AUTHORS: Dr Rodney Rate (presenter)

Business Phone	(08) 92738273
Fax	(08) 92738167
e-mail	r.rate@cowan.edu.au
Address	Edith Cowan University
	Pearson St
	Churchlands WA 6018

Mr Robert Rispoli

Business Phone	(08) 92738234
Fax	(08) 92738167
e-mail	r.rispoli@cowan.edu.au
Address	Edith Cowan University
	Pearson St
	Churchlands WA 6018

AUTHORS: Dr Rodney Rate (presenter)

Business Phone	(08) 92738273

Dr Rate is the Manager of the Institutional and Research Branch at Edith Cowan University, a position he has held since its inception in 1988.

Mr Rispoli is a Research and Statistics Officer in the Branch who has been attached to the Branch for the same period.

TITLE: **Providing Information to Managers:**
An Alternative to the Engine Room Model

Abstract

Increasingly, management and faculty are demanding ready-access to information that will maximise institutional-decision making within the University sector. Planning and Statistics Offices are being placed under pressure to provide vast arrays of data with reduced timelines and in an environment where human resources in many management units have been cut quite severely. Alternative procedures for providing timely and easy access to data have had to be examined. Clearly, OLAP products such as Seagate Holos, the UniOn endorsed product, are available but the cost of providing software and user-training is a serious impediment to their widespread use.

At Edith Cowan University, the Institutional Research and Statistics (IRS) Branch has turned to the features and capacity of the World Wide Web and Microsoft Excel to present data in a format which allows users to easily interrogate critical databases. In particular, Branch staff have created Excel-based PivotTables® which summarise the information contained within the databases and which enable the end-users to customise their own reports. The tables offer the facility of viewing the data from alternative perspectives and are powerful analytical devices. While there are some limitations, they provide ready-access to the most recent data and include appropriate time-series comparisons, with minimal cost. This has resulted in a significant reduction in ad-hoc query demands particularly in heavy work periods such as the enrolment period of January though February.

This paper describes some of the PivotTables which have been made available to the University community through the home page of the IRS Branch; their construction, special features and relationship to the data sources. Particular attention is given to the support mechanisms provided from within the Branch which make the process user-friendly.

Providing Information To Managers:

An Alternative To The Engine-Room Model

Dr Rodney Rate
Manager
Institutional Research and Statistics Branch
Edith Cowan University
Perth, Western Australia
e-mail   r.rate@cowan.edu.au
fax   (08) 92738167
phone   (08) 92738273

Mr Robert Rispoli
Research and Statistics Officer
Edith Cowan University
Perth, Western Australia
e-mail   r.rispoli@cowan.edu.au
fax   (08) 92738167
phone   (08) 92738234

Introduction

The provision of reliable and valid data upon demand and its presentation in a clear and informative manner is a critical function of University offices which are variously named Statistics, Institutional Research, Policy and Planning etc. In Australia, no two organisational units charged with responsibility to supply management information appear to have exactly the same structure or function but they all have the common charter to respond to requests for data which will help management and faculty to plan and promote appropriate policies. Invariably the answers to questions raised simply stimulate further queries which often require data sources to be re-visited or re-created. This is often time-consuming and frustrating and, of course, inefficient. Most of the problems are created through inadequate scoping of the specific problem being examined.

In introducing this project to you, the authors would like to share the following scenario:

Scene One

Academic on Notice with Nowhere-to-go (Ann) calls the Institutional Research and Statistics Branch (IRS) in a panic-stricken state looking for favourable statistics on attrition in her course which could save her job …

Ann: Hello, is that the Stats Branch? I need some stats on my students. Can you help?

IRS: Sure, what do you need?

Ann: Well I'm the coordinator of Course X and I'm concerned about the drop-out rate. We seem to lose an awful lot. I'm sure we had a lot more at the start of the year.

IRS: Well yes, we have data that will assist you but what do you mean by drop-out? From the course, the Faculty or the University. And what about students who have officially deferred? Do we include them or exclude them as drop-outs?

Ann: Oh! I hadn't really thought about that. What do you suggest?

sometime later after much discussion …

IRS: That seems to have it pinned down then doesn't it. Is there anything else you might want to know? … No! … OK, we should be able to provide you with a report within three working days.

Ann: That would be wonderful.

Scene Two

Three days later, IRS staff, up-to their armpits in ad hoc requests and DEETYA returns, receive another call from Ann

Ann: Hello IRS. Thank you for the data. It was terrific. I just had another thought. Would you be able to separate out the fee paying overseas students?

IRS: (*slightly miffed*) No problem - we'd be pleased to do that but it will require us to re-run our analysis and you won't get the data for a couple of days.

Ann: You're wonderful. I can wait.

Scene Three

Two days later, with one staff member on annual leave, 425 fatal errors on the DEETYA return and eight unattended ad hoc requests, the IRS Branch takes yet another call from Ann …

Ann: Hello IRS. You're going to hate me for this! (*audible grunt from the other end of the phone*). The data I got today really made me think about possible explanations for the drop-out rate. Would you be able to give me some data for the previous three years? You would? Oh that would be fantastic. And on another point, do we know whether the drop-outs were school-leavers or not and whether those that weren't had previous Higher Education experience? … Hello … IRS Branch … Are you still there…Hello?

Been there? Done that? In the IRS Branch at Edith Cowan University (ECU) we certainly have and very often the demands and requests have arrived at most inopportune times such as the beginning of the year when staff are encouraged to take leave or around census dates when data cleansing and consolidation require higher priority.

The Setting

ECU is a 20,000 student institution which evolved from the West Australian College of Advanced Education in 1991 as an outcome of the introduction of the unified national system. It recently underwent a major restructure of its academic organisation. It now comprises three faculties

- Business, Legal Services and Public Administration
- Communications, Health and Science
- Community Services, Education and Social Sciences

and two additional large organisational units which are resourced and managed as "Faculties"

- the "Bunbury" campus
- the WA Academy of Performing Arts

The University is housed on three metropolitan teaching campuses and one country campus in the City of Bunbury, 150 kms south of Perth. Consequently courses and units of study are offered across a number of campuses as well as through distance education and on a number of "off-shore" sites. In addition only half the student population is enrolled on a full-time basis and a minority of the undergraduate commencers are school leavers. The greater proportion are mature age or transfers from other programs or institutions. Requests for data and information related to staff and student activity are therefore complicated by the physical structure and mission of the university.

At the time of writing, the IRS Branch comprises a Manager, three Data Analysts and an Administrative Assistant. Its mission is to continually develop and provide efficient and effective institutional research and statistical operations. Specifically, the goals of the Branch are to:

- collect, analyse and present statistical data on students, staff, finance, research, facility and environmental factors

- prepare and provide statistical publications, annual statistical returns to the Commonwealth Government and ad hoc statistical reports

- prepare projections, models and planning reports to support institutional planning and decision making with respect to enrolments and resources.

Scope of the Project

For many years at ECU, the large part of student management information has been produced through the development and application of SAS programmes which produce accurate but rather unattractive reports. Matters concerning staff data have been resolved by SQL queries of the University's Oracle database. Fortunately the Branch has skilled programmers in both areas and reports have been provided within reasonable response periods. However, Branch staff were keen to see the level of ad hoc enquiries reduced. The University's involvement in and commitment to the development of the UniOn Executive Information System (EIS) was seen to offer a quality solution. After three years and no sign of a definite release date, it was considered that a temporary alternative needed to be developed and implemented.

In developing the scope of the project the following factors were determined to be critical;

- data to be readily accessible to the end-user (the enquirer)

- implementation and access to be cost efficient

- access to data to be user-friendly

- users to be able to explore the data so as to extend queries

- users to be able to automatically print quality reports

- training in use of the system to be readily available.

After significant debate and review of alternatives the Branch identified Microsoft Excel and its PivotTable® features as the best solution in the short term. The decision was supported by the fact that all staff in the University had access to Excel and that Branch staff were well acquainted with PivotTables. The application selected as the "pilot development" was named the "Enrolment Tracking Data Module". Essentially its purpose was to make enrolment data readily available to the university community throughout the enrolment period of January and February. The data includes course intake quotas and March 31 targets. In previous years the Branch had created inflexible SAS summaries on a daily basis. The daily report consisted of over 100 pages of "ugly" SAS reports from which the main users were able to extract the information they required. However, non-regular users found the reports to be very cumbersome and difficult to interpret. Often, special reports would have to be prepared in a quite time-consuming and inefficient manner. It was also difficult to forward the reports electronically because of incompatible e-mail clients and desktop software.

The first PivotTable applications incorporated the following features and accompanying strategies;

- they were made available on the web but restricted to ECU staff

- the opening view represented that which we thought would be of interest to the majority of our users (eg Figure 1)
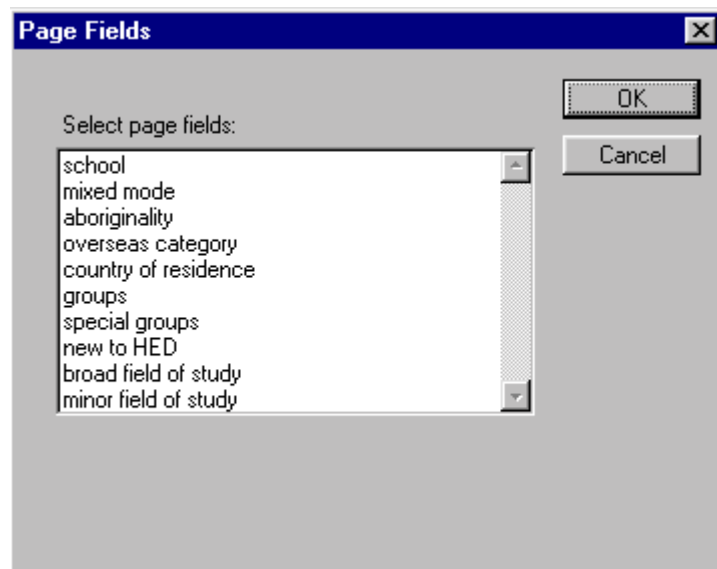
Figure 1: Enrolment Statistics - Opening View

Data refreshed: 8/10/98 10:56:01	Add Page Fields	Select Statistic(s)	Tidy Up
funding category	Operating Grant		
campus	(All)		
mode of study	(All)		
sex	(All)		
course level	(All)		
course	(All)		
year of study	(All)		

|		Data			
faculty	broad course level	Commencing Bodies	Commencing EFTSU	Total Bodies	Total EFTSU
Bus,LegalSrv,PubAdmn	Higher Degree Research	7	4.00	59	34.75
	Postgrad. Coursework	3	1.50	318	122.75
	Undergrad.(inc. Enabl.)	307	200.25	3198	2180.75
Bus,LegalSrv,PubAdmn Total		317	205.75	3575	2338.25
Comm Srv,Ed,Soc Sci	Higher Degree Research	31	14.33	237	126.81
	Postgrad. Coursework	13	5.25	654	405.04
	Undergrad.(inc. Enabl.)	596	390.90	5490	4180.64
Comm Srv,Ed,Soc Sci Total		640	410.48	6381	4712.50
Comms,Hlth,Science	Higher Degree Research	25	15.57	186	115.97
	Postgrad. Coursework			270	103.66
	Undergrad.(inc. Enabl.)	456	306.41	4503	3239.76
Comms,Hlth,Science Total		481	321.98	4959	3459.39
Cross-Inst./Non-award	Cross-Inst/Non-award	39	16.50	78	31.00
Cross-Inst./Non-award Total		39	16.50	78	31.00
WA Academy	Higher Degree Research			6	4.75
	Postgrad. Coursework			40	23.57
	Undergrad.(inc. Enabl.)	9	7.40	613	547.08
WA Academy Total		9	7.40	659	575.40
Grand Total		1486	962.11	15652	11116.53

- less critical page fields were hidden but made easily available with the addition of the "Add Page Field" button and associated dialogue box. This feature, in particular, is essential to the success of the application since long field names are often truncated by the PivotTable wizard included in Excel and short field names are, often, not very meaningful. (eg Figure 2)

Figure 2: Enrolment Statistics - Additional Page Fields

- users were given the opportunity to select appropriate data fields via the "Select Statistics" button and associated dialogue box (eg Figure 3)

Figure 3; Enrolment Statistics - Select Data Fields
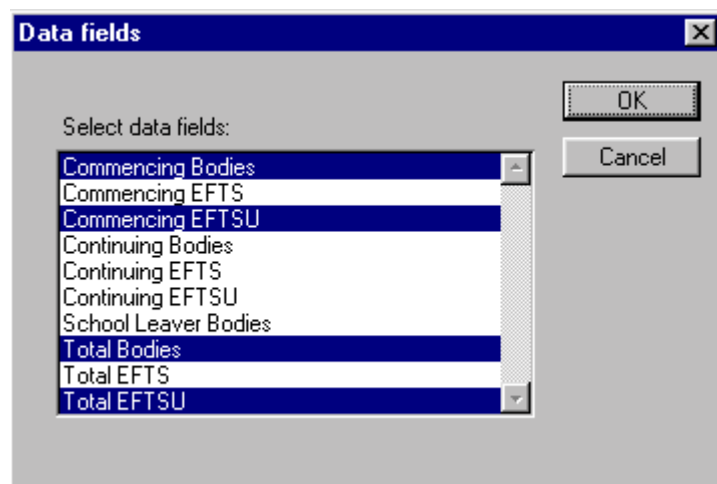
- an opening popup message box directed users to staff in the IRS Branch who could provide assistance in using the PivotTable

- an instructional video movie demonstrating PivotTable navigation was developed and made available on the web

- Deans, Heads of School and Senior Administration Officers in Faculties were invited to attend instructional sessions at which PivotTable navigation was demonstrated and discussed

- training was offered and provided to staff both in groups and individually

- an attractive print facility was in-built

The "Enrolment Tracking Data Module" was immediately successful. Staff at all levels across a wide cross-section of the university community sought assistance in learning how to find their way around the PivotTable. It became patently clear that the application was being used for the reasons that;

- it was accessible

- staff were given support, encouragement and training in table navigation

- reports generated were attractive and clear

  and, most importantly

- the data was always fresh, informative and extensive

Branch staff were greatly encouraged by the response and moved quickly to plan and present other databases as PivotTables. Since that time, PivotTables have been developed for;

- semester enrolment data

- taught load distribution

- 1994-1998 staff data

- course attrition - multiple modules

- current course demand (Tertiary Institution Service Centre data)

- Western Australian university course demand

- all course completions

- results of the GCCA graduate destination survey and course experience questionnaire

- university enrolment projections

Technical Issues

Background Information

PivotTables are manipulated by using several objects in Excel's object model.

When a PivotTable is initially created (or refreshed), all the underlying data is read into an invisible cache which Excel maintains in the computer's RAM. This data structure is represented by the PivotCache object.

When the visible PivotTable is manipulated to show various fields in various positions, data from the cache is summarized and displayed in the worksheet as requested. These operations are very fast since all the underlying data is already in RAM in the cache. In this way, the structure of the visible PivotTable can be changed radically over and over without going to the original source of the data. The visible PivotTable is represented by the PivotTable object.

Since PivotTables store all the underlying data in RAM in the PivotCache, performance can degrade with very large data sets.

Software Requirements

Microsoft Excel V5.0 or above

Getting the data into shape

To illustrate the processes used by the IRS Branch in the preparation of the Enrolment Tracking PivotTable, for example, we sequence the steps:

1. Run the Stats Extract program (COBOL) to create the Enrolments and Load flat statistics files from the Student Records System (SRS) based on current enrolments

2. FTP the files created in (1) to the PC

3. FTP updates to the Course (contains course names, faculty codes etc) flat files to the PC if necessary

4. FTP latest preferences/offers data from TISC

5. Ensure Quota and Course Projection .dbf files are accurate and complete

6. Ensure TISC course codes map to ECU course codes in a .dbf file

7. Run SAS job to bring it (all the above data) together into one delimited flat file ready for parsing in Excel

8. Refresh the PivotTable using the data from (7) above

Tips

Here are some worthwhile tips which we have gathered in the process of the creation of the many PivotTables we now offer our users. You'll note that some of the tips relate to the processes involved in the development of any EIS application:

- We never manipulate/amend any "raw" data in Excel itself! Nor are any code lookups done in Excel.

- Give meaningful descriptions to field classifications e.g.. do NOT display course codes alone - include course names!

- We never maintain our own lookup files! For example, it is much better to FTP a newer version of the Course file from our SRS rather than maintain a separate one on the PC.

- If you don't have SAS, can use SPSS etc. or even link the PivotTable to a database (e.g. Oracle, Access) via ODBC. Basically, use whatever tool you are familiar with.

- Never save the raw data as a worksheet in the same workbook as the PivotTable unless there is a problem with RAM and you wish to use the new features in Excel97 (see last tip). It is totally unnecessary and wasteful.

- If you have Excel V7.0 or below and your raw flat file data has more than 16384 rows, you will not be able to "open" it in a worksheet. Instead, import it into, say Access, and then use the "Get External Data" feature in Excel to refresh the PivotTable. This is, of course, not an issue if your source data is a remote database.

- Never clutter the screen with too many pagefields. Confuses the user and looks messy.

- If in doubt … put it IN! i.e. if in doubt about whether to include a particular field and, more importantly, it does not involve much more overhead, then put it in. For example, the inclusion of "field of study" as a field does not generate any more rows of data if "course" is already a field. Never underestimate the requirements of your users!

- The Cursor property of the Application object in Excel can only be changed in V7.0 or above. So, for example, the following VBA code will bomb in V5.0 :

  Application.Cursor = xlWait (sets the appearance of the mouse pointer to the hourglass pointer)

- There are two techniques available to manage large data sets in Excel97

  - Server-Based Page Fields

    In Excel97 PivotTables have a new feature called Server-Based Page fields which can be used to manage memory when dealing with large data sets. For PivotFields in the page position, setting the ServedBased property to True will cause the PivotTable to only query data into its cache that matched the currently displayed value in the page field. Thus, each time the selection is changed in the page field, the query is re-executed and the cache rebuilt. For large data sets, this can help manage the size of the cache. Of course, for small data sets, this will probably slow performance since it incurs a database hit every time the selection is changed. Having this property gives you the option of tuning the PivotTable cache according to the situation. Note that when a field is server based, the "(all)" selection is not available since the whole data set is never in memory at once.

- Limiting Data Read into the PivotCache

  Through thoughtful contruction of the SQL statement which is used to build the PivotCache itself, you can limit the number of records read into RAM. For example, if you only need to analyze data for a particular campus, restrict the data being read by using a WHERE clause on your original query.

Limitations

As indicated above (see *Background Information*), PivotTable size is limited by a workstation's available RAM.

One of the "biggest" PivotTables developed at ECU (the *1994-97 One Year Attrition* PivotTable) summarises data from a flat file which consists of over 60,000 rows by (17 fields + 6 statistics)

The fields consist of:

- Year (4 values)
- Course (329)
- Sex (2)
- Aboriginality (2)
- Rural/isolated status (4)
- TEE score (14)

- Faculty (7)
- Course level (14)
- Overseas Indicator (2)
- Campus (8))
- Basis for admission (7)
- Study status (16)

- Non-English speaking background (2)
- Mode of study (3)
- Age (6)
- Socio-economic status (4)
- Admission category (6)

This PivotTable can be successfully manupulated on a PC with 16 Mb RAM. However, Apple Macintosh users with similar RAM have reported problems. Our solution is to replace them with PCs! In fact, it is University policy that Apple desktops will no longer be supported by the central IT division.

The Future

There can be no doubt that the development of PivotTables as described above has had a dramatic impact on the workload of the members of the IRS Branch at ECU. Equally, there is no doubt that the rest of the university has endorsed and adopted them as primary data sources for management decision-making and policy formulation. The authors consider however that the UniOn EIS will eventually replace all other methods of statistical data reporting and analysis within ECU. It is based on a true OLAP software package (Seagate Holos) with data residing in an enterprise data warehouse.

Microsoft has, however, indicated that :

"Excel 2000 will support OLAP data sources (those that comply with the OLE DB for OLAP specification) in addition to traditional relational sources. Excel will consume such data in the PivotTable feature. The current PivotTable feature is being extended to accommodate OLAP data sources."(See http://www.microsoft.com/data/oledb/olap/excel.htm for further information. It's interesting to note that Holos data cubes comply with the OLE DB for OLAP spec.!)

In the interim, the IRS Branch intends to maximise the power of PivotTables and to extend their use throughout the university community.