

```

1      TITLE1 'Appendix of Simple linear Regression (SLR)';
2      dm'log;clear;output;clear';
3
4      ODS HTML style=minimal body='C:\EXST7005\Spring2010\SAS\Appendix08.html';
NOTE: Writing HTML Body file: C:\EXST7005\Spring2010\SAS\Appendix08.html
5      ODS RTF style=minimal body='C:\EXST7005\Spring2010\SAS\Appendix08.rtf' ;
NOTE: Writing RTF Body file: C:\EXST7005\Spring2010\SAS\Appendix08.rtf
6      ODS PDF style=minimal body='C:\EXST7005\Spring2010\SAS\Appendix08.PDF' ;
NOTE: Writing ODS PDF output to DISK destination
      "C:\EXST7005\Spring2010\SAS\Appendix08.PDF", printer "PDF".
7
8      *****;
9      *** EXST7005 Regression Appendix ***;
10     *** Redfin Pickerel, and other fish, accumulate parasites ***;
11     *** on their fins. These parasites attach and stay with ***;
12     *** the fish throughout its life until the fish is eaten ***;
13     *** and the parasite continues its life cycle. ***;
14     *** - - - - - ***;
15     *** If parasites are accumulated at a constant rate, older ***;
16     *** fish should have more parasites. Test this hypothesis. ***;
17     *** OBJECTIVES: ***;
18     *** 1) Determine if older fish have more parasites. ***;
19     *** 2) Estimate the rate of accumulation of parasites. ***;
20     *** 3) Place a confidence interval on this estimate ***;
21     *** 4) Estimate the intercept with confidence interval. ***;
22     *** 5) How many parasites a 10 year old fish would have. ***;
23     *** 6) Put a confidence interval on the 10 year old fish ***;
24     *** 7) Determine if a linear model is adequate. ***;
25     *** 8) An old published article states that the rate ***;
26     *** should be about 5 per year. Test this. ***;
27     *****;
28
29     options ps=256 ls=99 nocenter nodate nonumber nolabel;
30
31     DATA ONE; INFILE CARDS MISSOVER;
32     TITLE2 'Rate of parasite accumulation in Redfin Pickerel';
33     INPUT AGE PARASITE;
34     LABEL AGE = 'Fish age from scales reading';
35     LABEL PARASITE = 'Pectoral fin parasites / sq cm';
36     CARDS;
NOTE: The data set WORK.ONE has 18 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time          0.00 seconds
      cpu time           0.00 seconds
55     ;
56     PROC PRINT DATA=ONE;
57     TITLE3 'Data Listing for Fish Parasite Regression'; RUN;
NOTE: There were 18 observations read from the data set WORK.ONE.
NOTE: The PROCEDURE PRINT printed page 1.
NOTE: PROCEDURE PRINT used (Total process time):
      real time          0.07 seconds
      cpu time           0.00 seconds

```

**Example of Simple linear Regression (SLR)  
Rate of parasite accumulation in Redfin Pickerel  
Data Listing for Fish Parasite Regression**

| Obs | AGE | PARASITE |    |    |    |
|-----|-----|----------|----|----|----|
| 1   | 1   | 3        | 10 | 6  | 15 |
| 2   | 2   | 7        | 11 | 6  | 16 |
| 3   | 3   | 8        | 12 | 7  | 19 |
| 4   | 3   | 12       | 13 | 7  | 21 |
| 5   | 3   | 10       | 14 | 8  | 18 |
| 6   | 4   | 15       | 15 | 9  | 17 |
| 7   | 4   | 14       | 16 | 9  | 20 |
| 8   | 5   | 16       | 17 | 0  | .  |
| 9   | 6   | 17       | 18 | 10 | .  |

```

59      PROC REG DATA=ONE LINEPRINTER;
60          TITLE3 'Fish Parasite Appendix using REG with CLM';
61          MODEL PARASITE=AGE / clb; *** CLI CLM P R; ID AGE;
62          TEST AGE=5;
63          OUTPUT OUT=NEXT P=Predicted R=Resid STUDENT=student rstudent=rstudent
64              lcl=lcl lclm=lclm ucl=ucl uclm=uclm;
65      RUN;
65      !      OPTIONS PS=45; TITLE4 'Plots of raw data & residuals';
66          PLOT PREDICTED.*AGE='P' PARASITE*AGE='O' / OVERLAY;
67          PLOT RESIDUAL.*AGE='E' ;
68      RUN;
68      !      QUIT;

```

NOTE: The data set WORK.NEXT has 18 observations and 10 variables.  
NOTE: The PROCEDURE REG printed pages 2-5.  
NOTE: PROCEDURE REG used (Total process time):  
real time           0.29 seconds  
cpu time            0.10 seconds

**Appendix of Simple linear Regression (SLR)  
Rate of parasite accumulation in Redfin Pickerel  
Fish Parasite Appendix using REG with CLM**

The REG Procedure

Model: MODEL1

Dependent Variable: PARASITE

|  |    |
|--|----|
| Number of Observations Read                | 18 |
| Number of Observations Used                | 16 |
| Number of Observations with Missing Values | 2  |

| Analysis of Variance | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|----------------------|----|----------------|-------------|---------|--------|
| Source               |    |                |             |         |        |
| Model                | 1  | 301.94955      | 301.94955   | 54.86   | <.0001 |
| Error                | 14 | 77.05045       | 5.50360     |         |        |
| Corrected Total      | 15 | 379.00000      |             |         |        |

|                |          |          |        |
|----------------|----------|----------|--------|
| Root MSE       | 2.34598  | R-Square | 0.7967 |
| Dependent Mean | 14.25000 | Adj R-Sq | 0.7822 |
| Coeff Var      | 16.46299 |          |        |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr >  t | 95% Confidence Limits |
|----------|----|--------------------|----------------|---------|---------|-----------------------|
|----------|----|--------------------|----------------|---------|---------|-----------------------|

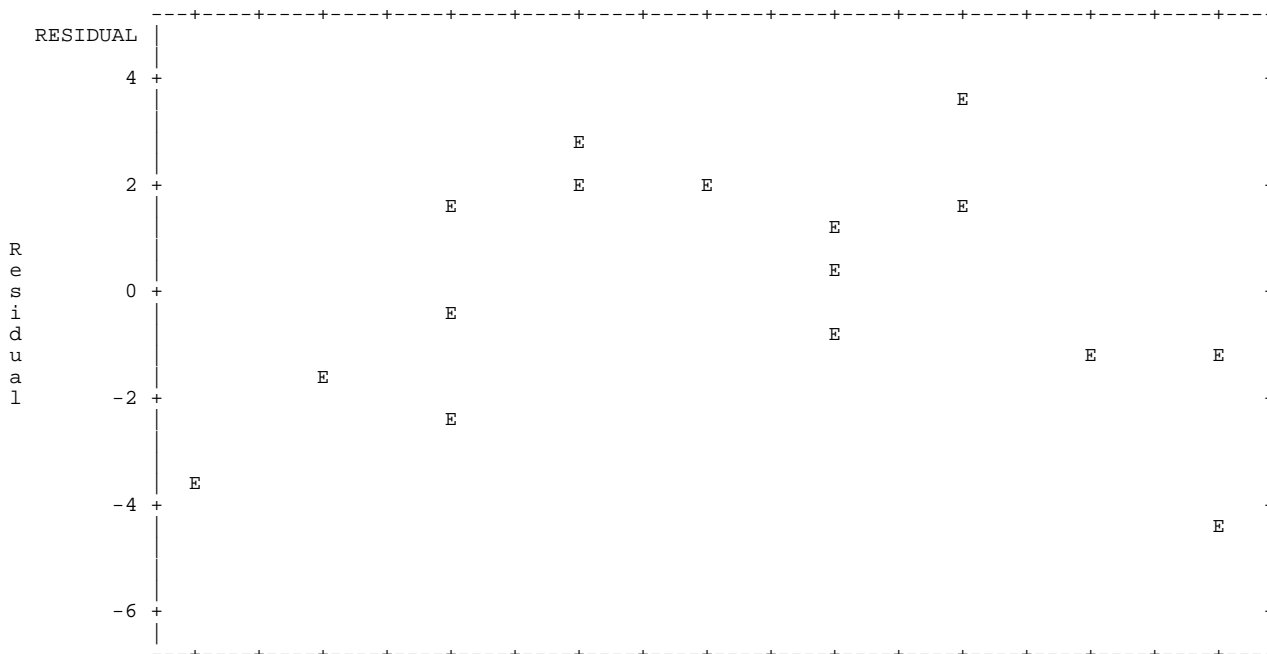
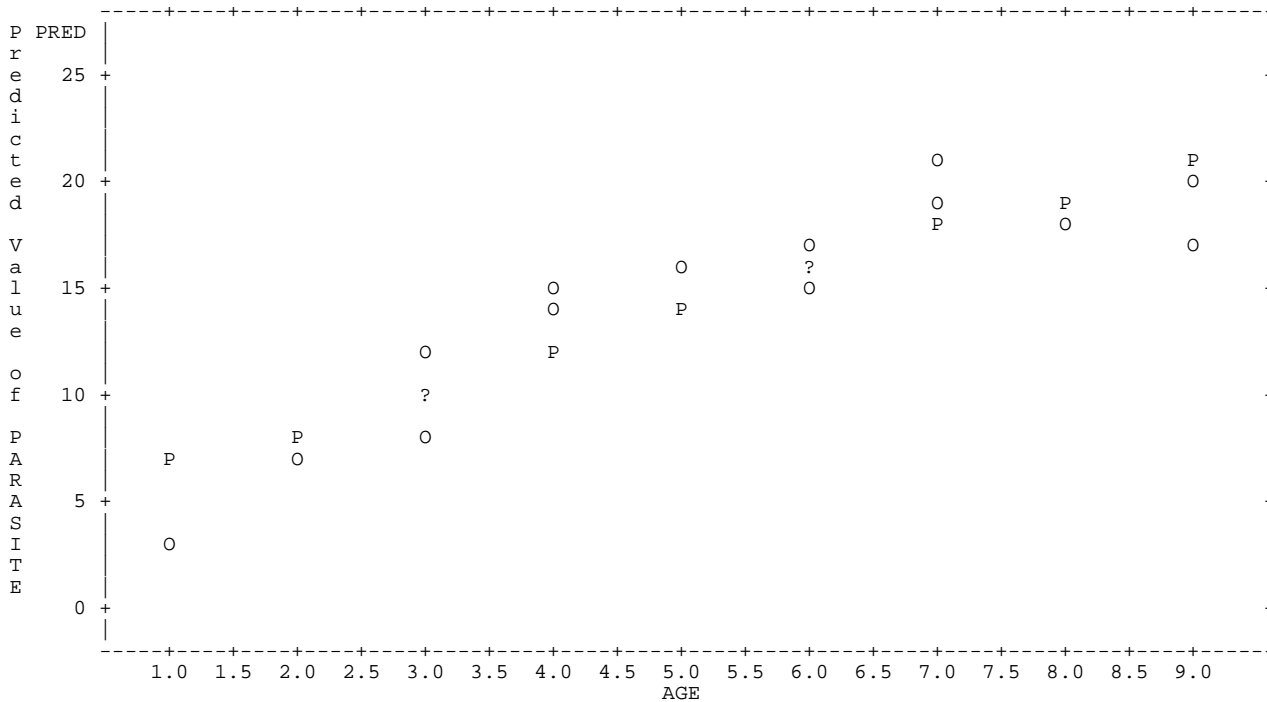
|           |   |         |         |      |        |         |         |
|-----------|---|---------|---------|------|--------|---------|---------|
| Intercept | 1 | 4.77125 | 1.40769 | 3.39 | 0.0044 | 1.75205 | 7.79045 |
| AGE       | 1 | 1.82723 | 0.24669 | 7.41 | <.0001 | 1.29813 | 2.35632 |

Appendix of Simple linear Regression (SLR)  
 Rate of parasite accumulation in Redfin Pickerel  
 Fish Parasite Appendix using REG with CLM

The REG Procedure  
 Model: MODEL1

Test 1 Results for Dependent Variable PARASITE

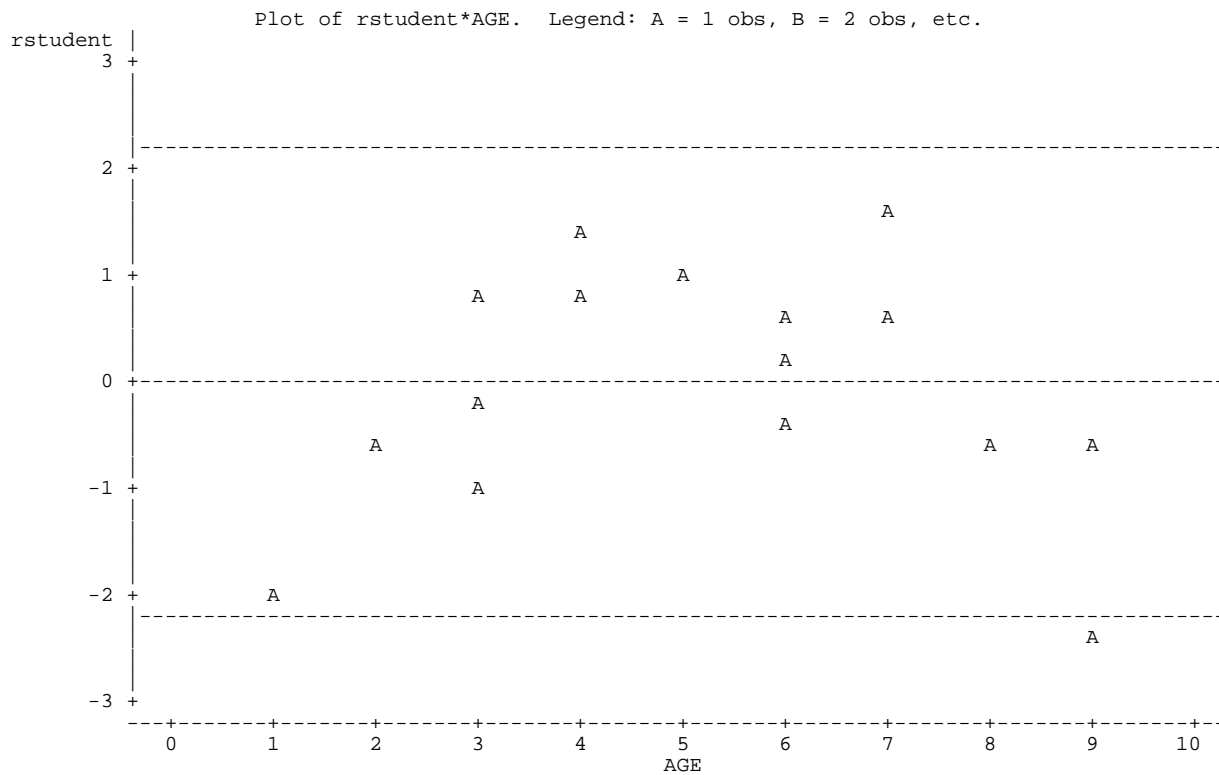
| Source      | DF | Mean Square | F Value | Pr > F |
|-------------|----|-------------|---------|--------|
| Numerator   | 1  | 910.38705   | 165.42  | <.0001 |
| Denominator | 14 | 5.50360     |         |        |



```

1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0
                                AGE
70   proc plot data=next; plot rstudent*age / vref = -2.145 0 +2.145;
71       TITLE4 'Plots of deleted standardized residuals with 95% interval';
72   run;
73
74   OPTIONS PS=256;
NOTE: There were 18 observations read from the data set WORK.NEXT.
NOTE: The PROCEDURE PLOT printed page 6.
NOTE: PROCEDURE PLOT used (Total process time):
      real time          0.09 seconds
      cpu time           0.03 seconds
    
```

Appendix of Simple linear Regression (SLR)  
 Rate of parasite accumulation in Redfin Pickerel  
 Fish Parasite Appendix using REG with CLM  
 Plots of deleted standardized residuals with 95% interval



NOTE: 2 obs had missing values.

```

75   proc print data=next;
76       TITLE4 'Listing of output from PROC REG';
77       var age parasite Predicted Resid student rstudent lcl ucl lclm uclm;
run;
NOTE: There were 18 observations read from the data set WORK.NEXT.
NOTE: The PROCEDURE PRINT printed page 7.
NOTE: PROCEDURE PRINT used (Total process time):
      real time          0.09 seconds
      cpu time           0.03 seconds
    
```

Appendix of Simple linear Regression (SLR)  
 Rate of parasite accumulation in Redfin Pickerel  
 Fish Parasite Appendix using REG with CLM  
 Listing of output from PROC REG

| Obs | AGE | PARASITE | Predicted | Resid    | student  | rstudent | lcl     | ucl     | lclm    | uclm    |
|-----|-----|----------|-----------|----------|----------|----------|---------|---------|---------|---------|
| 1   | 1   | 3        | 6.5985    | -3.59848 | -1.77879 | -1.94833 | 0.9586  | 12.2384 | 4.0507  | 9.1463  |
| 2   | 2   | 7        | 8.4257    | -1.42571 | -0.66902 | -0.65524 | 2.9719  | 13.8795 | 6.3218  | 10.5296 |
| 3   | 3   | 8        | 10.2529   | -2.25294 | -1.02107 | -1.02274 | 4.9389  | 15.5670 | 8.5436  | 11.9623 |
| 4   | 3   | 12       | 10.2529   | 1.74706  | 0.79180  | 0.78068  | 4.9389  | 15.5670 | 8.5436  | 11.9623 |
| 5   | 3   | 10       | 10.2529   | -0.25294 | -0.11464 | -0.11052 | 4.9389  | 15.5670 | 8.5436  | 11.9623 |
| 6   | 4   | 15       | 12.0802   | 2.91983  | 1.29626  | 1.33156  | 6.8558  | 17.3046 | 10.6741 | 13.4863 |
| 7   | 4   | 14       | 12.0802   | 1.91983  | 0.85231  | 0.84348  | 6.8558  | 17.3046 | 10.6741 | 13.4863 |
| 8   | 5   | 16       | 13.9074   | 2.09261  | 0.92144  | 0.91614  | 8.7200  | 19.0948 | 12.6456 | 15.1692 |
| 9   | 6   | 17       | 15.7346   | 1.26538  | 0.55925  | 0.54503  | 10.5304 | 20.9389 | 14.4053 | 17.0640 |
| 10  | 6   | 15       | 15.7346   | -0.73462 | -0.32468 | -0.31405 | 10.5304 | 20.9389 | 14.4053 | 17.0640 |
| 11  | 6   | 16       | 15.7346   | 0.26538  | 0.11729  | 0.11308  | 10.5304 | 20.9389 | 14.4053 | 17.0640 |
| 12  | 7   | 19       | 17.5619   | 1.43815  | 0.64577  | 0.63176  | 12.2875 | 22.8362 | 15.9801 | 19.1436 |
| 13  | 7   | 21       | 17.5619   | 3.43815  | 1.54382  | 1.63316  | 12.2875 | 22.8362 | 15.9801 | 19.1436 |
| 14  | 8   | 18       | 19.3891   | -1.38908 | -0.64222 | -0.62818 | 13.9934 | 24.7848 | 17.4406 | 21.3376 |
| 15  | 9   | 17       | 21.2163   | -4.21631 | -2.03920 | -2.34368 | 15.6514 | 26.7812 | 18.8391 | 23.5936 |
| 16  | 9   | 20       | 21.2163   | -1.21631 | -0.58826 | -0.57400 | 15.6514 | 26.7812 | 18.8391 | 23.5936 |
| 17  | 0   | .        | 4.7713    | .        | .        | .        | -1.0967 | 10.6392 | 1.7520  | 7.7905  |
| 18  | 10  | .        | 23.0435   | .        | .        | .        | 17.2657 | 28.8213 | 20.2035 | 25.8836 |

```
79 PROC UNIVARIATE DATA=NEXT NORMAL PLOT; VAR Resid;
80 TITLE4 'Residual analysis with PROC UNIVARIATE';
81 RUN;
```

NOTE: The PROCEDURE UNIVARIATE printed page 8.  
 NOTE: PROCEDURE UNIVARIATE used (Total process time):  
 real time 0.12 seconds  
 cpu time 0.04 seconds

Appendix of Simple linear Regression (SLR)  
 Rate of parasite accumulation in Redfin Pickerel  
 Fish Parasite Appendix using REG with CLM  
 Residual analysis with PROC UNIVARIATE

The UNIVARIATE Procedure  
 Variable: Resid

Moments

|                 |            |                  |            |
|-----------------|------------|------------------|------------|
| N               | 16         | Sum Weights      | 16         |
| Mean            | 0          | Sum Observations | 0          |
| Std Deviation   | 2.26642816 | Variance         | 5.13669661 |
| Skewness        | -0.3183952 | Kurtosis         | -0.7591259 |
| Uncorrected SS  | 77.0504492 | Corrected SS     | 77.0504492 |
| Coeff Variation | .          | Std Error Mean   | 0.56660704 |

Basic Statistical Measures

| Location |          | Variability         |         |
|----------|----------|---------------------|---------|
| Mean     | 0.000000 | Std Deviation       | 2.26643 |
| Median   | 0.006220 | Variance            | 5.13670 |
| Mode     | .        | Range               | 7.65446 |
|          |          | Interquartile Range | 3.24084 |

Tests for Location: Mu0=0

| Test        | -Statistic- | -----p Value----- |
|-------------|-------------|-------------------|
| Student's t | t 0         | Pr >  t  1.0000   |
| Sign        | M 0         | Pr >=  M  1.0000  |
| Signed Rank | S 4         | Pr >=  S  0.8603  |

Tests for Normality

| Test               | --Statistic--- | -----p Value----- |
|--------------------|----------------|-------------------|
| Shapiro-Wilk       | W 0.961962     | Pr < W 0.6975     |
| Kolmogorov-Smirnov | D 0.149185     | Pr > D >0.1500    |
| Cramer-von Mises   | W-Sq 0.038869  | Pr > W-Sq >0.2500 |
| Anderson-Darling   | A-Sq 0.248615  | Pr > A-Sq >0.2500 |

Quantiles (Definition 5)

| Quantile   | Estimate    |
|------------|-------------|
| 100% Max   | 3.43814789  |
| 99%        | 3.43814789  |
| 95%        | 3.43814789  |
| 90%        | 2.91983414  |
| 75% Q3     | 1.83344851  |
| 50% Median | 0.00621977  |
| 25% Q1     | -1.40739461 |
| 10%        | -3.59847961 |
| 5%         | -4.21630961 |
| 1%         | -4.21630961 |
| 0% Min     | -4.21630961 |

Extreme Observations

| -----Lowest----- |     | -----Highest----- |     |
|------------------|-----|-------------------|-----|
| Value            | Obs | Value             | Obs |
| -4.21631         | 15  | 1.74706           | 4   |
| -3.59848         | 1   | 1.91983           | 7   |
| -2.25294         | 3   | 2.09261           | 8   |
| -1.42571         | 2   | 2.91983           | 6   |
| -1.38908         | 14  | 3.43815           | 13  |

Missing Values

| -----Percent Of----- |                           |
|----------------------|---------------------------|
| Missing Value        | Count All Obs Missing Obs |
| .                    | 2 11.11 100.00            |

