What is data mining?

Data mining can almost literally be described as digging up information from databases. Information that was never explicitly put in, is brought to the surface by advanced techniques.

Without data mining, data analysis is limited to a small number of selected aspects. With data mining, all possible relations within the data are checked. By using artificial intelligence, statistical methods and visualisation, relations are discovered that would otherwise be easily overlooked. Furthermore, all relations that are found can be applied to make predictions.

“Using data mining in analysis results in more effective and goal-oriented control on strategic, tactical and operational levels within the police force.”
[Abrio, Management office of Dutch police]
How does DataDetective help?

In this digital era, police forces have access to quickly expanding sources of information. The enormous increase in the amount of available data has made the use of data mining techniques essential in finding important patterns.

DataDetective enables analysts to gain insight in thousands of variables from various police systems, census data, spatial information, weather and lifestyle data (Experian).

All the data is maintained in a data warehouse. This means that repeated data collection and cleaning is not required for every analysis. Automated procedures have been developed for the extraction, linkage and conversion of data in the data warehouse. These procedures are included in the product.

The data mining experience of police forces has resulted in best practices: successful methods of working with DataDetective. These ‘recipes’ can be accessed from within the software.

“DataDetective is being used more and more in our police work. It adds important value because insight in trends and developments is gained easily, enabling the organisation to react quickly and adequately. This leads to a more efficient, more focused and overall better tracing of criminals.”

[Mick Hoff - Police Inspector]

Applications of DataDetective

- Find series of incidents with similar MO and witness’ descriptions.
- Find possible suspects by comparing cases with incidents from the past and linking to the suspects of those incidents.
- Detect emerging crime trends in detail
- Find relations to explain a trend or certain behaviour, e.g. find causes for an increase in burglary, or find combinations of factors that account for a person becoming a repeat offender.
- Apply link analysis on networks of criminals and finding the ‘spiders in the web’.
- Easily define selections without the need for a query language such as SQL.
- Create data counts and graphs.
- Create cross-tables (OLAP) to compare data from various sources.
- Easily create a map (point map, hotspot map, difference map, area map).
- Make selections in the map to further analyse the data.
- Geographic profiling
- Create a forecast map of a specific type of crime in a given context: taking into account the time, day of the week, season, period (e.g. holiday), the weather and recent trends.
- Text mining, by including textual data, e.g. witness statements, in the analysis.
- Use fuzzy matching to get more results by looking for partial matches.
- Find combinations of location, time and circumstances where a specific type of crime is more likely to occur, to plan preventive measures.
- Build and apply self-learning models, for example to predict dangerous behaviour based on previous situations, or recognize domestic violence incidents that might escalate.

From DataDetective, information can easily be exported to Word, Excel, SPSS, MapInfo, Analyst’s Notebook, Google Earth and WEKA. The product can also be linked to web applications and external reporting tools, such as Cognos Reportnet.

Practice shows that police officers can use DataDetective to execute strategic, tactical and operational analysis from start to finish, within one integrated environment. Analysis and queries are performed in a fraction of the time normally required. Because of the greater depth that can be achieved in the analysis, surprising results are often found.
Example session

1. Specify a selection of all street robberies.
2. Create a chart of this selection over the past months.
3. Show the graph in Excel to include it in a report.
4. DataDetective finds a significant increase in the past two months.
5. Find an explanation for the increase by performing a profile analysis of the last two months against the months before. This immediately shows that there is a significant increase in street robberies performed by 2 persons on a scooter in the Central Station area, especially between 0:00 and 2:00 AM.
6. Zoom in on these incidents and create a map in MapInfo, showing the highest concentrations. It is decided to increase surveillance in those concentrations and time frames. Now, possible suspects can be determined.
7. Use fuzzy matching to search for incidents with similar methods, profiles and locations in the past. Find the suspects from those incidents and show their pictures to the robbery victims and patrolling officers.
8. By using network analysis, possible associates in their criminal environment are identified.
9. The resulting selection of persons is copied to a report in Microsoft Word, with pictures and personal data, to be used in briefings.
10. With Analyst’s Notebook, the data can be augmented with pictures and descriptions to provide more background information.

Success stories

- User experiments with DataDetective show an efficiency gain of factor 10 to 20.
- Field tests have shown that DataDetective has a 50% higher success rate when looking for suspects compared to standard methods.
- Forecast maps from DataDetective are twice as accurate as forecasting with conventional hotspot maps.
- A comparative study has shown that criminal activity dropped over 15% due to measures that were based on DataDetective results, while activity remained the same in districts that were not using DataDetective.
- The first application of DataDetective forecasts in Amsterdam led to the lowest report rate of street robbery ever.
- DataDetective is used to plan preventive stop-and-searches in Amsterdam, which has brought possession of weapons down by 27%.
- Reports created with DataDetective are distributed to hundreds of people within the police force on a weekly basis, resulting in a fast and uniform spread of information, saving analysts much tedious work.
Why DataDetective?

DataDetective satisfies the demands that Dutch police forces set for a data mining product:

- Special expertise or new personnel is not required.
- Options for advanced search.
- Support for searching text (e.g. statements).
- Integration with standard tools (e.g. MapInfo, Excel, Analyst’s Notebook).
- Spatial functions for prediction, time/pace clustering and profiling.
- Easy extensibility of the software.
- Use of open standards to access databases and to share results.
- Support of rich data structures.
- Ability to authorize functions per user group.
- Results need to be easy to reproduce.
- Proven in practice.

Basic functions

- Define selections in an intuitive way
- Select using fuzzy matching
- Charts and graphs
- Cross-tables and decision trees
- Profiling and segmentation
- Network analysis and clustering
- Risk prediction
- Geographic forecasting and creating maps

The rich collection of techniques makes this system applicable on various levels in the police force:

- On the **strategic level** crime trends can be visualised and explained in rich detail.
- On the **tactical level** forecasts can be used to take preventive measures and give relevant instructions.
- On the **operational level** all kinds of relations can be found between people and incidents, using similarity matching and linking.

The speed of the system and the fact that most data analysis is performed automatically using the optimal data mining techniques, makes DataDetective very accessible.

More information

**Sentient Information Systems**

*Pieter Bison*

Phone +31 20 5300 322  
info@sentient.nl

**Police Amsterdam:**

*Hedda Roos*

Phone +31 20 5598 495  
hedda.roos@amsterdam.politie.nl

**Police Mid- and West Brabant:**

*Harrie Oerlemans*

Phone +31 6 2954 3791  
harrie.oerlemans@mw-brabant.politie.nl

**Police Brabant-North:**

*Jacob Baas*

Phone +31 6 5338 1061  
jacob.baas@brabant-noord.politie.nl