MGT6203 - Team 46 - Final Project Report

Ruslan Ozerin, Christina Waldek, Min Hsuan Lin, Daniel Morace, Sayeed Ahmed

Evaluating Investment Decisions - Bitcoin, S&P500, Gold

1. Background Information and Framing of the Problem

Cryptocurrencies are digital currencies based on blockchain technology that keeps a ledger of transactions. Because cryptocurrencies do not require banks or other institutions, they often are cheaper and faster than standard bank transactions. They also are not tied to one specific nation's economy (Voigt and Rosen 2022). However, because they are a modern technology, they have not yet been globally accepted as currency and are subject to high volatility. Bitcoin currently is the first and largest cryptocurrency with a market cap of over $360 billion (Duggan 2022).

As economies cycle and vary from country to country, cryptocurrencies offer a virtual currency that is "theoretically immune to government interference or manipulation" (Duggan 2022). As cryptocurrencies are a recent technology, we will investigate the risks and benefits of Bitcoin, the largest cryptocurrency.

Bitcoin is a digital currency, which operates free of any central control or the oversight of banks or governments. It was created in 2009 by an anonymous figure named Satoshi Nakamoto who wanted to create an alternative to the banking system, where banks can create credit bubbles while keeping very little as a reserve thus breaching the trust of its customers (Saleem 2018). Bitcoin offers a way to invest into a deflationary asset with an ability to buy/sell at any moment as well as send/receive it to/from anyone in the world with a Bitcoin address, which is a unique identifier that serves as a virtual location where the cryptocurrency can be sent (although some countries banned cryptocurrency). This is a promising technology and coupled with high returns over the years, it makes Bitcoin a viable investment option for companies and individuals.

However, cryptocurrency is a relatively new and risky asset class, and we will evaluate its performance compared to the S&P 500 index and its stocks. Bitcoin is referred to as "digital gold", and we will also find out how gold compares to it as an investment option. As inflation rates continue to rise and the risk of recession increases, cryptocurrencies can help potentially hedge against these risks. Investing in Bitcoin could be a way to help protect a business's financial assets. Additionally, there is a potential benefit of cryptocurrencies for international companies since cryptocurrencies are decentralized and not tied to a specific country.

2. Primary Research Question

Does Bitcoin yield a higher short-term (one year or less) or long-term (more than one year) return and have a lower risk than S&P 500, its stocks, and commodities like gold?

3. Initial hypotheses

1) We anticipate that Bitcoin has high returns but high risk as well. S&P 500 stocks are a safer investment.

   - Cryptocurrency is a relatively new asset class with a large total market cap but no intrinsic value, while stocks are a share of ownership and have been around for centuries. It is hard to know what Bitcoin's real value is. On the other hand, Bitcoin has experienced astronomical returns and is seen as a promising groundbreaking technology. Bitcoin has seen large swings in price, both up and down. We anticipate that our statistical analysis will confirm it.

2) Bitcoin's price is correlated to S&P 500 more each year thus becoming less risky. Volatility is reduced each year.
   - As the market cap increases, volatility reduces. Bitcoin is being adopted by large companies and is seen as an alternative asset class. We expect that Bitcoin may behave like some S&P 500 stocks and react to the overall market price action.

3) We anticipate finding a company whose price action resembles Bitcoin the most (most likely a tech company).
   - Tech stocks and Bitcoin are seen as technology investments that have a lot of promise. Investors try to predict the market and invest in promising projects. We expect that Bitcoin behaves like a growth stock resembling some tech stocks.

4) We may conclude that Bitcoin is a suitable investment option nowadays, compared to more established stocks.
   - As Bitcoin becomes less risky and more adopted, it is a viable investment option with exposure to cryptocurrency, one of the world's most innovative technologies.

5) As Bitcoin has not experienced a true recession, we may conclude that it will be greatly affected during one when looking at stocks that compare to Bitcoin the most.
   - Cryptocurrency is a relatively young asset class and we do not have as much historical data as with the S&P 500 stocks, so we do not know how Bitcoin historically reacted to a recession. The last major recession was in 2007-2009, before Bitcoin's creation.

6) Bitcoin price action can be separated into several cycles.
   - This is caused by Bitcoin halving events that occur about every 4 years. We can see several distinctive cycles on the price chart and will analyze how each one resembles others.

4. Novelty of Our Project

Many of the studies on Bitcoin and cryptocurrency revolve around correlations with other assets. Our project aims to take this a step further and try to answer whether Bitcoin's differences from other assets yields higher returns, making it a worthwhile investment despite the risks. We will use the approach in the next section.

5. Data Cleaning & Processing

Several datasets have been considered for this initiative. The datasets were specifically selected for the purpose of comparative analysis between Bitcoin price, S&P 500 companies stock price, Gold and other commodity prices. Each dataset demanded several preparatory steps for data cleaning and preprocessing. S&P index and companies stock prices characterizes "dailies" weekdays datasets with some missing data, which was cleaned, transformed and imputed to hydrate the time series. Bitcoin datasets were more granular and included data collected at 1 minute and 30 minutes intervals. Each of the 2 datasets were transformed to daily data by aggregating followed by cleaning and imputation to fill in the missing data points and produce time series for relevant analyses. The functions *ddply* (*plyr* package) and *approxm* (*FreqProf* package) were found suitable for this scenario. The date range of 2012-Oct-22 to 2022-Sep-27 was selected for analysis.

Dataset	Characteristics		Initial Dimensions	Dimensions After Processing
BTC	Interval - Span -	30 minutes 2019-07 – 2022-09 (3 yrs)	55,382 x 13	1,156 x 10
BTC (large)	Interval - Span -	60 seconds 2011-12 – 2021-03 (10 yrs)	4,857,377 x 9	3,379 x 8
S&P 500 Index	Interval - Span -	Daily (weekdays) 2012-10 – 2022-10 (10 yrs)	2,517 x 2	3,652 x 2
500 companies stock prices	Interval - Span -	Daily (weekdays) 2010-01 – 2022-10 (12 yrs)	1,621,672 x 8	2,351,022 x 8
Commodities (Gold, Silver etc.)	Interval - Span -	Daily (weekdays) 2012-10 – 2022-10 (10 yrs)	5,888 x 24	8,328 x 24

7. Models:

a. Linear Regression Model

First, we are going to investigate the relationship between the response variable - Bitcoin and predictor variables - S&P 500 index, gold, and some S&P 500 stocks using linear regression. Let's pick 5 companies with the biggest market cap from different sectors to investigate if there may be a more linear relationship between Bitcoin and specific companies than others. We have Bitcoin as the response variable; predictor variables: S&P 500 index, Gold, and the following stocks: Apple (Technology sector), Google (Communication Services sector), Berkshire Hathaway (Financial Services sector), UnitedHealth (Healthcare sector), Exxon Mobil (Energy sector).

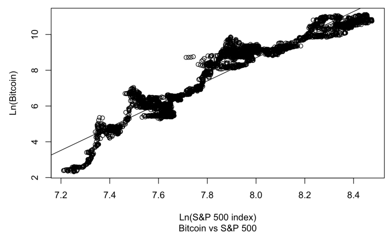
There are four assumptions to linear regression that we must follow: linearity, homoscedasticity, independence, normality. Looking at each variable, we see skewed histograms so we may not reasonably fit the data by a line. Applying natural log to each variable gives a more bell-shaped histogram for each variable except Exxon Mobil, but they all still appear to be bimodal and trimodal. There is no multicollinearity since we will build models with one predictor variable and should not build models including both S&P 500 index and its stocks, since they are included in the S&P 500 index. We could build a model with S&P 500 index and gold as predictor variables, but we will explore Bitcoin's relationship with each variable first.

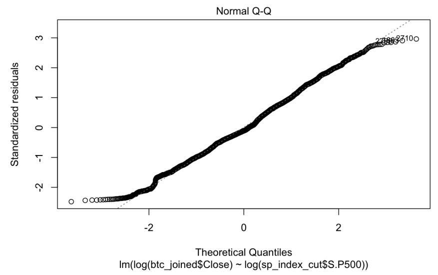
Initial analysis doing a simple linear regression on Bitcoin vs S&P 500 confirms issues, where the normal Q-Q plot has residuals way off the line on both ends, residuals vs fitted plot shows non-linear relationship between the response variable and predictors.

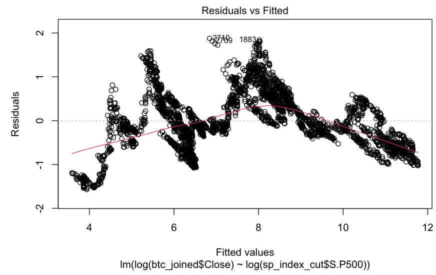
Variable transformation can be performed to have more linearity. It also makes sense with financial time series data as returns are compounded. We also tried Box-Cox variable transformation but at least for Bitcoin vs S&P 500 it is not a better model than the log-log model, and we also cannot easily interpret coefficients.

Ln (Bitcoin) ~ Ln (S&P 500)

Ln (Bitcoin) = -43.07 + 6.47 * Ln (S&P 500)







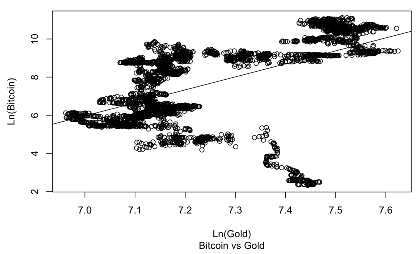
R-squared: 0.9114

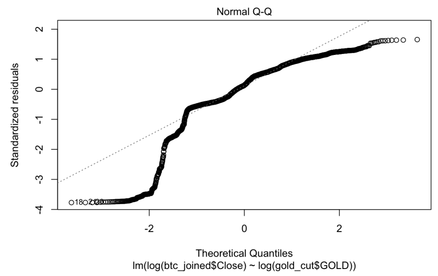
Normal Q-Q plot has residuals mostly following the line except for both ends, residual vs fitted plot shows some nonlinear pattern. We should not remove the outliers at the beginning and end because they encompass big time periods.

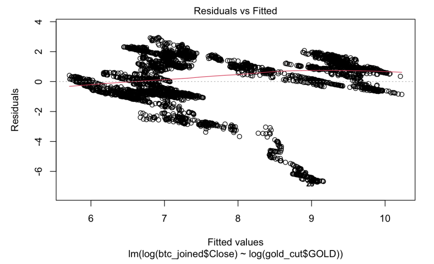
Despite the high R-squared value, there are some problems with this model, especially on both ends. We also observe several cycles, so a linear regression model may not be the best model.

Ln (Bitcoin) ~ Ln (Gold)

Ln (Bitcoin) = -41.77 + 6.82 * Ln (Gold)







R-squared: 0.3005

Normal Q-Q plot is off, residual vs fitted plot shows a nonlinear pattern.

This is not a good model and we cannot conclude there is a relationship between Bitcoin and gold.

For the five S&P 500 stocks:

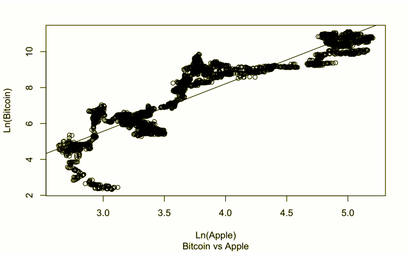
	Model	R-squared	Diagnostic Plots
Apple	Ln(BTC) ~ Ln(AAPL)	0.8139	Problems with earlier data
Google	Ln(BTC)~Ln(GOOGL)	0.9104	Plots look okay except for both ends like with the S&P 500 index model
Berkshire Hathaway	Ln(BTC) ~ Ln(BRK-B)	0.8998	Same as the Google model

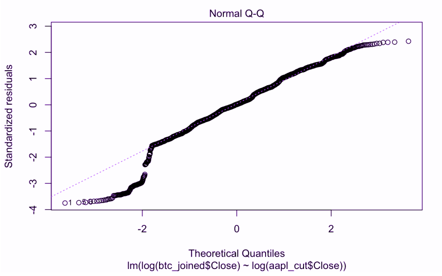
UnitedHealth	Ln(BTC) ~ Ln(UNH)	0.8932	Visible cycles and problems on both ends of data
Exxon Mobil	Ln(BTC) ~ XOM	0.3407	Lin reg assumptions are not followed

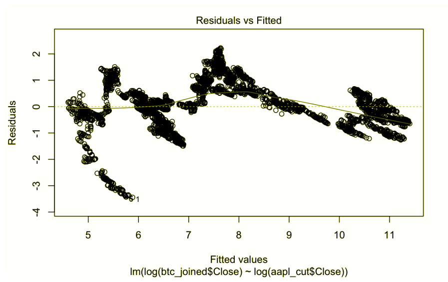
Diagnostic plots:

Ln (BTC) ~ Ln (AAPL)

Ln (BTC) = -2.38 + 2.65 * Ln (AAPL)

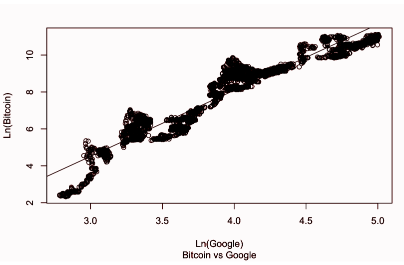


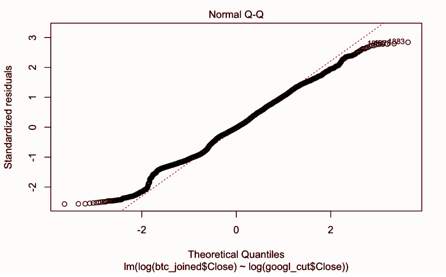


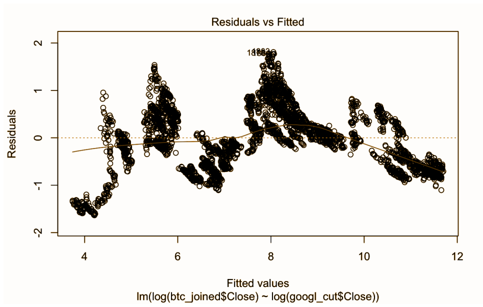


Ln (BTC) ~ Ln (GOOGL)

Ln (BTC) = -6.22 + 3.58 * Ln (GOOGL)

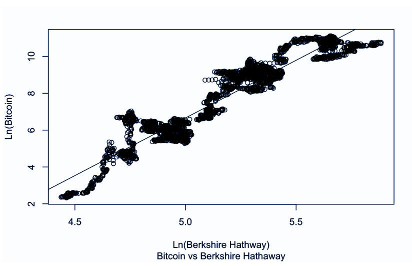


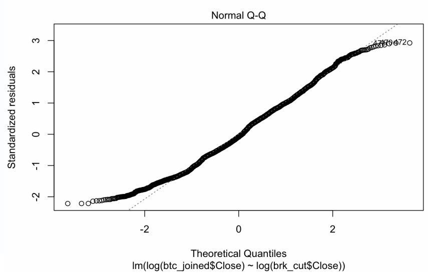


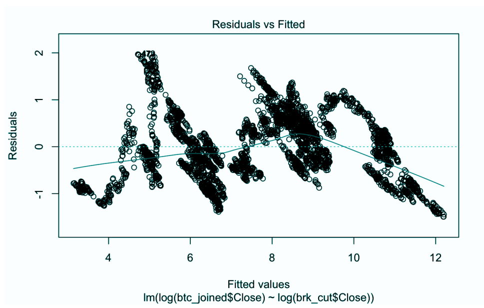


Ln (BTC) ~ Ln (BRK-B)

Ln (BTC) = -24.64 + 6.26 * Ln (BRK-B)

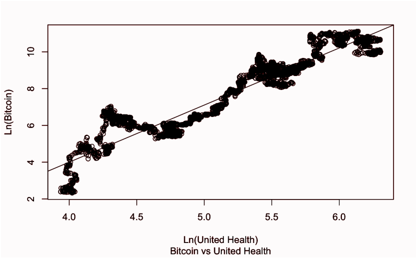


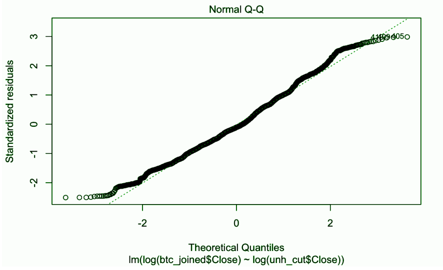


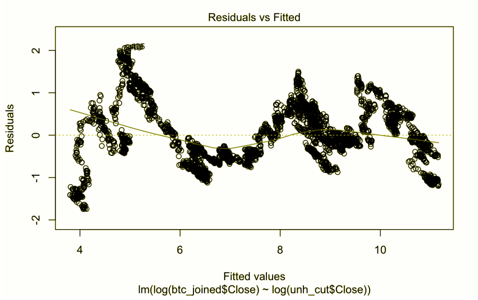


Ln (BTC) ~ Ln (UNH)

Ln (BTC) = -8.42 + 3.11 * Ln (UNH)

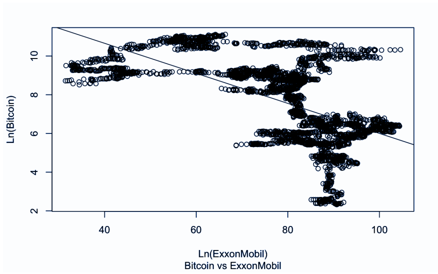


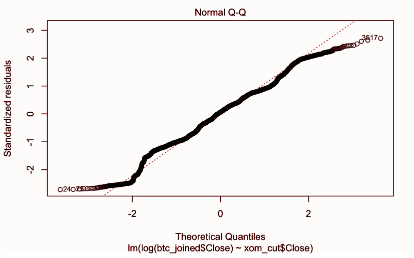


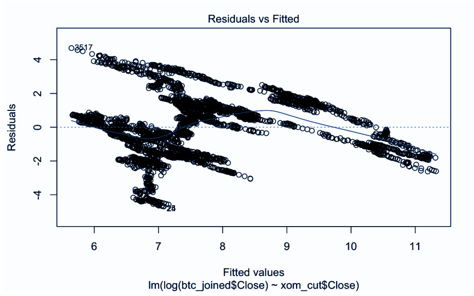


Ln (BTC) ~ XOM

Ln (BTC) = 13.77 + -0.078 * XOM







Conclusions:

Linear regression models may not be suitable to characterize Bitcoin's relationship with our predictor variables. There are obvious cycles in Bitcoin's price where it greatly differs from the fitted line, and linear regression models do not account for that unless we break data into several time periods. This is caused by Bitcoin's halving events, which are deflationary. Applying natural log helps with linearity but there are issues on both ends of data. One may argue that the log-log Bitcoin~S&P 500 index, Bitcoin~Google, Bitcoin~Berkshire Hathaway models could be useful in describing the variable relationship. Using these 3 models we can accept some initial hypotheses:

- Bitcoin yields a higher return than the S&P 500 index: when S&P 500 index increases by 1%, Bitcoin is expected to increase by **6.47%**.
- Some stocks' prices have a relationship with Bitcoin's price - Google and Berkshire Hathaway.
- Bitcoin's price has cycles. It is visible in the plots above where even after variable transformation Bitcoin smoothly fluctuates up and down.

We will investigate which other models will work better than linear regression especially since we deal with time series data that has seasonality.

b. GARCH Models

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are used to describe the variance of the errors in time series data based on the variances in previous periods. They are often used to describe the volatility of financial markets (Keton 2020). We used GARCH models to compare the volatility of the returns on the S&P 500 index, Bitcoin, and gold.

We created 8 different types of GARCH models and compared the AIC and BIC to determine which model was best.

1. Standard GARCH model with a Normal Distribution
2. Standard GARCH model with a Skewed Normal Distribution
3. Standard GARCH model with a Student T-Distribution
4. Standard GARCH model with a Skewed Student T-Distribution
5. GJR-GARCH model with a Skewed Student T-Distribution
    - The GJR-GARCH model adds a coefficient for the leverage effect.
6. eGARCH with a Student T-Distribution
    - The eGARCH model is an Exponential GARCH model that allows the effects of positive and negative returns to be asymmetrical.
7. fGARCH with a Student T-Distribution
    - The fGARCH model is a Family GARCH model that nests a variety of other GARCH models.
8. iGARCH with a Student-T Distribution
    - The iGARCH model is an Integrated GARCH model. The model is unit-root with persistence in volatility.

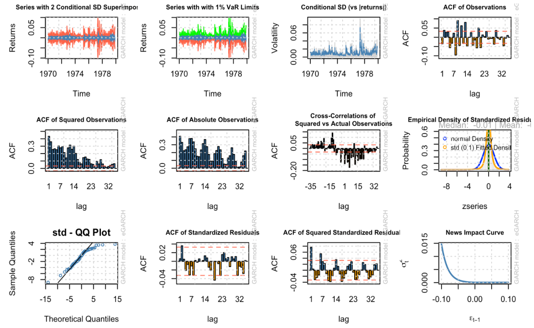
The AIC and BIC for the 8 models are:

		S&P 500	**Bitcoin**	**Gold**	**S&P 500**	**Bitcoin**	**Gold**
Model	**Distribution**	**AIC**	**AIC**	**AIC**	**BIC**	**BIC**	**BIC**
sGARCH	Normal	-7.264	-4.270	-7.431	-7.258	-4.263	-7.425
sGARCH	Skewed Normal	-7.273	-4.270	-7.435	-7.265	-4.261	-7.426
sGARCH	Student T	-7.495	-4.398	-7.651	-7.486	-4.388	-7.643
sGARCH	Skewed Student T	-7.495	-4.398	-7.652	-7.484	-4.386	-7.641
gjrGARCH	Skewed Student T	-7.530	-4.398	-7.651	-7.518	-4.385	-7.639
eGARCH	Student T	-7.548	-4.410	-7.655	-7.538	-4.398	-7.645
fGARCH	Student T	-7.495	-4.398	-7.651	-7.486	-4.389	-7.643
iGARCH	Student T	-7.496	-4.399	-7.652	-7.489	-4.391	-7.645

The model with the lowest AIC and BIC for all three datasets is the eGARCH model with a Student T-Distribution. This is the model we used to analyze the three datasets. The coefficients are mu, the mean; omega, the intercept; alpha1, the ARCH term; beta1, the GARCH term; gamma1, the leverage effect; and shape, the shape parameter of the student t-distribution.

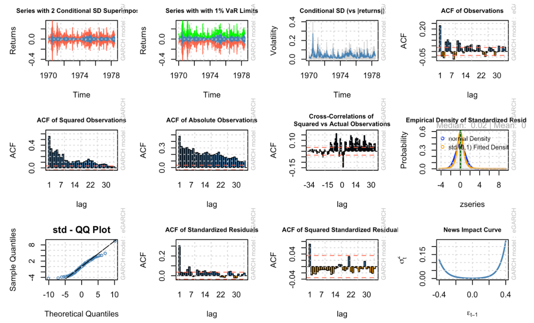
S&P 500 estimated coefficients and plots:

Parameter	Est.	Std. Error	t-value	Pr(>|t|)
mu	0.000365	0.000043	8.2286	0
omega	-0.220834	0.004874	-45.3084	0
alpha1	-0.250980	0.024900	-10.0795	0
beta1	0.976451	0.000329	2971.8243	0
gamma1	0.233994	0.035231	6.6418	0
shape	2.382051	0.093503	25.4756	0



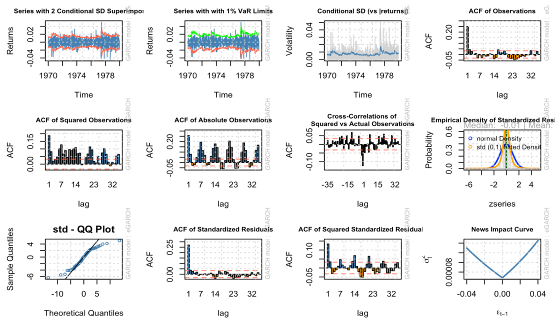
Bitcoin estimated coefficients and plots:

Parameter	Est.	Std. Error	t-value	Pr(>|t|)
mu	0.001599	0.00319	5.0115	0.000001
omega	-0.218218	0.062407	-3.4967	0.000471
alpha1	0.029360	0.017067	1.7203	0.085383
beta1	0.967710	0.008924	108.4334	0.000000
gamma1	0.44915	0.048302	9.2732	0.000000
shape	3.224048	0.238691	13.5072	0.000000



Gold estimated coefficients and plots:

Parameter	Est.	Std. Error	t-value	Pr(>|t|)
mu	0.000112	0.000067	1.66602	0.095711
omega	-0.056345	0.001925	-29.27297	0.000000
alpha1	0.007778	0.008399	0.92601	0.35440
beta1	0.994046	0.000184	5414.6389	0.000000
gamma1	0.084340	0.016502	5.11103	0.000000
shape	2.548048	0.144625	17.61829	0.000000



Based on these model coefficients, we calculated the long-run volatility using the formula: $\sqrt{\frac{omega}{1-beta1}}$ with the following results:

Long Run Volatility	Sqrt(exp(omega/(1-beta1))
S&P 500	0.92%
Bitcoin	3.41%
Gold	0.88%

Conclusions:

Based on these results, the volatility of Bitcoin is the highest, followed by the S&P 500 and lastly gold. These results were expected. Additionally, looking at the New Impact Curve, which shows the impact of previous positive or negative results, the impact of negative returns is high with positive returns flat for the S&P 500, while positive previous returns have a higher impact on Bitcoin. The positive and negative returns are equally impactful for gold.

Some challenges faced with the GARCH models are that we did not have an efficient way to test various parameters, and therefore could not find a model where all coefficients were statistically significant. Also, in the goodness of fit tests, we rejected the null hypothesis in all cases, suggesting the models are not a good fit. Looking at the QQ plots for each of the models, there appear to be tails in all three, however the tails were minimized compared to the standard GARCH model.

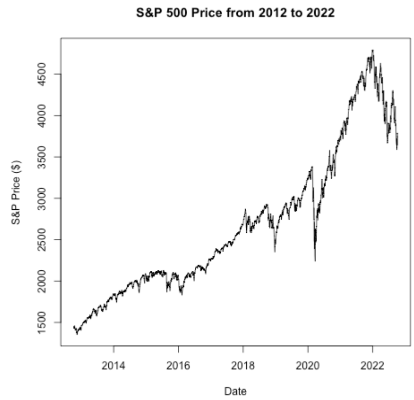
Given more time we would fine tune the parameters to get better fitting models. We would also want to look at various time periods as well as analyzing weekly and monthly returns to see how the volatility changes.

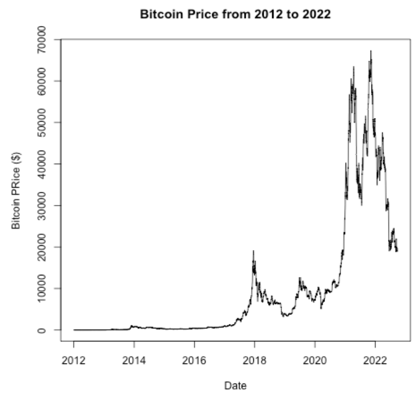
c. ARIMA Model

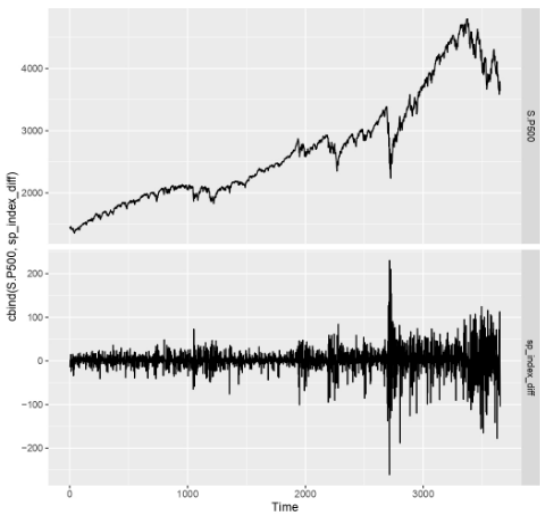
Autoregressive Integrated Moving Average (ARIMA) models utilize past time series data to forecast future outcomes. Rather than using residual errors of the data to predict future values, ARIMA models use auto correlations and moving averages from the historical data. One common use of an ARIMA model is forecasting future market prices. We have used the ARIMA model to predict the future price of two sectors of the market, S&P 500 and Bitcoin.

To begin, we will look at the original S&P 500 and Bitcoin datasets. To be able to create an ARIMA model, we needed to create stationary time-series data. The image on the right is the original time-series data for the S&P 500. The second image on the next page is the original time-series data for bitcoin. Based on both datasets, there is a clear increasing trend in the data as well as seasonality throughout the datasets.

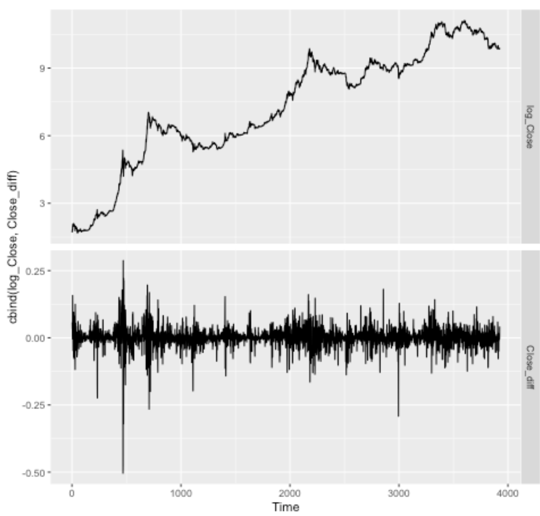
In order to remove both trends, we performed a log-transformation on both datasets. To remove the seasonality in each dataset, we performed differencing on lag 1. Below is each stationary dataset after the changes mentioned above.



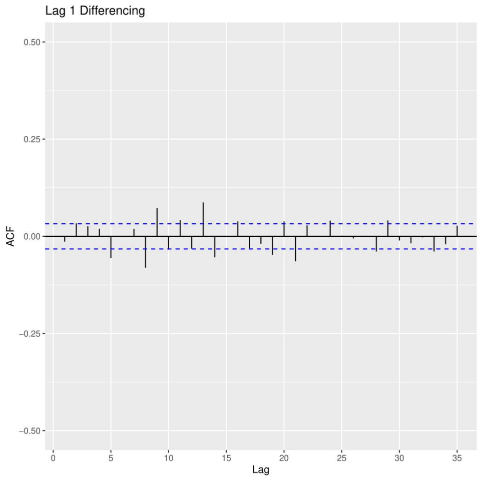


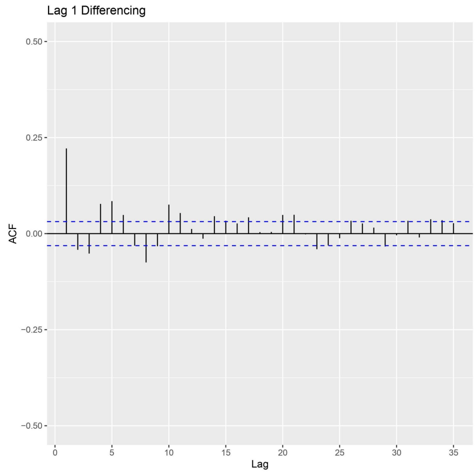


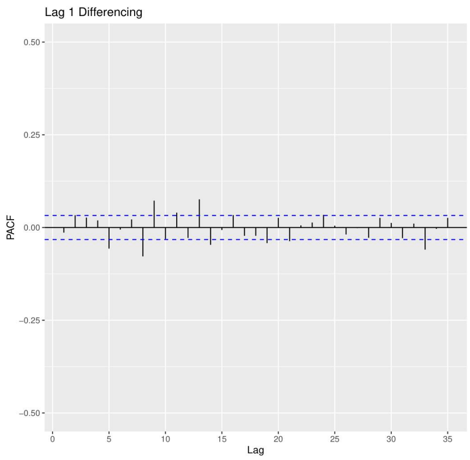
S&P 500 Original vs Stationary data

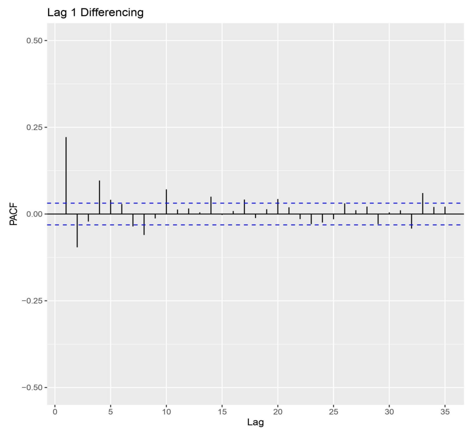


Bitcoin Original vs Stationary data









ACF and PACF S&P 500

ACF and PACF Bitcoin

Now that the time-series datasets are stationary, we can create each ARIMA model. To create the models, we used the auto. Arima function to find the best (p,d,q) values. auto.arima determines the best model according to the model's AIC value. The models that we chose are the following:

S&P 500 Model: ARIMA (2,1,2)

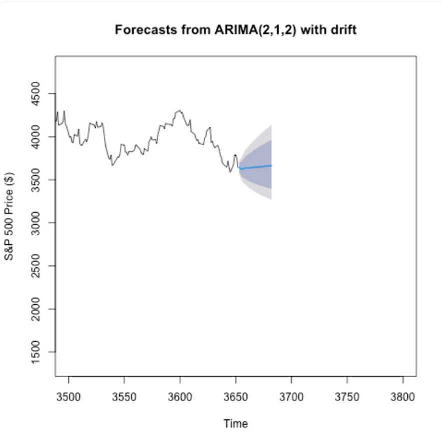
Coefficients	Estimate	Standard Error
Ar1	0.9065	0.2141
Ar2	-0.6375	0.1759
Ma1	-0.9300	0.2023
Ma2	0.6924	0.1647

Bitcoin Model: ARIMA(5,1,0) with Drift

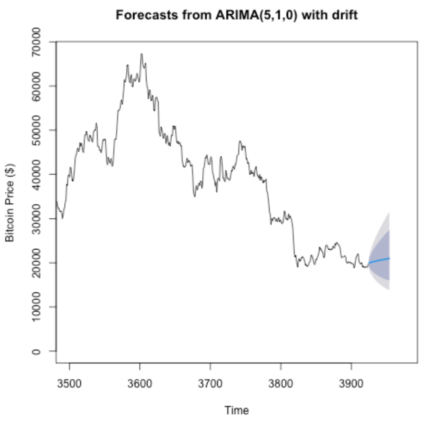
Coefficients	Estimate	Standard Error
AR1	0.2375	0.0160
AR2	-0.0699	0.0164
AR3	-0.0294	0.0164
AR4	0.0721	0.0164
AR5	0.0359	0.0160
Drift	0.0034	0.0014

Forecasting:

For each model created, we generated forecasts to predict the price for the next thirty days. Below are the forecasts for each dataset. The blue line is the predicted price for each day. The shaded areas are the upper and lower values for each forecasted day.



Forecast for S&P 500 Price



Forecast for Bitcoin Price

Conclusions:

Based on the two ARIMA models above, bitcoin's ARIMA model seems to have a better fit to the data due to the significantly lower standard error values for each model variable. To further support this idea, we performed the Ljung-Box test for both models. The corresponding R-squared values are:

- S&P 500: 0.6428
- Bitcoin: 0.9414

Bitcoin's r-squared value is very high which is another indicator of a strong fit to the data.

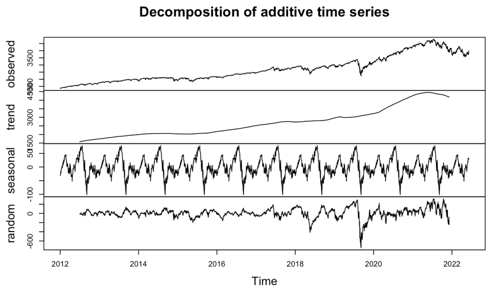
Based on the forecasts, the price of both the S&P 500 and bitcoin will slightly increase over the next thirty days. For each forecast, the model struggled to make a prediction without unreasonable upper and lower values beyond thirty days. However, there seems to be more volatility in the price of bitcoin over the forecasted period.

For further research, we would fine tune the parameters beyond auto.arima to try to improve each model. We could also further improve the forecasting of each dataset by improving the fit of each model. In addition to improving the fit, we could approach the problem differently by changing the period discussed. Instead of a 10-year sample from 2012 to 2022, we could look at data from the last five years.

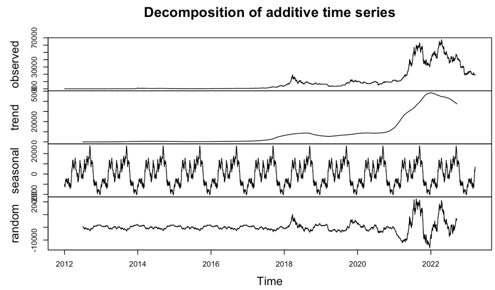
d. Holt-Winters Model

Holt-Winters model can be used to describe time series datasets with additive behavior, those with increasing or decreasing trend with or without seasonality. There are three parameters guarding the model: alpha, beta, and gamma. Alpha is used to describe level and is also the base value; higher alpha value puts more weight on more recent observations. Beta is used to describe trend; higher beta value means the trend slope is more reliant on recent trend values. Gamma is used to describe seasonality; higher gamma value means more emphasis on recent seasonal cycles.

Using the above stationary time-series datasets, we can create another prediction model such as Holt-Winters. To create the models, we first had to decompose the time series data into 3 components: trend, seasonality, and randomness. The plots for both S&P 500 and Bitcoin prices can be seen here:



Decomposition of S&P 500 price



Decomposition of Bitcoin prices

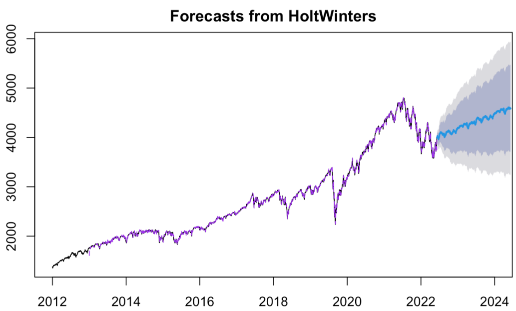
As we compare the decomposition of the models, we can visually see how each of the components has affected our initial data and can pick out the significant factors that will contribute greatly to our predictions. Fitting both datasets to the Holt-Winters model in R gave the following best parameters for S&P 500 and Bitcoin data:

Data	Coefficients	Additive	Seasonal
S&P	Alpha	0.938	0.9112
	Beta	0	0
	Gamma	1	1
Bitcoin	Alpha	0.2	0.7956
	Beta	0.1	0
	Gamma	0.1	1

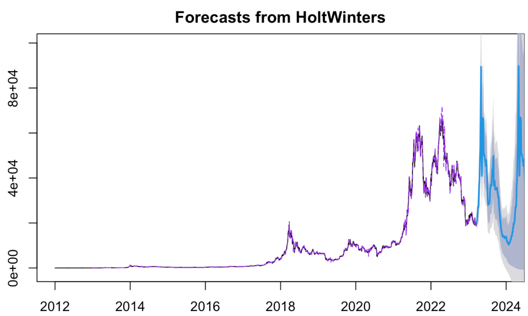
With the parameters above, the S&P dataset can be best modeled with additive behavior while the Bitcoin dataset can be best modeled with seasonal/multiplicative behavior.

Forecast:

For each model created with the best fit, we generated forecasts to predict the price for the next 2 years. The predicted values can be seen below. The blue line is the predicted price, with the light shaded area being the 80% confidence interval and the dark shaded area being the 95% confidence interval.



Forecast of S&P 500 price



Forecast of Bitcoin prices

Conclusions:

The two prediction models exhibited above have very drastic differences when it comes to the prices for the next 2 years. S&P 500 shows additive behavior and constant changes, the prices can be seen to increase steadily. In contrast, Bitcoin exhibits multiplicative behavior and inconsistent trends and thus the predicted prices can be seen to fluctuate. This furthers the initial hypothesis that Bitcoin operates in different cycles and that it is seen to be more volatile than S&P 500.

Our next steps for this model and for the future would be to investigate the prices in more detail and try to predict how each of the datasets would operate 10, 15, or 20 years in the future.

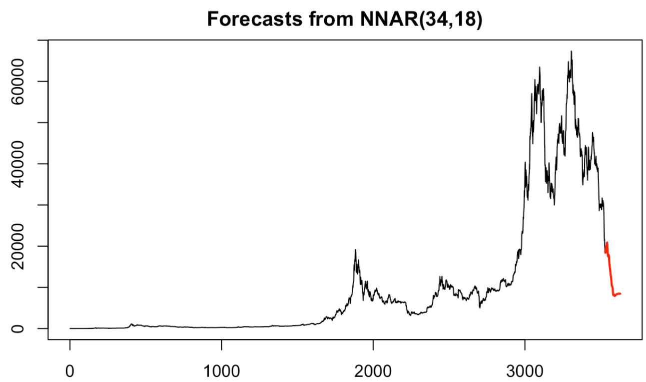
e. Neural Network Model

We have used the feed-forward neural network model with a single hidden layer and lagged inputs for forecasting the time series data for the given stock prices.

Like the earlier models, forecasting in the neural network is performed after detection and removal of trend and seasonality.

For this approach we have chosen the default parameters, like number of nodes in hidden layer is 1 + half of input nodes, lambda = NULL, inputs not scaled etc.

The data was split into train and test, where 'test' represented the last 100 days.

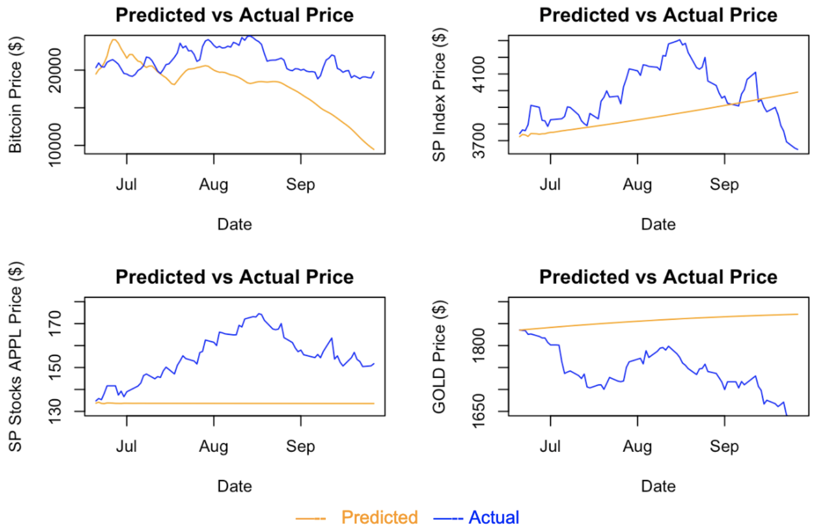


The neural network models were fitted to the different time series. Adjacent is the graph for the full time series including the forecast (in red) towards the end.

Focusing only on the forecasted period (06/20/22 - 09/27/22) revealed the following trends.
Based on the scale of Y axes, it is visually observed that the predictions of SP Index, APPL and GOLD are accurate while BTC predictions are way off. This is also an indicator of high

volatility in BTC prices since close to 10 years of training data fails to predict with decent accuracy.

This is expected, since the prediction window of 100 days is quite large. So, we analyzed the accuracies in different sized windows as follows.




```
------------- 100 days prediction accuracy -----------

Accuracy for BTC forecast
               ME     RMSE      MAE      MPE     MAPE      ACF1 Theil's U
Test set 3196.313 4452.721 3782.329 15.04641 17.95446 0.9521615  8.498662
Accuracy for SP Index forecast
               ME     RMSE      MAE      MPE     MAPE      ACF1 Theil's U
Test set 124.4822 214.4839 174.2719 2.975139 4.311117 0.9328616  4.946705
Accuracy for AAPL stocks forecast
               ME     RMSE      MAE      MPE     MAPE      ACF1 Theil's U
Test set 21.45699 23.83612 21.45699 13.44332 13.44332 0.9579501  10.46937
Accuracy for GOLD price forecast
                ME     RMSE      MAE       MPE     MAPE      ACF1 Theil's U
Test set -115.4676 129.1309 115.4765 -6.732418 6.732904 0.9370056  12.12627
```

```
------------- 50 days prediction accuracy -----------

Accuracy for BTC forecast
               ME    RMSE      MAE      MPE     MAPE
Test set 1049.521 2546.68 2221.553 4.306166 10.12228
Accuracy for SP Index forecast
               ME     RMSE      MAE     MPE    MAPE
Test set 148.5307 177.8744 148.5307 3.70455 3.70455
Accuracy for AAPL stocks forecast
               ME     RMSE      MAE      MPE     MAPE
Test set 15.75518 18.38235 15.75518 10.18856 10.18856
Accuracy for GOLD price forecast
                ME     RMSE      MAE       MPE     MAPE
Test set -84.32063 96.48018 84.33846 -4.848828 4.849799
```

```
------------- 5 days prediction accuracy -----------

Accuracy for BTC forecast
               ME     RMSE      MAE       MPE     MAPE
Test set 85.86183 715.5604 649.6718 0.4247498 3.139444
Accuracy for SP Index forecast
               ME    RMSE      MAE      MPE     MAPE
Test set 62.07913 83.9049 62.07913 1.612946 1.612946
Accuracy for AAPL stocks forecast
              ME     RMSE     MAE      MPE     MAPE
Test set 3.37618 4.229759 3.37618 2.428169 2.428169
Accuracy for GOLD price forecast
               ME     RMSE      MAE        MPE      MAPE
Test set -5.10668 7.178474 5.284933 -0.2794996 0.2892105
```

```
------------- 1 days prediction accuracy -----------

Accuracy for BTC forecast
               ME     RMSE      MAE      MPE     MAPE
Test set 847.3606 847.3606 847.3606 4.168192 4.168192
Accuracy for SP Index forecast
               ME     RMSE      MAE      MPE     MAPE
Test set 18.27546 18.27546 18.27546 0.488348 0.488348
Accuracy for AAPL stocks forecast
               ME     RMSE      MAE       MPE      MAPE
Test set 1.021778 1.021778 1.021778 0.7580379 0.7580379
Accuracy for GOLD price forecast
                ME      RMSE       MAE        MPE       MAPE
Test set 0.4456318 0.4456318 0.4456318 0.02427737 0.02427737
```

Conclusions:

Thiel's U coefficient being much larger than 1 implies Naive forecasting is preferred to the neural network prediction on 100 days. Upon reducing the prediction window/interval, we observed significant improvement in the prediction accuracies for each time series. This improvement observed is consistent in SP Index, AAPL and GOLD prices, while BTC predictions experience worsening when moved from 5 day to 1 day interval. This is another indication of the volatility in BTC prices.

In the future we would explore and utilize other parameters and further tune the model to enhance predictive capabilities. We would also evaluate correlations between BTC and other stocks to rank the stocks with closer resemblance to BTC. We also intend to analyze and evaluate the time series on a weekly, monthly and annual time scales to derive further insights into short-term and long-term investing.

8. Model Comparison

Since the models implemented are dissimilar in structure, metrics like AIC and BIC may not be helpful in comparing their performance. We leverage metrics like MAE, MAPE, MASE etc. to draw comparison of forecast accuracies.

ARIMA

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-8.720688	562.054	222.5957	-0.05173699	2.293156	0.9965193	-0.03834819

HoltWinter

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-1297.46930	1297.48431	1297.46930	-48.02531	48.02531	93.31498	-0.50000

Neural Network

ME	RMSE	MAE	MPE	MAPE	ACF1
3196.313	4452.721	3782.329	15.04641	17.95446	0.9521615

With the understanding that all models require further fine tuning and optimizations, the metrics for the models trained so far reveal that ARIMA performs best in terms of predicting price for BTC since it has the lowest RMSE, ME, MAE and MAPE.

9. Future Scope

Owing to the limited amount of time for this initial analysis and implementation, we intend to accomplish the following tasks in the future.

a. Fine tune selected models and parameters to arrive at an optimized set of models that are better at predictions.
b. Analyze different time periods to see how the models change in different macro-environments.
c. Create models using weekly, monthly and annual returns and generating weekly, monthly and annual forecasts.
d. Check each sector in the S&P 500 companies list and analyze how each one performs vs Bitcoin.
e. Analyze previous S&P 500 bear markets (the dot-com bubble in 2000-2002 and the great recession in 2007-2009) and suggest how Bitcoin could perform when a global bear market comes again since Bitcoin has never experienced one before.

10. Conclusion

Through our analyses, we have determined that Bitcoin is more volatile than more traditional investments. In the short term, Bitcoin can exceed the returns of the S&P 500 in markets that are growing. However, given the extreme volatility of Bitcoin, more traditional investments are preferred for long-term benefits.

Works Cited

Chan, Wing Hong, Minh Le, and Yan Wendy Wu. “Holding Bitcoin Longer: The Dynamic Hedging Abilities of Bitcoin.” Quarterly Review of Economics & Finance 71 (February 2019): 107–13. doi:10.1016/j.qref.2018.07.004.

Duggan, Wayne. “Is Bitcoin Worth Investing In?” Edited by Rachel McVearry. U.S. News, September 22, 2022. https://money.usnews.com/investing/cryptocurrency/articles/is-bitcoin-worth-investing-in.

Kapar, Burcu, and Jose Olmo. “Analysis of Bitcoin Prices Using Market and Sentiment Variables.” World Economy 44, no. 1 (January 2021): 45–63. doi:10.1111/twec.13020.

Kenton, Will. 2020. “What Is the GARCH Process? How It’s Used in Different Forms.” Investopedia. October 20, 2020. https://www.investopedia.com/terms/g/generalalizedautogregressiveconditionalheteroskedasticity.asp.

Klein, Tony, Hein Pham Thu, and Thomas Walther. “Bitcoin is Not the New Gold – A Comparison of Volatility, Correlation, and Portfolio Performance.” International Review of Financial Analysis 59 (July 2018): 105-16. Doi: 10.1016/j.irfa.2018.07.010.

Saleem, Shaikh Zoaib. “The 2008 Global Meltdown and the Birth of Bitcoin.” mint, September 13, 2018. https://www.livemint.com/Money/YTYMYUD7dytGK5PGSpdRTN/The-2008-global-meltdown-and-the-birth-of-Bitcoin.html.

Voigt, Kevin, and Andy Rosen. “Cryptocurrency: Definition and Current Prices.” Edited by Michael Randall. NerdWallet, September 30, 2022. https://www.nerdwallet.com/article/investing/cryptocurrency.