

# Time Series Forecasting: Cardano(ADA)

...

Springboard Capstone Project 3  
Hastings Reeves  
August 18, 2021

# Overview: What is Cardano?

-Proof of Stake blockchain platform:  
founded on peer-reviewed research and  
developed through evidence-based  
methods.

-Seeks to combine pioneering  
technologies with security and  
sustainability through it's decentralized  
applications, systems, and societies.



# Use Cases for Cardano

## Example 1: Education

-ATALA: ID and Credentials solution built on Cardano Blockchain.

-Provide secure method for storing immutable academic transcripts that students or institutions can store and retrieve from a tamper-free ecosystem.

## Example 2: Retail

-ATALA: Anti-counterfeit solutions through the blockchain technology.

-Provide a method to determine whether or not a good is fake, taking an ongoing problem of counterfeit goods (\$300bn loss in 2020) by determining authenticity so consumers can buy with confidence

## Example 3: Agriculture

-ATALA, EMURGO: Traceability and transparency solution through the Cardano Blockchain.

-Provide farmers, haulers and retailers with a way to confidently track and manage their goods from field to store. Decreases interruptions of the supply chain.

**Project Objective:**  
Analyze price action for  
Cardano's native token, ADA,  
and predict future price action.

## Data:

- I was able to download csv files for BTC, ETH and ADA off of the Kaggle Competition website.
- The date range began in 2013 for BTC, 2015 for ETH and 2017 for ADA. 2013 was the first major bull run for BTC where the market grew over 1000% in under a year.

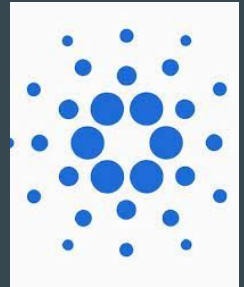
## Libraries:

- EXPLORATORY DATA ANALYSIS: Pandas, Numpy, Matplotlib.pyplot, Seaborn
- PREPROCESSING: Statsmodels(api, tsa.stattools, ARIMA), sklearn.metrics
- MODELING: pmdarima.arima, ADFTTest, fbprophet

## Features:

- SNo : Serial number
- Date : Date of observation
- Open : Opening price on a given day
- High : Highest price on a given day
- Low : Lowest price on a given day
- Close : Closing price on a given day
- Volume : Volume of transactions on a given day
- Market Cap : Market capitalization in USD

# Understanding the market: Bitcoin, Ethereum and Cardano





# Bitcoin

## The Original Cryptocurrency:

- Deflationary Asset
- Challenges the legacy banking system and governmental currencies.
- Early '21 marked the first time that total market cap valuation exceeded \$1.0 Trillion.

---

# Ethereum



Blockchain technology:

- Thousands of Decentralized applications are run through the Ethereum Blockchain.
- Created a platform for NFT's, decentralized finance, buying and trading native tokens for different projects.

---





# Cardano

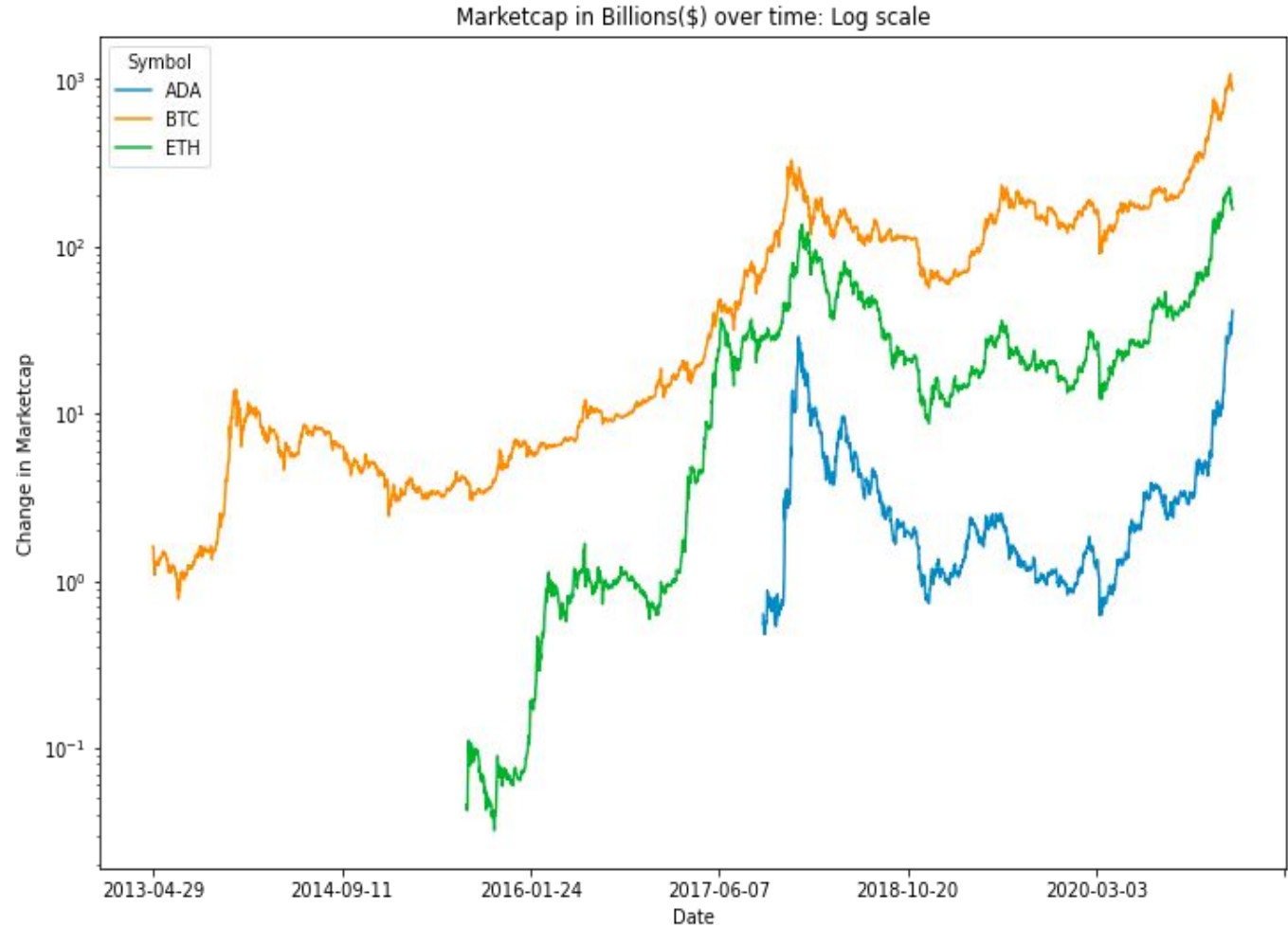
## Blockchain Competitor:

- Seeks to reduce energy consumption by using proof of stake to incentivize ADA miners to retain their tokens in order to retain their influence on it's blockchain.
  - Addresses similar issues that Ethereum is attempting to solve.
-

**Market Cap:** As BTC moves, so moves the market.

**Implications:**

- As we can see, BTC's fluctuation in Market Cap value has a tremendous impact on ETH and ADA.
- Market Cap Values:
  - BTC: \$861.32 B
  - ETH: \$372.4 B
  - ADA: \$66.09 B



# Market trends: Macro moves

## Transaction Volume:

As BTC moves, so moves the market.

### Implications:

- Again, BTC leads the way for all native tokens in transaction volume per day.

## Price:

Percentage of Price Change over time

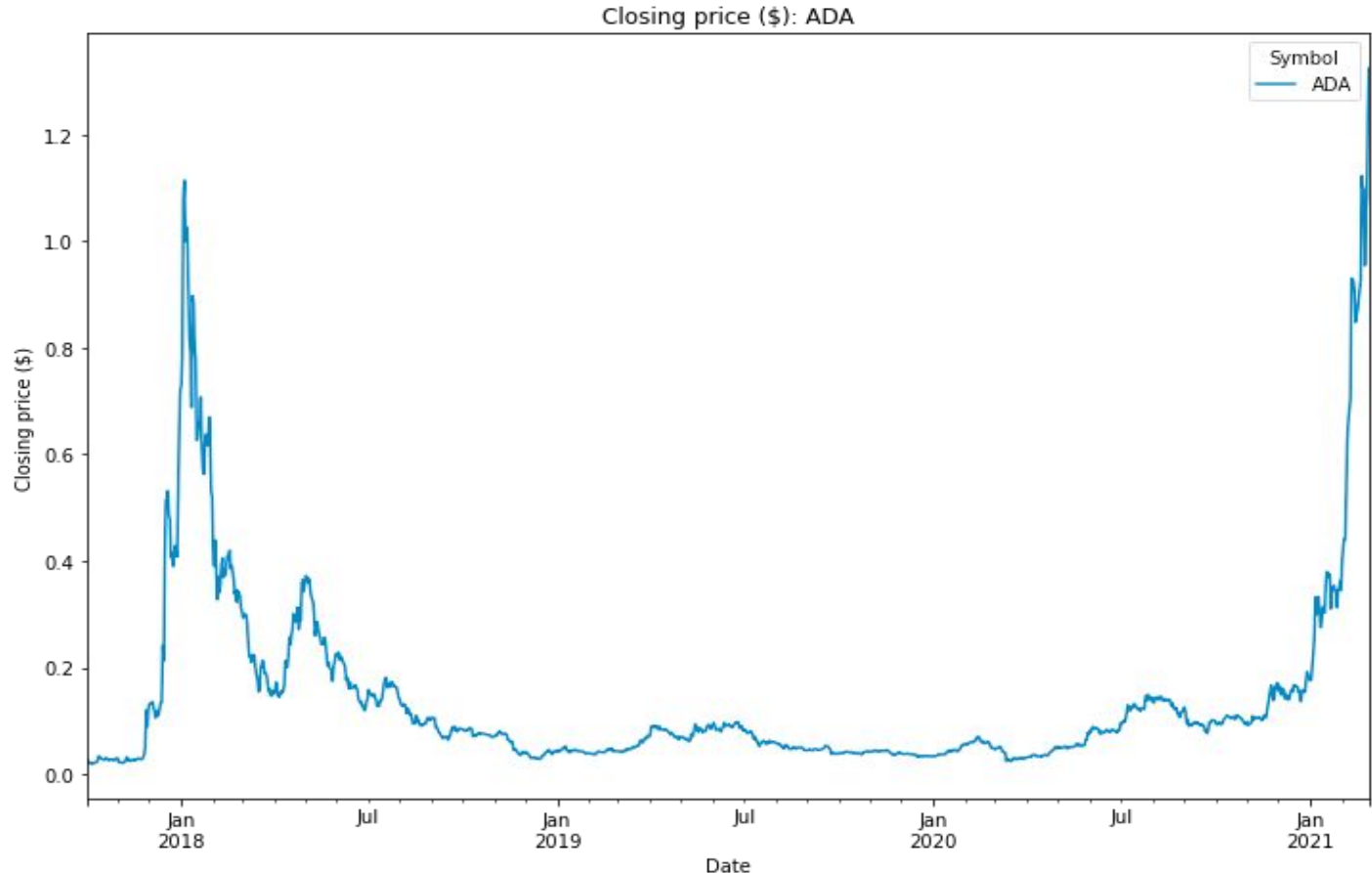
### Shifting of big gains:

- While BTC experienced astronomical returns early on, ETH and ADA look to follow a similar path and provide an opportunity to return more value on an investment.

# Predictive analysis: Prepare data for ARIMA model

## Constraints:

- Non-stationary data
- Large peaks, volatile price movement
- Lack of clear trend
- Lack of seasonality



# Pre-processing:

## Gaps or duplicates in Date Column

- Check to make sure there are no gaps in dates
- Compare DF date column with a set of expected dates.

## Trend, Seasonality, Noise

- There is an upward trend, driven by two massive peaks
- There is no seasonality
- Variance in the price is high and volatile.

## Stationary data

- KPSS(y) where y is the target variable
- Indicated a need to difference the data in order to achieve stationarity. Difference = 1

## Train/Test Split

- I chose a 80/20 split for my data.

# Building the Baseline ARIMA Model

## Determining P, D, Q values:

1. Evaluate\_Arima\_Model function that uses the Mean Absolute Error as an indicator for Best Model.
  - a. Pass this function into Evaluate\_Models function that iterates through a range of 0-3 to determine which P, D, Q values performance best based off of their MAE.
  - b. Result: ARIMA(1,0,2)
    - i. d should be 1
2. Set Difference to 1
  - a. Use Lag plots to determine P and Q values.
  - b. Result: Inconclusive on lag plots.
    - i. ACF for P, PACF for Q.
    - ii. Random Walk model.
      1. Hyndman
    - iii. ARIMA(0,1,0)

# Baseline Results:

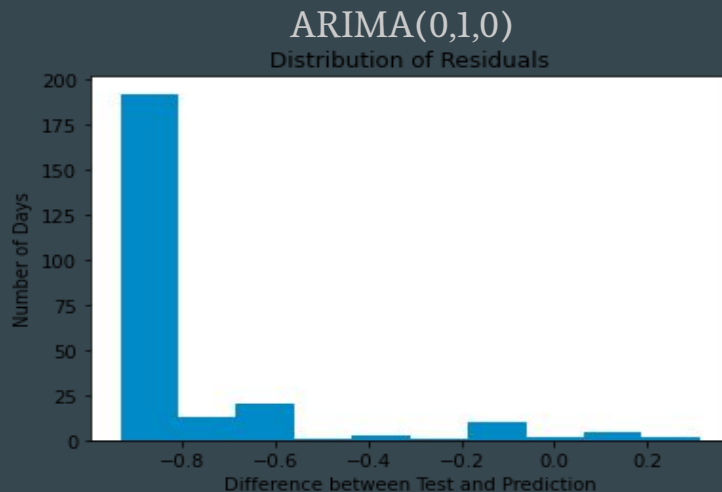
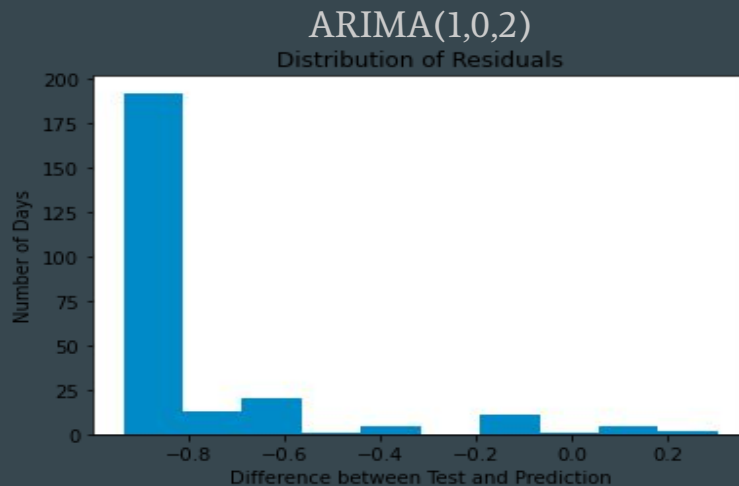
Model 1: ARIMA(1,1,2):

MAPE = 6.284

Model 2: ARIMA(0,1,0)

MAPE = 6.267

Conclusions: As expected, both models had a hard time determining future price action.



# Modeling:

## Time Frames

- Focus on the first peak, and the middle 'valley'
- For each model: Full TF, Peak TF, Valley TF

## MAPE and Distribution of Residuals

- Use MAPE as a performance metric
- Plot Distribution of Residuals

## PMD ARIMA

- Use PMD ARIMA and compare P, D, Q values with Baseline model

## FB Prophet

- Use Facebook Prophet on outlined Time Frames



# PMD ARIMA

## Steps:

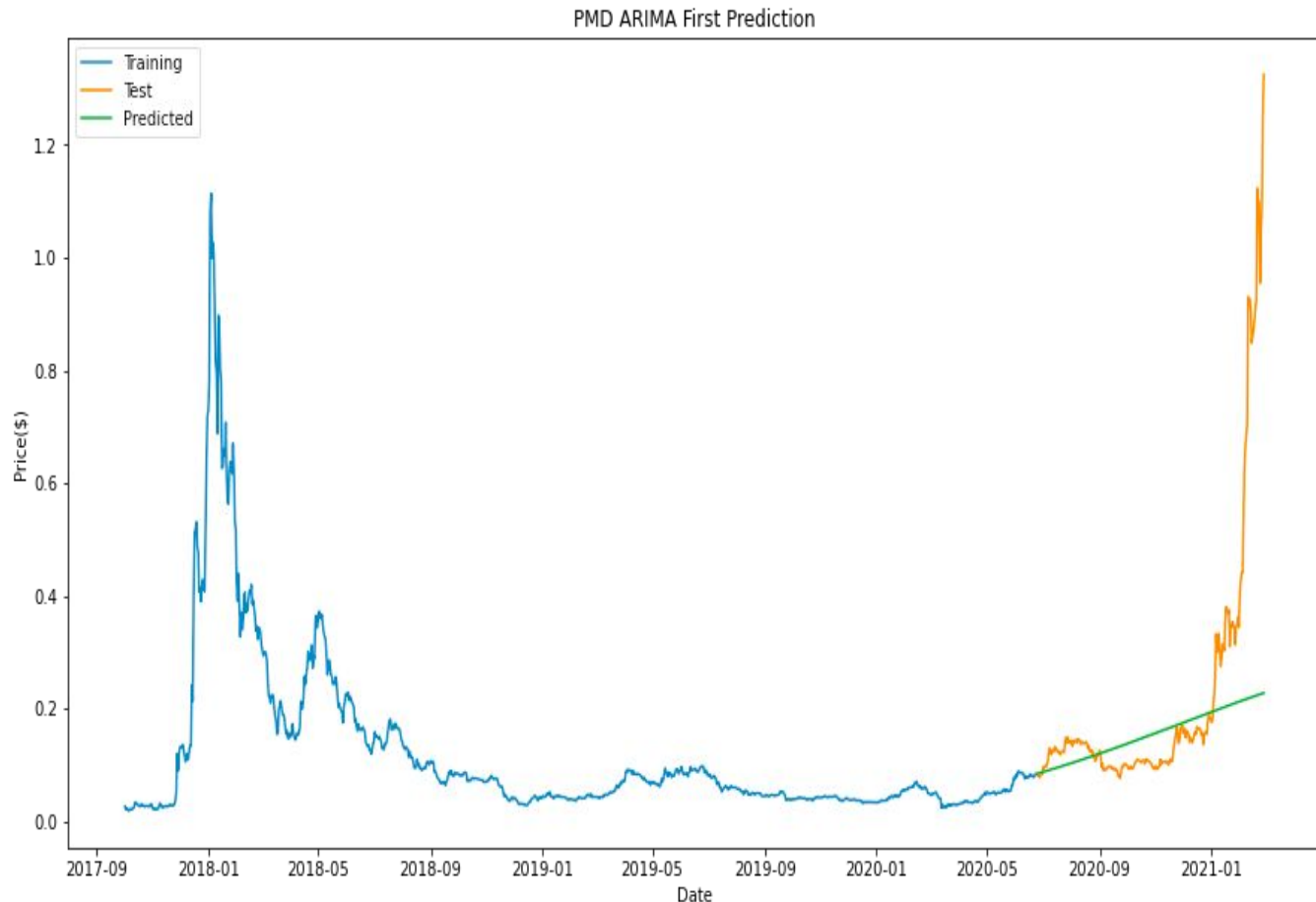
-Import PMD Arima: Start  $p, q = 0$  and  $d=2$ , max  $p,d,q = 5$ .

-Results: Best Model --  
ARIMA(0,2,4)

-Interesting that the  $d$  value was set to 2

Full TF Results:

MAPE = .34



# PMD ARIMA: Valley

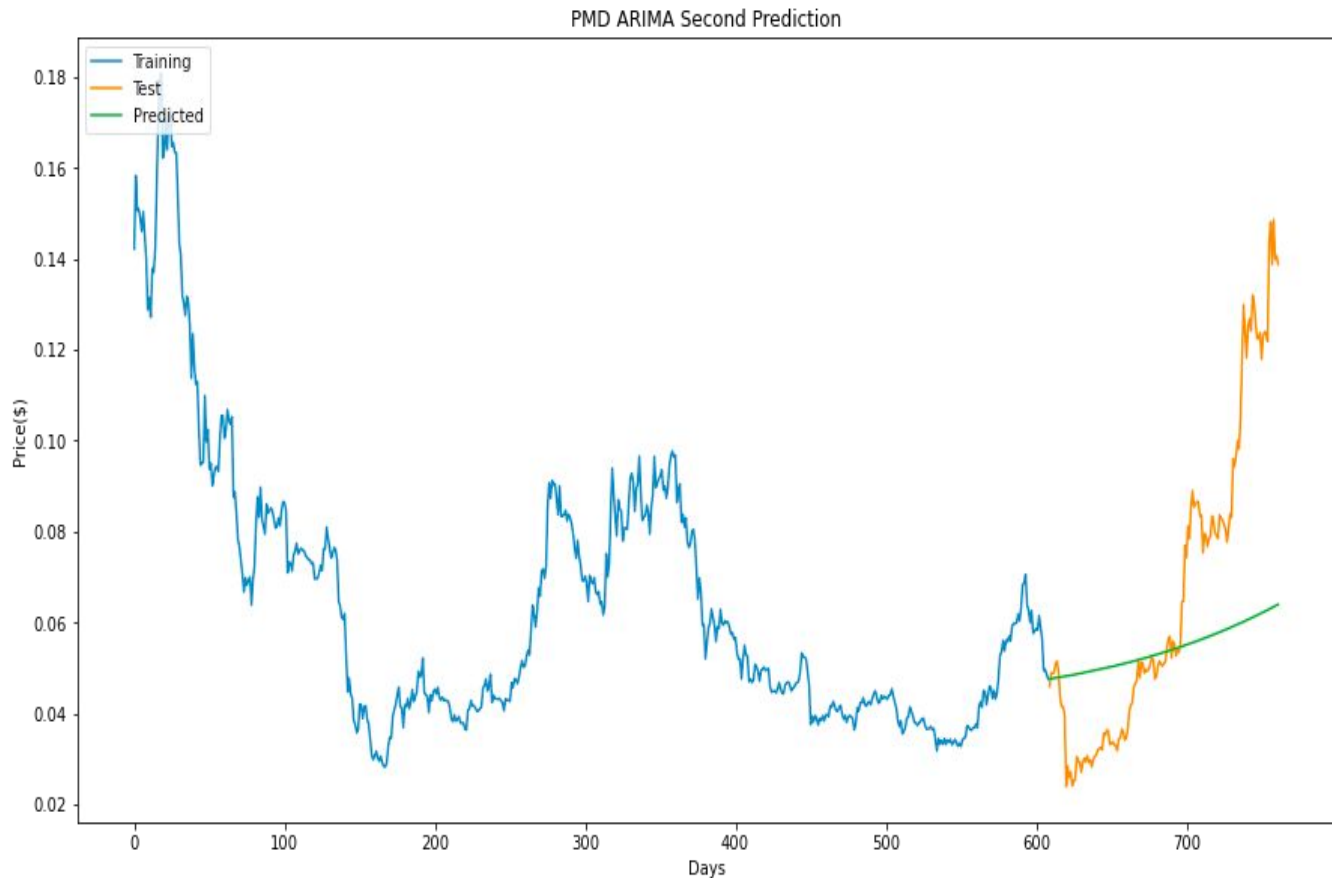
ARIMA(2,2,3)

Valley Time Frame:  
6/30/2018 - 7/21/2020

Does the prediction fit the  
actual values any better?

Valley TF Results:

MAPE = .352



# PMD ARIMA: Peak TF

ARIMA(3,2,1)

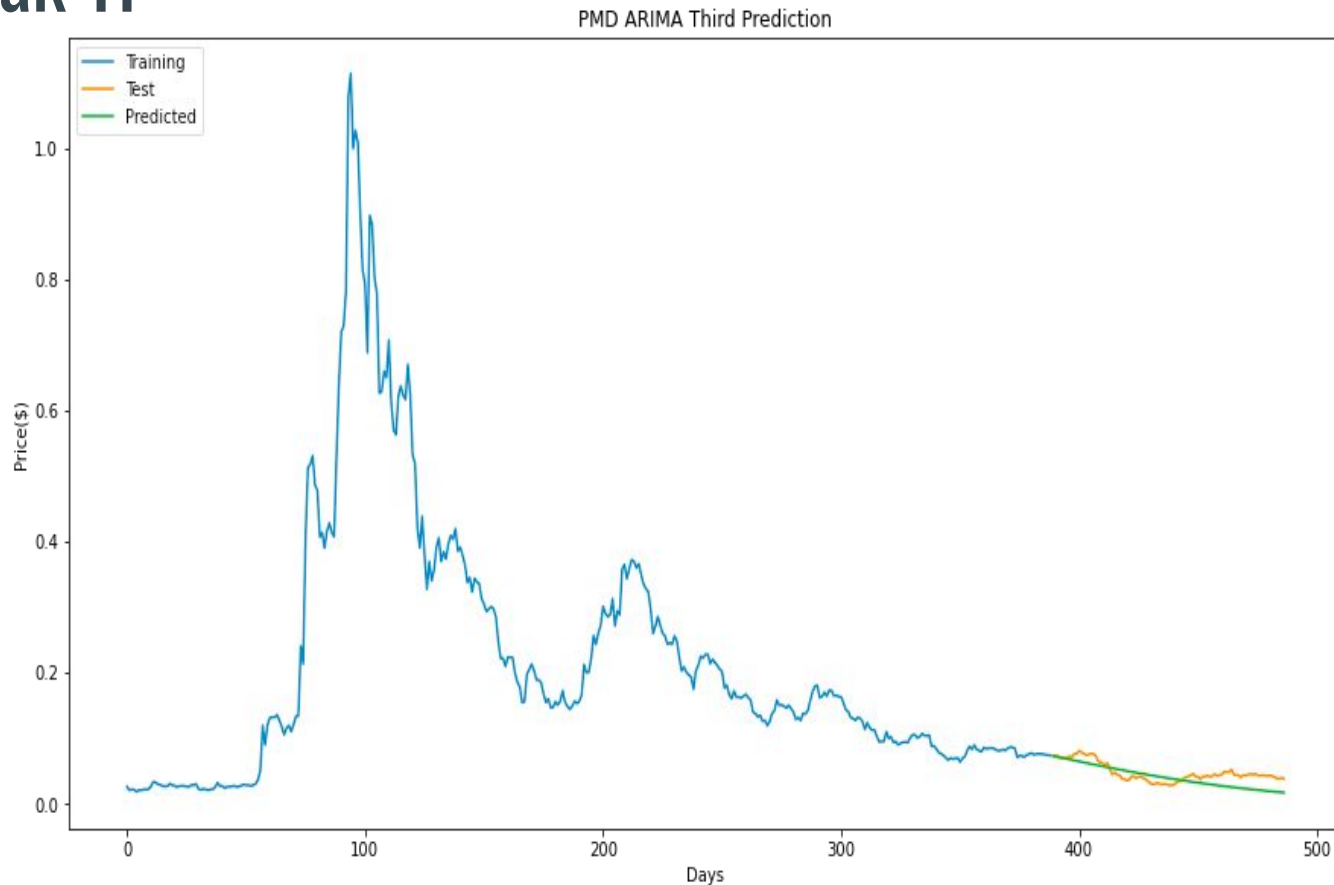
Peak Time Frame:

10/01/2017 - 1/31/2019

Does the prediction fit the  
actual values any better?

Peak TF Results:

MAPE = .283



# FB Prophet: Full TF

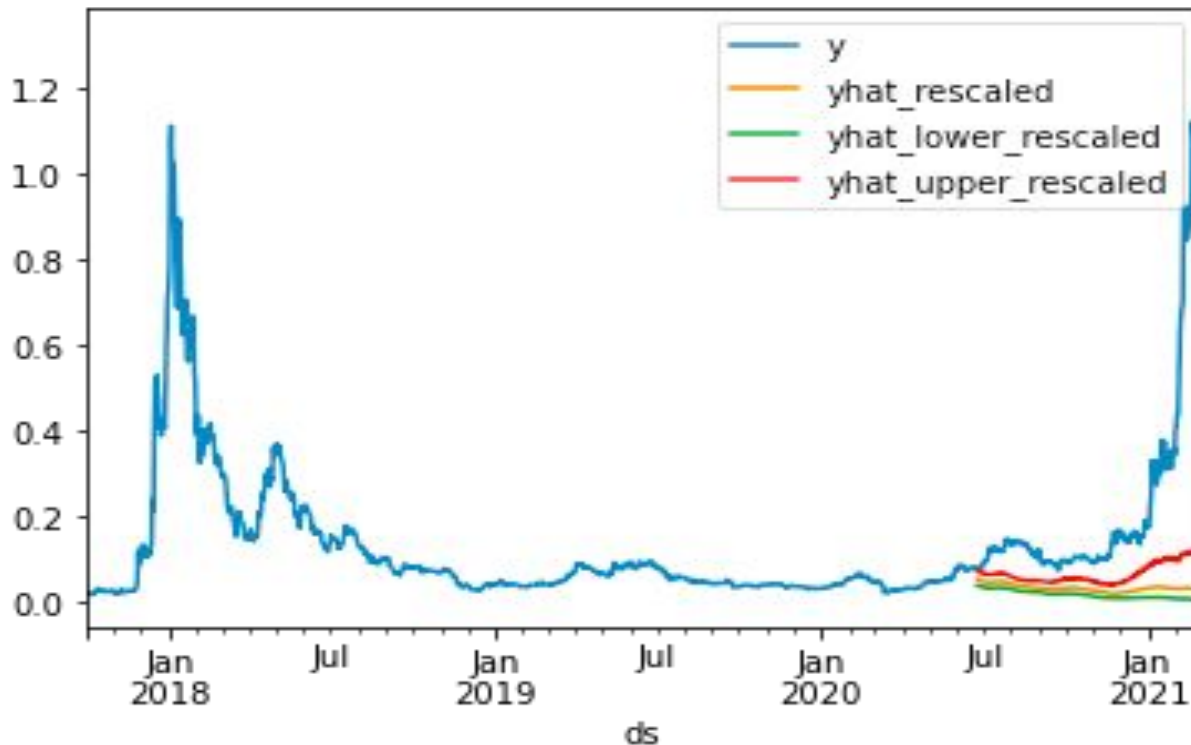
Really easy to use  
compared to other models

Does the prediction fit the  
actual values any better?

Full TF Results:

MAPE = .74

PMD ARIMA fit to the full  
time frame better

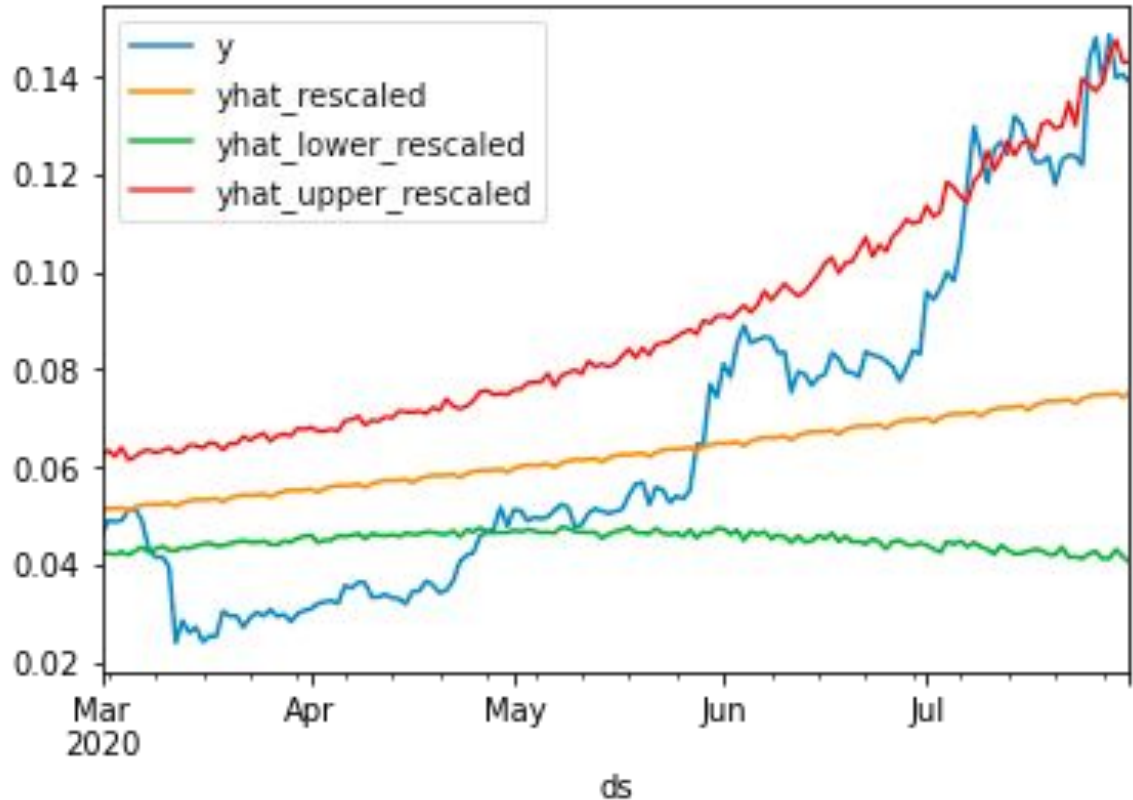


# FB Prophet: Valley TF

Does the prediction fit the actual values any better?

Valley TF Results:

MAPE = .39

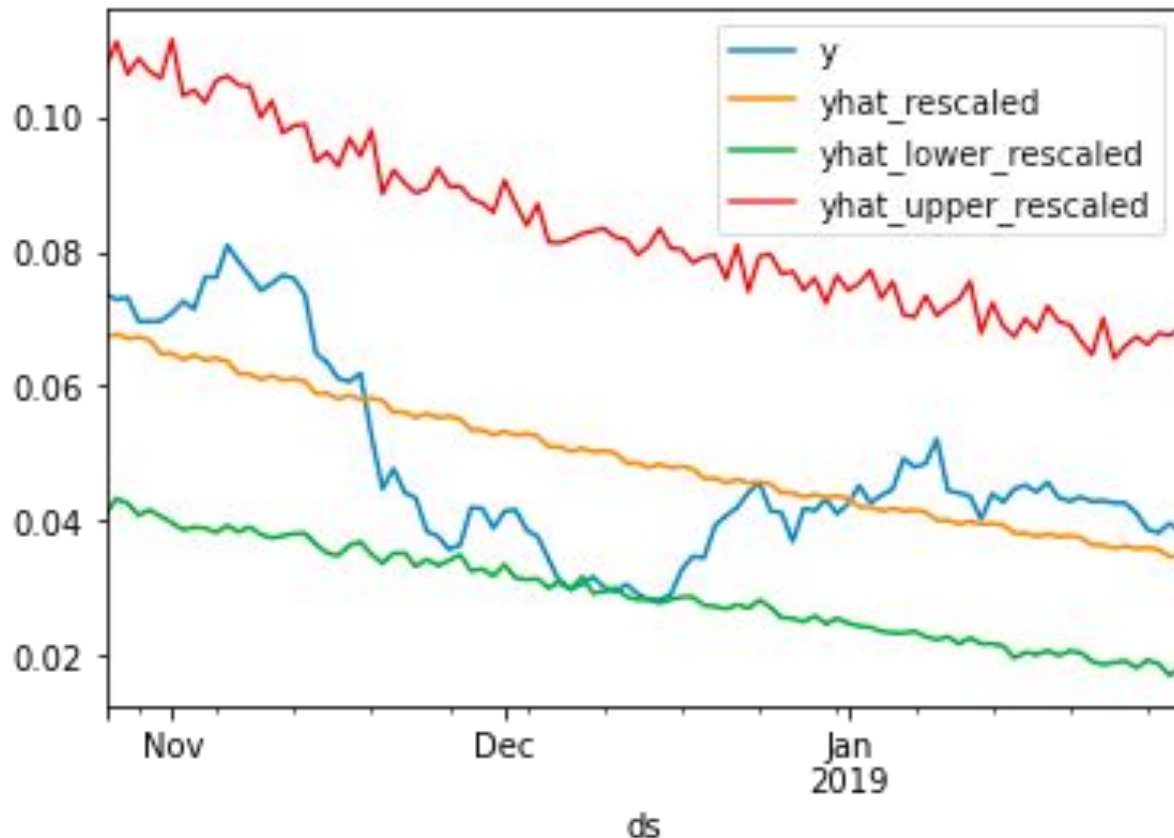


# FB Prophet: Peak TF

Does the prediction fit the  
actual values any better?

Peak TF Results:

MAPE = .22



# Modeling Results:

Peak Time Frame produced best results.

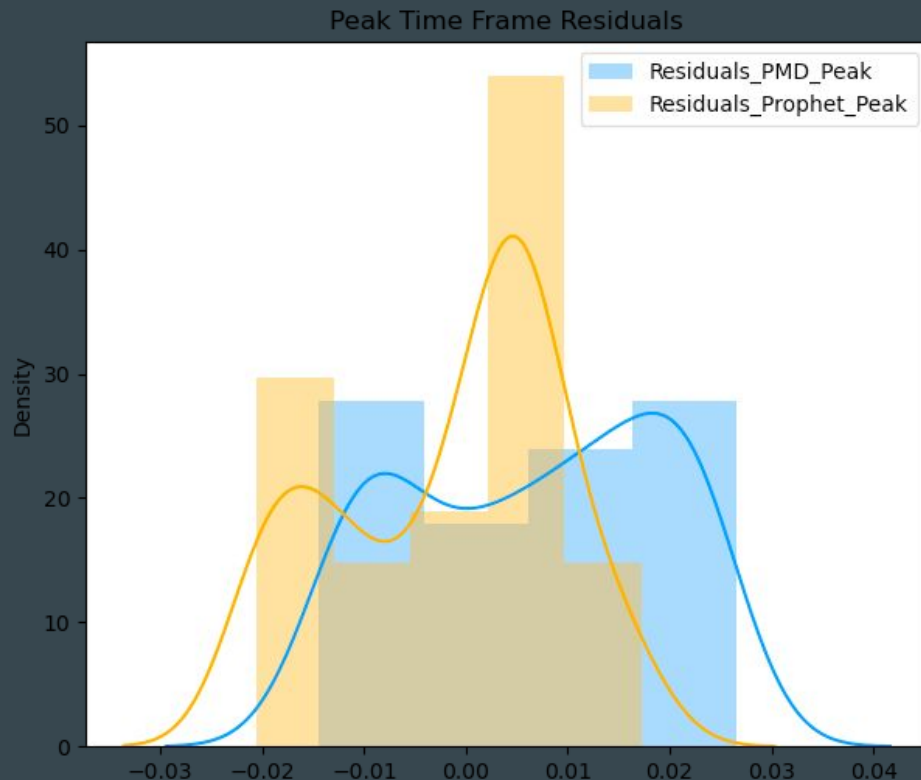
The only distribution of the Residuals that was close to a Normal Distribution were the Peak TFs.

Focusing on a smaller time frame yielded better results than attempting to train the entirety of this data set.

MAPE:

PROPHET\_PEAK\_TF = .22

PMD\_PEAK\_TF = .28



# Conclusion and Recommendations:

## Smaller Time Frames

It is clear from the results, that the clearer the trend and the smaller the time frame, the better the performance of the model.

It is also clear that even on a smaller time frame, if the variance in the data is too great, the model will not fit predicted values well.

## Transaction Volume

It would be worthwhile to observe the relationship between price action and Transaction volume on a given day.

Introducing a multivariate time series forecasting approach would be interesting.

## BTC and ETH

When investing in any cryptocurrency it is imperative to keep tabs on the price action of BTC and ETH.

It is possible to predict well-defined trends with FB Prophet and PMD ARIMA and make informed guesses based on BTC and ETH movements on when trends may reverse.