Report

Data series – 12

Fig 1.1 – Actual data plot – No seasonality observed

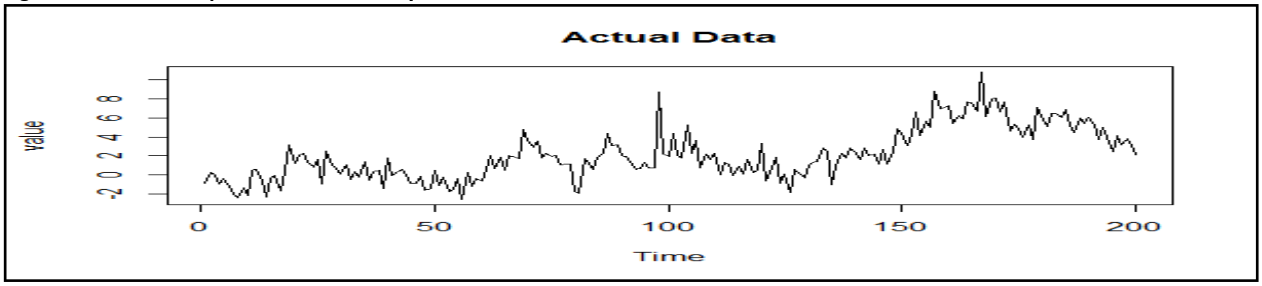
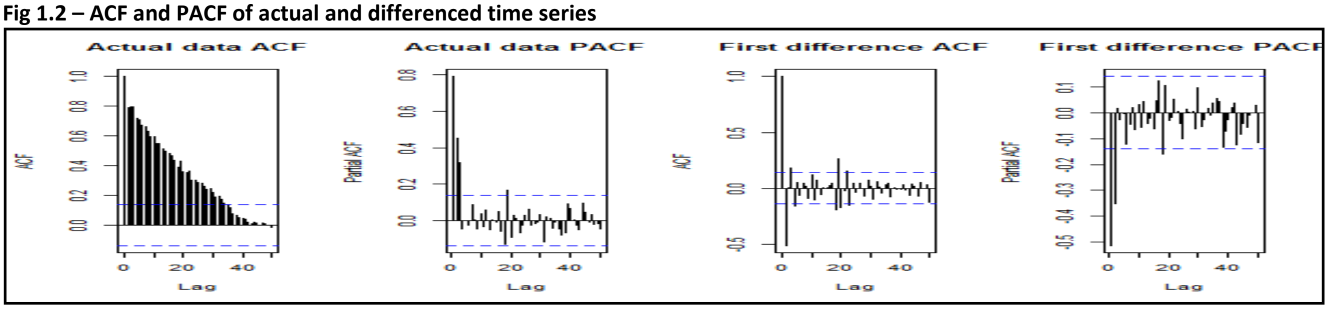


Table 1.1 – Basic statistics

Statistic	Value	Comment
Mean	2.134	Standard deviation is larger than mean i.e. data is more spread out indicating volatility is high. Coefficient variation is 125% approximately.
Standard deviation	2.663	
Skewness	0.653	Data is positively skewed i.e. mean is greater than mode and as per kurtosis, data is platykurtic. This indicates data is not normally distributed. Using JB test, normality is rejected with p value 0.0007009.
Kurtosis	2.882	

<u>Exploratory Analysis</u>

Fig 1.2 – ACF and PACF of actual and differenced time series



Actual data - Since ACF is geometric and PACF cuts off at lag 3, this suggests AR(3) process. ADF test with p value 0.2124 suggests that data is not stationary. Using tbats and !is.null(fit$seasonal) function, it is proved that there is no seasonality. Hence data is only detrended by first order differencing.
First order difference data – ACF and PACF cuts off at lag 1 suggesting ARIMA (2,1,1). ADF test with p value 0.01 and unit root test with p value 0.0009 suggests first order differenced data series is stationary. AR order based on OLS method with BIC criteria also suggests AR(2) process. With AR(2) process, solution of characteristic root is 1.687 which suggests stationarity.

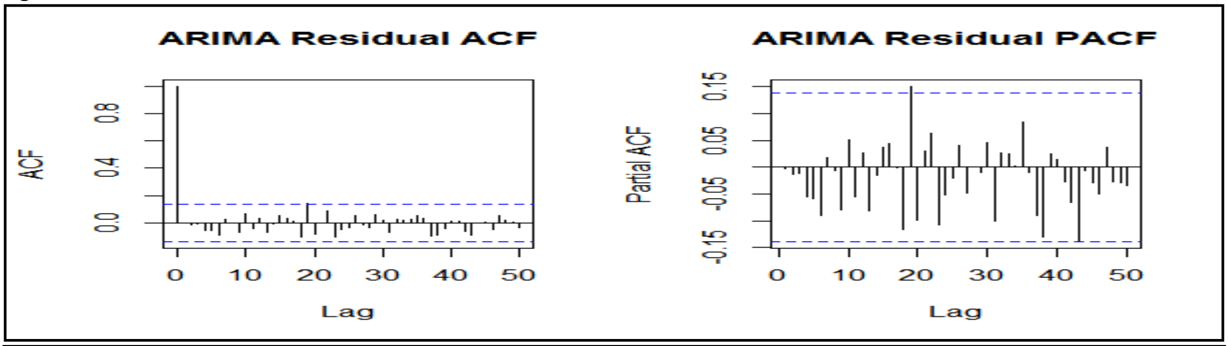
<u>Table 1.2 - Summary of ARIMA</u> - ARIMA (2,1,1) is better model than model estimated by auto.arima function [ARIMA(0,1,1)].

Model	AR(1)	AR(2)	MA(1)	Adj R^2	AIC	Overall MSE	Last 15 obs MSE
0,1,1			-0.6217	72.40%	703.14	1.94	1.06
S.E			0.0523				
2,1,1	-0.7413	-0.375	0.0530	73.64%	697.9	1.86	0.97
S.E	0.2006	0.115	0.2185				

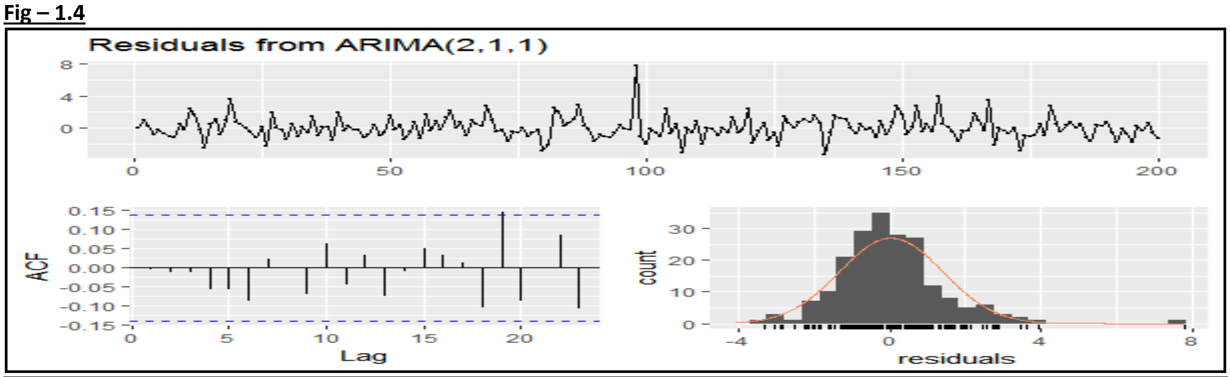
<u>Table 1.3 - Residual diagnosis (ARIMA – 2,1,1)</u> – ADF test and unit root test suggests that residual is stationary. Run test suggests that residuals are random in nature. White noise test suggests that residual is white noise process. Ljung-Box test suggest that there is no significant serial correlation.

Test	Unitroot Test	ADF Test	Run Test	White noise Test	Ljung-Box
p value	0.0009	0.01	0.4784	0.420	0.501

Fig 1.3



<u>Fig – 1.4</u>



<u>Model 2: HoltWinters Exponential Smoothing method</u> – Holt Winter exponential smoothing method's model parameters with trend has alpha 0.46 and beta 0.11. This is the best model w.r.t. minimum MSE. For without trend mode, alpha is 0.37. Overall in sample forecast model without trend is better than with trend (Beta). However, out of sample point forecast in without trend will be flat as slope is not available. Since objective of the assignment is in sample forecast accuracy, model without trend is selected.

<u>Table 1.4 – Coefficient and Summary Holt-Winters Exponential smoothing model</u>

Model	α	β	Overall MSE	Last 15 obs MSE
With trend	2.62	-0.21	2.26	0.94
Without trend	2.97	-	1.96	1.06

Fig 1.5 – Holt-Winters Actual vs fitted values

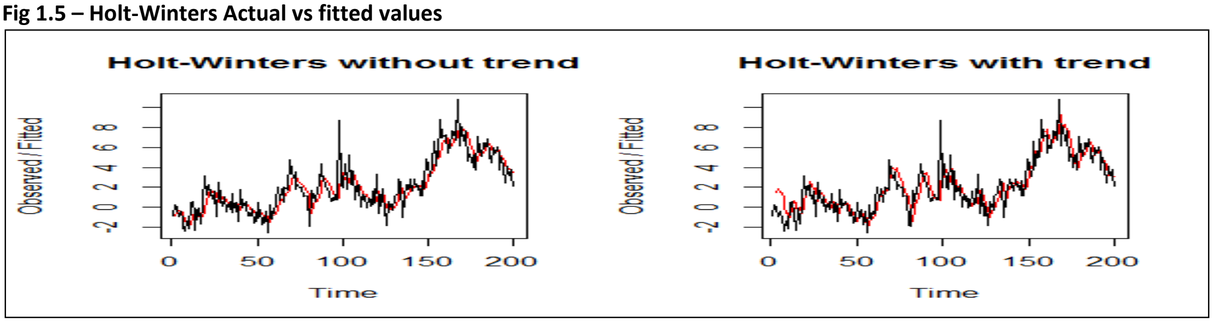


Table 1.5 - Residual diagnosis (HoltWinter) – For both the models**,** ADF test and unit root test suggests that residual is stationary. Run test suggests that residuals are random in nature. White noise test suggests that residual is white noise process. Ljung-Box test suggest that there is significant serial correlation.

P value	ADF Test	Run Test	White noise Test	Ljung-Box
Without trend	0.01	0.15	0.80	0.036
With trend	0.01	0.19	0.69	0.016

Table 1.6 - Summary of all models – ARIMA (2,1,1) is best model w.r.t Overall MSE and AIC. HoltWinters outperforms rest model if accuracy of last 15 observations are only considered. However, HoltWinter model has autocorrelation in residuals, hence ARIMA model is more adequate than HoltWinter exponential smoothing for forecasting.

Model	Parameters	Overall MSE	MSE - Last 15 Obs
ARIMA	0,1,1	1.94	1.06
ARIMA	2,1,1	1.86	0.97
Holt-Winters	α, β	2.26	0.94
Holt-Winters	α	1.96	1.06

Data series – Bharti Infratel (01-Jun-2014 till 30-Sep-2017)

Fig 2.1 – Actual data plot – No seasonality observed

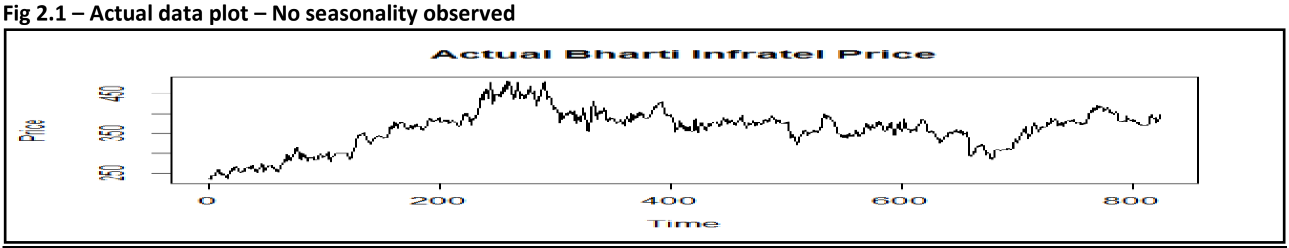
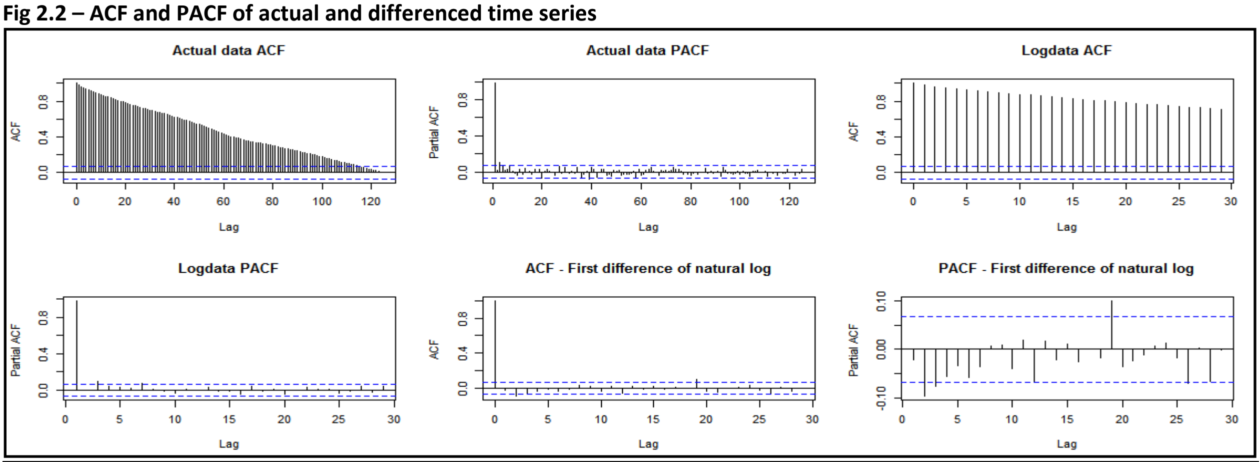


Table 2.1 – Basic statistics of Actual data

Statistic	Value	Comment
Mean	361.38	Standard deviation is smaller than mean i.e. data is not much spread out. Coefficient variation is 13.6% approximately.
Standard deviation	49.115	
Skewness	-0.420	Data is negatively skewed i.e. mode is greater than mean and as per kurtosis, data is platykurtic. This indicates data is not normally distributed. Using JB test, normality is rejected with p value 0.00000395.
Kurtosis	0.1180	

Exploratory Analysis

Fig 2.2 – ACF and PACF of actual and differenced time series



Actual data - Since ACF is geometric and PACF cuts off at lag 3, this suggests AR(3) process. ADF test with p value 0.4324 and 0.4114 for actual and natural log respectively, suggests that data is not stationary. Using tbats and !is.null(fit$seasonal) function, it is proved that there is no seasonality. Hence data is only detrended by first order differencing of natural log.
First difference data of natural log– First order differencing is done on natural log data. ADF test with p value 0.01 and unit root test 0.0009 suggests differenced data series of natural log is stationary. ACF cuts off at lag 1 suggesting MA(1) and PACF cuts off at lag 3 suggesting AR(3).

<u>**Table 2.2 - Summary of ARIMA**</u> – ARIMA (3,1,1) is selected as best model based on AIC and MSE for forecast.

Model	AR(1)	AR(2)	AR(3)	MA(1)	Adj R^2	AIC	In Sample MSE	Out of Sample MSE
3,1,1	0.36	-0.10	-0.04	-1.40	1.00	-3834.34	73.36	1010.90
S.E	0.58	0.04	0.07	0.58				
Unit root	2.16	2.82	2.16					

<u>**Table 2.3 - Residual diagnosis of ARIMA(3,1,1)**</u> – ADF test suggests that residual is stationary. Run test suggests that residuals are random in nature. White noise test suggests that residual is white noise process. Ljung-Box test suggest that there is no significant serial correlation. Unit root test suggests stationary.

Test	ADF Test	Run Test	White noise Test	Ljung-Box
p value	0.01	0.3642	0.0876	0.171

Fig 2.3 – ARIMA (6,1,2) Residual ACF and PACF

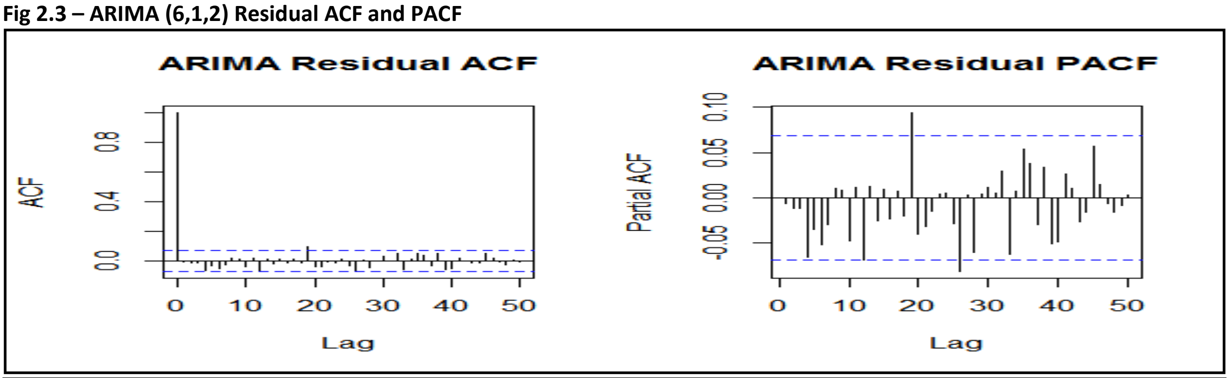
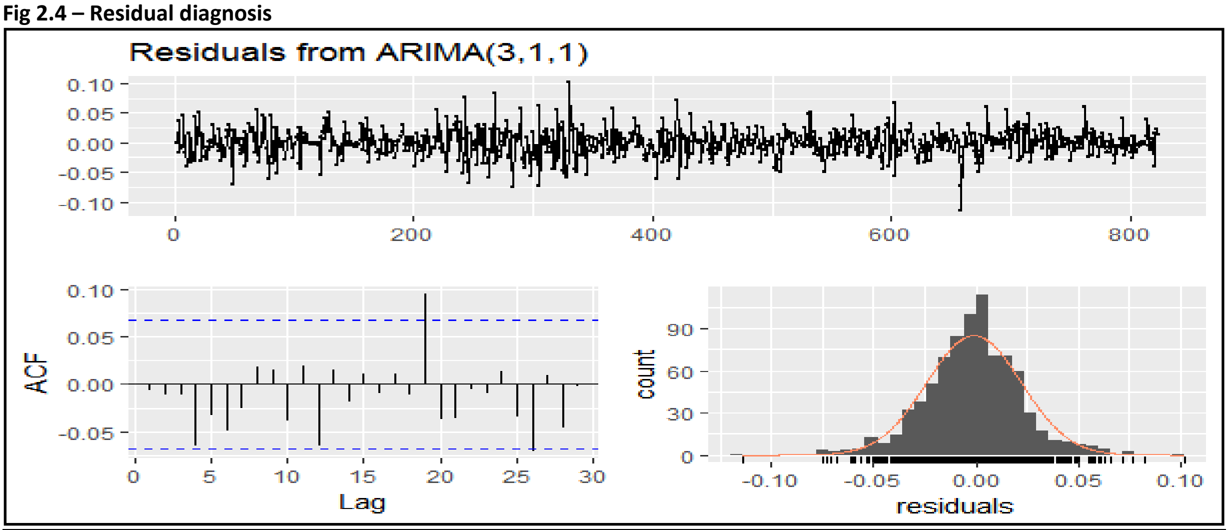


Fig 2.4 – Residual diagnosis



<u>**Model 2: HoltWinters Exponential Smoothing method**</u> -

Table 2.4 – Coefficient and Summary Holt-Winters Exponential smoothing model – Accuracy is better by using differenced log data. With trend model parameters are alpha 0.32 and beta 0.19. This is the best model w.r.t. out of sample forecast accuracy. For without trend mode, alpha is 0.012. Overall in sample forecast model without trend is better than with trend (Beta). However, out of sample point forecast in without trend will be flat as slope is not available. Since objective of the assignment is out of sample forecast accuracy, model with trend is selected.

Model	α	β	In Sample MSE	Out of Sample MSE
With trend	0.008	0.001	111.84	329.97
Without trend	0.0008		75.87	-

Table 2.5 - Residual diagnosis (HoltWinter) – For both the models**,** ADF test and unit root test suggests that residual is stationary. Run test suggests that residuals are random in nature. White noise test suggests that residual is white noise process. Ljung-Box test suggest that there is significant serial correlation.

P value	ADF Test	Run Test	White noise Test	Ljung-Box
Without trend	0.01	0.40	0.49	0.002
With trend	0.01	0.62	0.13	0.007

Table 2.6 - Summary of all models – ARIMA (3,1,1) gives better result for in sample forecast as compare to Holt Winters. But, Holt Winter outperforms ARIMA in out of sample forecast for this data series. However, Holt Winters residuals are autocorrelated, hence from model adequacy perspective ARIMA outperforms HoltWinter model.

Model	Parameters	Overall MSE	Out of sample MSE
ARIMA	3,1,1	73.36	1010.90
Holt-Winters	α, β	111.84	329.97

Fig 2.5 – Out of sample forecast comparison