STAT 500

Model Diagnostics for
the SLR analysis of Forbes data

```
/*  Compute least squares estimates of regression coefficients
    and output the residuals and predicted values to a file
    that can be used used by other SAS procedures.

     The following code includes the plots=diagnostics(unpack)
     option that outputs diagnostic plots as individual graphs,
     If you want to put the diagnostic plots into a panel of
     plots, remove this option.  */

proc reg data=set1 plots=diagnostics(unpack);
  model y = x /  p clm cli;
  output out=set2 p=yhat r=residual student=stdres stdp=stdp;
  run;
```

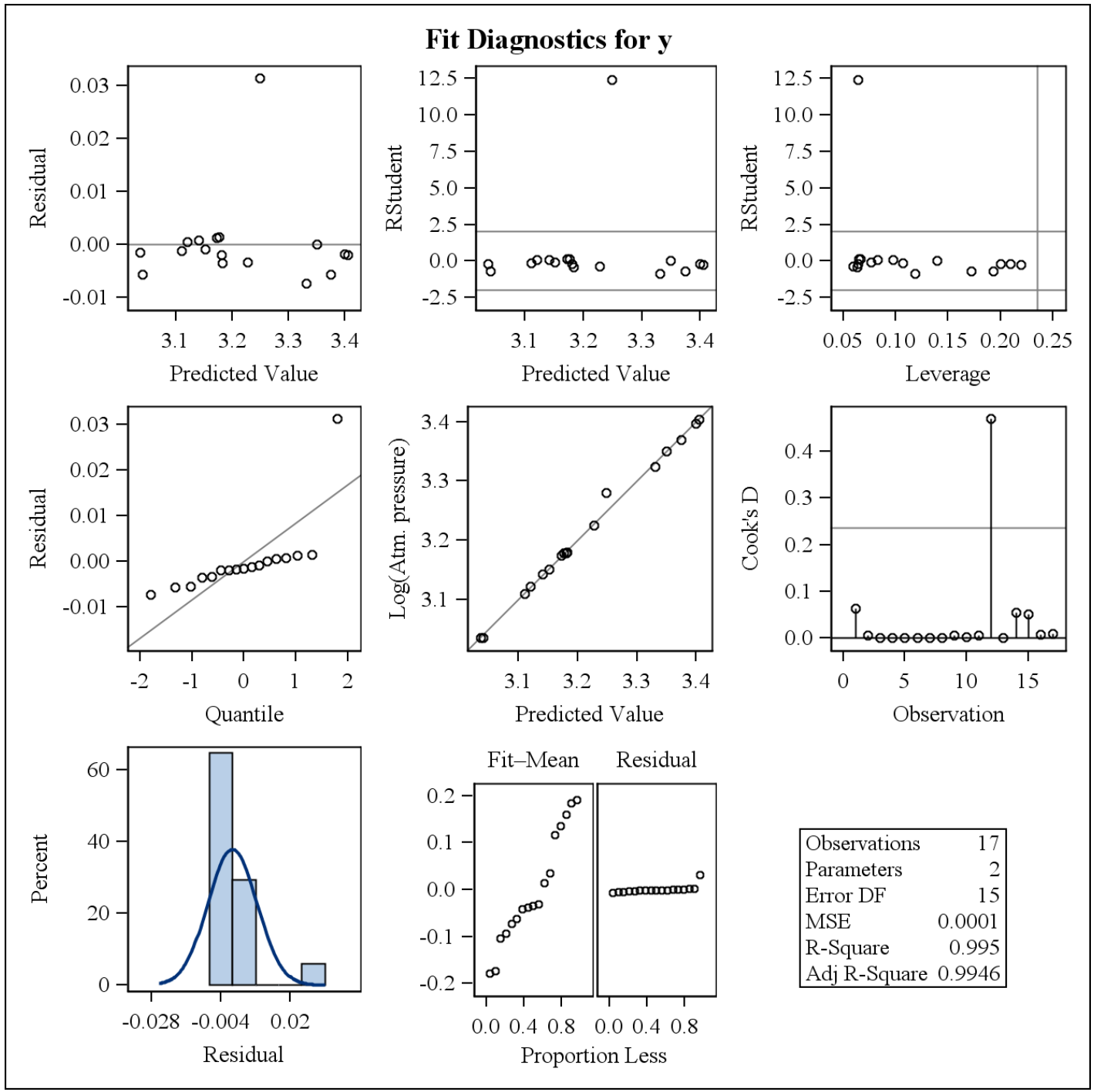
The REG Procedure
Dependent Variable: y Log(Atm. pressure)

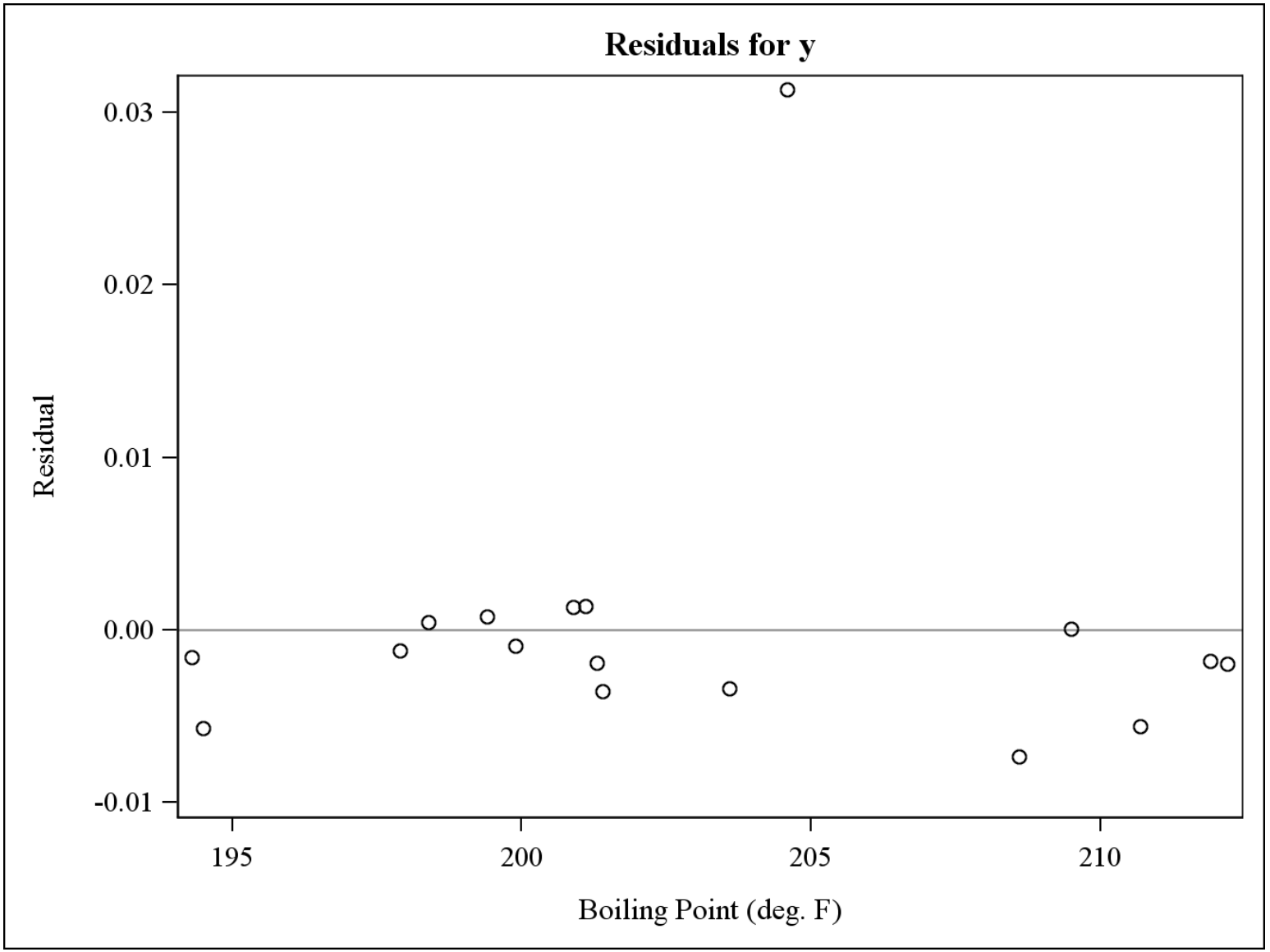
Number of Observations Read	17
Number of Observations Used	17

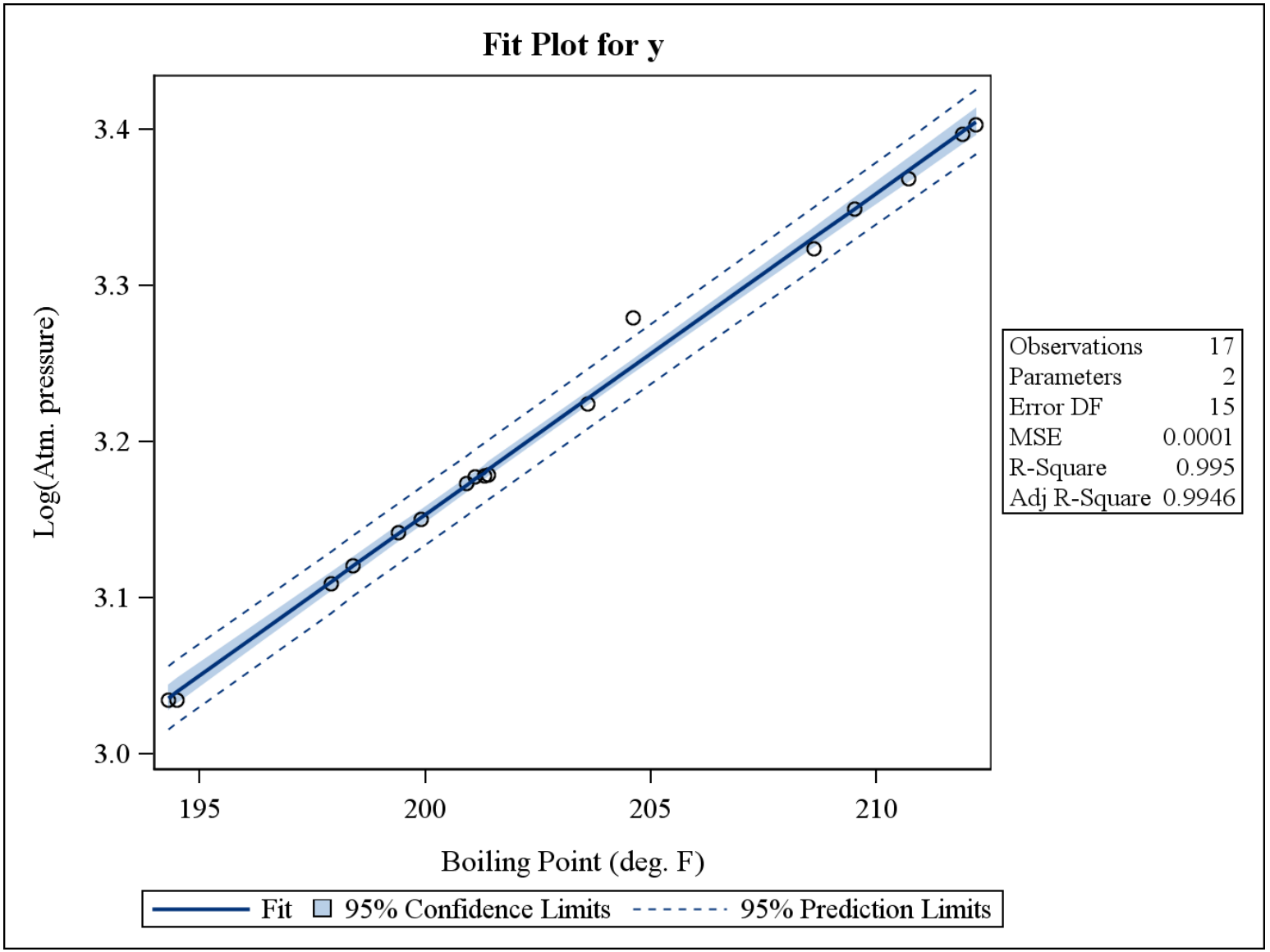
Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	0.22573	0.22573	2961.55	<.0001
Error	15	0.00114	0.00007622		
Corrected Total	16	0.22688			

Root MSE	0.00873	**R-Square**	0.9950
Dependent Mean	3.21450	**Adj R-Sq**	0.9946
Coeff Var	0.27160		

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	Intercept	1	-0.97087	0.07694	-12.62	<.0001
x	Boiling Point (deg. F)	1	0.02062	0.00037895	54.42	<.0001

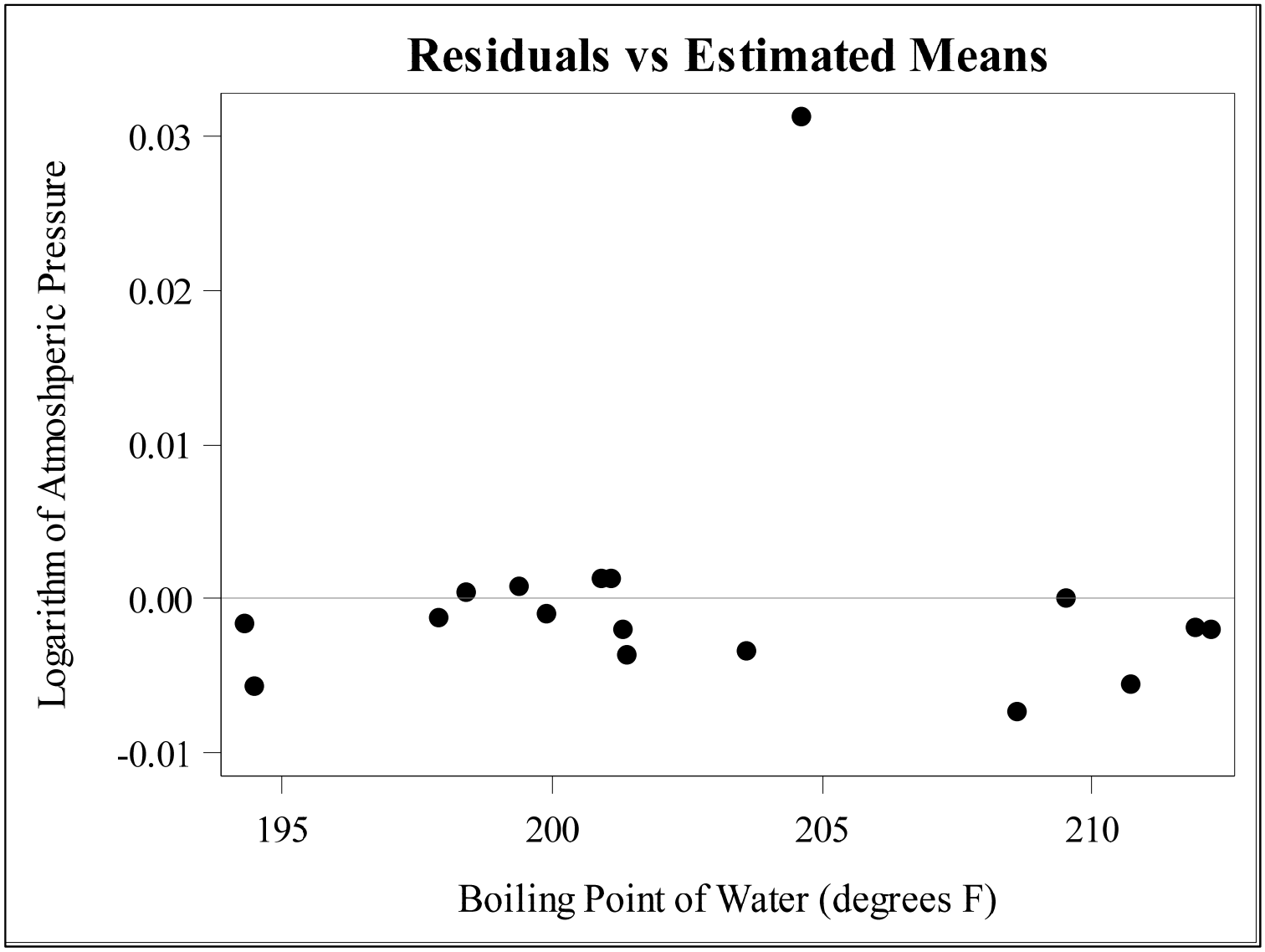

Fit Diagnostics for y
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log(Atm. pressure)
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 17
Parameters 2
Error DF 15
MSE 0.0001
R-Square 0.995
Adj R-Square 0.9946


Residuals for y
0.03
0.02
0.01
0.00
-0.01
Residual
195
200
205
210
Boiling Point (deg. F)


Fit Plot for y
Log(Atm. pressure)
3.4
3.3
3.2
3.1
3.0
195
200
205
210
Boiling Point (deg. F)
Observations 17
Parameters 2
Error DF 15
MSE 0.0001
R-Square 0.995
Adj R-Square 0.9946
Fit
95% Confidence Limits
95% Prediction Limits

```
/*  Plot residuals against the explanatory variable to check
    for homogeneity of variance and patterns in the residuals
    that would suggest that the model is inadequate  */

title h=2 'Residuals vs Estimated Means';
proc sgplot data=set2;
  scatter x=x y=residual /
     markerattrs=(size=12 symbol=CircleFilled color=black);
  refline 0 / axis=y;
  xaxis label="Boiling Point of Water (degrees F)"
          labelattrs=(size=14) valueattrs=(size=13);
  yaxis label="Logarithm of Atmospheric Pressure"
          labelattrs=(size=14) valueattrs=(size=13);
  run;
```


Residuals vs Estimated Means
Logarithm of Atmoshperic Pressure
0.03
0.02
0.01
0.00
-0.01
195
200
205
210
Boiling Point of Water (degrees F)

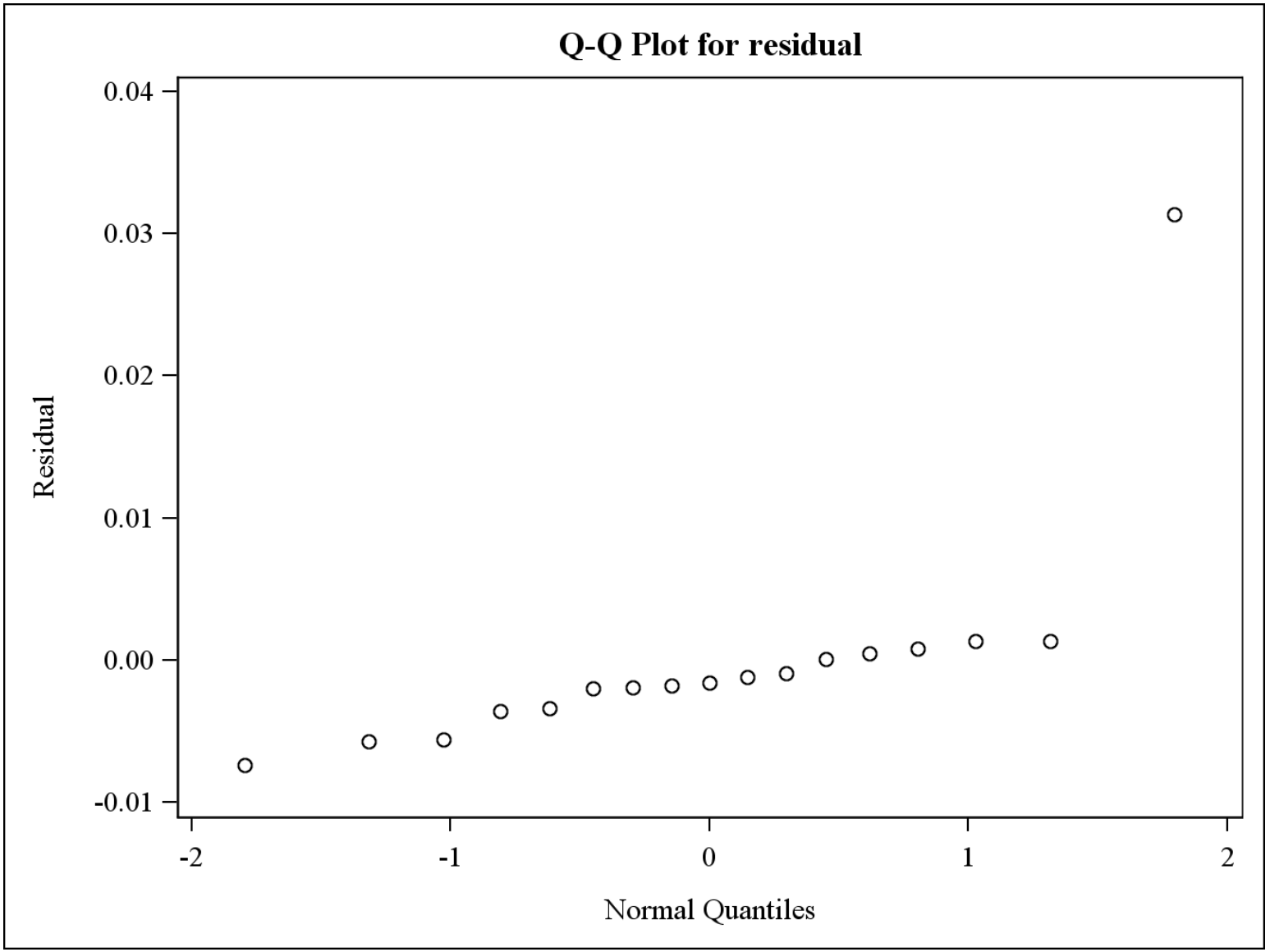
```
/*  Compute normal probability plot for residuals */

title " ";
proc univariate data=set2 normal;
  var residual;
  qqplot;
  run;
```

The UNIVARIATE Procedure
Variable: residual

Moments			
N	17	**Sum Weights**	17
Mean	0	**Sum Observations**	0
Std Deviation	0.00845323	**Variance**	0.00007146
Skewness	3.49882176	**Kurtosis**	13.5869581
Uncorrected SS	0.00114331	**Corrected SS**	0.00114331
Coeff Variation	.	**Std Error Mean**	0.00205021

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.548298	**Pr < W**	<0.0001
Kolmogorov-Smirnov	**D**	0.377899	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.496863	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2.736351	**Pr > A-Sq**	<0.0050


Q-Q Plot for residual
Residual
0.04
0.03
0.02
0.01
0.00
-0.01
-2
-1
0
1
2
Normal Quantiles