BIO2426 Assignment 3

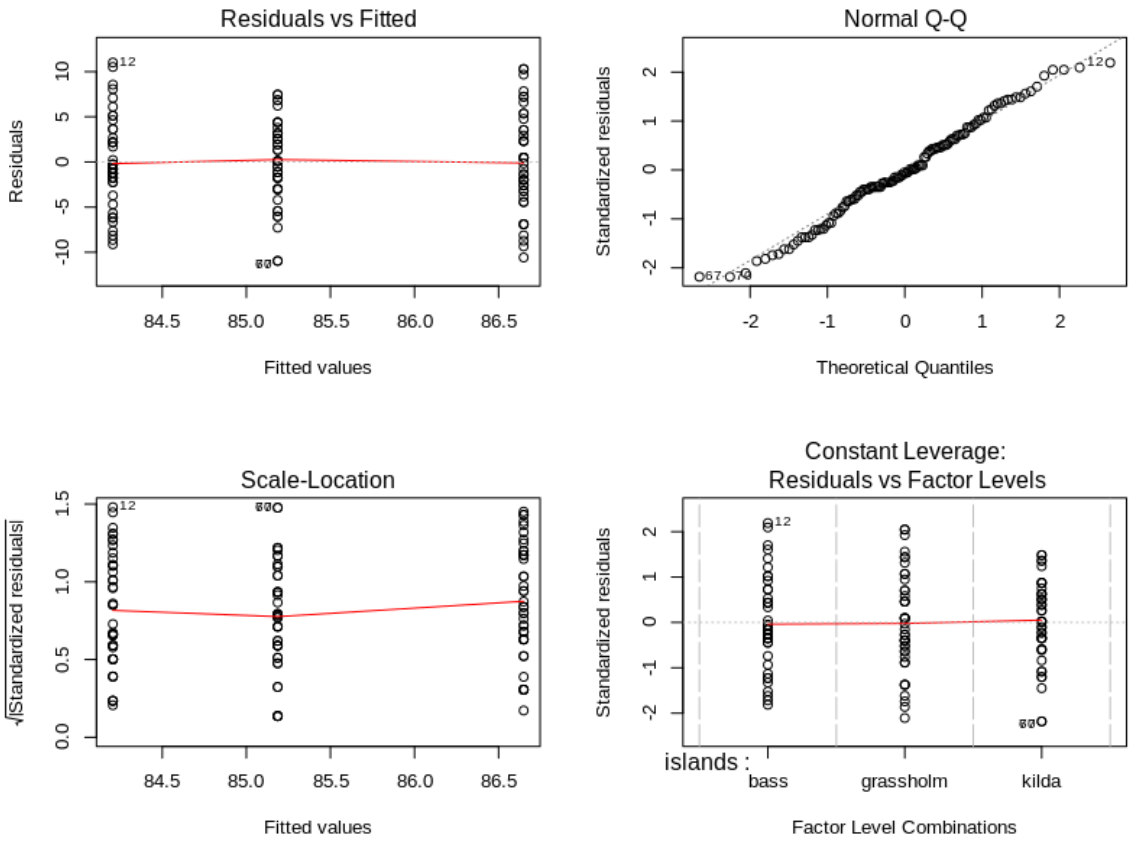
In this practical, we are interested in whether hatchling gannets on three different islands vary in their hatching weight. To address this question we will use linear models with multiple explanatory variables. Download the data "gannets.txt" from ELE for use in the practical.

Perform each task below and put your answer into the spaces provided. Any requested graphics should be copied from R and pasted. To paste images from the R server to Word, in the R server, go to Export and Copy to Clipboard, then reshape image using the bottom right corner, right click on image and 'Copy image'. Then go to Word and go to Home>Paste Special and select Bitmap or Device independent bitmap. Other parts of answers should be provided as you would give them in a report, i.e. give the test statistic, the degrees of freedom or sample size and the p-value, in the style in which you would report them in a paper – with reference to the biological question, and direction of any effect. Do not paste unmodified output from R, and don't include your R-code. There is ample space to report your answers and graphs, given in the boxes underneath each question. Avoid unnecessary precision in your answers (e.g. r=0.1083176) **round all values to 3 decimal places.**

Part 1: Comparing means across several groups

We'd like to know whether the hatchling weight is different on the three different islands. Use *lm* to fit a linear model of hatchling weight among islands, remembering to store the model as an object and to use model plots to check assumptions. Use *anova(model)* to request the output from the model.

1. Give diagnostic plots of the model using the *plot(model)* command and comment on the plots (1 mark)



Residual vs Fitted:	Residuals are equally spread around a horizontal line indicating there are not non-linear relationships.
Normal Q-Q:	Residuals are normally distributed as they do not deviate too much from the line.
Scale-Location:	Residuals appear randomly spread either side of the line in distinguishable groups showing that variance in equal.
Constant Leverage:	There are no extreme outliers that cross Cook's Line so the model is not being influenced in an extreme way by outliers.

2. Use *anova(model)* to test whether the differences in hatchling weight amongst the islands are statistically significant. Report your results below as you would in a paper. (1 mark)

   There was no significant difference in the weights of Gannet hatchlings at each of the islands (one-way ANOVA, $F_{2,123} = 2.456$, $P = 0.090$)

3. Display the differences in hatchling weight among islands using a boxplot, remembering to add a figure legend (usually provided at the bottom of the figure and in Word) and label the axes. The figure legend should describe what the different parts of the plots (ie boxes, whiskers etc) show (1 mark)

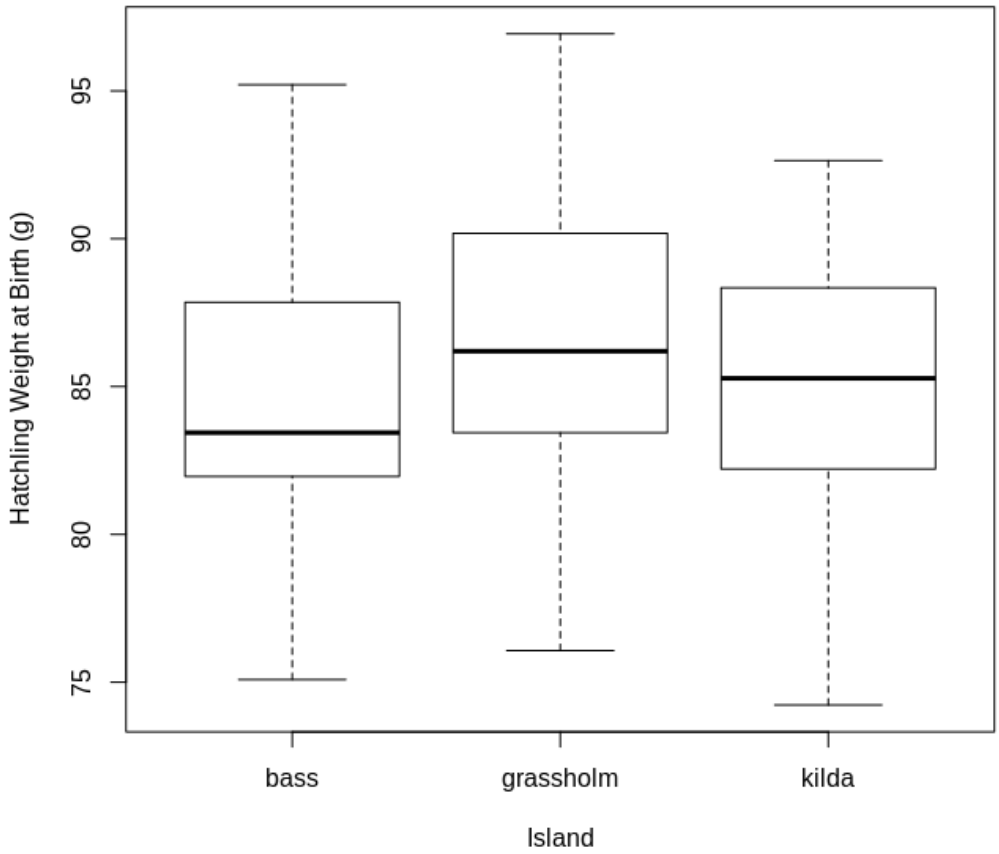


Figure 1 - The weights of hatchlings from three islands, the thick black line shows the median weight, while each box displays the hatchling weights data within the 1st and 3rd quartile for each island and the whiskers of each plot show the range of data for each island within 1.5x the interquartile range of either the 1st or 3rd quartile.

Part 2: LM with a covariate

You decide that what you're really interested in is not just whether the hatchlings differ in their weight between islands, but how much they weigh relative to their parent's body size.

1. Use a linear model (use the *lm* command) to examine hatchling weight for each island, controlling for parent's (adult) weight, by including it in the model. You will be looking at hatchling weight as a function of both island and parent's body size. Report the result from your final model as you would in a research paper (2 marks).

While Parent's weight affects the hatchling weight ($F_{(1,122)}$ = 61.978, P < 0.001), hatchling weight was significantly affected by the island they were born on, hatchlings being larger if they were born on Grassholm ($F_{(2,122)}$ = 3.098, P = 0.049) when accounting for Parent weight. There was no significant interaction between the island hatchlings were born on and the weight of the parents, ($F_{(2,120)}$ = 0.309, P = 0.735).

While mother's pre-pregnancy weight affects the birth weight ($F(1,81) = 20.296$, $p < 0.001$), birth weight was significantly affected by whether the mother smoked or not, producing smaller birth weights if the mother smoked ($F(1,81) = 8.672$, $p = 0.004$) when accounting for mother's pre-pregnancy weight. There was no significant interaction between smoking and mother's pre-pregnancy weight ($F(1,80) = 0.007$, $p = 0.935$).

There is no significant interaction effect between smoking and the mother's weight prior to pregnancy, (F(1,80)=0.007,P=0.935).

2. Use *summary* to investigate the explanatory power (R-squared) of your model e.g. *summary(model)*. Report this as a percentage, explaining what it means (1 mark).

   Adjusted R-Squared: 34.67%
   Shows how well the data fit this model, in this case it means that the variables in the model explain 34.67% of the variance in Hatchling Weight..

3. Use the model output to predict the mean hatchling weight of a gannet chick on Grassholm whose parent's body weight is 3.2kg. (1 mark)
   =90.073g

Part 3: Factorial LM

You are interested in whether or not hatchling weight varies between the sexes and the different islands.

1. Conduct a linear model to investigate the effect of sex and island on hatchling weight. Investigate whether an interaction exists, and describe the results of your analysis as you would in a paper. Display the difference in hatchling weight between sex and islands using a boxplot with appropriate axis labels and a figure legend to explain the plot. Ensure the axis labels are easily readable (3 marks). *HINT: To change the font size of an axis label, use the argument cex.axis=…. The value given will be proportionate to the standard size (i.e. cex.axis = 1 will be the usual size, cex.axis = 0.5 will be half the standard size).*

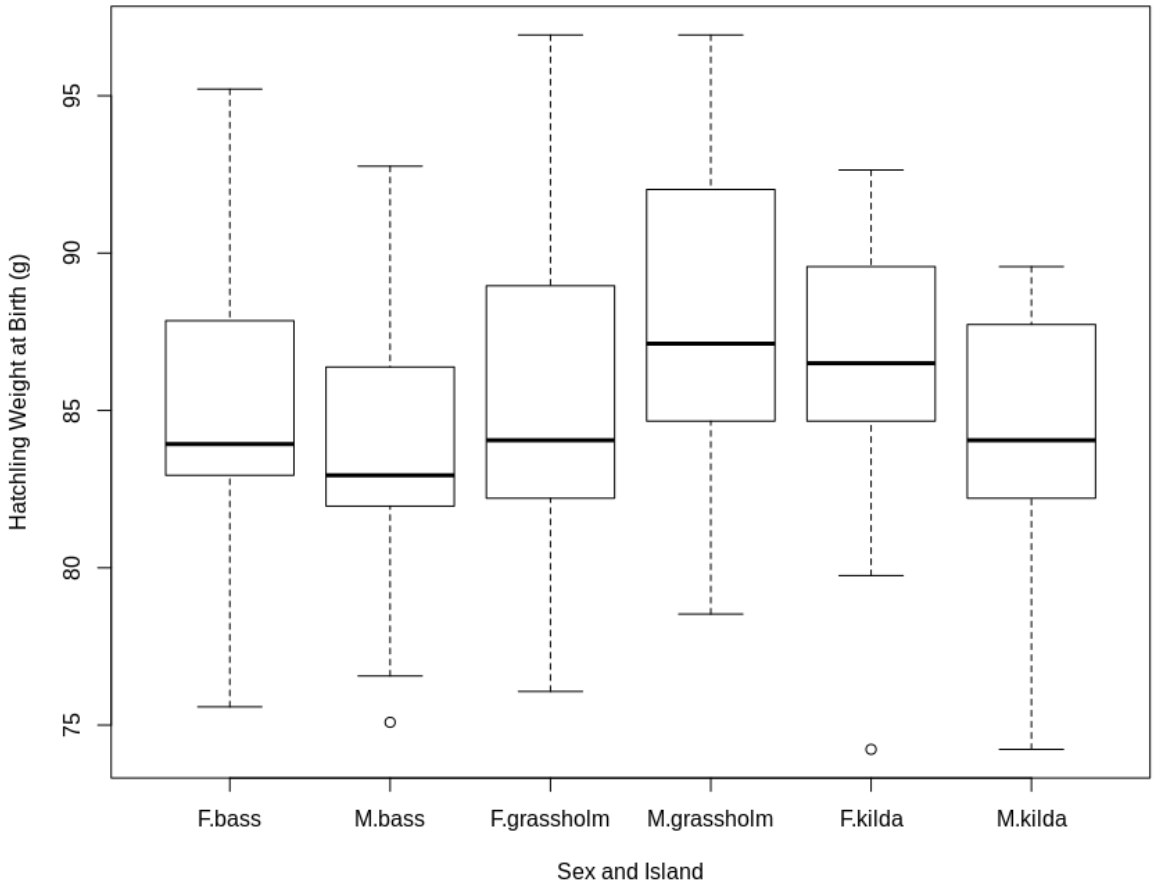


Figure 2 - The weights of male and female hatchlings from three islands, the thick black line shows the median weight, while each box displays the data between first and third quartile of the hatchling weights for each island and the whiskers of each plot show the data within 1.5x the interquartile range from the 1st and 3rd quartile. The dots display outliers that are not within 1.5x interquartile range from the 1st and 3rd quartile.

There is a significant interaction effect between the island hatchlings were born on and the sex of the hatchling on hatchling weight ($F_{2,120} = 3.854$, $p = 0.024$). Male chicks born on the island of Grassholm weighed significantly more than female chicks born on Grassholm; on the islands of Bass

and Kilda however, there was no significant difference between the weights of male and female hatchlings.