

# Design Of Experiment (DOE) & Response Surface Methodology (RSM)

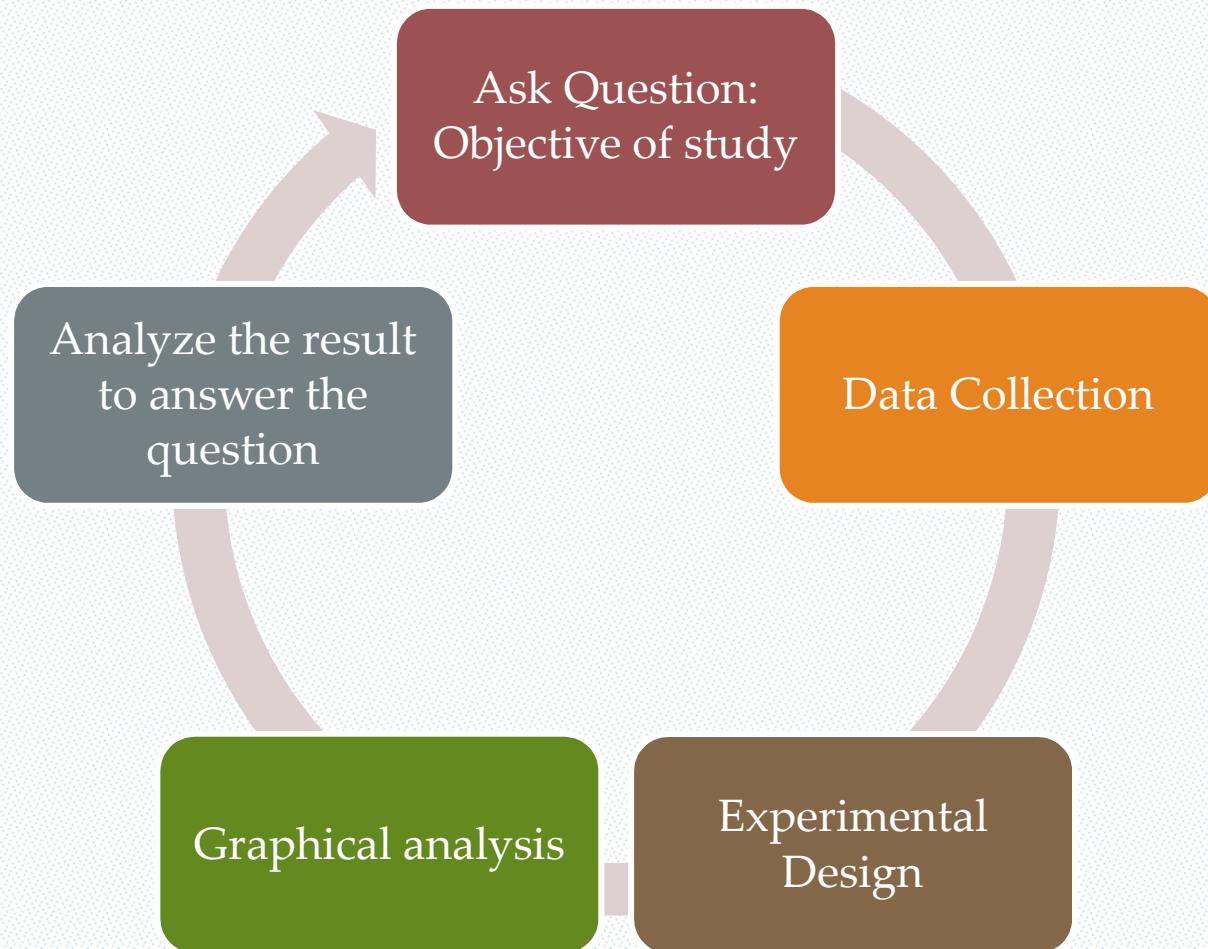
Present by:  
Wan Nor Nadyaini Wan Omar,  
B.Eng (Chem),M.Eng (Chem)  
*Faculty of Chemical Engineering*  
(wnnadyaini@gmail.com)

Date: 12 Aug 2015  
Place: N29,,Faculty of Chemical Engineering

# IF??



# Research Cycle process



# DATA COLLECTION

To clarify the objective of experiment

The experimenter must determine

- What data to be collected?
- How to measure it?
- How the data relates to process performances and experimental objective?

The experimenter must ensure the data collected is represented the process

- Thus, the data could lead to correct conclusion

The experimental design must related to experimental objectives

# Experimental Design

## Conventional Method

- One factor at a time (OFAT)
- Time consuming
- Cannot interpret the interaction between 2 or more variables

## Statistical method

- Known as Design of Experiment (DOE)
- Apply factorial concept
- Use the modelling to predict the behavior of process variables
- RSM, ANN etc.
- Could explain the interaction between the process variables
- Reduce lead time and improve efficiency

## What is DOE?

A collection of predetermined process variables setting

## What is RSM?

Response Surface Methodology (RSM) is a **statistic techniques** employed a **regression analysis** to performed for the collective data.

## What is STATISTICA, Design Expert, MiniTab and etc?

- is a **tools** to help we designs our experiment and analyses our data.
- RSM is one of the technique that have been programmed in that software.

# What DOE & RSM can do?

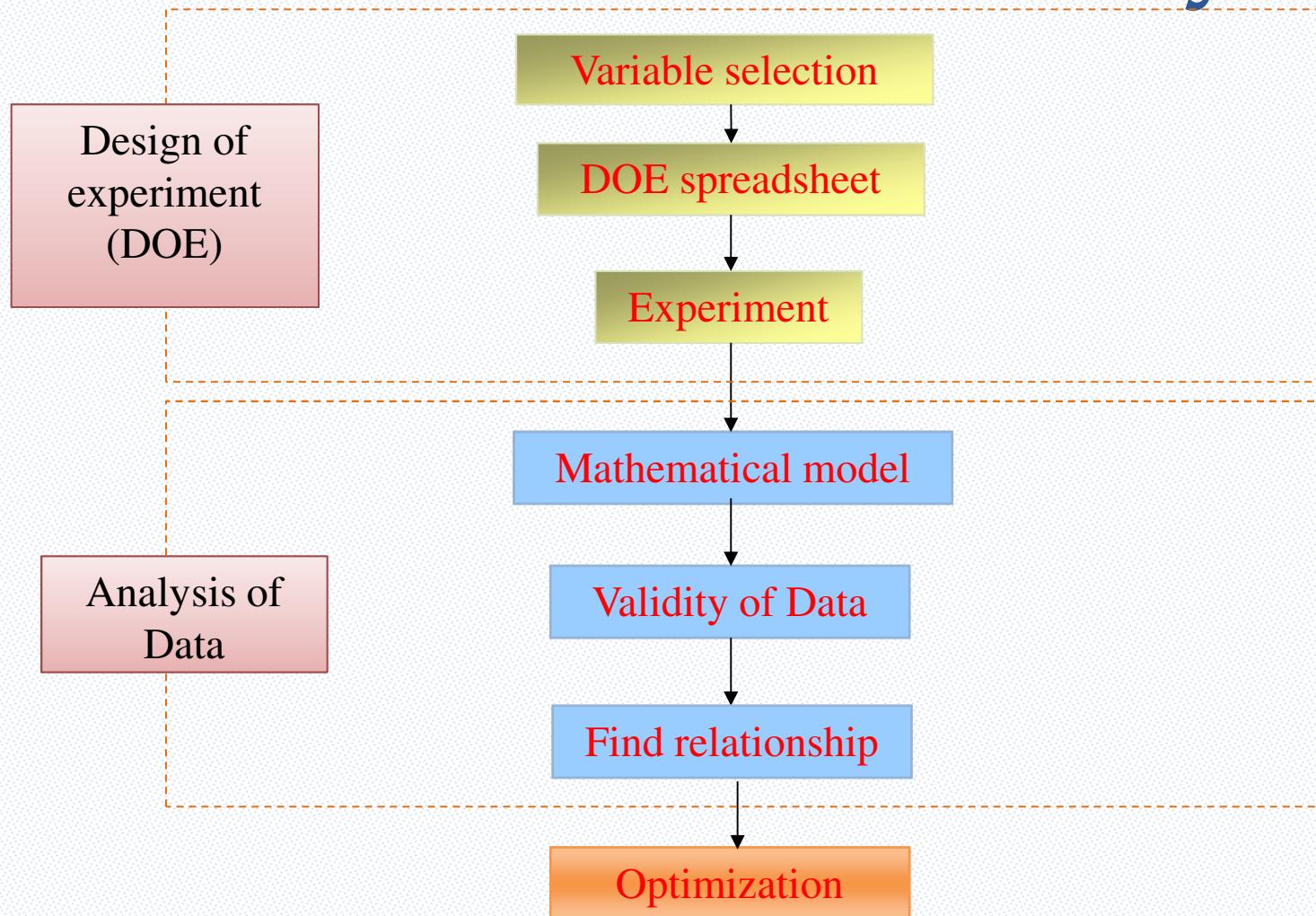
## PREDICTION

- Could predict the relationship/interaction between the values of some measurable response variable(s) and those of a set of experimental factors presumed to affect the response(s)
- Predict the response value at various process condition

## OPTIMIZATION

- Could find the values of the factors that produces the best value or values of the response(s).

# Flow of RSM study



# Step in RSM study

Before: Select the  
variable-Design the  
experiment



After: Analyze the  
data.

During: The actual  
experiment will be  
carried out

# Before Experiment

• • •

1. Selecting the process variables
2. Selecting the level and range for each process variable
3. Selecting the design of experiment (DOE)

# Preparing for RSM study

## RESPONSE

- What response variables are to be measured, how they will be measured, and in what sequence?

## PROCESS VARIABLE

- Which factor are most important and therefore will be included in the experiment, and which are least important and can these factors be omitted? With the important factors, can the desired effects be detected?

## DISTURBANCE

- What extraneous or disturbing factors must be controlled or at least have their effects minimized?

## EXPERIMENTAL UNIT

- What is the experimental unit, that is to say, what is the piece of experimental material from which a response value is measured? How are the experimental units to be replicated, if at all?

## DESICION

- The choice of the factors and level determined the type, size and experimental region. The no. of levels at each factor as well as the no. of replicated experiment units represent the total no. of experiments.

# Design of experiment (DOE) Process

## Objective

- Screening
- Prediction
- Optimization

academia@DahliaOmar

## Factor

- No of Independent Var.
- Block
- Level
- Range

## Type of design

- Full factorial
- Fractional factorial
- Placket Burman
- CCD
- Box-behnken
- Taguchi
- Etc.

# Objective

## Screening

- To identify significant main effect of factors from a list of many potential ones
- Not identified the interaction effect
- Type of design: 2-level with resolution III or IV, fractional factorial, Plackett-Burman

## Optimization

- To identify the best process performance, interaction effect, and significant of factors
- Type of design: CCD or BBD

# 1) Selecting the Parameter

## Factors:

- Process conditions influence the value of response variable
- Can be qualitative or quantitative
- Qualitative-**blocking** variables
- Quantitative –normally considered in RSM

## Responses:

- The measureable quantity whose value is assumed to be affected by changing the levels of the factors and most interested in optimizing.

## 2) Selecting the Level

Two level ( $2^k$ ) - (-1,+1)-first order,

- Two-level factorial design is each factor is evaluated at a “low” setting and at “high” setting.

Three level ( $3^k$ ) – (-1,0,+1) second or higher order

- Three-level factorial design is each factor is evaluated at a “low”, “center” and at “high” setting.

Five-level ( $5^k$ )-(- $\alpha$ , -1,0,+1, - $\alpha$ ) second or higher order

- Five-level factorial design is each factor is evaluated at a “Star low”, “low”, “center”, “high” and “star high” setting.

# Experimental region

- The region of conceivable factor level values that represents the factor combinations of potential interest.
- Need to determined before the experiment by finding the range of variables.
- If at the end of analysis, the factor value or optimum is out of the range, the experiment need to repeat with the new range.

Factors	Symbol	Range and Levels		
		-1	0	+1
Molar ratio methanol: oil	$X_1$	20:1	30:1	40:1
Catalyst loading, wt%	$X_2$	2	3	4
Reaction Time, min	$X_3$	120	180	240
Reaction Temperature	$X_4$	90	120	150

# 3) Selecting the Type of DOE

- Full factorial
- Fractional factorial
- Placket Burman
- CCD
- Box-behnken
- Taguchi
- Etc.

The most popular is CCD and box-behnken design

What are different between CCD and box-behnken design???

# Factorial design

Easy to be used by simply following relatively simple design

Able to meet the majority of the experimental needs and its data analysis can be performed by graphical methods

Require relatively few runs at a reasonable size

If large number of factors is selected, the fractional factorial design can be employed to keep the experimental run at a reasonable size

# Full factorial & fractional factorial

- Two level Full factorial  $(-1,+1) = 2^k$
  - Three level full factorial  $(-1,0,+1) = 3^k$
  - Fractional factorial (two level)  $= 2^{k-m}$ ,  $m < k$ 
    - $\frac{1}{2} = 2^{k-1}$
    - $\frac{1}{4} = 2^{k-2}$
    - $\frac{1}{8} = 2^{k-4}$
  - Fractional factorial (three level)  $= 3^{k-m}$ ,  $m < k$
- } Rotatable      } Orthogonal

# Matrix Arrangement (2-level)

2-level fractional factorial design  
(resolution IV)

	A	B	C	D
1	-1	-1	-1	-1
2	1	-1	-1	1
3	-1	1	-1	1
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

2-level factorial design (full)

	A	B	C	D
1	-1	-1	-1	-1
2	1	-1	-1	-1
3	-1	1	-1	-1
4	1	1	-1	-1
5	-1	-1	1	-1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	-1
9	-1	-1	-1	1
10	1	-1	-1	1
11	-1	1	-1	1
12	1	1	-1	1
13	-1	-1	1	1
14	1	-1	1	1
15	-1	1	1	1
16	1	1	1	1

# Matrix Arrangement (3-level)

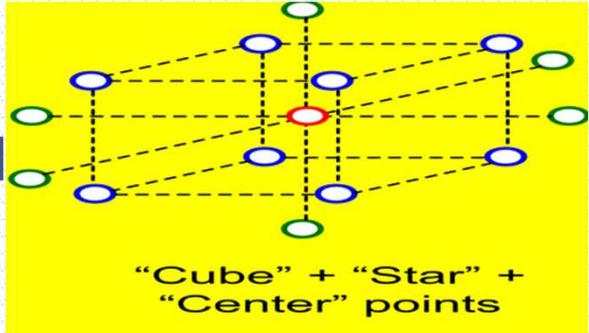
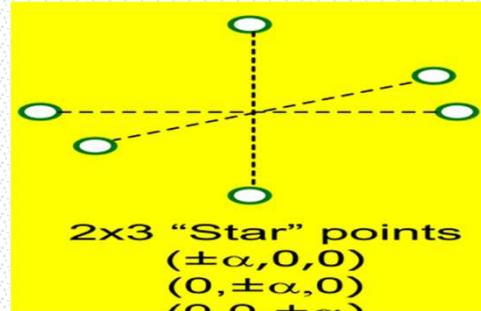
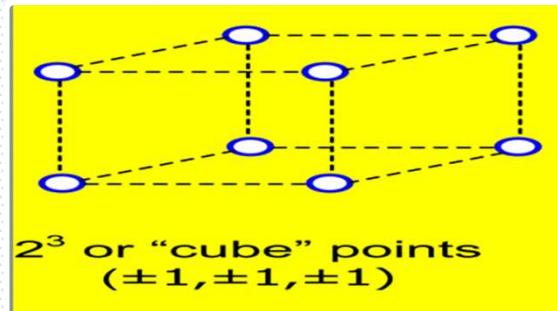
3** <sup>(4-0)</sup> full factorial design, 1 block , 81 runs (Spreadsheet1)				
	A	B	C	D
1	-1	-1	-1	-1
2	<b>-1</b>	<b>-1</b>	<b>-1</b>	0
3	-1	-1	-1	1
4	-1	-1	0	-1
5	-1	-1	0	0
6	-1	-1	0	1
7	-1	-1	1	-1
8	-1	-1	1	0
67	1	0	0	-1
68	1	0	0	0
69	1	0	0	1
70	1	0	1	-1
71	1	0	1	0
72	1	0	1	1
73	1	1	-1	-1
74	1	1	-1	0
75	1	1	-1	1
76	1	1	0	-1
77	1	1	0	0
78	1	1	0	1
79	1	1	1	-1
80	1	1	1	0
81	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

3** <sup>(4-1)</sup> fractional factorial design, 9 blocks, 27 runs (Spreadsheet1)					
	Bloc k	A	B	C	D
1	1	0	1	0	1
2	1	1	0	1	0
5	2	1	0	0	1
6	2	0	1	-1	-1
7	3	1	0	-1	-1
8	3	-1	-1	0	1
11	4	-1	0	-1	1
12	4	1	1	1	-1
13	5	-1	0	1	-1
14	5	1	1	0	0
15	5	0	-1	-1	1
16	6	1	1	-1	1
17	6	0	-1	1	-1
18	6	-1	0	0	0
19	7	-1	1	-1	0
20	7	0	0	0	-1
21	7	1	-1	1	1
22	8	0	0	-1	0
23	8	1	-1	0	-1
24	8	-1	1	1	1
25	9	0	0	1	1
26	9	1	-1	-1	0
27	9	-1	1	0	-1

3** <sup>(4-1)</sup> fractional factorial design, 3 blocks, 27 runs (Spreadsheet1)					
	Block	A	B	C	D
1	<b>1</b>	<b>1</b>	<b>0</b>	<b>-1</b>	<b>-1</b>
2	1	0	1	0	1
3	1	-1	-1	0	0
4	1	0	1	-1	1
5	-1	0	0	0	0
6	-1	0	1	-1	-1
7	-1	1	-1	0	0
8	-1	1	0	0	-1
9	-1	1	1	1	1
10	0	-1	-1	1	1
11	0	-1	0	0	0
12	0	-1	1	-1	-1
13	0	-1	1	1	-1
14	0	-1	0	0	0
15	2	0	-1	-1	1
16	2	1	1	0	0
17	2	-1	0	-1	1
18	2	1	1	1	-1
19	3	0	0	1	1
20	3	0	0	0	-1
21	3	-1	1	-1	0
22	3	0	0	-1	0
23	3	1	-1	1	1
24	3	1	-1	0	-1
25	3	1	-1	-1	0
26	3	-1	1	1	1
27	3	-1	1	0	-1

3** <sup>(4-1)</sup> fractional factorial design, 1 block , 27 runs (Spreadsheet1)				
	A	B	C	D
1	<b>-1</b>	<b>-1</b>	<b>-1</b>	<b>-1</b>
2	-1	-1	0	1
3	-1	-1	1	0
4	-1	0	-1	1
5	-1	0	0	0
6	-1	0	1	-1
7	-1	1	-1	0
8	-1	1	0	-1
9	-1	1	1	1
10	0	-1	-1	1
11	0	-1	0	0
12	0	-1	1	-1
13	0	-1	1	0
14	0	-1	0	0
15	2	0	-1	1
16	2	1	1	0
17	2	-1	0	-1
18	2	1	1	-1
19	1	-1	-1	0
20	1	-1	0	-1
21	1	-1	1	1
22	1	0	-1	-1
23	1	0	0	1
24	1	0	1	0
25	1	1	-1	1
26	1	1	0	0
27	1	1	1	-1

# Central Composite Design (CCD)

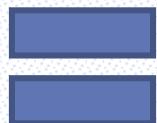


$2^k$  vertices of a  $k$ -dimensional "cube" ( $2$ -level full factorial design or  $2^{k-m}$  fractional design)  $\rightarrow$  coded as  $\pm 1$

$2k$  vertices of a  $k$ -dimensional "star"  $\rightarrow$  coded as  $\pm \alpha$

$n_0 \geq 1$  "center" point replicates  $\rightarrow$  coded as 0

Providing the estimate of pure error and curvature



Total run =  $2^k + 2k + n_0$  or  $2^{k-m} + 2k + n_0$

# CCD

The most common design (for the 2<sup>nd</sup> degree model)

Can be orthogonal or rotatable design

- orthogonal: the term of model have to redefined
  - : normally used if the blocking variable is considered.
- Rotatable: related to the precision of the predicted value
  - : archieved by selecting appropriate values for  $n_o$  ( $>0$ ) and  $a=4\sqrt{M}$ ,  $M=2^k$

# Box-behnken design

The equivalent in the case of  $3^{(k-p)}$  designs (3-level full factorial with incomplete block) are the so-called Box-Behnken designs (Box and Behnken, 1960).

These designs do not have simple design generators (they are constructed by combining two-level factorial designs with incomplete block designs), and have complex confounding of interaction.

However, the designs are economical and therefore particularly useful when it is expensive to perform the necessary experimental runs.

# DOE Matrix Arrangement

-1.00	-1.00	-1.00	-1.00
-1.00	-1.00	-1.00	1.00
-1.00	-1.00	1.00	-1.00
-1.00	-1.00	1.00	1.00
-1.00	1.00	-1.00	-1.00
-1.00	1.00	-1.00	1.00
-1.00	1.00	1.00	-1.00
-1.00	1.00	1.00	1.00
1.00	-1.00	-1.00	-1.00
1.00	-1.00	-1.00	1.00
1.00	-1.00	1.00	-1.00
1.00	-1.00	1.00	1.00
1.00	1.00	-1.00	-1.00
1.00	1.00	-1.00	1.00
1.00	1.00	1.00	-1.00
1.00	1.00	1.00	1.00
-2.00	0.00	0.00	0.00
2.00	0.00	0.00	0.00
0.00	-2.00	0.00	0.00
0.00	2.00	0.00	0.00
0.00	0.00	-2.00	0.00
0.00	0.00	2.00	0.00
0.00	0.00	0.00	-2.00
0.00	0.00	0.00	2.00
0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00

● academia@Dahlia@mar  
CCD, 26 run

-1.00	-1.00	-1.00	-1.00
-1.00	-1.00	0.00	1.00
-1.00	-1.00	1.00	0.00
-1.00	0.00	-1.00	1.00
-1.00	0.00	0.00	0.00
-1.00	0.00	1.00	-1.00
-1.00	1.00	-1.00	0.00
-1.00	1.00	0.00	-1.00
-1.00	1.00	1.00	1.00
0.00	-1.00	-1.00	1.00
0.00	-1.00	0.00	0.00
0.00	-1.00	1.00	-1.00
0.00	0.00	-1.00	0.00
0.00	0.00	0.00	-1.00
0.00	0.00	1.00	1.00
0.00	1.00	-1.00	-1.00
0.00	1.00	0.00	1.00
0.00	1.00	1.00	-1.00
1.00	-1.00	-1.00	0.00
1.00	-1.00	0.00	-1.00
1.00	-1.00	1.00	1.00
1.00	0.00	-1.00	-1.00
1.00	0.00	0.00	1.00
1.00	0.00	1.00	0.00
1.00	1.00	-1.00	1.00
1.00	1.00	0.00	0.00
1.00	1.00	1.00	-1.00

3 fractional factorial, 27 run

-1.00	-1.00	0.00	0.00
1.00	-1.00	0.00	0.00
-1.00	1.00	0.00	0.00
1.00	1.00	0.00	0.00
0.00	0.00	-1.00	-1.00
0.00	0.00	1.00	-1.00
0.00	0.00	-1.00	1.00
0.00	0.00	1.00	1.00
0.00	0.00	0.00	0.00
-1.00	0.00	0.00	-1.00
1.00	0.00	0.00	-1.00
-1.00	0.00	0.00	1.00
1.00	0.00	0.00	1.00
0.00	-1.00	0.00	0.00
0.00	1.00	0.00	0.00
0.00	-1.00	0.00	-1.00
0.00	1.00	0.00	-1.00
0.00	0.00	1.00	0.00
0.00	0.00	1.00	-1.00
0.00	0.00	-1.00	0.00
-1.00	0.00	1.00	0.00
1.00	0.00	1.00	0.00
0.00	0.00	0.00	0.00
-1.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00
-1.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00

12/8/2015 ● 25  
BBD, 27 run

# CCD vs BBD

Criteria	CCD	BBD
Design	2-level factorial, with star point	3-level fractional factorial,
Block	Up to researcher	Limited
Mean effect	Not considered	Considered
Interaction	Linear	Linear, quadratic
Optimization	Yes	Yes

# DOE Table

Run s	Manipulated Variables						Response s
	$X_1$	$X_2$	$X_3$	Molar Ratio (meOH: oil)	Reaction time,t (h)	Level	
	Operating temperature,T(°C)	Level					Yield, $Y_1$ (%)
1	50	-1	3	-1	2	-1	91.90
2	50	-1	3	-1	4	+1	84.60
3	50	-1	10	+1	2	-1	65.15
4	50	-1	10	+1	4	+1	95.95
5	70	+1	3	-1	2	-1	63.90
6	70	+1	3	-1	4	+1	94.95
7	70	+1	10	+1	2	-1	87.60

# Addition note

## Randomisation

- The order of run is random
- Can protect us from bias caused by unaware factors, and validates our analysis based on normal mode assumptions

## Replication

- Repetition of experiments
- It is essential feature to increase the degree of belief

## Blocking

- The experimental units are grouped into homogeneous clusterd in an attempt to improve the comparison of treatments with greater precision by randomly allocating the treatments withing each cluster or 'block'
- To detect the effect of treatment from background noise caused by non-homogeneous experimental unit.
- Can eliminate a source of variability from analysis

# REMEmBERS

There are none  
software that can help  
you IF your DOE is worst  
or wrong.

1. Insert the complete data into DOE
2. Develop the empirical/predicted model
  3. Statistic analysis of empirical model
4. Find the importance of process variables
5. Investigate the influence of process variables
  6. Optimization of process variables

• • •

## After experiment

# 1) Complete data

• • •

Insert the collected data into the software.  
(refer to tutorial 2)

- First Order polynomial
- Second Order polynomial

• • •

## 2) Develop the model

# Mathematical Model (empirical/predicted model)

To

- represent the relationship of response function and the factor level
- Predict the response at various combination of process variables

Can be shown by

- First order model (linear,  $X$ )
- Second order model (Quadratic,  $X^2$ )
- Third model (Cubic,  $X^3$ ), *if using design expert.*

Analysis

- the least square method was employed to estimate the response surface model.

Test of significance of model

- t and F-test
- Parity plot
- Pareto chart
- Probability plot

Random Error (Pure error)- are normal distributed

# First Order Model

- Fit for
  - Limited for small experimental region (two level)
  - Response surface is hyperlane
  - First approximation of the surface
  - Cost of experimentation are held to a minimum
  - To locate higher value of the response-steepest ascent
  - Screening for the important factor

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

**Y**

: predicted response (response function)

**$\beta_0$**

: intercept coefficient (offset)

**$\beta_1$ ,  $\beta_2$  and  $\beta_3$**

: linear terms (first order)

**$X_1$ ,  $X_2$  and  $X_3$**

: uncoded independent variables

# Second order model (polynomial)

- Normally used for optimization since it is consider the center point.

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 \\ + \beta_{23} X_2 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2$$

$Y$	: predicted response (response function)
$\beta_0$	: intercept coefficient (offset)
$\beta_1, \beta_2$ and $\beta_3$	: linear terms (first order)
$\beta_{11}, \beta_{22}$ and $\beta_{33}$	: quadratic terms (second order)
$\beta_{12}, \beta_{13}$ and $\beta_{23}$	: interaction terms
$X_1, X_2$ and $X_3$	: uncoded independent variables

- Regression analysis
  - ANOVA
  - Hypothesis testing

• • •

### 3) Statistic analysis of model

# Validity of model

The adequacy of the fitted model is checked by *ANOVA (Analysis of Variance) using Fisher F-test*

The fit quality of the model can also be checked from their *Coefficient of Correlation (R)* and *Coefficient of Determination (R<sup>2</sup>)*

# Observed and predicted table

$x_1$	$x_2$	$Y_u$	$\hat{Y}_u$	$Y_u - \bar{Y}$	$\hat{Y}_u - \bar{Y}$	$Y_u - \hat{Y}_u$
-1	-1	7.0	6.8375	0.6625	0.5000	0.1625
-1	-1	6.9	6.8375	0.5625	0.5000	0.0625
-1	+1	5.2	5.4125	-1.1375	-0.9250	-0.2125
-1	+1	5.4	5.4125	-0.9375	-0.9250	-0.0125
+1	-1	7.1	7.2625	0.7625	0.9250	-0.1625
+1	-1	7.2	7.2625	0.8625	0.9250	-0.0625
+1	+1	6.1	5.8375	-0.2375	-0.5000	0.2625
+1	+1	5.8	5.8375	-0.5375	-0.5000	-0.0375
		$\Sigma Y_u = 50.7$		$\Sigma(Y_u - \bar{Y})^2 = 4.59875$	$\Sigma(\hat{Y}_u - \bar{Y})^2 = 4.4225$	$\Sigma(Y_u - \hat{Y}_u)^2 = 0.17625$
		$\bar{Y} = 6.3375$				

$$SSR = \sum (\hat{Y}_u - \bar{Y})^2 = 4.4225$$

$$SST = \sum (Y_u - \bar{Y})^2 = 4.59875$$

academia@DahliaOmar

$$SST = \sum (Y_u - \hat{Y})^2 = 0.17625$$

12/8/2015 38

# Coefficient of Determination (R-square, $R^2$ )

Coefficient of Determination ( $R^2$ ): a proportion of total variation of the observed values of activity ( $Y_i$ ) about the mean explained by the fitted model

- $R^2 = \text{SSR}/\text{SST}$

Coefficient of Correlation ( $R$ ) : an acceptability about the correlation between the experimental and predicted values from the model.

Adjusted  $R^2$ : Measure the drop of magnitude of the estimate of the error variance

- adj  $R^2 = 1 - \text{Msresidual}/(\text{SST}/N)$ -more smaller more better

# How to interpret the $R^2$ ?

R<sup>2</sup> value is always in between 0 to 1

The value of 1, indicated the empirical/predicted model explains all of the variability in the data

The value of 0, indicated that none of the variability in the data can be explained by predicted model.

$R^2$  is closer to 1, the predicted model is more reliable

$R^2 > 0.75$  acceptable (Haaland), however,  $> 0.8$  is much better

# Analysis of variance (ANOVA)

- The  $F$ -value is a measurement of variance of data about the mean based on the ratio of mean square (MS) of group variance due to error.
- $F\text{-value} = \frac{MS_{\text{regression}}/MS_{\text{residual}}}{(SSE/DF_{\text{residual}})} = \frac{(SSR/DF_{\text{regression}})}{(SSE/DF_{\text{residual}})}$
- F table =  $F(p-1, N-p, a)$ 
  - $p-1 : DF_{\text{regression}}$
  - $N-p : DF_{\text{residual}}$
  - $N = \text{total exp}$
  - $P = \text{no of term in fitted model}$
  - $a = \text{value: level of significant}$
- the calculated  $F$ -value should be greater than the tabulated  $F$ -value to reject the null hypothesis,

*where:*

*SSR = sum of square of regression*

*SSE = sum of square of error/residual*

*DF<sub>regression</sub> = degree of freedom of regression*

*DF<sub>residual</sub> = degree of freedom of residual*

# Hypothesis testing (F value)

- There are 2 statement is comparing at **significant confident level (95%,  $\alpha= 0.05$ )**

F table Can be  
find online

Null hypothesis,  $H_0$ : All the coefficient ( $\beta$ ) are zero

Alternative hypothesis,  $H_1$ : At least one of coefficient ( $\beta$ ) is not zero.

The surface is plane

Conclusion: The null is

True:  $F_{cal} < F_{table}$ , cannot be rejected

OR

Rejected:  $F_{cal} > F_{table}$ .

The surface is  
twisted.

- T-value and p-value
  - Pareto chart
  - Probability plot

• • •

## 4) Importance/significant of process variables

# Significant of the model coefficient

## T-Value:

- Measure how large the coefficient is in relationship to its standard error
- $T\text{-value} = \text{coefficient} / \text{standard error}$

## P-value

- is an observed significance level of the hypothesis test or the probability of observing an F-statistic as large or larger than one we observed.
- The small values of p-value → the null hypothesis is not true.

Can be visualized

- Pareto Chart
- Normal Probability plot

# Interpretation?

- If a  $p$ -value is  $\leq 0.01$ , then the  $H_0$  can be rejected at a 1% significance level → “*convincing*” evidence that the  $H_A$  is true.
- If a  $p$ -value is  $0.01 < p\text{-value} \leq 0.05$ , then the  $H_0$  can be rejected at a 5% significance level → “*strong*” evidence in favor of the  $H_A$ .
- If a  $p$ -value is  $0.05 < p\text{-value} \leq 0.10$ , then the  $H_0$  can be rejected at a 10% significance level. → it is in a “*gray area/moderate*”
- If a  $p$ -value is  $> 0.10$ , then the  $H_0$  cannot be rejected. → “*weak*” or “*no*” evidence in support of the  $H_A$ .

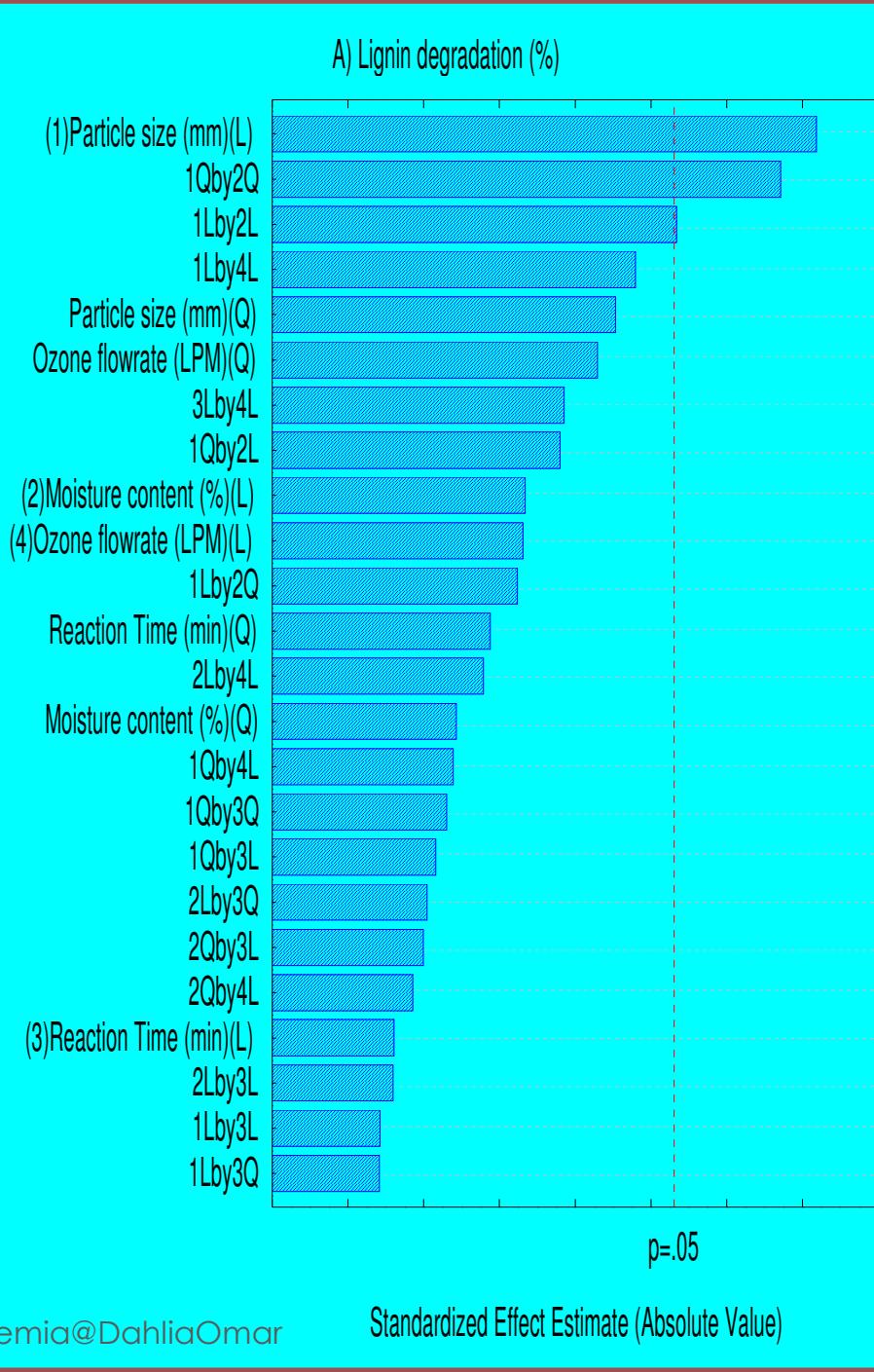
Design expert:P-value sometimes known as  $P_{prob}$

• academia@DahliaOmar

Statistica: visualize using pareto chart

12/8/2015 • 45

# Pareto Chart



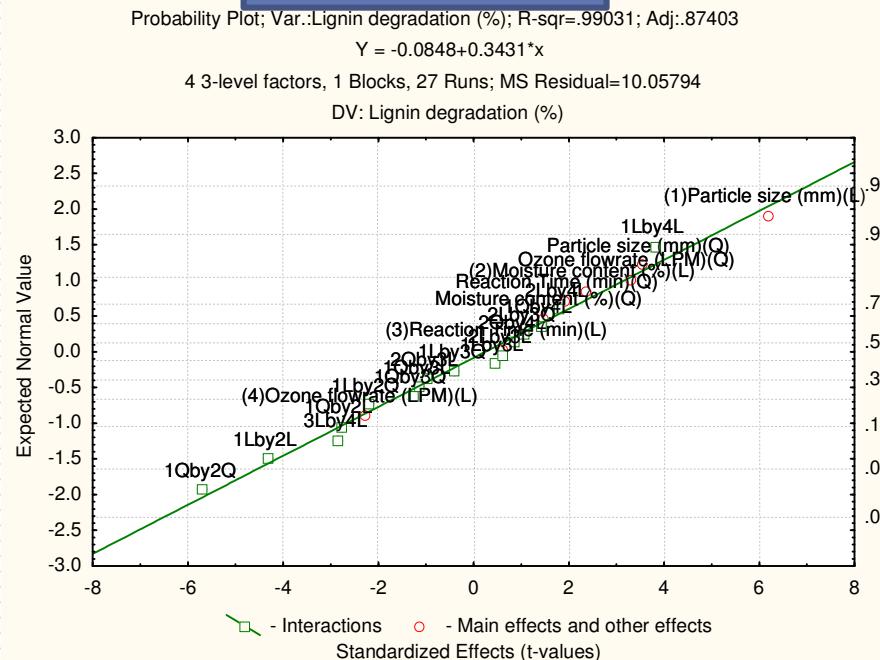
ANOVA effect estimates are sorted from largest to small value

The magnitude of each effect is represented by a column

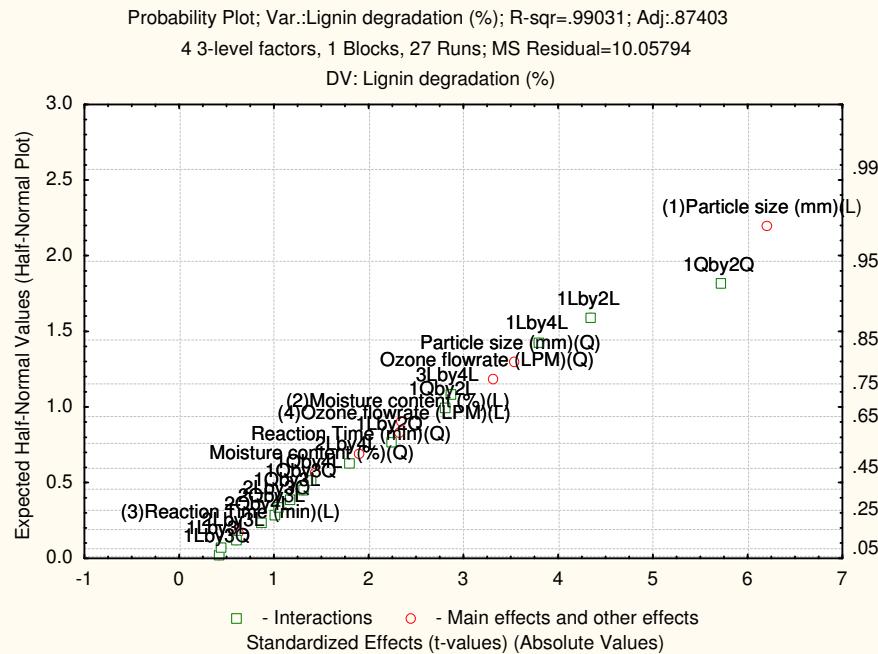
A line going across the column indicates how large and effect has to be statistically significant

# Probability plot

Normal



Half- Normal



- To assess how closely a set of observed values follow a theoretical distribution
- if all values fall onto straight line, the residual follow the normal distribution
  - The parameter were rank –ordered.

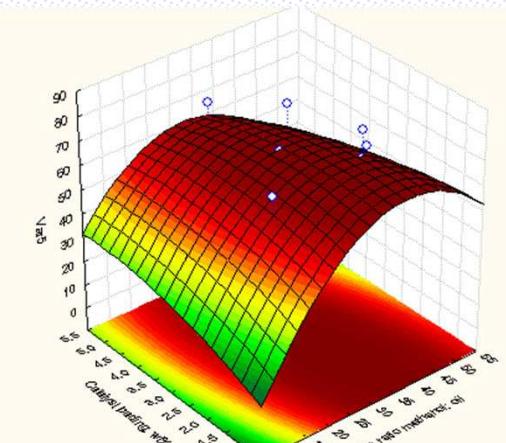
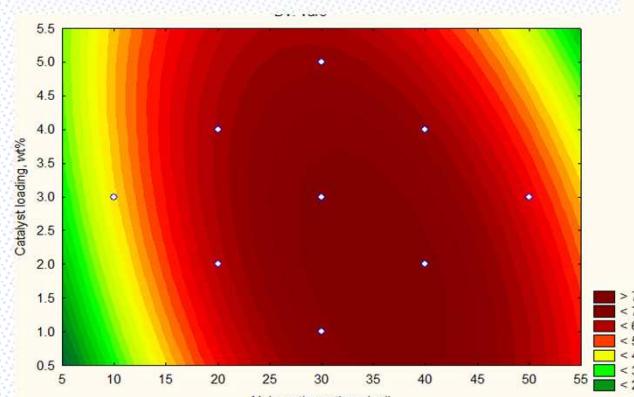
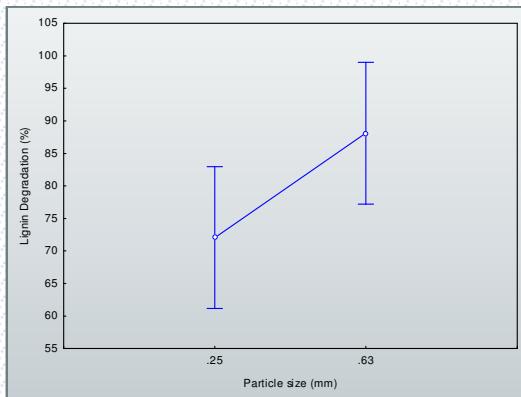
- Contour plot
- Surface (3D)
- Single parameter

• • •

## 5) Interaction/influence of process variables

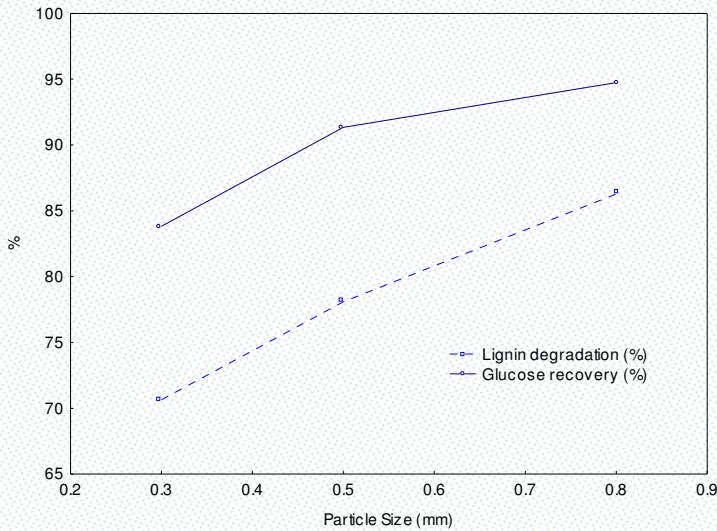
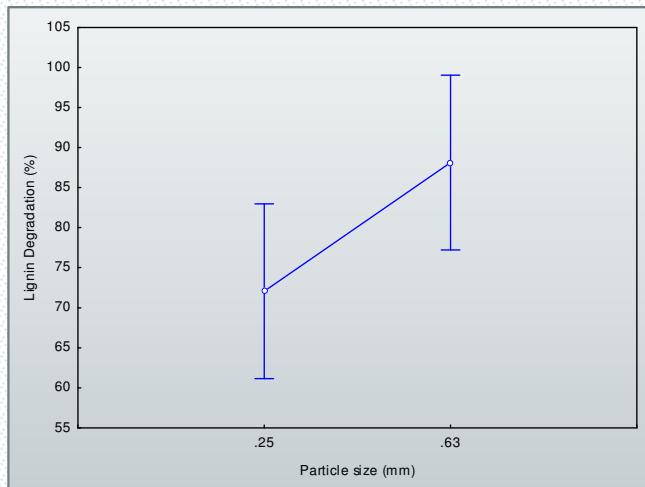
# Visualize the result

- Predicted response function ( $\hat{Y}$ ) (read “Y hat”):
  - Predict the value of response
- Response surface:
  - Represent the relationship between predicted response function and factor
  - Is visualized in 3D, contour, single parameter;



# Single Variables

- The graph is plot the predicted Mean of value of process variables



# Contour

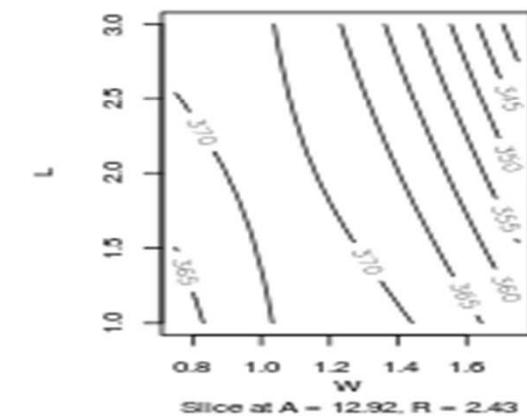
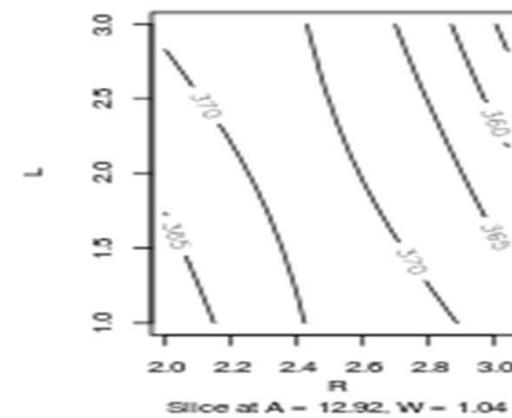
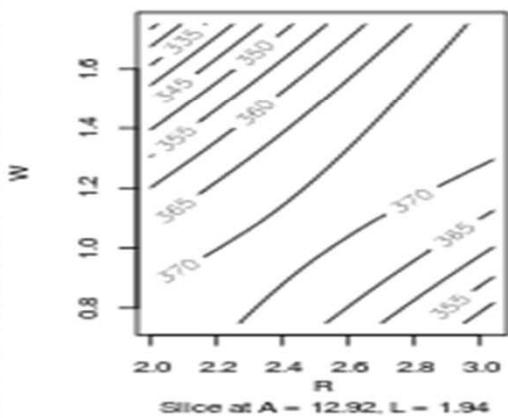
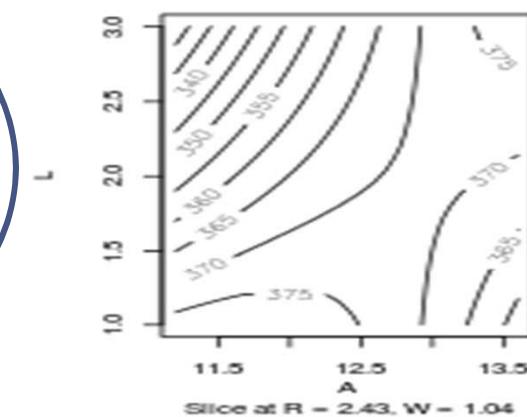
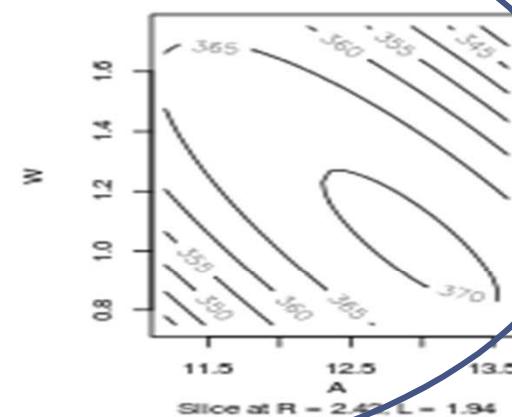
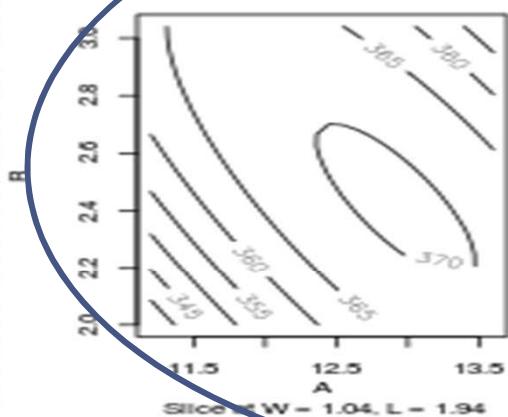
Visualized the shape of the 3D response surface

Line or curves (known as contour) represent the surface of response value are drawn on graph or plane whose coordinates represent the level of the factor.

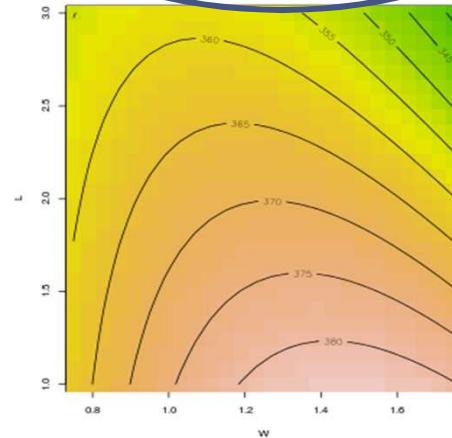
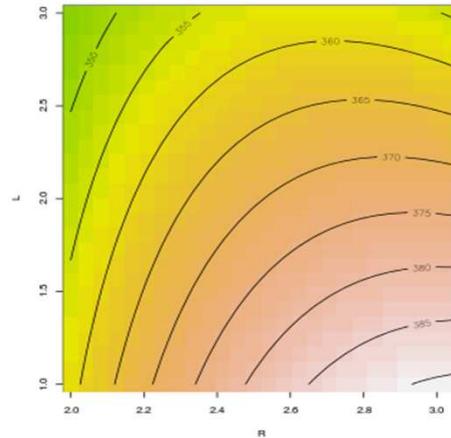
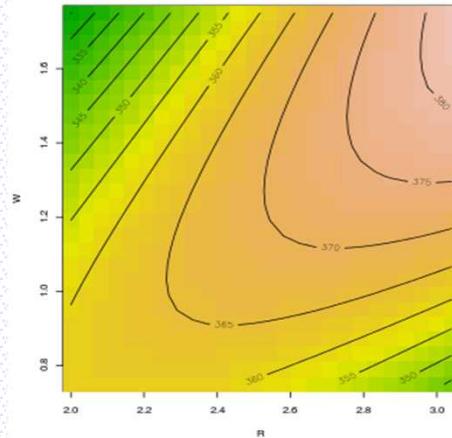
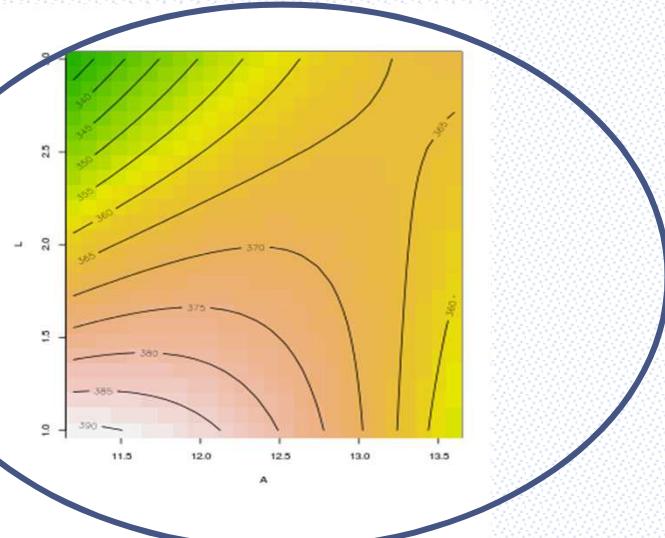
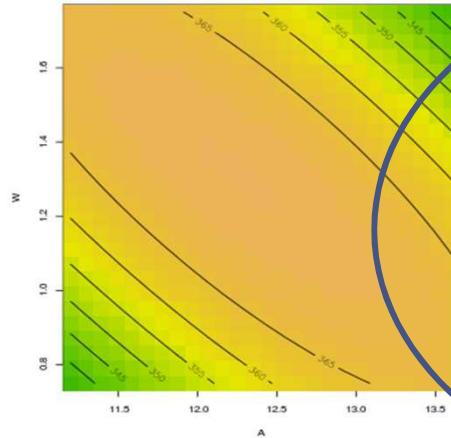
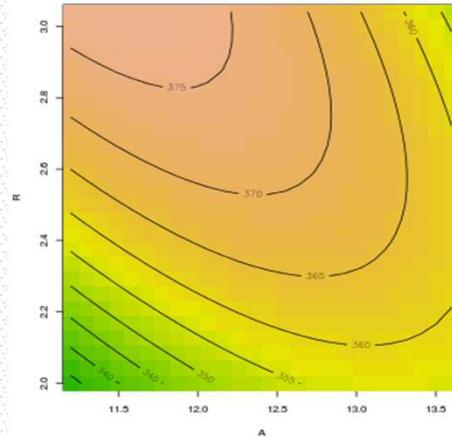
The direction of contour can be used to explained the behavior of interaction for both parameter

Ellipses, circular or saddle point

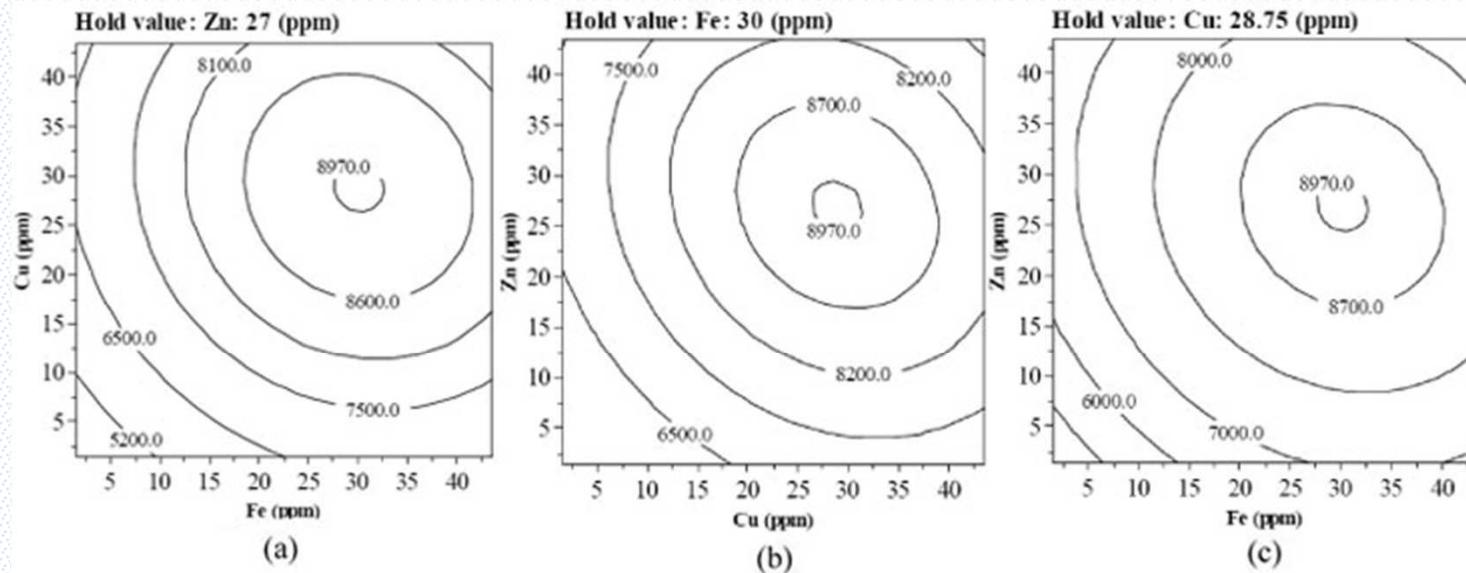
# Examples (Ellipse)



# Examples (saddle point)



# Example (Circular)



**Figure 3:** Contour plots (two-dimensional surface plots) of the model equation fitted to the data. (a) interaction of Fe<sup>3+</sup> and Cu<sup>2+</sup> concentration, (b) interaction of Fe<sup>3+</sup> and Zn<sup>2+</sup> concentration, (c) interaction of Cu<sup>2+</sup> and Zn<sup>2+</sup> concentration. (Cu<sup>2+</sup> (copper, ppm); Fe<sup>3+</sup> (iron, ppm); Zn<sup>2+</sup> (zinc, ppm))

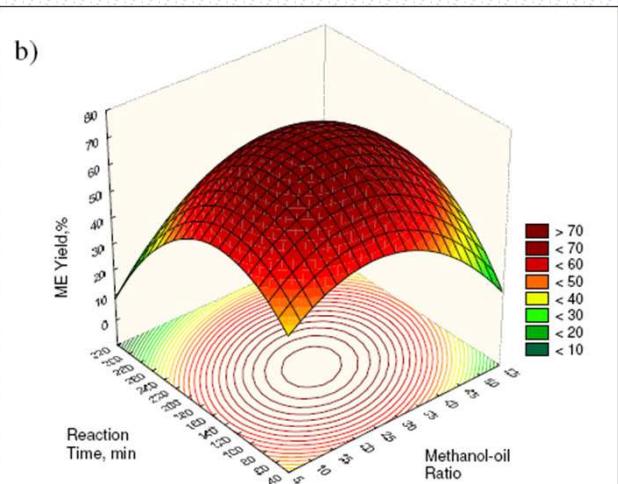
# 3D surface plot

Shows the interaction between two process variables as function of factors.

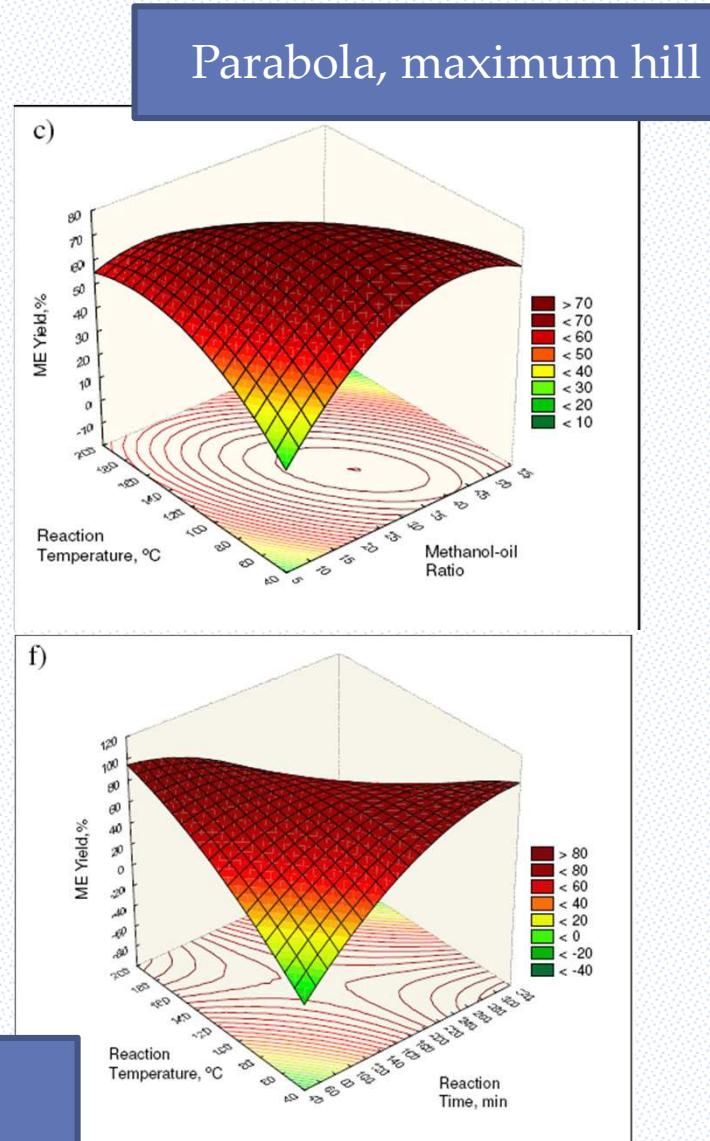
## Shape

- Minimum: basin
- Maximum: hill
- Saddle: saddle shape

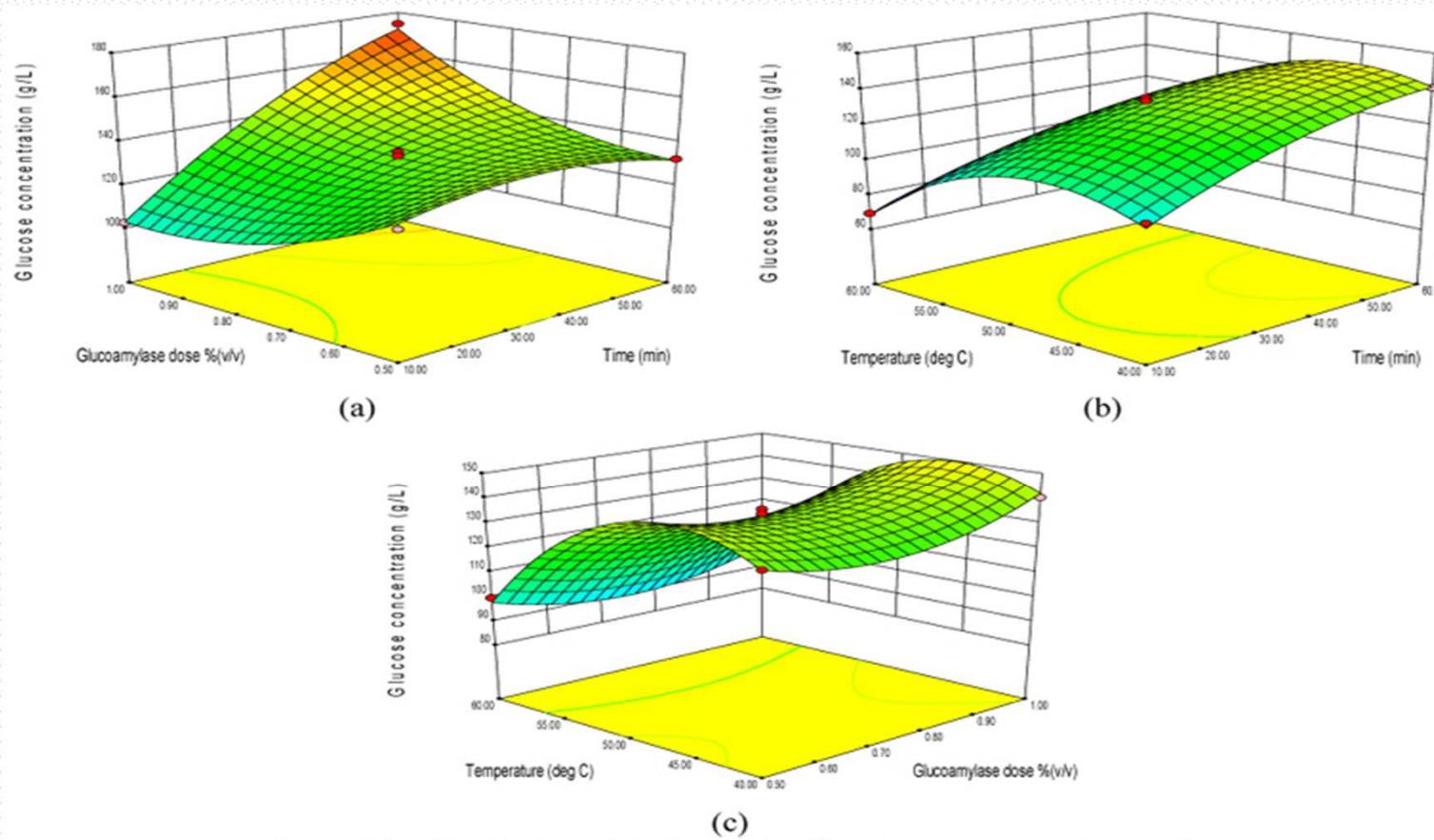
# Surface



Hyperbola, maximum hill

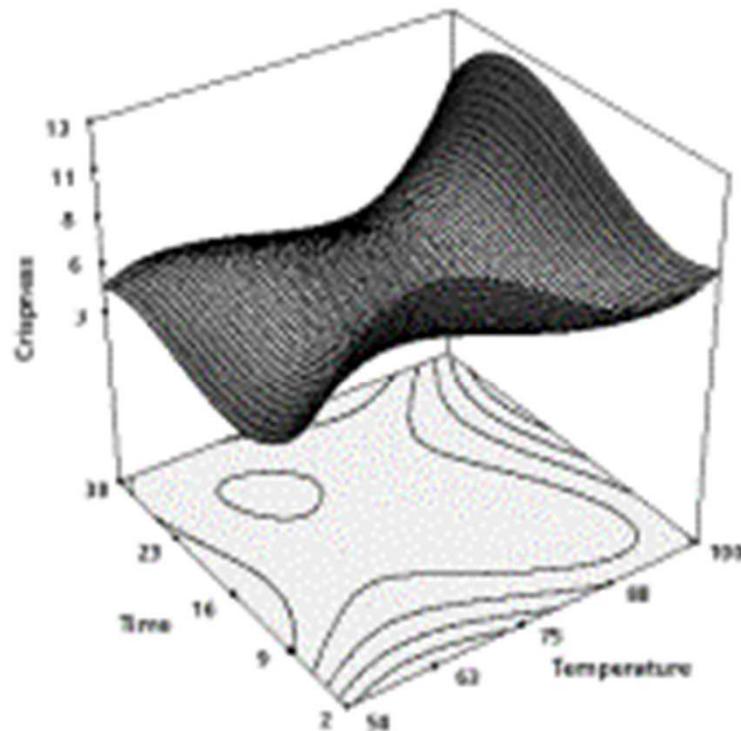


Saddle-shaped

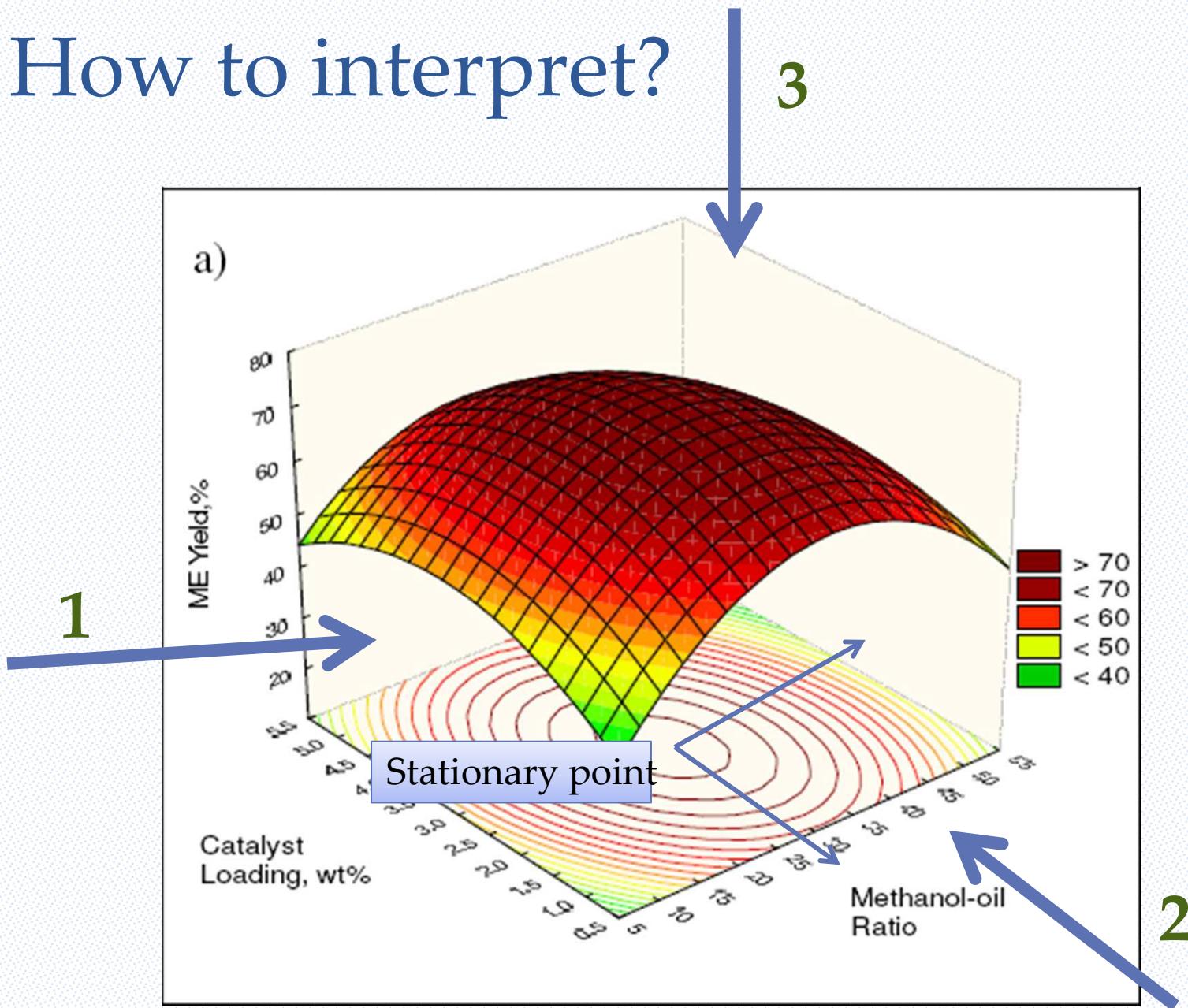


**Figure 2(a)-(c):** Surface plots for saccharification of sweet potato peel.  
 (a) effect of glucoamylase dose, time and their reciprocal interaction on glucose concentration.  
 (b) effect of temperature, time and their reciprocal interaction on glucose concentration.  
 (c) effect of temperature, glucoamylase dose and their reciprocal interaction on glucose concentration.

# Quadratic interaction



# How to interpret?



# 6) Optimization

• • •

- Single response
- Multi response

# Optimization: Single response-Critical value

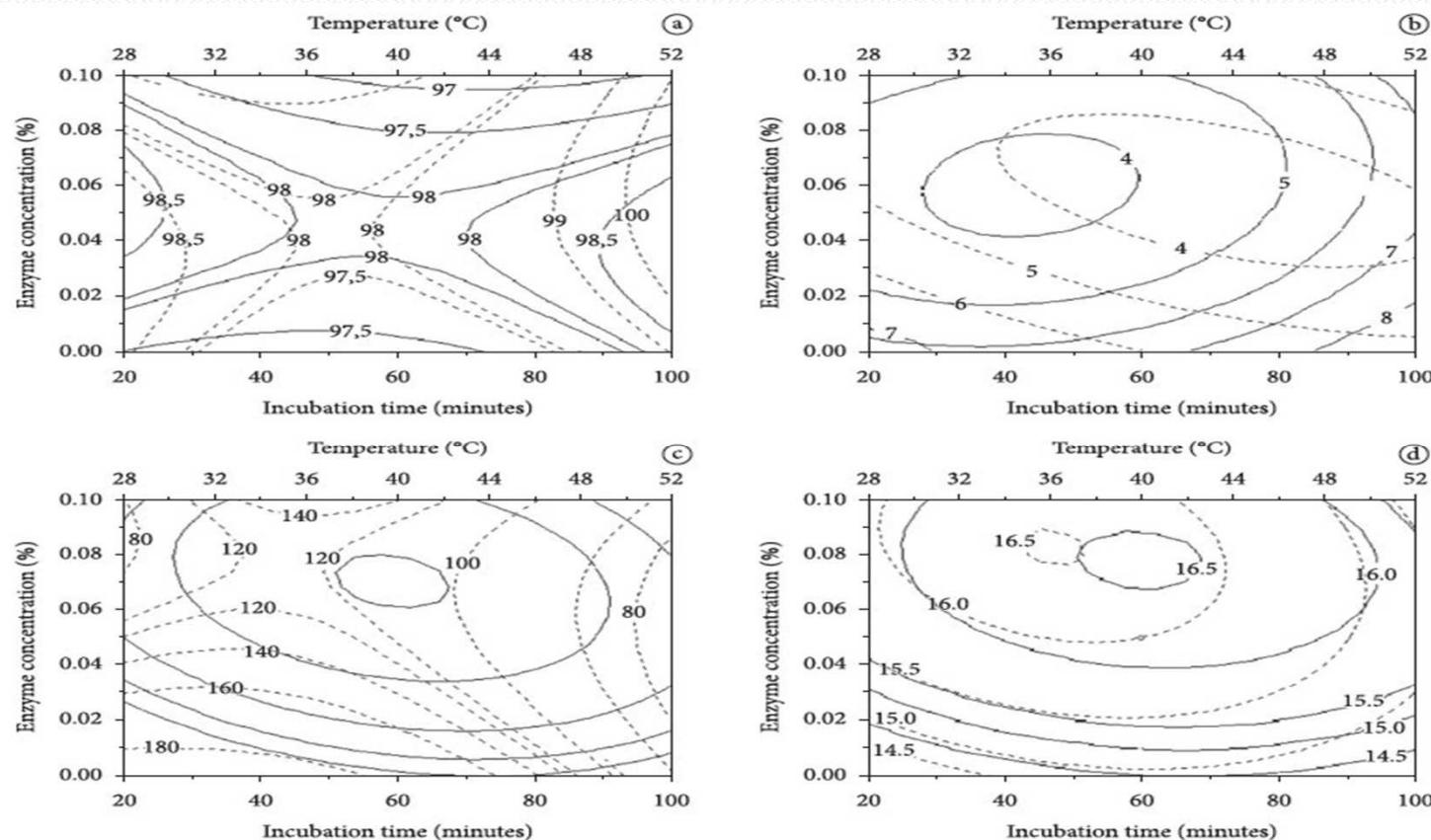
Will identified the point on the quadratic response surface either it the minimum, maximum, or saddle point of the surface.

The critical values for the independent variables are the coordinates of the origin of the quadratic response surface.

Shown the predicted value of the dependent variable (response) at the critical values for each of the independent variables.

# OPTIMIZATION: Multi-response

- Superimpose of two contour plot.



● academica@DanillaOMar  
Figure 6. Optimum regions for the contour plots (a-  $L^*$  value; b- viscosity; c- turbidity; d- yield, ---- time, ——temperature).

# OPTIMIZATION: Multi-response via Desirability Function

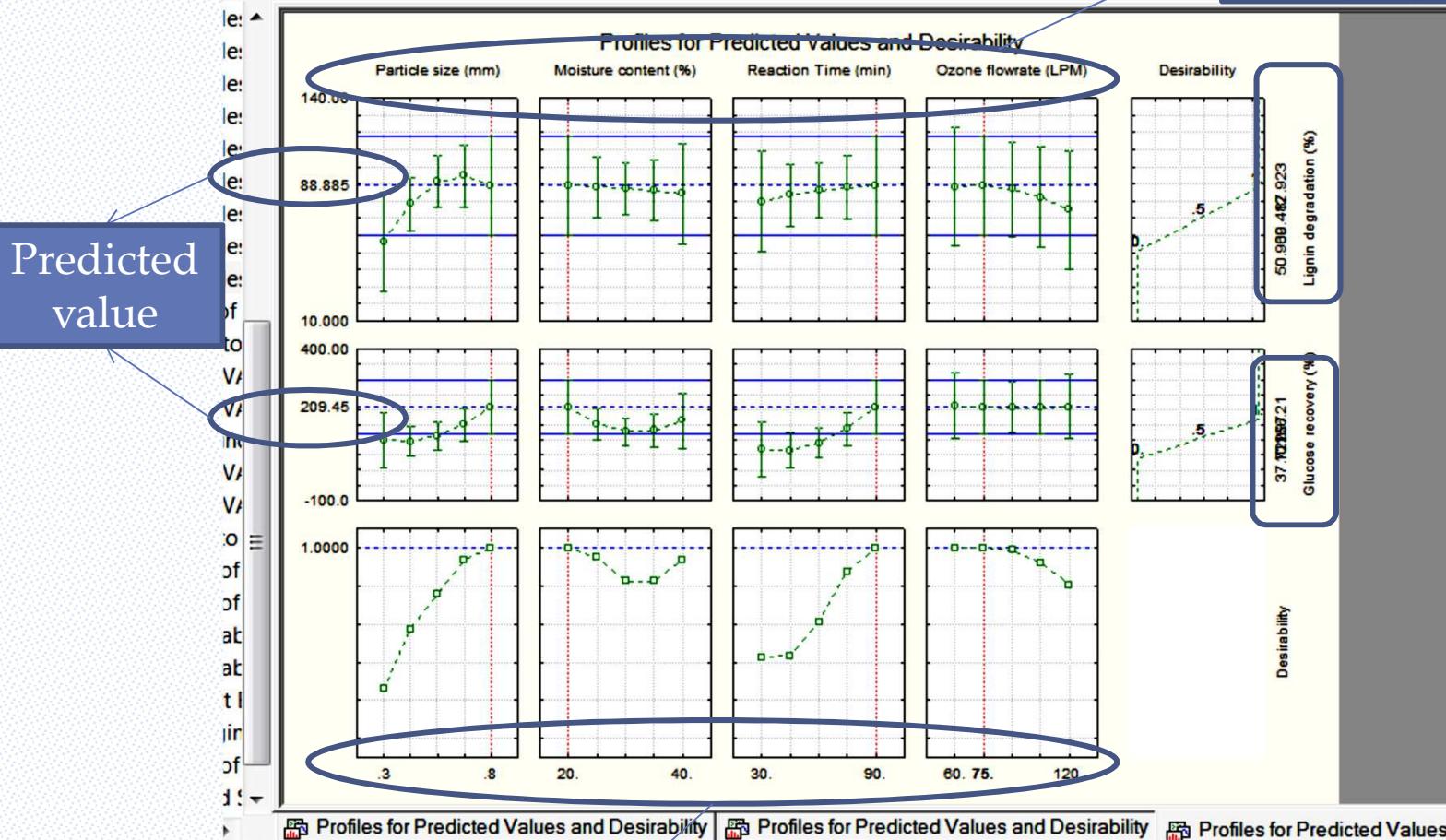
A popular and established technique for simultaneous determinization of optimum settings of input variables that can determine optimum performance levels for one or more responses

Converting the estimated response model ( Y) into individual desirability function (d) that are then aggregated into a composite function (D).

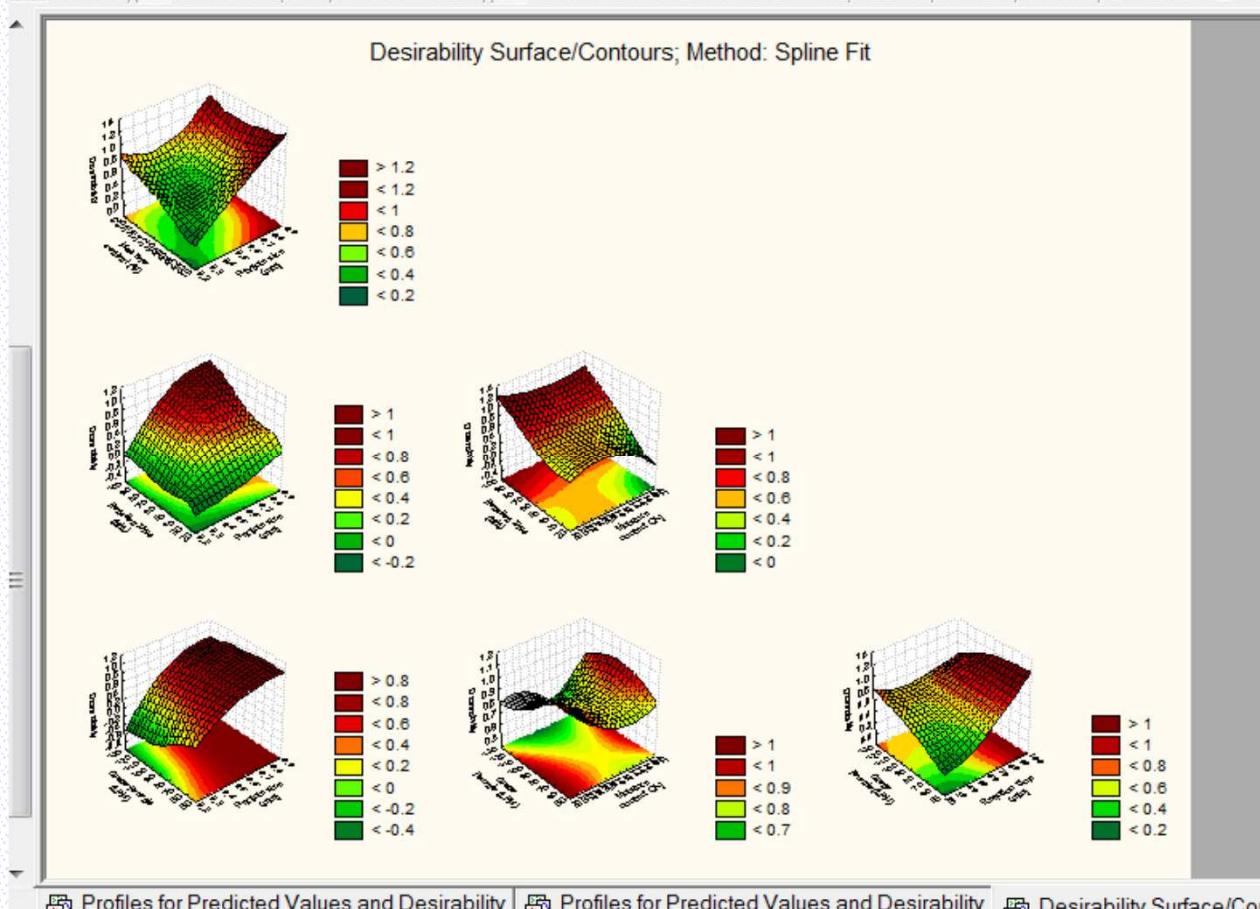
This composite function is usually a geometric or an arithmetic , which will be maximized or minimized, respectively.

# Desirability Profile

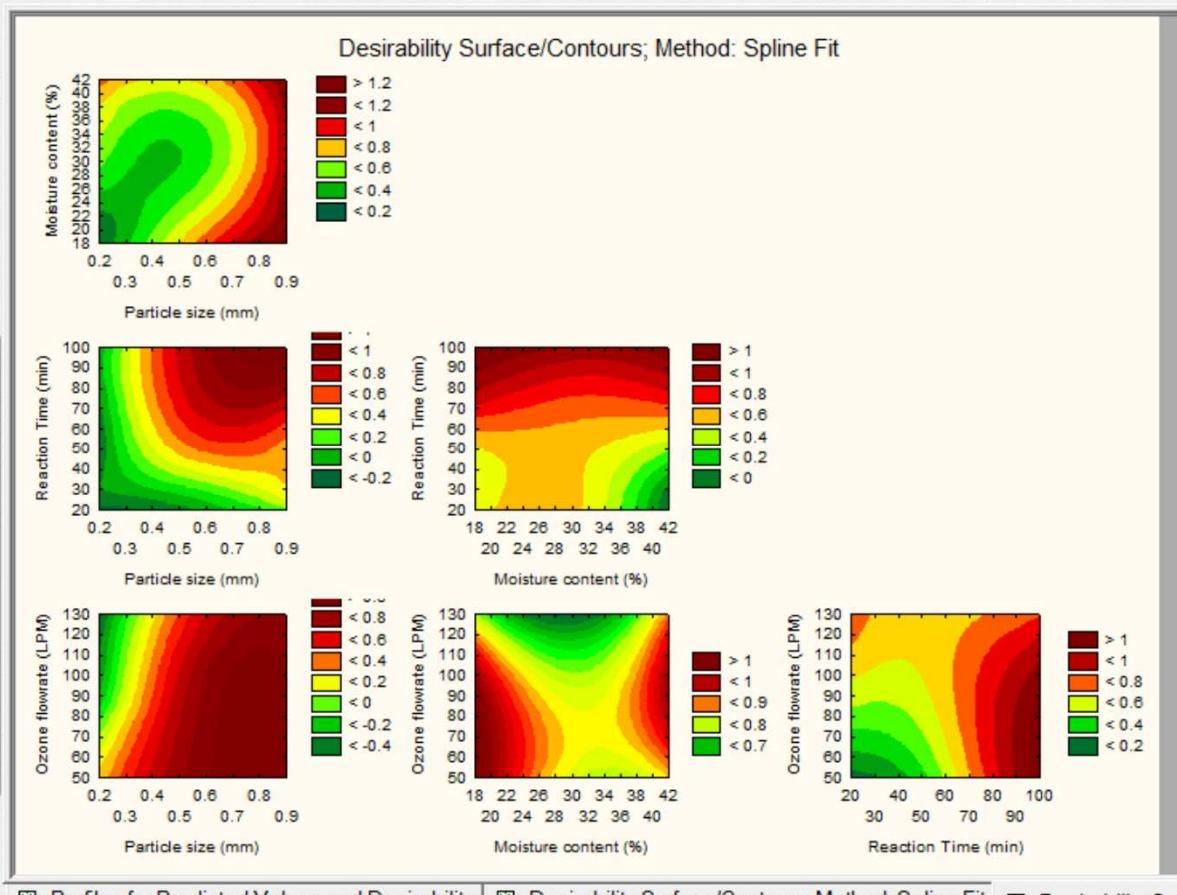
Process variable



# 3D surface plot



# Contour plot



# Conclusion

## DOE and RSM

- A powerful method for design of experimentation, analysis of experimental data, and optimization.

## Advantages

- design of experiment, statistical analysis, optimization, and profile of analysis in one step
- Produce empirical mathematical model

## Disadvantage

- The prediction only can be determined in range of study.

# references

- Montgomery, D. C. 1997. *Design and Analysis of Experiment*. Fifth Edition. Wiley, Inc., New York, USA.
- Brown, S. R. and Melemend, L. E. 1990. *Experimental Design and Analysis Quantitative Application in the Social Science*. Sage Publication, California. 74.
- Cornell J.A. 1990. *How to Apply Response Surface Methodology*. America Society For Quality Control: Statistic Devision . US.
- Haaland, P. D. 1989. *Experimental Design in Biotechnology*. Marcel Dekker Inc., New York.

# Slide can be found at

• • •

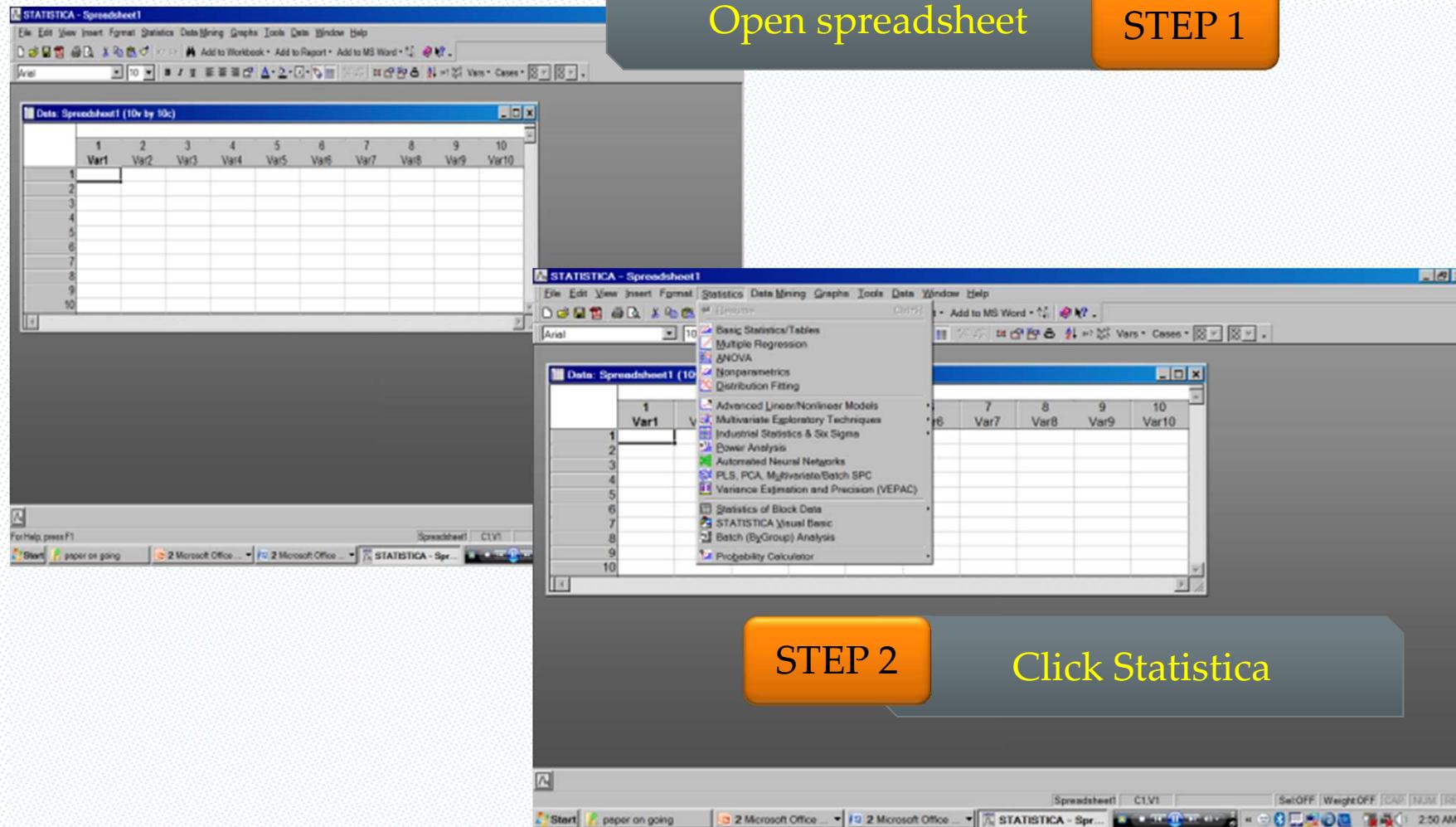
<https://teknologimalaysia.academia.edu/DahliaOmar>

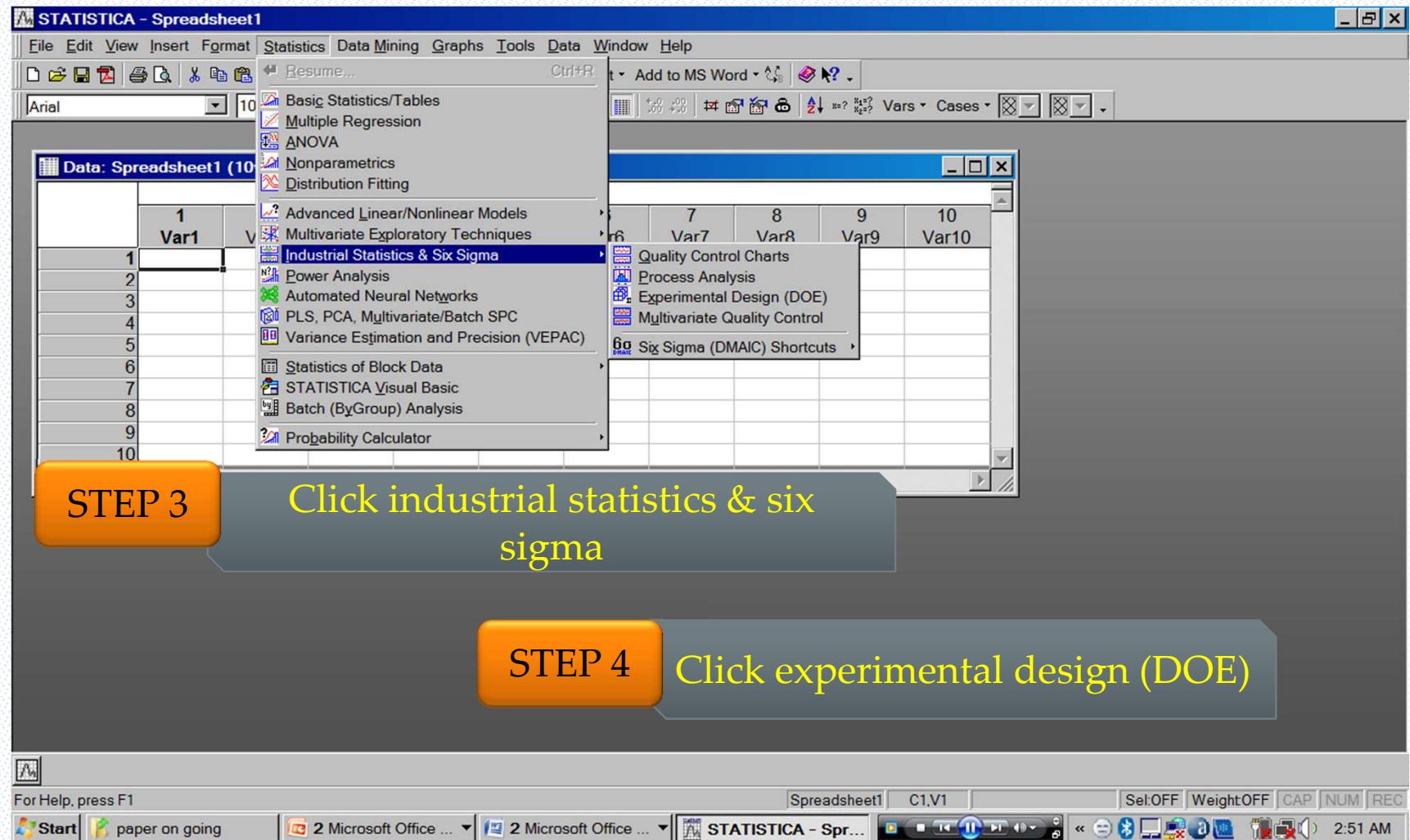
# Statistica Tutorial 1

• • •

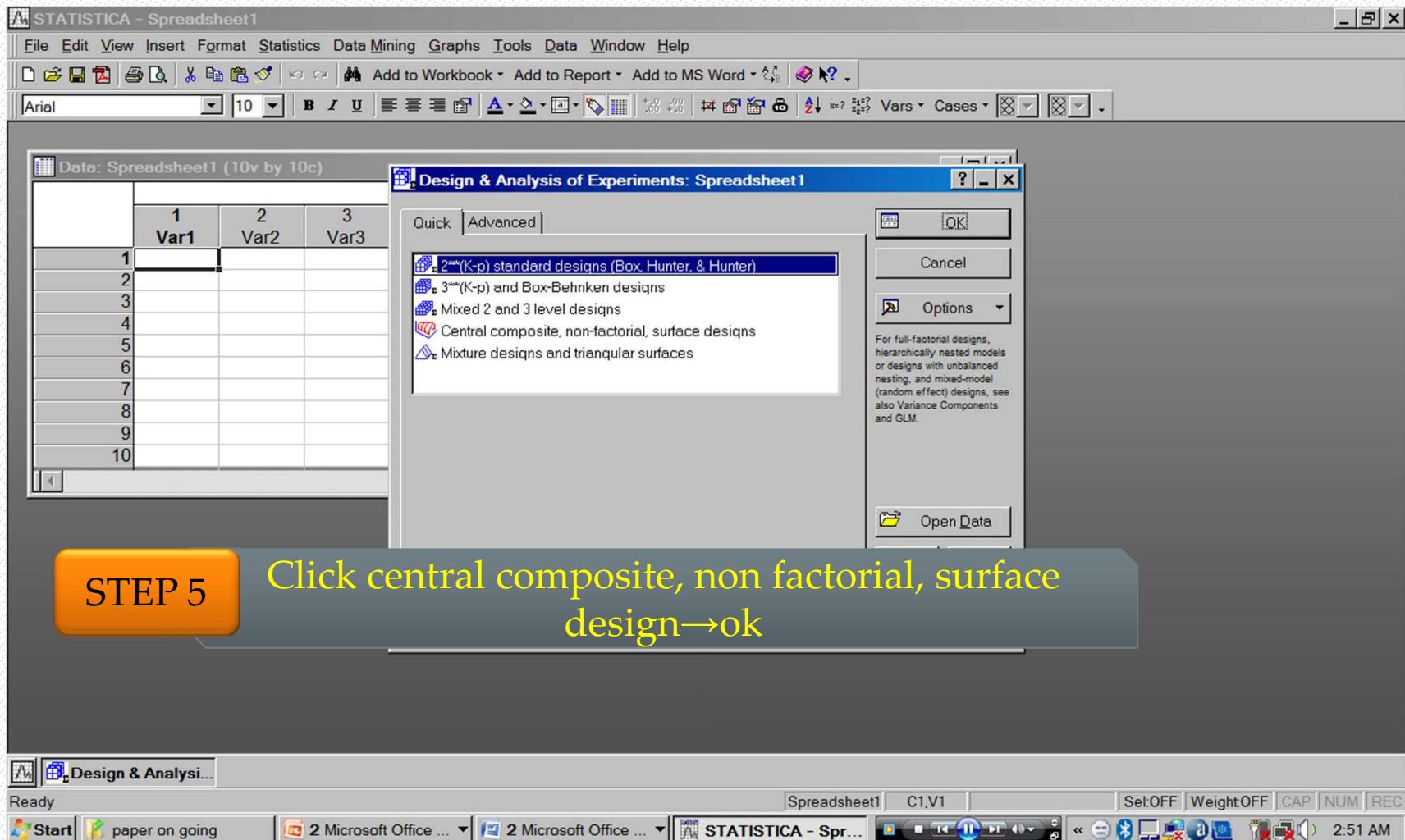
## DESIGN OF EXPERIMENT

# DOE spreadsheet





# Design & analysis of experiment windows



STATISTICA - Spreadsheet1

Data: Spreadsheet1 (10v by 10c)

Design & Analysis of Central Composite (Response Surface) Experiments : Spreadsheet

**STEP 6** Pick suitable design →ok

CCD

	1	2
Var1	1	2
1		
2		
3		
4		
5		
6		
7		

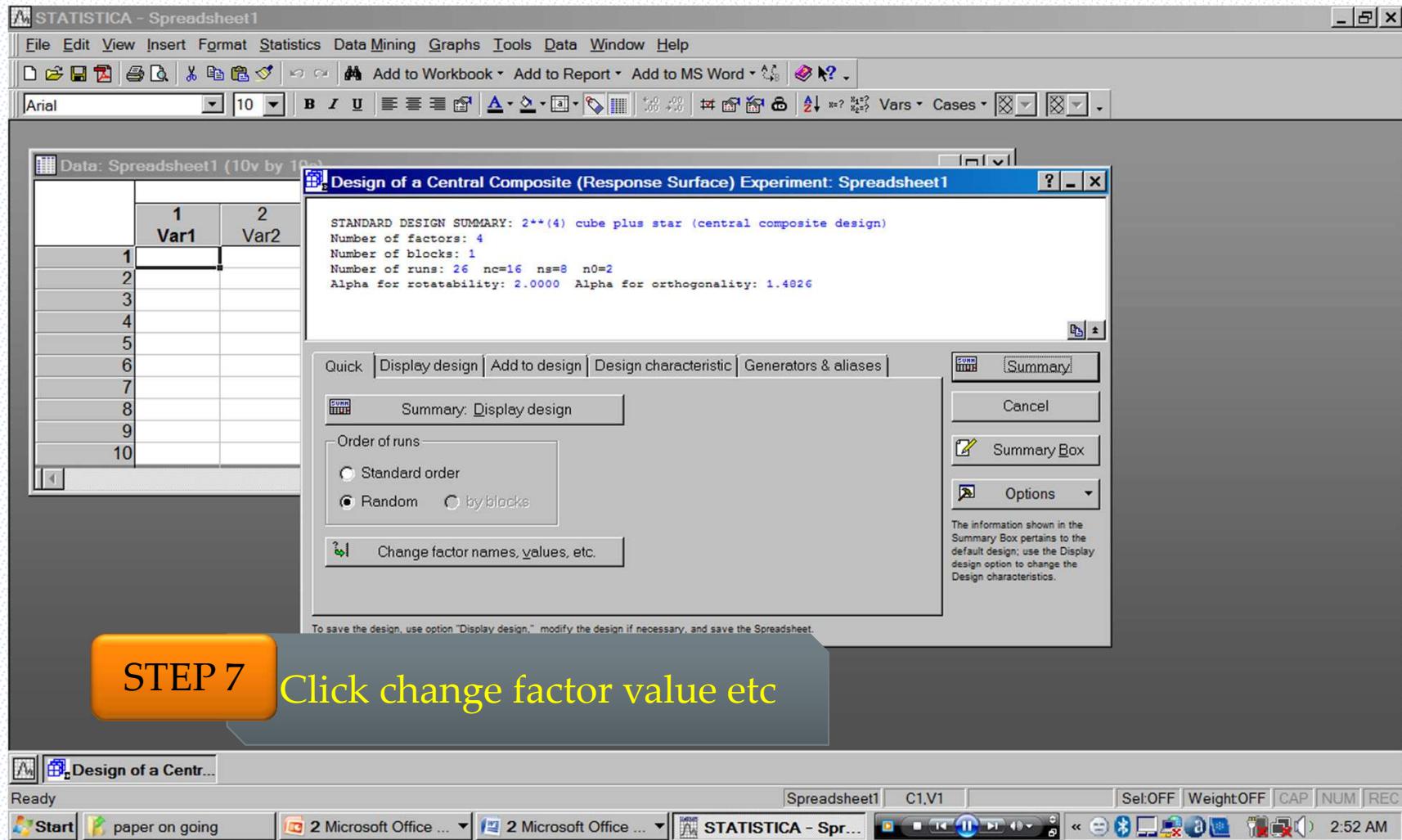
Factors/blocks/runs:

1/1/10	5/1/44	3/1/82
2/2/10	5/2/44	8/2/82
3/1/16	5/5/47	8/5/82
3/2/16	6/1/46	
3/3/17	6/2/46	
4/1/26	6/3/47	
4/2/26	7/1/80	
4/3/27	7/2/80	
5/1/27	7/5/83	
5/2/28	7/9/87	

Factors/blocks/runs:

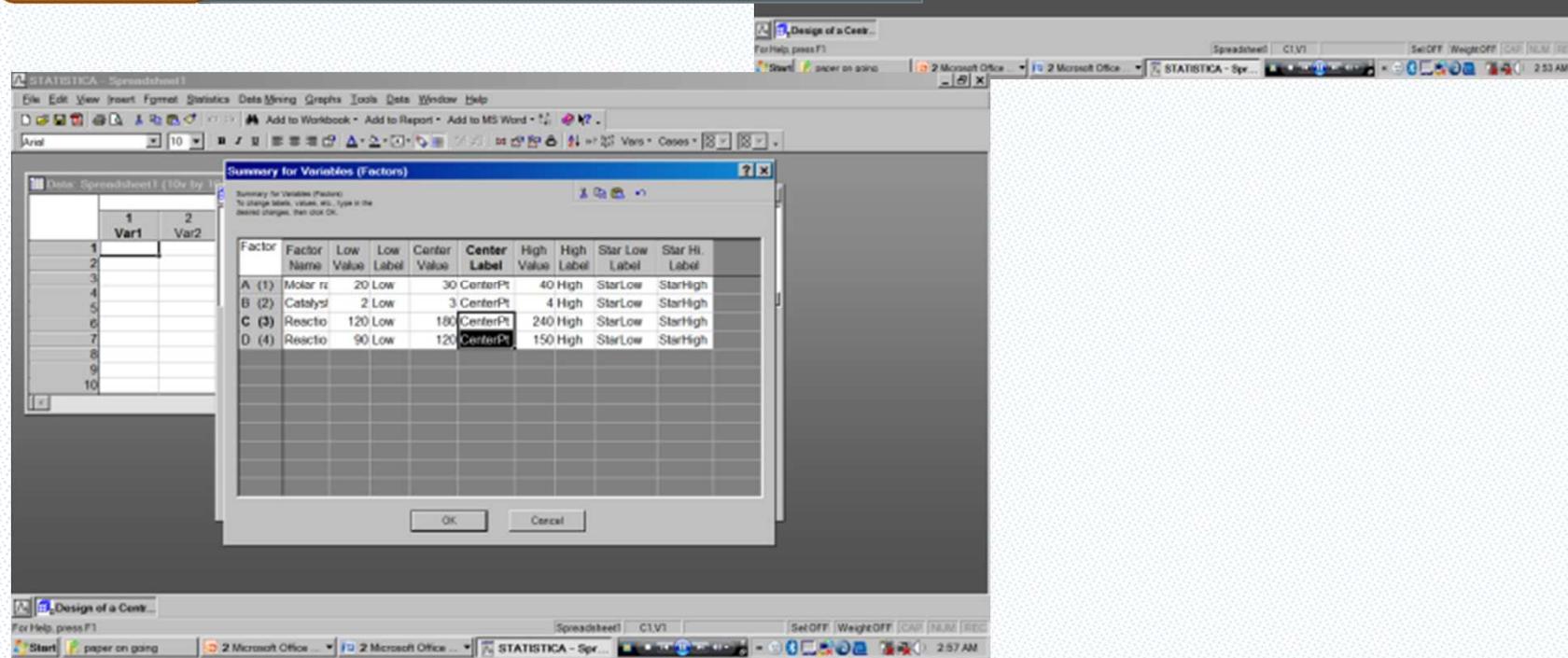
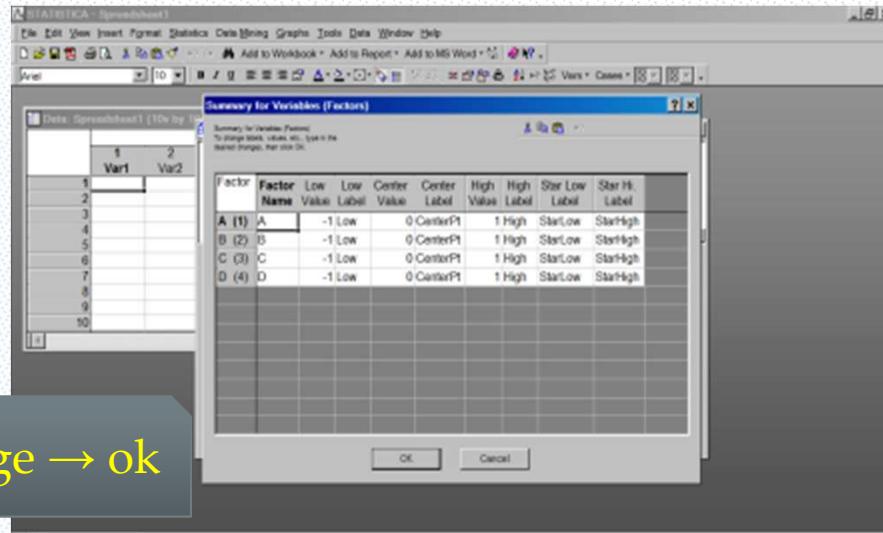
4/1/18
4/2/18
5/1/24
5/2/24
6/1/30
6/2/30
7/1/40
7/2/40
8/1/54
8/2/54

OK Cancel Options SELECT CASES



# Change value

STEP 8 Insert the variable and range → ok



## STEP 9

Click design display  
(standard order)

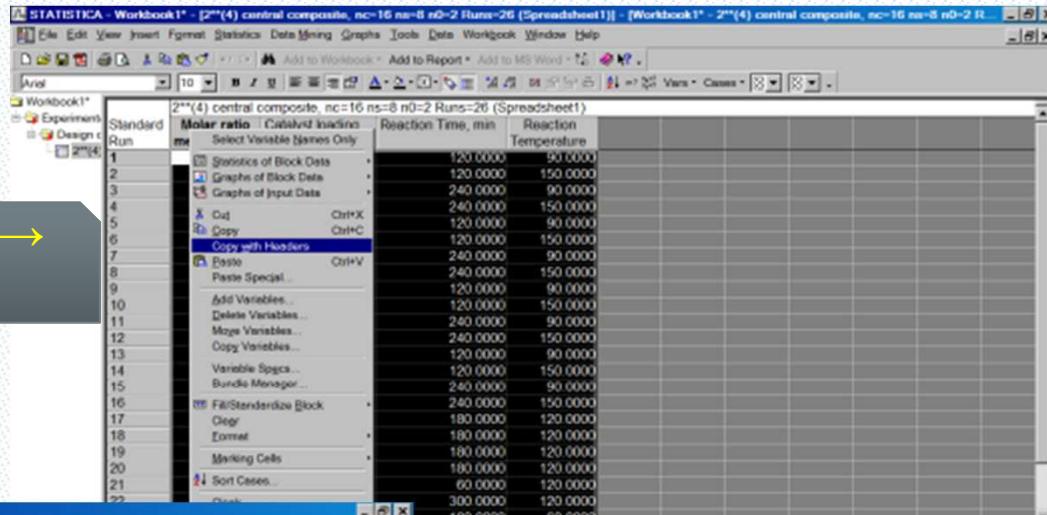
The screenshot shows the STATISTICA software interface. At the top, there is a menu bar with File, Edit, View, Insert, Format, Statistics, Data Mining, Graphs, Tools, Data, Workbook, Window, Help. Below the menu is a toolbar with various icons. A dialog box titled "Design of a Central Composite (Response Surface) Experiment: Spreadsheet1" is open, showing a summary of the design: "STANDARD DESIGN SUMMARY: 2\*\*4 cube plus star (central composite design)", "Number of blocks: 1", "Number of runs: 24", "alpha = 0.05", "alpha for irreversibility: 1.4500", and "alpha for nonorthogonality: 1.4500". The dialog has tabs for Quick, Display design, Add to design, Design characteristic, Generate & alias, and Summary. Under "Order of runs", the "Standard order" radio button is selected. A note at the bottom says "To review the design, use option 'Display design', modify the design if necessary, and save the Spreadsheet". In the background, there is a main window titled "STATISTICA - Workbook1\* - 2\*\*4 central composite, nc=16 ns=8 n0=2 Runs=26 (Spreadsheet1)" showing a table of experimental runs. The table has columns for Standard Run, Molar ratio methanol: oil (wt%), Catalyst loading, Reaction Time, min, and Reaction Temperature. The runs are numbered from 1 to 26. The data shows a central composite design with varying levels of catalyst loading and reaction time.

Design display on  
workbook windows

# Copy DOE to spreadsheet

STEP 1

Select all → right click→  
copy with headers



A screenshot of Microsoft Excel showing a spreadsheet titled 'Spreadsheet1'. The data consists of 10 rows of experimental runs. The first row contains headers: 'Molar ratio', 'Catalyst loading.', 'Reaction Time, min', 'Reaction', 'Var5', 'Var6', 'Var7', and 'Var8'. Rows 2 through 10 contain data points. The 'Molar ratio' column has values like 20.000000, 20.000000, 20.000000, 20.000000, 20.000000, 20.000000, 20.000000, 40.000000, 40.000000, and 40.000000. The 'Catalyst loading.' column has values like 2.000000, 4.000000, 4.000000, 4.000000, 4.000000, 4.000000, 4.000000, 2.000000, 2.000000, and 2.000000. The 'Reaction Time, min' column has values like 120.0000, 120.0000, 240.0000, 240.0000, 120.0000, 240.0000, 240.0000, 120.0000, 120.0000, and 120.0000. The 'Reaction' column has values like 90.0000, 150.0000, 90.0000, 150.0000, 90.0000, 150.0000, 150.0000, 90.0000, 150.0000, and 150.0000. The remaining columns (Var5, Var6, Var7, Var8) are empty.

STEP 2

Paste on spreadsheet

# Edit, save & print spreadsheet

The screenshot shows the STATISTICA Spreadsheet interface with three orange callout boxes labeled STEP 1, STEP 2, and STEP 3.

**STEP 1:** Right click on the column→edit

**STEP 2:** click file→save

**STEP 3:** click file→print

The spreadsheet contains the following data:

	Molar ratio	Catalyst loading	Reaction Time, min	Reaction	conversion	ME yield	Var7	Var8	Var9
1	20.00000	2.000000	120.0000			90.0000			
2	20.00000	2.000000	120.0000			150.0000			
3	20.00000	2.000000	240.0000			90.0000			
4	20.00000	2.000000	240.0000			150.0000			
5	20.00000	4.000000	120.0000			90.0000			
6	20.00000	4.000000	120.0000			150.0000			
7	20.00000	4.000000	240.0000			90.0000			
8	20.00000	4.000000	240.0000			150.0000			
9	40.00000	2.000000	120.0000			90.0000			
10	40.00000	2.000000	120.0000			150.0000			
11	40.00000	2.000000	240.0000			90.0000			
12	40.00000	2.000000	240.0000			150.0000			
13	40.00000	4.000000	120.0000			90.0000			
14	40.00000	4.000000	120.0000			150.0000			
	to mmmm	x mmmm	120.0000			150.0000			

# Statistica tutorial 2

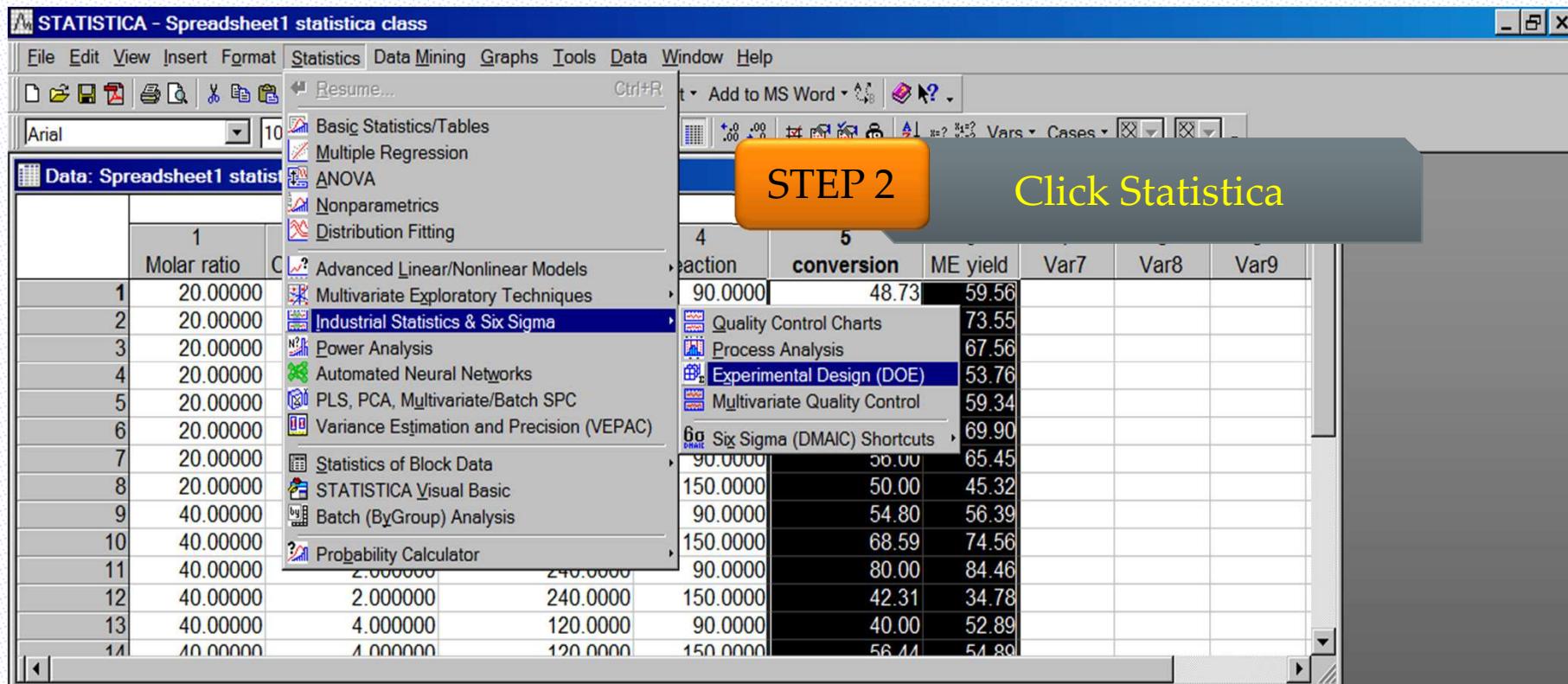
• • •

## INSERT AND ANALYSIS THE DATA

# Insert the result into spreadsheet

The screenshot shows the STATISTICA software interface. A main window titled "Data: Spreadsheet1 statistica class\* (10v by 26c)" displays a table with 14 rows and 9 columns. The columns are labeled 1 through 9, with sub-labels: Molar ratio, Catalyst loading, Reaction Time, min, Reaction, conversion, ME yield, Var7, Var8, and Var9. The data includes various numerical values such as 20.00000, 120.0000, 90.0000, etc. A yellow callout box points to the text "Open spreadsheet and insert the result". To the right of the callout is a yellow button labeled "STEP 1". Below the main window, a status bar shows the text "For Help, press F1", "Page Load Error - ...", "STATI", "Eight OFF CAP NUM REC", and the time "12:01 PM".

	1 Molar ratio	2 Catalyst loading,	3 Reaction Time, min	4 Reaction	5 conversion	6 ME yield	7 Var7	8 Var8	9 Var9
1	20.00000	2.000000	120.0000	90.0000	48.73	59.56			
2	20.00000	2.000000	120.0000	150.0000	69.02	73.55			
3	20.00000	2.000000	240.0000	90.0000	50.00	67.56			
4	20.00000	2.000000	240.0000	150.0000	40.00	53.76			
5	20.00000	4.000000	120.0000	90.0000	43.88	59.34			
6	20.00000	4.000000	120.0000	150.0000	53.00	69.90			
7	20.00000	4.000000	240.0000	90.0000	56.00	65.45			
8	20.00000	4.000000	240.0000	150.0000	50.00	45.32			
9	40.00000	2.000000	120.0000	90.0000	54.80	56.39			
10	40.00000	2.000000	120.0000	150.0000	68.59	74.56			
11	40.00000	2.000000	240.0000	90.0000	80.00	84.46			
12	40.00000	2.000000	240.0000	150.0000	42.31	34.78			
13	40.00000	4.000000	120.0000	90.0000	40.00	52.89			
14	40.00000	4.000000	120.0000	150.0000	56.44	51.80			

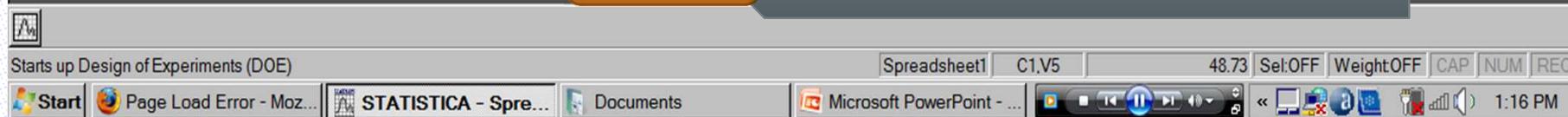


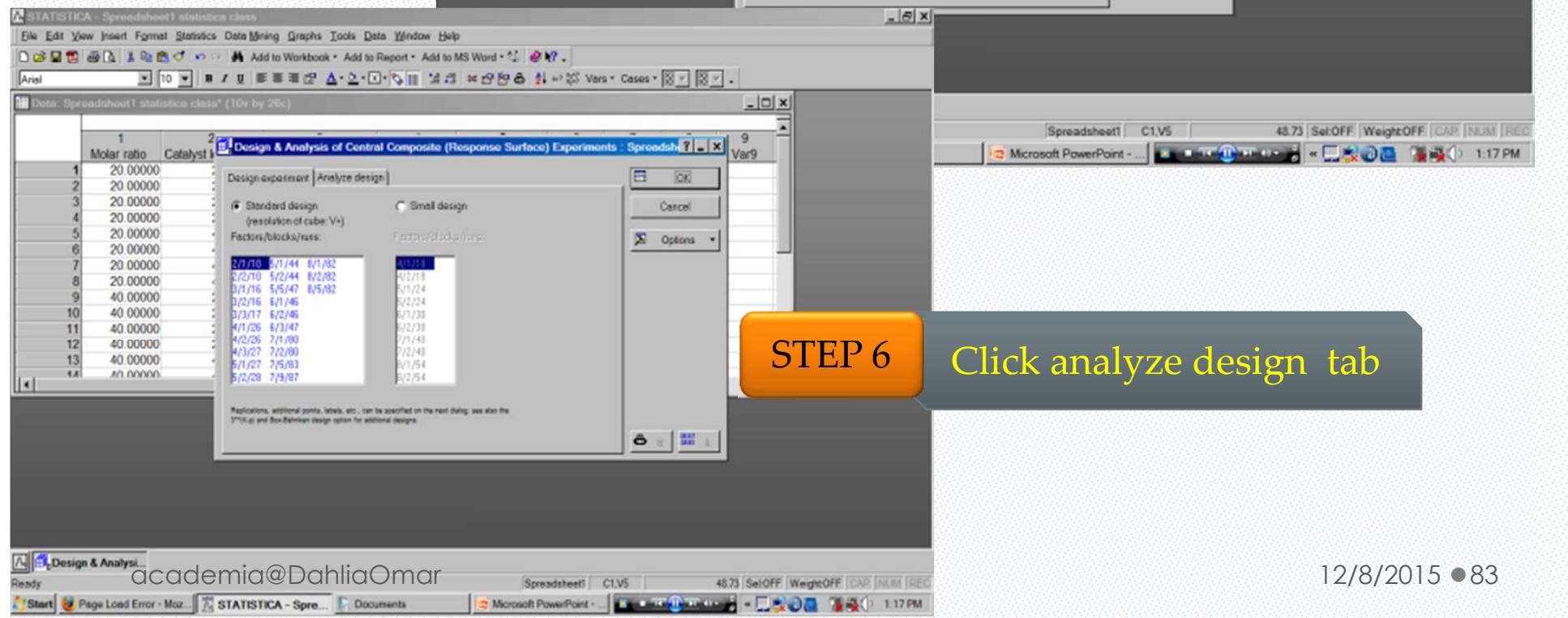
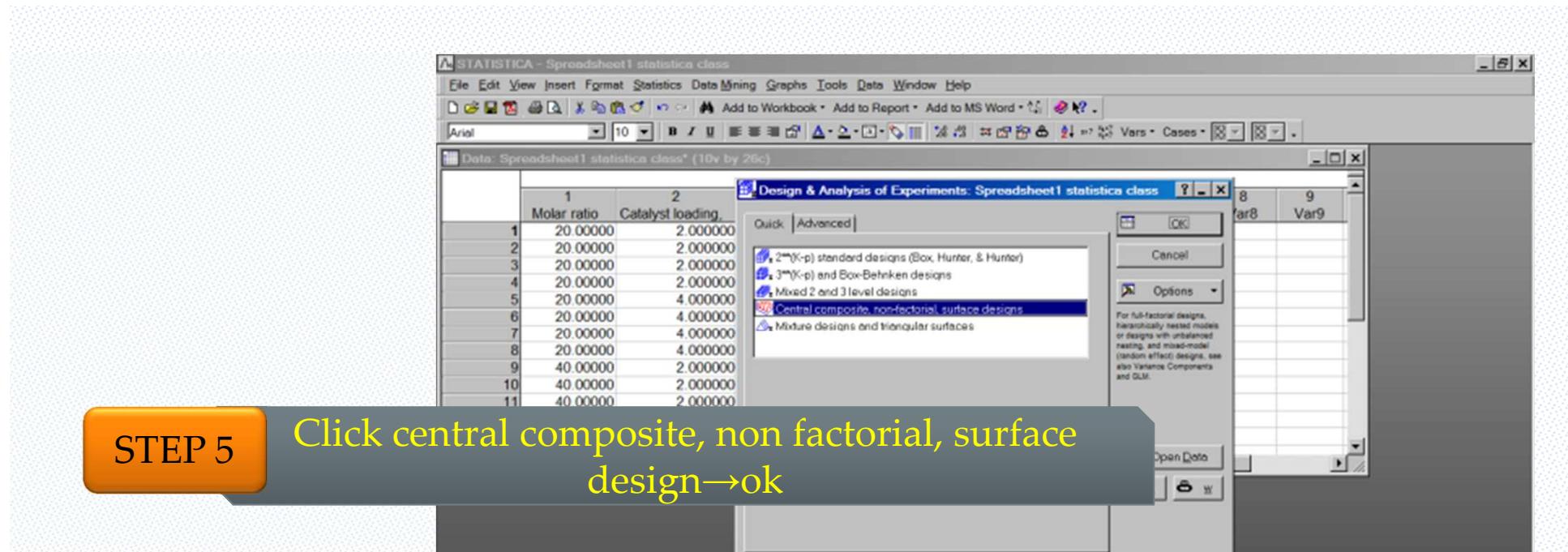
STEP 3

Click industrial statistics & six sigma

STEP 4

Click experimental design (DOE)





STATISTICA - Spreadsheet1 statistica class

Data: Spreadsheet1 statistica class\* (10v by 26c)

Design & Analysis of Central Composite (Response Surface) Experiments : Spreadsheet1

STEP 8 Click variables

1 Molar ratio 2 Catalyst level 3

	Molar ratio	Catalyst level	Var9
1	20.00000		
2	20.00000		
3	20.00000		
4	20.00000		
5	20.00000		
6	20.00000		
7	20.00000		
8	20.00000		
9	40.00000		
10	40.00000		
11	40.00000		
12	40.00000		
13	40.00000		
14	40.00000		

OK Cancel Options

Dependent: none

Independent (factors): none

Blocking variable: none

To recode factor values (levels), use

Automatically determined factor levels from file

User-defined high/low factor values

Factor levels are recoded as  $x=(value-avg.)/(range/2)$ , where range=HighValue-LowValue, and avg.=(HighValue+LowValue)/2; press F1 or click ? for more info.

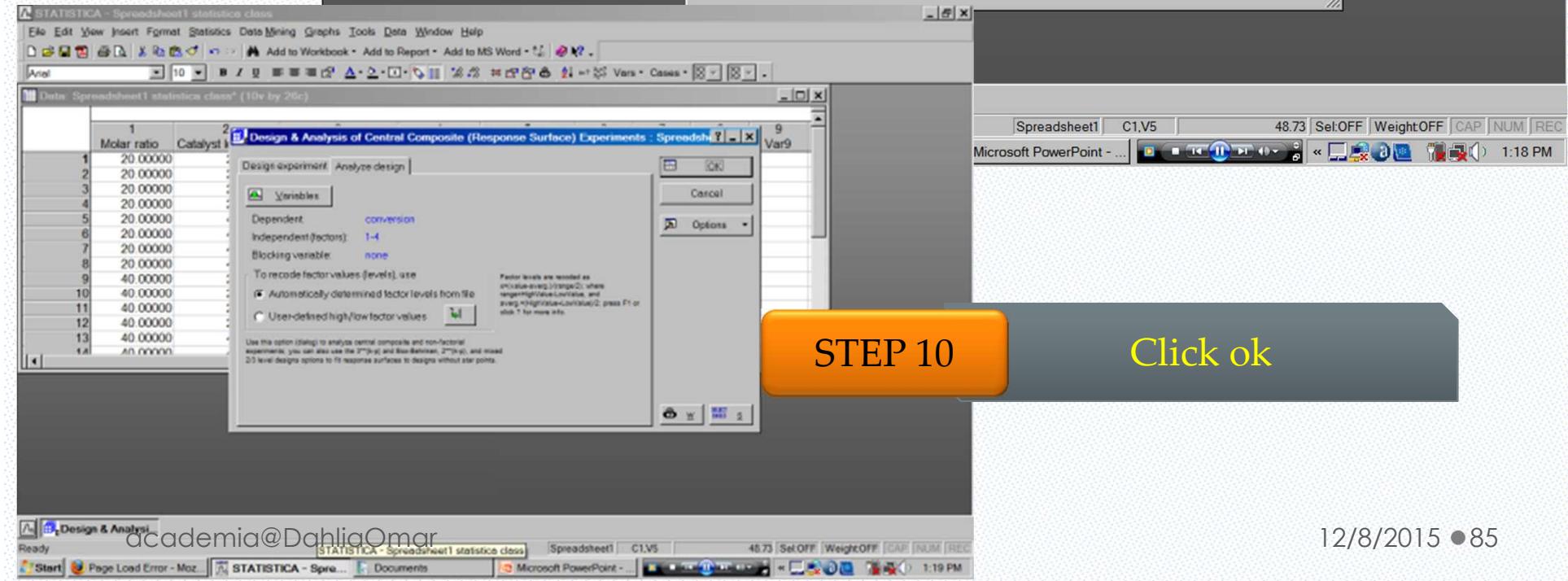
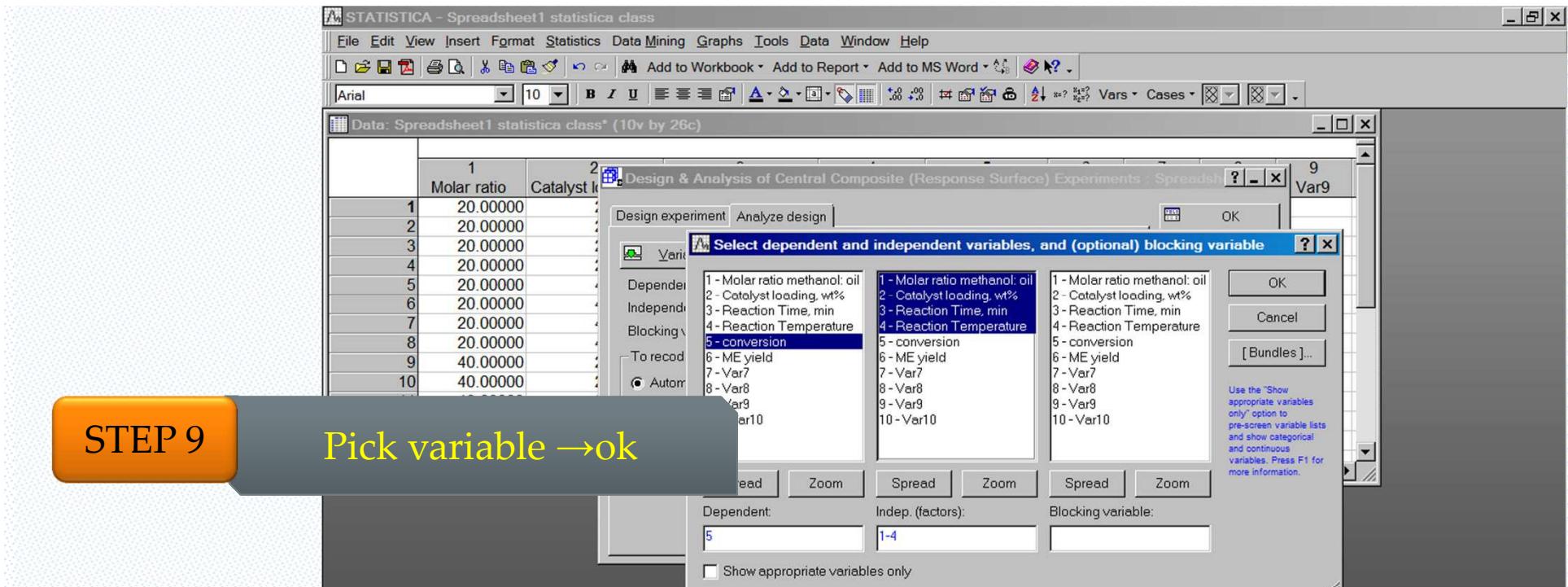
Use this option (dialog) to analyze central composite and non-factorial experiments; you can also use the  $3^{(k-p)}$  and Box-Behnken,  $2^{(k-p)}$ , and mixed 2/3 level designs options to fit response surfaces to designs without star points.

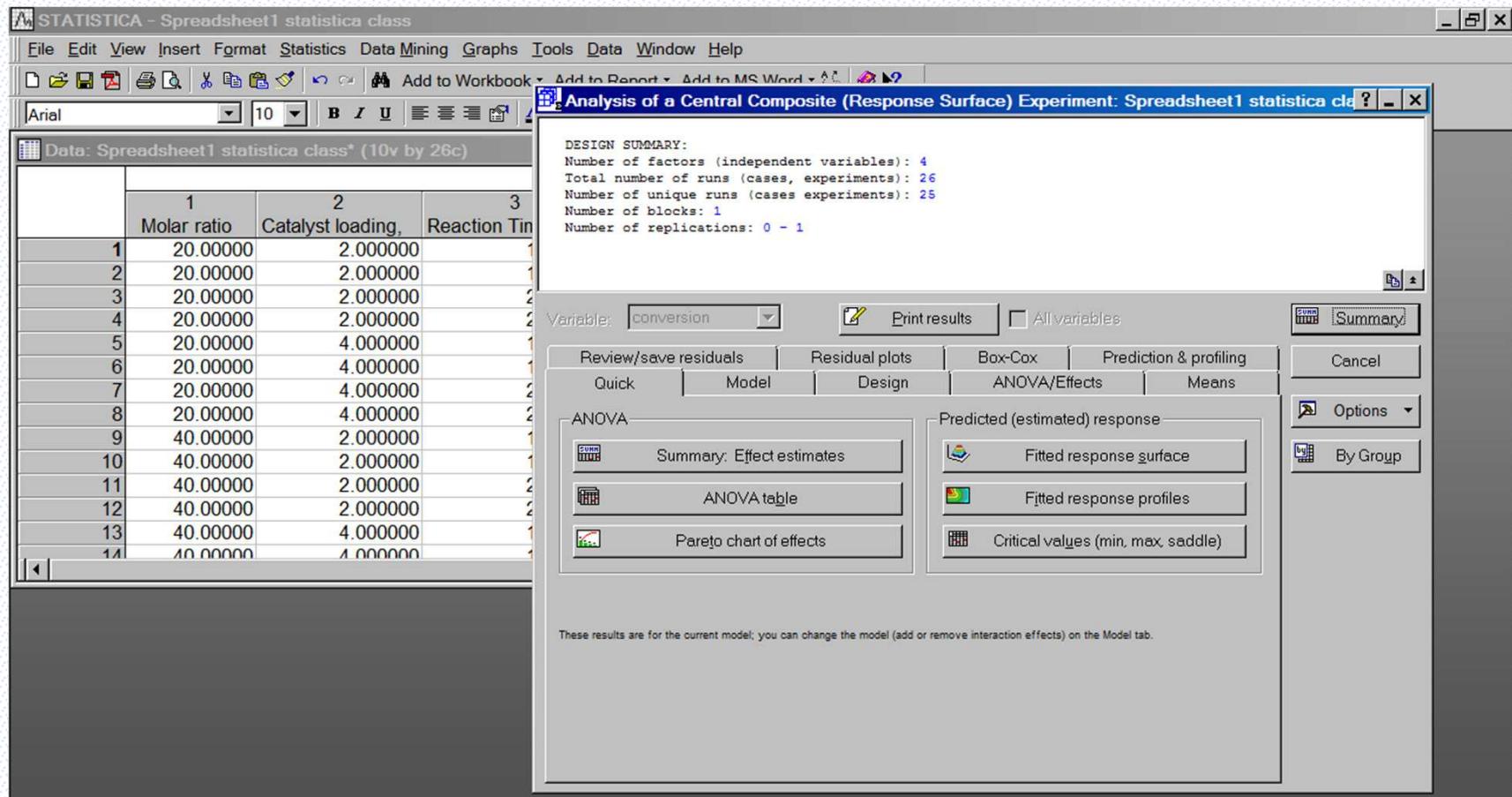
SELECT CASES

Ready

Spreadsheet1 C1,V5 48.73 Sel:OFF Weight:OFF CAP NUM REC

Start Page Load Error - Moz... STATISTICA - Spre... Documents Microsoft PowerPoint - ... 1:18 PM





**Analysis of the central composite (response surface) experiment windows opened.**  
 (note: this windows is an important for analysis since it display all information needed)

# Save as project

The screenshot shows the STATISTICA software interface. The window title is "STATISTICA - [Data: final spreadsheet adjusted (14v by 16c)]". The "File" menu is open, and the "Save Project As..." option is highlighted with a blue selection bar. A large orange button labeled "STEP 11" is overlaid on the left side of the menu. To the right of the menu, a dark gray callout box contains the text "Click file → save project as". The main workspace displays a data grid with 16 columns and approximately 20 rows of data. The columns are labeled: 3 article size (mm), 4 Reaction time (min), 5 Ozone flowrate (LPM), 6 Solid Recovery (%), 7 Lignin Degradation (%), 8 WHC (g H<sub>2</sub>O/g), 9 Mass swollen (g/g OD), 10 Solubility (%), 11 Glucose (g/g material), 12 Xylose (g/g material), and 13 Glucos (%)

3 article size (mm)	4 Reaction time (min)	5 Ozone flowrate (LPM)	6 Solid Recovery (%)	7 Lignin Degradation (%)	8 WHC (g H <sub>2</sub> O/g)	9 Mass swollen (g/g OD)	10 Solubility (%)	11 Glucose (g/g material)	12 Xylose (g/g material)	13 Glucos (%)
0.25	30	90	74.80	85.05	7.23	8.03	25.00	1.47	2.19	109
0.25	30	60	68.00	66.74	8.32	9.20	13.21	0.83	0.75	56
0.25	30	60	59.80	48.60	8.20	9.23	-3.23	0.53	0.62	31
0.25	30	90	74.06	56.82	8.23	9.00	30.43	1.10	1.68	81
0.63	30	60	59.20	98.73	5.93	6.90	3.45	1.96	3.56	115
0.63	30	90	63.60	89.11	6.67	7.67	0.00	2.20	3.50	139
0.63	30	90	63.40	73.76	7.63	8.50	15.38	2.10	2.71	133
0.63	30	60	67.40	89.94	7.17	8.17	0.00	1.90	2.26	128
0.63	60	60	61.70	67.04	5.93	6.93	0.00	1.34	2.56	82
0.63	60	60	54.40	95.54	6.87	7.90	-3.23	2.10	3.13	114
0.63	60	90	63.60	92.74	6.67	7.67	-0.00	2.33	3.64	148

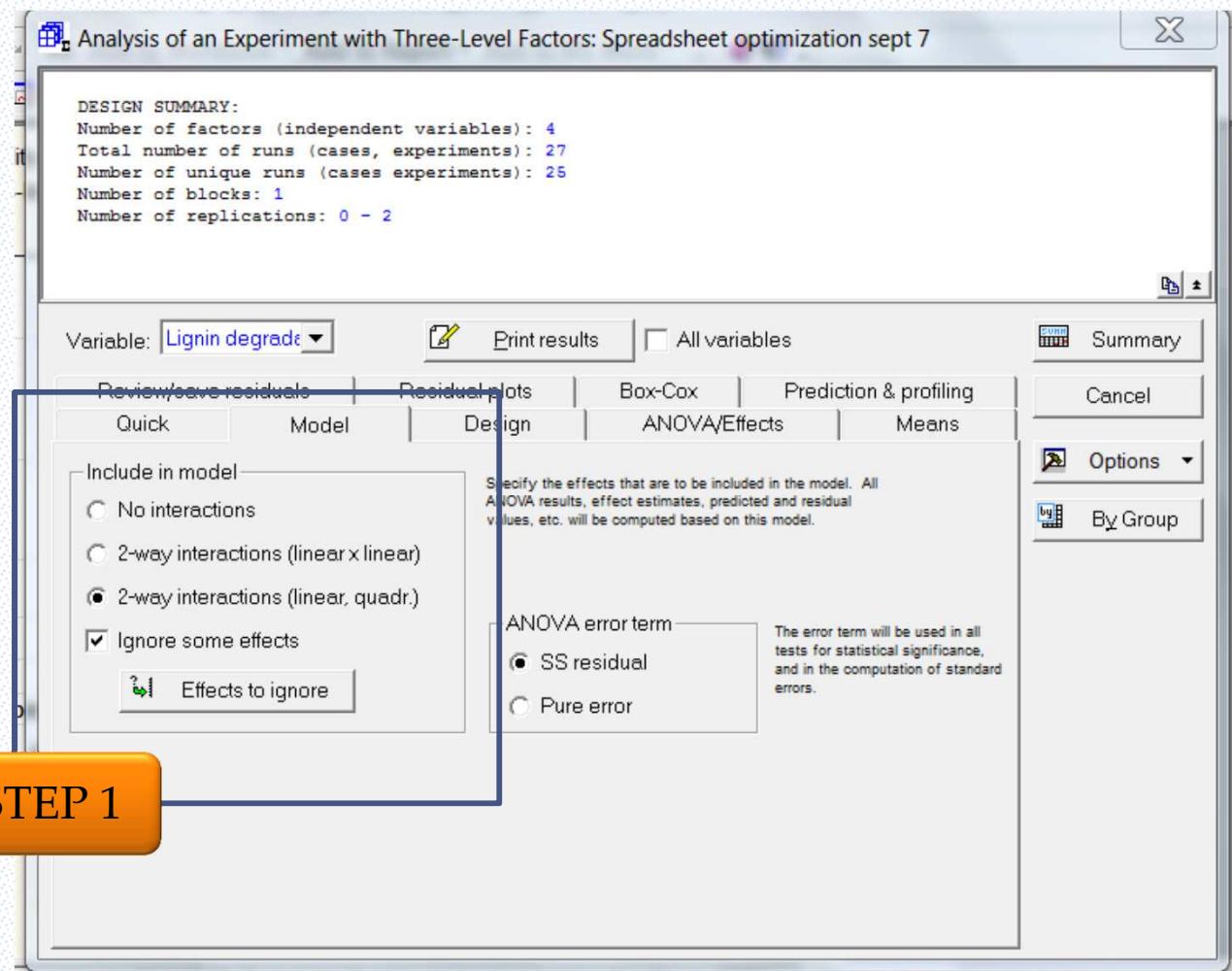
This statistica project file can be opened anytime and the analysis and workbook could be resume.

# STATISTICA TUTORIAL 3

• • •

## PREDICTED/EMPIRICAL MODEL

# Model (Coefficient selection)



**STEP 2**

Click Anova/effect  
→ regression coefficient

The screenshot shows the STATISTICA interface. On the left, a data spreadsheet titled "Anel" is displayed with columns for Block, Molar ratio methanol: oil, Catalyst loading, wt%, and Reaction Time. The right side shows the "Analysis of a Central Composite (Response Surface) Experiment" dialog box with tabs for DESIGN SUMMARY, Variable (set to conversion), and ANOVA/Effects (selected). Below it is a "Regr. Coefficients" dialog box showing a table of regression coefficients for various factors like Molar ratio methanol: oil(L) and Reaction Temperature(L).

Factor	Regressn Coeff.	Std. Err.	t(11)	P	-95. % Cnf. Limit	+95. % Cnf. Limit
Mean/Interc.	-192.048	74.73568	-2.56970	0.026064	-356.540	-27.5562
(1)Molar ratio methanol: oil(L)	4.623	1.55151	2.97971	0.012526	1.208	8.0379
Molar ratio methanol: oil(Q)	-0.050	0.01771	-2.83881	0.016116	-0.089	-0.0113
(2)Catalyst loading, wt%(L)	-0.266	15.51509	-0.01711	0.986652	-34.414	33.8830
Catalyst loading, wt%(Q)	-0.843	1.77108	-0.47580	0.643528	-4.741	3.0554
(3)Reaction Time, min(L)	0.774	0.25858	2.99253	0.012242	0.205	1.3430
Reaction Time, min(Q)	-0.001	0.00049	-2.52885	0.028031	-0.002	-0.0002
(4)Reaction Temperature(L)	1.978	0.57997	3.41055	0.005820	0.702	3.2545
Reaction Temperature(Q)	-0.003	0.00197	-1.52765	0.154828	-0.007	0.0013
1L by 2L	-0.198	0.19204	-1.03170	0.324361	-0.621	0.2245
1L by 3L	0.004	0.00320	1.10005	0.294801	-0.004	0.0108
1L by 4L	-0.011	0.00640	-1.74316	0.109151	-0.025	0.0029
2L by 3L	0.056	0.03201	1.76399	0.105449	-0.014	0.1269
2L by 4L	0.001	0.06401	0.01562	0.987816	-0.140	0.1419
3L by 4L	-0.005	0.00107	-4.75235	0.000597	-0.007	-0.0027

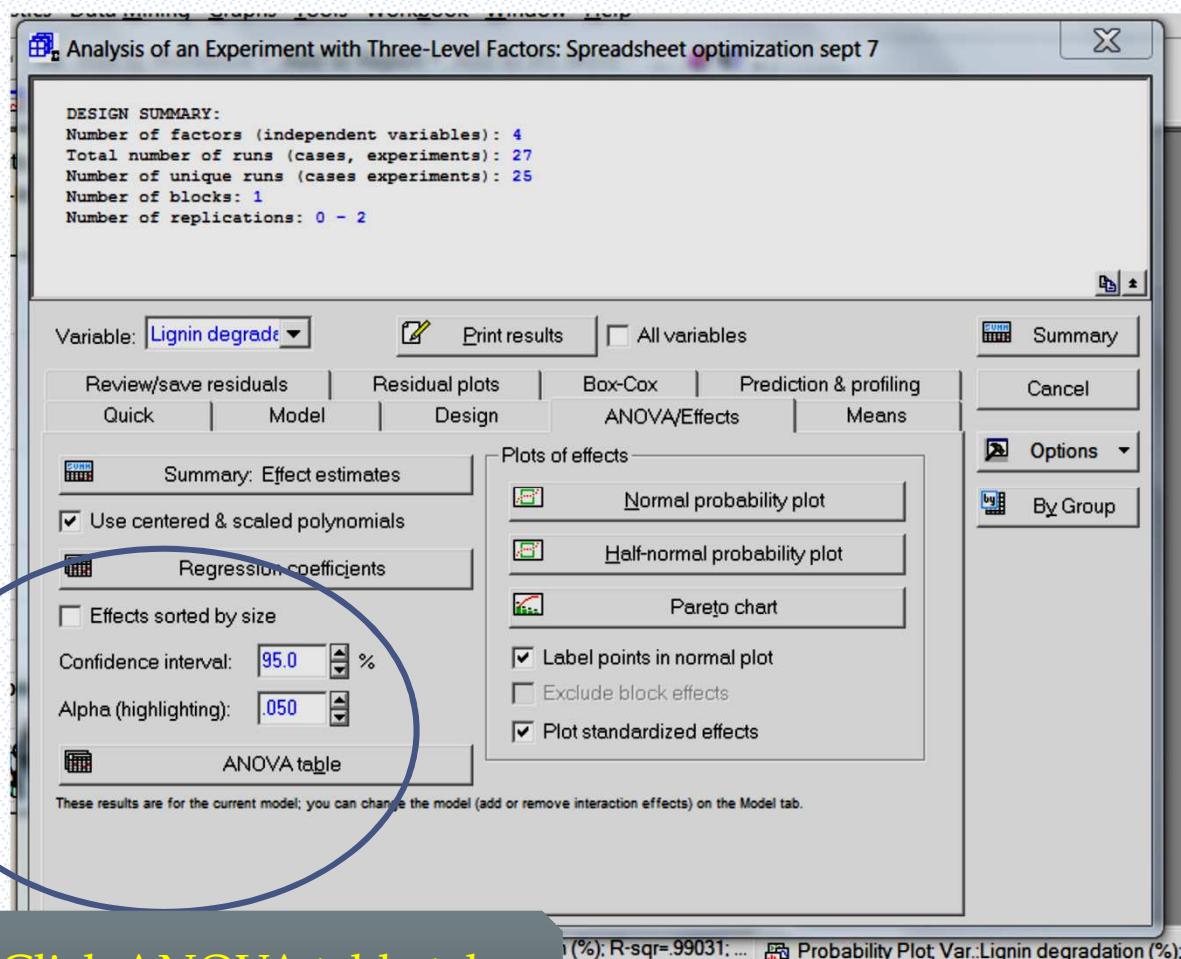
$$Y_2 = -192.048 + 4.623X_1 - 0.266X_2 + 0.774X_3 + 1.978X_4 - 0.050X_1^2 - 0.843X_2^2 - 0.001X_3^2 - 0.003X_4^2 - 0.198X_1X_2 + 0.004X_1X_3 - 0.011X_1X_4 + 0.056X_2X_3 + 0.001X_2X_4 - 0.005X_3X_4$$

# STATISTICA TUTORIAL 4

• • •

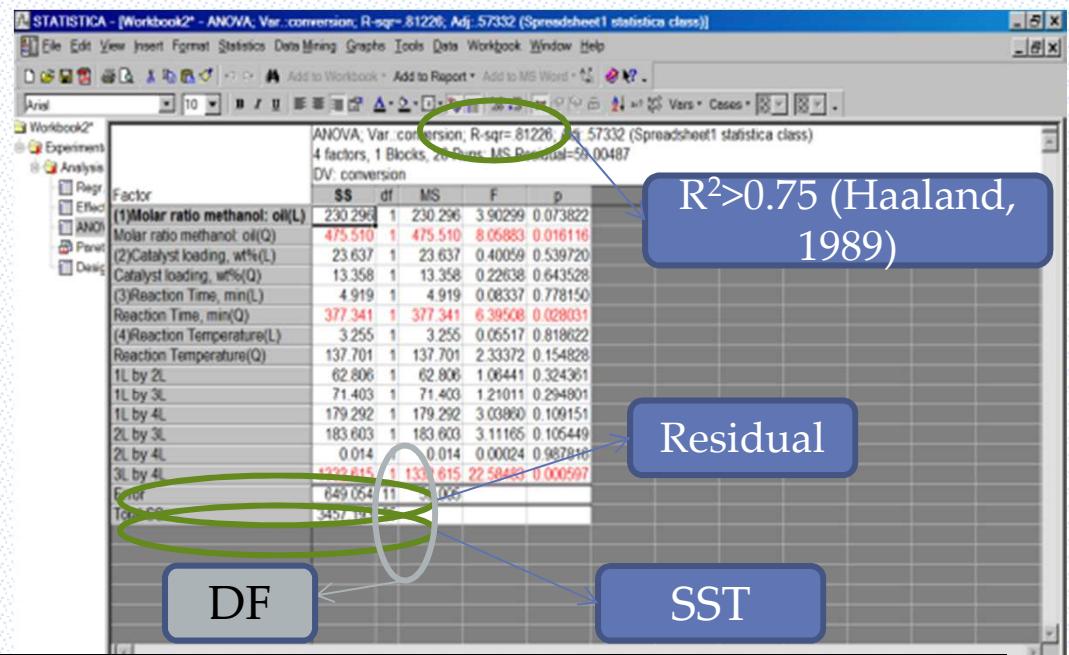
## ANOVA

# ANOVA/Effects



STEP Click ANOVA table tab

# ANOVA table



Sources	Sum of Squares(SS)	Degree of Freedom(d.f)	Mean Squares (MS)	F-value	$F_{0.05}$
Regression (SSR)	2807.32	14	200.52	3.39	>2.74
Residual	649.87	11	59.08		
Total (SST)	3457.29	25			

$$SSR = SST - \text{residual}$$

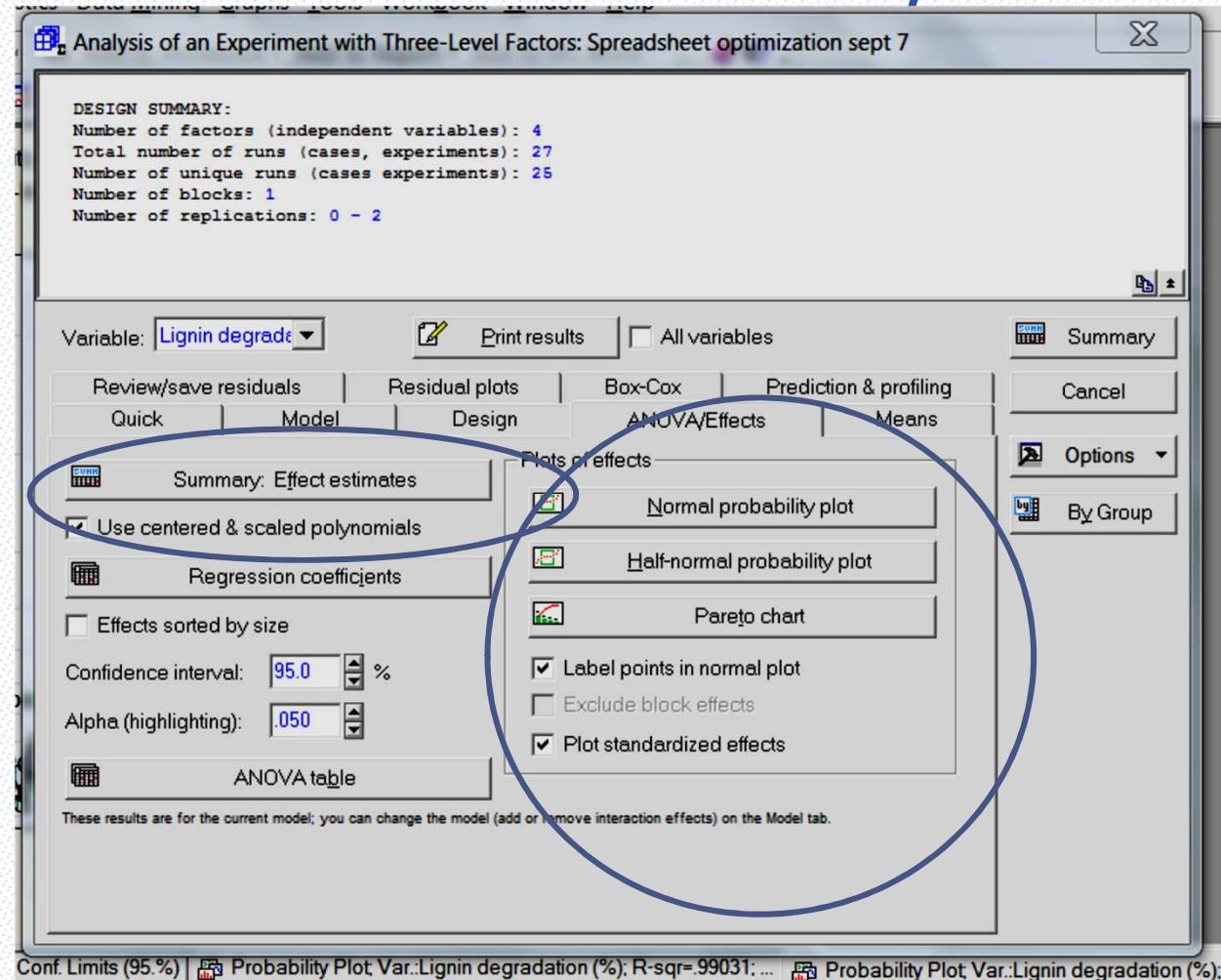
$$DF_{SSR} = DF_{SST} - DF_{\text{residual}}$$

# STATISTICA TUTORIAL 5

• • •

## Effects

# Tab of ANOVA/Effects



# Effect estimates

Effect Estimates; Var.:Lignin degradation (%); R-sqr=.99031; Adj.:87403 (Spreadsheet optimization sept 7)  
 4 3-level factors, 1 Blocks, 27 Runs; MS Residual=10.05794  
 DV: Lignin degradation (%)

Factor	Effect	Std.Err.	t(2)	p	-95.% Cnf.Limit	+95.% Cnf.Limit	Coeff.	Std.Err.	-95.% Cnf.Limit	+95.% Cnf.Limit
<b>Mean/Interc.</b>	<b>74.7109</b>	<b>1.618268</b>	<b>46.16721</b>	<b>0.000469</b>	<b>67.7481</b>	<b>81.67374</b>	<b>74.71089</b>	<b>1.618268</b>	<b>67.7481</b>	<b>81.67374</b>
(1)Particle size (mm)(L)	<b>19.6213</b>	<b>3.171426</b>	<b>6.18691</b>	<b>0.025144</b>	<b>5.9758</b>	<b>33.26686</b>	<b>9.81066</b>	<b>1.585713</b>	<b>2.9879</b>	<b>16.63343</b>
Particle size (mm)(Q)	8.1603	2.310647	3.53162	0.071665	-1.7816	18.10225	4.08017	1.155324	-0.8908	9.05112
(2)Moisture content (%) (L)	7.4338	3.178466	2.33880	0.144277	-6.2420	21.10963	3.71690	1.589233	-3.1210	10.55482
Moisture content (%) (Q)	2.8125	1.965687	1.43078	0.288787	-5.6452	11.27014	1.40623	0.982844	-2.8226	5.63507
(3)Reaction Time (min) (L)	1.9354	3.178466	0.60890	0.604542	-11.7405	15.61119	0.96768	1.589233	-5.8702	7.80560
Reaction Time (min) (Q)	3.6935	1.965687	1.87896	0.201020	-4.7642	12.15112	1.84673	0.982844	-2.3821	6.07556
(4)Ozone flowrate (LPM) (L)	-7.3450	3.178466	-2.31086	0.147050	-21.0208	6.33083	-3.67250	1.589233	-10.5104	3.16542
Ozone flowrate (LPM) (Q)	5.4365	1.650462	3.29390	0.081112	-1.6649	12.53783	2.71823	0.825231	-0.8325	6.26891
1L by 2L	<b>-13.7577</b>	<b>3.171426</b>	<b>-4.33802</b>	<b>0.049247</b>	<b>-27.4033</b>	<b>-0.11217</b>	<b>-6.87886</b>	<b>1.585713</b>	<b>-13.7016</b>	<b>-0.05609</b>
1L by 2Q	-5.0159	2.242537	-2.23672	0.154775	-14.6648	4.63292	-2.50797	1.121268	-7.3324	2.31646
1Q by 2L	-6.3382	2.264851	-2.79849	0.107490	-16.0830	3.40669	-3.16909	1.132425	-8.0415	1.70335
1Q by 2Q	<b>-9.1473</b>	<b>1.601491</b>	<b>-5.71174</b>	<b>0.029311</b>	<b>-16.0380</b>	<b>-2.25663</b>	<b>-4.57365</b>	<b>0.800746</b>	<b>-8.0190</b>	<b>-1.12832</b>
1L by 3L	1.3490	3.171426	0.42537	0.711968	-12.2965	14.99456	0.67451	1.585713	-6.1483	7.49728
1L by 3Q	-0.9315	2.242537	-0.41536	0.718199	-10.5803	8.71739	-0.46573	1.121268	-5.2902	4.35870
1Q by 3L	-2.6225	2.264851	-1.15793	0.366486	-12.3674	7.12234	-1.31126	1.132425	-6.1837	3.56117
1Q by 3Q	-2.0845	1.601491	-1.30163	0.322788	-8.9752	4.80611	-1.04227	0.800746	-4.4876	2.40306
1L by 4L	12.0341	3.171426	3.79453	0.062964	-1.6115	25.67960	6.01703	1.585713	-0.8057	12.83980
1Q by 4L	3.1461	2.264851	1.38910	0.299255	-6.5988	12.89098	1.57305	1.132425	-3.2994	6.44549
2L by 3L	1.8843	3.171426	0.59416	0.612662	-11.7612	15.52988	0.94217	1.585713	-5.8806	7.76494
2L by 3Q	2.3447	2.242537	1.04556	0.405508	-7.3041	11.99357	1.17236	1.121268	-3.6521	5.99679

Probability Plot Var.:Lignin degradation (%); R-sqr=.99031; ... Probability Plot Var.:Lignin degradation (%); R-sqr=.99031; ... Effect Estimates: Var:Lignin degradation (%); R-sqr=.99031

# STATISTICA TUTORIAL 6

• • •

## Mean Effect

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾ Print results  All variables

Review/save residuals | Residual plots | Box-Cox | Prediction & profiling  
 Quick | Model | Design | ANOVA/Effects | Means

Observed design and means  
 Display design and observed means  
 Show text labels instead of factor values

Predicted (estimated) means  
 Response desirability profiling  
 By Group

Statistics Data Mining Graphs Tools Data Workbook Window Help

Observed marginal means  
 Display  Means plot  
 Display/plot weighted means  
 Show confidence intervals

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

Effect Estimates: Max Lignin degradation (%) R-sqr=.99031

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾

Compute marginal means...  Summary |  Cancel  
 Options |  By Group

Factor  
 Particle si  
 Moisture co  
 Reaction Ti  
 Ozone flower

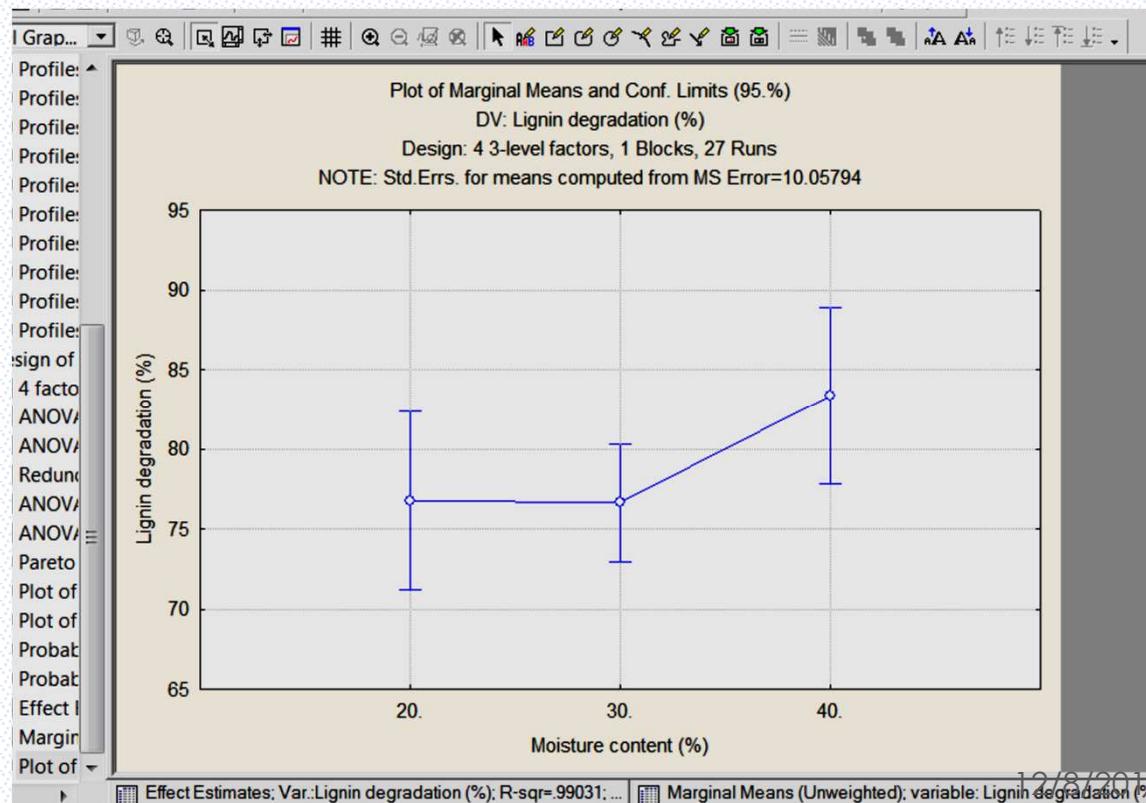
Select the factors for the marginal means table or plot.

Observed design and means  
 Display design and observe  
 Show text labels instead of factor values

Observed marginal means  
 Display  Means plot  
 Display/plot weighted means  
 Show confidence intervals

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

Marginal Means (Unweighted); variable: Lignin degradation (%) (Spreadsheet optimization sept 7)								
Design: 4 3-level factors, 1 Blocks, 27 Runs								
NOTE: Std.Errs. for means computed from MS Error=10.05794								
Moisture content (%)	Means	Pooled Std.Dev.	Overall Std.Dev.	N	Std.Err. for Mean	-95.% Cnf.Limit	+95.% Cnf.Limit	
20.	76.75631	0.000000	13.80098	6	1.294729	71.18554	82.32708	
30.	76.62517	3.171426	7.92173	15	0.856745	72.93890	80.31145	
40.	83.36216	0.000000	2.74082	6	1.294729	77.79139	88.93293	



# STATISTICA TUTORIAL 7

• • •

## Contour plot

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾ Print results  All variables

Quick Model Design ANOVA/Effects Means  
Review/save residuals Residual plots Box-Cox Prediction & profiling

Response desirability profiling

Predicted vs. observed values  
 Surface plot (fitted response)  
 Contour plot (fitted response)

Show fitted function  
 Show area contours  
 Critical values, minimum, maximum

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

Var.Lignin degradation (%); R-sqr=99.031; ... Marginal Means (Unweighted); variable:Lignin degradation (%)... Plot of Marginal

**Statistics Data Mining Graphs Tools Workbook Window Help**

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾ Select factors for 3D plot

X Axis Factor Y Axis Factor OK Cancel

Response desired  
Surface plot (fitted)  
Contour plot (fitted)

Particle size (m)  
Moisture content  
Reaction Time  
Ozone flowrate

Particle size (m)  
Moisture content  
Reaction Time  
Ozone flowrate

Select the factors for 3D plot; specify values for other factors on the next dialog.

Summary Cancel Options By Group

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾ Select factor values

Reaction Time 80 OK Cancel  
Ozone flowrate 90 Common Value 0 Apply

Critical values, minimum, maximum

Print results  All variables  
Prediction & profiling  
Response desired  
Surface plot (fitted)  
Contour plot (fitted)  
 Show fitted function  
 Show area contours  
 Critical values, minimum, maximum

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

# STATISTICA TUTORIAL 8

• • •

## 3D Surface

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**  
 Number of factors (independent variables): 4  
 Total number of runs (cases, experiments): 27  
 Number of unique runs (cases experiments): 25  
 Number of blocks: 1  
 Number of replications: 0 - 2

Variable: Lignin degrade ▾ Print results  All variables

Quick | Model | Design | ANOVA/Effects | Means  
 Review/save residuals | Residual plots | Box-Cox | Prediction & profiling

**Response desirability profiling**

- Surface plot (fitted response)
- Contour plot (fitted response)
- Show fitted function
- Show area contours
- Critical values, minimum, maximum

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

**Predicted vs. observed values**

**Predict dependent variable values**

**Summary** | **Cancel** | **Options** ▾ | **By Group**

**Select factors for 3D plot**

X Axis Factor: Particle size (m)  
 Y Axis Factor: Particle size (m)  
 OK | Cancel

Moisture content  
 Reaction Time  
 Ozone flowrate

Moisture content  
 Reaction Time  
 Ozone flowrate

Select the factors for 3D plot; specify values for other factors on the next dialog.

**Summary** | **Cancel** | **Options** ▾ | **By Group**

**Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7**

**DESIGN SUMMARY:**  
 Number of factors (independent variables): 4  
 Total number of runs (cases, experiments): 27  
 Number of unique runs (cases experiments): 25  
 Number of blocks: 1  
 Number of replications: 0 - 2

Variable: Lignin degrade ▾ Print results  All variables

Quick | Model | Design | ANOVA/Effects | Means  
 Review/save residuals | Residual plots | Box-Cox | Prediction & profiling

**Select factor values**

Reaction Time: 80 | Ozone flowrate: 90 | Common Value: 0 | OK | Cancel | Apply

Critical values, minimum, maximum

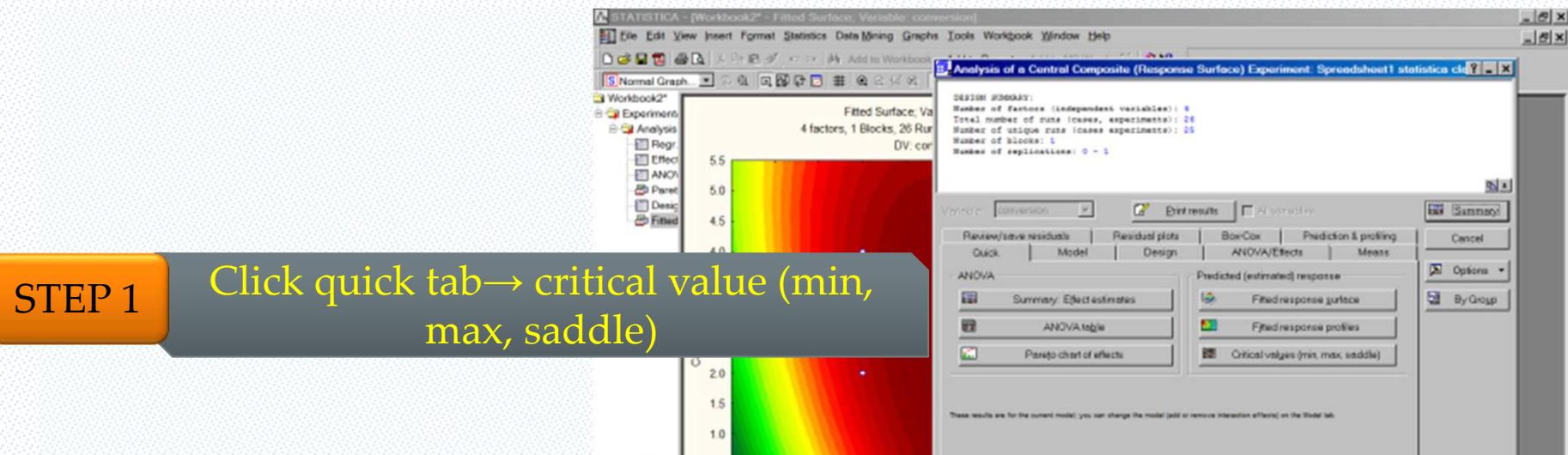
These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

# STATISTICA TUTORIAL 9

• • •

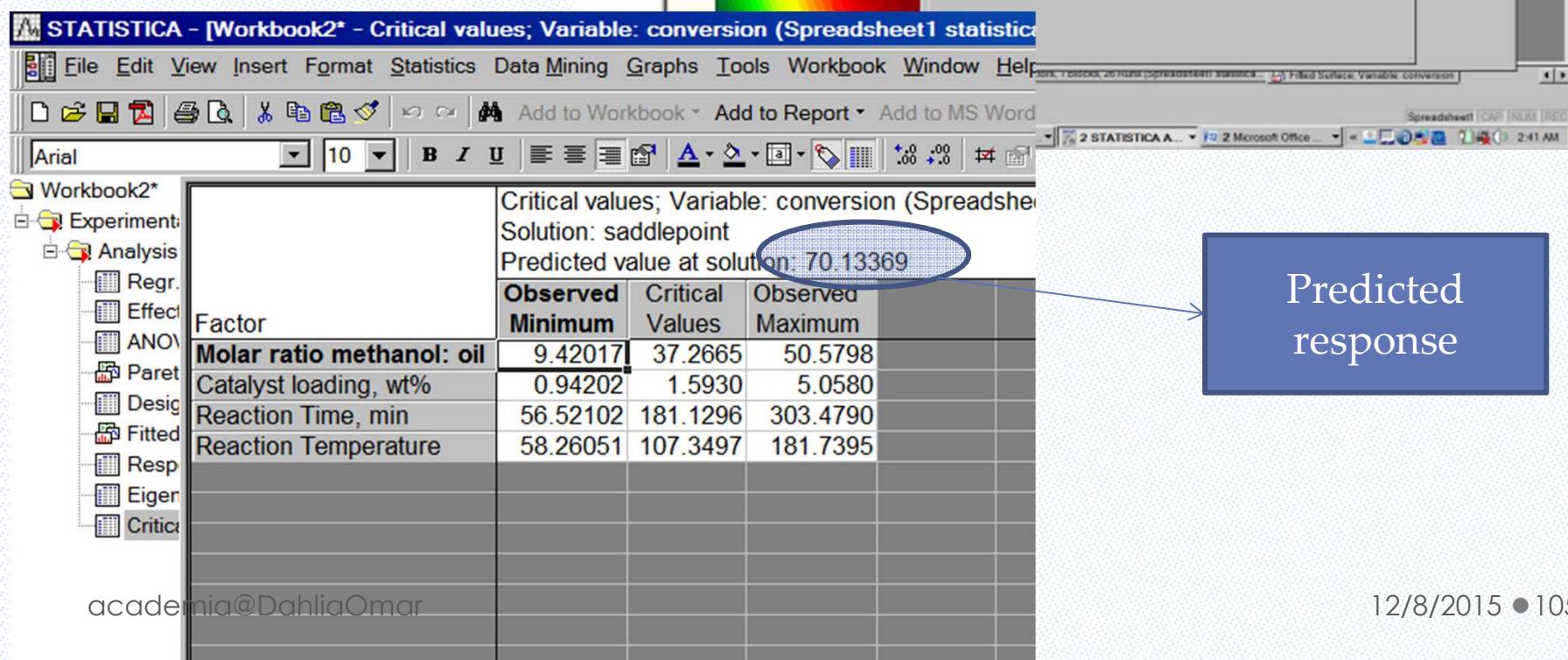
## Optimization: Single response

# OPTIMIZATION: Single responses



STEP 1

Click quick tab → critical value (min, max, saddle)



# STATISTICA TUTORIAL 10

• • •

## Optimization: Desirability function

Analysis of an Experiment with Three-Level Factors: Spreadsheet optimization sept 7

**DESIGN SUMMARY:**

- Number of factors (independent variables): 4
- Total number of runs (cases, experiments): 27
- Number of unique runs (cases experiments): 25
- Number of blocks: 1
- Number of replications: 0 - 2

Variable: Lignin degrade ▾ Print results  All variables

Quick Model Design ANOVA/Effects Means  
Review/save residuals Residual plots Box-Cox Prediction & profiling

**Response desirability profiling**  Predicted vs. observed values  Predict dependent variable values

Surface plot (fitted response) Contour plot (fitted response)  
 Show fitted function  Show area contours  Critical values, minimum, maximum

These results are for the current model; you can change the model (add or remove interaction effects) on the Model tab.

Add to report Add to MS Word X

Profiler: Spreadsheet optimization s... X

Dep. vars: 7.11 Cancel

View 1 2

Quick Save/Open Options

Set factors at

Mean values  User specified

At optimum value Block: 1

Factor grid: Particle size (mm)

Desirability function

Show desirability function

Variable: Lignin degradation (%)

Value	Desirability
Low: 50.9003	0.00
Medium: 69.4117	.50
High: 87.9231	1.00
s (curvature, low):	1.00
t (curvature, high):	1.00

Apply to all vars Reset specs for all vars 2/8/2015 107

