as opposed to the other assets. The findings also reported that META, Google, and NVDA also had a moderate correlation with each other. This is also based on the assets falling in almost the same industry, also influenced by almost similar forces.

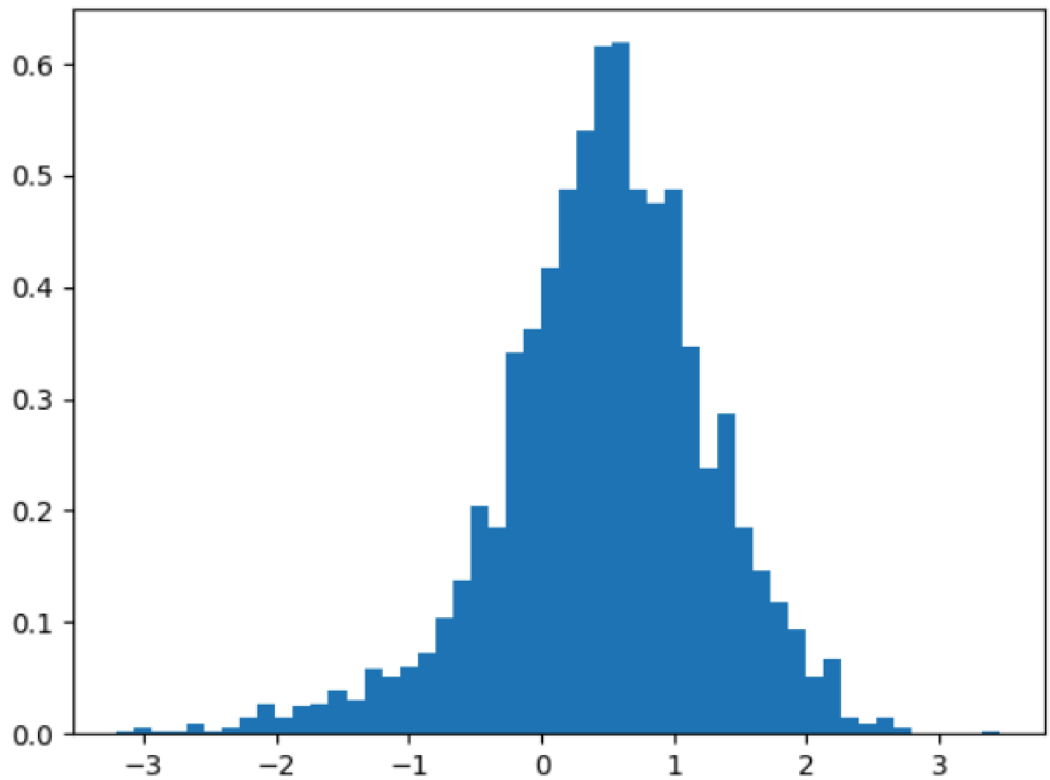


Fig 3: Histogram of the returns

The histogram indicates that the series was bi or tri-modal based on the heaps witnessed. This informs that the log of the returns informs the essence of regime-switching models, to address the different regimes witnessed by the selected companies.

Volatilities:

Cryptocurrencies tend to bounce around a lot in value - they're super speculative, so their prices go up and down a lot. For example, in 2020, when COVID hit, crypto went haywire up and then way down, and tech stocks also swung pretty wildly, just less than crypto.

Another thing is volatility comes in waves - like clusters. When COVID first slammed the markets in the spring of 2020, volatility spiked across assets. Then it calmed down a bit before blasting off again when the recovery started kicking in 2021. The persistence of volatility was clear when we looked at autocorrelation functions. Volatility has a memory and hangs around longer than you might think.

Default Probabilities (Proxy)

Since direct default probabilities are not available for the selected assets, so we use jump numbers. Our observation is shown in Figure 1. The average default probability was higher in crypto. Bonds were steadier, representing low risk.

Step 2: Regime Change and Models

2.1 Visualization and Regime Identification

Time Series Plotting

We plot the time series for the assets under analysis from January 2019 through September 2022 and just look at the basic price trends and movement. This is to see see how things have changed, maybe some big crashes or spikes.

The graph shown in Figures 6a and 6b shows potential regime shifts. Figure 6b was plotted using log scale to improve the readability of Figure 6a. We'll see some clear actions around when COVID-19 hit and how things tanked but recovered. Tech and crypto actually did pretty well.

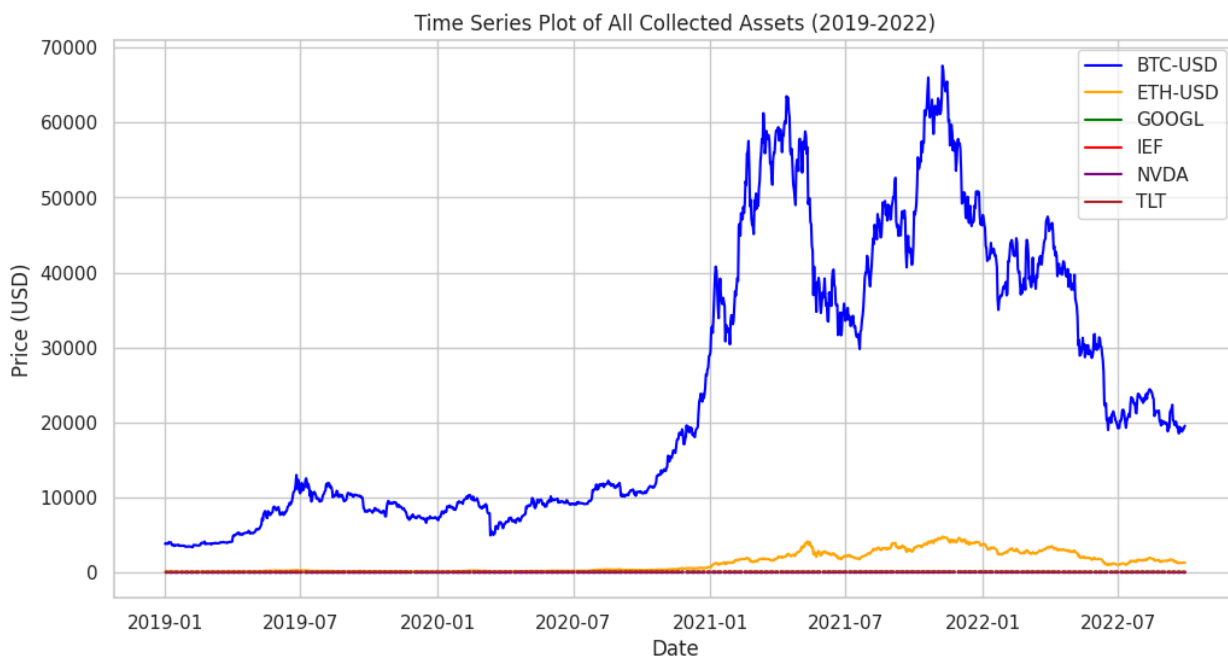


Figure 4a: Time Series plot for all assets from 2019 - 2022

Regime Changes and Approximate Dates

Figure 4b highlights key regime shifts that were caused by the following significant dates:

- **March 2020**: The COVID-19 market crash.
- **June 2020**: The CVID-19 recovery begins.
- **January 2022**: The period of inflation-driven volatility and interest rate hikes.

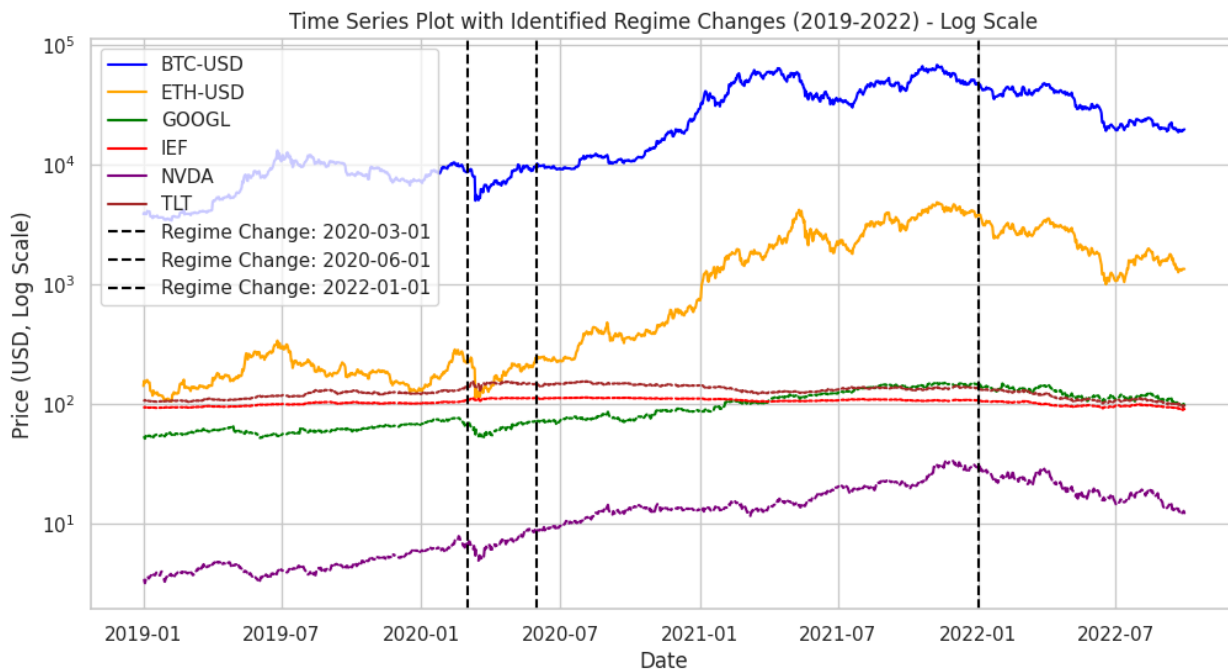


Figure 4b: A logarithm representation of the price in Figure 4a

Selection of Time Series for Modeling

Criteria for Selection

- **Volatility and Regime Shifts**: we prefer highly volatile assets and distinct regime shifts.
- **Market Relevance**: We want big players that are influential, trendy, and widely traded
- **Data Characteristics**: Which asset shows clear and diverse regimes (e.g., speculative bubbles, corrections, crashes).
- **Potential Insights:** Does the asset have room for meaningful insights, especially in dynamic and evolving sectors?

Justification for choosing BTC-USD

Volatility and Regime Shift:

Bitcoin makes a lot of sense to study using regime-switching models because it tends to change the way it behaves at different times. For example, when the pandemic hit, everything went crazy. Prices spiked like crazy in 2021 when speculation went wild, then crashed hard late that year into 2022. So, we can see that Bitcoin has clear boom-and-bust cycles.

Market Relevance:

BTC-JUS is the leader of crypto - the one everyone watches to see where things might go next, and when Bitcoin sneezes, the rest of the crypto market catches a cold. People are glued to what Bitcoin prices are doing both on Wall Street and Main Street.

Diverse Regime Characteristics:
With Bitcoin, you've also got obvious times when investor sentiment switched, like between speculative bubbles, crashes, corrections, etc. This is a great chance to model how speculative assets can have such different moods and phases. It might help figure out how and why they swing so wildly compared to stocks and bonds that don't have such crazy ups and downs.

Potential for Insight:
We figure taking a closer look can give us some clues about what makes speculative markets tick during the highs and lows, especially with something as headline-grabbing as Bitcoin.

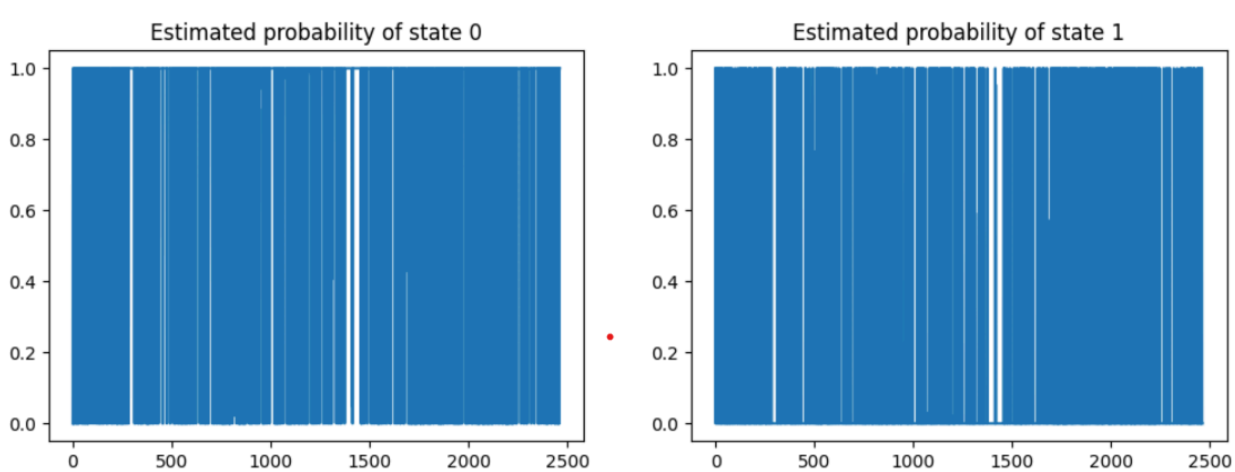
2.2 Markov-Regime Switching Model Estimation

After selecting BTC-USD (Bitcoin) to use for the regime-change model, the next move is to estimate a Markov-Regime Switching Model with a couple of different numbers of states. Basically, we want to see how well models with two states or three states can capture Bitcoin's shifts between different regimes and analyze both models to compare which one better shows when Bitcoin switches from a bull market to a bear market. More states might capture more nuance but could also overcomplicate things. So we'll have to check the fits. The goal is to find the simplest model that reflects the significant shifts.

Models with Different Numbers of States

We will compare a two-state model and a three-state model to see which provides a better fit for Bitcoin's historic price behavior.

- Two-state model (typically capturing "high-volatility" and "low-volatility" regimes)
- Three-state model (capturing "bullish," "bearish," and "neutral or correction" regimes)



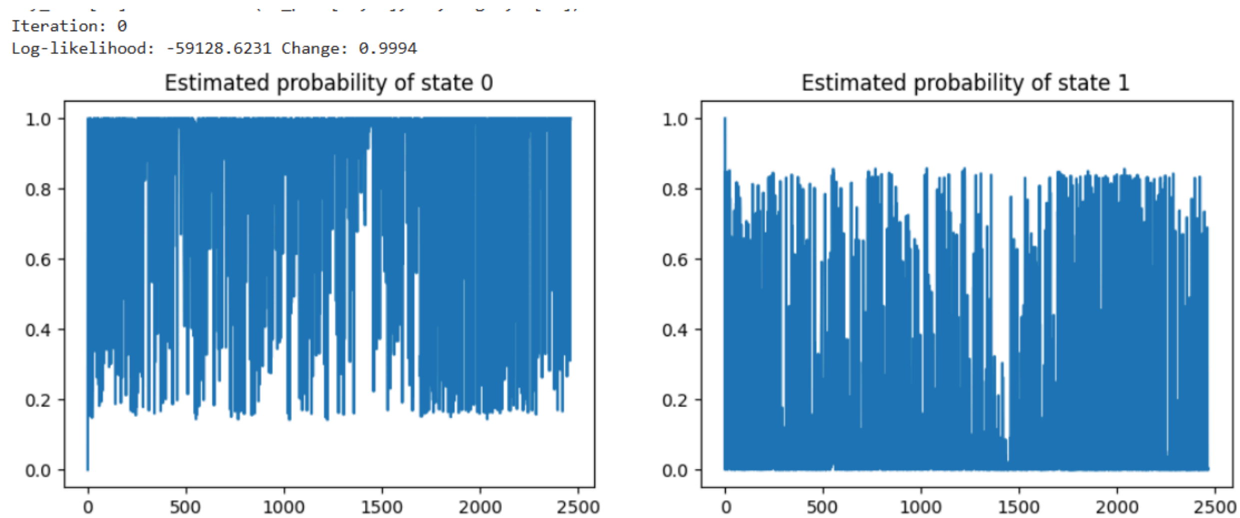


Fig 5-14: Markov switching graphs for Regime switching Model Identification

The regime change graphs revealed that change was taking place between state 0 and state 1. This indicates that as we progress, the model is more able to capture

2b. Model estimation

The regime change reported 14 iterations. The resultant log-likelihood for the model indicated that the change was 0.0 with a log-likelihood of -3494.38. The final estimates informed that the log likelihood was -3494.3807, Akaike being 7008.7615, and Schwarz being 7066.8528.

```
                         Markov Switching Model Results
================================================================================
Dep. Variable:                      y   No. Observations:                 2463
Model:               MarkovRegression   Log Likelihood               -2683.299
Date:                Wed, 23 Oct 2024   AIC                           5378.599
Time:                        00:22:25   BIC                           5413.454
Sample:                             0   HQIC                          5391.262
                               - 2463
Covariance Type:               approx
                             Regime 0 parameters
================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------
const          0.5785      0.014     40.780      0.000       0.551       0.606
sigma2         0.3404      0.012     29.078      0.000       0.317       0.363
                             Regime 1 parameters
================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------
const          0.2186      0.041      5.371      0.000       0.139       0.298
sigma2         1.2736      0.065     19.592      0.000       1.146       1.401
                         Regime transition parameters
================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------
p[0->0]        0.9996      0.000   2401.728      0.000       0.999       1.000
p[1->0]        0.0005      0.001      0.755      0.450      -0.001       0.002
================================================================================

Warnings:
[1] Covariance matrix calculated using numerical (complex-step) differentiation.
```

Regime Switching Model constant parameters for period 0 was 0.4268, increasing to 1.1477 in state 1. From the visualization, Mu: [0.4658 1.0977] and Sigma: [0.8124 0.0233]. These figures were close to what was obtained in the actual model, informing the relevance of the estimated Regime switching model to the actual model realized. Furthermore, all the model parameters were statistically significant at 5% level of significance as witnessed. The model is similar to the one obtained from the regime switching graphs, based on the iterations.

(i) Different states (N=4)

```
Iteration: 2
Log-likelihood: nan Change: nan
Final estimates:
Log-Likelihood: nan Akaike: nan Schwarz: nan
Mu: [0.2725 1.4487 3.4595 3.4595]
Sigma: [0.7241 0.4188 0.     0.    ]
Transition Matrix:
[8.379e-01 1.616e-01 5.000e-04 0.000e+00]
[0.8311 0.1689 0.     0.    ]
[0. 1. 0. 0.]
[0. 1. 0. 0.]
Initial probabilities: [0. 1. 0. 0.]
```

The regime change reported 2 iterations. The resultant log-likelihood for the model indicated that the change was nan, with a change of nan. The Akaike value was nan and Schwarz being nan. State 2 and 3 had the highest average and the least sigma than the other states.

(ii) Different expected realization and constant sigma

```
Iteration: 2
Log-likelihood: nan Change: nan
Final estimates:
Log-Likelihood: nan Akaike: nan Schwarz: nan
Mu: [0.2725 1.4487 3.4595 3.4595]
Sigma: [0.7241 0.4188 0.     0.    ]
Transition Matrix:
[8.379e-01 1.616e-01 5.000e-04 0.000e+00]
[0.8311 0.1689 0.     0.    ]
[0. 1. 0. 0.]
[0. 1. 0. 0.]
Initial probabilities: [0. 1. 0. 0.]
```

The regime change reported 2 iterations. The resultant log-likelihood for the model indicated that the change was nan. The Akaike value was nan and Schwarz being nan. States 2 and 3 had the highest average and the least sigma than the other states.

(iii) Different variance and constant mu

```
Iteration: 5
Log-likelihood: -3494.8748 Change: 0.0
Final estimates:
Log-Likelihood: -3494.8748 Akaike: 7045.7496 Schwarz: 7208.4054
Mu: [0.466  1.9587 1.9582 1.9571]
Sigma: [0.8123 0.1315 0.1318 0.1319]
Transition Matrix:
[1. 0. 0. 0.]
[1. 0. 0. 0.]
[1. 0. 0. 0.]
[1. 0. 0. 0.]
Initial probabilities: [1. 0. 0. 0.]
```

The regime change reported 5 iterations. The resultant log-likelihood for the model indicated that the change was -3494.8748, with a change of 0.0. The Akaike value was 7045.7496 and Schwarz being 7208.4054. State 1 had the highest average and the least sigma than the other states.

(iv) Different variance and mu

```
Iteration: 11
Log-likelihood: -3498.1515 Change: 0.0
Final estimates:
Log-Likelihood: -3498.1515 Akaike: 7052.303 Schwarz: 7214.9588
Mu: [0.4648 1.0985 1.411  3.0194]
Sigma: [0.8111 0.0011 0.0801 0.4398]
Transition Matrix:
[9.998e-01 0.000e+00 0.000e+00 2.000e-04]
[1. 0. 0. 0.]
[9.991e-01 0.000e+00 9.000e-04 0.000e+00]
[0.2452 0.     0.2784 0.4764]
Initial probabilities: [0. 1. 0. 0.]
```

The regime change model reported 11 iterations. The resultant log-likelihood for the model indicated that the change was -3498.1515, with a change of 0.0. The Akaike value was 7052.303 and Schwarz was 7214.9588. State 3 had the highest average and the least sigma than the other states.

Table 2: Regime models assumption outcomes' comparison

Assumption	AIC	Schwarz	Log
Different states (N=4)	NAN	NAN	NAN
Different mu and constant sigma	NAN	NAN	NAN
Different sigma and constant mu	7045.7496	7208.4054	-3494.8748
Different mu and sigma	7052.303	7214.9588	-3498.1515

The analysis indicated that regime switching model with different sigma and constant mu had a lower AIC and Schwarz values, with a higher log likelihood value. This indicates that the model performed better than the others.

Step 3: Model Performance

(a) Model for different means

```
Iteration: 14
Log-likelihood: -3494.3807 Change: 0.0
Final estimates:
Log-Likelihood: -3494.3807 Akaike: 7008.7615 Schwarz: 7066.8528
Mu: [0.4658 1.0977]
Sigma: [0.8124 0.0233]
Transition Matrix:
[1. 0.]
[1. 0.]
Initial probabilities: [0. 1.]
```

The regime-switching model reported 14 iterations. The resultant log-likelihood for the model indicated that the change was -3494.3807, with a change of 0.0. The Akaike value was 7008.7615 and Schwarz being 7066.8528. State 1 had a higher average and a lower sigma than state 0.

(b) Model for different sigma values

```
Iteration: 4
Log-likelihood: -3494.8719 Change: 0.0
Final estimates:
Log-Likelihood: -3494.8719 Akaike: 7009.7438 Schwarz: 7067.8352
Mu: [1.9596 0.466 ]
Sigma: [0.1194 0.8123]
Transition Matrix:
[0. 1.]
[0. 1.]
Initial probabilities: [0. 1.]
```

The regime-switching model reported 4 iterations. The resultant log-likelihood for the model indicated that the change was -3494.8719, with a change of 0.0. The Akaike value was 7009.7438 and Schwarz was 7067.8352. State 0 had a higher average and a lower sigma than state 0.

(c) Model for different means and sigma values

```
Iteration: 18
Log-likelihood: -3494.7755 Change: 0.0001
Final estimates:
Log-Likelihood: -3494.7755 Akaike: 7009.551 Schwarz: 7067.6424
Mu: [0.4657 1.0022]
Sigma: [0.8124 0.2242]
Transition Matrix:
[9.997e-01 3.000e-04]
[9.999e-01 1.000e-04]
Initial probabilities: [0. 1.]
```

The regime-switching model reported 18 iterations. The resultant log-likelihood for the model indicated that the change was -3494.7755, with a change of 0.0001. The Akaike value was 7009.6424 and Schwarz was 7067.6424. State 1 had a higher average and a lower sigma than state 0.

(d) Comparison of the models

Table 3: Model comparison

Model	AIC	Schwarz	Log-Likelihood
Different mu	7008.7615	7066.8528	-3494.3807
Different sigma	7009.7438	7067.8352	-3494.8719
Different mu and sigma	7009.551	7067.642	-3494.7755

The analysis indicated that model with different m had the least AIC, Schwarz and highest log-likelihood value. This informs that it is the most preferred model. This is followed by model with different mu and sigma as the last model being that with different sigma in that order.

Step 4: AR Models with AR components changing with states

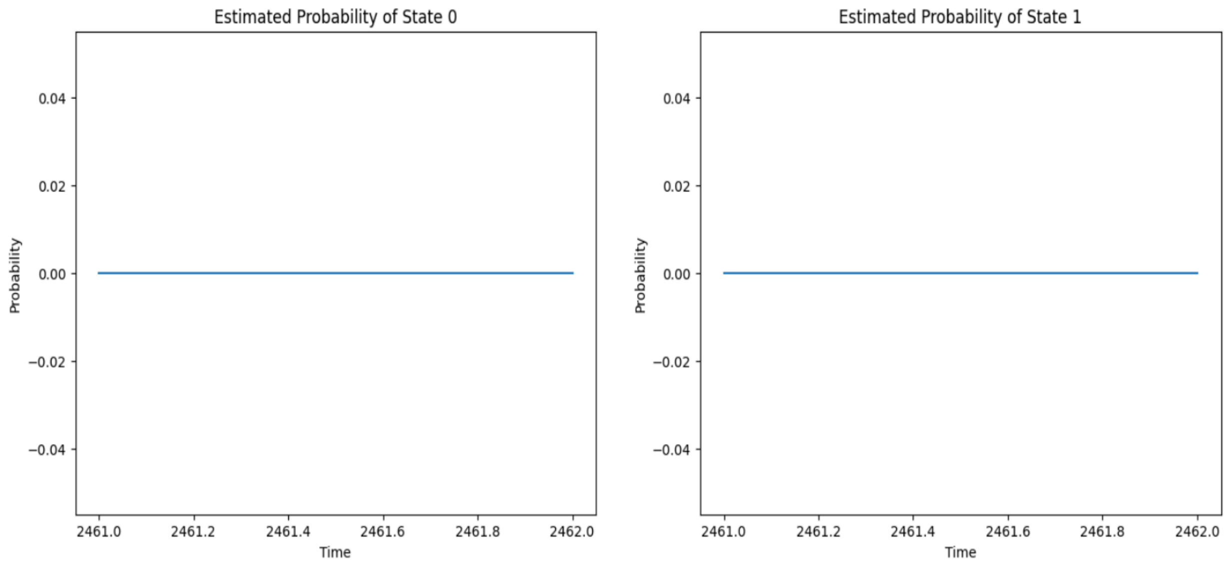
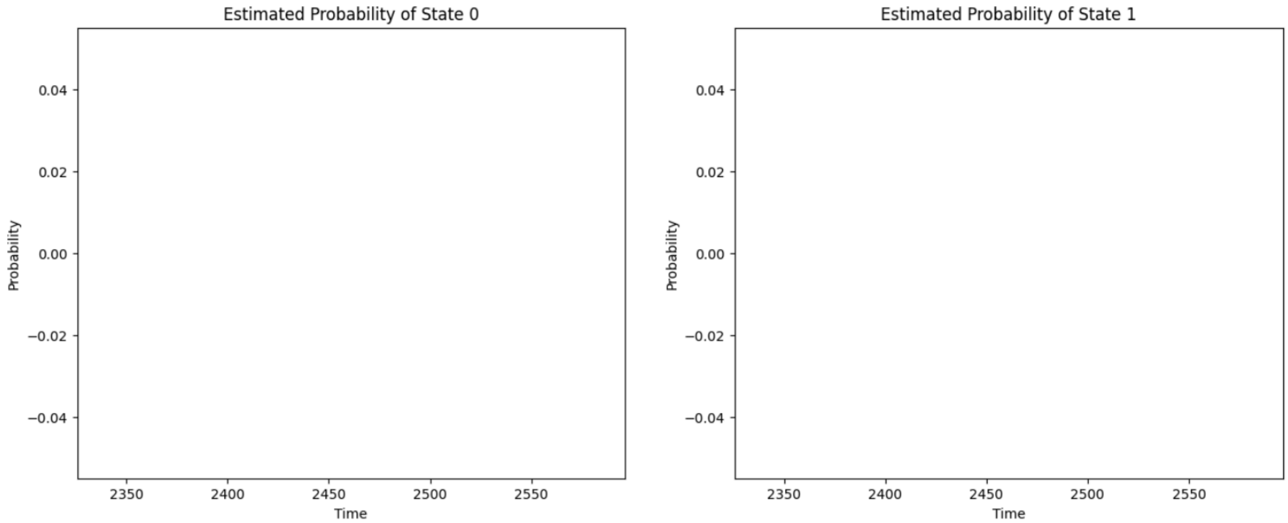


Figure 6: Estimated probabilities from the different states from the baseline model

```
Iteration: 2
Log-likelihood: 0.6625 Change: -0.3336
Final estimates
Log-Likelihood: 0.6625 Akaike: 18.6751 Schwarz: 76.7665
Mu: [0.6849 0.6118]
Sigma: [0.2893 0.0529]
Rho: [0.1 0.1]
Transition Matrix:
[0.8972 0.7665]
[0.5134 0.5833]
Initial Probabilities: [0.5 0.5]
```

The analysis informed that there was clarity on how the states were separated. However, there was some similarity witnessed on the two states (0 and 1). The log-likelihood for the model was 0.6625 with a negative change of -0.336 from the 2 iterations. The final estimates reported an AIC of 18.675 and a Schwarz of 76.7665. The estimated means for the model were 0.6849 for state 0 and 0.6118 for state 1. This informs that the mean for state 1 is lower than for state 0. On the sigma, it was reported that state 0 had a higher standard deviation of 0.2893 as compared to state 1 at 0.0529. On the rho, it was noted that there was not change. The transition probabilities report changes across the different states, from initial probabilities that were similar.

Figure 7: AR model with AR components changing with states



WLS model for comparison

```
Iteration: 2
Log-likelihood: 0.3749 Change: -0.5236
Final estimates
Log-Likelihood: 0.3749 Akaike: 19.2501 Schwarz: 77.3415
Mu: [0.0682 0.9798]
Sigma: [0.4876 0.4905]
Rho: [0.1 0.1]
Transition Matrix:
[0.4953 0.7856]
[0.8913 0.3673]
Initial Probabilities: [0.5 0.5]
```

Weighted Least Squares model

The results showed not much difference between state 0 and state 1's probabilities. The analysis reported a log-likelihood value of 0.3749 with a negative change of -0.5236. The final estimates revealed that the log-likelihood value was 0.3749, Akaike was 19.2501, and Schwarz was 77.3415. The analysis showed that the mean for state 1 (0.9798) was higher than that for state 0 (0.06820. The standard deviations were also in the same fashion at 0.4905 for state 1 and 0.4876 for state 0.

In comparing the two models, the baseline model had lower Akaike and Schwarz values and a higher log-likelihood value. This showed that the baseline model (AR model) where the time series component emanates from the autoregressive process, was considered a better fit for the data.

References

1. Sutton, Richard S., and Andrew G. Barto. Reinforcement Learning: An Introduction. MIT Press, 2018. http://incompleteideas.net/sutton/book/the-book.html
2. Bertsekas, Dimitri P. Dynamic Programming and Optimal Control. 3rd ed., Athena Scientific, 2005.
3. Hamilton, James D. "Regime switching models." Macroeconometrics and time series analysis. London: Palgrave Macmillan UK, 2010. 202-209.
4. Chang, Yoosoon, Yongok Choi, and Joon Y. Park. "A new approach to model regime switching." Journal of Econometrics 196.1 (2017): 127-143.
5. Bazzi, Marco, et al. "Time-varying transition probabilities for Markov regime switching models." Journal of Time Series Analysis 38.3 (2017): 458-478.
6. Kim, Chang-Jin, Jeremy Piger, and Richard Startz. "Estimation of Markov regime-switching regression models with endogenous switching." Journal of Econometrics 143.2 (2008): 263-273.
7. Savku, E., and G-W. Weber. "A regime-switching model with applications to finance: markovian and non-markovian cases." Dynamic Economic Problems with Regime Switches (2021): 287-309.
8. Shen, Yang, Jiaqin Wei, and Qian Zhao. "Mean–variance asset–liability management problem under non-Markovian regime-switching models." Applied Mathematics & Optimization 81 (2020): 859-897.
   Statsmodels Documentation. *Markov Switching Models* in Statsmodels. https://www.statsmodels.org/stable/tsa.html#markov-switching-models. Accessed 22 October 20
9. Yahoo Finance. “BTC-USD Bitcoin to US Dollar.” *Yahoo Finance*, https://finance.yahoo.com/quote/BTC-USD/history/. Accessed 22 October 2024.
10. **Anku, Eric K., Odero, Gabriel O., Kamara, Alpha K., .** *MScFE 622 Stochastic Modeling: Group Work Project #2*. Google Colab, 22 October 2024, https://colab.research.google.com/drive/14N4UK0q52ShZKrxnlwaRO6WCJfqLyWVk?usp=sharing