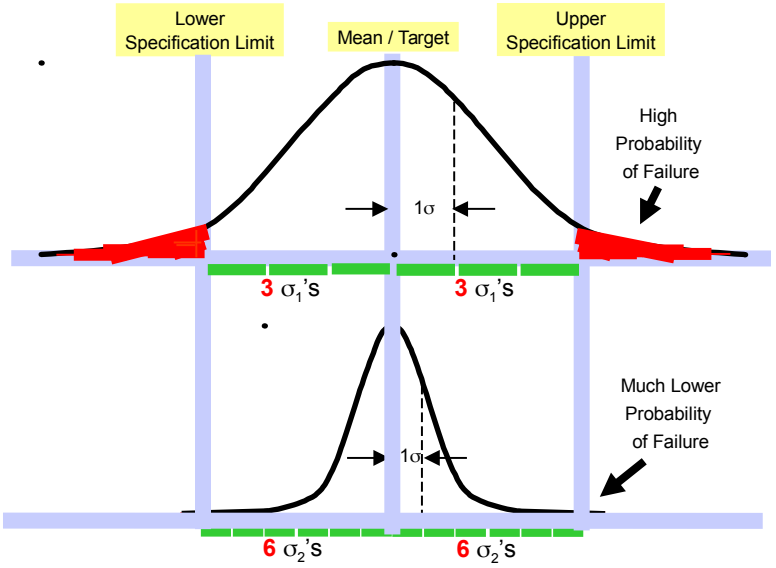
Minitab Training



Learning Objectives

- Understand the structure of Minitab
- Understand Tools in Minitab.
- Understand correct data structure for analysis in Minitab.
- Be able to create and interpret basic graphs in Minitab
- Be able to interpret Tools used in Minitab in DMAIC phases

Statistical software

- Minitab

Training Outline

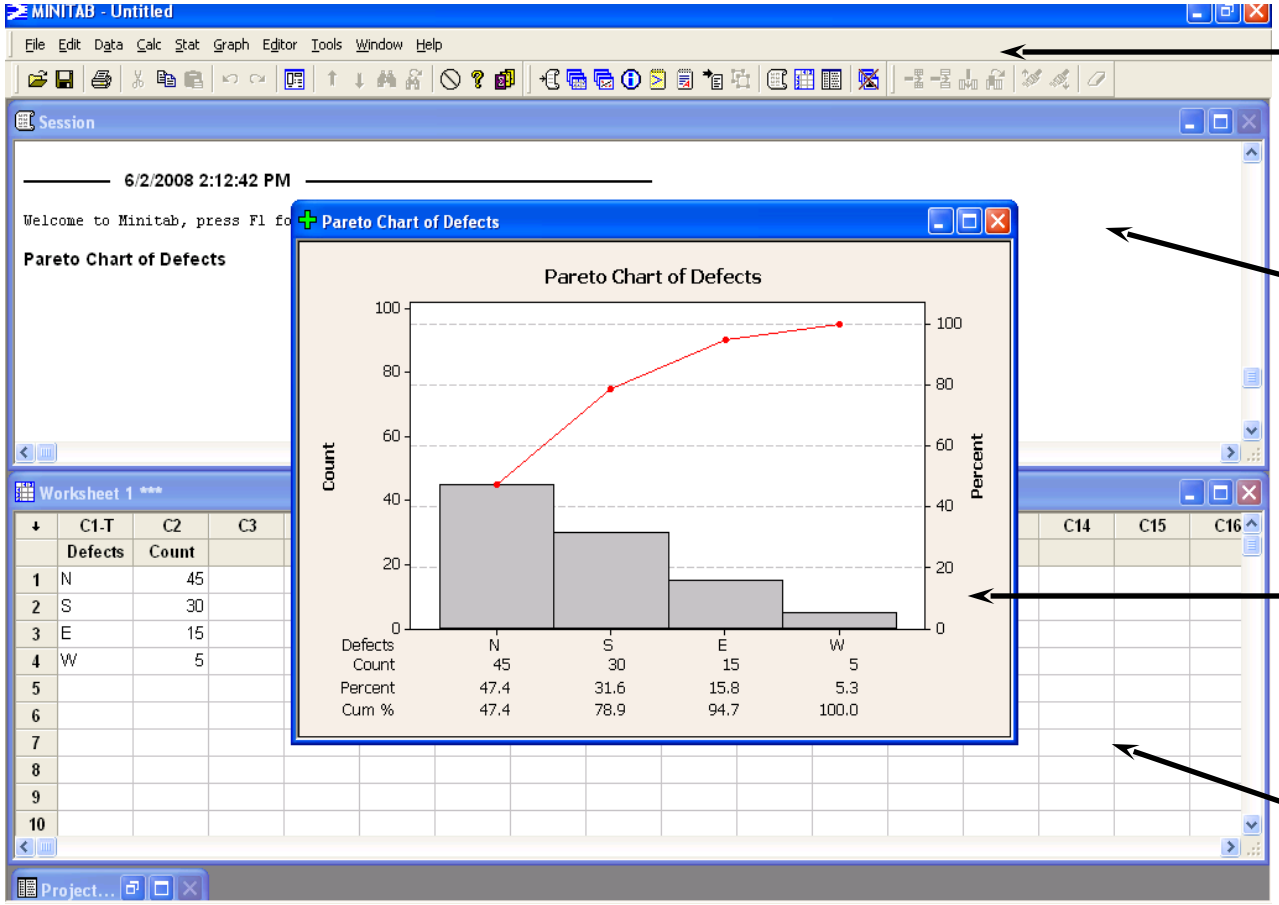
Date	Timings	Subject / Topics
Day I	09. 00 to 11.00	Introduction to Statistical Software –Minitab Menus in Minitab , Data Menu
	11.00 to 11.15	Tea Break
	11.15 to 12.30	Descriptive Statistics, Graphical Tools – Pareto, Box Plot, Dot Plot, Scatter Plot, Time Series, Run Chart, Histogram
	12.30 to 14.00	Lunch Break
	14.00 to 15.00	Normality Test, Individual Distribution Identification, Box Cox Transformation
	15.00 to 15.15	Tea Break
	15.15 to 18.00	Process Capability, MSA (ANOVA Method)

Training Outline

Date	Timings	Subject / Topics
Day II	09.00 to 10.00	Hypothesis Test – 1T test, 2T Test, Test for Equal Variance (F-Test)
	10.00 to 11.00	Hypothesis Test -ANOVA (1 way, 2way)
	11.00 to 11.15	Tea Break
	11.15 to 12.30	Hypothesis Test – 1P,2P, Chi square
	12.30 to 14.00	Lunch Break
	14.00 to 15.00	Multi vari charts
	15.00 to 15.15	Tea Break
	15.15 to 18.00	Regression -SLR, MLR SPC -Control Charts – I &MR, Xbar-R, P,C,U,NP

Minitab - Introduction

Worksheet Format and Structure



Menu Bar

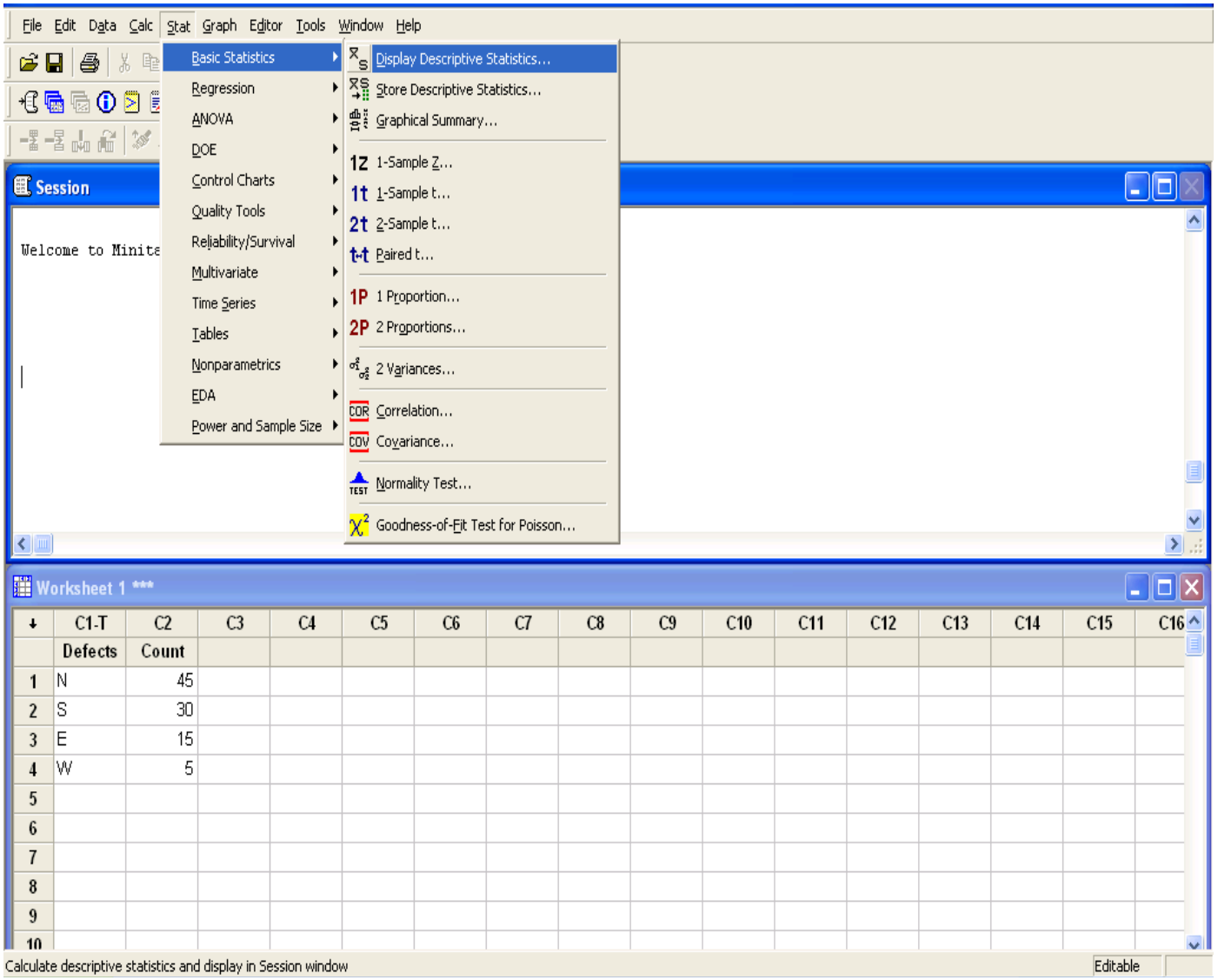
Session Window:
Analytical Output

Graph Window

Data Window:
A Worksheet

Minitab – Introduction

Step 2: Select menu command

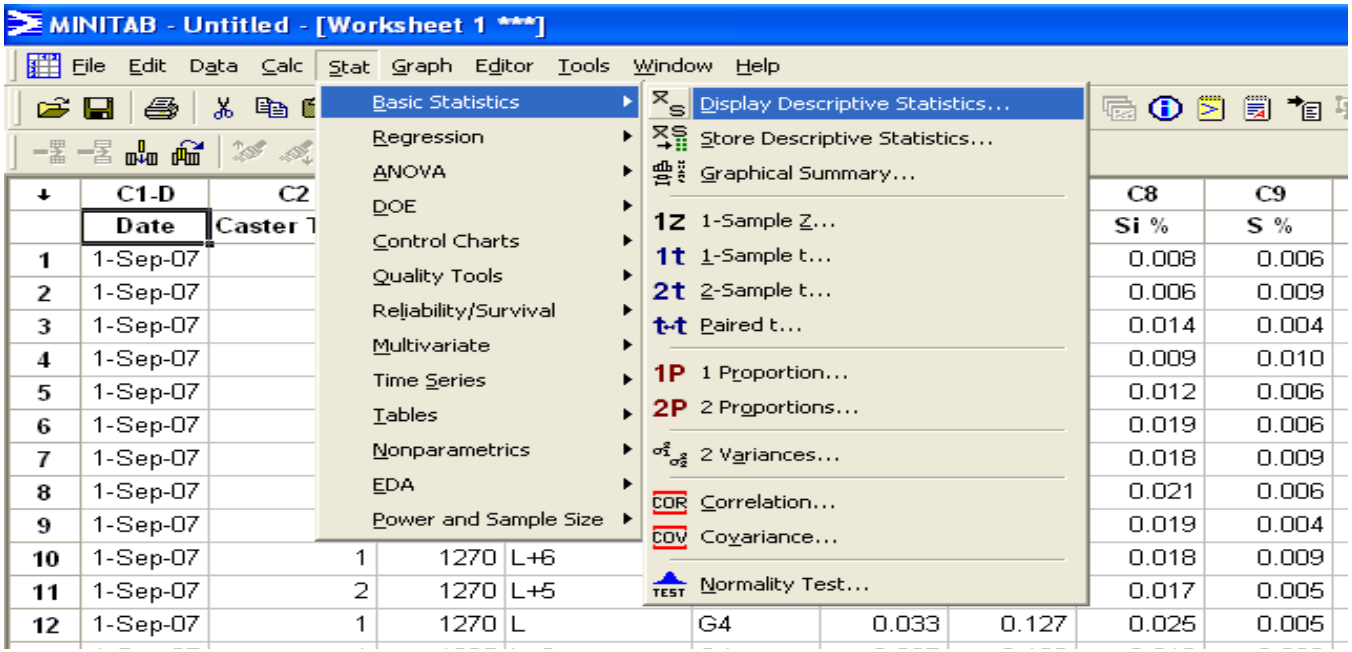


Stat→ Basic Statistics →Display Descriptive Statistics

Example - Descriptive statistics of Actual Average speed

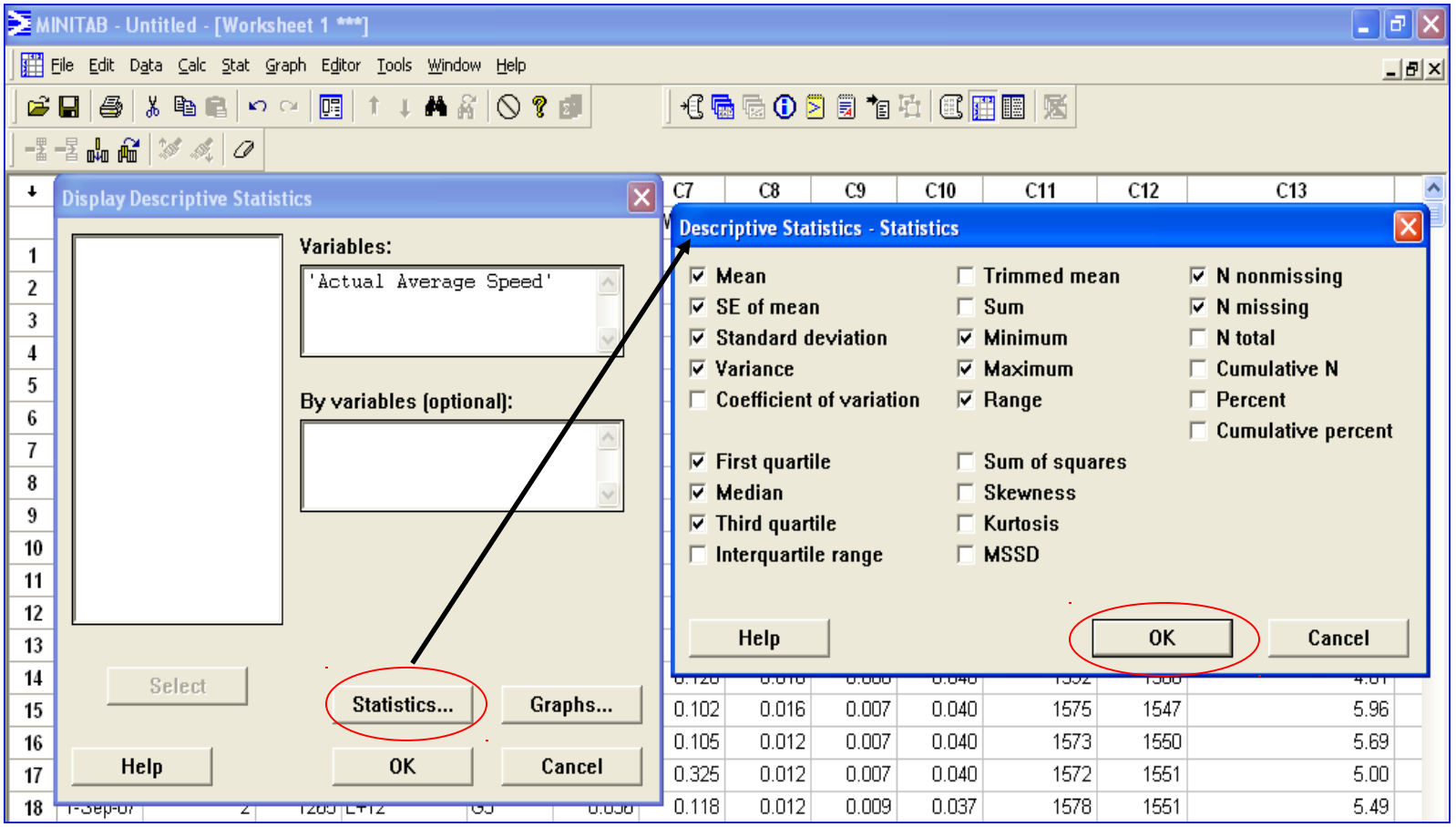
File Name – Minitab training

Worksheet Name – Descrip Statistics



Stat→ Basic Statistics →Display Descriptive Statistics

Example - Descriptive statistics of Actual Average speed



Stat→ Basic Statistics →Display Descriptive Statistics - Output

Example - Descriptive statistics of Actual Average speed

Descriptive Statistics: Actual Average Speed

Variable	N	N*	Mean	SE Mean	StDev	Variance	Minimum	Q1
Actual Average S	123	0	5.3896	0.0321	0.3565	0.1271	4.4500	5.1400

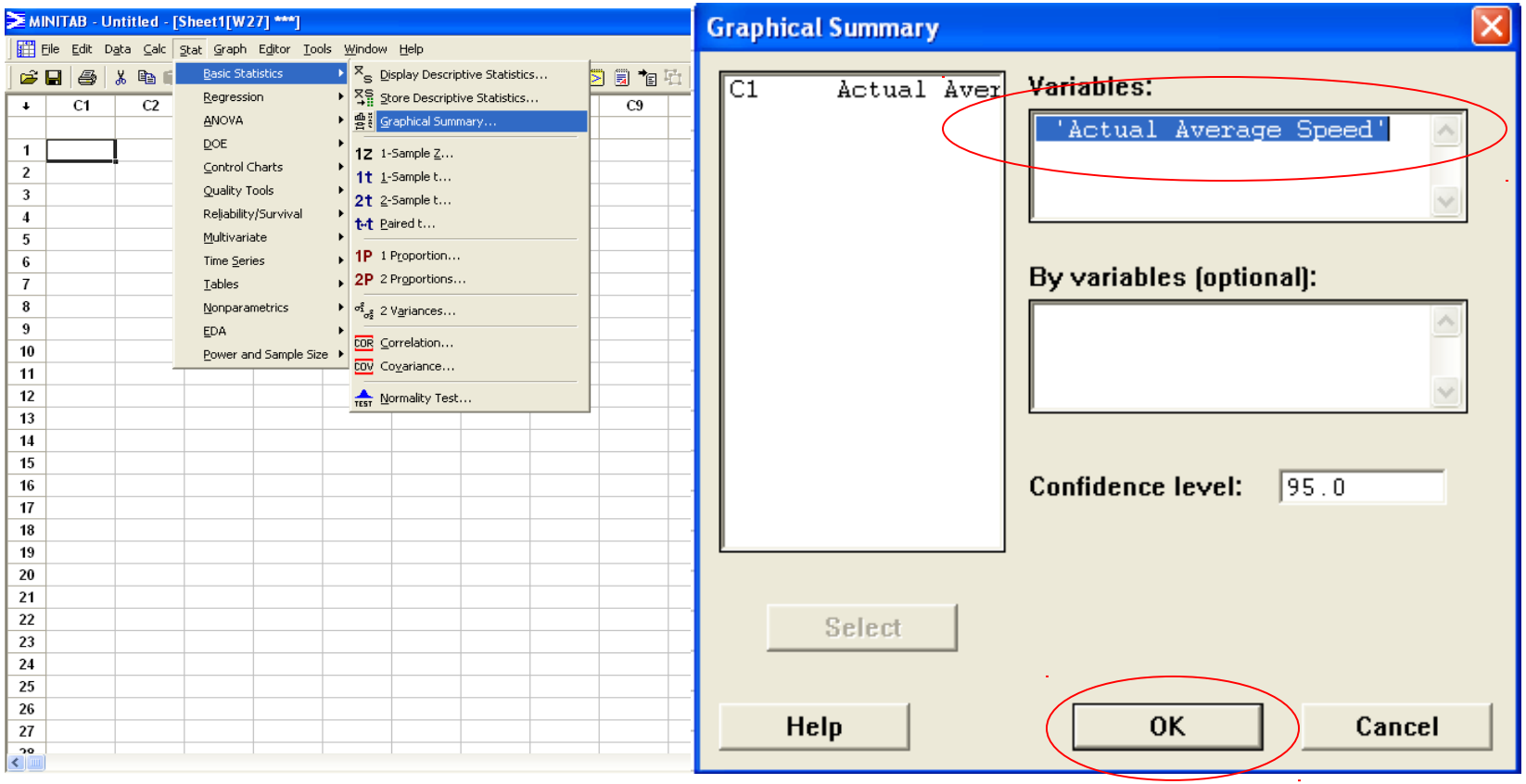
Variable	Median	Q3	Maximum	Range
Actual Average S	5.4500	5.6500	5.9600	1.5100

Interpreting the results

The session window displays the statistics selected. The Q1 and Q3 are the First and third Quartile. The Minimum and Maximum and the Range shows the Range of the Actual Average Speed with the Median greater than Mean showing the Negative skew ness of the actual average speed.

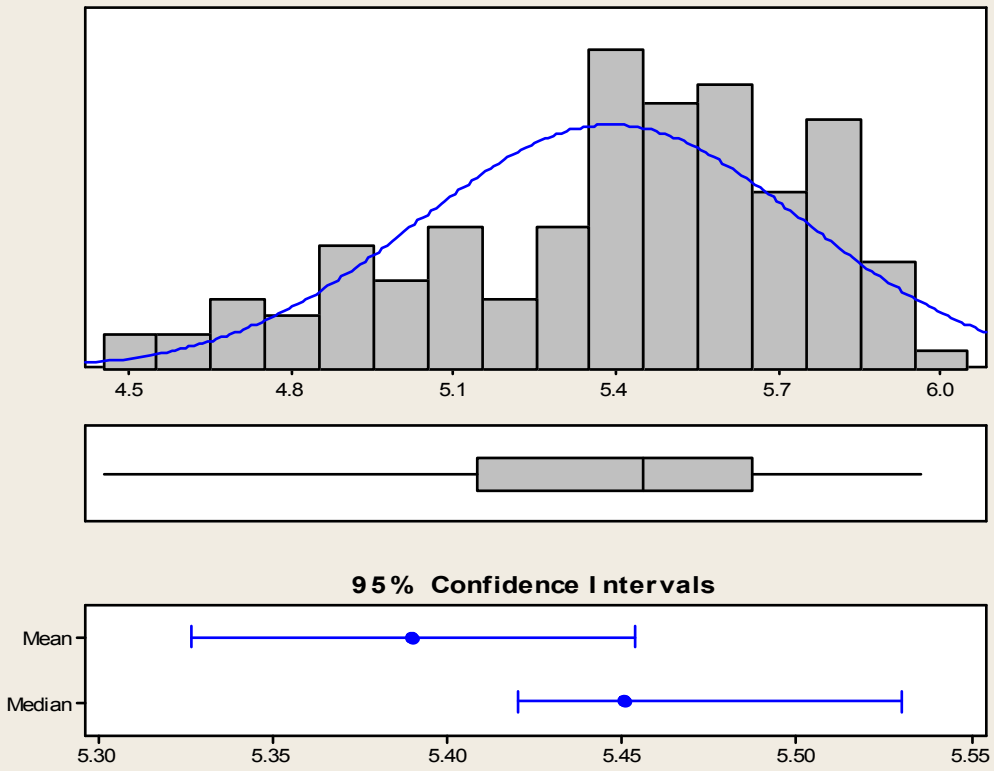
Stat→ Basic Statistics →Graphical Summary

Example – Graphical summary of Actual Average speed



Stat→ Basic Statistics →Graphical Summary - Output

Example - Descriptive statistics of Actual Average speed



Anderson-Darling Normality Test	
A-Squared	2.34
P-Value <	0.005
Mean	5.3896
StDev	0.3565
Variance	0.1271
Skewness	-0.712751
Kurtosis	-0.263699
N	123
Minimum	4.4500
1st Quartile	5.1400
Median	5.4500
3rd Quartile	5.6500
Maximum	5.9600
95% Confidence Interval for Mean	
5.3260	5.4532
95% Confidence Interval for Median	
5.4200	5.5300
95% Confidence Interval for StDev	
0.3169	0.4076

Kurtoisis:. Negative values indicate a distribution that is flatter than normal. Positive values indicate a distribution with a sharper than normal peak.

Stat→ Basic Statistics →Graphical Summary - Output

Interpreting the results

The P-Value is less than 0.05 indicates that data do not follow a Normal distribution.

Mean is less than Median showing that data is Negatively skewed

Kurtosis of a curve is a measure of how different is a distribution from the Normal Distribution.

95% Confidence Interval of Mean indicates that the true mean can vary from 5.3260 and 5.4532.

Stat→ Basic Statistics - Class Exercise

Class Exercise –

Conduct a) Descriptive statistics of “Lifting Temp” and b) Graphical Summary of Lifting Temp with the help of the data from the below file using Minitab

File Name – Minitab training

Worksheet Name – Descrip statistics

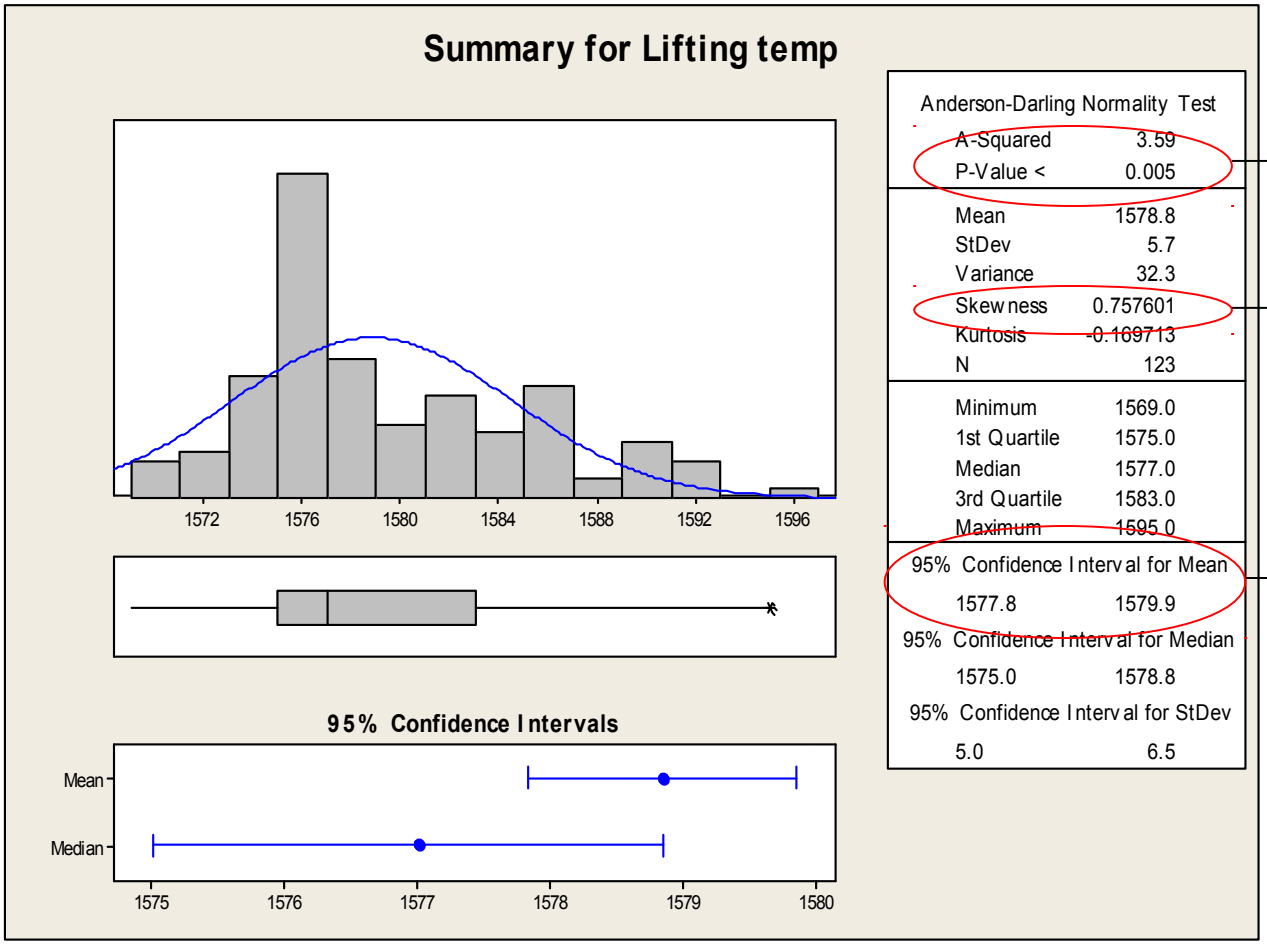
Stat→ Basic Statistics - Class Exercise - Output

Descriptive Statistics: Lifting temp

Variable	N	N*	Mean	SE Mean	StDev	Variance	Minimum	Q1
Lifting temp	123	0	1578.8	0.512	5.68	32.3	1569.0	1575.0

Variable	Median	Q3	Maximum	Range
Lifting temp	1577.0	1583.0	1595.0	26.0

Stat→ Basic Statistics - Class Exercise - Output



Indicates data is not Normally distributed

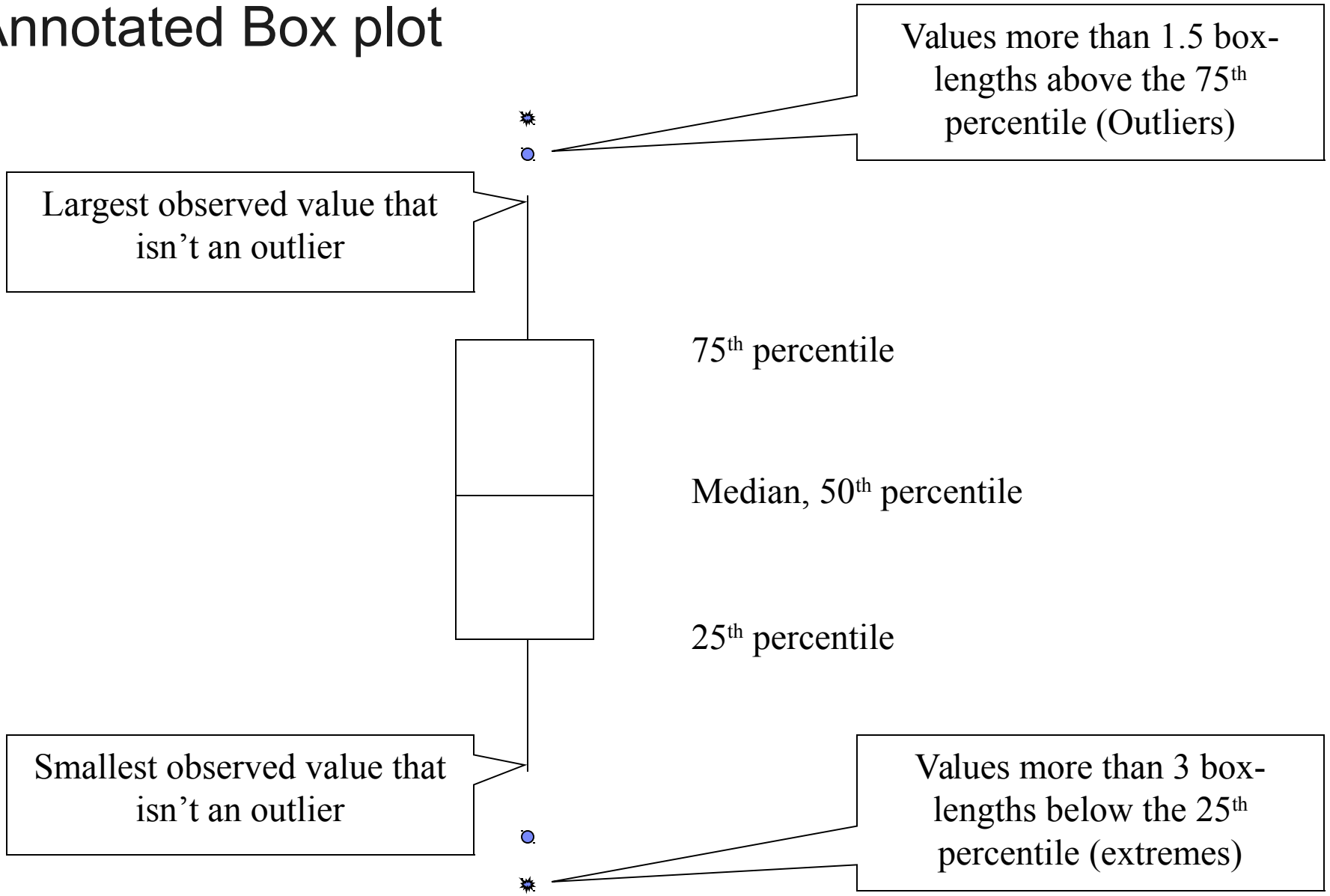
Indicates Positive skewness of the data

Indicates confidence interval of true mean

Box Plot

- Useful for comparing the distribution of values in several group
- Make it easy to see the different properties of the distribution
  - Location
  - Variability
  - Shapes

Annotated Box plot

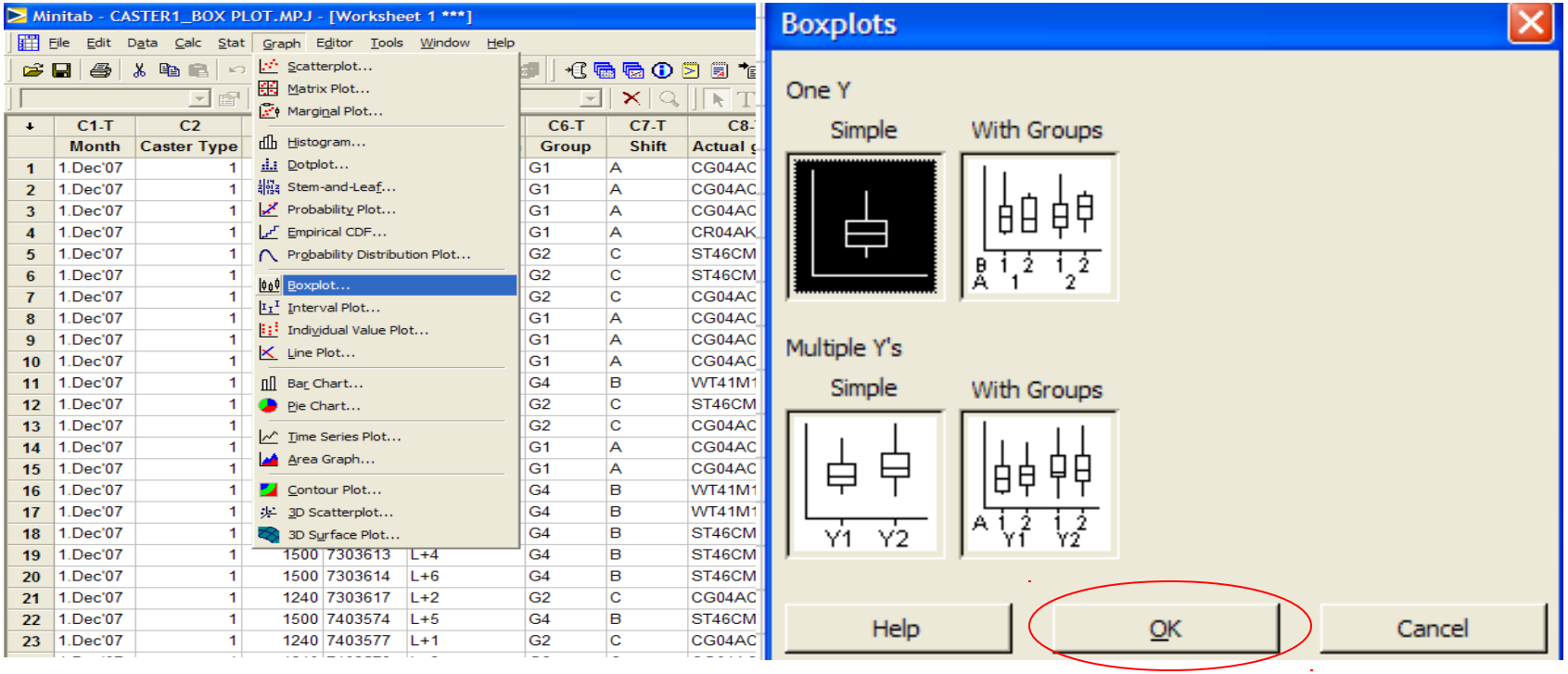


Graph → Boxplot - Example

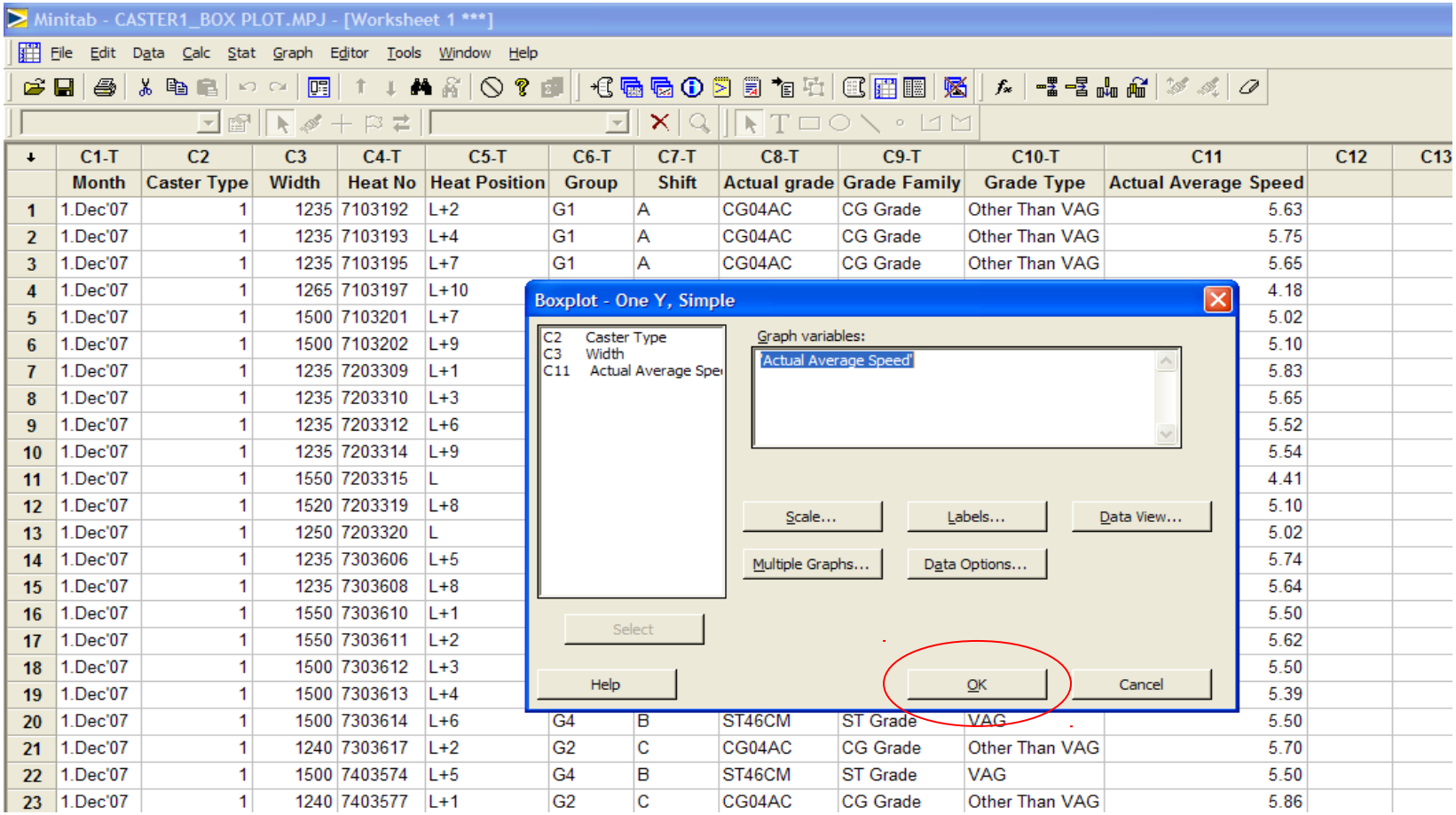
Example – Input the data from the following worksheet

File Name – Minitab training

Worksheet Name – Box Plot



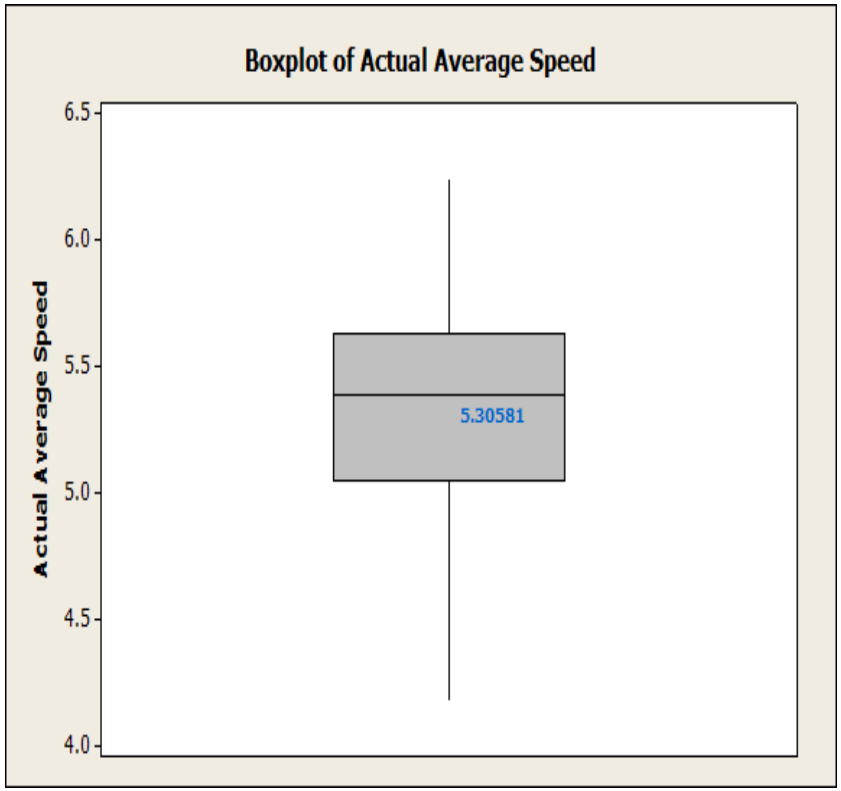
Graph → Boxplot - Example



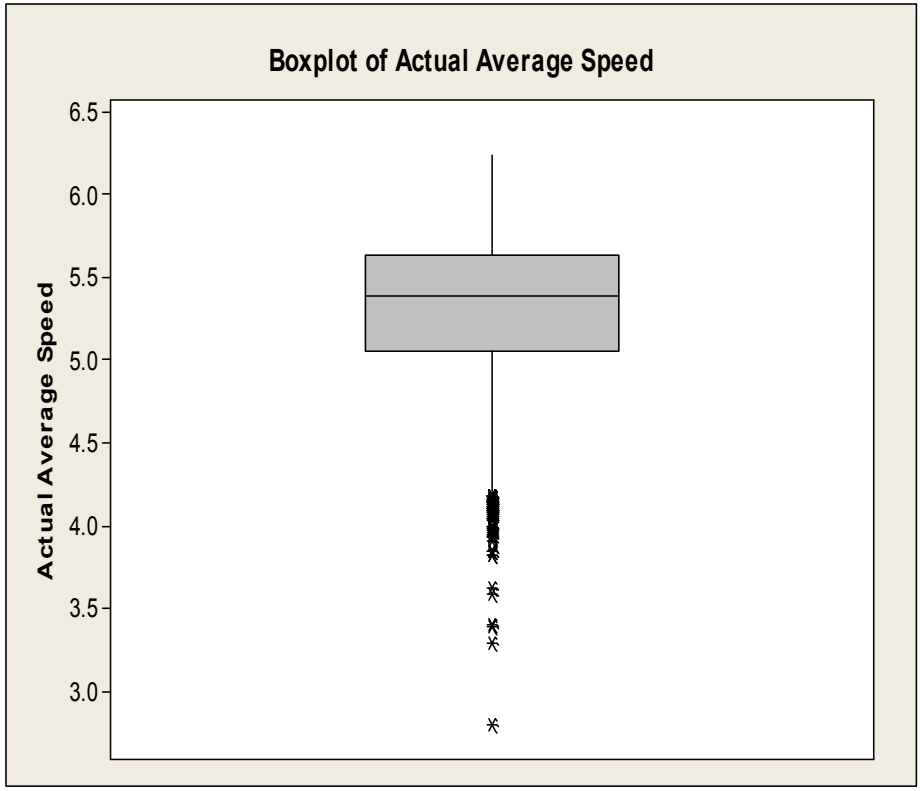
	C1-T	C2	C3	C4-T	C5-T	C6-T	C7-T	C8-T	C9-T	C10-T	C11	C12	C13
	Month	Caster Type	Width	Heat No	Heat Position	Group	Shift	Actual grade	Grade Family	Grade Type	Actual Average Speed		
1	1.Dec'07	1	1235	7103192	L+2	G1	A	CG04AC	CG Grade	Other Than VAG	5.63		
2	1.Dec'07	1	1235	7103193	L+4	G1	A	CG04AC	CG Grade	Other Than VAG	5.75		
3	1.Dec'07	1	1235	7103195	L+7	G1	A	CG04AC	CG Grade	Other Than VAG	5.65		
4	1.Dec'07	1	1265	7103197	L+10						4.18		
5	1.Dec'07	1	1500	7103201	L+7						5.02		
6	1.Dec'07	1	1500	7103202	L+9						5.10		
7	1.Dec'07	1	1235	7203309	L+1						5.83		
8	1.Dec'07	1	1235	7203310	L+3						5.65		
9	1.Dec'07	1	1235	7203312	L+6						5.52		
10	1.Dec'07	1	1235	7203314	L+9						5.54		
11	1.Dec'07	1	1550	7203315	L						4.41		
12	1.Dec'07	1	1520	7203319	L+8						5.10		
13	1.Dec'07	1	1250	7203320	L						5.02		
14	1.Dec'07	1	1235	7303606	L+5						5.74		
15	1.Dec'07	1	1235	7303608	L+8						5.64		
16	1.Dec'07	1	1550	7303610	L+1						5.50		
17	1.Dec'07	1	1550	7303611	L+2						5.62		
18	1.Dec'07	1	1500	7303612	L+3						5.50		
19	1.Dec'07	1	1500	7303613	L+4						5.39		
20	1.Dec'07	1	1500	7303614	L+6	G4	B	ST46CM	ST Grade	VAG	5.50		
21	1.Dec'07	1	1240	7303617	L+2	G2	C	CG04AC	CG Grade	Other Than VAG	5.70		
22	1.Dec'07	1	1500	7403574	L+5	G4	B	ST46CM	ST Grade	VAG	5.50		
23	1.Dec'07	1	1240	7403577	L+1	G2	C	CG04AC	CG Grade	Other Than VAG	5.86		

Graph → Boxplot - Output

Without Outlier Symbol



With Outlier Symbol



Box Plot – Interpretation of Result

- the Central tendency, location of distribution
  - From the median
- the spread, variability of distribution
  - From the length of the Box
- Skew ness
  - If the median is not in the centre
- Positive Skew ness
  - If median is closer to the bottom of the Box
- Negative Skew ness
  - If the median is closer to top of the box
- the length of the tail
  - From the whiskers

Graph → Boxplot –> With Group

Class Exercise

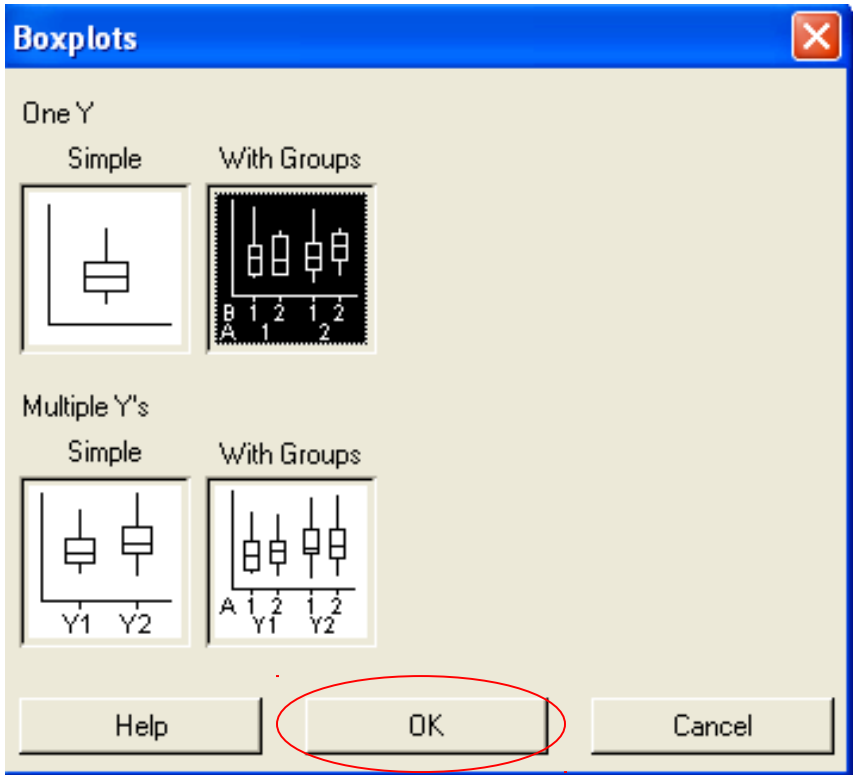
Draw the box plot for Casting speed Group wise and find which group is performing better.

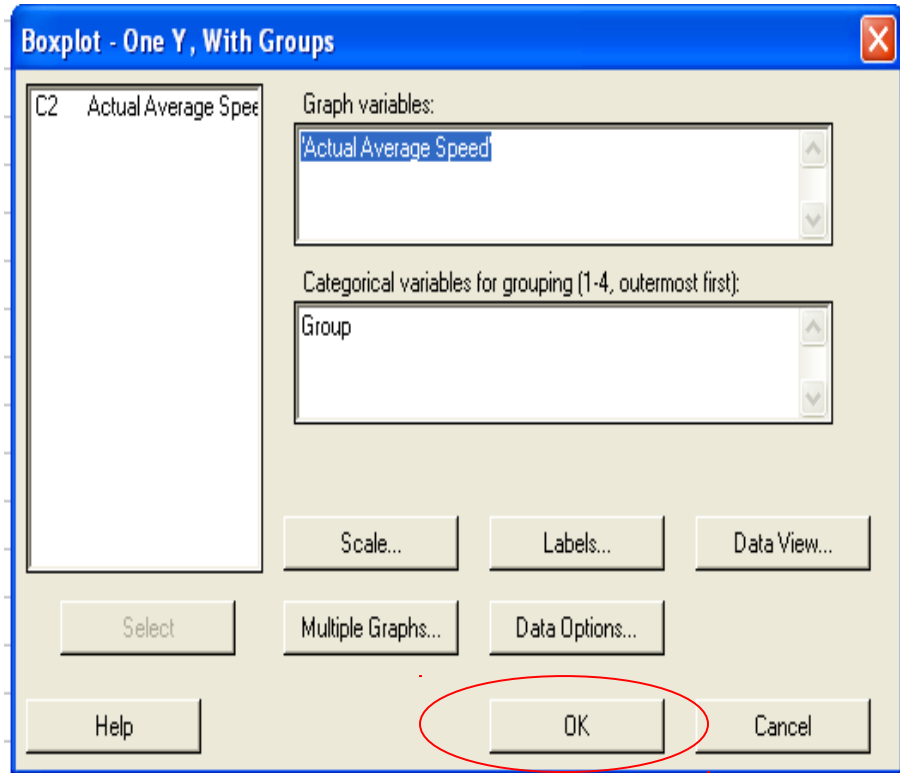
File Name – Minitab training

Worksheet Name – Boxplot

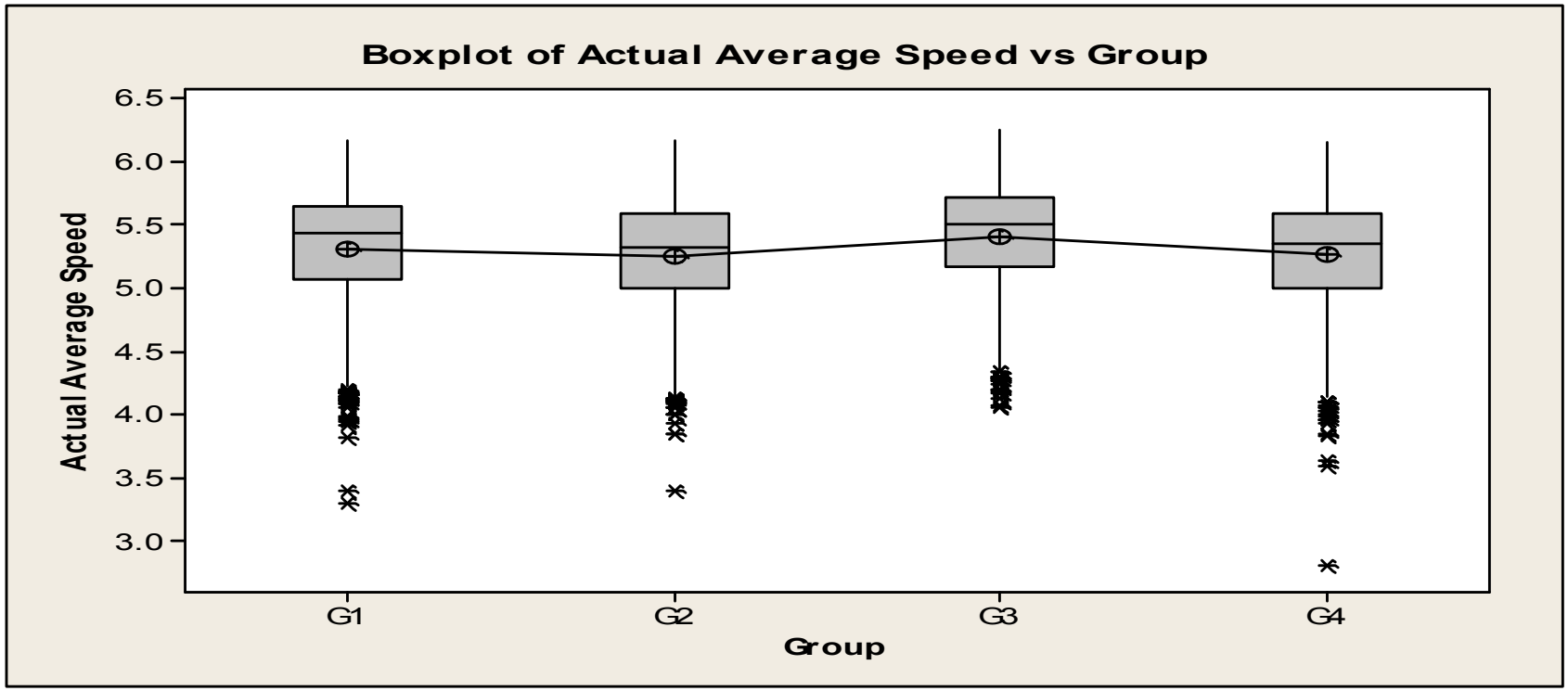
Graph → Boxplot –> With Group

Class Exercise





Graph → Boxplot – Output



Interpreting the results

Box Plot output indicates that Group G3 is having a higher Mean and Median than the other groups

The outliers are far spread in Group G1,G2 and G4 as compared to G3

Pareto Chart

Example

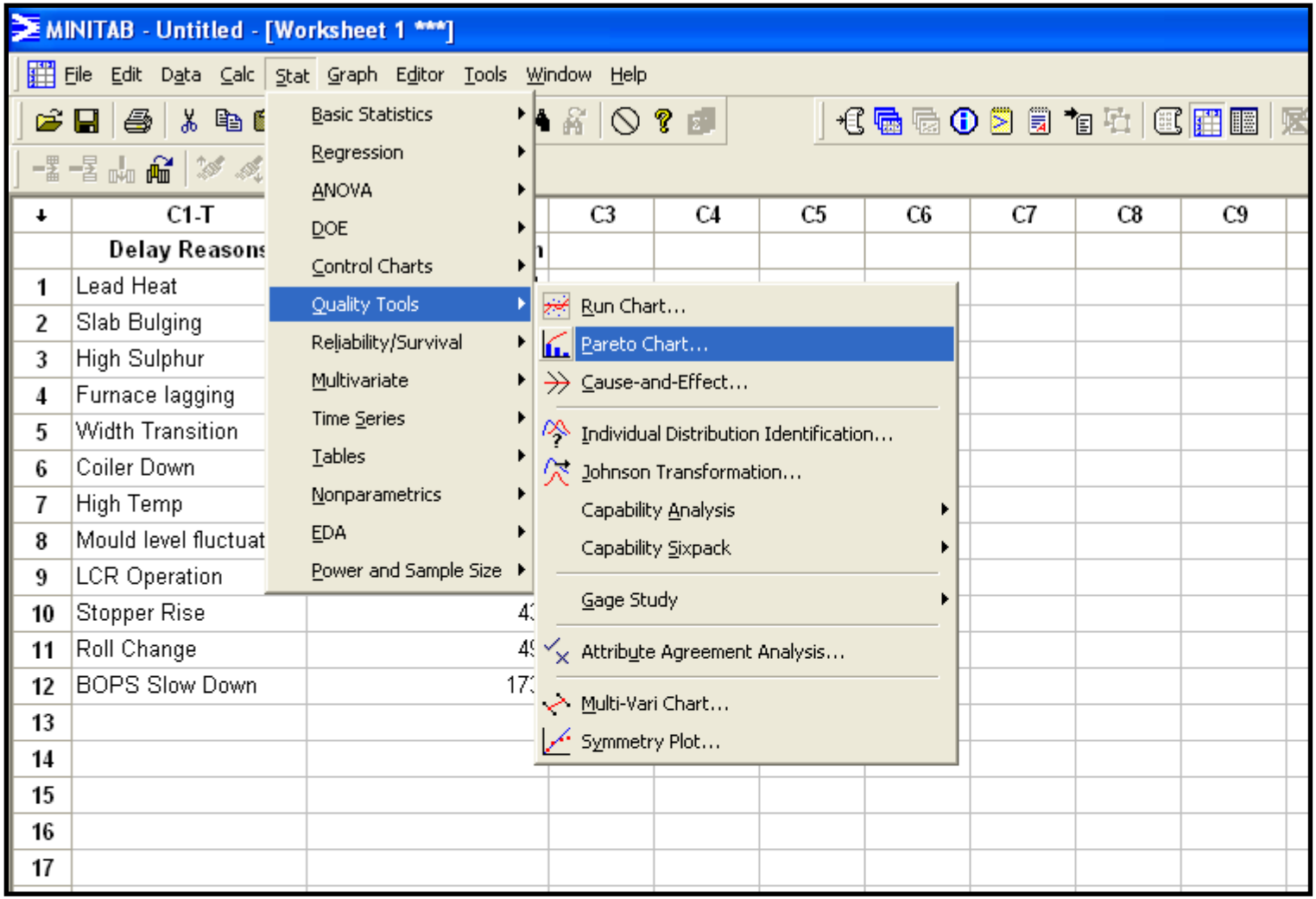
The following table consist of the delay reasons which causing delay at Caster2 in the month of July'08 and the corresponding sum of delays in minute. The management wants to take decision based upon the following data to reduce the delay.

File Name – Minitab training

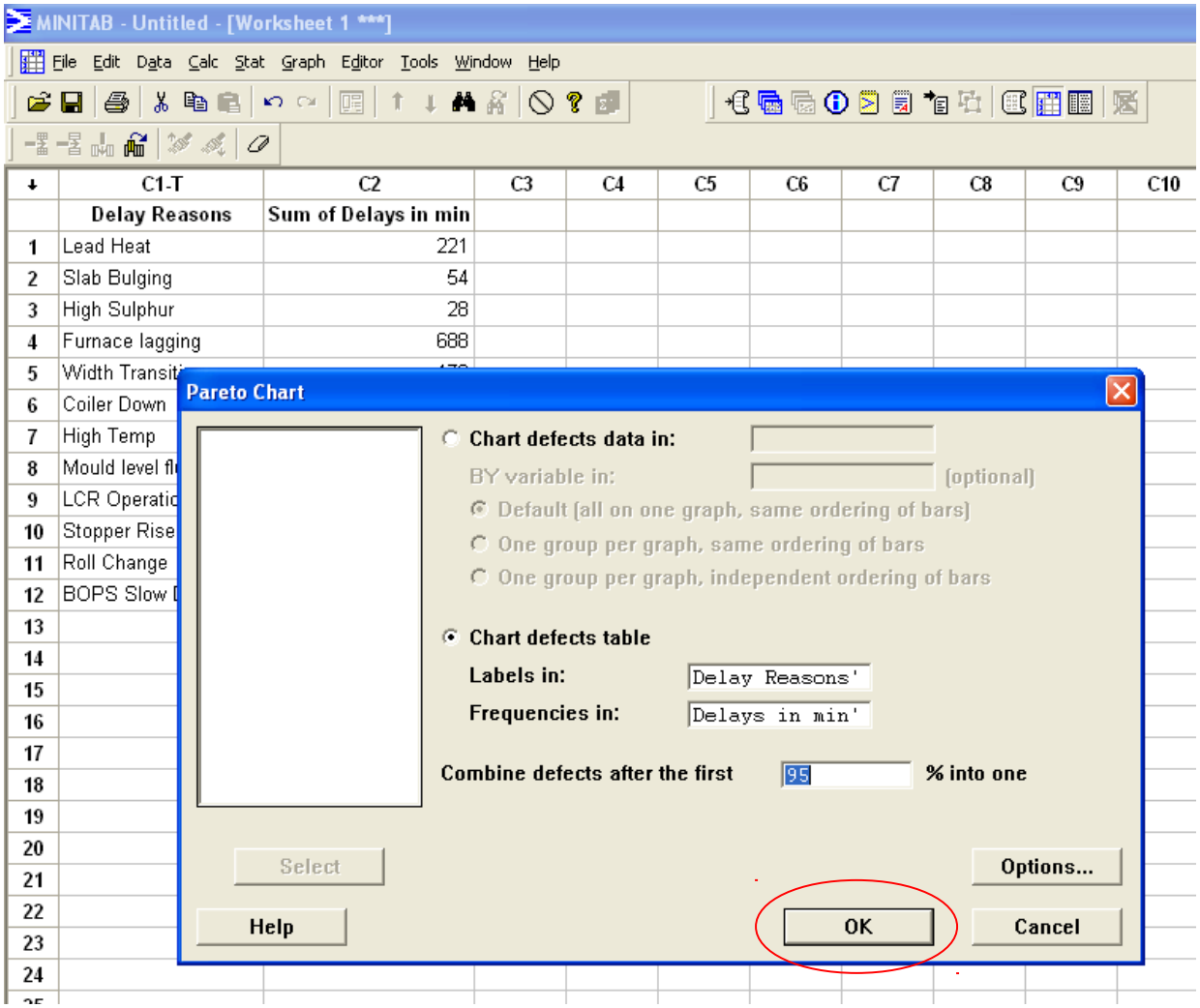
Worksheet Name – Pareto-s

Delay Reasons	Sum of Delays in min	Count of Delays	Agency
Lead Heat	351	29	Caster
Slab Bulging	54	25	Caster
High Sulphur	28	7	LF
Furnace lagging	752	120	EAF
Width Transition	125	52	Marketing
Coiler Down	67	11	Mill
High Temp	58	32	LF
Mould level fluctuatuon	68	26	Caster
LCR Operation	79	11	Mill
Stopper Rise	43	9	LF
Roll Change	49	16	Mill
BOPS Slow Down	194	25	Caster

Stat→ Quality tools →Pareto chart

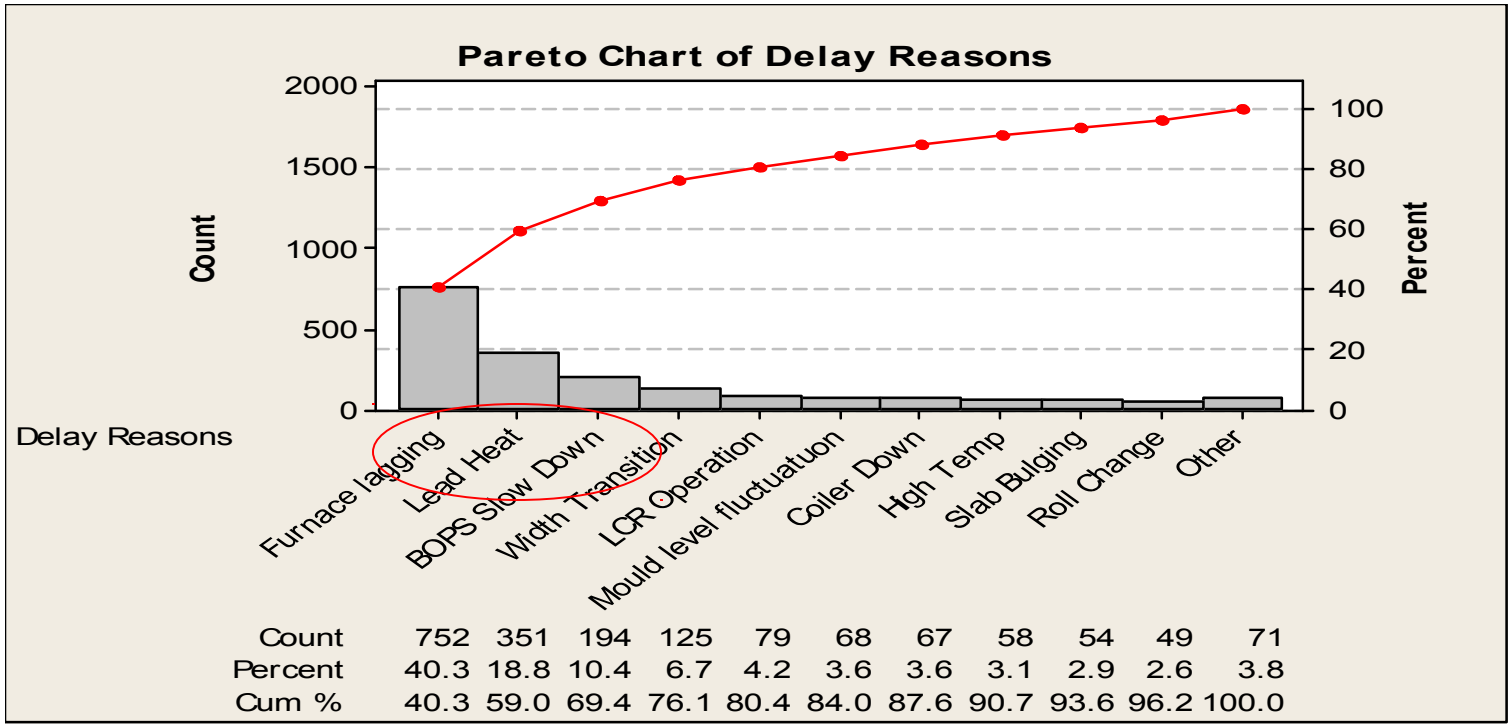


Stat→ Quality tools →Pareto chart



	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Delay Reasons	Sum of Delays in min								
1	Lead Heat	221								
2	Slab Bulging	54								
3	High Sulphur	28								
4	Furnace lagging	688								
5	Width Transiti									
6	Coiler Down									
7	High Temp									
8	Mould level fl									
9	LCR Operatic									
10	Stopper Rise									
11	Roll Change									
12	BOPS Slow [									
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										

Stat→ Quality tools →Pareto chart - Output



Interpreting the results

The first three reasons of delay contributes to 69.4% of the total delay.

You can suggest management to eliminate/reduce the causes of the delay of first four to reduce the overall delay by 76%

Stat→ Quality tools →Pareto chart – Class Exercise

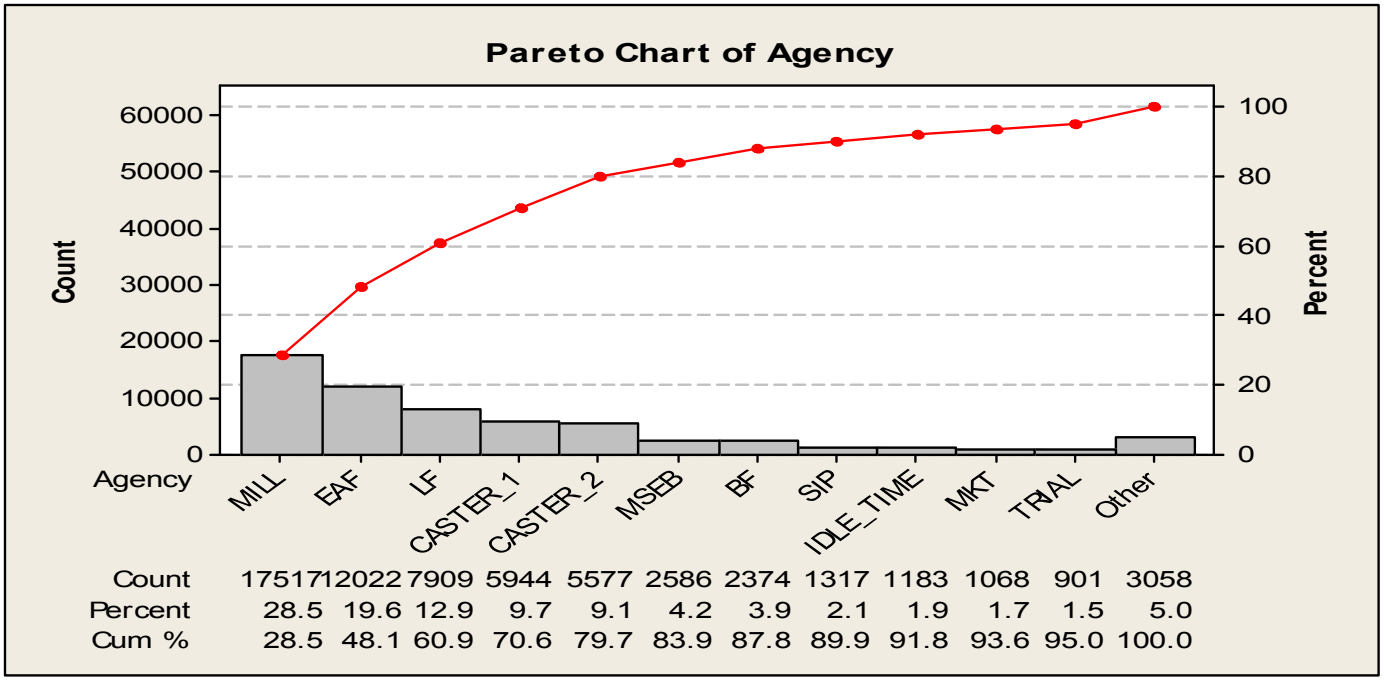
Class Exercise

Draw the Pareto Chart for the following data in the worksheet . The Pareto chart to be drawn between ‘Agency’ and ‘delays in min’ data

File Name – Minitab training

Worksheet Name – Pareto-C

Stat→ Quality tools →Pareto chart – Class Exercise



Interpreting the results

The first four parameters contributes to 70.6% of the total delay.

As a second step of data analysis, we can further identify the factors affecting the above parameters.

Stat→ Quality tools →Pareto chart – Class Exercise

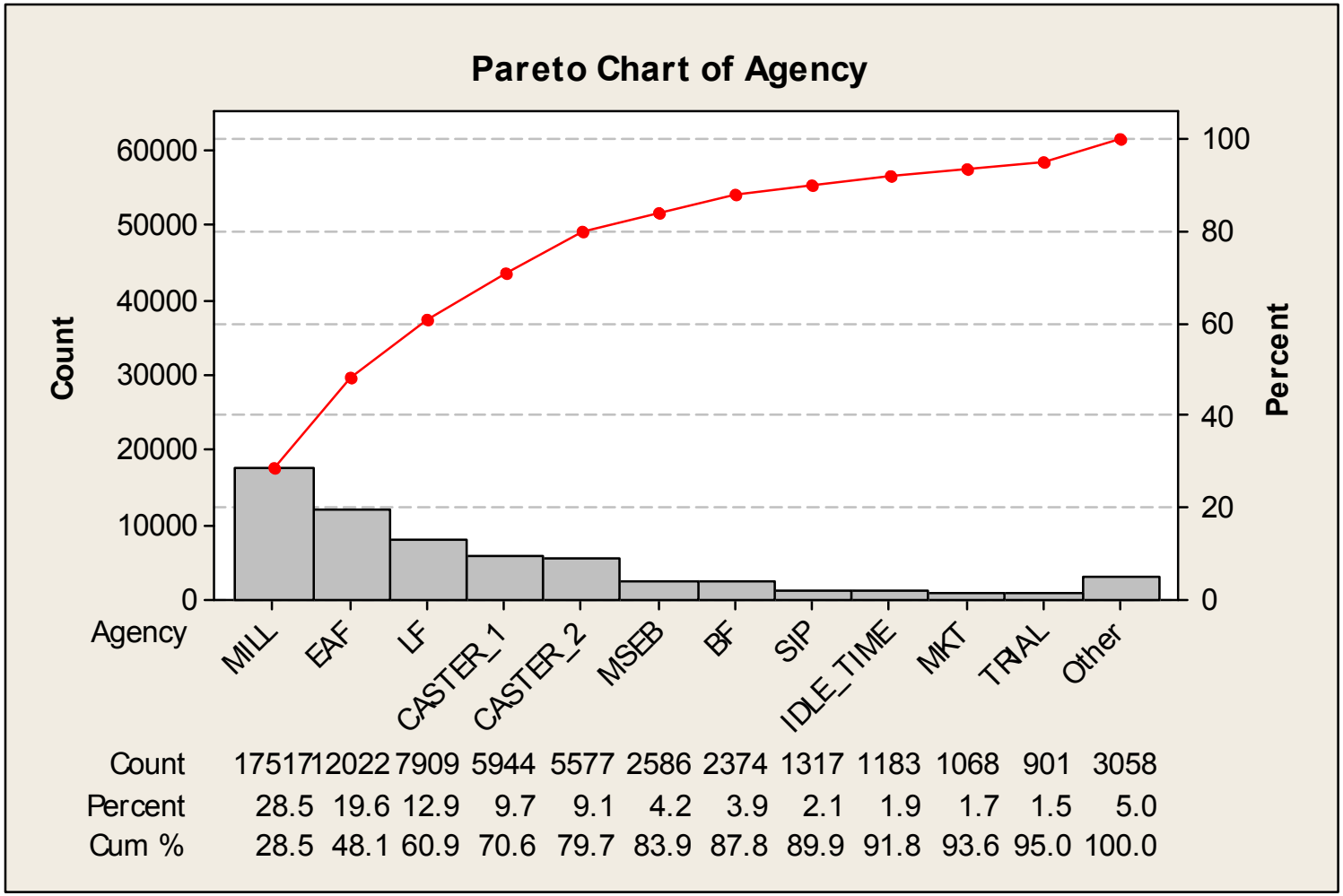
Class Exercise

Draw the Pareto Chart for the following data in the worksheet . The Pareto chart to be drawn between ‘Agency1’ and ‘delays’ data

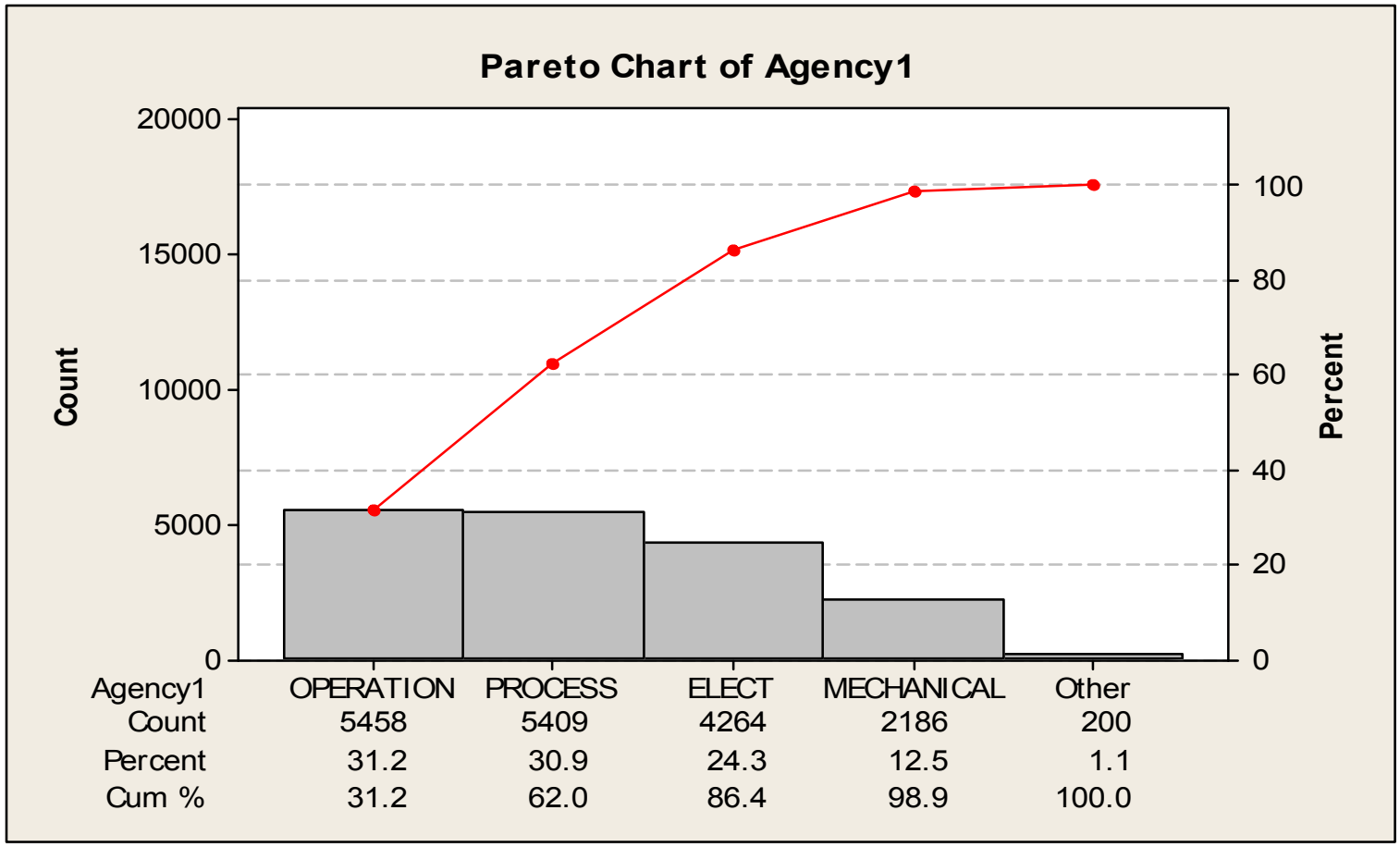
File Name – Minitab training

Worksheet Name – Pareto-C

Class Exercise –

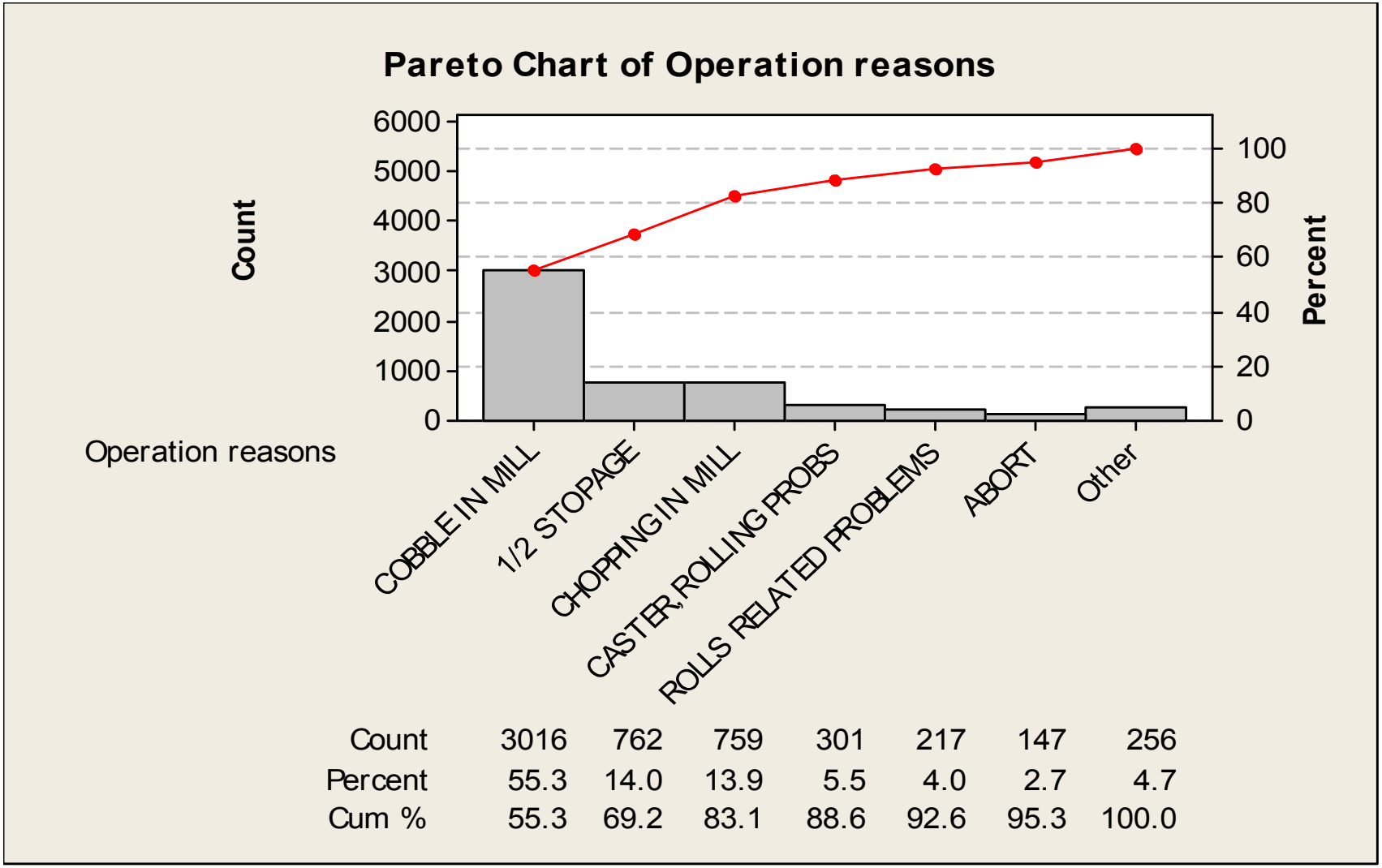


Stat→ Quality tools →Pareto chart – Class Exercise



Agency1	OPERATION	PROCESS	ELECT	MECHANICAL	Other
Count	5458	5409	4264	2186	200
Percent	31.2	30.9	24.3	12.5	1.1
Cum %	31.2	62.0	86.4	98.9	100.0

Class Exercise –



Count	3016	762	759	301	217	147	256
Percent	55.3	14.0	13.9	5.5	4.0	2.7	4.7
Cum %	55.3	69.2	83.1	88.6	92.6	95.3	100.0

Scatter Plot

The scatter diagram graphs pairs of numerical data, with one variable on each axis, to look for a relationship between them.

If the variables are correlated, the points will fall along a line or curve.

The better the correlation, the tighter the points will hug the line.

When to Use a Scatter Plot

- When you have paired numerical data.
- When your dependent variable may have multiple values for each value of your independent variable.
- When trying to determine whether the two variables are related, such as…
  - When trying to identify potential root causes of problems.
  - After brainstorming causes and effects using a fishbone diagram, to determine objectively whether a particular cause and effect are related.
  - When determining whether two effects that appear to be related both occur with the same cause.
  - When testing for autocorrelation before constructing a control chart.

Generating Scatter Plots in Minitab - Example

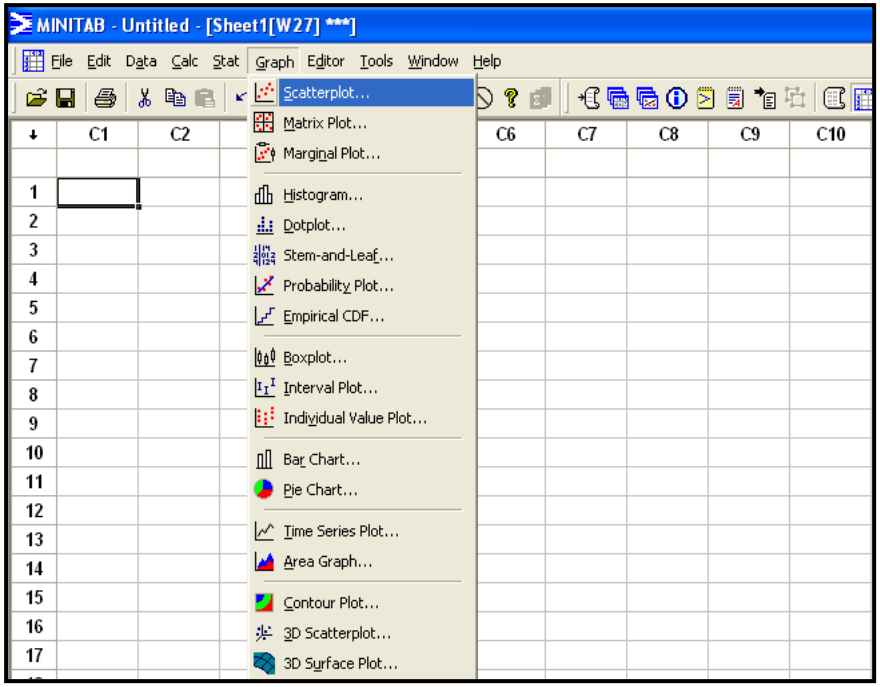
- For generating scatter plots, we must have both X & Y as continuous
- Data points for both should be entered in Minitab
- Let's say we have following data points on 'X' & its corresponding 'Y'

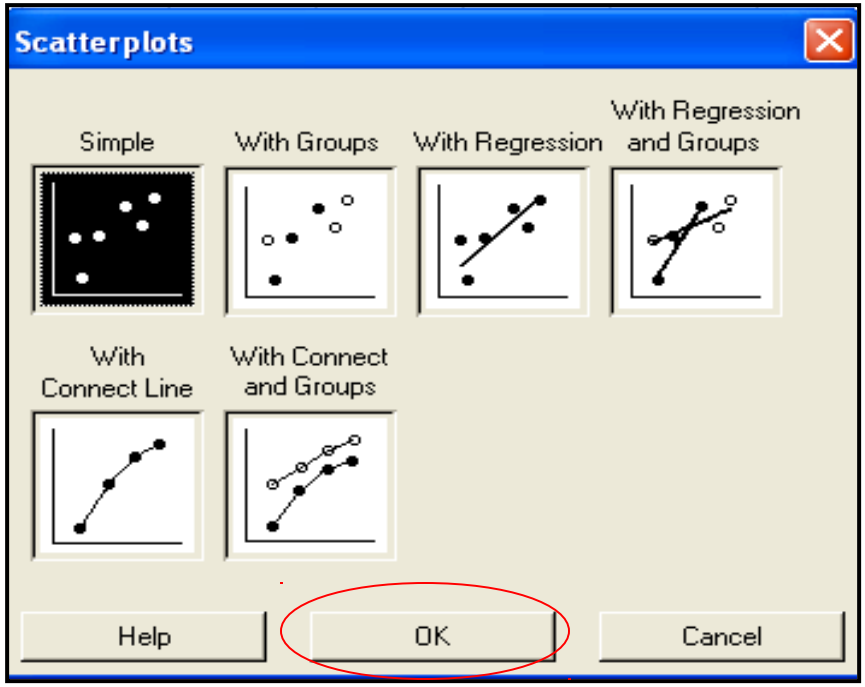
Employee	Experience (X)	Call Length (Y)
1	6	58
2	9	56
3	9	52
4	13	48
5	15	53
6	18	48
7	19	43
8	26	45
9	26	38
10	26	38
11	26	38
12	27	33
13	28	32
14	29	41
15	30	22

File Name – Minitab training

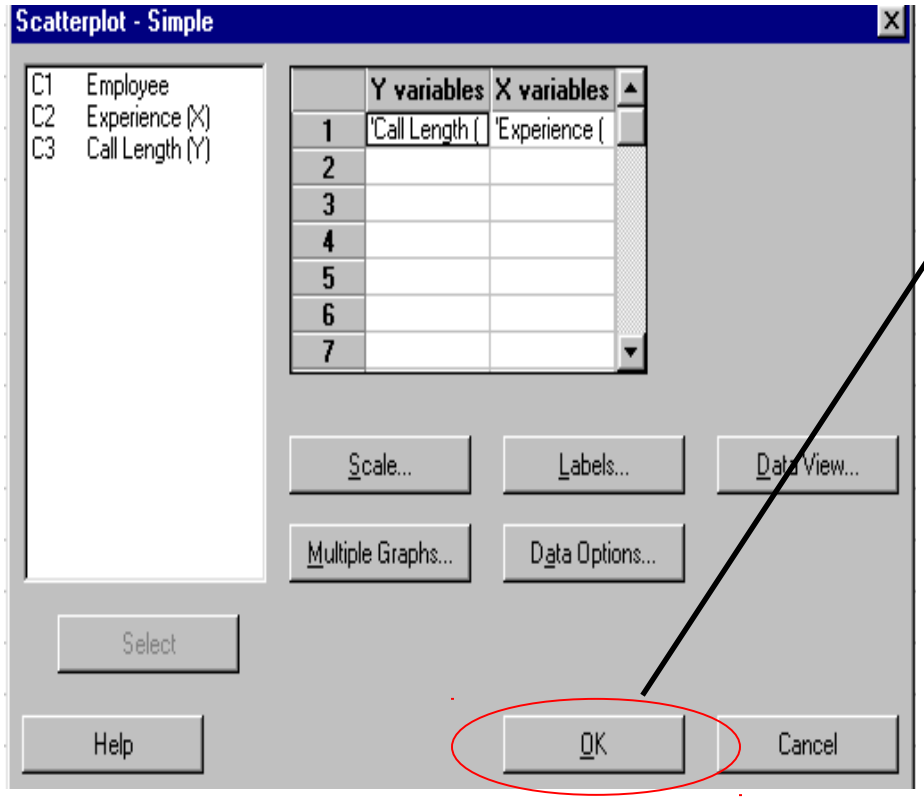
Worksheet Name – Scatter plot-S

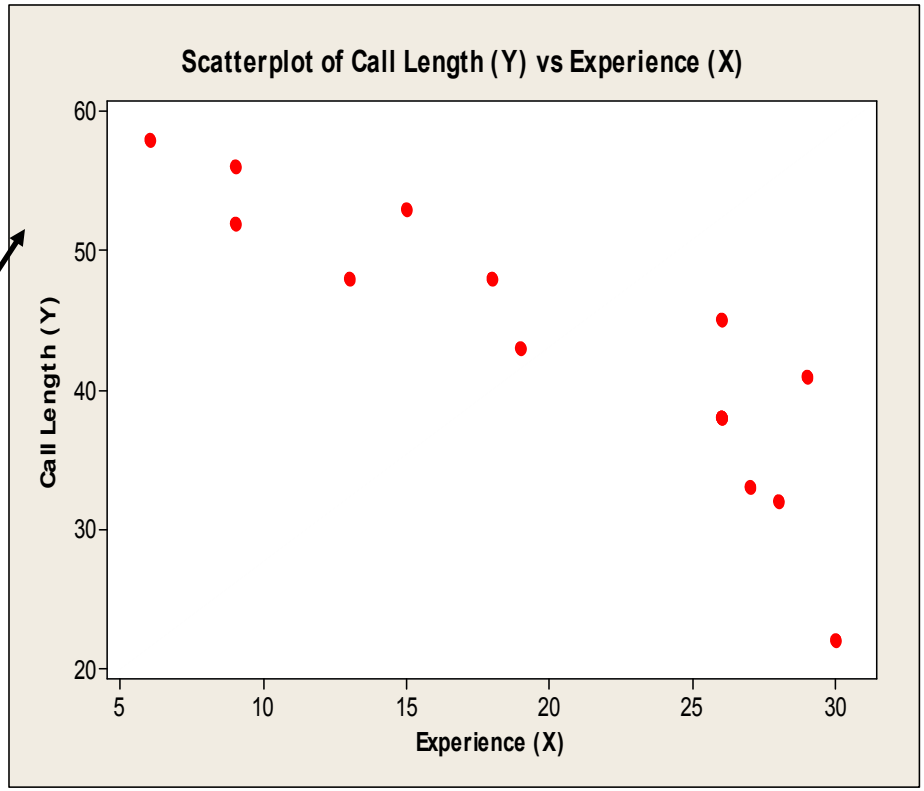
Graph→ Scatter Plot →Simple





Graph→ Scatter Plot →Simple





Interpreting the results

The higher the experience the lower the call length time tends to be.

Graph→ Scatter Plot →Simple

Class Exercise

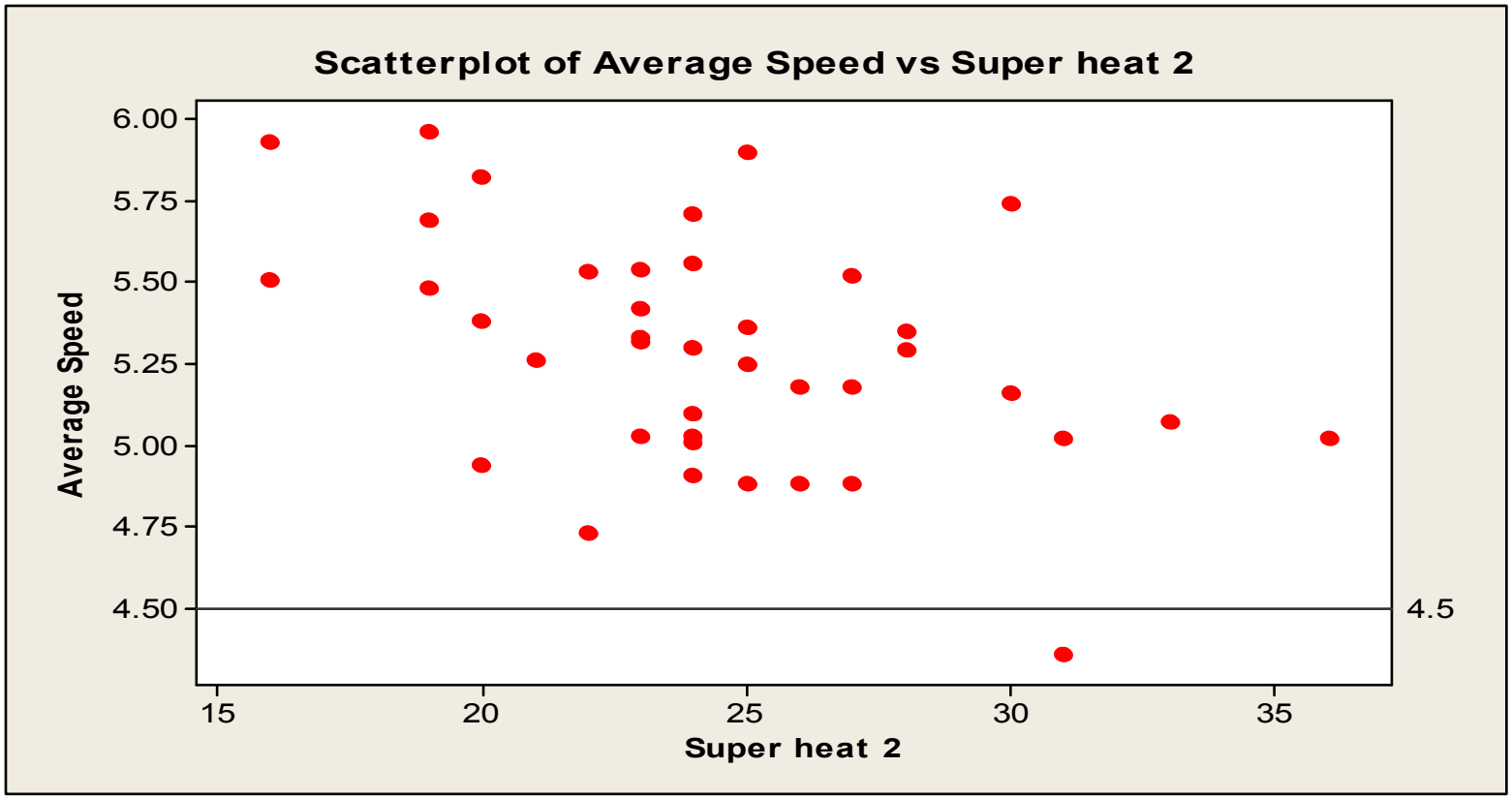
- Create a scatter plot to investigate the possible relationship between the variables Speed vs Superheat
- Include a reference line at the critical Average speed of 4.5 m/min

File Name – Minitab training

Worksheet Name – Scatter plot-C

Graph→ Scatter Plot →Simple

Class Exercise - Output



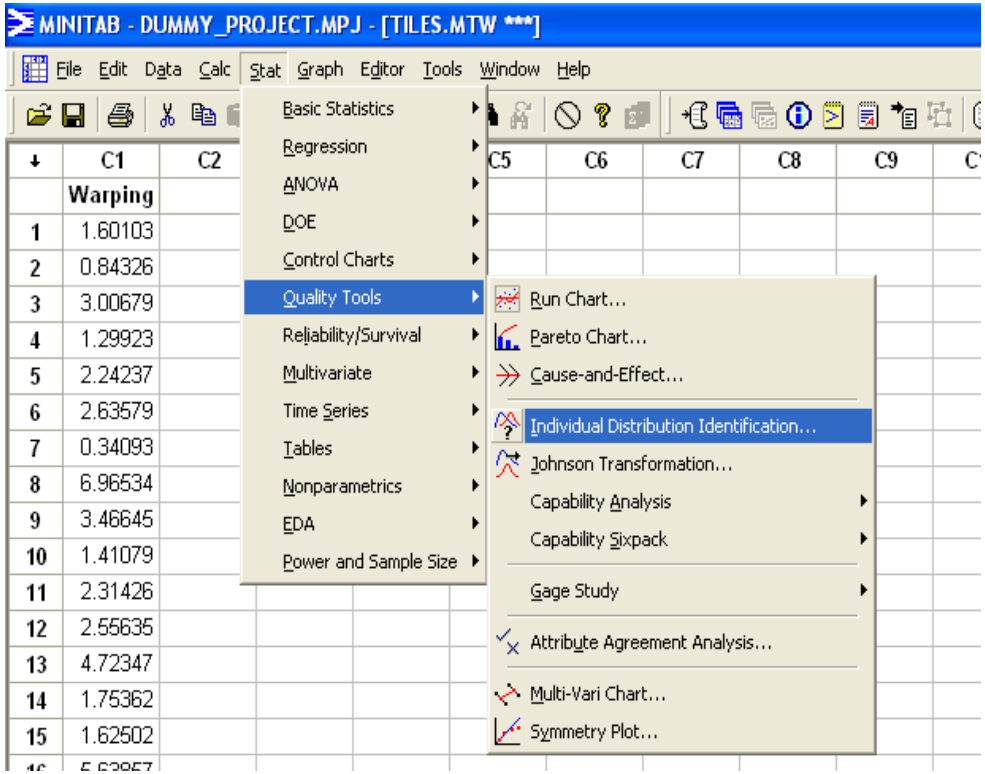
Scatter plot :Inferences & Precautions

- Even if the scatter diagram shows a relationship, do not assume that one variable caused the other. Both may be influenced by a third variable.
- When the data are plotted, the more the diagram resembles a straight line, the stronger the relationship.
- If the diagram shows no relationship, consider whether the independent (x-axis) variable has been varied widely. Sometimes a relationship is not apparent because the data don't cover a wide enough range.
- If the scatter diagram shows no relationship between the variables, consider whether the data might be stratified.
- If a line is not clear, statistics (N and Q) determine whether there is reasonable certainty that a relationship exists. If the statistics say that no relationship exists, the pattern could have occurred by random chance.

Individual distribution Identification-Step1

Example–

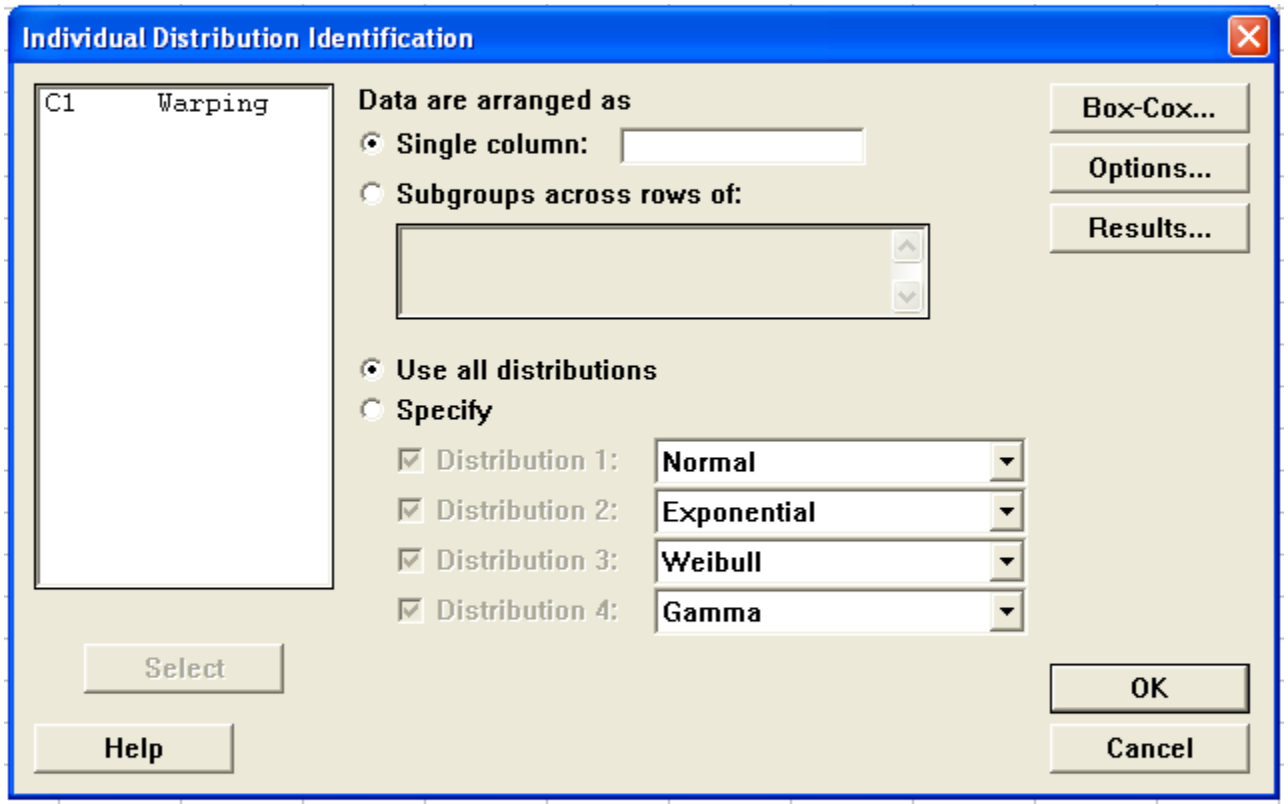
Suppose you work for a company that manufactures floor tiles, and are concerned about warping in the tiles. To ensure production quality, you measured warping in 10 tiles each working day for 10 days. The distribution of the data is unknown. Conduct the Individual Distribution Identification to identify the best fit.



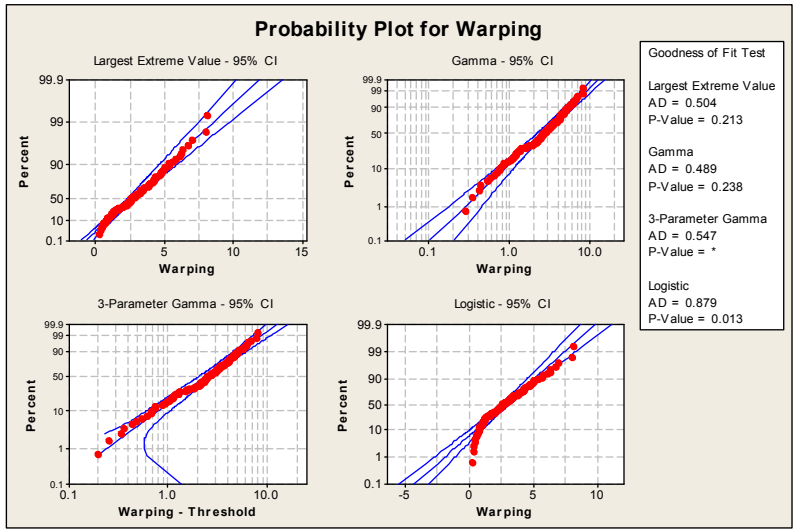
File Name – Minitab training

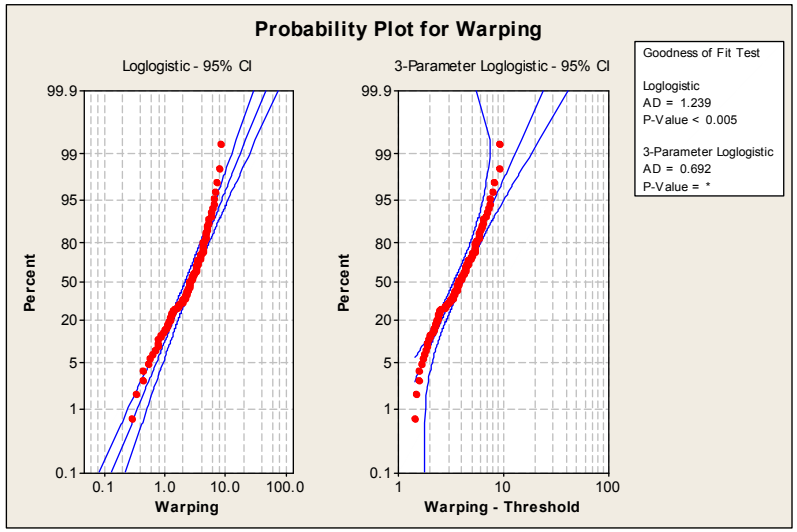
Worksheet Name – IDE_S

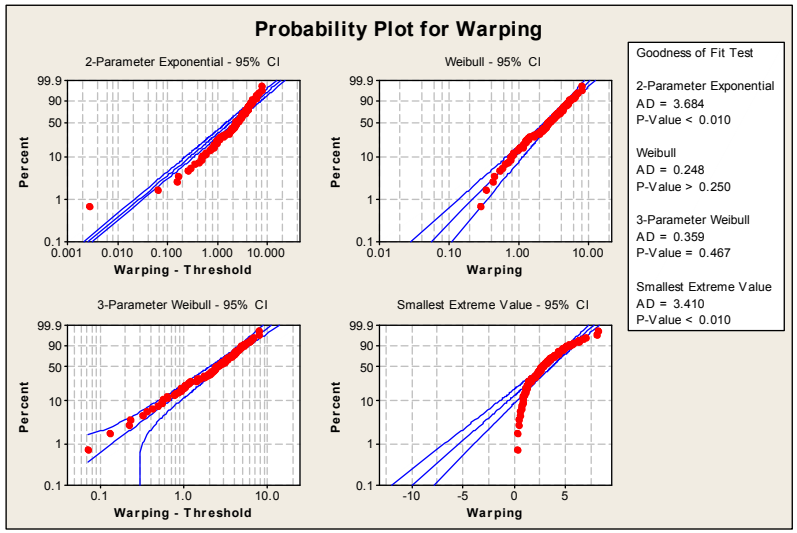
Individual distribution Identification-Step2

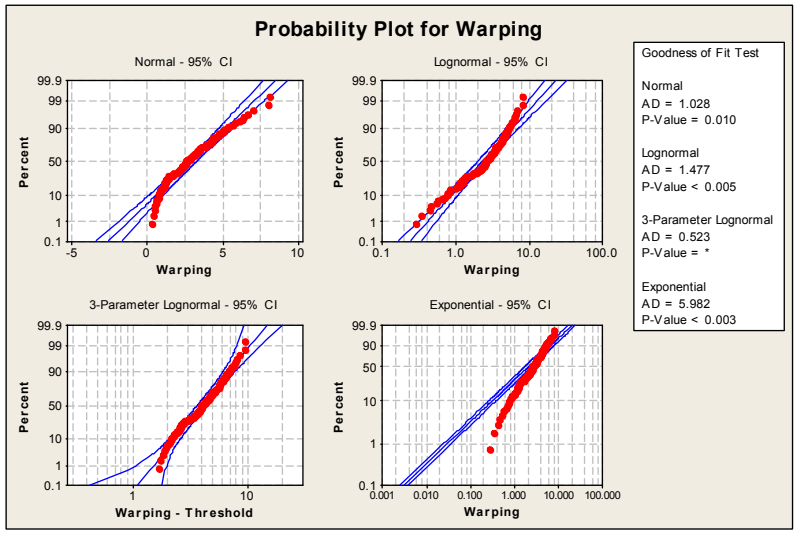


Individual distribution Identification-Graph Output









Individual distribution Identification-Output

Distribution ID Plot for Warping

Descriptive Statistics

N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
100	0	2.92307	1.78597	2.60726	0.28186	8.09064	0.707725	0.135236

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	1.028	0.010	
Lognormal	1.477	<0.005	
3-Parameter Lognormal	0.523	*	0.007
Exponential	5.982	<0.003	
2-Parameter Exponential	3.684	<0.010	0.000
Weibull	0.248	>0.250	
3-Parameter Weibull	0.359	0.467	0.225
Smallest Extreme Value	3.410	<0.010	
Largest Extreme Value	0.504	0.213	
Gamma	0.489	0.238	
3-Parameter Gamma	0.547	*	0.763
Logistic	0.879	0.013	
Loglogistic	1.239	<0.005	
3-Parameter Loglogistic	0.692	*	0.085

Non Normal Data- Box Cox Transformation

Purpose

Use a Box-Cox transformation to transform the data to follow a normal distribution and store the transformed data for further.

Example – Transform the below data using Box-Cox transformation

File Name – Minitab training

Worksheet Name – IDE_S

Non Normal Data- Box Cox Transformation

Class Exercise – Conduct a Box-Cox Transformation of the Following data and check the transformed data for Normality

File Name – Minitab training

Worksheet Name – IDE_E

Measurement System Analysis

Measurement Analysis

- An evaluation of the measurement system MUST be undertaken to ensure effective analysis of any subsequent data generated for a given process/product characteristic

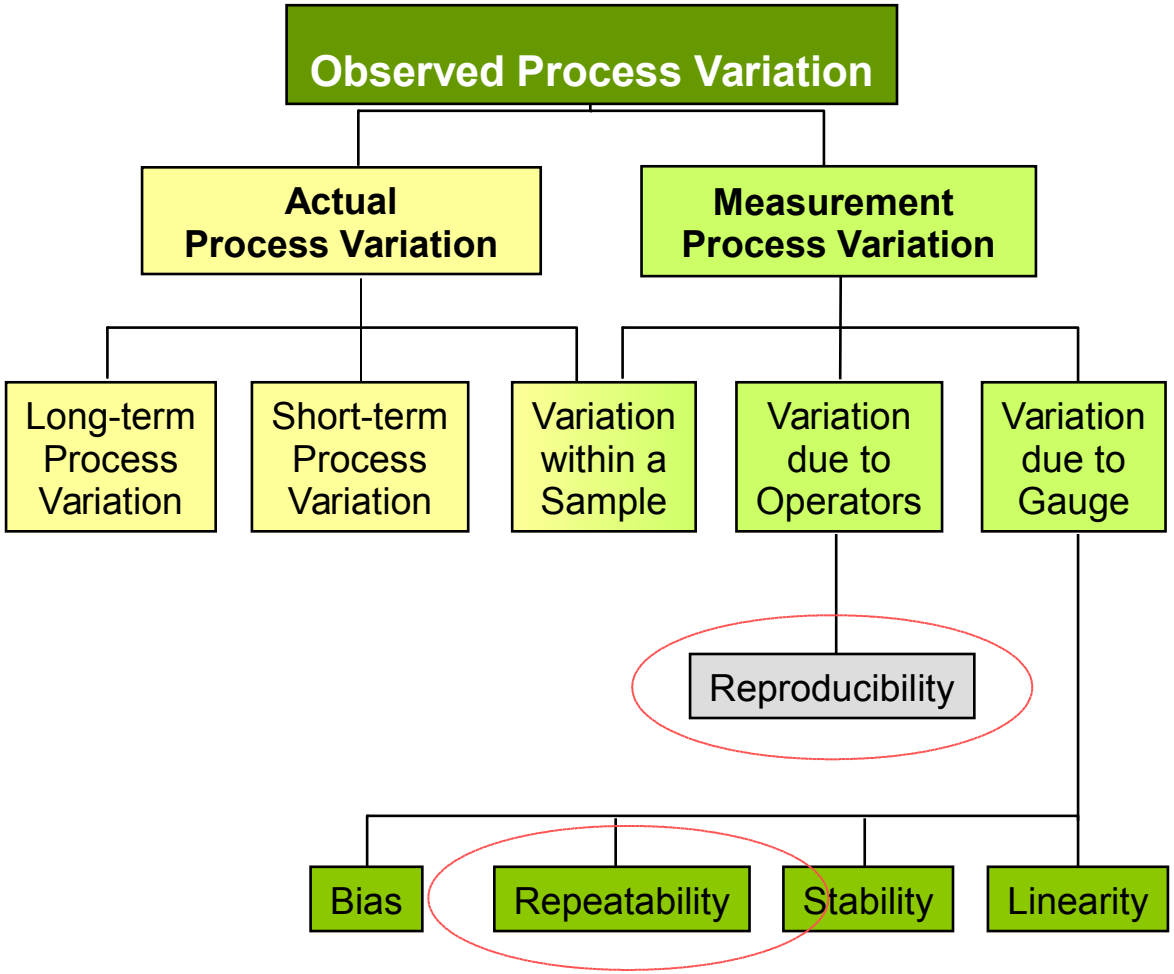
Observed Value	=	True Value	+/-	Measurement Error

- Measurement error is a statistical term meaning the net effect of all sources of measurement variability that cause an observed value to deviate from the true value

True Variability	=	Process Variability	+	Measurement Variability

- Both process and measurement variability must be evaluated and improved together
- If we work on process variability first and our measurement variability is large, we can never conclude that the improvement made was significant, or correct

MSA – Continuous Data



Methods of Performing GRR Studies

- **$\bar{X}$-R (Xbar R) Method**
  - Typically used in automobile industry
  - Extreme values affect the method
  - Short method & Long method
    - Short method does not measure operator & equipment variability separately
    - Long method measures operator & equipment variability separately, but does not measure combined effect
- **ANOVA Method**
  - Measures operator & equipment variability separately with combined effect as well that better defines causality
  - More effective when extreme values are present

Data Collection

- **Number of Trials**
  - General sampling techniques should be used to represent the population
  - Each unit is to be measured 2-3 times by each operator (Number of trials)
  - When 10 parts are used in the study even 2 trials are sufficient
  - When 5 parts are used in the study we should take three trials
  - All operators should measure all units in random order

Conducting the Study

- Select part, operator and trials
- Calibrate the gauge
- Mark all the part with serial numbers (1-10, 1-5)
- Ask the first operator to check all the parts and record the readings
- Change the number of parts & randomize (alphabets may be used)
- Ask the second operator to check all the parts second time and record the readings
- Repeat the above steps(3 through 5) with other two operators
- Compile the data in the template
- Use consistent units of measure. Avoid conversions (i.e. Inches to millimeters)

Gage RR – ANOVA Method

Example – Conduct a Gage RR study through ANOVA method for the following data

File Name – Minitab training

Worksheet Name – MSA_S

Note–

Structure your data so that each row contains the part name or number, operator (optional), and the observed measurement. Parts and operators can be text or numbers.

The ANOVA method is more accurate than the X and R method, in part, because it considers the operator by part interaction

Gage RR – ANOVA Method-Output

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Part	4	39647.9	9911.97	18.7691	0.000
Operator	2	70.5	35.27	0.0668	0.936
Part * Operator	8	4224.8	528.10	0.3478	0.939
Repeatability	30	45548.0	1518.27		
Total	44	89491.2			

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Part	4	39647.9	9911.97	7.56748	0.000
Operator	2	70.5	35.27	0.02693	0.973
Repeatability	38	49772.8	1309.81		
Total	44	89491.2			

Gage RR – ANOVA Method-Output

Gage R&R Study - ANOVA Method

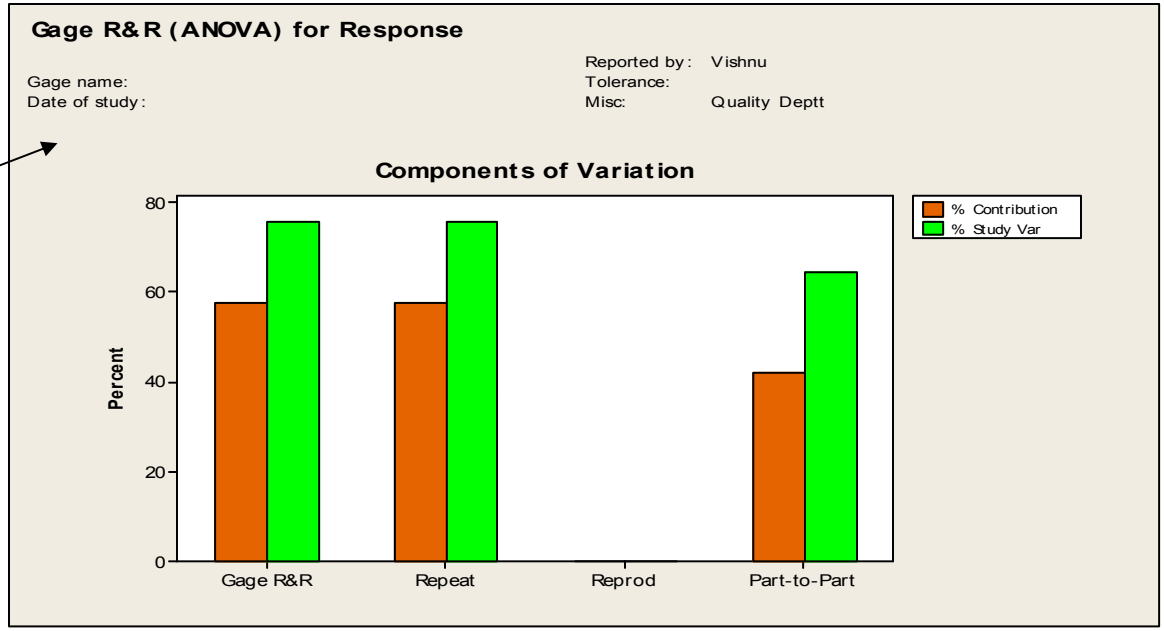
Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	1309.81	57.81
Repeatability	1309.81	57.81
Reproducibility	0.00	0.00
Operator	0.00	0.00
Part-To-Part	955.80	42.19
Total Variation	2265.61	100.00

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)
Total Gage R&R	36.1913	217.148	76.03
Repeatability	36.1913	217.148	76.03
Reproducibility	0.0000	0.000	0.00
Operator	0.0000	0.000	0.00
Part-To-Part	30.9159	185.496	64.95
Total Variation	47.5984	285.590	100.00

Number of Distinct Categories = 1

Gage RR – ANOVA Method-Graph Output



This is a graphical representation of the Gage R&R section of the Session Window output.

The sources of variation that are represented in the graph are:

·Total Gage R&R, is the variation due to the measuring system including multiple operators using the same gage.

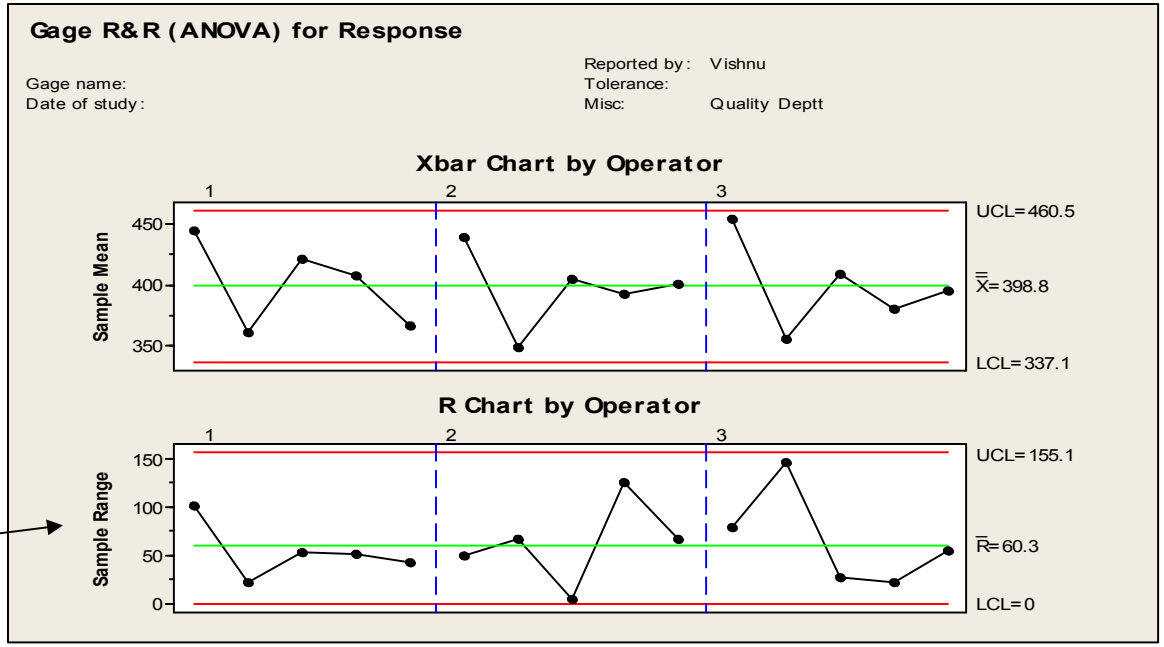
·Repeatability, is the variability in measurements obtained when the same part is measured multiple times by the same operator.

·Reproducibility, is the variability in measurements obtained when the same part is measured by different operators.

·Part-to-Part, which is the variability in measurements across different parts.

In a good measurement system, the largest component of variation is Part-to-Part variation. If instead you have large amounts of variation attributed to Gage R&R (Repeatability and/or Reproducibility), corrective action is needed.

Gage RR – ANOVA Method-Graph Output



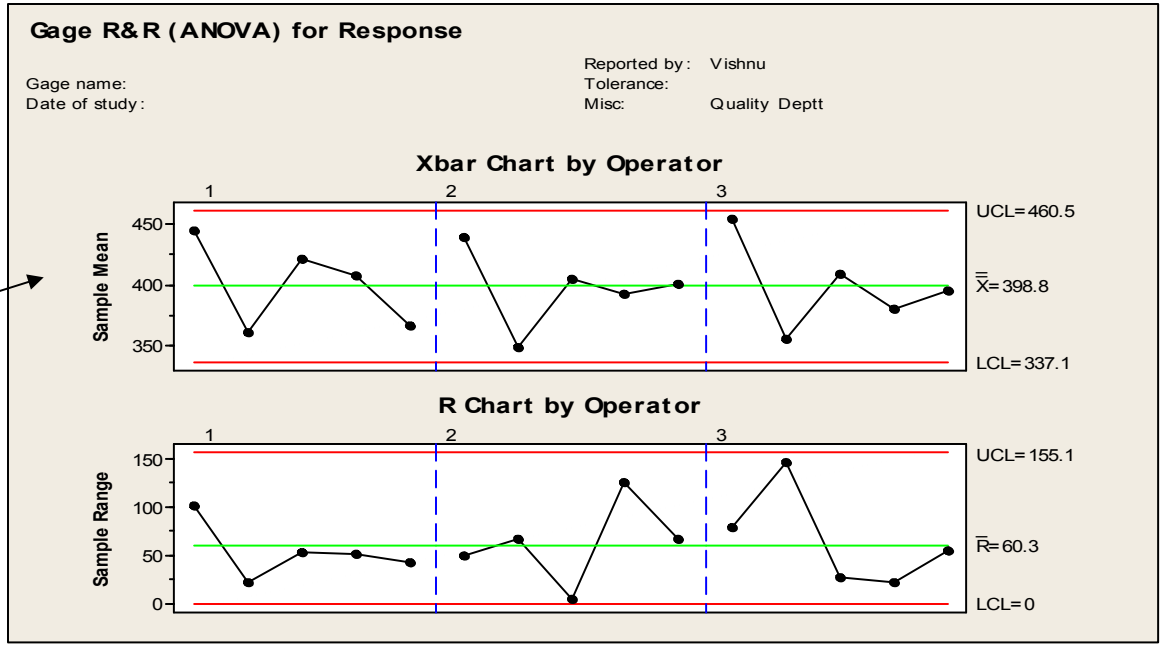
Plotted points represent, for each operator, the difference between the largest and smallest measurements on each part. Because the points are arranged by operator, you can see how consistent each operator is.

·Green center line, is the grand average for the process (average of all the subgroup ranges).

·Red control limits, which represent the amount of variation expected for the subgroup ranges. These limits are calculated using the variation within subgroups.

If any of the points on the graph go above the upper control limit (UCL), then that operator is having problems consistently measuring parts. The UCL value takes in account the number of measurements by an operator on a part and the variability between parts. If the operators are measuring consistently, then these ranges should be small relative to the data and the points should stay in control. All of the parts data are "in control" indicating that all operators are measuring consistently.

Gage RR – ANOVA Method-Graph Output



Plotted points, which represent, for each operator, the average measurement on each part.

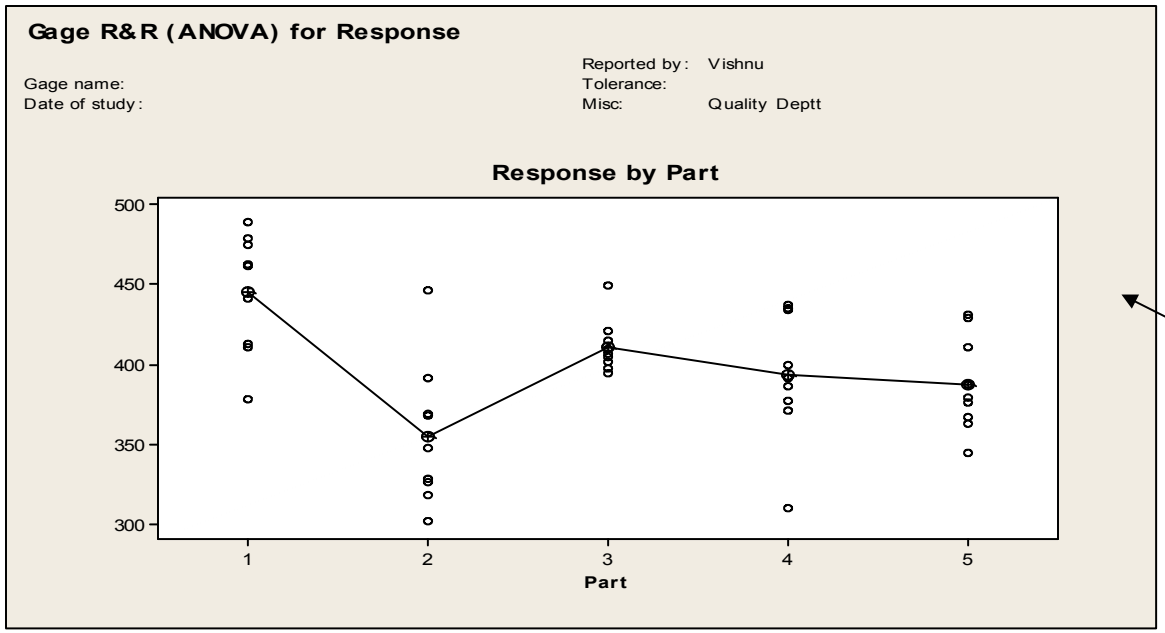
Green center line, which is the overall average for all part measurements by all operators.

Red control limits (UCL and LCL), which are based on how much variability there is between parts and the number of measurements in each average.

Because the parts chosen for a Gage R&R study should represent the entire range of possible parts, this graph should ideally show lack-of-control. Lack-of-control exists when many points are above the upper control limit and/or below the lower control limit.

For the parts data, there are many points withinthe control limits, which indicates that the measurement system is inadequate.

Gage RR – ANOVA Method-Graph Output



Shows all of the measurements taken in the study, arranged by part. The measurements are represented by dots; the means are represented by the circle-cross symbol. The black line connects the average measurements for each part.

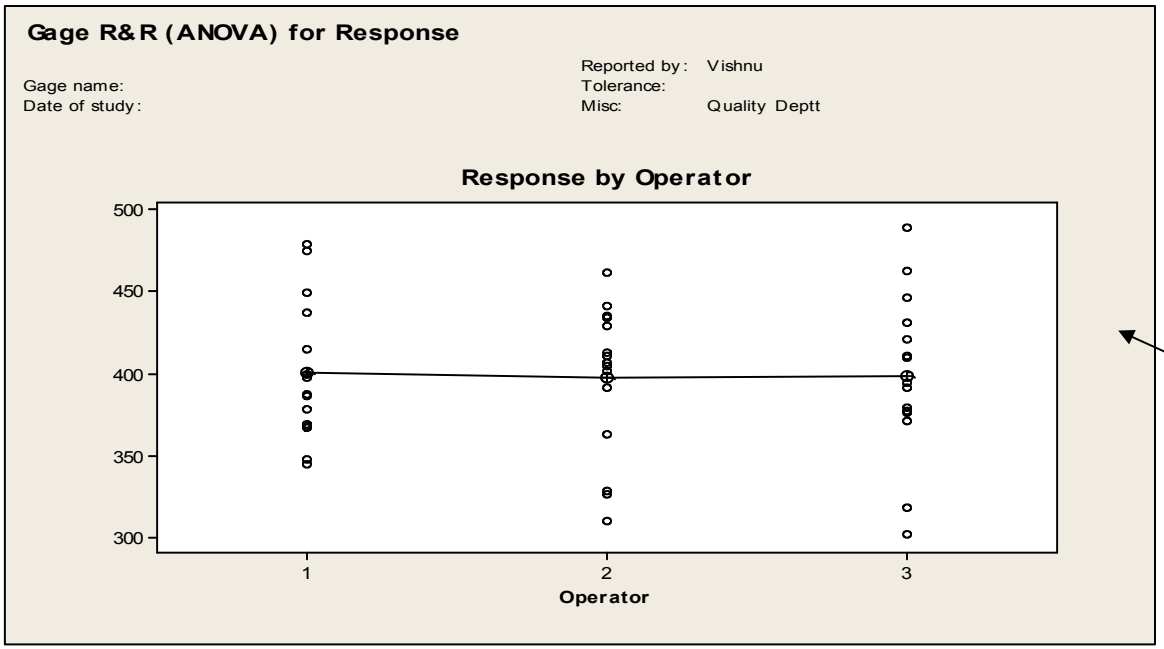
Ideally,

·the multiple measurements for each individual part will vary as little as possible (the dots for one part will be close together)

·the averages will vary enough that differences between parts are clear

For the parts data, the measurements for all part vary quite a bit. This variation may be due to the system's (operator and/or gage) inability to consistently measure that part. The variation in averages is not significant. This could be because of the parts chosen for the study are not representing the entire range of possible parts.

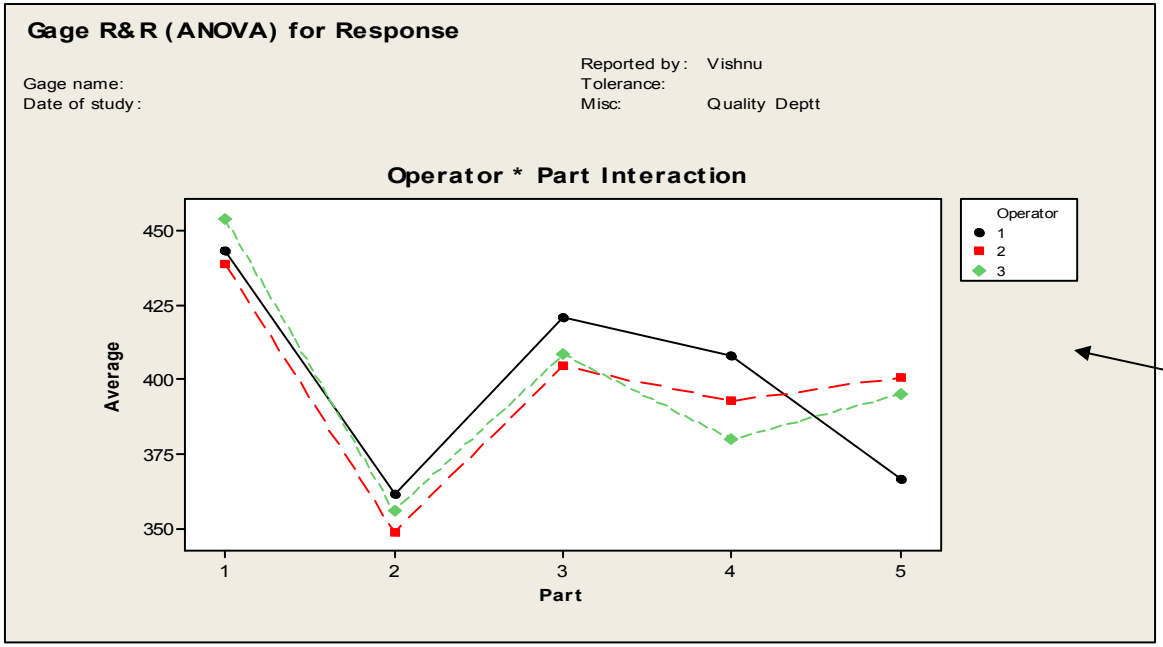
Gage RR – ANOVA Method-Graph Output



Shows all of the measurements taken in the study, arranged by operator. The measurements are represented by dots; the means by the circle-cross symbol. The line connects the average measurements for each operator.
Ideally,
·the measurements for each operator will vary an equal amount
·the part averages will vary as little as possible & the connecting line will be almost horizontal.

For the parts data, this seems to be the case. There will always be some variation, but it appears the operators are measuring consistently.

Gage RR – ANOVA Method-Graph Output



Shows the average measurements taken by each operator on each part in the study, arranged by part. Each line connects the averages for a single operator.

Ideally,

·the lines will follow the same pattern

·the part averages will vary enough that differences between parts are clear

For the parts data, the lines follow each other fairly well. The operators may have a problem consistently measuring part s

Interpretation of GRR Results

- GRR as a % of Contribution to Variation.
  - If GRR as % of contribution is less than 1% of total variation - excellent
  - If GRR as % of contribution is > 1% & < 10% of the total variation - acceptable
  - If GRR as % of contribution is > 10% of the total variation - unacceptable

- Number of Distinct Categories
  - If number of distinct categories is >= 4 - acceptable
  - If number of distinct categories is < 4 - unacceptable
- If measurement system is found to be unacceptable as per the above criteria, do not proceed to the next step
- If tolerance was known, GRR as a % of Tolerance should be used for decision as explained in the next slide

Interpretation of GRR Results

- GRR as a % of Tolerance
  - If GRR as % of tolerance is less than 10% - excellent measurement system
  - If GRR as % of tolerance is between 10% to 30% - acceptable measurement system

However, discretion may be needed depending upon application of the process / equipment

  - If GRR as % of tolerance is above 30% - unacceptable measurement system

You should not proceed to next DMAIC step. Simplify process / explore root cause

Gage RR – ANOVA Method-Class Exercise

- Conduct Gage RR study for the following data and interpret the output

File Name – Minitab training

Worksheet Name – MSA_E

Process capability

- Let's understand the concept of short & long term variations. Below is the data given on the % Metallic Iron content in DRI spread over 5 days. Each day's data can form a sub-group.

% Fe (M) in DRI				
Day 1	Day 2	Day 3	Day 4	Day 5
85.00	84.50	86.30	86.00	86.50
84.40	85.40	86.10	85.70	85.40
84.70	84.50	86.90	86.70	85.70
84.60	84.60	86.60	86.60	86.40
84.60	84.00	85.20	86.10	86.10
85.20	84.30	85.60	86.50	85.10
85.10	83.00	86.10	85.00	85.00
84.20	84.60	85.10	85.60	85.10
84.00	84.50	85.30	86.20	85.00
84.50	84.60	85.80	86.00	86.00
85.00	84.00	85.50	85.70	86.20

Calculating Process Variations

$$\text{StDev short term } S_{ST} = \sqrt{\frac{\text{Total variation within sub-groups}}{(\text{Total sample size} - \text{number of subgroups})}}$$

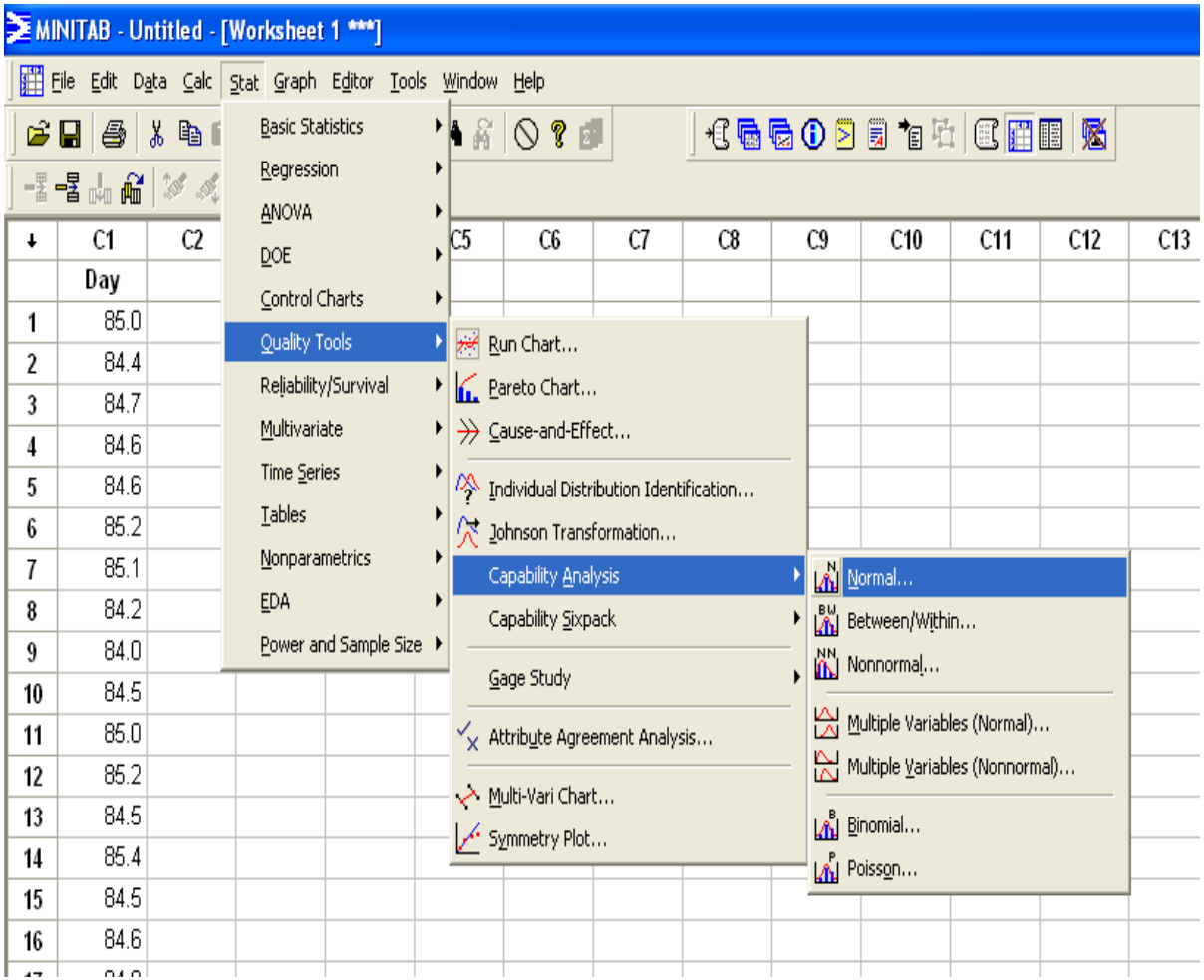
$$\text{StDev long term } S_{LT} = \sqrt{\frac{\text{Overall variation}}{(\text{Total sample size} - 1)}}$$

Process Capability

Stat > Quality tools > capability Analysis> Normal

File Name – Minitab training

Worksheet Name-Capability-S

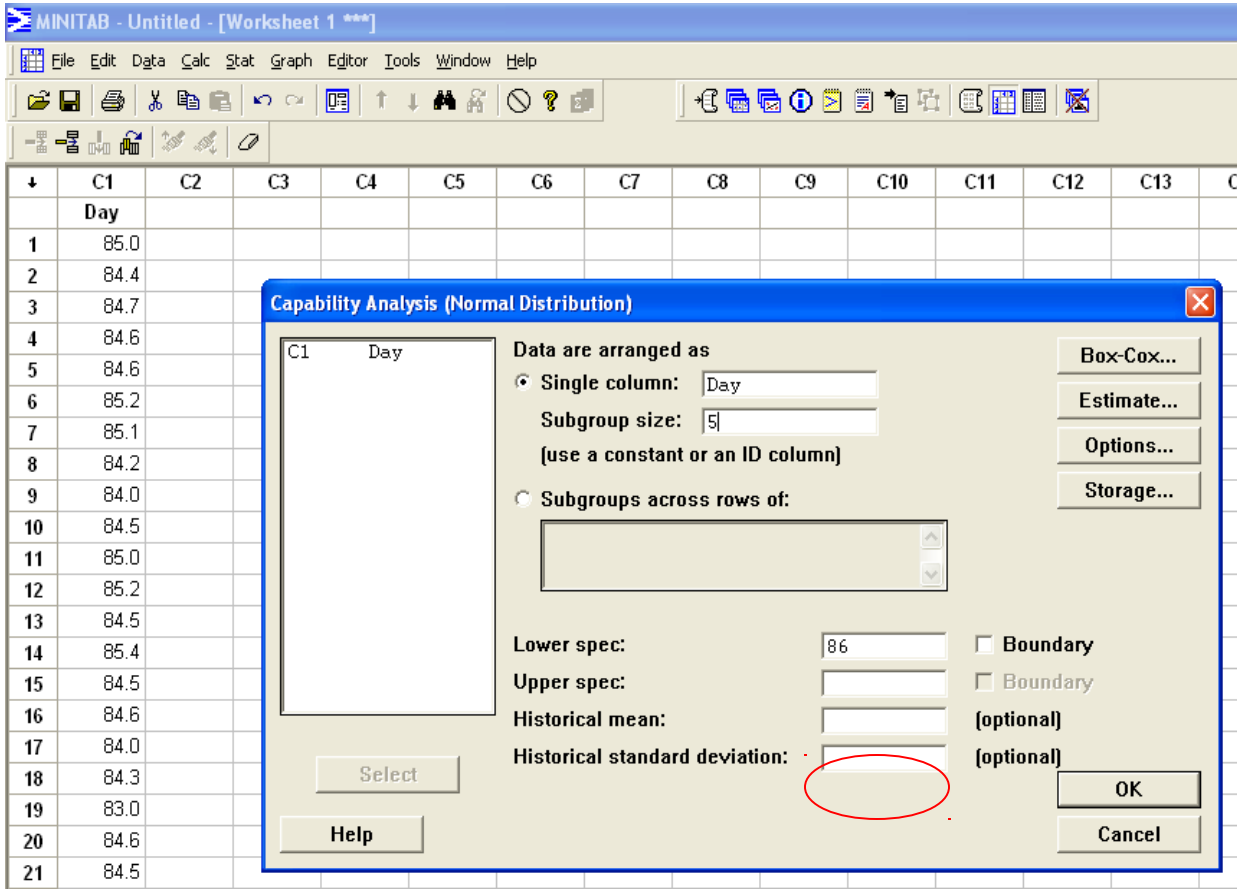


Process Capability

Stat > Quality tools > capability Analysis> Normal>single column "Day">Subgroup size "5">Lower spec "86">OK

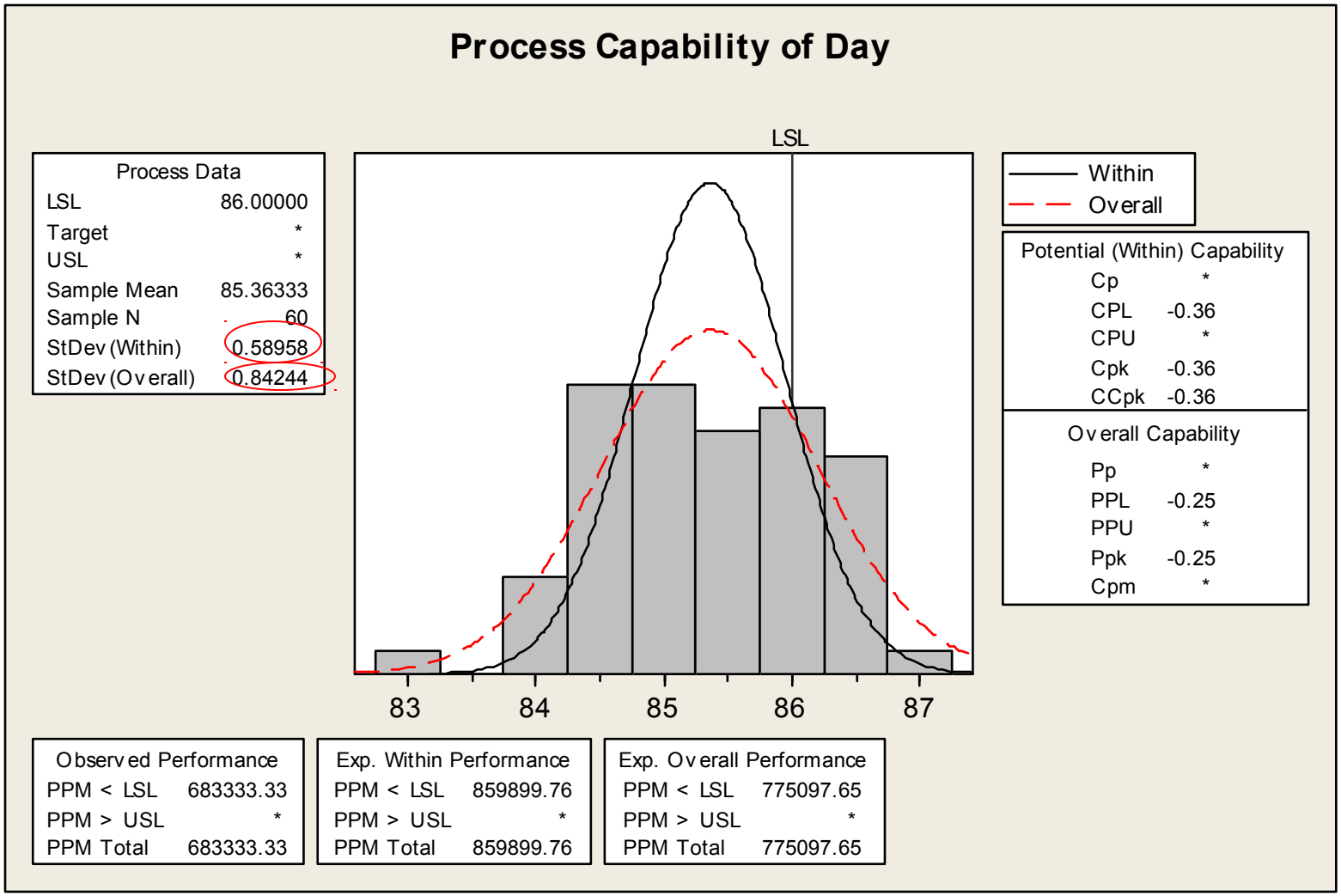
File Name – Minitab training

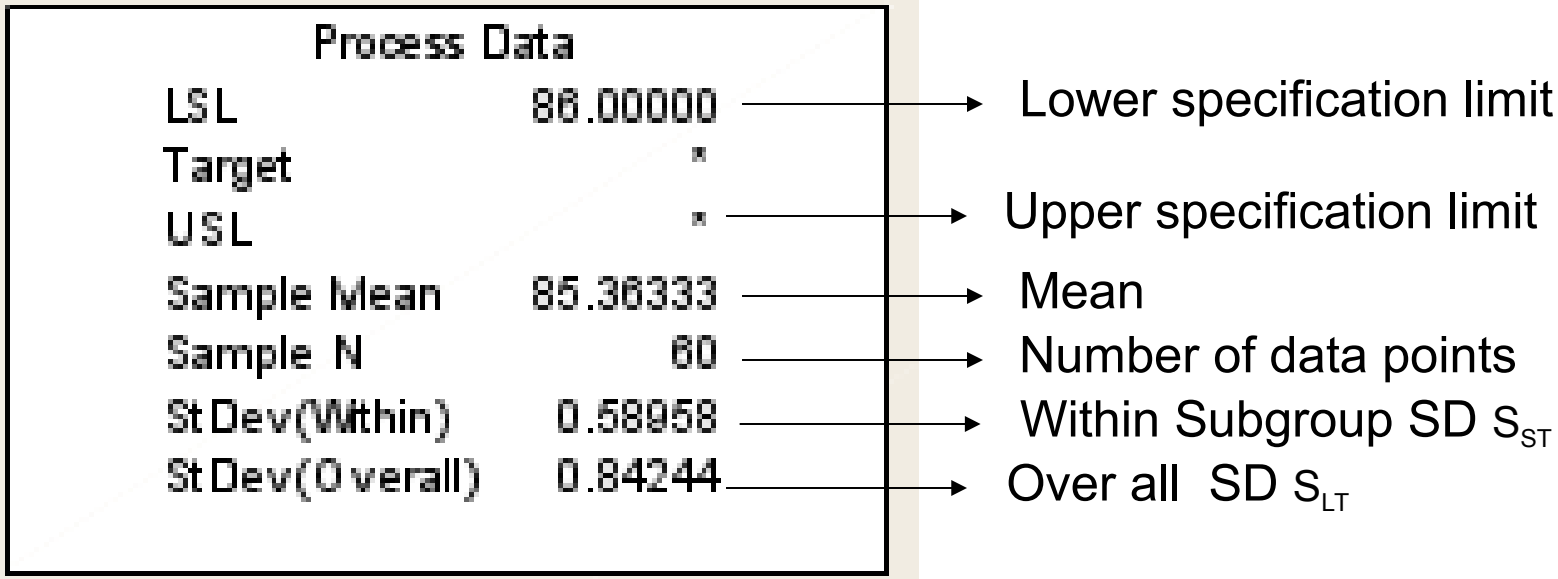
Worksheet Name-Capability-C



Process Capability

Minitab output




Process Data
LSL 86.00000
Target *
USL *
Sample Mean 85.36333
Sample N 60
StDev(Within) 0.58958
StDev(Overall) 0.84244
Lower specification limit
Upper specification limit
Mean
Number of data points
Within Subgroup SD S_{ST}
Over all SD S_{LT}

Process Capability

$$\text{Process Capability } C_P = \frac{USL - LSL}{6\,S_{ST}}$$

$$C_{PL} = \frac{\bar{Y} - LSL}{3\,S_{ST}}$$

$$C_{PK} = \text{Minimum}\,(C_{PU}, C_{PL})$$

$$C_{PU} = \frac{USL - \bar{Y}}{3\,S_{ST}}$$

Potential (Within) Capability	
Cp	*
CPL	-0.36
CPU	*
Cpk	-0.36
CCpk	-0.36
Overall Capability	
Pp	*
PPL	-0.25
PPU	*
Ppk	-0.25
Cpm	*

$$P_{PU} = \frac{USL - \bar{Y}}{3\,S_{LT}}$$

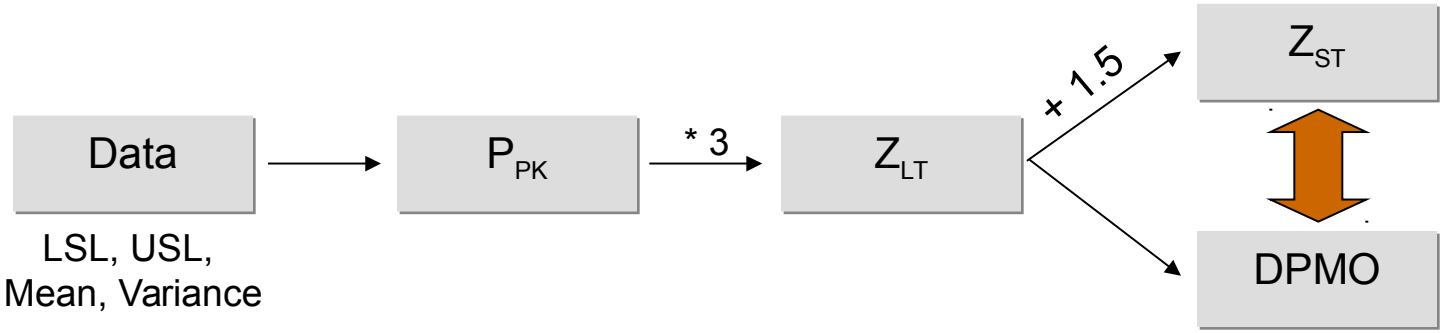
$$P_{PL} = \frac{\bar{Y} - LSL}{3\,S_{LT}}$$

$$P_{PK} = \text{Minimum}\,(P_{PU}, P_{PL})$$

The number of data points below LSL is 41 out of total data points 60 which gives the observed performance 683333.33PPM	From the area under the normality curve	From the area under the normality curve
(PPM <LSL) =41/60*1000000 =683333.33	**C_{PK} = -0.36** Z_{ST} = 3* **C_{PK} = -1.08**	**P_{PK} = -0.25** Z_{LT} = 3* **P_{PK} =- 0.75** $Z_{ST} = Z_{LT} + 1.5$ **= 0.75**
	Area under the curve corresponding to Z ST = 1.08 is 8599288	Area under the curve corresponding to Z ST = 0.75 is 773372

Key Concepts

- Capability is an internal measurement of the process behavior measured on a short term view
- Performance is an external view of the process behavior measured on a long term view
- Due to limitations of multiple shift factors & C_P, process sigma multiple calculations for continuous data start from P_{PK}



Z Calculation Worksheet – Discrete Data

Z Calculation for Discrete Data

Type of defect	Defects	Units	OFE
Total:	0	n/a	n/a

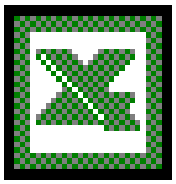
TOP	DPU	DPO	DPMO	Yield	Z_{ST}	Z_{LT}
0						
0						
0						
0						
0						
0						
0	n/a	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

DPMO to Z Conversion

Fill DPMO here →		Zst	
		Yield	

Z to DPMO Conversion

Fill Zst here →		DPMO	
		Yield	



Yield = (1 – DPO)

Process Capability

File Name – Minitab training

Worksheet Name-Capability-C

Find the process capability of the thickness of the HRC.

Thank you