



Applying the R Language in Streaming
Applications and Business Intelligence

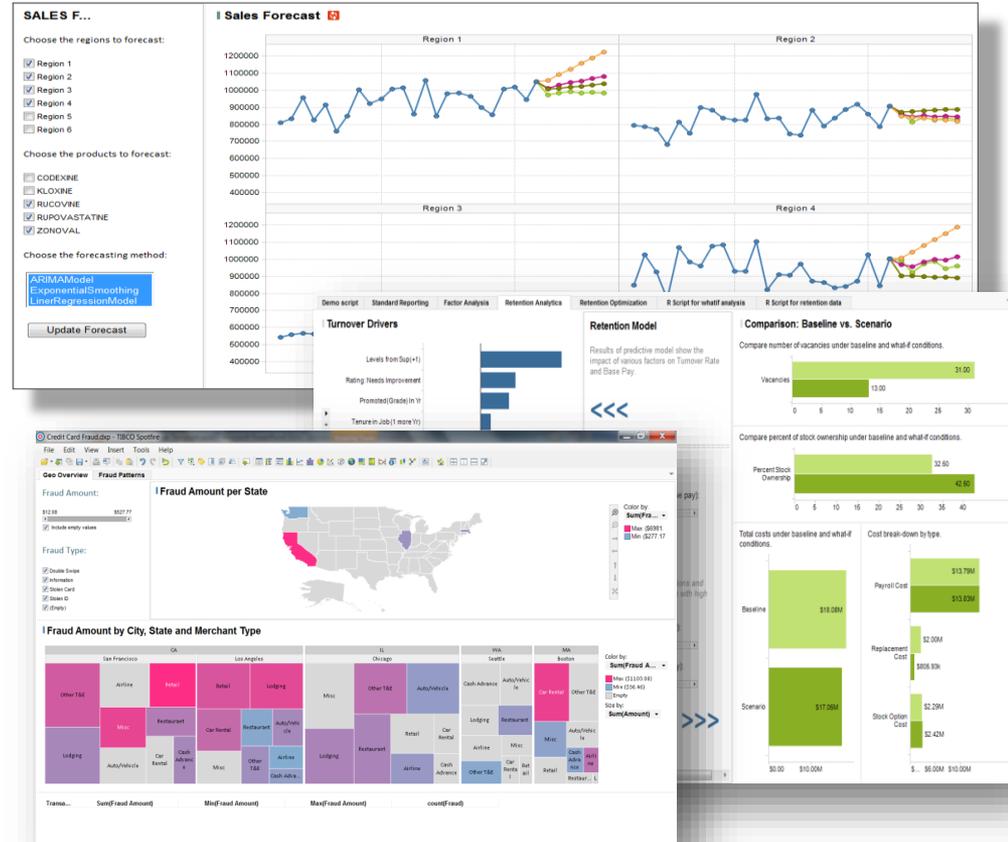
Lou Bajuk-Yorgan
Sr. Dir., Product Management
TIBCO Spotfire
lbajuk@tibco.com
@loubajuk

- TIBCO's History with S+ and R
- Challenges of R for Enterprise applications
- TIBCO Enterprise Runtime for R (TERR)
- Applications of R/TERR for BI and Streaming
- Learn more and try it yourself

- John Chambers developed the S language at Bell Labs
 - Starting in the mid 70's
- Insightful founded (as StatSci) to commercialize S as S+ in 1987
 - The “plus”: statistical libraries, documentation, and support
 - Focused on commercial users, ease of use, server integration
- R: development begun by Ross Ihaka and Robert Gentleman at University of Auckland in mid 90's
- Insightful joined TIBCO in 2008
 - Spotfire (for Data Discovery and Visualization) joined in 2007
- Focus on applying Advanced Analytics in Spotfire & TIBCO real-time applications
 - Embraced R language as part of a broader Advanced Analytic Ecosystem
- Ongoing commitment to the R Community
 - useR sponsor for 7 years running
 - Founding member of the R Consortium



- Spotfire Statistics Server
 - Integration of R & S+ into Spotfire applications
 - Later SAS® & MATLAB®
 - Leverage the interactive visualizations, data access and application development of Spotfire
- Well received—but our Enterprise customers needed more
 - Clear customer-feedback on the Open Source R engine’s shortcomings for Enterprise usage



- R provides tremendous benefits to statisticians
 - Enterprises are challenged to leverage that value
- Enterprises need:
 - R's expansive flexibility to develop the most powerful and relevant predictive models
 - Enterprise-friendly R
 - Better performance than open source R → faster, more memory efficient, greater scalability
 - Fully supported R environment
 - Tightly integrated into enterprise applications
 - Agility to use R through the entire prototyping-to-production process
- Vendors & application developers need:
 - An embeddable R engine that is not open-source GPL licensed
 - Alternative to loose integrations, manual configuration and poor user experience

- Unique, enterprise-grade statistics engine, architected from the ground up by TIBCO
 - Based on TIBCO's long history and expertise with S+
 - Better performance and memory management than open source R
- Designed for R language compatibility
 - Wide range of built-in analytic methods
 - Extensible through R community packages
- Designed for commercial embeddability
 - TIBCO licensed & supported product
 - Not GPL, not a repackaging of the Open source R engine
- TERR extends the reach of R in the enterprise
 - Develop code in open source R
 - Deploy on a commercially-supported and robust platform
 - Without the delay and cost of rewriting your code



Example 1: Embedded TERR in Spotfire

- Spotfire: Data Discovery and Visualization platform for Business Users and Analysts
 - Separate analytics platform, independent of TERR/R
- Easily enhance Spotfire analyses and applications with R language scripts
 - Extend the impact of the Data Scientist/R by making their analytic insights available to a wider audience

Register TERR Function

Function name: Result type:

Description: Performs a linear fit.

Custom Expression

Available properties for column: Functions Category:

TERR S

Write R code directly in Spotfire; TERR executes locally or on server

Market Segment: String, Product Category 1: String, Product Category 2: String, Product Category 3: String, Product Category 4: String, Ship Container: String, Ass Order Equip: String, Ass Pages Equip: String, Ass Panels Equip: String, Ass Pumps Equip: String, Document, Document, Document, Document

Expression: `[Gross Profit], lmfit([Sales Total], [Gross Profit])`

Recent expressions: `[Gross Profit], TERR, Real/Conv, Display name.`

Manage TERR analytics locally or in Server to reuse across community

Portfolio Optim_TERR.dxp - TIBCO Spotfire (6.5.0 ENGR 23, 64-bit RELEASE)

File Edit View Insert Tools Help

Introduction Power to Predict (The New TERR) Price/Return Analysis Portfolio Optimization

Please select the equity name for the x-axis of the correlation plot:

184.02 195.69

Correlation Plot

Color by: Return of Select...

Correlation Coefficient $a = -30.24$

Price Information

Color by: (Column...)

Deploy TERR-powered applications to the web

165 of 165 rows 0 marked 17 columns inputData

Cover Page The Data Predictor Importance Model Comparison

Using the model, TERR generates predictions by store and maximizes campaign revenue lift.

What If Analysis?

Number of Stores in Campaign (ranked by predicted revenue lift):
81

Campaign Duration (days):
4

Fixed Cost per Store (\$):
2000.00

Cost per Extended Offer (\$):
0.20

Cost per Accepted Offer (\$):
1.95

Offer Acceptance Rate (%):
1.25

Average gift_cards Purchase When Offer is Accepted (\$):
20.00

Model ID: h8t7q6z

Predicted Revenue

Predicted Cumulative Revenue

Edit Parameters

Name: [TERR] Merge Store Hierarchy and Predictions

Refresh function automatically Run Location: Default

Input Output

Input parameters:

Name	Description	Type	Required	Input	Data Types
store.hierarchy		Table	Yes	[Store Hierarchy ...	Integer, Real, SingleReal, String, Date, Time, ...
predictions.by.store		Table	Yes	[Scored Result by...	Integer, Real, SingleReal, String, Date, Time, ...

Input handler: Columns Expression None

Data table: Store Hierarchy with Predictions

Columns:

- TXN_PER_DAY
- STORE_CITY
- STORE_STATE
- LONGITUDE
- LATITUDE
- STORE_NM

Search expression:

Example input:

```
[Store Hierarchy with Predictions].[TXN_PER_DAY]
[Store Hierarchy with Predictions].[STORE_CITY]
[Store Hierarchy with Predictions].[STORE_STATE]
```

Help OK Cancel

\$46,002 Revenue Lift
Avg \$568 Lift Per Store

Customer Churn:

- Retain your most profitable customers
- Increase upsell, decrease churn

Fraud Detection:

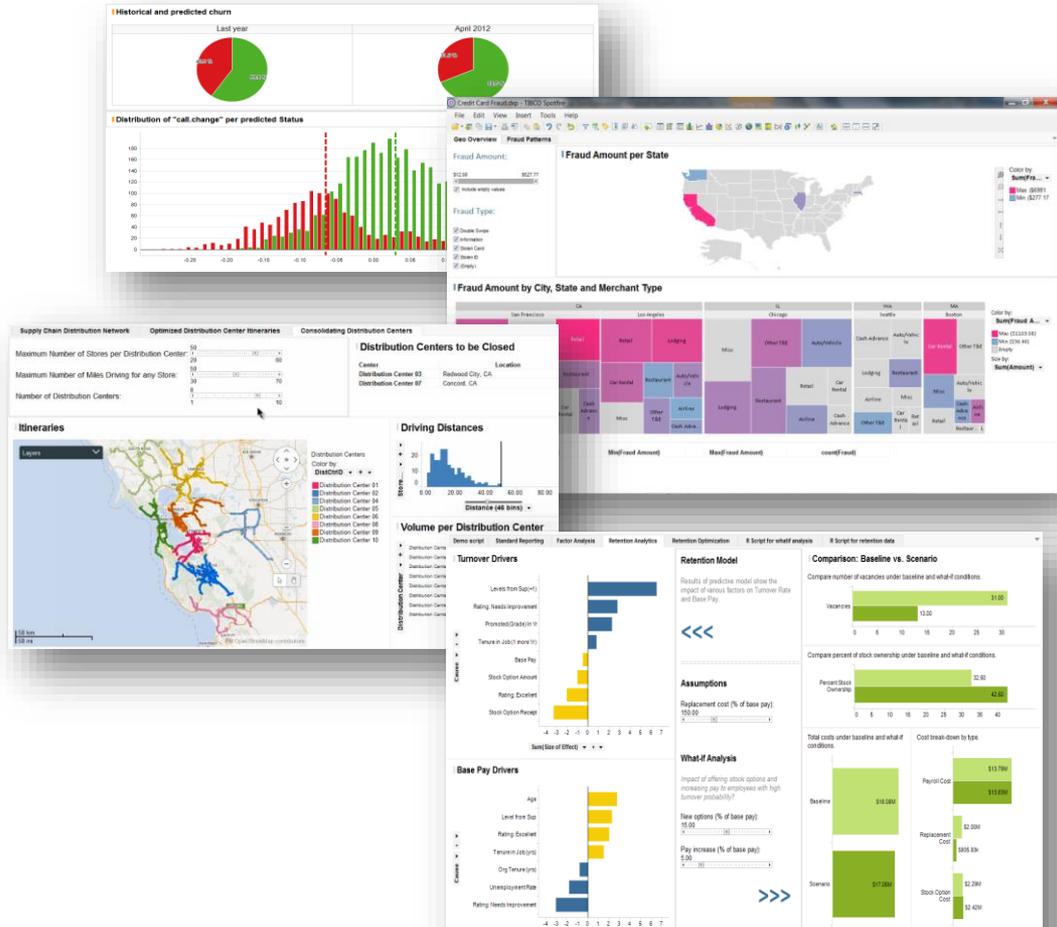
- Reduce losses due to fraudulent transactions

Supply Chain Optimization:

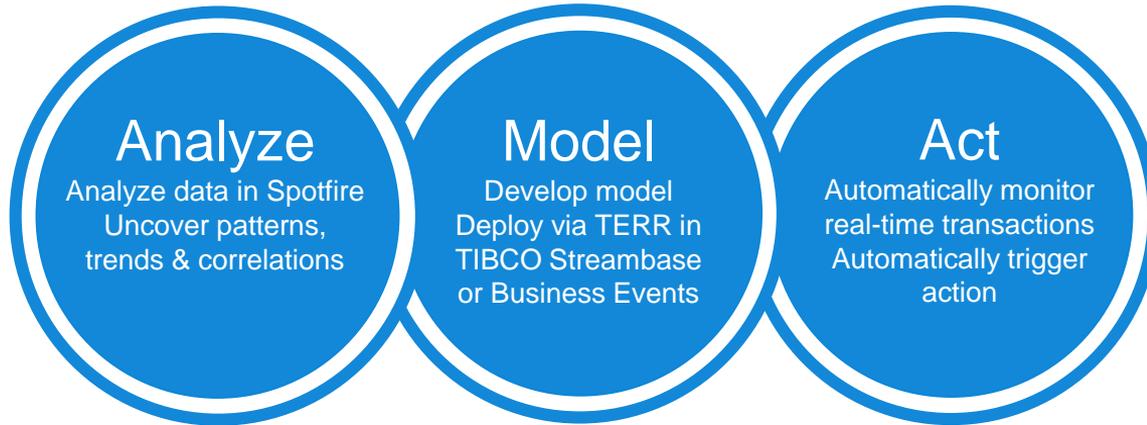
- Anticipate peaks and lulls
- Optimize distribution centers

HR Planning:

- Predict employee attrition and optimize retention



- TERR powers real-time advanced analytics in TIBCO “Fast Data”
 - When an event is identified, the CEP application applies a predictive model, and then can trigger an automated business process
 - E.g., extend a mobile offer to a customer; stop a fraudulent transaction in process



- Port Congestion Detection
 - Real time system triggers TERR
 - Analyzes port congestion
 - Recommends reduction of speed if no berths available
- Maritime Abnormality Detection
 - Based on Automatic Identification System info, TERR calculates likelihood of deviation from normal sailing routes
 - Alerts carrier & operator

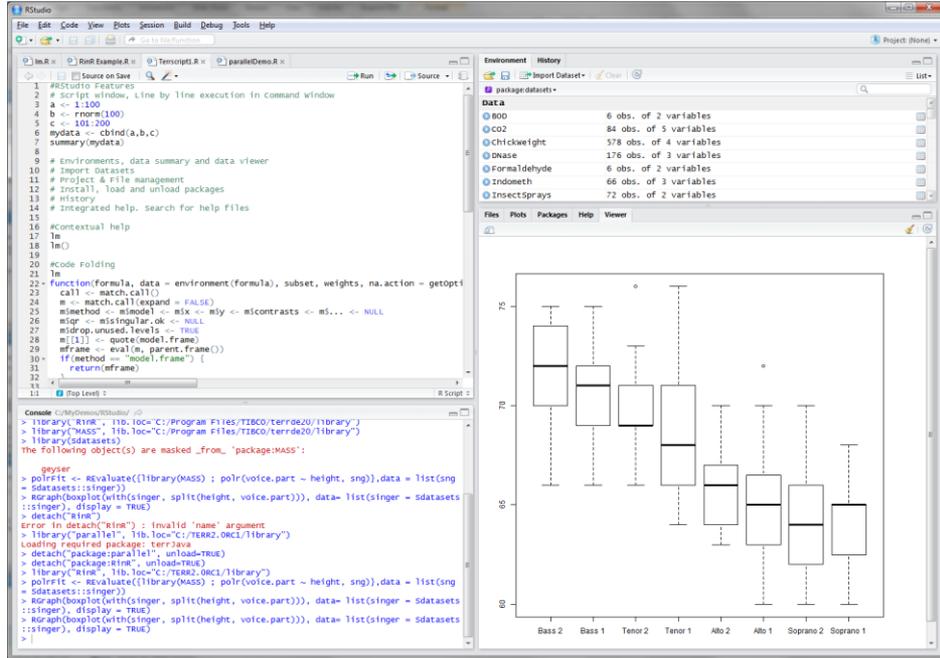
TIBC[®]

CargoSmart Uses Big Data to Transform Shipping
Processes and analyzes big data events to gain competitive advantage, valuable insight, and vastly improved customer service



- Oil & Gas Extraction
 - Maintenance Downtime and Equipment failures are costly
 - Engineers track sensor data to find leading indicators
 - Temperature, vibration, etc.
- Engineers usually use ad hoc rules on leading indicators
 - R/TERR used to develop predictive models for preventative maintenance
 - Deployed in real-time systems, alert when maintenance recommended

- Empower R users
 - Enterprise platform for the deployment and integration of your work—without having to rewrite it!
- TERR Developer Edition
 - Full version of TERR engine for testing code prior to deployment
 - Compatible with RStudio & ESS Emacs
 - Free for non-production use
 - Supported through Community site
 - Available at Tap.tibco.com



- TIBCO makes the power of R available across the enterprise
 - Ad-hoc Predictive Analytics for Data Scientists
 - Spotfire & TERR
 - Guided Analytic Applications
 - Spotfire & Statistics Services predictive ecosystem, starring TERR
 - Event-driven, real-time analytics
 - Spotfire Event Analytics, TIBCO CEP & TERR
- TERR is R for the Enterprise
 - Enables customers to develop code in open source R, then deploy their R code on commercially-supported and robust platforms—without recoding, saving time & money
 - Enables TIBCO, partners & customers products to tightly & efficiently embed R language functionality

- [TERR Community](#) at TIBCOmmunity.com
 - Resources, Documentation, FAQs, Forums
 - More info at spotfire.tibco.com/terr
- TERR Developer Edition
 - Full version of TERR engine for testing code prior to deployment
 - Supported through TIBCOmmunity, [download via tap.tibco.com](http://tap.tibco.com)
- Spotfire Free Trial: <http://spotfire.tibco.com/trial>
- Presentations: <http://www.slideshare.net/loubajukyorgan/presentations>
 - Will post slides to twitter @loubajuk #user2015
- Stop by our booth
- R Consortium Founding Member www.r-consortium.org

- Lavastorm Analytics

- Visual workflow tool for data management and analysis
- Embedding TERR for R scripting and predictive tools

- RStudio IDE

- Free, open source IDE widely used by the R Community
- Fully compatible with TERR Developer Edition

- KNIME

- Free, open source workflow tool for data management and analysis
- TERR fully compatible with KNIME Interactive R Statistics Integration nodes

