Agile BI: How Eandis is using SAS® Visual Analytics for energy grid management.

Olivier Goethals, Enterprise Architect and BICC Manager, Eandis cvba, Belgium

ABSTRACT

Eandis, a Belgian energy distribution grid operator, is using Visual Analytics to explore data from various sources, including SAP Business Warehouse, to better plan maintenance and investments in the grid. To achieve this, a new agile way of thinking about Business Intelligence was needed.

INTRODUCTION

Eandis is a rapidly growing energy distribution grid operator in the heart of Europe, with requirements to manage power distribution on behalf of 229 municipalities in Belgium. With a legacy SAP data warehouse and other diverse data sources, business leaders at Eandis faced challenges with timely analysis of key issues such as power quality, investment planning, and asset management. To face those challenges, a new agile way of thinking about Business Intelligence (BI) was necessary. A sandbox environment was introduced where business key-users could explore and manipulate data. It allowed them to have approachable analytics and to build prototypes. Many pitfalls appeared and the greatest challenge was the change in mindset for both IT and business users. This paper addresses those issues and possible solutions.

THE PAST, THE REALITY, THE WISH AND THE SOLUTION

THE PAST

At Eandis business intelligence was limited to classical reporting. Each day, week or month data was transferred through ETL towards a large data warehouse and each month a number of reports were extracted containing lists with indicators and numbers.

To develop those reports business needed to contact ICT who then had to develop the ETL and fill up the data warehouse. The only method for doing this was a classical waterfall development methodology where ICT analysts interviewed business representatives to capture the requirements, analyze those requirements, develop the extractions and fill up the data warehouse and finally build the reports and deliver them to the business for testing and deployment.

THE REALITY

Many actors were not happy with this situation. In reality there were many shortcomings to this approach:

First of all it has been apparent for years that a single "enterprise data warehouse" was an illusion. It was a holy grail that many have sought but no one ever reached. A single enterprise data warehouse requires a single enterprise wide information or data model. And this was unachievable. Many applications had their own datamodel, and attempts to map it to a common canonical datamodel took so long that the canonical datamodel had changed by the time development was done. Even more difficult were data coming from external sources or from other players in the market did not adhere to the common datamodel. In reality trying to force every single application into a common model remained only an idea.

In addition, the waterfall method for development proved troublesome for certain BI development. Often reporting was considered an "additional" or "nice-to-have" requirement as part of larger projects. The implementation of a new platform or system, for instance a new CRM system, required several months if not more than a year. And during the initial requirements analysis phase, end users had to come up with reporting requirements more than a year in advance of actually seeing the system or getting to know the

data. There were three possible results: either the users gave random requirements which were correctly built and delivered but in the end turned out to be useless, or the requirements were formulated so broad ("just give me a report containing all the data") that IT had difficulty delivering it, but more often the third possible outcome was that the development of reporting was pushed towards the end of the project and when delays jeopardized the release data reporting was skipped and would be done "later".

Business users then sought other means of gathering insight, which in most cases resulted in shadow-IT, homemade ETL, and excel reporting.

THE WISH

The business had several demands for a new approach.

First of all they wanted more self-service. Not only did they want the ability to make or change the reports, but certain key users needed the ability to upload or even change and transform certain source data. Especially in the analytical area, classical reporting that only provided OLTP data or at best, historical data, possibilities of analysis, exploration and experimenting was limited.

Secondly, the usability and interactivity of the current reporting environment was too restrictive.

And finally, a profound new way of thinking about the reporting development and lifecycle was needed. A more agile approach was requested.

THE SOLUTION

A public tender was issued and after a selection process, SAS® Visual Analytics was selected as a cornerstone of the new BI environment that was proposed as a solution.

This new environment consisted of three main parts:

Figure 1: Three components of the solution.

SAP Business Warehouse: the current reporting solution remained. Not only was all the development re-used, but even more so, it was upgraded, shortcomings where corrected and performance was enhanced. Because Eandis has a rather large SAP application footprint, the most logical go-to solution for classical reporting was SAP BW in many cases. As a result SAS® Data Surveyor for SAP was acquired to link SAS® Visual Analytics to SAP BW.

SAS ® Visual Analytics was introduced as a new approachable analytics platform, to be used in addition to SAP BW solution. SAS® Enterprise Guide® was introduced for a limited set of key users to allow them to perform end-user data-transformation. The IT department uses SAS® Data Integration Studio for industrialization of the ETL processes if required. (Remark: for centralized data scientist in BICC also a few instances of SAS ® Enterprise Miner ™ and SAS® Visual Statistics are used for specific types of analytics)

A **BI competence center** was launched to allow an end-user community and IT specialist to collaborate and help each other in a more agile approach.

SANDBOX

A key part of this new agile approach was the so called "Sandbox" environment.

As with all systems in use in Eandis a minimum of 3 logical environments were needed: a development environment, a test environment and a production environment.

Figure 2: Basic environments
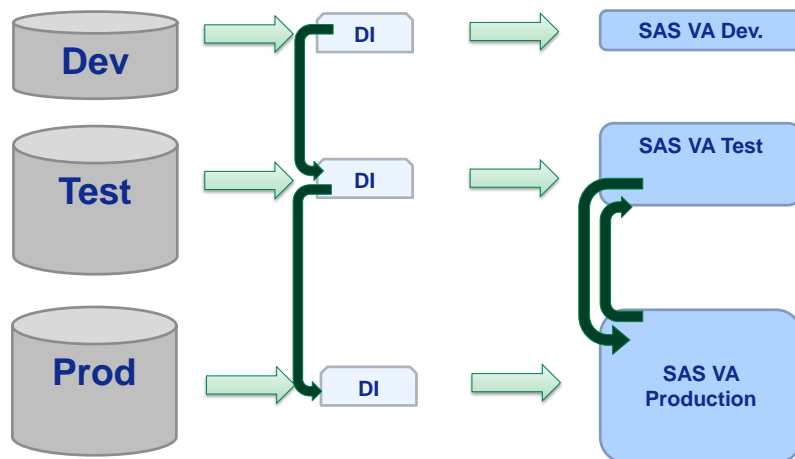
The source data was mainly SAP BW, but also other possible sources of data, such as copies of Oracle database are possible and used. The data is transformed and uploaded to the SAS® Server using SAS® Data Integrator (in addition to the SAS® Data Surveyor for SAP). All this is executed by IT.

Migration from development to test and finally into production is also executed by IT.

Because end-users have the ability to develop their own reports and dashboards in the production environment, it was required that these production reports could be migrated to the test environment to allow end-users to test their reports against enhanced or newly developed source systems and source data.

In addition to these three classical environments a fourth environment was created: a Sandbox, to allow exploration.
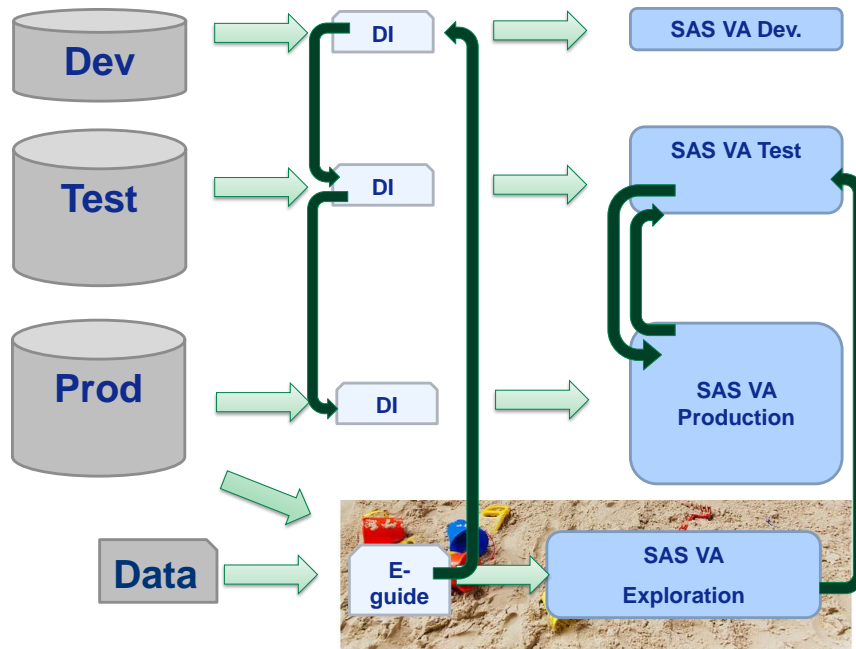
Figure 3: the extra exploration environment (sandbox principle)

In this exploration environment, certain key-users were allowed to use SAS® Enterprise Guide to transform production data or upload their own data sets and use SAS® Visual Analytics or SAS® Visual Statistics or even SAS® Enterprise Miner ™ on these data sets. If the resulting transformations and reports are deemed necessary to be used by a large population of end-users in the production environment, the transformations and reports are handed over to IT to go through the normal development cycle.

WHEN TO USE THE SANDBOX:

The sandbox is used for two purposes:

Prototyping

The sandbox can be used as a prototyping environment during the agile reporting development cycle. Even before source data is available, end-users can upload dummy data using excel and already simulate and experiment with the type of reporting and visualizations they would like to have.

Exploration and analytics environment

The main use of the sandbox environment however is as an exploration environment to allow end-users to perform more advanced analytics.

Typical analytics require data to be transformed, enhanced or mixed with external data sources. It is however not always certain in which direction the exploration will take you. It is clearly an iterative approach where data is uploaded or transformed, correlations and patterns are detected, and predictive or descriptive analysis models are constructed, tested and enhanced.

The sandbox environment allows key-users to access all production data (it is sized identical to the production environment) and allows them to upload own data sources. In addition, using enterprise-guide, key users have the possibility to transform the data and add the typically required analytical columns to build analytical columns. They can test a certain hypothesis, build a model and do explorations.

Most of our key-users do not have, nor require, advanced and highly complicated analytical processes. SAS® VA and SAS® Enterprise Guide were chosen exactly because of their approachable and user-

friendly interfaces allowing non-statistical end users (so called data-citizens instead of data scientists) to perform basic analytics. From within the BICC more advanced statistical support can be given if required.

The usage of the sandbox environment is not unlimited. Not only do the users of the sandbox need proper training, they also need to justify the use case and the usage of the sandbox is limited in time and size depending on their use case.

BICC

The introduction of the sandbox, the agile development life-cycle and the ability for end-users to upload and transform data has created serious resistance and required a rather big change, not only in the ICT population but also in certain business departments. The BI Competence Center was formed to help overcome some of this resistance and to address certain risks and issues put forward by the management of Eandis.

The goals of the BICC are:

- Knowledge sharing and community building
    - o In general end-users and ICT share what is happening in the BI area, who is doing what and why, so that everyone knows what data sources are exposed and how they are used.
    - o Training: end-users who are using the self-service exploration environment get training, not only about the tools, but also about the concepts of BI and analytics.
- Governance and management of roles & responsibilities (access rights, exploration environment…)
    - o Not everyone gets access to the sandbox or has the ability to change reports in production. The BICC manages these roles and responsibilities.
    - o Governance of BI initiatives is aligned with other information initiatives such as enterprise information management, master data management, enterprise content management or identity and access management. The strategic governance of all information related project has been centralized under the control of one steering group to assure alignment.
- Style guide
    - o To help end-users build their own reporting and visualizations but also, to avoid a kaleidoscope of colors a style guide was developed to guide the usage of colors and charts.
    - o The different environments and especially the sandbox environment have a distinct style and color to avoid confusion between environments.
- Quality assurance
    - o To address concerns about the quality of reporting, the results of the sandbox before being put into production get a quality label. Everything is allowed, but a clear quality label indicates if the report has gone through due-diligence and quality checks by the BICC or not. Tests include data quality of the source data, awareness and agreement of the information owner, adherence to the style guide etc…
- Usage monitoring and reporting
    - o The usage of the reports, the performance of the queries and the usage (and volume) of the sandbox are being monitored and reported on. This allows key-users to measure the success of their reports but also helps to manage the exploration environment, occasionally clean the sandbox and in the future possible introduce time-sharing or even cross-charging of the exploration environment if conflicts should arise.

POWER QUALITY MEASUREMENTS OF TRANSFORMATION STATIONS

All transformation stations providing electricity infeed in the power grid of Eandis are constantly measured. Every 15 minutes energy flows are measured at several points. This results in close to 10 million lines of measurements each year, that need analysis and monitoring. Processing this information took typically 9 to 12 months, using several excel sheets and aggregations to cope with the large amount of data.

Monitoring the load of these transformation stations became more important over the recent years due to the increased local production. Analyzing patterns, monitoring active and reactive power and carefully planning investments required a more detailed insight in the load distribution on these transformation stations.

After getting a crash course in SAS® VA and putting the power quality expert next to the key-user, using SAS® Enterprise Guide and SAS® VA, these two users were able to produce visualizations and insight in a matter of hours instead of months, without help from IT.

Their analysis included profiling of transformation station patterns, analysis of behavior of substations and predictions of future evolution.

All measurement can be represented in profile heatmaps showing usage per hour, per day, per month, seasonal patterns or weekly patterns.

In the following images you can see a typical industrial area with more load during the day, a residential area with more load during the evening and finally an area near the coast that shows typical weekend housing.

6
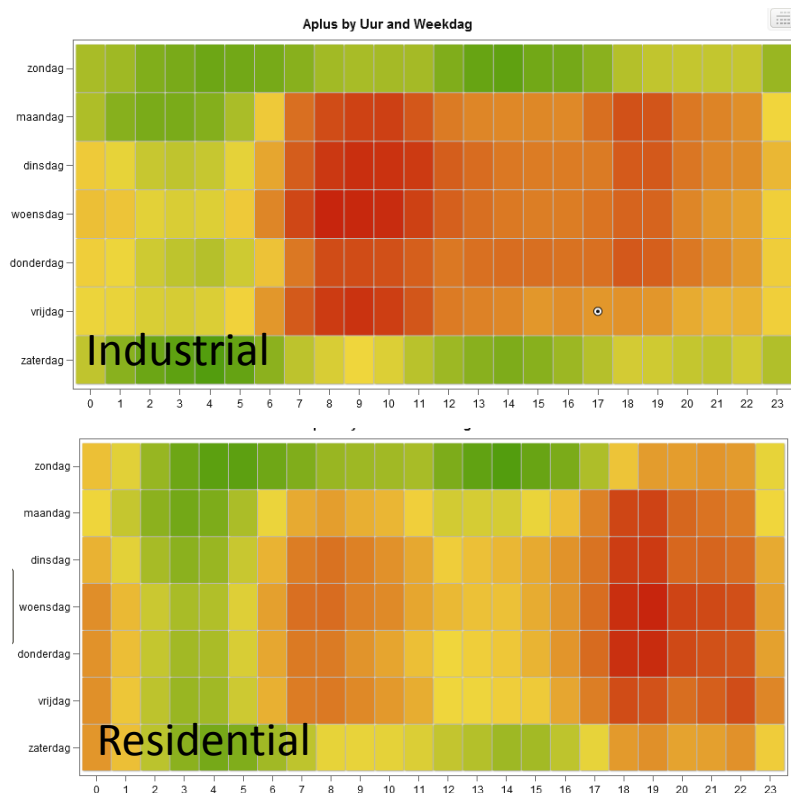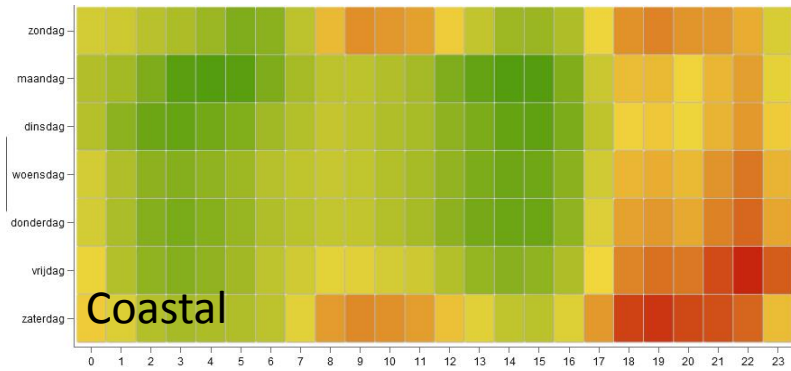
Figure 4: Profiling of substations

In addition to this profiling, a more thorough and user friendly time series analysis was possible as well as a detailed analysis of the behavior of each substation.
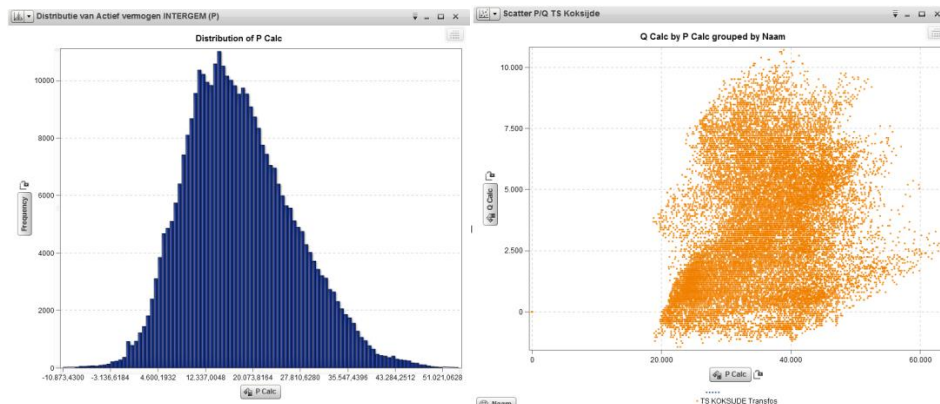
Figure 5: Detailed analysis of substation

Imbalances in the active and reactive power of each substation could be visualized per time slice or globally. And imbalances can be shown on geographic locations.

A key business user commented after the project: "For me and my team a new era of working has started. Now we can help investment planning with detailed and sound advice in a matter of hours on each individual case, with precise estimates and forecasts, instead of spending months on a general advice based on hunches and aggregate data".

CONCLUSION

The use of a separate sandbox exploration environment, production sized, with all production data available, and allowing key-users to add and change data using SAS® Enterprise Guide in this sandbox environment is a powerful tool for business experts. However, it requires a big change in mentality and attitude for both IT and Business. A BI Competence Center needs to address this change by introducing governance, help, knowledge sharing and support.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Olivier Goethals
Eandis
+3292638427
Olivier.goethals@eandis.be