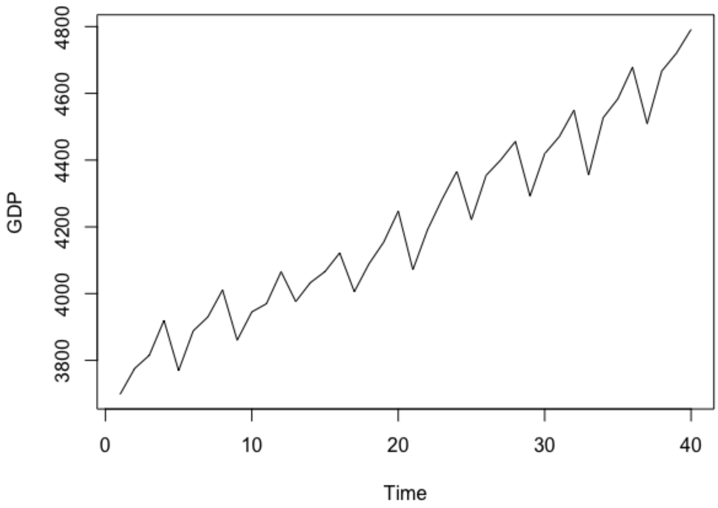
1) Timeseries plot:

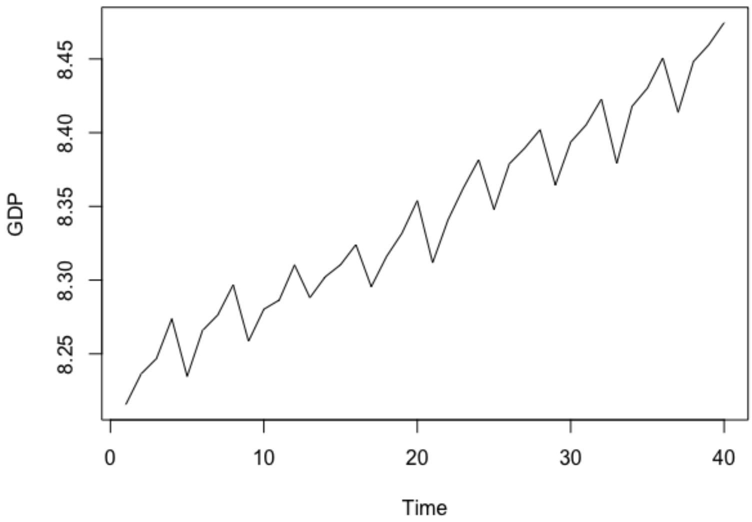
```
plot.ts(timeseriesdata[3])
```



➔ As you can see from the timeseries plot, the mean GDP is increasing as time increases, this could be the possible trend component.

2) Log transformed timeseries plot:

```
plot.ts(log(timeseriesdata[3]))
```



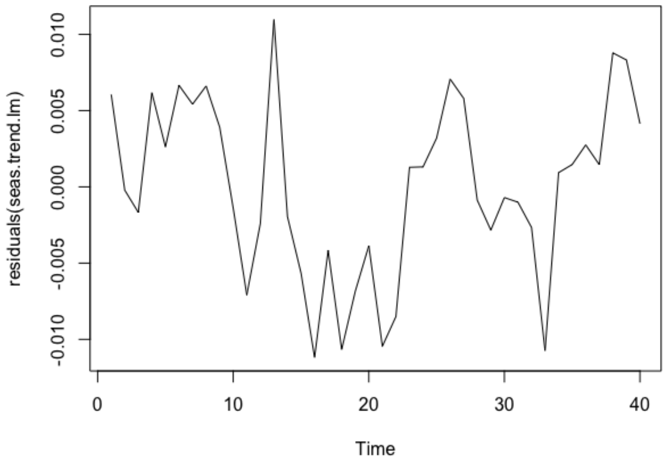
➔ The main difference in the plots are the units of the y-axis. This plot also shows the trend component of how mean GDP increases as time increases.

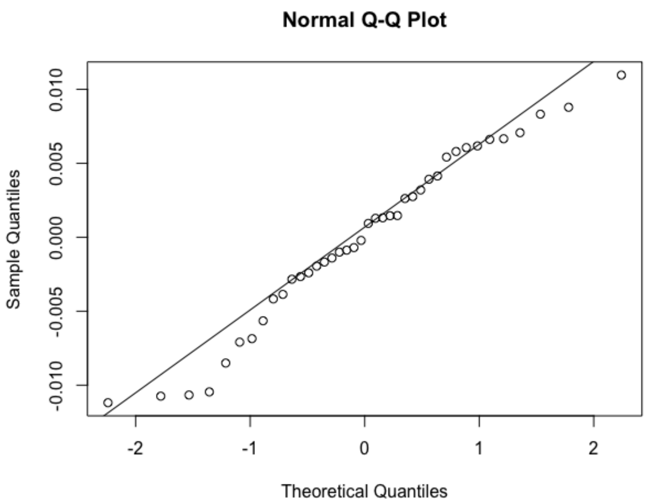
3) Linear model with seasonal and linear trend accounted:

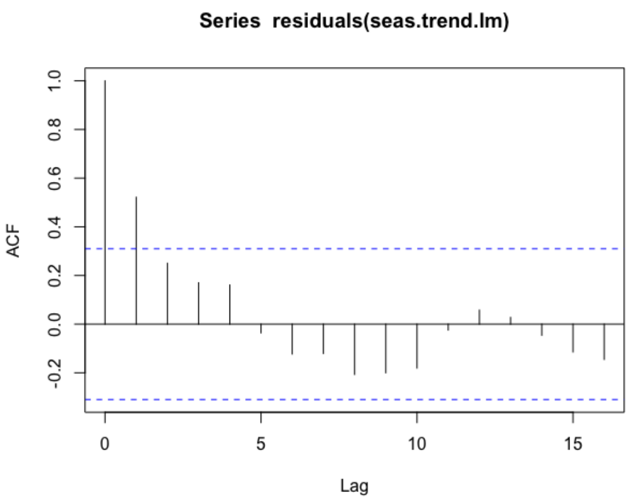
```
> quarter=c(rep(c('Q1','Q2','Q3','Q4'),length=40))
> quarter=factor(quarter,levels=c('Q1','Q2','Q3','Q4'))
> GDP = log(timeseriesdata[3])
> GDP = unlist(GDP)
> seas.lm = lm(GDP ~ quarter)
> t=c(1:40)
> seas.trend.lm=lm(GDP~quarter+t)
> summary(seas.trend.lm)

Call:
lm(formula = GDP ~ quarter + t)

Residuals:
       Min         1Q     Median         3Q        Max 
-0.0111854 -0.0030939  0.0003585  0.0044590  0.0109620 

Coefficients:
             Estimate Std. Error  t value Pr(>|t|)    
(Intercept) 8.204e+00  2.536e-03 3234.947  < 2e-16 ***
quarterQ2   2.145e-02  2.765e-03    7.757 4.14e-09 ***
quarterQ3   2.770e-02  2.769e-03   10.002 8.44e-12 ***
quarterQ4   4.123e-02  2.776e-03   14.854  < 2e-16 ***
t           5.635e-03  8.505e-05   66.256  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.00618 on 35 degrees of freedom
Multiple R-squared:  0.9928,	Adjusted R-squared:  0.992 
F-statistic:  1214 on 4 and 35 DF,  p-value: < 2.2e-16
```

➔ The linear model is very well fit. Accounting for seasonal (quarter) and trend terms explains almost all the variation in GDP in the model. This can be seen by the R^2_{adj} = 0.992 and the very low small p-value associated with the linear term.

4) time series plot of the residuals, QQ plot, and correlogram:







- ➔ Looking at the time series of residuals it seems like there exists a quadratic trend. The correlogram of the residuals indicate that the trend component is <u>not</u> completely accounted for. Because of the large value at lag=1. The normality of errors assumption also seems to be not completely met from the QQ plot.

5) After adding quadratic trend component:

```
> tsq = t^2
> seas.quadtrend.lm=lm(GDP~quarter+t+tsq)
> summary(seas.quadtrend.lm)

Call:
lm(formula = GDP ~ quarter + t + tsq)

Residuals:
       Min         1Q     Median         3Q        Max 
-0.0112731 -0.0027532 -0.0003484  0.0040286  0.0128477 

Coefficients:
             Estimate Std. Error  t value Pr(>|t|)    
(Intercept) 8.211e+00  3.041e-03 2700.454  < 2e-16 ***
quarterQ2   2.150e-02  2.433e-03    8.836 2.51e-10 ***
quarterQ3   2.774e-02  2.437e-03   11.387 3.79e-13 ***
quarterQ4   4.123e-02  2.442e-03   16.882  < 2e-16 ***
t           4.644e-03  3.054e-04   15.207  < 2e-16 ***
tsq         2.418e-05  7.221e-06    3.348    0.002 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.005438 on 34 degrees of freedom
Multiple R-squared:  0.9946,    Adjusted R-squared:  0.9938 
F-statistic:  1256 on 5 and 34 DF,  p-value: < 2.2e-16
```

- ➔ The model is very well fit. Accounting for seasonal (quarter) and quadratic trend terms explains almost all the variation in GDP in the model. The variation explained by the model has increased from the linear trend model. This can be seen by the increase in R^2_{adj} = 0.9938 and the very low small p-value associated with the quadratic term.

6)

- ➔ Based on the estimates and associated p-values, Q4 had highest GDP during the period of the data.
- ➔ Looking at both the positive trend term estimates and their low p-values, it is consistent that the mean GDP was increasing as time increases.

7) Forecasted GDP for Q1 of 2019:

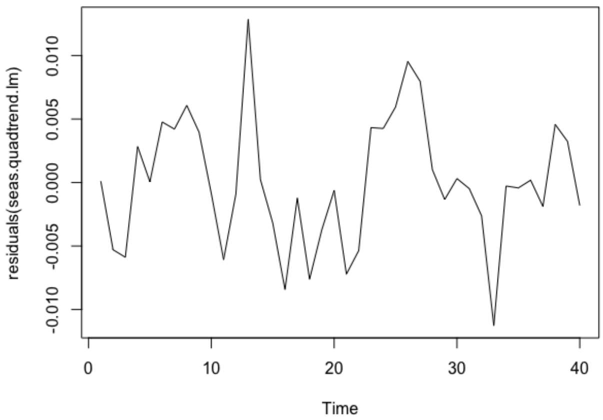
$y_{41} = \beta_0 + \beta_4(41) + \beta_5(41)^2$

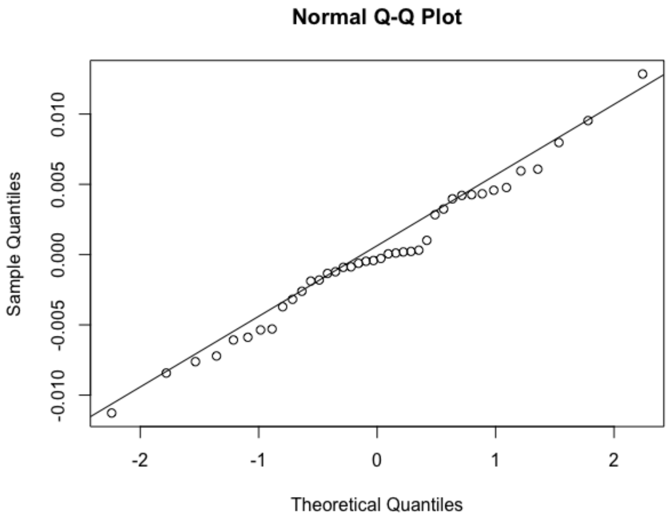
```
> 8.211e+00 + 41*4.644e-03 + (41^2) * 2.418e-05
[1] 8.442051
```

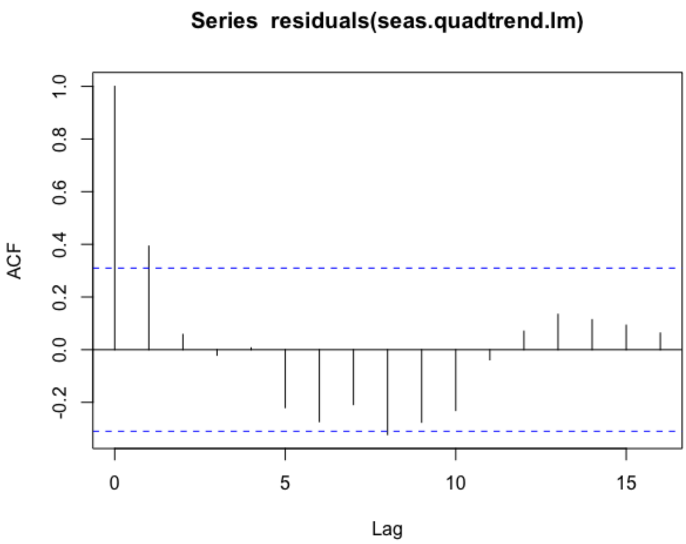
```
> exp(8.442051)
[1] 4638.058
```

- ➔ Forecasted GDP for Q1 of 2019 = \$4638.058 * 10^9

8) Replotting plots of 4):







➔ The model is a much better fit now with the quadratic trend component. Now the residuals don't seem to be exhibiting any trend, the independence assumption seems to be satisfied. Normality assumption of errors seems to stand. The trend component seems to be reduced too in the correlogram but still exists. The assumptions seem to be satisfied better in this new model.

9) $DW = 2 - 2r_1 = 2 - 0.8 = 1.2$
Alpha = 0.05, n = 40, k =4
$DW_L = 1.285$
$DW_U = 1.721$
Above values from Durbin-Watson table.

Since $1.2 < DW_L$ we conclude that true autocorrelation > 0.

10) DW test:

```
> dwtest (seas.quadtrend.lm)

        Durbin-Watson test

data:  seas.quadtrend.lm
DW = 1.2102, p-value = 0.00336
alternative hypothesis: true autocorrelation is greater than 0
```

➔ Consistent with our previous conclusion in 9).

11) AR (1) model:

```
> ar1.sim.arima=arima(residuals(seas.quadtrend.lm),c(1,0,0))
> ar1.sim.arima

Call:
arima(x = residuals(seas.quadtrend.lm), order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.3850     0.0000
s.e.  0.1432     0.0012

sigma^2 estimated as 2.124e-05:  log likelihood = 158.36,  aic = -310.71
> predict(ar1.sim.arima)
$pred
Time Series:
Start = 41
End = 41
Frequency = 1
[1] -0.0007119621

$se
Time Series:
Start = 41
End = 41
Frequency = 1
[1] 0.0046085
```

12)

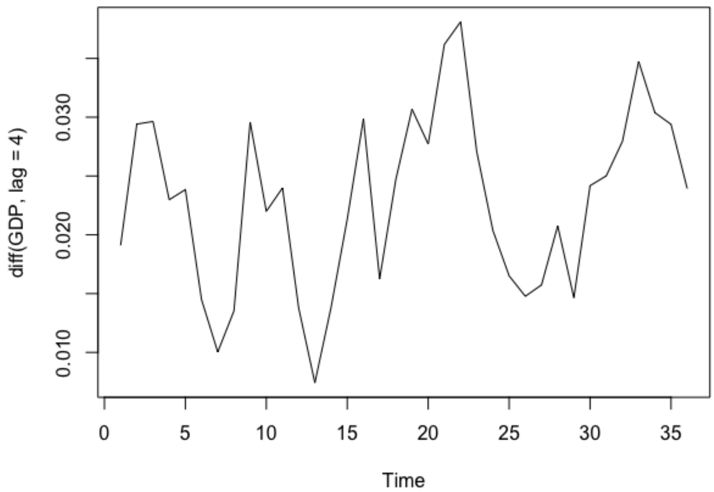
```
> exp(0.0046085)
[1] 1.004619
```

Revised forecasted GDP = 4638.058 + 1.004619 = $4639.063

13)

a)

```
plot.ts(diff(GDP,lag = 4))
```



➔ The plot seems to have no seasonal component, but a positive trend component is visible.

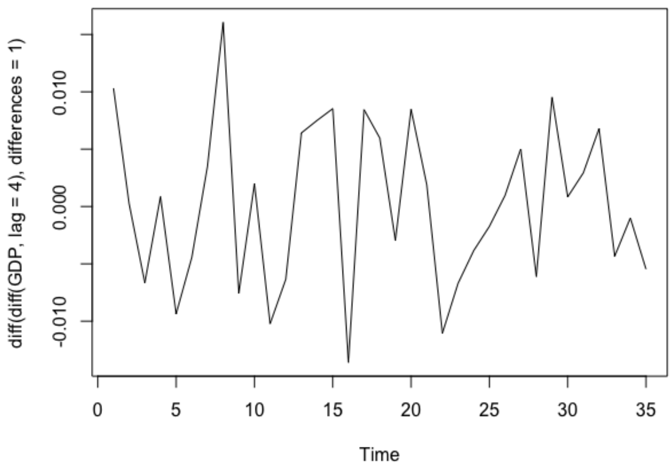
b)

```
> diff(diff(GDP,lag = 4),differences = 1)
```

This removes the trend on the seasonally differenced series.

c)

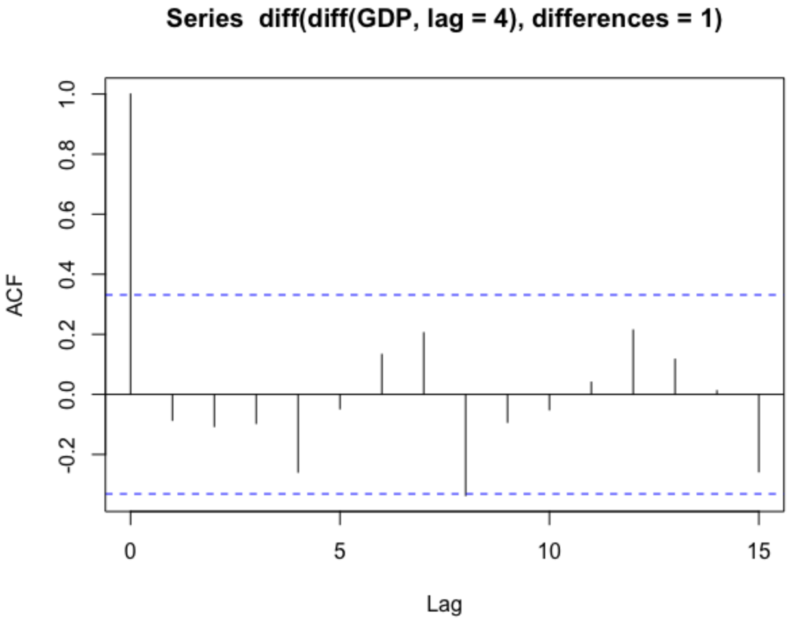
```
> plot.ts(diff(diff(GDP,lag = 4),differences = 1))
```



➔ Yes, the trend has been eliminated.

d)

```
acf(diff(diff(GDP,lag = 4),differences = 1))
```



➔ There's <u>no</u> significant trend or seasonal autocorrelation in the time series data after differencing. This can be seen from the above correlogram.