Lot Skipping Report

1 Reviewing the Old Presentation

Johnston wire came up with a method where they began skipping tests on lots. The old presentation aimed to correlate the data between Johnston wire and PLP, so PLP could implement this new solution.

The presentation took the minimum breaking load and minimum tensile strength [Figure 1] for both companies for the same products. If the data values were similar enough then Johnston Wire's testing method could be correlated and ideally PLP would be able to implement skip lot testing the same way Johnston Wire has.

Johnston Wire Certification Validation

Wire Diameter (inches)	Minimum Breaking Load (lbs.)	Minimum Tensile Strength (psi)	Maximum Tensile Strength (psi)
0.159 PLP Testing Threshold	3,595 (≤3,630)	193,000 (≤195,000)	223,000 (≥221,000)
0.119 PLP Testing Threshold	2,100 (≤2,120)	202,000 (≤204,000)	232,000 (≥230,000)
0.138 PLP Testing Threshold	2,720 (≤2,750)	193,000 (≤195,000)	223,000 (≥221,000)
0.241 PLP Testing Threshold	7,500 (≤7,580)	170,000 (≤172,000)	200,000 (≥198,000)
0.100 PLP Testing Threshold	1,570 (≤1,590)	217,000 (≤219,000)	242,000 (≥240,000)
0.128 PLP Testing Threshold		180,000 (≤182,000)	
0.188 PLP Testing Threshold	4,815 (≤4,860)	183,000 (≤185,000)	213,000 (≥211,000)

Figure 1

The presentation attempted to correlate the data by using a Null-Hypotheses test. A Null hypothesis test is a test that proposes there is no difference between certain characteristics of a population. In this scenario, the Null hypothesis states that the mean of the two sampling distributions is zero. In other words, taking the mean of the breaking load of all the different PLP wire diameters and the mean of all the Johnston wire diameters, the end result would be statistically negligible, and thus the two sets of data could be effectively treated as belonging to the same population.

1.1 Probability Plotting

The presentation began by showing a probability plot of the break loads of Johnston wire and PLP [Figure 2]. These two graphs are just to show that both sets of data are normally distributed and that we can apply the Null-Hypothesis test (a Null-hypothesis can only be applied to normal data). A set of data will follow the normal distribution, if the probability plot of the break load is linear. As seen in the figure below, both plots are fairly linear and can be interpreted to be normal.

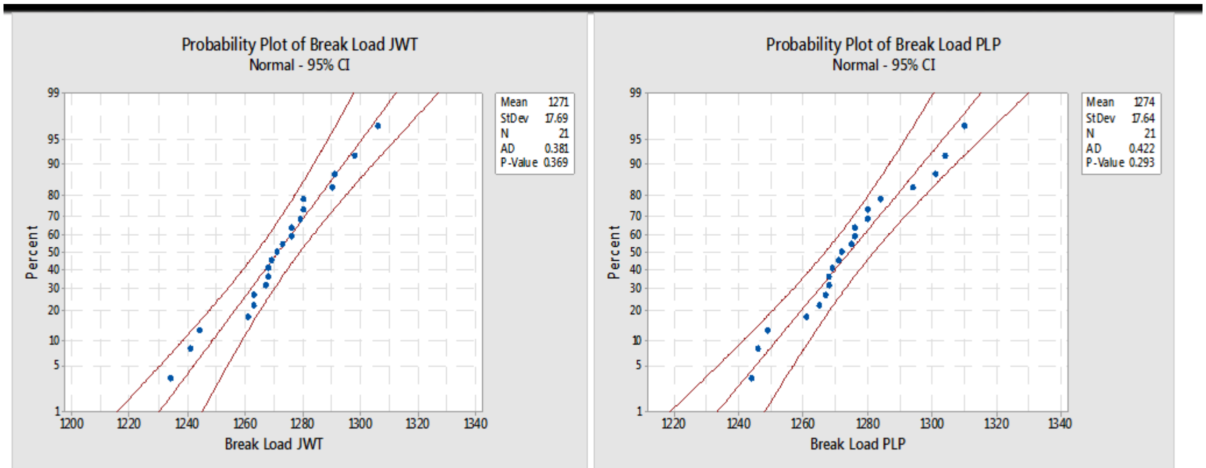


Figure 2

The percentage that is being plotted against in the graph above is the confidence interval. A confidence interval is the percentage you are certain the interval will contain the true mean of the data. So, in the confidence interval above it is saying you can be 95 percent certainty the range of values above will contain the true mean. A 95 percent confidence interval is standard when applying probability analysis/calculations. In the context of the graph above the confidence interval is irrelevant, it could have just been plotted against the regular normal distribution. See figure 3 for an example of a confidence interval on a regular normal distribution.

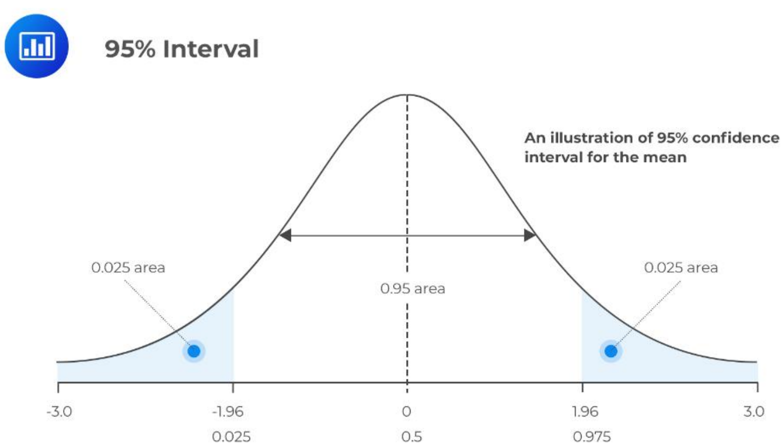


Figure 3

1.2 Null Hypothesis Testing

Now that the data has been established as normal, the presentation then moves on to the Null-Hypothesis test. In this scenario they used a two sample T-test. A two-sample t-test is used to test the difference between two population means. The most common application is to determine whether the means are equal. The data they layout in the presentation can be seen in figure 4.

Two-Sample T-Test and CI: Break Load JWT, Break Load PLP

Method

μ_1: mean of Break Load JWT
μ_2: mean of Break Load PLP
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Break Load JWT	21	1271.3	17.7	3.9
Break Load PLP	21	1274.3	17.6	3.8

Estimation for Difference

Difference	95% CI for Difference
-2.95	(-13.98, 8.07)

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
-0.54	39	0.591

Individual Value Plot of Break Load JWT, Break Load PLP

Boxplot of Break Load JWT, Break Load PLP

Figure 4

1.3 T-Values

There is a lot of background information that the presentation omits on this slide. So, I will do my best to describe it here. When performing a two sample T-test, you need to find a t-value and using the t-value you need to find a p-value. A t-value measures the size of the difference relative to the variance in your sample data. In most cases it is just a tool to find the P-value which we are actually concerned about. The t-value is calculated with the formula below in figure 5. X1 and x2 are the sample means, s1 and s2 are the standard deviations and n1 and n2 are the population sizes.

$$t = \frac{(x_1 - x_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

Figure 5

1.4 P-Values Background

The P-value is the probability of obtaining test results at least as extreme as the results observed.

A good example to illustrate this is if you have a friend who claims to have an 80% free throw rate in basketball. You are doubtful of him, so you tell him to try 10 free throws in front of you to prove it. Your friend hits 7 out of 10. This is less than the 80 percent originally claimed but it's close enough to not call your friend a liar. This 7 out of 10 free throw rate has a P-value of 0.32. This means that if your friend is telling the truth about hitting 80% of free throws, he would 32% of the time, hit 7 or fewer free throws in 10 attempts. This is fairly likely in the world of P testing. Now let's say for he only hits one out of ten baskets. It is still possible for him to have an 80 % free throw rate, but the odds are so low that you can confidently call your friend a liar. This has a P-value of 0.000004, so if your friend is telling the truth he would land one free throw out of ten 0.0004% of the time, making it near impossible.

I hope this example does a good job helping visualize what a p-value is. I found it's difficult to just look at it on a normal distribution and understand what it is representing.

1.5 P-Value Fallacies

Note a common fallacy when using P-values and the biggest issue that this presentation has, is the assumption that you can accept a Null Hypothesis. P-values are based on the outcome that your initial hypothesis is true. So, we can only reject or fail to reject a Null Hypothesis based on P-values alone. We can only say based on our test statistics and the assumption that our initial condition is true, that there is a certain chance of obtaining results as likely as our initial experiment. Figure 6 below shows how a P-value would look on the normal distribution

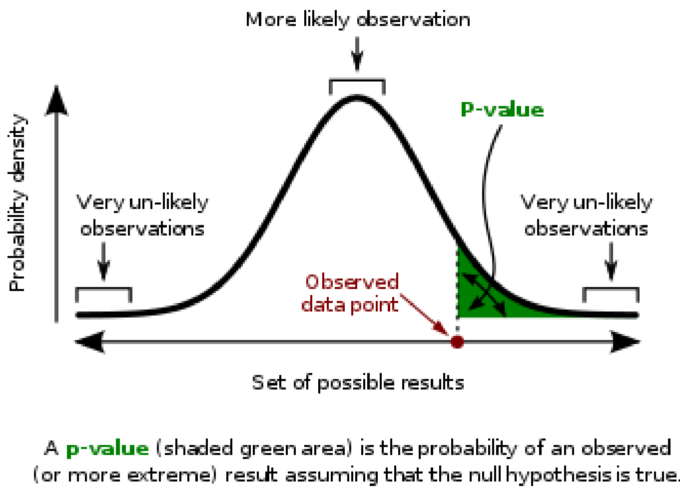


Figure 6

1.6 Presentation P-Values Analysis

When looking to reject a null hypothesis, it is common practice to use a significance level of 0.05. This means that if the P-value is below 0.05 then we can reject the null and if it is above 0.05, we fail to reject the null hypothesis. P-values are found by usually looking it up in a table, or in this case using an excel formula.

In the presentation I believe they just used the excel data analysis formula with the two tailed t-test assuming unequal variances [Figure 7]. I used the excel formula with their numbers and came to the same result as they did on the slide.

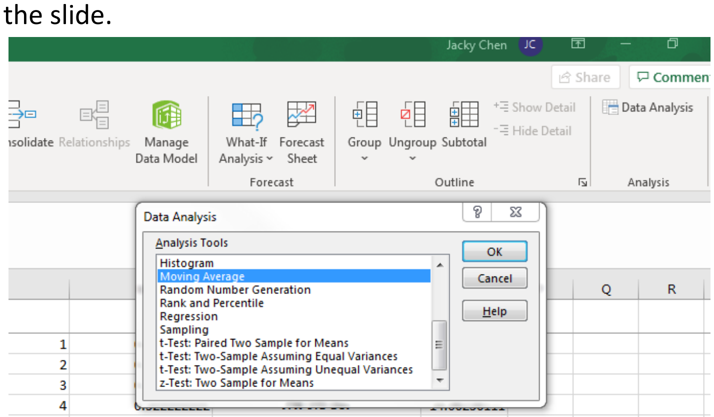


Figure 7

They came to a P-value of 0.591, which means that we cannot reject the null hypothesis, but says nothing about whether or not our null hypothesis is correct.

As an aside, a good resource to use when knowing how to use P-values and how to contextualize them is the American statistical association. They have put out six principles that should be adhered to when using P-values.

The statement's six principles, many of which address misconceptions and misuse of the *p*-value, are the following:

1. *P-values can indicate how incompatible the data are with a specified statistical model.*
2. *P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.*
3. *Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.*
4. *Proper inference requires full reporting and transparency.*
5. *A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.*
6. *By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.*

1.7 Presentation Issues

Obviously, the biggest problem is the assumption that the null hypothesis of both means being equal is true. All we know from the current analysis is that we cannot reject the hypothesis of it being untrue based on the current test data performed.

I think in an attempt to prove that the means of both sets of data are the same, two graphs were produced [Figure 8] that show how similar the means are of the sets of data, thereby saying the Null Hypothesis is true.

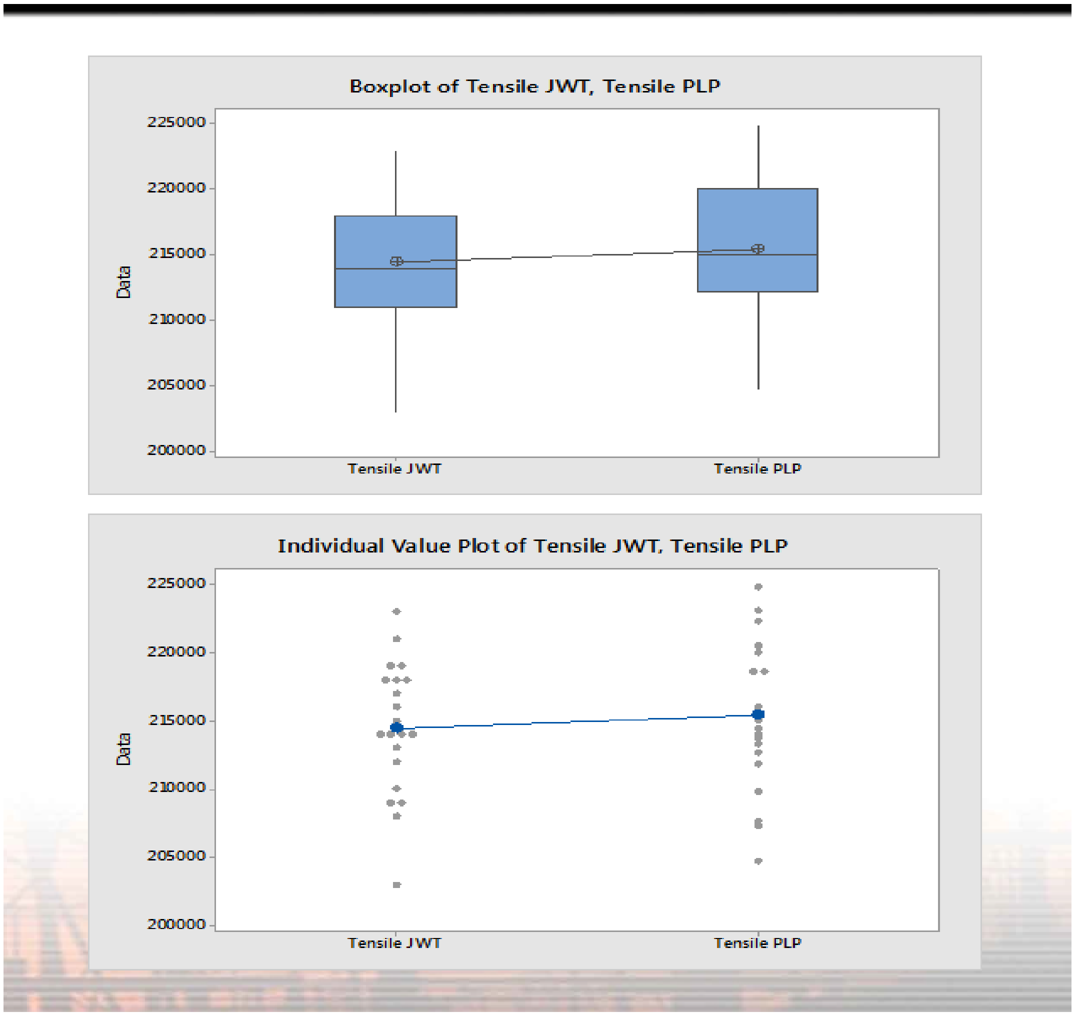


Figure 8

I also don't like how they preformed the t-test with the assumption of unequal variances. A rule of them when assuming equal or unequal variances, is that you take your larger standard deviation and divide it by your smaller one. If the result is greater than two you assuming non equal variance and other wise you assume equal. I believe they should have assumed equal variance here.

Finally, I don't like how they used one big calculation for all their different wire sizes. While in theory it should work, I think it is more reliable to look at each wire size individually and do an individual analysis.

2 Next Steps

So, after analyzing the data in the presentation, Ian and I decided to take some further steps in contextualizing the data and trying to come to a conclusion. We took the wire break loads from 2019, and 2020 and did the analysis for each individual part number that had enough volume for statistical significance.

This would come in three parts; the first part is confirming that each of the individual part numbers are normally distributed. The second is performing the two sample T-test again to confirm that each of them do not reject the Null Hypothesis. The final and third step is to look at the mean value for each of the wire lengths to gather supporting evidence that our initial Null Hypothesis was true.

2.1 Example of A Wire Analysis (11-086)

The probability plot is shown below [Figure 9], as we can see part number 11-086 is fairly linear and can be concluded as normally distributed.

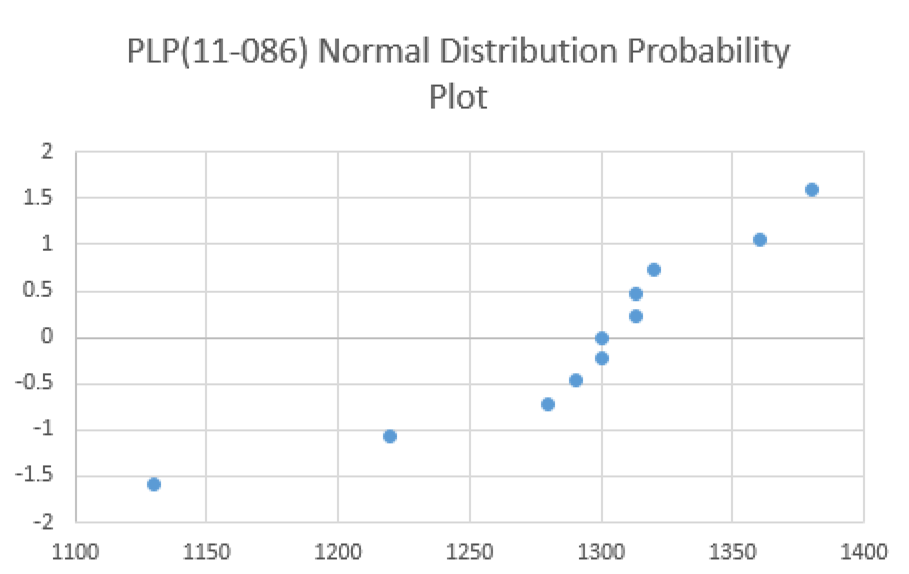


Figure 9

The means and p-value can also be seen below [Figure 10]. The P-value is higher than 0.05 so we cannot reject the null hypothesis, and the means are 1320 and 1291. These in my opinion are fairly similar, although I don't have enough background knowledge in what they are used for to come to a fair conclusion about their similarities

t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	1291.484545	1320
Variance	4574.904627	197.75
Observations	11	11
Hypothesized Mean Diffe	0	
df	11	
t Stat	-1.368978531	
P(T<=t) one-tail	0.099155115	
t Critical one-tail	1.795884819	
P(T<=t) two-tail	0.19831023	
t Critical two-tail	2.20098516	

Figure 10

The analysis done for the other parts can be seen in the appendix. They all turned out to be normally distributed, failed to reject the null hypothesis test and had similar mean values.

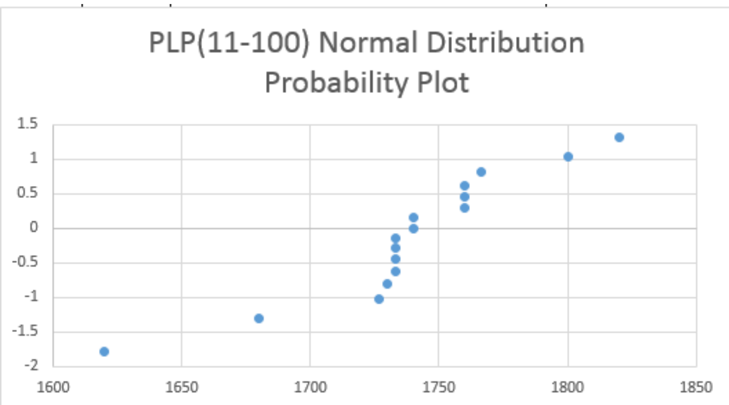
2.2 Conclusion

Given the normality of each set of data, how all of them fail to reject the null hypothesis, and the similarity in the mean values of each set of data, we can come to a conclusion that our Null Hypothesis is true. Since our Null Hypothesis is true, we can confidently say we can implement the lot skipping process.

I will once again state that I do not know the context the break load values mean when compared to each other. So, in reality the mean values could be very different I just don't realize it. But if they are as similar, as they look, we can say our Null Hypothesis is correct.

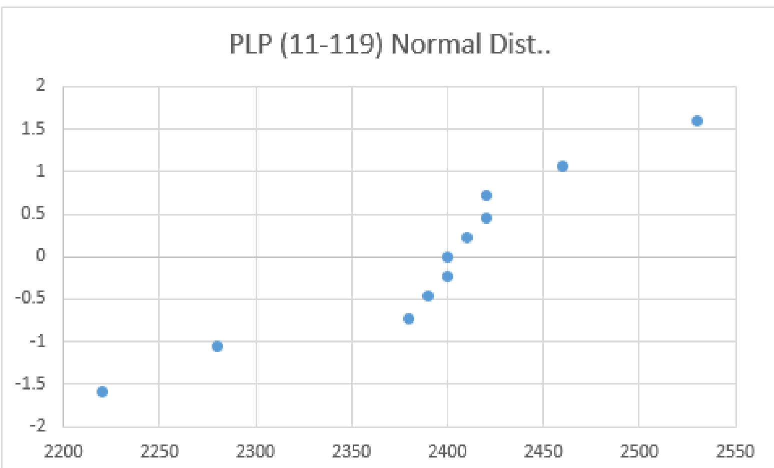
Appendix

Part Number 11-100



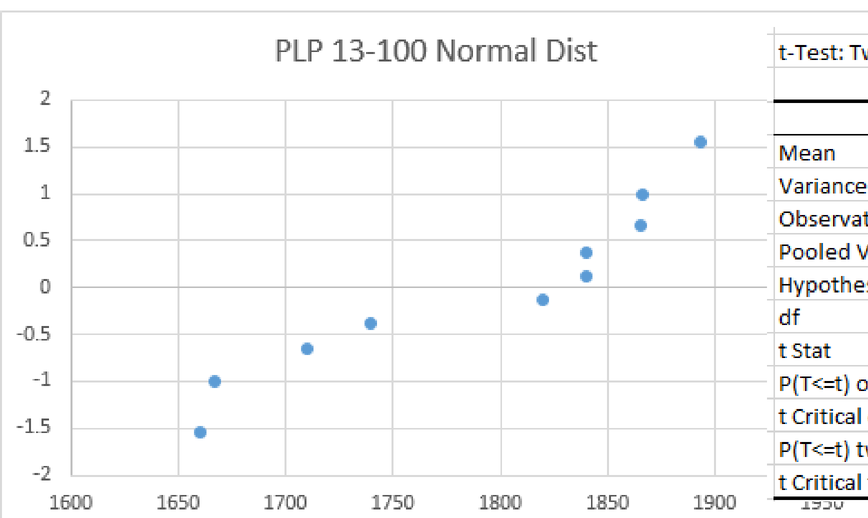
t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	1744.508235	1750.313529
Variance	2276.322228	837.7823618
Observations	17	17
Hypothesized Mean Diffe	0	
df	26	
t Stat	-0.428925722	
P(T<=t) one-tail	0.335755195	
t Critical one-tail	1.70561792	
P(T<=t) two-tail	0.671510389	
t Critical two-tail	2.055529439	

Part Number 11-119



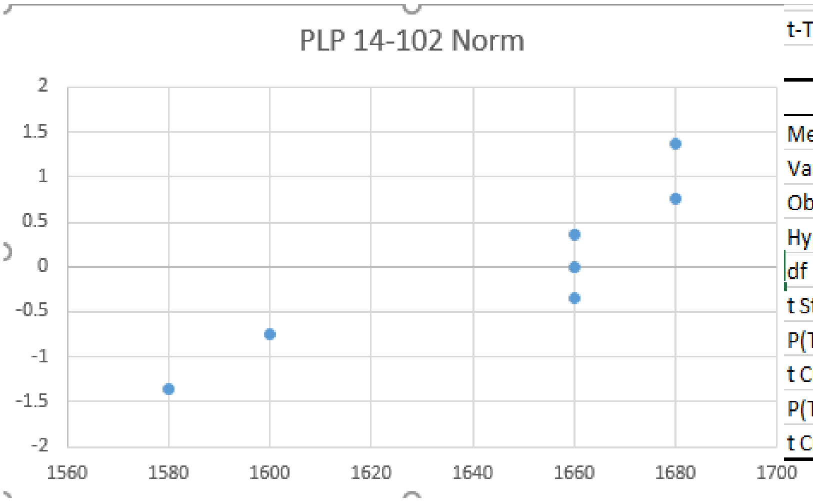
t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	2391.818182	2381.196364
Variance	6796.363636	1193.192145
Observations	11	11
Hypothesized Mean Diffe	0	
df	13	
t Stat	0.394124914	
P(T<=t) one-tail	0.34993892	
t Critical one-tail	1.770933396	
P(T<=t) two-tail	0.69987784	
t Critical two-tail	2.160368656	

Part Number 13-100



t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	1790.165	1763.8
Variance	7676.814206	7492.9
Observations	10	10
Pooled Variance	7584.857103	
Hypothesized Mean Diffe	0	
df	18	
t Stat	0.676922697	
P(T<=t) one-tail	0.253530976	
t Critical one-tail	1.734063607	
P(T<=t) two-tail	0.507061952	
t Critical two-tail	2.10092204	

Part Number 14-102



t-Test: Two-Sample Assuming Unequal Variances

	Variable 1	*Variable 2*
Mean	1645.714286	1633.428571
Variance	1561.904762	1466.285714
Observations	7	7
Hypothesized Mean Diffe	0	
df	12	
t Stat	0.590687576	
P(T<=t) one-tail	0.282844484	
t Critical one-tail	1.782287556	
P(T<=t) two-tail	0.565688967	
t Critical two-tail	2.17881283	