Time Series Analysis on Indian Car Manufacturers' Stock Performance

GROUP-5

DIKSHANT JOSHI

NAVEEN K KRISHNASAMY

RAJASHEKAR REDDY VEMULA

Agenda

INTRODUCTION TO THE DATASET

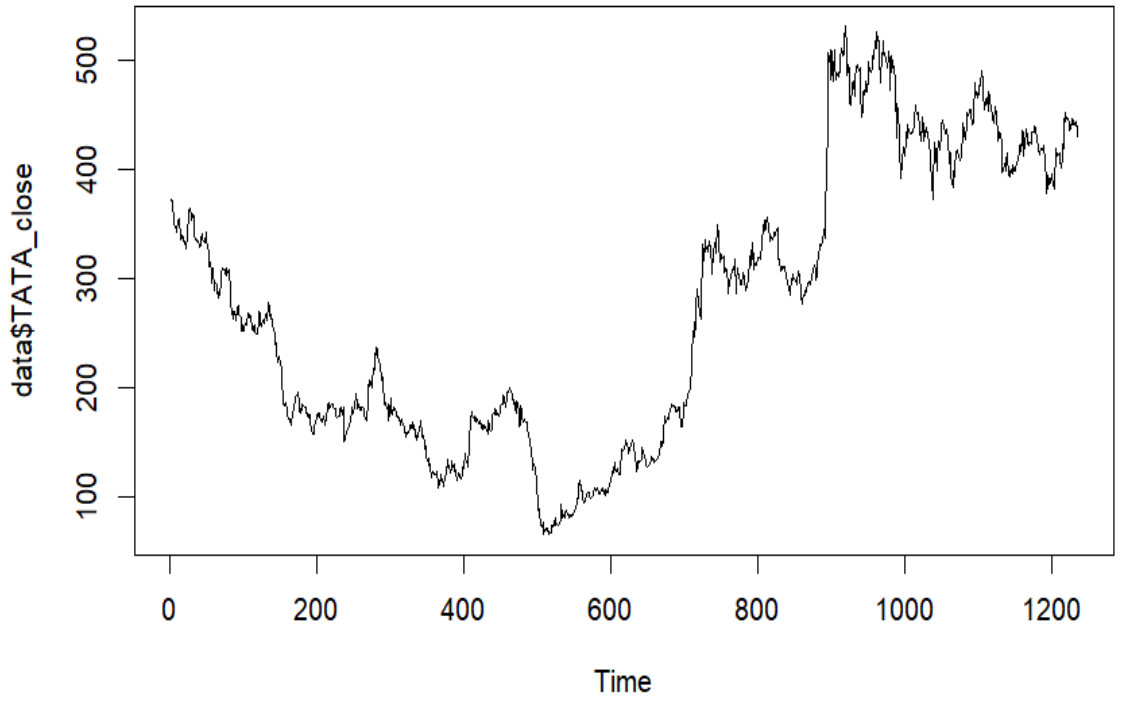
DATA EXPLORATION

TIME SERIES MODELS USED FOR PREDICTION

FUTURE ENHANCEMENTS

Dataset

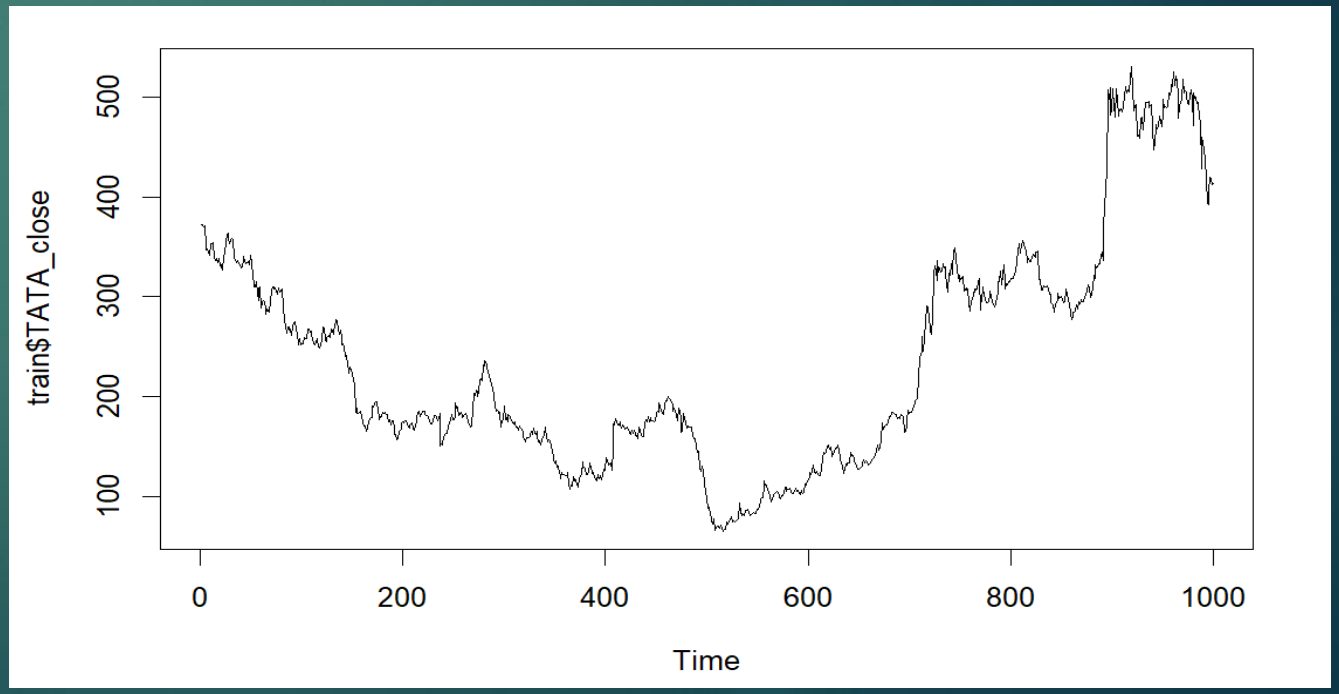
- Stocks considered for study:
  - TATA Motors
  - Maruti Suzuki
  - Mahindra & Mahindra
- Source: Yahoo Finance
- Data period: 5 years [02/26/2018 - 02/22/2023]
- Variables:
  - Closing Price
  - Daily returns
  - Volume Traded
- Variable of interest:
  - TATA Motors Closing Price



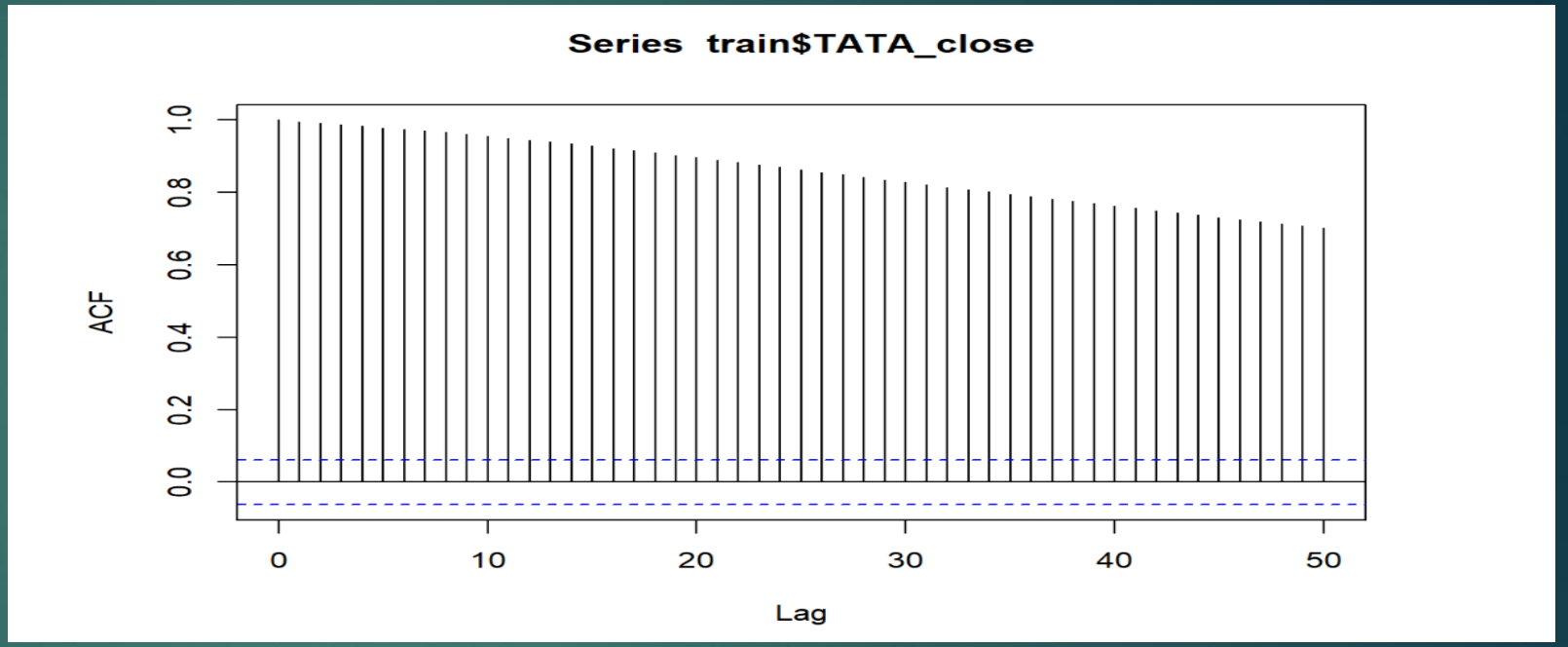
Data Exploration

- Training data:
  - 4 years data (1000 entries)
- Test Data:
  - 1 year data (last 235 entries)

- Inspecting Time series plot



TS & ACF of Actual data



- ACF Inference:
  - Non- Stationary data
  - Hence not a white noise

Box –Ljung Test

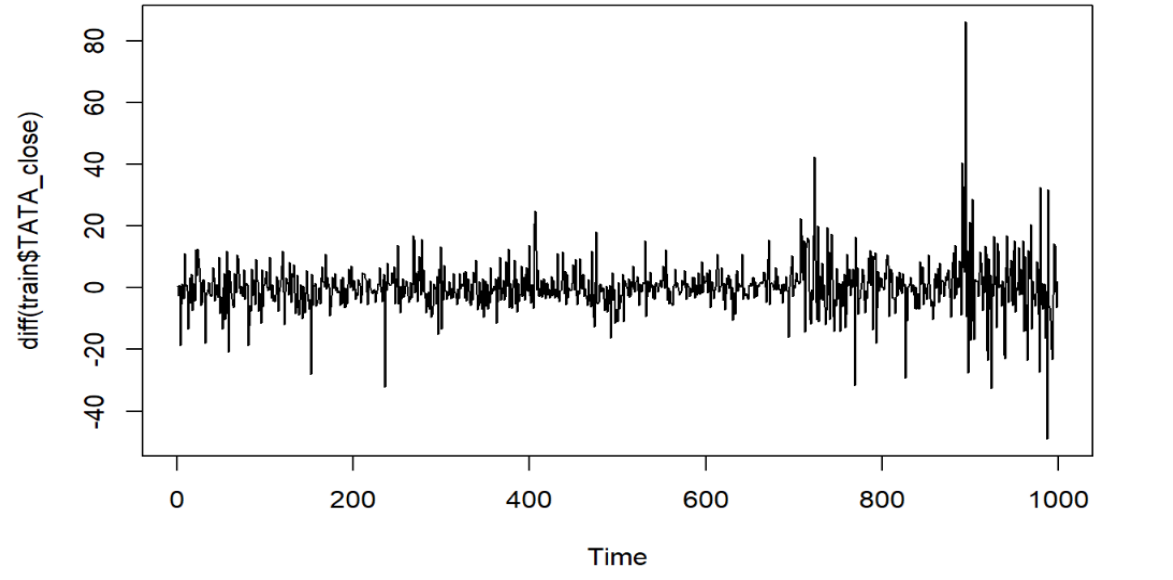
H_0: White Noise

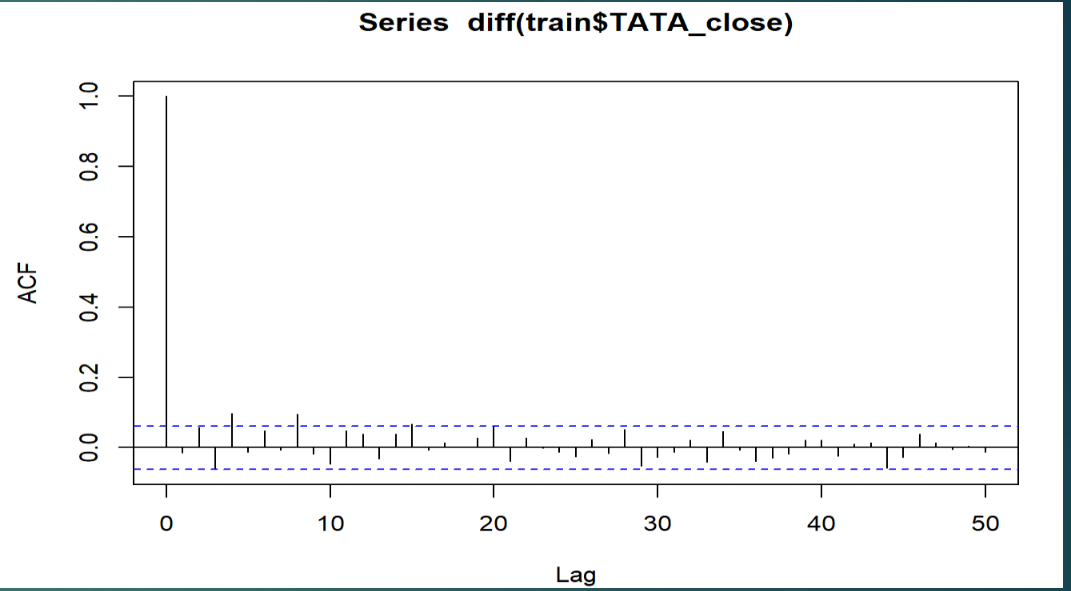
H_a: Not White Noise

```
##
##  Box-Ljung test
##
## data:  train$TATA_close
## X-squared = 37992, df = 50, p-value < 2.2e-16
```

Inducing Stationarity (1st Differencing)

The 1st difference of dependent variable shows stationarity





Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

```
##
##  Box-Ljung test
##
## data:  diff(train$TATA_close)
## X-squared = 46.072, df = 20, p-value = 0.0007878
```

TIME SERIES MODELS

- Multiple Regression Model
- Polynomial Model
- AR Model
- MA model
- Arima model
- ARCH/GARCH model

MULTIPLE LINEAR REGRESSION

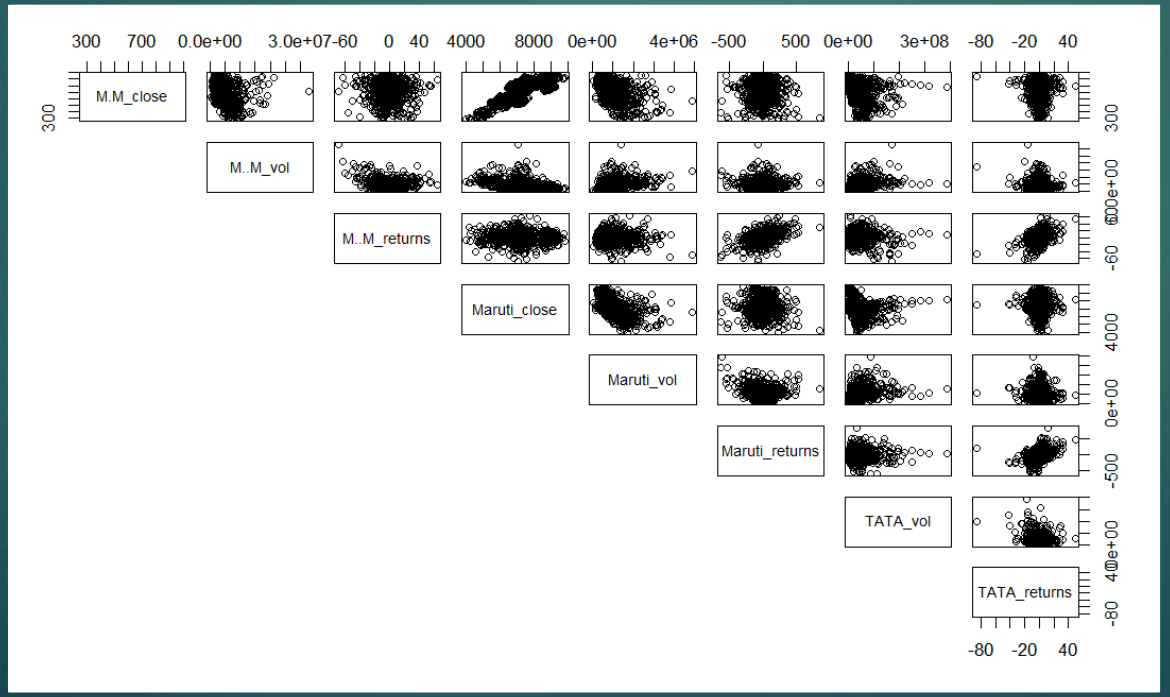
Model	RSquared	ResidualError	AIC	BIC
MLR Model	0.5873	74.32	11465.63	11514.7

```
##
## Call:
## lm(formula = TATA_close ~ ., data = newdata)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -137.119  -50.272   -9.492   36.047  194.245
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.345e+02  2.334e+01  -5.759 1.13e-08 ***
## M.M_close       6.213e-01  2.598e-02  23.918  < 2e-16 ***
## M..M_vol       -1.083e-06  9.888e-07  -1.095   0.2738
## M..M_returns    3.012e-01  2.069e-01   1.455   0.1459
## Maruti_close   -7.996e-03  4.035e-03  -1.981   0.0478 *
## Maruti_vol     -3.062e-06  5.144e-06  -0.595   0.5519
## Maruti_returns -7.536e-03  2.019e-02  -0.373   0.7090
## TATA_vol       -7.957e-09  7.047e-08  -0.113   0.9101
## TATA_returns   -5.004e-01  3.551e-01  -1.409   0.1591
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 74.32 on 991 degrees of freedom
## Multiple R-squared:  0.5906, Adjusted R-squared:  0.5873
## F-statistic: 178.7 on 8 and 991 DF,  p-value: < 2.2e-16
```

Multicollinearity Check

Correlation Matrix

	M.M_close	M..M_vol	M..M_returns	Maruti_close	Maruti_vol	Maruti_returns	TATA_vol	TATA_returns
M.M_close	1.00000000	-0.26122153	-0.04849320	0.79215140	-0.45389489	-0.01221074	-0.22851005	-0.03599352
M..M_vol	-0.26122153	1.00000000	-0.23895641	-0.35924779	0.42858538	-0.05661977	0.43579032	-0.12325773
M..M_returns	-0.04849320	-0.23895641	1.00000000	-0.04769979	-0.03356211	0.47714156	-0.08860927	0.47481060
Maruti_close	0.79215140	-0.35924779	-0.04769979	1.00000000	-0.48959527	-0.06158669	-0.27933125	-0.02386186
Maruti_vol	-0.45389489	0.42858538	-0.03356211	-0.48959527	1.00000000	-0.13999610	0.36666850	-0.04221662
Maruti_returns	-0.01221074	-0.05661977	0.47714156	-0.06158669	-0.13999610	1.00000000	-0.04267306	0.40372765
TATA_vol	-0.22851005	0.43579032	-0.08860927	-0.27933125	0.36666850	-0.04267306	1.00000000	-0.21722555
TATA_returns	-0.03599352	-0.12325773	0.47481060	-0.02386186	-0.04221662	0.40372765	-0.21722555	1.00000000



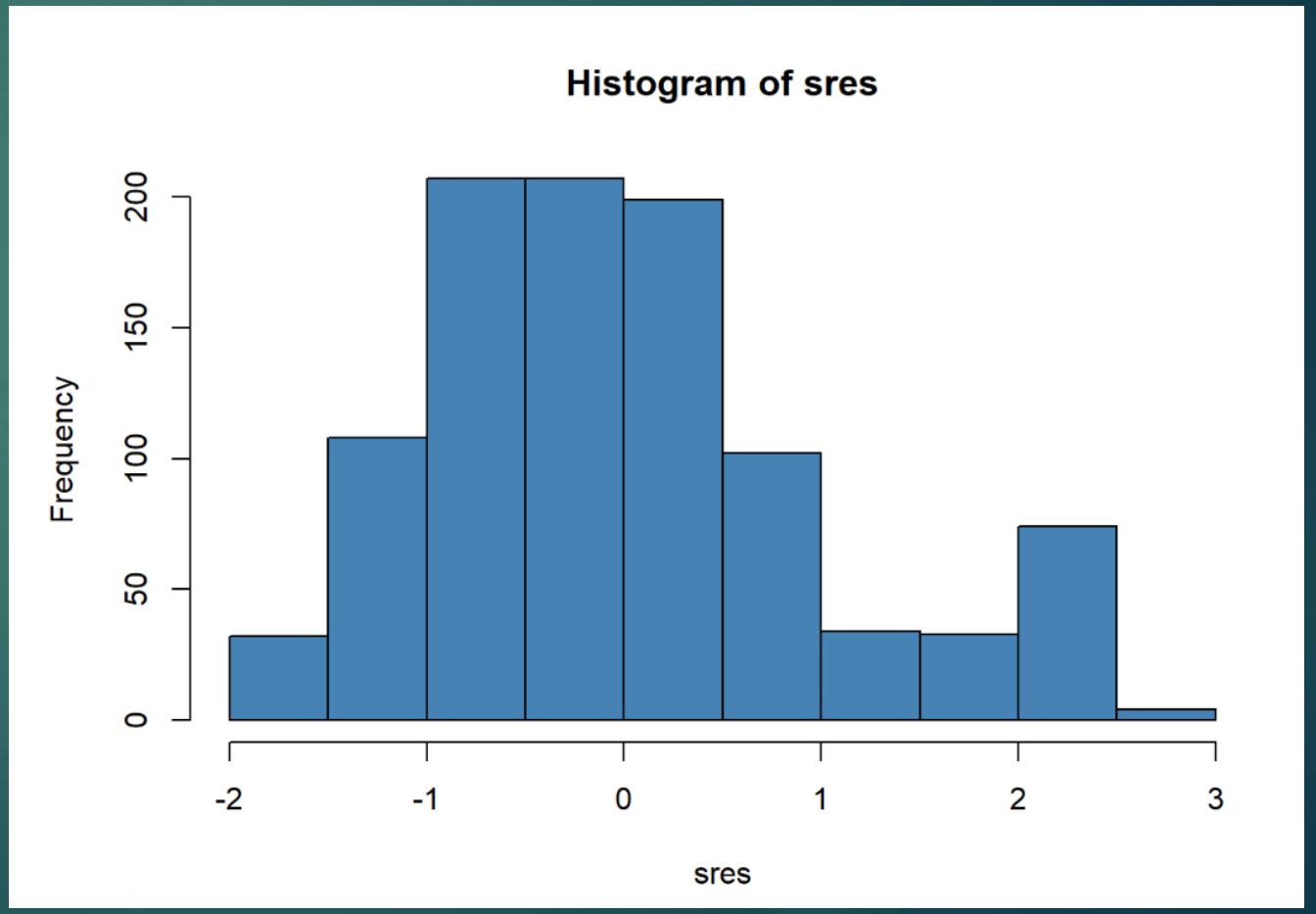
Normality Test

Hypothesis:
H_0 : Series are normally distributed
H_a : Series are not normally distributed

Shapiro-Wilk Normality Test

```
#Normality test
library(nortest)
shapiro.test(sres)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sres
## W = 0.94742, p-value < 2.2e-16
```



Heteroscedasticity Test

Hypothesis:
H_0 : No Heteroscedascity (constant variance)
H_a : Heteroscedasticity(non−constant variance)

White Test

```
white(mlr,interactions = TRUE)
```

```
## # A tibble: 1 × 5
##   statistic  p.value parameter method       alternative
##       <dbl>    <dbl>     <dbl> <chr>        <chr>
## 1      399. 8.54e-59        44 White's Test greater
```

Auto Correlation Test

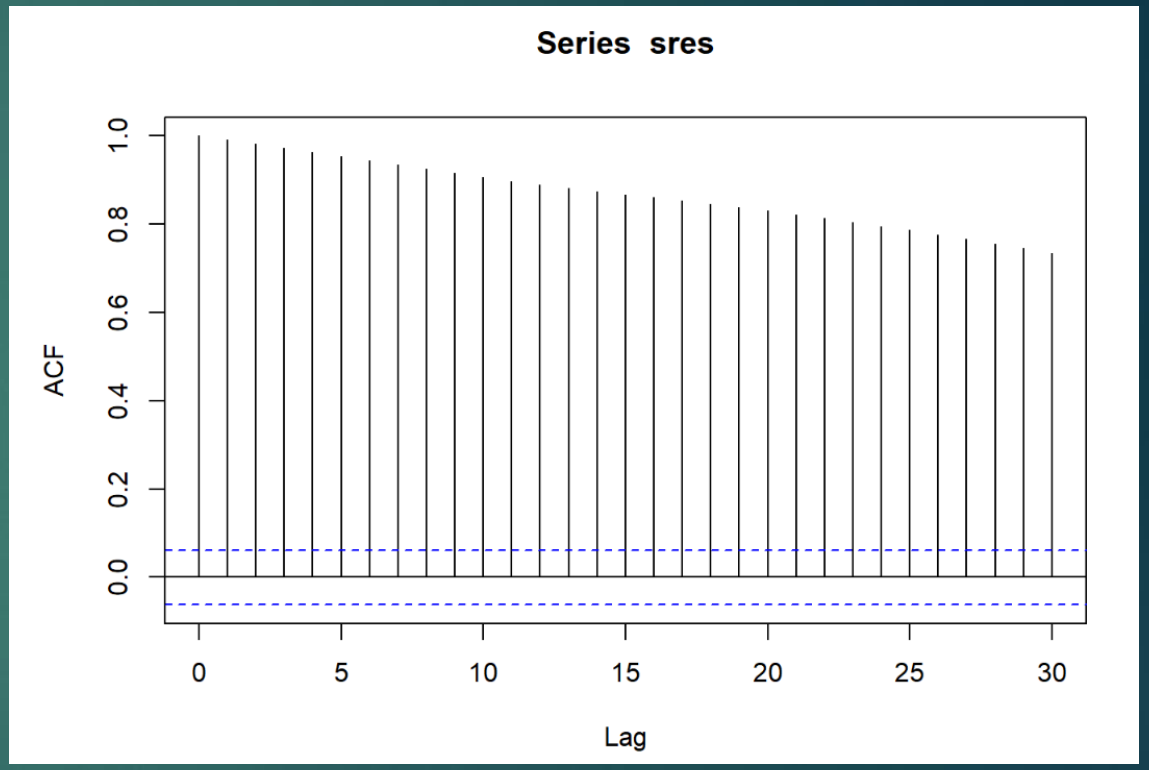
- Visual Test
  - Autocorrelation exists (the residuals are not white noise)

Hypothesis:
H_0 : All Auto Correlations are 0
H_a : : At least one autocorrelation is different than 0

Box-Ljung Test

```
Box.test(sres,type = "Ljung", lag=20)
```

```
##
##  Box-Ljung test
##
## data:  sres
## X-squared = 16668, df = 20, p-value < 2.2e-16
```



Deterministic Time Series Model


```
#Deterministic time series model
#hours=rep(c(1:24),times=412)[1:9886]
time=seq(1:length(diff(train$TATA_close)))
fit<-lm(diff(TATA_close)~time,data=train)
summary(fit)
```

```
##
## Call:
## lm(formula = diff(TATA_close) ~ time, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -50.003  -3.355  -0.013   2.963  85.269
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.8943236  0.5016593  -1.783   0.0749 .
## time         0.0018719  0.0008691   2.154   0.0315 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.922 on 997 degrees of freedom
## Multiple R-squared:  0.004631,   Adjusted R-squared:  0.003633
## F-statistic: 4.639 on 1 and 997 DF,  p-value: 0.03149
```

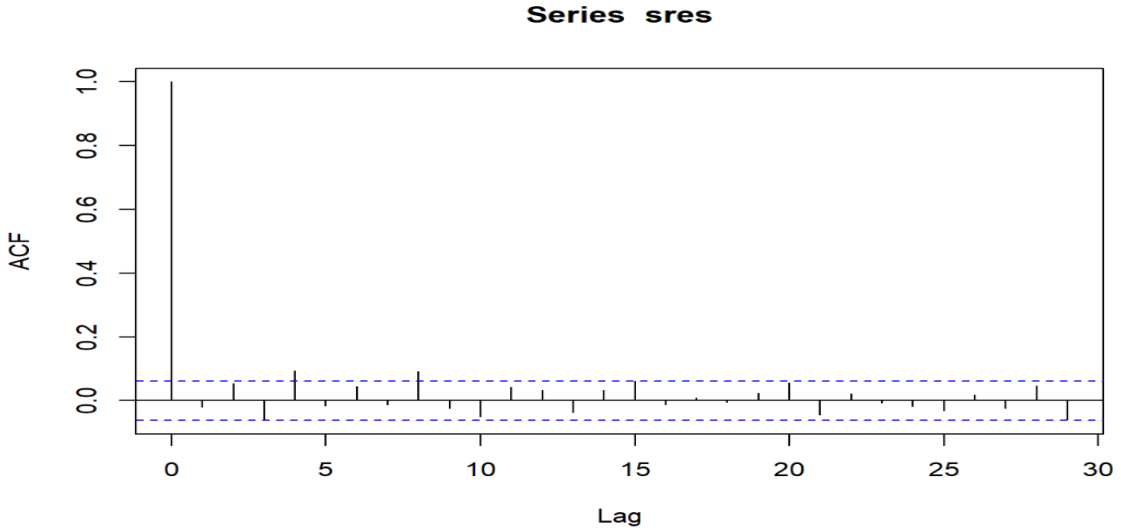
Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

```
Box.test(sres,type = "Ljung", lag=20)
```

```
##
##  Box-Ljung test
##
## data:  sres
## X-squared = 43.105, df = 20, p-value = 0.00198
```



Polynomial Model

K=8

```
Call:
lm(formula = diff(TATA_close) ~ poly(time, k), data = train)

Residuals:
    Min      1Q  Median      3Q     Max
-45.869  -3.326  -0.081   2.940  84.460

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.04164    0.25007   0.167   0.8678
poly(time, k)1  17.06268    7.90399   2.159   0.0311 *
poly(time, k)2  -8.71831    7.90399  -1.103   0.2703
poly(time, k)3 -10.07358    7.90399  -1.274   0.2028
poly(time, k)4 -13.24779    7.90399  -1.676   0.0940 .
poly(time, k)5  -8.62130    7.90399  -1.091   0.2756
poly(time, k)6  -4.89378    7.90399  -0.619   0.5360
poly(time, k)7 -10.53521    7.90399  -1.333   0.1829
poly(time, k)8 -12.62817    7.90399  -1.598   0.1104
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.904 on 990 degrees of freedom
Multiple R-squared:  0.01611,   Adjusted R-squared:  0.008162
F-statistic: 2.027 on 8 and 990 DF,  p-value: 0.04056
```

K=1

```
##
## Call:
## lm(formula = diff(TATA_close) ~ poly(time, k), data = train)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -50.003  -3.355  -0.013   2.963  85.269
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.04164    0.25064   0.166   0.8681
## poly(time, k) 17.06268    7.92201   2.154   0.0315 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.922 on 997 degrees of freedom
## Multiple R-squared:  0.004631,   Adjusted R-squared:  0.003633
## F-statistic: 4.639 on 1 and 997 DF,  p-value: 0.03149
```

Polynomial – White Noise Check

- ACF Plot
  - There is no white noise, but need to run test

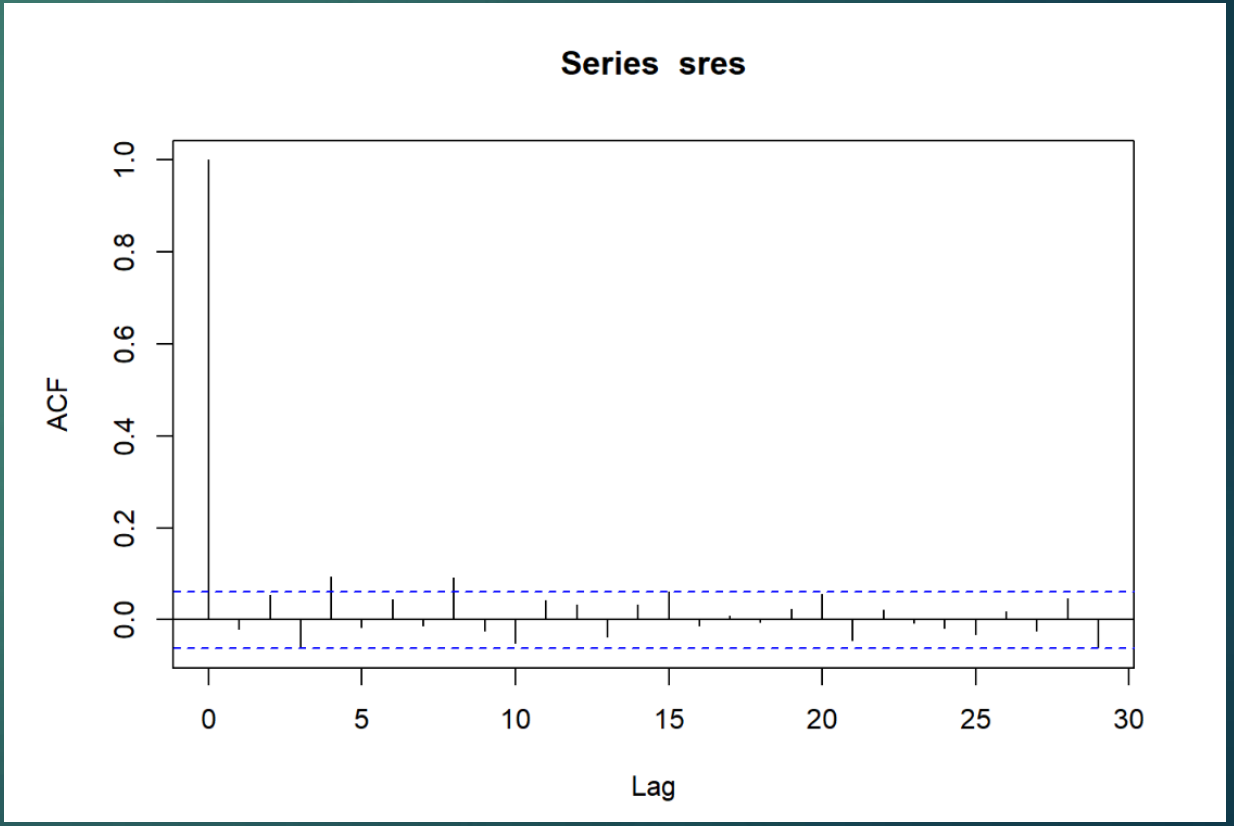
Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

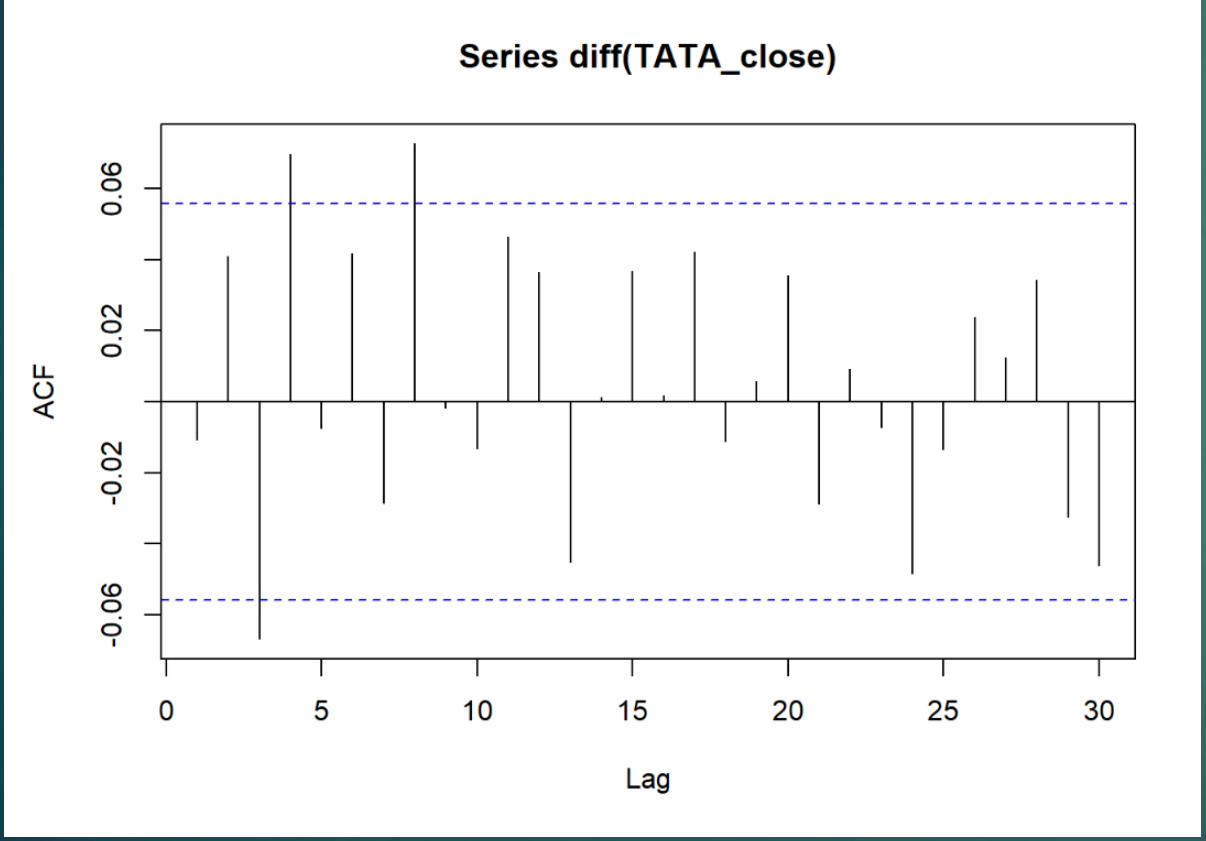
```
Box.test(sres,type = "Ljung", lag=20)
```

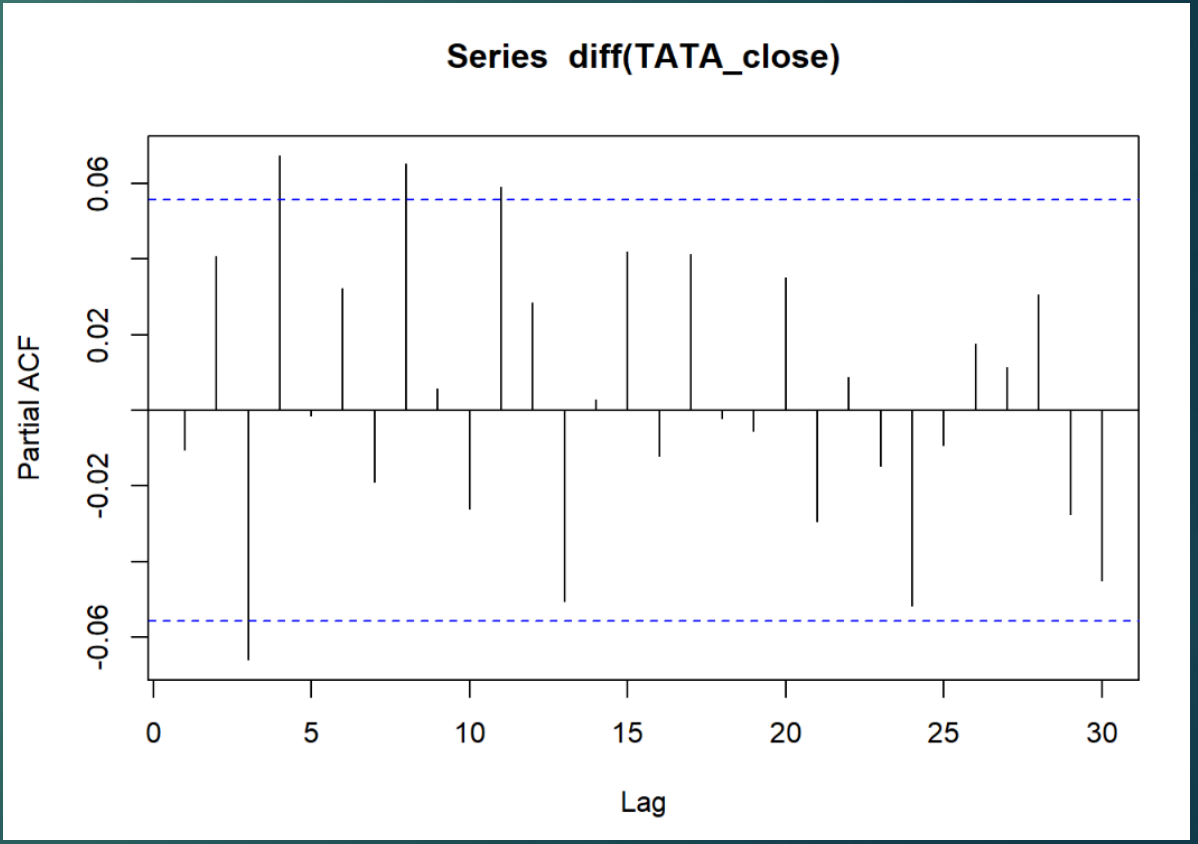
```
##
##  Box-Ljung test
##
## data:  sres
## X-squared = 43.105, df = 20, p-value = 0.00198
```



Stochastic time series model

Check for AR/MA model

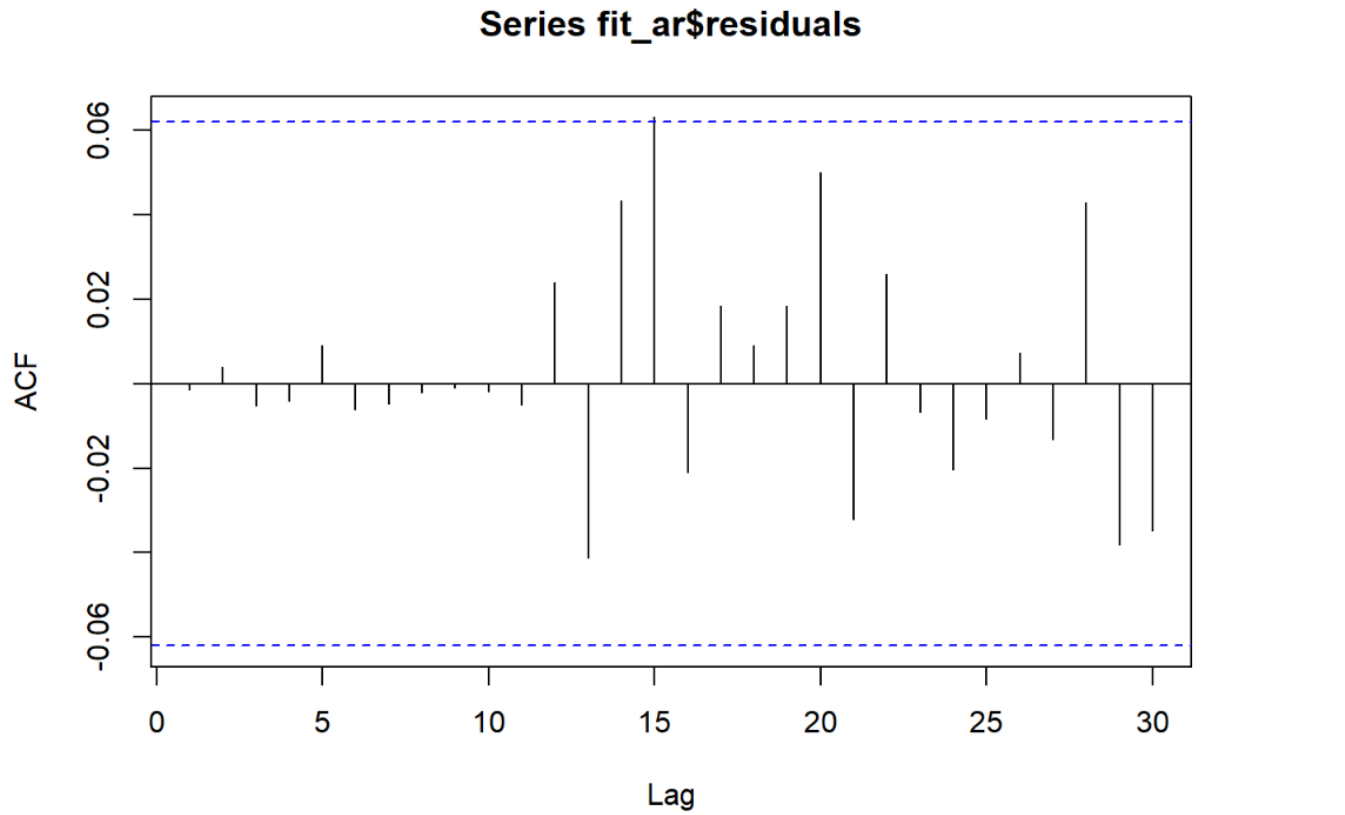




AR (11) Model

- ACF Plot
  - There is white noise

```
fit_ar = arima(train$TATA_close,order= c(11,1,0))
acf(fit_ar$residuals)
```



Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

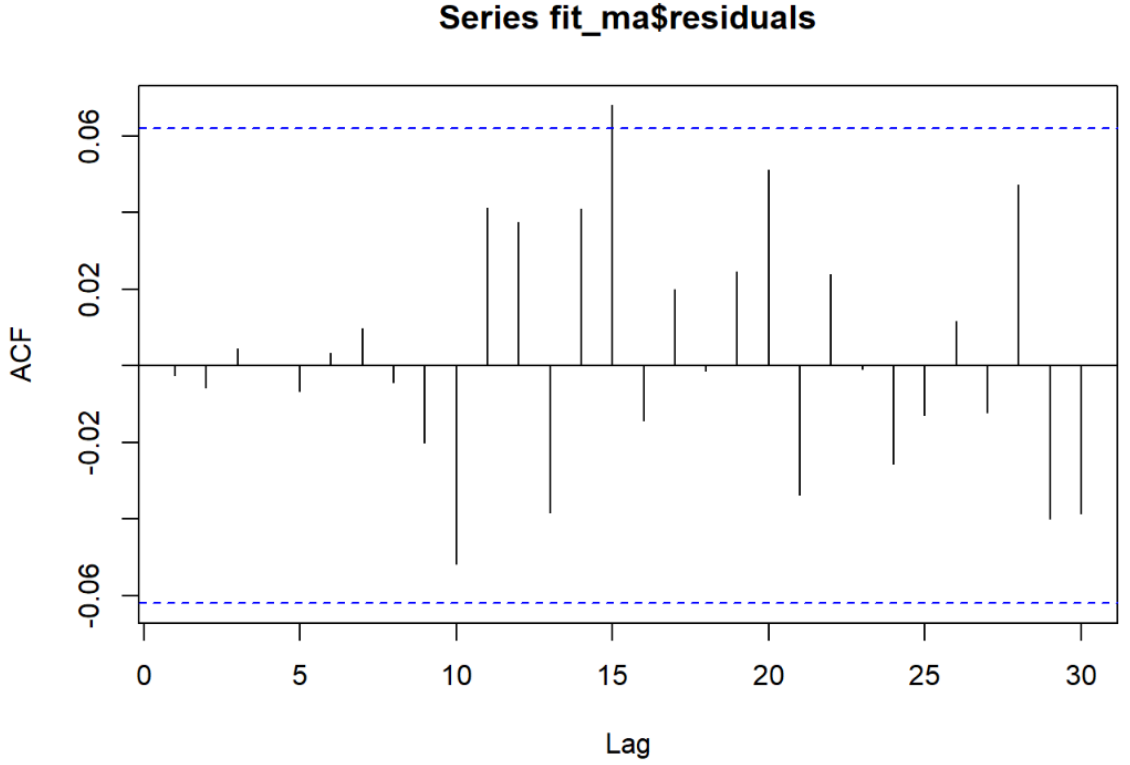
```
Box.test(fit_ar$residuals, lag=20)
```

```
##
## data:  fit_ar$residuals
## X-squared = 12.074, df = 20, p-value = 0.9135
```

MA(8) Model

- ACF Plot
  - There is a spike, but we will do Box-Ljung test

```
fit_ma = arima(train$TATA_close, order = c(0,1,8))
acf(fit_ma$residuals)
```



Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

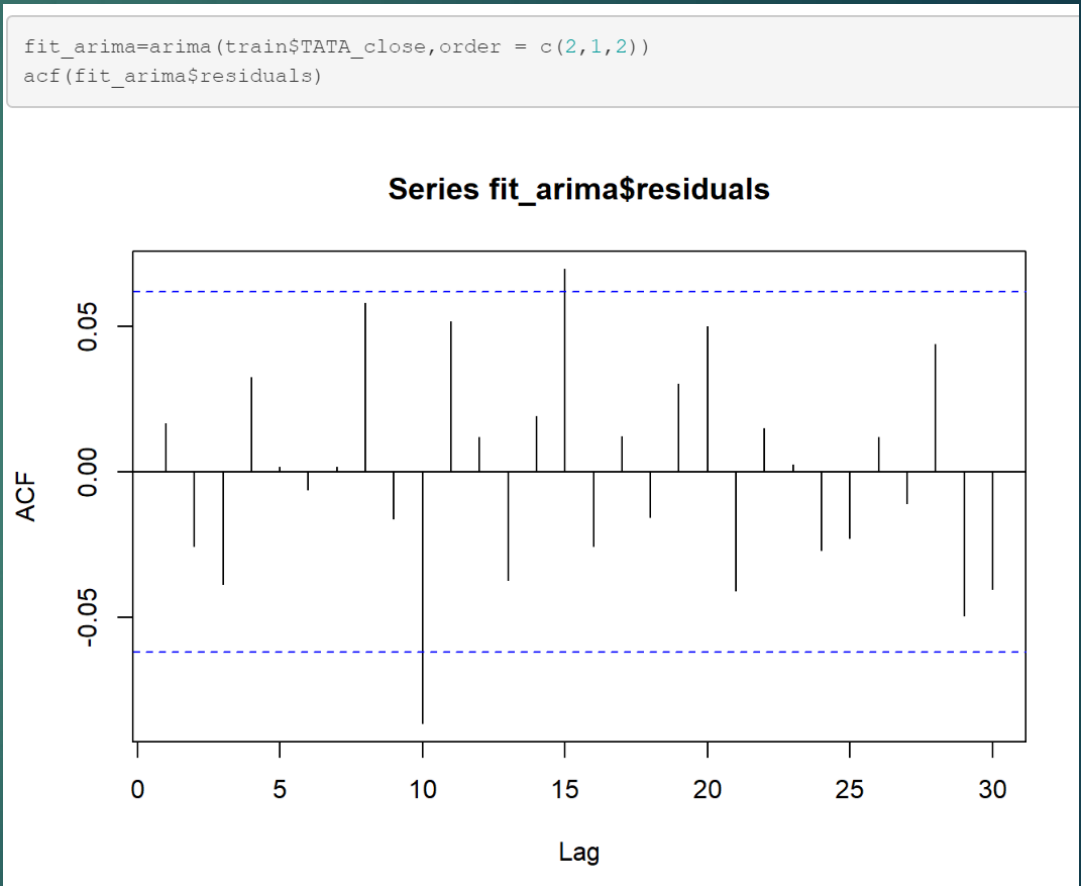
```
Box.test(fit_ar$residuals, lag=20)
```

```
##
## data:  fit_ar$residuals
## X-squared = 12.074, df = 20, p-value = 0.9135
```

ARIMA(2,2) model

- ACF Plot
  - There is a spike, but we will do Box-Ljung test

```
fit_arima=arima(train$TATA_close,order = c(2,1,2))
acf(fit_arima$residuals)
```



Box –Ljung Test

H_0: White Noise

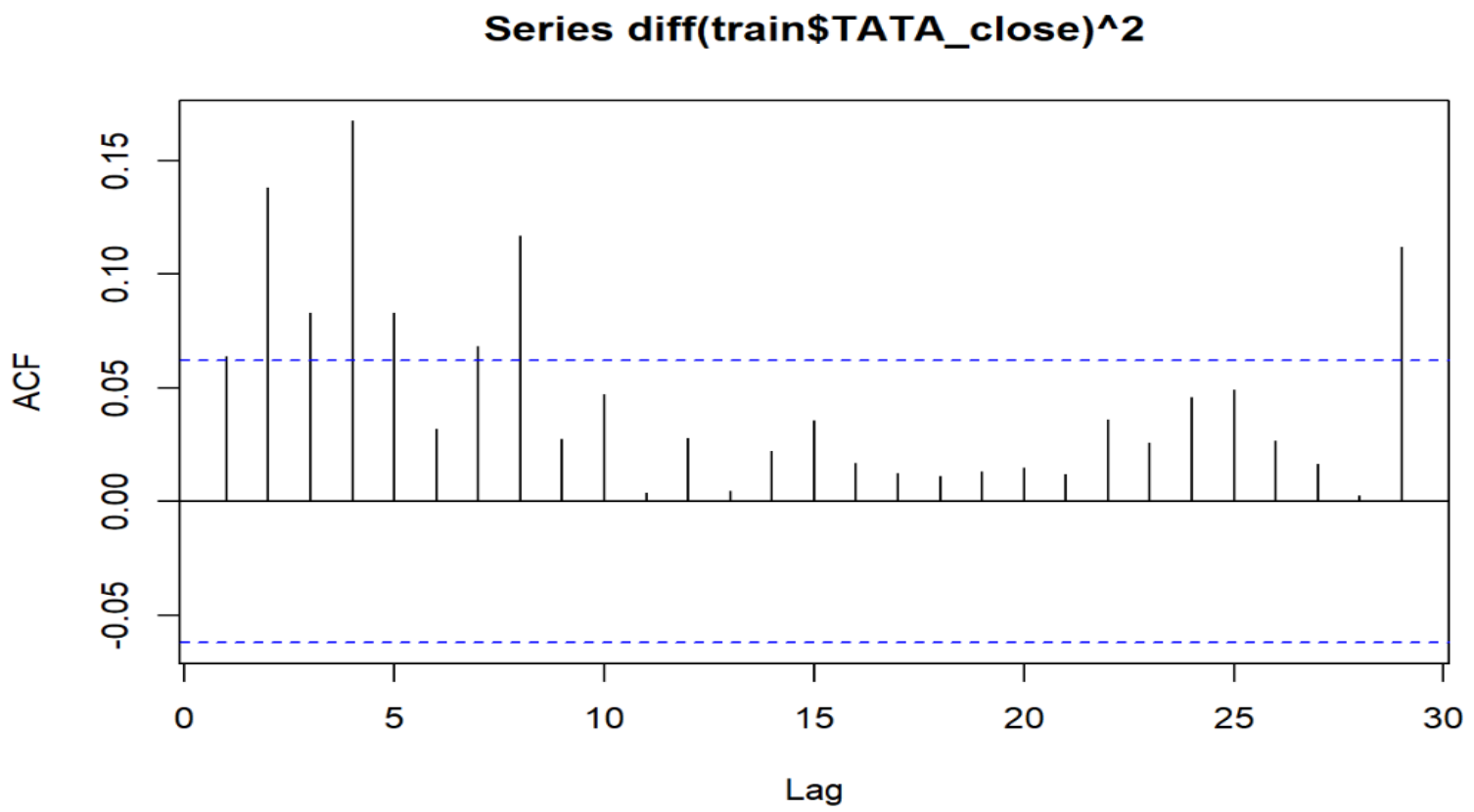
H_a: Not White Noise

```
Box.test(fit_arima$residuals, lag=20)
```

```
##
##  Box-Pierce test
##
## data:  fit_arima$residuals
## X-squared = 28.635, df = 20, p-value = 0.0952
```

CV test for GARCH


```
acf(diff(train$TATA_close)^2)
```



GARCH Model

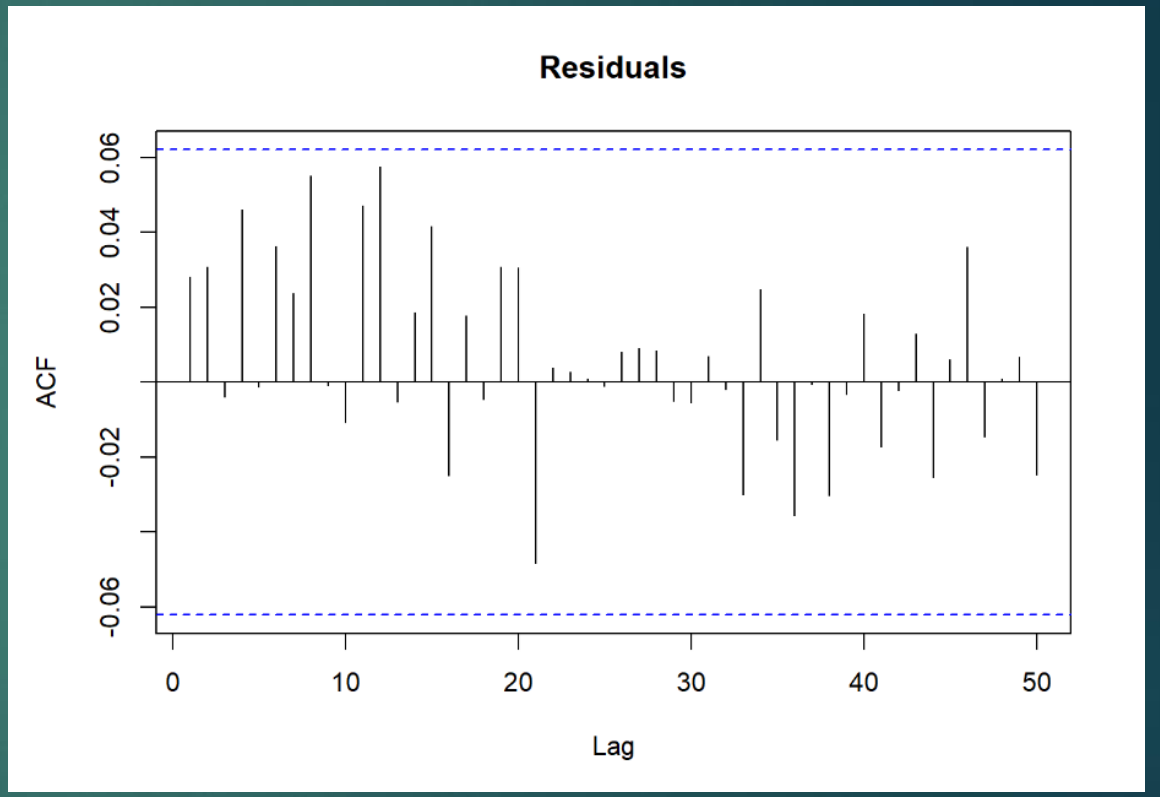

```
garch.fit<-garchFit(~garch(1,1),data=diff(train$TATA_close), trace=FALSE)
summary(garch.fit)
```

```
##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~garch(1, 1), data = diff(train$TATA_close),
##     trace = FALSE)
##
## Mean and Variance Equation:
##  data ~ garch(1, 1)
## <environment: 0x0000016d5d12fb58>
##  [data = diff(train$TATA_close)]
##
## Conditional Distribution:
##  norm
##
## Coefficient(s):
##        mu      omega     alpha1      beta1
## -0.142156   1.985680   0.095764   0.874475
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      -0.14216     0.19813   -0.717  0.47307
## omega    1.98568     0.75290    2.637  0.00836 **
## alpha1   0.09576     0.02189    4.375 1.22e-05 ***
## beta1    0.87447     0.03106   28.152  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Log Likelihood:
##  -3341.858    normalized:  -3.345203
##
## Description:
##  Wed Mar  8 00:36:19 2023 by user: Dikshant
##
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  1748.675  0
##  Shapiro-Wilk Test  R    W      0.925103  0
##  Ljung-Box Test     R    Q(10)  8.967408  0.5351995
##  Ljung-Box Test     R    Q(15)  16.6957   0.3373786
##  Ljung-Box Test     R    Q(20)  19.58289  0.4842785
##  Ljung-Box Test     R^2  Q(10)  10.45774  0.4012919
##  Ljung-Box Test     R^2  Q(15)  14.45905  0.4910372
##  Ljung-Box Test     R^2  Q(20)  15.51204  0.7463975
##  LM Arch Test       R    TR^2   11.66329  0.473086
##
## Information Criterion Statistics:
##      AIC      BIC      SIC     HQIC
## 6.698414 6.718061 6.698382 6.705881
```

Residual test for GARCH

- ACF Plot
  - There is white noise



Box –Ljung Test

H_0: White Noise

H_a: Not White Noise

```
Box.test(garch.fit@residuals/garch.fit@sigma.t, lag=50)
```

```
##
## data:  garch.fit@residuals/garch.fit@sigma.t
## X-squared = 29.586, df = 50, p-value = 0.9905
```

Squared Residual test for GARCH

- ACF Plot
  - There is a spike, but we will do Box-Ljung test

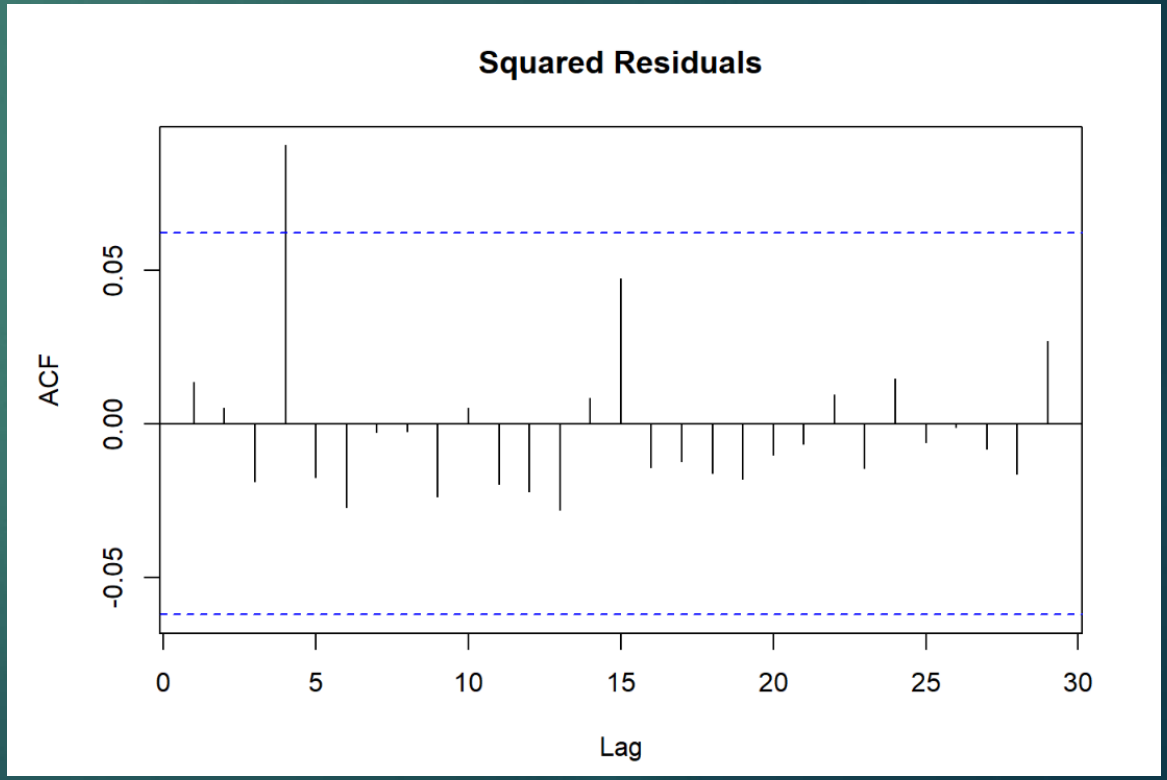
Box –Ljung Test

H_0: White Noise

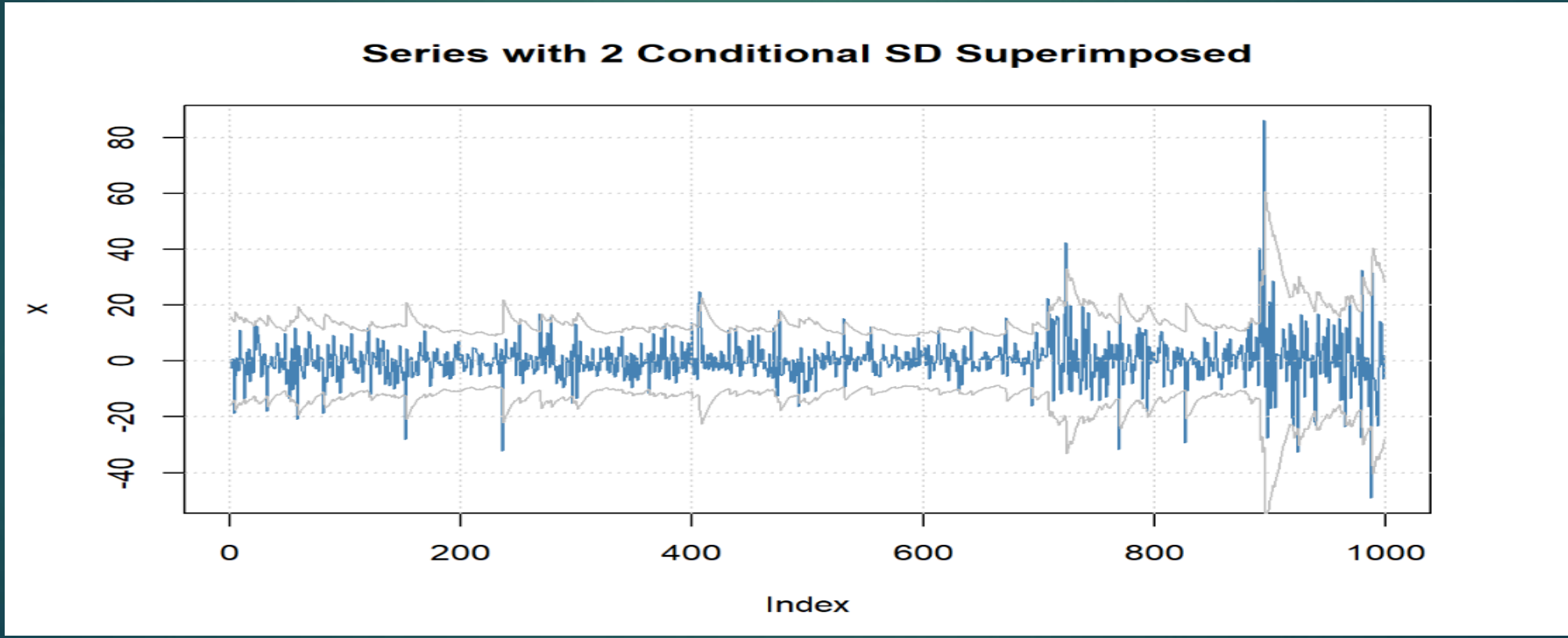
H_a: Not White Noise

```
Box.test((garch.fit@residuals/garch.fit@sigma.t)^2, lag=10)
```

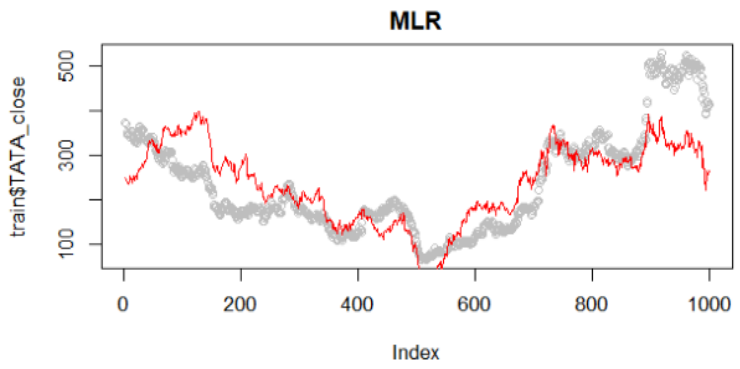
```
##
## data:  (garch.fit@residuals/garch.fit@sigma.t)^2
## X-squared = 10.391, df = 10, p-value = 0.4069
```

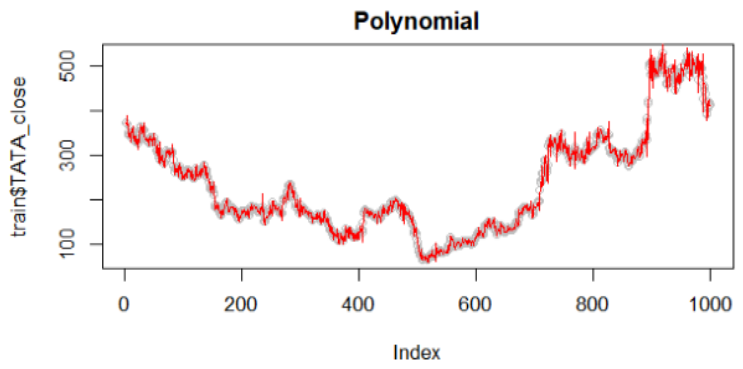


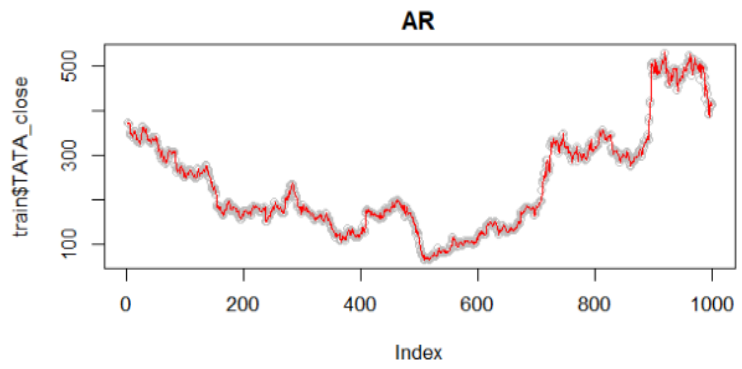
GARCH Fit plot

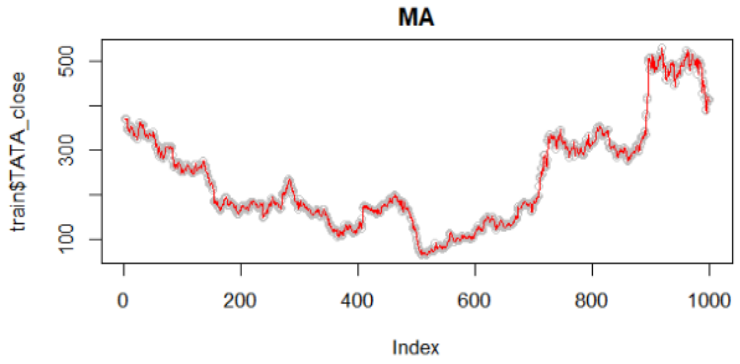


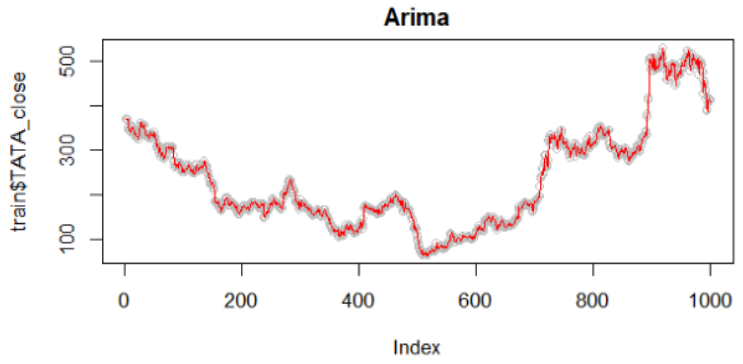
Prediction Plots









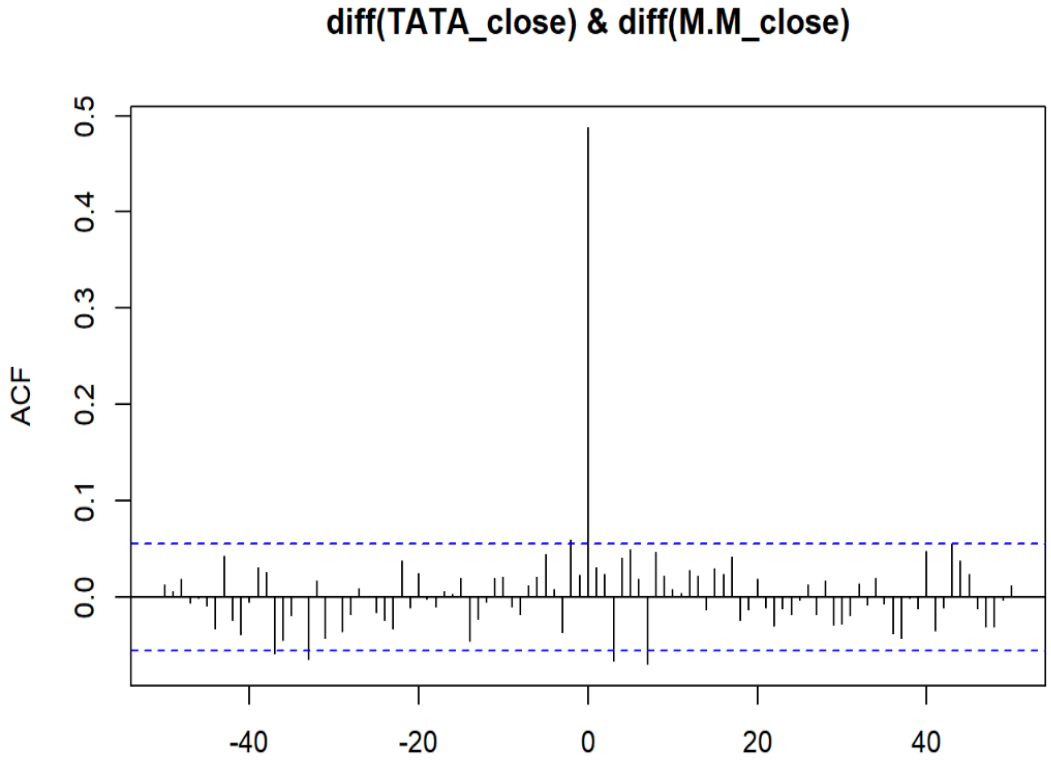


VARMA (1,0) Model


```
Number of parameters:  6
initial estimates:  0.1554 0.906 0.9972 8e-04 0.0056 0.9975
Par. lower-bounds:  -1.478 -2.2091 0.9913 -0.0023 -0.0055 0.9915
Par. upper-bounds:  1.7889 4.0212 1.003 0.004 0.0168 1.0035
Final   Estimates:  0.1554464 0.9060499 0.9971843 0.0008350342 0.005644754 0.9974993

Coefficient(s):
            Estimate  Std. Error  t value Pr(>|t|)
TATA_close  0.155446    0.815889    0.191    0.849
M.M_close   0.906050    1.556291    0.582    0.560
TATA_close  0.997184    0.002919  341.656   <2e-16 ***
M.M_close   0.000835    0.001579    0.529    0.597
TATA_close  0.005645    0.005567    1.014    0.311
M.M_close   0.997499    0.003013  331.098   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
---
Estimates in matrix form:
Constant term:
Estimates:  0.1554464 0.9060499
AR coefficient matrix
AR( 1 )-matrix
        [,1]    [,2]
[1,] 0.99718 0.000835
[2,] 0.00564 0.997499

Residuals cov-matrix:
         [,1]      [,2]
[1,] 64.12775  59.73347
[2,] 59.73347 233.24693
----
aic=  9.350165
bic=  9.375034
```



MAPE & AIC

Comparison of Mean of Absolute percentage Error(MAPE) & AIC for all the models

Model	MAPE	AIC
MLR	0.2935636	11465.627
Polynomial	0.9971848	6974.188
AR(11)	0.0502220	6965.529
MA(8)	0.0498548	6965.154
ARIMA(2,2)	0.0617765	6967.379
Garch	0.0500251	6698.000

Future Scope

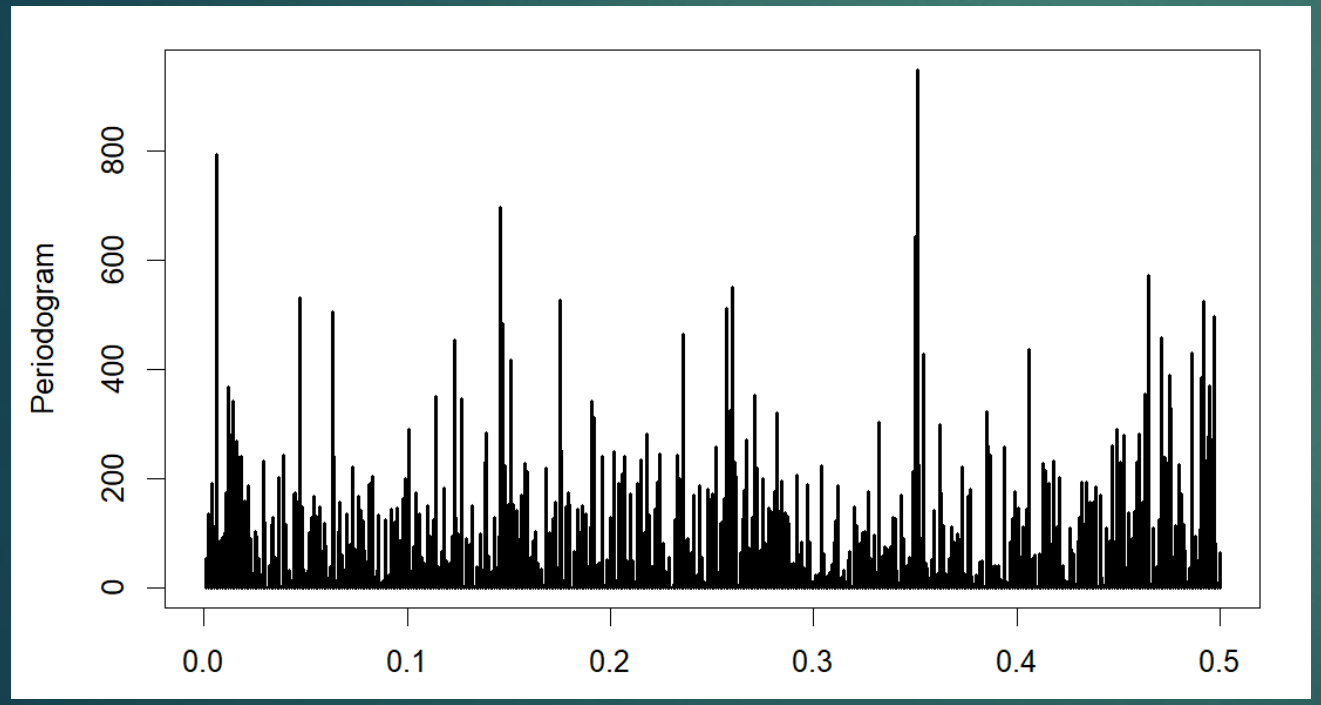
We would try and predict daily closing price for next few months for other automobile stocks in the study.

Thank you

Any questions?

Thank you

Any questions?


Periodogram
0
200
400
600
800
0.0
0.1
0.2
0.3
0.4
0.5