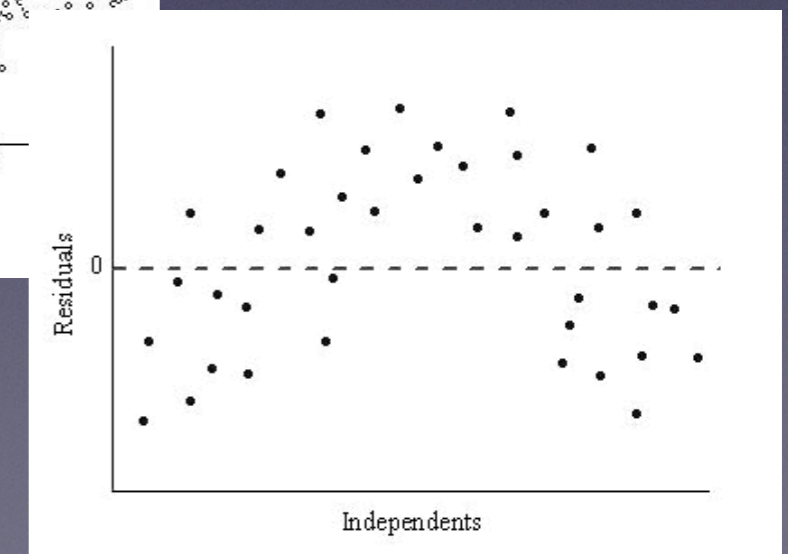
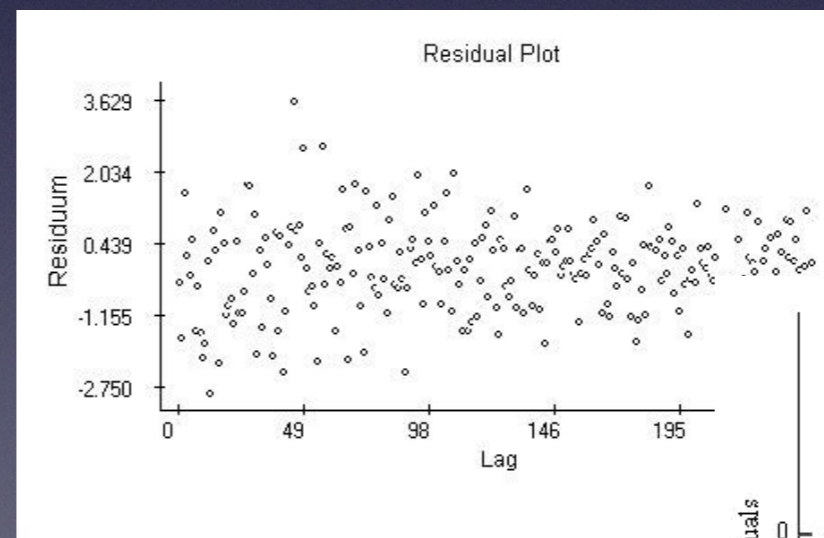
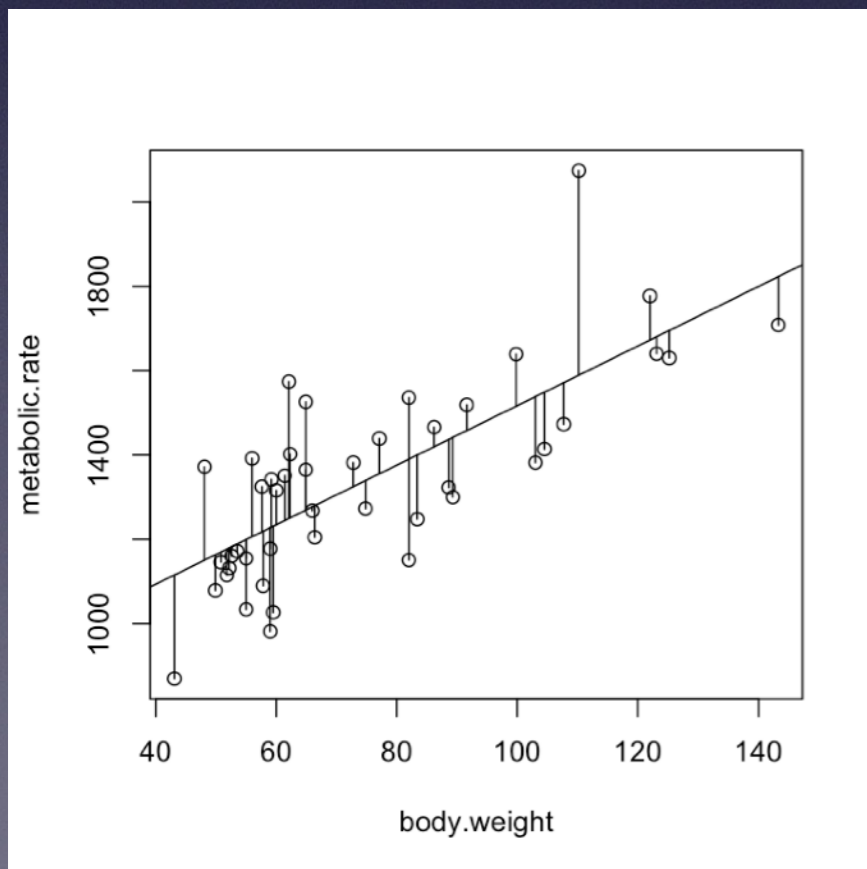


Residuals

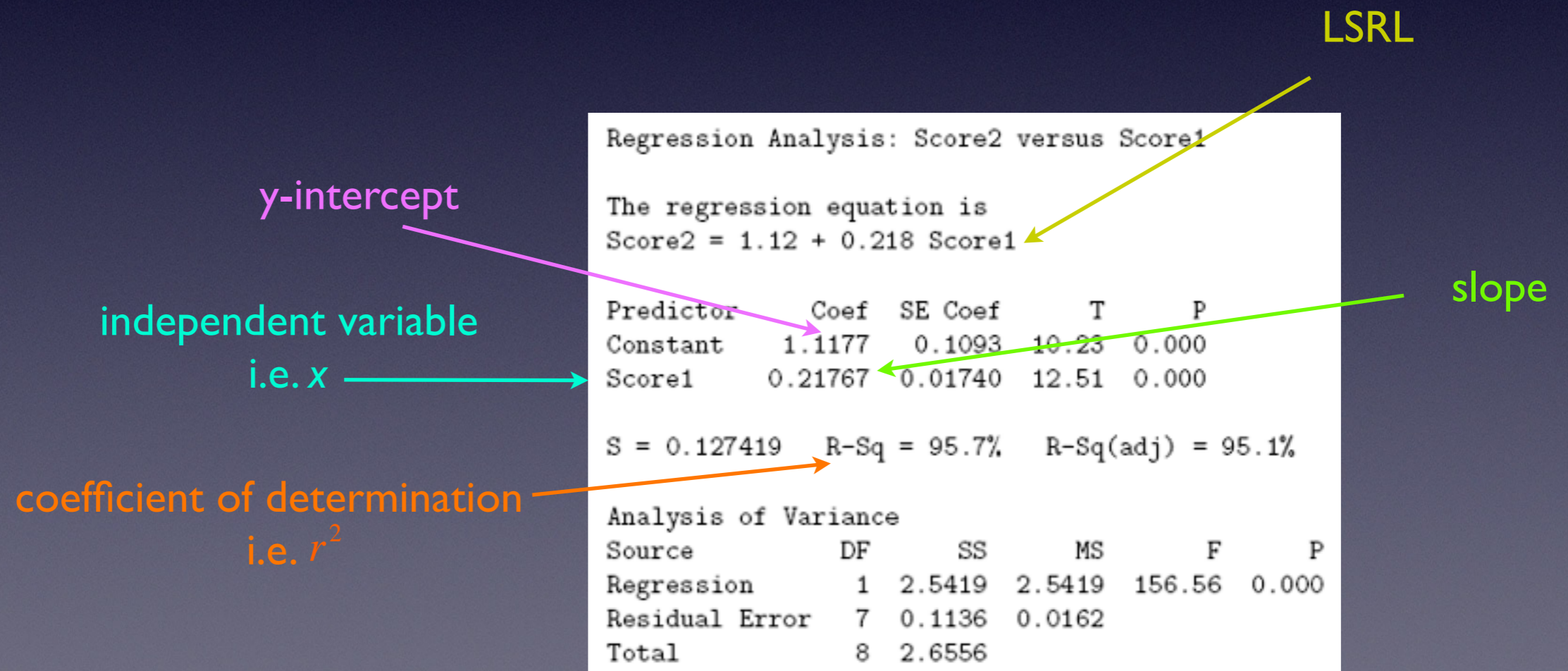
- To find a residual, subtract the predicted y -value from the actual y -value $\text{residual} = y - \hat{y}$.
- The mean of the residuals is 0.
- The best fit, or least squares, line minimizes the sum of the squares of the residuals.



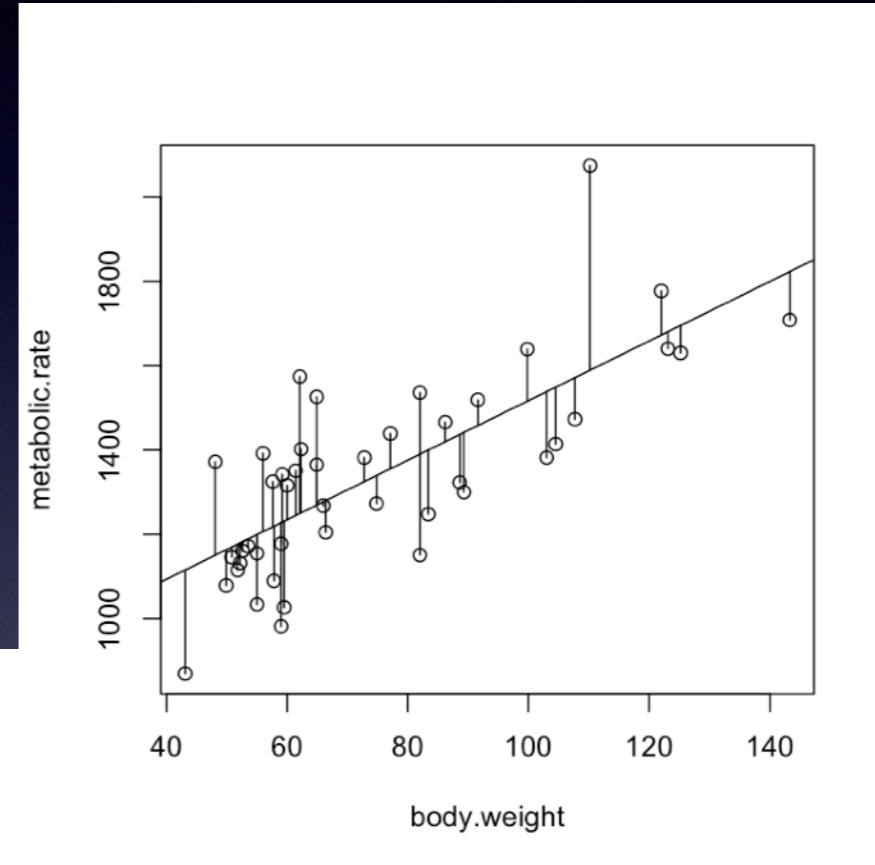
r^2 Coefficient of Determination

Coefficient of Determination: % of the variability in (context of y) can be explained by (context of x)

Example: Example: y = height of a plant in cm, x = age in months, $r = 0.945$, $r^2 = .893$
89.3% of the variability in the height of the plant can be explained by the age of the plant.



S_e - standard deviation about the LSRL, i.e., the average residual length



Regression Analysis: Score2 versus Score1

The regression equation is
Score2 = 1.12 + 0.218 Score1

Predictor	Coef	SE Coef	T	P
Constant	1.1177	0.1093	10.23	0.000
Score1	0.21767	0.01740	12.51	0.000

S = 0.127419 R-Sq = 95.7% R-Sq(adj) = 95.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2.5419	2.5419	156.56	0.000
Residual Error	7	0.1136	0.0162		
Total	8	2.6556			

y-intercept

slope

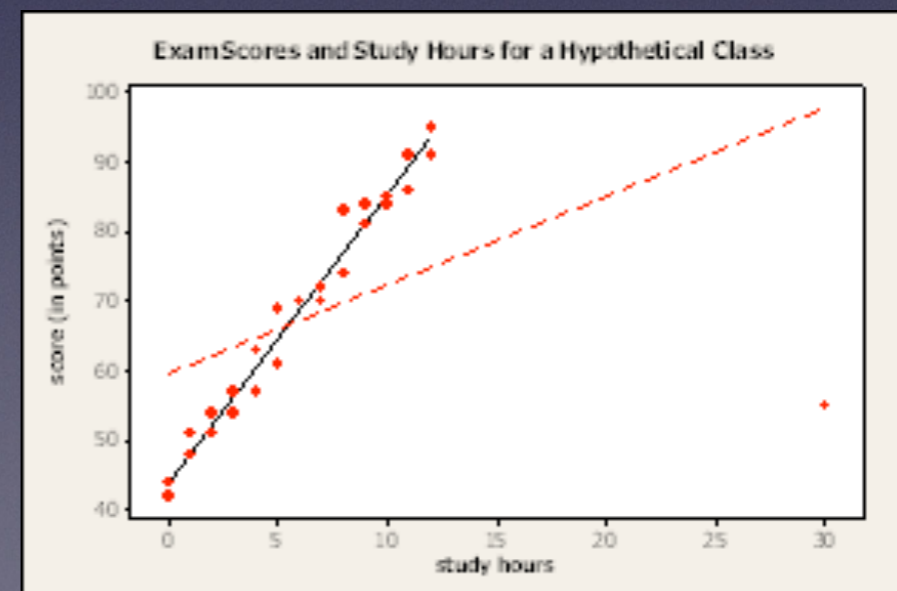
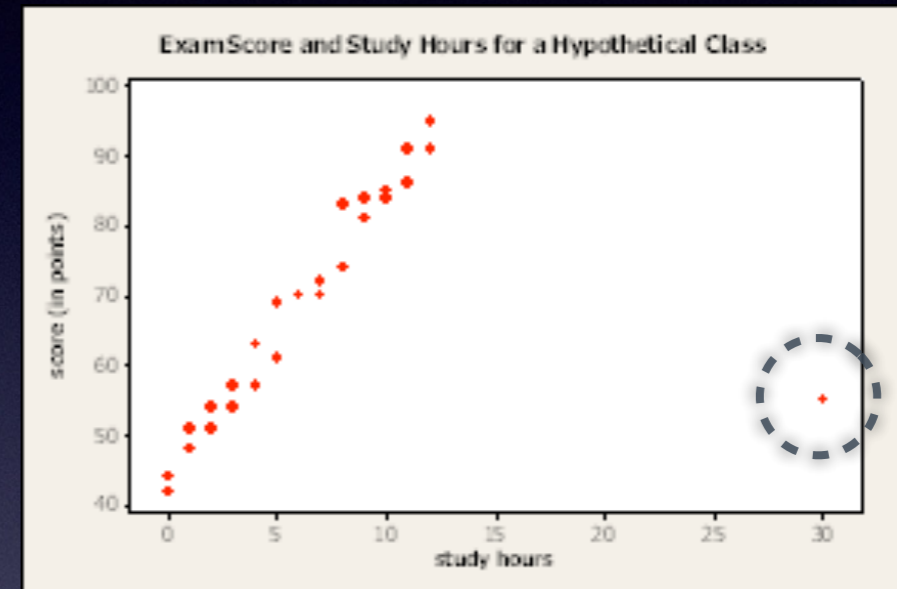
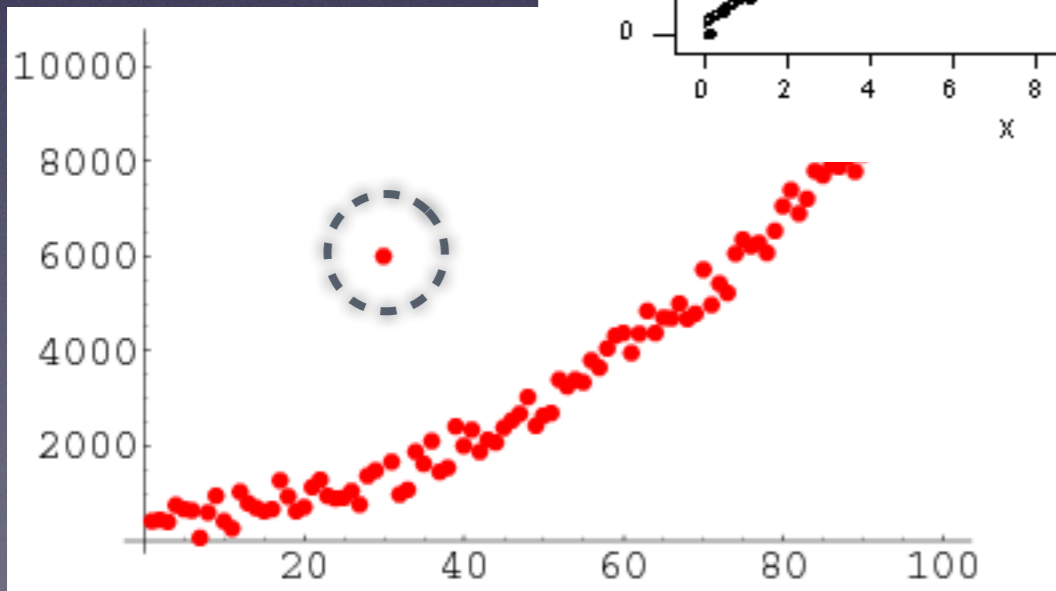
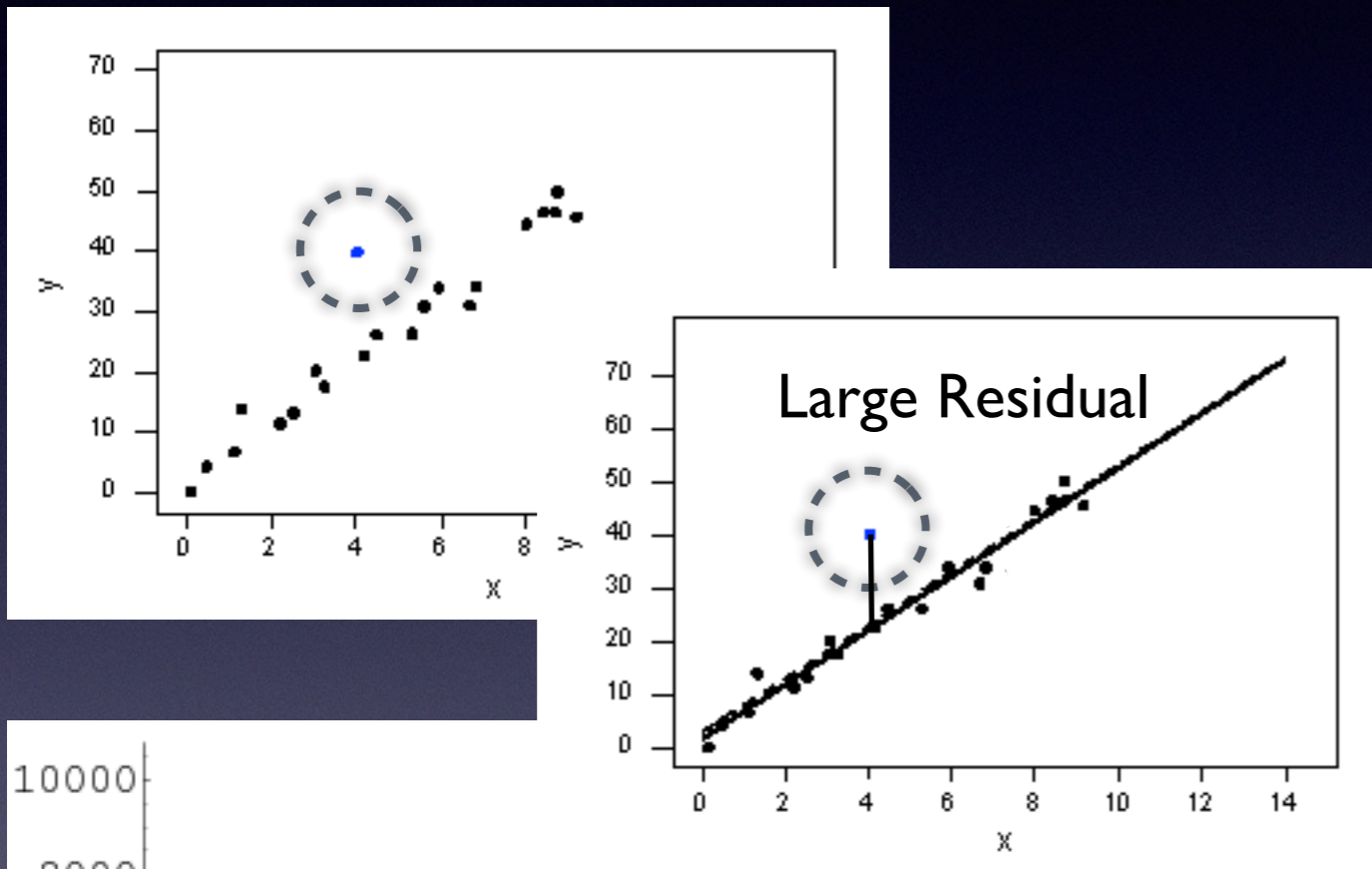
S_e

coefficient of determination
i.e. r^2

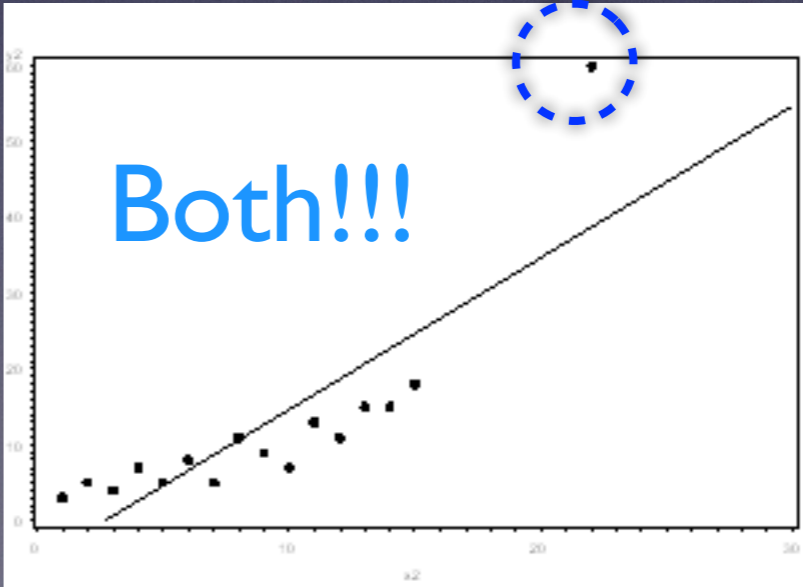
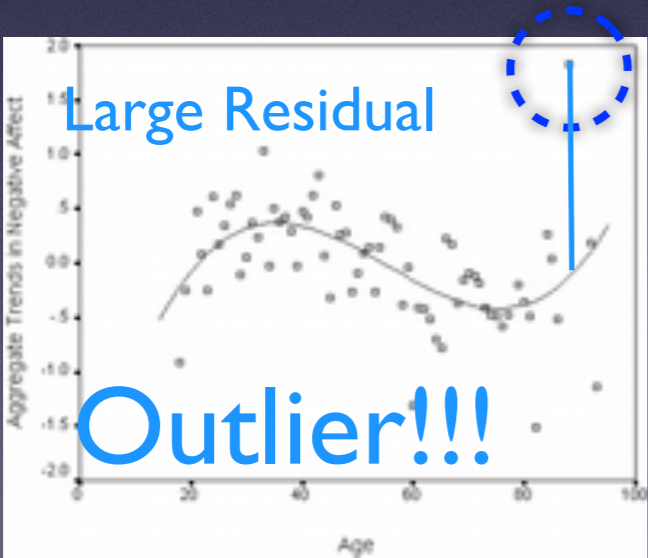
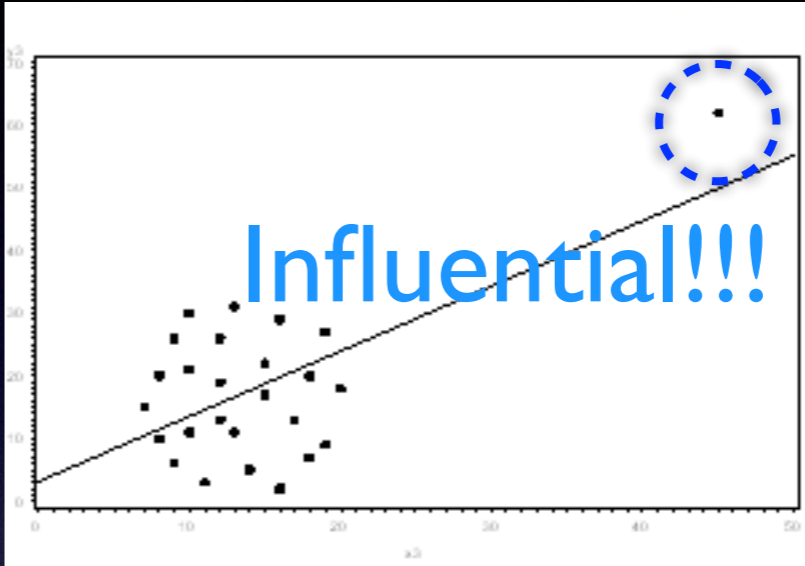
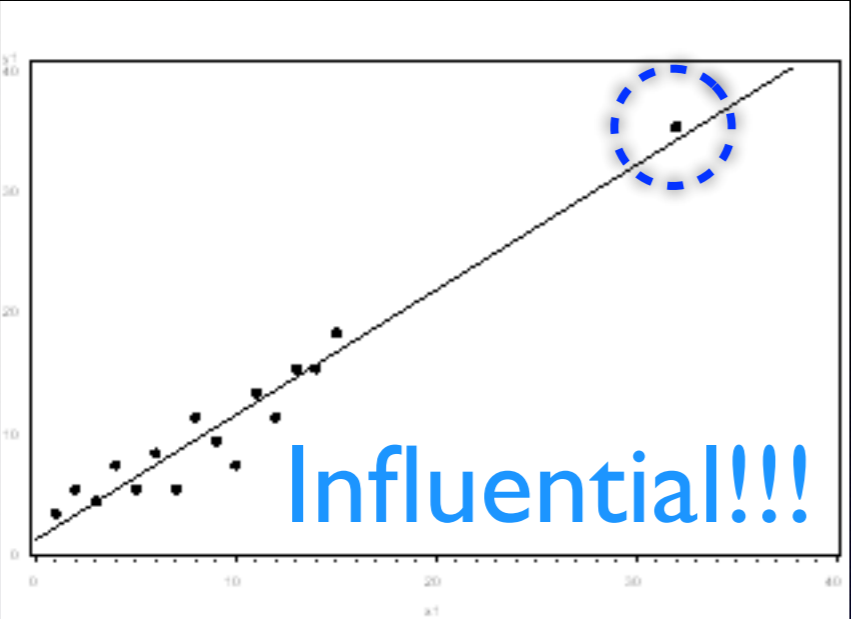
Outliers vs. Influential Points

An **outlier** is a data point whose response y does not follow the general trend of the rest of the data.

A data point is **influential** if it unduly influences any part of a regression analysis, such as the predicted responses, the estimated slope coefficients, or the hypothesis test results



Influential or Outlier?



Is this model a good fit?

Four factors go into your decision -

1. Does the scatterplot look linear?
2. r - should be close to $+1$ or -1 for a good fit
3. Residual plot - should be scattered for a good fit
4. s_e (the standard deviation of the residuals i.e. the average residual size) - should be small for a good fit...you can only tell if s_e is small if there is more than one MINTAB output to compare s_e 's