

Asian Productivity Organization

Feb 2019

6σ

series

THIẾT KẾ THỬ NGHIỆM
DOE

MY FIRST DESIGN OF EXPERIMENT

***Delivered by:
Azman Hussain***



Hint of the Module



***It is a knowledge.
It is never
a failure...***



MY FIRST DOE – *Main Topics*

1. Correlation Analysis
2. Simple Linear Regression
3. About DOE
4. Workshops
 - Statistical Tests
 - DOE Simulation 1
 - DOE Simulation 2
5. DOE Terminologies

1. Phân tích tương quan
2. Hồi quy tuyến tính đơn giản
3. Về DOE
4. Xưởng sx thử nghiệm
 - Thử nghiệm thống kê
 - Mô phỏng DOE 1
 - Mô phỏng DOE 2
5. Các thuật ngữ DOE



Workshop 1 TRAILER

(Statistical Test)

David thực hiện các tỉ lệ pha màu khác nhau và đo lường nhiều lần cho mỗi tỉ lệ pha. Sử dụng công cụ đồ họa và phép thống kê giúp David có quyết định tốt về tỉ lệ pha và xác suất lỗi rất nhỏ (tỉ lệ màu < 0.5 cho kết quả tốt hơn)

David is a senior chemist. Every time a new product is developed he and his team will perform multiple tests to confirm the amount of pigment used that would best suit the intended design of the product. Use relevant graphical tools and/or statistical tests to help David to make good decision with small probability of making a mistake (color reading of < 0.5 is better).

Phép thử Test #	Tỉ lệ pha 5% Pigment 5%	Tỉ lệ pha 8% Pigment 8%	Tỉ lệ pha 10% Pigment 10%	Tỉ lệ pha 15% Pigment 15%	Tỉ lệ pha 20% Pigment 20%
1	0.69	0.60	0.40	0.29	0.45
2	0.71	0.57	0.36	0.35	0.27
3	0.85	0.59	0.41	0.37	0.47
4	0.79	0.57	0.44	0.29	0.30

Mong muốn:
Tìm ra tỉ lệ màu sắc tốt nhất

Công cụ đồ họa: Box plot, Individual value plot,...

Công cụ thống kê: F-test, ANOVA,...

Goal: To find the best amount of pigment used for intended purpose that would give the best outcome (lowest colour unit number with small variability)

Graphical tools: Box plot, Individual value plot, etc)

Statistical tools: F-test (test for variance), ANOVA, etc

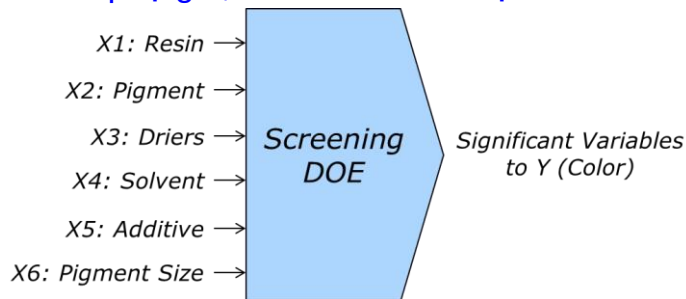


Workshop 2 TRAILER

(DOE Simulation 1)

- ◆ Being a senior chemist, David has always thought of manipulating the product formula hoping to see remarkable results rather than the normal expected outcomes. With the help from other chemists in the department, they perform an experiment manipulating the product formula systematically. Each combination is run 5 times as suggested to estimate errors present in each combination run. *The IPO chart is shown below.*

6 yếu tố ảnh hưởng: X1 Nhựa, X2 tỉ lệ màu, X3 Máy sấy, X4 Dung dịch, X5 Chất phụ gia, X6 Kích thước hạt màu



Goal: To find strong factors out of the six factors in impacting the colour quality. Factors that are least strong shall be left out in next level experiments.

Graphical tools: Marginal means
Statistical tools: ANOVA, Regression analysis, etc

Mong muốn: Tìm ra yếu tố nào ảnh hưởng mạnh nhất tới chất lượng màu sắc của sp.

Công cụ đồ họa: Marginal means

Công cụ thống kê: ANOVA, Regression analysis, ...

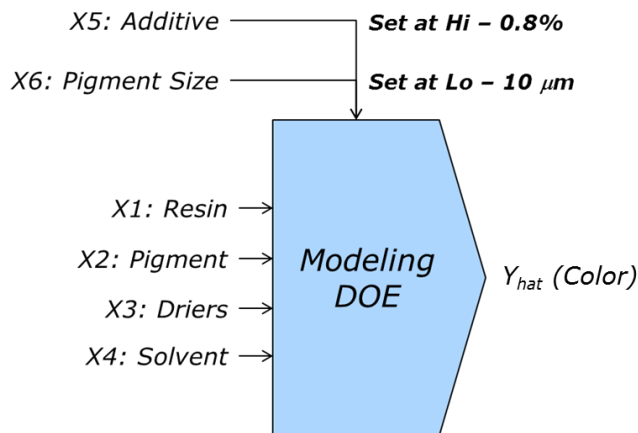


Workshop 3 TRAILER

(DOE Simulation 2)

- ◆ After completing Screening DOE, David is clearer about the significant variables in his process in resulting the Color level. He proceeds with Modeling DOE leaving the dropped variables to their optimized level based on the finding in Screening DOE.

Sau khi hoàn thành DOE Screening, David hiểu rõ các yếu tố ảnh hưởng mạnh tới quá trình pha màu và mức độ màu sắc. David tiếp tục chuyển sang bước DOE Modeling, bỏ đi các yếu tố ảnh hưởng ít.



Goal: To find the mathematical equation $[Y=f(X)]$ that will represent the relationship between the input factors and the process response (colour unit).

Graphical tools: Marginal means

Statistical tools: ANOVA, Regression analysis, t-test, F-test, capability analysis, etc

Mong muốn: tìm ra hàm số $Y = f(x)$ thể hiện mối quan hệ giữa các yếu tố đầu ra và kết quả quá trình (đơn vị màu sắc)

Công cụ đồ họa: marginal means

Công cụ thống kê: ANOVA, Regression analysis, t-test, F-test, capability analysis,...

Correlation Analysis

PHÂN TÍCH TƯƠNG QUAN



Are they from the same family?



PHÂN TÍCH MỐI QUAN HỆ

CORRELATION ANALYSIS

Simple linear regression

HỒI QUY TUYẾN TÍNH ĐƠN GIẢN

About DOE

VỀ DOE

1. *Correlation & Scatter Diagrams*
2. *Data Tabulation Type*
3. *Correlation Coefficient*
4. *Stratification*
5. *Spurious Correlation*
6. *Application Steps*

1. Biểu đồ mối quan hệ và Biểu đồ phân tán
2. Bảng phân loại dữ liệu
3. Hệ số tương quan
4. Sự phân tầng dữ liệu
5. Tương quan giả
6. Các bước áp dụng

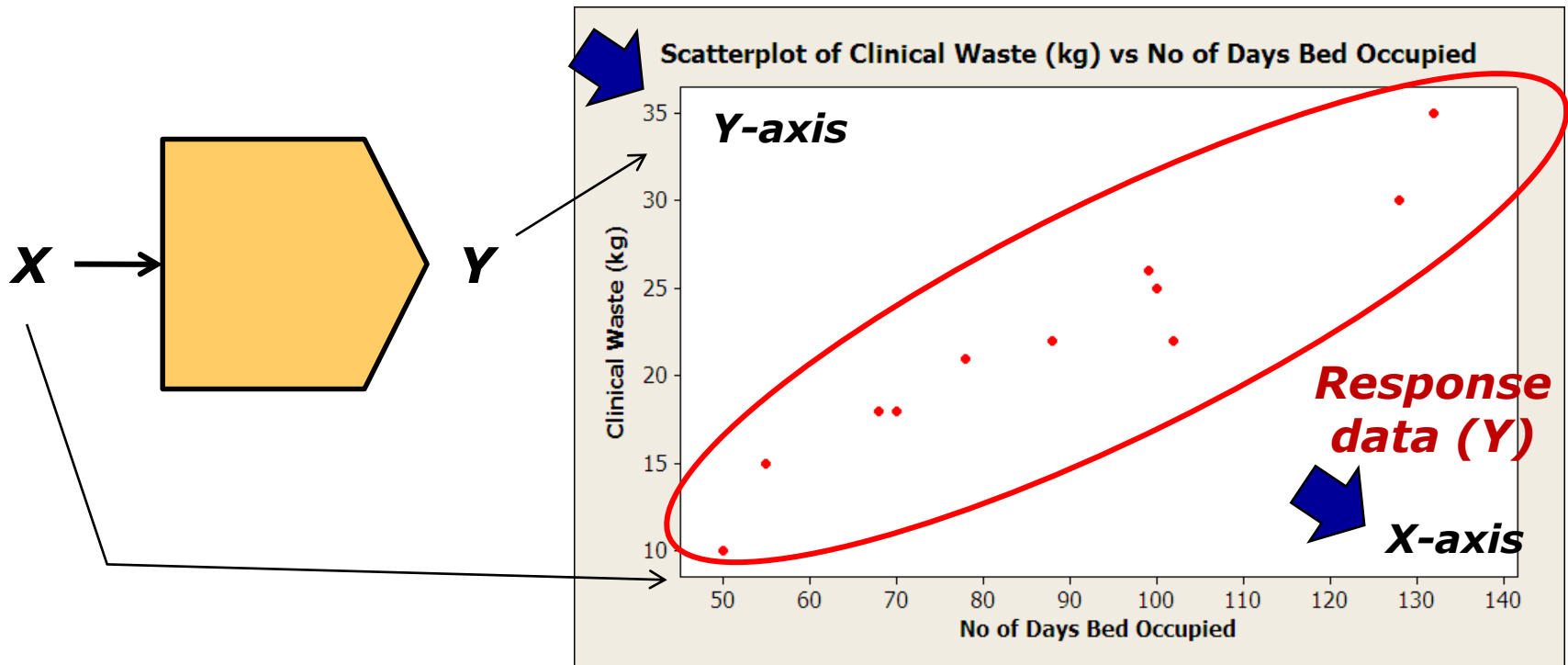


1-1. Correlation & Scatter Diagrams

1-1. Biểu đồ mối quan hệ và Biểu đồ phân tán

Biểu đồ phân tán chỉ ra sự phân loại của hai biến số (thường là 1 biến đầu vào và 1 biến kết quả)

A Scatter Diagram shows a tabulation of two variables (normally an input and its response)

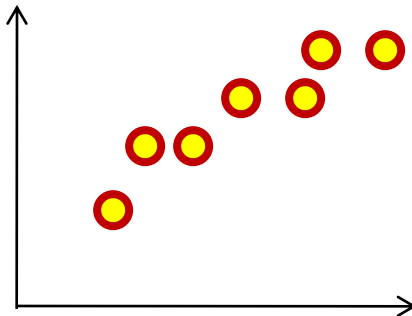




1-2. Phân loại dữ liệu **Data Tabulation Type**

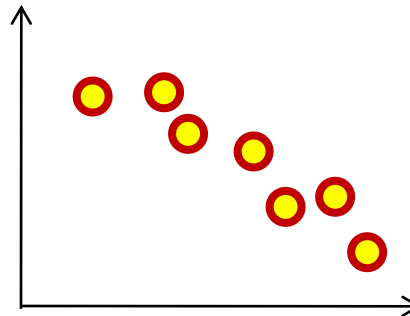
Mối quan hệ giữa hai biến có thể được phân loại thành:

The relationship or correlation of two variables can be classified as follows:-



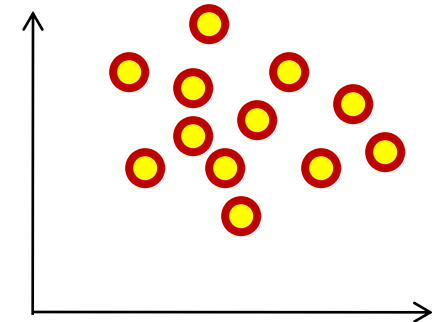
***Positively
Correlated***

Quan hệ hướng lên



***Negatively
Correlated***

Quan hệ hướng xuống



***No
Correlation***

Không có tương quan



Yếu tố nào quyết định tới
tương quan giữa các biến số?

Discussion 1-1:

What Determines the Correlation of Variables?

◆ Why in some situation the variables are positively correlated, some are negatively correlated, and some not correlated at all?

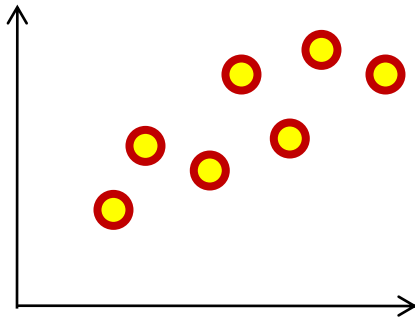
- ... Tại sao trong một tình huống nào đó, các biến số lại có
- ... mối tương quan hướng lên, hướng xuống hoặc không
- ... có mối tương quan nào?
- ...



Pop Quiz 1-1:

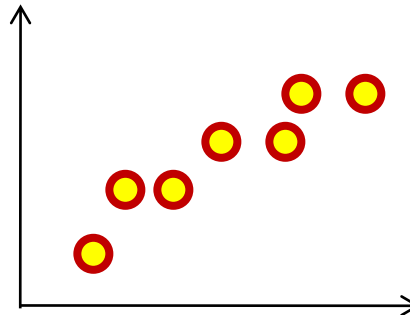
Relationship Strength Between Variables

Độ mạnh tương quan giữa các biến số



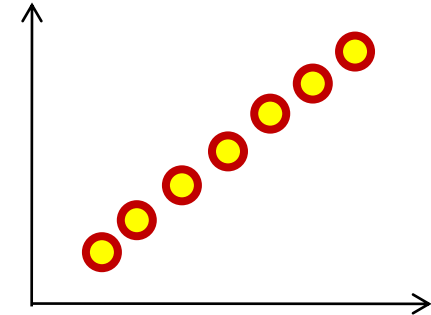
***Weak
Correlation?***

Tương quan yếu



***Moderately
Strong
Correlated?***

Tương quan mạnh vừa



***Strongly
Correlated?***

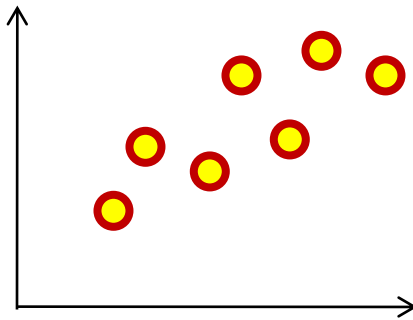
Tương quan mạnh



1-3. Hệ số tương quan **Correlation Coefficient**

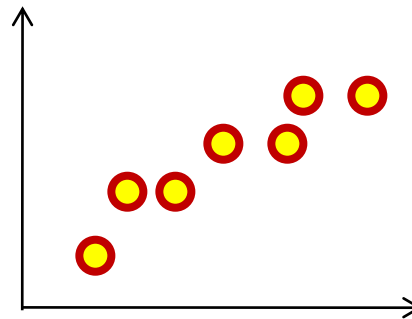
Độ mạnh tương quan giữa hai bộ dữ liệu được đặc trưng bởi hệ số tương quan, r

- ◆ The strength of correlation of two data sets is represented by using correlation coefficient, r
Hệ số tương quan nằm trong khoảng -1 đến +1
- ◆ Correlation coefficient lies between -1 and +1



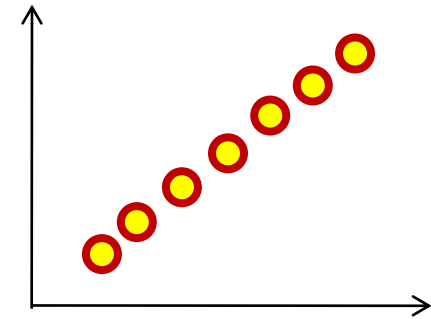
**Weak
Correlation**

$$|r| \leq 0.7$$



**Moderately
Strong
Correlated**

$$0.7 < |r| < 0.9$$



**Strongly
Correlated**

$$|r| \geq 0.9$$



Correlation Coefficient

Hệ số tương quan - công thức ***Mathematical Formula***

Công thức tính hệ số tương quan

*Mathematical equation for
Correlation Coefficient*

$$r = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$



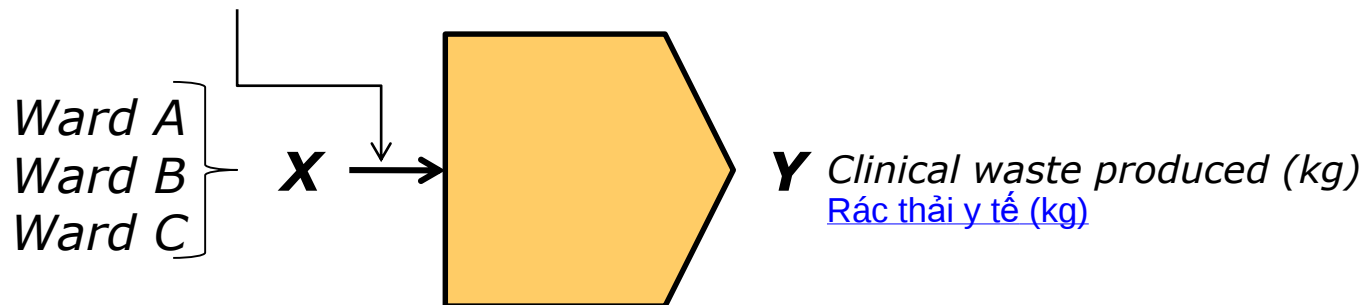
1-41. Phân tầng dữ liệu **About Stratification**

Phân tầng dữ liệu được sử dụng khi đầu vào từ nhiều nguồn khác nhau và đặt trong những trường hợp khác nhau (nhưng lại cùng loại dữ liệu). Các trường hợp khác nhau có thể hoặc không tác động lên kết quả. Phân tầng dữ liệu giúp chúng ta nhận biết các ảnh hưởng.

Stratification is used when the inputs are coming from different sources or situations (but in same category). This different situation may or may not affect the output – doing stratification will let us know the impact

Số ngày trong một tuần mà giường có bệnh nhân nằm

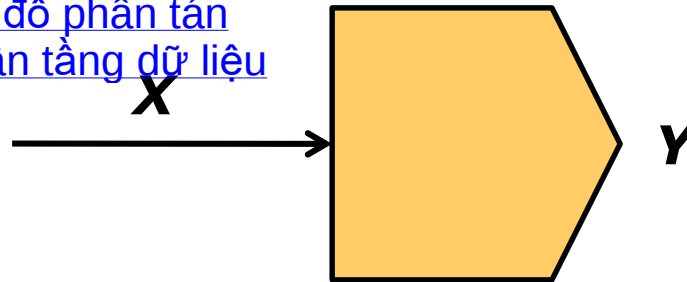
No of days bed occupied in a week





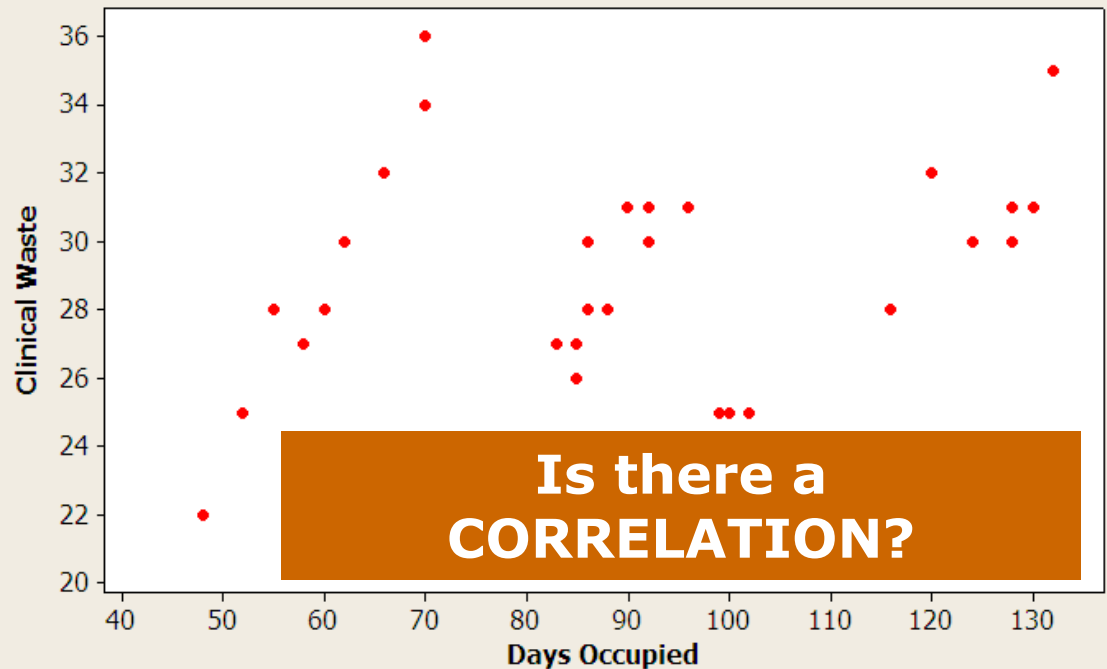
1-42. Un-Stratified Scatter Diagram

1-42. Sơ đồ phân tán chưa phân tầng dữ liệu



Wad	Days	Waste
Wad A	99	25
Wad B	83	27
Wad C	48	22
Wad A	100	25
Wad B	85	26
Wad C	52	25
Wad A	102	25
Wad B	85	27
Wad C	55	28
Wad A	116	28
Wad B	86	28
Wad C	58	27
Wad A	120	32
Wad B	86	30
Wad C	60	28
Wad A	124	30
Wad B	88	28
Wad C	62	30
Wad A	128	30
Wad B	90	31
Wad C	66	32
Wad A	128	31
Wad B	92	30
Wad C	66	32
Wad A	130	31
Wad B	92	31
Wad C	70	34
Wad A	132	35
Wad B	96	31
Wad C	70	36

Rác thải y tế theo ngày có giường bệnh nhân
Scatterplot of Clinical Waste vs Days Occupied

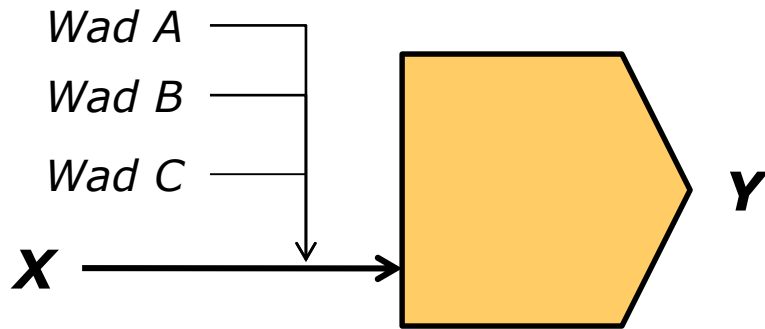




1-43. Dữ liệu được phân tầng. **Stratifying Data**

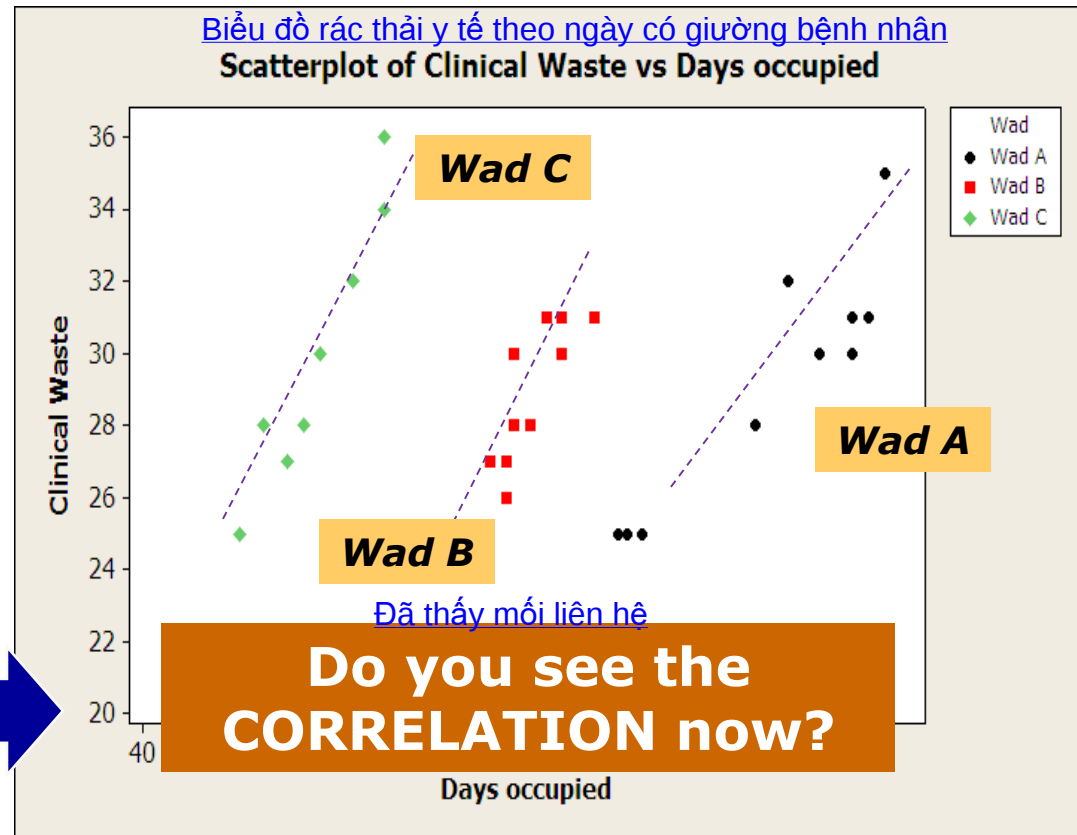
Cần hoàn thành việc gì để phân tầng dữ liệu?

- ◆ What should have been done is to stratify the data analysis.



Minitab allows Stratified Scatter diagram as shown here

Minitab cho phép vẽ được sơ đồ phân tầng dữ liệu





1-5. About Spurious Correlation

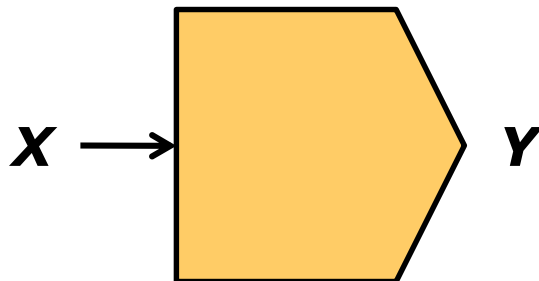
- If statistically a correlation exists between y and x, that does not mean y and x do have correlation
- Other variables also may have caused the correlation between y and x to exist

Nếu từ thống kê tìm ra mối quan hệ giữa y và x, nhưng không có nghĩa là giữa y và x có tương quan thật

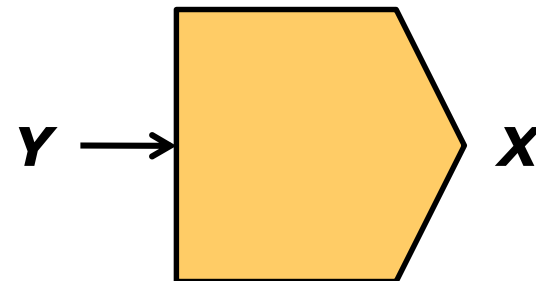
Các biến số khác cũng có thể tác động làm cho giữa y và x hình thành mối quan hệ



X produces Y



Y produces X



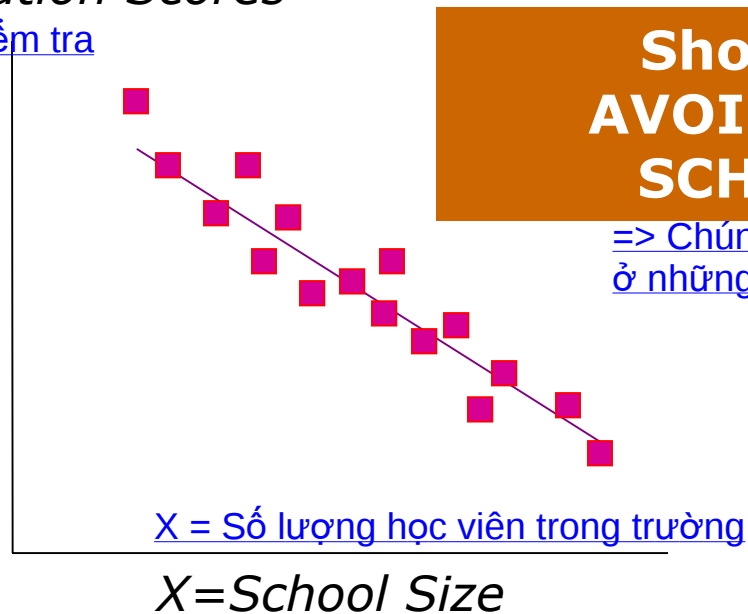


Discussion 1-2: Spurious Correlation

- ◆ Data showed that as the school size increases, the examination scores of students dramatically reduces.
Dữ liệu cho thấy khi số lượng học sinh trong trường tăng lên, thì kết quả kiểm tra của các học viên giảm đáng kể.

$Y = \text{Examination Scores}$

$Y = \text{Kết quả kiểm tra}$



**Should we
AVOID LARGE
SCHOOLS?**

$\Rightarrow$ Chúng ta có nên tránh học
ở những trường lớn ko?

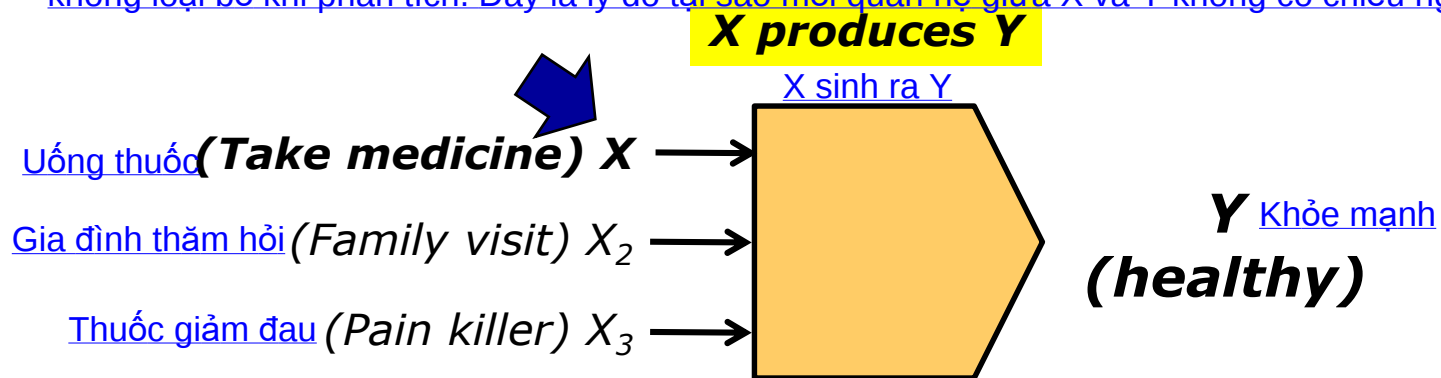


1-5. (cont) About Spurious Correlation

1-5. Tương quan giả

- In A Scatter Diagram, we only consider one X variable. There are other Xs' which we assume do not move at the time of study. This is why this $X \Leftrightarrow Y$ relationship is not vice-versa.

Trong sơ đồ phân tán, chúng ta chỉ xem xét được một biến x. Có những biến X khác được giả định nhưng không loại bỏ khi phân tích. Đây là lý do tại sao mối quan hệ giữa X và Y không có chiều ngược lại.

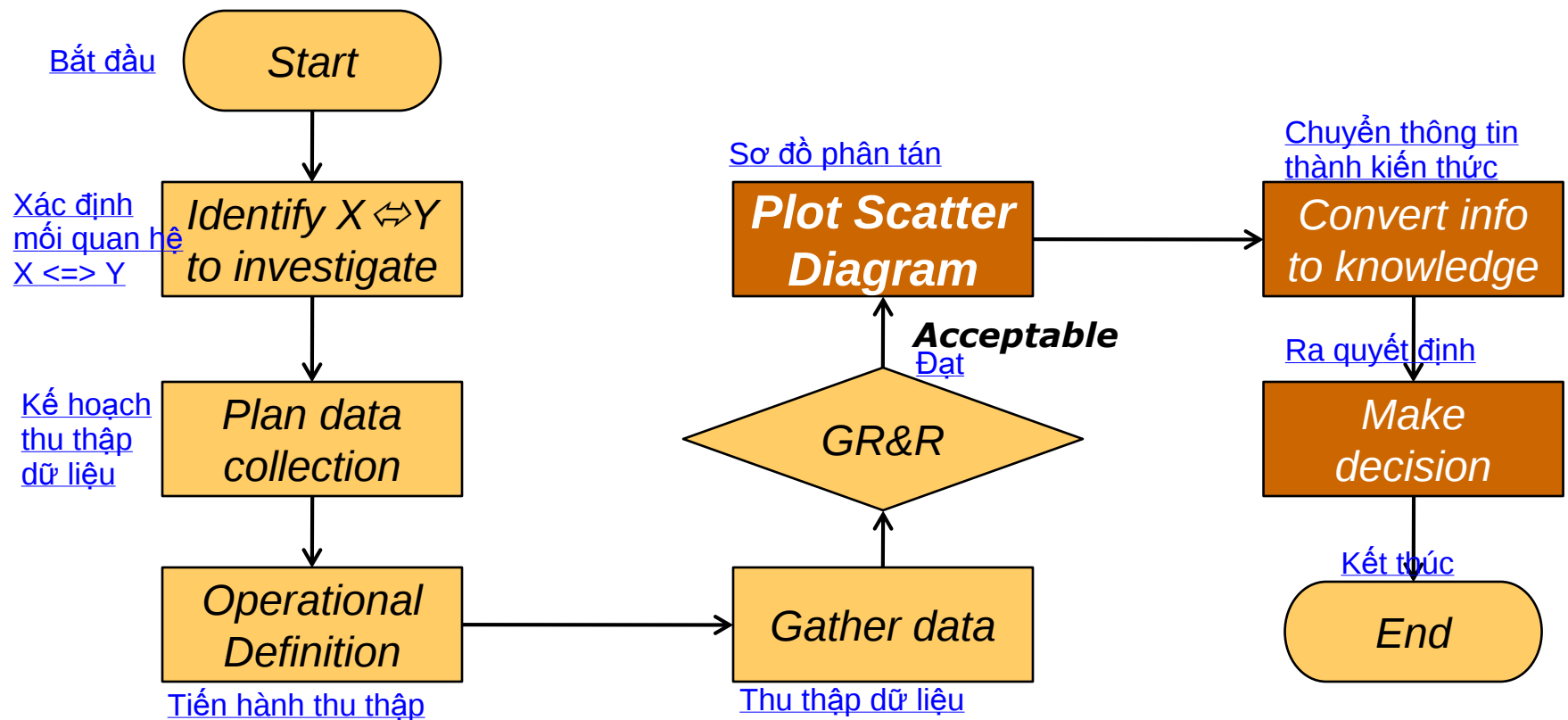


If taking prescribed medicine will cure the patient; does it mean when the patient is cured, the patient must have taken the prescribed medicine?

VD: uống thuốc theo kê đơn có thể cứu được bệnh nhân, thì không có nghĩa là khi bệnh nhân được cứu, thì bệnh nhân chắc hẳn đã uống thuốc theo kê đơn.



1-6. Application Steps





Discussion 1-3: Scatter Diagram Applications

◆ We can find out the relationship of the following variables:-

- Holiday package vs No of tickets sold (which one is X and which one is Y?)
- (give a few examples)
- .. Chúng ta có thể tìm ra mối quan hệ giữa các biến số sau:
 - Gói ngày nghỉ lễ và số lượng vé bán được (biến nào là X, biến nào là Y, ...)
 - Đưa ra vài ví dụ
 - ...
- ..



Summary of Correlation Analysis

Kết luận về Phân tích mối quan hệ

- ◆ What is the purpose of correlation analysis?

Mục đích của việc phân tích mối quan hệ?

- ◆ What does R value tell you in correlation analysis?

Giá trị R (hệ số tương quan) nói lên điều gì?

- ◆ What are the steps in carrying out correlation analysis?

Các bước để tiến hành công việc phân tích mối quan hệ?

- ◆ Where correlation analysis will lead you to?

Việc phân tích mối quan hệ dẫn bạn đến đâu?

- ◆ How to avoid spurious correlation from affecting your decision?

Làm thế nào để tránh được tương quan giả để không ảnh hưởng đến các phán đoán, kết luận?

Simple Linear Regression

HI QUY TUYẾN TÍNH NGUYỄN

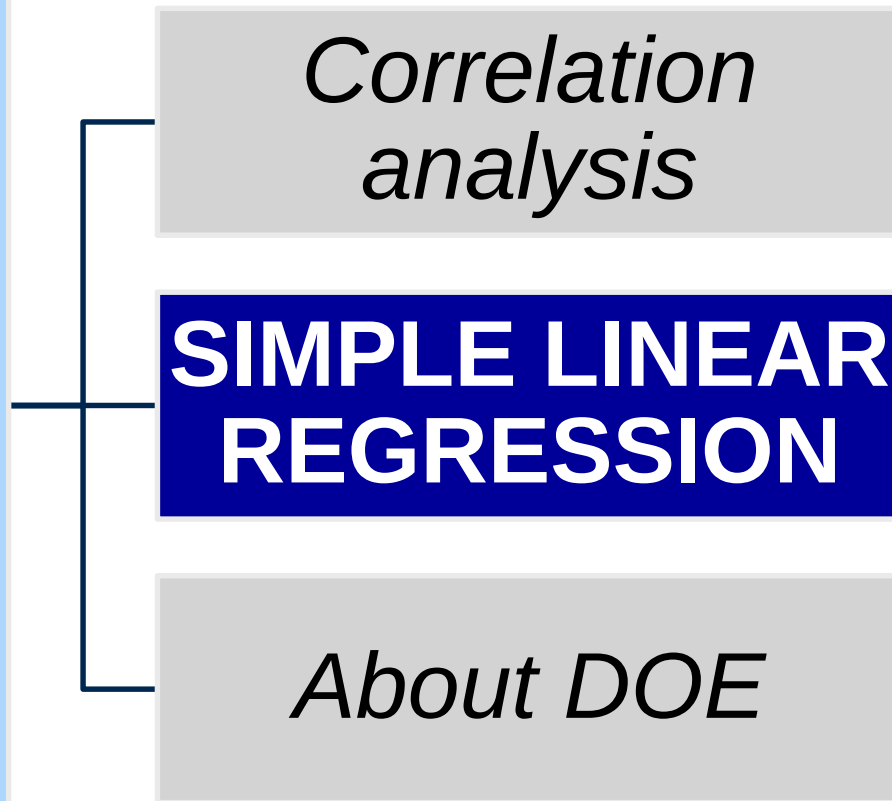


Innovation starts when you can model something...

KHI BẠN MÔ HÌNH HÓA MỘT ĐIỀU GÌ ĐÓ,
TỨC LÀ BẠN ĐANG SÁNG TẠO



MY FIRST DOE



1. *What is Regression Analysis?*
2. *When to Use Regression Analysis?*
3. *The Road to Regression*
4. *Why Use Regression Analysis?*
5. *Key Point in Analysis*
6. *Linear Regression Analysis*

1. Phân tích hồi quy là gì?
2. Khi nào sử dụng phân tích hồi quy?
3. Con đường dẫn đến hồi quy
4. Tại sao sử dụng phân tích hồi quy?
5. Điểm chính trong phân tích
6. Phép phân tích hồi quy tuyến tính



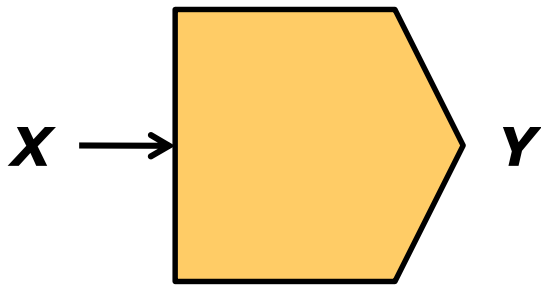
2-1. What is Regression Analysis?

2-1. Phân tích hồi quy là gì?

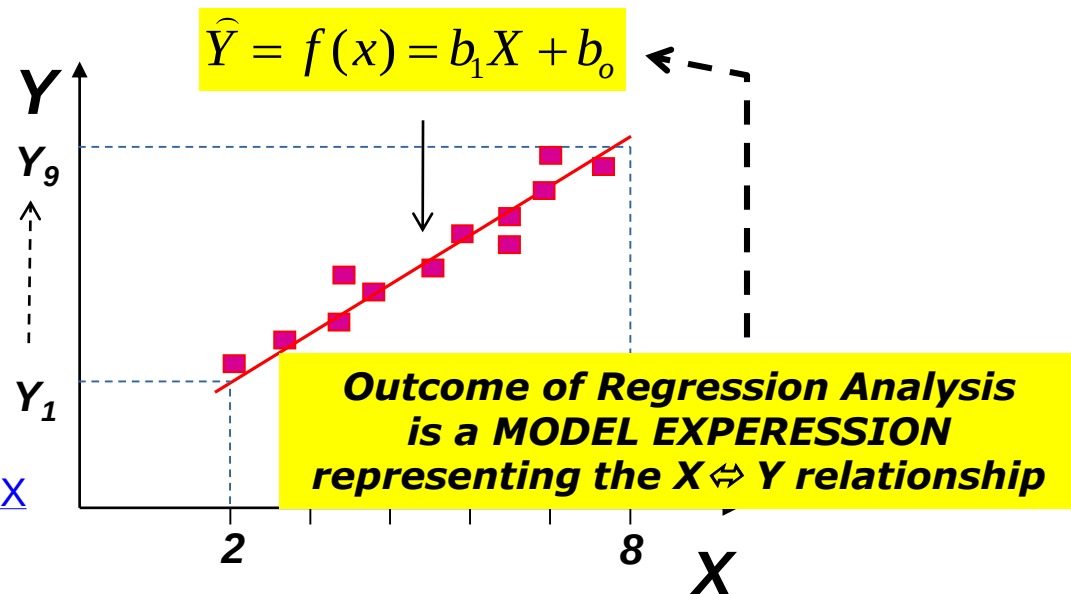
◆ Recall Simple regression analysis

Là kỹ thuật thống kê được sử dụng để phân tích mối quan hệ giữa hai hay nhiều biến số.

Regression analysis is a statistical technique used to analyze the relationship between two or more variables.



Kết quả của phép phân tích hồi quy:
công thức mô tả mối quan hệ giữa Y và X



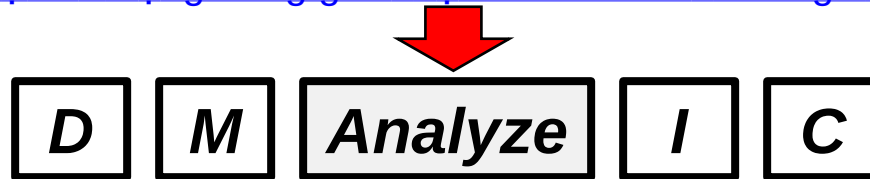


2-2. When to Use Regression

2-2. Khi nào sử dụng phép phân tích hồi quy? **Analysis?**

- Often used during the Analyse phase of a DMAIC approach projects

Thường được sử dụng trong giai đoạn Phân tích của công cụ DMAIC



Chỉ nên dùng 1 lần

- Should only be employed once:

2

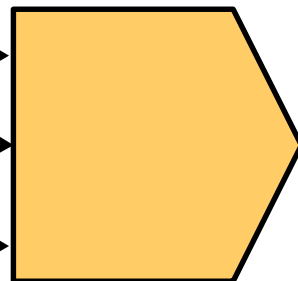
**The KPIVs
(Key Input Variables)
are identified**

Xác định các biến đầu vào

X_1 →

X_2 →

X_3 →



Y

1

**The KPOV
(Key Output Variable) is
identified**

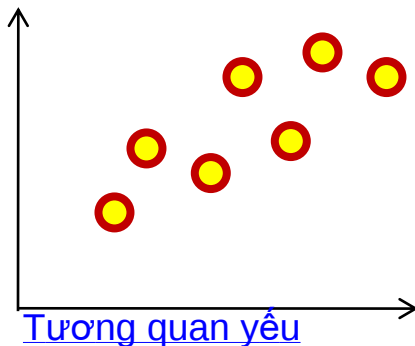
Xác định biến đầu ra



2-2. (cont) When to Use Regression Analysis?

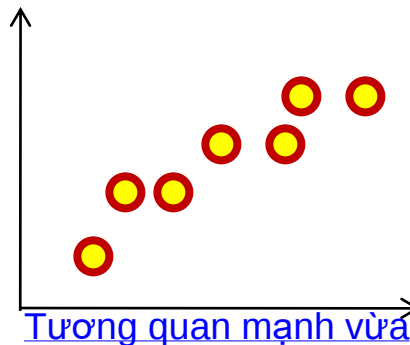
Biểu diễn mối quan hệ giữa Y và các biến X theo biểu đồ phân tán

- ◆ Any potential relationship between the Y and X's should always first be tested using a scatter plot



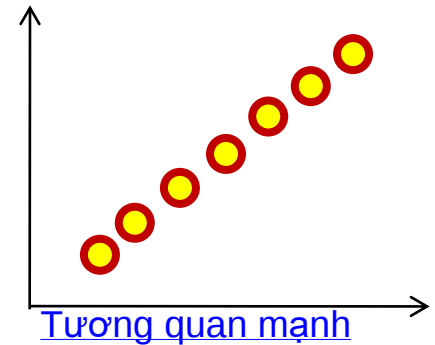
**Weak
Correlation**

$$r \leq 0.7$$



**Moderately
Strong
Correlated**

$$0.7 < r < 0.9$$



**Strongly
Correlated**

$$r \geq 0.9$$

***Which kind of relationship do you have?
Is it worth to proceed with Regression Analysis?***

Kết quả của bạn thu được là loại tương quan nào?
Sử dụng Phép phân tích hồi quy có đáng giá không?

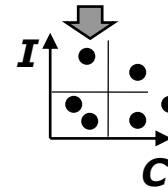


2-3. Phương pháp hồi quy **The Road to Regression**

Thiết lập hàm Y

1

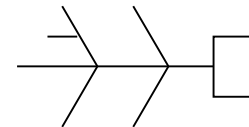
Establish Y's



Phân tích nguyên nhân và kết quả

2

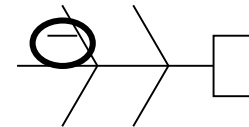
Cause & Effect Analysis



Xác định các biến số đầu vào tiềm ẩn

3

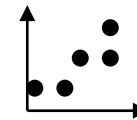
Identify Possible KPIVs'



Vẽ biểu đồ phân tán về sự phụ thuộc của kết quả đầu ra theo các biến đầu vào

4

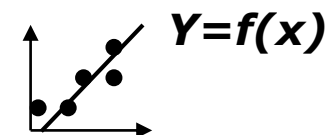
Scatter Plot KPOVs vs.. KPIVs



Phân tích hồi quy

5

Regression Analysis





2-41. Thử nghiệm giả thuyết **To Test Hypothesis**

Để kiểm tra xem giả thuyết về mối quan hệ X và Y

◆ To test an hypothesis about a relationship

- H_a (More tickets sold around Christmas) vs H_o (Most tickets sold in all festive seasons – no particular festive is significant). Which one is TRUE?





2-42. **To Make a Forecast**

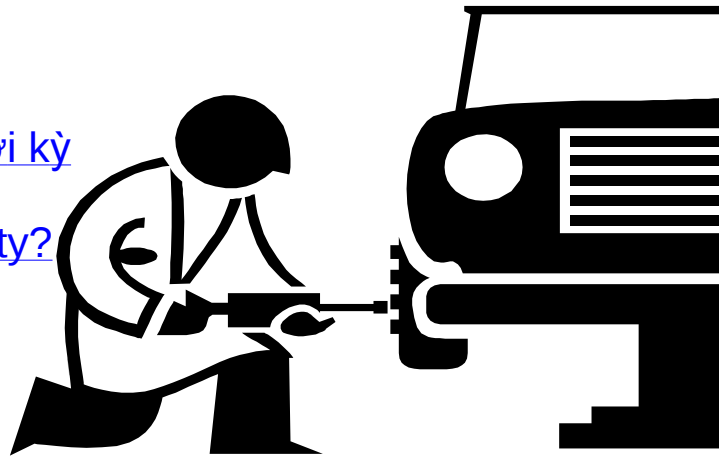
2-42. Dự đoán

Dự đoán dựa trên dữ liệu hữu ích thu được

◆ To make a forecast based on available data

- The Chief Engineer is making a prediction on company car service cost. He is keeping a detail record on cost of service against age of company cars. When a company car should be replaced?

Kỹ sư trưởng thực hiện dự báo về giá dịch vụ ô tô. Ông ấy giữ các tài liệu chi tiết về giá cả dịch vụ theo từng năm, từng thời kỳ của công ty.
Khi nào nên thay ô tô của công ty?





2-43. Setting KPOV on to a Target

2-43. Thiết lập KPOV hướng đến mục tiêu

Đặt các thông số đầu ra (giá trị trung bình) để đạt mục tiêu

◆ To set an output (average) to the target value

- We have only one driver [optimum Y] who can make 100 deliveries/d only. If there are >100 parcels/d we have to outsource the work.



Chúng ta chỉ có 1 người gửi hàng, người này chỉ có thể gửi 100 bưu kiện mỗi ngày. Nếu dung lượng > 100 bưu kiện mỗi ngày, chúng ta phải thuê ngoài.



2-5. Key Points in Data

2-5. Các điểm chính trong Phân tích Dữ liệu
(mục đích để phân tích hồi quy)

Analysis

(for Regression purpose)

◆ Always scatter-plot data first.

Luôn luôn vẽ biểu đồ phân tán trước

◆ Only do regression if it is meaningful (adds value) to the scatter plot.

Chỉ thực hiện phép hồi quy khi việc này có ý nghĩa đối với việc vẽ biểu đồ phân tán

◆ Don't forget that association itself is no guarantee of causation.



2-61. The Linear Regression

2-61. Mô hình hồi quy tuyến tính

Model

$$y = a + bx + \text{residuals}$$

Số dư

y – the *dependent variable* (KPOV) Biến độc lập

x – the *independent variable* (KPIV) Biến thụ động

Note that there are important assumptions about the Residuals.

Số dư: được giả định

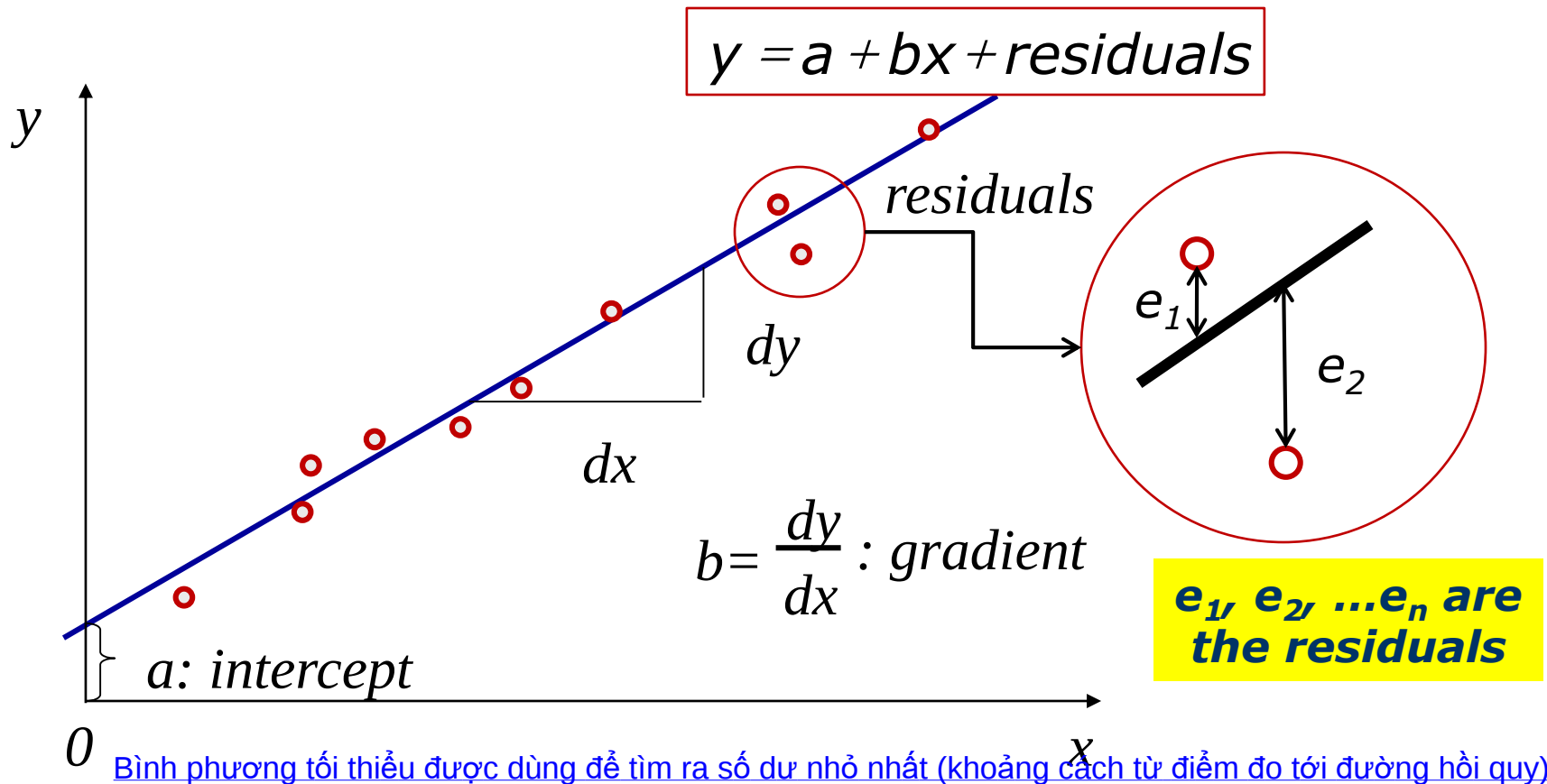
What are the Residuals?

Số dư là gì?



2-62. Linear Regression – A Graphical Illustration

2-62. Hồi quy tuyến tính - Đồ họa



The least square method is used to find a line that will minimize residuals $\sum_{i=1}^n e_i^2$ (the distances from the plotted points to the line)



The Linear Regression Model

Mô hình hồi quy tuyến tính

a: The intercept at Y axis (when x=0)

a: giá trị Y khi X=0

b: The slope of the graph

b: độ dốc của đường hồi quy

$$y = a + bx + residuals$$

$$b = \frac{N \sum XY - \sum X \sum Y}{\left(N \sum X^2 - (\sum X)^2 \right)}$$

$$a = \frac{\sum Y - b \sum X}{N}$$



1-63. Regression Analysis

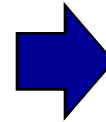
1-63. Các bước áp dụng hồi quy

Steps

Vẽ biểu đồ
phân tán

1

*Scatter –
Plot data*



MSA

Tìm hàm số
phù hợp nhất

2

*Calculate
best-fit line*

Phân tích số dư

3

Analyze residuals

Ước lượng
các điểm
không chắc chắn
trong phân tích

4

*Estimate uncertainty
in analysis*

Sử dụng
kết quả hồi quy

5

*Use the
regression result*

**MSA exercise is not
shown here – it is
assumed to have
been carried out
whenever necessary**



Simple Linear Regression

Tốc độ chảy (g/s)	Độ dày (mm)
Flow Rate (g/s)	Thickness (mm)
x	y
5.0	1.43
5.1	1.47
5.2	1.45
5.3	1.48
5.4	1.57
5.5	1.59
5.6	1.58
5.7	1.65
5.8	1.65
5.9	1.70

Bảng số liệu trên là khi thay đổi tốc độ dòng chảy thì độ dày của film (màng mỏng) thay đổi theo. Làm theo các bước trong trang sau, và hoàn thành Phân tích hồi quy

These observations were made by deliberately adjusting the flow-rate and measuring the consequent thickness of the film (it is like doing 1 factor experiment).

Follow the steps shown in next slides, and complete the Regression Analysis

Note that the MSA on thickness and flow rate were completed satisfactorily

Chú ý: MSA trên độ dày và tốc độ dòng chảy đã hoàn toàn đạt y/c

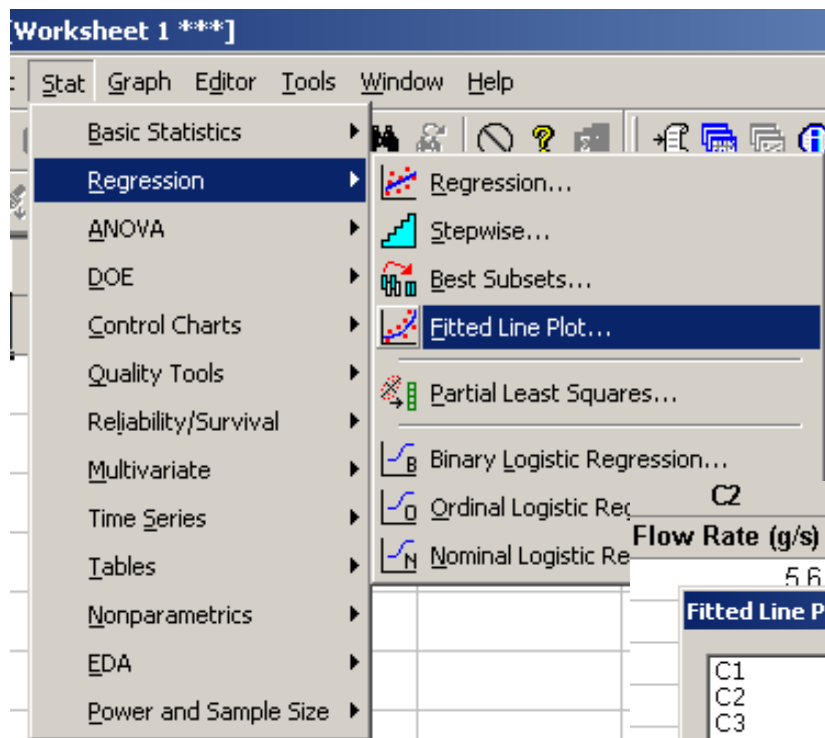


Obtaining a Fitted Line Plot

Sử dụng Minitab để tìm được đồ thị hàm số phù hợp

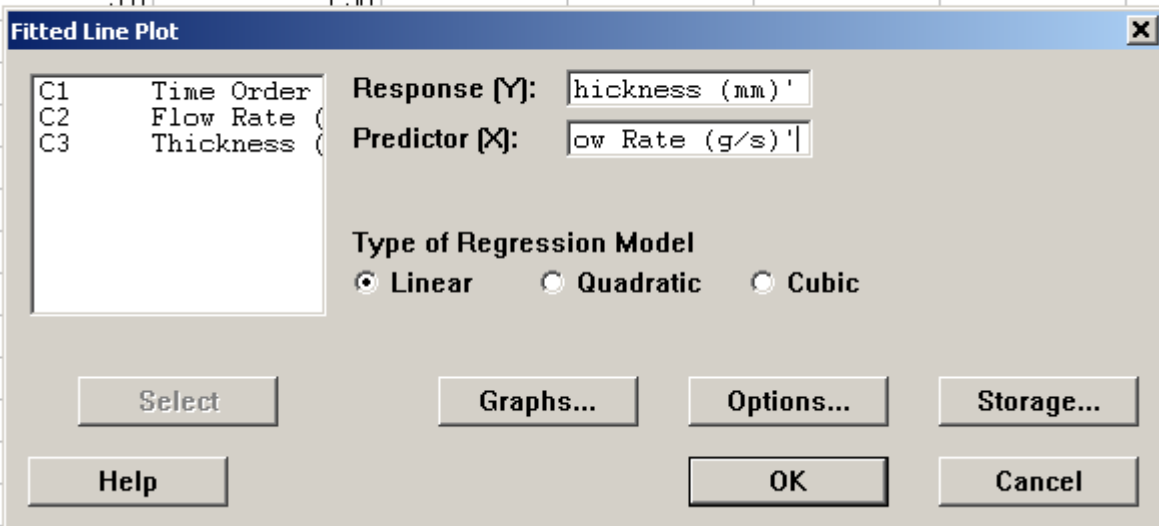
Using Minitab

Minitab can easily produce a fitted line plot - note the choice of regression models that can be used (the resin data will be used here to illustrate the method)



Sử dụng Minitab có thể dễ dàng tìm ra đồ thị hàm số phù hợp

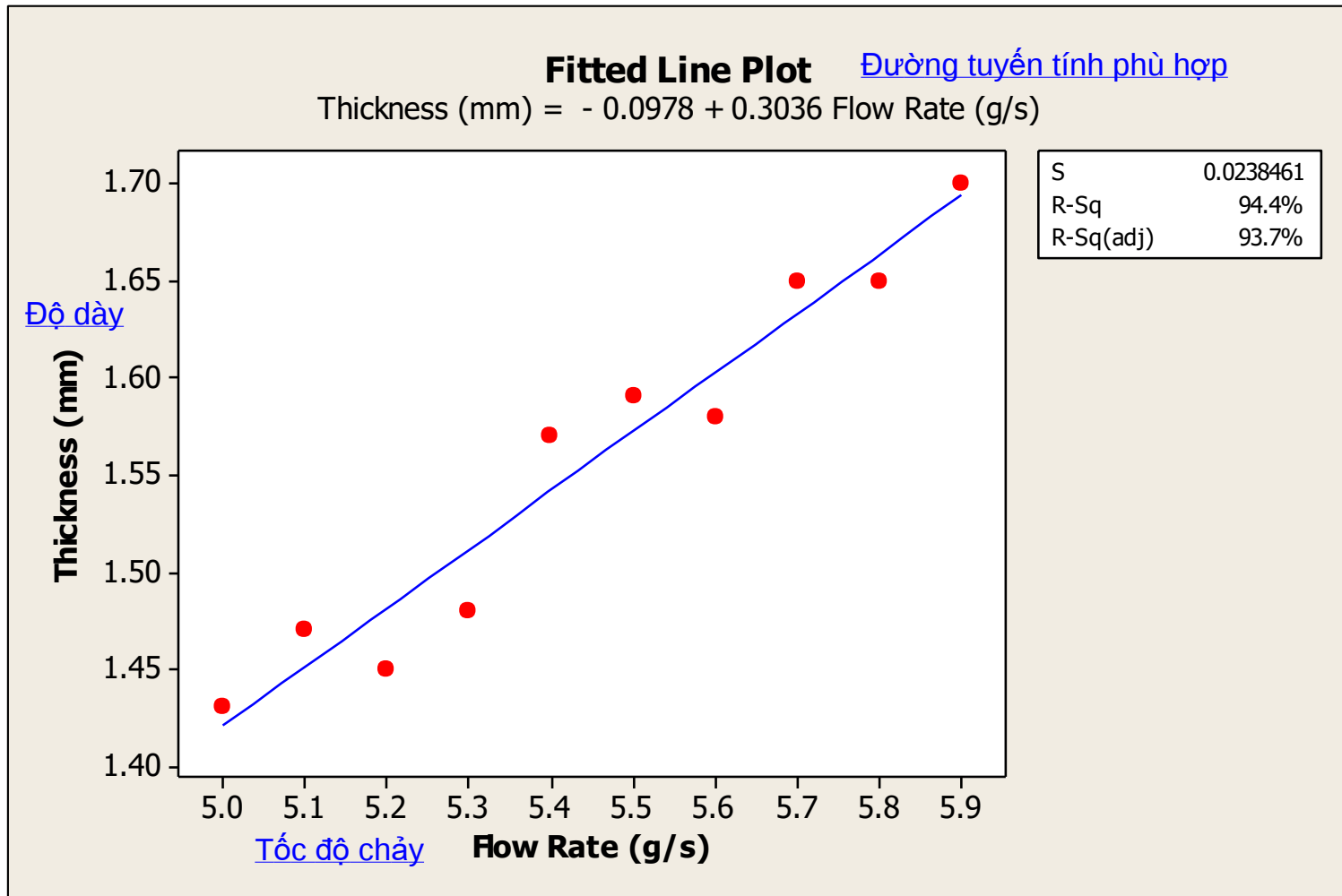
Lưu ý lựa chọn mô hình hồi quy





2-631. Fitted Line Plot – Using Minitab

2-631. Sử dụng Minitab để tìm ra đồ thị hàm số phù hợp





2-632. Regression – Using Minitab

2-632. Sử dụng Minitab - Phép hồi quy

MINITAB - Untitled - [EX-5-RESIN VS FLOW RATE.MTW ***]

File Edit Data Calc Stat Graph Editor Tools Window Help

Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables
Nonparametrics
EDA
Power and Sample Size

Regression...
Stepwise...
Best Subsets...
Fitted Line Plot...
Partial Least Squares...
Binary Logistic Regression...
Ordinal Logistic Regression...
Nominal Logistic Regression...

	C1	
	Time Order	Flow
1	1	
2	2	
3	3	
4	4	
5	5	
6	6	
7	7	
8	8	
9	9	
10	10	5.1 1.47
11		
12		
13		
14		
15		
16		
17		

Note: data entered in time order here

Scatter – Plot data

Vẽ biểu đồ phân tán

Calculate best-fit line

Tính toán đường hàm số phù hợp

Analyze residuals

Phân tích số dư

Estimate uncertainty in analysis

Ước lượng không chắc chắn

Use the regression result

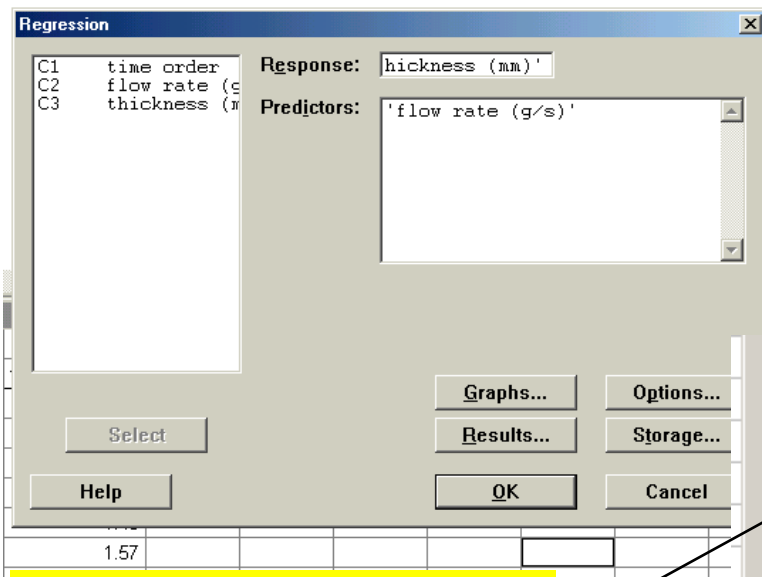
Sử dụng kết quả hồi quy



2-632. (cont) Regression –

2-632. Sử dụng Minitab - Phép hồi quy

Using Minitab

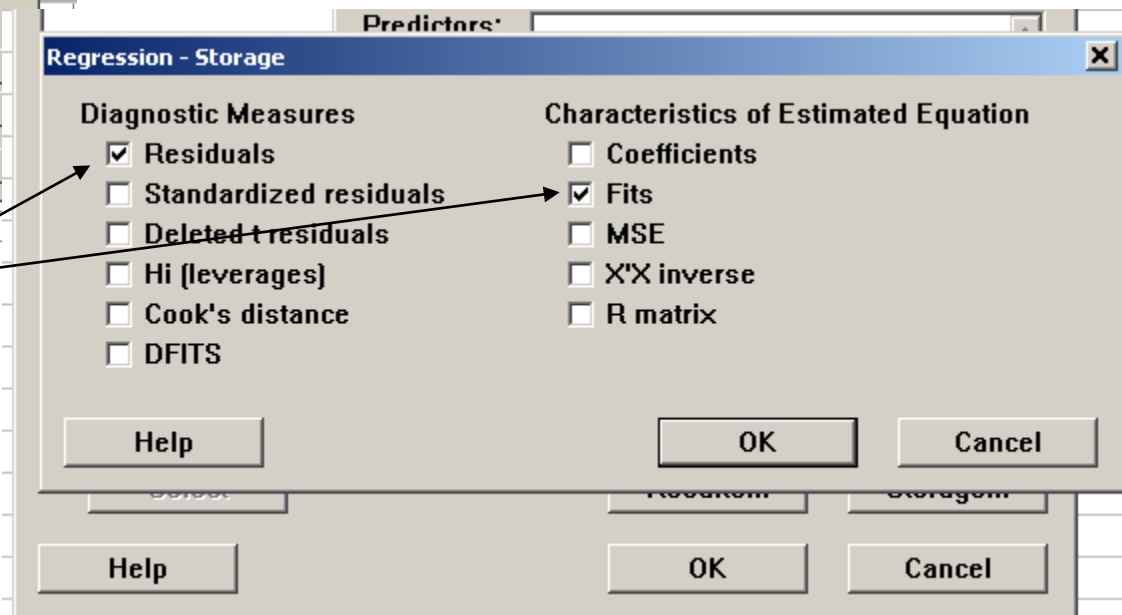


The regression dialogue boxes need to be filled in as shown

Phải điền đủ thông tin trong Hộp hội thoại hồi quy

Check these boxes for later use

Kiểm tra các box này sau khi sử dụng





2-633. Minitab's Session Window

2-633. Cửa sổ Minitab

Regression Analysis [Phân tích hồi quy](#)

The regression equation is
thickness (mm) = - 0.098 + 0.304 flow rate (g/s)

[Phương trình độ dày và tốc độ dòng chảy](#)

Predictor	Hệ số Coef	StDev	T	P
Constant	-0.0978	0.1433	-0.68	0.514
flow rat	0.30364	0.02625	11.57	0.000

S = 0.02385

R-Sq = 94.4%

R-Sq(adj) = 93.7%

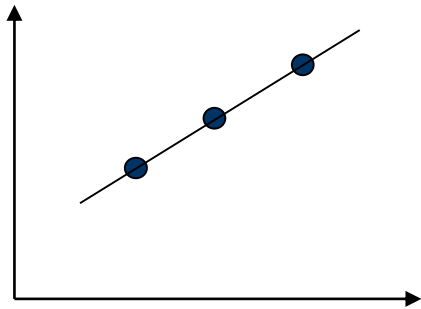
Analysis of Variance [Phân tích biến số](#)

Nguồn	Source	DF	SS	MS	F	P
Hồi quy	Regression	1	0.076061	0.076061	133.76	0.000
Lỗi số dư	Residual Error	8	0.004549	0.000569		
Tổng	Total	9	0.080610			

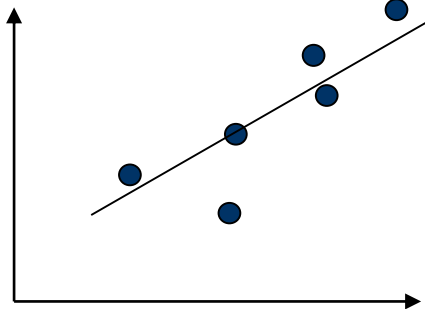


2-64. R^2 Value – *Example*

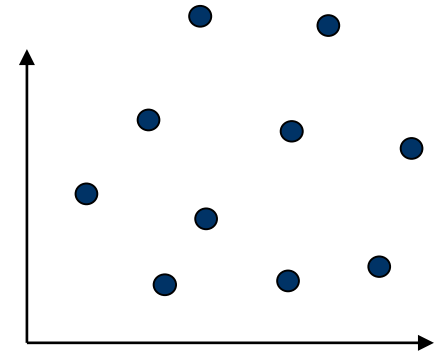
2-64. Giá trị R^2 - Ví dụ



$$R^2 = 1$$



$$R^2 = .8$$



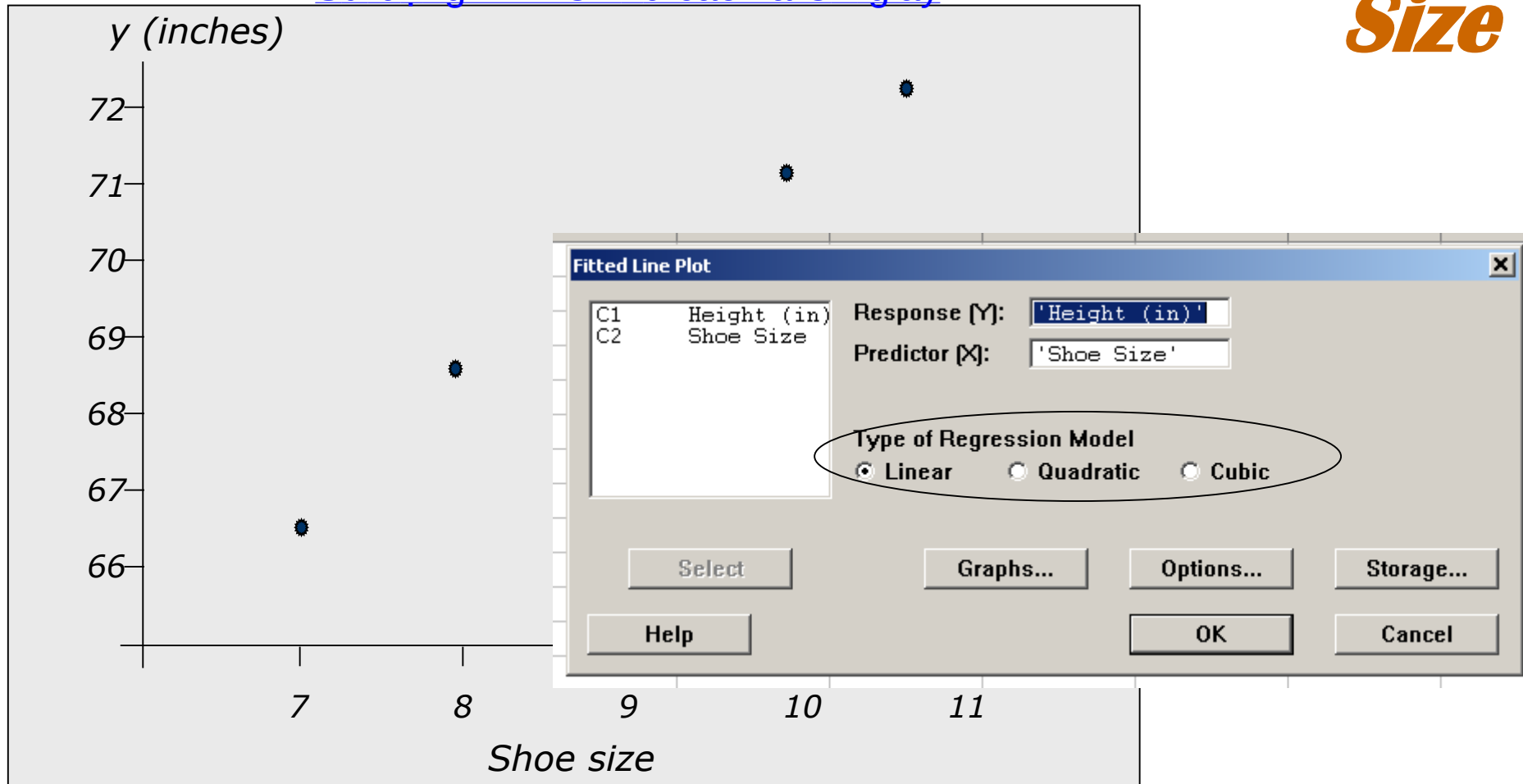
$$R^2 = 0$$



Using R^2 (adj) – Height vs Shoe

Sử dụng R^2 - Chiều cao và Cỡ giày

Size



Which is the best model for the height versus shoe size data?

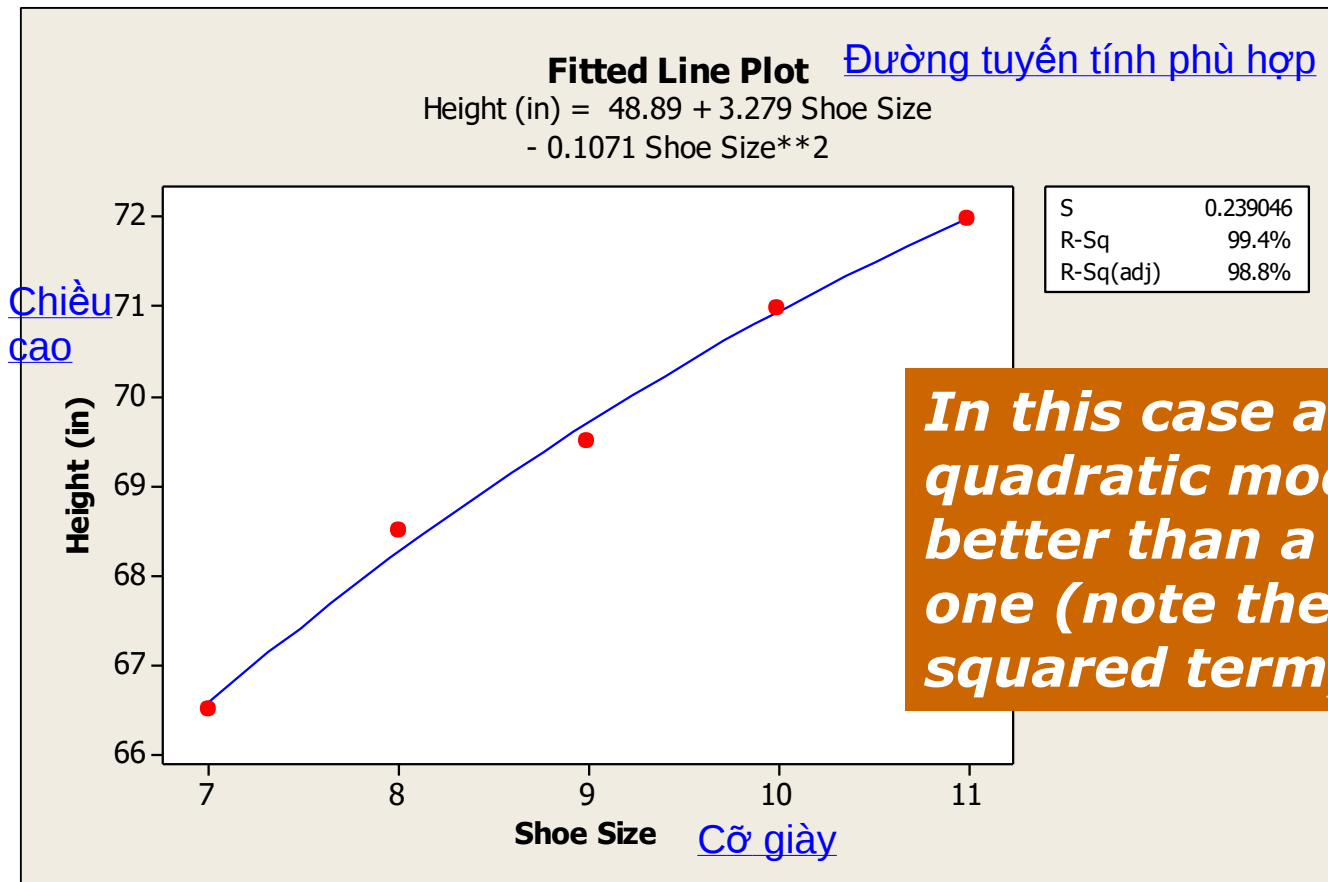
Mô hình nào tốt nhất cho dữ liệu về chiều cao và cỡ giày?



Height vs Shoe Size – A

Chiều cao và cỡ giày - Mô hình toàn phương

Quadratic Model



Sử dụng mô hình toàn phương tốt hơn hàm số tuyến tính



2-65. Khoảng cách tin cậy và
khoảng cách giới hạn

2-65. Confidence and Prediction Intervals (*Fitted Line Plot*)

Fitted Line Plot

Response [Y]: 'Height (in)'
Predictor [X]: 'Shoe Size'

Type of Regression Model
☐ Linear ☒ Quadratic ☐ Cubic

Options... Storage...

Fitted Line Plot - Options

Transformations
☐ Logten of Y ☐ Display logscale for Y variable
☐ Logten of X ☐ Display logscale for X variable

Display Options
☒ Display confidence bands
☒ Display prediction bands

Confidence level: 95.0

Title:

Help OK Cancel

	C1	C2	C3	
	Height (in)	Shoe Size	Shoe Size Sqd	shoe size
1	66.5	7	49	
2	68.5	8	64	
3	69.5	9	81	
4	71.0	10	100	
5	72.0	11	121	
6				
7				
8				
9				

Re-run the 'Fitted Line Plot', but this time select 'Options' and check the boxes shown

Vẽ lại Đường đồ thị phù hợp nhất
Nhưng lần này, chọn "options" trong hộp hội thoại ở bên



(*cont*) Confidence and Prediction Intervals



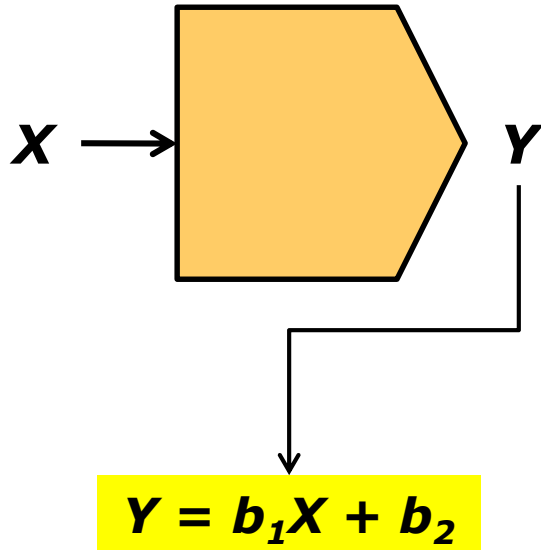


2-66. So What is Regression Modeling?

2-66. Mô hình hồi quy là gì?

Single Regression MODELING

Hồi quy đơn biến

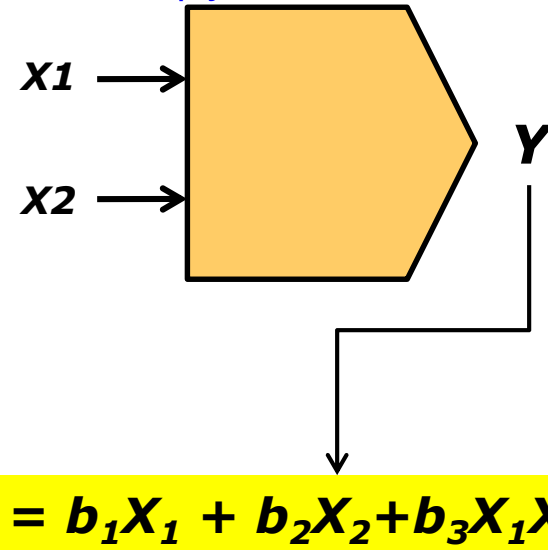


**Use the Regression
technique you
had just learned**

Sử dụng Phép hồi quy vừa học

Multiple Regression MODELING

Hồi quy đa biến



**Select Multiple
Regression feature
in Minitab**

Sử dụng Hồi quy đa biến trên Minitab



Exercise 2-2: *Catapult Shootings*

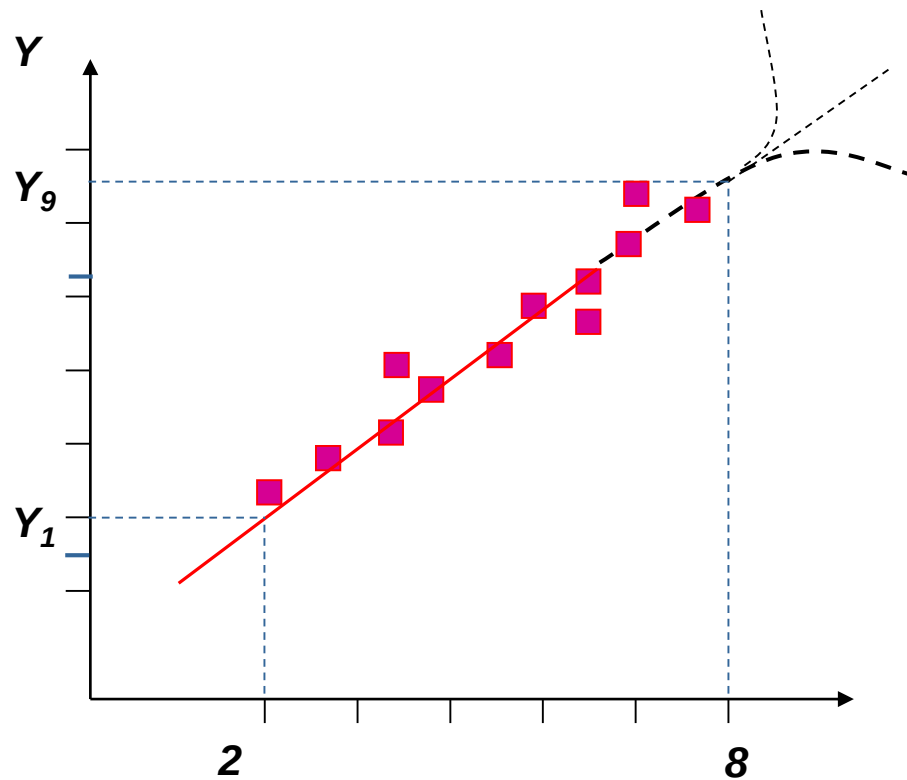
- ◆ Let's build a simple regression model for the catapult shown below. Analyze how does pullback angle affect the distance the projectile travels. The facilitator will demonstrate the catapult shooting and give detail explanation of the tasks.



Hãy xây dựng một mô hình hồi quy đơn giản cho máy bắn đá dưới đây. Phân tích xem góc kéo ra phía sau tác động thế nào đến đường đi của đạn đá. Học viên giải thích chi tiết cho mô hình này.



2-67. Phép ngoại suy **Beware of Extrapolation**

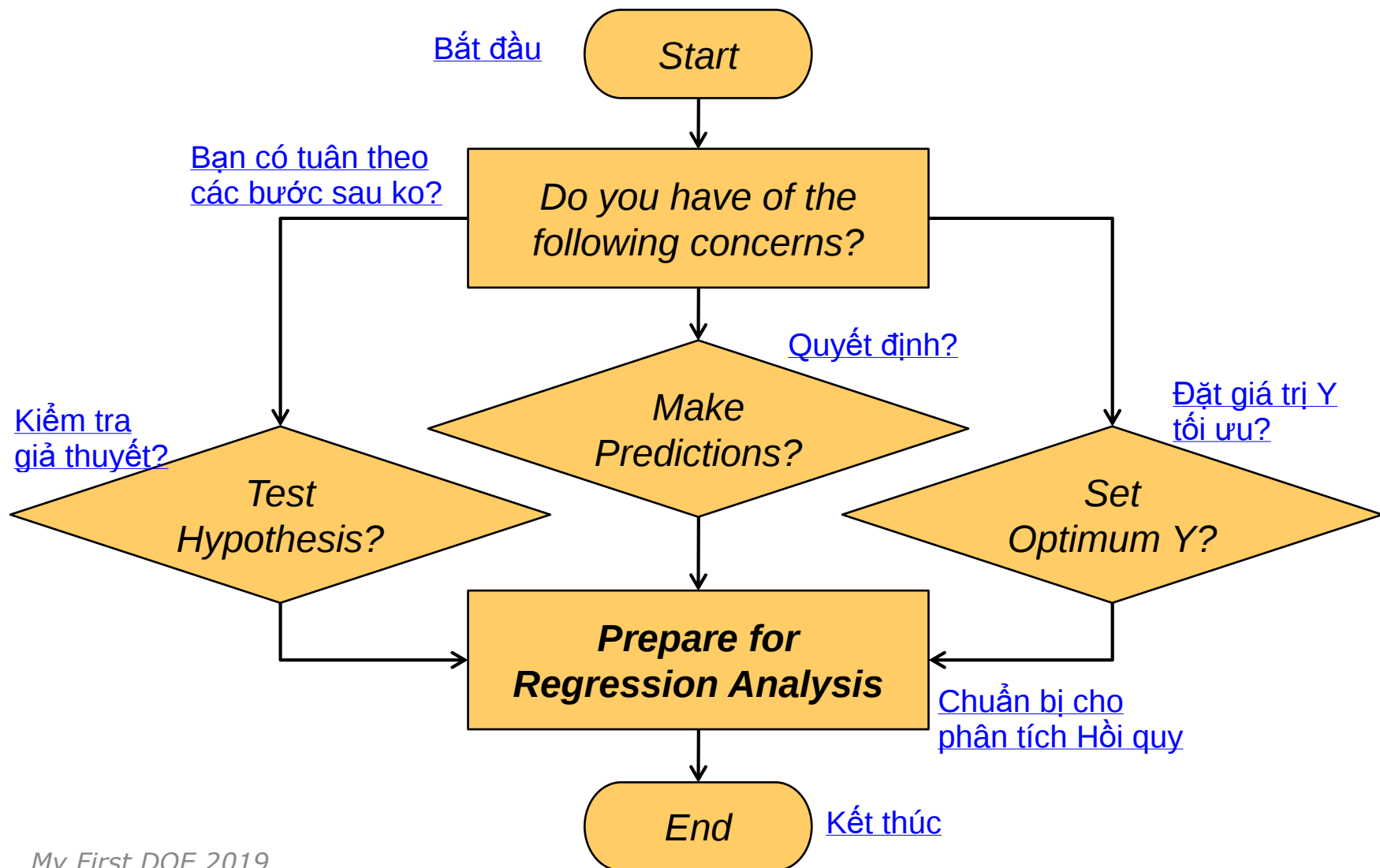


Liệu đường hồi quy có đúng với $X > 8$?

Does the regression line holds true for $X > 8$?



Thảo luận 2-1: Các bước áp dụng *Discussion 2-1:* *Application Steps*





Summary of Simple Linear Regression

Tổng kết về Phép hồi quy tuyến tính

- ◆ What is the purpose of regression analysis?

Mục đích của phép phân tích hồi quy?

- ◆ What does R^2 value tells you in regression analysis?

Giá trị R^2 có ý nghĩa gì trong phân tích hồi quy?

- ◆ How does correlation analysis link to regression analysis?

Phân tích tương quan có liên hệ gì với phân tích hồi quy?

- ◆ What is one the risk in regression analysis?

Rủi ro trong phân tích hồi quy?

- ◆ How many input variable are there in Simple Linear Regression?

Có bao nhiêu biến đầu vào trong Phân tích hồi quy tuyến tính đơn giản?

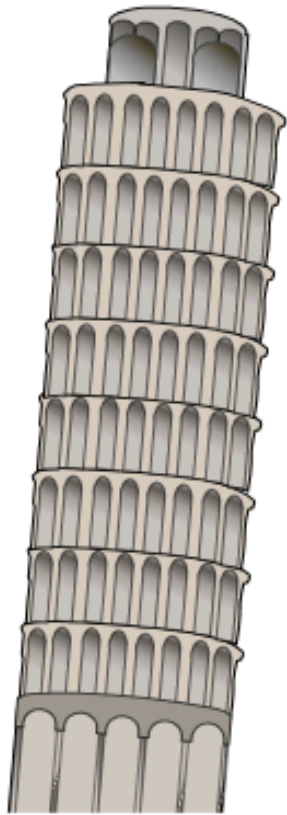
About DOE

VỀ DOE

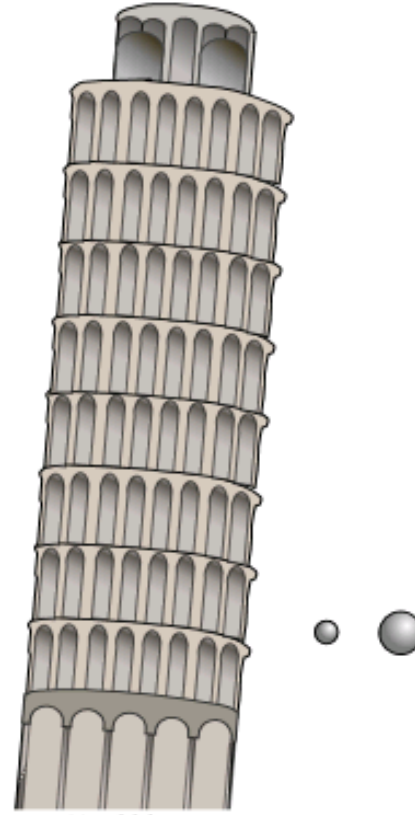


Theory that requires future generation to prove it...

LÝ THUYẾT ĐẶT RA VÀ ĐÒI HỎI THẾ HỆ TƯƠNG LAI CHỨNG MINH NÓ



Old idea



Galileo

MY FIRST DOE

Simple Linear regression

ABOUT DOE

Workshops

1. Think Experiments...

2. What is Design of Experiment?

3. DOE Learning Objectives

1. What will You Learn?

4. Purpose of Experiment

5. Components of DOE

6. Design Families

7. Procedures of Running DOE

8. Do's & Don'ts in DOE

1. Nghĩ về các thí nghiệm/thử nghiệm

2. Thiết kế thí nghiệm là gì?

3. Mục đích của việc học DOE

4. Mục đích của thí nghiệm

5. Các phần của DOE

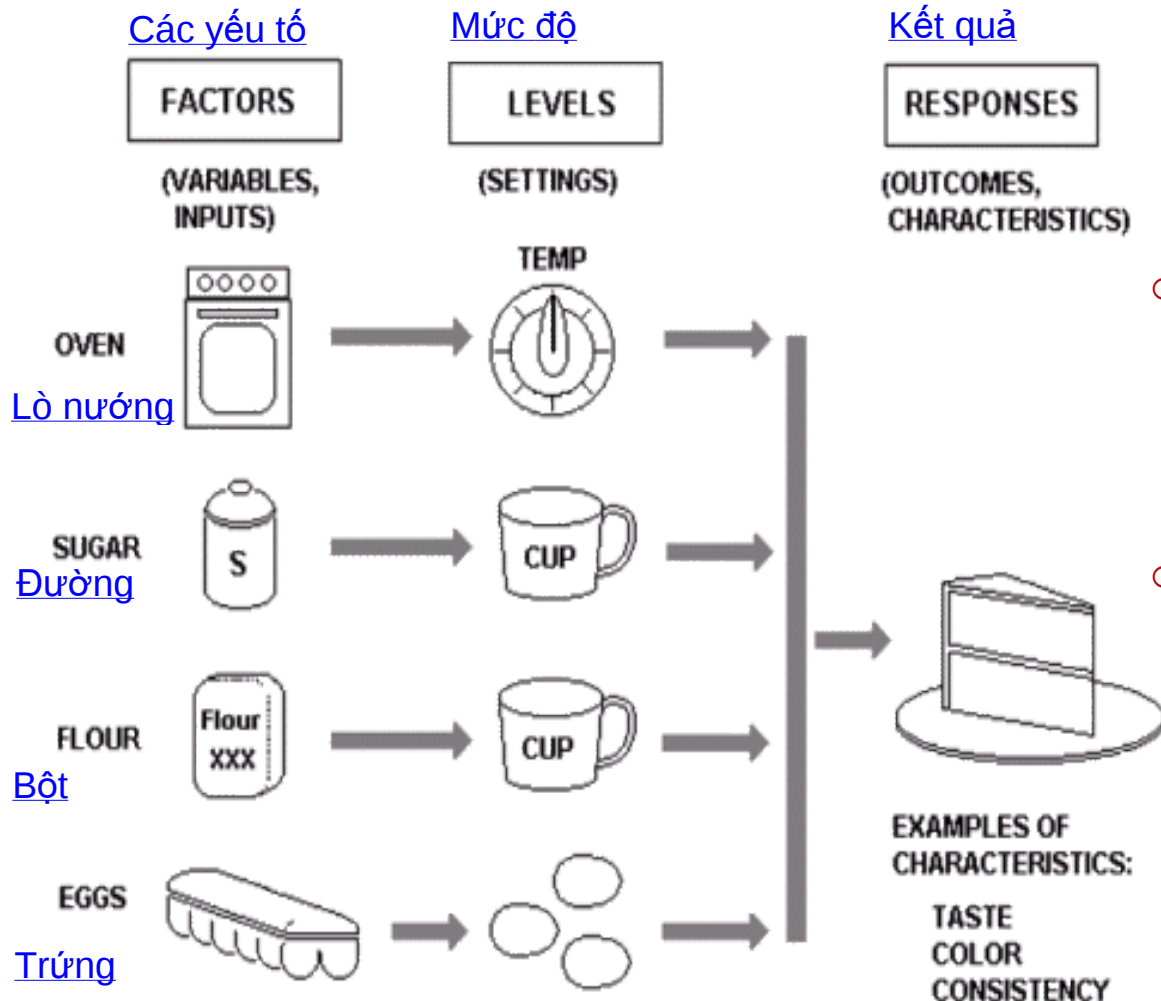
6.

7. Quy trình của chạy DOE

8. Nên và Không nên trong DOE



3-1. Nghĩ về các thí nghiệm **Think Experiments..**



Phải có thành phần đặc biệt, tuân theo công thức nghiêm ngặt để đạt được sự đồng đều mùi vị của các bánh

- ***There is a special ingredient ... must follow strictly for taste consistency***
- ***Let's make cakes that***
 - ***taste better***
 - ***softer***
 - ***..***

Hãy làm các bánh có:

- vị ngon hơn
- mềm xốp hơn
- ...



3-21. A Breakthrough Tool

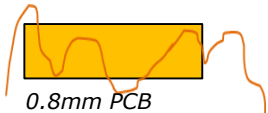
3-21. Công cụ đột phá

i Defects 90% reduced!

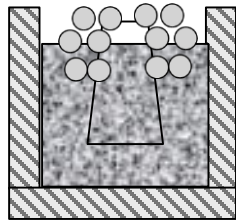
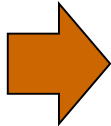
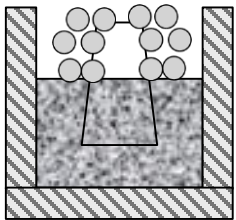
Giảm 90% khuyết tật

Quá trình hàn sóng

Solder wave performance



0.8mm PCB



Trước - 13 công nhân
BEFORE - 13 workers



Sau - 1 công nhân
AFTER - 1 worker



i Robust machine parameter setting

Thiết lập các thông số máy móc



i Widen paint application

Mở rộng các ứng dụng sơn

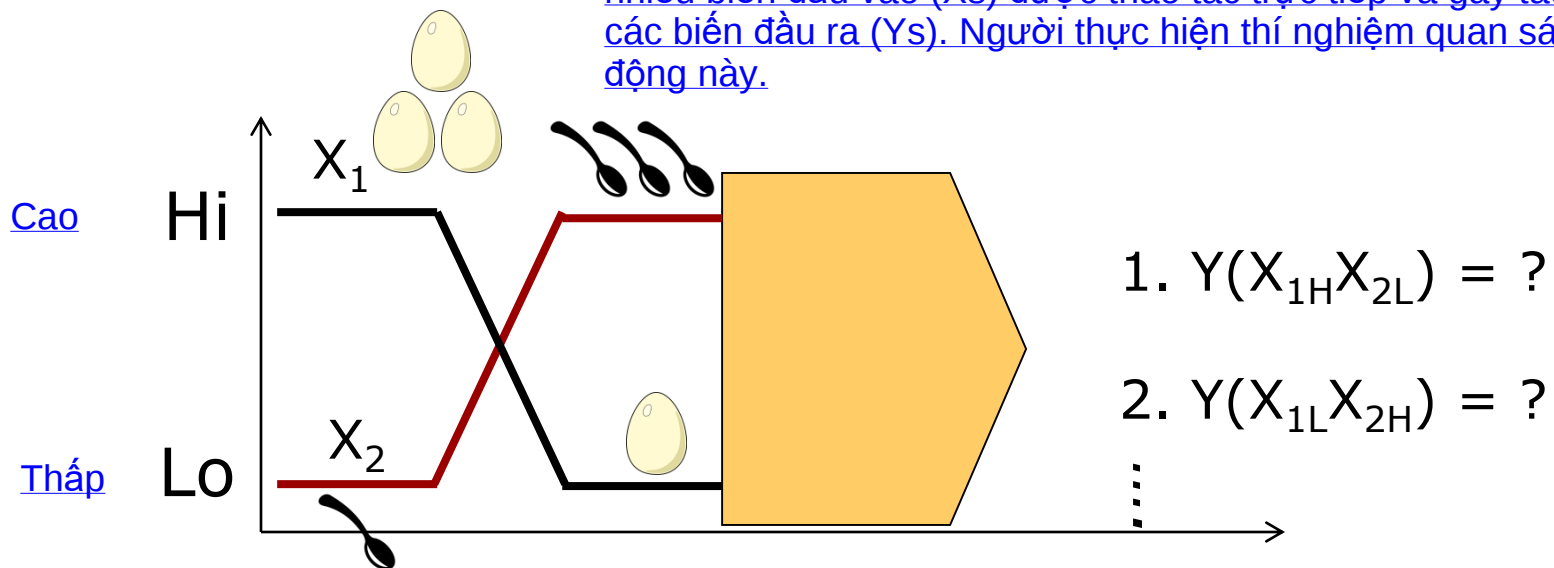


3-22. Systematic Series of Tests

3-22. Hệ thống các thí nghiệm

- ◆ A series of tests (systematically arranged) in which various Input Variables (X's) are manipulated directly, and the effects on the Output Variables (Y's) are observed.

Một chuỗi các thí nghiệm (được sắp xếp có hệ thống), trong đó có nhiều biến đầu vào (Xs) được thao tác trực tiếp và gây tác động lên các biến đầu ra (Ys). Người thực hiện thí nghiệm quan sát các tác động này.

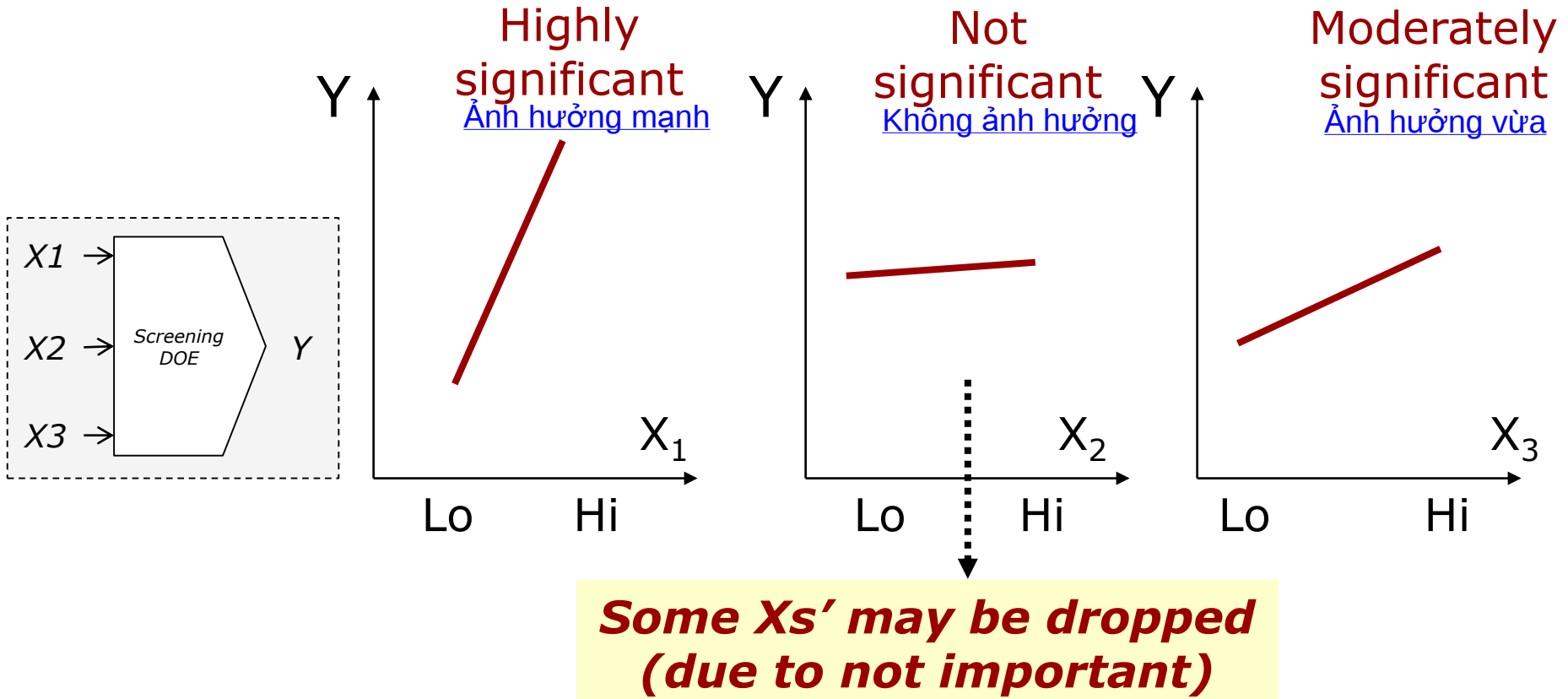




3-23. Xác định ảnh hưởng của Xs lên Y **Determine the Significance of Xs' to Y?**

Theo dõi thí nghiệm để xác định Xs nào ảnh hưởng lớn nhất lên Y

- ◆ To determine which Xs' most affect Y (average & variation)



3-24. Determine the Xs' Combination

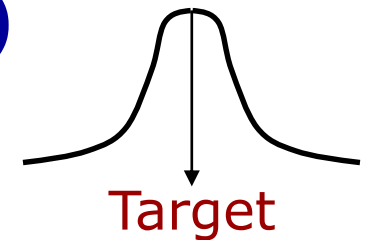
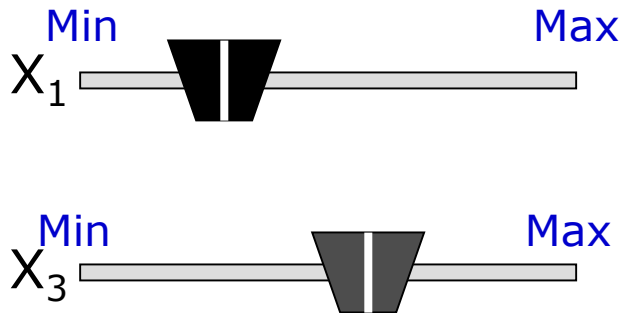
3-24. Xác định sự kết hợp của Xs

◆ Optimized setting for Tối ưu hóa để:

Tìm ra giá trị Xs mà tại đó giá trị Y đạt đúng mục tiêu Y-hat

- determining where to set the influential X's to center the Y on the target (**Y-hat**)

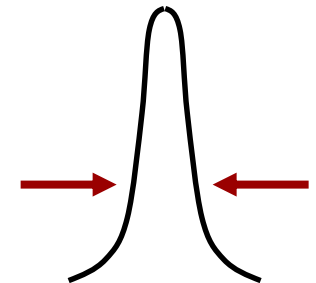
Optimized Setting



**Desired
Y-hat, S-hat**

- determining where to set the influential X's to minimize the variability of Y (**S-hat**)

Tìm ra giá trị Xs mà tại đó giá trị Y thấp nhất (S-hat)





3-3. DOE Learning Objectives

3-3. Mục đích học DOE

◇ Introduce participants the **concept of designed experiments.**

Giới thiệu cho học viên các khái niệm về thiết kế thí nghiệm

◇ **Provide** insights into ‘inference space’ and ‘experimental validity’.

Cung cấp cái nhìn sâu về "không gian giao thoa" và "hiệu lực thí nghiệm"

◇ Introduce **roadmap to planning an experiment.**

Giới thiệu cách thức xây dựng một thí nghiệm.



3-31. What will you Learn?

3-31. Học viên sẽ học gì?

◆ Concepts of experimental design (what)

Khái niệm về thiết kế thí nghiệm
(Cái gì?)

◆ Demonstration of doing DOE technique at workplace (how)

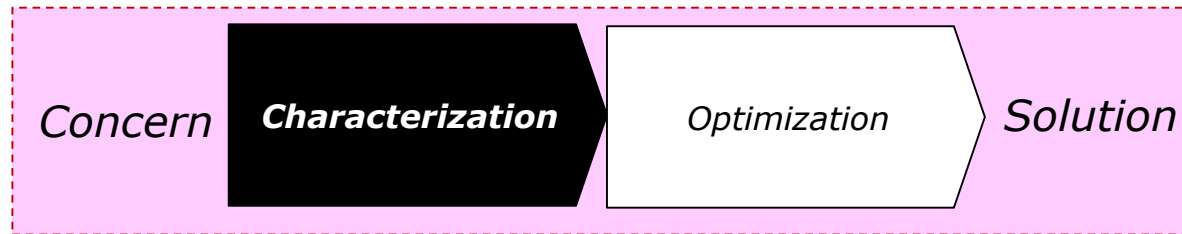
Thực hiện 1 kỹ thuật DOE (thiết kế thí nghiệm) tại nơi làm việc
(Làm như thế nào?)

◆ The in-depth knowledge of the tools (discussion only)

Kiến thức sâu hơn về các công cụ
(chỉ dành cho thảo luận)



3-41 Characterizing a Process



◆ To determine Để xác định:

Biến X nào gây ảnh hưởng lớn nhất lên kết quả Y

- **Which Xs' (controllable) most affect Y**

Tìm ra những biến nào cần điều khiển thận trọng trong quá trình (biến đó nhạy cảm, tới hạn, ...)

- Identifies variables that need to be carefully controlled in the Process (sensitive, critical, highly significant variables)

Ảnh hưởng của những nguyên liệu ít tốn kém lên tính năng sản phẩm

- The effects of less expensive materials on product performance

Nguồn gốc các biến trong quá trình

- Sources of variation in a process

Ảnh hưởng của sự thay đổi nguyên vật liệu lên độ tin cậy của sp

- The effects of material variation on product reliability

Ảnh hưởng của việc thay đổi máy móc/người vận hành máy lên sp

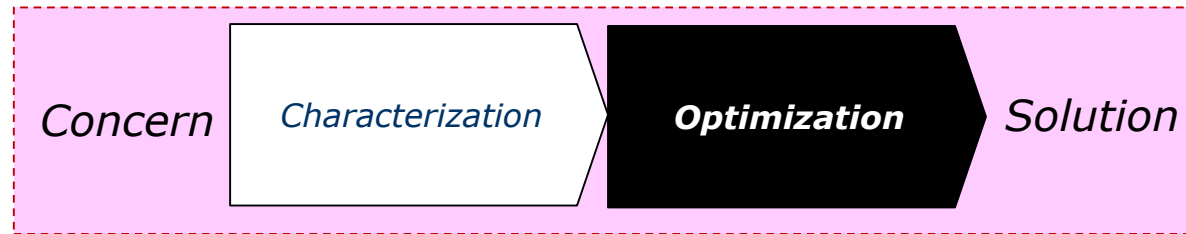
- The impact of operator variation on the product

Mối quan hệ (nhân quả) giữa đầu vào quá trình và đặc tính sp $[Y = f(x)]$

- **The relationships (cause & effect) between process inputs and product characteristics $[Y=f(X)]$**



3-42 Optimizing a Process



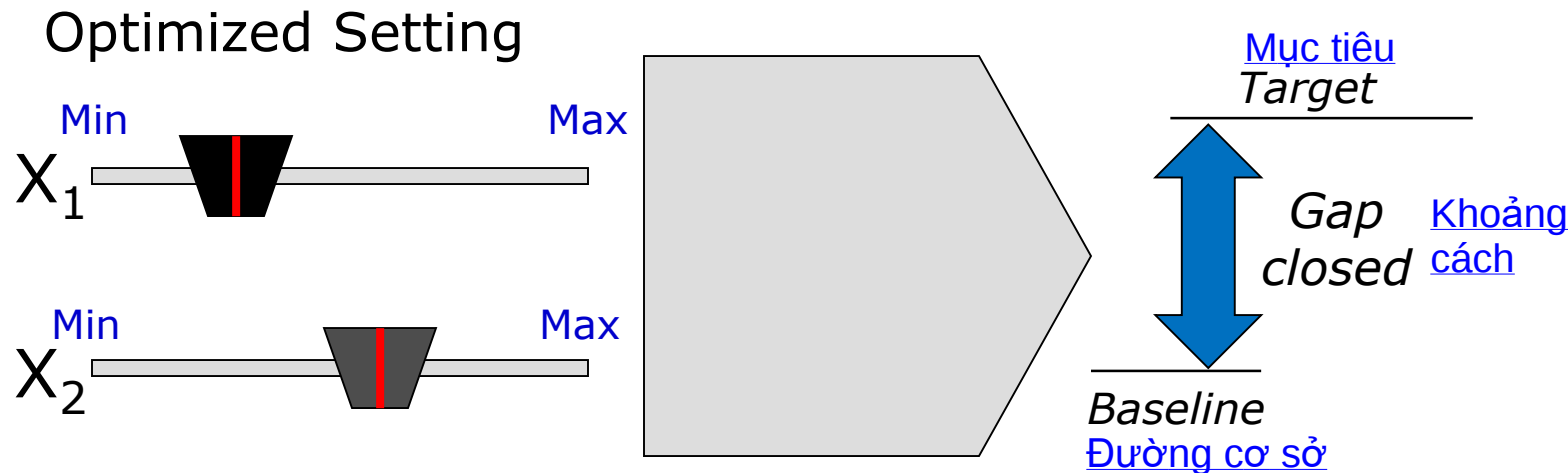
◆ To determine Để:

Tìm ra điểm đầu vào tới hạn cần thiết lập

- **Where the critical inputs should be set.**

Tìm ra phạm vi năng lực thực sự

- **The “real” capability ranges.**





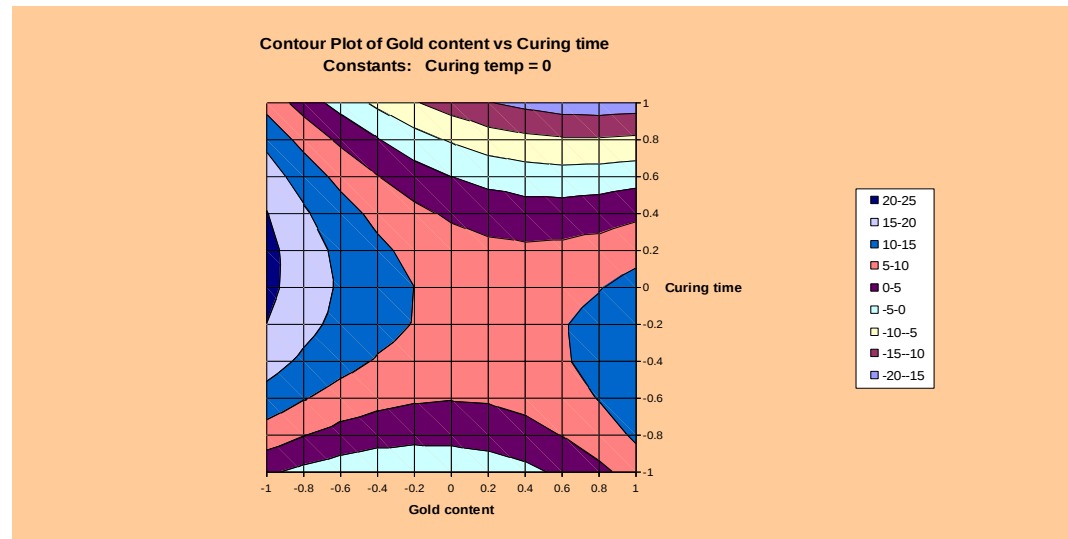
3-43. Hỗ trợ thiết kế sản phẩm **Product Design Aid**

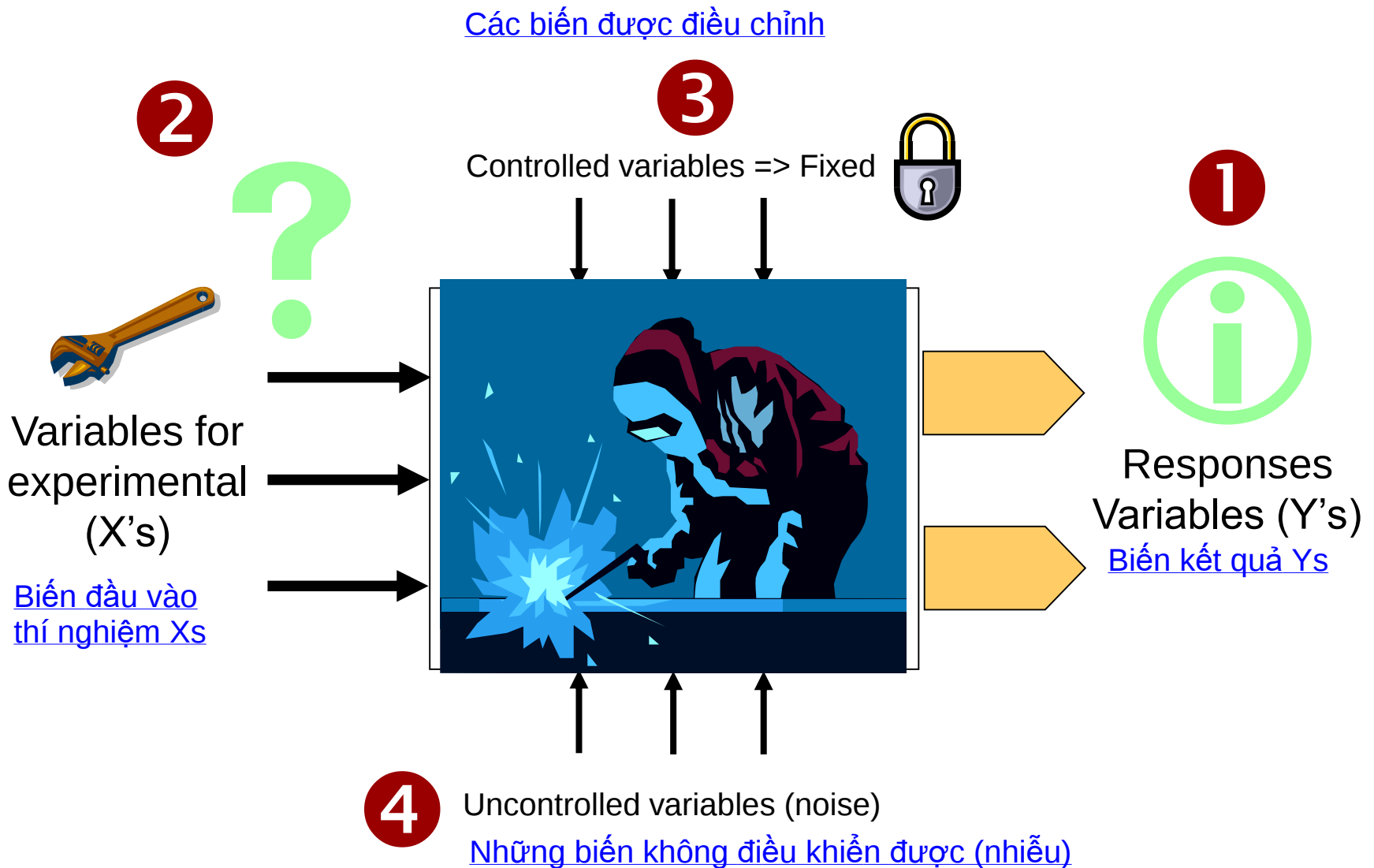
Hỗ trợ trong việc tìm hiểu các biến Xs ngay từ đầu quá trình thiết kế (phạm vi các thông số X)

- ◆ Aids in understanding X's early in the design process (**ranges of X parameters**).

Hướng dẫn thiết kế nhanh gọn (sử dụng phương pháp Taguchi)

- ◆ Provides direction for “robust” designs (**use of Taguchi method**).







3-521. Knob Variables

Biến knob là biến đầu vào có thể điều khiển được

◆ Controllable Input Variables = Knob Variables.

◆ About Knob Variables

Khi được điều chỉnh, có tác động (tích cực hoặc tiêu cực) lên kết quả đầu ra

- When adjusted, **have an effect** (positive or negative) on the output response(s).

Có thể không tác động lên các biến khác (biến độc lập)

- May **not** be affected by other variables (independent state).

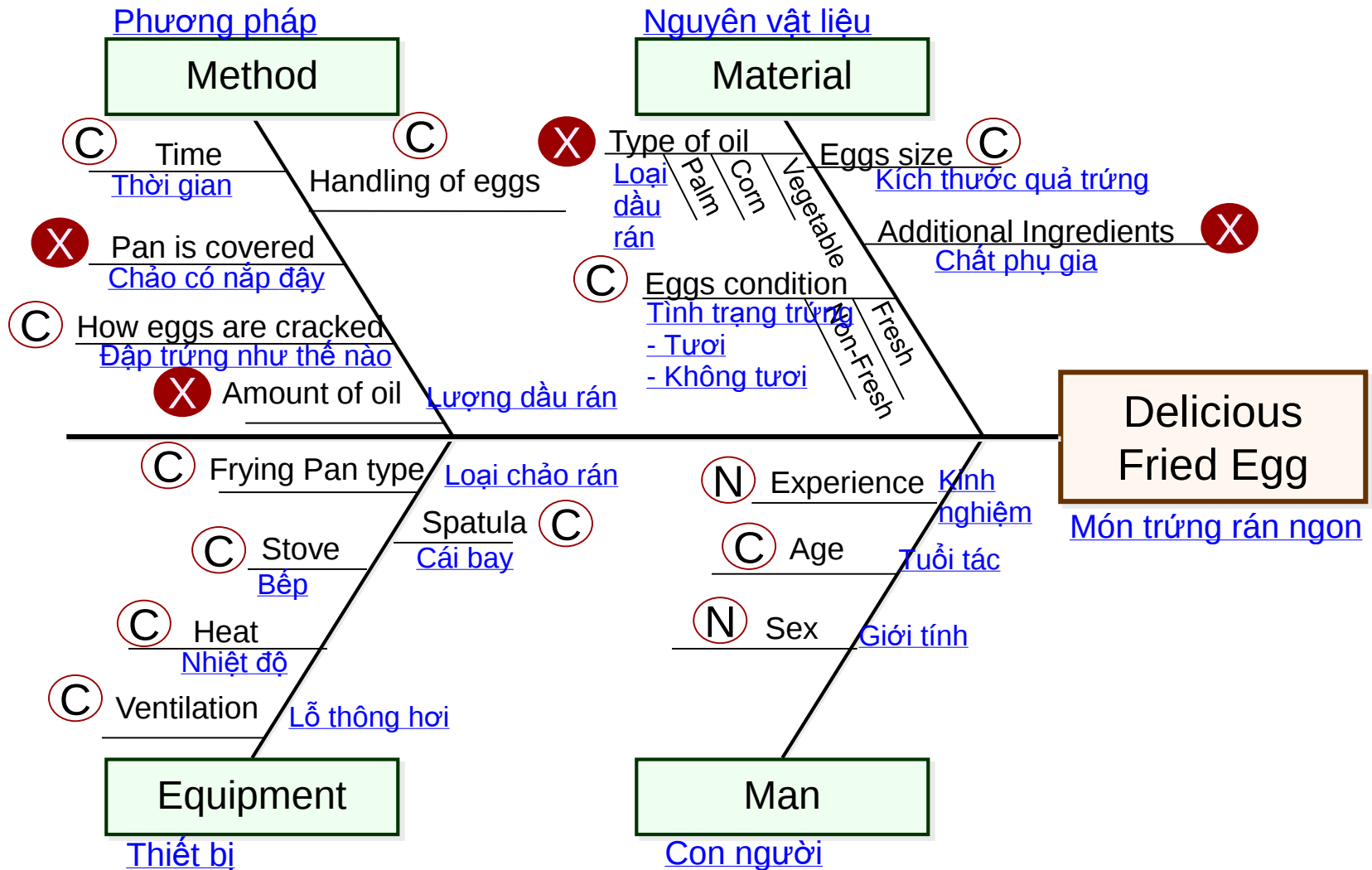
Có thể đo lường được, điều chỉnh được theo cách nào đó.

- **Can be measured**, adjusted or acted on in some way.



3-522. Generation of CEDAC- CNX

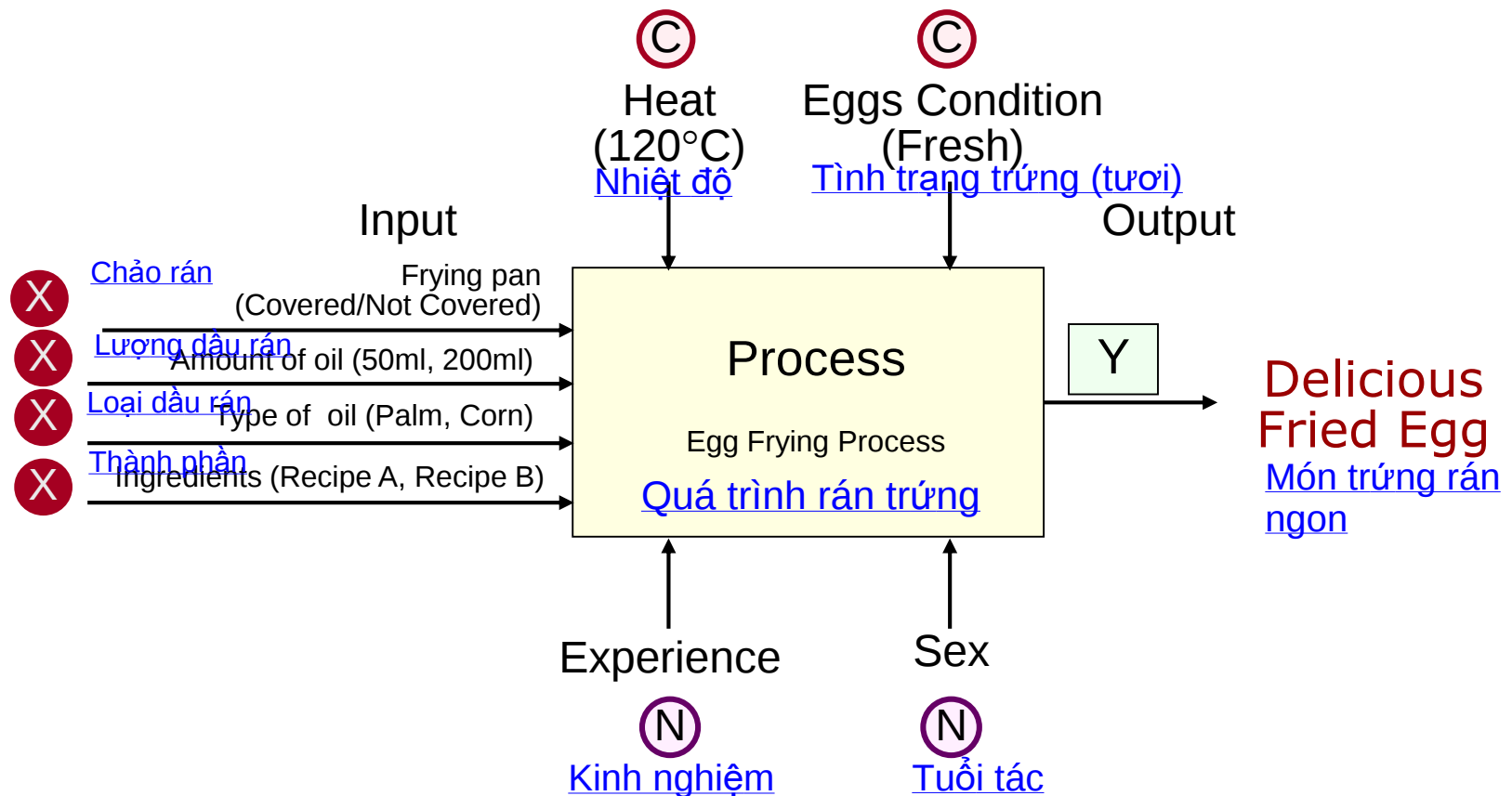
3-522. CEDAC-CN





Vẽ lại CEDAC trên sơ đồ IPO

◆ Classified CEDAC redrawn on IPO chart





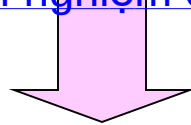
3-5241. About Experimental

3-5241. Các thông số thí nghiệm

Factors

Experimental factors are factors that are critical to the Y

Các thông số thí nghiệm có tác động tới hạn tới Y



INPUT

- X1 Chảo rán
Frying Pan
- X2 Lượng dầu rán
Amount of oil
- X3 Loại dầu rán
Type of oil
- X4 Thành phần
Ingredient

<u>Thấp/ Loại 1</u>		<u>Cao/ Loại 2</u>	
Lo/Cat 1		Hi/Cat 2	
→	Covered <u>Đậy nắp</u>	Not covered <u>Không đậy nắp</u>	
→	50ml	200ml	
→	Palm <u>Dầu cọ</u>	Corn <u>Dầu ngô</u>	
→	Recipe A <u>Công thức A</u>	Recipe B <u>Công thức B</u>	



3-5242. Coded Inputs

Khi tiến hành thí nghiệm, thì các biến đầu vào được mã hóa có giá trị trong khoảng -1 đến +1

- ◆ When experimenting the ranges in the input are coded with -1 and +1.

	Thấp/Loại 1	Cao/Loại 2		Thấp/Loại 1	Cao/Loại 2
	Lo/Cat 1	Hi/Cat 2		Lo/Cat 1	Hi/Cat 2
X1	Not Covered <u>Không đậy nắp</u>	Covered <u>Có đậy nắp</u>		-1	+1
X2	50ml	200ml		-1	+1
X3	Palm <u>Dầu cọ</u>	Corn <u>Dầu ngô</u>		-1	+1
X4	Recipe A <u>Công thức A</u>	Recipe B <u>Công thức B</u>		-1	+1

Example 3-1: How to Convert Coded $\leftrightarrow$ Actual Value?

- ◆ Let say we are testing a temperature from 40 ~ 80°C.

Về nhiệt độ 40 ~ 80°C

● (-1) ~ (+1) : 40°C ~ 80°C

Sau thí nghiệm nhiệt độ được tối ưu tới giá trị mã hóa là 0.8. Vậy giá trị thật của nhiệt độ là bao nhiêu?

- ◆ If after experiment, temperature is optimized at code value of 0.8. What is the actual temperature value?

Công thức tính giá trị nhiệt độ thật theo giá trị nhiệt độ mã hóa:

$$\text{Actual} = \left[\frac{Hi + Lo}{2} \right] + \left[\frac{Hi - Lo}{2} \right] \text{Coded}$$

Determine act
temp for code
0.3, -0.8)

$$= \left[\frac{80 + 40}{2} \right] + \left[\frac{80 - 40}{2} \right] 0.8$$

$$= [60] + [20] 0.8 = 60 + 16 = \underline{\underline{76^\circ C}}$$



3-5243. Xác định phạm vi. **Determination of Range**

Đầu vào rời rạc (được phân loại)

◆ Discrete inputs (by category)

Lựa chọn các thông số cần thí nghiệm.

- Choice of items to be tested.

Ví dụ: Loại 1 (-1) và Loại 2 (+1)

- *Example: Type 1 (-1) and Type 2 (+1).*

2 yếu tố

2 Factors

-1	+1
Type A	Type B
Part 1	Part 2
Round Tròn	Rectangular Chữ nhật

> 2 yếu tố

> 2 Factors

(wt variable output)

- Graphical analysis:
Pareto charts, Dotplots
- Stat test: **t-test, ANOVA**
- DOE: **Logistic DOE (for attribute output)**



3-5243. (cont) Determination of Range

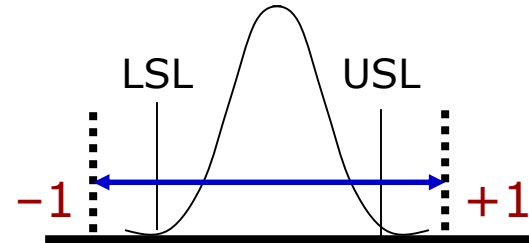
3-5243. Xác định phạm vi

Đầu vào liên tục

◆ Continuous inputs

Đủ lớn cho phép tìm ra điều mới

- Large enough ($>$ normal specs) to allow learning or discovering new things.

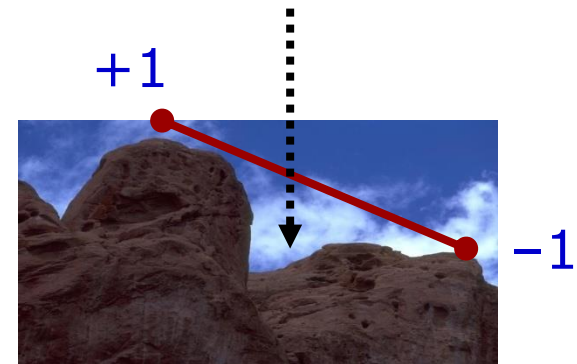


Không nên quá lớn để tránh bỏ lỡ kết quả nằm trong phạm vi cần tìm

- Should not be too large that one might loose the information in between ranges.

Nếu phạm vi quá lớn, nên chia thành các thí nghiệm khác nhau (hoặc sử dụng RSM)

*If the range is very large, it is advisable to **break it into several experiments** (or use RSM)*





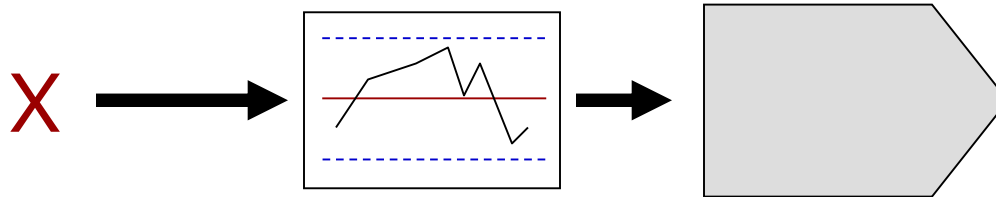
Controlling the Lo & Hi

Biến đầu vào cần được điều khiển một cách có hệ thống

◆ The **input variable** must be in statistical control

Biến X có triệu chứng mất kiểm soát khi để tự do

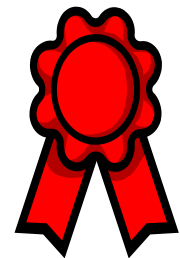
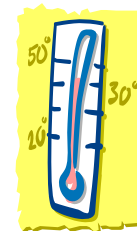
- Free from any out-of-control symptoms



Đo lường giá trị Lo và Hi

◆ Measurement of Lo & Hi must be adequate

- Within acceptable GR&R%

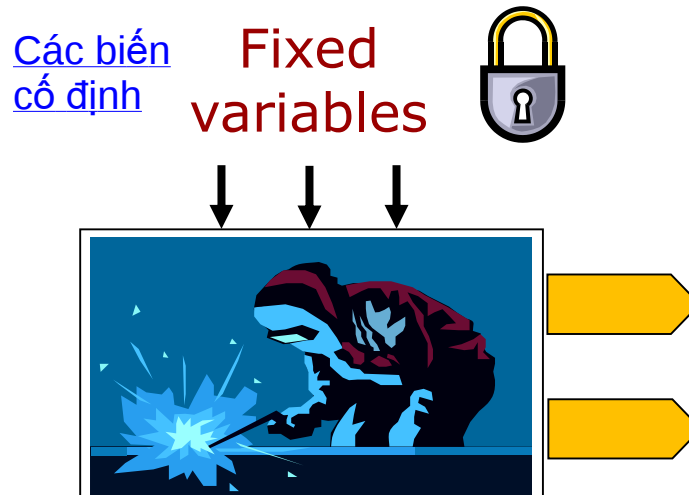




3-525. Cố định các biến đầu vào **Fixed Input Variables**

Có thể cố định các biến đầu vào (đặt các hằng số trong suốt thí nghiệm)

- ◆ Can be fixed with little variance (**put to constant through out experiment**).
 - ◆ **Not wishes to be experimented** (effects known, or to be experimented later).
- Thiết lập để thu được các giá trị tối ưu (giá trị kinh tế)
- ◆ Set to **optimized setting** (or economical setting).





3-5261. What Noise Means?

3-5261. Nhiễu là gì?

Các yếu tố quá đắt hoặc quá khó để điều khiển (thời tiết, môi trường)

- Factors which are **too expensive or too difficult to control** (ie weather, environment).

Có thể hoặc không thể tác động lớn đến thí nghiệm (điều này chúng ta ko biết)

- It may or **may not have great effect** to the experiment (we do not know).

Giống như các bậc tự do, nhiễu tồn tại mọi nơi (các biến thí nghiệm bị dính nhiễu)

- Like variance, noise exists everywhere (experimental variables too has noise attached).



3-5262. What to do with Noise?

Phải điều khiển được (càng nhiều càng tốt nếu chúng ta xác định rõ được loại nhiễu)

- **Must be controlled (as much as possible if we can identify the noise).**

Cố gắng điều chỉnh nhiễu tới mức thấp nhất

- **Put to minimum.**

Tìm ra phương án tối ưu để loại bỏ nhiễu

- **Find optimum setting that is robust to noise (noise effect can be disregard).**



3-5263. How to Control Noise?

3-5263. Làm thế nào để điều khiển nhiễu?

Ngẫu nhiên

● Randomization

Giữ các biến nhiễu ở hằng số

● Keep all noise variables constant

Cùng máy móc thiết bị

◆ Same machine

Cùng người vận hành

◆ Same operator

Cùng ngày

◆ Same day

Blocking - Để cả nhiễu trong thí nghiệm

● Blocking – Include noise factors inside the experiment



3-531. What is Design Matrix?

3-531. Ma trận thiết kế là gì?

◆ Design matrix: Method of combining & varying input variables

Ma trận thiết kế: Phương pháp kết hợp và thay đổi các biến đầu vào

Blue + Pink + Stirs...
lam + hồng + khuấy ...





3-532. Types of Experiments

3-532. Các loại thí nghiệm

Chạy thử và kiểm tra lỗi

Thử nghiệm 1 phép thử 1 lần (OFAT)

Full factorials

Fractional factorials

Phương pháp phản hồi RSM



- *Trial & error*
- *One factor at a time (OFAT)*
- *Full factorials*
- *Fractional factorials*
- *Response surface method (RSM)*



3-533. Trial & Error – Quiz 3-1

3-533. Chạy thử và kiểm tra lỗi

◆ Concern: **How to reach office in Hanoi on time within 30mins traveling time (clock in time is AM9)?** Vấn đề: Làm thế nào để đến văn phòng ở Hà Nội đúng giờ trong vòng 30' (giờ mục tiêu 9h sáng)

● Try Các hành động nên làm:

◆ Get out early from home
Ra khỏi nhà sớm

◆ Take commuter
Bắt tàu điện

◆ Take cab (maybe the driver know the short cuts)
Bắt taxi

◆ Use motorbike
Đi xe máy

◆ Move to Hanoi
Đến Hà Nội

● Result Kết quả

◆ One or more methods may work,
1 hoặc nhiều phương án có thể sử dụng được

◆ None is suitable
Cũng có thể không có phương án nào phù hợp





3-5341. OFAT Example

Vấn đề: Làm thế nào để đạt được lượng tiêu thụ dầu tốt nhất?

◆ **Concern: How to get the best oil consumption?**

Biến thử nghiệm

Thấp

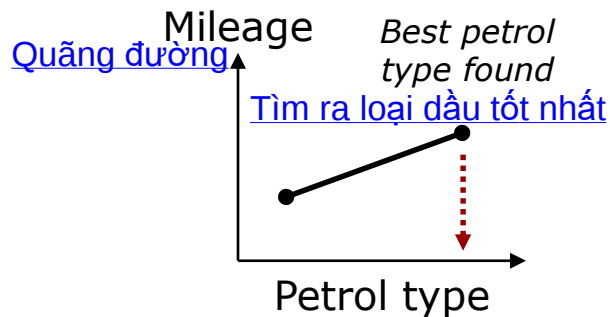
Cao

Loại dầu

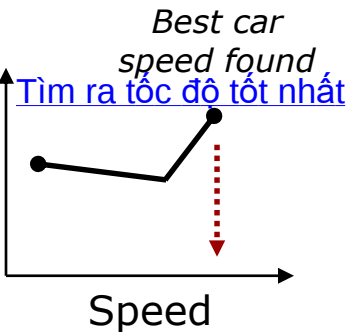
Tốc độ

Áp lực bánh xe

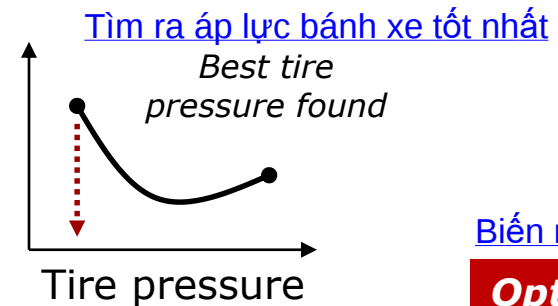
Test Variables	Lo	Hi
Petrol type	PETRON	CELL
Speed	60 km/hr	120 km/hr
Tire pressure	18	24



- Vary Petrol type
- Fix Speed, Tire pressure



- Fix Petrol type, Tire pressure
- Vary Speed



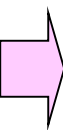
- Fix Petrol type, Speed
- Vary Tire pressure

Biến nào tối ưu?

Optimum?

- Petrol type
- Speed
- Tire pressure

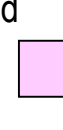
Thay đổi loại dầu
Tốc độ và áp lực bánh xe: ko đổi



Loại dầu, áp lực bánh xe: ko đổi
Thay đổi tốc độ



Loại dầu, tốc độ: ko đổi
Thay đổi áp lực bánh xe

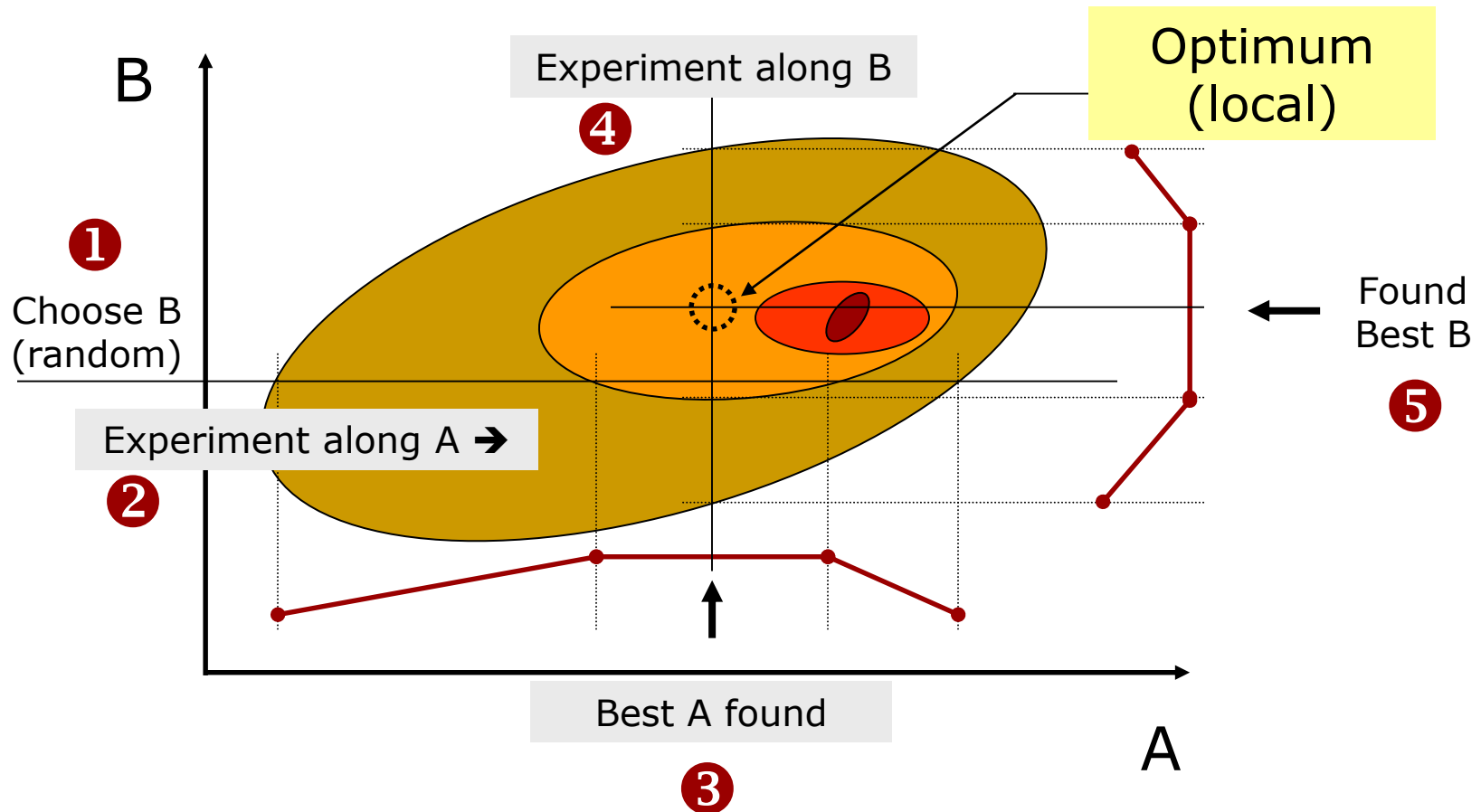




The Real World is not so Symmetric

OFAT bỏ lỡ peak như thế nào? (peak-kết quả tối ưu)

◆ How OFAT misses the peak (optimum result).





3-5343. Number of Experiments Needed

Đối với 3 yếu tố đầu vào

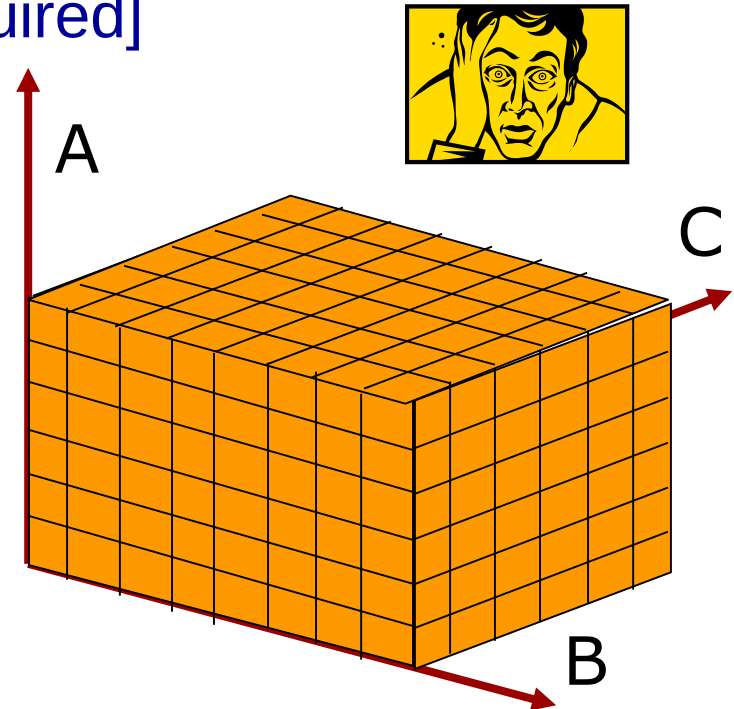
◆ For 3 input factors (example):

Số thí nghiệm cần thiết = $A \times B \times C$. Trong đó A, B, C là các mức thay đổi của 3 biến

- $A \times B \times C$ number of runs needed to find the right (optimum) combination [the higher the resolution, the more number of runs required]

Example

- A: Petrol type [2 types only]
 - B: Tire pressure (18~24, every increment of 2) [4 ranges]
 - C: Speed (60~120, every increment of 20) [4 ranges]
- ⇒ Number of runs = $2 \times 4 \times 4 = 32$ runs!
- ⇒ If add another factor [3 ranges], 96 runs needed!!





3-5351. How the Lo & Hi are

3-5351. Lo và Hi được sắp xếp thế nào?

Arranged?

Complete Lo & Hi combination Hoàn thiện Tổ hợp Lo và Hi

Run
↓

Full tank Đổ đầy bình xăng

Result
↓

Loại dầu

Tốc độ

Áp lực bánh xe

Quãng đường

①

Petron

60 km/hr

24

462 km

②

Petron

60 km/hr

18

455km

③

Petron

120 km/hr

24

512km

④

Petron

120 km/hr

18

490 km

⑤

Cell

60 km/hr

24

460 km

⑥

Cell

60 km/hr

18

466 km

⑦

Cell

120 km/hr

24

549 km

⑧

Cell

120 km/hr

18

511km



3-5352. Full Factorial = Simplified OFAT

Mã hóa đơn giản

Run	Ma trận thiết kế Design Matrix			Kết quả Response
	Factor A	Factor B	Factor C	Y(mean)
1	-1	-1	-1	
2	-1	-1	+1	
3	-1	+1	-1	
4	-1	+1	+1	
5	+1	-1	-1	
6	+1	-1	+1	
7	+1	+1	-1	
8	+1	+1	+1	



3-5353. Số lượng thí nghiệm cần thiết

3-5353. Number of Experiments Needed

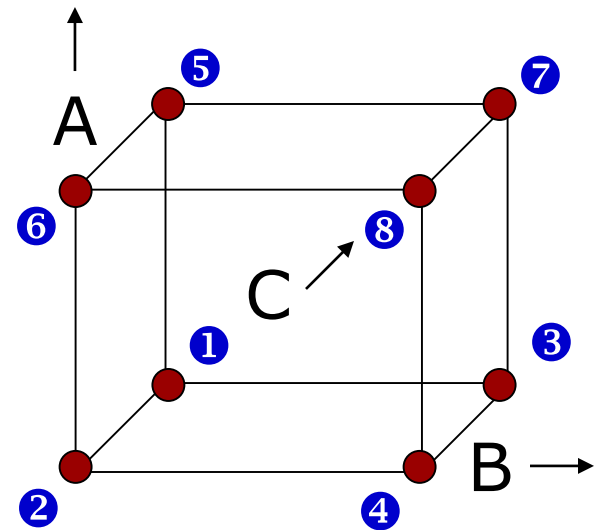
Đối với 3 biến đầu vào

◆ For 3 input factors (example):

8 thí nghiệm

8 runs only

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



OFAT 32 runs vs Full Factorial 8 runs

Which experiment would you choose?

Với OFAT 32 thì chỉ cần chạy 8 thí nghiệm



3-5361. What is Randomization?

◆ Running the experiment runs in random order

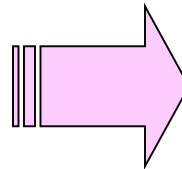
Chạy các thí nghiệm theo thứ tự ngẫu nhiên

Thứ tự chuẩn

Standard order

↓

Run #	A	B	C
1	-1	-1	-1
2	-1	-1	+1
3	-1	+1	-1
4	-1	+1	+1
5	+1	-1	-1
6	+1	-1	+1
7	+1	+1	-1
8	+1	+1	+1



Thứ tự ngẫu nhiên

Random order

↓

Run #	A	B	C
6	+1	-1	+1
2	-1	-1	+1
3	-1	+1	-1
8	+1	+1	+1
1	-1	-1	-1
4	-1	+1	+1
7	+1	+1	-1
5	+1	-1	-1



3-5362. Purpose of Randomization

- ◆ To reduce the effect of noise to the experiment
Để giảm ảnh hưởng của nhiễu tới thí nghiệm

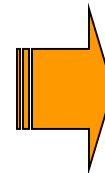
Ngẫu nhiên hóa
giúp trộn nhiễu với nhau

Randomization will mix the noise effect (more natural)

Thứ tự chuẩn
Standard order



		Run #	A	B	C	Yhat
<u>Sáng</u> Morning	{	1	-1	-1	-1	Y_{1m}
		2	-1	-1	+1	Y_{2m}
<u>Chiều</u> Afternoon	{	3	-1	+1	-1	Y_{3A}
		4	-1	+1	+1	Y_{4A}
		5	+1	-1	-1	Y_{5A}
		6	+1	-1	+1	Y_{6A}
<u>Tối</u> Night	{	7	+1	+1	-1	Y_{7N}
		8	+1	+1	+1	Y_{8N}



Thứ tự ngẫu nhiên
Random order

Run #	A	B	C	Yhat
6	+1	-1	+1	Y_{3A}
2	-1	-1	+1	Y_{7N}
3	-1	+1	-1	Y_{6A}
8	+1	+1	+1	Y_{1m}
1	-1	-1	-1	Y_{8N}
4	-1	+1	+1	Y_{5A}
7	+1	+1	-1	Y_{2m}
5	+1	-1	-1	Y_{4A}



3-541. About Output Variables

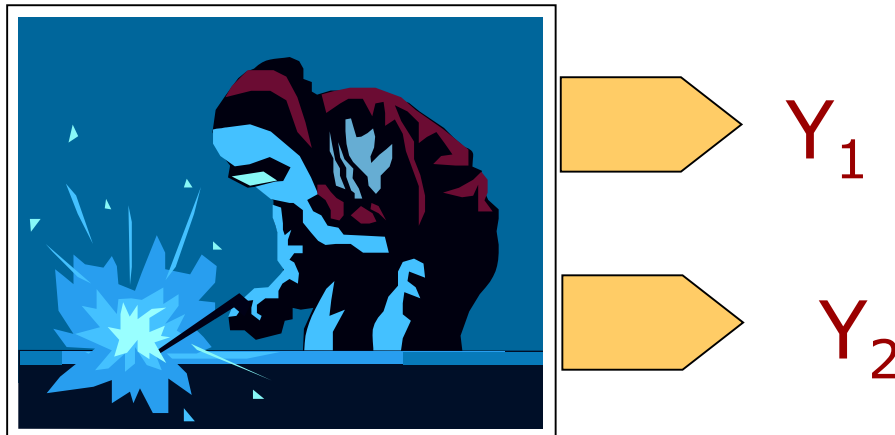
3-541. Về các biến đầu ra

Là Kết quả Y của thí nghiệm

- ◆ Response of (= Y) of an experiment.

Một thí nghiệm cần phải thu được ≥ 1 các giá trị đầu ra (tức là có các giá trị $Y_1, Y_2, \dots$)

- ◆ **An experiment can have ≥ 1 output variables ($Y: \geq 1 \Rightarrow Y_1, Y_2, \dots$)**





3-542. Defining Output Variables

Biến đầu ra phải:

◆ The output variable(s) must be:

Giá trị có ý nghĩa

- **Meaningful** index.

Định lượng được - đo lường được

- **Quantified** – measurable (qualitative or quantitative).

Được đo lường theo cách phù hợp

- Measured adequately – **verified by GR&R gage.**

Có phân bố Normal

- **Normal distribution** (prior to experiment)

- Statistically **in control** (prior to experiment)



3-543. Làm gì sau khi xác định được Y?

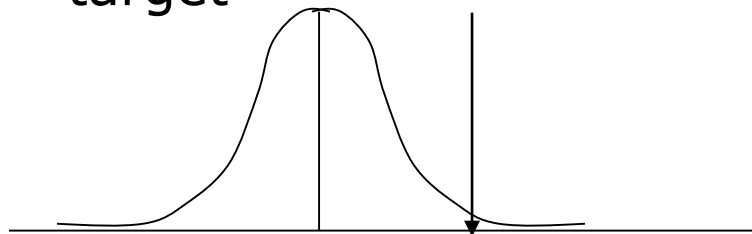
What to do After Defining Y?

Đo lường trạng thái hiện tại

1. Measure **current status**

Khoảng cách từ giá trị trung bình tới giá trị mục tiêu

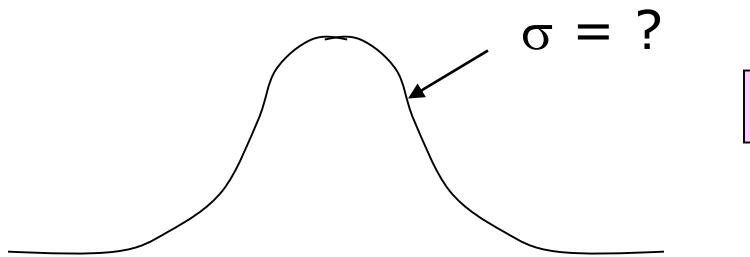
- How far is the Mean from target



Mean Target

Độ lệch chuẩn là bao nhiêu?

- How big is the Standard deviation

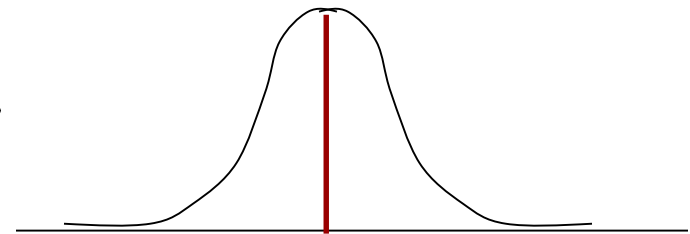
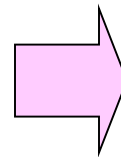


Đặt ra mục tiêu cần đạt được

2. Set target to achieve

Đặt giá trị trung bình làm mục tiêu

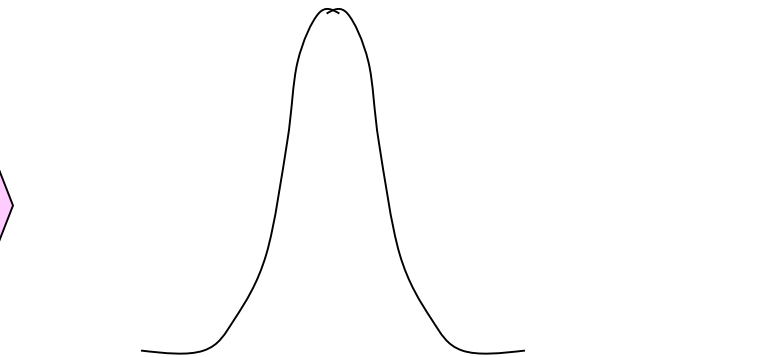
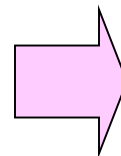
- Centering the mean



Mean = Target

Giảm phương sai

- Variance reduction





3-544. Repetition & Replication

Repetition: quan sát sự thay đổi ngắn hạn

◆ Meaning of **Repetition** (observes short-term variability)

Thử nghiệm vài mẫu trong 1 thí nghiệm

- **Running several samples during one experimental setup run.**

Độ lệch tiêu chuẩn
Standard deviation

Run #	A	B	C	Y1	Y2	Y3	Ybar	S
1	+1	+1	+1	509	513	512	511.3	2.082

Thí nghiệm 1

1. Run # 1

Lặp lại lần 1

2. Repeat Run # 1

lặp lại lần 2

3. Repeat Run # 1

Average of Y1 ~ Y3
Giá trị trung bình của các kết quả đo



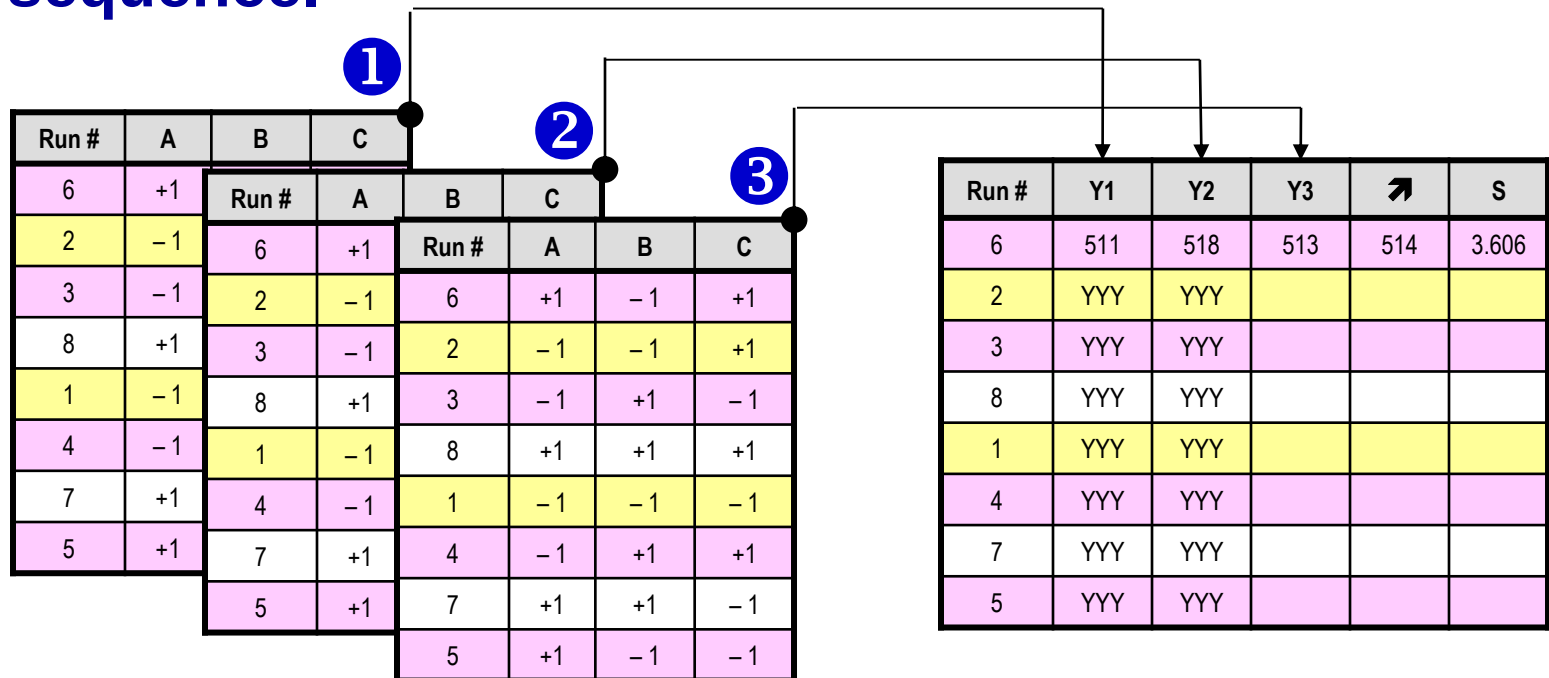
3-544. (cont) Repetition & Replication

Replication: quan sát sự thay đổi dài hạn

◆ Meaning of **Replication** (observes long-term variability)

Thực hiện thí nghiệm toàn diện

- Replicating the entire experiment in a time sequence.





3-545. Purpose of Repetition & Replication

Mục đích

◆ Purposes:

Repetition và Replication cho phép ước lượng sự thay đổi (tự nhiên) trong thí nghiệm

- **Repetition and replication provide estimates of the natural variability in the experiment.**

Run #	Y1	Y2	Y3	$\bar{Y}$	S
3	511	518	513	514	3.606
7	YYY	YYY			
6	YYY	YYY			
1	YYY	YYY			
8	YYY	YYY			
5	YYY	YYY			
2	YYY	YYY			
4	YYY	YYY			

Dữ liệu dùng để phân tích sự thay đổi trong thí nghiệm

Data used to analyze variability in the experiment

- Both Repetition and Replication can be used in combination in the same experiment.

Có thể kết hợp cả Repetition và Replication trong cùng một thí nghiệm



3-546. Size of Repetition & Replication

3-546. Cỡ Repetition và Replication

◆ The size depends on Cỡ phụ thuộc vào:

Số lần tiến hành thí nghiệm

● **Number of runs** in the experiment

Mức độ tin cậy mong muốn của thí nghiệm

● **Confidence level** desired in the experiment



3-61. Screening vs Modeling

3-61. Screening và Modeling

Hai giai đoạn thiết kế

◆ Two stages of design:

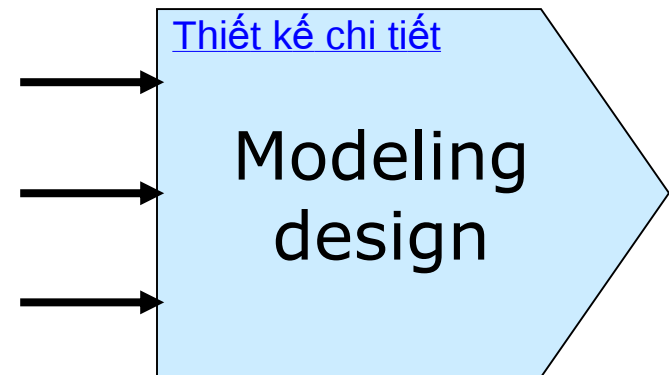
1



Design matrix

- Taguchi
- Fractional Factorial

2



Design matrix

- Full Factorial
- CCD
- Box Behnken



3-62. Thiết kế định hình. **Screening Design**

Mục đích

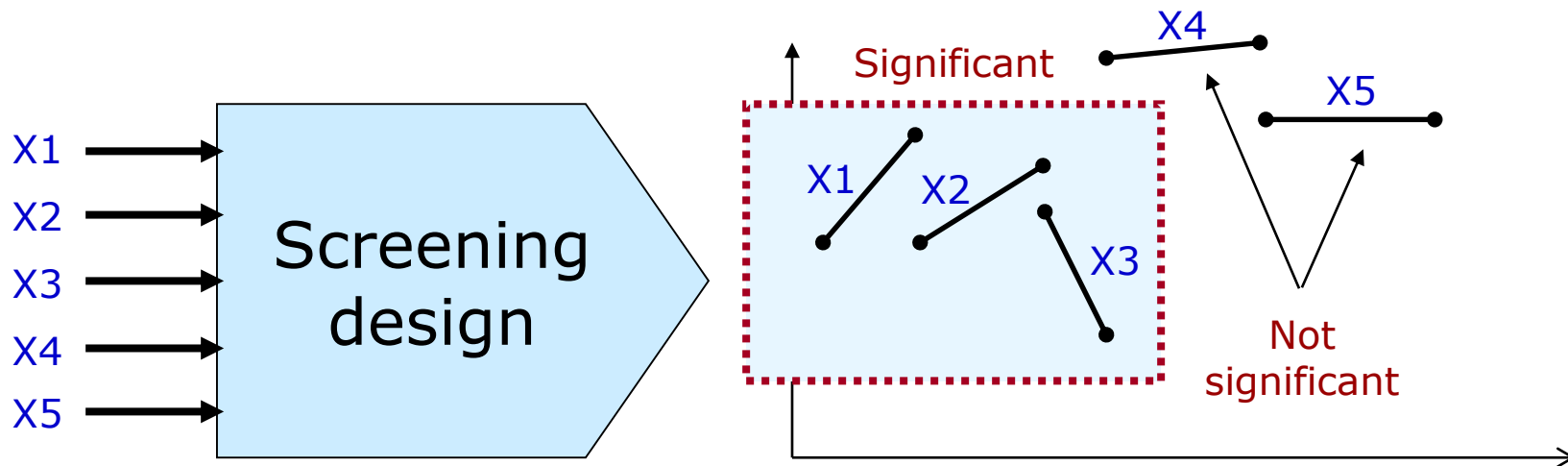
◆ Purposes:

Vẽ/mô phỏng ra những điểm quan trọng trong rất nhiều các yếu tố được khảo sát

- To screen out the **vital few from the trivial** many (investigates a relatively large number of factors).

Để bước đầu tìm ra mức độ của các ảnh hưởng chính

- To assess primarily **the significance of main effects**.





Mục đích

◆ Purpose:-

Tìm ra mối quan hệ giữa các yếu tố

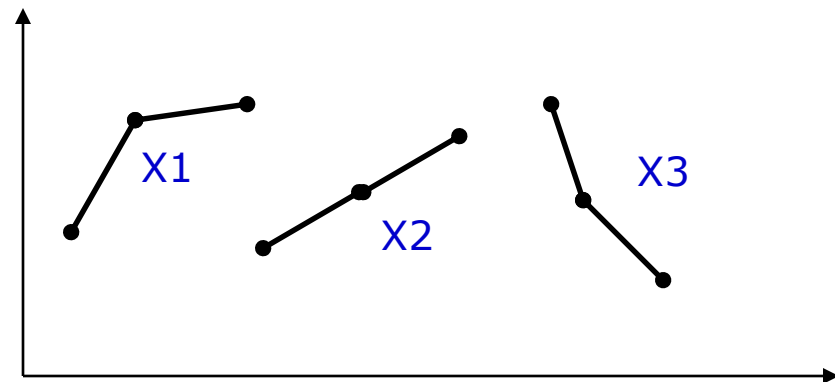
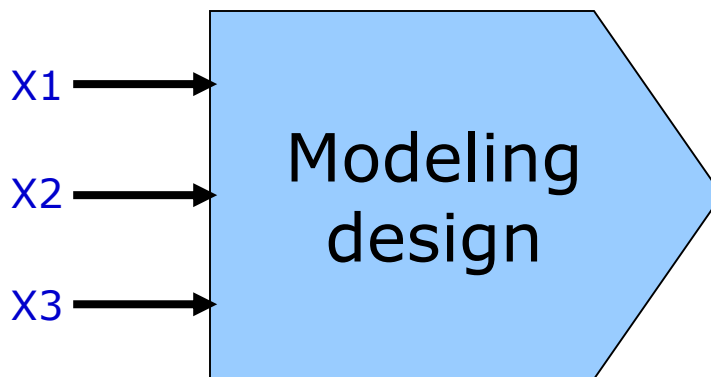
- To find the **interaction (relationship)** between factors.

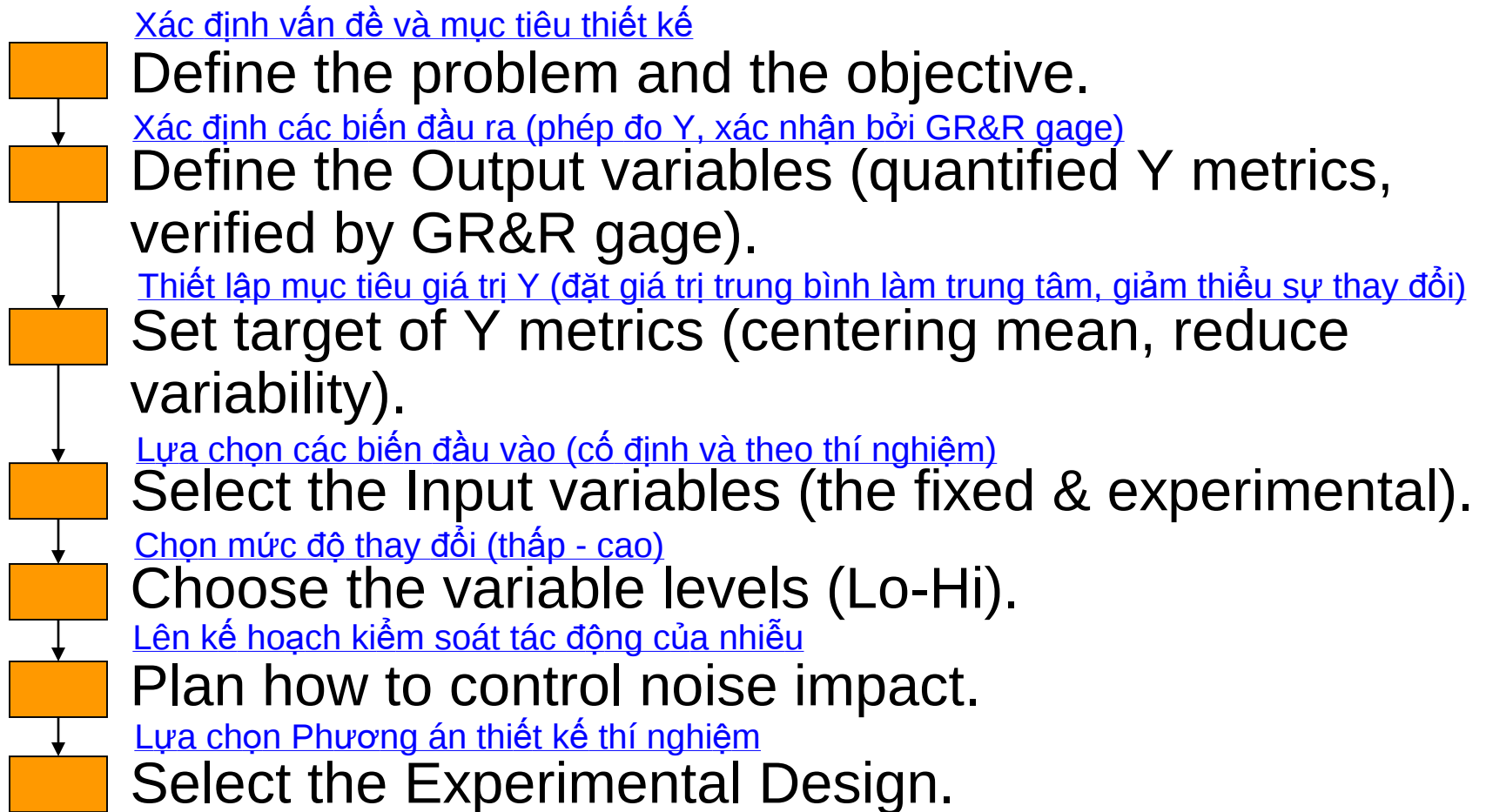
Tìm ra liệu có đường cong nằm trong phạm vi khảo sát

- To find **if there is a curve** within the investigated range.

Tìm ra biểu thức mô tả mối quan hệ đầu vào - đầu ra để tìm ra đặc điểm và tối ưu hóa

- To find the **model expression** for characterization and optimization. $[Y = AX_1X_2 + BX_1 + CX_2 + K]$

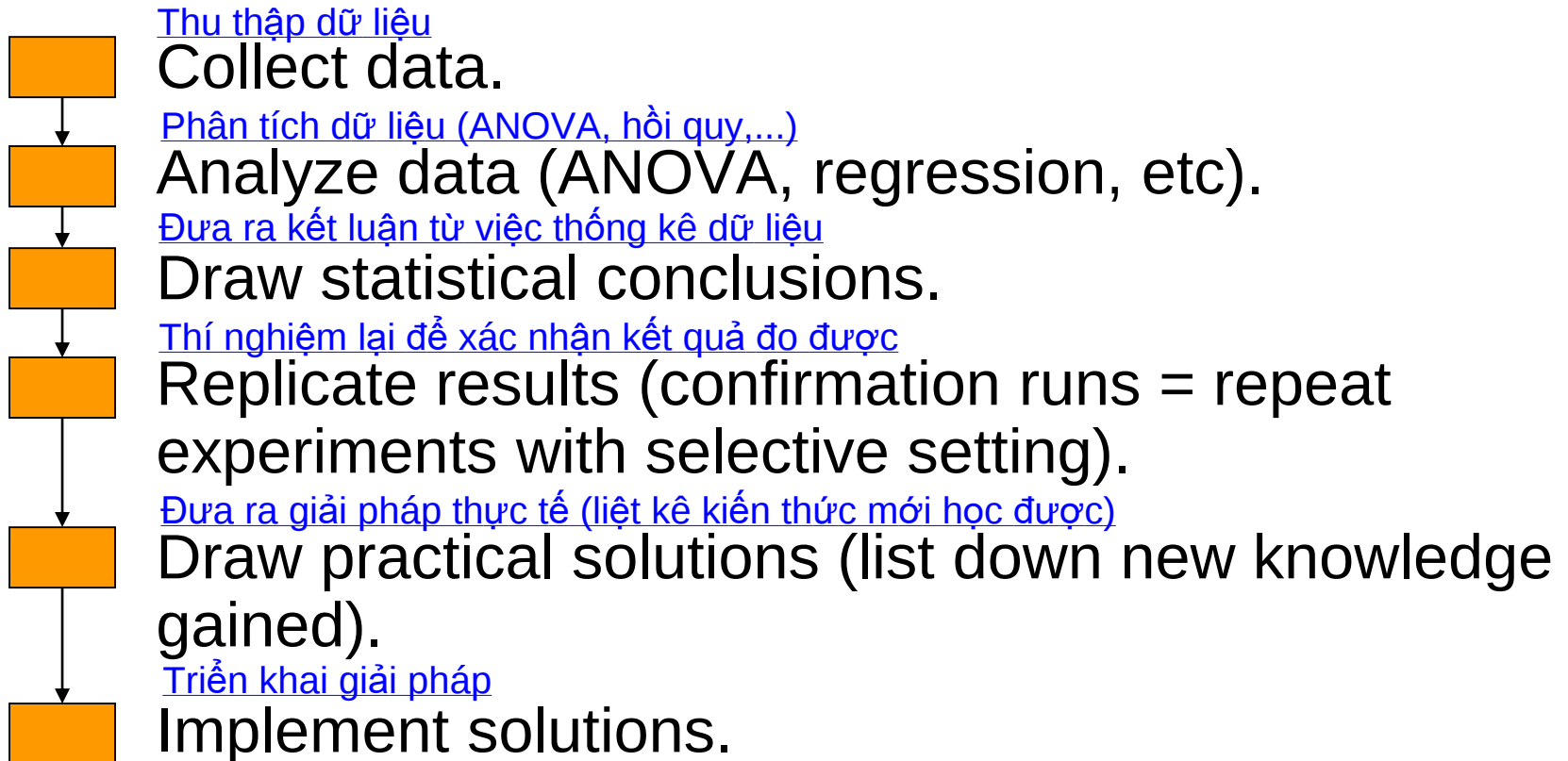






3-72. Running the Experiment

3-72. Triển khai thí nghiệm





3-73. Nhân rộng phương án mới

Expansion of New Solutions



Make sure solid implementation of new solution (at designated processes).

Đảm bảo việc triển khai chắc chắn phương án mới

Review the suitability of the solutions to the new processes or environment (modify if necessary).

Xem xét lại tính phù hợp của giải pháp đối với quá trình mới hoặc môi trường mới (điều chỉnh giải pháp nếu cần)

Perform confirmation run (with slight adjustment to suit the new environment).

Nếu chạy thí nghiệm để xác nhận mà thấy giải pháp mới phù hợp, thì áp dụng giải pháp đó một cách toàn diện trong khâu theo dõi về sau

If confirmation runs show the new solutions are suitable, apply at full scale with close monitoring.

Liệt kê những kiến thức thu được

List down new knowledge gained.

Chạy thí nghiệm mới (trong môi trường thông thường) chỉ khi thực sự cần thiết

Run new DOE (for common environment) only if it is really necessary.



3-91. Điều nên làm

3-81. The Do's

Sẵn sàng thay đổi để thu được kinh nghiệm hữu ích

- ◆ Be ready for changes for new best practices.

Đảm bảo rằng những kết quả thu được phải liên hệ chặt chẽ với dự án th.kế

- ◆ Make sure measurable potential business gains are tied to the project.

Viết thành văn bản

- ◆ Document initial information.

Phân công công việc thu thập dữ liệu và phân tích dữ liệu cho người phù hợp

- ◆ Allocate sufficient resources for data collection and analysis.

Đảm bảo xác nhận tất cả các phép đo phải theo cách hợp lý

- ◆ Make sure all measurement are verified adequately.

Lên kế hoạch triển khai tại tất cả các giai đoạn thiết kế

- ◆ Plan execution at all phases.

Bắt đầu thiết kế hai mức với cỡ mẫu nghiêm ngặt/chính xác

- ◆ Start with two-level designs with rigorous sample size determination.



3-81. Điều nên làm **3-81. (cont) The Do's**

Chi phí ít hơn 25% ngân quỹ cho thí nghiệm đầu tiên

- ◆ Spend less than 25% of budget on the first experiment.

Thí nghiệm cần một đội có các thành viên có kiến thức tốt nhất

- ◆ DOE requires team involvement where members prior knowledge is maximize.

Theo dõi và ghi lại bất kỳ nguyên nhân bên ngoài nào tác động lên sự thay đổi

- ◆ Watch for and record any extraneous sources of variation.

Thường xuyên xác nhận kết quả bằng các thí nghiệm xác nhận

- ◆ Always verify results with a confirmation runs (narrow
➔ broad inference).

Đưa ra kết luận cho mỗi thí nghiệm

- ◆ Must draw conclusions after every experiment.



- ❖ Don't start a new experiment design until the previous one is finished.
- ❖ Don't try to answer all the questions in one study – several experiments may be needed.
- ❖ Don't expect perfect experiment in the first DOE – many things will be learned along the way. Imperfect DOE is a lesson of new caution points.

Không nên bắt đầu thiết kế thí nghiệm mới cho đến khi kết thúc thí nghiệm trước đó

Không nên cố gắng trả lời tất cả các câu hỏi trong một thí nghiệm (vì cần thêm thí nghiệm để xác nhận)

Không nên mong đợi một thí nghiệm hoàn hảo - vì có nhiều điều cần phải học dài lâu. Một thí nghiệm chưa hoàn hảo là bài học kinh nghiệm



Discussion 3-1: Advantages of Big Hi-Lo Range

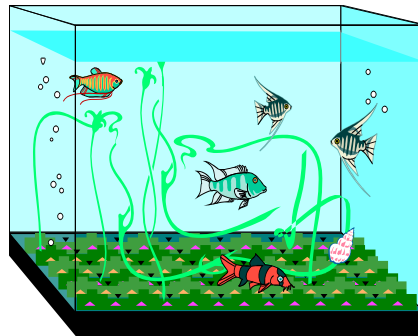
Thảo luận 3-1:
Ưu điểm của Phạm vi Hi-Lo lớn

Tiết kiệm thời gian

- **Saves time.**

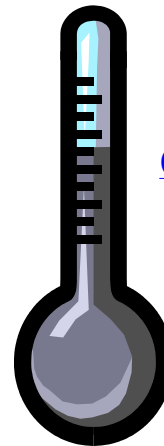
Có thể tìm ra được thông tin mới khác nhiều so với phạm vi hẹp thông thường của thông số

- **May be able to find new information beyond the normal narrow range of specification.**



Điều gì xảy ra tại 100°C?

What happen at 100°C?



Phạm vi thí nghiệm

Experiment range

Cao Hi = 100°C

Thấp Lo = 0°C



Discussion 3-2: Disadvantages of Small Lo-Hi Range

Thảo luận 3-2:
Nhược điểm của Phạm vi Lo-Hi hẹp

Tiêu tốn thời gian và nguồn lực

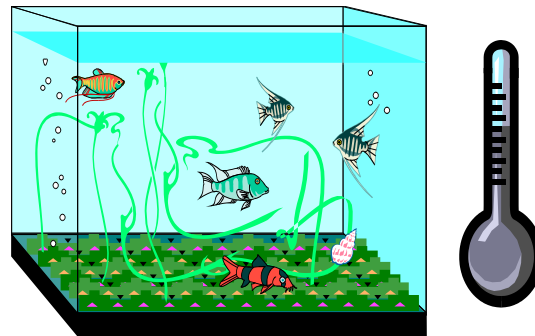
• Time & resources consuming

Các thí nghiệm khác nhau cho ra những kết quả tương tự

• Outcome of each experiment series is similar and expected

Phạm vi thí nghiệm

Experiment range



Run	Range (°C)
1	0 ~ 5
2	6 ~ 10
3	11 ~ 15
4	16 ~ 20

⋮



◆ Benefits of applying DOE at workplace:

Áp dụng DOE trong công việc có nhiều lợi ích:

Giúp đầu óc sáng tạo hơn



It helps creating innovative mind

Biến dữ liệu thành kiến thức



It transforms data to knowledge

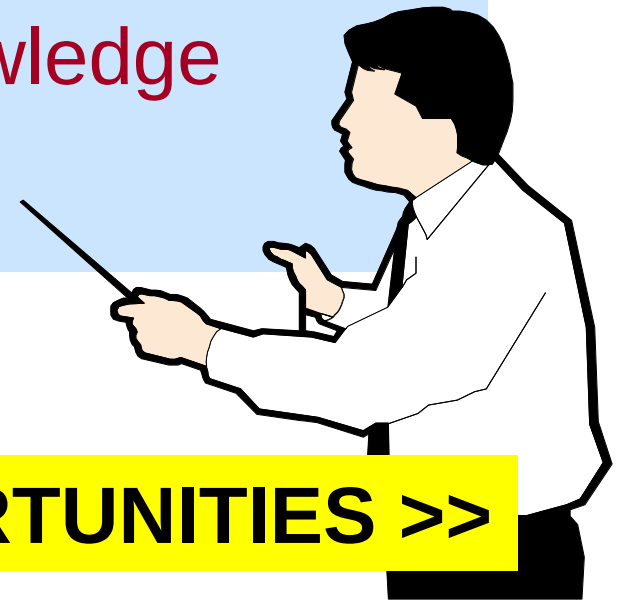
Tạo ra đột phá



It creates breakthrough...

Tạo ra cơ hội kinh doanh mới

>> NEW BUSINE\$\$ OPPORTUNITIES >>





Tổng kết về DOE *(cont) Summary of DOE*

- ◆ What is the purpose of DOE?

Mục đích của DOE là gì?

- ◆ What does R^2 value tells you in DOE?

Giá trị R^2 nói lên điều gì trong DOE?

- ◆ What does P-value in individual input variable tell you in DOE?

Giá trị P trong biến đầu vào nói lên điều gì trong DOE?

- ◆ How many stages are there in DOE?

Có bao nhiêu bước thực hiện trong DOE?

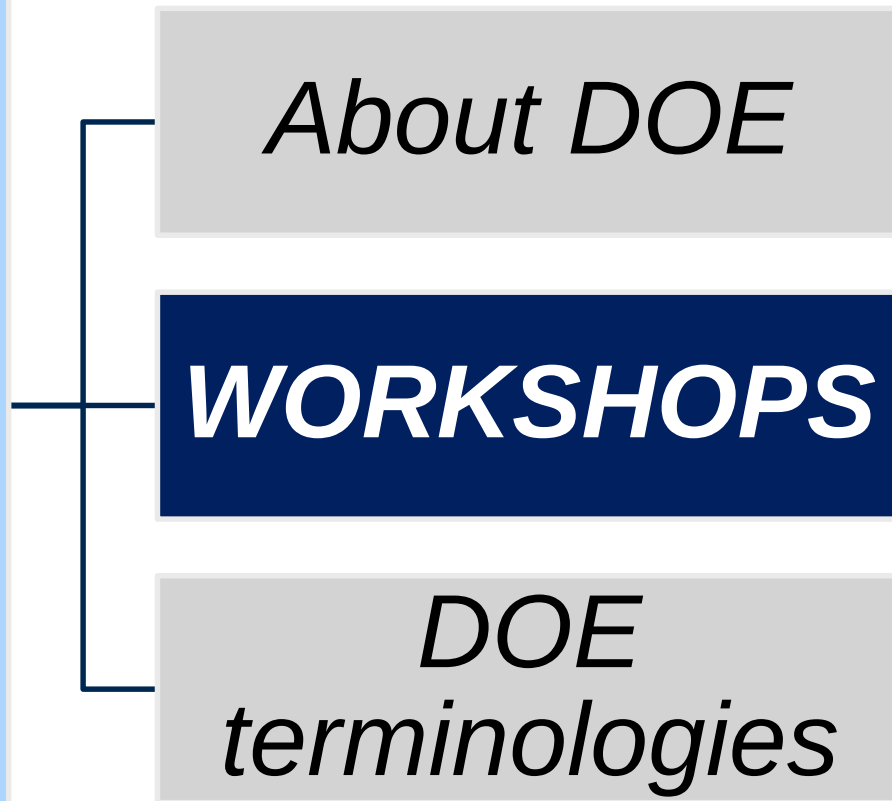
- ◆ In what case we can directly perform Screening DOE?

Trong trường hợp nào thì nên thực hiện trực tiếp DOE Screening?

Workshops

THỰC HIỆN THÍ NGHIỆM

MY FIRST DOE



1. Workshop 1: Statistical Test
 2. Workshop 2: Screening DOE
 3. Workshop 3: Modeling DOE
- [1. Thử nghiệm](#)
 - [2. DOE Screening](#)
 - [3. DOE Modeling](#)



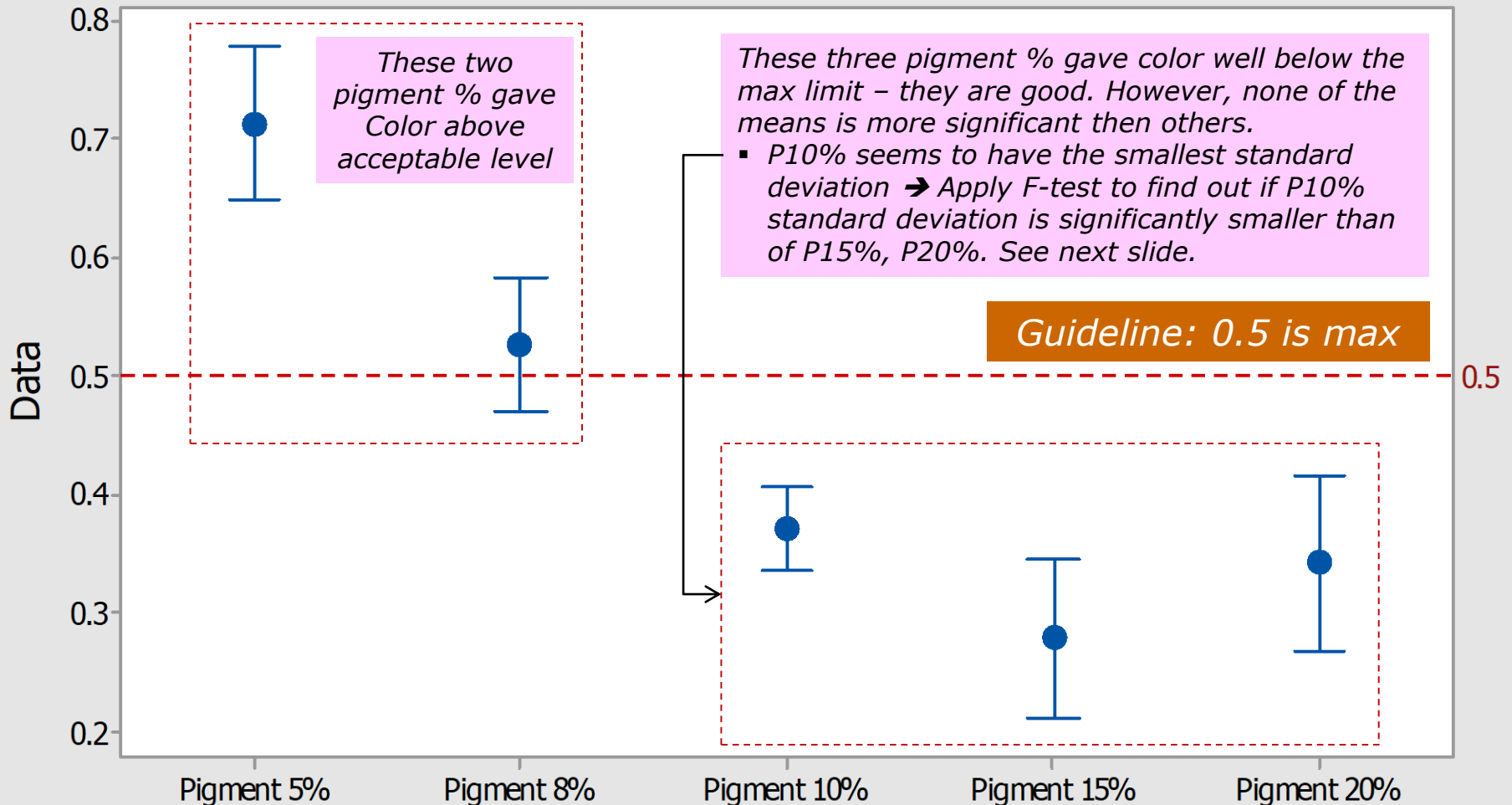
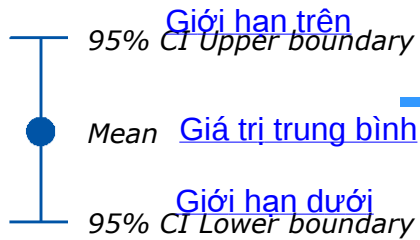
David thực hiện nhiều thí nghiệm để tìm ra tỉ lệ pha màu tối ưu

- David is a senior chemist. Every time a new product is developed he and his team will perform multiple tests to confirm the amount of pigment used that would best suit the intended design of the product. Use relevant graphical tools and/or statistical tests to help David to make good decision with small probability of making a mistake (color reading of < 0.5 is better).



<u>Thí nghiệm</u>	<u>Tỉ lệ 5%</u>	<u>Tỉ lệ 8%</u>	<u>Tỉ lệ 10%</u>	<u>Tỉ lệ 15%</u>	<u>Tỉ lệ 20%</u>
Test #	Pigment 5%	Pigment 8%	Pigment 10%	Pigment 15%	Pigment 20%
1	0.69	0.60	0.40	0.29	0.45
2	0.71	0.57	0.36	0.35	0.27
3	0.85	0.59	0.41	0.37	0.47
4	0.79	0.57	0.44	0.29	0.30
5	0.72	0.46	0.36	0.19	0.41
6	0.68	0.51	0.31	0.14	0.27
7	0.67	0.51	0.34	0.26	0.33
8	0.60	0.41	0.35	0.34	0.24

Interval Plot of Pigment Tests (95% CI for the Mean)



Individual standard deviations were used to calculate the intervals.



Descriptive Statistics for All

Bảng thống kê các tỉ lệ pha màu

Pigment%

Biến

Giá trị TB

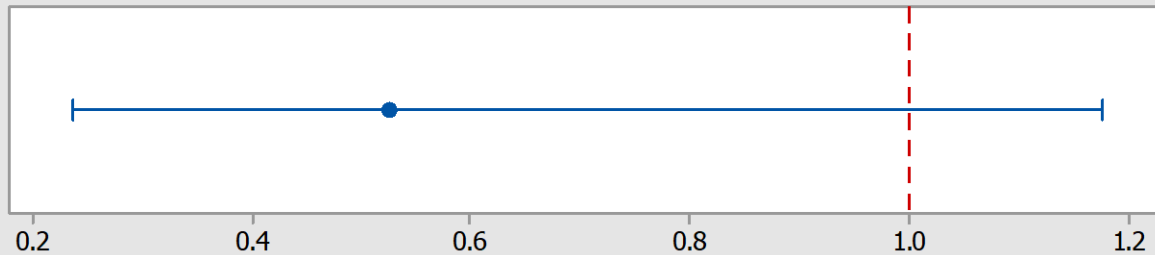
Variable	Mean	StDev	Minimum	Median	Maximum
Pigment 5%	0.7137	0.0765	0.6000	0.7000	0.8500
Pigment 8%	0.5275	0.0673	0.4100	0.5400	0.6000
Pigment 10%	0.3712	0.0422	0.3100	0.3600	0.4400
Pigment 15%	0.2788	0.0801	0.1400	0.2900	0.3700
Pigment 20%	0.3425	0.0889	0.2400	0.3150	0.4700



Thí nghiệm với
tỉ lệ pha 10% và 15%

F-test (Variance test) for Pigment 10% vs Pigment 15%

95% CI for $\sigma(\text{Pigment 10\%}) / \sigma(\text{Pigment 15\%})$

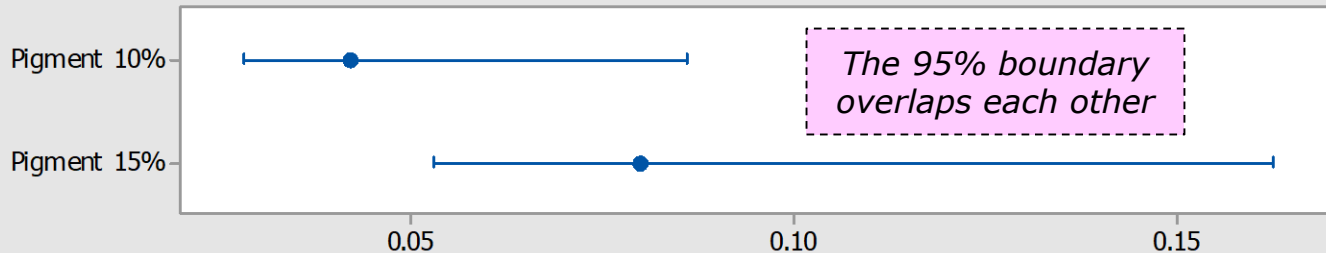


F-Test

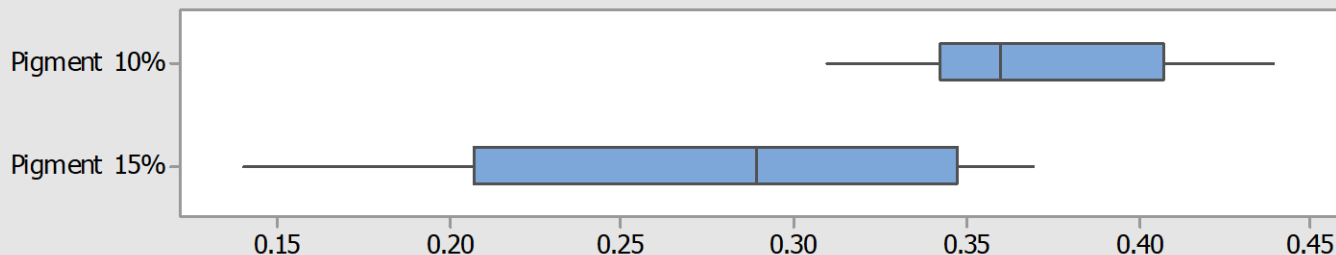
P-Value 0.113

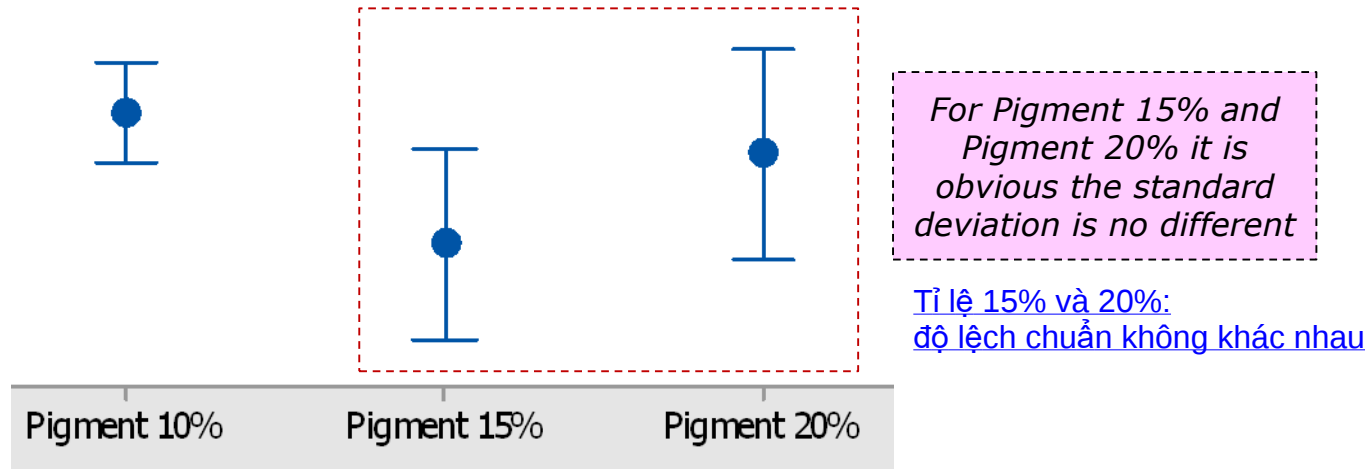
*P-value > 0.05
→ Means there is no
significant difference
between the two
standard deviation*

95% Chi-square CIs for StDevs



Boxplot of Pigment 10%, Pigment 15%





*If cost of pigment is of no issue, **Pigment 15%** shall be the winner. It gives lowest Color index, and Pigment 10% is not significantly better in terms of variability.*

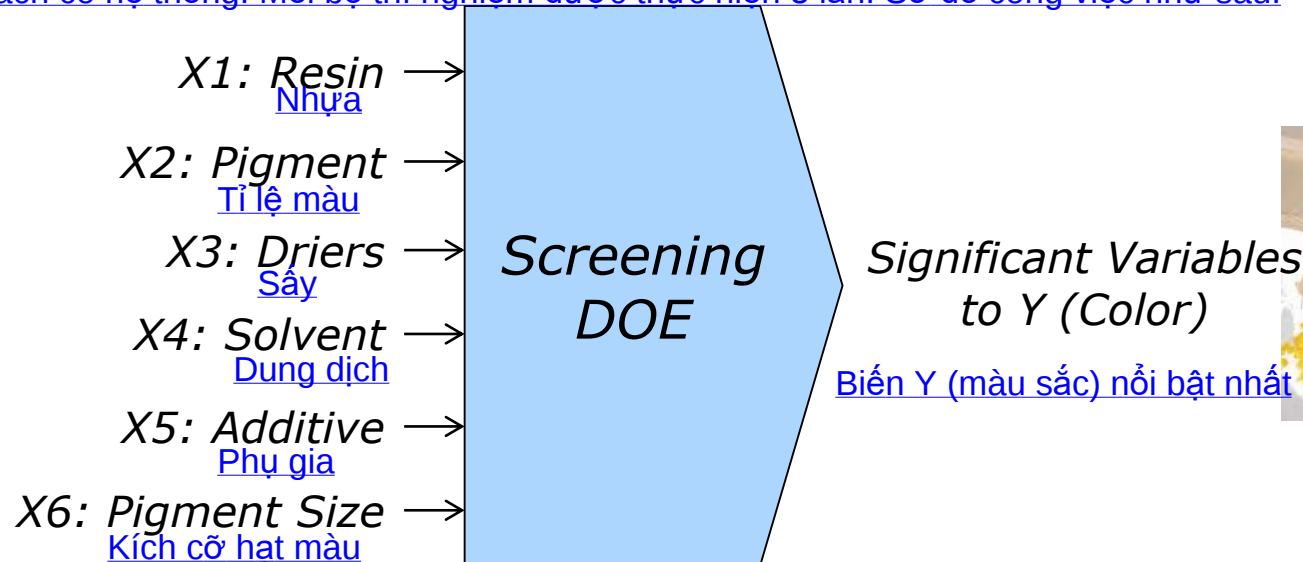
Nếu ko tính đến chi phí, tỉ lệ 15% là tỉ lệ tối ưu. Tỉ lệ này cho kết quả chỉ số màu sắc thấp nhất. Và tỉ lệ 10% so với tỉ lệ 15% không có chỉ số nào tối ưu hơn.



Workshop 2: Screening DOE

- Being a senior chemist, David has always thought of manipulating the product formula hoping to see remarkable results rather than the normal expected outcomes. With the help from other chemists in the department, they perform an experiment manipulating the product formula systematically. Each combination is run 5 times as suggested to estimate errors present in each combination run. *The IPO chart is shown below.*

David luôn nghĩ về việc thay đổi công thức màu để tìm ra kết quả tối ưu, tốt hơn kết quả hiện có. Các thí nghiệm được tiến hành một cách có hệ thống. Mỗi bộ thí nghiệm được thực hiện 5 lần. Sơ đồ công việc như sau:





Fractional Factorial Design

Matrix and Data Collection

Ma trận tổ hợp
và thu thập dữ liệu

Fractional Factorial, 6 factors, 8 runs, 5 replicates

6 biến đầu vào, 8 thí nghiệm, 5 phép lặp

A	B	C	D	E	F	COLOR						
					Pigment							
<u>Resin</u>	<u>Pigment</u>	<u>Driers</u>	<u>Solvent</u>	<u>Additive</u>	<u>Size</u>	<u>Y1</u>	<u>Y2</u>	<u>Y3</u>	<u>Y4</u>	<u>Y5</u>	<u>Y bar</u>	<u>S</u>
40	5	0.5	30	0.8	20	0.69	0.71	0.85	0.79	0.72	0.752	0.066483
40	5	2	30	0.1	10	0.4	0.36	0.41	0.44	0.36	0.394	0.034351
40	20	0.5	5	0.8	10	0.45	0.27	0.47	0.3	0.41	0.38	0.09
40	20	2	5	0.1	20	0.57	0.46	0.51	0.51	0.41	0.492	0.060166
60	5	0.5	5	0.1	20	0.6	0.57	0.59	0.57	0.46	0.558	0.056303
60	5	2	5	0.8	10	0.29	0.35	0.37	0.29	0.19	0.298	0.070143
60	20	0.5	30	0.1	10	0.44	0.36	0.31	0.34	0.35	0.36	0.048477
60	20	2	30	0.8	20	0.3	0.41	0.27	0.33	0.24	0.31	0.065192



Y-hat Model		Color			
Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		0.44300	0.0000		
A	Resin	-0.05750	0.0000	1	X
B	Pigment	-0.06950	0.0000	1	X
C	Driers	-0.06150	0.0000	1	X
D	Solvent	0.08500	0.0000	1	X
E	Additive	0.01100	0.2796	1	X
F	Pigment Size	-0.00800	0.4297	1	X
ABC		-0.03250	0.0027	1	X
R ²		0.8643	1		
Adj R ²		0.8346			
Std Error		0.0633			
F		29.1123			
Sig F		0.0000			
F _{LOF}		NA			
Sig F _{LOF}		NA			
Source		SS	df	MS	
Regression		0.8	7	0.1	
Error		0.1	32	0.0	
Error _{Pure}		0.1	32	0.0	
Error _{LOF}		0.0	0	NA	
Total		0.9	39		

Chất phụ gia và Kích thước hạt màu ko phải là yếu tố ảnh hưởng lớn $P \geq 0.05$

2 Additive & Pigment size are not significant factors ($P\text{-value} \geq 0.05$)

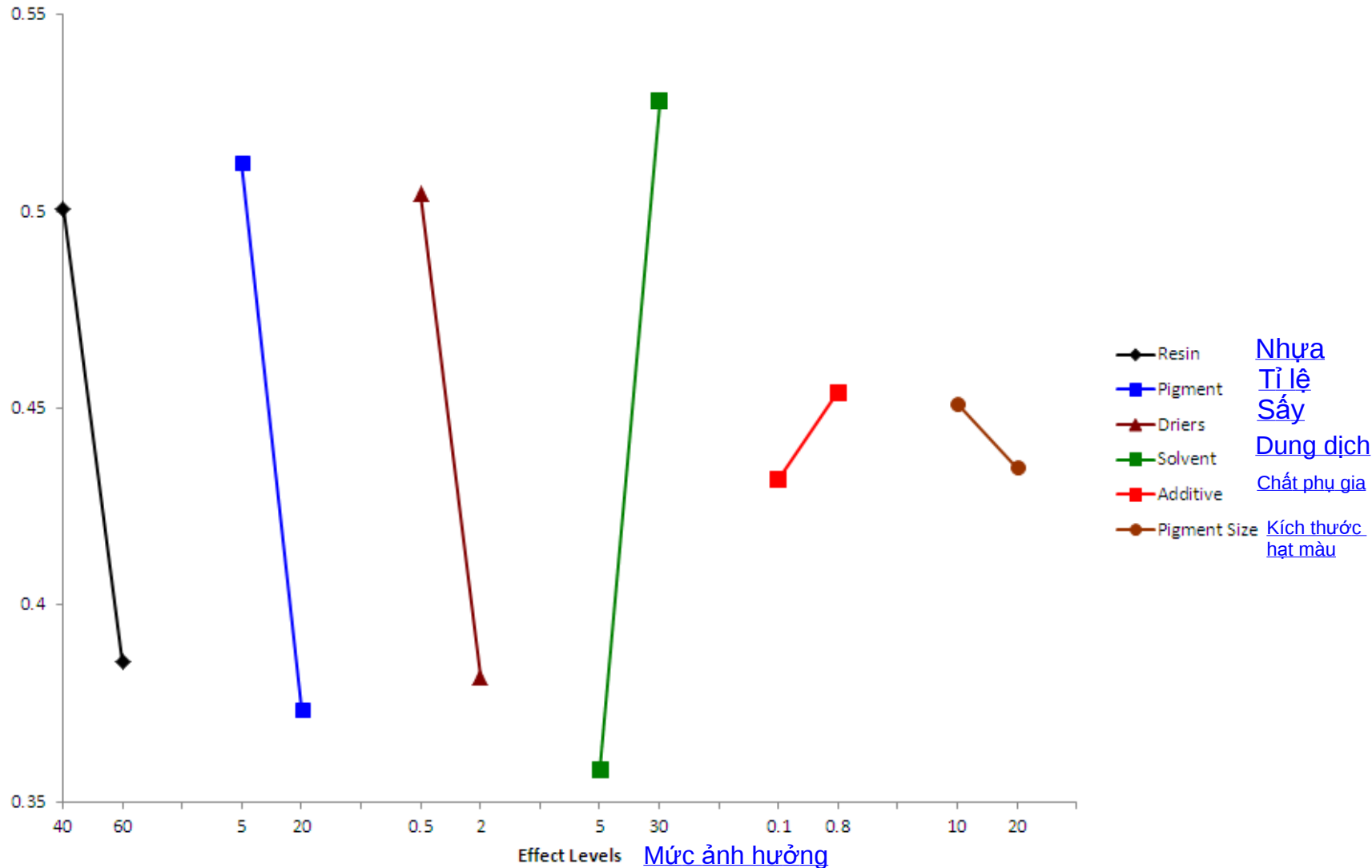
$R^2' > 0.7$ shows the regression is a reliable one

$R^2' > 0.7$ cho thấy phép hồi quy đáng tin cậy



Marginal Means Plot (Y-hat)

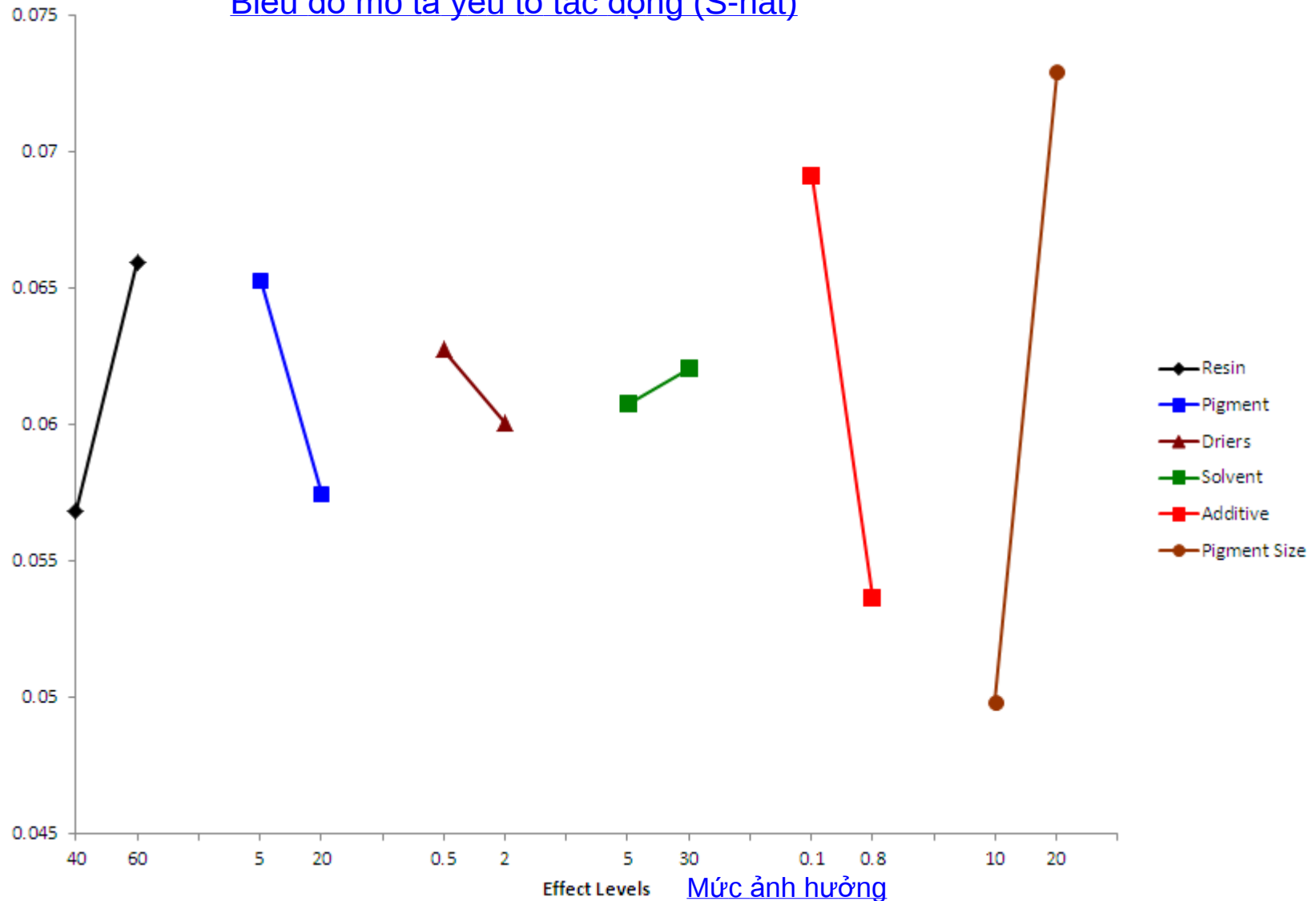
Biểu đồ mô tả yếu tố tác động (Y-hat)





Marginal Means Plot (S-hat)

Biểu đồ mô tả yếu tố tác động (S-hat)





IPO Chart for Modeling DOE

Sơ đồ IPO cho giai đoạn DOE Modeling

*Variables dropped
from Modeling
experiment at put at
Optimum Level*

Hai biến ko ảnh hưởng lớn
sẽ bị lược bỏ khi tối ưu hóa

Phụ gia

X5: Additive

Set at Hi – 0.8%

K.thước
hạt màu

X6: Pigment Size

Set at Lo – 10 μm

*Variable to be
experimented
further using Full
Factorial Matrix
Các yếu tố ảnh
hưởng lớn*

X1: Resin →

X2: Pigment →

X3: Driers →

X4: Solvent →

**Modeling
DOE**

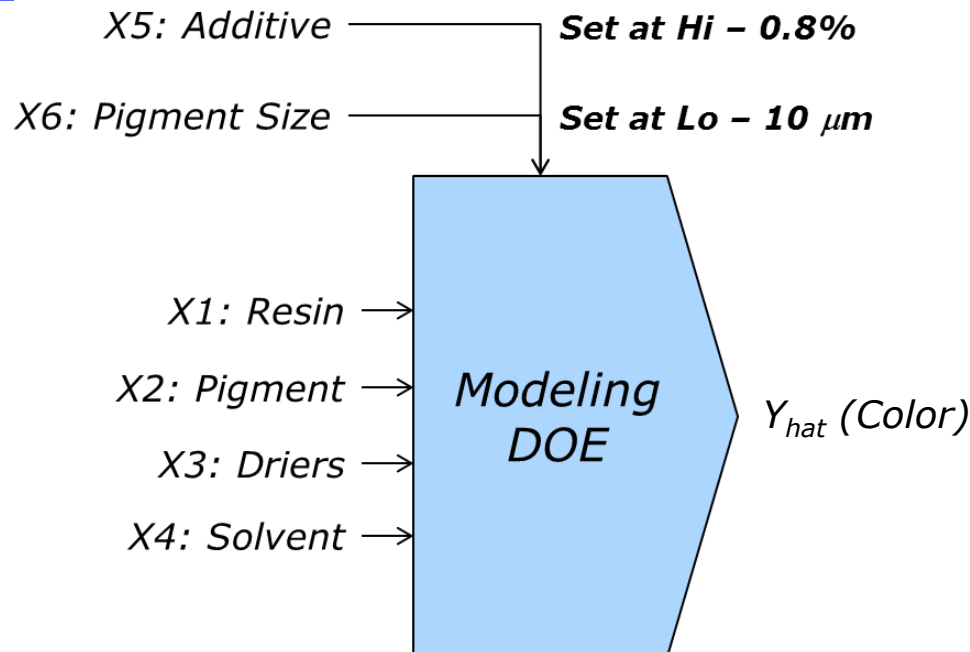
*Formulation for Y: Color
công thức cho Y (màu sắc)*



Modeling Experiment

- After completing Screening DOE, David is clearer about the significant variables in his process in resulting the Color level. He proceeds with Modeling DOE leaving the dropped variables to their optimized level based on the finding in Screening DOE.

Sau khi tìm được ảnh hưởng của từng yếu tố, David loại bỏ 2 yếu tố ít ảnh hưởng trong bước tiếp theo.





IPO Chart for Modeling DOE

Sơ đồ IPO cho DOE Modeling

*Variables dropped
from Modeling
experiment at put at
Optimum Level*

X5: Additive

Set at Hi – 0.8%

X6: Pigment Size

Set at Lo – 10 μm

*Variable to be
experimented
further using Full
Factorial Matrix*

X1: Resin →

X2: Pigment →

X3: Driers →

X4: Solvent →

**Modeling
DOE**

Y_{hat} S_{hat}



Full Factorial Design Matrix and Data Collection

Full Factorial, 4 factors, 16 runs, 3 replicates

Thực hiện thí nghiệm tối ưu: 4 yếu tố, 16 thí nghiệm, 3 phép lặp

Factor	A	B	C	D					
Row #	Resin	Pigment	Driers	Solvent	Y1	Y2	Y3	Y bar	S
1	40	5	0.5	5	0.41	0.45	0.43	0.43	0.02
2	40	5	0.5	30	0.70	0.70	0.68	0.693333	0.011547
3	40	5	2	5	0.46	0.46	0.48	0.466667	0.011547
4	40	5	2	30	0.68	0.70	0.69	0.69	0.01
5	40	20	0.5	5	0.25	0.24	0.21	0.233333	0.020817
6	40	20	0.5	30	0.65	0.64	0.68	0.656667	0.020817
7	40	20	2	5	0.24	0.22	0.24	0.233333	0.011547
8	40	20	2	30	0.58	0.62	0.64	0.613333	0.030551
9	60	5	0.5	5	0.45	0.40	0.48	0.443333	0.040415
10	60	5	0.5	30	0.65	0.68	0.71	0.68	0.03
11	60	5	2	5	0.44	0.36	0.42	0.406667	0.041633
12	60	5	2	30	0.68	0.68	0.72	0.693333	0.023094
13	60	20	0.5	5	0.20	0.30	0.26	0.253333	0.050332
14	60	20	0.5	30	0.65	0.58	0.64	0.623333	0.037859
15	60	20	2	5	0.18	0.22	0.25	0.216667	0.035119
16	60	20	2	30	0.62	0.65	0.60	0.623333	0.025166



Phân tích hồi quy lần 1 *1st Regression Analysis*

Y-hat Model		Color			
Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		0.49729	0.0000		
A	Resin	-0.00479	0.2587	1	X
B	Pigment	-0.06563	0.0000	1	X
C	Driers	-0.00438	0.3016	1	X
D	Solvent	0.16188	0.0000	1	X
AB		0.00229	0.5861	1	X
AC		-0.00313	0.4587	1	X
AD		0.0006250	0.8817	1	X
BC		-0.00563	0.1865	1	X
BD		0.03563	0.0000	1	X
CD		0.0002083	0.9604	1	X
ABC		0.00396	0.3492	1	X
ABD		-0.00396	0.3492	1	X
ACD		0.01063	0.0158	1	X
BCD		-0.00104	0.8042	1	X
ABCD		-0.0006250	0.8817	1	X
R ²		0.9829			
Adj R ²		0.9749			
Std Error		0.0289			
F		122.9345			
Sig F		0.0000			
F _{LOF}		NA			
Sig F _{LOF}		NA			
Source		SS	df	MS	
Regression		1.5	15	0.1	
Error		0.0	32	0.0	
Error _{pure}		0.0	32	0.0	
Error _{LOF}		0.0	0	NA	
Total		1.6	47		

Factor	Name	Low	High	Exper
A	Resin	40	60	50
B	Pigment	5	20	12.5
C	Driers	0.5	2	1.25
D	Solvent	5	30	17.5

2

Multiple Response Prediction

	Y-hat	S-hat	99% Confidence Interval	
			Lower Bound	Upper Bound
Color	0.4973	0.0263	0.418	0.576

Variables with P-value ≥ 0.05 are to be left out

Những biến cho $P \geq 0.05$ sẽ bị loại

$R^2 > 0.7$ shows the regression is a reliable one

$R^2 > 0.7$ cho thấy phép hồi quy đáng tin cậy

S-hat Model		Color			
Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		0.02628	NA		
A	Resin	0.00917	NA	1	X
B	Pigment	0.00275	NA	1	X
C	Driers	-0.00270	NA	1	X
D	Solvent	-0.00265	NA	1	X
AB		-0.00108	NA	1	X
AC		-0.00150	NA	1	X
AD		-0.00377	NA	1	X
BC		-0.0007347	NA	1	X
BD		0.00222	NA	1	X
CD		0.00127	NA	1	X
ABC		-0.00204	NA	1	X
ABD		-0.00140	NA	1	X
ACD		-0.00197	NA	1	X
BCD		0.00142	NA	1	X
ABCD		-0.0000908	NA	1	X
R ²		1.0000			
Adj R ²		NA			
Std Error		NA			
F		NA			
Sig F		NA			
F _{LOF}		NA			
Sig F _{LOF}		NA			
Source		SS	df	MS	
Regression		0.0	15	0.0	
Error		0.0	0	NA	
Error _{pure}		NA	0	NA	
Error _{LOF}		NA	0	NA	
Total		0.0	15		



Phân tích hồi quy cuối *Final Regression Analysis*

Y-hat Model		Color			
Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		0.49729	0.0000		
A	Resin	-0.00479	0.2305	1	X
B	Pigment	-0.06563	0.0000	1	X
C	Driers	-0.00438	0.2729	1	X
D	Solvent	0.16188	0.0000	1	X
BD		0.03563	0.0000	1	X
ACD		0.01063	0.0101	1	X
R ²		0.9805			
Adj R ²		0.9776			
Std Error		0.0273			
F		343.4010			
Sig F		0.0000			
F _{LOF}		0.5114			
Sig F _{LOF}		0.8553			
Source		SS	df	MS	
Regression		1.5	6	0.3	
Error		0.0	41	0.0	
Error _{Pure}		0.0	32	0.0	
Error _{LOF}		0.0	9	0.0	
Total		1.6	47		

Factor	Name	Low	High	Exper
A	Resin	40	60	50
B	Pigment	5	20	12.5
C	Driers	0.5	2	1.25
D	Solvent	5	30	17.5

Multiple Response Prediction				
99% Confidence Interval				
	Y-hat	S-hat	Lower Bound	Upper Bound
Color	0.4973	0.0263	0.418	0.576

Model Expression

$$\begin{aligned}
 Y_{hat} = & 0.497 \\
 & - 0.005(\text{Resin}) \\
 & - 0.066(\text{Pigment}) \\
 & - 0.004(\text{Driers}) \\
 & + 0.162(\text{Solvent}) \\
 & + 0.036(\text{Pigment})(\text{Solvent}) \\
 & + 0.011(\text{Resin})(\text{Driers})(\text{Solvent})
 \end{aligned}$$



Optimization (Color = 0.3 unit)

Tối ưu hóa (Màu sắc = 0.3 đơn vị)

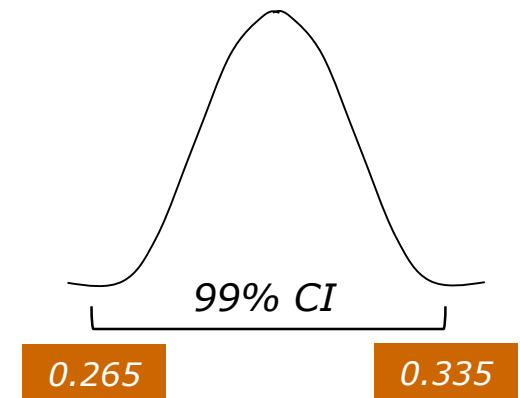
Factor yếu tố	Name Tên	Low Ngưỡng dưới	High Ngưỡng trên	Exper K.quả thí nghiệm
A	Resin	40	60	40
B	Pigment	5	20	15.94135008
C	Driers	0.5	2	2
D	Solvent	5	30	5

Màu sắc = 0.3 đơn vị
Setting that gives
Color = 0.3 unit

Dự đoán kết quả

Multiple Response Prediction

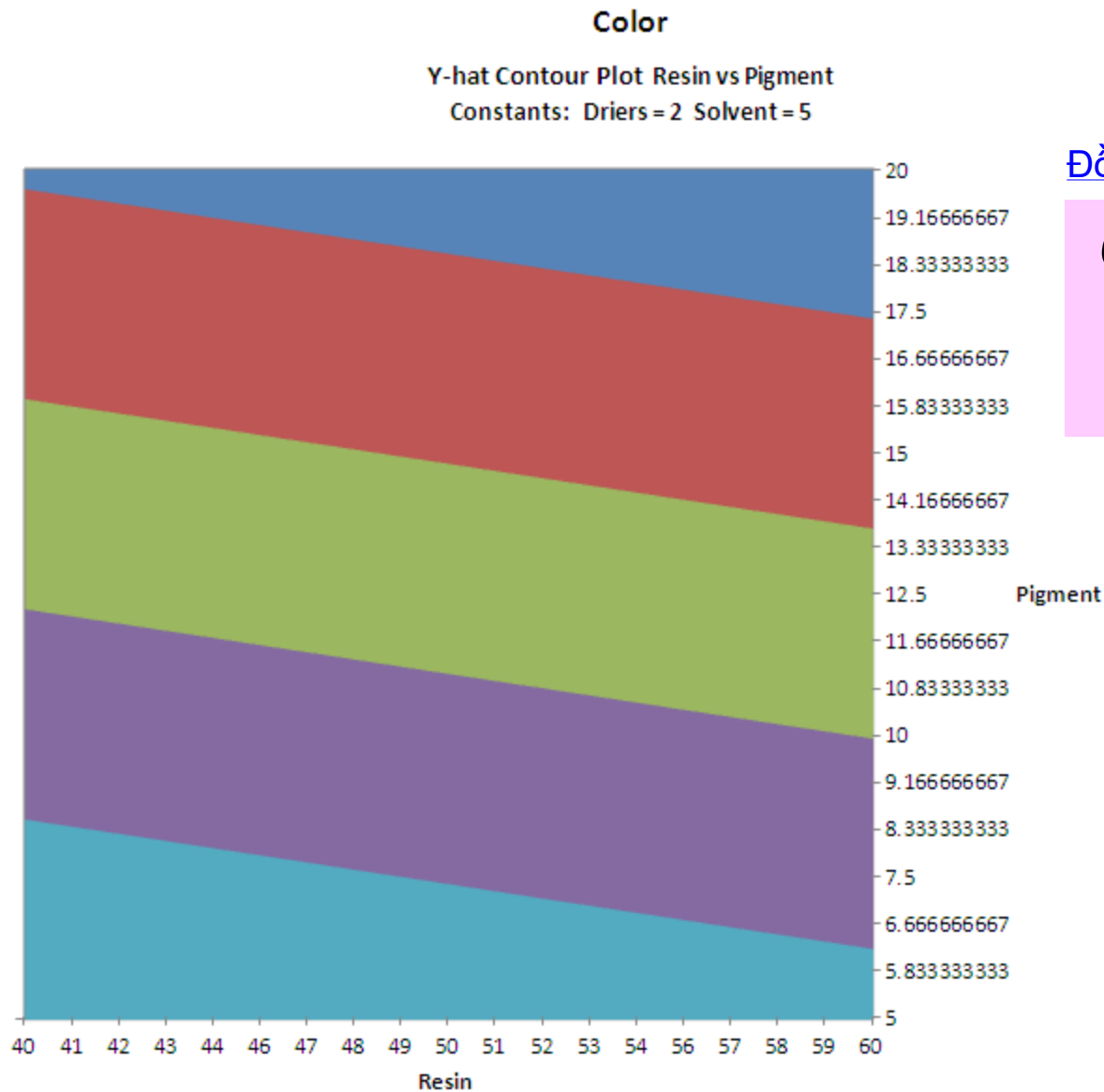
	Y-hat	S-hat	99% Confidence Interval	
			Giới hạn dưới Lower Bound	Giới hạn trên Upper Bound
Color	0.3000	0.0115	0.265	0.335





Đồ thị mô tả

Contour Chart Showing Options



Đồ thị cho thấy vùng cần lựa chọn

Contour chart shows region of choices



Factor	Name	Low	High	Exper
A	Resin	40	60	40
B	Pigment	5	20	15.94135008
C	Driers	0.5	2	2
D	Solvent	5	30	5

Run this setting ie 30x, and check if your result is as predicted

Chạy phép thử, và kiểm tra kết quả xem có gần với dự đoán không

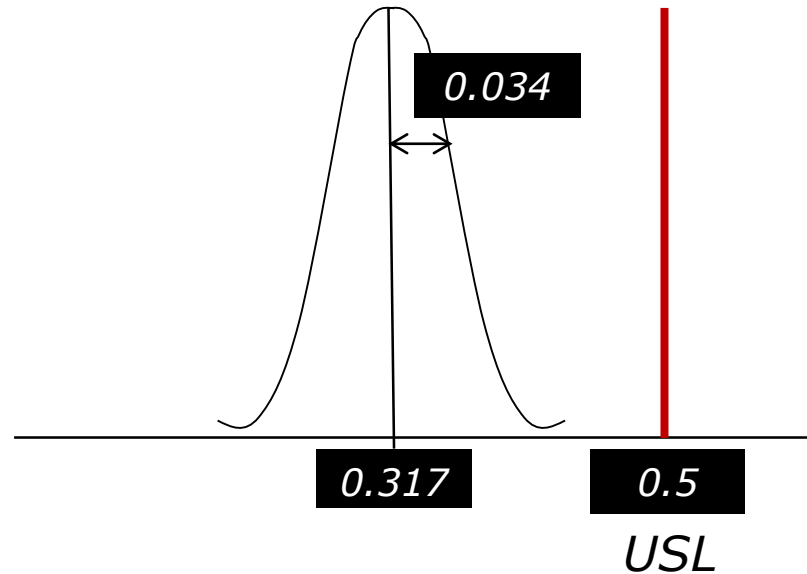
Multiple Response Prediction				
99% Confidence Interval				
	Y-hat	S-hat	Lower Bound	Upper Bound
Color	0.3000	0.0115	0.265	0.335



David chạy 10 phép thử để xác nhận, phần mềm sử dụng là Expert Optimizer.

- ◆ For validation purpose David performs 10 confirmation runs using setting recommended by Expert Optimizer.

#	Color Reading
1	0.31
2	0.33
3	0.36
4	0.29
5	0.34
6	0.3
7	0.3
8	0.28
9	0.28
10	0.38
Average	0.317
Std Dev	0.034
UCL	0.50
LCL	0.00
Cp	2.43
Cpk	1.78



Statistically, for C_p , C_{pk} at this level, the performance is very good. **Practically**, we keep the setting (and make pigment 16% for better control)

Theo lí thuyết, Giá trị C_p , C_{pk} tại mức này đều rất tốt.

Theo thí nghiệm, chúng ta giữ tỉ lệ màu 16% để điều chỉnh hiệu quả hơn

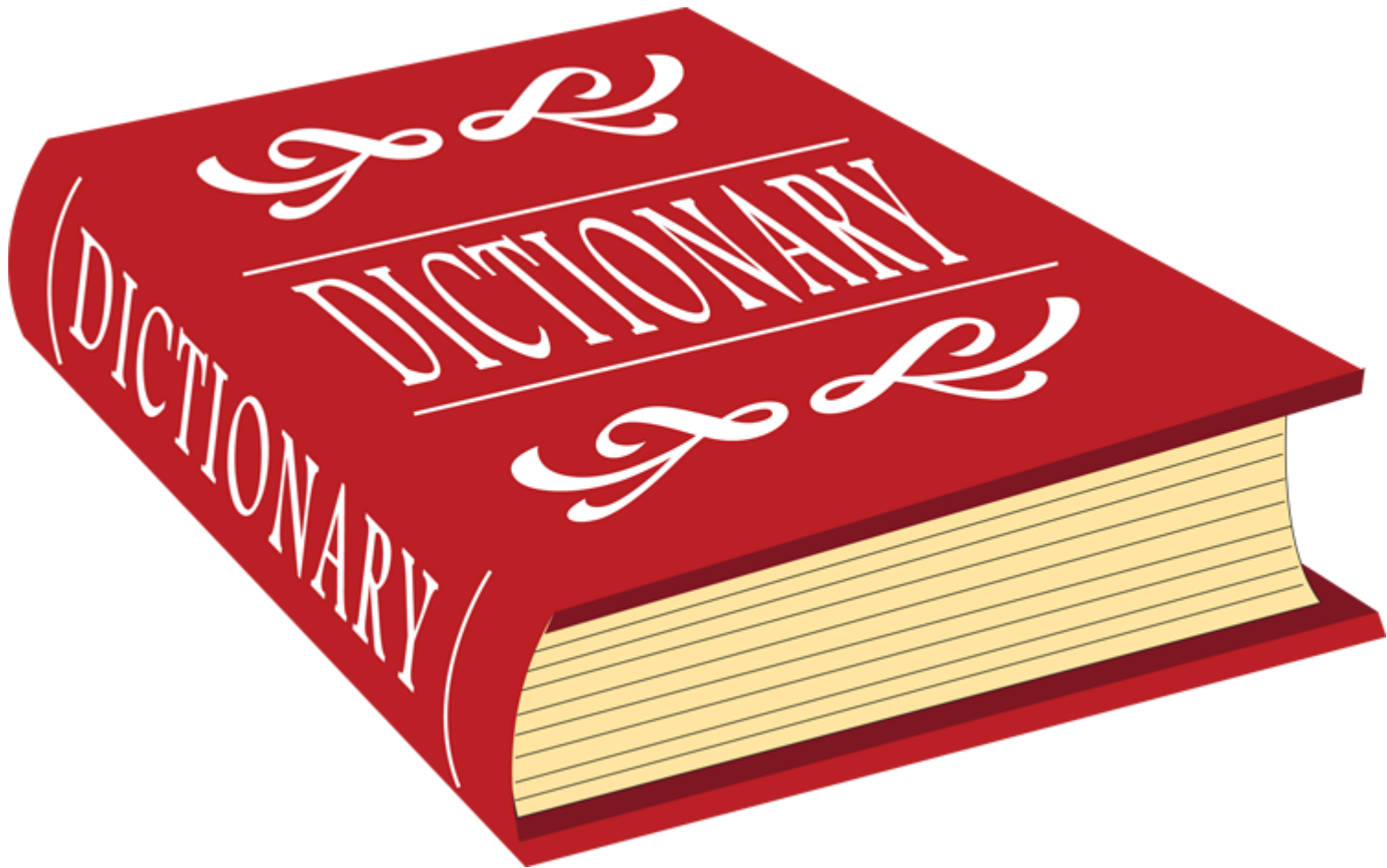
DOE *Terminologies*

THUẬT NGỮ DOE



DOE Operational Definitions..

Các định nghĩa



MY FIRST DOE

About DOE

Workshops

**DOE
TERMINOLOGIES**

1. *ANOVA*
2. *Factor Coding*
3. *Design Matrix*
4. *Main Effect*



[DOE Term 1] ANOVA

ANOVA: phân tích sự thay đổi, là quá trình toán học dùng để phân tách các kết quả trong một nhóm kết quả quan sát theo các nguyên nhân và tìm ra các thí nghiệm có sự thay đổi lớn nhất.

- ◆ **Analysis of Variance (ANOVA):** A mathematical process for separating the variability of a group of observations into assignable causes and setting up various significance tests.

David khảo sát xem thương hiệu dầu nhờn nào tốt nhất cho xe

David wants to find factors that give him better mileage for his new car. He is investigating the petrol brand (Petras or Cell)



Thực hiện thí nghiệm đơn giản

Simple systematic experiment

↓
*Use **ANOVA** to analyze the experimental data*

Sử dụng ANOVA để phân tích dữ liệu thí nghiệm

↓
ANOVA will tell either

- Petrol brand is the dominant factor, or
- Petrol type is not the dominant factor (meaning other untested factors are more dominant, ie. road condition, car speed)

ANOVA cho kết quả:

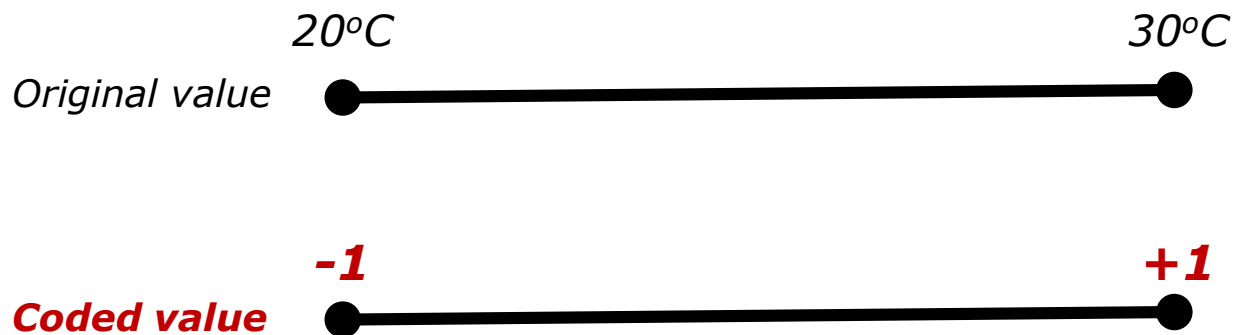
- loại dầu nhờn nào có ảnh hưởng lớn
- loại dầu nhờn nào ko gây ảnh hưởng mấy



[DOE Term 2] Factor Coding

[Mã hóa mức/giá trị của các yếu tố/biến đầu vào](#)

- ◆ **Coding Factor Levels:** Transforming the scale of measurement for a factor so that the high value becomes +1 and the low value becomes -1. Coding is a simple linear transformation of the original measurement scale along the experimental range. If we are experimenting painting at temperature (the X factor) ranged from 20° to 30°C, then 20°C is X_{lo} (-1) and 30°C is X_{hi} (+1).





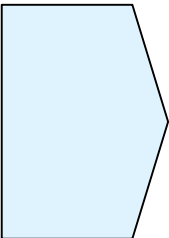
[DOE Term 3] Design Matrix

- ◆ **Design Matrix:** DOE is a systematic experiments. Unlike other experiments, in DOE the input factors are manipulated systematically following certain combination disciplines. The most commonly used combination or design matrix in Fractional Factorial and Full Factorial.

Fractional Factorial

(8 runs, Half Fraction, 4 factors)

X1
X2
X3
X4

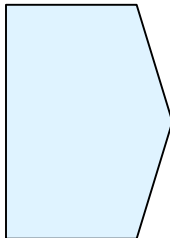


Run	X1	X2	X3	X4
1	Lo	Lo	Lo	Lo
2	Lo	Lo	Hi	Hi
3	Lo	Hi	Lo	Hi
4	Lo	Hi	Hi	Lo
5	Hi	Lo	Lo	Hi
6	Hi	Lo	Hi	Lo
7	Hi	Hi	Lo	Lo
8	Hi	Hi	Hi	Hi

Full Factorial

(8 runs, 3 factors)

X1
X2
X3



Run	X1	X2	X3
1	Lo	Lo	Lo
2	Lo	Lo	Hi
3	Lo	Hi	Lo
4	Lo	Hi	Hi
5	Hi	Lo	Lo
6	Hi	Lo	Hi
7	Hi	Hi	Lo
8	Hi	Hi	Hi

ma trận thiết kế: DOE là hệ thống các thí nghiệm. Biến đầu vào có thể điều chỉnh có hệ thống theo nguyên tắc tổ hợp. 2 nguyên tắc thường dùng là: Fractional Factorial và Full fractional



[DOE Term 4] Main Effect

Ảnh hưởng chính (của 1 biến đầu vào): Khi DOE hoàn thành, phần mềm DOE sẽ thiết lập công thức và cho ra mối quan hệ giữa biến đầu vào và đầu ra trong thí nghiệm (hoặc đơn giản là 1 phương trình toán học)

- ◆ **Main Effect (of a factor):** When DOE is completed the DOE software formulates the model or derives the input factors relationship with the experiment response into what is called model expression (or simply the mathematical equation).

$$\hat{Y} = k + k_1X_1 + k_2X_2 + k_3X_1X_2$$

Predicted Response

Single component biến đơn

Combined component biến phức

Tác động chính Main Effect

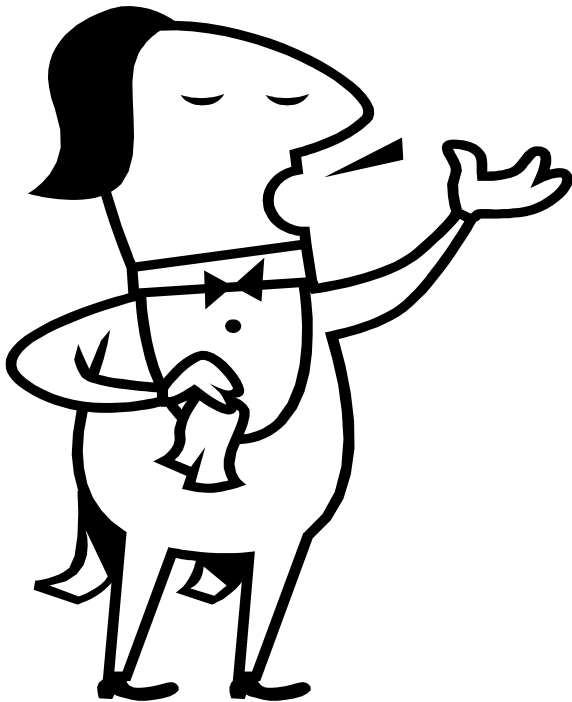
Hiệu ứng tương tác Interaction Effect

- ◆ In Screening experiment, due to its simplified number of runs, it puts focus on determining the strength of the main effects (and less focus on the interaction). **Main effect tells the effect size the individual input factor influences the response.**

Trong thí nghiệm Screening, vì đơn giản hóa số thí nghiệm, chỉ tập trung vào việc xác định độ mạnh của các tác động chính. Các ảnh hưởng chính cho ta biết độ mạnh của tác động của các biến đầu vào lên kết quả đầu ra.



MODULE OBJECTIVES DELIVERED..



*Can I take
your last order Sir?*