

A2-Q2

Zhaoqi Li-1006324639

2023/2/16

1.

```
library(TSA)
```

```
##
```

```
## Attaching package: 'TSA'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      acf, arima
```

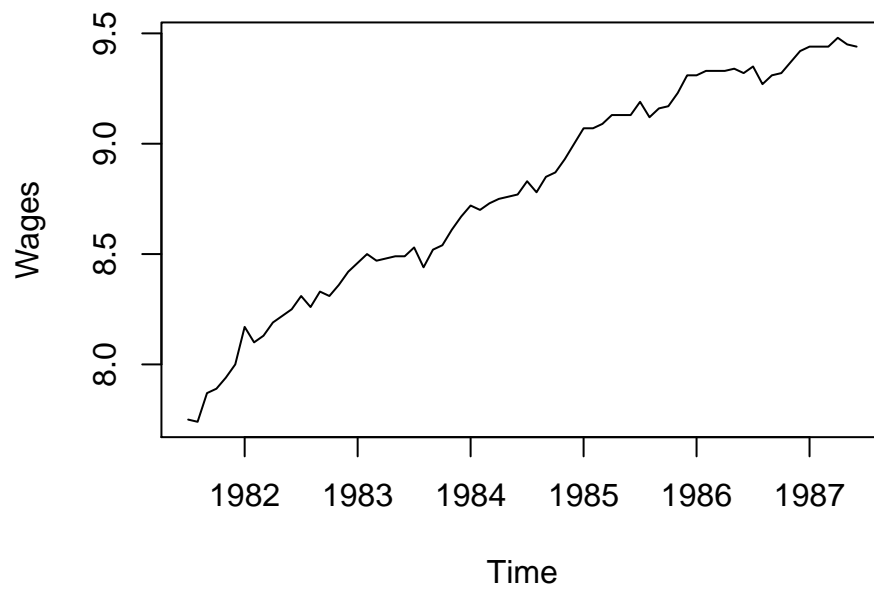
```
## The following object is masked from 'package:utils':
```

```
##
```

```
##      tar
```

```
data("wages")
```

```
plot.ts(wages)
```



There is a strong up trend of the plot

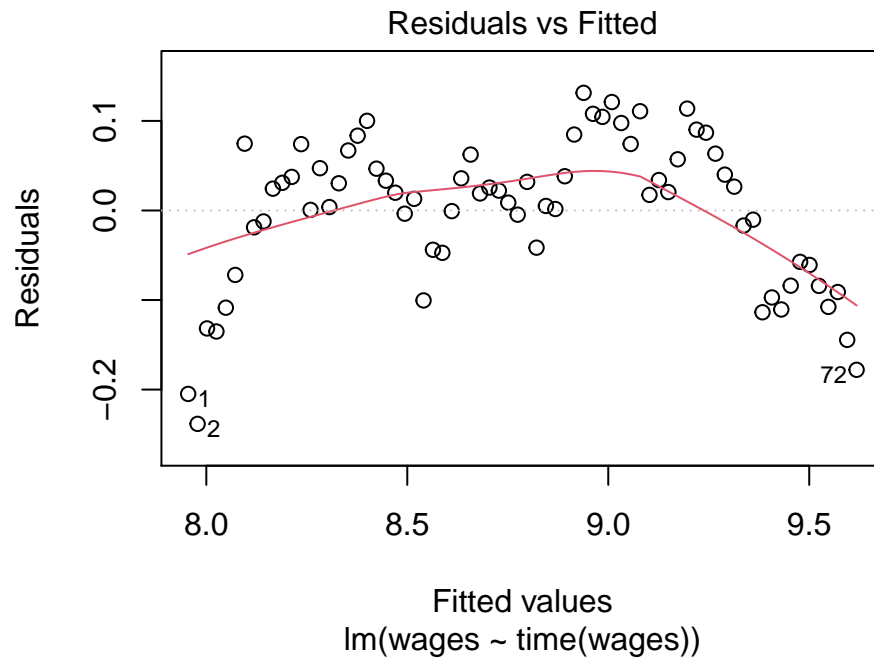
2.

```
reg1 <- lm(wages ~ time(wages))
summary(reg1)

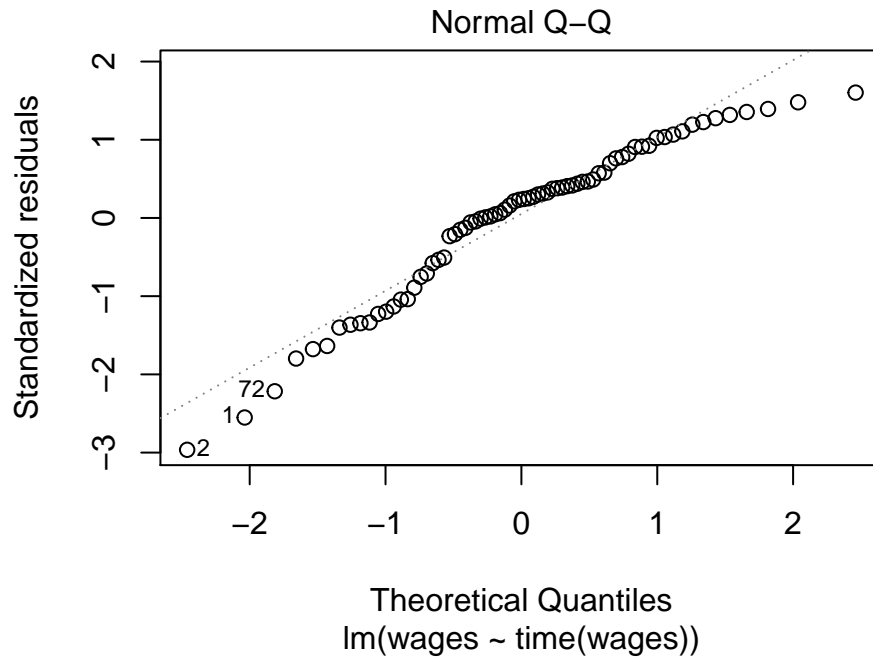
##
## Call:
## lm(formula = wages ~ time(wages))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23828 -0.04981  0.01942  0.05845  0.13136
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.490e+02  1.115e+01  -49.24  <2e-16 ***
## time(wages)  2.811e-01  5.618e-03   50.03  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08257 on 70 degrees of freedom
## Multiple R-squared:  0.9728, Adjusted R-squared:  0.9724
## F-statistic: 2503 on 1 and 70 DF, p-value: < 2.2e-16
```

3.

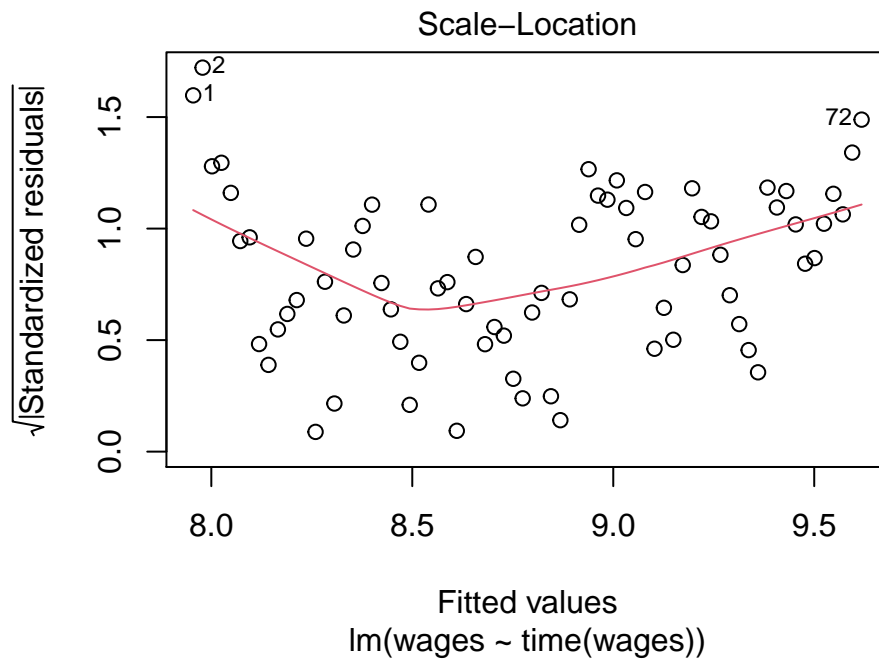
```
plot(reg1,1)
```



```
plot(reg1,2)
```



```
plot(reg1,3)
```



According to residual plots, there is a systematic pattern in the plot, so they are not random. It means the linearity is not satisfied. For the normality, there are obvious deviations on both sides and the points of middle do not fall on the line. Therefore, the normality is also not satisfied.

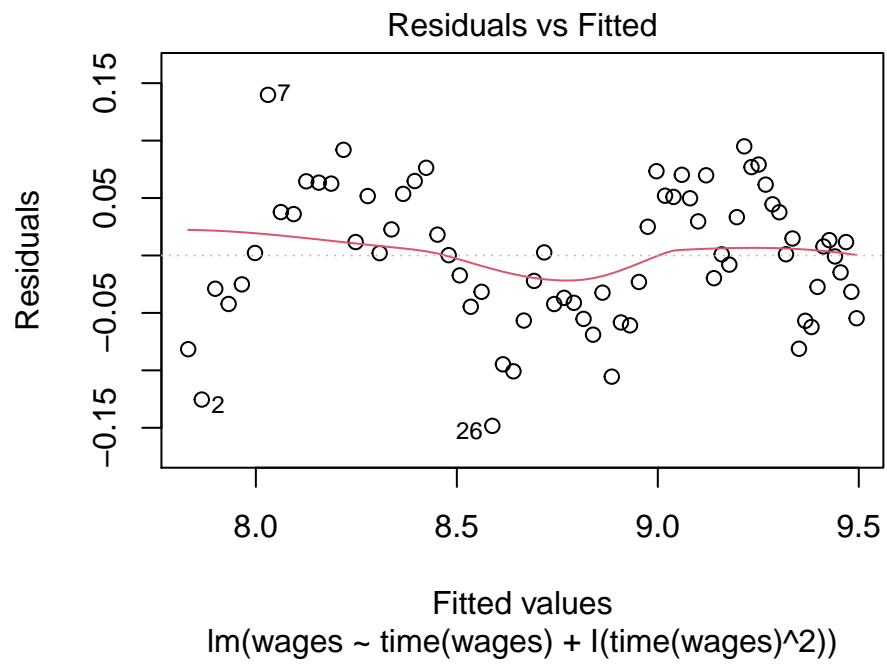
4.

```
reg2 <- lm(wages ~ time(wages) + I(time(wages)^2))
summary(reg2)

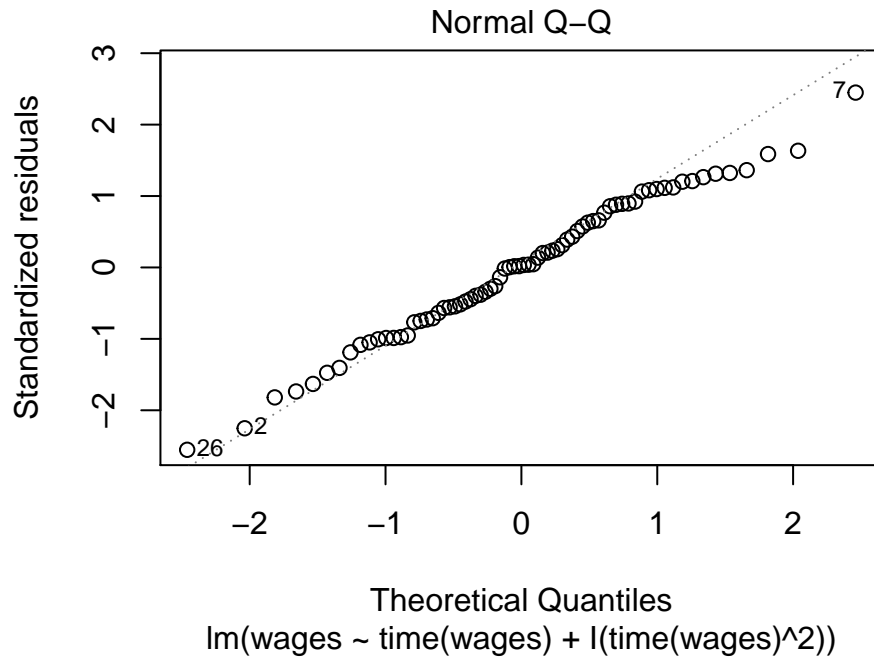
##
## Call:
## lm(formula = wages ~ time(wages) + I(time(wages)^2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.148318 -0.041440  0.001563  0.050089  0.139839
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -8.495e+04  1.019e+04  -8.336 4.87e-12 ***
## time(wages)    8.534e+01  1.027e+01   8.309 5.44e-12 ***
## I(time(wages)^2) -2.143e-02  2.588e-03  -8.282 6.10e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05889 on 69 degrees of freedom
## Multiple R-squared:  0.9864, Adjusted R-squared:  0.986
## F-statistic: 2494 on 2 and 69 DF,  p-value: < 2.2e-16
```

5.

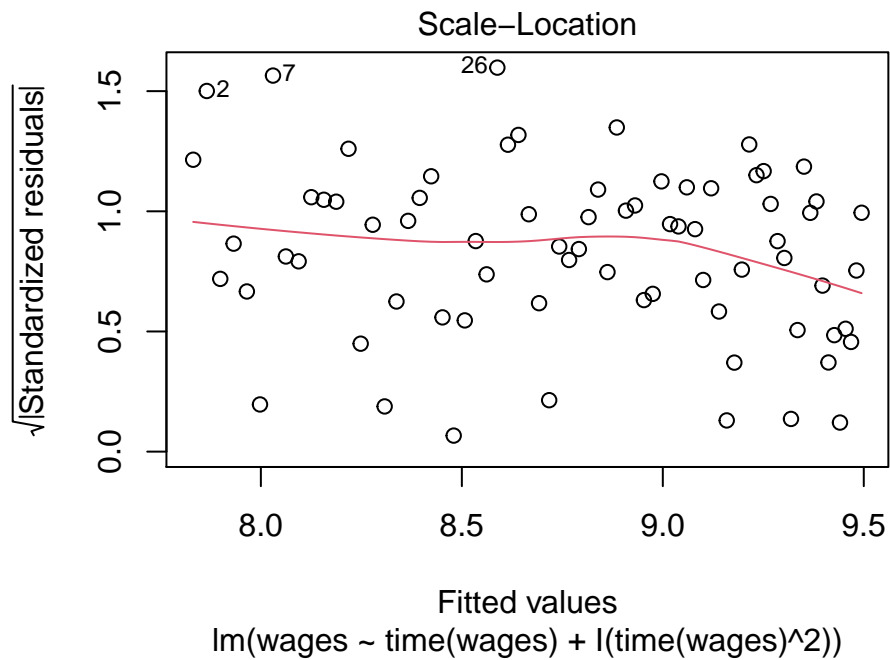
```
plot(reg2,1)
```



```
plot(reg2,2)
```



```
plot(reg2,3)
```



The assumptions are still not valid, but better than the first model. For the linearity, there is a systematic pattern in the plot like m shape, so they are not random. For the normality, there are obvious deviations on both sides. Therefore, this model do not satisfy linearity and normality.

6.

```
anova(reg1,reg2)
```

```
## Analysis of Variance Table
##
## Model 1: wages ~ time(wages)
## Model 2: wages ~ time(wages) + I(time(wages)^2)
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      70 0.47721
## 2      69 0.23932  1   0.23789 68.586 6.104e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$H_0 : \beta_2 = 0$$

$$H_a : \beta_2 \neq 0$$

According to the result, the test statistic is 68.586 and the p-value is 6.104e-12 which is less than 0.05. Therefore, we reject H_0 , and β_2 is not equal to 0.