Define.Measure.Analyze.Improve.Control

Asian Productivity Organization

Feb 2019

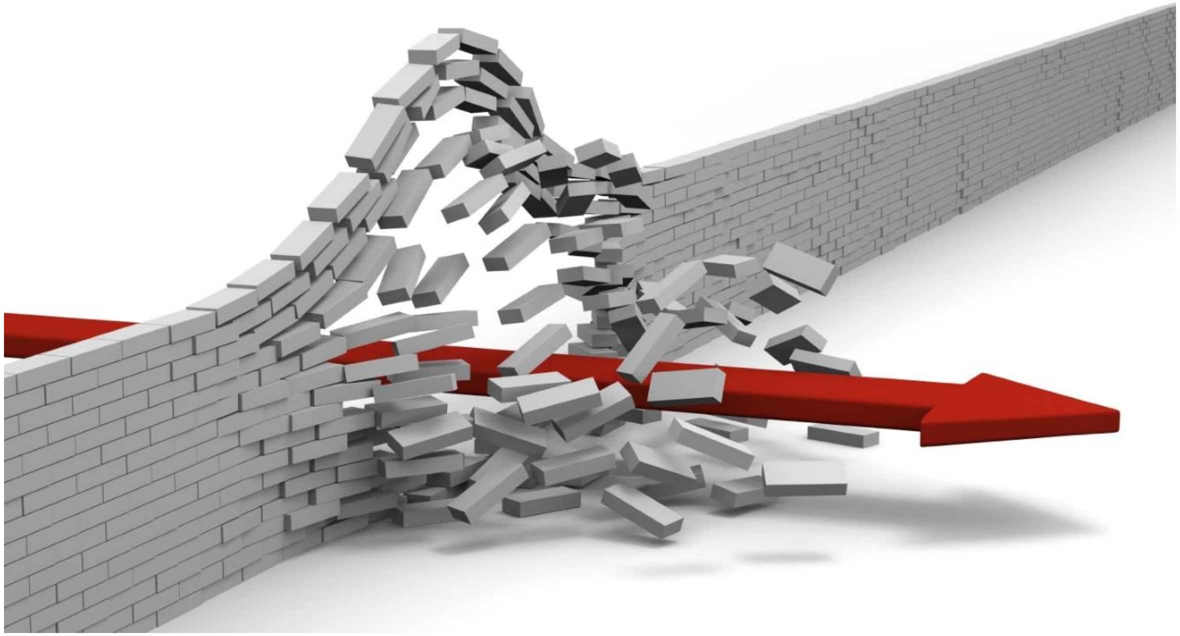
6σ

series

DOE FOR PRACTITIONERS

Delivered by:
Azman Hussain

Hint of the Module



BREAKTHROUGH

made easy by DOE methodology

DOE FOR PRACTITIONERS – *Main Topics*

0. About the Course
1. Multiple Regression
2. Screening DOE
3. Modelling DOE
4. Overviewing DOE Mixtures

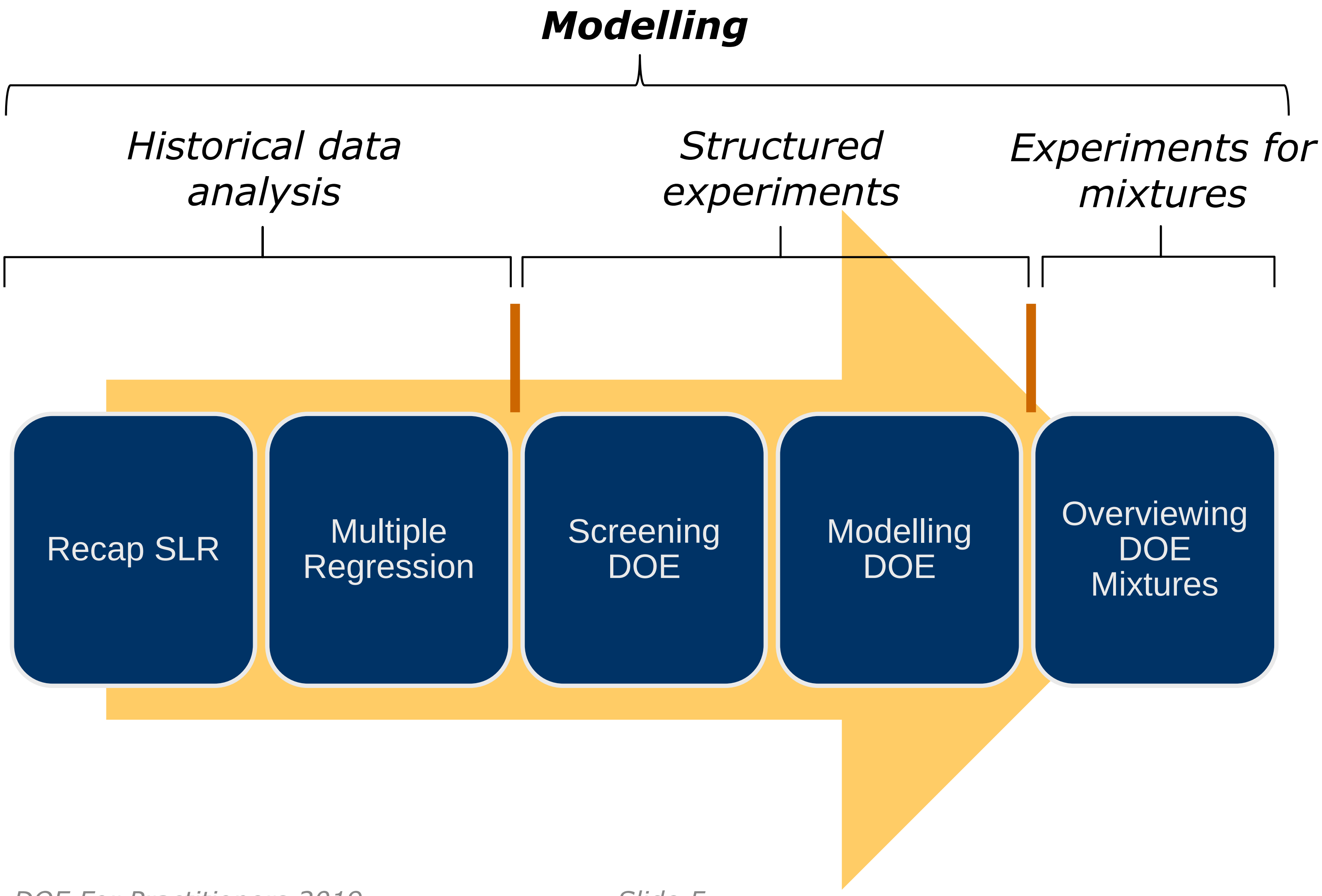
DOE for Practitioners

ABOUT THE COURSE

Multiple regression

Screening DOE

0. Course Layout



0-1. 3-days Workshop to Follow DOE Course

- DOE for Practitioners
  - Workshop 1: **Screening DOE**
  - Workshop 2: **Modelling DOE**
  - Workshop 3: **Validation test**

- *DOE Experts*
  - *Workshop 1:* ***Mixture DOE***
  - *Workshop 2:* ***3 level DOE***
  - *Workshop 3:* ***Taguchi method***

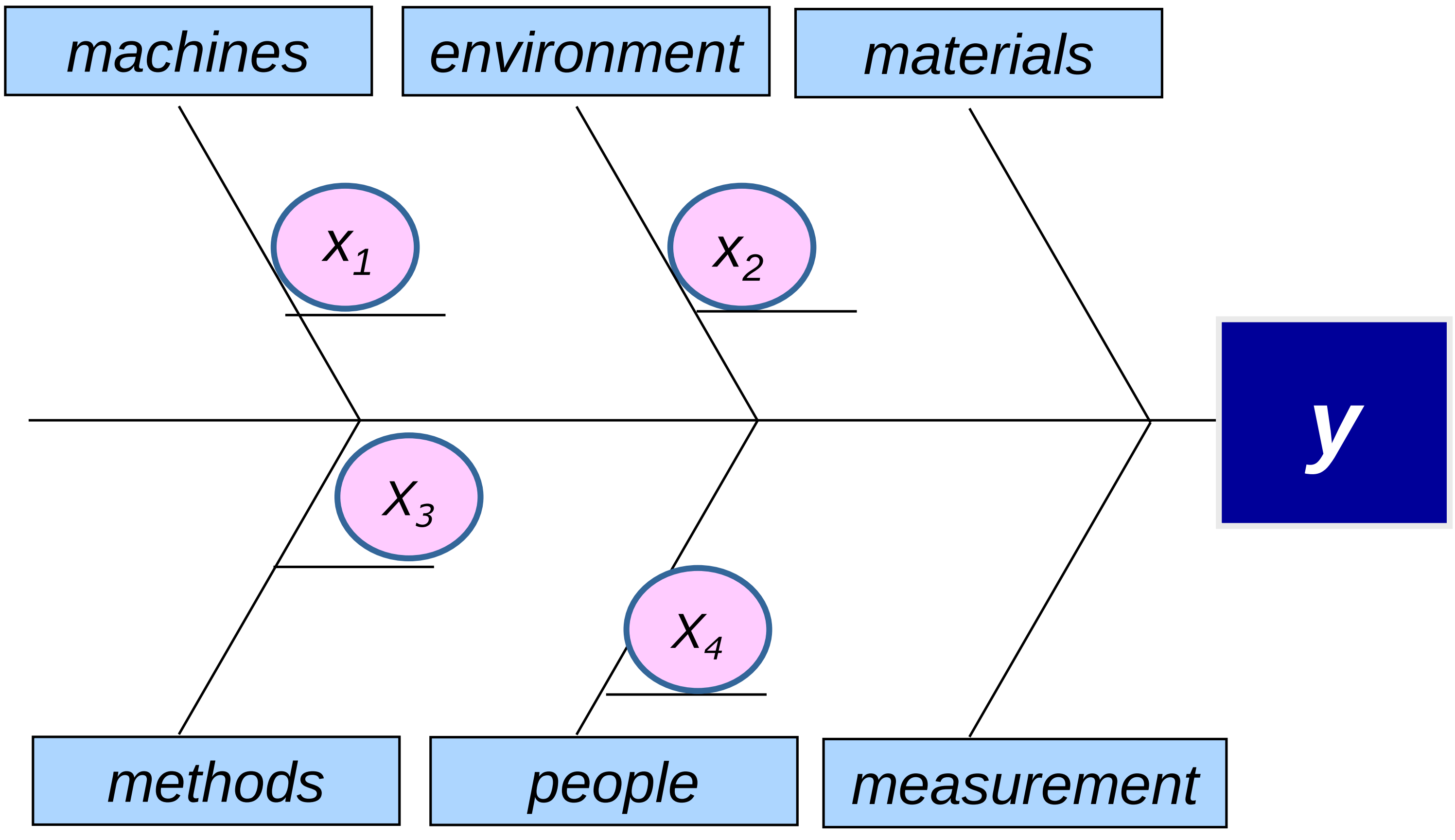
DOE for Practitioners

About the course

MULTIPLE REGRESSION

Screening DOE

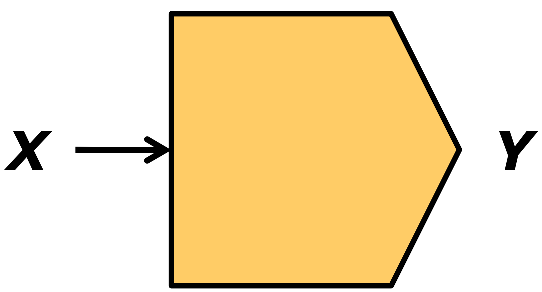
Hint about the Module...

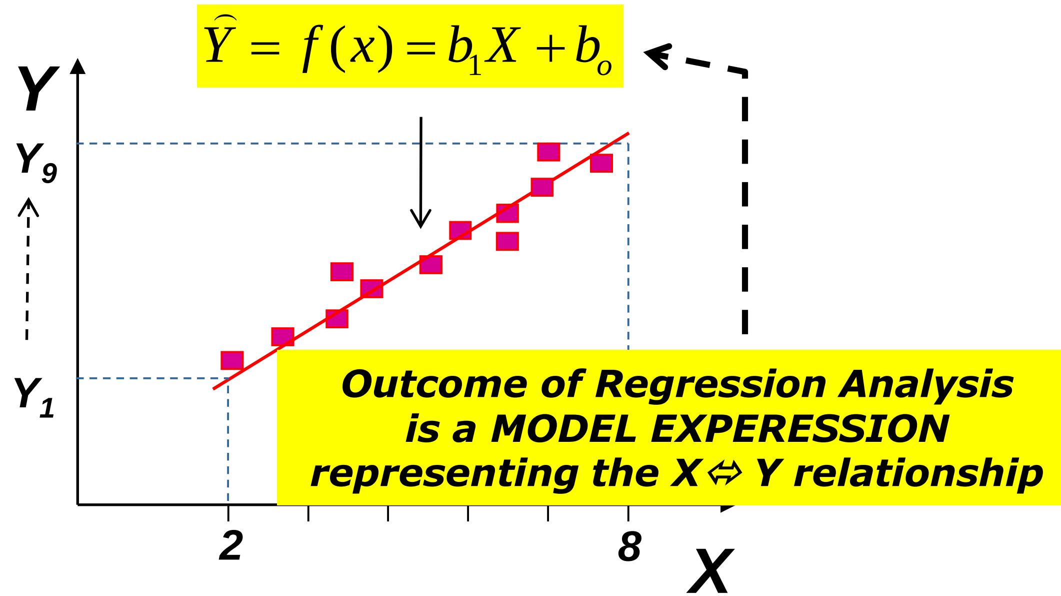


1-1. Recall Simple Linear Regression Model

- Simple regression analysis

Simple linear regression *is a technique in parametric statistics that is commonly used for analysing mean response of a variable Y which changes according to the magnitude of an intervention variable X.*

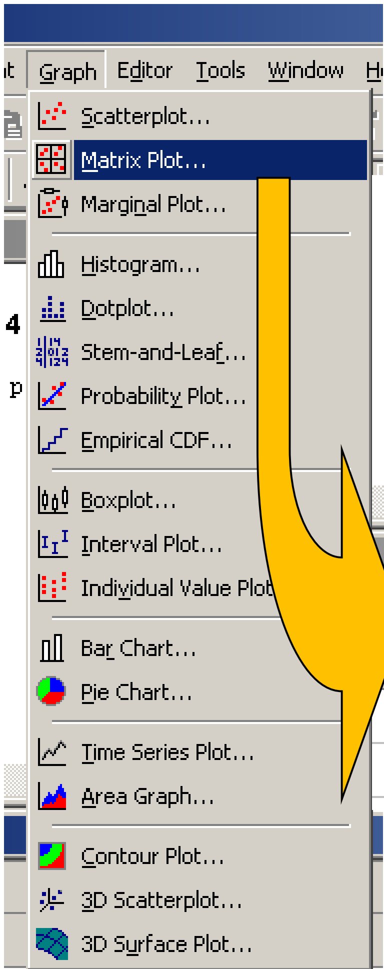


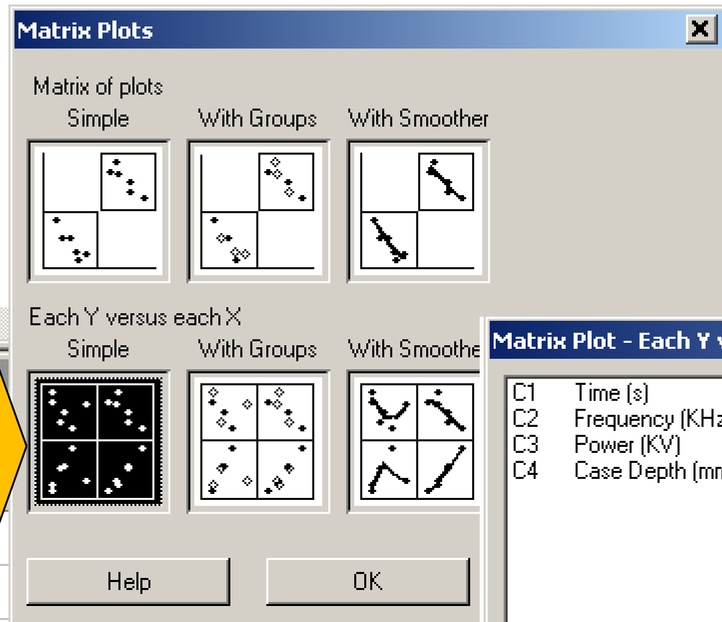


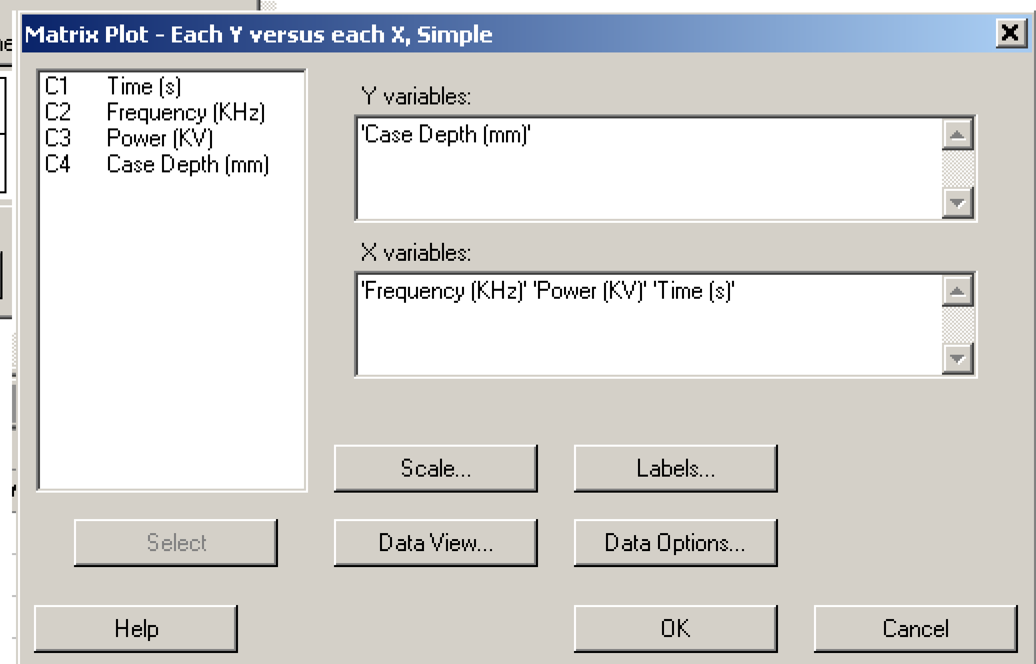
1-211. Multiple Regression – *Heat Treatment Example*

- A production engineer has observed that a number of settings of an induction-heating process are continually adjusted by the operators and wishes to understand the effect this has on variation in case depth
- The settings are:
  - Heating time
  - Power
  - Coil frequency

1-212. Visualising the Data – The Draftsman's Plot

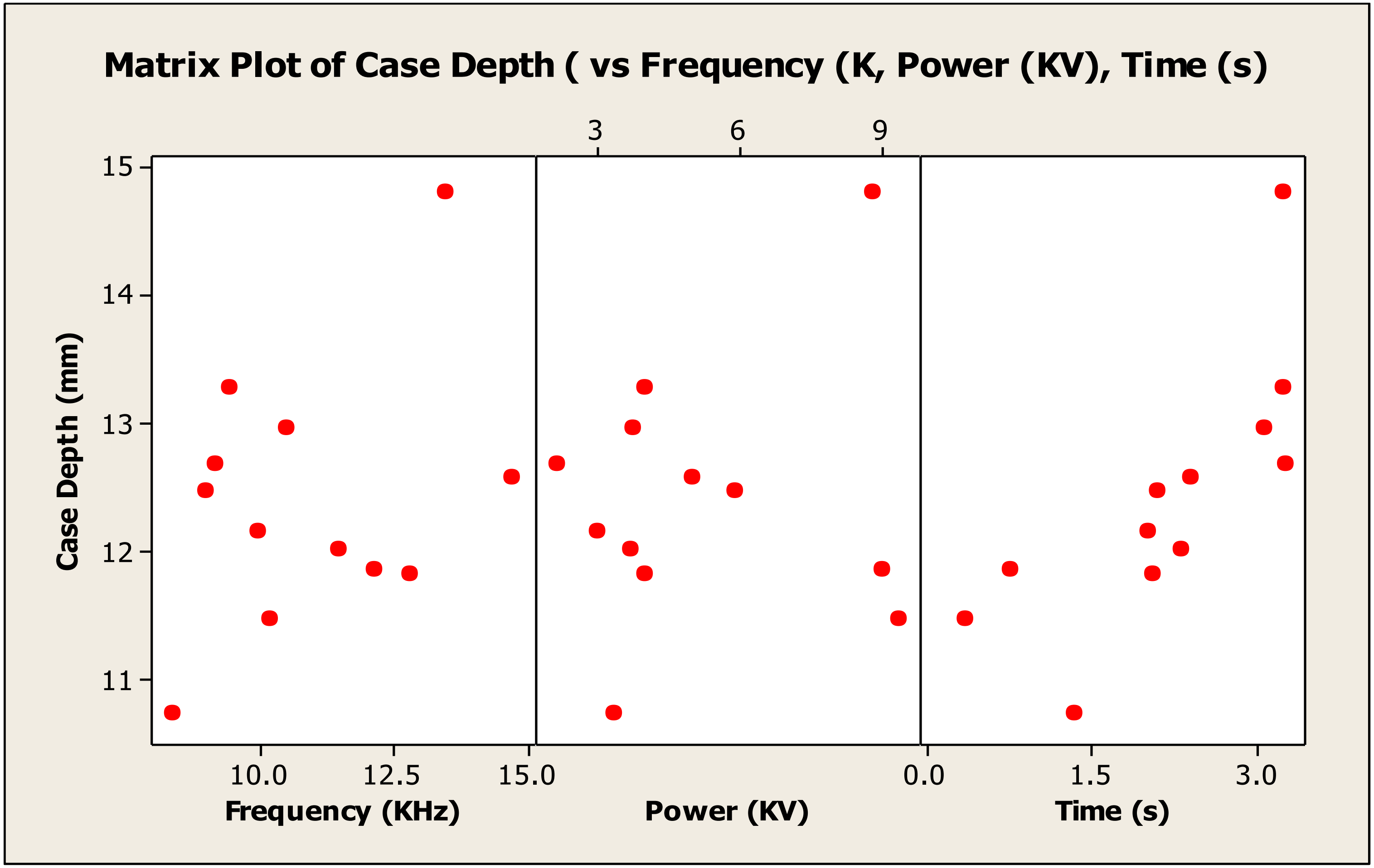






Remember always plot the data first

1-213. Minitab's Draftsman's Plot



1-214. Heat Treatment Study – Single Variable

Regression Analysis: Case Depth (mm) versus Time (s)

The regression equation is
Case Depth (mm) = 10.7 + 0.791 Time (s)

Predictor	Coef	SE Coef	T	P
Constant	10.6833	0.5087	21.00	0.000
Time (s)	0.7905	0.2153	3.67	0.004

S = 0.6967   R-Sq = 57.4%   R-Sq(adj) = 53.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	6.5429	6.5429	13.48	0.004
Residual Error	10	4.8539	0.4854		
Total	11	11.3969			

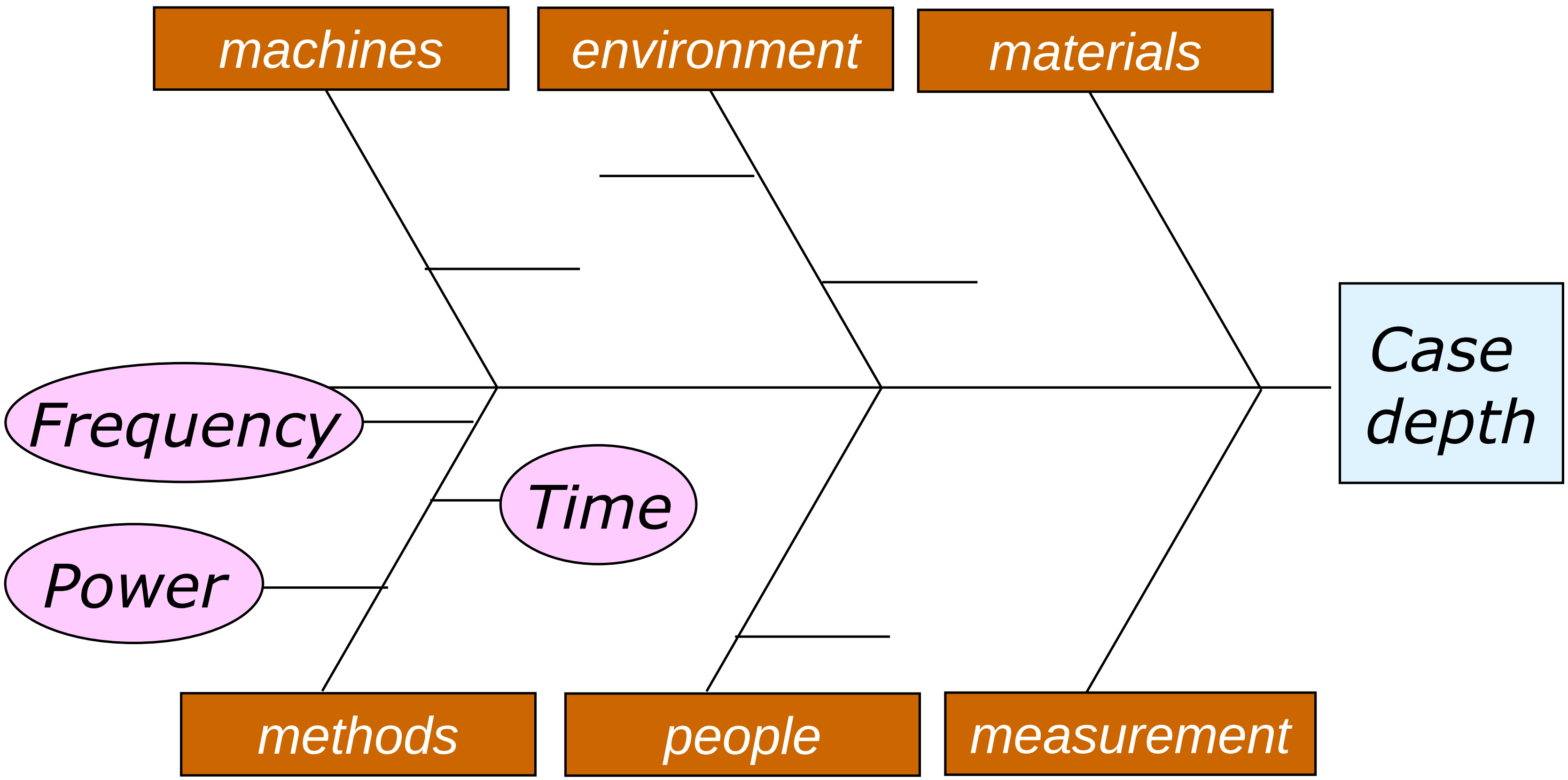
Unusual Observations

Obs	Time (s)	Case Dep	Fit	SE Fit	Residual	St Resid
10	3.24	14.795	13.245	0.306	1.551	2.48R

R denotes an observation with a large standardized residual

Looking at 'time', for example, as a single predictor

1-215. Back to the Fishbone



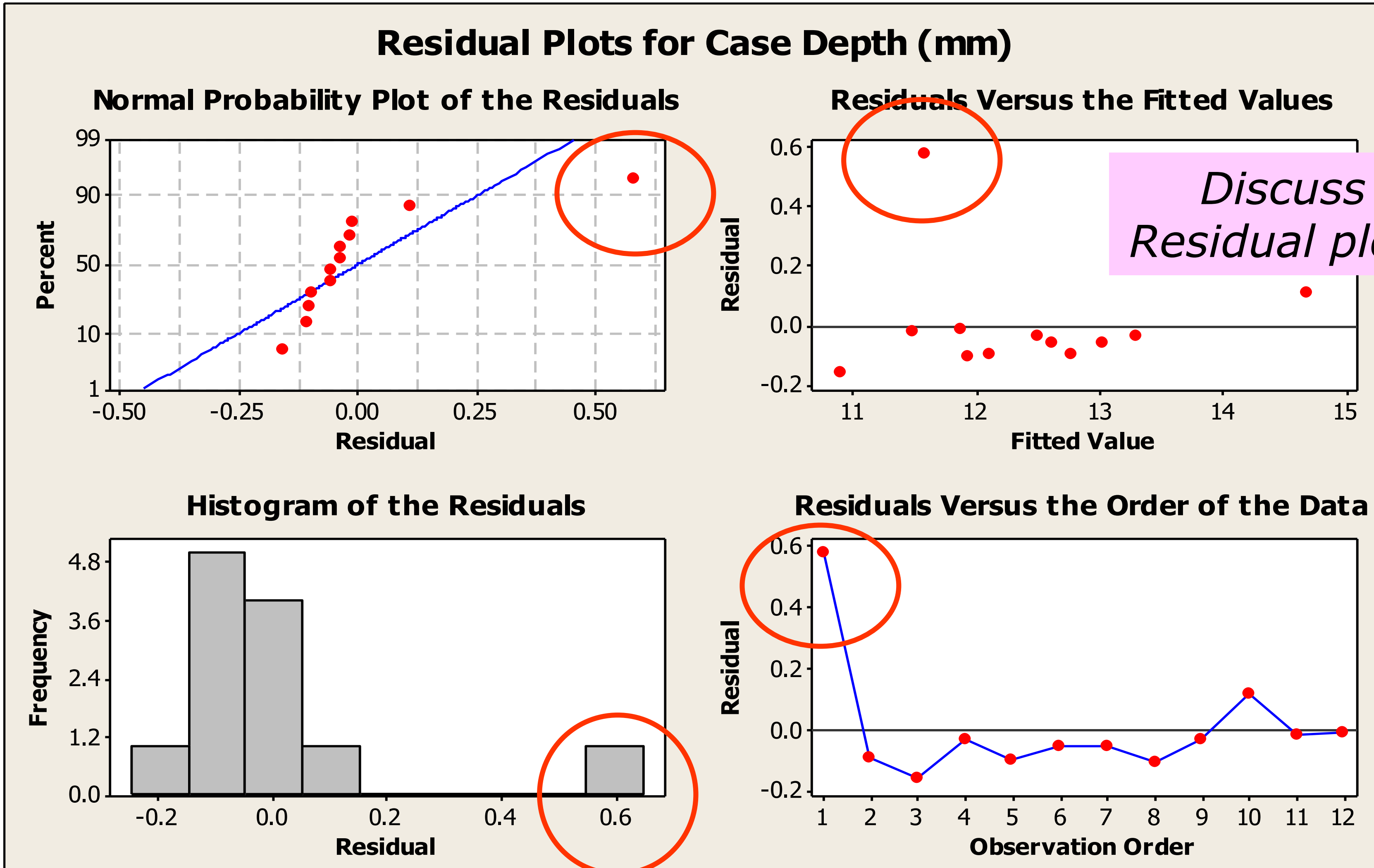
1-216. Regression – Using Multi Variable

Regression Analysis: Case Depth (versus Time (s), Frequency (K, ...

```
The regression equation is
Case Depth (mm) = 8.39 + 1.16 Time (s) + 0.0016 Frequency (KHz)
           + 0.288 Power (KV)

Predictor        Coef    SE Coef       T      P
Constant       8.3878     0.4139   20.27  0.000
Time (s)      1.15637    0.08566   13.50  0.000
Frequenc      0.00159    0.03931    0.04  0.969
Power (K      0.28763    0.03539    8.13  0.000

S = 0.2288     R-Sq = 96.3%     R-Sq(adj) = 94.9%

Analysis of Variance

Source           DF        SS       MS      F      P
Regression        3   10.9780   3.6593  69.90  0.000
Residual Error    8    0.4188   0.0524
Total            11   11.3969

Source      DF   Seq SS
Time (s)     1   6.5429
Frequenc     1   0.9769
Power (K     1   3.4582

Unusual Observations
Obs  Time (s)  Case Dep      Fit   SE Fit  Residual  St Resid
  1      2.00   12.1611  11.5794   0.1003    0.5817     2.83R

R denotes an observation with a large standardized residual
```

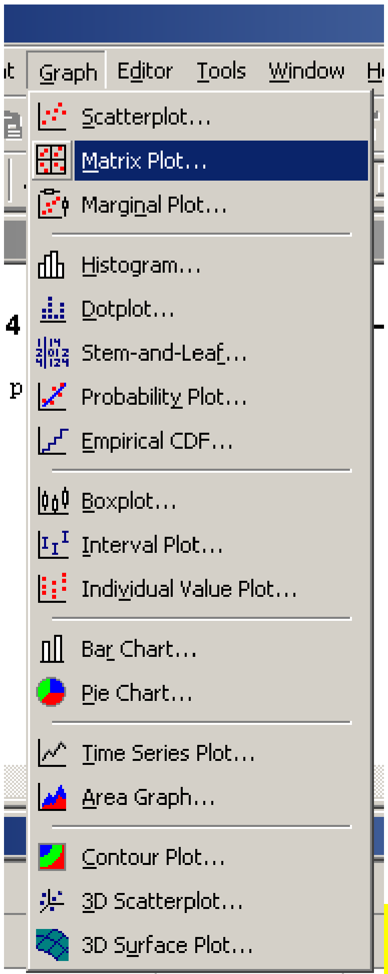
Definitely a better model!

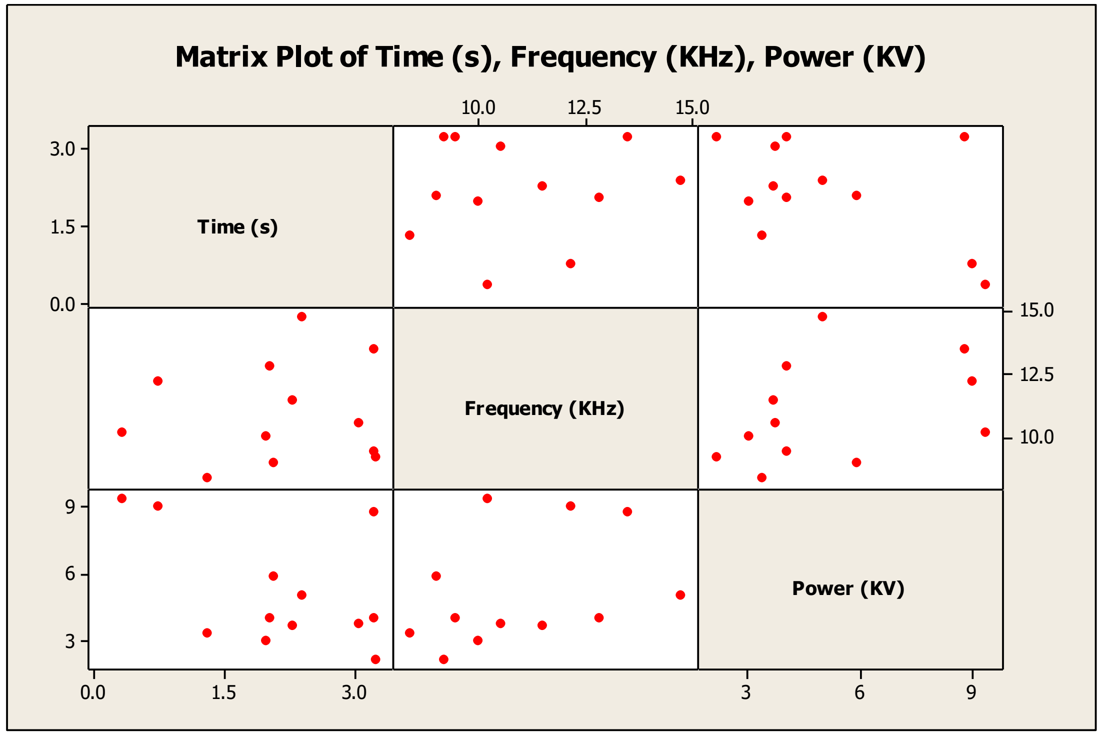
1-217. Residual Plots



Discuss Residual plots

1-218. Predictor Relationships – Back to the Matrix Plot Function





Relationships between predictors can be visualised using Minitab's matrix plots option

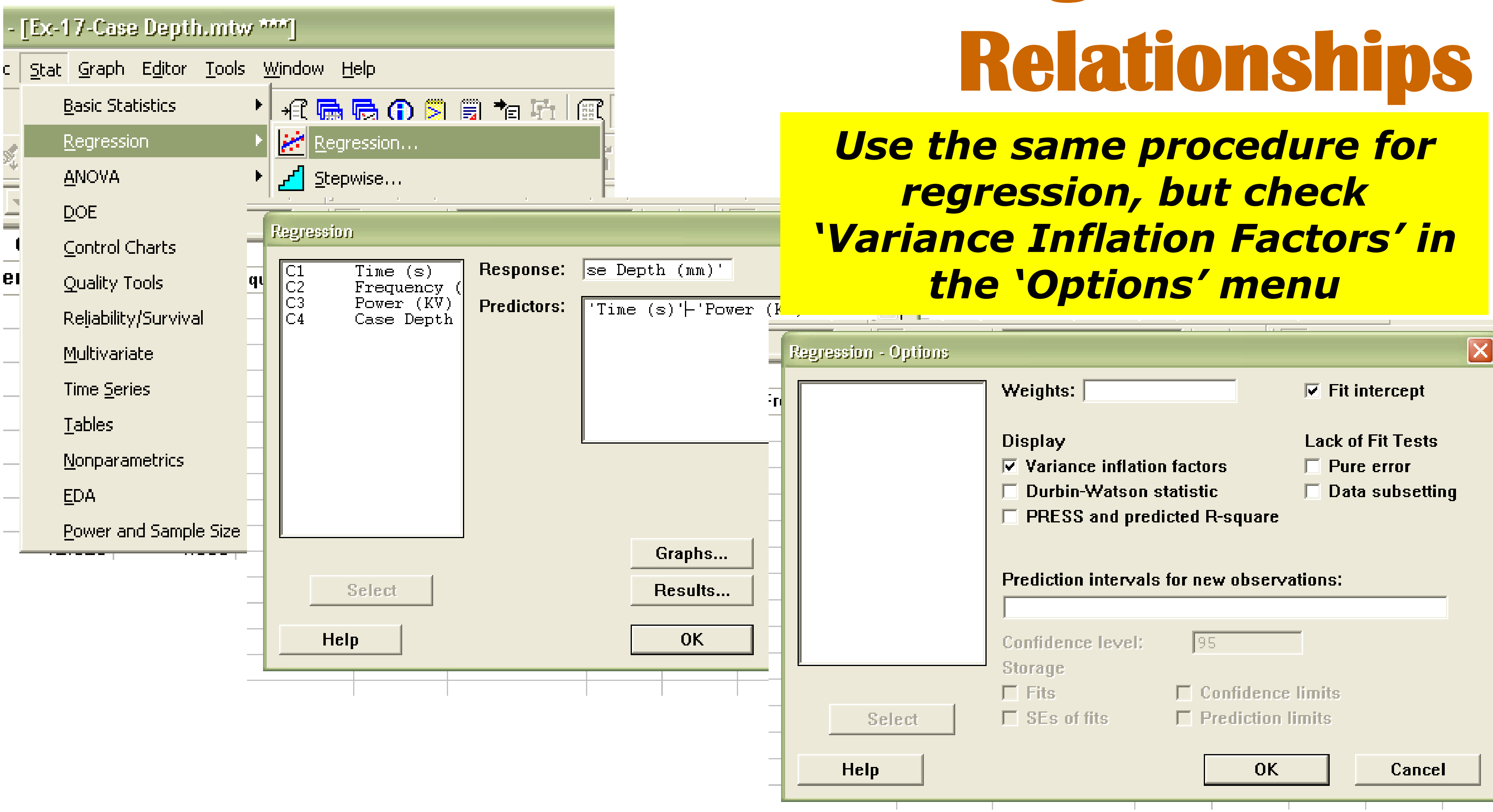
1-219. Improving the Model

Regression Analysis: Case Depth (mm) versus Time (s), Power (KV)

```
The regression equation is
Case Depth (mm) = 8.40 + 1.16 Time (s) + 0.288 Power (KV)

Predictor        Coef    SE Coef       T      P
Constant       8.3994     0.2821   29.78  0.000
Time (s)      1.15748    0.07654   15.12  0.000
Power (K      0.28830    0.02953    9.76  0.000

S = 0.2157     R-Sq = 96.3%     R-Sq(adj) = 95.5%

Analysis of Variance

Source           DF        SS       MS       F      P
Regression        2   10.9780   5.4890  117.93  0.000
Residual Error    9    0.4189   0.0465
Total            11   11.3969

Source       DF    Seq SS
Time (s)      1    6.5429
Power (K      1    4.4350

Unusual Observations
Obs  Time (s)  Case Dep      Fit   SE Fit  Residual  St Resid
  1      2.00   12.1611  11.5793   0.0945    0.5818     3.00R

R denotes an observation with a large standardized residual
```

As is evident here the model can be further improved by taking out frequency as a predictor as it was not proven significant in the earlier study

1-221. Variance Inflation Factors – Assessing Predictor Relationships



Use the same procedure for regression, but check 'Variance Inflation Factors' in the 'Options' menu

1-222. Using VIFs

Regression Analysis: Case Depth (mm) versus Time (s), Power (KV)

```
The regression equation is
Case Depth (mm) = 8.40 + 1.16 Time (s) + 0.288 Power (KV)

Predictor         Coef      SE Coef         T        P      VIF
Constant        8.3994       0.2821     29.78    0.000
Time (s)       1.15748      0.07654     15.12    0.000      1.3
Power (K       0.28830      0.02953      9.76    0.000      1.3

S = 0.2157      R-Sq = 96.3%      R-Sq(adj) = 95.5%

Analysis of Variance

Source            DF         SS          MS         F        P
Regression         2    10.9780      5.4890    117.93    0.000
Residual Error     9     0.4189      0.0465
Total             11    11.3969

Source        DF     Seq SS
Time (s)       1     6.5429
Power (K       1     4.4350

Unusual Observations
Obs    Time (s)    Case Dep        Fit     SE Fit    Residual    St Resid
  1        2.00     12.1611    11.5793     0.0945      0.5818       3.00R

R denotes an observation with a large standardized residual
```

VIFs close to 1.0 therefore little multicollinearity (<5 is OK)

P value is about the significance of the total regression

With no multi-collinearity, partition of SS does not depend on order

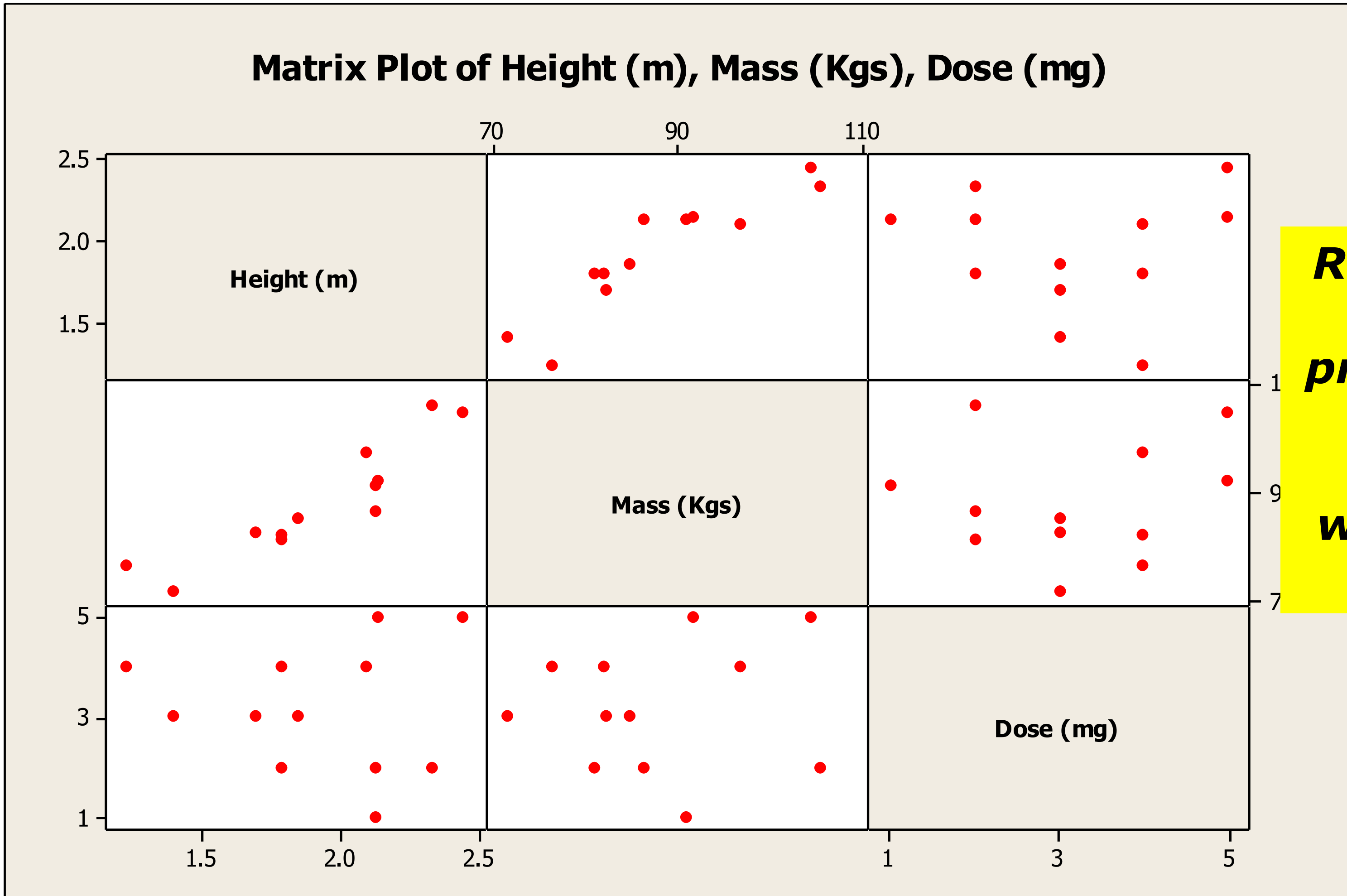
1-223. What High VIFs Shows – Example: Drug Concentration

Regression Analysis: Drug Concentrati versus Height (m), Mass (Kgs), ...

```
The regression equation is
Drug Concentration (mg/ml) = 44.3 + 0.00010 Height (m) - 0.199 Mass (Kgs)
           + 5.96 Dose (mg)

Predictor         Coef      SE Coef          T        P      VIF
Constant       44.3446       0.0021   21176.55    0.000
Height (      0.000095     0.001459       0.07    0.949      6.7
Mass (Kg     -0.198886     0.000051   -3918.84    0.000      6.8
Dose (mg       5.95631      0.00017   34718.51    0.000      1.1

S = 0.0006825   R-Sq = 100.0%    R-Sq(adj) = 100.0%

Analysis of Variance

Source            DF         SS         MS         F        P
Regression         3     633.43     211.14 4.533E+08    0.000
Residual Error     8       0.00       0.00
Total             11     633.43

Source        DF      Seq SS
Height (       1       44.42
Mass (Kg       1       27.60
Dose (mg       1      561.41
```

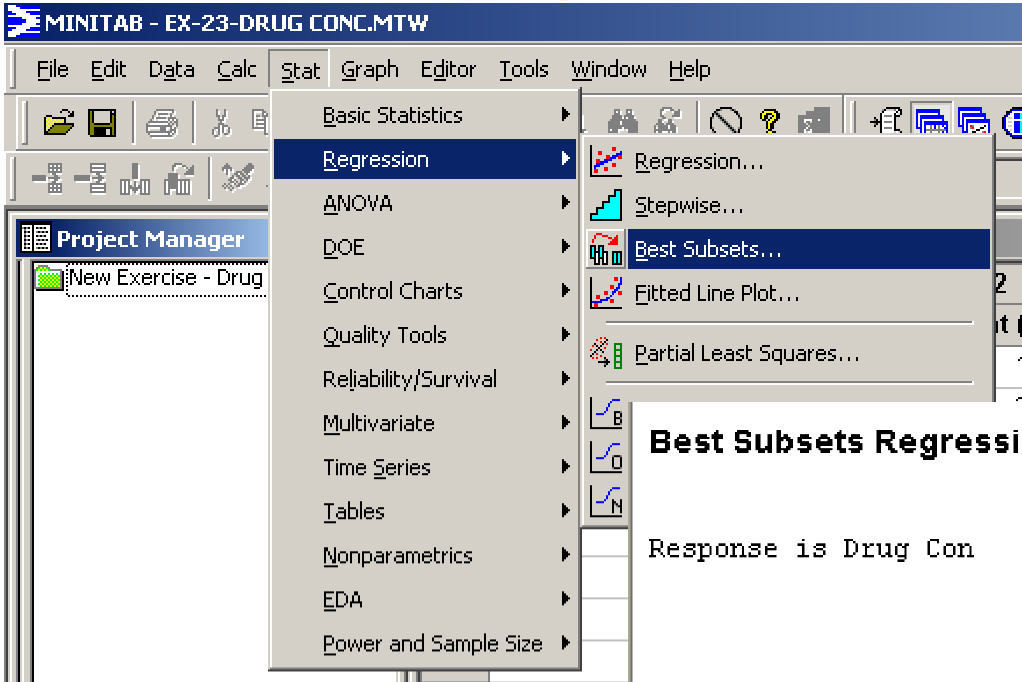
Although the three predictors are found to explain 100% of the variation in the response the effect of height and mass can not be assigned unambiguously

1-231. Plotting Matrix Plots – Example: Drug Concentration



Relationships between predictors are easily visualised with a Matrix Plot

1-232. Best Subsets Regression – Session Window Output



The model incorporating mass and dose is selected as the best one to use

Best Subsets Regression: Drug Concent versus Height (m), Mass (Kgs), ...

Response is Drug Con

Vars	R-Sq	R-Sq(adj)	C-p	S	Height (	Mass (Kg	Dose (mg
1	92.5	91.7	1E+08	2.1851			X
1	7.0	0.0	1E+09	7.6747	X		
2	100.0	100.0	2.0	0.0006436		X	X
2	98.9	98.6	2E+07	0.89149	X		X
3	100.0	100.0	4.0	0.0006825	X	X	X

1-232. *(cont)* Best Subsets Regression – Session Window Output

25/06/2003 15:22:36

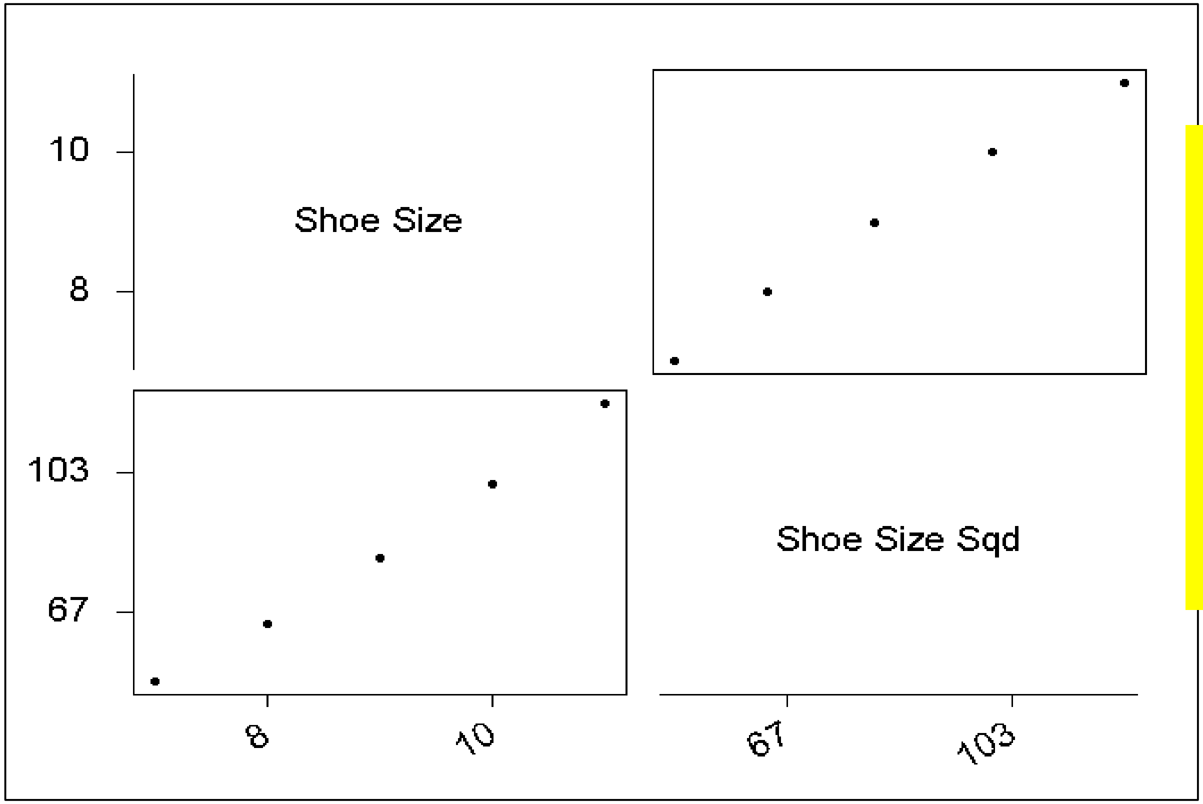
Regression Analysis: Drug Concentrati versus Mass (Kgs), Dose (mg)

```
The regression equation is
Drug Concentration (mg/ml) = 44.3 - 0.199 Mass (Kgs) + 5.96 Dose (mg)

Predictor        Coef     SE Coef          T       P     VIF
Constant      44.3446      0.0017   26823.53   0.000
Mass (Kg    -0.198883    0.000019  -10736.46   0.000     1.0
Dose (mg      5.95631     0.00015   38618.13   0.000     1.0

S = 0.0006436   R-Sq = 100.0%    R-Sq(adj) = 100.0%

Analysis of Variance

Source            DF        SS        MS          F       P
Regression         2    633.43    316.71  7.646E+08   0.000
Residual Error     9      0.00      0.00
Total             11    633.43

Source       DF     Seq SS
Mass (Kg      1      15.67
Dose (mg      1     617.76
```

The numbers now look good - we've got our model!

Multi-Collinearity & Data Transformations

Note that in the previous two examples higher ordered terms have not been considered (initially always use 'Fitted Line Plots' to visualise the form of relationships).



Predictor correlation between linear and quadratic terms, however, can be observed in the height versus shoe size study

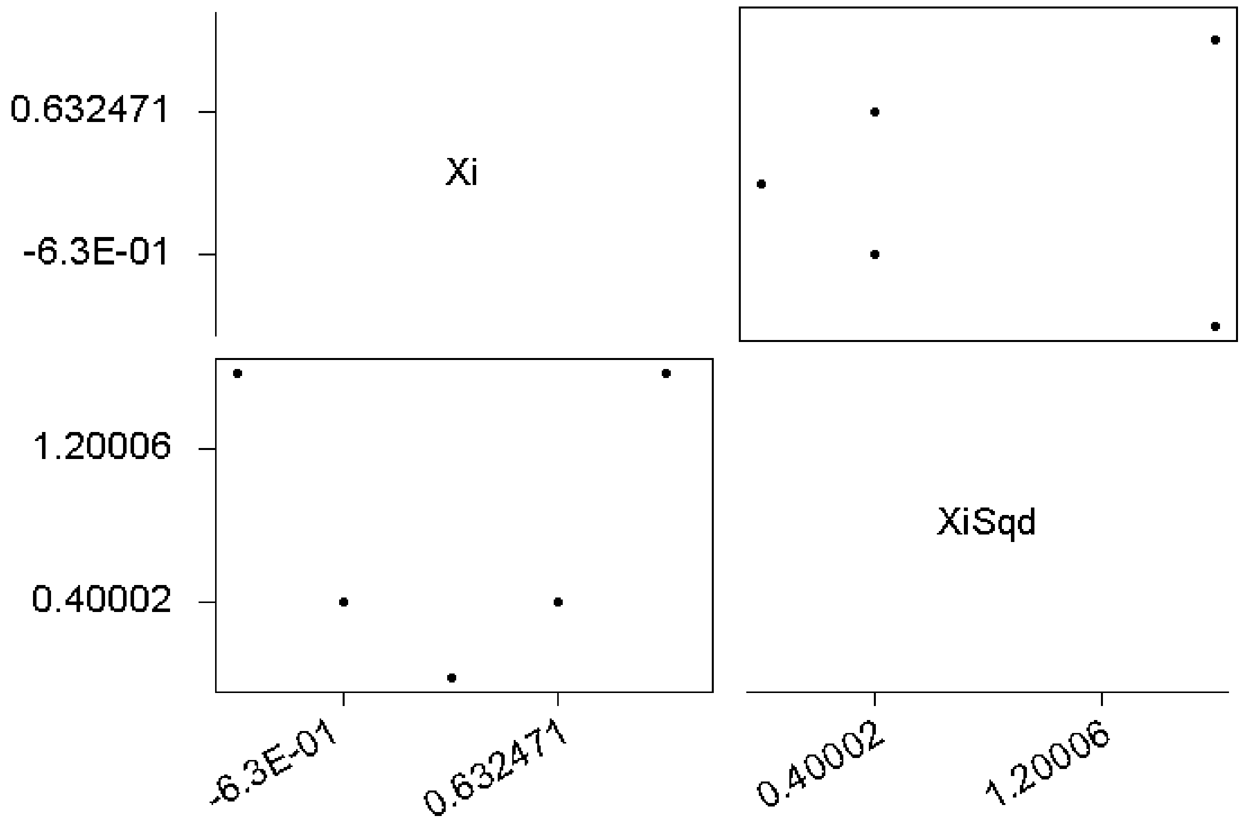
1-233. *(cont)* Multi-Collinearity & Data Transformations

Regression Analysis: Height (in) versus Shoe Size, Shoe Size Sqd

```
The regression equation is
Height (in) = 48.9 + 3.28 Shoe Size - 0.107 Shoe Size Sqd

Predictor        Coef     SE Coef        T        P      VIF
Constant       48.886       5.094     9.60    0.011
Shoe Siz        3.279       1.152     2.84    0.105    232.4
Shoe Siz     -0.10714     0.06389    -1.68    0.236    232.4

S = 0.2390      R-Sq = 99.4%     R-Sq(adj) = 98.8%

Analysis of Variance

Source            DF         SS         MS        F        P
Regression         2    18.3857     9.1929   160.88    0.006
Residual Error     2     0.1143     0.0571
Total              4    18.5000

Source       DF     Seq SS
Shoe Siz      1    18.2250
Shoe Siz      1     0.1607
```

For this study although the model may be sufficient for prediction, the effects of shoe size and (shoe size)^2 are obviously not easy to separate - the two predictors are certainly strongly related!

1-234. Reducing Multi-Collinearity Between Predictors



The X data has now been transformed to X_i and the squared term recalculated, where $X_i = X\text{-}X_{bar}/s$ (for the predictor column)

1-234. *(cont)* Reducing Multi-Collinearity Between Predictors

Regression Analysis: Height (in) versus Xi, XiSqd

The regression equation is
Height (in) = 69.7 + 2.13 Xi - 0.268 XiSqd

Predictor	Coef	SE Coef	T	P	VIF
Constant	69.7143	0.1666	418.46	0.000	
Xi	2.1345	0.1195	17.86	0.003	1.0
XiSqd	-0.2678	0.1597	-1.68	0.236	1.0

S = 0.2390 R-Sq = 99.4% R-Sq(adj) = 98.8%

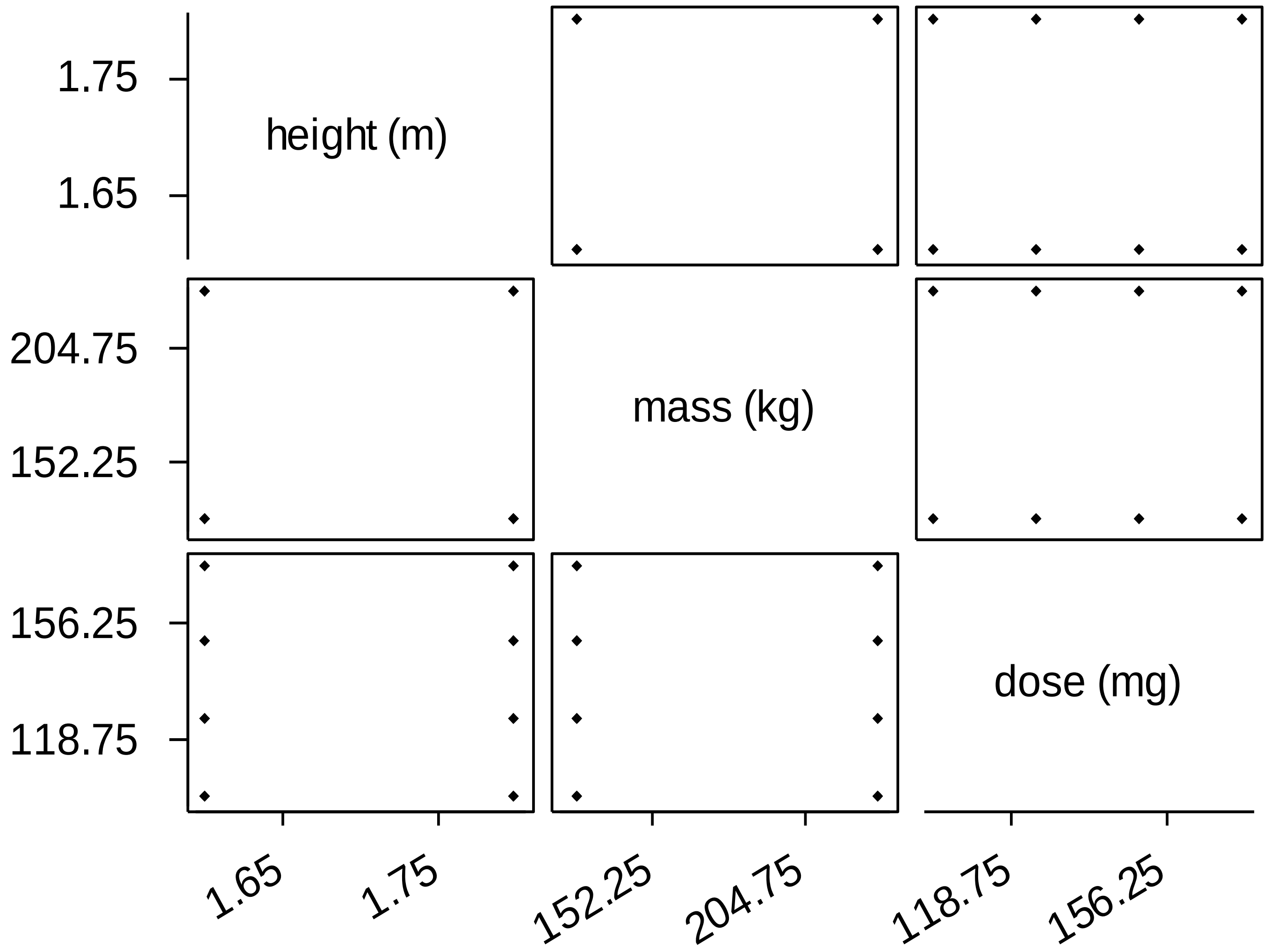
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	18.3857	9.1929	160.88	0.006
Residual Error	2	0.1143	0.0571		
Total	4	18.5000			

Source	DF	Seq SS
Xi	1	18.2250
XiSqd	1	0.1607

Note the VIFs have reduced to 1.0 (indicating no multi-collinearity) and the individual effects of the transformed variables can be more clearly understood

1-235. Orthogonal Study Design – Towards DOE



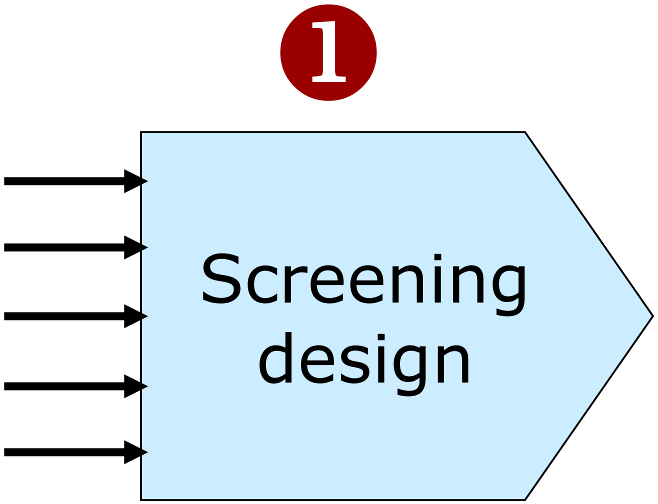
DOE for Practitioners

- *Multiple regression*
- **SCREENING DOE**
- *Modeling DOE*

Filtering out the less important

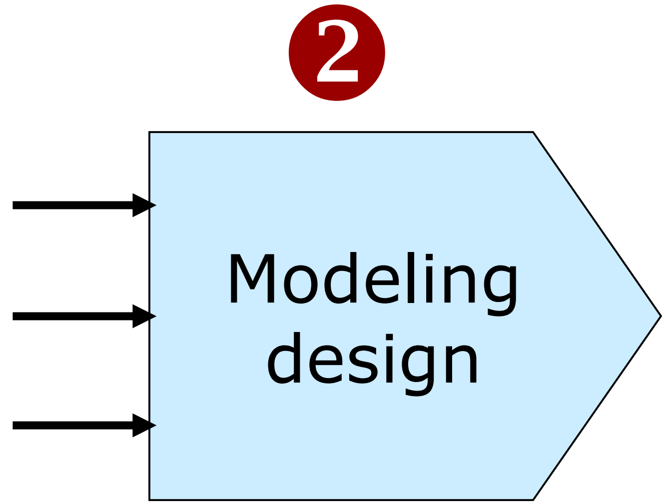
2-1. Recap DOE Methodology

◈ Two stages of design:



Design matrix

- Taguchi
- Fractional Factorial

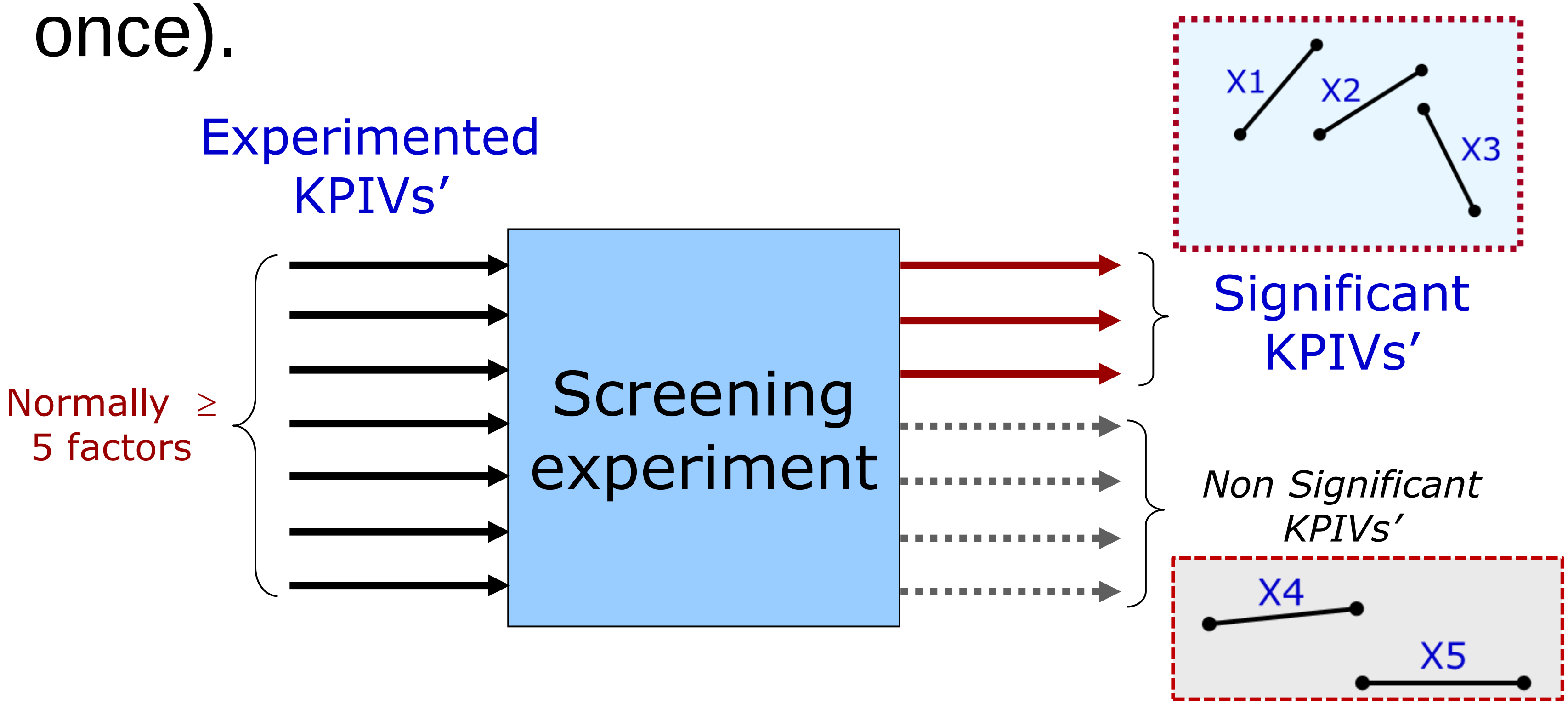


Design matrix

- Full Factorial
- CCD
- Box Behnken

2-2. Objective of Screening Experiments

- To screen out the non significant input variables (to select only the significance once).



2-21. Identifying The Key Players

- In screening experiment we are mainly investigating the main effects (interaction between factors is investigated in Modeling DOE).
  - What are main effects?
    - **The effect of an input factor is defined as the change in the output when the input factor changes.**

2-22. Strength of Main Effects – Example 2-1

Example 2-1:

Run	Creamer amount	Sugar amount	Y (response)
1	Level 1	Level 1	25
2	Level 1	Level 2	35
3	Level 2	Level 1	42
4	Level 2	Level 2	50

Y increases by 16 units

Creamer effect: When it moves from Lev 1 to Lev 2

$$Effect_{Creamer} = \frac{(42+50)-(25+35)}{2} = 16$$

Sugar effect: When it moves from Lev 1 to Lev 2

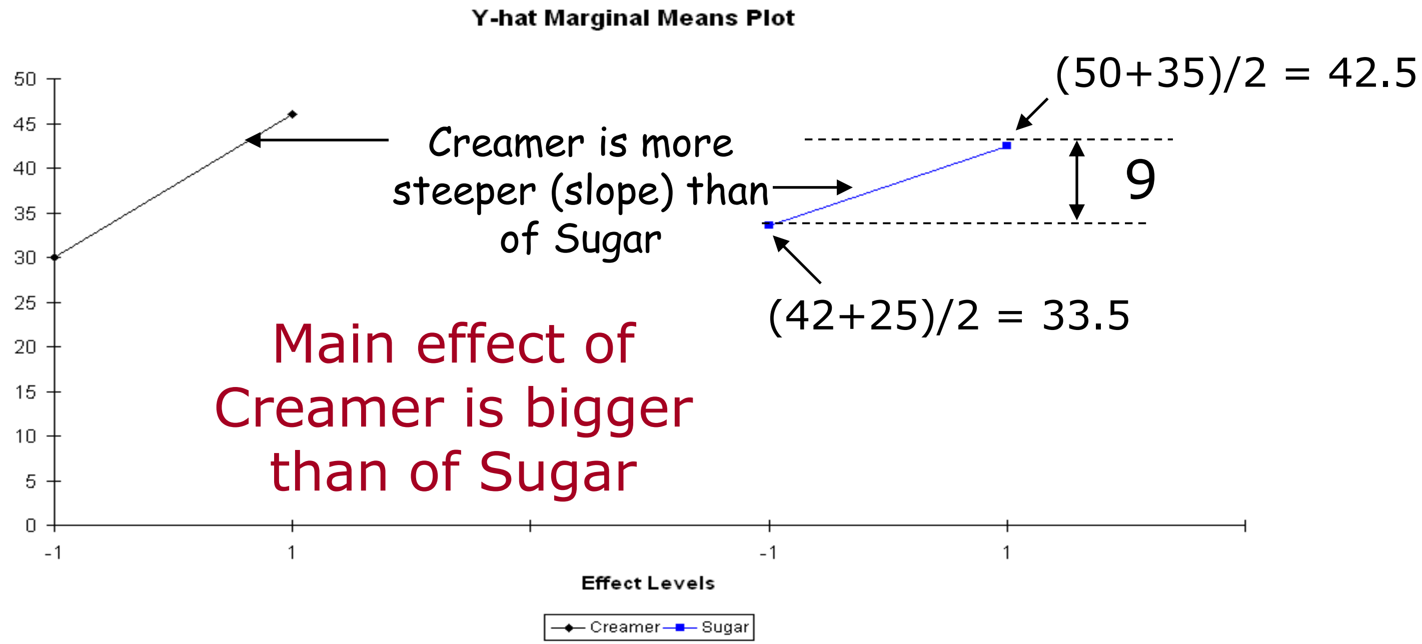
$$Effect_{Sugar} = \frac{(50+35)-(42+25)}{2} = 9$$

Y increases by 9 units

2-22. *(cont)* Strength of Main Effects

Example 2-1:

- The Effect of an input factor is defined as the change in the Output when the input factor changes.



In general, the effect magnitude of each factor is unique

2-31. Application of DOE – Recall

- Manufacturing – *A fixed level of solder materials in a pot was discovered. It makes other parameters to be robust.*
- Design – *A special PWB pattern was found to be the most stable in terms of catching consistent amount of solder.*
- Logistic – *A new shipping control had enhanced shortest delivery time from plant to customers.*

Either in manufacturing or services, DOE method can always be applied.

2-32. A Scenario

A mechanic is investigating ways to gain extra mileage in his driving.

2-33. Components of an Experiment (any experiments)

- Output variables (Responses)
  - **Is the response Continuous or Discrete?**
- Controllable input variables (Factors)
  - **Have the variables been selected correctly?**
- Noise (Background) variables
  - **Can the noise be controlled? (or accounted for if possible).**
  - **In most cases, the variability of these variables should initially be reduced.**

2-34. Planning an Experiment – Recall

1. Define the problem and the objective.
2. Define the Key Output Variables (quantified Y metrics, verified by GR&R gage).
3. Set target of Y metrics (centering mean, reduce variability).
4. Select the Key Input Variables (the fixed & experimental).
5. Choose the variable levels (Lo-Hi).
6. Plan how to control noise impact.
7. Select the Experimental Design.

2-341. Step 1

1 → 2 → 3 → 4 → 5 → 6 → 7

1 Define the problem and the objective.

Objective:-

To get the best oil consumption.

2-342. Step 2

2

Define the Output variables (quantified Y metrics, verified by GR&R gage).

3

Output:-

4

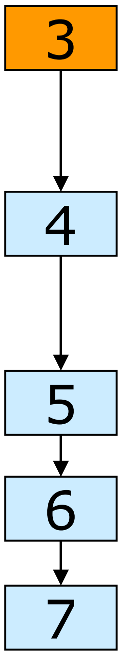
Amount of petrol consumed in 100 km driving.

5

6

7

2-343. Step 3



Set target of Y metrics (centering mean, reduce variability).

Target:-

Minimize Y (variability is not a priority).

2-344. Step 4

4 → 5 → 6 → 7

Select the Key Input Variables (the fixed & experimental).

(Study guidelines on selecting KPIV correctly)

2-3441. Factors (KPIVs') Selection

- Sources for KPIV selection:
  - Process Map
  - Cause and Effects Matrix
  - FMEA
  - Multi-vari study results
  - Brainstorming (for a simple processes)
  - Literature review
  - **Engineering knowledge**
  - Operator experience
  - Scientific theory
  - Customer/Supplier input

2-3442. Final KPIVs’

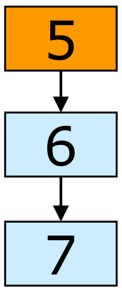
- From Engineering knowledge:

#	KPIV
1	Engine cc size
2	Spark plugs type
3	Petrol brand
4	Tire pressure
5	Driving speed
6	Temperature (filling petrol)
7	Engine oil level

KPIVs’ to be experimented

Other variables kept fixed

2-345. Step 5



Choose the variable levels (Lo-Hi).

#	KPIV	Lo	Hi
1	Engine cc size	**1.3**	**1.8**
2	Spark plugs type	**N**	**B**
3	Petrol brand	**Petron**	**Cell**
4	Tire pressure	**15**	**22**
5	Driving speed	**70**	**130**
6	Temperature (filling petrol)	**20^{o}C**	**25^{o}C**
7	Engine oil level	**Min**	**Max**

2-3451. Lo~Hi = Inference Space

◈ About **inference space**

- **Area within which you can draw your conclusions or explore new things**
- Two classifications:
  1. **Narrow inference**
  2. **Broad inference**

What happens to the taste between 1 ~ 2 tp of sugar?

2-34511. Narrow Inference

- Narrow Inference
  - Experiment focused on specific subset of overall operation.
  - Narrow inference studies normally are less affected by noise variables.
  - Examples: Only same operator, one shift, one machine, one building, one batch, etc..

2-34512. Broad Inference

- Broad Inference
  - Usually covers entire process (all machines, all shifts, all operators, etc..).
  - More data must be taken over a longer period of time (to let all changes to occur to check on robustness).
  - Broad inference studies are affected by Noise variables (most of the times purposely done).

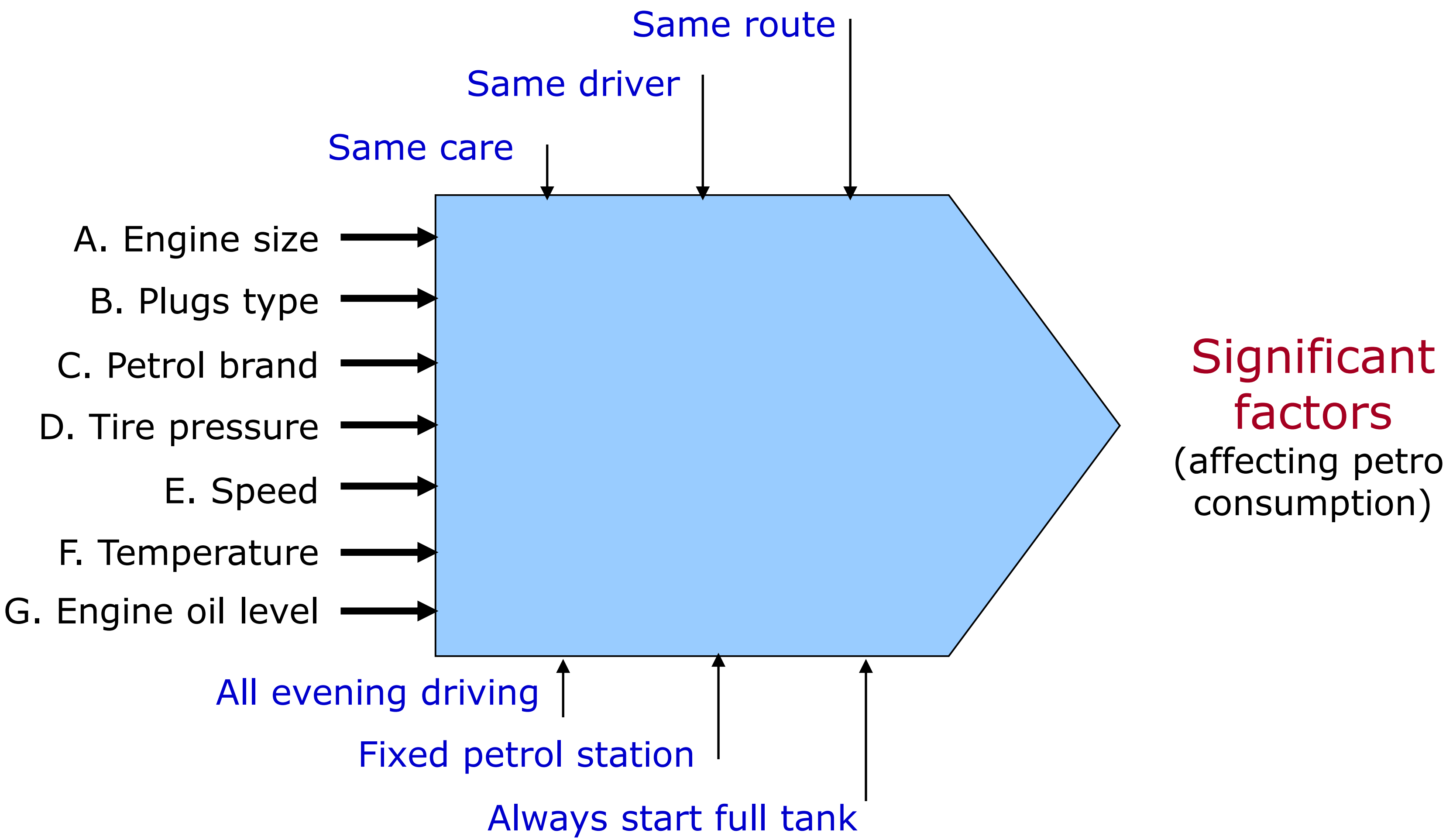
2-346. Step 6

6 → 7

Plan how to control noise impact

- Same route of driving
- Same driver
- Same car (for same cc – 2 cars altogether)
- All evening driving
- Fill petrol at same gas station (one station for each petrol brand)
- Start with full tank

2-3461. The IPO Diagram



2-347. Step 7

7 Select experimental design (design matrix)

Taguchi L-12 ➔

7 factors

12 runs

Factor	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Row #	A	B	C	D	E	F	G
1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	-1	1	1
3	-1	-1	1	1	1	-1	-1
4	-1	1	-1	1	1	-1	1
5	-1	1	1	-1	1	1	-1
6	-1	1	1	1	-1	1	1
7	1	-1	1	1	-1	-1	1
8	1	-1	1	-1	1	1	1
9	1	-1	-1	1	1	1	-1
10	1	1	1	-1	-1	-1	-1
11	1	1	-1	1	-1	1	-1
12	1	1	-1	-1	1	-1	1

2-35. Running an Experiment – Recall

1. Collect data.
2. Analyze data (ANOVA, regression, etc).
3. Draw statistical conclusions.
4. Replicate results (confirmation runs = repeat experiments with selective setting).
5. Draw practical solutions (list down new knowledge gained).
6. Implement solutions.

2-351. Step 1

1 → 2 → 3 → 4 → 5 → 6

Collect data.

Design matrix

3 Replication or Repetition

Factor	A	B	C	D	E	F	G							
Row #	A	B	C	D	E	F	G		Y1	Y2	Y3		Y bar	S
1	-1	-1	-1	-1	-1	-1	-1						#DIV/0!	#DIV/0!
2	-1	-1	-1	-1	-1	1	1						#DIV/0!	#DIV/0!
3	-1	-1	1	1	1	-1	-1						#DIV/0!	#DIV/0!
4	-1	1	-1	1	1	-1	1						#DIV/0!	#DIV/0!
5	-1	1	1	-1	1	1	-1						#DIV/0!	#DIV/0!
6	-1	1	1	1	-1	1	1						#DIV/0!	#DIV/0!
7	1	-1	1	1	-1	-1	1						#DIV/0!	#DIV/0!
8	1	-1	1	-1	1	1	1						#DIV/0!	#DIV/0!
9	1	-1	-1	1	1	1	-1						#DIV/0!	#DIV/0!
10	1	1	1	-1	-1	-1	-1						#DIV/0!	#DIV/0!
11	1	1	-1	1	-1	1	-1						#DIV/0!	#DIV/0!
12	1	1	-1	-1	1	-1	1						#DIV/0!	#DIV/0!

Random Order = 5 ,12 ,9 ,8 ,3 ,4 ,6 ,7 ,11 ,1 ,2 ,10

2-351. (cont) Step 1

1
2
3
4
5
6

Collect data.

Design matrix

Factor	A	B	C	D	E	F	G							
Row #	A	B	C	D	E	F	G		Y1	Y2	Y3		Y bar	S
1	-1	-1	-1	-1	-1	-1	-1		8.11	7.86	8.01		7.993333	0.125831
2	-1	-1	-1	-1	-1	1	1		8.22	8.15	7.88		8.083333	0.179536
3	-1	-1	1	1	1	-1	-1		7.711	7.52	7.56		7.597	0.100732
4	-1	1	-1	1	1	-1	1		7.98	7.88	7.89		7.916667	0.055076
5	-1	1	1	-1	1	1	-1		6.95	7.62	7.81		7.46	0.451774
6	-1	1	1	1	-1	1	1		7.25	7.53	7.55		7.443333	0.16773
7	1	-1	1	1	-1	-1	1		8.23	8.15	8.23		8.203333	0.046188
8	1	-1	1	-1	1	1	1		8.26	8.19	8.14		8.196667	0.060277
9	1	-1	-1	1	1	1	-1		8.23	8.5	8.77		8.5	0.27
10	1	1	1	-1	-1	-1	-1		8.23	8.45	8.4		8.36	0.115326
11	1	1	-1	1	-1	1	-1		8.15	8.21	8.18		8.18	0.03
12	1	1	-1	-1	1	-1	1		8.18	8.16	8.21		8.183333	0.025166

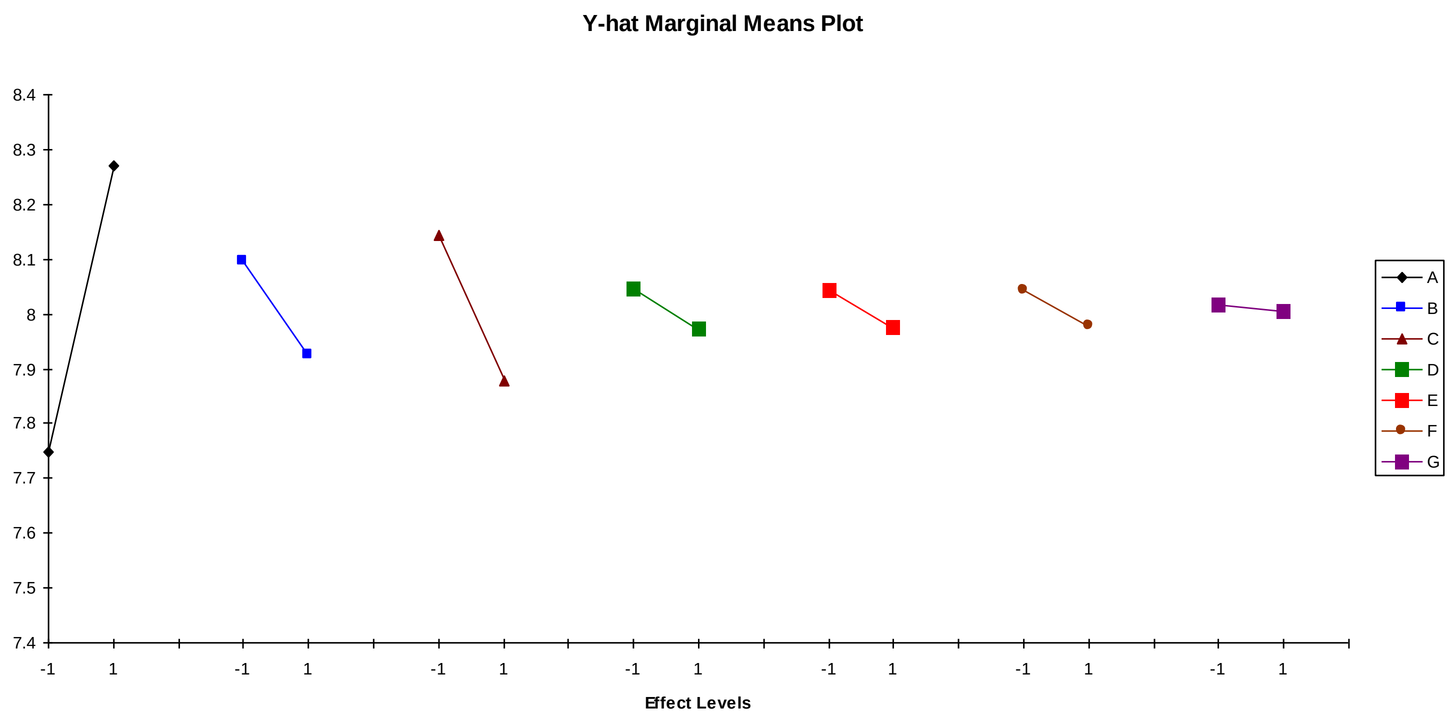
Random Order = 5 ,12 ,9 ,8 ,3 ,4 ,6 ,7 ,11 ,1 ,2 ,10

2-352. Step 2

2 → 3 → 4 → 5 → 6

Analyze data (ANOVA, regression, etc)

Marginal Means plot



2-352. *(cont)* Step 2

2 → 3 → 4 → 5 → 6

Analyze data (ANOVA, regression, etc)

Regression Analysis

Multiple Regression Analysis

Y-hat Model

Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		8.00975	0.0000		
A	A	0.26081	0.0000	1	X
B	B	-0.08586	0.0191	1	X
C	C	-0.13303	0.0006	1	X
D	D	-0.03636	0.3012	1	X
E	E	-0.03414	0.3311	1	X
F	F	-0.03253	0.3541	1	X
G	G	-0.00531	0.8789	1	X

Rsq	0.7434
Adj Rsq	0.6793
Std Error	0.2071
F	11.5891
Sig F	0.0000

Source	SS	df	MS
Regression	3.5	7	0.5
Error	1.2	28	0.0
Total	4.7	35	

Factor	Name	Low	High	Exper
A	A	-1	1	0
B	B	-1	1	0
C	C	-1	1	0
D	D	-1	1	0
E	E	-1	1	0
F	F	-1	1	0
G	G	-1	1	0

Prediction

Y-hat	8.00975
S-hat	0.1356363

99% Prediction Interval

Lower Bound	7.602841
Upper Bound	8.416659

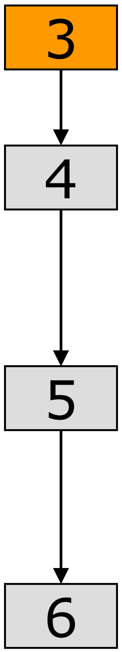
S-hat Model

Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		0.13564	0.0167		
A	A	-0.04448	0.2643	1	X
B	B	0.00521	0.8866	1	X
C	C	0.02137	0.5669	1	X
D	D	-0.02402	0.5223	1	X
E	E	0.02487	0.5085	1	X
F	F	0.05758	0.1684	1	X
G	G	-0.04664	0.2454	1	X

Rsq	0.6605
Adj Rsq	0.0663
Std Error	0.1188
F	1.1116
Sig F	0.4876

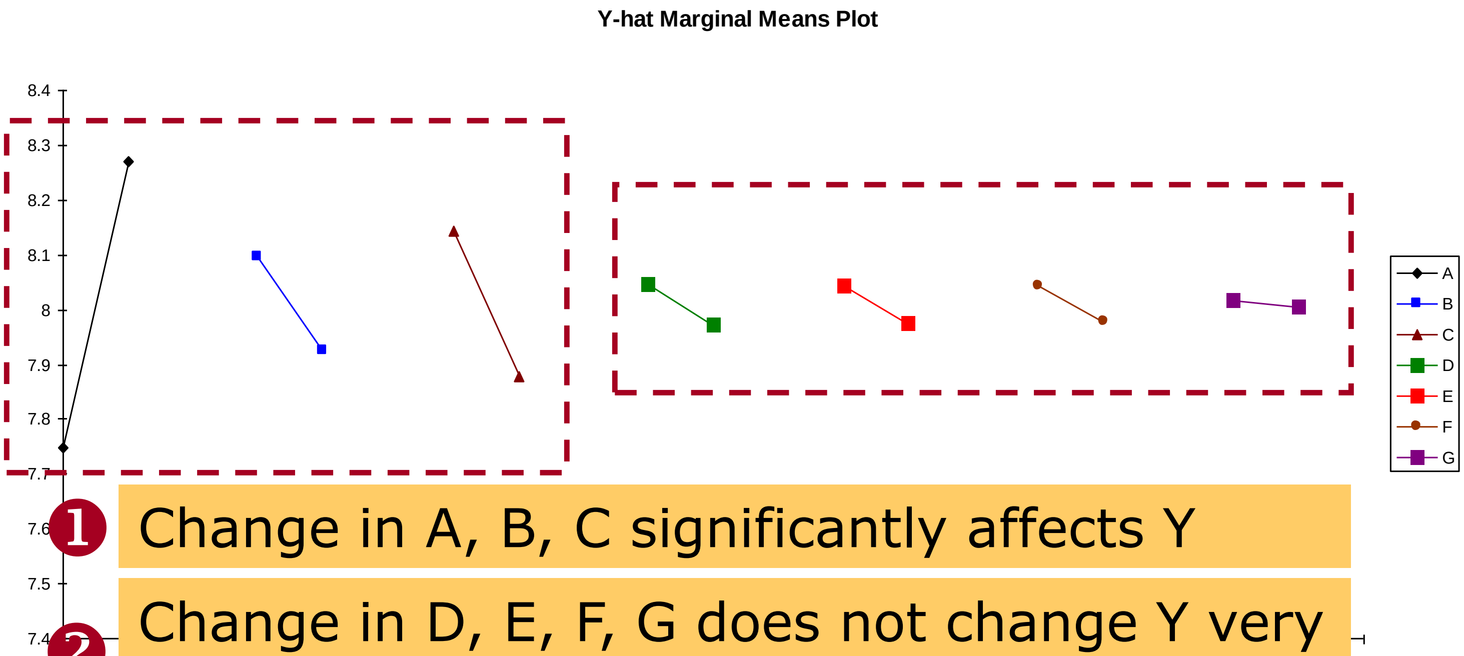
Source	SS	df	MS
Regression	0.1	7	0.0
Error	0.1	4	0.0
Total	0.2	11	

2-353. Step 3



Draw statistical conclusions

Marginal Means Plot



❶ Change in A, B, C significantly affects Y

❷ Change in D, E, F, G does not change Y very much (insignificant factors)

2-353. *(cont)* Step 3

3 Draw statistical conclusions

4

5

6

Regression Analysis

Y-hat Model

Factor	Name	Coeff	P(2 Tail)	Tol	ctive
Const		8.00975	**0.0000**		
A	A	0.26081	**0.0000**	1	
B	B	-0.08586	**0.0191**	1	
C	C	-0.13303	**0.0006**	1	
D	D	-0.03636	0.3012	1	
E	E	-0.03414	0.3311	1	X
F	F	-0.03253	0.3541	1	X
G	G	-0.00531	0.8789	1	X

Rsq	0.7434
Adj Rsq	0.6793
Std Error	0.2071
F	11.5891
Sig F	0.0000

Source	SS	df	MS
Regression	3.5	7	0.5
Error	1.2	28	0.0
Total	4.7	35	

❷ P value of < 0.05 signifies the significance of a factor

❶ R squared value of > 0.7 shows the experiment is a reliable one (predictable) – recall Coefficient of Determination (r^2)

2-353. *(cont)* Step 3

3 Draw statistical conclusions

4

5

6

Regression Analysis

- Factors A, B, C are significant factors – Next stage of DOE (Modeling DOE) should consider these factors.
- Factor A (engine size) is the most significant, and proportional to Y response (+ve direction).
- Factor C (petrol brand) and factor B (plugs type) are also significant, and they are inversely proportional to the Y response (–ve direction).

2-354. Step 4 ~ 6

4 Replicate results (confirmation runs = repeat experiments with selected setting).

5 Draw practical solutions (list down new knowledge gained).

6 Implement solutions.

These steps are recommended for Modeling DOE (next module)

2-36. Next Steps After Screening Experiments

- Run Modeling DOE
  - Procedures:-
    - **Planning the experiment**
    - **Run the experiment**
    - **Implement the solutions**

Similar as the procedures in screening experiment

2-4. Summary

- Screening Experiments is performed when we want to single out important variables out of many variables.
- It mainly looks for main effects (not interactions).
- If there are not so many variables in an experiment, one can go straight away to Modeling DOE (and look for interactions).

2-4. *(cont)* Summary

- What is screening experiments?

- What is it for?

- How to start screening experiment?

2-4. *(cont)* Summary

- How to use screening experiment to monitor the significant factors?
- What are the precautions when running screening experiment?
- What are the differences between screening experiments and modeling experiments?

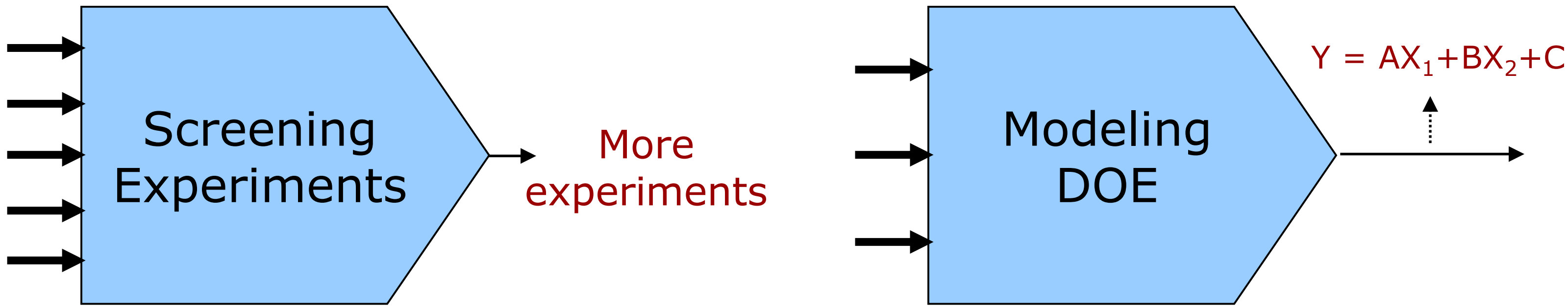
Discussion 2-1: Outcome of Screening Experiments

◈ *What is expected out of screening experiments?*

Screening Experiments

- Significant factors identified (with large main effect). **These factors are candidates for future experiments.**
- Non significant factors are also identified. **These factors should be dropped in future experiments, and should be set to economical setting as the impact to the output is almost negligible.**

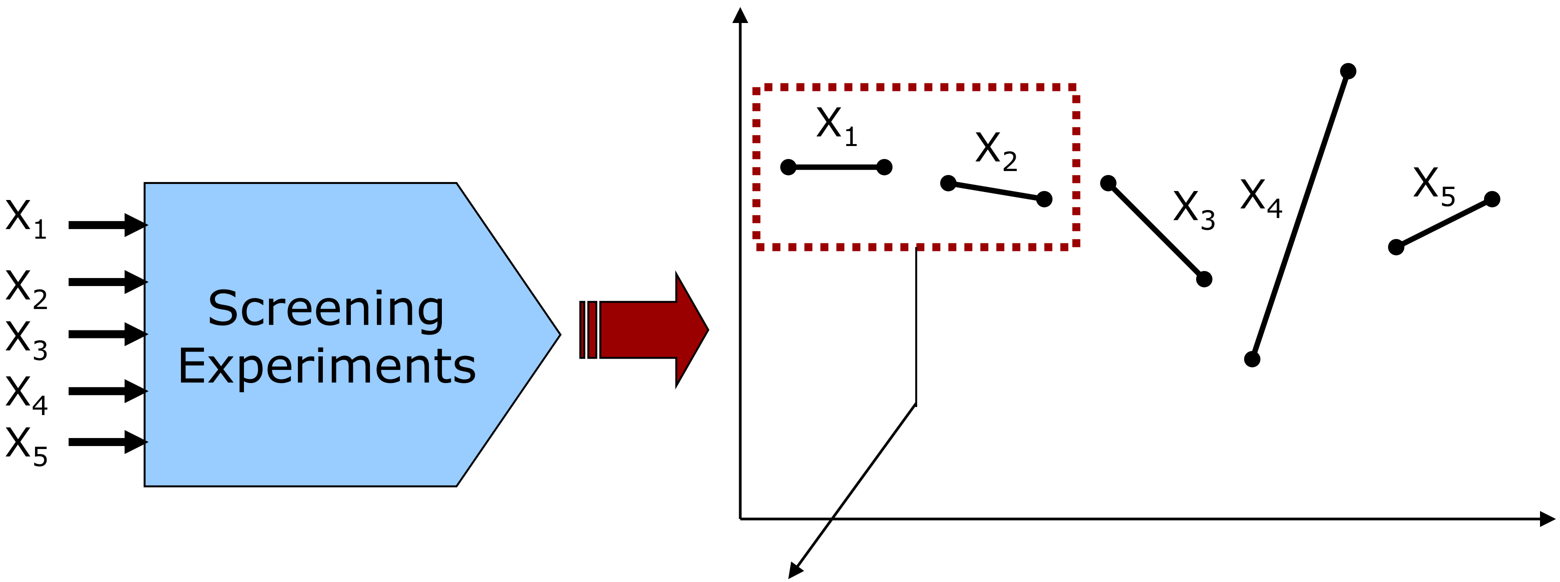
Discussion 2-2: Screening Experiments vs. Modeling DOE



- Many input factors (normally ≥ 5 variables)
- Initial experiments
- 2 level experiments (-1, +1)
- Focuses on main effects

- A few factors (normally < 5 variables)
- Done after Screening Experiments
- 3 levels experiment (-1, 0, +1)
- Focuses on interaction between factors, model expression, optimization

Discussion 2-3: Meaning of Small Main Effects



Small main effects means a change in the input variable (–ve ➔ +ve) make very little change in the Y response. The factors can be ignored.

DOE for Practitioners

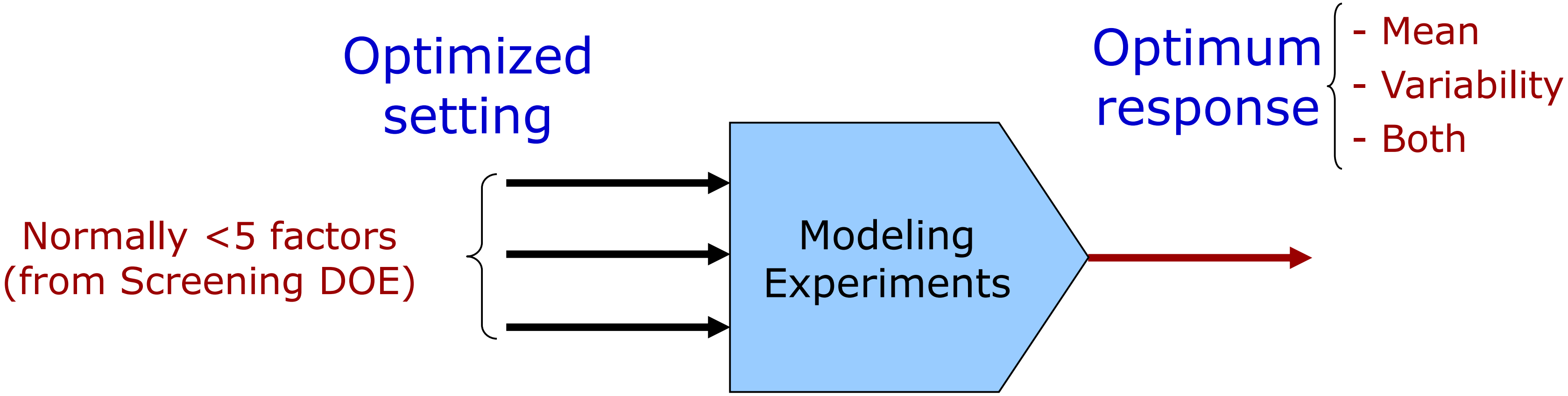
- *Screening DOE*
- **MODELING DOE**
- *Overviewing DOE Mixtures*

Formulate anti aging drink …

3-11. Process Optimization

Process optimization

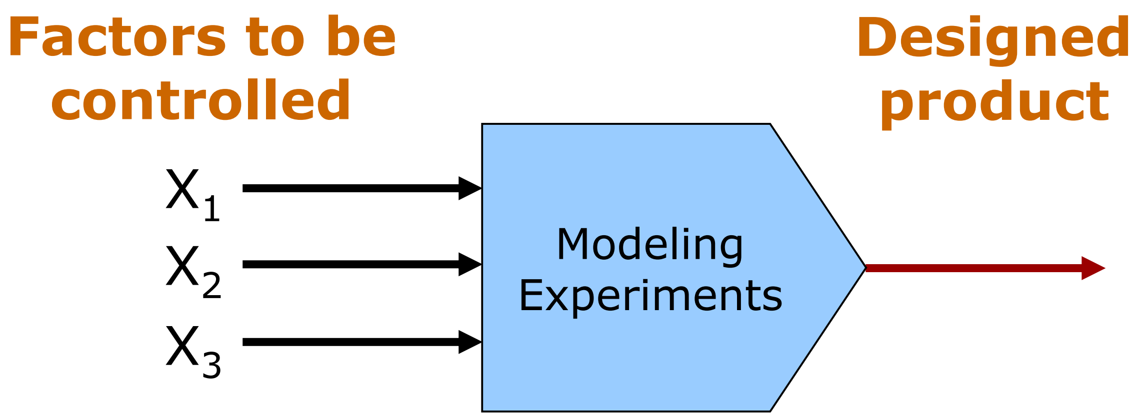
- Determining setting of the critical inputs (to get optimum response).
- Determining “real” specification limits (by knowing the mathematical model or expression).



3-12. Product Design

◈Product design

- Aids in understanding X's early in the design process.
- Provides direction for "robust" designs.



3-13. Interaction Effects

- **In modeling experiment we are investigating both the main effects and <u>interaction</u> between factors.**
  - What is an interaction?
    - **The change in response does not solely depend on the change of one input factor, but it is also depends on the change of other inputs.**

Example 3-1: Interaction Effects

Run	Creamer amount	Sugar amount	Creamer.Sugar	Y (response)
1	- 1	- 1	+ 1	25 ⇦
2	- 1	+ 1	- 1	35 ⇦
3	+ 1	- 1	- 1	42 ⇦
4	+ 1	+ 1	+ 1	50 ⇦

When Creamer.Sugar at −1,
the effect = (35+42)/2

When Creamer.Sugar at +1,
the effect = (25+50)/2

$$Effect_{CreamerSugar} = \frac{(25+50)-(35+42)}{2} = -1 \longleftarrow$$

Y response for Creamer.Sugar

(cont) Example 3-1: Interaction Effects

- Another way to calculate interaction effect:

Run	Creamer amount	Sugar amount	Y (response)
1	- 1 ⇦	- 1	25
2	- 1 ⇦	+ 1	35
3	+ 1 ⇦	- 1	42
4	+ 1 ⇦	+ 1	50

When Creamer at −1,
Sugar effect = (35-25)/2

When Creamer at +1,
Sugar effect = (50−42)/2

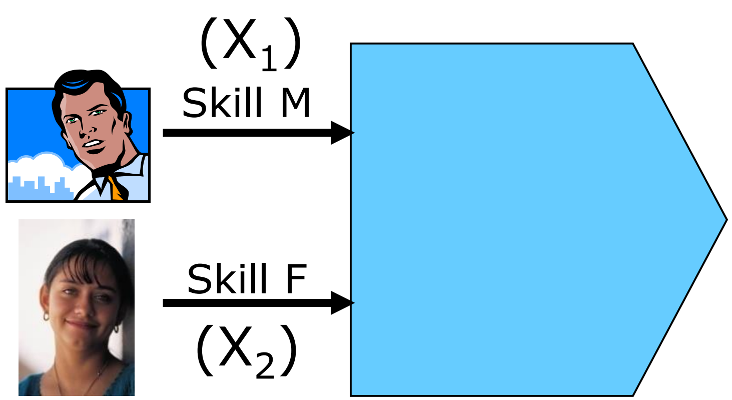
$$Effect_{Creamer.Sugar} = \frac{(50-42)-(35-25)}{2} = -1$$

Y response for Creamer.Sugar

Please try calculating Creamer effect when Sugar moves from −1 to +1. Do you get the same answer?

3-131. Definition of Interaction

- **Interaction means how the input factors work together to affect the output or response** (the input factors can affect the response differently at their different levels).



$$\hat{Y} = \overbrace{AX_1 + BX_2}^{\text{Main effects}} + \overbrace{CX_1X_2}^{\text{Interaction}} + K$$

C magnitude shows how well the two input factors work together (whether they like or dislike each other)

3-2111. Design Matrix

2 level 2 factor FF

$2^2 = 4$

Run	X_1	X_2
1	-	-
2	-	+
3	+	-
4	+	+

2 level 3 factor FF

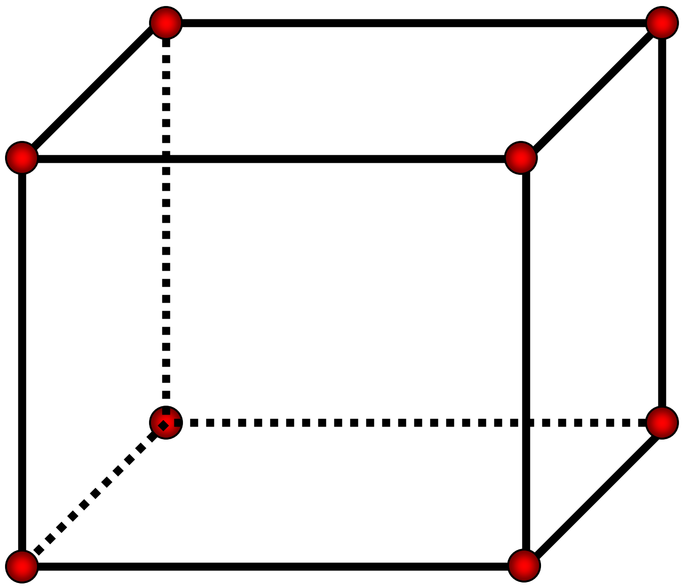
$2^3 = 8$

Run	X1	X2	X3
1	-	-	-
2	-	-	+
3	-	+	-
4	-	+	+
5	+	-	-
6	+	-	+
7	+	+	-
8	+	+	+

Number of runs (2 levels):

=> 2^n (n: # of factors)

◈ In Full Factorial each level of every factor in the experiment is investigated ***(all corners in a box)***



3-2113. Purposes of Using Full Factorial

- To understand the advantages of factorial experiments.
- To determine how to analyze general factorial experiments.
- To understand the concept of statistical interaction.
- To analyze two and three factor experiments.
- To use diagnostic techniques to evaluate the "goodness of fit" of the statistical model.
- To identify the most important or critical factors in the experiments.

3-2114. The Advantages

- More efficient than One-Factor-at-a-Time (OFAT) experiments.
- Allows the investigation of the combined effects of factors (Interactions).
- Covers a wider experimental region than OFAT studies.
- Identify critical Factors (Inputs).
- More efficient in estimating effects of both input and noise variables on the output.

3-2115. FF vs. OFAT – Example

Example 3-2: 2 levels with 2 input factors

Decided after run 1 ~ 2

OFAT

Run	Sugar	Creamer
1	Amount 1	Amount X
2	Amount 2	Amount X
3	Opt Amount	Amount 1
4	Opt Amount	Amount 2

- 4 runs
- 1 replication for each factor level
- No information about interaction

Full Factorial

Run	Sugar	Creamer
1	Amount 1	Amount 1
2	Amount 1	Amount 2
3	Amount 2	Amount 1
4	Amount 2	Amount 2

- 4 runs
- 2 replications for each factor level
- Provides some information about interaction

3-21151. What OFAT Misses – Example

(cont) <u>Example 3-2: 2 levels with 2 input factors</u>

Decided after run 1 ~ 2

OFAT

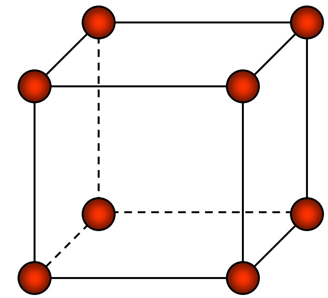
Run	Sugar	Creamer
1	Amount 1	Amount X
2	Amount 2	Amount X
3	Opt Amount	Amount 1
4	Opt Amount	Amount 2

- Hold Creamer at fixed amount while testing Sugar at Amount 1 and Amount 2.
- Found the best amount of Sugar => set as optimum amount.
- Hold Sugar at optimum amount; test Creamer at each amount
- Found the best amount of Creamer at the optimum amount of Sugar

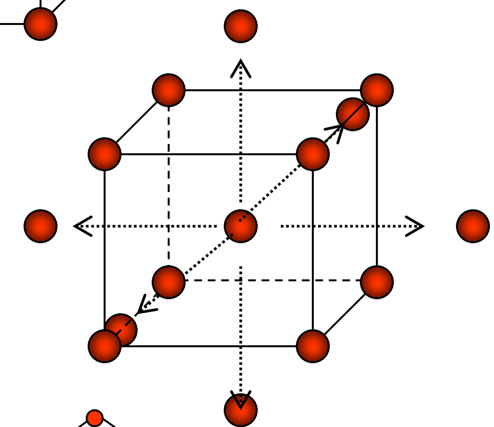
→ *Local optimum (OFAT misses the model optimum)*

3-2121. 2 Levels

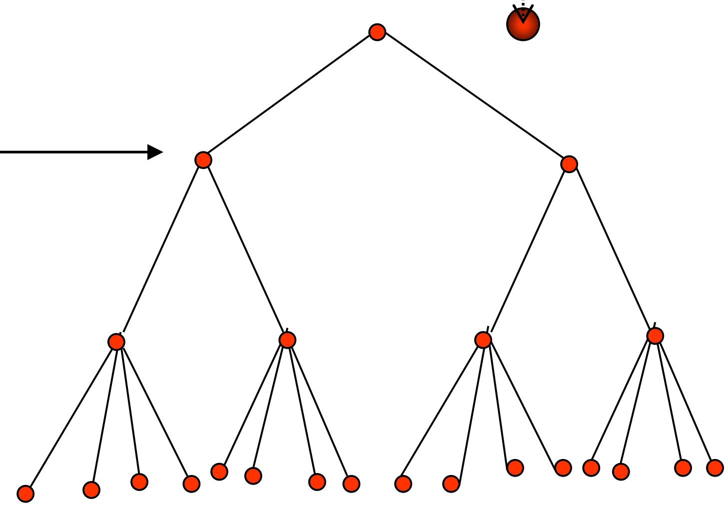
- Full Factorial
- Fractional Factorial



- Central Composite
- Taguchi L_8, Taguchi L_{12}
- 12 run Hadamard / Plackett-Burman
- Fold-over Koshal

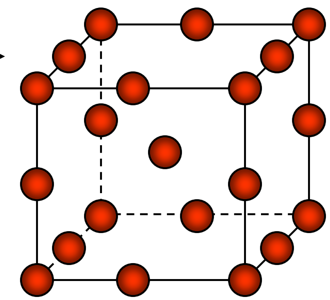


- Nested
- Mixture

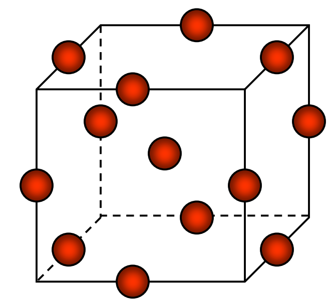


3-2122. 3 Levels

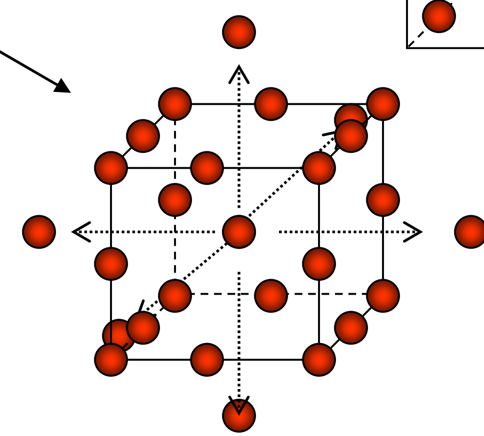
- Full Factorial



- Taguchi L_{18} (screening)
- Box Behnken



- Central Composite (CCD)



- D-Optimal

3-2123. 2 Levels vs. 3 Levels

<u>*2 level – 2 factors*</u>

#	A	B
1	-1	-1
2	-1	+1
3	+1	-1
4	+1	+1

Experiment at −1 and +1 levels

<u>*3 level – 2 factors*</u>

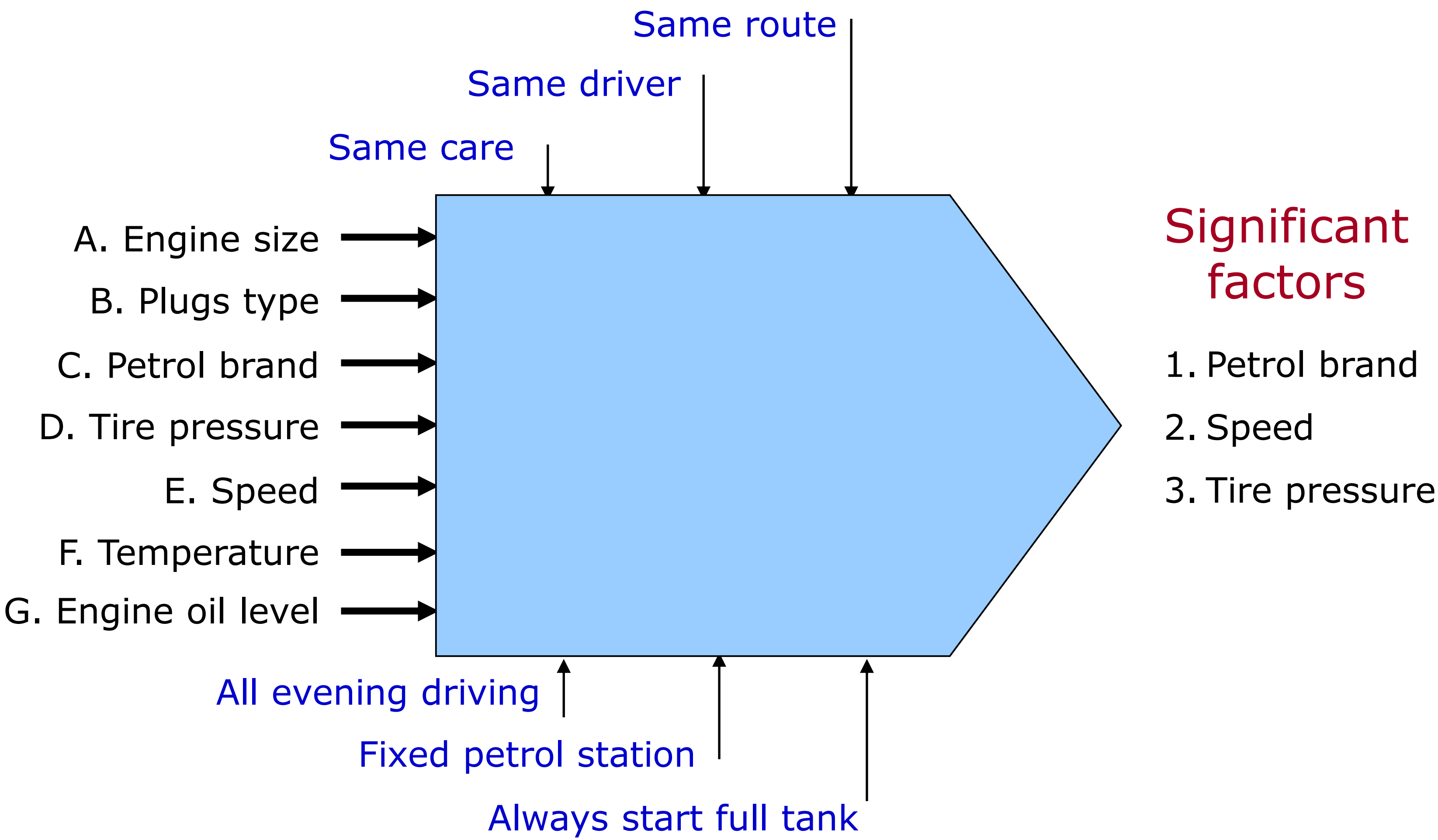
#	A	B
1	-1	-1
2	-1	0
3	-1	+1
4	0	-1
5	0	0
6	0	+1
7	+1	-1
8	+1	0
9	+1	+1

Experiment at −1, 0, +1 levels

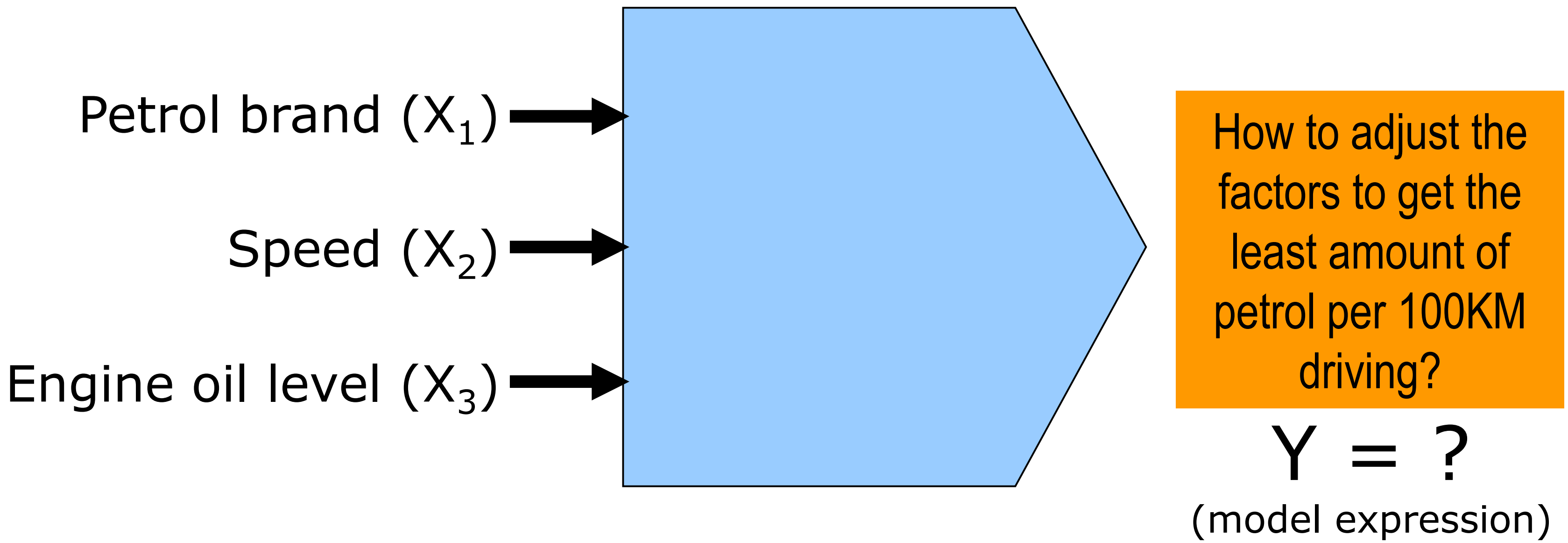
3-221. The Mechanic Case

A mechanic is investigating ways to gain extra mileage in his driving.

3-222. Finding in Screening DOE



3-222. *(cont)* Finding in Screening DOE

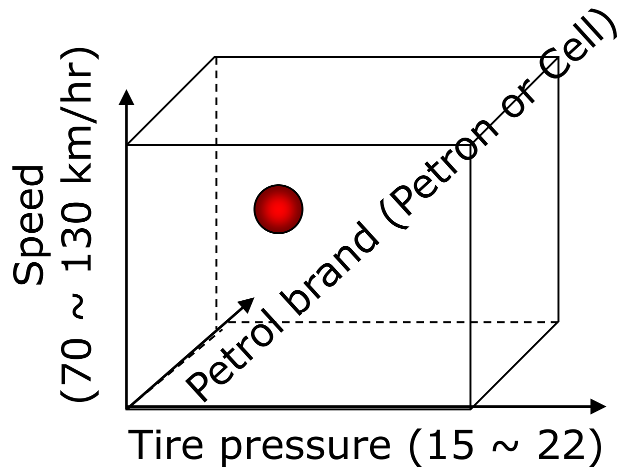


3-231. The Steps – Recall

1. Define the problem and the objective.
2. Define the Key Output Variables (quantified Y metrics, verified by GR&R gage).
3. Set target of Y metrics (centering mean, reduce variability).
4. Select the Key Input Variables (the fixed & experimental).
5. Choose the variable levels (Lo-Hi).
6. Plan how to control noise impact.
7. Select the Experimental Design.

Step 5: Lo-Hi Level Setting

5 Choose the variable levels (Lo-Hi).



The optimum point might be inside the experimented ranges – for modeling DOE (refining) it is advisable not to test outside previous ranges.

6 Plan how to control noise impact.

7 Select the Experimental Design.

3-233. Step 7: Design Selection

✔ 6 Plan how to control noise impact. (as in earlier experiment)

➡ 7 Select the Experimental Design.

- 2 levels
- 1 categorical factor (petrol brand)
- 2 continuous factors (speed, tire pressure)

Taguchi design allows a mix of 3 levels design and 2 level design (ie. categorical factor)

⇩

Full Factorial (wt one categorical factor)

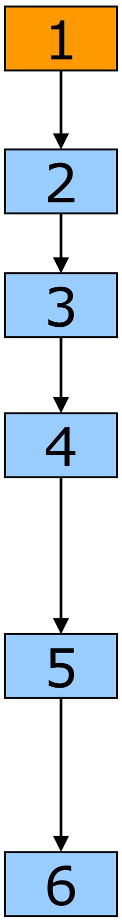
3-233. *(cont)* Step 7: Design Selection

Factor	A	B	C							
Row #	Petrol brand	Car speed	Tire pressure		Y1	Y2	Y3		Y bar	S
1	-1	70	150						#DIV/0!	#DIV/0!
2	-1	70	220						#DIV/0!	#DIV/0!
3	-1	130	150						#DIV/0!	#DIV/0!
4	-1	130	220						#DIV/0!	#DIV/0!
5	1	70	150						#DIV/0!	#DIV/0!
6	1	70	220						#DIV/0!	#DIV/0!
7	1	130	150						#DIV/0!	#DIV/0!
8	1	130	220						#DIV/0!	#DIV/0!
Random Order = 8 ,7 ,2 ,3 ,5 ,6 ,1 ,4										

3-241. The Steps – Recall

1. Collect data.
2. Analyze data (ANOVA, regression, etc)
3. Draw statistical conclusions.
4. Replicate results (confirmation runs = repeat experiments with selective setting).
5. Draw practical solutions (list down new knowledge gained).
6. Implement solutions.

3-242. Step 1: Run the Experiment



Collect data.

Factor	A	B	C		Y1	Y2	Y3		Y bar	S
Row #	Petrol brand	Car speed	Tire pressure		Y1	Y2	Y3		Y bar	S
1	-1	70	150		10.8	10.96	10.3		10.68667	0.344287
2	-1	70	220		11.8	11.01	11.77		11.52667	0.447698
3	-1	130	150		10.74	11	10.68		10.80667	0.170098
4	-1	130	220		11.45	11.4	11.75		11.53333	0.189297
5	1	70	150		11.1	11.5	11		11.2	0.264575
6	1	70	220		12.05	12.22	12.3		12.19	0.127671
7	1	130	150		12.6	11.97	12.45		12.34	0.32909
8	1	130	220		12.76	12.75	12.56		12.69	0.112694

Random Order = 8 ,7 ,2 ,3 ,5 ,6 ,1 ,4

3-243. Step 2-1: Regression Analysis

2 Analyze data (ANOVA, regression, etc)

3

4

5

6

Multiple Regression Analysis

Y-hat Model

Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		9.35083	0.0000		
A	Petrol brand	-0.05333	0.0468	1	X
B	Car speed	0.95667	0.0000	1	X
C	Tire pressure	0.27417	0.0000	1	X
AB		-0.00417	0.8685	1	X
AC		0.45500	0.0000	1	X
BC		0.29000	0.0000	1	X
ABC		-0.47583	0.0000	1	X

Rsq	0.9935
Adj Rsq	0.9907
Std Error	0.1213
F	352.0697
Sig F	0.0000

Source	SS	df	MS
Regression	36.3	7	5.2
Error	0.2	16	0.0
Total	36.5	23	

Factor	Name	Low	High	Exper
A	Petrol brand	-1	1	0
B	Car speed	70	130	100
C	Tire pressure	150	220	185

Prediction

Y-hat	9.35083333
S-hat	0.10390159

99% Prediction Interval

Lower Bound	9.03912856
Upper Bound	9.6625381

Fac					
Co					
A	Petrol brand	-0.02930	Not Avail	1	X
B					
C					
AB					
AC					
BC					
ABC					

Rsq	1.0000
Adj Rsq	Not Avail
Std Error	Not Avail
F	Not Avail
Sig F	Not Avail

Source	SS	df	MS
Regression	0.0	7	0.0
Error	0.0	0	Not Avail
Total	0.0	7	

Confirm (for the model to be acceptable):

1. R-sqrd value > 0.7
2. Adj R-sqrd ≈ R-sqrd

Remove non significant factors by deleting active [X] mark (for p-value > 0.1)

3-243. *(cont)* Step 2-1: Regression Analysis

2 Analyze data (ANOVA, regression, etc)

3

4

5

6

Repeat regression analysis after removing non significant factors

Multiple Regression Analysis

Y-hat Model

Factor	Name	Coeff	P(2 Tail)	Tol	Active
Const		9.35083	0.0000		
A	Petrol brand	-0.05333	0.0404	1	X
B	Car speed	0.95667	0.0000	1	X
C	Tire pressure	0.27417	0.0000	1	X
AC		0.45500	0.0000	1	X
BC		0.29000	0.0000	1	X
ABC		-0.47583	0.0000	1	X

Rsq	0.9935
Adj Rsq	0.9913
Std Error	0.1178
F	435.6437
Sig F	0.0000

Source	SS	df	MS
Regression	36.3	6	6.0
Error	0.2	17	0.0
Total	36.5	23	

Factor	Name	Low	High	Exper
A	Petrol brand	-1	1	0
B	Car speed	70	130	100
C	Tire pressure	150	220	185

Prediction

Y-hat	9.35083333
Std Error	0.11777762

99% Prediction Interval

Lower Bound	8.99750046
Upper Bound	9.7041662

3-244. Step 2-2: Interaction Study

2 Analyze data (interaction analysis)

3

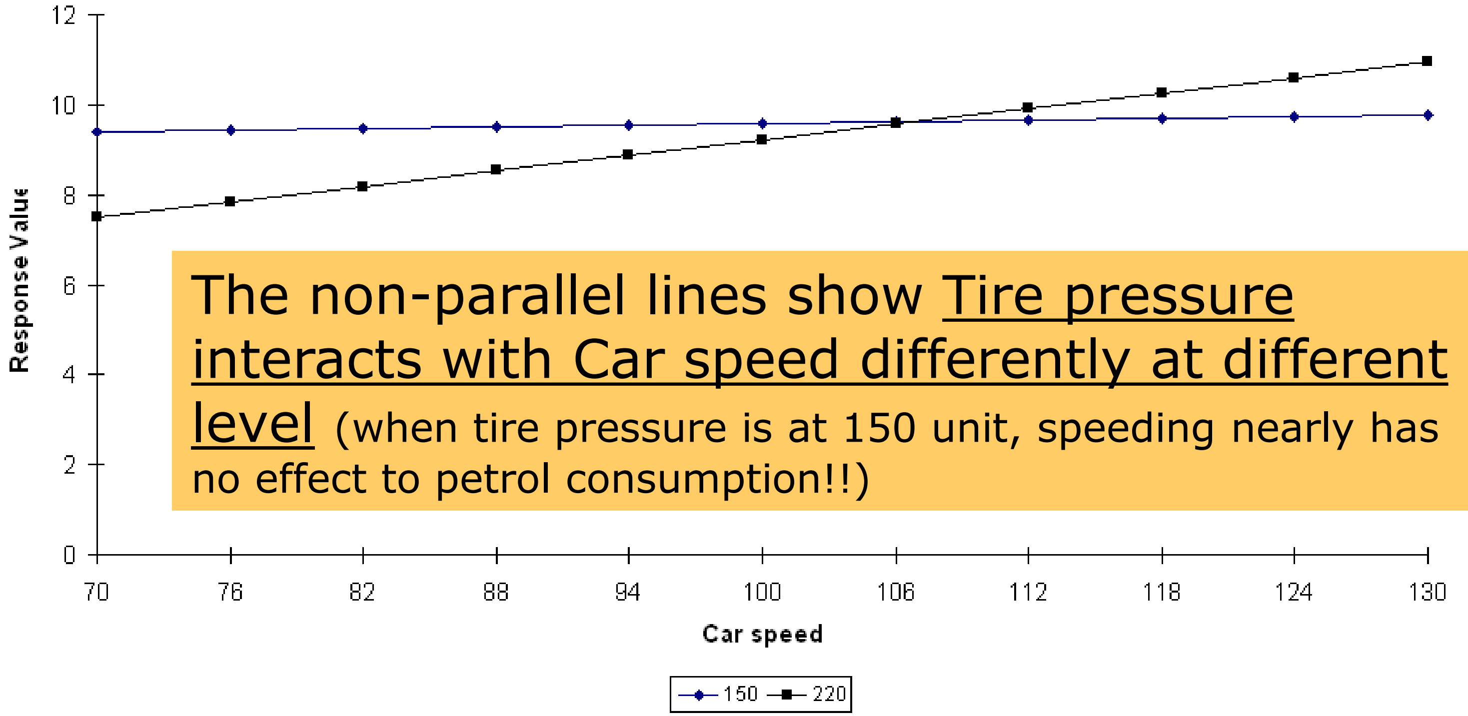
4

5

6

Interaction Plot of Car speed vs Tire pressure

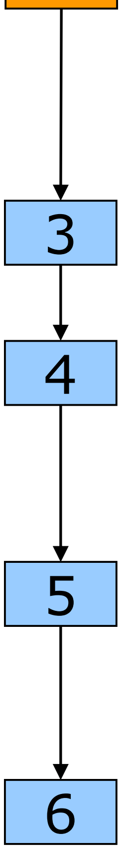
Constants: Petrol brand = -1

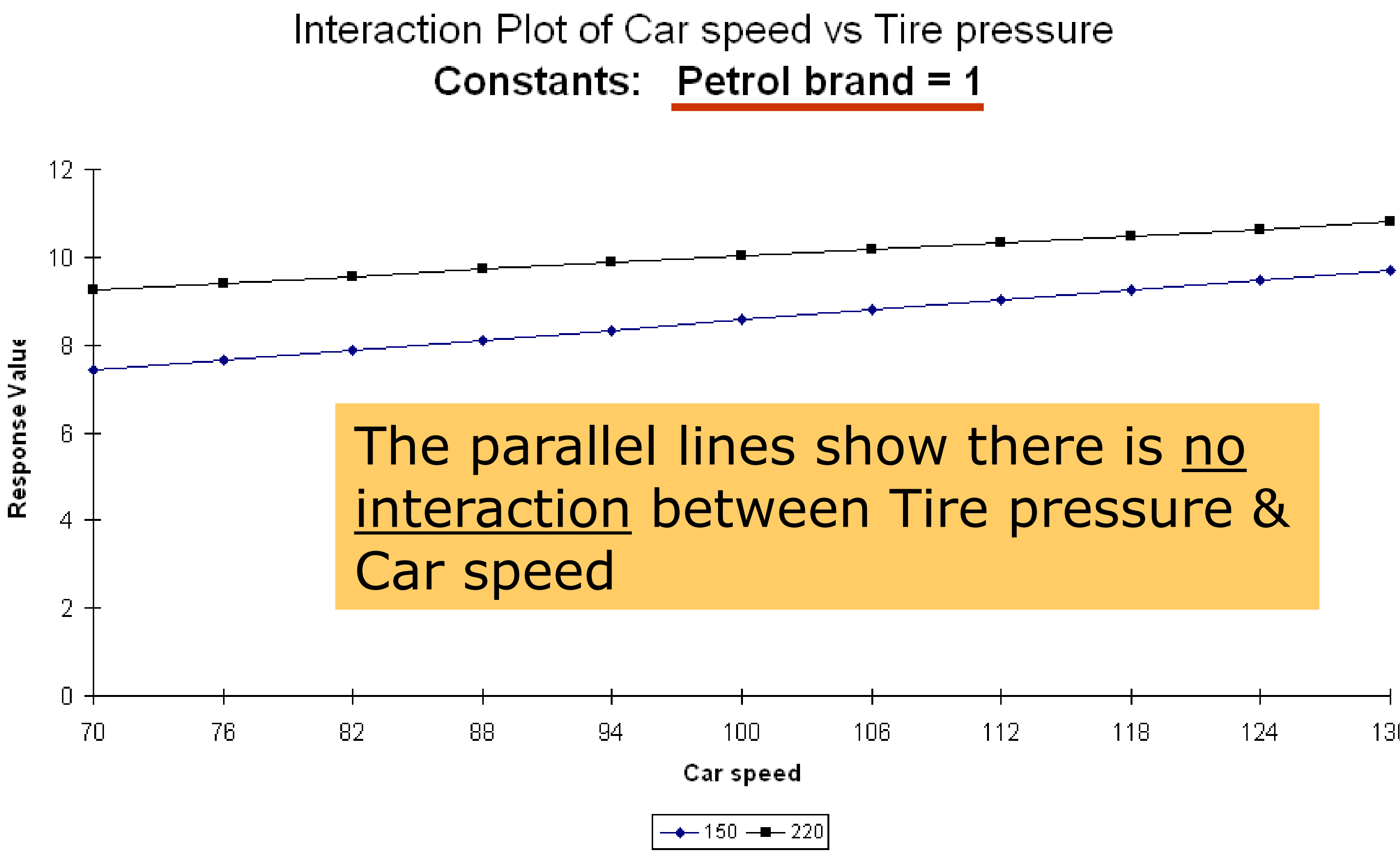


The non-parallel lines show Tire pressure interacts with Car speed differently at different level (when tire pressure is at 150 unit, speeding nearly has no effect to petrol consumption!!)

3-244. *(cont)* Step 2-2: Interaction Study

2 Analyze data (interaction analysis)





3-245. Step 2-3: Contour Plot

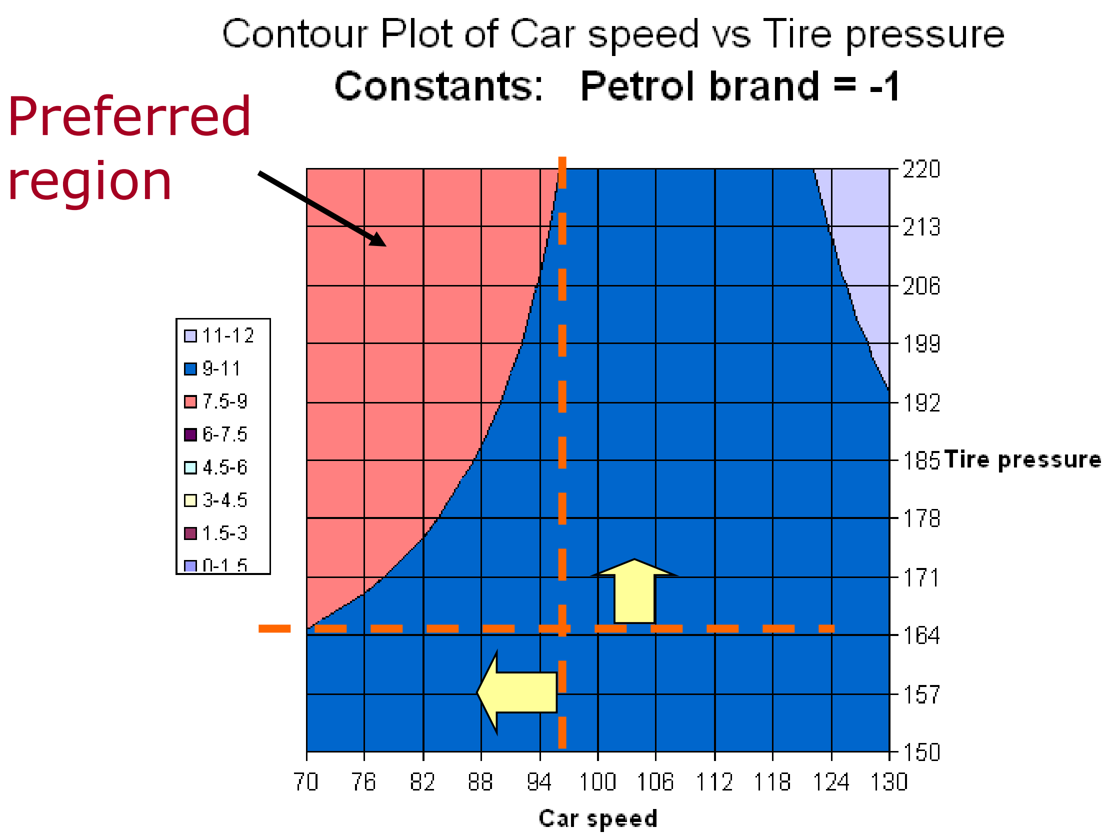
2 Analyze data (contour plot analysis)

3

4

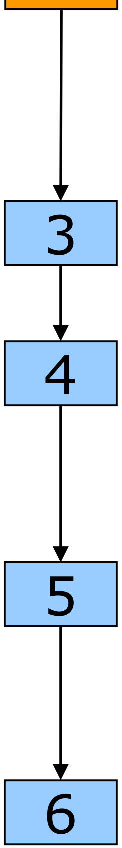
5

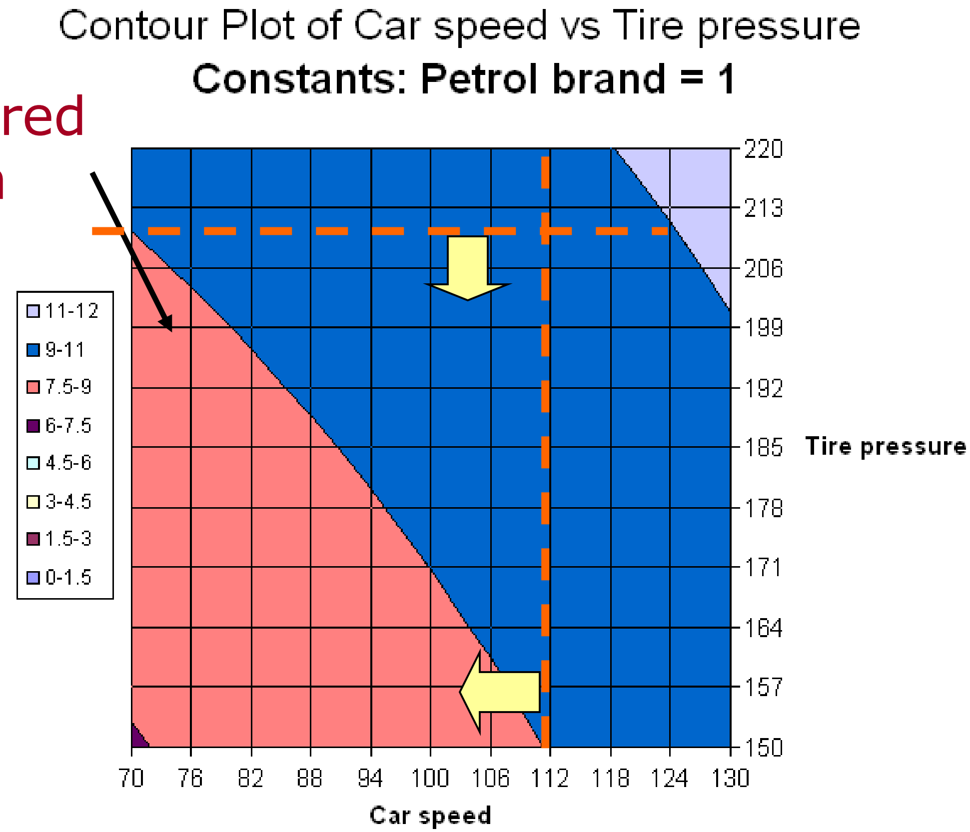
6



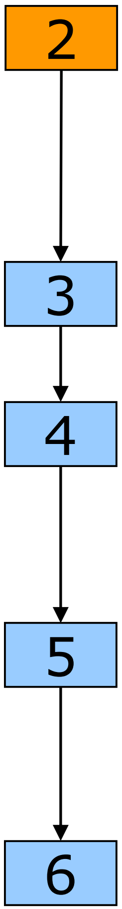
3-245. *(cont)* Step 2-3: Contour Plot

2 Analyze data (contour plot analysis)

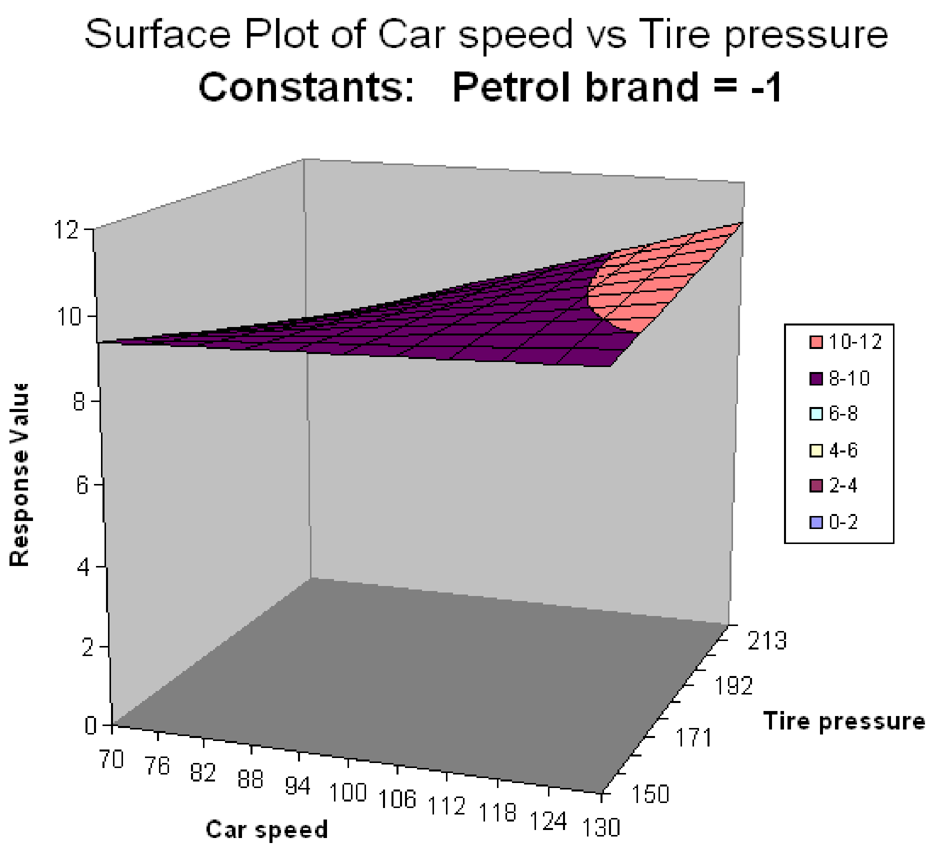




Step 2-4: Surface Plot



Analyze data (Surface plot)



3-246. *(cont)* Step 2-4: Surface Plot

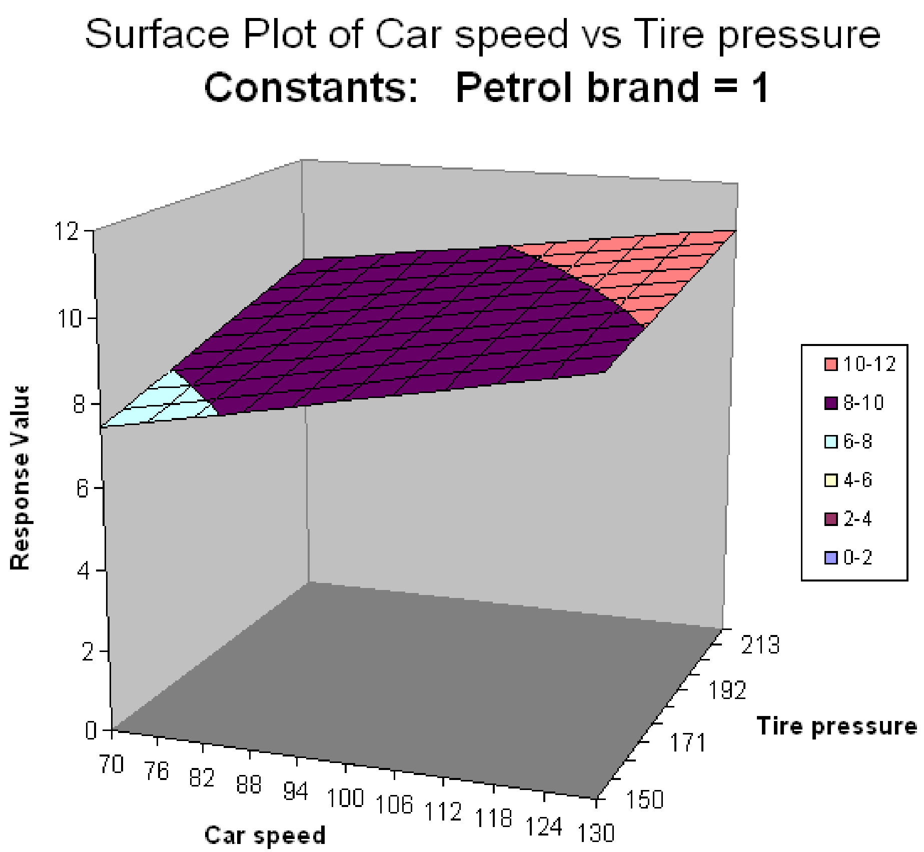
2 Analyze data (Surface plot)

3

4

5

6



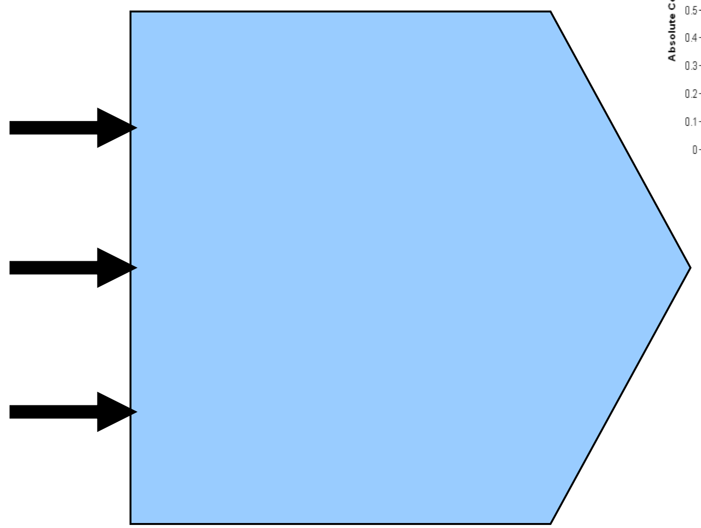
3-247. Step 3: Statistical Conclusions

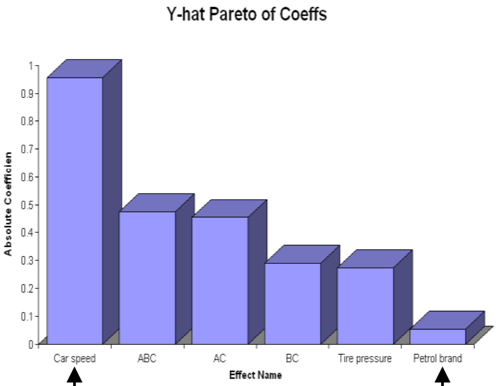
3 Draw statistical conclusions.

4

5

6





Car speed – the most significant factor

Petrol brand does not matter much

$$\hat{Y} = 9.35 - 0.05X_1 + 0.96X_2 + 0.27X_3 + 0.46X_1X_3 + 0.29X_2X_3 - 0.48X_1X_2X_3$$

(model expression)

3-248. Step 4: Confirmation Runs

4

5

6

Replicate results (confirmation runs = repeat experiments with selective setting).

Optimized setting

Factor	Name	Low	High	Exper
A	Petrol brand	-1	1	1
B	Car speed	70	130	70
C	Tire pressure	150	220	150

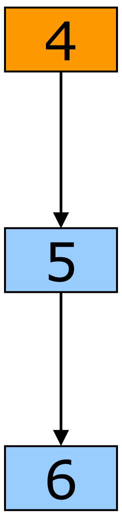
Confirmation setting

Factor	Name	Low	High	Exper
A	Petrol brand	-1	1	1
B	Car speed	70	130	90
C	Tire pressure	150	220	190

Prediction	
Y-hat	9.09162698
Std Error	0.11777762
99% Prediction Interval	
Lower Bound	8.73829411
Upper Bound	9.44495986

$\pm 3\sigma$ range (estimated)

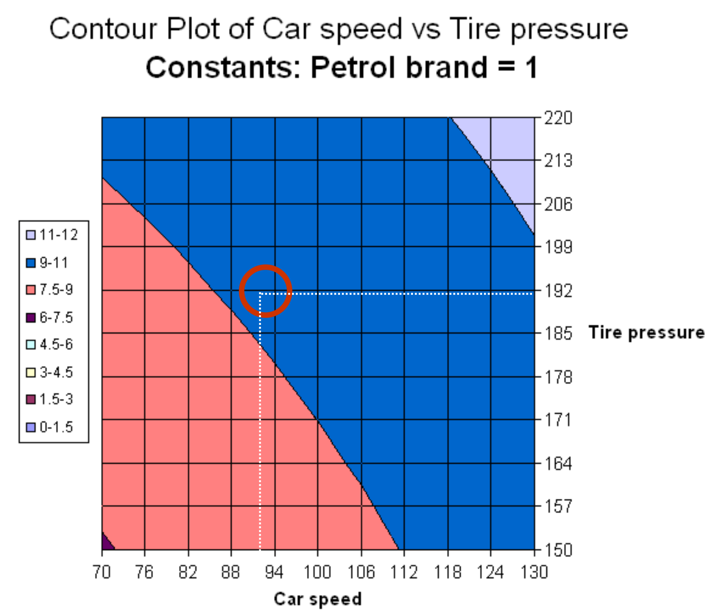
3-248. *(cont)* Step 4: Confirmation Runs



Replicate results (confirmation runs = repeat experiments with selective setting).

Confirmation setting

Factor	Name	Low	High	Exper
A	Petrol brand	-1	1	1
B	Car speed	70	130	90
C	Tire pressure	150	220	190



3-248. *(cont)* Step 4: Confirmation Runs

4

Replicate results (confirmation runs = repeat experiments with selective setting).

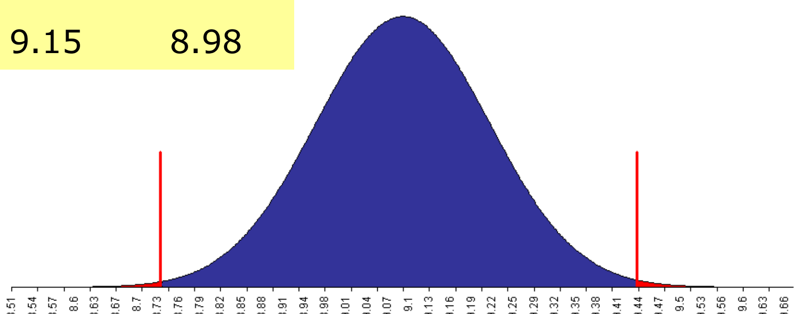
5

6

Confirmation runs should be run in Broad Inference space

9.19	9.32
9.10	9.23
8.96	9.03
9.03	8.93
9.15	8.98

Cpk Analysis



Is it acceptable?

Process Capability Analysis Of Confirmation Runs	
Upper Spec Limit	9.44
Lower Spec Limit	8.73
Mean	9.0909
Standard Deviation	0.128997373
Sigma Capability	4.011946252
Cpk	0.9021
Cp	0.9173
Defects Per Million	5976

3-2481. Inference Space

- Narrow (for experiment)
  - 1 lot of material
  - 1 day
  - 1 machine
  - 1 operator
  - 1 supplier

- Broad (for confirmation run)
  - Several lots
  - Several months
  - Five machines
  - Many operators
  - Several sources

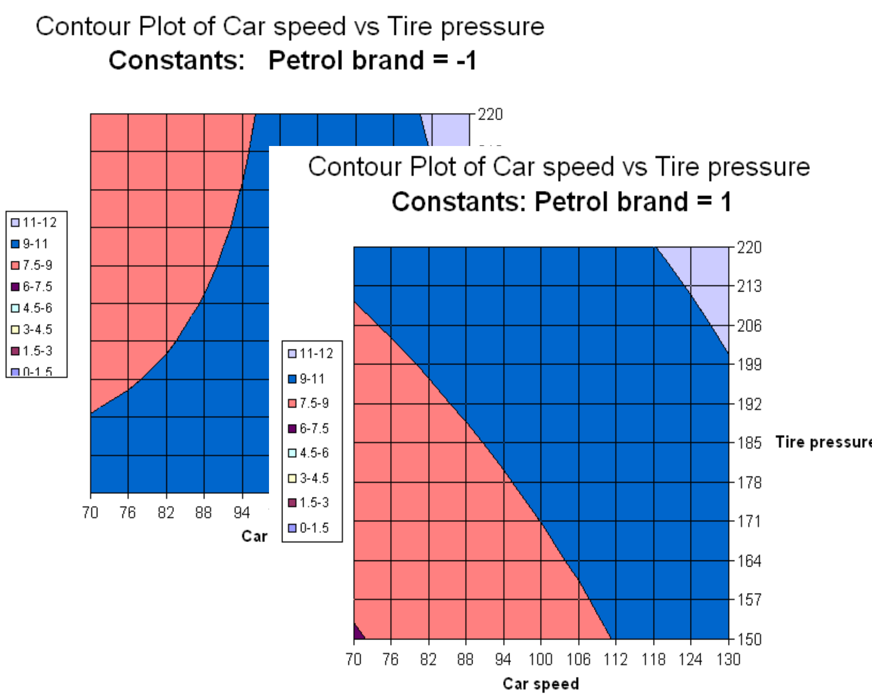
Initial studies are done under narrow inference to control noise variables

Broad inference studies are used to verify results of the narrow inference studies

3-249. Step 5: Practical Solution

5 Draw practical solutions (list down new knowledge gained).

6



Optimized setting

1. Use Cell brand petrol
2. Drive at 90 km/hr
3. Maintain tire pressure at 190 kPA

3-249. *(cont)* Step 5: Practical Solution

5 Draw practical solutions (list down new knowledge gained).

6

Economical setting

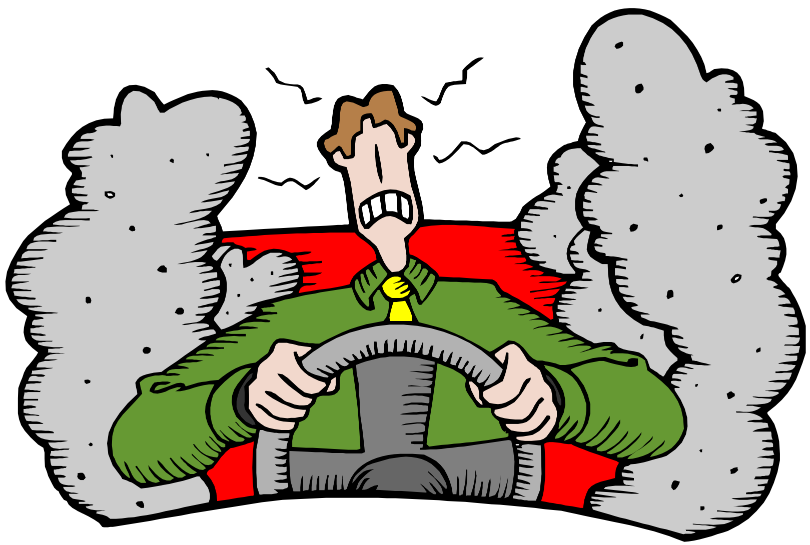
1. Drive small cc car
2. Fill petrol in at non-busy hour
3. Maintain engine oil at 'Hi' mark
4. Use normal spark plugs

Information from earlier experiment, and passed knowledge

3-2410. Step 6: Implementation

6 Implement solutions.

Most economical driving!!



SOP

1. Cell petrol only
2. 90 km/hr only
3. Tire pressure: 190 kPA
4. …

3-25. Next Steps After Modeling Experiment

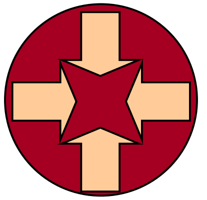
- Maintain optimum setting and handling method learned in Modeling Experiments by:
  - Control charts
  - SOPs

Gather more knew knowledge while implementing new solutions

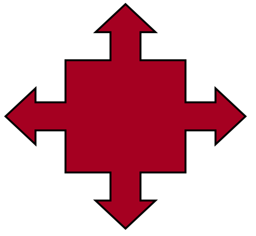
- Expand the new knowledge implemented to other operation areas (rerun DOE only when it is necessary).

3-3. Validating Experiments

◈ Notes on internal validity



- Randomization of experimental runs “spreads” the noise across the experiment
- Holding noise variables constant eliminates the effect of that variable (but it limits broad inferences)



◈ Notes on external validity

- Include samples that represent
  - Supplier variability
  - Different operation (shifts, days, section, products)
  - Different product lot

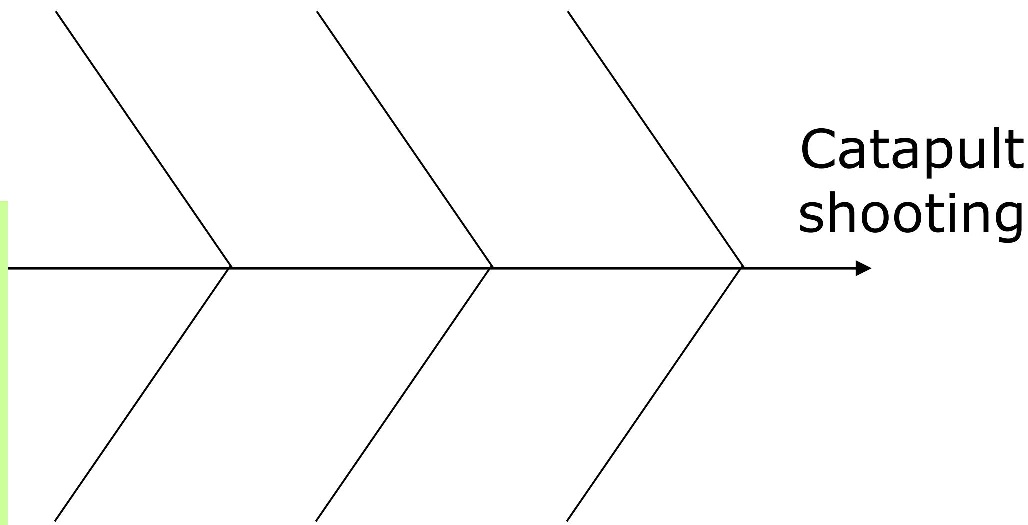
3-31. Reason if not Conforming

Confirmation Run

NG → Revisit CEDAC – study missing items

OK ↓

Practical solution



Possibilities:

- Some important factors are left out
- Too much noise
- Inappropriate level selection
- Inappropriate range selection

3-41. About Fractional Factorial

- Use in Screening experiments.
- Simplified from Full Factorial.
- Designed for main effect analysis.
- Not recommended for interaction analysis.

3-42. Number of Runs in Full Factorial

◈ Consider a Full Factorial matrix (2 levels)

2 factors

#	Sugar	Creamer
1	-	-
2	-	+
3	+	-
4	+	+

2^2 => 4 runs

3 factors

#	Sugar	Creamer	Coffee
1	-	-	-
2	-	-	+
3	-	+	-
4	-	+	+
5	+	-	-
6	+	-	+
7	+	+	-
8	+	+	+

2^3 => 8 runs

5 factors

#	S	C	→	W
1	-	-		-
2	-	-		+
3	-	-		-
4	-	-		+
5	-	-		-
6	-	-		+
7	-	-		-
8	-	-		+
↓				
32	+	+		+

2^5 => 32 runs →

As number of factors increases, number of runs rapidly increases

3-43. Full Factorial vs. Fractional Factorial

- Consider 2 levels – 5 factors experiment
  - Full Factorial: # of runs = 2^5 = 32 runs
  - Fractional Factorial:
    - ½ Fraction: ½ $2^5 = 2^{5-1} = 2^4$ = 16 runs (instead of 32 runs)
    - ¼ Fraction: ¼ 2^5 = ½ x ½ $2^5 = 2^{5-2} = 2^3$ = 8 runs

Fractional Factorial reduces number of runs by its selected fraction (eg. ½, ¼, …)

3-44. Derivation Method from a Full Factorial

◈ Consider a 2 levels – 3 factors Full Factorial experiment:

Additional factor normally placed at the highest order of interaction → Factor D

A	B	C	AXB	AXC	BXC	AXBXC
-1	-1	-1	1	1	1	-1
1	-1	-1	-1	-1	1	1
-1	1	-1	-1	1	-1	1
1	1	-1	1	-1	-1	-1
-1	-1	1	1	-1	-1	1
1	-1	1	-1	1	-1	-1
-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1

When column ABC interaction is replaced with Factor D, the ABC is aliasing with D. ABC interaction can no longer be estimated.

3-45. Examples of New Derived Design Matrix

- A new 2 levels – 4 factors Half Fraction ($2^{4–1} = 2^3$) experiment:

A	B	C	D
-1	-1	-1	-1
1	-1	-1	1
-1	1	-1	1
1	1	-1	-1
-1	-1	1	1
1	-1	1	-1
-1	1	1	-1
1	1	1	1

Instead of 16 runs (as in Full Factorial), Half Fraction requires only 8 runs.

3-45. *(cont)* Examples of New Derived Design Matrix

- Consider a 2 levels – 5 factors experiment:

Full Factorial

Factor	A	B	C	D	E
Row #	A	B	C	D	E
1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	1
3	-1	-1	-1	1	-1
4	-1	-1	-1	1	1
5	-1	-1	1	-1	-1
6	-1	-1	1	-1	1
7	-1	-1	1	1	-1
8	-1	-1	1	1	1
9	-1	1	-1	-1	-1
10	-1	1	-1	-1	1
11	-1	1	-1	1	-1
12	-1	1	-1	1	1
13	-1	1	1	-1	-1
14	-1	1	1	-1	1
15	-1	1	1	1	-1
16	-1	1	1	1	1
17	1	-1	-1	-1	-1
18	1	-1	-1	-1	1
19	1	-1	-1	1	-1
20	1	-1	-1	1	1
21	1	-1	1	-1	-1
22	1	-1	1	-1	1
23	1	-1	1	1	-1
24	1	-1	1	1	1
25	1	1	-1	-1	-1
26	1	1	-1	-1	1
27	1	1	-1	1	-1
28	1	1	-1	1	1
29	1	1	1	-1	-1
30	1	1	1	-1	1
31	1	1	1	1	-1
32	1	1	1	1	1

Half Factorial

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

Qtr Factorial

	A	B	C	D=AB	E=BC
1	-1	-1	-1	+1	+1
2	-1	-1	+1	+1	-1
3	-1	+1	-1	-1	-1
4	-1	+1	+1	-1	+1
5	+1	-1	-1	-1	+1
6	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1
8	+1	+1	+1	+1	+1

3-46. What is Aliasing?

- Two factor effects (ie ABC & D) that are represented by the same comparison (column ABC & column D are the same) are **aliases** to one another.
- Two effects that are aliases of one another are **confounded** (confused) with one another.

3way interaction column in Full Factorial (3 factors)

ABC
-1
1
1
-1
1
-1
-1
1

Calculation and estimation of ABC interaction is no longer possible

D
-1
1
1
-1
1
-1
-1
1

4th factor column in Half Fraction (4 factors)

3-47. The Cost of Reducing Runs

- By using reduced runs in Half Fraction (or other Fractional Fraction), we <u>loose information of higher order interaction</u>.

Is it okay to loose higher order interaction?

Study design Resolution

3-48. Resolution

- Resolution III Designs:
  - No main effects are aliased with other Main Effects. **Main Effects aliased with two-factor interactions.**
- Resolution IV Designs
  - No Main Effect aliased with other Main Effects or with two-factor interactions.
  - **Two-factor interactions aliased with other two-factor interactions.**
- Resolution V Designs
  - Main Effects okay, **Two-factor interactions aliased with 3-factor interactions**

3-481. Selecting Design – Example

Example 3

- Identify type of experiment: **2 level – 5 factors Screening experiment**
- **Choose number of runs.** Say we choose to run 8 trials only – 8 run trial comes from 3 factors Full Factorial (5 factors Full Factorial requires 32 runs).

A	B	C	*AB*	*AC*	*BC*	*ABC*

3 factors Full Factorial

A	B	C	D=AB	E=AC	*BC*	*ABC*

Fractional Factorial

3-481. *(cont)* Selecting Design – *Example*

A	B	C	D=AB	E=AC	*BC*	*ABC*

Fractional Factorial

- Factor D aliases with AB interactions; Factor E aliases with AC interaction (we loose information on AB & AC interactions – **this normally is not a problem in Screening experiment as we are mainly looking for main effects).**

3-481. *(cont)* Selecting Design – *Example*

◈ ¼ Fractional matrix for the experiment

	A	B	C	D=AB	E=AC
1	-1	-1	-1	+1	+1
2	-1	-1	-1	+1	+1
3	-1	-1	+1	+1	-1
4	-1	-1	+1	+1	-1
5	-1	+1	-1	-1	-1
6	-1	+1	-1	-1	-1
7	-1	+1	+1	-1	+1
8	-1	+1	+1	-1	+1

This is a Resolution III design.

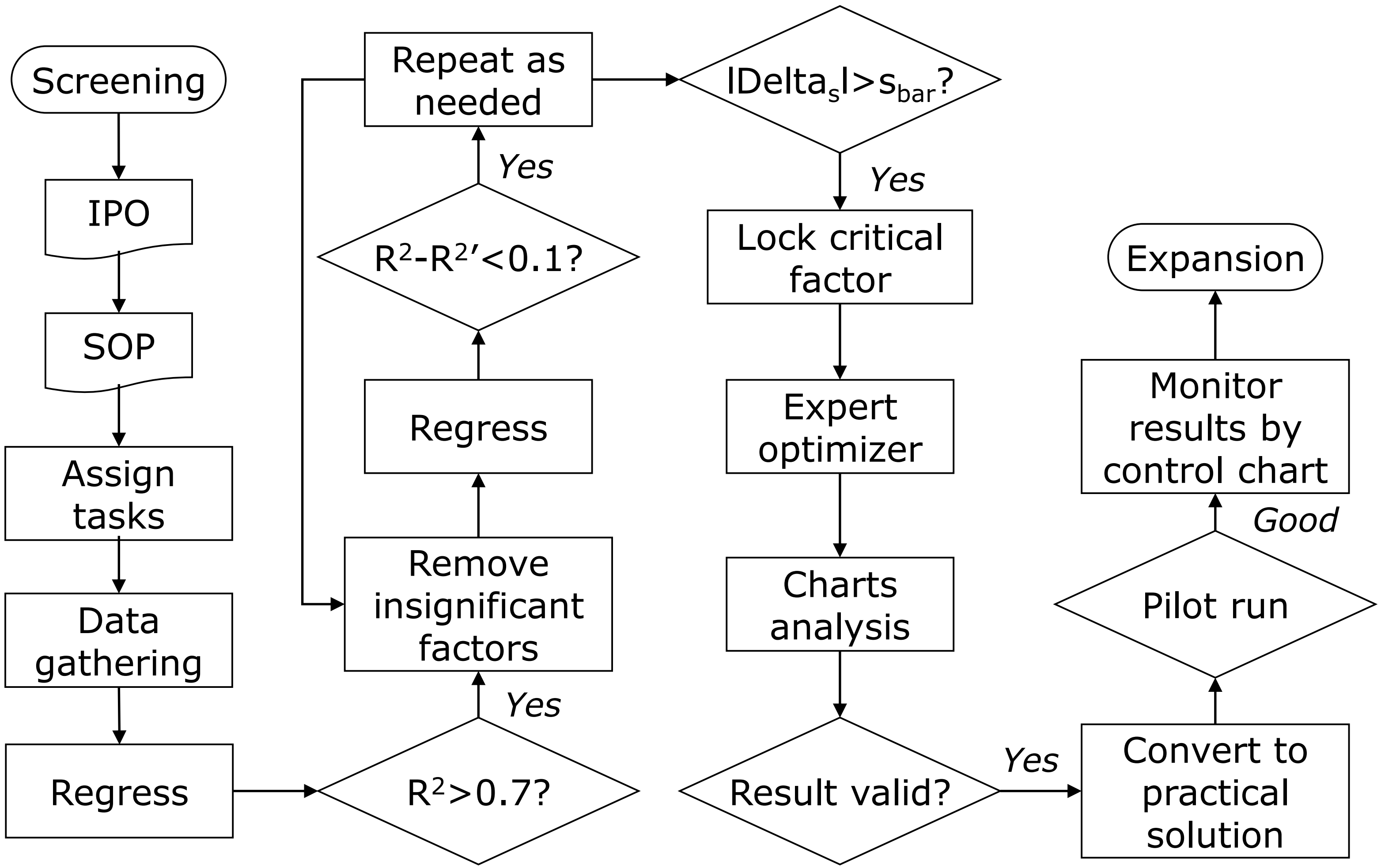
- No main effects are aliases with other Main Effects. Main Effects aliases with two-factor interactions.

3-49. Notes on Fractional Factorial

- Advantages of Fractional Factorial
  - If the higher order interactions (ABCD, ABC, ABD, BCD …) are assumed negligible, a fraction of the full factorial can still give good estimates of low-order interactions.
- Disadvantages of Fraction Factorial
  - Loose information about interaction of higher order.
- Fractional Factorial is normally used at Screening experiments (at early stage of improvement project).

Summary

- Steps in Modeling design



(cont) Summary

- Objective of modeling DOE
  - Process optimization
  - Interaction study
  - Model expression
- In Full Factorial design, all corners (of a box) are investigated.
- OFAT may find the local optimum, but most likely to miss the model optimum.
- Fractional Factorial has the advantage of having reduced number of runs and still comes out with a good estimation.
- Resolution describes the aliasing of factors with other columns (ie with 2way, 3way interaction).

(cont) Summary

- What is modeling experiments?

- What is it for?

- How to start modeling experiment?

(cont) Summary

- How to use modeling experiment to model the significant factors?
- What are the precautions when running modeling experiment?
- What are the differences between screening experiments and modeling experiments?

Discussion Y1: Outcome of Modeling Experiments

Modeling Experiments

- Identification of main effects **(magnitude of each factors effect).**
- Identification of interaction **(how the factors work to affect the response).**
- The model expression (the equation of Y = …)
- Optimization **(the main purpose of Modeling Experiments)**
- Confirmation runs **(to test new finding)**

Discussion Y2: Advantages of Fractional Factorial

- Can test many factors with small number of runs (reduced runs), and still can analyze the main factors effect effectively.

#	A	B	C	D=AB	E=AC	BC	ABC
1	-1	-1	-1	-1	+1		
2	-1	-1	+1	+1	+1		
3	-1	+1	-1	+1	-1		
4	-1	+1	+1	-1	-1		
5	+1	-1	-1	+1	-1		
6	+1	-1	+1	-1	-1		
7	+1	+1	-1	-1	+1		
8	+1	+1	+1	+1	+1		

¼ Fraction (5 factors for only 8 runs)

Discussion Y3: Calculation of Runs in Full Factorial

2 levels

- Number of runs
  - $= 2^n$ (n: # of factors)
- *Example*
  - n: 3; Trials $= 2^3 = 8$ runs

3 levels

- Number of runs
  - $= 3^n$ (n: # of factors)
- *Example*
  - n: 3; Trials $= 3^3 = 27$ runs

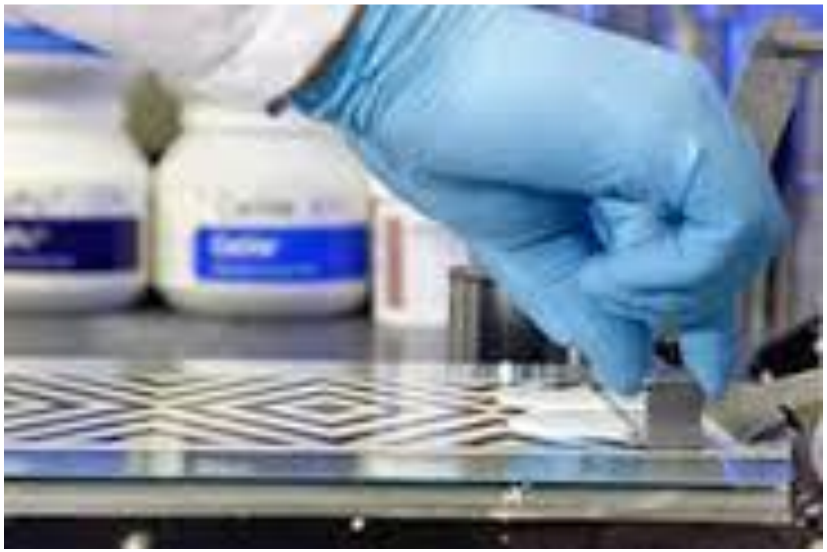
DOE for Practitioners

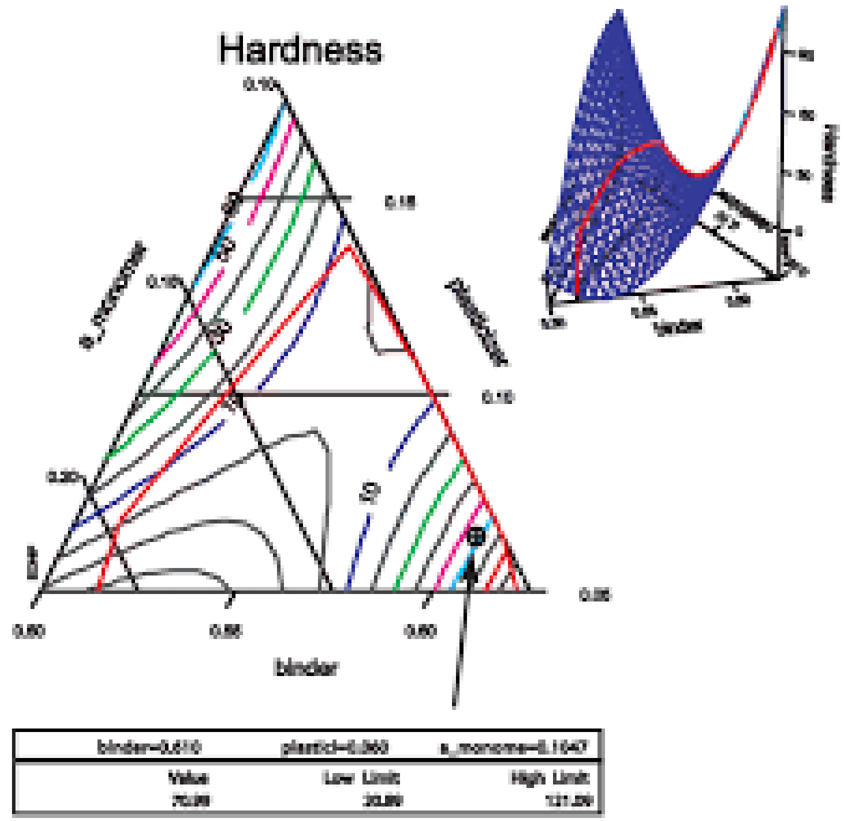
Screening DOE

Modeling DOE

OVERVIEWING DOE MIXTURES

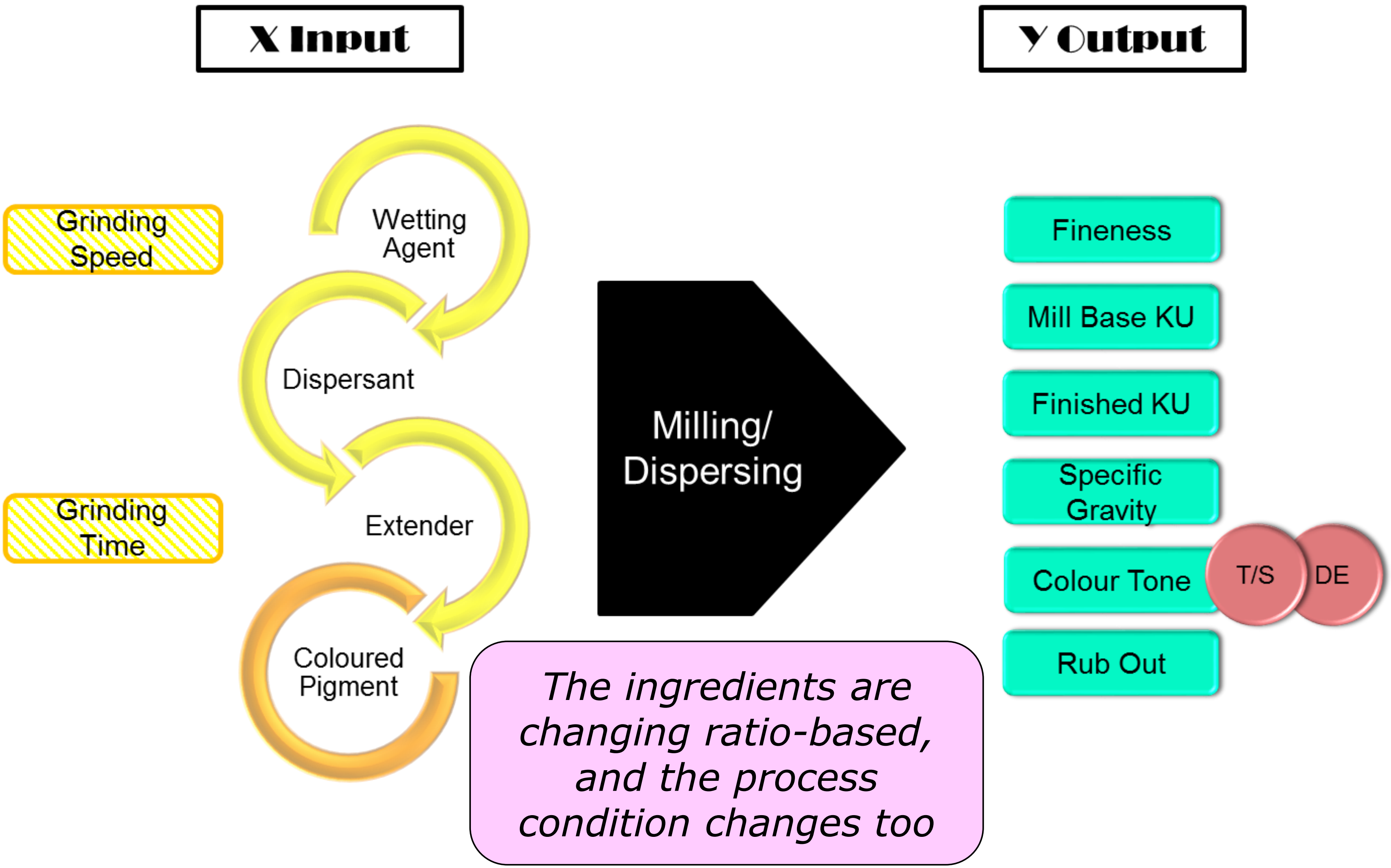
Hint of the Module...





Awesome!!

4-1. Scenario of DOE Mixtures



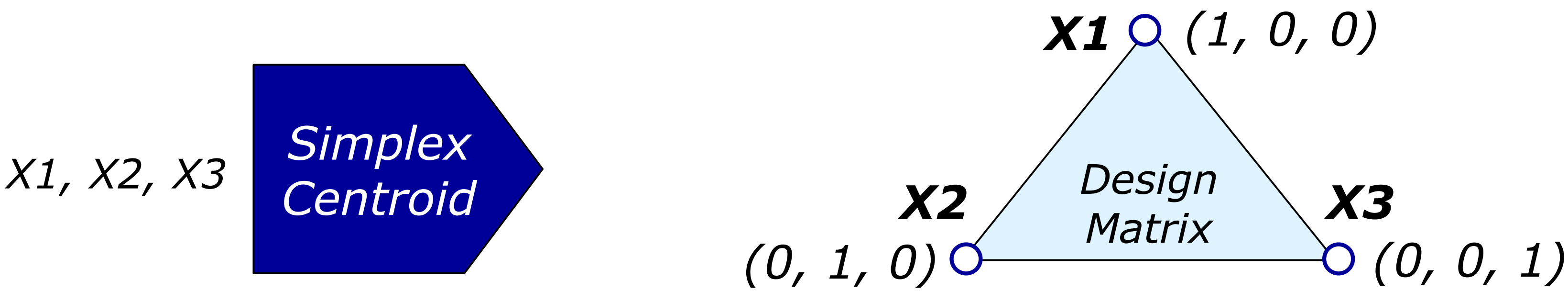
4-2. Different Types of DOE Mixtures

Types	Response depends on …	Example
Mixture	the relative proportions of the components only	the taste of lemonade depends *only* on the proportions of lemon juice, sugar, and water
Mixture Amount	the relative proportions of the components, and the amount of the mixture	the yield of a crop depends on the proportions of the insecticide ingredients *and* the amount of the insecticide applied
Mixture – Process variable	the relative proportions of the components *and* process variables. Process variables are factors that are not part of the mixture but may affect the blending properties of the mixture	the taste of a cake depends on the cooking time and cooking temperature, *and* the proportions of cake mix ingredients

4-31,2. Mixture Design 1, 2: Simplex Centroid & Simplex Lattice

Simplex centroid and simplex lattice designs

Mixture designs in which the design points are arranged in a uniform manner (or lattice) over an L-simplex. An L-simplex is similar to and has sides parallel to the triangle shown below:



For both simplex centroid and simplex lattice designs, you can add points to the interior of the design space. These points provide information on the interior of the response surface; thereby, improving coverage of the design space.

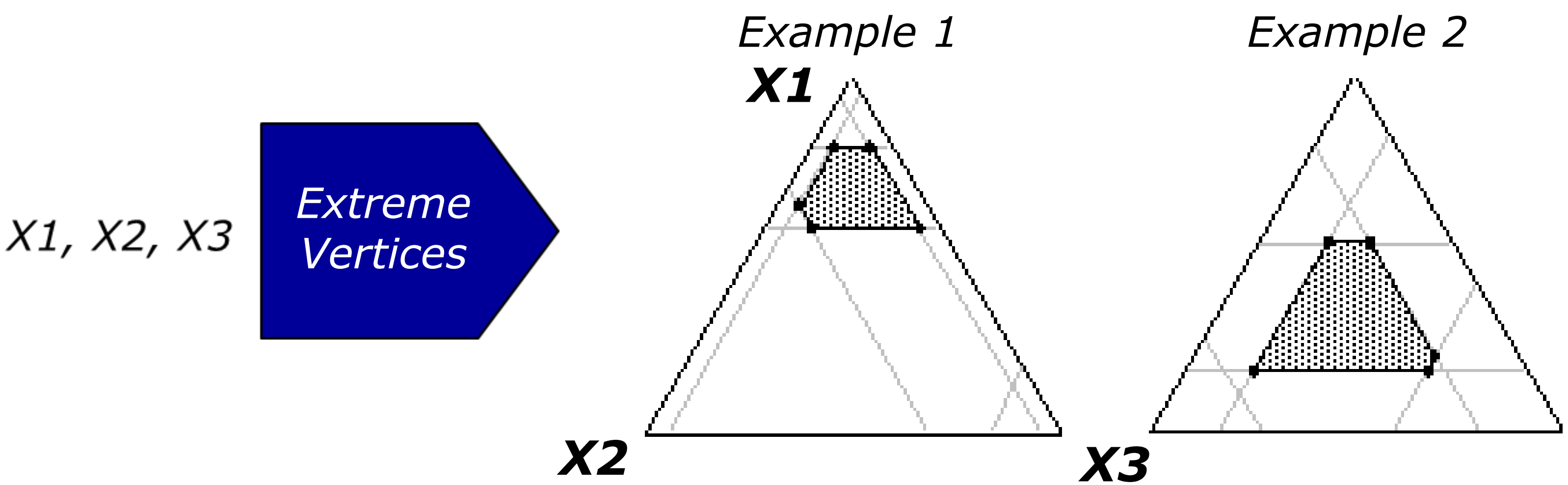
4-33. Mixture Design 3: Extreme Vertices

<u>Extreme vertices designs</u>

Mixture designs that cover only a subportion or smaller space within the simplex. These designs must be used when your chosen design space is not an L-simplex. <u>The presence of both lower and upper bound constraints on the components often create this condition</u>. For example, you need to determine the proportions of flour, milk, baking powder, eggs, and oil in a pancake mix that would produce an optimal product based on taste. Because previous experimentation suggests that a mix that does not contain all of the ingredients or has too much baking powder will not meet the taste requirements, you decide to constrain the design by setting lower bounds and upper bounds.

4-33. *(cont)* Mixture Design 3: Extreme Vertices

<u>The goal of an extreme vertices design is to choose design points that adequately cover the design space</u>. The illustration below shows the extreme vertices for two three-component designs with both upper and lower constraints:



The light grey lines represent the lower and upper bound constraints on the components. The dark grey area represents the design space. <u>The points are placed at the extreme vertices of design space.</u>

4-341. Specifying Design Space for NP Case – Colorant

Number of Boundaries for Each Dimension

Point Type	1	2	3	4	5	6	0
Dimension	0	1	2	3	4	5	6
Number	7	21	35	35	21	7	1

Number of Design Points for Each Type

Point Type	1	2	3	4	5	6	7	0	-1
Distinct	7	0	0	0	0	0	0	1	7
Replicates	1	0	0	0	0	0	0	1	1
Total number	7	0	0	0	0	0	0	1	7

Bounds of Mixture Components

	Amount		Proportion		Pseudocomponent	
Comp	Lower	Upper	Lower	Upper	Lower	Upper
A	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
B	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
C	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
D	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
E	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
F	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
G	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

Only one constraint specified

Linear Constraints of Mixture Components

Constraint	Lower	A	B	C	D	E	F	G	Upper	
1		3.0000	1.0000	2.0000	2.0000	5.0000	6.0000	8.0000	50.0000	**

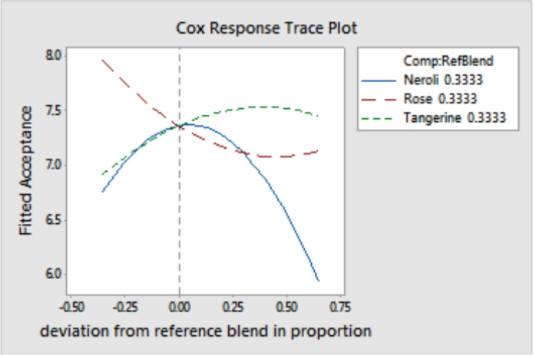
4-342. Combination for the Colorant Screening Experiment

StdOrder	RunOrder	PtType	Blocks	Wetting 1	Wetting 2	Dispersing 1	Dispersing 2	Dispersing 3	Extender 1	Pigment 1
1	1	1	1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000
2	2	1	1	1.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3	3	1	1	0.00000	1.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	4	1	1	0.00000	0.00000	1.00000	0.00000	0.00000	0.00000	0.00000
5	5	1	1	0.00000	0.00000	0.00000	1.00000	0.00000	0.00000	0.00000
6	6	1	1	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000	0.00000
7	7	1	1	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000
8	8	0	1	0.14286	0.14286	0.14286	0.14286	0.14286	0.14286	0.14286
9	9	-1	1	0.07143	0.07143	0.07143	0.07143	0.07143	0.07143	0.57143
10	10	-1	1	0.57143	0.07143	0.07143	0.07143	0.07143	0.07143	0.07143
11	11	-1	1	0.07143	0.57143	0.07143	0.07143	0.07143	0.07143	0.07143
12	12	-1	1	0.07143	0.07143	0.57143	0.07143	0.07143	0.07143	0.07143
13	13	-1	1	0.07143	0.07143	0.07143	0.57143	0.07143	0.07143	0.07143
14	14	-1	1	0.07143	0.07143	0.07143	0.07143	0.57143	0.07143	0.07143
15	15	-1	1	0.07143	0.07143	0.07143	0.07143	0.07143	0.57143	0.07143

4-35. Next Steps

Design Table (randomized)

Run	Type	A	B	C	D	E
1	1	0.425000	0.300000	0.050000	0.100000	0.125000
2	2	0.437500	0.300000	0.050000	0.112500	0.100000
3	-1	0.429687	0.329687	0.031250	0.104688	0.104688
4	2	0.450000	0.300000	0.025000	0.125000	0.100000
5	2	0.425000	0.300000	0.025000	0.125000	0.125000
6	2	0.425000	0.325000	0.025000	0.125000	0.100000
7	2	0.462500	0.300000	0.037500	0.100000	0.100000
8	1	0.425000	0.300000	0.025000	0.150000	0.100000
9	2	0.425000	0.312500	0.050000	0.100000	0.112500
10	2	0.425000	0.337500	0.037500	0.100000	0.100000
11	-1	0.429687	0.304688	0.043750	0.117188	0.104688
12	2	0.425000	0.300000	0.050000	0.112500	0.112500
13	-1	0.429687	0.304688	0.031250	0.104688	0.129688
14	2	0.450000	0.325000	0.025000	0.100000	0.100000
15	2	0.425000	0.312500	0.050000	0.112500	0.100000
16	2	0.437500	0.312500	0.050000	0.100000	0.100000
17	-1	0.429687	0.304688	0.031250	0.129688	0.104688
18	0	0.434375	0.309375	0.037500	0.109375	0.109375
19	-1	0.454687	0.304688	0.031250	0.104688	0.104688
20	1	0.425000	0.325000	0.050000	0.100000	0.100000
21	2	0.437500	0.300000	0.050000	0.100000	0.112500
22	1	0.425000	0.300000	0.050000	0.125000	0.100000
23	2	0.425000	0.300000	0.037500	0.137500	0.100000
24	2	0.425000	0.300000	0.037500	0.100000	0.137500
25	1	0.425000	0.300000	0.025000	0.100000	0.150000
26	-1	0.429687	0.304688	0.043750	0.104688	0.117188
27	2	0.425000	0.325000	0.025000	0.100000	0.125000
28	-1	0.442187	0.304688	0.043750	0.104688	0.104688
29	-1	0.429687	0.317188	0.043750	0.104688	0.104688
30	1	0.475000	0.300000	0.025000	0.100000	0.100000
31	1	0.450000	0.300000	0.050000	0.100000	0.100000
32	1	0.425000	0.350000	0.025000	0.100000	0.100000
33	2	0.450000	0.300000	0.025000	0.100000	0.125000



Create design space → Fitting a model → Creating a response trace plot → Creating a contour plot & a surface plot

Regression for Mixtures: Acceptance versus Neroli, Rose, Tangerine

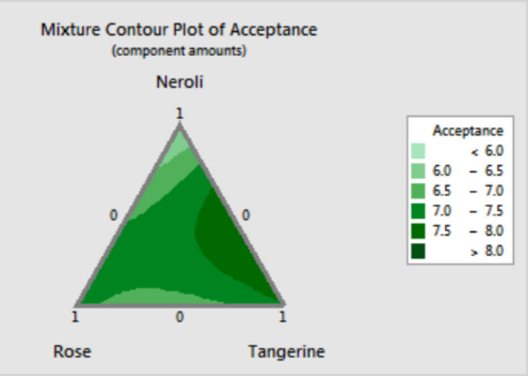
Estimated Regression Coefficients for Acceptance (component proportions)

Term	Coef	SE Coef	T	P	VIF
Neroli	5.856	0.4728	*	*	1.964
Rose	7.141	0.4728	*	*	1.964
Tangerine	7.448	0.4728	*	*	1.964
Neroli*Rose	1.795	2.1791	0.82	0.456	1.982
Neroli*Tangerine	5.090	2.1791	2.34	0.080	1.982
Rose*Tangerine	-1.941	2.1791	-0.89	0.423	1.982

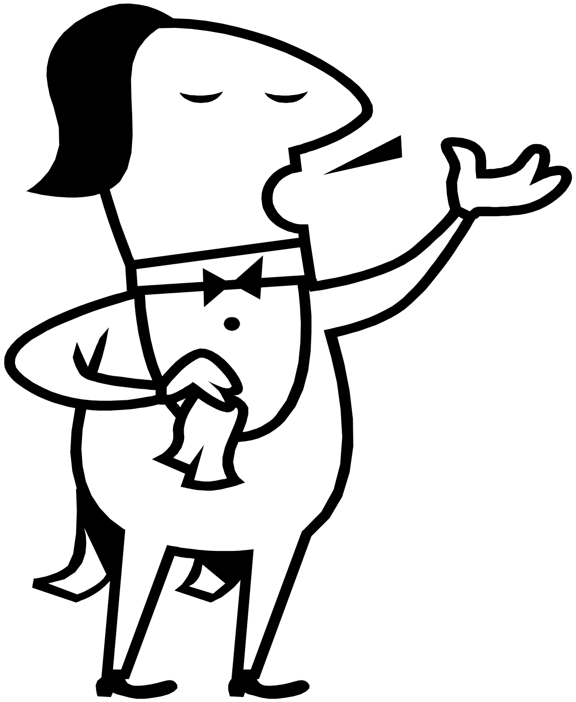
S = 0.490234   PRESS = 11.4399
R-Sq = 73.84%   R-Sq(pred) = 0.00%   R-Sq(adj) = 41.14%

Analysis of Variance for Acceptance (component proportions)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	2.71329	2.71329	0.54266	2.26	0.225
Linear	2	1.04563	1.56873	0.78437	3.26	0.144
Quadratic	3	1.66766	1.66766	0.55589	2.31	0.218
Neroli*Rose	1	0.15963	0.16309	0.16309	0.68	0.456
Neroli*Tangerin	1	1.31728	1.31109	1.31109	5.46	0.080
Rose*Tangerin	1	0.19075	0.19075	0.19075	0.79	0.423
Residual Error	4	0.96132	0.96132	0.24033		
Total	9	3.67461				



MODULE OBJECTIVES DELIVERED..



Can I take your last order Sir?

Appendix: List of Training Aids in DOE - Modeling

Training Aids	Title
Example 1	Interaction Effects
Example 2-1	FF vs. OFAT
Example 2-2	What OFAT misses
Case Study 1	The Mechanic Case

Training Aids	Title