Regression and Transformed Data: Section 4.1

1. Type these into L1 and L2.

|  |  |  |  |
| --- | --- | --- | --- |
| **x** | **y** | **ln x** | **ln y** |
| 7.3891 | 54.598 | 2.000006 | 3.999997 |
| 22.198 | 492.75 | 3.100002 | 6.200002 |
| 27.113 | 735.1 | 3.300013 | 6.600007 |
| 33.115 | 1096.6 | 3.499986 | 6.99997 |
| 49.402 | 2440.6 | 3.899991 | 7.799999 |
| 60.34 | 3641 | 4.099995 | 8.200014 |
| 73.7 | 5431.7 | 4.300003 | 8.600007 |
| 90.017 | 8103.1 | 4.499999 | 9.000002 |

1. Draw a rough scatterplot.



1. Find the LSRL. Is it a good idea to find the LSRL? What are r and r2? Interpret these. Make a prediction of y when x = 70.

r = .97 and r2=.94. Well, these seem pretty good. They are not perfect, but they indicate that this might be a decent model. (Of course, The scatterplot **definitely** shows us that this is not true.)

I can’t put this in context…there is no context.

1. What is the residual for the first point (x = 7.3891)? Draw a rough sketch of the residual plot. Interpret it.

y–= 54.598 –( = 54.598 – (–974.3127) =1028.9107



This residual plot shows a **clear** pattern. This contradicts the information provided by r2, but confirms what we saw in the scatterplot…we should not be using this model to make predictions for this data!

5. Perhaps we could transform this data. Take the log of both x and y. Fill in the chart at the top of the page.

Done.

1. Draw a rough scatterplot of this data.



You were not told to interpret, BUT here is an interpretation anyway. The scatterplot of the transformed data is much more linear than our original scatterplot. That one point that’s a bit far away is not really that far since we are using a scale of 0.5.

1. Find the LSRL of the transformed data. Is this a good idea? What are r and r2?

Both r and r2 are 1. This supports our decision to calculate the LSRL of the transformed data (go us!). It appears that our model is a good fit.

1. Draw a rough residual plot. What is the sum of the residuals? What should it be?

The sum of the residuals is ~0 (which is what it should be).

This residual plot is **definitely** better than our first plot. There looks like there may be a pattern, but this is clearly a better model. Also, look how tiny those residuals are!

1. Use your regression line to find a prediction if x = 70.

I can’t put this in context…there is no context.

1. "Untransform" your LSRL to find a regression equation in terms of x and y instead of log x and log y.

