Bitcoin Analysis - R Notebook

Does Bitcoin have a cycle?



- **Disclaimer:** I am not a professional investment advisor, and this material does not constitute investment advice. Please conduct your own research and due diligence before making any investment decisions.
- ChatGPT was utilized to assist with coding tasks during this analysis, including picture generation, code generation, and code debugging. However, all generated code was thoroughly inspected and tailored to meet the specific goals of the analysis.

APPASA

In this analysis, I follow the APPASA (Ask, Prepare, Process, Analyze, Share, Act) outline in the Google Data Analytics course. The breakdown of each section is as follows:

- Ask
 - Key Intrigues
- Prepare/Process
 - ETL
 - Examining The Data
- Analyze/Share
 - Examining the Data
 - Normalized Pricing
 - Statistical Significance
 - Forecasting Performance
 - Monetary Performance
 - Current Cycle
- Act
 - Conclusion

Background

A few months ago, I was listening to one of my favorite podcasts, All-In. The hosts' conversation was progressing as normal, talking about the top news stories of that week. That is until the chat changed topics to crypto-currency (All-In Podcast, 2024).

https://www.youtube.com/watch?v=R6hJh-OwoZw&t=3184s

One on the hosts, Chamath Palihapitiya, started speaking about Bitcoin and the halving periods that occurs every four years. These halving dates are when Bitcoin's reward for mining decreases by half and the difficulty to 'mine' a new Bitcoin increases by 2. (Conway, 2024) Hence the term halving. This helps to ensure that Bitcoins appreciate in value overtime as supply is constrained. The dates of when a Bitcoin halving has occurred are:

- November 28th, 2012
- July 9th, 2016
- May 11th, 2020

• April 19th, 2024

As shown the, latest one happened in April of 2024, the current year of this analysis. Chamath received a tip from one of his investor acquaintances to study Bitcoin's behavior after halving. He then referred to some graphs created by one of his associates, which focused on Bitcoin's performance over 18-month periods after every halving. Chamath aimed to demonstrate a cyclical pattern in the pricing behavior of Bitcoin in these time frames (All-In Podcast, 2024).

However, the graphs struck me as being inconclusive and not particularly impactful. The comparison of price movements over time wasn't scaled or normalized. Resulting in the earlier dates (when Bitcoin was valued in the hundreds of dollars) to appear flat compared its recent performance (in the tens-of-thousands). Additionally, he attempted to show bar charts that suggest price correlation, but it was difficult to find conclusive evidence in the analysis.

Key Intrigues

Overall, I finished the podcast but felt the answer of Bitcoin's cyclical nature had not been resolved. This is what led me to ask the following questions:

- Does Bitcoin have a cyclical pattern in relation to each halving period?
 - If so, can the cyclical pattern be observed in the price movements?
 - What does the average cycle look like?
 - Is the cyclical pattern only limited to 18 months or span the entire 4 years (48 months)?
 - Is it statistically significant?
- If a cyclical pattern exists, can I forecast the ending price of the current halving period, 2024-2028?
 - What is the rate of growth over the last halving periods?
 - What is the monetary gain/loss, including predicted?
 - What is the range of potential future price values with error?
- If halving period cyclical cycles exist, then do yearly prices also reflect these cycles?
- If halving period cyclical cycles exist, where is the current period, 2024-2028, in the cycle?

ETL

Libraries

```
#Load Packages
library(ggpattern) # For Visual
library(scales) # For formatting
library(ggplot2)
library(readr)

##
## Attaching package: 'readr'

## The following object is masked from 'package:scales':
##
## col_factor
```

library(dplyr)

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
## filter, lag

## The following objects are masked from 'package:base':
##
intersect, setdiff, setequal, union
```

Data Loading & Transform

Here, I load in the data. The data can be found and download for free from CoinMarketCap (Bitcoin Price Today, BTC to USD Live Price, Marketcap and Chart | CoinMarketCap, n.d.). This data has daily prices of Bitcoin starting in July 14th, 2010 to present (September 22nd, 2024).

```
#Load the Data
file_path <- "data/Bitcoin_5_13_2010-7_12_2010_historical_data_coinmarketcap.csv"
df <- read.csv(file_path, sep = ";", stringsAsFactors = FALSE)
# Reverse Data Order, Oldest to Newest
df <- df[nrow(df):1, ]
# Create standardized timestamp column for filtering
df$timestamp <- as.Date(df$timestamp, format="%Y-%m-%d")
head(df)</pre>
```

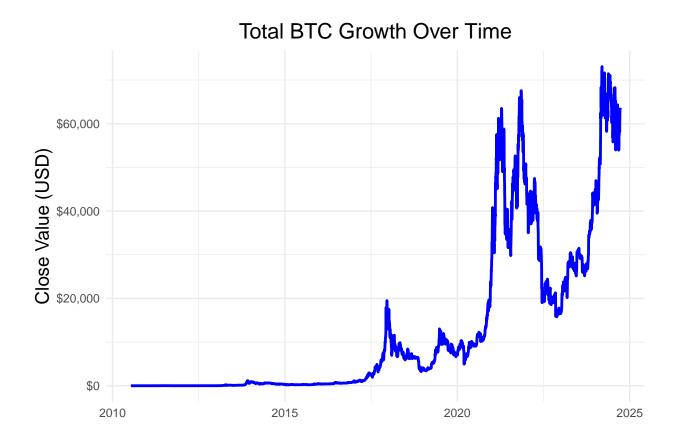
```
timeOpen
                                                timeClose
                                                                           timeHigh
## 5185 2010-07-14T00:00:00:00.000Z 2010-07-14T23:59:59.999Z 2010-07-14T00:34:00.000Z
## 5184 2010-07-15T00:00:00:00.000Z 2010-07-15T23:59:59.999Z 2010-07-15T11:39:00.000Z
## 5183 2010-07-16T00:00:00:00.000Z 2010-07-16T23:59:59.999Z 2010-07-16T02:11:00.000Z
## 5182 2010-07-17T00:00:00:00.000Z 2010-07-17T23:59:59.999Z 2010-07-17T06:51:00.000Z
## 5181 2010-07-18T00:00:00:00.000Z 2010-07-18T23:59:59.999Z 2010-07-18T17:38:00.000Z
## 5180 2010-07-19T00:00:00:00.000Z 2010-07-19T23:59:59:999Z 2010-07-19T14:09:00.000Z
##
                         timeLow name
                                            open
                                                        high
## 5185 2010-07-14T19:24:00.000Z 2781 0.05815725 0.06158806 0.04864654 0.05640216
## 5184 2010-07-15T00:41:00.000Z 2781 0.05640261 0.06795437 0.05396921 0.05756808
## 5183 2010-07-16T00:24:00.000Z 2781 0.05800138 0.07222029 0.05748353 0.06649170
## 5182 2010-07-17T16:21:00.000Z 2781 0.06649990 0.07773513 0.05741781 0.06599255
## 5181 2010-07-18T00:28:00.000Z 2781 0.06608795 0.08085810 0.06422061 0.07881380
## 5180 2010-07-19T09:31:00.000Z 2781 0.07879913 0.08360209 0.06903579 0.07416855
        volume marketCap timestamp
##
## 5185 261.54 190259.6 2010-07-14
## 5184 445.80 195982.1 2010-07-15
## 5183 497.25 228047.4 2010-07-16
## 5182 19.99 226904.8 2010-07-17
## 5181 75.13 271669.2 2010-07-18
## 5180 573.24 256302.4 2010-07-19
```

Examining The Data

Bitcoin's Price

Starting, I wanted to examine Bitcoin's price performance over the entire dataset. This allowed me to verify the veracity of the data. Additionally, I can observe if there's any obvious pattern. However, at this scale, the earliest years, 2010 to 2017, appear flat. The later years, after 2017, appear to have random movements in price. At this observational scale, it's difficult to ascertain any usable insight.

```
# Plot the actual close value of Bitcoin over time
p <- ggplot(df,</pre>
        aes(
         x = as.Date(timestamp),
         y = close)
        ) +
  geom_line(
    color = "blue",
    linewidth = 1
  labs(title = "Total BTC Growth Over Time",
       x = "",
       y = "Close Value (USD)") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 16),
    axis.title.x = element_text(size = 14),
    axis.title.y = element_text(size = 14)
  scale_y_continuous(labels = dollar_format())
print(p)
```



Creating Halving Periods Dataframes

In this section, I created the dataframes that would be used later for analysis. I filtered the data to match the weekly reported close price value on CoinMarketCap (Bitcoin Price Today, BTC to USD Live Price, Marketcap and Chart | CoinMarketCap, n.d.). This is to reduce random noise of day-to-day price movements while still allowing for potential to show cyclical patterns. Every weekly close price is reported on a Sunday, hence why it's filtered to only day 0. Additionally, for ease of date use, I re-formmatted the dates into timestamps.

```
# Create a column for day of the week
df$day_of_week <- as.POSIXlt(df$timestamp)$wday

# Create DataFrame filtered for weekly close value, recorded on Sundays (0)
sundays_df <- df %>% filter(day_of_week == 0)

# Extract the timestamp column and close columns only
weekly_data <- sundays_df %>% select(timestamp, close)

# Function for filtering data based on halving periods with start & end dates
filter_halving_period_manual <- function(start_date, end_date, data) {
    start_date <- as.Date(start_date)
    end_date <- as.Date(end_date)

# Ensure that timestamp is properly formatted and filter by dates
filtered_data <- data %>%
```

```
filter(as.Date(timestamp) >= start_date & as.Date(timestamp) < end_date)

return(filtered_data)
}

# Apply the filter to extract last three halving periods for weeks
halving_2012_to_2016_weekly <- filter_halving_period_manual("2012-11-28", "2016-07-09", weekly_data)
halving_2016_to_2020_weekly <- filter_halving_period_manual("2016-07-09", "2020-05-11", weekly_data)
halving_2020_to_2024_weekly <- filter_halving_period_manual("2020-05-11", "2024-04-19", weekly_data)</pre>
```

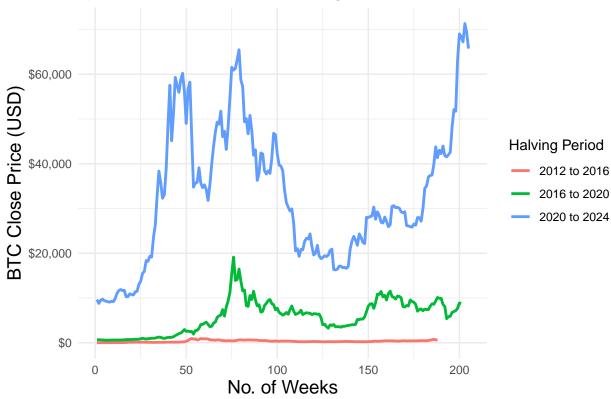
Comparing Prices over Halving Periods

This analysis nearly recreates the chart shown by Chamath Palihapitiya in the All-In podcast (All-In Podcast, 2024). This shows the problem I faced in the viewing the original analysis. Without proper scaling or normalization, its difficult to ascertain if price movements mimic each other over each halving period.

```
# Use weeks as index to ensure graphical compatability
halving_2012_to_2016_weekly$index <- seq_len(nrow(halving_2012_to_2016_weekly))
halving_2016_to_2020_weekly$index <- seq_len(nrow(halving_2016_to_2020_weekly))
halving_2020_to_2024_weekly$index <- seq_len(nrow(halving_2020_to_2024_weekly))
# Labels to identify the halving periods
halving_2012_to_2016_weekly$period <- "2012 to 2016"
halving 2016 to 2020 weekly period <- "2016 to 2020"
halving_2020_to_2024_weekly$period <- "2020 to 2024"
# Combine data for plotting
combined weekly close data <- rbind(
  halving_2012_to_2016_weekly,
 halving_2016_to_2020_weekly,
 halving_2020_to_2024_weekly
# Plot the dollar close value of the halving periods over each other
p <- ggplot(combined_weekly_close_data,</pre>
            aes(
              x = index,
              y = close,
              color = period)
            ) +
  geom line(linewidth = 1) +
  labs(title = "Weekly BTC Values Over Halving Periods (Overlayed)",
       x = "No. of Weeks",
       y = "BTC Close Price (USD)",
       color = "Halving Period") +
  theme_minimal() +
  theme(
   plot.title = element_text(hjust = 0.5, size = 16),
   axis.title.x = element_text(size = 14),
   axis.title.y = element_text(size = 14)
  ) +
  scale_y_continuous(labels = dollar_format())
```







Determining Price Patterns

As an acid test, to determine if the prices have any similarities, whatsoever, I compared the price value of Bitcoin in each period to each other in a side-by-side comparison. There seems to be some similarities, with 2012-2016 and 2020-2024 having similar price spikes and troughs. 2016 to 2020 appears to have less commonalities, but still shares similar movements around the same time frames. This passed the acid test and prompted further investigation.

```
theme(
   plot.title = element_text(hjust = 0.5, size = 16),
   axis.title.x = element_text(size = 14),
   axis.title.y = element_text(size = 14)
) +
facet_wrap(~ period, scales = "free_y") +
scale_y_continuous(labels = dollar_format())

print(p)
```

Weekly BTC Values Over Halving Periods (Side-by-Side)



Normalized Pricing

Comparing pricing directly has a downside because of the difference in value. This makes it difficult to see patterns across huge value jumps. In order to counteract this, I normalized the pricing using the following formula:

 $Normalized Value = \frac{ClosePrice-Minimum of Halving Period}{Range of the Halving Period}$

The range is give by:

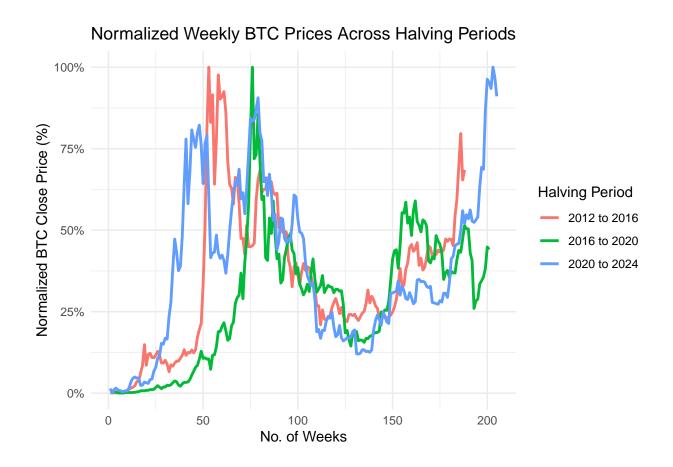
Range of the Halving Period = Maximum of Halving Period - Minimum of Halving Period

This resulted in a percentage value instead of a price value. Every price is scaled from 0% (the *Minimum of Halving Period*) to 100% (the *Maximum of Halving Period*). Now the prices of each halving period could be overlaid on the same graph and same y-axis without fear that differences in values could cause distortion.

Comparing the Normalized Prices

In this section, I built and applied the normalization function to Bitcoin's weekly close prices. Now that all the prices were percentages and could be graphed on the same axis, I examined the overlaid line patterns for similarities. As the graph more clearly shows, the pattern of prices have considerable similarity. The key differences are in timing, with 2020-2024 period peaking earlier than the others. However, the overall pattern of peaks and troughs can be seen very conclusively.

```
# Function to normalize the price close values by scaling them in percentage form, removing NA values
normalize_close <- function(data) {</pre>
  close_min <- min(data$close, na.rm = TRUE)</pre>
  close_max <- max(data$close, na.rm = TRUE)</pre>
  close_range <- close_max - close_min</pre>
  data$close_normalized <- ((data$close - close_min) / close_range) * 100</pre>
  return(data)
# Apply the normalization function for each halving period
halving_2012_to_2016_weekly <- normalize_close(halving_2012_to_2016_weekly)
halving_2016_to_2020_weekly <- normalize_close(halving_2016_to_2020_weekly)
halving_2020_to_2024_weekly <- normalize_close(halving_2020_to_2024_weekly)
# Combine all normalized data
combined weekly data <- rbind(</pre>
 halving_2012_to_2016_weekly[, c("index", "close_normalized", "period")],
 halving_2016_to_2020_weekly[, c("index", "close_normalized", "period")],
 halving_2020_to_2024_weekly[, c("index", "close_normalized", "period")]
p <- ggplot(combined_weekly_data,</pre>
            aes(
              x = index,
              y = close_normalized,
              color = period)
  geom_line(linewidth = 1) +
  labs(
    title = "Normalized Weekly BTC Prices Across Halving Periods",
    x = "No. of Weeks",
    y = "Normalized BTC Close Price (%)",
    color = "Halving Period"
  theme_minimal() +
  scale_y_continuous(labels = scales::label_percent(scale = 1))
print(p)
```



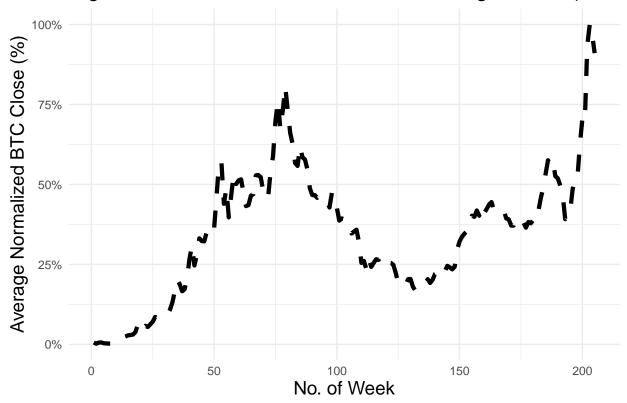
Finding the Average of the Cycles

Now the potential for a cyclical pattern has been confirmed, I wanted to find the average cycle. This would allow for a "baseline" performance to be assessed.

```
# Calculate the average normalized value for each week between halving periods
average_normalized_data <- combined_weekly_data %>%
  group_by(index) %>%
  summarize(average_close_normalized = mean(close_normalized, na.rm = TRUE))
p <- ggplot(average_normalized_data,</pre>
            aes(
              x = index,
                = average_close_normalized)
            ) +
  geom_line(
    color = "black",
    linewidth = 1.5,
    linetype = "dashed") +
  labs(
    title = "Average Normalized BTC Prices Across Halving Periods (Weekly)",
    x = "No. of Week",
    y = "Average Normalized BTC Close (%)") +
  theme_minimal() +
```

```
theme(
   plot.title = element_text(hjust = 0.5, size = 16),
   axis.title.x = element_text(size = 14),
   axis.title.y = element_text(size = 14)
) +
   scale_y_continuous(labels = scales::label_percent(scale = 1))
print(p)
```

Average Normalized BTC Prices Across Halving Periods (Weel-

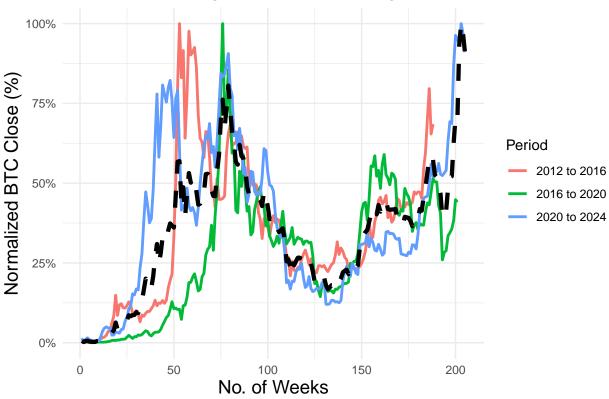


Showing Cyclicality

Again, overlaying the normalized prices, I also included the average performance line to compare. All three halving periods show that the value of Bitcoin adheres to the average pattern. This helps to visualize the cycles with a singular reference point.

```
geom_line(
    data = average_normalized_data,
      aes(
        x = index,
        y = average_close_normalized
    color = "black",
    linewidth = 1.5,
    linetype = "dashed"
  labs(
    title = "Comparison: Halving Periods vs Average Normalized",
    x = "No. of Weeks",
    y = "Normalized BTC Close (%)",
    color = "Period") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 16),
    axis.title.x = element_text(size = 14),
    axis.title.y = element_text(size = 14)
  ) +
  scale_y_continuous(labels = scales::label_percent(scale = 1))
print(p)
```

Comparison: Halving Periods vs Average Normalized



Statistical Significance

In this section, the cyclicality of Bitcoin is tested for statistical significance. This will determine whether the cyclical patterns observed in the normalized data are related to the week (index) of the cycle or just coincidence.

Given that this test is for cyclicality, the most appropriate model is the sinusoidal model, which uses sine waves to demonstrate patterns of growths, and contractions. Linear models wouldn't work here as they are used to demonstrate a 1:1 relationship.

The key factors of the sinusoidal model are:

- Amplitude (a): Shows the magnitude of price swings.
- Frequency (b): Indicates how often these cycles occur.
- Phase Shift (c): Suggests that the cycles may start slightly earlier or later than expected.
- Vertical Shift (d): Represents the average level around which Bitcoin's price oscillates.

Percentage Change

These first two sections are for testing percentage change. This is to show if Bitcoin's price movements have any causal relation to the supposed cycles.

```
# Function to calculate percentage change of BTC Prices weekly, using weekly here instead of daily to m
calculate_percentage_change_original <- function(data) {</pre>
  data$percent_change <- c(NA, diff(data$close) / data$close[-length(data$close)] * 100)
  return(data)
}
# Apply percentage change to BTC price data in dataframes
halving_2012_to_2016_weekly <- calculate_percentage_change_original(halving_2012_to_2016_weekly)
halving_2016_to_2020_weekly <- calculate_percentage_change_original(halving_2016_to_2020_weekly)
halving 2020 to 2024 weekly <- calculate_percentage_change_original(halving 2020 to 2024 weekly)
combined_percentage_change_data <- rbind(</pre>
  halving_2012_to_2016_weekly[, c("timestamp", "close_normalized", "percent_change", "period", "index")
 halving_2016_to_2020_weekly[, c("timestamp", "close_normalized", "percent_change", "period", "index")
  halving_2020_to_2024_weekly[, c("timestamp", "close_normalized", "percent_change", "period", "index")
# Removing missing values caused by percentage change calculation
filtered_data <- combined_percentage_change_data %>%
  filter(!is.na(percent_change))
# Fit the sinusoidal model using inital guess of a=1, b=2, c= 0, d=0, 1500 Interations
sinusoidal_model_weekly <- nls(</pre>
  percent_change ~ a * sin(b * index + c) + d,
  data = filtered_data,
  start = list(a = 1, b = 2 * pi / 208, c = 0, d = 0),
  control = nls.control(maxiter = 1500)
```

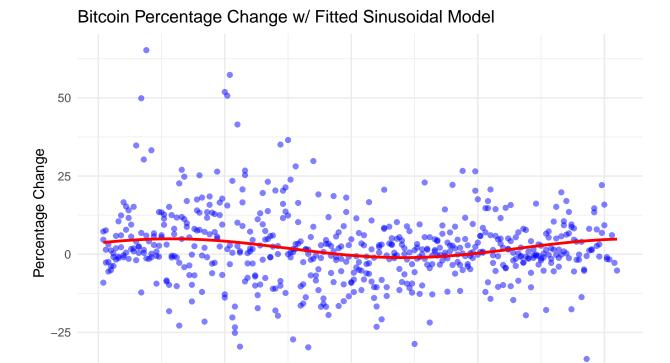
```
s <- summary(sinusoidal_model_weekly)</pre>
print(s)
##
## Formula: percent_change ~ a * sin(b * index + c) + d
##
## Parameters:
   Estimate Std. Error t value Pr(>|t|)
##
## a 3.003400  0.660800  4.545 6.67e-06 ***
## b 0.034225  0.004787  7.149  2.60e-12 ***
## c 0.584750  0.572143  1.022 0.307185
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 11.41 on 587 degrees of freedom
##
## Number of iterations to convergence: 24
## Achieved convergence tolerance: 6.335e-06
```

Interpreting the results, all variables except Phase Shift show significant p-values (< 0.01). This suggests that Bitcoin's cyclicality plays a significant role in price fluctuations. However, the Residual Standard Error is considerable at 11.41, indicating that while price variability is influenced by cyclicality, it's not the sole explanatory factor of price movement.

```
# Added predicted values to dataframe
filtered_data$predicted_values <- predict(sinusoidal_model_weekly)

# Plot predicted (red) values versus actual values (blue) to grade fit
p <- ggplot(filtered_data, aes(x = index)) +
    geom_point(aes(y = percent_change), color = "blue", alpha = 0.5) +
    geom_line(aes(y = predicted_values), color = "red", linewidth = 1) +
    labs(
        title = "Bitcoin Percentage Change w/ Fitted Sinusoidal Model",
        x = "Weeks",
        y = "Percentage Change") +
    theme_minimal()

print(p)</pre>
```



The graph illustrates the distribution of percentage changes and how closely they align with the predicted sinusoidal cycle. While the overall pattern suggests a cyclical behavior in Bitcoin's price movement, there's still substantial variability.

50

100

Weeks

150

200

Normalized Change

0

The following sections test normalized price changes to determine if the value of Bitcoin has any causal relation to the observed cyclical patterns.

```
# Remove missing rows again
filtered_data <- combined_percentage_change_data %>%
    filter(!is.na(close_normalized))

# Fit the sinusoidal model using inital guess of a=1, b=1, c= 0, d=0, 1500 Interations
sinusoidal_model_weekly_norm <- nls(
    close_normalized ~ a * sin(b * index + c) + d,
    data = filtered_data,
    start = list(a = 1, b = 1 * pi / 208, c = 0, d = 0),
    control = nls.control(maxiter = 1500)
)

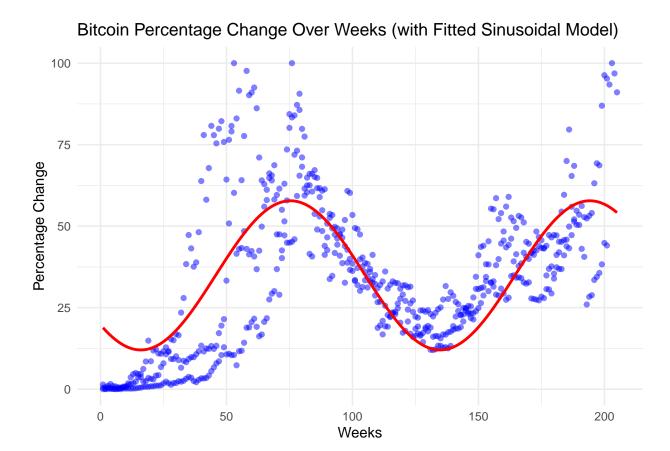
s <- summary(sinusoidal_model_weekly_norm)
print(s)</pre>
```

```
## Formula: close_normalized ~ a * sin(b * index + c) + d
##
## Parameters:
      Estimate Std. Error t value Pr(>|t|)
##
## a -2.288e+01 9.074e-01 -25.22
                                    <2e-16 ***
## b -5.289e-02 8.064e-04 -65.58
                                    <2e-16 ***
## c 2.417e+00 9.245e-02 26.14
                                    <2e-16 ***
## d 3.491e+01 6.798e-01 51.34
                                    <2e-16 ***
## ---
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' 1
## Residual standard error: 16.06 on 590 degrees of freedom
## Number of iterations to convergence: 32
## Achieved convergence tolerance: 7.221e-06
```

Again, the results show significant p-values (< 0.01) across all variables, Amplitude, Frequency, Phase, and Shift. This is suggesting that Bitcoin's cyclicality significantly impacts its price value. The Residual Standard Error is 16.06, reflecting the variability is not fully accounted for by cyclicality alone.

```
# Added predicted values to dataframe
filtered_data$predicted_values_norm <- predict(sinusoidal_model_weekly_norm)

# Plot predicted (red) values versus actual values (blue) to grade fit
p <- ggplot(filtered_data, aes(x = index)) +
    geom_point(aes(y = close_normalized), color = "blue", alpha = 0.5) +
    geom_line(aes(y = predicted_values_norm), color = "red", linewidth = 1) +
    labs(
        title = "Bitcoin Percentage Change Over Weeks (with Fitted Sinusoidal Model)",
        x = "Weeks",
        y = "Percentage Change") +
    theme_minimal()</pre>
```



The graph shows the distribution of the normalized values and how closely they adhere to the predicted sinusoidal cycles. The fit of Bitcoin's normalized price closely follows the prediction, yet still has notable variability.

Both tests indicate existence of causal cyclical patterns in Bitcoin's price movement. This is confirmed by the significance of the p-values. However, the residual errors are also considerable too, at 11.41 and 16.06. These indicates that price value and movement variability isn't fully explained by cyclicality. This is as expected as no equity's or currency's price can be entirely predicted. Prices are often influenced by random occurences and variables in the market including inflation, news, stock market performance, political events, and investor sentiment. The key takeaway is that Bitcoin's price movement does exhibit a statistically significant cyclical pattern.

Forecasting Performance

Calculating Halving Period Percentage Change.

This code section is needed for forecasting a percentage return for the current halving period (2024-2028). The start and finish values of each halving period are needed to show how much Bitcoin may have grown. The resulting dataframe shows that the greatest growth occurred in 2012 to 2016 at over 5000%. The more recent 2020 to 2024 is the lowest at only 650% gain.

```
# Function to filter daily data, with error catch
filter_halving_period_daily <- function(start_date, end_date) {
   start_date <- as.Date(start_date)</pre>
```

Bibliography

All-In Podcast. (2024, May 31). Trump verdict, COVID Cover-up, Crypto Corner, Salesforce drops 20%, AI correction? [Video]. YouTube. https://www.youtube.com/watch?v=R6hJh-OwoZw

 $Bitcoin\ price\ today,\ BTC\ to\ USD\ live\ price,\ marketcap\ and\ chart\ /\ CoinMarketCap.\ (n.d.).\ CoinMarketCap.\ https://coinmarketcap.com/currencies/bitcoin/historical-data/$

Chen, J. (2023, September 8). *Elliott Wave Theory: What It Is and How to Use It.* Investopedia. https://www.investopedia.com/terms/e/elliottwavetheory.asp

Conway, L. (2024, August 8). Bitcoin Halving: What It Is and Why It Matters for Crypto Investors. Investopedia. https://www.investopedia.com/bitcoin-halving-4843769