RAPID BUSINESS BENEFITS FOR CLIENT RELATIONSHIP MANAGEMENT THROUGH DATA WAREHOUSING AND MINING

Nigel Wigram

INTRODUCTION
Life assurers exist and make profits from people investing growing sums of money in life assurance policies and other financial products, and in leaving the money there. Our revenue essentially comes from having money under management. There are some ongoing administrative costs, but the largest parts of our variable costs are associated with starting money flow.

The cost benefit ratio is central to all operations. If we can reduce costs or improve benefits we will be able to make our products more attractive to our clients - making it easier to sell them more or to keep the ones we have.

The growth of client portfolios in force thus both drives cost/benefit relationships and is driven by them.

As life assurance is a relatively unsought good, we are dependent on various sales channels, and particularly a force of more than 5000 commission-earning salesmen to go out and bring in the business. In theory a commission based system should mean that costs of sales are directly proportionate to the volume of sales, but in practice there is a huge overhead from the cost of recruiting, training and establishing salesmen. Salesmen with low productivity leave the business thus no longer contributing benefits to offset the sunk costs. Any improvement in salesmen’s productivity is immediately reflected in a reduction in the staff turnover rate. A few more sales now and many more sales into the future. Other sales channels like direct marketing also have break-even targets so that improved productivity allows additional markets to be reached cost effectively.

Finally, the life assurance industry is plagued by the problem of early discontinuances - clients who purchase a policy only to discontinue it after a very short time. Most of the life assurers’ costs are incurred at the start of the policy and will generally not have been fully recouped from an early discontinuance. A reduction in discontinuances immediately saves costs, but also grows the portfolios proportionately.

Commission on new business is generally spread over the first one or two years of the policy’s existence, so if the policy is discontinued early, the salesman is not rewarded for his effort with a negative effect on his productivity and probability of continuing in the industry. Conversely, as we shall show later, intermediaries who leave the industry are generally associated with discontinued policies. Actions can also be taken in regard to later discontinuances.

It is recognised that the proper use of a marketing database can drive profitability and Old Mutual’s top management have identified its client database as its most important strategic asset.
Old Mutual considers itself something of a leader in database management at least in South Africa.
For the last 20 years we have had an operational marketing database which we have used extensively both to underpin direct marketing campaigns, and to support our salesforce.

We have 2-3 million current premium paying clients about whom we know a certain amount, and with whom we can be in regular contact.

Not all life assurers share this advantage since some keep their clients at arm’s length and only communicate through intermediaries such as independent brokers.

We also have information on several million other people either because they were clients or are connected to contracts in some way … often being family members of existing clients. Unfortunately the information which we have about these extra people may be somewhat skimpy, although reliable name and date of birth information provides good matching keys to external data sources.

Our name is very well known and trusted by virtually all the economically viable people living in the countries in which we operate - probably 40% of these households are on our current clientbase. Although much of our business comes from our existing clients, we are only getting about half the additional business which they purchase.

The latest brand theory suggests that good brand image puts you in the clients acceptable set but actions come from being in contact with your client at the right moment, whilst productivity comes from not wasting contacts unnecessarily.

**OPERATIONAL SYSTEMS**
Our administrative systems are based on silo’s of different classes of life assurance products each running under its own operating systems for day to day contract level administration. These were largely developed independently and often on a wide range of technologies. The majority are however COBOL systems operating against IDMS data.

In front of these we have a client system which carries personal and address information on all the people associated with a contract … One person may be associated with many contracts, and each contract may be associated with several people. In our individual life business we have 2.5 million current premium payers owning 3.5 million current policies, but there are over 6 million people with a total of about 10 million contract relationships.

Associated with these contracts are our sales and servicing intermediaries.

For transactional purposes individual clients contact our sales or service points who use a workbench or interface system to access the current information in real time.

**MARKETING DATABASE (PORTFOLIO EXTRACT SYSTEM)**
Each month we swept the operational databases to populate the marketing database … This database was designed and tuned to provide batch extracts of data based on a portfolio views of our clients of our major lines of business. This system was originally written in COBOL/IDMS and subsequently rewritten in Huron. The extract systems were written in COBOL.

The database can thus be accessed in terms of any combination of client, product and intermediary attributes as they are reflected at the time of the sweep. At the time of an extract a campaign marker is set which is marked back to the client records so that it can be used for subsequent analysis purposes.

The basic selection or extraction systems were unable to handle the complexity of many of the requests that were put to it, particularly when there were complex exclusions and the need to establish matching samples.

This led to the development of a 2 stage system whereby there was an initial extraction on the basis of simple rules which was then passed to a second system (the VM machine) for further refinement. This has the additional disadvantage that we lost the ability to mark extracts accurately back onto the database for subsequent tracking.

The various outputs could then be prepared in electronic or paper form and passed on to the users - either our sales intermediaries or direct mail or direct marketing systems. We use Compuset as a formatting tool.

The database is used as the data source for a variety of analysis systems.

**REASONS FOR CHANGE**

Marketing users have been asking for new systems promising to help grow the business, whilst the I.T. department wanted to spend most of their effort on the basic systems to keep the business running.

General Management had to balance these pressure whilst also allocating resources to enable the organisation to go into new lines of business.

Early last year we finally came to a cross-road. Although the existing systems gave us considerable functionality, they were inherently slow and inefficient consuming vast amounts of computer power. This wasn’t really a concern when we were able to run extracts at night and over weekends on otherwise idle machines, but became more and more problematic as the available batch windows became smaller. By the beginning of 1997 some of the very large jobs were taking several days to complete, and a backlog of extraction requests was building up.

We found that many of the systems that ran on the VM machine were not year 2000 compliant. This, coupled with the disproportionately high cost of running this environment, prompted a decision to rewrite the systems on other platforms and remove the VM environment.

**RULE 1:** To create a successful new system, top management, user...
Whilst the operational priorities demanded that we create a new data system, other developments suggested that the new system could do a much bigger job. General Management already under pressure for improved productivity and recognising the potential value from the client sector, authorised us to start work on a “Client Relationship Information System” which was to be based on the broad principle of an enterprise wide data warehouse.

**RULE 2: Have a big vision but start small.**

Although the theory of datawarehousing says that you should build an enterprise wide datawarehouse containing all conceivable data items, and to meet all needs of all users, experience suggests that this is an impossible dream. More success has been achieved by focusing on the more immediate needs of a limited user group, although the basic vision and design should be open ended.

As the Client Relationship Manager of the Individual Life company I was designated as the user owner and primary beneficiary on a project funded by the life assurance company, but with the intention to sell benefits to other lines of business. The warehouse thus has a marketing focus.

**THE DECISION TO START IMPLEMENTATION WITH THE SAS SYSTEM**

The SAS System appeared to offer a relatively low risk and low cost opportunity to test the viability of new directions.

At no stage have we had to take a major decision which committed us irrevocably to this particular route, and even the hardware purchases have a payback period measured in months - or less.

The SAS System route provides a series of building blocks. Each step which is completed forms the basis for the next without prescribing what form the next step will take. Although the SAS System has a natural advantage in ease of building on the existing material.

The SAS system certainly appears to be hardware independent, and we have been able to test and use a variety of platforms.

- **MVS** - Virtually all our data is derived from MVS systems, and it is the logical place to crunch big volumes of data.
- **SERVER** - initially launched with NT but could move to Unix at some later stage. Each box is quite low cost and can be used for other purposes so experimentation does not require a capital commitment. We have however found some problems downloading large volumes of data from the mainframes to servers. Legacy architectures and bandwidth constraints ensured that both shared data access and data transport were problematic.
- **PC** - these are used both as front ends and for analysis purposes.
The initial decisions were perhaps easier because we already had most of the SAS System components licensed somewhere in the business.

My user teams had two previous experiences with SAS System which caused some initial concerns.

In our market research area we had been using the SAS System as a data analysis tool but our experience was that the SAS System was an extremely unfriendly language that was difficult to work with. We subsequently discovered that the version of the SAS System that our Research team was using was not the latest or most friendly. In the meantime we found an extremely simple point and click windows based graphical analysis system which was very simple to understand and use. I have yet to see a SAS System data visualisation equivalent.

The second unfortunate experience was a premature attempt to build a SAS System based geographic information system interlocking with our client database. The contractor was also the promoter of the system and somewhat over-reached himself being unable to deliver on his promises.

**RULE 3: GET FULL USER MANAGEMENT INVOLVEMENT.**

There are a number of lessons that can be learnt from these experiences. Technology is a moving target and end users must be kept up-to-date with new releases of software. Over-promising leads to disappointment; and a project which only delivers benefits at the end is not recognised when it is 95% complete. If you don’t get proper user management commitment you will probably be seen to fail - no matter how good you are.

**VALIDATING THE PROCESS**

The first steps towards a SAS based datawarehouse were to convert the existing Huron based snapshot database system into SAS datasets and to demonstrate the efficiency of the SAS System as a data selection tool.

This was done as a prototype by a SAS Institute contractor in very few weeks, and quickly demonstrated that the SAS System could easily replicate the existing system and produce comparable output using a fraction of the computer usage - one tenth of the elapsed time and one thirtieth of the CPU. There was also the promise that the SAS System extraction could, in a single pass, do the complex 2-phase extracts that previously required transfer onto the V.M. machine.

The user has now got access to a tool which allows him to compose an “extract request” without technical assistance. This is converted by a “source code generator” within the system, allowing multiple complex requests in a single pass, including segmentation for test samples.
This gave us the confidence to immediately “industrialise” the process relief from bottlenecks. The industrialisation is an ongoing process, but we were able to switch off the old system in December 1997.

The initial benefits of the move to the SAS System were that the new systems were very much faster than the ones that they replaced. Operational savings in computer usage would immediately recovered the costs of moving to the new system. Much of the building work has been done by the team that would otherwise have been supporting the Portfolio Extract System. The costs of contractors and SAS Institute support is covered by ongoing savings and even the purchase of a new NT Server recovered its costs in weeks.

The real gains however came from being able to handle the extraction requests much more rapidly and comprehensively - Turning this into user benefits we were able to promise our users that all requests would be handled within 8 days, as opposed to the previous 6-8 week promise, and that urgent requests would be delivered within 48 hours. These promises are sufficient to change the whole way in which the users can treat extraction needs.

**RULE 4: ENSURE THAT YOU DEMONSTRATE REAL BENEFITS FOR ALL PARTIES - QUICKLY.**

Better defined needs and sales situations mean that we can generate more focused leads which immediately mean that we can be more selective and improve the quality of leads at least by generating less “unsuitable” leads. This both enhances value and reduces costs.

This is where datamining and data analysis play a role. The more we are able to understand the dynamics of our marketplace, the better we are able to focus attention on the right places. We will expand on datamining below.

In March 1998 the upgraded system was run in parallel on the MVS System and on an NT server. Whilst we have not yet fully explored all the issues associated with running the SAS System code on the NT platform the initial tests showed the server to be somewhat slower than the mainframe, but this may be counterbalanced by the fact that server operations are essentially less expensive, and that users have much more flexibility in scheduling operations. Even in the MVS system many of the extracts can be handled in parallel with production runs. The six to eight days for standard requests can now come down to 48 hours.

The test also demonstrated the total portability of the SAS System code across the various platforms.

**BUILDING THE STORE (DATAWAREHOUSE)**

The success of the initial project and the immediate operational relief it provided also gave us an opportunity to launch the datawarehouse process under slightly less pressure, although the project manager, Tommy Freeborough committed himself to deliver a basic working system in 4 months.
The first stage of our new datawarehouse has been to extend the range of business covered by the system. Old Mutual has moved away from being just a life assurer and introduced a number of other product fields - including Unit Trusts. These products are managed through separate subsidiary companies and were not linked to the Portfolio Extract System. They were however often the same clients and sold through the same intermediaries.

With the rebuilding of the system we have been able to ensure that we bring across more information from the source files. This is especially true of information about our intermediaries and includes derived information about their sales habits, as well as linkage to their personal portfolios. The relationship between their personal purchasing behaviour and success in selling specific product types is interesting.

Clients can also be viewed as households so that we can look at the combined portfolios and interactions within a household or family unit. Whilst we expect that this will give a huge boost to the value of contacts, and to our understanding of the dynamics of the purchase of life assurance. Householding doesn’t mean that all our activities should be based on one per household.

Next we have changed the way the datawarehouse is refreshed. Whilst initially the concept of a full monthly sweep of operational data is maintained, key items will be updated frequently - the first being death claims and changes of address which are updated daily. This allows us the facility to ensure that any campaign is aimed at latest address and that we exclude all deceased persons. This should give a huge saving on contacts which would otherwise have gone to the wrong address.

The original Portfolio Extract System only contained the single snapshot of data, but the new system is also carrying the date-stamped history of the changes enabling us to look at both the before and after situation.

The next phase is to bring this part of the system up to industrial strength, and to give it a full set of user access tools.

Then we must set up a number of partitions or datamarts giving each user line of business access to the full information about the clients and their families who are associated with that line of business - not just their own data, but fully enriched from the other sources.

Although we thought that our databases were all in reasonable condition, the act of bringing them together has identified many inconsistencies.

In May 1998 the whole system will receive a major boost as we will be doing a major datacheck on the more than 3 million clients who will benefit from a free share distribution when Old Mutual demutualises.

The database is essentially open ended in that we will be adding new data sources and facilities over time.
The upgrade of the system is not limited to the actual datawarehouse itself, but will also extend to the formatting of the output and to the delivery to the workbenches in electronic form rather than simple paper forms such as direct mail, paper leads or control lists.

MARKETING SUPPORT

RULE 5: GOOD COMPUTER SYSTEMS ARE FACILITIES NOT PANACEAS

It is most important at this stage that marketing teams are involved teaching end users the real benefits of the new information.

Stage 1 Test the data and make sure that it is really telling you what you think it does.

Stage 2 Give the data to a few intelligent users and study how they use it … work with them both to refine the way you present the data and to develop tips that you can pass on to others.

Stage 3 Extend the test process to new - more representative, users and ensure that the tips work in practice.

RULE 6: TEST MARKET AND MEASURE THE BENEFITS.

Test marketing also gives us time to stabilise systems before we stress test them. We can live with a testmarketing system which stumbles occasionally.

Stage 4 Once you have proven results launch the material to the rest of your potential users.

Stage 5 Look for other opportunities to generate similar data.

Change of Address

When a client changes his address it is normally a sign of a change in circumstances and time to review the insurance portfolio. Under the old system we would immediately send an intermediary a “change of address lead”.

Each family member was treated separately, and each line of business acted independently. If an address change was not recorded for one family member nobody would know. The absence of a change of address may be an omission or it may be the splitting up of a family.

Under the new system we can issue a comprehensive statement indicating all the members of the household for all lines of business. This shows where a change of address has been recorded and where not, putting the intermediary into a
strong position to offer a real service and to uncover sales opportunities from any one of several prospects. The fact that changes of address are now updated daily also means that the lead can be much more current.

Since far fewer pieces of paper are issued, and fewer intermediaries try to make contact, the company gives a much better image of efficiency ... and at much lower costs. This is an important consideration when one is handling more than 1000 changes every day.

The new lead also lists all current servicing agents and key information about them to help guide the manager in the new area to allocate the lead to an appropriate intermediary.

**DATAMINING**

Whilst our main team commenced the building of the new datawarehouse, a second team was given the opportunity of testing and experimenting with the Beta version of SAS Institute's Enterprise Miner™.

This new tool was tested alongside our existing analysis package, and demonstrated that although it is slower and more difficult to use, it is also capable of much more complex analysis and of providing more reliable and more useful information.

**RULE 7 ENSURE THAT THE TEAM HAVE APPROPRIATE SKILLS FOR DATAMINING**

At a very early stage we identified that if Enterprise Miner™ - or indeed any advanced statistical technique - is used without proper knowledge it can provide amazingly misleading information.

We have got our most useful results using a combined team. An “expert” from SAS Institute to introduce us to the techniques and to guide us in the use of the tool. A Professor of Statistics from a local university to help us understand the true mathematics.

Our resident research analysts who are relatively used to the data, where it comes from and what it means.

The end users who had to “discover” new guidelines and then turn them into profitable business interventions. The end users were under particular pressure from top management to be seen to be delivering valuable activity.

**RULE 8: MAKE SURE YOU USE THE RIGHT TOOL FOR ACTIONABLE RESULTS**

Mining in relation to new business persistency has allowed us to identify a number of different sets of circumstances which require different management responses.
We have identified some indicators that suggest that salesmen should be discouraged from being active in particular market segments. It is generally impractical to impose a blanket ban, even though a segment may have a high discontinuance rate, since that might be discriminatory against the few who would be lasting prospects and might also impact on the commission earning agent’s livelihood. Discouragement can take the form of a suspension of commission until the client has demonstrated his viability.

Similarly if a client is assessed as having a poor probability of becoming viable it may not be worth instituting any action to rescue the case if premiums are missed.

This highlights an important difference in the application of datamining tools.

If we wish to discourage activity in a particular market we must be able to identify a few clearly recognisable features which can be described to the intermediary. Only certain classes of product should be sold to new clients aged under 28 unless they are family members of longstand clients ....

Where we are going to withhold follow-up actions we only have to establish a probability of long term viability, without being too concerned about the underlying dynamics, and can use less descriptive tools such as neural networks and also "confidential" information which might not be available to the intermediary.

**RULE 9: EXPERIMENT AND MEASURE BEFORE MAKING LARGE COMMITMENTS**

In doing this work we learned some valuable lessons in how to use the tools and how not to use them... and this was all done before the main warehousing initiative began to deliver any functionality.

The new systems are also making it easier to attach additional external data, particularly in regard to the client. This includes credit information which appears to have a high indicator value in the probability of the sale remaining viable. By doing this in the context of a datamart set up for mining purposes we can assess the true value of potential additional data before going to the expense of making it an integral part of the full datawarehouse.

**RULE 10: MAKE SURE YOU REALLY KNOW WHAT THE DATA MEANS**

Fields need to be properly labelled so that the people working with them know what is in them ... but you must be very careful that easy to understand descriptions don’t become misleading by providing a false sense of certainty.

Metadata is generally recognised as important but users don’t go and look at the metadata if they think that the answer is obvious. It is also important to recognise that past actions will change the current trends. Actions taken in aspect of past symptoms may mask the underlying dynamics.
QUICKER ACCESS TO SAMPLING
Power sampling has allowed us to look at some situations much earlier and easier than we had expected. We know that in an average month only between 1 and 2 per cent of our clients purchase from us, so we would need to look at sales over a considerable period to be able to do conventional analyses.

However a single month’s “repeat sales” amount to 20-30 000 clients, and a power sample comparison against an appropriate number of non-purchasers is quite sufficient to get us started - Obviously we couldn’t look at seasonality without a year’s business but there are a lot of other pointers that can be established.

The SAS System also gave us another facility which has allowed us to create instant history whilst still using the initial snapshot databases.

Looking at our current snapshot of our data base we have been able to identify groups of policyholders who have made some significant change to their portfolio - buying a new policy, discontinuing a policy, etcetera.

We have then been able to reload the snapshot of the database as it existed a few months previously and identify the situation before the change occurred.

Taking this further we can even begin to establish a time series over many snapshots.

Naturally this “history” makes it necessary to create a whole lot of new axes or dimensions to represent the changes which have taken place in regard to the client’s circumstances, the intermediary’s circumstances and the portfolio.

Utilizing these additional axes in datamining immediately highlights the fact that the true predictors of changes often relate to the dynamics of the intermediary and to changes in the client’s personal circumstances rather than to absolutes of any of these values.

RULE 11: USE A GUIDED MISSILE, NOT A RIFLE OR SHOTGUN.

It is often said that proper use and understanding of your database allows you to use a rifle rather than a shotgun - you can select your target and aim directly at it with a high degree of accuracy. In reality I believe that a proper tool set allows you to have a guided missile. Once you have a reasonable idea where your target is you can fire off a missile in the general direction, and then gradually adjust its course to take into account experience and improved knowledge as you go along.

RULE 12: PROVIDE A CONTINUING STREAM OF BENEFITS - IT KEEPS UP THE ENTHUSIASM AND SUPPORT
Let me summarise what Old Mutual has achieved over the past year.

We started the proof of concept in July 1997. This involved converting the Portfolio Extract database into SAS sets, and creating new selection systems. The test delivered results in August which signalled significant cost and time savings. We started industrialisation in September. By December the new system was sufficiently robust for us to use it and to switch off the old system.

By this stage, we had started to receive real benefits - lower cost, fast turnaround, less queuing, complex extracts completed in a single pass and the markback of history.

We continued to tune the system on an as-and-when basis and by the end of March were able to have the alternative of running on the server without any change of code - To the user this means greater flexibility and again potential cost savings.

In the meantime the new team was set up in December to start building the new warehouse, and full requirements were gathered. Then after the disruptions of Christmas and New Year development started in January 1998, and the first generation of data was available in March.

This gave us the full information on the combined portfolios of three lines of business and the ability to look at households. Trickle feeding of changes commenced at that time. This means that any mailings or extracts generated from this file would have the latest information, thus avoiding a lot of wastage. Work immediately commenced on industrialising the output systems from this new datasource, and the segregation of information for the different lines of business.

The Beta version of Enterprise Miner™ was also loaded towards the end of January 1998 and immediately set to work on the initial persistency analysis. Work on purchasing patterns based on matched back history commenced in April; as did the exploration of the new combined database.

CONCLUSION
I think that this has been a real success story although things definitely didn’t go quite as easily as we hoped. We made many mistakes - often temporally negating the promised savings, and many of our processes still don’t work nearly as efficiently as we would like.

We make no claims to have a best in class solution, but have justified the presentation title “Rapid Business Benefits for Client Relationship Management through datawarehousing and mining”.