Form 712A Non-Government Disclosure (Contractor, Academic, UARC etc.)

1. Check the Event Website for the form deadline at www.mors.org.
2. If using the MORS Presenter Center all forms must be uploaded in the Presenter Center.
3. For other events e-mail the completed form to liz.marriott@mors.org.
4. If the presentation classification and or Disclosure Statement below is different than those entered in the MORS Presenter Center, correct the Presenter Center entries to ensure proper labeling of the Presentation.

PART I

Author Request - The following author(s) request authority to disclose the following presentation at the MORS Event below with subsequent publication in the MORS Event Report and posting on the MORS website if applicable.

Principal Author Thomas A. Donnelly	Other Author(s)
Principal Author's Organization and complete mailing address SAS Institute Inc. 27 Farmingdale Ln Newark, DE 19711	Principle Author Phone 302-489-9291 Principle Author FAX Principle Author E-mail tom.donnelly@jmp.com
Principal Author's Signature X Thomas A. Donnelly	Date 3 June 2021
MORS Event 89th MORSS	Event Date(s) 21-24 June 2021

Presentation Type ☐ Plenary ☐ Course ☒ Tutorial ☐ Special Session ☐ Poster ☐ Demonstration

☐ Working/Composite/Distributed or Focus Group List All _____ ☐ Other _____

Title of Presentation Exploratory Data Analysis & Root Cause Analysis – "What can you do when you don't have a designed experiment?"	Presentation ID (if assigned) 56939

Classification ☐ SECRET ☐ SECRET//REL TO FVEY ☐ CONFIDENTIAL ☐ CONFIDENTIAL//REL TO FVEY ☒ UNCLASSIFIED ☐ UNCLASSIFIED W/FOUO ☐ Other _____

Distribution Statement ☒ A (Publicly Releasable) ☐ B ☐ C ☐ D ☐ E (see side 2 for definitions)

A.	This work was performed in connection with a government contract.	☐ YES (Complete Parts I, II & III)
B.	This presentation is based on material developed by the author as part of company/organization approved research e.g. IR&D and was NOT done under a government contract.	☐ YES (Complete Parts I, II & III)
C.	This presentation was NOT done under a government contract, contains no government information, is my own work and is approved for public release.	☒ YES (Complete Part I only)

Copyright © 2013, SAS Institute Inc. All rights reserved.

EXPLORATORY DATA ANALYSIS AND ROOT CAUSE ANALYSIS

"What can you do when you don't have a designed experiment?"

89th MORSS
Webcast Tutorial 56939
June 21, 2021

Tom Donnelly, PhD, CAP
Principal Systems Engineer
JMP Defense & Aerospace Team
SAS Institute Inc.
tom.donnelly@jmp.com

Copyright © 2013, SAS Institute Inc. All rights reserved.

TOPICS

Exploratory Data Analysis and Root Cause Analysis

- Dynamic linking – look for correlation between behavior and possible causes – make lots of graphs
  - Just don't think that **correlation** necessarily means **causation**!
- Use data filters to isolate factors and levels
- Preparing the data – outliers, missing, recoding
- Example data sets
  - Pharmaceutical – continuous & binary – root cause
  - Breast cancer – binary – PCA – Penalized Regression
  - Cyber Attack – Multinomial - binary
  - Diamonds – lots of categorical classifications
  - FAA Data – daily records
  - Hundreds of factors
  - Positional – latitude & longitude (GPS, satellite, image)
    - GPS, SST, Drone
  - What Type of Data Do You Have?

jmp
Copyright © 2013, SAS Institute Inc. All rights reserved.
sas THE POWER TO KNOW.

TOPICS

Exploratory Data Analysis and Root Cause Analysis

- Just how much are my factors correlated?
  - Matrix – VSS cont/cat response
  - PCA – breast cancer
- Create a design of experiments from candidate set –
  - Phamaceutical Tablet Data
  - Diamonds Data
  - FAA Data
- Use Data Mining Tools
  - Use a Robust “Honest Assessment Method” – Train, Validate & Test
  - Other Cross-Validation approaches – K-fold, Leave-One-Out
  - Penalization criteria – BIC, AICc, ERIC, Adjusted R-Square

jmp
Copyright © 2013, SAS Institute Inc. All rights reserved.
sas THE POWER TO KNOW.

TOPICS

Exploratory Data Analysis and Root Cause Analysis

- Decision tree methods (Useful when too many columns for regression)
- Regression – Linear, Generalized Linear (logistic), Stepwise (not always), Penalized (GenReg)
- Neural

- CKPS Today? Tomorrow?
  - Physics (timeless) vs People (shifty/adaptive)
- Case Study - FAA Data/Models
  - Cross-Validation approaches – K-fold, Leave-One-Out
  - Penalization criteria – BIC, AICc, ERIC, Adjusted R-Square

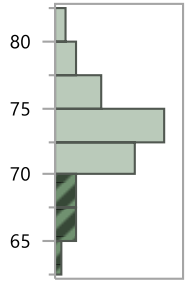
Copyright © 2013, SAS Institute Inc. All rights reserved.

INTERACTIVE SELECTION

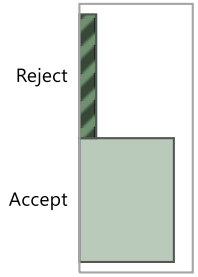
LOOK FOR CORRELATIONS BETWEEN MEASURED RESPONSES AND POTENTIAL FACTORS

Distributions

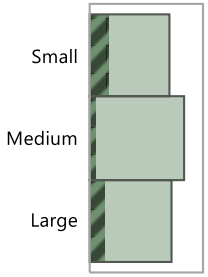
Dissolution



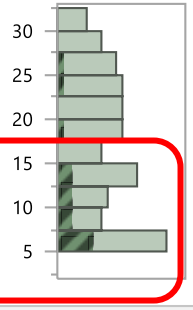
Lot Acceptance



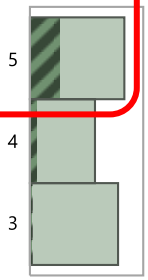
API Particle Size



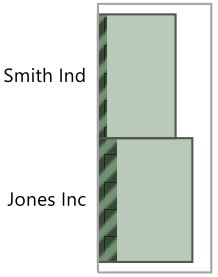
Mill Time



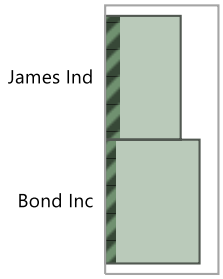
Screen Size



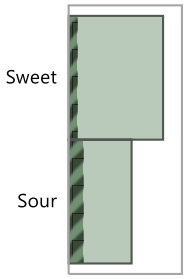
MgSt Supplier



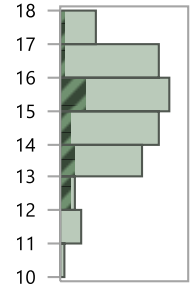
Lactose Supplier



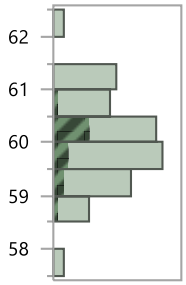
Sugar Supplier



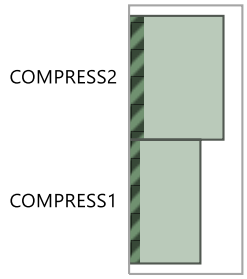
Blend Time



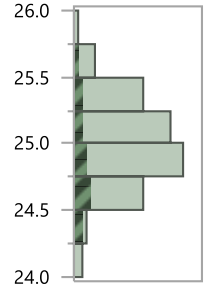
Blend Speed



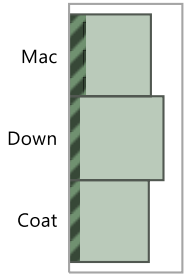
Compressor



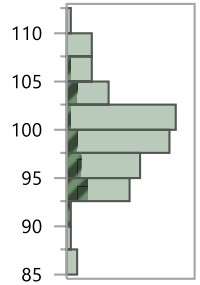
Force



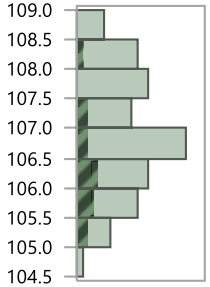
Coating Supplier



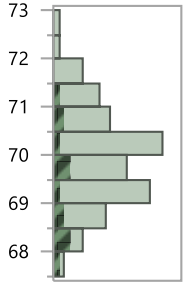
Coating Viscosity



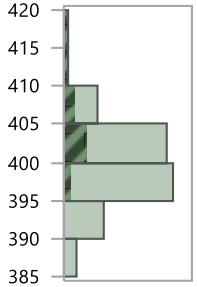
Inlet Temp



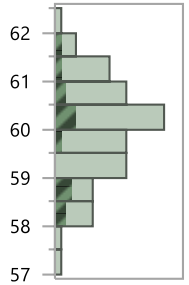
Exhaust Temp



Spray Rate



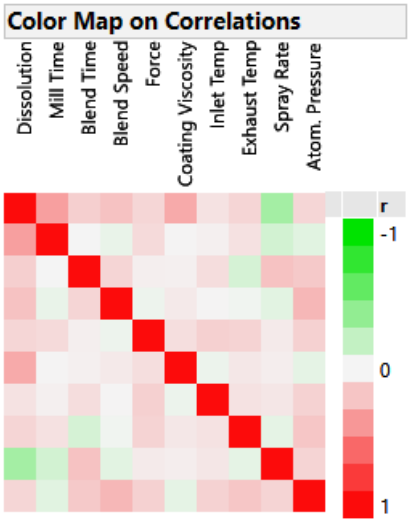
Atom. Pressure

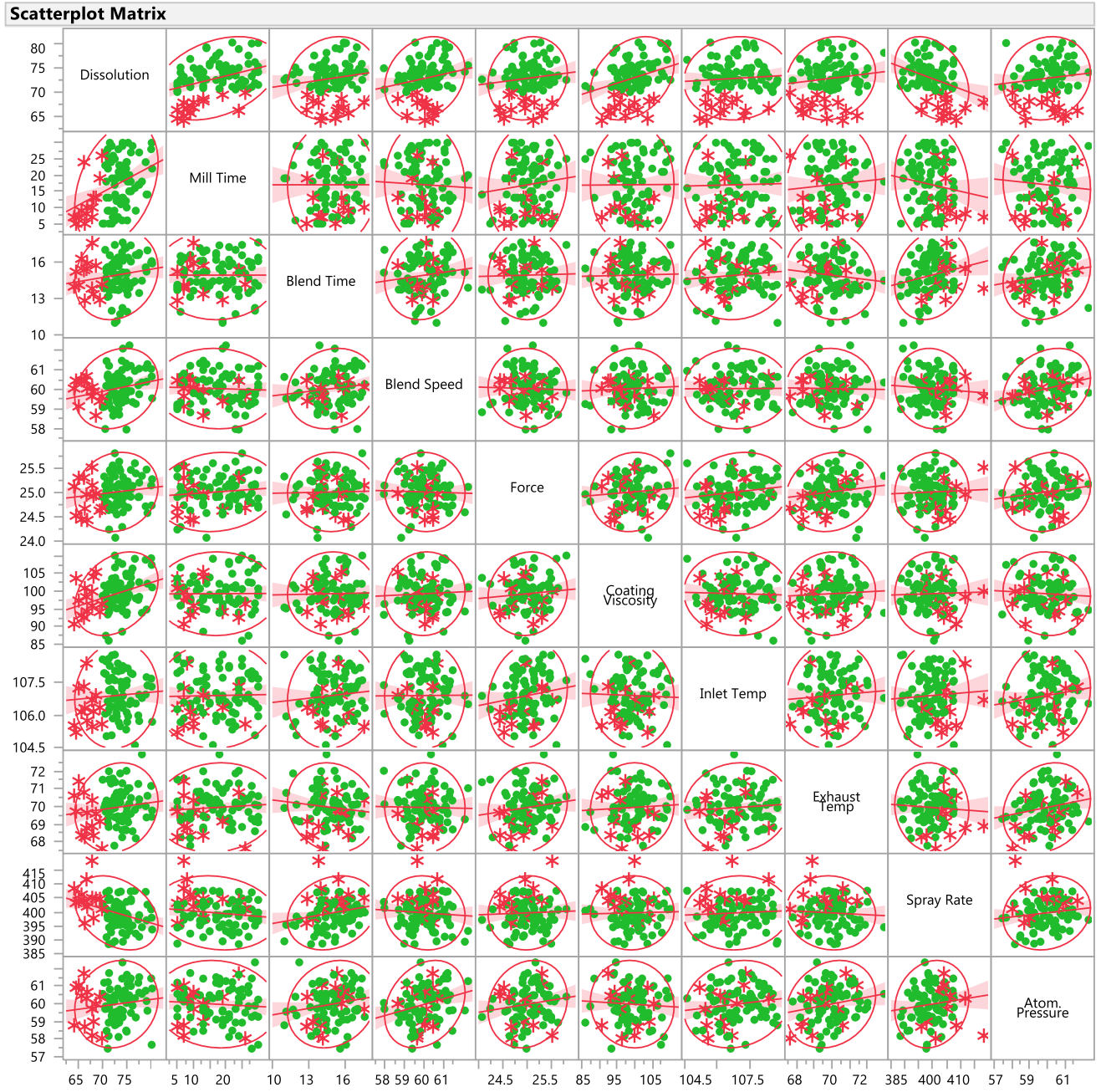


Copyright © 2013, SAS Institute Inc. All rights reserved.

CORRELATION OF CONTINUOUS COLUMNS

Look for correlations between responses and potential factors, as well as factors with each other

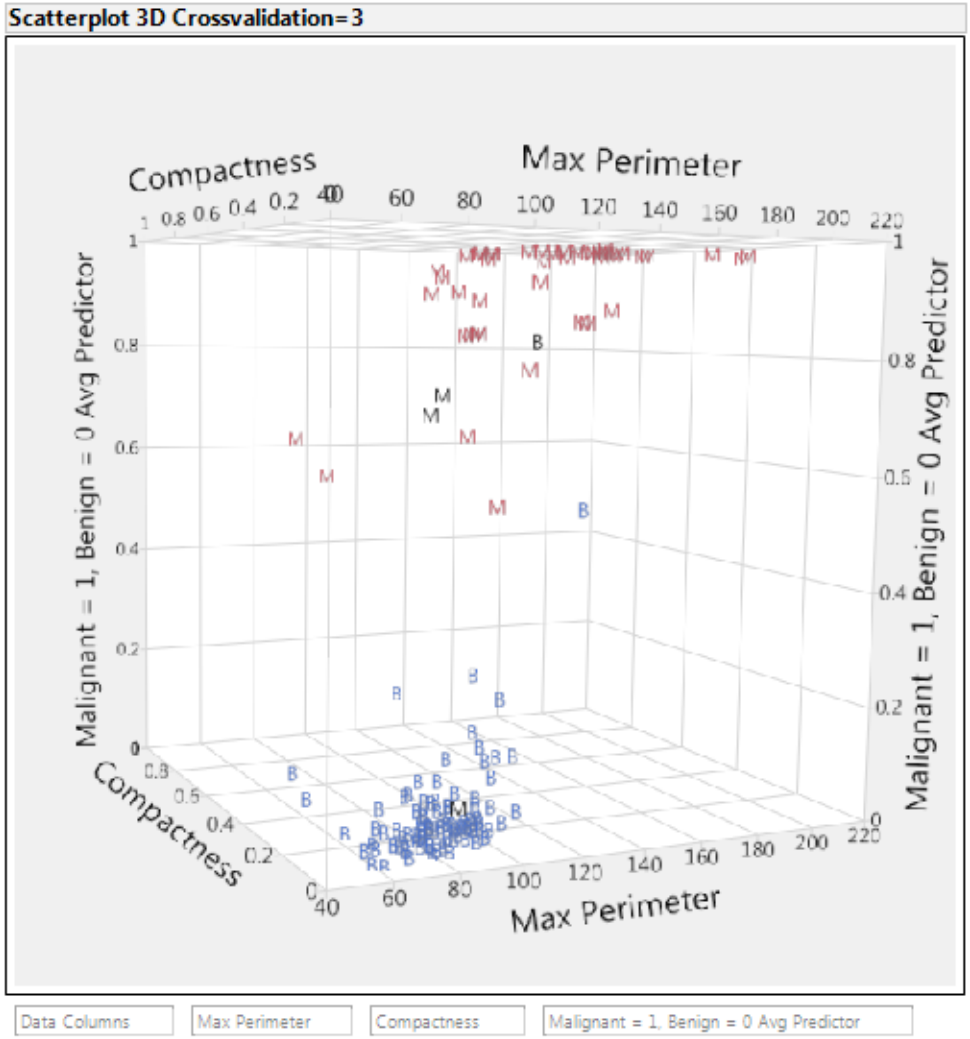




Copyright © 2013, SAS Institute Inc. All rights reserved.

TEST DATA

ACTUAL DIAGNOSIS (M OR B) PROBABILITY OF MALIGNANT AND TWO MAIN FACTORS



Copyright © 2013, SAS Institute Inc. All rights reserved.

JMP DEFENSE & AEROSPACE TEAM … AND HOW TO CONTACT US

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 & Co-Insurrectionist
tom.donnelly@jmp.com
302-489-9291

Copyright © 2013, SAS Institute Inc. All rights reserved.

Thanks.
Questions or comments?

TOM.DONNELLY@JMP.COM

Copyright © 2010 SAS Institute Inc. All rights reserved.