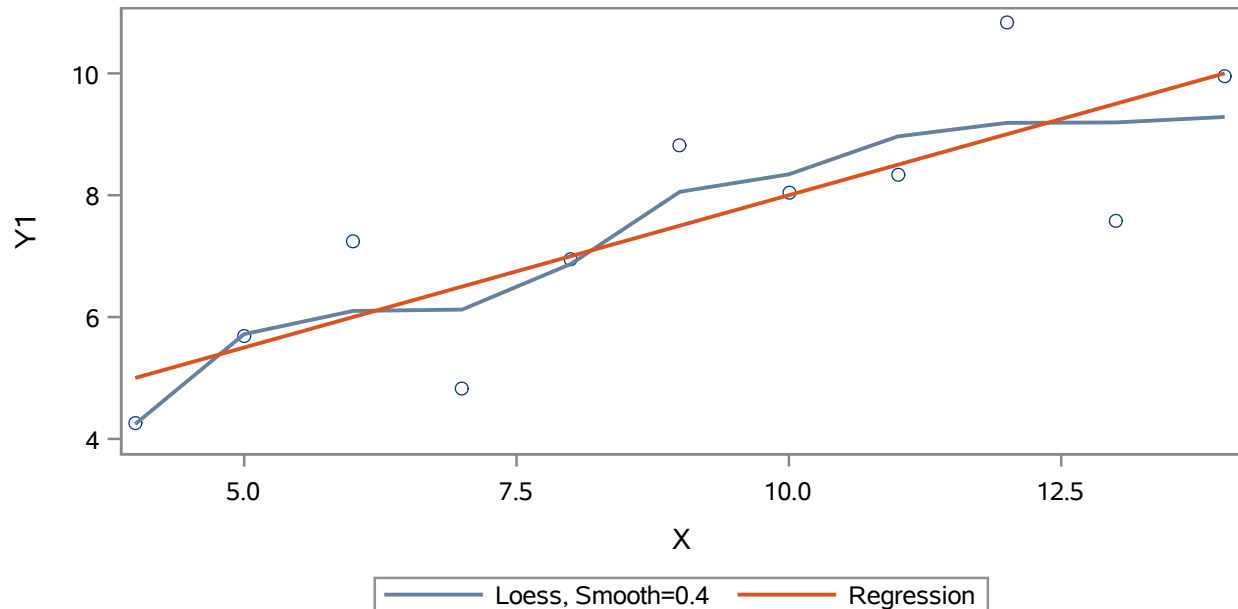


**An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.**

The Anscombe Data

| Obs | X | X4 | Y1 | Y2 | Y3 | Y4 |
|-----|----|----|-------|------|-------|-------|
| 1 | 10 | 8 | 8.04 | 9.14 | 7.46 | 6.58 |
| 2 | 8 | 8 | 6.95 | 8.14 | 6.77 | 5.76 |
| 3 | 13 | 8 | 7.58 | 8.74 | 12.74 | 7.71 |
| 4 | 9 | 8 | 8.81 | 8.77 | 7.11 | 8.83 |
| 5 | 11 | 8 | 8.33 | 9.26 | 7.81 | 8.47 |
| 6 | 14 | 8 | 9.96 | 8.10 | 8.84 | 7.04 |
| 7 | 6 | 8 | 7.24 | 6.13 | 6.08 | 5.25 |
| 8 | 4 | 19 | 4.26 | 3.10 | 5.39 | 12.50 |
| 9 | 12 | 8 | 10.84 | 9.13 | 8.15 | 5.56 |
| 10 | 7 | 8 | 4.82 | 7.26 | 6.42 | 7.91 |
| 11 | 5 | 8 | 5.68 | 4.74 | 5.73 | 6.89 |

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.
Plots of the Anscombe Quartet



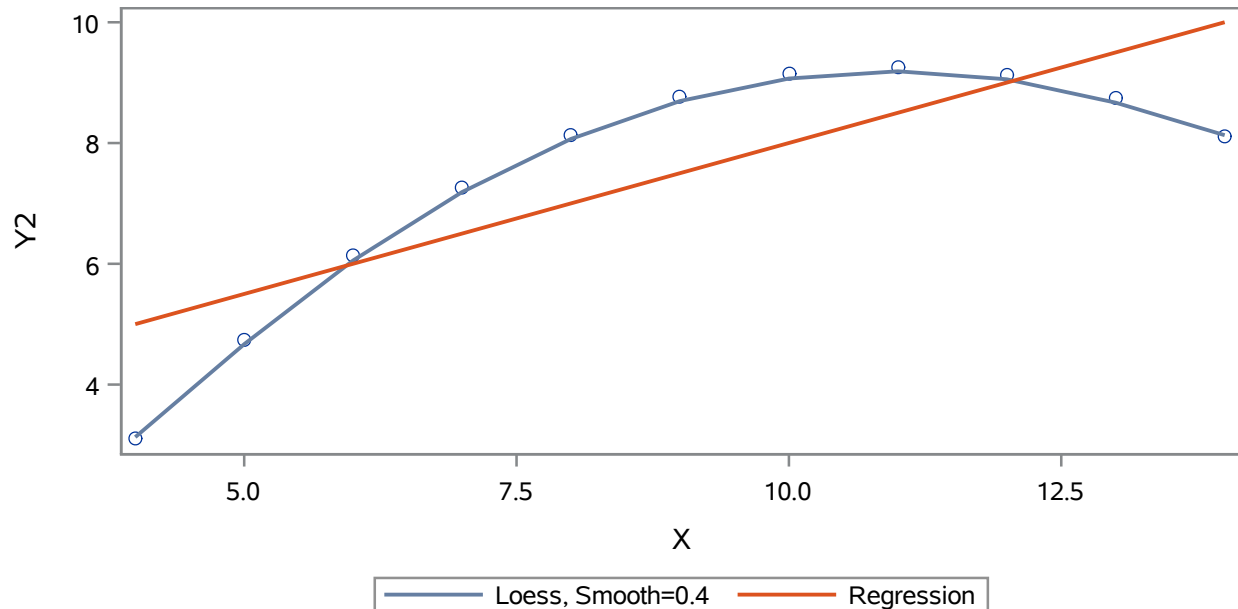
Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

FOR LOESS see https://en.wikipedia.org/wiki/Local_regression

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.
Plots of the Anscombe Quartet



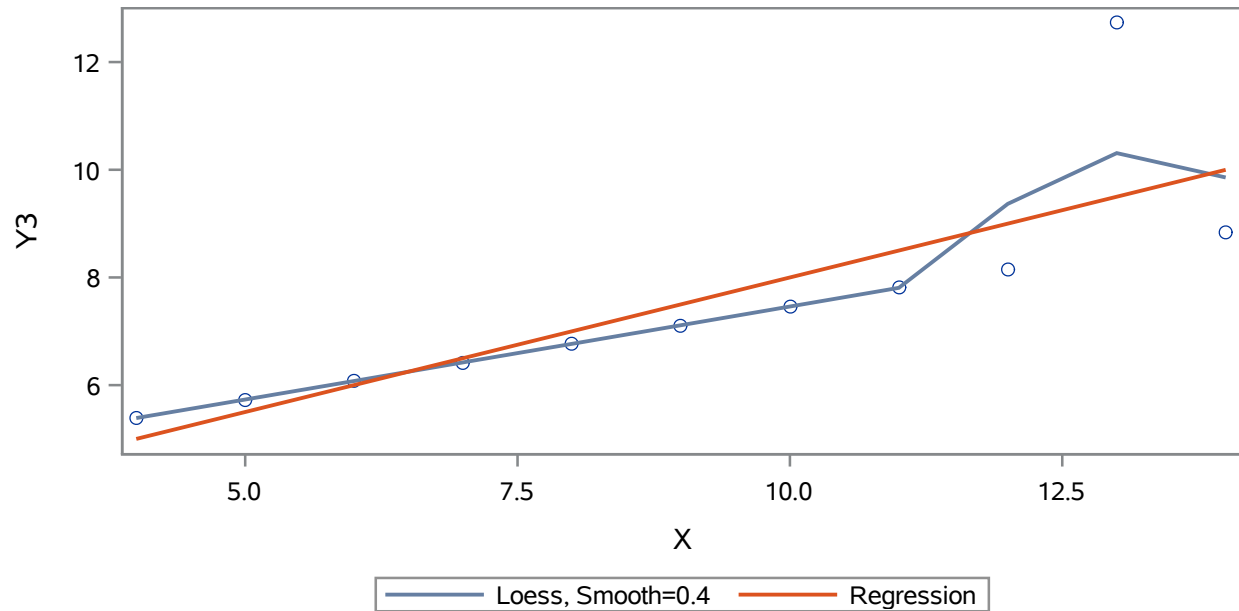
Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

FOR LOESS see https://en.wikipedia.org/wiki/Local_regression

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.
Plots of the Anscombe Quartet



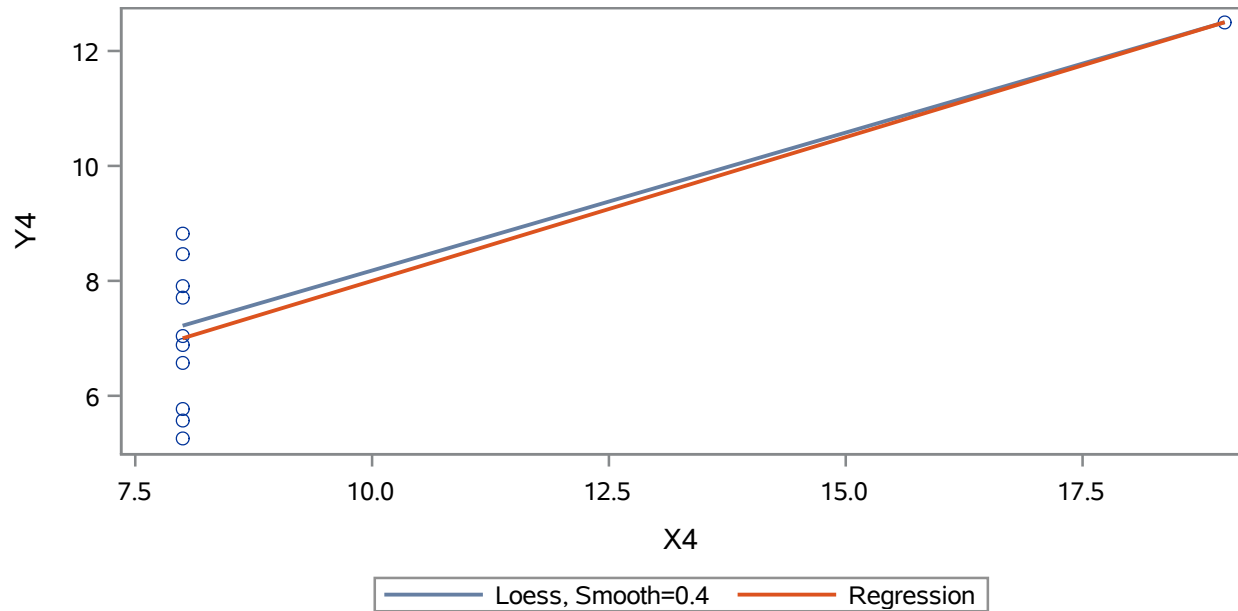
Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

For LOESS see https://en.wikipedia.org/wiki/Local_regression

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.
Plots of the Anscombe Quartet



Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

FOR LOESS see https://en.wikipedia.org/wiki/Local_regression

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 6

The REG Procedure

| | |
|-----------------------------|----|
| Number of Observations Read | 11 |
| Number of Observations Used | 11 |

| Descriptive Statistics | | | | | |
|------------------------|----------|---------|-------------------|----------|-----------------------|
| Variable | Sum | Mean | Uncorrected SS | Variance | Standard Deviation |
| Intercept | 11.00000 | 1.00000 | 11.00000 | 0 | 0 |
| X | 99.00000 | 9.00000 | 1001.00000 | 11.00000 | 3.31662 |
| Y1 | 82.51000 | 7.50091 | 660.17270 | 4.12727 | 2.03157 |
| Y2 | 82.51000 | 7.50091 | 660.17630 | 4.12763 | 2.03166 |
| Y3 | 82.50000 | 7.50000 | 659.97620 | 4.12262 | 2.03042 |
| X4 | 99.00000 | 9.00000 | 1001.00000 | 11.00000 | 3.31662 |
| Y4 | 82.50000 | 7.50000 | 659.95580 | 4.12058 | 2.02992 |

Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see

https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 7

The REG Procedure
Model: MODEL1
Dependent Variable: Y1

| | |
|-----------------------------|----|
| Number of Observations Read | 11 |
| Number of Observations Used | 11 |

| Analysis of Variance | | | | | |
|----------------------|----|----------------|-------------|---------|--------|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 27.51000 | 27.51000 | 17.99 | 0.0022 |
| Error | 9 | 13.76269 | 1.52919 | | |
| Corrected Total | 10 | 41.27269 | | | |

| | | | |
|----------------|----------|----------|--------|
| Root MSE | 1.23660 | R-Square | 0.6665 |
| Dependent Mean | 7.50091 | Adj R-Sq | 0.6295 |
| Coeff Var | 16.48605 | | |

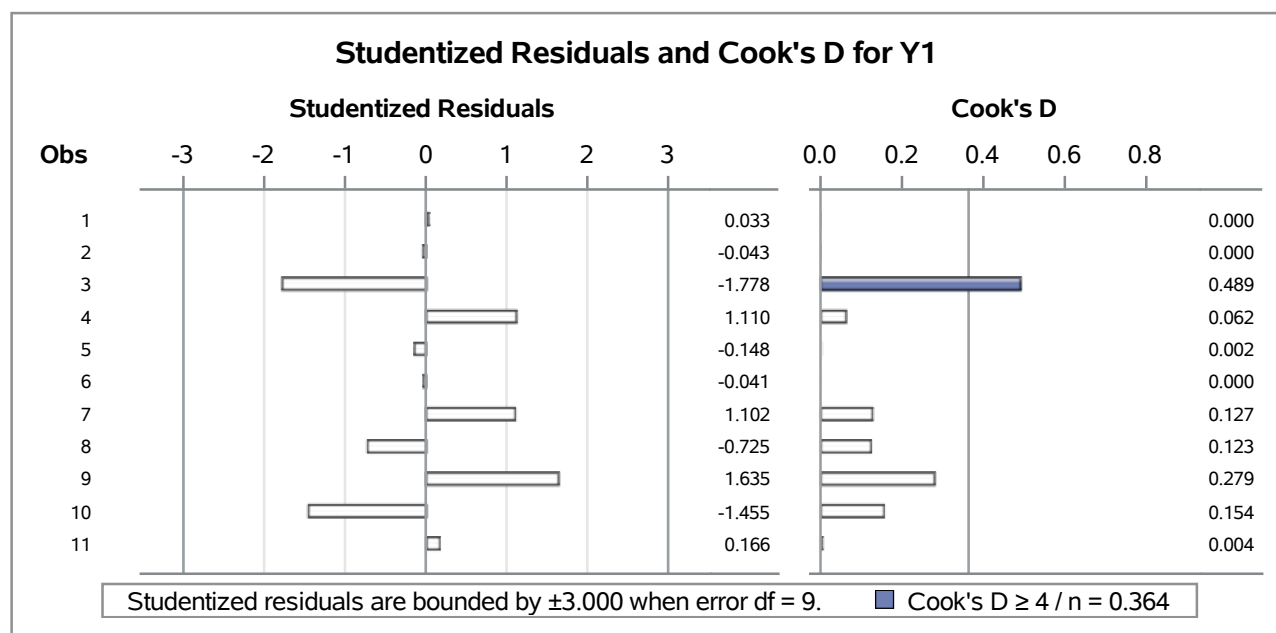
| Parameter Estimates | | | | | |
|---------------------|----|--------------------|----------------|---------|---------|
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > t |
| Intercept | 1 | 3.00009 | 1.12475 | 2.67 | 0.0257 |
| X | 1 | 0.50009 | 0.11791 | 4.24 | 0.0022 |

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 8

The REG Procedure
Model: MODEL1
Dependent Variable: Y1

| Output Statistics | | | | | | | |
|-------------------|----------|----------|------------|-----------|---------|-----------|---------|
| Obs | Residual | RStudent | Hat Diag H | Cov Ratio | DFFITS | DFBETAS | |
| | | | | | | Intercept | X |
| 1 | 0.0390 | 0.0313 | 0.1000 | 1.4059 | 0.0104 | 0.0003 | 0.0032 |
| 2 | -0.0508 | -0.0408 | 0.1000 | 1.4057 | -0.0136 | -0.0082 | 0.0041 |
| 3 | -1.9213 | -2.0811 | 0.2364 | 0.6976 | -1.1578 | 0.6189 | -0.9083 |
| 4 | 1.3091 | 1.1268 | 0.0909 | 1.0369 | 0.3563 | 0.1181 | 0.0000 |
| 5 | -0.1711 | -0.1398 | 0.1273 | 1.4431 | -0.0534 | 0.0120 | -0.0285 |
| 6 | -0.0414 | -0.0382 | 0.3182 | 1.8556 | -0.0261 | 0.0162 | -0.0221 |
| 7 | 1.2394 | 1.1170 | 0.1727 | 1.1449 | 0.5104 | 0.4541 | -0.3513 |
| 8 | -0.7405 | -0.7046 | 0.3182 | 1.6457 | -0.4813 | -0.4691 | 0.4068 |
| 9 | 1.8388 | 1.8383 | 0.1727 | 0.7561 | 0.8400 | -0.3434 | 0.5781 |
| 10 | -1.6807 | -1.5685 | 0.1273 | 0.8483 | -0.5990 | -0.4699 | 0.3202 |
| 11 | 0.1795 | 0.1568 | 0.2364 | 1.6472 | 0.0872 | 0.0825 | -0.0684 |



| | |
|-------------------------------|----------|
| Sum of Residuals | 0 |
| Sum of Squared Residuals | 13.76269 |
| Predicted Residual SS (PRESS) | 20.57925 |

Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see

https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 9

The REG Procedure
Model: MODEL2
Dependent Variable: Y2

| | |
|-----------------------------|----|
| Number of Observations Read | 11 |
| Number of Observations Used | 11 |

| Analysis of Variance | | | | | |
|----------------------|----|----------------|-------------|---------|--------|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 27.50000 | 27.50000 | 17.97 | 0.0022 |
| Error | 9 | 13.77629 | 1.53070 | | |
| Corrected Total | 10 | 41.27629 | | | |

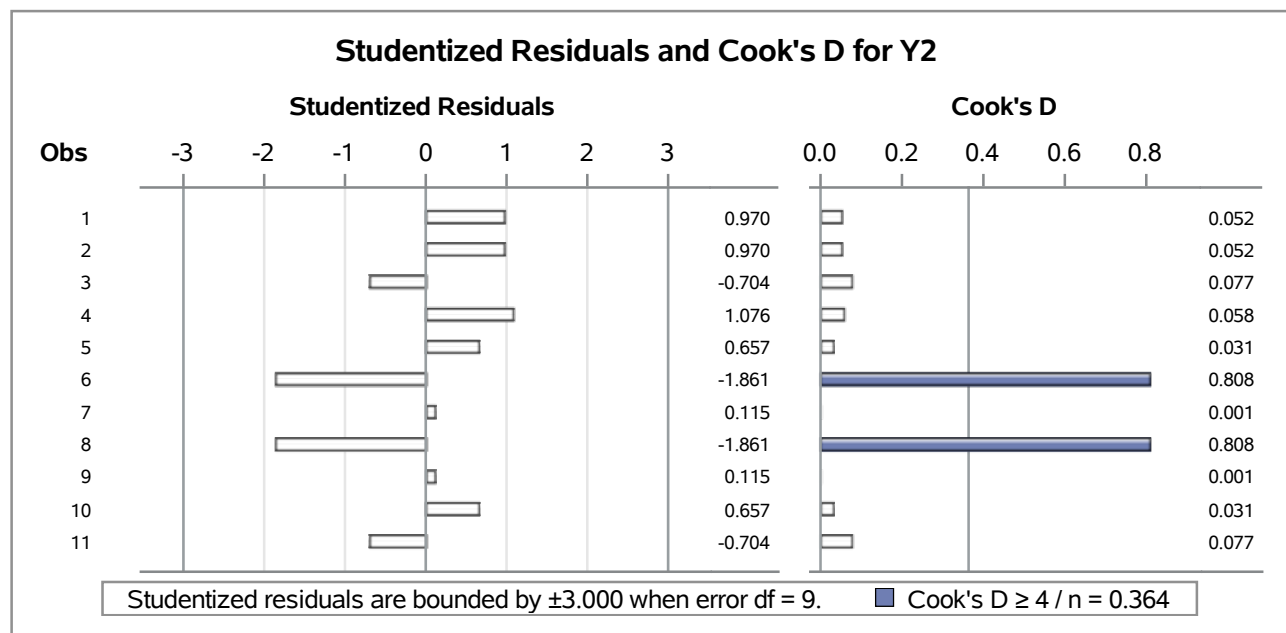
| | | | |
|----------------|----------|----------|--------|
| Root MSE | 1.23721 | R-Square | 0.6662 |
| Dependent Mean | 7.50091 | Adj R-Sq | 0.6292 |
| Coeff Var | 16.49419 | | |

| Parameter Estimates | | | | | |
|---------------------|----|--------------------|----------------|---------|---------|
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > t |
| Intercept | 1 | 3.00091 | 1.12530 | 2.67 | 0.0258 |
| X | 1 | 0.50000 | 0.11796 | 4.24 | 0.0022 |

An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

The REG Procedure
Model: MODEL2
Dependent Variable: Y2

| Output Statistics | | | | | | | |
|-------------------|----------|----------|------------|-----------|---------|-----------|---------|
| Obs | Residual | RStudent | Hat Diag H | Cov Ratio | DFFITS | DFBETAS | |
| | | | | | | Intercept | X |
| 1 | 1.1391 | 0.9670 | 0.1000 | 1.1273 | 0.3223 | 0.0102 | 0.0972 |
| 2 | 1.1391 | 0.9670 | 0.1000 | 1.1273 | 0.3223 | 0.1936 | -0.0972 |
| 3 | -0.7609 | -0.6826 | 0.2364 | 1.4800 | -0.3798 | 0.2030 | -0.2979 |
| 4 | 1.2691 | 1.0866 | 0.0909 | 1.0571 | 0.3436 | 0.1139 | 0.0000 |
| 5 | 0.7591 | 0.6346 | 0.1273 | 1.3145 | 0.2423 | -0.0543 | 0.1295 |
| 6 | -1.9009 | -2.2365 | 0.3182 | 0.7028 | -1.5278 | 0.9475 | -1.2912 |
| 7 | 0.1291 | 0.1082 | 0.1727 | 1.5254 | 0.0495 | 0.0440 | -0.0340 |
| 8 | -1.9009 | -2.2365 | 0.3182 | 0.7028 | -1.5278 | -1.4889 | 1.2912 |
| 9 | 0.1291 | 0.1082 | 0.1727 | 1.5254 | 0.0495 | -0.0202 | 0.0340 |
| 10 | 0.7591 | 0.6346 | 0.1273 | 1.3145 | 0.2423 | 0.1901 | -0.1295 |
| 11 | -0.7609 | -0.6826 | 0.2364 | 1.4800 | -0.3798 | -0.3591 | 0.2979 |



| | |
|-------------------------------|----------|
| Sum of Residuals | 0 |
| Sum of Squared Residuals | 13.77629 |
| Predicted Residual SS (PRESS) | 24.24599 |

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 11

The REG Procedure
Model: MODEL3
Dependent Variable: Y3

| | |
|-----------------------------|----|
| Number of Observations Read | 11 |
| Number of Observations Used | 11 |

| Analysis of Variance | | | | | |
|----------------------|----|----------------|-------------|---------|--------|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 27.47001 | 27.47001 | 17.97 | 0.0022 |
| Error | 9 | 13.75619 | 1.52847 | | |
| Corrected Total | 10 | 41.22620 | | | |

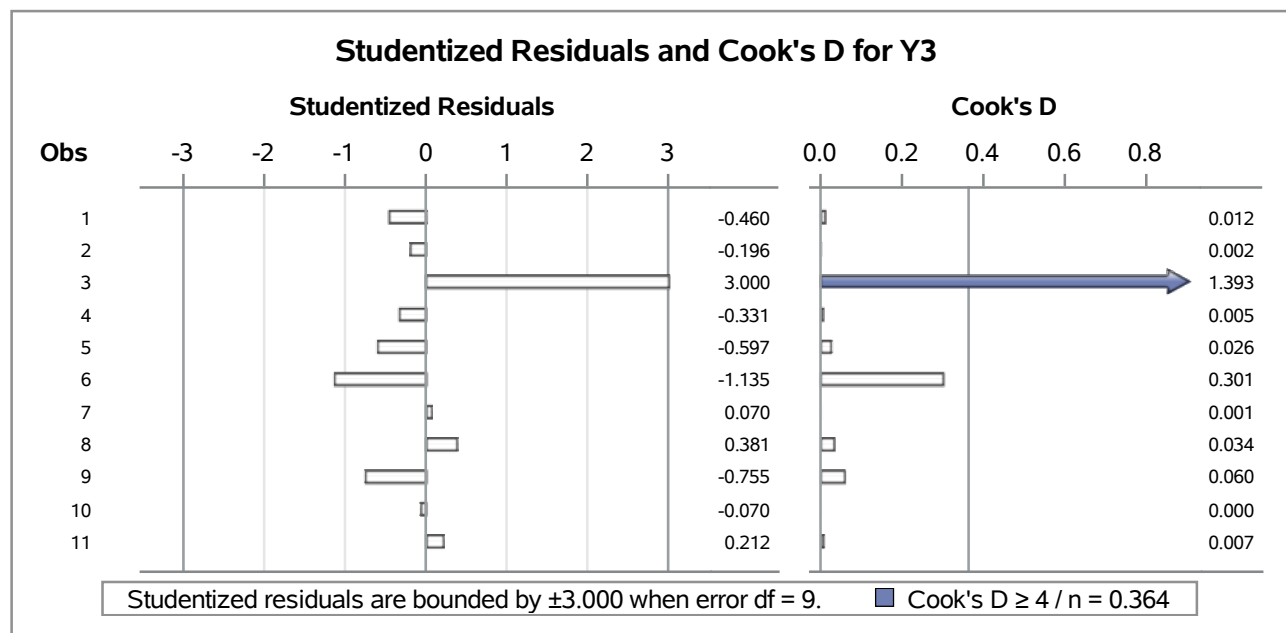
| | | | |
|----------------|----------|----------|--------|
| Root MSE | 1.23631 | R-Square | 0.6663 |
| Dependent Mean | 7.50000 | Adj R-Sq | 0.6292 |
| Coeff Var | 16.48415 | | |

| Parameter Estimates | | | | | |
|---------------------|----|--------------------|----------------|---------|---------|
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > t |
| Intercept | 1 | 3.00245 | 1.12448 | 2.67 | 0.0256 |
| X | 1 | 0.49973 | 0.11788 | 4.24 | 0.0022 |

An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

The REG Procedure
Model: MODEL3
Dependent Variable: Y3

| Output Statistics | | | | | | | |
|-------------------|----------|----------|------------|-----------|----------|-----------|----------|
| Obs | Residual | RStudent | Hat Diag H | Cov Ratio | DFFITS | DFBETAS | |
| | | | | | | Intercept | X |
| 1 | -0.5397 | -0.4391 | 0.1000 | 1.3409 | -0.1464 | -0.0046 | -0.0441 |
| 2 | -0.2303 | -0.1855 | 0.1000 | 1.3942 | -0.0618 | -0.0371 | 0.0186 |
| 3 | 3.2411 | 1204 | 0.2364 | 0.0000 | 669.5875 | -357.910 | 525.2677 |
| 4 | -0.3900 | -0.3138 | 0.0909 | 1.3585 | -0.0992 | -0.0329 | 0.0000 |
| 5 | -0.6895 | -0.5743 | 0.1273 | 1.3376 | -0.2193 | 0.0492 | -0.1172 |
| 6 | -1.1586 | -1.1560 | 0.3182 | 1.3629 | -0.7897 | 0.4897 | -0.6674 |
| 7 | 0.0792 | 0.0664 | 0.1727 | 1.5282 | 0.0303 | 0.0270 | -0.0209 |
| 8 | 0.3886 | 0.3619 | 0.3182 | 1.7969 | 0.2472 | 0.2409 | -0.2089 |
| 9 | -0.8492 | -0.7357 | 0.1727 | 1.3421 | -0.3362 | 0.1374 | -0.2314 |
| 10 | -0.0805 | -0.0658 | 0.1273 | 1.4486 | -0.0251 | -0.0197 | 0.0134 |
| 11 | 0.2289 | 0.2003 | 0.2364 | 1.6409 | 0.1114 | 0.1054 | -0.0874 |



| | |
|-------------------------------|----------|
| Sum of Residuals | 0 |
| Sum of Squared Residuals | 13.75619 |
| Predicted Residual SS (PRESS) | 23.62104 |

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 13

The REG Procedure
Model: MODEL4
Dependent Variable: Y4

| | |
|-----------------------------|----|
| Number of Observations Read | 11 |
| Number of Observations Used | 11 |

| Analysis of Variance | | | | | |
|----------------------|----|----------------|-------------|---------|--------|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 27.50000 | 27.50000 | 18.06 | 0.0021 |
| Error | 9 | 13.70580 | 1.52287 | | |
| Corrected Total | 10 | 41.20580 | | | |

| | | | |
|----------------|----------|----------|--------|
| Root MSE | 1.23404 | R-Square | 0.6674 |
| Dependent Mean | 7.50000 | Adj R-Sq | 0.6304 |
| Coeff Var | 16.45393 | | |

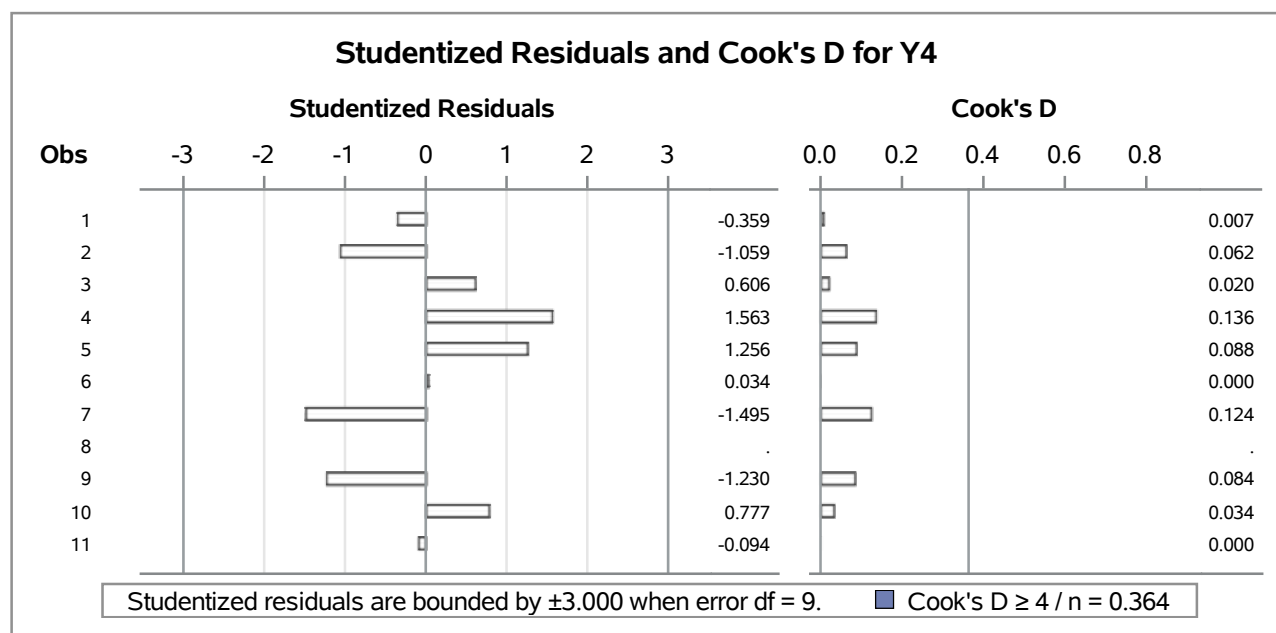
| Parameter Estimates | | | | | |
|---------------------|----|--------------------|----------------|---------|---------|
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > t |
| Intercept | 1 | 3.00000 | 1.12242 | 2.67 | 0.0255 |
| X4 | 1 | 0.50000 | 0.11766 | 4.25 | 0.0021 |

Regression is fun if you are careful!
An exploration of influential observations on regression outcomes.
Or, just because you get results doesn't mean they are worth anything.

16:06 Friday, August 14, 2020 14

The REG Procedure
Model: MODEL4
Dependent Variable: Y4

| Output Statistics | | | | | | | |
|-------------------|----------|----------|------------|-----------|---------|-----------|---------|
| Obs | Residual | RStudent | Hat Diag H | Cov Ratio | DFFITS | DFBETAS | |
| | | | | | | Intercept | X4 |
| 1 | -0.4200 | -0.3407 | 0.1000 | 1.3663 | -0.1136 | -0.0682 | 0.0342 |
| 2 | -1.2400 | -1.0673 | 0.1000 | 1.0775 | -0.3558 | -0.2137 | 0.1073 |
| 3 | 0.7100 | 0.5838 | 0.1000 | 1.2937 | 0.1946 | 0.1169 | -0.0587 |
| 4 | 1.8300 | 1.7267 | 0.1000 | 0.7463 | 0.5756 | 0.3456 | -0.1735 |
| 5 | 1.4700 | 1.3035 | 0.1000 | 0.9567 | 0.4345 | 0.2609 | -0.1310 |
| 6 | 0.0400 | 0.0322 | 0.1000 | 1.4059 | 0.0107 | 0.0064 | -0.0032 |
| 7 | -1.7500 | -1.6255 | 0.1000 | 0.7947 | -0.5418 | -0.3254 | 0.1634 |
| 8 | 0 | . | 1.0000 | . | . | . | . |
| 9 | -1.4400 | -1.2715 | 0.1000 | 0.9732 | -0.4238 | -0.2545 | 0.1278 |
| 10 | 0.9100 | 0.7588 | 0.1000 | 1.2238 | 0.2529 | 0.1519 | -0.0763 |
| 11 | -0.1100 | -0.0886 | 0.1000 | 1.4035 | -0.0295 | -0.0177 | 0.0089 |



| | |
|-------------------------------|----------|
| Sum of Residuals | 0 |
| Sum of Squared Residuals | 13.70580 |
| Predicted Residual SS (PRESS) | 16.92074 |

Author: Steven C. Myers, econdatascience.com

Source: Anscombe Quartet, https://en.wikipedia.org/wiki/Anscombe%27s_quartet

For interpretation of Influence Statistics see

https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_reg_details26.htm&docsetVersion=15.1&locale=en