MONTHLY PRODUCTION OF CLAY BRICKS
Grp4 - Project

Claudius Taylor, Tom Wilson, Junpu Zhao

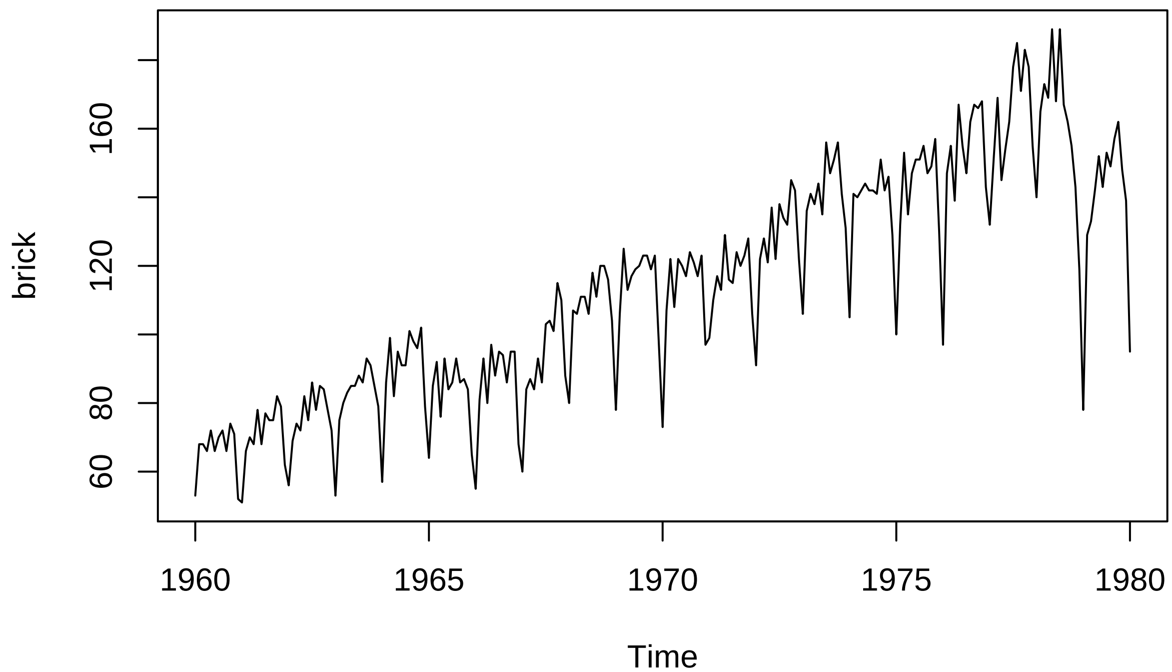
12/12/2018

INTRODUCTION:

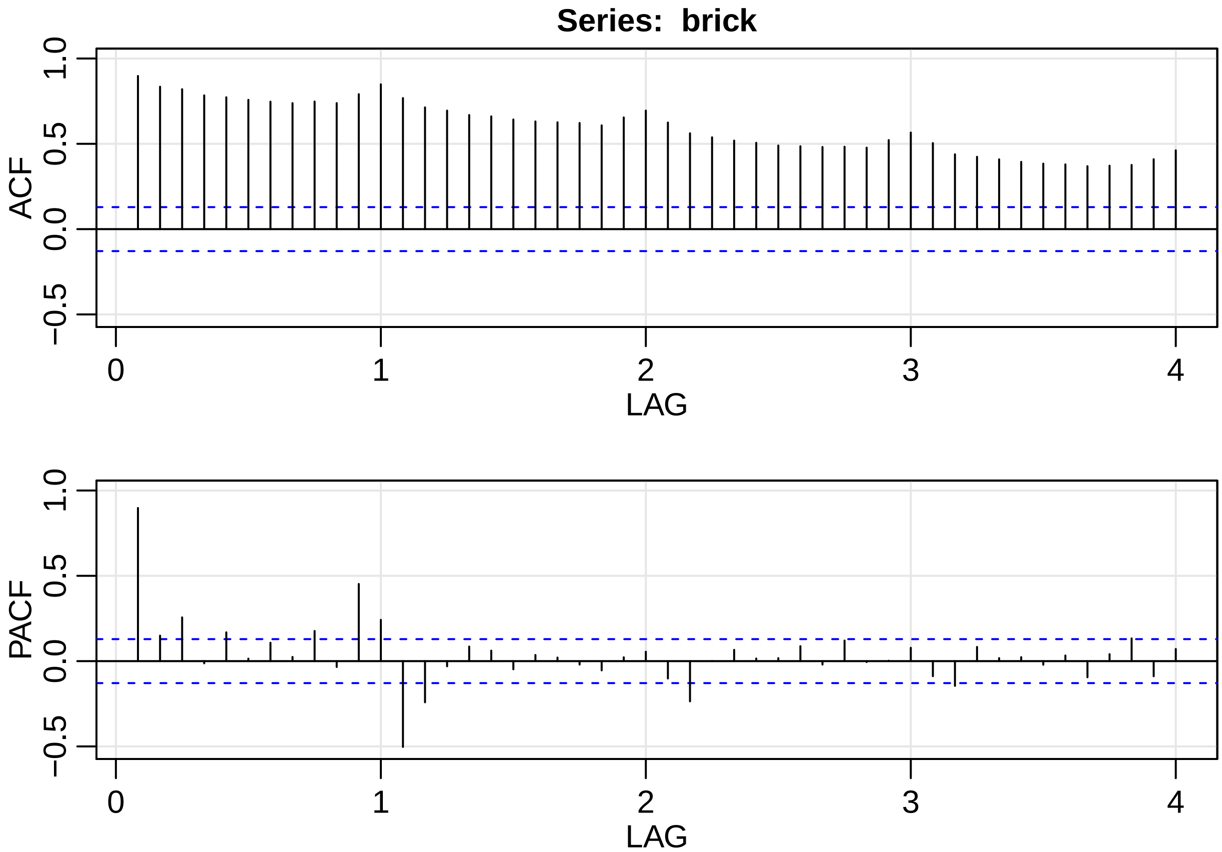
The aims of this study are to identify and forecast a model best fitting brick production data in the United States. The method of maximum likelihood was used to estimate the parameters and to forecast the number of production in the future. The data is a twenty year period from 1960 to 1980 and was obtained from the Time Series Data library at datamarket.com website.

This project is of utmost importance and relevance because bricks are used for building and pavement all throughout the world. Being made from clay and shale, brick is most abundant and natural material on earth. In the USA, bricks were once used as a pavement material, and now it is more widely used as a decorative surface rather than a roadway material. A healthy living environment especially requires the use of the right building material. In general building materials are strongly influencing the indoor climate and quality of living.

```r
# loading the data
data = ("./data/claybrick.csv")
columnNames = c("month", "production")
brick = read.csv(file = data,
                 comment.char = "",
                 header = TRUE,
                 col.names = columnNames)
brick = ts(brick$production, start = 1960, end = 1980, frequency = 12)
plot(brick)
```




```r
acf2(brick)
```

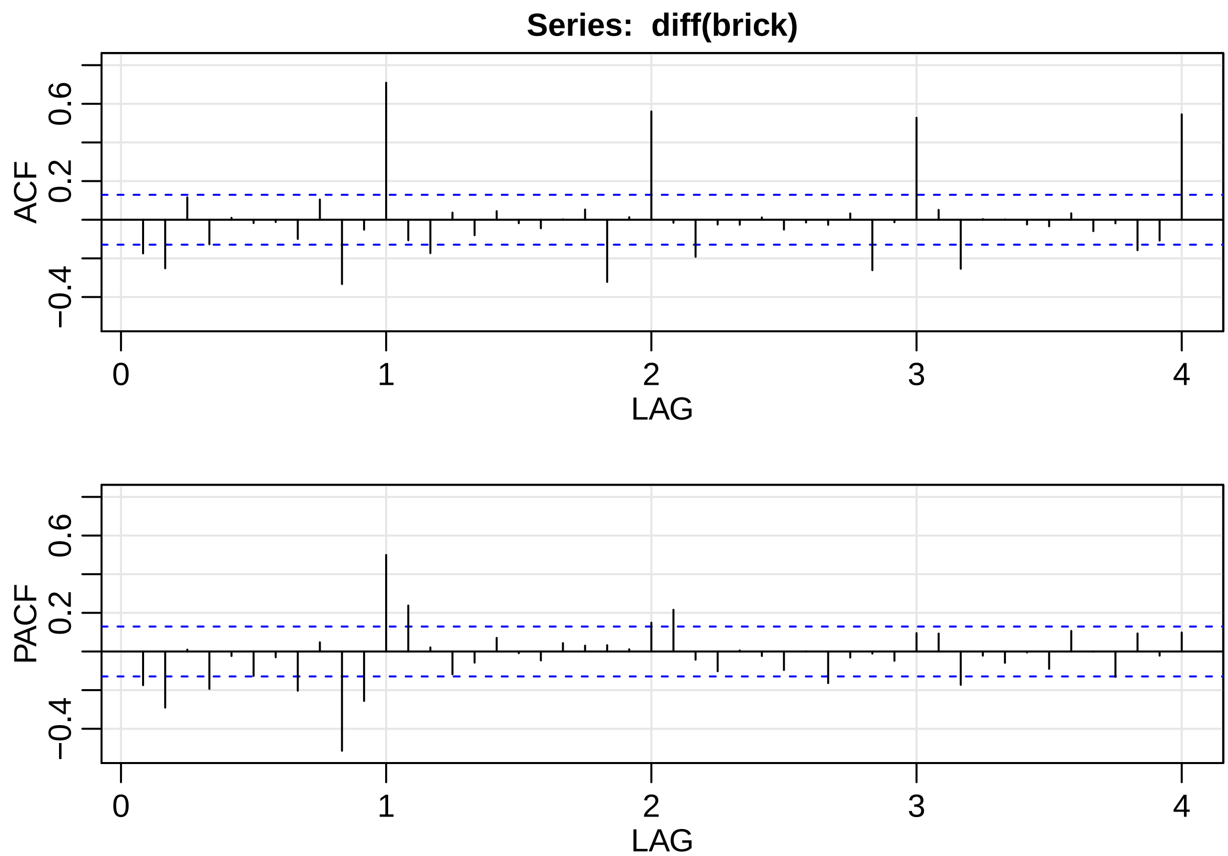



```
##        ACF  PACF
##  [1,] 0.90  0.90
##  [2,] 0.83  0.15
##  [3,] 0.82  0.26
##  [4,] 0.78 -0.01
##  [5,] 0.77  0.17
##  [6,] 0.76  0.01
##  [7,] 0.75  0.11
##  [8,] 0.74  0.02
##  [9,] 0.75  0.18
## [10,] 0.74 -0.03
## [11,] 0.79  0.45
## [12,] 0.85  0.24
## [13,] 0.77 -0.50
## [14,] 0.71 -0.24
## [15,] 0.69 -0.03
## [16,] 0.67  0.09
## [17,] 0.66  0.06
## [18,] 0.64 -0.05
## [19,] 0.63  0.04
## [20,] 0.63  0.02
## [21,] 0.62 -0.02
## [22,] 0.61 -0.05
## [23,] 0.65  0.02
## [24,] 0.69  0.06
## [25,] 0.62 -0.10
```

```
## [26,] 0.56 -0.24
## [27,] 0.54  0.00
## [28,] 0.52  0.07
## [29,] 0.51  0.02
## [30,] 0.49  0.02
## [31,] 0.49  0.09
## [32,] 0.48 -0.02
## [33,] 0.48  0.12
## [34,] 0.48 -0.01
## [35,] 0.52  0.00
## [36,] 0.57  0.08
## [37,] 0.50 -0.09
## [38,] 0.44 -0.15
## [39,] 0.42  0.08
## [40,] 0.41  0.02
## [41,] 0.39  0.02
## [42,] 0.38 -0.02
## [43,] 0.38  0.03
## [44,] 0.37 -0.09
## [45,] 0.37  0.04
## [46,] 0.38  0.13
## [47,] 0.41 -0.09
## [48,] 0.46  0.07
```

The graph shows a trend and seasonal variations in the number of coal production every year. A distinct trough is shown in 1979. The random fluctuations seem constant over time.

```r
acf2(diff(brick), 48)
```

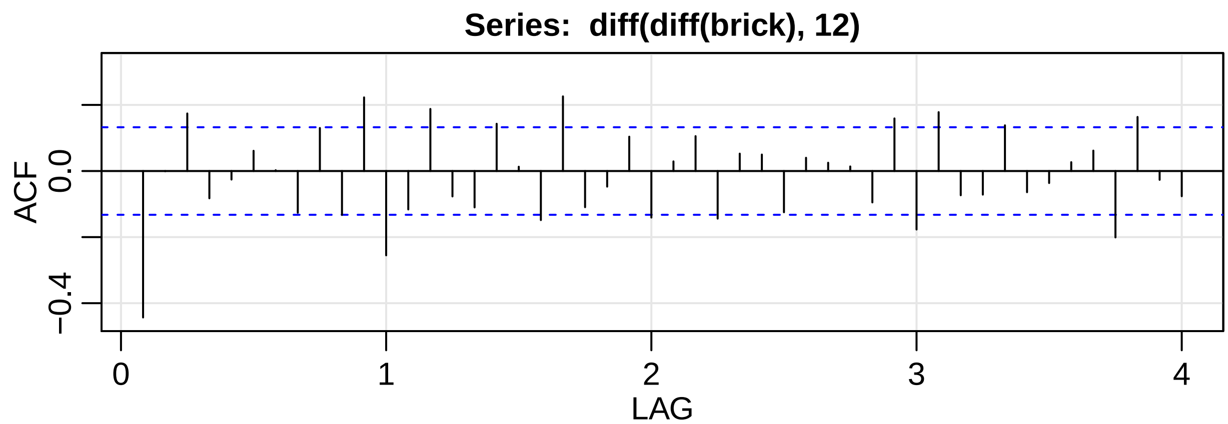


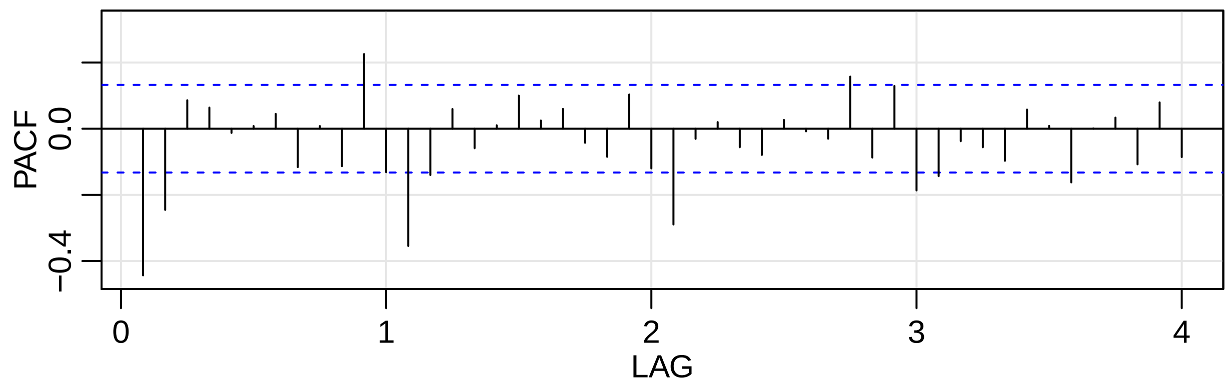

```
##          ACF  PACF
##  [1,] -0.17 -0.17
##  [2,] -0.25 -0.29
##  [3,]  0.12  0.01
##  [4,] -0.13 -0.19
##  [5,]  0.01 -0.02
##  [6,] -0.02 -0.13
##  [7,] -0.01 -0.03
##  [8,] -0.10 -0.20
##  [9,]  0.10  0.05
## [10,] -0.33 -0.51
## [11,] -0.05 -0.26
## [12,]  0.71  0.50
## [13,] -0.11  0.24
## [14,] -0.17  0.02
## [15,]  0.04 -0.12
## [16,] -0.08 -0.06
## [17,]  0.04  0.07
## [18,] -0.02 -0.01
## [19,] -0.04 -0.05
## [20,]  0.00  0.04
## [21,]  0.05  0.03
## [22,] -0.32  0.03
## [23,]  0.01  0.01
## [24,]  0.56  0.15
## [25,] -0.02  0.22
```

```
## [26,] -0.19 -0.04
## [27,] -0.02 -0.10
## [28,] -0.03  0.01
## [29,]  0.01 -0.02
## [30,] -0.05 -0.10
## [31,] -0.01  0.00
## [32,] -0.03 -0.16
## [33,]  0.03 -0.03
## [34,] -0.26 -0.01
## [35,] -0.01 -0.05
## [36,]  0.53  0.09
## [37,]  0.05  0.09
## [38,] -0.25 -0.17
## [39,]  0.00 -0.02
## [40,]  0.00 -0.06
## [41,] -0.02 -0.01
## [42,] -0.03 -0.09
## [43,]  0.03  0.11
## [44,] -0.06  0.00
## [45,] -0.02 -0.13
## [46,] -0.16  0.09
## [47,] -0.11 -0.02
## [48,]  0.55  0.10
```

Even with the first order of differencing, we observe that there is still slow residual decay in the ACF plots at a seasonal lag period of 12. This thus suggest a seasonal difference to be applied.

```r
acf2(diff(diff(brick), 12), 48)
```






```
##         ACF  PACF
##  [1,] -0.44 -0.44
##  [2,]  0.00 -0.25
##  [3,]  0.17  0.09
##  [4,] -0.08  0.06
##  [5,] -0.03 -0.01
##  [6,]  0.06  0.01
##  [7,]  0.00  0.04
##  [8,] -0.13 -0.12
##  [9,]  0.13  0.01
## [10,] -0.13 -0.11
## [11,]  0.22  0.23
## [12,] -0.26 -0.13
## [13,] -0.12 -0.35
## [14,]  0.19 -0.14
## [15,] -0.08  0.06
## [16,] -0.11 -0.06
## [17,]  0.14  0.01
## [18,]  0.01  0.10
## [19,] -0.15  0.02
## [20,]  0.23  0.06
## [21,] -0.11 -0.04
## [22,] -0.05 -0.08
## [23,]  0.10  0.10
## [24,] -0.14 -0.12
## [25,]  0.03 -0.29
```

```
## [26,]  0.11 -0.03
## [27,] -0.14  0.02
## [28,]  0.05 -0.06
## [29,]  0.05 -0.08
## [30,] -0.12  0.03
## [31,]  0.04 -0.01
## [32,]  0.03 -0.03
## [33,]  0.01  0.16
## [34,] -0.10 -0.09
## [35,]  0.16  0.13
## [36,] -0.18 -0.19
## [37,]  0.18 -0.14
## [38,] -0.07 -0.04
## [39,] -0.07 -0.06
## [40,]  0.14 -0.10
## [41,] -0.06  0.06
## [42,] -0.04  0.01
## [43,]  0.03 -0.16
## [44,]  0.06  0.00
## [45,] -0.20  0.03
## [46,]  0.16 -0.11
## [47,] -0.03  0.08
## [48,] -0.08 -0.09
```

From the seasonal lag perspective, we can see that the ACF cuts off at the 2nd seasonal lag, while the PACF appears to tail off. This would suggest a SARMA model of (0,2).

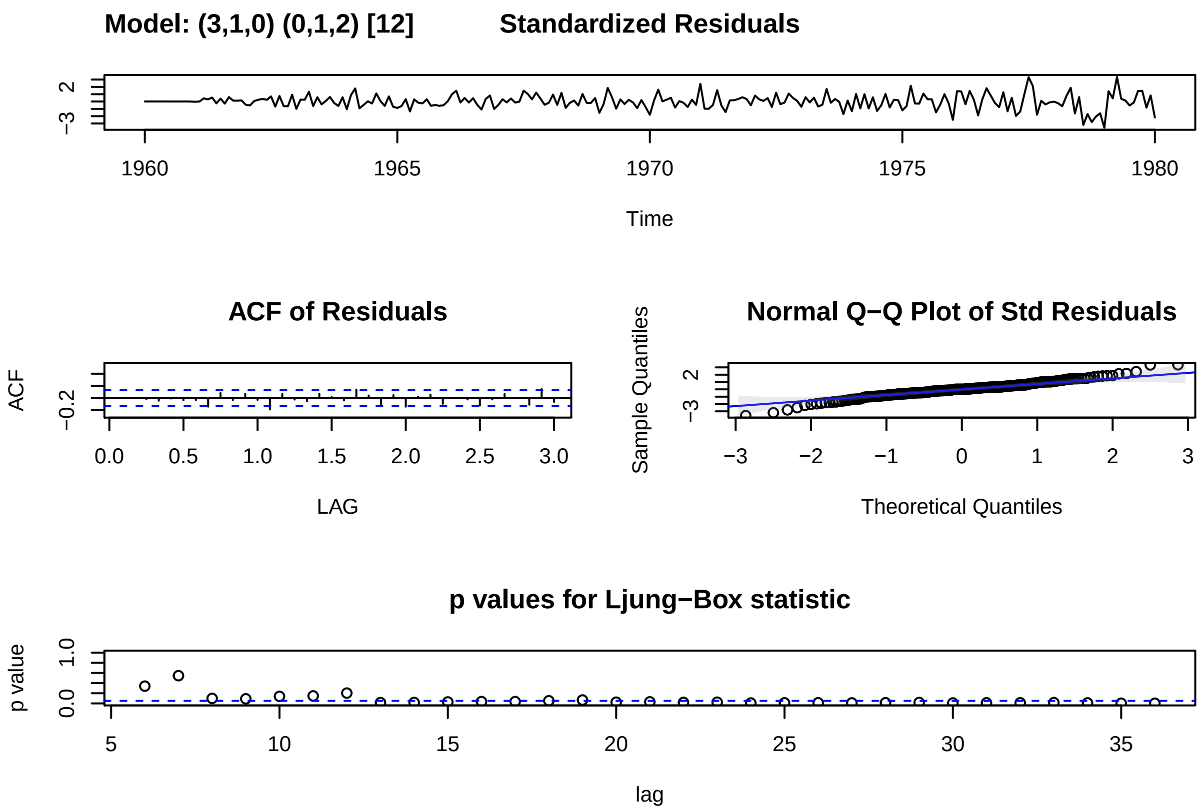
Within the first seasonal cycle, it can be seen that PACF appears to be cutting off at lag = 3, while the ACF tails off.

Thus a proposed model can be ARMA (3,0) x (0,2)__12 for the differenced time series.

```r
sarima(brick, 3,1,0, 0,1,2, 12)
```

```
## initial  value 2.316501
## iter   2 value 2.082344
## iter   3 value 2.018921
## iter   4 value 1.986378
## iter   5 value 1.964898
## iter   6 value 1.960672
## iter   7 value 1.959212
## iter   8 value 1.959089
## iter   9 value 1.958890
## iter  10 value 1.958870
## iter  11 value 1.958867
## iter  12 value 1.958867
## iter  12 value 1.958867
## final  value 1.958867
## converged
## initial  value 1.969522
```

```
## iter   2 value 1.969380
## iter   3 value 1.969078
## iter   4 value 1.968929
## iter   5 value 1.968904
## iter   6 value 1.968903
## iter   6 value 1.968903
## iter   6 value 1.968903
## final  value 1.968903
## converged
```




```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, optim.control = list(trace = trc,
##     REPORT = 1, reltol = tol))
##
## Coefficients:
##           ar1      ar2     ar3     sma1     sma2
##       -0.5935  -0.2042  0.1479  -0.7591  -0.0367
## s.e.   0.0684   0.0762  0.0682   0.0866   0.0869
##
## sigma^2 estimated as 48.63:  log likelihood = -772.43,  aic = 1556.86
##
## $degrees_of_freedom
## [1] 223
```

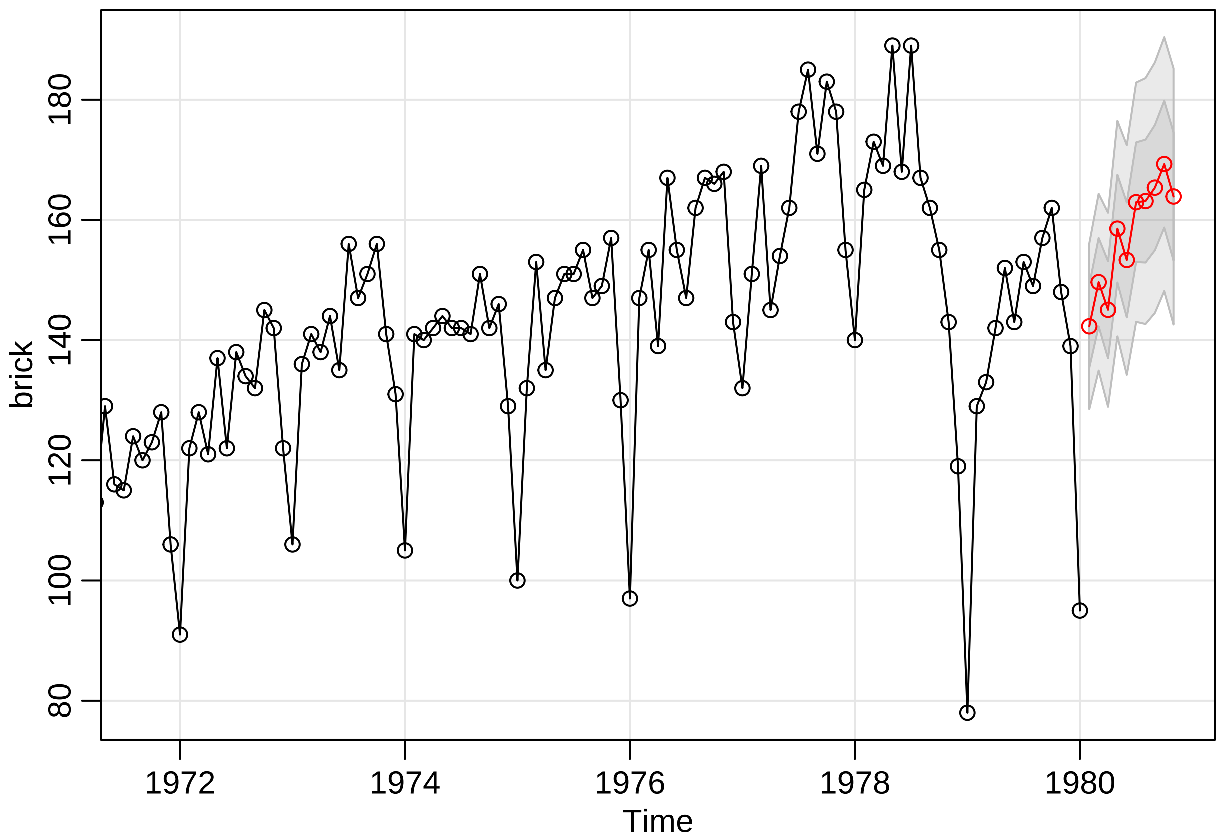
```
##
## $ttable
##      Estimate     SE t.value p.value
## ar1   -0.5935 0.0684 -8.6713  0.0000
## ar2   -0.2042 0.0762 -2.6788  0.0079
## ar3    0.1479 0.0682  2.1668  0.0313
## sma1  -0.7591 0.0866 -8.7688  0.0000
## sma2  -0.0367 0.0869 -0.4228  0.6729
##
## $AIC
## [1] 4.925686
##
## $AICc
## [1] 4.935474
##
## $BIC
## [1] 3.997985
```

Looking at the model diagnostics, we can see that the model does fit fine for earlier lags, although there might still be some outliers in the data with unexplained variance (as shown in the Normal QQ plot, and the standardised residuals).

```r
auto.arima(brick)
```

```
## Series: brick
## ARIMA(1,0,3)(0,1,2)[12] with drift
##
## Coefficients:
##          ar1      ma1     ma2     ma3     sma1     sma2   drift
##       0.8452  -0.4738  0.1683  0.1588  -0.6485  -0.1605  0.3894
## s.e.  0.0566   0.0915  0.0830  0.0910   0.0758   0.0748  0.0546
##
## sigma^2 estimated as 49.01:  log likelihood=-773.31
## AIC=1562.62   AICc=1563.27   BIC=1590.09
```

```r
sarima.for(brick, 10, 1,0,3, 0,1,2, 12)
```




```
## $pred
##           Feb      Mar      Apr      May      Jun      Jul      Aug
## 1980 142.2943 149.6362 145.0447 158.5480 153.3325 162.9443 163.1305
##           Sep      Oct      Nov
## 1980 165.3662 169.2914 163.8926
##
## $se
##            Feb       Mar       Apr       May       Jun       Jul       Aug
## 1980  6.894003  7.354126  8.070724  8.965617  9.553698  9.952536 10.227917
##            Sep       Oct       Nov
## 1980 10.420162 10.555319 10.650802
```

RESULT AND DISCUSSION:

By using the method of maximum likelihood, the present study identified that the ARMA (,) model fits the

CONCLUSION:

Along with the predicted data, there is the prediction bounds (+- 1 standard error represented by the darker gray bands and +-2 standard errors boundaries represented by the lighter gray bands). As the time progresses beyond the first predicted point, the uncertainty increases and thus the prediction boundaries increase in amplitude.

REFERENCES:

https://www.datamarket.com: SOURCE OF DATA

https://www.eia.gov/totalenergy/data/annual

www.claybrick.org

"Trends in Brick Plant Operation," The American Ceramic Society Bulletin. 1992, pp.69-74

"Brick Manufacturing from Past to Present," The American Ceramic Society Bulletin. May, 1990, pp.807-813