



TIBCO Data Science: The Great Analytics Migration People, process, and technology



Introduction

What do monarch butterflies, sockeye salmon, wildebeest, sperm whales, and red crabs have in common? They have all undertaken a great migration fraught with challenge and peril, but one that ensures survival.

Whether your organization is migrating 10 users or 10,000, in six weeks or six years, we think you'll end up asking questions such as the ones in this ebook. It describes the approaches to people, process, and technology that contribute to the success of a migration to the TIBCO Data Science data science solution.

To give you insight into the scope and objectives of one migration project we have experience with, here are a few high-level results:

- Hundreds of users migrated worldwide
- Substantial, bottom-line impact from saved fees
- More than 300 projects across multiple business units migrated to a TIBCO Data Science solution
- Migration project team consisting of 12 points of contact for users
- Overall project duration of less than nine months, with user migration in less than six months

Based on real experiences, this ebook describes an enterprise migration project and provides advice and lessons learned.

Of course, every organization is different. The project we describe here achieved these results in so little time because of the commitment and extraordinary efforts of teams all around the company. We can't guarantee similar savings from every migration to TIBCO Data Science software, but we have seen a bottom-line impact that similar companies making a similar transition could expect to see. In any event, we invite you to learn from these lessons.



Chapter 1: People





First, why migrate to TIBCO Data Science?

Without a good answer to this question,

the others don't matter.

Your mileage may vary, but the company described here was spending a great deal of money every year to keep the legacy product. Of course, everybody wanted to save money, but everybody also knew the migration was going to involve a lot of risk, work, and change, so they had to justify the effort.

You want to enable analytics in the enterprise.

We've all read about it in books, magazines, and news articles — we need to do something about analytics and big data. Organizations that embed analytics within You want your analysts to analyze data, not write code. all parts of their business to make faster decisions and improve decision-making, Many platforms are powerful, but you may have to hire people to write and maintain planning, and forecasting have a distinct competitive advantage. Unfortunately, there code and administer the complex system to get the most out of it. Many companies is a skills shortage, so we need a software package that plays well with existing IT are having increasing trouble finding good replacements for legacy product specialists investments and is sufficiently easy to use. The goal is to enable all users — experts who are leaving or retiring, so the cost of keeping a legacy platform can rise beyond and line-of-business users alike — to make the most of their data. annual licensing fees.

You want to lower the cost of software licensing.

You want analytics software that is less expensive. That's no surprise, but more important, you want software that makes analytics accessible to more people.

But you probably don't know what a migration project would entail and like most companies, you are concerned about the downside of an unsuccessful one.

You can find better value for money.

A midsize to large company derives significant value and competitive advantage from applying analytics in areas such as marketing, price optimization, forecasting, technical support, supply chain optimization, preventive maintenance, and financial credit risk analysis. We believe you can get better value for less money with the full range of analytics muscle and ease of use of TIBCO Data Science tools.

How many people are you going to affect? Where are they?

As with any software product, most of your users are working with a subset of product functions to accomplish their daily tasks. Then you have a group of power users who eat, sleep, and breathe the product.

You probably can quickly identify hundreds of users who need to move to a full featured, modern platform like TIBCO Data Science software. Some companies have a team that provides analytics expertise and support to different functional organizations that do not have their own internal analytics expertise — for example, finance, the customer service center, marketing and sales, supply chain, operations, pricing and product management.



Today most marketing operations teams rely on business intelligence and analytics tools, as do technical support, supply chain, and financial services. In short, analytics that is pervasive is a big part of your competitive advantage. Migrating to the TIBCO Data Science platform will affect data analysts and users in all of those groups. But more important, if analytics is already ubiquitous at your company and embedded in your systems and processes, there are many more people who consume and rely on the output. Any adverse change to this output could drastically impact your business.



How do your people use analytics?

In almost every company, analytics is finding its way into marketing to help make sense of what customers and prospects are saying, tweeting, looking for, and buying. We use analytics to personalize offers, attract prospects, and retain customers. We study things like customer churn, cross-sell/upsell opportunities, customer sentiment and satisfaction.

In daily operations, companies apply analyses to improve the quality of manufacturing processes. In professional services, service teams embed it in customer solutions as part of service engagements. For customer support of products, analytics for predictive/ prescriptive/preventive maintenance might be used to improve up-time and service.

We also use analytics to assess and control the potential cost and risk of our decisions.

Many companies rely heavily on analytics for modeling, assessing credit risk, and detecting fraud. Their models are closely tied to forecasts and bank rates, so statistical analysis is part of what these teams do day in and day out.

Analytics is used to make decisions. But since not all decisions are perfect, analytics is also used to assess and control the potential cost and risk of decisions.



Analytics use case: Customer churn

Suppose that a logistic regression model tells us we're in danger of losing five customers to churn. We decide to offer each of them a \$50 discount on a product, and they take it. So it costs us \$250 to keep them as customers. That sounds like a good use of \$250. But suppose it wasn't the right model to use, and the customers weren't really in danger of churn. Then we've wasted \$250.

With the right model and analytics, it might only cost you \$50 to save \$500.

Now look at the flip side. Suppose a random forest model tells you that a particular customer is happy and not in danger of churning, so you don't offer them a discount. If it turns out the customer buys another company's product, you might lose \$500 of revenue. With the right model and analytics, it would have cost you only \$50 to save \$500.

So, you might say that the random forest is a better model because it results in just one error, but it's a more costly error than the five combined from the logistic regression model. Models need to take that into account, and these trade-offs of model accuracy and costs associated with incorrect predictions are easy to illustrate with the TIBCO Data Science platform.

Analytics is also used to evaluate models based not only on the overall error rate but also on the type of error that the model is making. It allows you to make more-precise, better-informed decisions about the models you apply to business factors such as customer churn, product upsell, service-level agreements (SLAs), competitive pricing, and delivery dates.

Analytics use case: Direct mail campaign

Assumptions

- \$2 to mail each prospect
- •1 out of 100 will buy
- \$220 profit for each response



The value of a prediction

Assumptions

- Analytics model output:
 - 25% of the entire list is 3× more likely to respond



Adapted from Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie or Die by Siegel



How will people in your organization react?

You'll know the answer to this question in almost no time, because word travels fast. Most people's reactions land in one of three buckets:

1. "But we've never used TIBCO Data Science."

Since your users have been using the legacy analytics platform for many years, they aren't familiar with how robust an analytics platform like TIBCO Data Science software is, so naturally, they are skeptical. You pay them to be skeptical.

They know their work consists of mission-critical analytics in the previous product and assumed (incorrectly, as it turned out) that the TIBCO Data Science platform wasn't up to it.

2. "We've spent years writing thousands of lines of code. We don't want to just throw that away."

Users anticipate a great deal of work in trying to replicate in the TIBCO Data Science platform the programs they have built in your legacy product, so naturally, they balk. Who wouldn't feel that way?

3. "We consider ourselves the legacy company's professionals and analysts first, and employees second. For career longevity and our ability to do our jobs, we believe that it's really important to continue using the current platform."

That's a tough one. A number of heavy users who have been working with the legacy product for many years are comfortable using it, and they have grown, evolved, and become pretty good with it over much of their career. Asking them to switch to something they don't know is a huge disruption for them.



Most users have never heard of TIBCO Data Science analytics — let alone used it — so it is the devil they know versus the devil they don't know. Plus, many feel an emotional attachment to a tried and true product.

You can address their reactions by having your TIBCO Data Science migration leads sit down and show them that their long years of work would not be simply discarded. The leads examine the techniques and functions your users have worked with in the legacy product, such as K-means clustering, polynomial regression, GLM, ARIMA, and neural networks, then demonstrate how to replicate and enhance them in the TIBCO Data Science platform. Nearly all the techniques they use are easier to implement in TIBCO Data Science software without the need to write thousands of lines of code. Users simply drag and drop icons onto the TIBCO Data Science workspace in its easyto-use graphical user interface (GUI). You'll see that making the migration a

success requires consistent communication, executive buy-in, technical support, training, and encouragement, mostly from peers.

How will you get them on board?

Put the legacy product in its proper context. Explaining your migration project is one thing; getting all your users on board with it is another.

Why use a Ferrari to deliver a load of dirt? You might find that, for lack of a better (and lower-priced) tool, many people have grown accustomed to using the legacy product in cases where it wasn't the best fit and where it was overkill. They may Your users and business units may say they need more than a year to migrate have applied it to functions such as data movement (extract, transform, load [ETL]), successfully to the TIBCO Data Science platform. There are several approaches to and data aggregation and preparation, which don't require advanced analytics make the project palatable to users: technologies.

Make sure you have identified all current users and properly communicate the change to them.

By centralizing data movement tasks in officially supported environments and Nobody wants to find out about something like this at the water cooler. Your migration processes, you lower internal risk and improve performance. In cases of ad hoc data project team should compile a list of active users of the legacy platform and identify movement, migration leads can show users how to accomplish the same task with the affected executives and managers who would be responsible for keeping the users TIBCO Data Science analytics or with tools such as Toad Data Point to get them on apprised of project status. Be mindful that many users will initially regard the change board with the migration. as a difficult one, so keep communications consistent, positive, and frequent so they can digest the information completely. Holding several workshops early in the project Launch a contest. to field and address users' concerns is often a great help.

Give people early access to TIBCO Data Science analytics.

It's easy to overlook this as a migration project starts up, but as the message goes out and you talk to group managers about the timeline, they naturally will want access to the tool if they are expected to hit the schedule you've planned. Show them and their users what they are in for and let them start working with TIBCO Data Science tools. That exercise will kick off the next set of conversations about setting up infrastructure, rolling out the product internally, and paying for it.

Allay the concerns of the dyed-in-the-wool legacy company professionals.

Using a new tool does not strip people of their expertise or marketability, but they consider themselves part of a community that extends beyond your company. You may not be able to address that on an emotional level, so address it on a technical level by deeply discussing their analytics tasks and demonstrating how the TIBCO Data Science offering can satisfy them.

Guide users away from inefficiencies and shadow IT.

To inspire teams to take TIBCO Data Science software with both hands and apply it, consider running an internal contest. In these situations, teams of one to five people have been known to spring up around the company, build analytic models to address real-world questions, and compete against one another. The contest can attract a lot of attention to TIBCO Data Science tools but also helps you solve business problems.

Every organization is a different ship, steered differently. You can try a mindset like "Between now and the end of the year, we're switching to a tool you don't know and which you can't yet access, and in the meantime, you have to do your day job," but often a carrot works better than a stick.

If your migration enjoys the support of your executive leadership, whenever your migration team encounters a substantial hurdle, its members should get the support they need. In general, you can overcome most of obstacles with additional training, support, and one-on-one sessions customized to the teams that need them.

In short: communication, communication, communication. Add executive support, cultivate the employees who are not afraid of the challenge, and persist.



Who gets to be an exception, and why? Lesson learned.

When you look back, you'll see where you had to change your original plan and how you had to change it for the people factor.

There are plenty of lessons about how people deal with an extensive tool change in an organization. The bigger lesson, though, is about allowing for exceptions on a migration project.

For example, the financial team might be able to make the case that they need more time to migrate their most sensitive functions, and so they might receive an extension to continue using the legacy product for some months. More important is that they demonstrate that any interruption or fluctuation in their models would introduce greater risk than it would be worth to meet the initial deadline.

That makes sense. Some groups use analytics to predict, say, staffing levels for IT call centers. If their predictions are wrong, customers may have to wait longer, but the immediate impact is not as prominent as a hiccup in financial modeling. It might be unfortunate that not every user is on the TIBCO Data Science platform by your deadline, but the downside might be even more unfortunate, so the project team should work with finance (or whatever team is asking for more time) on a specific plan that extends the migration timeline for that small group.

Other than that, it speaks well to the quality of employees (and of the product) that the transition is smooth for the vast majority, especially considering that you may be asking them to achieve migration in a timeline that many think will take longer. Help people over their emotional obstacles to using a different tool in a different way. What matters is arriving at the result.



Chapter 2: Process



How do you eat an elephant?

One bite at a time, of course.

That may not be the exact question you ask yourself as you move from thinking about people (as we described in Chapter 1) to thinking about process, but the questions in this phase boil down to that.

How are you going to take all of the people in the organization who depend on your analytics platform — possibly hundreds of users worldwide — and move them to a new and different platform?

You can avoid the shock of turning off your legacy analytics product one night and then turning on the TIBCO Data Science environment the next morning. In general, to help ease the transition and ensure that your analytics processes keep running without disrupting the business, consider a high-level plan to stand up a parallel TIBCO Data Science environment and provide users what they need to move onto it.

You can run the two platforms in parallel for several months so users can see that, with training and support, they will have the time, opportunity, and motivation to move their work into the TIBCO Data Science platform. In fact, you can make it easy for them to compare the two products, satisfy their own curiosity, and develop their analytical muscles. Clearly communicate a goal date for them to stop using the legacy product altogether and start using the TIBCO Data Science solution exclusively.



One plan might be to stand up a parallel TIBCO Data Science environment and provide users what they need to move onto it.



Where do you start?

Imagine that your company has been using a product from Tool Supplier A for many years. It could be a chemical, an electronic device, or a gas-powered turbine — anything you and your co-workers use to get your work done. You check your email one morning in mid-March and discover that your company has acquired Tool Supplier B, which sells a competing product.

"Hmm," you think. "I wonder if that means we're going to stop using Product A and start using Product B."

Around April 1, you stop wondering because executive staff has decided that the entire company is going to migrate to Product B by the end of the year.

Nobody looks forward to the perceived learning curve of moving from a tool of long-standing use to a new one that few employees have ever used.

You may not have to switch that quickly, but you will have to figure out the best place for your organization to start. In most cases, that will be to develop timelines for rolling out the software, hardware, and infrastructure you'll need, and for putting in place the education and training, centers of excellence, user community, and help desk resources you'll need for success.

How do you schedule IT delivery?

Your mileage may certainly vary. Here's an example timeline: Laying the groundwork

April — Where are all the pieces?

First, engage every available resource from the TIBCO Data Science team to help your IT staff understand the software configuration, infrastructure, and security protocols for the hardware on which you will run the TIBCO Data Science platform. Then work out which IT skills and resources you need to put all of those in place.

Your technical and IT staff should spend a fair amount of time with the TIBCO Data Science development team to understand how TIBCO Data Science software is going to interact with all of the other software your company runs. A lot of April will be spent developing a delivery schedule for the hardware resources.

Begin identifying the analytics user base, the "customers." Who are they? Where do they work? Who are the contacts you need to work with so you can start putting together a customer plan for the migration?

May — How long will this take?

Once you identify the users, ask how long they think it would take them to migrate. If yours is like many companies, most will reply, "Let me start using TIBCO Data Science analytics, and I'll let you know. But if I had to guess, I'd say it will take at least a year."

The question will kick off a couple of weeks of dealing with purely logistical questions from the user community: Who's going to pay for this? How do we get the hardware? Which data center will host them? Do we need multiple data centers? How will we accommodate off-network users? From those answers, come up with an estimate of the hardware needed for the migration, then secure approval to proceed based on the estimate.

By the end of May, a solid plan should emerge so you can present to IT and business leaders and say, "Here's what we are going to do."



June – How many users really need to migrate?

The IT team will need June to iron out technical details around security and configuration.

Meanwhile, on the business side, perform a valuable exercise in rationalization with the centers of excellence (CoEs) for each of your business functions — finance, professional services, customer support, marketing operations, and sales. You might ask the CoEs to identify their analytics users. But many users run analytics for non-analytic features like data management and ETL, and others may not actively be using analytics anymore.

Ask the CoEs to examine more precisely which users really need analytical functions and confirm validated candidates.

Migration

Training continues, CoEs established, users onboard

August-December

Deadline

Installation and training

December

Migration complete (extended to June 30 for 16 users in Finance)



Knowing exactly how many people need to migrate and who they are will help you estimate the number of servers you need, how to configure them, and which groups you will have to set up for a successful transition.

July — How fast can you get the hardware up and running?

With approval to purchase and install the hardware and configure the software for delivery to users, you can start the procurement process with your eye on the deadline. In this example, we are using late August as the target date. B sure to let everyone in the supply chain know the urgency of the project.

When your colleagues really put their collective minds to a task, you may be surprised and relieved. You need to secure the physical servers, install them in a data center, get them online, and configure the software so the TIBCO Data Science platform is available for delivery to users. Whatever your schedule, transparent communication of the plan, the timelines, and the urgency of meeting specific deadlines will help everyone. When you empower people with information, they are more willing and able to do really good things.



How do you train and support users?

With hardware and software in place, you can put TIBCO Data Science software in the hands of your users, but don't forget that you need comprehensive training and technical support for them to successfully migrate to a new technology, especially when you're on a tight schedule.

Training

Some teams don't wait to be invited for training; they get a jump on it.

They know the TIBCO Data Science system will replace the legacy product and begin thinking about ways to use it. People have been known to obtain a TIBCO Data Science trial license and invite regional TIBCO Data Science experts to their offices to provide an introduction to the product's capabilities well before IT gets it up and running. Using that foresight, they get an informal orientation long before the formal migration project was announced.

On the example timeline we're using (deadline is December 31st), beginning in mid-July, have TIBCO Data Science subject matter experts (SMEs) facilitate three rounds of training:

- Train-the-trainer sessions Ask each of the teams that will be using the new platform to send two people for a two-week training session in July, while IT is still installing hardware and configuring software. TIBCO Data Science SMEs can instruct users from all over the world in train-the-trainer sessions. The goal is to create champions and advocates in different groups around the company and around the world.
- In-person, deep-dive training Later, to meet the particular needs of remote users, conduct a formal session.
- Web conference You can also record virtual sessions and make them available as knowledge transfer resources for all migration candidates. The recordings remain available to all appropriate employees.

The train-the-trainer and in-person sessions give attendees their first opportunity to learn about and use the TIBCO Data Science platform, even before it is available at their desks.

Technical support

As more users begin adopting TIBCO Data Science analytics, you might create a separate, specific site (SharePoint is one way) where migrating users could submit support issues or questions. You can build up a knowledge repository with frequently asked questions and how-to articles, and the TIBCO Data Science team can assign several sales engineers to provide answers, both electronically and in person.

One method is to keep the information separate from the standard tech support queue. There are three reasons why you want to do this:

- You can apply an SLA and expectation of response different from those applied to the standard queue.
- You can monitor the queue, escalate urgent items, and keep an eye out in case a significant need for widespread training should suddenly come to light.
- If a group has questions or concerns about a particular TIBCO Data Science feature or area, you can switch from technical support to a model more like a professional services engagement and a workshop.

With IT assets, training, and support in place, you are ready to embark on the actual migration work.



What goes on during the actual migration?

The important thing to keep in mind is that the process of

migration happens behind the scenes — almost invisibly — as users steadily adopt the new tool and cease using the old one. In the example case we're using here, although IT manages the project and keeps executive staff apprised of overall status, it is the business leads in the CoEs who know how close or far each user is from completely migrating to the TIBCO Data Science tool. Because of all the moving parts inside the actual migration, this might be difficult to discern, even for several months.

Here are some possible scenarios you may encounter:

You bump into resource constraints.

During migration, users spend time converting their daily tasks — processes, models, algorithms, and so on — to the TIBCO Data Science environment while continuing to perform their normal functions with the legacy product. These employees have already allocated 75 to 90 percent of their time, so when asked to take on the additional work of migration, they begin to wonder (aloud, mostly) if they can meet deadline.

You spend time double-checking.

Users run tasks on both products to ensure they are getting the same result from both tools. (That's the healthy skepticism we hire them for.) They know how their models and outputs should look in the old product, and they figure out how to get them to look that way in the TIBCO Data Science system as well. That is the result of all the resources you dedicated to the project, such as training, technical support, and informal knowledge transfer.

You herd cats.

Individual business groups run their own analytics business processes, so they formulate their own migration plans and start executing on them at their own pace to meet the deadlines. Of course, many of the timelines and migration plans were developed with input from the TIBCO Data Science SME team.





You stop to correct inefficiencies.

The last thing you want to think about in the middle of migration is process improvement, but inefficient processes come out of the woodwork, and you have to deal with them. Let's say, data extraction tasks are discovered that take hours to accomplish with existing tools because they had never been properly configured or had not been updated for changing requirements. Correcting inefficient processes was probably not initially part of the migration plan, but you can't ignore them, particularly when the analytics innovation team sees the chance to whittle a four-hour data extraction process down to a few minutes with the TIBCO Data Science tool, for example.

You eliminate over-dependence on the incumbent software platform.

As described above, CoEs may discover that many users have been using the legacy product for non-analytic functions such as data management and ETL. Users may soon find that the TIBCO Data Science system does not yet support some of these data management functions as well or in the same manner as the legacy programming language.

Besides showing the product development team features to add or enhance in future versions, this type of discovery can open up another migration opportunity: steer users toward products designed for data management and ETL. Purpose-build tools can allow teams to meet their deadline for migrating off the legacy product at a fraction of the cost.

You run a contest.

To encourage your power users to showcase their talents, you might host a contest in which your analytics teams submit business cases with real-world examples of how it is easier, more appropriate, or more robust to deploy certain analytics processes using TIBCO Data Science software. Teams can develop advanced analytics processes and workflows that highlight their migration story through production applications, including customer lifetime value, data manipulation, pricing optimization, and predicting rates of hard drive failure.



How do you stick to the deadline?

The middle of the timeline is often the smoothest part of the process. It's usually at the beginning and at the end of the project where you encounter the most resistance.

With a December 31 deadline, the biggest obstacles will be the U.S. holidays reasons for not relinquishing until necessary, so IT can probably allow them to retain of Thanksgiving (late November) and Christmas. People will be planning holiday access until the deadline. absences, so that one-week tasks and approval cycles are going to start dragging As described in Chapter 1, 16 financce users obtained permission from the company's out to two- and three-week delays. Users still have to access both products, and it's executive staff to continue using the legacy product through the following June. common for many to be reluctant to relinquish licenses to the legacy product until they absolutely have to. To avoid blowing your deadline, appeal to the CoE leads The main lesson about sticking to the deadline: If nobody is telling you that you have during early December. to stop using the incumbent analytics software platform right now, and if retaining

Pressure from the business leads

Tell the CoE leads, "Your people are doing great, but the number of users who have not yet migrated and relinquished access voluntarily is awfully high considering how little time remains on the project plan."

Ask them to identify a lead on each functional team to tell their users pointedly, "If you're using the TIBCO Data Science platform and you haven't touched the old product in a while, then I need you to give up your access."

You will probably find that most of those users simply say, "OK," and relinquish their licenses. During the two weeks after Thanksgiving, expect a dramatic increase in the number of users switching to exclusive use of TIBCO Data Science analytics. By mid-December, hopefully the vast majority of validated migration candidates will have relinquished access.

Pressure from IT

IT project management can also make phone calls to the final group of users to obtain specific reasons why they want to retain access until the December 31 deadline.

Holdouts might be regional offices that plan to remain open during the last week of December and continue working on year-end analysis. They have sound business

access to it doesn't appear to cost anything, then naturally you'll keep it in your back pocket until the very end, just in case. This eleventh-hour cleanup operation is an integral part of sticking to the deadline.

What else do you learn during a migration project?

You learn many of your lessons while the project is still und way. Others don't occur to you until the dust settles and people begin talking about the project in the past tense.

Unexpected dividends

As described in "What goes on during the actual migration?", a great deal more efficiency and process improvement than anticipated may be realized. Migration projects include unexpected opportunities to turn over a lot of rocks and scrutin how you have been doing things for a long time. You didn't specifically set out to make a better company, but when the chances arise, you can't pass them up.

The importance of CoEs

IT management of a migration project is usually the method used when software a hardware are involved. IT may be responsible for knowing where each team was in process of relinquishing the old product, but not for knowing status at the user lev

If you have analytics and business intelligence CoEs within finance, professional services, customer support, marketing operations, and/or sales, IT project manages should ask each CoE to make a single migration lead accountable for its populate of users. The CoE enables IT to break the user population down into smaller unit have the organizations that depend on analytics drive the migration.

If you're migrating more than a handful of users, you need an organizational structure that is segmented into manageable sized user groups. CoEs can do that because they understand what their users do with the incumbent product, what they wish they could do with it, and what they need in the new product. While IT deploys the application and puts infrastructure in place, the CoE handles the migration at the user level, a task for which IT is not set up.

ize	
	As described above, many users perform double duty by not only building new models and processes in the TIBCO Data Science system, but also maintaining existing models and processes in the legacy product. During the train-the-trainer phase of the formal education process, some users perform triple duty.
	For train-the-trainer sessions, each business unit should identify a small set of power users who will undergo in-depth training. These trained users interact regularly with TIBCO Data Science SMEs and, in theory, are the first line of reference for general questions within their business or functional areas.
	In practice, however, teams may still need support from the TIBCO Data Science SMEs for three reasons:
and h the /el.	 Some teams are not physically co-located, so the trained users have to help co-workers in other offices or time zones. When applied to an intricate analytics modeling problem, even a software solution as robust and easy to use as the TIBCO Data Science platform does not always offer a simple solution.
gers	 The trained users also have to perform their day-to-day tasks while migrating
tion tis and ticture e h	Ultimately, you may want to conduct formal classroom training for all users (see the section "How do you train and support the users?"). Train-the-trainer sessions are a good idea, but it's easy to underestimate the amount of effort required for these individuals to learn the new product, act as the primary TIBCO Data Science point contact for their team, and perform normal analytics tasks while migrating.
Э	

Users with three iobs

Chapter 3: Technology



What will your new architecture look like?

By June, when you specify the architecture for your TIBCO Data Science implementation, estimate on the high side for several reasons:

- Assuming your project goal is to migrate hundreds of active users by December 31, time is of the essence.
- Not only will you want to avoid bottlenecks from inadequate system resources, you will also want to take full advantage of new hardware.
- Once word spreads around the company about the product's ease of use and analytics capabilities, envision a steady increase in the user population much greater than the user population that had licenses for the legacy product.
- Beyond the basic configuration, real-time TIBCO Data Science Monitoring and Alerting Servers (MAS), and servers for data management and data manipulation, may be needed.

See "What is the right amount of infrastructure? Lesson learned." for minimum infrastructure requirements.





Here's how you did it before. How do you make it work in the TIBCO Data Science platform?

Let's say your legacy analytics product was codeoriented, and many of your users spent years developing and maintaining code for their projects. Conversely, TIBCO Data Science analytics is GUIoriented, and your users fear they will not be able to perform their duties without using the legacy code they had written. With hundreds of analytics projects across the company, it is understandable that some users will have concerns.

However, there is no need to rewrite code in all projects:

Conduct a survey of procedures currently in use Assemble a team to send out an internal survey to discover which procedures analysts use on the legacy product and how frequently they use them. The results will vary, such as enterprise exploratory data analysis (EDA) and analytic modeling (see Figure 1).

Map existing procedures to TIBCO Data Science functions Have the team correlate each existing procedure to the corresponding process in the TIBCO Data Science platform. The resulting map will represent the procedures they need to understand and use.

EXPLORATION FUNCTIONS	% OF USAGE	MODELING FUNCTIONS	% OF USAGE
Means	86%	REG	65%
Freq	81%	LOGISTIC	61%
Contents	73%	ARIMA	40%
Summary	72%	SURVEYSELECT	33%
Univariate	67%	GLM	27%
Corr	58%	TIMESERIES	25%
Tabulate	28%	ARIMAX	20%
Gplot	21%	MIXED	11%
SQL	7%	PHREG	11%
Transpose	3%	MODEL	11%
Print	1%	MI	10%

Figure 1: Breakdown of frequently used functions

REG	65%
LOGISTIC	61%
ARIMA	40%
SURVEYSELECT	33%
GLM	27%
TIMESERIES	25%
ARIMAX	20%
MIXED	11%
PHREG	11%
MODEL	11%
MI	10%
GLMSELECT	8%
МСМС	5%
GLMMIX	4%
X12	4%
UCM	4%
FASTCLUS	2%
FORECAST	2%
VARCLUS	1%

Validate use cases in a test environment

After training analytics SMEs for about two days, ask the team to take a sample of documented, current use cases and verify that they run correctly in a TIBCO Data Science test environment. In certain instances, this requires an internal SME on the old product and a TIBCO Data Science SME to interpret functional capabilities between the products.

Set up TIBCO Data Science production environment

When the software and hardware for TIBCO Data Science analytics in place (in this example, by August 1), the migration team can set up its production environment, obtain access for users, and begin migrating the identified projects to the TIBCO Data Science platform.

Move projects one by one

The team should chose the most efficient tool — TIBCO Data Science software, SQL, open source R, Toad — and begin the process of converting and testing, gradually moving each project over.





How do you perform data management and manipulation?

Whichever data repository you use—Teradata, SQL Server, MySQL, or other data source—you still need to aggregate, cleanse, and prepare data before you can perform analysis and modeling on it.

As described in the "What goes on during the actual migration?" section in Chapter 2, for much of their data management and manipulation steps, many users have come to rely on non-analytic functions in the previous product. That is an expensive proposition for relatively common functions. There is probably room to refine, optimize, and automate years of legacy code written by many people.

You want to use external tools designed specifically for data integration and data management, rather than try to embed them in your analytics platform. You also want to use GUI-based rather than programming-based products. This approach will likely prove the best choice because you now have data integration experts focused on data management and analytics experts focused on analytics. You may want to do this in two separate stages:

- First, data management and manipulation outside of the TIBCO Data Science platform, handled by data integration experts
- Then data import for analysis and modeling using the TIBCO Data Science platform, handled by analytics professionals

This structure keeps experts, products, and responsibilities in their respective wheelhouse.



Should you migrate users from other tools in the same project?

You run the risk of distracting yourself from the original goal when you decide to extend your migration effort to other problems you discover along the way. But as described in "What goes on during the actual migration?" in Chapter 2, extending the effort can pay off handsomely.

If the old product has been used (or misused) for data management for a lot of years, what other tools have people been using (or misusing) that you can consolidate or replace? Run a survey to find out about widespread usage of other tools across the stages of data management:

- ETL process automation (for example, Microsoft SQL Server Integration Services, Microsoft Visual Studio)
- Data extraction (for example, Microsoft SQL Server, D3 JavaScript library, Adobe Site Catalyst)
- Reporting (for example, Microsoft SQL Server Reporting Services, Microsoft Access)
- Statistical analysis (for example, JMP)

Perhaps you find that dozens of unique tools have been at work, a common inefficiency in large organizations. Identify the top tools for migration using appropriate technologies.

A TIBCO Data Science migration project presents the opportunity to steer users toward products better suited to data management and ETL.

Fit-for-purpose tools can help you meet your migration deadline, yet still cost less than typical licensing for similar capabilities on a legacy tool.





How do you onboard users throughout the project?

If you're accustomed to having the onboarding process resi in IT, you'll see a big difference during migration that will pa dividends long after your project is finished.

With a defined number of licenses available to business users, you probably had become accustomed to having IT grant access. Three burdensome restrictions become a way of life:

- "Sorry, no more licenses. We can't add anybody." A finite number of licens means a finite ability to onboard users and allow them to use the product. Unless someone makes an unpleasant decision about who must give up his her license, teams have to contend with this as yet another moving part in migration project.
- License swapping Even if someone does give up a license for the few week of the migration project, the team still needs to have IT deactivate and reacti the license. Every change represents another request and multiple touchpoin
- Adjudication When teams move completely off the old product, other use sometimes request the relinquished licenses. Not knowing who else in the company might suddenly want access to them, IT has to allocate the availa licenses carefully. That requires asking a high-value data scientist to determine whether the user's proposed work justifies the request for short-term access to a product you are leaving behind.

side ay	Unless you purchase the requested licenses to the old product, plus an undetermined buffer for safety, this model needlessly hamstrings centralized onboarding of the migration progress. During migration to TIBCO Data Science software and its flexible licensing model, those restrictions disappear.
	First, with the TIBCO Data Science platform, the model for onboarding users switches from being IT-centric to more of a self-service model with flexible licensing.
ses 5 or their	Second, we recommend making the business intelligence centers of excellence (see the "How do you schedule IT delivery?" section in Chapter 2) responsible for their own strategy for deploying advanced analytics to their users. This includes determining which users require access to applications like the TIBCO Data Science system. The change takes IT out of the business of determining why somebody needs access. Users now contact a person in their own business unit who understands the scope of their duties and decides who gets access.
ivate nts. sers	By pushing the vetting process out of IT and into the business units, you remove an unwelcome variable during migration, and in the long term, it's more efficient, internal customer satisfaction is higher, and the business unit is better off owning it.
able	



What is the right amount of infrastructure? Lesson learned.

A few important variables determine the answer to that question.

First, the architecture and infrastructure requirements for the TIBCO Data Science platform are different from those of older analytics products. The TIBCO Data Science system is designed to meet the needs of even small analytics teams, and it uses a fraction of the hardware required for most advanced analytics products. A basic TIBCO Data Science configuration can include these components:

- Microsoft Windows Server (64-bit) 2008 R2 or later with 2 GB of memory, 5 GB of disk space, and a dual-core processor
- Microsoft Internet Information Services, which comes with Windows Server
- A standard database, such as Microsoft SQL Server or Oracle, acting as the repository for TIBCO Data Science metadata

Users can work on even more modestly equipped Windows desktops.

Scaling up successfully to meet the needs of TIBCO Data Science users depends on how resource-intensive their collective workload is. Several factors influence resource requirements:

- Number of concurrent users
- Types of users (for example, analytics producers vs. information consumers)
- Data volumes
- Analytic complexity (such as simple linear regressions vs. an automated neural network)
- Desired and required response times

Conclusion

As in the real-world example presented, beginning with a decision in March, through December 31st, a successful migration to TIBCO Data Science analytics of hundreds of analytics projects and all but a small percentage of hundreds of users was accomplished with a relatively small migration team working worldwide.

As described in Chapter 1, a small group of users made the business case to continue working on the old product for a few additional months. They needed more time to migrate their most sensitive functions and demonstrated that any interruption of their models would introduce unnecessary risk. But after a few months, those 16 users moved entirely to the TIBCO Data Science platform.

No longer dependent on a legacy analytics product, the company—and you, too—can now realize a substantial, bottom-line benefit from saved fees.

This TIBCO Data Science migration project was a success because of the people, processes, and technology applied to it. Migration is a challenge requiring the attention, support, and engagement of C-level executives. Cooperation of business units and teams throughout the company is an integral part of making any migration a success.

You can do it. And we have the experience to show you how. You have nothing to lose except high licensing fees.

Learn more at www.tibco.com/products/data-science