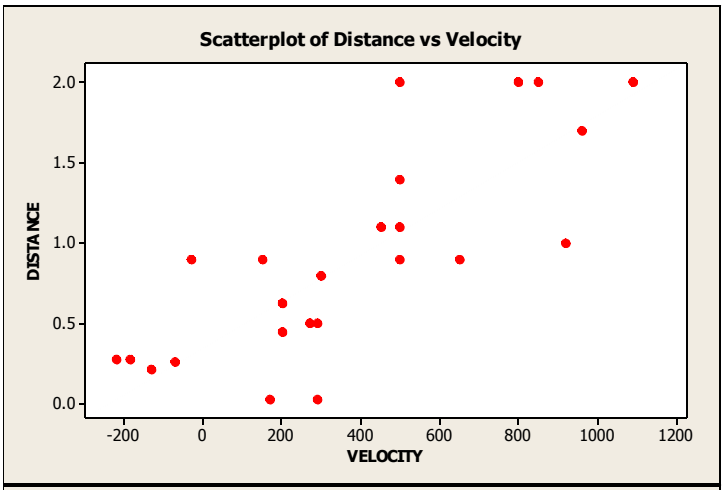
Chapter 7 - Minitab Details

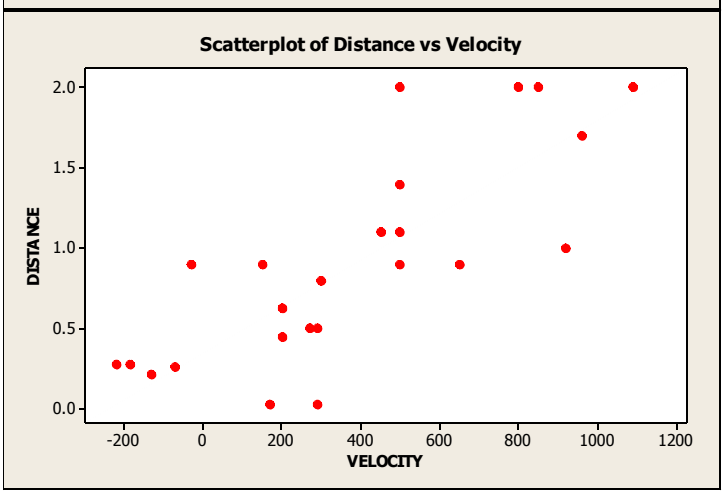
Case 7.1.1. The Big Bang – An Observational Study *R&S p. 177-179*

Step 1: Copy the data into a Minitab Worksheet: use these steps:

File → Open Worksheet →Browse your local directory and upload the csv file case0701.csv. The data will appear as two columns in Minitab with titles VELOCITY(Recession Velocity in km/sec) and DISTANCE (megaparsecs). See the first three columns of ***R&S Display 7.8***.

Step 2: To construct a Scatterplot of DISTANCE versus VELOCITY, go to Graph and select Simple graph option. Select DISTANCE into Y variables and select VELOCITY into X variables and click OK. You can add a title by going into Labels option. This produces this graph.





Step 3: Correlation: Go to Stat →Basic Statistics → Correlation; select variables DISTANCE and VELOCITY and click ok to produce this output.

Correlations: VELOCITY, DISTANCE

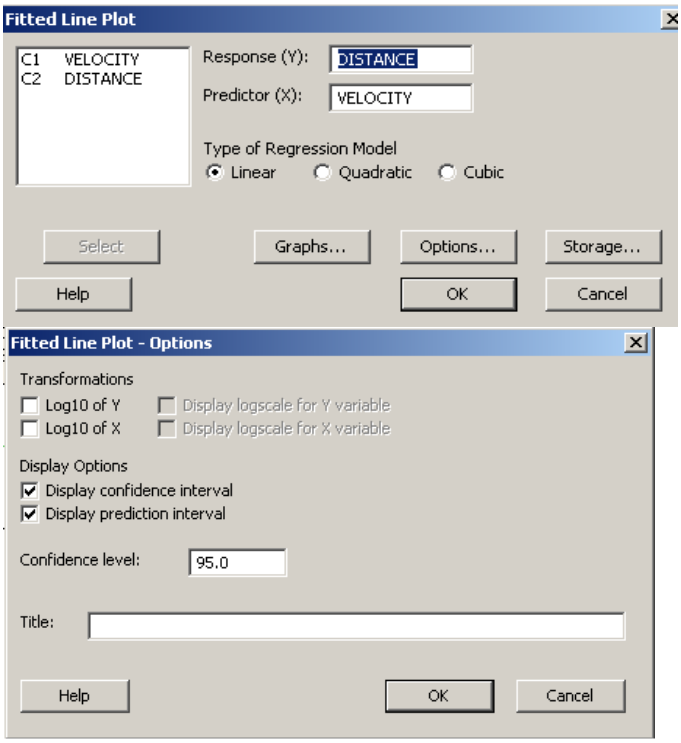
```
Pearson correlation of VELOCITY and DISTANCE = 0.790
P-Value = 0.000
```

Step 4: To produce a fitted simple linear regression, go to Stat → Regression → Regression (first option for a linear regression); select Distance into Response (Y) and Velocity into Predictors (X), and click Ok to produce these results.

Regression Analysis: Distance versus Velocity

```
Analysis of Variance

Source         DF  Adj SS  Adj MS  F-Value  P-Value
Regression      1  5.9709  5.9709    36.29    0.000
  Velocity      1  5.9709  5.9709    36.29    0.000
Error          22  3.6198  0.1645
  Lack-of-Fit  17  2.8031  0.1649     1.01    0.548
  Pure Error    5  0.8166  0.1633
Total          23  9.5907


Model Summary

       S    R-sq  R-sq(adj)  R-sq(pred)
0.405630  62.26%     60.54%      56.01%


Coefficients

Term          Coef   SE Coef  T-Value  P-Value   VIF
Constant     0.399     0.119     3.36    0.003
Velocity  0.001372  0.000228     6.02    0.000  1.00


Regression Equation

Distance = 0.399 + 0.001372 Velocity


Fits and Diagnostics for Unusual Observations

                                Std
Obs  Distance    Fit  Resid  Resid
 21     2.000  1.085  0.915   2.31  R

R  Large residual
```

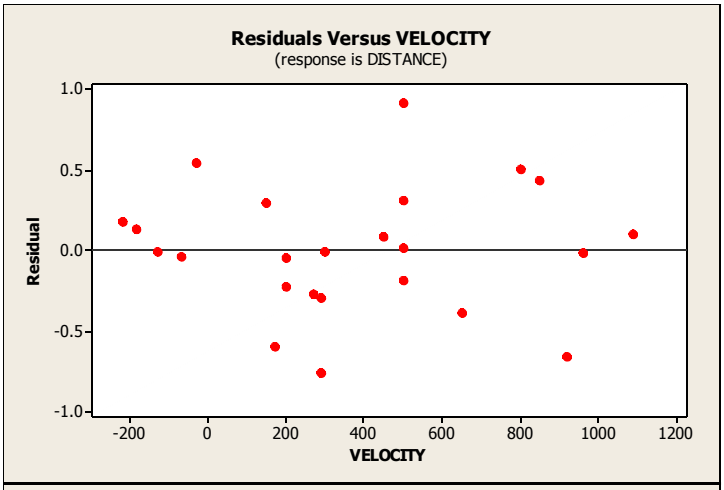
Step 5: Go to Stat → Regression → Fitted Line Plot; select Distance into Response (Y) and Velocity into Predictors (X), select Type of Model to be Linear; → Options and under Display Options, click on Display Confidence Interval and Display Prediction Interval, select Confidence level of 95.0 and click OK, to produce this graph. Also, click on Graphs to produce the residual plots.

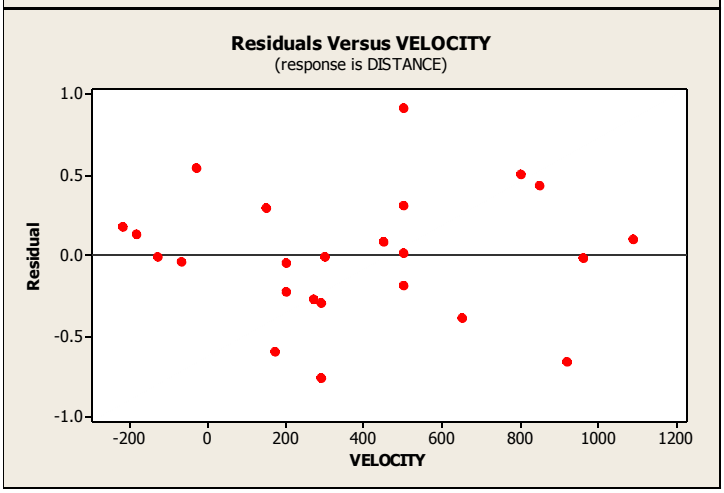


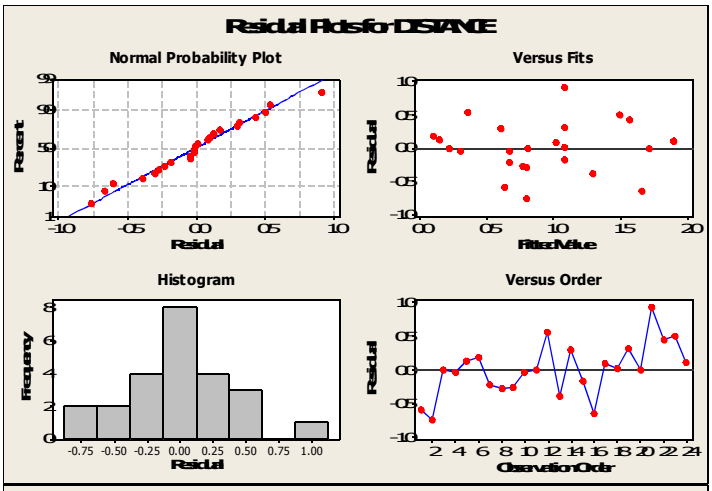
The output for Case 7.1.1 are shown below; we show Standardized Residuals and Fitted Values.

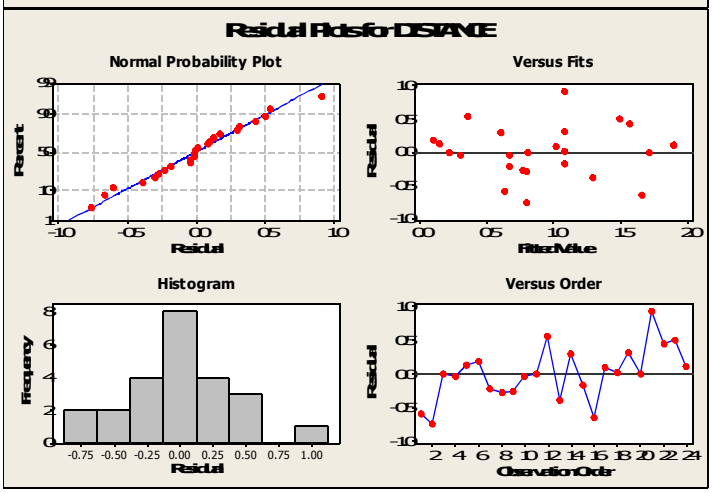
Row	VELOCITY	DISTANCE	SRES1	FITS1
1	170	0.03200	-1.52515	0.63250
2	290	0.03400	-1.92749	0.79725
3	-130	0.21400	-0.01743	0.22062
4	-70	0.26300	-0.10431	0.30299
5	-185	0.27500	0.34587	0.14511
6	-220	0.27500	0.47736	0.09705
7	200	0.45000	-0.56705	0.67369
8	290	0.50000	-0.75067	0.79725
9	270	0.50000	-0.68174	0.76979
10	200	0.63000	-0.11074	0.67369
11	300	0.80000	-0.02772	0.81098
12	-30	0.90000	1.40553	0.35791
13	650	0.90000	-1.00028	1.29151
14	150	0.90000	0.75021	0.60504
15	500	0.90000	-0.46934	1.08557
16	920	1.00000	-1.75923	1.66220
17	450	1.10000	0.20977	1.01692
18	500	1.10000	0.03651	1.08557
19	500	1.40000	0.79527	1.08557
20	960	1.70000	-0.04585	1.71712
21	500	2.00000	2.31279	1.08557
22	850	2.00000	1.13795	1.56609
23	800	2.00000	1.30751	1.49745
24	1090	2.00000	0.28892	1.89560

Residual Plots

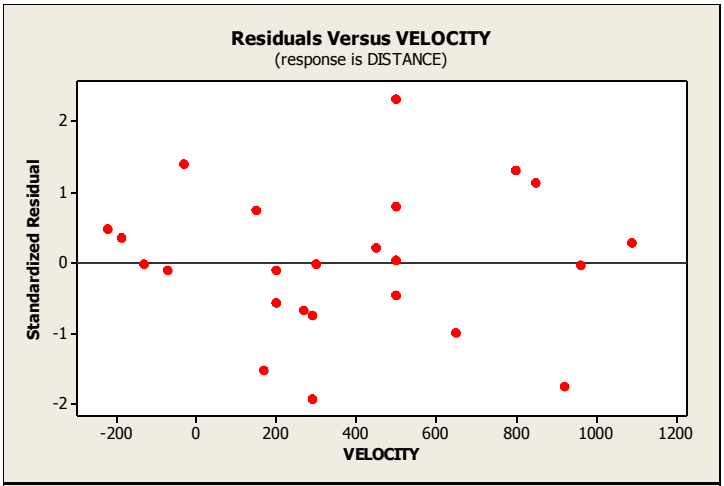


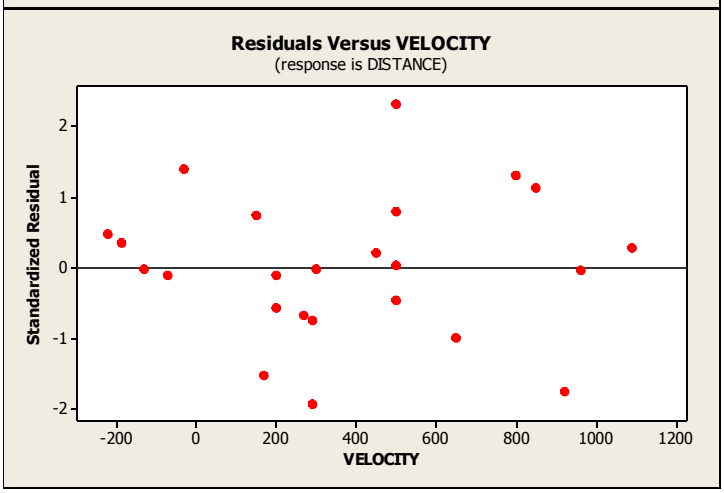


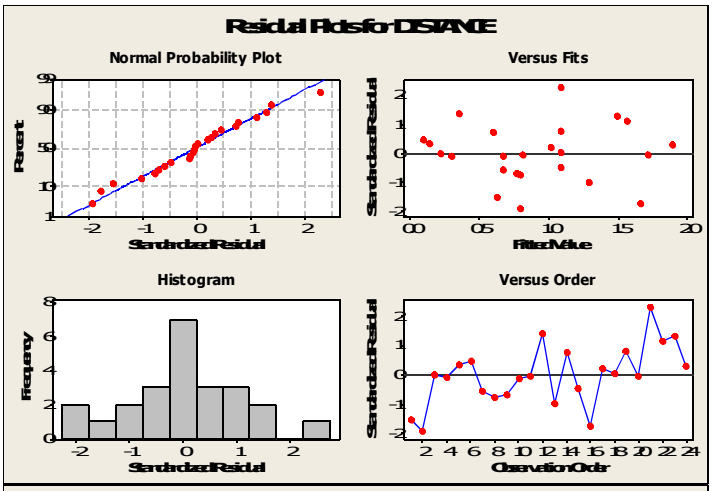

Residual Plots for DISTANCE
Normal Probability Plot
Versus Fits
Histogram
Versus Order
Percent
Residual
Fitted Value
Frequency
Observation Order

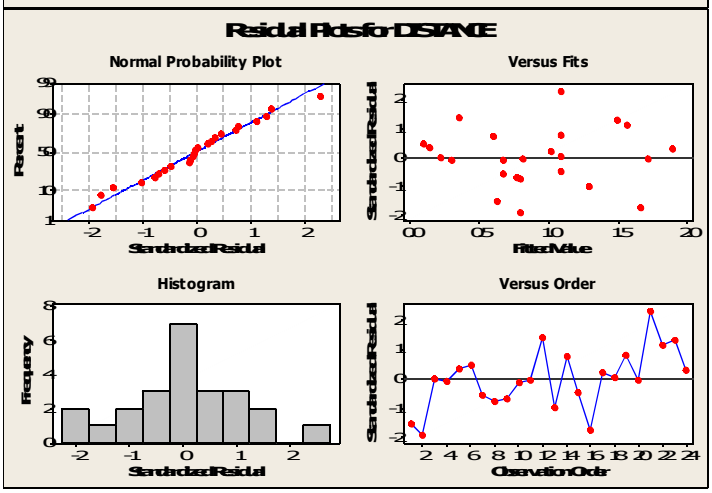

Residual Plots for DISTANCE
Normal Probability Plot
Versus Fits
Histogram
Versus Order
Percent
Residual
Fitted Value
Frequency
Observation Order

Standardized Residual Plots






Residual Plots for DISTANCE
Normal Probability Plot
Versus Fits
Histogram
Versus Order
Percent
Standardized Residual
Fitted Value
Frequency
Observation Order


Residual Plots for DISTANCE
Normal Probability Plot
Versus Fits
Histogram
Versus Order
Percent
Standardized Residual
Fitted Value
Frequency
Observation Order

Case 7.1.2. Meat Processing and pH – A Randomized Experiment *R&S p. 179-180*

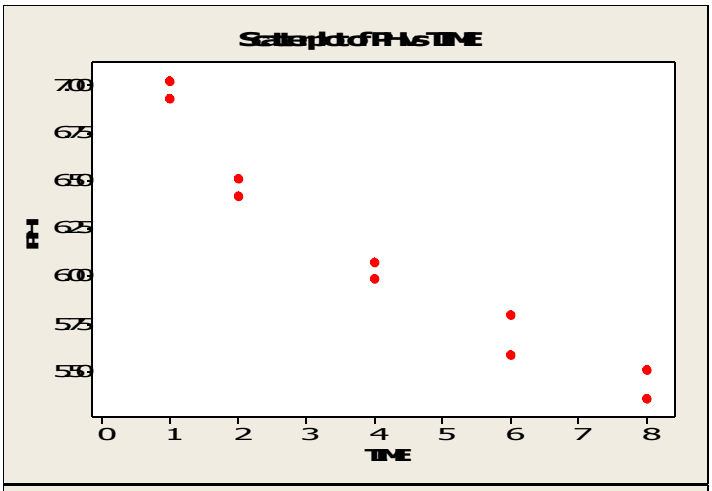
Step 1: Copy the data into a Minitab Worksheet: use these steps:

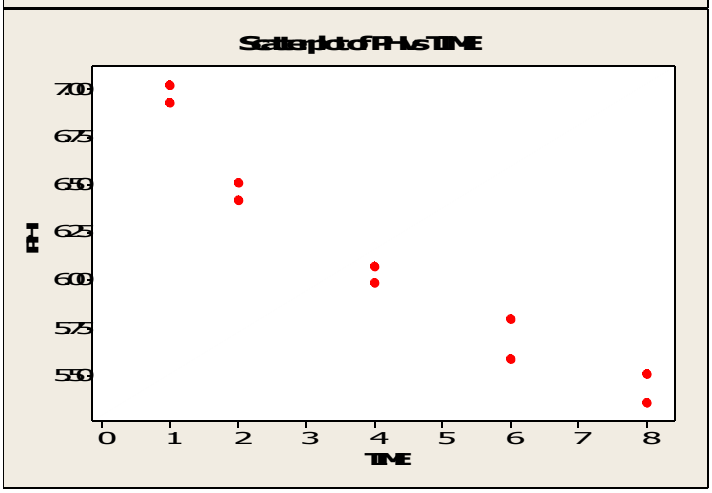
File → Open Worksheet →Browse your local directory and upload the csv file case0702.csv.The data will appear as two columns in Minitab with titles TIME and PH. Next, save LOG TIME as well. See the first four columns of ***R&S Display 7.3***.

Data Display

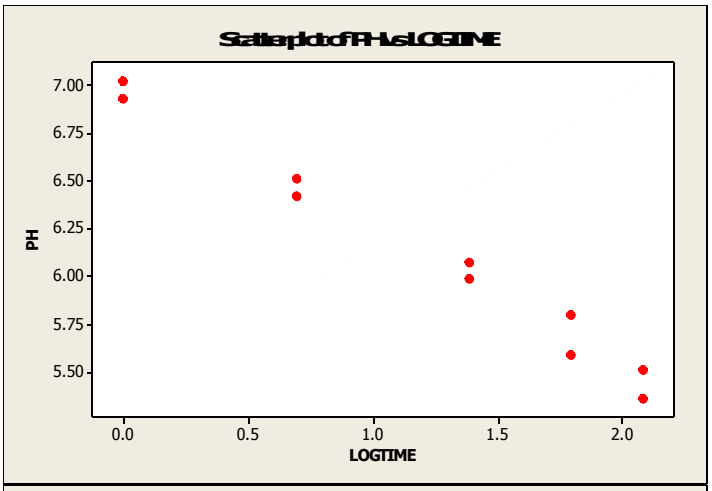
Row	PH	TIME	LOGTIME
1	7.02000	1	0.00000
2	6.93000	1	0.00000
3	6.42000	2	0.69315
4	6.51000	2	0.69315
5	6.07000	4	1.38629
6	5.99000	4	1.38629
7	5.59000	6	1.79176
8	5.80000	6	1.79176
9	5.51000	8	2.07944
10	5.36000	8	2.07944

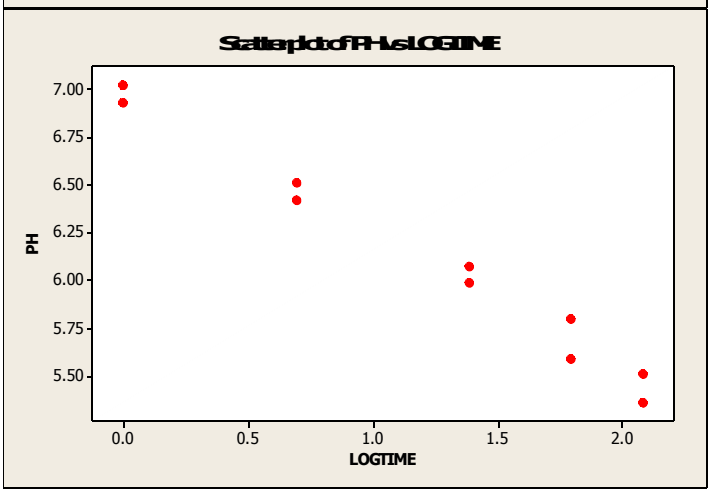
Step 2: To construct a Scatterplot of PH versus TIME, go to Graph and select Simple graph option. Select PH into Y variables and select TIME into X variables and click OK. This produces this graph.





Repeat to get a plot of pH versus LOG TIME, which is saved in C3





Step 3: Correlation: Go to Stat →Basic Statistics → Correlation; select variables PH and LOGTIME and click ok to produce this output.

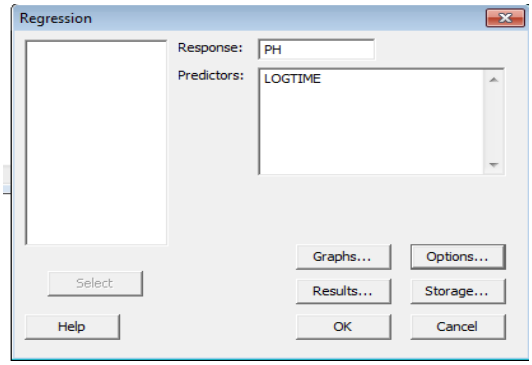
Correlations: PH, LOGTIME

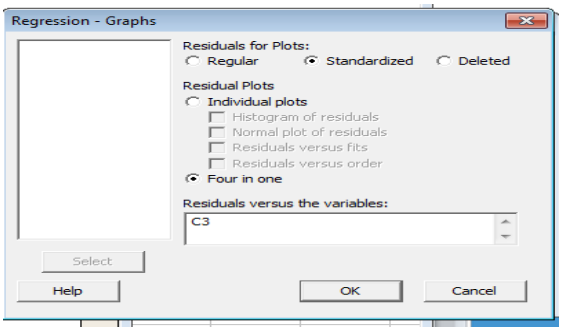
```
Pearson correlation of PH and LOGTIME = -0.991
P-Value = 0.000
```

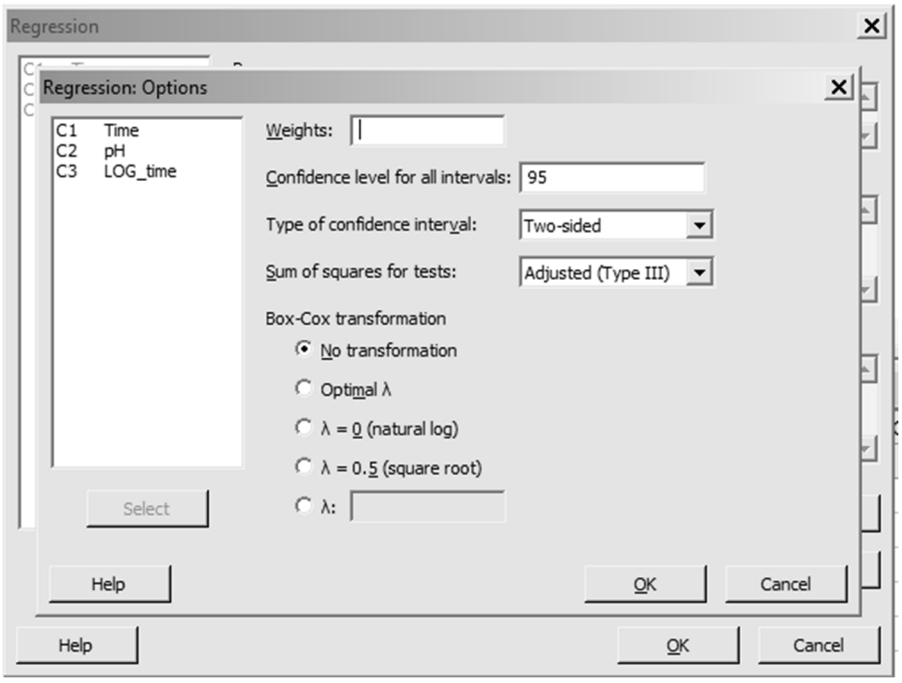
Step 4: To produce a fitted simple linear regression, go to Stat → Regression → Regression (first option for a linear regression); select PH into Response (Y) and LOGTIME into Predictors (X); then click on Options. Enter the X0=1.386 (which is LN(4)) value into the box below and click on Fits, SEs of Fits, Confidence Limits (for Case 1), and Prediction limits (for Case 2), and click OK to produce these results. Helps to obtain prediction results for

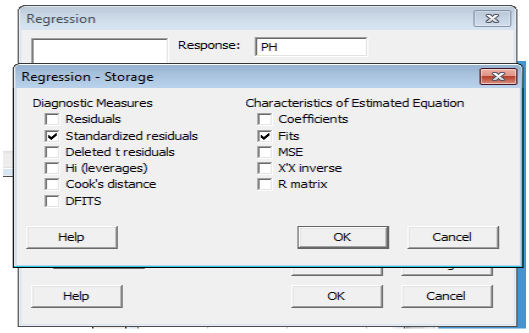
Case 1: Predict the ***mean*** Y response at a given value of $X=X_0$

Case 2: Predict an ***individual*** Y response at a given value of $X=X_0$


Regression
Response:
PH
Predictors:
LOGTIME
Graphs...
Options...
Select
Results...
Storage...
Help
OK
Cancel


Regression - Graphs
Residuals for Plots:
Regular
Standardized
Deleted
Residual Plots
Individual plots
Histogram of residuals
Normal plot of residuals
Residuals versus fits
Residuals versus order
Four in one
Residuals versus the variables:
C3
Select
Help
OK
Cancel


Regression
Regression: Options
C1 Time
C2 pH
C3 LOG_time
Weights:
Confidence level for all intervals: 95
Type of confidence interval: Two-sided
Sum of squares for tests: Adjusted (Type III)
Box-Cox transformation
No transformation
Optimal λ
λ = 0 (natural log)
λ = 0.5 (square root)
λ:
Select
Help
OK
Cancel
Help
OK
Cancel

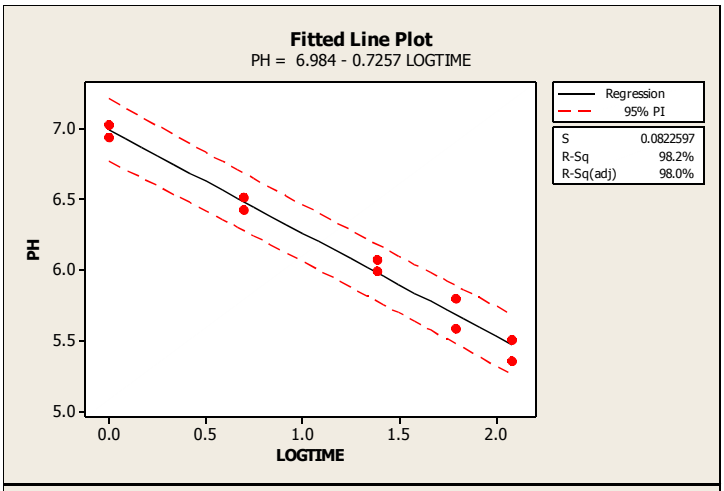


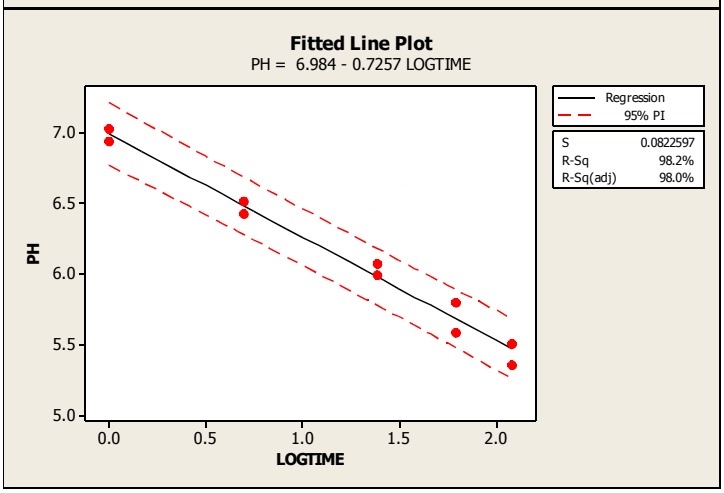
Regression Analysis: pH versus LOG_time

```
Analysis of Variance

Source         DF   Adj SS   Adj MS  F-Value  P-Value
Regression      1  3.00647  3.00647   444.31    0.000
  LOG_time      1  3.00647  3.00647   444.31    0.000
Error           8  0.05413  0.00677
  Lack-of-Fit   3  0.00953  0.00318     0.36    0.788
  Pure Error    5  0.04460  0.00892
   Total        9  3.06060


Model Summary

        S    R-sq  R-sq(adj)  R-sq(pred)
0.0822597  98.23%     98.01%      97.27%


Coefficients

Term         Coef  SE Coef  T-Value  P-Value   VIF
Constant   6.9836   0.0485   143.90    0.000
LOG_time  -0.7257   0.0344   -21.08    0.000  1.00


Regression Equation
```

pH = 6.9836 - 0.7257 LOG_time

Step 5: Go to Stat → Regression → Fitted Line Plot; select PH into Response (Y) and LOGTIME into Predictors (X), select Type of Model to be Linear; → Options and under Display Options, click on Display Prediction Interval, select Confidence level of 95.0 and click OK, to produce this graph. Also, click on Graphs to produce the residual plots (you can do this here, or while fitting the Least Square model)


Fitted Line Plot
PH = 6.984 - 0.7257 LOGTIME
Regression
95% PI
S 0.0822597
R-Sq 98.2%
R-Sq(adj) 98.0%
PH
7.0
6.5
6.0
5.5
5.0
0.0
0.5
1.0
1.5
2.0
LOGTIME


Fitted Line Plot
PH = 6.984 - 0.7257 LOGTIME
Regression
95% PI
S 0.0822597
R-Sq 98.2%
R-Sq(adj) 98.0%
PH
7.0
6.5
6.0
5.5
5.0
0.0
0.5
1.0
1.5
2.0
LOGTIME