TIME SERIES MODEL-SRIMA
Seasonal Box-Jenkins Models(SARIMA)

Seasonal ARIMA Models

- ARIMA (Auto Regressive Integrated Moving Average) models are regression models that use lagged values of the dependent variable and/or random disturbance term as explanatory variables

- **Seasonal ARIMA (Often abbreviated as SARIMA) Model is formed by including seasonal terms in the ARIMA model**

- **In addition to p,d,q seasonal ARIMA includes P,D,Q**

- Several real world time series have a seasonal component. Some examples are: Sales of woolen clothes, demand for fertilizers, electricity consumption, etc.

Seasonal ARIMA Models

The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model. Shorthand notations for the model is

p = AR order,
d = Order of differencing,
q = MA order,

P = seasonal AR order,
D = Order of seasonal differencing,
Q = seasonal MA order, and

A five-step iterative procedure

1) Stationarity Checking and Differencing
2) Model Identification
3) Parameter Estimation
4) Diagnostic Checking
5) Forecasting

Sales Data Snapshot

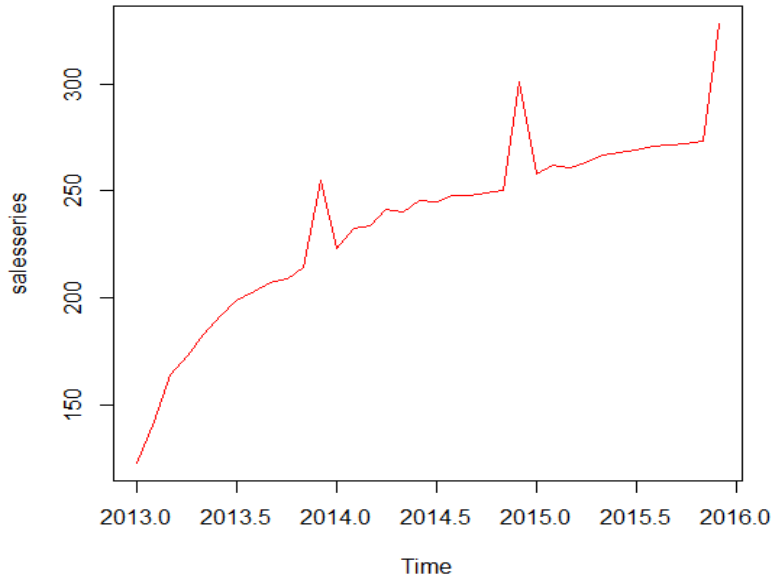
Year	Month	Sales
2013	Jan	123
2013	Feb	142
2013	Mar	164
2013	Apr	173
2013	May	183
2013	Jun	192
2013	Jul	199
2013	Aug	203
2013	Sep	207
2013	Oct	209
2013	Nov	214
2013	Dec	255
2014	Jan	223
2014	Feb	232.5
2014	Mar	233.8
2014	Apr	241.1
2014	May	240
2014	Jun	245.6

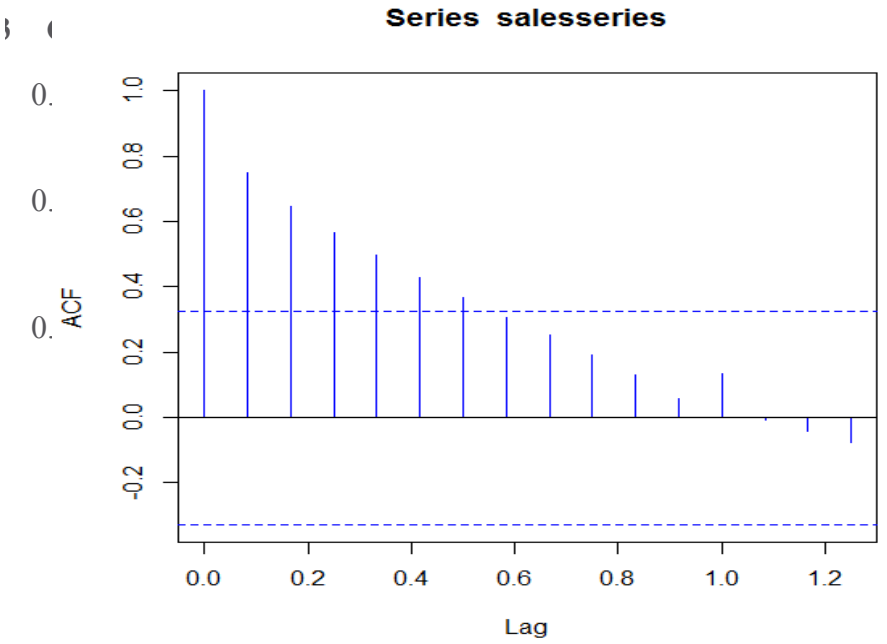
Monthly sales figures for 3 years

Step One: Stationarity Checking

Plot Time Series and ACF

```
salesdata<-read.csv(file.choose(),header=T)
salesseries<-ts(salesdata$Sales,start=c(2013,1),end=c(2015,12),frequency=12)
plot(salesseries,col="red")
acf(salesseries,col="blue")
```





DATA SCIENCE INSTITUTE

How Many Times Should Time Series Be Differenced to Make Stationary?

```
library(forecast)
ndiffs(salesseries)
```

```
> ndiffs(salesseries)
[1] 1
```

```
library(urca)
summary(ur.df(salesseries,lag=0))
```

```
Value of test-statistic is: 1.6212
Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

```
salesdiff<-diff(salesseries,differences=1)
summary(ur.df(salesdiff,lag=0))
```

```
Value of test-statistic is: -6.8914
Critical values for test statistics:
     1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

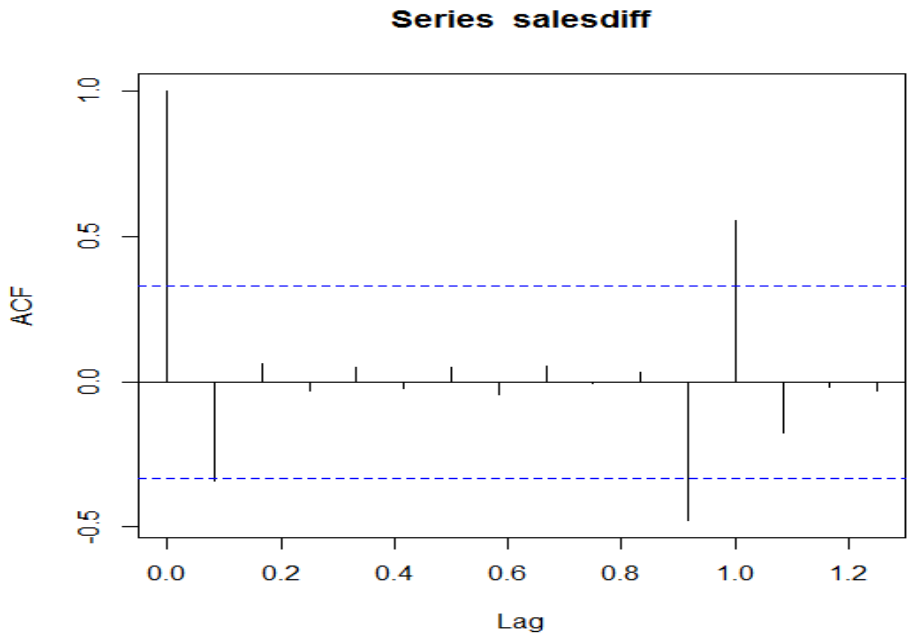
Reject H0.
First difference time series is stationary.

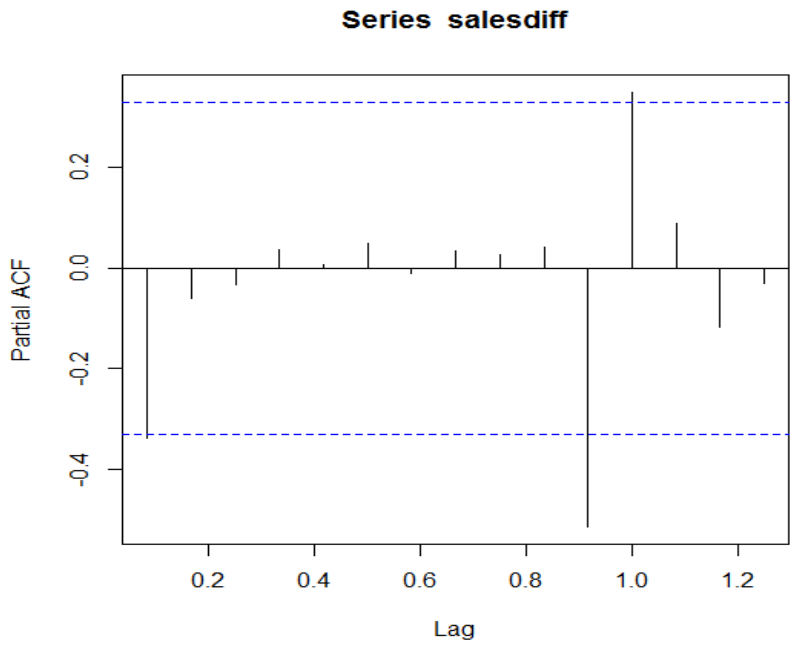
DATA SCIENCE INSTITUTE

Step Two: Model Identification

SARIMA Model in R…
Model Identification

acf(salesdiff)
pacf(salesdiff)





The acf and pacf plots indicate p=1 and q=1. However, it is recommended to set max.p and
Max.q to higher values. In this case we will set both to 2.
Max.P and max.Q can be set to 1 with D=1

DATA SCIENCE INSTITUTE

Step Three: Parameter Estimation

SARIMA Model in R…
Parameter Estimation

```
library(forecast)
salesmodel<-
    auto.arima(salesseries,d=1,D=1,max.p=2,max.q=2,max.P=1,max.Q=1,
trace=TRUE,ic="aic")
```

```
 ARIMA(2,1,2)(1,1,1)[12]              : 157.0397
ARIMA(0,1,0)(0,1,0)[12]               : 156.1096
ARIMA(1,1,0)(1,1,0)[12]               : 159.1271
ARIMA(0,1,1)(0,1,1)[12]               : 159.4057
ARIMA(0,1,0)(1,1,0)[12]               : 157.6536
ARIMA(0,1,0)(0,1,1)[12]               : 157.6536
ARIMA(0,1,0)(1,1,1)[12]               : 159.6536
ARIMA(1,1,0)(0,1,0)[12]               : 157.1806
ARIMA(0,1,1)(0,1,0)[12]               : 157.6069
ARIMA(1,1,1)(0,1,0)[12]               : 154.6016
ARIMA(1,1,1)(1,1,0)[12]               : 161.3531
ARIMA(1,1,1)(0,1,1)[12]               : 156.6
ARIMA(1,1,1)(1,1,1)[12]               : 163.3298
ARIMA(2,1,1)(0,1,0)[12]               : 152.5563
ARIMA(2,1,1)(1,1,0)[12]               : 153.1015
ARIMA(2,1,1)(0,1,1)[12]               : Inf
ARIMA(2,1,1)(1,1,1)[12]               : 155.0459
ARIMA(2,1,0)(0,1,0)[12]               : 151.6156
ARIMA(2,1,0)(1,1,0)[12]               : 151.8961
ARIMA(2,1,0)(0,1,1)[12]               : Inf
ARIMA(2,1,0)(1,1,1)[12]               : 153.864
```

Best model: ARIMA(2,1,0)(0,1,0)[12]

d=1 indicates differencing order
D=1 indicates seasonal differencing order
p=Number of AR terms
q= Number of MA terms
P=Number of seasonal AR terms
Q=Number of seasonal MA terms

SARIMA Model in R…
Parameter Estimation

```
#Type model object and display coefficients
salesmodel

Series: salesseries
ARIMA(2,1,0)(0,1,0)[12]

Coefficients:
         ar1     ar2
      0.1583  0.6353
s.e.  0.1545  0.1856

sigma^2 = 34.14:  log likelihood = -72.81
AIC=151.62   AICc=152.88   BIC=155.02
```

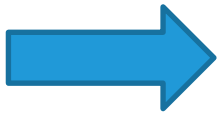
Step Four: Diagnostic Checking

SARIMA Model in R…
Diagnostic Checking

```
resi<-residuals(salesmodel)
Box.test(resi)
plot(resi,col="red")
```

Box-Pierce test

data: resi
X-squared = 0.78553, df = 1, p-value = 0.3755



Do not reject Ho. Errors follow white noise process.

Step Five: Forecasting

SARIMA Model in R… Forecasting

predict(salesmodel,n.ahead=3)

- predict(salesmodel,n.ahead=3)

```
$pred
          Jan       Feb       Mar
2016 285.6334 292.1748 292.0954

$se
           Jan         Feb         Mar
2016  5.842612  8.940863   13.887043
```

Intuitive Explanation of ARIMA/SARIMA Model

p=2 ,d=1 and q=2

December 2022 sales forecast depends on November 2022 sales, October 2022 sales ,error in November 2022 forecast and error in October 2022 Forecast.

In addition if P=1 , D=1 and Q=1 then

December 2022 sales forecast also depends on December 2021 sales and error in December 2021 forecast

THANK YOU!!

DATA SCIENCE
INSTITUTE