JMP Topics and Capabilities for
Getting Beyond Spreadsheet Analytics with

Visual Data Discovery

Move from Data to Decision – Faster!

Tom Donnelly, PhD, CAP
Principal Systems Engineer
Defense & Aerospace Team
JMP Statistical Discovery LLC

Desk 919-531-1380
Mobile 302-489-9291
www.jmp.com/fedgov
tom.donnelly@jmp.com

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Exploratory Data Analysis

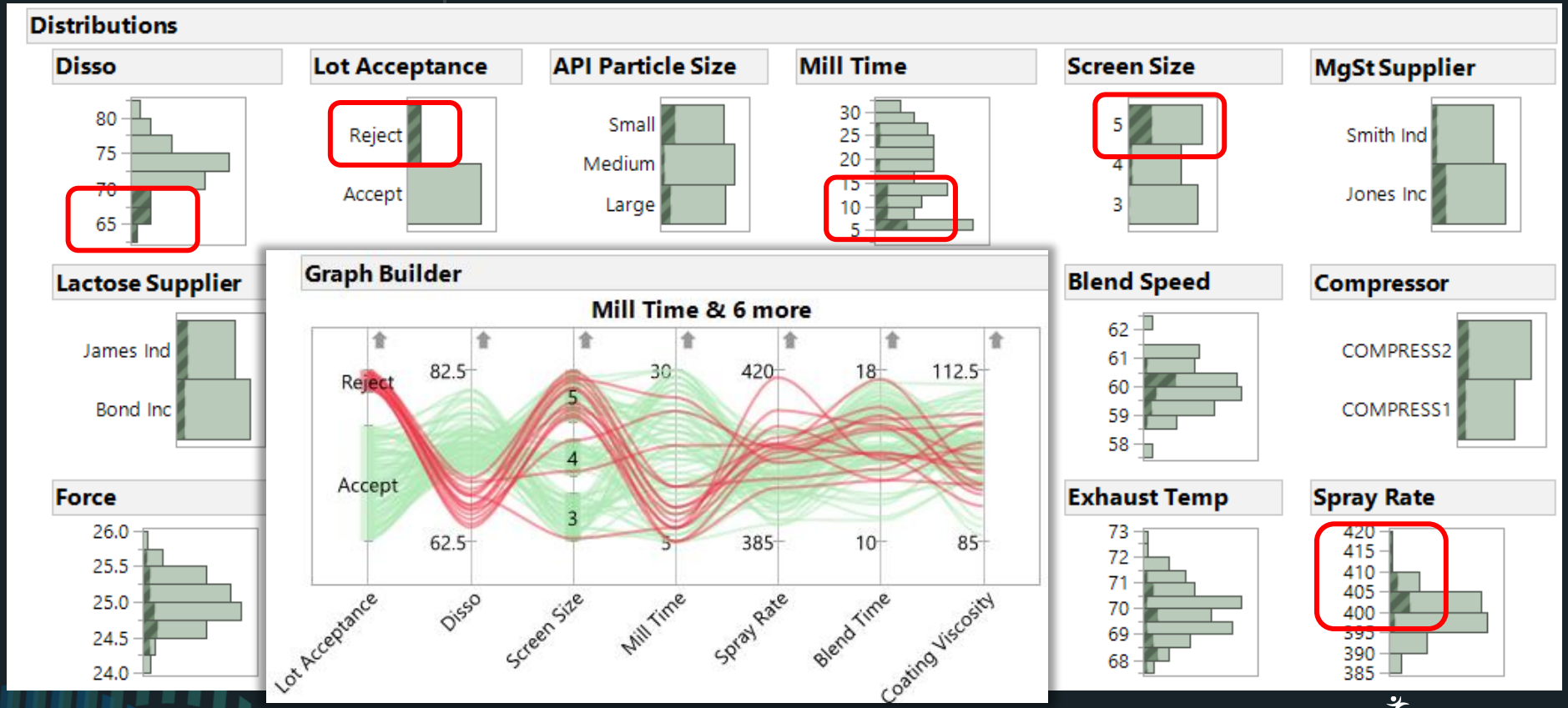
DYNAMIC LINKING AMONG ALL GRAPHS AND DATA TABLE ALLOWS SELECTION OF SUSPICIOUS DATA IN ONE GRAPH AND SEE POSSIBLE CORRELATIONS IN ALL!

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Exploratory Data Analysis

DYNAMIC LINKING AMONG ALL GRAPHS AND DATA TABLE ALLOWS SELECTION OF SUSPICIOUS DATA IN ONE GRAPH AND SEE POSSIBLE CORRELATIONS IN ALL!

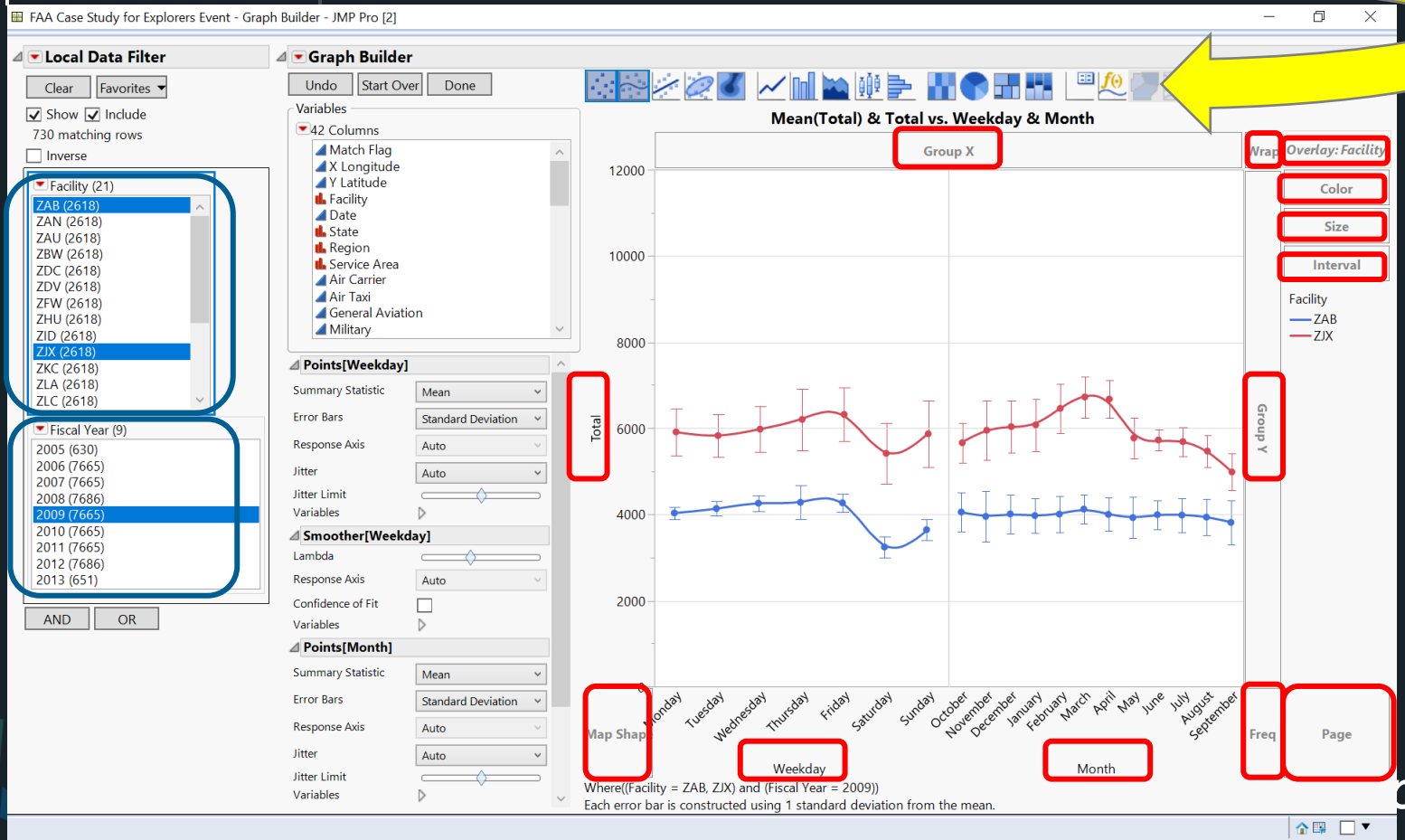


Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Highly Interactive Graph Builder

DRAG & DROP VARIABLES TO DROP ZONES TO EXPLORE RELATIONSHIPS FILTER DATA TO DRILL DOWN TO FIND INSIGHTS, MORE THAN A DOZEN *GRAPH ELEMENTS*



STATISTICAL DISCOVERY

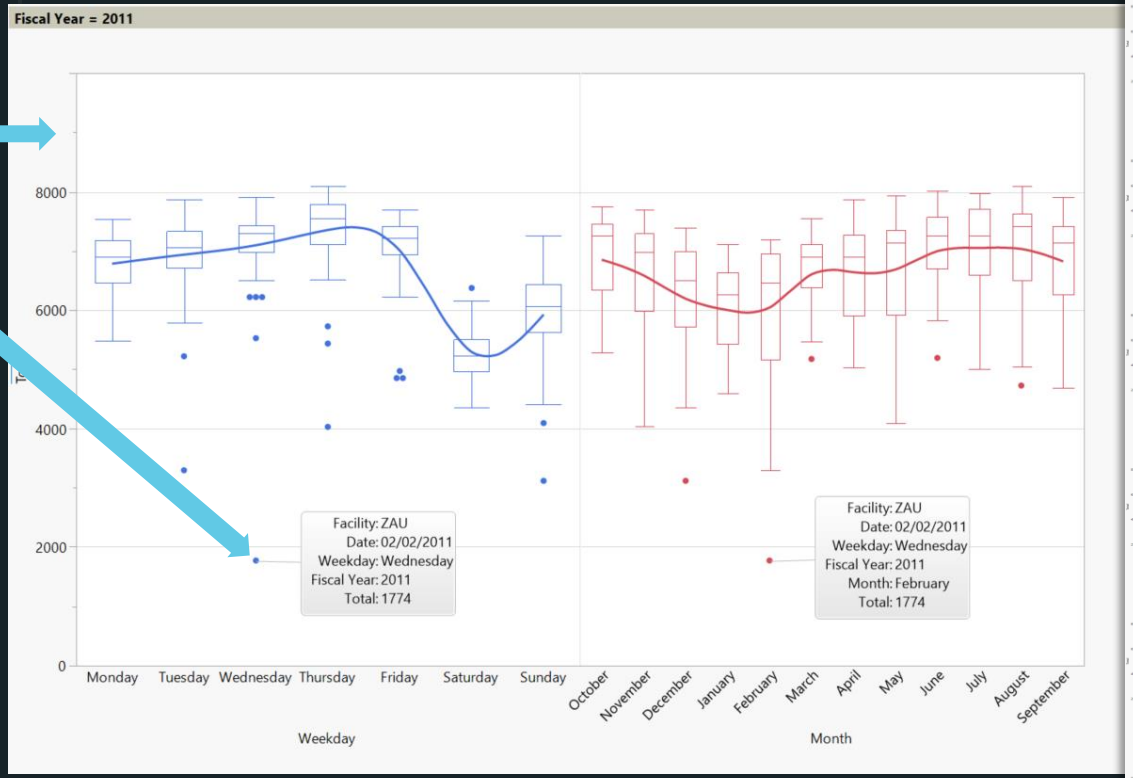
Highly Interactive Graph Builder

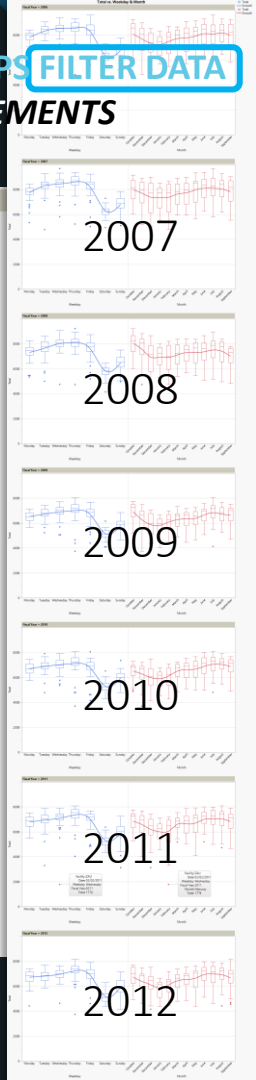
DRAG & DROP VARIABLES TO DROP ZONES TO EXPLORE RELATIONSHIPS FILTER DATA
TO DRILL DOWN TO FIND INSIGHTS, MORE THAN A DOZEN *GRAPH ELEMENTS*

Using ***Outlier Box Plot*** Graph Element
Filter ***Facility*** on ZAU (Chicago/Aurora)
Drop ***Fiscal Year*** in Page Zone to get

Select most extreme outlier

Search on "2/2/2011 & Chicago"
Top hit is "Groundhog Day Blizzard"

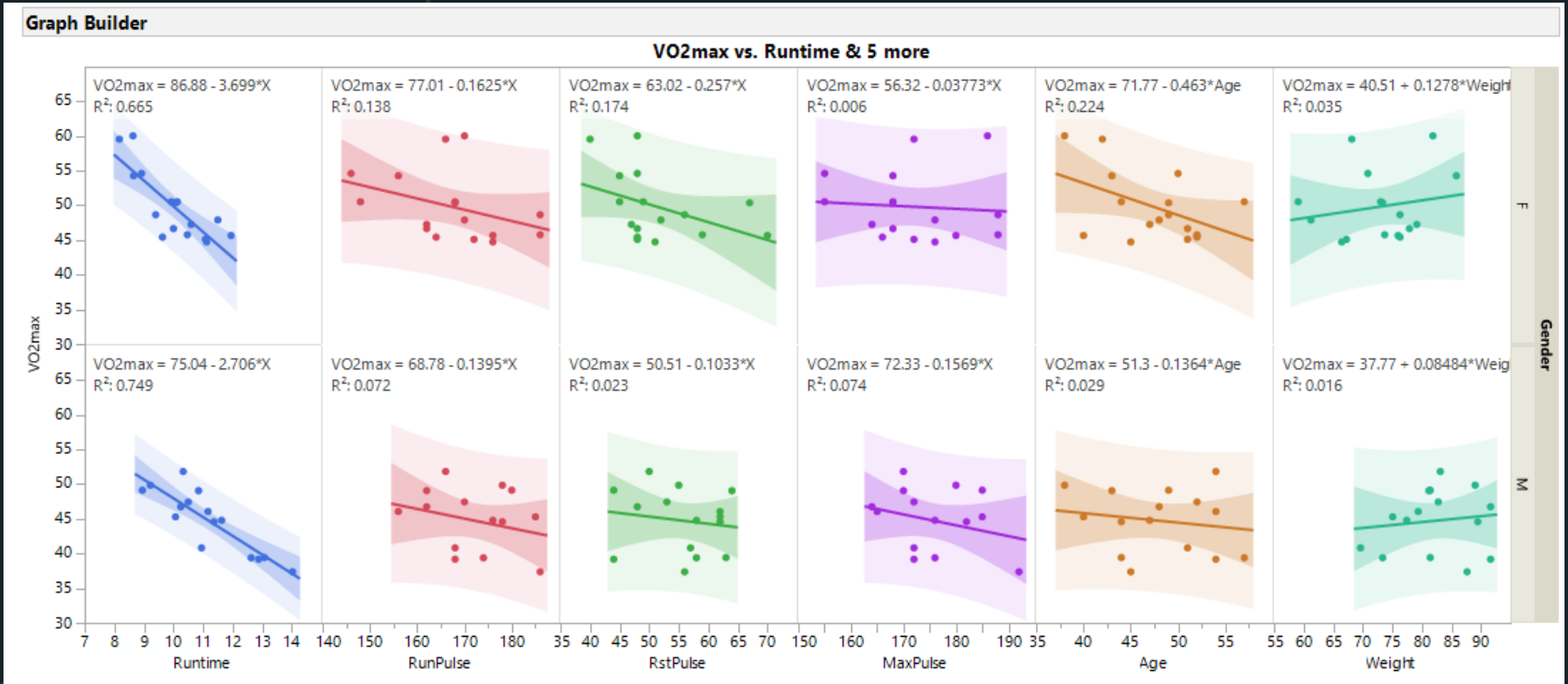




Copyright © JMP Statistical Discovery LLC. All rights reserved.

Multivariate Visual Modeling

DRAG & DROP INDIVIDUAL FITS OF EACH CONTINUOUS FACTOR BY GENDER

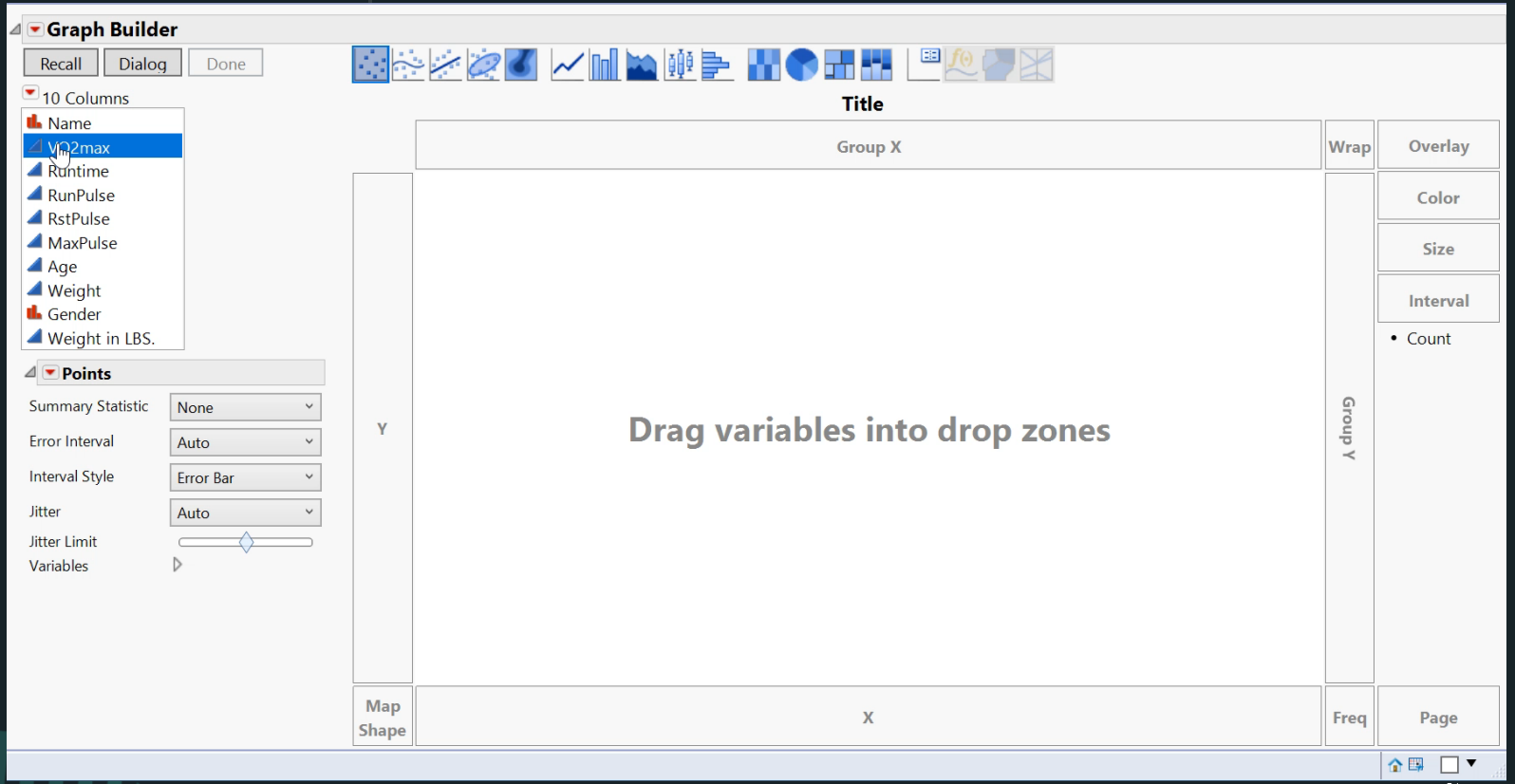


Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Multivariate Visual Modeling

DRAG & DROP INDIVIDUAL FITS OF EACH CONTINUOUS FACTOR BY GENDER

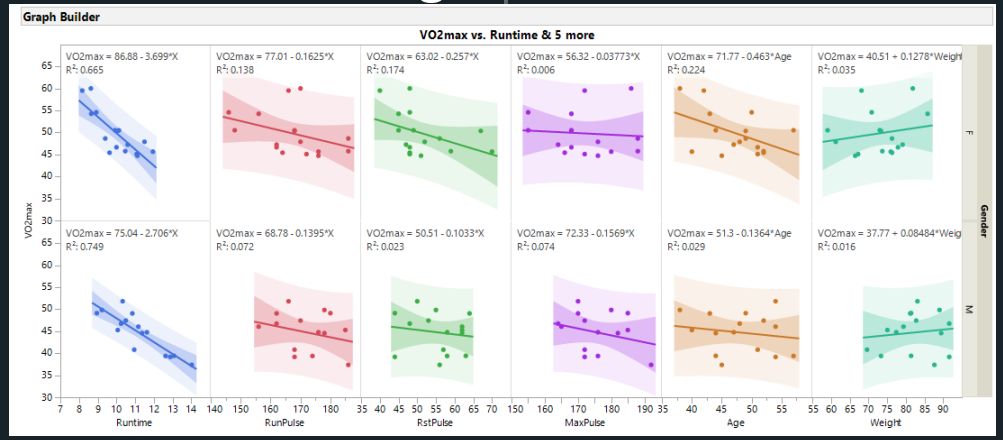


Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Multivariate Visual Modeling

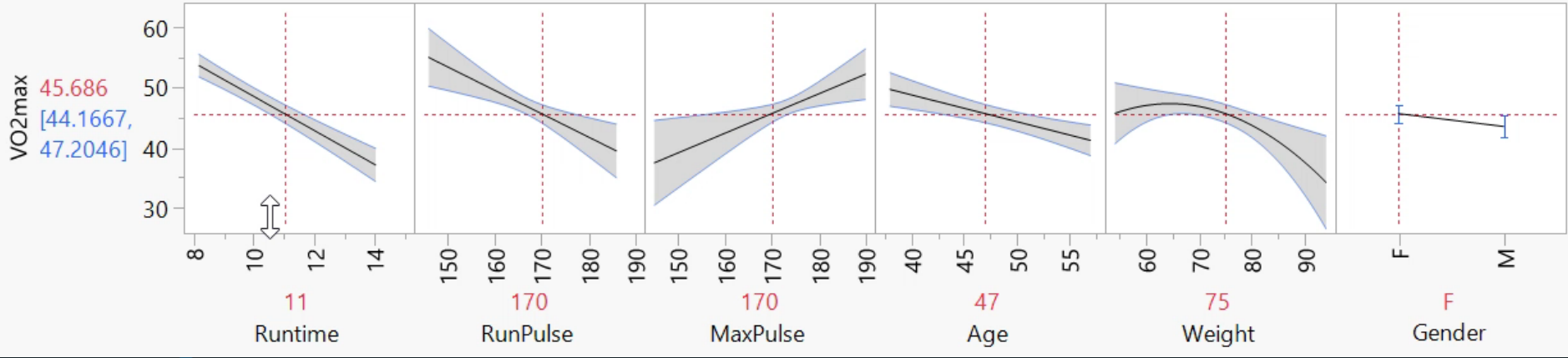
VISUALLY IDENTIFY MOST LIKELY TERMS IN FINAL MODEL



Effect Summary

Source	LogWorth	PValue
Runtime	8.405	0.00000
RunPulse	2.823	0.00150
Weight*Gender	2.741	0.00181
Age	2.198	0.00634
Age*Gender	2.179	0.00663
MaxPulse	1.963	0.01088
Weight*Weight	1.392	0.04056
Weight	0.865	0.13652 ^
Gender	0.161	0.69039 ^

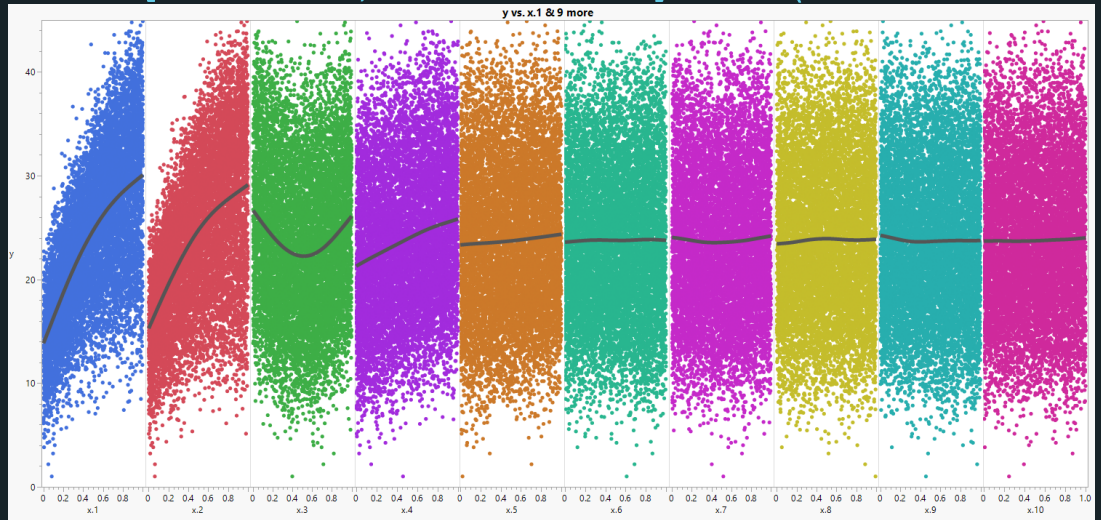
Prediction Profiler



Visual Data Discovery with 200 Continuous Factors!

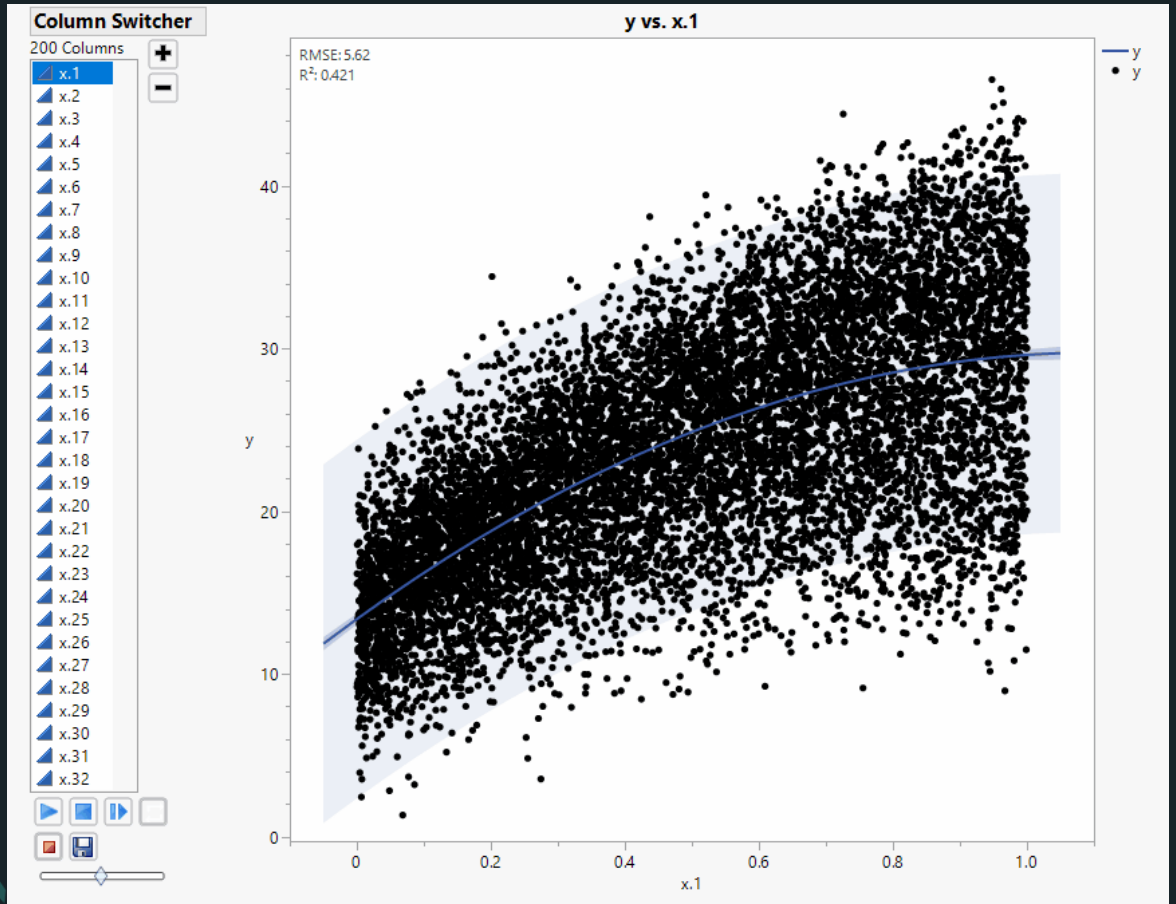
This function has only 5 active factors (*x1, x2, x3, x4, & x5*).

$y = 7$	(intercept)
$+ 20*x1 + 17*x2 + (-20)*x3 + 10*x4 + 6*x5$	(5 main effects - *slopes*)
$+ (-13)*x1*x1 + (-13)*x2*x2 + 20*x3*x3$	(3 squared effects - *curvature*)
$+ 19*x1*x2 + (-10)*x4*x5$	(2 interaction effects - *twists*)
+ Random Normal [Mean = 0, StdDev = 2.8]	(Random error)



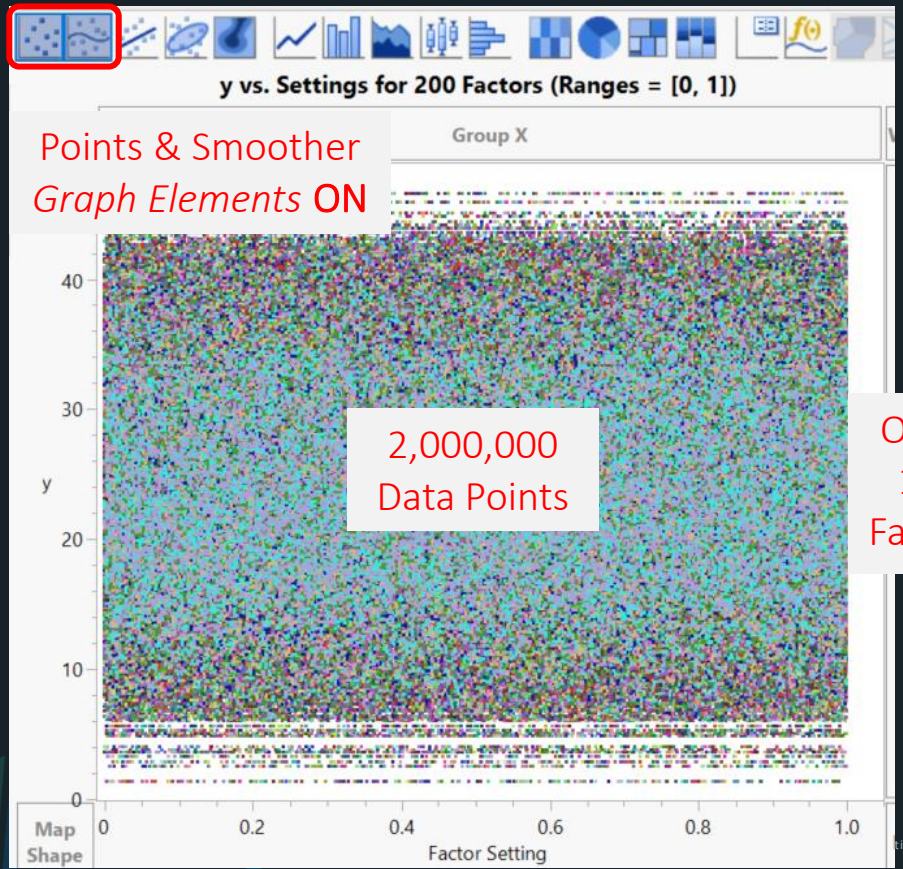
jmp STATISTICAL DISCOVERY

Animate Column Switcher to View Each Plot of y vs x_i

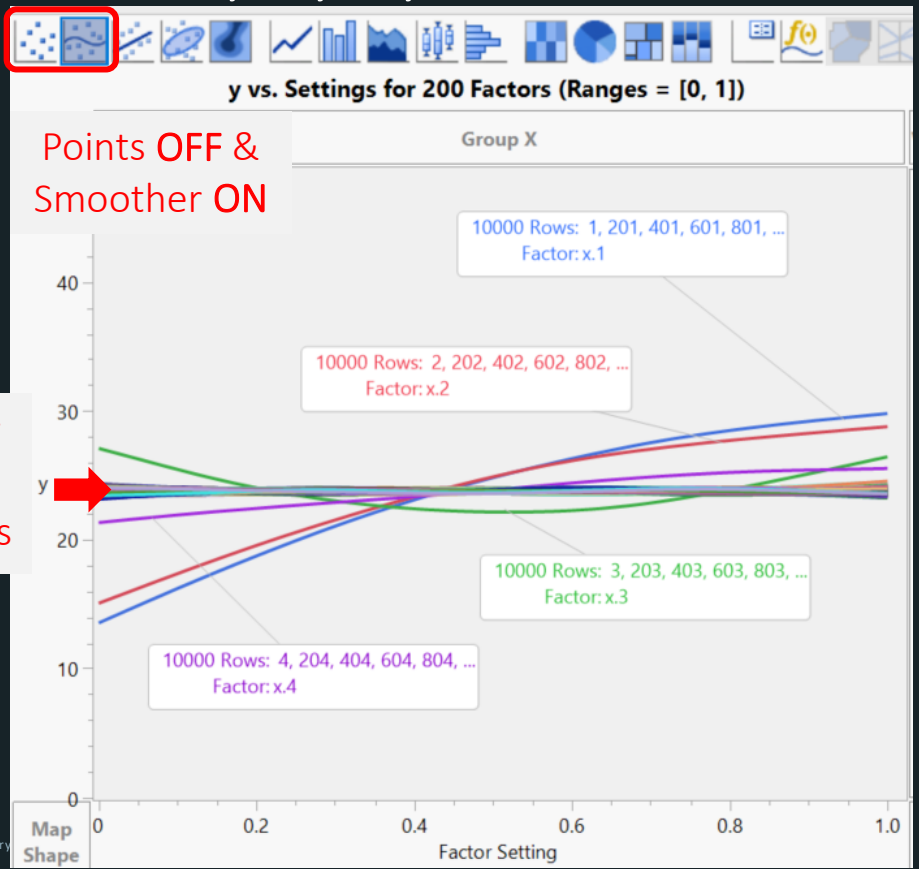


jmp STATISTICAL DISCOVERY

Stack ALL Factors (200 X 10K = 2M Rows), Then *Overlay* by Factor, Eliminate the Data Points to View the 200 Smoothers, to See the 4 Dominant Factors *x1*, *x2*, *x3*, & *x4*



Other 196 Factors



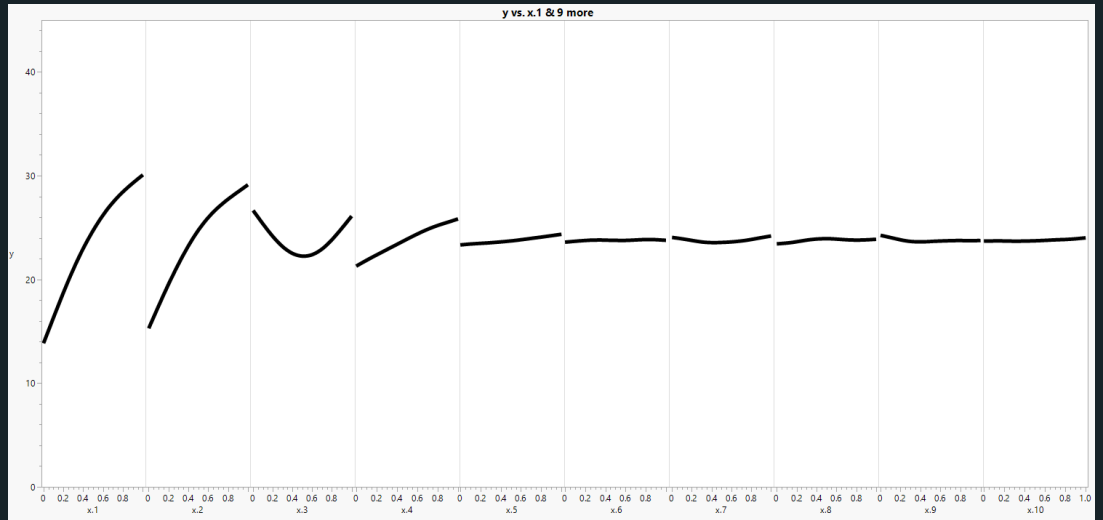
Further **Visual Data Discovery** with 200 Continuous Factors!
This function has only 5 active factors (*x1, x2, x3, x4, & x5*).

$$y = 7 \quad \text{(intercept)}$$

$$+ 20*x1 + 17*x2 + (-20)*x3 + 10*x4 + 6*x5 \quad \text{(5 main effects - } slopes\text{)}$$

$$+ (-13)*x1*x1 + (-13)*x2*x2 + 20*x3*x3 \quad \text{(3 squared effects - } curvature\text{)}$$

$$+ 19*x1*x2 + (-10)*x4*x5 \quad \text{(2 interaction effects - } twists\text{)}$$



jmp STATISTICAL DISCOVERY

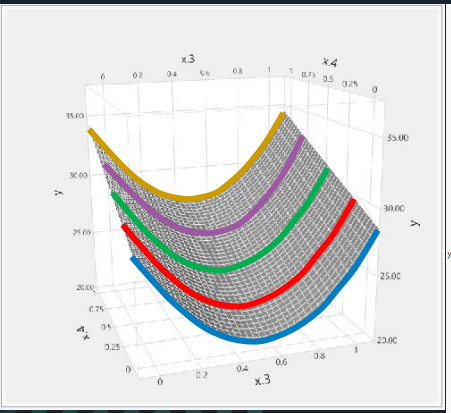
Overlaying by *x4* to See *Curvature* in *x3* without an Interaction between *x3* & *x4*

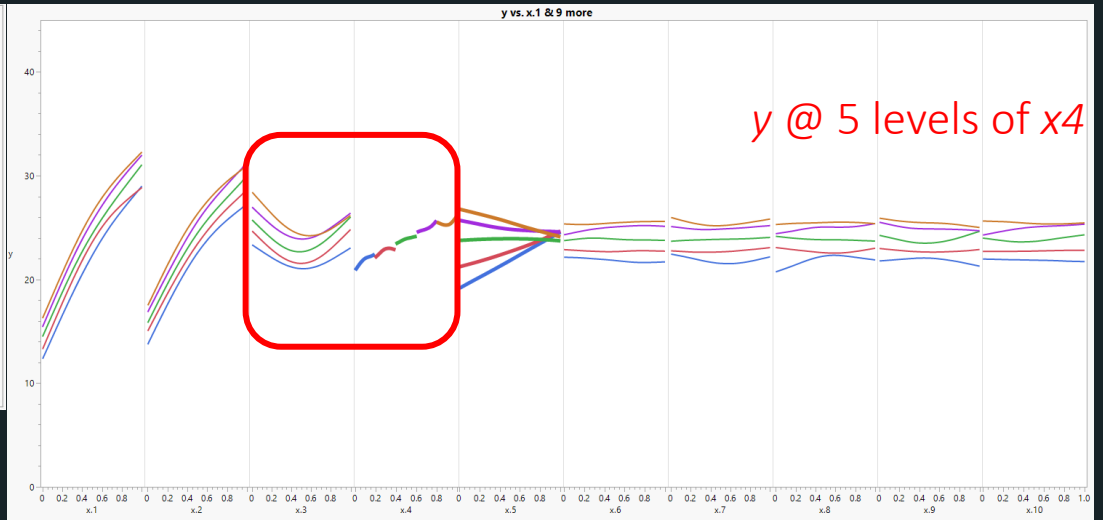
$y = 7$ (intercept)

$+ 20*x1 + 17*x2 + (-20)*x3 + 10*x4 + 6*x5$ (5 main effects - *slopes*)

$+ (-13)*x1*x1 + (-13)*x2*x2 + 20*x3*x3$ (3 squared effects - *curvature*)

$+ 19*x1*x2 + (-10)*x4*x5$ (2 interaction effects - *twists*)





x3	lin	sq	Sum int+lin+sq
0	0	0	0
0.1	-2	0.2	-1.8
0.2	-4	0.8	-3.2
0.3	-6	1.8	-4.2
0.4	-8	3.2	-4.8
0.5	-10	5	-5
0.6	-12	7.2	-4.8
0.7	-14	9.8	-4.2
0.8	-16	12.8	-3.2
0.9	-18	16.2	-1.8
1	-20	20	0

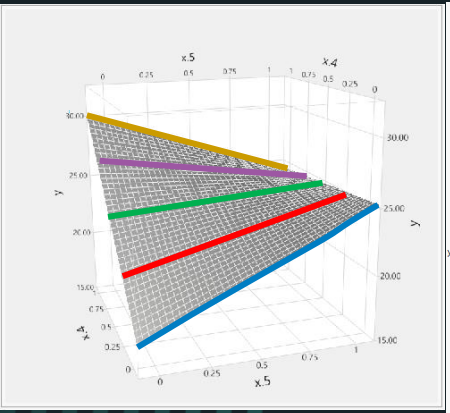
Overlaying by *x4* to See *Interaction* with *x5* without any Curvature in *x4* & *x5*

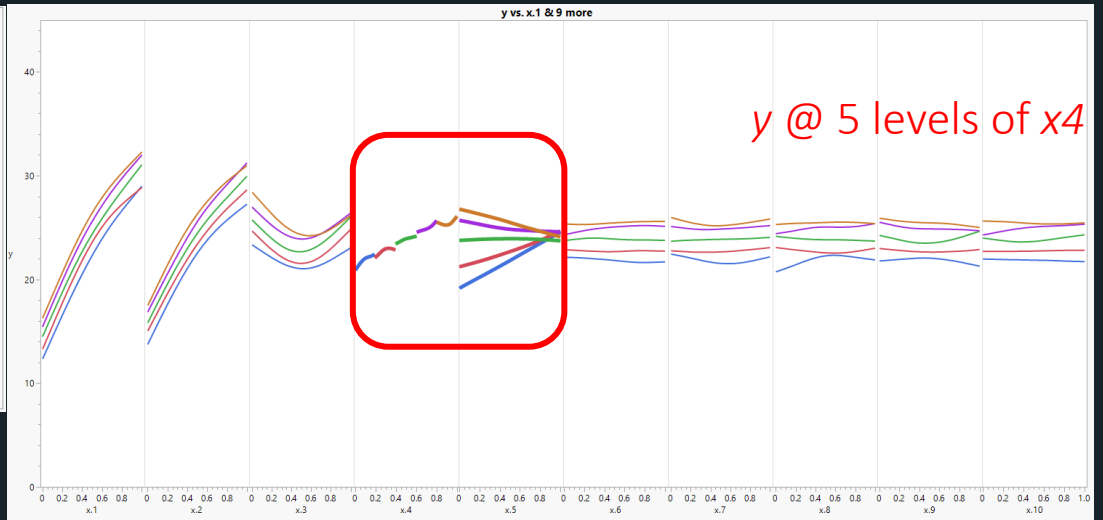
y = 7 (intercept)

+ 20**x1* + 17**x2* + (-20)**x3* + **10****x4* + **6****x5* (5 main effects - *slopes*)

+ (-13)**x1***x1* + (-13)**x2***x2* + 20**x3***x3* (3 squared effects - *curvature*)

+ 19**x1***x2* + (**-10**)**x4***x5* (2 interaction effects - *twists*)





jmp STATISTICAL DISCOVERY

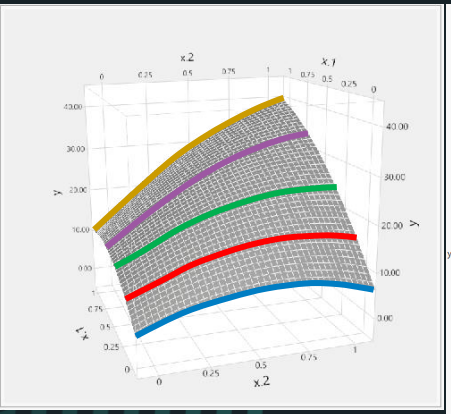
Overlaying by *x1* to See *Interaction* with *x2* and *Curvature* in both *x1* & *x2*

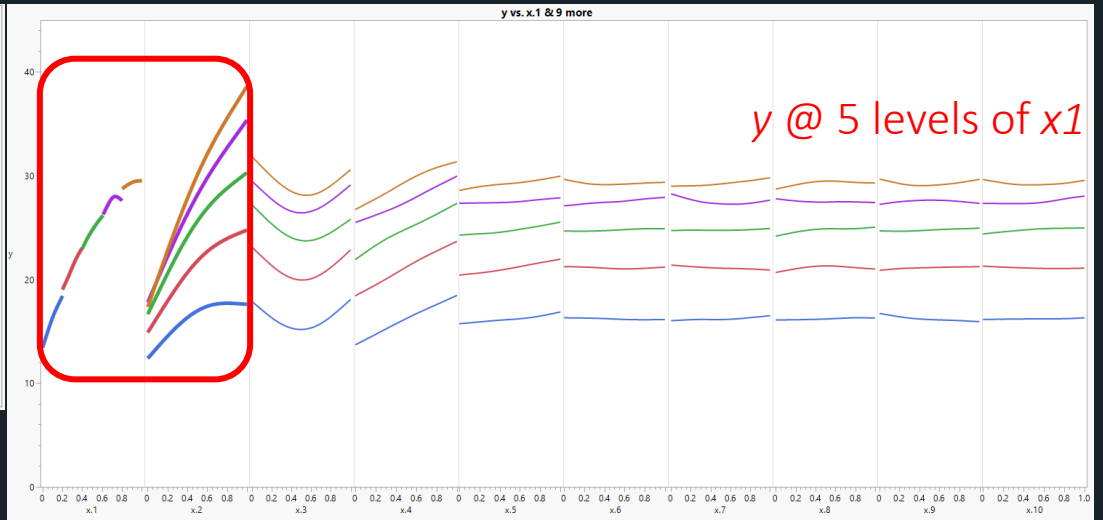
y = 7 (intercept)

+ 20**x1* + 17**x2* + (-20)**x3* + 10**x4* + 6**x5* (5 main effects - *slopes*)

+ (-13)**x1***x1* + (-13)**x2***x2* + 20**x3***x3* (3 squared effects - *curvature*)

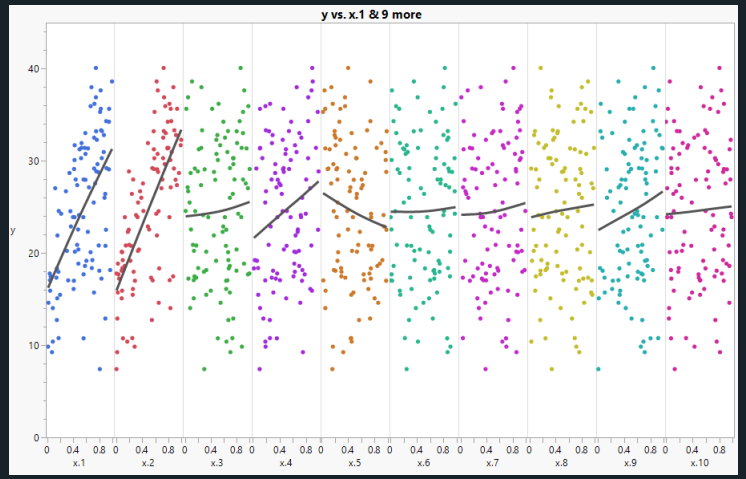
+ 19**x1***x2* + (-10)**x4***x5* (2 interaction effects - *twists*)

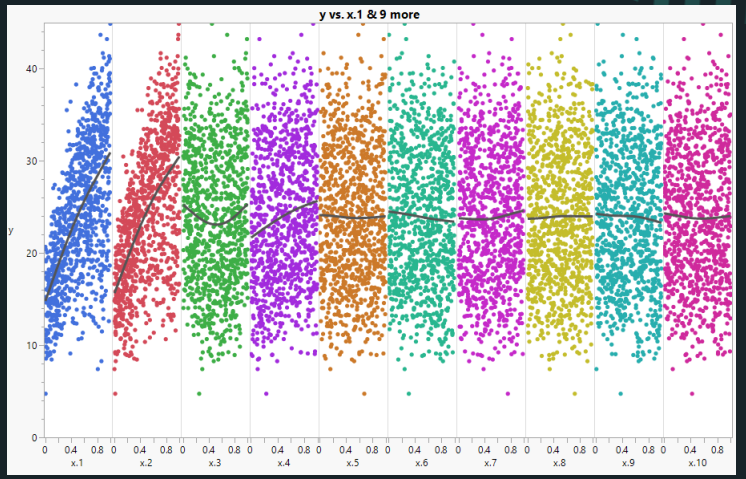


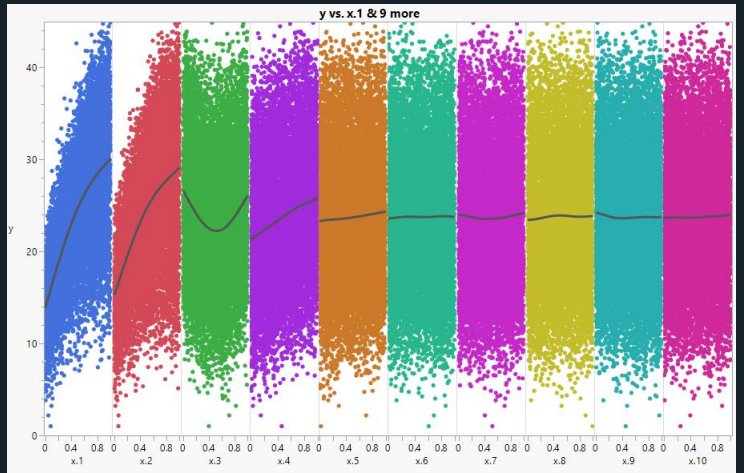


jmp STATISTICAL DISCOVERY

Compare Results for 100, 1,000, and 10,000 Data Points







Generating Function for Data Points:

$y = 7$

$+ 20*x1 + 17*x2 + (-20)*x3 + 10*x4 + 6*x5$

$+ (-13)*x1*x1 + (-13)*x2*x2 + 20*x3*x3$

$+ 19*x1*x2 + (-10)*x4*x5$

+ Random Normal [Mean = 0, StdDev = 2.8]

The more data used, the better the spline fit approximates the generating function

jmp STATISTICAL DISCOVERY

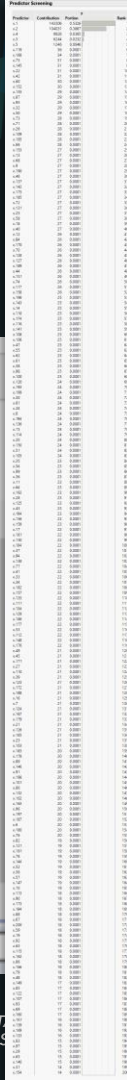
After Doing **Visual Data Discovery,** Run ***Predictor Screening*** to Quickly Rank the 200 Factors. There are only 5 active factors (*x1, x2, x3, x4, & x5*), ALL of which are Correctly Identified.

$$y = 7 + 20*x1 + 17*x2 + (-20)*x3 + 10*x4 + 6*x5 + (-13)*x1*x1 + (-13)*x2*x2 + 20*x3*x3 + 19*x1*x2 + (-10)*x4*x5$$

Predictor Screening

Predictor	y Contribution	y Portion	Rank
x.1	143306	0.5326	1
x.2	104031	0.3867	2
x.4	9826	0.0365	3
x.3	6244	0.0232	4
x.5	1246	0.0046	5
x.119	36	0.0001	6
x.168	34	0.0001	7
x.75	31	0.0001	8
x.145	31	0.0001	9
x.22	31	0.0001	10

Top 10 of 200 Shown



Copyright © JMP Statistical Discovery LLC. All rights reserved.

With Reduced Factor List, Easy to Use Machine Learning in JMP® Pro

- Partition (Regression/Decision Tree) – very fast
- Bootstrap Forest* used in Predictor Screening Averages Many Trees
  - Fast selection of important factors when you have many
  - Bootstrapping randomly picks factors so dominant ones are sometimes absent so they do not overshadow/obscure less important factors
- Neural Network – one or two layers, and boosted
  - Often best predictor, but tends to overfit
  - More computationally intensive than trees, especially with many factors
- Regression – Ordinary Least Squares, Logistic, Stepwise, or Penalized
  - More computationally intensive than trees, especially with many factors
  - Penalized regression can perform variable selection even with correlated data
  - Analyses provide confidence intervals

*JMP name for Random Forest™

Copyright © JMP Statistical Discovery LLC. All rights reserved.

Model Screening with Many Factors (250) and Many Rows of Data (10K)

- Model Screening results (left) with 250 factors and 10,000 rows of data is too big a problem for some methods (regression, SVM) or times out for others (neural network).
- After reviewing results for fast tree methods (e.g., Bootstrap Forest) use top ranked factors and rerun Model Screening (right) to obtain more and better outcomes.

Test

Method	N	RSquare	RASE
Boosted Tree	2001	0.8380	2.9795
Bootstrap Forest	2001	0.8039	3.2782
Decision Tree	2001	0.7649	3.5895
Neural Boosted	2001	0.5467	4.9838
K Nearest Neighbors	2001	0.1832	6.6900

Select Dominant Run Selected Save Script Selected

Support Vector Machines avoided when number of rows > 10000.Linear Model methods skipped because the number of parameters, 753, exceeded the limit of 450.Sum Freq and Sum Weight are suppressed when they are the same as N.

Method	Time Limit Exceeded Count
Neural Boosted	1

Test

Method	Details	N	RSquare	RASE
Fit Stepwise	2FI Quad	2001	0.8541	2.8272
Fit Least Squares	2FI Quad	2001	0.8541	2.8280
Generalized Regression Lasso	2FI Quad	2001	0.8541	2.8280
Neural Boosted		2001	0.8500	2.8666
Boosted Tree		2001	0.8424	2.9382
Bootstrap Forest		2001	0.8332	3.0237
K Nearest Neighbors		2001	0.8217	3.1259
Decision Tree		2001	0.7744	3.5161
Fit Least Squares		2001	0.7218	3.9040
Generalized Regression Lasso		2001	0.7218	3.9046
Fit Stepwise		2001	0.7217	3.9053

Select Dominant Run Selected Save Script Selected

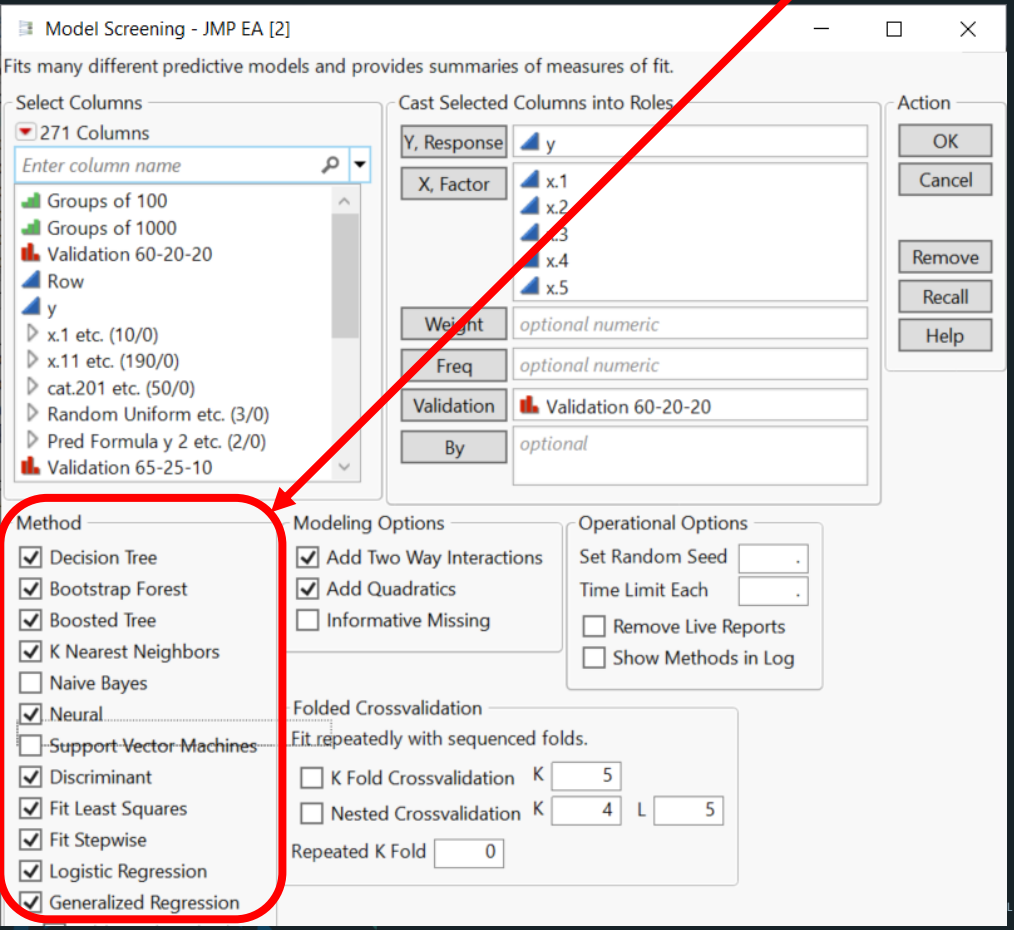
Support Vector Machines avoided when number of rows > 10000.Sum Freq and Sum Weight are suppressed when they are the same as N.

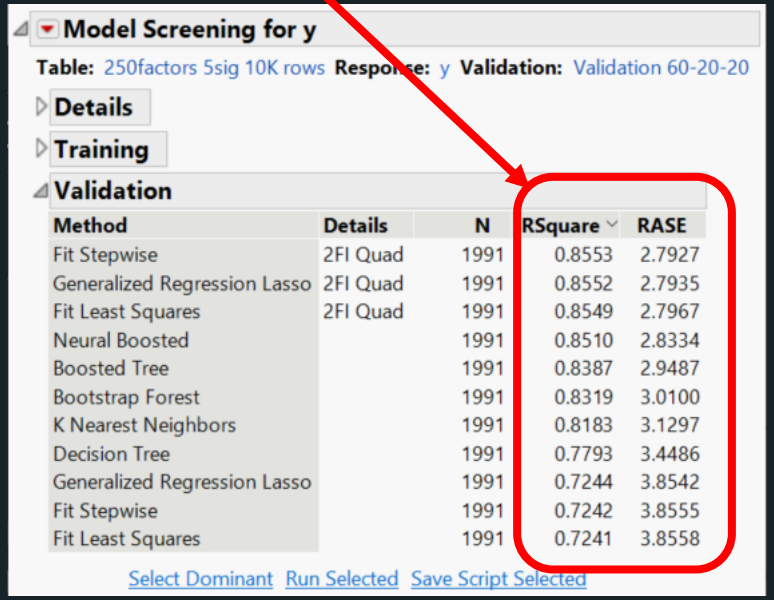
Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Model Screening

Run 10 selected models all at once (about half a minute in this case)
then view ranked performance



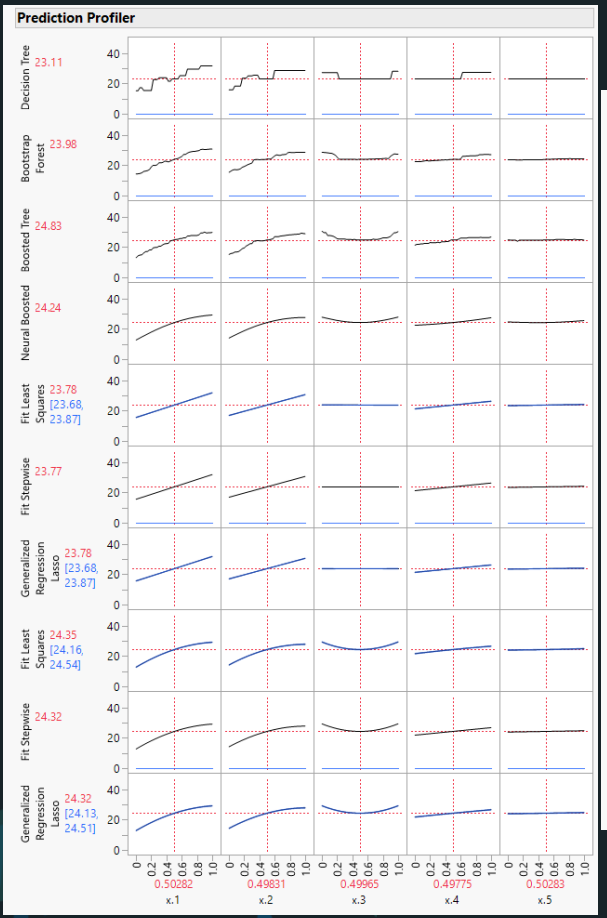


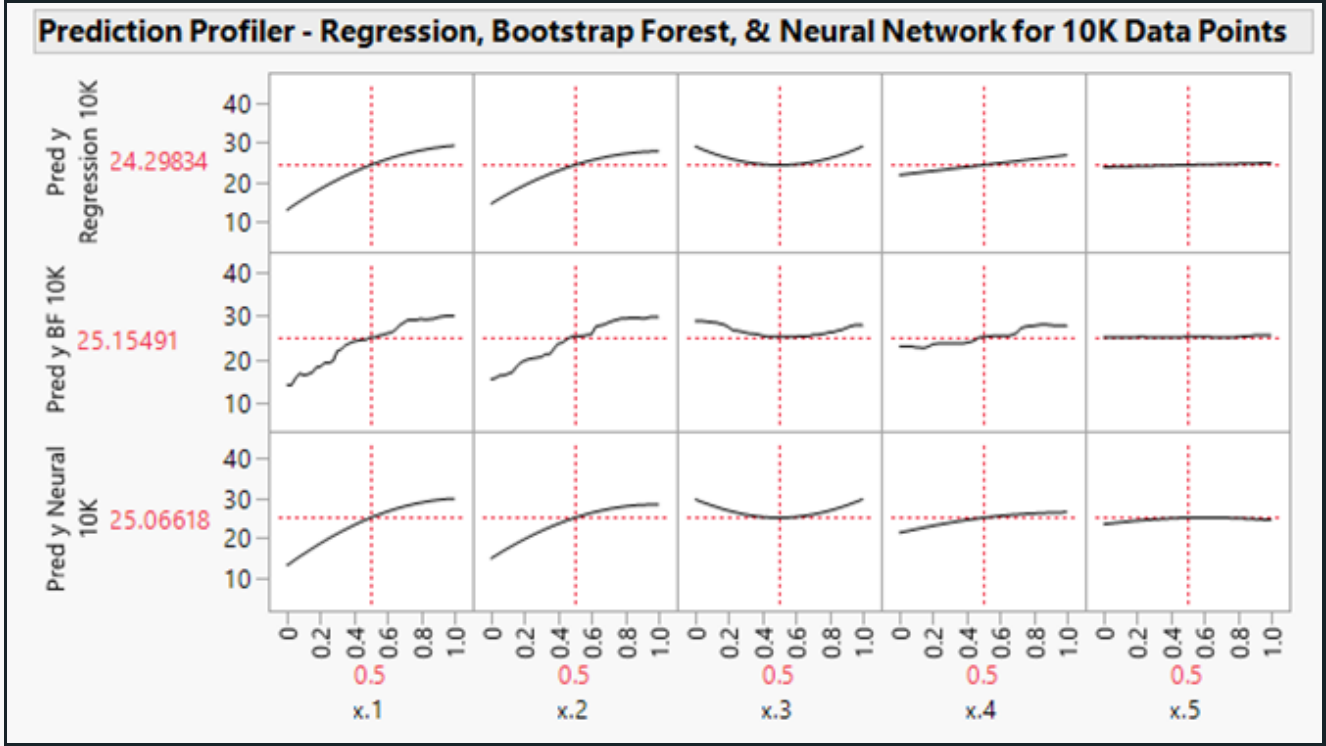
Publish models in:
Python, C, SAS, SQL, & JavaScript

LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Profilers for All 10 Models (left) & 3 Selected Models (right)

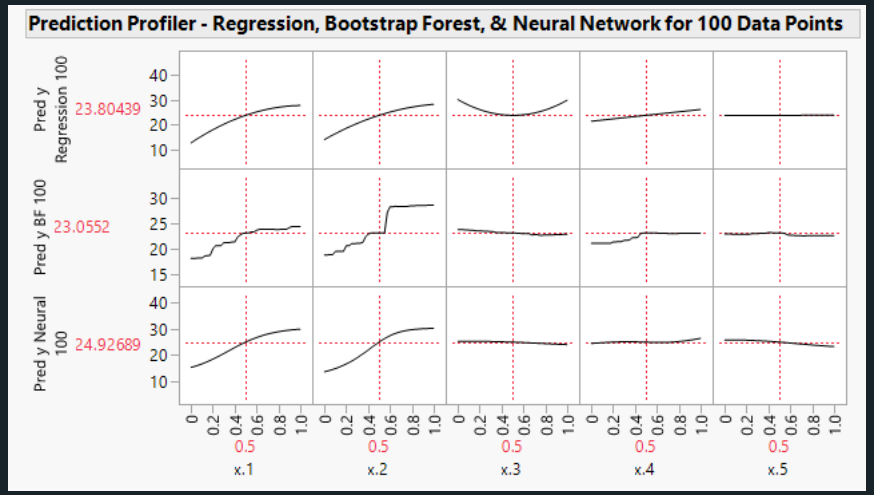


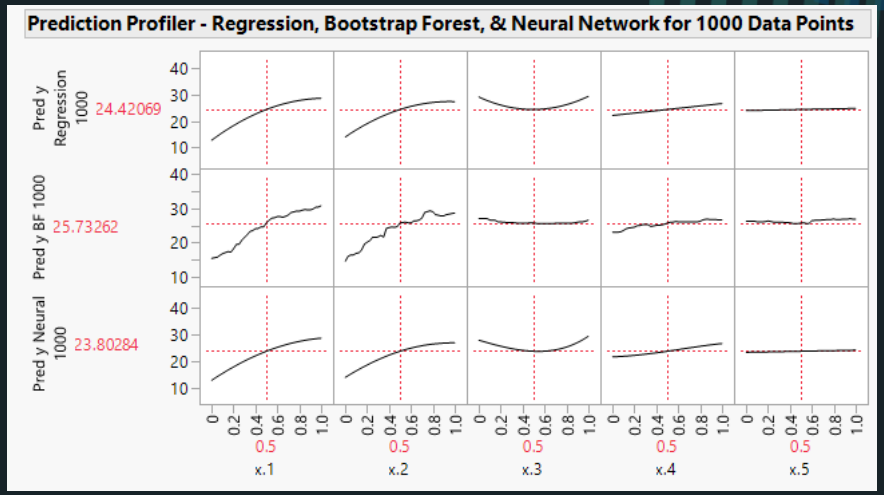


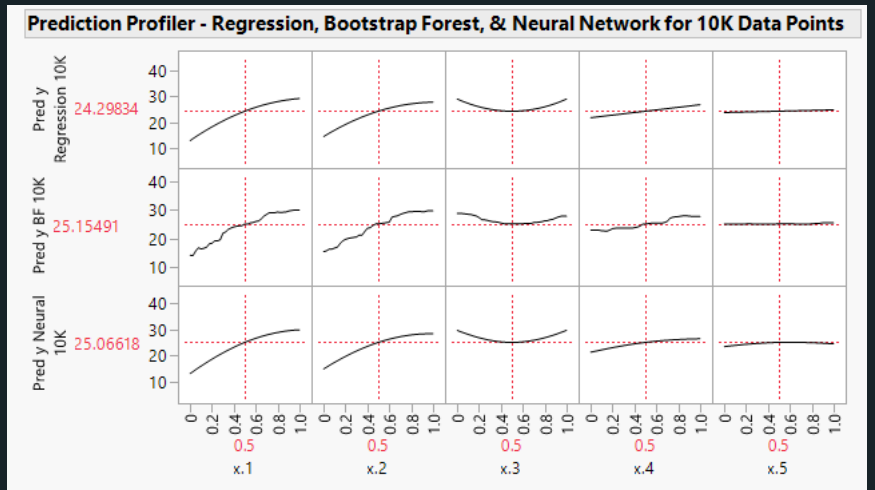
Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Compare Results for 100, 1,000, and 10,000 Data Points







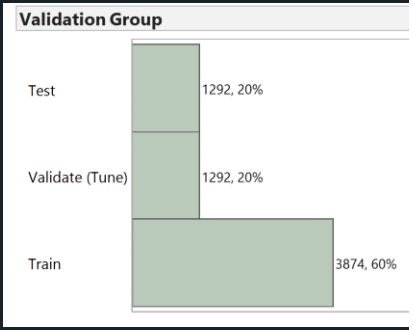
- Three ML methods per group
- Same amount of data in each group

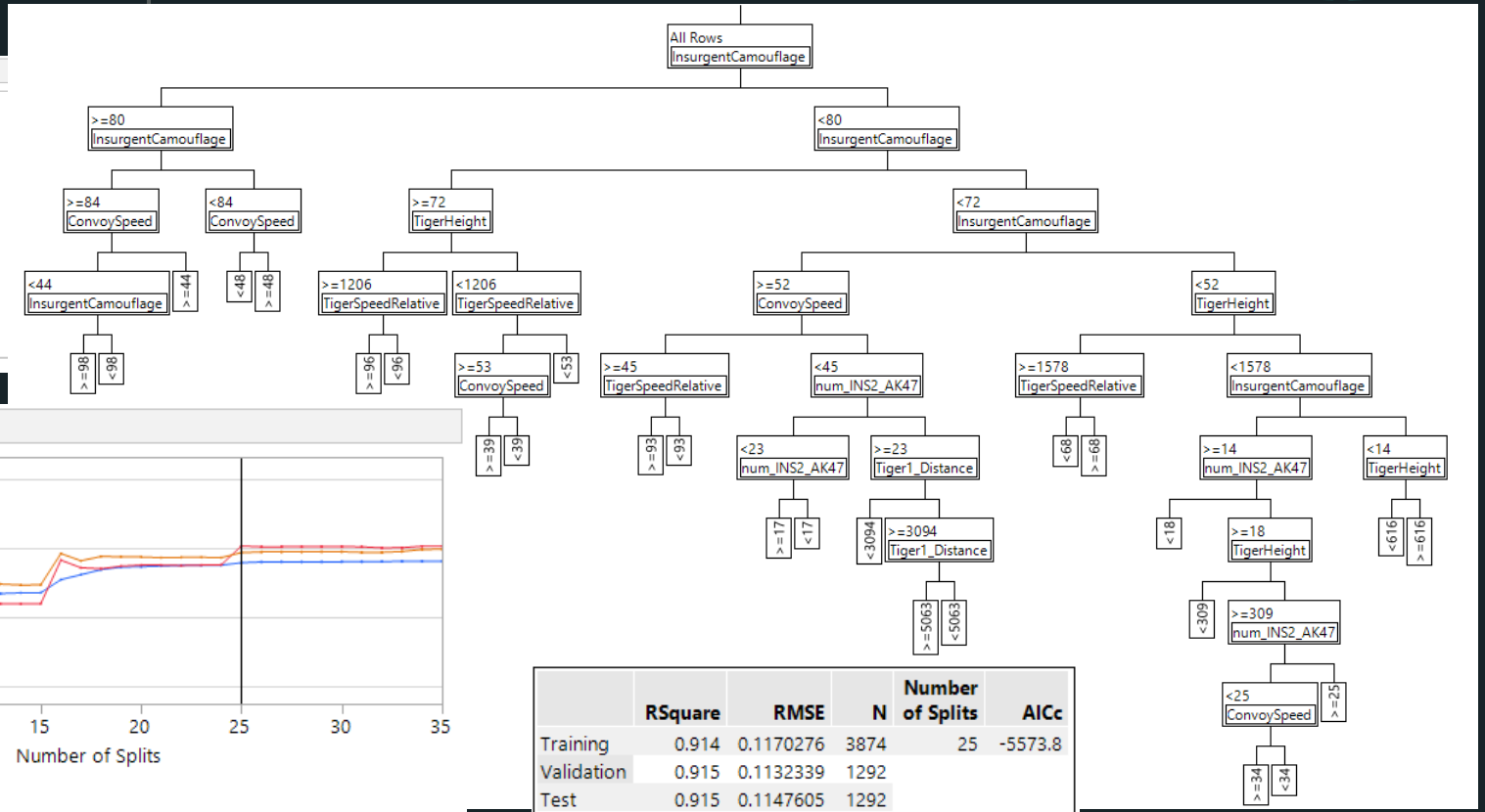
With enough data all three ML methods well approximate the generating function

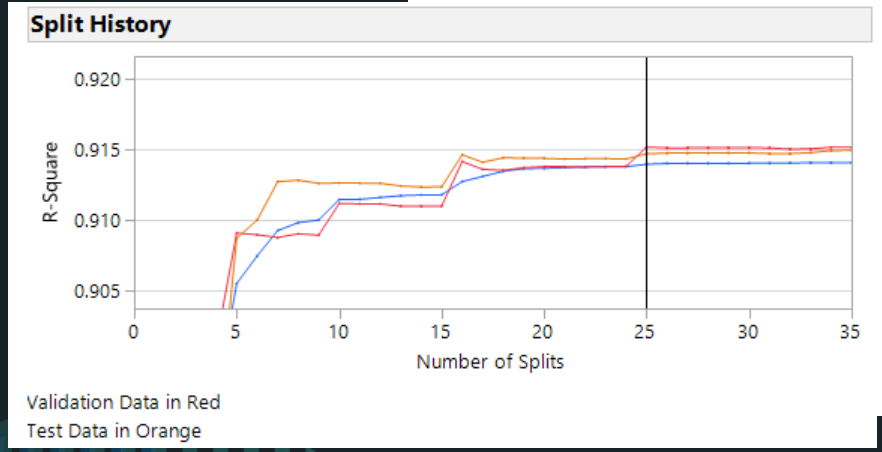
jmp STATISTICAL DISCOVERY

Honest Assessment

SUBSET DATA TO CREATE *TRAIN*, *VALIDATE*(*TUNE*), & *TEST* GROUPS
USE VALIDATE(TUNE) GROUP TO PREVENT OVER-FITTING DATA MINING MODELS







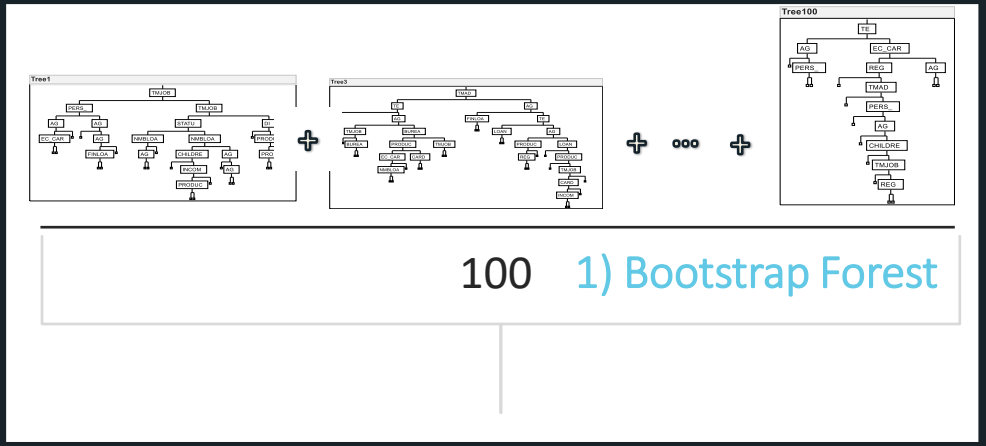
	RSquare	RMSE	N	Number of Splits	AICc
Training	0.914	0.1170276	3874	25	-5573.8
Validation	0.915	0.1132339	1292		
Test	0.915	0.1147605	1292		

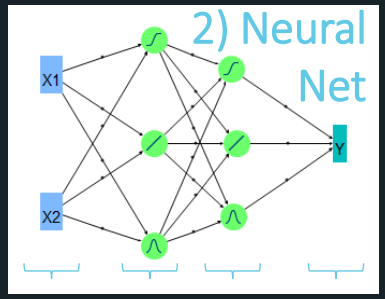
Copyright © JMP Statistical Discovery LLC. All rights reserved.

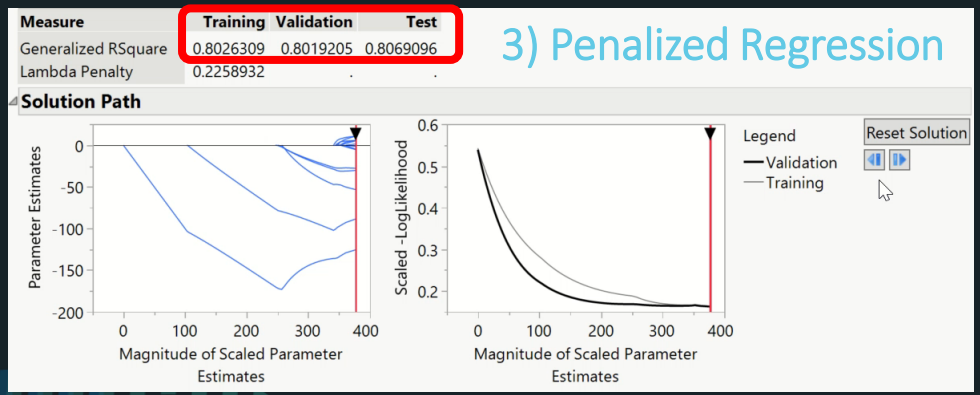
JMP STATISTICAL DISCOVERY

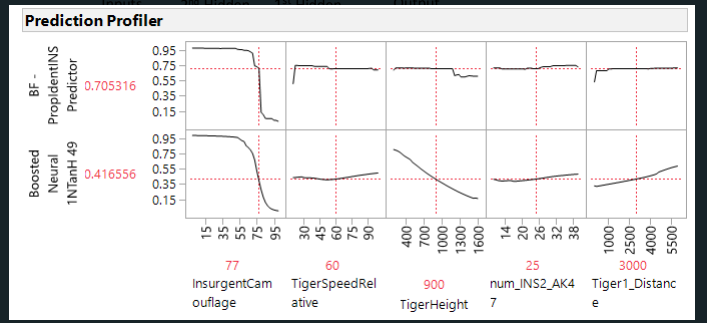
Robust Strategy for Machine Learning

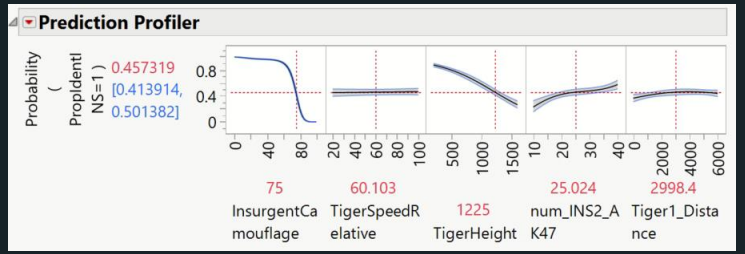
1) Bootstrap Forest – *FAST even w/many Xs – Unlikely to Miss Less Dominant Xs*
2) Neural Network – *Often Most Flexible & Best Predictor – Tendency to Overfit*
3) Penalized Regression – *Often More Interpretable Model + Confidence Intervals*









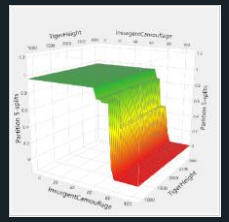


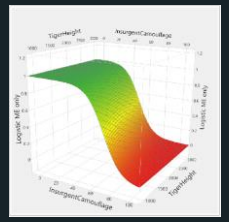
NOT just better prediction, but better understanding!

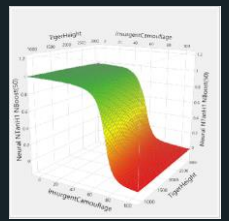
Copyright © JMP Statistical Discovery LLC. All rights reserved.

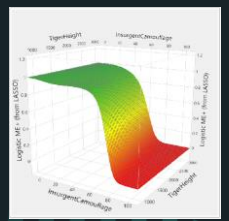
Visually Compare Multiple Machine Learning Models

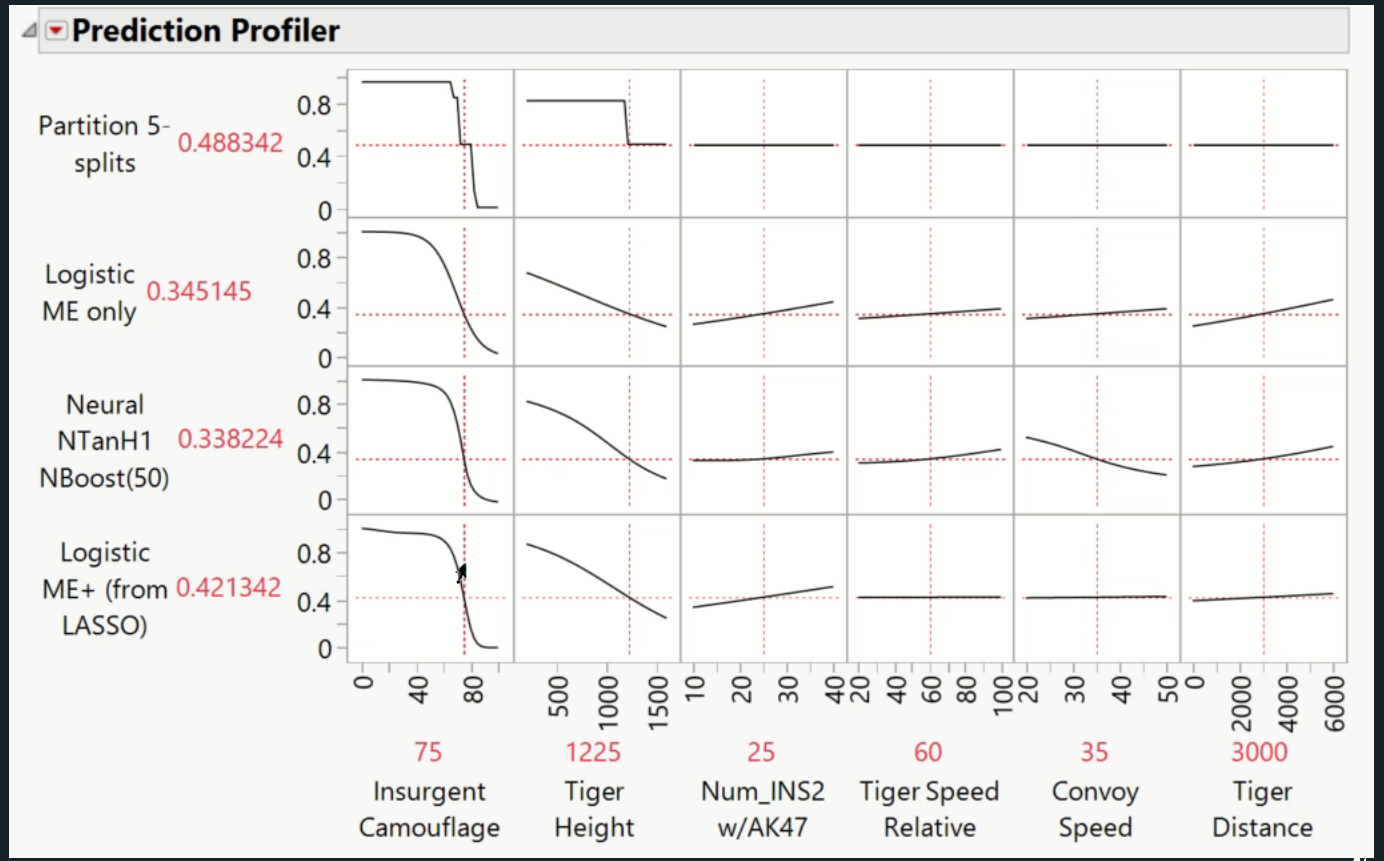
Partition with 5-Splits, Logistic Regression, Neural Network, & LASSO (Binomial)











Simplest

Classic

Best Predictor

More Interpretable

"LASSO" stands for **L**east **A**bsolute **S**hrinkage and **S**election **O**perator

Copyright © SAS Institute Inc. All rights reserved.

jmp STATISTICAL DISCOVERY

Actual vs. Prediction Plots for 4 Surrogate Models for *Test Data ONLY* (Not used in fitting or tuning the models)

Column Switcher

4 Columns

- Partition 5-splits
- Logistic ME only
- Neural NTanH1 NBoost(50)
- Logistic ME+ (from LASSO)

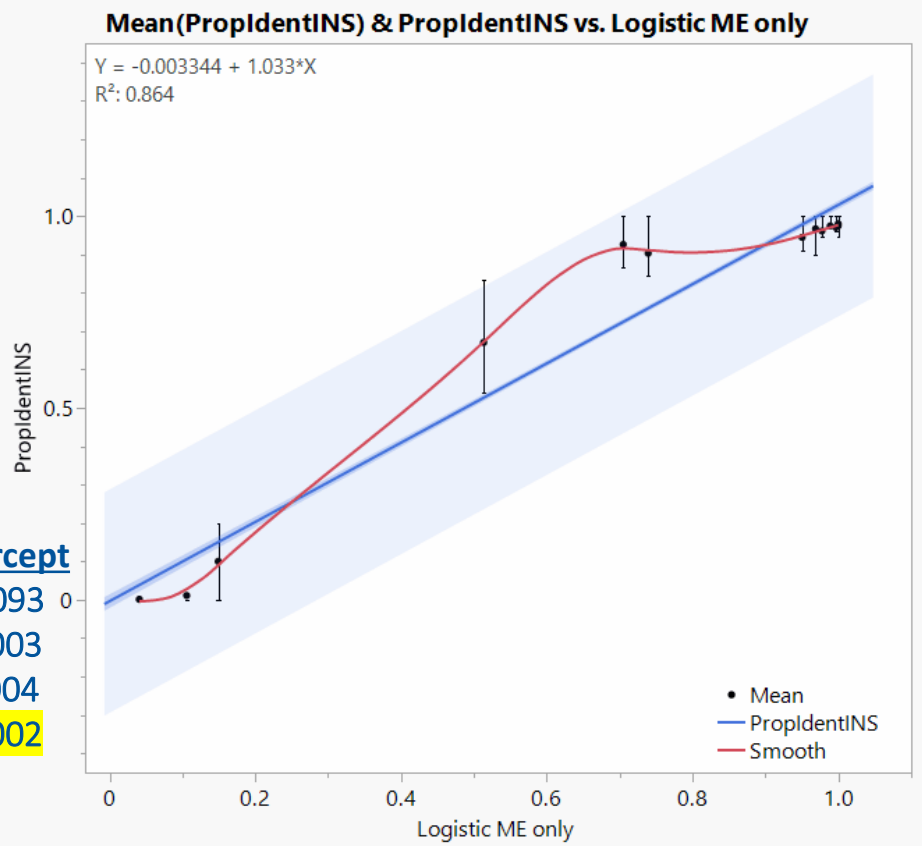
Local Data Filter

1296 matching rows

☐ Inverse

Group-Strat Val (3)

Training	3876
Validation	1286
Test	1296

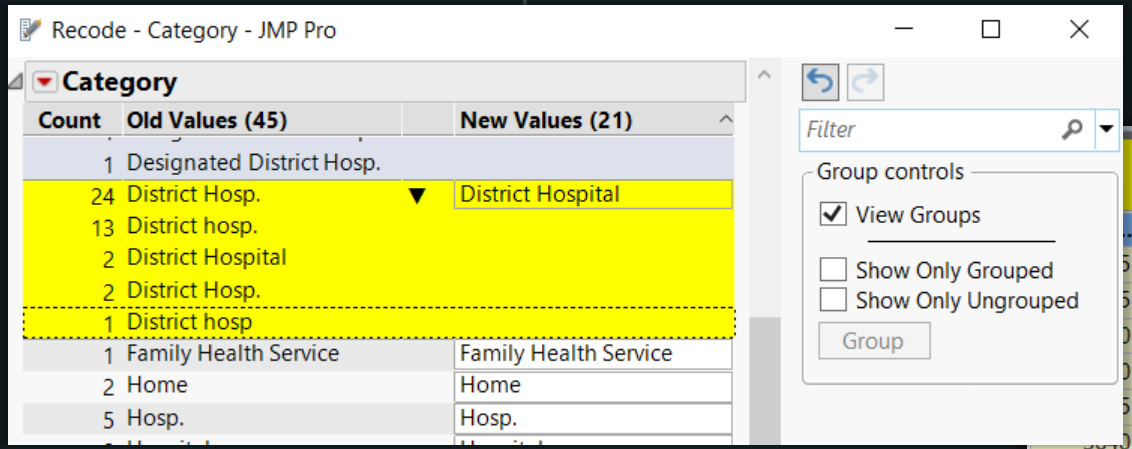


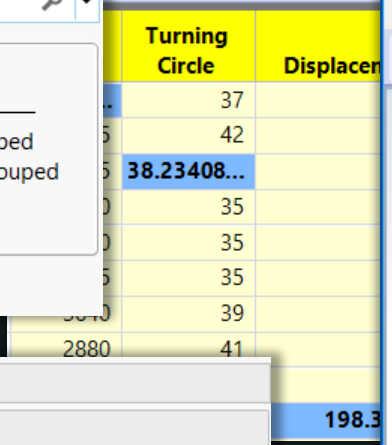
Models	R^2	Slope	Intercept
Partition with 5-Splits	0.847	1.06	-0.093
Logistic Regression	0.864	1.03	-0.003
Neural Network	0.914	0.99	0.004
LASSO Binomial	0.913	1.00	-0.002

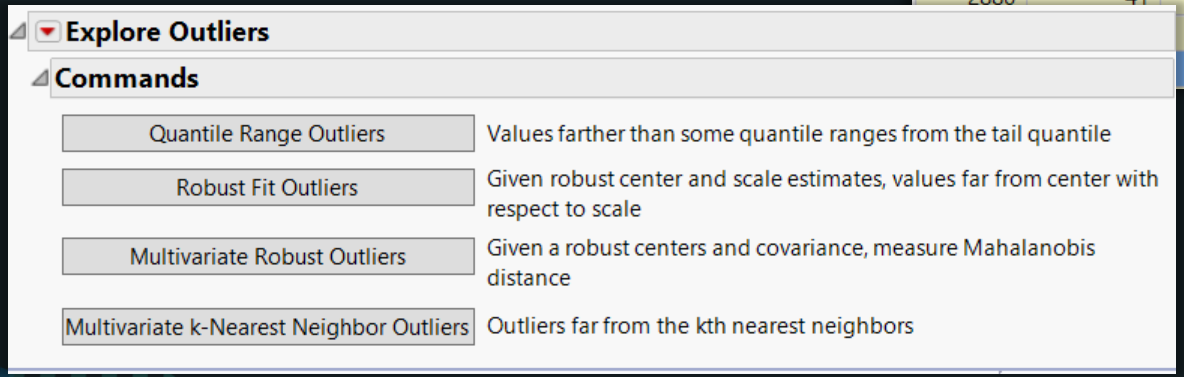
Data Curation

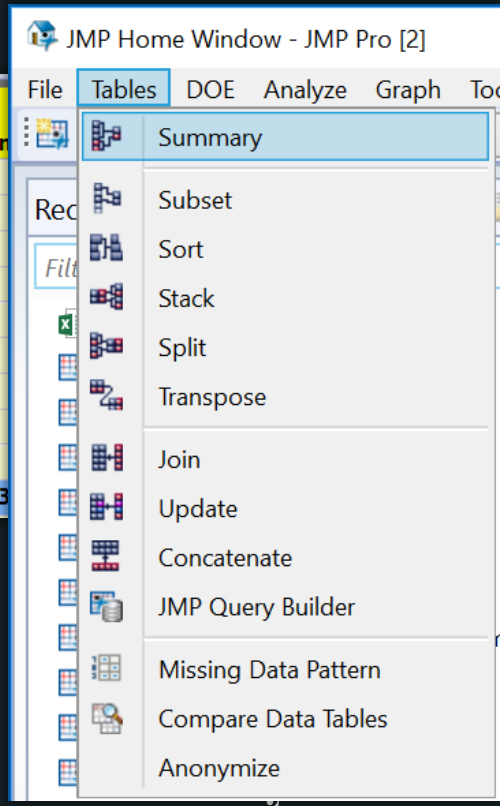
RECODE, OUTLIER DETECTION, AND IMPUTE MISSING VALUES, STACK, SPLIT, ETC.

"WHAT YOU JUST DID IN MINUTES TAKES ME HOURS (DAYS!) IN EXCEL…"









Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Design of Experiments

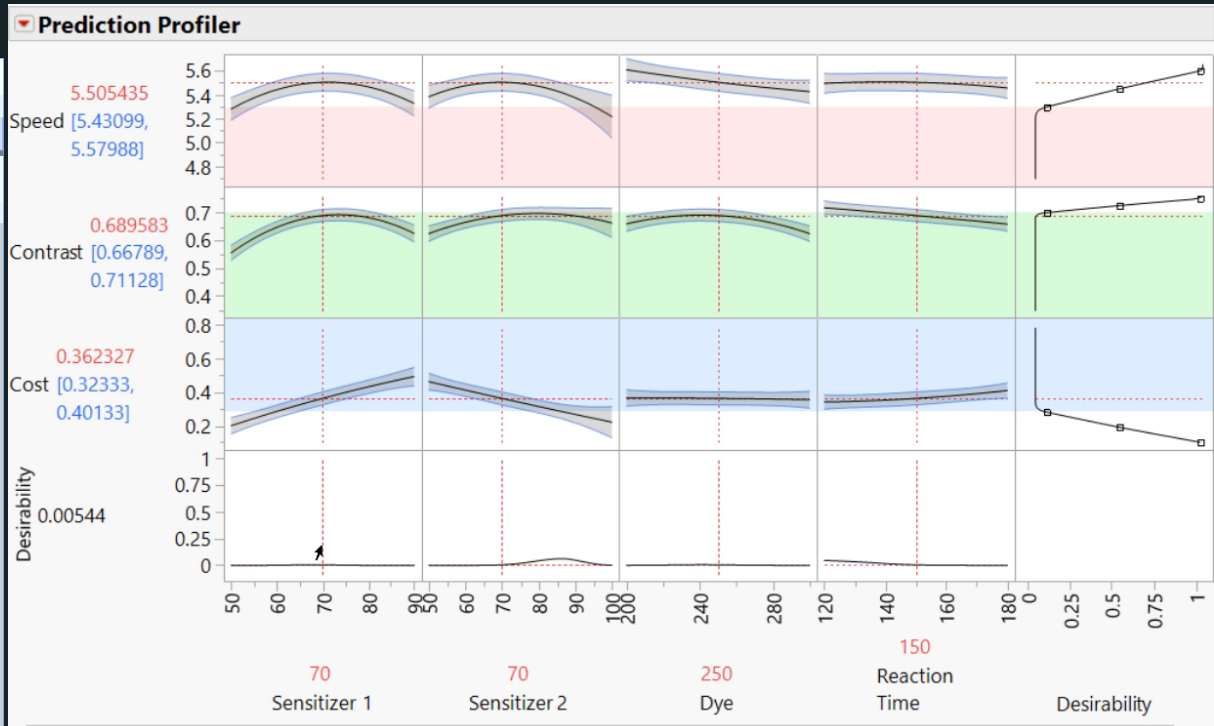
DERIVE MAXIMUM INFORMATION FROM FEWEST TESTS – YIELDING "INTERACTIVE" OPTIMIZATION AND TRADE-SPACE ANALYSIS

Photo_Cost27 - JMP Pro [3]

File Edit Tables Rows Cols DOE Analyze Graph Tools Add-Ins View Window Help

7/4 Cols
27/0

	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost
1	50	50	250	120	5.36	0.616	0.198
2	50	50	200	180	5.39	0.537	0.175
3	90	70	200	120	5.31	0.623	0.447
4	50	90	200	150	5.13	0.431	0.177
5	70	70	250	180	5.37	0.643	0.445
6	50	90	300	120	4.79	0.375	0.231
7	90	90	200	180	5.45	0.626	0.471
8	90	50	250	150	5.00	0.470	0.670
9	50	50	300	150	5.22	0.478	0.283
10	70	90	200	120	5.41	0.668	0.226
11	90	90	250	120	5.33	0.734	0.310
12	50	50	250	120	5.32	0.574	0.257
13	70	50	200	150	5.49	0.596	0.456
14	50	70	250	180	5.22	0.558	0.166
15	70	70	250	150	5.57	0.689	0.390
16	90	90	300	150	5.26	0.653	0.226
17	70	70	250	150	5.47	0.688	0.356
18	70	70	300	120	5.42	0.657	0.337
19	50	70	200	120	5.43	0.518	0.222
20	50	50	300	150	5.15	0.505	0.287
21	90	70	200	120	5.33	0.661	0.457
22	50	90	300	120	4.97	0.411	0.191
23	90	50	300	120	5.09	0.492	0.588
24	90	50	300	180	5.03	0.358	0.733
25	70	70	250	150	5.59	0.707	0.318
26	70	90	300	180	5.25	0.605	0.290
27	50	90	200	150	5.24	0.476	0.177



Remembered Settings

Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost	Desirability
Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014	0.347702
Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274	0.004875
Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539	0.214425
Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116	0.264298

Design of Experiments

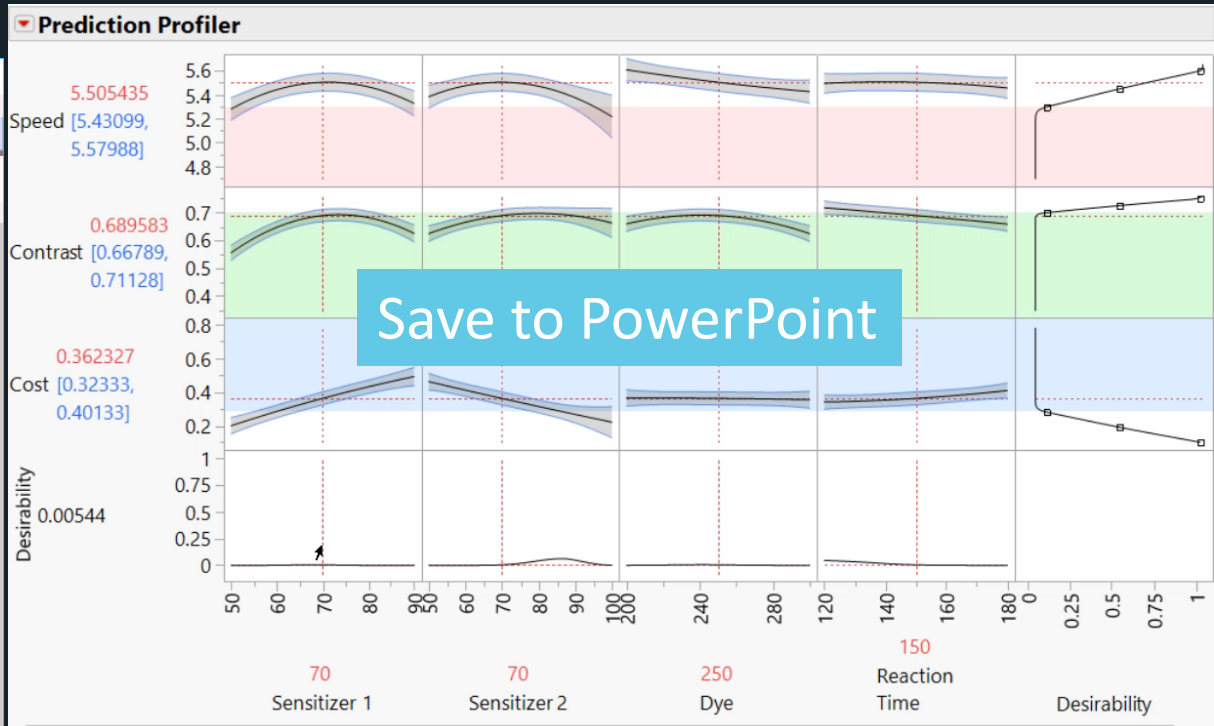
DERIVE MAXIMUM INFORMATION FROM FEWEST TESTS – YIELDING "INTERACTIVE" OPTIMIZATION AND TRADE-SPACE ANALYSIS

Photo_Cost27 - JMP Pro [3]

File Edit Tables Rows Cols DOE Analyze Graph Tools Add-Ins View Window Help

7/4 Cols
27/0

	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost
1	50	50	250	120	5.36	0.616	0.198
2	50	50	200	180	5.39	0.537	0.175
3	90	70	200	120	5.31	0.623	0.447
4	50	90	200	150	5.13	0.431	0.177
5	70	70	250	180	5.37	0.643	0.445
6	50	90	300	120	4.79	0.375	0.231
7	90	90	200	180	5.45	0.626	0.471
8	90	50	250	150	5.00	0.470	0.670
9	50	50	300	150	5.22	0.478	0.283
10	70	90	200	120	5.41	0.668	0.226
11	90	90	250	120	5.33	0.734	0.310
12	50	50	250	120	5.32	0.574	0.257
13	70	50	200	150	5.49	0.596	0.456
14	50	70	250	180	5.22	0.558	0.166
15	70	70	250	150	5.57	0.689	0.390
16	90	90	300	150	5.26	0.653	0.226
17	70	70	250	150	5.47	0.688	0.356
18	70	70	300	120	5.42	0.657	0.337
19	50	70	200	120	5.43	0.518	0.222
20	50	50	300	150	5.15	0.505	0.287
21	90	70	200	120	5.33	0.661	0.457
22	50	90	300	120	4.97	0.411	0.191
23	90	50	300	120	5.09	0.492	0.588
24	90	50	300	180	5.03	0.358	0.733
25	70	70	250	150	5.59	0.707	0.318
26	70	90	300	180	5.25	0.605	0.290
27	50	90	200	150	5.24	0.476	0.177

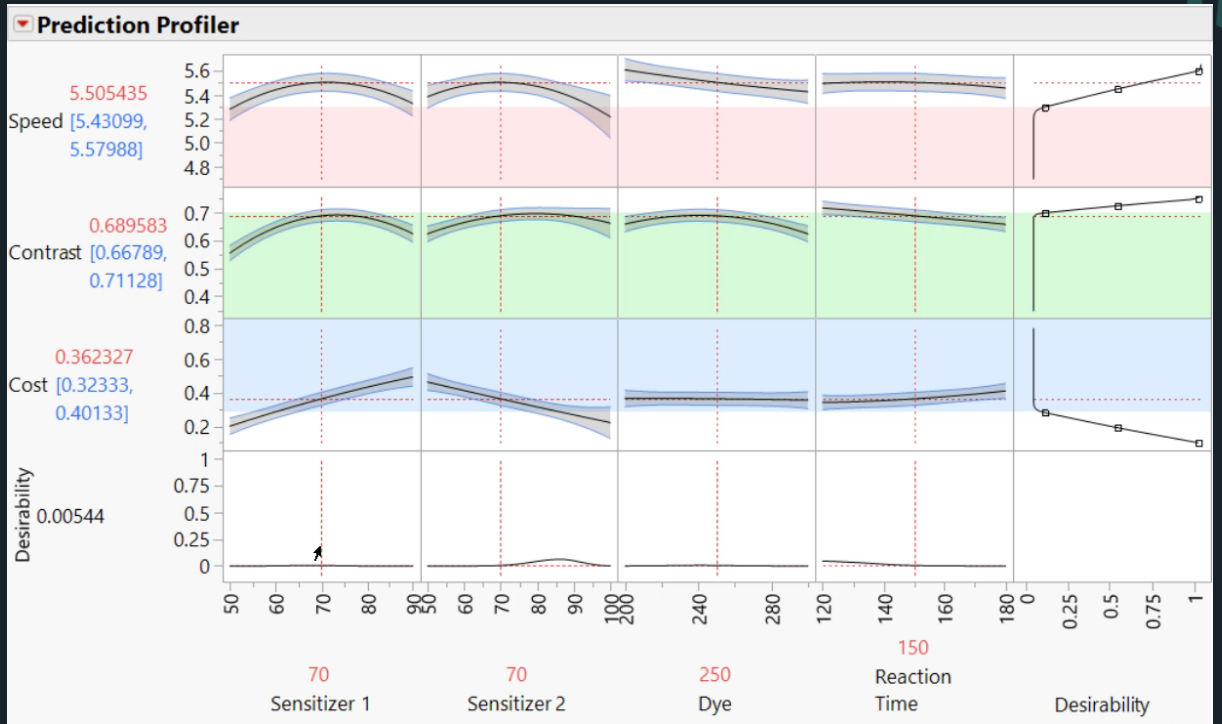


Remembered Settings

Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost	Desirability
Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014	0.347702
Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274	0.004875
Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539	0.214425
Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116	0.264298

Informing Decision Discussions

OPTIMIZATION AND TRADE-SPACE ANALYSIS



Remembered Settings

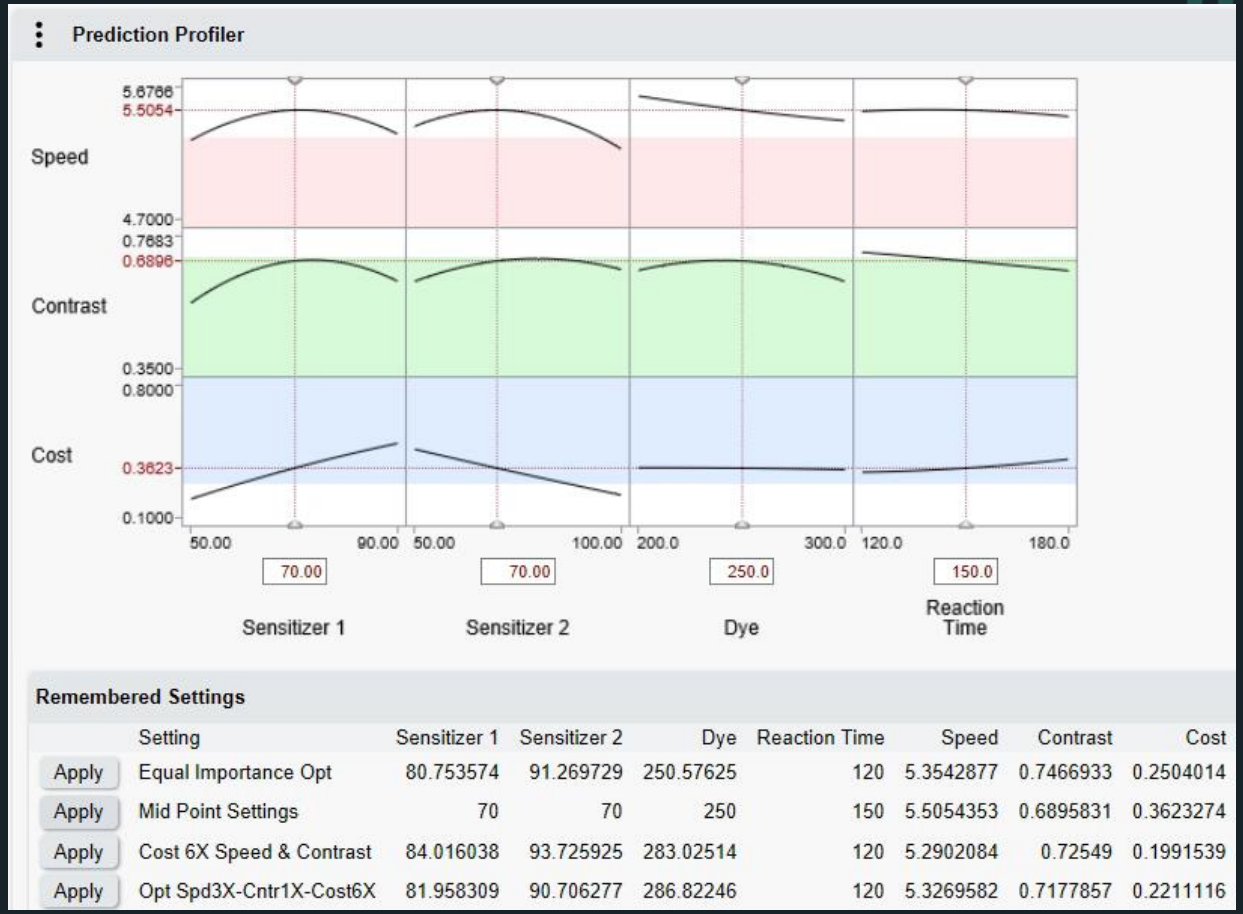
	Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost	Desirability
○	Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014	0.347702
○	Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274	0.004875
○	Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539	0.214425
○	Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116	0.264298

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Share Analysis via HTML

View optimizations on your phone.
Scan the QR code to launch browser, then use finger to interact with the Prediction Profiler and to “Apply” saved settings.



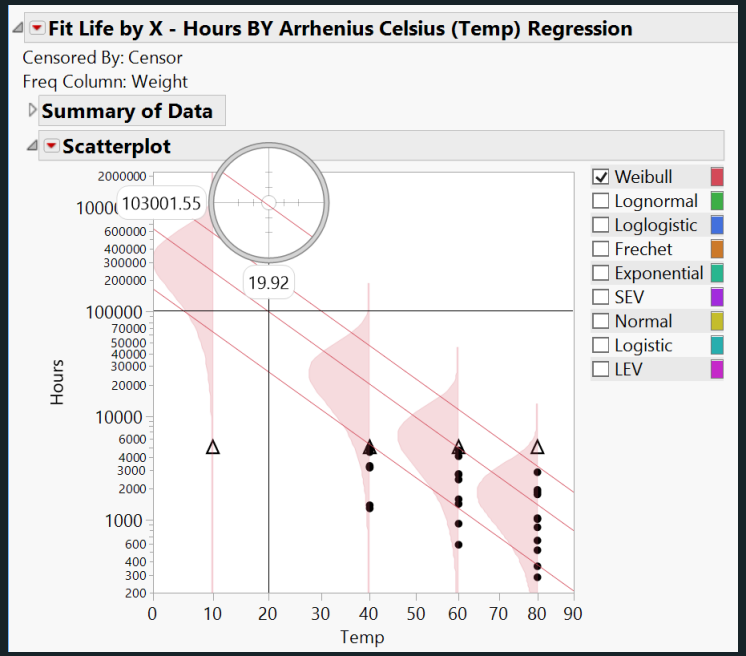
Remembered Settings

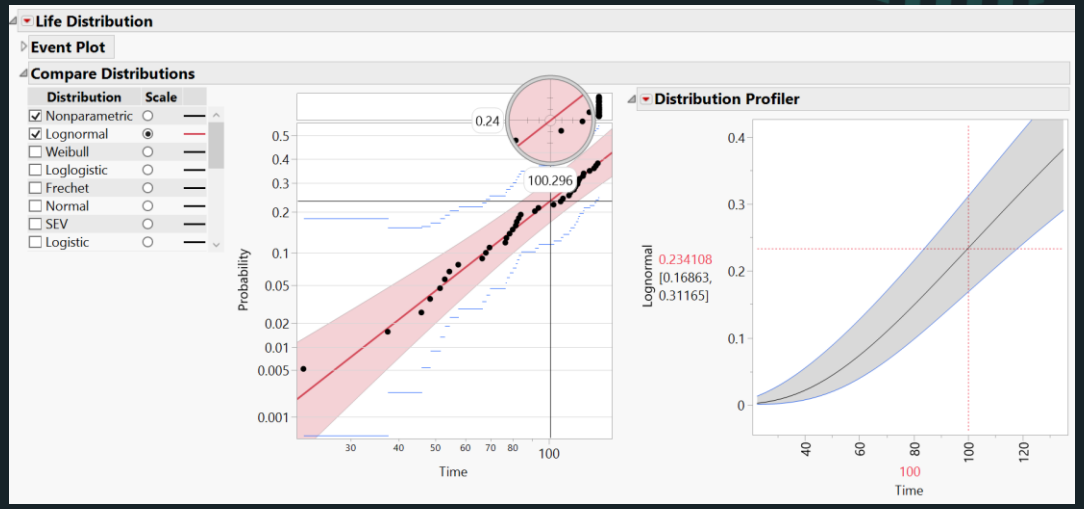
	Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost
Apply	Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014
Apply	Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274
Apply	Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539
Apply	Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116

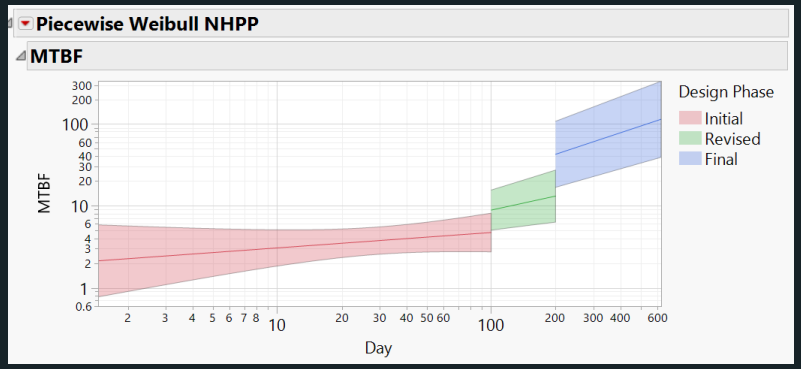
Copyright © JMP Statistical Discovery LLC. All rights reserved.

Reliability

PROBABILITY OF FAILURE, ACCELERATED LIFE TEST, RELIABILITY GROWTH





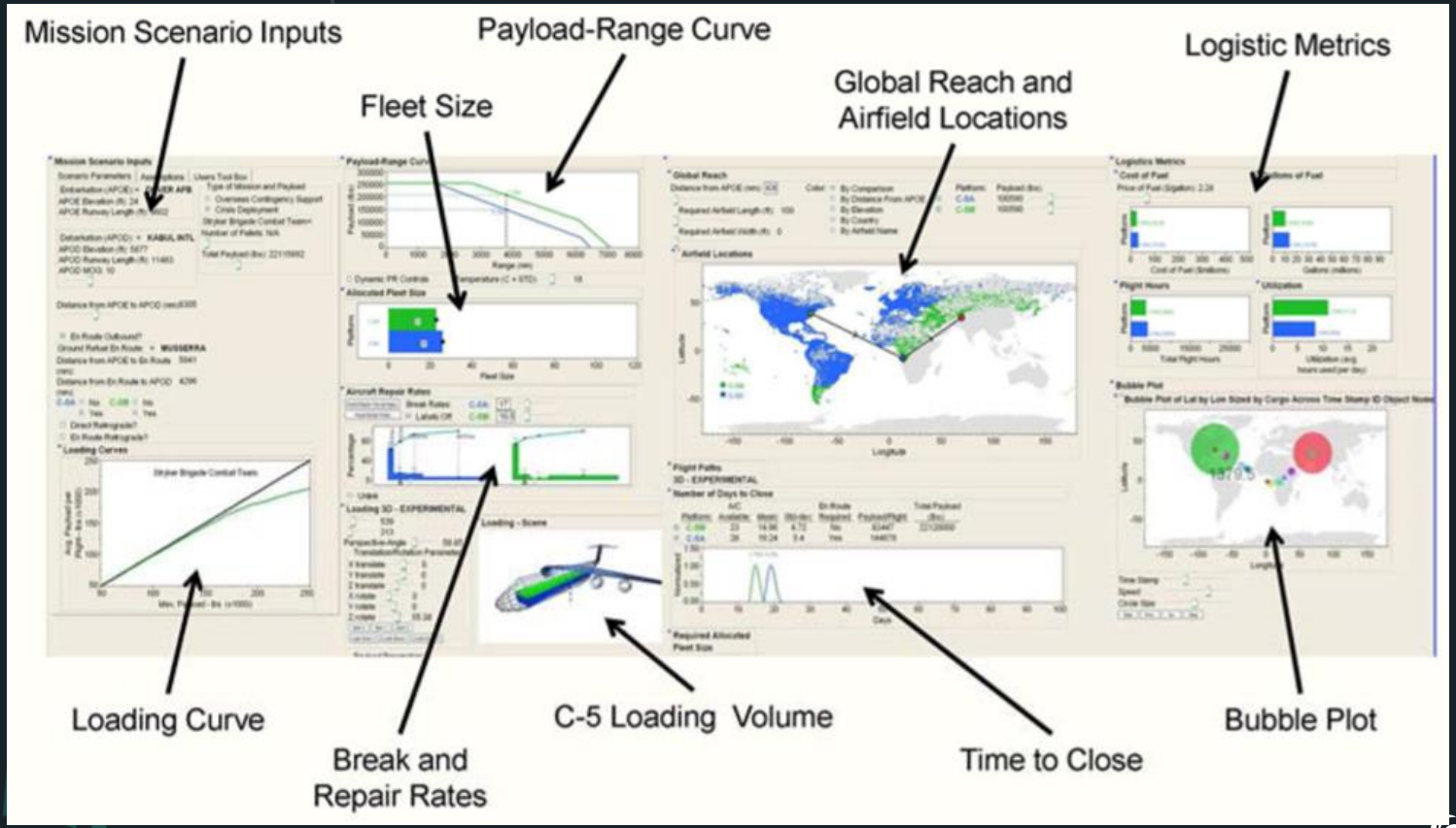


Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Fast Surrogate Modeling

BY MODELING THE "RIGHT NUMBER" OF LONG RUNNING SIMULATIONS (THINK SEQUENTIAL DOE) , AN ACCURATE & INSTANTANEOUS PREDICTION CAN BE MADE FOR ANY NEW SCENARIO!



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Download & Recording

- 16 Factors
- 50,000 unique cases
- Each 1,000 times
- 50 Million Simulations
- Neural Network for Surrogate Models

1.6 Comparative Assessment and Decision Support System for Strategic Military Airlift Capability

Comparative Assessment and Decision Support System for Strategic Military Airlift Capability

John Salmon, Curtis Iwata, Dimitri Mavris and Neil Weston
Georgia Institute of Technology
john.salmon@asdl.gatech.edu, curtis.iwata@asdl.gatech.edu, dimitri.mavris@aerospace.gatech.edu, neil.weston@ae.gatech.edu
Philip Fahringer
Lockheed Martin Company
philip.fahringer@lmco.com

ABSTRACT

The Lockheed Martin Aeronautics Company has been awarded several programs to modernize the aging C-5 military transport fleet. In order to ensure its continuation amidst budget cuts, it was important to engage the decision makers by providing an environment to analyze the benefits of the modernization program. This paper describes an interface that allows the user to change inputs such as the scenario airfields, take-off conditions, and reliability characteristics. The underlying logistics surrogate model was generated using data from a discrete-event simulation. Various visualizations, such as intercontinental flight paths illustrated in 3D, have been created to aid the user in analyzing scenarios and performing comparative assessments for various output logistics metrics. The capability to rapidly and dynamically evaluate and compare scenarios was developed enabling real-time strategy exploration and trade-offs.

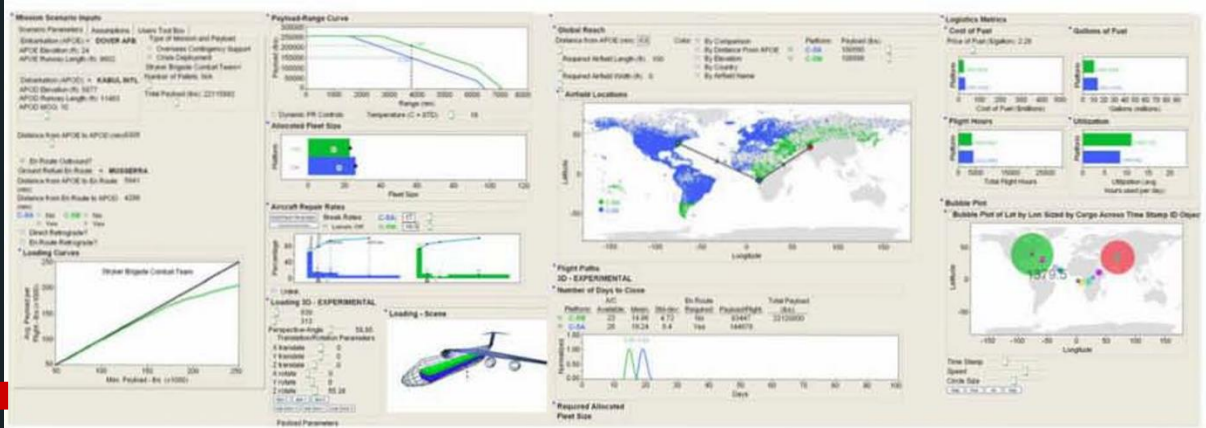


Figure 2. Strategic Airlift Comparison Tool Layout

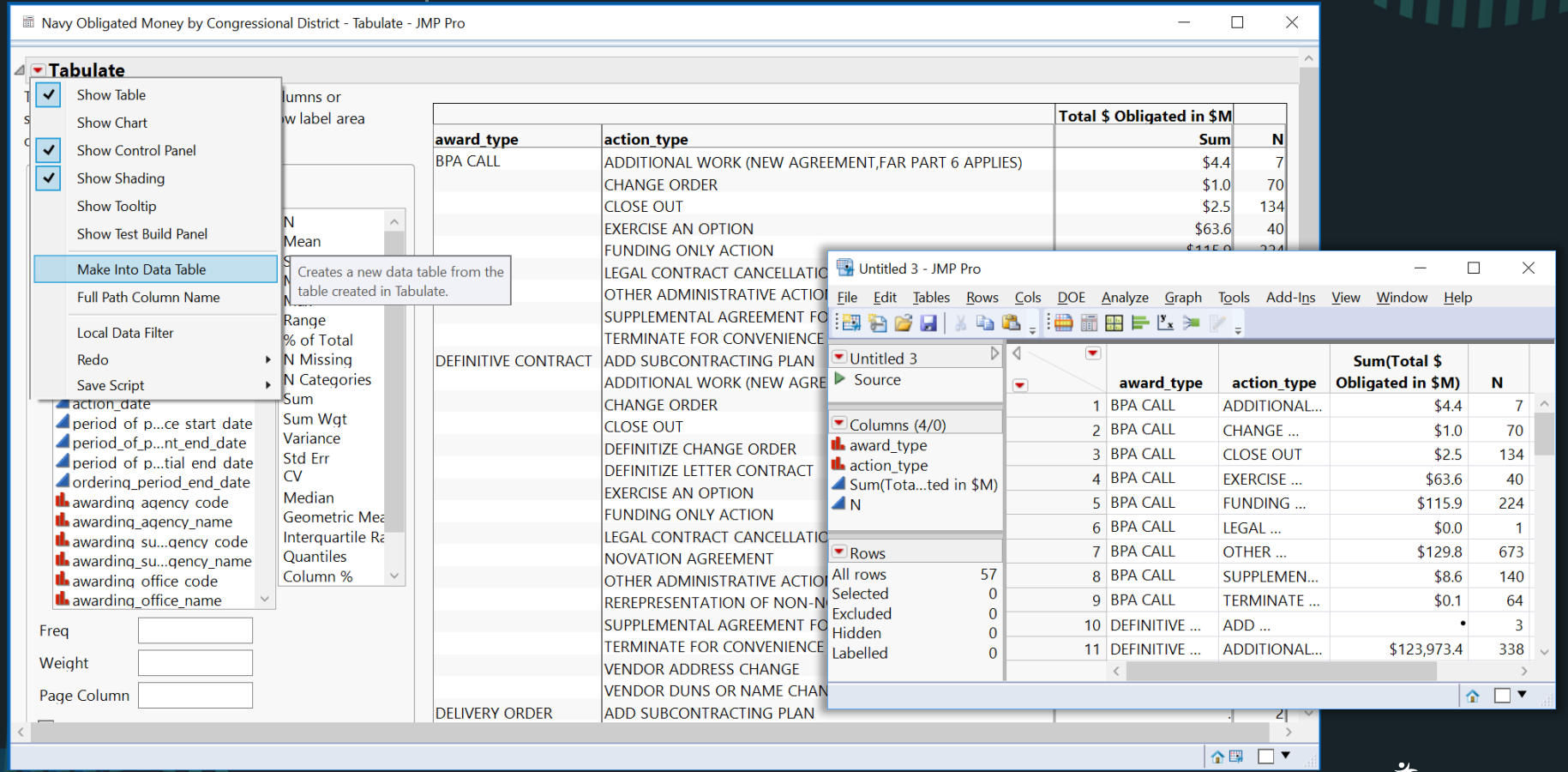
Download Document
https://ntrs.nasa.gov/search.jsp?R=20110012110

Recording

https://community.jmp.com/t5/US-Federal-Government-JMP-Users/VIDEO-of-C-5-Decision-Support-Tool-Dashboard/ta-p/34364

TABULATE

JMP'S DRAG & DROP PIVOT TABLE – EASILY MAKE INTO NEW DATA TABLE



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Graph Builder – Visual Pivot Table

CONDITIONAL DATA FILTERING "ACTION TYPE" WITHIN "AWARD TYPE"

Local Data Filter

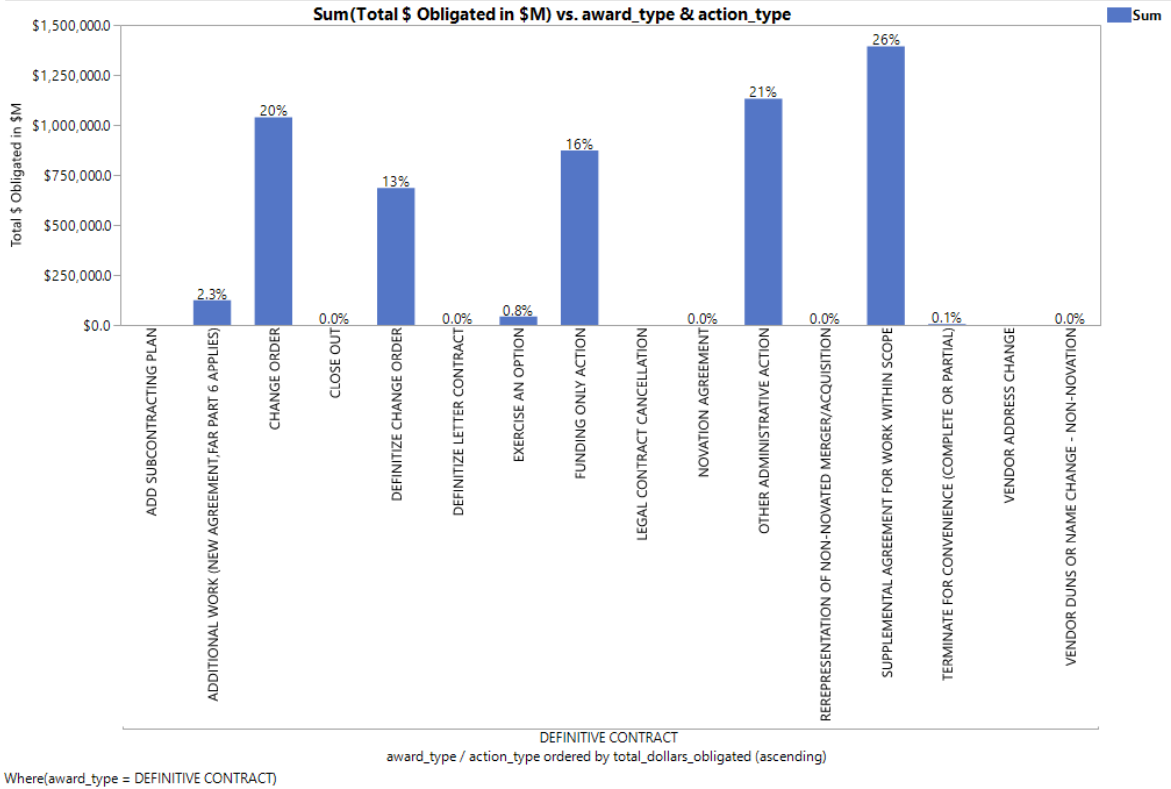
24898 matching rows

Inverse

[1] award_type (4)
- BPA CALL (6652)
- DEFINITIVE CONTRACT (24898)
- DELIVERY ORDER (111867)
- PURCHASE ORDER (53115)

action_type (17)
- ??? (2561)
- ADD SUBCONTRACTING PLAN (3)
- ADDITIONAL WORK (NEW AGREEMENT,FAR PART 6 APPLIES) (338)
- CHANGE ORDER (1580)
- CLOSE OUT (183)
- DEFINITIZE CHANGE ORDER (419)
- DEFINITIZE LETTER CONTRACT (9)
- EXERCISE AN OPTION (2081)
- FUNDING ONLY ACTION (5602)
- LEGAL CONTRACT CANCELLATION (1)
- NOVATION AGREEMENT (5)
- OTHER ADMINISTRATIVE ACTION (6485)
- REREPRESENTATION OF NON-NOVATED MERGER/ACQUISITION (3)
- SUPPLEMENTAL AGREEMENT FOR WORK WITHIN SCOPE (5566)
- TERMINATE FOR CONVENIENCE (COMPLETE OR PARTIAL) (55)
- VENDOR ADDRESS CHANGE (1)
- VENDOR DUNS OR NAME CHANGE - NON-NOVATION (6)

Graph Builder



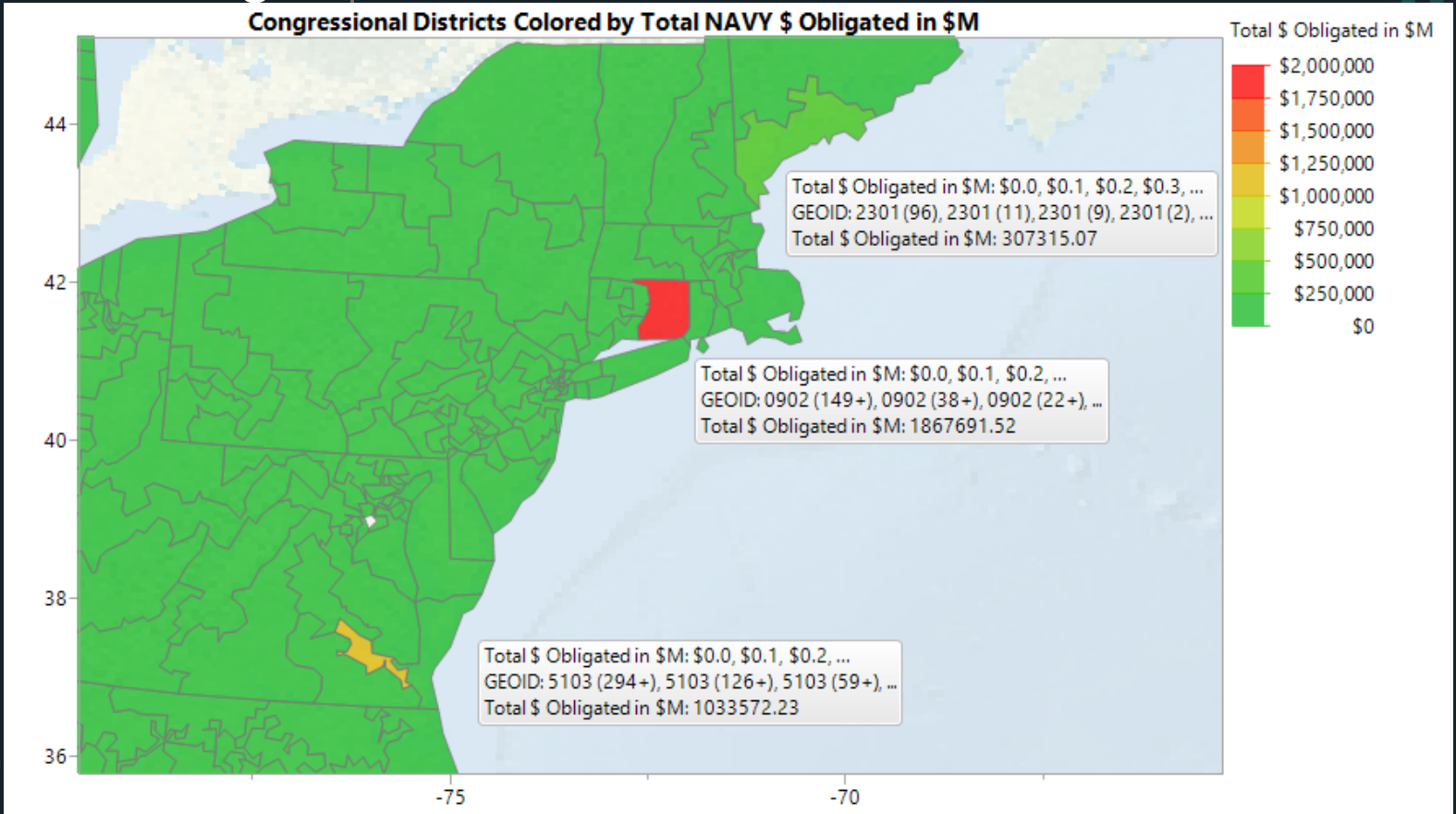
Where(award_type = DEFINITIVE CONTRACT)

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Maps and images

NAVY OBLIGATED SPENDING BY CONGRESSIONAL DISTRICT IN $M

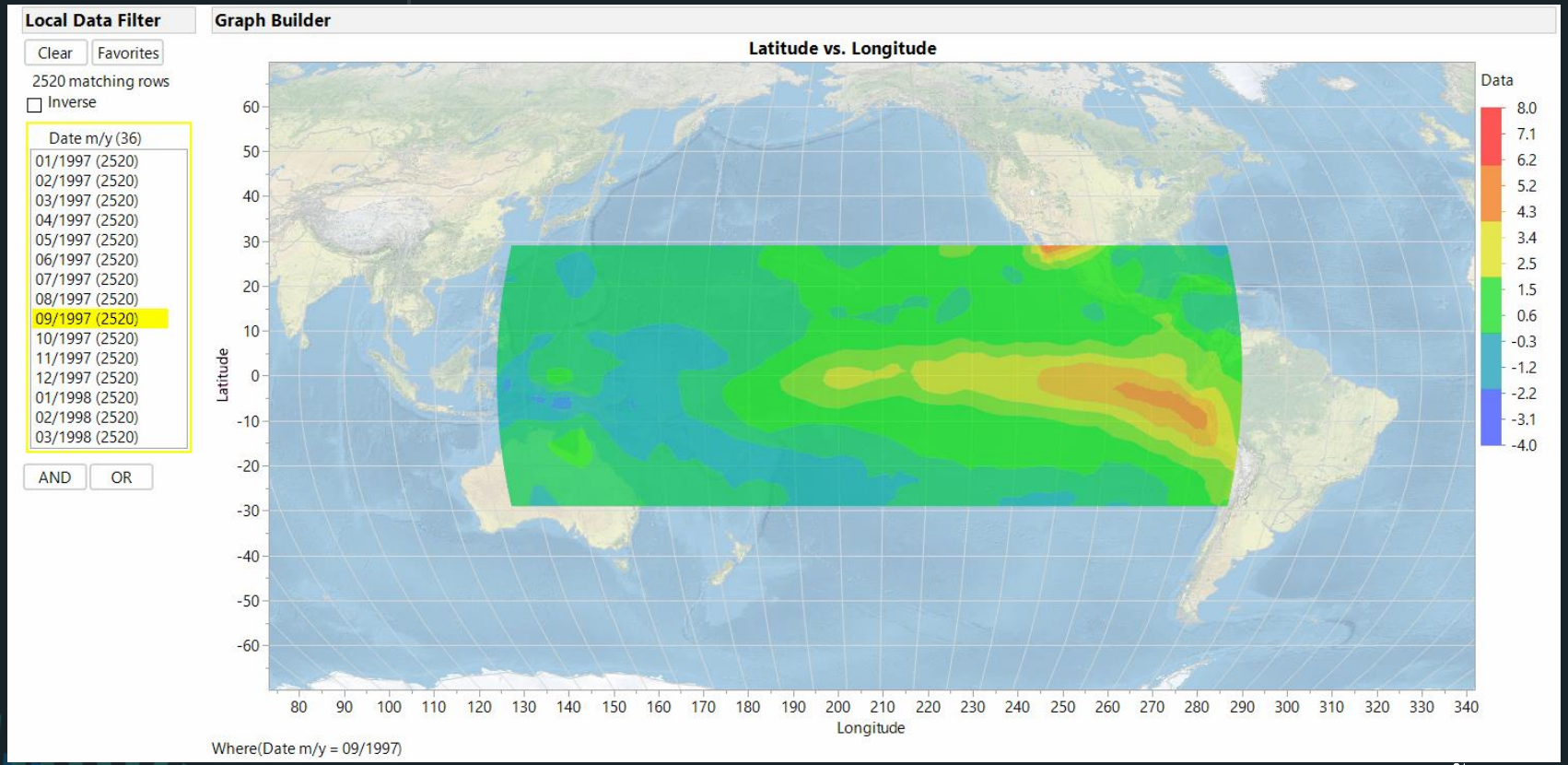


Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Maps and images

- **MAPS ANIMATED OVER TIME AND SAVED AS A .GIF CAN BE DROPPED INTO PPT**



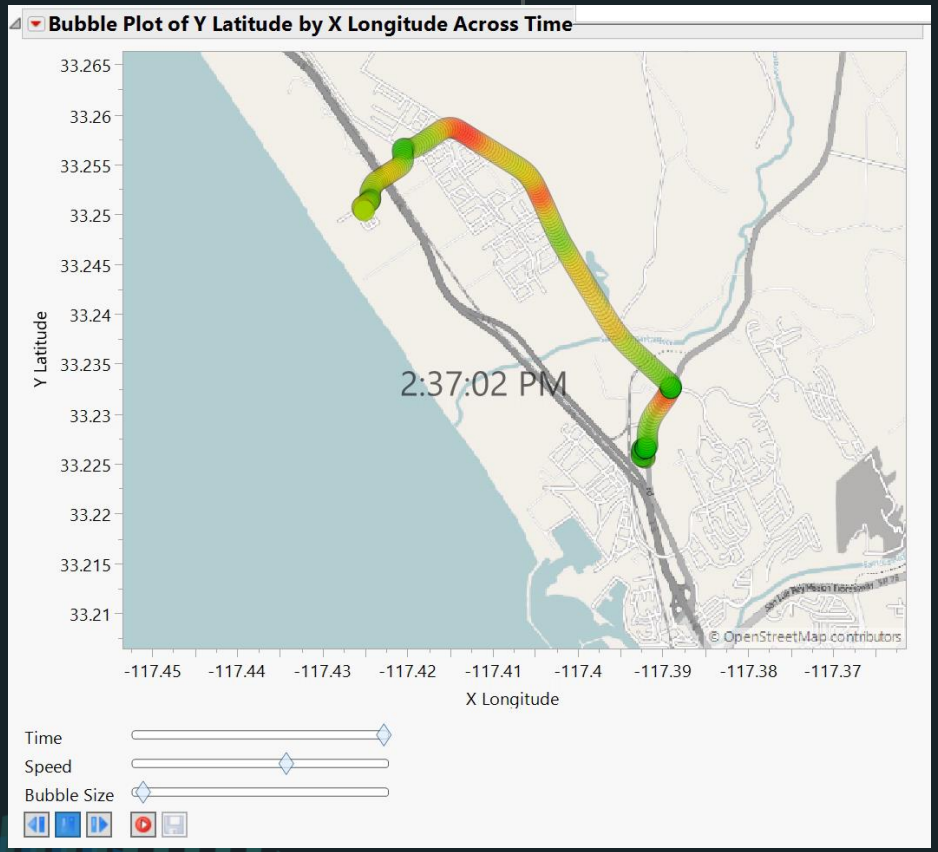
Copyright © JMP Statistical Discovery LLC. All rights reserved.

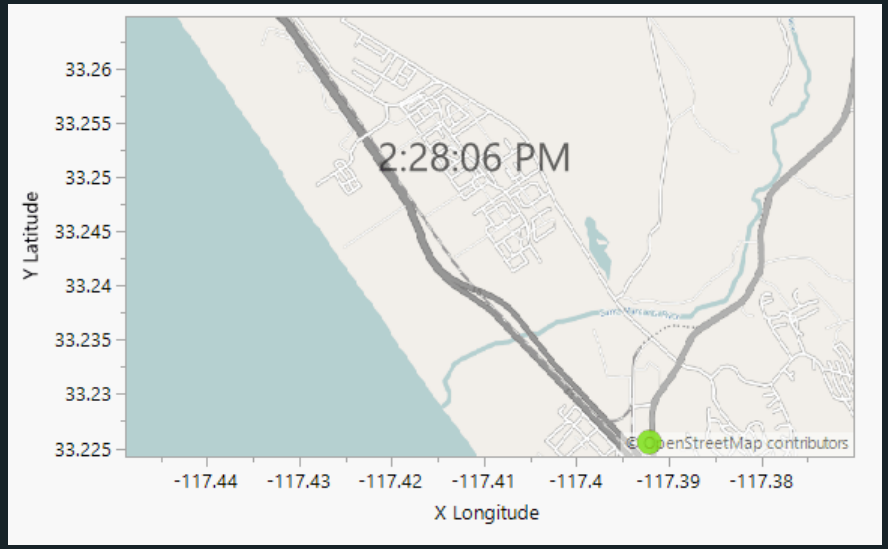
DROP GIF HERE

Copyright © JMP Statistical Discovery LLC. All rights reserved.

Maps and images

MAPS ANIMATED OVER TIME: SAVED AS PPT VIDEO (LEFT) ANIMATED GIF (RIGHT) OR HTML (USE QR CODE TO VIEW AT PUBLIC.JMP.COM ON YOUR PHONE)





Copyright © JMP Statistical Discovery LLC. All rights reserved.

Map of Printing Plate

HOVER TO PULL UP IMAGE OF DEFECT OR PLOT OF SENSOR OVER TIME

Note random data line chart - Graph Builder - JMP Pro

Graph Builder

Undo | Start Over | Done

Variables

5 Columns

- Note Location
- Random Data
- Images
- Defect Type
- Line Chart

Map Shapes

Note Location colored by Random Data

Group X

100
80
0
25
50

Defect Type: Misalignment Cut
Mean(Random Data): 0.62
Line Chart:

State = MA

Avg Retail Elec Price

16
14
12
10
8

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013

Year Char / Month Char

Defect Type: Ink Smear
Mean(Random Data): 0.60
Images:

evaluations done

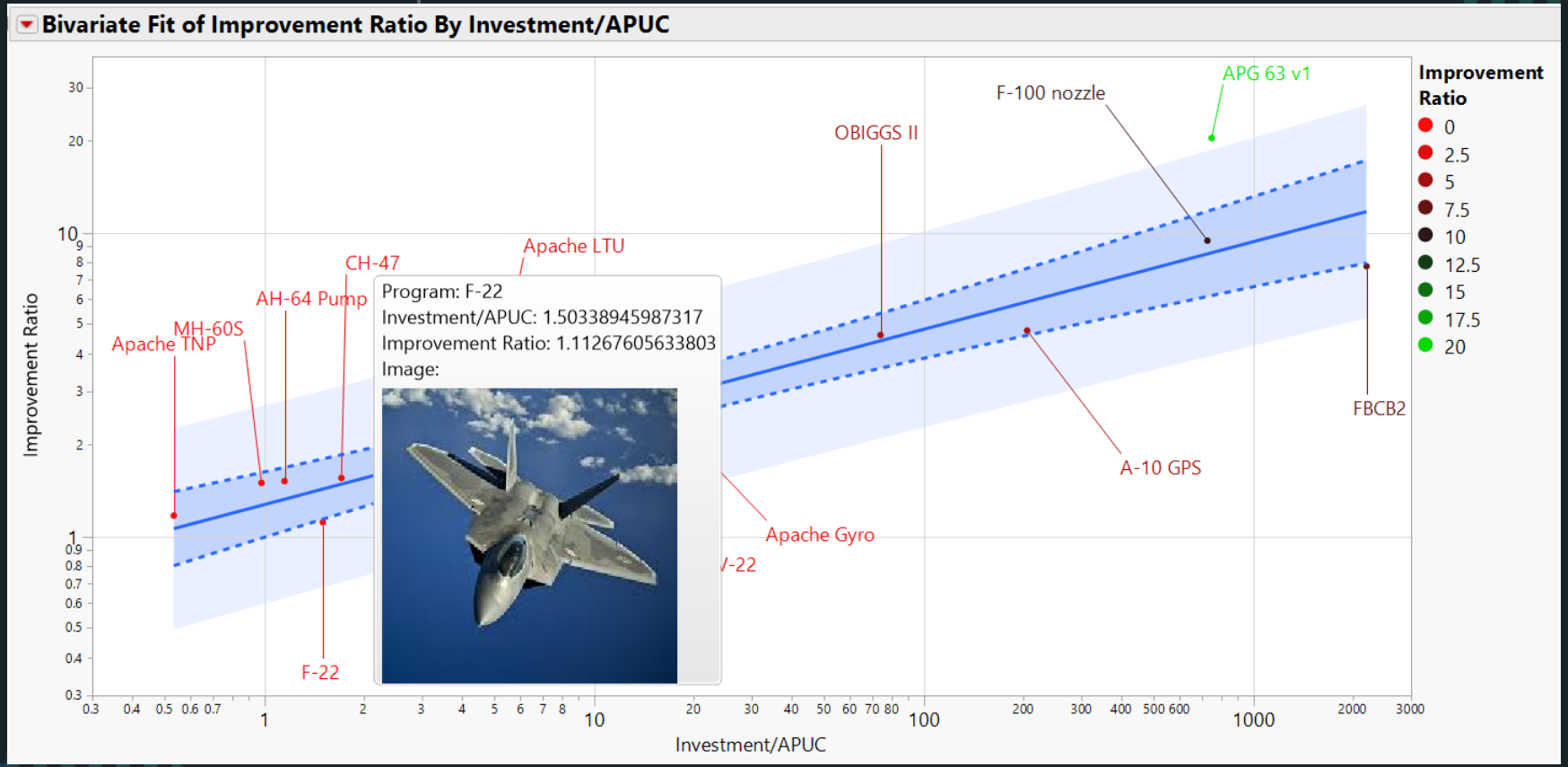
4:42 PM

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Maps and images

IMAGES CAN BE IMPORTED IN BULK AND SHOWN BY HOVERING ON A DATA POINT

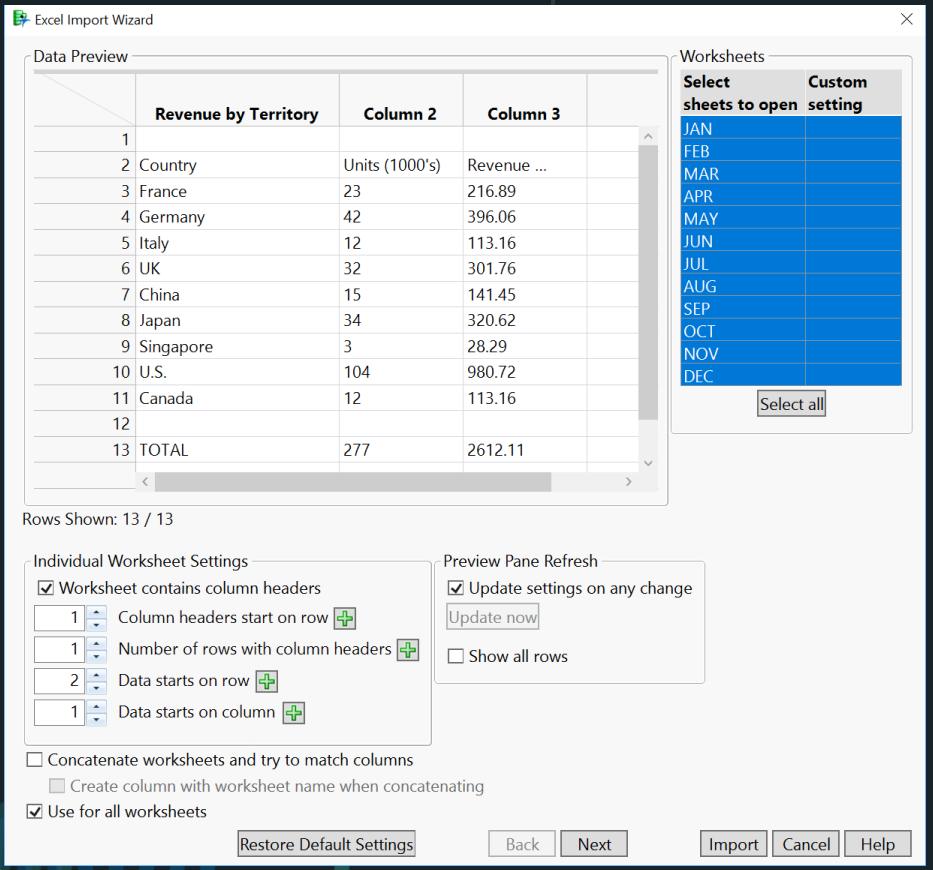


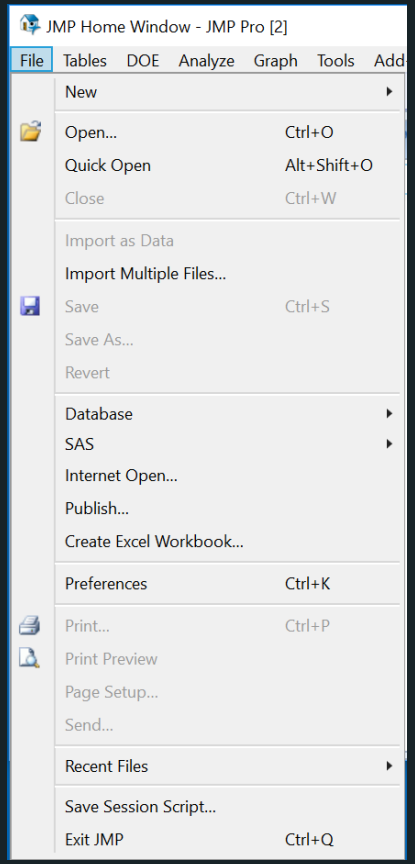
Copyright © JMP Statistical Discovery LLC. All rights reserved.

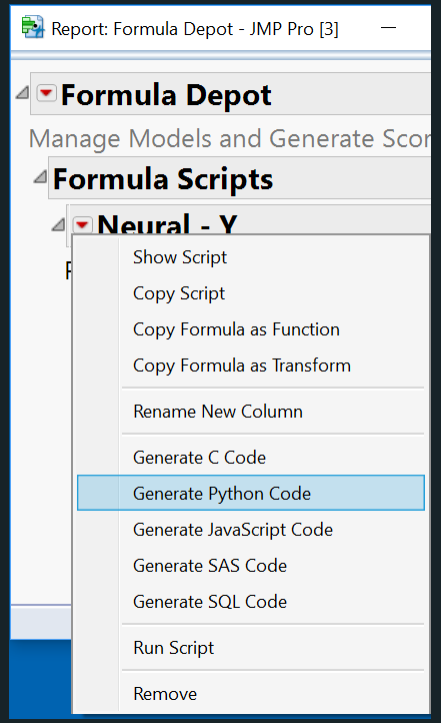
JMP STATISTICAL DISCOVERY

JMP as Data HUB

EASILY IMPORT DATA FROM EXCEL, TXT, CSV, PDF, SAS, DATA BASES, WEB, &
EASILY EXPORT MODELS (WITH 13 & 14 PRO) TO: SAS, PYTHON, JAVASCRIPT, C, SQL
EASILY PUBLISH REPORTS TO WEB (HTML); SAVE REPORTS TO POWERPOINT







Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

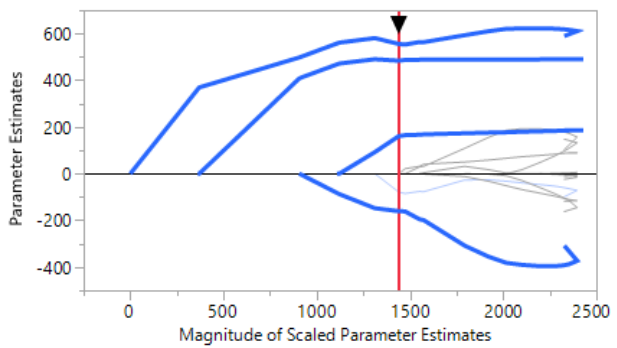
Fast modern regression (JMP 12)

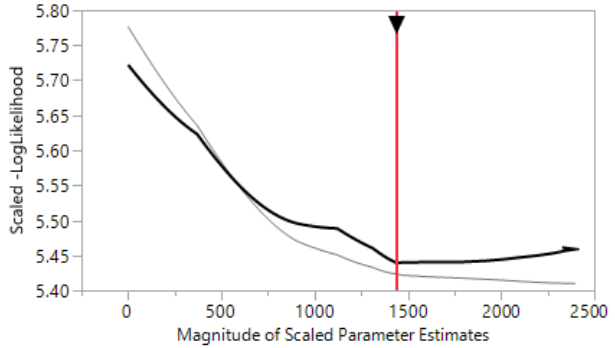
INTERACT WITH ADAPTIVE LASSO MODEL – VARIABLE & MODEL SELECTION FOR DATA THAT'S HIGHLY CORRELATED AND/OR NON-NORMAL

Measure	Training	Validation	Test
Number of rows	294	74	74
Sum of Frequencies	294	74	74
-LogLikelihood	1594.5348	402.5022	399.46519
Number of Parameters	7	7	7
BIC	3228.8547	835.13285	829.05883
AICc	3203.4613	820.70136	814.62735
RSquare	0.506121	0.4320668	0.5061904
RMSE	54.852588	55.703388	53.440629
Lambda Penalty	6030.1572	.	.

Near Machine Learning accuracy, AND provides a more interpretable model with confidence intervals

Solution Path



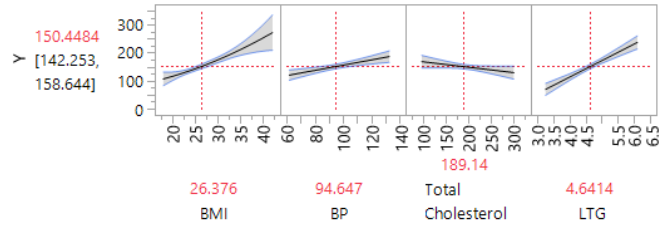


Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Intercept	-302.705	39.488631	58.761791	<.0001*	-380.1013	-225.3087
BMI	6.3716072	0.8995747	50.167619	<.0001*	4.6084731	8.1347413
BP	0.6863572	0.2729626	6.3225768	0.0119*	0.1513604	1.2213541
Total Cholesterol	-0.253761	0.1183203	4.5997185	0.0320*	-0.485664	-0.021857
HDL	-0.355974	0.28592	1.5500595	0.2131	-0.916367	0.2044185
LTG	61.913643	8.8560351	48.87582	<.0001*	44.556133	79.271153

(BP-94.647)*(Total Cholesterol-
(Total Cholesterol-189.14)*(LDL
(Total Cholesterol-189.14)*(HDL
(Total Cholesterol-189.14)*(TCH
(HDL-49.7885)*(TCH-4.07025)
(Total Cholesterol-189.14)*(LTG
LDL
TCH

Prediction Profiler



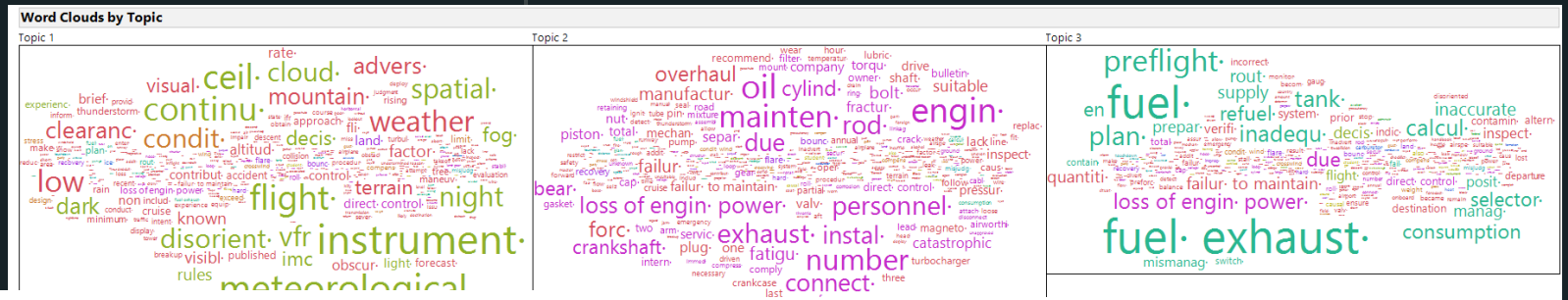
Copyright © JM

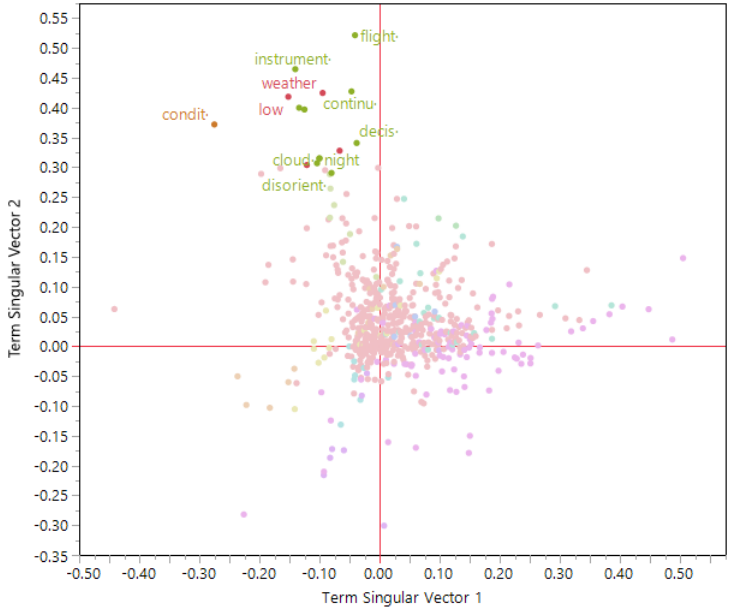
JMP® PRO

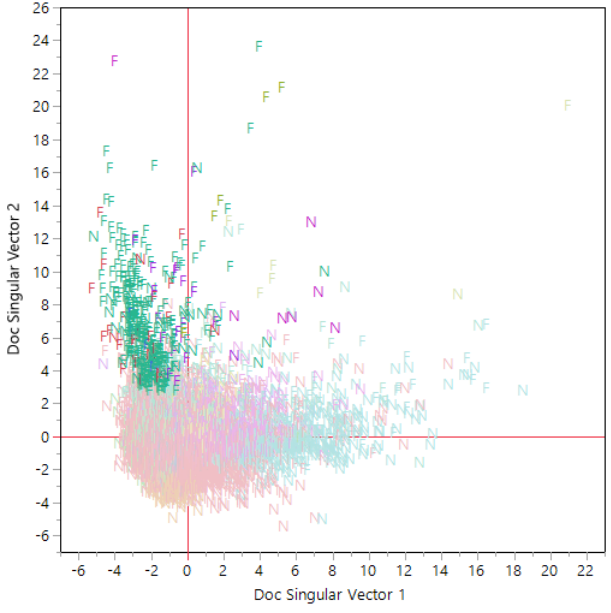
jmp STATISTICAL DISCOVERY

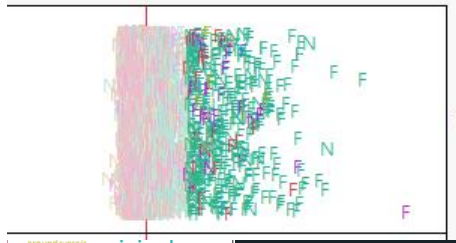
Exploratory Text Analysis (JMP 13)

DIMENSION REDUCTION OF SPARSE DOCUMENT TERM MATRIX INTO DOCUMENT AND TERM VECTORS – ALSO CLUSTERING OF DOCUMENTS AND TOPICS









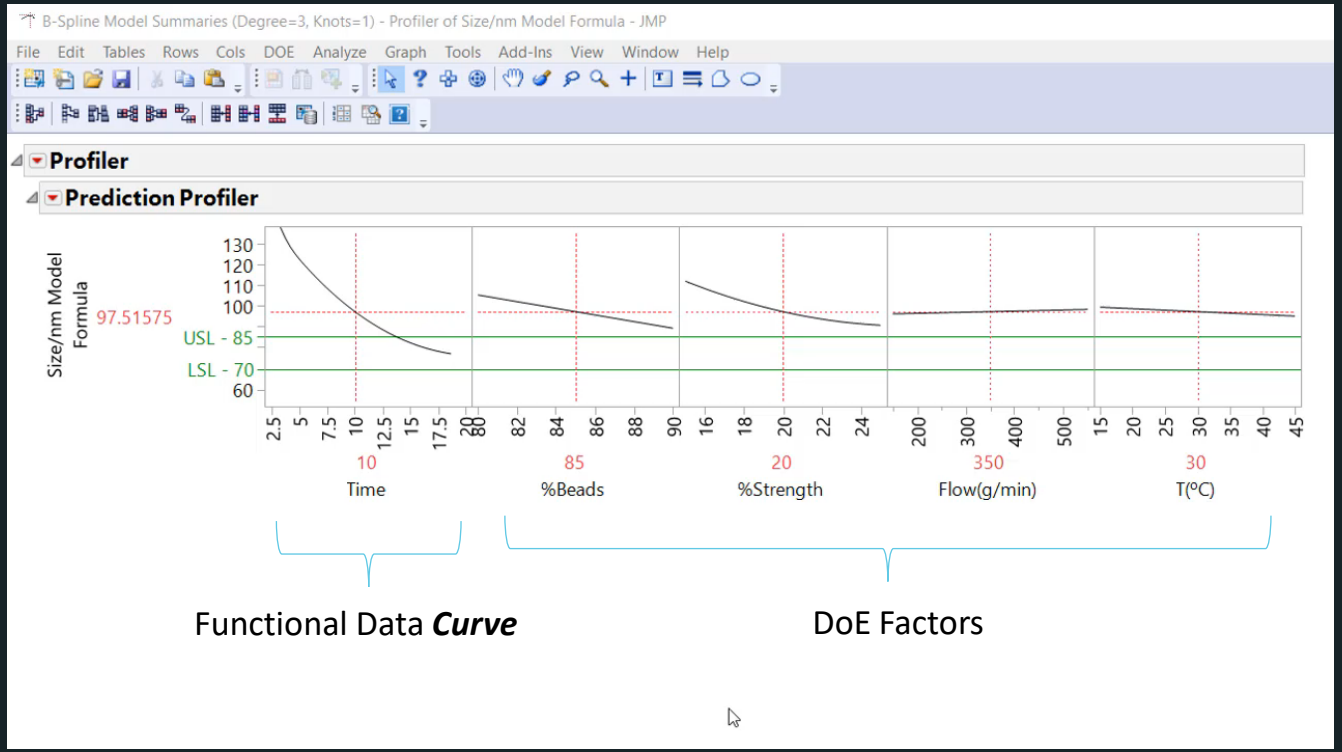
Topic 1

JMP® PRO

jmp STATISTICAL DISCOVERY

Functional Data Analysis (JMP 14)

MODELING THE "SHAPE" OF A STREAM OF DATA – SHAPE IS THE FUNDAMENTAL UNIT OF OBSERVATION – DIMENSION REDUCTION WITH FUNCTIONAL PCA ABLE TO CONTROL AND PREDICT SHAPE AS FUNCTION OF DOE FACTORS



Copyright © JMP Statistical Discovery LLC. All rights reserved.

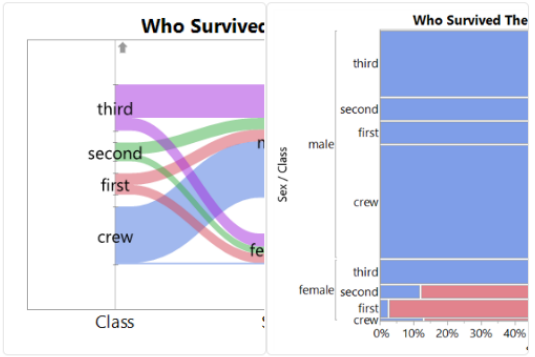
JMP Public & Live

HTTPS://PUBLIC.JMP.COM AND HTTPS://DEVLIVE16.JMP.COM USE TO SHARE RESULTS *DYNAMICALLY* WITH NON-JMP USERS.

https://public.jmp.com/featured

Home | Featured | What's New | Search | Tom Donnelly

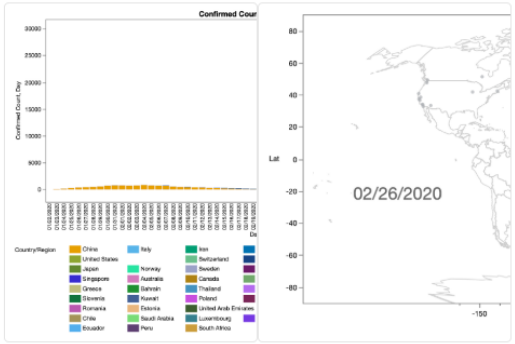
Who Survived the Titanic?
Phil Kay



Interactive plots visualise Titanic survivor data.

Views: 87 — Wed Mar 06, 2019 4:49 am

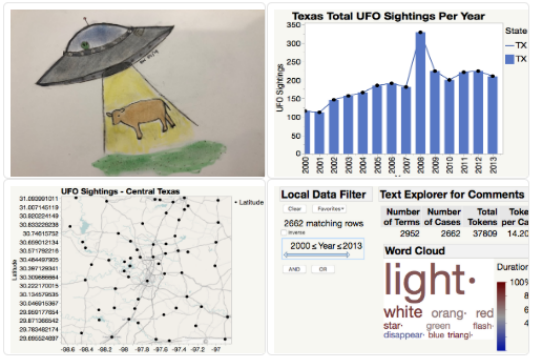
Coronavirus COVID-19 Global Cases
JMP_Byron



Time series data represented in multiple graph formats

1379 0 — Sat Mar 21, 2020 2:54 pm

Texas UFO Sightings
Scott Wise



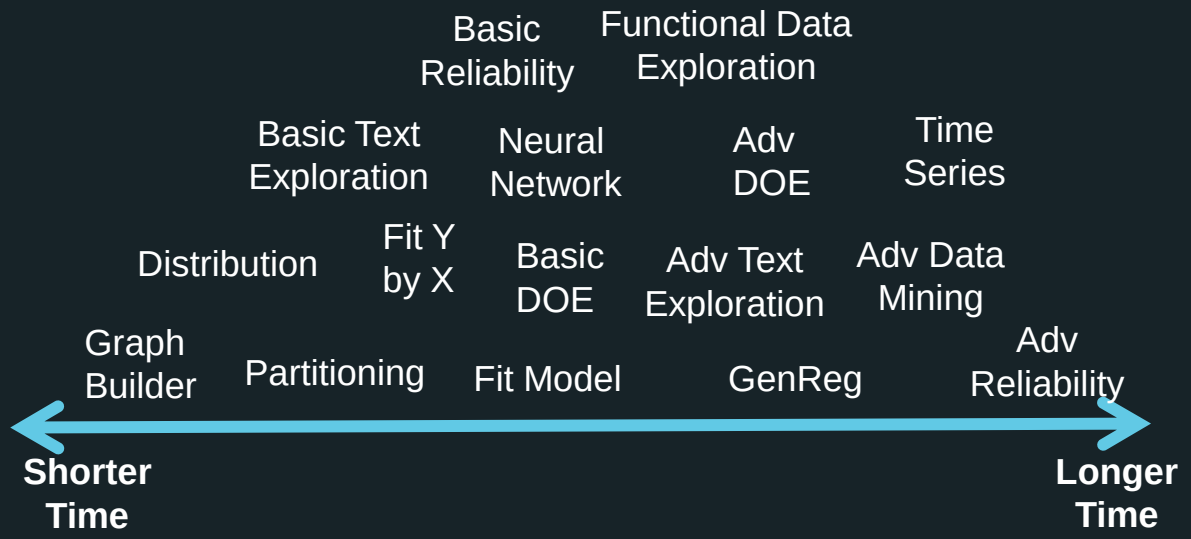
I visualized the number and location of UFO sightings in Texas from 2000 to 2014. I also explored the text of comments in sighting reports.

Views: 65 — Mon Mar 04, 2019 9:57 pm

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

How long does it take to learn JMP?
Depends… on your level of statistical know how



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Summary: JMP capabilities that can help you

Move from data to decision – *Faster!*

- Exploratory Data Analysis with Dynamic Linking & Filtering
- Design of Experiments – Real & Simulated Data
- Reliability Analysis, Accelerated Life Test, Reliability Growth
- Surrogate modeling of simulations
- Visual Multivariate Modeling
- Robust Data Mining using Honest Assessment stopping criteria when using Decision Trees, Neural Networks, Regression, Non-Linear modeling
- Maps and Images – animation over time
- Data Curation – recoding, impute missing, find outliers, import, export
- Fast and Modern Penalized Regression – interact with model (12 Pro)
- Exploratory Text Analysis – dimension reduction – Topic Vectors (13 Pro)
- Functional Data Analysis – modeling the “shape” of a data stream (14 Pro)

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Resources

My Recorded Tutorials & Slide Decks at www.jmp.com/fedgov

These 9 videos cover predicitve analytics (including text exploration), data visualization, and "What's New in JMP 14?" topics.

Building Better Models Overview and Use of Honest Assessment	Neural Networks - Single Layer, Dual Layer, Boosted	All Graphs are Wrong - Some are Useful - Or view Xan Gregg's Original 2015 Discovery Summit Presentation
Regression Linear, Stepwise, Logistic, & All Possible	Generalized Regression Near Machine Learning Accuracy – More Explainable Model	What's New in JMP 14? JMP Learning Resources
Decision Trees Simple Partition, Bootstrap Forest, & Boosted Tree	Text Exploration Analyze Unstructured Free Text	Functional Data Explorer Modeling a "Stream" of Data – New in JMP 14

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Questions or comments?

Thank you.

jmp.com

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP Defense & Aerospace Team

Anna-Christina De La Iglesia
Program Manager
anna-christina.delaiglesia@jmp.com
919-531-2593

Procurement, Upgrades, License Renewals…

Sam Tobin
Senior Account Representative
sam.tobin@jmp.com
919-531-0640

Technical Questions, Getting Started, Tutorials, Mentoring…

Tom Donnelly, PhD, CAP
Principal Systems Engineer
tom.donnelly@jmp.com
302-489-9291

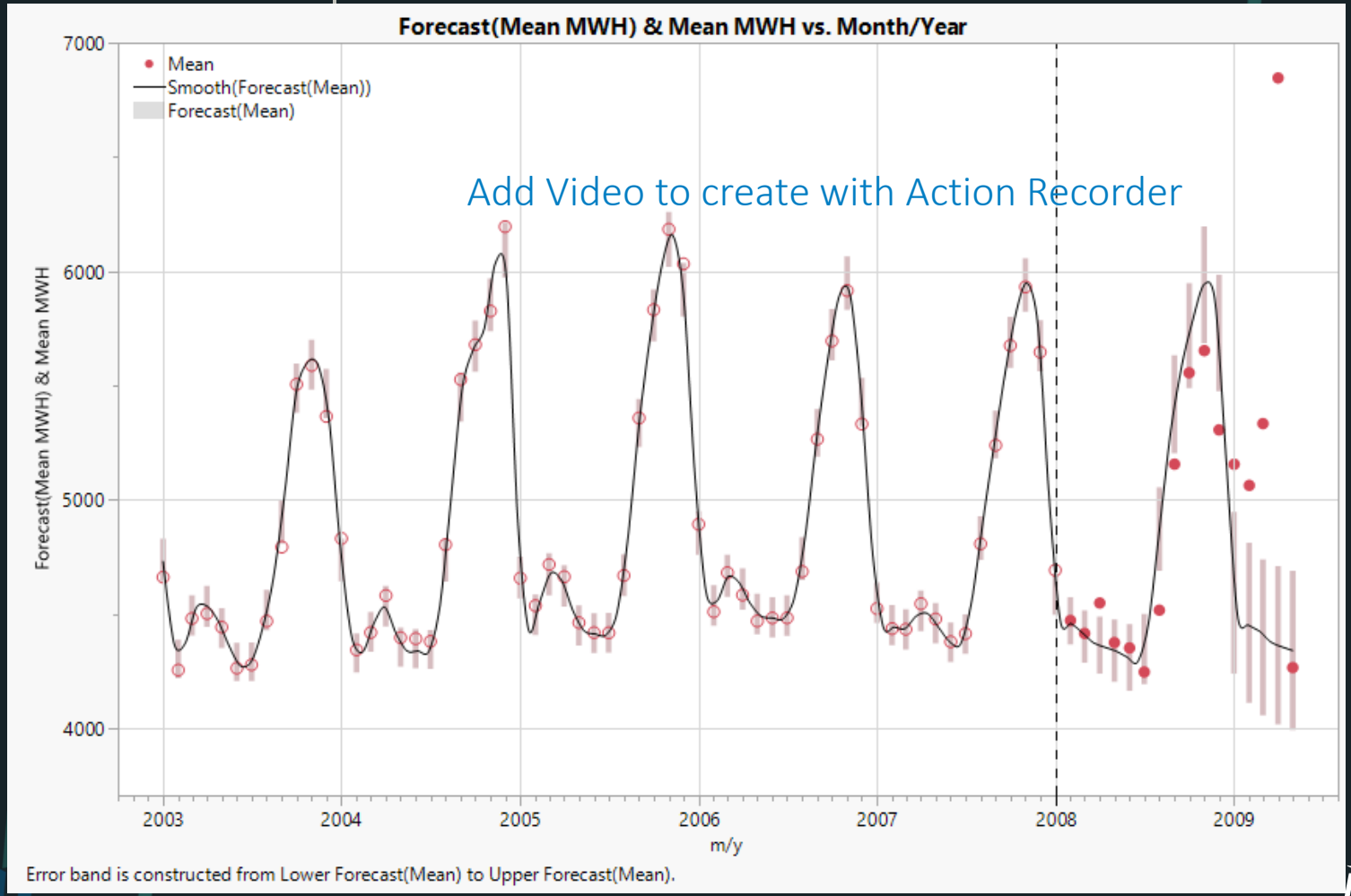
Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

ADDITIONAL TOPICS

- Import WGI data Set
- Time series for Army bases - Action Recorder MFG History
- DOE for Paint coating of shell
- Fuse testing simulation via model
- SFO Crime Dashboard
- Action Recorder MFG History

Copyright © JMP Statistical Discovery LLC. All rights reserved.



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY