YEAR 7

ANIMALS IN THE RAIN FOREST

DATABASE PROJECT

Objectives

This project introduced Year 7 to some of the practical decisions which have to be made when setting up computer databases for analysing "real world" data.

They were already studying the Rain Forest, so we decided to create an IT project which would help them understand about the different types of animals which lived in the Rain Forest regions.

The project involved:

- collecting data about animals that live in the rain forest,
- deciding what categories of information were most significant,
- designing a simple database to hold this information efficiently
- entering the data and creating display graphs and tables to summarize the results.

Year 7 were encouraged to think about what might be the most suitable program to present these results, and a number of constructive suggestions (database package, word processor, cardfile etc.) were made.

In practice, we used the Excel spreadsheet. This is a very good tool to help understand some of the underlying principles of data organisation.

In a commercial relational database, data is organised into logical "tables" of "rows and columns", which can be easily understood by looking at the grid structure in Excel.

They were also asked to imagine that they had far greater quantities of data (e.g. millions of diffent bits of information, instead of the hundreds which they were able to collect), and how this might affect the way they would organise it.

The Project had four sessions (explained below):

1. Planning the Project
2. Organising and Collecting Data
3. Input and Presentation
4. Feedback

1. Planning the Project

When you want to understand lots of data (for example, information about a group of children), you can organise it into *categories*, e.g...

- how many boys and how many girls ?
- how many have black hair and how many have fair hair ?
- how many like football, or computers etc. ?

Similarly, with Animals in the Rain Forest, we could use many different categories, e.g.

- where does each animal live (the emergent layer, canopy, understory or floor) ?
- what type of animal is it (mammal/reptile/bird etc.) ?
- what type of food does it eat (is it carnivorous, herbivorous or omnivorous) ?
- what colour is it (brightly coloured or camouflaged) ?
- is it solitary, or does it live in groups ?
- how does it give birth to it's young ?
- does it sleep during the day or night ?
- how big is it ?

We had to decide what categories were most useful to help understand the animals in the rain forest.

We organised the Year into teams of between 4 and 6.

Each team had to decide *what categories of data to research*, how they would *collect, organise* and *present* the results, and who would co-ordinate the team

2. Organising and Collecting Data

Pie Charts, Bar Charts, Graphs etc are used in many different types of books and periodicals. For example:

- Advertising (to show how cost-effective things are),
- Computer Magazines (to show which products are best),
- Newspapers and Magazines (we even had an example showing what type of people attended a party).
- in Geography to analyse data about our world

So the concept of entering data into a computer and producing charts is not confined just to Computers, or just to Geography.

Each team had to decide what *categories* of data were the most significant.

The next question was *how to collect* this data. Some suggestions were the school books, the library, encyclopedia etc. It may be possible to get some data from PC Software such as CD-Rom Encyclopedia, or access The Internet.

In a *team project*, it was important to allocate the workload. Some teams gave each member a set of animals to research, and find data on the chosen categories.

Other teams approached this a different way by giving each team member a category to research, and find data on the chosen animals.

The next issue was how to *represent or store* this data. For example, if animals were classified by "type" (mammal / reptile / bird etc.), they could be represented by a simple grid, as follows...

Animal	*Mammal*
Jackamar	Bird
WeaverBird	Bird
AntEater	Mammal
Macaw	Bird
Monkey	Mammal
Bird-Eating Spider	Reptile

But this would make it difficult to ask questions like "how many birds are there ?"

Alternatively, they could have a more complex grid with "ticks" against each attribute, as follows...

Animal	*Mammal*	*Bird*	*Fish*	*Reptile*
Jackamar		1		
WeaverBird		1		
AntEater	1			
Macaw		1		
Monkey	1			
Bird-Eating Spider				1

This led to the concept of having *sub-categories* to split the data into.

We discussed that advantages of having simple numbers, ("1" mean "yes", "0" means "no"), rather than entering words. It is quicker to put in a number "1" than write in a complete word.

A computer is better suited to adding up numbers than working out the meanings of words. If we had numbers instead of words, we could get the computer to quickly add up the number of animals in each category, and put a "total" amount at the bottom.

We had to put the data into the computer in a different way.

If we chose an animal's "colour" as a category, we could have sub-categories such as "red", "green", "black", but what would happen if we found an animal that is coloured "brown" ? The trouble with having rigid sub-categories is that they may not cover every possible value. We might need to add new sub-categories as we went along.

This discussion led to the idea of having an *"other" sub-category* in many cases. We now understand why parents often have to fill in forms with categories including "other" !

3. Input and Presentation

We had data to enter into the computer.

It was important that we remembered that planning, and doing the research, is just as important as putting information into the computer.

We discussed the idea of *data samples*. If we took a sample of one child from the class, and asked his opinion (on, say, the best football team), would that be a good representation of the entire class ?

In a real world database like our rain forest, we might need to get thousands or millions of different bits of data.

If we have wanted to *summarise* the data, (such as "how many mammals are there ?"), this could be done fairly easily with a small amount of data. But it was much easier if we could get the computer to do this. So we used the "sum()" function to add columns of data for us.

At this stage, we had lots of data, but it was difficult to see any pattern. So we used the *charting* functions within Excel to create bar charts, pie charts, graphs etc. In order to make sense, we had to ensure that charts had correctly labelled axis, legends and titles.

When we had completed and printed the graphs and charts, we needed to look at how to *interpret* them. What did a a particular shaped chart mean ? Did it tell us anything about what the Rain Forest was really like ? What conclusions could we draw from the diagrams ?

At the same time, this project also involved some practice with Excel. We needed to use some of Excel's more advanced features to solve some practical problems...

- Freeze Cells so that we could enter data on one part of a spreadsheet at the same time as reading the headings
- Altering column and cell formats so that the data could be shown easily.
- Colour Filling for Text and Background to present the data well.
- The AutoSum function to add columns of data.
- Selecting Data Ranges for formatting or charting.
- The Charting Wizard.

4. Feedback

Once data has been organised in a data base, it is possible to represent it using pie charts, bar charts etc, to answer some simple questions, such as...

- what is the most common type of animal (mammal/reptile/bird etc.) ?
- are most rainforest animals nocturnal ?
- what is the most common colouring of rainforest animals ?
- which part of the rainforest (emergent/canopy/understory/floor) has most animals ?

Using more advanced database techniques, it would also be possible to analyse it in a more complex way, by looking at correlations. For example...

- Do most mammals live in the understory ?
- In the canopy, what is the most common type of animal (mammal/bird/reptile etc.) ?
- Which layer (canopy / emergant / understory / floor) do most of the nocturnal animal live in ?

But these sort of questions may need to be left until another day.

Reviewing the project, a number of questions could be asked ...

- Did we chose the correct categories of data ? If not, which categories should we have chosen and why ?
- How did we research our data, and were there any other places where we could have got useful information from ?
- Did our data fairly represent what the rainforest is really like ? If not, what have we learnt about data samples, etc.
- What sort of graphs, pie charts, bar charts etc were most helpful in understanding the animals in the rainforest, and why ?
- Was Excel the best computer program to use ? Why ? If not, what would have been a better program, and why ?
- Was a computer necessary for this project, or could it have been done some other way ? What have we leant about computers being a useful tool for organising information ?
- If we had millions of bits of information, how would that have affected the way we approached this project ?

Dennis Adams BSc.
Database Systems Consultant
Member British Computer Society
April 1998.

Some Hints and Tips when using Excel

Use the “Red” option of Excel, which gives access to the Formatting Options, additional Toolbars, and Charting Wizard.

Don’t forget - put time into planning what you are going to do, before entering data.

Make sure that your work has a title (including the name of your team), and properly labelled X & Y axis and descriptions.

Toolbars

Excel has a number of toolbars available. The toolbars we require are **Standard** and **Formatting**. If you right-click on any toolbar, you can chose which toolbars to display.
Alternatively, you can display the appropriate toolbars by chosing the Menu Option **View / Toolbars**

Freeze Panes

Excel is capable of handling thousands of rows and columns of data. If you enter a lot of data, you may notice that the first row and/or column disappear as Excel scrolls to another cell.

We can “freeze” these initial cells, so we can enter data on one part of a spreadsheet at the same time as reading the headings. Position the cursor just inside the area which you want to “Freeze” and chose the Menu Options **Window / Freeze Panes**

	A	B	C	D	E	F	G	H
1	(Examples)	Animals we have found in the Rain F						
2		What they Eat		Where they Live				V
3	Animal	Herbivore	Carnivore	Floor	Understory	Canopy	Emergent	Mami
4	Jackamar							
5	Vulture							
6	WeaverBird							
7	AntEater							
8	Macaw							
9	Butterfly							

This is the region that will scroll, while the "frozen panes" will be visible all the time.

Formatting Cells

Sometimes, columns are not wide enough to display the data which you have entered. If you position the cursor on the column headers, you will see a column width cursor, which enables you to change the width of each column to display the data better.

To make some text fill across several columns of data, select the range, and then click on the "centre across columns" icon on the toolbar.

Colour Filling

To change the colour of a cell, the colour of it's text, or it's border, highlight the cell, and click on the appropriate "borders", "color" or "font color" icons...

Auto Sum

The AutoSum function can be used to automatically add columns of data. Position the cursor where you want the total to appear. Click on the AutoSum icon.

Then use the cursor to highlight the range to be summed, and press the ENTER key.

Selecting Ranges of Data

Ranges of data can be selected by highlighting the area with the mouse cursor, and keeping the mouse button down at the same time.

If you need a *discontinuous* range (such as when you want to include chart titles as well as totals), hold down the “Ctrl” key at the same time as you select the other part of the range.

The Charting Wizard.

The charting wizard is used to “build” a chart or graph. Select a range of data you wish to display, then click on the wizard icon. 

Pick up the icon and place it over the cells which you want to display.