The following time series is taken from Time Series Data Library (TSDL)

TSDL was created by Rob Hyndman

Professor of Statistics at Monash University, Australia.

==========

====== Daily total female birth in California, 1959 =======

Data is exported as csv file to the wroking directory

Link: https://datamarket.com/data/list/?q=cat:fwy%20provider:tsdl (https://datamarket.com/data/list/?q=cat:fwy%20provider:tsdl)

In [ ]:
```
library(astsa)

# read data to R variable
birth.data<-read.csv("daily-total-female-births-in-cal.csv")

# pull out number of births column
number_of_births<-birth.data$Daily.total.female.births.in.California..1959

# use date format for dates
birth.data$Date <- as.Date(birth.data$Date, "%m/%d/%Y")
```

In [ ]:
```
plot.ts(number_of_births,main='Daily total female births in california, 195
```

In [ ]:
```
# Test for correlation
Box.test(number_of_births, lag = log(length(number_of_births)))
```

In [ ]:
```
# Plot the differenced data
plot.ts(diff(number_of_births), main='Differenced series', ylab = '')
```

In [ ]:
```
# Test for correlation in the differenced data
Box.test(diff(number_of_births), lag = log(length(diff(number_of_births))))
```

In [ ]:
```
# acf and pacf of the differenced data

acf(diff(number_of_births), main='ACF of differenced data', 50)
pacf(diff(number_of_births), main='PACF of differenced data', 50)
```

In [ ]:
```
# Fit various ARIMA models


model1<-arima(number_of_births, order=c(0,1,1))
SSE1<-sum(model1$residuals^2)
model1.test<-Box.test(model1$residuals, lag = log(length(model1$residuals)))

model2<-arima(number_of_births, order=c(0,1,2))
SSE2<-sum(model2$residuals^2)
model2.test<-Box.test(model2$residuals, lag = log(length(model2$residuals)))

model3<-arima(number_of_births, order=c(7,1,1))
SSE3<-sum(model3$residuals^2)
model3.test<-Box.test(model3$residuals, lag = log(length(model3$residuals)))

model4<-arima(number_of_births, order=c(7,1,2))
SSE4<-sum(model4$residuals^2)
model4.test<-Box.test(model4$residuals, lag = log(length(model4$residuals)))

df<-data.frame(row.names=c('AIC', 'SSE', 'p-value'), c(model1$aic, SSE1, mod
               c(model2$aic, SSE2, model2.test$p.value), c(model3$aic, SSE3,
               c(model4$aic, SSE4, model4.test$p.value))
colnames(df)<-c('Arima(0,1,1)','Arima(0,1,2)', 'Arima(7,1,1)', 'Arima(7,1,2)



format(df, scientific=FALSE)
```

In [ ]:
```
# Fit a SARIMA model

sarima(number_of_births, 0,1,2,0,0,0)
```

In [ ]:
```

```