



Detection and analysis of abnormal behaviour

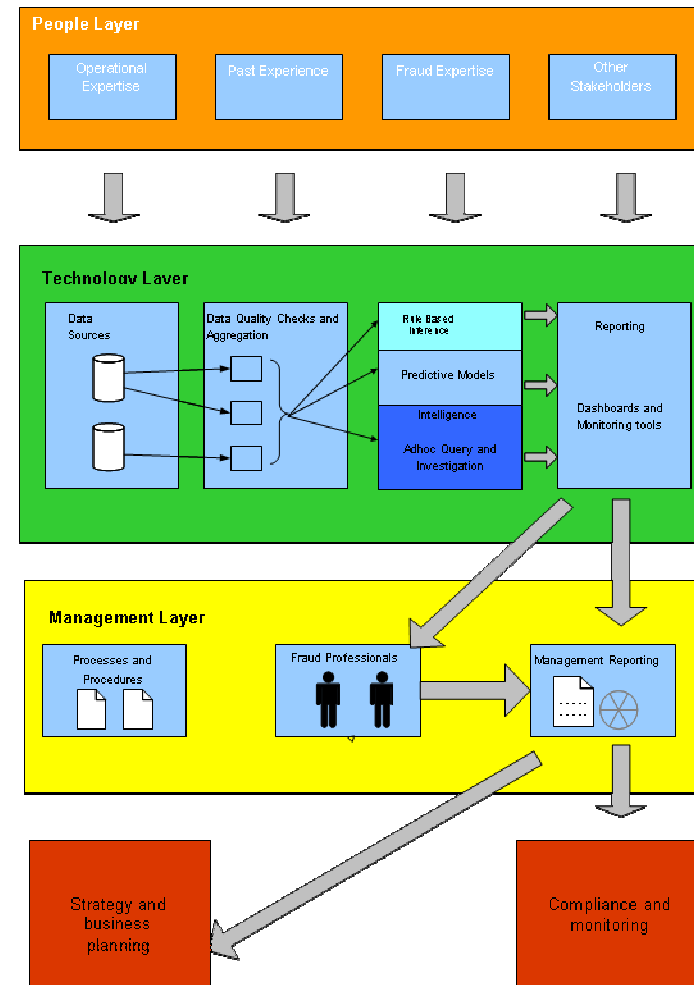
Insurance and banking examples

Detecting fraud

Fraud is an ongoing problem
Companies need to manage fraud to:

- ▶ Meet business objectives and performance requirements
- ▶ Comply with regulations

Unknown unknowns are the hardest aspect to address



Unknown unknowns

“There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns -- the ones we don't know we don't know. ... it is the latter category that tend to be the difficult one”

Donald Rumsfeld, 12 February 2002

Standard fraud control framework:

- ▶ Examine patterns from previously identified frauds
- ▶ Define internal controls to detect similar frauds in the future

Prevents or detects repeats of existing fraud mechanisms

But what about new mechanisms?

- ▶ Need a method to detect the “Unknown unknowns”

Anomaly detection

Don't focus on detecting fraud

Focus on characterising behaviour

- ▶ Typical behaviours will have similar characteristics
- ▶ Anomalies or abnormal behaviours will have different characteristics
- ▶ Anomalies are defined in contrast to normal behaviour

But, anomalies are not necessarily fraud

- ▶ Can be a rare form of behaviour
- ▶ Expert investigation is still required
- ▶ Analysis helps focus limited resources on appropriate areas

Characterising behaviours

- ▶ A level for analysis has to be selected:
 - ▶ Customer
 - ▶ Employee
 - ▶ Account
 - ▶ Policy
- ▶ Need not be related to fraud controls
- ▶ Should bring behavioural differences to the surface
 - ▶ Ratios
 - ▶ Entropies
- ▶ Should avoid spurious differences
- ▶ Not too many, not too few
- ▶ Consider transformations (log, square root)

Seeing the forest for the trees

In all analysis, it is important to be able to visualise the data:

- ▶ Helps to develop a feel for the data set
- ▶ Avoids mistakes

But how do we visualise hundreds of variables?

- ▶ Principal components analysis (PROC FACTOR, PROC PRINCOMP)
- ▶ Variable clustering (PROC VARCLUS)
- ▶ Other dimension reduction techniques

Tips:

- ▶ Check proportion of variance explained!
- ▶ Ensure variables are scaled!

Classifying types of behaviour

Typical data sets will have multiple types of normal behaviour

Differentiating between these can improve anomaly detection

Clustering techniques can be used:

- ▶ PROC CLUSTER – ok for smaller numbers of instances (e.g. analysing staff)
- ▶ PROC FASTCLUS – better for larger numbers of instances (e.g. analysing customers or accounts)

Manual methods can help as well

Always perform visual check

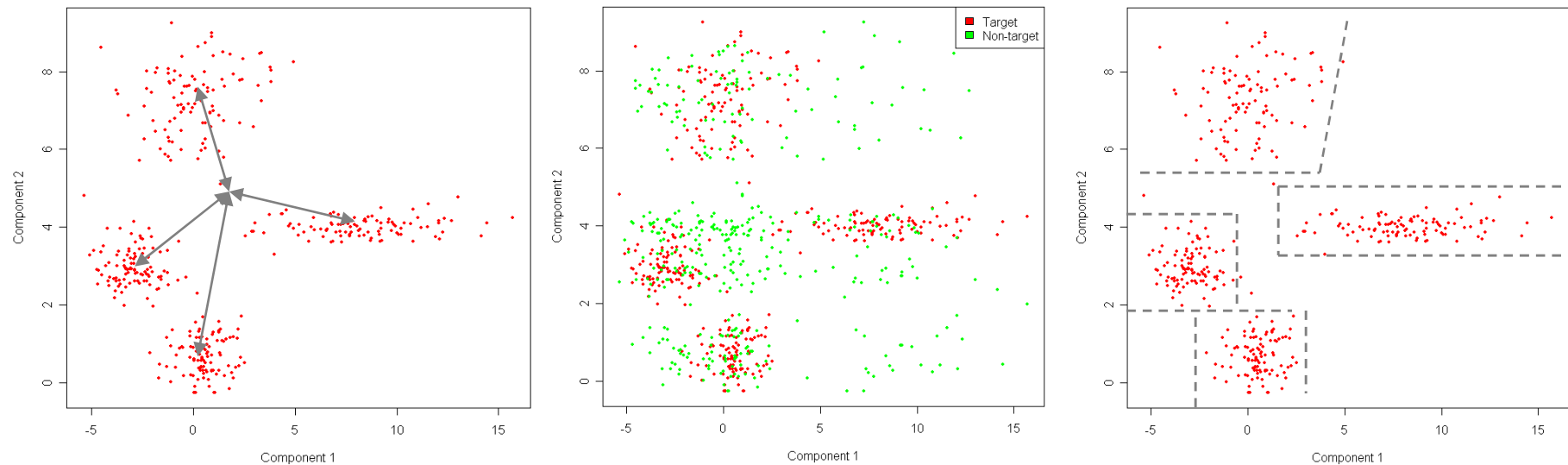
Identifying anomalies

Distances to cluster centres

Classification rules based on cluster allocation

Classification against artificial “background” data

Manual rule sets



Interpreting the results

The process so far has only identified anomalies – these are not necessarily frauds

A fraud expert should review the anomalies

- ▶ Without business contextualisation they are not likely to be implemented so this is an essential part of the analysis

Data required:

- ▶ Overview of each behavioural cluster
- ▶ Summary details for each anomaly
- ▶ Detailed data for drill down

Consider developing a set of “fingerprint” variables

Putting it into production

Data preparation process

- ▶ Consider production during development
- ▶ Could make use of SAS stored processes

Detection of anomalies

- ▶ Consider ease of implementation

Data volumes – how many instances can be reviewed with current resources?

Schedule for updates

- ▶ How quickly does the business change
- ▶ Review as frauds are detected

Insurance example

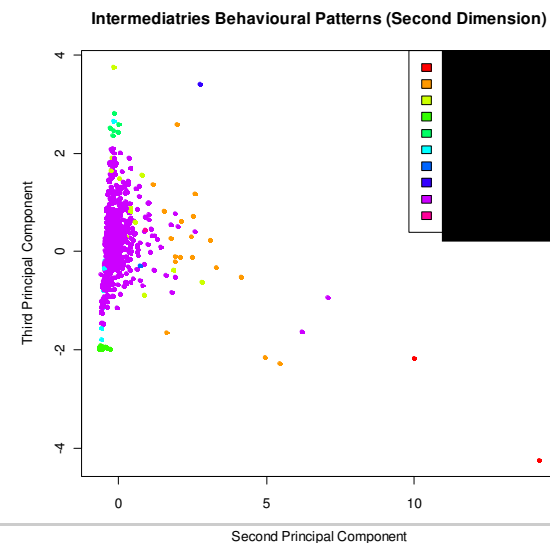
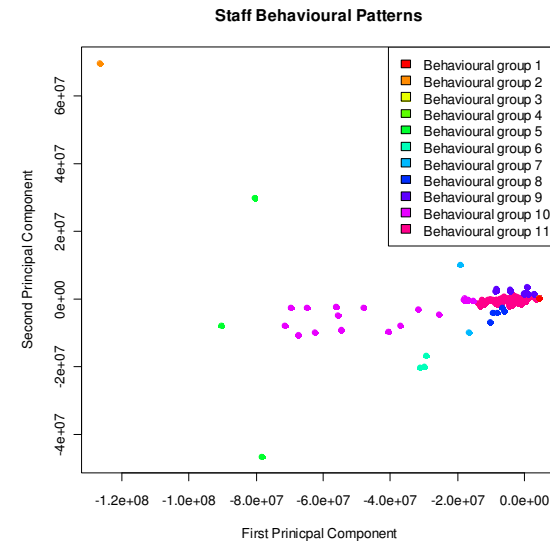
Pilot program for an insurance client

Focus on small to medium business insurance product

Investigation carried out at two levels:

- ▶ Staff member
- ▶ Intermediary

Detailed follow up on key anomalous behaviours identified
some areas for further review



Banking example

Analysis of transactions in suspense accounts

Unit of analysis was the transaction batch processed by the backend system

Key desired outcomes:

- ▶ Understanding of the major types of suspense account behaviour
- ▶ Understanding what behaviours are worth investigating further

