

Time Series Modelling: A Comprehensive Analysis of ASX All Ordinaries, Gold, Crude Oil, and Copper Prices (2004-2017)

Nikki Lau

Introduction

This report focuses on the monthly averages of ASX All Ordinaries (Ords) price index, gold (AUD), crude oil (Brent, USD/bbl), and copper (USD/Tonne) prices from January 2004 to December 2017. The aim of the project is to investigate and analyse the existence of nonstationarity, the impact of the components of the time series data, and to provide the most accurate and suitable distributed lag model for the ASX price index. This will be conducted through descriptive analyses, time series modelling and diagnostic checking.

Dataset

The dataset contains 161 observations with the following 4 variables:

- ASX Price Index
- Gold Price
- Crude Oil Price
- Copper Price

Method

```
#Load all libraries
library(readr)
library(tseries)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo
```

```
library(forecast)
library(TSA)
```

```
## Registered S3 methods overwritten by 'TSA':
##   method      from
##   fitted.Arima forecast
##   plot.Arima  forecast
```

```
##
## Attaching package: 'TSA'
```

```
## The following object is masked from 'package:readr':  
##  
##     spec
```

```
## The following objects are masked from 'package:stats':  
##  
##     acf, arima
```

```
## The following object is masked from 'package:utils':  
##  
##     tar
```

```
library(x12)
```

```
## Loading required package: x13binary
```

```
## x12 is ready to use.
```

```
## Use the package x12GUI for a Graphical User Interface.
```

```
## By default the X13-ARIMA-SEATS binaries provided by the R package x13binary
```

```
## are used but this can be changed with x12path(validpath)
```

```
## -----
```

```
## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/x12/issues
```

```
library(dLagM)
```

```
## Loading required package: nardl
```

```
## Loading required package: dynlm
```

```
## Loading required package: zoo
```

```
##  
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':  
##  
##     as.Date, as.Date.numeric
```

```
##  
## Attaching package: 'dLagM'
```

```
## The following object is masked from 'package:forecast':  
##  
##     forecast
```

```
library(urca)
```

Importing Data

```
#Load dataset  
asx = read_csv("/Users/nikz/Desktop/ASX_data.csv")
```

```
## Rows: 161 Columns: 4  
## — Column specification ——————  
## Delimiter: ","  
## dbl (2): ASX_price, Crude_Oil  
## num (2): Gold_price, Copper  
##  
## i Use `spec()` to retrieve the full column specification for this data.  
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(asx)
```

ASX_price <dbl>	Gold_price <dbl>	Crude_Oil <dbl>	Copper <dbl>
2935.4	611.9	31.29	1650
2778.4	603.3	32.65	1682
2848.6	565.7	30.34	1656
2970.9	538.6	25.02	1588
2979.8	549.4	25.81	1651
2999.7	535.9	27.55	1685

6 rows

Data Pre-processing

With time series, the dataset needs to be converted to time series objects before implementing time series analysis. Datasets are typically loaded as a matrix which will need to be vectorised to convert into a time series object.

```
#Check dataframe structure  
str(asx)
```

```

## spc_tbl_ [161 x 4] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ ASX_price : num [1:161] 2935 2778 2849 2971 2980 ...
## $ Gold_price: num [1:161] 612 603 566 539 549 ...
## $ Crude_Oil : num [1:161] 31.3 32.6 30.3 25 25.8 ...
## $ Copper     : num [1:161] 1650 1682 1656 1588 1651 ...
## - attr(*, "spec")=
##   .. cols(
##     ..  ASX_price = col_double(),
##     ..  Gold_price = col_number(),
##     ..  Crude_Oil = col_double(),
##     ..  Copper = col_number()
##   .. )
## - attr(*, "problems")=<externalptr>

```

```

#Vectorise values
asx_p <- as.vector(asx$`ASX_price`)
gold_p <- as.vector(asx$`Gold_price`)
crude_p <- as.vector(asx$`Crude_Oil`)
copper_p <- as.vector(asx$`Copper`)

```

```

#Convert vectors into time series objects
asx_ts <- ts(asx_p, frequency = 12, start = c(2004,1))
head(asx_ts)

```

```

##           Jan    Feb    Mar    Apr    May    Jun
## 2004 2935.4 2778.4 2848.6 2970.9 2979.8 2999.7

```

```

gold_ts <- ts(gold_p, frequency = 12, start = c(2004, 1))
head(gold_ts)

```

```

##           Jan    Feb    Mar    Apr    May    Jun
## 2004 611.9 603.3 565.7 538.6 549.4 535.9

```

```

crude_ts <- ts(crude_p, frequency = 12, start = c(2004, 1))
head(crude_ts)

```

```

##           Jan    Feb    Mar    Apr    May    Jun
## 2004 31.29 32.65 30.34 25.02 25.81 27.55

```

```

copper_ts <- ts(copper_p, frequency = 12, start = c(2004,1))
head(copper_ts)

```

```

##           Jan    Feb    Mar    Apr    May    Jun
## 2004 1650 1682 1656 1588 1651 1685

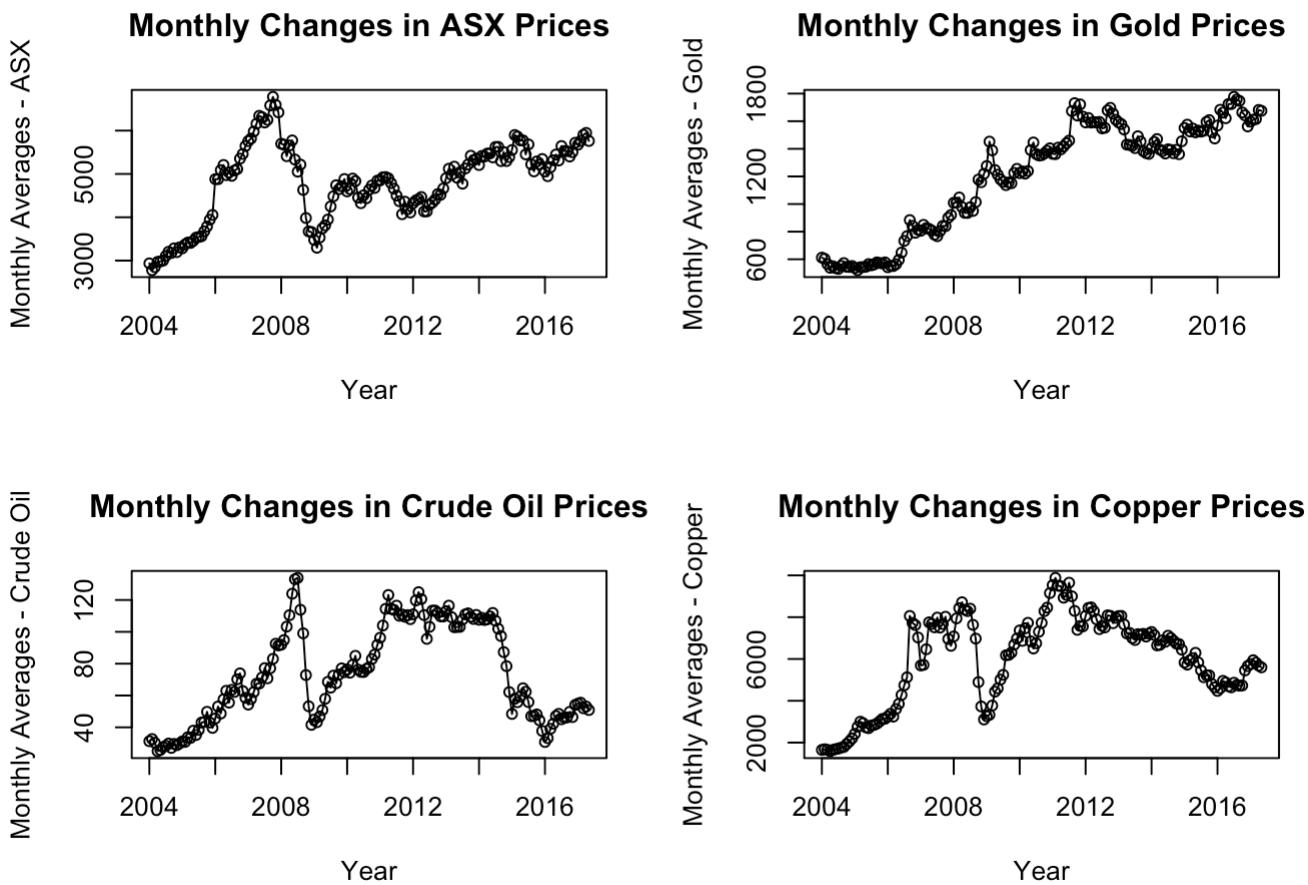
```

Visualisation of Time Series

Time series plots are useful for visualising trends of data across time periods. In this case, through the visualisation of the time series plot, ASX index, gold, crude oil and copper prices can be analysed over the monthly periods from 2004-2017. From these time series plots, the existence of trend, seasonality, changing variance over time, behaviour and intervention are displayed which will assist in analysing the patterns of the time series plots.

Time Series Plots

```
par(mfrow=c(2,2))
plot(asx_ts, type = 'o', ylab = "Monthly Averages - ASX", xlab = "Year", main = "Monthly Changes in ASX Prices")
plot(gold_ts, type = 'o', ylab = "Monthly Averages - Gold", xlab = "Year", main = "Monthly Changes in Gold Prices")
plot(crude_ts, type = 'o', ylab = "Monthly Averages - Crude Oil", xlab = "Year", main = "Monthly Changes in Crude Oil Prices")
plot(copper_ts, type = 'o', ylab = "Monthly Averages - Copper", xlab = "Year", main = "Monthly Changes in Copper Prices")
```



ASX:

- Trend: There is a definite upward trend from 2004 - 2008, with a declining trend in 2008-2009. The trend displays an upward trend again after 2010
- Seasonality: There is no seasonality present as there are no regular repeating patterns in every month or year
- Intervention: There is an obvious intervention at 2008
- Changing variance: There is no sign of changing variance
- Behaviour: The data points are related to each other as the previous data point affects the next data points

Gold:

- Trend: There is an overall upward, positive trend
- Seasonality: There is no seasonality present as there are no regular repeating patterns in every month or year
- Intervention: There is no sign of an intervention
- Changing variance: There is no sign of changing variance
- Behaviour: The data points are related to each other as the previous data point affects the next data points

Crude Oil:

- Trend: There is a definite upward trend from 2004-2008 where it is met with a steep downward trend. The trend increases before plateauing and declining
- Seasonality: There is no seasonality present as there are no regular repeating patterns in every month or year
- Intervention: There is an obvious intervention at 2008
- Changing variance: There is no sign of changing variance
- Behaviour: The data points are related to each other as the previous data point affects the next data points

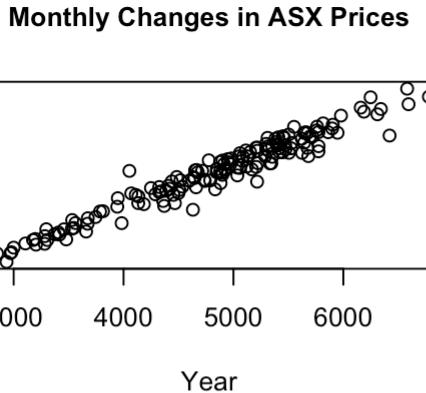
Copper:

- Trend: There is a definite positive trend which changes over time to a negative trend
- Seasonality: There is no seasonality present as there are no regular repeating patterns in every month or year
- Intervention: There is an obvious intervention at 2008
- Changing variance: There is no sign of changing variance
- Behaviour: The data points are related to each other as the previous data point affects the next data points

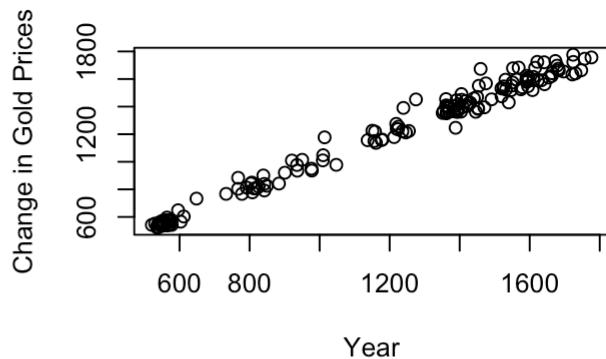
Scatterplots

```
par(mfrow=c(2,2))
plot(y = asx_ts, x = zlag(asx_ts),
      ylab = "Change in ASX Prices",
      xlab = "Year",
      main = "Monthly Changes in ASX Prices",
      cex.main = 1)
plot(y = gold_ts, x = zlag(gold_ts),
      ylab = "Change in Gold Prices",
      xlab = "Year",
      main = "Monthly Changes in Gold Prices",
      cex.main = 1)
plot(y = crude_ts, x = zlag(crude_ts),
      ylab = "Change in Crude Prices",
      xlab = "Year",
      main = "Monthly Changes in Crude Prices",
      cex.main = 1)
plot(y = copper_ts, x = zlag(copper_ts),
      ylab = "Change in Copper Prices",
      xlab = "Year",
      main = "Monthly Changes in Copper Prices",
      cex.main = 1)
```

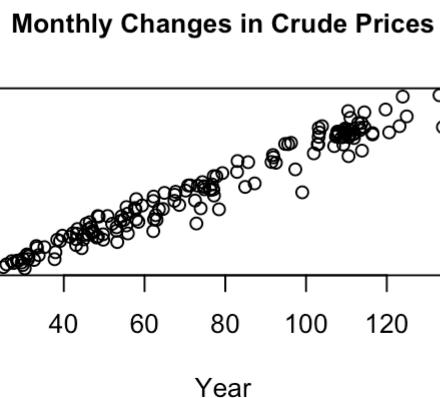
Change in ASX Prices



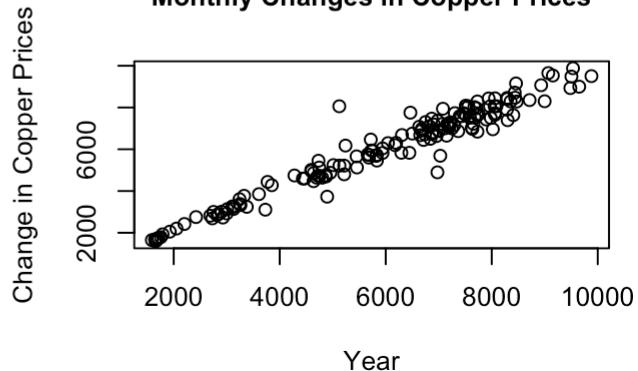
Change in Gold Prices



Change in Crude Prices



Change in Copper Prices



```
y = asx_ts
x = zlag(asx_ts)
exclude_na = 2:length(x)
asx_corr = cor(y[exclude_na],x[exclude_na])
print(paste("ASX Corr:", asx_corr))
```

```
## [1] "ASX Corr: 0.974060525146871"
```

```
y = gold_ts
x = zlag(gold_ts)
exclude_na = 2:length(x)
gold_corr = cor(y[exclude_na],x[exclude_na])
print(paste("Gold Corr:", gold_corr))
```

```
## [1] "Gold Corr: 0.99141659359352"
```

```
y = crude_ts
x = zlag(crude_ts)
exclude_na = 2:length(x)
crude_corr = cor(y[exclude_na],x[exclude_na])
print(paste("Crude Corr:", crude_corr))
```

```
## [1] "Crude Corr: 0.977821840210335"
```

```
y = copper_ts
x = zlag(copper_ts)
exclude_na = 2:length(x)
copper_corr = cor(y[exclude_na],x[exclude_na])
print(paste("Copper Corr:", copper_corr))
```

```
## [1] "Copper Corr: 0.974238698133264"
```

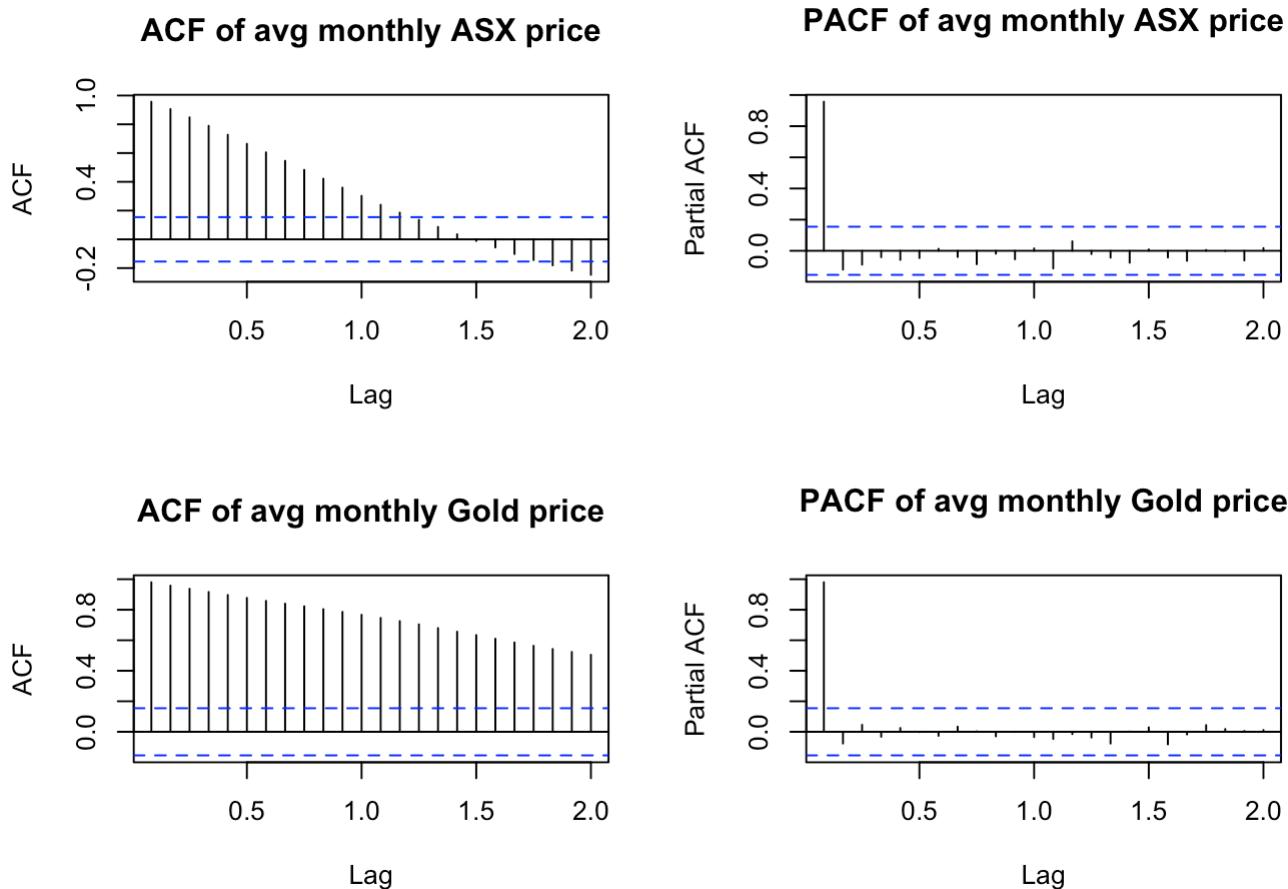
The scatterplots for ASX, gold, crude and copper prices all show a strong, linear, positive trend. The correlation values (0.9741, 0.9914, 0.9778, 0.9723), respectively, further indicates a strong correlation between the values of consecutive monthly ASX, gold, crude or copper prices. This indicates that there is a strong correlation between consecutive ASX prices, and therefore, the prices of ASX of one month impacts the ASX price of the next month. This is the same for gold, crude oil and copper prices.

Residual Analysis

ACF and PACF

```
#ACF and PACF
par(mfrow=c(2,2))
acf(asx_ts, lag.max = 24, main = "ACF of avg monthly ASX price")
pacf(asx_ts, lag.max = 24, main = "PACF of avg monthly ASX price")

acf(gold_ts, lag.max = 24, main = "ACF of avg monthly Gold price")
pacf(gold_ts, lag.max = 24, main = "PACF of avg monthly Gold price")
```

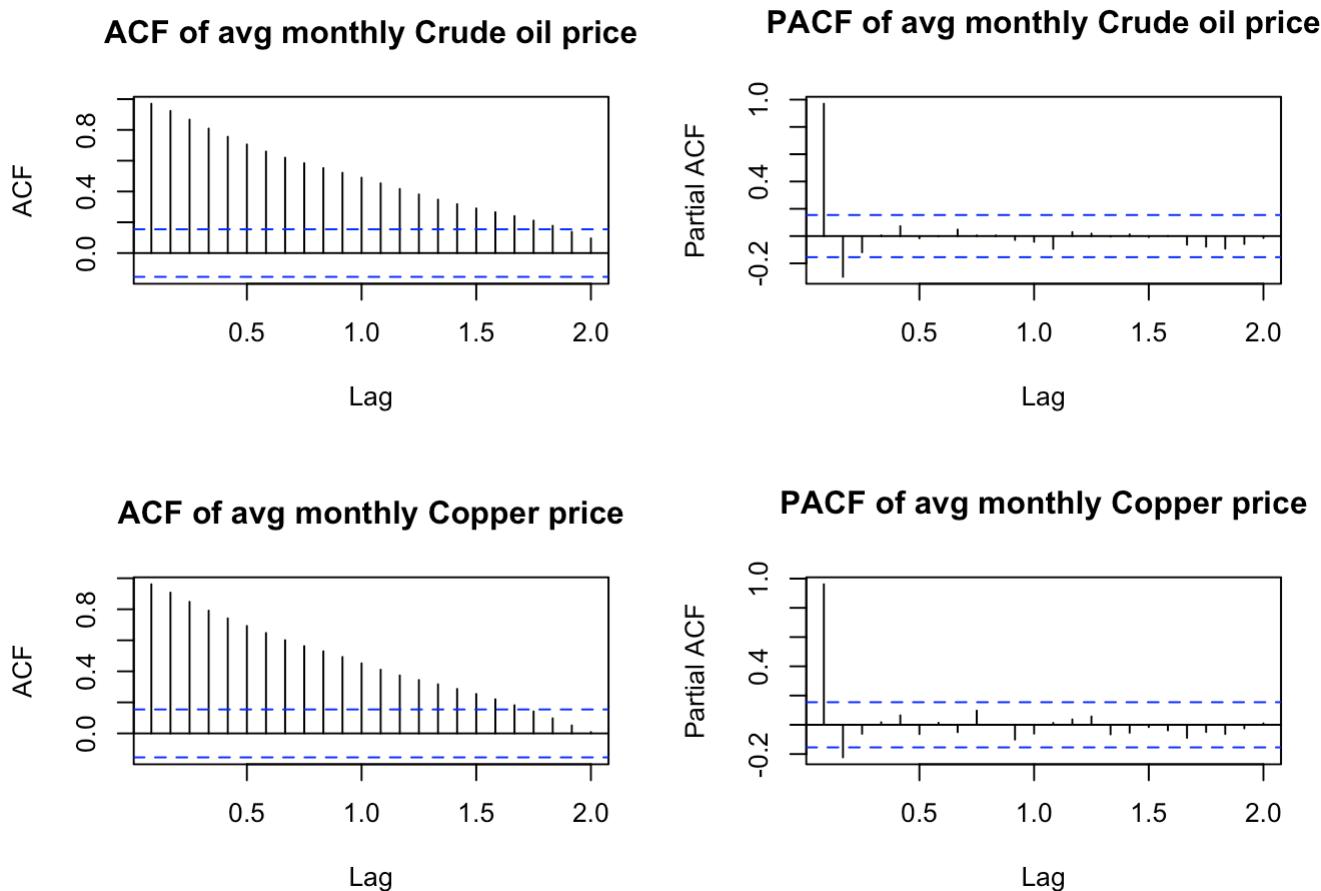


ASX Price: From the ACF plot, as the lags are not within the boundary lines, it supports that there is an existence of a trend in the series. Furthermore, with a very high first autocorrelation seen in the PACF plot, the series is considered as nonstationary.

Gold Price: From the ACF plot, as the lags are not within the boundary lines, it supports that there is an existence of a trend in the series. Furthermore, with a very high first autocorrelation seen in the PACF plot, the series is considered as nonstationary.

```
par(mfrow=c(2,2))
acf(crude_ts, lag.max = 24, main = "ACF of avg monthly Crude oil price")
pacf(crude_ts, lag.max = 24, main = "PACF of avg monthly Crude oil price")

acf(copper_ts, lag.max = 24, main = "ACF of avg monthly Copper price")
pacf(copper_ts, lag.max = 24, main = "PACF of avg monthly Copper price")
```



Crude Oil Price: From the ACF plot, as the lags are not within the boundary lines, it supports that there is an existence of a trend in the series. Furthermore, with a very high first autocorrelation seen in the PACF plot, the series is considered as nonstationary.

Copper Price: From the ACF plot, as the lags are not within the boundary lines, it supports that there is an existence of a trend in the series. Furthermore, with a very high first autocorrelation seen in the PACF plot, the series is considered as nonstationary.

Unit Root Tests

In order to further confirm whether the series are stationary or nonstationary, unit root test statistics can be conducted. However in order to perform unit root tests, the lag length of each series needs to be determined to specify the regression model that will be used for testing. After selecting the lag length, various unit root test statistics can be performed.

```
k_asx = ar(diff(asx_ts))$order
k_asx
```

```
## [1] 0
```

```
k_gold = ar(diff(gold_ts))$order
k_gold
```

```
## [1] 2
```

```
k_crude = ar(diff(crude_ts))$order  
k_crude
```

```
## [1] 1
```

```
k_copper = ar(diff(copper_ts))$order  
k_copper
```

```
## [1] 1
```

```
adf.test(asx_ts, k = k_asx)
```

```
##  
##  Augmented Dickey–Fuller Test  
##  
## data: asx_ts  
## Dickey–Fuller = -1.9684, Lag order = 0, p-value = 0.5895  
## alternative hypothesis: stationary
```

```
adf.test(asx_ts, k = k_gold)
```

```
##  
##  Augmented Dickey–Fuller Test  
##  
## data: asx_ts  
## Dickey–Fuller = -2.442, Lag order = 2, p-value = 0.392  
## alternative hypothesis: stationary
```

```
adf.test(asx_ts, k = k_crude)
```

```
##  
##  Augmented Dickey–Fuller Test  
##  
## data: asx_ts  
## Dickey–Fuller = -2.224, Lag order = 1, p-value = 0.4829  
## alternative hypothesis: stationary
```

```
adf.test(asx_ts, k = k_copper)
```

```
##  
##  Augmented Dickey–Fuller Test  
##  
## data: asx_ts  
## Dickey–Fuller = -2.224, Lag order = 1, p-value = 0.4829  
## alternative hypothesis: stationary
```

```
PP.test(asx_ts, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: asx_ts  
## Dickey-Fuller = -2.2284, Truncation lag parameter = 4, p-value = 0.4811
```

```
PP.test(gold_ts, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: gold_ts  
## Dickey-Fuller = -2.1044, Truncation lag parameter = 4, p-value = 0.5328
```

```
PP.test(crude_ts, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: crude_ts  
## Dickey-Fuller = -1.7567, Truncation lag parameter = 4, p-value = 0.6778
```

```
PP.test(copper_ts, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: copper_ts  
## Dickey-Fuller = -2.0718, Truncation lag parameter = 4, p-value = 0.5464
```

The lag lengths for the asx, gold, crude oil and copper price series were determined to be 0, 2, 1 and 1, respectively. From these lag lengths, the augmented dickey-fuller test (ADF) were performed for each series, where the null and alternative hypotheses are as follows:

- Null hypothesis: There is a unit root
- Alternative hypothesis: There is stationarity

Based on the ADF test results for ASX, gold, crude oil and copper prices indicates nonstationarity at a 5% level of significance as $p > 0.05$.

Furthermore, the Phillips-Perron (PP) test was performed to further confirm that stationarity of the series. According to the PP test, the series are nonstationary at a 5% level of significance as $p > 0.05$.

```
shapiro.test(asx_ts)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: asx_ts  
## W = 0.9709, p-value = 0.001786
```

```
shapiro.test(gold_ts)
```

```
##
## Shapiro-Wilk normality test
##
## data: gold_ts
## W = 0.8886, p-value = 1.202e-09
```

```
shapiro.test(crude_ts)
```

```
##
## Shapiro-Wilk normality test
##
## data: crude_ts
## W = 0.93463, p-value = 9.927e-07
```

```
shapiro.test(copper_ts)
```

```
##
## Shapiro-Wilk normality test
##
## data: copper_ts
## W = 0.94632, p-value = 8.282e-06
```

The shapiro-wilk test shows that the normality of residuals are supported for ASX, gold, crude, copper prices, at a 5% level of significance as $p > 0.05$.

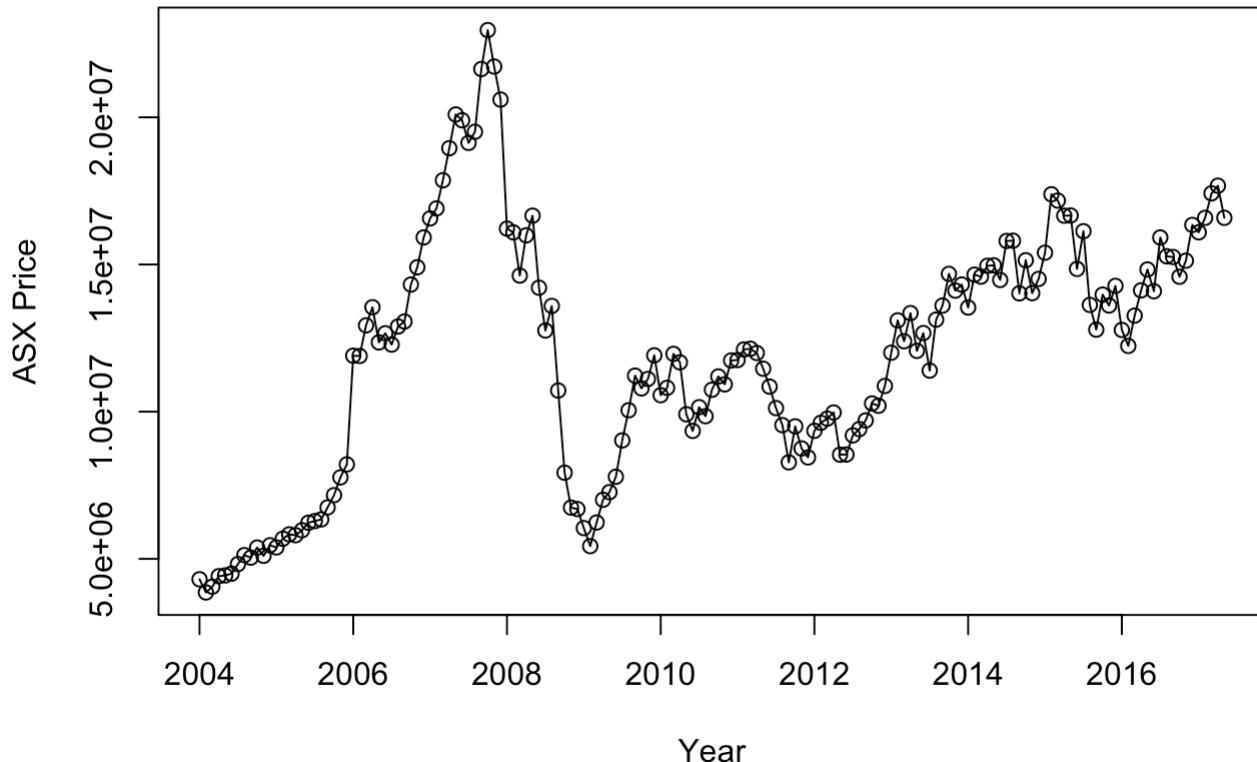
Transformation

```
#Lambda values
asx_lambda = BoxCox.lambda(asx_ts)
gold_lambda = BoxCox.lambda(gold_ts)
crude_lambda = BoxCox.lambda(crude_ts)
copper_lambda = BoxCox.lambda(copper_ts)
cbind(asx_lambda, gold_lambda, crude_lambda, copper_lambda)
```

```
##      asx_lambda gold_lambda crude_lambda copper_lambda
## [1,]    1.999924     0.976695    -0.8304136     0.9336783
```

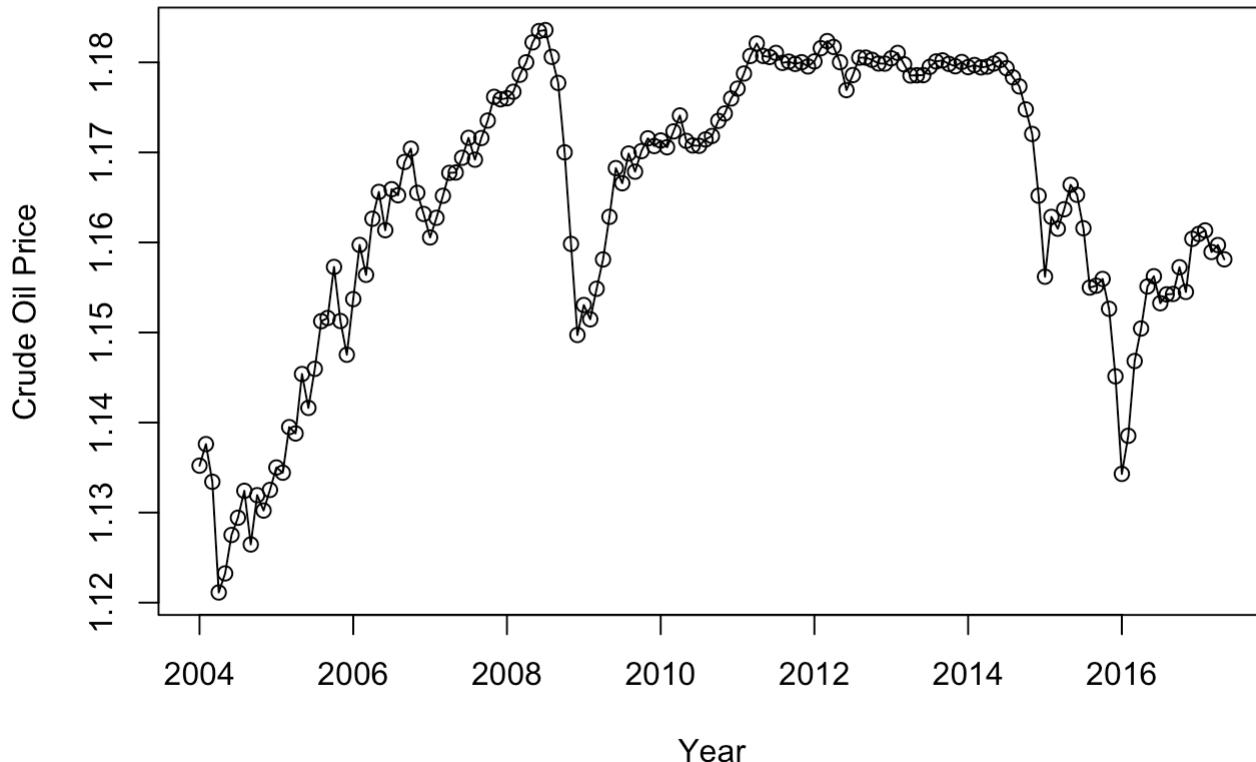
```
#ASX transformation
asx.bc = ((asx_ts^(asx_lambda)) - 1) / asx_lambda
plot(asx.bc, ylab = "ASX Price", xlab = "Year", type = 'o', main = "Box-Cox Transformed ASX Price Series")
```

Box-Cox Transformed ASX Price Series



```
#Crude oil transformation  
crude.bc = ((crude_ts^(crude_lambda)) - 1) / crude_lambda  
plot(crude.bc, ylab = "Crude Oil Price", xlab = "Year", type = 'o', main = "Box-Cox T  
ransformed Crude Oil Price Series")
```

Box-Cox Transformed Crude Oil Price Series



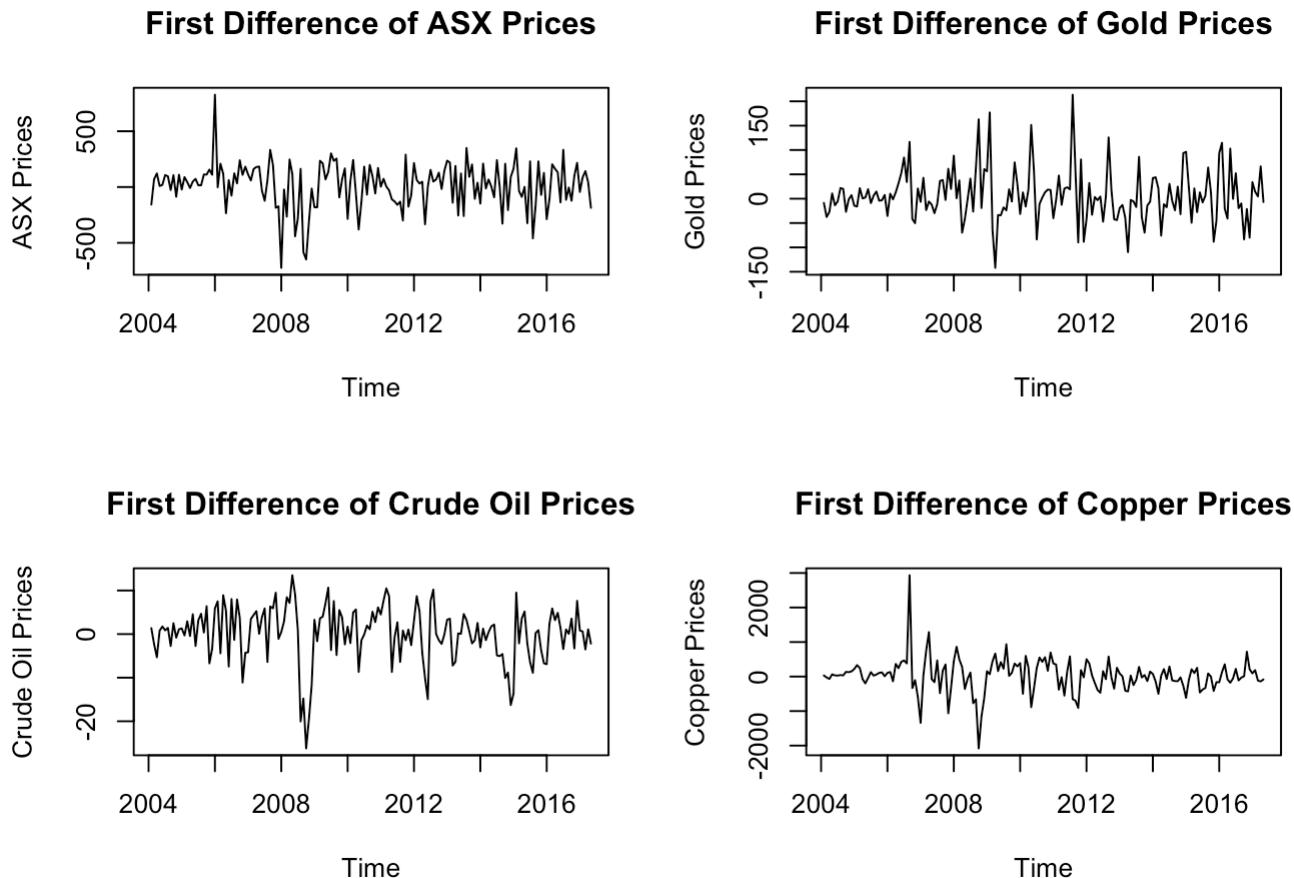
From the lambda values, it shows that for ASX price index, the lambda value is close to 2, hence a transformation is required. In addition, as the lambda value for crude oil price is negative, a transformation is also required. For the others, gold and copper prices, as their lambda values are close to 1, no transformation is required.

As the lambda value for ASX price index is close to 2, a (Y^2) transformation was used, while a (Y^{-1}) transformation was used for crude oil due to its negative lambda value. A time series plot for ASX price index and crude oil prices were used to visualise the effects of the transformation. As the time series plot appear to be identical to the original time series plot for ASX price index and crude oil price, and hence still appear nonstationary, the transformations were not applied to the series.

Differencing

```
asx_diff = diff(asx_ts)
gold_diff = diff(gold_ts)
crude_diff = diff(crude_ts)
copper_diff = diff(copper_ts)
```

```
par(mfrow=c(2,2))
plot(asx_diff, ylab = "ASX Prices", xlab = "Time", main = "First Difference of ASX Prices")
plot(gold_diff, ylab = "Gold Prices", xlab = "Time", main = "First Difference of Gold Prices")
plot(crude_diff, ylab = "Crude Oil Prices", xlab = "Time", main = "First Difference of Crude Oil Prices")
plot(copper_diff, ylab = "Copper Prices", xlab = "Time", main = "First Difference of Copper Prices")
```



```
adf.test(asx_diff)
```

```
## Warning in adf.test(asx_diff): p-value smaller than printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: asx_diff
## Dickey-Fuller = -4.5543, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

```
adf.test(gold_diff)
```

```
## Warning in adf.test(gold_diff): p-value smaller than printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: gold_diff
## Dickey-Fuller = -5.8718, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

```
adf.test(crude_diff)
```

```
## Warning in adf.test(crude_diff): p-value smaller than printed p-value
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: crude_diff  
## Dickey-Fuller = -5.4261, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

```
adf.test(copper_diff)
```

```
## Warning in adf.test(copper_diff): p-value smaller than printed p-value
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: copper_diff  
## Dickey-Fuller = -5.478, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

```
PP.test(asx_diff, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: asx_diff  
## Dickey-Fuller = -11.674, Truncation lag parameter = 4, p-value = 0.01
```

```
PP.test(gold_diff, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: gold_diff  
## Dickey-Fuller = -11.014, Truncation lag parameter = 4, p-value = 0.01
```

```
PP.test(crude_diff, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: crude_diff  
## Dickey-Fuller = -8.2597, Truncation lag parameter = 4, p-value = 0.01
```

```
PP.test(copper_diff, lshort = TRUE)
```

```
##  
## Phillips-Perron Unit Root Test  
##  
## data: copper_diff  
## Dickey-Fuller = -9.1092, Truncation lag parameter = 4, p-value = 0.01
```

As the series appear to be non-stationary, the series will be differenced and the stationarity of the series will be observed again.

After first differencing the series, the series appear stationary. This was confirmed with the unit root tests.

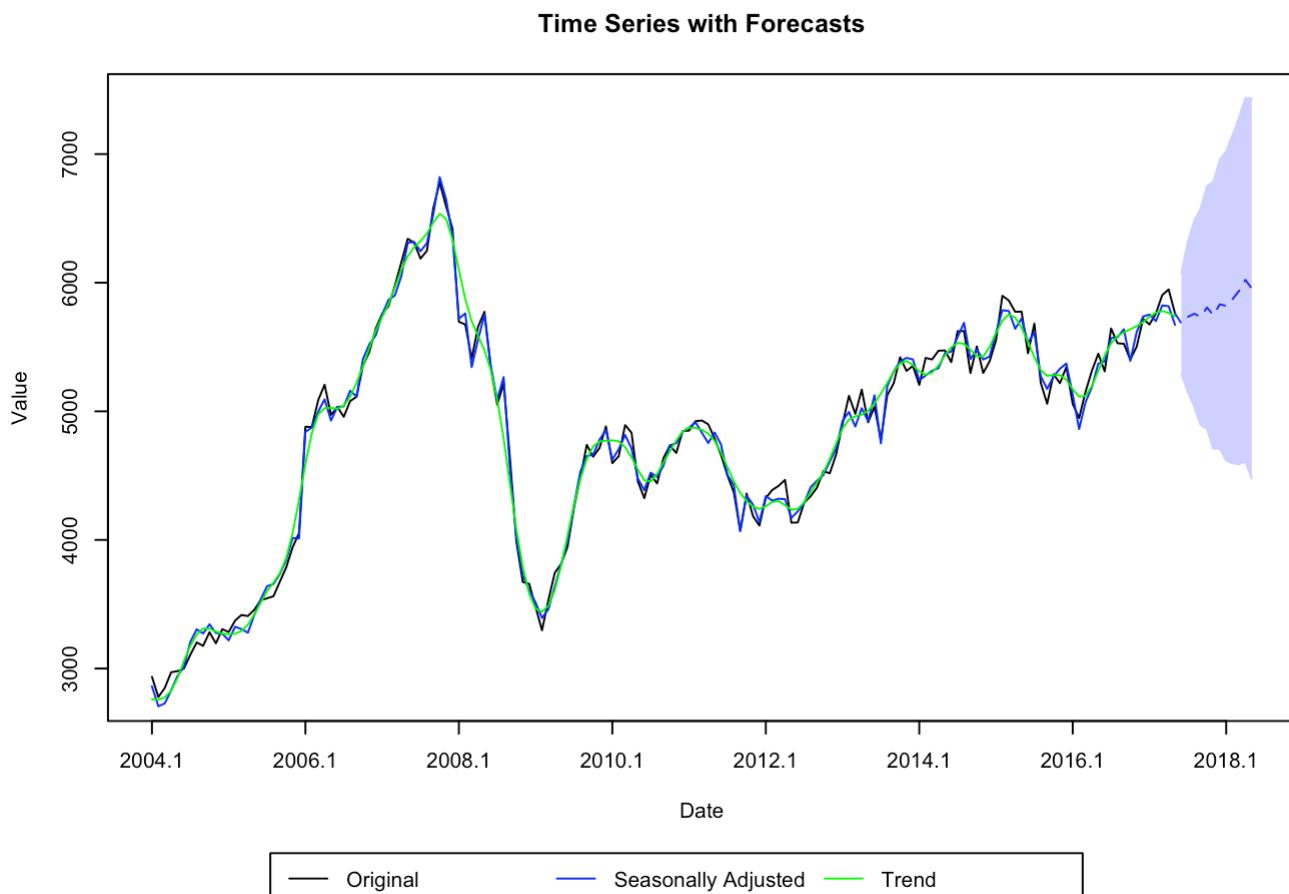
From the ADF test, it confirms that the series is now stationary at 5% level of significance after first differencing as $p < 0.05$.

The stationarity of the series are further confirmed through the PP tests, where the series are stationary at 5% level of significance as $p < 0.05$.

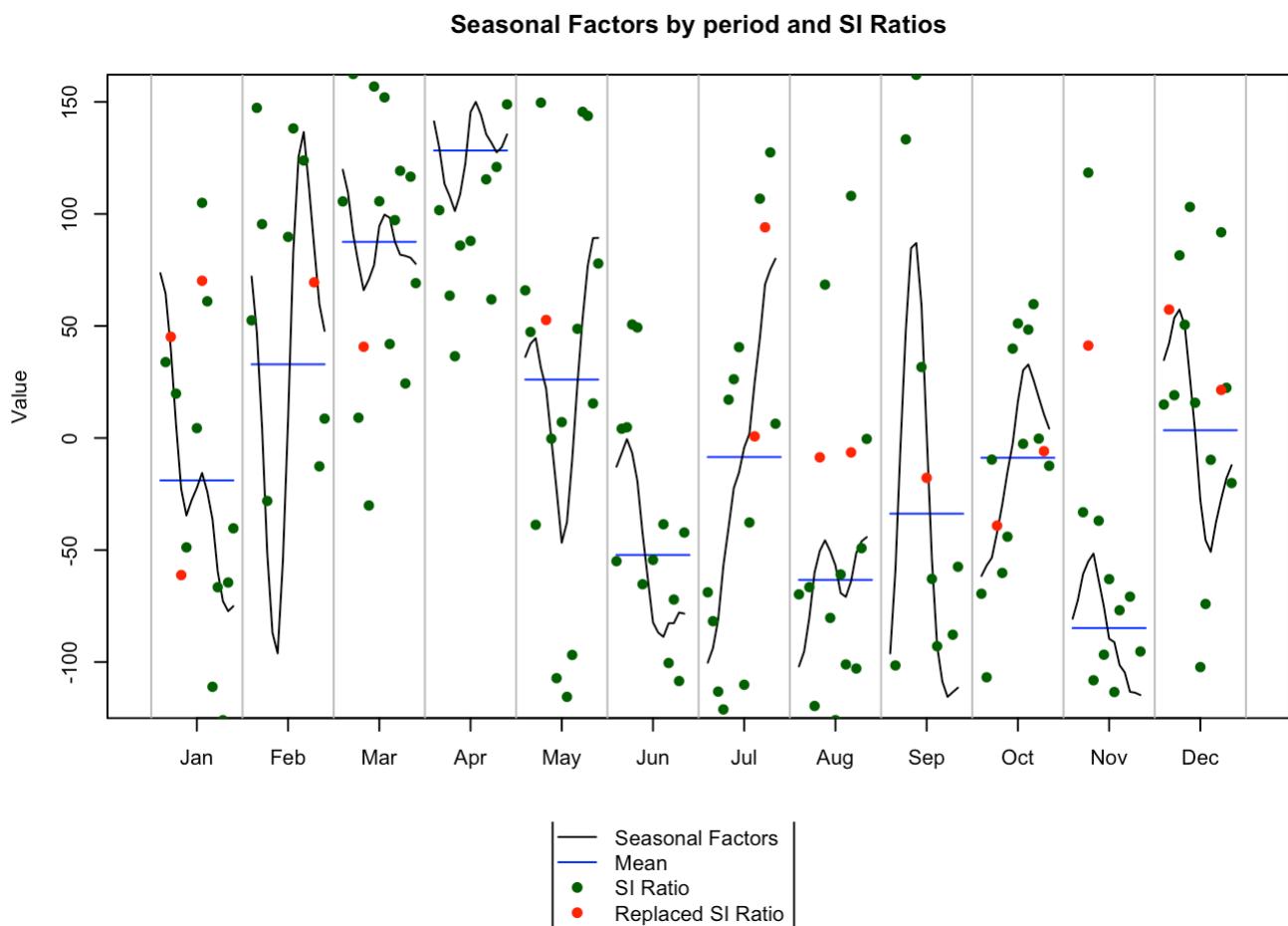
Decomposition

ASX Price Series

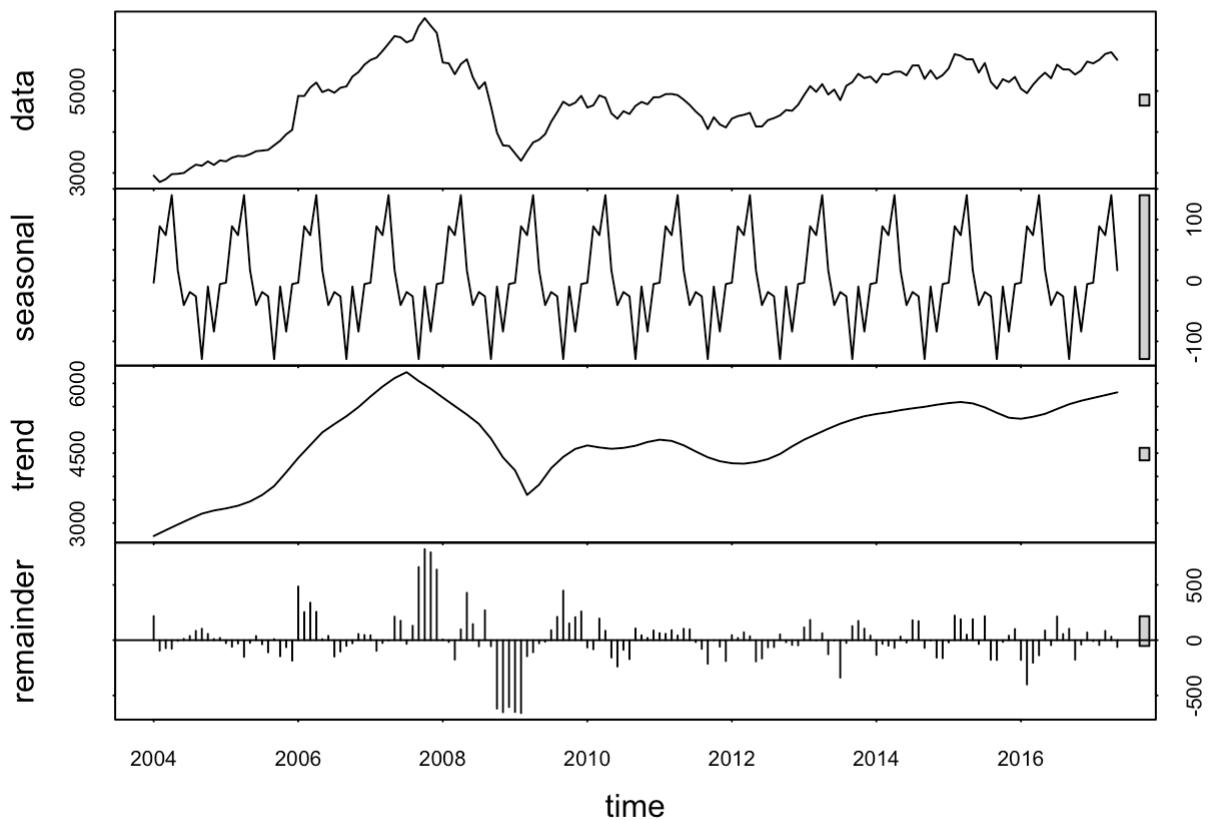
```
asx_decomp = x12(asx_ts)  
plot(asx_decomp, sa = TRUE, trend = TRUE, forecast = TRUE)
```



```
plotSeasFac(asx_decomp)
```



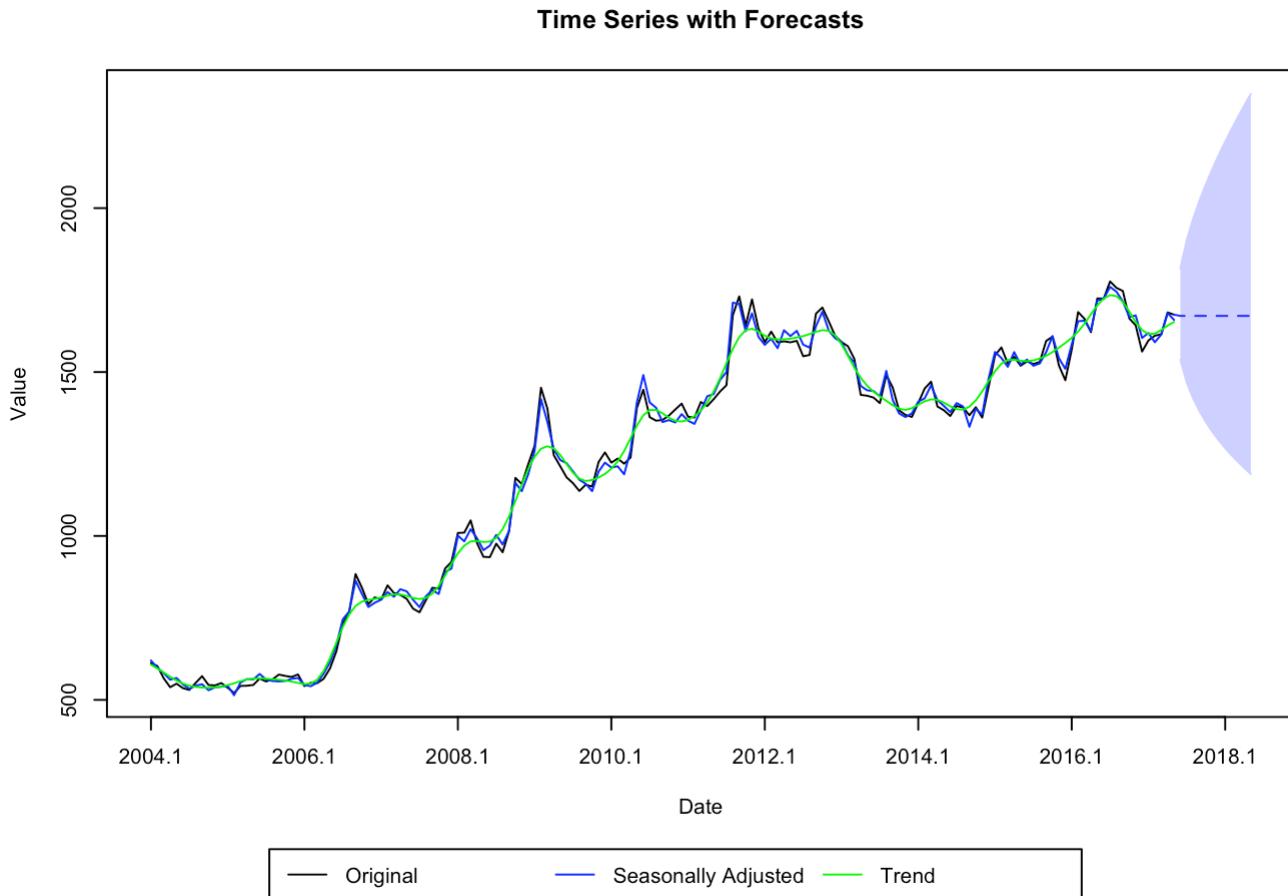
```
fit.asx <- stl(asx_ts, t.window = 15, s.window = "periodic", robust = TRUE)
plot(fit.asx)
```



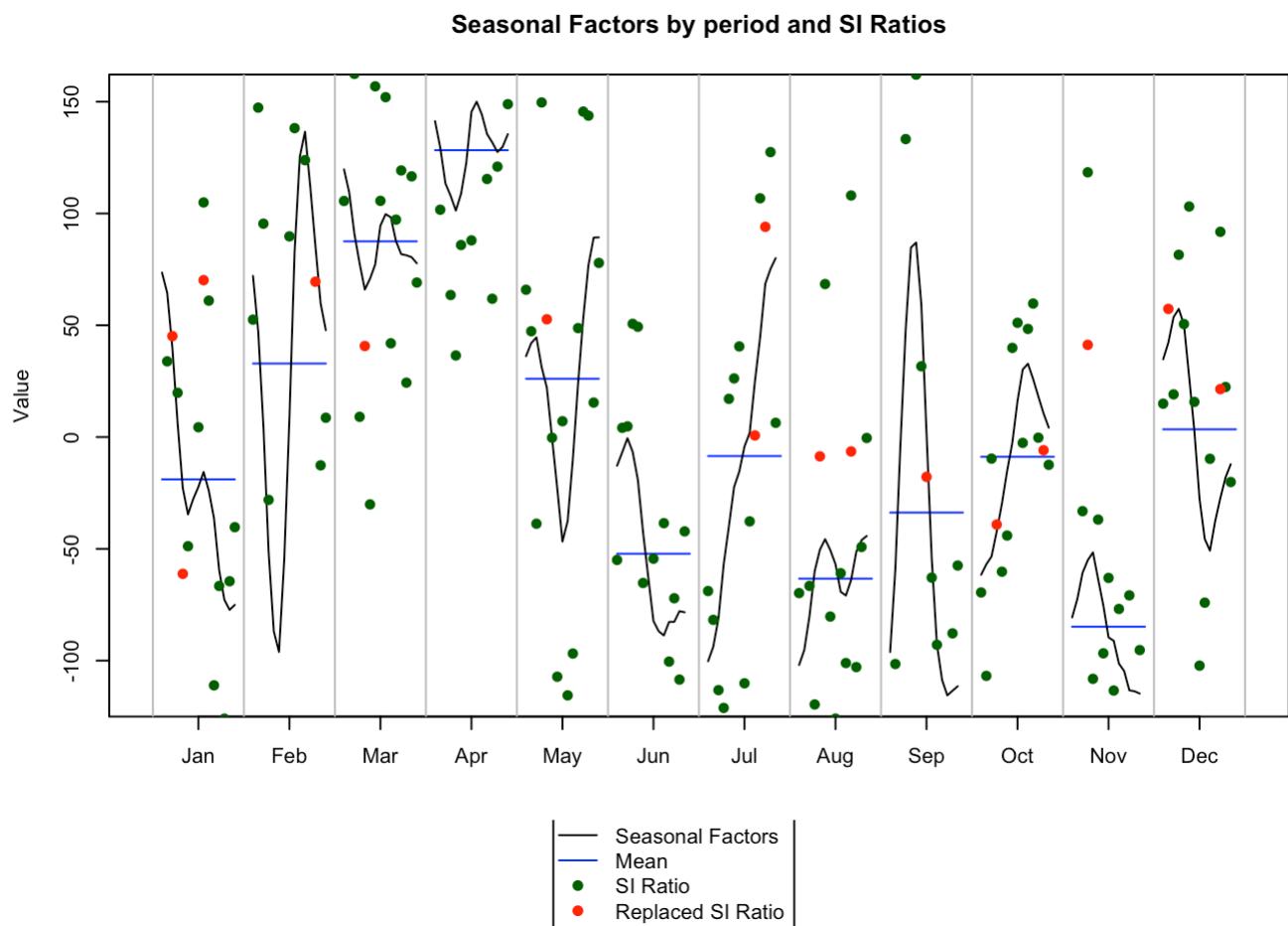
- From the x-12 decomposition, it is observed that the pattern of the seasonally adjusted series follows the original series closely, however there are slight deviations of a seasonal effect. There is also minor deviations from the original series, which displays that it has an effect on the series. Furthermore, the forecast reveals that the seasonally adjusted series follows the original series
- From the seasonal factors plot, the seasonal factors are diverting from the mean level from January to December
- The STL decomposition plot displays a trend that follows the overall pattern of the original series - that is, an upward trend, followed by a decline and a gradual upward trend again.

Gold Price Series

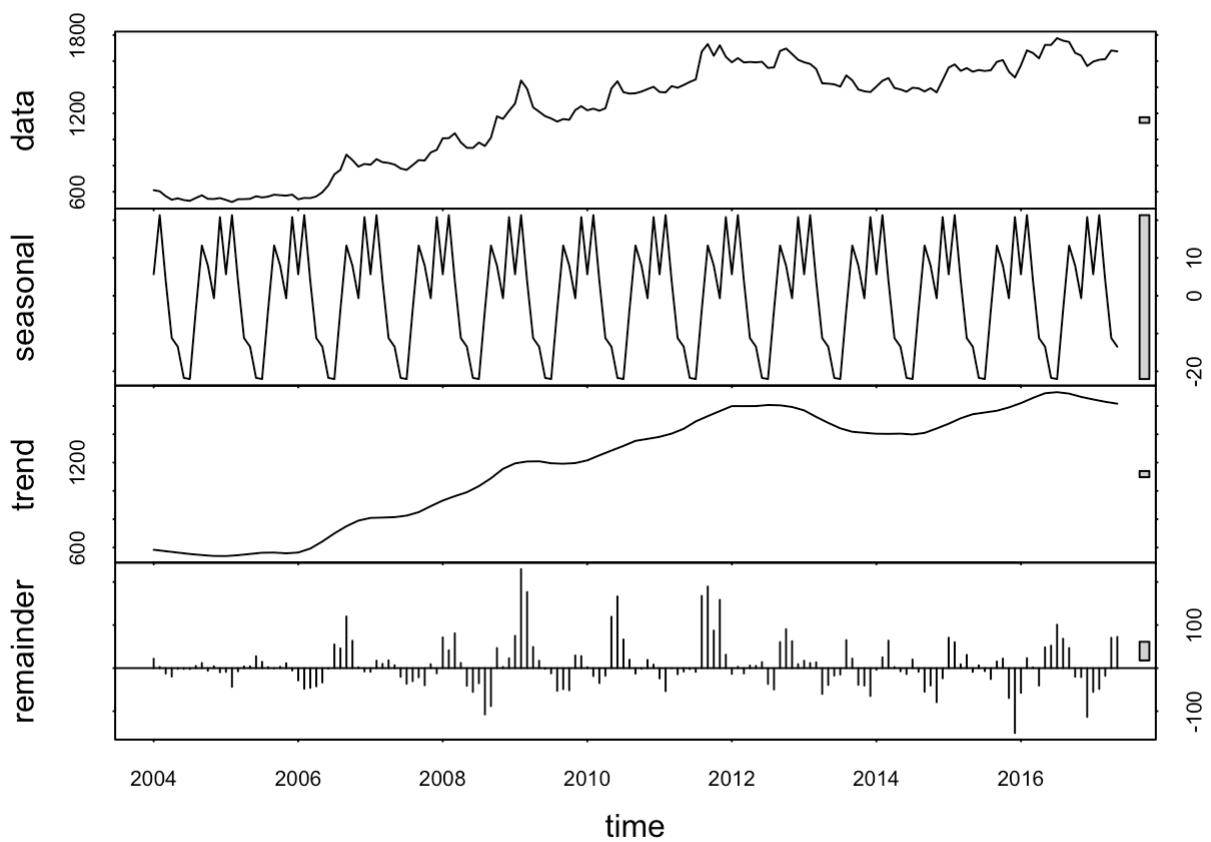
```
gold_decomp = x12(gold_ts)
plot(gold_decomp, sa = TRUE, trend = TRUE, forecast = TRUE)
```



```
plotSeasFac(asx_decomp)
```



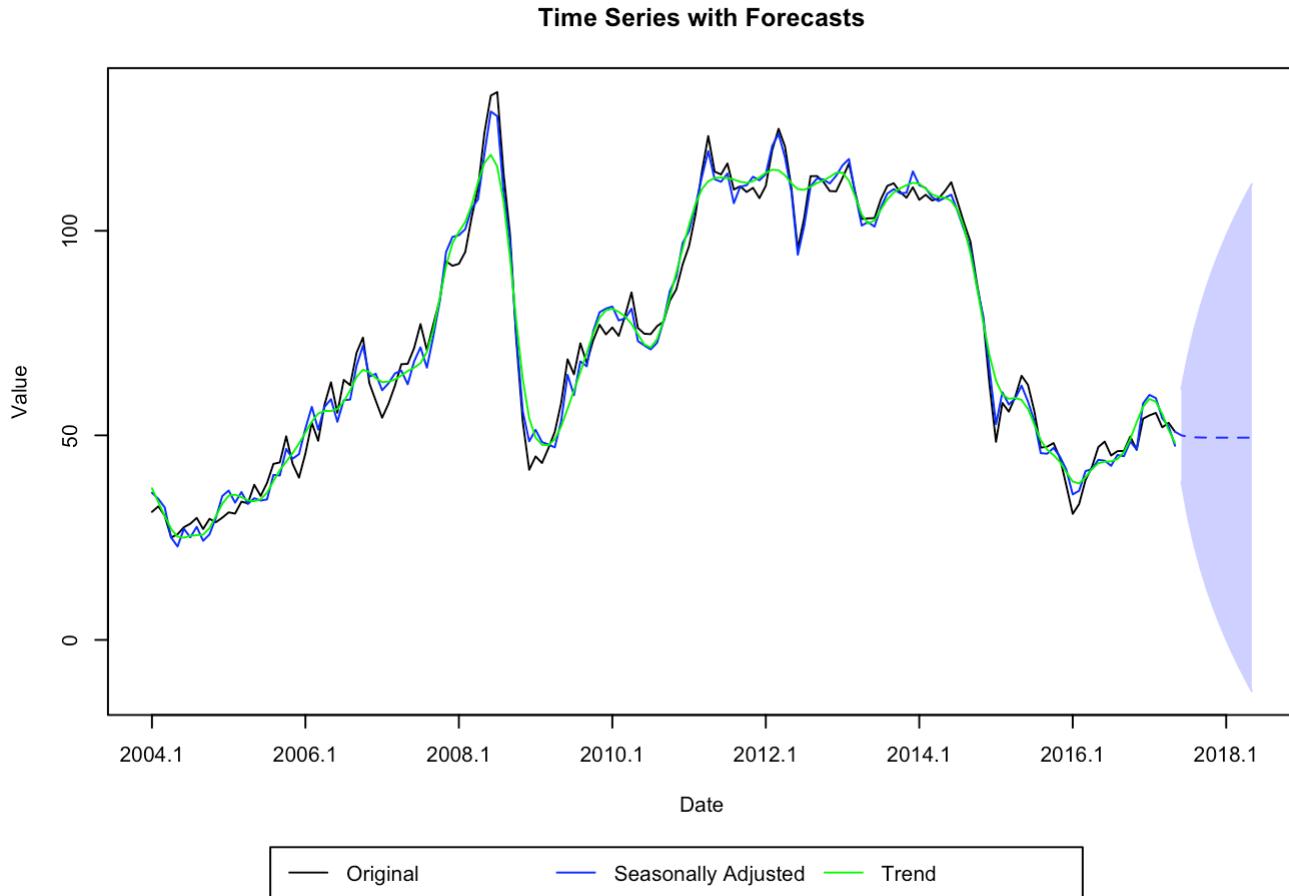
```
fit.gold <- stl(gold_ts, t.window = 15, s.window = "periodic", robust = TRUE)
plot(fit.gold)
```



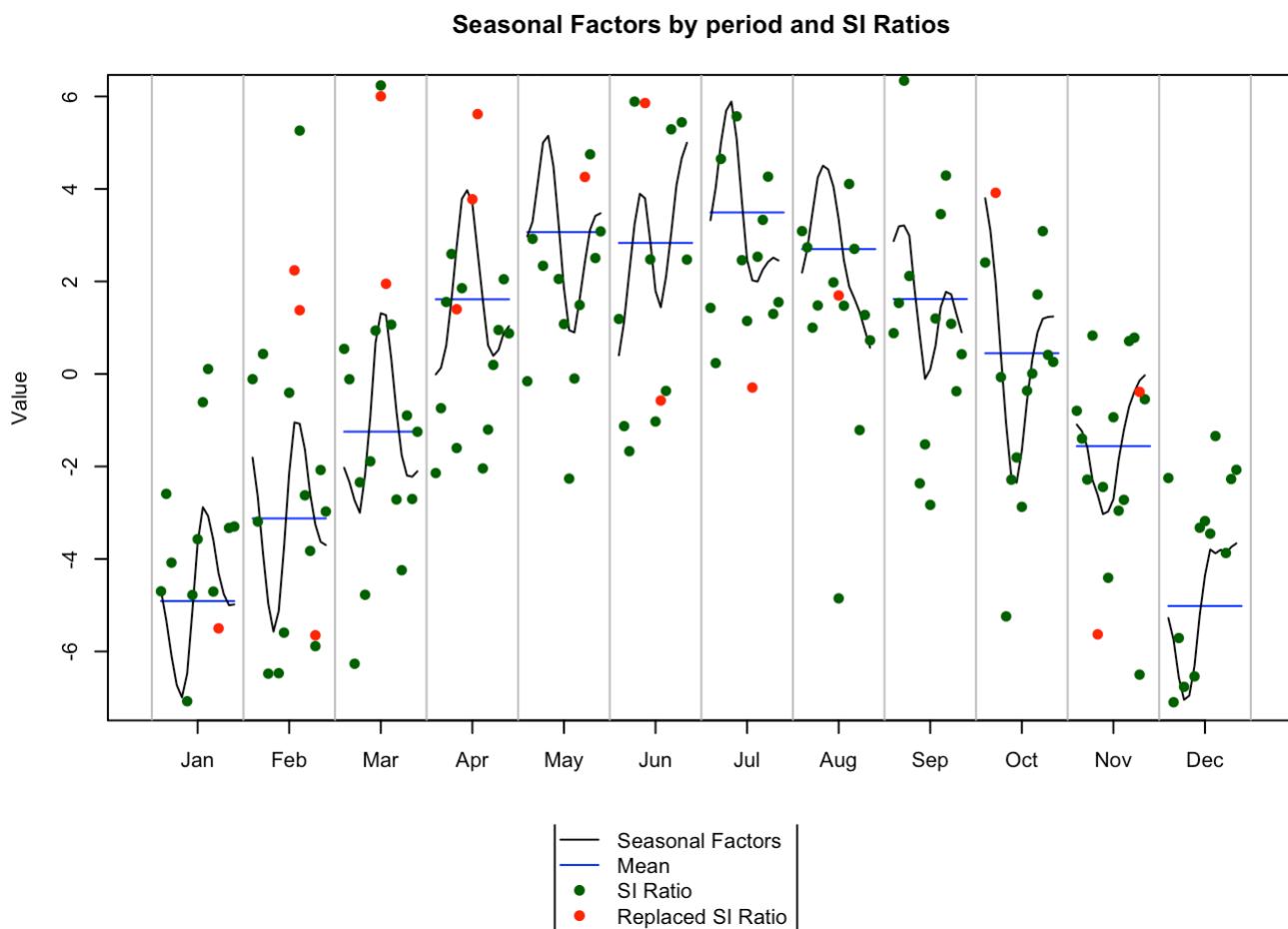
- From the x-12 decomposition, it is observed that the pattern of the seasonally adjusted series follows the original series closely, however there are slight deviations. Furthermore, the forecast reveals that the series is to remain constant in the future years
- From the seasonal factors plot, the seasonal factors are diverting from the mean level from January to December
- The STL decomposition plot displays a trend that follows the overall pattern of the original series - that is, an upward trend

Crude Oil Series

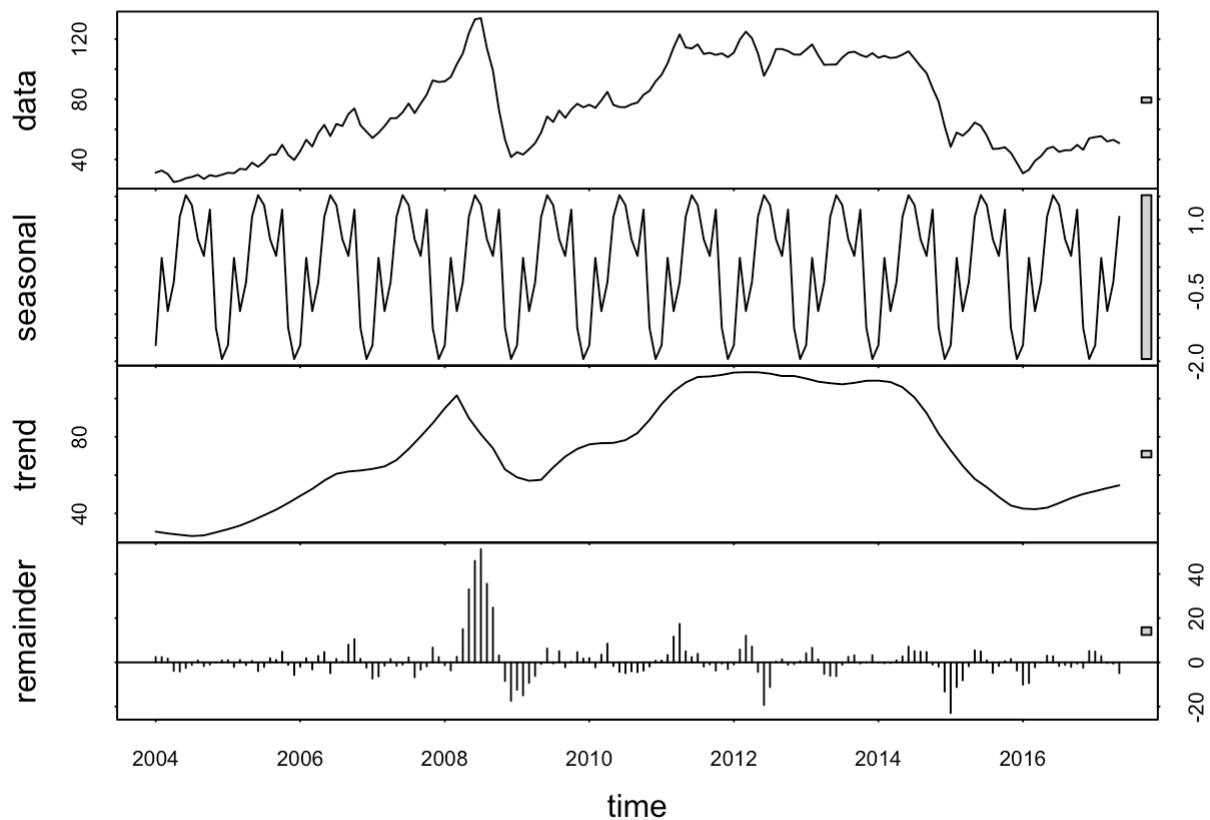
```
crude_decomp = x12(crude_ts)
plot(crude_decomp, sa = TRUE, trend = TRUE, forecast = TRUE)
```



```
plotSeasFac(crude_decomp)
```



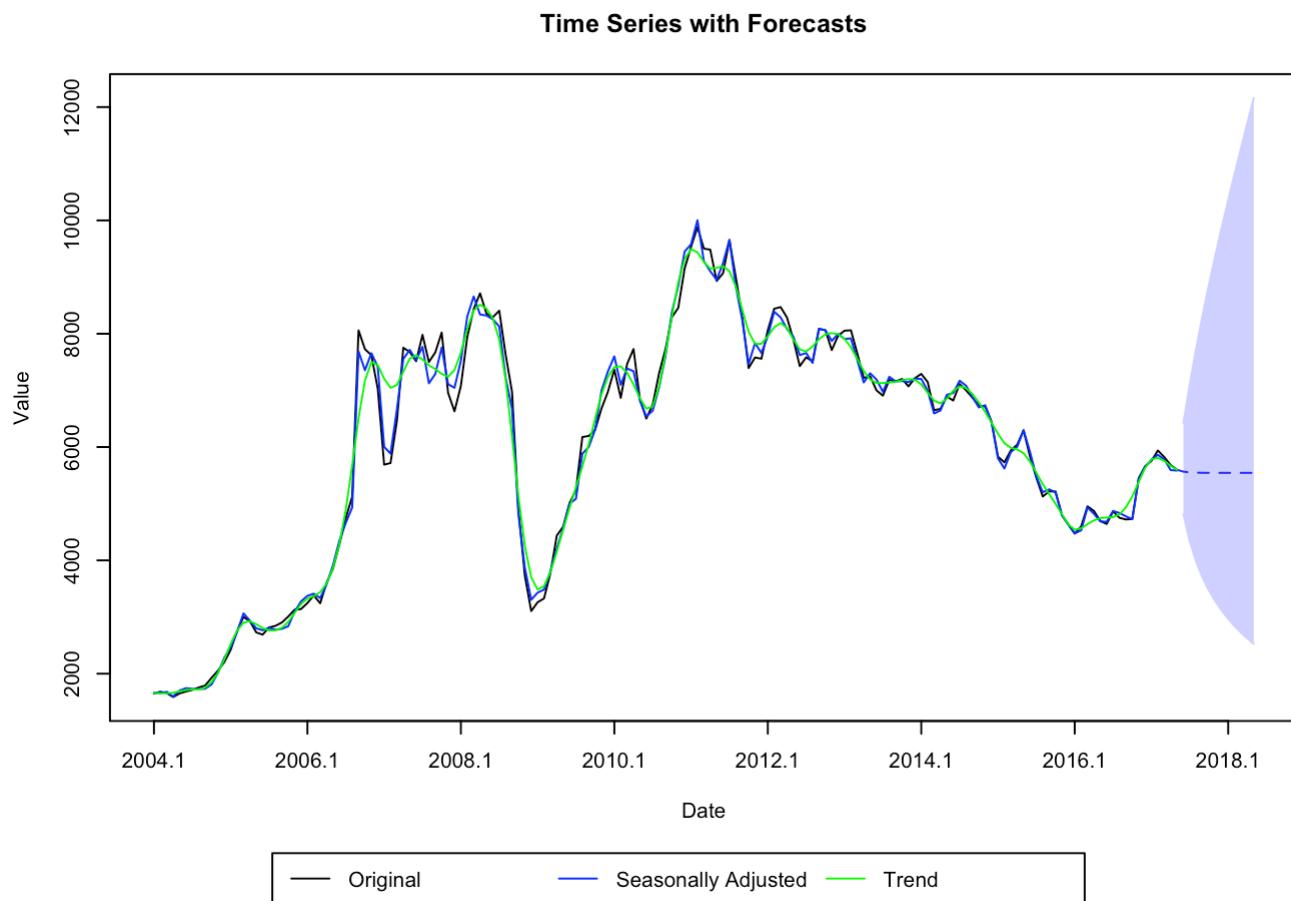
```
fit.crude <- stl(crude_ts, t.window = 15, s.window = "periodic", robust = TRUE)
plot(fit.crude)
```



- From the x-12 decomposition, it is observed that the pattern of the seasonally adjusted series follows the original series closely from 2004-2006. After 2006, there are significant deviations from the original series from 2006-2011. However, after 2011, the series follow the original series and the forecast reveals that the series is to remain constant in the future years
- From the seasonal factors plot, the seasonal factors are diverting from the mean level from January to December
- The STL decomposition plot displays a trend that follows the overall pattern of the original series. The intervention points are also clearly depicted in 2008-2010. Furthermore, the plot shows variation in seasonality.

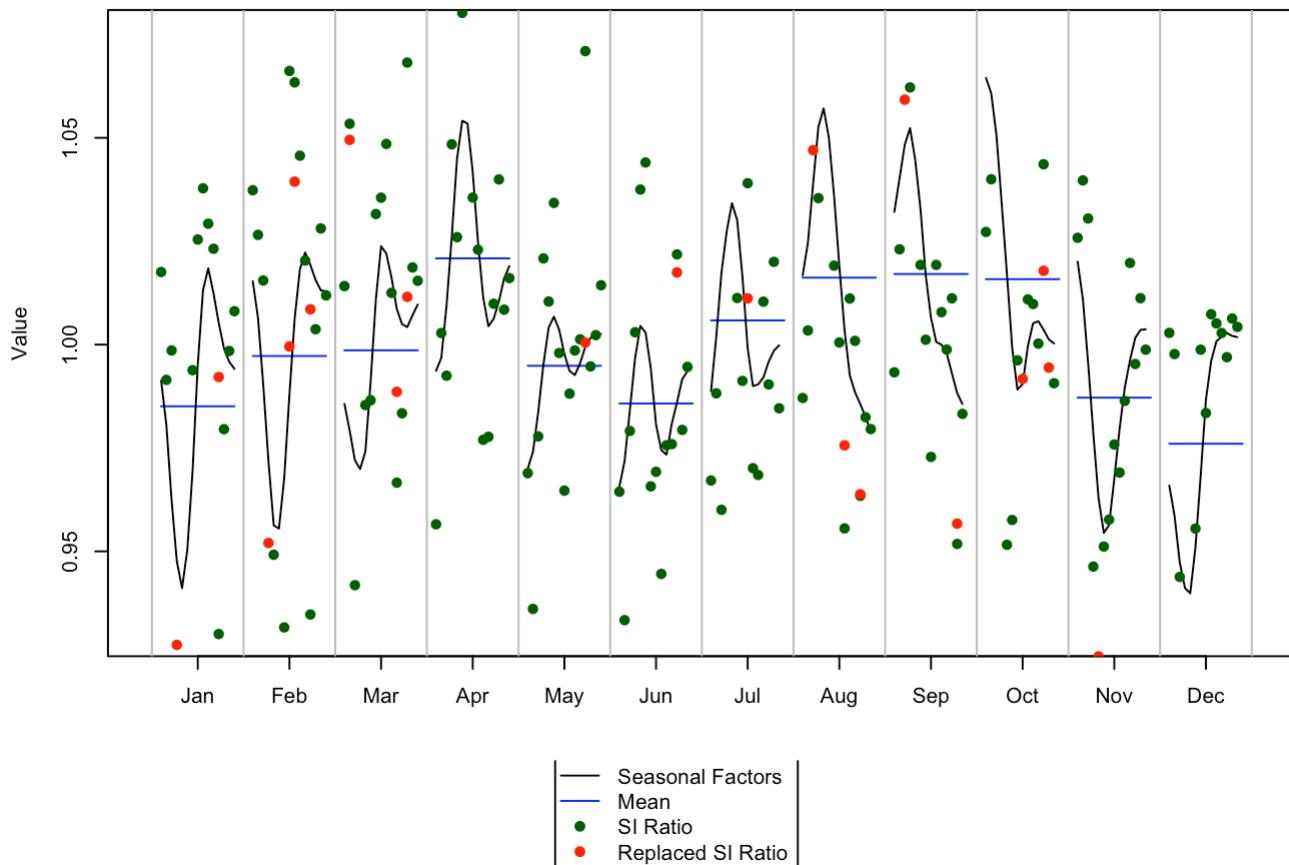
Copper Price Series

```
copper_decomp = x12(copper_ts)
plot(copper_decomp, sa = TRUE, trend = TRUE, forecast = TRUE)
```

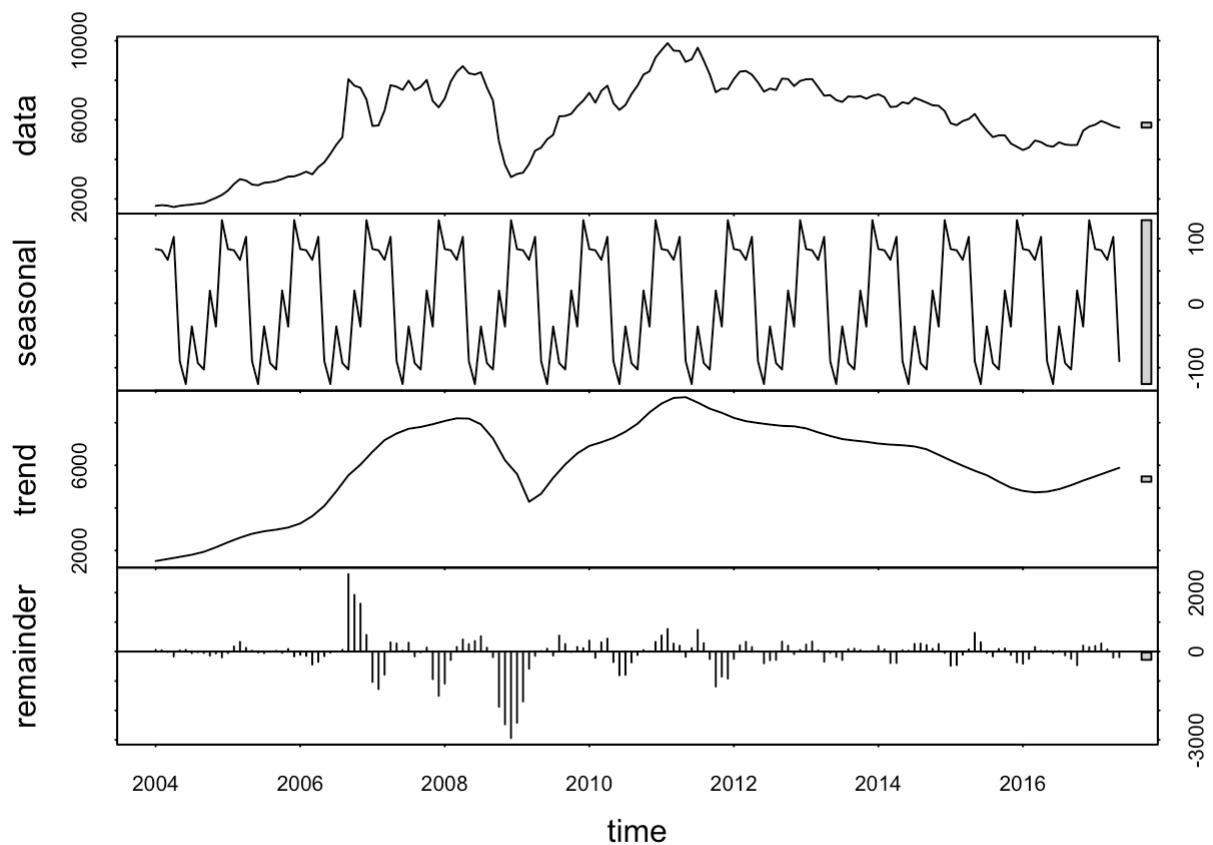


```
plotSeasFac(copper_decomp)
```

Seasonal Factors by period and SI Ratios



```
fit.copper <- stl(copper_ts, t.window = 15, s.window = "periodic", robust = TRUE)
plot(fit.copper)
```



- From the x-12 decomposition, it is observed that the pattern of the seasonally adjusted series follows the original series closely throughout. The forecast reveals that the series is to remain constant in the future years.
- As there is not much change in the seasonality, the seasonal factors are surrounding the mean.
- The STL decomposition plot displays a trend that follows the overall pattern of the original series. The intervention points are also clearly depicted in 2007-2009.

Model Fitting

```
asx_bind <- cbind(asx_ts, gold_ts, crude_ts, copper_ts)
cor(asx_bind)
```

```
##           asx_ts   gold_ts  crude_ts copper_ts
## asx_ts     1.000000 0.3431908 0.3290338 0.5617864
## gold_ts    0.3431908 1.0000000 0.4366382 0.5364213
## crude_ts   0.3290338 0.4366382 1.0000000 0.8664296
## copper_ts  0.5617864 0.5364213 0.8664296 1.0000000
```

A correlation test is performed to evaluate the correlation between the variables. As the ASX price index is considered the dependent variable, Y, the other variables, gold price, crude oil price and copper price, will be the independent variables, X. From the correlation test, ASX price index and copper show the strongest correlation (0.5618) compared to the other variables. The variable with the least correlation to ASX price index is crude oil price (0.3290).

In order to determine the most suitable model for the series, multiple lag models such as finite, polynomial, koyk and auto-regressive distributed lag model, were performed and evaluated for all independent variables (gold price, crude oil price and copper price) to dependent variable (ASX price index).

ASX Price Index and Gold Prices

```
#Finite lag model - ASX & Gold
model1_gold = dlm( x = as.vector(gold_ts), y = as.vector(asx_ts), q = 9)
summary(model1_gold)
```

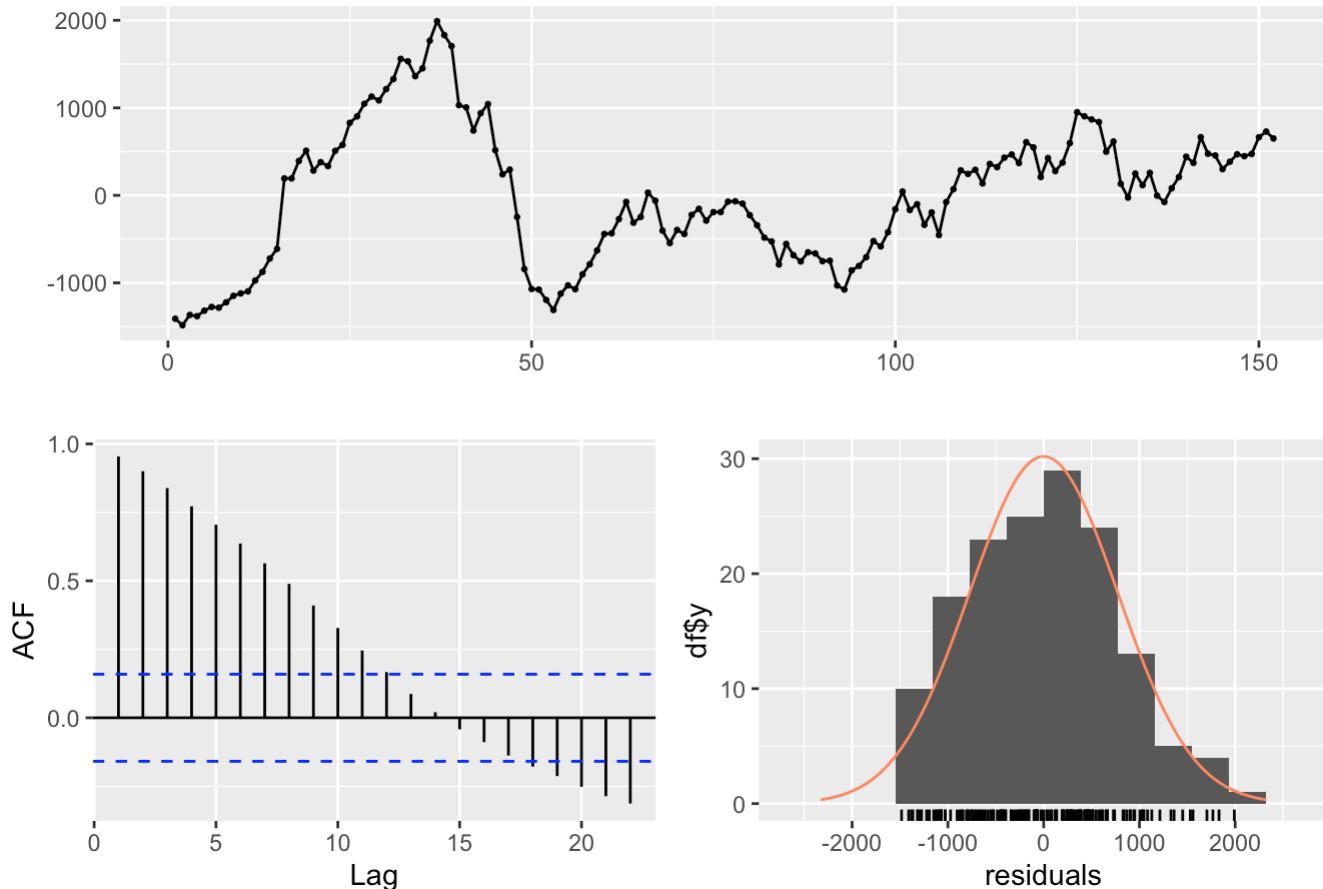
```

## 
## Call:
## lm(formula = model.formula, data = design)
##
## Residuals:
##    Min      1Q  Median      3Q     Max
## -1484.01 -590.23   13.97  473.71 1991.28
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4452.20100  223.27538 19.940 <2e-16 ***
## x.t        -0.35288   1.27113 -0.278  0.782
## x.1        -0.10591   1.88058 -0.056  0.955
## x.2         0.02149   1.91708  0.011  0.991
## x.3        -0.07972   1.93268 -0.041  0.967
## x.4        -0.25836   1.93610 -0.133  0.894
## x.5         0.36953   1.93817  0.191  0.849
## x.6         0.16311   1.95298  0.084  0.934
## x.7         0.70404   1.94616  0.362  0.718
## x.8        -0.21513   1.91991 -0.112  0.911
## x.9         0.16715   1.29373  0.129  0.897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 802.4 on 141 degrees of freedom
## Multiple R-squared:  0.05795,   Adjusted R-squared: -0.008866
## F-statistic: 0.8673 on 10 and 141 DF,  p-value: 0.5654
##
## AIC and BIC values for the model:
##       AIC     BIC
## 1 2476.983 2513.27

```

```
checkresiduals(model1_gold$model)
```

Residuals



```
##  
## Breusch-Godfrey test for serial correlation of order up to 14  
##  
## data: Residuals  
## LM test = 140.36, df = 14, p-value < 2.2e-16
```

```
# Finite lag length based on AIC and BIC values  
for (i in 1:10) {  
  model1.gold = dlm(x = as.vector(gold_ts), y = as.vector(asx_ts), q = i)  
  cat("q = ", i, "AIC = ", AIC(model1.gold$model), "BIC = ", BIC(model1.gold$mo  
del), "\n")  
}
```

```
## q = 1 AIC = 2613.609 BIC = 2625.91  
## q = 2 AIC = 2596.292 BIC = 2611.637  
## q = 3 AIC = 2579.215 BIC = 2597.59  
## q = 4 AIC = 2562.296 BIC = 2583.69  
## q = 5 AIC = 2544.887 BIC = 2569.286  
## q = 6 AIC = 2527.575 BIC = 2554.966  
## q = 7 AIC = 2510.535 BIC = 2540.905  
## q = 8 AIC = 2493.885 BIC = 2527.22  
## q = 9 AIC = 2476.983 BIC = 2513.27  
## q = 10 AIC = 2460.345 BIC = 2499.57
```

From the finite lag model for ASX price index and gold price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is -0.008 which indicates that additional predictors will not improve the model.

Furthermore, the model is insignificant at 5% level of significance with a p-value of 0.5654. For the Breusch-Godfrey test, the null hypothesis and alternative hypothesis are as follows:

- Null hypothesis: There is no autocorrelation at some order less than or equal to p
- Alternative hypothesis: There is autocorrelation at some order less than or equal to p

Based on the Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 14. This is further confirmed by the ACF plot. Based on the time series and histogram of standardised residuals, there are very high residual values that violate general assumptions. In addition, both the AIC and BIC values suggest the use of lag length 10 as it has the lowest AIC and BIC values. From these results, it shows that the model does not fit the data well.

```
#Polynomial lag model - Gold & ASX
model2_gold = polyDlm(x = as.vector(gold_ts), y = as.vector(asx_ts), q = 2, k = 2, show.beta = TRUE)
```

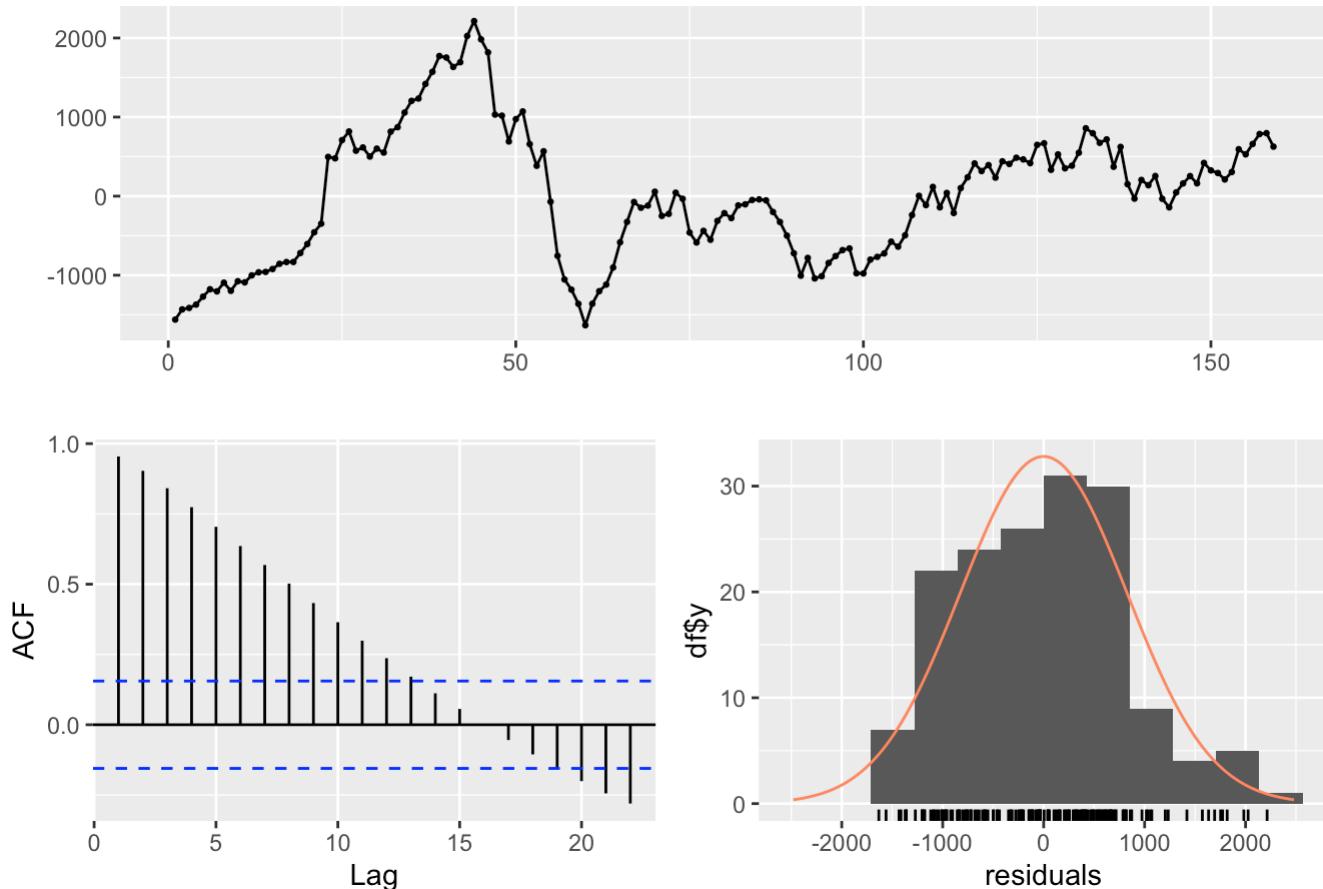
```
## Estimates and t-tests for beta coefficients:
##           Estimate Std. Error t value P(>|t|)
## beta.0     0.396      1.28   0.310  0.757
## beta.1    -0.233      1.90  -0.122  0.903
## beta.2     0.536      1.27   0.421  0.675
```

```
summary(model2_gold)
```

```
##
## Call:
## "Y ~ (Intercept) + X.t"
##
## Residuals:
##      Min       10      Median       30       Max
## -1632.50  -700.82      4.61    549.72  2213.87
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3998.3587  212.3161 18.832 <2e-16 ***
## z.t0        0.3958    1.2767   0.310  0.757
## z.t1       -1.3268    5.7723  -0.230  0.819
## z.t2        0.6983    2.8546   0.245  0.807
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 834.5 on 155 degrees of freedom
## Multiple R-squared:  0.1015, Adjusted R-squared:  0.08409
## F-statistic: 5.835 on 3 and 155 DF,  p-value: 0.0008385
```

```
checkresiduals(model2_gold$model)
```

Residuals



```
## 
## Breusch-Godfrey test for serial correlation of order up to 10
## 
## data: Residuals
## LM test = 147.24, df = 10, p-value < 2.2e-16
```

From the polynomial lag model for ASX price index and gold price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is 0.084 which indicates that additional predictors will not improve the model. However, the model is significant at 5% level of significance with a p-value of 0.00083. Based on the Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot. Based on the time series and histogram of standardised residuals, there are very high residual values that violate general assumptions and observe non-random patterns observed. From these results, it shows that the model does not fit the data well.

```
#Koyk lag model - Gold & ASX
model3_gold = koyckDlm(x = as.vector(gold_ts), y = as.vector(asx_ts))
summary(model3_gold)
```

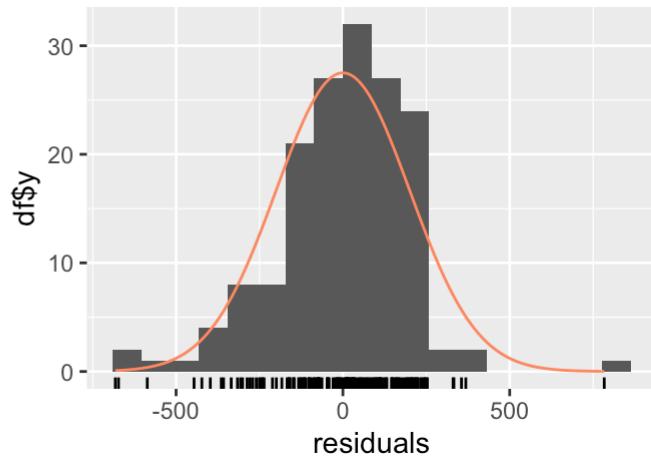
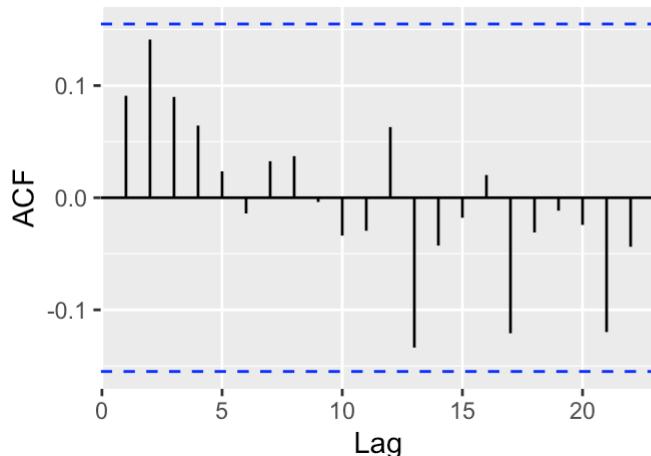
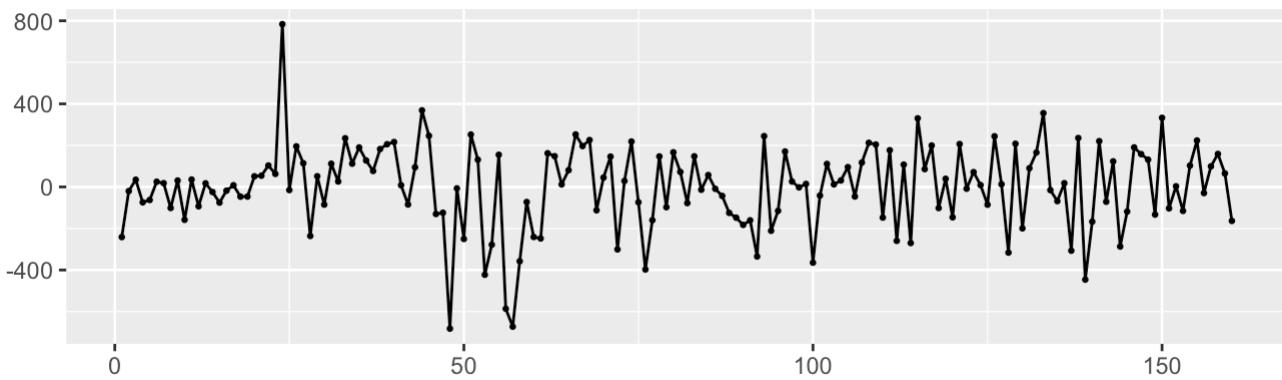
```

## 
## Call:
## "Y ~ (Intercept) + Y.1 + X.t"
## 
## Residuals:
##    Min     10   Median     30      Max 
## -682.19 -105.44    15.86  135.04  783.60 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 1.902e+02  8.958e+01   2.123   0.0353 *  
## Y.1         9.635e-01  1.909e-02  50.469  <2e-16 ***  
## X.t        2.595e-03  4.304e-02   0.060   0.9520    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 201.4 on 157 degrees of freedom
## Multiple R-Squared: 0.9488, Adjusted R-squared: 0.9481 
## Wald test: 1454 on 2 and 157 DF, p-value: < 2.2e-16
## 
## Diagnostic tests:
## NULL
## 
##          alpha       beta       phi  
## Geometric coefficients: 5205.15 0.002595168 0.9634602

```

```
checkresiduals(model3_gold$model)
```

Residuals



```
##  
## Ljung-Box test  
##  
## data: Residuals  
## Q* = 7.3793, df = 10, p-value = 0.6892  
##  
## Model df: 0. Total lags used: 10
```

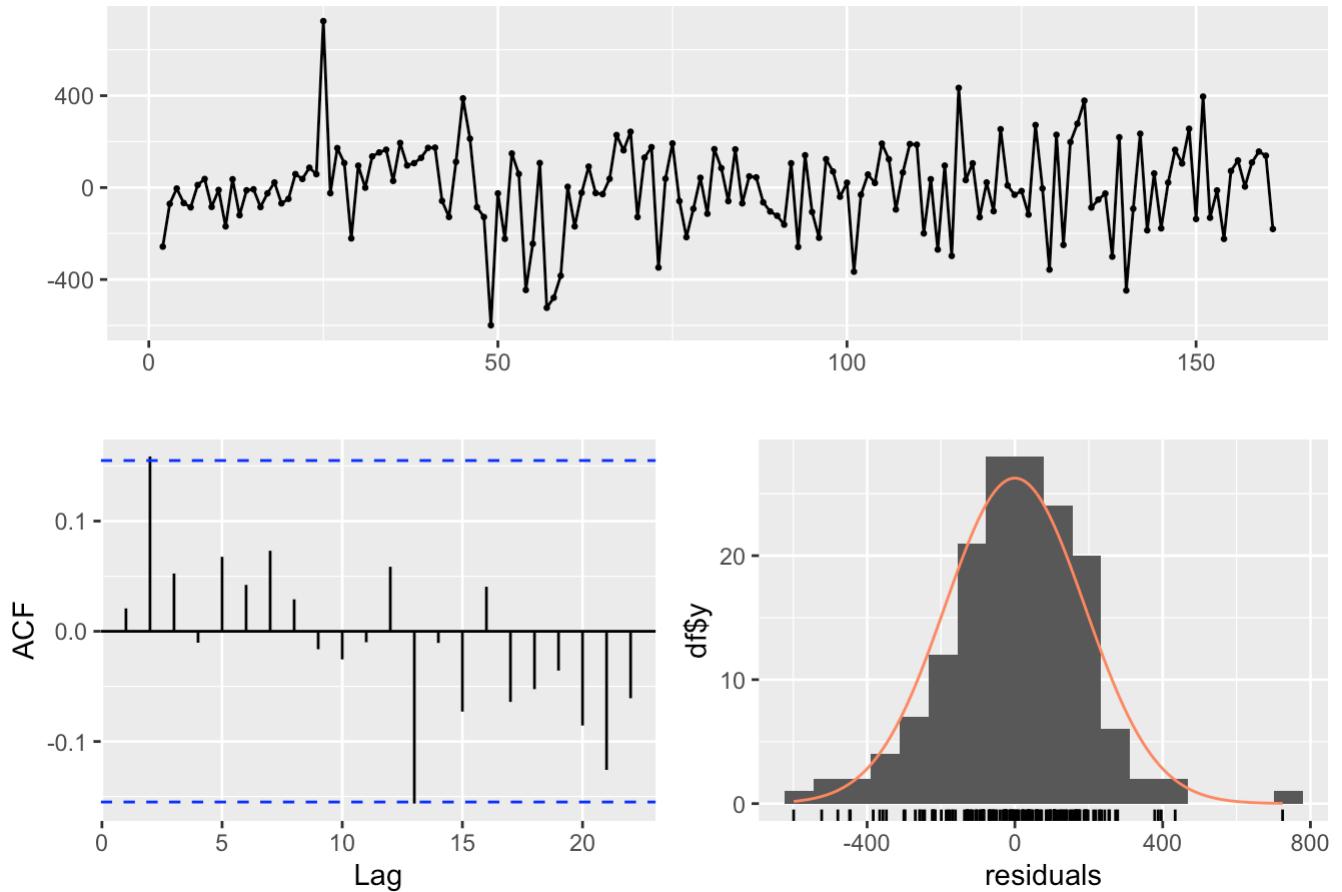
From the koyck lag model for ASX price index and gold price, some of the coefficients is significant ($p < 0.05$) and the adjusted R-square value is 0.9481 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. The ACF plot shows that there are no significant correlations as all the lags are well within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays a left-skewed distributed. From these results, it shows that the model does fit the data well.

```
#Autoregressive distributed lag model - Gold & ASX  
model4_gold = ardlDlm(x = as.vector(gold_ts), y = as.vector(asx_ts))  
summary(model4_gold)
```

```
##  
## Time series regression with "ts" data:  
## Start = 2, End = 161  
##  
## Call:  
## dynlm(formula = as.formula(model.text), data = data, start = 1)  
##  
## Residuals:  
##      Min    1Q Median    3Q   Max  
## -598.9 -102.9   10.4  119.5  724.6  
##  
## Coefficients:  
##             Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 182.80452  85.05812  2.149  0.0332 *  
## X.t         -1.24911   0.29162 -4.283 3.21e-05 ***  
## X.1          1.23169   0.28976  4.251 3.66e-05 ***  
## Y.1          0.97172   0.01812 53.624 < 2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 191.1 on 156 degrees of freedom  
## Multiple R-squared:  0.9542, Adjusted R-squared:  0.9533  
## F-statistic: 1083 on 3 and 156 DF,  p-value: < 2.2e-16
```

```
checkresiduals(model4_gold$model)
```

Residuals



```
##  
## Breusch-Godfrey test for serial correlation of order up to 10  
##  
## data: Residuals  
## LM test = 7.9182, df = 10, p-value = 0.6368
```

```
for (i in 1:5) {  
  for (j in 1:5) {  
    model4.gold = ardlDlm(x = as.vector(gold_ts), y = as.vector(asx_ts),  
    p = i, q = j)  
    cat("p = ", i, "q = ", j, "AIC = ", AIC(model4.gold$model), "BIC = ",  
    BIC(model4.gold$model), "\n")  
  }  
}
```

```

## p = 1 q = 1 AIC = 2140.897 BIC = 2156.273
## p = 1 q = 2 AIC = 2128.524 BIC = 2146.938
## p = 1 q = 3 AIC = 2113.99 BIC = 2135.428
## p = 1 q = 4 AIC = 2102.754 BIC = 2127.204
## p = 1 q = 5 AIC = 2092.194 BIC = 2119.643
## p = 2 q = 1 AIC = 2128.627 BIC = 2147.04
## p = 2 q = 2 AIC = 2130.523 BIC = 2152.005
## p = 2 q = 3 AIC = 2115.89 BIC = 2140.39
## p = 2 q = 4 AIC = 2104.694 BIC = 2132.2
## p = 2 q = 5 AIC = 2094.14 BIC = 2124.639
## p = 3 q = 1 AIC = 2118.109 BIC = 2139.547
## p = 3 q = 2 AIC = 2120.027 BIC = 2144.528
## p = 3 q = 3 AIC = 2117.305 BIC = 2144.868
## p = 3 q = 4 AIC = 2105.731 BIC = 2136.293
## p = 3 q = 5 AIC = 2095.264 BIC = 2128.812
## p = 4 q = 1 AIC = 2107.002 BIC = 2131.452
## p = 4 q = 2 AIC = 2108.914 BIC = 2136.42
## p = 4 q = 3 AIC = 2106.276 BIC = 2136.839
## p = 4 q = 4 AIC = 2107.456 BIC = 2141.074
## p = 4 q = 5 AIC = 2097.01 BIC = 2133.608
## p = 5 q = 1 AIC = 2094.908 BIC = 2122.357
## p = 5 q = 2 AIC = 2096.86 BIC = 2127.359
## p = 5 q = 3 AIC = 2094.144 BIC = 2127.692
## p = 5 q = 4 AIC = 2095.425 BIC = 2132.023
## p = 5 q = 5 AIC = 2097.324 BIC = 2136.972

```

From the autoregressive distributed lag model for ASX price index and gold price, all the coefficients are significant ($p < 0.05$) and the adjusted R-square value is 0.9533 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. Based on the Breusch-Godfrey test, it shows that as the p-value is greater than 0.05, we cannot reject the null hypothesis and conclude that autocorrelation does not exist among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot which shows that there are no significant correlations as all the lags are within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays an approximate normal distribution. In addition, $p = 1$ and $q = 5$ gives the smallest AIC and BIC values. From these results, it shows that the model does fit the data well.

ASX Price Index & Crude Oil Price

```

#Finite Lag Model
model1_crude = dlm(x = as.vector(crude_ts), y = as.vector(asx_ts), q = 9)
summary(model1_crude)

```

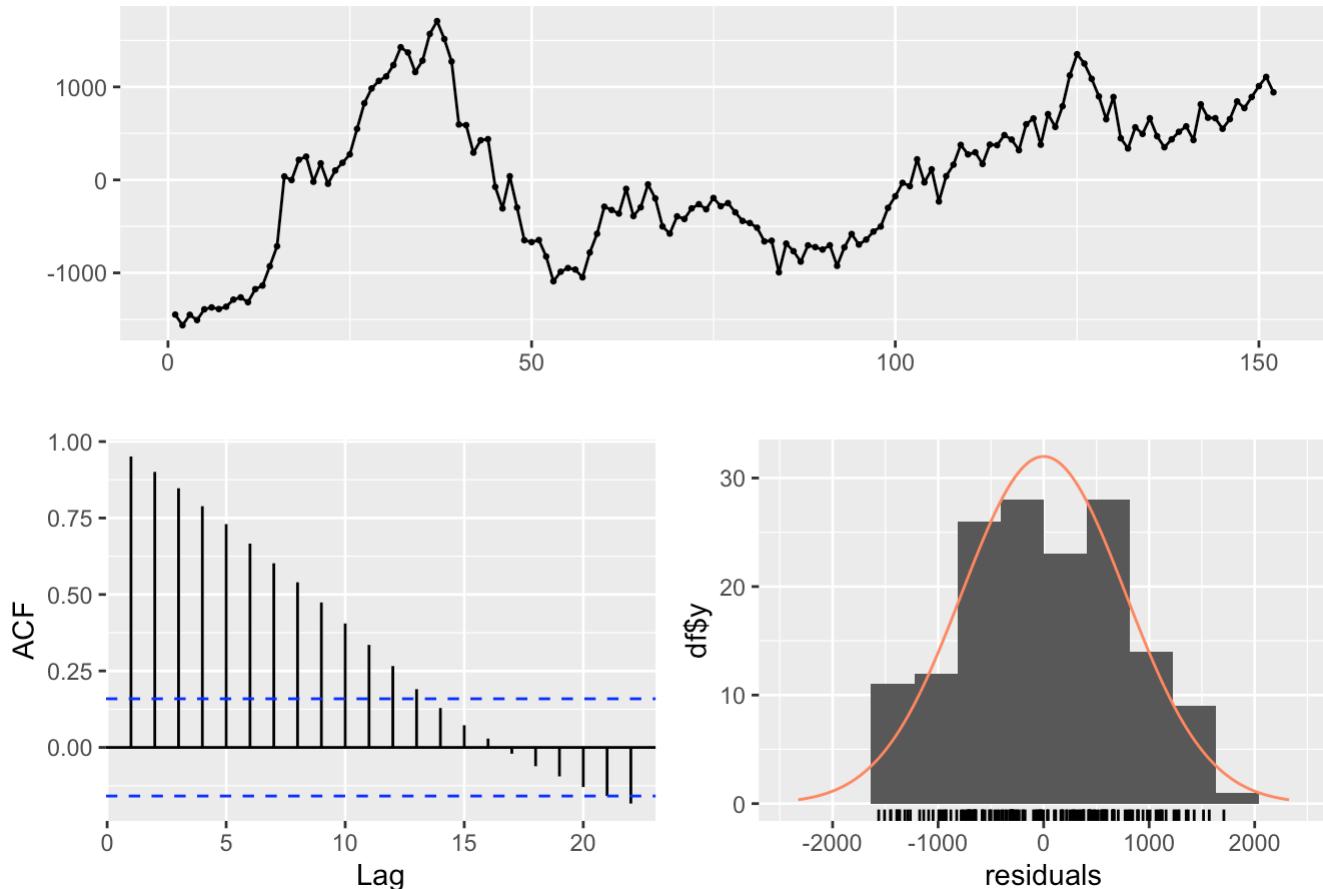
```

## 
## Call:
## lm(formula = model.formula, data = design)
##
## Residuals:
##    Min      1Q  Median      3Q     Max
## -1563.08 -643.18 -11.17  571.05 1707.84
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4641.7675   196.4286  23.631 <2e-16 ***
## x.t          8.8037    11.3647   0.775   0.440    
## x.1          2.8911    19.0087   0.152   0.879    
## x.2         -2.1607    19.2799  -0.112   0.911    
## x.3          2.9569    19.4045   0.152   0.879    
## x.4         -7.0213    19.3913  -0.362   0.718    
## x.5          0.7784    19.3589   0.040   0.968    
## x.6          1.7254    19.4044   0.089   0.929    
## x.7         -2.5268    19.5635  -0.129   0.897    
## x.8         -2.7015    19.3336  -0.140   0.889    
## x.9          0.8312    11.3876   0.073   0.942    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 802.1 on 141 degrees of freedom
## Multiple R-squared:  0.05864,   Adjusted R-squared: -0.008123
## F-statistic: 0.8783 on 10 and 141 DF,  p-value: 0.5551
##
## AIC and BIC values for the model:
##       AIC      BIC
## 1 2476.871 2513.158

```

```
checkresiduals(model1_crude$model)
```

Residuals



```
## 
## Breusch-Godfrey test for serial correlation of order up to 14
## 
## data: Residuals
## LM test = 140.06, df = 14, p-value < 2.2e-16
```

```
# Finite lag length based on AIC and BIC values
for (i in 1:10) {
  model1.crude = dlm(x = as.vector(crude_ts), y = as.vector(asx_ts), q = i)
  cat("q = ", i, "AIC = ", AIC(model1.crude$model), "BIC = ", BIC(model1.crude
$model), "\n")
}
```

```
## q = 1 AIC = 2614.698 BIC = 2626.998
## q = 2 AIC = 2596.715 BIC = 2612.059
## q = 3 AIC = 2579.101 BIC = 2597.477
## q = 4 AIC = 2561.888 BIC = 2583.281
## q = 5 AIC = 2544.936 BIC = 2569.335
## q = 6 AIC = 2527.701 BIC = 2555.091
## q = 7 AIC = 2510.754 BIC = 2541.124
## q = 8 AIC = 2493.914 BIC = 2527.249
## q = 9 AIC = 2476.871 BIC = 2513.158
## q = 10 AIC = 2459.842 BIC = 2499.066
```

From the finite lag model for ASX price index and crude oil price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is -0.008 which indicates that additional predictors will not improve the model. Furthermore, the model is insignificant at 5% level of significance with a p-value of 0.5551. Based on

the Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 14. This is further confirmed by the ACF plot. Based on the time series and histogram of standardised residuals, there are very high residual values that violate general assumptions. In addition, both the AIC and BIC values suggest the use of lag length 10 as it has the lowest AIC and BIC values. From these results, it shows that the model does not fit the data well.

```
#Polynomial Lag Model
model2_crude = polyDlm(x = as.vector(crude_ts) , y = as.vector(asx_ts) , q = 2 , k =
2 , show.beta = TRUE)
```

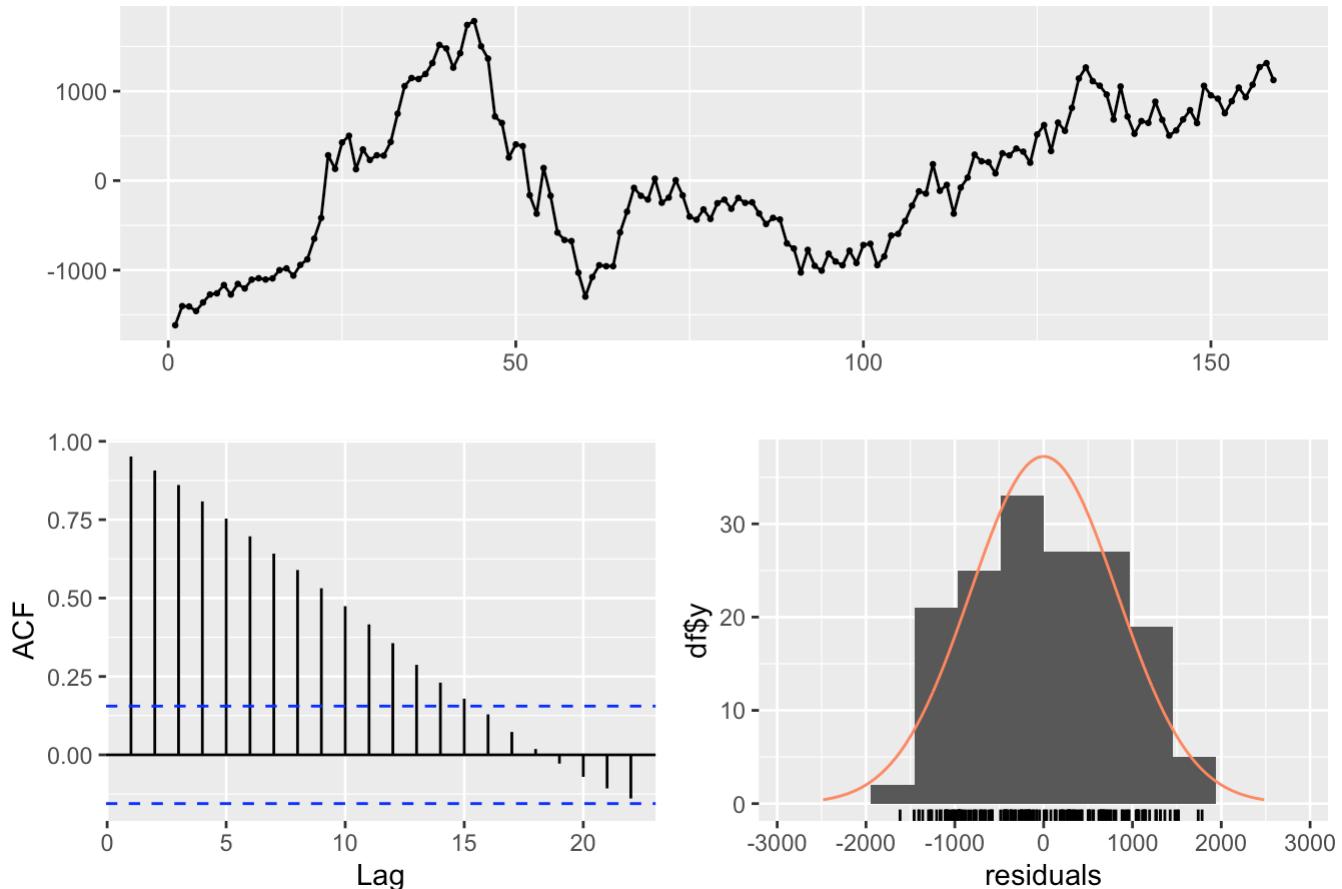
```
## Estimates and t-tests for beta coefficients:
##           Estimate Std. Error t value P(>|t|)
## beta.0     13.70     11.6   1.180  0.239
## beta.1      3.18     19.1   0.167  0.868
## beta.2     -8.41     11.5  -0.730  0.467
```

```
summary(model2_crude)
```

```
##
## Call:
## "Y ~ (Intercept) + X.t"
##
## Residuals:
##    Min      1Q  Median      3Q      Max
## -1616.4  -703.8  -77.9   657.5  1783.6
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4209.0519  179.6342  23.431 <2e-16 ***
## z.t0        13.6870   11.5777   1.182  0.239
## z.t1       -9.9541   57.9232  -0.172  0.864
## z.t2       -0.5483   28.7836  -0.019  0.985
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 835.6 on 155 degrees of freedom
## Multiple R-squared:  0.09909,    Adjusted R-squared:  0.08165
## F-statistic: 5.683 on 3 and 155 DF,  p-value: 0.001019
```

```
checkresiduals(model2_crude$model)
```

Residuals



```
## 
## Breusch-Godfrey test for serial correlation of order up to 10
## 
## data: Residuals
## LM test = 146.42, df = 10, p-value < 2.2e-16
```

From the polynomial lag model for ASX price index and crude oil price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is 0.0816 which indicates that additional predictors will not improve the model. However, the model is significant at 5% level of significance with a p-value of 0.001. Based on the Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot as the lags are over the boundary lines. Based on the time series and histogram of standardised residuals, there are no consistent upward and downward trend and non-random patterns are observed. From these results, it shows that the model does not fit the data well.

```
#Koyk Lag Model
model3_crude = koyckDlm(x = as.vector(crude_ts) , y = as.vector(asx_ts))
summary(model3_crude)
```

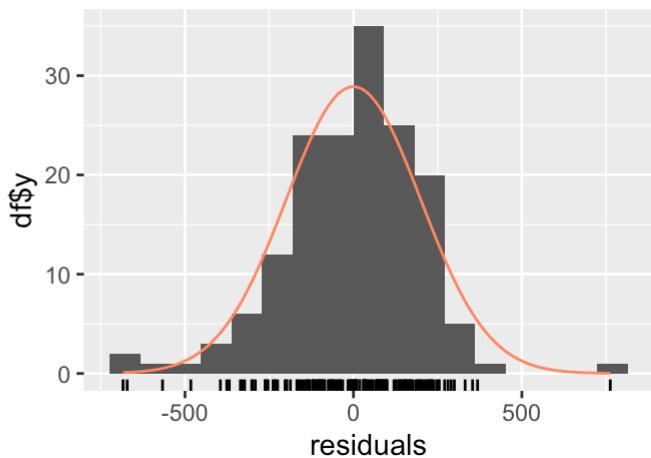
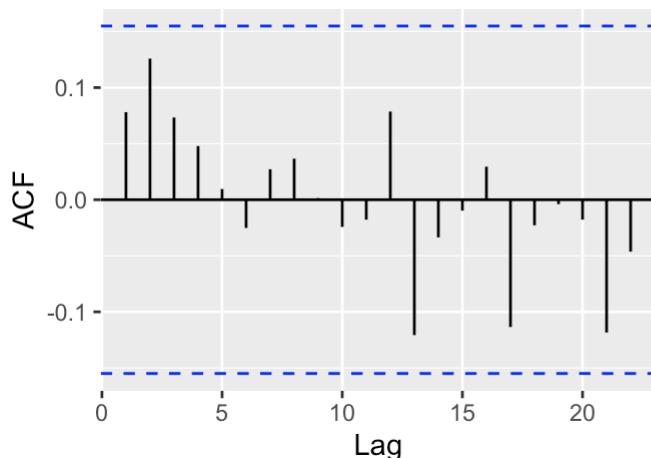
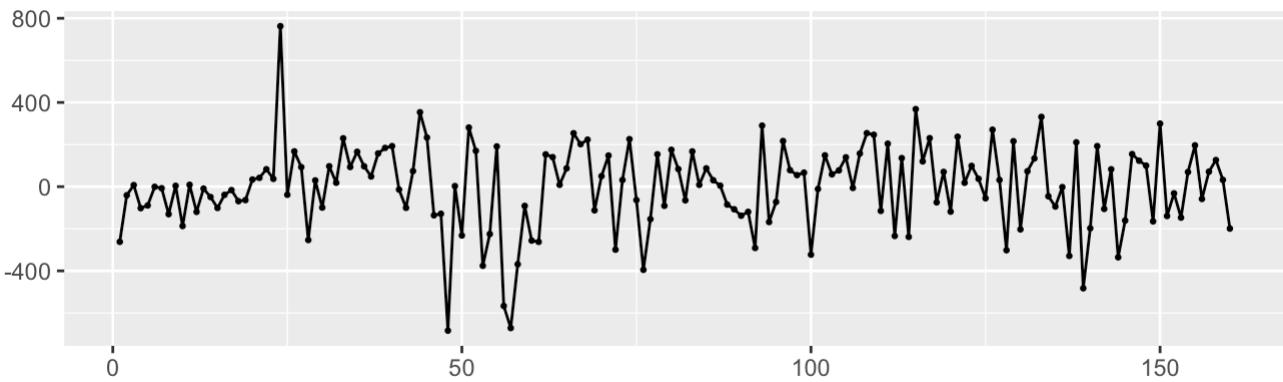
```

## 
## Call:
## "Y ~ (Intercept) + Y.1 + X.t"
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -683.91 -108.66   13.68  139.77  762.55 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 209.89536   87.89368   2.388   0.0181 *  
## Y.1          0.97537    0.01905  51.193   <2e-16 *** 
## X.t         -0.99907    0.58045  -1.721   0.0872 .  
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 201.1 on 157 degrees of freedom 
## Multiple R-Squared: 0.949,   Adjusted R-squared: 0.9483 
## Wald test: 1461 on 2 and 157 DF, p-value: < 2.2e-16 
## 
## Diagnostic tests:
## NULL
## 
##                  alpha        beta        phi 
## Geometric coefficients: 8522.034 -0.9990694 0.9753703

```

```
checkresiduals(model3_crude$model)
```

Residuals



```
##  
## Ljung-Box test  
##  
## data: Residuals  
## Q* = 5.4432, df = 10, p-value = 0.8597  
##  
## Model df: 0. Total lags used: 10
```

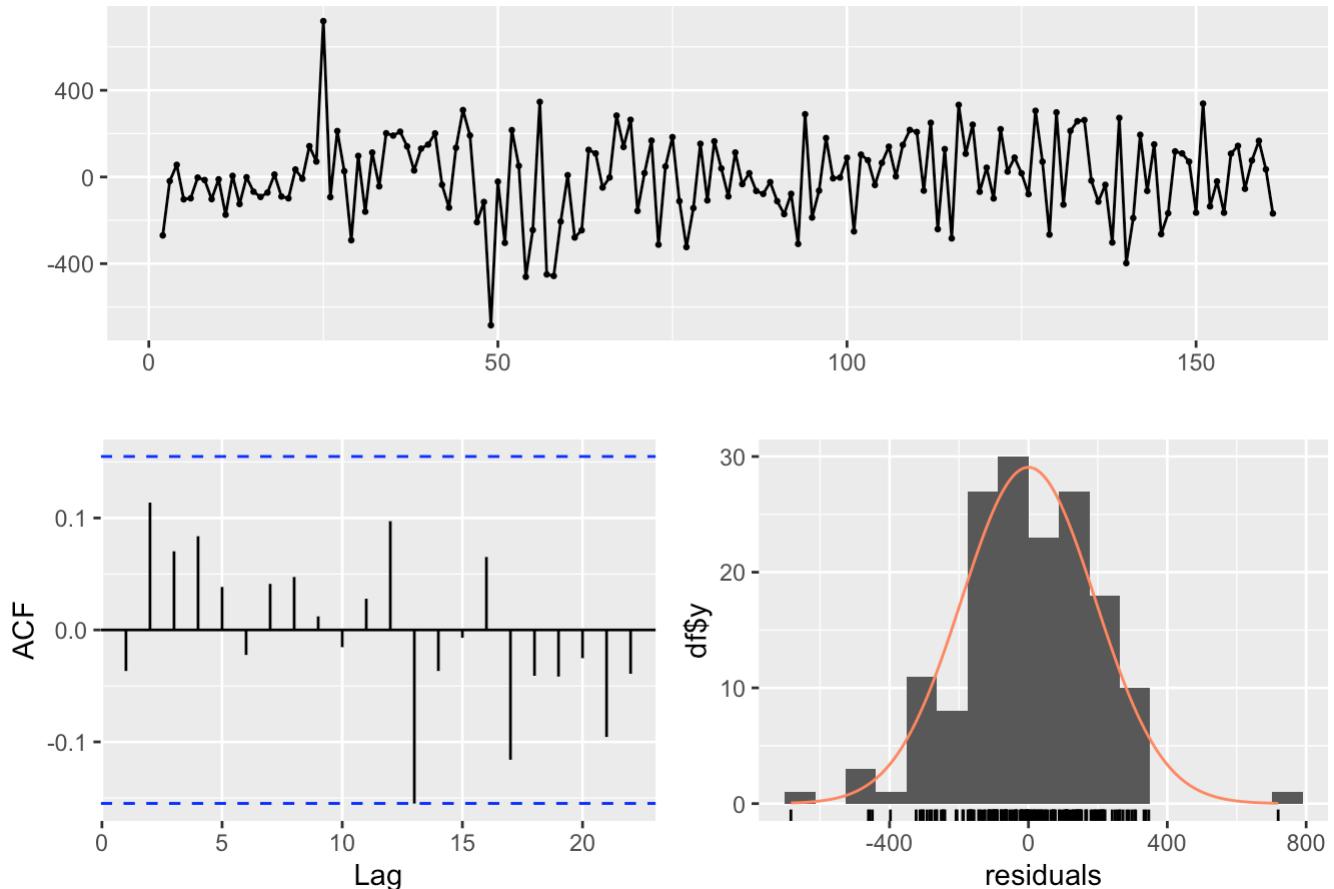
From the koyck lag model for ASX price index and crude oil price, some of the coefficients is significant ($p < 0.05$) and the adjusted R-square value is 0.9483 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. The ACF plot shows that there are no significant correlations as all the lags are well within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays a left-skewed distributed with an approximate normal distribution and presence of outliers. From these results, it shows that the model does fit the data well.

```
#Autoregressive distributed model  
model4_crude = ardlDlm(x = as.vector(crude_ts) , y = as.vector(asx_ts))  
summary(model4_crude)
```

```
##  
## Time series regression with "ts" data:  
## Start = 2, End = 161  
##  
## Call:  
## dynlm(formula = as.formula(model.text), data = data, start = 1)  
##  
## Residuals:  
##      Min      1Q Median      3Q     Max  
## -684.06 -111.57   -1.77  138.90  719.35  
##  
## Coefficients:  
##             Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 211.45397  85.03705  2.487  0.01395 *  
## X.t          7.45111   2.46199  3.026  0.00290 **  
## X.1         -8.17917   2.44422 -3.346  0.00103 **  
## Y.1          0.97067   0.01837 52.827 < 2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 194.5 on 156 degrees of freedom  
## Multiple R-squared:  0.9525, Adjusted R-squared:  0.9516  
## F-statistic: 1044 on 3 and 156 DF,  p-value: < 2.2e-16
```

```
checkresiduals(model4_crude$model)
```

Residuals



```
##  
## Breusch-Godfrey test for serial correlation of order up to 10  
##  
## data: Residuals  
## LM test = 6.0252, df = 10, p-value = 0.8131
```

```
for (i in 1:5) {  
  for (j in 1:5) {  
    model4.crude = ardlDlm(x = as.vector(crude_ts), y = as.vector(asx_ts), p = i, q = j)  
    cat("p = ", i, "q = ", j, "AIC = ", AIC(model4.crude$model), "BIC = ",  
      BIC(model4.crude$model), "\n")  
  }  
}
```

```

## p = 1 q = 1 AIC = 2146.524 BIC = 2161.9
## p = 1 q = 2 AIC = 2134.107 BIC = 2152.521
## p = 1 q = 3 AIC = 2121.07 BIC = 2142.508
## p = 1 q = 4 AIC = 2109.4 BIC = 2133.85
## p = 1 q = 5 AIC = 2098.335 BIC = 2125.784
## p = 2 q = 1 AIC = 2132.312 BIC = 2150.726
## p = 2 q = 2 AIC = 2134.235 BIC = 2155.718
## p = 2 q = 3 AIC = 2122.356 BIC = 2146.857
## p = 2 q = 4 AIC = 2110.793 BIC = 2138.299
## p = 2 q = 5 AIC = 2099.752 BIC = 2130.251
## p = 3 q = 1 AIC = 2121.919 BIC = 2143.357
## p = 3 q = 2 AIC = 2123.835 BIC = 2148.335
## p = 3 q = 3 AIC = 2124.324 BIC = 2151.887
## p = 3 q = 4 AIC = 2112.401 BIC = 2142.963
## p = 3 q = 5 AIC = 2101.35 BIC = 2134.899
## p = 4 q = 1 AIC = 2111.383 BIC = 2135.832
## p = 4 q = 2 AIC = 2113.294 BIC = 2140.8
## p = 4 q = 3 AIC = 2113.805 BIC = 2144.367
## p = 4 q = 4 AIC = 2114.384 BIC = 2148.003
## p = 4 q = 5 AIC = 2103.342 BIC = 2139.94
## p = 5 q = 1 AIC = 2097.076 BIC = 2124.525
## p = 5 q = 2 AIC = 2099.041 BIC = 2129.54
## p = 5 q = 3 AIC = 2099.518 BIC = 2133.066
## p = 5 q = 4 AIC = 2099.845 BIC = 2136.443
## p = 5 q = 5 AIC = 2100.917 BIC = 2140.566

```

From the autoregressive distributed lag model for ASX price index and crude oil price, all the coefficients are significant ($p < 0.05$) and the adjusted R-square value is 0.9516 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. Based on the Breusch-Godfrey test, it shows that as the p-value is greater than 0.05, we cannot reject the null hypothesis and conclude that autocorrelation does not exist among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot which shows that there are no significant correlations as all the lags are within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays an approximate normal distribution. In addition, $p = 5$ and $q = 1$ gives the smallest AIC and BIC values. From these results, it shows that the model does fit the data well.

ASX Price Index & Copper Price

```

#Finite Lag Model
model1_copper = dlm( x = as.vector(copper_ts) , y = as.vector(asx_ts), q = 9 )
summary(model1_copper)

```

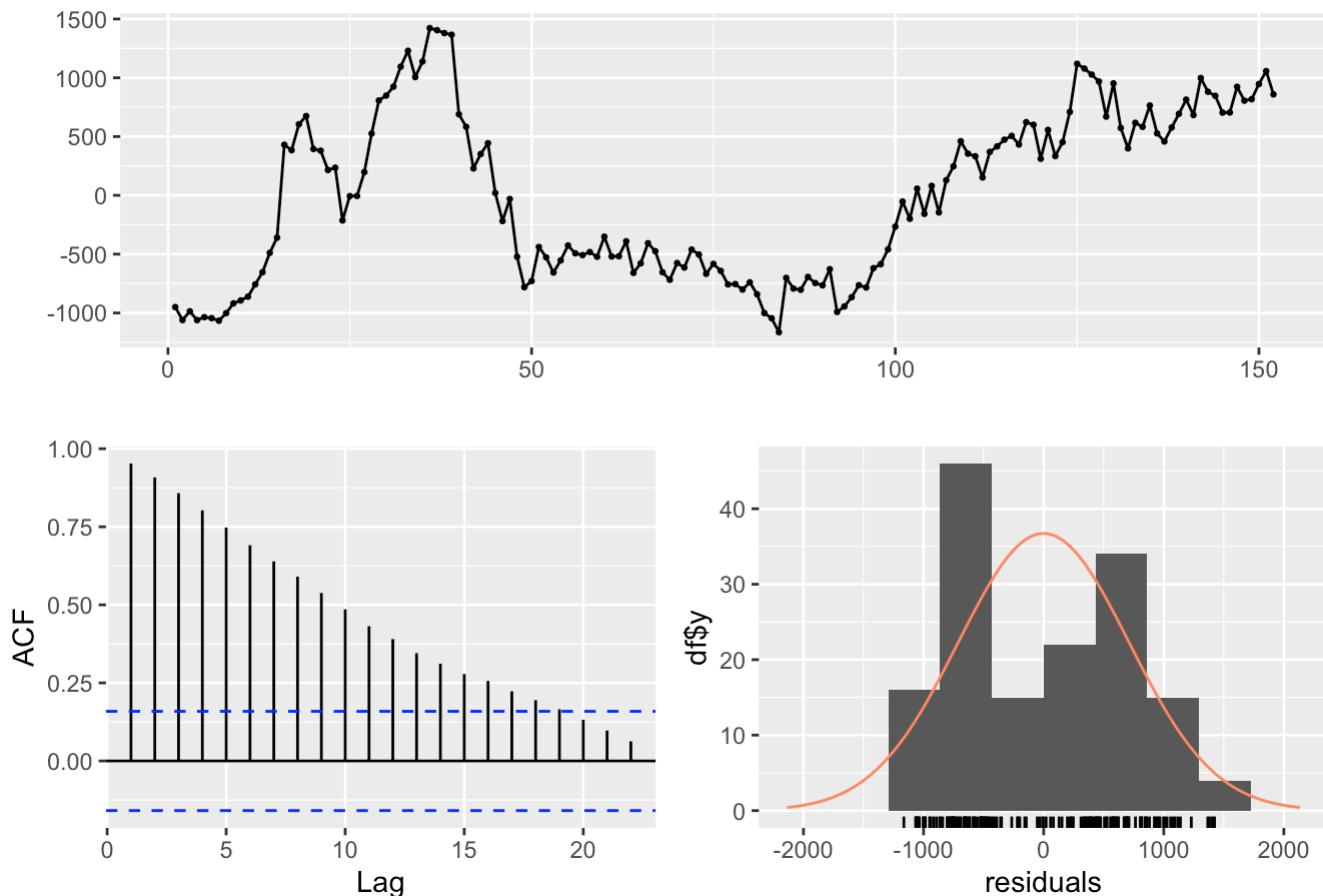
```

## 
## Call:
## lm(formula = model.formula, data = design)
##
## Residuals:
##    Min      1Q  Median      3Q     Max 
## -1163.95 -653.62   -5.48   601.46 1422.56 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 3.921e+03 2.114e+02 18.546 <2e-16 ***
## x.t         1.576e-01 1.351e-01  1.166   0.245    
## x.1         1.829e-02 2.177e-01  0.084   0.933    
## x.2         4.688e-02 2.177e-01  0.215   0.830    
## x.3         2.755e-02 2.164e-01  0.127   0.899    
## x.4         2.061e-02 2.157e-01  0.096   0.924    
## x.5        -5.263e-02 2.157e-01 -0.244   0.808    
## x.6         3.688e-02 2.165e-01  0.170   0.865    
## x.7        -5.357e-03 2.186e-01 -0.025   0.980    
## x.8        -2.372e-04 2.195e-01 -0.001   0.999    
## x.9        -9.203e-02 1.337e-01 -0.688   0.493    
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 736.7 on 141 degrees of freedom
## Multiple R-squared:  0.2059, Adjusted R-squared:  0.1496 
## F-statistic: 3.656 on 10 and 141 DF,  p-value: 0.000233 
## 
## AIC and BIC values for the model:
##       AIC      BIC  
## 1 2451.016 2487.302

```

```
checkresiduals(model1_copper$model)
```

Residuals



```
## 
## Breusch-Godfrey test for serial correlation of order up to 14
## 
## data: Residuals
## LM test = 140.43, df = 14, p-value < 2.2e-16
```

```
# Finite lag length based on AIC and BIC values
for (i in 1:10) {
  model1.copper = dlm(x = as.vector(copper_ts), y = as.vector(asx_ts), q = i)
  cat("q = ", i, "AIC = ", AIC(model1.copper$model), "BIC = ", BIC(model1.copper$model), "\n")
}
```

```
## q = 1 AIC = 2574.488 BIC = 2586.789
## q = 2 AIC = 2559.356 BIC = 2574.7
## q = 3 AIC = 2544.155 BIC = 2562.531
## q = 4 AIC = 2528.895 BIC = 2550.289
## q = 5 AIC = 2513.265 BIC = 2537.664
## q = 6 AIC = 2497.775 BIC = 2525.166
## q = 7 AIC = 2481.988 BIC = 2512.357
## q = 8 AIC = 2466.511 BIC = 2499.846
## q = 9 AIC = 2451.016 BIC = 2487.302
## q = 10 AIC = 2436.164 BIC = 2475.389
```

From the finite lag model for ASX price index and copper price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is 0.1496 which indicates that additional predictors will not improve the model. However, the model is significant at 5% level of significance with a p-value of 0.0002. Based on the

Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 14. This is further confirmed by the ACF plot as the lags are not within the boundary lines. Based on the time series and histogram of standardised residuals, there are very high residual values that violate general assumptions. In addition, both the AIC and BIC values suggest the use of lag length 10 as it has the lowest AIC and BIC values. From these results, it shows that the model does not fit the data well.

```
#Polynomial Lag Model
model2_copper = polyDlm(x = as.vector(copper_ts) , y = as.vector(asx_ts) , q = 2 , k = 2 , show.beta = TRUE)
```

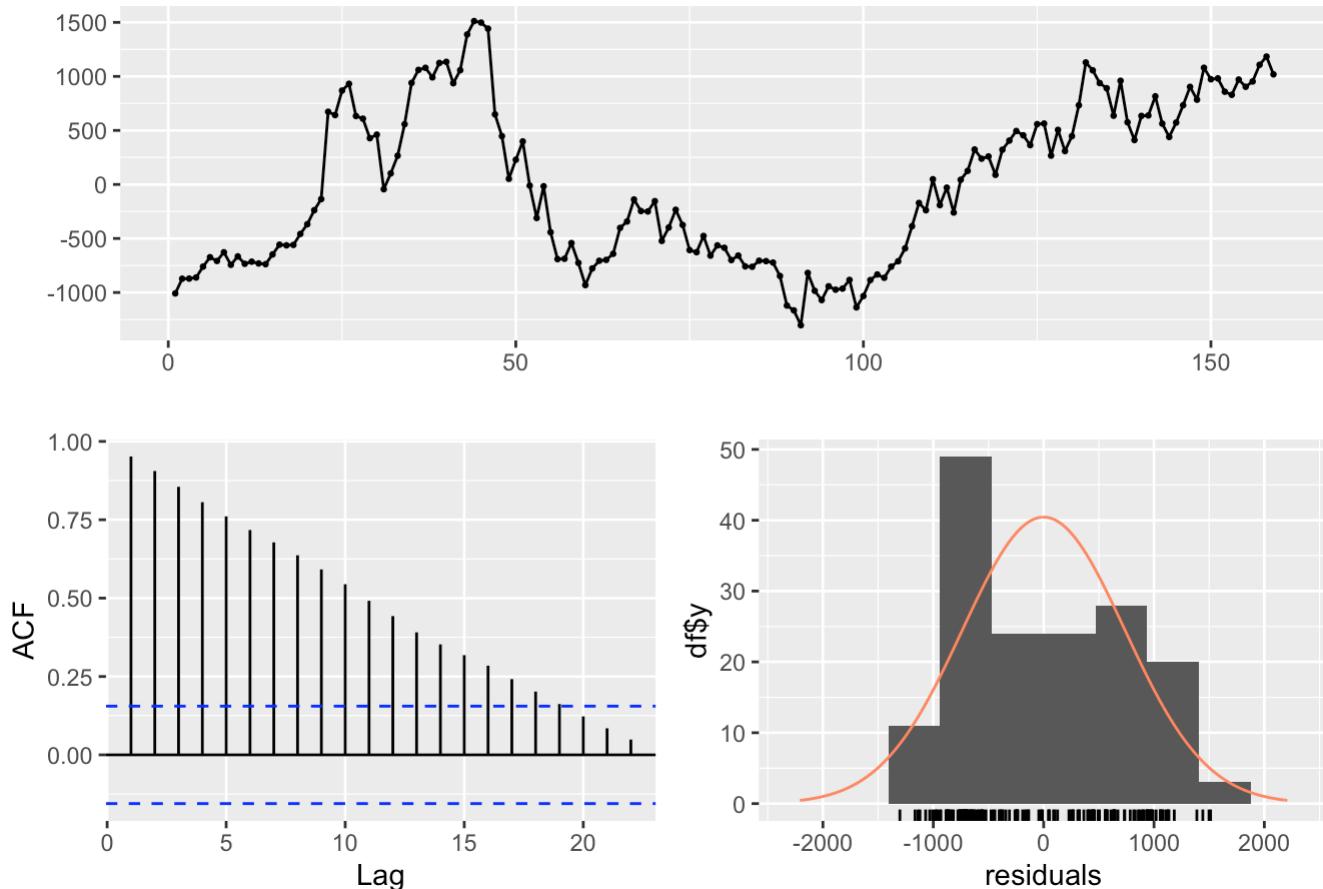
```
## Estimates and t-tests for beta coefficients:
##           Estimate Std. Error t value P(>|t|)
## beta.0    0.17800   0.131   1.3600  0.176
## beta.1    0.05290   0.207   0.2550  0.799
## beta.2   -0.00654   0.129  -0.0507  0.960
```

```
summary(model2_copper)
```

```
##
## Call:
## "Y ~ (Intercept) + X.t"
##
## Residuals:
##    Min      1Q  Median      3Q     Max
## -1302.7 -694.4 -135.3  635.5 1512.0
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3484.26812 181.89988 19.155 <2e-16 ***
## z.t0        0.17781   0.13067  1.361   0.176
## z.t1       -0.15771   0.62898 -0.251   0.802
## z.t2        0.03277   0.31191  0.105   0.916
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 743 on 155 degrees of freedom
## Multiple R-squared:  0.2877, Adjusted R-squared:  0.274
## F-statistic: 20.87 on 3 and 155 DF,  p-value: 2.065e-11
```

```
checkresiduals(model2_copper$model)
```

Residuals



```
## 
## Breusch-Godfrey test for serial correlation of order up to 10
## 
## data: Residuals
## LM test = 146.69, df = 10, p-value < 2.2e-16
```

From the polynomial lag model for ASX price index and copper price, all the coefficients are insignificant ($p > 0.05$) and the adjusted R-square value is 0.274 which indicates that additional predictors will improve the model. However, the model is significant at 5% level of significance with a p-value less than 0.05. Based on the Breusch-Godfrey test, it shows that as the p-value is less than 0.05, we can reject the null hypothesis and conclude that autocorrelation exists among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot as the lags are over the boundary lines. Based on the time series and histogram of standardised residuals, there are no consistent upward and downward trend and non-random patterns are observed. From these results, it shows that the model does not fit the data well.

```
# Koyck Lag Model
model3_copper = koyckDlm(x = as.vector(copper_ts) , y = as.vector(asx_ts))
summary(model3_copper)
```

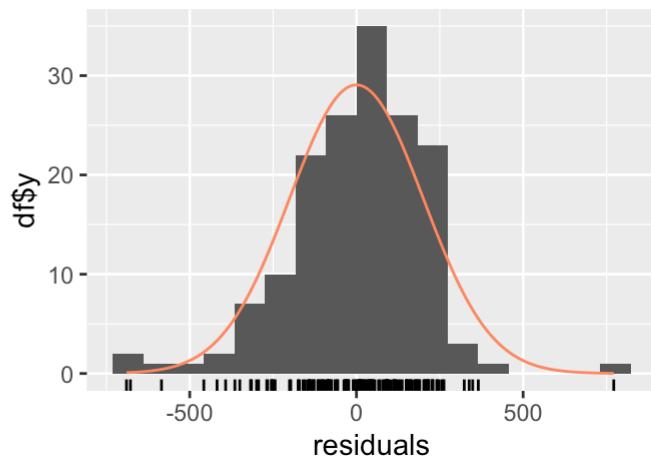
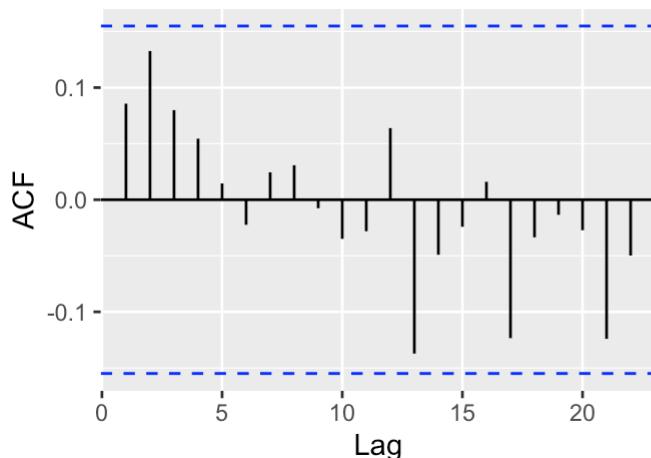
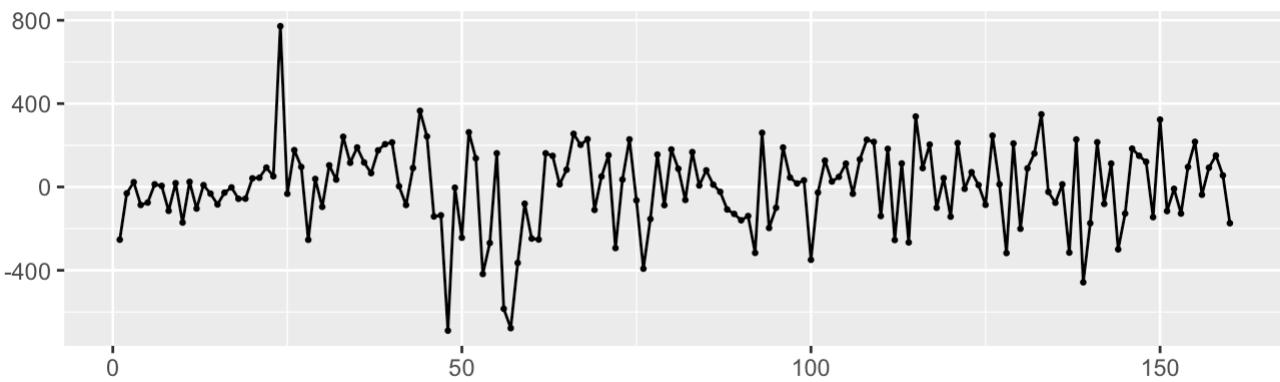
```

## 
## Call:
## "Y ~ (Intercept) + Y.1 + X.t"
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -689.64 -108.62   12.78  140.20  771.79 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 189.368812  87.644648   2.161   0.0322 *  
## Y.1          0.971621   0.021895  44.376  <2e-16 ***  
## X.t         -0.005864   0.009517  -0.616   0.5387    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 201.9 on 157 degrees of freedom
## Multiple R-Squared: 0.9485, Adjusted R-squared: 0.9479 
## Wald test: 1448 on 2 and 157 DF, p-value: < 2.2e-16
## 
## Diagnostic tests:
## NULL
## 
##                  alpha          beta          phi
## Geometric coefficients: 6672.885 -0.005863623 0.9716211

```

```
checkresiduals(model3_copper$model)
```

Residuals



```
##  
## Ljung-Box test  
##  
## data: Residuals  
## Q* = 6.2327, df = 10, p-value = 0.7953  
##  
## Model df: 0. Total lags used: 10
```

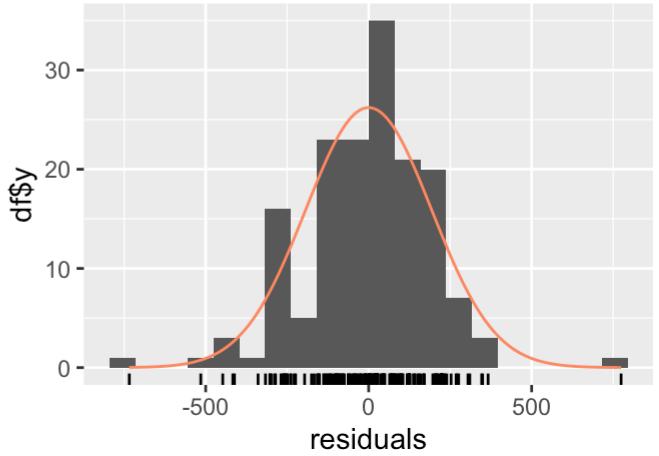
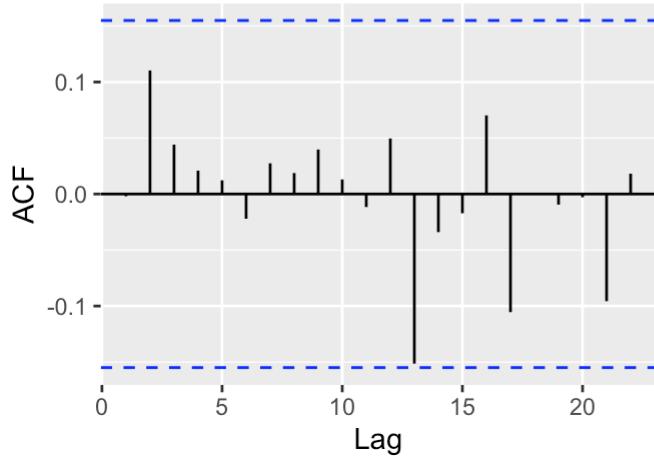
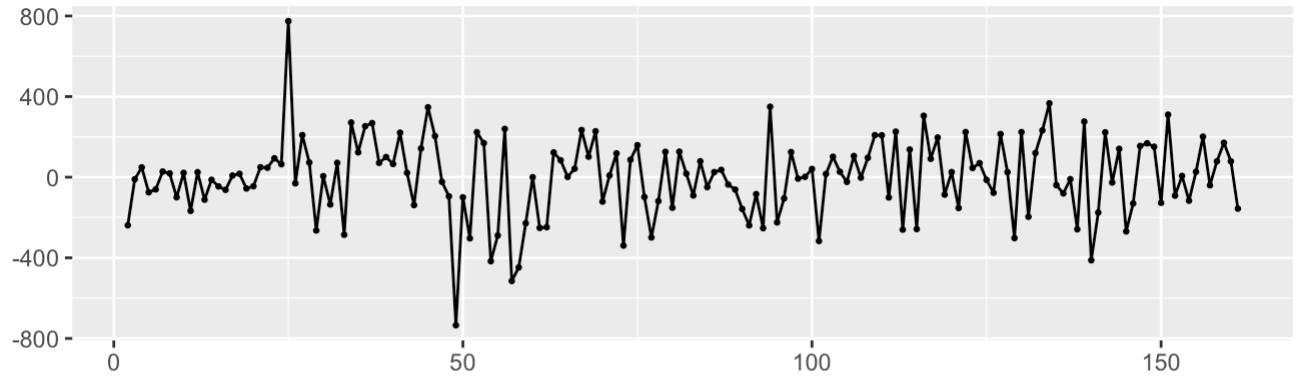
From the koyck lag model for ASX price index and copper price, some of the coefficients is significant ($p < 0.05$) and the adjusted R-square value is 0.9479 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. The ACF plot shows that there are no significant correlations as all the lags are well within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays a left-skewed distributed with an approximate normal distribution and presence of outliers. From these results, it shows that the model does fit the data well.

```
# Autoregressive Distributed Lag Model  
model4_copper = ardlDlm(x = as.vector(copper_ts) , y = as.vector(asx_ts), p = 1 , q =  
1 )  
summary(model4_copper)
```

```
##  
## Time series regression with "ts" data:  
## Start = 2, End = 161  
##  
## Call:  
## dynlm(formula = as.formula(model.text), data = data, start = 1)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max  
## -734.34 -101.82   16.56  123.05  774.55  
##  
## Coefficients:  
##             Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 173.61066   84.94572   2.044  0.04266 *  
## X.t          0.10629    0.03258   3.263  0.00136 **  
## X.1         -0.10695    0.03228  -3.313  0.00115 **  
## Y.1          0.96784    0.02103  46.027 < 2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 195.2 on 156 degrees of freedom  
## Multiple R-squared:  0.9522, Adjusted R-squared:  0.9513  
## F-statistic: 1035 on 3 and 156 DF,  p-value: < 2.2e-16
```

```
checkresiduals(model4_copper$model)
```

Residuals



```
##  
## Breusch-Godfrey test for serial correlation of order up to 10  
##  
## data: Residuals  
## LM test = 3.6178, df = 10, p-value = 0.9629
```

```
for (i in 1:5) {  
  for (j in 1:5) {  
    model4.copper = ardlDlm(x = as.vector(copper_ts), y = as.vector(asx_ts), p = i, q = j)  
    cat("p = ", i, "q = ", j, "AIC = ", AIC(model4.copper$model), "BIC = ",  
      BIC(model4.copper$model), "\n")  
  }  
}
```

```

## p = 1 q = 1 AIC = 2147.741 BIC = 2163.116
## p = 1 q = 2 AIC = 2135.4 BIC = 2153.813
## p = 1 q = 3 AIC = 2121.12 BIC = 2142.558
## p = 1 q = 4 AIC = 2109.759 BIC = 2134.209
## p = 1 q = 5 AIC = 2099.056 BIC = 2126.505
## p = 2 q = 1 AIC = 2130.043 BIC = 2148.456
## p = 2 q = 2 AIC = 2132.038 BIC = 2153.52
## p = 2 q = 3 AIC = 2119.241 BIC = 2143.741
## p = 2 q = 4 AIC = 2107.649 BIC = 2135.155
## p = 2 q = 5 AIC = 2097.021 BIC = 2127.52
## p = 3 q = 1 AIC = 2117.307 BIC = 2138.745
## p = 3 q = 2 AIC = 2119.247 BIC = 2143.748
## p = 3 q = 3 AIC = 2119.696 BIC = 2147.259
## p = 3 q = 4 AIC = 2108.537 BIC = 2139.1
## p = 3 q = 5 AIC = 2097.832 BIC = 2131.38
## p = 4 q = 1 AIC = 2105.916 BIC = 2130.366
## p = 4 q = 2 AIC = 2107.774 BIC = 2135.28
## p = 4 q = 3 AIC = 2108.608 BIC = 2139.17
## p = 4 q = 4 AIC = 2110.085 BIC = 2143.704
## p = 4 q = 5 AIC = 2099.454 BIC = 2136.052
## p = 5 q = 1 AIC = 2095.118 BIC = 2122.566
## p = 5 q = 2 AIC = 2096.96 BIC = 2127.459
## p = 5 q = 3 AIC = 2097.887 BIC = 2131.436
## p = 5 q = 4 AIC = 2099.497 BIC = 2136.095
## p = 5 q = 5 AIC = 2101.419 BIC = 2141.067

```

From the autoregressive distributed lag model for ASX price index and copper price, all the coefficients are significant ($p < 0.05$) and the adjusted R-square value is 0.9513 which indicates that additional predictors will improve the model. Furthermore, the model is significant at 5% level of significance with a p-value less than 0.05. Based on the Breusch-Godfrey test, it shows that as the p-value is greater than 0.05, we cannot reject the null hypothesis and conclude that autocorrelation does not exist among the residuals at some order less than or equal to 10. This is further confirmed by the ACF plot which shows that there are no significant correlations as all the lags are within the reference boundary lines. Based on the time series, there is no consistent upward or downward trend and the histogram of standardised residuals displays an approximate normal distribution. In addition, $p = 5$ and $q = 2$ gives the smallest AIC and BIC values. From these results, it shows that the model does fit the data well.

Discussion/Conclusion

The monthly averages of ASX All Ordinaries (Ords) price index, gold (AUD), crude oil (Brent, USD/bbl), and copper (USD/Tonne) prices from January 2004 to December 2017. The trend of these variables throughout the time period were evaluated through time series and scatterplots. From the time series plots, overall, there are upward and downward trends with no presence of seasonality. Based on the scatterplots, all the series display a strong, linear, positive trend which is further confirmed by the high correlation values.

The stationarity of the series are analysed through ACF and PACF plots which displays that as the lags are not within the boundary in the ACF plots, there is an existence of a trend in the series. This is further confirmed by the PACF plots where a very high first autocorrelation is seen. The ACF and PACF plots can conclude that the series for ASX price index, gold, crude oil and copper prices are nonstationary. Further testing such as two unit root test statistics were used to confirm the presence of nonstationarity in the series. Based on the ADF and PP tests, they both confirm that the series are all nonstationary. Furthermore, Shapiro-Wilk test was performed to analyse the normality of the residuals and the results show that the normality of residuals are supported for the series.

In order to evaluate the necessity of transforming the series, the lambda values for each series were determined. From these lambda values, it shows that only two series required transformation, ASX price index and crude oil price, as the lambda value is close to 2 and -1, respectively. The appropriate transformation technique were applied to the series and coupled with a time series plot to display the transformation, the trend of the series remain the same as the original series. Due to this, the transformation of the series were not applied to both ASX price index and crude oil price.

As the series are non-stationary, first differencing was applied to each of the series to transform the stationarity of the series. After applying first differencing, the time series plot display a more stationary trend, however this was further confirmed by unit root test statistics such as ADF and PP tests. From these two tests, it shows that the series are stationary after first differencing.

Decomposition of the series were performed to further analyse the trend and seasonal factors of the time series. This will improve understanding of the series as well as forecasting accuracy. The results from the decompositon show that the pattern of the seasonally adjusted series follows the original series closely however there are slight deviations of a seasonal effect.

Multiple distributed lag models were used to determine the best fitted model for the series, where the dependent variable, Y, is ASX price index, and the independent variables, X, are gold, crude oil and copper prices. Such distributed lag models include definite, polynomial, koyk and auto-regressive distributed lag model. Diagnostic checks are performed after the fitting of the model to determine the suitability of each model to the series. Comparing all the different models utilised, the worst model for each series is the finite distributed lag model. This is due to the lack of significant coefficients found in the model, the significantly low adjusted R-squared value, and the high p-value of the model ($p > 0.05$) indicates that the model is a poor fit for the series. Furthermore, the diagnostic checks confirms that autocorrelation exists as the p-value is less than 0.05 in the Breusch-Godfrey test and the lags are out of the boundary lines in the ACF plot. The time series of standardised residuals display an overall upward trend and the histogram of standardised residuals violates general assumptions. Based on these results, it can be concluded that the definite distributed lag model is the worst model for the series. On the other hand, the best model for the series appear to be the auto-regressive distributed lag model. Based on the results of this model, most of the coefficients are significant ($p < 0.05$), and the model has a significantly high adjusted R-squared value, followed by a significantly low p-value ($p < 0.05$), which indicates the significant suitability of the model for the series. Furthermore, from the diagnostics checks, the high Breusch-Godfrey test indicates that there is no autocorrelation present in the series and this is confirmed by the ACF plot where the lags are well within the boundary lines. From the time series plot of the standardised residuals, there are no definite upward and downward trend and the histogram of standardised residuals appear approximately normally distributed. Based on these results, it can be said that the best performing model for the series is the auto-regressive distributed model.