# Homework Solutions
## Applied Regression Analysis

## WEEK 4

**Exercise Two:** Equations and Plotting

**Using the output from exercise one, obtain the following:**

1. The estimated equation of the straight-line (degree 1) regression of Y on X

> To generate an estimated equation, we must fit the regression in order to obtain the least squares estimates by typing "regress" into the command window, followed by the dependent and independent variables. From the output, you can obtain the coefficient for the slope ($\beta_1$) as well as the intercept ($\beta_0$) in the bottom right corner of the output in the "Coef." column. You can then substitute these values into an equation for the estimated straight line regression (see output below).

> After you fit the regression, type "predict yhat" into the command window in order to generate fitted values for Y under the given model. This will generate a new variable called "yhat" (Note: as you continue to fit several regressions, you must change the name of "yhat" for each regression (eg. predict yhat2, predict yhat3, predict yhat4....). Otherwise, you will get an error as you have already generated a variable under the same name). Graph the fitted regression line onto the scatterplot of the data by typing "scatter yhat" followed by the independent and independent variables, then type ",connect (1 .) symbol (i o)". This latter command will connect the yhat to form a regression line, while plotting the data as dots.

```
. regress concen time

      Source |       SS       df       MS              Number of obs =      18
-------------+------------------------------           F(  1,    16) =   43.70
       Model |  12.7054079      1  12.7054079          Prob > F      =  0.0000
    Residual |  4.65200556     16  .290750347          R-squared     =  0.7320
-------------+------------------------------           Adj R-squared =  0.7152
       Total |  17.3574134     17  1.02102432          Root MSE      =  .53921

------------------------------------------------------------------------------
      concen |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   .2459714   .0372092     6.610   0.000     .1670914    .3248514
       _cons |  -1.931797   .4285795    -4.507   0.000    -2.840345   -1.023249
------------------------------------------------------------------------------
```

Estimated equation of the straight-line regression:

$$y = -1.931797 + 0.245714 * time$$

```
. predict yhat
(option xb assumed; fitted values)
```

```
. scatter yhat concen time, connect(l .)symbol(i o)
```
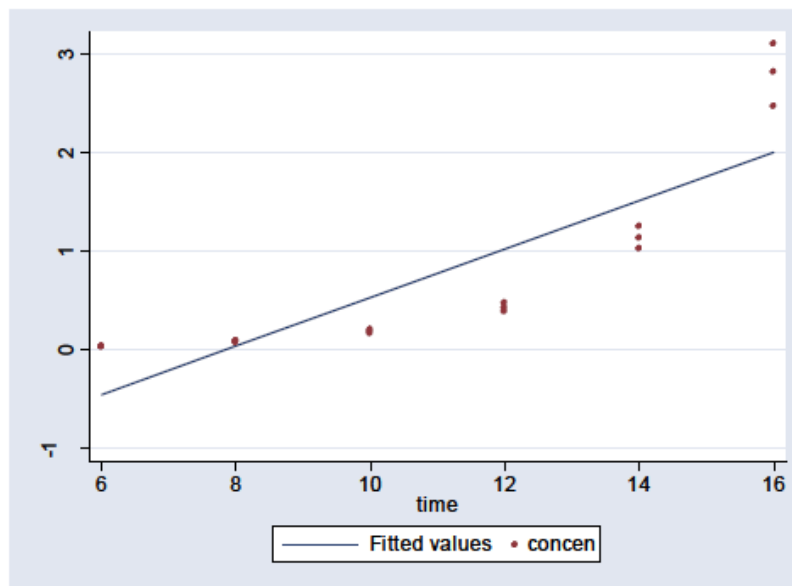
2. The estimated equation of the quadratic (degree 2) regression of Y on X

Recall that quadratic regressions include both X and $X^2$ into the model. Thus, you must first generate a variable for $X^2$ by typing "gen time2=time^2" into the command box. You can then fit the quadratic regression by typing "regress concen time time2" into the command box. Proceed to then substitute these both slope coefficients into the estimated equation for the straight line regression, making sure to include $time^2$ as the second variable (see output below).

```
. gen time2=time^2
. regress concen time time2

    Source |       SS       df       MS                  Number of obs =      18
-----------+------------------------------               F(  2,    15) =  166.79
     Model | 16.6104749       2  8.30523743              Prob > F      =  0.0000
  Residual |  .74693859      15  .049795906              R-squared     =  0.9570
-----------+------------------------------               Adj R-squared =  0.9512
     Total | 17.3574134      17  1.02102432              Root MSE      =  .22315

------------------------------------------------------------------------------
    concen |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
      time |  -.7810226   .1169892    -6.676   0.000    -1.030379   -.5316661
     time2 |   .0466815   .0052714     8.856   0.000     .0354458    .0579173
     _cons |   3.172052   .6030163     5.260   0.000     1.886754    4.457351
------------------------------------------------------------------------------
```

Estimated equation of the quadratic regression:

$$y = 3.172052 - 0.7810226*time + 0.0466815*time^2$$

```
. predict yhat2
(option xb assumed; fitted values)

. scatter yhat2 concen time, connect(l .) symbol(i o)
```
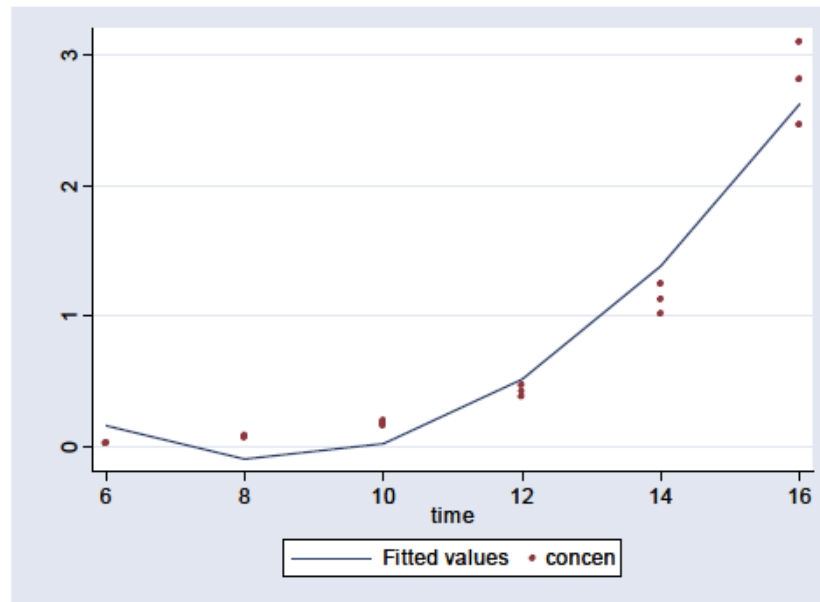


alternatively,

```
. twoway (scatter concen time, sort) (qfit yhat2 time, sort)
```

3. The estimated equation of the straight-line (degree 1) regression of lnY on X

```
. regress lnconc time

    Source |       SS        df       MS                   Number of obs =       18
-----------+------------------------------               F(  1,    16) =  4555.99
     Model | 42.7523927        1   42.7523927            Prob > F      =   0.0000
  Residual | .150140289       16   .009383768            R-squared     =   0.9965
-----------+------------------------------               Adj R-squared =   0.9963
     Total | 42.902533        17   2.52367841            Root MSE      =   .09687

------------------------------------------------------------------------------
    lnconc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
      time |   .4512015   .0066847    67.498   0.000     .4370307    .4653724
     _cons |  -6.209981   .0769945   -80.655   0.000    -6.373202   -6.04676
------------------------------------------------------------------------------
```

Estimated equation of the straight-line regression:

$$\ln(y) = -6.209981 + 0.4512015 * time$$

```
. predict yhat3
(option xb assumed; fitted values)

. scatter yhat3 lnconc time, connect(l .) symbol(i o)
```
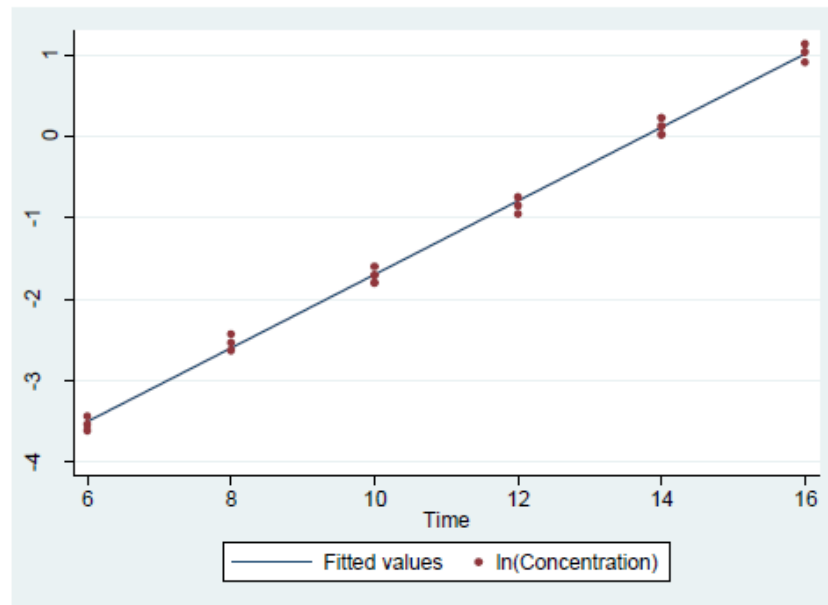


4. Plots of each of these fitted equations on their respective scatter diagrams.

Each scatter plots are located as an image in each of the 3 estimated equations above