data meat;
  input time ph;
  logtime = log(time);
cards;
1 7.02
1 6.93
2 6.42
2 6.51
4 6.07
4 5.99
6 5.59
6 5.80
8 5.51
8 5.36
;
run;

/* Here's how to get and plot predictions for other X values */
/* read those time values into a new data set, meat2 */
/* the ph = . means that the pH value is missing in these observations */
/* ph needs to be missing here because we don't want these X values */
/* to contribute to the regression fit */
/* the variable names here (time and logtime) must be the same */
/* as in the original data set */
data meat2;
  input time ;
  logtime = log(time);
  ph = .;
datalines;
1
1.5
2
2.5
3
3.5
4
4.5
5
;
/* Here is a second way to create a regularly spaced list */
/* of time values */
data meat2;
  do time = 1 to 10 by 0.25;
    logtime = log(time);
    ph = .;
    output;
  end;
run;

/* the do statement is the beginning of a loop */
/* the end; statement is the end */
/* the 1 is the first value of time, the 10 is the last, */
/* the 0.5 is the increment */
/* the output statement says 'write an observation to the data set */
proc print data=meat2(obs=10);
  title 'prediction points';
run;
/* create a new data set called ALL by concatenating the two data sets */

data all;
  set meat meat2;
run;

/* SAS will read all the observations from meat, then all the observations from meat2 */
/* if you wanted to do other things to the data set, those commands would go between */
/* the set and the run; (just like when we log transformed means */

/* Fit regression to all values - those with missing Y's are ignored */

/* The output statement produces a data set */
/* called resid, with lots of additional variables. */
/* you can use or omit keywords and variable names as needed */
proc glm data = all;
  model ph=logtime;
  output out=resid
    residual=resid /* residuals (or r=) */
    predicted=yh /* predicted values (or p=) */
    stdp = seline /* s.e. of predicted mean (the line) */
    stdi = seobs /* s.e. on predicted observation */
    lclm = lciobs /* lower 95% c.i. limit for pred. mean */
    uclm = uciobs /* upper 95% c.i. limit for pred. mean */
    lcl = lciobs /* upper 95% c.i. limit for pred. obs */
    ucl = uciobs /* upper 95% c.i. limit for pred. obs */
  ;
  title 'regression on all data';

/* all the information from the output command goes into a data set */
/* print it to see the stuff you need */
proc print data = resid (obs=15);
  title 'output from all obs';
  run;

/* can also plot stuff in many ways */
/* the following plots three pieces of information together: */
/* the predicted line and the upper and lower bounds of the */
/* 95% c.i. for predicted observations */
/* the where statement restricts the plot to those obs. where ph is missing */
/* this restricts the plot to those points in the meat2 data set */
/* (i.e. the additional regularly spaced time points) */
/* more importantly, the points in meat2 are in order of increasing X values */
/* the points in all are not.*/
proc sgplot;
  where ph = .;
  band x=time lower=lciobs upper=uciobs /transparency = 0.5
    legendlabel='Prediction Interval';
  /* draws a light band. larger values of transparency = lighter line */
  band x=time lower=lcioline upper=ucioline /legendlabel='Confidence Interval for line'
    fillattrs=(color=Red);
  /* draws a darker band in Red */
  series x=time y=yhat /legendlabel='Predicted line'
    lineattrs=(color=Green thickness=3); /* draws a line */
  /* draws a line */
run;
/* the order of drawing commands matters! */
/* Each command draws over what is already on the plot */
/* Start with the 'biggest' thing, then the next biggest, etc. */
/* If you draw the line first, it is drawn over */

/* compare above plot to this one */
proc sgplot;
  where ph = .;
  series x=time y=yhat / legendlabel='Predicted line'; /* draws a line */
  band x=time lower=lciline upper=uciline / legendlabel='Confidence Interval for line';
  band x=time lower=lciobs upper=uciobs / transparency = 0.5 legendlabel='Prediction Interval';
run;

/* and here's another way to plot just one set of limits */
/* refline adds a reference line at the specified value(s) */
/* /axis=y draws a horizontal line, /axis=x draws a vertical line */
proc sgplot;
  where ph = .;
  series x=time y=yhat / legendlabel='Predicted line';
  scatter x=time y=yhat / yerrorlower = lciline yerrorupper = uciline;
  refline 6 /axis=y;
run;
### output from all obs

<table>
<thead>
<tr>
<th>Obs</th>
<th>time</th>
<th>logtime</th>
<th>ph</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.0000</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>1.25</td>
<td>0.22314</td>
<td>.</td>
</tr>
<tr>
<td>3</td>
<td>1.50</td>
<td>0.40547</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>1.75</td>
<td>0.55962</td>
<td>.</td>
</tr>
<tr>
<td>5</td>
<td>2.00</td>
<td>0.69315</td>
<td>.</td>
</tr>
<tr>
<td>6</td>
<td>2.25</td>
<td>0.81093</td>
<td>.</td>
</tr>
<tr>
<td>7</td>
<td>2.50</td>
<td>0.91629</td>
<td>.</td>
</tr>
<tr>
<td>8</td>
<td>2.75</td>
<td>1.01160</td>
<td>.</td>
</tr>
<tr>
<td>9</td>
<td>3.00</td>
<td>1.17865</td>
<td>.</td>
</tr>
<tr>
<td>10</td>
<td>3.25</td>
<td>1.17865</td>
<td>.</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>36</td>
<td>9.75</td>
<td>2.27727</td>
<td>.</td>
</tr>
<tr>
<td>37</td>
<td>10.00</td>
<td>2.30259</td>
<td>.</td>
</tr>
</tbody>
</table>

### PROC GLM output – most deleted, since saw last week

<table>
<thead>
<tr>
<th>Number of Observations Read</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Used</td>
<td>10</td>
</tr>
</tbody>
</table>

Notice only 10 obs (from meat) used to fit the regression. The other 37 are the prediction points from meat2.