Online Visual Analysis Tools

Diab Mahmoud, Minwegen Jürgen, Perrogon Georg and Plieschnegger Markus

Institute for Information Systems and Computer Media (IICM),
Graz University of Technology
A-8010 Graz, Austria

21 May 2014

Abstract

The amount of accessible data is constantly increasing. Therefore the role of tools to process and present information is more and more important. Within this paper the description of four of the leading data visualization programs, namely Tableau, QlikView, Spotfire, and Omniscope, is presented. Also a very short overview about the meaning of Data Visualization is given. To show some of the capabilities of said programs, a data set provided by the city of Graz, is chosen and a small insight about the financial transaction of 2012 and 2013 is provided. Although those business intelligence products are quite expensive they are worth their money. The amount of time saved is quite impressive. All four programs better suited for different kinds of usage whereas Tableau stands out with its intuitive handling and overall good selection of features.

Contents

Chapter 1

Introduction

1.1 Motivation

More and more data needs to be processed and presented in an easy to understand way. Today business intelligence solutions are a common part in a lot of companies, especially if they depend on marketing strategies. Therefore data mining is vital but looses relevance as long as there is no suitable post processing that allows any kind of understanding of the collected information. Finding relevant correlations in data accumulations that are not processible by hand any more because of their size is a non-trivial but essential task in these environments. Applications like recommendation systems, pattern recognition, search engines, etc. use statistical tool to handle these situations. While automated handling of big data is absolutely crucial these days, it only covers a part of the whole topic: The possibility to visualize data in a perspective as needed and interact in a way to find exactly these points of interest without having to scan the whole data set by hand is a necessary ability for applications like decision support systems. Programs that provide such functionalities are extremely powerful tools that get even more powerful when they are easy to use, scaling well to the situation in which they are used and supporting the user as good as they can without interfering the work flow.

Like Colin Ware stated: "Humans acquire more information through vision than through all other senses combined." For this purpose there are many different programs. Some of them specify on presentation while others have a lot more to offer. Basically Information Visualization can be divided into two different categories: Data presentation and data interaction. While data presentation solely has the purpose of presenting information to viewers data interaction makes it possible to navigate through data and to concentrate on different sectors. It also helps to establish connections between data, find similarities way faster and structure data which builds the base for visual analysis.

1.2 Structure of Work

First a look into the significance of data visualization tools is given in Chapter 2, describing the basic usage of those programs. The Sections following list four examples of those tools and gives a detailed insight on each one of them. A resume is provided subsequently in Chapter 3. After that Chapter 4 treats the used data set followed by a comparison of two versions of this data set. Finally, in Chapter 5 a summary of the tools and a short outlook is given.

Chapter 2

Data Visualization

Creating plots and info graphics can be very stress full and nobody really wants to look at black and white tables anymore. Because of that there is a steadily growing number of tools available which make data visualization a lot faster and easier. Those range from very basic ones like Microsoft Excel, and Open Office Calc onto homepage analyze tools like Google Analytics and on the top end there are professional tools like Tableau and Spotfire. These programs give a huge amount of possibilities to process and visualize almost any kind of data. One of the main reasons to use these visualization tools is that it is not necessary to have a degree in statistics or art. Satisfying and good-looking results can be created within a few seconds or clicks. The most interesting ones are probably the professional tools since they offer the broadest amount of possible use. Those give the possibility not to just visualize the presented data in almost every thinkable way, they also give the opportunity to interact with the data. There are a few different ways to interact with a data set. [Yi et al., 2007] Those are:

- Select important things

- Show different data

- Reconfigure the arrangement

- Show another presentation style

- Show more or less detail

- Filter specific information

- Show related items

This gives the user the opportunity to gain information on any data set presented to him.

In the following subsections four of the most used and well known programs are described in detail.While Tableau is the leading software analytics software, QlikView, Spotfire and Omniscope are the most important competitors. The approaches are slightly different, so although the products try to achieve the same goals the solutions tend to be very different. The following sections should provide a deeper insight how Tableau (see Section 2.1), QlikView(see Section 2.2), Spotfire(see Section 2.3) and Omniscope(see Section 2.4) work in detail before doing a comparison in Chapter 4.

Tableau Desktop Private	$999	30 Day Trial	Windows
Tableau Desktop Professional	$1999	30 Day Trial	
		1 Year Student	Windows
Tableau Server	via contact	30 Day Trial	Windows
Tableau Online	$500 annual	30 Day Trial	most common browsers
			iOS
			Android
Tableau Public	Free		Windows
Tableau Reader	Free		Windows

Table 2.1: Different Tableau versions [Tableau, 2014]

2.1 Tableau

2.1.1 Introduction

Tableau is a Data Visualization Tool provided by Tableau Software which comes in a variety of different versions. Sadly there is no support for any kind of plugins to extend the functionality. Also no API is provided therefore it is not possible to create own programs.

2.1.2 Product Overview

All of the versions provided by Tableau Software with their pricing, trial time, and supported operating system are listed in Table 2.1.

Tableau Server

Tableau Server is the server solution provided by Tableau Software. It can be installed on any windows server. Because of the high performance server specifications it is possible to process huge amounts of data in an even shorter time. It connects to any other Tableau Software product. You can simply subscribe to a spreadsheet and get real time updates on your desktop, tablet or phone. It allows you to directly connect to databases or bring data in memory to speed it up even more. The administrations of the Server is possible through an online interface and for the authentication of users Tableau offers a variety of options: [Tableau, 2013]

- Local: Accounts stored in Tableau Server

- Active Directory: User names imported into Tableau Server Authentication done directly against AD

- SAML: Enterprises with central identity management Web-Based Single Sing On

- Trusted Tickets: Establish trust with another server for authentication

- Guest: No authentication necessary

Tableau Online

Tableau Online is a cloud-based solution therefore in order to work with it, it is necessary to have a working version of Tableau Server. Since it runs on the cloud, accessing on-premise data is not possible but you can push data to Tableau Online on a schedule. It is the only Version of Tableau which can run on every possible device since it supports almost any kind of web browser. Since Tableau is heavily reliant on drag and drop, handling the software is similar to the Desktop Version. Tableau Software also provides a mobile application for Apple and Android products.

Supported web browsers are:

- Android Browser (Android 3.2 or later)

- Apple Safari 3.x or later

- Microsoft Internet Explorer 7 or later

- Google Chrome, including on Android devices

Tableau Public

This is the only completely free version of Tableau which enables you to also open your own data in files. Since it is free there are a few restrictions. With this version it is only possible to save to the Tableau Public Website and not locally. Therefore everyone with a version of Tableau Public is able to see your data. It is also not possible to open data in databases. Even exporting results locally is not available. The amount of data you are able to process is limited to 1,000,000 rows.

Tableau Reader

With this version it is only possible to open Tableau Desktop Packaged Workbooks. Therefore the ability to import own data sets is not given. Like the label implies it only works as a simple reader of Tableau workbooks.

- View

- Filter

- Sort

- Export

- Print

2.1.3   GUI Overview

A quick overview about the Tableau User Interface can be seen in Figure 2.1.

1. Access point for data. It is divided into Dimensions and Measures

    (a) Dimensions: Text field
    (b) Measures: Numerical values for calculation

2. Assembly point for filters.

3. Marks gives the opportunity to change displayed data and captions

4. Legend for tables is shown here

5. Values for columns and rows of the table can be changed here

6. Menu for quick access to the different plot types

7. Main window in which the resulting plot is shown

8. Tabs for different sheets and dashboards

    (a) Sheet: Single plot
    (b) Dashboard: collection of sheets (plots)

9. Standard menu to access all functionality of the program

2.1.4 Features

Filtering

Tableau features a variety of filters. All there is to do is to drag and drop the field you want to filter into the filter area within the user interface and a pop up will appear which gives the possibility to choose the amount of data displayed. The most basic filter is a range filter for numerical values. Tableau lets the user decide if he wants to add the numbers directly into a text field or wants to change the output with the help of sliders. The slider gives the possibility to choose a maximal value and a minimal value. Or the maximum range can be chosen and the slider can be positioned at free will. Another filter would be a textual filter. With this one it is possible to search the data set for specific words. The next filter would be a check box filter which gives the ability to disable or enable different data sets. Also a condition filter is featured and if the user just wants to work with a specific amount of data it is possible to choose the desired number of values and at what position in the data set it should start.

With the so called quick-filters it is possible to include every of the above mentioned filters into your presentation. A little window appears and can be positioned everywhere within the grid system. There is also the possibility to create preset filters. After creating them they can be easily enabled or disabled with just a click on the adjacent check box. A already build in quick-filter is the calendar control. With this filter a specific date or a range of time can be chosen and the spreadsheets changes according to your selection. This raises the quality of presentations and also saves a lot of time.

Handling

The handling of Tableau is very intuitive and really fast to learn. It is mostly based on drag and drop actions. Almost everything this program has to offer can be accessed with simple mouse interaction.

Presentation Mode

One of the best features Tableau has to offer is the built-in presentation mode. This gives you a very easy way to present your data to others in a simple and fast way. The main windows is stretched to the whole screen and the background changes to white. Leaving almost no trace of the interface. The presentation mode can be activated anytime. The possibility to change in between the created sheets and dashboards makes it easy to focus on different charts at a time. the spreadsheets are fully interactive and the data information can be brought to sight with the help of the mouse-over pop up. With a simple right mouse click the functionality is enhanced even further. Simple things like showing the basic data set for your selection is only the start. It is possible to group selections, exclude them from the spreadsheet, keep only the selected part, and even some functions like show trend lines are within the range of possibility. An overview is given in 2.2.
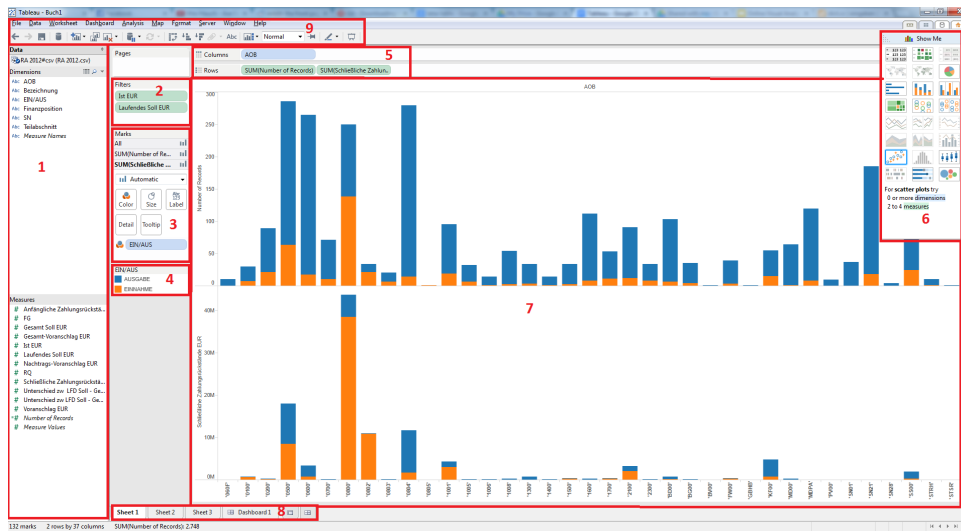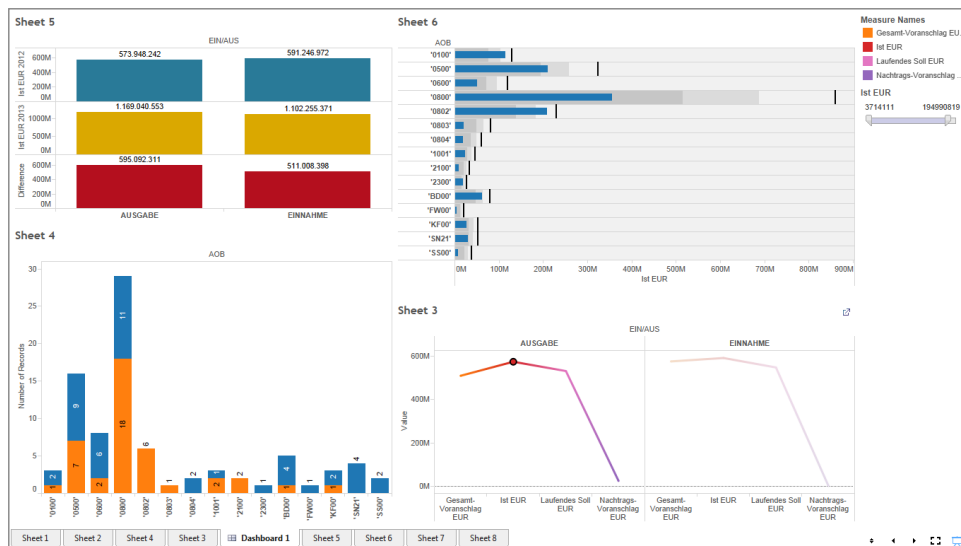
Figure 2.1: Tableau GUI after including data

Figure 2.2: Tableau GUI of the presentation mode

2.2 QlikView

2.2.1 Introduction

QlikView is a Business Intelligence Software to analyze and visualize professionally large amounts of data. QlikView is one of the most famous tools in the market that perform such tasks nowadays. QlikView is developed by the Qlik Company, which was founded in 1993 in Sweden and its main field was to help organizations to solve critical problems and make right decisions. QlikView has some features which make it preferable for some users and distinctive against the other data analysis tools. [Qlik, 2014d]

2.2.2 QlikView Products

The QlikView Business Discovery platform has 3 main components:

1. 1. QlikView Developer: Will be handeld in this paper. The tool which will be focussed on later is a local client and is called Personal Edition. With this product the developers or even the normal users can extract data and transform it and use it to create meaningful charts and visualizations.

2. 2. QlikView Server (QVS) : It is a server side Product which control the communication between the clients and server with a Management Console for Administration.

3. 3. QlikView Publisher: it is used to load data from a database or text file which is defined in the connection strings and store it.

[Qlik, 2011]

2.2.3 GUI Overview

QlikView has a user friendly interface, which can be learned and memorized easily allowing the user to be fast and creative in his working.

According to Figure 2.3 the Menu Bar is in the top of the interface and then Tool Bars which contain buttons for the commands in the menus that makes the work easy and fast. The sheet tab is considered to be the work place which contains the sheet objects like the bar chart and the list box. The sheet objects can be minimized to make more place for other objects. The places of the objects can be changed by holding and dragging. QlikView supports a lot of chart types like Bar charts, Pie charts, Scatter charts, Line charts, Combo charts, Radar charts, Grid charts, Gauge charts, Straight tables and Pivot tables.

The edit script window (see Figure 2.4) is used to load the data base tables or the table files that will be worked with. Everything that will be made graphically will be converted to a script. The Button Connect is used to open other window to set the parameters needed to connect to a database. The button Select is used to select the tables and its fields that will be used. The Reload button is used to access the source data and load it into the QlikView data model. The most important two things to create a chart like a bar chart with QlikView is the dimension and the size of the bars which will be determined with some expression, I will speak about the expressions later. In our example two dimensions are be used AOB and EIN/AUS with the following expression Count ([EIN/AUS]).

2.2.4 The Associative Search Approach

QlikView allows the user to see the data during the try to analyze and visualize it. And this will make a deep comprehension to the data. Understanding and seeing the relations between the data directly helps the user to make a professional analysis and professional visualization for the data. QlikView allows the user to connect to a wide variety of databases with the edit script window which looks like a wizard to convert the user actions to a script and can be opened from the menu file. With the table viewer button in the edit script window the user can see the tables and its relations in the associative data model created by QlikView. QlikView allows the user
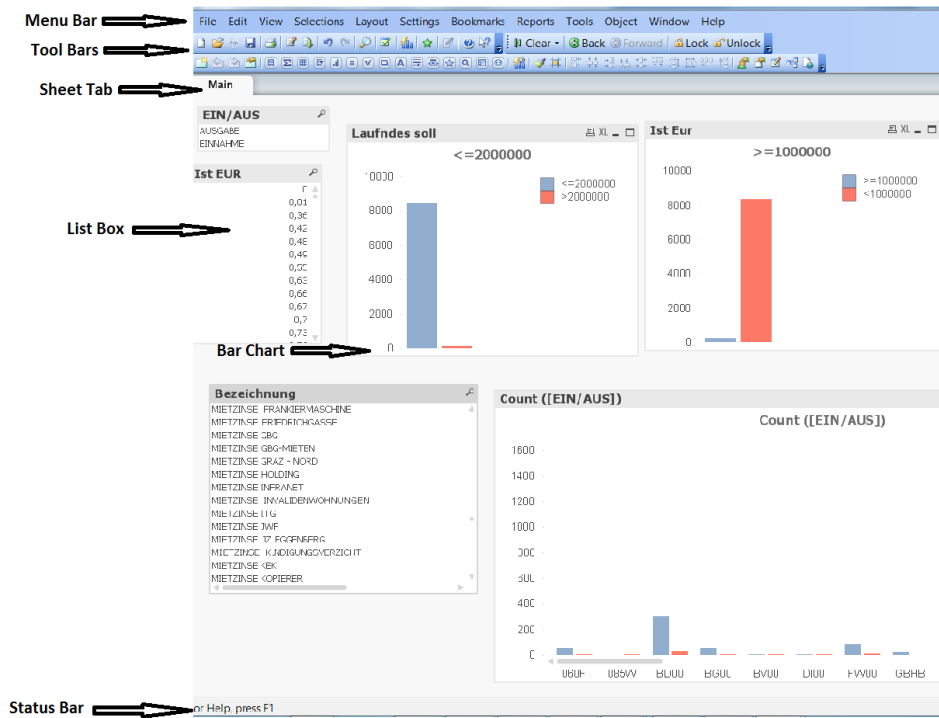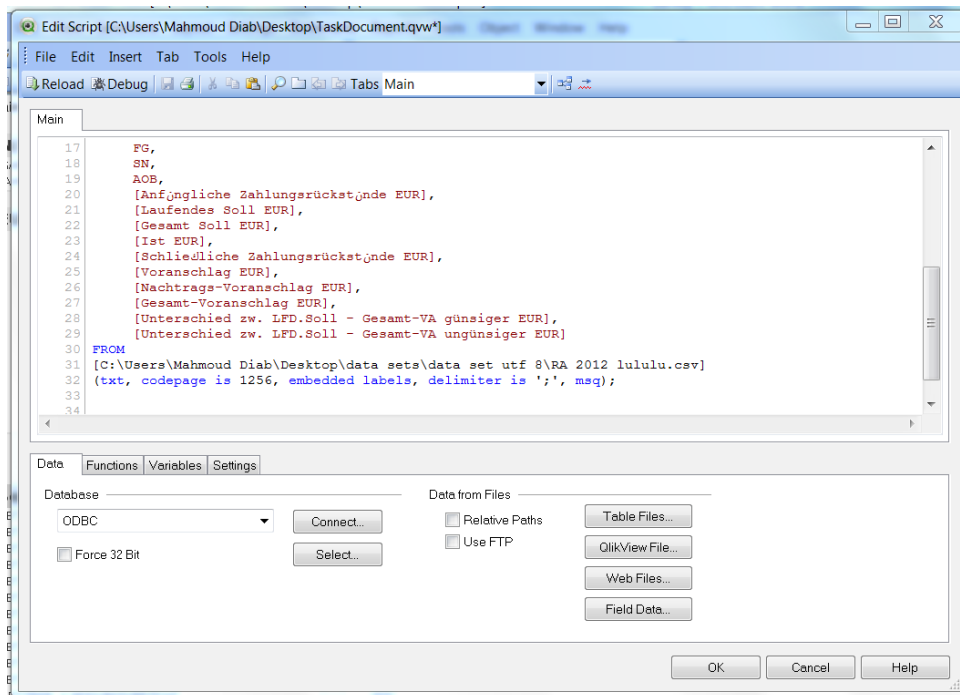
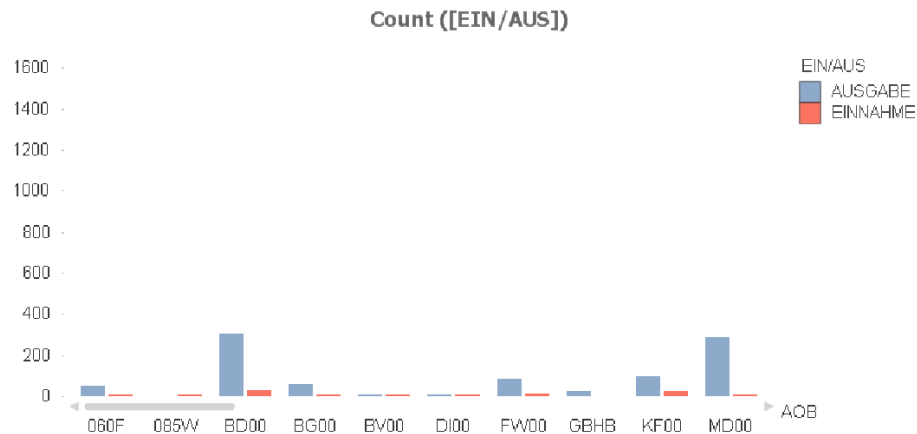**Figure 2.3:** QlikView interface

Figure 2.4: The Edit Script Window

Figure 2.5: Hierarchical visualization about the fields "AOB" and "EIN/AUS" (how many in which category of "AOB", how do these distribute in "EIN/AUS"
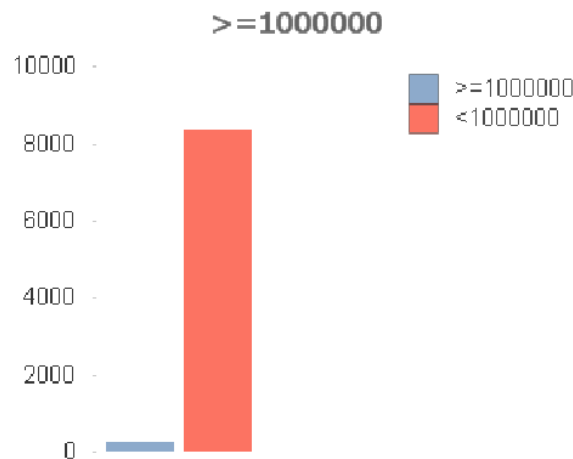

Figure 2.6: Number of records with "Ist EUR" >= 1000000

to load the data in list boxes. Meaning, every field can load its values in a list box. If the user clicks on one value he will see the associated values in the other fields selected in the white color and the excluded values will be gray. He can also notice the focus on the selected value on the charts viewed by the same sheet. This approach makes the search and query process very easy for the user who even does not have knowledge about the databases. The search process will be accomplished without writing anything. Except when the user does not want to use a simple search, he can use the Set Analysis and expressions. [Qlik, 2014c] [Qlik, 2014b]

2.2.5 QlikView Scripting and set Analysis

QlikView uses SQL-like scripting environment where the user can write scripts to connect to the data or transform the data like renaming a field. Using scripts the user can restrict the data or showing the desired data. Set Analysis is a QlikView function used to control the data viewed by a chart and has its own syntax. Example for Set Analysis is the task made using the data set to count the number of records with "Ist EUR" >= 1000000. In this Example the following two Expressions is used for the two bars.

- Count(<[Ist EUR] = ">=1000000"> [Ist EUR])

- Count(<[Ist EUR] = "<1000000"> [Ist EUR])

[James, 2014]

2.2.6   QlikView as in-memory Technology

QlikView is considered to be an in-memory Business Intelligence which makes its performance higher than the other tools. In in-memory technology all data will be loaded to the memory and processed from there. QlikView with this technology has an inference engine to maintain the association between the data. Such privilege does not exist in the normal Business Intelligence platforms which requires hand coding and rebuilds. The in-memory Technology gives QlikView other privileges like calculating aggregations on the fly and compressing the data. [Qlik, 2014a] [Qlik, 2014c]

Product	Pricing information
Spotfire Desktop	$1999 perpetual
Spotfire Cloud	$30 per month or $300 per year
Spotfire Cloud Work Group	$200 per month
Spotfire Cloud Enterprise	depending on company
Spotfire Mobile Metrics	Spotfire Cloud licensing
Spotfire GeoAnalytics	Spotfire Cloud licensing

Table 2.2: Different Spotfire versions [TIBCO, 2014a]

2.3 Spotfire

2.3.1 Introduction

Spotfire is a data analysis and visualization tool offered by the company TIBCO. TIBCO Software Inc. is an US American company located in Palo Alto, California [Wikipedia, 2014a]. The company was founded 1997.

The base of the software was engineered in the year 1990. so even before TIBCO Software was founded. This was done at the University of Maryland, College Park, by the Swedish guest student Christopher Ahlberg and his colleague Ben Shneiderman. 1996 a company named "Spotfire" was founded by Ahlberg which was bought by TIBCO in the year 2007. The current version of the software is 6.5 and was released in May 2014. [Wikipedia, 2014b]

TIBCO focuses on offering a flexible and easy to use business intelligence solution. Spotfire features a lot of different visualizations and a modern interface to the user. The software gained some fame when it was used by the International Federation of Association Football (FIFA) during the World Cup 2010 to publish visualizations about past teams' performances [Wikipedia, 2014b].

2.3.2 Product Overview

TIBCO offers Spotfire in different variations and different pricing strategies. Spotfire Desktop is the only program that is not included in the Spotfire Cloud licensing scheme. The price for Spotfire Desktop is payed once and the license is not expiring [TIBCO, 2014b]. Spotfire Cloud is an alternative data analytics software that is offered in different versions and paid on a monthly or yearly (customer's choice) base [TIBCO, 2014c]. This licensing scheme is also including the products "Mobile Metrics" and "GeoAnalytics". See table 2.2 for more details. Furthermore TIBCO offers 30-Day trials which are not limited in terms of features for their products. The following sections will focus on Spotfire Desktop.

2.3.3 GUI Explanation

Spotfire Desktop starts a welcome screen that is comparable to modern office suite's opening screens. It offers a search bar that is linked to different libraries and recent local files and the most important features to start a new project or continue an existing one. Spotfire keeps track of recent analysis (dpx files) and recent data sources (different formats) separately. Worth mentioning is the importing dialog (see Figure 2.7): CSV is a common exchange format but it lacks unity. Spotfire allows the user to select all format details, so preprocessing the file should never be necessary. This is very important as preprocessing is always a possible error source and should be avoided if possible.

Data Conjunction

Very important for a modern data analysis software is the possibility to fetch information from different sources and bind these tables together. In case of additional rows the columns of both tables have to be assignable somehow to get one combined table. The more interesting case is when columns are about to be added: In this
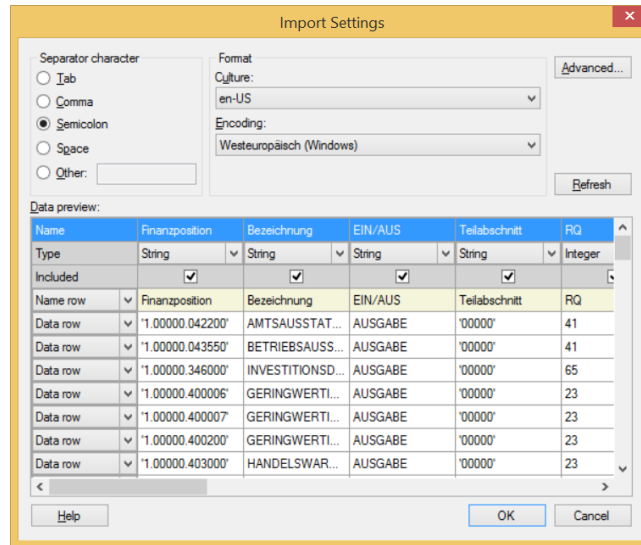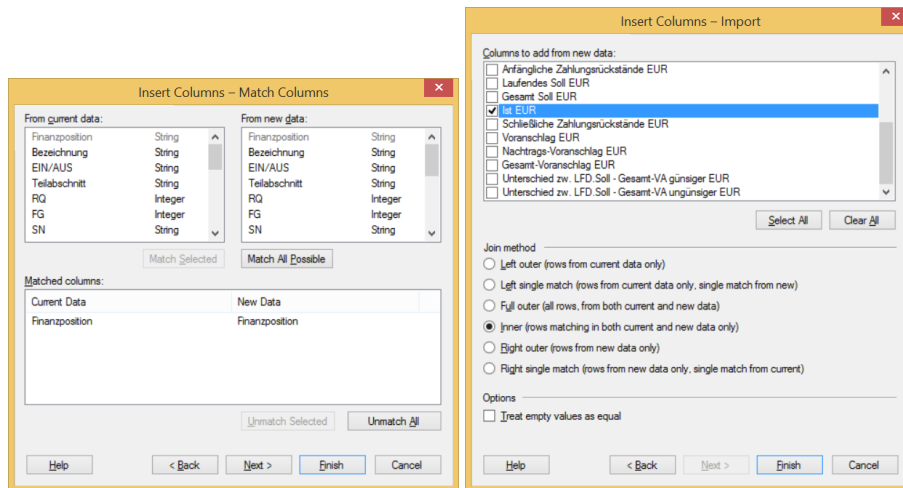
Figure 2.7: Spotfire import dialog

Figure 2.8: Data conjunction of additional columns is done by selecting a source, shared keys and a joining strategy

scenario shared keys are used to identify matching data rows of both tables. Fortunately Spotfire offers all tools to solve both tasks. Figure 2.8 shows how a column is imported to an existing data set: First of all one or more shared keys have to be selected. Like in relational databases different variants of joins (inner, left, right, full outer) can be used. This is necessary to handle situations when both tables are not complete, so they are not sharing the exact set of primary keys using for the conjunction. A left join for example uses only the data rows from the already imported data set and joins the columns from the new data set if there is a matching key in the new data set. Otherwise these columns will be empty. Right join is the exact opposite. For an inner join both tables must provide a key, so the resulting rows will not contain empty cells. Full outer join will result in a table having all rows from both source tables and these combined into one row that had a matching key.

Overview

After importing the data Spotfire will present a basic bar chart (see Figure 2.9). Therefore it will select some column (it is always the same for the same data set) for the horizontal axis and it will put the function "Row Count" on the vertical axis. The toolbar offers the most important features like data tables and different types of visualizations. Furthermore it will show all filters on the right area of the screen. Below the filters box it
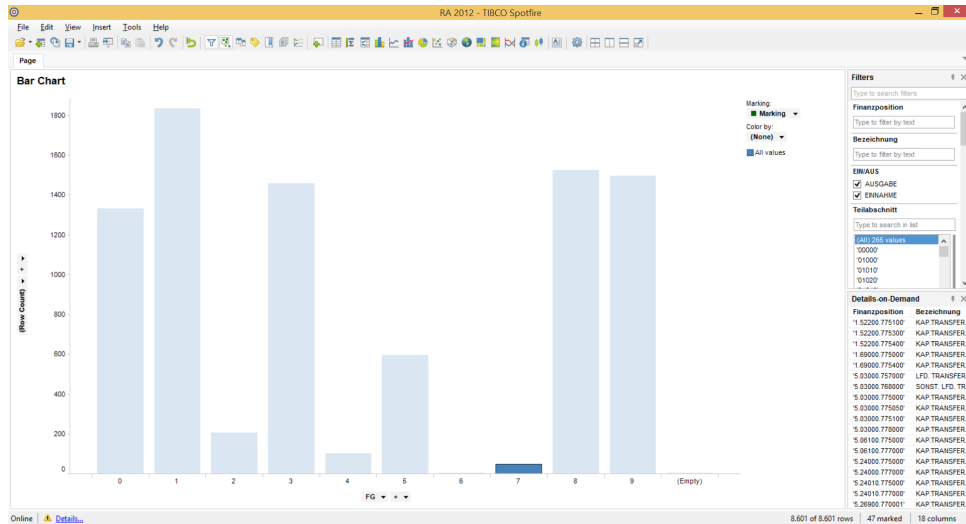
Figure 2.9: Spotfire GUI after importing data and selecting one category (7)

Filter type	Data types applicable to	Description
Range Filter	Any	Lower and upper bound selection
Item Filter	Any	Exact selection of a specific item
Checkbox Filter	Categories	Exact multi selection while having overview
Radio Button Filter	Categories	Exact selection while having overview
Text Filter	Any	Text search
List Box Filter	Any	Multi selection list box combined with filtering text box while having overview

Table 2.3: Available filter types in Spotfire and the different data types they are applicable to

shows a feature named "Details-on-Demand". "Details-on-Demand" is simply a data table that shows all data rows that are currently selected. In this table the columns can be re-arranged, single columns can be removed for convenience.

Filtering

Spotfire offers one filter for every column in the data set. It supports six different filter types that are freely switchable as long as the data type of the column is supported by the filter (see Table 2.3 for more details) and preset depending on the data type of the values of the column. If it is enough to get the amount of data sets applying to the conjunction of the filter expressions the status bar will already solve the task: It always prints: "A of B rows | C marked | D columns". While A is the amount of rows applying to the filter rules, B is the amount of rows on total. C returns the number of marked rows and D is the number of columns in the data set. While updating the filters the resulting output is applied to all visualizations automatically. There is no further assignment necessary. This is very helpful because the the results of filtering can be watched while for example moving a range slider around and having a bar chart on the screen. Filters can be reset by using the context menu in the filtering area. For convenience all filters can be reset at once. Figure 2.10 shows all filter types next to each other.

Visualizations

Spotfire offers a lot of different visualizations and further optimization. They all have in common that the data are arranged on axis and additional fields like coloring information. The big advantage of this architecture is that the visualizations can be switched around without having to reassign the data to the specific axis and
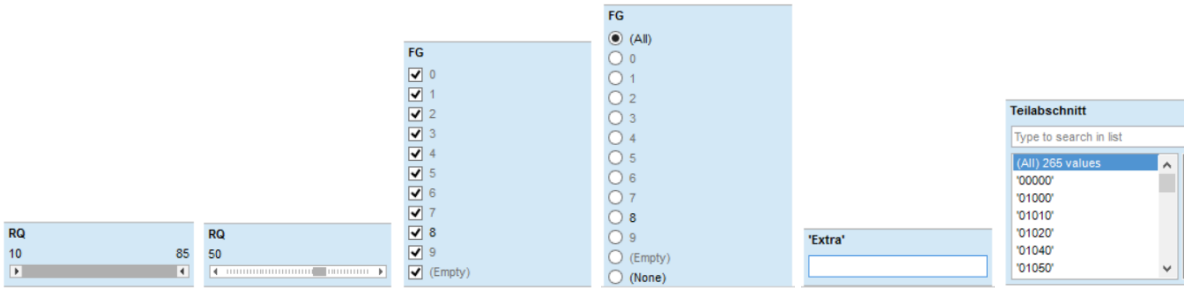
Figure 2.10: Spotfire offers six different filter types (from left to right): Range Filter, Item Filter, Checkbox Filter, Radio Button Filter, Text Filter, Listbox Filter
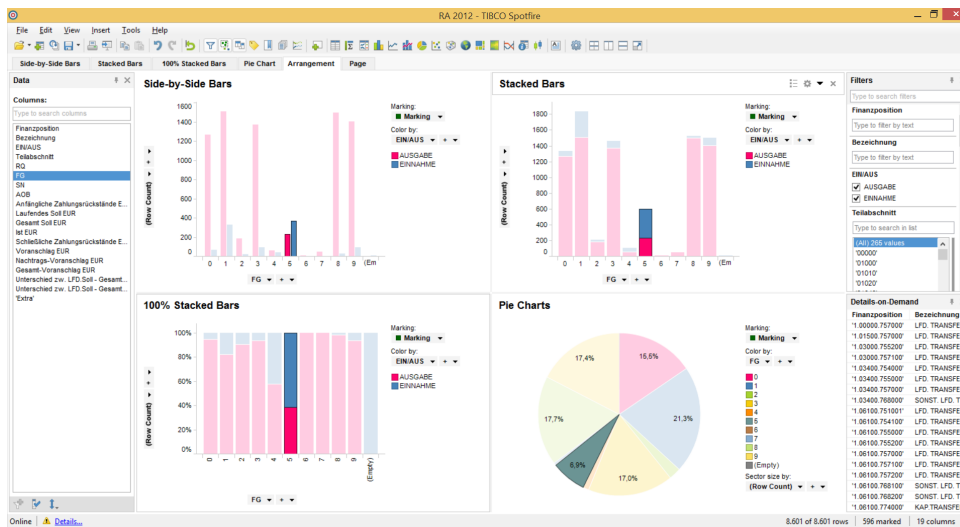

Figure 2.11: Four different visualizations in Spotfire: Side-by-Side Bars (upper left), Stacked Bars (upper right), 100% Stacked Bars (bottom left) and a Pie Chart (bottom right)

fields. Spotfire uses different layers to manage visualizations. On top it uses pages. Pages are arranged like tabs in browsers, they are always next to each other and on the top of the screen. After starting Spotfire one page is created automatically. Empty pages can be created, existing pages can be duplicated or deleted. Within pages visualizations can be placed. If only one visualization is present on a page it will take all the space by default. However, visualizations are much more flexible, they can be arranged next to each other (horizontally or vertically), within a grid, or in a way that the active visualization is maximized while all others are minimized. Like pages visualizations can be named, duplicated, created empty and removed. They can also be moved around between pages. Figure 2.11 shows four example visualizations arranged in a grid. It is worth mentioning that only one category was selected in one visualization, the others where selected automatically by Spotfire. Arranging visualizations in a clever way can show interesting portions of data very easily this way. This can be understood as a very simple but still use able application of data interaction.

Calculated Columns

Another essential feature is the possibility to create new columns by combining existing ones and using specific functions without touching the original data source. To define these columns Spotfire uses a expression language. Fortunately for the user the syntax can be built by using a dialog (see Figure 2.12) that features all necessary fields by simply using drag and drop. The functions available are categorized, a few of them are: Date and Time, Logical, Math, Statistical and Text. This feature gets especially powerful when additional data sources are added to an existing project like described in Section 2.3.3.

Figure 2.13 shows a data set with one column of four different data sets of other years imported in. These
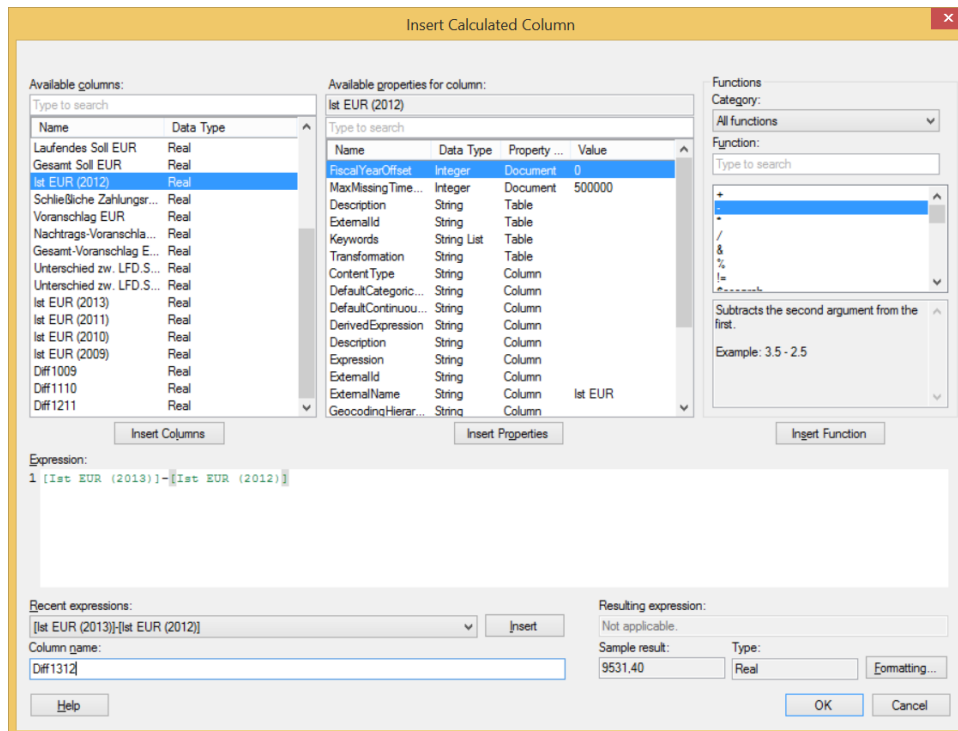
Figure 2.12: Spotfire grants the user the possibility to create calculated columns using existing columns and specific functions.

columns are called "IST Eur (year)". Four calculated columns were inserted and they were always building the difference of one year to year before. The new columns were saved using the naming schema "Diffabxy". Afterwards a new page was created and a Cross Table was used to show the summed differences of the calculated columns over the years for the specific categories of the column "FG". The same data was visualized using a Bar Chart on the right side. This simple example shows a very powerful tool for data analysis especially if more complex functions and more columns are used.

2.3.4 Summary

Spotfire offers a very user friendly and very well thought out data analysis and visualization solution. All parts of the program seem to be consistent, no feature seem to be lost anyhow. Although the user interface is easy to use this would not be because of a lack of complexity, the opposite is more true: TIBCO accomplished to placed the features where they belong to, the complexity rises step by step and the user does not get overwhelmed by a lot of features. A good example is the calculated columns dialog mentioned in Section 2.3.3: Most of the features are use able without ever using this dialog, charts can be rendered, tables can be created. When having to use combined columns the dialog still offers an easy enough solution: The operands and the operator are assignable via drag and drop, still this a very flexible system, it is an expression language, complex combinations of columns could be created using this dialog. A lot of different functions are categorized, the user is not limited to basic mathematical functions. Spotfire also offers a lot of visualizations and a lot of customizations.

A typical work flow in Spotfire is to create a new page, create a visualization, whether it is a table or a chart, this is not so important in most cases at creation time and then clone duplicate things: Visualizations and pages. On the one hand this is a pretty nice thing as the data assignment to axis is kept and nice pages can be arranged very easily, the only downside is the lack of the possibility to change the visualization type with one click. This has to be done by using the context menu and can be a little time-consuming when a visualization has to duplicated a couple of times and switched to different types of plots.

Nevertheless Spotfire is best described by using these two words: Grown up. It is a reliable solution that offers a presentable portfolio of features in an innovative and intuitive way. From a user experience perspective
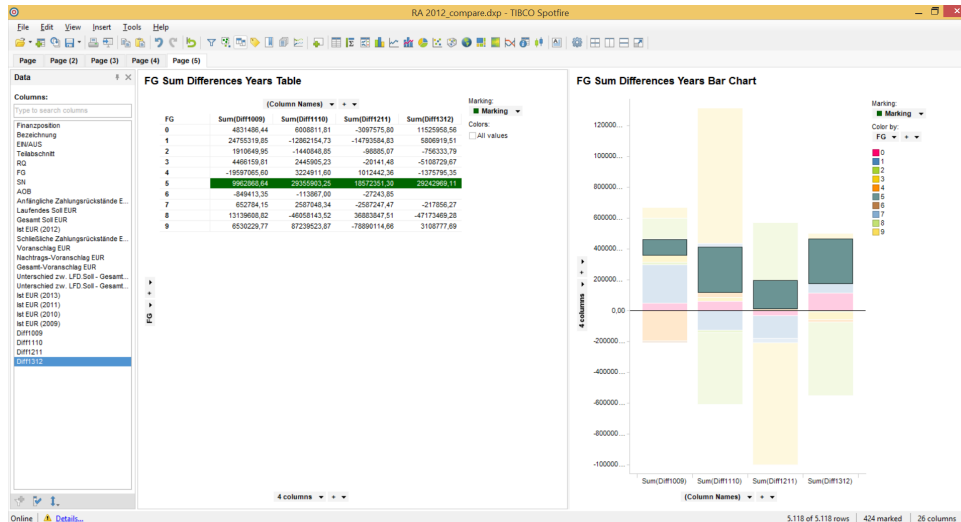
Figure 2.13: Four different calculated columns summed up in a Cross Table (left) and in a Stacked Bar
Chart (right). Both were summed up for the same categorical column "FG"

slight optimizations could maybe help the user to create charts in a more convenient way and even faster. One
feature definitely missing is a recommendation of which visualization is suitable for the information depending
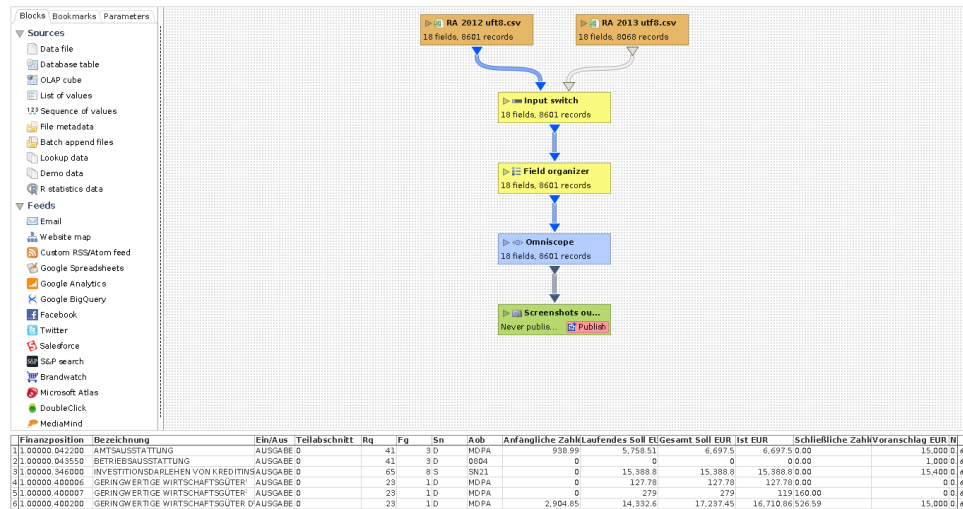on how the axis are assigned on the current screen.

Figure 2.14: Omniscope let the user handle the source and output data in a simple way.

2.4 Omniscope

2.4.1 Introduction

Visokio is a small British company. It is founded 2002 and its headquarter is in London [Visokio, 2014a]. But there is also an office in St. Helier Jersey. It took them three years to release Omniscope 1.0.

Omniscope is written in Java. Therefore it supports the Operating Systems Windows, Mac OS X and Linux. There are four different editions [Visokio, 2014b]:

- **Viewer**: The Viewer supports navigating and exploring existing projects. The files have to be created by other versions. But with the viewer it is possible to browse through the data, filter it and create screen shots. It is free.

- **Desktop**: The Desktop Version let user create and edit data. There is a 30 days trial for this version.

- **Server**: With the Server it is possible to run the source files as data-server. The data refreshes automatically when the source is changed.

- **ServerPlus**: The ServerPlus extends the server with the possibility to let Viewer access the data

For this work the Desktop Version is used.

2.4.2 Graphical User Interface

Omniscope uses tabs to subdivide different functionality. There are at least two tabs in use, the *Data Manager* and the *DataExplorer*. The *Data Manager* (see figure 2.14)contains information about the data. Here it is possible to set the source files. More than one can be added but only one data-pool is used. It is possible to combine them or simply switch them

The *Data Explorer* is the working space. All charts and data-information are handled here. This is the place which is used to work with the data. If needed other *DataExplorer*-Tabs can be created.

By default the *DataExplorer* is split into two panels for exploring the data and creating charts. A third panel which contains the filter-methods is placed on the right side. Omniscope is loose with its panels, only the filter-panel is fix. It is possible to add more panels, move them around and close them. To facilitate the work on a chart a panel can get maximized to the window-size. The information about the view do not get lost by this.

The panels are built up the same way. On the head there are the operations:

- selector to change the chart,

Figure 2.15: Three different filters for different data-types and searching.

- data to show (filter-in, filter-out, selected),

- functionality for the current chart,

- selector for the columns to show,

- highlight the data (coloring, sizing).

The body contains the chart itself. Here it is possible to explore the data.

2.4.3   Working with Omniscope

Handle Input and Output

Omniscope can handle different kind of files as source: plain text, Excel, XML, CSV and several more. As location beside the hard-disc it supports FTP, Dropbox and Google Drive. It can understand a different location format, for example ',' as decimal-separator.

Like mentioned in chapter 2.4.2 handling the source data is done in the *DataManager*. This is a very powerful tool. Here the user can manipulate the data before the work starts. Two or more sources can be combined. The data can get prefiltered. If Omniscope is not able to identify the data type correctly (for example numbers should be handled like strings) the user can change the it. Not only the data input is handled here. The User can define some output rules. After the work is finished the user can create screen shots, upload the the data file or send at to other users just with one click.

Filtering

Depending on the data type Omniscope offers three different types of filtering (see Figure 2.15):

- range slider for numbers,

- check boxes for categories,

- input fields for strings.

With the range sliders the user can define the upper and the lower limit just by dragging it. For huge values it is very complicated to find the wished value. Therefore Omniscope let the user type the wished limit directly.

Check boxes are checked by default. So the user has to unselect the categories which should not be shown. If there is just one category which should be shown the user can simple click on label. This unselected all other categories automatically.

The input field do not need the full string. Just entering a sub string is enough. Omniscope gives some suggestion about what the user can think of.

It is possible to filter by search. The user enters a search string and Omniscope searches all columns for it. There is an auto-complete too like in the input field. For better results the user can start a custom search only in a few columns.
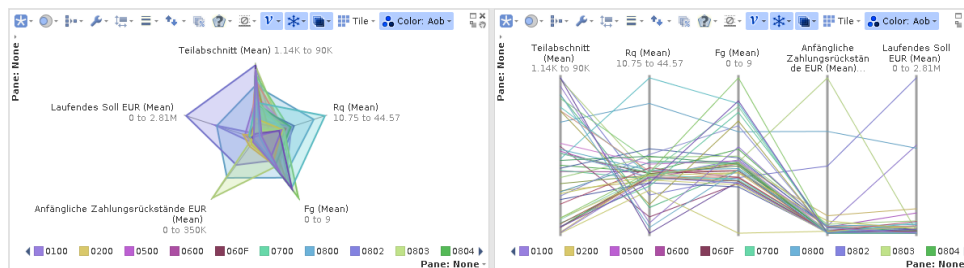
Figure 2.16: Stacked bar chart maximized to 100%.

Figure 2.17: In Omniscope parallel charts are part of the radar chart.

Create Charts

Omniscope provides many different charts. Bar chart, pie chart, tile chart, network chart and radar chart are just some of them. Many charts are combined into one main chart selector. The bar/line selector for example contains bar chart, line chart and markers. These are charts which actually shows their information in a similar way. But there can be found more. The stacked bar chart (Figure 2.16) is also a part of the bar/line selector.

Another example for such a combination can be found in the Radar selector. Omniscope provides different method to present the data as a radar chart. It can be shaped, lined or combined both. But there is also the possibility to create a parallel coordinate chart (Figure 2.17). There is a similarity between radar chart and parallel coordinate. But to find it there is unexpected.

Some charts are a bit trickier to create. A treemap is a sub-chart of the tile chart. As standard a tilemap represents all data as tile with the same size. No one can read any information from this. There are actually two ways for creating a treemap depending on what the user want to show. If just information of a single column should be represented it is enough to set this column as size source. In Figure 2.18 a treemap is created which should show the relations between two columns and give an overview about the sum of values which correspond the respective conditions. At first the values are grouped based on specific categories. The user now has the information about which category has assigned the most values. Now it is needed to set the tile base so that Omniscope knows which column it is actually about to work with. On the same time it is useful to set the coloring to the same value so that the user gets an optical representation. The Treemap is finished by giving a sort so within the group tiles the values are grouped by the second column. Omniscope does not require this specific order. So the user can create it right the other way too.
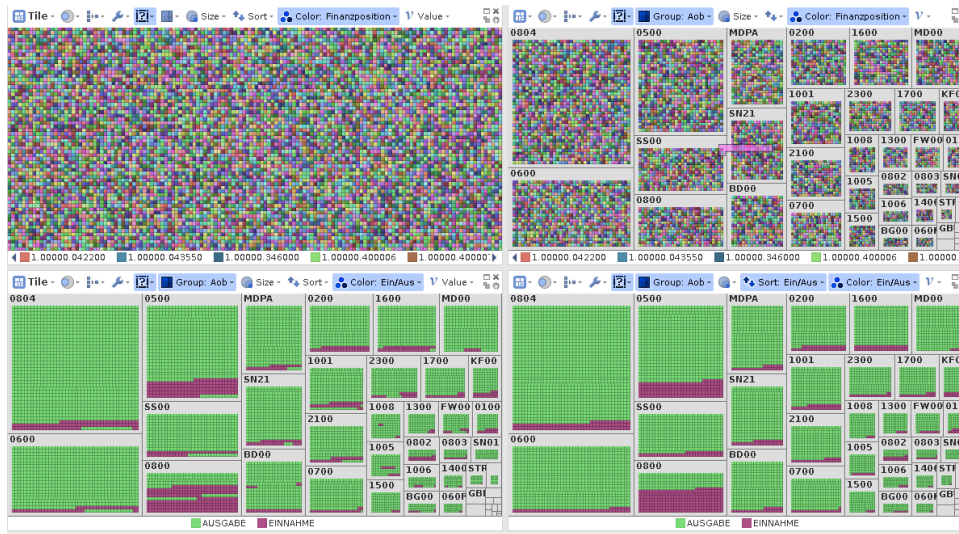
Figure 2.18: Create a treemap. Step one select tile chart. Step two select group. Step three select tile content and coloring. Step four sort it.

Handle Charts

In Omniscope the *DataExplorer* is a tab which can shows charts and data. A DataExplorer can be split into numerous panels (see Figure 2.19). All of them are showing charts with different settings. It is not possible to copy the settings of an chart into a new other one. But Omniscope keep in memory the old settings. So if a chart was created and set up before the new chart would be set up with the old settings.

All these panels share the filter and data selection. If the user selects data in one chart it will be shown on other charts. The selected data is shown like normally but all the other data is grayed out. Charts can use selected data as filter. This is only useful if there are more than one chart to work with. So selecting data on another chart let this chart only show the data. For example one chart shows an overview and the other one shows the details.

Based on the selection and filtering the user can create so called *Queries*. Queries save the filtering and selecting data and can be used as subset of data for a single chart. With this method charts with different subsets can be used in one tab.

If the user wants to create views with a different filter a new *DataExplorer*-tab can be created. The old tab can get duplicated so the user can start the new work based on the old one. But a completely fresh tab can be created as well.

2.4.4 Summary

Omniscope is easy to get start with. It is very simple to use. Just some clicks are needed to create some wonderful charts. There is no need to know the program to do the first steps. The base handling is clearly. It is no problem to try it out by trial and error.

But looking at into detail the handling gets a bit awkward. There is no suggestion for the charts. They have to be set up manually. The control-buttons for setting up are confusing. Depending on the chart there are many buttons in the control-panel. Even if the user knows exactly what to do it is still search work to find the right option-button.

The *DataManager* seems to be very powerful. At first the user would ignore it because the benefit is not that manifestly. But as soon as Omniscope has problems with the data to recognize it correctly or the user has to handle different data sets the strengths get clearly visible. The publish functionality is very impressive. After work is finished one click is enough to upload the results, create screen shots and PDFs and send an email.
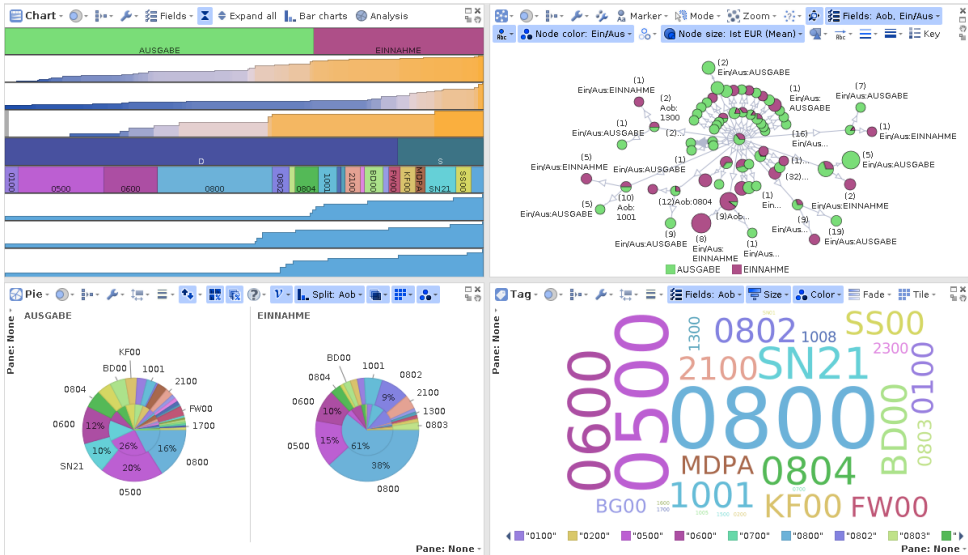
Figure 2.19: DataExplorer can be split into different panels. Each of them shows a chart.

Chapter 3

Comparison

3.1 Similarities

The purpose of all of these programs is to give the user the possibility to visualize and analyze data. Still the approaches achieving this are different and this results in completely different products that offer different solutions and philosophies. Nevertheless some aspects are still comparable. Tableau and Spotfire are the two products that are most visualization oriented and most similar products. Omniscope is a little bit in between QlikView and the rest, this software is more data interaction oriented and offer some features that other products are missing like the network browsing. QlikView on the other hand is much more expression based than the competitors and tries to offer dialogs to be as user friendly as possible, but the philosophy is a completely different. The similarities are displayed in Figure 3.1.

As Tableau and Spotfire are most similar to each other it is necessary to compare these two products in more detail: The possibility to switch between visualizations by clicking once and suggesting suitable visualizations makes Tableau the even more user friendly software than Spotfire. Still also Spotfire fulfills the requirements of a successful implemented user interface very well.

3.2 Data Import

CSV is a very common data exchange format. As CSV lacks unity data analysis software have to grant provide some flexibility when importing CSV files. Tableau for example offers a drop down menu for the comma separation character and another selection for the encoding. Unfortunately the encoding drop down list does only contain ANSI, OEM, UTF-8, UTF-16 and an entry called "Others". When this entry is selected a text box appears where a number can be entered representing the character encoding but there is no further explanation what this number means. Importing data sets which are encoded in some ISO format can lead into problems in Tableau. Another downside of Tableau is that it is not showing the table as it would be imported if the options were kept until confirming. There is no selection of the regional setting in Tableau (this would be used for the decimal separator) but Tableau seem to detect numbers correctly, even if the regional settings in the operating system and the decimal separator in the CSV file do not match.

The import dialog of QlikView is a little bit more advanced compared to Tableau. It shows a preview table, the character encoding drop down selection offers a lot more options than Tableau. There is also no selection
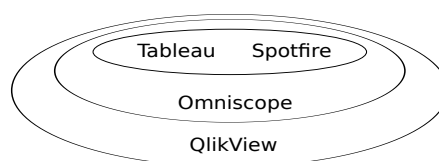

Figure 3.1: A simple clustering visualization trying to demonstrate similarities between the programs.

for the region as well, QlikView detects the numbers as Tableau does.

Spotfire and Omniscope both offer a locale selection that influence the decimals of numbers. Both programs offer a large list of encodings. Omniscope does not show a preview while Spotfire does. Another slight disadvantage in Omniscope is the fact that the option to get this import dialog is very inconspicuous and easy to overlook. If the corresponding option is not selected the data set is imported with default options which can for example lead into wrong locale options. On the other hand with the DataManager Omniscope offers a tool which allows to change the option and reload the data. Here the user can check the correctness of the import with the help of a preview.

3.3   Searching and Making Data Filters

QlikView uses the associative approach to allow the user to search for some value and its associated values. In QlikView the user can view all the values of each field in a list box. For example a database which contains data about sales, customers and countries could be used to be queried for all customers in Austria. Therefore in QlikView it is enough to click on Austria in the country list box and this value Austria will be green. The associated values in the customer list box i.e. the customers in Austria will be selected automatically and become white. The excluded values will be gray. In QlikView users can use also expressions to search. The expression can be written at the top of each list box. For example, the following expression can be written at the top of the country list box =sum(Sales)>100000 to search for the countries with sum of sales greater than $100000. The expressions can be more complicated than this one. The clear button is used to clear the selection to start another search process.

Tableau, Spotfire and Omniscope act completely different than QlikView in terms of filtering. If a data element is clicked in some visualization in one these programs no filtering is done. Instead of filtering the clicked element and all corresponding data elements in other visualizations are highlighted. Filtering is done in specific areas.

3.4   Creating and Managing Visualizations

Tableau and Spotfire have in common that dragging and dropping columns on axis is the default way of preparing visualizations. Duplicating them and switching the type for arrangements is also a typical way of working in both programs. The presentation mode in Tableau surpasses the arrangement features of Spotfire because it grants even more space by removing everything else from the screen than the visualizations. Selecting data columns for a visualization in Omniscope however is not more complicated than in Tableau or Spotfire although it is not solved by drag and drop: The available columns are offered in the toolbar and the user has to simply choose by clicking them. The problem here is that this is represented by one of many buttons that is expanding like almost all the others next to it do. This can be very confusing as the icons on these buttons are not really describing them well. The user can get overwhelmed by this amount of features very easily. Furthermore it is not possible to switch the visualization type in Omniscope without having to reassign the data columns.

With QlikView the user can create a wide variety of visualizations. QlikView saves its document in a .qvw file and every document can have more than one sheet. Sheets behaves like the tabs. The easiest visualization offered by QlikView is to load the columns values in a list box where the user can see all the values of some field and can see also their frequencies. To create a new chart the user can use New Sheet Object from the float menu or create chart from the design toolbar or uses the Quick Chart wizard from the standard toolbar. The most important two things to create chart are the dimensions and Expressions. Dimension is the field which has string values like countries but expression is used to state the size of the bars representing some dimension values. In QlikView a conversion from a type of chart to another can be done easily but in some cases the process needs a little bit more modifications. The properties of the chart can be used to make it more readable for the user. In QlikView the charts can be minimized to make more space for the other charts and objects on the work place. They can also be enlarged or reduced by dragging.

	Tableau	Spotfire	Omniscope	Qlikview
Heat Map	🟩	🟦	🟧	
Highlite Table	🟩	🟦	🟧	🟥
Maps	🟩	🟦	🟧	
Pie charts	🟩	🟦	🟧	🟥
Bar charts	🟩	🟦	🟧	🟥
TreeMap	🟩	🟦	🟧	🟥
Circle charts	🟩		🟧	🟥
Line charts	🟩	🟦	🟧	🟥
Area charts	🟩			🟥
Scatter Plot	🟩	🟦	🟧	🟥

	Tableau	Spotfire	Omniscope	Qlikview
Box Plot	🟩	🟦		
Gantt charts	🟩		🟧	
Bullet Graph	🟩	🟦		
Bubble Graph	🟩			🟥
Parallel Coord. P.		🟦	🟧	🟥
Network			🟧	
Cross Table		🟦		🟥
Tag	🟩		🟧	
Radar Chart			🟧	🟥

Figure 3.2: A variety of plot types and their supports in the different data analysis tools.

3.5 Visualization Types

While the basic chart types like bar charts and pie charts are supported by all products the more complex visualization types like radar charts or network graphs are not available in every software. There are also differences between visualization details. Stacking bars for example is a feature that all products support. Showing line charts as an area is not possible in every solution. Figure 3.2 shows an overview of available plot types in the different software solutions. Some visualization types are of course implemented partially divergent from others. Some trens are also worth a glance: TreeMaps for example became a common way to compare values by showing them as areas next to each other. These TreeMaps are supported by all products investigated by default.

Using Maps on the other hand is not a default option in QlikView and an extension must be used. Google Maps can be used with QlikView with some instructions. Tableau, Spotfire and Omniscope do have Map implementations by default.

Chapter 4

Dataset

4.1 Data Description

In this paper the data set *Rechnungsabschluss der Stadt Graz 2012* is used. It describes the financial data of the city of Graz. There are 8.601 financial items in it.

The data set contains 18 columns:

- **Finanzposition:** It is kind of an unique id containing the section and a number.

- **Bezeichnung:** The name of the data.

- **EIN/AUS:** There are two different states telling if it is income or output.

- **Teilbereich:** The section number.

- **Voranschlagsquerschnitt (RQ):** It is an classification number for the type of the income or outgoing.

- **Finanzwirtschaftliche Gebarung (FG):** An identification number for the financial management.

- **Sammelnachweis (SB):** A classification which should combine similar income and outgoing.

- **Anordnungsbefugnis (AOB):** Another kind of classification. It is an identification string informs about authority.

- **Anfängliche Zahlungsrückstände EUR:** Opening arrears

- **Laufendes Soll EUR:** The running target

- **Gesamt Soll EUR:** The final target. In this target it is the sum of the opening arrears and the running target.

- **Ist EUR:** The actual amount of income and outgoing

- **Schließliche Zahlungsrückstände EUR:** ending arrears

- **Voranschlag EUR:** Estimate

- **Nachtrags-Voranschlag EUR:** Supplementary estimates

- **Gesamt-Voranschlag EUR:** Total estimate

- **Unterschied zw. LFD.Soll - Gesamt-VA günsiger EUR:** The positive difference between the running target and the total estimate. On income the running target should be higher than the total estimate. On outgoing the total estimate should be higher than the running target.

- **Unterschied zw. LFD.Soll - Gesamt-VA ungünsiger EUR:** The negative difference between the running target and the total estimate.
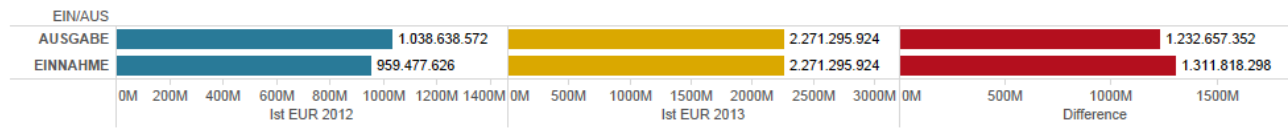
Figure 4.1: The income and outgoing of 2012 and 2013 and its difference.

Figure 4.2: The final target and the amount of income and outgoing from 2012.

4.2 Compare Data of 2012 and 2013

Figure 4.1 presents the difference between the income and the outgoing of 2012 and 2013. Both the income and the outgoing in 2013 are higher than 2012. In 2013 there was an outgoing of 1.169.040.553€ which is 130.401.982€ more than 2012. The income in 2013 was 1.102.255.371€ which is 142.777.745€ more. So in 2013 there was an higher growth of income than of outgoing.

In 2012 the city of Graz planned with 1.079.146.506,49€ of income and outgoing (see Figure 4.2). Their outgoing was not that high. They missed it by 40.507.935€. But they also missed their planned income by 119.668.881€.

In 2013 they had a planned with a larger budget (see Figure 4.3). They planned with an amount of 1.241.424.002€ as outgoing but with an amount 1.241.447.420€ as income.So they actually planned with greater income than outgoing. It did not work. The outgoing was 72.382.469€ lower than planned one. But the income was even about 139.192.050€ lower than planned.
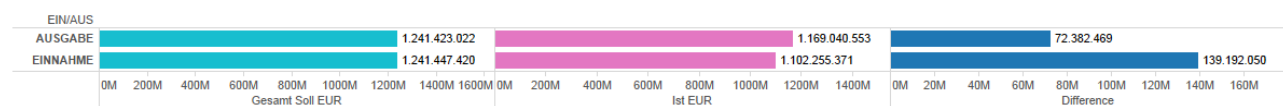

Figure 4.3: The final target and the amount of income and outgoing from 2013.

Chapter 5

Conclusion

Overall Tableau is the easiest to start with. With its handling it really helps users without any experience to work with it. On the other side QlikView is too hard for inexperienced users.

Unfortunately the financial data set of the city of Graz is not good to work with. There is no good explanation for the data schema. The values are not explained either. For example *Sammelnachweis (SB)* has two different values: *D* and *S*. But it is not possible to get out what they mean.

There are just a few meaningful comparison which could be done with this data set. It is possible to create charts that shows income or outgoing according *Anordnungsbefugnis (AOB)* or other categories. But what should be said about its meaning if these categories are just random numbers for people without corresponding meaning.

All in all they are good programs. It is impressive what can be done with them. There are good charts for different assignments. There are also nearly useless charts which are just good looking, like pie chart (which is part of every program), bubble chart or tag chart. It is a pity that with that data set it is hardly possible to show the real advantage of the programs in an ordinary way.

Bibliography

James, Mark [2014]. *Set Analysis A Beginners Guide*. Jan. 14, 2014. `http://community.qlik.com/servlet/JiveServlet/downloadBody/5695-102-1-6926/Set%20Analysis%20Beginners%20Guide%20v2.0.docx` (cited on page 11).

Qlik [2011]. *QLIKVIEW DEVELOPMENT AND DEPLOYMENT ARCHITECTURE A QlikView Technical Brief*. Feb. 1, 2011. `http://qlik.com/gr//~/media/Files/resource-library/global-us/direct/datasheets/DS-Technical-Brief-Dev-and-Deploy-EN.ashx` (cited on page 9).

Qlik [2014a]. *QLIKVIEW ARCHITECTURE AND SYSTEM RESOURCE USAGE QlikView Technical Brief*. Apr. 1, 2014. `http://qlikview.com/jp/~/media/Files/resource-library/global-us/direct/datasheets/DS-Technical-Brief-QlikView-Architecture-and-System-Resource-Usage-EN.ashx` (cited on page 12).

Qlik [2014b]. *QlikView Associative Search*. May 19, 2014. `http://qlik.com/us/videos/how-to/qlikview-associative-search` (cited on page 11).

Qlik [2014c]. *What makes QlikView unique*. Jan. 1, 2014. `http://qlik.com/us/explore/resources/whitepapers/what-makes-qlikview-unique` (cited on pages 11, 12).

Qlik [2014d]. *Who is Qlik?* May 19, 2014. `http://qlik.com/us/company` (cited on page 9).

Tableau [2013]. *Tableau Server v8 1 SAML Overview*. Nov. 12, 2013. `https://youtube.com/watch?v=_UFkMGDwcbc` (cited on page 5).

Tableau [2014]. *Tableau Software*. May 13, 2014. `http://tableausoftware.com` (cited on page 5).

TIBCO [2014a]. *TIBCO Spotfire*. May 15, 2014. `http://spotfire.tibco.com/` (cited on page 13).

TIBCO [2014b]. *TIBCO Spotfire® Desktop Beautiful Visualizations*. May 19, 2014. `http://spotfire.tibco.com/discover-spotfire/spotfire-overview/spotfire-desktop` (cited on page 13).

TIBCO [2014c]. *Which Spotfire is the right fit?* May 19, 2014. `https://spotfire.cloud.tibco.com/tsc/#!/buyspotfire` (cited on page 13).

Visokio [2014a]. *Company*. May 13, 2014. `http://visokio.com/company` (cited on page 19).

Visokio [2014b]. *Omniscope Editions*. May 13, 2014. `http://visokio.com/omniscope-editions` (cited on page 19).

Wikipedia [2014a]. *TIBCO Software*. May 15, 2014. `http://en.wikipedia.org/wiki/TIBCO_Software` (cited on page 13).

Wikipedia [2014b]. *TIBCO Software Analytics*. May 15, 2014. `http://de.wikipedia.org/wiki/TIBCO_Spotfire_Analytics` (cited on page 13).

Yi, Ji Soo et al. [2007]. *Towards a Deeper Understanding of the Role of Interaction in Information Visualization*. Oct. 27, 2007. `http://cc.gatech.edu/~stasko/papers/infovis07-interaction.pdf` (cited on page 3).