Anexas Consultancy Services

Welcome to

ANEXAS EUROPE

Part B

Master Black Belt Training

Amitabh Saxena

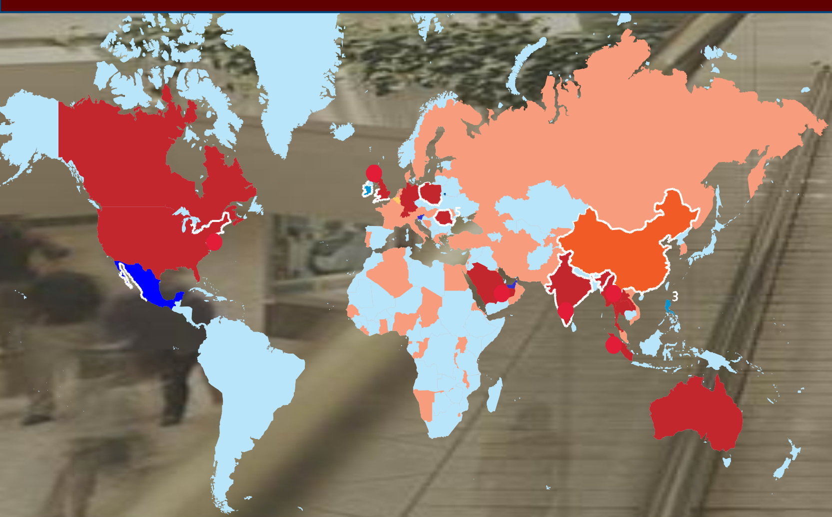
Introductions and Expectations

Introduce yourself to the group:

- Name
- Job
- Any one thing not many people know about you!

About Anexas

Enabling individuals and organizations achieve excellence since 2006



- Anexas is a global network of lean six sigma and project professionals serving the wide spectrum of industries. We operate in 10 countries and have 25 professionals in the team.
- Our mission is to help organizations and individuals achieve excellence.
- Trained more than 300,000 professionals in Lean and Six Sigma, Project Management and quality related trainings across the world from various industries.
- Lean Green Belts and Black Belts certified by Anexas have completed more than 5000 successful projects under our guidance.

Introductions and Expectations

Introduce yourself to the group:

- Name
- Job
- Any one thing not many people know about you!
- Expectations for the session

Anexas Consultancy Services

Chapter 1:
Change Management

Change Management is Critical

- It is not the strongest of the species that survive, nor the most intelligent, but the ones most responsive to change.
  - Charles Darwin

- We cannot solve our problems with the same thinking we used when we created them."
  - Albert Einstein

Change Management

$$Q \times A = E$$

Quality Acceptance Effectiveness

Change Management - Overview

D + V + F > R

Dissatisfaction

+

Vision

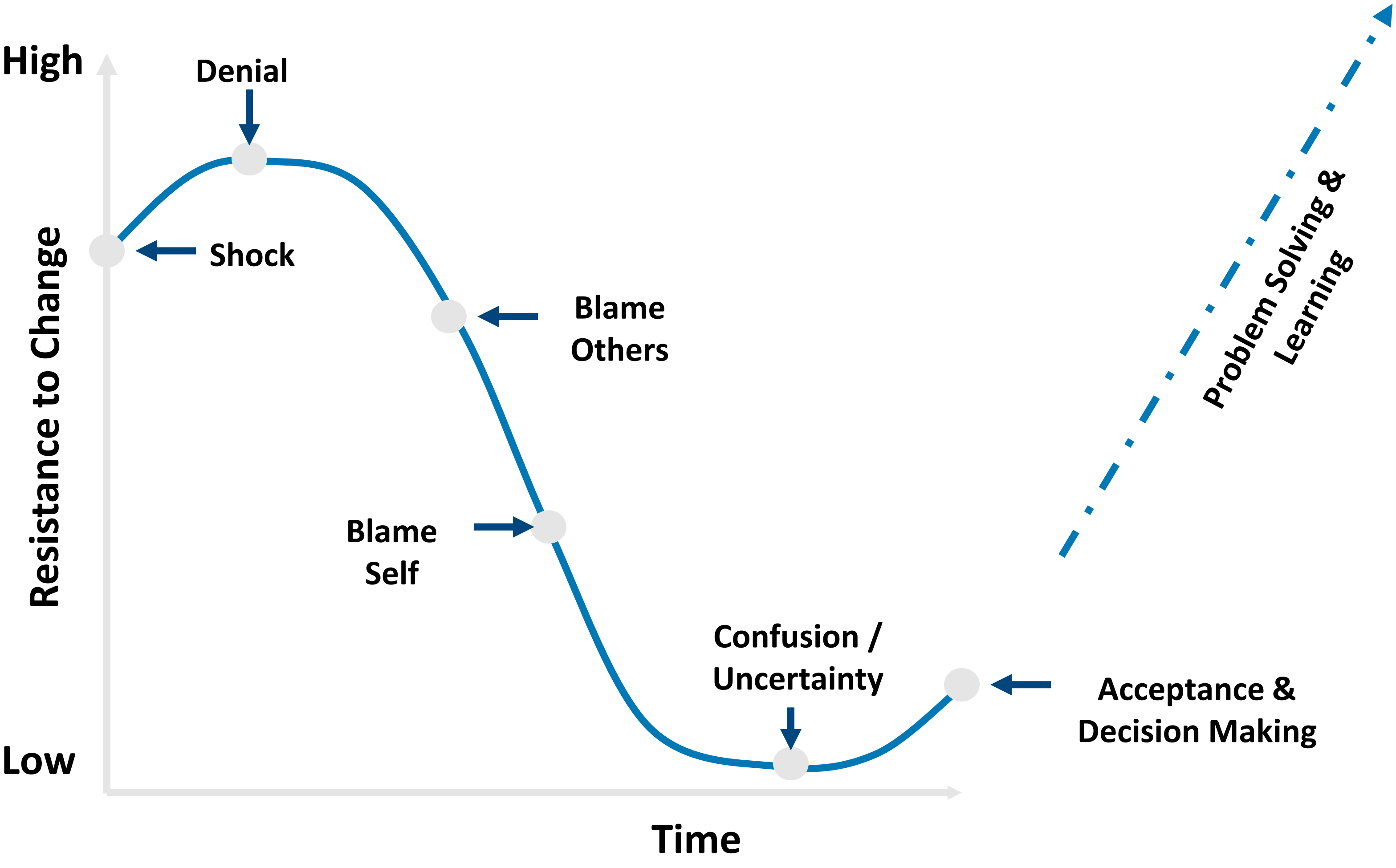
+

First Step

>

Resistance

Change Management – Change Curve



Stakeholders Analysis

Analyzing Stakeholders Position

Names	Strongly Against	Moderately Against	Neutral	Moderately Supportive	Strongly Supportive

Steps:

1. Plot where individuals currently are with regard to desired change (= current)
2. Plot where individuals need to be (X = desired) in order to successfully accomplish desired change – identify gaps between current and desired positions
3. Indicate how individuals are linked to each other, draw lines to indicate an influence link using an arrow ( ) to indicate who influences whom
4. Plan action steps for closing gaps

Stakeholders Analysis

Analyzing Sources of Resistance

TPC matrix

Source of Resistance	Definition of causes of resistance	Examples from this change
Technical	▪ Habit and inertia ▪ Difficulty in learning new skills ▪ Lack of skills	
Political	▪ Threats to old guard from new guard ▪ Relationships and networks ▪ Loss of power or authority ▪ Self-preservation	
Cultural	▪ Selective perception ▪ Old "mindset" ▪ Strong values hampering change	

Change Management Actions

Collecting Elements of Proof

3 Ds matrix

Variety Of Approaches...	Elements of proof
Produce Data: ▪ Internal sources ▪ External networks	
Show Demonstration: ▪ Finding examples ▪ Best practices ▪ Visiting other organizations	
Generate Demand: ▪ Leadership modeling ▪ High standards ▪ “Walk the talk”	

Change Management is Critical

- Change Management is Critical
  - Question is not whether change is needed but rather how much and how often
  - Healthcare is a complicated system that demands change quickly to aim at being a high reliability organization
  - High reliability is desired in healthcare. When caring for patients, we must take aim at reducing harm in an extremely complex environment
- Critical factors to assess change
  - Limits of human performance
  - System's actual capacity to handle change

Scope of change

- Scope of change
  - Change is moving from an existing state through a transition state to a future state
  - Change is inevitable and essential for growth
- Different models and strategies are required, depending on:
  - Type of change
  - People involved
  - Scope of change

Change Management Process

- Change Management Process
  - Existing State
  - Transition
  - Future State

Change Management Process

- Significant Change
  - Significant change, may cause distress and disruption (Pandemic requiring full PPE, causing high census and increased ICU cases).Transition
- Second order: Complex change
  - Complex, requires significant change in behavior (new computer system or new clinical guidelines).
- First order: small change
  - Small, requires minimal effort (new form, obtaining same product from a different vendor).

Change Management (continued)

- An organization's ability to handle frequent change is dependent upon individuals and leaders
- Successful transformation requires Transformational Leadership.
- Role of leaders is to establish the culture of change, and model behaviors needed to adapt to change
- Sense of loss must be acknowledged in the early stage of change to help people transition
- The engagement of an interdisciplinary team of staff and stakeholders is key so that change seen positively is valued

Change Management (continued)

- Resilience is the process of adapting well in the face of adversity or significant stress
- Flexibility of individuals is a critical element in an organization's ability to:
  - Make changes quickly
  - Rebound from one change to another
- Individuals can build the skill of resilience

Change Management (continued)

- Change is not a linear process
  - Individuals will be at different stages of accepting and implementing change
- The quality professional's role is as change agent, improvement advisor, and facilitator
  - It is critical to continuously assess where staff is in their acceptance of the change

Assessing Change Readiness

- In this activity, you'll be tasked with assessing organizational readiness for change. There are 7 categories and choose which of the options best indicates maximum readiness for change.
- Category
    1. LEADING CHANGE
    2. CREATING SHARED NEED
    3. SHAPING A VISION
    4. MOBILIZATION COMMITMENT
    5. MONITORING PROGRESS
    6. FINISHING THE JOB
    7. ANCHORING THE CHANGE

Strategies to Support Rapid Change

- Leadership
  - Active, visible, supportive of change
  - Makes focus or goal of change clear
  - Strong communicator of change
  - Accountable for change (owns, doesn't delegate)
  - Involves all leaders and expects results
  - Manages perceptions and expectations

Strategies to Support Rapid Change

- Structure
  - Develops strong team infrastructure
  - Team members are interdependent with respect to information, resources and skills
  - Empowers team members
  - Encouraged to use their knowledge, expertise and creativity as they work toward a successful change
  - Makes resources available
  - Includes the human, financial, technological resources, as well as others as needed

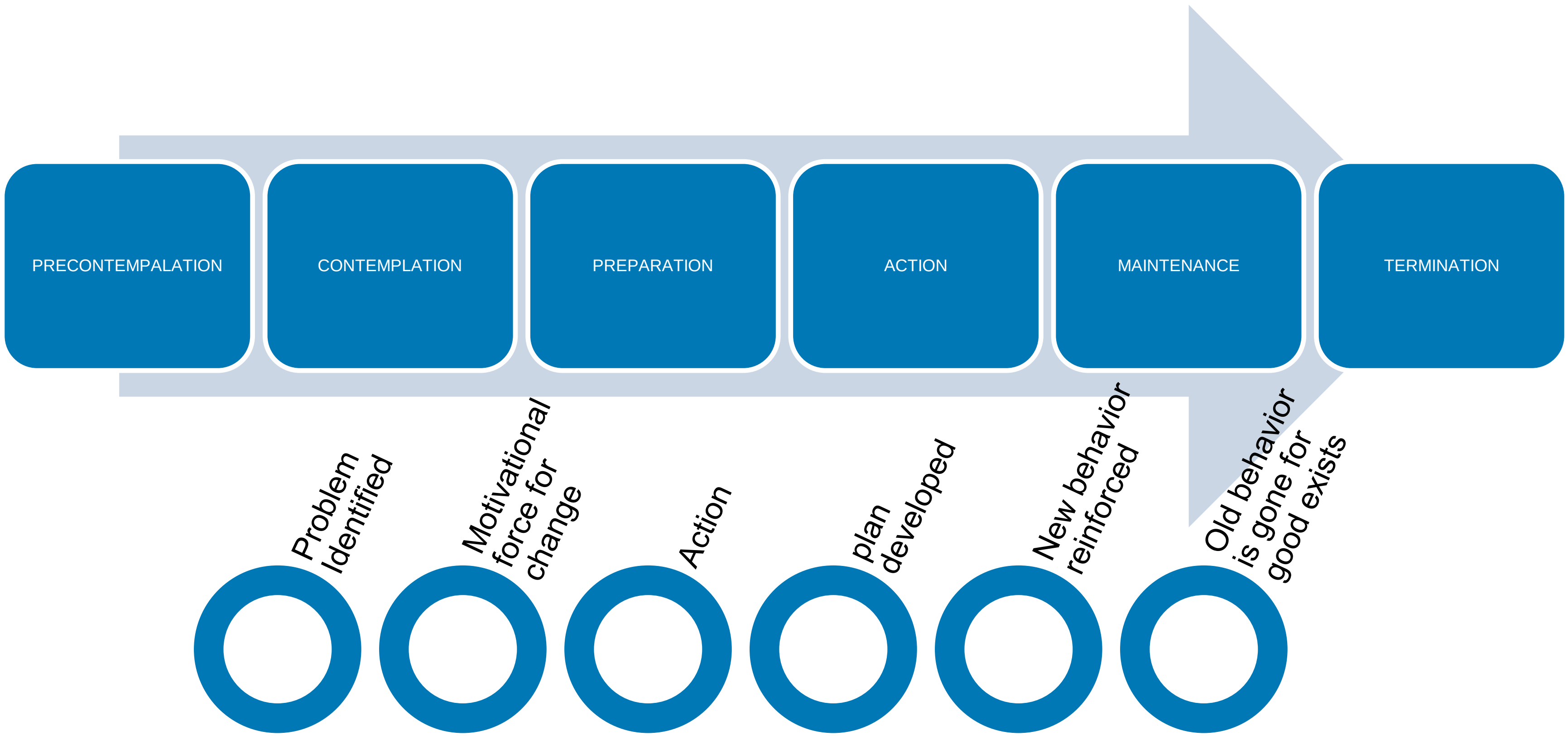
Strategies to Support Rapid Change

- Culture
  - Assesses readiness for change
  - Supports creativity, innovation, and risk-taking
  - Accepts failures, successes, and encourages ideas
  - Supports participative structure (including staff levels)
  - Focuses/values group learning and improvement
  - Encourages diversity
  - Focuses on systems/processes
  - Rewards individuals and teams

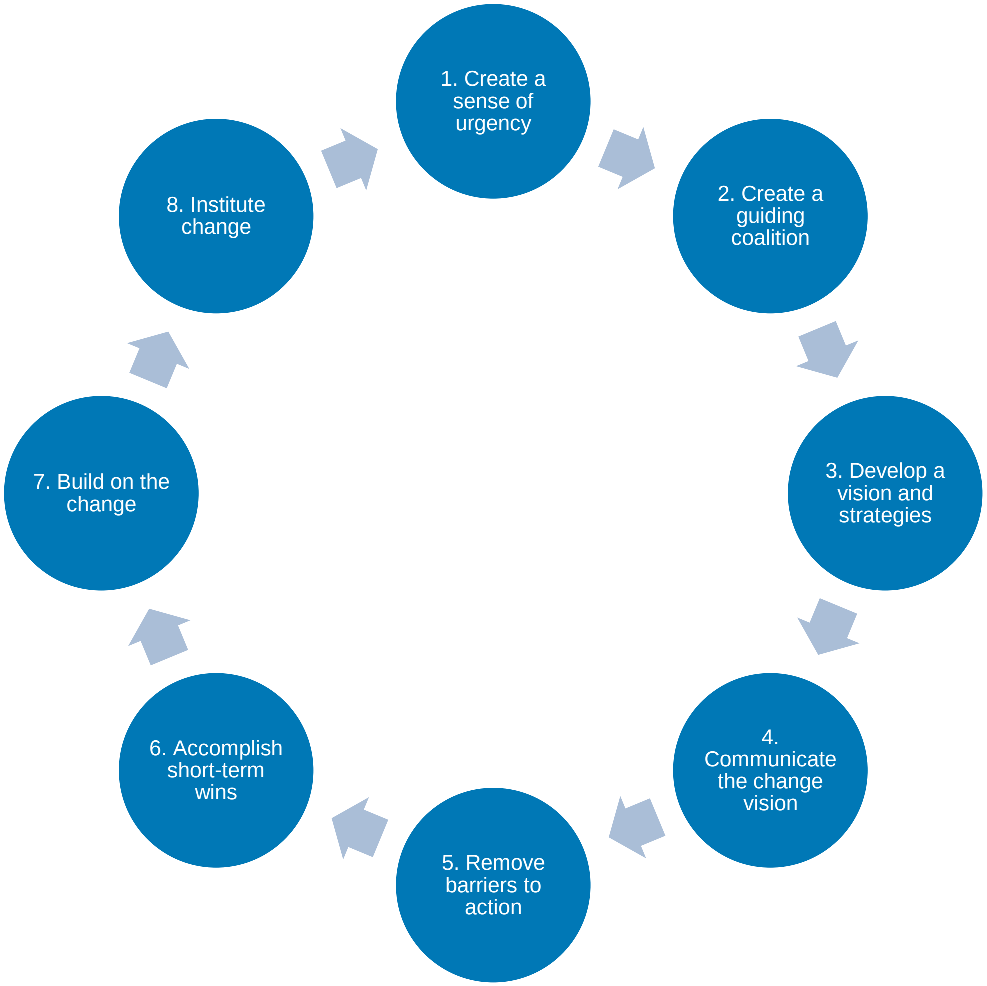
Strategies to Support Rapid Change

- Techniques
  - Make tools available for teams
  - Use available tools and technology
  - Apply change models and concepts
  - Replace old ways with new customs/norms
  - Use training to reinforce the change
  - Put procedures in place to reinforce the change
  - Link promotion and pay to desired behaviors
  - Simplify, standardizes, and uses technology

Prochaska's Stages in Changing Behavior



Kotter's Eight-stage Process for Major Change



Kotter's Heart of Change Model

- Pitfalls in implementing Kotter's model
  - Allowing too much complacency
  - Failing to create a sufficiently powerful guiding coalition
  - Underestimating the power of vision
  - Under communicating the vision
  - Permitting obstacles to block the new vision
  - Failing to create short-term wins
  - Declaring victory too soon
  - Neglecting to anchor changes firmly in the corporate culture

Change Management Models

Common elements of change management models include:

- A description of movement from a current state to a future state
- The impetus for transition
    - This may be dissatisfaction with current state, a need to reduce cost or increase revenue, a re-allocation of resources, **changes in regulations**, or other outside influences
- Factors to identify readiness for change
- Communication strategies for change
- Leadership characteristics for change
- The approach to planning for change
- Characteristics of the change and factors of how this makes it more or less likely to be accepted

Force Field Analysis Change Model

DRIVING FORCES

RESTRAINING FORCES

Example: Proposed change is to allow families 24-hour visiting access for patients in the intensive care unit.

Driving forces:

- Families provide comfort and reassurance to patients during ICU stays. Long periods without family may increase stress.
- Patients have a right to have family present during illnesses.
- Families have variable work schedules and cannot always meet the hospital's schedule.

Restraining forces:

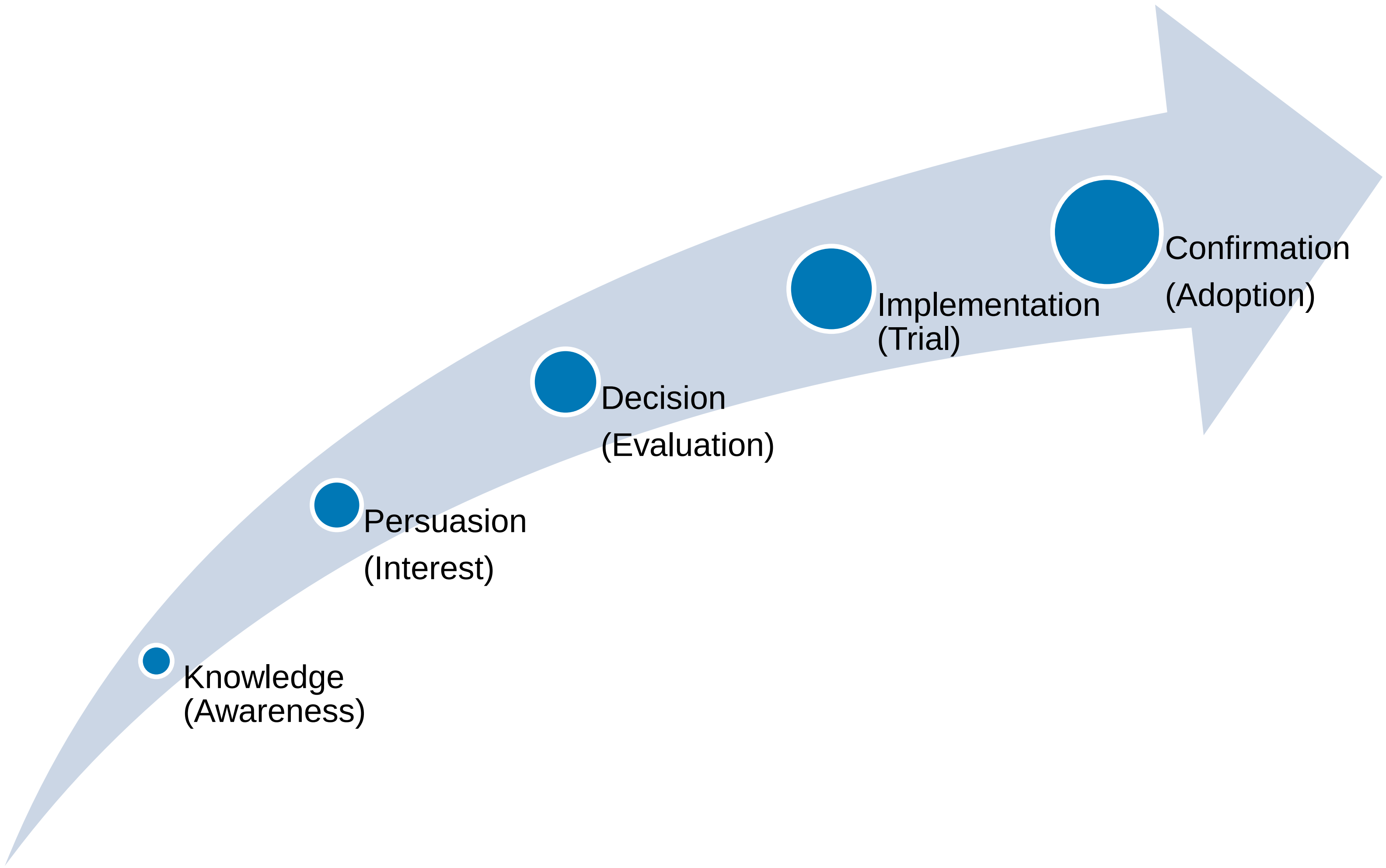
- Chief medical officer finds the open visiting policy disruptive to patient care.
- Nursing staff finds the open visiting policy disruptive to nursing routines and getting their work done.
- The open visiting hours will tire patients and not all sufficient rest.
- More families will stay overnight and crowd waiting rooms, sleeping on chairs, and couches.
- More families present in the hospital throughout the night increases security risks for the hospital.

Force Field Analysis Change Model

- CQI education program

STRENGTHS	BARRIERS
Management supports program	Lack of time
Physicians support program	Lack of commitment of all participants
Have committed time and other resources to complete program	Viewed as not important

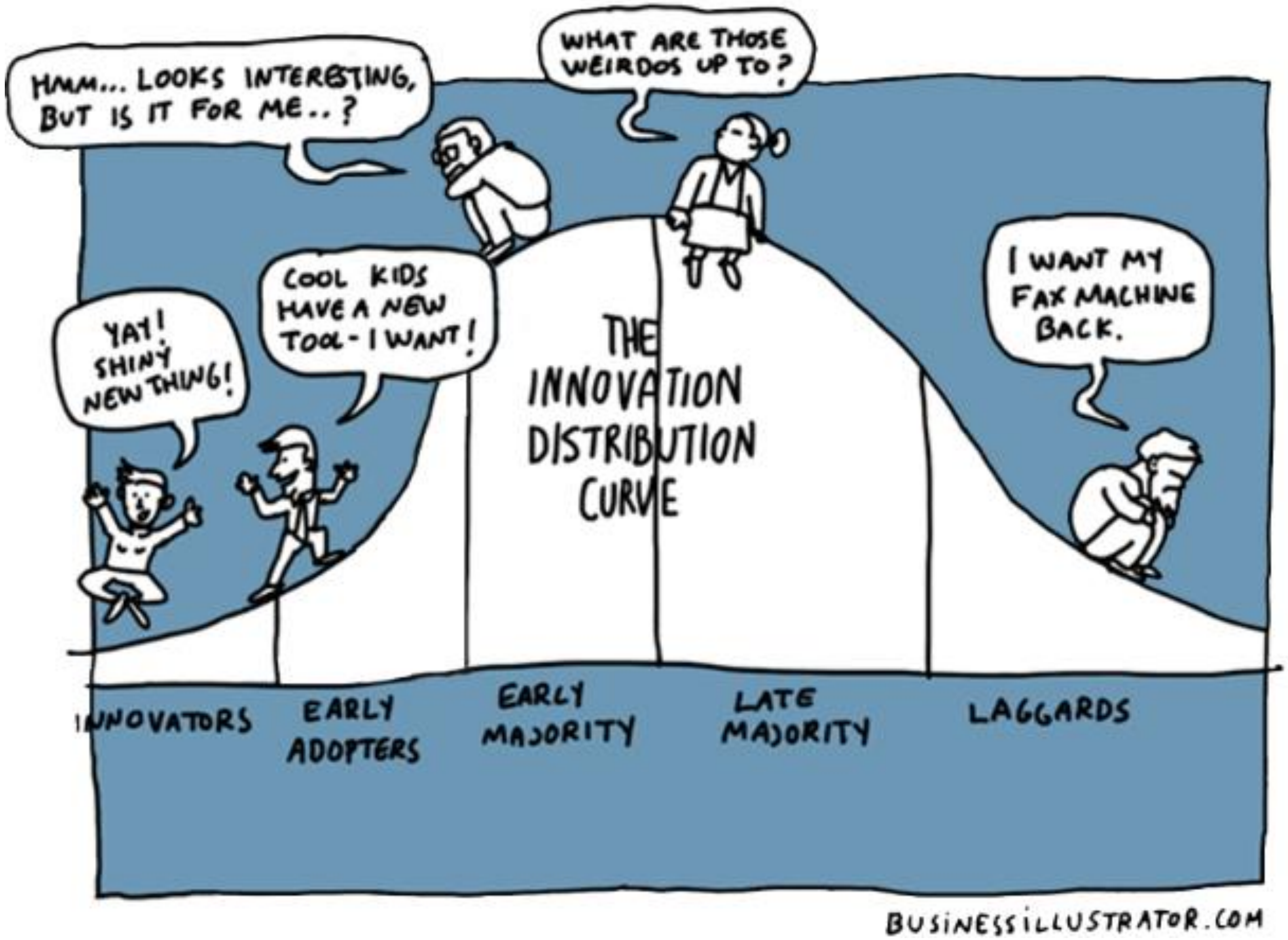
Diffusion of Innovation Model



Diffusion of Innovation Model

- Knowledge (Awareness)
  - Socioeconomic characteristics, personality variables, communication behavior.
- Persuasion(Interest)
  - Attending to perceived characteristics of innovation such as relative advantage, compatibility, complexity, visibility, and uncertainty.
- Decision(Evaluation)
  - Adoption or rejection.
- Implementation (Trial)
  - Direct application, reinvention.
- Confirmation(Adoption)
  - Evaluation of effectiveness.

Change Adoption



Reducing Resistance

SITUATION	STRATEGIES
Fear of Change	Clearly identify effect of change effort and communicate openly with those involved
Not willing to make the change (who)	Set goals, measure performance, provide coaching and feedback, reward and recognize positive efforts
Not able to perform the change (how)	Provide education and training in new skills and use various management techniques
Lack of necessary knowledge to make the change (why)	• Communicate what, why, how, when, and who of change process; present positive outlook • Have clear focus and goal for change for expectations • Be adaptable • Use structured approach to manage ambiguity and confusion • Plan and coordinate change • Use proactive rather than reactive approach

Outcomes of Successful Change

- Check your items off the checklist.
- Think of what factors are or are not present in your organization
  - Leadership systems are designed for results
  - Strategy is simple, aligned, and deployed
  - Design of organizational culture is intentional
  - Mission and vision are clearly understood
  - Rapid response is employed
  - Desired results are defined, measured, and aligned
  - Decisions are based on sound data
  - Customer focus is the foundation
  - Measurement is deployed at all levels
  - Organizational learning is valued
  - Human resource practices support culture
  - Employees are involved
  - Innovation is valued
  - Partnerships are created
  - Improvements are integrated into daily work
  - Focus is on improving employee knowledge

Chapter 2:
Multiple Regression

Patient Discharge Time – Recap

Find out if patient discharge summary delay impacting the discharge time. The data is given in Anexas hospital case study file.

Discharge summary delay	Discharge Time
100	300
45	245
160	360
85	280
95	300
130	330
120	320
110	310
95	300
100	302
90	287
100	294
110	310
125	320
105	300

What would you do?

What questions might you ask?

Data in ***Anexas Hospital study .mtw***

Patient Discharge Time

- Practical Problem
- Is there a relationship between discharge summary delay and discharge time?
- If there a relationship, how strong is that relationship?

- Statistical Problem
- Are the variables correlated?
- Null hypothesis: Discharge Summary delay and Discharge time are not correlated
- Alternate hypothesis: Discharge Summary delay and Discharge time are correlated

Patient Discharge Time

State the Hypotheses and Significance Level

H_o: $\rho = 0$

H_a: $\rho \neq 0$

$\alpha = 0.05$

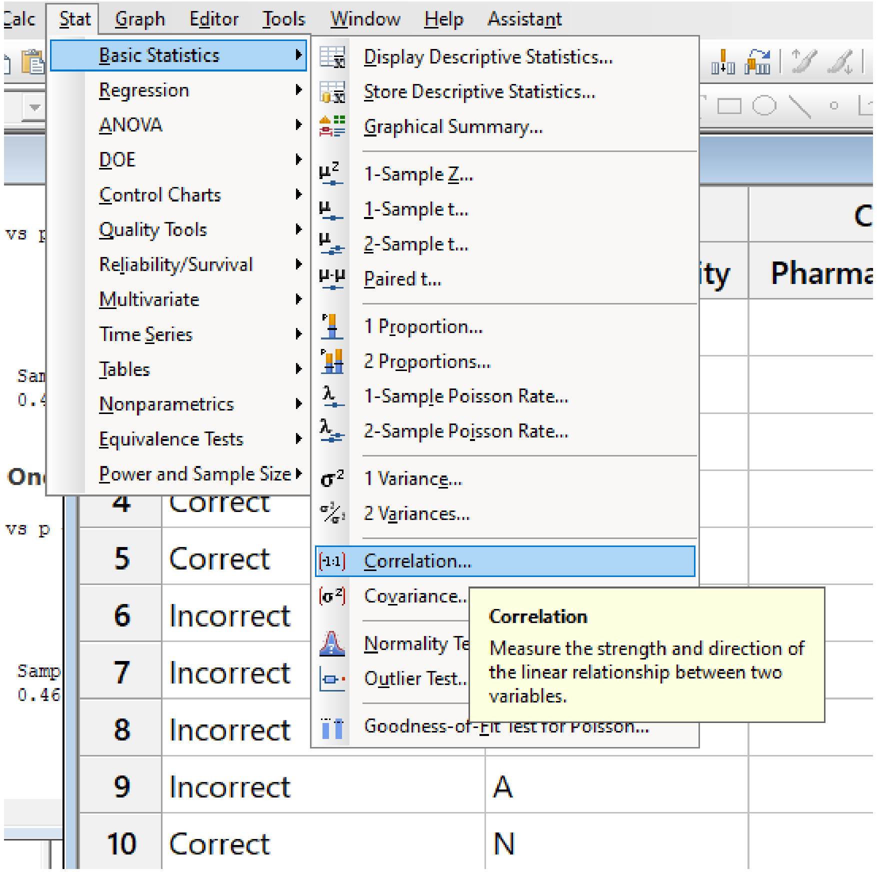
Notice that the hypotheses are about a population parameter

What Hypothesis Test is Appropriate?

- These hypotheses deal with correlation coefficient
- Make decisions based on Pearson correlation coefficient and ‘p-value’

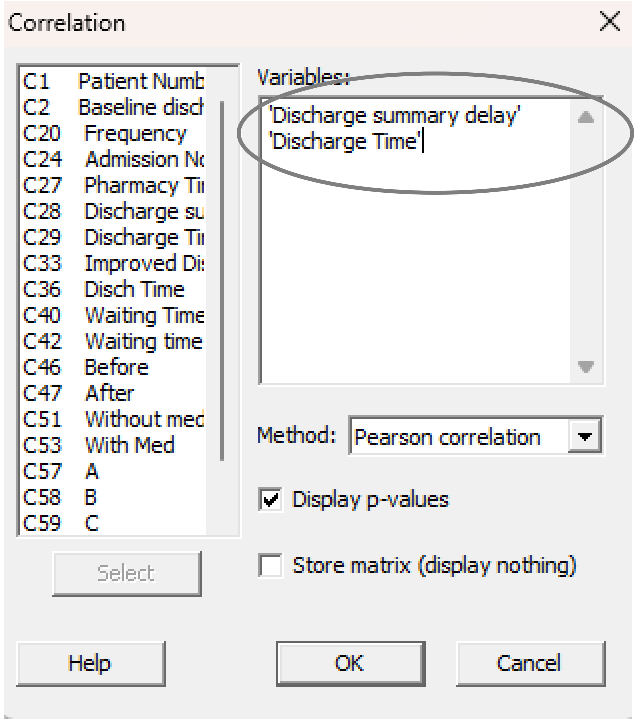
Patient Discharge Time

Tool Bar Menu > Stat > Basic Statistics > Correlation



Patient Discharge Time

Tool Bar Menu > Stat > Basic Statistics > Correlation



Patient Discharge Time

Correlation: Discharge summary delay, Discharge Time

```
Pearson correlation of Discharge summary delay and Discharge Time = 0.991
P-Value = 0.000
```

- What is the Decision?

- Pearson correlation or correlation coefficient for the sample, r = 0.991
- Does that mean 'ρ' is greater than zero? Or could it be that r = 0.991 due to chance variation while 'ρ' is still zero?
- Answer this question using p value

- What about
- the p-value?

Patient Discharge Time

- What is the statistical interpretation?
- p-value (0.000) < α-risk (0.01): reject the null hypothesis
- Infer H_a: sufficient evidence that there is a correlation between discharge summary delay and discharge time

Regression

Correlation and Regression

- Correlation tells how much linear association exists between two variables
- The Regression process creates a line that best resembles the relationship between process input and output and also provides an equation describing the nature of relationship.

Different types of Mathematical Models

- The regression process can fit several different types of line, since the linear relationship won't be applicable to all situations. The alternatives are

Correlation and Regression

- Correlation tells how much linear association exists between two variables
- Regression provides an equation describing the nature of relationship

Regression Terminology

- Response Variable
  – This is the uncontrolled variable - also known as dependent variable, output variable or Y variable
- Regressor Variable
  – Response depends on these variables - also known as independent variables, input variables, or X variables
- Noise Variable
  – Input variables (X) that are not controlled in the experiment
- Regression Equation
  – Equation that describes the relationship between independent variables and dependent variable
- Residuals
  – Difference between predicted response values and observed response values

Regression Objectives

- Determination of a Model
  - Explore the existence of relationship
- Prediction
  - Describe the nature of relationship using an equation and use the equation for prediction
- Estimation
  - To assess the accuracy of prediction achieved by the regression equation
- Determination of KPIV (x factor)
  - Screen variables and determine which variable has the biggest impact on the response variable

Anexas Consultancy Services

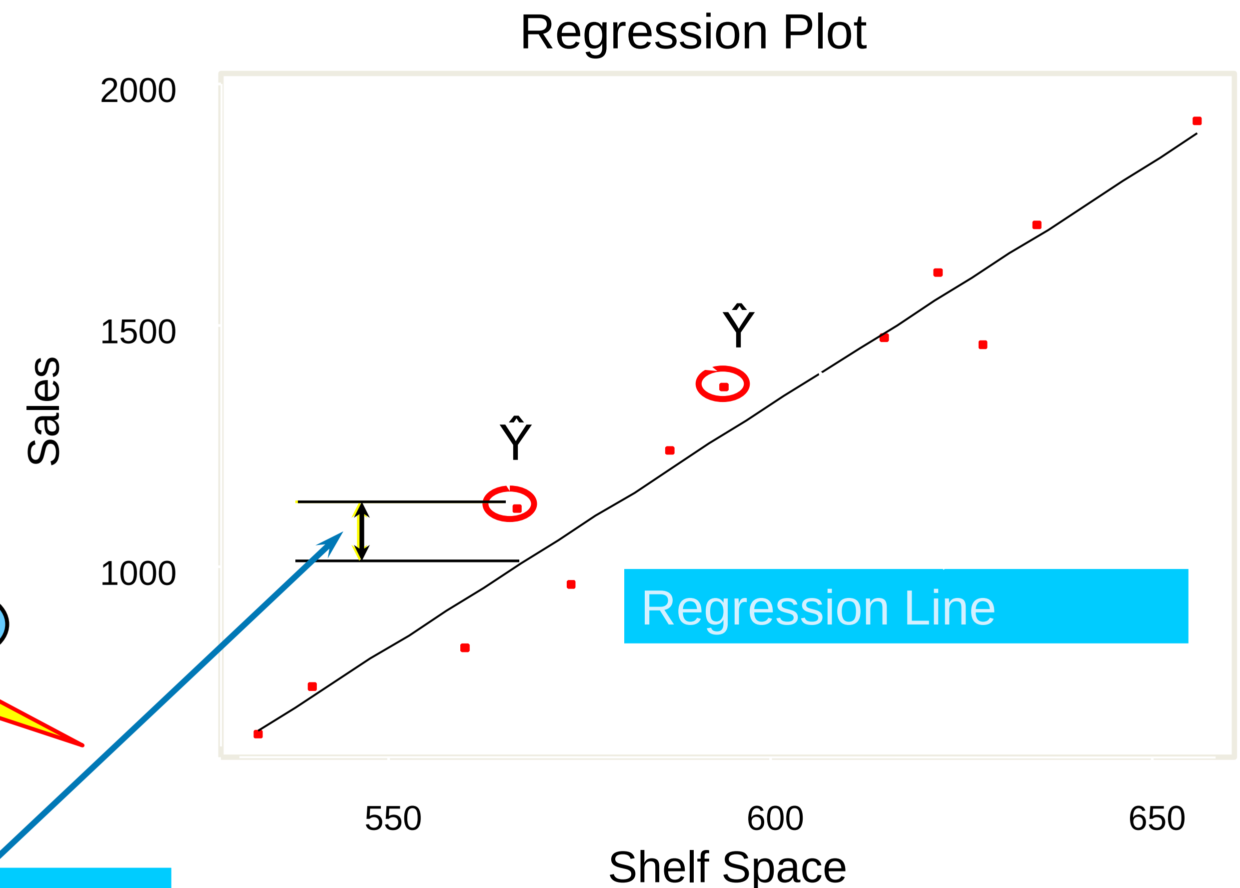
Simple Linear Regression

Method of Least Squares

Objective:

Find a line that will minimize sum of squares of residuals

Residuals are the error of prediction



Residual = $Y - \hat{Y}$

Patient Discharge Time

A Black Belt is interested in predicting the discharge time generated from information on patient discharge summary delay. As a result she conducted a study and collected data from 15 patients. Find out if patient discharge summary delay is impacting the discharge time. The data is given in Anexas hospital case study file

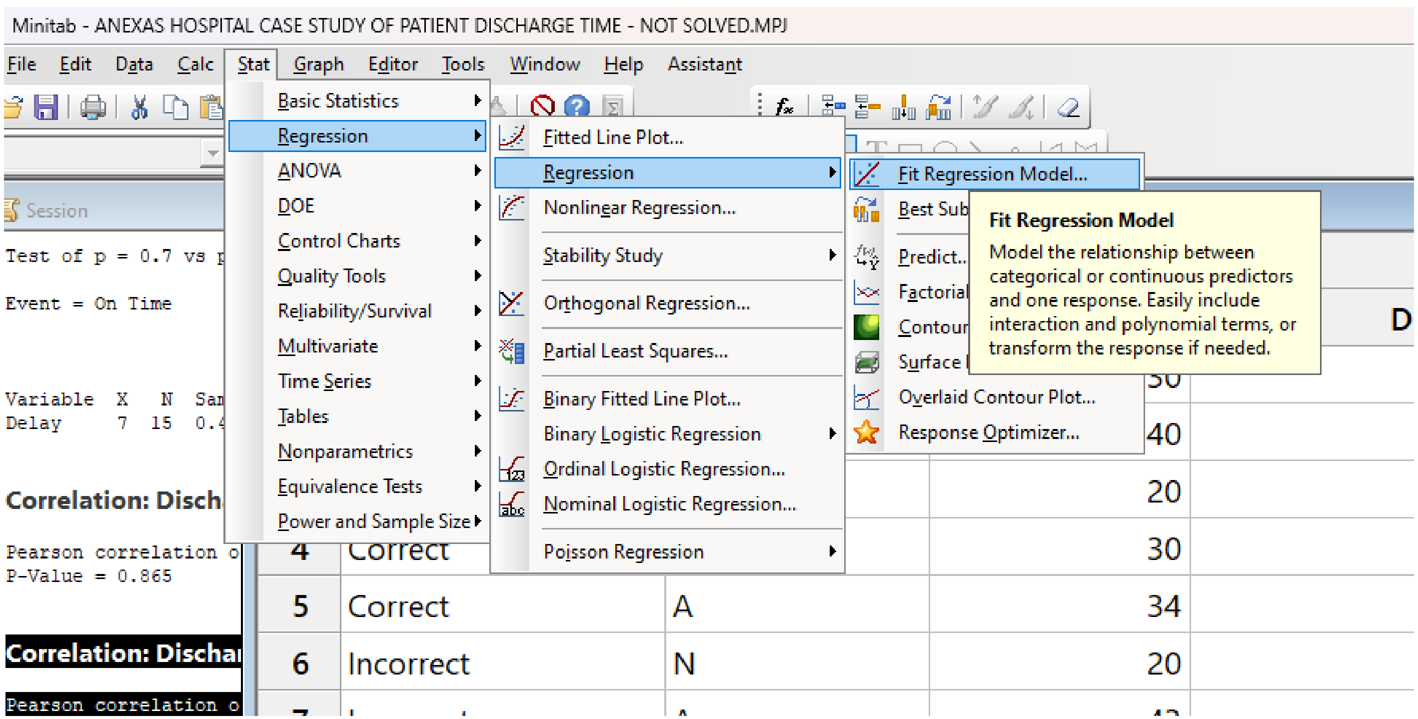
Discharge summary delay	Discharge Time
100	300
45	245
160	360
85	280
95	300
130	330
120	320
110	310
95	300
100	302
90	287
100	294
110	310
125	320
105	300

How will we create a simple linear regression model for the two variables?

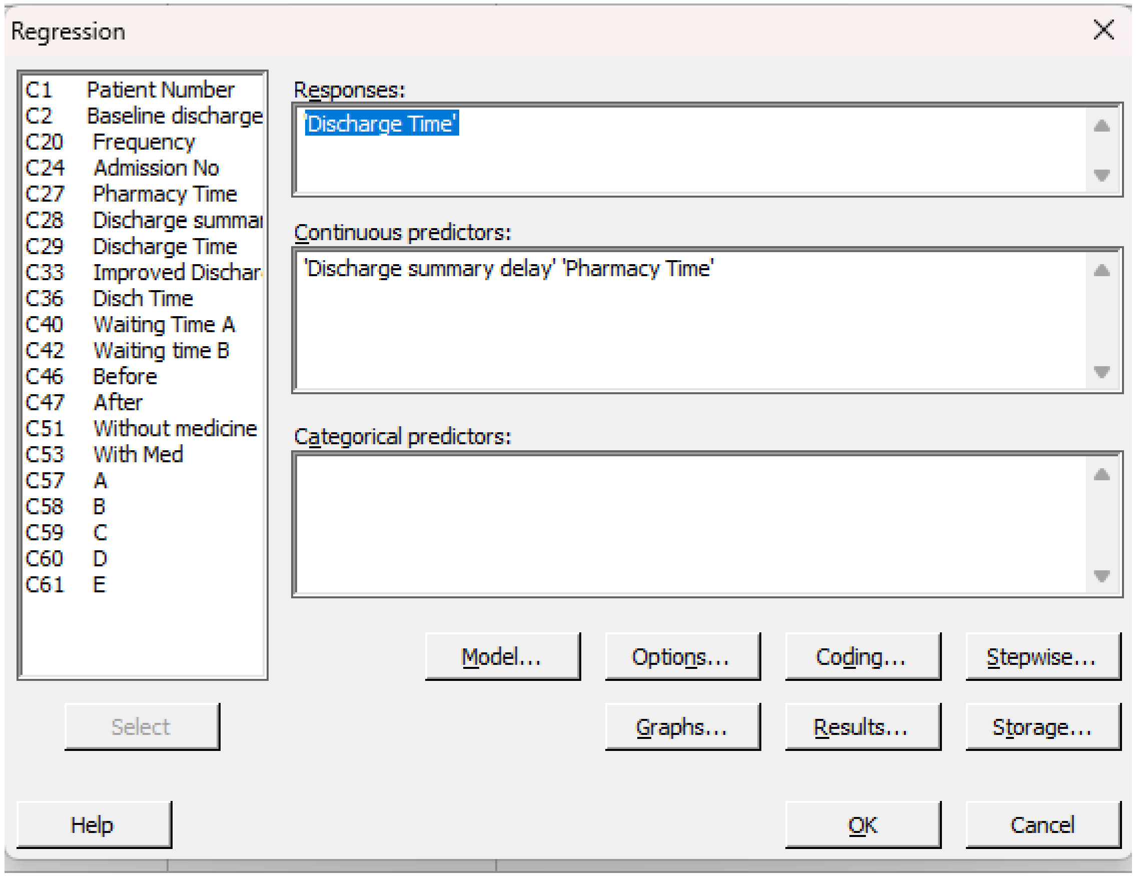
Data in ***<u>Anexas Hospital study .mtw</u>***

Patient Discharge Time

Tool Bar Menu > Stat > Regression > Fit Regression Model



Patient Discharge Time



Patient Discharge Time

Session Output from Minitab

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.69837	98.16%	97.85%	97.43%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	200.22	7.97	25.12	0.000	
Discharge summary delay	0.9937	0.0453	21.95	0.000	1.32
Pharmacy Time	-0.010	0.126	-0.08	0.936	1.32

Discharge summary delay is significant

Regression Equation

Discharge Time = 200.22 + 0.9937 Discharge summary delay - 0.010 Pharmacy Time

What About R-squared?

- R-squared is a measure describing the quality of regression
- Measures the proportion of variation that is explained by the regression model
- $R^2 = SS_{regression} / SS_{total} = (SS_{total} - SS_{error}) / SS_{total} = 1 - [SS_{error}/SS_{total}]$

Source	DF	SS	MS	F	P
Regression	1	1709656	1709656	224.511	0.000
Error	10	76150	7615		
Total	11	1785806			

$R^2 = 1709656 / 1785806 = 95.74\%$

In this example, 95.7% of variation in Y can be explained by variation in X

Example: Patient Discharge Time

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.55428	98.16%	98.01%	97.78%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	199.66	4.07	49.11	0.000	
Discharge summary delay	0.9956	0.0378	26.31	0.000	1.00

Regression Equation

Discharge Time = 199.66 + 0.9956 Discharge summary delay

Assumptions for Regression

Tool Bar Menu > Stat > Regression > Regression

- To use the results of regression, assumptions about residuals must be satisfied
- What are the assumptions about residuals?
  - Residuals are normally distributed with mean of zero
  - Residuals show no pattern (random)
  - Residuals have constant variance (homogeneous variance or no heteroscedasticity)
  - Residuals are independent of the values of regressor (x) variables
  - Residuals are independent of each other

Assumptions for Regression

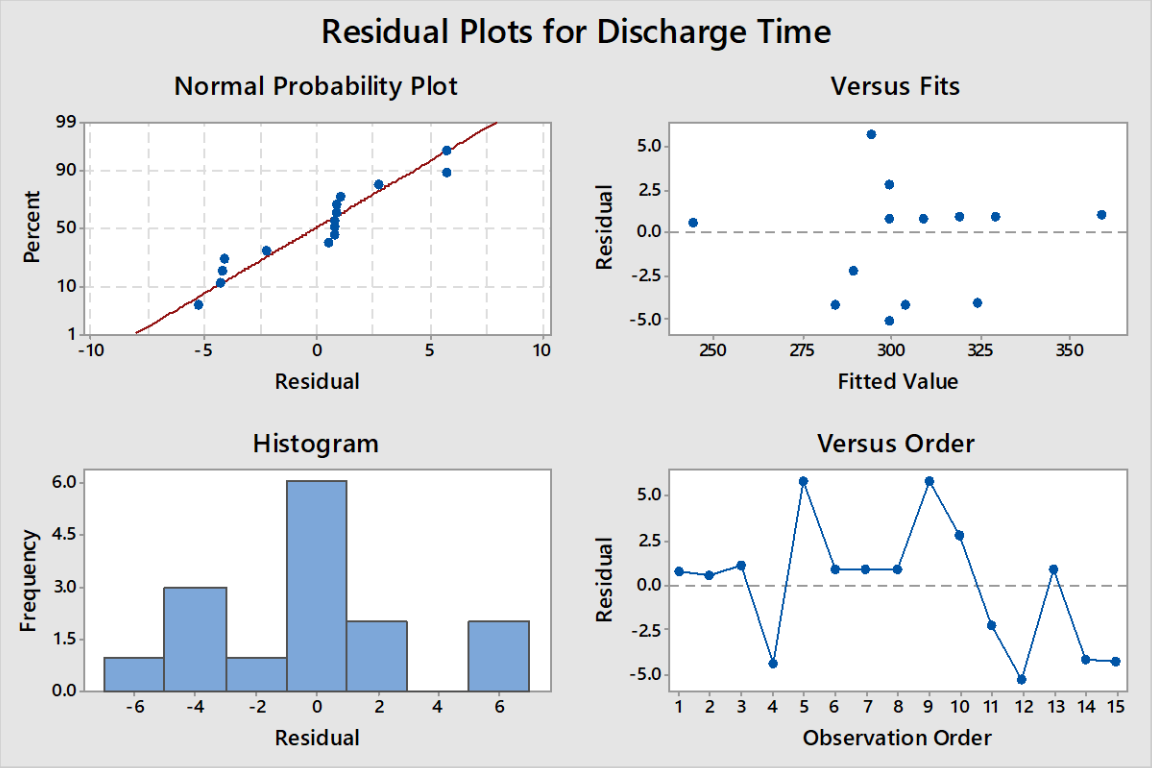
Residuals are error in the fit of regression line

- Difference between the observed value of response variable and fitted value

Residuals are normally distributed around mean of zero

Histogram of residuals resemble a normal distribution

with mean of zero



Residuals do not exhibit heteroscedasticity (they have homogenous variance)

Residuals are randomly distributed

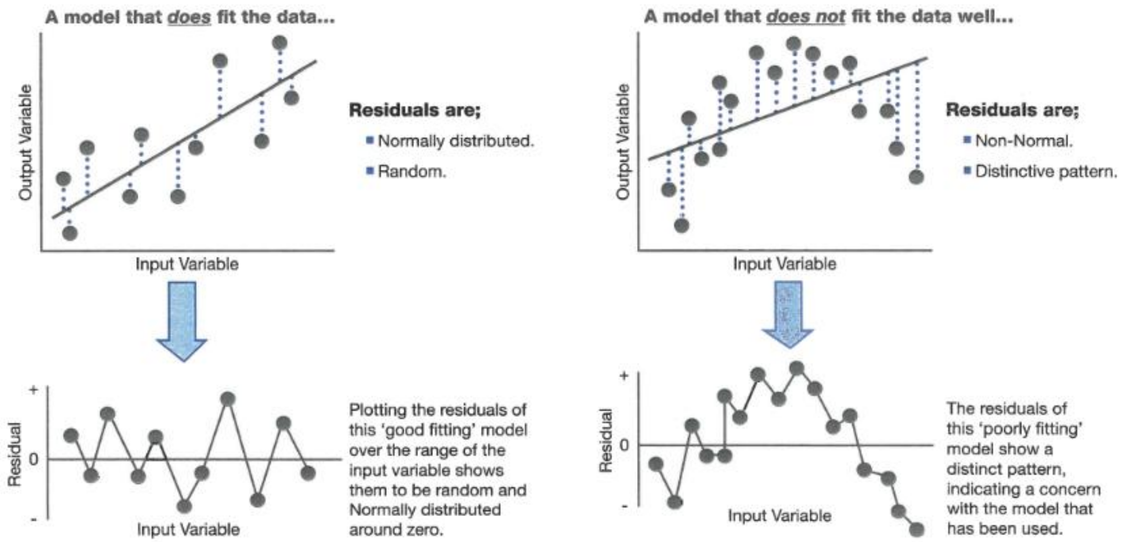
Residuals are independent of X variable

Fit Regression Model
Graphs - Four-in-One Residual Plot

- The four-in-one residual plot displays four different residual plots together in one graph window. This layout can be useful for comparing the plots to determine whether your model meets the assumptions of the analysis. The residual plots in the graph include:
- Histogram - indicates whether the data are skewed or outliers exist in the data
- Normal probability plot - indicates whether the data are normally distributed, other variables are influencing the response, or outliers exist in the data
- Residuals versus fitted values - indicates whether the variance is constant, a nonlinear relationship exists, or outliers exist in the data
- Residuals versus order of the data - indicates whether there are systematic effects in the data due to time or data collection order

Checking the Model – Analysis of Residuals

- Confirming that your regression model is a reasonable fit can be done visually with simple, regression, because you are able to see the fitted line plot. However this is not possible wit Multiple regression since it is more than two dimensional. Instead, the residual errors between the model and the data points can be analysed in order to decide if the model is a good for the data. This possible because the behavior of residuals for a gods fitting is well established as described below.



Module Objectives

By the end of this module participant will be able to:

- Determine, for a given response variable, the key process input variables from a set of multiple input variables
- Perform multiple linear regression for a given set of response variable using several input variables
- Perform model diagnostics and validate assumptions
- Use regression model to predict the value of a response variable for given values of predictor variables

Why Learn Multiple Regression?

- Explore the existence of relationship between a dependent variable and several independent variables
- Screen multiple input variables and determine which variables have the biggest impact on the response variable
- Describe the nature of relationship with an equation and use it for prediction

What is Multiple Regression?

- Procedure of establishing relationship between a continuous type response variable and two or more independent variables
- Multiple regression equation can be used to predict a response based on values of predictor variables
- Multiple regression equation takes the form

$$Y = f\,(x_1, x_2, x_3, \ldots.)$$

Types of Multiple Regression

Multiple Linear Regression

Multiple regressor (x) variables such as x_1, x_2, x_3 and model linear with respect to coefficients

Example1: $y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + error$

Example2: $y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_2^2 + error$

Multiple Non-Linear Regression

Multiple regressor (x) variables such as x_1, x_2, x_3 and model non-linear with respect to coefficients

Example: $y = (a_0 + a_1 x_1) / a_2 x_2 + a_3 x_3 + error$

This module focuses on multiple linear regression applying general least squares method

Multicollinearity

- A condition in which two or more independent variables (x variables) are correlated (pairwise and more complex linear relationships)
- When used in multiple regression model, they contribute to redundant information
- For example, fuel economy of a truck = f (truck load, engine horse power)
- But truck load may be correlated with engine horse power
- Truck load and horse power provide some overlapping information leading to potential problems

Problems Due to Multicollinearity

- Multicollinearity can cause severe problems
  - calculations of coefficients and standard errors are affected (unstable, inflated variances)
  - difficulty in assessing any particular variable's effect
  - opposite signs (from what is expected) in the estimated parameters
  - if two input variables x_1 and x_2 are highly correlated, then p-value for both might be high

Detecting Multicollinearity

- High values of pairwise correlation (generally > 0.8) provide warnings of potential multicollinearity problems
- If the above two variables are strongly correlated, one of them should be removed from regression model

Variance Inflation Factor

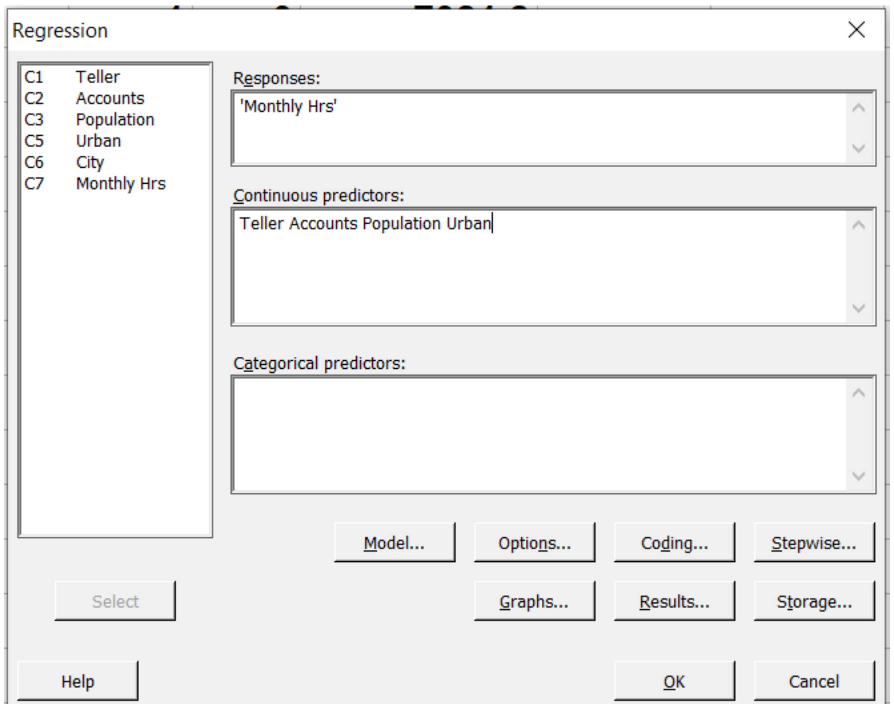
- A metric, called variance inflation factor (VIF) calculates the degree of multicollinearity

$$VIF = \frac{1}{1 - R_i^2}$$

- R_i^2 is the R^2 value obtained when X_i is regressed against other X
- A large VIF implies that at least one variable is redundant
- VIF > 10: high degree of multicollinearity - cause for serious concern ($R_i^2 > .9$)
- VIF > 5: moderate degree of multicollinearity ($0.8 < R_i^2 < 0.9$)
- Guideline: Ensure that VIF < 5 when possible

Calculating VIF

Minitab displays VIF values in the session window through *Stat > Regression > Regression> Fit Regression Model* menu



Regression Analysis: Monthly Hrs versus Teller, Accounts, Population, Urban

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	34308448	8577112	5345.10	0.000
Teller	1	2719647	2719647	1694.83	0.000
Accounts	1	41	41	0.03	0.876
Population	1	8948965	8948965	5576.83	0.000
Urban	1	123813	123813	77.16	0.000
Error	12	19256	1605		
Total	16	34327704			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
40.0583	99.94%	99.93%	99.90%

Coefficients

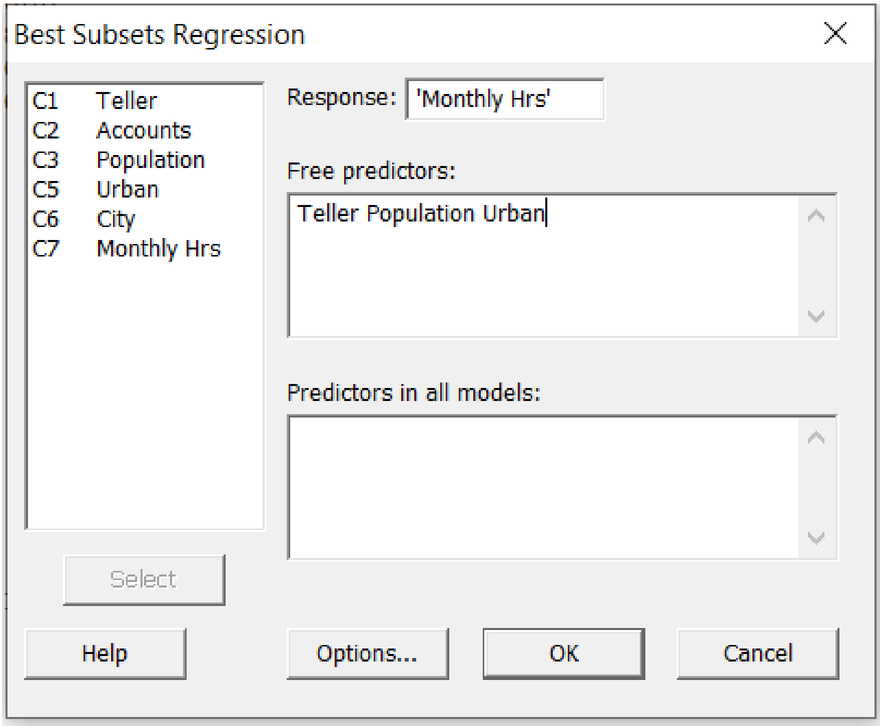
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1094	148	7.41	0.000	
Teller	0.9510	0.0231	41.17	0.000	10.59
Accounts	0.0091	0.0571	0.16	0.876	8.31
Population	0.056381	0.000755	74.68	0.000	1.04
Urban	238.5	27.2	8.78	0.000	1.95

Predictor Variable Selection

- What combination of predictor variables is best for the regression model?
- Three options in Minitab:
  - Stepwise: procedure to add and remove variables to the regression model to produce a useful subset of predictors
  - Best Subsets: procedure to give best fitting regression model that can be constructed with one variable, two variable, three variable, etc. models
  - Regression: once the best model is selected, use Regression to get more detailed diagnostics

Best Subsets

Tool Bar Menu > Stat > Regression > Regression> Best Subsets



Use 'Best Subsets' technique to select a group of likely models for further analysis.

Best Subsets Statistics

- Select the smallest subset that fulfills certain statistical criteria
- Minitab displays R^2, R^2 (adjusted), C-p, and s statistics
  - R^2 (large R^2 is desired; use to compare models with the same number of terms)
  - adjusted R^2 (large is desired; use to compare models with different number of terms)
  - s (standard deviation of error terms; small is desired)
  - Mallows' C-p statistic (small is desired; Guideline: want C-p $\leq$ number of terms in model)

Putting It All Together

Multiple regression objective: Establish a model with high prediction ability and minimum multicollinearity

Multiple regression steps:

1. Remove variables contributing to multicollinearity from the predictors
2. Use remaining variables and apply Best Subsets to evaluate best predictor candidates for the model
3. Choose the best candidate and complete regression analysis
4. Perform model diagnostics to identify outliers and unusual observations
5. Analyze residuals for violation of assumptions
6. Assess predictive capability using new observations

Example: Monthly Labor Hours

A hospital wants to produce an empirical equation that will estimate personnel needs in its branches. The following data was collected from its existing branches at various locations. The response variable (Y) for the study was monthly labor hours. The input variables were average number of average revenue (x1), average count of total patients enrolled(x2), location of branch (x3), population within 10 Km radius (x4). The data is recorded in the file ***Multiple Regression.mtw.***

Example: Labor Hours

- Establish a multiple regression model to predict monthly labor hours using the predictor variables
- Perform model diagnostics to detect any outliers or unusual observations
- Validate any assumptions used for creating the model

Example: Monthly Labor Hours

- Output variable: Monthly Hrs
- Input variables: Revenue, Patientss, Population, Location
- Practical questions about the data and process?

C1	C2	C3	C4-T	C5
Average Revenue	Patients	Population	Location	Monthly Hrs
2842.8	2967	58800	Urban	7381.8
2549.2	3243	65200	Urban	7465.0
2718.4	3519	70900	Urban	7925.3
3683.2	3818	77400	Urban	9265.2
5916.8	4439	79300	City	11282.6
5825.6	4347	81000	City	11156.2
4785.6	4025	71900	City	9714.0
5410.4	4278	63900	Urban	10123.0
5062.4	4370	54500	Urban	9269.0
5647.6	4301	39500	City	8746.2
6676.4	4485	44500	City	9963.6
5828.4	4738	43600	Urban	9346.3
5447.6	4554	56000	Urban	9734.0
5830.0	4416	64700	City	10390.2
5822.4	4393	73000	City	10773.6
7429.2	4600	78900	City	12673.2
6247.2	4600	79400	Urban	11776.1

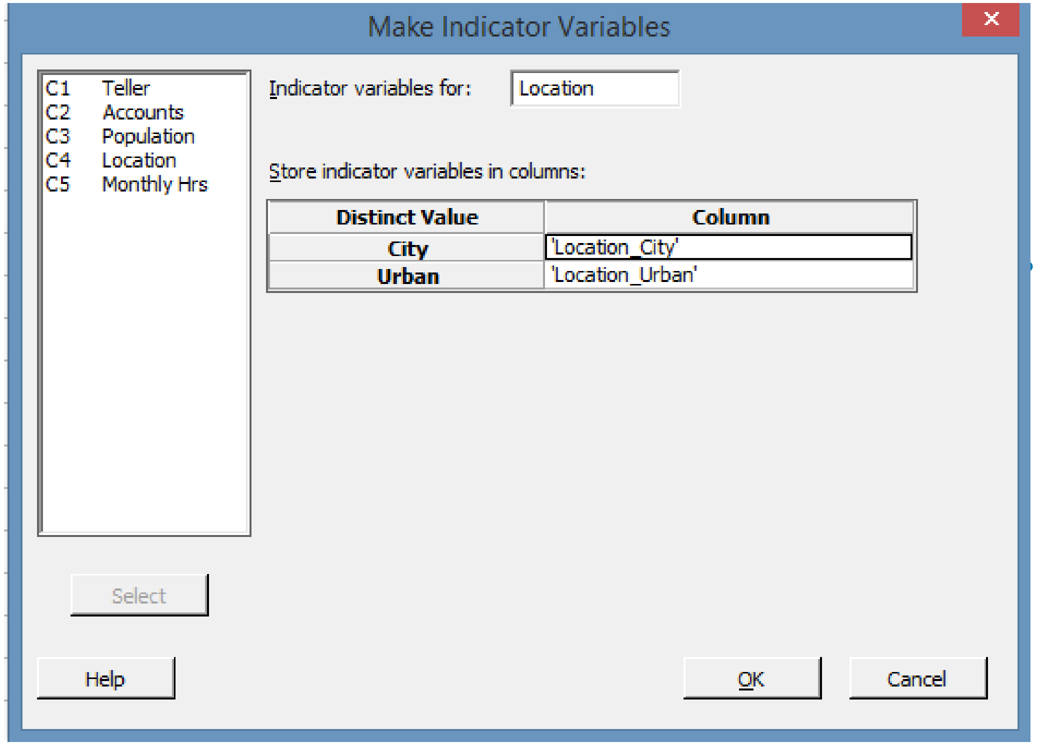
Example: Labor Hours

- ***Dealing with indicator variables (dummy variables)***

  - To use independent variables which are categorical (e.g, location, gender) in regression, first create “indicator” variables (dummy variables)
  - Indicators are simply 1’s and 0’s which are used like binary code.
  - Each level of a categorical variable is assigned a column.
  - If a row of data is associated with that level of the variable, the value in that column for that row of data will be 1. If not associated with that level, the value will be 0.

Example: Labor Hours

1. Create an indicator variable for “Location”
   - Toolbar>Calc> Make indicator Variables



Example: Labor Hours

- Result should follow the pattern below
- Order in which the columns are named is important. For alphanumeric data, the columns are created in the order specified as the value order (alphabetic, order of appearance, user defined).

Location	Urban	City
Urban	1	0
Urban	1	0
Urban	1	0
Urban	1	0
City	0	1
City	0	1
City	0	1
Urban	1	0
Urban	1	0
City	0	1
City	0	1
Urban	1	0

Example: Labor Hours

- ***Create an indicator variable for "Location"***
- When categorical variables are used in the regression model, one column of indicator values is NEVER included.
- In the example, use either "Urban" or "City" since it would be redundant to use both. (If "Urban"=0, then the observation must be "City")

Example: Labor Hours

- ***Multiple regression steps:***
- Remove variables contributing to multicollinearity from the predictors
- Identify if multicollinearity is a problem by using variance inflation factors (VIF) measurements
- 1. *Stat> Regression> Regression> Fit Regression Model*
- 2. Fill in the dialog as shown below:

Example: Labor Hours

Serious multicollinearity problem!
Want VIF <5

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1094	148	7.41	0.000	
Average Revenue	0.9510	0.0231	41.17	0.000	10.59
Patients	0.0091	0.0571	0.16	0.876	8.31
Population	0.056381	0.000755	74.68	0.000	1.04
Urban	238.5	27.2	8.78	0.000	1.95

Regression Equation

Monthly Hrs = 1094 + 0.9510 Average Revenue + 0.0091 Patients
+ 0.056381 Population
+ 238.5 Urban

Example: Labor Hours

b. Identify pairwise correlations between x variables

1. *Stat> Basic Statistics> Correlation..*
2. Fill in the dialog as shown below:
3: Press Ok

Correlation: Average Revenue, Patients, Population, Urban

	Average Revenue	Patients	Population
Patients	0.918		
Population	-0.024	-0.082	
Urban	-0.573	-0.372	-0.126

Example: Labor Hours

Next step: remove Revenue and check multicollinearity

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-2789	1303	-2.14	0.052	
Patients	2.184	0.247	8.85	0.000	1.18
Population	0.05735	0.00865	6.63	0.000	1.04
Urban	-456	244	-1.87	0.084	1.20

Regression Equation

Monthly Hrs = -2789 + 2.184 Patients + 0.05735 Population - 456 Urban

All VIF <5

Example: Labor Hours

- ***Multiple regression steps:***

2. Use remaining variables and apply Best Subsets to evaluate best predictor candidates for the model

1. Stat> Regression> Best Subsets..
2. Fill in the dialog as shown below:

3: Press Ok

Example: Labor Hours

Best Subsets Regression: Monthly Hrs versus Patients, Population, Urban

Response is Monthly Hrs

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	Patients	Population	Urban
1	59.4	56.7	49.6	53.2	964.18	X		
1	25.7	20.8	4.9	108.0	1303.6			X
2	89.9	88.4	81.6	5.5	498.22	X	X	
2	65.0	60.0	49.0	46.0	926.06	X		X
3	92.0	90.2	84.0	4.0	459.00	X	X	X

Example: Labor Hours

```
Regression Equation

Monthly Hrs = -2789 + 2.184 Patients
+ 0.05735 Population - 456 Urban
```

Chapter 3:
Design of Experiments

Module Objectives

By the end of this module, the participant will:

- Understand the strategy behind Design of Experiments (DOE)
- Create a design matrix
- Code factors
- Understand the limitations of OFAT experiments
- Interpret interactions

Why Learn About Design of Experiments?

- Properly designed experiments will improve
- Efficiency in gathering information
  - Planning
  - Resources
- Predictive knowledge of the process
- Ability to optimize process
  - Response
  - Control input costs
- Capability of meeting customer CTQs

What is Design of Experiments ?

- A DOE is a planned set of tests on the response variable(s) (KPOVs) with one or more inputs (factors) (PIVs) each at two or more settings (levels) which will:

- Determine if any factor is significant
- Define prediction equations
- Allow efficient optimization
- Direct activity to rapid process improvements
- Create significant events for analysis

DOE Terminology

- Response (Y, KPOV): the process output linked to the customer CTQ
- Factor (X, PIV): uncontrolled or controlled variable whose influence is being studied
- Level: setting of a factor (+, -, 1, -1, hi, lo, alpha, numeric)
- Treatment Combination (run): setting of all factors to obtain a response
- Replicate: number of times a treatment combination is run (usually randomized)
- Repeat: non-randomized replicate
- Inference Space: operating range of factors under study

DOE Objectives

- Learning the most from as few runs as possible; efficiency is the objective of DOE
- Identifying which factors affect mean, variation, both or none
- Screening a large number of factors down to the vital few
- Modeling the process with a prediction equation:

  Y = f (A,B,C …)
- Optimizing the factor levels for desired response
- Validating the results through confirmation

Anexas Consultancy Services

DOE Strategy

Strategy of Experimentation

Planning Phase

- Define the problem
- Set the objective
- Select the response(s), Ys
- Select the factors(s), Xs
- Select the factor levels
- Select the DOE design
- Determine sample size
- Run the experiment

Execution Phase

- Collect data
- Analyze data
- Validate results
- Evaluate conclusions
- Revise SOPs
- Train affected workforce
- Document results
- Consider next experiment

Anexas Consultancy Services

Planning Phase

The purpose of an experiment is to better understand the real world, not to understand the experimental data…

William Diamond
Statistician, IBM

Define the Problem
Step 1

- The problem statement should:
- Be complete and detailed
- Contain no conclusions
- Contain no solutions
- Contain no causes
- Be specific

Well begun is half done…

Mary Poppins

Issues With Problem Definition

- Common problem definition issues
- Response not well defined
- Response not quantifiable
- Response does not link to CTQs
- Specifications not customer focused
- Quantification not based upon data
- Stated as a predetermined solution

Guide your team to unambiguous, executable experiments

Set the Objectives
Step 2

- The upfront objectives should
- Define what you want to learn from the experiment
- Support screening, modeling or both
- Look at main effects, interactions or both
- Set boundaries for optimization
- Know resource requirements

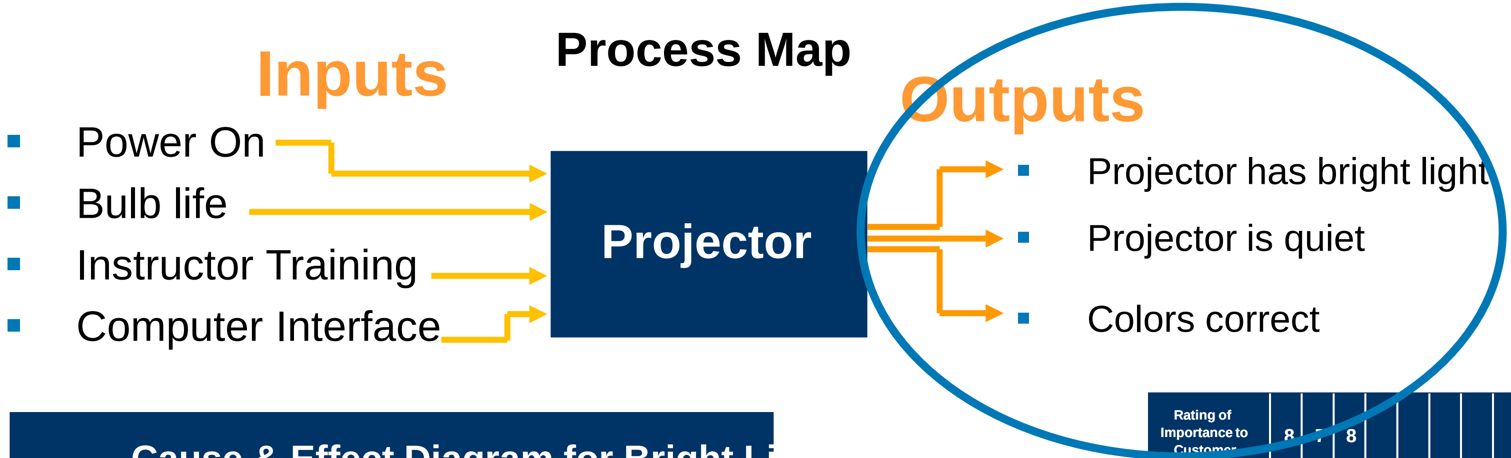
Do not use all of your resources on the first experiment

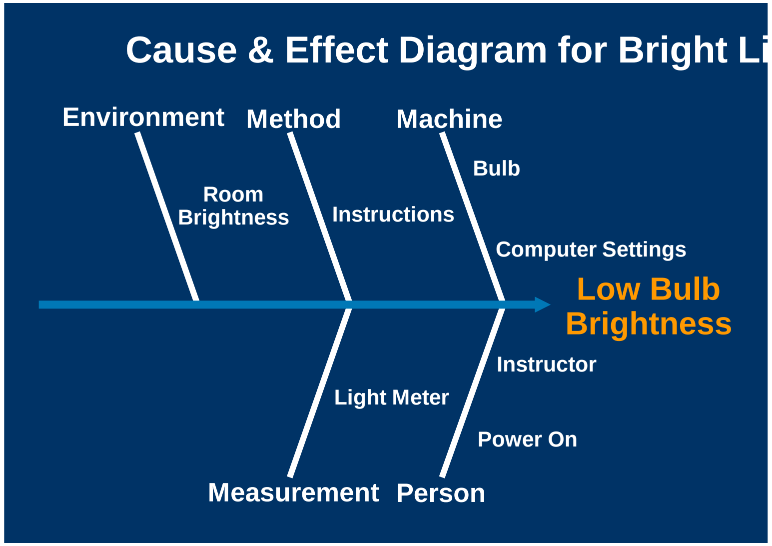
Selecting the Response
Step 3

A response is the output linked to the customer CTQ

- Usually quantitative (continuous)
  - Yield
  - Level of purity
  - Effort Variance
  - Energy consumed
- May be qualitative (discrete)
  - Bubbles in glass plate
  - Defective Coding
- Comes from Define and Measure Phase
  - Voice of the Customer
  - C&E, FMEA, process map and other

Response Relationship to Other Tools





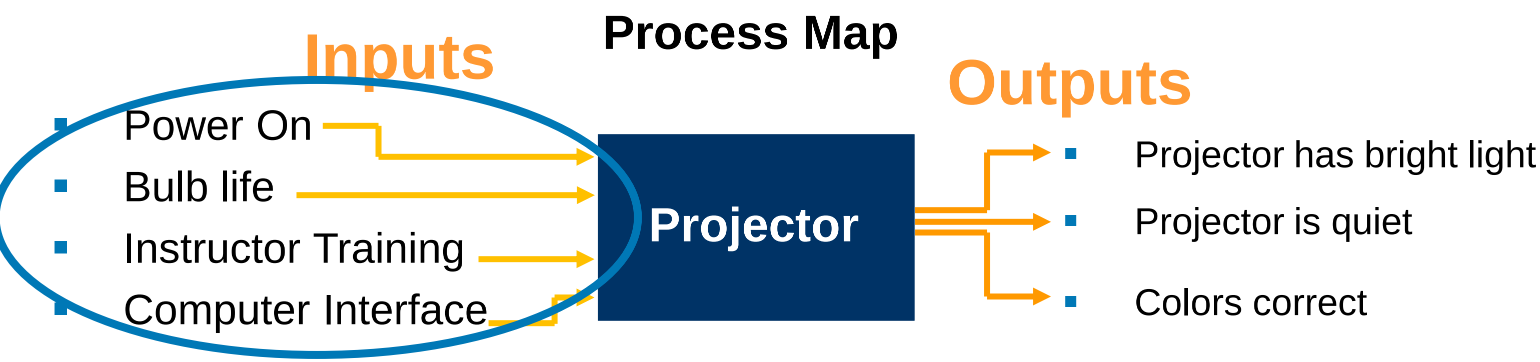
	Rating of Importance to Customer	8	7	8												
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	Process Inputs	Bright Light	Quiet	Color Correct													Total
1	Power on	9	3	2													119
2	Bulb Life	8	1	3													95
3	Instructor	1	1	1													23
4	Computer	2	1	9													95
5																	
6																	
7																	

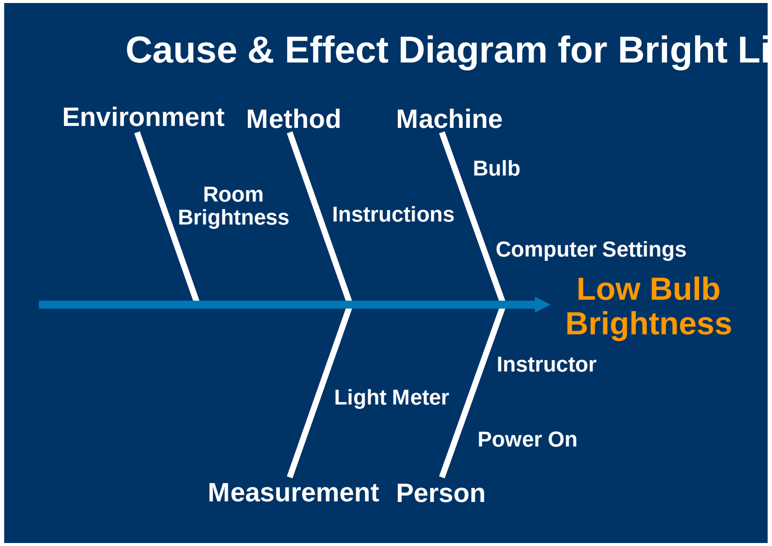
Selecting the Factors
Step 4

A factor is one of the inputs suspected of influencing the response

- May be qualitative (discrete)
  - Different operators
  - Different shifts
  - Different Resources
- May be quantitative (continuous)
  - Pressure
  - Pull back angle
- Comes from Measure and Analyze Phase
  - C&E, FMEA, process map and other
  - Hypothesis testing

Response Relationship to Other Tools





	Rating of Importance to Customer	8	7	8													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	Process Inputs	Bright Light	Quiet	Color Correct													Total
1	Power on	9	3	2													119
2	Bulb Life	8	1	3													95
3	Instructor	1	1	1													23
4	Computer	2	1	9													95
5																	
6																	
7																	

Selecting the Factor Levels
Step 5

The factor levels are the range of values of the input under study

- Qualitative (discrete) inputs
  - operator 1, operator 2, operator 3…
  - shift 1, shift 2, shift 3 …
  - pin loc 1, 2
- Quantitative (continuous) inputs
  - 50 , 75, 100 psi
  - 160, 170, 180 degrees
- Process knowledge/inference space

Consider what works, what could work and what won't hurt

Select the DOE Design
Step 6

The selected design should align with the objectives of the experiment and resource commitment

- OFAT (One Factor at a Time)
- 2k-n Fractional Factorial Designs
- 2k Full Factorial
- Full Factorial (replication)
- RSM (response surface method)

Experimental Design Considerations

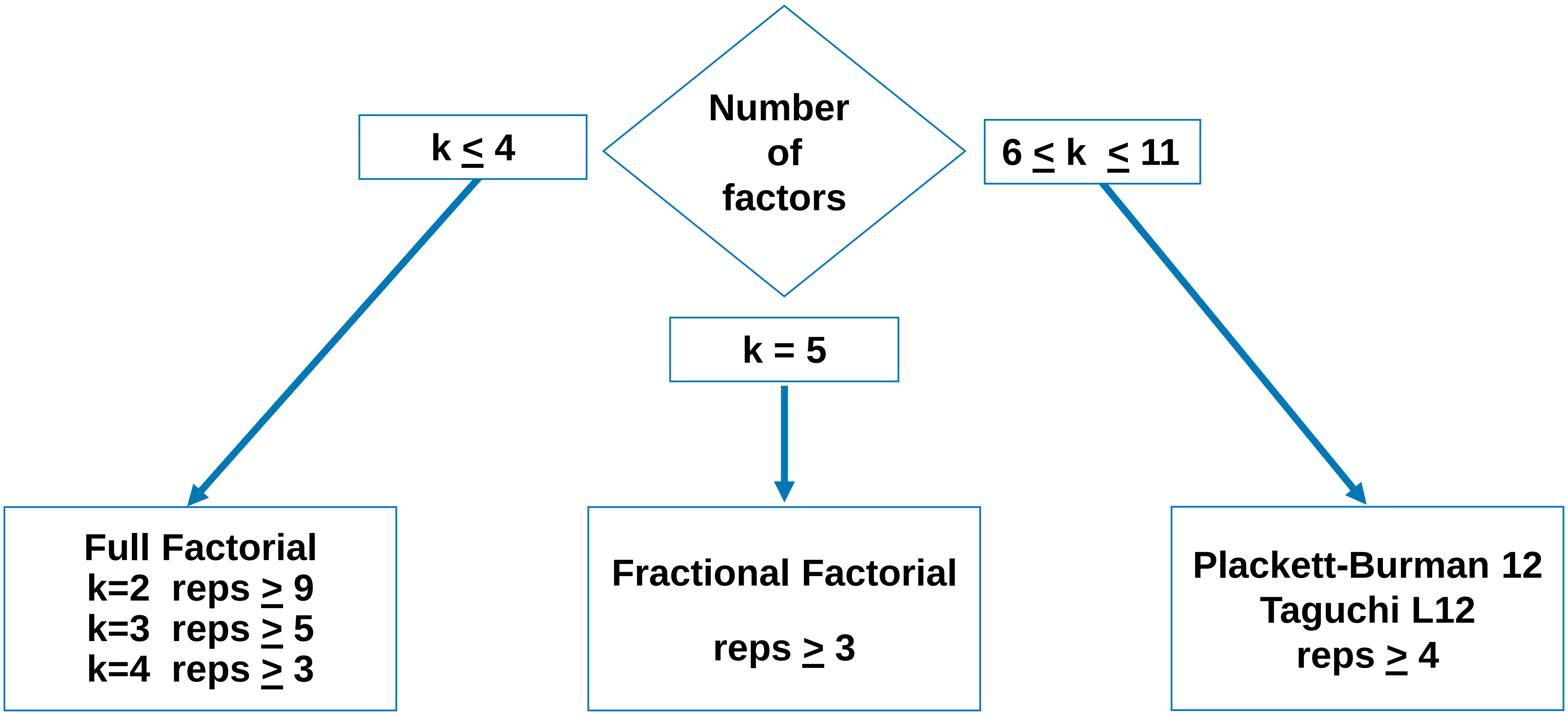
KPIV's	Experimental Objective	Experimental Design Types	Knowledge	Cost
Many	*Screen*	OFAT (One Factor at a Time)	Less	Less
		Fractional Designs		
		2^k Full Factorial		
	Optimize / *Model*	2^k Full Factorial (w center points/replication)		
		Full Factorial (replication)		
Few		RSM (Response Surface Method)	More	More

Higher complexity designs offer greater knowledge at a higher price

Determine Sample Size
Step 7

- The type of design determines minimum number of samples (replicates/repeats). Considerations for replicates
  - Multiple runs help identify data and measurement errors
  - Screening for main effects requires fewer replicates
  - Modeling designs require more replicates to have confidence in prediction models
  - Detectable effect
  - Resources

Two Level Sample Size



Provides greater than 99% confidence in mean and 95% confidence in variance prediction models

Run the Experiment
Step 8

- Manage resources
  - Assign tasks
  - Stage supplies
- Have contingency plan
  - Personnel changes
  - Business changes
  - Botched runs
- Statistical common sense
  - Randomize runs if possible
  - Replicate and/or repeat

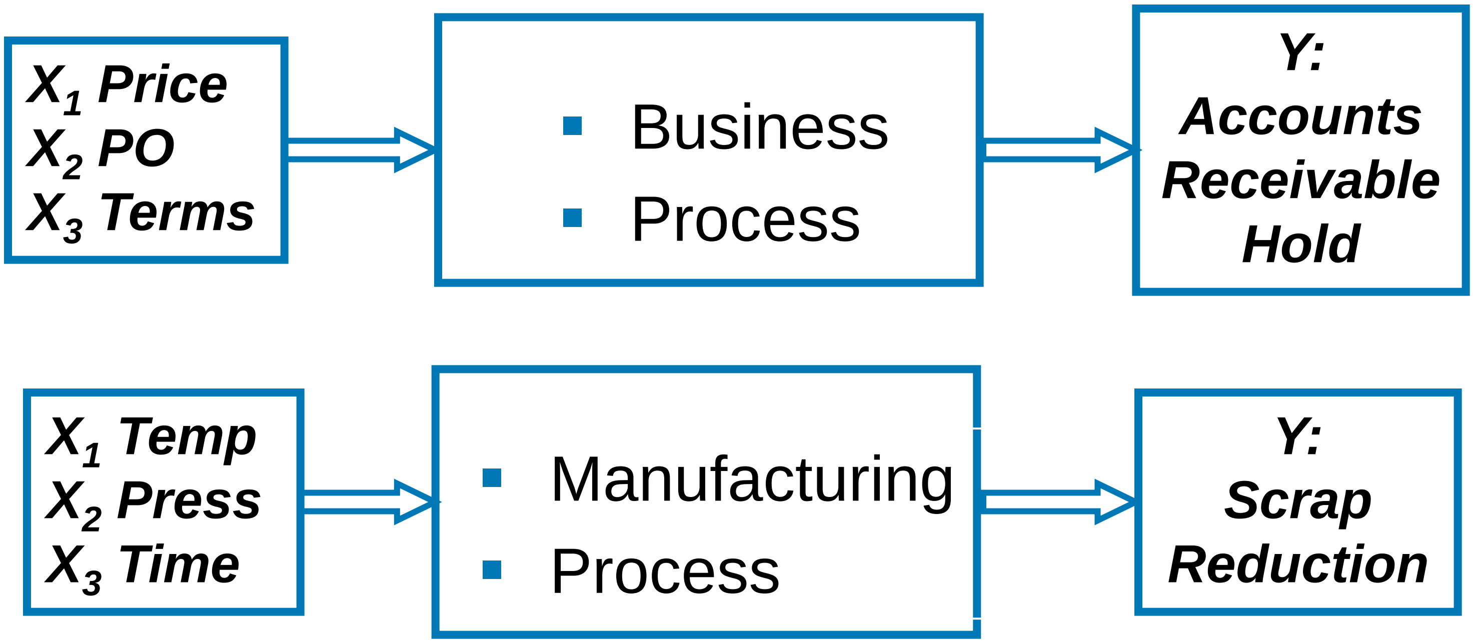
DOE Strategy Exercise

- Break into table teams. Select two projects at each table (one industrial process and one business process). Using flip charts list the Planning Phase elements for each project.
- Be prepared to discuss your learnings.

Anexas Consultancy Services

Factors

Factors and the Process



DOE can be applied to both business and industrial processes

Organizing the Factors

- The best conceived experiment will fail if executed poorly. Defining the factor settings and capturing the process response(s) properly is paramount to a successful DOE.

	INPUT FACTORS			OUTPUT RESPONSE REPLICATES					STATISTCAL ANALYSIS	
RUN	X_1	X_2	X_3	Y_1	Y_2	-	-	Y_n	Ybar	s
1										
2										
3										
4										
-										
-										
-										
n										

A matrix that clearly defines the input settings to the output response showing the flow of the experiment is a must

Factors at Two Levels

Many experiments will be quite successful when only two levels of each factor are studied.

	INPUT FACTORS			OUTPUT RESPONSE REPLICATES					STATISTCAL ANALYSIS	
RUN	X_1	X_2	X_3	Y_1	Y_2	-	-	Y_n	Ybar	s
1	234	1	0.0345						11	1.46
2	116	1	0.0737						14	2.03
3	234	2	0.0737						9	2.18
4	116	2	0.0345						16	4.22
-										
-										
-										
n										

Experimental Data Placed Here

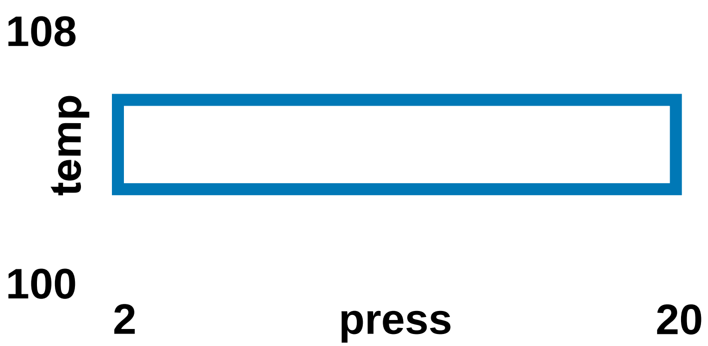
The matrix of inputs should be arranged to reduce or eliminate correlation between factors

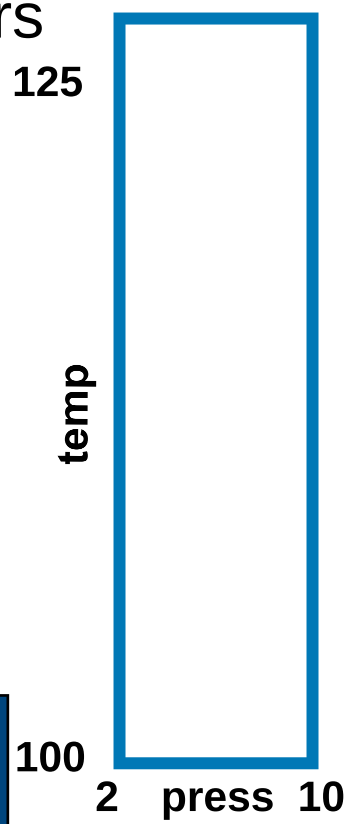
Anexas Consultancy Services

Coding

Factors and Scale of Units

Representation of the inference space is often related to the linear scale of units of the factors





It would be advantageous to normalize the range of factors to a common scale

Concept of Coding

Low
-1

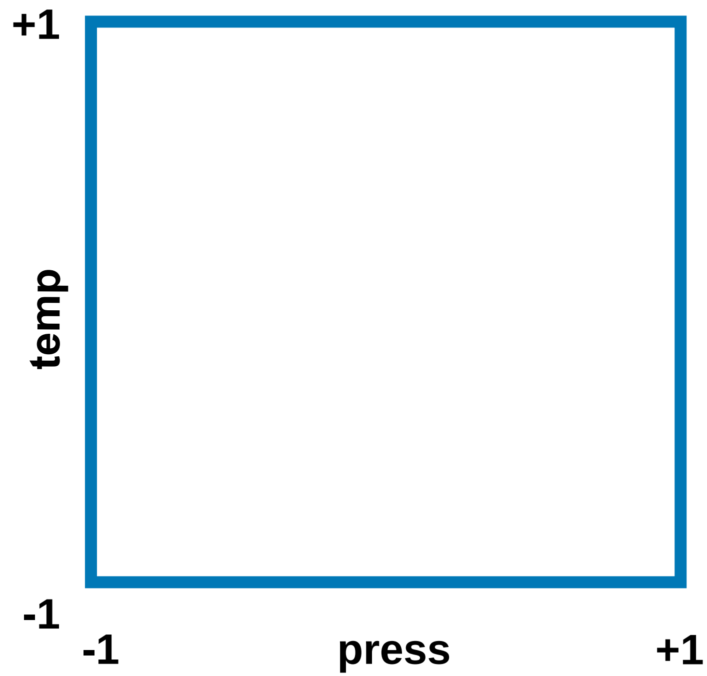
High
+1

Press = 2

Temp = 10

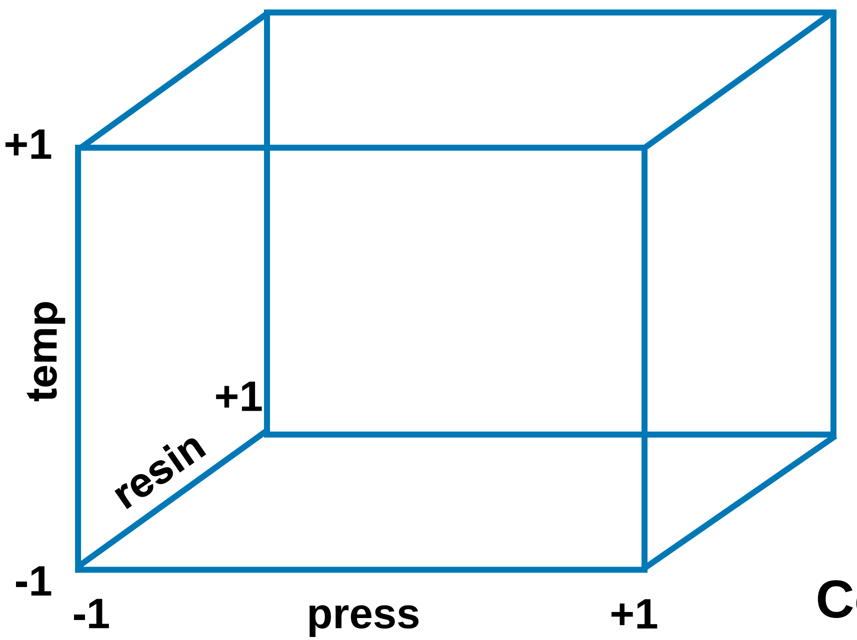
Press = 10

Temp = 125



Coding Three Factors

Low -1	High +1
Press = 2	Press = 10
Temp = 10	Temp = 125
Resin = 56	Resin = 3560



Coding convention often uses "+" for high and "-" for low

Class Coding Exercise

Transform the uncoded factor levels to coded factor levels using "+" and "-"

	UNCODED INPUT FACTORS			CODED INPUT FACTORS			OUTPUT RESPONSE			STATISTCAL ANALYSIS	
RUN	X_1	X_2	X_3	X_1	X_2	X_3	Y_1	Y_2 …	Y_n	Ybar	s
1	234	1	0.0345							11	1.46
2	116	1	0.0737							14	2.03
3	116	2	0.0737							9	2.18
4	234	2	0.0345							16	4.22
-											
-											
-											
n											

Why Code?

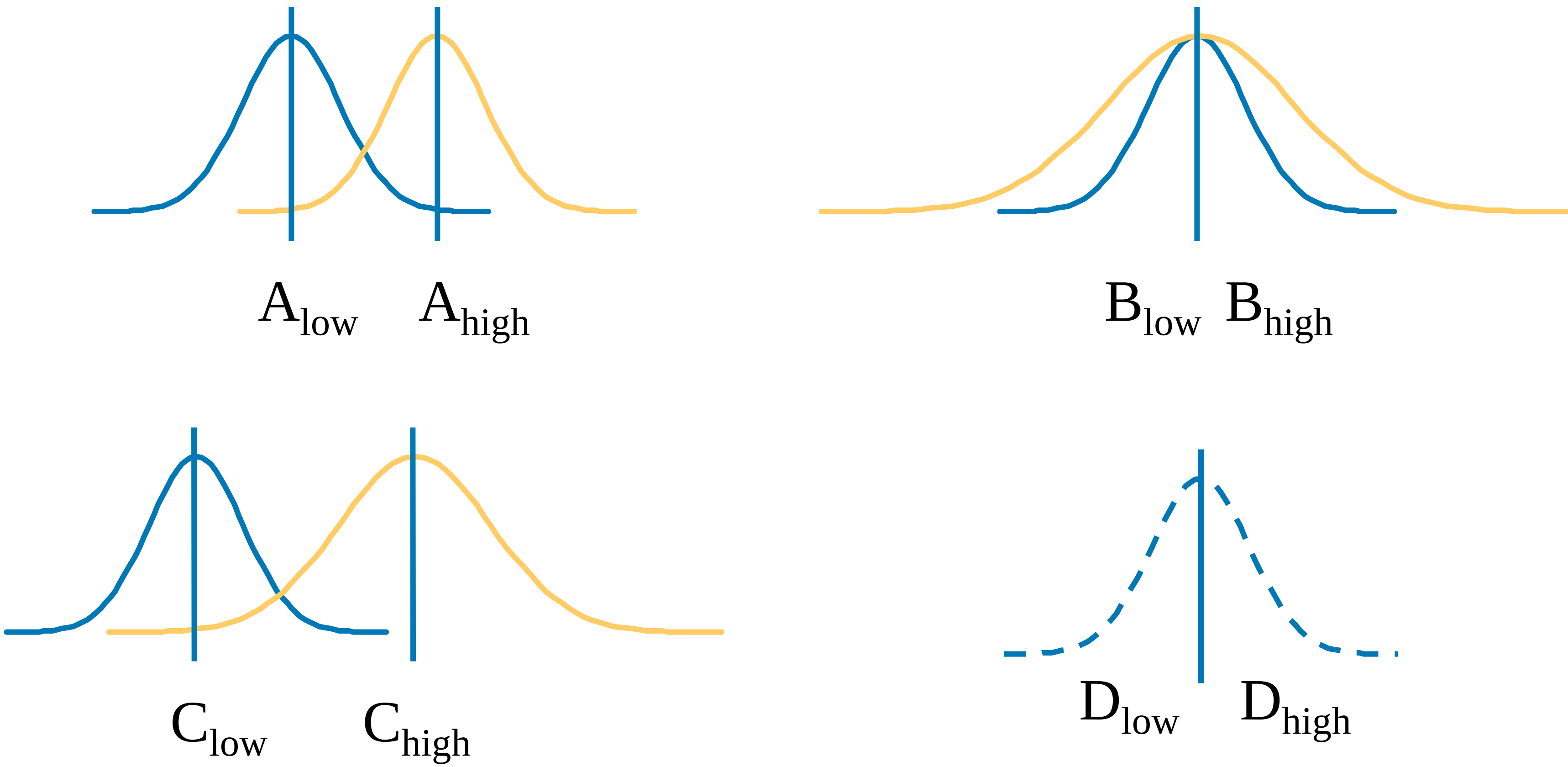
- Allows experimenter to focus on the design
- Shows the design generation clearly
- Permits easier mathematical analysis
- Produces fewer data entry errors
- Reduces analysis errors
- Manages magnitude of scale

Coding is the standard practice for experimental design

Anexas Consultancy Services

Exploring Factor Effects

How Factors Affect a Response



Which factors affect mean or variation?

One Factor at a Time (OFAT) Experiments

One Factor at a Time

- A process has two variable inputs: flow and temperature. Historically the process has run with flow = 18 and temp = 160 with a yield of 81.7%.

Flow	Temp	Yield
18	140	71.4
18	145	74.8
18	150	78.3
18	155	79.4
18	160	81.7
18	165	82.8
18	170	84.0
18	175	84.5
18	180	84.2
18	185	83.3
18	190	80.5
18	195	78.3
18	200	76.0

One Factor at a Time Data Analysis First Step

Flow	Temp	Yield
18	140	71.4
18	145	74.8
18	150	78.3
18	155	79.4
18	160	81.7
18	165	82.8
18	170	84.0
18	175	84.5
18	180	84.2
18	185	83.3
18	190	80.5
18	195	78.3
18	200	76.0

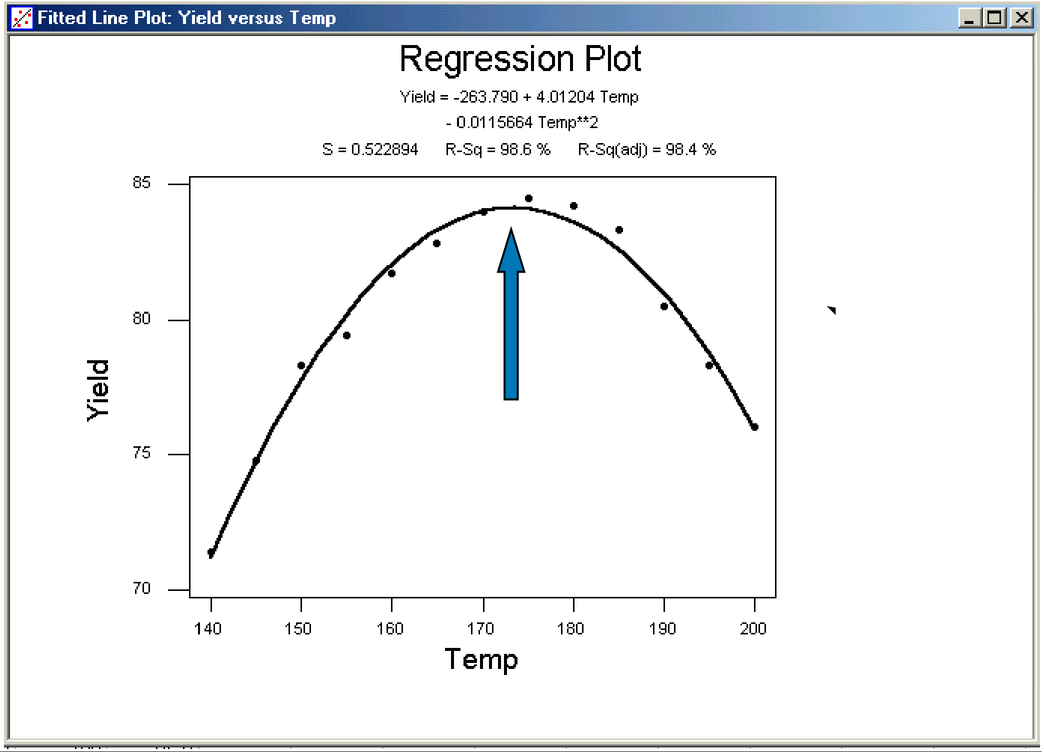
Old setting → 160 / 81.7

New setting → 175 / 84.5

Higher yield is obtainable, but at a higher temperature (cost)

Data is used for a regression analysis

One Factor at a Time Data Analysis Second Step



$$Y = -263.79 + 4.012Temp - 0.011566Temp^2$$

Regression analysis provides an equation which can be solved for temperature

Another test is planned to study flow at the "optimum" temperature of 174.4^{0}C

One Factor at a Time Data Analysis Third Step

Flow	Temp	Yield
2	174.4	65.5
4	174.4	65.5
6	174.4	77.8
8	174.4	84.8
10	174.4	90.8
12	174.4	91.8
14	174.4	89.8
16	174.4	88.8
18	174.4	84.5
20	174.4	82.8
22	174.4	74.2
24	174.4	58.8
26	174.4	40.8

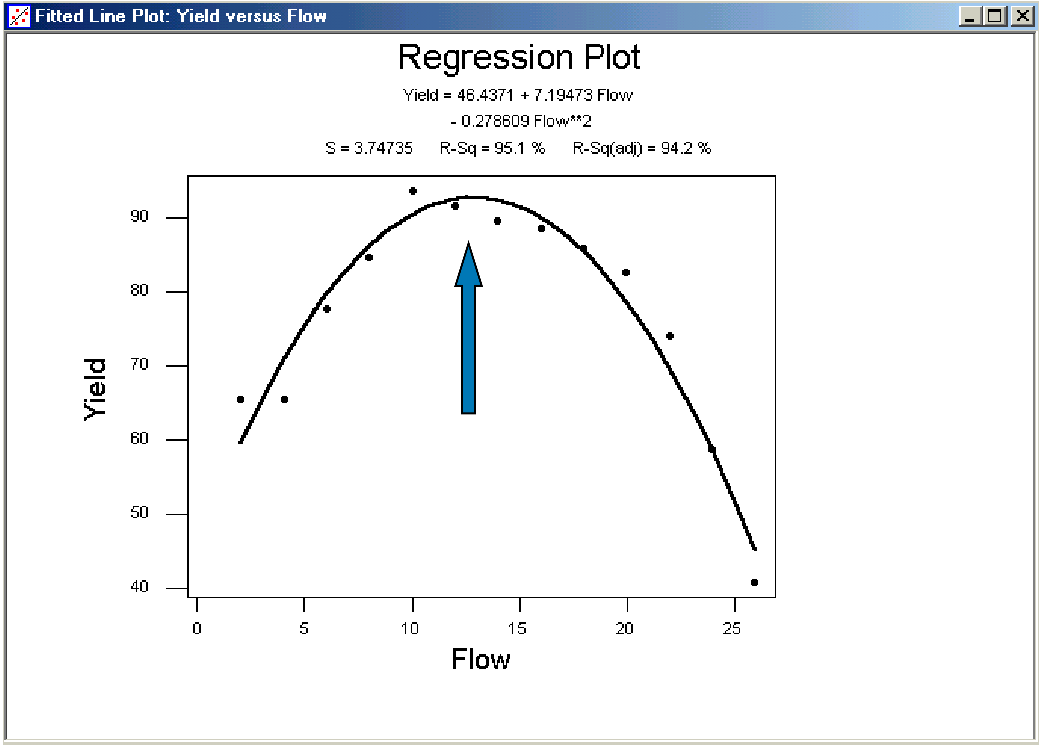
New setting (Flow 12, Yield 91.8)

Old setting (Flow 18, Yield 84.5)

Higher yield at lower flow is found, but is this the true optimum?

Data is used for a regression analysis

One Factor at a Time Data Analysis Fourth Step



Regression analysis provides equation which can be solved for flow

$$Y = 46.437 + 7.195Flow - 0.2786Flow^2$$

Another test is planned to study temp at the “optimum” flow of 12.9 lpm

One Factor at a Time Data Analysis Fifth Step

- Repeat steps 1 – 4 until satisfied “optimum” is found
- Each cycle will produce an equation for yield versus temperature
- Each cycle will produce an equation for yield versus flow
- Solve equations iteratively
- Repeat cycle, obtaining NEW equations

OFAT Equations

- Each cycle produces a set of independent, simultaneous equations describing the process
- First cycle:

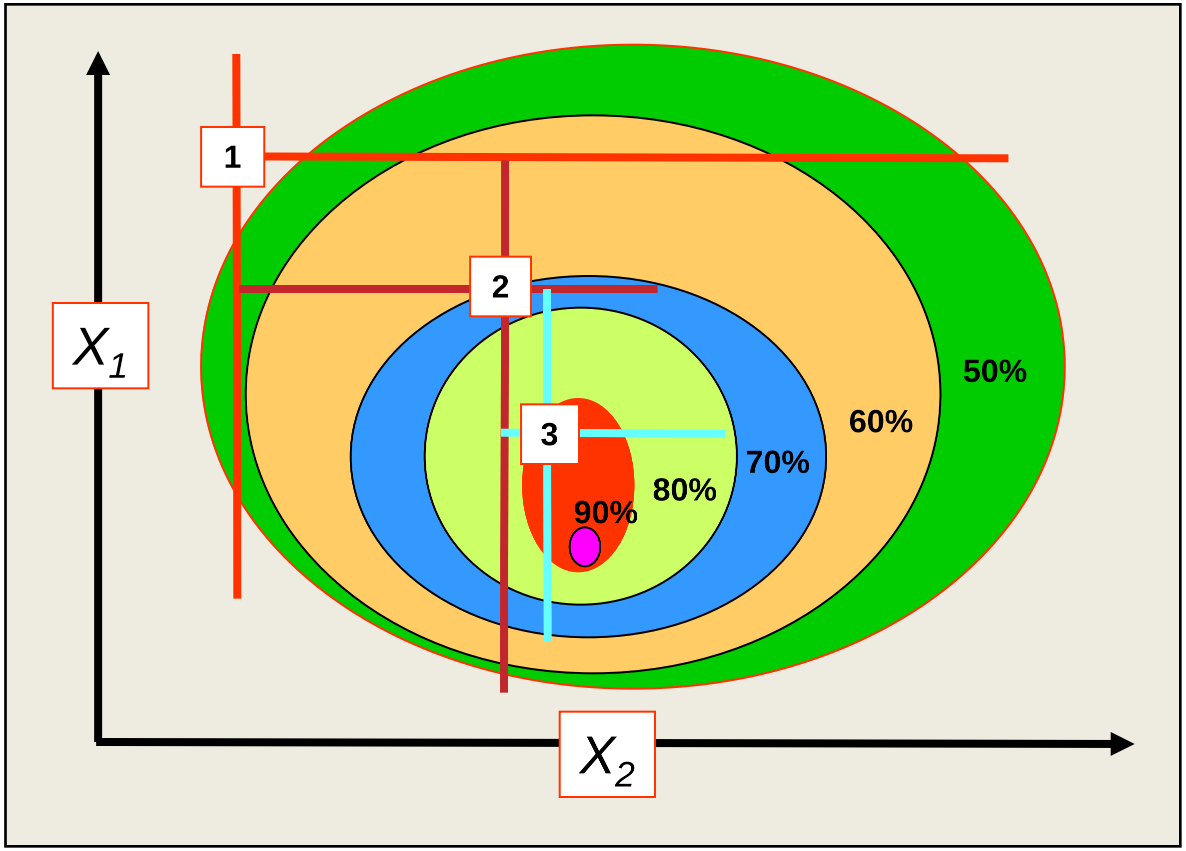
$$Y = -263.79 + 4.012Temp - 0.011566Temp^2$$
$$Y = 46.437 + 7.195Flow - 0.2786Flow^2$$

- Subsequent cycles:

$$Y = f(Temp)$$
$$Y = f(Flow)$$

OFAT will generate a family of independent prediction equations that do not readily lead to the optimum solution

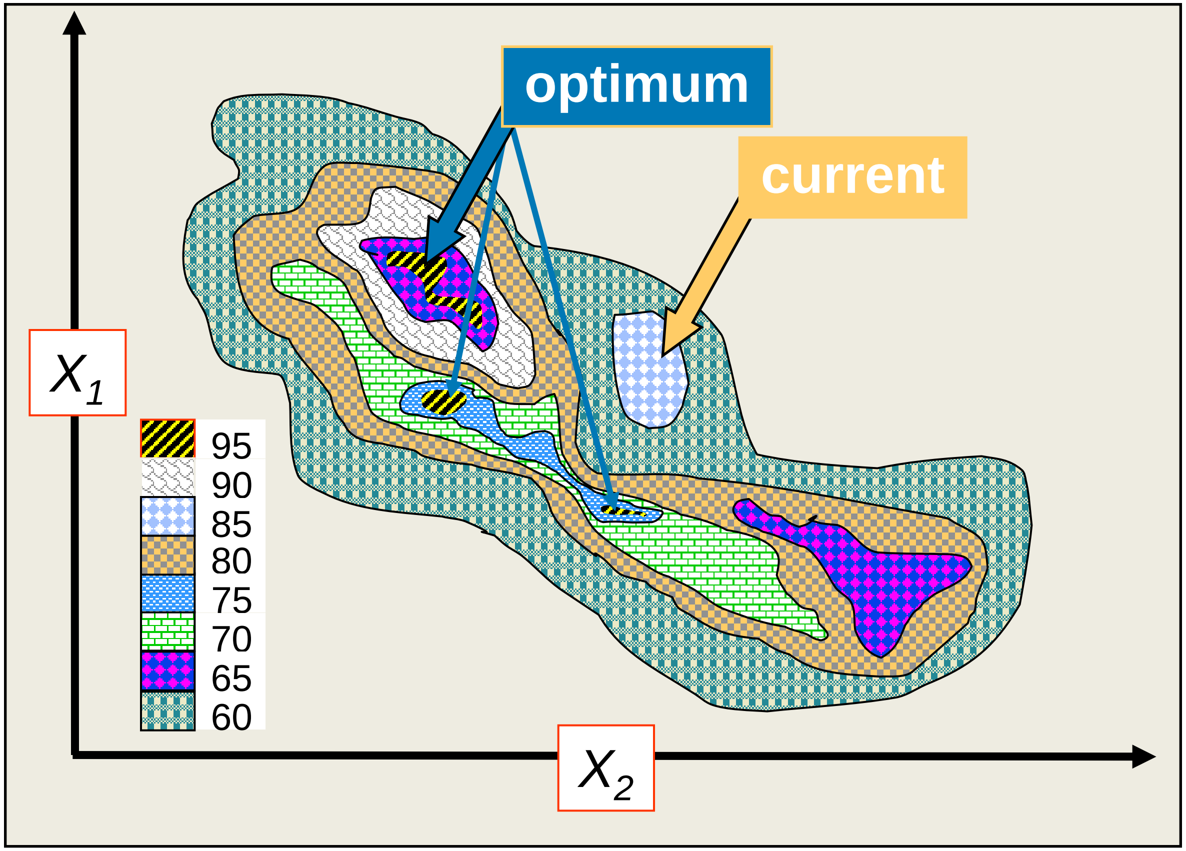
OFAT Prediction Contours



One factor at a time presumes a nested family of responses

Factor interactions are missed with OFAT

Process Response Contours



More often than not, the response contours of a process contain many peaks and valleys created by the interaction of factors

OFAT optimization can become inefficient and sometimes misleading

Prediction Equations

- More useful and efficient is one equation that predicts the output as a function of all factors

$$\hat{Y} = f(temp, press)$$

$$\hat{Y} = f(A, B, C...)$$

- This type of equation has multiple factors. It will contain:
  - Main effects
  - Interactions

General Form of DOE Prediction Equations

- The general full equation has all combinations and power terms

$$\hat{Y} = b_0 + b_1 Flow + b_2 Temp + b_3 Flow * Flow + b_4 Flow^2 + \ldots$$

- Most DOE experiments will use the linear solution of the general equation to define the response

$$\hat{Y} = b_0 + b_1 A + b_2 B + b_3 C + b_4 AB + b_5 AC + \ldots + b_7 ABC$$

DOE equations will deal with a linear subset of the general equation

Capturing Interactions

Statistically based designs provide information about factors and interactions

- 2k-n Fractional Factorial Designs
- 2k Full Factorial
- Full Factorial (replication)
- RSM (response surface method)

Revealing information about main effects and interactions is key to process understanding

Prepared to Explore DOE

DoE 3 Factor 2 Level ***

↓	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	A	B	C	
1	6	1	1	1	1	-1	1	
2	3	2	1	1	-1	1	-1	
3	2	3	1	1	1	-1	-1	
4	1	4	1	1	-1	-1	-1	
5	8	5	1	1	1	1	1	
6	7	6	1	1	-1	1	1	
7	4	7	1	1	1	1	-1	
8	5	8	1	1	-1	-1	1	
9								

Minitab will be the tool for creating and analyzing DOE experiments in the Breakthrough Strategy ™

Key Learning Points

Objectives Review

By the end of this module, the participant should:

- Understand the strategy behind Design of Experiments (DOE)
- Create a design matrix
- Code factors
- Understand the limitations of OFAT experiments
- Interpret interactions

Anexas Consultancy Services

Full Factorial Design of Experiments

Module Objectives

By the end of this module, the participant will:

- Generate a full factorial design
- Look for factor interactions
- Develop coded orthogonal designs
- Write process prediction equations (models)
- Set factors for process optimization
- Create and analyze designs in Minitab™
- Evaluate residuals
- Develop process models from Minitab™ analysis
- Determine sample size

Why Learn About Full Factorial DOE?

- A Full Factorial Design of Experiment will
- Provide the most response information about
  - Factor main effects
  - Factor interactions
- Provide the process model's coefficients for
  - All factors
  - All interactions
- When validated, allow process to be optimized

What is a Full Factorial DOE

- A full factorial DOE is a planned set of tests on the response variable(s) (KPOVs) with one or more inputs (factors) (PIVs) with all combinations of levels

- ANOVA analysis will show which factors are significant
- Regression analysis will provide the coefficients for the prediction equations
  - Mean
  - Standard deviation
- Residual analysis will show the fit of the model

DOE Terminology

- **Response** (Y, KPOV): the process output linked to the customer CTQ
- **Factor** (X, PIV): uncontrolled or controlled variable whose influence is being studied
- **Level**: setting of a factor (+, -, 1, -1, hi, lo, alpha, numeric)
- **Treatment Combination** (run): setting of all factors to obtain a response
- **Replicate**: number of times a treatment combination is run (usually randomized)
- **Repeat**: non-randomized replicate
- **Inference Space**: operating range of factors under study

Full Factorial DOE Objectives

- Learning the most from as few runs as possible..
- Identifying which factors affect mean, variation, both or have no effect
- Modeling the process with prediction equations,

$$\hat{Y} = f(A, B, C \ldots)$$

$$\hat{s} = f(A, B, C \ldots)$$

- Optimizing the factor levels for desired response
- Validating the results through confirmation

Linear Combinations of Factors for Two Levels

Combinations of Factors and Levels

- A process whose output Y is suspected of being influenced by three inputs A, B and C. The SOP ranges on the inputs are

A	15 through 25, by 1
B	200 through 300, by 2
C	1 or 2

- A DOE is planned to test all combinations

Is testing all combinations possible, reasonable and practical?

Combinations of Factors and Levels cont'd

- Setting up a matrix for the factors at all possible process setting levels will produce a really large number of tests.
- The possible levels for each factor are

  A = 11

  B = 51

  C =   2
- How many combinations are there?

A	B	C
15	200	1
16	200	1
17	200	1
18	200	1
19	200	1
20	200	1
21	200	1
22	200	1
23	200	1
24	200	1
25	200	1
15	202	1
16	202	1
17	202	1
.	.	.
.	.	.
.	.	.
.	.	.
22	300	2
23	300	2
24	300	2
25	300	2

We must make assumptions about the response in order to manage the experiment

Linear Response for Factors at Two Levels

- The team decides, from process knowledge, that the response is close to being linear throughout the range of factor level settings (inference space).
- A reasonable assumption for most processes
- The levels of the factors for the test would then be

  A 15 and 25

  B 200 and 300

  C 1 and 2

The design becomes much more manageable!

The Three Factor Design at Two Levels

- The revised experiment consists of all possible combinations of A, B and C each at the chosen low and high settings:

A	B	C
15	200	1
15	200	2
15	300	1
15	300	2
25	200	1
25	200	2
25	300	1
25	300	2

This is a 2^3 full factorial design (pronounced two to the three). It consists of all combinations of the three factors each at two levels

Naming Conventions

- The naming convention for full factorial designs has the level raised to the power of the factor:

$$\textbf{\textit{level}}^{\text{(factor)}}$$

- and is called “a (level) to the (factor) design”

- What would a two level, four factor design be called?
- How many combinations (runs) are in a 23 design?

Class Exercise

- Write the total number of combinations for the following designs

  2^3

  2^4

A	B	C	D

- Assume factors are named A, B, C, D, etc. and the levels are low "-" and high "+".

Did we all generate the same designs?

The Yates Standard Order

A method to generate experimental designs in a consistent and logical fashion was developed by Frank Yates.

The Yates Standard Order
Step 1

Create a matrix with factors along the top, runs down the left side (2^3 example shown)

	Factors		
Runs	**A**	**B**	**C**
1			
2			
3			
4			
5			
6			
7			
8			

For a 2^3 there will be 8 runs

The Yates Standard Order
Step 2

Starting with the first factor, insert its low value in the first row followed by its high value in the second row. Repeat through the last row.

	Factors		
Runs	**A**	**B**	**C**
1	**15**		
2	**25**		
3	**15**		
4	**25**		
5	**15**		
6	**25**		
7	**15**		
8	**25**		

The Yates Standard Order
Step 3

Move to the next factor and place its low value in the first two rows, followed by its high value in the next two rows. Repeat through the last row.

	Factors		
Runs	**A**	**B**	**C**
1	15	200	
2	25	200	
3	15	300	
4	25	300	
5	15	200	
6	25	200	
7	15	300	
8	25	300	

The Yates Standard Order
Step 4

Move to the next factor and place its low value in the first four rows, followed by its high value in the next four rows. Repeat through the last row.

	Factors		
Runs	**A**	**B**	**C**
1	15	200	1
2	25	200	1
3	15	300	1
4	25	300	1
5	15	200	2
6	25	200	2
7	15	300	2
8	25	300	2

The Yates Standard Order
Step N

Continue the pattern until all factors are included

Yates Design Generator

Factor	Row	Pattern
1st	1 low	1 high
2nd	2 low	2 high
3rd	4 low	4 high
•		
•		
nth	2n-1 low	2n-1 high

The Yates Standard Order Summary

- The Yates design generator yields all possible combinations of factor levels.
- This is a full factorial two level design for three factors.

	Factors		
Runs	**A**	**B**	**C**
1	15	200	1
2	25	200	1
3	15	300	1
4	25	300	1
5	15	200	2
6	25	200	2
7	15	300	2
8	25	300	2

Class Exercise

- Write the total number of combinations for the following designs using the Yates Standard order design generator.
- **2^3**
- **2^4**
- Assume factors are named A, B, C, D, etc. and the levels are low = “-” and high =“+”.

Did we all generate the same designs?

Anexas Consultancy Services

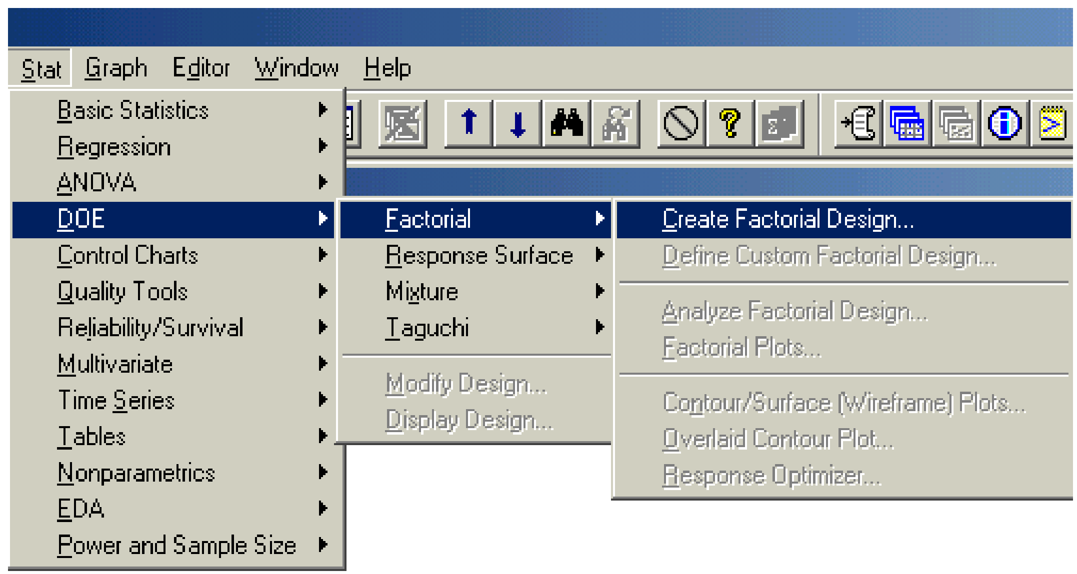
Creating a Full Factorial Design in Minitab®

Creating a Full Factorial Design in Minitab Step 1

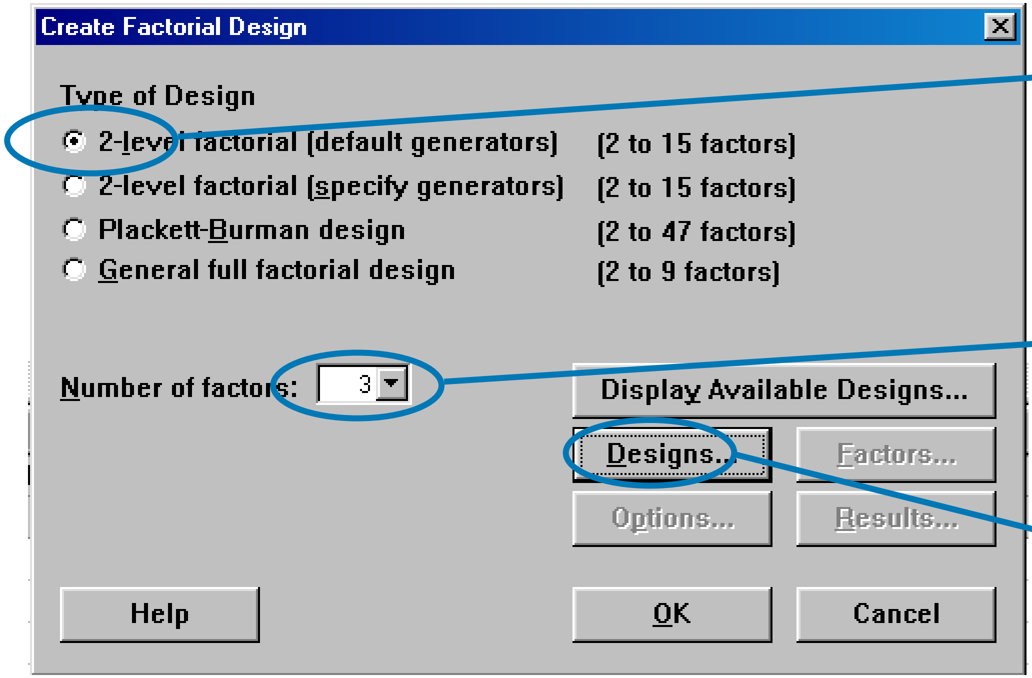
Tool Bar Menu > Stat > DOE > Factorial > Create Factorial Design

- Create a Minitab® design for
  - A - pressure
  - B - temperature
  - C - material vendor

2^3 design



Creating a Full Factorial Design in Minitab Step 2

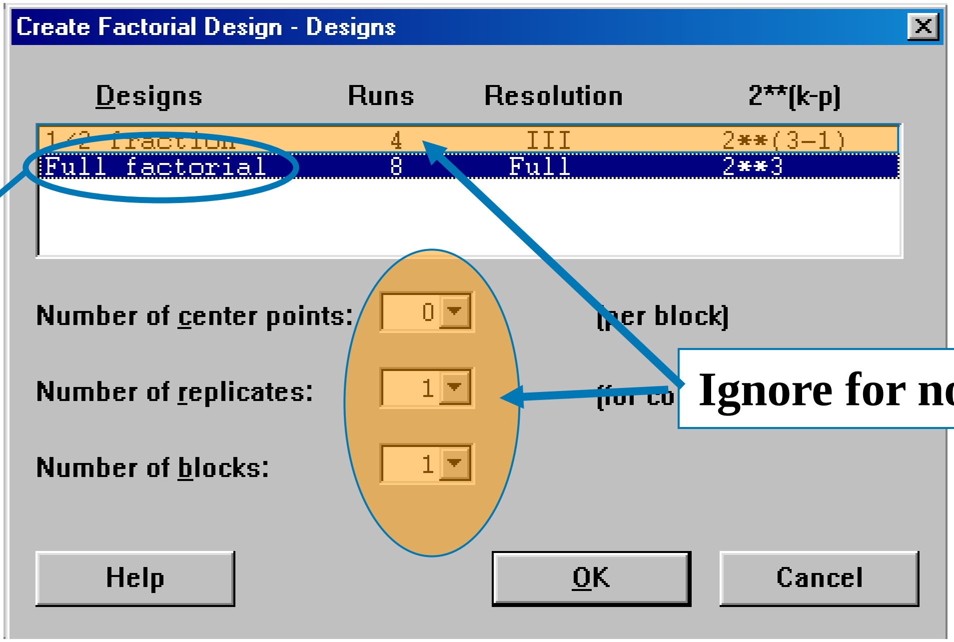


1. Select default generators (Yates is default)

2. Set number of factors to three

3. Go to Designs

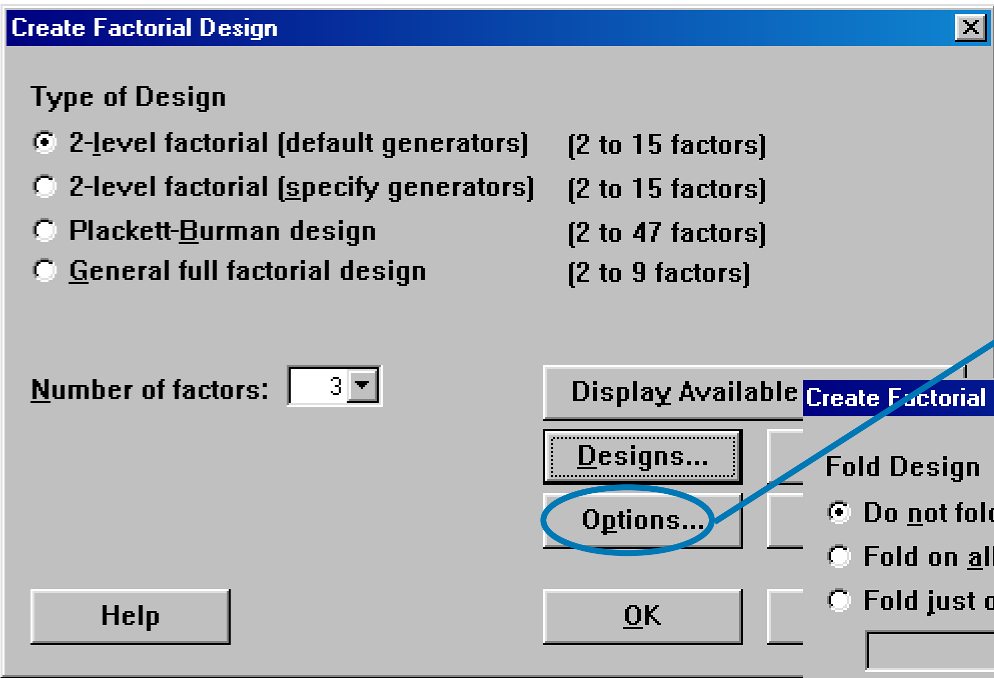
Creating a Full Factorial Design in Minitab Step 3

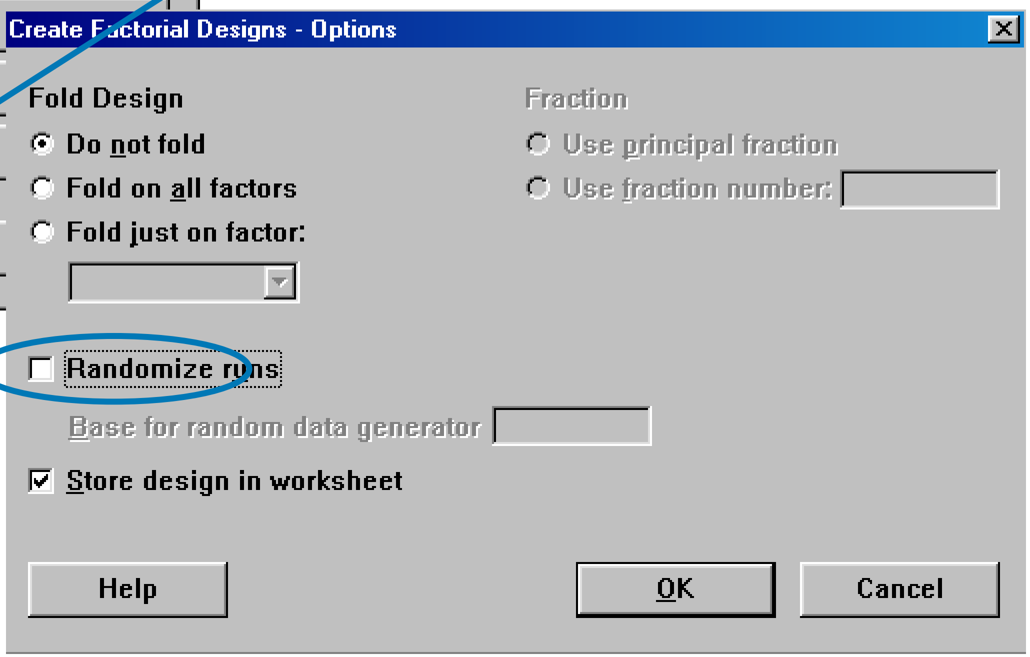


Select Full factorial

Ignore for now

Creating a Full Factorial Design in Minitab Step 4

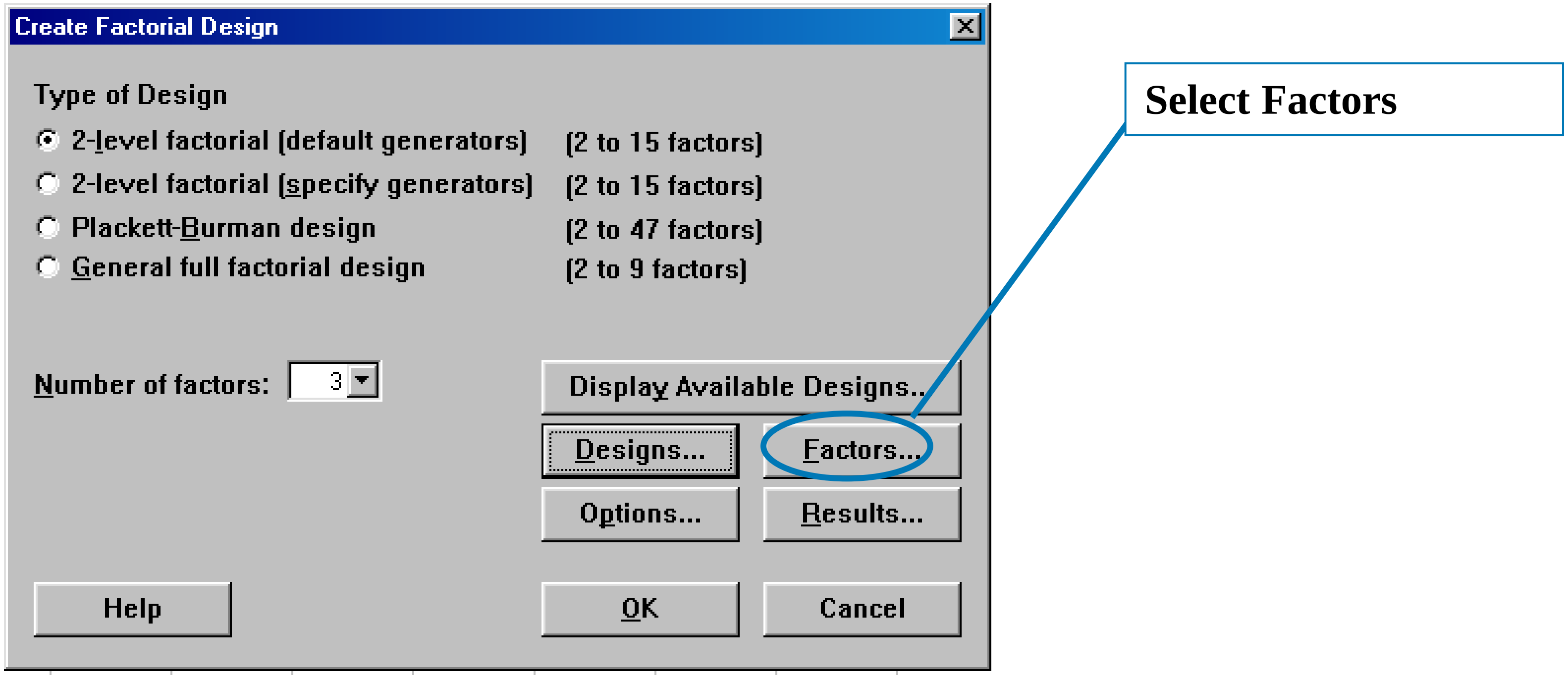




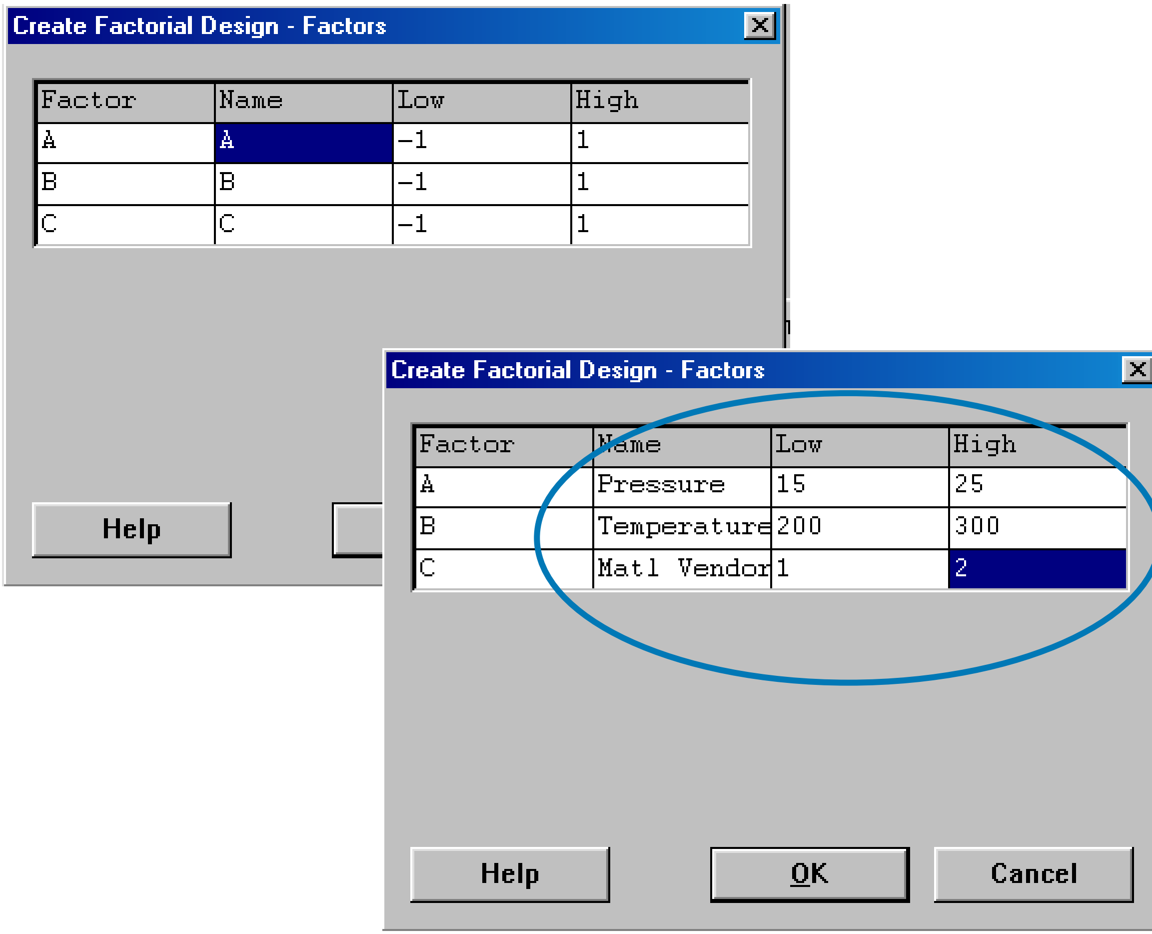
1. Select Options

2. Un-check Randomize runs

Creating a Full Factorial Design in Minitab Step 5

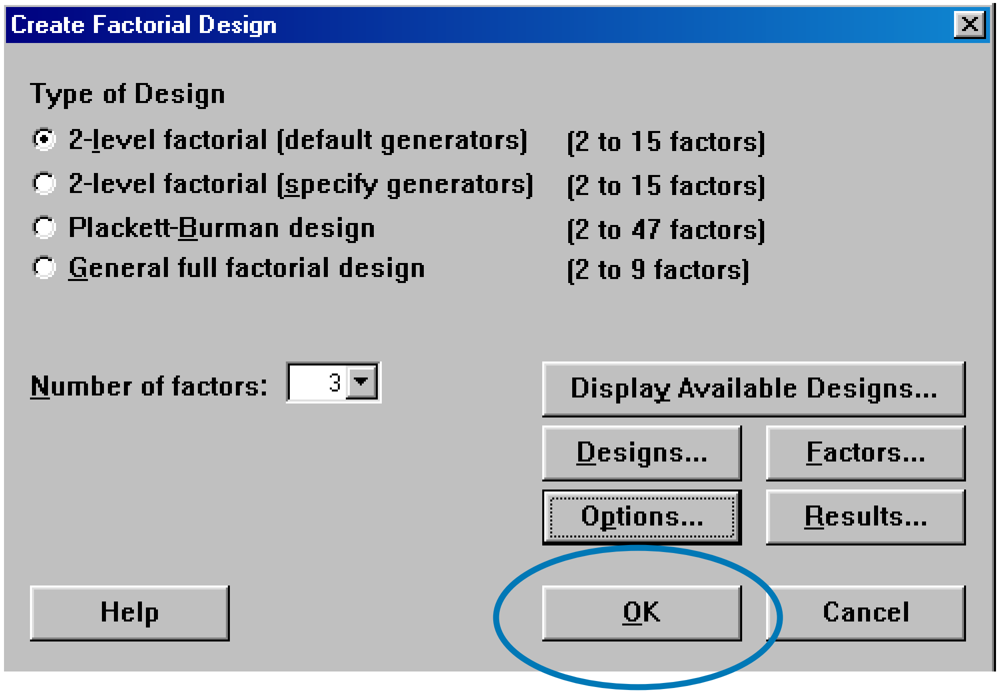


Creating a Full Factorial Design in Minitab Step 6



Input the Factor Names and Levels

Creating a Full Factorial Design in Minitab Step 7



Design parameters are entered; ready to create the design in the worksheet.

Creating a Full Factorial Design in Minitab Step 8

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	Pressure	Temperature	Matl Vendor	
1	1	1	1	1	15	200	1	
2	2	2	1	1	25	200	1	
3	3	3	1	1	15	300	1	
4	4	4	1	1	25	300	1	
5	5	5	1	1	15	200	2	
6	6	6	1	1	25	200	2	
7	7	7	1	1	15	300	2	
8	8	8	1	1	25	300	2	
9								

This is the Yates Standard order for a 2^3 uncoded design

Std Order Column of Minitab Design

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	Pressure	Temperature	Matl Vendor	
1	1	1	1	1	15	200	1	
2	2	2	1	1	25	200	1	
3	3	[illegible]	[illegible]	1	15	300	1	
4	4	[illegible]	[illegible]	1	25	300	1	
5	5	[illegible]	[illegible]	1	15	200	2	
6	6	[illegible]	[illegible]	1	25	200	2	
7	7	[illegible]	[illegible]	1	15	300	2	
8	8	8	1	1	25	300	2	
9								

The StdOrder Column is the Yates Standard Order

Run Order Column of Minitab Design

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	Pressure	Temperature	Matl Vendor	
1	1	1			15	200	1	
2	2	2			25	200	1	
3	3	3			15	300	1	
4	4	4			25	300	1	
5	5	5			15	200	2	
6	6	6			25	200	2	
7	7	7			15	300	2	
8	8	8			25	300	2	
9								

RunOrder Column is the sequence of the runs

StdOrder is created by the design choice RunOrder is created by randomize runs choice

Class Exercise Part 1

- Generate the previous design with the randomize runs tic box selected
- Compare results with your neighbors

Class Exercise Part 2

Create a non random Minitab design for the following:

The factors are

	NAME	LEVEL (Low	High)
A	Customer Type	Retail	Consumer
B	System	Legacy	SAP
C	Warehouse	Atlanta	Dallas
D	Overtime Hours	0	100

The response is on-time delivery.

Be prepared to present your results.

Anexas Consultancy Services

Replicates and Repeats

What are Replicates and Repeats?

- Replicate
- Total run of all treatment combinations
  - Usually in random order
- Requires factor level change between runs
- All experiments will have one replicate
  - Two replicates are two complete experiment runs
- Statistically best experimental scenario

Repeat (also repetition)

- Additional run without factor level change

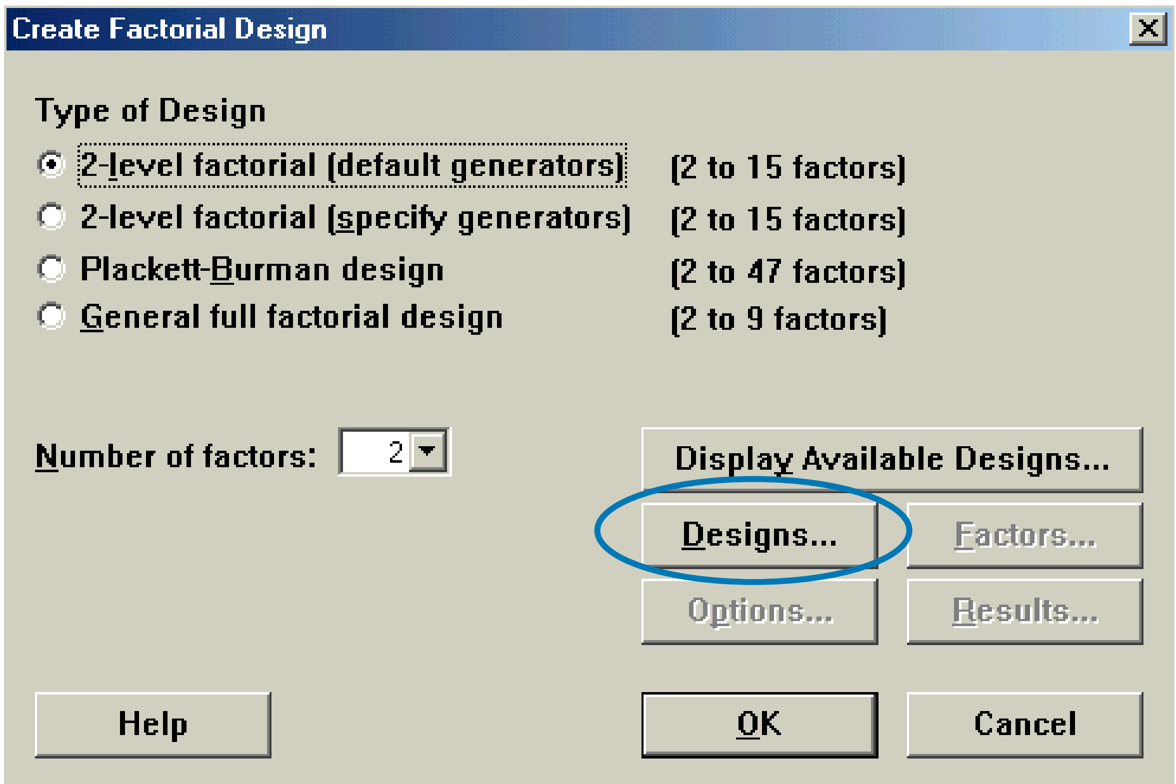
Minitab Design Replication

- Minitab easily handles replicating the design
- Replicate or repeat is treated same in design
- Actual factor level change between runs is at the discretion of the experimenter
  - Minitab provides treatment combination
  - Randomization or information needed is part of strategy of experiment

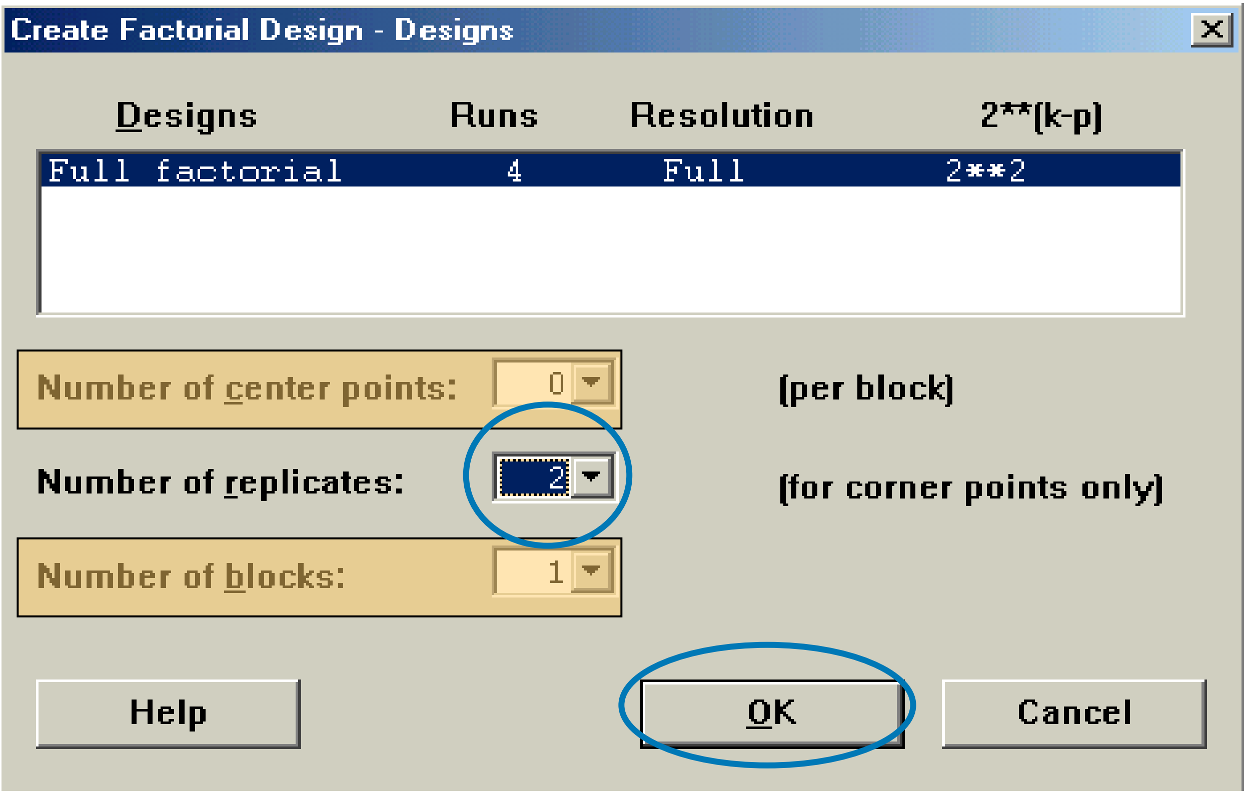
Replication in Minitab
Step 1

Tool Bar Menu > Stat > DOE > Factorial > Create Factorial Design

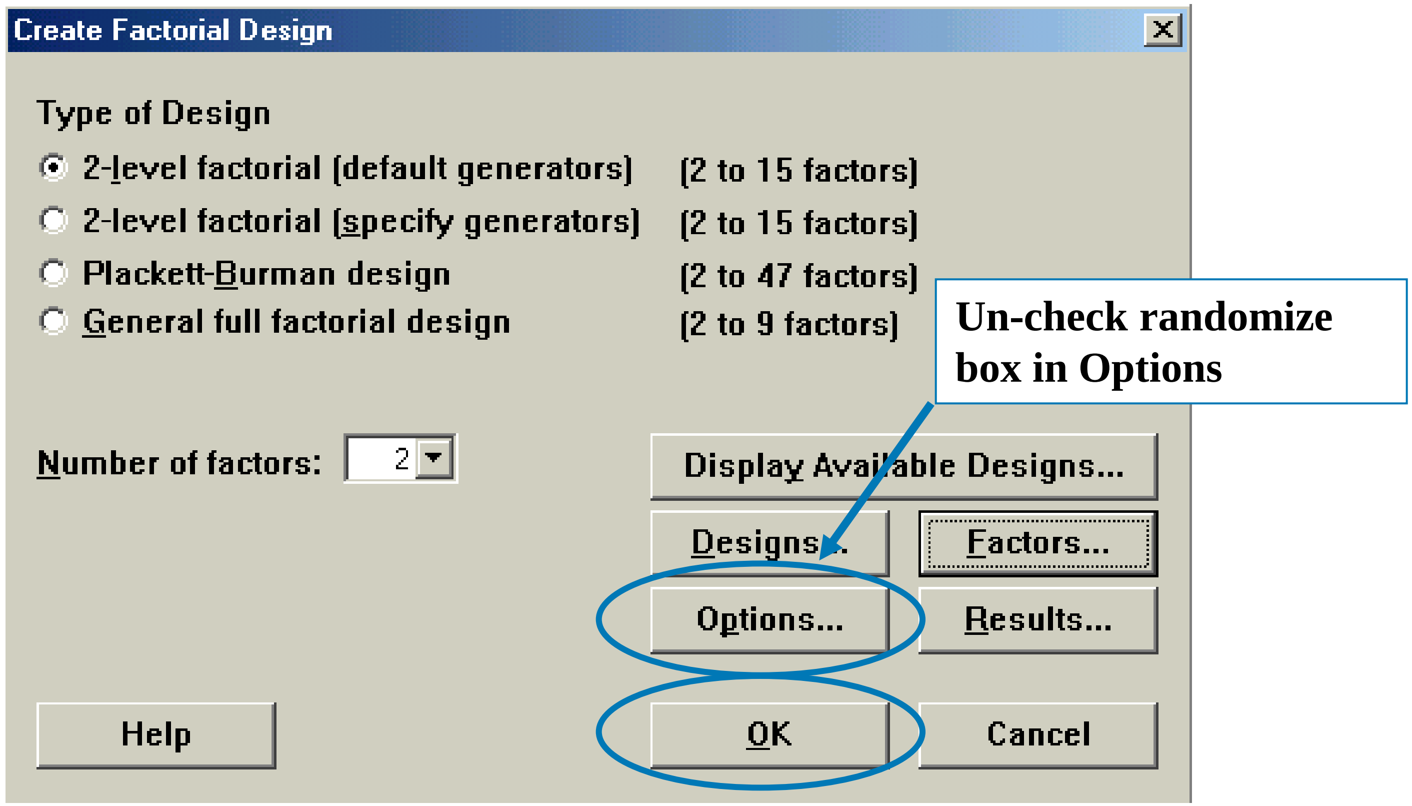
Create a 2^2 with two non randomized replicates



Replication in Minitab
Step 2



Replication in Minitab
Step 3



Skip the other dialog options

Replication in Minitab
Step 4

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6	C7
	StdOrder	RunOrder	CenterPt	Blocks	A	B	
1	1	1	1	1	-1	-1	
2	2	2	1	1	1	-1	
3	3	3	1	1	-1	1	
4	4	4	1	1	1	1	
5	5	5	1	1	-1	-1	
6	6	6	1	1	1	-1	
7	7	7	1	1	-1	1	
8	8	8	1	1	1	1	
9							

First Replicate

Second Replicate

Response Y would be placed in C7

Randomized Replication in Minitab for Step 4

Worksheet 2 ***

↓	C1	C2	C3	C4	C5	C6	C7
	StdOrder	RunOrder	CenterPt	Blocks	A	B	
1	7	1	1	1	-1	1	
2	4	2	1	1	1	1	
3	5	3	1	1	-1	-1	
4	3	4	1	1	-1	1	
5	8	5	1	1	1	1	
6	2	6	1	1	1	-1	
7	1	7	1	1	-1	-1	
8	6	8	1	1	1	-1	
9							

Response Y would be placed in C7

Coding the Design

Coding the design by transforming the low factor level to a “-1” and the high factor level to a “+1” offers analysis advantages

Coding Review Exercise

Fill in the coded design based upon the uncoded design

	Uncoded Factors			Coded Factors		
Runs	A	B	C	A	B	C
1	15	200	1			
2	25	200	1			
3	15	300	1			
4	25	300	1			
5	15	200	2			
6	25	200	2			
7	15	300	2			
8	25	300	2			

Any uncoded design can be transformed into a coded design

Requirement of Factor Independence

- Factors are mathematically independent when only the response is a function of the factors
- A factor is not a function of another factor
- The coded design is orthogonal
  - Factors will be independent

DOE analysis requires that the factors be independent

Anexas Consultancy Services

Orthogonal Designs

Determining Orthogonality

Consider the 2^2 design

	Coded Factors		OutPut
Runs	X_1	X_2	Y
1	-1	-1	(y_1)
2	1	-1	(y_2)
3	-1	1	(y_3)
4	1	1	(y_4)

The run outputs, Y_n, can be described by

where b and c are coded settings of the factors

It can be shown that for X_1 to be independent of X_2

$$\sum_{i=1}^{n} b_i c_i = 0$$

Calculating Orthogonality

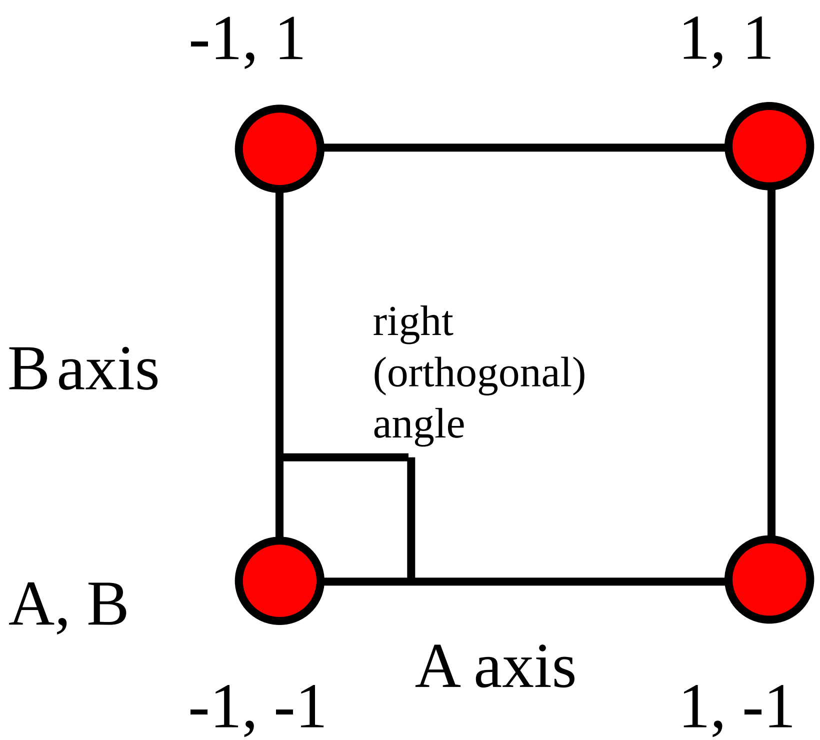
	Coded Factors		Coef1*coeff2
Runs	X1	X2	X1*X2
1	-1	-1	1
2	1	-1	-1
3	-1	1	-1
4	1	1	1
		$\sum$	0

Satisfies

$$\sum_{i=1}^{n} b_i c_i = 0$$

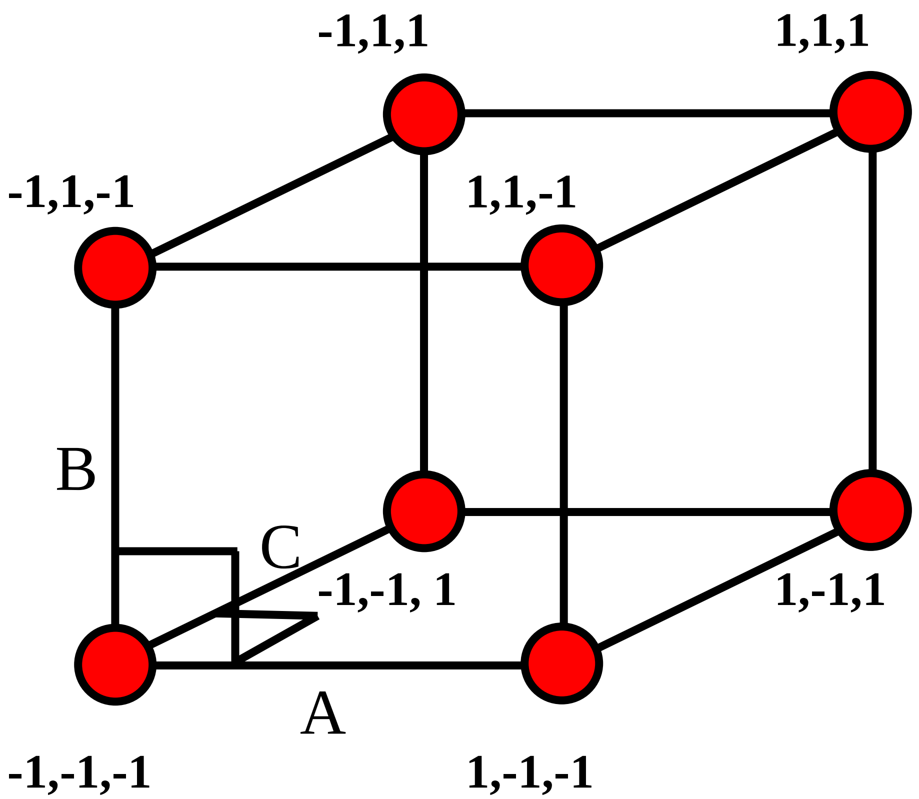
Therefore orthogonal and independent

Coding and Orthogonality for a 2^2 Design



Orthogonal designs can be represented as a geometric (mathematical) figure

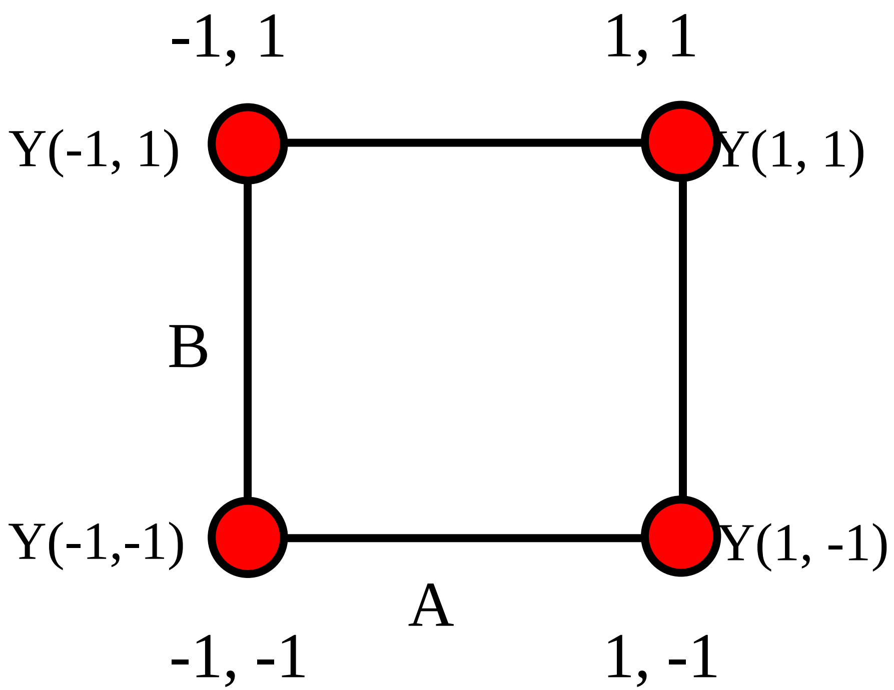
Coding and Orthogonality for a 2^3 Design



The vertices are the response measurement points; the volume within is the inference space

Response and Orthogonality

The response can be measured at each corner of the design, as represented by Y(A,B).



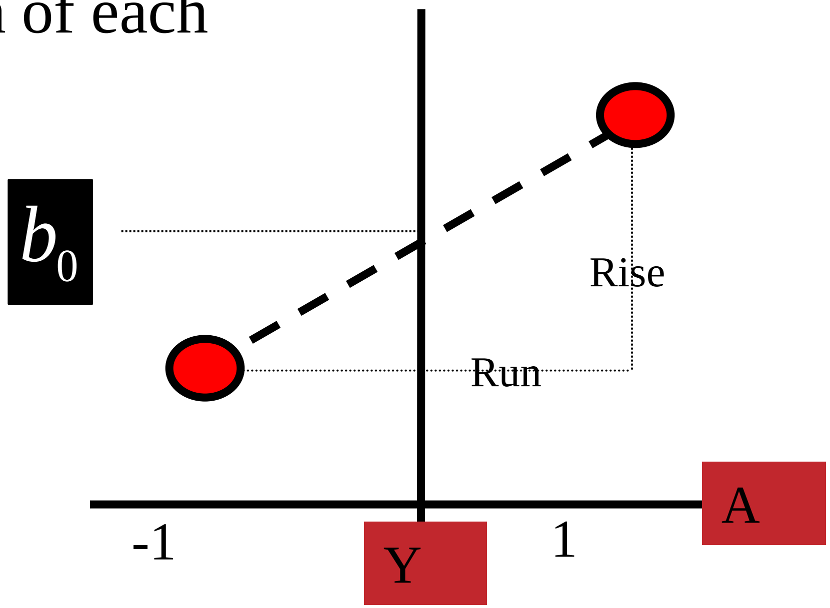
Coding and the Linear Model

Since the factors are orthogonal, Y is a linear combination of each factor

The linear equation,

$$Y = b_0 + b_1 A$$

describes the line between the points



$b_0 =$ Intercept

$b_1 =$ Rise/Run

Fitting a linear model to the coded design becomes very easy

Main Effects and Interactions

Calculating Main Effects

A DOE is run:

	Coded Factors		Response
	A	B	Y
	-1	-1	48
	1	-1	96
	-1	1	72
	1	1	36
Y_{ave} at $FACTOR_{high}$	66	54	
Y_{ave} at $FACTOR_{low}$	60	72	
Effect	6	-18	

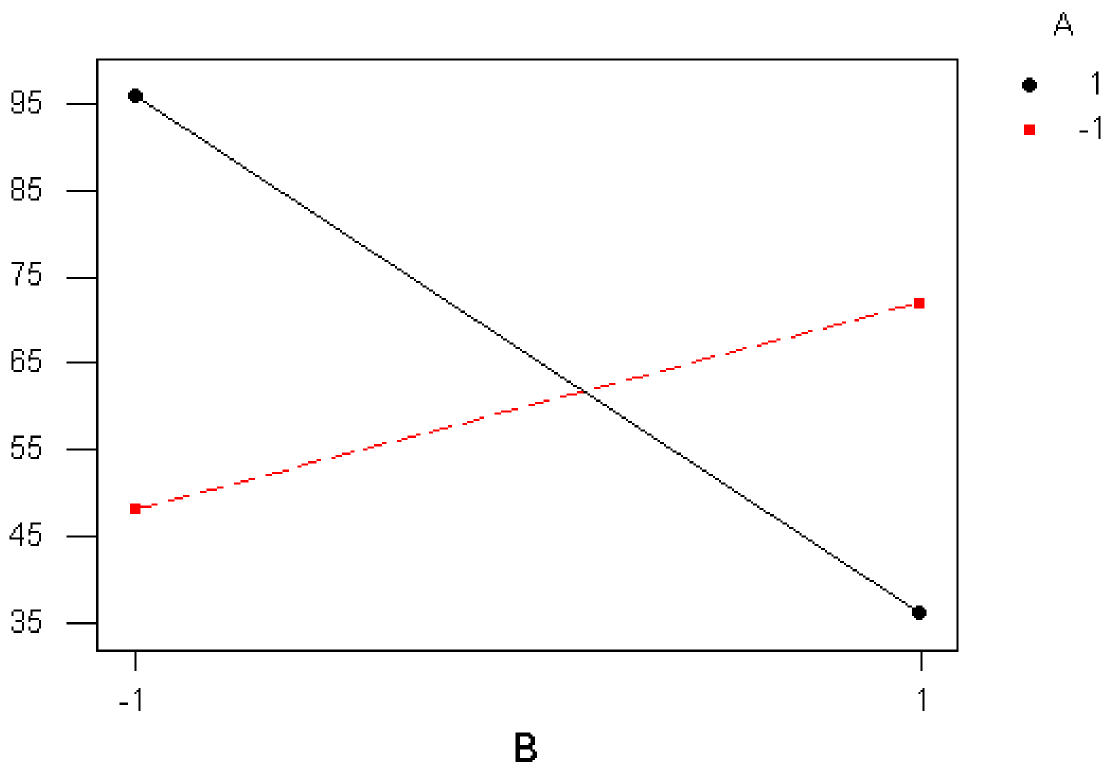
$$\frac{96+36}{2}=66$$

What does a non-zero effect mean?

$$effect = \overline{Y}(@\ factorhigh) - \overline{Y}(@\ factorlow)$$

The factor (or main) effects are easily calculated

Discovering Interactions



A response change due to both A and B changing is called an interaction

Linear Prediction Equation With Interaction Term

- The interaction will add a term to the linear equation

$$Y = b0 + b1\ A + b2\ B + b3\ AB$$

- A and B are the main effects
- AB is the interaction
- b0 is the grand mean (intercept)
- b1, b2, and b3 are the term coefficients

Coded Values for Interactions

The coded value matrix for an interaction is the product of the factor coded values:

A	B
-1	-1
1	-1
-1	1
1	1

-1*-1 = 1
1*-1 = -1
-1*1 = -1
1*1 = 1

AB
1
-1
-1
1

Finding the interaction levels in a coded design is as simple as multiplying one times one

Calculating Interaction Effects

The DOE again

	Coded Factors		Interaction	Response
This is the Yates Standard Order for a 2^2 coded design with interaction	A	B	AB	Y
	-1	-1	1	48
	1	-1	-1	96
	-1	1	-1	72
	1	1	1	36
Y_{ave} at $FACTOR_{high}$	66	54	42	
Y_{ave} at $FACTOR_{low}$	60	72	84	
Effect	6	-18	-42	←

The interaction is set by the design (math) based upon the factor settings

Anexas Consultancy Services

Main Effects, Interactions and Cube Plots in Minitab

Create the Experiment

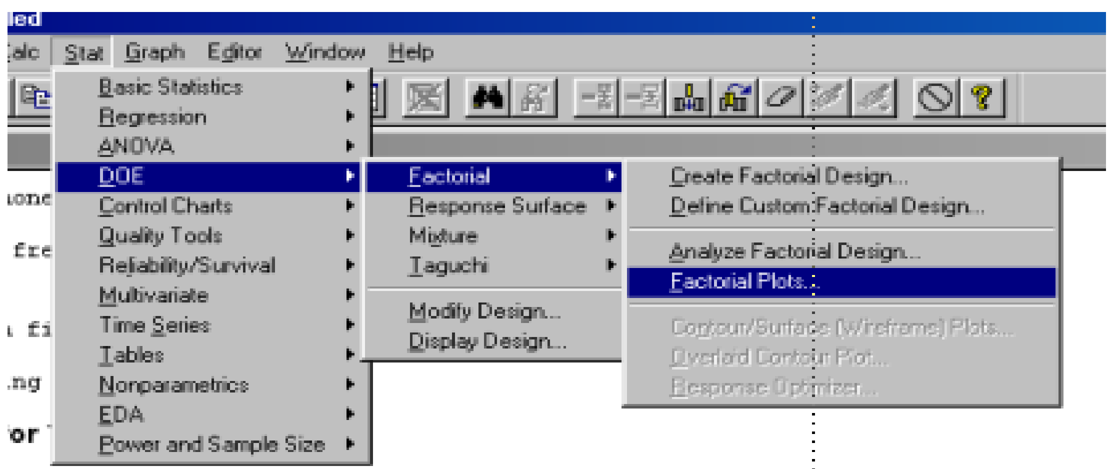
Tool Bar Menu > Stat > DOE > Factorial > Factorial Plots

Create a 2^2 coded design for factors A and B.

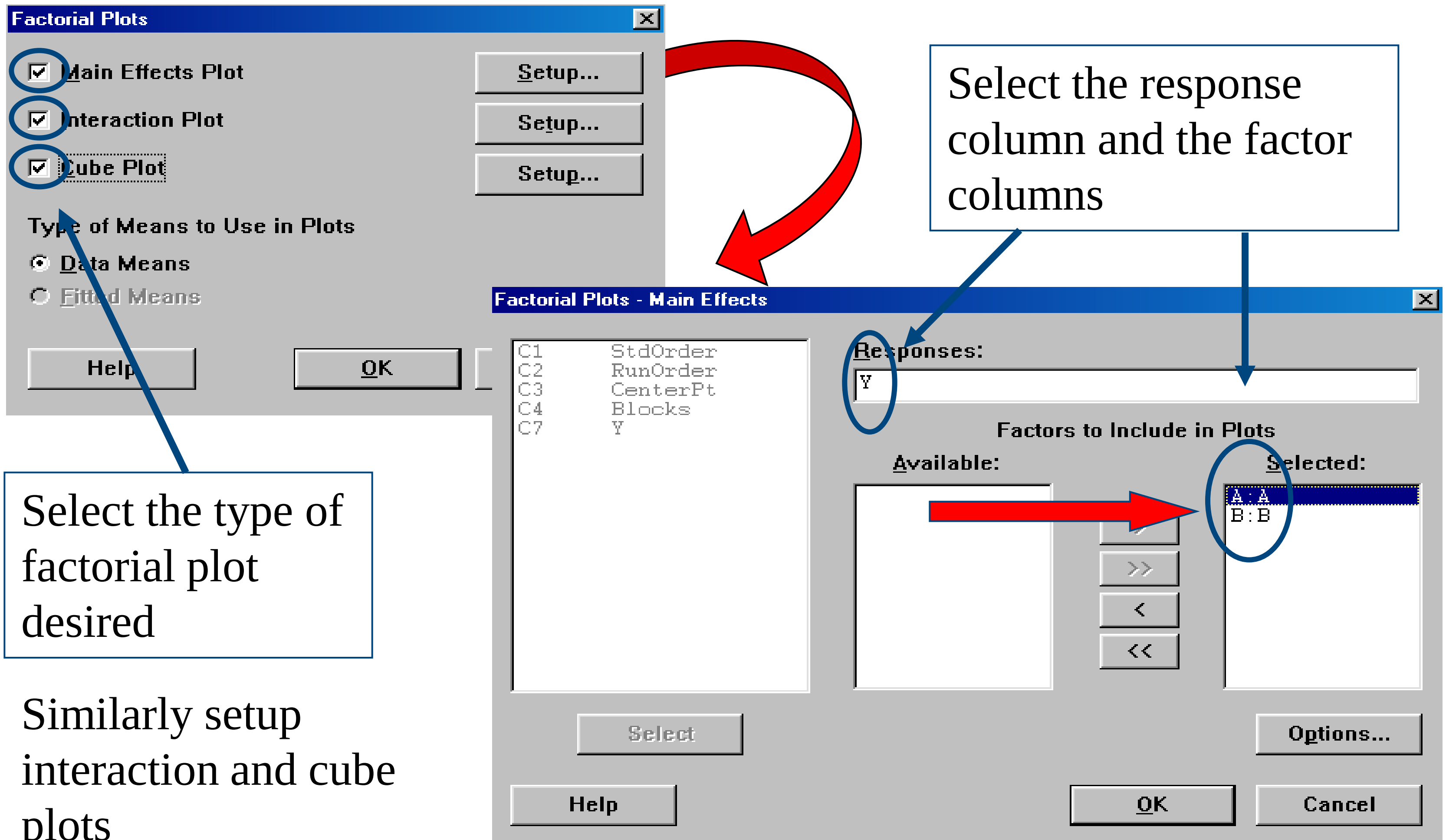
Worksheet 4 ***

↓	C1	C2	C3	C4	C5	C6	C7
	StdOrder	RunOrder	CenterPt	Blocks	A	B	Y
1	1	1	1	1	-1	-1	48
2	2	2	1	1	1	-1	96
3	3	3	1	1	-1	1	72
4	4	4	1	1	1	1	36
5							

Input the Y response



Factorial Plots in Minitab
Step 1

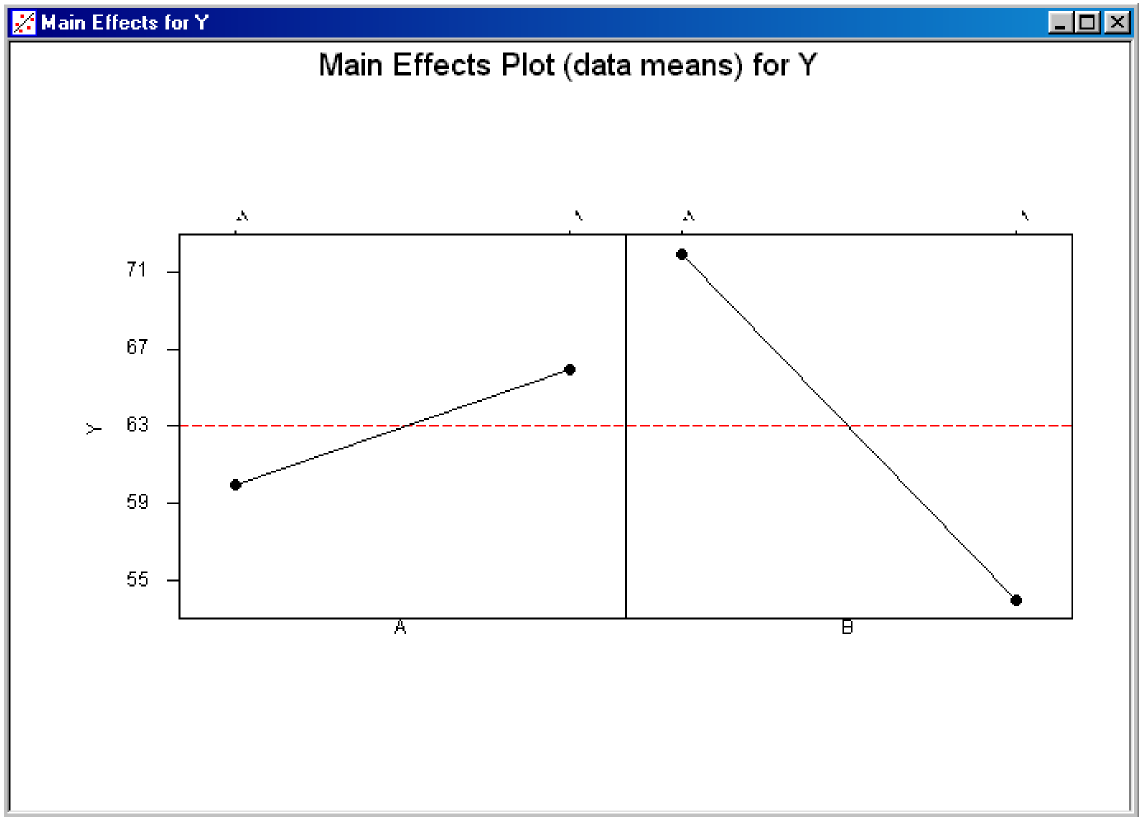


Select the response column and the factor columns

Select the type of factorial plot desired

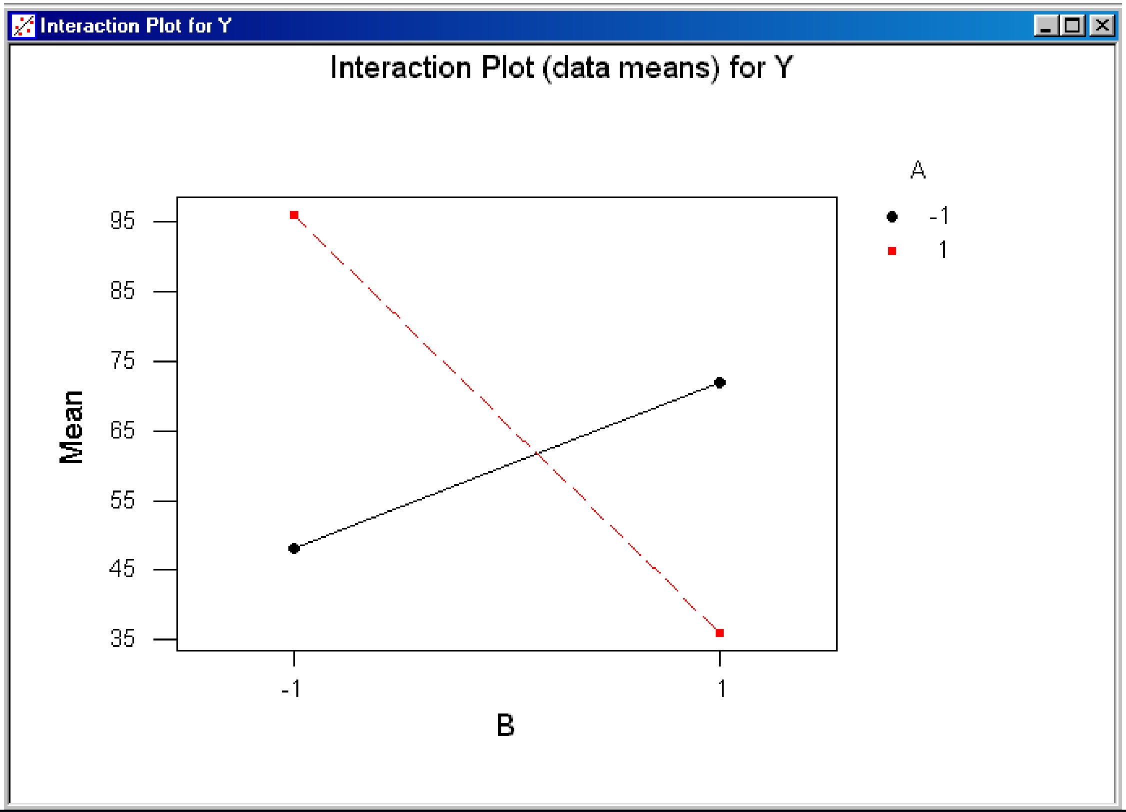
Similarly setup interaction and cube plots

Main Effects Plot



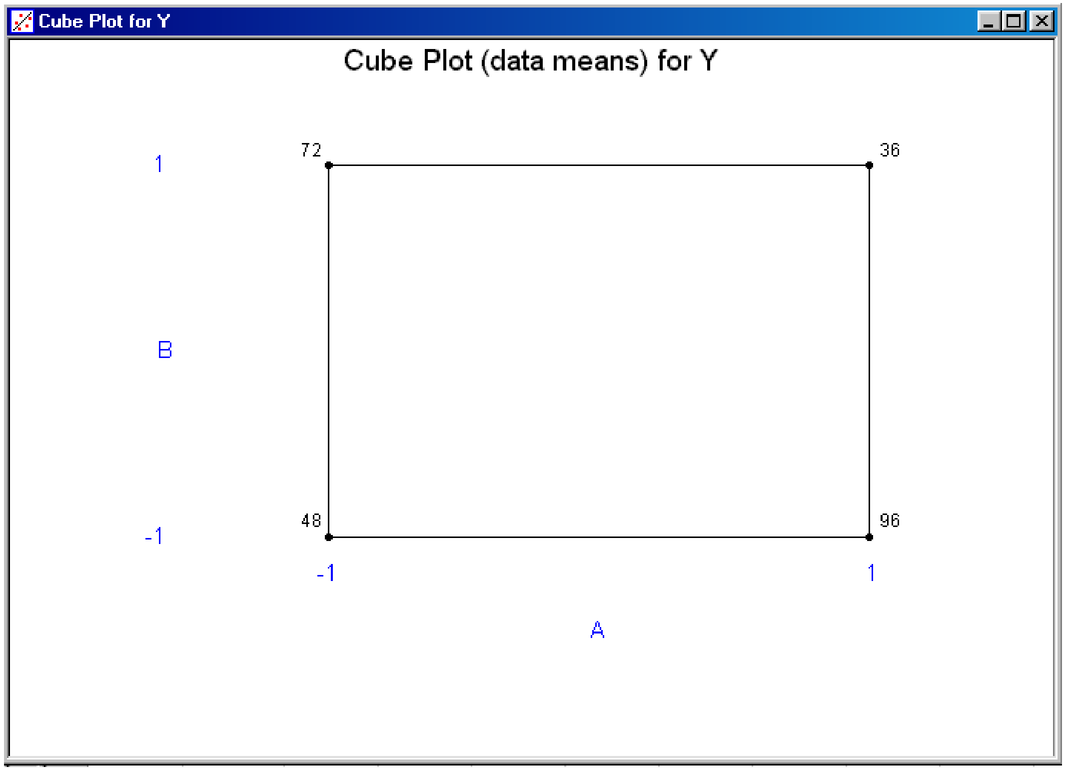
The response is plotted by factor from low level to high level

Interaction Plot



Parallel lines indicate no interaction; the less parallel, the higher the degree of interaction

Cube Plot



The response is plotted on the orthogonal factor axis

Class Exercise

Using the Minitab file to create all three factorial plots.

Be prepared to discuss your results.

Chapter 4:
Project closure

Goal

- Identify ways to bring a project to closure in your own organization

Importance of Closure

- Recognize the considerable time and effort that went into the initiative
- Capture the learnings from the initiative:
  - About the problem or process being studied
  - About the improvement process itself
- Hand over responsibilities for standardization and monitoring to the appropriate people

Project Closure

- Improvement must be continuous, but individual initiatives and project teams come to an end
- Learn when it is time to say good-bye
- Effective project closure weaves together the themes of:
  - Project purpose
  - Improvement methods
  - Team skills and structures
- Develop managerial systems to capture learnings and enable the organization to address system issues
- Documentation and recognition are two critical aspects of project team closure

Exercise: Closure

Objective: Build awareness of how to successfully bring a project to a close

Instructions:

- Work in pairs or small groups to answer the following questions:
  - What would indicate that a team had ended well? What would indicate
    that a team had not formally closed or had not done it well? Explain
  - Think of teams that have completed their work—what methods of closure were used by these teams? How effective were they? How do you know?
- Be prepared to share your answers with the whole group

Time: 10 minutes

Closure Checklist

6 steps to close a Project: Project Closure Process

1. Avoid needless continuation
2. Summarize learnings:
   - About the work process
   - About the team's process
   - About your results
3. Finalize documentation on improvements
4. Summarize future plans and recommendations
5. Communicate the ending
6. Celebrate

1. Avoiding Needless Continuation

2. Summarizing Learnings

- About your results
- About the work process
- About the team’s process

3. Finalize Documentation on Improvements

- Finish your storyboard:
  - It should contain your final results and conclusions
- Present the completed document to:
  - The sponsors
  - The people whose jobs are changing as a result of the work
  - Customers of the change
  - Other interested people
  - Collect, catalog, and make the documentation available to others

4. Summarize Future Plans and Recommendations

- Have your team discuss the following issues and compile recommendations that you submit to your sponsor or guidance team:
  - What are your recommendations for maintaining the gains already put in place? What role would you like your team to play?
  - How much improvement is still needed to achieve the business goals?
  - What portions of the problem are still left to address? Which of these are the most urgent to address?
  - What would you and your team like to work on next, if approved by management? Where do you think management should devote resources next?

5. Communicate the Ending

- Communicating the team's results is a joint task for the project team and its sponsor or guidance team:
  - If not done already, identify the people who will be involved in implementing the improved methods
  - Which employees could benefit from your lessons learned?
  - How can you communicate to management? To the rest of the organization?
  - How can the end of this project sow the seeds for future projects?

6. Celebration and Recognition

- What is the appropriate way to celebrate this closure (lunch/dinner/dessert)?
- How will you say good-bye?

6. Celebration and Recognition, cont.

- Recognition is an important aspect of celebration and should reinforce intrinsic sources of satisfaction and motivation:
  - Examples:
    - A small souvenir related to the project work
    - A pizza lunch for everyone who was involved in the initiative, including those whose jobs have changed as a result of the initiative

Exercise: Closure in Your Organization

Objective: Understand what parts of closure may be most challenging in your organization

Instructions:

- Work alone or in small groups
- Review at least one of the closure checklist items and discuss how it would work in your organization
- Be prepared to discuss your results with the whole group

Time: 10 minutes

Presentation exercise

- Please present the topic allotted to you

Thank you!

Question time

Stay Connected!

- You can subscribe to Anexas Europe YouTube channel to get access to all the training material on video. https://m.youtube.com/channel/UCAk6lJsnPCsz7-rZTVx90uw
- Please like our page on Facebook to stay updated about the latest videos on Training, Quality, Process Excellence and Project management https://www.facebook.com/denmarkanexas

Stay Connected!

- Please join Anexas Europe LinkedIn Page for updates about latest developments https://www.linkedin.com/school/9379319/
- Connect with the trainer– Amitabh Saxena on his LinkedIn account to see his latest videos on training quality topics https://www.linkedin.com/in/amitabhsaxena/

Thank You!

- Congratulations on completing a milestone in your life!
- Best wishes for your MBB journey!

- Amitabh Saxena

Trainer