Gold ETFs Price Forecast, Trends, & 2 Year Predictions

Deepanshu sharma

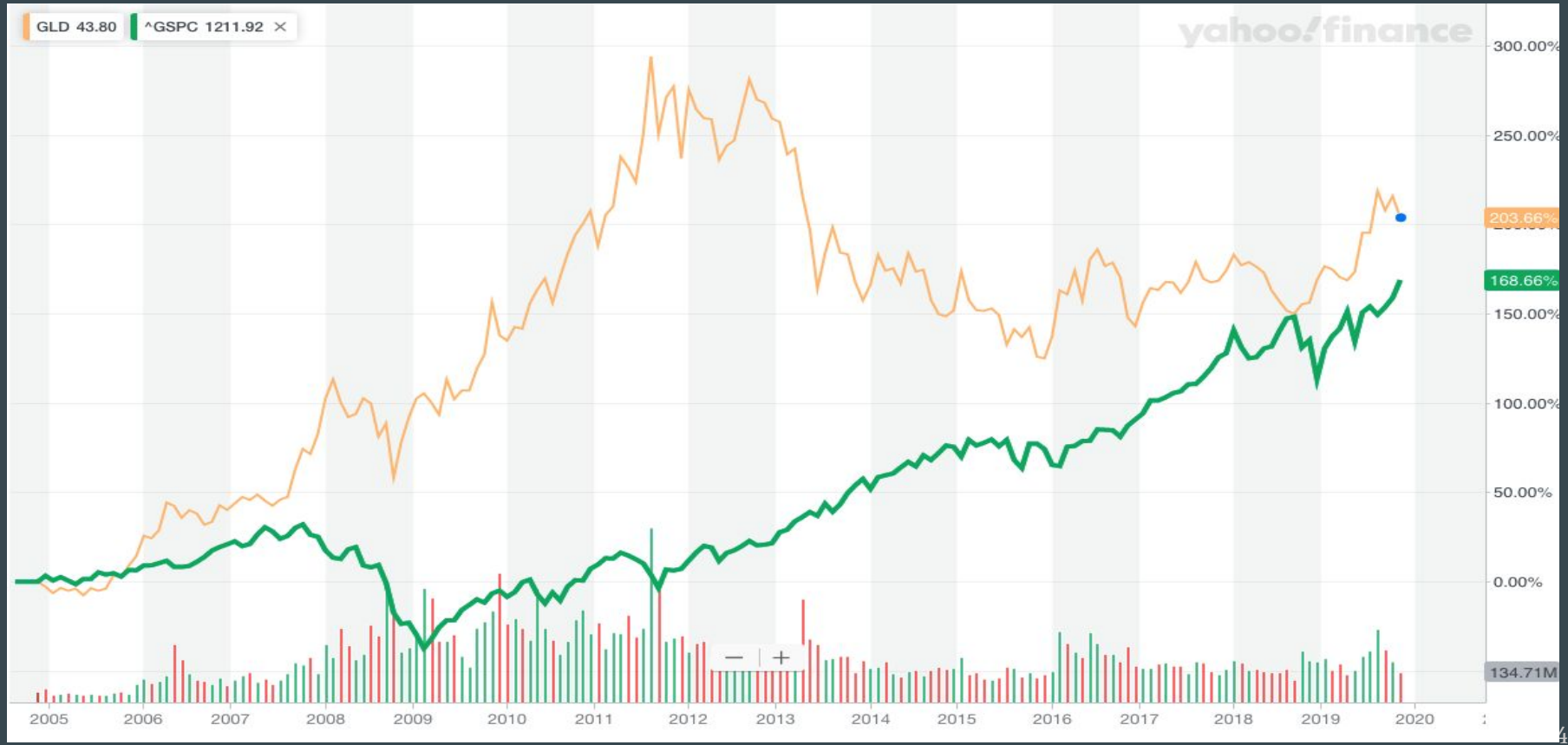
SPDR Gold Shares (GLD)

This analysis is on SPDR® Gold Shares (NYSE Arca : GLD) which is an exchange traded fund.

GLD is a cost-effective and convenient way to invest in gold without buying the real gold

Is Gold a Good Investment In General?

Performance of GLD vs SPY from 2004 to 2019 (15 Years)



Performance of GLD vs SPY - 15 Years Performance

GLD started trading on 11/18/2004. Let's compare the performance of GLD against the S&P500 index (SPY) for the last 15 years. If you've invested $10,000 in both, here are the results:

	GLD	SPY
Gain/loss since 11/18/2004	204%	168%
What will the $10K investment be worth today?	$30,400	$26,800
Compounded annual growth rate	7.69%	6.79%

GLD seems to have a better rate of return if you've invested in it 15 years ago.

Performance of GLD vs SPY from 2015 to 2019 (5 Years)

Performance of GLD vs SPY - Last 5 Years Performance

Let's compare the performance of GLD against the S&P500 index (SPY) for the last 5 years. If you've invested $10,000 in both, here are the results:

	GLD	SPY
Gain/loss since 1/1/2015	20%	52%
What will the $10K investment be worth today?	$12,000	$15,200
Compounded annual growth rate	3.71%	8.73%

GLD vs SPY - 10 Year Risk Analysis

	GLD	SPY
Standard Deviation - the price of something varies from its average over a given period of time (lower the better)	16.59	12.48
Sharpe Ratio- the average return earned in excess of the risk-free rate per unit of volatility or total risk (higher the better)	0.25	1.04

Source:
https://finance.yahoo.com/quote/GLD/risk?p=GLD
https://finance.yahoo.com/quote/SPY/risk?p=SPY&.tsrc=fin-srch

Gold Is A Volatile Investment

- The performance analysis has shown a big swing in gains depending on when you've invested in GLD. Investing 15 years ago, would have a CAGR of 7.69%, whereas investing 5 years ago, the CAGR would have dropped to 3.71%, that's a 51% drop!

- The Sharpe Ratio which measures the average return earned in excess of the risk-free rate is also much lower than SPY, which indicates it's a riskier investment than SPY.

Gold is Risky, But....

Why Do People Still Buy Gold?

- Diversification - Gold and stocks don't always do the same thing at the same time, e.g when stock market is doing well, gold often lags behind. Correlation between Gold and the stock market over the last 10 years is 0.04. They are not correlated.

- Protection - Many investors buy gold as a safe haven to protect themselves against a possible catastrophe and against inflation.

- Recession-Proof Portfolio - With recent fears of an economic recession looming in the distance, investors are looking to recession-proof their portfolios

GLD - To Buy or Not To Buy?

Forecast GLD Prices For The Next 2 Years

Where Is GLD Data From?

The historical prices of SPDR® Gold Shares (NYSE Arca : GLD) was downloaded from Yahoo. Data spans from the inception of this share from 11/18/2004 to the date of download, 11/22/2019.

Source:
https://finance.yahoo.com/quote/GLD/history?p=GLD

Exploratory Data Analysis

There are 3780 rows of data from 11/18/2004 to 11/22/2019. However, there are missing data. That's because when the stock market is closed on weekends and holidays, there's no trading activities.

Missing 2004-11-20
2004-11-21

	Date	Open	High	Low	Close	Adj Close	Volume
0	2004-11-18	44.430000	44.490002	44.070000	44.380001	44.380001	5992000
1	2004-11-19	44.490002	44.919998	44.470001	44.779999	44.779999	11655300
2	2004-11-22	44.750000	44.970001	44.740002	44.950001	44.950001	11996000
3	2004-11-23	44.880001	44.919998	44.720001	44.750000	44.750000	3169200
4	2004-11-24	44.930000	45.049999	44.790001	45.049999	45.049999	6105100
5	2004-11-26	45.250000	45.599998	45.060001	45.290001	45.290001	3097700
6	2004-11-29	45.099998	45.500000	45.080002	45.400002	45.400002	3759000
7	2004-11-30	45.369999	45.410000	44.820000	45.119999	45.119999	3857200
8	2004-12-01	45.279999	45.590000	45.259998	45.380001	45.380001	2037500

Exploratory Data Analysis

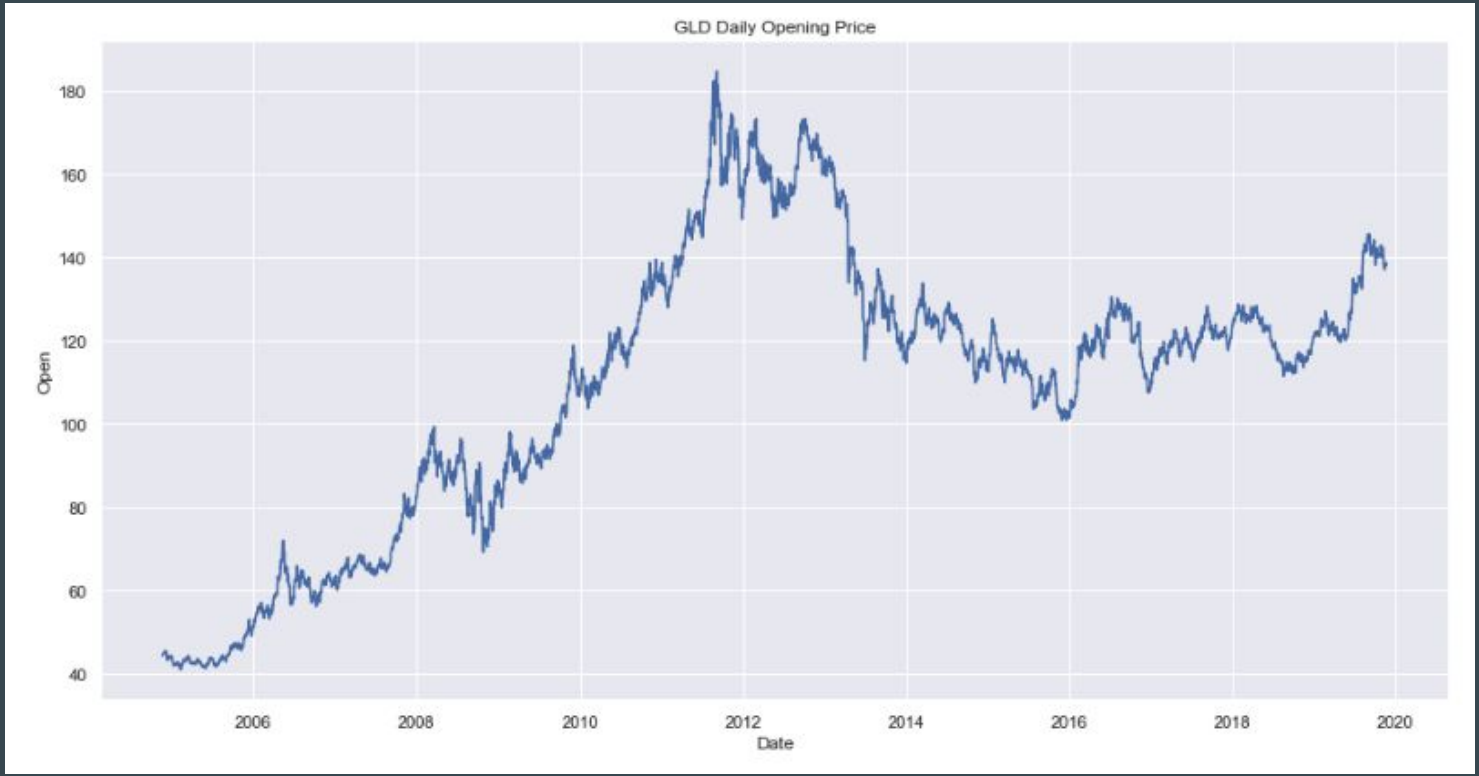
Perform a re-sample by day and do a forward fill to fill in the missing values from previous day.

these 2 missing dates are added to the dataframe

	Date	Open	High	Low	Close	Adj Close	Volume
0	2004-11-18	44.430000	44.490002	44.070000	44.380001	44.380001	5992000
1	2004-11-19	44.490002	44.919998	44.470001	44.779999	44.779999	11655300
2	2004-11-20	44.490002	44.919998	44.470001	44.779999	44.779999	11655300
3	2004-11-21	44.490002	44.919998	44.470001	44.779999	44.779999	11655300
4	2004-11-22	44.750000	44.970001	44.740002	44.950001	44.950001	11996000
5	2004-11-23	44.880001	44.919998	44.720001	44.750000	44.750000	3169200

After this transformation, the dataframe now has 5483 rows of data.

Plot of GLD Daily Opening Price



Make Time-Series Data Stationary

Perform Dicky-Fuller Test to check if the data is stationary.

Dicky-Fuller Results

- ADF Statistic: -1.80251749817130 52
- p-value: 0.3792125439071432

Based on the results above, the data is not stationary because the p-value is greater than 0.05.

Make The Time Series Data Stationary

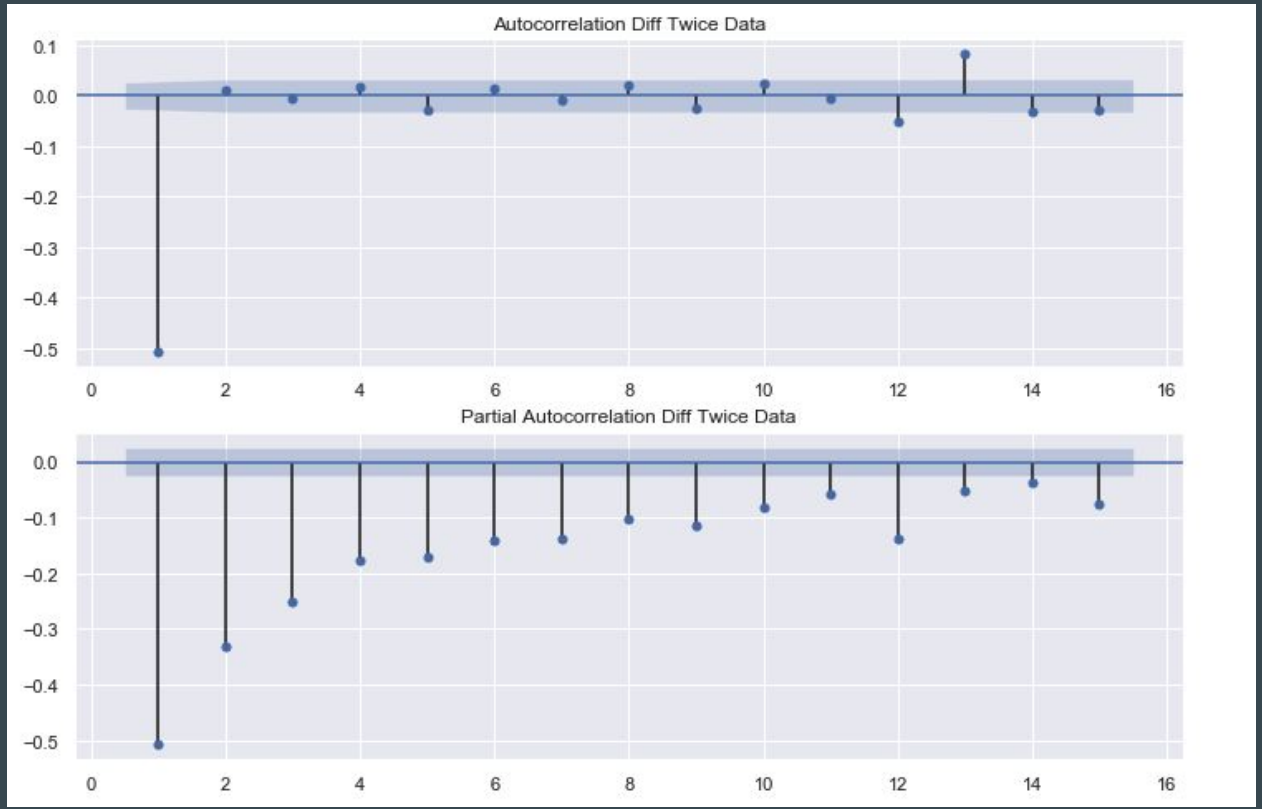
Use the difference once method, difference twice and square root method and pick the best method.

	ADF Statistic	P-Value
Difference Once	-16.738250093306753	1.3650617859601889e-29
Difference Twice	-21.13672510562556	0.0
Square Root	-2.0279705213801678	0.27444523490828154

Both Differencing once and twice methods produced a p-value less than 0.05 but Differencing Twice produced a much more negative ADF Statistic. That's what we want, the more negative the better.

How To Choose The Right Model?

ACF and PACF



The initial observation shows that the ACF cuts off after lag 1 and PACF tails off over time. This may indicate that it's a Moving Average model with an order of 1, that's MA(1)

AIC and BIC

Results Sorted by AIC

	p	q	aic	bic
1	0	1	15870.084038	15883.302124
2	0	2	15871.949100	15891.776229
4	1	1	15871.977921	15891.805050
7	2	1	15873.328543	15899.764715
5	1	2	15873.984951	15900.421122
8	2	2	15875.843240	15908.888454
6	2	0	17419.185255	17439.012384
3	1	0	18053.677637	18066.895723
0	0	0	19683.644070	19690.253113

Results Sorted by BIC

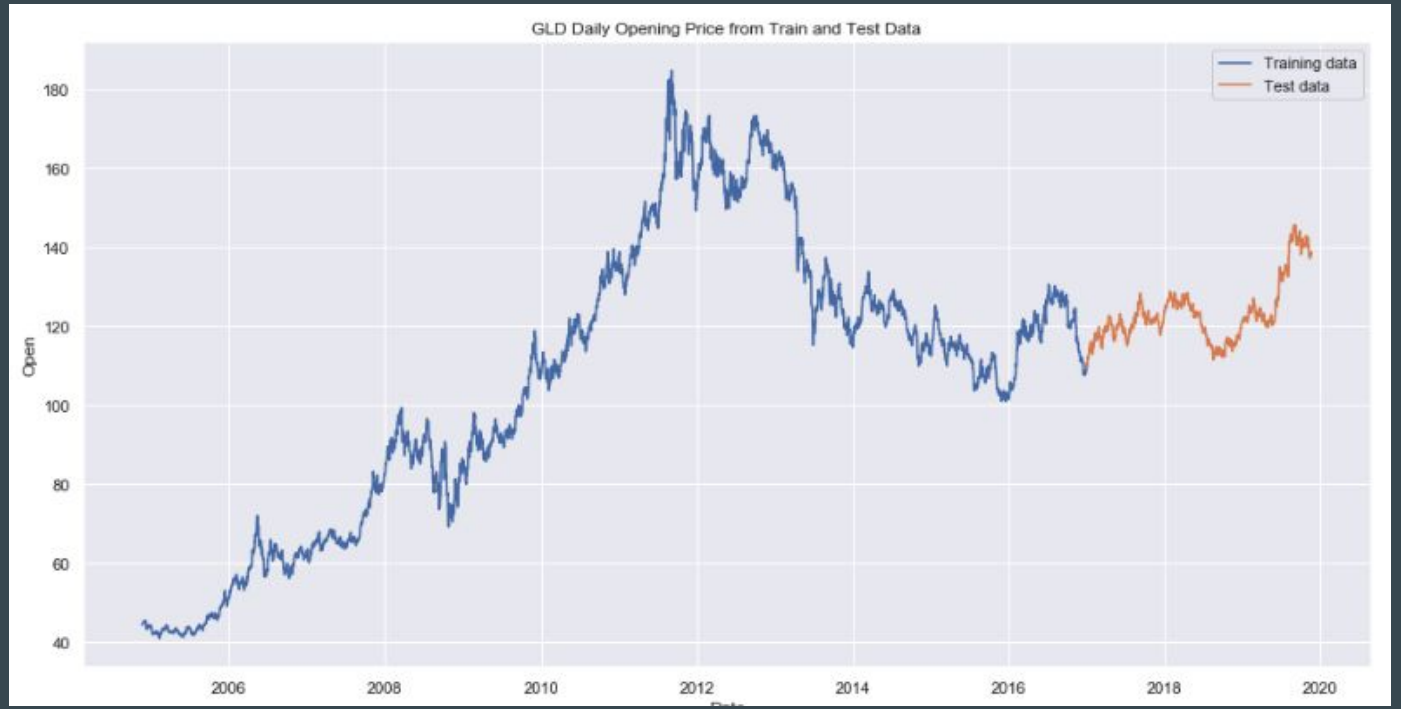
	p	q	aic	bic
1	0	1	15870.084038	15883.302124
2	0	2	15871.949100	15891.776229
4	1	1	15871.977921	15891.805050
7	2	1	15873.328543	15899.764715
5	1	2	15873.984951	15900.421122
8	2	2	15875.843240	15908.888454
6	2	0	17419.185255	17439.012384
3	1	0	18053.677637	18066.895723
0	0	0	19683.644070	19690.253113

We want to select the results with the lowest AIC and BIC. Both AIC and BIC selected the same model order in this case. This is the same results as the ACF and PACF analysis where we determined it was a MA(1) model order.

Train-Test Data Split

Splitting Data Into Training and Test Sets

There are 15 years of data. We are going to use the first 12 years (2004 - 2016) as training data and the last 3 years (2017 - 2019) as test data.



Modeling Using ARIMA

Modeling Method 1

ARIMA Model with No Seasonality & One-Step Ahead Forecast

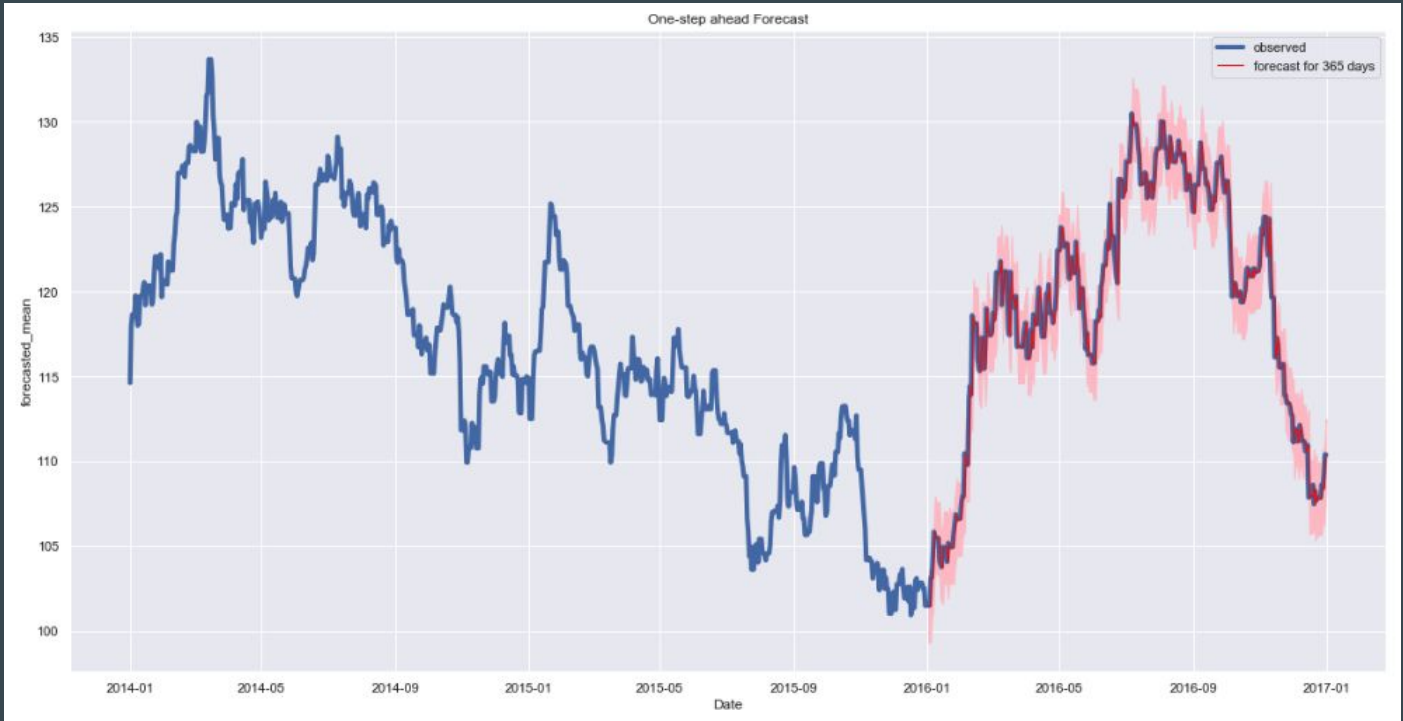
ARIMA Model with No Seasonality & One-Step Ahead Forecast

We use SARIMAX model here as it accounts for seasonality parameters if seasonalty exists.

```
model = SARIMAX(df, order=(p,d,q), seasonal_order=(P,D,Q,S))
```

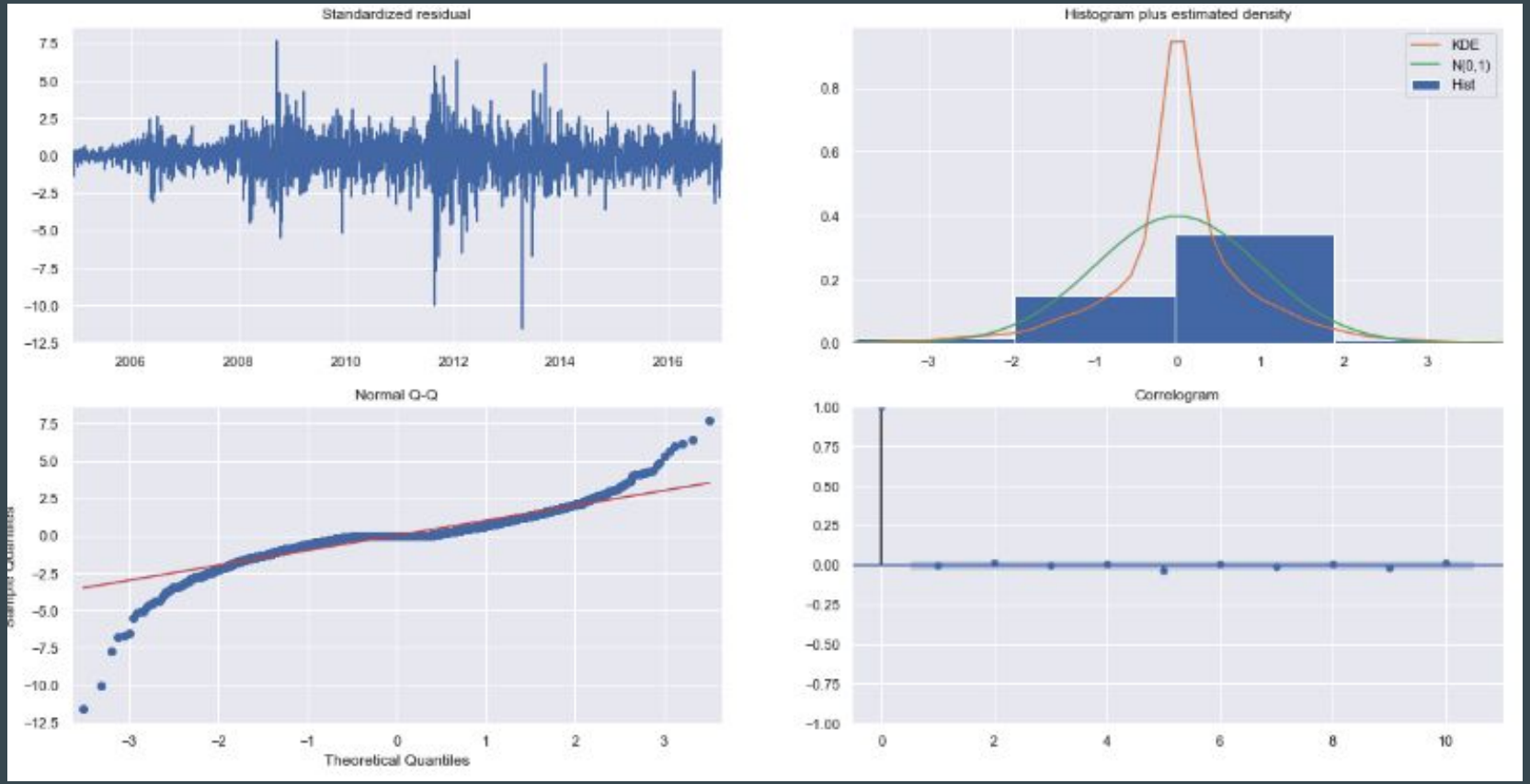
```
model = SARIMAX(train_data['Open'], order=(0,2,1), trend= 'c')
```

Plot Predicted Price Against Last 365 Days of Train Data



The above forecasts aligns very well with the true values for the last 365 days of the training data and it falls within the confidence intervals

Model Diagnostics Results



Model Diagnostics Results

Standardized Residual Plot - The graph doesn't seem to show a trend. That's what we want.

Histogram Plus Estimated Density - This shows the distribution of the residuals. The green line shows a normal distribution and the orange line needs to be as close to the green line. The 2 lines are very different in this case. This model might need tweaking.

Normal Q-Q - This shows how the distribution of the residuals compared to a normal distribution. Most of the residuals are on the line except the ends.

Correlogram - ACF plot of the residuals. 95% of the data where lag > 0 should not be significant. That means, they need to be within the blue shaded area. Based on the graph, it looks OK as 95% of the data is not significant, they are within the blue shaded area.

Model Diagnostics Results

	ARIMA Model with No Seasonality & One-Step Ahead Forecast
Mean Absolute Error (MAE)	0.64
Mean Squared Error (MSE)	1.04
Root Mean Squared Error (RMSE)	1.02

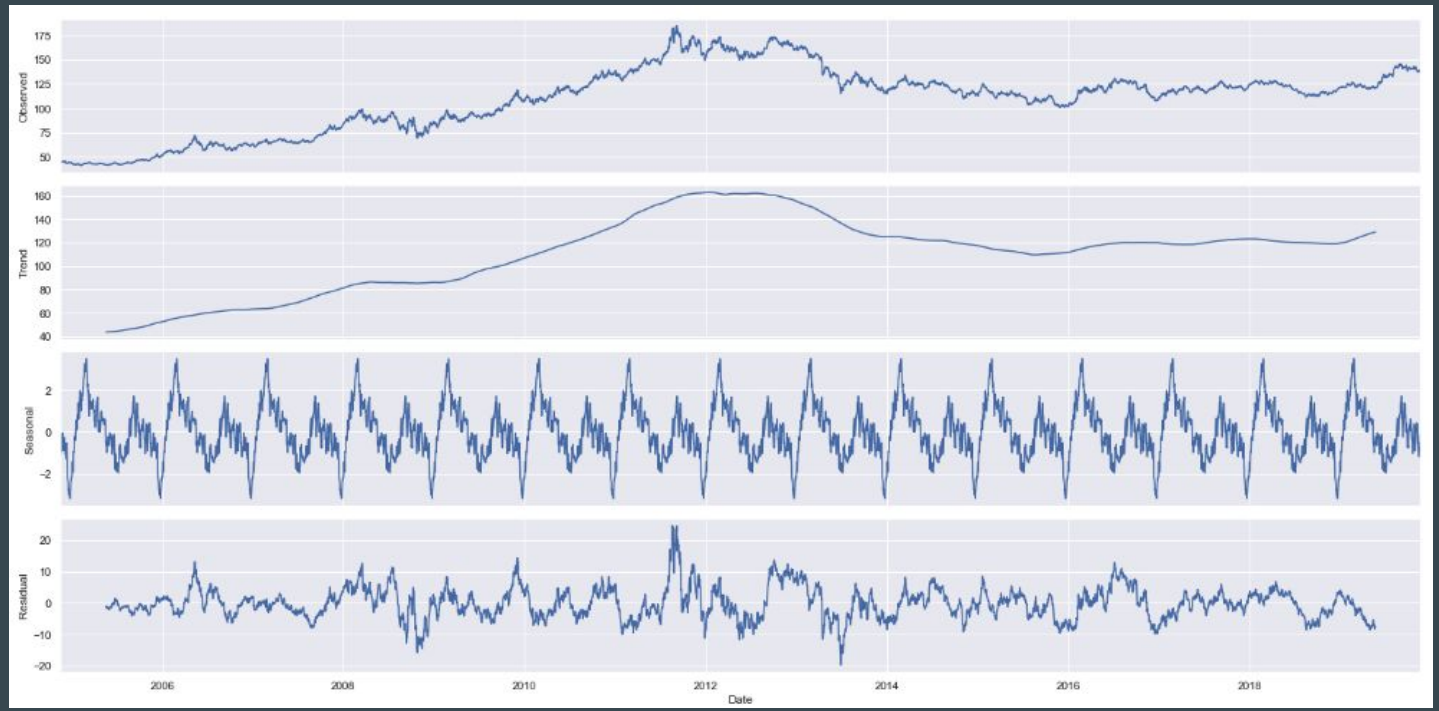
This model forecasted the average daily open price in the training set is within $1.02 of the real open prices.

Modeling Method 2

Auto ARIMA Model with Seasonality & One-Step Ahead Forecast

Seasonal Decompose

Use seasonal_decompose to check for any seasonality



The plots above shows that the trend in prices of gold is not consistent but there is some obvious seasonality

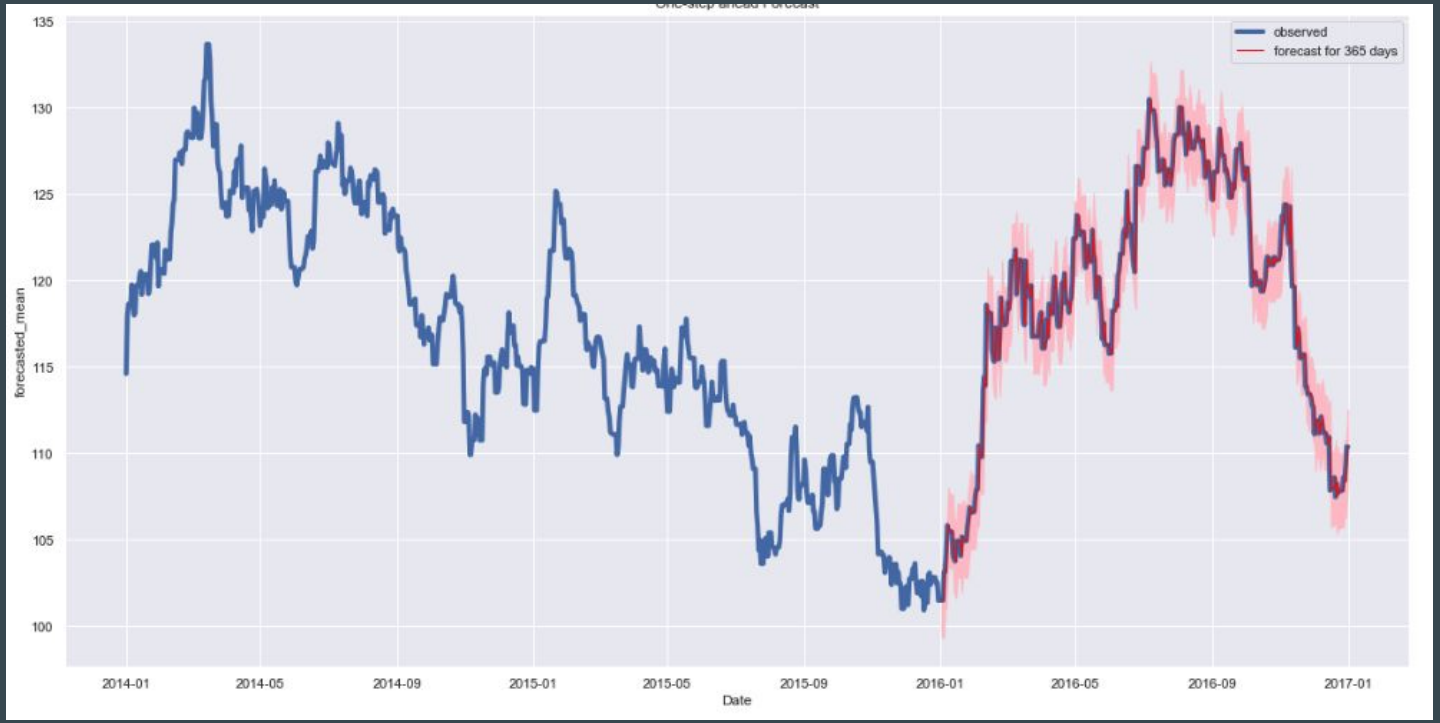
Auto Arima

Auto Arima is like an automatic grid search.

Based on the results generated by auto_arima that produced the lowest AIC score, the Best Fit ARIMA is: order=(0, 1, 0) seasonal_order=(0, 0, 0, 7)

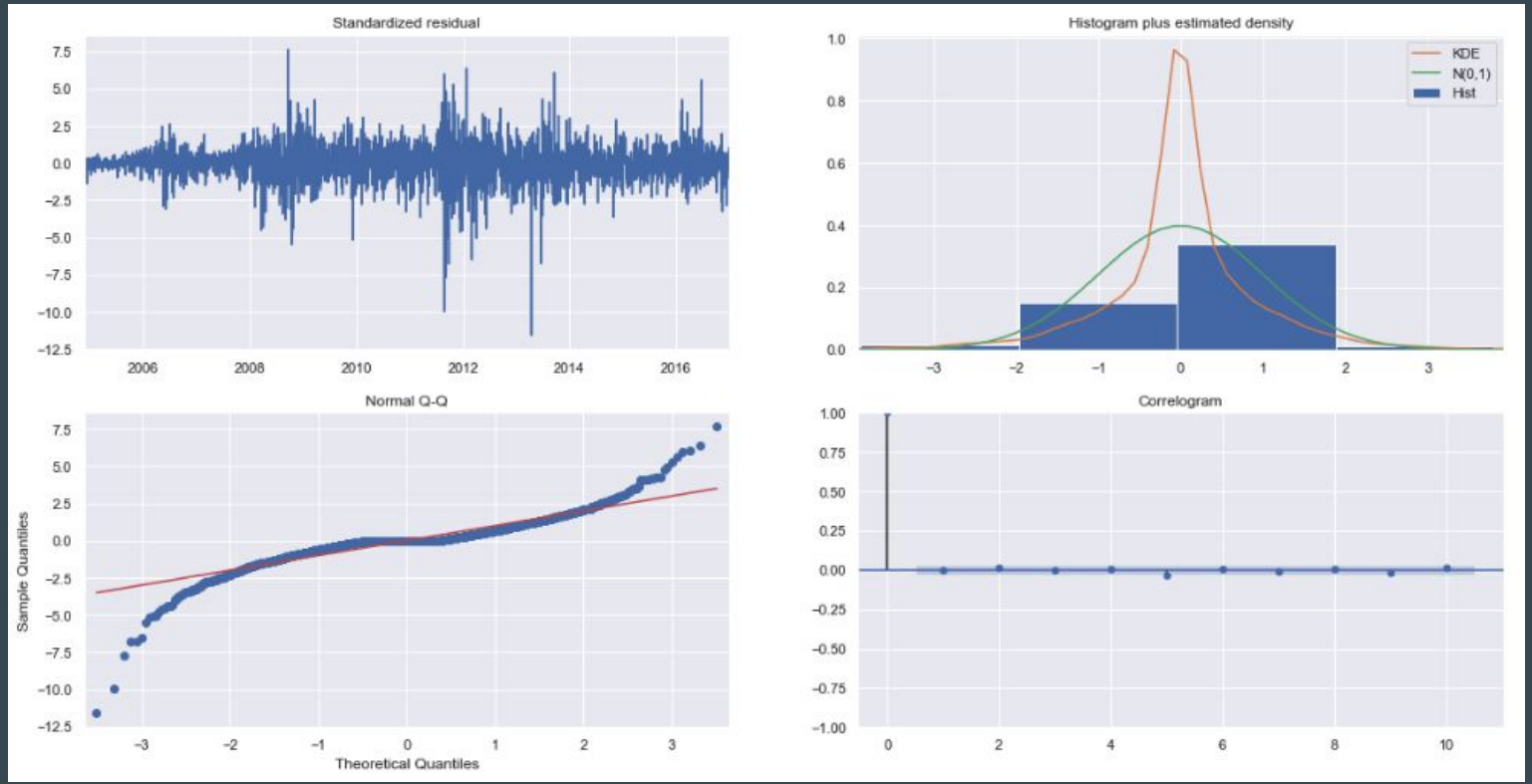
Auto Arima didn't detect any seasonality and suggested the differencing is ONLY ONCE and not TWICE which we concluded earlier using the Dicky-Fuller test.

Plot Predicted Price Against Last 365 Days of Train Data



The above forecasts aligns very well with the true values for the last 365 days of the training data and it falls within the confidence intervals

Model Diagnostics Results



These 4 diagnostic plots shows no alarming difference from Method 1 diagnostic plots.

Model Diagnostics Results

	Method 1 - ARIMA Model with No Seasonality & One-Step Ahead Forecast	Method 2 - Auto ARIMA Model with Seasonality & One-Step Ahead Forecast
Mean Absolute Error (MAE)	0.64	0.63
Mean Squared Error (MSE)	1.04	1.04
Root Mean Squared Error (RMSE)	1.02	1.02

Method 2 of MAE performed slightly better than Method 1. However, there's no improvement in MSE and RSME.

Modeling Method 3

ARIMA Model with Seasonality & One-Step Ahead Forecast Using A Manual Grid Search

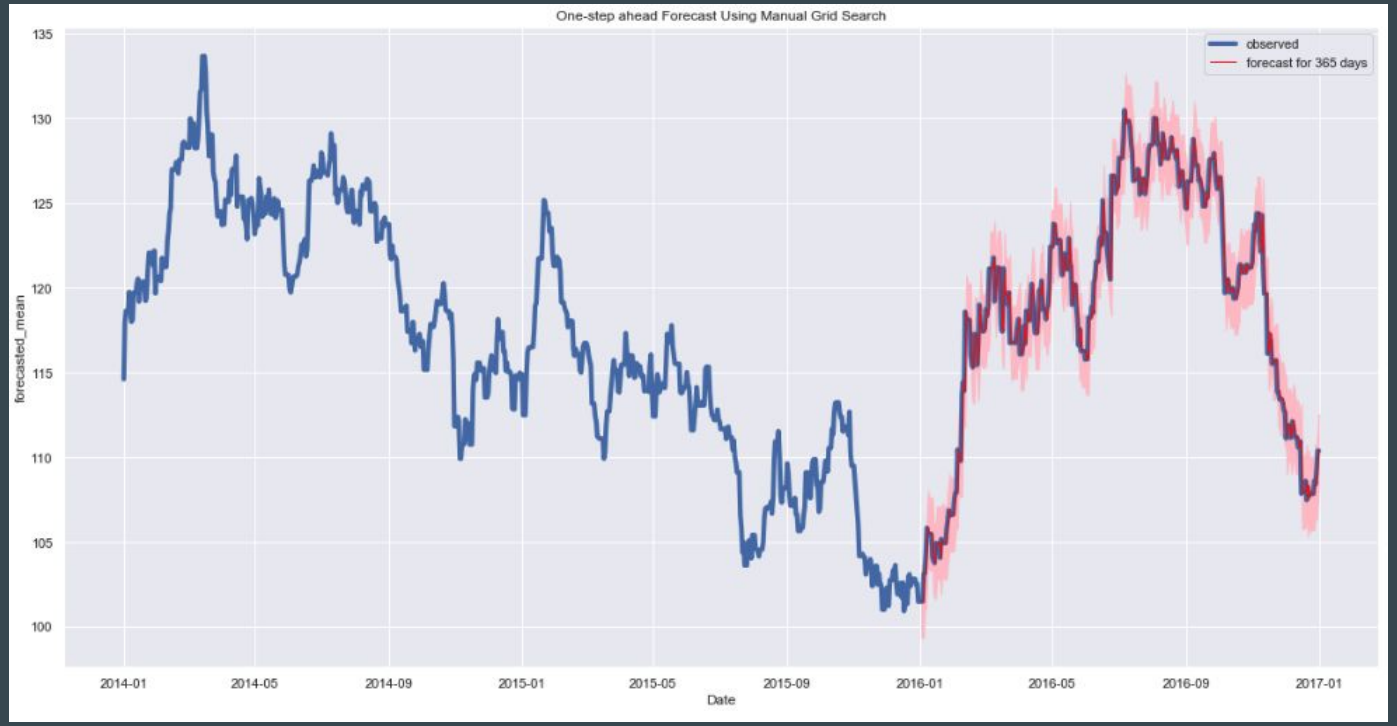
Arima with a Manual Grid Search

Use manual grid search to see if a manual grid search can outperform auto_arima

The optimal parameters found by the manual grid search: ARIMA(2, 1, 2)x(0, 0, 2, 7)

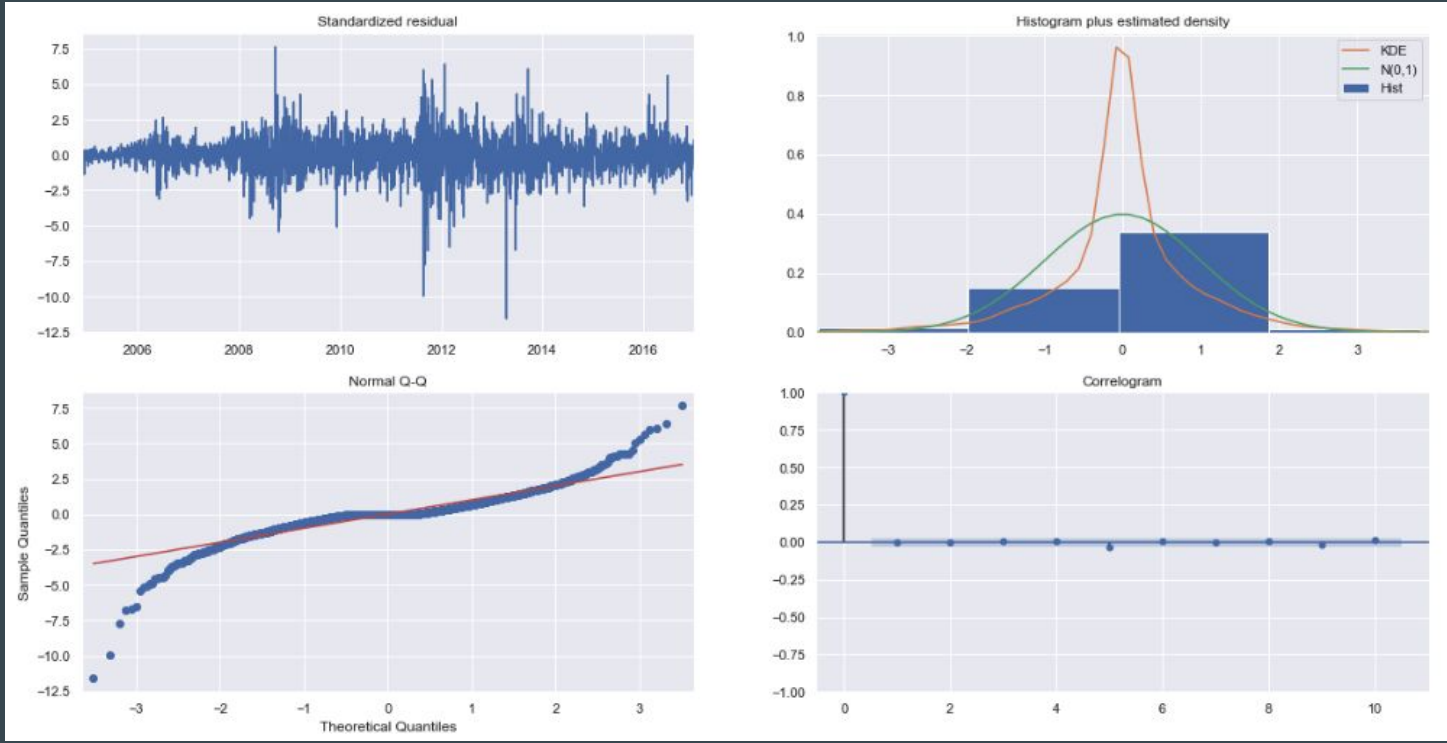
This is different from auto_arima which recommended order=(0, 1, 0) seasonal_order=(0, 0, 0, 7). In this case, the non-seasonal parameters are different and there is seasonality detected.

Plot Predicted Price Against Last 365 Days of Train Data



The above forecasts aligns very well with the true values for the last 365 days of the training data and it falls within the confidence intervals

Model Diagnostics Results



These 4 diagnostic plots shows no alarming difference from Method 1 and 2 diagnostic plots.

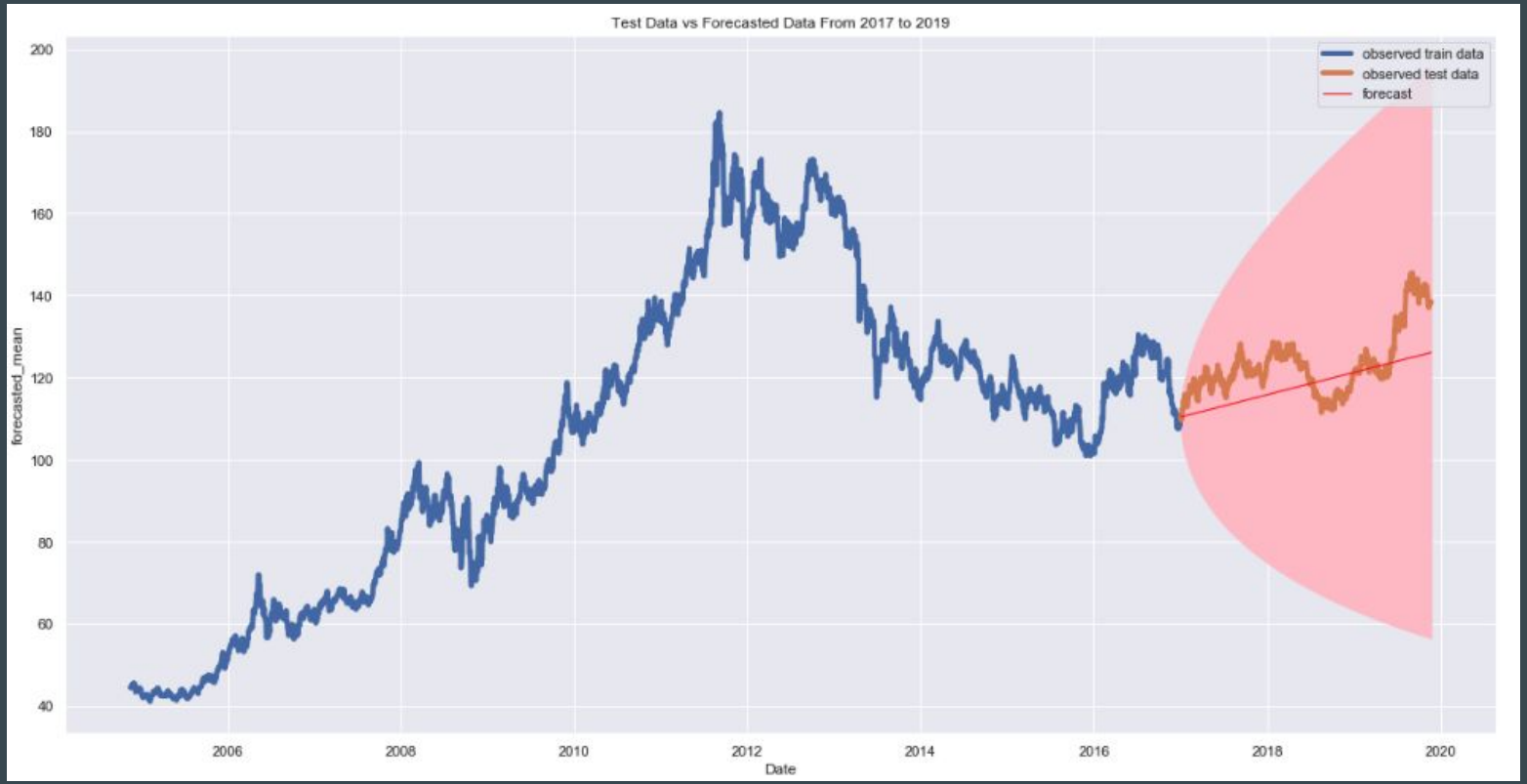
Model Diagnostics Results

	Method 1 - ARIMA Model with No Seasonality & One-Step Ahead Forecast	Method 2 - Auto ARIMA Model with Seasonality & One-Step Ahead Forecast	Method 3 - Auto ARIMA Model with Seasonality & One-Step Ahead Forecast Using Manual Grid Search
Mean Absolute Error (MAE)	0.64	0.63	0.63
Mean Squared Error (MSE)	1.04	1.04	1.04
Root Mean Squared Error (RMSE)	1.02	1.02	1.02

There no difference in the results between Auto Arima and Manual Grid Search methods. MAE for Method 2 and 3 is slightly lower than Method 1. We are going to use the Auto Arima method for forecasting of future data.

Forecasting and Compare Results with Test Data

Use Auto Arima to Forecast Opening Prices and Compare With Test Data



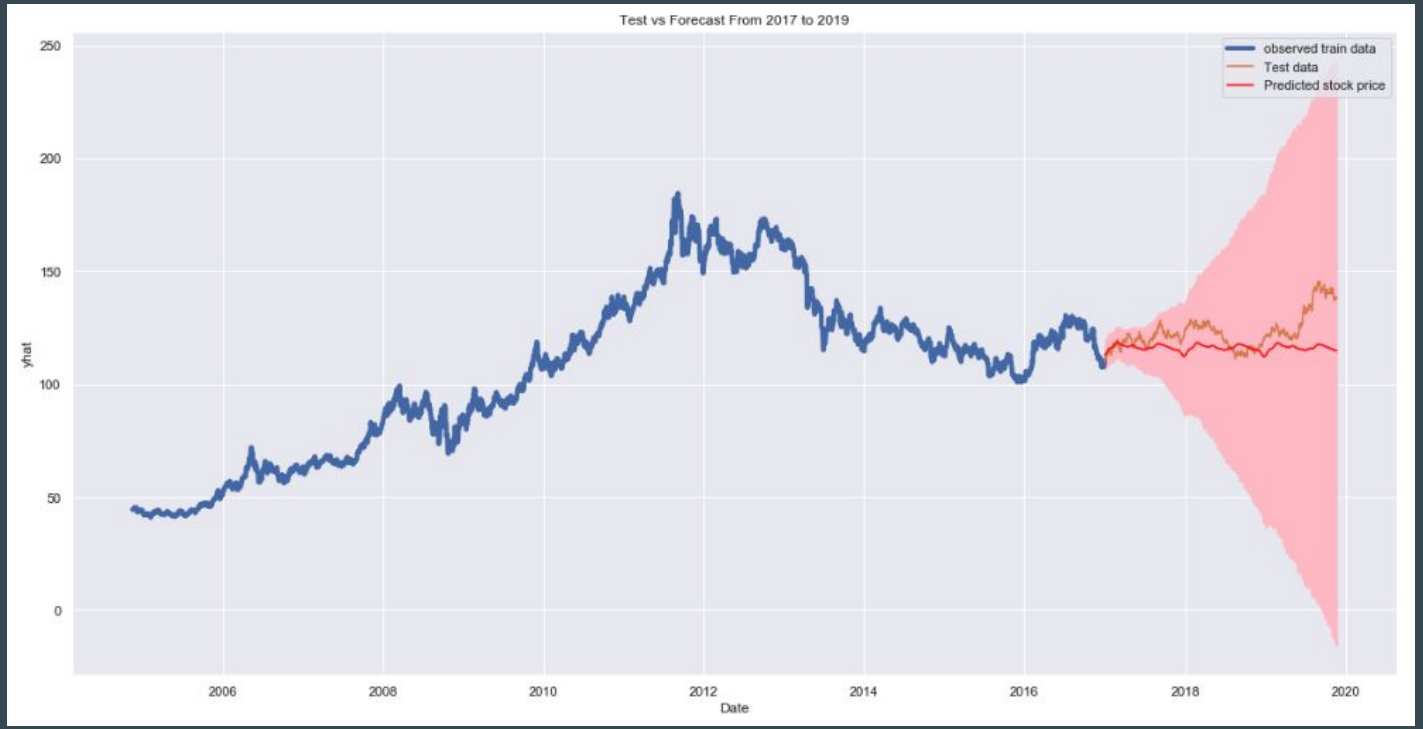
ARIMA Conclusion

- Mean Absolute Error (MAE): 6.70
- Mean Squared Error (MSE): 64.36
- Root Mean Squared Error (RMSE): 8.02

By looking at the plot, the forecasted data showed an upward trend which is aligned with the test data. It correctly predicted that Gold Prices will go up from 2017 - 2019. It is also within the confidence interval. However, the interval is large. This shows that it's hard to predict the prices of gold day-to-day but it's able to predict a general trend over time.

Modeling Using Facebook Prophet

Plot Predicted Price Against Last 365 Days of Train Data



The forecasted data shows a flat line. The forecasted data is not aligned with the test data when the test data shows an upward trend

Facebook Prophet Conclusion

	Auto Arima Forecasting	Facebook Prophet Forecasting
Mean Absolute Error (MAE)	6.70	7.71
Mean Squared Error (MSE)	64.36	108.88
Root Mean Squared Error (RMSE)	8.02	10.43

Prophet does not seem to be as accurate as ARIMA model. The MAE, MSE and RMSE of Prophet are also higher than the results of ARIMA model

Predicting GLD Price For The Next 2 Years

Out-Of-Sample Forecasting

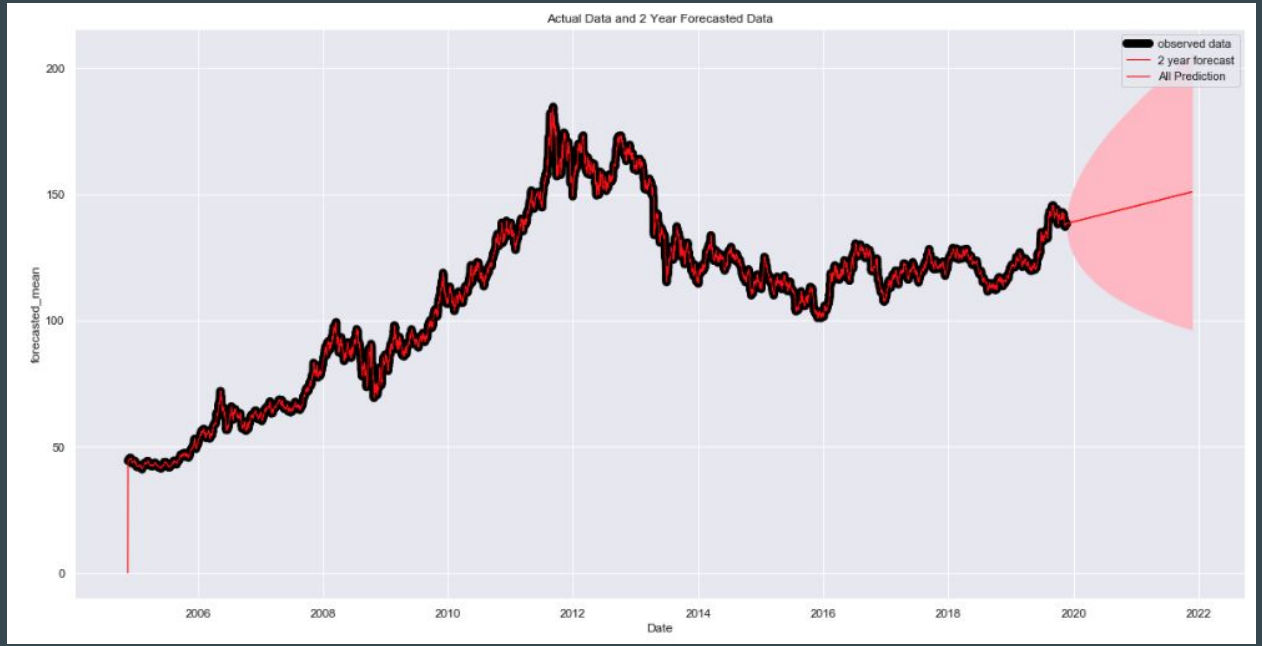
Having validated our forecast results with our test data, we are going to perform an out of sample forecasting using the auto_arima method

Forecast the price of gold for the next 2 years from 11/23/2019 - 11/21/2021.

Model Parameters:

```
all_auto_arima_model = SARIMAX(arima_data,
                               seasonal=True,
                               order=(0,1,0),
                               seasonal_order=(0,0,0,7),
                               trend='c')
```

Use Model to Predict Prices from 11/18/04-11/21/19 and Forecast 2 Years Out to 11/21/21



The red line is the prediction results from 11/18/2004 - 11/21/2021. You can see that predicted prices are very well aligned to the actual prices as shown in the black line. The 2 year forecast does indicate an upward trend in gold price ETFs in the next 2 years.

Forecasting Conclusion

- Between ARIMA and Facebook Prophet, ARIMA shows a better fit between actual data and predicted data.
- In the out-of-sample forecast, the ARIMA model shows an upward trend in gold prices for the next 2 years, forecasting the price of gold to be at $150.80 by 11/21/2021. That's 9.28% increase from today's price of $138.
- GLD seems to be a safe asset to buy now.

Disclaimer: This report is for educational purposes. It’s not meant to be any form of financial advice.