# 1. Review the quality of the data, list any potential errors, and propose corrected values. Please list each quality check error and correction applied.

From a visual inspection of the Data, we can see that there are a number of anomalies and extreme values in the ETF price data. Some of these extremes are present in the Close Price, but not the open Price, and Vice Versa. To locate and eliminate anomalies, the Facebook Prophet model is applied on the Close Price.

Prices outside the 99% confidence interval are replaced with the mean predicted price as shown below (Fig2).

**The Close Price is used for analysis throughout the project.**

**Fig 1**

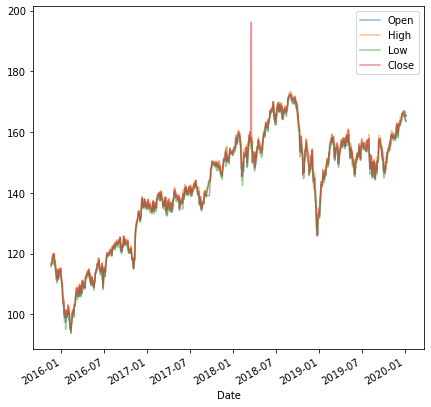
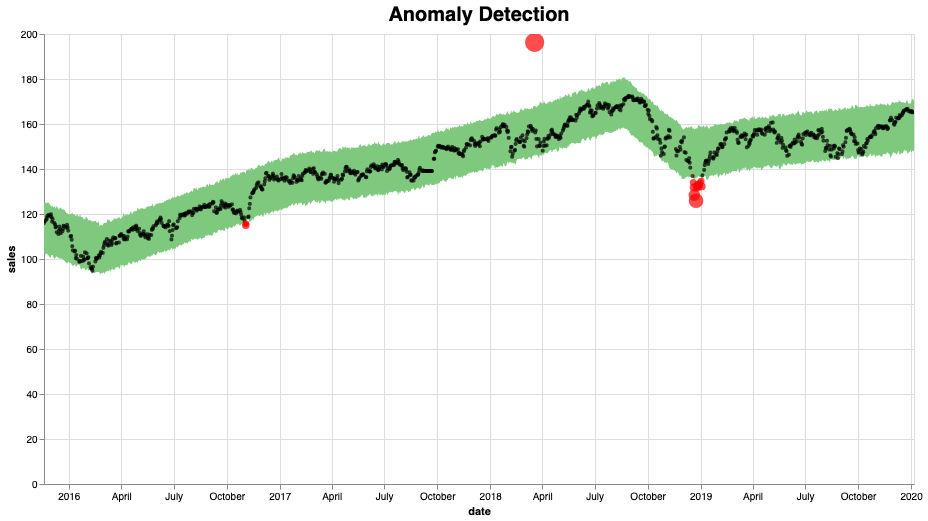
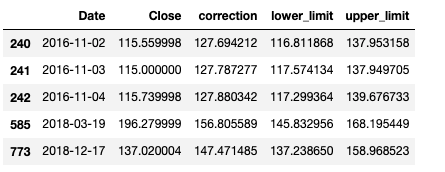


Fig 2



**Fig 3: Corrections for Anomalous Data**



# 2. Please analyze the signal’s effectiveness or lack thereof in forecasting ETF price, using whatever metrics you think are most relevant.

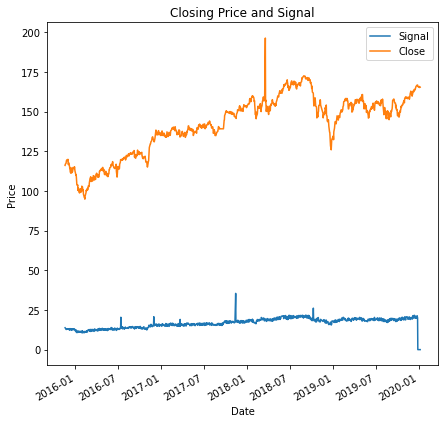
The signal is off from the ‘Close’ Price in terms of predicting the mean Price(bias) and changes in Price. This is seen in the figure below. In order to check the performance of the Signal on the mean normalized price, I performed the following transformations on the Close price and Signal.

mean\_normalized\_price = Close price - mean(Close Price)

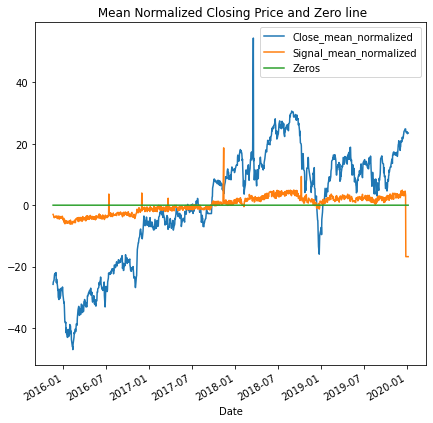
mean\_normalized\_signal = Signal - mean(Signal)

From Fig 5, we see that the Signal only slightly outperforms the Zero line in terms of predicting Closing Price (Close).

**Fig 4**



**Fig 5 Closing Price and Zero line.**



RMSE: Signal-Normalized 16.113204044707746

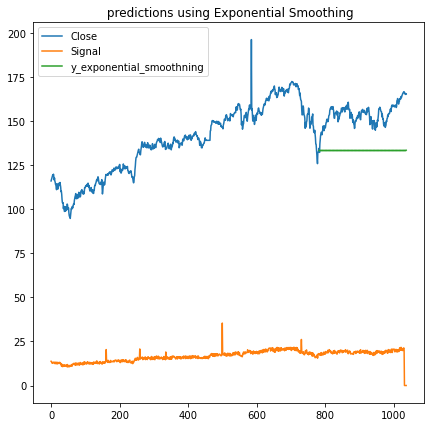
MAPE: Signal 0.8845531093170288

RMSE: Zero Line 18.488098021819695

MAPE: Zero Line 1.0

# 3. (Extra credit) Write a 1-2 paragraph summary for the Portfolio Manager addressing your observations about the efficacy and believability of the product, and recommendation for next steps.

**Fig 6: Predictions using Signal, Mean and Exponential Smoothing**



MAPE: Exponential Smoothing 0.13297683518918785

MAPE: mean model 0.1037292477617683

MAPE: Signal 0.8789067912721177

The Signal is only slightly better than the Naive Prediction. Thus it will have to be discarded. There are two major errors - 1)The average price level indicated by the Signal (16) is significantly different from the actual price (141). 2) For the mean normalized comparison of Prices between the Signal and actual Price, the Signal does a slightly better than a naive prediction (mean of the timeseries). Thus the Signal is unable to capture the Changes in Price effectively.

In Fig 6 we see that the Signal (MAPE: 87%) performs worse than the mean model (MAPE: 10%) and the exponential smoothing model (MAPE: 13%). Although, in this case, the exponential smoothing model performs worse than the mean model, through better training and using more sophisticated models, improvements over the mean model could be achieved.

I would recommend to develop basic models against which one can check performance of the Product, including models like moving average, exponential smoothening and ARIMA. As can be seen above, a simple exponential smoothing model performs better than the custom model

# 