



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



MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga

2006

Internal Assessment Resource

Subject Reference: **Statistics and Modelling 3.5**

Resource Title: **“Animal Antics”**

Achievement standard: **90645 version 2**

Standard title: *Select and analyse continuous bi-variate data*

Credit: 3

This resource has been trialled in a school and includes annotated examples of assessed student work. There are eight documents in this resource:

Task and schedule	<input type="checkbox"/>	Student 1 EXCELLENCE	<input type="checkbox"/>
Assessment guidelines	<input type="checkbox"/>	Student 2 MERIT	<input checked="" type="checkbox"/>
Teaching notes	<input type="checkbox"/>	Student 3 ACHIEVED	<input type="checkbox"/>
		Student 4 ACHIEVED	<input type="checkbox"/>
		Student 5 NOT ACHIEVED	<input type="checkbox"/>

Internal assessment resource reference number:
Statistics and Modelling/3/5 X version 1

Date version published:
Ministry of Education
Quality assurance status

July 2006
For use in internal assessment
from 2006

EXAMPLE OF ASSESSED STUDENT WORK

ASSESSMENT COVER SHEET FOR STUDENT 2

MERIT

	Achievement Criteria	Code	Evidence	Judgement (refer to Assessment Schedule for judgement statements)	Sufficiency
Achievement	Select and analyse bi-variate continuous data	A	Purpose stated.	✓	All four of code A
		A	Scatterplot drawn.	✓	
		A	Regression line obtained.	✓	
		A	Relationship explained in context.	✓	
Achievement with Merit	Carry out an in-depth analysis of bi-variate data	M	Relationship between two pairs compared with explanation.		Achievement plus THREE of code M
		M	Regression equations used to obtain predictions.	✓	
		M	Appropriateness of regression model(s) discussed.	✓	
		M	R ² values interpreted correctly.		
		M	Difference between correlation and causality explained.	✓	
Achievement with Excellence	Report on the validity of the analysis	E	Assumptions about the data stated.	✓	Merit plus THREE of code E
		E	Limitations of the model given.		
		E	Piecewise or other models proposed and justified, and/or outliers identified and an approach to dealing with them suggested.		
		E	Relevance and usefulness of the evidence explained.		
		E	Applicability of findings stated.		

Final Grade Awarded

N		A		M	✓	E	
----------	--	----------	--	----------	---	----------	--

EXAMPLE OF ASSESSED STUDENT WORK

STUDENT 2

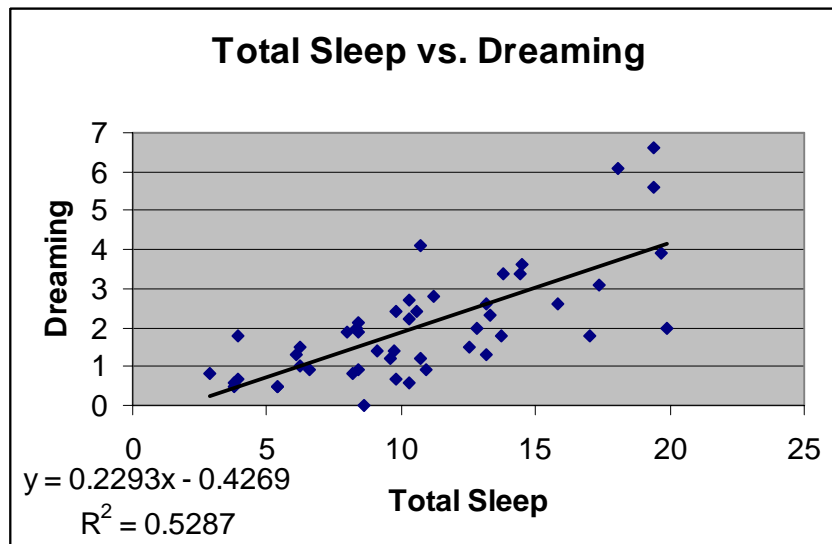
MERIT

Purpose

The purpose of this investigation is to use regression to investigate the relationship between:

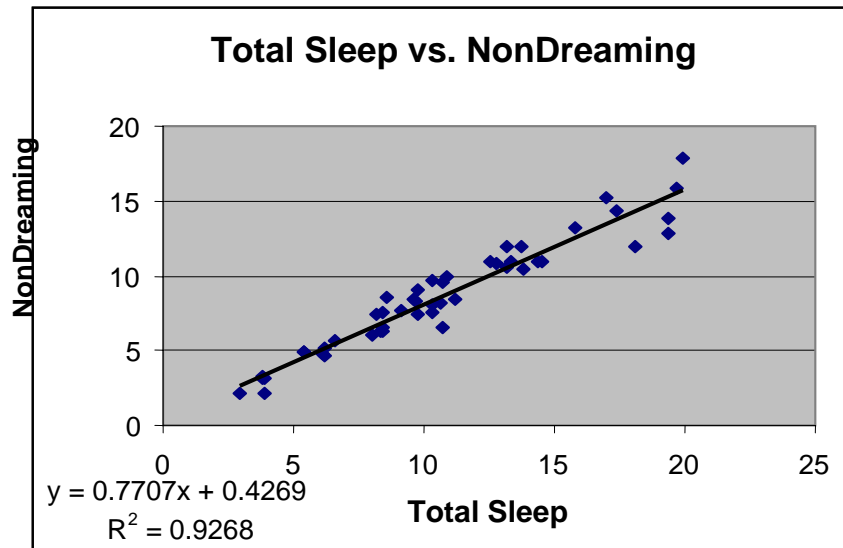
- Total Sleep vs. Dreaming
 - Total Sleep vs. Non dreaming
- And to compare the two.

Purpose is stated. The predictor/explanatory and predicted/response variables are implied on the graph. The purpose, as worded, pre-supposes a relationship exists and would be better worded "To determine if there is any relationship between , and if any relationship is found to investigate its nature". **A**



The strong R2 value shows us that the relationship between Dreaming and Total Sleep is strong and that as the amount of sleep increases the amount of dreaming sleep also increases.

Scatter graph drawn and regression line found. The relationship is not interpreted in context. This is an ideal situation to obtain further clarifying evidence through conferencing. **A**



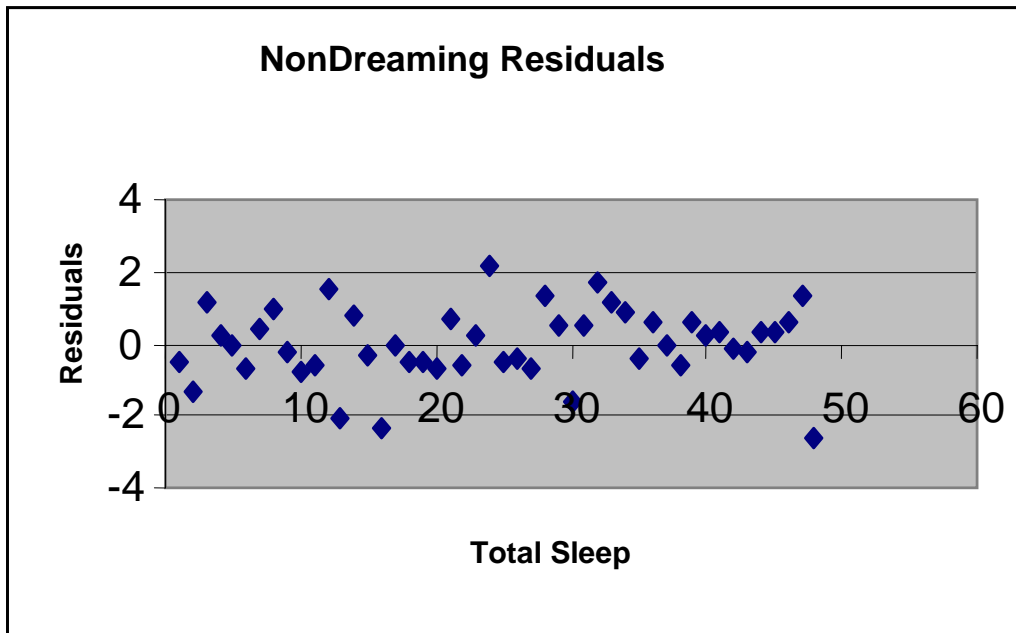
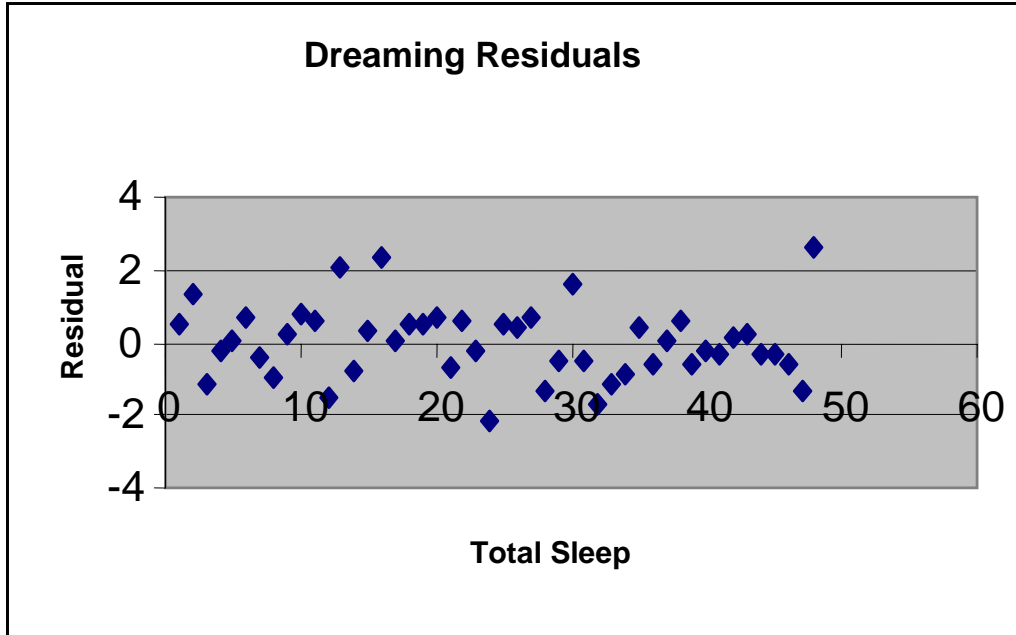
The very strong R2 value in this data shows that the relationship between NonDreaming and Total Sleep is very strong and as the amount of Total Sleep increases the amount of NonDreaming sleep increases at a similar rate.

This is incorrect - the rates are different. The statement probably arises from a visual inspection of the graphs on which the gradients look the same; however the scales are different and they have different numerical values.

Comparing the Relationship between variables

The linear relationship with Total Sleep and Dreaming is strong, but is weaker than the relationship between Total Sleep and NonDreaming. You can tell this because the R2 Value for Total Sleep vs. NonDreaming is 0.9268 which is a lot higher than the R2 Value for Total Sleep vs. Dreaming which is 0.5287. This is confirmed by the scatter graph, by the points on the NonDreaming Graph being a lot closer to the regression line than in the Dreaming graph.

Residuals



You can see from the Residual graphs above that there is no pattern in the data and so the Regression line is an accurate measure of the data.

Relationships compared, with explanation. However the residual plot for total sleep is incorrect, and there is insufficient evidence to make a contribution towards Merit.

Predictions

Total Sleep vs. Dreaming

Total sleep	Dreaming	Predicted dreaming
14.5	3.6	2.89795
8.4	1.9	1.49922

This shows that if an animal gets 14.5 hours of Total Sleep then this animal's Predicted Dreaming Sleep is 2.89795 hours but in actual fact it got 3.6 hours of Dreaming Sleep.

I got these results from using this formula:

$$\text{Dreaming Sleep} = 0.2293 * \text{Total Sleep} - 0.4269$$

Total Sleep vs. NonDreaming

Total sleep	Non-dreaming	Predicted non-dreaming
14.5	10.9	11.60205
8.4	6.5	6.90078

This shows that if an animal gets 14.5 hours of Total Sleep then this animal's Predicted NonDreaming Sleep is 11.60205 hours but in actual fact it had 10.9 hours of NonDreaming Sleep.

I got these results from using this formula:

$$\text{NonDreaming Sleep} = 0.7707 * \text{Total Sleep} + 0.4269$$

Predictions made. It is unwise to compare predicted values and data values for isolated cases as the cases chosen could be atypical.

M

How Appropriate the Regression Model is

I do not think that the Total Sleep vs. Dreaming graph is as appropriate for analysing data as the Total Sleep vs. NonDreaming graph. I feel that the Total Sleep vs. NonDreaming graph is a more accurate description of the data because of the very strong R2 value and so the points are closer to the regression line.

I noticed that the data ended after 20 hours which I think is a good representation for I do not think that many animals would sleep for more much or else they would only be awake for a few hours per day, and this would not be enough time to look for food, breed and defend their territory.

There is marginal evidence for the limitations aspect of Excellence.

So both regression models are fine for measuring interpolation for the data is mainly clumped in the middle anyway and shows a good measure of it. And the extrapolation for the regression models are quite accurate as well for the

lowest mark is 2.6hours and the highest is 19.7hours and so gives quite a wide range for the data to be within.

Appropriateness of models discussed well. Interpolation discussed appropriately. Extrapolation is not well understood. The discussion would be stronger if there had been some comment that the predictions are less reliable for larger values of total sleep time as the points on the scatter graph are more dispersed for large values than for small values. **M**

Causality vs. Correlation

Causality is a factor that makes the data change (causes it to change) such as if an animal has more sleep it will have more Dreaming and or NonDreaming sleep, where as correlation is a relationship with the data, but this may just be a coincidence to the data and it doesn't actually affect it as such, such as if the animal is hungry when it goes to sleep it will dream about food.

So I think that all of my data does have causal relationships from the research I have done that proves that the more Total Sleep you get the more Dreaming Sleep you can have up until your body has replenished itself, and also with the NonDreaming Sleep as well. I.e. You cannot have 3 hours of dreaming sleep if you only have 2 ½ hours total sleep.

Marginal comment on causality. **M**

Assumptions made

In the experiment there were some assumptions made:

- That each animal was in the same environment with the same amount of Predators.
- All the research was done at the same time of year – such as in winter the animals may sleep longer.
- That the animals were of the same age.
- That these animals were not in hibernation.

Appropriate comments. **E**

Limitations

There were also a few limitations to this data:

- The data was collected in 1976 and so their equipment is outdated.
- In the data we were given we were shown the sample mean (average) value of each of the species sampled and so do not get the exact data for each species.
- Some of the data had to be deleted because certain results were unavailable (NA) and so certain species were missed out all together.
- We do not know if these animals were in the same environment, such as some may have been more in danger than the others and so didn't have as much Total Sleep.
- Also we do not know how stressed out each animal was from the attention/experiments that the scientist was doing on them so this may have affected there sleep and dreaming patterns.

- We do not know if the Total Sleep for certain animals was quick naps, or long periods at a time, this makes it a limitation because to be able to dream animals have to be in a certain state of rest i.e. be asleep for a certain period of time and so for those animals who just had quick naps they wouldn't have dreamed as much.
- The scientist only measured the brain activity in these animals and so would not know for sure if they were dreaming or not for certain animals may have slower brain waves than others.
- The fact that in this investigation in my findings I did not take into account the other variables such as Brain Weight, Gestation or even the particular animals the sleeping pattern was based on and so this may have affected my data otherwise.
- They did not sample every mammal and so does not give an accurate account of all mammals sleeping and dreaming patterns.

Relevance/usefulness of evidence

This information is very useful for scientists in the study of the human brain and sleeping disorders/problems for they can learn things from animals that they aren't allowed to experiment on in humans. It would also help them to find out what our ancestors sleep patterns would have been like before everyday modern life now with its stresses and schedules.

Then again this data was collected 30 years ago and so this may have been outdated because of the technology used back then and so it may not be as useful and the modern day experiments of the same scale.

These three sections show some good thinking and are sufficient to contribute to the assumptions aspect for Excellence. **E**

The limitations do not refer to the model.

The relevance and usefulness section could be the starting point for an appropriate answer but are not developed sufficiently.