TIBCO Spotfire: A Comprehensive Primer

Second Edition

Building enterprise-grade data analytics and visualization solutions

Andrew Berridge and Michael Phillips

Packt
www.packt.com

TIBCO Spotfire: A Comprehensive Primer *Second Edition*

Building enterprise-grade data analytics and visualization solutions

Andrew Berridge
Michael Phillips

Packt>

BIRMINGHAM - MUMBAI

TIBCO Spotfire: A Comprehensive Primer
Second Edition

Copyright © 2019 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the authors, nor Packt Publishing or its dealers and distributors, will be held liable for any damages caused or alleged to have been caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

Commissioning Editor: Amey Varangaonkar
Acquisition Editor: Meeta Rajani
Content Development Editor: Princia Dsouza
Technical Editor: Sayali Thanekar
Copy Editor: Safis Editing
Project Coordinator: Nusaiba Ansari
Proofreader: Safis Editing
Indexer: Manju Arasan
Graphics: Jisha Chirayil
Production Coordinator: Arvindkumar Gupta

First published: February 2015
Second edition: April 2019

Production reference: 1300419

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham
B3 2PB, UK.

ISBN 978-1-78712-132-4

www.packtpub.com

To my wife, Jenna, and my three children for their undying love and their support for my career.
To my father, John, for teaching me so many life lessons and being the inspiration for my career.
To my managers at TIBCO for enabling and encouraging me to write this book.

To Michael Phillips—the author of the first edition of this book, thank you for letting me carry the torch in this edition!

– Andrew Berridge

mapt.io

Mapt is an online digital library that gives you full access to over 5,000 books and videos, as well as industry leading tools to help you plan your personal development and advance your career. For more information, please visit our website.

Why subscribe?

- Spend less time learning and more time coding with practical eBooks and Videos from over 4,000 industry professionals
- Improve your learning with Skill Plans built especially for you
- Get a free eBook or video every month
- Mapt is fully searchable
- Copy and paste, print, and bookmark content

Packt.com

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.packt.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at customercare@packtpub.com for more details.

At www.packt.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.

Contributors

About the authors

Andrew Berridge is a data scientist at TIBCO Software Ltd. He has 10 years' experience with TIBCO Spotfire—first as an end user and latterly (for nearly 8 years) as a TIBCO employee. He is a world-renowned Spotfire expert and has assisted with selling Spotfire and implementing it for many customers. He has also developed parts of the latest versions of Spotfire, being on the team that implemented the AI-Driven Recommendations feature in Spotfire X. Prior to his role at TIBCO, Andrew was an internal consultant at one of the world's largest pharmaceutical companies.

In his spare time, Andrew restores classic cars. He is also an orchestral horn player. He is dedicated to his family and enjoys spending time with his wife and children.

*Andrew would like to thank **Peter Shaw**, PhD., staff data scientist at TIBCO Software for additionally reviewing Chapter 9. He would also like to thank **Thomas Blomberg**, product manager at TIBCO Software for reviewing Chapter 10.*
Andrew is very grateful to Michael Phillips, who was the original author of the first edition of this book. He acknowledges the amazing contribution and the legacy of the first edition and thanks Michael for handing the project over to him and for his advice and council during the authoring process.

Michael Phillips is the original author of the first edition of *TIBCO Spotfire: A Comprehensive Primer*. He is an eClinical product innovator specializing in informatics solutions that support the drug development process generally, and clinical risk management in particular. He is a creative analyst with over 14 years' experience in IT and business intelligence and 5 years' experience in clinical informatics. He has a background in medicine and general science publishing, and a PhD in biochemistry (drug metabolism).

About the reviewer

Colin Gray has worked in industries such as pharmaceuticals, environmental, and IT, and has over 15 years of experience in data analysis and informatics. Throughout this time, he has led data analysis and informatics projects, has worked on developing methods to make better use of data, and has had a key focus on how to communicate data better to others. To this aim, he has heavily employed web-based technologies and statistical packages. Having worked with TIBCO Spotfire for many years, he is a keen enthusiast in using Spotfire to further data science in all areas of business.

Packt is searching for authors like you

If you're interested in becoming an author for Packt, please visit `authors.packtpub.com` and apply today. We have worked with thousands of developers and tech professionals, just like you, to help them share their insight with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Table of Contents

Preface 1

Section 1: Section 1: Introducing Spotfire

Chapter 1: Welcome to Spotfire 7
- **Getting started with TIBCO Spotfire** 7
  - Launching Spotfire Analyst 9
  - Logging in to TIBCO Cloud Spotfire 11
  - Downloading Spotfire Cloud Analyst 13
  - Installing the mobile apps and logging in to TIBCO Cloud 15
  - TIBCO Spotfire for macOS 18
  - Logging in to the Spotfire web clients (on-premises) 21
  - Spotfire licenses 21
- **Getting started with loading data** 22
  - Importing Excel spreadsheets into Spotfire 22
- **Introduction to the data panel** 25
  - The Spotfire recommendations engine 28
- **Saving Spotfire files** 33
  - Saving a file in analyst clients 33
  - Saving a file in Spotfire Cloud or Business Author 34
- **Producing a useful interactive dashboard** 35
  - Coloring 36
  - Proportionality with bar charts and pie charts 36
  - Drilling in to the data – details visualizations 43
    - Insights from details visualizations 49
  - Using filters 51
  - Trellising 53
- **Summary** 56

Chapter 2: It's All About the Data 57
- **Technical requirements** 58
- **Understanding the basic row/column structure of a data table** 58
- **Exploring key data types** 60
- **Building a scatter plot with real-world data** 61
  - Gathering population data 61
  - Gathering country and region data 62
  - Combining the data 63
  - What is a scatter plot? 66
  - Building the scatter plot using population growth versus child mortality 67
  - Working with natural hierarchies in data 76

Gaining insight from the scatter plot 82
Displaying information quickly in tabular form 85
Enriching visualizations with color categorization 90
Visualizing hierarchical data using treemaps 94
Summary 100

Chapter 3: Impactful Dashboards! 101
KPI charts 102
Constructing a KPI chart 102
Coloring, sorting, and other customization of the KPI chart 112
Enabling end users to configure the KPI chart interactively 116
Framing your analysis using text areas 117
Spotfire custom expressions 117
Spotfire document properties 119
Spotfire property controls 121
Bringing it all together – interactively configuring the KPI chart 122
Drilling in to the KPI chart 132
Marking in Spotfire 132
Building a details visualization using marking 133
Deep dive – what insights can the dashboard reveal? 139
Publishing the dashboard to Spotfire Web 150
Summary 157

Chapter 4: Sharing Insights and Collaborating with Others 158
Bookmarks 158
Advanced topic – key columns 167
Annotations 170
Conversations 172
Summary 181

Section 2: Section 2: Spotfire In Depth

Chapter 5: Practical Applications of Spotfire Visualizations 183
Bar charts 184
Things to watch out for with bar charts 185
Bar chart summary 190
Combination charts 190
Cross tables 190
Cross table grand totals 191
Grand totals – underlying values versus sum of cell values 192
Underlying values 193
Sum of cell values 193
Cross table summary 195
Scatter plots 195
Line ordering 203
Density plots 205

Other types of visualization with scatter plots 207
Scatter plot summary 208
Line charts 209
Pie charts 212
Box plots 216
How to further interpret a box plot 220
More box plot options 221
Box plot summary 222
Treemaps 223
Waterfall charts 225
Graphical tables 228
Graphical table visualization types 229
Graphical table summary 230
Taking action from KPI charts or graphical tables 230
Other visualization types 232
Summary 232

Chapter 6: The Big Wide World of Spotfire 234
An overview of the Spotfire platform 235
Spotfire server 236
Spotfire database 236
Nodes and services 237
Spotfire Web Player(s) and Business Author 237
Spotfire Automation Services 237
Spotfire Analyst clients and mobile clients 237
Spotfire Statistics Services 238
Spotfire Web Player scalability 238
A quick guide to administration manager 240
Users 241
Groups and licenses 242
Special user groups 245
Preferences 246
Using the library administration interface 247
Folder permissions 247
Import and export 249
Automating tasks using Automation Services 250
Running Automation Services jobs 252
The Spotfire administration console 253
Analytics 254
Library 254
Users and groups 254
Scheduling and routing 255
Nodes and services 260
Deployments and packages 262
Monitoring and diagnostics 265

Automation Services 267
Summary 274

Chapter 7: Source Data is Never Enough 275
Technical requirements 276
Creating metrics using calculated columns 276
Basic metric 277
Dynamic metric 282
Categorizing continuous numerical data using binning functions 285
Slicing and dicing data using hierarchy nodes 288
LastPeriods 290
PreviousPeriod 294
ParallelPeriod 295
NavigatePeriod 297
AllPrevious and AllNext 298
Previous and next 298
Intersect 299
Over methods in calculated columns versus axis expressions 301
Over method summary 302
Other calculations in Spotfire 303
Data manipulation – where, why, and how? 303
Merging (joining) data from multiple sources 304
Tall tables versus wide tables 308
Tall tables 309
Wide tables 310
Transforming data structure through pivots and unpivots 310
Unpivot 311
Pivot 315
Other transformations 316
Summary 317

Chapter 8: The World is Your Visualization 318
Data relations (between tables) 319
Setting up relationships between data tables 320
Configuring marking and filtering between related data tables 324
Mashing up data from different tables in a single visualization 327
Comparing subsets of data 329
Showing/hiding items of data 332
Annotating visualizations with reference lines, fitted curves, and error bars 337
Fitted curves 337
Forecasting 343
Error bars 348
Visualizing categorical information and trends together in combination charts 350

Visualizing complex multidimensional data using heat maps 353
Heat maps 353
Dendrograms 356
Profiling your data using parallel coordinate plots 359
Summary 364

Chapter 9: What's Your Location? 365
Map chart layers 366
Getting started with map charts 366
Coordinate reference systems 373
Using geocoding to position data on a map 373
Geocoding using a zip code 374
Case study – when geocoding doesn't work 379
Feature layers 385
Using a data function to assist with geocoding 391
Adding Web Map Service data to a map chart 400
Creating custom maps using Tile Map Service layers 404
Using the map chart for non-geographic spatial analysis 407
Mapping an airport 408
Process mapping using a map chart 409
TIBCO GeoAnalytics 416
Summary 416

Section 3: Section 3: Databases, Scripting, and Scaling Spotfire

Chapter 10: Information Links and Data Connectors 418
In-memory versus in-database analytics 419
Information links 420
Data sources 421
Column 422
Joins 422
Filters 422
Procedures 422
Information links 423
Loading data from an information link 423
Writing back to a database using an information link 424
Data connectors 428
Loading data from a connector 429
Using data on demand to retrieve raw data 435
Troubleshooting data on demand 442
Custom expressions with in-database data 444
Important points to note with in-database (external) data 446
Connection credentials for Web Player users and Automation Services jobs 447

Saving connections to the Spotfire library 458
Streaming data 459
Summary 462

Chapter 11: Scripting, Advanced Analytics, and Extensions 464
Scripting in Spotfire – the why, how, and what 465
Automating actions using IronPython 465
Customizing the Spotfire user interface using JavaScript and HTML in text areas 480
Spotfire Developer Tools 487
Script trust 491
TIBCO Enterprise Runtime for R and open source R 492
Python in Spotfire 495
Statistica and Spotfire 496
Extending Spotfire 496
The JavaScript API 498
Spotfire Server APIs 498
Summary 499

Chapter 12: Scaling the Infrastructure; Keeping Data up to Date 500
Context-aware load management with scheduling and routing 501
Some definitions 501
Server topology 502
Scaling the Spotfire Server using clustering 503
Spotfire Sites 503
Spotfire Web Player scaling 504
Nodes, services, and resource pools 505
Spotfire Web Player scalability – Scheduling & Routing 513
Routing rules – definitions 515
File 515
Group 515
User 516
Routing rules – practical examples 516
Scenario 1 – manufacturing sensor data analysis 516
Scenario 2 – separating critical functions and everyday operations 520
Verifying the example rules and schedules 523
Checking the schedule 524
Checking the routing rules for the scenarios 524
Triggering an update from an external system 527
Notes on scheduling and routing 527
Summary 528

Chapter 13: Beyond the Horizon 530
Natural language search 531
Canvas styling and theming 531

Exporting from Spotfire 532
Visualization export 532
Exporting data to a file or a library 532
PDF export 533
JavaScript visualizations 535
Alerting in Spotfire 537
TIBCO Data Virtualization 538
TIBCO Data Science 538
KNIME and Spotfire 539
TIBCO StreamBase and Spotfire Data Streams 540
TIBCO Spotfire Cloud Enterprise 541
Summary 541
Other Books You May Enjoy 542
Index 545

Preface

Welcome to TIBCO Spotfire: A Comprehensive Primer—Second Edition—Building enterprise-grade data analytics and visualization solutions. This book will introduce you to Spotfire and how to build analysis and visualization solutions using Spotfire. You'll learn how to determine what is going on in your data and then how to drill into it to determine why something is happening! You'll find out how to get to data-driven insights fast, with the help of real-world examples and visualizations. The book doesn't just cover the basics—if you're a seasoned user of Spotfire, there are many advanced topics that will extend and enhance your knowledge.

Who this book is for

This book is for those who wish to learn about how to use Spotfire and how to work with analytics in general. It is also for those who would like to evaluate Spotfire to help determine whether it's suitable for their own or their organization's analytics needs. Finally, advanced users will benefit from reading the book—there are many topics, technical hints, and tips that experienced Spotfire users will find very useful. You don't need any prior experience of Spotfire or analytics to get the best out of the book, although some background in how to work with data will be helpful.

What this book covers

`Chapter 1`, *Welcome to Spotfire*, introduces Spotfire, showing the different Spotfire clients and how to get started with them. It jumps right into analytics with a real-world example of analyzing data from the Titanic disaster.

`Chapter 2`, *It's All About the Data*, covers how to work with data in Spotfire by way of building a scatter plot showing world population growth. Along the way, the chapter discusses how to build detailed visualizations to gain additional insights.

`Chapter 3`, *Impactful Dashboards!*, shows how to construct attention-grabbing dashboards using KPI charts.

`Chapter 4`, *Sharing Insights and Collaborating with Others*, discusses how to share insights with others and how to use Spotfire's collaboration features.

`Chapter 5`, *Practical Applications of Spotfire Visualizations*, is a tour through some of the most frequently used visualizations in Spotfire's toolbox. It gives useful hints and tips as to how to use visualizations to their best effect and details common pitfalls to avoid.

`Chapter 6`, *The Big Wide World of Spotfire*, introduces the Spotfire platform. It covers scaling in brief and shows some of Spotfire's administration tools. It also introduces Automation Services.

`Chapter 7`, *Source Data is Never Enough*, covers data manipulation in Spotfire—creating calculated columns, working with custom expressions, and transforming data.

`Chapter 8`, *The World is Your Visualization*, rounds off what has been learned so far. It covers data relationships, adding smoothing and forecasting lines, error bars, and working with Spotfire's subsets and show/hide features. It finishes off by covering the final two visualization types that haven't yet been discussed.

`Chapter 9`, *What's Your Location?*, covers map charts and geo-analytics in Spotfire. It also introduces the concept of data functions and TIBCO Enterprise Runtime for R.

`Chapter 10`, *Information Links and Data Connectors*, shows you how to work with information links and data connectors in Spotfire. These features are important if you're working with anything other than flat files. In-database versus in-memory data is discussed, along with all the tools you'll need to work with both. The chapter also covers streaming (live) data.

`Chapter 11`, *Scripting, Advanced Analytics, and Extensions*, introduces how to work with IronPython and JavaScript in Spotfire. It covers how to work with the many and varied data science platforms that integrate with Spotfire and explains how to get started with developing custom extensions.

`Chapter 12`, *Scaling the Infrastructure; Keeping Data up to Date*, is designed primarily for Spotfire administrators. It shows how to scale up the Spotfire infrastructure to support as many users as you need. It also details the various mechanisms for keeping data in Spotfire files up to date, showcasing schedules and rules.

`Chapter 13`, *Beyond the Horizon*, covers searching in Spotfire, styling and theming, exporting, conditional alerting, JavaScript visualizations, and some additional products that work nicely with Spotfire.

To get the most out of this book

You can follow a large number of the examples in this book by just using a web browser, since you can sign up for a trial account for TIBCO Cloud Spotfire and use Spotfire in the cloud. However, the more in-depth examples require you to use Spotfire Analyst (more on this in `Chapter 1`, *Welcome to Spotfire*). You can download a version of Spotfire Analyst once you have signed up for a Cloud account. You'll need a reasonably modern PC running Microsoft Windows 7 or later. It should be a 64-bit computer system, as 32-bit systems are only suitable for analyzing small datasets or working with in-database analytics.

You can use Spotfire on a Mac computer too, but the Mac client isn't as fully-featured as the PC-based clients.

This book is specifically targeted toward Spotfire version 10.0.0 and later. You can still take advantage of the book, even if you're using Spotfire version 7, but just be aware that the menu selections won't correspond exactly, and some new functionality was released with Spotfire X (Spotfire 10) that wasn't available in previous versions. Specifically, AI-driven recommendations, natural language search, and streaming data are all new to Spotfire X.

If you're a new user to Spotfire, I recommend you start at the beginning of the book and work your way through it, sequentially, to `Chapter 4`, *Sharing Insights and Collaborating with Others*, or `Chapter 5`, *Practical Applications of Spotfire Visualizations*. Then, you should dip in and out of the rest of the book as you see fit. `Chapter 10`, *Information Links and Data Connectors*, is a must-read for when you want to connect Spotfire up to any data source other than flat files. `Chapter 9`, *What's your Location?*, is essential reading if you want to perform any kind of geoanalytics or other location-based analytics. `Chapter 13`, *Beyond the Horizon*, covers several interesting topics, from theming, through to export, search, and more.

A lot of the URLs provided in the book are quite long, so I used `https://bitly.com/` for link shortening (in most cases). `https://bitly.com/` does track how many clicks each of the links has received, but not who clicked them. Of course, you can always just use a search engine to find each of the references, but I have included the links so you know you're always looking at exactly the right topic on the web!

Although the author, Andrew Berridge, works for TIBCO Software Ltd., it must be stressed that any views or opinions expressed within the book are solely Andrew's. They are not sanctioned in any way by TIBCO Software Ltd. and do not represent the views or opinions of TIBCO Software. Any apparently forward-facing statements or opinions should not be used for purchasing decisions for Spotfire or any other TIBCO product.

All product names, trademarks, and registered trademarks are property of their respective owners. All company, product and service names used in this book are for identification purposes only. Use of these names and trademarks does not imply endorsement.

Download the color images

We also provide a PDF file that has color images of the screenshots/diagrams used in this book. You can download it here: `https://www.packtpub.com/sites/default/files/downloads/9781787121324_ColorImages.pdf`.

Conventions used

There are a number of text conventions used throughout this book.

`CodeInText`: Indicates code words in text, database table names, folder names, filenames, file extensions, pathnames, dummy URLs, user input, and Twitter handles. Here is an example: "Also, try changing the left-hand line chart to show `Avg(Volume)` rather than `Avg(Close)`."

A block of code is set as follows:

```
if (Document.Properties['SelectedMetrics'] ==
'Descriptions' ):
```

Bold: Indicates a new term, an important word, or words that you see onscreen. For example, words in menus or dialog boxes appear in the text like this. Here is an example: "Spotfire automatically recognizes that we have a zip-code column and so will load the **USA Zip Codes** geocoding table."

Warnings or important notes appear like this.

Tips and tricks appear like this.

Get in touch

Feedback from our readers is always welcome.

General feedback: If you have questions about any aspect of this book, mention the book title in the subject of your message and email us at `customercare@packtpub.com`.

Errata: Although we have taken every care to ensure the accuracy of our content, mistakes do happen. If you have found a mistake in this book, we would be grateful if you would report this to us. Please visit `www.packt.com/submit-errata`, selecting your book, clicking on the Errata Submission Form link, and entering the details.

Piracy: If you come across any illegal copies of our works in any form on the Internet, we would be grateful if you would provide us with the location address or website name. Please contact us at `copyright@packt.com` with a link to the material.

If you are interested in becoming an author: If there is a topic that you have expertise in and you are interested in either writing or contributing to a book, please visit `authors.packtpub.com`.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions, we at Packt can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about Packt, please visit `packt.com`.

1
Section 1: Introducing Spotfire

The first section of this book is a general introduction to Spotfire—it shows how to load data and how to get started with Spotfire visualizations. Here, you will see how quick and easy it is to get started with producing insightful and impactful visualizations and how you can use Spotfire to collaborate with others.

In this section, the following chapters will be covered:

- `Chapter 1`, *Welcome to Spotfire*
- `Chapter 2`, *It's All About the Data*
- `Chapter 3`, *Impactful Dashboards!*
- `Chapter 4`, *Sharing Insights and Collaborating with Others*

1
Welcome to Spotfire

Welcome to the world of TIBCO Spotfire®! This book will take you on a journey through data discovery and visualization, advanced analysis, and beyond. It will show you how to get started really easily, then how to progress on to more advanced topics.

When you start Spotfire for the first time, your first task is to load some data. This data can be loaded from a wide variety of sources, from files through to database and big data repositories. It's then really easy to visualize and explore the data, gaining fresh insight and understanding all the time.

The following topics are covered in this chapter:

- Introduction to TIBCO Spotfire
- Getting started with TIBCO Spotfire
- The different Spotfire clients
- Importing and loading data into Spotfire
- The Spotfire recommendations engine
- Simple visualization types
- Building useful visualizations
- Details visualizations
- Introduction to marking and filtering
- Gaining insights from your data

Getting started with TIBCO Spotfire

In this section, we are going to explore loading and visualizing data using the Spotfire rich desktop client and the Spotfire web authoring clients.

There are a few different Spotfire client applications. As of Version X (pronounced 10), these are as follows:

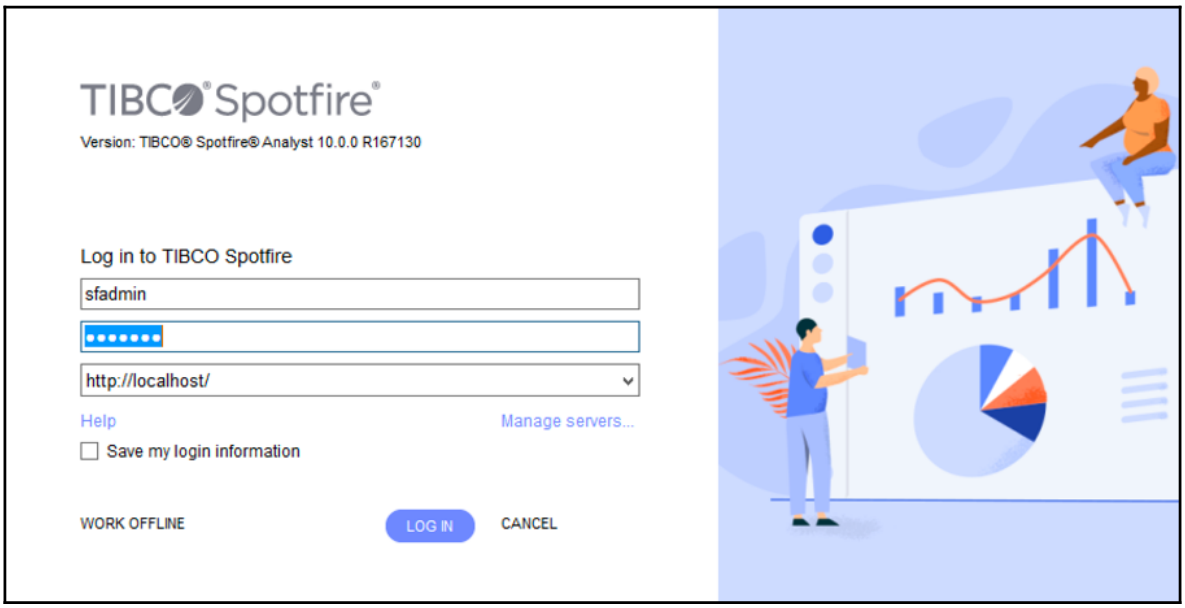
1. Desktop clients (rich, installed):
    - **Spotfire Analyst**: This fully featured application connects to an on-premises Spotfire server that's installed at your organization and is sometimes called Spotfire professional.
    - **Spotfire Cloud Analyst**: A fully featured desktop application that connects to a cloud-based Spotfire server, you can download it from your Spotfire Cloud account if you have one.
    - **Spotfire for macOS**: This is a hybrid application—it's essentially a wrapper around the Spotfire web clients (detailed later), but it is installed on desktop machines.
2. Web clients (thin, accessible via a web browser):
    - **Spotfire Consumer**: This is a standard, read-only Spotfire web client, usually available on-premises. You can consume existing visualizations and data. You cannot create new visualizations, load data, or change the configuration of visualizations (unless explicitly enabled by the author of a Spotfire file).
    - **Spotfire Business Author**: This is an advanced web client. In addition to the features of consumer, it allows loading of data, configuration of visualizations, and many other Spotfire authoring capabilities. You can check if it's available to you by logging in to Spotfire web (Consumer) and checking if there's an option to edit or create an existing analysis—more on this later.
    - **TIBCO Cloud Spotfire**: This has all the features of Business Author and is accessible via `https://cloud.tibco.com/`. You can even sign up for a free trial account—I recommend you do this if you just want to explore Spotfire before you purchase it.
3. Mobile clients:
    - **Spotfire Analytics for iOS**: Provides Spotfire Consumer-type functionality for iPhone® and iPad® devices.
    - **Spotfire Analytics for Android**: Provides Spotfire Consumer-type functionality for Android phones and tablet computers.
    - **Important**: You cannot create new analyses or modify existing analysis files using the mobile clients.

Launching Spotfire Analyst

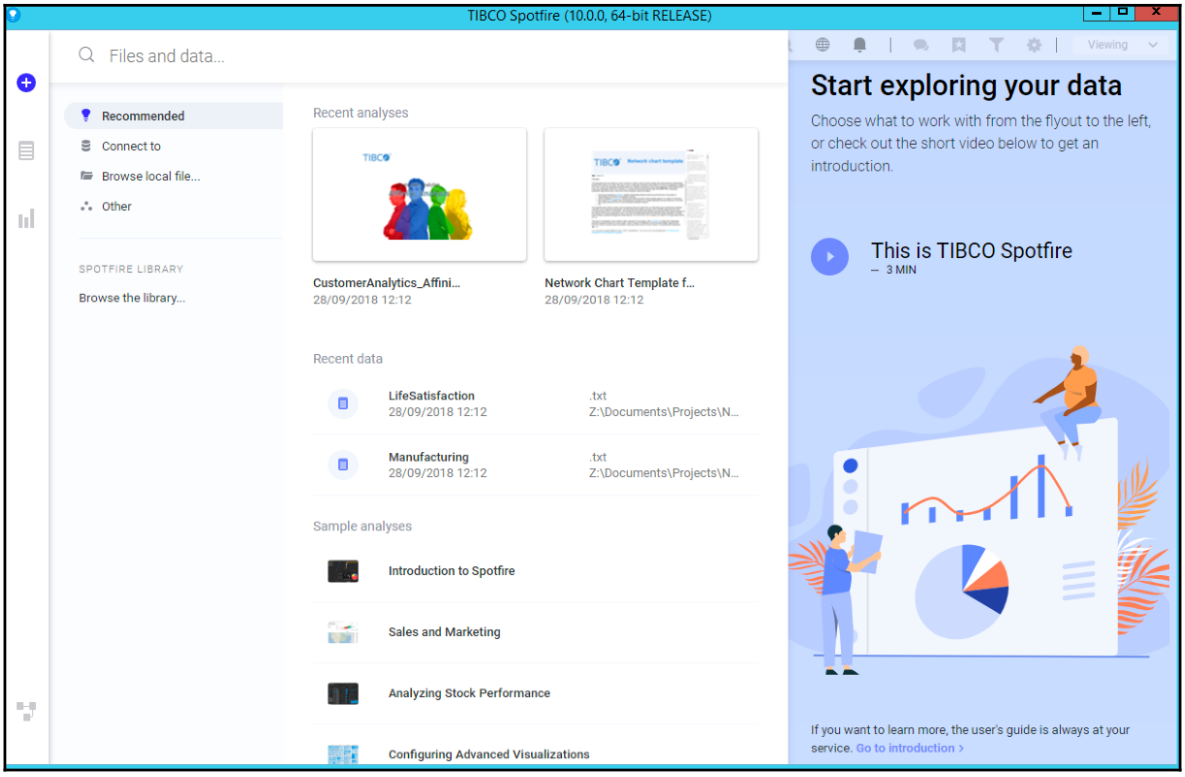
Spotfire Analyst is a rich desktop client and can be launched like any other Windows application. There are usually two shortcuts to Spotfire once it's installed. The first is as follows:

The other is as follows:

The first shortcut is the one you will use the most often. The **(show login dialog)** shortcut is useful if you've previously asked Spotfire to remember your server details and login information and you want to update these for any reason. By clicking on **Manage servers...**, you will be able to select which server to connect to (or, potentially, work offline, by selecting the corresponding option at the bottom-left of the page):



When you launch the Spotfire client, you will be presented with a home page that has shortcuts to various frequently used features, such as loading recently used analyses, loading recently used data, or looking at sample Spotfire files. This is the **Files and Data...** view, which can be shown or hidden by pressing the **+** sign on the top-left of the page. This page is the starting point for connecting to data and analyses, whether this be on the local system or in the Spotfire library:

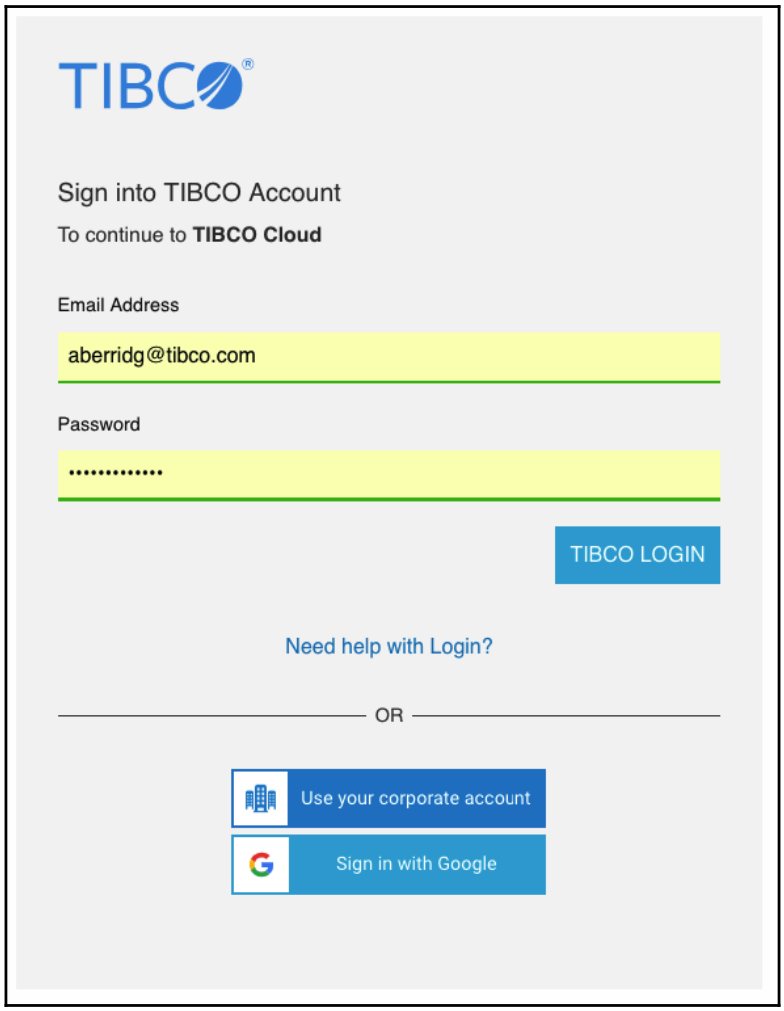


Feel free to explore some of the sample analysis files. I hope they inspire you to see what you can do with visualizations in Spotfire. Of course, you can always use the search box—it's really useful for finding files or connections to databases, and so on.

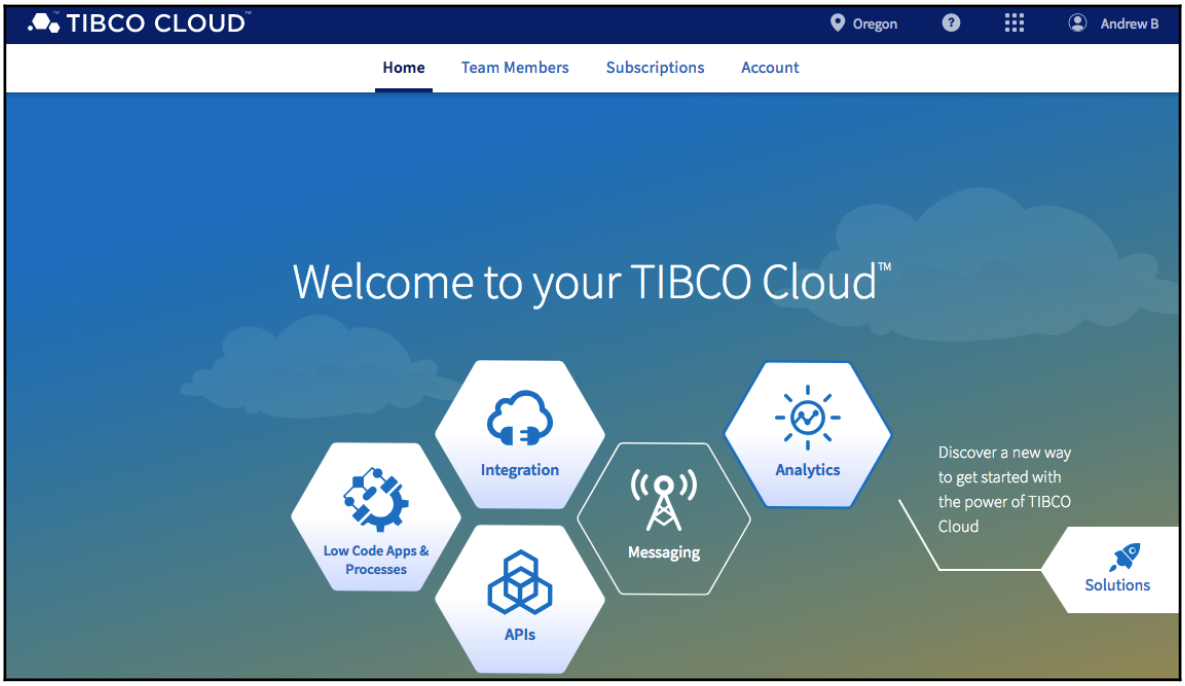
If you're interested, view the sample video! The link opens a YouTube video—while you're there, there are lots of other Spotfire videos that you can browse. Many of them cover some advanced topics, but hopefully they will inspire you to learn more and you'll come back repeatedly as there is always new content being published.

Logging in to TIBCO Cloud Spotfire

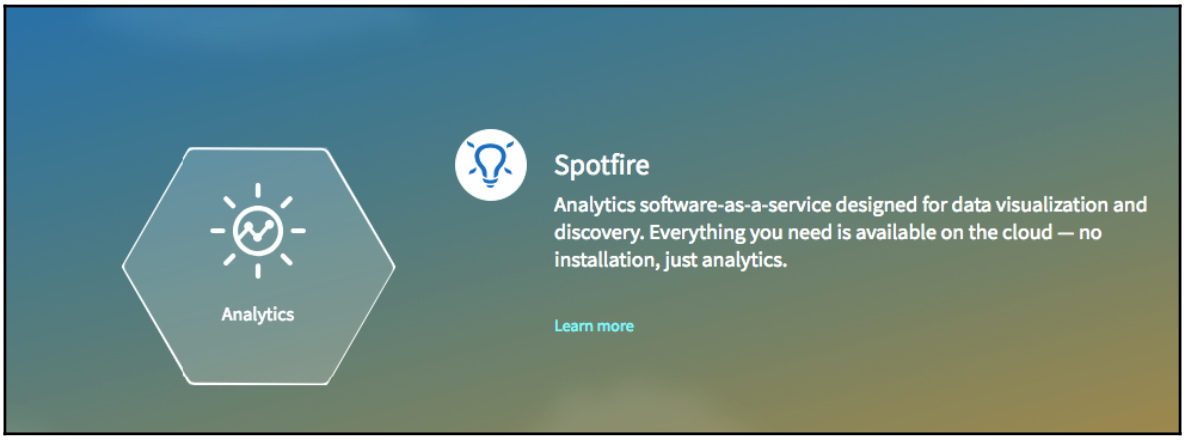
Visit `https://cloud.tibco.com/` and sign up for an account, if you don't have one already. Once you have an account, you can log in the usual way:



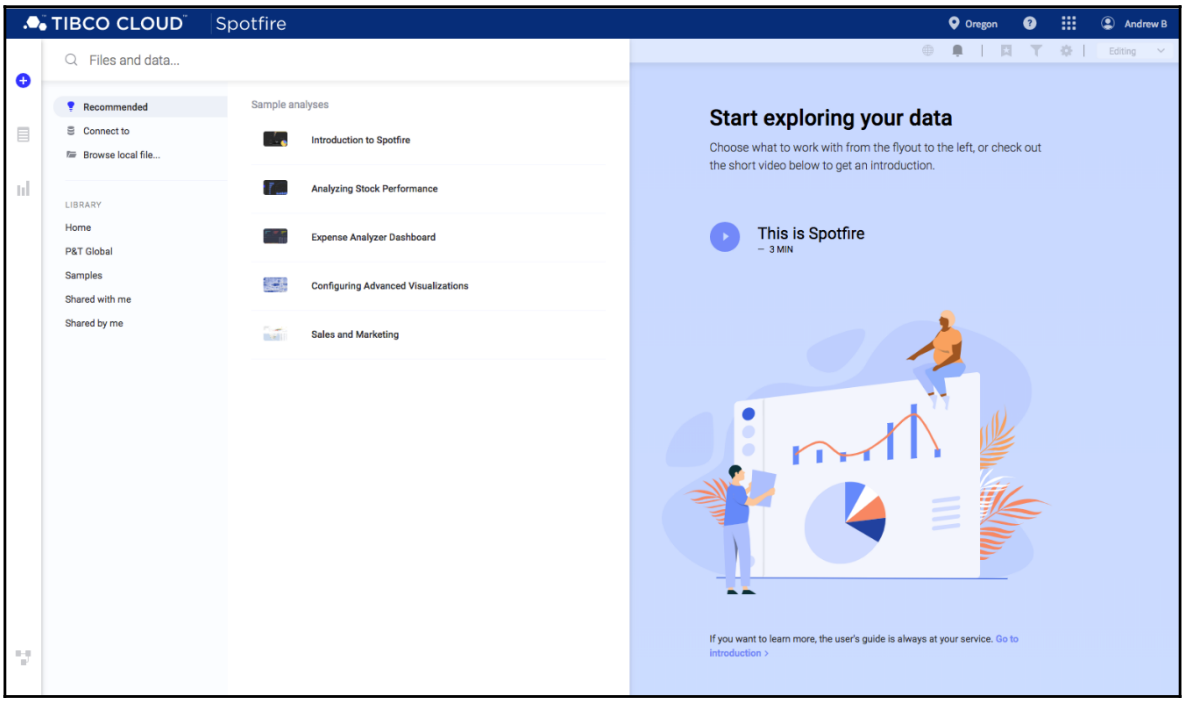
You will now be able to select which application(s) you would like to work with. In this case, choose **Analytics**:



Click **Spotfire**:



Spotfire will open and be shown to you. The web-based client looks very similar to the desktop (analyst) client, but has less functionality when it comes to building data workflows and various other authoring functions:

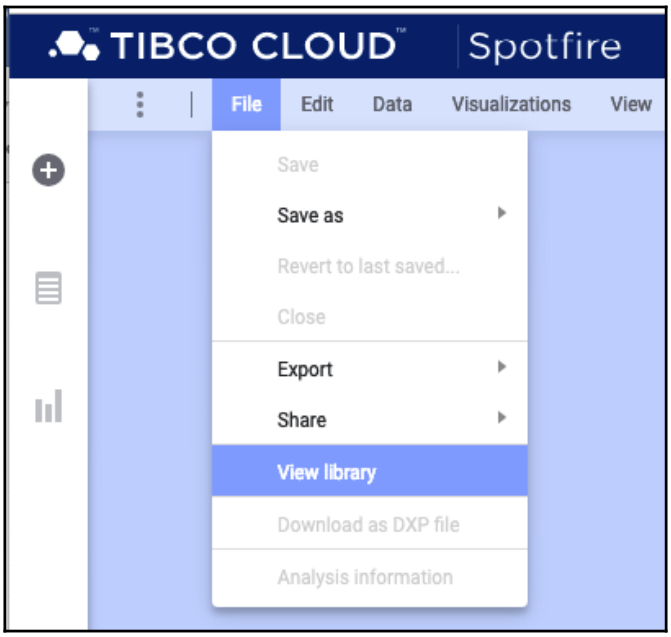


Downloading Spotfire Cloud Analyst

If you have a Spotfire Cloud account you can download Spotfire Cloud Analyst—this is a full-featured version of Spotfire that works on your desktop. I recommend that you download it as soon as you can—it's more powerful than the web-based clients, and some of the exercises in this book can only be performed using an analyst client.

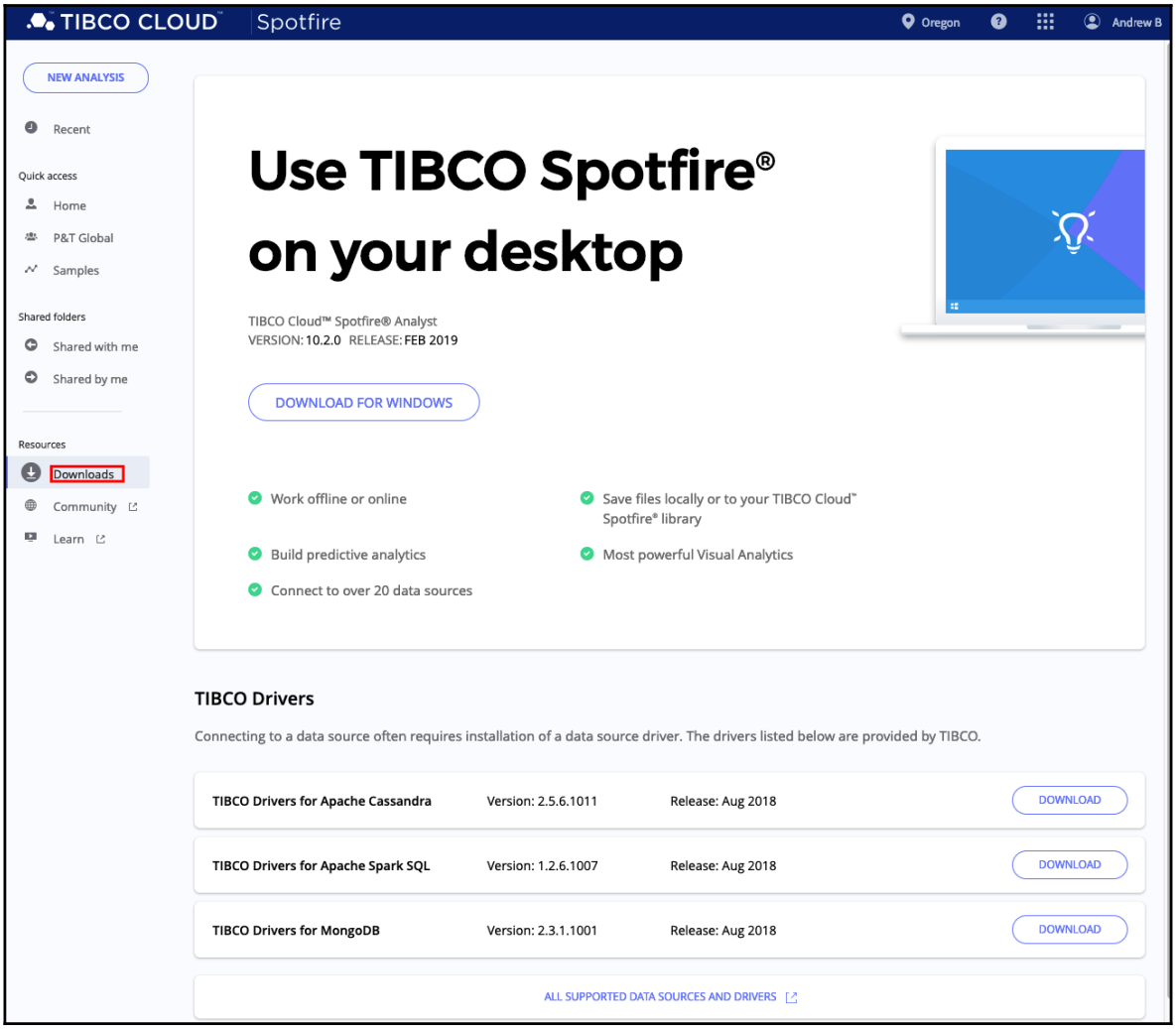
To download and install Spotfire Cloud Analyst, follow these steps:

1. Make sure you're logged in to your TIBCO Cloud account and Spotfire has been launched (like we did previously).
2. Dismiss the left-hand panel by clicking anywhere on the right-hand side of the window.
3. You should be left with a blank Spotfire screen—from the **File** menu, choose **View library**:



Hint: If the menu bar is not visible, click the three vertical dots in order to expand it!

4. Click **Downloads** under **Resources:**



5. Click **DOWNLOAD FOR WINDOWS**.
6. Once the download has finished, install the application as you would do so with any other Windows application.

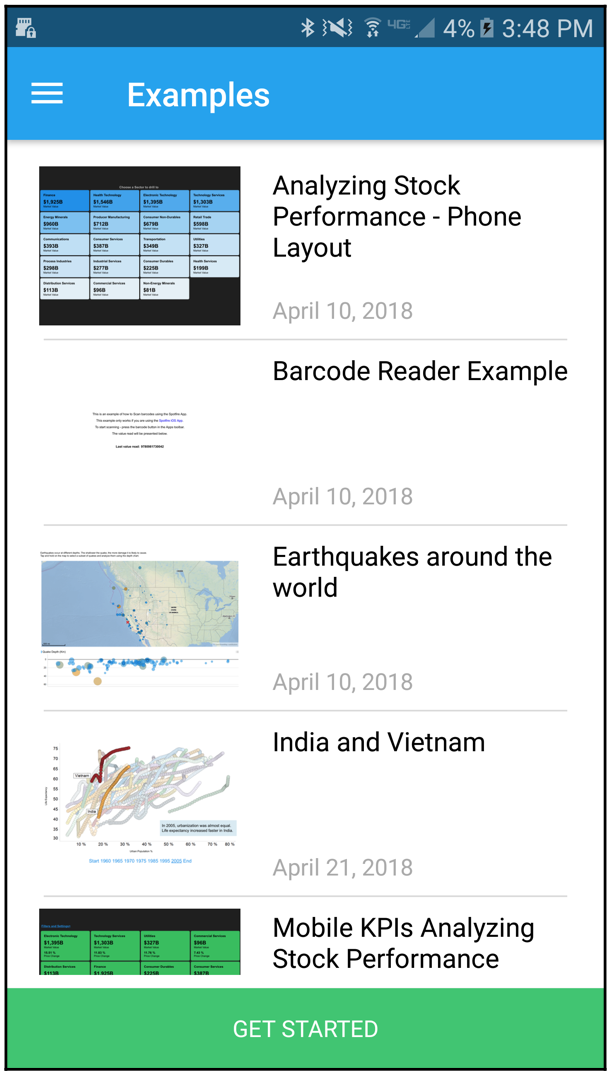
Installing the mobile apps and logging in to TIBCO Cloud

The Spotfire Analytics app is available for Android and iOS phones and tablets from the Google Play Store and the iTunes App Store, respectively.

Just search for Spotfire Analytics and install the app!

I have an Android phone, so I am going to use that to demonstrate this functionality. iOS is broadly similar.

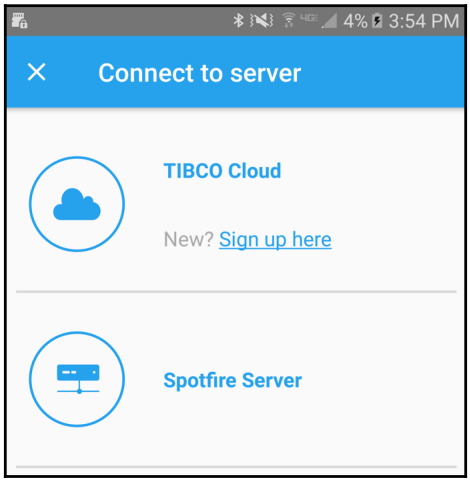
Once you have opened the app, you should get something that looks a bit like this:



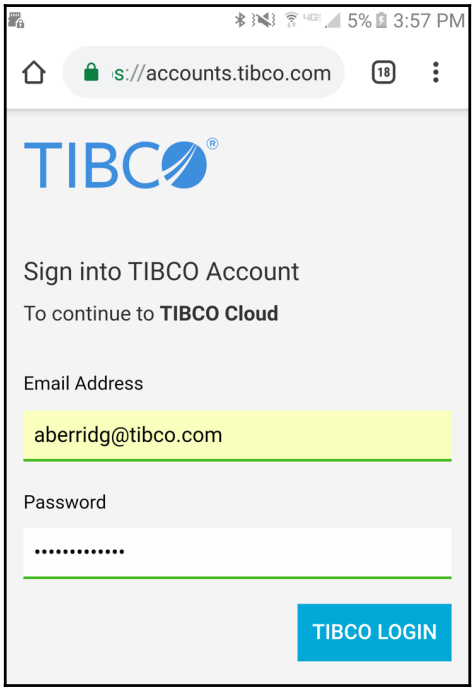
You can start off by looking at some of the examples—in fact, it's a great way to get started without even needing to log in anywhere!

However, let's get started by logging in to a Cloud account:

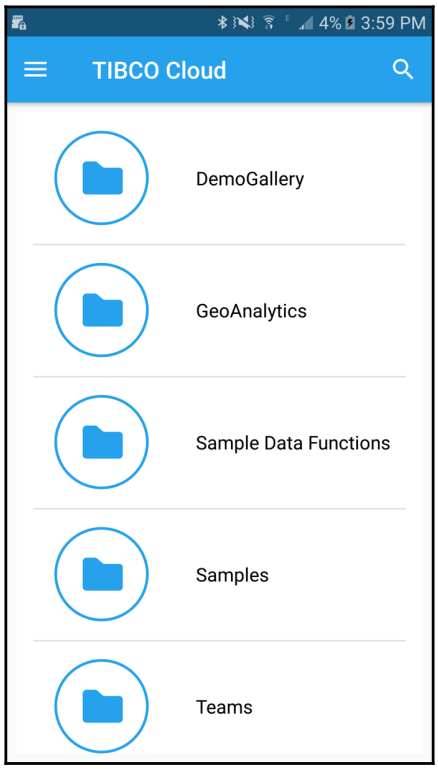
1. Tap the **Get Started** button.
2. Tap the **TIBCO Cloud** button (or **Sign up here** if you don't have an account yet):



3. Log in to your TIBCO Cloud account:



4. Once you have logged in, you should be able to see Spotfire's library browser. This is what I get:



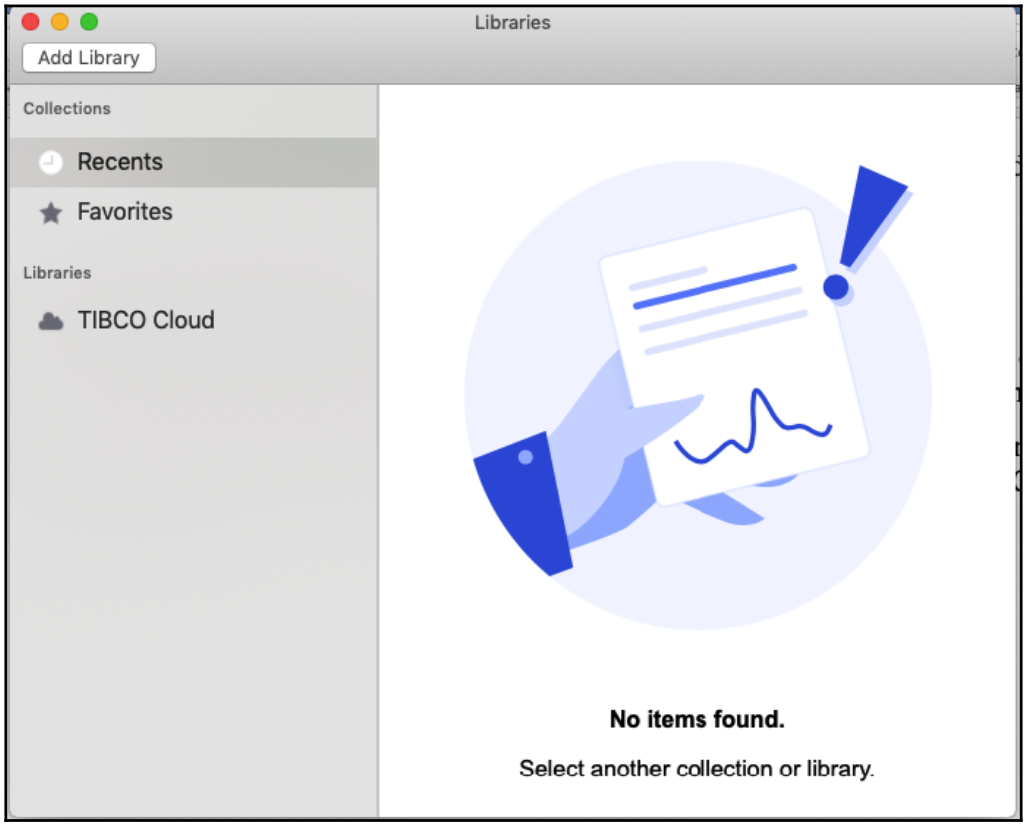
5. Feel free to browse the existing analyses. It's important to remember that you cannot create new analysis files or edit existing ones using the mobile apps.

You can also log in to your on-site Spotfire server if you have the requisite login and permissions (and potentially VPN access).

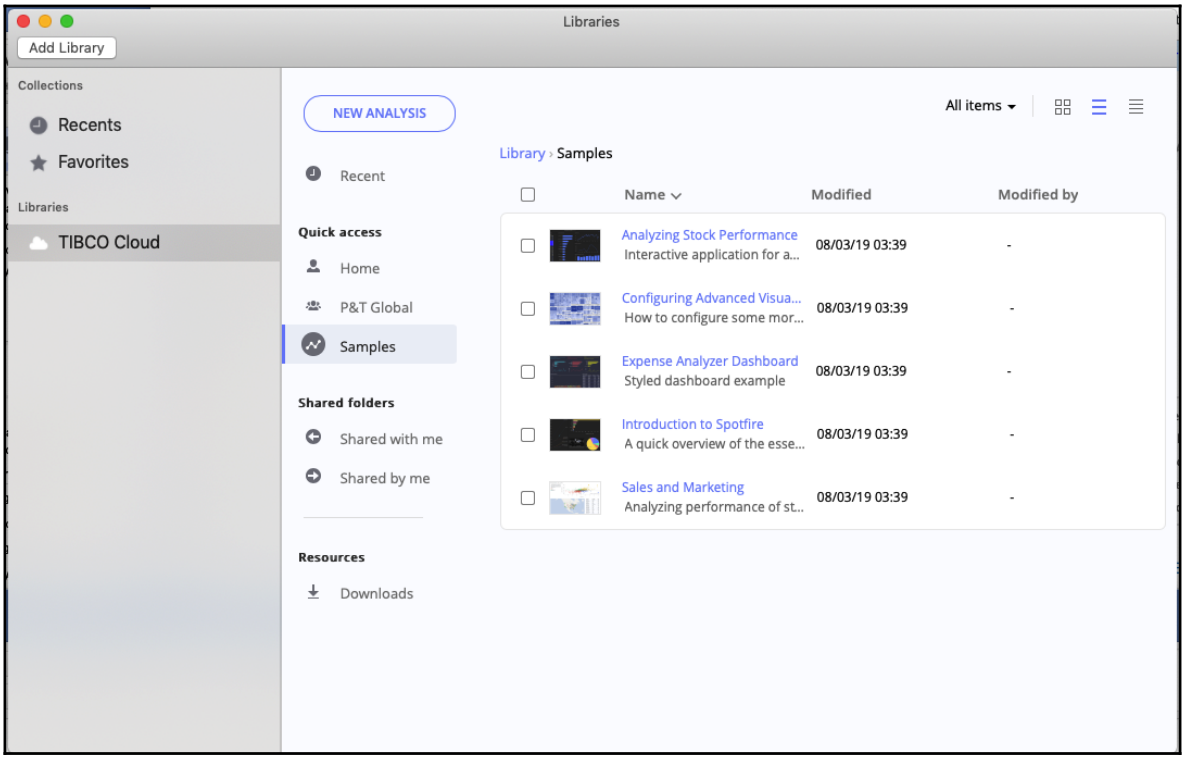
TIBCO Spotfire for macOS

As we mentioned previously, this is a lightweight wrapper around Spotfire's web clients, so you will need access to a Spotfire web server. This can either be in the cloud or on-premises.

Install Spotfire for macOS by searching for it on the App store. Once the application is installed, the following window is shown:



You can sign into your organization's Spotfire web server by clicking **Add Library** or log in to your Cloud account by clicking on **TIBCO Cloud**. In my case, I have logged in to my TIBCO Cloud account. If I select **Samples**, this is what I see:



Logging in to the Spotfire web clients (on-premises)

Logging in to the Spotfire web clients is straightforward. Just navigate to the URL provided by your server administrator. Once you are logged in, you will be presented with a blank Spotfire client, a view of the Spotfire library, or with a blank analysis, depending on your access levels.

If you are using Spotfire Consumer, the ability to create a new analysis will not be available to you. You'll need to get access to Spotfire Business Author or one of the analyst clients (for example, Cloud Analyst) in order to be able to author Spotfire analyses, as per most of the examples in this book.

Spotfire licenses

Before we get started with loading data and doing some analysis, I'd just like to briefly cover the very important topic of Spotfire licenses. Spotfire licenses are not software licenses in the normal sense (where you have to buy a license in order to use the software). You can think of Spotfire licenses as permissions to perform certain functions in the application. Almost every function in Spotfire has an associated license. Licenses are assigned via user groups. Assigning licenses is an administrative function and, as such, will be controlled via your Spotfire administrator.

A TIBCO Cloud account will probably have the most licenses assigned to it, so you should be able to do most things with Cloud Analyst or the web (business) author.

However, if you struggle to follow along with the examples in this book—the options don't seem to be there, or you just can't even get started–it's possible that you don't have the required license(s) to perform the functions that are suggested. Please contact your Spotfire administrator to get this fixed.

To emphasise: Licenses grant permissions to perform functions in TIBCO Spotfire. If you don't seem to be able to do what you want or cannot even get started, please contact your Spotfire administrator!

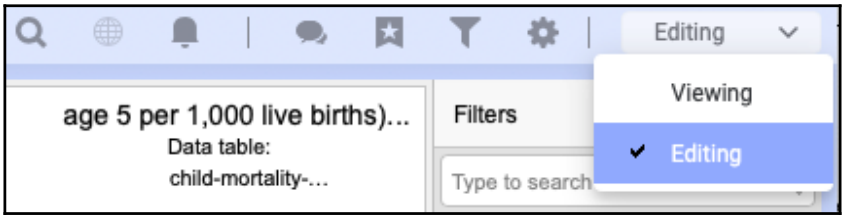
Getting started with loading data

The simplest way to get started with loading data into Spotfire is to import some data from a file such as an Excel spreadsheet, so that's what this tutorial will cover.

Important! Before you start, you must be in **Editing** mode. I will periodically remind you of this throughout this book. The analyst client defaults to **Editing** mode—other clients, such as the web client, may not. So, beware!

To switch to **Editing** mode, follow these steps:

1. In the top right-hand corner of the application, click the dropdown.
2. Choose **Editing**:

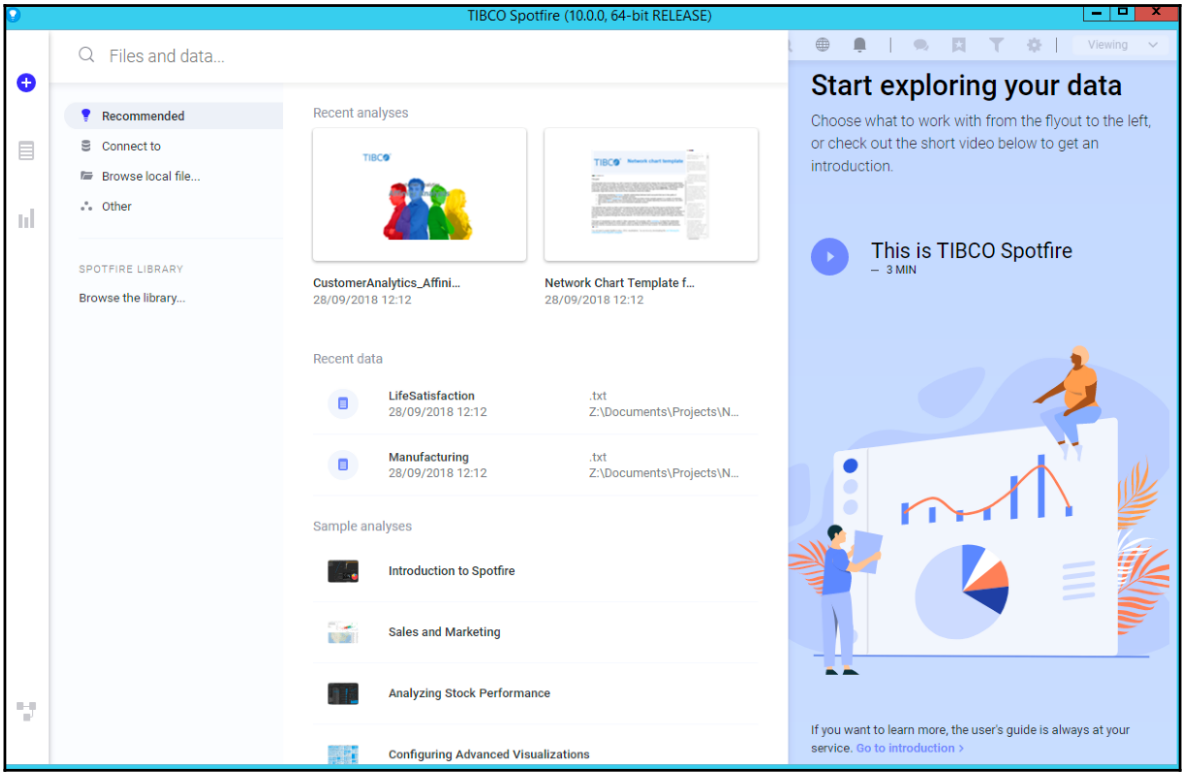


3. If **Editing** mode is not available to you, it means that you do not have the correct permissions (license) for editing, or that the Spotfire client itself does not support it (for example, the Android or IOS clients).

Importing Excel spreadsheets into Spotfire

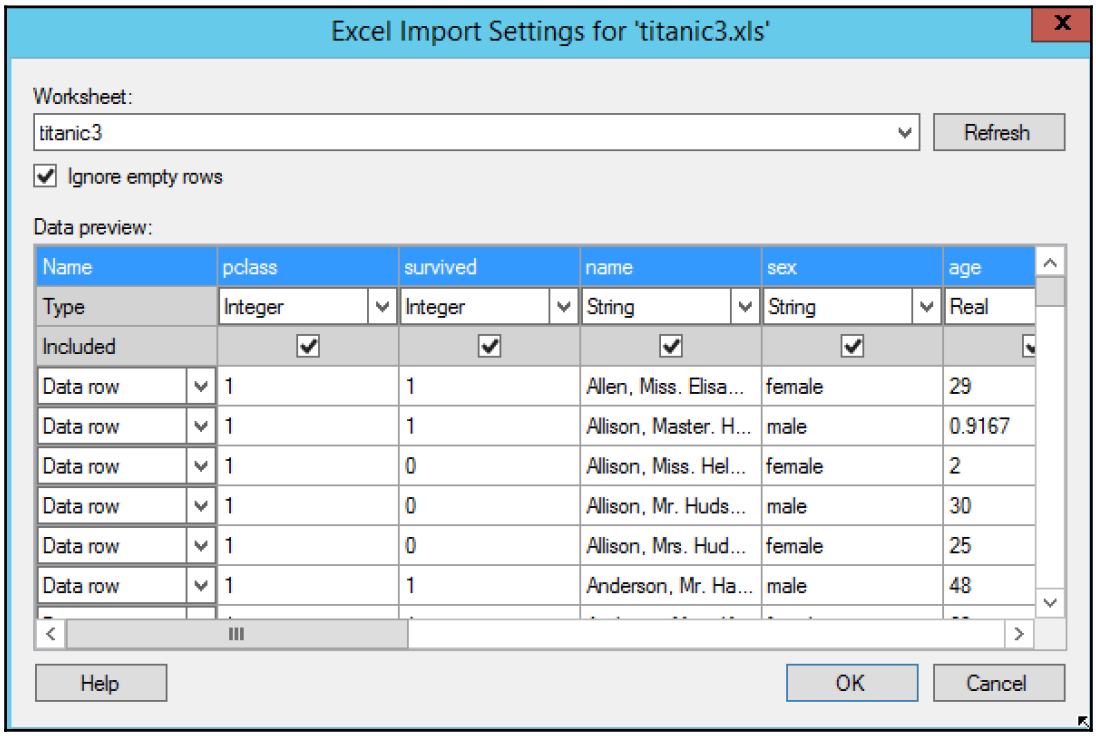
The procedures for importing **comma-separated values** (**CSV**) files or Microsoft Excel spreadsheets in Spotfire are essentially identical:

1. From the Spotfire home page (shown initially when launching the Analyst or web clients), select **Browse local file...**:

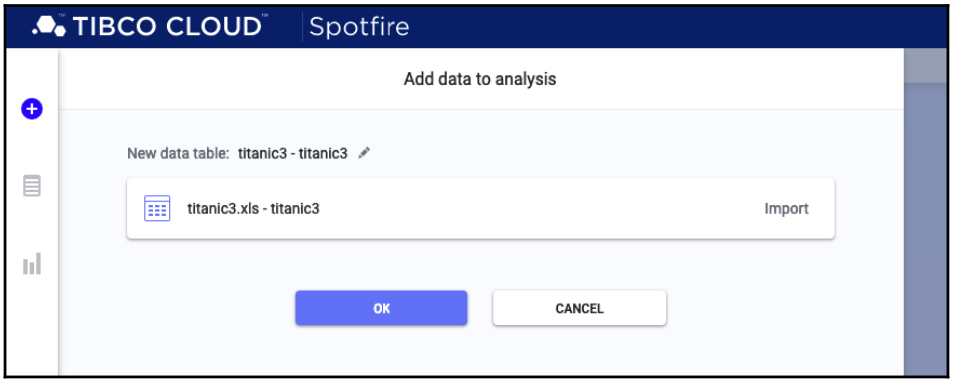


2. You will be presented with a standard **Open File** dialog, allowing you to navigate to the file that you want to load. For this example, let's use some publicly available data on the Titanic disaster—it can be downloaded from the following link: `http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3.xls` or: `http://bit.ly/2HStQ7R`.

3. In the Analyst client, Spotfire will open a dialog, which will allow you to define the import settings:



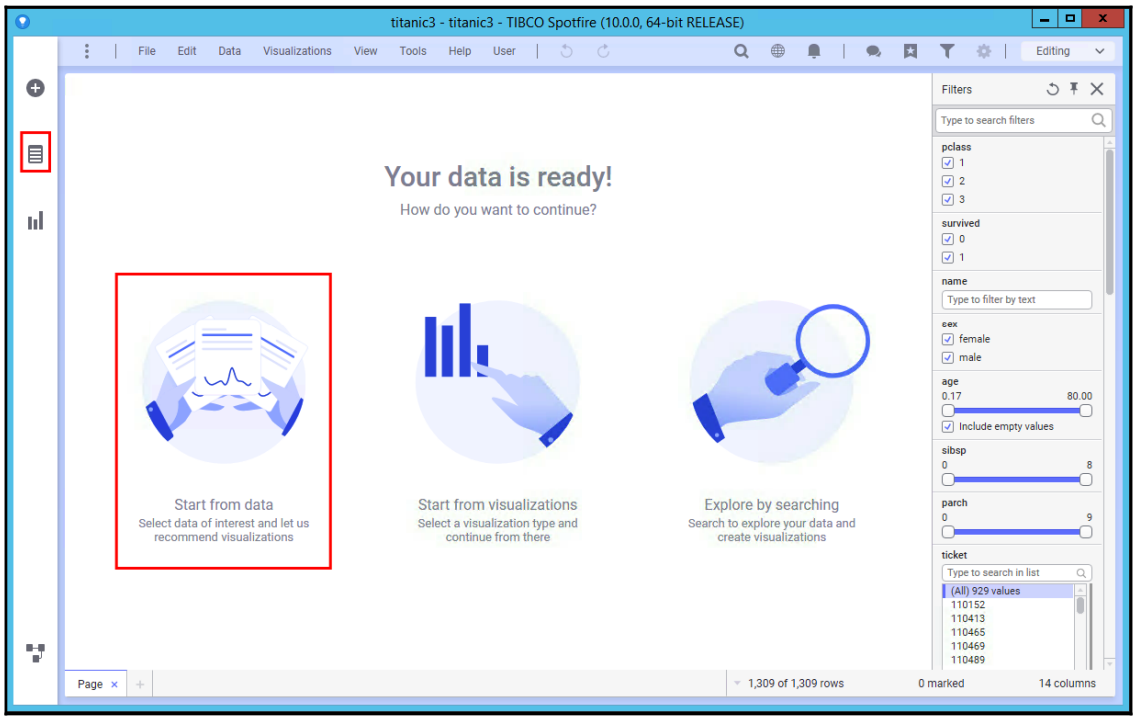
4. The first thing to notice is the **Worksheet** selection dropdown at the very top of the dialog window. Spotfire can only import one worksheet at a time. There is only one sheet in our file, so we don't need to do anything with this option.
5. The next thing to notice is the preview of the data and its structure. Spotfire will automatically detect and assign column headers and data types, but you can change any of these settings. You can also tell Spotfire not to import specific columns or rows.
6. We want to open the file with all defaults, so we're just going to click **OK**, but please do explore the drop-down options for columns and rows and experiment with the settings. The core philosophy of Spotfire is discovery, so start as you mean to continue and explore some of the options.
7. Once you click **OK**, Spotfire will show the **Add data to analysis** display:



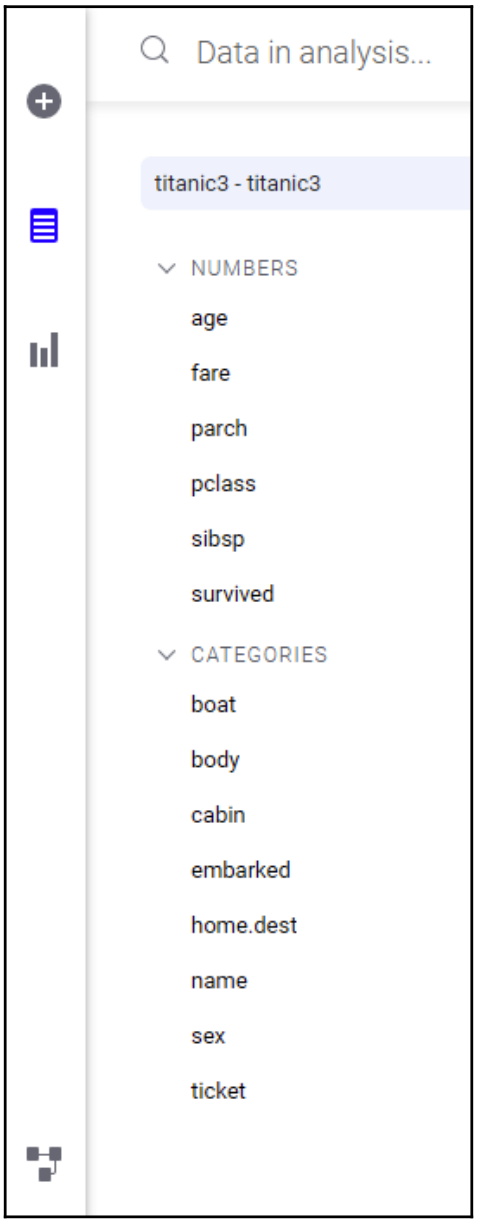
8. Click **OK** to load the data.

Introduction to the data panel

You'll find that a lot of work is done in Spotfire via the data panel. You can show the data panel by clicking on the big icon in the middle of the Spotfire window, or by pulling it out by clicking on the data icon on the left-hand side of Spotfire:



The data panel shows all the data tables and columns that are available in the analysis. Spotfire has already classified the columns into different groups of numerical and categorical columns:



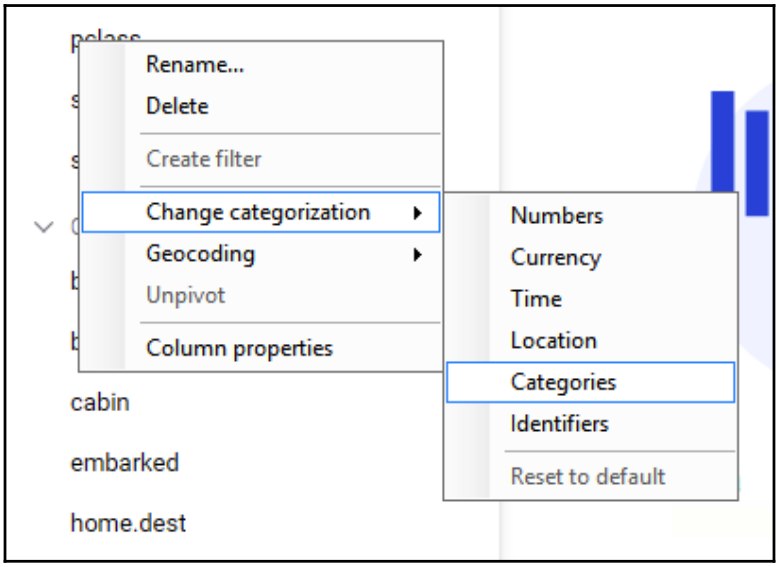
In this particular dataset, some categorical columns have been loaded as numeric columns. It's not Spotfire's fault—it's just that some of the data columns are integers in the data, and represent categories. Think of the column called **survived**. This is a 1 or 0, indicating whether the passenger, died or survived. Similarly, passenger class (**pclass**), the class of passenger should be categorical since it is either 1, 2, or 3, and taking any kind of aggregation of this (average, max, and so on) probably doesn't make much sense. You can read more about the dataset and its data dictionary here:
`http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3info.txt`

Or:

`http://bit.ly/2uDg6oX`

Next, we are going to use Spotfire's recommendations engine to build a visualization, but in order to get the best results from it, it can sometimes be a good idea to change the categorization of columns in order to give Spotfire some hints about how to display or analyse the data. So, let's do this first:

1. Right-click the **pclass** column and change its categorization to **Categories**:



2. Do the same with the **survived** column.
3. Now, we can get started with building visualizations!

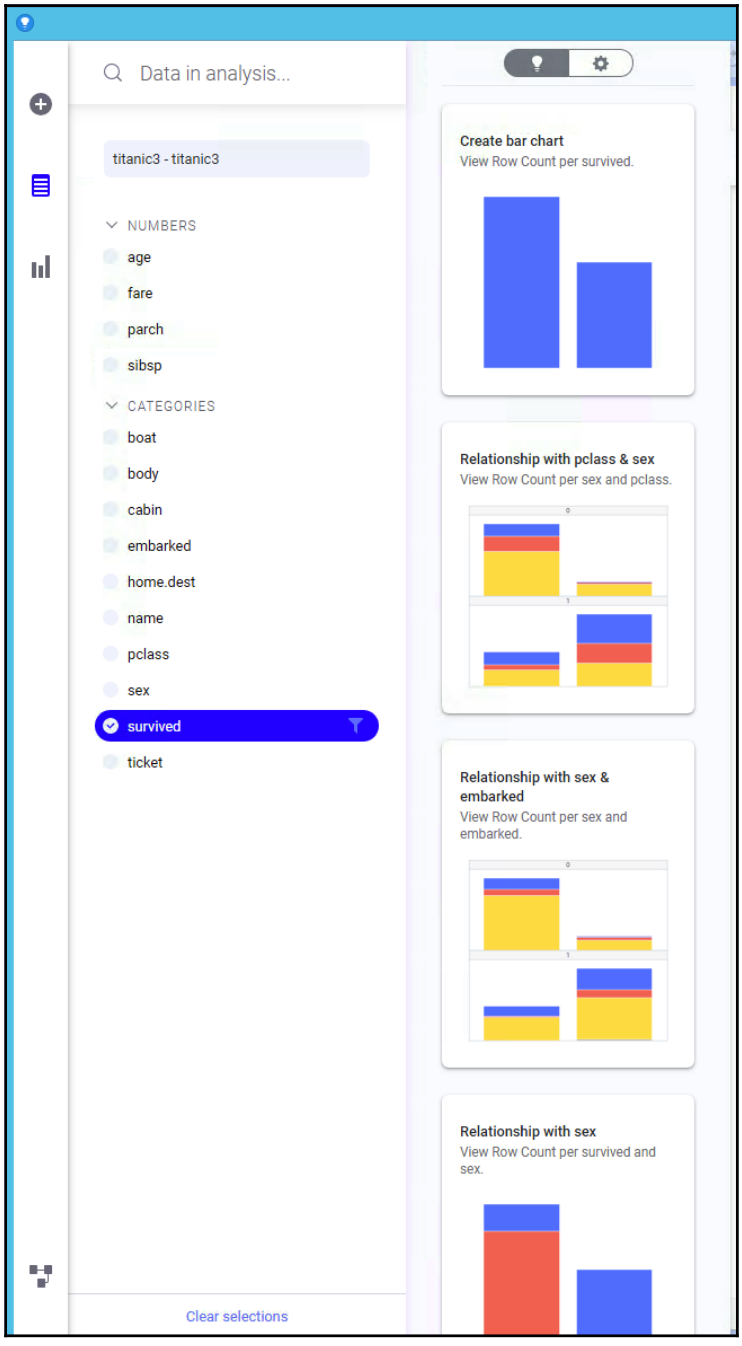
The Spotfire recommendations engine

The recommendations engine gives you instant insight into your data and some suggestions of which visualizations to use. Spotfire's analyst clients have an advanced feature called **AI-Powered Suggestions**. This is Spotfire's new way of helping make sense of any type of data, regardless of its size or shape. The basic premise is that you should select a "target" column in the data panel and Spotfire will do the rest. Spotfire runs a specialized algorithm over all the columns in the data and selects those that most strongly drive or influence the target. Those columns are called "predictors." It then produces suggested visualizations for the target and selected predictors.

The recommender is available in the web clients too, but (at the time of writing) the web clients do not have the AI element, where predictor columns are automatically selected. I hope that the feature will be made available at some point. In the meantime, if you're using the web clients, you can follow a slightly different path to create visualizations. I'll point out how to do that along the way.

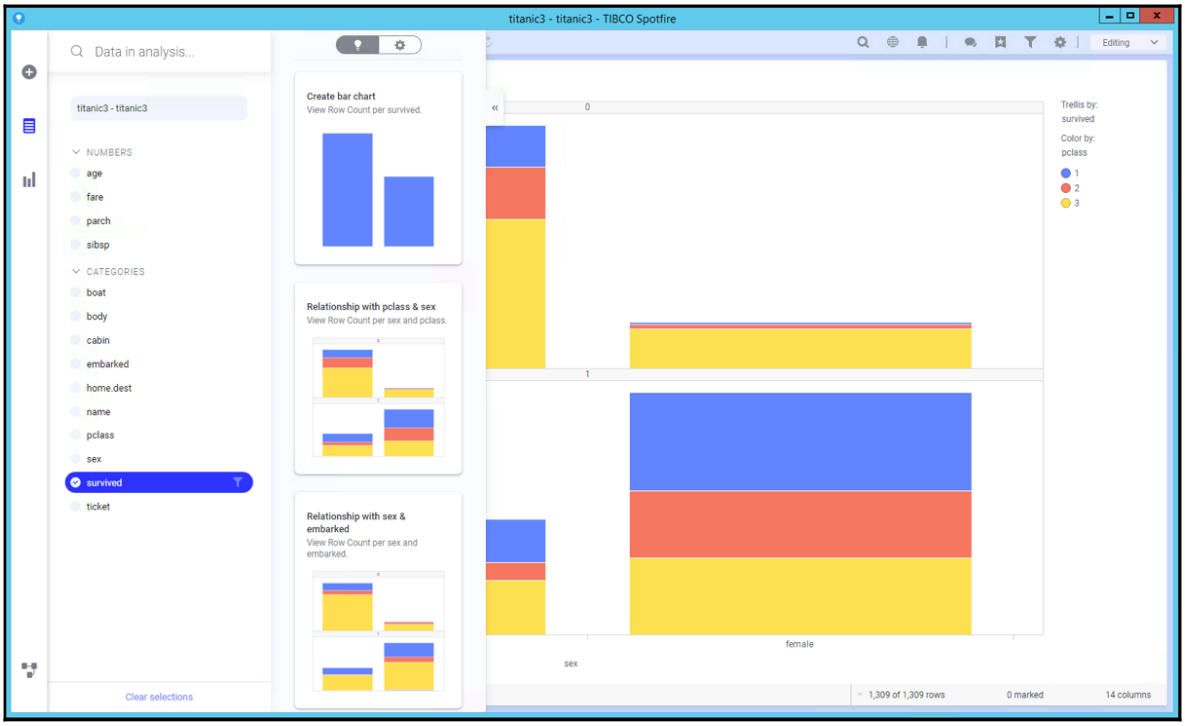
Let's get started! In the case of the Titanic data, the most obvious target column is **survived**. In other words, we'd like to know which columns best predict, influence, or explain whether passengers survived the Titanic disaster or not:

1. In the data panel, select the **survived** column. In analyst clients, Spotfire will produce something that looks like this:

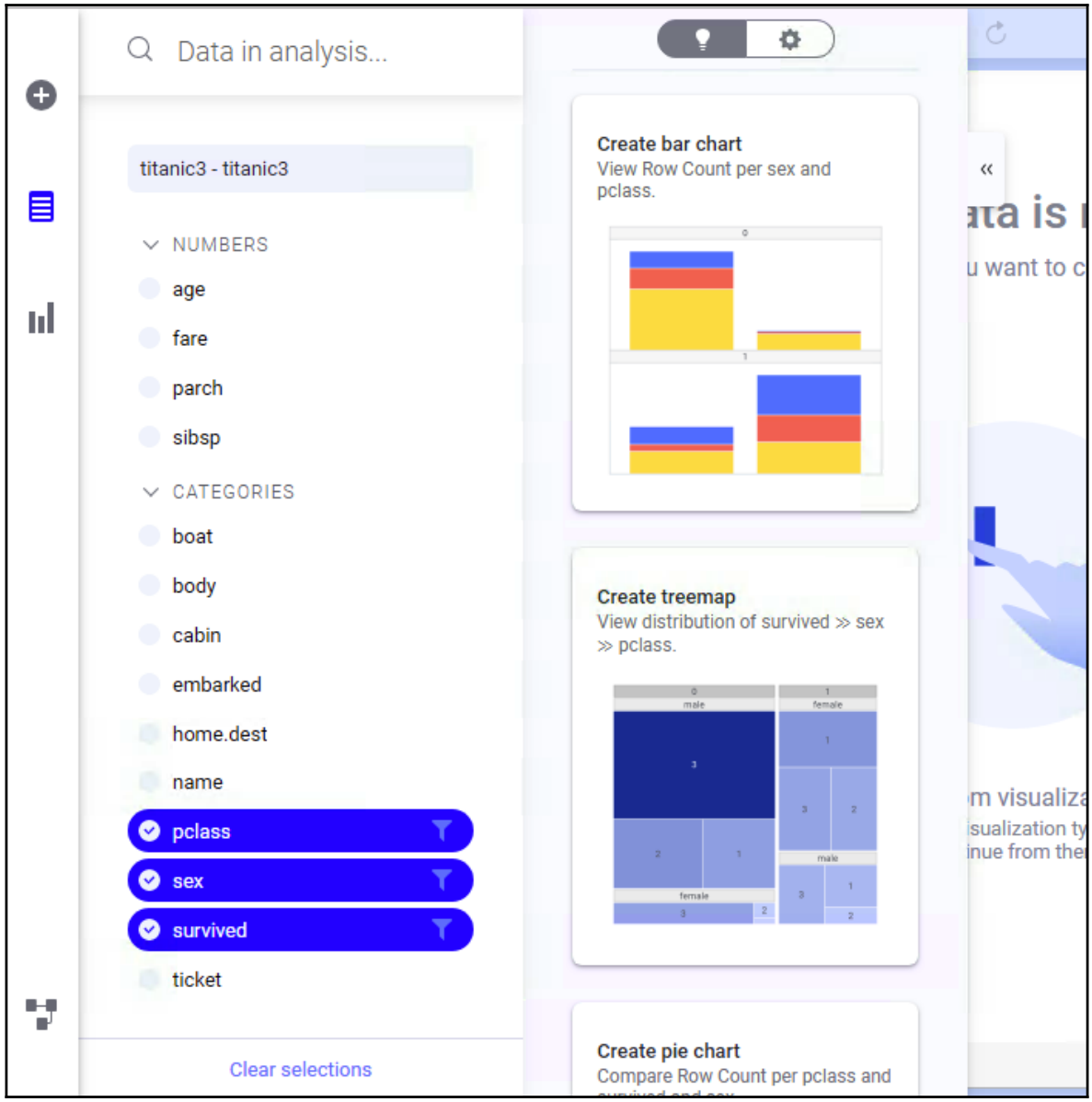

Data in analysis...
titanic3 - titanic3
NUMBERS
age
fare
parch
sibsp
CATEGORIES
boat
body
cabin
embarked
home.dest
name
pclass
sex
survived
ticket
Clear selections
Create bar chart
View Row Count per survived.
Relationship with pclass & sex
View Row Count per sex and pclass.
Relationship with sex & embarked
View Row Count per sex and embarked.
Relationship with sex
View Row Count per survived and sex.

Interesting! Immediately, we can see that the strongest predictors of survival are **pclass** and **sex**. The very first visualization is always just the row count of each of the values in the target column, so the second visualization is the one that begins to explain the target.

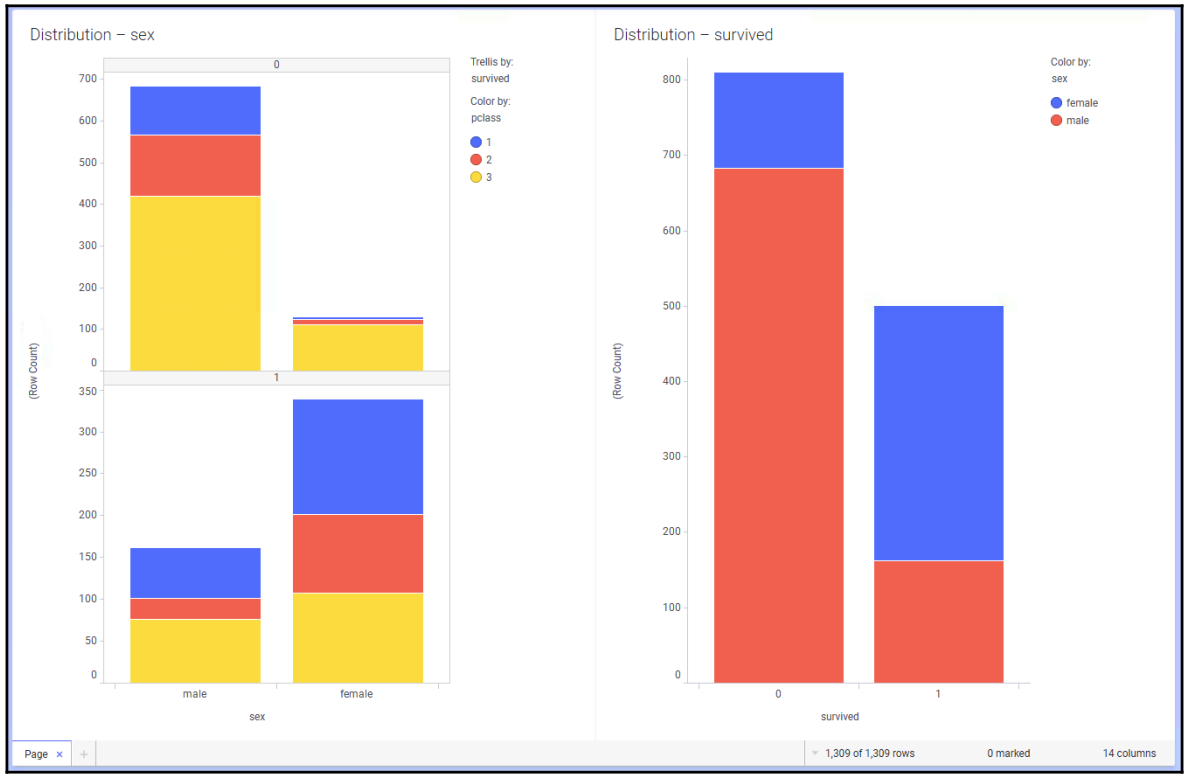
2. To add the visualization to your analysis, just click on the visualization that shows the relationship with **pclass** and **sex**. Your Spotfire session should now look something like the following screenshot:



3. You can produce the same effect in Spotfire web clients by selecting the **survived** column, the **pclass** column, and the **sex** column. Hold down the *Ctrl* key while clicking to select multiple columns:


Data in analysis...
titanic3 - titanic3
NUMBERS
age
fare
parch
sibsp
CATEGORIES
boat
body
cabin
embarked
home.dest
name
pclass
sex
survived
ticket
Clear selections
Create bar chart
View Row Count per sex and pclass.
Create treemap
View distribution of survived » sex » pclass.
Create pie chart
Compare Row Count per pclass and

4. I think it's also interesting to explore the male/female ratio on board the Titanic, so we need to add a bar chart visualization that shows just **survived** and **sex**. The AI recommender will choose these columns—just scroll down the panel a bit to find the visualization that shows this relationship, or manually select the columns in the web clients, then click the visualization to add it to your analysis. Now, let's pause and look at what we've created. In a few clicks, we have loaded some data and created two visualizations that really explains a lot of findings (insights) all in one go! Here is the analysis without the data panel (collapse it by clicking the double arrow toward the top-right corner of the panel):



Here are my notes on interpreting these visualizations. Of course, what I say is only a matter of opinion, so feel free to draw your own conclusions:

1. Notice how many more men there were on board than women?
2. Look at how many more men died than survived! The survival rate of women was much higher—this would be borne by the "women and children first" policy of filling the lifeboats.

3. Look at the left-hand visualization. It is trellised by **survived**—this means that Spotfire has split it into panels, one for each data value.
4. First-class female passengers had the greatest chance of survival—suggesting that class and socioeconomic factors played a role in the survival rates.
5. More third-class males survived than first- or second-class ones. Might they have had stronger fighting instincts or been pushier? Were they more willing to cram into overcrowded lifeboats? All are potential explanations.

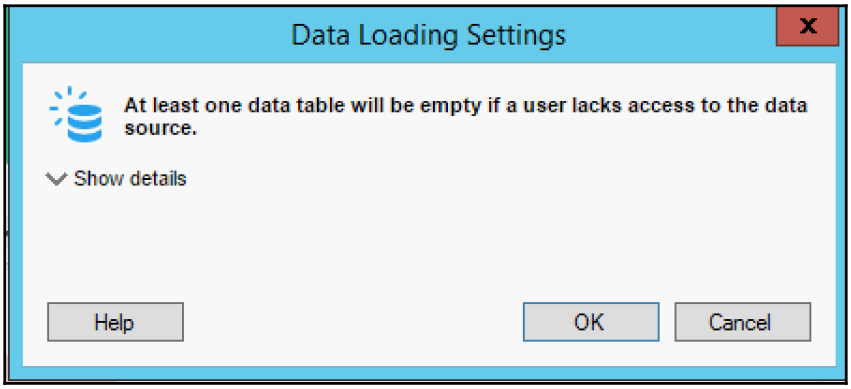
Saving Spotfire files

It's a good idea to get into the habit of saving your Spotfire analysis files. There's no auto-save capability at the time of writing, so save regularly!

Saving a file in analyst clients

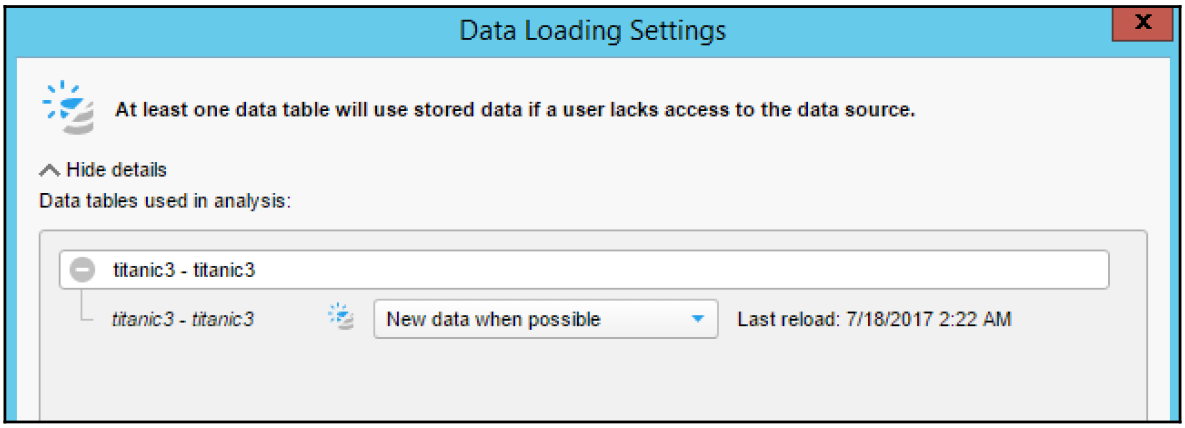
When saving a file in Spotire Analyst, you have a lot more options than in web-clients. Here's how to save a file in Spotfire Analyst, with a couple of tips along the way:

1. You'll notice that when you save the file (**File** | **Save as** | **File**) for the first time, Spotfire will show this prompt:



2. This is an important dialog to discuss. If you want to share your Spotfire file with another user, they will not be able to view it if they don't have access to the original data file. We can fix this now.

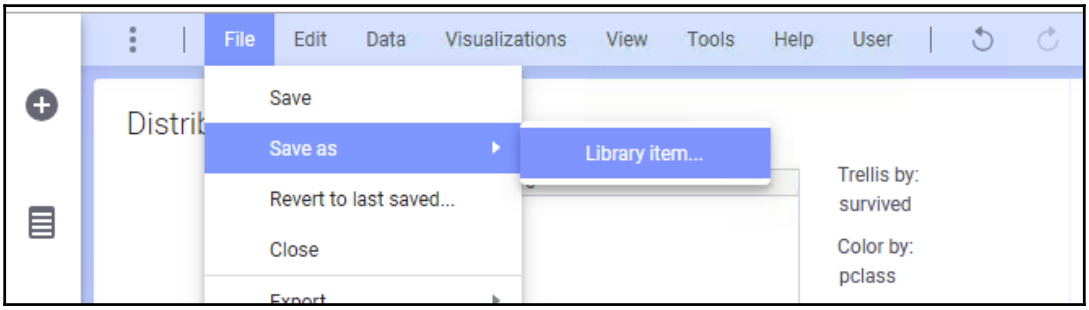
3. Click **Show details**. You'll see a dialog that allows you to choose what happens with the data. In this case, it's most appropriate to select **New data when possible**. This means that Spotfire will load new data if it's available (for example, if you updated the original Titanic data file), otherwise it will use data stored (embedded) in the Spotfire file:



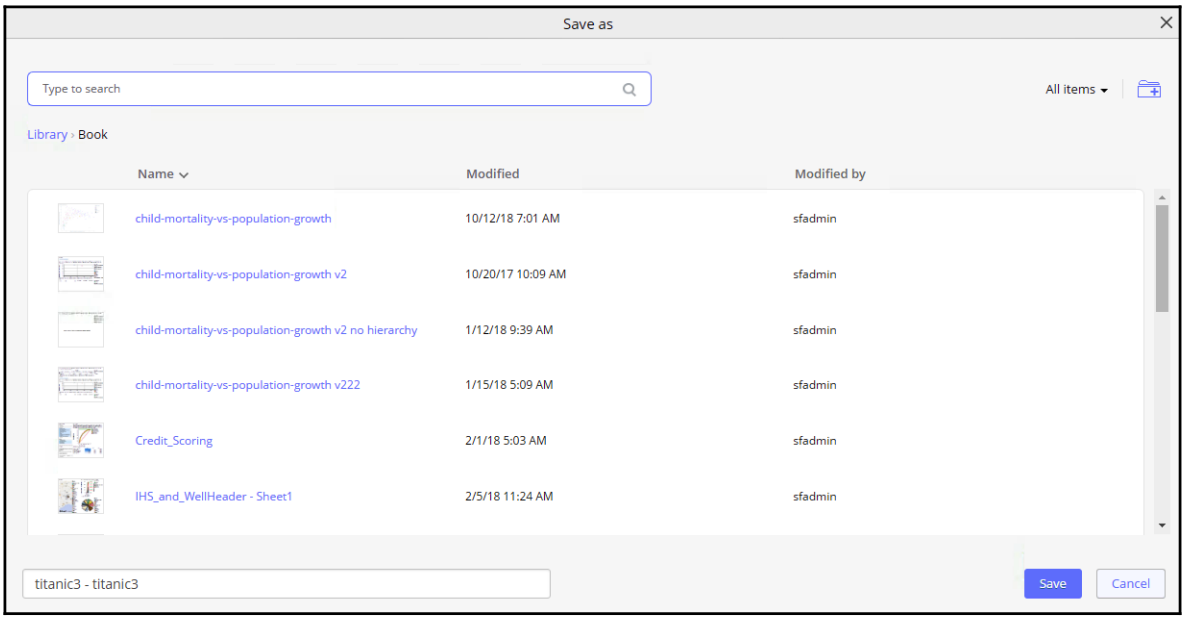
Saving a file in Spotfire Cloud or Business Author

Saving a file in Spotfire Cloud or Business Author is really straightforward. Saving a file will save it into the Spotfire library, where you and others can access the file (if folder permissions are set correctly—there is more on this later in this book):

1. Click the **Save** menu at the top of the screen and select **Save As** | **Library item...**:



2. The Spotfire web client will prompt you to choose the destination folder in the Spotfire library:



3. You can create a new folder using the + icon if you want.
4. Name the file as you see fit. Spotfire won't prompt you for the data saving settings (as it did in the analyst client) because the data is always embedded in the analysis.
5. When you're finished, click **Save**.

Producing a useful interactive dashboard

Now that we have produced some visualizations from the data, let's turn them into a useful interactive dashboard that we can use to gain more insight from the data.