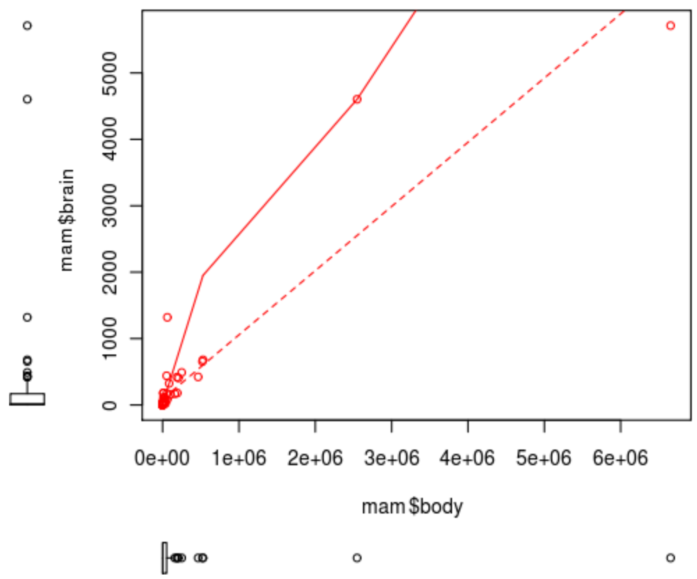
David O Neill 0813001

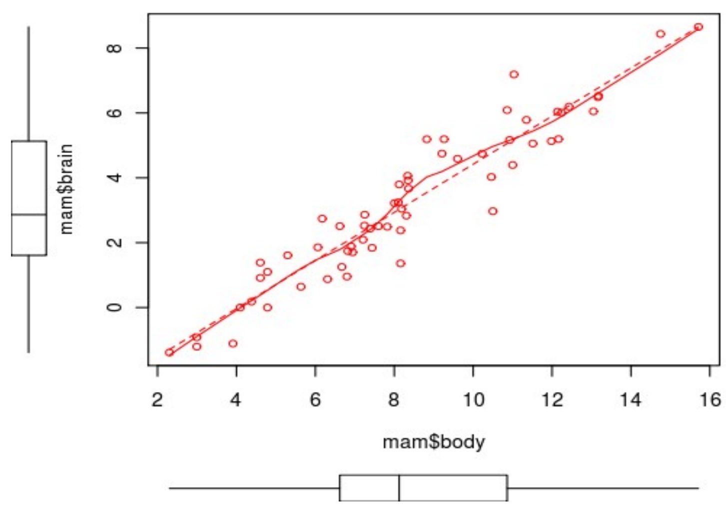
Q1

Firstly we convert the kilograms to grams so that were working with the same units for the body, brain weight relation.

It is reasonable to use the brain weight as a response to body weight because by calculating the linear regression model we can see that 87.04% (R-squared: 0.8704) of the intercepts (60 degrees of freedom) are explained by the initial model.
The initial scatter plot is difficult to see a relationship.



By applying the natural logarithm we can see a better relation but exposes a greater problem with the model. Lesser short-tailed shrew brain size is too small -Inf (y) -1.9661129 (x).



Q2

Fitted Regression Line

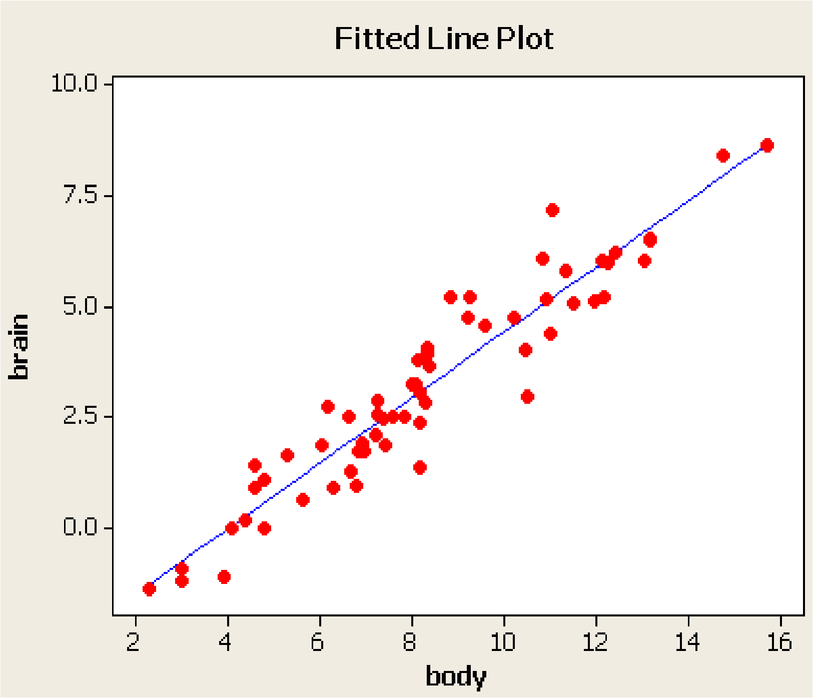
Total numbers DF	60
slope(b)	0.74
y-intercept(a)	-3
regression equation (y)	brain(y) = -2.998 + 0.7415 body(x)

Coefficients and Standard error

	Coefficients	Standard Error of Coefficients
Brain	-2.997800	0.2793
Body	0.74151	0.051730

S and R^2

S-Value	0.739680
R-squared	90.40%



Q3

Two key assumptions are made, X (body) has some relation to Y (brain) and the data is presumed error free. The statistical relationship between the error terms and the regressors plays an important role in determining whether an estimation procedure has desirable sampling properties such as being unbiased and consistent.

The variances of the error terms may be equal across the *n* units (homoscedasticity) or not (heteroscedasticity). Some linear regression (as seen in the above plots, linearity) estimation methods give less precise parameter estimates and misleading inferential quantities such as standard errors when substantial heteroscedasticity is present.

The arrangement, or probability distribution of the predictor variables *x* has a major influence on the precision of estimates of β.

Q4

95% Confidence Interval

	Ln Min	Ln max	E^ Min	E^ max
brain	2.588000	3.800000	13.3 g	44.7 g
body	7.573000	9.128000	1944.96 g	9209.58 g

Q5

Let μ_d = mean of the differences for two sets of brain weights (brain, 3/4)

Hypotheses conclusion

Ho: $\mu_d = 0$ no difference for two sets of data

Ha: $\mu_d \neq 0$ differences exist between the two sets (brain, 3/4)

One-Sample T: diff

Test of mu = 0 vs not = 0

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
diff	50	1.821	1.054	0.149	(1.522, 2.120)	12.22	0.000

P-value is too small to accept the null hypotheses therefore we reject this and accept the alternative hypotheses

Q6

Fitted Regression Line **primates**

Total numbers	10
slope(b)	0.87
y-intercept(a)	-3.24
regression equation (y)	Brain(y) = - 3.244 + 0.8722 body(x)

Fitted Regression Line **non primates**

Total numbers	49
slope(b)	0.72
y-intercept(a)	-3.04
regression equation (y)	Brain(y) = - 3.035 + 0.7245 body(x)

Statistically primates have higher brain mass than their non primate counterparts of the same weight. This is because of the slope of the fitted line equations for primates is greater than the non primates.

Q7

Estimated brain size for pen-tailed tree shrew **1.079147118 grams**.

Y(brain) = intercept +- 1.96 *S + slope * body(x)

Brain(y) = -3.244 - (1.96 * 2.669) + (0.872245 * 3.806662) = 0.005771068 min

Brain(y) = -3.244 +(1.96 * 2.669) + (0.872245 * 3.806662) = 201.861515609 max