TIBCO Spotfire for Developers

Create, manipulate, and deploy rich analytical visualizations using the TIBCO Spotfire platform

Manuel Xavier

[PACKT] enterprise 88
PUBLISHING   professional expertise distilled

TIBCO Spotfire for Developers

Create, manipulate, and deploy rich analytical
visualizations using the TIBCO Spotfire platform

Manuel Xavier

[PACKT] enterprise
PUBLISHING
professional expertise distilled

BIRMINGHAM - MUMBAI

TIBCO Spotfire for Developers

First published: October 2013

Credits

Author
Manuel Xavier

Reviewer
Sri Prakash

Acquisition Editor
Usha Iyer

James Jones

Commissioning Editor
Govindan K.

Technical Editors
Iram Malik

Shali Sasidharan

Project Coordinator
Romal Karani

Proofreader
Chris Smith

Indexer
Tejal Soni

Graphics
Abhinash Sahu

Production Coordinator
Melwyn D'sa

Cover Work
Melwyn D'sa

About the Author

Manuel Xavier is an experienced IT Architect (certified SOA Architect and TOGAF 9), specialized in the creation of business-driven and vendor-neutral Service Oriented Architectures (SOA) and Business Process Management (BPM) solutions.

He has been in IT for over 10 years, having participated in many high-profile projects across several countries, namely Portugal, Egypt, UAE, and the Netherlands. His experience was accumulated in distinct vertical markets, from Financial Services and Insurance to Telecommunications, Energy, and Utilities.

Besides SOA and BPM, other reference areas for Manuel are Complex Event Processing (CEP), Case Management, Business Activity Monitoring (BAM), and Visual Analytics.

Throughout his career, he has employed Analytics in multiple scenarios, the most relevant being: measurement of process performance, monitoring of Service Level Agreement expiration, and inspection and evaluation of product sales.

Regarding technologies, he has know-how in several platforms, such as Tibco, Oracle, webMethods, and Progress.

About the Reviewer

Sri Prakash is a successful Management Consultant and Project Manager who heads E-Com Canada Inc., a company dedicated to providing Information Technology consulting services in Canada. He has an MBA from the Schulich School of Business in Toronto and also a Master's Degree in Information Systems.

He brings to the table leadership earned in global giants, such as General Electric, EDS, Rogers-AT&T, and the Provincial Government of Ontario. He has personally led multi-million dollar projects for these organizations and believes in leadership by example.

Under his leadership, E-Com Canada has successfully grown over the years to deliver large and medium projects for reputed companies in and around the Greater Toronto Area. Its clients include the Ontario Real Estate Association (OREA), Stelco (US Steel), Canadian Red Cross, and the Healthcare of Ontario Pension Plan (HOOPP) to name a few.

www.PacktPub.com

Support files, eBooks, discount offers and more

You might want to visit `www.PacktPub.com` for support files and downloads related to your book.

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at `www.PacktPub.com` and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at `service@packtpub.com` for more details.

At `www.PacktPub.com`, you can also read a collection of free technical articles, sign up for a range of free newsletters and receive exclusive discounts and offers on Packt books and eBooks.

`http://PacktLib.PacktPub.com`

Do you need instant solutions to your IT questions? PacktLib is Packt's online digital book library. Here, you can access, read and search across Packt's entire library of books.

Why Subscribe?

- Fully searchable across every book published by Packt
- Copy and paste, print and bookmark content
- On demand and accessible via web browser

Free Access for Packt account holders

If you have an account with Packt at `www.PacktPub.com`, you can use this to access PacktLib today and view nine entirely free books. Simply use your login credentials for immediate access.

Instant Updates on New Packt Books

Get notified! Find out when new books are published by following `@PacktEnterprise` on Twitter, or the *Packt Enterprise* Facebook page.

Table of Contents

Preface

In 2012, 2.5 quintillion bytes of data were produced each day. Much of it has the potential to guide corporations and businesses to improve the offered services and to better focus on customer satisfaction by helping to understand behaviors and patterns, to identify and minimize waste, and to have process visibility. However, the analysis of such large chunks of data, known as **Big Data**, is a huge challenge for the IT world, and only recently have hardware and software started to offer options for such burdensome tasks.

Visual Analytics tools rose as a partial solution for this complex Big Data problem, allowing humans to have a better insight into their owned data. The TIBCO Spotfire platform is one of the pioneers of this world, very well known as a reference in the market. This platform is a leader in both Gartner 2013 Magic Quadrant for Business Intelligence and Analytics Platforms and Forrester WaveTM: Big Data Predictive Analytics Solutions, Q1 2013.

Unfortunately, there are few publications about such an essential tool, whether from TIBCO or others. This book is an opportunity to challenge this fact. I hope it not only fulfills its purpose as a fast-paced introduction to the platform, but also serves as a trigger for the publication of more TIBCO software books.

I sincerely hope the chosen topics and content simplicity make this book mandatory reading for anyone wishing to triumph with such a fabulous tool, and I wish all readers may reach the end of the book with a positive outlook and with enough knowledge to embrace new challenges in the world of TIBCO Spotfire.

What this book covers

Chapter 1, *The Spotfire Architecture Overview* introduces the Spotfire Platform products, defining their independent roles in the platform's architecture. The platform's main features are also described in this chapter.

Chapter 2, *Setting Up a Development Environment* covers the installation pre-requisites, the installation process, and the post-installation tasks for the creation of a local standalone environment, with the capability for creation and deployment of meaningful visualizations.

Chapter 3, *Creating Visualizations* details the processes of loading external data in the Spotfire Platform and of creating visualizations. Several visualization types are presented, such as Bar Chart, Pie Chart, and Scatter Plot.

Chapter 4, *Mastering Visualizations* deepens the knowledge of the previous chapter, by introducing the concepts of filtering, data preparation and transformation, and libraries. This chapter closes with an introduction to the displaying of visualizations through Spotfire Web Player Server.

What you need for this book

The following software are required to build the examples presented in this book:

- Microsoft Windows 2008 64 bit, with Microsoft IIS
- Oracle XE 11g Database
- Oracle Data Provider for .NET (ODP.NET)
- TIBCO Spotfire Server 5.5
- TIBCO Spotfire Professional 5.5
- TIBCO Spotfire Web Player Server 5.5

Who this book is for

This book is intended for everybody who is interested in Visual Analytics and in one of the best Analytics platforms available—TIBCO Spotfire.

Although TIBCO tools have a tendency to be intended for users with a technical background (from Information Technology), the Spotfire Platform is a positive exception, expanding the range of possible clients to other areas and backgrounds, such as Statistics, Science, Banking, and Pharmaceuticals. The installation process and the creation of visualizations (graphs) are very straightforward and should not impose any major difficulties to common users.

Conventions

In this book, you will find a number of styles of text that distinguish between different kinds of information. Here are some examples of these styles, and an explanation of their meaning.

Code words in text are shown as follows: "First, confirm `sqlplus` is in the environment `%PATH%`."

Any command-line input or output is written as follows:

```
sqlplus is not recognized as an internal or external command, operable
program or batch file
```

New terms and **important words** are shown in bold. Words that you see on the screen, in menus or dialog boxes for example, appear in the text like this: "Alternatively, the menu option **File | Add Data Tables** or the **Add Data Tables** button in the toolbar can be used".

> Warnings or important notes appear in a box like this.

> Tips and tricks appear like this.

Reader feedback

Feedback from our readers is always welcome. Let us know what you think about this book—what you liked or may have disliked. Reader feedback is important for us to develop titles that you really get the most out of.

To send us general feedback, simply send an e-mail to `feedback@packtpub.com`, and mention the book title via the subject of your message.

If there is a topic that you have expertise in and you are interested in either writing or contributing to a book, see our author guide on `www.packtpub.com/authors`.

Customer support

Now that you are the proud owner of a Packt book, we have a number of things to help you to get the most from your purchase.

Downloading the color images of this book

We also provide you a PDF file that has color images of the screenshots/diagrams used in this book. The color images will help you better understand the changes in the output. You can download this file from: `http://www.packtpub.com/sites/default/files/downloads/9832EN_ColoredImages.pdf`

Errata

Although we have taken every care to ensure the accuracy of our content, mistakes do happen. If you find a mistake in one of our books—maybe a mistake in the text or the code—we would be grateful if you would report this to us. By doing so, you can save other readers from frustration and help us improve subsequent versions of this book. If you find any errata, please report them by visiting `http://www.packtpub.com/submit-errata`, selecting your book, clicking on the **errata submission form** link, and entering the details of your errata. Once your errata are verified, your submission will be accepted and the errata will be uploaded on our website, or added to any list of existing errata, under the Errata section of that title. Any existing errata can be viewed by selecting your title from `http://www.packtpub.com/support`.

Piracy

Piracy of copyright material on the Internet is an ongoing problem across all media. At Packt, we take the protection of our copyright and licenses very seriously. If you come across any illegal copies of our works, in any form, on the Internet, please provide us with the location address or website name immediately so that we can pursue a remedy.

Please contact us at `copyright@packtpub.com` with a link to the suspected pirated material.

We appreciate your help in protecting our authors, and our ability to bring you valuable content.

Questions

You can contact us at `questions@packtpub.com` if you are having a problem with any aspect of the book, and we will do our best to address it.

1

The Spotfire Architecture Overview

The companies of today face innumerable market challenges due to the ferocious competition of a globalized economy. Hence, providing excellent service and having customer loyalty are the priorities for their survival.

In order to best achieve both goals and have a competitive edge, companies can resort to the information generated by their digitalized systems, their IT. All the recorded events from **Human Resources** (**HR**) to **Customer Relationship Management** (**CRM**), **Billing**, and so on, can be leveraged to better understand the health of a business.

The purpose of this book is to present a tool that behaves as a digital event analysis enabler, the **TIBCO Spotfire platform**. In this first chapter, we will list the main characteristics of this platform, while also presenting its architecture and describing its components.

TIBCO Spotfire

Spotfire is a visual analytics and business intelligence platform from TIBCO software. It is a part of new breed of tools created to bridge the gap between the massive amount of data that the corporations produce today and the business users who need to interpret this data in order to have the best foundation for the decisions they make.

In my opinion, there is no better description of what TIBCO Spotfire delivers than the definition of visual analytics made in the publication named *Illuminating the Path: The Research and Development Agenda for Visual Analytics*.

> *"Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces"*.

> *–Illuminating the Path: The Research and Development Agenda for Visual Analytics, James J. Thomas and Kristin A. Cook, IEEE Computer Society Press*

The TIBCO Spotfire platform offers the possibility of creating very powerful (yet easy to interpret and interact) data visualizations. From real-time **Business Activity Monitoring** (**BAM**) to **Big Data**, data-based decision making becomes easy – the what, the why, and the how becomes evident.

Spotfire definitely allowed TIBCO to establish itself in an area where until recently it had very little experience and no sought-after products.

The main features of this platform are:

- More than 30 different data sources to choose from: Several databases (including Big Data Teradata), web services, files, and legacy applications.
- Big Data analysis: Spotfire delivers the power of MapReduce to regular users.
- Database analysis: Data visualizations can be built on top of databases using information links. There is no need to pull the analyzed data into the platform, as a live link is maintained between Spotfire and the database.
- Visual join: Capability of merging data from several distinct sources into a single visualization.
- Rule-based visualizations: The platform enables the creation and tailoring of rules, and the filtering of data. These features facilitate emphasizing of outliers and foster management by exception. It is also possible to highlight other important features, such as commonalities and anomalies.
- Data drill-down: For data visualizations it is possible to create one (or many) details visualization(s). This drill-down can be performed in multiple steps as drill-down of drill-down of drill-down and so on.
- Real-time integration with other TIBCO tools.

Spotfire platform 5.x

The platform is composed of several intercommunicating components, each one with its own responsibilities (with clear separation of concerns) enabling a clustered deployment. As this is an introductory book, we will not dive deep into all the components, but we will identify the main ones and the underlying architecture.

A depiction the platform's components is shown in the following diagram:
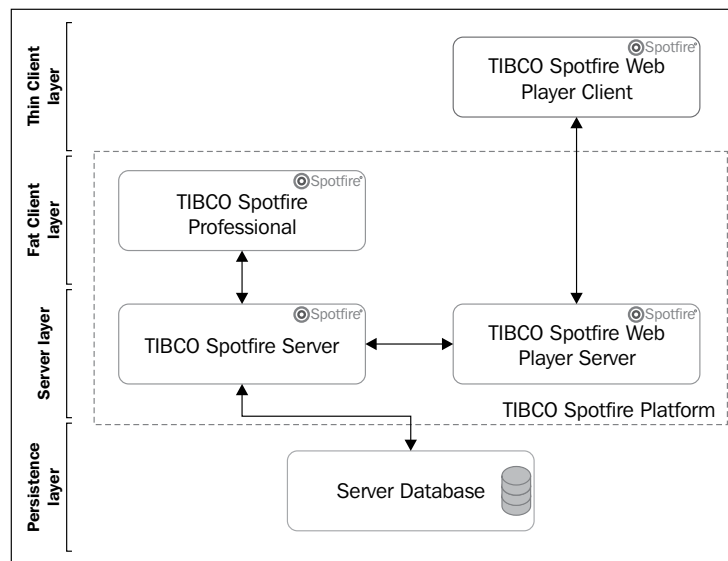
The descriptions of each of the components in the preceding diagram are as follows:

- **TIBCO Spotfire Server**: The Spotfire server makes a set of services available to the analytics clients (**TIBCO Spotfire Professional** and **TIBCO Spotfire Web Player Server**):
    - User services: It is responsible for authentication and authorization
    - Deployment services: It handles the consistent upgrade of Spotfire clients
    - Library services: It manages the repository of analysis files
    - Information services: It persists information links to external data sources

    The Server component is available for several operating systems, such as Linux, Solaris, and Windows.

- **TIBCO Spotfire Professional**: This is a client application (fat client) that focuses on the creation of data visualizations, taking advantage of all of the platform's features. This is the main client application, and because of that, it has enabled all the data manipulation functionalities such as use of data filters, drill down, working online and offline (working offline allows embedding data in the visualizations for use in limited connectivity environments), and exporting visualizations to MS PowerPoint, PDF, and HTML. It is only available for Windows environment.

- **TIBCO Spotfire Web Player Server**: This offers users the possibility of accessing and interacting with visualizations created in TIBCO Spotfire Professional. The existence of this web application enables the usage of an Internet browser as a client, allowing for thin client access where no software has to be installed on the user's machine. Please be aware that the visualizations cannot be created or altered this way. They can only be accessed in a read-only mode, where all rules are enabled, as well as data is drill down. Since it is developed in ASP.NET, this server must be deployed in a Microsoft IIS server, and so it is restricted to Microsoft Windows environments.

- **Server Database**: This database is accessed by the TIBCO Spotfire Server for storage of server information. It should not be confused with the data stores that the platform can access to fetch data from, and build visualizations. Only two vendor databases are supported for this role: Oracle Database and Microsoft SQL Server.

- **TIBCO Spotfire Web Player Client**: These are thin clients to the Web Player Server. Several Internet browsers can be used on various operating systems (Microsoft Internet Explorer on Windows, Mozilla Firefox on Windows and Mac OS, Google Chrome on Windows and Android, and so on). TIBCO has also made available an application for iPad, which is available in **iTunes**.

> For more details on the iPad client application, please navigate to:
> `https://itunes.apple.com/en/app/spotfire-analytics/`
> `id417436823?mt=8`

Summary

In this chapter, we introduced the main attributes of the Spotfire platform in the scope of visual analytics, and we detailed the platform's underlying architecture. We will now proceed to the next chapter, where you will learn the steps to install and configure a development environment.

2

Setting Up a Development Environment

This chapter details the steps for the creation of a **TIBCO Spotfire 5.5** development environment. It will focus solely on the installation of the components presented in the previous chapter, in a single standalone Windows environment.

Installation prerequisites

Next we will present the main hardware and software installation prerequisites according to what is specified by TIBCO. If you require a more extensive and detailed list, please navigate to the following location: `http://support.spotfire.com/sr.asp`.

The requisite values presented for hardware are to be considered for multiuser environments, such as test and production. For local development environments, less powerful configurations can be used.

Hardware

Following are the hardware requirements for the three Spotfire platform components covered in this chapter. All the values were defined by TIBCO, and are available at the following location: `http://support.spotfire.com/sr_spotfire55.asp`

- **TIBCO Spotfire Server**: The requirements for this component are:
    - Processor: Intel Pentium 4 or equivalent, 2 GHz or higher
    - RAM: 4 GB recommended (2 GB minimum)
    - Hard disk: 1.5 GB

- **TIBCO Spotfire Professional**: The requirements for this component are:
    - ° Processor: 2 GHz, Multicore, 64-bit (1 GHz, 32 or 64-bit minimum)
    - ° RAM: 4 GB recommended (2 GB minimum)
    - ° Hard disk: 1.4 GB

- **TIBCO Spotfire Web Player Server**: The requirements for this component are:
    - ° Processor: 2 GHz, 6 physical cores or more, 64-bit (2 GHz, 64-bit minimum)
    - ° RAM: 12 GB recommended (6 GB minimum)
    - ° Hard disk: 1.2 GB

Software

Following are the software requirements for the three Spotfire platform components covered in this chapter. Just like the hardware requirements, all the values were defined by TIBCO, and are available at the following location: `http://support.spotfire.com/sr_spotfire55.asp`

- **TIBCO Spotfire Server**: The requirements are as follows:
    - ° Operating system: Microsoft Windows Server 2012; Microsoft Windows Server 2008 R2; Microsoft Windows Server 2008; Oracle (Sun) Solaris 10; Oracle (Sun) Solaris 9; Red Hat Enterprise Linux 6; Red Hat Enterprise Linux 5; SUSE Linux Enterprise Server 11
    - ° Server database: Oracle 11g Release 2 (11.2.0.x); Oracle 11g Release 1 (11.1.0.x); Oracle 10g Release 2 (10.2.0.x); Oracle 10g Release 1 (10.1.0.x); Oracle 10g XE (for development environments); Microsoft SQL Server 2012 (Datacenter, Parallel Data Warehouse, Enterprise, Standard, Workgroup, Express, Developer); Microsoft SQL Server 2008 R2 (Datacenter, Parallel Data Warehouse, Enterprise, Standard, Workgroup, Express, Developer); Microsoft SQL Server 2008 (Enterprise, Standard, Workgroup, Express, Developer); Microsoft SQL Server 2005 SP2 and higher (Enterprise, Standard, Workgroup, Express, Developer)

- **TIBCO Spotfire Professional**: The requirements are as follows:
    - ° Operating system: Microsoft Windows 8; Microsoft Windows 7; Microsoft Windows Vista SP2; Microsoft Windows XP SP3; XP x64 Edition SP2
    - ° .NET framework: Microsoft .NET Framework version 4.x

- **TIBCO Spotfire Web Player Server**: The requirements are as follows:
    - ° Operating system: Microsoft Windows Server 2012; Microsoft Windows Server 2008 R2; Microsoft Windows Server 2008 (64-bit versions required)
    - ° .NET framework: Microsoft .NET Framework version 4.x
    - ° Web server: Microsoft Internet Information Services Versions 7.0 through 8.0, with ASP.NET 4.x enabled

Installing the Spotfire environment

The step-by-step installation of a Spotfire platform environment presented in this chapter, is an example of Spotfire 5.5.0 installed in Windows 2008 64 bit (on the `C:` drive), and using a local Oracle 11*g* XE database (previously installed in `C:\oraclexe\app`). Please note that components can also be installed in separate machines, as long as the software and hardware prerequisites are respected.

For this book, a standalone Windows environment was chosen, not only because it is the simplest environment to configure and use, but also because of the operating system requirements for two of the TIBCO Spotfire components, TIBCO Spotfire Professional and TIBCO Spotfire Web Player Server.

TIBCO Spotfire Server

We will start by describing the installation process for Spotfire server, which will include the verification of the setup of the supporting database, and an overview of the product's main configuration tasks.

Verifying the DB installation

Before starting with the installation of the server, we must make sure the server database is correctly installed and configured.

> Because Oracle XE is not supported in 64-bit Windows environments, you might encounter some issues during the installation process. Two workarounds to overcome this limitation are:
>
> - During the installation, open `SQLNET.ORA` (located in `$ORACLE_HOME/network/admin`), and change the parameter `SQLNET.AUTHENTICATION_SERVICES` to `NONE`.
> - Ignore any error related to the registry, as the installation will succeed anyway.
>
> When the installation process finishes, Oracle XE will start and function correctly.

First, confirm `sqlplus` is in the environment `%PATH%`. This can be done by opening a command line and typing the following command followed by *Enter*:

```
sqlplus
```

If you get a result as the one presented next, then please make sure you add `sqlplus` to the `%PATH%`:

```
sqlplus is not recognized as an internal or external command, operable program or batch file.
```

As we opted for using Oracle XE, we must first tune the database to cope with the load generated by the Spotfire server. We must execute the following SQL instructions as SYSDBA:

```
alter system set processes=150 scope=spfile;

alter system set sessions=300 scope=spfile;

alter system set transactions=330 scope=spfile;

Shutdown immediate;

Startup;
```

> A user guide for SQL*Plus can be found at: `http://docs.oracle.com/cd/B19306_01/server.102/b14357/toc.htm`.
>
> Among others, it contains instructions of how to connect to a database as SYSDBA.

Downloading Spotfire Server

Now that we have confirmed that our database is configured and running properly, we will proceed with the installation of the Spotfire server.

Start by downloading the server installer `TIB_ss_5.5.0_win.zip` from the TIBCO download site (`http://download.tibco.com`, **Products Home | TIBCO Spotfire Server-5.5.0 | TIBCO Spotfire Server Software 5.5.0**) and unpack it to `C:\temp\ TIB_ss_5.5.0_win`.

Preparing the server database

Before actually installing the Spotfire Server software, we need to prepare the server database by creating all the required server objects. For that, we will copy the database creation scripts of `C:\temp\TIB_ss_5.5.0_win\scripts\oracle_ install` to a temporary folder for update and execution. We can use the folder `C:\ temp\oracle_install`.

Open the script `C:\temp\oracle_install\create_databases.bat` in a text editor, and locate the following section:

```
rem Set these variable to reflect the local environment:
set ROOTFOLDER=<ROOTFOLDER>
set CONNECTIDENTIFIER=<SID>
set ADMINNAME=system
set ADMINPASSWORD=<ADMINPASSWORD>
set SERVERDB_USER=<SERVERDB_USER>
set SERVERDB_PASSWORD=<SERVERDB_PASSWORD>
set SERVER_DATA_TABLESPACE=SPOTFIRE_DATA
set SERVER_TEMP_TABLESPACE=SPOTFIRE_TEMP
```

We must perform several configuration changes, by replacing the current values:

- `set ROOTFOLDER=C:\oraclexe\app\oracle\oradata\XE`: It sets the location for the creation of Spotfire Server tablespace files

- `set CONNECTIDENTIFIER=XE`: Here, we set the **System Identifier for the Database** instance (**SID**) of the database

- `set ADMINNAME=system`: Here, we set the **Database Administrator** (**DBA**) user. We will leave it as it is

- `set ADMINPASSWORD=spotfire`: Here, we set the DBA user password. My `SYSTEM` user has password `spotfire`

- `set SERVERDB_USER=spotfire`: Here, we set the schema (user) to create TIBCO Spotfire Server

- `set SERVERDB_PASSWORD=spotfire`: Here, we set the password of TIBCO Spotfire Server user

- `set SERVER_DATA_TABLESPACE=SPOTFIRE_DATA`: Here, we set the name of tablespace to be created

- set `SERVER_TEMP_TABLESPACE=SPOTFIRE_TEMP`: Here, we set the name of the temporary tablespace to be created

There is also a special configuration to set up the demo data, but as we have no intention to install it, we will not update the following section:

```
rem Demo data parameters
set INSTALL_DEMODATA=no
set DEMODB_USER=spotfire_demodata
set DEMODB_PASSWORD=spotfire_demodata
```

We will leave `INSTALL_DEMODATA` with the value `no`.

Now, we will save and close the script, and thus the changes are complete.
To execute it, please use a command line for the following commands:

```
cd C:\temp\oracle_install
create_databases.bat
```

During execution and on completion, the following output will be presented:

```
Creating Spotfire Server table spaces and user
Creating Spotfire Server tables
Populating Spotfire Server tables
Spotfire Server demo database user and data will not be created
----------------------------------------------------------------
Please review the log file (log.txt) for any errors or warnings!
```

After the script terminates its execution, we should make sure it succeeded.
We will do this by logging into the newly created schema, validating the presence of the new tablespaces, and checking the script log.

- You can log in into the database with the new `spotfire` user through SQL*Plus
  (or another tool).

- After logging in, you can list the database tablespaces with the SQL command `SELECT * FROM user_tablespaces;`. Tablespaces `SPOTFIRE_DATA` and `SPOTFIRE_TEMP` must be listed.

- Finally, you will find the script log under `C:\temp\oracle_install\log.txt`. Check it for errors.

If all three validation steps were successful, the Spotfire Server database is now prepared. As we will not require the database creation scripts again, we can remove `C:\temp\oracle_install`.
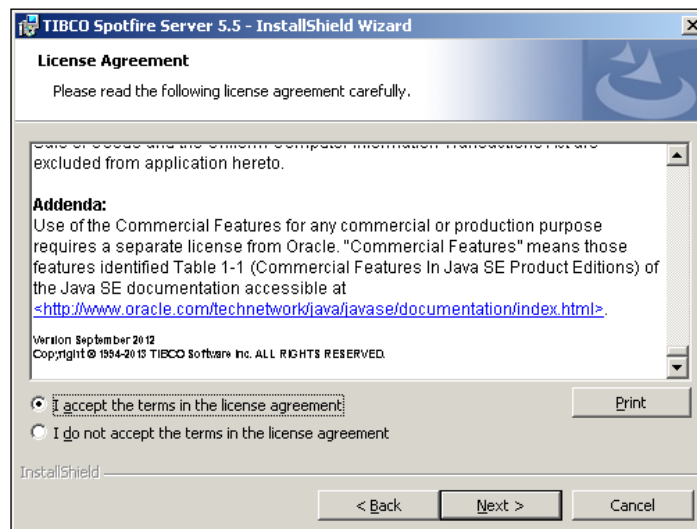
Installing Spotfire Server

The easiest installation mode for TIBCO products in Windows development machines tends to be the interactive one. To trigger it, double-click on `C:\temp\TIB_ss_5.5.0_win\setup-win64.exe` (or `setup-win32.exe` for 32-bit OS).
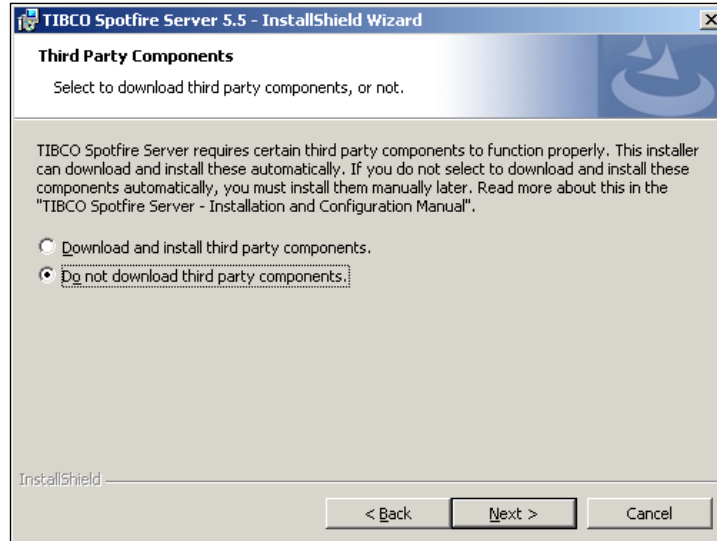
The following dialog box will be presented:

Click on **Next >**. The license agreement will be displayed, as shown in the following screenshot:

Read the conditions, accept them, and click on **Next >**. The **Third Party Components** screen will be pop up, as shown in the following screenshot:

The third-party components are required solely for configuring **NT LAN Manager** (**NTLM**). As we will not use this authentication mode, we can skip this installation step. Select **Do not download third party components** and click on **Next >**.

The screen for installation location will be presented, as shown in the following screenshot:
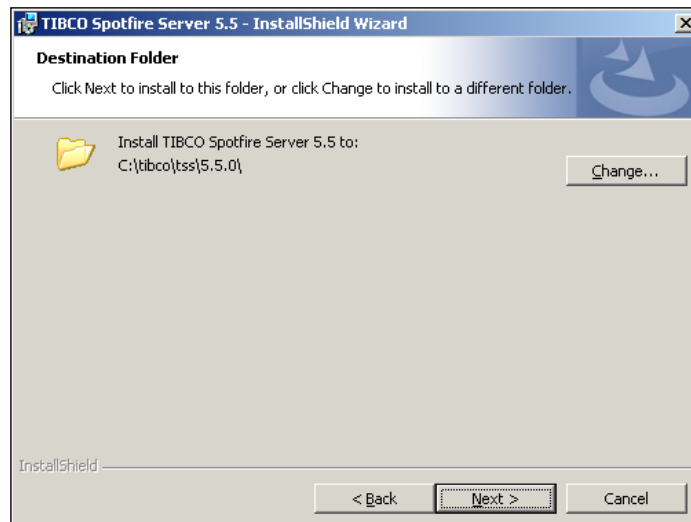
We will keep the default installation path. Click on **Next**, and the following screenshot appears:
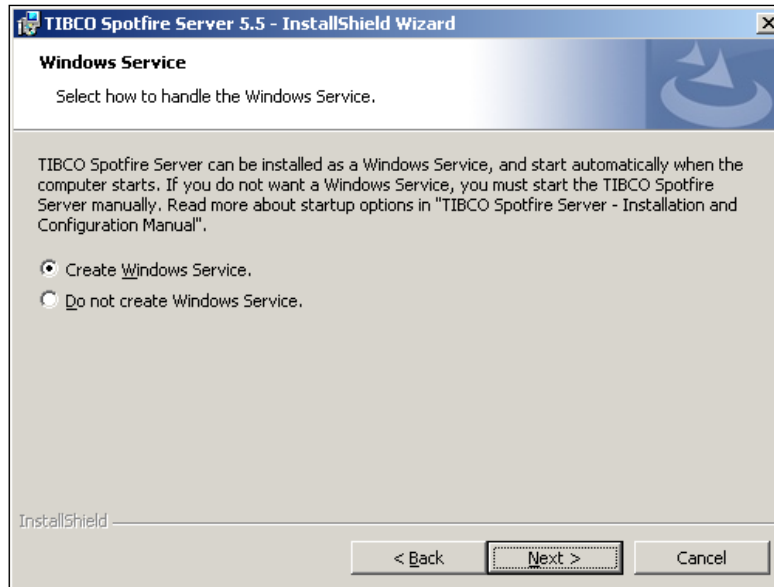
Keep the option **Create Windows Service** selected and click on **Next >**. The **Spotfire Server Port** selection screen will pop up, as shown in the following screenshot:
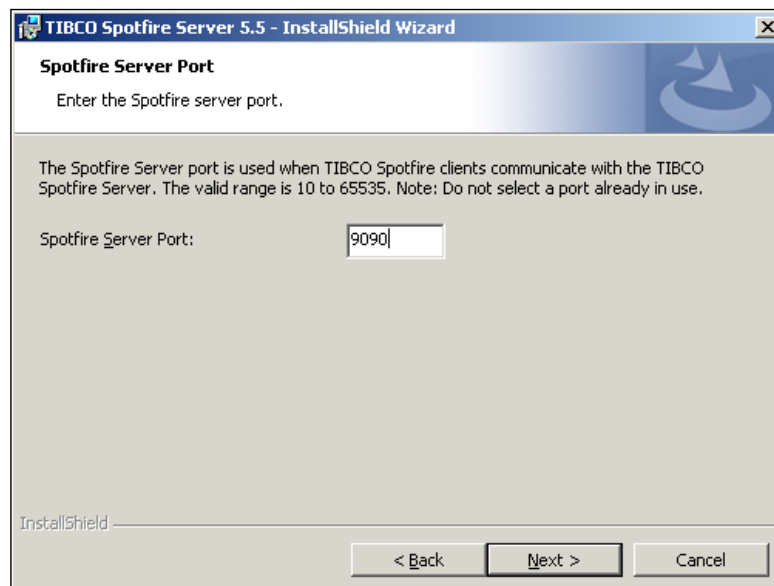
Please change the **Spotfire Server Port** value to `9090`.

Another free port can also be used. Just make sure that you do not choose port 80, as it belongs to Microsoft IIS (used later for TIBCO Spotfire Web Player Server). Also, be aware that Oracle XE uses port 8080 for its TNS listener.

When you are done, click on **Next >**. You will be presented with an installation confirmation screen, as shown in the following screenshot:
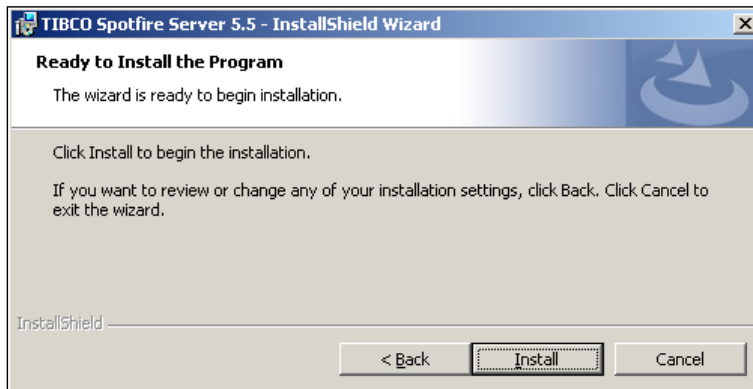
The installation configuration has been set. Proceed by clicking on **Install**.

When the installation concludes, you will be shown the following screenshot:
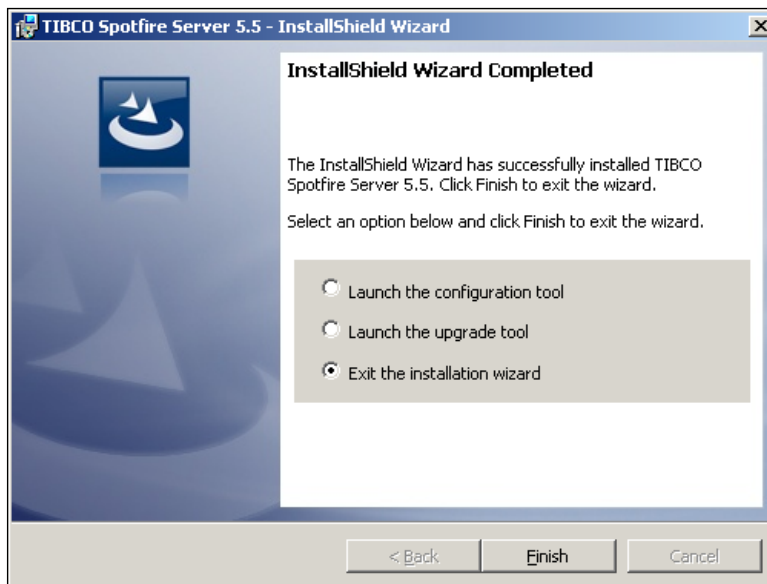
The installation is now complete. Select **Exit the installation wizard** as we will explicitly launch the configuration tool later. Click on **Finish**.

> After installation completion, the TIBCO provided server installation manual can be found under `C:\tibco\tss\5.5.0\tomcat\webapps\spotfire\support\manuals\TIB_SSDoc_ServerInstallationManual.pdf`.

The **Launch the upgrade tool** option is used for upgrading existing Spotfire versions. This is out of the scope of this book and will not be discussed further.

> Here is a neat trick to force the installed Spotfire Windows service to start automatically only after the database startup completes. Our XE database also has a Windows service with automatic startup. We will create a startup dependency between both services.
>
> Run (as administrator) the following command in the command prompt:
>
> **`sc config Tss550 depend= OracleServiceXE`**

The server is now installed, so we can remove `C:\temp\TIB_ss_5.5.0_win`.

Installing database vendor drivers

TIBCO recommends users to install the **JDBC** drivers of the vendor database (although DataDirect drivers are embedded in the Spotfire product).

The Oracle XE installation already contains the required JDBC drivers to access it. You can find `ojdbc6.jar` at: `C:\oraclexe\app\oracle\product\11.2.0\server\jdbc\lib`.

As an alternative, you can download it from `http://www.oracle.com/technetwork/database/features/jdbc/index-091264.html`.

Place the JAR file in the `C:\tibco\tss\5.5.0\tomcat\lib` folder.

Configuring Spotfire Server

Even though the Spotfire Server software was correctly installed, we are still unable to start it. It will fail with an error.

This is due to the lack of server configuration, as we only performed a few configuration tasks during the installation. The bulk of configuration must be done after installation, through the server configuration tool.

To start this configuration tool, run `C:\tibco\tss\5.5.0\tomcat\bin\uiconfig.bat`. The following console is started:
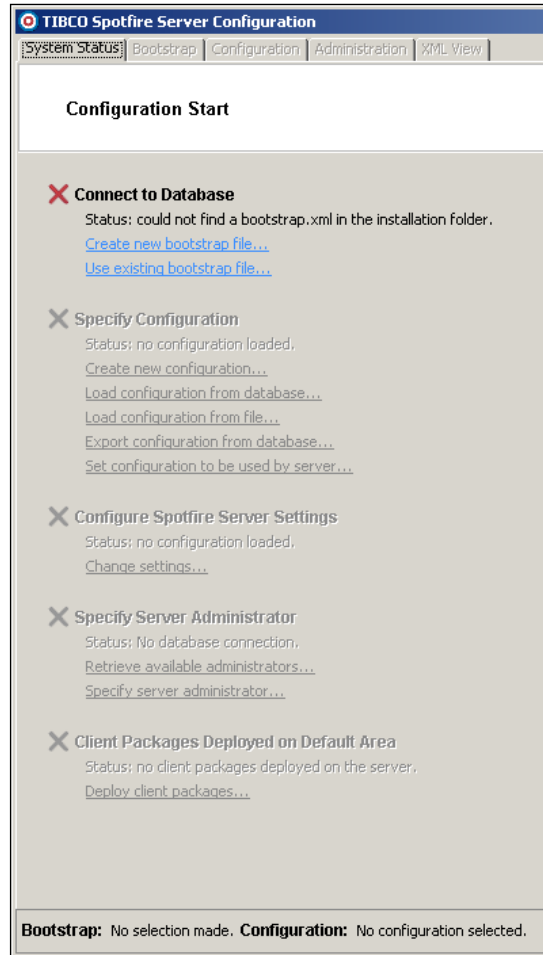
> As an alternative, `C:\tibco\tss\5.5.0\tomcat\bin\config.bat` can be used. But please keep in mind that this alternative is script oriented.

The first configuration must be the creation of a bootstrap file. As specified in configuration console, the bootstrap file specifies connection details to the Spotfire database and is needed for further configuration.

Start by selecting the option **Create new bootstrap file**. When configuring this file (`bootstrap.xml` stored in the XML format), please pay attention to the following settings:

- JDBC driver: Replace **Bundled DataDirect** with **Oracle thin** (which we installed before)

- Database username and password: these will be respectively `SERVERDB_USER` and `SERVERDB_PASSWORD` from the schema creation scripts. In our case, it is `spotfire` for both

- Hostname: Local database, so `localhost` is fine

- SID: **XE**, our local database is `SID`

- Configuration tool password: Select a password to protect the configuration. In our case it will be `spotfire` again

- Server name: Select a name for the server. In our case, it will be named `SpotfireDev`

After filling in all the required values, click on the button **Create Bootstrap**.

A success pop up will be displayed, as shown in the following screenshot, and the following three tabs will become enabled in the configuration tool: **Bootstrap**, **Configuration**, and **XML View**. We will now focus on the **Configuration** tab.
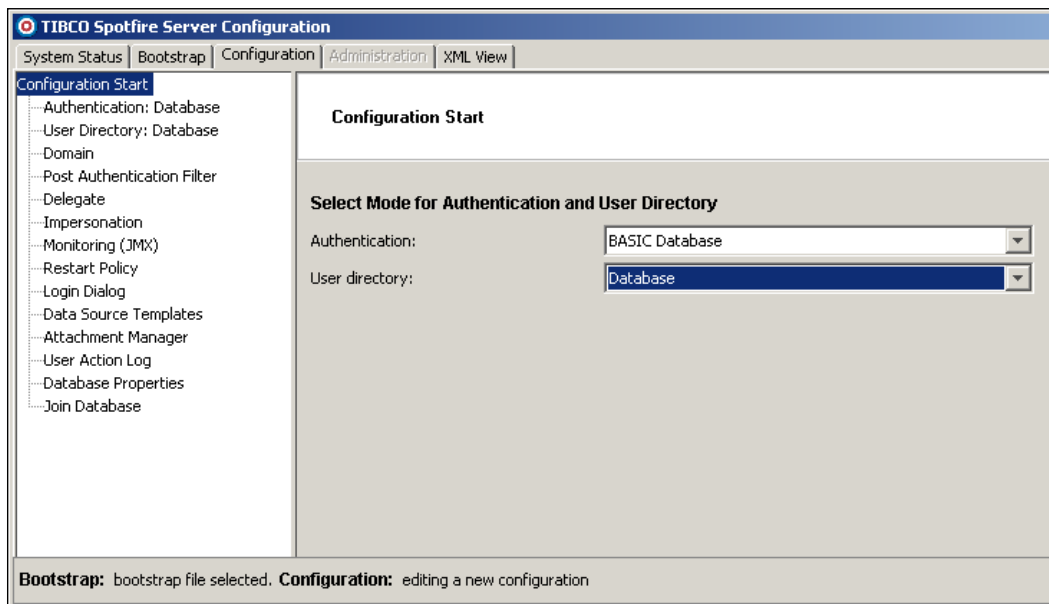
The simplest alternative for authentication mode is **BASIC Database**, and that is the one we will use. This means Spotfire Server credentials will be stored in the Spotfire Server database.

For our development environment, all other default configuration settings are suitable.

Click on the **Save configuration** button, and then select **Database** as destination for these configurations.

For any configuration changes to become active we must restart our server manually (this can be changed using the tab **Configuration | Restart Policy**). As our server is still not running, we will not worry about this now.

If we go back to the **System Status** tab in the server configuration tool, we will see the following status:
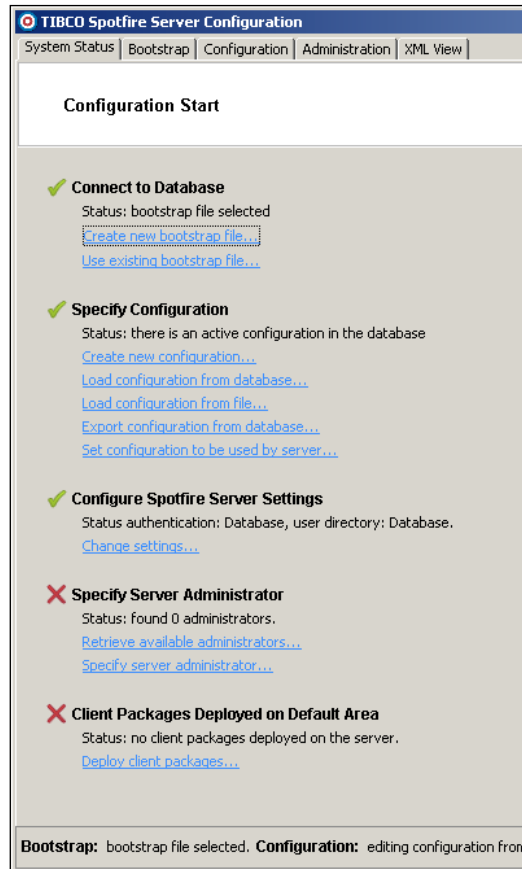
We are still missing the creation of a Spotfire Server administrator and we are still required to deploy the client packages in the server (for the client updates mentioned in the previous chapter).

For the creation of an administrator user, we need to go to the **Administration** tab and create a user. For simplicity reasons, we will create a user with the same credentials as our Spotfire Server DB schema (user `spotfire`), as shown in the following screenshot:
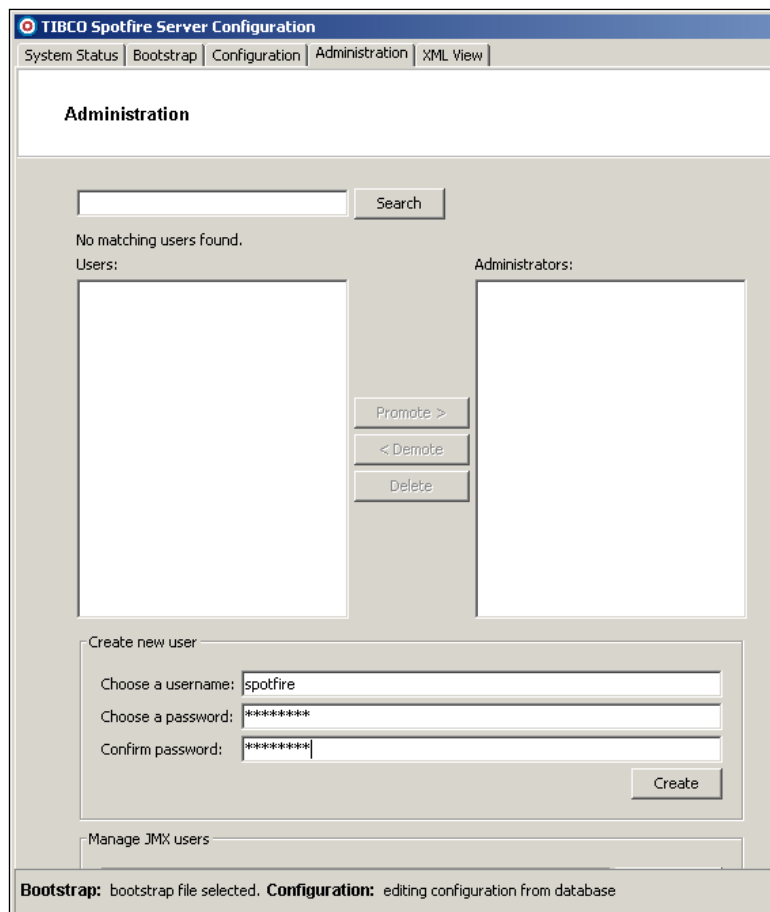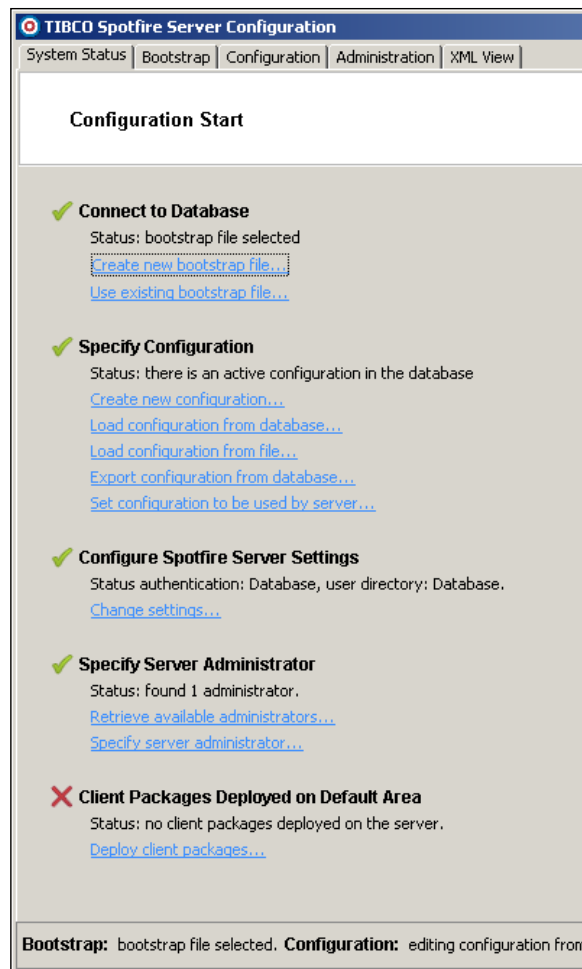

After defining `spotfire` as username and password, click on **Create**.

Now select the newly created user and promote it to an administrator. You will end up with a configuration as shown in the following screenshot:
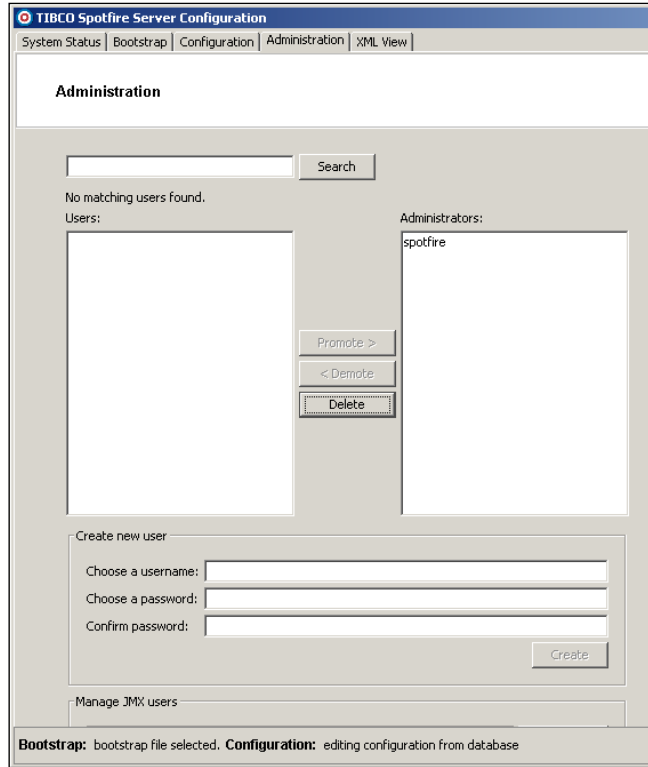
Our changes were persisted to the server database, and if we once again navigate to the **System Status** tab, we have one less issue, that is we now have a server administrator. Your configuration will now be similar to the following screenshot:

TIBCO Spotfire Server Configuration

System Status | Bootstrap | Configuration | Administration | XML View

Configuration Start

✔ **Connect to Database**
Status: bootstrap file selected
Create new bootstrap file...
Use existing bootstrap file...

✔ **Specify Configuration**
Status: there is an active configuration in the database
Create new configuration...
Load configuration from database...
Load configuration from file...
Export configuration from database...
Set configuration to be used by server...

✔ **Configure Spotfire Server Settings**
Status authentication: Database, user directory: Database.
Change settings...

✔ **Specify Server Administrator**
Status: found 1 administrator.
Retrieve available administrators...
Specify server administrator...

✘ **Client Packages Deployed on Default Area**
Status: no client packages deployed on the server.
Deploy client packages...

Bootstrap: bootstrap file selected. **Configuration:** editing configuration from

We will deal with the missing **Client Packages Displayed on Default Area** during the installation of the next component – TIBCO Spotfire Professional. This is the only server configuration activity that we will pick up later.

> We have to handle the installation of the client packages during the installation of Spotfire Professional, because the client packages' source binary comes bundled with the source binary of Spotfire Professional.

You should now be able to successfully start the TIBCO Spotfire 5.5 Windows service. For 32 bit environments, some adjustments might be required on the JVM settings (such as reducing the 1536 MB memory allocation to 1024 MB). For detailed instructions, refer to page 138 of the tool's installation manual under *Chapter 11.3*, *Modifying the Virtual Memory*, `TIB_SSDoc_ServerInstallationManual.pdf`.

In case of any errors on service startup, check the logs under `C:\tibco\tss\5.5.0\tomcat\logs`.

TIBCO Spotfire Professional

Spotfire Server is installed, and we will now proceed with the installation of Spotfire Professional.

Downloading Spotfire Professional

Download the Spotfire Professional installer `TIB_spotdeploy_5.5.0_win.zip` from the TIBCO download (`http://download.tibco.com`, **Products Home | TIBCO Spotfire Professional-5.5.0 | TIBCO Spotfire Deployment Kit Software 5.5.0**) and extract it to `C:\temp\TIB_spotdeploy_5.5.0_win`.

> The TIBCO provided deployment administration manual can be found under `C:\temp\TIB_spotdeploy_5.5.0_win\Documentation\TIB_SpotDeploy_5.5.0_DeployAdminManual.pdf`. Please move it to a different directory as you might need to refer to it in the future.

Installing Spotfire Server client packages

The second step in the fat client installation is the wrap up of server installation (as stated in a previous section). We still have a pending task to complete, which is the deployment of client packages on the default area (work area).

If you have closed the server configuration console, open it again, switch to the **System Status** tab, and select the option **Deploy client packages...**.

On the pop up, click on the **Browse** button and navigate to `C:\temp\TIB_spotdeploy_5.5.0_win\Products\TIBCO Spotfire Distribution\Spotfire.Dxp.sdn`.

Leave **Production** as work area—we will rename it later, when we approach the subject of **Server Administration**. Click on **Deploy**.

The **System Status** window on the Spotfire Server configuration console should now look as the following screenshot:
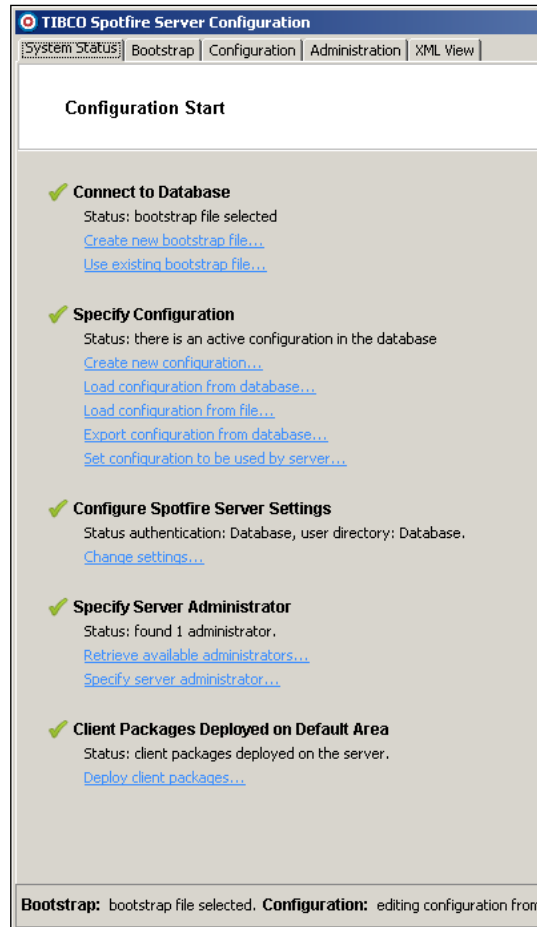

The Server configuration has been finished; we will now proceed with the installation of Spotfire Professional.

For an alternative mode of deploying the client packages (using the Server Administration console), please refer to page 8 of *Chapter 3.2*, `TIB_SpotDeploy_5.5.0_DeployAdminManual.pdf`

Installing Spotfire Professional

To trigger the installation of Spotfire Professional, double-click on `C:\temp\TIB_ spotdeploy_5.5.0_win\Products\TIBCO Spotfire Installer\setup-5.5.0.exe`. This will begin the procedure, and the following screenshot will be presented:
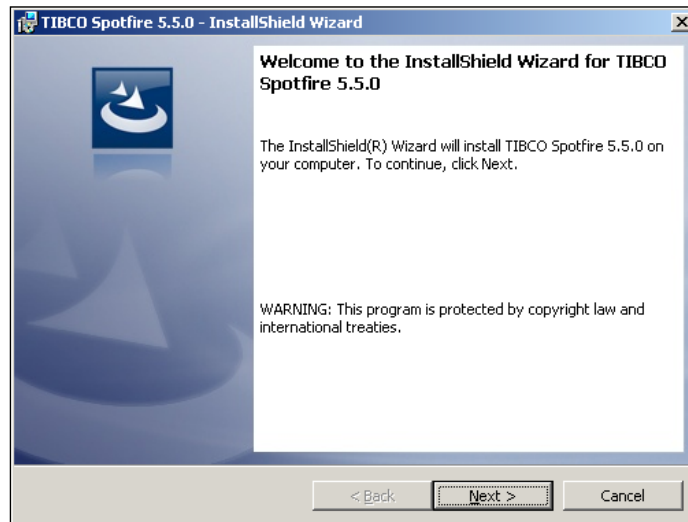
Click on **Next >**. The license agreement will be displayed, as shown in the following screenshot:
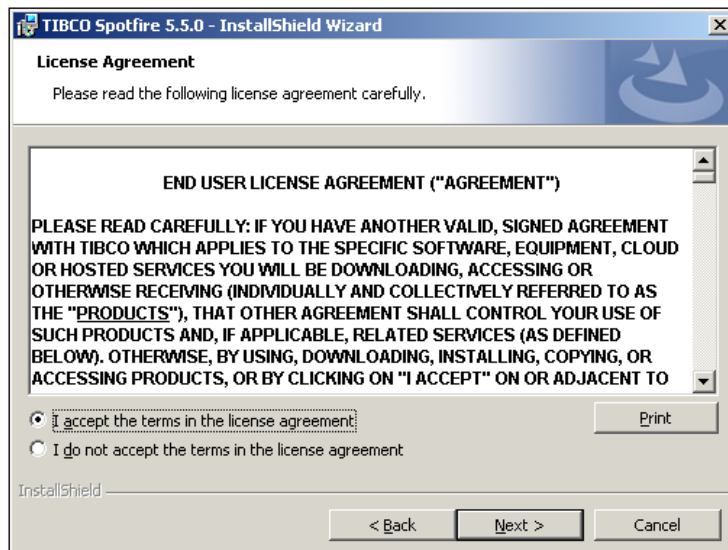
Read the conditions, accept them, and click on **Next >**. The Spotfire Server URL configuration screen will pop up, as shown in the following screenshot:
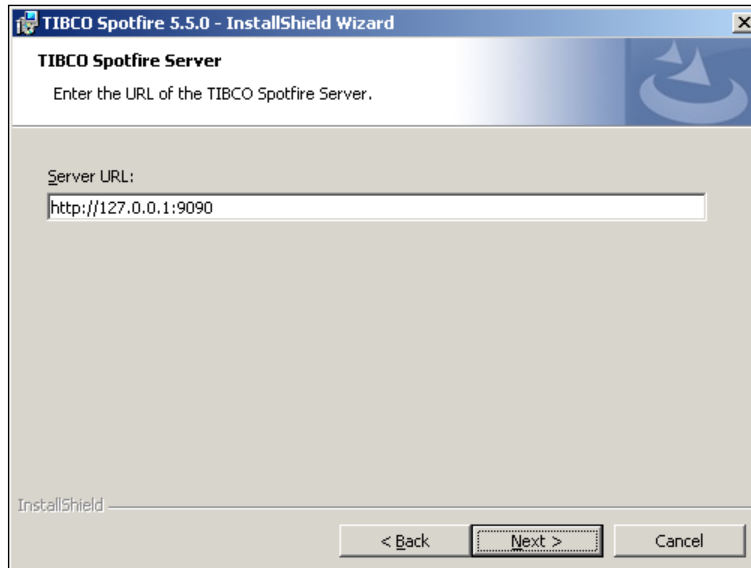
Fill the **Server URL** field with `http://127.0.0.1:9090` or `http://localhost:9090`, and click on **Next >**. You will be presented with an installation confirmation screen, as shown in the following screenshot:
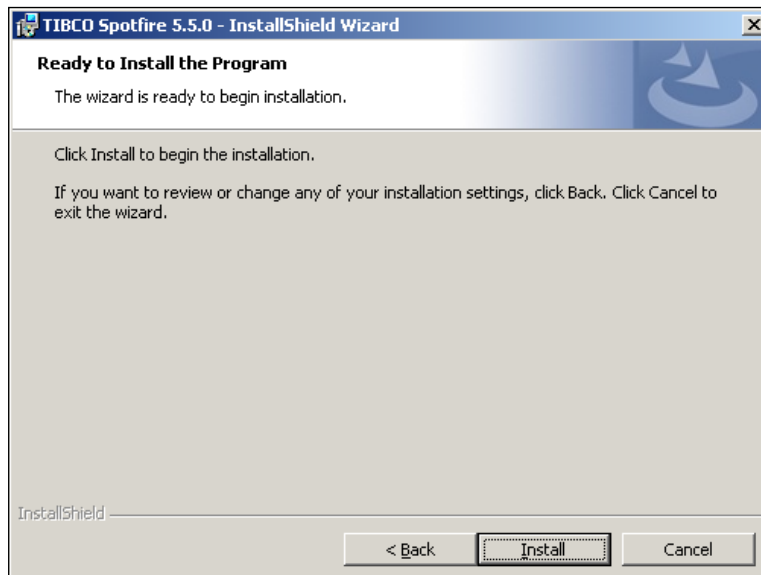
Proceed by clicking on **Install**.

When installation concludes, you will be shown the following screenshot:
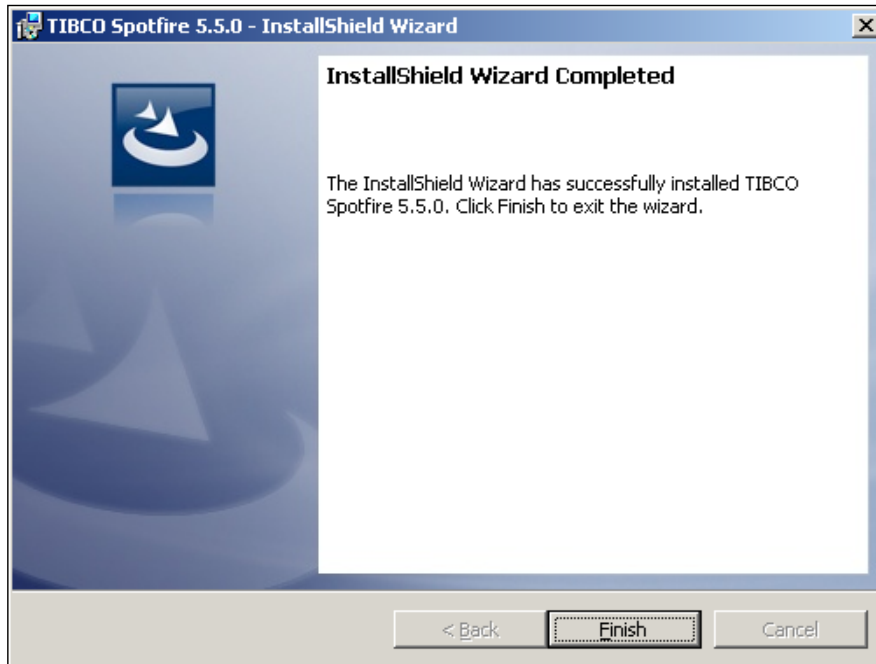
You can now click on **Finish** to exit the installation wizard.

Spotfire Professional is now installed. Since we will not need it anymore, we can remove `C:\temp\TIB_spotdeploy_5.5.0_win`.

The best way to test the installation is to start the Spotfire Professional tool. There should be one shortcut on the Desktop, and one in the Windows Start menu.

> For logging in, make sure that the selected server has URL `http://127.0.0.1:9090` (or `http://localhost:9090`). If it is set to a different value, use the option **Manage Servers** to change it to the correct one.

If you get any pop ups suggesting updates, proceed by clicking on **Update Now**. They will be related to the client packages existing on the server, which are newer versions than the software we just installed.

An update pop-up looks similar to the following screenshot:
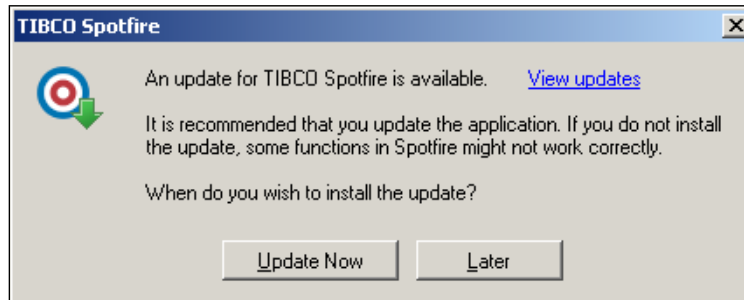
The installation procedure of TIBCO Spotfire Professional — the fat client — is now finished.

TIBCO Spotfire Web Player Server

Since Spotfire Professional installation is concluded, we can now proceed with Spotfire Web Player Server.

Configuring Microsoft IIS

The Spotfire Web Payer Server will be deployed in Microsoft IIS; so before proceeding with this task, we must make sure that the IIS is configured according to the requirements of TIBCO.

The first configuration task consists of enabling the active scripting. For this, open Local Group Policy Editor and run `gpedit.msc`. Navigate to **Local Computer Policy** | **Computer Configuration** | **Administrative Templates** | **Windows Components** | **Internet Explorer** | **Internet Control Panel** | **Security Page** | **Internet Zone**, right-click on **Allow active scripting**, and click on **Edit**.

You will be presented with a pop up as shown in the following screenshot:
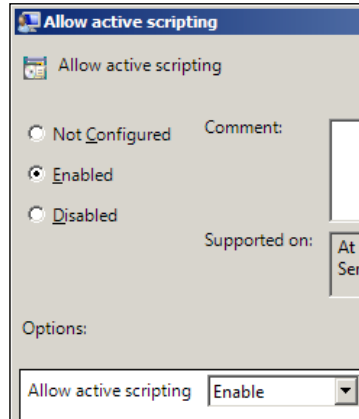
Select the **Enabled** option and also change the **Allow active scripting** option to **Enable**. Complete the step by clicking on **OK**.

With active scripting enabled, we just need to enable IIS. Commence by starting Windows 2008 Server Manager. When it is running, right-click on **Roles** and then click on **Add Roles**. Click on **Next** on the **Before you begin** screen.

In the **Select Server Roles** screen, select **Web Server (IIS)**. Click on **Next** twice.

In the **Select Role Services** screen, besides the default selections, also select the following:

- **ASP.NET**
- **.NET Extensibility**
- **ISAPI Extensions**
- **ISAPI Filters**
- **Basic Authentication**
- **Windows Authentication**
- **IIS Management Scripts and Tools**
- **IIS 6 Metabase Compatibility**
- **IIS 6 WMI Compatibility**
- **IIS 6 Scripting Tools**

Click on **Next**, followed by **Install**. Finally, when done, click on **Close**.

For other operating systems, please refer to the product's documentation
`TIB_SpotWebPlay_5.5.0_InstallationManual.pdf` for the instructions
related to installation.

Downloading Spotfire Web Player Server

Download the Spotfire Web Player Server installer `TIB_spotwebplay_5.5.0_win.`
`zip` from the TIBCO download site (`http://download.tibco.com`, navigate to
Products Home | TIBCO Spotfire Web Player-5.5.0 | TIBCO Spotfire Web Player
Software 5.5.0) and extract it to `C:\temp\TIB_spotwebplay_5.5.0_win`.

> The TIBCO provided Web Player Server installation manual
> can be found under `C:\temp\TIB_spotwebplay_5.5.0_`
> `win\Documentation\TIB_SpotWebPlay_5.5.0_`
> `InstallationManual.pdf`. Please move it to a different
> directory as you might need to refer to it in the future.

Installing Spotfire Web Player Server

Double-click on `C:\temp\TIB_spotwebplay_5.5.0_win\Products\TIBCO`
`Spotfire Web Player Installer\setup-5.5.0.exe`. This will trigger the
installation procedure.

The following screenshot will be presented:
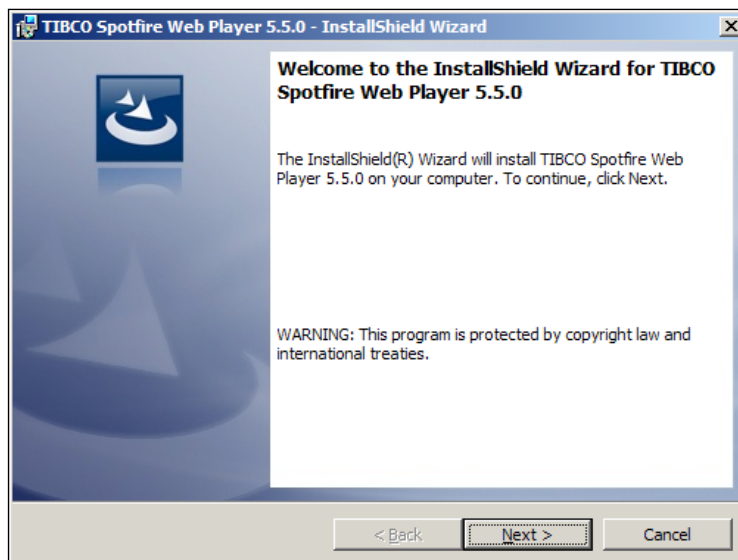
Click on **Next >**. The license agreement will be displayed, as shown in the following screenshot:

Read the conditions, accept them, and click on **Next >**. The installation location screen, as shown in the following screenshot, will then be presented:
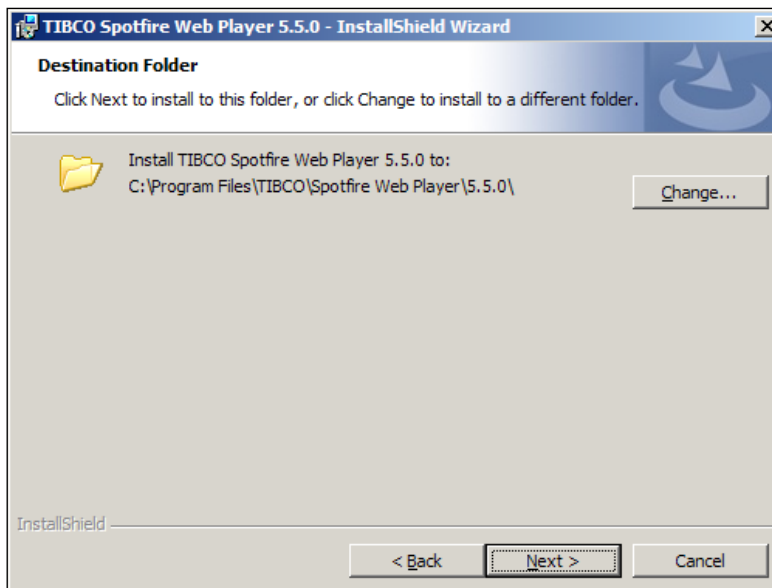
We will maintain the default installation path, so click on **Next >**. The screen for IIS configuration will be displayed as shown in the following screenshot:
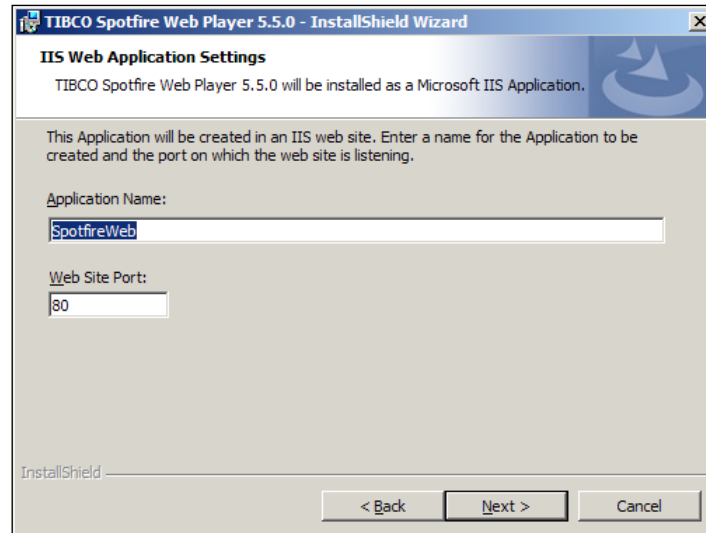

We will keep the default **Application Name** (which is recommended by TIBCO) and IIS is using port 80, so we will keep the value of **Web Site Port** as well. Click on **Next >**.

The Spotfire Server URL configuration screen will pop up, as shown in the following screenshot:
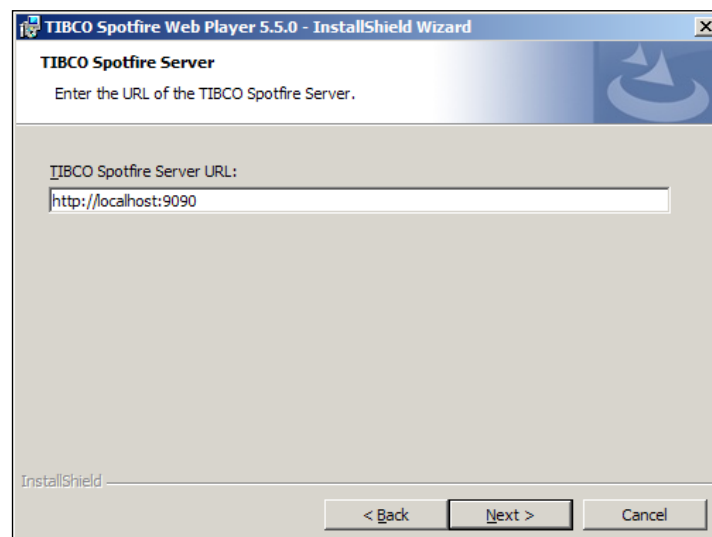
Fill the **TIBCO Spotfire Server URL** field (with the value `http://127.0.0.1:9090` or `http://localhost:9090`) and click on **Next >**. The **Spotfire Server Administrator Email** configuration screen will pop up, as shown in the following screenshot:
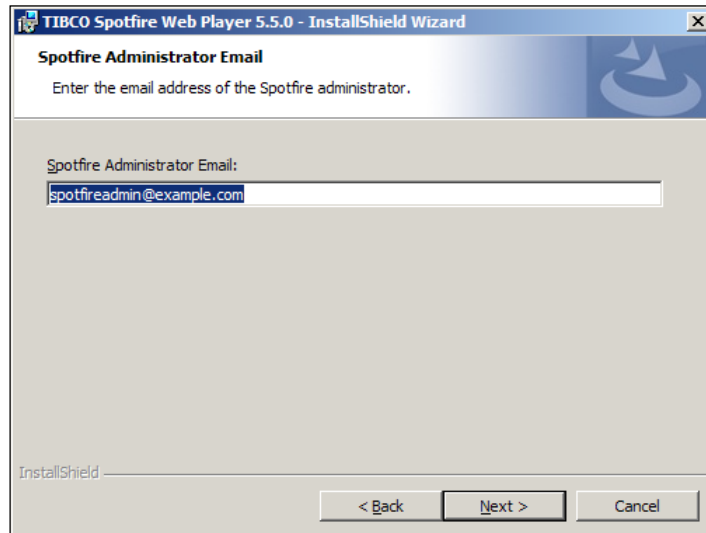
We will not configure an e-mail account for the Spotfire Administrator, so you can leave the configuration the way it is and click on **Next >**.

You will be presented with an installation confirmation screen, as shown in the following screenshot:
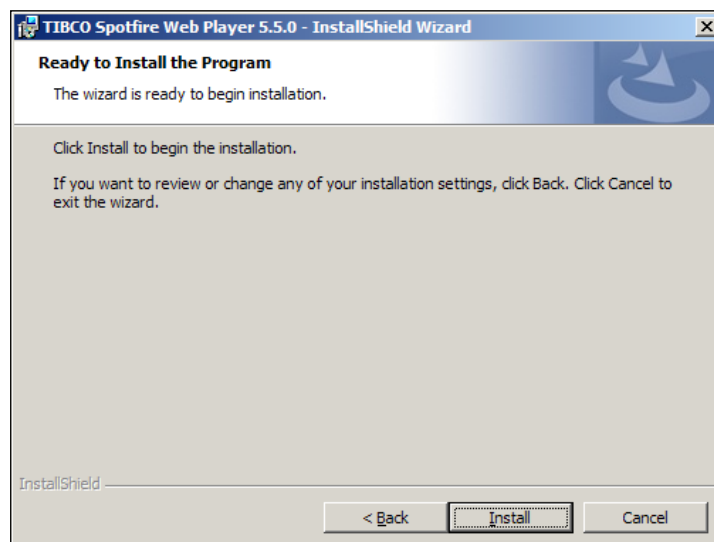
The installation configuration has been set. Proceed by clicking on **Install**.

When installation concludes, you will be shown the following screenshot:
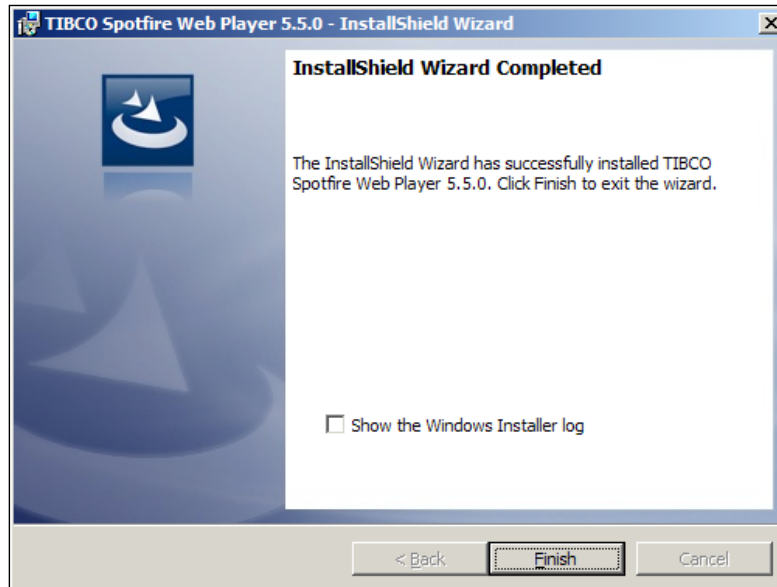
The installation is now complete; click on **Finish** to terminate the installer.

At this point, the TIBCO Spotfire Web Player Server is installed in the machine, configured with username and password authentication type. Although this is enough for our development environment, for a live environment, other alternatives should be used (please refer page 27 of `TIB_SpotWebPlay_5.5.0_ InstallationManual.pdf` for more on the options available to you).

To test the installation, open a web browser and access the following URL: `http://localhost/SpotfireWeb`. A Spotfire Web Server login page should be displayed.

Since we no longer require `C:\temp\TIB_spotwebplay_5.5.0_win`, we can remove it.

> If you get an HTTP error while accessing `http://localhost/SpotfireWeb`, you may require to perform an extra step to complete the ASP.NET installation.
>
> Please run the following command at the command prompt, as administrator:
>
> `C:\Windows\Microsoft.NET\Framework64\v4.0.30319\aspnet_regiis.exe /i`

Final installation status

If you followed the previous recommended installation configuration, then your environment should have the following characteristics:

- Oracle XE:
    - Connection details: server = localhost; port = 1521; SID = XE
    - System users: `SYS` [password = `spotfire`]; `SYSTEM` [password = `spotfire`]
    - Spotfire Server user = `spotfire` [password = `spotfire`]
- Spotfire Server:
    - URL: `http://localhost:9090`
    - Administration user: `spotfire` [password = `spotfire`]
    - Bootstrap configuration file password: `spotfire`
    - Administration console: `http://localhost:9090/spotfire/`
- Spotfire Web Player Server:
    - URL: `http://localhost/SpotfireWeb/` [port 80]

Configuration and administration post installation

Now that we have finished the installation process for the platform components, we will move on to configuration and administration tasks.

We will only focus on Spotfire Server and Spotfire Web Player Server configuration, as only these offer an extended range of options.

TIBCO Spotfire Server

Focusing on the server, we will now look at the topics of extra configuration steps, startup administration, and logging.

Configuration

The topic of server configuration already came up during the installation. In my opinion, it is an intrinsic part of the server installation procedure and this is why I placed it in that section of the book. Just keep in mind how to start the server configuration console:

- Method 1: Executing (double-clicking) `C:\tibco\tss\5.5.0\tomcat\bin\uiconfig.bat`
- Method 2: Navigating to the Windows Start menu: **Start** | **All Programs** | **TIBCO Sportire Server 5.5** | **Configure TIBCO Spotfire Server**

Administration

You might remember that during the installation of the Spotfire Server, we opted for the creation of a Server Windows service. You may use it to start and stop the Spotfire Server (named as **Tss550)**. Also, beware that it was created with **Startup type** as **Automatic**, so it will start up whenever you boot your PC.

However, if you want to manually start up the application, you have the alternative of executing the server's Tomcat start-up script `C:\tibco\tss\5.5.0\tomcat\bin\startuptomcat`.

After starting the server, to access its Administration console, navigate to `http://localhost:9090/spotfire`, and the following screenshot will be shown:
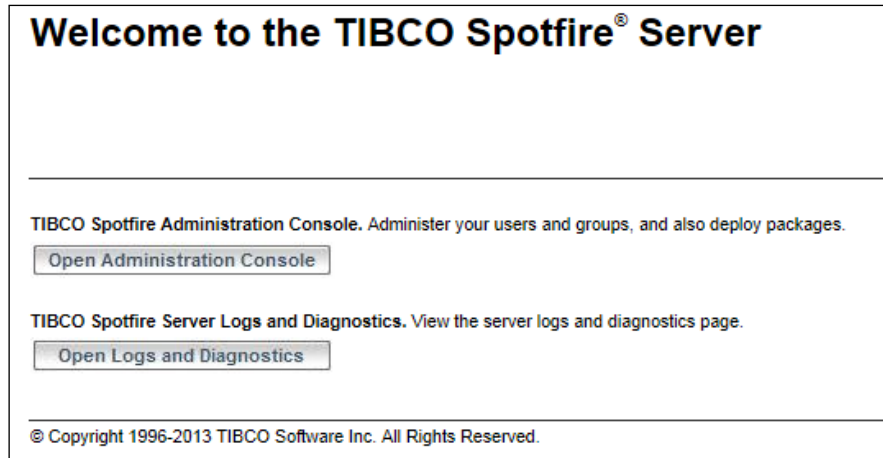

The user is presented with two distinct options: **Open Administration Console** or **Open Logs and Diagnostics**.

By selecting any of these options, you will be requested to log in. Feel free to do so, resorting to the created administration user: `spotfire`.

Administration console

The direct link for this option is `http://localhost:9090/spotfire/administration/`, and after logging in, you will be presented with the menu shown in the following screenshot:
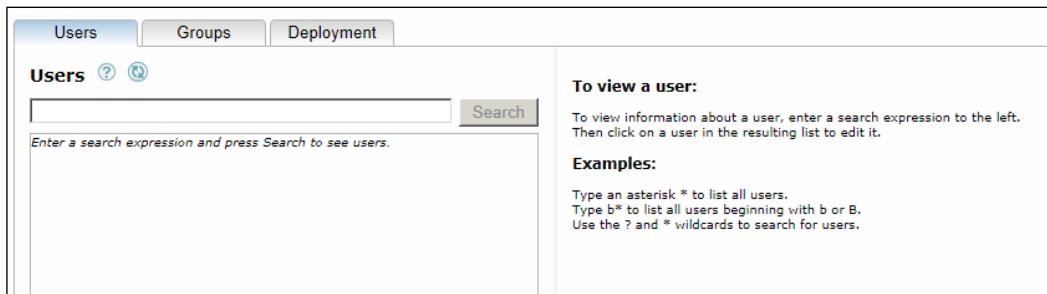
Three distinct operations can be performed in the administration console: user management, group management, and deployment management.

Try out the process of creating a user by clicking on **New User**. Create a user named **testUser** with the password **testUser**. The result will be the displayed as shown in the following screenshot:
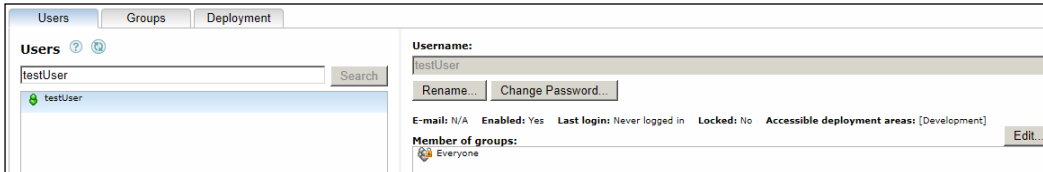
You can also manage the group to which a user belongs; to try it out, click on the **Edit** button. When done, close the **Edit Groups** dialog box in order to return to the main screen.

Now focusing on group management, this feature allows Spotfire Server administrators to bundle Spotfire permissions (also known as **licenses**) into new applicable sets (or groups).

To add a group, please select the **Groups** tab and click on **New Group**. Create a group named **testGroup**, at the top level. Click on **OK** when done, and then try to add our **testUser** to the newly created group. You can do so by using the **Add Users** button at the bottom of the **Members of Group** list. Do not forget to assign the user to a deployment area by clicking on the **Assign Deployment Area** button and selecting an available area from the dialog box shown. The topic of deployment areas will be explained later.

The result should be similar to what is shown in the following screenshot:
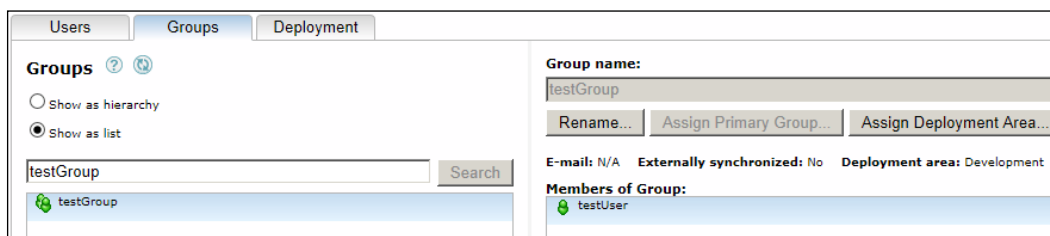
For more information on how to manage users and groups, refer to document `TIB_SSDoc_ServerInstallationManual.pdf`, or document `admin.chm` (which you can download from: `https://docs.tibco.com/products/tibco-spotfire-5-5-0`). By default, only the group **Administrator** has access permission (license) for Spotfire Professional and Spotfire Web Player Server. Users belonging to other groups will not be able to use both the tools.

The management of licenses (assignment to groups) can only be done in Spotfire Professional, in a special dialog box named **Administration Manager**. For more information about this topic, please refer to the *Help* documentation of Spotfire Professional.

Accessing the **Deployment** tab will allow us to manage **Deployment areas** and **Software packages**. After selecting this tab, you will be presented with the following screenshot:

You might have noticed we already have one deployment area (which was created by default during the installation process). This area is named **Production**.

The procedure for creating new areas is pretty straightforward—we only need to make sure the client packages are installed in the new area.

Since ours is a development environment, we will not do it; we will only rename the current **Production** area to **Development**. To rename the area, select it and click on **Rename**.

Regarding **Software Packages**, the presented list is related to the client packages we installed earlier; we do not have much to do here.

Opening logs and diagnostics

The direct link for log files and diagnostics option is `http://localhost:9090/spotfire/console/admin.jsp`. After accessing it, you will be presented with the following screenshot:
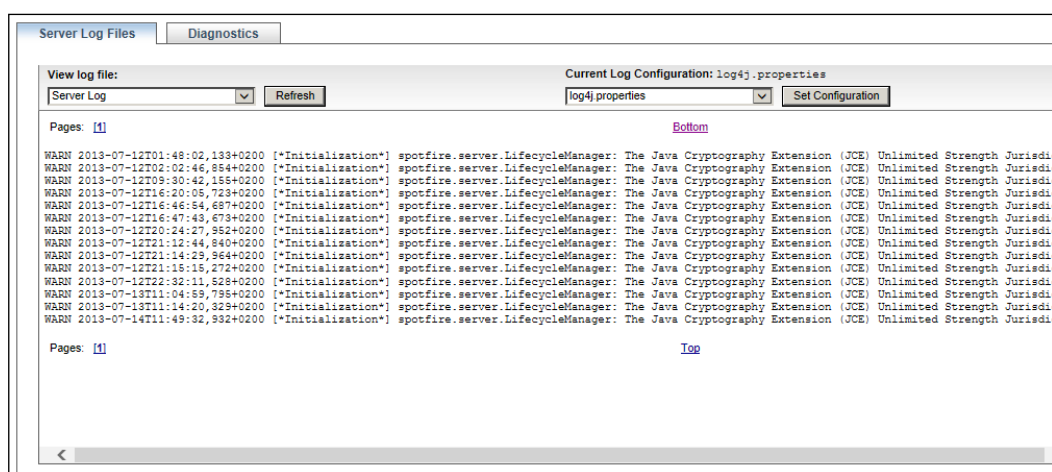
There are two tabs in the console: **Server Log Files** and **Diagnostics**.

The **Server Log Files** option allows you to access the Spotfire Server's logs through a web user interface. All logs presented in the list of possible logs to choose from for display are fetched from `C:\tibco\tss\5.5.0\tomcat\logs`, and can be accessed directly with any text editor.

Some examples of these selectable/viewable logs are: `catalina.log`, `server.log`, `startup.log`, and `tss550-stderr.*.log`.

The **Diagnostics** option allows you to access the server environment diagnostics, for instance database server, operating system, hardware and hardware configuration, and so on.

For more information on monitoring, refer to page 109 of *Chapter 6*, `Monitoring TIB_SSDoc_ServerInstallationManual.pdf`. It explains how to configure and use **JMX** and **TIBCO Hawk**. For more advanced logging options, you should also refer to page 113 of *Chapter 7*, `Log User Action TIB_SSDoc_ ServerInstallationManual.pdf`. It explains how to configure and enable the new **Log User Action** feature of Spotfire, which is basically a centralized log for all the platform's components log events, which can be stored in file, database, or both.

TIBCO Spotfire Web Player Server

Regarding the Web Player, we will now focus on the topics of configuration, diagnostics, logging, and monitoring.

Configuration

You may find the application's configuration file `Web.config` under the directory `C:\Program Files\TIBCO\Spotfire Web Player\5.5.0\webroot`. In this file you can set the authentication mode and more advanced features such as enabling JavaScript API, Spotfire administrator e-mail address, Spotfire Server URL, and so on. For a full list of settings, please refer to *Chapter 6.2*, `Advanced Web.Config settings TIB_SpotWebPlay_5.5.0_InstallationManual.pdf`.

Diagnostics

To access the Web Player Diagnostics console, open a browser window and navigate to `http://localhost/SpotfireWeb/Administration/Diagnostics.aspx`. After logging in (with an **Administrator** group user), the following tabs will become available:

- **Spotfire Server**: Displays information about Spotfire Server and the machine where it is installed (hardware, operating system, and other software)
- **Web Server**: Displays information about the IIS server, including operating system, hardware status, and so on
- **Web Application**: Displays information about the Spotfire Web Player configuration; these configuration items are fetched from the application's `Web.config` file
- **Loaded Assemblies**: Lists the `.dll` (assemblies) loaded by the web application
- **Site**: Displays statistics concerning the accesses to the Spotfire Web Player web application

- **Web Server Log**: Displays the Web Server Player log entries. The log file used as source can be found here: `C:\Program Files\TIBCO\Spotfire Web Player\5.5.0\Logfiles\Spotfire.Dxp.Web.log`

Logging and monitoring

To alter the logging level of `Spotfire.Dxp.Web.log`, you can edit the XML tag `level` (`<level value="INFO"/>`) of file `C:\Program Files\TIBCO\Spotfire Web Player\5.5.0\webroot\app_data\log4net.config`. This file contains the configuration of the `log4net` of the web application; the following levels are available: `DEBUG`, `INFO`, `WARN`, `ERROR`, and `FATAL`.

For more information on `log4net` and the logging levels, visit `http://logging.apache.org/log4net/` and refer to page 101 of *Chapter 6.12 Logging and Monitoring*, `TIB_SpotWebPlay_5.5.0_InstallationManual.pdf`.

Also, be aware that it is possible to enable performance trackers for the Web Player Server. For Performance monitoring using the management console of Windows, refer to page 111 of *Chapter 6.1.3*, *Performance*, `TIB_SpotWebPlay_5.5.0_InstallationManual.pdf`.

Summary

In this chapter, we learned how to create a Spotfire development environment, by installing the following TIBCO Spotfire platform products: TIBCO Spotfire Server, TIBCO Spotfire Professional, and TIBCO Spotfire Web Player Server.

All the components were installed in a single machine, where a database instance (for the server) was previously installed and configured, enabling the creation of an environment without any external dependencies.

After the installation procedure, we focused on the subjects of environment configuration, administration, and management.

In the next chapter, we will begin using our environment to explore the creation and manipulation of rich data visualizations using Spotfire Professional.

3
Creating Visualizations

This chapter provides a how-to for the creation of Spotfire Professional Visualizations. By the end of it, the reader should be an independent user of the Spotfire platform, having the required skills to create meaningful Analytics visualizations. However, before actually diving deep into the usage of Spotfire Professional, I would like to emphasize the purpose of Analytics tools.

The discipline of Visual Analytics allows us to understand the what, the why, and the how of raw data. It does so by fostering intuitive data representations, where the multiple dimensions of data are reduced and simplified.

The power and success of Analytics tools is closely related to their capability to summarize big chunks of data, providing interactive visualizations, where filters, rules, and data drill-down can be applied, always in a very attractive and captivating environment.

Users from many disciplines (IT, Scientists, and Statisticians) are allowed to find the correct answers for their most challenging questions, and complex data becomes understandable through analysis, integrated with computational analytics and statistical tools. Reports can be quickly created and distributed, being available on the Web, ready for access from several device types, including mobile.

To demonstrate the power of Visual Analytics, I will now provide an example where a set of data becomes dissectible through the creation of visualizations. This will reinforce the idea of the benefits of Visual Analytics.

The next screenshot is of a Microsoft Excel file containing sales information. The records contain data, such as customer age, customer gender, city, number of purchases, total spent in purchases, and type of goods.

Customer age	Age group	Gender	Business location	Number of purchases	Receipt average	Total	Electronics	Furniture	Garden	Groceries	Toys
68	60	Female	New York	60	700	42082	4860	5865	5429	5394	18918
41	40	Male	Boston	9	577	5196	885	1362	1431	1419	0
56	50	Female	New York	12	483	5799	216	524	467	4286	304
77	70	Female	New York	5	185	926	0	238	0	684	0
61	60	Female	Los Angeles	42	926	38866	10013	3488	6999	5165	11266
45	40	Female	Seattle	59	405	23964	6396	5774	7156	4449	185
62	60	Male	Los Angeles	1	153	153	153	0	0	0	0
44	40	Female	Los Angeles	22	398	8748	1139	825	2373	3532	0
52	50	Female	Boston	20	200	4038	1033	584	1582	649	185
18	10	Female	Seattle	14	414	5799	0	417	0	5061	0
74	70	Female	Los Angeles	3	342	1026	436	0	467	122	0
55	50	Female	Boston	20	552	11043	5586	1369	731	1369	354
75	70	Female	New York	20	213	4260	474	379	1626	1478	298
44	40	Female	Los Angeles	88	135	11768	492	91	464	1431	0
66	60	Female	New York	33	514	16987	2593	4377	5535	2223	2216
56	50	Female	Boston	41	257	10610	3846	219	4154	1164	662
42	40	Male	Los Angeles	6	248	1485	555	282	122	521	0
58	50	Female	Boston	5	433	2163	797	609	0	0	756

To the human eye, this raw data is readable, but further analysis would be complex and lengthy. However, importing it to Spotfire Professional allows me to generate in a very short amount of time a set of three visualizations (displayed in the next three screenshots), where distinct trends can easily be identified. These are as follows:

- *Which city has the biggest spender? Is this a man or woman?*
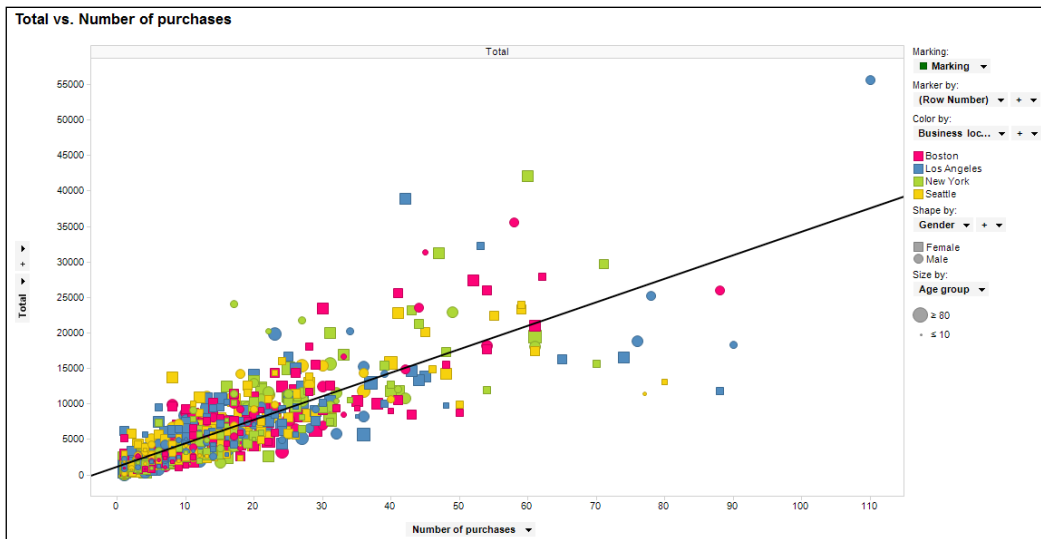
We can clearly see the biggest spender was a man from Los Angeles—top-right corner of the screenshot. He was not only the biggest spender, but he also did the most purchases (this is clearly the case of an outlier).

- *Who buys more often – men or women?*
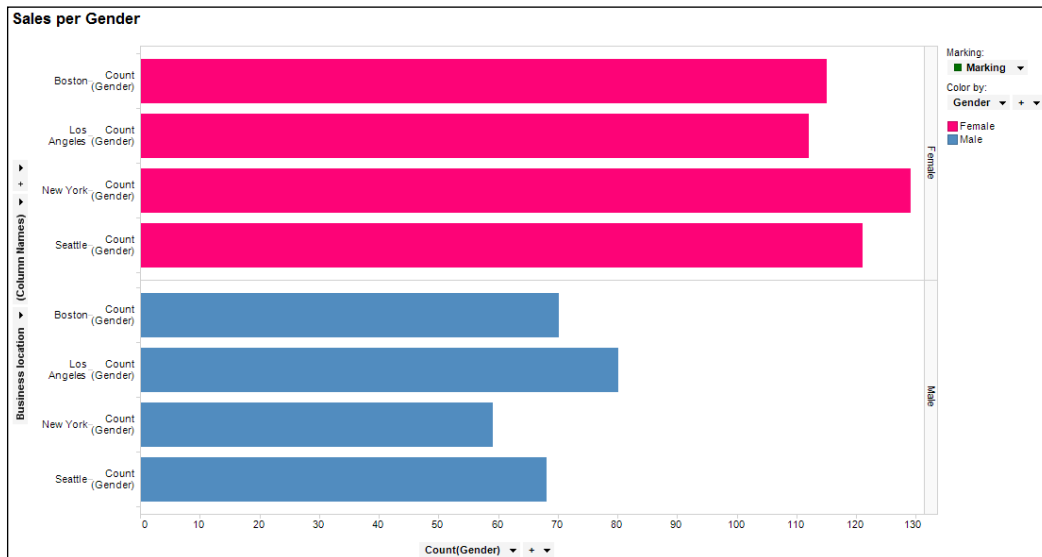
Segregating by gender, we realize we have more records from women—women bought more often. This conclusion is true for the four analyzed locations.

Please also note that although women bought more often, this does not necessarily mean they spent more as well.

- *Which age group spends the most?*

Sum(Total) per Age group
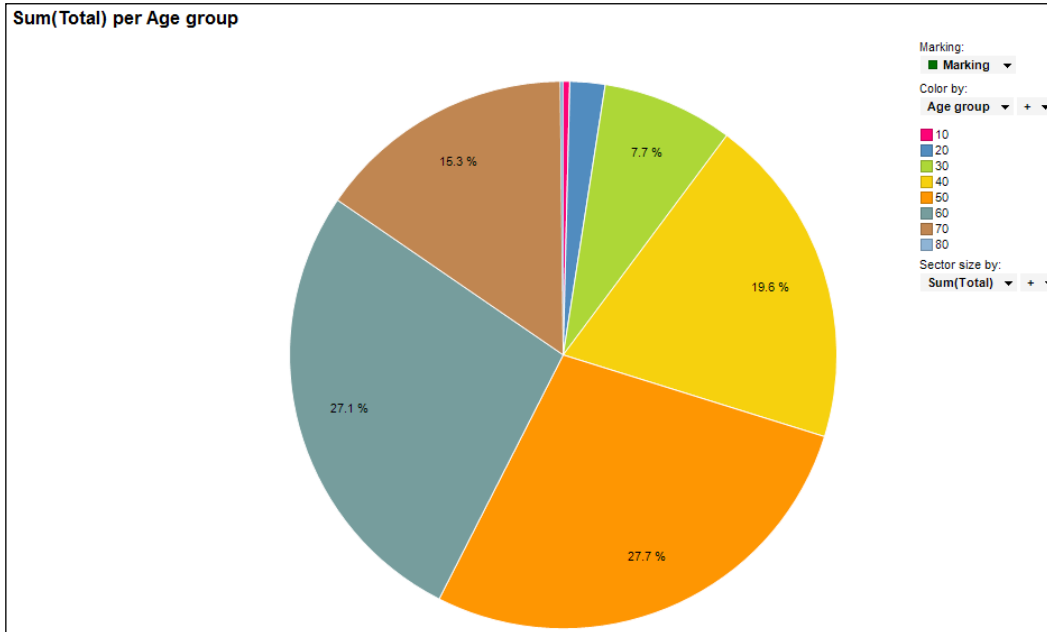
> This pie chart shows the distribution of expenditure between the different age groups. The groups that spend the most are 50s and 60s. Respectively, these groups spent 27.7% and 27.1% of the total spent.

These three visualizations will be recreated in the *Visualization types* section so the reader can realize how simple their creation was.

I hope that the examples shown here have increased your appetite for the tool, as we will now start with the technical details behind the creation of visualizations. In the next section, we will focus on data sources—which ones can we access and how do we do it?

Data

The first step to start working with Spotfire Professional consists of the loading of data into the application. This data will be fetched from an external data source, and internally (in Spotfire), it will be defined and known as a data table.

Spotfire can create data tables by accessing several types of external data sources:

- **Clipboard**: Performed through copy and paste operations.

- **Text files**: These could be plain text (`.txt`), comma separated (`.csv`), or fixed format.

- **MS Excel files**: The supported versions are MS Excel 2000 or later, with extensions `xslx` and `xls`.

- **SAS files**: File extensions for SAS files are `sas7bdat` and `sd2`.

- **Database**: Database tables or views, represented by an SQL query (in SQL-92). Columns can be filtered, and tables joined, allowing for a prehandling of data before the loading. The following vendor databases are supported:

  ○ Microsoft SQL Server
  ○ Microsoft SQL Server Analysis Services
  ○ Oracle
  ○ Teradata
  ○ ADS Composite Information Server
  ○ IBM Netezza
  ○ Oracle MySQL
  ○ PostgreSQL
  ○ SAP® Business Information Warehouse (SAP BW)

- **Information link**: This consists of a database query, specifying tables, columns, and filters, created in the tool **Information Designer** (from the **Tools** menu of Spotfire Professional). The queries are created visually, as opposed to the programmatically in SQL.

- **Custom file type**: An example of this is XML. Spotfire can be configured to open custom file types.

> Please note that to connect to an Oracle Database, TIBCO recommends users should have installed the **Oracle Data Provider for .NET** (**ODP.NET**). To connect to an SQL database, users should have installed the **SqlClient Data Provider**.
>
> Also, if using a Windows machine, users can create an ODBC data sources in the Windows ODBC Data Source Administrator.

It is also important to differentiate between the the two modes in which Spotfire Professional internally handles the loaded data:

- **In-Memory data**: In this mode, the tool will load all the data into the working memory, allowing for quicker (and independent) access.

- **In-Database data**: In this mode, Spotfire Professional does not load all the data into the working memory, but takes advantage of the system that owns this data for processing power (all queries will be run in the source system). This mode should be used in scenarios where there are extremely large data sets, too big to fit into the application's working memory (with this mode, data loading into Spotfire becomes virtually unlimited).

> In-Database data mode has a few inherent limitations. For more details, please refer to the page **Data** | **Data in Spotfire** | **Working With In-Database Data** of Spotfire Professional's **Help**.

Data loading

In this section, examples of data loading will be provided. The following data sources will be used: text file, MS Excel file, and database table.

Loading a text file

There are two different ways to load the data of a text file into Spotfire. The first consists of using the menu option **File** | **Open** or the **Open** button in the toolbar. Alternatively, the menu option **File** | **Add Data Tables** or the **Add Data Tables** button in the toolbar can be used.

To exemplify, we will load the file `C:\Program Files (x86)\TIBCO\Spotfire\5.5.0\Example Data\StoreSales\StoreSales.txt`. We will start by clicking on the **Add Data Tables** toolbar button. A dialog box named **Add Data Tables** will be opened; this dialog box is presented in the following screenshot:

Add Data Tables

Data tables:

Name	Source	Transformations

Add ▼

Remove

Settings

Name:

⊙ Show transformations (no transformation steps added)

Help Manage Relations... OK Cancel
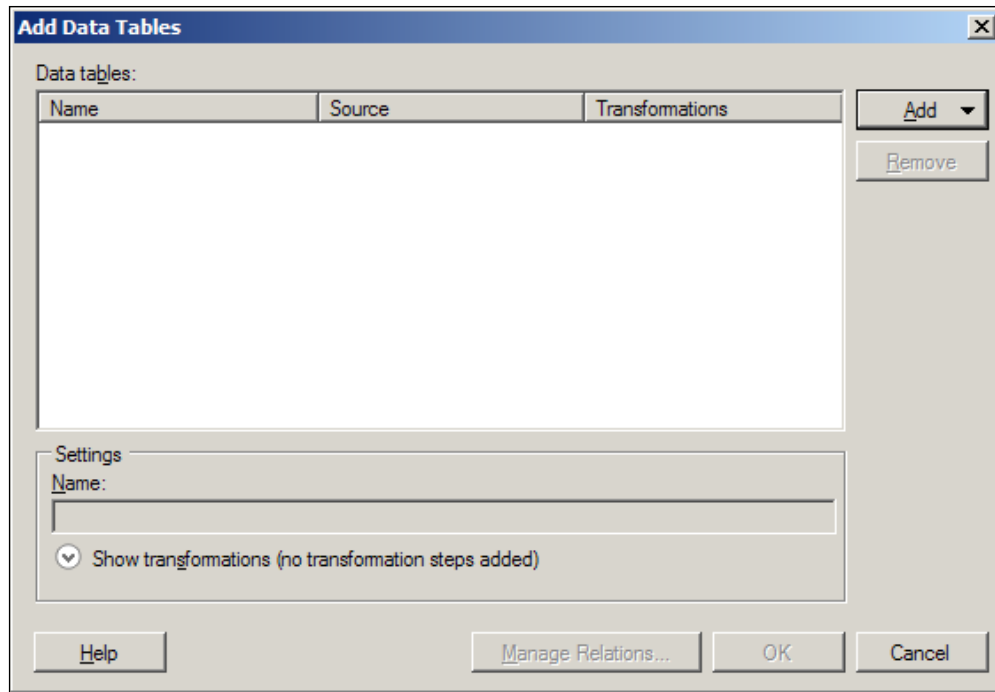
> There are several examples of data files and related visualizations
> in the directory `C:\Program Files (x86)\TIBCO\`
> `Spotfire\5.5.0\Example Data`. These come bundled with the
> product Spotfire Professional.

Navigate to **Add | Files**. In the file selection window, navigate to `C:\Program`
`Files (x86)\TIBCO\Spotfire\5.5.0\Example Data\StoreSales\StoreSales.`
`txt` and click on **Open**.

The **Import Settings** dialog box will be presented as shown in the
following screenshot:
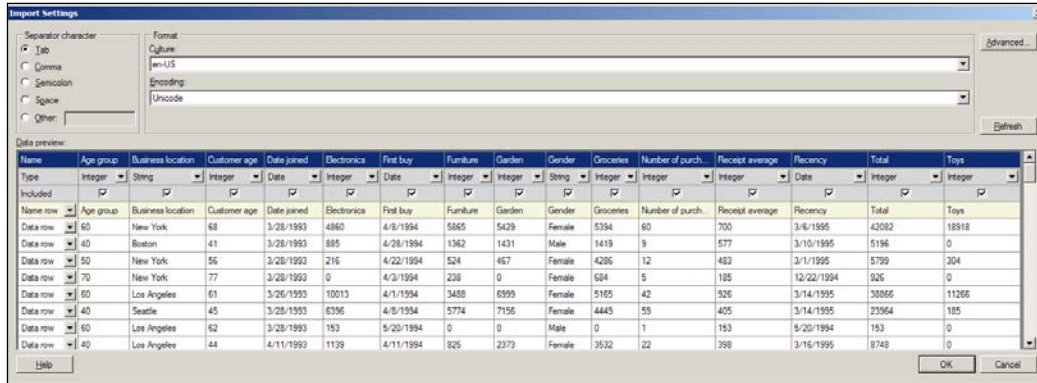
All the contents of the file will be displayed in the section named **Data preview:**,
and several options will be available to the user:

- Defining the separator character (manageable in the **Separator character**
  section). This is a generic customization.

- Defining the file encoding (manageable in the **Format** section). This is a
  generic customization.

- Defining the data types of the columns. The default data type of a column
  can be redefined by the user. To apply any change, the user must click
  on the button **Refresh**; if there are any issues, such as the impossibility of
  application of a new data type, an error will be displayed in the problematic
  records. This is a column related customization. The following screenshot
  lists the available types to choose from and associate with a column:
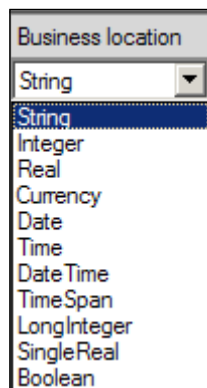
- Deciding on the inclusion of a column (manageable in the checkbox on top of each column). This is a column related customization.

- Choosing the definition of the row content. It is manageable in the first column of the **Data preview:** section. This is a row related customization. The following screenshot lists the available types:
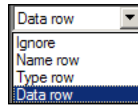
Their meaning is the following:

- ° **Ignore**: This means do not include the row in the data table.
- ° **Name row**: This indicates that the row content is part of the column name.
- ° **Type row**: This indicates that the row content defines the type of the column.
- ° **Data row**: This indicates that the row content is actual loadable data.

By clicking on **Advanced...** (on the top-right corner of the **Import Setting** dialog box), the user can access the advanced import settings, as for instance, the definition of *null* values and the first row to read from.

Please leave all the default options and click on **OK**. A new dialog box will appear, allowing for the setting of a name for this new data table. This is shown in the following screenshot:
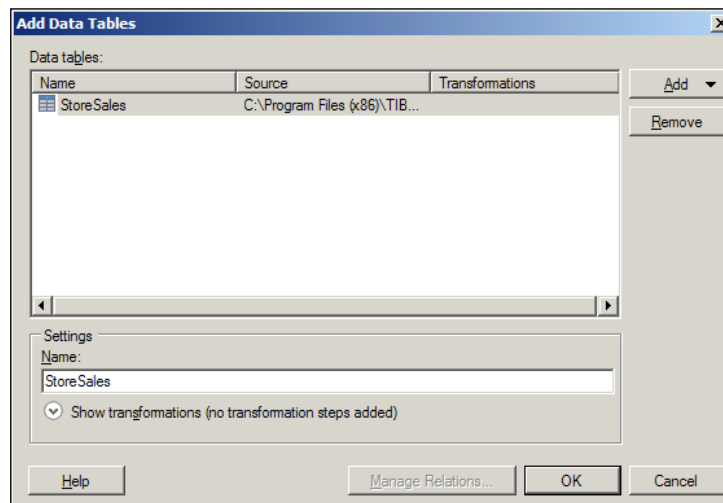
Please leave the default name `StoreSales` and click on **OK**. Immediately, a Scatter Plot visualization will be generated; this graph can be altered or replaced with a different chart type.

The following screenshot presents this initial default graph:
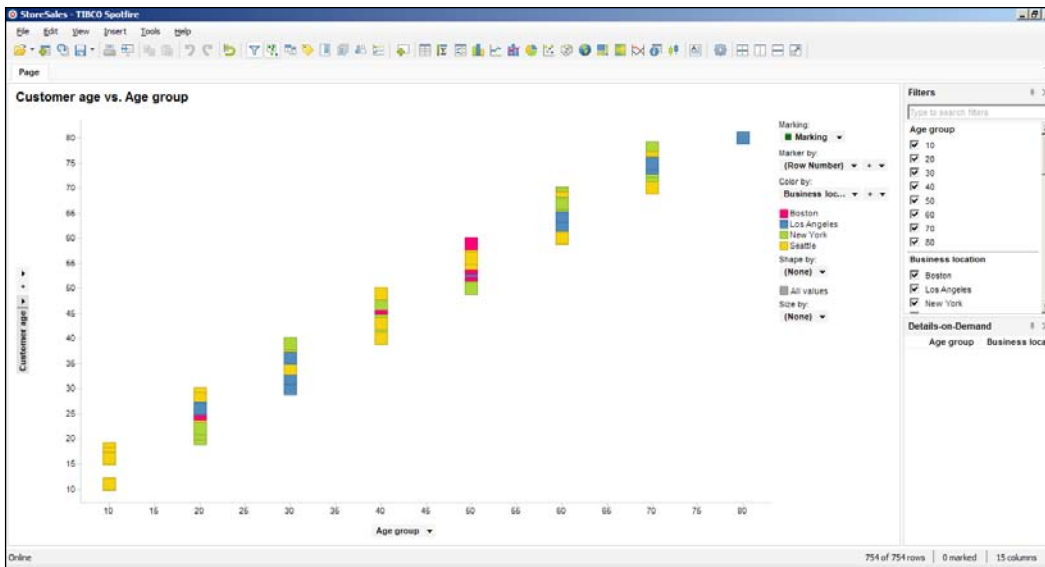
Visualizations generated from the loaded data are created in a new page with a default name (each page is a tab below the toolbar). To rename a page, right-click on top of its name.

In this example, the new page has name **Page**; rename it to `StoreSales`.

Please save the visualization and the data table, as we will continue building it on later. You can save it by navigating to the menu option **File** | **Save** or the **Save** button in the toolbar.

> TIBCO Spotfire Professional generates by default a Scatter Plot for each added data table.

Loading an Excel file

Start by preparing an Excel worksheet similar to the one presented in the following screenshot. Name it `Performance.xslx`.

	A	B	C	D
	Race	Runner	Gender	Time
	1	Marc	Male	11,3
	1	John	Male	12,1
	1	Mary	Female	13,4
	1	Wanda	Female	13,7
	2	Marc	Male	12,1
	2	John	Male	12,4
	2	Mary	Female	13,9
	2	Wanda	Female	13,8
	3	Marc	Male	11,6
	3	John	Male	11,7
	3	Mary	Female	13,9
	3	Wanda	Female	13
	4	Marc	Male	12,1
	4	John	Male	11,8
	4	Mary	Female	13,9
	4	Wanda	Female	13,7

Loading the data from an Excel sheet works in a similar way to loading a plain text file (using the option **Add Data Tables**). Taking this into consideration, please load the data.

You may notice the pop-up is now named **Excel Import**, and one of the import options available is that the user can now select the worksheet to import. Please see the following screenshot for example:
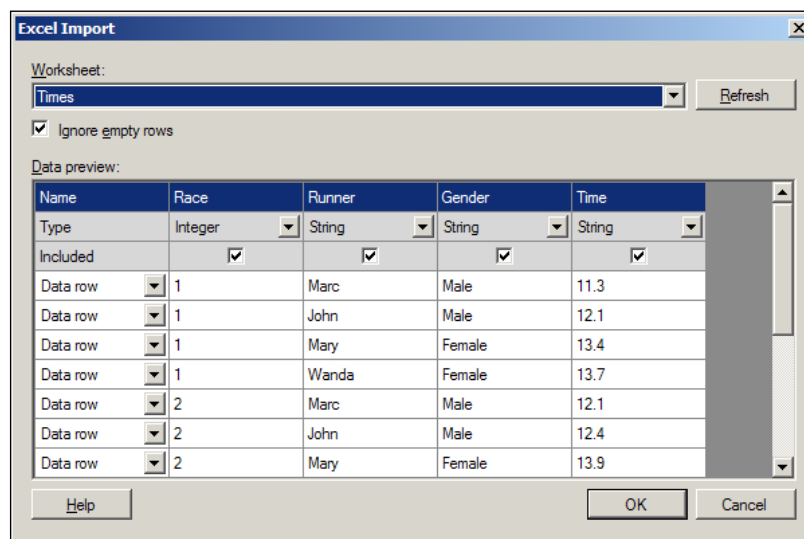
Click on **OK**, and next, accept the default name for the data table by clicking on **OK** again. Once again, a default Scatter Plot visualization will be automatically generated for us. Please rename the created page to `Performance`.

Save the visualization and the data table, as we will continue building it on later.

Loading from databases

To be able to load data from a database, you will need to have the right connectors. In this example, we will connect to our local Oracle XE instance, so we will need to have the Oracle Client ODP.NET installed. The 64-bit version installer can be downloaded from `http://download.oracle.com/otn/other/ole-oo4o/ODAC1120320Xcopy_x64.zip`.

For this example, we will need to have some loadable data in our local XE database. For simplicity, we will use the *Human Resources (HR)* example, which comes bundled with the Oracle XE installation (the demo setup scripts can be found at `C:\oraclexe\app\oracle\product\11.2.0\server\demo\schema\human_resources`).

Although there are many scripts at the specified location, users are solely required to run `hr_main.sql`. Please make sure that the database objects are created in a schema other than the one created for the Spotfire Server: `spotfire` schema.

The following screenshot presents the input values that should be used in the dialog generated while running the script `hr_main.sql`.

If you used this configuration, a new database user named `hr` will be created (with the password `spotfire`). All the demo objects will be created in this user's schema.

To trigger the loading of the data from the database, please click on the **Add Data Tables** toolbar button. You will be presented with the dialog box **Add Data Tables**. Navigate to **Add | Database**. The dialog box **Open Database**, as shown in the following screenshot, will be presented.
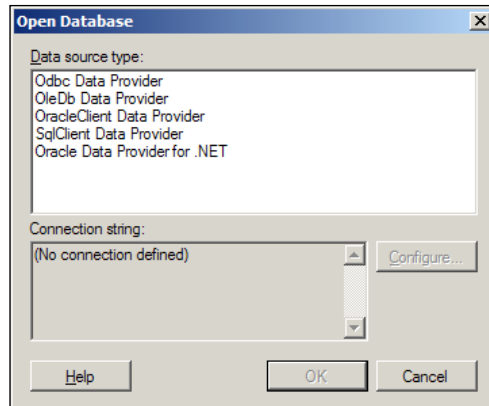
Three drivers will be available for accessing Oracle:

- **ODBC Data Provider**: Windows configured ODBC System or User data sources.
- **OracleClient Data Provider**: Oracle client driver.
- **Oracle Data Provider for .NET**: ODP .NET driver, which is faster and more suitable to use with Spotfire.

In certain Windows versions, the ODP for .NET driver will be listed as **ODP .NET, Managed Driver**. Please see the following screenshot:
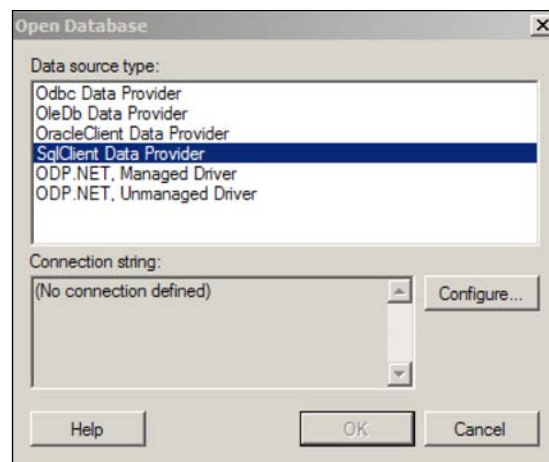
Choose **Oracle Data Provider** for .NET, and click on **Configure....** A pop-up, named **Configure Data Source Connection**, will be shown. This is presented in the following screenshot:
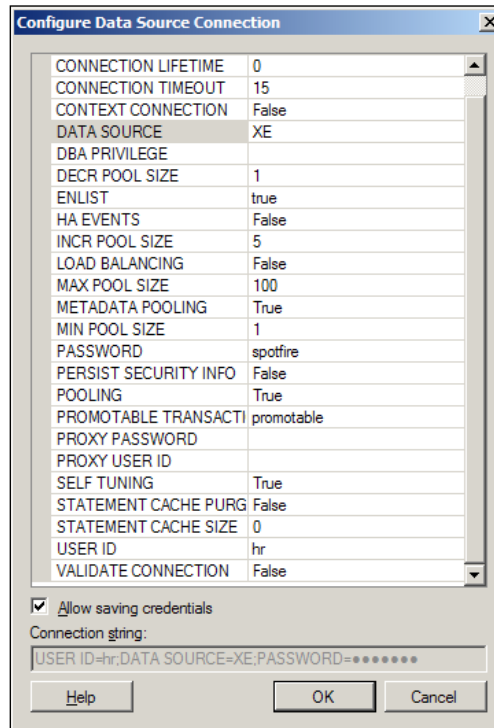
Fill in the following parameters:

- `DATA SOURCE: XE`

- `USER ID: hr`

- `PASSWORD: spotfire`

Tick **Allow saving credentials** and click on **OK**.

The dialog box will close and focus will return to the **Open Database** dialog. You can push **OK**.

A dialog box named **Specify Tables and Columns** will be presented; see the following screenshot. This will allow us to choose which database tables or views we want to load into Spotfire.
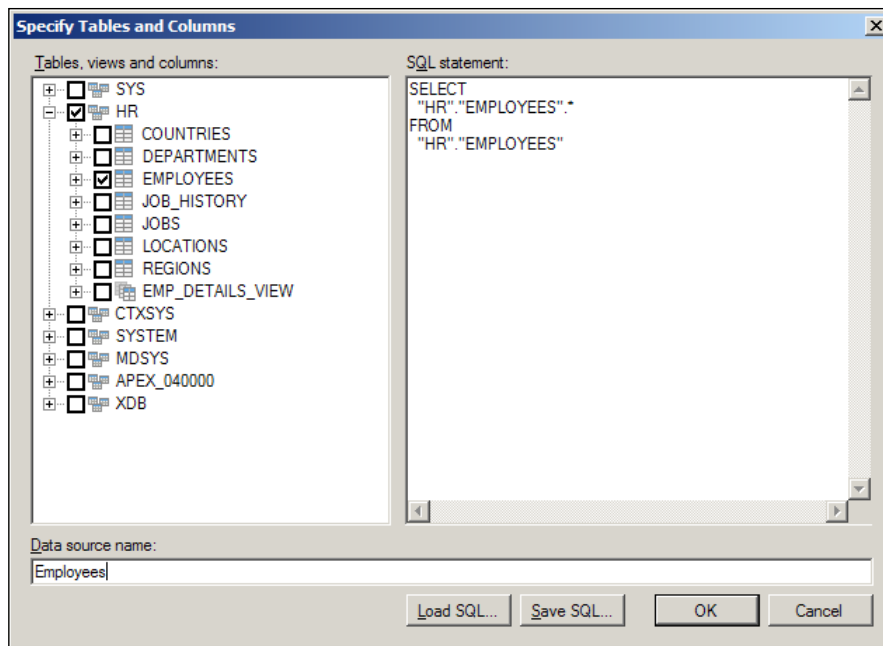
Drill down **HR** and select **EMPLOYEES** (we will load all the columns from the table). Name the data source as `Employees`, and click on **OK**.

Once again, a default Scatter Plot visualization will be automatically generated. Please rename the new page to `Employees`.

Please save the visualization, so we can build it on later.

> Saving an analysis project does not save the underlying data (this data is not embedded in the project). Users have, however, the option to embed it explicitly while saving.
>
> The impacts of non-embedding (or linking) data are related to the fact that, for visualizations to be available, the underlying data sources also have to be available (this also applies to the system files).

So far we loaded three distinct sources of data into Spotfire Professional, and we made no changes to the default visualizations generated. In the next section, we will start working with several visualization types and we will leverage these default Scatter Plots that we have.

Visualization types

We will now look at the topic of visualizations creation. We will start with the type Table, and we will then look at other types, such as Bar Chart, Pie Chart, Line Chart, and Scatter Plot. All the presented theory will be accompanied by illustrations and instructions for the reader to achieve the same end result.

Tables

As stated previously, for the three data tables we created, the default Scatter Plot visualizations were created. This is a default characteristic of Spotfire Professional; however, a Table visualization should be the best place to start looking at the data.

Tables present the records of the underlying data tables, where each table row corresponds to a data table record.

In order to study this type of visualization, our first task will be to create a Table. Since the type of all created visualizations can be changed at any point in time, we can leverage one of the Scatter Plots created earlier; we will use Scatter Plot generated when we created the `StoreSales` data table.

Please access the analysis project and select the page `StoreSales`. Then right-click on the Scatter plot, select **Switch Visualization To** and select **Table**. A table, as displayed in the following screenshot, will be presented:

StoreSales							
Age group	Business locat...	Customer age	Date joined	Electronics	First buy	Furniture	Garden
60	New York	68	3/28/1993	4860	4/8/1994	5865	5429
40	Boston	41	3/28/1993	885	4/28/1994	1362	1431
50	New York	56	3/28/1993	216	4/22/1994	524	467
70	New York	77	3/28/1993	0	4/3/1994	238	0
60	Los Angeles	61	3/26/1993	10013	4/1/1994	3488	6999
40	Seattle	45	3/28/1993	6396	4/8/1994	5774	7156
60	Los Angeles	62	3/28/1993	153	5/20/1994	0	0
40	Los Angeles	44	4/11/1993	1139	4/11/1994	825	2373
50	Boston	52	4/11/1993	1033	4/25/1994	584	1582
10	Seattle	18	4/8/1993	0	4/6/1994	417	0
70	Los Angeles	74	3/28/1993	436	5/18/1994	0	467
50	Boston	55	3/28/1993	5586	4/7/1994	1369	731
70	New York	75	4/5/1993	474	5/17/1994	379	1626
40	Los Angeles	44	4/4/1993	492	5/2/1994	91	464
60	New York	66	4/5/1993	2593	4/9/1994	4377	5535
50	Boston	56	4/10/1993	3846	4/21/1994	219	4154
40	Los Angeles	42	4/10/1993	555	5/24/1994	282	122
50	Boston	58	4/8/1993	797	4/13/1994	609	0
60	Seattle	66	3/31/1993	4688	4/13/1994	4452	1186
50	Seattle	59	3/31/1993	527	4/30/1994	0	0

As you can confirm, the records shown correspond to the lines of the file that generated the `StoreSales` data table (it contains customer sales records).

To verify and manipulate the visualization's properties, right-click (with the mouse) on top of the table and select **Properties**. This will bring up a dialog box with the title **Data Table name Properties**. The dialog box is shown in the following screenshot:
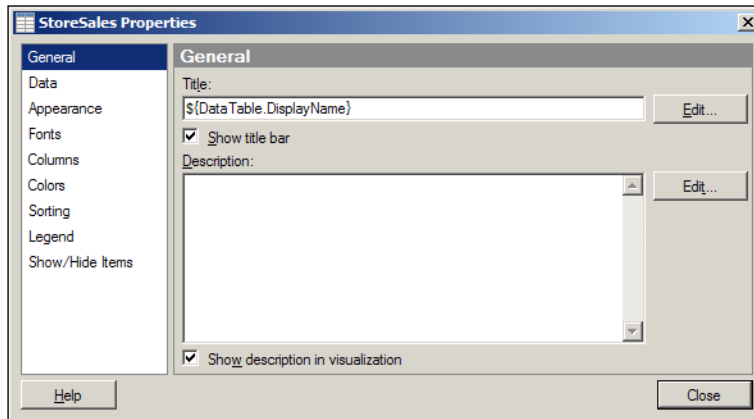
For tables, several customizations can be achieved in this dialog box, including:

- **Data**: The data table associated with the visualization can be swapped.
- **Columns**: The columns selected to be displayed as part of the table can be changed, allowing for the hiding of data. In our example, we can remove the columns `Electronics`, `Furniture`, `Garden`, `Groceries`, and `Toys`. The result will be the following (see following screenshot):
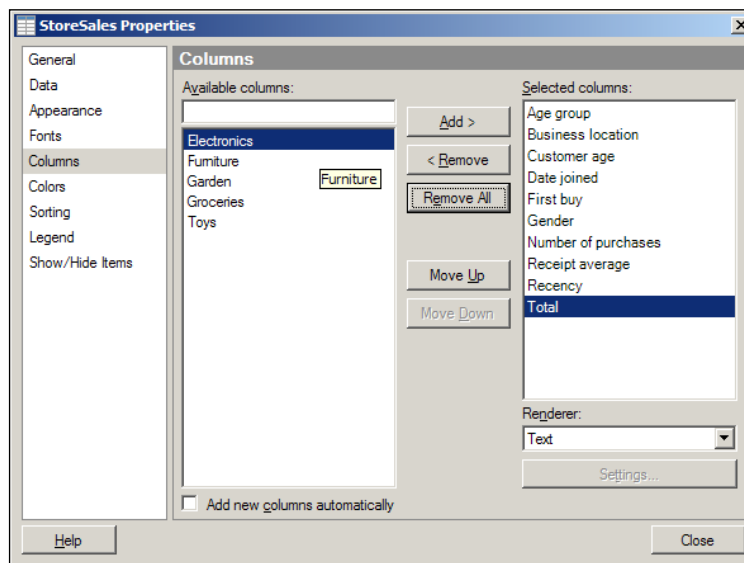
The order of display of the columns and the rendering can also be changed.

> The **Renderer**: option allows for the loading of images as a replacement for the data. The options available to the user are: **Link**, **Text**, **Image from URL**.

- **Colors**: Color schemes can be defined for a table. The schemes can be defined with multiple color modes: **Gradient**, **Segments**, **Fixed**, and **Unique values**. The following two screenshots display the definition of the **Unique values** scheme for the **Age group** and the **Gradient** scheme for **Number of purchases**.
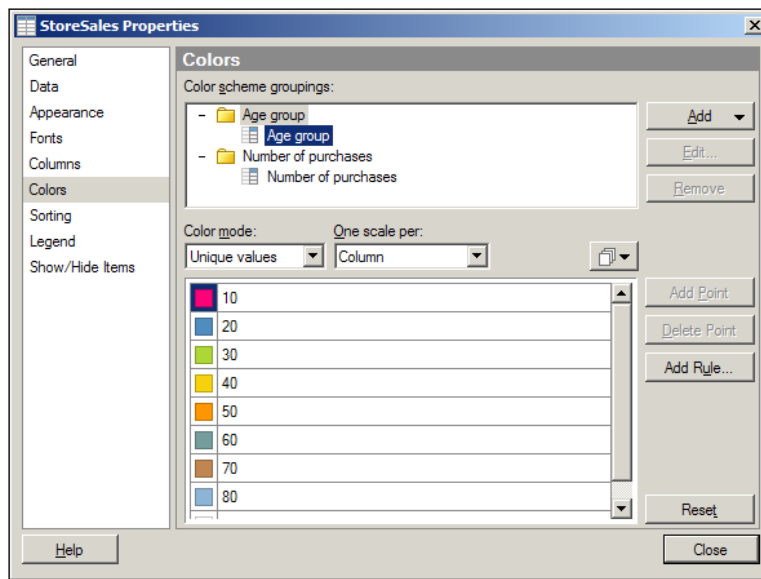
> The coloring modes available for a certain column will depend on the data type of the column. Integers have selectable **Gradient**, **Segments**, **Fixed**, and **Unique values**, while strings have selectable **Fixed** and **Unique values**.
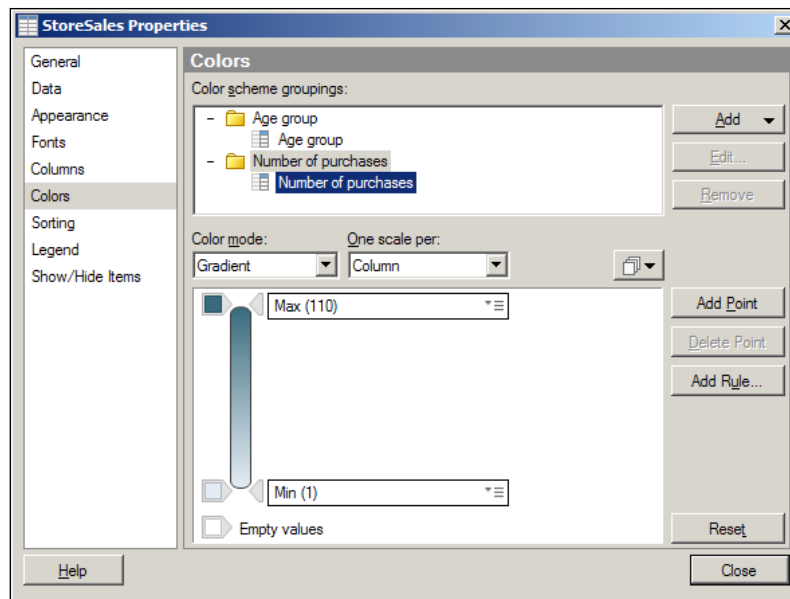
For the **Gradient** and **Segments** color modes, points can be defined to better emphasize a coloring scale. The following screenshot shows the definition of a middle point, and redefined coloring from red (low number of purchases), to yellow (medium number of purchases), and to green (high number of purchases).
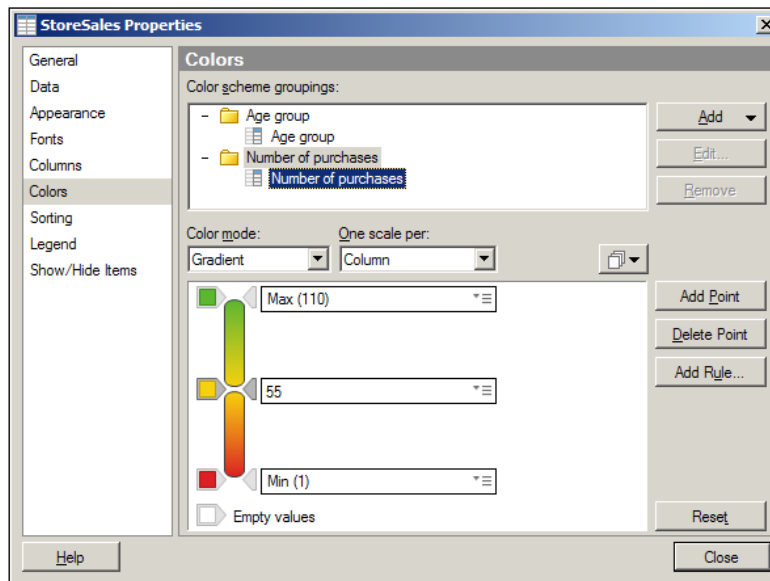
To achieve the same result, please click on the button **Add Point**. Colors can be changed by clicking on the colored squares.

For all the color modes, coloring rules can be created. By clicking on **Add Rule** in the dialog, we can create a rule to mark all the values greater than 105 as blue; this is shown in the following screenshot:
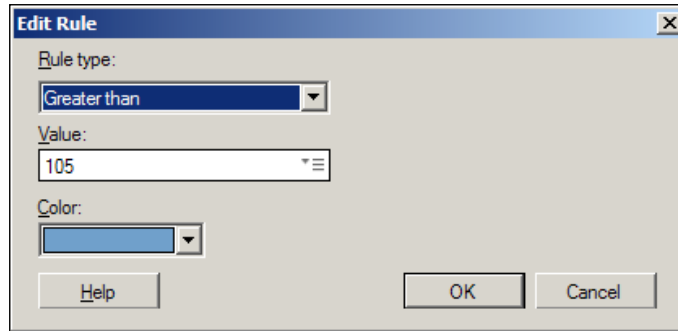
- **Sorting**: Here we can select the column to sort by a parameter. In our example, we can sort the table by age group. Select **Age group** in the **Sort by:** drop-down list and leave the default selection **Ascending**. The following screenshot is the example of this:
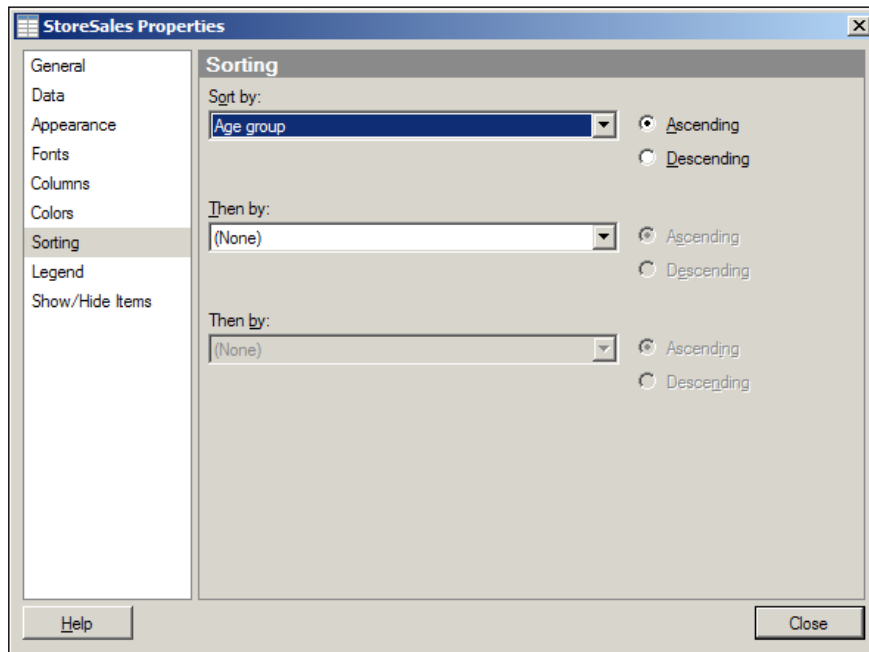
- **Show/Hide items**: Custom filters can also be applied to the existing records, showing or hiding data accordingly. In our example, we can create a rule that shows all the records where the customer's age is greater than or equal to 16. Please see the following screenshot:

The end result table (after all the manipulated properties) is presented in the following screenshot:

StoreSales

Age group ▲	Business locat...	Customer age	Date joined	First buy	Gender	Number of pur...
10	Seattle	18	4/8/1993	4/6/1994	Female	14
10	Los Angeles	16	6/29/1993	4/20/1994	Female	8
10	Seattle	17	4/21/1993	5/4/1994	Female	2
10	Boston	16	12/7/1994	6/26/1994	Female	4
10	Boston	16	7/10/1994	7/10/1994	Female	1
10	Seattle	16	12/12/1994	12/18/1994	Female	1
20	Los Angeles	27	4/12/1993	4/5/1994	Female	20
20	Seattle	29	5/23/1993	4/13/1994	Male	12
20	Los Angeles	28	6/7/1993	5/4/1994	Female	35
20	New York	28	6/1/1993	4/6/1994	Female	10
20	New York	20	9/9/1993	4/3/1994	Female	16
20	Seattle	26	6/8/1993	4/18/1994	Male	19
20	Boston	26	9/22/1993	1/19/1995	Male	3
20	New York	25	11/21/1993	4/5/1994	Female	3
20	Boston	28	12/22/1993	5/12/1994	Male	6
20	Los Angeles	27	2/22/1994	4/11/1994	Female	9
20	Seattle	27	8/15/1994	8/15/1994	Female	4
20	Boston	22	11/30/1994	11/30/1994	Female	7
20	New York	21	1/7/1995	1/22/1995	Male	5

Feel free to navigate through the table and analyze the presented data, especially when it comes to the shown data, the sorting, the coloring, and the filtering in/out.

Please save the analysis project.

Bar Charts

The Bar Chart is one of the most common forms of visualization, where values are represented by the height of a bar, which can symbolize many distinct summarized values, such as averages and counts.

Bar Charts also offer the possibility of having side-by-side bars (where several series are examined side-by-side in multiple bars) and stacked bars (where several series are arranged vertically in a single bar).

To create a Bar Chart, we can switch the type of the visualization existing in the **Performance** page to **Bar Chart**; this page contains a Scatter Plot analysis generated when we created the `Performance` data table. The procedure is similar to what was done for the Table. A defaulted Bar Chart graph is created, where the displayed information is not meaningful, but which will be the base for our work.

Please change the x axis to variable **Race** and y axis to variable **Avg(Time)** (with the function `Average`). Finally, change **Color by:** to **Gender**.

These customizations will give us a stacked bar chart (having the stack categories as male and female gender (from **Color by:**)), where the average times per race are displayed . This is shown in the following screenshot, where the red arrows indicate the place to switch the variables.
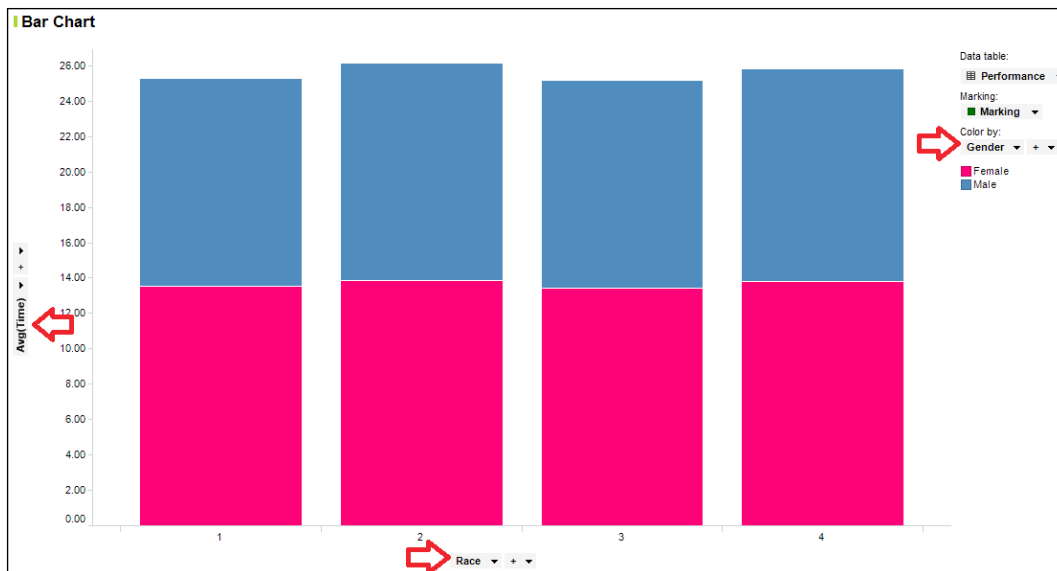
> To view the details of any chart element, whether it is a bar, a dot, or a line, you only have to place your mouse cursor over the element.

Often, stacked columns are not the best way to compare categories as the series of the stack have different starting points. Visually, the series sizes become increasingly difficult to compare.

The solution for these cases can be the use of side-by-side bar charts, which is what we will now demonstrate. Start by right-clicking with the mouse cursor on top of the new visualization and choose the option **Duplicate Visualization**. You will end up with a horizontally split screen, containing two identical charts. We will leave the one on top as it is, and we will focus on the one beneath.

Right-click (with the mouse) on top of the chart and select **Side-by-Side Bars**. At this point, your visualization should be similar to the one presented in the following screenshot:
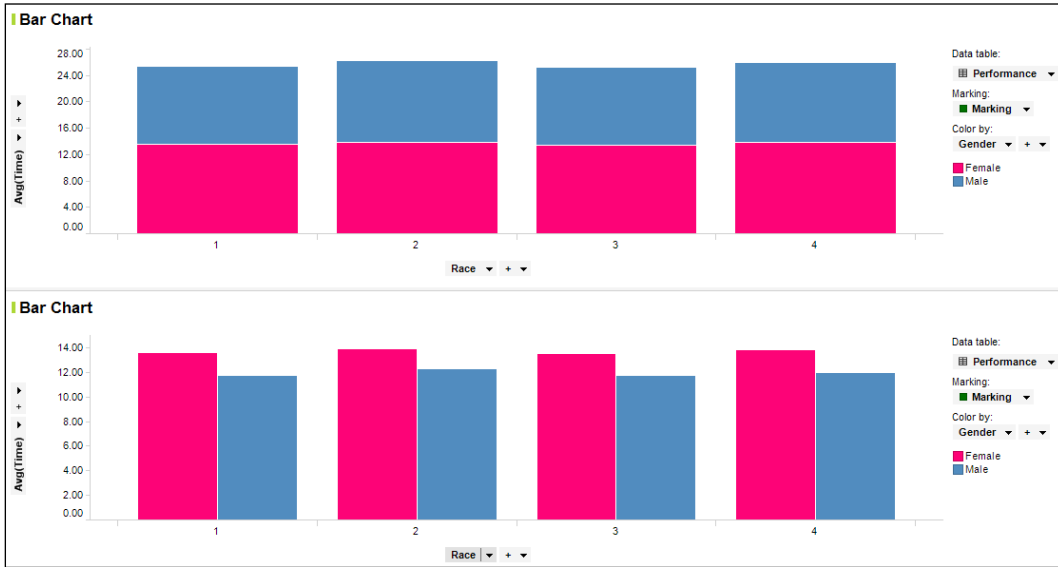
From the previous task, you probably realized Spotfire bar charts offer several display possibilities, such as **Vertical Bars** or **Horizontal Bars**, and **Side-by-Side Bars**, **Stacked Bars** or **100% Stacked Bars**. To switch between any of these, the procedure will always be the same—right-click on top of a bar chart.

The following screenshot displays these available options:
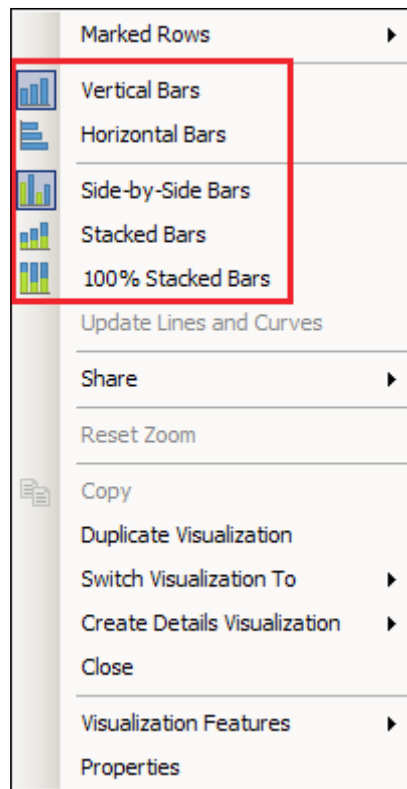
Regarding the available properties of Bar Charts, besides the graph appearance, formatting, fonts, colors, and show/hide, one can also manipulate the axis (for instance, by incorporating sliders and gridlines) and trellis.

A trellis fosters the breaking of categories into separate mini charts. This allows for a better separation of data when too much of it is available. This type of display represents a big improvement, especially when compared to stacked bar charts, where, as explained previously, if many segments (or series) are available, visual comparison becomes challenging.

To add a gender trellis to our Stacked Bars graph, access the properties of the page's top graph and select the option **Trellis**. Then, select **Gender** as a columns trellis. Please see the following screenshot:
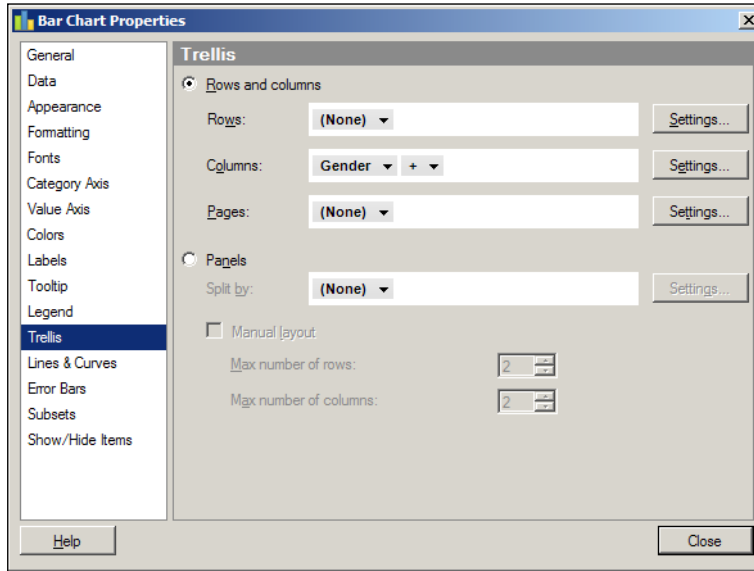
The end result should be similar to the one displayed in the following screenshot (there are now separate graphs for female and male).
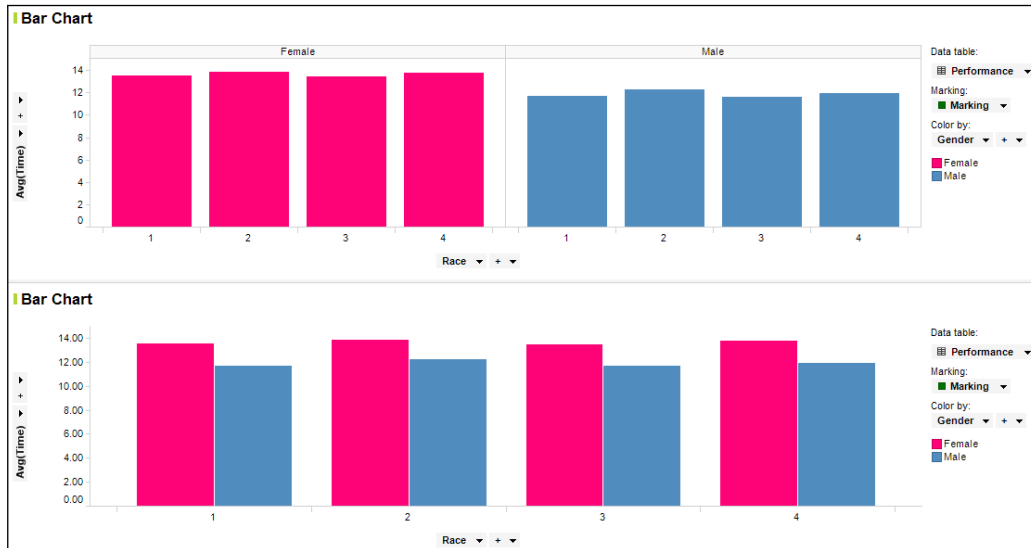
Other types of trellis can also be selected, such as **Rows:**, **Pages:**, and **Panels**.

So far, we have worked solely with one variable per axis, but there is no such restriction in Spotfire Professional. The addition of variables to both the axes helps to create analysis categories, but it will have an impact in the number of displayed columns. This means there is a tradeoff between the available information and the chart complexity.

Focusing again on the following chart of the **Performance** page, we can add the gender as a second variable on the x axis—click on the plus (**+**) button in the x axis variables and select **Gender**. Afterwards, click on **Gender** and move it before **Race**. This way, the first category displayed becomes **Gender** and the second one becomes **Race**. Your visualization should now look like the following screenshot:
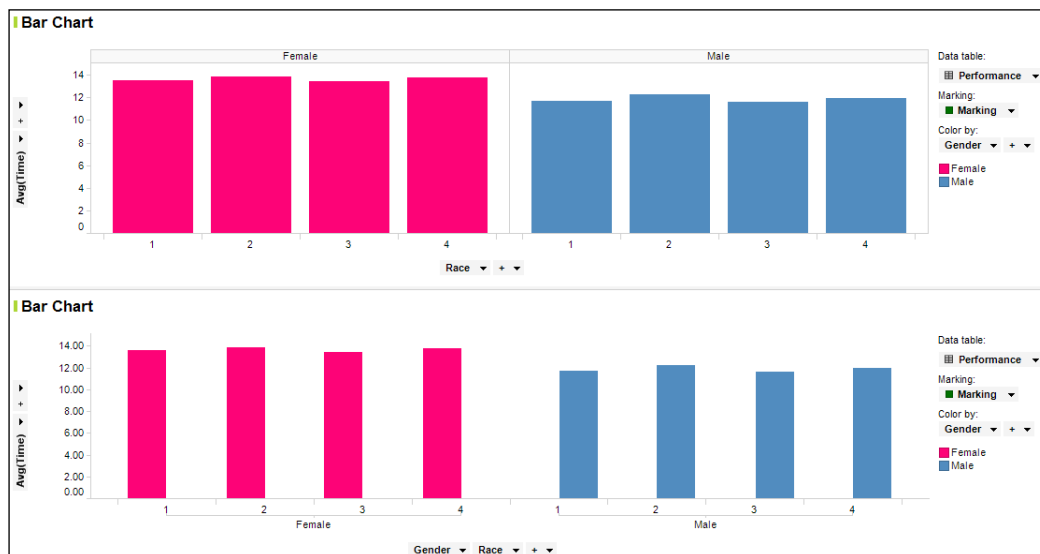
The previous screenshot represents the average times, categorized first by gender and then by race number.

You may have noticed that the visualizations are interactive and chart components are selectable (in the case of this example, chart bars are selectable (individually or by group)). Selecting a bar will display the related data table records in an auxiliary view named **Details-on-Demand**. Please see the following sample screenshot:
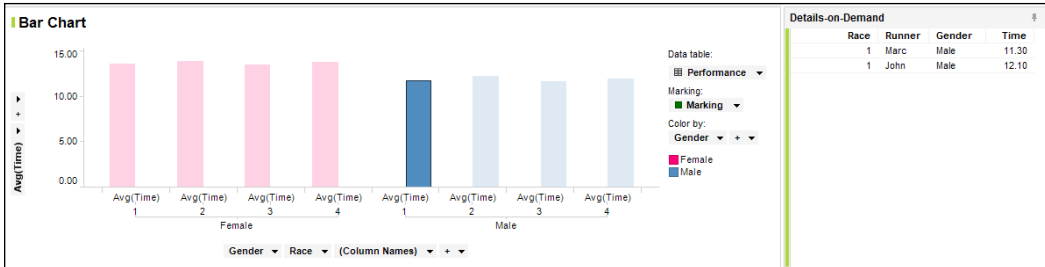
Selection is a great usability feature, which is available in both modes, single and multiple. This feature will be fully described further ahead in the book.

To recap, the y axis of bar charts has to display a summarized (aggregated) variable. In our examples so far, we always used **Average** (of race times), however, Spotfire offers an extensive list of aggregation functions.

Besides **Avg (Average)**, the user can choose **Sum**, **Count**, **UniqueCount**, **Min**, **Max**, **Median**, **StdDev (Standard Deviation)**, and so on. A complete list can be found in Spotfire Professional's **Help Topics** documentation on the **Aggregations Overview** page.

The tool offers extra flexibility by not restricting users to the listed aggregations and allowing them to build their own expressions. The following screenshot is an example of a custom-built aggregation expression:
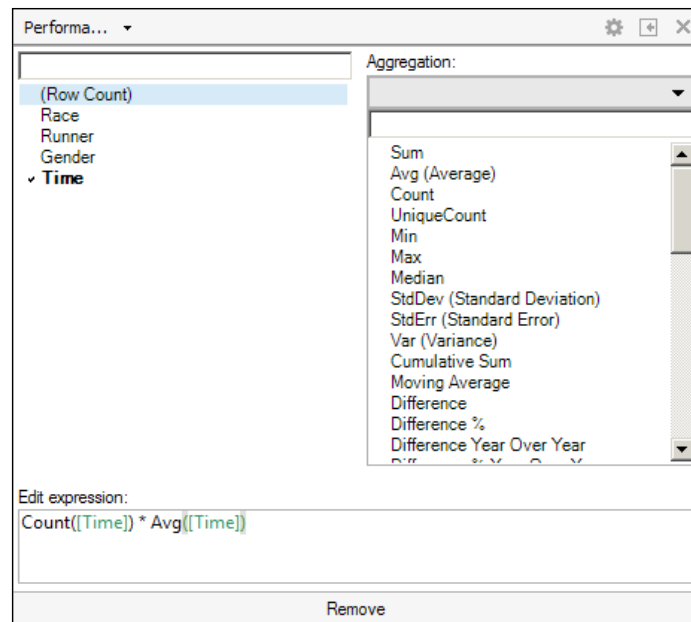
Please note aggregation functions are only available for numeric variables. For string variables, the possible aggregation functions are **Count**, **UniqueCount**, and **CountBig**. An example of this limitation is shown in the following screenshot:
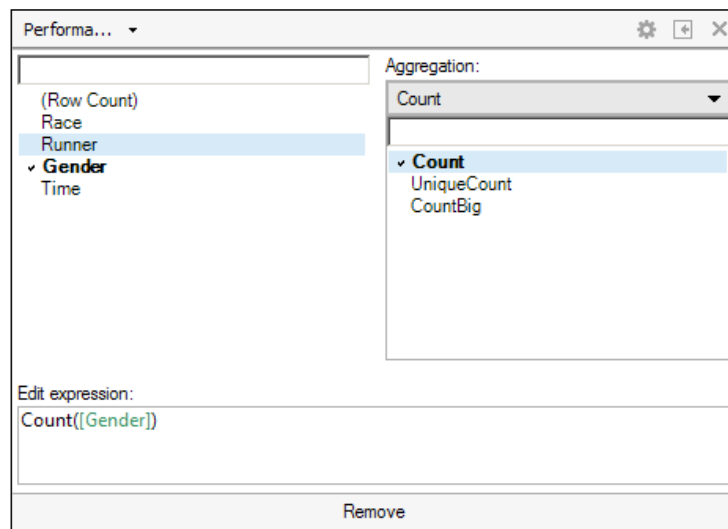
Please save the analysis project.

Line Charts

The Line Chart connects data points together, forming a line. They are used massively in finance, fostering the visualization of trends over periods of time.

In terms of configuration, these graphs are very similar to the Bar Charts, allowing the definition of multiple variables on the x axis, and of multiple aggregated variables on the y axis.

To create Line Chart, start by switching the **Employee Scatter Plot** visualization of the page **Employees** to **Line Chart**.

The Employees data table contains information about employees, namely the hire date, their salary, and their department—one good analysis to create would be the distribution of new hire salaries through the timeline, segregated by department. For this, make sure the following configuration is used:

- The variable on the x axis is chosen by navigating to **HIRE_DATE | Year >> Month** (see the following screenshot for details)
- The variable on the y axis is **Sum(SALARY)**
- Set **Line by:** as **DEPARTMENT_ID**
- Set **Color by:** as **DEPARTMENT_ID**
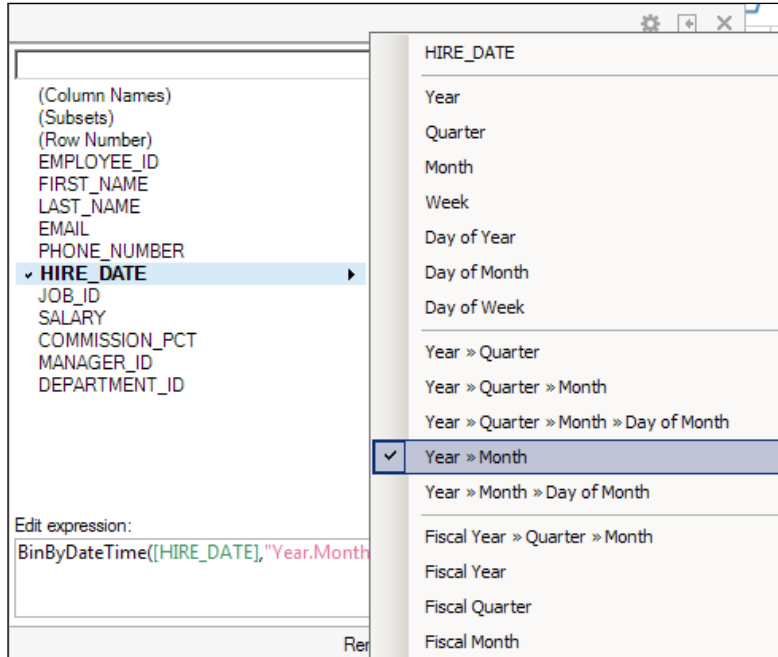
The resulting graph still needs some work as **DEPARTMENT_ID** is defined as an integer and Spotfire colors integers using a gradient interval. Because we did not import this column as a string, we will have to work around it. Please open the chart's **Properties**, navigate to **Colors**, click on **Settings**, and change the **Axis** mode from **Continuous** to **Categorical**. The integer value will now be interpreted as a category instead of an interval.

The created Line Chart should look like the following screenshot:
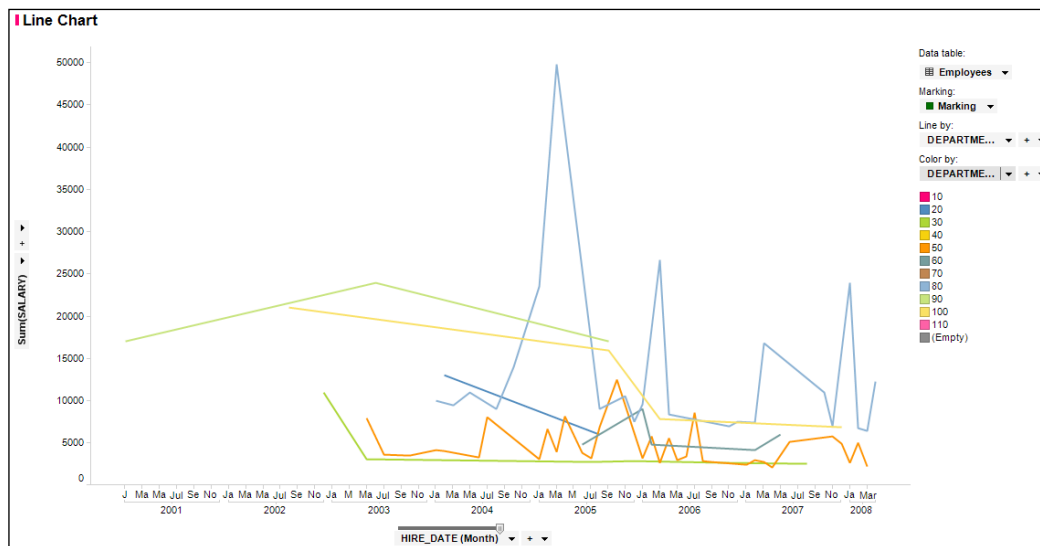
We can finally visualize the sum of the salaries paid to new hires, distributed by department, and month. One spike is noticeable, which is in March 2005; five new sales employees with high salaries joined the company.

Regarding the topic of customization, several chart property values can be manipulated in the chart's **Properties** dialog box. Besides the changes in the x axis, y axis, **Line by:**, and **Color by:**, users can also manipulate **Appearance**, **Formatting**, **Fonts**, **Labels**, **Legend**, **Trellis**, **Show/Hide Items**, and so on.

Please save the analysis project.

Pie Charts

A Pie Chart is a circular graph divided into sections (or sectors), where the size of a section in comparison to the size of the other sections represents the relative quantity.

> This type of chart is named pie due to its similarities with a sliced pie.

To create a Pie Chart, please navigate to the page **StoreSales** of you Spotfire Professional analysis and switch the visualization type to **Pie Chart**.

Regarding the displayed data, we will create a simple example: we will display the total purchases per business location. For that, we are required to configure the following:

- Define variable **Color by:** with value **Business location**
- Define variable **Sector size by:** with variable **Sum(Total)**

> As stated in the *Bar Chart* section, the graph is interactive and single or multiple sections can be selected for the **Details-On-Demand** visualization.

Once again, there are several available properties for users to customize. More information can be found in Spotfire Professional's **Help Topics** documentation.

To improve the chart using properties, we can, for instance, define the following:

- **Pie size by Sum(Total)** (size properties)
- **Gender** as **Columns Trellis** (trellis properties)

This will give us a notion of the sales split not only per city, but also per gender: **Gender Trellis**.

The resulting visualization will be similar to the following screenshot:
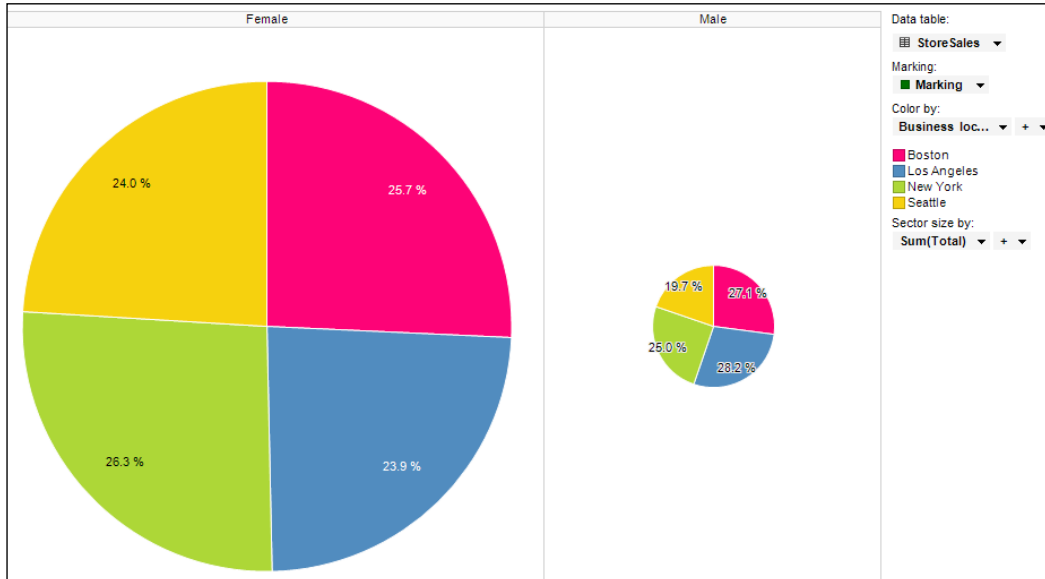
There is a somewhat similar sales distribution through the analyzed cities, although men spend significantly less than women. This difference is depicted due to the **Pie Size by property** setting. Further effort could be placed in realizing why Seattle attracts significantly less spending by men than other cities.

Also, it is clear that men spent about a third of what women spent. Extra work could also be placed in understanding why this happens and what can be done to change it.

Please save the analysis project.

Scatter Plots

The Scatter Plots are built to correlate two variables by plotting their distribution. A straight line is often included in the charts to help identify the distribution's correlation—this line is named the best fit line.

A set of plotted data can have the following distribution types:
- If the points are distributed from the lower left corner to the upper right corner (X increases and Y increases), there is a positive correlation
- If the points are distributed from the upper left corner to the lower right corner (X increases and Y decreases), there is a negative correlation
- If it is hard to draw a line following the distribution of points, then the variables are probably not correlated

To create a Scatter plot, we are going to create a new graph in a new page. Follow the following instructions to do so:

1. Right-click on the page bar (beneath the toolbar) and select the option **New Page**.
2. On the toolbar, click on the button **New Scatter Plot**.
3. Place the new page **Page** in the last position, done by dragging and dropping the page to the last position.
4. Rename the **Page** page to `StoreSales - Total Sales`.

Our aim will be to create the Scatter Plot visualization from the beginning of the chapter, where we identify the city with the biggest spender. To achieve this, please perform the following actions:

1. Select **Total** as variable for the y axis.
2. Select **Number of purchases** as x axis.
3. Leave **Color by:** with the default value of **Business Location**.
4. Select **Gender** as **Shape by**; records will be depicted according to the gender.
5. Select **Age group** as **Size**; the older the group, the bigger the graph element.

To add a best fit line, please access the **Properties** dialog box, and in property **Lines & Curves**, check the box **Straight Line Fit**.

If all steps were executed, the resulting visualization should be similar to the one in the following screenshot:
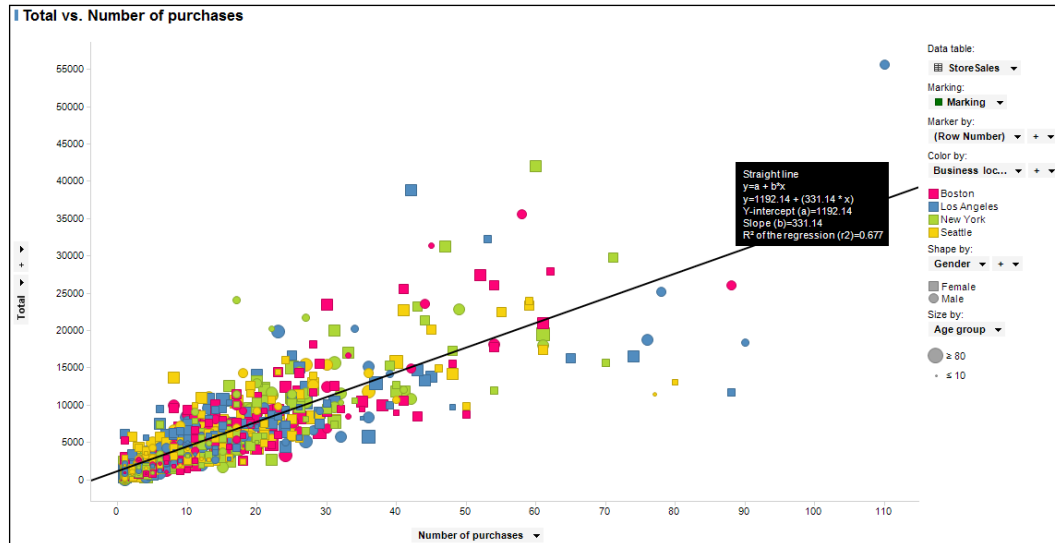
To view the correlation equation and coefficient, please place your mouse cursor over the best fit line.

Correlation (r) varies between -1 (negative correlation) and 1 (positive correlation). The squared correlation coefficient (r^2) (also known as the coefficient of determination) is one of the best means for evaluating the strength of a relationship; in our case Spotfire gives us $r^2 = 0.677$, so r is 0.823. Because $r^2 = 0.677$, 68 percent of our records are directly correlated.

To further investigate the usage of correlation, we will now create a second visualization using uncorrelated variables to generate the points:

1. Please duplicate the plot just created, by right-clicking on top of it and selecting **Duplicate Visualization**.
2. Change the x axis variable to **Customer Age**.
3. Change the y axis variable to **Receipt average**.

The new coefficient of determination is zero. This means the selected data variables are not related at all.

The following screenshot contains both the created scatter plots, side-by-side:
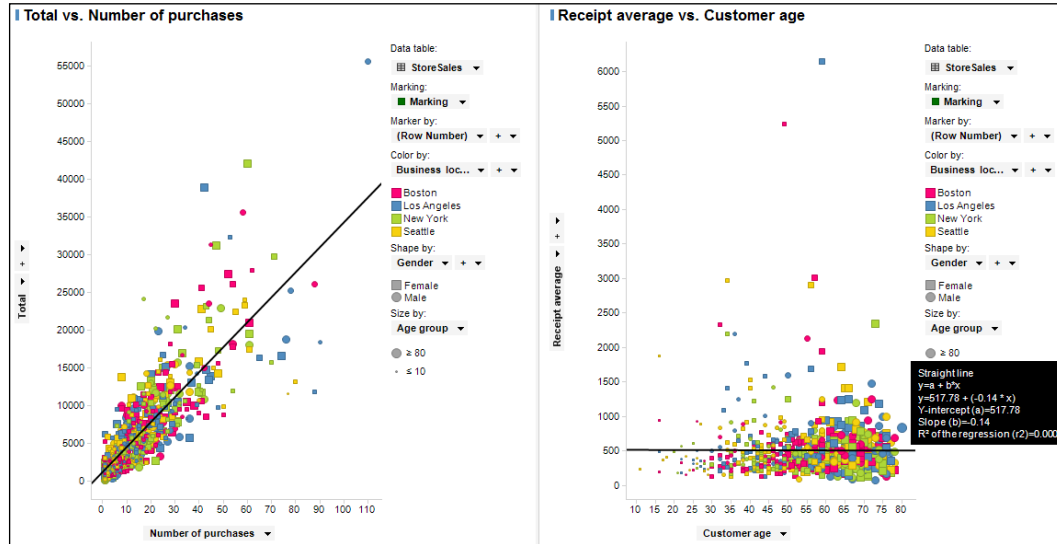
Please save the analysis project.

Box Plots

The Box Plot is a graph used to evaluate the variation of a variable amongst different sets of data. It is formed by three values:

- Upper quartile (Q3)
- Lower quartile (Q1)
- Median

There are also two extra values depicted in each set:

- Upper adjacent value
- Lower adjacent value

The lines from the Upper adjacent value to the Upper quartile and from the Lower adjacent value to the Lower quartile are calculated in the following manner:

Upper adjacent value = Q3 - 1.5 (Q3 - Q1)

Lower adjacent value = Q1 - 1.5 (Q3 - Q1)

Values greater than the Upper adjacent value and values lower than Lower adjacent value are considered Outside values.

The size of each box (boxing Q1 to Q3) depicts the variation of data, so bigger boxes have more varied data.

To create a Box Plot chart, please create a new page named `StoreSales - Age Group comparison` and add to it one chart of type Box Plot with the following characteristics:

- The underlying data table must be `StoreSales`.
- The x axis must be configured to variable **Age group**.
- The y axis must be configured to variable **Total**.
- **Color by:** must be configured to **Age group**.

In the **Properties** dialog box, property group **Lines & Curves**, select the checkbox **Horizontal Line: Avg(Y)**. This will display a horizontal line with the average.

In the **Properties** dialog box, select the property group **Appearance**, and then select the checkbox **show 95% confidence interval of the mean**—it is 95 percent certain that the actual mean for the underlying data is actually within the black line next to the box.

In the **Properties** dialog box, select the property group **Appearance**, and then select the checkbox **show comparison circles**—the circles have their centers corresponding to the mean value of the box to which they belong. Overlapping circles mean the relative boxes are similar.

The resulting graph should look similar to the one shown in the following screenshot:
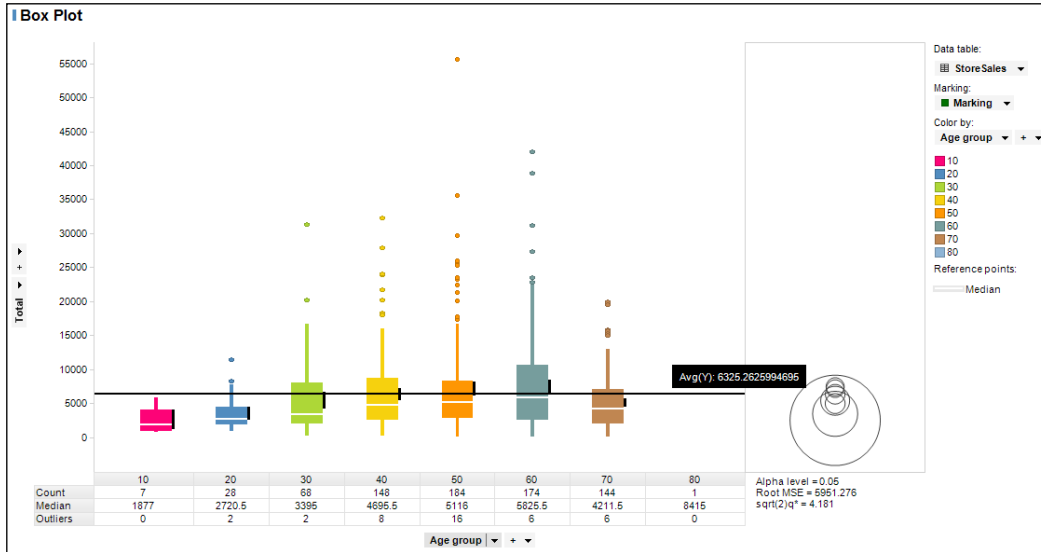
All the outside values are independently depicted as dots. We are also presented with a table listing the `Count`, `Median`, and `Outliers` of each plotted box. This table is named `statistics table` and more data can be displayed in this table, as it is customizable. To access its configuration, please go to the visualization's **Properties** dialog box, property group **Statistics Table**.

To demonstrate the table's customization, we will add the following properties:

- `StdDev` (standard deviation)
- `L95` (lower endpoint of 95 percent confidence interval)
- `U95` (upper endpoint of 95 percent confidence interval)
- `P10` (tenth percentile)
- `P90` (ninetieth percentile)

The resulting table will be similar to the following screenshot:

	10	20	30	40	50	60	70	80
Count	7	28	68	148	184	174	144	1
Median	1877	2720.5	3395	4695.5	5116	5825.5	4211.5	8415
Outliers	0	2	2	8	16	6	6	0
StdDev	2025.62	2533.92	5234.98	5581.36	7263.53	6759.13	3966.35	(Empty)
L95	1109.86	2597.19	4177.79	5426.84	6122.51	6417.64	4404.52	8415
U95	4111	4474.31	6666.3	7225.24	8221.53	8426.24	5700.16	8415
P10	767.6	1466.2	1123.7	1407.2	1682.3	1147.4	1120.2	8415
P90	4847.4	7294	10311.8	12261.4	15452.4	14546.1	10505.4	8415

For a full list of available statistical values and their meanings, please consult the Spotfire Professional's **Help Topics** documentation, page **Box Plot Properties – Statistics Table**.

Please save the analysis project.

The text area

As the name indicates, the text area is not a graph per se. Its purpose is to provide users with customization capabilities, such as text, buttons, and images.

Text areas can be created just like regular graphs, and can occupy a full page or just a section of it.

The **Text Area** palette contains the following items:

- **Text**: This is the main focus of the text area. Text can be added and formatted just as in a regular word processing tool: font, color, alignment, lists, indentation, and so on.
- **Image**: Images (in the formats `bmp`, `jpg`, `gif`, and `png`) can also be added to the text layout.
- **Web links**: These links the URLs to outside web content.
- **Filters**: The existing filter configurations from all available filtering schemes can be a part of the text area layout.
- **Action control**: This allows the association of actions to three control types: **Button**, **Link**, and **Image**.
- **Property control**: This allows the association of controls to property values. Property controls can be of the following types: **Label**, **Input field**, **Input field (multiple lines)**, **Drop-down list**, **List box**, **List box (multiple select)**, and **Slider**.
- **Dynamic items**: These are simplified visualizations to be embedded in the text area. The following types are available to users: **Sparkline**, **Calculated value**, **Icon**, and **Bullet Graph**.

Text areas have two display modes: selected **Edit Text Area** (visualization will be in edit mode) and not selected **Edit Text Area** (visualization will be in view mode). To toggle between modes, please right-click on top of the visualization name bar and select **Edit Text Area**.

The following screenshot presents an example of a text area visualization with one image, one filter, and one dynamic item. This **Text Area** is in view mode.
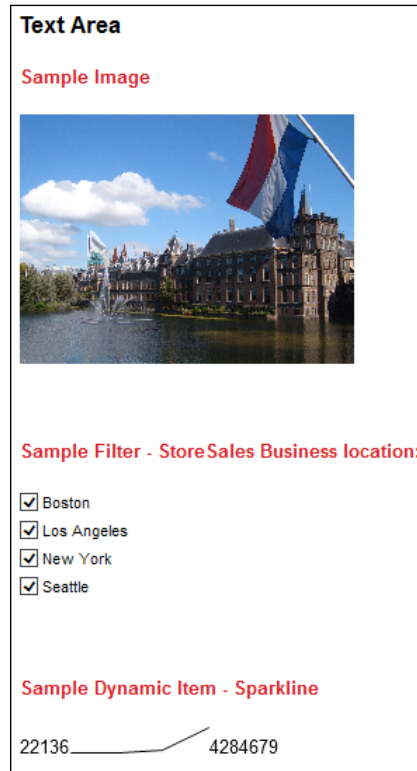
Any changes to the embedded filter will update all visualizations that depend on the filtered data, including any dynamic item.

The steps for replicating it are very simple.

Start by making sure you have an image of one of the compatible types. From the palette, select the button **Insert Image**; a dialog box for selecting the image file will be presented. Select your image and click on **Open**. If the image is too big, it can be resized using the mouse.

To add the filter, select the button **Insert Filter** from the palette; a dialog box for the filter selection will be presented. Please choose **Filter Business location** from the data table StoreSales.

Finally, for the creation of the dynamic item, select the button **Insert Dynamic Item | Sparkline**; a dialog box for graph property selection will be presented. In the **Appearance** properties, select **Start point** and **End point** (their purpose is to display the x axis value at the beginning and end of the depicted line). In the **Axes** properties, for the x axis, select **Recency | Year >> Quarter**, and for the y axis, select **Sum(Total)**. This chart displays the total sales per quarter.

Text can be added where the user sees it fit, and it can be formatted with the palette formatting options.

If you followed the previously listed steps, please do not forget to save the analysis project.

Details visualizations

Details visualizations are a great feature available in Spotfire. They act as daughter visualizations of one or more parent graphs, where the displayed data will have, as source, the marked (Marking) data of the parent(s). The details visualization data table will be the subset of data of the parent's data table selected by the user, which in turn will correspond to the records available to the users in the **Details-on-Demand** view.

Users can derive details visualizations from other details visualizations in a theoretically unlimited cascade.

We will create one details visualization for the StoreSales page pie chart. This chart displays the percentage of total spending in each location, segregated by men and women. This new graph will also be a pie chart, and will further drill down in to the spending details, by providing information about the customer's age group.

To create the details graph, please follow the following instructions:

1. Right-click on top of the **Pie Chart** visualization and navigate to **Create Details Visualization | Pie Chart**. We will create a details pie chart.

2. Change **Color by:** to **Age group**. Drill down on age group.

3. Access the chart's **Properties** dialog box, select the property group **Data**, and select the value **Empty visualization** in **If no items are marked in the master visualizations, show**. With this option, the details table is filled only if there is marked data in the parent visualization (the default option shows all available data if nothing is marked).

4. Access the chart's **Properties** dialog box, select the property group **Labels**, and then select the checkbox **Show in labels** under **Sector category**, in order to display the **Color by:** category value in the pie slices.

5. Access the chart's **Properties** dialog box, property group **Size**, and set the **Pie size** to the maximum, to maximize the chart's size.

If you followed all indications, your visualization should be similar to the one in then select following screenshot:
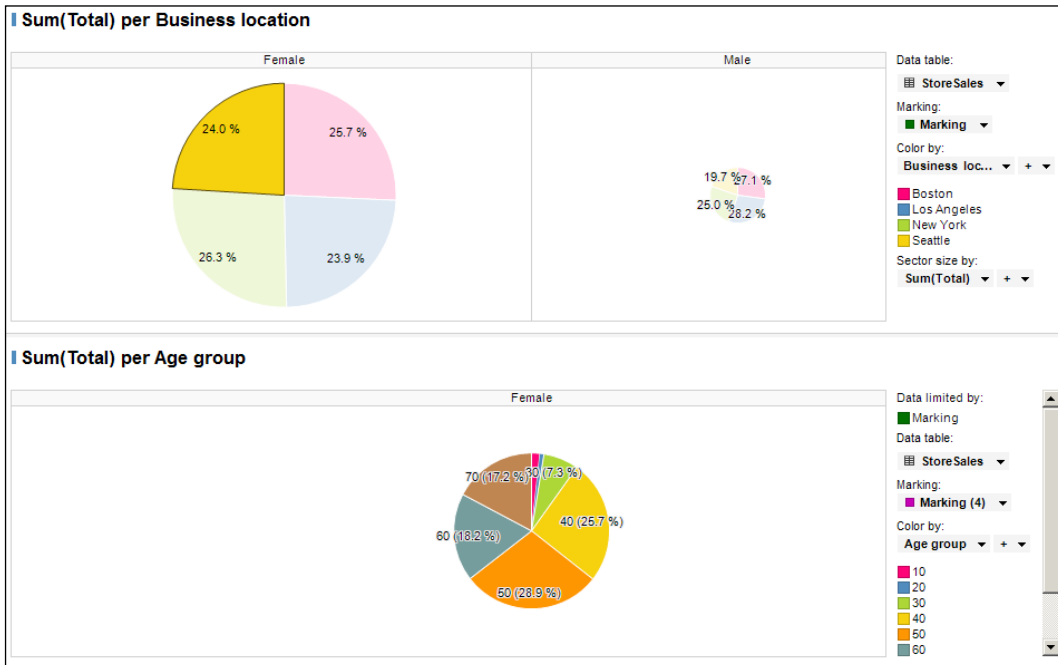
This detailed visualization displays the total spent by the different age groups, solely for the females from Seattle.

To demonstrate the cascading details, we will now create one details visualization for the just created details visualization. It will be a combination chart with the purpose of counting the number of records for each customer age. This chart will be rendered individually for each age group, as it will be a detailed visualization of it.

To create the graph, please follow the following instructions:

1. Create a new page to host the new chart as adding a chart to the page **StoreSales** would make it too confusing. Rename it to `StoreSales - Age Details`.

2. Back on the StoreSales page, create a details visualization (of type **Combination**) for the details pie chart. This new chart will be placed under the parent one.

3. Drag the new details chart to the created page (StoreSales - Age Details) and drop it there (charts are dragged by the bar where the chart name is displayed).

4. For the x axis, select the variable **Customer age**.

5. For the y axis, select the variable **(Row Count)**.

6. Define **Customer age** as **Series by:**. This assigns a different color to each age of the age group.

If you followed all indications, your visualization should be similar to the following screenshot:
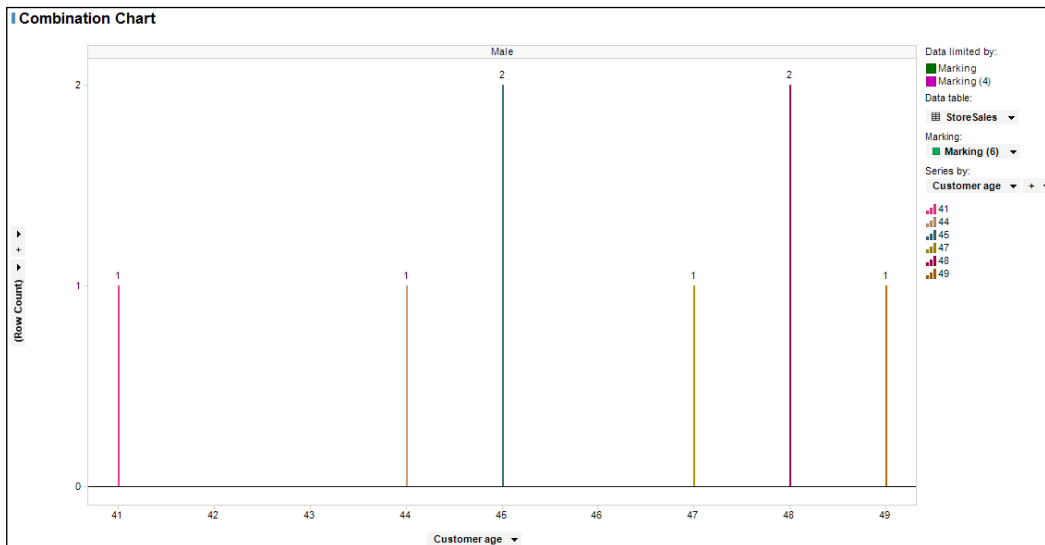
Please note that this chart will only present data if there are markings in its parent chart, which in its turn will also only present data if there are markings in its parent chart.

This details visualization corresponds to the males from Boston marking of the grandparent visualization, and the age group 40 of the parent visualization.

Please save the analysis project.

Summary

We have now reached the end of this chapter. Its intent was to introduce the main capabilities of Spotfire as a tool to create analysis projects. We saw how to load data in the platform, how to create several visualization types (Tables, Bar Charts, Line Charts, Pie Charts, and so on), and how to customize existing visualizations.

Finally, we learned what a details visualization is and how it can be configured.

In the next chapter, we will continue to build up our Spotfire knowledge, by looking at more advanced features, such as filtering, data preparation and transformation, and Spotfire Web Player usage.

4
Mastering Visualizations

In this chapter, we will continue improving our Spotfire skills. We will focus on activities such as data filtering, data manipulation, and library creation.

We will also focus on analysis visualization and interaction, resorting to Spotfire Web Player Server — the reader will become familiarized with the tool's main features. We will focus a bit on this product, as it is a great tool to share analysis with others.

Filtering

The function of filters is to restrict the data shown in visualizations. This means that the users can limit the records of the visualization's data table that are displayed or used in the visualization itself. The filtering out of data has an impact on visualization aggregations and axis scales, and its application is immediate, as any selection/deselection will be rendered without any delays.

By default, a filter is created by Spotfire for each column of each existing data table, and all filters are available in a panel named **Filters**, that is presented on the top-right corner of the Spotfire Professional display.

Filters are grouped by data table, and these tables will be sorted chronologically by ascending creation date. Focus will be on the first data table that is used in a visualization of the current page.

The following screenshot is an example of the **Filters** panel for the visualization of the page **Store Sales - Total Sales – StoreSales** data table.
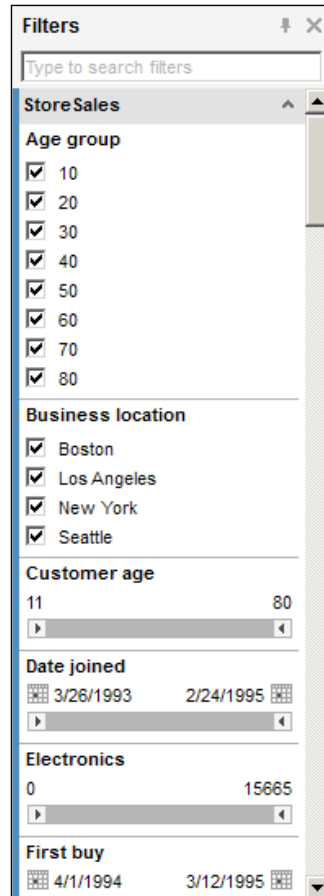
This panel also has a search field to allow a quicker finding of filters by their name.

Filter names not matching with the search keyword will be hidden from the user and cannot be applied. To display all filters again, the user must remove all the keywords from the search box.

The following screenshot displays a filter search for the keyword **gender**.
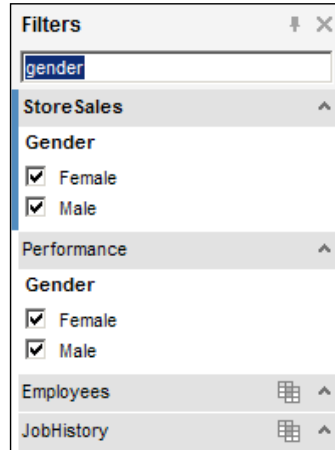
Two distinct filters were found which match the keyword: one from the data table **StoreSales** and the other from the data table **Performance**.

The **Filters** panel can be moved, closed, and opened. To open and close the panel, users can navigate to the **View | Filters** menu, or the **Filters** button on the toolbar. Both the enumerated options are presented in the following two screenshots respectively:
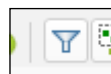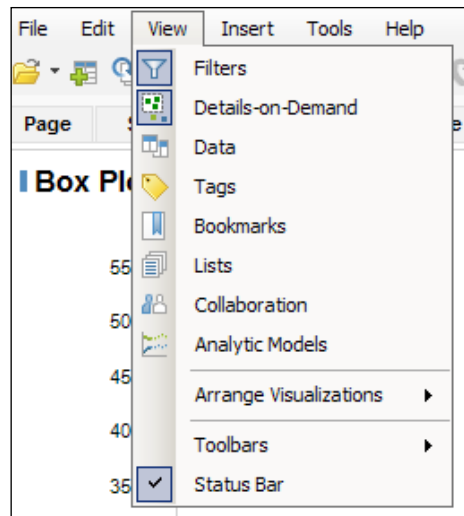
Filter types

Just as the data columns have an underlying data type, filters also make available several interaction types to improve and facilitate usage. Spotfire offers the following filter types:

- **Item filter**: This is represented by a slider, which allows the user to select one, none, or all the values. When the slider is on the left, all values are displayed; when the slider is on the right, no values are displayed; and when it is in between, unique values are displayed. The next screenshot presents an example of such a filter type:
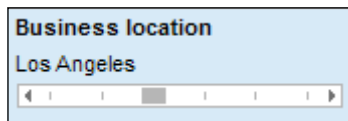
- **Checkboxes**: There is a set of checkboxes, one for each value of the filter. A user can select none or more checkboxes to filter the displayed values. If there are empty values in the column, an **(Empty)** checkbox will be available as well. The following screenshot presents an example of such a filter type:
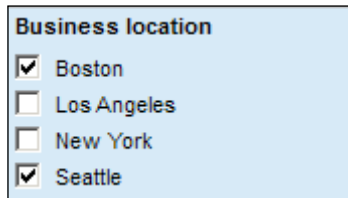
- **Radio buttons**: These buttons have the same behavior as an item filter, but the display is different. Because radio buttons allow only one choice, the user will also have **(All)** and (**None)** options. The following screenshot presents an example of such a filter type:

- **Range filter**: This type of filter operates with a slider, where two handles exist: one on the left corner and other on the right corner. These handles allow the user to set the ranges that are being filtered. Because this filter type involves the setting of an interval with minimum and maximum value, it is best suited for columns of the type `integer` or `date(time)`. The label values above the slider (that represent the interval limits) can be double-clicked and set manually (for date formats, users can simply click on the calendar icon and choose the date from the pop up). The following screenshot presents an example of such a filter type:
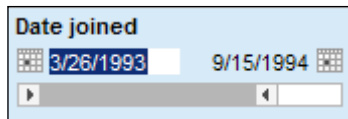
- **Text filter**: This type of filter allows the user to type in a string, which will be used for matching with filter values. Values that do not match will be filtered out. This search is non-case sensitive, white spaces will filter out values without all search strings (OR can also be used for disjunction), and the wildcard `*` can also be used. The following screenshot presents an example of such a filter type:
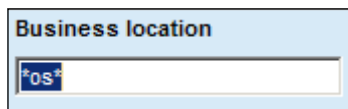
  This filtering will filter out the records with **Business locations** other than **Boston** and **Los Angeles**, as only these two match with the expression `*os*` (only these two contain `os` somewhere in their value).

- **Listbox filter**: This allows the selection of filter values to apply from a list of values contained in a column box. Multiple selections can be made at a time using the *Shift* and *Ctrl* keys. This type of filter also contains an input search field that allows the user to filter the values displayed in the box.
  The following screenshot presents an example of such a filter type:
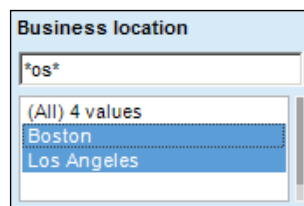
[
The initial type of a filter is defined by Spotfire Professional, based on the number of records of the underlying data table and the data type of the related column.
]

Customizing the filters

By right-clicking on a filter, the user is presented with a filter options menu, which allows manipulation of filters.

The following options are available:

- **Filter type**: For a certain filter, we can toggle between different filter types.
- **Reset Filter**/**Reset all Filters**: It includes options to reset the filters.
- **Rename**: It is used to change the name of a filter.
- **Hide Filter**: This option allows the user to temporarily remove unnecessary filters, in order to reduce the amount of data shown in the **Filters** panel. If this option is used, a warning will appear in the **Filters** panel as **Some filters are hidden. Show all**. Clicking on **Show all** will unhide all the filters. For a finer-grained hiding/showing, users should use the option **Organize Filters**.
- **New Group**: Users can create groups to better organize the existing filters. Hierarchically, groups can only be created under data tables, which means each group can only organize filters of one data table.
- **Move to Group**: This allows the moving of a filter to an existing group.
- **Organize Filters**: This is a dialog box, which hierarchically displays the data tables, groups, and filters of the analysis, and allows the user to set their sorting and visibility settings. Group creation and deletion are also possible here (although it is not possible to add filters to groups). The following screenshot displays such a dialog box:
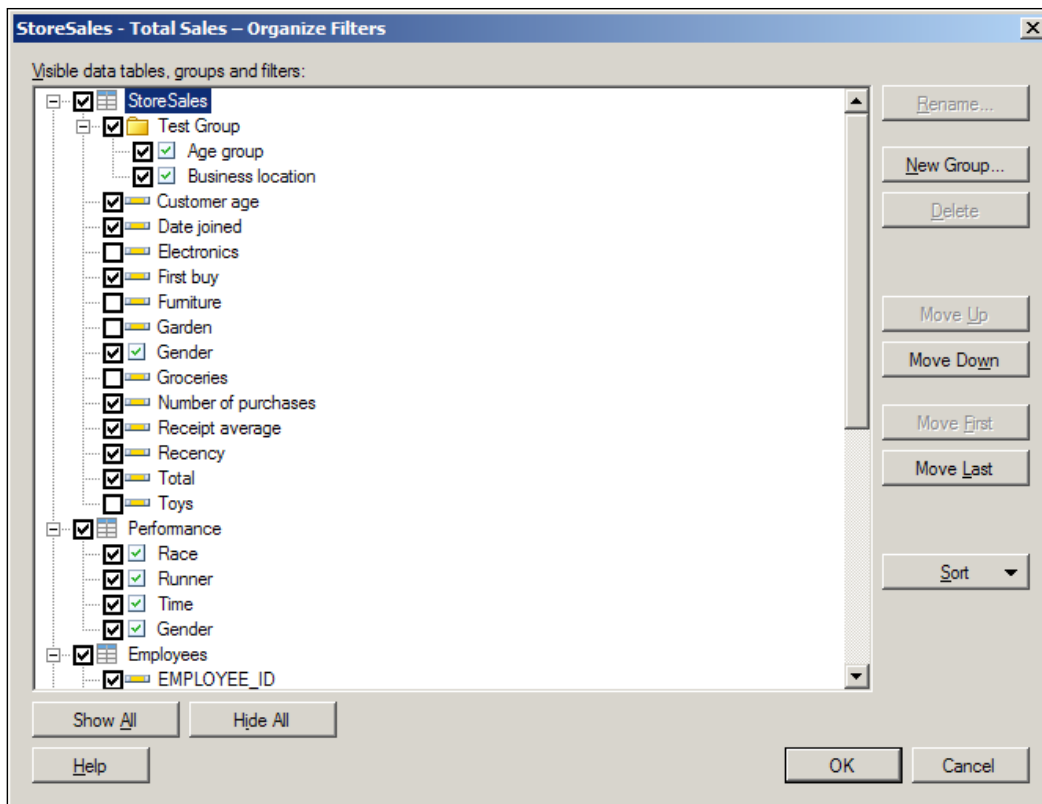
The order in which the data tables should be sorted can also be set within this dialog.

Filtering schemes

Spotfire places a lot of emphasis on its data filtering capabilities.

For increased flexibility, users can create their own **schemes**, where filters can be applied independently of other schemes.

All schemes will always contain all available filters. It is the application of filters that is independent between the schemes. The following two screenshots demonstrate this independent behavior; although the scheme **Filtering scheme 1** has values **10**, **20**, **30**, and **40** filtered out from the filter **Age group** (of the data table **StoreSales)**, the scheme **Filtering scheme 2** has all the values included.
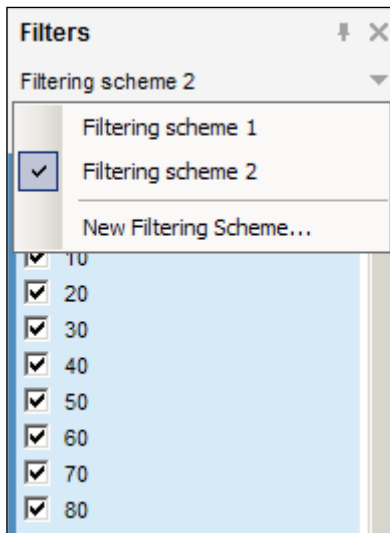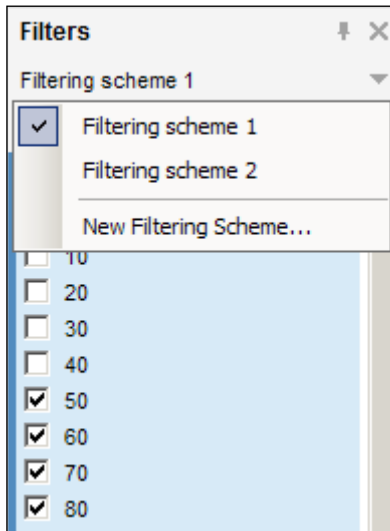
Any change between schemes has an immediate impact on the displayed visualizations.

> To add filtering schemes, please navigate to the menu **Edit |
> Document Properties**, and click on the **Filtering schemes** tab.
>
> To make a drop down available in the **Filters** panel that allows
> selection of a scheme, the checkbox **Show filtering menu in Filters
> panel** must be selected.

Applying filters

To illustrate the application of filters, we will now see the difference between
two visualizations with and without filters. The following two screenshots will
demonstrate this.

The first screenshot contains the visualizations of the page **Store Sales - Total Sales**
without any filtering:
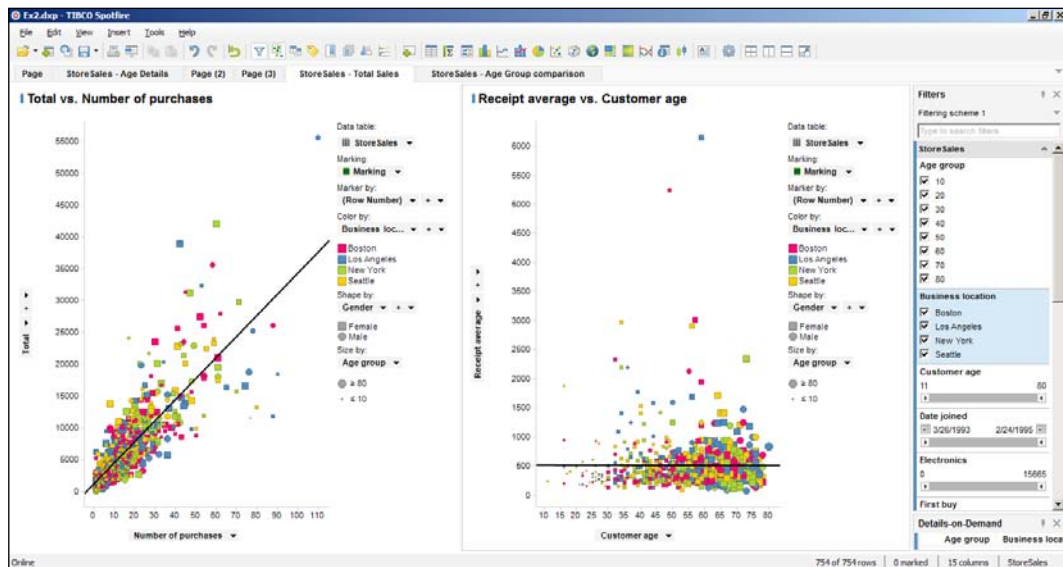
The following screenshot presents the same visualizations with the following filters applied:

- **Age group**: **50**
- **Business location**: **Seattle**
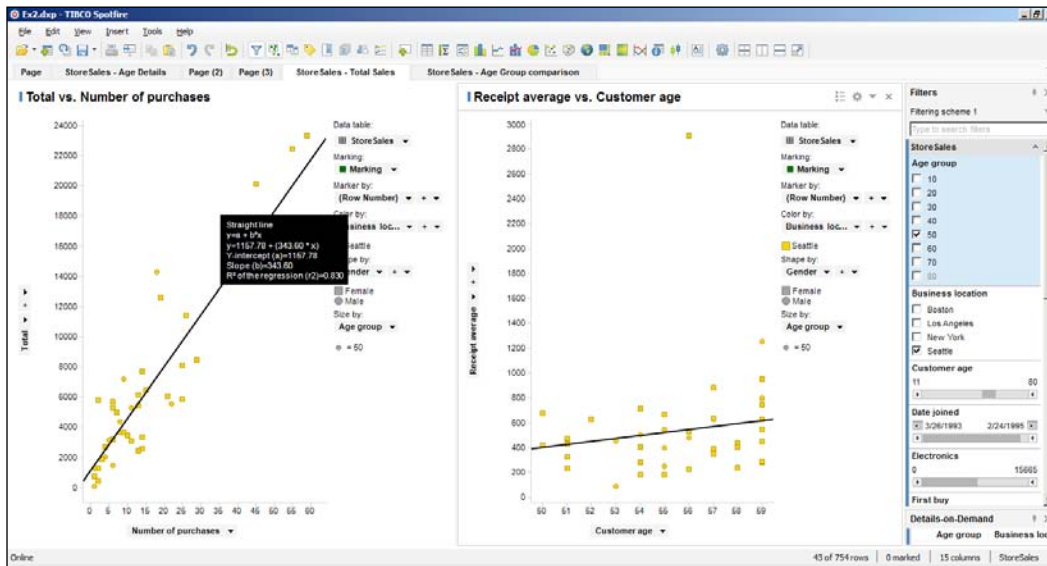
It is curious to realize that the correlation between the variables **Total** and **Number of purchases** increases with the application of these filters. With the filtered out data, the direct correlation is of 83 percent.

Hierarchies

Hierarchies are a special type of filter, where the filter itself is displayed in a hierarchical format. Conceptually, these filters behave as data trees.

An example of a hierarchy is the definition of a **Calendar**, divided into years, quarters, months, weeks, and days.

The simplest way to demonstrate this Spotfire feature is to create **Date Hierarchy Filter** from a **Date Range filter type**. We will use the **Date joined** filter from the **StoreSales** data table.

Right-clicking on the **Date joined** filter, and then selecting the **Insert Hierarchy** option, will present a dialog box as shown in the following screenshot:
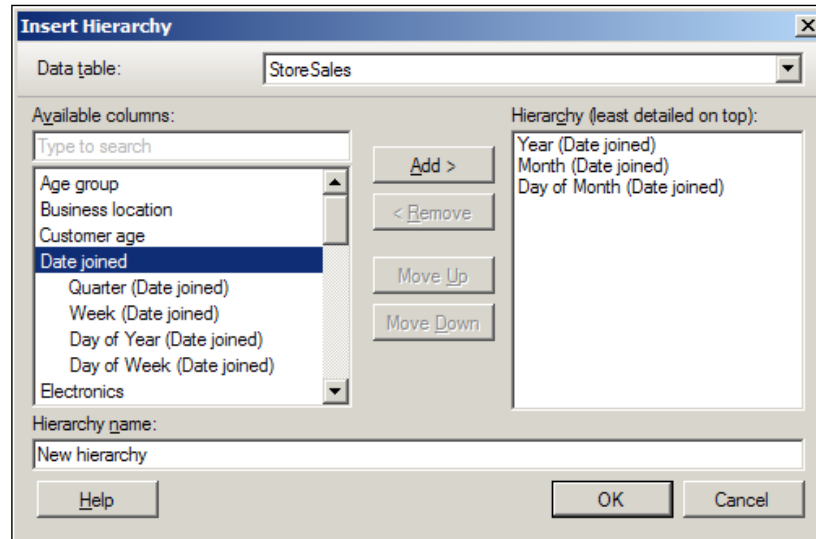
The default selected columns for hierarchy creation are **Year**, **Month**, and **Day of Month**. The remaining selectable columns are **Quarter**, **Week**, **Day of Year**, and **Day of Week**. We can add the **Quarter** column and move it up between **Year** and **Month**. The hierarchy configuration will look similar to the following screenshot:
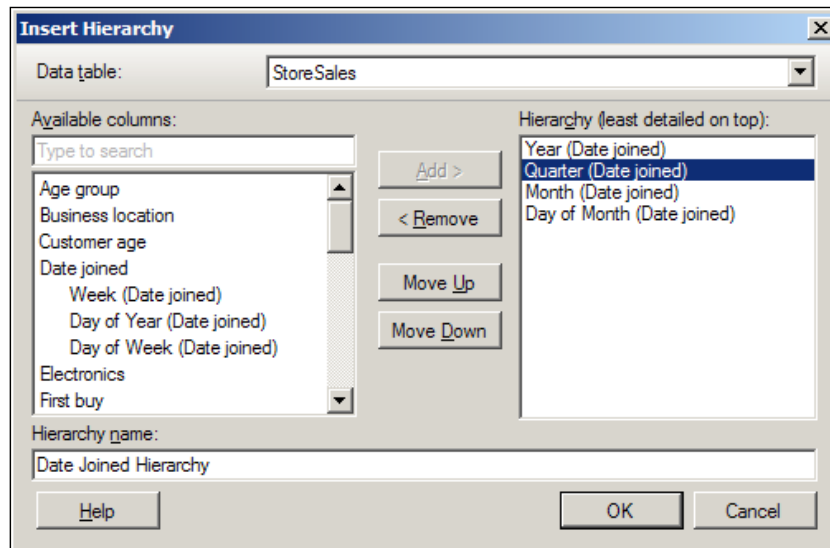
The addition of a column can be done by selecting the column and clicking on **Add**, and shifting of the column through the hierarchy can be done using the buttons **Move Up** and **Move Down**.

Hierarchies also have a name, which can be set during the creation process (in the dialog box) or changed later (by right-clicking on top of the hierarchy).

The created hierarchical filter does not replace the original one. Both will coexist in the analysis. In our example, both the **Date joined** filter and the hierarchy named **Date Joined Hierarchy** will be shown to the user.

The following screenshot presents the created hierarchy:
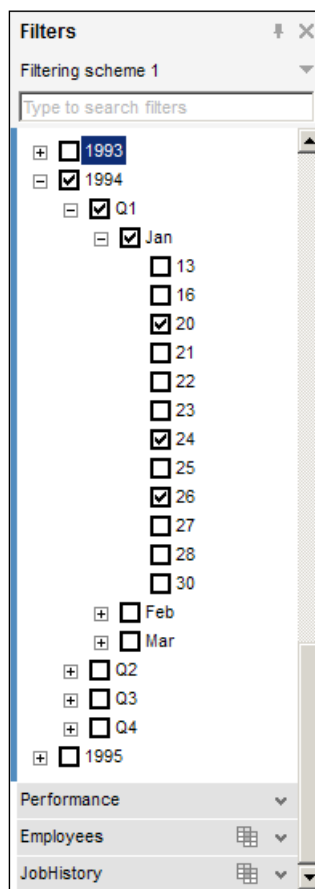
Hierarchies can also be created using two or more data table columns. The only existing limitation is related to the fact that only one data table can be used in one hierarchy.

Marking and Details-on-Demand

As already specified, graph elements (displayed data) can be selected, and this is called **marking**. This action has several purposes, such as updating the details visualization (in the **Details-on-Demand** panel), copying records, and so on.

The user has a varied set of options available for marking: click on the item, *Ctrl* + click on the item, *Shift* + click, *Ctrl* + *Shift* + click, *Alt* + click, and so on. A complete list with detailed descriptions can be found in the Spotfire Professional's *Help Topics* documentation page, *Marking in Visualizations*.

To unmark a selection, users must right-click on the top of the visualization and select **Marked Rows | Unmark**. Other options will also be offered to the user such as **Invert**, **Delete**, **Filter To**, and **Filter Out**. The following screenshot lists these options:
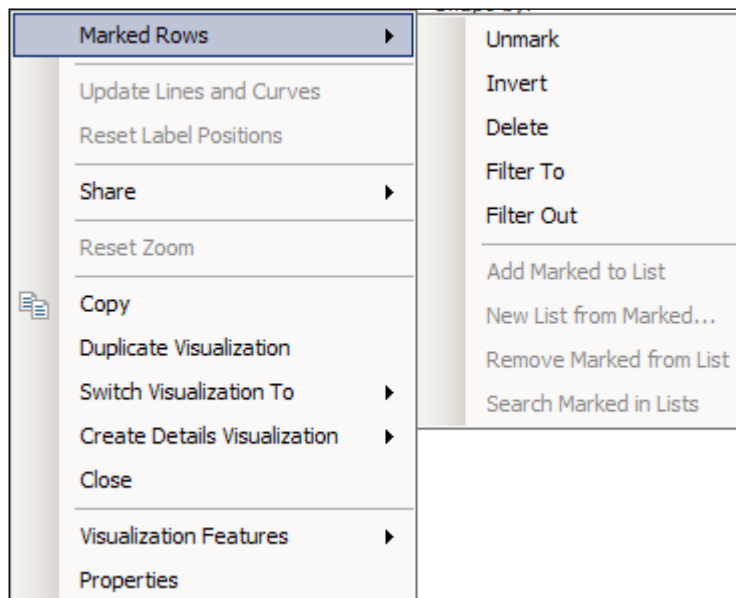
The data table records of marked data will be listed in the **Details-on-Demand** panel of Spotfire Professional. This panel is located by default on the right lower corner of the screen, just below the **Filters** panel. And just like the filters, it can be moved, closed, and opened. To open and close the panel, navigate to the **View | Details-on-Demand** menu, or the **Details-on-Demand** button on the toolbar.

This panel can be customized, by using its **Properties** dialog box (obtained by right-clicking on the window):

- **Appearance**: It allows setting the appearence of the panel, in terms of number of rows for header, number of rows for data, show or hide cell borders, and column orientation
- **Fonts**: It allows selection of font for table text and headers
- **Columns**: It shows data table columns and their sorting (by default, all are shown). Also allows the specification of a renderer for each column (**Link**, **Text**, or **Image from URL**)
- **Sorting**: It defines the sorting order by a maximum of three columns in ascending or descending order

Preparing and transforming the data

As already emphasized, in Spotfire, the data loaded into data tables is crucial for successful visualization creation. It must be in a workable format (or a format suitable for aggregation) as Spotfire's data table data manipulation capabilities are somewhat limited.

This does not mean, however, that Spotfire cannot perform some work on the loaded data. The tool has the following loading and transformation capabilities:

- Data loading: It includes:
    ° Adding a data table
    ° Inserting data table columns
    ° Inserting data table rows

- Data transformation: It includes:
    ° Pivoting the data
    ° Unpivoting the data
    ° Calculating and replacing the column
    ° Calculating a new column
    ° Changing the column name
    ° Changing the column data type

- Replacing data table

All these capabilities will be described in the following sections.

Adding a data table

This topic was already presented in the *Data loading* section of *Chapter 3*, *Creating Visualizations*. The theory was accompanied by the demonstration of loading from one of the three different data source types: text file, MS Excel file, and database.

Inserting data table columns

Spotfire Professional allows its user to load extra (or new) columns to an already existing data table.

This is useful due to the fact that the visualizations have only one main data table. The records of main data tables are the base for the creation of graphs, defining rows in nonaggregated visualizations, and items in aggregated ones.

Users are able to insert new columns from a second data source, extending the original data table. More importantly, the second data source can be totally unrelated to the first data source (of the already loaded data).

To add a new column, please access the menu **Insert | Columns**. If more than one data table exists in the analysis, a dialog box will be launched for the user to choose the table to which add the columns (dialog box **Select Destination**). In our case, we will select **Employees**. After clicking on **Next**, the dialog box will allow the user to configure the data source, as shown in the following screenshot:
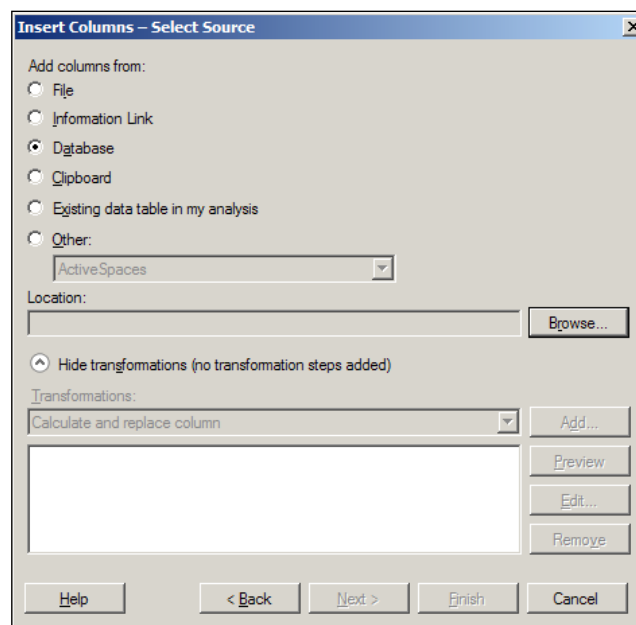
We will add the columns from the JOBS table of the hr database schema; the configuration steps are identical to the *Loading from Databases* scenario previously described in *Chapter 3, Creating Visualizations.*

After configuring the source, the user will be presented with the **Match Columns** dialog box (here we should configure the matching columns between both tables). The **Match All Possible** option handles this task for us, as in our case, both columns to relate share the name and can be automatically identified by the application.

Refer to the following screenshot where the **Column Match** dialog box is depicted:
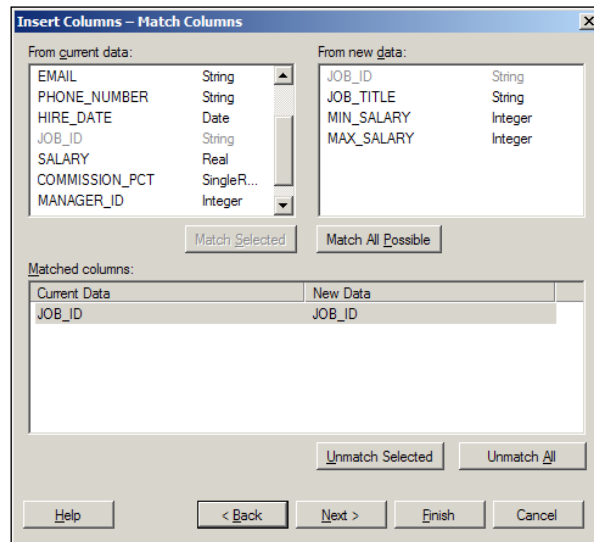
By clicking on **Next**, the user is moved to a column selection dialog box; although several columns were selected while loading the data source, the user might not wish to add all to the existing data table. In our case, we will select **JOB_TITLE** (refer to following screenshot for details).
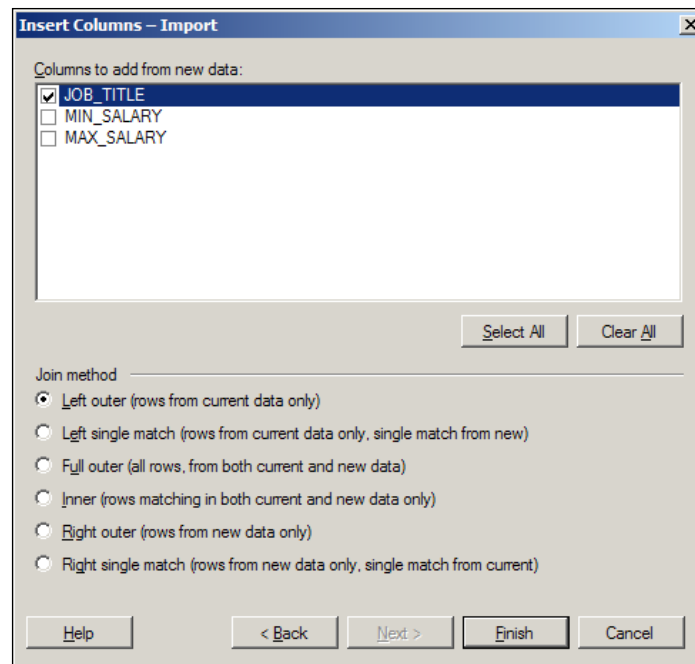
There are also some options related to the definition of which data is going to be included in the resulting data table—**Join method**. We will leave the default option.

Clicking on **Finish** terminates the dialog box and imports the data.

> To confirm the new column, we can go to the menu **View | Data**. A new panel is opened on the left-hand side of the screen. Choosing the data table **Employees** will reveal that there is in fact a new column.
>
> To further analyze the data, it is also a good idea to create a new page with a new table visualization. Besides the job ID, there is now also a job description column.

Inserting data table rows

Adding rows to a data table is very similar to adding columns to a data table. And also, as happens for columns, this feature serves as a solution for extending any main data table that exists in the analysis.

The added rows need to match the existing data table columns if automatic matching is to be done. If not, manual matching has to be performed.

Create the following MS Excel worksheet so that we can import the data:

	A	B	C	D	E
	Race	Name	Sex	Seconds	Age
	5	Marc	Male	11,7	18
	5	John	Male	11,6	17
	5	Mary	Female	13,1	18
	5	Wanda	Female	13,2	16

To add new rows, access the menu **Insert | Rows**, and we will use the data table **Performance** as the destination for the new rows.

The automatic matching of columns (from the **Match All Possible** button) will only work for **Race**. All other columns have to be matched by hand as the names differ. The resulting matching should be similar to the following screenshot:
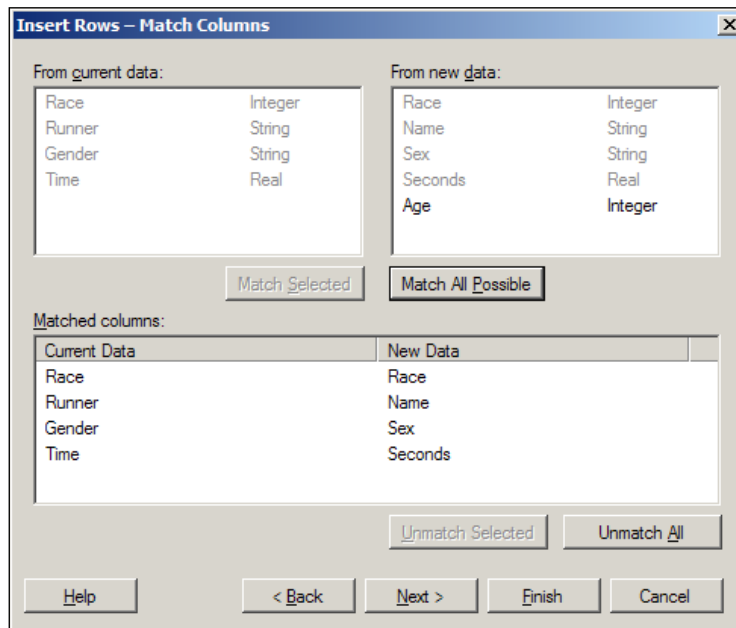
Since **Age** (or similar) does not exist in the **Performance** data source, we will not include this column. Its addition would lead to the creation of empty column values for the records already in the **Performance** data table.

> A simple way to confirm the addition of new rows is by creating a new visualization of the type Table and by selecting the new data table as main data table.
>
> In our scenario, if you navigate to the **Performance** page, you will see that **Race 5** was added to the page's visualizations.

Data transformation

While loading data into Spotfire, the user is allowed to modify it. The aim of this feature is to permit users to clean up the data, fix data errors or omissions, or even change formats and types of the columns.

This task has to be specified while loading the data, which as specified previously, occurs solely in the following scenarios: **Add Data Tables**, **Insert Data Table Columns**, **Insert Data Table Rows**, and **Replace Data Table**. Then, the transformations have to be defined in the **Add Data Tables** dialog box. Refer to the **Transformations** section on the following screenshot:
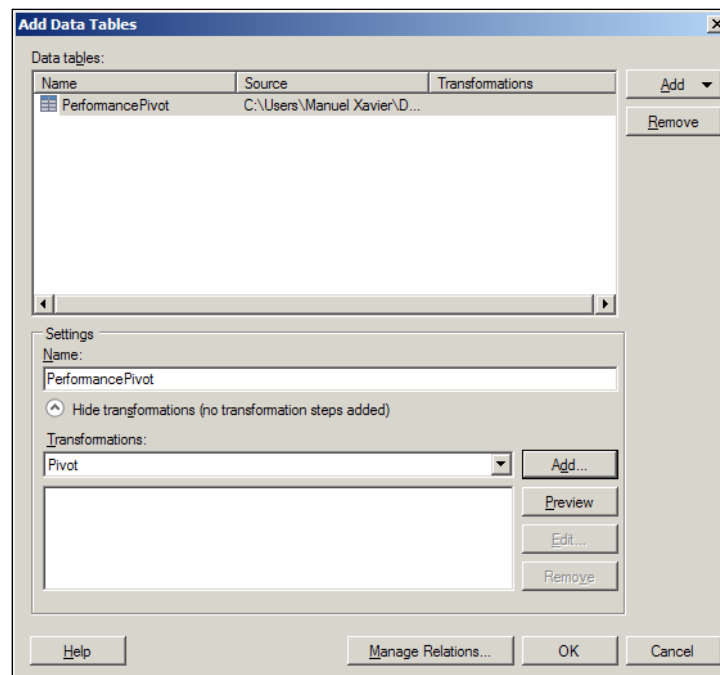
Pivoting data

A pivot table in Spotfire has the same meaning as in MS Excel. The purpose of these tables is to summarize the data.

This transformation mutates a tall and skinny data table into a short and wide one, where the new columns are usually comprise of aggregated values.

As an example, we will pivot the data of **Performance** into a new **PerformancePivot** data table.

It is during the creation of the data table, that the user will have the opportunity to define its pivot transformation. Our aim will be to pivot [for each **Runner**] **Time** per **Race** (without any selected aggregation function).

To achieve this, follow these steps:

1. Start by creating a new data table named **PerformacePivot**, by loading the **Performance** Excel worksheet. From the **Transformations** drop-down list, select the **Pivot** option and click on **Add**. A dialog for the **Pivot** definition will be presented as shown in the following screenshot.

2. Define **Runner** as **Row identifiers** (for each runner), **Race** as **Column titles** (per race), and **Time** as **Values** (value of the records).
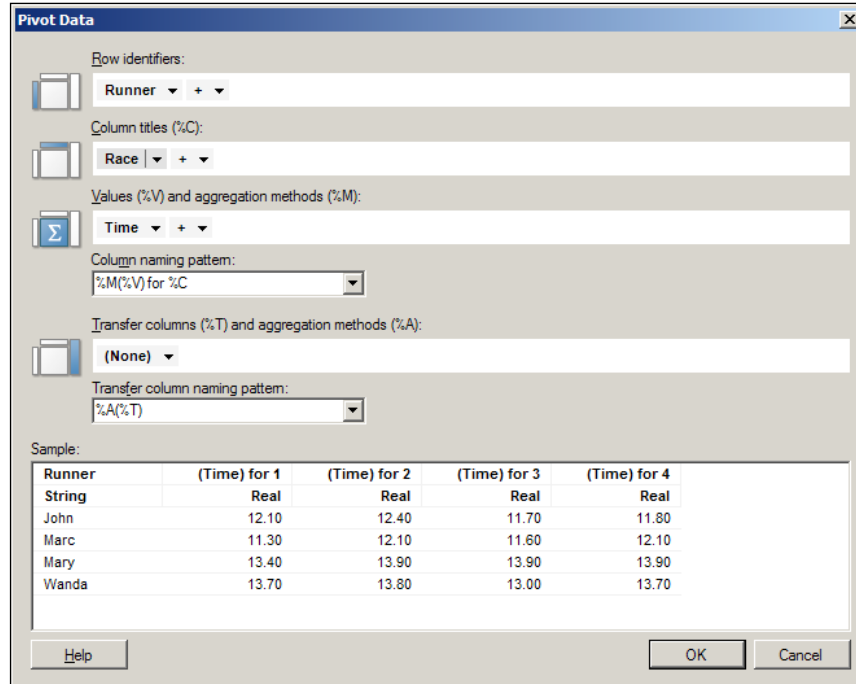
Comparing the pivot data table with the original format, we can easily visualize the transformation from tall and skinny to short and wide (both tables are shown side-by-side in the following screenshot):

Unpivoting data

Unpivoting a table transforms it in the opposite way of pivoting: from short and wide to tall and skinny.

As an example, we will unpivot the **StoreSales** data into a new **StoreSalesUnpivot** data table.

The procedure to unpivot data is similar to the procedure to pivot data; a new data table needs to be created with an unpivot transformation. Our objective will be to unpivot the sales department (**Furniture**, **Toys**, and so on) into a **Department Id** column, and department totals in a **Department Total** column.

To achieve this, please create a data table named **StoreSalesUnpivot** loading the StoreSales.txt file. Choose transformation of the type **Unpivot** and define the transformation as presented in the following screenshot. The meaning of the unpivot configuration is as follows:

- **Columns to pass through**: It lists the columns that are not affected by the transformation, but which will exist in the new data table.
- **Columns to transform**: It lists the columns that are to be merged in a single category column. In our example, we will unpivot the several departments (**Furniture**, **Toys**, and so on) into a single column named **Department Id**.
- **Available Columns**: It defines columns that will not exist in the new data table.

- **Category column name**: It defines the name of the new category column.
- **Value column name**: It defines the aggregated value for the result of transforming multiple [integer] rows into one category row. This value is calculated by aggregating the different values from the excluded columns. In our example, this value is the result of adding the department total for the several distinct records of: **Customer age**, **Date joined**, **First buy**, **Recency**, **Number of purchases**, and **Receipt average**.
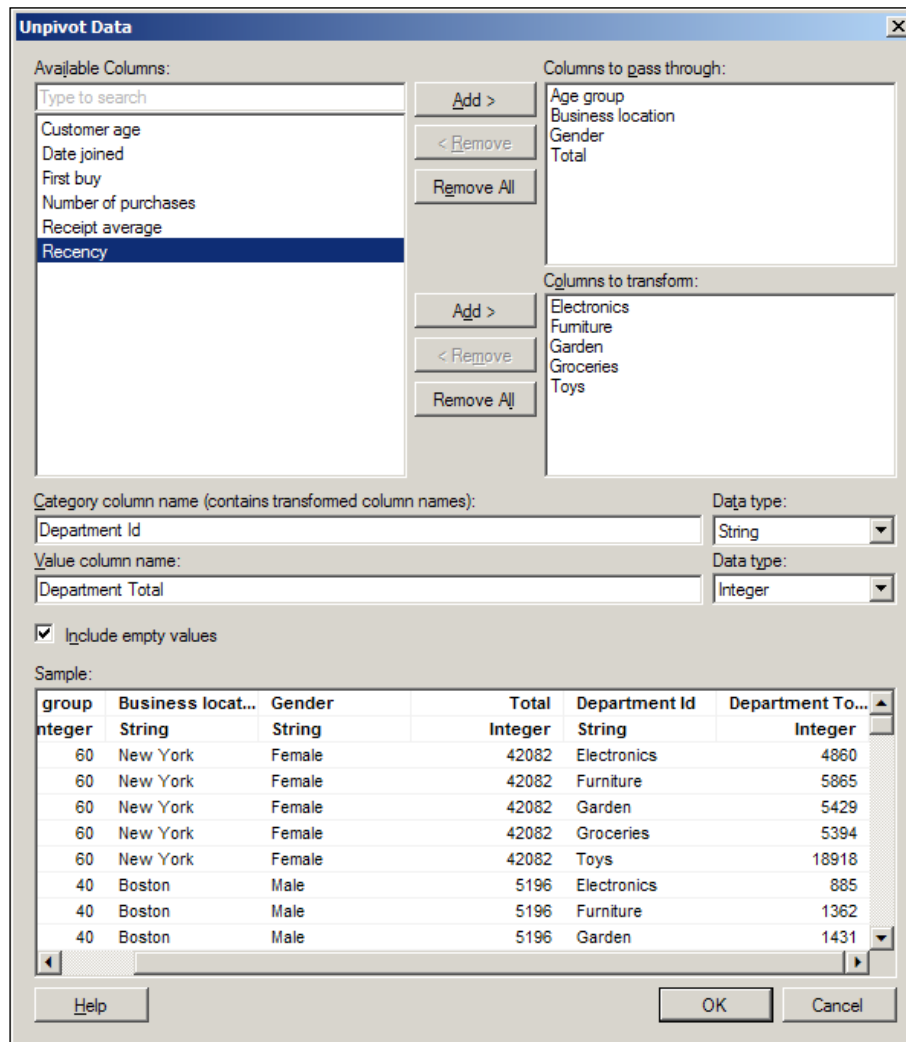
Note the **Department Id** and **Department Total** columns are the new unpivoted columns.

Calculating and replacing a column

Spotfire allows users to replace columns with others whose values are calculated through the use of an **Expression**. This transformation is known as **Calculate and Replace Column**, and as it happens with all the data transformations, it can only occur on data loading.

The **Calculate and Replace Column** dialog box offers users several customizations, such as **Column to replace**, **Available columns** (to use in the expressions), **Functions** (to create the expressions), **Expression** builder, and also various **Formatting** options.

The following screenshot demonstrates the creation of a column named **Date joined span**, which replaces **Date joined** (for the **StoreSales** data). This new column contains the difference between now and **Date joined** in days.
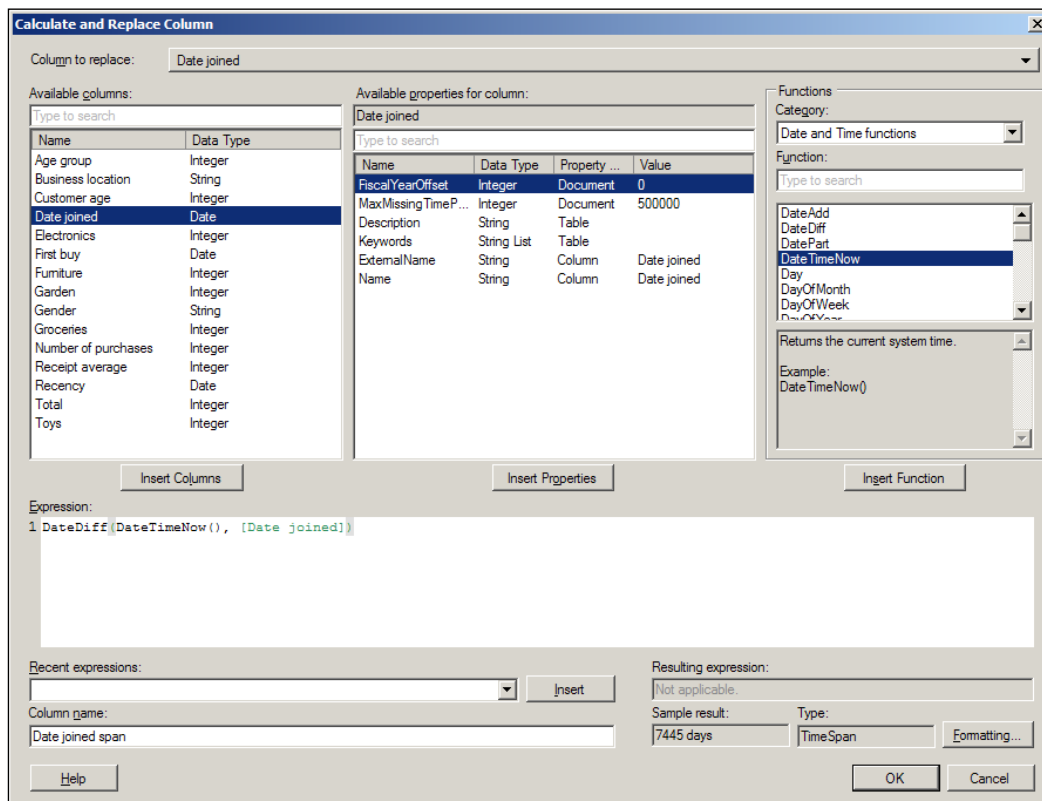
The calculated column will have values as shown in the following screenshot:

Date joined span
7445 days
7445 days
7445 days
7445 days
7447 days
7445 days
7445 days
7431 days
7431 days
7434 days

Calculating a new column

The difference between the transformations **Calculate New Column** and **Calculate and Replace Column** is solely that the first adds a new column to the data table, while the second replaces a column in the data table.

In terms of configuration, both behave in the same way, each having the same rich feature set.

Changing the column name

Change Column Names is one more transformation available for the user. The new column name does not need to be a defined hardcoded value, as it can also be calculated using the **Expression** builder.

The following screenshot displays the available functions for building a new column name.

Change Column Names

Available columns:

Type to search

Business location
Customer age
Date joined
Electronics
First buy
Furniture
Garden
Gender
Groceries
Number of purchases
Receipt average
Recency
Total
Toys

Add >

< Remove

Remove All

Columns to rename ([%C]):

Age group

Function:

Upper

Upper
Right
Length
Left
Lower
Mid
Find
Trim
Substitute
Regular Expression Replace
~=
Repeat
Concatenate
NameEncode
NameDecode

Insert >

Expression:

60
40
50

Help   OK   Cancel

Changing the column data type

The default data type of a loaded column can also be manipulated: `Date` columns can be loaded as `Strings`, `Integer` columns can be loaded as `Currency`, and so on.

The formatting of the columns is also possible, as the application provides a very rich set of formats. The following screenshot lists the available formatting options of a **Date** category.

Besides some predefined possibilities, the user can inclusively define a format from scratch using the **Custom** format.
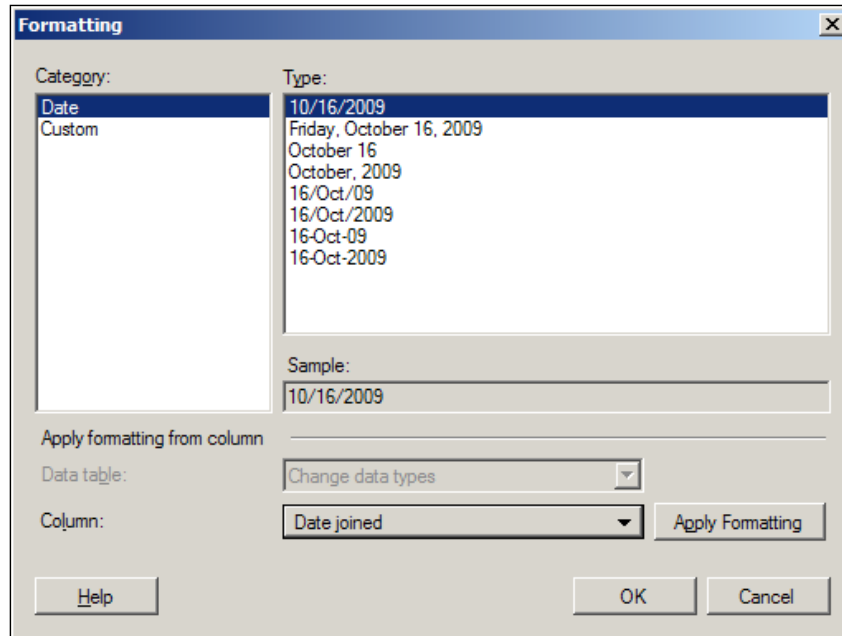
Replacing the data table

In some scenarios, the users might want to keep one or more visualizations, but replace the underlying data with a new set. An example of such a case could be a consulting company that analyzes their quarterly results through a set of custom visualizations. It is expected that at the end of every quarter, they will want to load the quarterly data into Spotfire and run the analysis on the new data.

For such cases, Spotfire Professional offers the possibility of replacing the data table content. Users just need to provide a new set of data which has the same columns with the same data types as the original data.

To experiment with this function, navigate to the menu **File | Replace Data Table**. A dialog box named **Replace Data Table – Select Data Table** will be displayed, as shown in the following screenshot:
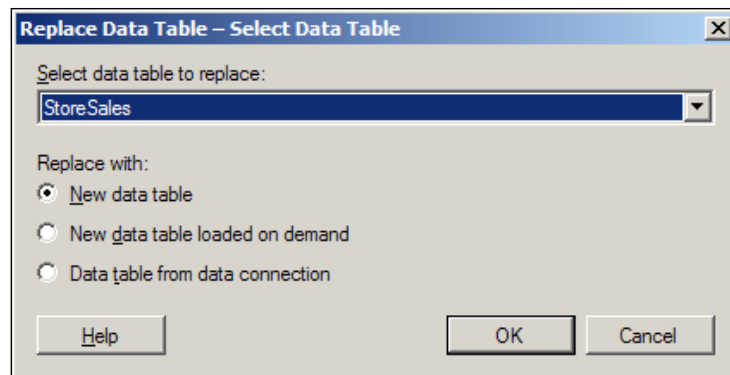
In the **Select data table to replace:** dropdown, users are able to select the data table in which they want to load the new data, while the **Replace with:** radio button provides extra loading configuration options.

Column matches

So far we have seen the creation of visualizations from a single data table. This is the most common task; however, the tool is not limited to it. If this limitation existed, it might lead the users to create massive merged tables, full of repeating data, instead of the traditional relational data models.

The Spotfire feature **Column Matches**, enables the matching of columns between distinct data tables, thus fostering the use of multiple data tables in a single visualization. In this aspect, Spotfire is extremely friendly, as it makes two matching modes available: automatic and manual.

The automatic mode

Spotfire considers that the two columns match when the names and data types match. This is not always necessarily true, however; matched columns are not necessarily connected and hence do not need to be leveraged.

To access the analysis's **Column Matches**, navigate to menu **Edit | Data Table Properties** and a dialog box will be shown. Accessing the tab **Column Matches** displays the matched columns for each of the selectable data tables.

The following screenshot lists the automatic matches for the **StoreSales** data table:
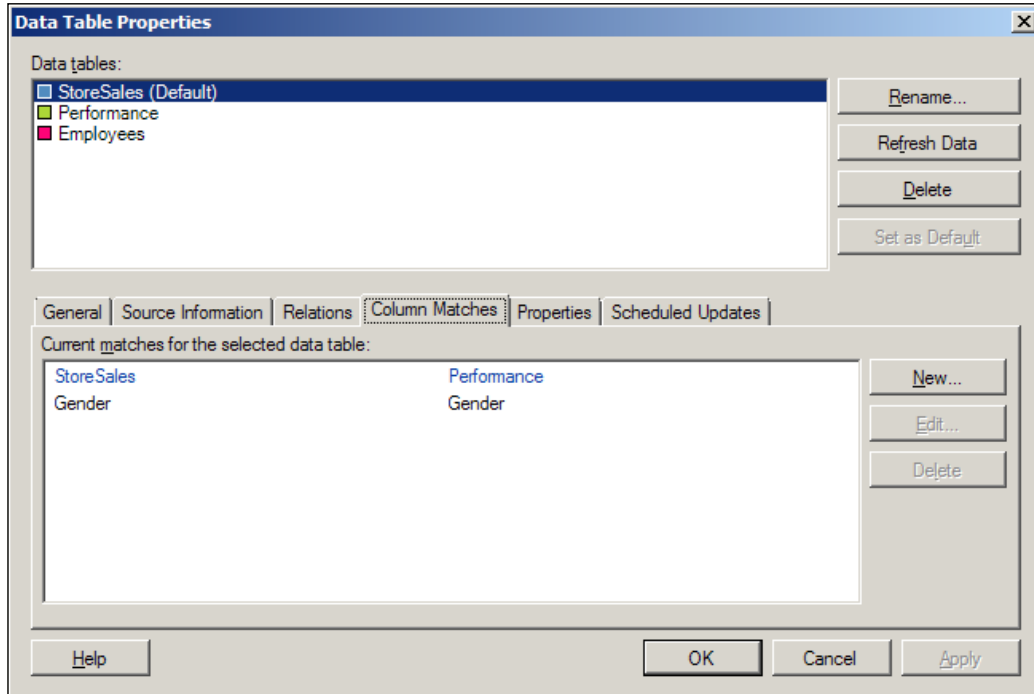
This match is actually incorrect as the data tables are unrelated, and therefore we should delete it. To delete it, select the match in the current matches list and click on the **Delete** button next to the list. Now, our data tables are no longer connected.

The manual mode

If Spotfire does not identify the columns as matching, the user can force this, by defining new matches.

To do so, users should click on the **New** button in the **Data Table Properties** dialog box (the **Column Matches** tab), and select both the data table and the columns to match as shown in the following screenshot:
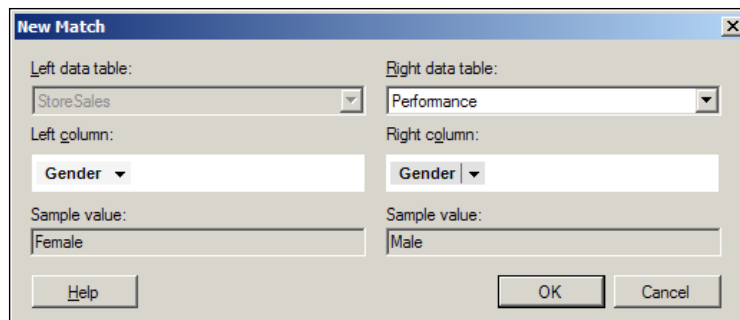
Relations

Relations are the ties between columns of two distinct data tables. They are similar in principle to **Column Matches** but their function is different; their aim is to propagate filtering and marking between visualizations.

For instance, if two visualizations have defined a relation between two columns of their data tables, marking data in the first visualization will mark related records in the second visualization.

The columns chosen to participate in a relation should be identifier columns in both the data tables, and more than one column can be related between two data tables.

To exemplify the creation of a relation between two data tables, we will first need to import new data. Add the table JOB_HISTORY (from the Oracle XE database and user hr) as an analysis data table named **JobHistory**. Before closing the **Add Data Tables** dialog box, make sure you click on the **Manage Relations** button to add a relation to data table **Employees**. Refer to the following screenshot for the **Manage Relations** dialog box:
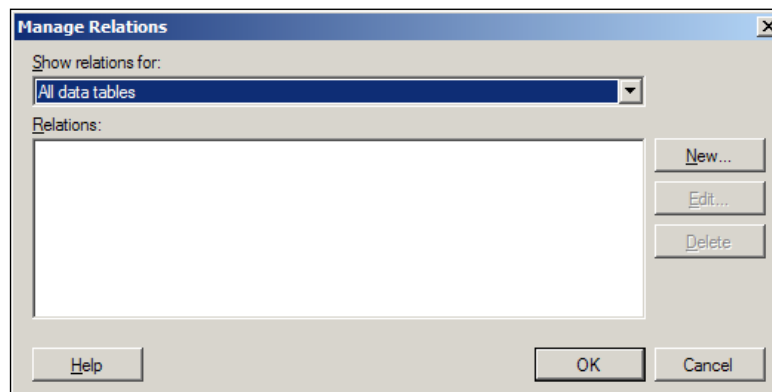
Click on **New** and the **New Relation** dialog box will be opened. Refer to the next screenshot for details:
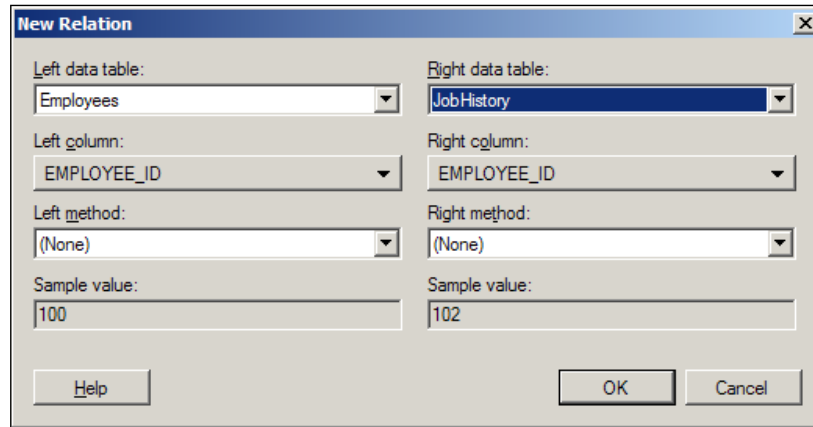
To configure a new relation, perform the following steps:

1. Define the **Employees** data table as **Left data table**: and leave **EMPLOYEE_ID** as **Left column**:

2. Define the **JobHistory** data table as **Right data table**: and leave **EMPLOYEE_ID** as **Right column**:

3. Click on **OK**.

A new relation is now created between `Employees.EMPLOYEE_ID` and `JobHistory.EMPLOYEE_ID`, which is displayed in the **Manage Relations** dialog box shown in the following screenshot:
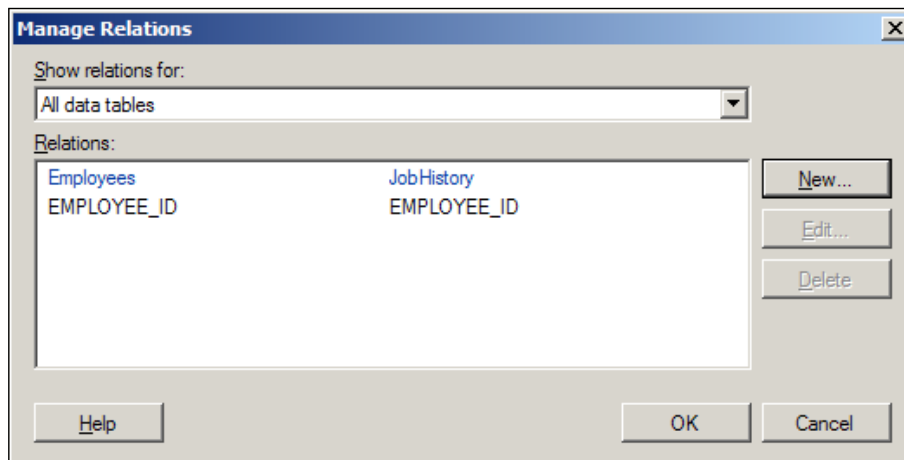
Click on **OK** twice to close the remaining the dialog boxes. You may also close the default scatter plot created for the new **JobHistory** data table.

Going back to the **Data Table Properties** dialog box, and validating **Column Matches** of the data table **Employees**, you can verify that three new **Column Matches** were created between the data tables **Employees** and **JobHistory**: **DEPARTMENT_ID**, **EMPLOYEE_ID**, and **JOB_ID**. Remove **DEPARTMENT_ID** and **JOB_ID**, as only **EMPLOYEE_ID** is a linking key between the data tables.

By deleting the match on the **Employees** data table, the symmetrical match (from **JobHistory** to **Employees**) will also be removed.

The following screenshot displays the **Data Table Properties** dialog box with the final result:
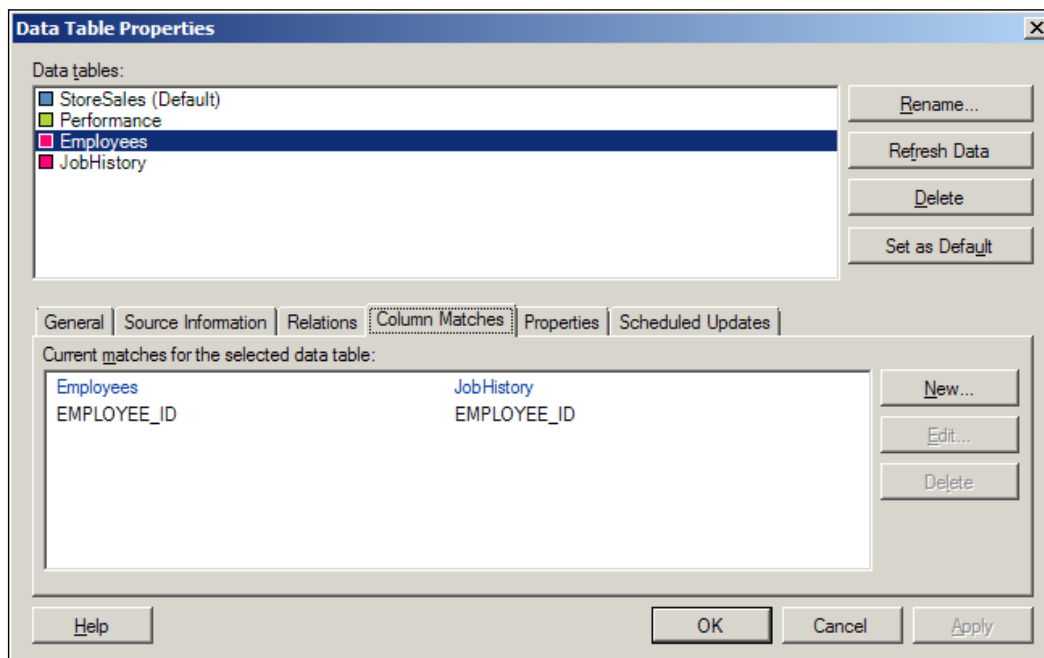
Multiple data tables in one visualization

As stated in the *Column matches* section, the purpose of data table column matching is to enable the use of multiple data tables in a single visualization. In that section, we also demonstrated how to create column matches between data tables.

Using multiple data tables in one visualization is similar to using just a single data table, although in the first we will have the concept of additional data tables besides the already known main data table:

- **Main data table**: It has the standard main data table behavior, defining the rows of an unaggregated visualization

- **Additional data tables**: These are viewed as extra columns, which can be added to visualizations. These columns can only be used in aggregation axes (y axis), not grouping (or categorical) axes (x axis).

As an example of this feature, using the main data table **Employees** and additional data table **JobHistory**, we will represent the total number of employees that are or have been in each department. We will create a bar chart with the list of departments (**DEPARTMENT_ID**) as x axis, and the count of employees as y axis. This count will be retrieved from both the data tables—the main and the additional.

Please be aware that chosen grouping columns (x axis) should belong to both the main and the additional data tables (they have to be matched columns) in order for the aggregation to work. Due to this limitation and because we want to use **DEPARTMENT_ID** as x axis, we must first match this column of data table **Employees** with the matching column **DEPARTMENT_ID** of data table **JobHistory**.

As specified earlier, column matching is done using the data table **Properties** editor from the **Edit menu**. In the editing tool, matching of columns can be done in the **Column Matches** tab.

To add a variable to a chart from a data table other than the main one, users have a drop-down list available in the variable selection pop up. Only data tables that have matched columns to the main data table will be listed in this dropdown.

The following screenshot presents an example of this pop up:
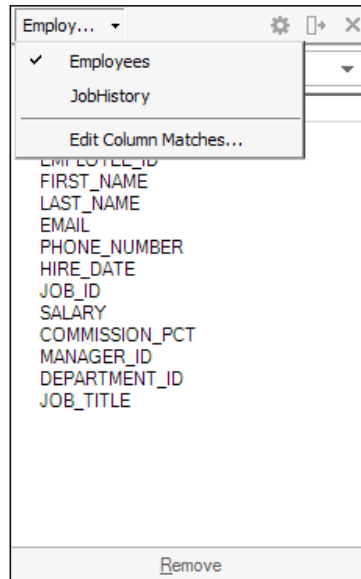
This example belongs to a visualization with the main data table **Employees**. Since this data table has matched columns solely in the data table **JobHistory**, only these two are listed.

To create an example graph for multiple data tables, open your analysis project and execute the following steps:

1. Create a new page named **Employees - Multiple DT**.
2. Add a bar chart to the new page.
3. Define **Employees** as the chart main data table.
4. Select **DEPARTMENT_ID** as the x axis variable of the chart.
5. Select **Count(DEPARTMENT_ID)** as the y axis variable of the chart.
6. Select **Count(DEPARTMENT_ID)** from the data table **JobHistory** as a second variable of the y axis of the chart.

The resulting graph will look similar to the one presented in the following screenshot:
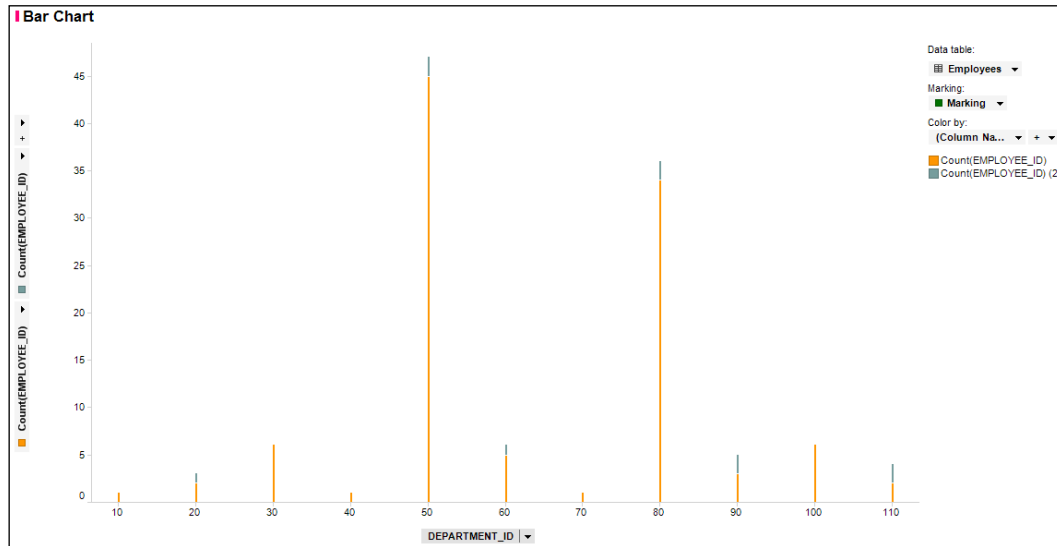
We can see only departments with IDs **20**, **50**, **60**, **80**, **90** and **110** had former employees in the past (the coloring of data is different for records for each of the underlying data tables).

Save the analysis project.

The visualization layout

Several visualizations can be placed in a single page. The disposition of these is customizable, and the size of each can also be manipulated.

Each visualization that is added to a page will be placed on the top, above any already existing visualization. The layout can then be manipulated by the user and visualizations can be shifted around and arranged in a user chosen fashion.

Spotfire offers the possibility for visualizations to be arranged vertically, horizontally, or in a mixed grid. Just be aware that the screen size limits the available area.

Visualizations can be shifted around, just by dragging-and-dropping. To grab a visualization, select (with the mouse) the visualization bar which contains the visualization name. From there, users can drag-and-drop it anywhere on the screen. Moving visualizations can be done at any point in time. When dragging a visualization around a page, grayed drop areas will appear to guide the users to drop the visualization in a proper location.

Spotfire offers the possibility of auto-arranging the visualizations, where a user can simply select one of the four predefined arrangement types. To access this feature, open the menu **View | Arrange Visualizations**. The following screenshot presents the list of available possibilities:
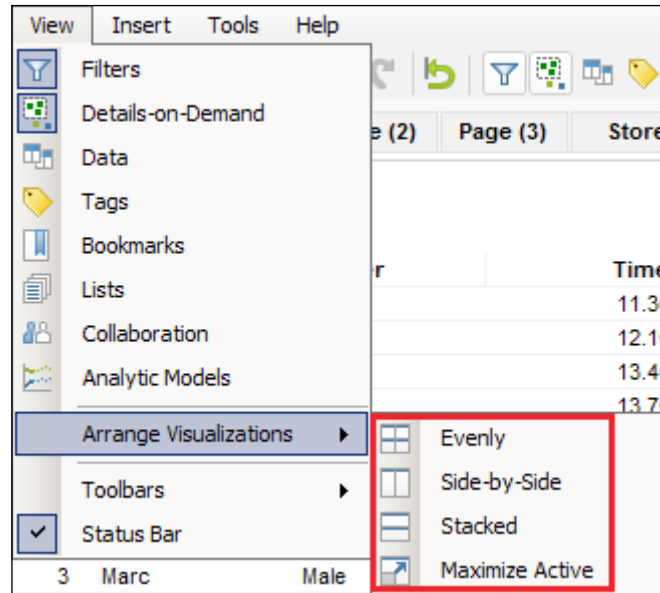
The possibilities are described as follows:

- **Evenly**: It distributes the visualizations evenly throughout the page.
- **Side-by-side**: It distributes the visualizations horizontally (side-by-side).
- **Stacked**: It distributes the visualizations vertically stacked.
- **Maximize Active**: It expands the currently selected visualization, minimizing all remaining ones. The minimized visualizations are placed at the bottom of the page, and only their title bar will be visible.

Pages

Pages foster the creation of Spotfire analysis that includes many visualizations functionally segregated and distributed across different containers. Users can create pages organized in a logical sequence to better shape their analysis.

> Pages do not limit in any way the usage of data tables. One page can contain visualizations of many sources.

Spotfire allows navigation between pages in three different ways: using titled tabs, step-by-step, and history arrows.

- **Titled tabs**: This is the mode used by us so far. Pages are represented by selectable tabs on top of the screen, just below the toolbar. Also, there is a drop-down arrow available on the right-hand side of the tabs bar, which allows the user to select to switch tab. The following screenshot displays the tab bar with the drop-down selection option.

- **Step-by-step**: In step-by-step mode, users will see numbered links as the navigator between pages (instead of tabs). The following screenshot exemplifies this:
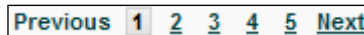
- **History arrows**: History arrows are similar to step-by-step links, but no numbered links are available. Only history arrows will be available to users. Navigation has to be enabled using text areas or dynamic items on graphical tables. More information can be found in Spotfire Professional's *Help Topics* documentation.

> To toggle between navigation modes, right-click on the **Pages** toolbar and select **Page Navigation**.

Libraries

Libraries allow users to share analysis using a central repository. Analysis assets created by a user to be shared include not only the pages with their visualizations, but also information links, data tables, and so on.

Spotfire Professional offers a library management tool, which can be initiated by accessing the menu **Tools | Library Administration,** as shown in the following screenshot:
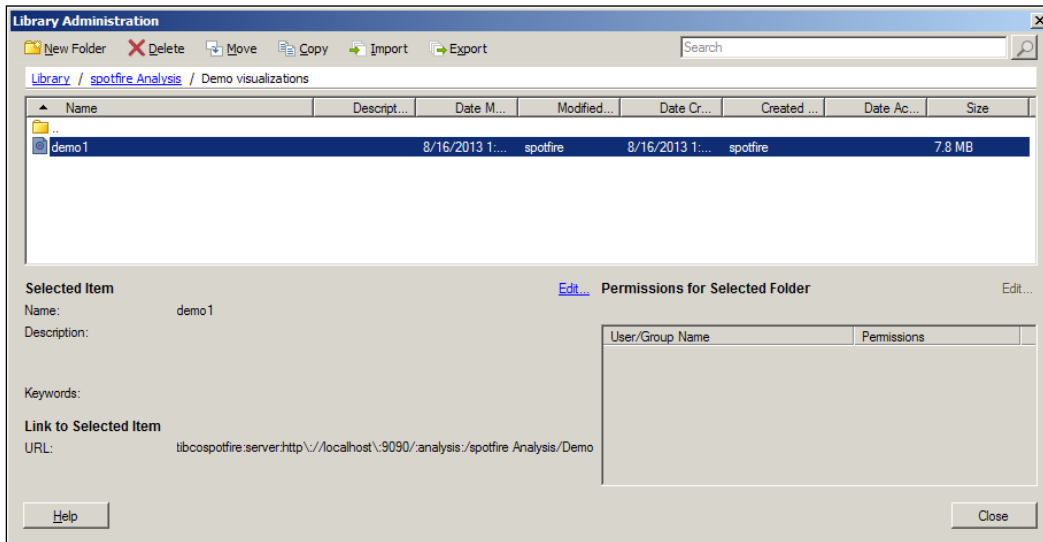
Administration tasks include folder structure creation and management, importing and exporting library files, and permission setting.

To save an analysis project to the library, users can use the menu **File | Save (As) | Library Item** or **File | Save Copy (As) | Library Item**. There, it will be possible to define the folder where the analysis will be stored, and it will also be possible to choose other configuration values such as the embedding of data in the analysis. All shared analysis will be available not only for Spotfire Professional users, but also for Spotfire Server Web Player clients.

> Embedding data in libraries has the same behavior as embedding data in an analysis file. It means that the source data of the data tables is stored with the analysis itself and is always available for loading.
>
> Not embedding data means that the source data of the data tables has to be available when an analysis is loaded. If the data is not available, users will fail to load the analysis. This configuration has the advantage of keeping analysis files smaller in size.

To create a library containing the analysis created throughout this book, perform the following steps:

1. Access the **File** menu, and select the option **Save As** | **Library Item**.
2. Create a new folder named `spotfire Analysis`, by clicking on the button **New Folder**.
3. Double-click on the newly created folder and inside it create a new folder named `Demo visualizations`.
4. Once again, double-click on the newly created folder to enter it.
5. In the **name** input box, name the library as `demo1`. Click on the **Next** button.
6. At this point, a description and several keywords can be associated with our library. We will leave both blank. Click on **Next**.
7. In this last configuration step, the user can set the data tables configuration for the library (embed or not the data). This is done by using the options **Edit** and **Override these settings and embed all in-memory data**. We will tick the checkbox **Override these settings** and embed all in-memory data. Click on **Finish**.
8. The following pop up will be presented. You can close it by clicking on the **Close** button.
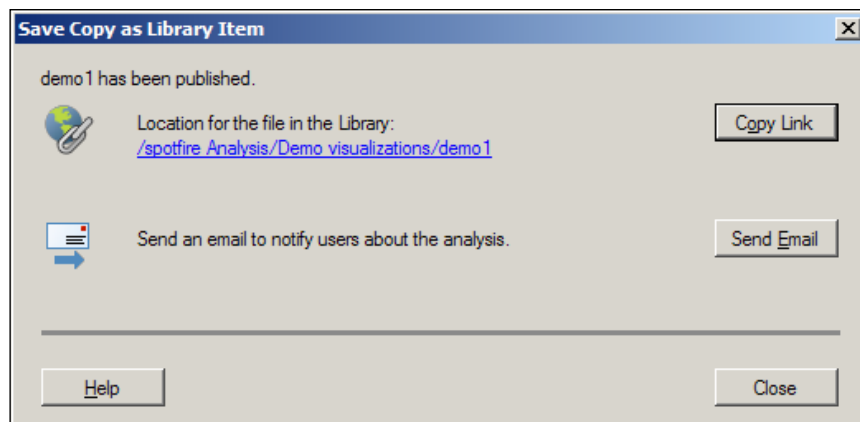
Web Player Server

Web Player Server is the Spotfire platform component which enables web browser access to the shared analysis projects stored in the library. This product consists of a web application, which is integrated with the Spotfire Server, and can read the analysis deployed in it.

Because it supports solely thin clients, Web Player Server makes available only a subset of the features of Spotfire Professional. Users can only read analyses; they cannot create or modify them.

The following features are available in the tool:

- Visualizations panel
- Marking of data for drill down and for display of details visualization
- **Filters** panel without filtering schemes and filter customization
- **Details-on-demand** panel, in which the data presented comes from standard marking
- Libraries

To access the Web Player, open one of the Spotfire Web Player supported web browsers, and navigate to `http://localhost/SpotfireWeb/`.

After logging in with the administration `spotfire` user, the following library screen will be presented:
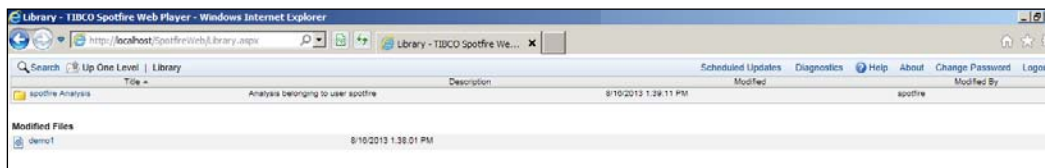
Here are listed the libraries that are visible to the logged-in user. The library **demo1** was saved as an example for the *Libraries* section of this chapter.

To open the library, just click on it. If we click on **demo1**, we will be presented with the following screen:
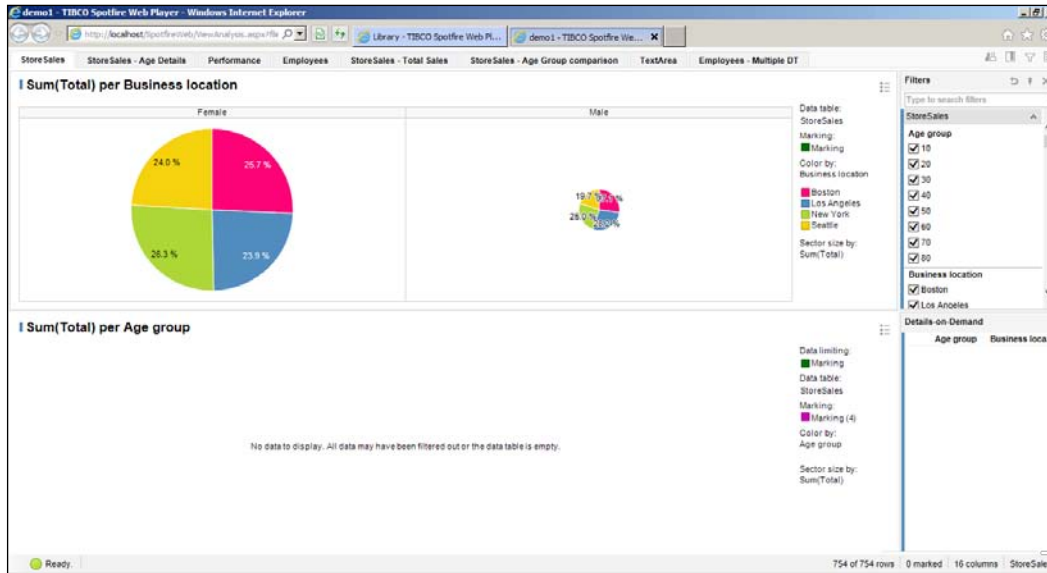
Computational analytics

The Spotfire platform has a set of computational analytics functionalities that are distributed among several statistical tools.

One of these functionalities is **Predictive modeling**, which is the mostly used form of data mining, and allows the prediction of the future by forecasting the probabilities and trends. This statistical model is widely used in IT, for instance in e-mail spam filters, to identify the probability of a message being spam.

In Spotfire, predictive modeling is composed by three distinct steps:

- **Fitting the model**: In this step, the model options are chosen, such as statistical method to use, source data table, column to predict, predictor columns, and so on. The output of this step is a model page.

- **Evaluating the model** :The model page is evaluated against new data, into an evaluation page.

- **Predicting the model**: Additional prediction data becomes available, which can be added to the source data table.

Two different models are supported for predictive modeling:

- **Regression modeling**: It offers both the **linear regression** method and the **regression tree** method.
- **Classification modeling**: It offers both the **logistic regression** method and the **classification tree** method.

Each model has a different tool, which can be started by accessing the menu tools.

This book will focus on regression modeling or the linear regression method. For extra information regarding predictive modeling and the other models, refer to Spotfire Professional's *Help Topics* documentation on the page *Predictive Modeling*.

Regression modeling/the linear regression model

In this model, the scores on one variable are predicted from the scores on a second variable. The variable we are predicting is known as **Response** and the variable we are basing our predictions is known as **Predictor**. When there is only one predictor variable, the prediction method is called **Simple regression**. And, when there are many predictor variables, the prediction method is called **Multiple regression**.

Before attempting to use a linear regression model on the existing data, the user should first determine whether or not there is a relationship between the response and the predictor variables.

In the visualization of the page **StoreSales - Total Sales**, we created a scatter plot that correlates the **StoreSales** variable **Number of purchases** with the variable **Total**. Both the visualization's correlation value of 68 percent and the **Straight Line Fit** direction indicate that there is in fact a positive correlation between the variables.

The **StoreSales** data is therefore suitable for the creation of an example of the linear regression model, but since there is only one set of data available, we will need to break it into two; we will separate the data into two files according to the year in the **Recency** column – **1994** or **1995**. This can be achieved using MS Excel or any other tool that will help in the easy location and deletion of lines with certain content. In the new data files, columns **Customer age**, **Date joined**, **First buy,** and **Recency** can also be removed as they are not necessary for our example.

Import both the files and create two data tables named **StoreSales_1994** (for the **1994** file) and **StoreSales_1995** (for the **1995** file). If you kept the columns **Customer age**, **Date joined**, **First buy**, and **Recency**, beware of the date-time formats while importing the files. When done, you can close the default created scatter plots and delete their respective pages.

Our objective with this predictive model will be to predict the sales **Total**, based on the **Number of purchases**. This will be a simple linear regression.

According to the three steps of predictive modeling, we must first fit the model. To do so, we will access the **Tools** menu and select **Regression Modeling**. This action will open the regression modeling configuration tool.

Our model will be named **StoreSalesRegressionModel**, and the remaining configuration should be the following:

- **Comment**: `Predict Sales Total based on Number of purchases`
- **Model Method**: **Linear Regression**
- **Data table**: **StoreSales_1994**
- The **General** tab | **Response** column: **Total**
- The **General** tab | **Predictor** column: **Number of purchases**

Click on the **Add** button.

The filled configuration should look similar to the following screenshot:

Click on **OK**. After a few seconds, a page named **StoreSalesRegressionModel** will be generated, containing the information about our model. The following screenshot shows an example of such a page:
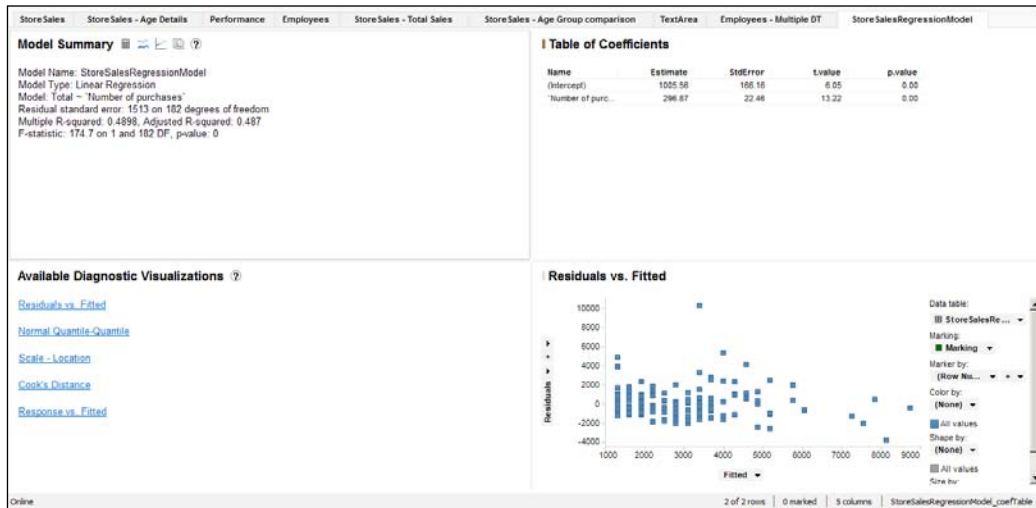
The generated model page contains 4 areas:

- **Model Summary**: It lists the name of the model, the model type, and the model formula
- **Table of Coefficients**: It lists the model coefficients for the regression model
- **Available Diagnostic Visualizations**: It lists the available visualizations of the predictive model, which help to assert its validity
- Visualizations area: It is the area where the diagnostic visualizations are presented

Proceeding to the second step of our predictive model, we will now evaluate the model. To do so, click on the **Evaluate model** button from the **Model Summary**'s toolbar. Refer to the following screenshot for details:
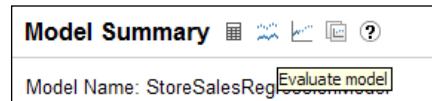
A dialog box will be presented, where we can specify the data table on which we want to run the model. We will run it on **StoreSales_1995**.

Make sure that you have the configuration as shown in the following screenshot:
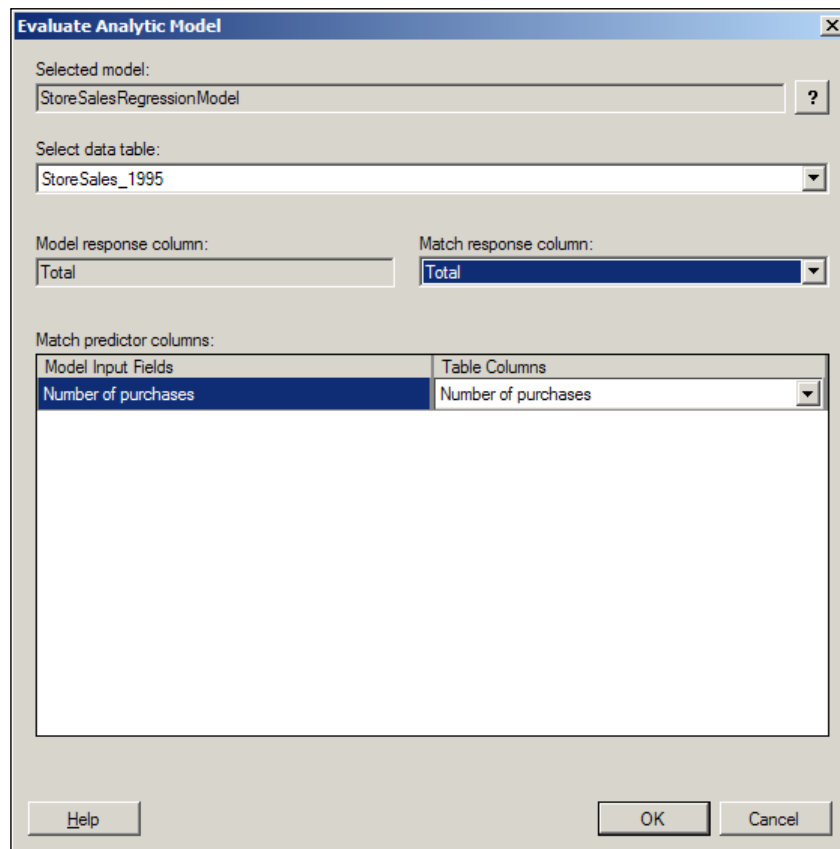
Click on **OK**. After a few seconds, a page named **Evaluation (StoreSalesRegressionModel)** will be generated, containing the information about our evaluation. The following screenshot shows an example of such a page:

The generated model page contains three areas:

- **Evaluation Summary**: It lists name of the model, the data table used in the evaluation, and the model formula

- **Available Diagnostic Visualizations**: It lists the available visualizations of the predictive model, which help to assert its validity

- Visualizations area: It is the area where the diagnostic visualizations are presented

As a final step of the predictive model, we will add the predicted data to the **StoreSales_1994** data table . For that, click on the **Predict from model** button from the **Model Summary**'s toolbar. Refer to the following screenshot for details:
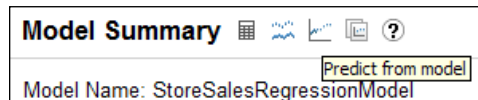
A dialog box will be presented, where we can specify the data table where to add the predicted data. We will choose **StoreSales_1994**. Also, make sure that you have the configuration presented on the following screenshot:
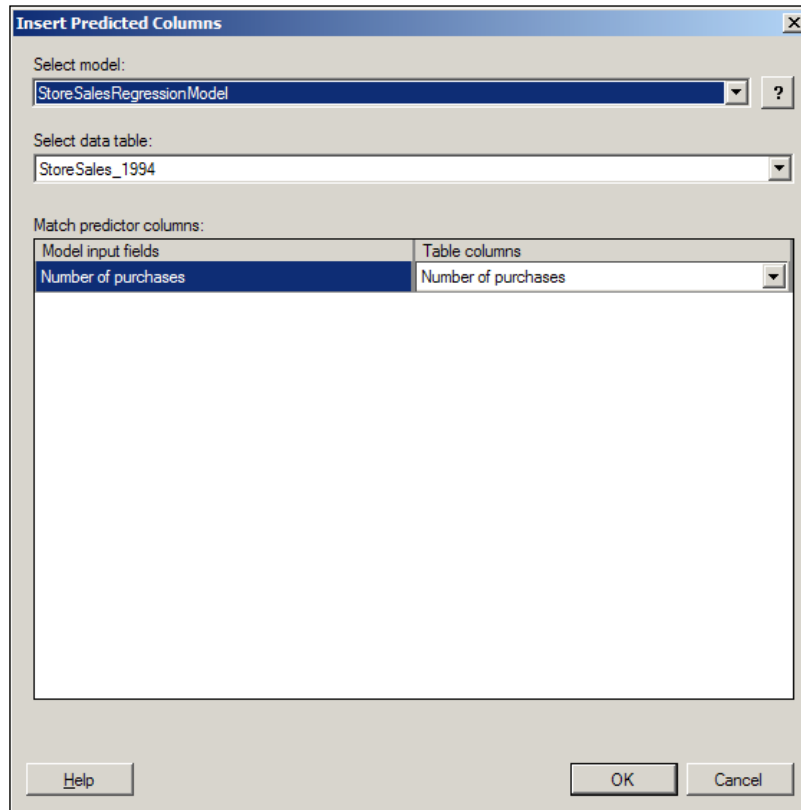
A new column will be added to the **StoreSales_1994** data table named **predicted**.

Please save the analysis project.

Information designer

The **Information designer** is a Spotfire tool to create information links. These links are database queries created from simple elements (columns, filters, procedures, and joins) which are composed into complex queries. The purpose of information links is their later usage as a data table data source. This tool can be found under the **Tools** menu.

After starting **Information Designer**, the first configuration step is to define **Data Source**. This can be done by selecting the **Setup Data Source** option (or by using the **New** dropdown on the top left corner). A configuration as shown in the following screenshot specifies the creation of a connection to our installed database, with user `hr`.
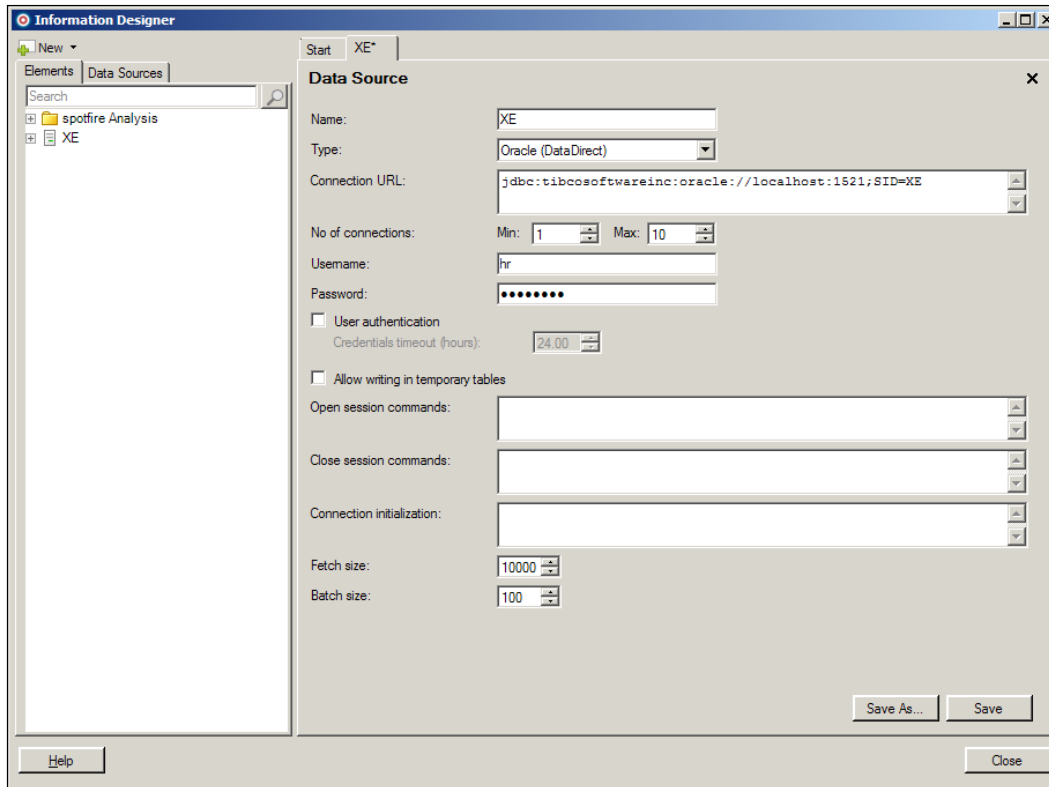
After configuring it, please save it by clicking on the **Save As** button.

After creating a data source, and going back to the **Start** tab, users can then define the **Elements** (columns, filters, and so on) required for the creation of the information link.

Our objective will be to create two **Multiple Columns** elements: one with the columns **DEPARTMENT_ID** and **DEPARTMENT_NAME** from the `HR` table **DEPARTMENTS**; and a second one with the columns **DEPARTMENT_ID**, **FIRST_NAME** and **LAST_NAME** from `HR` table **EMPLOYEES**. The following screen demonstrates the configuration for the first element:
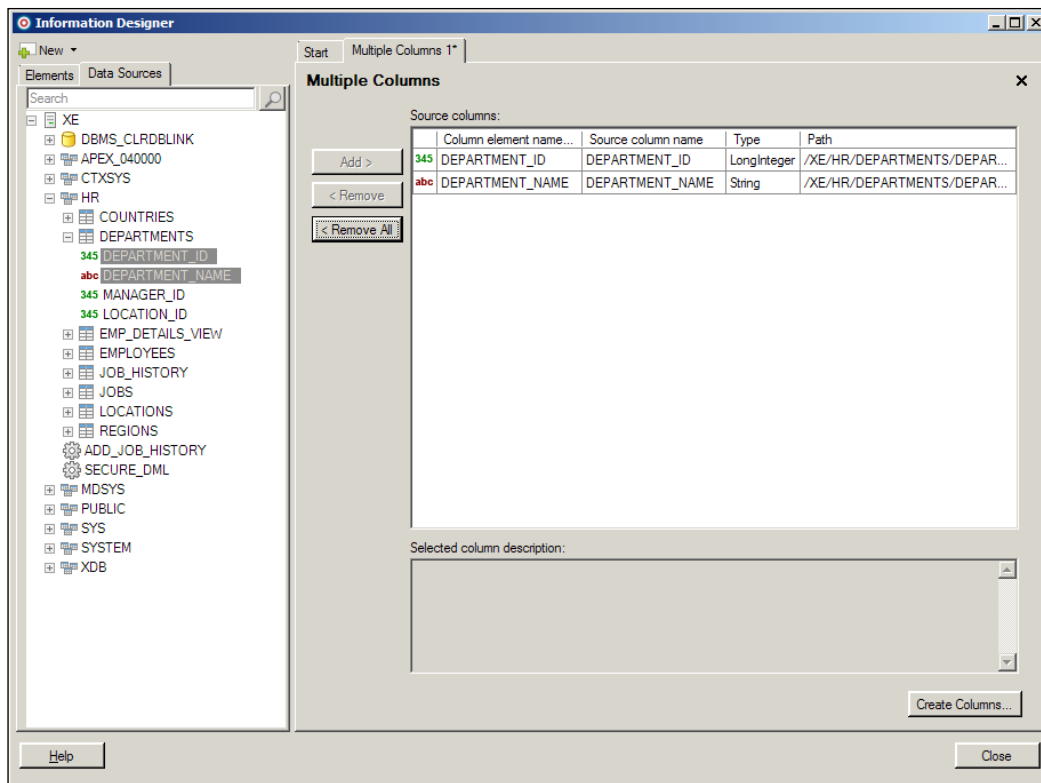
For starting the creation procedure, it is necessary to select the option **Multiple Columns** under the **Start** tab of the **Create Elements** option. Also click on the **Create Columns** button.

Since the columns will be stored as individual elements, it is a good idea to create the following folder structure to store them; refer to the following screenshot:
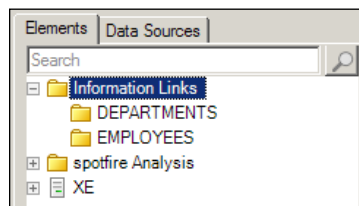
This can be achieved using the **New** drop-down menu.

The result of the creation of both **Multiple Columns** should be similar to the following screenshot:
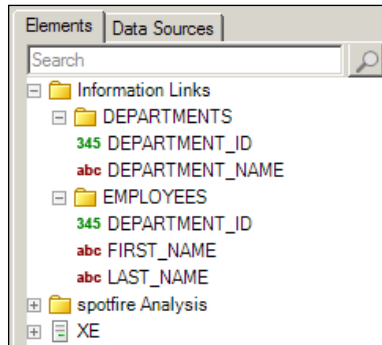
Next, we will be creating a join between both **DEPARTMENT_ID** columns (in both the **DEPARTMENTS** and **EMPLOYEES** tables). The **Create Join** option of the **Start** tab triggers the configuration of joins between the **Data Sources** tables.

To create such a join in the **Join** configuration, both the **DEPARTMENT_ID** columns must be added. The resulting **Join** should be saved under the **Information Links** folder.

Our last step will be the creation of the **Information Link** itself. This procedure can be triggered by clicking on the **Create Information Link** button in the tool's **Start** tab (or by using the **New** dropdown). The link's elements will be **EMPLOYEES | FIRST_NAME, EMPLOYEES | LAST_NAME, EMPLOYEES | DEPARTMENT_ID** and **DEPARTMENTS | DEPARTMENT_NAME**, and **Join path** will be the created join **DEPARTMENTS - EMPLOYEES**.

At this point, the **Information Link**'s configuration should look similar to the following screenshot:
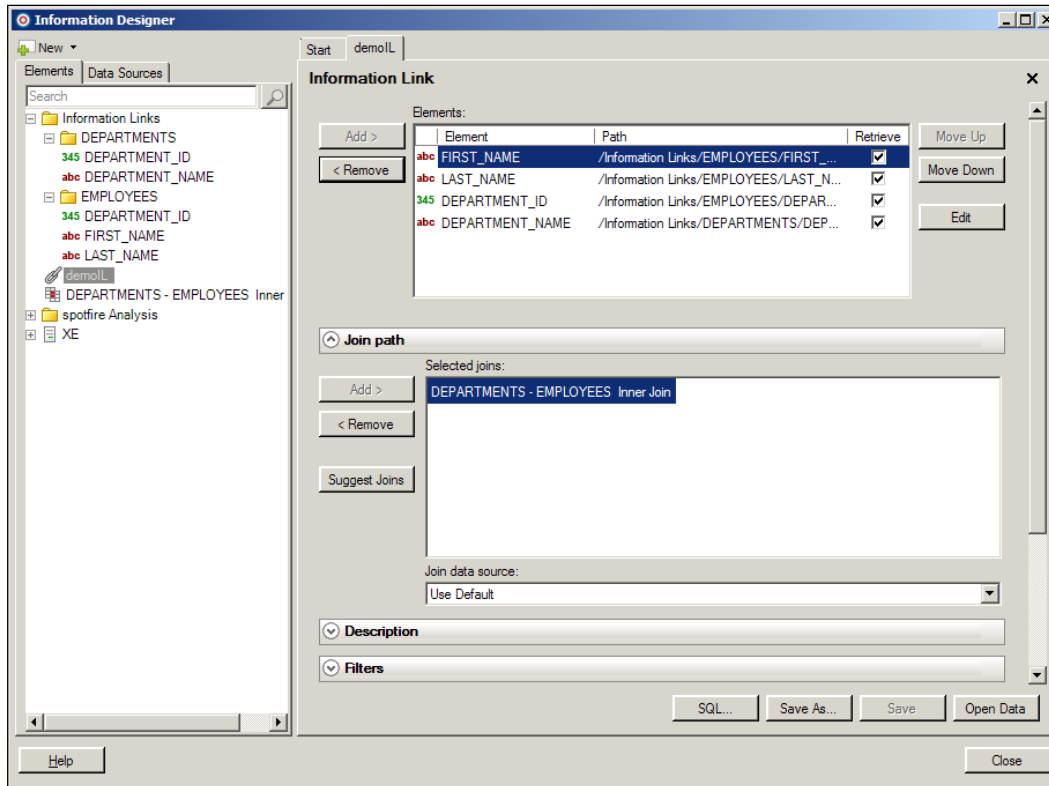
Please save it as **demoIL** under the **Information Links** folder, and close the I**nformation Designer** window.

At this point, a data table can be created, having **demoIL Information Link** as the source. Please do it and feel free to create a visualization with this main data table.

You should also be aware that all the configured **Elements**, **Joins**, and **Information Links** will be available for sharing in the library.

Summary

Together with the previous chapter, this chapter contains the main functionalities a user requires when he/she wants to embrace the Spotfire world. It is, however, not a comprehensive training guide, as the platform can do more than what was presented. I recommend users to consolidate their gained experience and further enrich their knowledge, by reading and studying the tool's documentation.

We augmented our knowledge by studying filtering and drill down on the data, and how to better prepare and transform the loaded data. We also tried out the Spotfire Web Player and described a few similarities between it and Spotfire Professional. Also we covered some of the limitations of this tool, and identified its advantages.

This was the final chapter of the book, and I sincerely hope it served as a good introduction to the world of Analytics and Spotfire. It is my expectation that you will carry on working with the tool and proceed with topics such as 3D scatter plots and map chart visualizations, and detail computational analytics and statistical tools.

Index

Thank you for buying
TIBCO Spotfire for Developers

About Packt Publishing

Packt, pronounced 'packed', published its first book "Mastering phpMyAdmin for Effective MySQL Management" in April 2004 and subsequently continued to specialize in publishing highly focused books on specific technologies and solutions.

Our books and publications share the experiences of your fellow IT professionals in adapting and customizing today's systems, applications, and frameworks. Our solution based books give you the knowledge and power to customize the software and technologies you're using to get the job done. Packt books are more specific and less general than the IT books you have seen in the past. Our unique business model allows us to bring you more focused information, giving you more of what you need to know, and less of what you don't.

Packt is a modern, yet unique publishing company, which focuses on producing quality, cutting-edge books for communities of developers, administrators, and newbies alike. For more information, please visit our website: `www.packtpub.com`.

About Packt Enterprise

In 2010, Packt launched two new brands, Packt Enterprise and Packt Open Source, in order to continue its focus on specialization. This book is part of the Packt Enterprise brand, home to books published on enterprise software – software created by major vendors, including (but not limited to) IBM, Microsoft and Oracle, often for use in other corporations. Its titles will offer information relevant to a range of users of this software, including administrators, developers, architects, and end users.

Writing for Packt

We welcome all inquiries from people who are interested in authoring. Book proposals should be sent to author@packtpub.com. If your book idea is still at an early stage and you would like to discuss it first before writing a formal book proposal, contact us; one of our commissioning editors will get in touch with you.

We're not just looking for published authors; if you have strong technical skills but no writing experience, our experienced editors can help you develop a writing career, or simply get some additional reward for your expertise.
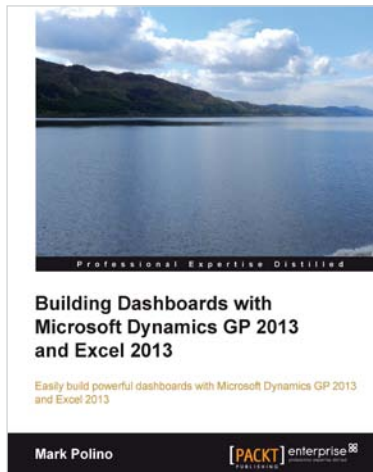
Building Dashboards with Microsoft Dynamics GP 2013 and Excel 2013

ISBN: 978-1-84968-906-9   Paperback: 268 pages

Easily build powerful dashboards with Microsoft Dynamics GP 2013 and Excel 2013

1. Build a dashboard using Excel 2013 with information from Microsoft Dynamics GP 2013

2. Make Excel a true business intelligence tool with charts, sparklines, slicers, and more

3. Utilize PowerPivot's full potential to create even more complex dashboards

IBM Cognos Business Intelligence

ISBN: 978-1-84968-356-2   Paperback: 318 pages

Discover the practical approach to BI with IBM Cognos Business Intelligence

1. Learn how to better administer your IBM Cognos 10 environment in order to improve productivity and efficiency.

2. Empower your business with the latest Business Intelligence (BI) tools.

3. Discover advanced tools and knowledge that can greatly improve daily tasks and analysis.

Please check **www.PacktPub.com** for information on our titles

Creating Data Models with PowerPivot How-to

ISBN: 978-1-84968-956-4   Paperback: 58 pages

Build better business intelligence with this practical guide to creating Excel data models with PowerPivot

1. Learn something new in an Instant! A short, fast, focused guide delivering immediate results

2. Detailed, step-by-step interactive tutorial guide to learning PowerPivot

3. Carefully organized topics for users of all levels

4. Learn how to make your data accessible and attractive
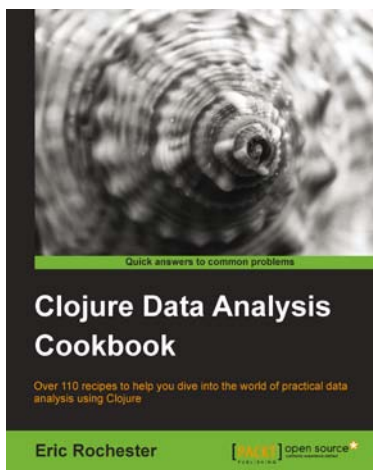
Clojure Data Analysis Cookbook

ISBN: 978-1-78216-264-3   Paperback: 342 pages

Over 110 recipes to help you dive into the world of practical data analysis using Clojure

1. Get a handle on the torrent of data the modern Internet has created

2. Recipes for every stage from collection to analysis

3. A practical approach to analyzing data to help you make informed decisions

Please check **www.PacktPub.com** for information on our titles