



LEAN
SOLUTIONS
ACADEMY

A group of five professionals (three men and two women) in business attire are gathered around a table, smiling and looking at a laptop screen. They appear to be in a collaborative meeting or training session. The image is partially obscured by blue geometric overlays.

Lean Six Sigma Training **Black Belt**

brought to you by:



**PATRICK
ADAMS**
CONSULTING

Measure Phase

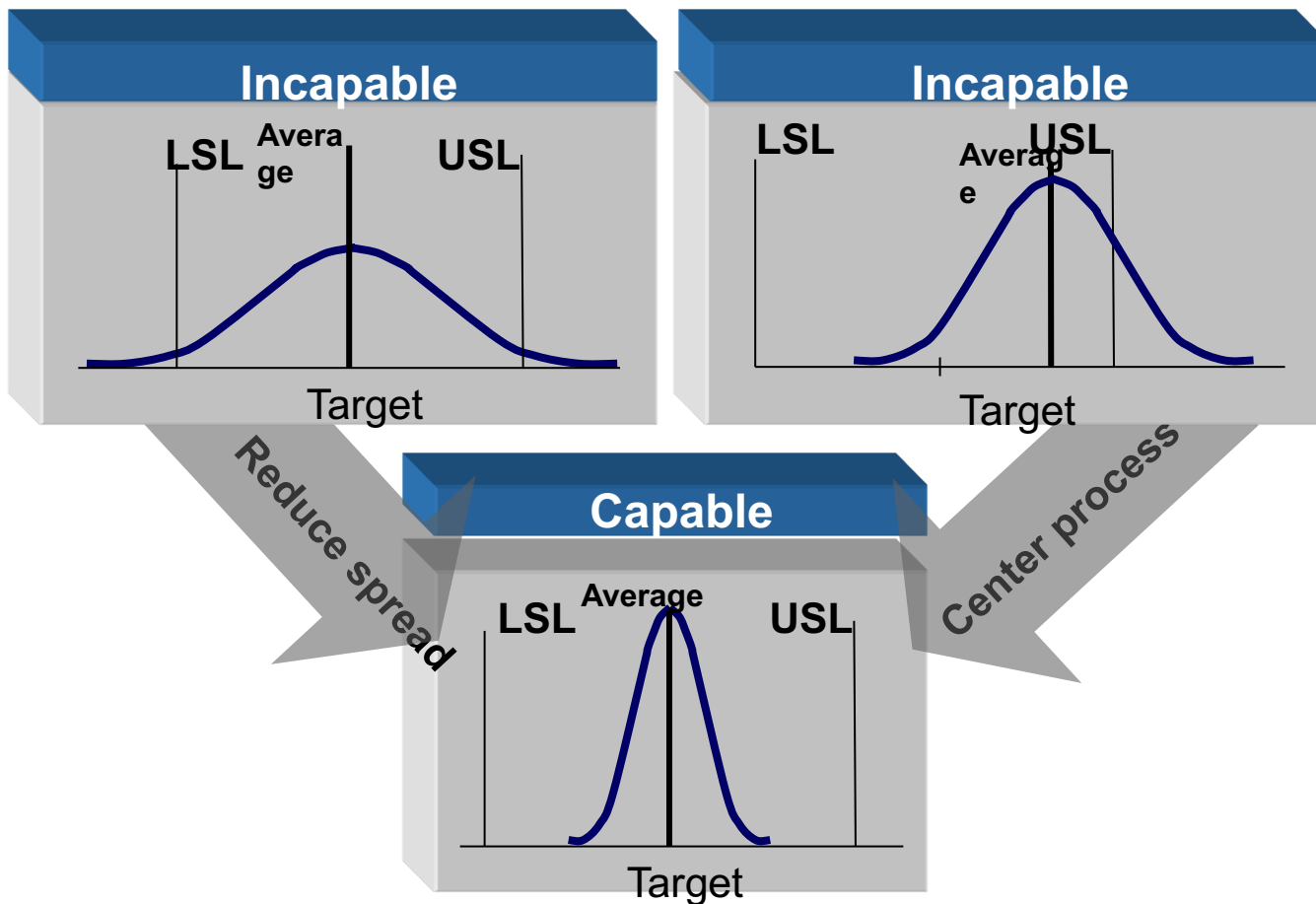
Process Capability

Understanding Process Capability

Process Capability / Process Performance :

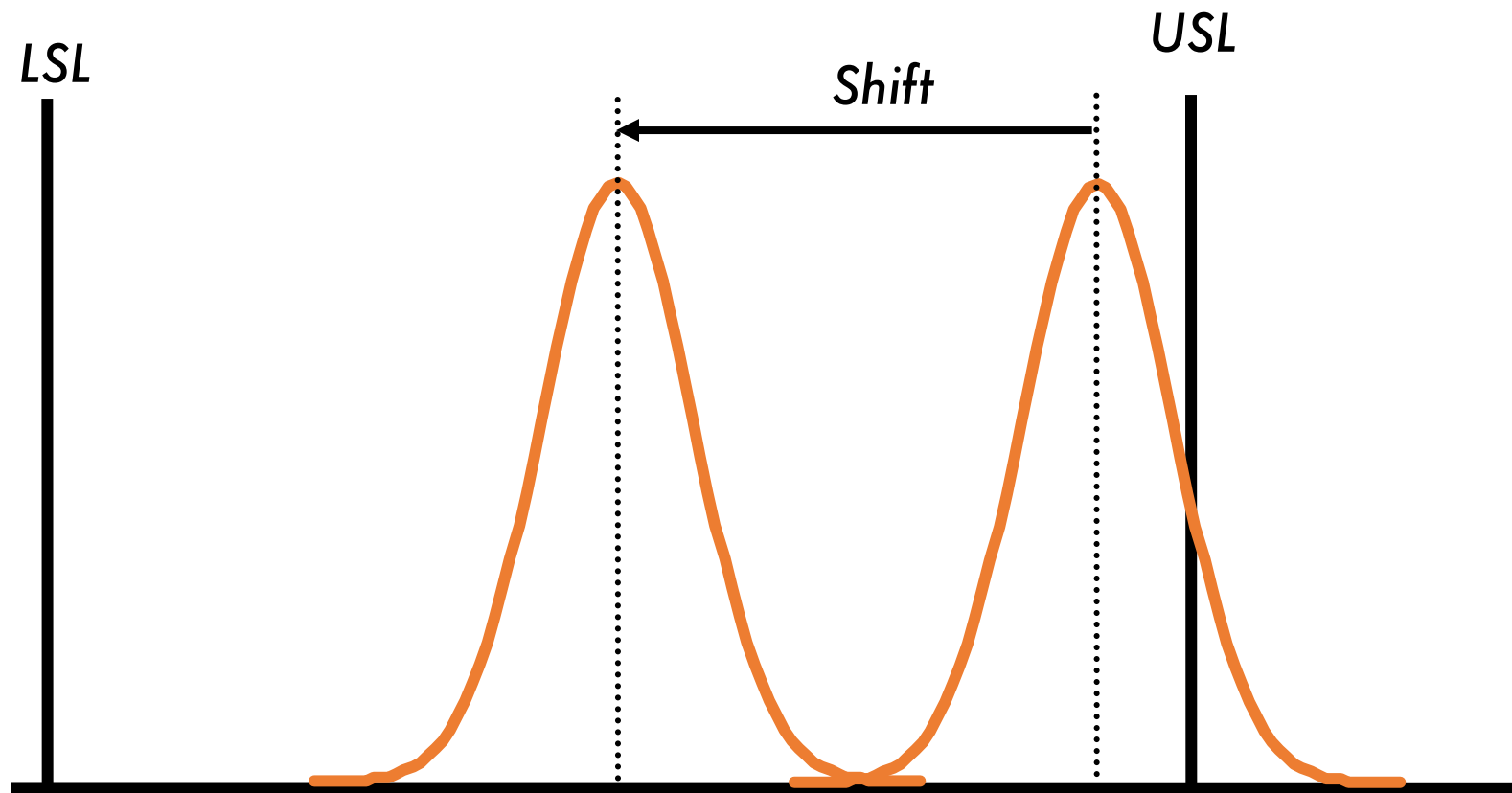
- The ability of a process to meet the expectations of the customer
- Provides insight as to whether the process has a:
 - Centering Issue (relative to specification limits)
 - Variation Issue
 - A combination of Centering and Variation
 - Inappropriate specification limits
- Provides a baseline metric for improvement.

Capability Defined



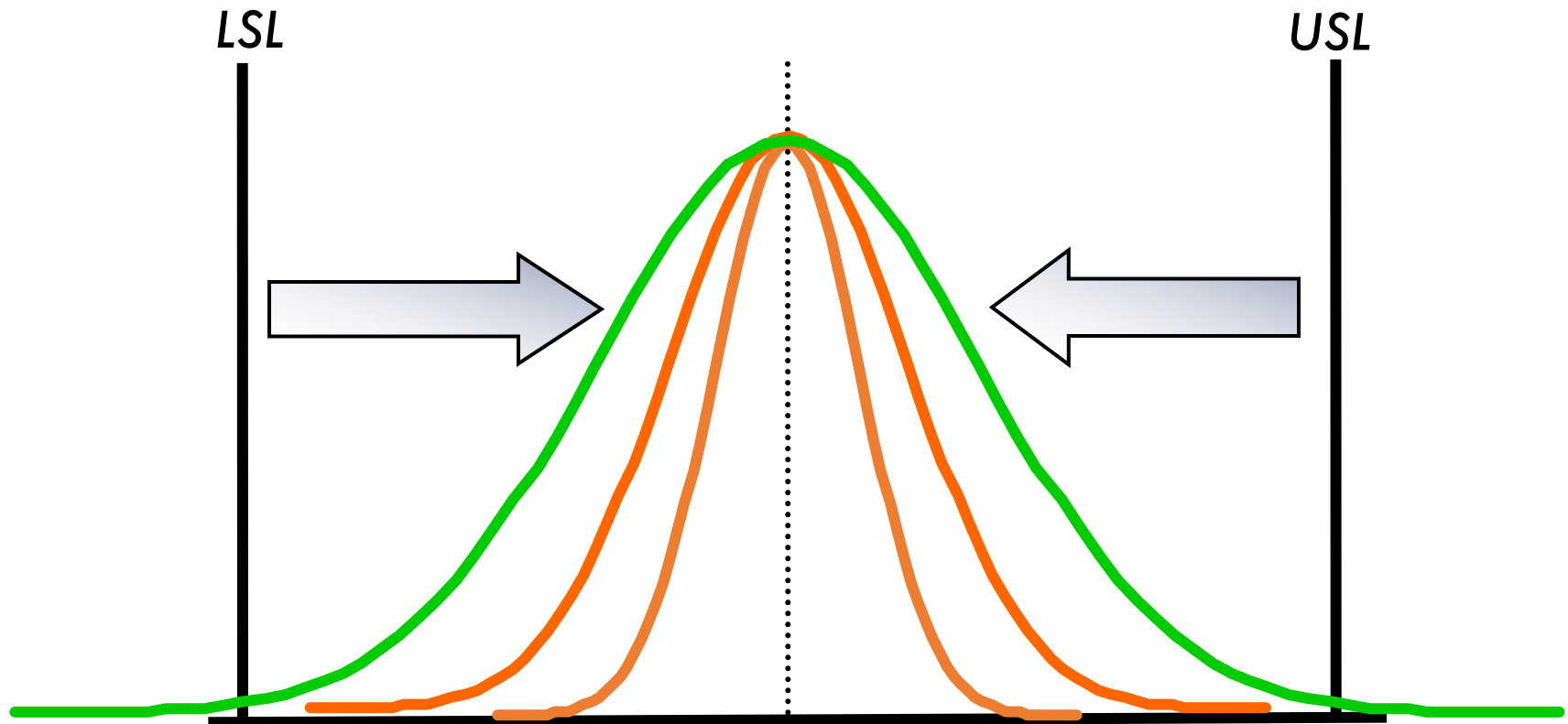
Step 1 – Shift the Mean

This involves finding the variables that will shift the process to the target. This is what is done first.



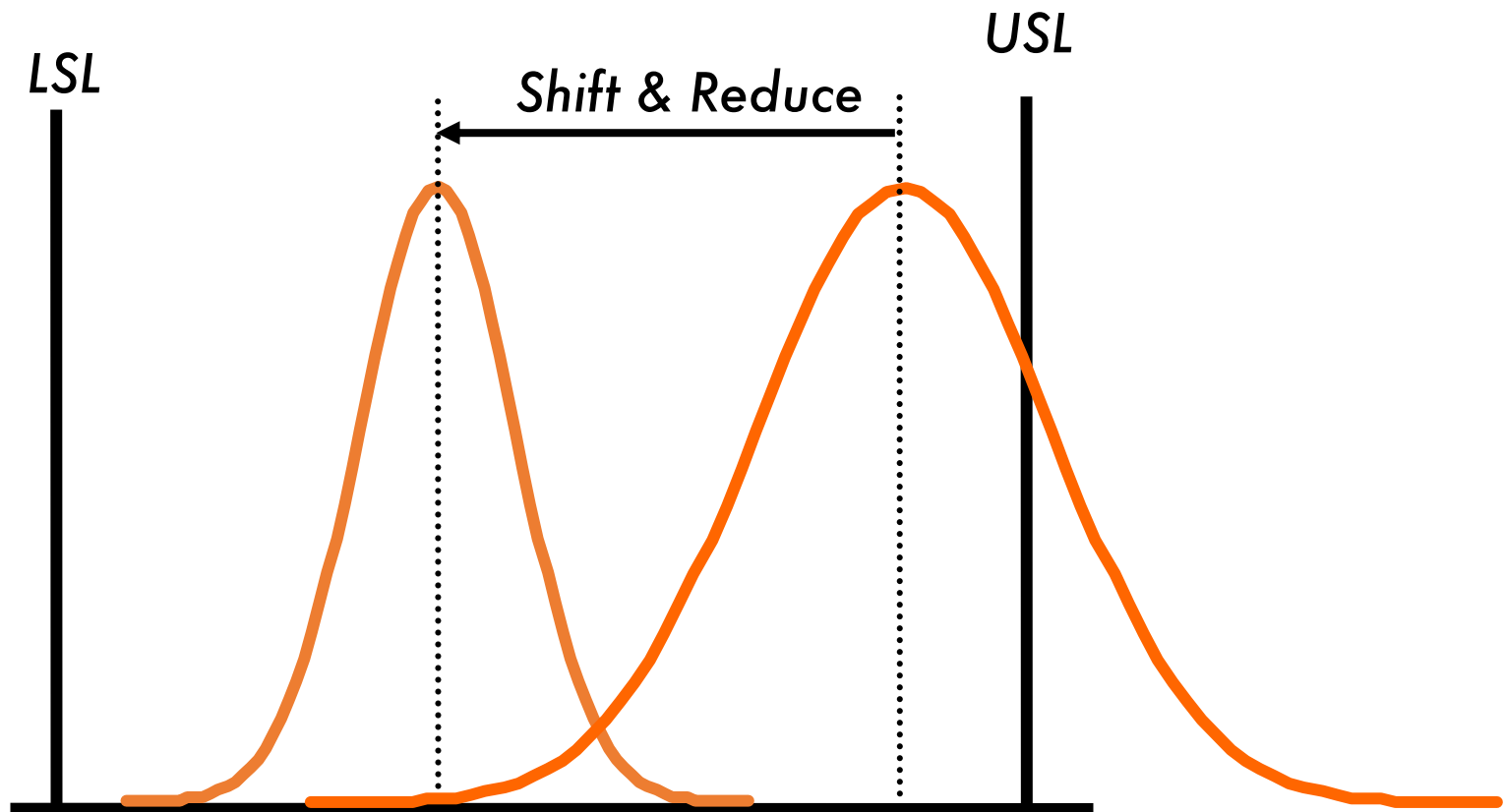
Step 2 – Reduce Variation

Reducing Variation is typically not so easy to accomplish and is a main focus of Six Sigma projects.



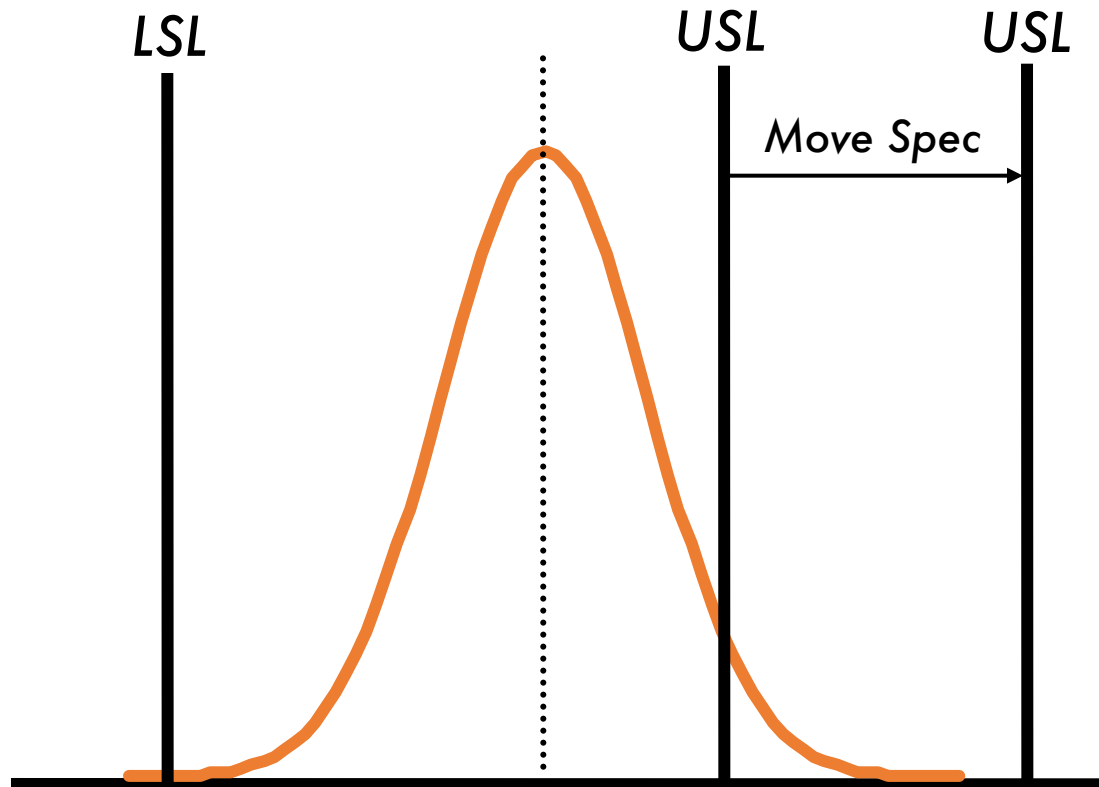
Shift Mean & Reduce Variation

This is the aim of most Six Sigma projects.



Moving Customer Specification limits

Customers usually do not go for this option but if they do...it is the easiest way to increase capability !



Capability Studies

Capability Studies:

- Are intended to be regular, periodic, estimations of a process's ability to meet customer requirements.
- Can be conducted on both Variable and Attribute Data.
- Are most meaningful when conducted on stable, predictable processes.
- Are commonly reported as either
 - Cp
 - Cpk
 - Pp
 - Ppk
 - Process Sigma Level
 - Percent of Defectives
 - DPMO
 - DPU

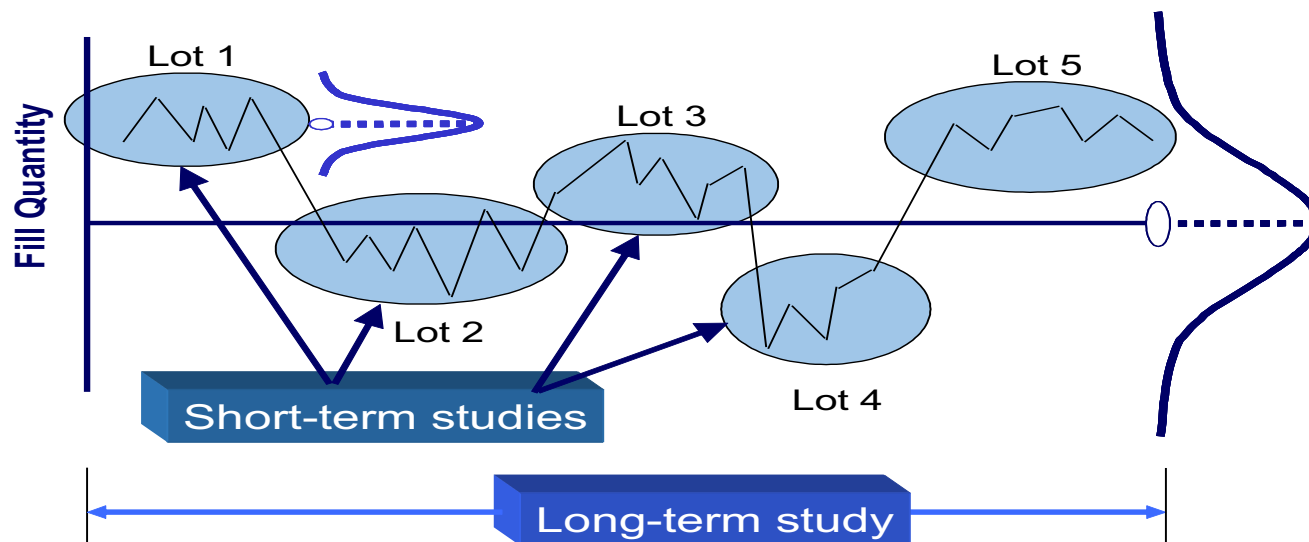
Data Collection

Short-term data: (C_p and C_{pk})

- Collected across a narrow inference space.
 - Daily, weekly; for one shift, machine, operator, etc.
- Is potentially free of Special Cause variation.
- Often reflects the optimal performance level.
- Typically consists of 30 – 50 data points.

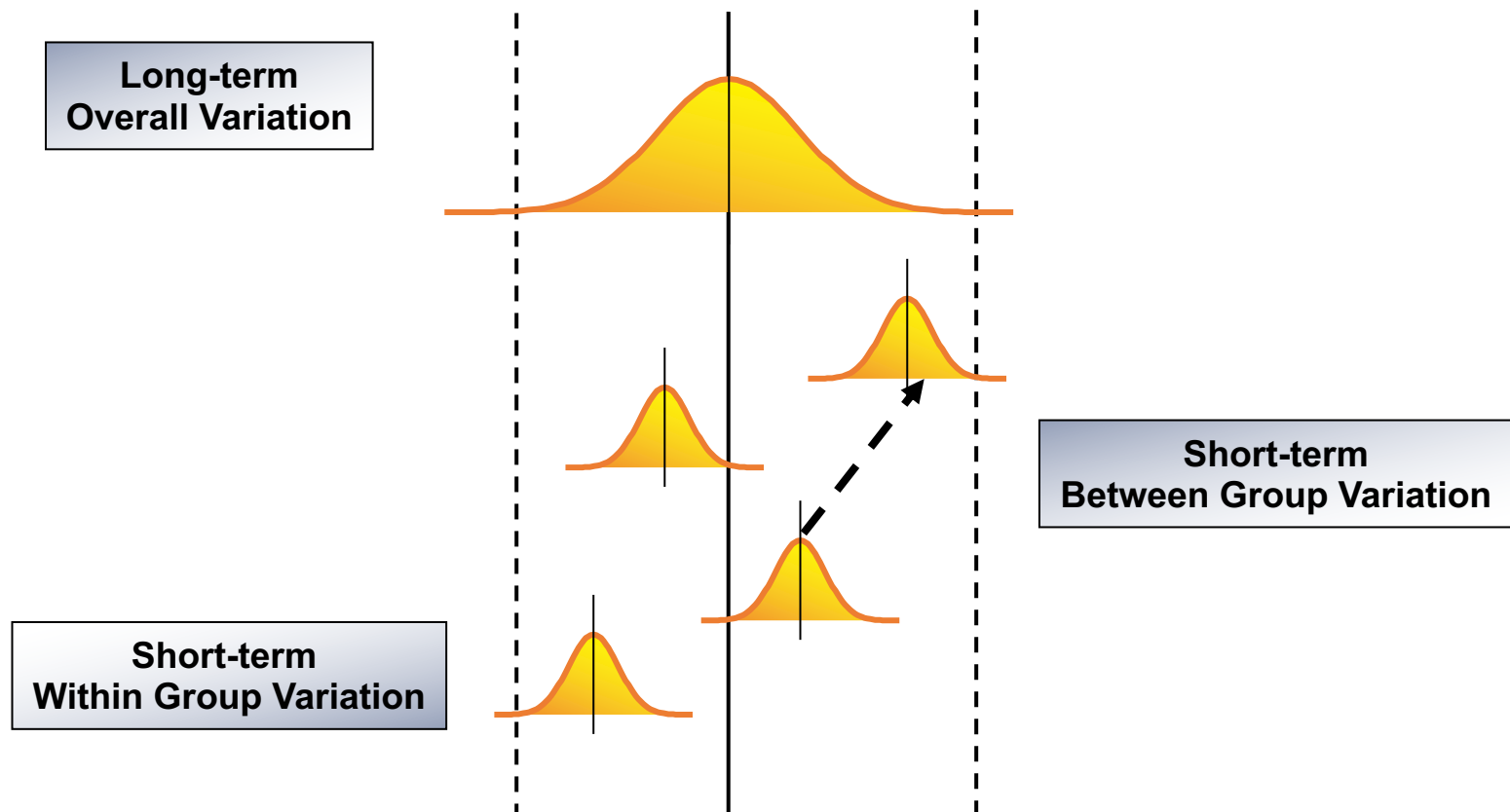
Long-term data: (P_p and P_{pk})

- Is collected across a broader inference space.
 - Monthly, quarterly; across multiple shifts, machines, operators, etc
- Subject to both Common and Special Causes of variation.
- More representative of process performance over a period of time.
- Typically consists of at least 100 – 200 data points.



Components of Variation

Many processes will drift and shift over time by as much as 1.5 Standard Deviations.



Measures of Capability

Potential



Cp and Pp

- What is Possible if your process is perfectly Centered
- The Best your process can be
- Process Potential (Entitlement)

Reality



Cpk and Ppk

- The Reality of your process performance
- How the process is actually running
- Process Capability relative to specification limits

Difference between Cp,Cpk, Pp and Ppk

Cp and Cpk (Process Capability)

Uses Estimated Standard Deviation calculated from the mean of standard deviations of short term samples

- This “Average over Samples” tend to “smooth” the variation and so the capability appears slightly better
- Results in a slightly higher capability estimate than Pp or Ppk

Pp and Ppk (Process Performance)

Uses the actual Standard Deviation calculated from all the individual measurements over long term

- This usually results in a higher standard deviation (variation)
- Result is a lower (more conservative) capability number.

Capability Formulas

$$C_p \text{ or } P_p = \frac{(USL - LSL)}{6s}$$

Six times the sample
Standard Deviation

$$C_{pk} \text{ or } P_{pk} = \frac{(\text{Closest Spec} - \bar{X})}{3s}$$

Sample Mean

Three times the sample
Standard Deviation

LSL – Lower specification limit

USL – Upper specification limit

s – Standard Deviation

Closest Spec = Closest Specification Limit to the Mean

Camshaft Supplier Capability example

- ABC Engine Manufacturing Ltd buys its camshafts for its engines from 2 different suppliers.
- Supplier 1 and Supplier 2
- ABC wants to know what the CAPABILITY is of each supplier to meet its requirements for length of the camshaft.
- The requirements set by ABC are defined by its Specification Limits



LSL = 598 mm

USL = 602 mm

A CAPABILITY STUDY is run in order to investigate.

SigmaXL Example

Create a **Capability Analysis** for the 2 suppliers of Camshafts to compare Capabilities.
Use the worksheet Camshaft.xlsx.

Note the subgroup size for this example is 5.

LSL=598 mm USL=602 mm

Process Capability > Capability Combination Report (Subgroups)

Supplier 1

Supplier 2

☒ Stacked Column Format (1 Numeric Data Column & Subgroup Size or Column)
☐ Subgroups across Rows (2 or More Numeric Data Columns)

Numeric Data Variable (Y) >>

Supplier 1

OK >>

Subgroup Column or Size >>

5

Cancel

Optional X-Axis Labels >>

Help

<< Remove

☒ X-Bar & R Charts
☐ X-Bar & S Charts

☒ Calculate Control Limits
☐ Historical Control Limits

☐ Tests for Special Causes

X-Bar

R

UCL

CL

LCL

Spec Limits for Capability

USL

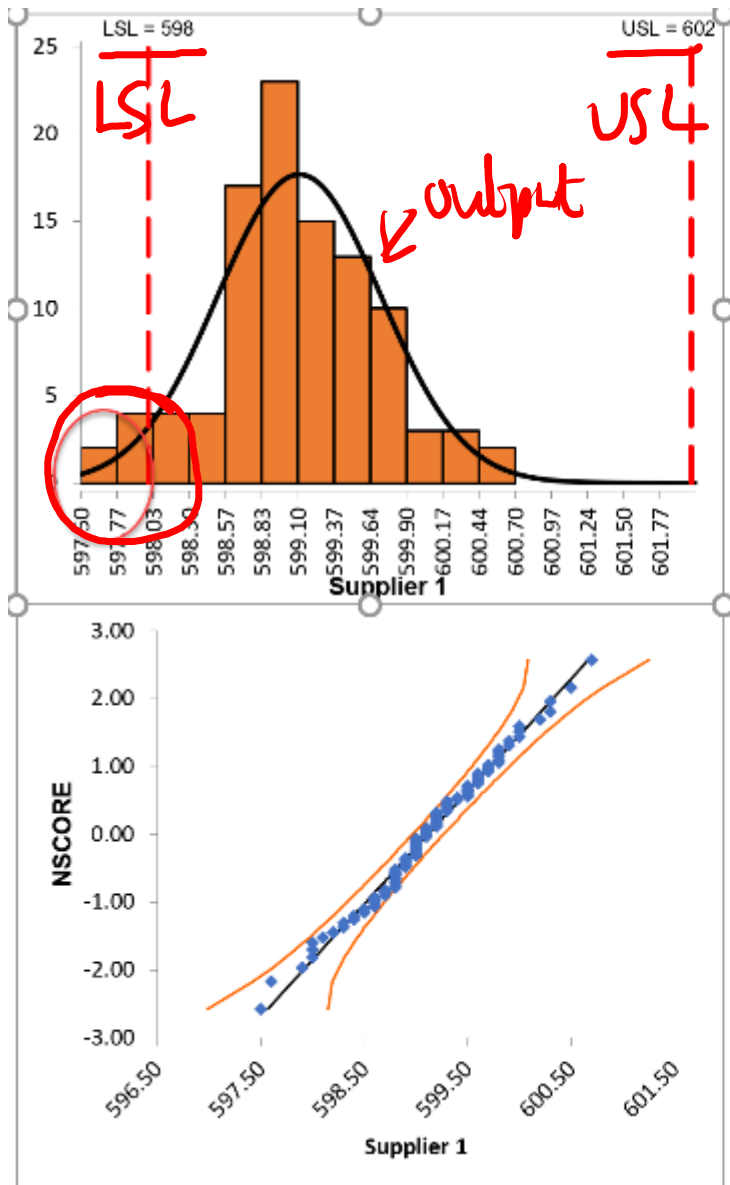
602

Target

LSL

598

SigmaXL Example



Process Capability Report: Supplier 1	
Count	100
Mean	599.12
StDev (Overall, Long Term)	0.602583
StDev (Within, Short Term)	0.563199
USL	602
Target	
LSL	598

Capability Indices using Overall StDev	
Pp	1.11
Ppu	1.60
Ppl	0.62
Ppk	0.62
Cpm	

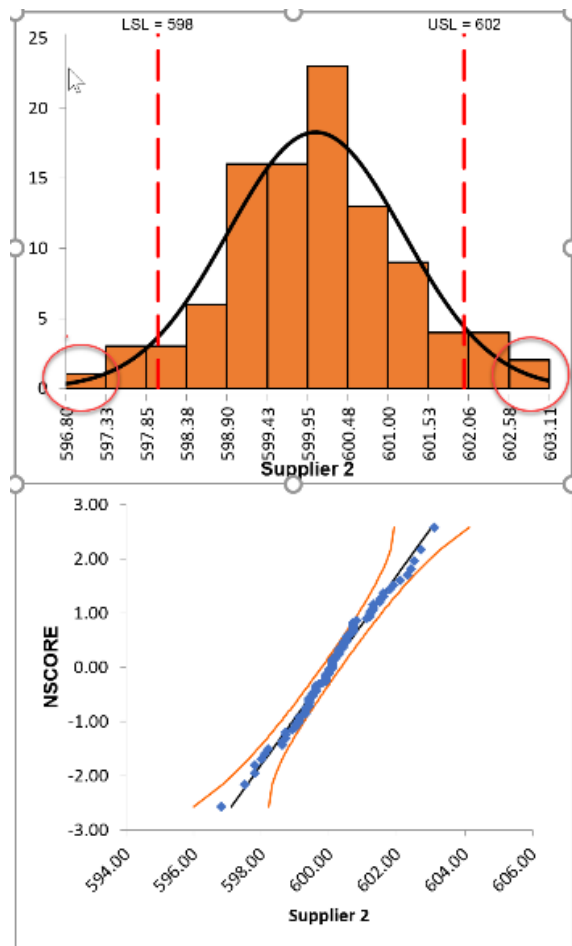
Potential Capability Indices using Within StDev	
Cp	1.18
Cpu	1.71
Cpl	0.66
Cpk	0.66

Expected Overall Performance	
ppm > USL	0.8
ppm < LSL	32130.3
ppm Total	32131
% > USL	0.00%
% < LSL	3.21%
% Total	3.21%

Actual (Empirical) Performance	
% > USL	
% < LSL	3.00%
% Total	3.00%

0 → 1 → 2
 ↑
 Min acceptable

SigmaXL Example



Process Capability Report: Supplier 2	
Count	100
Mean	600.06
StDev (Overall, Long Term)	1.146
StDev (Within, Short Term)	0.988822
USL	602
Target	
LSL	598

Capability Indices using Overall StDev	
Pp	0.58
Ppu	0.56
Ppl	0.60
Ppk	0.56
Cpm	

Potential Capability Indices using Within StDev	
Cp	0.67
Cpu	0.65
Cpl	0.69
Cpk	0.65

Expected Overall Performance	
ppm > USL	45337.8
ppm < LSL	36065.2
ppm Total	81403
% > USL	4.53%
% < LSL	3.61%
% Total	8.14%

Actual (Empirical) Performance	
% > USL	6.00%
% < LSL	4.00%
% Total	10.00%

Summary

At this point you should be able to:

- Understand what is Capability Analysis
- Select the appropriate method for Capability Analysis based on the type of data and the data distribution of your process
- Run a Capability Study on Normal Data
- Interpret the output of SigmaXL Capability functions
- Understand the difference between C_p , C_{pk} , P_p , P_{pk}

Measure Phase Overview - The Goal

Remember The goal of the Measure Phase was to:

- Define, explore and classify “X” variables using a variety of tools.
 - Detailed Process Mapping
 - Fishbone Diagrams
 - X-Y Matrixes
 - FMEA
- Acquire a working knowledge of Basic Statistics to use as a communication tool and a basis for inference.
- Understand the concepts of Measurement System Analysis
- Perform Measurement Capability studies on output variables.
- Evaluate the “As-Is” Capability of the process.

Measure Phase Deliverables

Listed here are the **Measure Deliverables you should have :**

- Primary Metric
- Secondary Metric(s)
- Process Map – detailed
- Fish Bone Diagram
- Pareto Analysis
- FMEA
- X-Y Matrix
- Basic Statistics on Y's and X's
- MSA
- Capability Analysis

Measure Phase Overview - The Goal

Remember The goal of the Measure Phase was to:

- Define, explore and classify “X” variables using a variety of tools.
 - Detailed Process Mapping
 - Fishbone Diagrams
 - X-Y Matrixes
 - FMEA
- Acquire a working knowledge of Basic Statistics to use as a communication tool and a basis for inference.
- Understand the concepts of Measurement System Analysis
- Perform Measurement Capability studies on output variables.
- Evaluate the “As-Is” Capability of the process.

Measure Phase Deliverables

Listed here are the **Measure Deliverables you should have :**

- Primary Metric
- Secondary Metric(s)
- Process Map – detailed
- Fish Bone Diagram
- Pareto Analysis
- FMEA
- X-Y Matrix
- Basic Statistics on Y's and X's
- MSA
- Capability Analysis

Analyze Phase

The Analyze Phase

*The Analyze Phase is for determining the **ROOT CAUSE** of the problem that we are trying to solve and the optimal settings for our input variable (X) to produce the optimal values of our outputs (Y).*

The Analyze Phase uses many statistical tools and concepts such as:

- Value Stream Mapping for the 'To-Be' future state
- Understanding Correlations between Inputs and Outputs
- Regression Analysis
- Designed Experiments (DOE) for optimization of processes
- Hypothesis Testing

The “I-P-O” model for your business



$X's$

Transformation

Y

X_n

$f(x_n)$

Focus of Six Sigma on the Inputs and the process

$$Y = f(x_n)$$

Y

- Dependent Variable
 - Output
 - Effect
 - Symptom
 - Response

X1 . . . X_n

- Independent Variable
 - Input
 - Cause
- Root of the Problem
 - Factor
 - Predictor

Analyze Phase

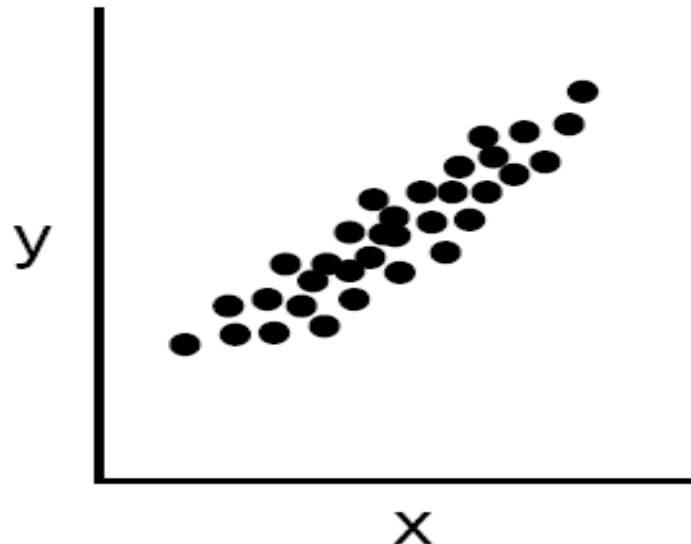
Correlation and Regression

Part 1

Scatter Diagrams

Purpose:

- To identify the correlations that might exist between a quality characteristic and a factor that might be driving it
- A scatter diagram shows the **correlation** between two variables in a process



Scatter Plot Example

Graphical Tools > Scatter Plots...

Scatter Plots

Numeric Response (Y) >>

PGM concentrate (g/ton)

OK >>

Numeric Predictor (X1) >>

Agitator RPM

Cancel

Group Category (X2) >>

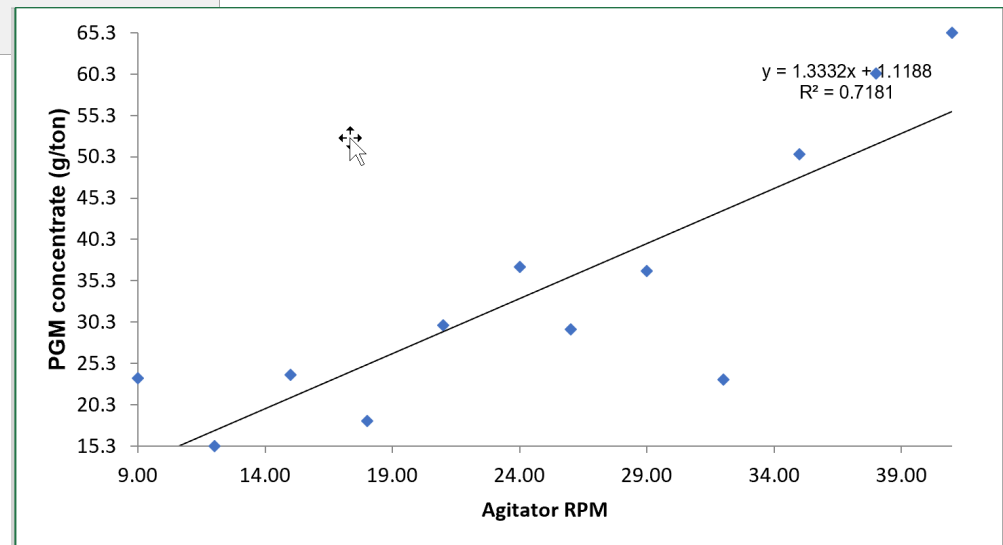
Help

<< Remove

Display Options

☒ Trendline
 ☐ 95% Confidence Interval
 ☐ 95% Prediction Interval

☐ Add Title



Correlation

- The primary purpose of correlation analysis is to measure the strength of association between two variables (X and Y).
- The correlation is positive when Y tends to increase with an increase in X and negative when Y tends to decrease with an increase in X.
- If the ordered pairs (X, Y) tend to follow a straight line path there is a linear correlation.

Correlation Coefficient

The **Correlation Coefficient R** is a value between -1 and $+1$.

R-squared, R-Sq, R^2 is a statistical measure of how close the data are to the fitted regression line.

It is also known as the **coefficient of determination**

0% indicates that the model explains NONE of the variability of the response data around its mean

100% indicates that the model explains ALL of the variability of the response data around its mean

Correlation Example

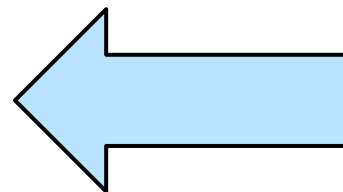
The Correlation Coefficient ***R***

- Is a positive value if one variable increases as the other variable increases.
- Is a negative value if one variable decreases as the other increases.

Pearson's Correlation Formula

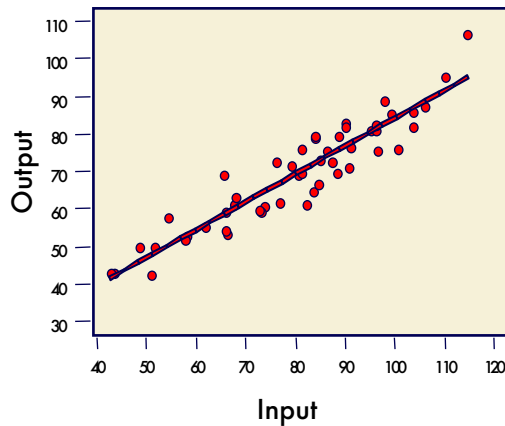
$$R = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma(X_i - \bar{X})^2 \Sigma(Y_i - \bar{Y})^2}}$$

You don't need
To Remember
this or use it in
the exam !!

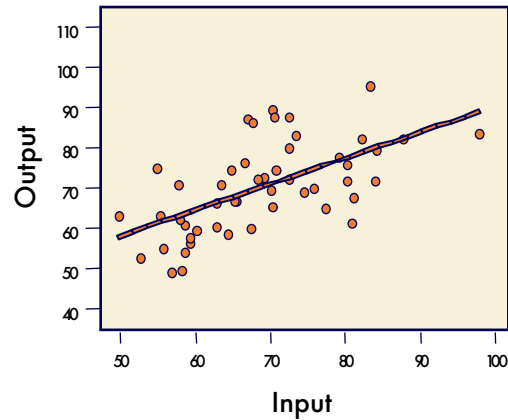


Types and Magnitude of Correlation

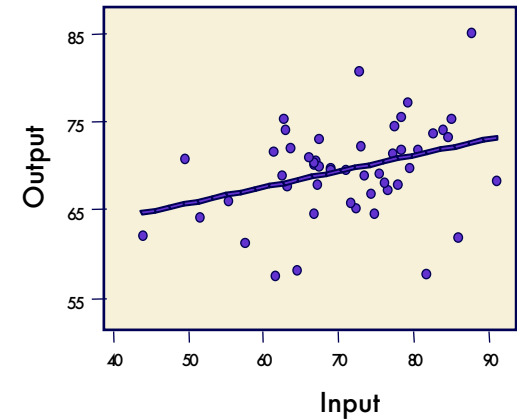
Strong Positive Correlation



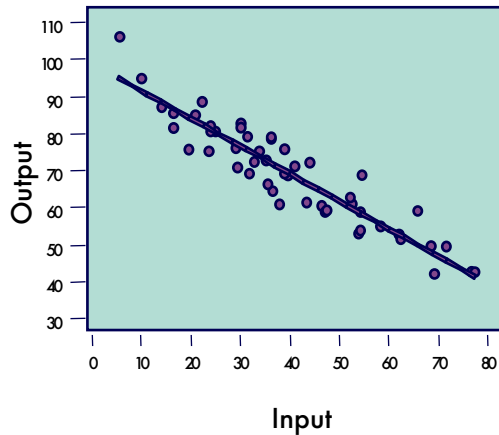
Moderate Positive Correlation



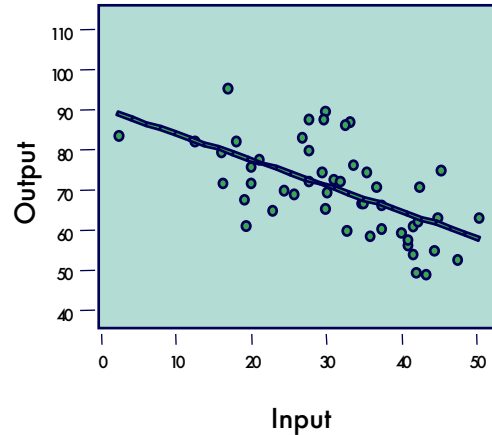
Weak Positive Correlation



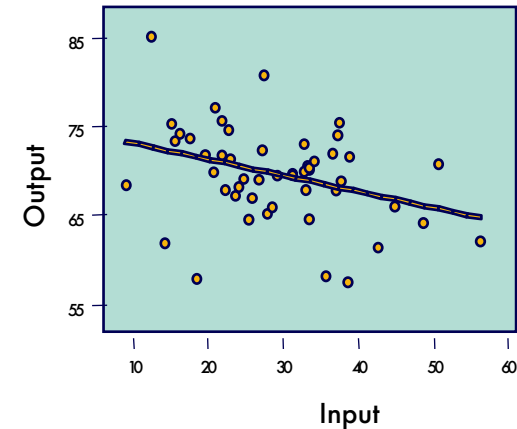
Strong Negative Correlation



Moderate Negative Correlation

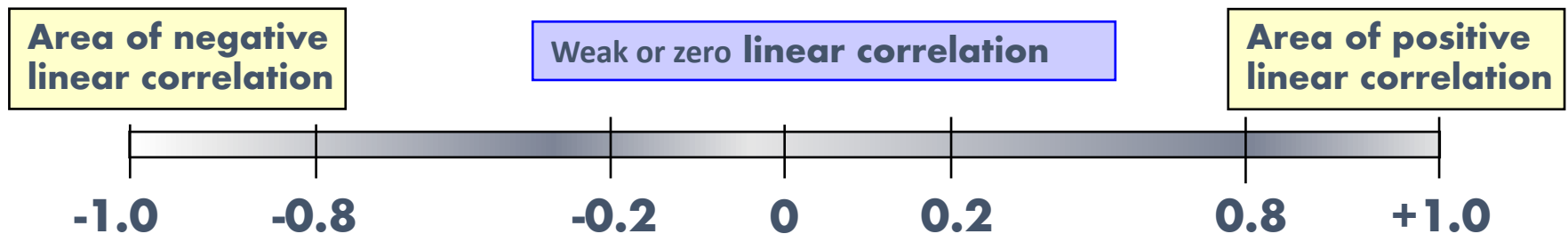


Weak Negative Correlation



Correlation Measures

- Guidelines for correlation strength are as follows:
 - If $R\text{-Sq} > 0.80$ or 80% , relationship is a **Strong Correlation**
 - If $R\text{-Sq} < 0.20$ or 20% relationship is a **weak or zero Correlation**



Cause and Effect

Correlation and regression DO NOT imply a causal relationship between the variables.

- Designed experiments (DOE) allow for true cause and effect relationships to be identified. (Next section)

Which are Cause and Effect vs Non-Cause and Effect ?

Shoe size vs Height?

Cigarette Smoking vs Lung Cancer?

Weight vs Daily Food Intake?

Stress Level vs Age of Death?

Blood Alcohol Level vs Reaction Time?

Call Centre Answering Times Project

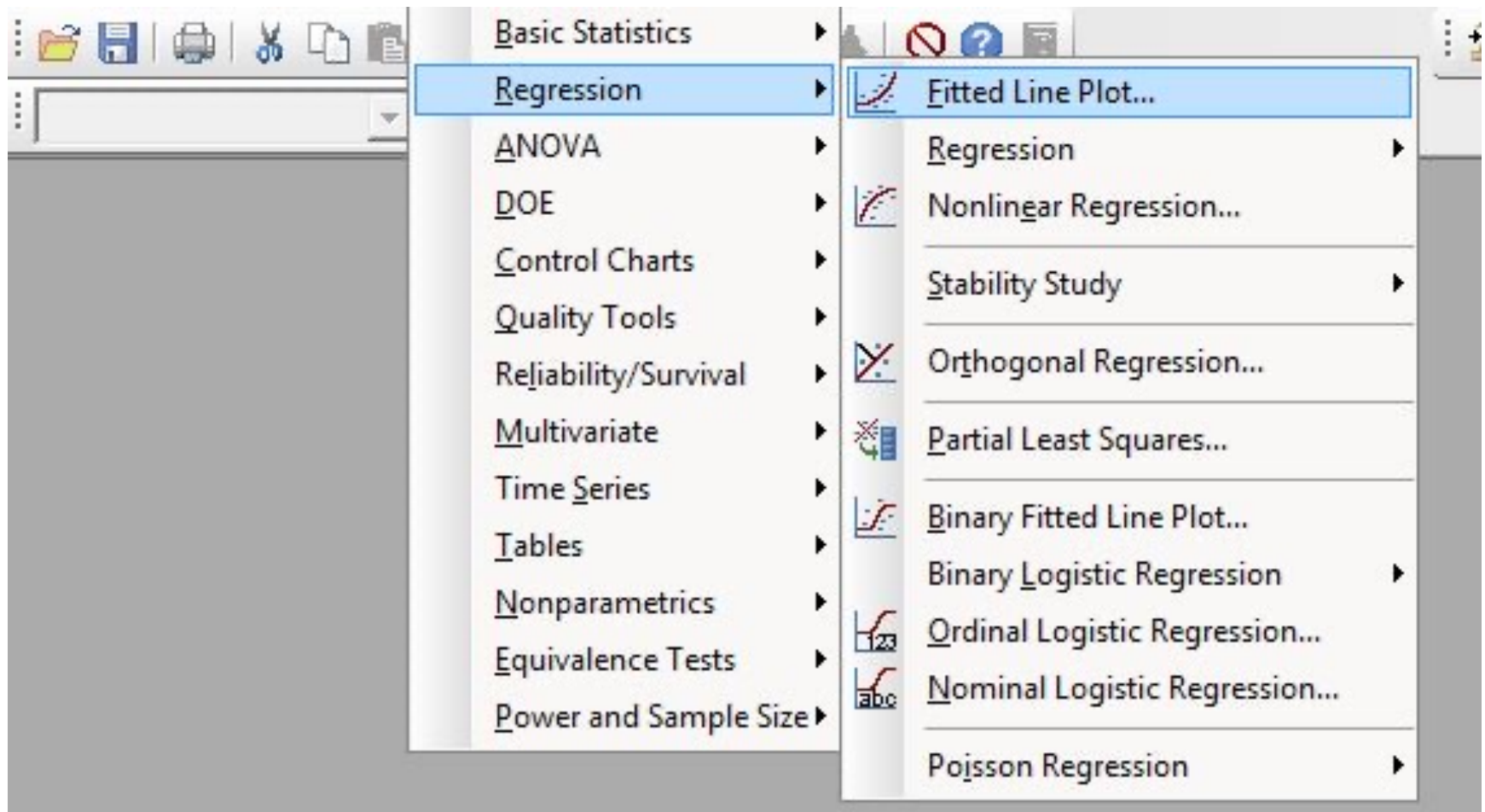
EXAMPLE : Call Answer time (pick-up time) for a Call Centre

- The project is focussing on the answer time for telephone calls to the centre.
- The average answer time was recorded for 80 shifts
- $Y =$ Answer time for calls (Pick-up time)
- $X_1 =$ Number of Personnel
- $X_2 =$ Number of Calls per hour (Average)
- $X_3 =$ Time per call (Average)

Call Answering time Project

↓	C2	C4	C5	C7
	Answer Time (Average) (secs)	Number of personnel	Calls per hour (average)	Time per call (average)
1	26	111	1357	6.1
2	24	114	1265	6.2
3	27	111	1336	5.9
4	16	118	1319	6.7
5	21	115	1290	6.4
6	22	112	1325	5.7
7	15	117	1396	6.9
8	21	116	1263	7.0
9	38	103	1319	5.8
10	48	102	1317	5.8
11	39	107	1358	5.2
12	50	101	1329	5.7
13	18	116	1292	6.6
14	20	114	1378	6.4
15	28	110	1327	6.0

Simple Linear Regression



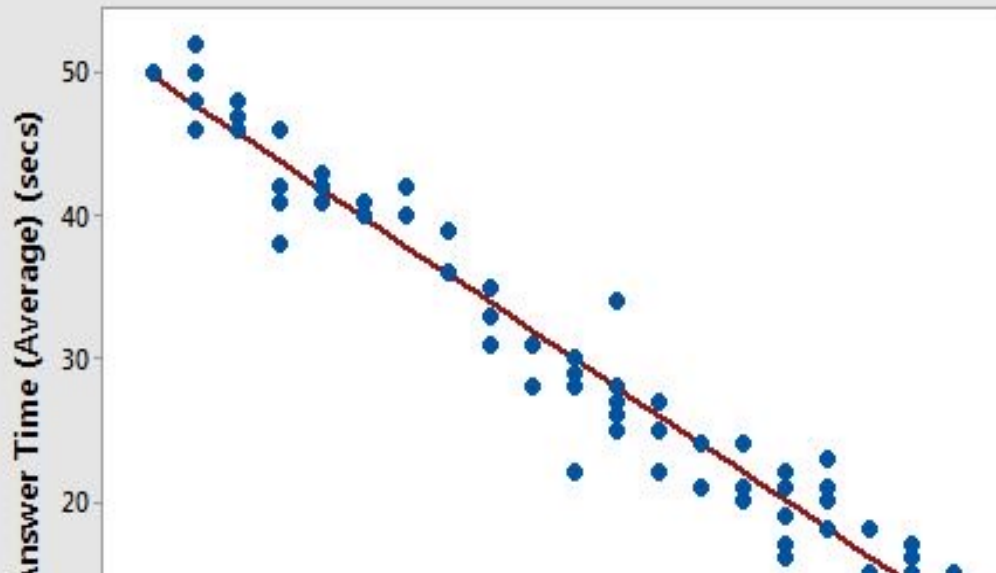
Simple Linear Regression

S = 2.49113 R-Sq = 95.4% R-Sq(adj) = 95.4%

✓ Fitted Line: Answer Time (Average) (secs) versus Number of personnel

Fitted Line Plot

Answer Time (Average) (secs) = 247.9 - 1.981 Number of personnel



S	2.49113
R-Sq	95.4%
R-Sq(adj)	95.4%

The Regression Equation

- The regression equation is :

$$\text{Answer Time} = 247.9 - 1.981 \times \text{Number of Personnel}$$

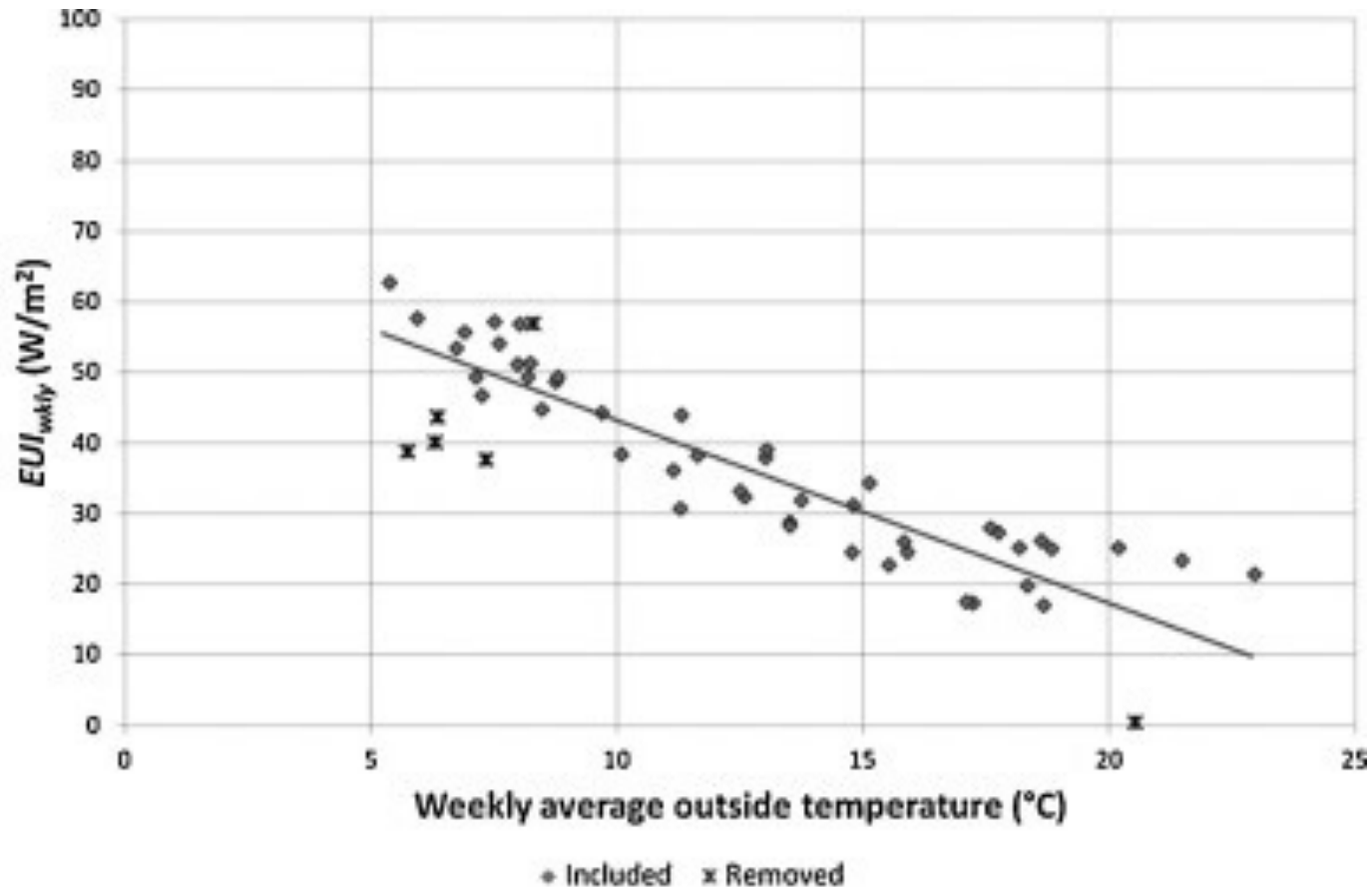
$$R^2 = 95\%$$

- This means that 95% of the change in answer time can be explained by the change in Number of Personnel

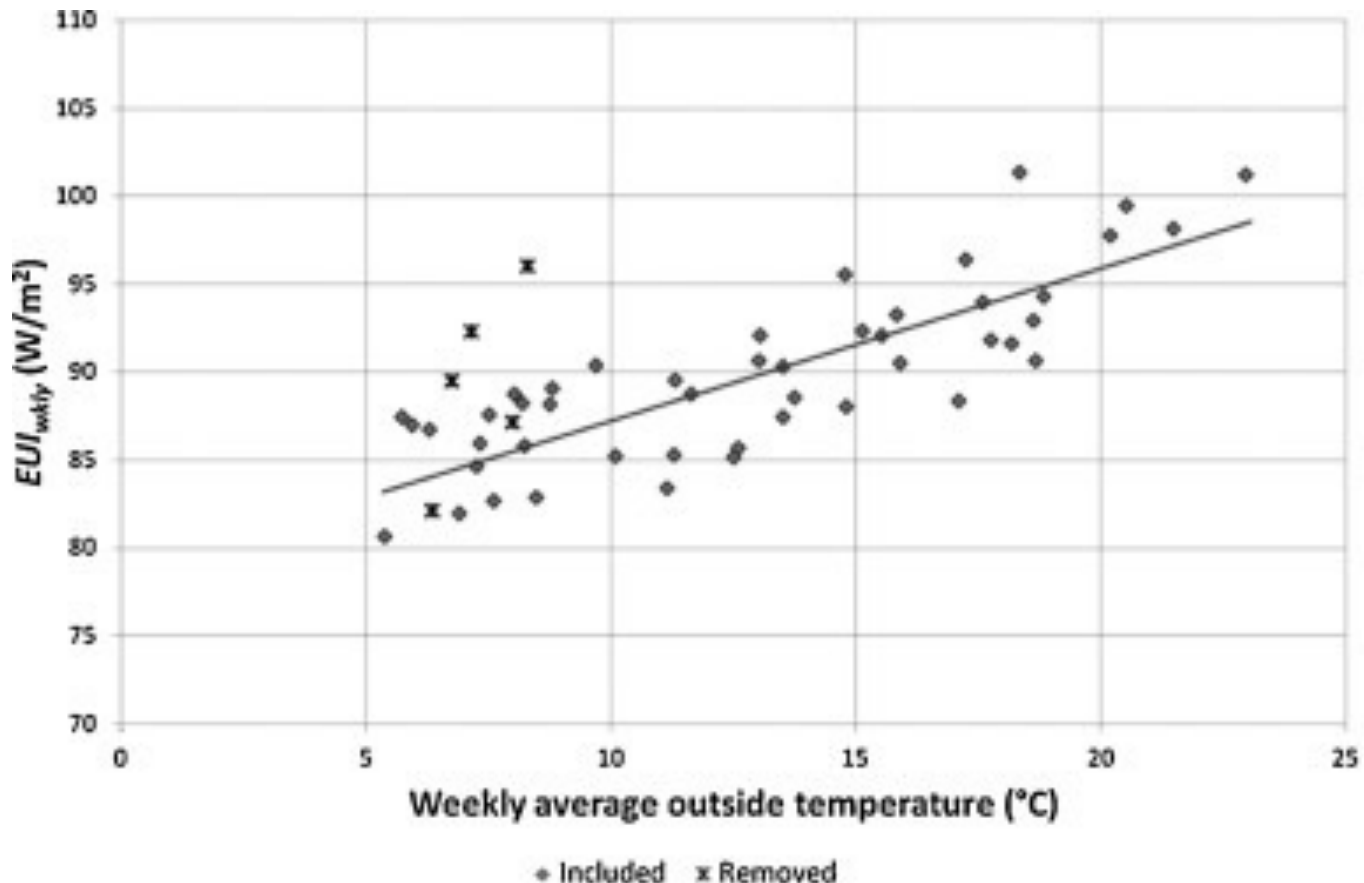
https://en.wikipedia.org/wiki/Energy_in_Switzerland

Mtoe = 11.63 TWh.

What is the Correlation ?



What is the Correlation ?



Electricity

Microsoft EXCEL

Minitab [**Minitab.com**](https://www.minitab.com)

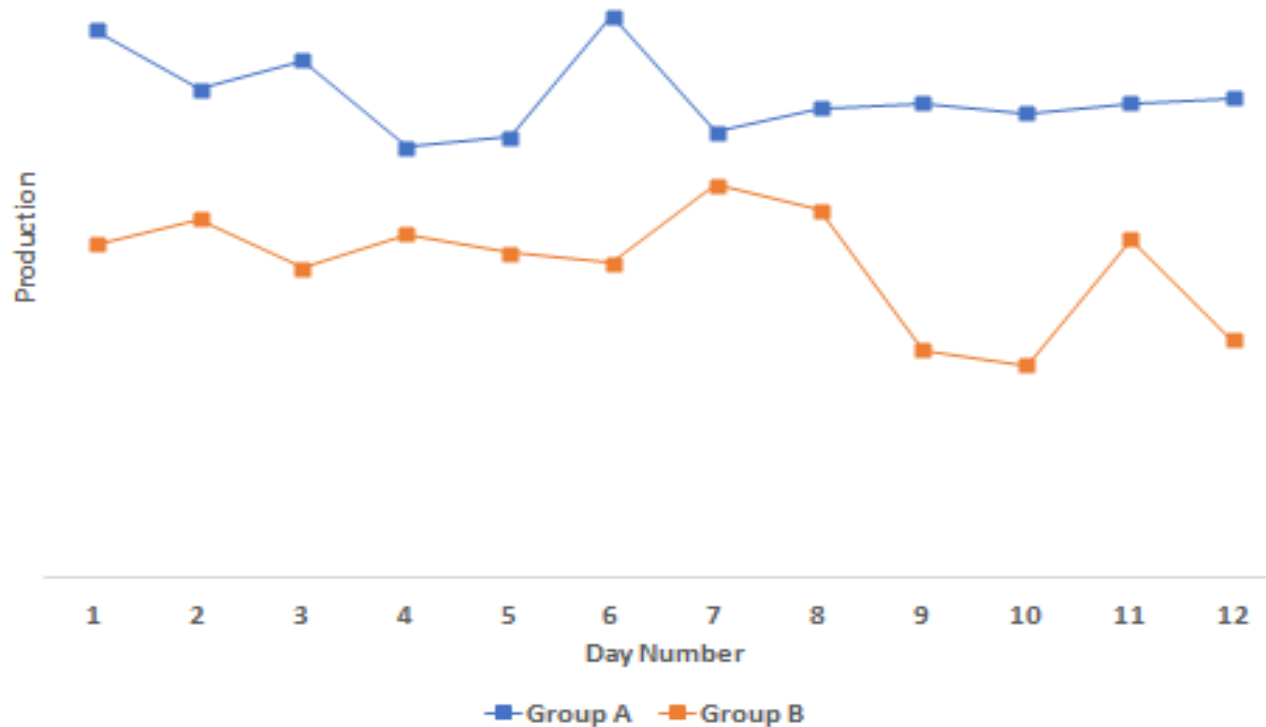
Sigma XL [**SigmaXL.com**](https://www.sigmaxl.com)

QI Macros [**QIMacros.com**](https://www.qimacros.com)

SPC XL [**SigmaZone.com**](https://www.sigmazone.com)

Stratification

Stratification is a method used to look for different relationships across different attributes (categories) of data.



Video : Correlation, Regression and R-Squared Summary



Precision Delivery Inc. Case Study

Precision Delivery Inc. (PDI) is a package pickup and delivery service for homes and small businesses. PDI specializes in packages 50 pounds or less and has a full-price rebate policy for any pickup or delivery made outside the customer-designated 15-minute window.

PDI advertising proudly states, “Pick up and delivery at your convenience, not ours.”

PDI has facilities at 2 locations, Downtown (D) and Suburbia (S) , each servicing customers within a 15-mile radius with pickup and deliveries made by truck or bicycle. Customers designate a 15-minute window for their packages to be picked up or delivered. PDI charges customers \$5 per package plus \$1 per pound (50-pound maximum). PDI Sales Operators receive pickup and delivery requests by phone from customers. Dispatchers issue instructions to Field Operators for pickups and deliveries. To ensure prompt service, Dispatchers plan for 30 minutes of travel time, one way, for each pickup or delivery and target arrival time for the start of the 15-minute window. Under current procedures, travel time in excess of 45 minutes will result in a rebate and travel time less than 30 minutes results in idle time for the Field Operator. Field Operators return directly to the dispatch facility after each pickup or delivery.

Class Exercise

Review the Case Study : **PRECISION DELIVERY INC**

Using the Data in the Minitab file : **Precision-Delivery-Data.mtw**

Look for Correlations using Scatter Graphs and Time Series Plots

Y Variable	X Variable	Stratification
Minutes per trip	Package Weight	none
Minutes per trip	Package Weight	By Vehicle type
Profit Dollars \$\$	Time	none
Profit Dollars \$\$	Time	By Vehicle type
Revenue \$\$	Package Weight	none
Revenue \$\$	Time	By Slot (AM or PM)
Profit Dollars \$\$	Time	By Facility (D or S)

Present your findings in a 5 minute presentation to the class.

What conclusions can you make from your analysis ?



Analyze Phase

Process Modeling

Correlation and Regression

Part 2

Types of Regression Analysis

Simple Linear Regression

- One X, One Y

Multiple Linear Regression

- Two or More X's, One Y with Powers of X = 1

Simple Non-Linear Regression

- One X, One Y Powers of X >1 or < 1

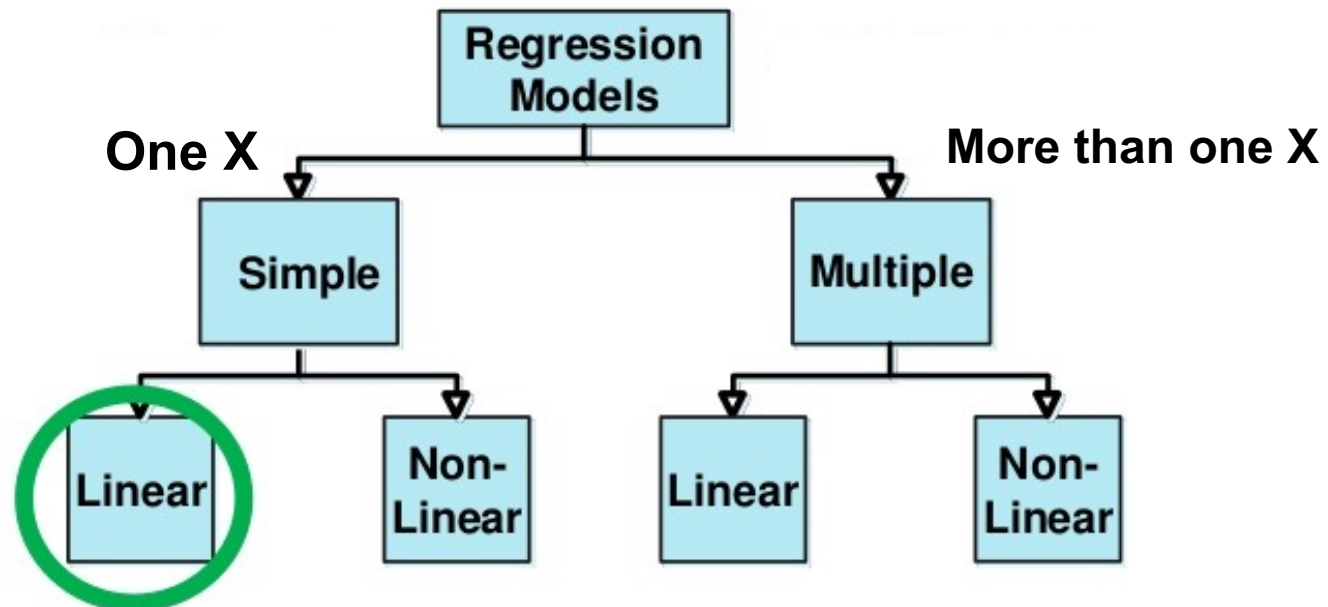
Multiple Non-Linear Regression

- Two or More X's, One Y Powers of X >1 or < 1

Types of Regression

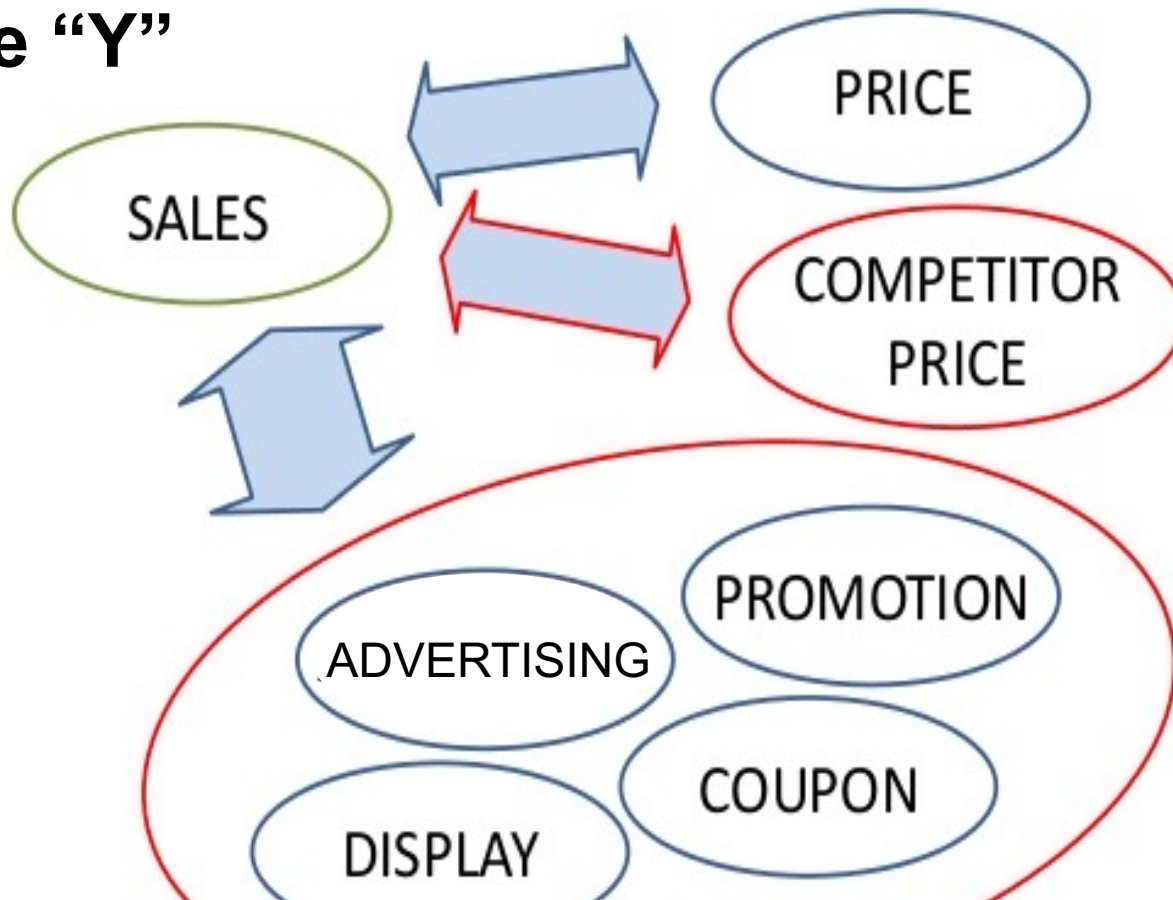
Types of Regression Analysis

3



Multiple Regression

The “Y”

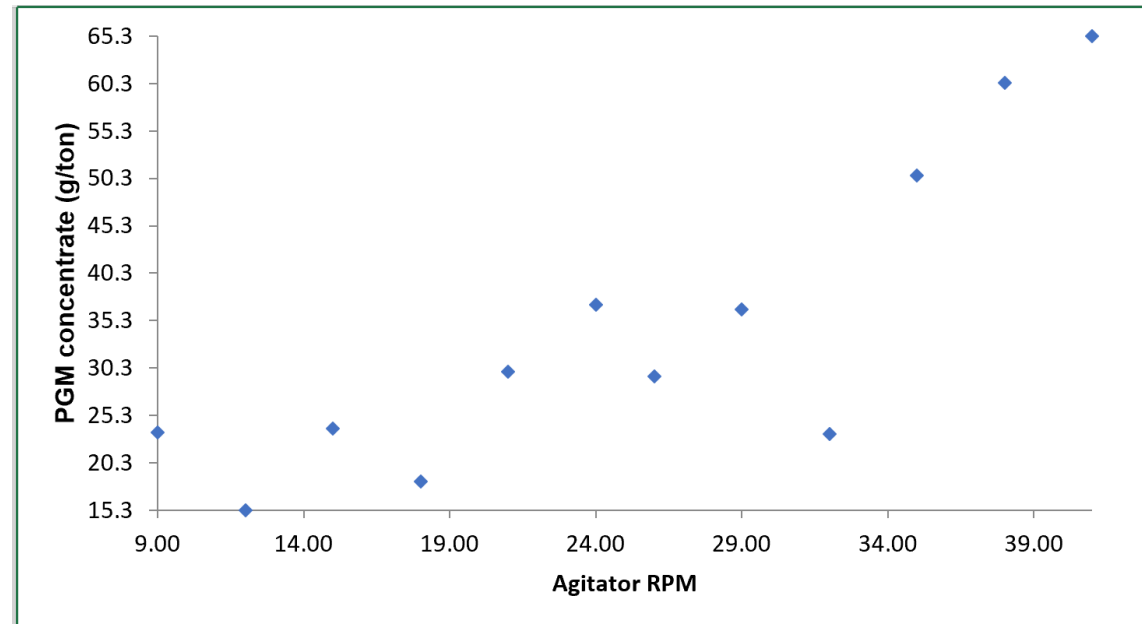
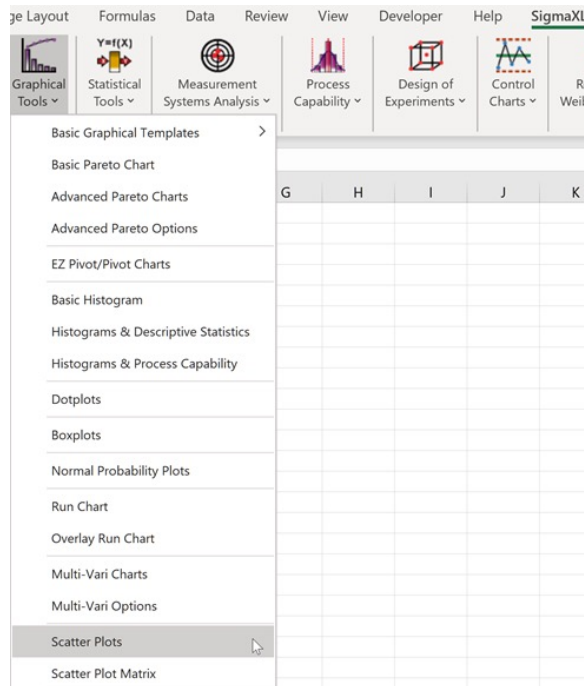


**“The
“X’s”**

Regression Example

This data set is from the mining industry. It is an evaluation of ore concentrators.

Graphical Tools > Scatter Plots



Use worksheet:
Concentrator.xlsx

Linear Regression Example

Multiple Regression Model: PGM concentrate (g/ton) = (1.119) + (1.333) * Agitator RPM

Model Summary:

R-Square	71.81%
R-Square Adjusted	68.99%
S (Root Mean Square Error)	9.082

Parameter Estimates:

Predictor Term	Coefficient	SE Coefficient	T	P	VIF	Tolerance
Constant	1.119	7.106	0.157450	0.8780		
Agitator RPM	1.333	0.264169	5.047	0.0005	1.00000	1

Analysis of Variance for Model:

Source	DF	SS	MS	F	P
Model	1	2101.1	2101.1	25.472	0.0005
Error	10	824.86	82.486		
Total (Model + Error)	11	2925.9	265.99		

Durbin-Watson Test for Autocorrelation in Residuals:

DW Statistic	2.205
P-Value Positive Autocorrelation	0.6192
P-Value Negative Autocorrelation	0.3214

The P-value < 0.05 therefore the Regression is significant.

Notice the unusual observation may indicate a Non-linear analysis may explain more of the variation in the data.

Linear and Non-Linear Regression Example

Linear Model

Regression Analysis: PGM concentrate (g/ton) versus Agitator RPM

The regression equation is

PGM concentrate (g/ton) = 1.119 + 1.333 Agitator RPM

S = 9.08220 R-Sq = 71.8% R-Sq(adj) = 69.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2101.07	2101.07	25.47	0.001
Error	10	824.86	82.49		
Total	11	2925.93			

More variation is explained using the Non-linear model since the R-Squared is higher and the S statistic is lower which is the estimated Standard Deviation of the error in the model.

Non- Linear Model

Polynomial Regression Analysis: PGM concentrate (g/ton) versus Agitator RPM

The regression equation is

PGM concentrate (g/ton) = 30.53 - 1.460 Agitator RPM + 0.05586 Agitator RPM**2

S = 7.61499 R-Sq = 82.2% R-Sq(adj) = 78.2%

Analysis of Variance

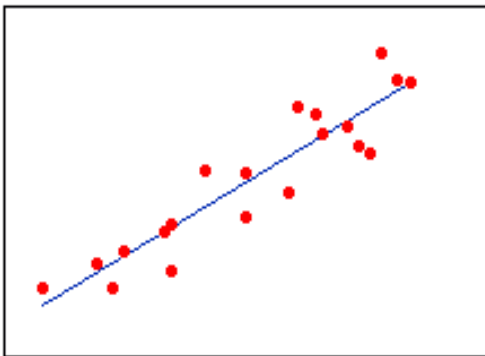
Source	DF	SS	MS	F	P
Regression	2	2404.04	1202.02	20.73	0.000
Error	9	521.89	57.99		
Total	11	2925.93			

Sequential Analysis of Variance

Source	DF	SS	F	P
Linear	1	2101.07	25.47	0.001
Quadratic	1	302.97	5.22	0.048

Nonlinear regression

Other ways to express the general forms of the regression equations

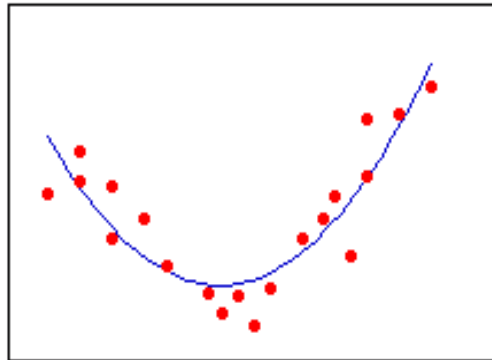


Linear

$$Y' = a + b_1X_1$$

$$Y' = a + b_1X_1 + b_2X_1^2$$

$$Y' = a + b_1X_1 + b_2X_1^2 + b_3X_1^3$$

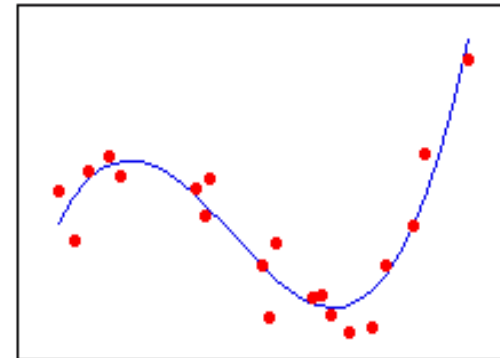


Nonlinear (Quadratic)

Linear

Quadratic

Cubic



Nonlinear (Cubic)

Flight Regression Example

An airplane manufacturer wanted to see what variables affect flight speed. The historical data available covered a period of 10 months.

AutoSave ☐ Off Flight Regression MLR.xlsx - Saved

File Home Insert Page Layout Formulas Data Review View Developer Help **SigmaXL** Sigma

Data Manipulation Templates and Calculators Graphical Tools Statistical Tools Measurement Systems Analysis Process Capability Design of Experiments Control Charts Reliability/Weibull Analysis

V21

	A	B	C	D	E	F	G	H	I	J	K	L
1	Flight Spee	Altitude	Turbine An	Fuel/Air ra	ICR	Temp						
2	618	783.35	33.53	40.55	16.66	10.06						
3	600	748.45	36.5	36.19	16.46	11.26						
4	542	684.45	34.66	37.31	17.66	13.34						
5	524	827.8	33.13	32.52	17.5	6.53						
6	572	860.45	35.75	33.71	16.4	7.15						
7	586	875.15	34.46	34.14	16.28	7.56						
8	600	909.45	34.6	34.85	16.06	8.42						
9	605	905.55	35.38	35.89	15.93	9.24						
10	520	756	35.85	33.53	16.6	6.7						
11	544	769.35	35.68	33.79	16.41	6.95						
12	586	793.5	35.35	34.72	16.17	7.69						
13	585	801.65	35.04	35.22	15.92	8.36						
14	607	819.65	34.07	36.5	16.04	9.6						
15	607	808.55	32.2	37.6	16.19	10.56						
16	590	774.95	34.32	37.89	16.62	11.4						
17	546	711.85	31.08	37.71	17.37	13.18						

ICR (Intermediary Compression Ratio)
Temp (deg F)
Altitude (10s of feet)
Flight speed (mph)

Flight Regression Example Model Selection

Open Worksheet : Flight Regression MLR.xlsx

Statistical Tools > Regression > Multiple Regression...

Multiple Regression

×

Flight Speed

Altitude

Turbine Angle

Fuel/Air ratio

ICR

Temp

Numeric Response (Y) >>

Continuous Predictors (X) >>

(Numeric Data)

Categorical Predictors (X) >>

(Text or Numeric Discrete Data)

<< Remove

OK >>

Cancel

Help

☒ Fit Intercept

☒ Display Residual Plots

Regular

Multiple Regression Model: Flight Speed = (770.39) + (0.153178) * Altitude + (5.806) * Turbine Angle + (8.696) * Fuel/Air ratio + (

Model Summary:

R-Square	89.89%
R-Square Adjusted	87.69%
S (Root Mean Square Error)	18.309

Parameter Estimates:

Predictor Term	Coefficient	SE Coefficient	T	P	VIF	Tolerance
Constant	770.39	229.71	3.354	0.0028		
Altitude	0.153178	0.066046953	2.319	0.0296	2.320	0.431008
Turbine Angle	5.806	2.843	2.042	0.0527	1.355	0.737912
Fuel/Air ratio	8.696	3.327	2.614	0.0155	3.175	0.314983
ICR	-52.269	6.157	-8.489	0.0000	2.612	0.382887
Temp	4.107	3.114	1.319	0.2001	5.371	0.186200

Analysis of Variance for Model:

Source	DF	SS	MS	F	P
Model	5	68522	13704	40.882	0.0000
Error	23	7709.9	335.21		
Total (Model + Error)	28	76231	2722.6		

Durbin-Watson Test for Autocorrelation in Residuals:

DW Statistic	1.580
P-Value Positive Autocorrelation	0.0390
P-Value Negative Autocorrelation	0.7218

Flight Regression Example Model Selection

Regression Analysis: Flight Speed versus Altitude, Turbine Angle, Fuel/Air ratio, Intermediate Compression Ratio, Outside Air Temperature

Multiple Regression Model: Flight Speed = (770.39) + (0.153178) * Altitude + (5.806) * Turbine Angle + (8.696) * Fuel/Air ratio + (-52.269) * ICR + (4.107) * Temp

Model Summary:

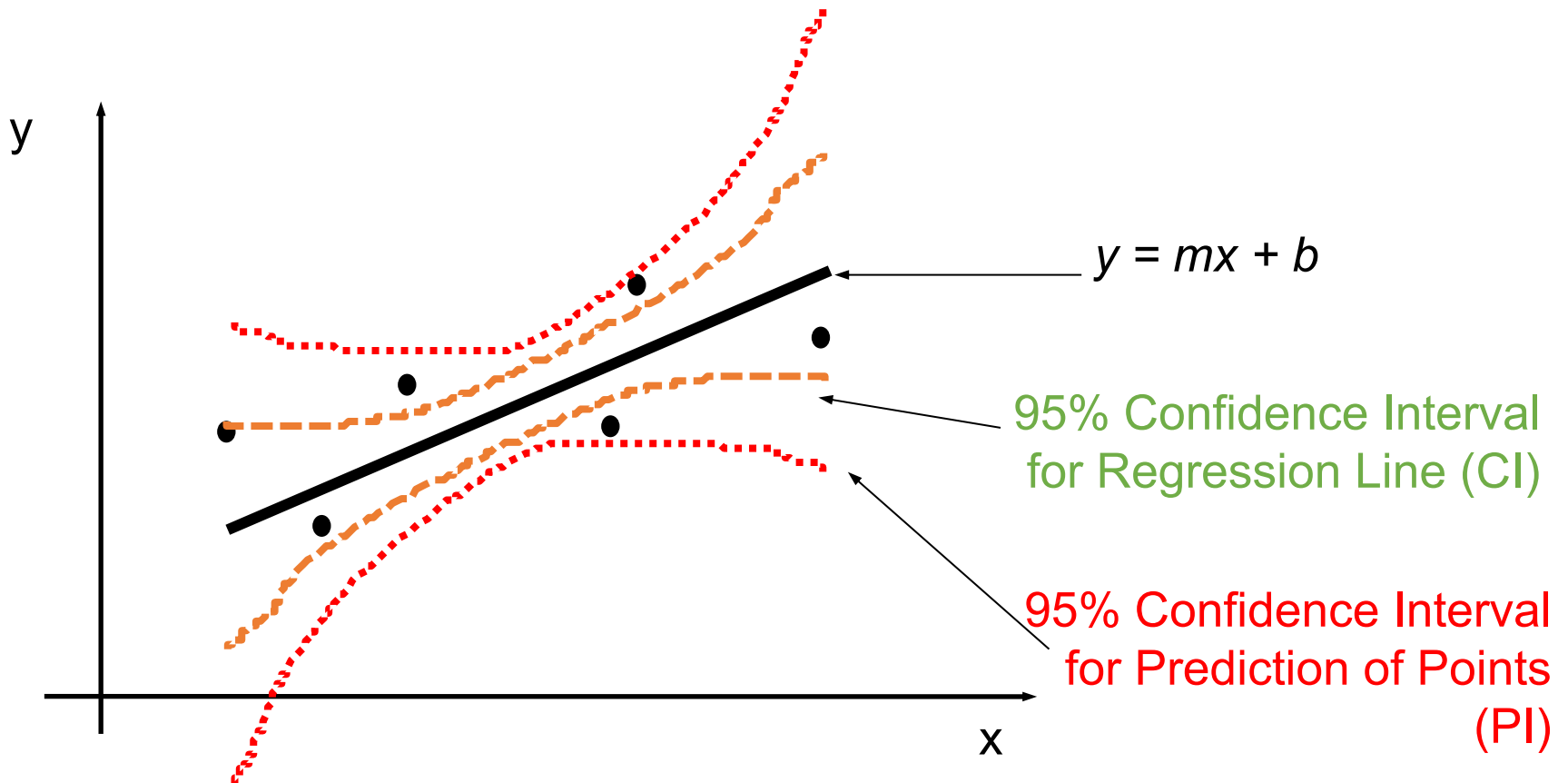
R-Square	89.89%
R-Square Adjusted	87.69%
S (Root Mean Square Error)	18.309

Parameter Estimates:

Predictor Term	Coefficient	SE Coefficient	T	P	VIF	Tolerance
Constant	770.39	229.71	3.354	0.0028		
Altitude	0.153178	0.066046953	2.319	0.0296	2.320	0.431008
Turbine Angle	5.806	2.843	2.042	0.0527	1.355	0.737912
Fuel/Air ratio	8.696	3.327	2.614	0.0155	3.175	0.314983
ICR	-52.269	6.157	-8.489	0.0000	2.612	0.382887
Temp	4.107	3.114	1.319	0.2001	5.371	0.186200

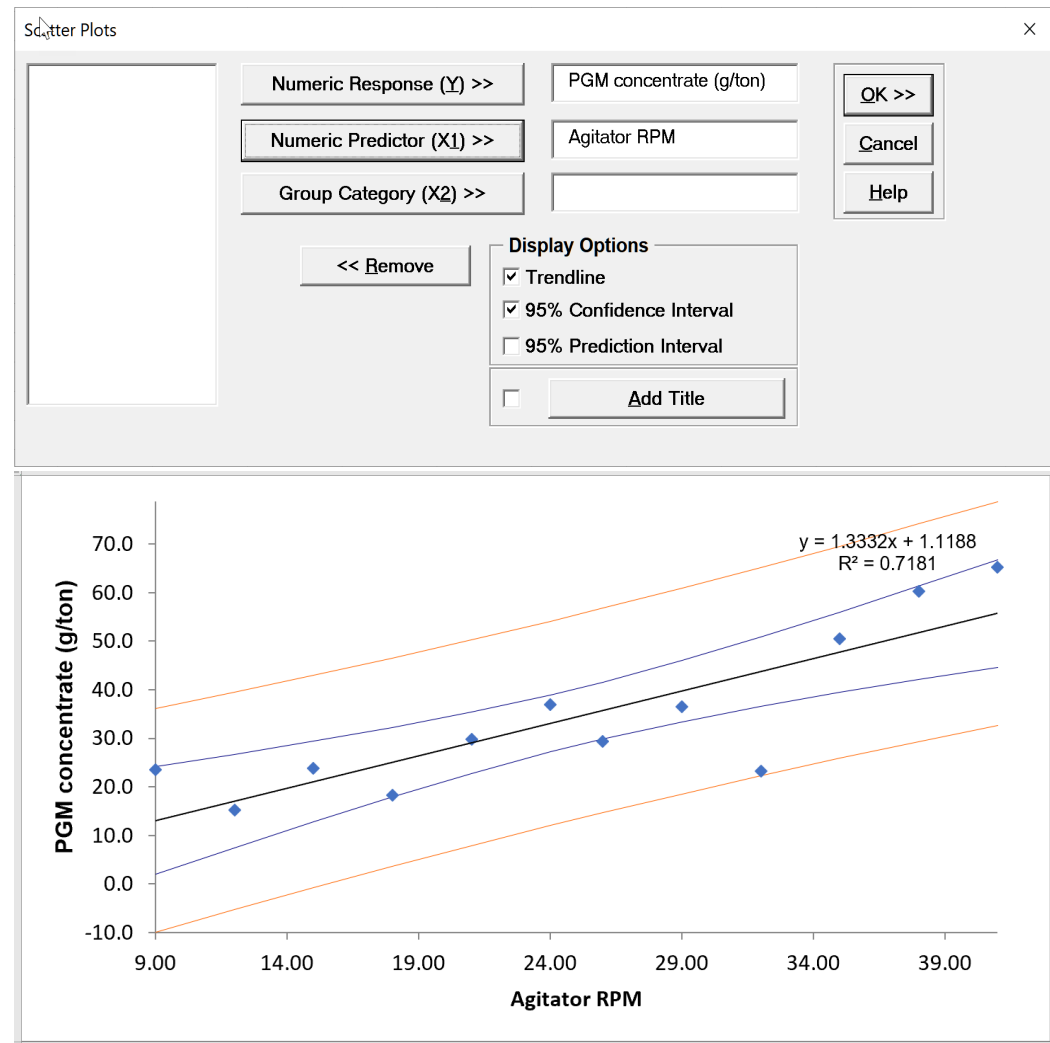
1. What is the dependent variable ?
2. What are the independent variables ?
3. What are the coefficients for the independent variables ?
4. Which independent variable has the largest effect on the dependent variable ?
5. What is the constant β_0 ?
6. Which Predictor coefficient is calculated with the greatest statistical confidence ?
7. What does the R-Sq value tell us about the strength of the relationship ?

Confidence & Prediction Intervals



Scatter Plot with 95% Confidence Intervals

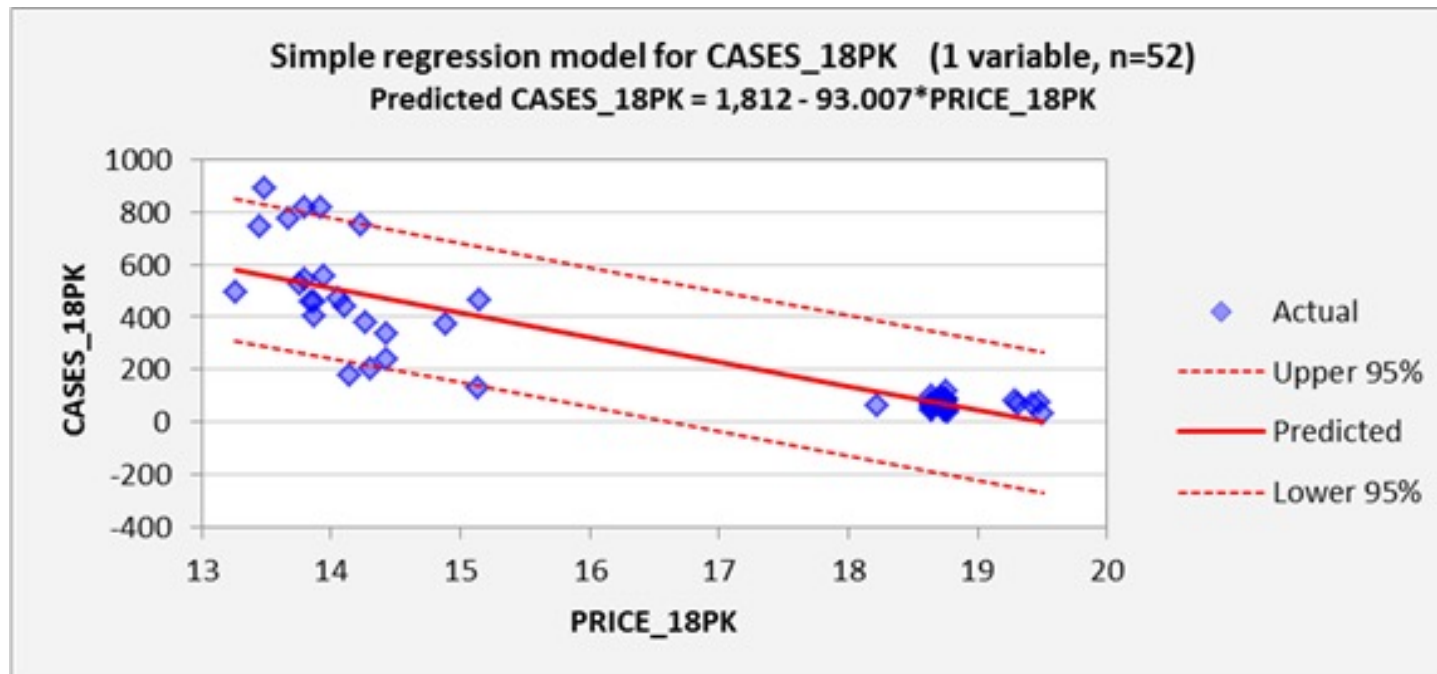
- 95% Confidence Intervals can be displayed for a regression line.
- They represent the “outer envelope” that the regression line could have.
- There will be a 95% confidence level that the true regression line will lie between these confidence interval lines.



Regression Curve Confidence Intervals

Class Discussion :

- Explain what type of Regression is shown.
- What the Regression is telling us about sales of Beer Packs of 18 Bottles and their Price ?
- What are the upper and lower 95% lines (Confidence Intervals) telling us?



Summary

At this point you should:

- Understand concept of Non-linear Regression Analysis
- Understand concept of Multiple Linear Regression Analysis (MLR)
- Understand the P value of a Regression Analysis
- Understand concept of the Confidence Interval of a Regression Curve (CI)

Analyze Phase

Design of Experiments (DOE)

History of Experiments

Thomas Edison / Joseph Swan



Aristotle

320 BC



OFAT

1880



Fisher

1925



Box

1950



Taguchi

1960

Scientific Method

Robust Design

Response Surface Method

Sir Ronald Fisher (Cambridge, UK) in 1920's was the first person to perfect and use the Designed Experiment method for investigating multiple factors that effect Agricultural Crop Growth and Yield. ('F-test' = Fisher Test)

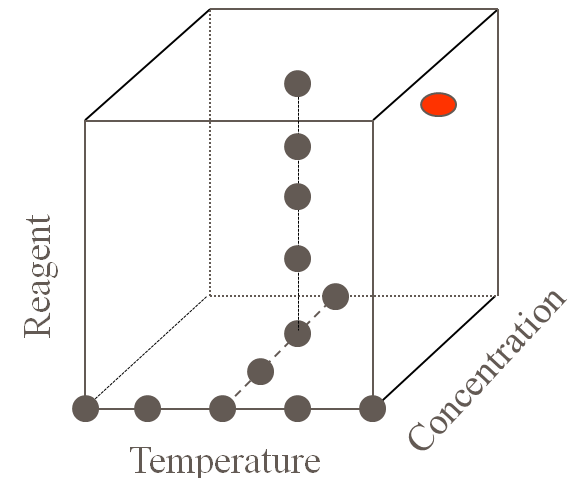
Trail and Error / One Factor at a Time (OFAT) is NOT a DOE

One Factor at a Time (OFAT) is an experimental style but not categorized as a planned or designed experiment (DOE).

OFAT works well only when there is only **ONE INPUT** effecting **ONE OUTPUT** or where **INPUTS ARE ALL INDEPENDENT**.

OFAT CHANGES THE INPUTS ONE AT A TIME

Another name for this is
Trial and Error !



Definition of Design of Experiments

- ***Design of Experiments*** (DOE) is an advanced methodology for determining relationships between Xs and Ys (Inputs and Outputs)
- DOE is a Proactive Method in which we actively change the Xs and then measure how much the Ys change as a result.
- DOE will reveal if a causal relationship exists between the Xs and the Ys (unlike Regression Analysis)
- DOE allows the input variables to be changed simultaneously (unlike OFAT)
- DOE will reveal the interaction effects as well as the main effects
- The type of DOE that is chosen for a particular experiment is based on which will produce the detail of information required at a minimum cost and time.

DOE Terminology

“Factors”

Factors = inputs = X's = Predictors

The input variables that we are interested in to see what effect they have on the output (response). For example these may be the “Store Location”, “Store opening hours”, “Number of Sales people”, “Type of Advertising in window”.

“Response”

Response = Output = Y

The value that we are interested in optimizing. For example this may be a KPI for a process such as ‘Number of Sales per day’.

“Level”

How many “states is a Factor (X) set to during the experiment . The simplest is 2 Level where each factor is used in two (2) states (On/Off, High/Low etc)



DOE Terminology (2)

“Replication”

To repeat (replicate) the same experiment more than once. The results of each replication are used in the analysis and this increases the statistical confidence in the results.

“Full Factorial” (Full Factorial Design)

Also called ‘Optimization Designs’ or ‘Modelling Designs’

Full Factorial means that every possible combination of X at every level is tested. This is often costly and time consuming.

“Fractional Factorial” (Fractional Factorial Design)

Also called ‘Screening Experiments’ or ‘Screening Designs’

Fractional Factorial means testing a subset of the combinations of Xs. This is less cost and time than Full Factorial however the downside is some loss of resolution and detail in determining which are the critical interactions.

Types of DOE Experimental methods

Fractional Factorial

- 5-30 input variables

Full Factorial

- 2-4 input variables

Response Surface Method (RSM) / Method of Steepest Ascent

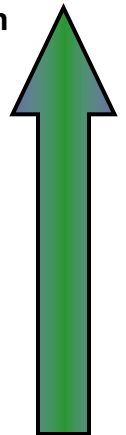
- 2-3 input variables

RESPONSE SURFACE METHODs
Steepest Ascent / Decent Design

Full Factorial

Fractional Factorials

KNOWLEDGE



Full Factorial Experiments

Nomenclature for Full Factorial Designed Experiments

The general notation used to designate a full factorial design is given by:

$$N^k$$

- Where k is the number of input variables or factors (X s).
- N is the number of “levels” that will be used for each factor in the experiment.
- Since most DOE experiments are run with only 2 levels, this becomes

$$2^k$$

Design of experiments

- *Roger Federer*



Maximize speed serve

FACTORS:

Racket Weight
Head Surface Size
Racket length

LEVELS :

Racket Weight (grams)
Head Surface Size (cm²)
Racket length (cm)

RESPONSE :

Speed of Serve (Km / hr)



Maximize speed serve

FACTORS :

Racket Weight (grams)
Head Surface Size (cm²)
Racket length (cm)



LEVELS

LOW	HIGH
300	350
620	645
70	75

Example

Example

A manufacturer of paint is examining the effects of inputs on the final thickness of the paint in order to control the thickness to best meet customer specification limits. They decide to do a DOE on the following inputs tested at the following levels:

X_1	= Solvent Type	Type 12A2	or	Type 04RR4
X_2	= Mixing Speed	1200 revs / min	or	1800 revs / min
X_3	= Temperature	30 degC	or	45 degC
X_4	= Pressure	1.1 bar	or	1.2 bar



Questions:

1. How many factors are being tested ?
2. How many levels are the factors being tested at ?
3. How many independent experiments are required for a Full Factorial DOE ?

Example 2

Example 2

The SB Glass Co makes float glass sheets. Recently their customers have been complaining that glass delivered from SB is prone to cracking.

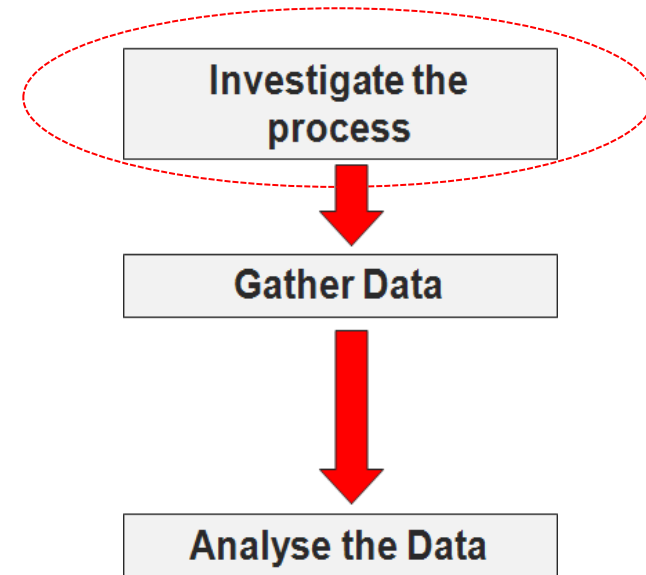
The company suspect that the cracking is influenced by a combination of input factors.

They decide to design an experiment to investigate this problem



Example 3 : Step 1 : Investigate the Process

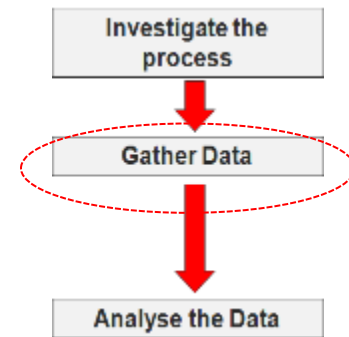
- The company investigates the process and identify what they consider to be critical inputs.
- In this case they identify:
 - Melt Temperature
 - Initial Float Temperature
 - The amount of sodium carbonate in the mix
- The critical Y is the “hardness” of the glass
 - Measured on a scale of 1-100 (100 hardest)



Example 3 : Step 2: Gather Data

- The team decides on a **3 Factor Full Factorial** DOE
 - Three Xs will involve eight combinations
- They define the values for the three Xs as follows:

Input	Low value	High value
Melt Temp (Deg C)	1500	1600
Float Temp (Deg C)	1100	1200
Sodium Carb	2%	5%



- The team build an initial SigmaXL DOE worksheet (next slide)
- The team then vary the Xs in the order selected in the worksheet – they record the corresponding hardness of the glass for each combination

Example 3 : Recording the data: Building the worksheet

- In SigmaXL you cannot use a standard worksheet to record the experiment data
- Design of Experiments → 2-Level Factorial/Screening → 2-Level Factorial/Screening Designs

1: Set up type of design and number of factors

2: Select the design

3: Configure the factors

How many times to repeat experiments

Level Factorial/Screening Design of Experiments

Number of Factors: 3

Select Design: 8-Run, 2**3, Full-Factorial

Number of Replicates: 2

Power Information (based on # of runs and replicates):

Very Low Power to detect Effect = 1*StDev (1-Beta < 0.5);
Medium Power to detect Effect = 2*StDev (0.8 <= 1-Beta < 0.95);
Very High Power to detect Effect = 3*StDev (1-Beta >= 0.99).

Number of Blocks: 1

Number of Center Points per Block: 0

☒ Randomize Runs

Factor Names and Level Settings:

	Name	Low	High
A:	Melt temp	1500	1600
B:	Float Temp	1100	1200
C:	Sod Carb	2	3

Number of Responses: 1

Response Name

Y1: Hardness

OK >> Cancel Help Reset

Example 3 : SigmaXL creates an initial DOE worksheet

Design of Experiments Worksheet

	Title:	
	Date:	
	Name of Experimenter:	
	Notes:	

Design Type:	3 Factor, 8-Run, 2**3, Full-Factorial
Number of Replicates:	2
Number of Blocks:	1
Number of Center Points per Block:	0
Number of Responses:	1

All 8 combinations
repeated twice
(replicates = 2)

All performed by the
same people on a
single shift hence
only one block

Run Order	Std. Order	Center Points	Blocks	A: Melt temp	B: Float Temp	C: Sod Carb	Hardness
1	4	1	1	1600	1200	2	
2	10	1	1	1600	1100	2	
3	13	1	1	1500	1100	3	
4	11	1	1	1500	1200	2	
5	8	1	1	1600	1200	3	
6	1	1	1	1500	1100	2	
7	5	1	1	1500	1100	3	
8	9	1	1	1500	1100	2	
9	12	1	1	1600	1200	2	
10	15	1	1	1500	1200	3	
11	3	1	1	1500	1200	2	
12	16	1	1	1600	1200	3	
13	2	1	1	1600	1100	2	
14	14	1	1	1600	1100	3	
15	7	1	1	1500	1200	3	
16	6	1	1	1600	1100	3	

Example 3 : Team completes the worksheet

Design of Experiments Worksheet

Title:	
Date:	
Name of Experimenter:	
Notes:	

Design Type:	3 Factor, 8-Run, 2**3, Full-Factorial
Number of Replicates:	2
Number of Blocks:	1
Number of Center Points per Block:	0
Number of Responses:	1

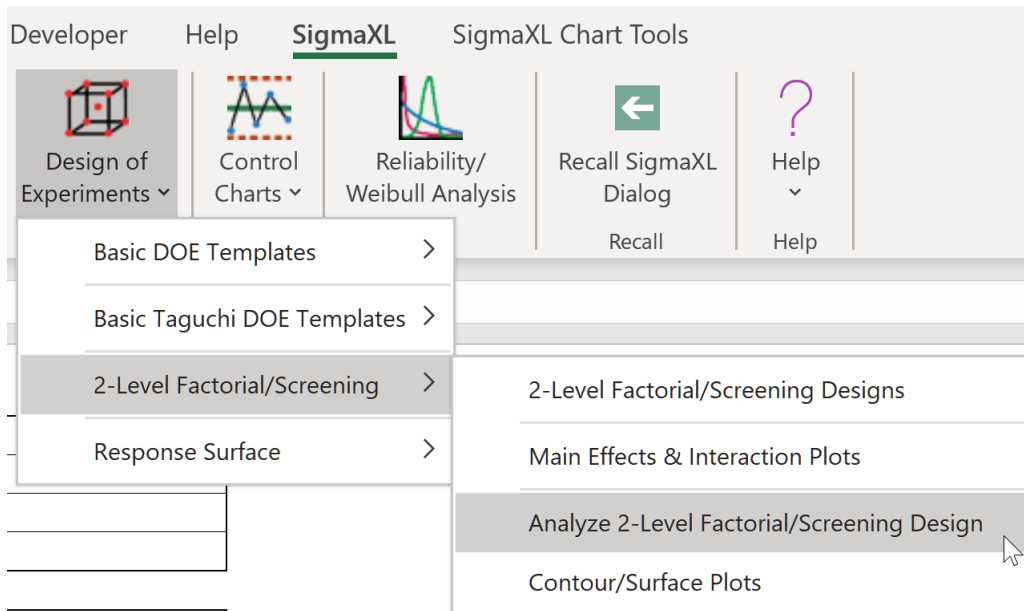
Run Order	Std. Order	Center Points	Blocks	A: Melt temp	B: Float Temp	C: Sod Carb	Hardness
1	13	1	1	1500	1100	5	65
2	2	1	1	1600	1100	2	60
3	7	1	1	1500	1200	5	60
4	16	1	1	1600	1200	5	70
5	4	1	1	1600	1200	2	55
6	5	1	1	1500	1100	5	65
7	15	1	1	1500	1200	5	60
8	8	1	1	1600	1200	5	70
9	3	1	1	1500	1200	2	45
10	6	1	1	1600	1100	5	75
11	1	1	1	1500	1100	2	50
12	11	1	1	1500	1200	2	45
13	14	1	1	1600	1100	5	75
14	12	1	1	1600	1200	2	60
15	10	1	1	1600	1100	2	60
16	9	1	1	1500	1100	2	50

The team have added an extra column (C8) in which they have listed the hardness of the glass for that particular combination

Completed DOE Worksheet – Note that when Randomize Runs is checked, Std. Order will be different for each sheet.

Example 3: Step 3: Analyse Data

Now the team have the necessary data captured in a SigmaXL worksheet they use DOE tool included in the software to investigate the process



Design of Experiments →

2-Level

Factorial/Screening →

Analyze 2-Level Factorial/Screening Design

Example 3 : SigmaXL Results



Design of Experiments Analysis

DOE Multiple Regression Model: Hardness = (55.52083333) + (5.52083333) * A: Melt temp + (-1.979166667) * B: Float Temp + (2.395833333) * C: Sod Carb + (0.520833333) * AB + (-0.104166667) * AC + (-0.104166667) * BC + (-0.104166667) * ABC

Title:	
Date:	
Name of Experimenter:	
Notes:	

Design Type:	3 Factor, 8-Run, 2**3, Full-Factorial
Number of Replicates:	2
Number of Blocks:	1
Number of Center Points per Block:	0
Response:	Hardness

Model Summary:

R-Square	99.09%
R-Square Adjusted	98.29%
S (Root Mean Square Error)	1.250

Parameter Estimates (Coded Units):

Term	Coefficient	SE Coefficient	T	P	VIF	Tolerance
Constant	55.52083333	0.375578258	147.83	0.0000		
A: Melt temp	5.520833333	0.375578258	14.700	0.0000	1.444444444	0.692307692
B: Float Temp	-1.979166667	0.375578258	-5.270	0.0000	1.444444444	0.692307692
C: Sod Carb	2.395833333	0.104166667	23	0.0000	1	1
AB	0.520833333	0.375578258	1.387	0.2029	1.444444444	0.692307692
AC	-0.104166667	0.104166667	-1	0.3466	1.444444444	0.692307692
BC	-0.104166667	0.104166667	-1	0.3466	1.444444444	0.692307692
ABC	-0.104166667	0.104166667	-1	0.3466	1.444444444	0.692307692

Analysis of Variance for Model:

Source	DF	SS	MS	F	P
Model	7	1360.9	194.42	124.43	0.0000
Error	8	12.500	1.563		
Pure Error	8	12.500	1.563		
Total (Model + Error)	15	1373.4	91.563		

Durbin-Watson Test for Autocorrelation in Residuals:

DW Statistic	2.000
P-Value Positive Autocorrelation	0.6060
P-Value Negative Autocorrelation	0.6060

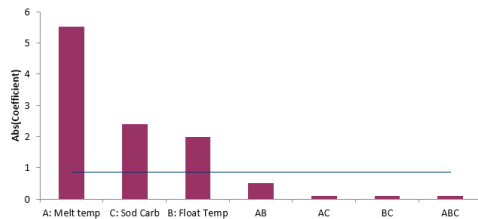
Predicted Response Calculator:

Predictors	Enter Actual Settings:	Coded Settings	Predicted Response	Lower 95% CI	Upper 95% CI	Lower 95% PI	Upper 95% PI
A:							
B:							
C:							

Note:

1. Enter settings for predictors. If the predictors are block terms, specify a 0 or 1 for each level.
2. Do not insert or delete rows or columns in this worksheet.

Pareto Chart of Coefficients for Hardness



Example 3 : The Catapult

Consider a 2^3 DOE to optimize a catapult for distance.

-Level Factorial/Screening Design of Experiments

Number of Factors: 3

Select Design: 8-Run, 2^{3-1} , Full-Factorial

Number of Replicates: 2

Power Information (based on # of runs and replicates):

Very Low Power to detect Effect = $1 \times \text{StDev}$ ($1 - \text{Beta} < 0.5$);
 Medium Power to detect Effect = $2 \times \text{StDev}$ ($0.8 \leq 1 - \text{Beta} < 0.95$);
 Very High Power to detect Effect = $3 \times \text{StDev}$ ($1 - \text{Beta} \geq 0.99$).

Number of Blocks: 1

Number of Center Points per Block: 0

☒ Randomize Runs

Factor Names and Level Settings:

	Name	Low	High
A:	Pull Back Angle	160	180
B:	Stop Pin	2	3
C:	Pin Height	2	2

Number of Responses: 1

Response Name

Y1: Distance

OK >>
Cancel
Help
Reset



1. What is the response being investigated ?
2. What are the 3 factors being investigated in this design?
3. What levels are being used in this DOE ?
4. How many runs are needed for Full Factorial DOE ?

Three Factor Full Factorial Design

Class exercise: Open **DOE** tab from worksheet and run an Analysis on it.

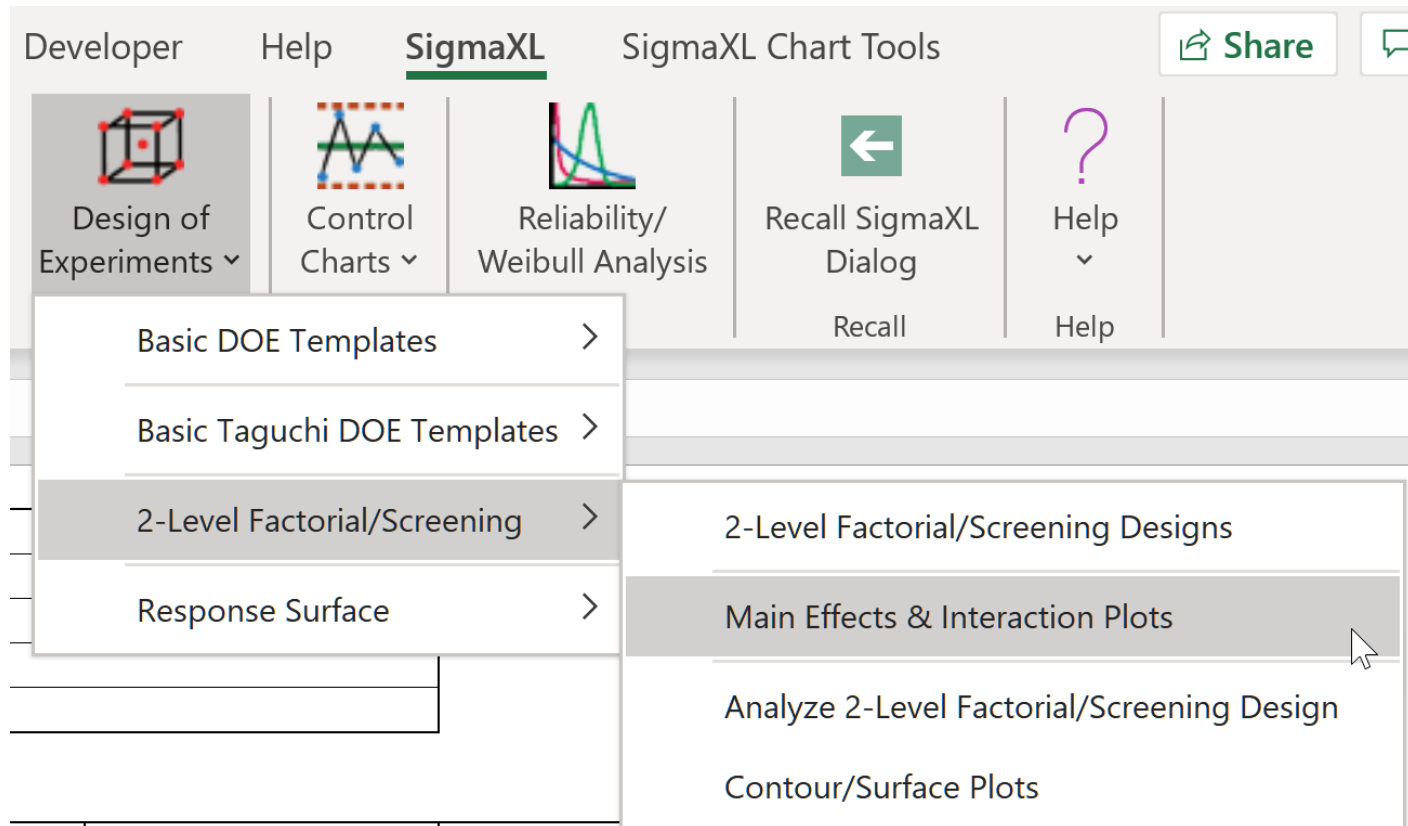
1. What factor has the biggest effect on the distance travelled ?
2. Which effects (main and interaction) are statistically significant ?

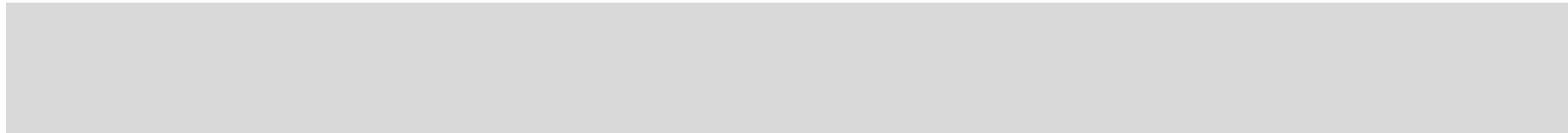
Design Type:	3 Factor, 8-Run, 2**3, Full-Factorial
Number of Replicates:	2
Number of Blocks:	1
Number of Center Points per Block:	0
Number of Responses:	1

Run Order	Std. Order	Center Points	Blocks	A: Pull Back Angle	B: Stop Pin	C: Pin Height	Distance
1	14	1	1	180	2	3	124
2	4	1	1	180	3	2	115
3	12	1	1	180	3	2	116
4	9	1	1	160	2	2	55
5	1	1	1	160	2	2	57
6	5	1	1	160	2	3	70
7	2	1	1	180	2	2	102
8	15	1	1	160	3	3	95
9	10	1	1	180	2	2	101
10	7	1	1	160	3	3	96
11	11	1	1	160	3	2	80
12	13	1	1	160	2	3	72
13	16	1	1	180	3	3	144
14	6	1	1	180	2	3	125
15	8	1	1	180	3	3	142
16	3	1	1	160	3	2	79

Graphical DOE Analysis - The Main Effects Plot

Main Effects Plots



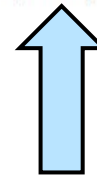


Design of Experiments Analysis

DOE Multiple Regression Model: Distance = (98.3125) + (22.8125) * A: Pull Back Angle + (10.0625) * B: Stop Pin + (10.1875) * C: Pin Height + (-1.9375) * AB + (2.4375) * AC + (0.6875) * BC + (0.4375) * ABC

Title:	Catapult Example (SigmaXL Version 6)
Date:	
Name of Experimenter:	
Notes:	

Design Type:	3 Factor, 8-Run, 2**3, Full-Factorial
Number of Replicates:	2
Number of Blocks:	1
Number of Center Points per Block:	0
Response:	Distance



**Regression Equation
(Model)**

Model Summary:

R-Square	99.93%
R-Square Adjusted	99.86%
S (Root Mean Square Error)	1.030776406



R-Squared Value

Model Summary:

R-Square	99.93%
R-Square Adjusted	99.86%
S (Root Mean Square Error)	1.030776406

Parameter Estimates (Coded Units):

Term	Coefficient	SE Coefficient	T	P	VIF	Tolerance
Constant	98.3125	0.257694102	381.51	0.0000		
A: Pull Back Angle	22.8125	0.257694102	88.526	0.0000	1	1
B: Stop Pin	10.0625	0.257694102	39.048	0.0000	1	1
C: Pin Height	10.1875	0.257694102	39.533	0.0000	1	1
AB	-1.9375	0.257694102	-7.519	0.0001	1	1
AC	2.4375	0.257694102	9.459	0.0000	1	1
BC	0.6875	0.257694102	2.668	0.0285	1	1
ABC	0.4375	0.257694102	1.698	0.1280	1	1

Analysis of Variance for Model:

Source	DF	SS	MS	F	P
Model	7	11773	1681.8	1582.9	0.0000
Error	8	8.500	1.0625		
Pure Error	8	8.500	1.0625		
Total (Model + Error)	15	11781	785.43		

Analyze Phase

Introduction to Hypothesis Testing

What is Hypothesis Testing ?

- ***A Hypothesis is a statement or claim*** about something that is made on the basis of limited evidence.
- ***Hypothesis Testing*** help us to make a data-based decision about whether to take action based on our Hypothesis and the amount of evidence.
- ***Hypothesis Testing*** can never tell that our Hypothesis is 100% true, however it can tell us how much risk we are taking in accepting the Hypothesis and acting upon it.
- ***Hypothesis Testing*** uses evidence to help us to make decisions about whether PERCEIVED differences or changes in data are STATISTICALLY SIGNIFICANT.

Null and Alternative Hypotheses

Every Hypothesis has two parts :

1. The “**null hypothesis**” H_0 assumes there are no differences or relationships and that there probably is no reason to take action.

This is the ‘Status Quo’ condition where nothing is changed

This is the default assumption of all statistical tests.

2. The “**alternative hypothesis**” H_a states there is a difference or relationship or real change and so there is probably a good reason to take action.

H_0 = there is no difference or relationship in the data

H_a = there is a difference or relationship in the data

Statistical Significant Difference

What is the NULL HYPOTHESIS ?

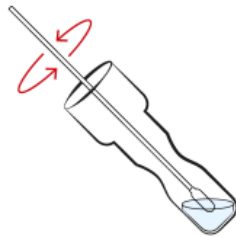
What is the ALTERNATIVE HYPOTHESIS ?

How much risk is there that we are wrong ?

How a rapid lateral flow test works



1
Swab taken from nose and throat



2
Swab sample added to extraction tube



3
Drop of extraction solution
added to test cartridge



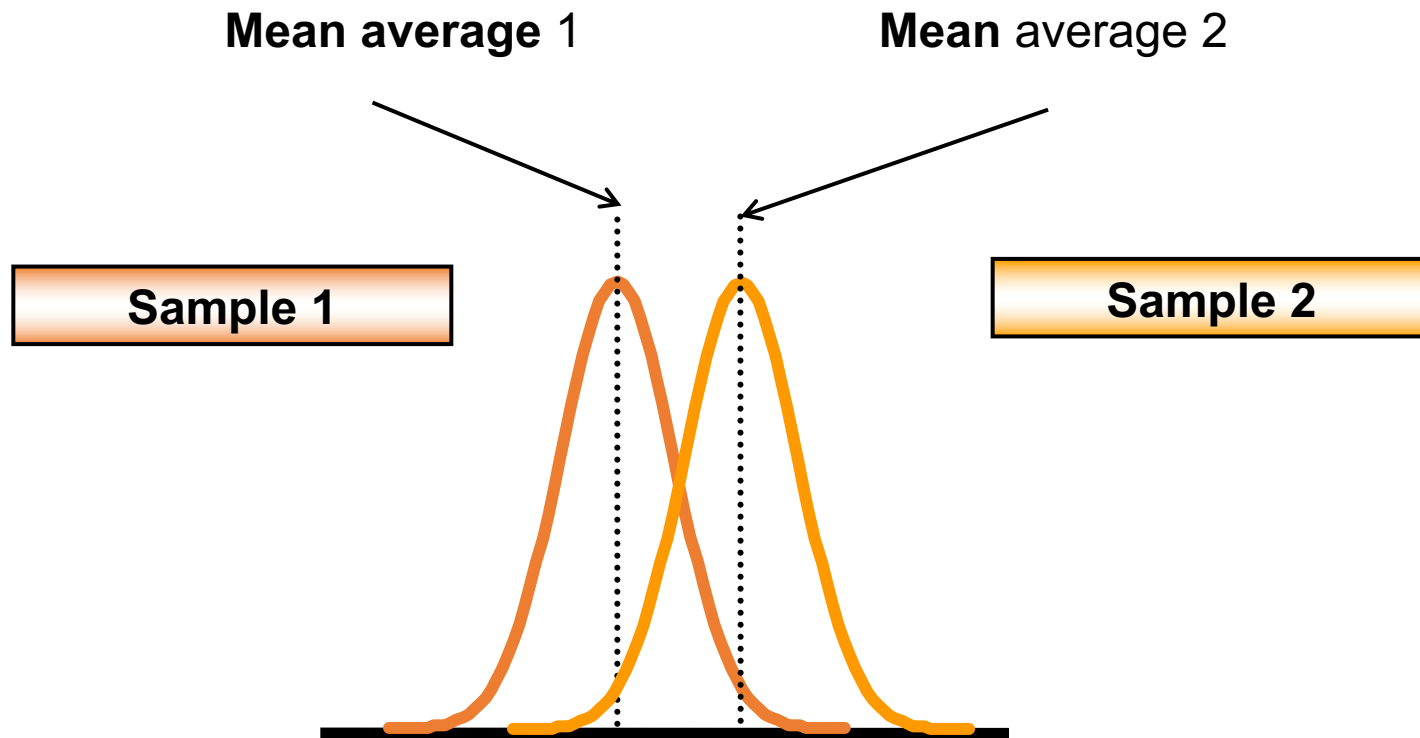
4
The test gives a result in
20-30 minutes



Statistical Significant Difference

Are the two samples “statistically significantly” different from each other?

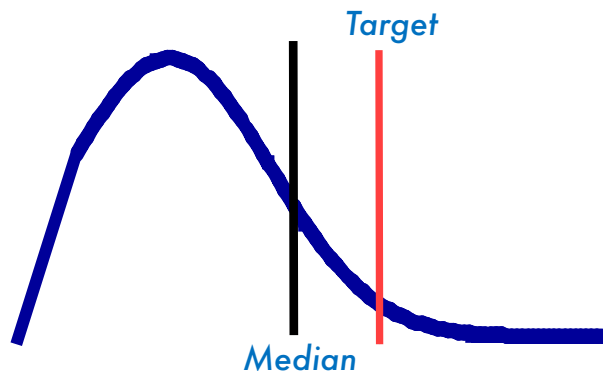
How sure are we of our decision? How much risk is there that we are wrong ?



Types of Hypothesis Test

A Hypothesis Test can be based on either a difference in the Means, the Medians, a Proportion or a variance / standard deviation.

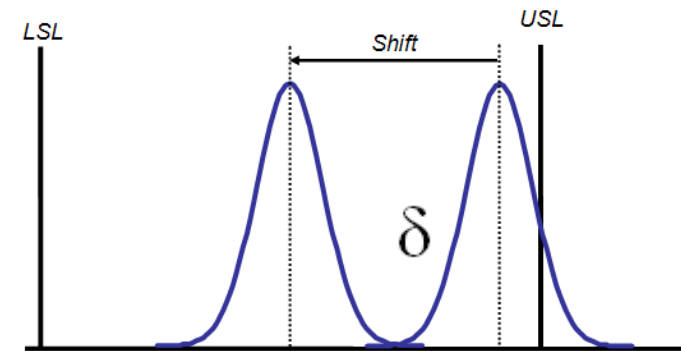
Median Difference ?



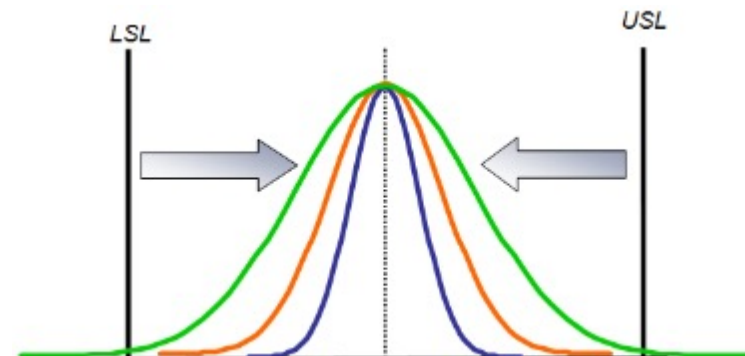
Proportion Difference ?

	Before	After
- Bad Parts	10	12
Sample Size	1000	1500

Mean Difference ?



Variation Difference ?



p-values

Hypothesis Tests uses Probability Values, called '**p-values**' to indicate **the Probability that a Change or Difference is just by chance (due to common cause variation)**.

p-values are expressed either as a percentage (0% to 100%)

Or expressed as a decimal (0.0 to 1.0)

EXAMPLE:

$p = 0.01$ is the same as saying $p = 1\%$

This means there is a 1% chance that the data is different only because of random chance or 'luck'. Therefore a 99% likelihood that there is a real difference and our Alternative Hypothesis should be followed (Null Hypothesis should be rejected)

Alpha Level of risk

The **Alpha Level** of a risk is set by the business (management). It is the amount of risk they are prepared to take in being WRONG about a decision.

Alpha Level represents the benchmark against which *p-values* are measured in a business.

The NULL HYPOTHESIS, H_0 is rejected if the ***p-value*** is less than the alpha level of risk.

$p > \text{alpha level of risk}$ accept H_0

$p < \text{alpha level of risk}$ reject H_0

The most commonly used alpha level in business is 5% (0.05)

$p > 0.05$ accept H_0

$p < 0.05$ reject H_0

Hypothesis Testing : Type 1 Error (False Positive)

Type 1 Error (False Positive) is the probability we could be wrong in saying something is “different from the Status Quo, different from the Null Hypothesis. It is an assessment of the likelihood that the result could have occurred by random chance.

“As of today, someone who gets a positive Lateral Flow Test result in (say) London has at best a 25% chance of it being a true positive, but if it is a self-reported test potentially as low as 10% (on an optimistic assumption about specificity) or as low as 2% (on a more pessimistic assumption).” Ben Dyson May 2021

Ben Dyson is an executive director of strategy at the UK health department and UK Government adviser on Covid19 strategy.

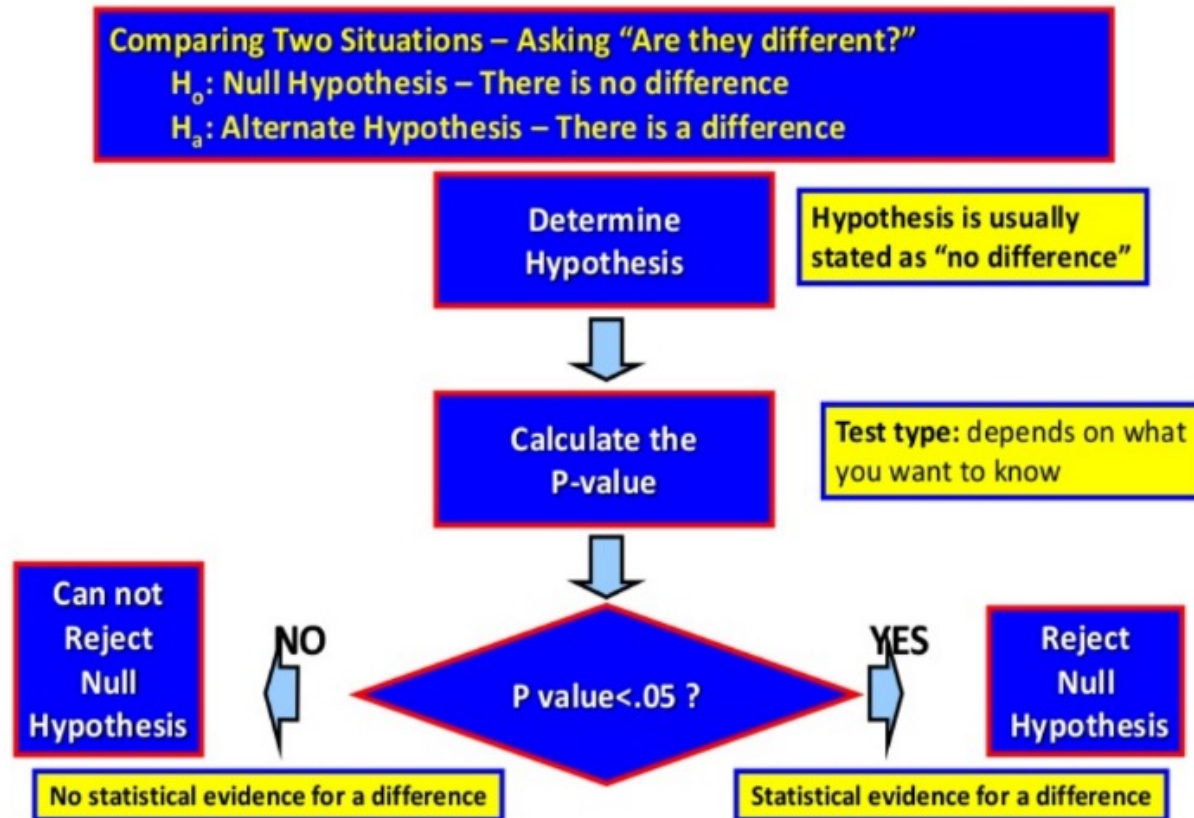
Hypothesis Testing : Type 1 Error (False Positive)

Type 1 Error (False Positive) is the probability we could be wrong in saying something is “different from the Status Quo” It is an assessment of the likelihood that the observed difference could have occurred by random chance.



Method of Hypothesis Testing

Hypothesis Testing



Which Test do I use ?

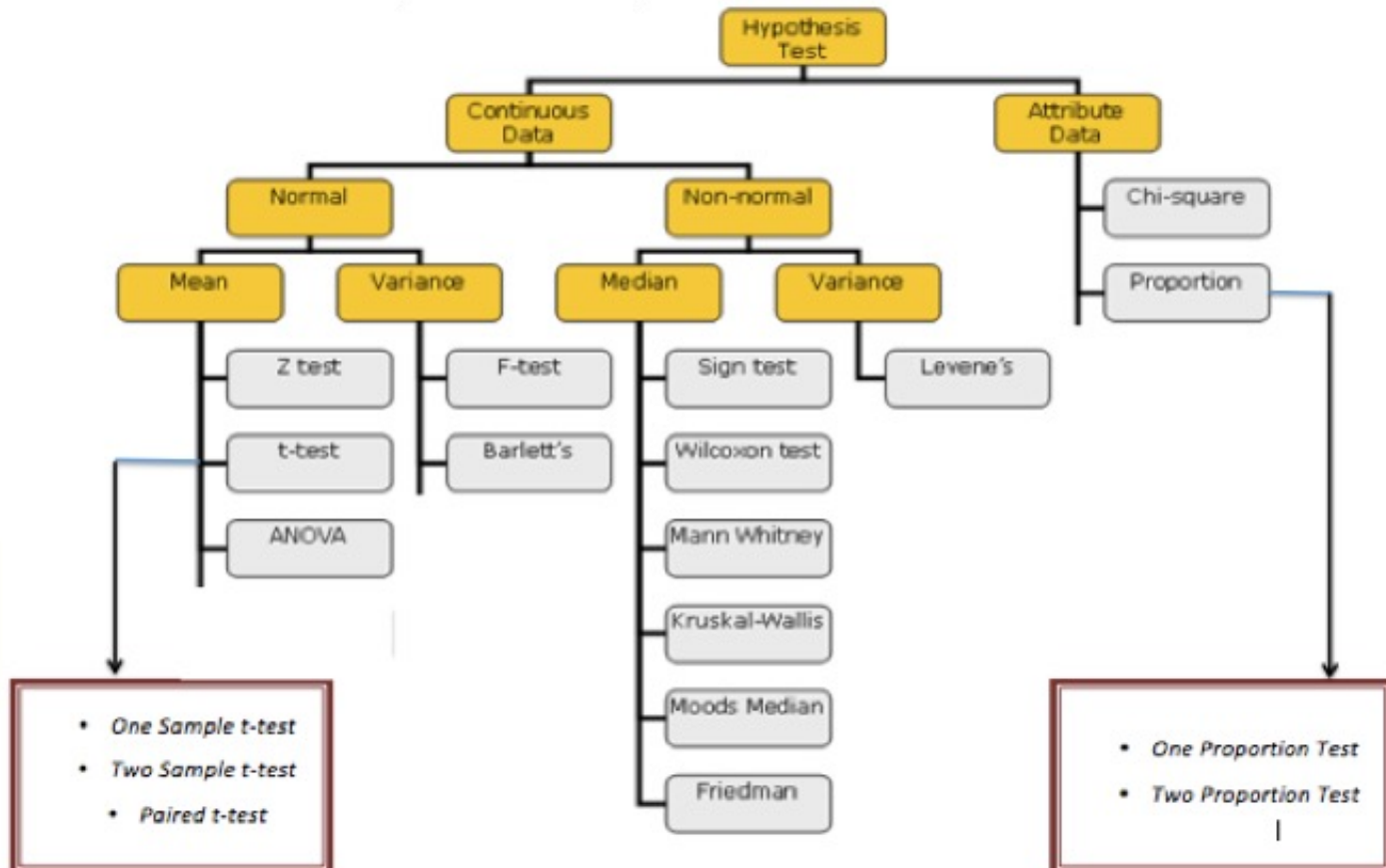
- There are many different statistical methods (Hypothesis tests) that each have different equations in order to calculate the “p-value”

Which test is used is governed by:

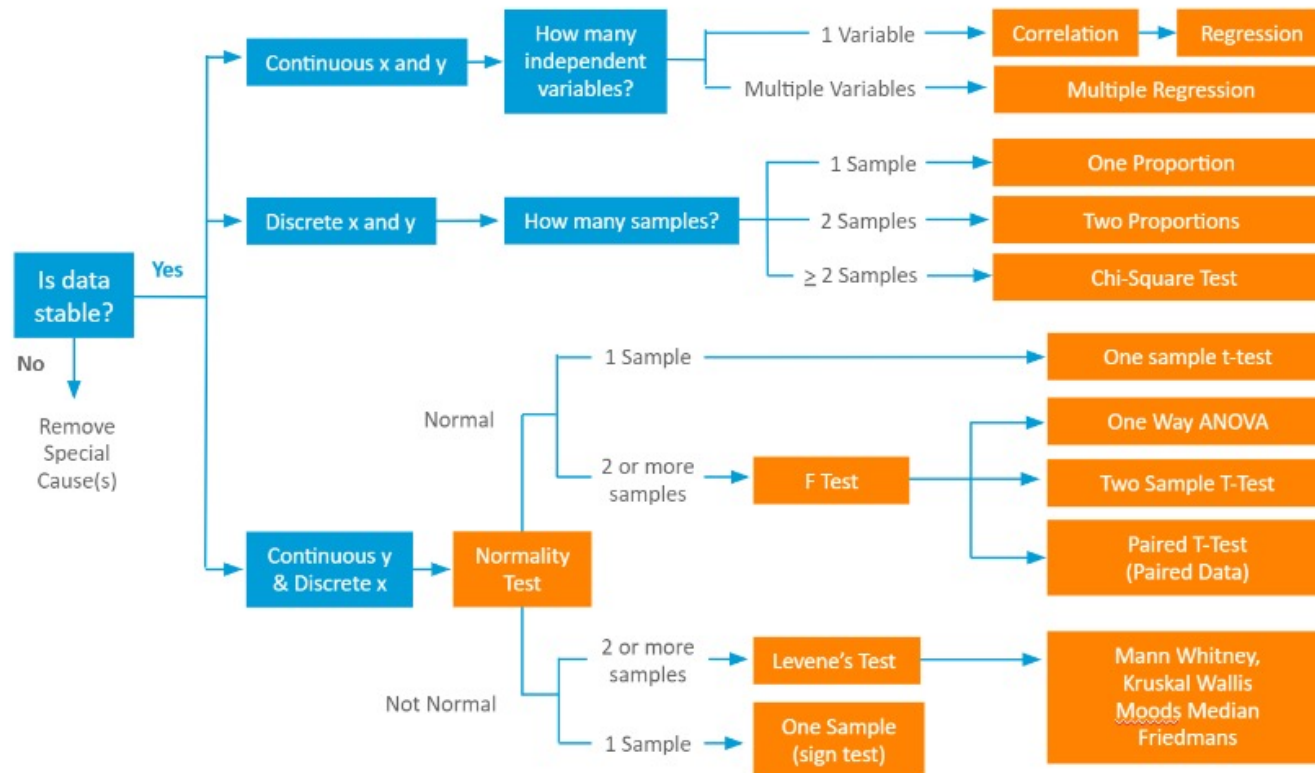
- Number of samples
- Type of data
- What we are trying to test
 - Difference in means
 - Difference in standard deviation (or variance)
 - Difference in medians
- We used DECISION CHARTS to choose the correct Hypothesis test to use
- (Step 4 of the Hypothesis Testing method)

Different Types of Hypothesis Test

Note: Minitab uses the phrase 'Non-parametric' instead of 'Non-normal'.



HYPOTHESIS TESTING ROADMAP



SigmaXL's Nonparametric tests

1-Sample Sign: performs a one-sample sign test of the Median and calculates the corresponding point estimate and confidence interval. Use this test as an alternative to one-sample Z and one-sample t-tests.

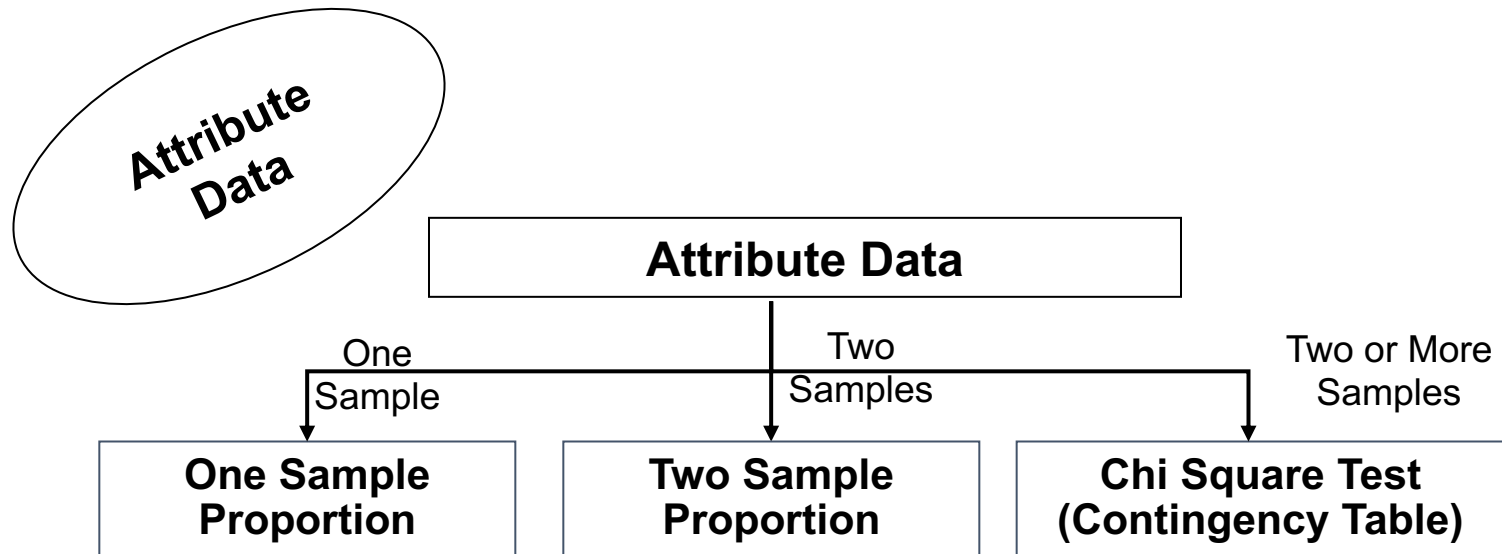
1-Sample Wilcoxon: performs a one-sample Wilcoxon signed rank test of the Median and calculates the corresponding point estimate and confidence interval (more discriminating or efficient than the sign test). Use this test as a nonparametric alternative to one-sample Z and one-sample t-tests.

Mann-Whitney: performs a Hypothesis Test of the equality of two population Medians and calculates the corresponding point estimate and confidence interval. Use this test as a nonparametric alternative to the two-sample t-test.

Kruskal-Wallis: performs a Hypothesis Test of the equality of population Medians for more than 2 samples. This test is more powerful than Mood's Median (the confidence interval is narrower, on average) for analyzing data from many populations, but is less robust to outliers. Use this test as an alternative to the one-way ANOVA.

Mood's Median Test: performs a Hypothesis Test of the equality of population Medians for more than 2 samples. Test is similar to the Kruskal-Wallis Test.

Hypothesis Testing Roadmap Attribute Data



Principles

While using Hypothesis Testing the following principles should be applied:

- The decision is about H_0 and NOT H_A .
- The null hypothesis H_0 is on trial.
- When a decision has been made:
 - Nothing has been proved.
 - All decisions can lead to errors (Types I and II).

Summary

At this point you should be able to:

- Explain the purpose of Hypothesis Testing
- Understand what is Alpha Risk and a Type I error
- Understand the concept of the CONFIDENCE LEVEL
- Be familiar with the types of Hypothesis Tests