

# Homework 12

Using the R programming language, I was able to calculate four linear regressions for the Kittiwakes data set: First, predicting population from Area; second, predicting  $\log(\text{population})$  from Area; third, predicting population from  $\log(\text{area})$ ; and fourth, predicting  $\log(\text{population})$  from  $\log(\text{area})$ .

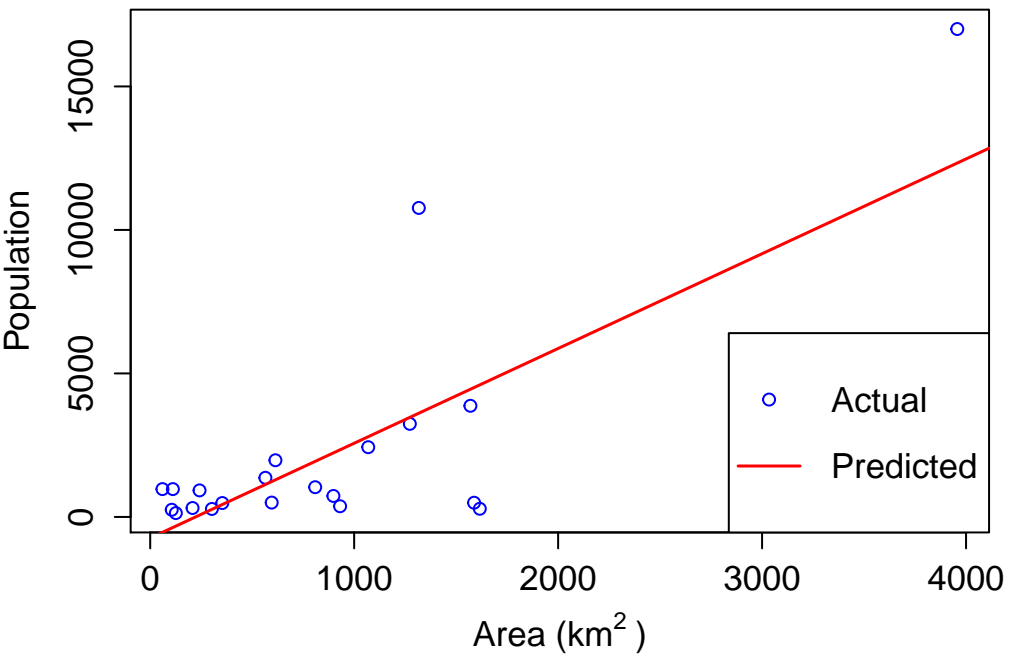
The graphs for all of the linear regressions has been attached on the following sheet, and illustrate the relations between the actual and predicted values, as well as describe an accuracy metric for the prediction, the adjusted R-squared value.

R-Squared Values	Population v. Area	Log(Population) v. Area	Population v. Log(Area)	Log(Population) v. Log(Area)
Normal R-Squared	0.616	0.473	0.315	0.364
Adjusted R-Squared	0.597	0.446	0.280	0.333

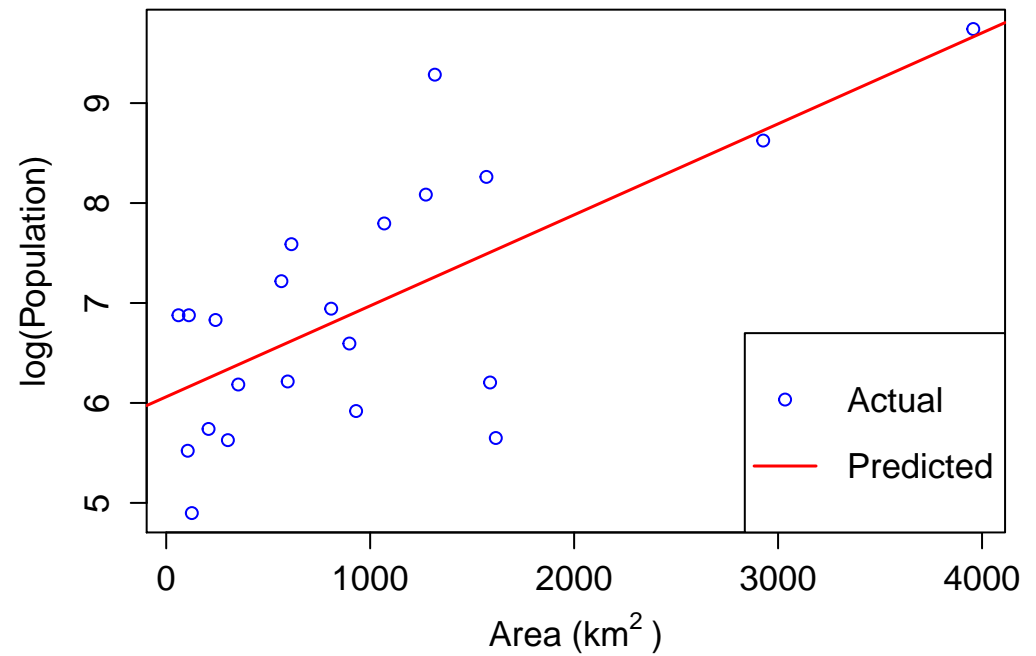
According to the above R-squared values (the higher the better), the *Population vs. Area* linear regression is the most accurate to the true values, and therefore would be the prediction which I would trust the most, since any new prediction would be as close as possible to the currently known data set.

*Graphs attached on next page*

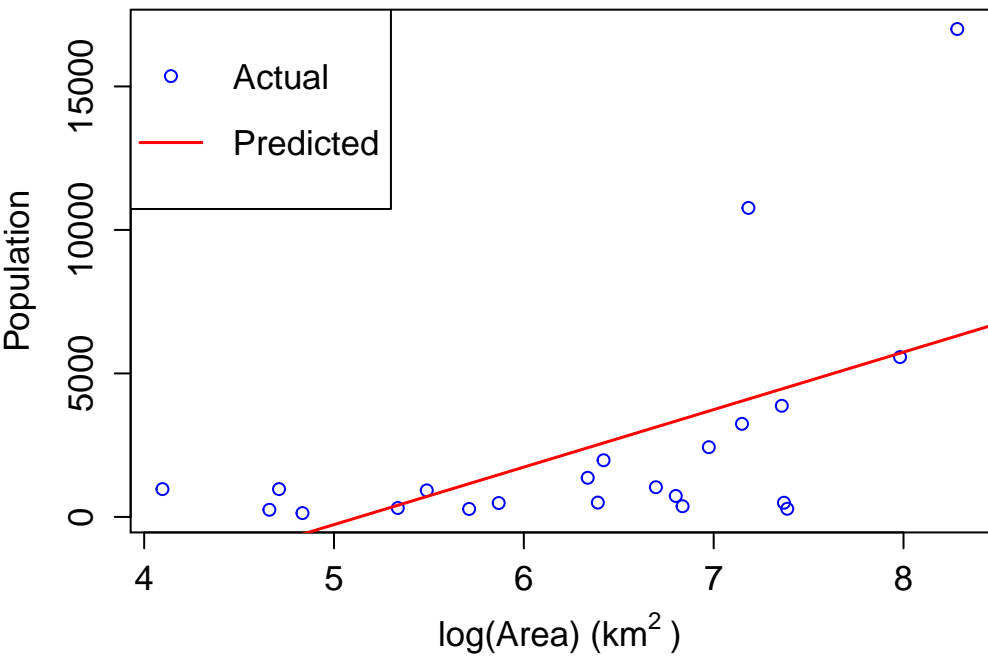
**Kittiwake Colonies ( $R^2_{\text{adj}} = 0.597$ )**



**Kittiwake Colonies ( $R^2_{\text{adj}} = 0.446$ )**



**Kittiwake Colonies ( $R^2_{\text{adj}} = 0.28$ )**



**Kittiwake Colonies ( $R^2_{\text{adj}} = 0.333$ )**

