OUTLIER SCREENING IN TEST OF AUTOMOTIVE SEMICONDUCTORS

USE OF JMP 12 PRO 'MULTIVARIATE ANALYSIS' PLATFORMS AND 'EXPLORE OUTLIERS' UTILITY

CORINNE BERGÈS
BLACK BELT FOR ANALOG & SENSOR QUALITY

15, FEB, 2016

SECURE CONNECTIONS FOR A SMARTER WORLD

Introduction

- Hundreds of parametric tests per part

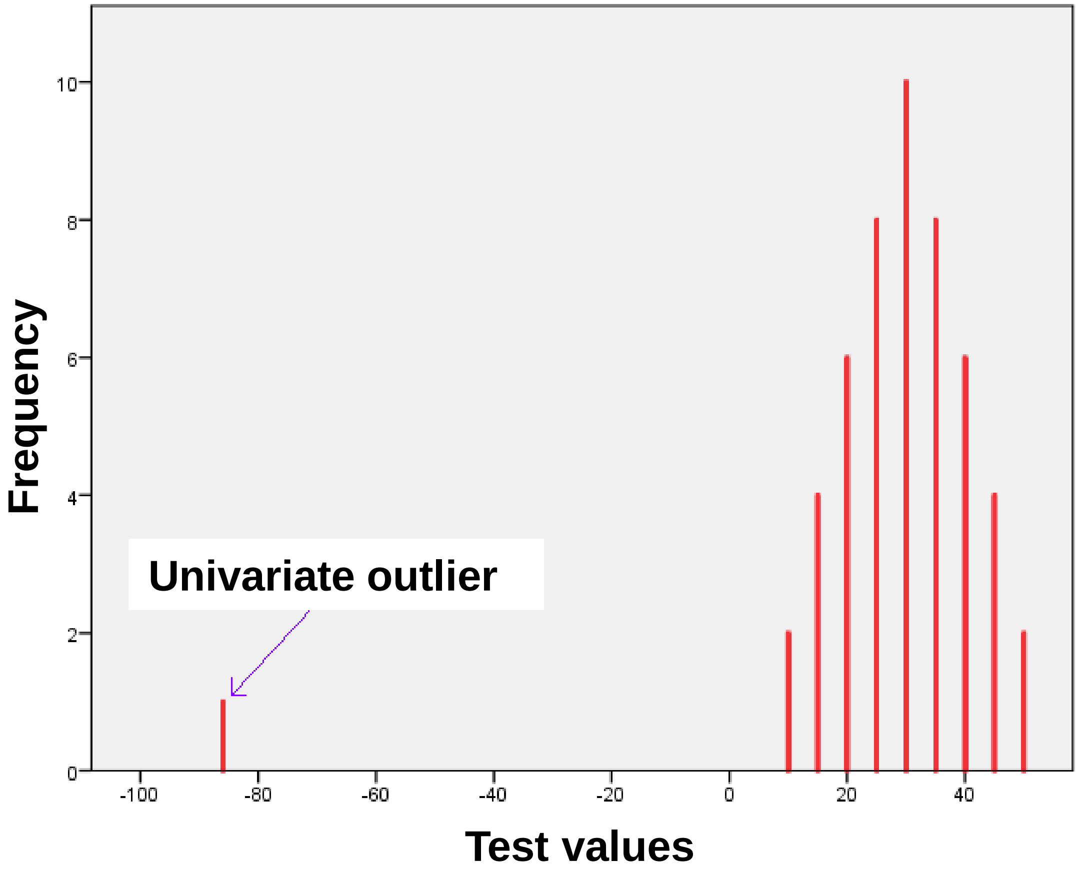
- Parts whose test results are far from normal process variability → Likely-to-fail parts / Outliers

- Test → Search for outliers

- Outlier detection means:
  - Univariate screening
  - Multivariate screening
    - methods without learning
    - methods with learning

Univariate analysis

Univariate outlier



Multivariate analysis with jmp

Compass

1. Presentation on some multivariate techniques that are possible to be run with jmp
   →Case study for an automotive valve driver

2. Considerations about space size

3. Efficiency and yield loss

4. ‘Explore Outliers’ jmp platform

1. Some multivariate technics with jmp

- **Methods without learning step**
- **Methods needing a learning step**

Types of multivariate analysis

- **Methods without learning step**
  - Based on a detection threshold that is direclty linked with yield loss
  - Challenge: setting of a threshold that detects returns with the lowest yield loss
  - Examples: Mahalanobis distance estimation, k-means clustering method, deviation estimation from a linear regression

Mahalanobis distance estimation

Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

K-near neighbors and clustering methods

Distance estimation from each observation to the K-near neighbors

Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last centroids by new ones including the last observation assigned

Deviation estimation from a linear regression

Bivariate method (2 tests) on tests highly correlated

Distance estimation from each point to the linear regression line between the 2 tests

Types of multivariate analysis

- **Methods needing a learning step**
  - Implementation step:
    - Learning on first well-known customer returns
    - Running of the first built model to detect outliers among the following manufactured parts
    - Improvement of the first model by new potential returns
  - Challenge: building of a model that does not stick to the part sample but that could be used to detect outliers and returns on other following part samples (overfitting risk)
  - Examples: discriminant analysis, partial least squares (pls)

Discriminant analysis

Membership prediction in a category (failed/not-failed) from observed values

Search for a test combination that provides a maximal Mahalanobis distance between the two groups

Entropy R-Square measures model fit

Partial Least Squares (PLS)

Trade-off between two purposes: to maximise:

Explained variance of the predictors

Correlation between variables and response

Main method advantage: to be run even if number of tests > number of parts

Two available algorithms: NIPALS and SIMPLS

Types of multivariate analysis

- **Methods without learning step**
  - Based on a detection threshold that is direclty linked with yield loss
  - Challenge: setting of a threshold that detects returns with the lowest yield loss
  - Examples: Mahalanobis distance estimation, k-means clustering method, deviation estimation from a linear regression

- **Methods needing a learning step**
  - Implementation step:
    - Learning on first well-known customer returns
    - Running of the first built model to detect outliers among the following manufactured parts
    - Improvement of the first model by new potential returns
  - Challenge: building of a model that does not stick to the part sample but that could be used to detect outliers and returns on other following part samples (overfitting risk)
  - Examples: discriminant analysis, partial least squares (pls)

→ *Success criteria: low yield loss*

How to do with jmp ?
Case study
for an automotive valve driver

Case study

- Product: automotive valve driver

- Failing of univariate detection for several customer returns

- Test of several multivariate analysis methods:
  - 745 tests (where standard deviation is not null)
  - 13 000 parts → only one is failing = the customer return
  - Around 20 wafers tested at final test (after part assembly)
- Space size: n = 13 000 parts; p = 745 tests

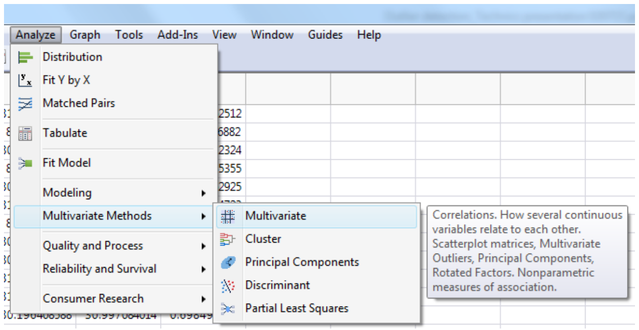
- Question: *what is the best multivariate method to detect the customer returns **with jmp** ?*

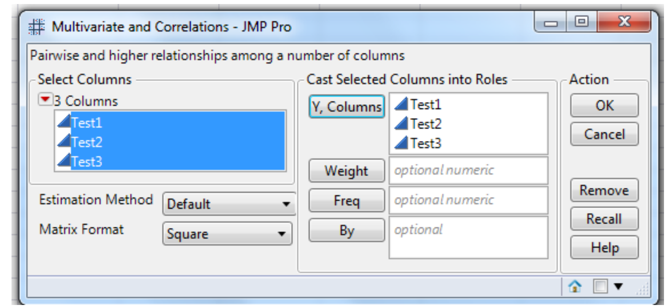
Methods without learning (1/3)

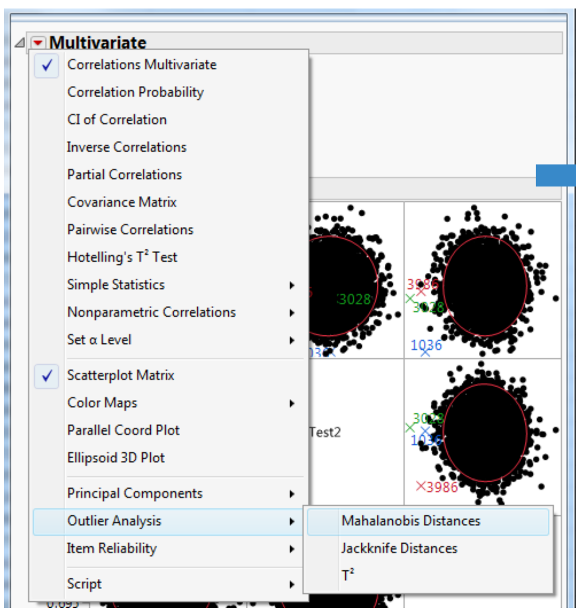
Mahalanobis distance estimation

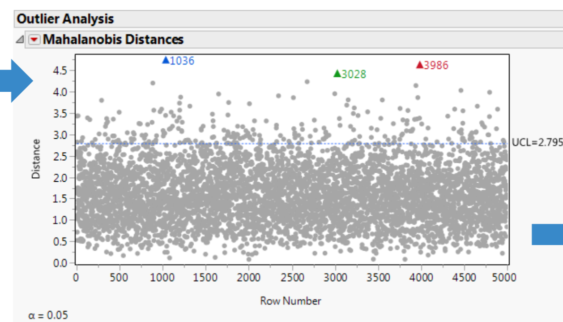
Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

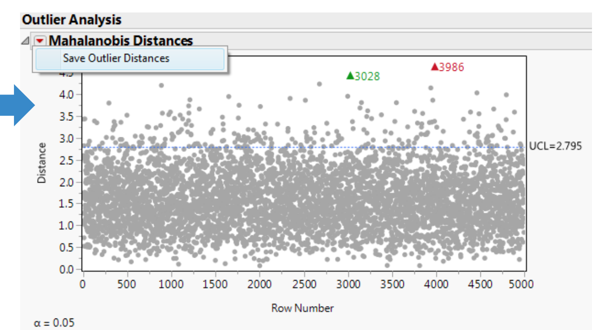
File: ‘Multivariate analysis.jmp’ → 3 tests











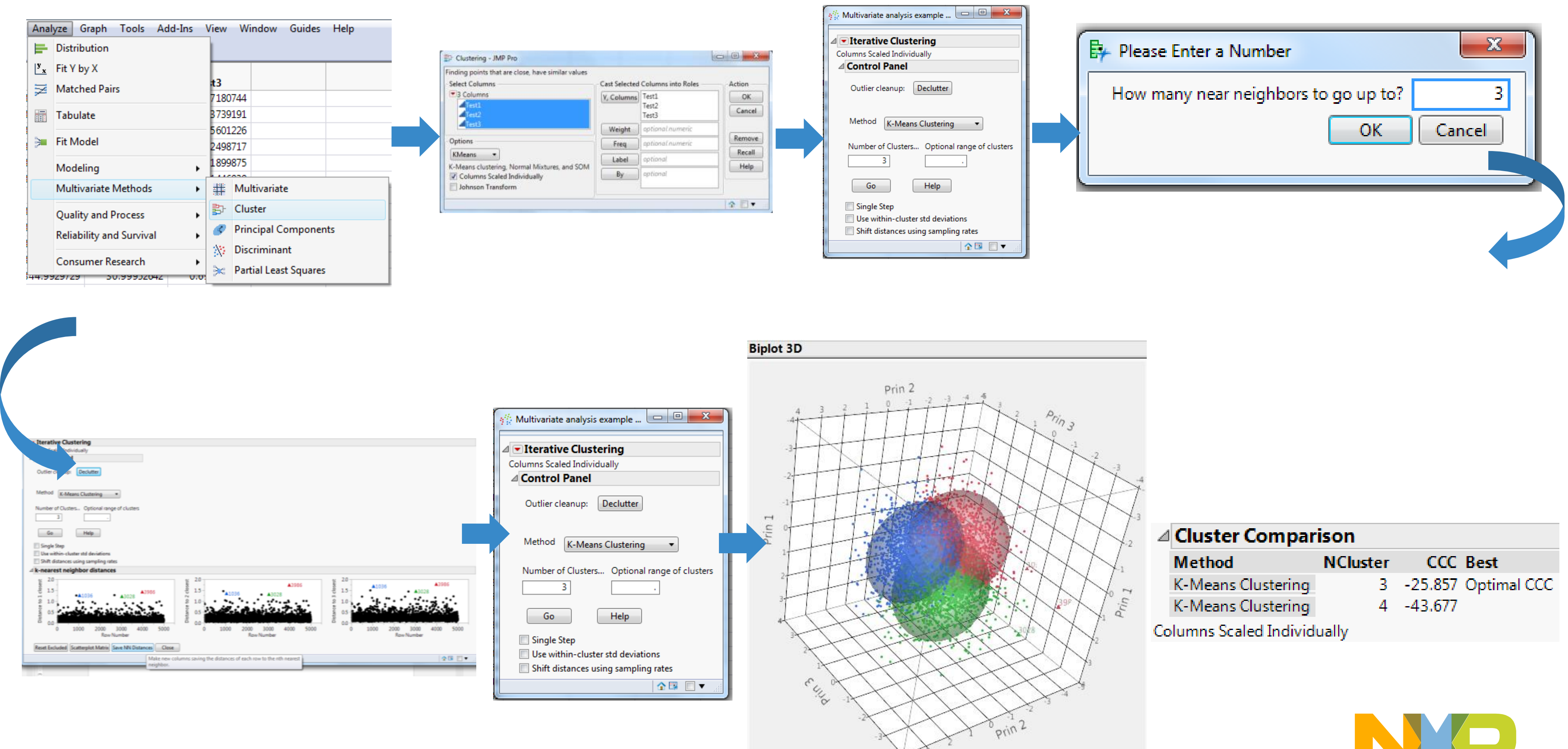
Methods without learning (2/3)

K-near neighbors and clustering methods

Distance estimation from each observation to the K-near neighbors

Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last centroids by new ones including the last observation assigned

File: ‘Multivariate analysis.jmp’ → 3 tests

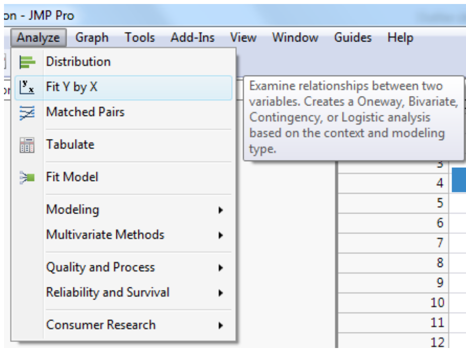


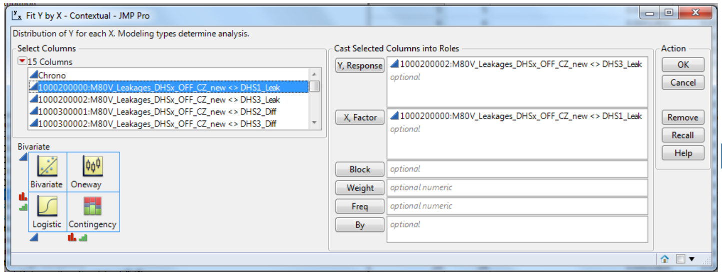
Methods without learning (3/3)

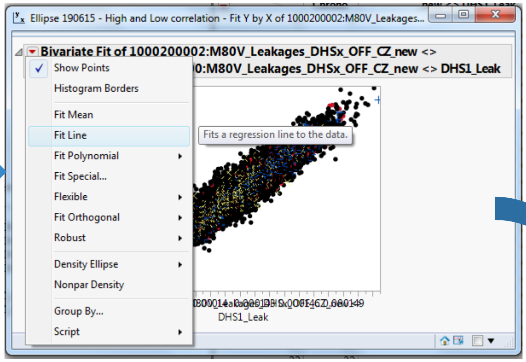
Deviation estimation from a linear regression

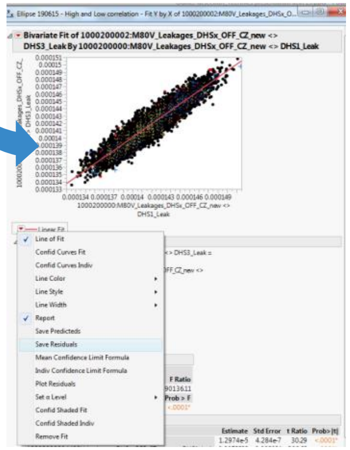
Bivariate method (2 tests) on tests highly correlated

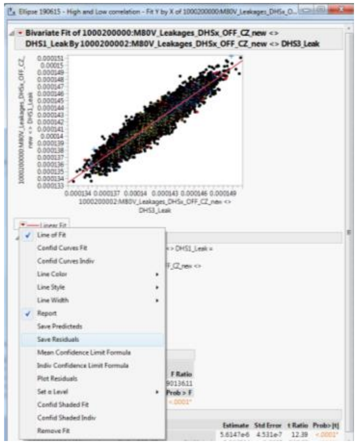
Distance estimation from each point to the linear regression line between the 2 tests

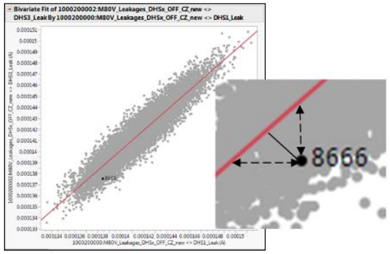












$$Distance\ (P(x,y), regression\ line) = \sqrt{[residuals(test2\ vs\ test1)]^2 + [residuals(test1\ vs\ test2)]^2/2}$$

File: ‘Bivariate and PCA.jmp’

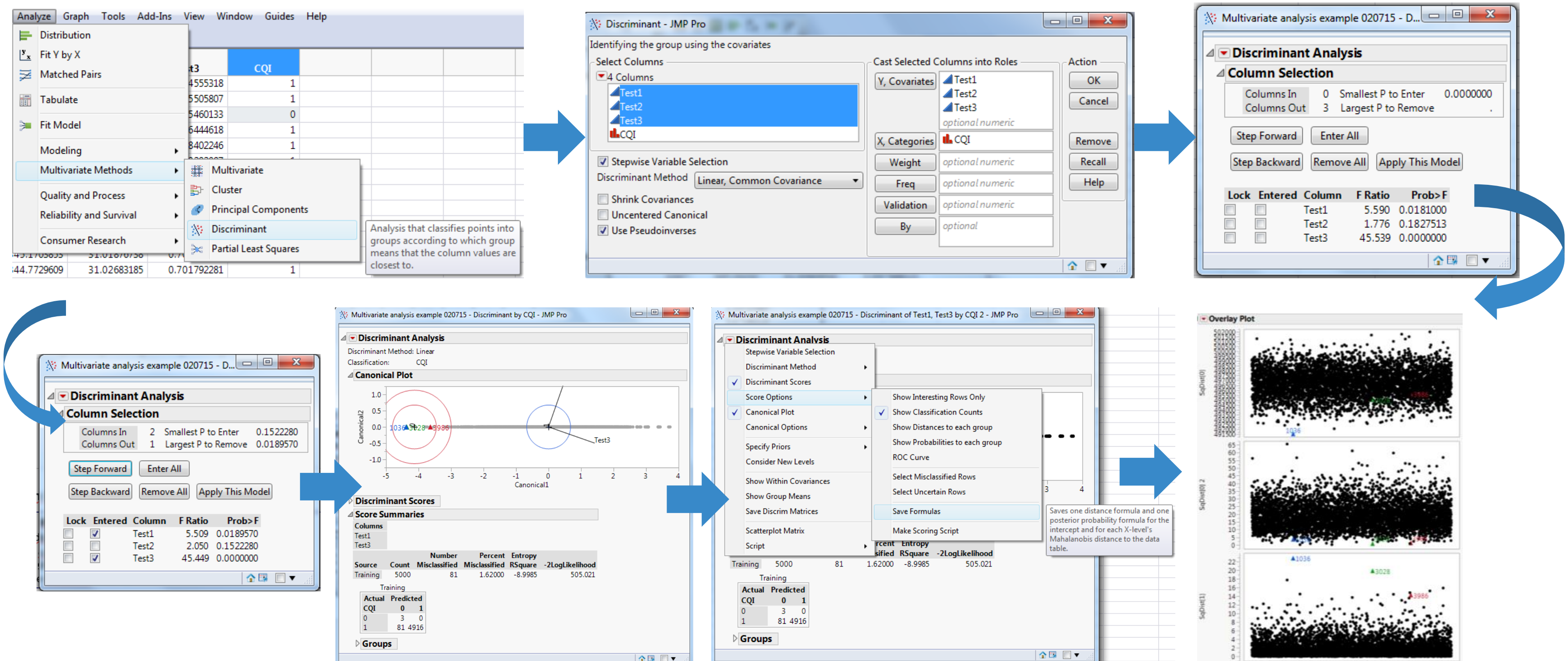
Methods with learning (1/2)

Discriminant analysis

Membership prediction in a category (failed/not-failed) from observed values

Search for a test combination that provides a maximal Mahalanobis distance between the two groups

Entropy R-Square measures model fit



File: ‘Bivariate and PCA.jmp’

Methods with learning (2/2)

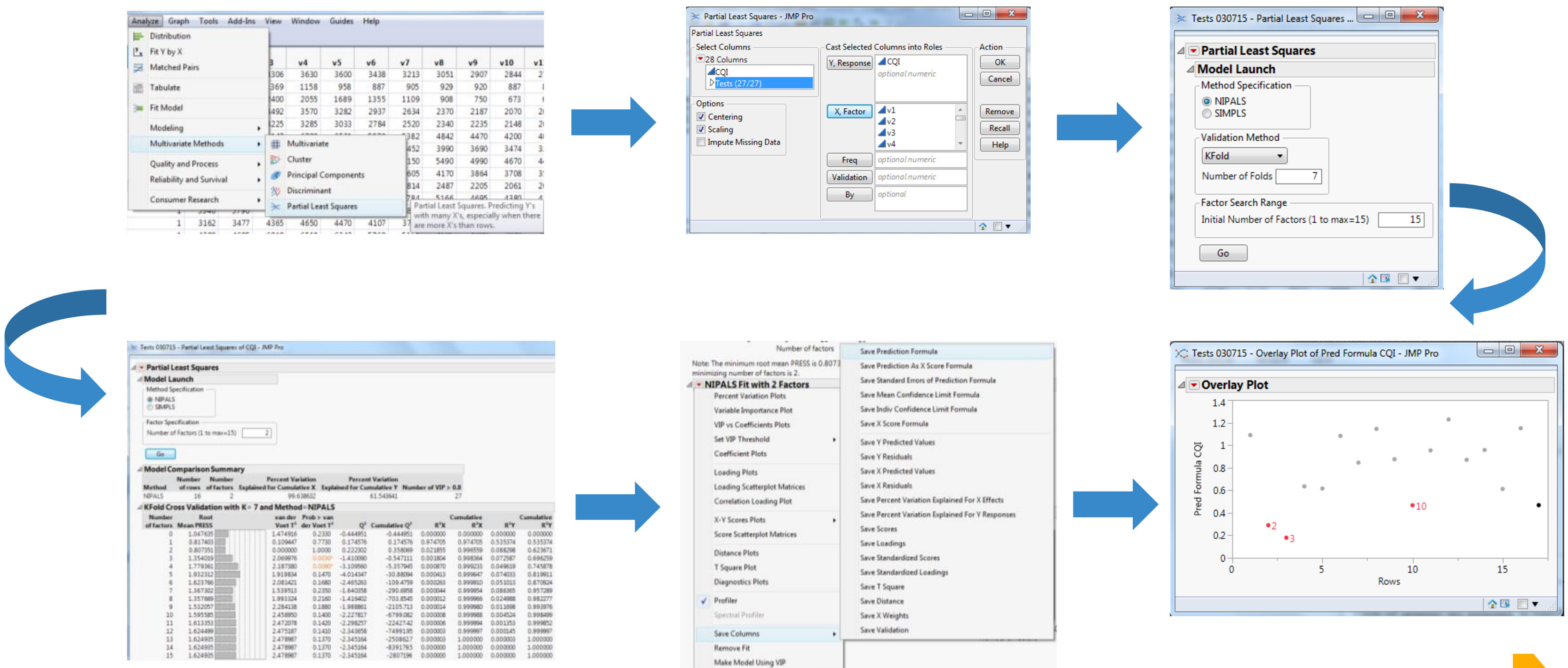
Partial Least Squares (PLS)

Trade-off between two purposes: to maximise:

- Explained variance of the predictors
- Correlation between variables and response

Main method advantage: to be run even if number of tests > number of parts

Two available algorithms: NIPALS and SIMPLS

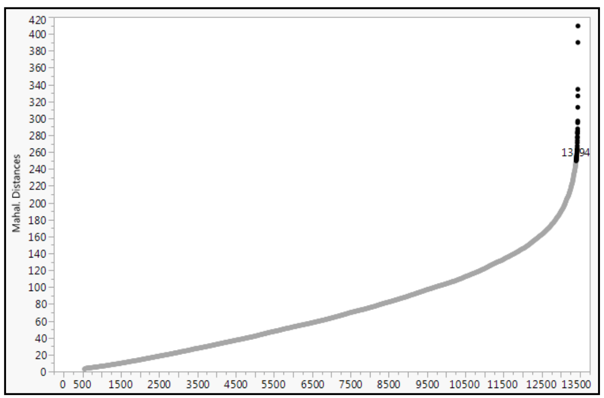


File: 'Partial Least Square.jmp'

Methods without learning

Mahalanobis distance estimation

Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

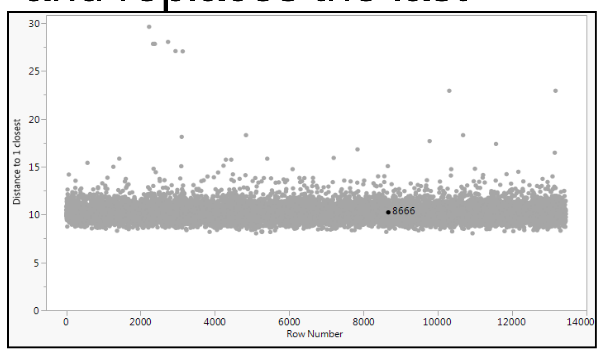


Mahalanobis distance plot for the 13000 parts

K-near neighbors and clustering methods

Distance estimation from each observation to the K-near neighbors

Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last

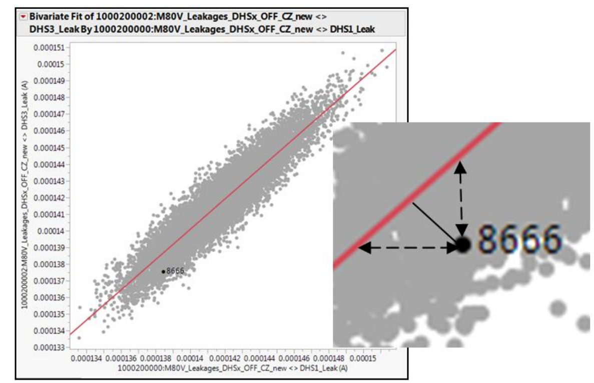


Distance of each part with the first nearest neighbor

Deviation estimation from a linear regression

Bivariate method (2 tests) on tests highly correlated

Distance estimation from each point to the linear regression line between the 2 tests



Distance computation for the return to regression line

Case study: best multivariate method → Mahalanobis distance with a yield loss = 0.36%

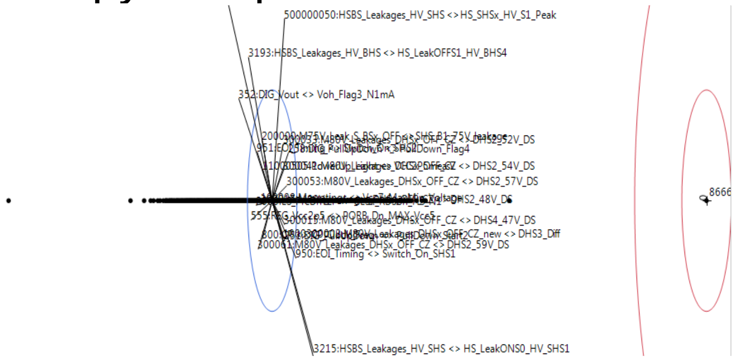
Methods with learning

- **Discriminant analysis**

  Membership prediction in a category (failed/not-failed) from observed values

  Search for a test combination that provides a maximal Mahalanobis distance between the two groups

  Entropy R-Square measures model fit



Mahalanobis distance plot in two-dimensional space

Case study: low entropy R-Square (0.39)

- **Partial Least Squares (PLS)**

  Trade-off between two purposes: to maximise:

  Explained variance of the predictors

  Correlation between variables and response

  Main method advantage: to be run even if number of tests > number of parts

  Two available algorithms: NIPALS and SIMPLS

Case study: Mahalanobis distance without learning stays the best multivariate method

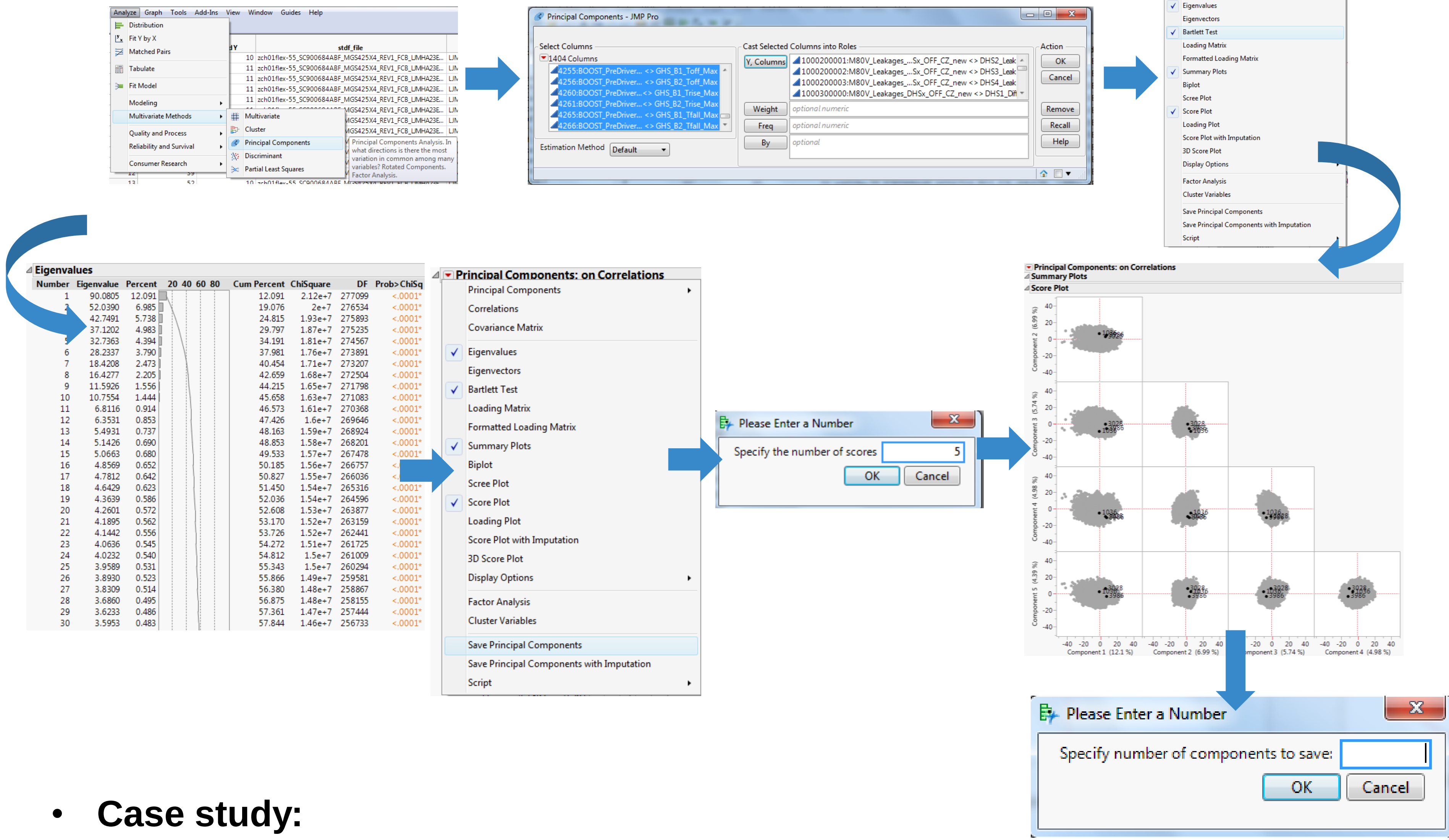
2. Space size

Considerations about space size

- **Size reduction motivation:**
  - Test reduction:
    - Better results on a reduced space and on correlated tests
    - For the analysis with learning, overfitting risk reduction
  - Part reduction: noise reduction on homogeneous data

- **Size reduction means:**
  - Statistical analysis: Principal Component Analysis (PCA)
  - Other selection criteria: functionality criteria
  - Part reduction: run on wafer lot

Principal Component Analysis (PCA)



Eigenvalues

Number	Eigenvalue	Percent	Cum Percent	ChiSquare	DF	Prob>ChiSq
1	90.0805	12.091	12.091	2.12e+7	277099	<.0001*
2	52.0390	6.985	19.076	2e+7	276534	<.0001*
3	42.7491	5.738	24.815	1.93e+7	275893	<.0001*
4	37.1202	4.983	29.797	1.87e+7	275235	<.0001*
5	32.7363	4.394	34.191	1.81e+7	274567	<.0001*
6	28.2337	3.790	37.981	1.76e+7	273891	<.0001*
7	18.4208	2.473	40.454	1.71e+7	273207	<.0001*
8	16.4277	2.205	42.659	1.68e+7	272504	<.0001*
9	11.5926	1.556	44.215	1.65e+7	271798	<.0001*
10	10.7554	1.444	45.658	1.63e+7	271083	<.0001*
11	6.8116	0.914	46.573	1.61e+7	270368	<.0001*
12	6.3531	0.853	47.426	1.6e+7	269646	<.0001*
13	5.4931	0.737	48.163	1.59e+7	268924	<.0001*
14	5.1426	0.690	48.853	1.58e+7	268201	<.0001*
15	5.0663	0.680	49.533	1.57e+7	267478	<.0001*
16	4.8569	0.652	50.185	1.56e+7	266757	
17	4.7812	0.642	50.827	1.55e+7	266036	
18	4.6429	0.623	51.450	1.54e+7	265316	<.0001*
19	4.3639	0.586	52.036	1.54e+7	264596	<.0001*
20	4.2601	0.572	52.608	1.53e+7	263877	<.0001*
21	4.1895	0.562	53.170	1.52e+7	263159	<.0001*
22	4.1442	0.556	53.726	1.52e+7	262441	<.0001*
23	4.0636	0.545	54.272	1.51e+7	261725	<.0001*
24	4.0232	0.540	54.812	1.5e+7	261009	<.0001*
25	3.9589	0.531	55.343	1.5e+7	260294	<.0001*
26	3.8930	0.523	55.866	1.49e+7	259581	<.0001*
27	3.8309	0.514	56.380	1.48e+7	258867	<.0001*
28	3.6860	0.495	56.875	1.48e+7	258155	<.0001*
29	3.6233	0.486	57.361	1.47e+7	257444	<.0001*
30	3.5953	0.483	57.844	1.46e+7	256733	<.0001*

- **Case study:**
  After PCA, return detected with a higher yield loss (1%)

File: ‘Bivariate and PCA.jmp’

3. Efficiency and yield loss

Efficiency and yield loss (1/5)

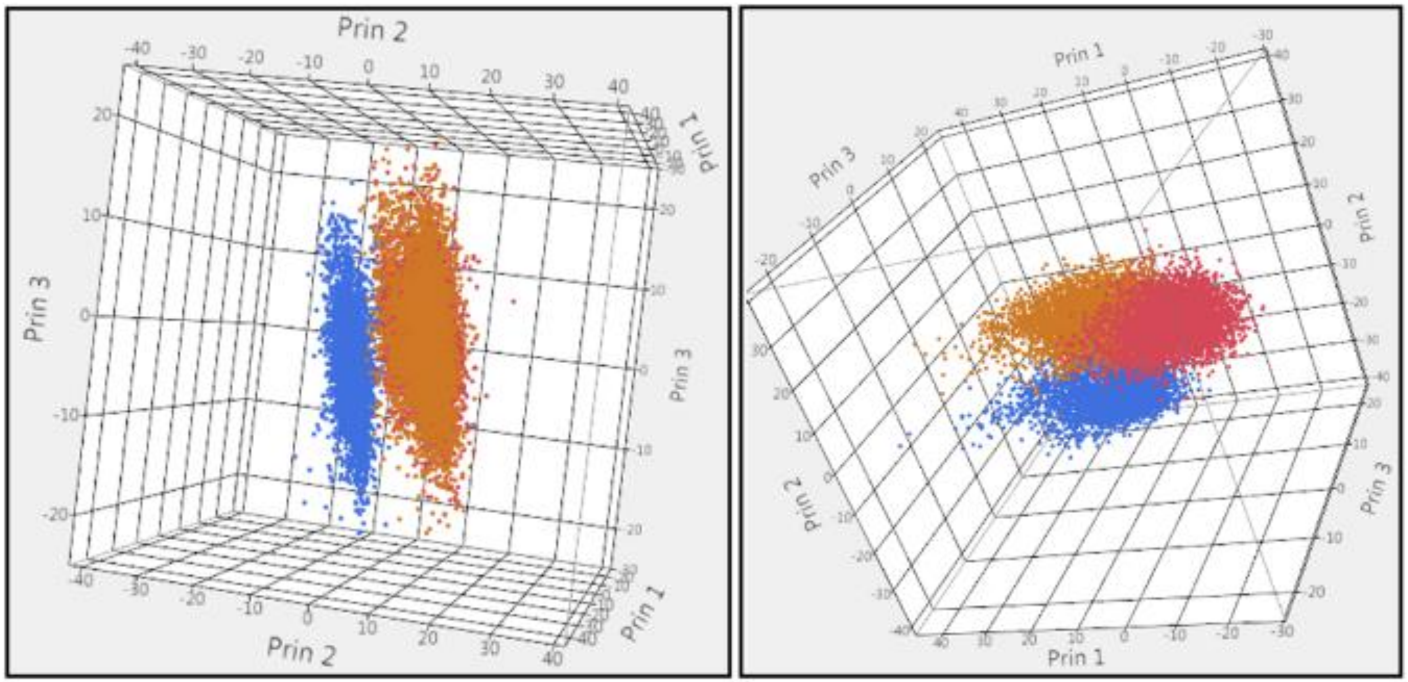
- Efficiency: outlier detection ability with minimal yield loss

- One mean to increase efficiency: noise reduction

- Case study: test performed on fours sites → multivariate analysis to vizualize and understand additional noise due to sites:
  - K-means clustering method
  - Contingency analysis
  - ANOVA

Efficiency and yield loss (2/5)

- **K-means clustering method**

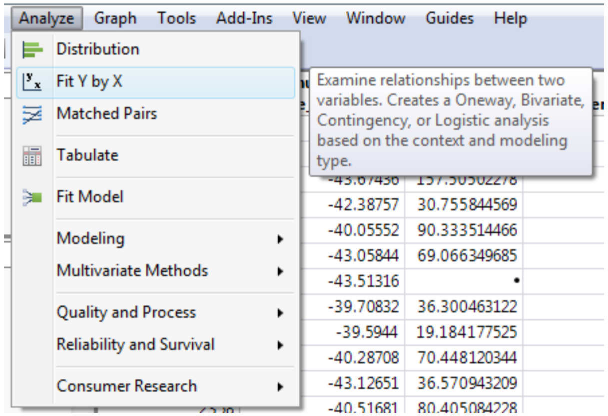


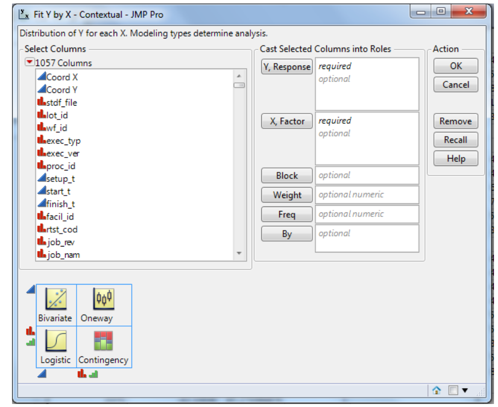
Two clusters observed:
-one for one site (blue) → *Cluster #10 for the following study*
-one gathering data from three sites → *Cluster #20 for the following study*

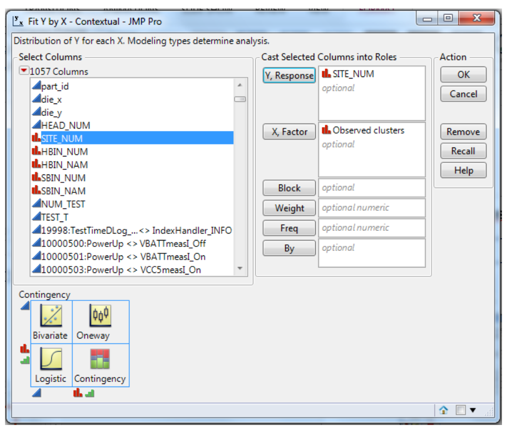
File: ‘Noise analysis.jmp’

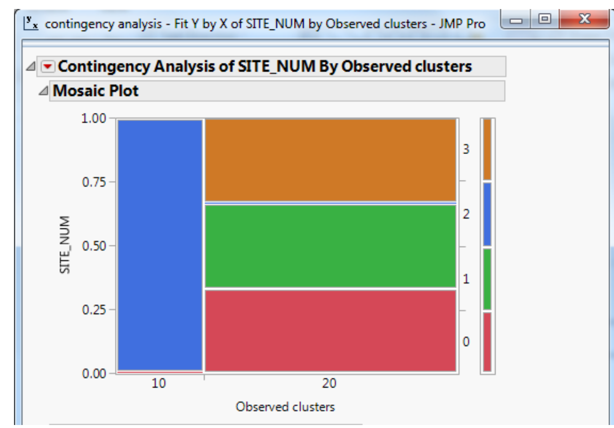
Efficiency and yield loss (3/5)

- **Contingency analysis**









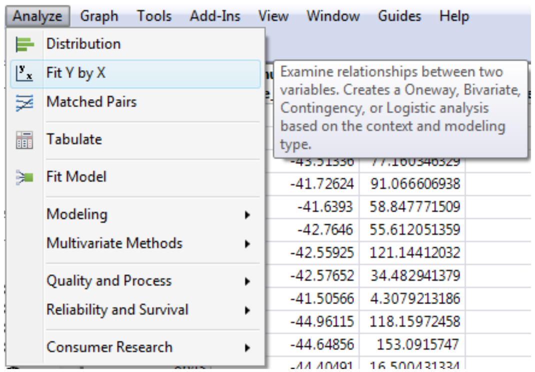
→ Cluster #10 contains site 2 data

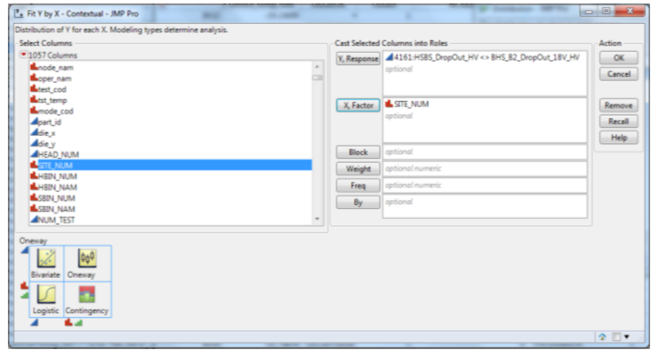
→ Cluster #20 contains data from sites 0, 1 and 3

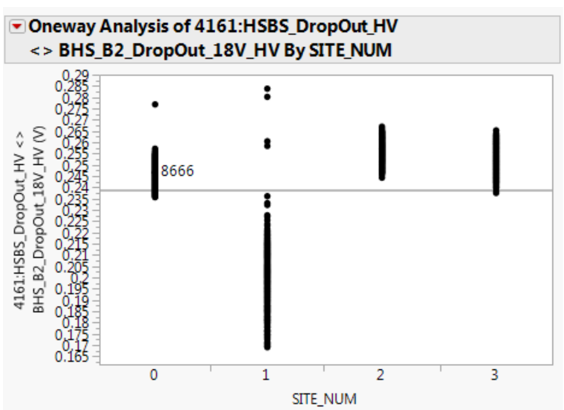
File: 'Noise analysis.jmp'

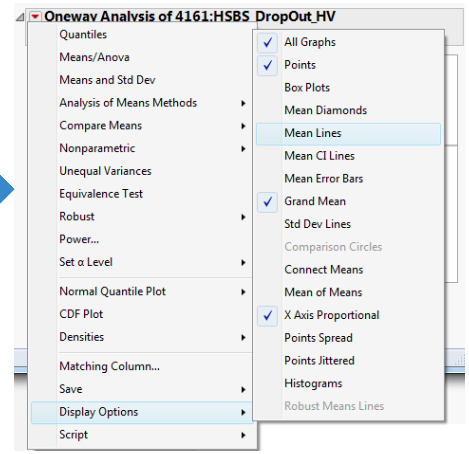
Efficiency and yield loss (4/5)

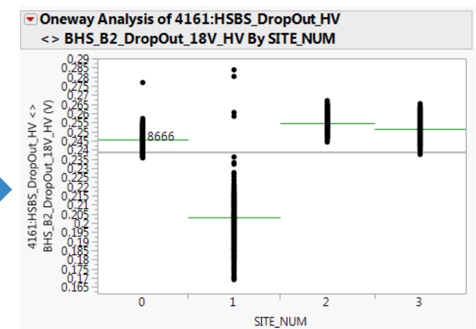
- **ANOVA** → *ANOVA of one test distribution by site*











Statistical test to compare means

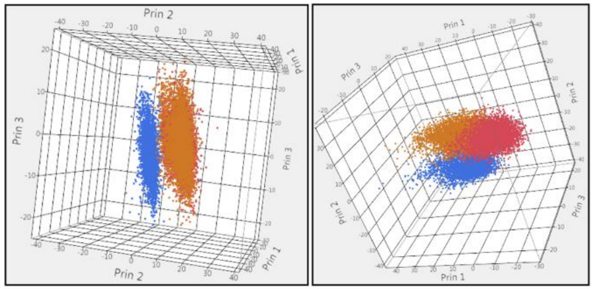
For this test, data from the site 1 are significantly different from the other sites

File: 'Noise analysis.jmp'

Efficiency and yield loss (5/5)

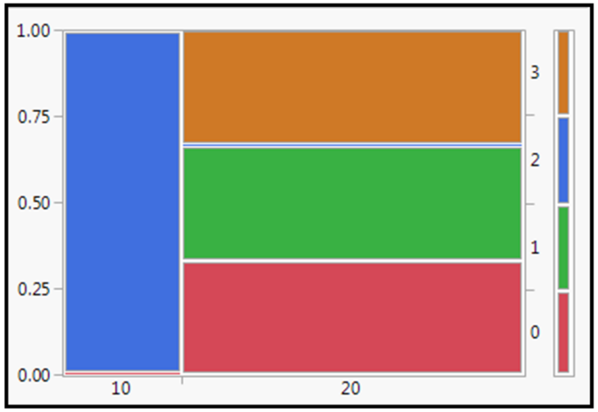
- Efficiency: outlier detection ability with minimal yield loss
- One mean to increase efficiency: noise reduction
- Case study: test performed on fours sites → multivariate analysis to vizualize and understand additional noise

- **K-means clustering method**



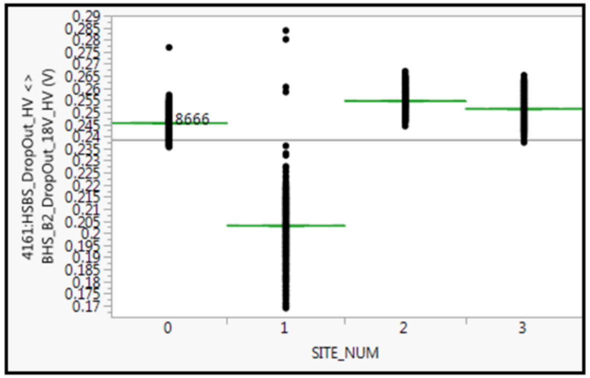
Two clusters:
-one for one site (blue)
-one gathering data from three sites

- **Contingency analysis**



Vizualisation of clustering method results

- **ANOVA**



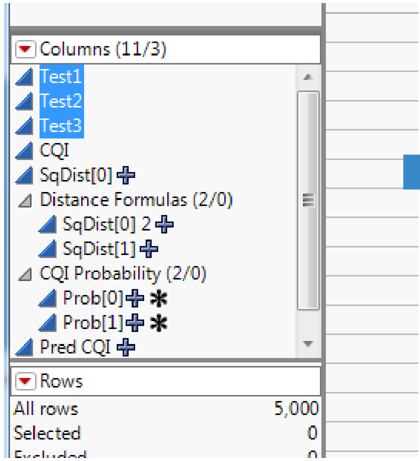
ANOVA of one test distribution by site

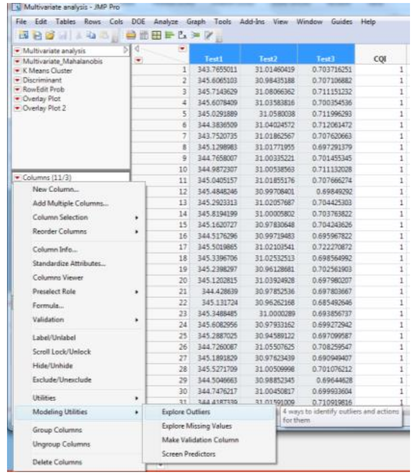
- **Noise elimination after part test: possibility to shift and align means of each site → *Yield loss decrease***

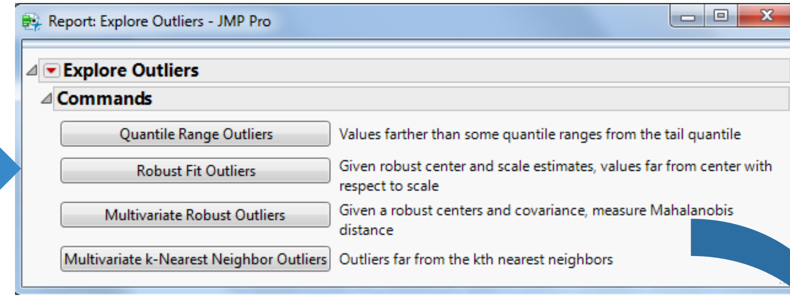
4. ‘Explore Outliers’ jmp platform

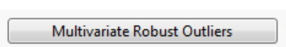
‘Explore Outliers’ jmp platform

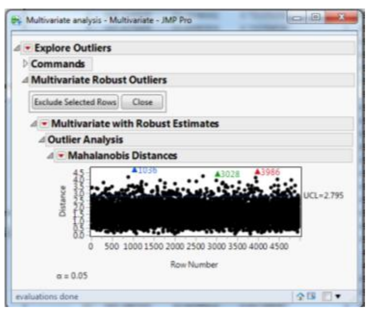
File: ‘Multivariate analysis.jmp’ → 3 tests

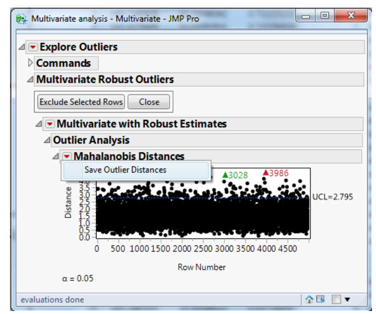


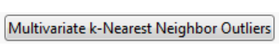


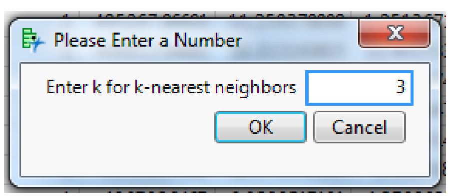


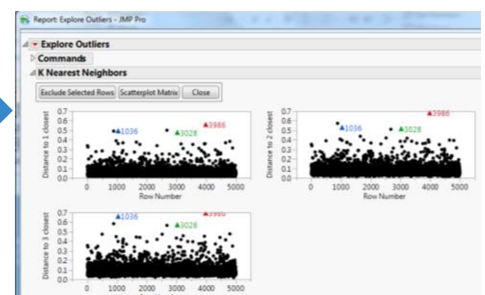


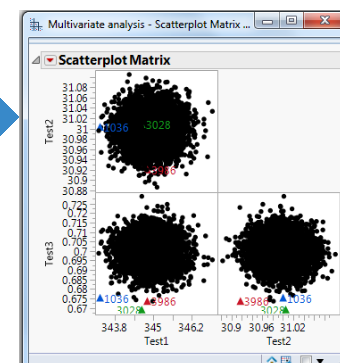












Conclusion

Conclusion

- Outlier detection in univariate analysis → Robust PAT for a better detection (real outliers) and a lower yield loss

- Outlier detection in multivariate analysis:
  - Many multivariate analysis based on the spatial Mahalanobis distance
  - Method without learning:
    - Useful data diluted in multidimensional space
    - High computation time and cost
    - In a reduced space, higher yield loss
  - Method with learning: reduced space but detection failing risk increase
  - One of the easiest method to be implemented: 'distance-to-regression-line estimation' method: Python used in the model design step / EWM in probe
  - Many other methods have to tested, in Python or in jmp, above all when a CQI happens

- In order to improve detection with a lower yield loss, a preliminary step has to be gage study / noise reduction and elimination → will benefit also the univariate analysis

Thank you for your attention

Any question ?

→ *Please, feel free to contact Corinne Bergès:*

corinne.berges@nxp.com

SECURE CONNECTIONS
FOR A SMARTER WORLD