Define Phase Exercises

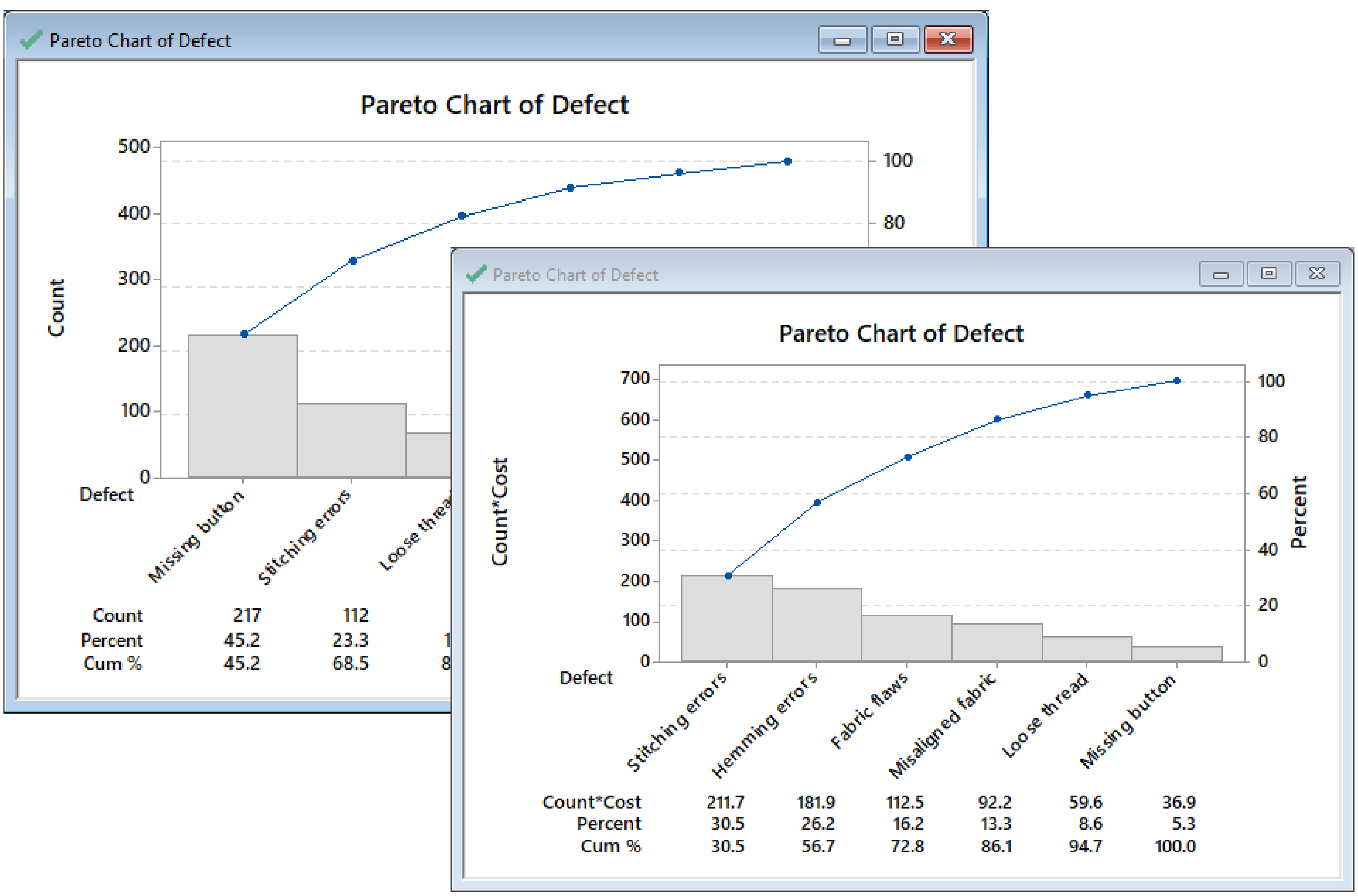
Pareto Exercise - Clothing Defects

Data File: Pareto.MTW

An inspector for a clothing manufacturer investigates the sources of clothing defects to prioritize improvement projects. The inspector tracks the number and type of defects in the process. Run a Pareto chart on the count of defects to determine the priority based on defects and then run another Pareto chart on the “Count*Cost” of defects. Based on your findings, what defect type would you recommend be the highest priority?

Solution Steps Screencast: https://youtu.be/uIYvHqy9XBQ

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab’s “Data Set Library” at https://support.minitab.com/en-us/datasets/

Pareto Chart Result Interpretation

The first Pareto shows us that the largest defect category is Missing Buttons, followed by Stitching Errors. These two defects represent 68.5% of all defects. If the manufacturer wanted to reduce total defects to improve customer satisfaction, addressing button and stitching errors would be the focus.

However, another piece of information added to the decision is the cost of the defects. When we consider cost combined with the count of defects, the prioritization changes. The second Pareto has Stitching Errors at the top followed by Hemming Errors. Missing Buttons fell to the bottom. If the manufacturer wanted to focus internally to improve the business, addressing stitching and hemming errors would be the optimal choices.

For the most balanced and impactful approach, with an eye toward both the business and the customer, addressing stitching errors would be the single most impactful project to work on first.

Rolled Throughput Yield

Data File: N/A

As the manager of the label production process of a sign company, you want to understand the full view of your process and determine the probability of producing defect-free labels. The process is made up of three process steps, and you are evaluating the defect rate of the production of 1,000 labels.

- **Process Step 1: Printing**. 1000 labels go through the printing process, and it is determined that 950 of the printed labels are acceptable.
- **Process Step 2: Lamination**. After printing, 950 good labels reach the laminating process, and 800 of the laminated labels are accepted by the quality reviewer.
- **Process Step 3: Trim**. Now, after printing and laminating, there are 800 labels that will go through the trimming process, and 700 of these trimmed labels are determined to be acceptable.

Assignment:

1. Determine the yield of each process step.
2. Calculate the rolled throughput yield of the entire process.

Solution:

Printing
DPU = 50/1000 = 0.05
Yield = 2.718^-0.05 = 0.951
Yied = 95.1 %

→

Lamination
DPU = 150/950 = 0.1583
Yield = 2.718^-0.158 = 0.854
Yied = 85.4 %

→

Trim
DPU = 100/800 = 0.125
Yield = 2.718^-0.125 = 0.883
Yied = 88.3 %

Rolled Throughput Yield

0.951 x 0.854 x 0.883 = 0.712

71.2%

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Measure Phase Exercises

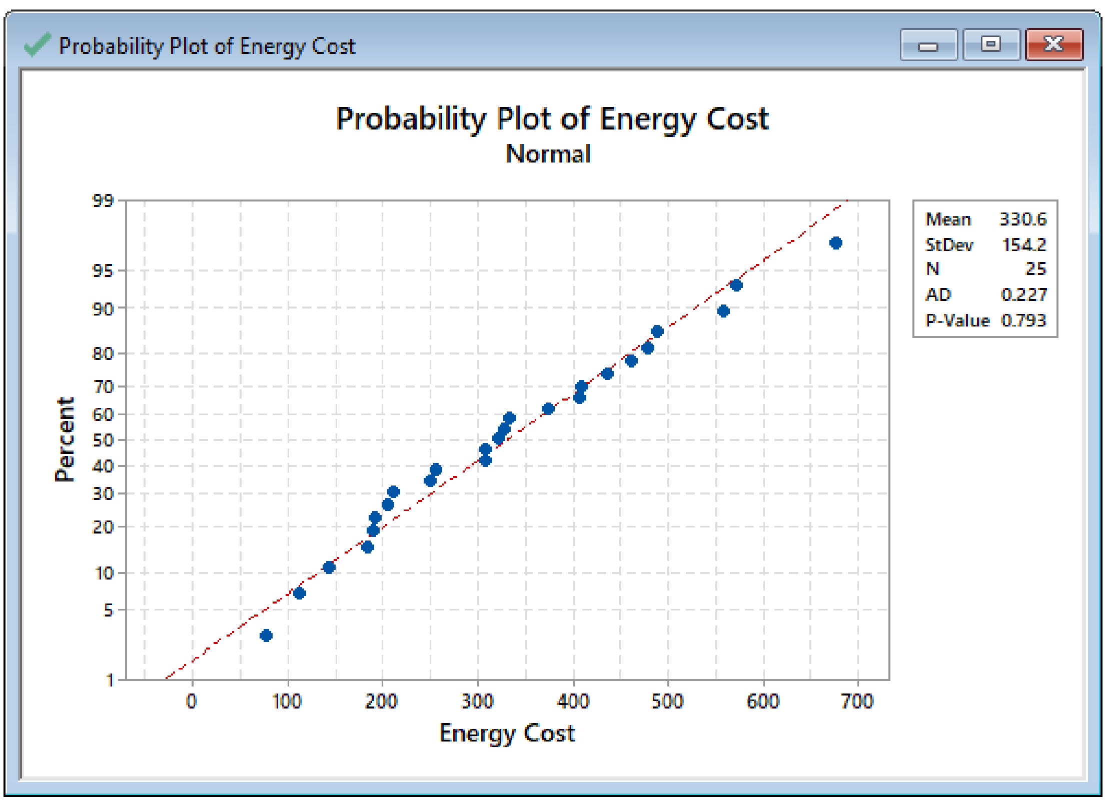
Normality Test Exercise – Family Energy Costs

Data File: [Normality.MTW]

An economist wants to determine whether the monthly energy cost for families has changed from the previous year, when the mean cost per month was $200. The economist randomly samples 25 families and records their energy costs for the current year. Before conducting any statistical comparison tests, the economist should first determine if the data are normal. Use the "Normality.MTW" data file to determine if the data are normally distributed.

Solution Steps Screencast: https://youtu.be/o8WvDOkd2e0

Solution Output Screenshot:



Normality Test Interpretation

The null hypothesis (H_0) for a normality test is that the data are normal. An Anderson-Darling test for normality yields a p-value of 0.793, so we fail to reject the null hypothesis and conclude that the data are normal.

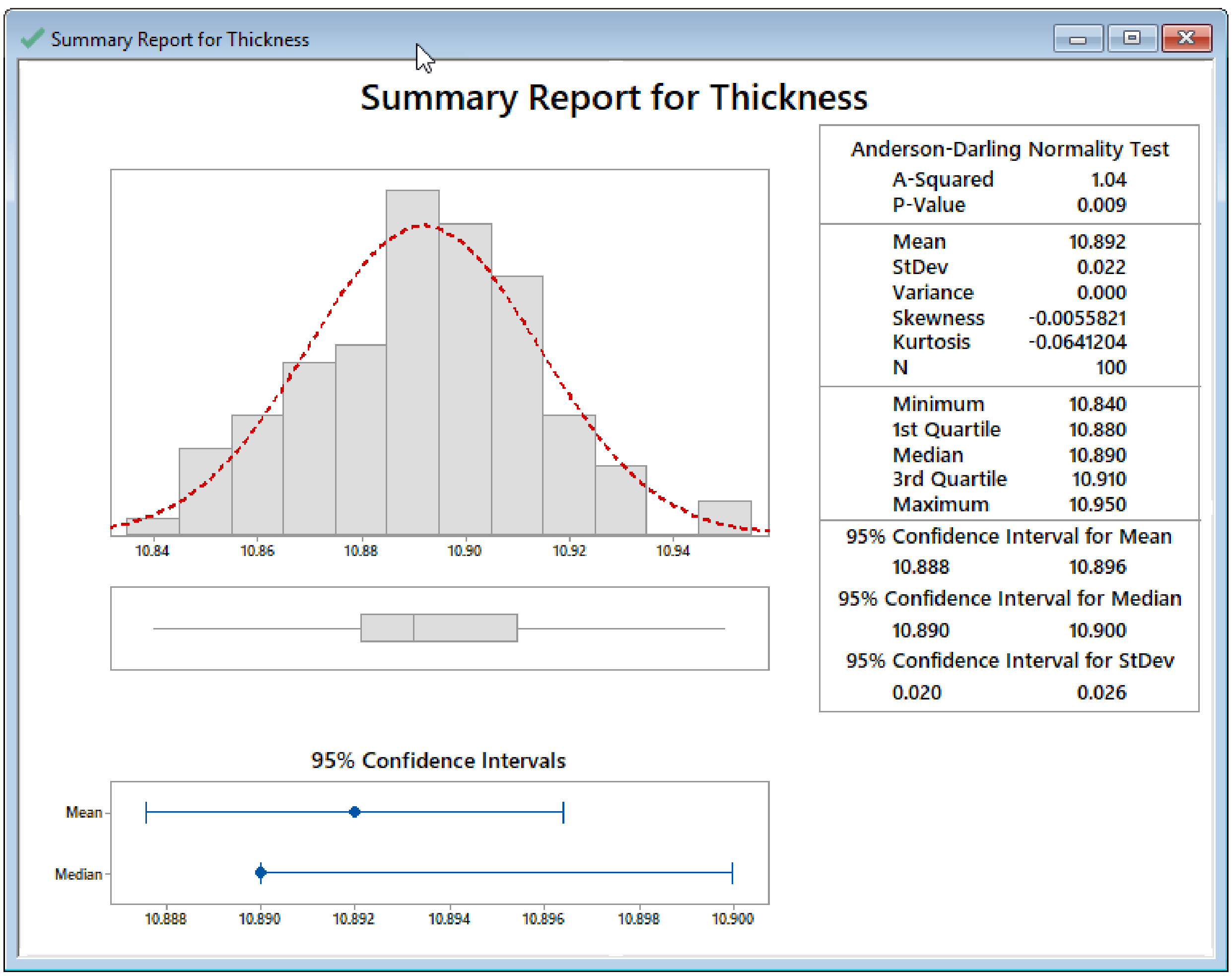
Graphical Summary Exercise – Support Beam Thickness

Data File: Graphical_Summary.MTW

A production engineer wants to investigate the capability of the process that manufactures support beams. The engineer thinks that the process capability might differ between the morning and evening shifts. The engineer measures the thickness of five samples out of 10 boxes from each shift. The thickness must be between 10.44 mm and 10.96 mm to meet the customer requirements. Before conducting capability analysis or hypothesis testing, let's look at the data using a graphical summary.

Solution Steps Screencast: https://youtu.be/pYc03H_bMaQ

Solution Output Screenshot:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

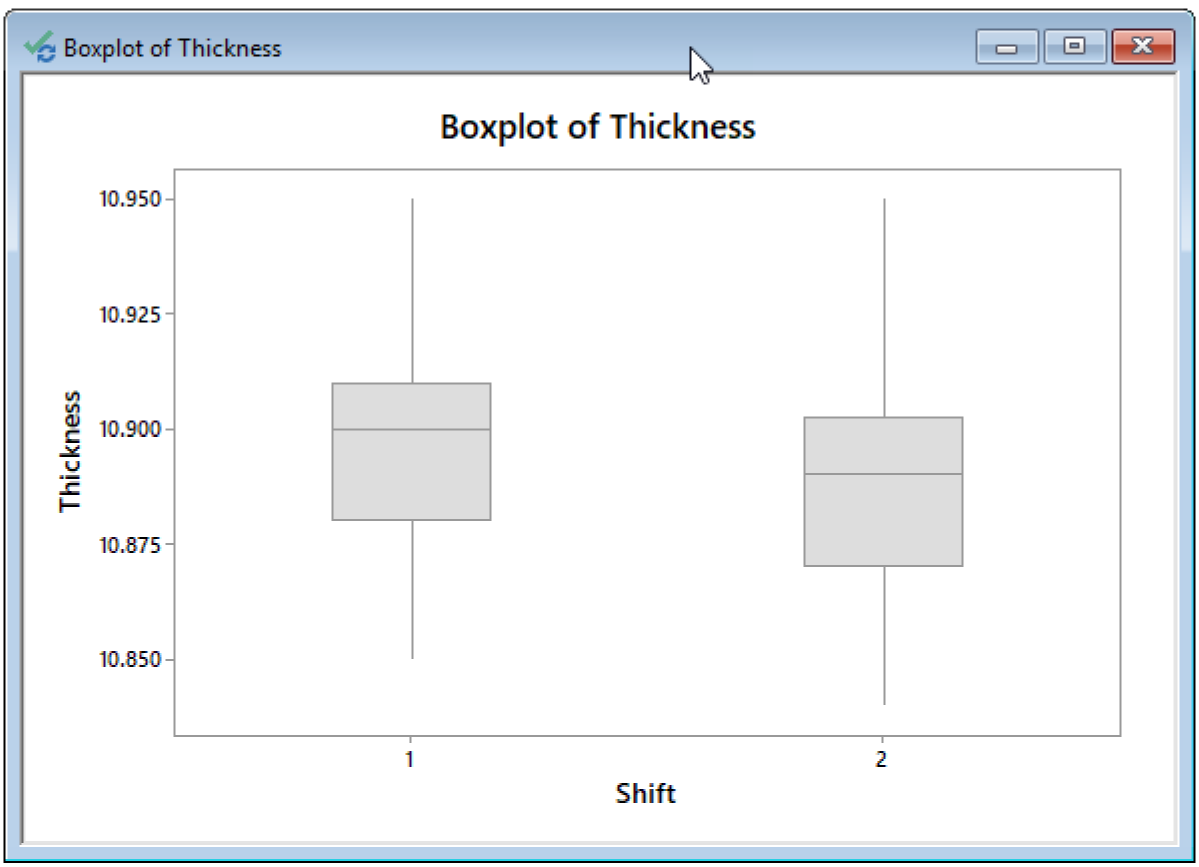
Box Plot Exercise – Support Beam Thickness by Shift

Data File: *Box_Plot.MTW*

A production engineer wants to investigate the capability of a process that manufactures support beams. The engineer thinks that the process capability might differ between the morning and evening shifts. The engineer measures the thickness of five samples out of 10 boxes from each shift. The thickness must be between 10.44 mm and 10.96 mm to meet the customer requirements. Before conducting capability analysis or hypothesis testing, perform a simple box plot of the thickness between the two shifts.

Solution Steps Screencast: https://youtu.be/WjbZ8pzyysM

Solution Output Screenshot:



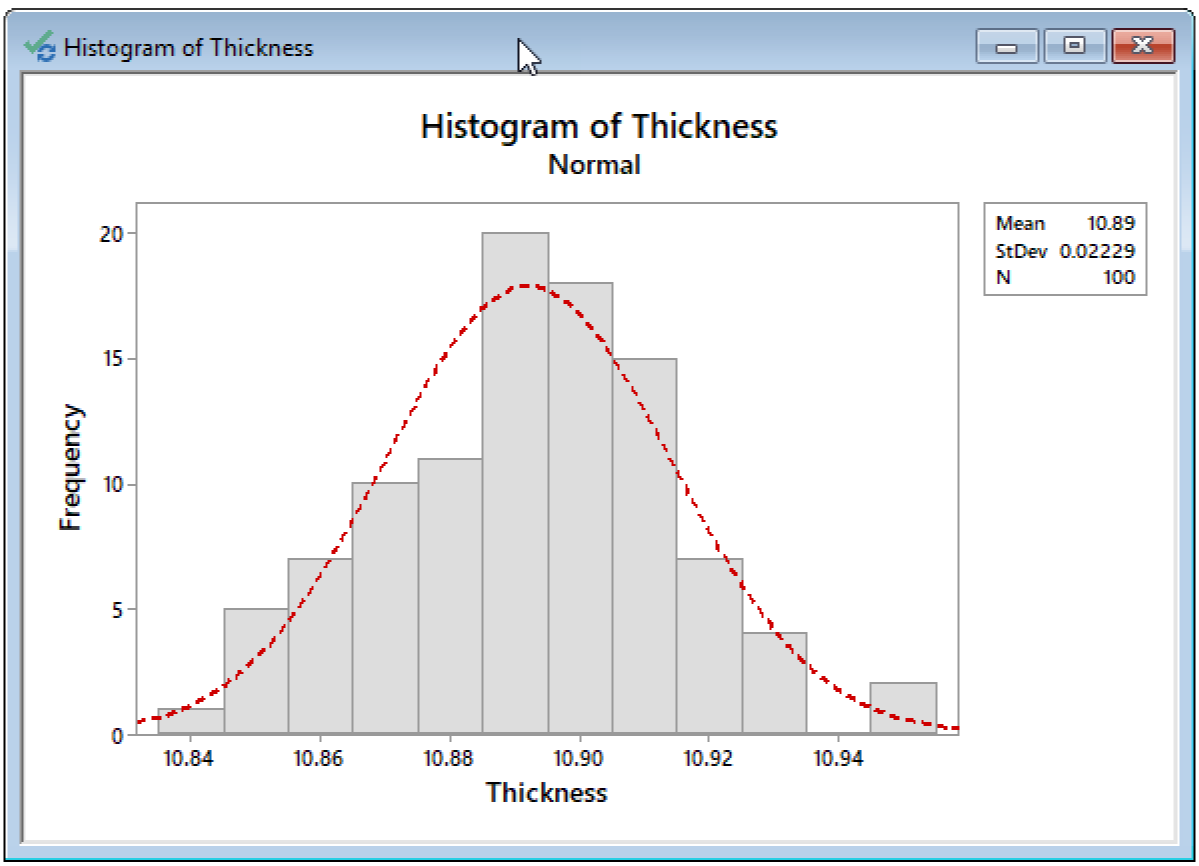
Histogram Exercise – Support Beam Thickness

Data File: Histogram.MTW

A production engineer wants to investigate the capability of the process that manufactures support beams. The engineer thinks that the process capability might differ between the morning and evening shifts. The engineer measures the thickness of five samples out of 10 boxes from each shift. The thickness must be between 10.44 mm and 10.96 mm to meet the customer requirements. Before conducting capability analysis or hypothesis testing, perform a simple histogram of beam thickness.

Solution Steps Screencast: https://youtu.be/f_HHsMWO09U

Solution Output Screenshot:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

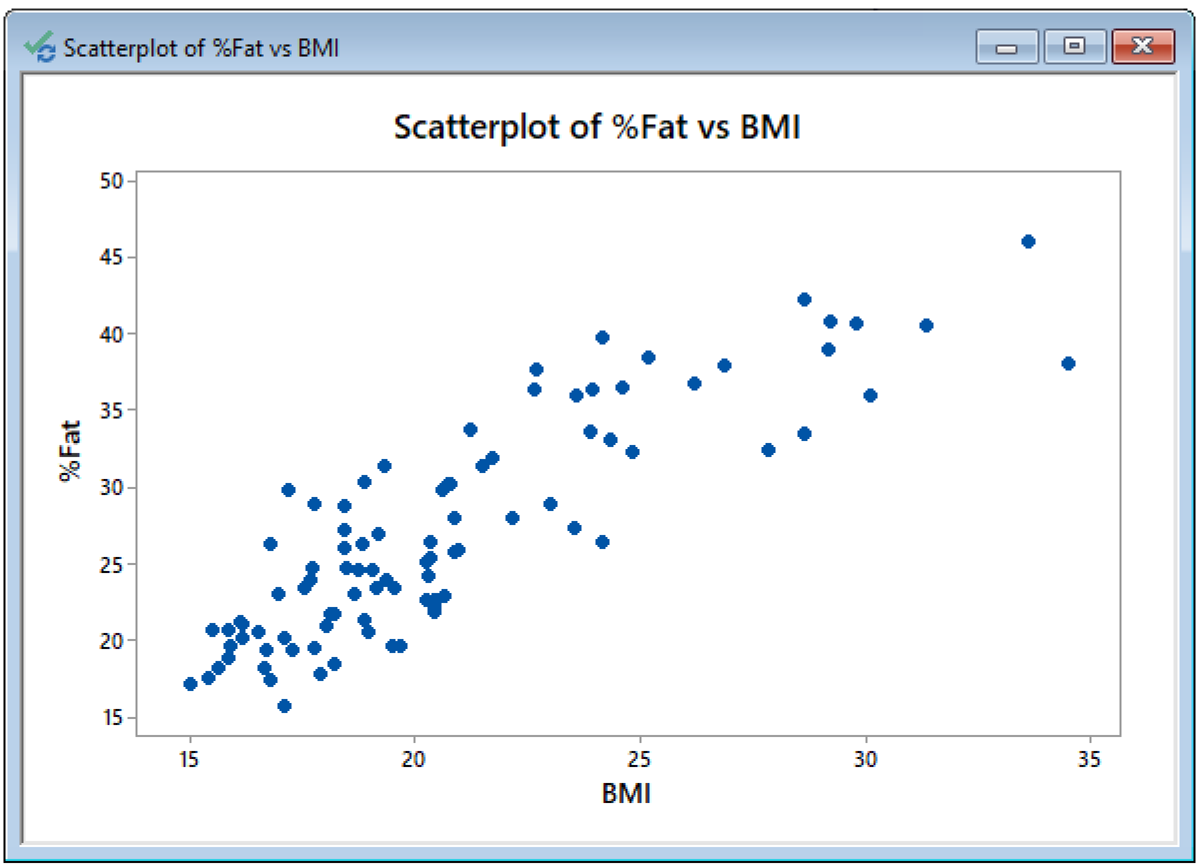
Scatterplot Exercise – Body Fat Percentage

Data File: [Scatterplot.MTW]

A medical researcher studies obesity in adolescent girls. Because body fat percentage is difficult and expensive to measure directly, the researcher wants to determine whether the body mass index (BMI), which is a measurement that is easy to take, will be a good predictor of body fat percentage. The researcher collects BMI, body fat percentage, and other personal variables of 92 adolescent girls. Use the "Scatterplot.MTW" data file to run a scatterplot on BMI vs. %Fat.

Solution Steps Screencast: https://youtu.be/bmb5w7zCluY

Solution Output Screenshot:



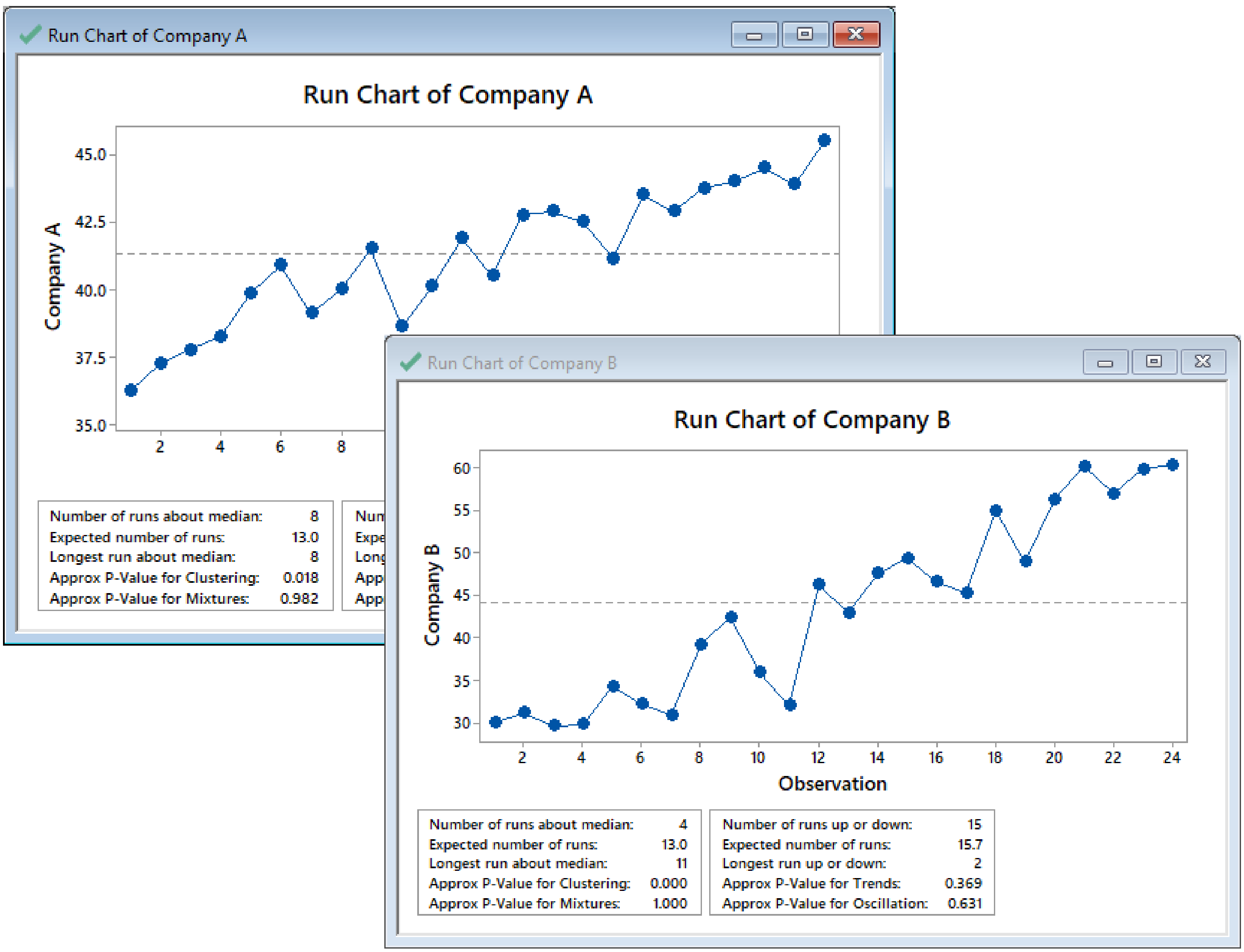
Run Chart Exercise – Stock Price

Data File: [Run_Chart.MTW]

A stock broker is reviewing the stock prices of two companies over the past 24 months and wants to create a visual chart that shows each stock performance over time. Use the "Run_Chart.MTW" data set to perform a run chart for each company stock price.

Solution Steps Screencast: https://youtu.be/rrXBTDRc2EQ

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Variable Gage R&R Exercise

Data File: [Variable_Gage_R&R.MTW]

An engineer selects 10 parts that represent the expected range of the process variation. Three operators measure the 10 parts, three times per part, in a random order. Use the "Variable_Gage_R&R.MTW" data file to perform a variable gage R&R to determine if the measurement system is of any value.

Solution Steps Screencast: https://youtu.be/WXLK6Ec584g

Solution Output Screenshots:

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Part	9	88.3619	9.81799	492.291	0.000
Operator	2	3.1673	1.58363	79.406	0.000
Part * Operator	18	0.3590	0.01994	0.434	0.974
Repeatability	60	2.7589	0.04598		
Total	89	94.6471			

α to remove interaction term = 0.05

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Part	9	88.3619	9.81799	245.614	0.000
Operator	2	3.1673	1.58363	39.617	0.000
Repeatability	78	3.1179	0.03997		
Total	89	94.6471			

Gage R&R

Variance Components

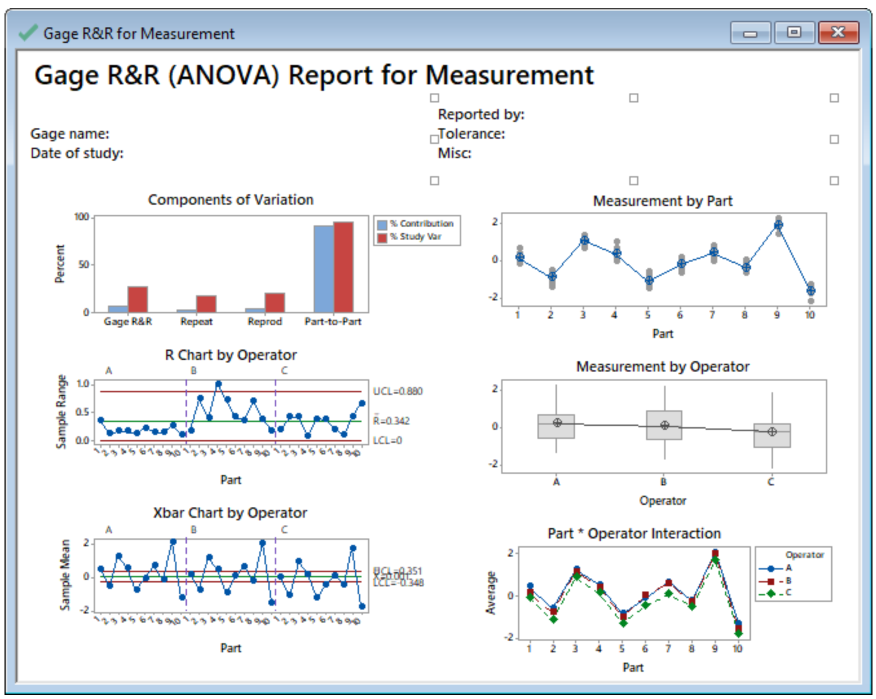
Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.09143	7.76
Repeatability	0.03997	3.39
Reproducibility	0.05146	4.37
Operator	0.05146	4.37
Part-To-Part	1.08645	92.24
Total Variation	1.17788	100.00

Gage Evaluation

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)
Total Gage R&R	0.30237	1.81423	27.86
Repeatability	0.19993	1.19960	18.42
Reproducibility	0.22684	1.36103	20.90
Operator	0.22684	1.36103	20.90
Part-To-Part	1.04233	6.25396	96.04
Total Variation	1.08530	6.51180	100.00

Number of Distinct Categories = 4

Gage R&R for Measurement



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Variable Gage R&R Result Interpretation

Components of Variation

The components of variation (% Contribution) show us that 92.24% of the variation is due to parts. This is OK, as we hope to see this measure better than 90% for a marginal measurement system but prefer it to be 99% for a good measurement system. This measure suggests that 92% of the variation found by the measurement system is due to part-to-part variation, which is where most of the variation should be found.

Study Variation

The percent study variation (% Study Var) is less than 30%, with a value of 27.86%. This again indicates a marginal result, as we would prefer to see this measure less than 10%. Anything over 30% would be considered a poor measurement system.

Distinct Categories

The Automotive Industry Action Group's (AIAG) standard for Gage R&R Distinct Category results suggest that an acceptable measurement system should have the ability to distinguish at least five distinct categories of parts. Anything less might indicate an insufficient measurement system. The result of this gage study shows distinct categories with a value of 4. We should use this information along with the other study results to draw a comprehensive conclusion.

Graphical Analysis

The Gage R&R report shows us six graphical displays that we can use to evaluate the measurement system. The two graphs of interest are the Range chart (R chart) and the Xbar chart. The R chart hints at which operator (operator B) has higher variation amoung the same part measurements. We want to see this chart in control. The Xbar chart, on the other hand, shows us each operator's mean part measurements, and this chart should be out of control.

Conclusion

Given that % Contribution and % Study variation are marginal and Distinct Categories is less than 5, we should conclude that this measurement system is marginal at best, and we should seek to improve it. Some businesses, however, may accept this measurement system depending on the criticality of what it is measuring.

Attribute Gage R&R Exercise

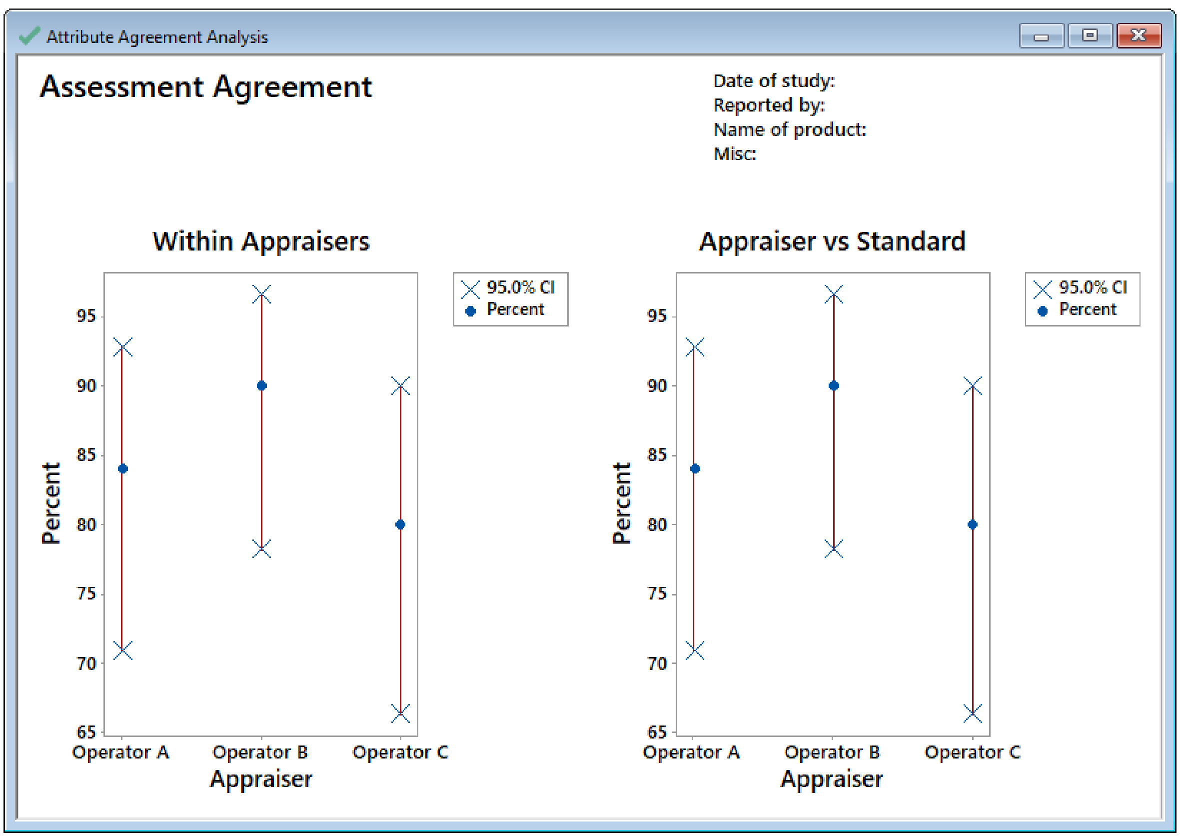
Data File: Attribute_Gage_R&R.MTW

An engineer uses an attribute measurement system that compares parts against a set of acceptance criteria. If the parts are within limits, they will be accepted. If the parts are out of bounds, they will be rejected.

To determine if the measurement system is of any value, three operators are asked to evaluate 50 parts, three times each. If the measurement system is good, the operators will agree with their own measurements (repeatability), with each other (reproducibility), and with a reference value that is the standard. Use the "Attribute_Gage_R&R.MTW" data set to perform an attribute gage R&R and determine if this engineer has an effective measurement system.

Solution Steps Screencast: https://youtu.be/NnSMHSfdOnY

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

LEAN SIGMA CORPORATION

ttribute Agreement Analysis for Assessment

Within Appraisers

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Operator A	50	42	84.00	(70.89, 92.83)
Operator B	50	45	90.00	(78.19, 96.67)
Operator C	50	40	80.00	(66.28, 89.97)

 # Matched: Appraiser agrees with him/herself across trials.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa
Operator A	0	0.760000
	1	0.760000
Operator B	0	0.845073
	1	0.845073
Operator C	0	0.702911
	1	0.702911

ach Appraiser vs Standard

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Operator A	50	42	84.00	(70.89, 92.83)
Operator B	50	45	90.00	(78.19, 96.67)
Operator C	50	40	80.00	(66.28, 89.97)

 # Matched: Appraiser's assessment across trials agrees with the known standard.

Assessment Disagreement

Appraiser	# 1 / 0	Percent	# 0 / 1	Percent	# Mixed	Percent
Operator A	0	0.00	0	0.00	8	16.00
Operator B	0	0.00	0	0.00	5	10.00
Operator C	0	0.00	0	0.00	10	20.00

 # 1 / 0: Assessments across trials = 1 / standar
 # 0 / 1: Assessments across trials = 0 / stand
 # Mixed: Assessments across trials are not id

Fleiss' Kappa Statistics

Appraiser	Response	Kappa
Operator A	0	0.880236
	1	0.880236
Operator B	0	0.922612
	1	0.922612
Operator C	0	0.774703
	1	0.774703

etween Appraisers

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with each other.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.793606	0.0235702	33.6698	0.0000
1	0.793606	0.0235702	33.6698	0.0000

All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with the known standard.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.859184	0.0471405	18.2260	0.0000
1	0.859184	0.0471405	18.2260	0.0000

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Attribute Gage R&R Result Interpretation:

Within Appraiser

The within-appraiser agreement percentages range from about 80% to 90%, so each appraiser agreed with themselves at least 80% of the time. This is a pretty good indication that the measurement system is sound in terms of repeatability. Also, the kappa statistics for within-appraiser agreement are all above 0.7, supporting this assessment.

Between Appraisers

The between-appraiser agreement, which is a measure of how well the appraisers agree with each other, is also showing a fair result: 78% of the time, the appraisers agreed with each other, and the Fleiss' kappa statistic is 0.79. This statistic will range from -1 to 1, with -1 being perfect disagreement and 1 being perfect agreement. Anytime this statistic is above 0.7, we should be able to conclude that there is strong agreement between appraisers.

Appraiser vs. Standard

When assessing the appraiser's ability to match the standard, they were all able to do so effectively at least 80% of the time. The kappa statistics are all greater than 0.7, ranging from 0.77 to 0.92.

Conclusion

In terms of repeatability and reproducibility, the engineer's overall measurement system should be considered acceptable. If we are looking for opportunities for improvement, we could start with Operator C, whose 95% confidence interval for repeatability ranges from 66% to 89%. This suggests that 95% of the time Operator C's repeatability will be between 66% and 89%. That range slips below our 70% threshold, and we might want to seek ways for that operator to improve their measurement process.

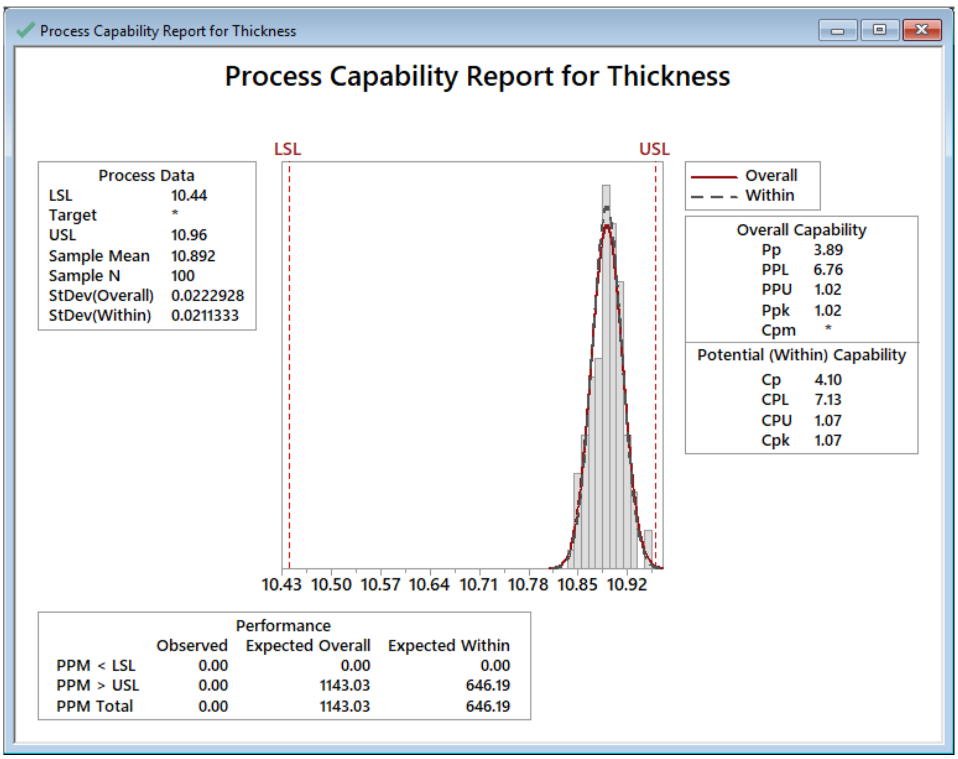
Capability Analysis Exercise – Support Beam Thickness

Data File: *Capability.MTW*

A production engineer wants to investigate the capability of a process that manufactures support beams. The engineer measures the thickness of five samples out of 10 boxes from each shift. The thickness must be between 10.44 mm and 10.96 mm to meet customer requirements. Using the “Capability.MTW” data file, conduct a capability analysis on the thickness across both shifts.

Solution Steps Screencast: https://youtu.be/LXX4pbyKrbc

Solution Output Screenshot:



Capability Analysis Interpretation

The capability analysis results show us a Ppk of 1.02 and a Cpk of 1.07. These results are below the AIAG recommendations of 1.67 and 1.33, respectively. However, both measures being greater than 1 indicates that the difference between the process mean and either the USL or LSL is greater than 3-sigma. Therefore, unless your business mandates something greater, we can conclude that the process is capable.

Analyze Phase Exercises

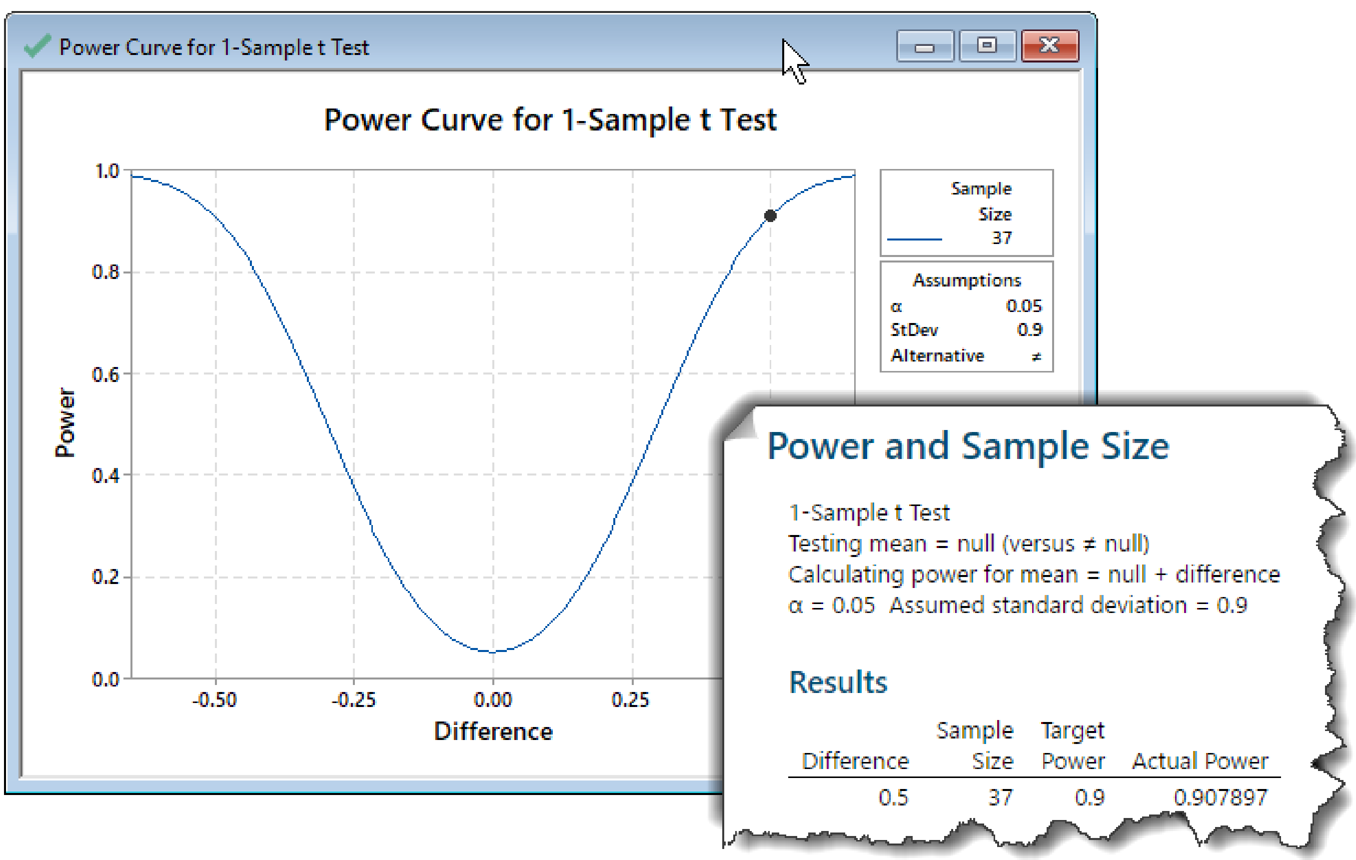
Sample Size Exercise – Beam Length

Data File: N/A

The manager of a lumber yard wants to assess the performance of a saw mill that cuts beams that are supposed to be 100 cm long with a tolerance of ±0.5 cm. The manager has some historical data that indicate a mean of 99.5 cm with a standard deviation of 0.9 cm. Using this information, determine the sample size necessary for the manager to perform a 1-sample t-test to determine if the beams are being cut at a length of 100 cm. The manager wants to have a 90% chance of detecting the difference between the population mean and the target value of 100 cm.

Solution Steps Screencast: https://youtu.be/EnsZLAzixmU

Solution Steps Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Sample Size Result Interpretation

Difference

In determining the proper sample size to assess the mill's ability to cut beams at 100 cm (±0.5 cm), we entered a difference of 0.5. Minitab calculates the minimum difference for which you can achieve the specified level of power for your sample size. Larger sample sizes enable the test to detect smaller differences. The mill manager wants to detect the smallest difference that has practical consequences for his purposes. Therefore, the manager chose 0.5 as his difference.

Sample Size

Minitab calculates how large the manager's sample must be for a test with the specified power to detect the specified difference. Because sample sizes are whole numbers, the actual power of the test might be slightly greater than the power value that the manager specified. If the manager were to increase the sample size, the power of the test would also increase. As a rule, you want enough observations in your sample to achieve adequate power. But you don't want a sample size so large that you waste time and money on unnecessary sampling to detect unimportant differences.

Power

Minitab will calculate the power of the test based on the specified difference and sample size. A power value of 0.9 is usually considered adequate. A value of 0.9 indicates that the manager will have a 90% chance of detecting the difference of 0.5 cm between the population mean and the target value when a difference exists. If the test has low power, the manager might fail to detect a difference and mistakenly conclude that no difference exists. Usually, when the sample size is too small, the test has less power to detect a difference.

Conclusion

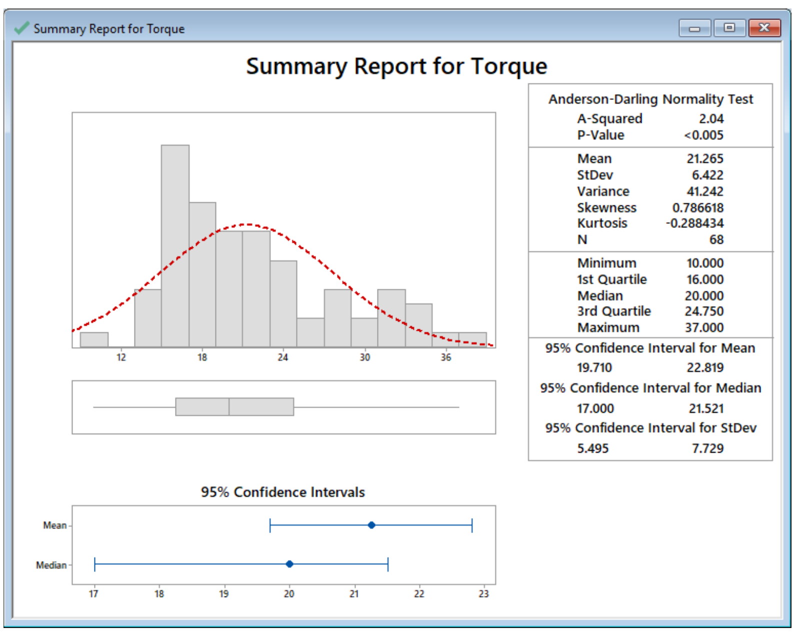
In this exercise, the important factors were difference, standard deviation, and power. No matter what the historical mean was, the saw mill manager wants to be able to detect if the mean is 100 cm ± 0.5 cm, and he wants 90% confidence that he can see a difference if one exists. Therefore, the proper sample size would be 37. With a sample size of 37, the manager can detect a 0.5 cm difference with 90% confidence and a test power of 0.9078.

Confidence Interval Exercise – Cap Torque

Data File: Confidence_Interval.MTW
A quality control engineer is responsible for ensuring that the caps on shampoo bottles are fastened correctly. If the caps are fastened too loosely, they may fall off or the bottles may leak during shipping. If the caps are fastened too tightly, they may be too difficult to remove. The target torque value for fastening the caps is 18. The engineer collects a random sample of 68 bottles and needs to first determine the 95% confidence interval of the mean torque required to remove the caps. Use the “Confidence_Interval.MTW” data set to determine the 95% confidence interval for mean torque.

Solution Steps Screencast: https://youtu.be/A16Ohp-0ncc

Solution Output Screenshots:



Confidence Interval Result Interpretation

The graphical summary output from Minitab provides three confidence intervals for the mean, median, and standard deviation. The quality control engineer wants to determine the 95% confidence interval for the mean, which was found to be between 19.7 and 22.8. Therefore, she is 95% confident the mean torque required to remove the caps as determined from her sample of 68 is between 19.7 and 22.8.

1-Sample t-Test Exercise – Family Energy Costs

Data File: [1-Sample-T.MTW]

An economist wants to determine whether the monthly energy cost for families has changed from the previous year, when the mean cost per month was $200. The economist randomly samples 25 families and records their energy costs for the current year.

Solution Steps Screencast: https://youtu.be/uGHTConE-CY

Solution Output Screenshots:

One-Sample T: Energy Cost

Descriptive Statistics

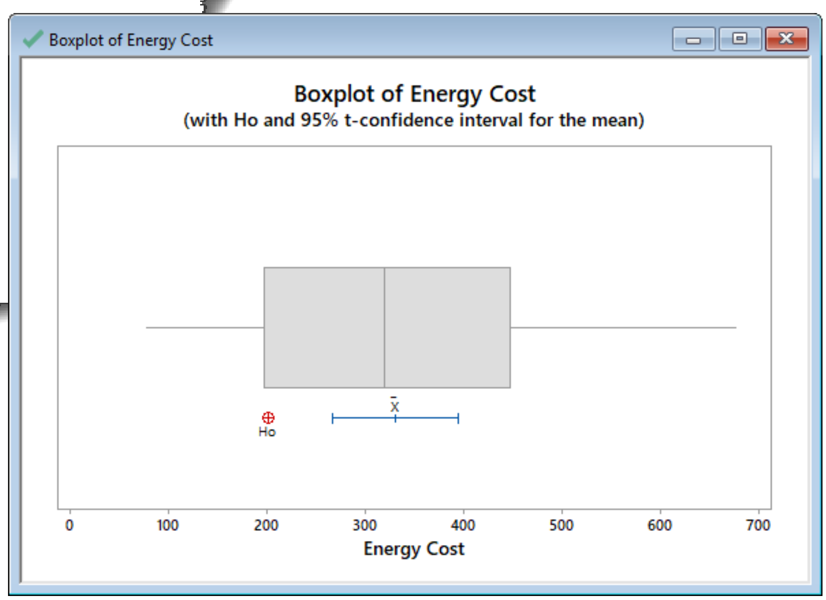
N	Mean	StDev	SE Mean	95% CI for μ
25	330.6	154.2	30.8	(266.9, 394.2)

μ: mean of Energy Cost

Test

Null hypothesis $H_0: \mu = 200$
Alternative hypothesis $H_1: \mu \neq 200$

T-Value	P-Value
4.23	0.000



1-Sample t-Test Result Interpretation

The 1-sample t-test results indicate that this year's family energy costs are statistically different from the prior year mean of $200. With a new mean of $330.6 and a p-value of 0.00, we reject the null hypothesis (H_0), which states that there is no difference, and conclude that there is a statistical difference.

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Test of Equal Variance Exercise – Calcium Intake

Data File: [*Equal_Variances.MTW*]()

A nutrition consulting company created an education program to increase the calcium intake in children ages 9 to 13. To measure the effectiveness of the program, an analyst performs an experiment in which 198 children are assigned randomly to either the control group (no education program) or the treatment group (with education program). The average daily dietary calcium intake is calculated from three-day diet records. Before the analyst can perform hypothesis tests, she needs to know if the variances of the two groups are equal or not so that she can determine what type of hypothesis test to use (parametric or non-parametric). Use the "Equal_Variances.MTW" data file to perform a test of equal variances and determine if the two groups have equal variances.

Solution Steps Screencast: https://youtu.be/qipJEeCN1ZE

Solution Output Screenshots:

st and CI for Two Variances: Calcium vs Group

Method

σ_1: standard deviation of Calcium when Group = Control
σ_2: standard deviation of Calcium when Group = Education
Ratio: σ_1/σ_2
The Bonett and Levene's methods are valid for any continuous distribution.

Descriptive Statistics

Group	N	StDev	Variance	95% CI for σ
Control	98	247.478	61245.340	(220.158, 283.865)
Education	100	183.999	33855.625	(162.194, 212.909)

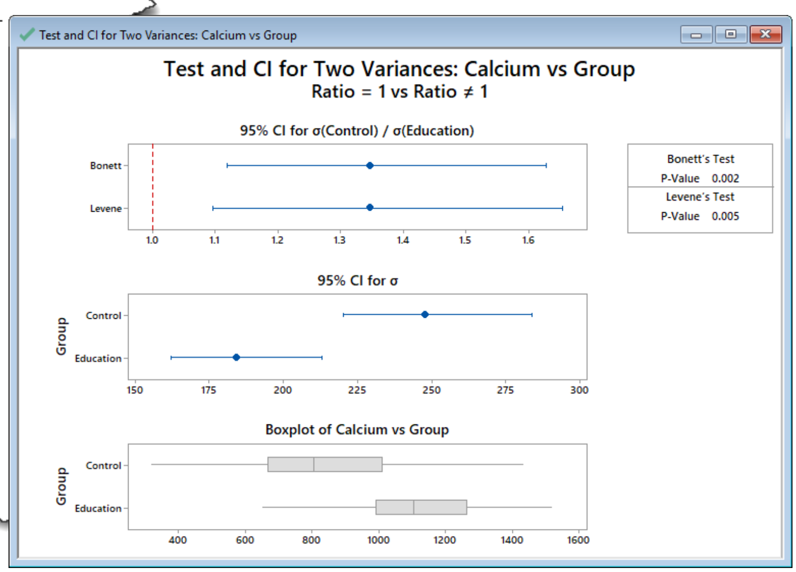
Ratio of Standard Deviations

Estimated Ratio	95% CI for Ratio using Bonett	95% CI for Ratio using Levene
1.34500	(1.118, 1.627)	(1.096, 1.654)

Test

Null hypothesis $H_0: \sigma_1 / \sigma_2 = 1$
Alternative hypothesis $H_1: \sigma_1 / \sigma_2 \neq 1$
Significance level $\alpha = 0.05$

Method	Test Statistic	DF1	DF2	P-Value
Bonett	*			0.002
Levene	7.99	1	196	0.005



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Test of Equal Variance Result Interpretation

Since Levene's test is less sensitive to non-normality and we have not determined if the data are normally distributed, we will use Levene's test statistics to draw our conclusion. From Minitab's output, we can see that Levene's test p-value is less than the significance level of 0.05. Therefore, we reject the null hypothesis, which states that the variances are equal, and conclude that the variances are not equal.

Graphically, we can also observe that the 95% confidence intervals for sigma of the two groups have no overlap, and this also suggests that the variances are not equal.

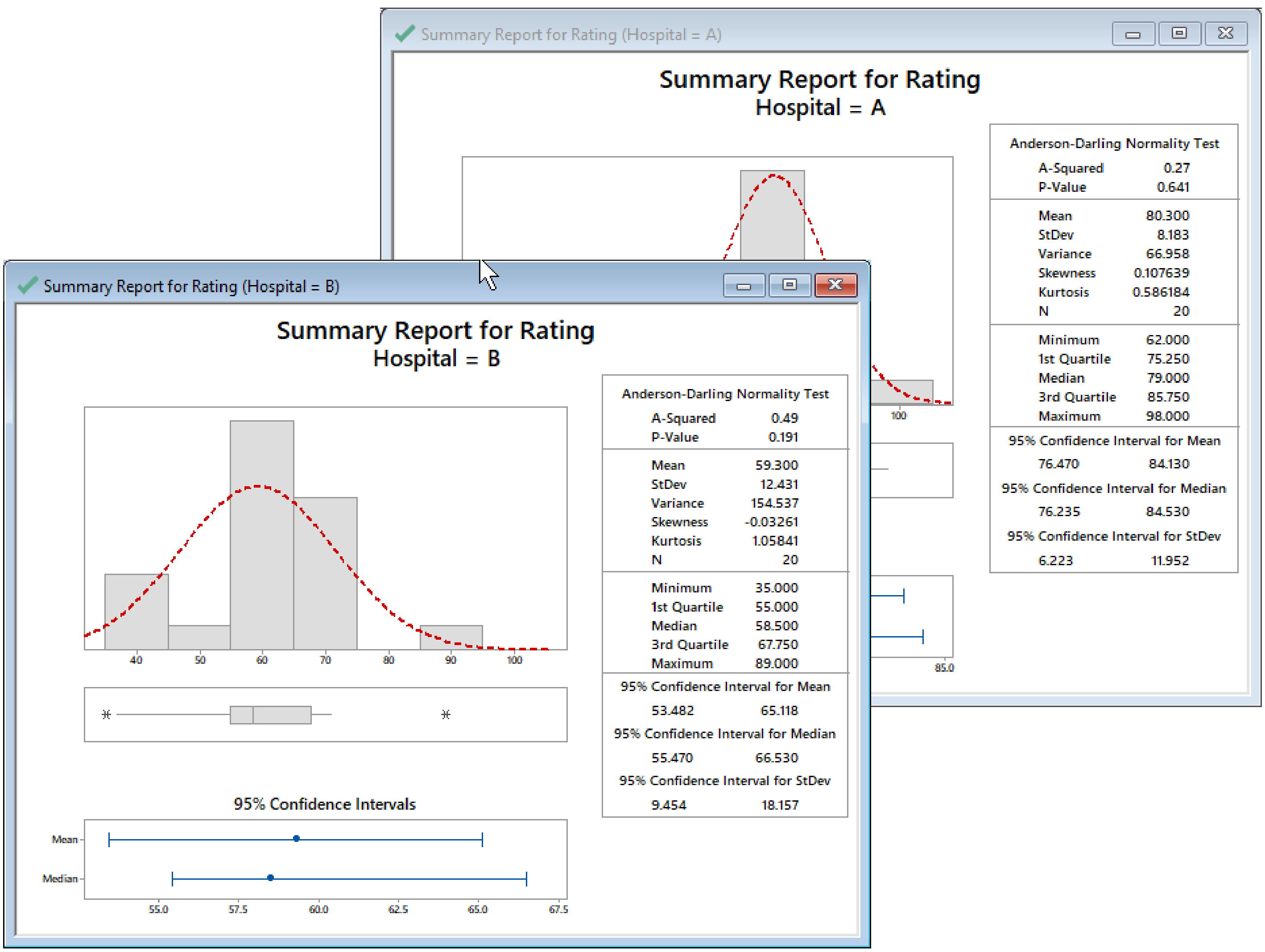
2-Sample t-Test Exercise – Hospital Satisfaction

Data File: [2-Sample-T.MTW]

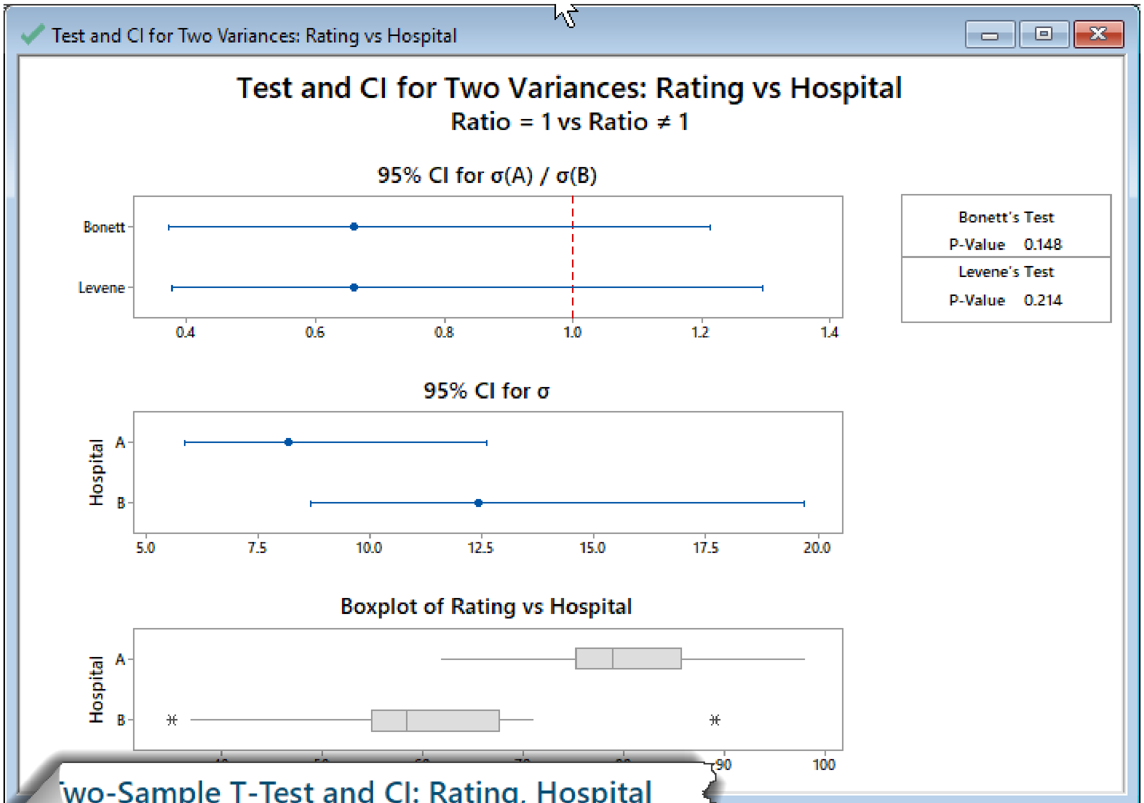
As a mid-level analyst at a healthcare conglomerate, you have been asked to determine if there is a difference in satisfaction ratings between two hospitals serving a particular geographical region. Your manager has emailed you 20 satisfaction data points from each hospital. After combining them in your data file "2-Sample-T.MTW," perform a 2-sample t-test to determine if there is a difference in satisfaction ratings.

Solution Steps Screencast: https://youtu.be/HUB8pVNhNbI

Solution Steps Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/



Two-Sample T-Test and CI: Rating, Hospital

Method

μ_1: mean of Rating when Hospital = A
μ_2: mean of Rating when Hospital = B
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics: Rating

Hospital	N	Mean	StDev	SE Mean
A	20	80.30	8.18	1.8
B	20	59.3	12.4	2.8

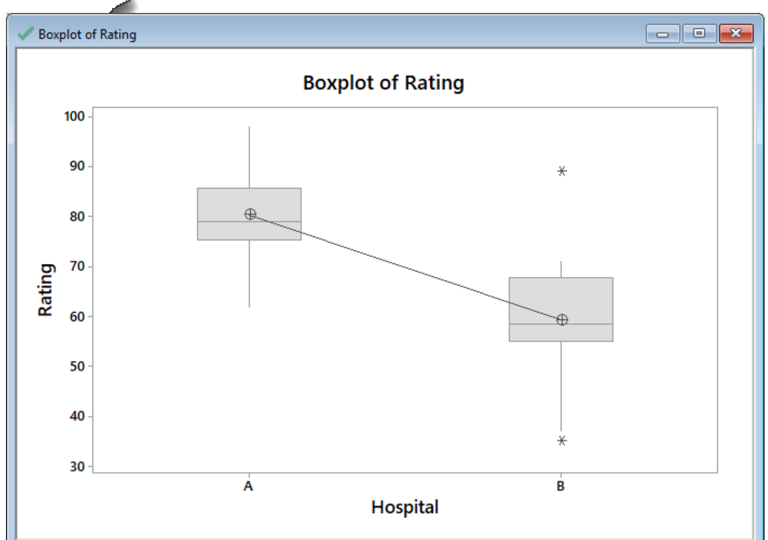
Estimation for Difference

Difference	Pooled StDev	95% CI for Difference
21.00	10.52	(14.26, 27.74)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
6.31	38	0.000



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

2-Sample t-Test Result Interpretation

Normality

Before conducting a 2-sample t-test, we must first determine if the data are normal. If not, a different hypothesis test may be necessary. In this case, both data sets from each hospital are normally distributed with p-values of 0.641 and 0.191 for hospitals A and B, respectively. As a result, we fail to reject the null hypothesis and conclude normality.

Equal Variances

Before performing the 2-sample t-test it is also important to determine if the two data sets have equal variances. If not, a selection of running a 2-sample t-test assuming unequal variances will be necessary. Based on the result of the test of equal variances, we fail to reject the null hypothesis and conclude equal variances.

2-Sample T-Test

We can now perform a 2-sample t-test assuming equal variances. Given the resulting p-value of 0.000, we reject the null hypothesis and conclude that there is a difference between the two hospitals satisfaction ratings (mean of hospital A = 80.3, mean of hospital B = 59.3).

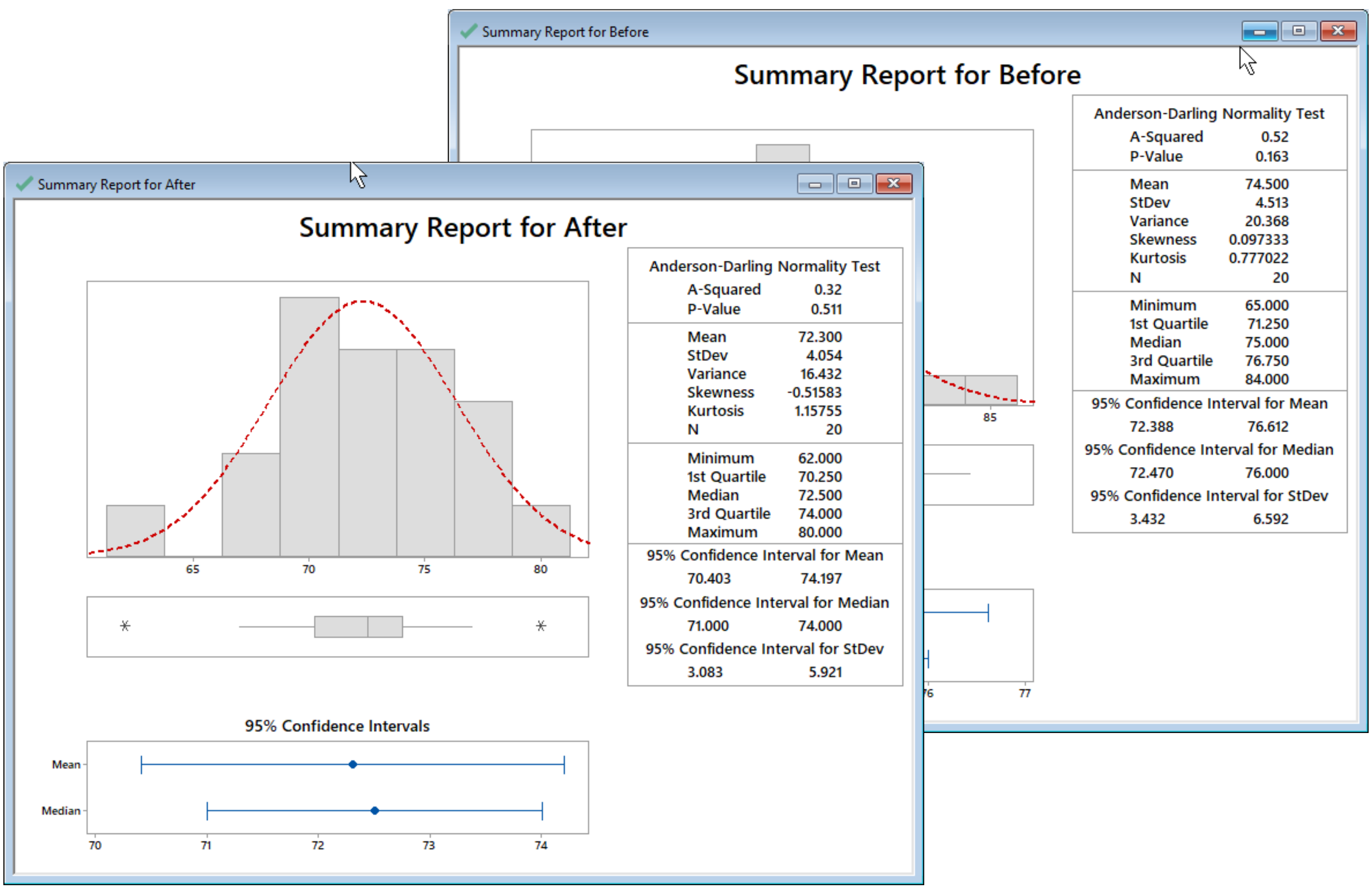
Paired t-Test Exercise – Resting Heart Rate

Data File: *Paired-t.MTW*

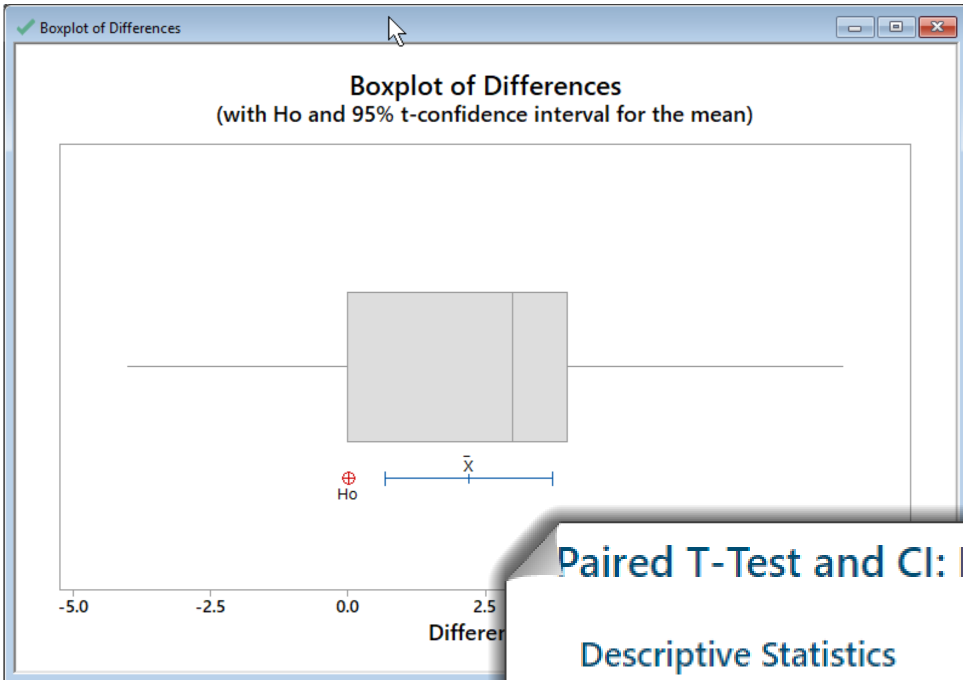
As a physiologist, you want to determine whether a running program influences resting heart rate. The heart rates of 15 randomly selected people were measured. The people were then put on the running program and measured again one year later. Thus, the "Before" and "After" measurements for each person are a pair of observations with obvious dependencies. Use the "Paired-t.MTW" data file to perform a paired t-test to determine whether the heart rates differ before and after the running program.

Solution Steps Screencast: https://youtu.be/IfDs26L8PZU

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/



Paired T-Test and CI: Before, After

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Before	20	74.50	4.51	1.01
After	20	72.30	4.05	0.91

Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for µ_difference
2.200	3.254	0.728	(0.677, 3.723)

µ_difference: mean of (Before - After)

Test

Null hypothesis H_0: µ_difference = 0

Alternative hypothesis H_1: µ_difference ≠ 0

T-Value	P-Value
3.02	0.007

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Paired t-Test Result Interpretation

Normality

An assumption of the paired t-test is normality. Therefore, we must first determine if each of the data sets is normally distributed. After performing a graphical summary, we can see the Anderson-Darling p-value for normality of each data set. They are both greater than 0.05, so we fail to reject the null hypothesis and conclude that the data are normally distributed.

Paired t-Test

After addressing the tests for normality, we can perform the paired t-test and determine if there is a statistically significant difference between the resting heart rates of the “Before” and “After” data sets.

The test output shows us that the 95% confidence interval of the mean difference is a range that does not include zero. Armed with this information and the p-value of 0.007, we must reject the null hypothesis, which states that there is no difference. So, we conclude that there is a statistically significant difference between the resting heart rates of the “Before” and “After” samples.

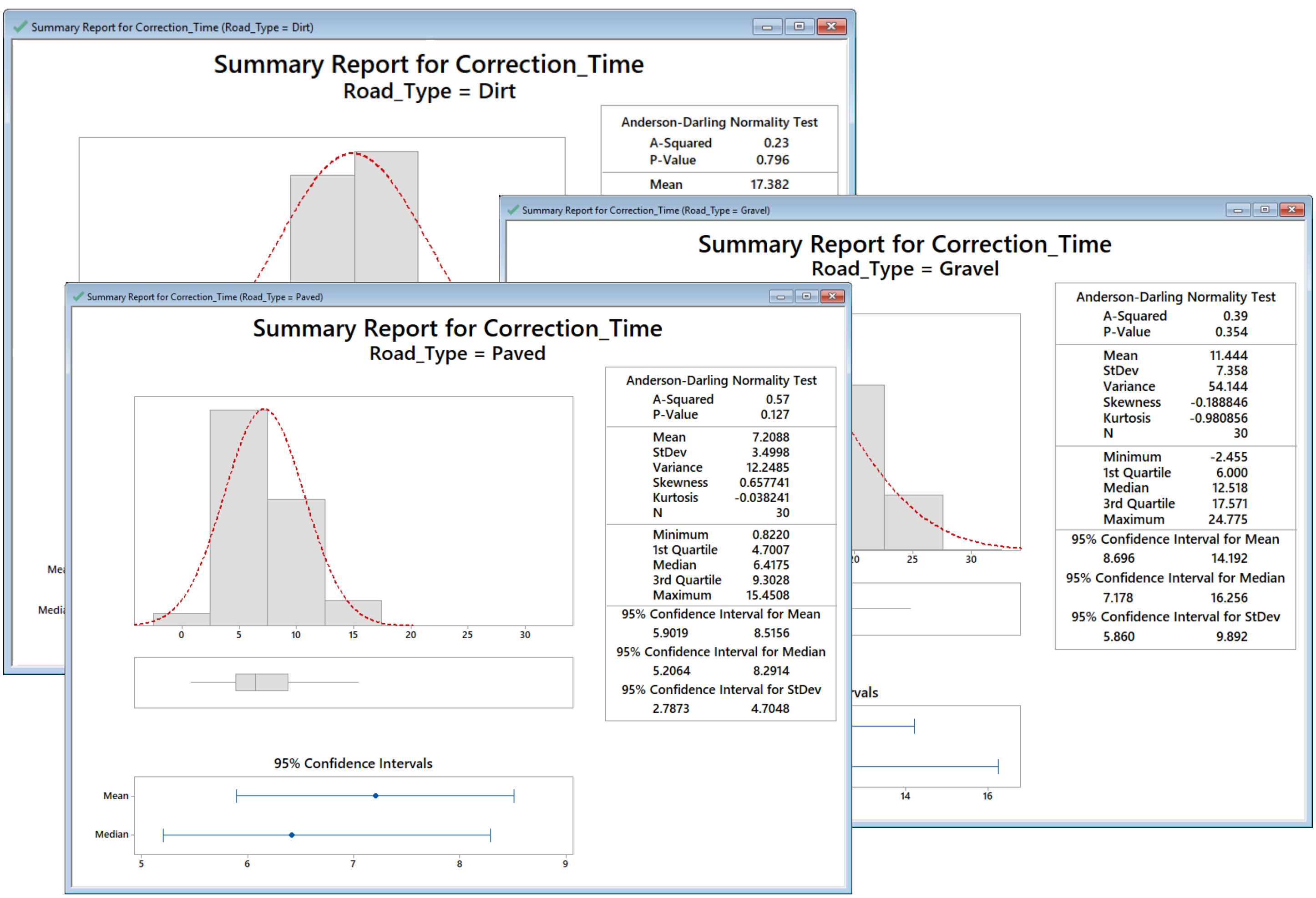
ANOVA Test Exercise – Road Surface Correction Times

Data File: [*ANOVA.MTW*]

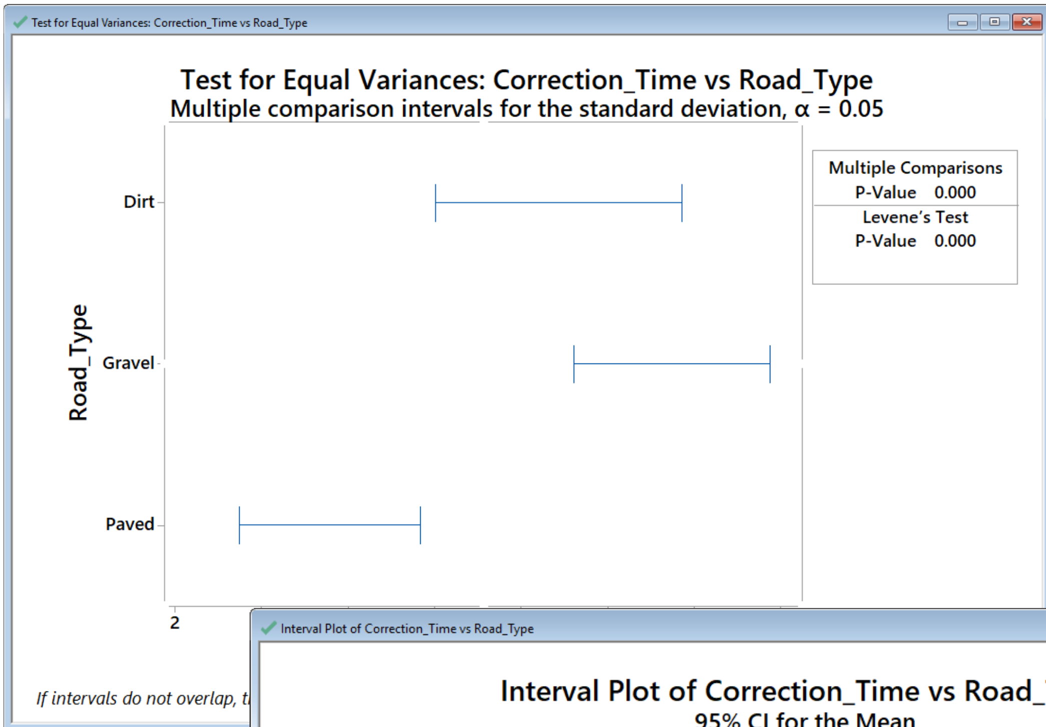
A safety analyst wants to compare how well drivers drive on three types of roads: paved, gravel, and dirt. To measure driving performance, the analyst records the time in seconds that each driver uses to make steering corrections on each type of road. All other variables (vehicle type, speed, tires, tire air pressure, etc.) are held constant. Use the "ANOVA.MTW" data file to perform an ANOVA test to determine which road surface type yields the best and worst correction times.

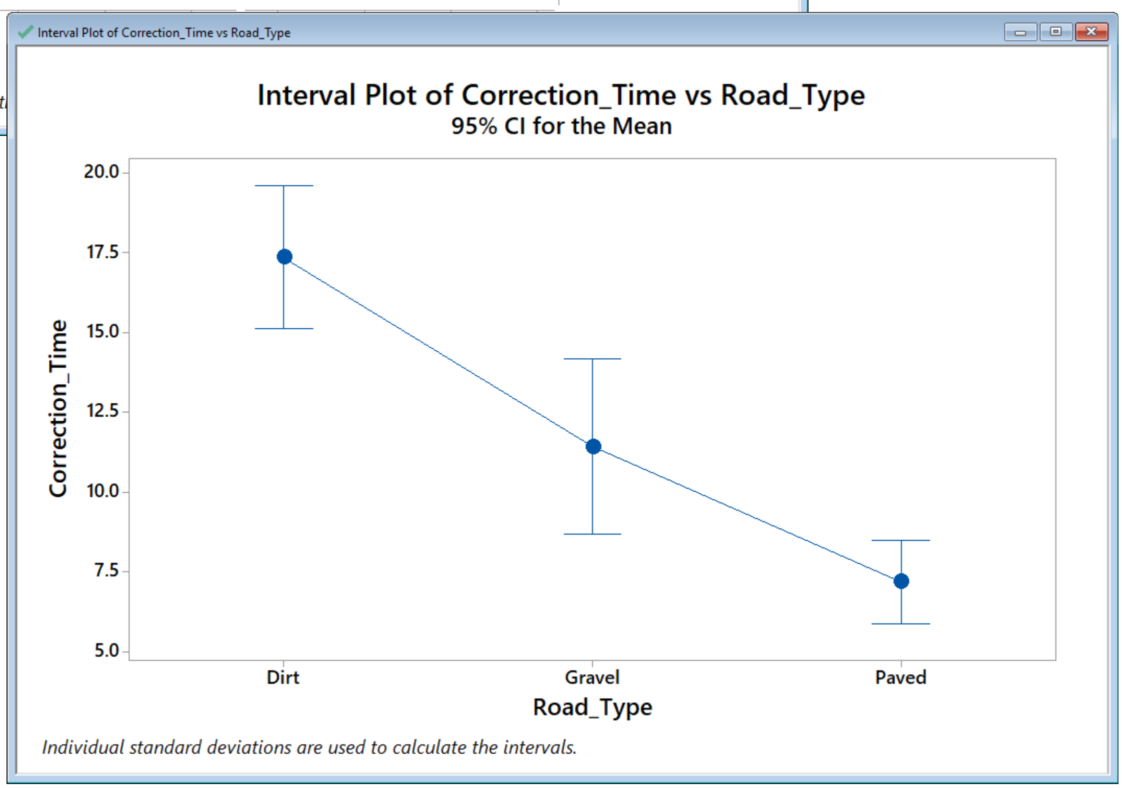
Solution Steps Screencast: https://youtu.be/XRgI-kIV0nA

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/


Test for Equal Variances: Correction_Time vs Road_Type
Test for Equal Variances: Correction_Time vs Road_Type
Multiple comparison intervals for the standard deviation, α = 0.05
Multiple Comparisons
P-Value 0.000
Levene's Test
P-Value 0.000
Road_Type
Dirt
Gravel
Paved
2
If intervals do not overlap, t


Interval Plot of Correction_Time vs Road_Type
Interval Plot of Correction_Time vs Road_Type
95% CI for the Mean
Correction_Time
20.0
17.5
15.0
12.5
10.0
7.5
5.0
Dirt
Gravel
Paved
Road_Type
Individual standard deviations are used to calculate the intervals.

LEAN SIGMA
CORPORATION

ne-way ANOVA: Correction_Time versus Road_Type

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0.05

Equal variances were not assumed for the analysis.

Factor Information

Factor	Levels	Values
Road_Type	3	Dirt, Gravel, Paved

Welch's Test

Source	DF Num	DF Den	F-Value	P-Value
Road_Type	2	52.3708	32.19	0.000

Model Summary

R-sq	R-sq(adj)	R-sq(pred)
34.46%	32.96%	29.87%

Means

Road_Type	N	Mean	StDev	95% CI
Dirt	30	17.38	6.03	(15.13, 19.63)
Gravel	30	11.44	7.36	(8.70, 14.19)
Paved	30	7.209	3.500	(5.902, 8.516)

ANVOA Test Result Interpretation

Before performing the ANOVA test to determine the correction times based on road surface type, we must understand the assumptions of this test.

First, we perform tests of normality to determine if the data are normal. In doing so, we find that all three data sets per road surface type are normally distributed.

Second, we determine if there are equal variances between road surfaces. The results of the test of equal variance tell us that variances are not equal. So, before running the ANOVA test, we must remove the assumption of equal variances. Therefore, we uncheck the “Assume equal variances” check box and continue running the ANOVA.

The results of the ANOVA demonstrate that there are statistically significant differences between correction times based on road surface type, with paved surfaces clearly showing better correction times than both gravel and dirt surfaces. Additionally, dirt surfaces are the worst performing in correction times.

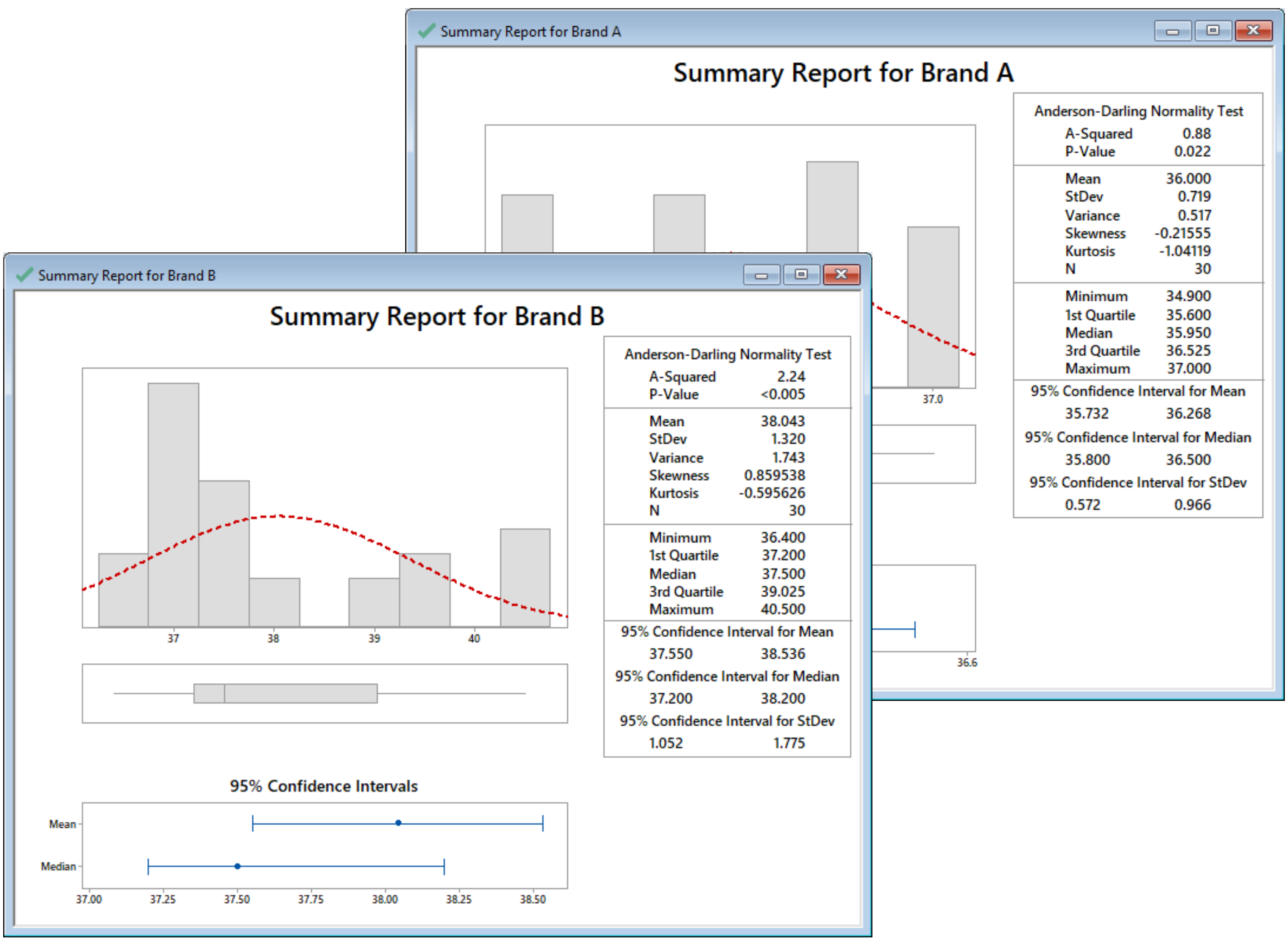
Mann–Whitney Test Exercise – Highway Paint

Data File: [Mann-Whitney.MTW]()

A state highway department uses two brands of paint for painting stripes on roads. A highway official wants to know whether there's a difference in the durability of the two brands of paint. For each paint type, the official researches and records the number of months the paint persists on the highway. Use the data in the "Mann-Whitney.MTW" data file to perform a test to determine whether the median number of months the paint persists differs between the two brands.

Solution Steps Screencast: https://youtu.be/mQ9CzXFWjlI

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Mann-Whitney: Brand A, Brand B

Method

η_1: median of Brand A
η_2: median of Brand B
Difference: $\eta_1 - \eta_2$

Descriptive Statistics

Sample	N	Median
Brand A	30	35.95
Brand B	30	37.50

Estimation for Difference

Difference	CI for Difference	Achieved Confidence
-1.7	(-2.3, -1.4)	95.16%

Test

Null hypothesis	H_0: $\eta_1 - \eta_2 = 0$
Alternative hypothesis	H_1: $\eta_1 - \eta_2 \neq 0$

Method	W-Value	P-Value
Not adjusted for ties	501.00	0.000
Adjusted for ties	501.00	0.000

Mann–Whitney Test Result Interpretation

Normality

You can use the Mann–Whitney test when you have non-normally distributed data. When distributions are not well represented by means, the use of a median test can be more appropriate. Therefore, we perform a test of normality on each of the paint persistence data sets and conclude that they are not normally distributed, because the Anderson-Darling p-values are below the alpha level of 0.05.

Mann–Whitney Test

The results of the Mann–Whitney test demonstrate that the difference between the medians of Brand A and Brand B is statistically significant. The p-value of 0.000, adjusted for ties, is less than the alpha level of 0.05. Therefore, we reject the null hypothesis, which states that there is no statistical difference, and conclude that a statistical difference exists between the two medians.

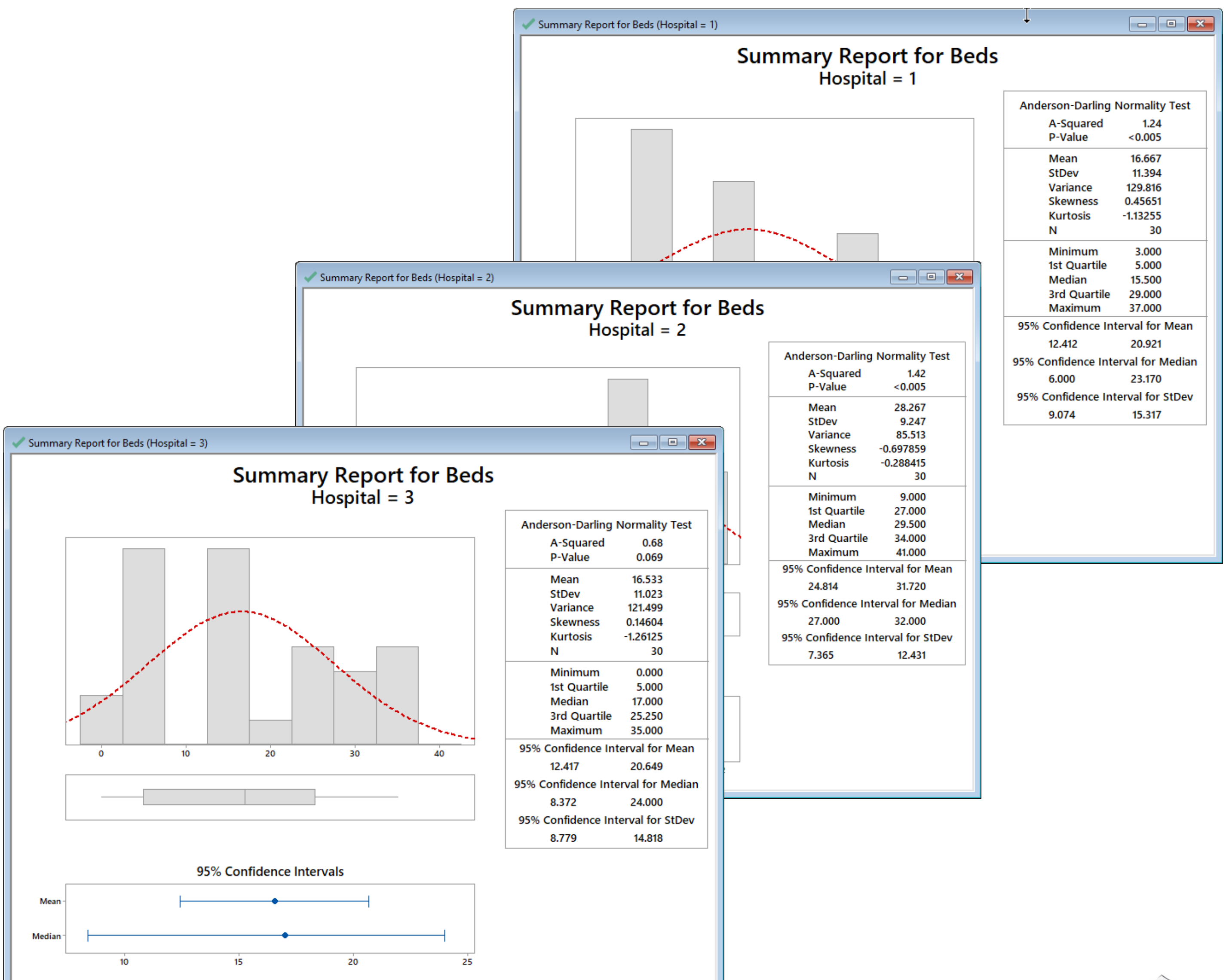
Kruskal–Wallis Test Exercise – Unoccupied Hospital Beds

Data File: [Kruskal-Wallis.MTW]()

A health administrator wants to compare the number of unoccupied beds for three hospitals in the same region. The administrator randomly selects 30 different days from the records of each hospital and enters the number of unoccupied beds for each day into the file "Kruskal-Wallis.MTW." The administrator then asks you to determine if there is a difference in the number of unoccupied beds between the three hospitals.

Solution Steps Screencast: https://youtu.be/PcBdUw6ouQY

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Kruskal-Wallis Test: Beds versus Hospital

Descriptive Statistics

Hospital	N	Median	Mean Rank	Z-Value
1	30	15.5	37.8	-1.98
2	30	29.5	61.8	4.18
3	30	17.0	37.0	-2.20
Overall	90		45.5	

Test

Null hypothesis	H_0: All medians are equal
Alternative hypothesis	H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	2	17.46	0.000
Adjusted for ties	2	17.52	0.000

Kruskal–Wallis Test Result Interpretation

Normality

You can use the Kruskal–Wallis test when you have non-normally distributed data. When distributions are not well represented by means, the use of a median test can be more appropriate. Therefore, instead of a one-way ANOVA, use the Kruskal–Wallis test to compare the medians of three or more groups. Here, we perform a test of normality on each of the hospital bed data sets and conclude that two of the three were not normally distributed. This is evidenced by the Anderson-Darling p-values below the alpha level of 0.05.

Kruskal–Wallis Test

After addressing the tests for normality, we can perform the Kruskal–Wallis test and determine if there is a statistically significant difference between the unoccupied beds of Hospitals 1, 2, and 3.

The test output demonstrates that the difference between the number of unoccupied beds between the three hospitals is significant. The p-value of 0.000 “adjusted for ties” is less than the alpha level of 0.05. Hence, we reject the null hypothesis and conclude that a statistical difference exists.

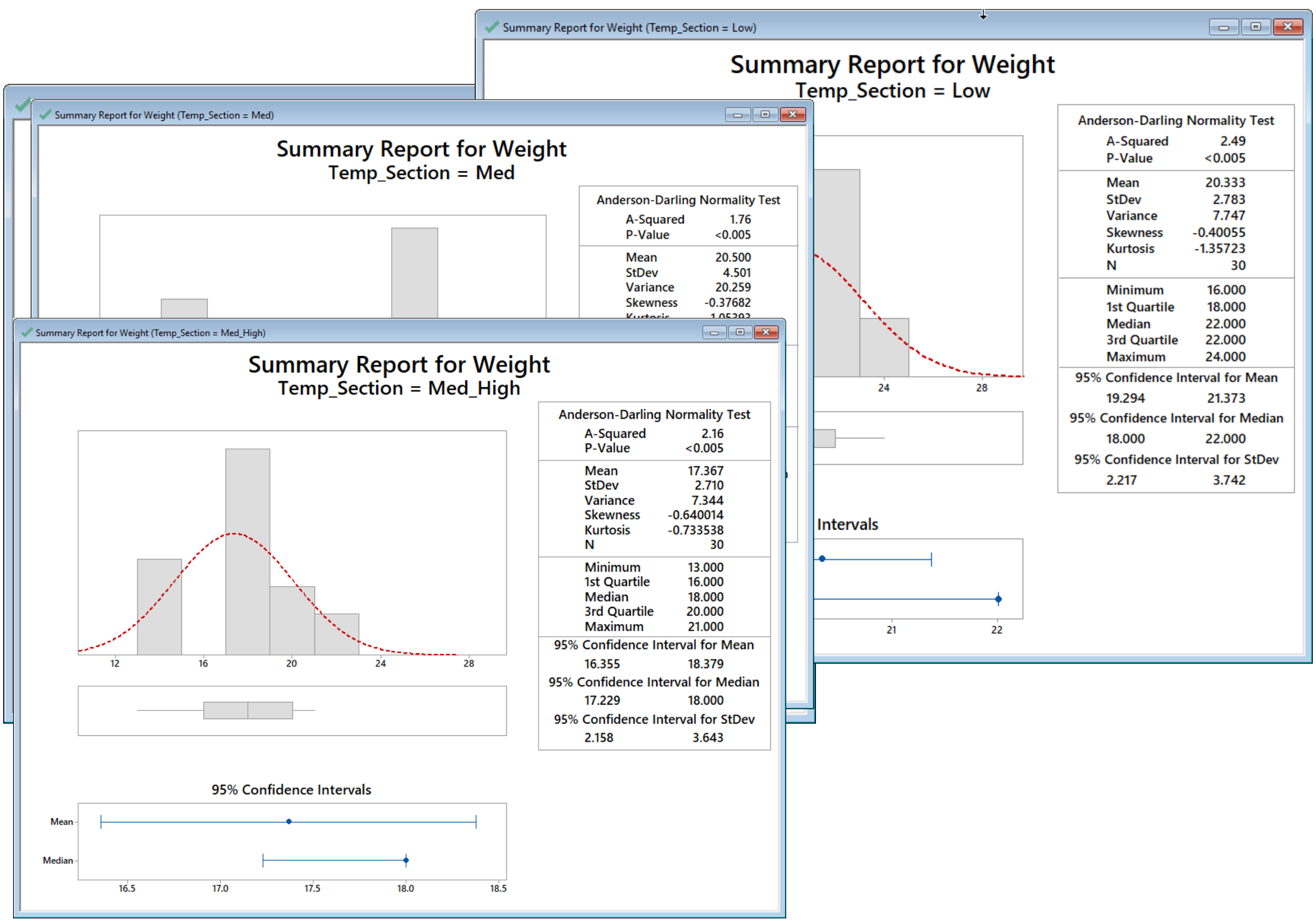
Moods Median Test Exercise – Fish Growth vs. Water Temp.

Data File: [Moods_Median.MTW]

An environmental scientist wants to determine whether the temperature changes in the ocean near a nuclear power plant affect the growth of fish. The scientist randomly divides 25 fish into four groups and places each group into a separate, simulated ocean environment. The simulated environments are identical except for temperature. Six months later, the scientist measures the weight of the fish. Use the “Moods_Median.MTW” data file to determine if water temperature influences the weight of fish.

Solution Steps Screencast: https://youtu.be/NsU2NGcMLe8

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab’s “Data Set Library” at https://support.minitab.com/en-us/datasets/

Mood's Median Test: Weight versus Temp_Section

Descriptive Statistics

Temp_Section	Median	N <= Overall Median	N > Overall Median	Q3 – Q1	95% Median CI
Low	22	12	18	4	(18, 22)
Low_Med	17	19	11	9	(17, 21)
Med	21	8	22	10	(19, 24)
Med_High	18	22	8	4	(17.2287, 18)
Overall	18				

Test

Null hypothesis H_0: The population medians are all equal
Alternative hypothesis H_1: The population medians are not all equal

DF	Chi-Square	P-Value
3	16.37	0.001

Mood's Median Test Result Interpretation

We commonly use Mood's median test when our data are non-normal and when comparing three or more factor levels. In this case, we are comparing fish growth as determined by weight vs. four water temperature environments.

Our first action is to determine if the data are normal. If they are, we will use ANOVA for this type of analysis. Otherwise, we will use Mood's median. The results of the normality tests for fish weight relative to each section of water temperature show that the data are not normal.

Therefore, we will use Mood's median test to compare fish weight vs. water temperature. The test results demonstrate that water temperature has a significant effect on the growth rate of fish. The Low and Med_High water temperatures show that growth rates lag those of Low_Med and Med water temperatures with a p-value of 0.001.

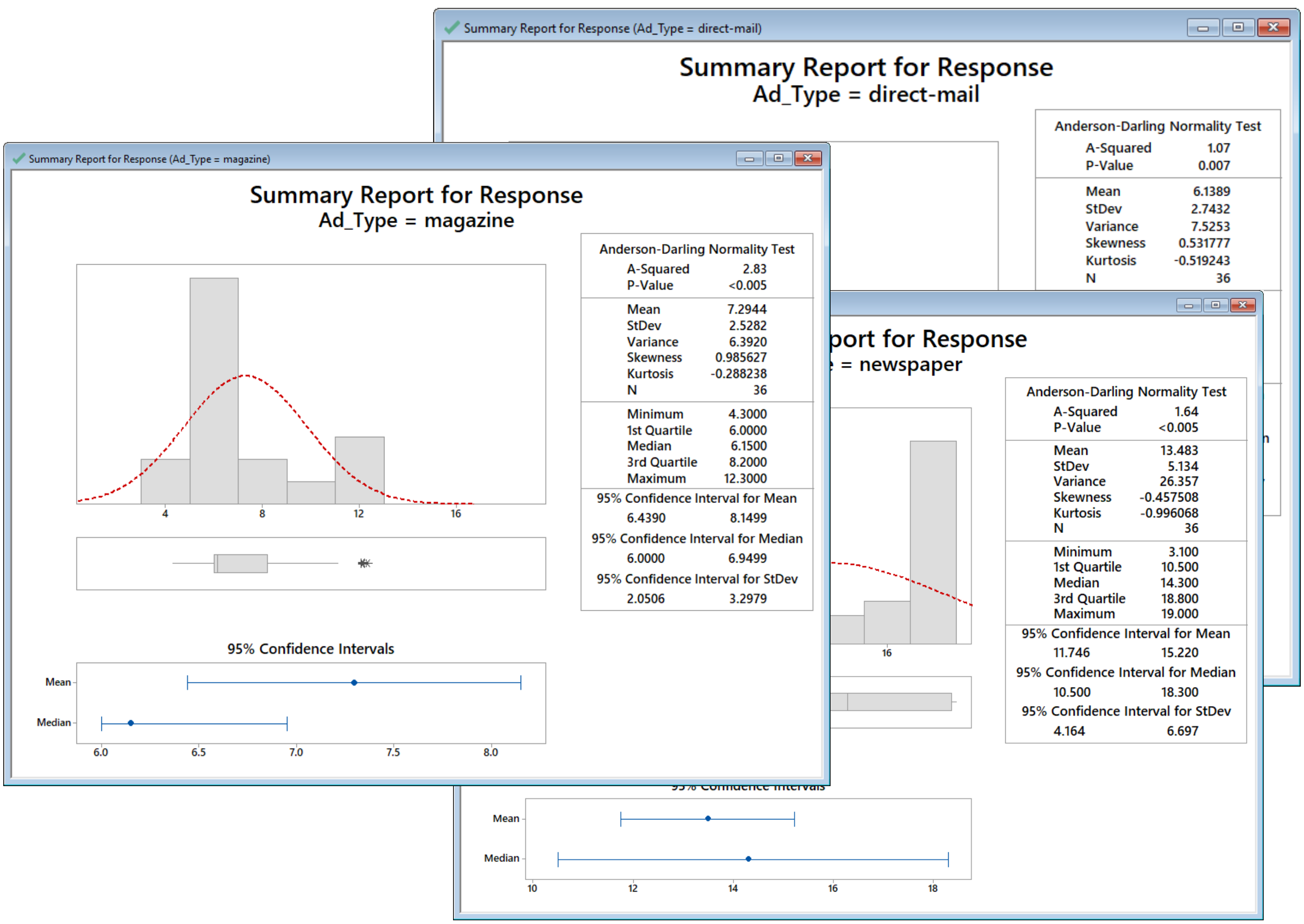
Friedman Test Exercise – Advertising Response Rates

Data File: Friedman.MTW

A marketing analyst wants to compare the relative effectiveness of three types of advertising: direct mail, newspaper, and magazine. The analyst performs a randomized block experiment. For 36 clients, the marketing firms used all three types of advertising over one year and recorded the year's percentage response to each type of advertising. Use the "Friedman.MTW" data file to determine if there is a difference in response rates between advertising methods.

Solution Steps Screencast: https://youtu.be/31DpV8izVkk

Solution Output Screenshots:



riedman Test: Response vs Ad_

Method

Treatment = Ad_Type
Block = Company

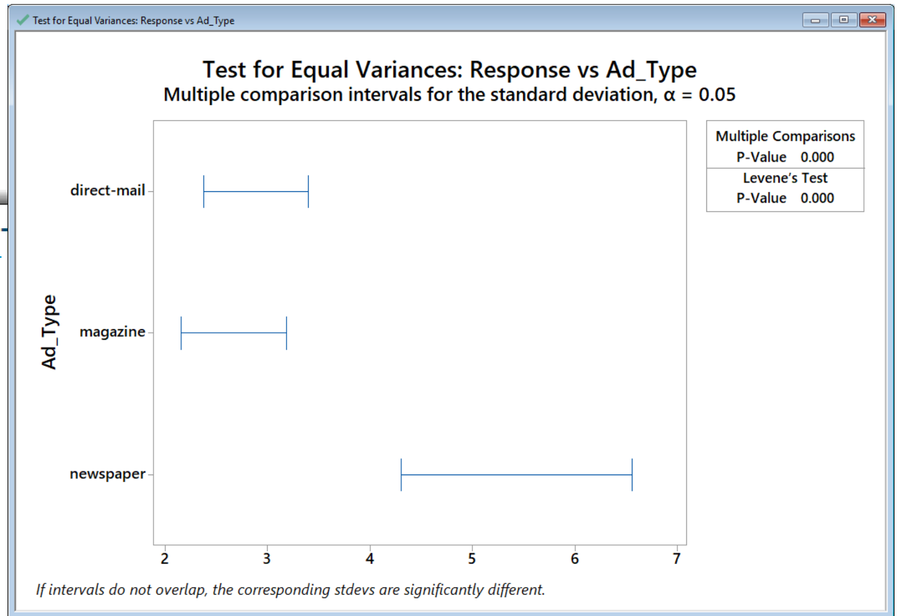
Descriptive Statistics

Ad_Type	N	Median	Sum of Ranks
direct-mail	36	5.9000	58.0
magazine	36	7.1667	64.0
newspaper	36	13.5333	94.0
Overall	108	8.8667	

Test

Null hypothesis H_0: All treatment effects are zero
Alternative hypothesis H_1: Not all treatment effects are zero

Method	DF	Chi-Square	P-Value
Not adjusted for ties	2	20.67	0.000
Adjusted for ties	2	20.96	0.000



Friedman Test Result Interpretation

We commonly use the Friedman test when our data are non-normal and when comparing the differences between the medians of various groups across multiple measures. Here, the data are non-normal, and variances are not equal. The results show that the differences in response rates between the three advertising types are significant, with newspaper having the highest median response rate.

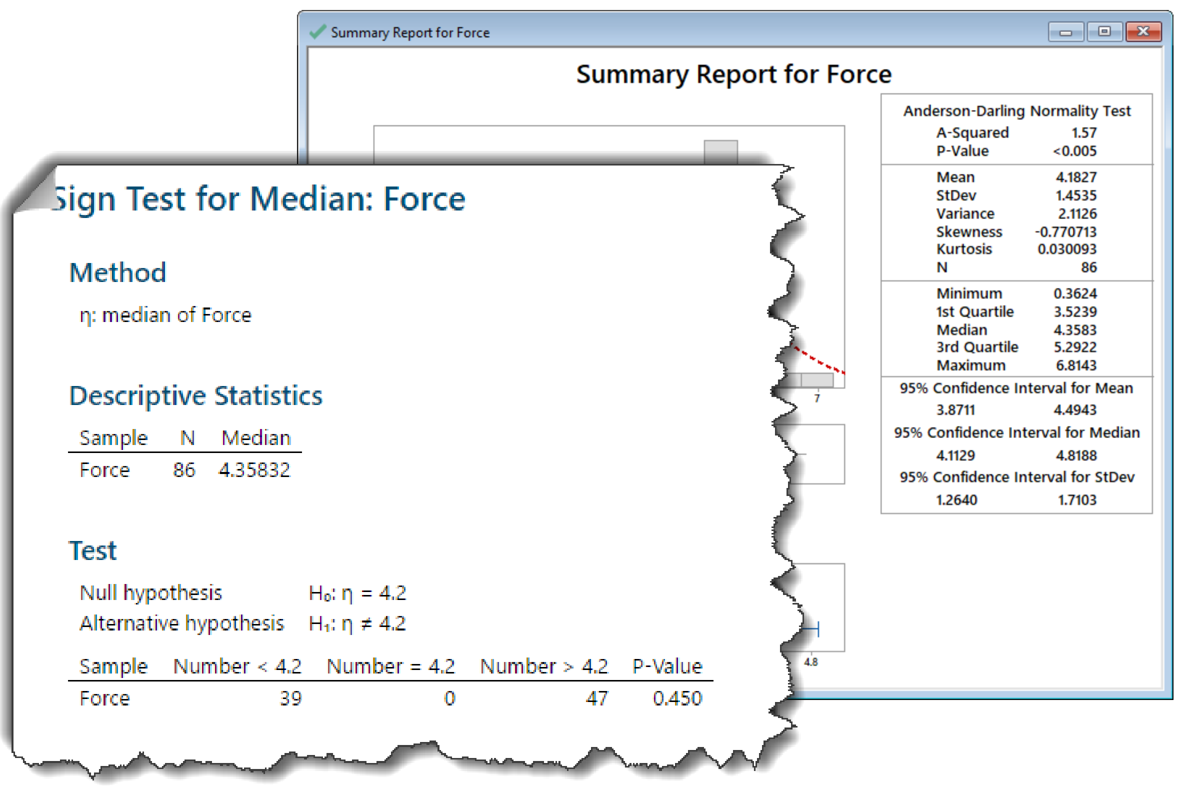
1-Sample Sign Test Exercise – Snack Bag Opening Force

Data File: 1-Sample_Sign.MTW

A packaging engineer wants to test a new method to seal snack bags. The target force that is expected to open the bags should be 4.2 N (Newtons). The engineer randomly samples 86 bags that are sealed using the new method and records the force that is required to open each bag. Use the "1-Sample_Sign.MTW" data file to determine if the new sealing method meets the target opening force.

Solution Steps Screencast: https://youtu.be/LQ0mi4vKC30

Solution Output Screenshots:



Summary Report for Force

Anderson-Darling Normality Test

A-Squared	1.57
P-Value	<0.005
Mean	4.1827
StDev	1.4535
Variance	2.1126
Skewness	-0.770713
Kurtosis	0.030093
N	86
Minimum	0.3624
1st Quartile	3.5239
Median	4.3583
3rd Quartile	5.2922
Maximum	6.8143

95% Confidence Interval for Mean: 3.8711 4.4943

95% Confidence Interval for Median: 4.1129 4.8188

95% Confidence Interval for StDev: 1.2640 1.7103

Sign Test for Median: Force

Method

η: median of Force

Descriptive Statistics

Sample	N	Median
Force	86	4.35832

Test

Null hypothesis H₀: η = 4.2
Alternative hypothesis H₁: η ≠ 4.2

Sample	Number < 4.2	Number = 4.2	Number > 4.2	P-Value
Force	39	0	47	0.450

1-Sample Sign Test Result Interpretation

The data are not normal; otherwise, we would have elected to use the 1-sample t-test. The result of the 1-sample sign test demonstrates that that we cannot reject the null hypothesis, which states that the median opening force is 4.2 N. Therefore, we must conclude that the median opening force is not different from 4.2 N.

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

1-Sample Wilcoxon Test Exercise – Antacid Reaction Time

Data File: 1-Sample_Wilcoxon.MTW

A chemist for a pharmaceutical company wants to determine whether the median reaction time for a newly developed antacid is less than 12 minutes. The chemist measures the reaction time for 16 samples of the antacid. Use the "1-sample_Wilcoxon.MTW" data file to test whether the median reaction time is less than 12 minutes.

Solution Steps Screencast: https://youtu.be/ovyrMHu9k14

Solution Output Screenshots:

Wilcoxon Signed Rank Test: Time

Method

η: median of Time

Descriptive Statistics

Sample	N	Median
Time	16	11.55

Test

Null hypothesis	H_0: η = 12
Alternative hypothesis	H_1: η < 12

Sample	N for Test	Wilcoxon Statistic	P-Value
Time	16	53.00	0.227

1-Sample Wilcoxon Test Result Interpretation

Based on the test results with a p-value of 0.227, which is greater than the alpha value of 0.05, we fail to reject the null hypothesis. The null statement in this case is that the median reaction time is 12 minutes (indicated by H_0: n=12). By failing to reject the null hypothesis, we cannot conclude that the reaction time is less than 12 minutes.

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

1 Sample Proportion Test Exercise – Satisfaction Rates

Data File: N/A

A training company provides classroom and online instruction to hundreds of people per month. The company conducts a satisfaction survey at the end of each course. The training company believes there might be a change in its satisfaction rates. During the prior calendar year, the satisfaction rate was 94.2%. Over the past six months, the number of satisfied responses was 824 out of 892 surveys. Use the information provided to perform a 1-sample proportion test to determine if there is a difference between the current satisfaction rate and that of the prior year.

Solution Steps Screencast: https://youtu.be/64GApegc42Q

Solution Output Screenshots:

st and CI for One Proportion

Method

p: event proportion
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
892	824	0.923767	(0.904354, 0.940319)

Test

Null hypothesis H_0: $p = 0.942$
Alternative hypothesis H_1: $p \neq 0.942$

P-Value
0.026

1-Sample Proportion Test Result Interpretation

The results of the 1-sample proportion test indicate that there is a difference between the last six months of satisfaction ratings and the prior calendar year. The p-value of 0.026 is less than 0.05; therefore, we reject H_0, which states there is no difference, and conclude there is a difference.

LEAN SIGMA
CORPORATION

2-Sample Proportion Test Exercise – Mortgage Defect Rates

Data File: N/A

A mortgage company performs reviews of its closing documents prior to the closing dates. This review may prompt changes or document additions that should have been in the closing package. Any closing package found with inaccuracies or known to be incomplete is deemed defective and requires rework prior to closing. In the prior calendar year, the defective rate for these reviews was 17.99% (783 defective packages out of 4,352 reviewed). During the current year, the year-to-date defective rate is 21.15% (180 of 870). Use the information provided to perform a 2-sample proportion test to determine if there is a difference between the current satisfaction rate and that of the prior year.

Solution Steps Screencast: https://youtu.be/nJkCCR22nxk

Solution Output Screenshots:

Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	4352	783	0.179917
Sample 2	870	180	0.206897

Test

Null hypothesis	H_0: $p_1 - p_2 = 0$
Alternative hypothesis	H_1: $p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Normal approximation	-1.81	0.071
Fisher's exact		0.062

2-Sample Proportion Test Result Interpretation

The results of the 2-sample proportion test indicate that there is no difference between the year-to-date defective rate and that of the prior calendar year.

The p-value of 0.062 is greater than 0.05; therefore, we fail to reject H_0, which states there is no difference, and conclude there is no difference.

LEAN SIGMA
CORPORATION

Improve Phase Exercises

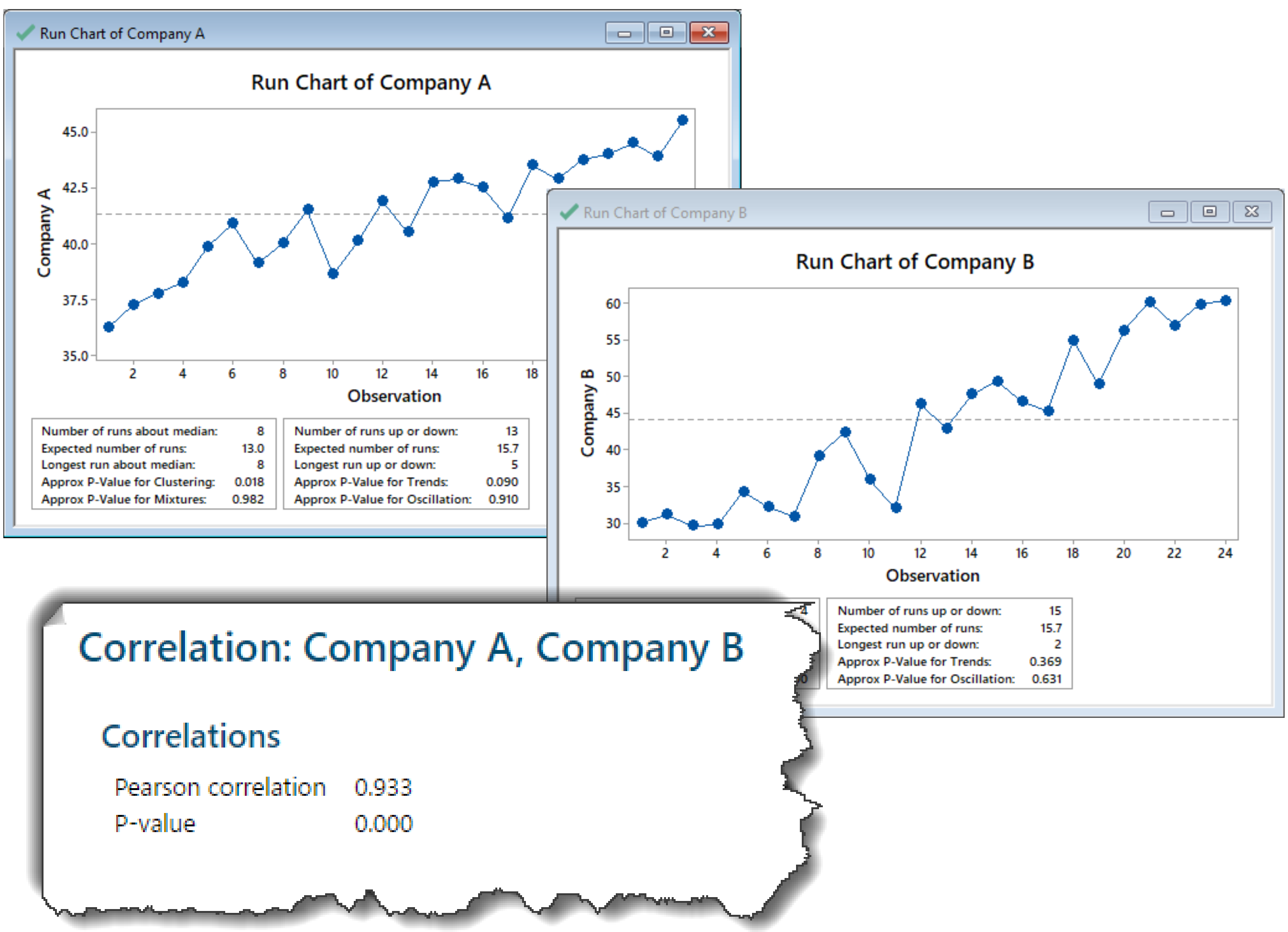
Correlation Exercise – Stock Price

Data File: Correlation.MTW

The stock broker reviewed the run charts pictured below of Company A and Company B, and based on the charts, now wants to quantify the strength of the linear relationship between the two companies. Use the “Correlation.MTW” data set to perform a correlation analysis between the two companies.

Solution Steps Screencast: https://youtu.be/f0eT_jUgJlg

Solution Output Screenshots:

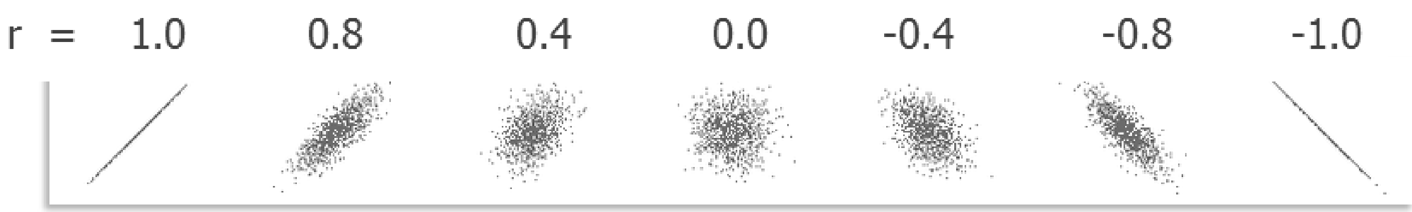


Attribution: A portion of the following exercises and data sets have been adapted from Minitab’s “Data Set Library” at https://support.minitab.com/en-us/datasets/

Result Interpretation

About Pearson's Correlation Coefficient

Pearson's correlation coefficient is a measure of the strength of a linear relationship. Its value will be between -1 (perfect inverse linear relationship) and +1 (perfect direct linear relationship). A value of zero indicates no linear relationship. Below are a few examples of scatterplots relative to a range of correlation coefficients between 1 and -1:



If the value of Pearson's correlation coefficient is zero, it does not mean that there is no relationship, only that there is no linear relationship. There may be a non-linear relationship that cannot be determined with Pearson's correlation coefficient. Below are few examples of non-linear relationships where Pearson's correlation coefficient is zero:

Conclusion

The correlation coefficient measuring the linear relationship between the stock prices of Company A and Company B is 0.933. This suggests that there is a strong positive linear relationship. When Company A's stock price moves up or down, Company B's stock price will tend to do the same. This relationship does not indicate any causal relationship, only that the two stock prices have a positive linear relationship.

Simple Linear Regression Exercise – Stock Price vs. 10 Yr. T

Data File: Regression.MTW

The stock broker now wants to determine if the driving factor for the stock price increase of Company A is due to the recent improvements in the 10 Yr. Treasury Note. Perform a simple linear regression between Company A and the 10 Yr. T (columns C1 and C3). Use the Minitab data file "Regression.MTW."

Solution Steps Screencast: https://youtu.be/ps5zbzH2asw

Solution Output Screenshots:

egression Analysis: Company A versus 10-Yr T

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	99.236	99.236	49.89	0.000
10-Yr T	1	99.236	99.236	49.89	0.000
Error	22	43.763	1.989		
Lack-of-Fit	18	33.874	1.882	0.76	0.697
Pure Error	4	9.889	2.472		
Total	23	142.999			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.41040	69.40%	68.01%	64.07%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	23.19	2.57	9.03	0.000	
10-Yr T	6.819	0.965	7.06	0.000	1.00

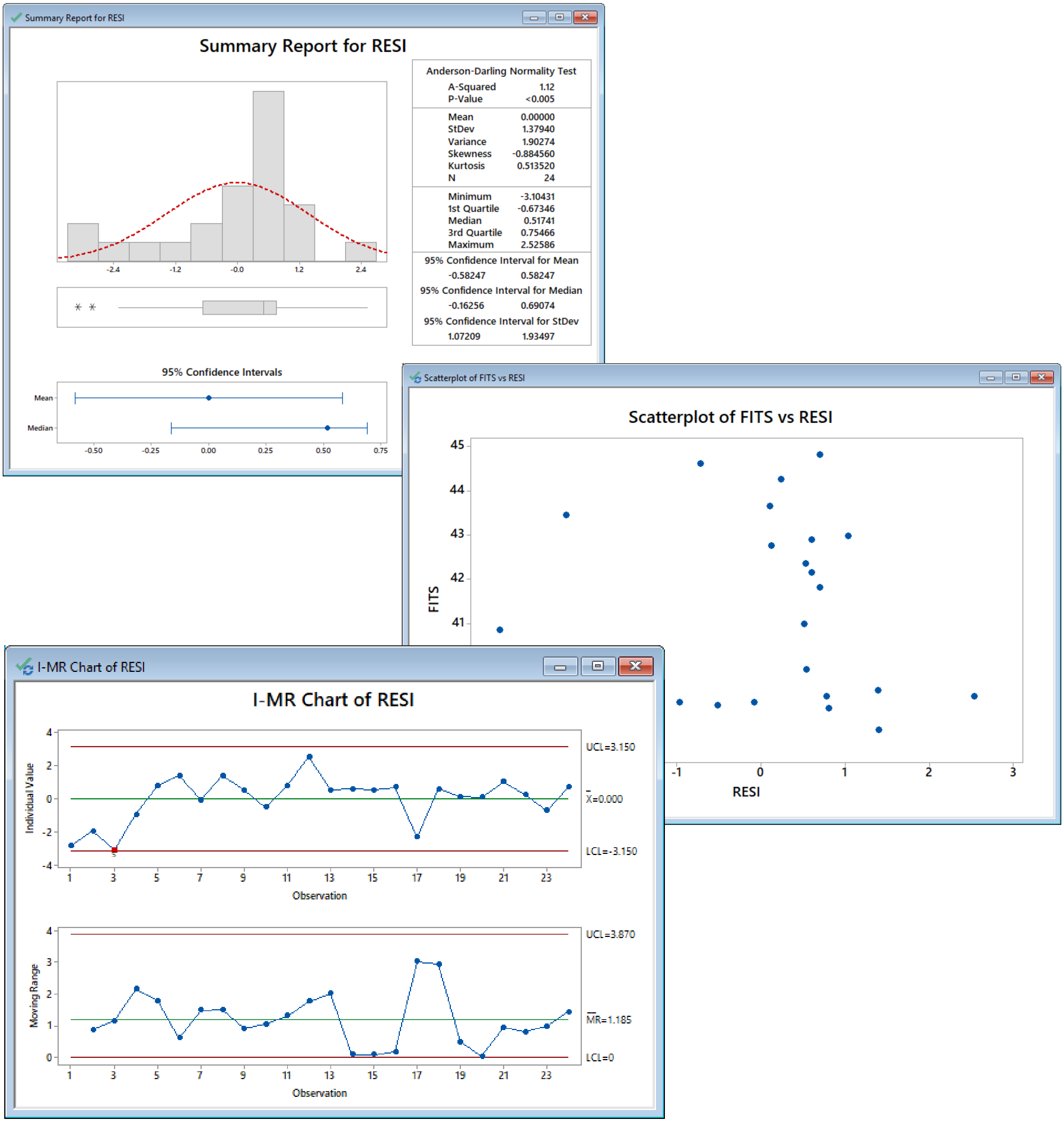
Regression Equation

Company A = 23.19 + 6.819 10-Yr T

Fits and Diagnostics for Unusual Observations

Obs	Company A	Fit	Resid	Std Resid	
1	36.250	39.081	-2.831	-2.10	R
3	37.750	40.854	-3.104	-2.25	R

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Simple Linear Regression Result Interpretation

About Simple Linear Regression

The result of the simple linear regression provides us with a significant p-value for our predictor (10 Yr. T), but the R-Sq (adjusted) is 68.01%, which is not bad but doesn't completely explain the movement in Company A's stock price. What the R-Sq (adj) tells us is that 68.01% of the variation in Company A's stock price can be explained by the 10 Yr. Treasury note.

Residual Analysis Normality

As with any regression analysis, we must pass a few assumptions to validate the model so that we can trust the data. There are a few further tests we must perform to accomplish this. When performing the regression analysis, we elected to store the residuals and fitted data in our worksheet so that we could determine if the residuals are normally distributed. Therefore, we perform a graphical analysis to determine normality. In this case, the residuals are not normally distributed. This suggests that we may need to use a better predictor and we could be missing an important predictor variable.

Residual Analysis Independence

By performing an I-MR control chart, we can determine if the residuals are independent. If the control chart is out of control we can conclude that the residuals are not independent and that an important factor may be missing from the model.

Residual Analysis Heteroscedasticity

Heteroscedasticity is the condition where the assumption of equal variance is violated and can lead us to believe a variable is a predictor when it is not. Therefore, another assessment of residuals is determining if there is equal variance. We perform a scatter plot to look for a random pattern in which residuals spread out randomly with a mean around zero. In this example, the residual values disperse randomly around with a mean of zero (also verified by the graphical summary output).

Conclusion

Although on the surface the model looks decent with an R-Sq (adj) of 68.01% and a predictor p-value of 0.000, when we evaluate the residuals we can see that the model is lacking an important predictor or just can't be reliably explained by the current predictor.

Multiple Linear Regression Exercise – Stock Price

Data File: Regression.MTW

The stock broker now wants to include additional predictor variables to determine the driving factors for the stock price increase of Company A. Perform a multiple linear regression analysis between Company A and the 10 Yr. T, GDP, and Unemployment (columns C1, C3, C4, and C5). Use the Minitab data file "Regression.MTW."

Solution Steps Screencast: https://youtu.be/RoTur28bW40

Solution Output Screenshots:

egression Analysis: Company A versus 10-Yr T, GDP ... ne

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	127.111	42.3702	53.34	0.000
10-Yr T	1	1.041	1.0406	1.31	0.266
GDP ($millions)	1	0.429	0.4291	0.54	0.471
Unemployment	1	2.645	2.6454	3.33	0.083
Error	20	15.888	0.7944		
Total	23	142.999			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.891291	88.89%	87.22%	84.06%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	50.3	10.1	4.97	0.000	
10-Yr T	1.70	1.48	1.14	0.266	5.90
GDP ($millions)	0.000000	0.000000	0.73	0.471	25.10
Unemployment	-378	207	-1.82	0.083	13.53

Regression Equation

Company A = 50.3 + 1.70 10-Yr T + 0.000000 GDP ($millions) - 378 Unemployment

Fits and Diagnostics for Unusual Observations

Obs	Company A	Fit	Resid	Std Resid	
10	38.630	40.345	-1.715	-2.07	R

Regression Analysis: Company A versus 10-Yr

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	126.682	63.3408	81.52	0.000
10-Yr T	1	5.629	5.6287	7.24	0.014
Unemployment	1	27.446	27.4460	35.32	0.000
Error	21	16.317	0.7770		
Total	23	142.999			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.881477	88.59%	87.50%	85.30%

Coefficients

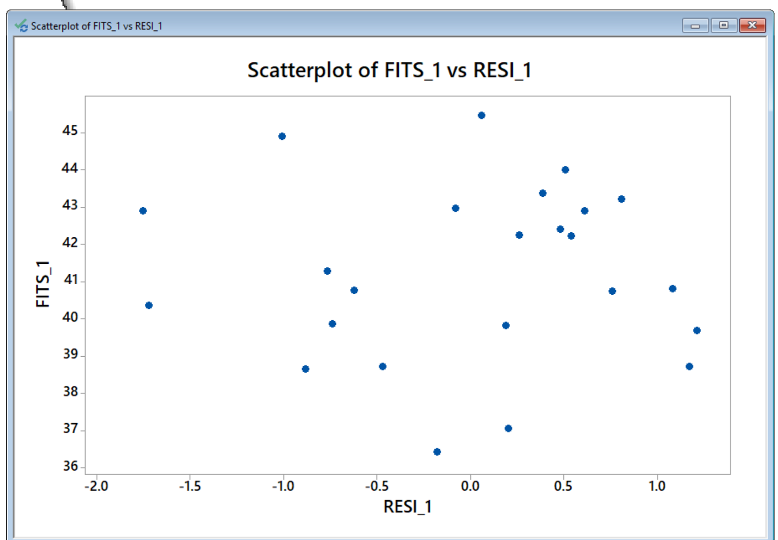
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	56.33	5.80	9.71	0.000	
10-Yr T	2.531	0.940	2.69	0.014	2.43
Unemployment	-516.0	86.8	-5.94	0.000	2.43

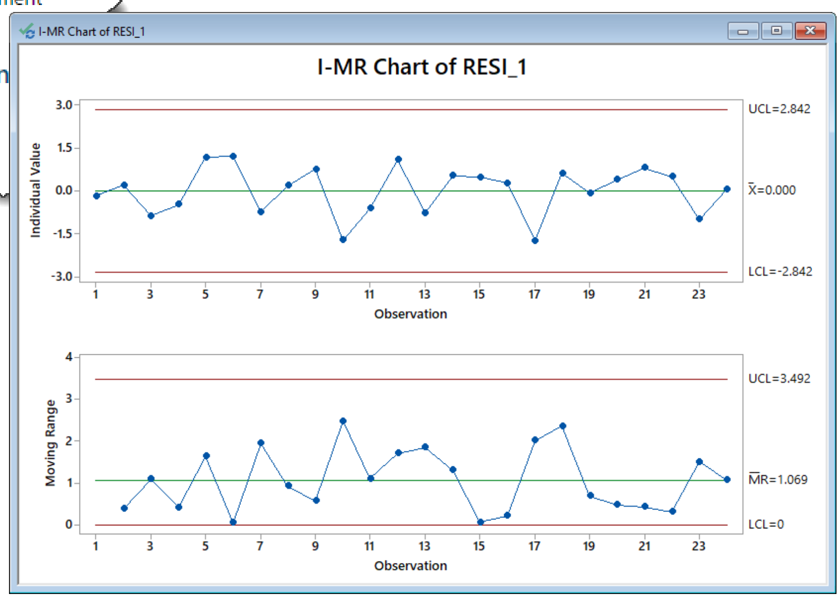
Regression Equation

Company A = 56.33 + 2.531 10-Yr T - 516.0 Unemployment

Fits and Diagnostics for Unusual Observation

Obs	Company A	Fit	Resid	Std Resid	
10	38.630	40.354	-1.724	-2.11	R
17	41.130	42.886	-1.756	-2.10	R





Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

Multiple Linear Regression Result Interpretation

About Multiple Linear Regression

The first result of the multiple linear regression with three predictor variables (10 Yr. T, GDP, and Unemployment) provides us with no significant p-values for any predictor but the R-Sq (adjusted), which is 87.22%, and we already know from our simple linear regression that the 10 Yr. T is significant. Therefore, we must look further into the results, which leads us to considering multicollinearity.

Multicollinearity

Multicollinearity is the situation when two or more independent variables in a multiple regression model are correlated with each other. Although multicollinearity does not necessarily reduce the predictability for the model, it may mislead the calculation for individual independent variables. To detect multicollinearity, we use the VIF (Variance Inflation Factor) to quantify its severity in the model. In reviewing the VIF values, we can see that all variables have a value greater than 5. When this occurs, we begin reducing the model by removing the variable with the highest VIF and re-running the model.

After removing GDP from the model, we now have what looks to be a very strong regression model with two statistically significant factors, high R-Sq (adj), at 87.5%, and low VIF. These are strong results. The next step is to validate our assumption about our residuals, which must be normally distributed, independent, with equal variances across our fitted values.

Residual Analysis Normality

When we performed our first regression, we stored both fits and residuals. They were stored in C6 and C7 and labeled FITS and RESI, respectively. After running our second analysis, Minitab again stored the fits and residuals but appended a value of _1 to each to avoid creating columns with the same name. These values were stored in columns C8 and C9 and were labeled FITS_1 and RESI_1, respectively. The second set of stored residuals is the set we now want to perform analysis on. These are the values relative to our second regression model after removing GDP.

Upon running the graphical analysis, we can see that the residuals are normally distributed with a mean of 0.00. This is a positive sign.

Residual Analysis Independence

By performing an I-MR control chart on RESI_1, we can determine if the residuals are independent. In this case, the control chart demonstrates no out-of-control conditions, and we can conclude that the residuals are independent. This is also a positive sign.

Residual Analysis Heteroscedasticity

Lastly, we perform a scatterplot of FITS_1 vs. RESI_1 to evaluate the possible condition of heteroscedasticity, which is the condition where the assumption of equal variance is violated and can lead us to believe a variable is a predictor when it is not. In this example, the residual values disperse randomly with a mean of zero (also verified by the graphical summary output).

Conclusion

After validating all residual assumptions, we can conclude that the regression model is a significant model (p-value 0.000) with two significant predictor variables and an R-Sq (adj) value of 87.5%. This is a sound and reliable predictor model.

What we have learned through this process is that 87.5% of the variation in Company A's stock price can be attributed to the 10 Yr. Treasury note and the Unemployment rate. The equation derived from this model is

Company A's Stock Price = 56.33 + (2.531*(10 Yr. T)) – (516*(Unemployment)).

Word of Caution

No model is perfect! But good models provide insight. Remember, the R-Sq (adj) is 87.5%, and there is still a fair amount of variation in the stock price that is not explained by the model.

Control Phase Exercises

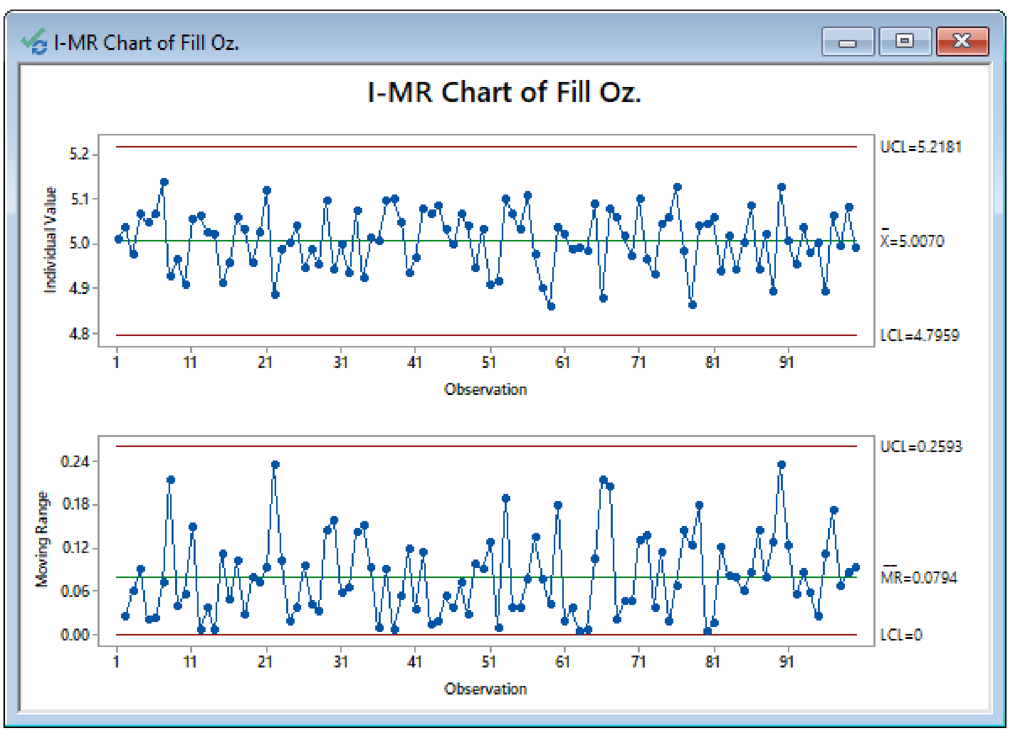
I-MR Chart Exercise – Bottle Fill

Data File: [SPC.MTW](SPC.MTW)

A small and fairly new pharmaceutical company has an over-the-counter liquid drug product that is packaged in bottles intended to contain 5 oz of medication. As the quality control engineer, you are tasked with monitoring the fill quantities using statistical process control. Doing so enables the company to make production and process adjustments if fill quantities go out of control. Use the "SPC.MTW" data file to run an I-MR chart on the 100 oz samples found in column C1 and determine if any process changes are necessary.

Solution Steps Screencast: https://youtu.be/DMJkWsX_lwM

Solution Output Screenshots:



Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

I-MR Chart Result Interpretation

The I-MR chart looks very good, with both the range chart and the individuals chart in control. All eight tests were performed with 100 data samples and a moving range of 2. At the current time, there should be no reason for the company to make process adjustments assuming the fill quantities are within specification.

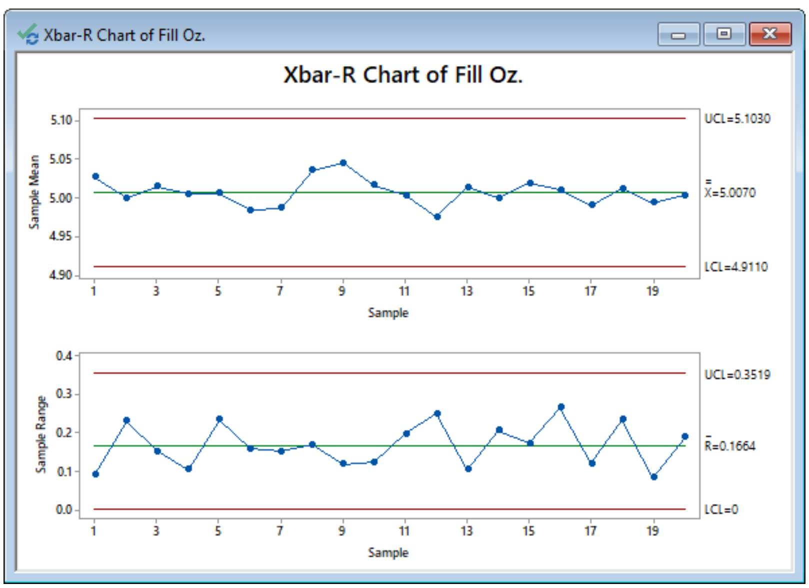
XbarR Chart Exercise – Bottle Fill

Data File: [SPC.MTW]

A small and fairly new pharmaceutical company has an over-the-counter liquid drug product that has five production lines packaging bottles intended to contain 5 oz of medication. As the quality control engineer, you are tasked with monitoring the fill quantities across production lines using statistical process control. In doing so, you will be able to alert the company to make production or process adjustments if fill quantities go out of control. Use the "SPC.MTW" data file to run an XbarR chart on the 100 oz samples, with the subgroups being the production lines. Determine if any process changes are necessary.

Solution Steps Screencast: https://youtu.be/smLIA1MJTf8

Solution Output Screenshots:



XbarR Chart Result Interpretation

The XbarR is used in this case because we are looking at the average and range of samples for each production line, with subgroup samples size less than 10. This XbarR chart looks very good, with both the range and sample mean charts in control. All eight tests were performed with 100 data samples over 20 subgroups. At the current time, there should be no reason for the company to make process adjustments assuming the fill quantities are within specification.

LEAN SIGMA
CORPORATION

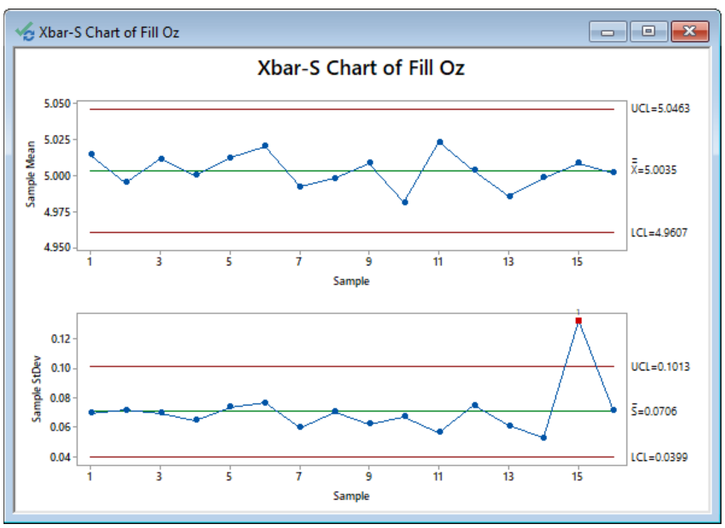
XbarS Chart Exercise – Bottle Fill

Data File: [XbarS.MTW]

A small and fairly new pharmaceutical company has an over-the-counter liquid drug product that has two shifts of production packaging bottles intended to contain 5 oz of medication. As the quality control engineer, you are tasked with monitoring the fill quantities within and between shifts using statistical process control. You have collected 25 samples per shift over the course of eight business days. Use the "XbarS.MTW" data file to run an XbarS chart on the 400 oz samples, with the subgroups being production shift. Determine if any out-of-control conditions exist.

Solution Steps Screencast: https://youtu.be/Buy91W9R_tg

Solution Output Screenshots:



XbarS Chart Result Interpretation

The XbarS chart is used in this case because we are looking at the average and standard deviation of samples for each production shift, with subgroup sample size greater than 10 (in this case 25). The results show the XbarS chart with an out-of-control condition in the S chart, suggesting an anomaly within a subgroup. Upon further research, you determine that it's the first shift on the eighth day and did not occur previously. You now need to investigate what happened and if the first-shift anomaly is a procedural or production issue and seek to ensure that the conditions that led to this special cause can be mitigated.

Attribution: A portion of the following exercises and data sets have been adapted from Minitab's "Data Set Library" at https://support.minitab.com/en-us/datasets/

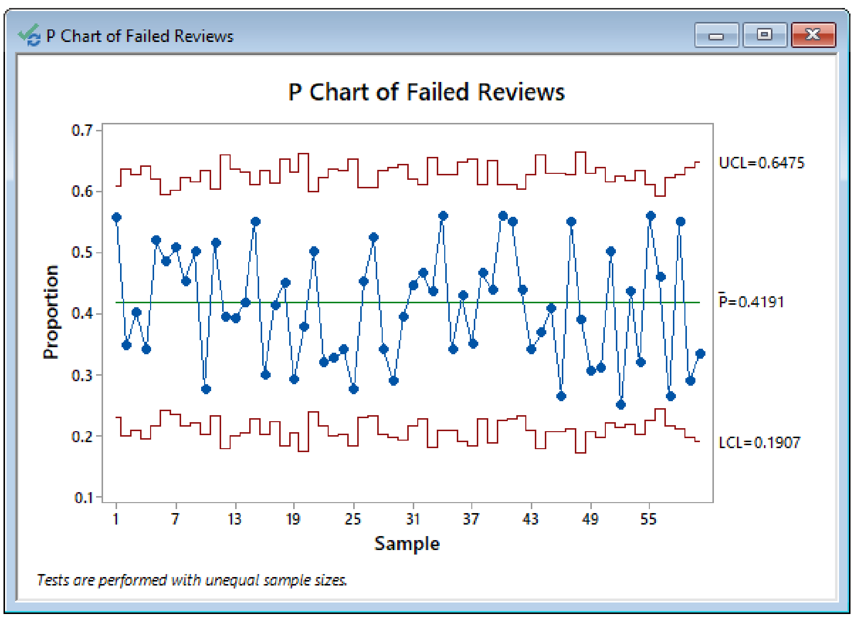
LEAN SIGMA CORPORATION

P Chart Exercise – Mortgage Closing Documents

Data File: P.MTW
A mortgage company performs daily reviews of its closing documents prior to each closing date. This review may prompt changes or document additions that should have been in the closing package. Any closing package found with inaccuracies or known to be incomplete fails the quality control check and requires rework prior to closing. The company is working to minimize closing document errors and has begun tracking their daily performance. Use the “P.MTW” file to run a P chart on the daily closing document failure rate.

Solution Steps Screencast: https://youtu.be/E4PCwPaTdW0

Solution Output Screenshots:



P Chart Result Interpretation

The P chart is a control chart monitoring the percentages of defectives. The P chart plots the percentage of defectives in one subgroup as a data point. It considers the situation when the subgroup size of inspected units is not constant. Since the sample sizes are not constant over time, the control limits are adjusted to different values accordingly. All the data points fall within the control limits and spread randomly around the mean. Although the failure rate is terrible, we conclude that the process is stable and in control.

Attribution: A portion of the following exercises and data sets have been adapted from Minitab’s “Data Set Library” at https://support.minitab.com/en-us/datasets/