Assignment#1

```r
#install.packages('tidyverse')
library(tidyverse)
```

```
## -- Attaching packages ------------------------------------------------------------------------

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts ---------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

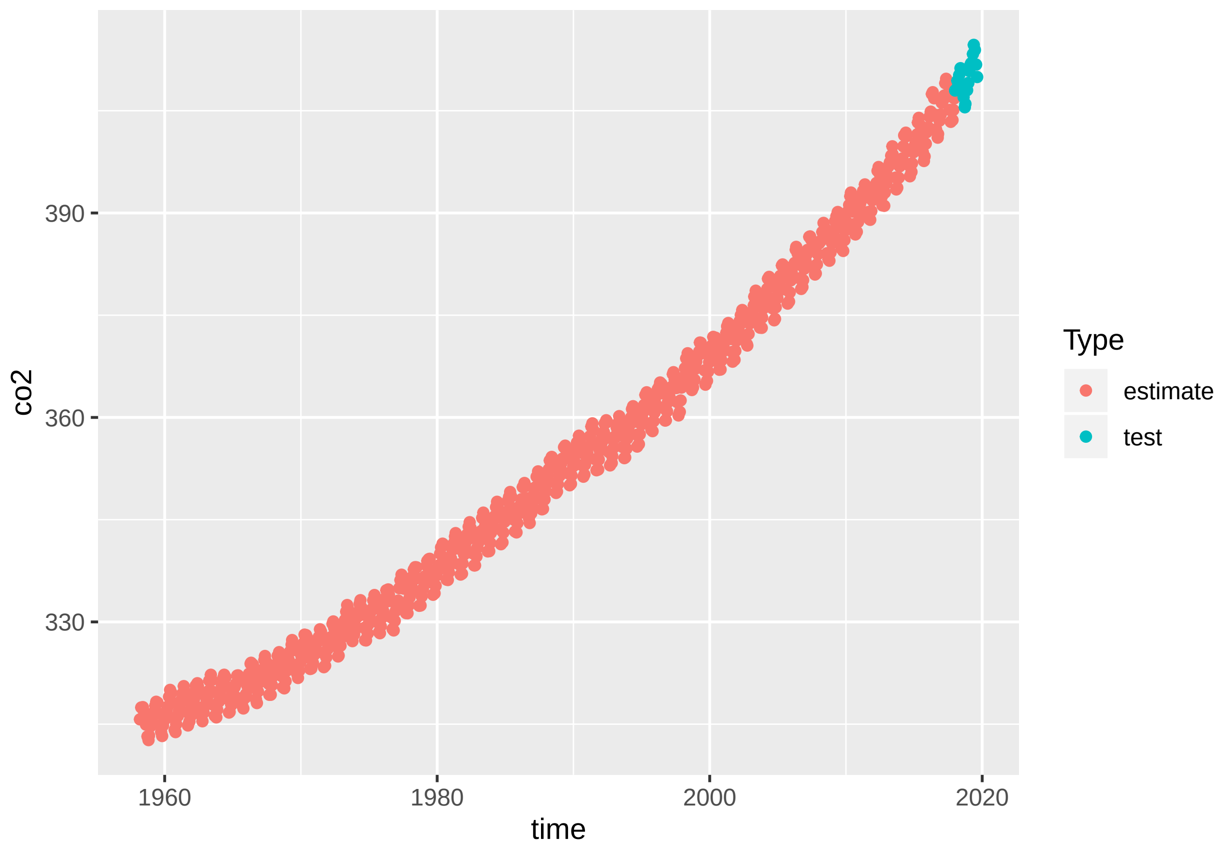
Here, the data set was imported. Then, the dataset was divided as the assignment requested and and the estimate data and the test data in two different color has been plotted.

```r
DATA <- read.csv('A1_co2.txt',sep = '')

DATA_est <-  DATA %>%
  filter(year !=2019 & year !=2018) %>%
  mutate(Type='estimate')
#tail(DATA_estimate)
DATA_test <- DATA %>%
  filter(year ==2019 | year ==2018) %>%
  mutate(Type='test')

DATA_est_tes <- bind_rows(DATA_est,DATA_test)

#
ggplot(data = DATA_est_tes,aes(x=time,y=co2)) +
  geom_jitter(aes(colour = Type))
```




```
# tail(Data)
```

The lm function was used here in order to perform the Ordinary Least Sqaure based on the the training data set provide in the previous section.

```
p <- 365
#x <- cbind(1,DATA_est$time,sin(2*pi*DATA_est$time/p),cos(2*pi*DATA_est$time/p))

lm1 <- lm(co2 ~ time + I(sin(2*pi*time/p)) + I(cos(2*pi*time/p)), data=DATA_est) ## Notice the I(...) t
```

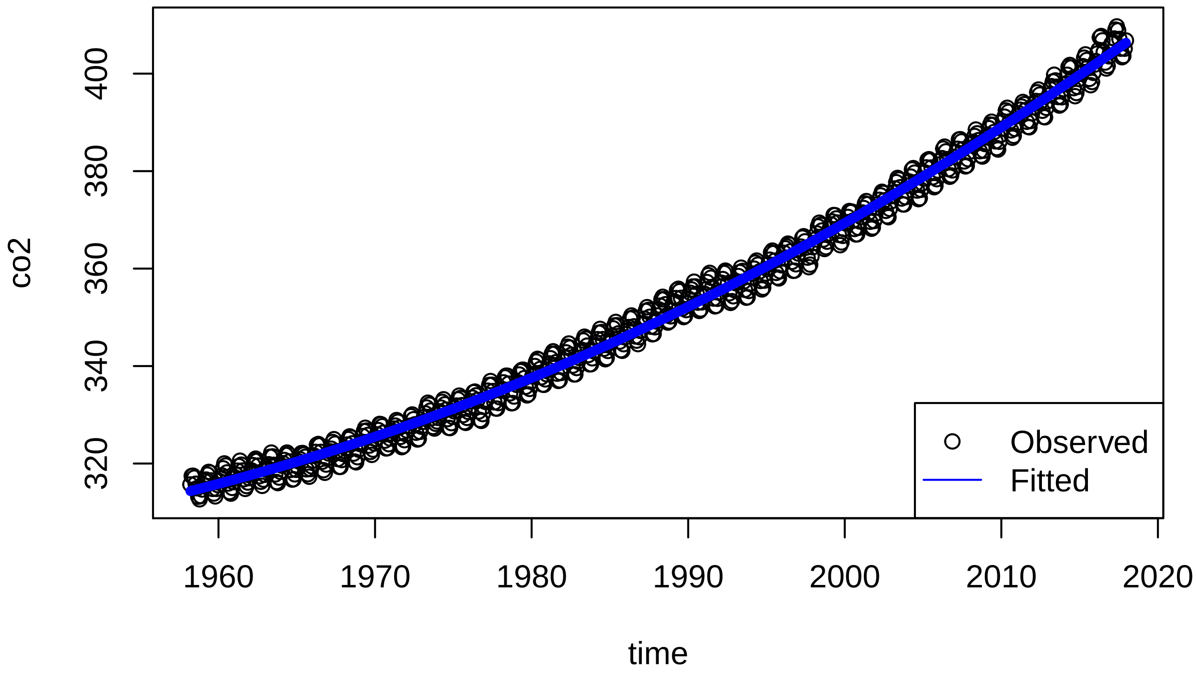
Here, the estimated parameters and as well the standard deviation of each parameters (4 in this case) were estimated.

```
sum_param=summary(lm1)$coefficient
row.names(sum_param) <- c('alpha','beta_t','beta_s','beta_c')
sum_param
```

```
##             Estimate  Std. Error   t value     Pr(>|t|)
## alpha   -2649.759905 814.0681612 -3.254961 1.187573e-03
## beta_t      1.551843   0.4094726  3.789859 1.634528e-04
## beta_s    -27.579930  23.0858185 -1.194670 2.326128e-01
## beta_c     81.721452   8.2633543  9.889622 1.062263e-21
```

Plot of the data Vs. Fitted values based on Ordinary Least Sqaure:

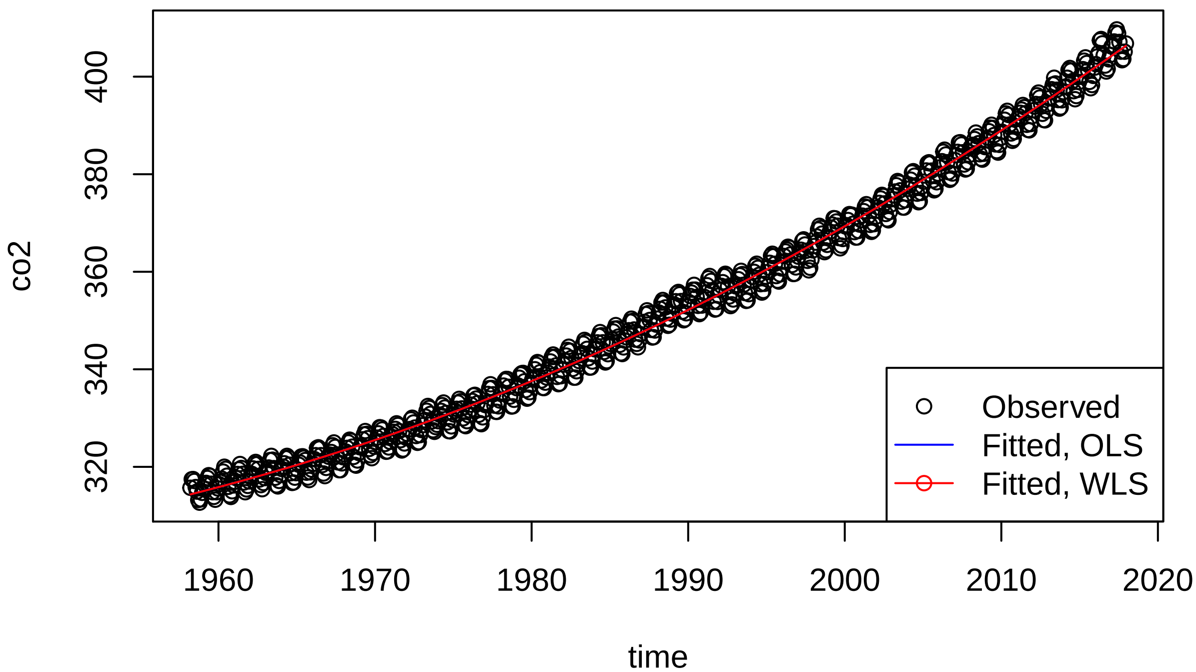
```
DATA_est$fit <- lm1$fitted.values
plot(co2~time, DATA_est)
lines(fit~time, DATA_est, type="l", lwd=5, col = "blue")
legend("bottomright", legend = c("Observed","Fitted"), col = c('black','blue'), pch=c(1,NA), lty=c(NA,1
```



here, the algorithem was provided in order to ro estimate the rho values after 5 iterations. At the end, the plot of the WLS method based on the estimated paramters (after 5 iteraions) were plotted:

```
x <- cbind(1,DATA_est$time,sin(2*pi*DATA_est$time/p),cos(2*pi*DATA_est$time/p))
sigma <-  diag(718)
Y <- DATA_est$co2
theta_relax <- c()
for (k in 1:5){
        theta_relax <-  solve((t(x)%*%solve(sigma)%*%x))%*%t(x)%*%solve(sigma)%*%Y
               Y_hat <-  x%*%theta_relax
               res <-  DATA_est$co2 - Y_hat
               res1 <-  res[1:718-1]
               res2 <-  res[2:718]
               rho <-  cor(res1,res2)
               J <- 1:718
               P <- rho^(J-1)
               sigma <-  toeplitz(P)
               Y <-  Y_hat
}
DATA_est$wfit <- Y_hat
plot(co2~time, DATA_est)
```

```r
lines(Y_hat~time, DATA_est, type="l", lwd=1, col = "blue")
points(fit~time, DATA_est, type="l", lwd=1, col = "red")
legend("bottomright", legend = c("Observed","Fitted, OLS","Fitted, WLS"), col = c('black','blue','red')
```



The difference between two plots are negligible and two methods almost rprovide identical solution.

The L value of the Linear + harmonic model in the case of the problem could be written as:

$$\begin{matrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & cos(2\pi/p) & sin(2\pi/p) \\ 0 & 0 & -sin(2\pi/p) & cos(2\pi/p) \end{matrix}$$

The f(0) as well could be written as:

$$\begin{matrix} 1 \\ 0 \\ 0 \\ 1 \end{matrix}$$