

# Discoveries Time-Series Worksheet

*Dr. Nicole M. Pamphilis*

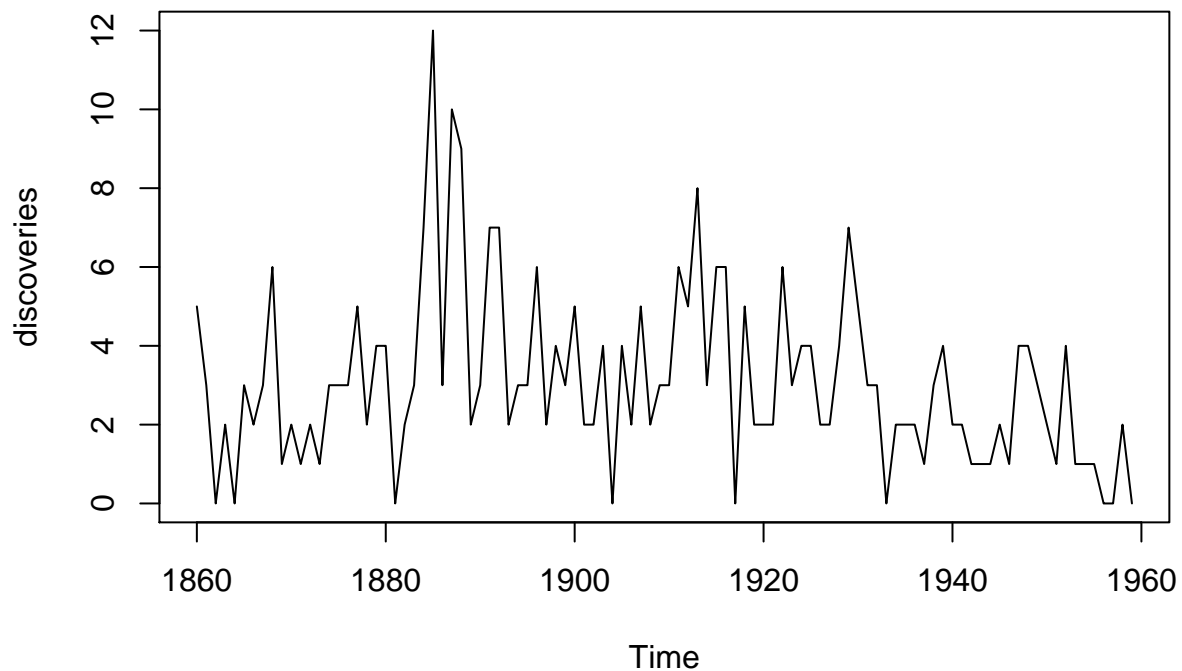
Load data-already time set

Number of discoveries per year from 1860 until 1959

```
data<-discoveries
```

Plot the data for first visual check

```
plot.ts(discoveries)
```



Check if data is stationary

Load library with time-series tests

```
library(tseries)
```

```
## Warning: package 'tseries' was built under R version 3.4.4
```

## Run checks to ensure stationary data

```
adf.test(discoveries)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: discoveries  
## Dickey-Fuller = -3.6235, Lag order = 4, p-value = 0.03462  
## alternative hypothesis: stationary
```

```
pp.test(discoveries)
```

```
## Warning in pp.test(discoveries): p-value smaller than printed p-value  
##  
## Phillips-Perron Unit Root Test  
##  
## data: discoveries  
## Dickey-Fuller Z(alpha) = -81.682, Truncation lag parameter = 3,  
## p-value = 0.01  
## alternative hypothesis: stationary
```

```
kpss.test(discoveries)
```

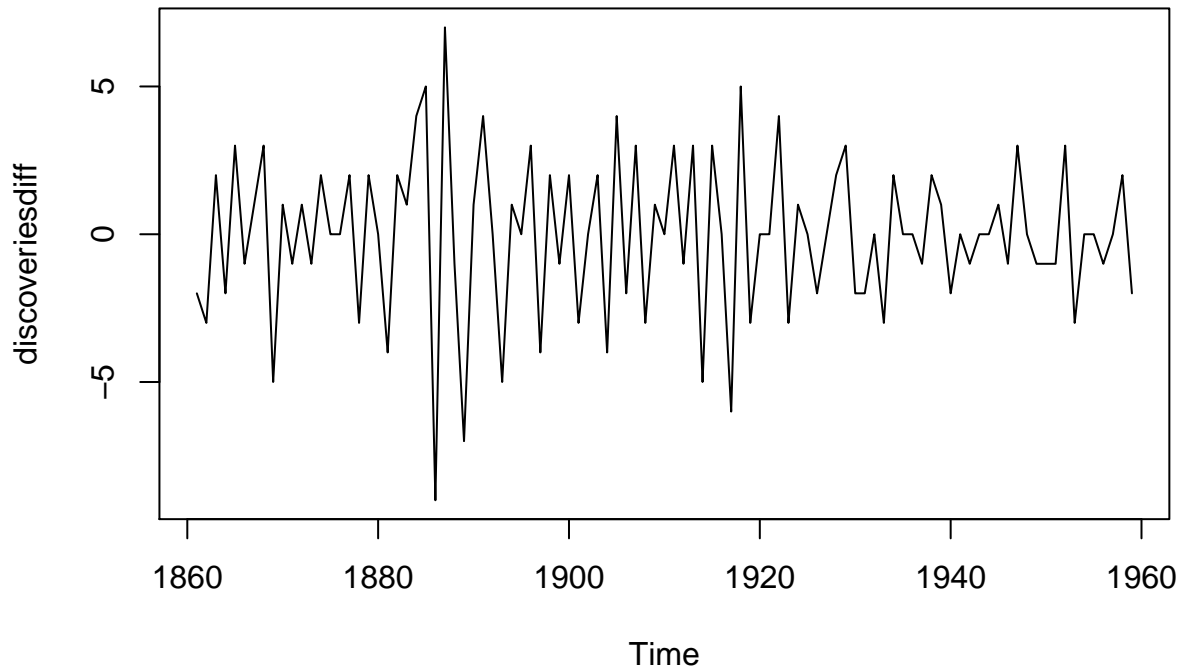
```
##  
## KPSS Test for Level Stationarity  
##  
## data: discoveries  
## KPSS Level = 0.54761, Truncation lag parameter = 2, p-value =  
## 0.03094
```

## Differencing the data

```
discoveriesdiff<-diff(discoveries, differences = 1)
```

## Plot differenced data

```
plot.ts(discoveriesdiff)
```



```
##Check for stationary
```

```
adf.test(discoveriesdiff)
```

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

```
##
```

```
## Augmented Dickey-Fuller Test
```

```
##
```

```
## data: discoveriesdiff
```

```
## Dickey-Fuller = -5.9543, Lag order = 4, p-value = 0.01
```

```
## alternative hypothesis: stationary
```

```
pp.test(discoveriesdiff)
```

```
## Warning in pp.test(discoveriesdiff): p-value smaller than printed p-value
```

```
##
```

```
## Phillips-Perron Unit Root Test
```

```
##
```

```
## data: discoveriesdiff
```

```
## Dickey-Fuller Z(alpha) = -127.59, Truncation lag parameter = 3,
```

```
## p-value = 0.01
```

```
## alternative hypothesis: stationary
```

```
kpss.test(discoveriesdiff)
```

```
## Warning in kpss.test(discoveriesdiff): p-value greater than printed p-value
```

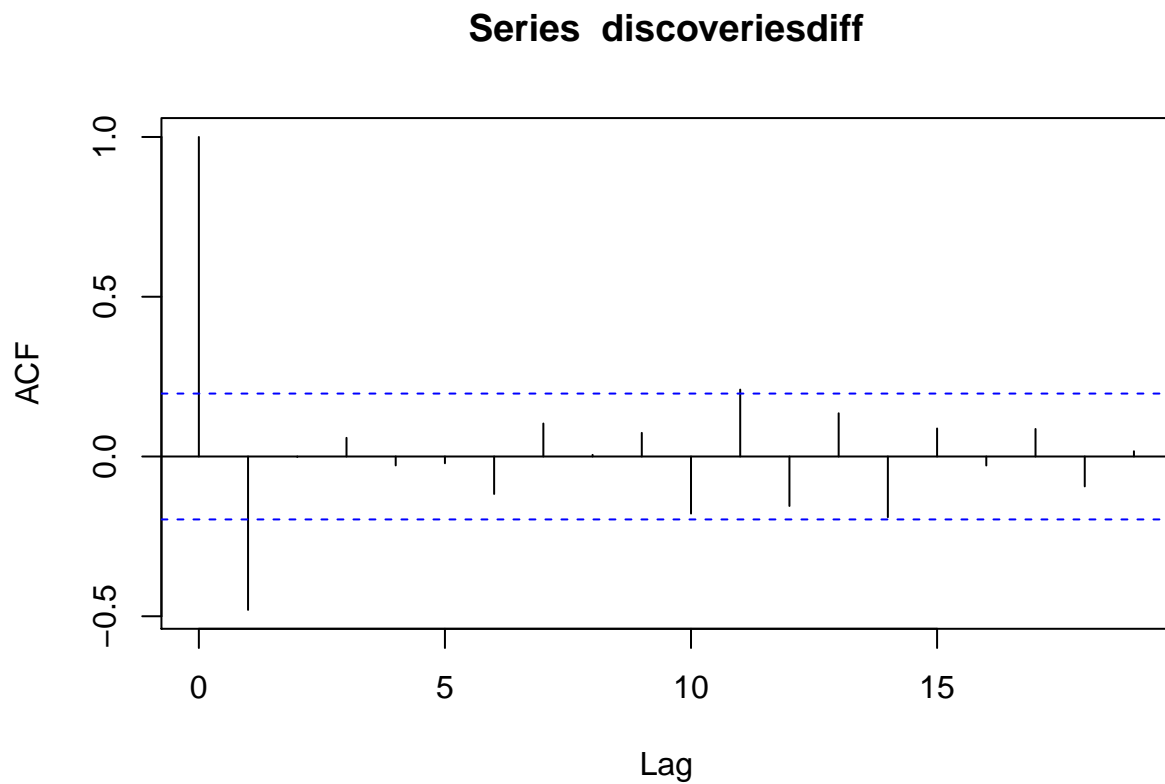
```
##
```

```
## KPSS Test for Level Stationarity
##
## data: discoveriesdiff
## KPSS Level = 0.024036, Truncation lag parameter = 2, p-value = 0.1
```

## Examine ACF/PACF

ACF

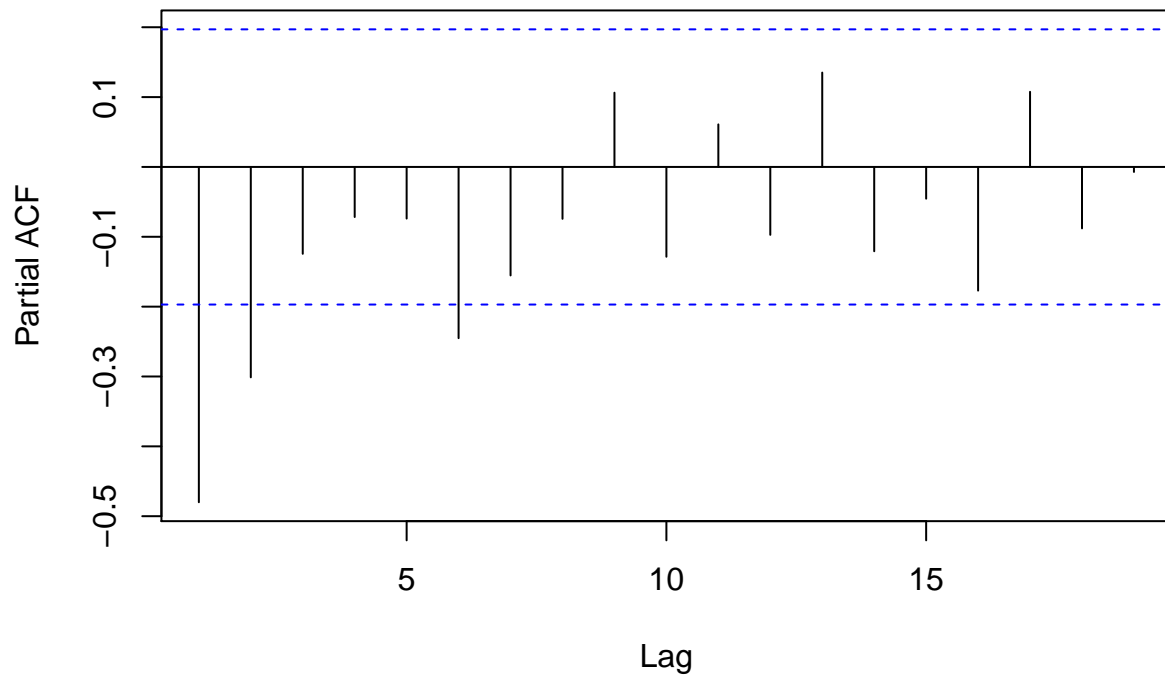
```
acf(discoveriesdiff)
```



PACF

```
pacf(discoveriesdiff)
```

## Series discoveriesdiff



```
##Run Possible Models
```

```
##AR1
```

```
discoveriesdiff100<-arima(discoveriesdiff, order=c(1,0,0))  
discoveriesdiff100
```

```
##
```

```
## Call:
```

```
## arima(x = discoveriesdiff, order = c(1, 0, 0))
```

```
##
```

```
## Coefficients:
```

```
##      ar1  intercept
```

```
##    -0.4804  -0.0376
```

```
## s.e.  0.0878   0.1605
```

```
##
```

```
## sigma^2 estimated as 5.554:  log likelihood = -225.47,  aic = 456.94
```

```
##ARIMA 110
```

```
discoveries110<-arima(discoveries, order=c(1,1,0))  
discoveries110
```

```
##
```

```
## Call:
```

```
## arima(x = discoveries, order = c(1, 1, 0))
```

```
##
```

```
## Coefficients:
```

```
##      ar1
```

```
##    -0.4803
```

```
## s.e.  0.0879
```

```

##
## sigma^2 estimated as 5.557: log likelihood = -225.5, aic = 455
##MA1
discoveriesdiff001<-arima(discoveriesdiff, order=c(0,0,1))
discoveriesdiff001

##
## Call:
## arima(x = discoveriesdiff, order = c(0, 0, 1))
##
## Coefficients:
##      ma1  intercept
##    -0.8117  -0.0164
## s.e.  0.0806   0.0426
##
## sigma^2 estimated as 4.592: log likelihood = -216.47, aic = 438.94
##MA2
discoveriesdiff002<-arima(discoveriesdiff, order=c(0,0,2))
discoveriesdiff002

##
## Call:
## arima(x = discoveriesdiff, order = c(0, 0, 2))
##
## Coefficients:
##      ma1      ma2  intercept
##    -0.7724 -0.0666  -0.0156
## s.e.  0.0971  0.1048   0.0368
##
## sigma^2 estimated as 4.569: log likelihood = -216.27, aic = 440.53

```

## Collect Residuals for each Model to test

```

discoveriesdiff100res<-residuals(discoveriesdiff100)
discoveriesdiff001res<-residuals(discoveriesdiff001)
discoveriesdiff002res<-residuals(discoveriesdiff002)

```

## Generate Q-Stats for each model

```

##AR1
Box.test(discoveriesdiff100res, type="Ljung-Box")

##
## Box-Ljung test
##
## data: discoveriesdiff100res
## X-squared = 2.2437, df = 1, p-value = 0.1342
##MA1
Box.test(discoveriesdiff001res, type="Ljung-Box")

##

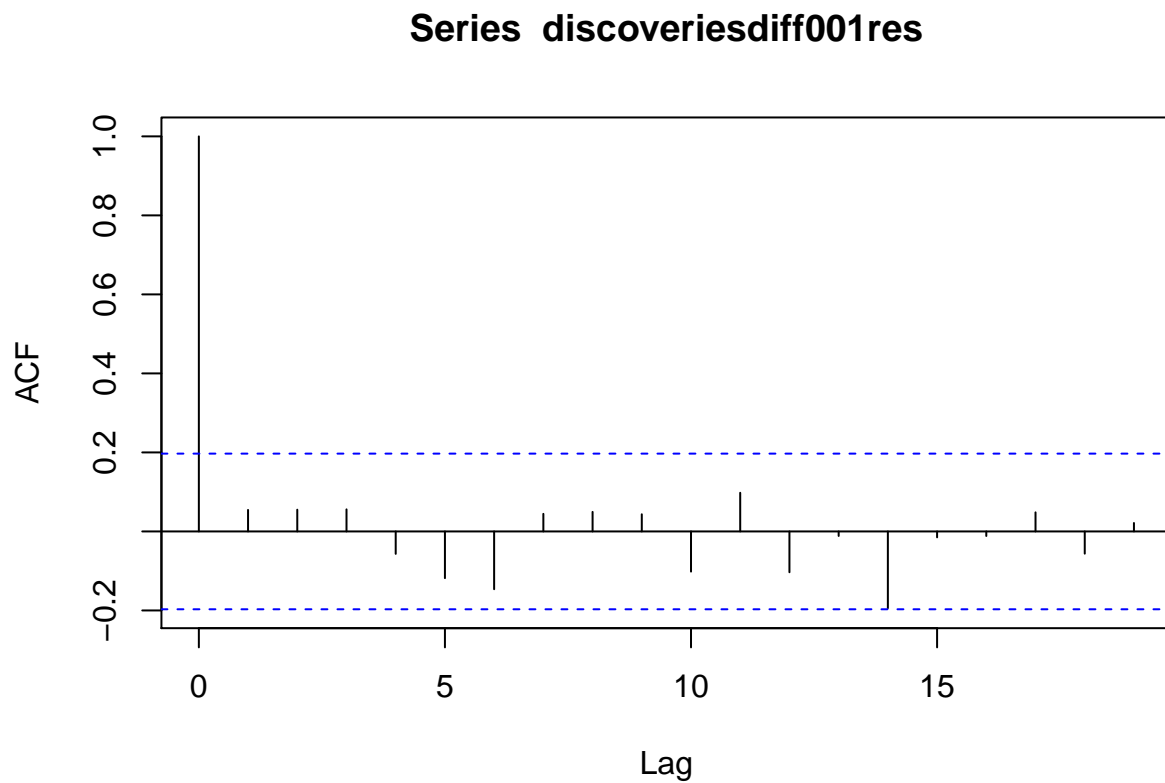
```

```
## Box-Ljung test
##
## data: discoveriesdiff001res
## X-squared = 0.30273, df = 1, p-value = 0.5822
##MA2
Box.test(discoveriesdiff002res, type="Ljung-Box")

##
## Box-Ljung test
##
## data: discoveriesdiff002res
## X-squared = 0.0070418, df = 1, p-value = 0.9331
```

## Visual Check for White-Noise

```
#ACF of Residuals
acf(discoveriesdiff001res)
```



```
#PACF of Residuals
pacf(discoveriesdiff001res)
```

### Series discoveriesdiff001res

