



National Certificate of Educational Achievement  
TAUMATA MĀTAURANGA Ā-MOTU KUA TĀEA

## Exemplar for Internal Assessment Resource Mathematics Level 3

### Resource title: Sub-species

This exemplar supports assessment against:

Achievement Standard 91581

Investigate bivariate measurement data

#### Expected responses

The moderators have developed expected student responses from a wide variety of sources

Date version published by  
Ministry of Education

December 2012  
To support internal assessment from 2013

For this task the emphasis is on investigating possible relationships. The choice of response and explanatory variables is not crucial. However in other tasks this could be important.

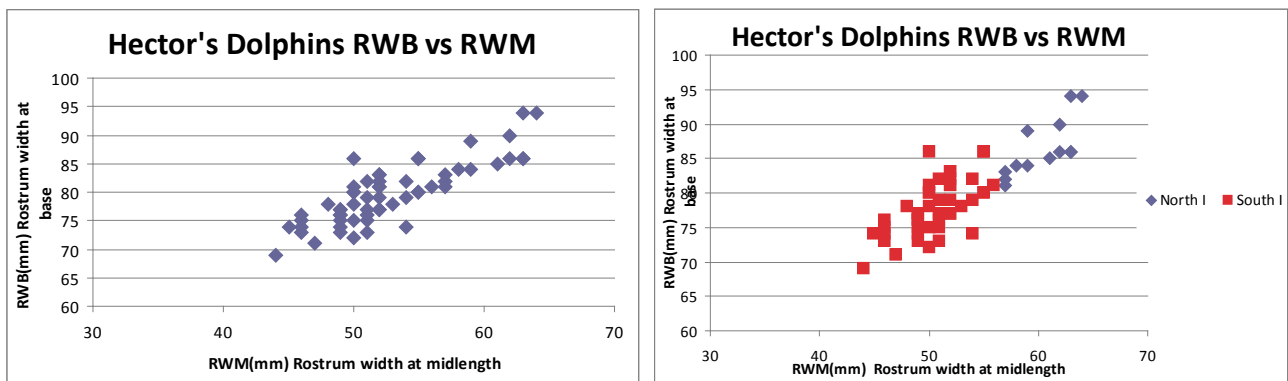
|    |   |
|----|---|
|    | Grade Boundary: Low Excellence  |
| 1. | <p>For Excellence the student is required to investigate bivariate measurement data with statistical insight. This involves integrating statistical and contextual knowledge throughout the investigation process, and may include reflecting about the process; considering other relevant variables; evaluating the adequacy of any models, or showing a deeper understanding of the models.</p> <p>An appropriate relationship question has been posed. Contextual aspects of the scientific report have been identified and related to the investigation. The student could have referenced details in the report which discussed possible reasons for differences, for example NI/SI separation in interglacial periods or differences in current flows (1).</p> <p>Appropriate displays have been selected and used. The choice of different pairs of variables demonstrates an understanding of the purpose of the investigation (2).</p> <p>Features of the data have been identified and the nature and strength of the relationship has been described in context. Comments indicate that the context of the scientific report has been considered, for example the discussion about outliers demonstrates a high level of understanding (3).</p> <p>An appropriate model has been found and it has been used to make predictions. The appropriateness of a linear model has been considered and the usefulness of predictions, in terms of the context of the investigation, has been discussed.<br/>The fact that the report said rostrum width at half-length and rostrum lengths were measurements which best distinguished the North and South island populations could have been identified (4).</p> <p>Findings have been communicated in a conclusion. Comments demonstrate a good understanding of the context and purpose of the investigation. Issues related to the gender and overall length of dolphins which are identified in the scientific report could have been identified and discussed. The use of 'bigger' is not completely clear (5).</p> <p>There is evidence of investigating bivariate measurement data with statistical insight in the integration of contextual knowledge which has been obtained from the report to a number of different aspects of the investigation.</p> <p>For a more secure Excellence some aspects of the scientific report which were overlooked would have been identified and discussed with respect to the analysis. The student could also have considered the relevance of the analysis to a wider population by acknowledging that these results only applied to Hector's Dolphins.</p> |

I need to decide if the classification of the North Island Hector's Dolphins as a new sub-species can be supported by looking at relationships. There are two types of variables in the data set, ones which measure lengths and ones that measure widths. So to cover different possibilities I will investigate three cases; length/length, length/width, width/width. This covers different possibilities; it could be that there is a relationship between the variables in one of these cases but not others.

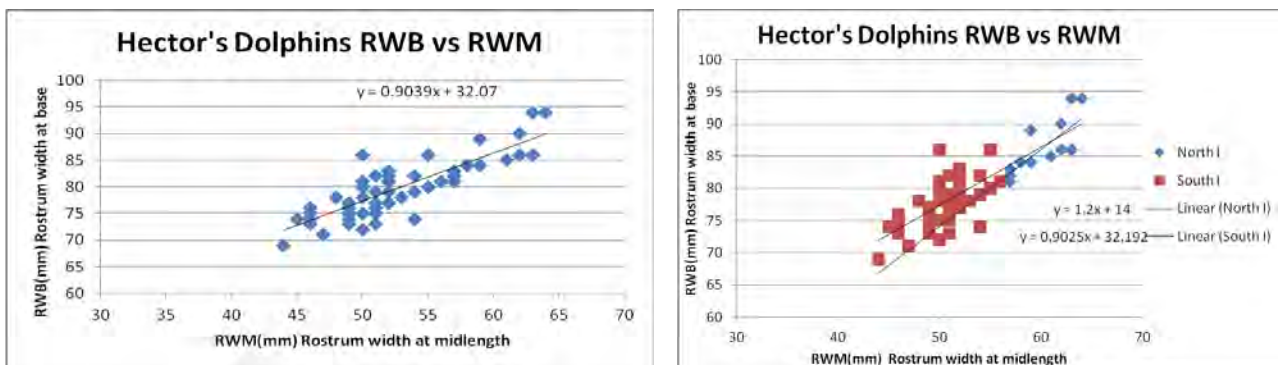
There are actually four different populations, three from the South Island and one from the North Island. While there might be differences in each of the populations the North Island dolphins were thought to be the most different and this is the reason for the classification. Measurements were taken from skeletal material in different museums. (1)

**Case 1: width/width**

Is there a relationship between RWM (rostrum width at midlength) and RWB (rostrum width at base)? My predictor variable is RWM and my response variable is RWB.

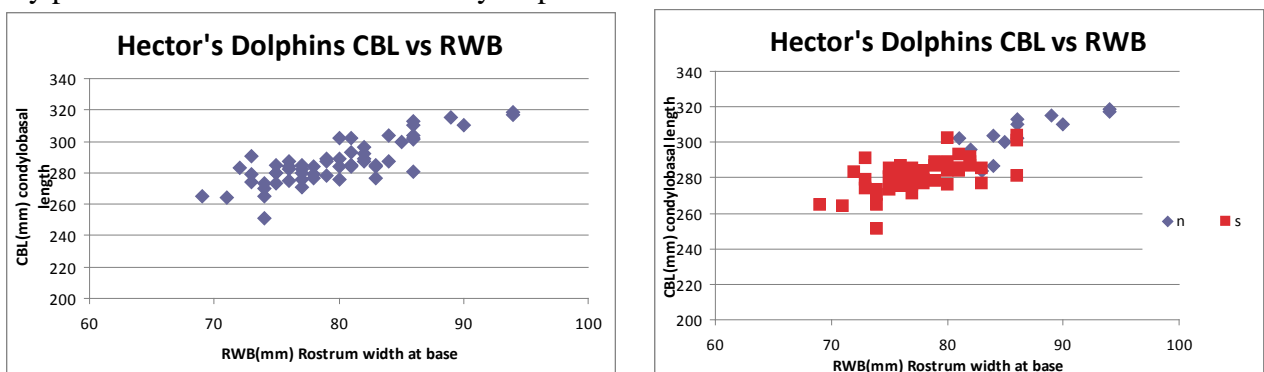


Looking at both of these graphs I believe that an appropriate model will be a linear one. There is nothing to suggest that any other type of model should be used.

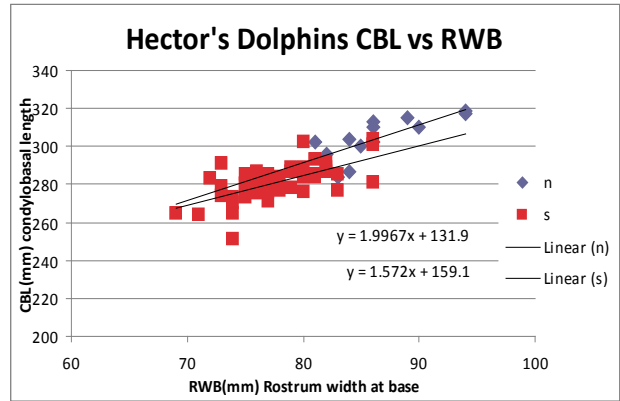
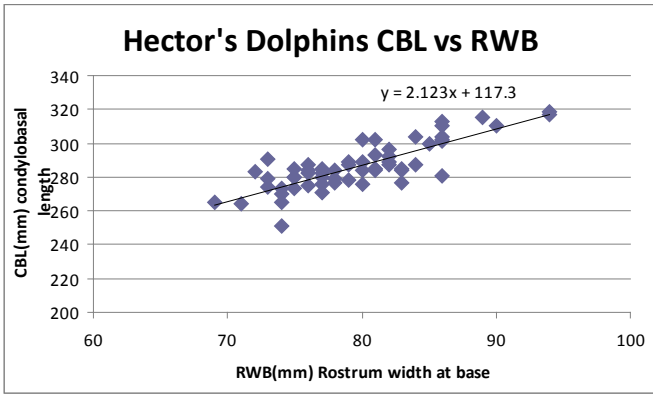


**Case 2: width/length**

Is there a relationship between RWB (rostrum width at base) and CBL (condylobasal length)? My predictor variable is RWB and my response variable is CBL.



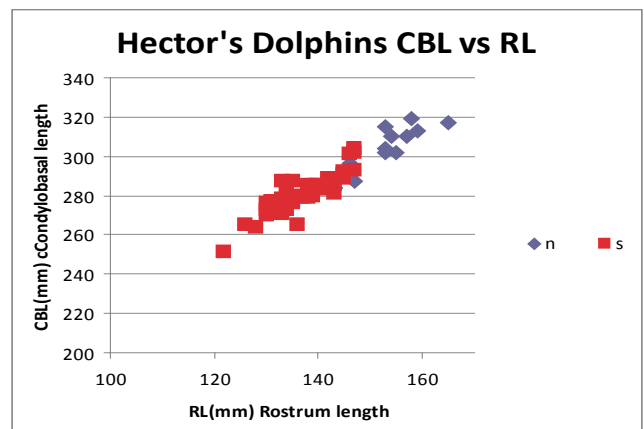
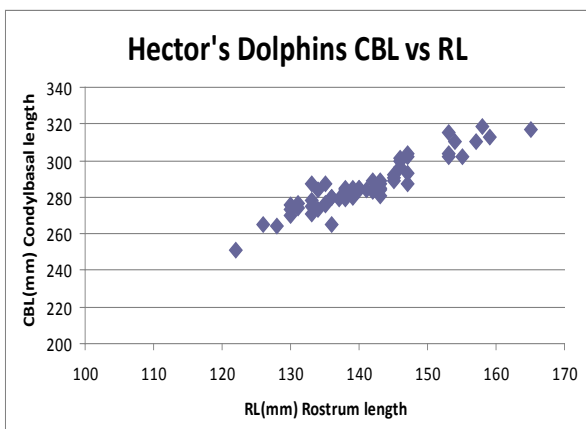
Again there is no reason to use anything other than a linear model.



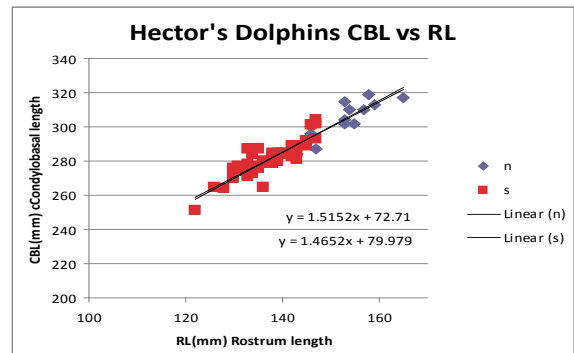
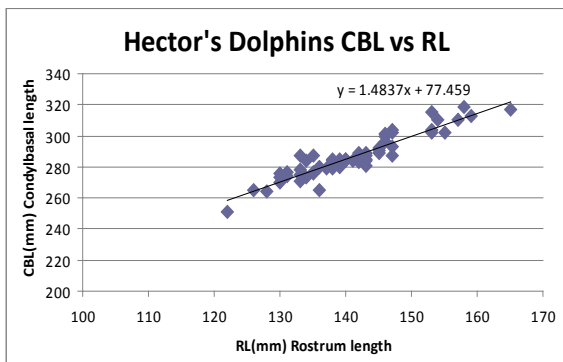
**Case 3: length/length**

Is there a relationship between RL (rostrum length) and CBL (condylobasal length)?

My predictor variable is RL and my response variable is CBL.



Again there is no reason to use anything other than a linear model.



(2)

**Relationships**

In all cases there is a positive relationship between the pairs of variables. For example in Case One as RWM increases RWB also tends to increase. I think this makes sense because for a bigger skull it is reasonable to have the variables increasing in some sort of proportion. It might not have happened, for example it is possible that, no matter what the RWB value is the dolphins all had nearly the same CBL, that the total length of the skull is the same no matter what the width was, but the graphs show that this is not the case.

In each case I think the relationships can be classified as strong ones. Most data points are close to the regression line. There are some points which are further away from the regression line, for example in Case One there is a dolphin with an RWB value of 50mm and an RWM value of 86mm but this is not likely to be unusual because it would be normal to have differences in measurements, there is no reason to be surprised that for an RWB value of 50mm there are variations in the RWM values.

In Case Three there appears to be two groups in the data, there is a gap for RL values at about 150mm. I initially thought that this was because of the North and South Island dolphins but then I noticed that while all of the higher values are for North Island dolphins there are also some North Island dolphins which overlap with some South Island dolphins. I do not think there is anything special about the gap, it is just that the data did not have any dolphins of this size.

The relationships for South Island dolphins do not appear to be very different from those I got using all of the data.

One reason for all the data points being close to the regression line is that some dolphins were removed from the data before the analysis began. Three North Island dolphins were removed because they were classified as outliers; two appeared to be dead South Island dolphins which had drifted onto North Island beaches. The third was from the Bay of Islands and it was unclear where it had come from. Also any dolphins which clearly were juveniles were removed from the data. For the analysis it seems reasonable to do this, the background of the dolphins were carefully checked before they were classified as outliers and removed from the data set.

Measurements for North Island dolphins are bigger than those for South Island dolphins. There is some overlap but not that much. (3)

## **Predictions**

### **Case 1.**

Predicting the RWB value if the RWM value is 55mm (I picked this because this could be either a North or South Island dolphin):

$y = 0.9039x + 32.07$  if all the data is used. RWB = 81.8mm.

$y = 0.9025x + 32.192$  if we used the South Island model. RWB = 82.6mm.

$y = 1.2x + 14$ . If we used the North Island model RWB = 80mm.

These predictions are not very different. I think the all data model could be used for predictions no matter what type of dolphin it is. There are only 12 North Island dolphins and this is not much for getting a reliable model. The all data model fits all the data well, no matter where the dolphin comes from.

### **Case 2.**

Predicting the CBL value if the RWB value is 85mm:

$y = 2.123x + 117.3$  if all the data is used. CBL = 298mm.

$y = 1.572x + 159.1$  if we used the South Island model. CBL = 293mm.

$y = 1.9967x + 131.9$  if we used the North Island model. CBL = 302mm.

These are not very different. When I looked at the all data model it appeared the data points for higher values (these are North Island dolphins) tend to be above the regression line and so it is possible that predictions for North Island dolphins with bigger measurements might be too low if the all data model is used in such a case.

### **Case 3.**

Predicting the CBL value if the RL value is 150mm:

$y = 1.4837x + 77.459$  if all the data is used. CBL = 300mm.

$y = 1.4652x + 79.979$  if we used the South Island model. CBL = 300mm.

$y = 1.5152x + 72.71$  if we used the North Island model. CBL = 300mm.

All the same. As you can see from the graphs there is very little difference in the models. So it might be best to use the all data model in all cases. (4)

### **Conclusion**

It is not that easy to decide if the classification of North Island dolphins as a new subspecies is supported by my investigation. There were only 12 North Island dolphins and so this makes any conclusions hard to be sure about. While it is easy to say we need more data it might not be that easy in real life. You need to have access to the dolphins and that may not be easy. By visiting different museums the researchers probably got all the data they could lay their hands on.

One thing that stands out from the graphs is that North Island dolphins tend to be bigger than South Island ones so if the classification depends on size then it can be clearly supported.

If the classification depends on other relationships then it is not that clear but the final classification was determined by using genetic evidence, the differences in measurements and knowledge of an absence of gene flow between populations. It goes well beyond an investigation into relationships. (5)

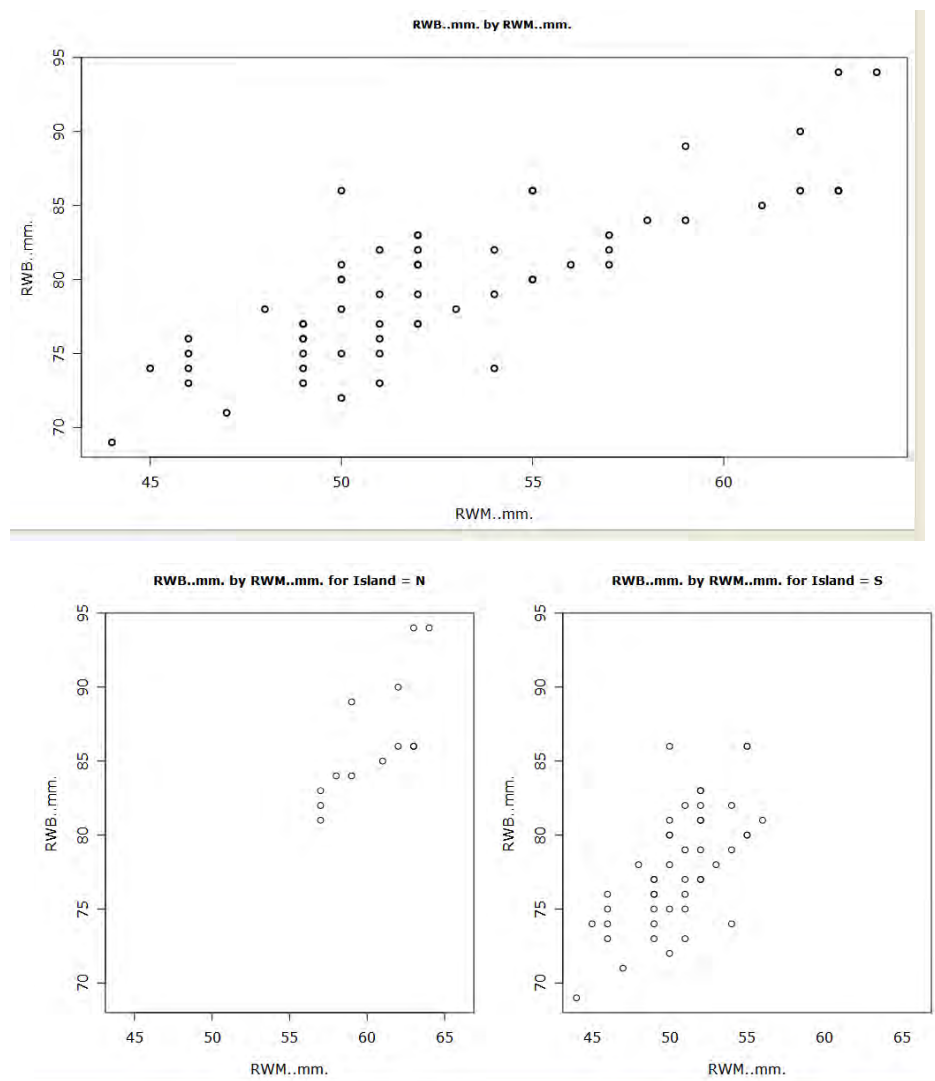
|           |  |
|-----------|--|
|           | <p>Grade Boundary: High Merit</p>  |
| <p>2.</p> | <p>For Merit the student is required to investigate bivariate measurement data with justification. This involves linking components of the statistical enquiry cycle to the context, and referring to evidence such as statistics, data values, trends, or features of visual displays in support of statements made.</p> <p>An appropriate relationship question has been posed. There is some contextual reflection on the nature of the data which shows the scientific report has been considered (1).</p> <p>Appropriate displays have been selected and used for two different pairs of variables (2).</p> <p>Features of the data have been identified and the nature and strength of the relationship has been described in context. There is contextual discussion on features and the relationship; while these are sensible they do not refer to possible reasons which are identified in the scientific report (3).</p> <p>A linear model has been found and its appropriateness has been considered. The model has been used to make predictions. There is contextual discussion which justifies the predictions (4).</p> <p>Findings have been communicated in a conclusion. There has been contextual consideration but this is not referenced to aspects of the scientific report (5).</p> <p>There is evidence of investigating bivariate measurement data with justification in the provision of contextual evidence to support statements for components of the statistical enquiry cycle.</p> <p>To reach Excellence a complete solution would have supported more aspects of the investigation with reflective comments demonstrating a contextual understanding consistent with the scientific report.</p> |

I am investigating if the classification of the North Island Hector's Dolphins as a new subspecies can be supported by looking at relationships.

I will use two sets of variables.

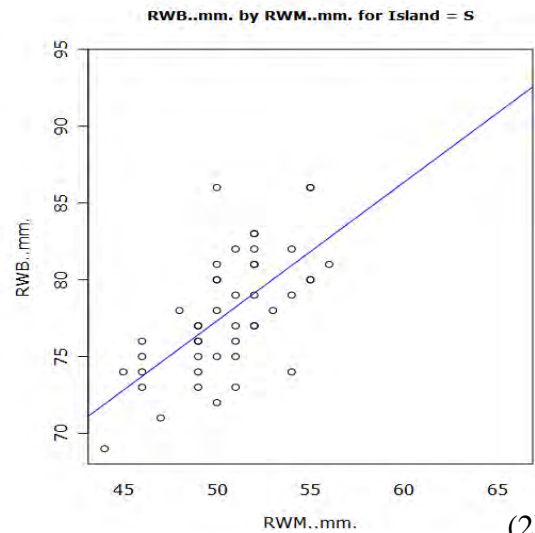
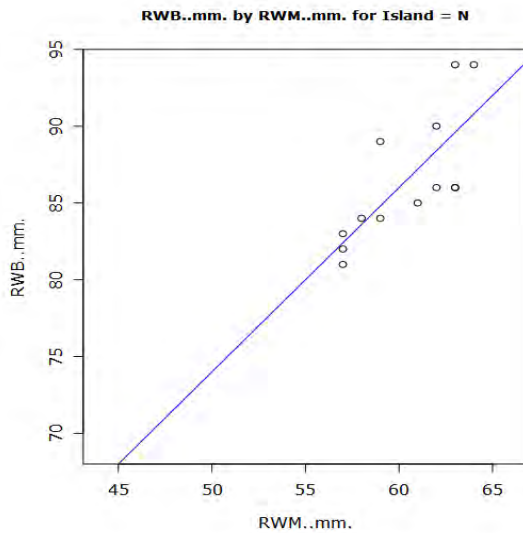
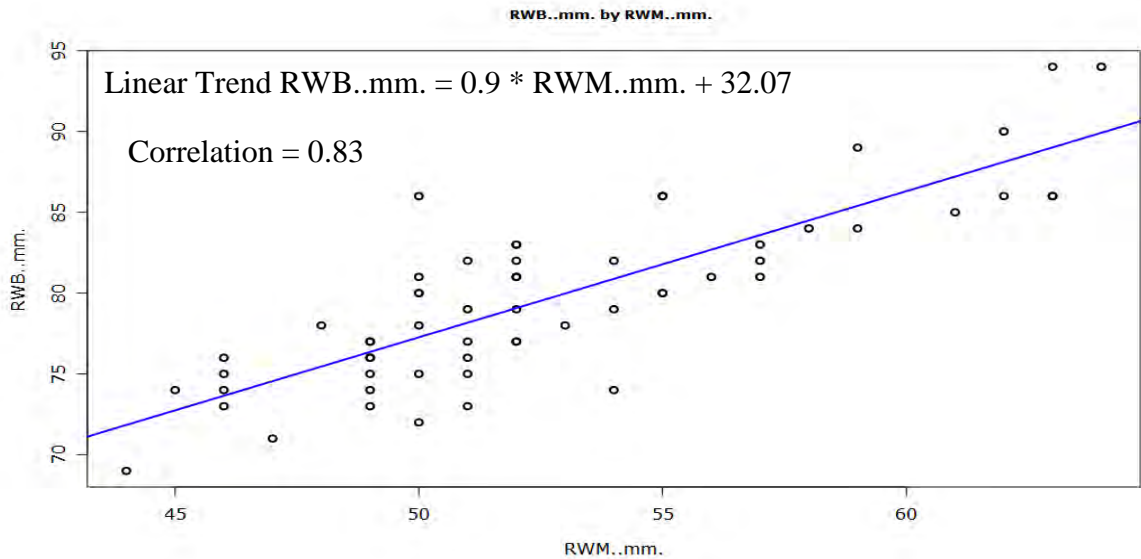
1. Is there a relationship between RWM (rostrum width at midlength) and RWB (rostrum width at base)? This will look for a relationship between widths.
2. Is there a relationship between RL (rostrum length) and CBL (condylobasal length)? This will look for a relationship between lengths.

Measurements were taken from skeletal material in different museums. These might have come into the museums at different times but I do not think dolphin measurements would have changed that much. This method must have been used because you cannot get measurements from live dolphins. (1)



It is clear that a straight line would be the best model if all of the data was used. It is not so clear when the islands are separated but there is nothing to say a non-linear model would be better.





(2)

Summary for Island = N

$$\text{RWB..mm.} = 1.2 * \text{RWM..mm.} + 14$$

Correlation = 0.76

Summary for Island = S

$$\text{RWB..mm.} = 0.9 * \text{RWM..mm.} + 32.19$$

Correlation = 0.65

There is a positive linear relationship between RWM and RWB. This is shown by the positive gradients of the regression lines. Dolphins with greater midlength rostrum widths tend to have longer base rostrum widths. This would make sense because if the rostrum is wider at the base you would also expect it to be wider at the midlength. The gradients of the regression lines are similar, it is a bit steeper for the North Island.

Points are scattered along the regression line with no obvious gaps and there is not a lot of deviation from it. This is likely to be because some outliers were removed from the data. While there is a small difference in correlation coefficients between the North and South Islands (the NI one is higher and this suggests slightly less scatter) when I look at the graphs the degree of scatter is not noticeably different.

The main thing about the graphs is that the North Island dolphins are at the higher end of the scale. Their measurements tend to be greater than for South Island dolphins. This could be

because North Island water temperature is usually higher than in the South Island or because the food is different in different parts of the country. (3)

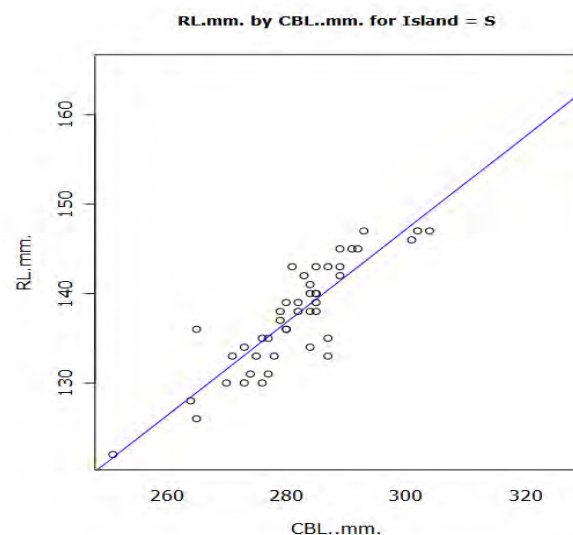
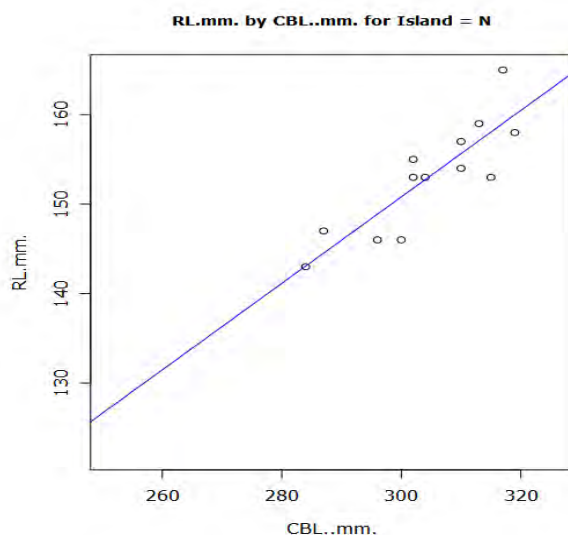
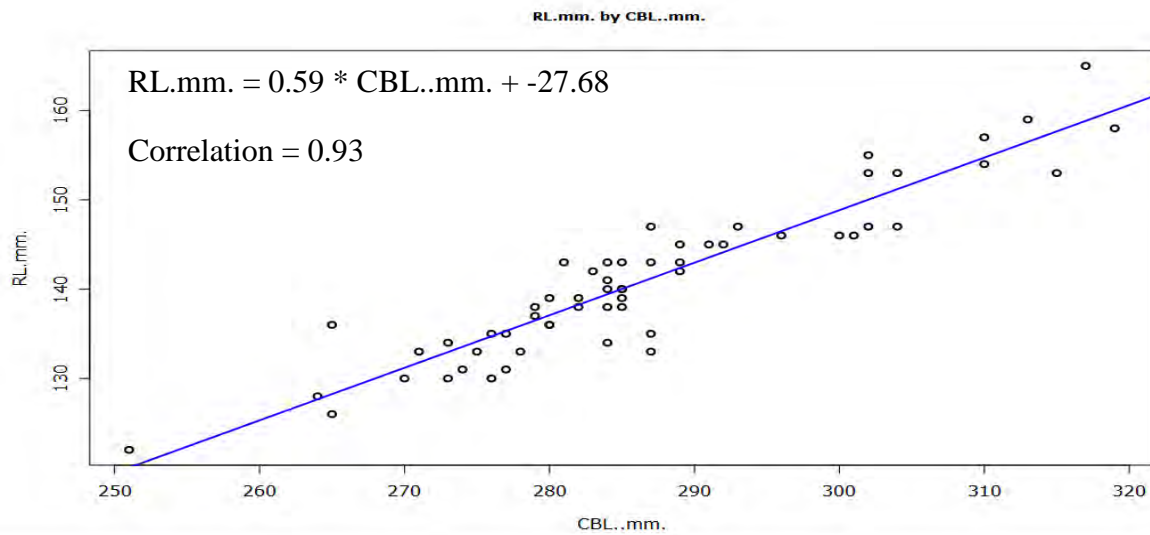
The small number of North Island dolphins means I am not so confident about predicting for them but I will use the NI equation because it is a bit different from the overall one.

For a North Island dolphin with an RWM of 64mm I would estimate the RWB to be 91mm.

Because there is not much difference in the overall equation and the South Island equation it would not matter which of these I use but there is enough data and the correlation is strong enough to be confident using the SI equation.

For a South Island dolphin with an RWM of 48mm I would estimate the RWB to be 75mm.

Because of the strength of the relationships I think these estimates would be quite good. If the NI were not a new sub-species then I would have to use the overall equation. For 64mm the estimate would be 90mm which is 1mm less and the estimate for 48mm would be 75mm which is the same. So it would not matter too much if you used the overall equation. (4)



Summary for Island = N

$$\text{RL.mm.} = 0.48 * \text{CBL..mm.} + 5.72$$

Correlation = 0.86

Summary for Island = S

$$\text{RL.mm.} = 0.52 * \text{CBL..mm.} + -9.08$$

Correlation = 0.87

These results are similar to the width ones. The positive gradients show that dolphins with greater condylobasal lengths tend to have greater rostrum lengths and again this makes sense in terms of the dolphins. There is not too much difference in the slopes of the regression lines, the line using all the data is a bit steeper.

Again points are scattered along the regression line with no gaps. Data points are close to the line and there is not a lot of difference in correlation coefficients.

North Island dolphins tend to have greater measurements, the reason would be the same as before. (3)

The North Island equation estimates a dolphin with a CBL of 300mm to have an RL of 150mm (the overall equation gives 149mm).

The South Island equation estimates a dolphin with a CBL of 300mm to have an RL of 147mm (the overall equation gives 149mm).

Again not a lot of difference. (4)

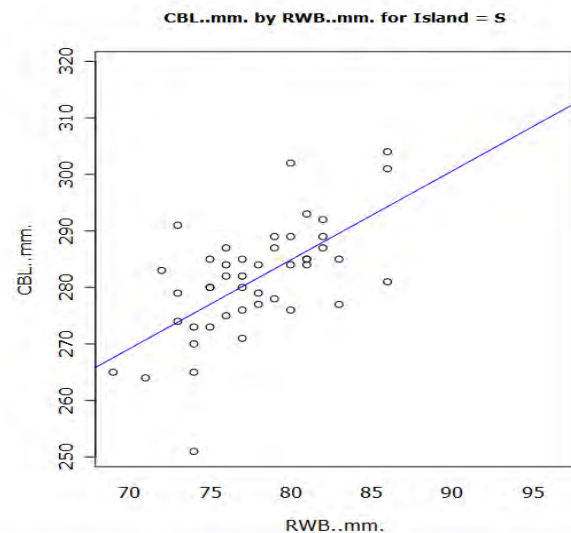
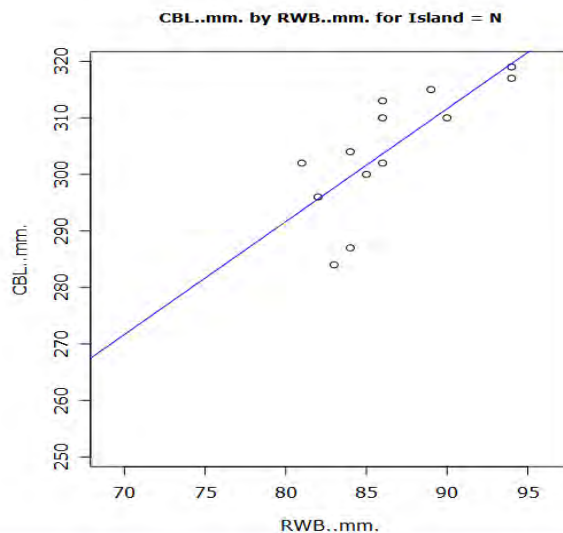
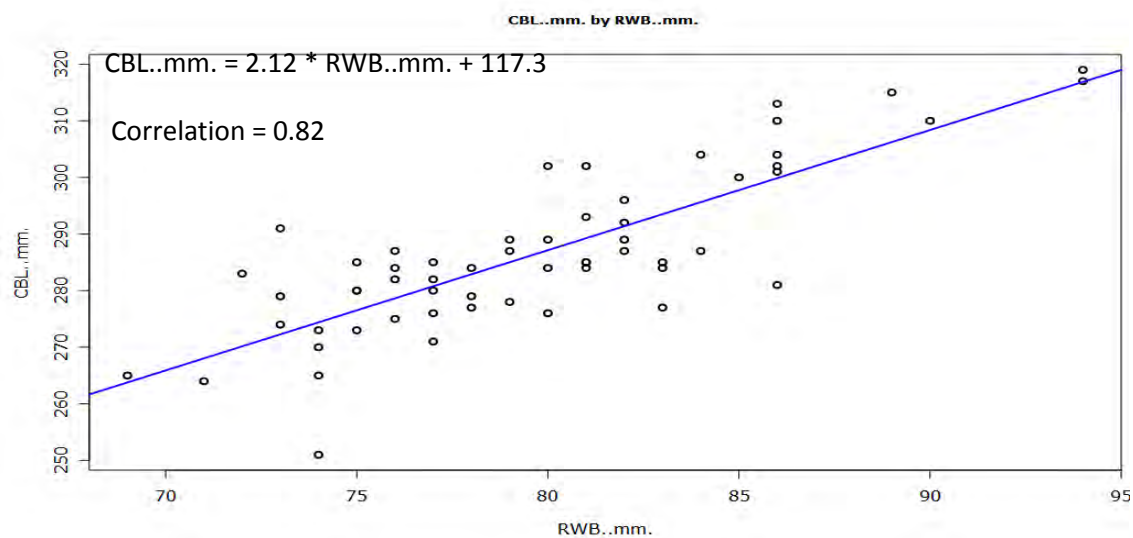
While the scientific research shows that the North Island dolphins can be classified as a new sub-species this relationship investigation does not clearly support it. The relationships between North and South Island are not that different and the estimates are similar if you use either the Island equation or the overall one. The reasons must be other than what I have done; for example the size and colouring of the dolphins could be different. (5)

|    |  |
|----|--|
|    | Grade Boundary: Low Merit  |
| 3. | <p>For Merit the student is required to investigate bivariate measurement data with justification. This involves linking components of the statistical enquiry cycle to the context, and referring to evidence such as statistics, data values, trends, or features of visual displays in support of statements made.</p> <p>A relationship question has been implied by the purpose statement. The variable selection has been justified and there is contextual consideration of the data source (1).</p> <p>Appropriate displays have been selected and used (2).</p> <p>Features of the data have been identified and the nature and strength of the relationship has been described in context. Some contextual reasons have been used to justify comments (3).</p> <p>A linear model has been found, its appropriateness has been considered, and it has been used to make predictions. There is a contextual comment about the predictions (4). Findings have been communicated in a conclusion (5).</p> <p>There is evidence of investigating bivariate measurement data with justification in the provision of contextual evidence to support statements for components of the statistical enquiry cycle.</p> <p>For a more secure Merit there would be more contextual discussion on key aspects, for example the small number of North Island dolphins has not been commented on.</p> |

I am investigating if the classification of the North Island Hector's Dolphins as a new subspecies can be supported by looking at relationships from the data set. I will investigate a possible relationship between RWB (rostrum width at base) and CBL (condylobasal length). I chose these two variables because I think it is pretty obvious that there will be relationships when both variables are widths or lengths.

Measurements were taken from skeletal material in different museums. The skeletal material came from dead dolphins which were washed ashore. (1)

Before I fitted the regression lines to the data I looked at the plots on the computer. It did not look like a model other than a straight line one should be used.



(2)

Summary for Island = N

CBL..mm. = 2 \* RWB..mm. + 131.9

Correlation = 0.76

Summary for Island = S

CBL..mm. = 1.57 \* RWB..mm. + 159.1

Correlation = 0.63

For both North and South Island dolphins those with greater midlength rostrum widths at base tend to have longer condylobasal lengths. This would make sense because if the rostrum is wider at the base you would expect the condylobasal length to also be longer. You would expect the head dimensions to be in some sort of proportion. The upward slope of the regression lines shows that the relationships are positive. The slope for North Island dolphins is greater than for South Island ones so this might be a difference in the relationships.

The relationships can be said to be strong because points are close to the regression lines and are scattered all along the regression line with no gaps. There is a small difference in correlation coefficients between the North and South Islands (the NI one is higher and this suggests slightly less scatter) and it does look like there is a bit more scatter for the South Island.

The main thing about the graphs is that the North Island dolphins have measurements which tend to be greater than for South Island dolphins. This could be because of the age or gender of the dolphins. (3)

I will make predictions for a dolphin which has an RWB value of 85mm. My reason for picking this value is that such a dolphin might be from either island so I can compare my answers.

North Island dolphin estimate is 302mm.

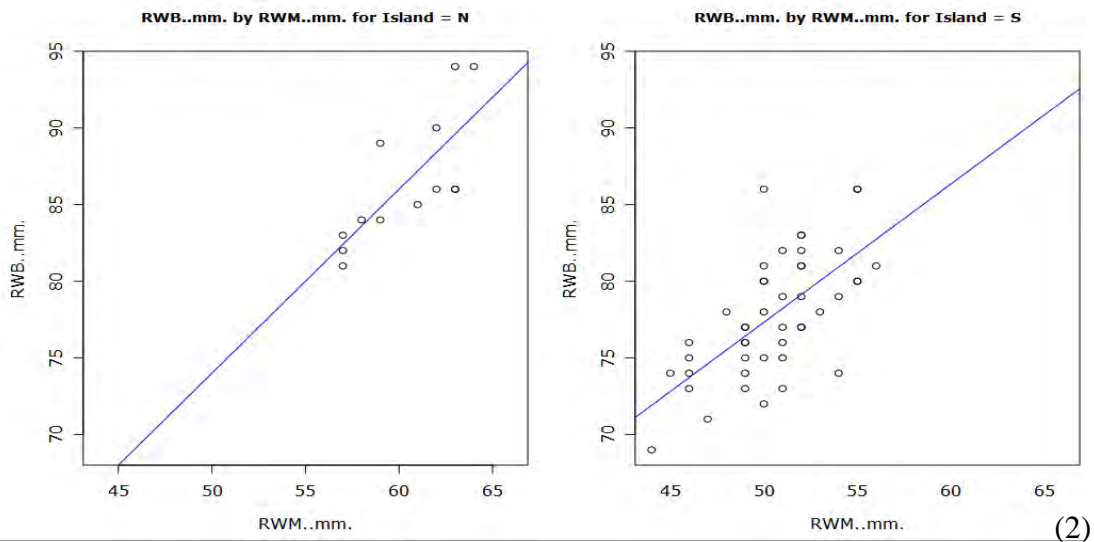
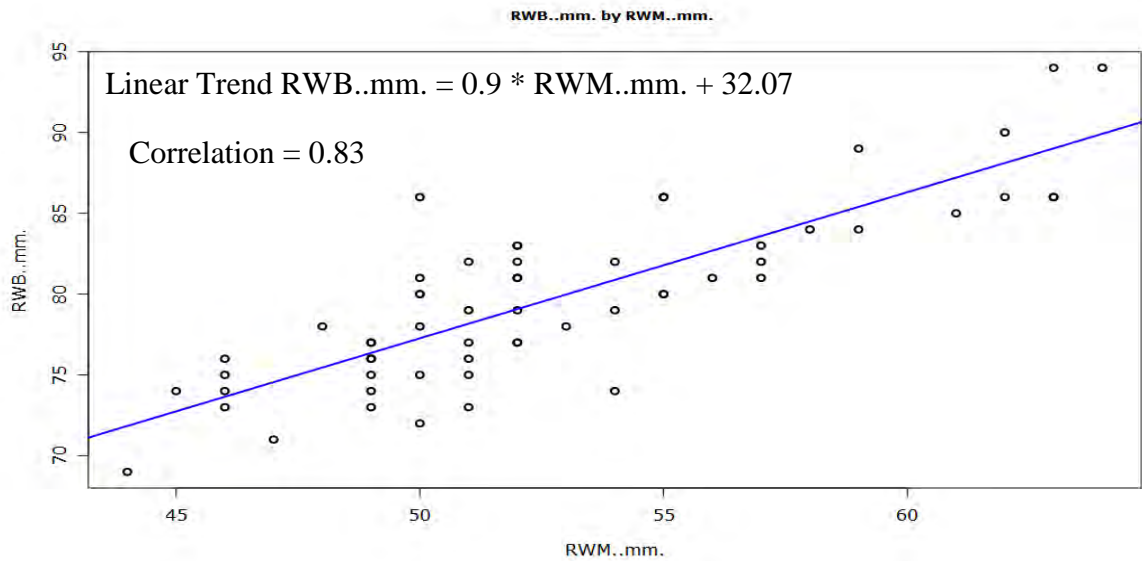
South Island estimate is 293mm.

The South Island estimate is a bit lower but this is consistent with the South Island dolphins being smaller. (4)

If relationships were used to support the classification then the reason would be the different slopes of two regression lines and also differences in scatter. But I think more research should be done if this was to be used, the results are not that clear. (5)

|           |  |
|-----------|--|
|           | <p>Grade Boundary: High Achieved</p>   |
| <p>4.</p> | <p>For Achieved the student is required to investigate bivariate measurement data. This involves showing evidence of using each component of the statistical enquiry cycle.</p> <p>An appropriate relationship question has been posed. Variables have been identified and the explanatory and response implied. A purpose which relates to the classification of the North Island population as a new sub-species has been identified (1).</p> <p>Appropriate displays have been selected and used (2).</p> <p>Features of the data have been identified and the nature and strength of the relationship has been described in context. One possible contextual reason has been given for a difference of the plots for the two Islands, but this is limited in depth (3).</p> <p>Appropriate models have been found and have been used to make predictions. There is some contextual discussion about the predictions (4).</p> <p>Findings have been communicated in a conclusion. There is some contextual reasoning about the classification, but it is not supported with the reasons which can be found in the scientific report (5).</p> <p>There is evidence of investigating bivariate measurement data in the use of each component of the statistical enquiry cycle.</p> <p>To reach Merit the student would need to provide more in the way of contextual references to evidence which support statements which have been made. There would need to be some evidence that the scientific report has been considered.</p> |

I wonder if relationships from the data base can support the classification of the North Island Hector's Dolphins as a new sub-species. I will use rostrum midlength width as my explanatory variable and rostrum width at base as my response variable. (1)



Summary for Island = N  
 $RWB..mm. = 1.2 * RWM..mm. + 14$   
 Correlation = 0.76

Summary for Island = S  
 $RWB..mm. = 0.9 * RWM..mm. + 32.19$   
 Correlation = 0.65

There is a positive linear relationship between RWM and RWB because the regression lines slope upwards and the correlation coefficients are positive. The line for North Island dolphins is steeper than for South Island dolphins but I cannot be sure that this is a true result or if it is because of the small number of North Island dolphins. The graphs show that dolphins with greater midlength rostrum widths tend to have longer base rostrum widths.

Both relationships are strong. The straight lines fit the data well and all of the points are close to the lines. The correlation coefficient for North Island dolphins is a bit higher than for South Island dolphins and means the relationship for the NI could be stronger than the SI. There are some SI points which appear to be a bit further away from the regression line but like I said before the small amount of NI data means this might not be true.



Other than the strengths and slopes there are no patterns in the data and the straight line seems to be a good fit.

North Island dolphins are bigger than South Island ones because the NI point are mostly to the right of the South Island ones. This could be because they have better food in the North Island. (3)

For a North Island dolphin with an RWM value of 50mm I predict the RWB value to be 74mm.

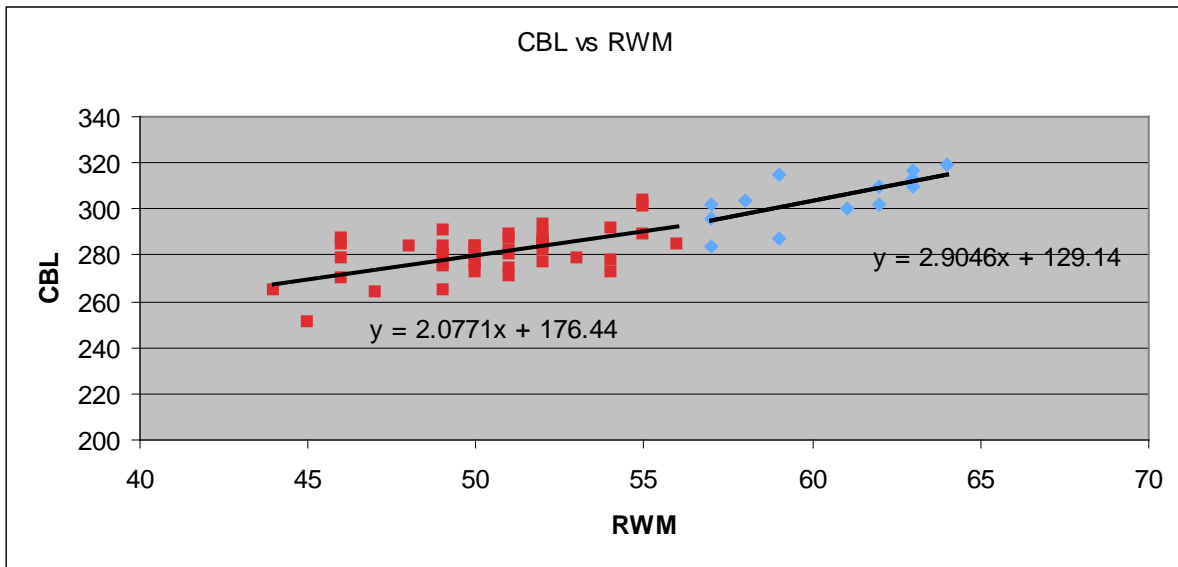
For a South Island dolphin with an RWM value of 50mm I predict the RWB value to be 77mm.

I think these predictions would not be too bad because the relationships are strong but an RWM of 50mm is not in the data range for a North Island dolphin and it is possible that they do not go this small. (4)

The classification of North Island dolphins as a separate sub-species is not supported by my relationships. They are both positive and have about the same strength. Size is what makes them different but that is not a relationship. Also the difference in size might just be because of their ages so that might not be a factor either. (5)

|    |  |
|----|--|
|    | Grade Boundary: Low Achieved   |
| 5. | <p>For Achieved the student is required to investigate bivariate measurement data. This involves showing evidence of using each component of the statistical enquiry cycle.</p> <p>An appropriate relationship question has been implied. Variables have been identified and the explanatory and response implied. A purpose which relates to the classification of the North Island population as a new sub-species has been identified (1).</p> <p>Appropriate displays have been selected and used (2).</p> <p>Features of the data have been identified and the nature and strength of the relationship has been described in context. The description of the direction and comments about features are limited in terms of contextual understanding (3).</p> <p>An appropriate model has been found and it has been used to make predictions. There is limited discussion about the predictions (4).</p> <p>Findings have been communicated in a conclusion. There is a basic link between the relationship and the classification of North Island dolphins as a new sub-species (5).</p> <p>There is evidence of investigating bivariate measurement data in the use of each component of the statistical enquiry cycle.</p> <p>For a more secure Achieved there would have been more depth in the description of the relationship and features of the display, particularly with respect relating descriptions to the context of the investigation.</p> |

I will investigate if there is a relationship between RWM and CBL and then will try to see if this helps show that North Island Hector's Dolphins are different from South Island ones. (1)



(2)

In general there is a relationship and it is that the longer the RWM (rostrum length at midlength), the longer the CBL (condylobasal length). This applies to both the North and South Island dolphins.

Both relationships are strong. The straight lines fit the data well and all of the points are close to the lines. There are no patterns in the data and it is evenly spread over all the RWM values. While the two graphs look quite similar there is a lot less data for the North Island so I would not be so sure about these results. The slope of the North Island line is steeper than the South Island line. The RWM values for the North Island dolphins are all bigger than the RWM values for South Island dolphins. The CBL values overlap a bit but in general North Island ones are bigger. (3)

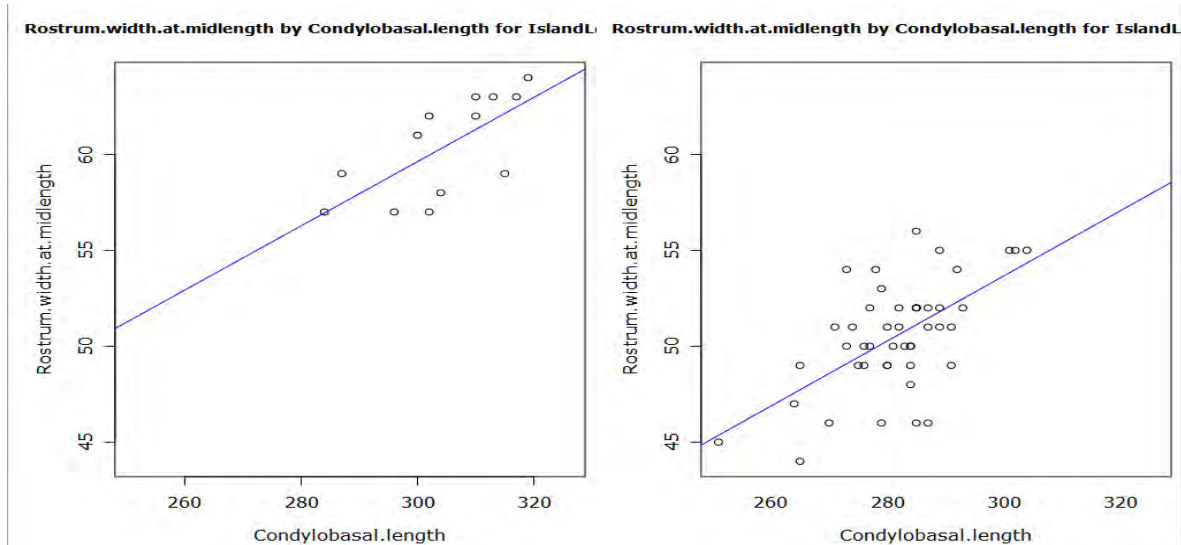
For an RWM value of 50mm I predict the CBL value to be 280mm. This would be for a South Island dolphin.

For an RWM value of 60mm I predict the CBL value to be 303mm. This would be for a North Island dolphin. (4)

The actual relationships are not that different because they look about the same for scatter and they both slope up. The difference is that the points for North and South Island are separated so that could be why the North Island dolphins are a new sub-species not because of the relationships so much. (5)

|           |  |
|-----------|--|
|           | <p>Grade Boundary: High Not Achieved</p>   |
| <p>6.</p> | <p>For Achieved the student is required to investigate bivariate measurement data. This involves showing evidence of using each component of the statistical enquiry cycle.</p> <p>An appropriate relationship question has been posed. Variables have been identified and the explanatory and response implied. A purpose which relates to the classification of the North Island population as a new sub-species has been identified (1).</p> <p>Appropriate displays have been selected and used (2).</p> <p>Some features of the data have been identified and the nature of the relationship has been described in context. The strength of the relationships relies on the value of the correlation coefficients and does not refer to visual aspects of scatter about the regression line (3).</p> <p>An appropriate model has been found and it has been used to make a prediction. Predictions should have been rounded to a whole number of millimetres (4).</p> <p>Findings have been communicated in a conclusion. There is a basic link between the relationship and the classification of North Island dolphins as a new sub-species (5).</p> <p>Requirements for Achieved have not been met because the description of the strength of the relationship needs to refer to visual aspects of scatter about the regression line, rather than just relying on the value of the correlation coefficient.</p> <p>To meet Achieved the student would need to describe the strength of the relationship by discussing visual aspects of scatter about the regression line in context and in more detail.</p> |

Is there a relationship between CBL and RWM? Does this help with classifying North Island dolphins as a new sub-species? (1)



(2)

Summary for Island Letter = S

Linear Trend

$$\text{Rostrum.width.at.midlength} = 0.17 * \text{Condylbasal.length} + 2.81$$

Correlation = 0.59

Summary for Island Letter = N

Linear Trend

$$\text{Rostrum.width.at.midlength} = 0.17 * \text{Condylbasal.length} + 9.4$$

Correlation = 0.7

The relationship for both graphs is that dolphins with longer condylbasal lengths tend to have longer rostrum widths at mid-length.

Both relationships are strong because of the correlation values but the North Island relationship is stronger because it has a higher value of 0.7.

Both regression lines have the same slope and there are no outliers or groupings. There are fewer points for the North Island dolphins so results may not be so reliable. North Island dolphins are higher up the graph than South Island ones. (3)

For a condylbasal length of 290mm I predict the rostrum width at mid-length to be 52.11mm for a NI dolphin and 58.7mm for a South Island dolphin. (4)

The difference in the relationships is that the South Island one is not so strong but the classification might be because of the higher North Island values. Also since the slopes of the lines are the same I do not think the relationship can be used to decide on a new sub-species. (5)