Description :

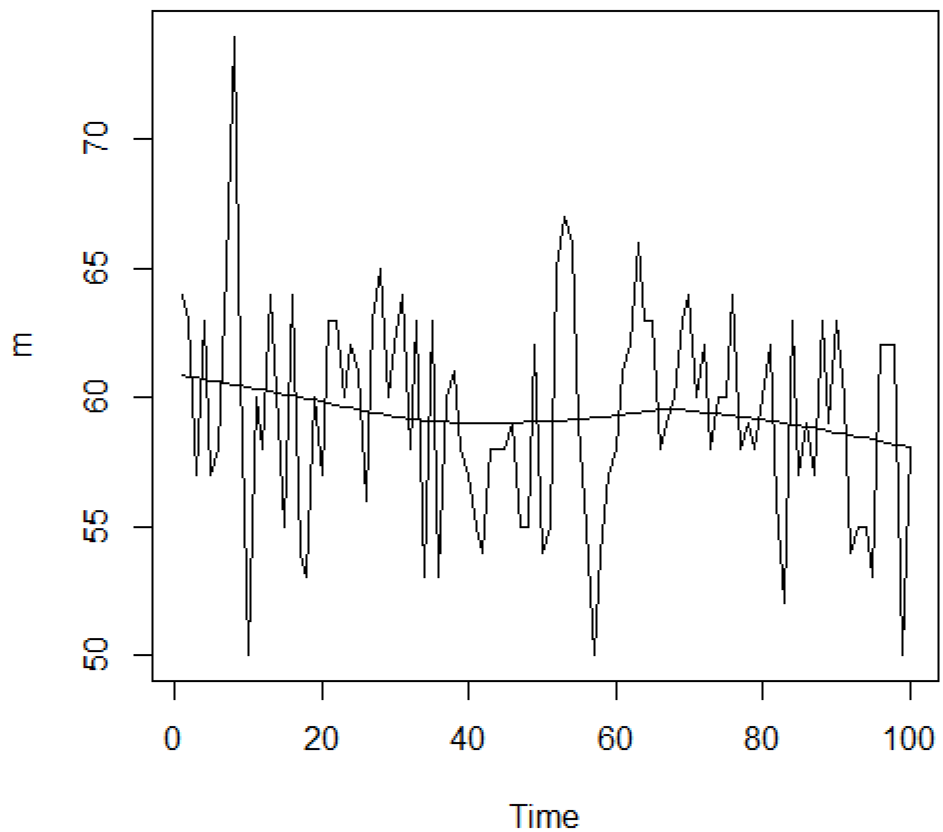
The Data describe Rockwell hardness (measured on rockwell –B– scale), 100 coils produced in sequence at a Chicago Steel Mill .

Subject :
“ Industry “

Source :
“ Roberts (1992) “

```
> m = tsdl [[601]]

> attributes(m)
```

Plot The Time Series :




```
> plot(n)

> lines(n)

> lines(lowess(n))
```

- We plot The time series to know is there trend in our data or not and to know the seasonality .

As we see In this data it appear there is no seasonal component in the set and its stationarity .

Testing Stationarity :

```
> adf.test(m)


	Augmented Dickey-Fuller Test


data:  m
Dickey-Fuller = -4.7037, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

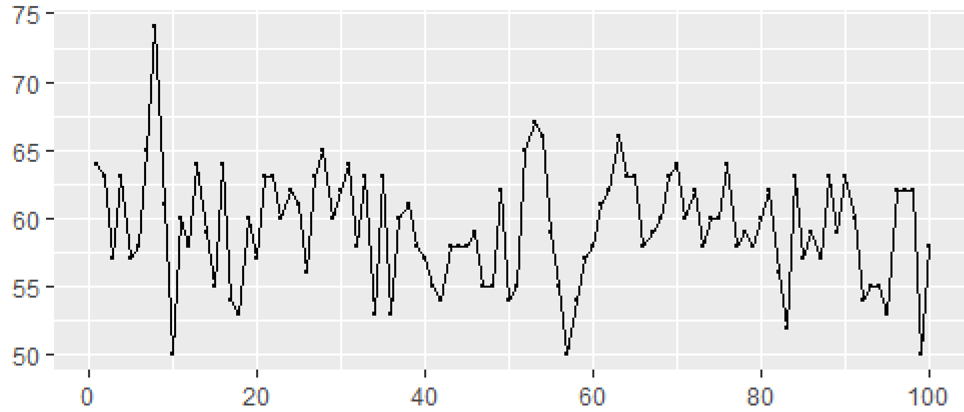
We can test the stationary of the series from this command > adf.test(m) and the Hypothesis will be :

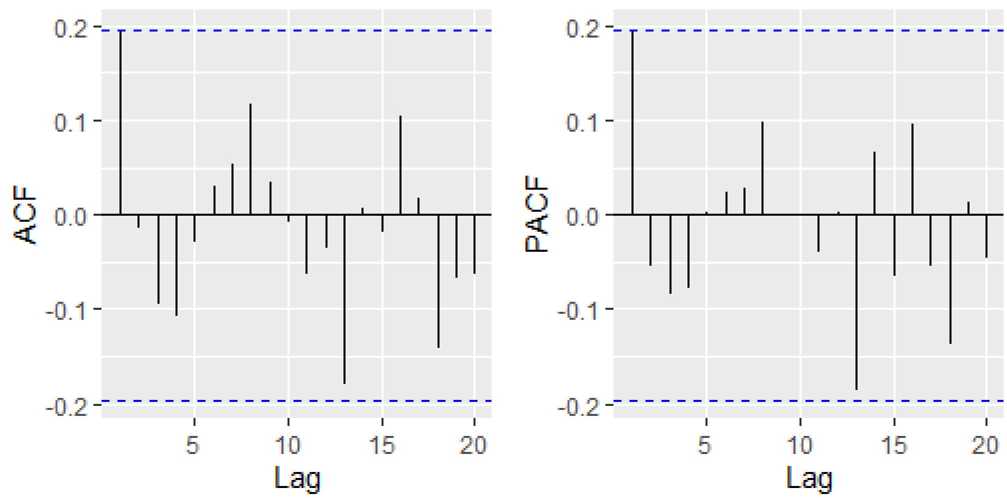
$$H_0 : The\ series\ is\ not\ stationary\ .$$

$$H_1 : The\ series\ is\ stationary\ .$$

So since the $\mathbf{P\,.\,Value < \alpha\,,P\,.\,Value = 0.01}$ and $\boldsymbol{\alpha} = \mathbf{0.05}$ so we Reject H_0 that's mean the series is stationary

Model Type :





From the AutoCorrelation (ACF) graph and the Parial AutoCorrelation (PACF) graph we can nominate two models are suggested :

AR(1) ,MA(1) , ARMA(1,1).

Testing Coeffiicients :

For model AR(1)

```
> fit1 = arima(m,order=c(1,0,0))
> coeftest(fit1)


z test of coefficients:


          Estimate Std. Error  z value Pr(>|z|)
ar1        0.19704    0.09816   2.0073  0.04471 *
intercept 59.35806    0.50167 118.3203  < 2e-16 ***
--- signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The Hypothesis will be :

$$H_0 : \emptyset_1 = 0 \quad vs \quad H_1 : \emptyset_1 \neq 0$$

Since the $\boldsymbol{P.value's\ in\ all\ parameters < \ \alpha = 0.05}$

So we reject H_0 , that’s mean the value is significantly different from zero .

For model MA(1):

```
> fit2 = arima(m,order=c(0,0,1))
> coeftest(fit2)


z test of coefficients:


          Estimate Std. Error  z value Pr(>|z|)
ma1       0.201523   0.095843   2.1026   0.0355 *
intercept 59.358901   0.484028 122.6353   <2e-16 ***
---Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The Hypothesis will be :

$$H_0 : \theta_1 = 0 \quad vs \quad H_1 : \theta_1 \neq 0$$

Since the $\boldsymbol{P.value's\ in\ all\ parameters} < \boldsymbol{\alpha} = \mathbf{0.05}$

So we reject H_0 , that's mean the value is significantly different from zero .

For Model ARMA(1,1) :

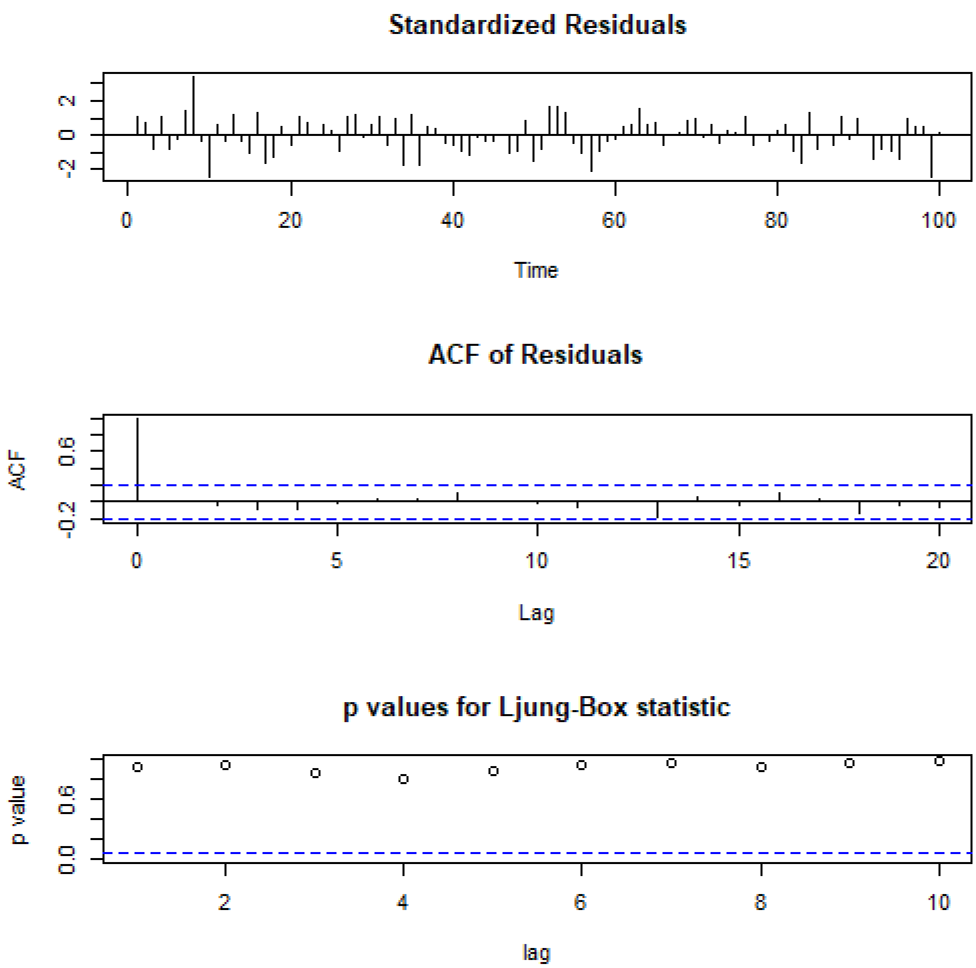
```
> fit3 = arima(m,order=c(1,0,1))
> coeftest(fit3)

z test of coefficients:

          Estimate Std. Error  z value Pr(>|z|)
ar1       0.048759   0.409339   0.1191   0.9052
ma1       0.155681   0.400283   0.3889   0.6973
intercept 59.358957   0.489301 121.3137   <2e-16 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$H_0 : \theta_{1,} \emptyset_1 = 0 \quad vs \quad H_1 : \theta_{1,} \emptyset_1 \neq 0$$

Since the $\boldsymbol{P.value=0.9052}$ & $\boldsymbol{0.6973 > \alpha=0.05}$
So we Accept H_0 , that's mean the value is not significantly .
So we eliminate this model of the suggested models .

Testing Residuals :

For Model AR(1) :



```
> tsdiag(fit1)

Box.test(res,lag=12,type=c("Ljung-Box"))
Box-Ljung test
data: res
X-squared = 3.517, df = 12, p-value = 0.2934
```

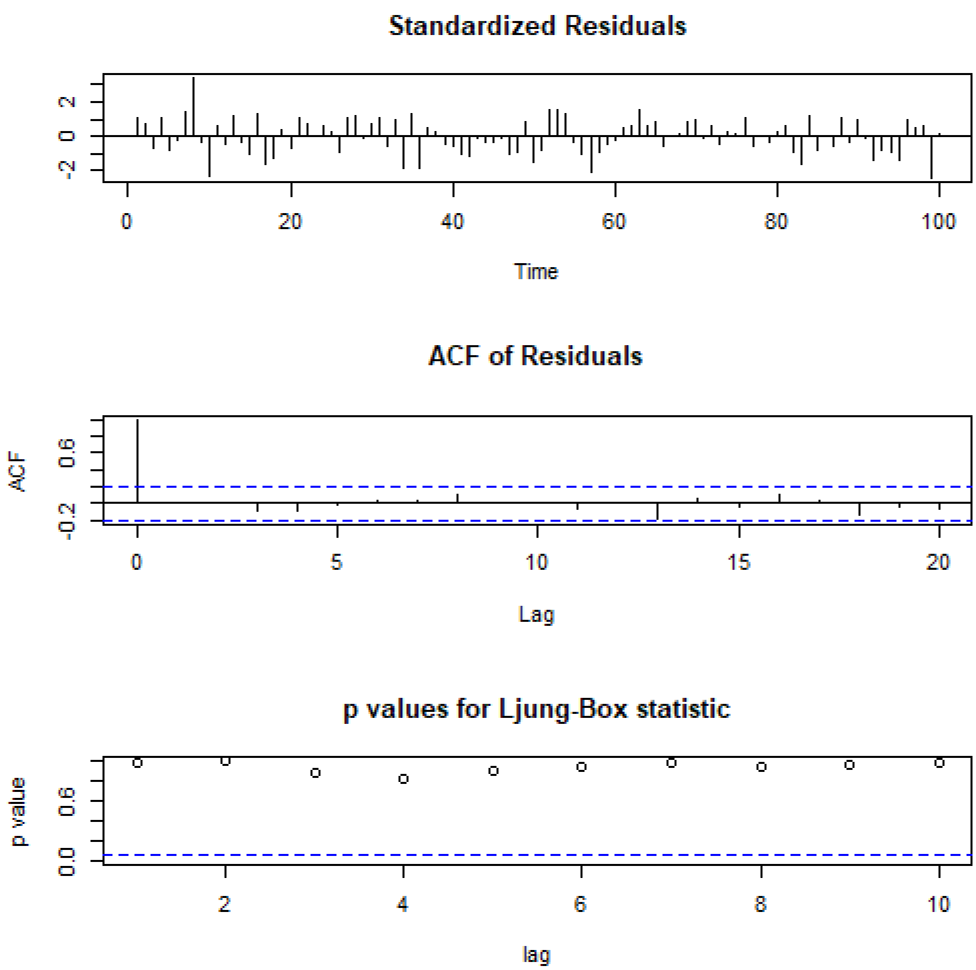
- We test the residuals are correlated or not :

$$\boldsymbol{H_0} : \boldsymbol{\rho_1} = \boldsymbol{\rho_2} = \cdots = \boldsymbol{\rho_k} = \mathbf{0}$$

$$\boldsymbol{H_1} : \boldsymbol{at\ least\ two} \neq \mathbf{0}$$

From the **grarph and p.value for Ljung-Box statistic we** notice all **the p.value's = 0.2934 > α = 0.05** , so we Accept H_0 , that's mean all the residuals are not correlated .

For Model MA(1) :



```
> tsdiag(fit2)


Box.test(res,lag=12,type=c("Ljung-Box"))
Box-Ljung test
data: res
X-squared = 3.517, df = 12, p-value = 0.9907
```

- We test the residuals are correlated or not :

$$\boldsymbol{H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0}$$
$$\boldsymbol{H_1 : at\ least\ two\ \neq 0}$$

From the **grarph p.value for Ljung-Box statistic** we notice all **the p.value's = 0.9907 > α = 0.05** , so we Accept H_0 , that's mean all the residuals are not correlated .

- Now we test are the Residuals random or not :

For model AR(1) :

$$H_0 \ : Residuals\ are\ random$$

$$H_1 \ : Residuals\ are\ not\ random$$

```
> res1=fit1$res

> runs.test(res1)


Runs Test
data: res1
statistic = 0.30729, runs = 51, n1 = 49, n2 = 48, n = 97, p-value = 0.7586
alternative hypothesis: nonrandomness
```

Since the $\boldsymbol{p.value = 0.7586 > \alpha = 0.05}$, so we Accept H_0 , that's mean the residuals are random .

For model MA(1) :

$$H_0 \ : Residuals\ are\ random$$

$$H_1 \ : Residuals\ are\ not\ random$$

```
> res2=fit2$res
> runs.test(res2)


Runs Test
data: res2
statistic = 0.40204, runs = 53, n1 = 50, n2 = 50, n = 100, p-value = 0.6877
alternative hypothesis: nonrandomness
```

Since the $\boldsymbol{p.value = 0.6877 > \alpha = 0.05}$, so we Accept H_0 , that's mean the residuals are random .

- Now we test the mean of the ε_t equal zero or not :

For model AR(1) :

$$H_0 : E(\varepsilon_t) = 0$$

$$H_1 : E(\varepsilon_t) \neq 0$$

```
> t.test(res1)


One Sample t-test
data: res1
t = -0.024777, df = 99, p-value = 0.9803
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-0.8152978 0.7951879
sample estimates:
mean of x
-0.01005497
```

Since the $\boldsymbol{p.value = 0.9803 > \alpha = 0.05}$, so we Accept H_0 , that's mean the expection of ε_t equal zero .

For model MA(1) :

$$H_0 : E(\varepsilon_t) = 0$$

$$H_1 : E(\varepsilon_t) \neq 0$$

```
> t.test(res2)



One Sample t-test
data: res2
t = -0.017248, df = 99, p-value = 0.9863
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-0.8116520 0.7976633
sample estimates:
mean of x
-0.006994388
```

Since the $\boldsymbol{p.value = 0.9863 > \alpha = 0.05}$, so we Accept H_0 , that's mean the expection of ε_t equal zero .

- Now we do inveritiblty analysis :

For Model MA(1) :

$$\theta_1 = 0.2015$$

$$|\theta_1| < 1 \checkmark$$

- Now we do stationary analysis :

For Model AR(1) :

$$\emptyset_1 = 0.1970$$

$$|\emptyset_1| < 1 \checkmark$$

- Now we test the Normality of the Residuals :

For model AR(1) :

```
> shapiro.test(res1)

Shapiro-Wilk normality test
data: res1
W = 0.98315, p-value = 0.2318
```

H_0 : *the Residuals follow Normal Distribution*

H_1 : *the Residuals Don't follow Normal Distribution*

Since the $\boldsymbol{p.value = 0.2318 > \alpha = 0.05}$, so we Accept H_0 , that's mean the Residual are follow Normal Distribution .

For model MA(1) :

H_0 : *the Residuals follow Normal Distribution*

H_1 : *the Residuals Don't follow Normal Distribution*

```
> shapiro.test(res2)

Shapiro-Wilk normality test
data: res2
W = 0.98297, p-value = 0.2247
```

Since the $\boldsymbol{p.value} = \mathbf{0.2247} > \boldsymbol{\alpha} = \mathbf{0.05}$, so we Accept H_0 , that's mean the Residual follow Normal Distribution , so we eliminate this model

AIC or BIC for Models:

Now we have Models AR(1) and MA(1) we calculate AIC for models and compare between them:

```
> fit1$aic
[1] 568.9728


> fit2$aic

[1] 568.8289
```

We choose the lowest **AIC** between models , so MA(1) more appropriate for our Data

Mathematical form of MA(1) :

$$y_t = \varepsilon_t - \theta_1 \varepsilon_t$$

$$y_t = \varepsilon_t - 0.2015\varepsilon_{t-1}$$

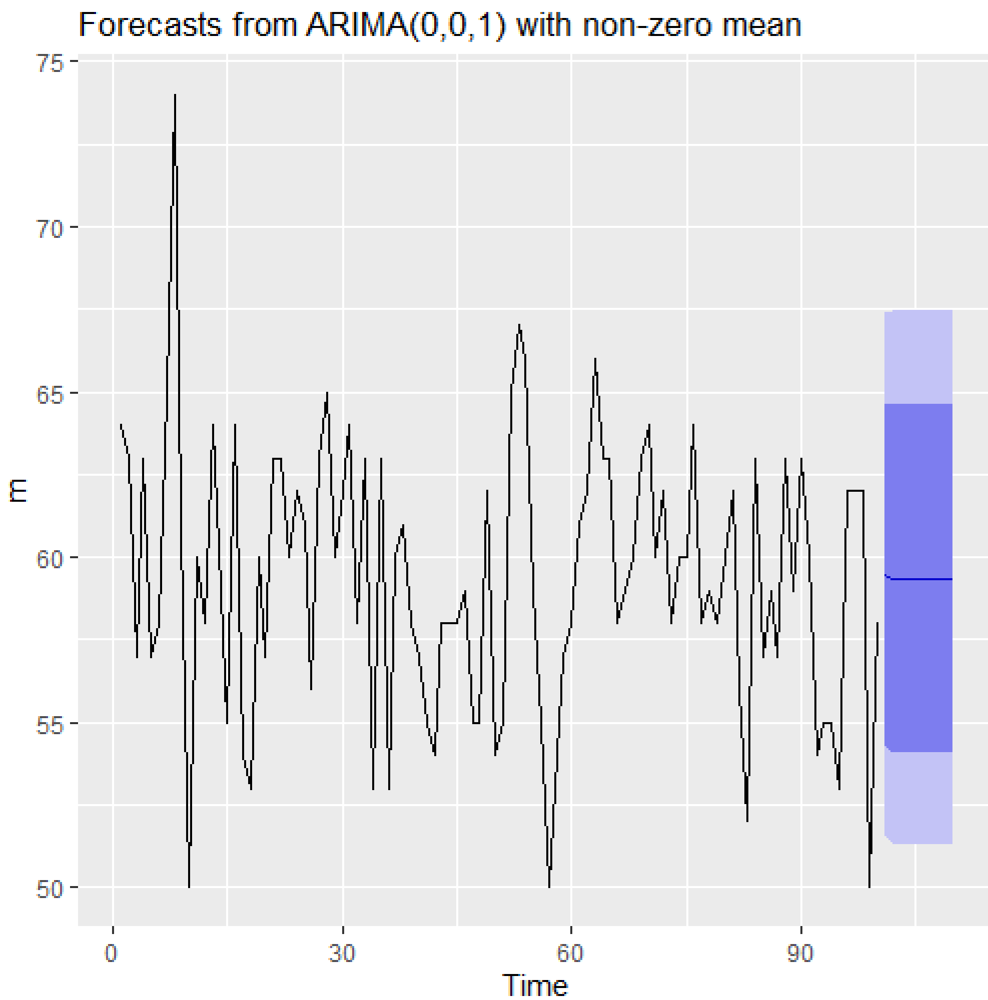
$$y_t = (1 - \theta_1 B)\varepsilon_t$$

```
> auto.arima(m)
Series: m
ARIMA(0,0,1) with non-zero mean
Coefficients:       ma1     mean
                 0.2015  59.3589
           s.e.  0.0958   0.4840
sigma^2 estimated as 16.61:  log likelihood=-281.41
AIC=568.83   AICc=569.08   BIC=576.64
```

Forecasting :

We ready now to forecast with our model we forecast the next 10 value MA(1) :

```
> f=forecast(fit2, h=10)
>  autoplot(f)
> f
  Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
101      59.48364 54.31262 64.65466 51.57525 67.39204
102      59.35890 54.08392 64.63388 51.29152 67.42628
103      59.35890 54.08392 64.63388 51.29152 67.42628
104      59.35890 54.08392 64.63388 51.29152 67.42628
105      59.35890 54.08392 64.63388 51.29152 67.42628
106      59.35890 54.08392 64.63388 51.29152 67.42628
107      59.35890 54.08392 64.63388 51.29152 67.42628
108      59.35890 54.08392 64.63388 51.29152 67.42628
109      59.35890 54.08392 64.63388 51.29152 67.42628
110      59.35890 54.08392 64.63388 51.29152 67.42628
```


Forecasts from ARIMA(0,0,1) with non-zero mean
75
70
65
60
55
50
m
0
30
60
90
Time

Attachments

The Packages and libraries are used :

```
install.packages("zoo")
install.packages("devtools")
devtools::install_github("FinYang/tsdl")
install.packages("nortest")
install.packages("fBasics")
install.packages("forecast")
install.packages("tseries")
install.packages("randtests")
install.packages("astsa")
install.packages("lmtest")
library("zoo")
library(tsdl)
library(nortest)
library(fBasics)
library(forecast)
library(tseries)
library(randtests)
library(astsa)
library(lmtest)
```

Codes Are used :

Code	Meaning
>attributes(m)	To describe the data for what
> plot.ts(m)	To plot your series
> adf.test(m)	To test the stationarity of series
>fit1 = arima(m,order=c(0,0,0),include.mean=T)	To identify your model
>coeftest(fit1)	To test the coefficients of model
>Box.test(res,lag=12,type=c("Ljung-Box"))	To test the residuals are correlated or not
> tsdiag(fit)	To show the graph and test residuals are correlated or not
> runs.test(res1)	To test the residuals are random or not
> t.test(res1)	To test the expection of residuals are zero or not
>Shapiro.test(res1)	To resdiuals are follow normal distribution
>aic$fit	To know what model have less AIC
> f=forecast(fit2, h=10) > autoplot(f)	To predict next values To plot the forecast
> m1=diff(m,diffrence=1)	To take difference fot the series

INDEX