Name: KEY

"Transformations to Achieve Linearity"

Consider the following data set:

x	y
.2	1.2
.4	1.3
.8	1.4
1.2	1.5
1.5	1.55
1.9	1.6
1.95	1.6
2.2	1.6
2.3	1.61
2.5	1.62
3	1.64

2

11. Is a linear fit appropriate? NO. Justify your answer.

3

← Scatterplot shows curved pattern.

Even though the lin. corr. coeff. is $r = .937$, which is very high, the residual plot shows a clear pattern, meaning lin. fit is not appropriate.

12. Two fits are being considered: a linear fit of log x to y^3, and a linear fit of log x to log y. Which fit is more appropriate, and why? Justify your answer with a scatterplot and two residual plots. Use reverse side if necessary.

2

Scatterplot

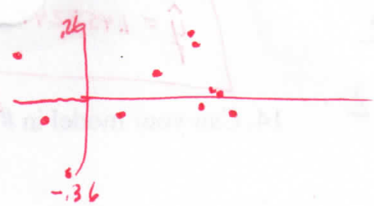
(1) $y^3 \approx a + b \log x$

$a = 3.265118833$

$b = 2.460056023$

$r = .9900909$

Resid. plots for lin. fit:



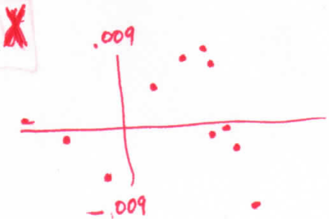
2+2 (resid. plots)

(2) $\log y \approx a + b \log x$

$a = .163829431$

$b = .1216224207$

$r = .9951644$



2

1 for ID of winner

Second fit is better, since r value is stronger, but mainly because the residuals are much smaller. Note that patterns are evident in both residual plots which is not good.

13. For the fit that you chose in #12, compute the equation for $\hat{y}$. Then use your equation to estimate y when $x = 0.92$.

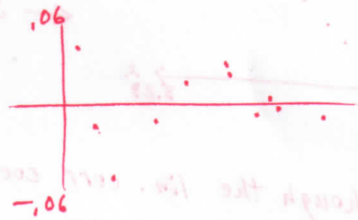
①

$y^3 \approx a + b \log x$

$\hat{y} = \sqrt[3]{a + b \log x}$

where $a = 3.265118833$, $b = 2.460056023$

Resid. plot:

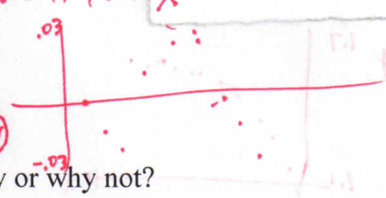


② $\log y \approx a + b \log x$ $a = .163829$, $b = .1216224207$

$\hat{y} = 10^{a + b \log x} = 10^{a} 10^{b \log x} = 1.458241424\, x^{.1216224207}$

2 $\boxed{\hat{y} = 1.45824124\, x^{.1216224207}}$

Resid. plot:



2 $\hat{y}(0.92) = 1.45824124(.92)^{.1216224207} = 1.444$

14. Can your model in #13 be used to predict y when $x = 3.3$? Why or why not?

2 $\boxed{\text{NO.}}$ That would be extrapolation, since

2 3.3 is outside the domain of x-values for which we have any data.