1. Least squares estimation

(a)

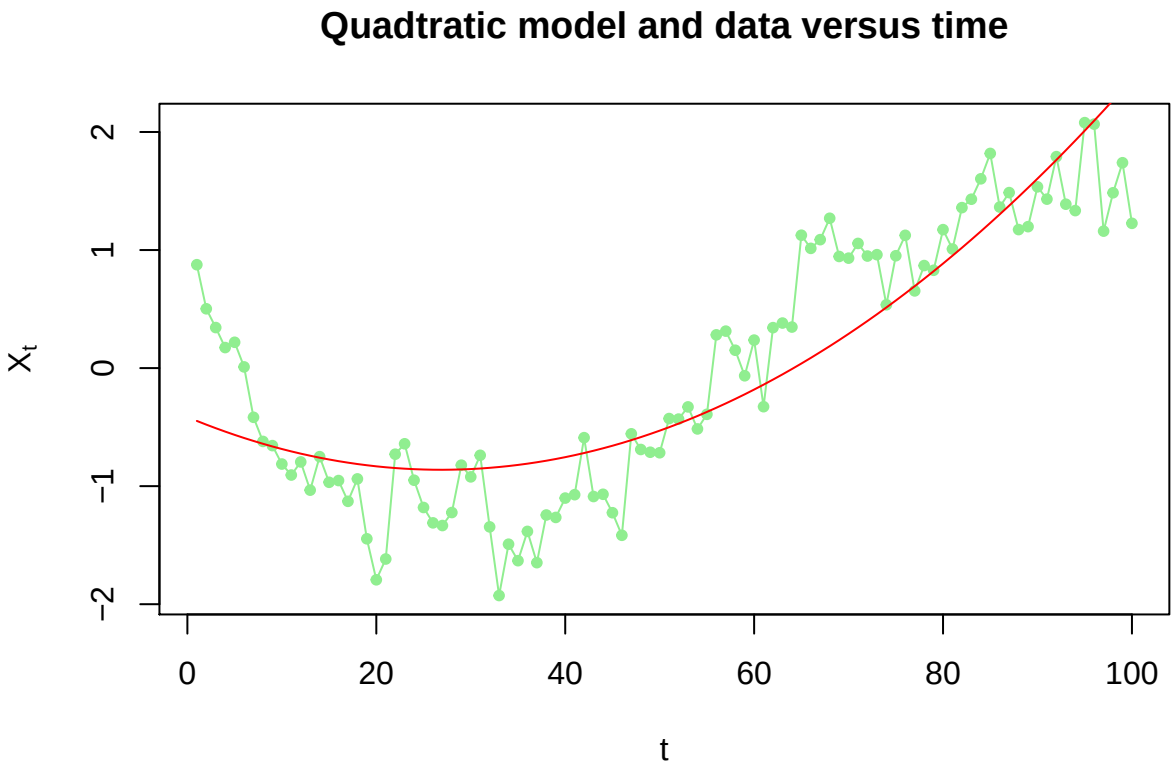
i

```
data = read.table('project4_data.txt')
time = 1:100
quad_mdl = lm(data$V1 ~ time + I(time^2))
coef(quad_mdl)
```

```
##   (Intercept)          time     I(time^2)
## -0.4147689817 -0.0332342602  0.0006186195
```

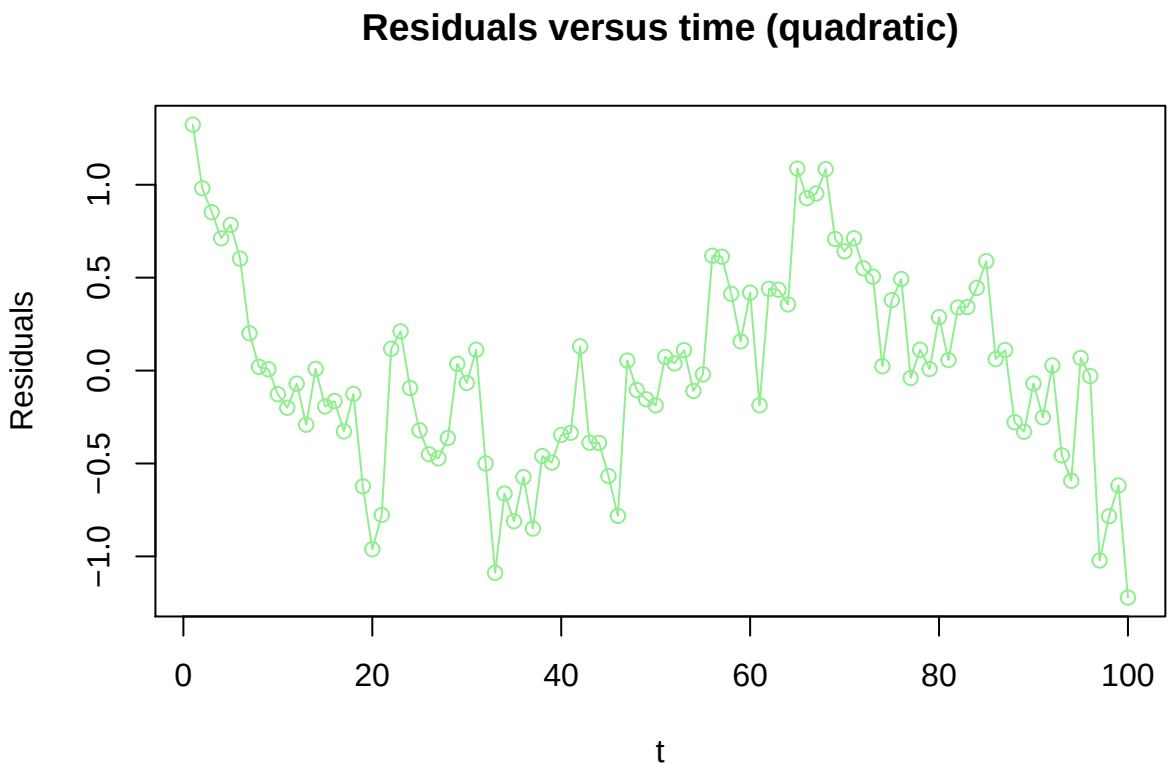
ii

```
plot(data$V1, type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'Quadtratic model and data versus time')
points(data$V1, pch = 20, col = 'lightgreen')
lines(fitted(quad_mdl), col = 'red')
```



iii

```
plot(resid(quad_mdl), type = 'o', col = 'lightgreen',
     xlab = 't', ylab = 'Residuals',
     main = 'Residuals versus time (quadratic)')
```



iv

The residuals show trend.

(b)

i

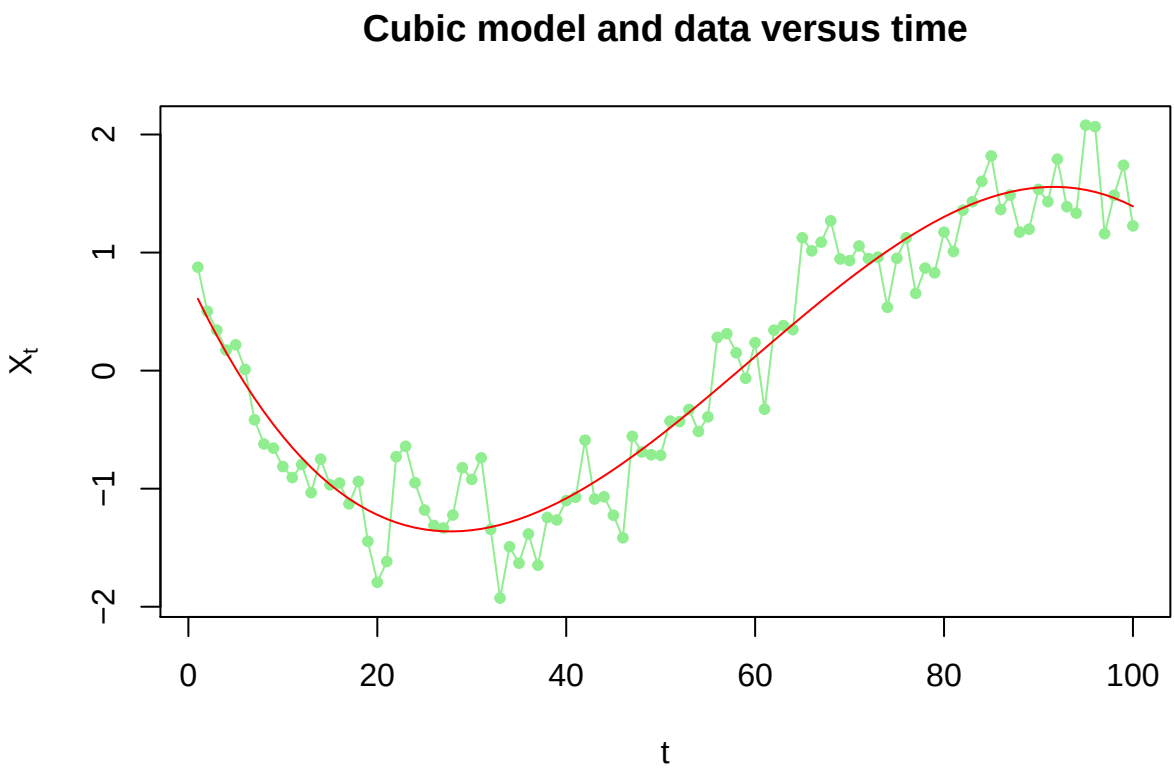
```
cubic_mdl = lm(data$V1 ~ time + I(time^2) + I(time^3))
coef(cubic_mdl)
```

```
##   (Intercept)          time     I(time^2)     I(time^3)
##  0.7760731619 -0.1712975472  0.0040190825 -0.0000224453
```

ii

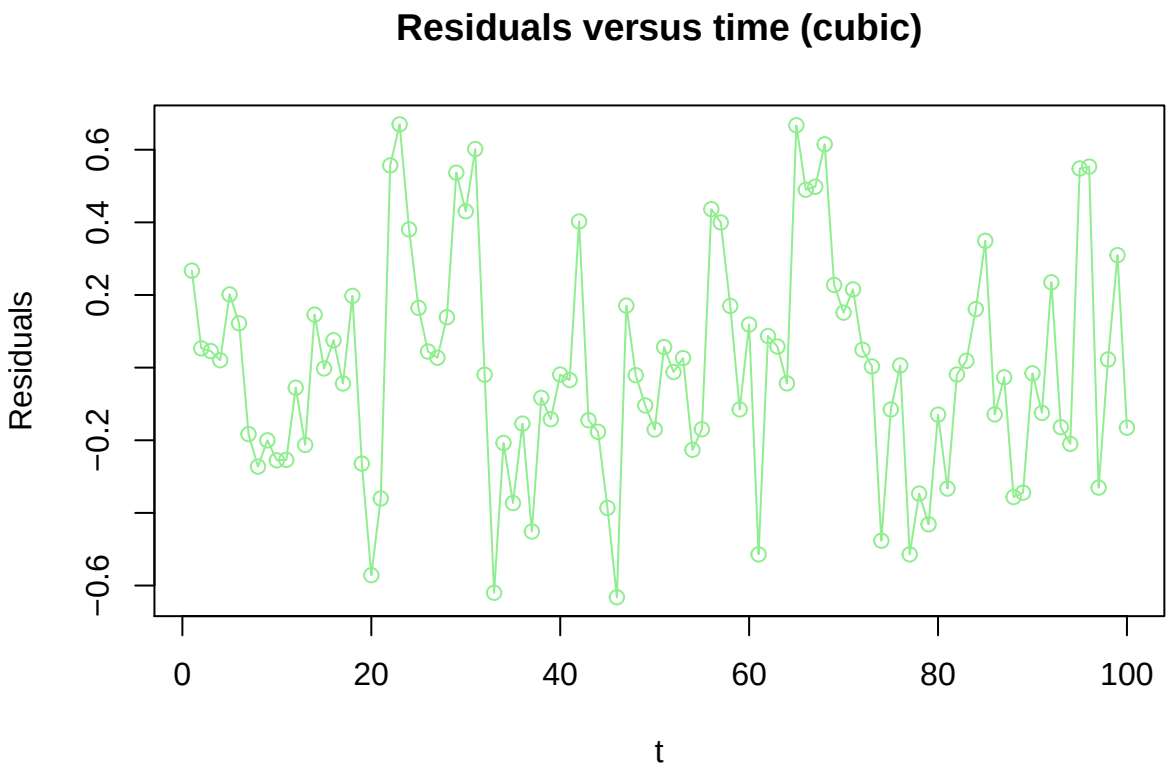
```
plot(data$V1, type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
```

```
     main = 'Cubic model and data versus time')
points(data$V1, pch = 20, col = 'lightgreen')
lines(fitted(cubic_mdl), col = 'red')
```



iii

```
plot(resid(cubic_mdl), type = 'o', col = 'lightgreen',
     xlab = 't', ylab = 'Residuals',
     main = 'Residuals versus time (cubic)')
```



iv

The residuals do not show trend.

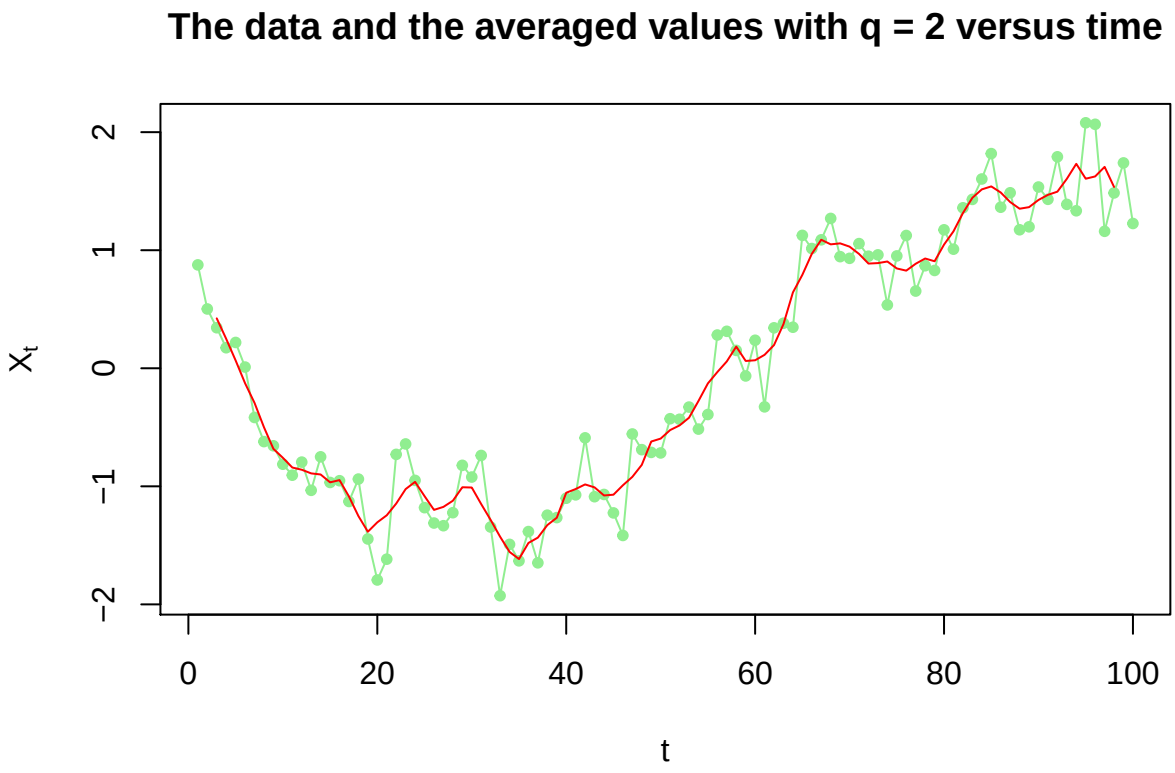
(c)

Cubic trend model is better.

2. Moving average

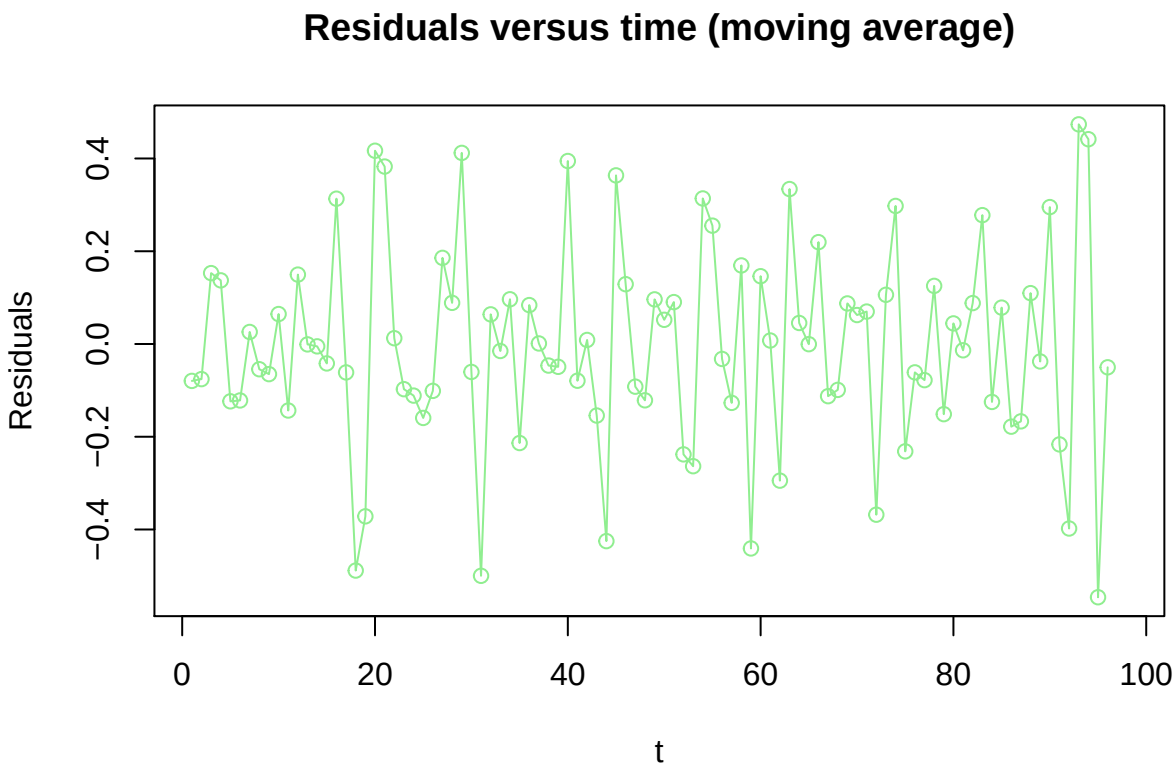
(a)

```
plot(data$V1, type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'The data and the averaged values with q = 2 versus time')
points(data$V1, pch = 20, col = 'lightgreen')
avgval = stats::filter(data$V1, filter = rep(1/5,5), sides = 2)
lines(avgval, col='red')
```



(b)

```
plot(data$V1[3:100]-avgval[3:100], type = 'o', col='lightgreen',
     xlab = 't', ylab = 'Residuals',
     main = 'Residuals versus time (moving average)')
```



(c)

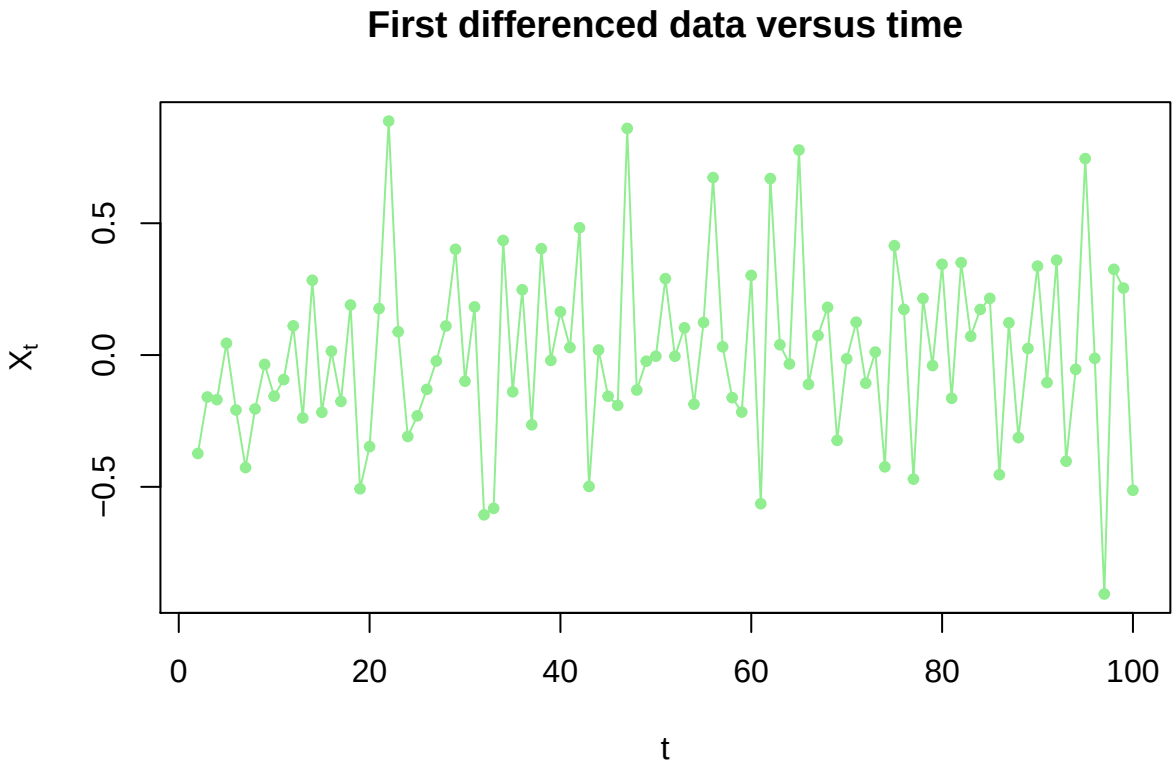
The residuals do not show trend.

3. Differencing

(a)

i

```
plot(x = 2:100, y = diff(data$V1), type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'First differenced data versus time')
points(x = 2:100, y = diff(data$V1), pch = 20, col = 'lightgreen')
```



ii

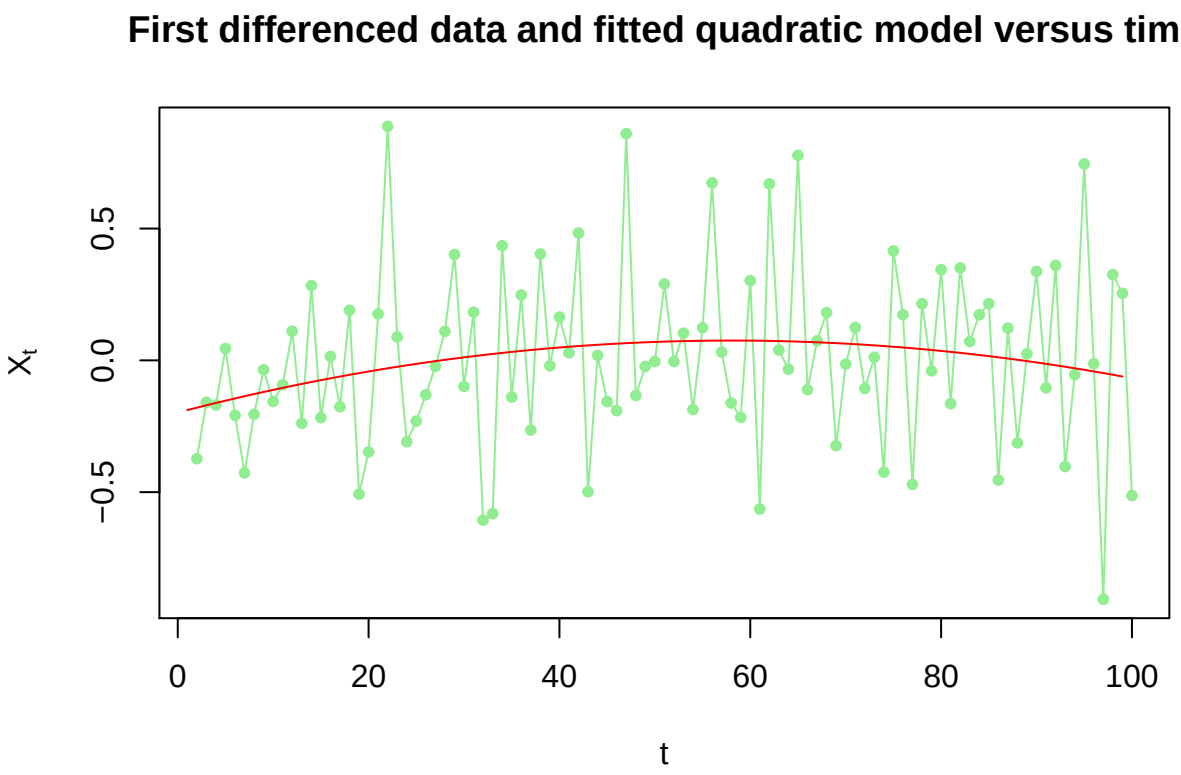
I do not think the plot of first order differencing show evidence of trend.

iii

```
time = 2:100
quad_1stdiff = lm(diff(data$V1) ~ time + I(time^2))
coef(quad_1stdiff)
```

```
##   (Intercept)          time     I(time^2)
## -0.2075238107  0.0095749939 -0.0000811257
```

```r
plot(x = 2:100, y = diff(data$V1), type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'First differenced data and fitted quadratic model versus time')
points(x = 2:100, y = diff(data$V1), pch = 20, col = 'lightgreen')
lines(fitted(quad_1stdiff), col = 'red')
```

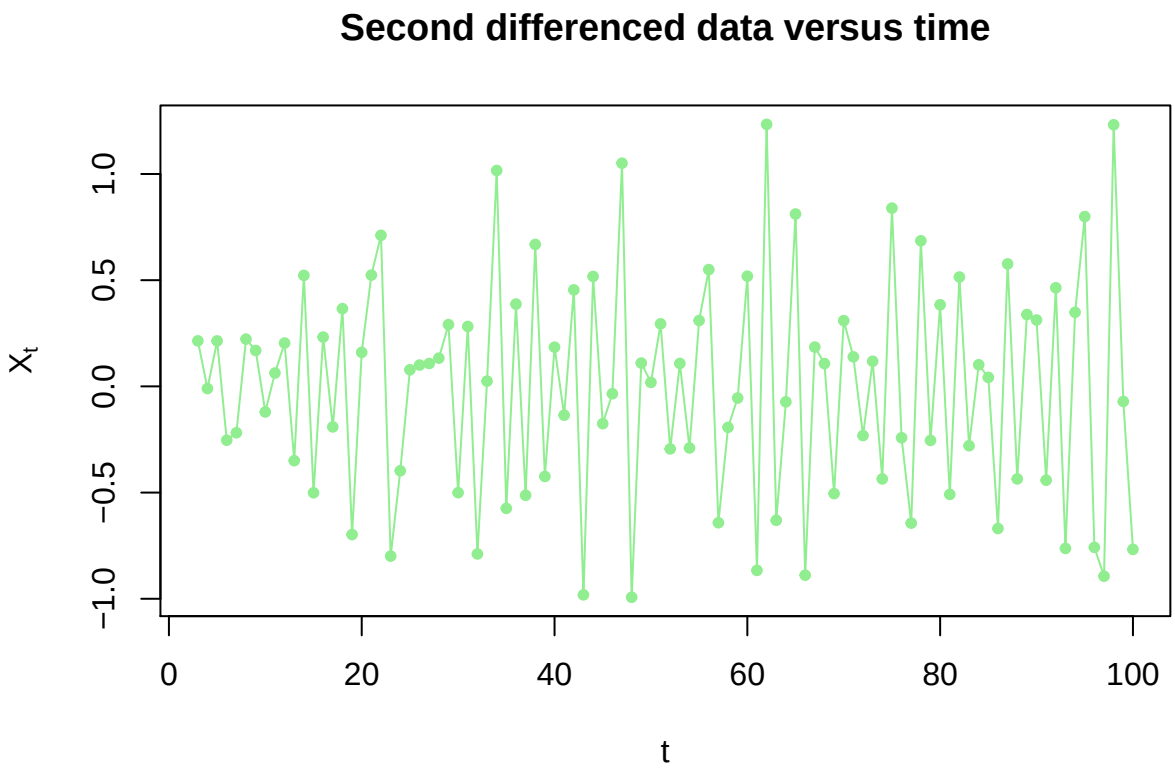


This suggest that the first order differencing have trend.

(b)

i

```r
plot(x = 3:100, y = diff(data$V1, differences = 2), type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'Second differenced data versus time')
points(x = 3:100, y = diff(data$V1, differences = 2), pch = 20, col = 'lightgreen')
```



ii

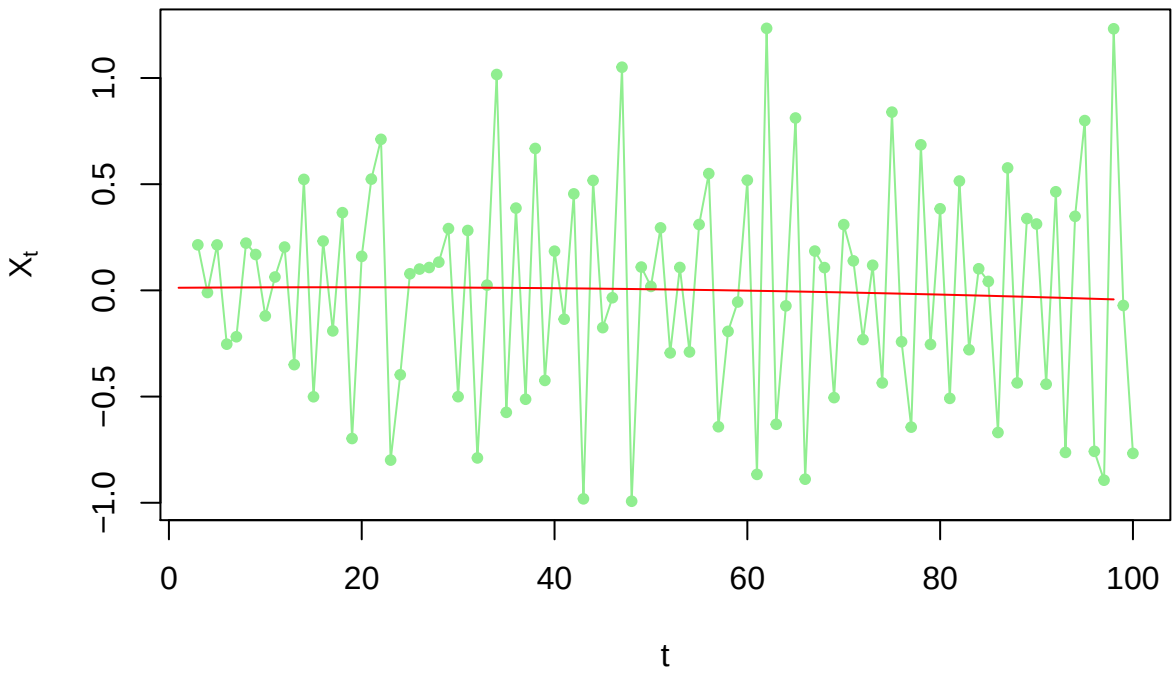
I do not think the plot of second order differencing show evidence of trend.

iii

```
time = 3:100
quad_2nddiff = lm(diff(data$V1, differences = 2) ~ time + I(time^2))
coef(quad_2nddiff)
```

```
##   (Intercept)          time     I(time^2)
##  1.136119e-02  3.404778e-04 -8.782542e-06
```

```
plot(x = 3:100, y = diff(data$V1, differences = 2), type = 'l', col = 'lightgreen',
     xlab = 't', ylab = expression(X[t]),
     main = 'Second differenced data and fitted quadratic model versus time')
points(x = 3:100, y = diff(data$V1, differences = 2), pch = 20, col = 'lightgreen')
lines(fitted(quad_2nddiff), col = 'red')
```



This suggest that the second order differencing does not have trend.