



From Insight to Action:
Analytics from Both Sides of the Brain

Michael O'Connell

Chief Analytics Officer


moconnell@tibco.com

@MichOConnell


about.me/moconnell

Insight to Action – from Both Sides of the Brain


Value



Grow Revenue



Reduce Risk



Increase Productivity
ROI

Both Sides of the Brain

- Fast & Slow

Insight to Action

1. Visual Analytics
2. Numerical Algorithms
3. Insight Execution

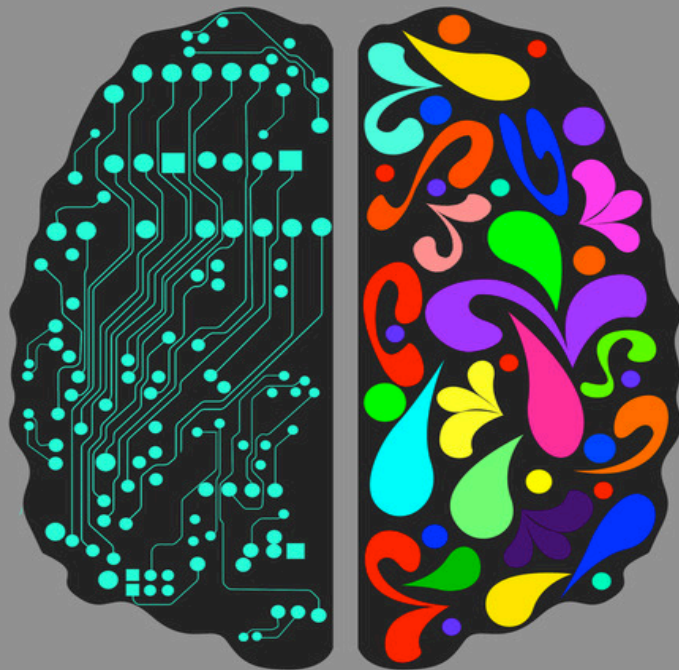
Insight to Action Case Studies

- Connected Equipment; IIoT
- Hi Tech Manufacturing

Come See the Demos

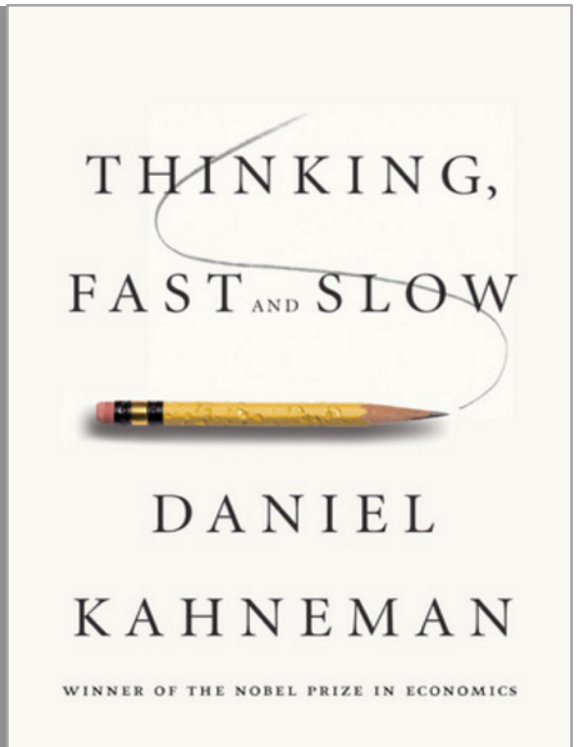
- Exhibition Hallway

Themes: Thinking Fast and Slow... from Both Sides of the Brain

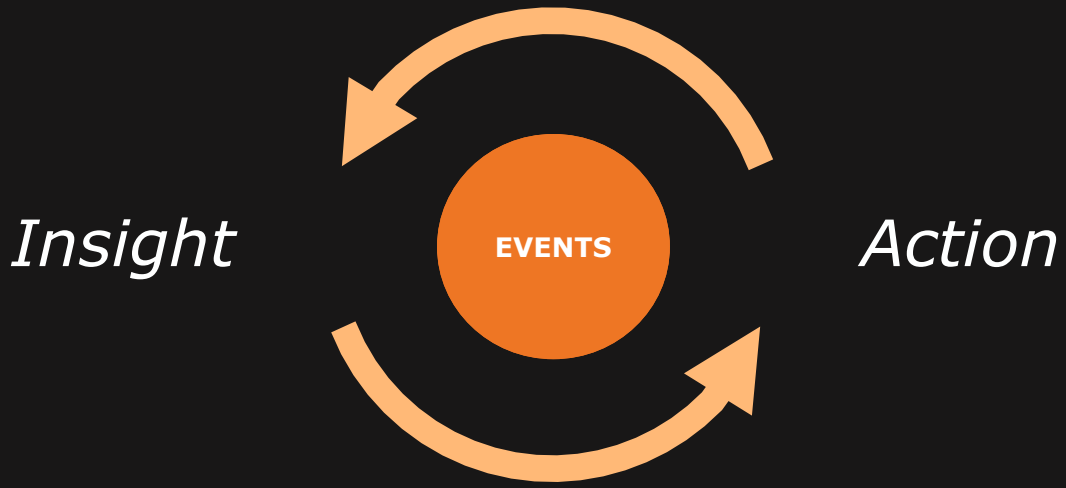


System 1: Association Engine

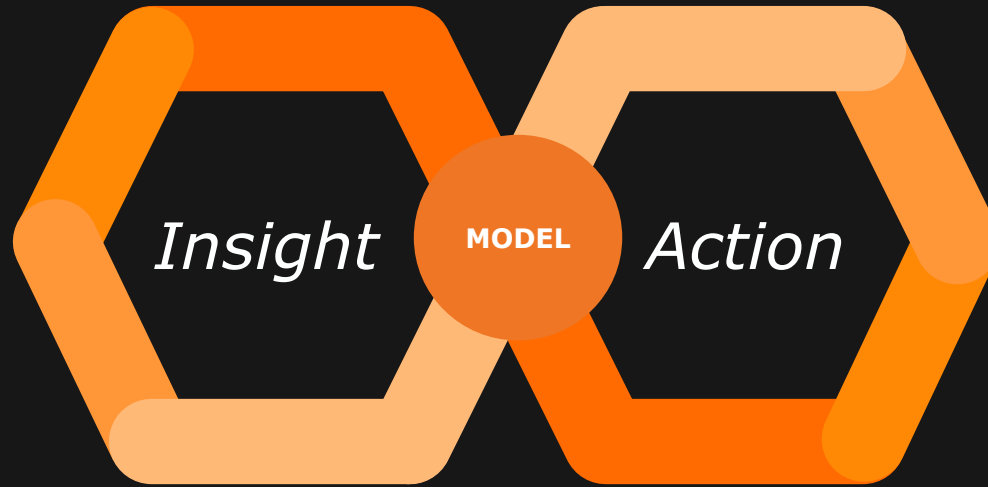
System 2: Monitor & Control



■ Making Sense of the World



■ Making Sense of the World – Some Key Steps



Smart Visual Analytics

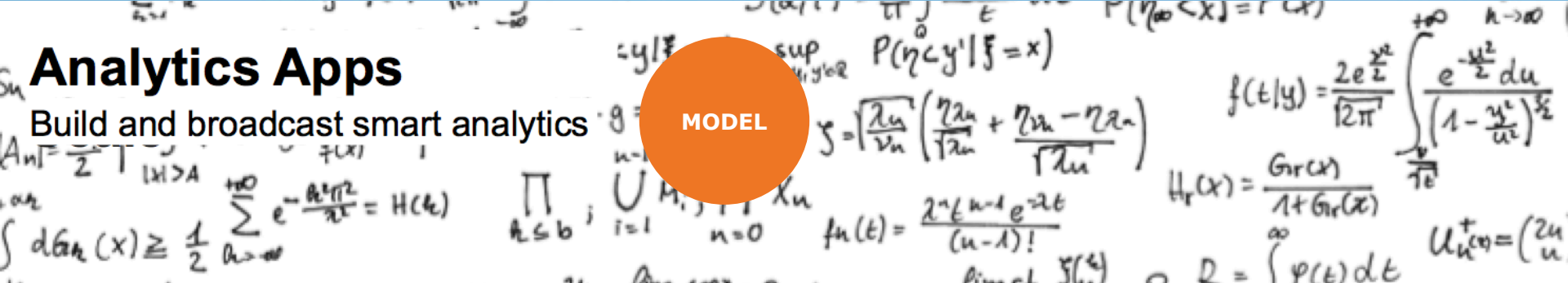
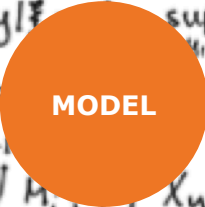
Be first to insight, first to action



Visual analytics is like a **bicycle** for your business mind.

Analytics Apps

Build and broadcast smart analytics



Streaming Analytics

Continuous algorithmic awareness and automation



#1. Smart Visual Analytics

Smart Visual Analytics

Be first to insight, first to action

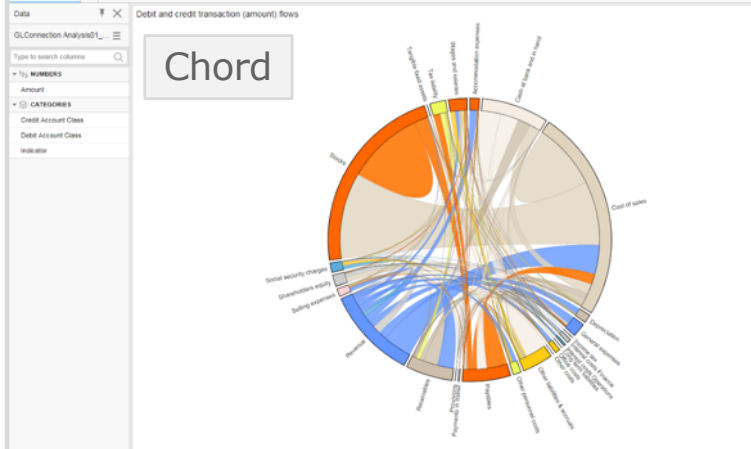
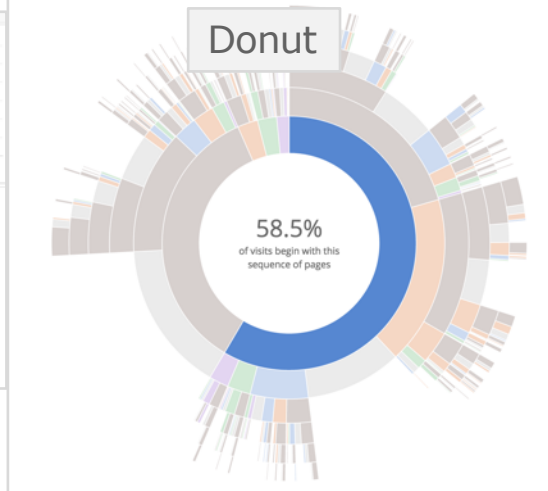
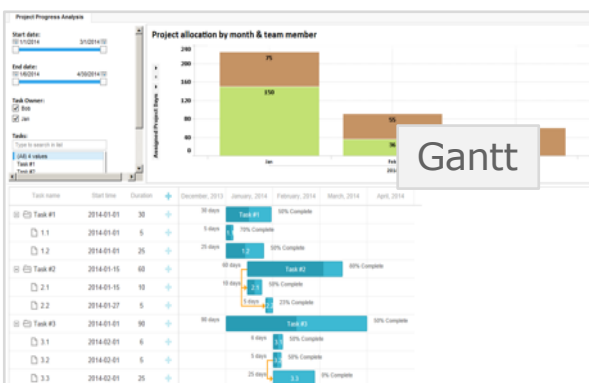
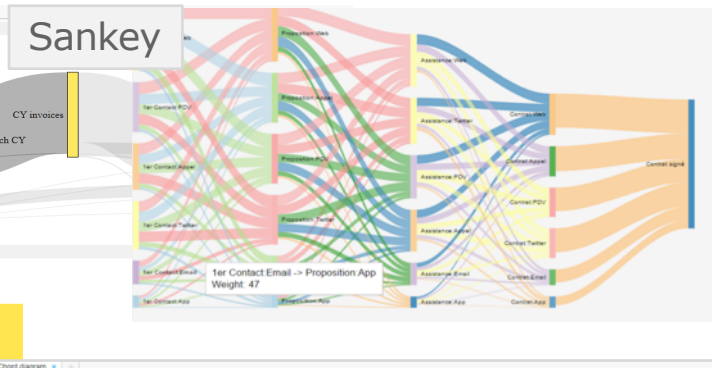
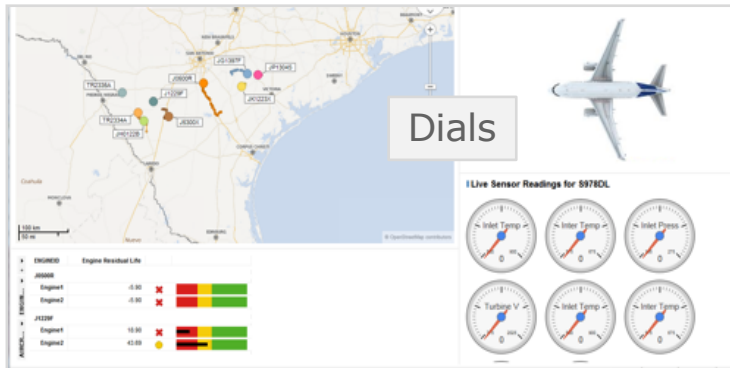


Visual analytics is like a **bicycle for your business mind.**

Visual Analytics – Interactive Spotfire visualization



Visual Analytics – Extending the Palette



Visual Analytics – Graph Configurations in Spotfire

Chip Contour Data Function

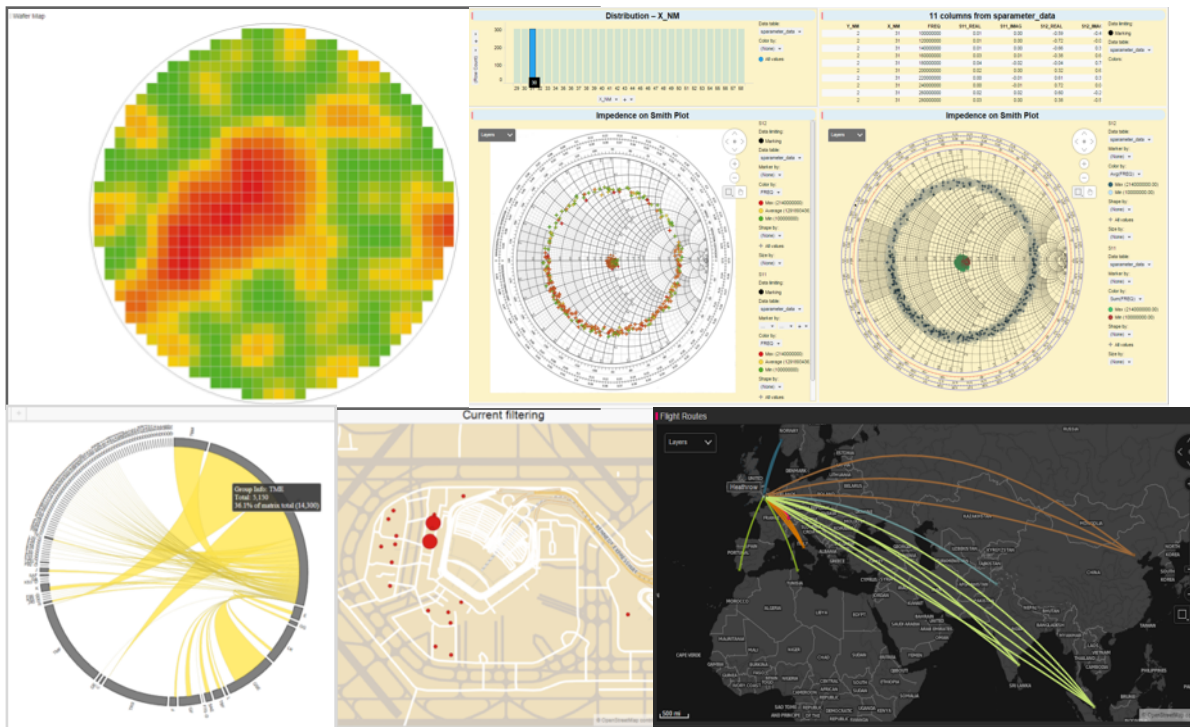
- Contour coloring
- Contour layers

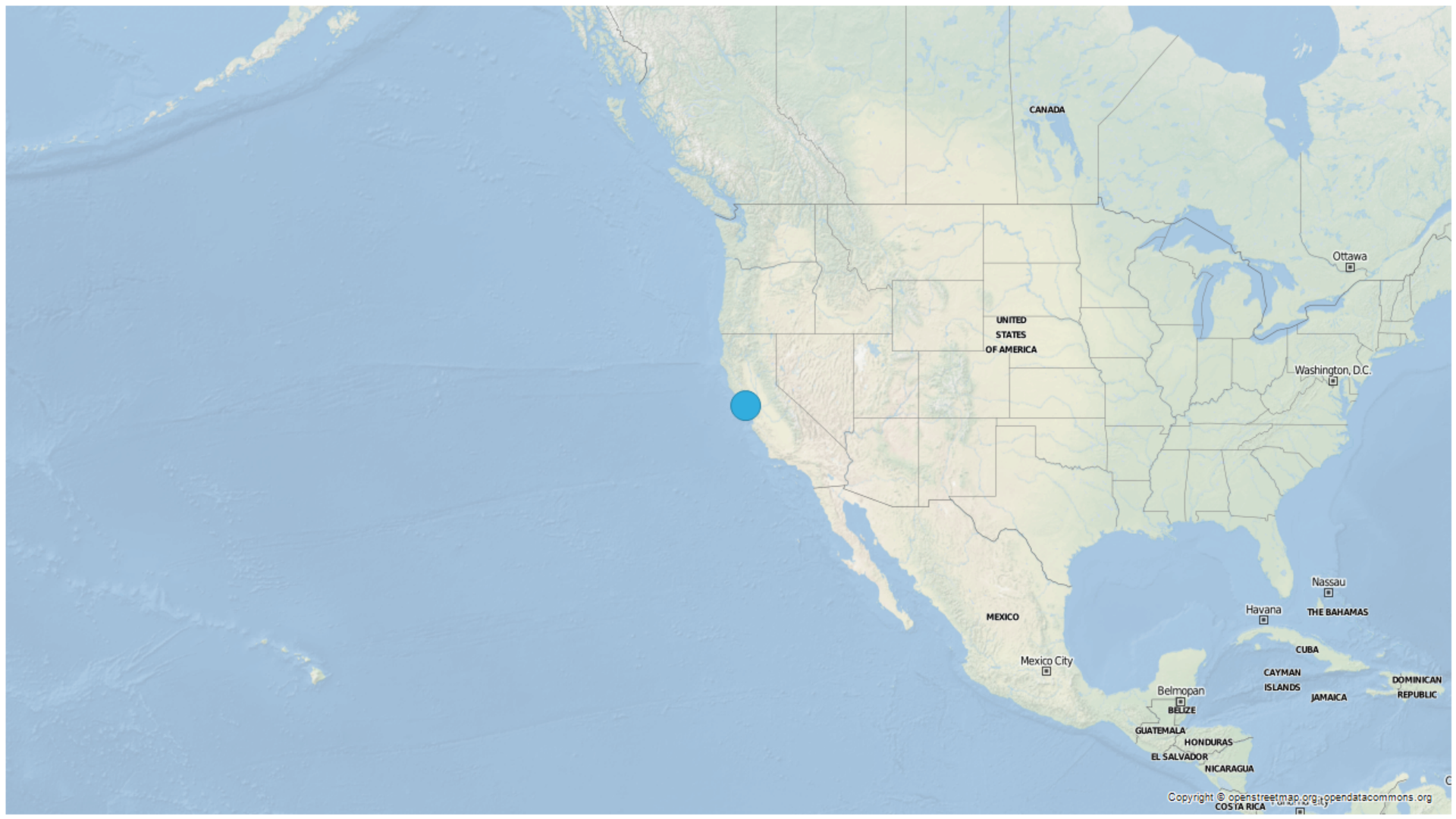
Auto Wafer Data Function

- Auto-generate based on chip location data
- Wafer border
- Wafer reticle shot

Background Image

- Register





CANADA

UNITED STATES OF AMERICA

MEXICO

Ottawa

Washington, D.C.

Nassau

THE BAHAMAS

Havana

CUBA

CAYMAN ISLANDS

JAMAICA

DOMINICAN REPUBLIC

Belmopan

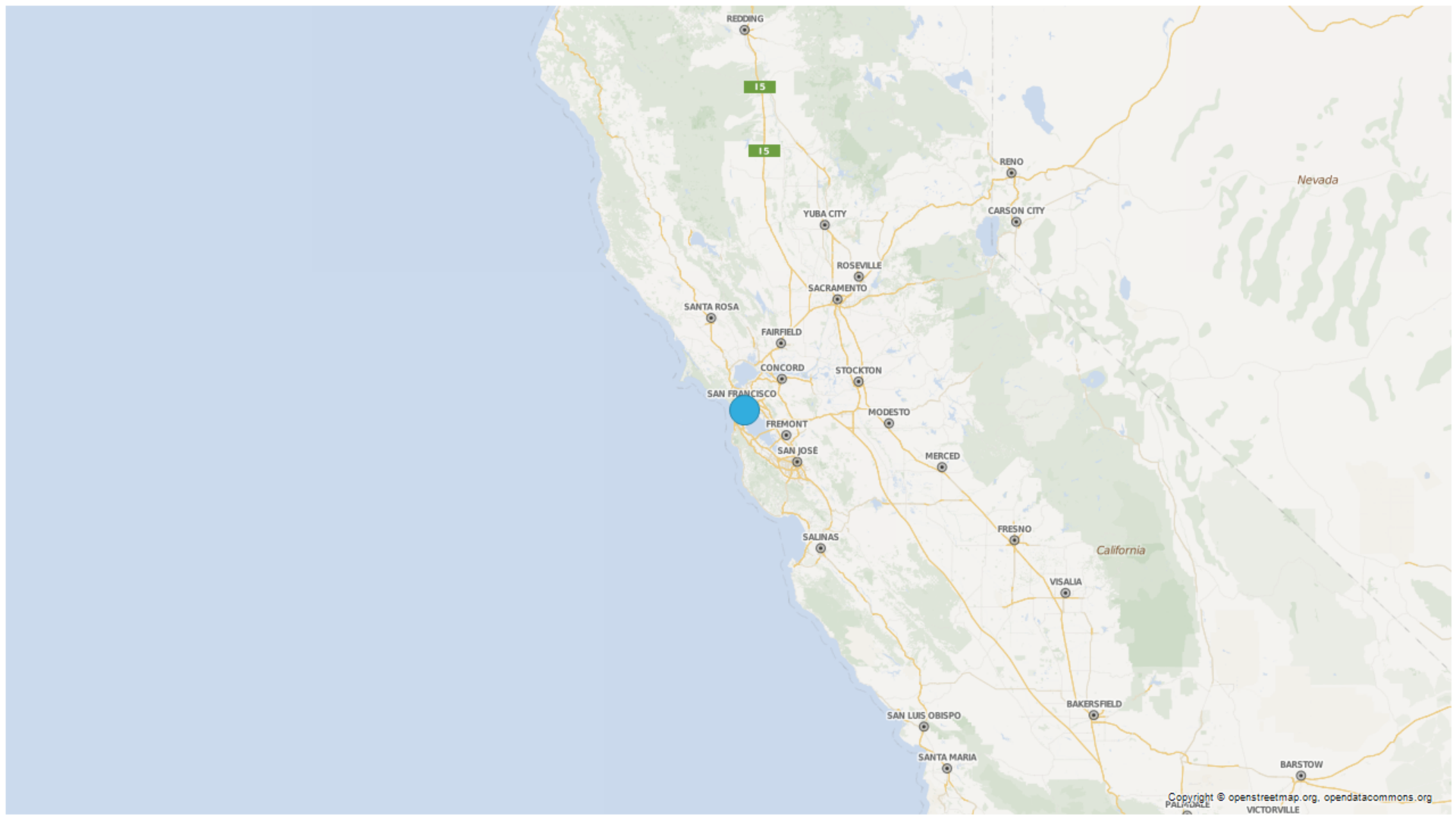
BELIZE

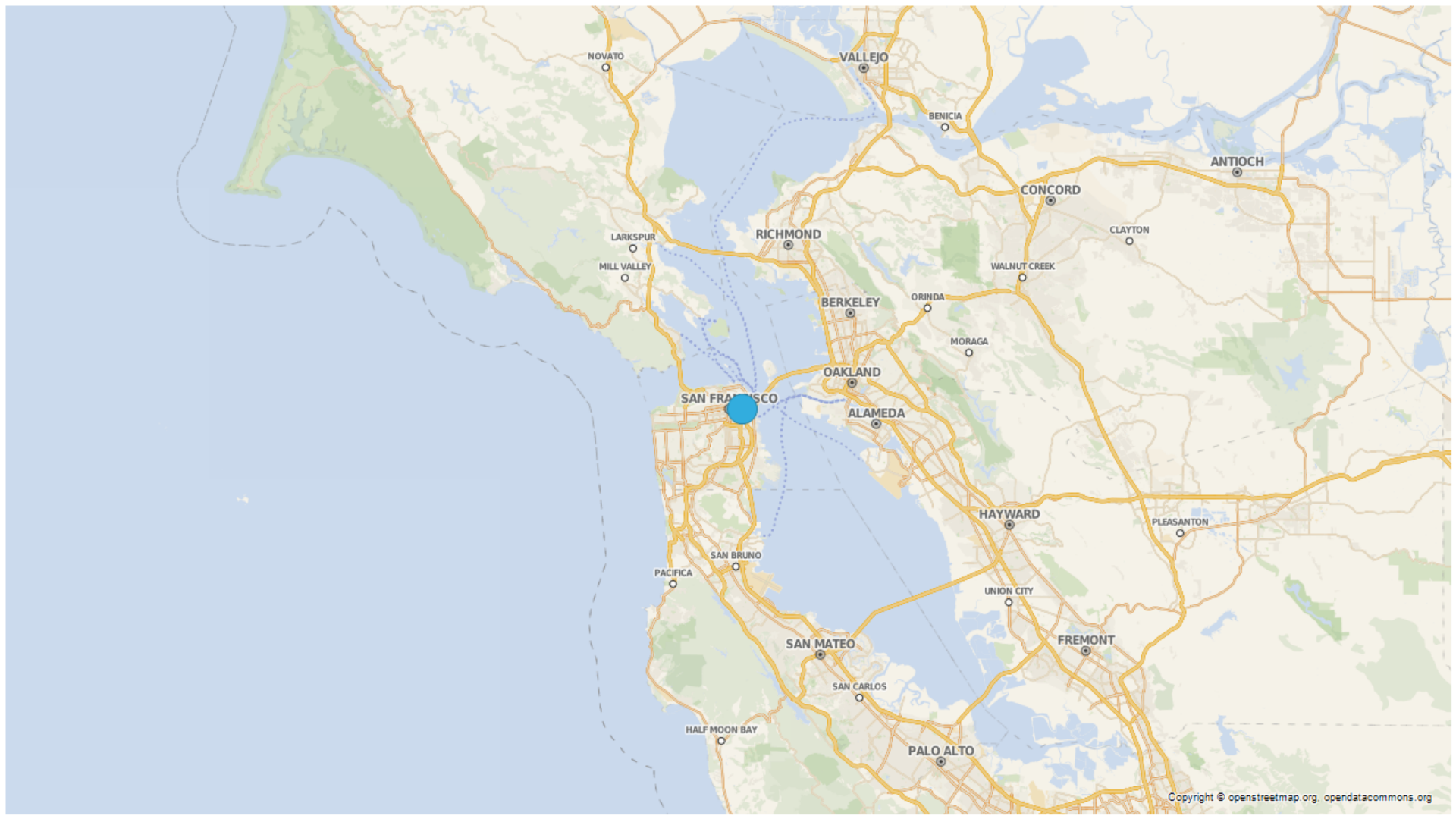
GUATEMALA

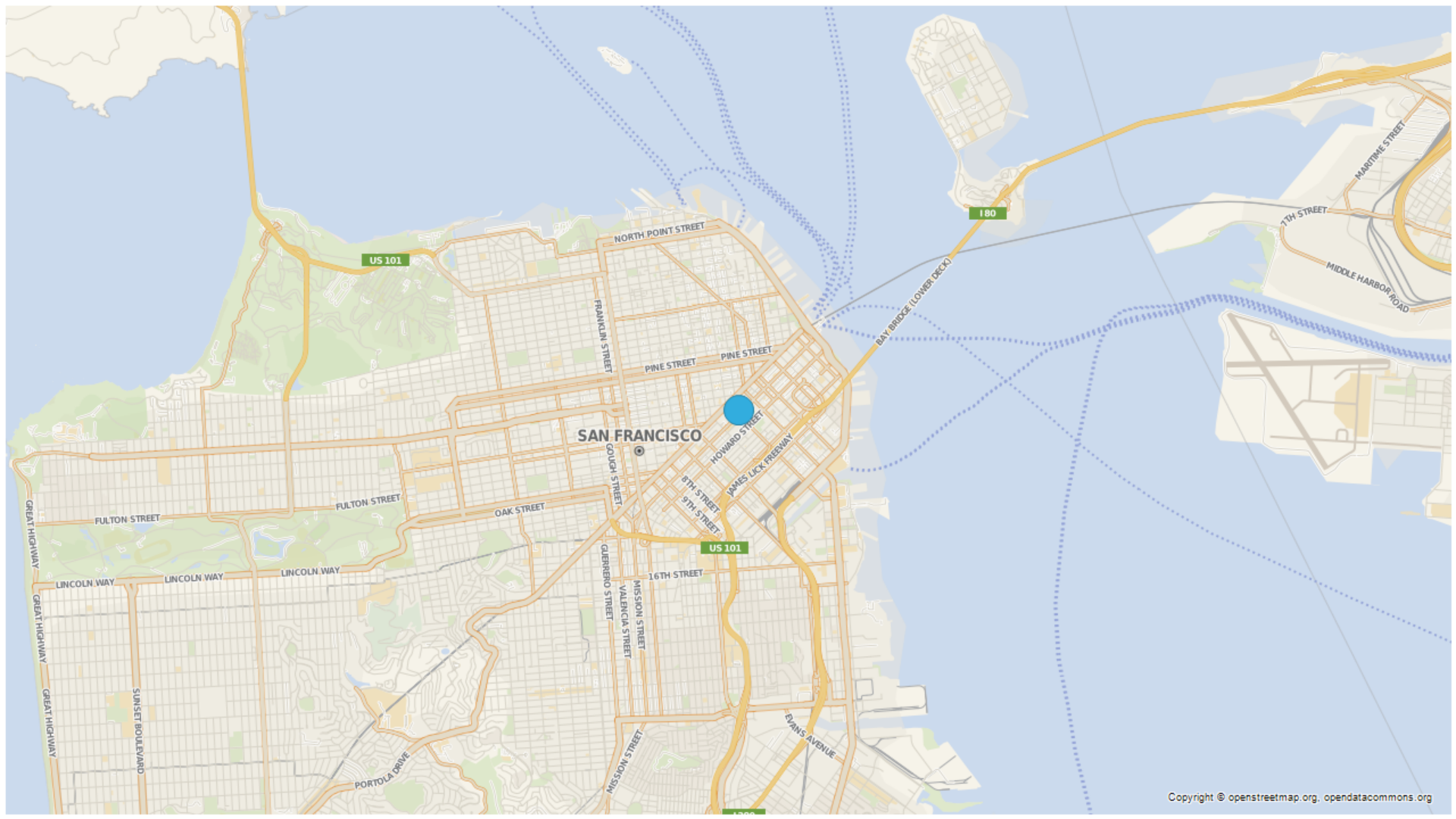
HONDURAS

EL SALVADOR

NICARAGUA





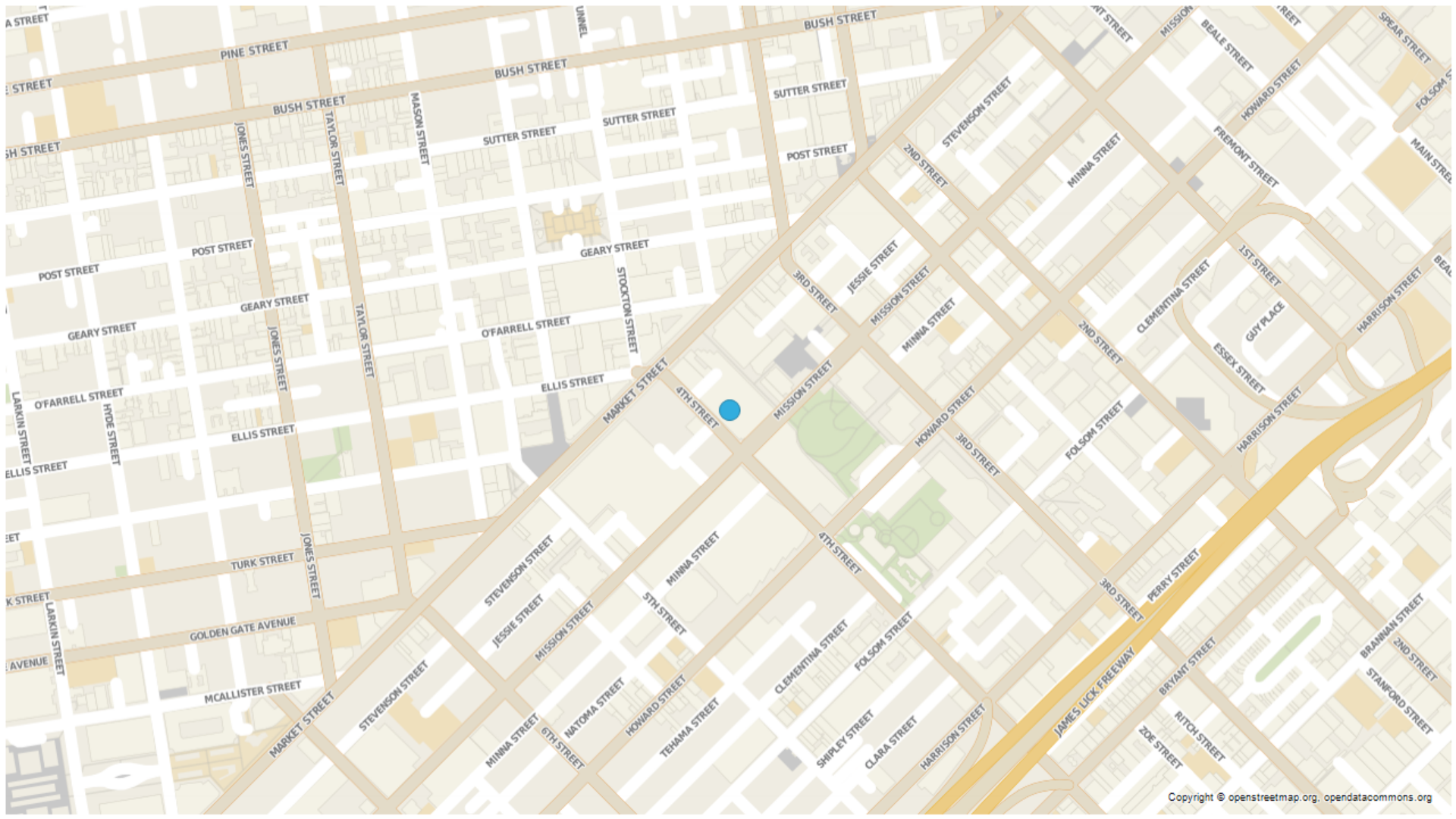


SAN FRANCISCO

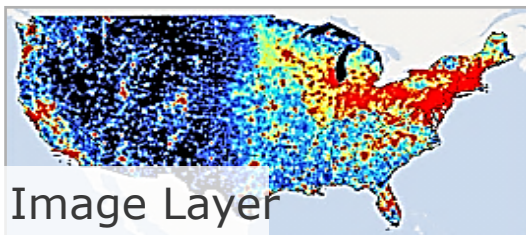
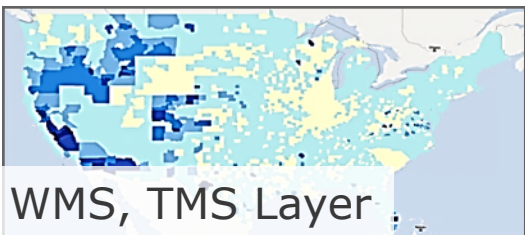
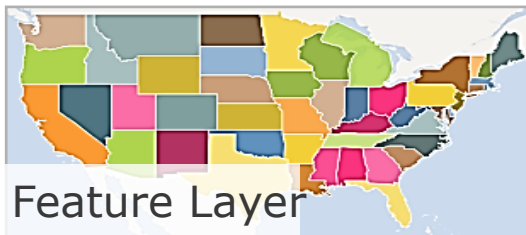
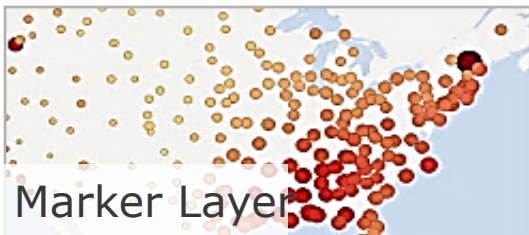
US 101

I 80

US 101



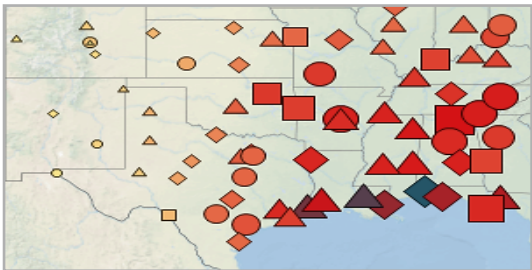
Map Layers



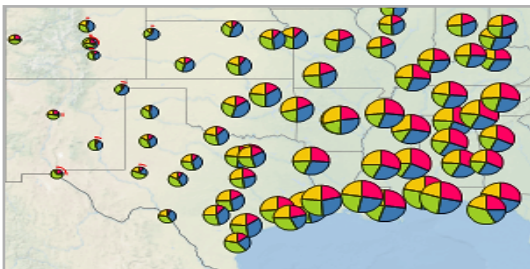
Map Elements

Marker Layer

- Color
- Shape
- Size

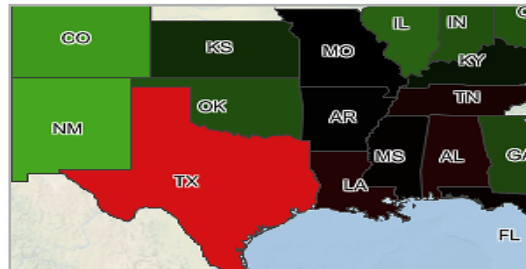


- Relative amounts
- Size



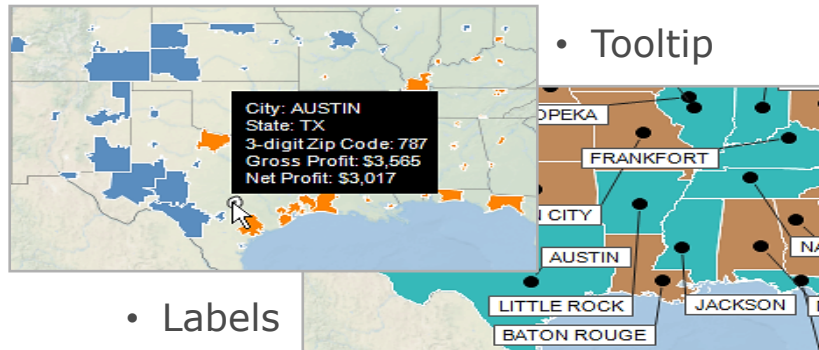
Feature Layer

- Color



Marker or Feature Layer

- Tooltip



- Labels

Spotfire : Recommendations

Recommended visualizations

Oil_gas_Recent_P

Type to search columns

NUMBERS

- gas
- water

TIME

- prodDate

LOCATION

- Latitude
- Longitude

CATEGORIES

- API - 14 Digit/UWI
- Operator
- Operator (grouped)

Filtered out at 10:52:37 AM

Operator

Operator (grouped)

Clear selections

Do not show recommendations automatically. Add default visualization instead.

TIBCO



Examples of Spotfire Recommendations in Action
Easy dashboard setup for business users,
dramatically faster creation of full-featured data
analysis applications for analysts

TIME IS OF THE ESSENCE

"With a dashboard, every unnecessary piece of information results in time wasted trying to filter out what's important, which is intolerable when time is of the essence."

The agile business intelligence market is growing rapidly, and as Gartner points out, the transition is toward platforms that can be rapidly implemented and used by analysts and business users to find insights quickly—as well as by IT staff to quickly build analytics content to meet business requirements and deliver more timely business benefits. This drive for speed is about business value: accuracy and speed of interpretation for decision-making, authoring, and development of data discovery applications, and task completion to enable developers to implement their ideas quickly and obtain accurate insights.

This paper describes a recommendation engine for the TIBCO Spotfire® interactive graphical analysis system. Spotfire Recommendations makes data discovery fast and easy for both analysts and business users. The system uses metadata typing and built-in graphics taxonomy to produce a collection of inherently sensible graphics choices applied to the data at hand. The user chooses from the suggestions, and the software builds a dashboard of linked, brushable, configurable graphics with supporting data filters and graphics controls that can be rapidly applied to the canvas.

1 Rita L. Sallam, Bill Hostmann, Kurt Schlegel, Joao Tapadinhas, Josh Parenteau, Thomas W. Oestreich, Magic Quadrant for Business Intelligence and Analytics Platforms, Gartner, February 23, 2015.

2 Stephen Few, Information Dashboard Design, Analytics Press, CA, 2015.



Do we have enough shelters for the homeless?

Homeless Count
610,042

Bed Count
730,376

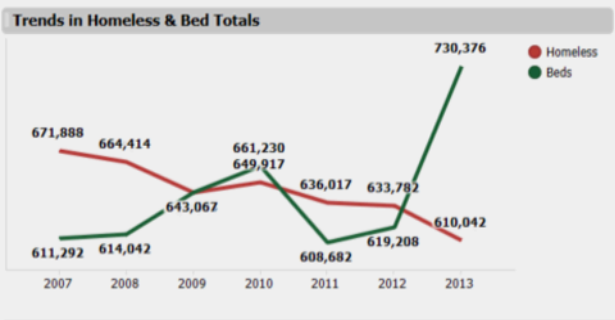
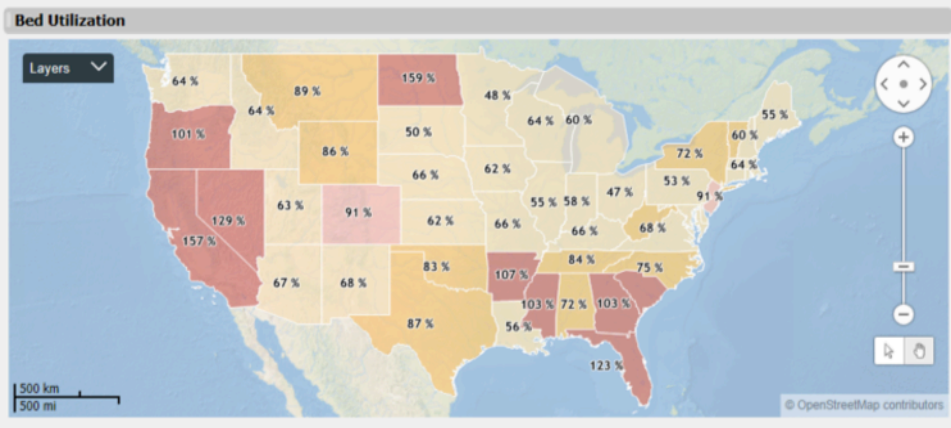
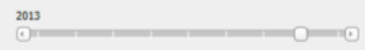
Bed Utilization
84 %



CoC Count
426

Bed Deficit CoC Count
108

Bed Deficit CoC %
25 %



Bed Utilization by CoC

State	CoC Number	Homeless	Beds	Bed Utilization %
AK	AK-501	824	1,068	77 %
AK	AK-500	1,122	1,619	69 %
AL	AL-507	905	710	127 %
AL	AL-505	399	398	100 %
AL	AL-501	493	603	82 %
AL	AL-503	586	737	80 %
AL	AL-502	223	326	68 %
AL	AL-504	515	806	64 %
AL	AL-500	1,469	2,583	57 %
AL	AL-506	99	314	32 %
AR	AR-503	1,678	771	218 %
AR	AR-512	82	60	137 %
AR	AR-501	668	563	119 %
AR	AR-504	254	362	70 %

Search by County in Cross Tab:

Dashboards in Spotfire



Do we have enough shelters for the homeless?

Homeless Count
19,029

Bed Count
29,800

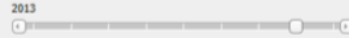
Bed Utilization
64 %



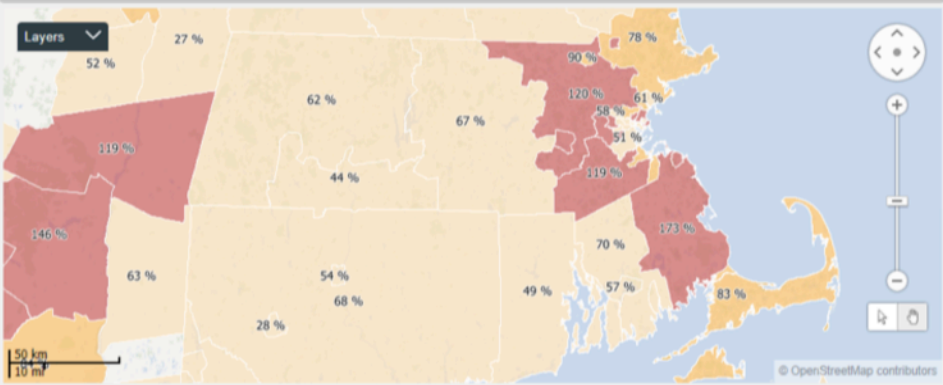
CoC Count
25

Bed Deficit CoC Count
7

Bed Deficit CoC %
28 %



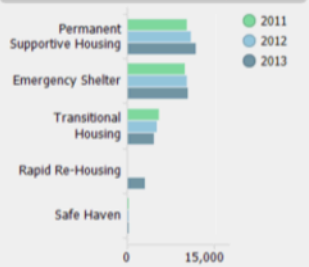
Bed Utilization



Trends in Homeless & Bed Totals



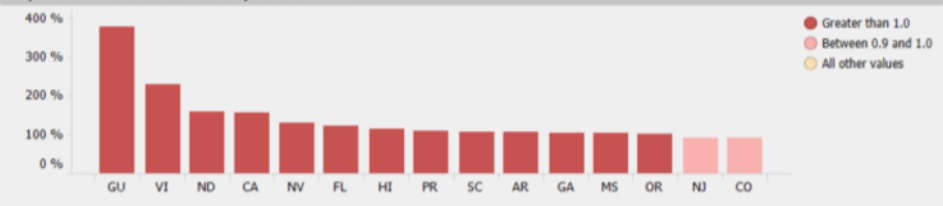
Trends in Beds per Shelter Type



Bed Utilization by CoC

State	CoC Number	Homeless	Beds	Bed Utilization -
MA	MA-518	707	379	187 %
	MA-520	1,296	749	173 %
	MA-510	1,286	978	131 %
	MA-516	1,593	1,349	118 %
	MA-508	559	624	90 %
	MA-503	514	621	83 %
	MA-513	352	482	73 %
	MA-511	394	545	72 %
	MA-519	251	359	70 %
	MA-506	1,657	2,486	67 %
	MA-507	678	1,086	62 %
	MA-502	468	768	61 %
	MA-517	169	292	58 %
	MA-515	235	410	57 %
MA, CoC		4716	6673	64 %

Top 15 States and Territories by Bed Utilization



Search by County in Cross Tab:



Intro Dashboard

Do we have enough shelters for the homeless?

Homeless Count
19,029

Bed Count
29,800

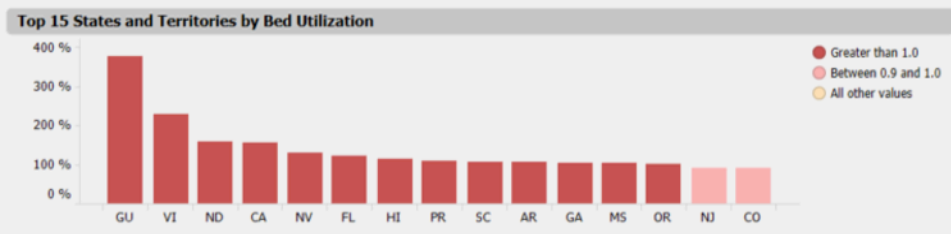
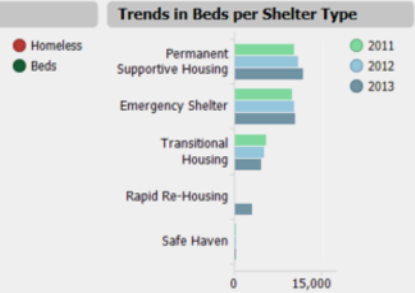
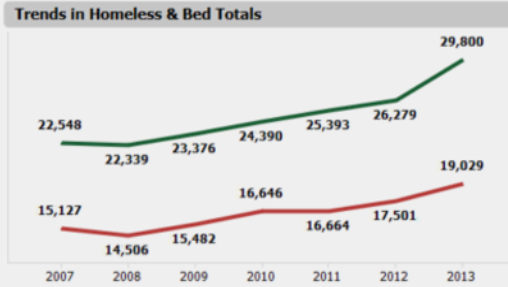
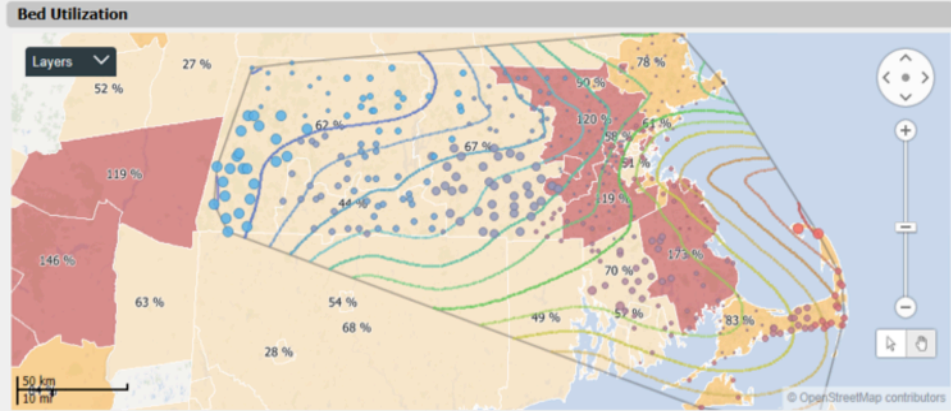
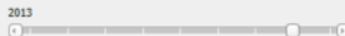
Bed Utilization
64 %



CoC Count
25

Bed Deficit CoC Count
7

Bed Deficit CoC %
28 %



Bed Utilization by CoC

State	CoC Number	Homeless	Beds	Bed Utilization %
MA	MA-518	707	379	187 %
	MA-520	1,296	749	173 %
	MA-510	1,286	978	131 %
	MA-516	1,593	1,349	118 %
	MA-508	559	624	90 %
	MA-503	514	621	83 %
	MA-513	352	482	73 %
	MA-511	394	545	72 %
	MA-519	251	359	70 %
	MA-506	1,657	2,486	67 %
	MA-507	678	1,086	62 %
	MA-502	468	768	61 %
	MA-517	169	292	58 %
	MA-515	235	410	57 %
	MA-509	536	653	54 %

Search by County in Cross Tab:

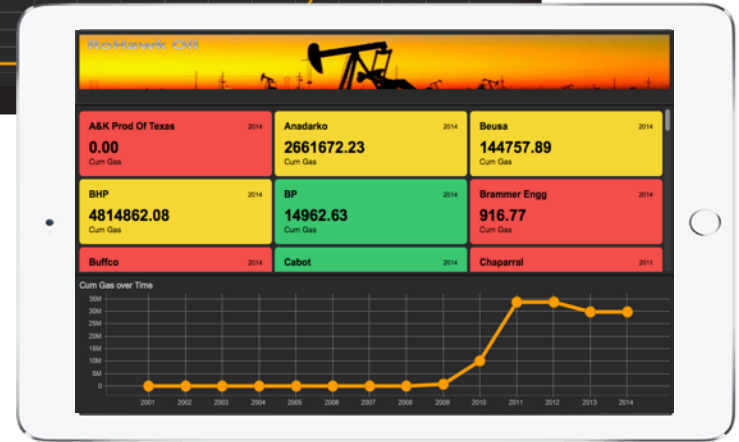
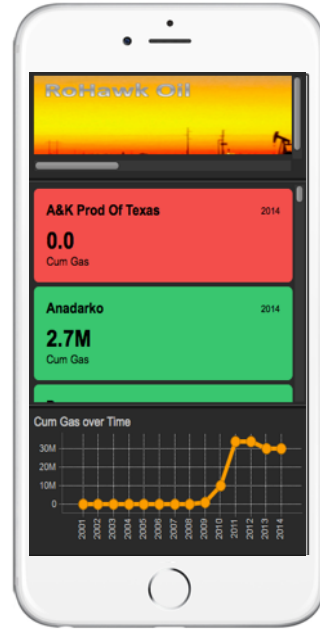
Mobile – Responsive Design

Responsive Design

- Responsive to real estate
- Laptop, Tablet, Phone

Deployment Kit

- White label apps



#2. Numerical Models – What's Needed

Analytics Apps

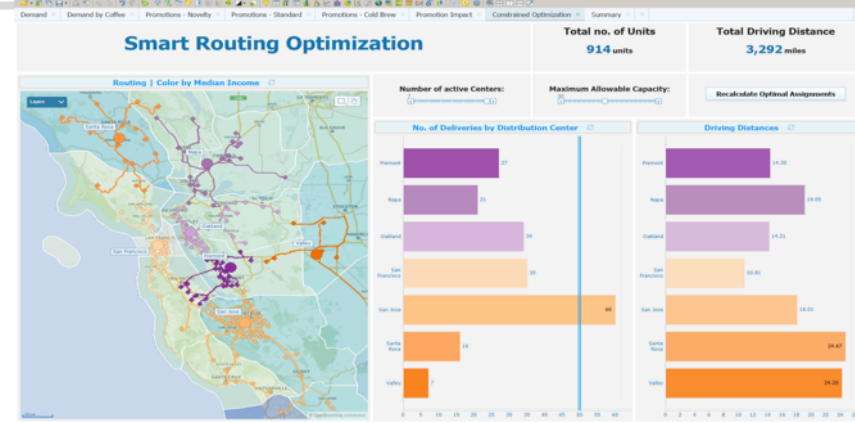
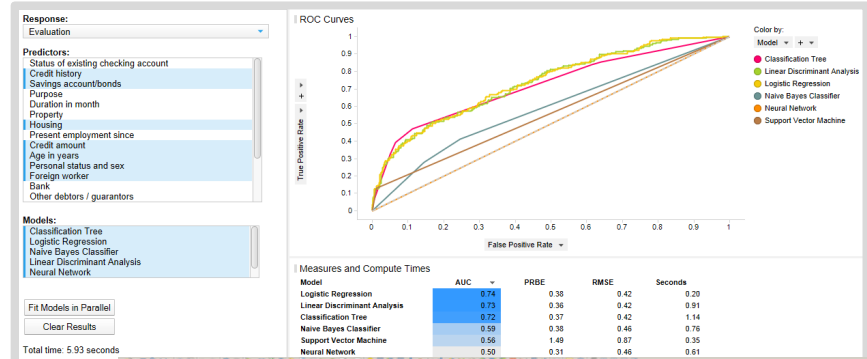
Build and broadcast smart analytics

Handwritten mathematical notes including:

- $y|f=x = \sup_{y' < y, y' \in R} P(\eta < y' | f=x)$
- $f(t|y) = \frac{2e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} \int_0^{\frac{y^2}{2t}} \frac{e^{-\frac{u^2}{2t}} du}{(1-\frac{u^2}{2t})^{\frac{3}{2}}}$
- $H_r(x) = \frac{G_r(x)}{1+G_r(x)}$
- $f_n(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}$
- $\lim_{n \rightarrow +\infty} \frac{f_n(t)}{n} = P_e$
- $R = \int_{-\infty}^{\infty} \varphi(t) dt$
- $U_n^+ = (2n)$
- $\log \varphi(t) = i\gamma t - c|t|^\alpha [1 + i\beta \frac{t}{|t|} \omega(t, \alpha)]$
- $B(u) = \sum_{k=1}^r \Psi^*(b_k u)$
- $Civ = \sum_{j=1}^n a_{ij} b_{jv}$
- $\lim P \left(\frac{\sum_{k=1}^n a_{kj} - n \log \frac{1}{q}}{\dots} \right)$
- $C_n(\alpha) \geq \frac{n!}{\Gamma(\dots)}$
- $\frac{\sin th}{th} [\varphi(t)]$
- $\frac{u}{\dots} \varphi(t) = \varphi$

Algorithms: Rules, Machine Learning & Optimization

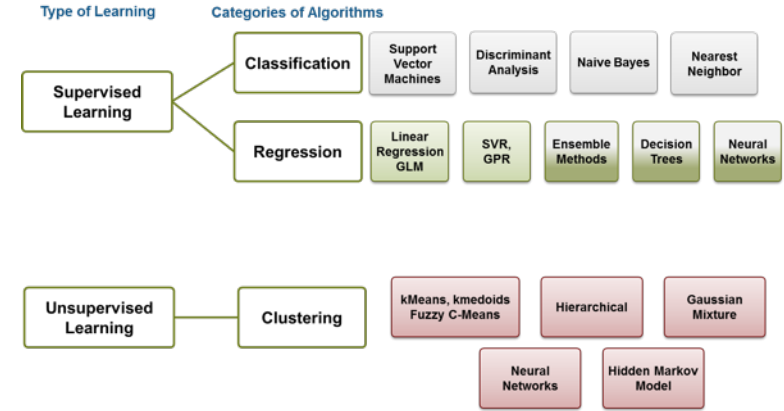
- Declarative & Heuristic Rules
- SPC and Anomaly Detection
- Machine Learning
 - Supervised
 - Unsupervised
 - Gradient Boosting Machines
 - Random Forests
 - Deep Learning
- Optimization
 - Linear & Quadratic Programming
 - Genetic Algorithms
 - Process optimization
 - Capacity constraints



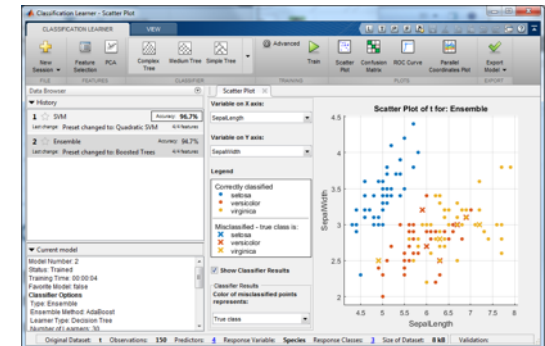
Algorithms: Machine Learning with MATLAB

Machine Learning finds predictive models in data without being told where to look

- **Supervised** – Solve known problems: $y=f(X)$
 - Build a model that predicts a condition (failure, success, ..)
 - *What factors are driving failures?*
- **Unsupervised** – Identify patterns, Detect anomalies *X only*
 - *Are there new patterns or failure modes emerging?*
- Easy to get started with MATLAB
 - Interactive, app-driven workflows
 - Work with business and engineering data (signal, images, financial, geospatial)
 - Deploy to IT systems or run on embedded systems
 - High quality libraries



Classification Learner App

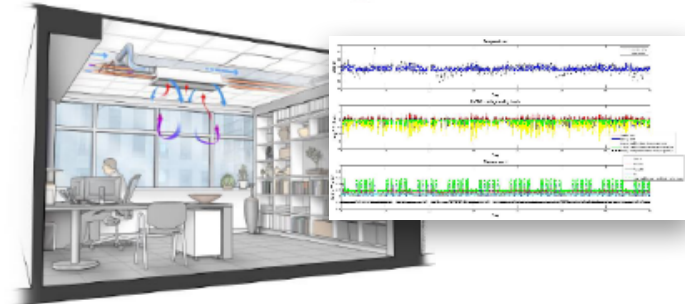


Algorithms: Optimization with MATLAB

- **Prescriptive Analytics** – Support Decision-making
 - Find best solution when there are constraints on the process
 - *What is the optimum allocation of resources for equipment maintenance? ...for energy production?*
- **Decision-making**
 - Linear, Quadratic, Mixed-integer, Nonlinear
- **Design**
 - Nonlinear
 - Global: multistart, genetic algorithm, particle swarm, pattern search, simulated annealing
- **Financial Applications**
 - Portfolio Optimization, Risk Analytics, Econometrics
- **Performance options**
 - Multi-threaded, symbolic
 - On-demand Amazon EC2 with MATLAB Parallel Cloud
 - Compute cluster with MATLAB Distributed Computing Server



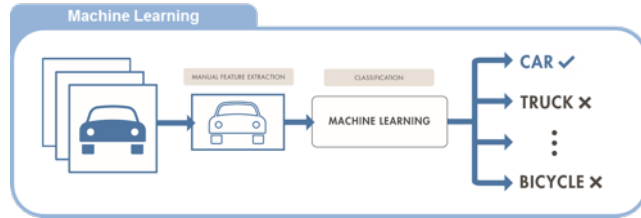
Predict and Optimize Energy production



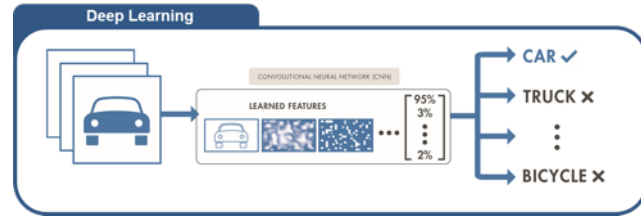
Online Optimization of Building Energy Use



Algorithms: Deep Learning with MATLAB



Machine Learning learns tasks using features extracted manually from data



Deep Learning learns both features and tasks directly from data

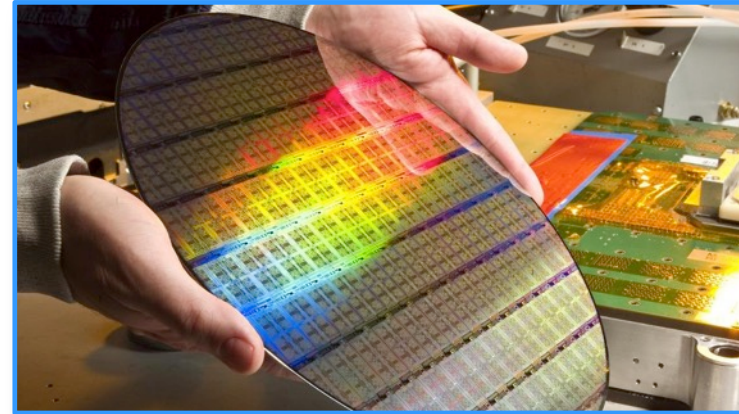
Deep learning – for image classification and computer vision

- Access to pre-trained models and datasets (eg ImageNet)
- Apps for data augmentation and labeling
- GPU for training acceleration
- High quality libraries: Autoencoders, CNNs

Modeling Yield and Quality

Goal: Predict Quality (e.g. Yield) as function of equipment and process attributes

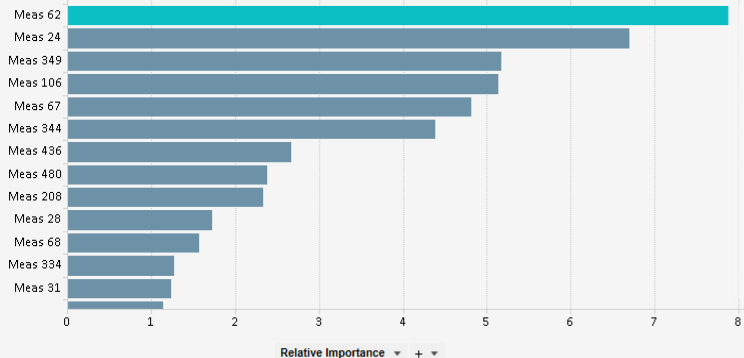
- **Response: Yield (continuous)**
- **Predictors: equipment and process attributes**
 - Machines, assemblers, operators, date ranges,
 - Sensor data: pressure, temperature, ...
 - Maintenance logs, control charts
 - Supplier data: electrical, chemical, physical characteristics
 - Defect inspection data
- **Big Data: many columns**
 - Wafer production: 1000 sensors * 1000 readings / sensor
 - Assembly: 1000-5000+ components in some assemblies
- **Models: Gradient Boosting Machine works well**
 - Root Cause / Fingerprints



Model: Gradient Boosting Machine

GBM Results

Predictor Importance - Effect on Yield



Predictor Interactions Summary Table

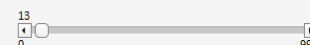
var1.name	var2.name	index
Meas 62	Meas 24	0.20
Meas 349	Meas 67	0.05
Meas 24	Meas 106	0.03
Meas 24	Meas 349	0.03
Meas 24	Meas 67	0.03
Meas 106	Meas 67	0.02
Meas 62	Meas 67	0.02
Meas 349	Meas 106	0.02
Meas 62	Meas 349	0.01
Meas 62	Meas 106	0.00

Heat Map Setup

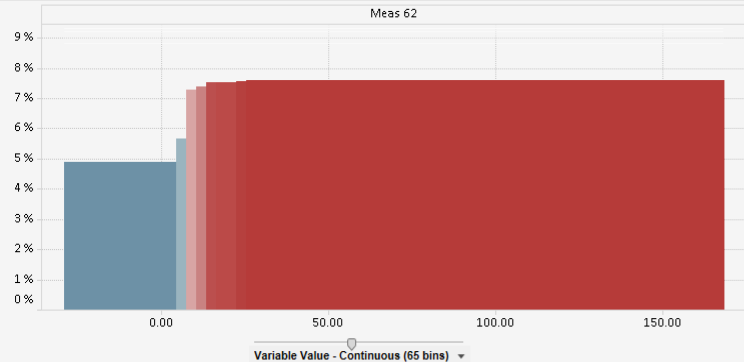
Bin Continuous variables into N groups for display:



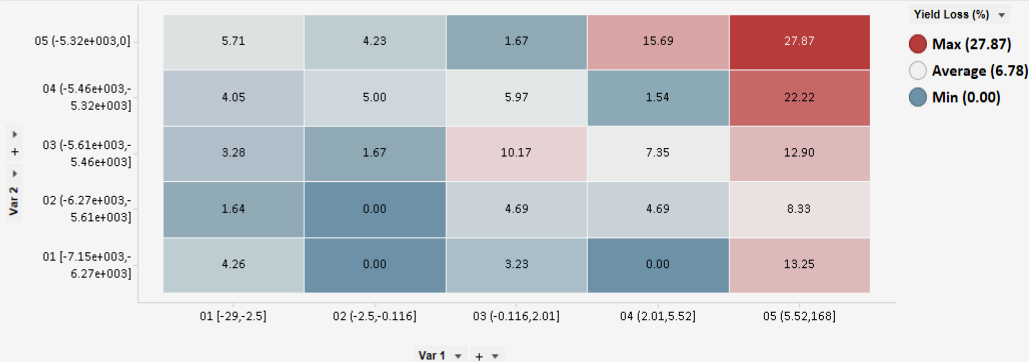
Eliminate small cells with unstable target average. Only show cells with n rows greater or equal to:



Predictor Effect on Yield Detail



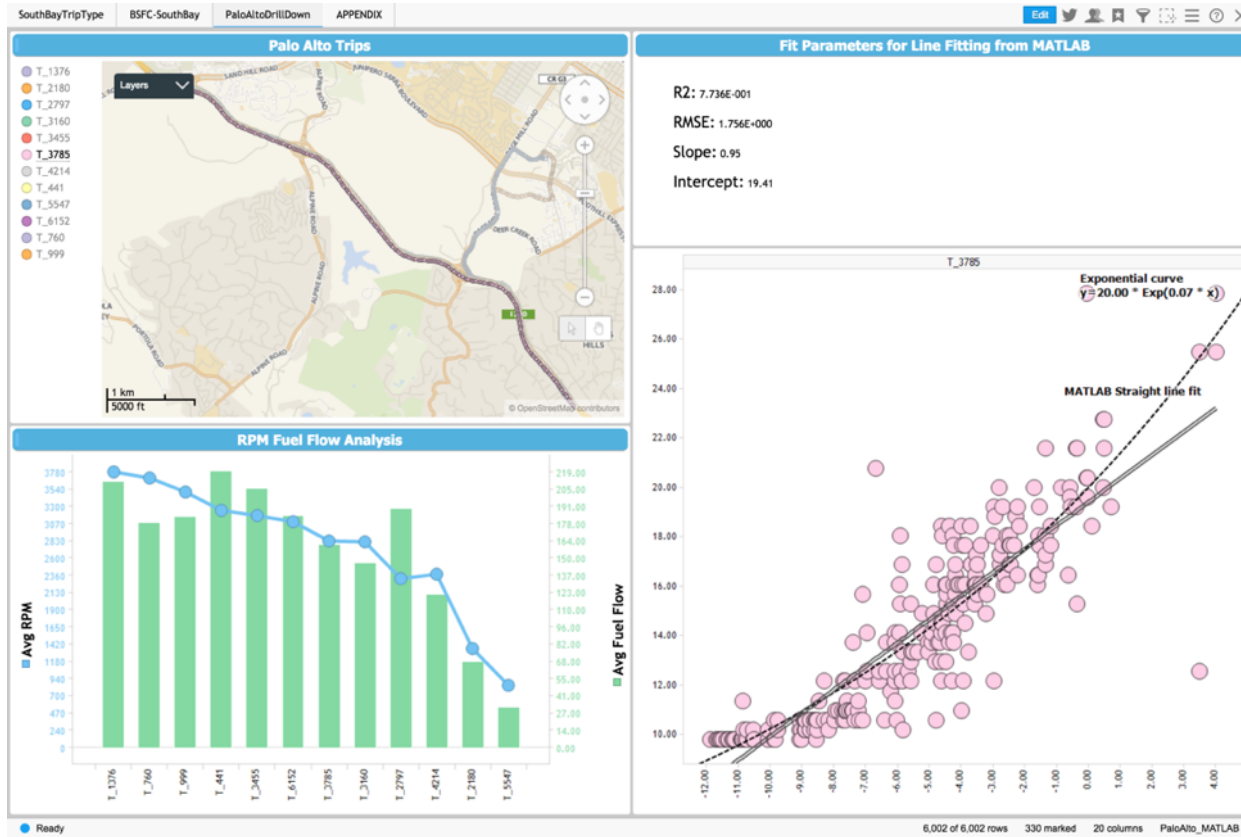
Predictor Interactions Detail



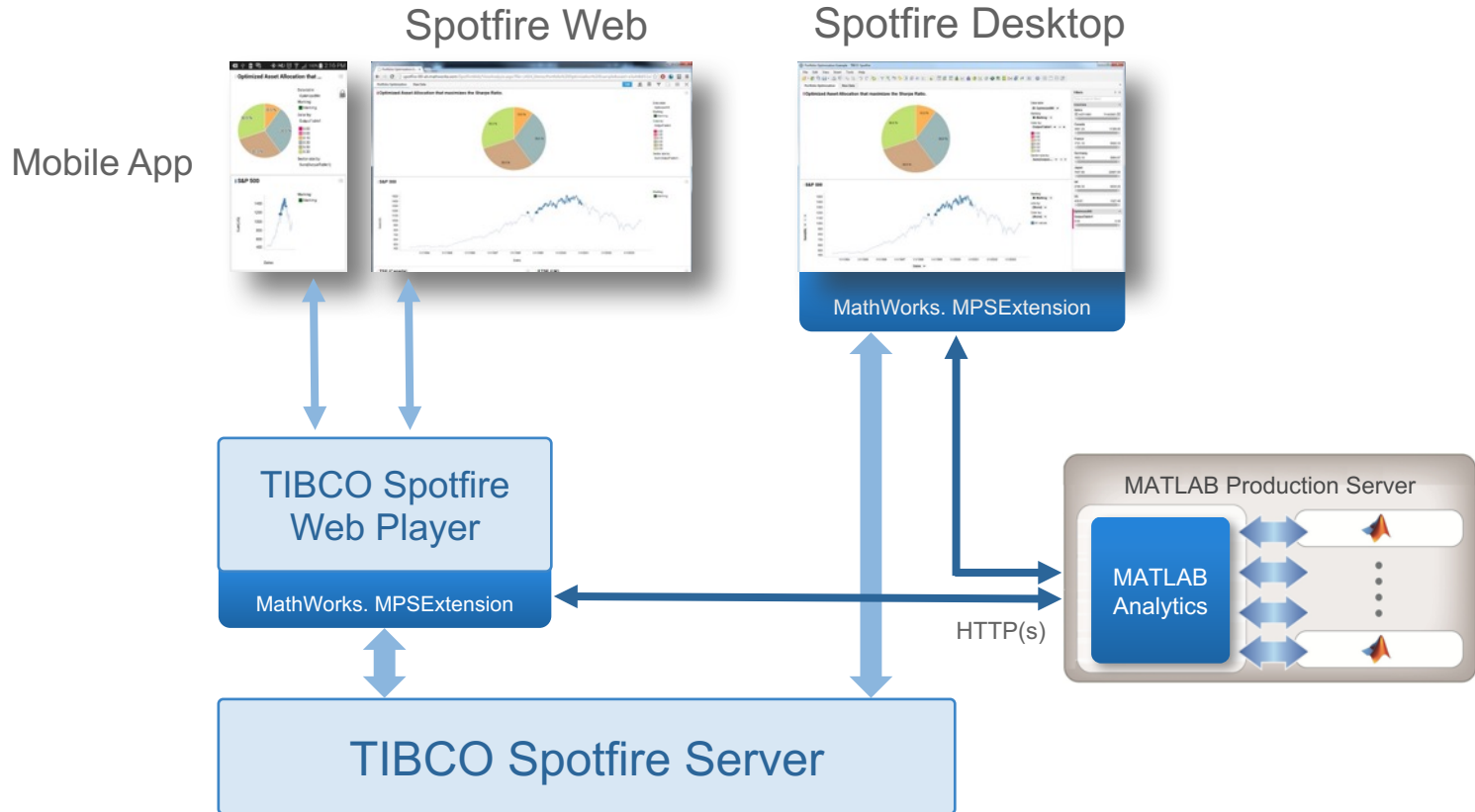
Demo



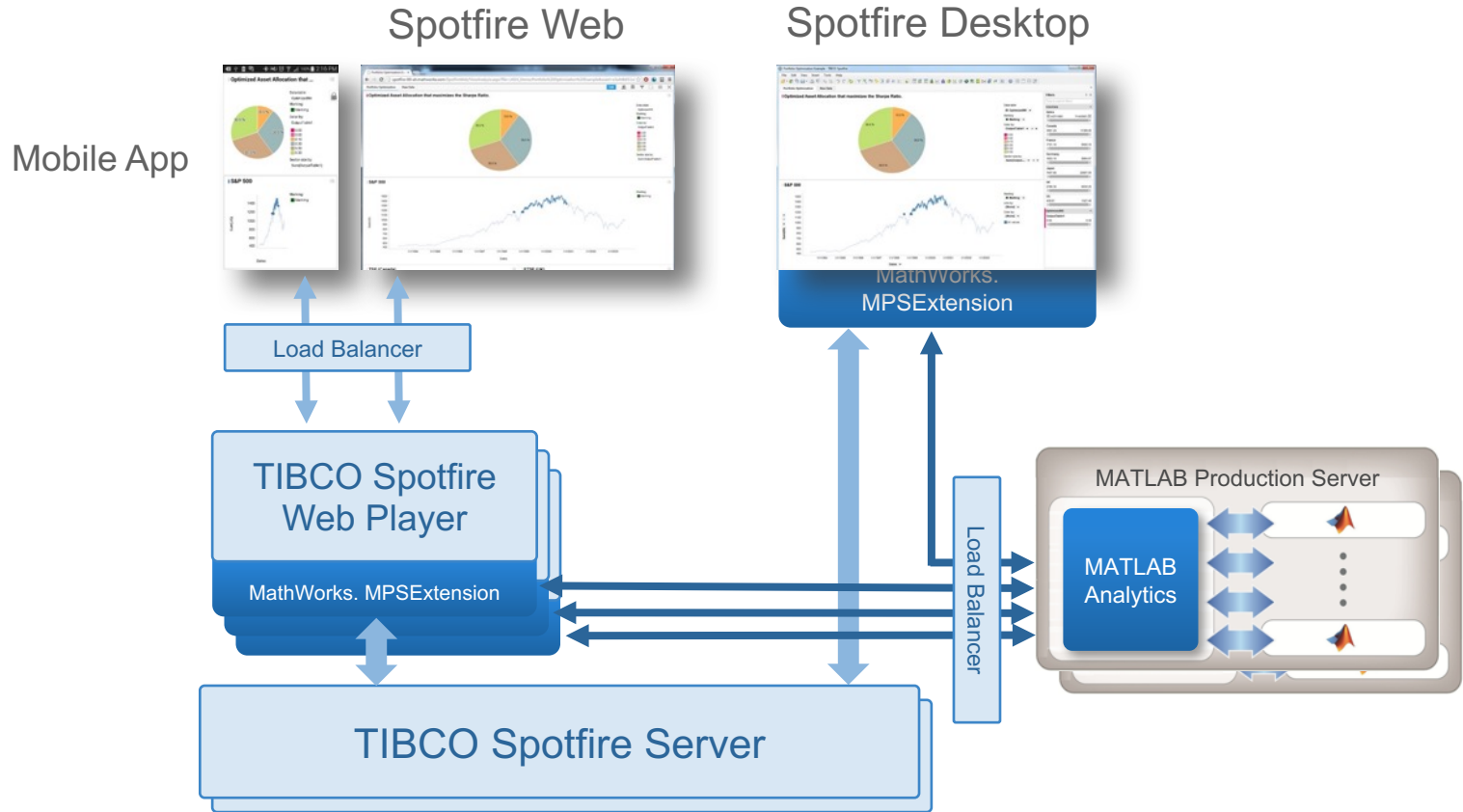
Demo



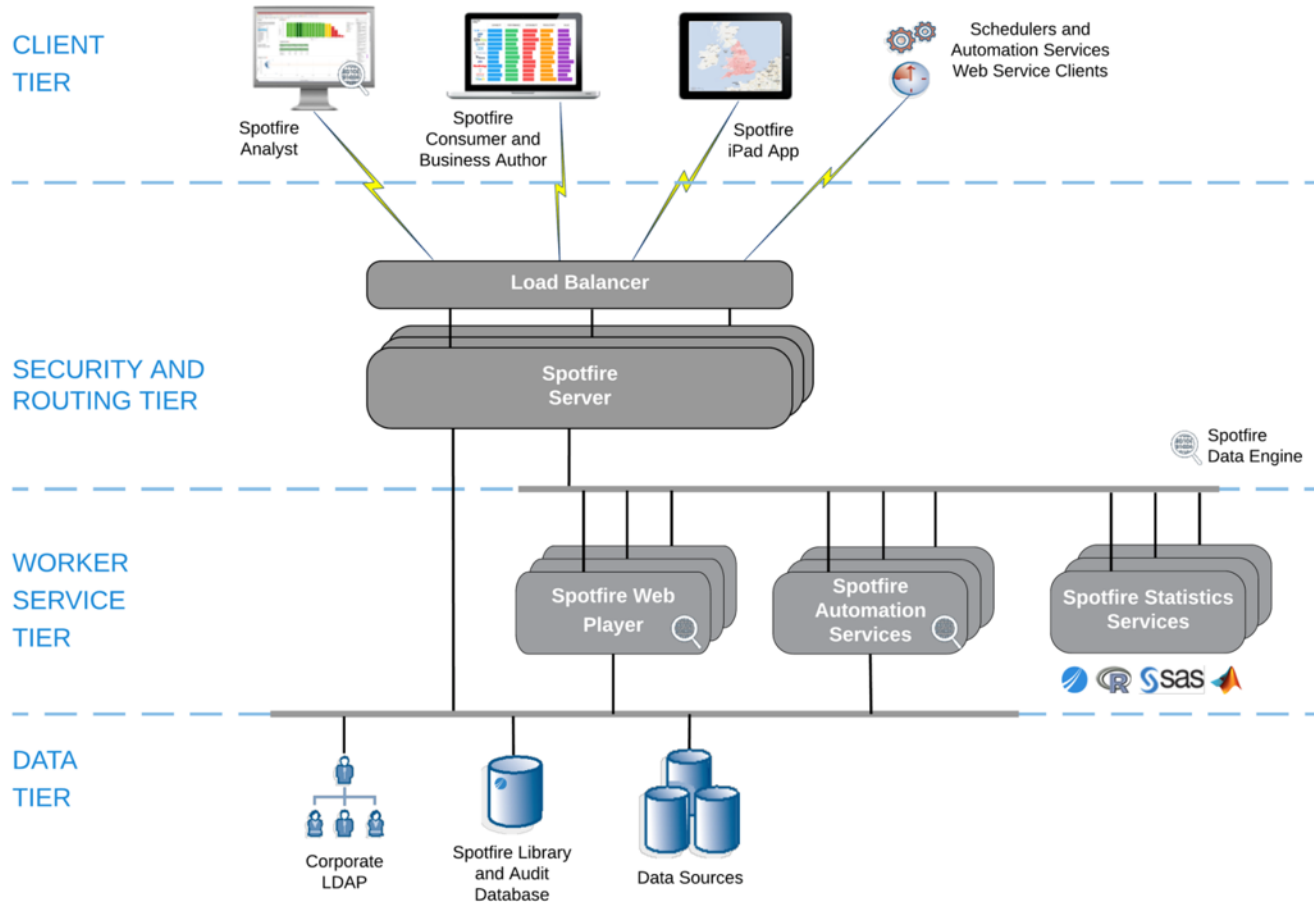
Reference Diagram



Increasing Capacity and Redundancy



Overall Spotfire Architecture



#3. Streaming Analytics

Streaming Analytics

Continuous algorithmic awareness
and automation

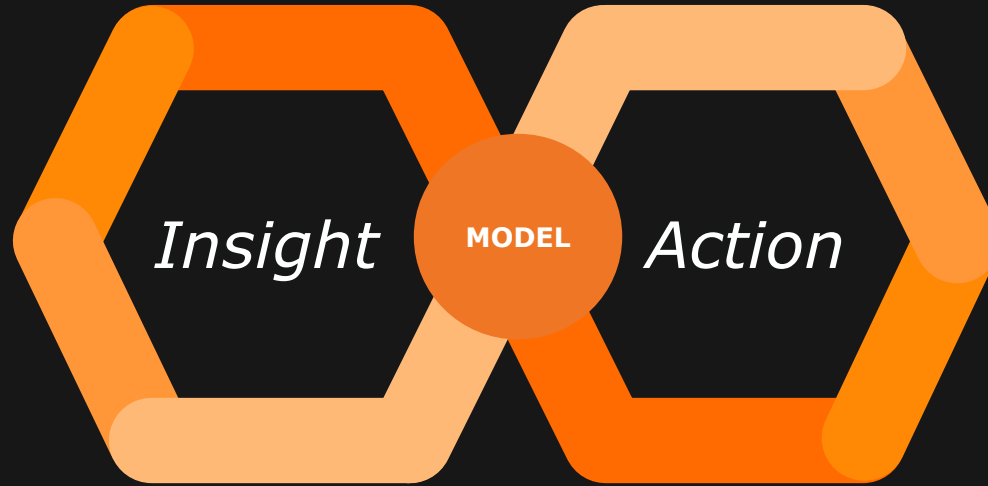


The TIBCO Insight Platform

TIBCO Spotfire

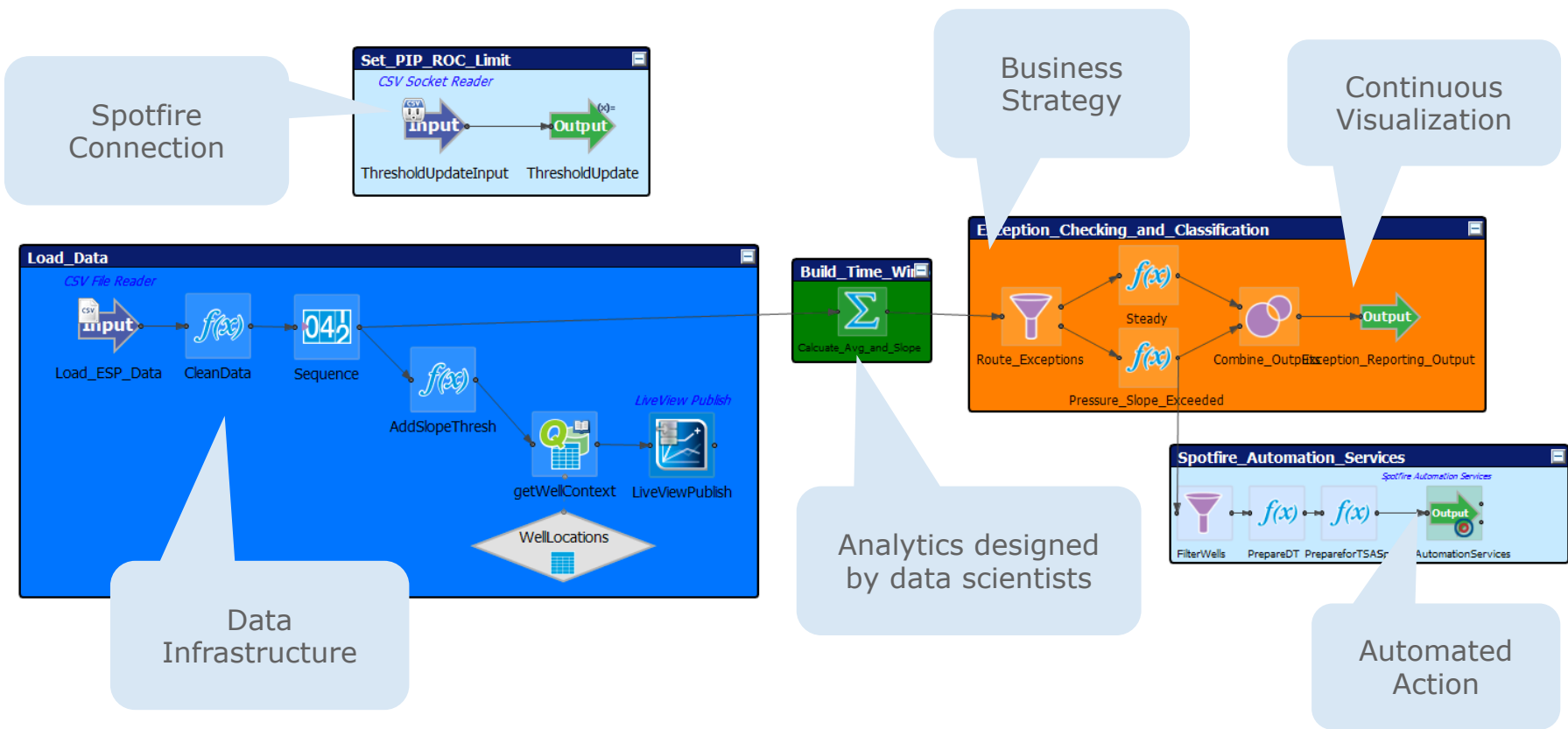
MATLAB

TIBCO Streambase

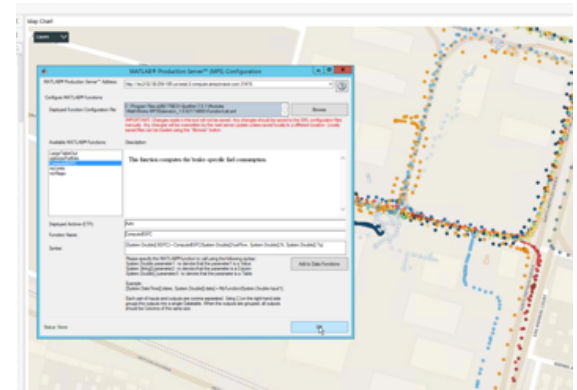
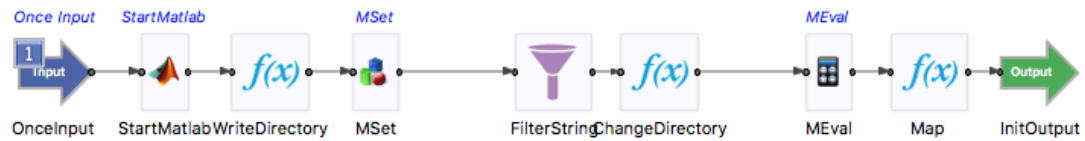
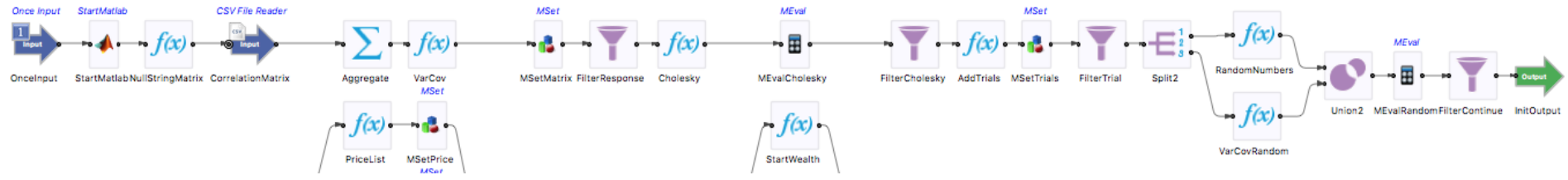


TIBCO EMS, BW, ...

Streaming Analytics with Streambase



Streaming Analytics – with MATLAB injection

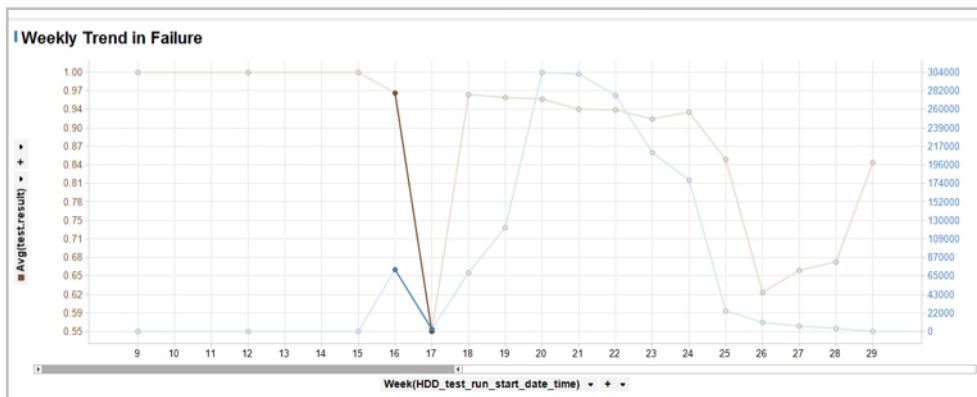


MATLAB Production Server

Example: Hard Disk Manufacturing

Example: Hard Drive Manufacturing

- Problem in week 17
 - Yield drops from 96% to 55%
 - Production reduced from 70K to 3K drives



Here, we use GBM, a Machine Learning Algorithms to identify which important variables within the control and test groups.

Step 1: Select the control and test groups

Step 2: Change the universe of variables that you want to explore if necessary.

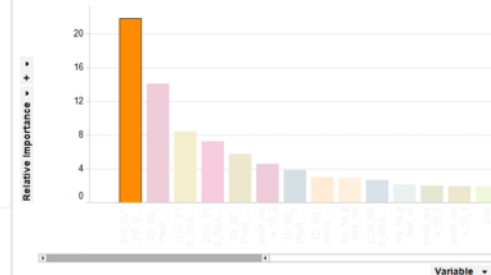
Step 3: Select the "Calculate" button.

Calculate

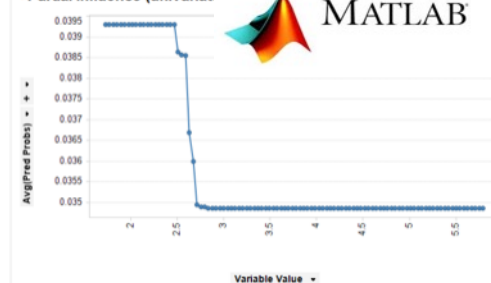
Variable Metadata

Variable	Role	Data Type
...	Response	NA
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test
...	Predictor	Continuous - Test

Variable Importance

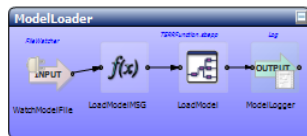
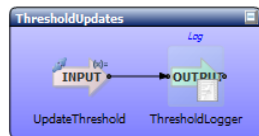


Partial Influence (univariate)



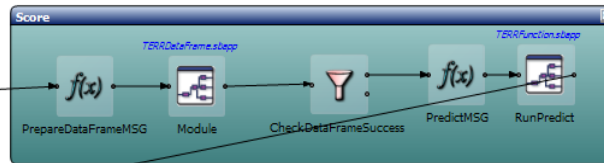
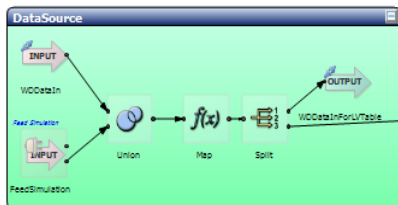
- Machine Learning Model
- Parameter linked to head is primary culprit
- Publish Model to Event Server to monitor

Example: Hard Disk Manufacturing



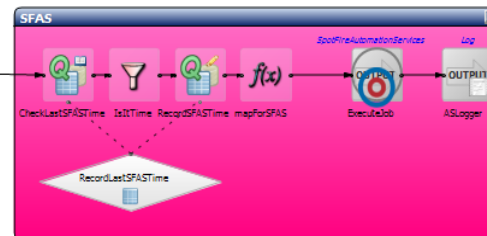
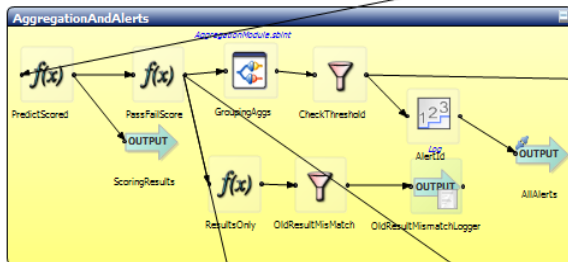
GBM Model published from MATLAB to Streambase

Data Refresh

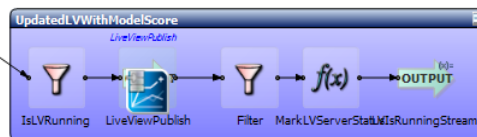
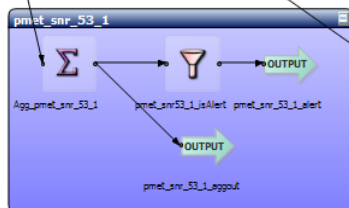


GBM Model Scores Data

Thresholds

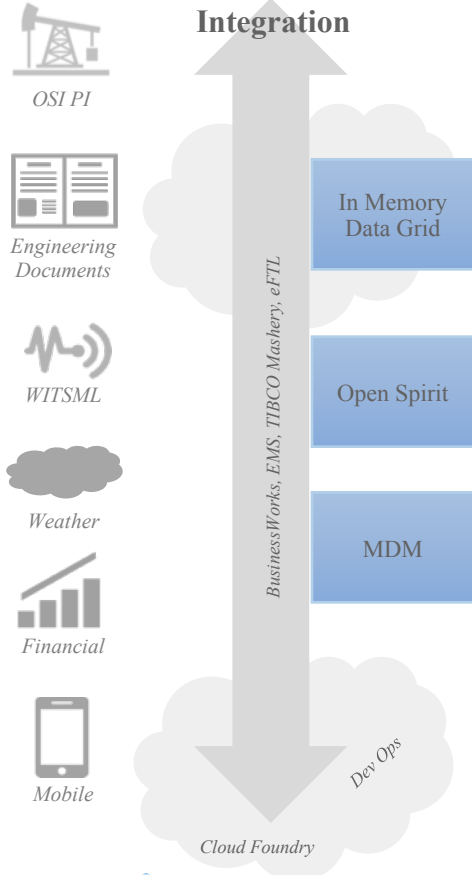


Notifications Interventions



\$10+ MM ROI generated

TIBCO Insight Platform



OSI PI



Engineering Documents



WITSML



Weather



Financial



Mobile

Spotfire Analytics

TIBCO Spotfire, R/TERR, MATLAB

Streaming Analytics

StreamBase, BusinessEvents
R/TERR, MATLAB

Digital Operations

TIBCO Live Datamart

Case Management

Contextual Information

I will deal with this incident.
27/11/2015, 12:51:15

I have requested coaches to transfer passengers from the next station
27/11/2015, 12:53:19

TIBCO BPM

ALERTS



Community Wiki & Exchange

<https://community.tibco.com/exchange>

Users Wiki Exchange

Extend Spotfire client

Embed Spotfire in web applications

Automation Services

IronPython scripts

JavaScript API

Custom tasks (C# API)

C# Extensions

TIBCO Community Products Answers Wiki

Data Access and Data Wrangling with TIBCO Spotfire®

Last updated: 8:07pm May 05, 2016

627 views

2 comments

Flag as Inappropriate

#Big data #Data wrangling + see all/edit tags

Edit This Page Create New Page

TIBCO Community Products Answers Wiki

Extending TIBCO Spotfire®

Last updated: 3:37pm Jun 06, 2016

Download Solutions or Build your own Tips & Tricks

http://www.tibco.com/blog/2016/07/29/dr-spotfires-live-online-office-hours/

Edit This Page Create New Page

TIBCO Community Products Answers Wiki Exchange

Dashboard Design Best Practices Examples

LARGE TITLES on each page: Calculated values in text area to display KPI's

Link to dpt: Warehouse Shipment Analysis

Link to the public analysis on Cloud: CFO Analysis

DARK THEMES for simple dashboard: **COORDINATED COLOURS:**

Link to the public analysis on Cloud: US Child Hunger

Link to the public analysis on Cloud: California Drought

CUSTOM DESIGN for customer:

Link to dpt: Time Series Forecasting Automotive Theme

Link to dpt: World Analysis

TIBCO Community Products Answers Wiki Exchange Documentation Try Now

Color Schemes

Here you can download pre-prepared color schemes and apply them to your visualizations in Spotfire. Click on the colour schemes to download them as zip files.

CATEGORICAL: **CONTINUOUS:**

Filter by

- Product
 - TIBCO Spotfire® (2)
 - TIBCO StreamBase® (2)
 - TIBCO Live Business (2)
- Category
 - Analytics
 - Event Processing (2)
- Tags
 - Energy Industry (3)
 - Machine Learning (3)
 - Manufacturing Industry (3)
 - Analysis Templates (4)
 - data functions (4)

For instructions visit the following Spotfire help pages:

- Color Schemes Overview: The entire setup of colors for a visualization is referred to as its color scheme. Which colors and threshold values you choose, as well as the color mode you select, are part of a visualization's color scheme. The current color scheme of a visualization can be viewed, changed, and on the Colors page of the Visualization Properties. This is also where you edit a color scheme.
- Copying a Color Scheme: You can open a color scheme that has previously been saved to disk, the library. You can also apply a color scheme from another visualization in the analysis, or open document color scheme. The color scheme must be in the same scale mode as the visualization which you want to use it.
- Applying a Color Scheme to Another: When you have set up a color scheme for a visualization, can apply it to another visualization, provided that the two visualizations are colored by a color the same scale mode.
- Saving a Color Scheme: You can save a color scheme for later reuse or to share it with others. If you save the color scheme to disk or in the library, you can use it in other analyses. If you save it as document color scheme, you can use it again within the analysis. A document color scheme can be selected for a specific visualization in the analysis from the Colors page of the visualization pro. You can also select it as the default color scheme to use for a specific column as well as for new visualizations in the analysis. See Column Properties - Properties, Column Properties Descriptors and Options - Visualization to learn more about using default color schemes.
- Coloring Overview

HTML in Text Areas

Using HTML in Spotfire Text Areas is a great way to enhance the look and feel of your dashboards

Events

Most Recent • Most Popular • Highest Rated

- Gradient Boosting Machine Regression - Data Function for TIBCO Spotfire®
Gradient boosting is an ensemble-decision-tree, machine learning data function that's useful to identify variables that best predict some outcome and build highly accurate predictive models. For example, a retailer might use a gradient-boosting algorithm to determine the propensity of customers to buy a product based on their buying histories.
Last Updated on 10/25am Sep 01, 2016 by NBCSO Software
Exchange #data-functions #Machine Learning #Energy Industry
- Clustering with Variable Importance Data Function for TIBCO Spotfire®
This data function clusters objects together based on similarities between the objects in each cluster. After identifying clusters, this function then ranks the variables according to their influence on cluster formation.
Last Updated on 10/5am Sep 01, 2016 by NBCSO Software
Exchange #Clustering #data-functions #Machine Learning
- Simple Decline Curve Analysis Data Function for TIBCO Spotfire®
This Data Function calculates a Hyperbolic Decline Curve Analysis using production oil and gas data.
Last Updated on 10/5am Sep 01, 2016 by NBCSO Software
Exchange #data-function #Oil and Gas #Analysis Template
- Accelerator for Apache Spark
Analyze your Big Data FAST with the use of this accelerator: Gain insights into your historical data and act in real-time on the current streams of data in conjunction with historical analysis to make crucial decisions when it matters.
Last Updated on 10/5am Sep 01, 2016 by NBCSO Software

TIBCO Spotfire® Community Wiki

Last updated: 8:41pm Jun 10, 2016

5,54k views

15 comments

Create New Page

This wiki provides a location for the TIBCO Spotfire® community to share how-to information with another.

Main Topics

- Getting Started
- Data Access and Wrangling
- Visualizations
- Maps
- Building Applications
- Extending Spotfire
- Administration
- Partners

Other Resources

- Spotfire Product Ideas Portal - submit or vote on ideas and suggestions for the Spotfire product
- List of upcoming Spotfire Webinars and live events with the Spotfire team - meet the experts
- TIBCO and Spotfire specific Meetups and User Groups - share with your fellow Spotfire users
- What's new in TIBCO Spotfire
- TIBCO Now 2016: Breakout Sessions | Keynote talks
- TIBCO Spotfire® Ranked #1 of 28 in Dresner 2016 BI Wisdom of the Crowds Study

Thank you!

Michael O'Connell, PhD

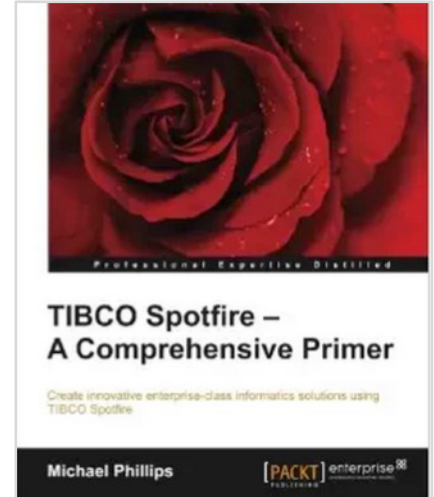
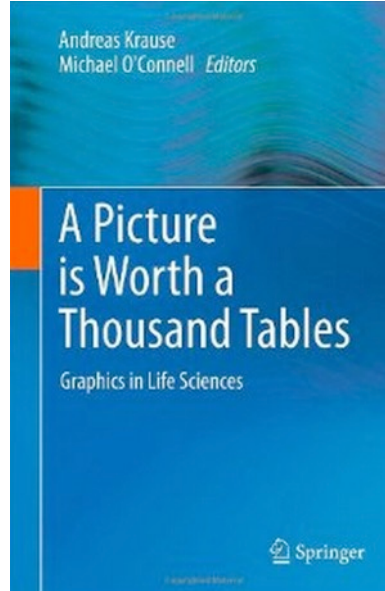
Chief Analytics Officer

TIBCO Fellow

moconnell@tibco.com

@MichOConnell

+1-919-7401560



I

TIBCO®