Abstract

We make use of time series to predict the future values of the Honeywell stock. We perform exponential smoothing forecast on Honeywell stock prices with varying value of parameters to find the best fit. To find the best fit we make use of SSE and MSE and compare all the values. We also perform the prediction using linear regression analysis and compare the results with exponential smoothing forecast. We find the coefficient of correlation and determination. We learn more about the residuals and their shapes when used in scatterplots. We also find the actual Honeywell stock price and compare it with all our forecasts.

Keywords: [Timeseries, forecast, exponential smoothing, prediction, linear regression, stock prices, MSE]

[Title Here, up to 12 Words, on One to Two Lines]

We are given with the historical stock prices of Honeywell corporation from 10/15/2017 to 4/15/2018. We perform exponential smoothing forecasts on Honeywell stock prices to forecast the price for 4/16/2018. We use of 0.15, 0,35, 0.55, and 0.75 for the smoothing parameter α and calculate the SSE and MSE to find out which value of α that has provided the most accurate forecast. We then use the smoothing parameter α = 0.75 and use 0.15, 0.25, 0.45, and 0.85 for the trend parameter β. We then find which value of β has yielded the most accurate forecast.

After using Holt-Winters exponential smoothing we perform the same analysis using linear regression. We then find the coefficients of correlation and determination and interpret their values. We then find the residuals and plot them with time and predicted stock values to study homoscedasticity.

Finally, we compare the actual stock price with our predictions to find which method has provided with the most accurate results.

The analysis has been performed using R studio.

**Methodology and Analysis**

We are required to perform exponential smoothing on Honeywell stock prices to forecast the price for 4/16/2018. We load the dataset into R studio and plot the timeseries.

A close up of text on a white background

Description generated with very high confidence

We then perform exponential smoothing using HoltWinters function.

mod1 <- HoltWinters(ts\_honey, alpha=0.15, beta=FALSE, gamma=FALSE)

We use 0.15, 0,35, 0.55, and 0.75 for the smoothing parameter α. We then find which smoothing parameter provided with the best results by using SSE.

Eg: mod1$SSE

We find that that alpha 0.75 generates the best result with SSE of 366.29 and the worst results are produced with alpha 0.15 with SSE of 980.51.

We also calculate the best alpha value which is alpha =0.96 which gives SSE of 351.85.

Such a high value of alpha indicates that recent values of the stock are the strongest indicator of the predicted price and does not depend much on the distant past values.

We also find the MSE of the values. It is found using - mean((ts\_honey[-1,] - temp1[,1]))^2. We get the lease MSE of alpha = 0.75 which comes to 0.00063 and the highest with alpha = 0.15 which comes to 0.00785.

We then plot the forecast with alpha 0.75 and get the following results:

A close up of a logo

Description generated with high confidence

Our forecast from HoltWinters comes to 143.4575 for 4/16/2018.

We are required to perform exponential smoothing on Honeywell stock prices to forecast the price for 4/16/2018. We load the dataset into R studio and plot the timeseries.

We then perform exponential smoothing using HoltWinters function with alpha 0.75 and we use 0.15, 0.25, 0.45, and 0.85 for the trend parameter β. We then find which value of β has yielded the most accurate forecast.

Eg: mod1 <- HoltWinters(ts\_honey, alpha=0.75, beta=0.15, gamma=FALSE)

We then find which smoothing parameter provided with the best results by using SSE.

Eg: mod1$SSE

We find that that beta 0.15 generates the best result with SSE of 396.3 and the worst results are produced with beta 0.85 with SSE of 513.29.

We also calculate the best beta value which is beta =0.012 which gives SSE of 361.

The value of beta which is close to zero indicates that the estimate of the slope b of the trend component is not updated over the time series. As the level changes a bit over the time series, but the slope b of the trend component roughly remains the same.

We also find the MSE of the values. It is found using - mean((ts\_honey[c(-1,-2),] - temp1[,1]))^2. We get the lease MSE with beta = 0.15 which comes to 0.0001 and the highest with beta = 0.85 which comes to 0.0011.

We then plot the forecast with beta 0.15 and get the following results:

A screenshot of a cell phone

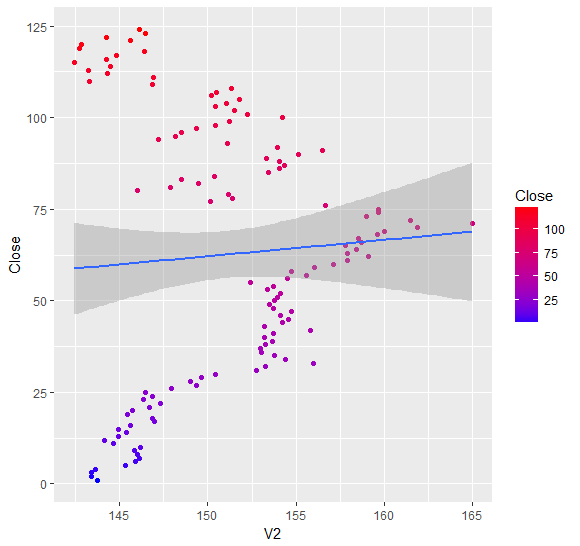
Description generated with high confidence

Our forecast from HoltWinters comes to 146.13 for 4/16/2018.

We need to perform the above steps using linear regression analysis. We proceed by using

df <- lm(ts\_honey$V2 ~ ts\_honey$Close)

We have plotted it using ggplot2 to get an interactive visual chart.



We then predict future values using prediction <- predict(df, newdata = new,se.fit = TRUE).

We get the predicted value for 04/16/2018 as 150.38.

We find the SSE and MSE of the regression using:

mean((ts\_honey[,2] - prediction$fit)^2)

sse <- sum(df$residuals^2)

We get the following values SSE = 3167.2 and MSE = 25.54.

The SSE from first part with alpha 0.75 and Beta not specified was 366.29 and MSE was 0.00063.

The SSE from second part with alpha 0.75 and Beta 0.15 was 396.3 and MSE was 0.0001.

Hence, we can see how badly the regression analysis has performed compared to timeseries analysis.

Using summary on the regression analysis we get,

A screenshot of a cell phone

Description generated with very high confidence

We can see only stock prices are significant which is obvious, the ts\_honey$close (I didn’t change column names but is actually the date) indicates the dates which are not significant for the analysis.

We get R squared 0.003981 and Adjusted R squared as -0.004183 which indicates how far the points are from the regression line and is a very bad model.

When we plot the residuals on a histogram we get the following results:

A screenshot of a cell phone

Description generated with high confidence

We can see most of the residuals are at a distance of +-5 from the regression line which is large distance and hence we have a bad model.

We then make the normal probability plot

A close up of a map

Description generated with very high confidence

We can see out sample data is slightly skewed and not a uniform distribution.

As some of the values of residuals are negative performing a chi-square test is not possible by removing negative sign of the values and performing chi-squared test we get following results:

A close up of a person

Description generated with high confidence

We then plot the residuals versus time and get the following results:

A close up of a piece of paper

Description generated with high confidence

As seen from the histogram we can see most the values of the residuals range from -5 to 5 and we can see many 10 values during 10th October to 18th October which have not been captured by the regression line effectively.

Let us plot the residuals with the predicted values:

A screenshot of a cell phone

Description generated with high confidence

We can see that the values range 150.4 to 151.4 with distance from regression line mostly ranging from -5 to 5. Homoscedasticity means having the same scatter. The points must be about the same distance from the regression line and the opposite is heteroscedasticity where points are at varying distance from the regression line.

We can see there is heteroscedasticity here as the points are scattered randomly.

Looking at the actual Honeywell stock price on 04/16/2018 we get the closing price as 146.75.

The predicted stock price from our analysis are 143.45 using the analysis from the first part. We get 146.13 prediction using the analysis explained in the second part and 150.38 using linear regression analysis.

Hence, we can see that the best estimation is given by exponential forecasting method with beta =0.15 and alpha = 0.75 which comes to 146.13 almost equal to the actual value of 146.75.

We can also see how badly the linear regression performed with the estimation of 150.38.

**Conclusion**

We have learned the implementation of time series and calculated the future values using exponential smoothing forecasting and linear regression analysis. We can see that we cannot use linear regression for everything and time series models are designed to handle such kind of data to make a prediction. We also used different combinations of alpha and beta to test which gives the best result. From our findings the values of stock are mostly dependent on the recent values and has very less weightage to the distant past values.

The best forecast was given as 146.13 which is very close to the actual value of the stock 146.75. We also learned about residuals, scatter plots, histograms, normal probability plots, chi squared test.

References

1) Using R for Time Series Analysis¶. (n.d.). Retrieved from <http://a-little-book-of-r-for-time-series.readthedocs.io/en/latest/src/timeseries.html#holt-winters-exponential-smoothing>

2) Homoscedasticity / Homogeneity of Variance/ Assumption of Equal Variance. (n.d.). Retrieved from <http://www.statisticshowto.com/homoscedasticity/>

3) Historical Stock Prices. (n.d.). Retrieved from <https://www.nasdaq.com/symbol/hon/historical>