Lab 2: Hypothesis Testing using JMP

Objectives:

- Assessing normality in JMP

- Hypothesis Testing in JMP
  - Introduction to the various types of comparative methods and associated hypothesis tests (one sample t- and z-test, paired test, two sample t-test)
  - Be able to construct a hypothesis statement
  - Understand the different types of risks associated with hypothesis tests

Note: There is no wet lab in this practice.

Follow the instruction steps in the order that they are written; do not jump to a later section before you have covered the previous ones.

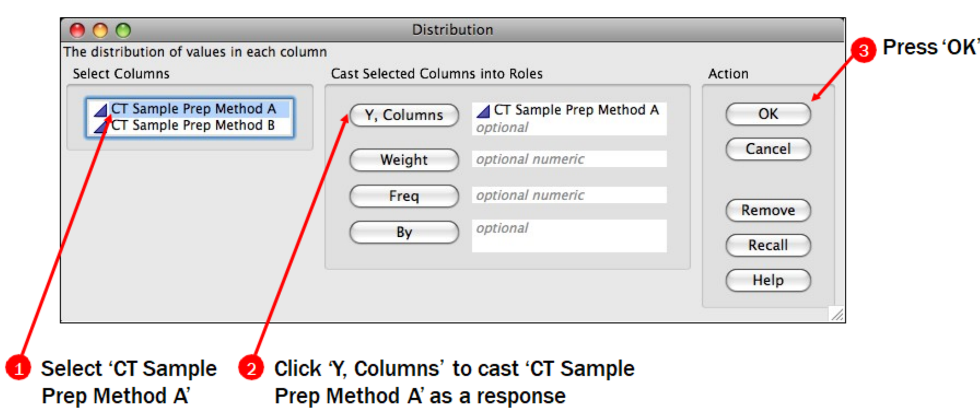
Part I: Assessing Normality

Fit a distribution to your data: use the **Continuous Fit** options in JMP to fit a distribution to a continuous variable. A curve is overlaid on the histogram, and a Parameter Estimates report is added to the report window. A red triangle menu contains additional options.

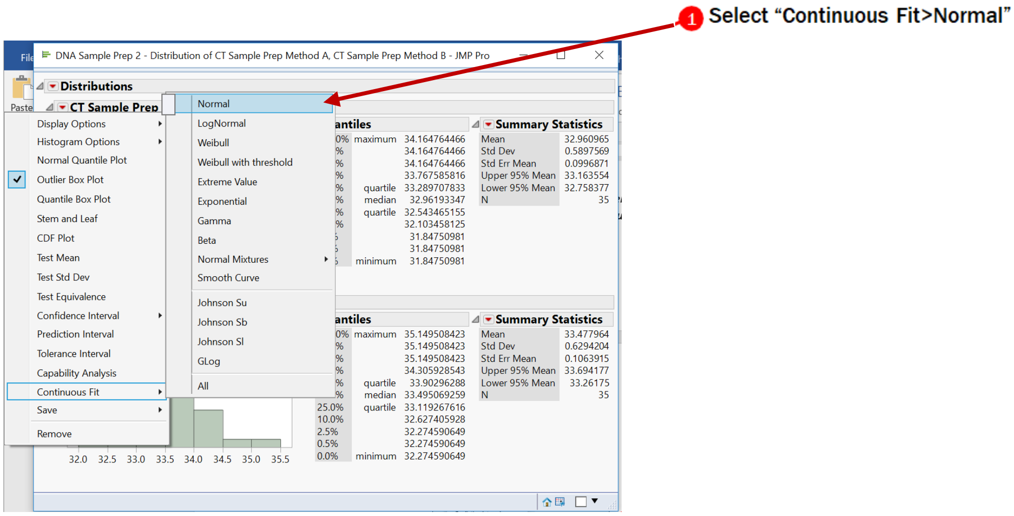
Instructions:

Follow the Example for continuous variable:

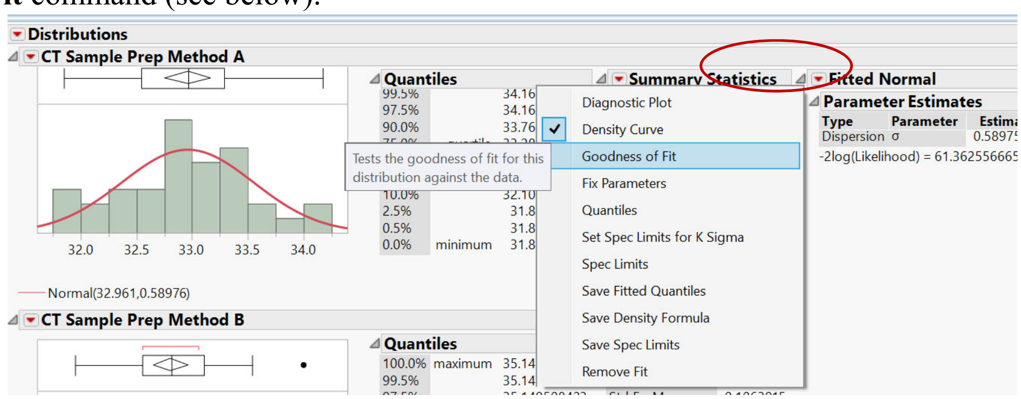
1. **Open** DNA Sample 2.jmp file, then go to **Analyze > Distribution** (see below). Select the two sample prep methods as Y columns (see below), then press OK. JMP should show the histograms and the summary statistics after this command.



2. **Next** click the downward pointing red icon (red triangle menu) next to the CT Sample Prep Method A. From the command menu select **Continuous Fit > Normal** (see below).

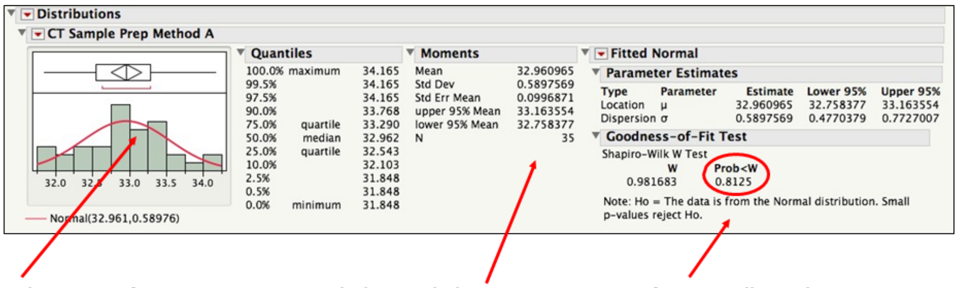


3. **Next** go to the "**Fitted Normal"** title and select/click from its sub-menu the **Goodness of Fit** command (see below).



The results from the Goodness-of-Fit Test (tests for normality by Shapiro-Wilk Goodness-of-Fit test) will show a report below the Fitted Normal title (see below)

4. Interpret the Distribution Report



Histogram of Data with Normal Curve Fit to Data

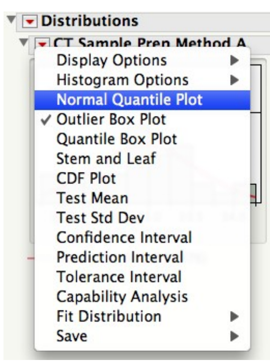
Box Plot of Data shown above histogram

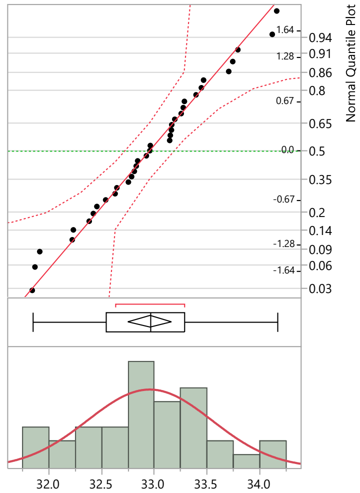
Descriptive Statistics for the data set

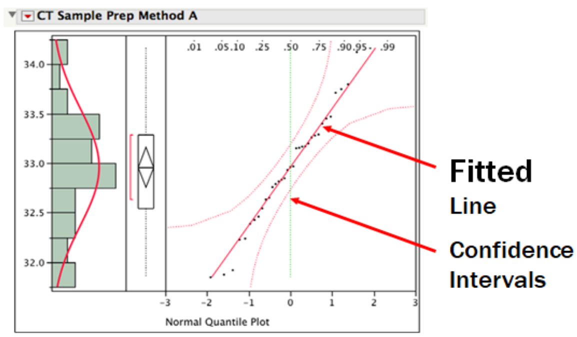
Test for Normality, Select Fitted Normal -> Goodness-of-Fit

If p-value is **lower** than the pre-determined **level of significance** (typically 0.05), then the data is **not normal**

5. **Next** click the downward pointing red icon (red triangle menu) next to the CT Sample Prep Method A and from the command menu select **the Normal Quantile Plot command** (see below). The **Normal Quantile Plot** is another visual test for normality of the data distribution







Distribution fits data if:

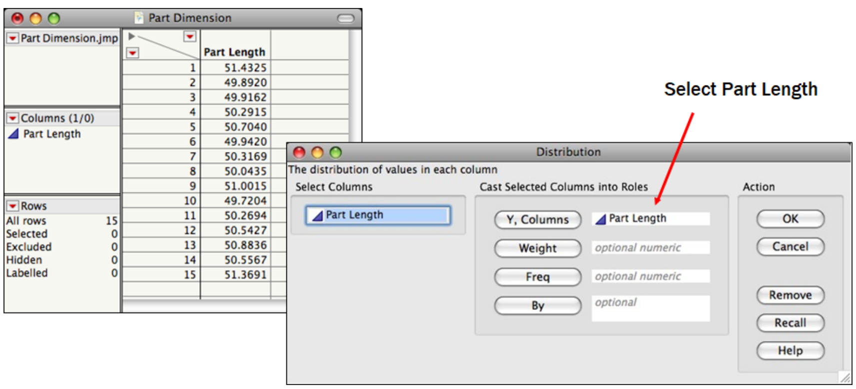
- **Plotted points roughly form a straight line**
- **Plotted points fall close to the fitted line**

<u>To complete the Lab Report, perform the following tasks and answer all Questions:</u>

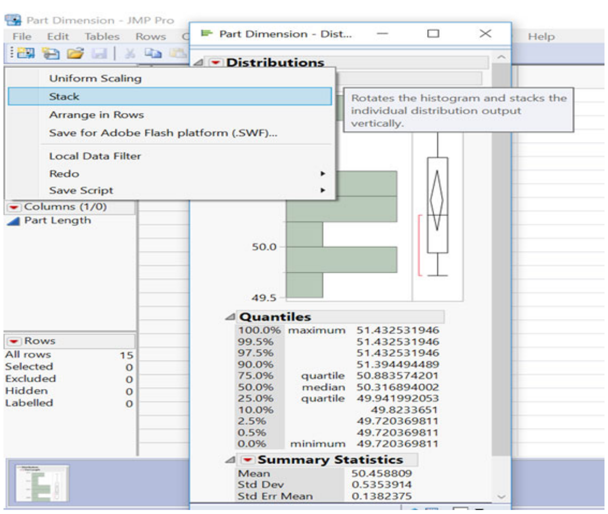
- Open JMP File: Assessing Normality.jmp
- The table contains four (4) sets of data
- For each set of data, answer these questions:
  - Is each data set normally distributed? ***[Type your answer here]***
  - If it is non-normal, what type of distribution is it (left skewed, etc.)? ***[Type your answer here]***
  - If it is non-normal, what are the possible root causes for the non-normality? ***[Type your answer here]***
  - How could you go about correcting the problem? ***[Type your answer here]***

Part II. Hypothesis Testing (statistical tools to detect differences in data)

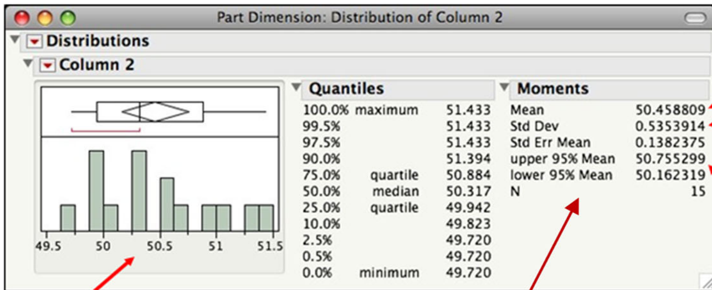
1. **One sample compared to a standard, when the standard deviation is known: z-test**
   Follow Example-1: Evaluating part dimensions
   The specification for the length of an injection-molded plastic part is 50 mm. An order of parts from the supplier was received with a reported mean = 50 and std. deviation = 0.50. An engineer selects 15 parts at random, measures the length and finds the mean to be equal to 50.45 and the standard deviation = 0.54. Is it likely that this sample of 15 came from a population whose true mean is 50 with a standard deviation of 0.50? Use JMP to find out.
   Follow the Instructions:
   a. **Open** file Part Dimension.jmp
   b. Go to **Analyze>Distribution**, select part Length for Y, click OK



c. Next go to downward pointing red icon (red triangle menu) next to the **Distribution title** and select from the dropdown window the **Stack** option. This should bring the histogram in a horizontal position for easier visualization.



d. **Look at the Distribution Report:**



Sample Statistics
Mean = 50.45
Standard Deviation = 0.53

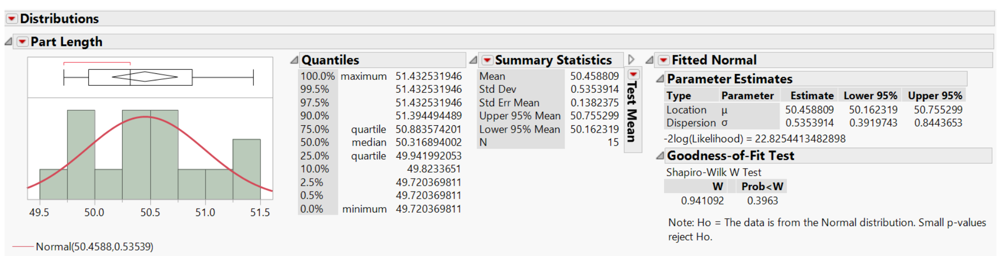
Notice the hypothesized mean of 50 is not contained in the confidence interval

Histogram of sample data. Data looks to be normally distributed

Look at the Confidence Interval (CI) defined by the "upper 95% Mean" and the "lower 95% Mean" (see above): from 50.76 to 50.16

Does it include the expected mean? ***[Type your answer here]***

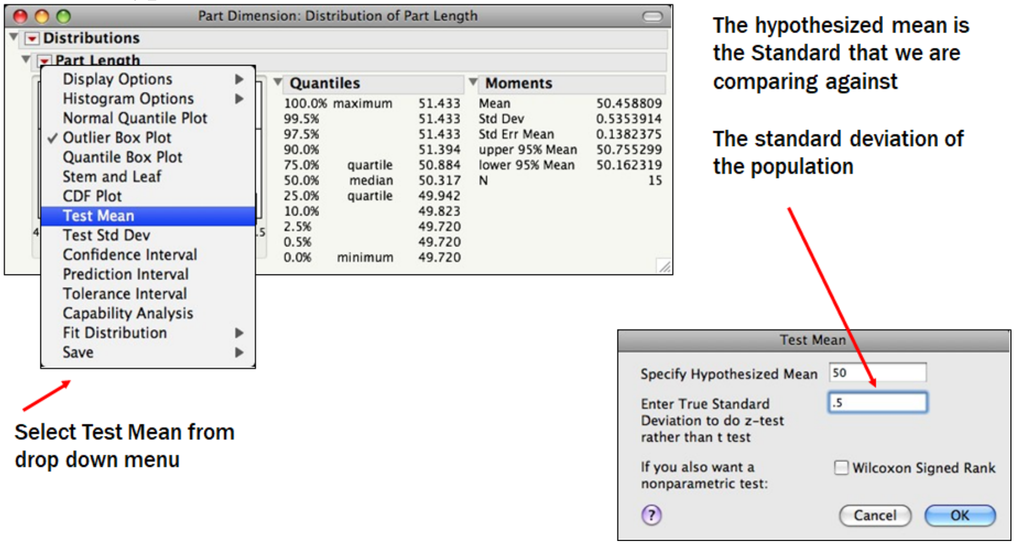
e. To check for **Normality,** perform the **Goodness of Fit test** as in Part I (see result below).



Answer the question: Is the distribution normal? Explain.

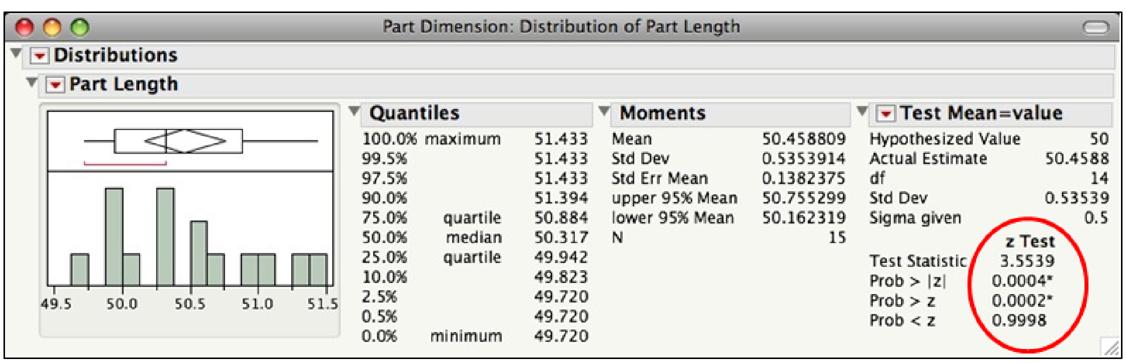
[Type your answer here]

f. Next ("test the mean" step) go to Distributions, select from dropdown window **Test Mean. Enter the hypothesized mean and std. deviation.** See below



Use the Test Mean window to specify options and perform **a one-sample test for the mean**. If you specify a value for the standard deviation, a **z-test** is performed. Otherwise, the sample standard deviation is used to perform a **t-test**. You can also request the nonparametric Wilcoxon Signed-Rank test (see the option above in the "Test Mean" window).

g. **Interpretation of One sample z-test results:**



H_o: $\mu = \mu_1$ There is no difference between standard and the sample mean

P-Value less than $\alpha = 0.05$, therefore reject the Null Hypothesis

What is your conclusion about the true mean? *[Type your answer here]*

Probability values in JMP:
Prob>|t|
The probability of obtaining an absolute t-value **by chance** alone that is greater than the observed t-value when the population mean is equal to the hypothesized value. This is the p-value for observed significance of the two-tailed t-test.

Prob>t
The probability of obtaining a t-value greater than the computed sample t ratio by chance alone when the population mean is not different from the hypothesized value. This is the p-value for an upper-tailed test.

Prob<t
The probability of obtaining a t-value less than the computed sample t ratio by chance alone when the population mean is not different from the hypothesized value. This is the p-value for a lower-tailed test.

h. **Next Calculate the z-value manually using the formula below:**

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Question: What is your z-value, is it the same as the one calculated by JMP?

[Type your answer here]

2. **One sample test (t-test) compared to standard when the standard deviation is unknown:**

Follow Example-1: Cover Heater Temperature from Vendor 1.
Open File: Heated Cover Design.jmp

The temperature of a plate cover heater for a diagnostic instrument must be greater than 100°C to prevent condensation of the sample during the thermal cycle. A new design is being evaluated. An engineer from **Vendor-1** assembles 10 prototypes and measures the temperature of the cover. The results for **Vendor-1** are in the data file: Heated Cover Design.jmp.

The engineer wants to know if the average temperature of the heater cover from Vendor-1 is meeting the 100°C requirement. Hence, he needs to apply hypothesis testing for the mean.

Follow the Instructions for the Hypothesis analysis for the data set <u>for Vendor-1</u>:

a. Construct the Hypothesis:
   - H_o: $\mu_1 = 100$ Mean cover temperature = 100°C
   - H_a: $\mu_1 \neq 100$ Mean cover temperature ≠ 100°C

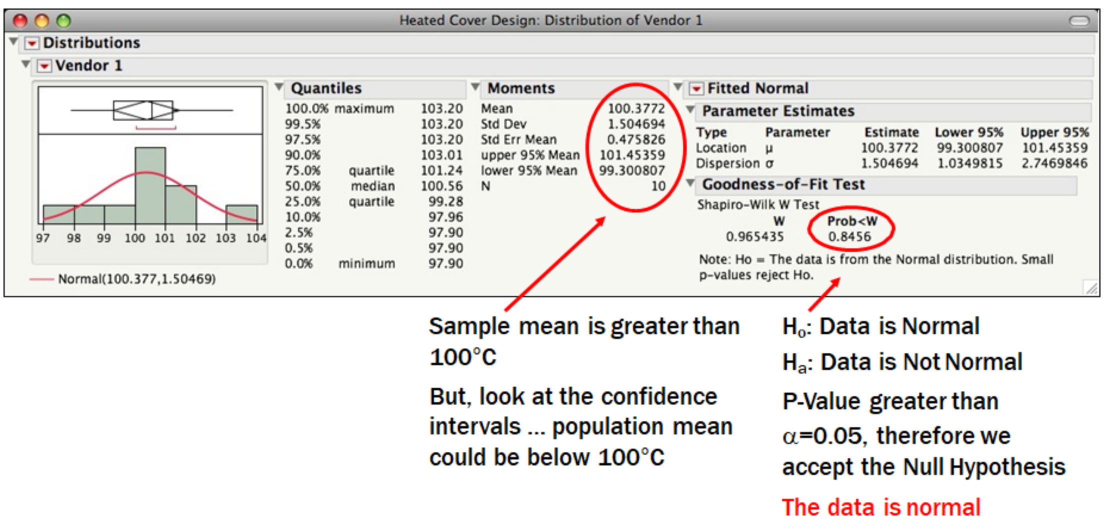
b. Test assumption of normality in JMP (see Part I): perform Goodness-of-Fit test

c. Calculate the test statistic (Test Mean) in JMP

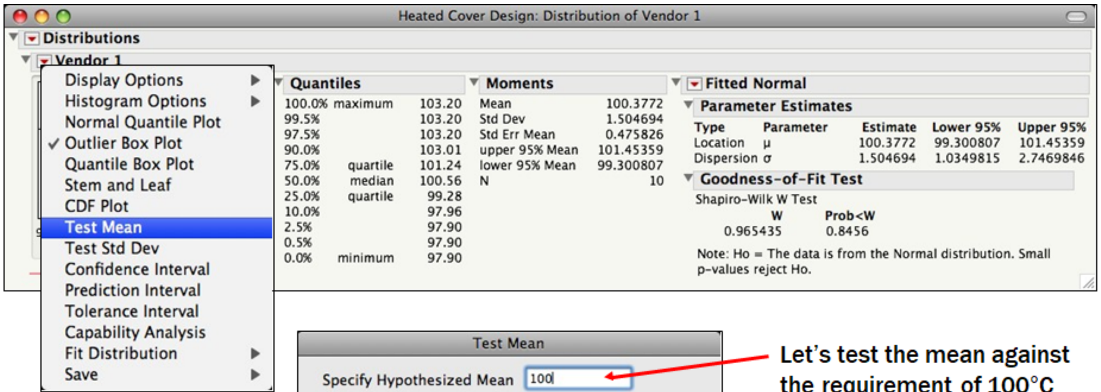
d. Compare the p-value to α(alpha). Do you reject or fail to reject the Null Hypothesis? ***[Type your answers here]***

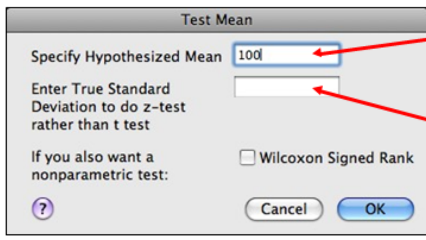
<u>Results for VENDOR-1:</u>

✓ **Normality Test results:**



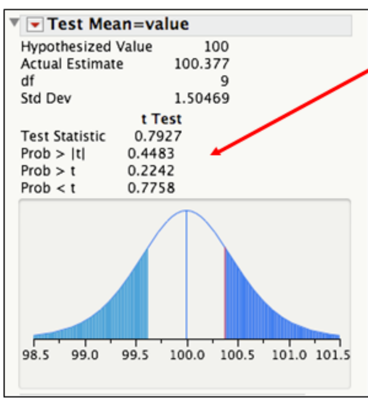
✓ **Test Mean Results: t statistic is 0.793; p-value (0.448) is greater than 0.05**





Let's test the mean against the requirement of 100°C

We don't know the true standard deviation of the populations, so we'll leave this blank



<u>Understanding P-Values</u>

Prob > ItI Probability of obtaining t-value greater than the absolute value buy chance assuming H_o is true (i.e. no difference in means)

Prob > t Probability of obtaining a t-value greater than the computed t-value by chance assuming H_o is true

Prob < t Probability of obtaining a t-value less than the computed t-value by chance assuming H_o is true

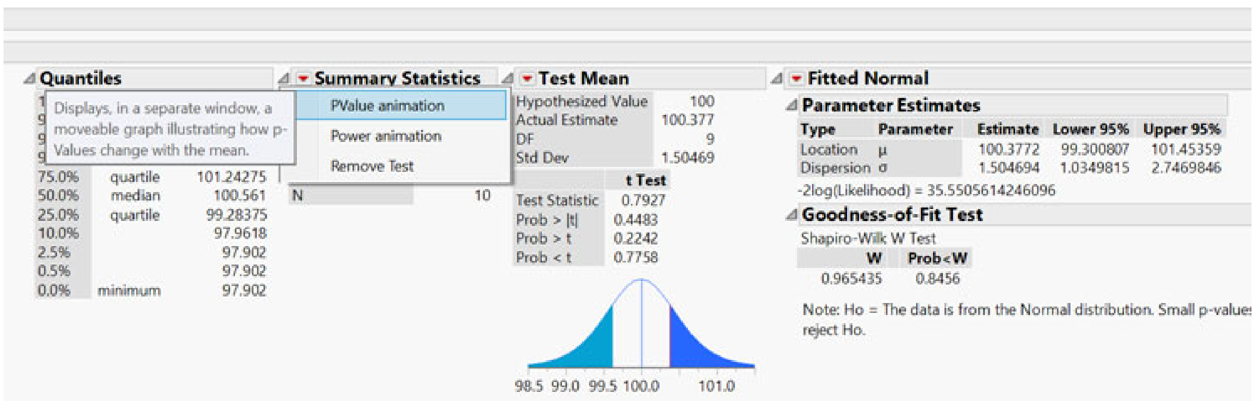
Since the P-Value is greater than α = 0.05, we accept the Null Hypothesis

We can not conclude that the mean temperature is greater than 100°C

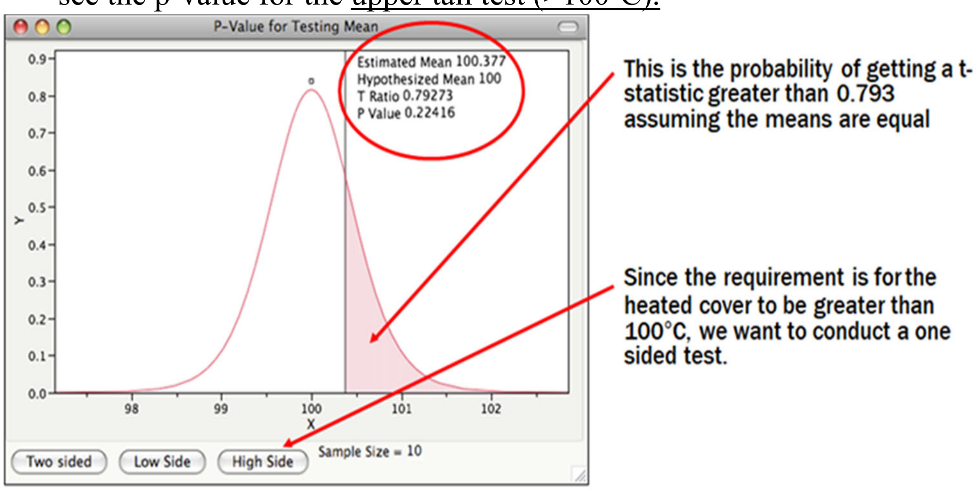
<u>What is the probability of getting a t statistic greater than 0.793 by chance assuming H_o is true?</u> *[Type your answer here]*

✓ **Using PValue Animation:** an interactive visual representation of the p-value. Enables you to change the hypothesized mean value while watching how the change affects the p-value.

Follow instructions below:
Go to the downward pointing red icon (red triangle menu) next to the **Test Mean** title and select from the dropdown window the **PValue animation** command as shown below.



Since we are interested in a temperature equal or <u>higher</u> than 100ºC, <u>click</u> on the **High Side** option from the P-value Animation window (see below), which will allow us to see the p-value for the <u>upper tail test (>100ºC).</u>



If we're willing to accept a 22.4% probability of being wrong, then we can claim that the mean temperature of the heated cover meets specification!

✓ Calculate the t-statistic manually using the formula below.

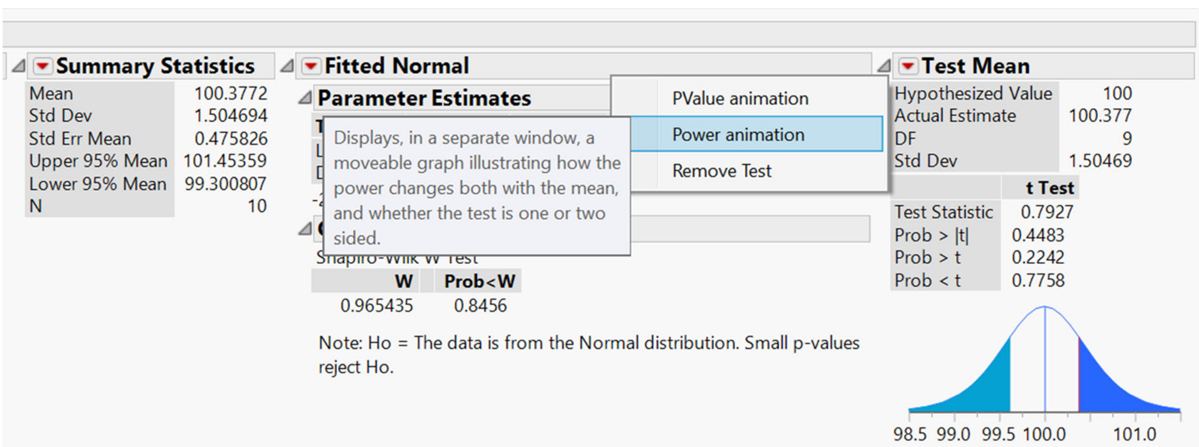
Question: What is your t value; is it the same as the one from the JMP software?

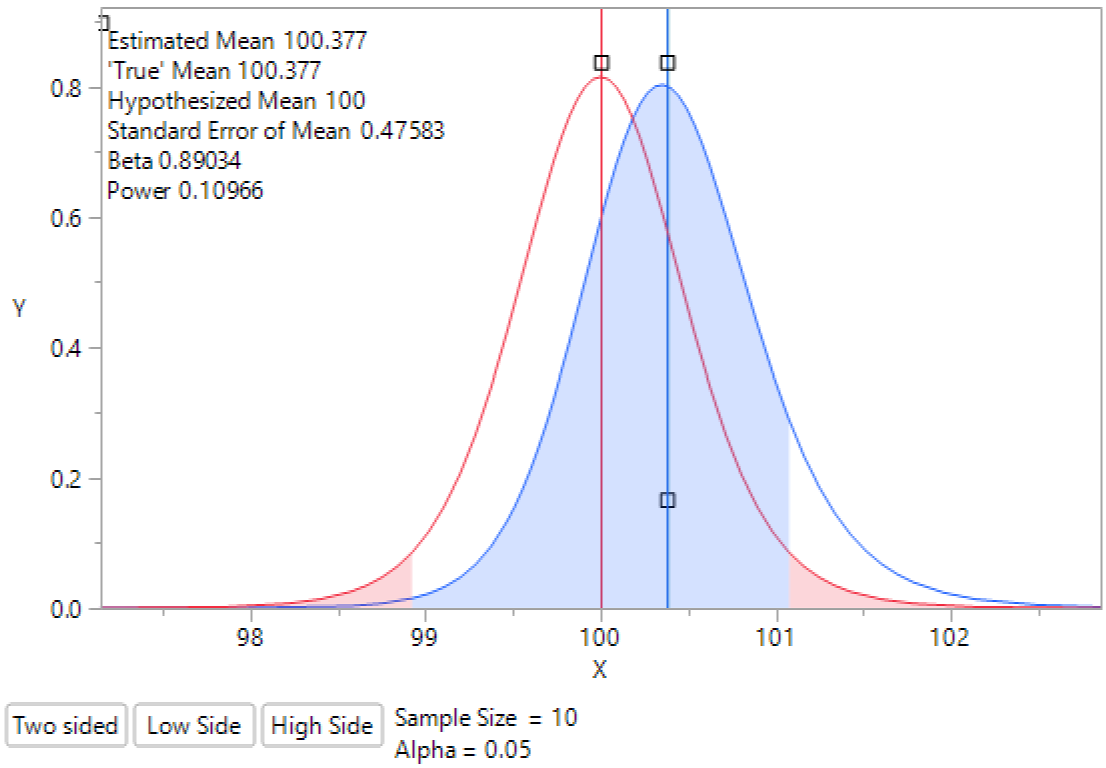
[Type your answer here]

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- ✓ **Power animation:** starts an interactive visual representation of power and beta type error. You can change the hypothesized mean and sample mean while watching how the changes affect power and beta.

__**Instructions:**__ Click the red triangle next to the **Test Mean** and select **Power animation** command from the sub-menu as shown below. This opens the interactive graph that illustrates the concept of power.





We see two distributions. The red one centered at 100 shows the sampling distribution based on the assumptions in the null hypothesis. The blue one shows the alternative sampling distribution based on the possibility that μ is not 100. We can relocate the blue curve and thereby see how the **power of the test** will change if the population mean were any value we choose.

The graph displays a two-sided test (but you can change this by clicking on the other options/commands) at the bottom of the window. In the red curve the pink shaded areas represent alpha (=0.05), that is each tail is shaded to have 2.5% with 95% of the curve between the shaded areas. If we find the sample mean in the pink shaded tails, we will reject the null. **The blue shaded area shows β.**

Question: Looking at the power animation graph above, how much is β ? Do you think we have enough power? Explain.

[Type your answer here]

<u>To complete the Lab Report, perform the following tasks and answer all Questions:</u>

- ✓ Change alpha from 0.05 to 0.01: click on 0.05, enter 0.01, and observe what happens to the graph (the red one and the blue curve)

  What happens to the pink shaded areas? ***[Type your answer here]***

  What happens to the blue shaded area? ***[Type your answer here]***

  What happens to the power of the test? ***[Type your answer here]***

- ✓ Restore alpha to 0.05 and then click on **sample size** and enter 20 (twice the original sample size of 10).

  What happens to β? ***[Type your answer here]***

  What happens to the power of the test? ***[Type your answer here]***

- ✓ Restore the Sample size to 10 and then grab the small square at the top of the blue curve and slide it to the right until **<u>the true mean is approximately 102</u>**. This action shows a vertical green line which represents our sample mean (estimated mean).

  What happens to the power and β ? How much are they? ***[Type your answer here]***

<u>To complete the Lab Report, perform the analysis for the second type of covers (for Vendor 2) and answer the following Questions:</u>

- ✓ **Is the data normally distributed?** ***[Type your answer here]***
- ✓ **Do you reject or fail to reject the Null hypothesis for a true mean of 100°C and why?** ***[Type your answer here]***
- ✓ **What is the power of your test? Do you have enough power?** ***[Type your answer here]***

<u>To complete the Lab Report, perform a hypothesis test for the mean on your sample volume data from Lab-1:</u> Go to the results from Lab-1 and test the two sets of 16-samples with the two pipettes (total 32 sample volumes, 16 for each pipette, from one operator only) for **<u>the true mean of 100ul (hypothesized mean value of 100ul).</u>** **<u>Answer the following Questions:</u>**

- ✓ **Do you reject or fail to reject the null hypothesis and why?** ***[Type your answer here]***
- ✓ **What is the power of your test? Do you have enough power?** ***[Type your answer here]***

3. The Paired t-test

When two groups are considered, there can be two different analyses:

- **Matched pairs**: the two responses form a pair of measurements coming from the **same experimental unit or subject** (before-and-after type of data from the same subject, responses are correlated)
- **Independent groups:** responses from the two groups are unrelated and statistically independent (responses are uncorrelated)

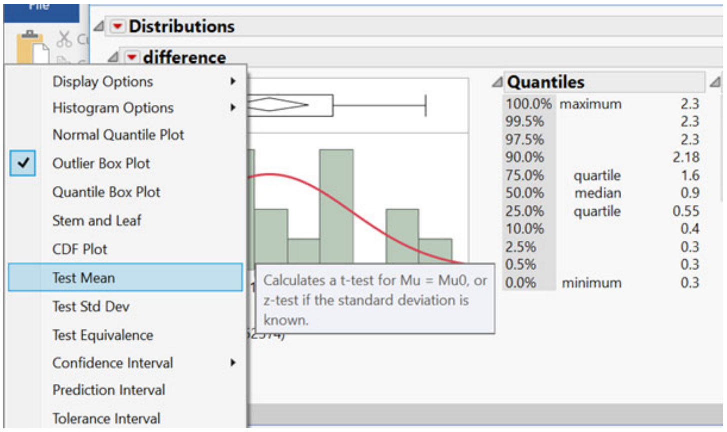
Further Assumptions for the paired t-test: the differences are Independent and Normally distributed.

Follow Example-1: Open data file designated as Therm-paired t-test Lab 2.jmp

A health care center suspected that temperature readings from a new ear drum probe thermometer were consistently higher than readings from the standard oral mercury thermometer. To test this hypothesis two temperature readings were taken, once with the ear-drum probe, and the other with the oral thermometer. The table has 20 observations and 4 variables. The two responses are the temperature readings taken orally and tympanically (by ear) on the same person (the name column) on the same visit. The fourth column contains the difference between the two responses.

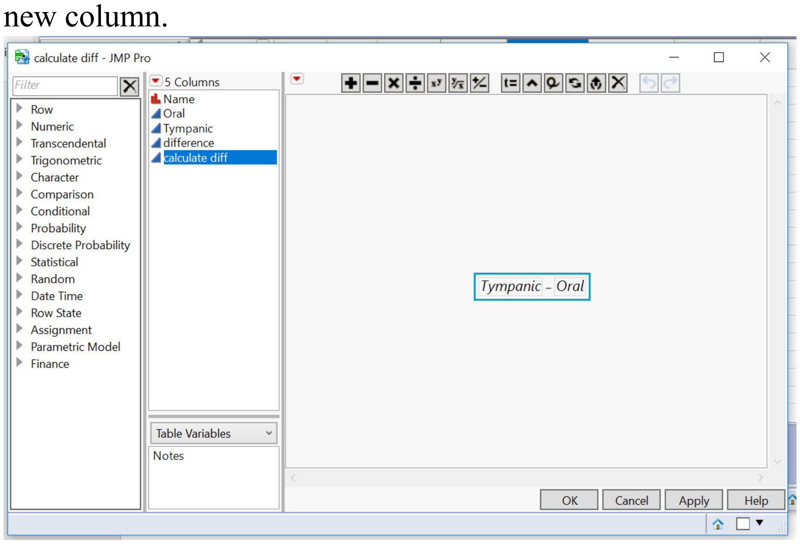
To complete your report, you need to perform the following tasks and answer the posted questions:

- Start with the **distribution analysis** (select oral and tympanic as Y)
- Look at the **mean, std deviation, boxplot, skewness, median**, **what can you conclude**? ***[Type your answer here]***
- Check **for normality**: perform goodness-of-Fit test to determine if you have normal distribution. **Is the distribution normal?** ***[Type your answer here]***
- Now analyze the distribution of **the Difference** (data in the 'difference' column) in the temperatures. **Is the distribution Normal for the Difference?** ***[Type your answer here]***
- What are the upper 95% and lower 95% mean limits (Confidence Interval) for the Difference? ***[Type your answer here]***
- Does the confidence interval (CI) of the means include zero? How is the zero value within the CI important? What would the presence of zero suggest? ***[Type your answer here]***
- Perform Test Mean t-test: choose Test Mean from the dropdown menu (see below) and select zero for the hypothesized value.



- Interpret the t-test for the hypothesized mean of the difference.
  a. What is the test statistic? ***[Type your answer here]***
  b. What is the p-value? ***[Type your answer here]***
  c. Do you reject or fail to reject the null? ***[Type your answer here]***
  d. Are the two temperature means different? ***[Type your answer here]***

Note: to calculate the difference by yourself in JMP create a separate new column in the data table with the temperature measurements and designate it as "calculate diff". Select this new column and go to Formula command. In the Formula window select one of the two responses, for example Tympanic, then select the minus sign (-) from the top menu in the formula window, then click on Oral as shown below. Click OK and the difference will populate the new column.

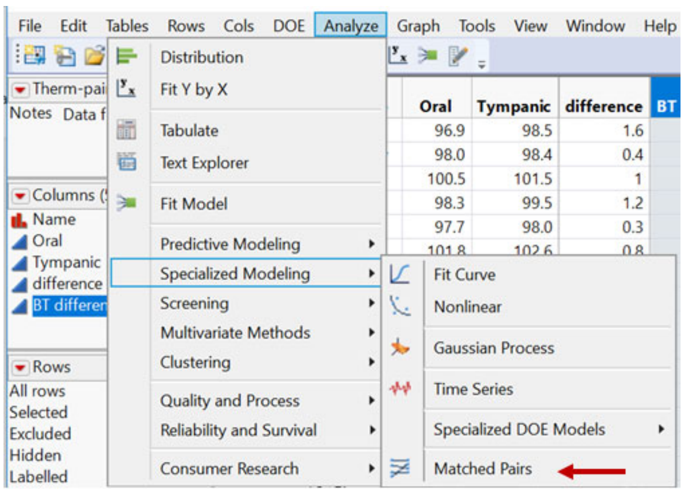


4. **Matched Pairs Platform in JMP:** another way of performing matched pairs analysis in JMP (the difference is not pre-calculated in a separate column)

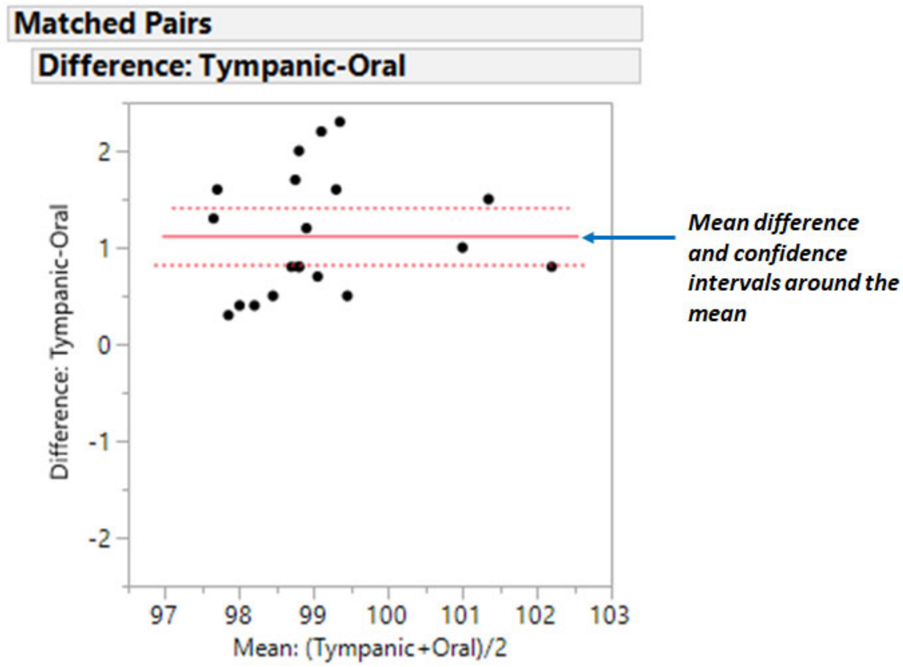
JMP offers a special platform for the analysis of paired data. The Matched Pairs platform compares means between two response columns using a paired t-test.

To use the platform:

Go to Analyze> Specialized Modeling> Matched Pairs (see below):



Use oral and tympanic as the paired responses. Click OK to see a scatterplot (see next page).



Tympanic	99.63	t-Ratio	8.030246
Oral	98.51	DF	19
Mean Difference	1.12	Prob > |t|	<.0001*
Std Error	0.13947	Prob > t	<.0001*
Upper 95%	1.41192	Prob < t	1.0000
Lower 95%	0.82808		
N	20		
Correlation	0.87389		

P-value < 0.05, reject the null

Confidence interval around difference does not contain 0

You can see here the correlation value, which shows strength of the correlation (dependence) between the two groups. The high correlation value shows that the two groups are not independent.

5. One sample compared to another sample: <u>2-sample t-test</u>

a. Two sample t-test with Equal variances. For systems with unknown variances that are assumed to be equal, the "measure of dispersion" is pooled sample standard deviation s_p

The test statistic t_o is calculated as:

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

t_0 has a t-distribution with $n_1 + n_2 - 2$ degrees of freedom

b. Two sample t-test with <u>Unequal variances</u>

When the variances are not equal the test statistic is:

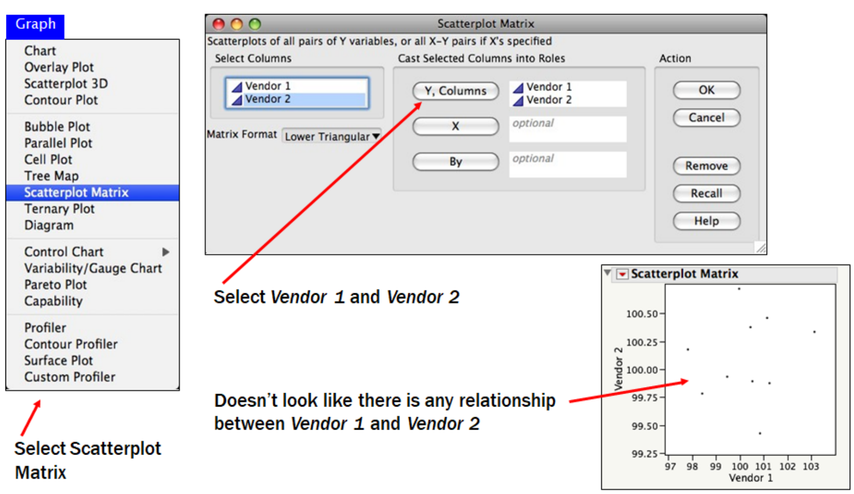
$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Example-1: We will use the Heated Cover Design data file (<u>Heated Cover Design.jmp)</u> from the one sample t-test. The engineer is evaluating now the two different vendors, Vendor-1 and Vendor-2, for the heated cover and wants to determine if there is a difference in the temperature between the two vendors. 10 prototypes with covers from each vendor are evaluated by measuring the temperature of the cover. We want to determine if the mean temperature from Vendor 1 is different from that of Vendor 2. The data is in file: <u>Heated Cover Design.jmp</u>

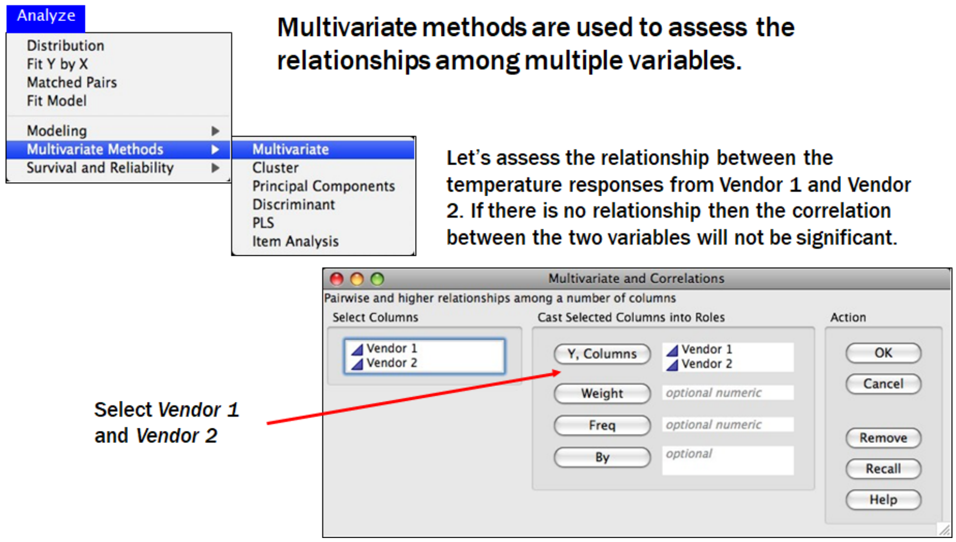
Follow the instructions/steps below:

1. **First step: test assumptions for normality (analyze distribution and goodness-of-fit)**
2. **<u>Assess independence</u> of the two data sets using scatterplot.** <u>To perform a 2-sample t-test the two populations need to be independent.</u>

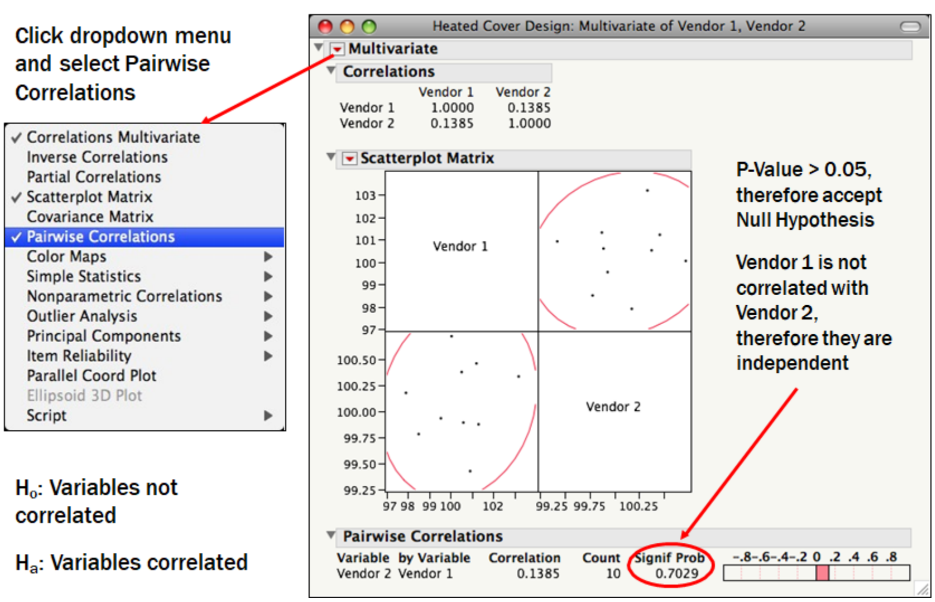
   a. Go to the **Graph** Command on the top menu and select **Scatterplot Matrix** from the sub-menu. Select for Y both data sets (Vendor-1 and Vendor-2) as shown below:



b. Go to **Analyze > Multivariate Methods > Multivariate** to perform a <u>Correlation analysis</u> and thus assess the data independence (see below)

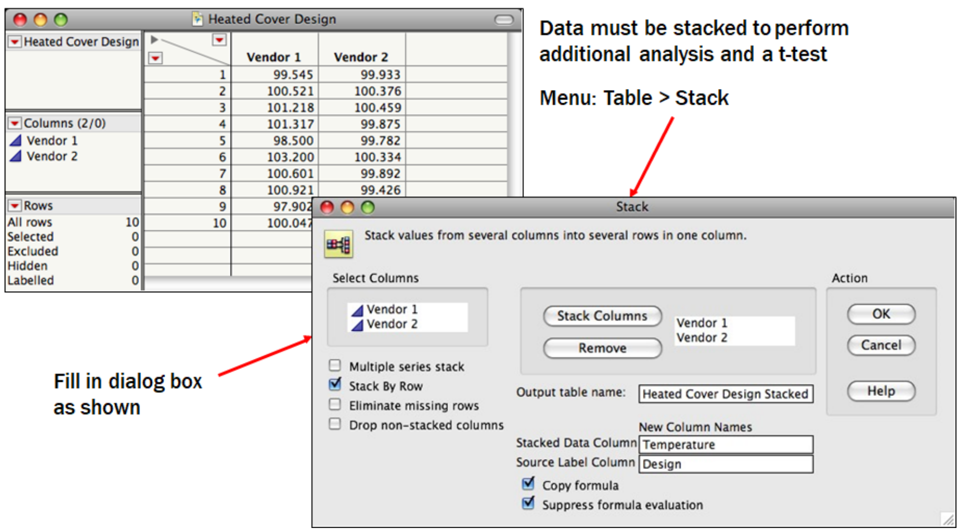


<u>To find the p-value (Signif Prob), go to the Multivariate command menu and select Pairwise Correlations as shown below:</u>

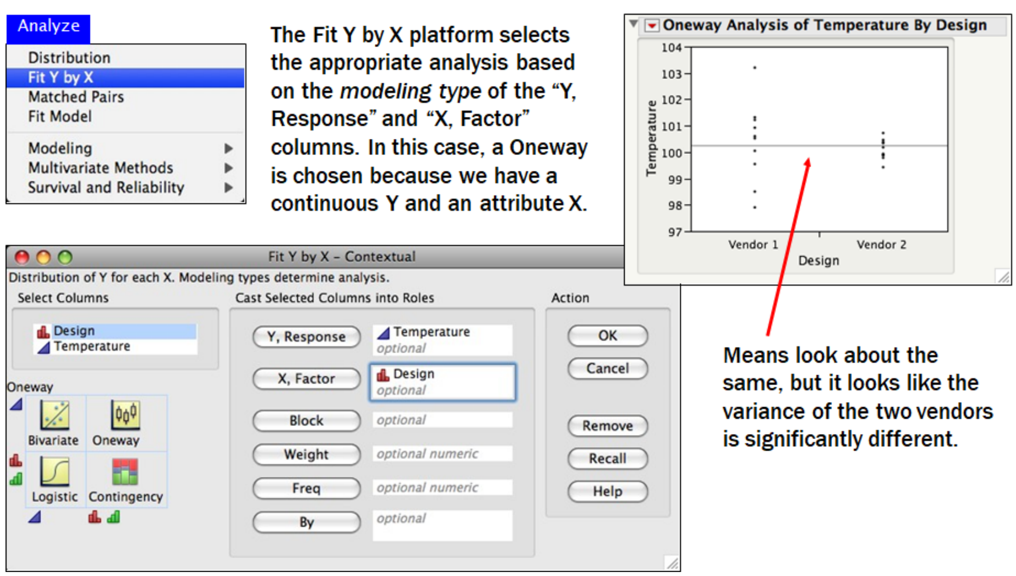


The correlation analysis shows that the two groups are not correlated (they are independent).

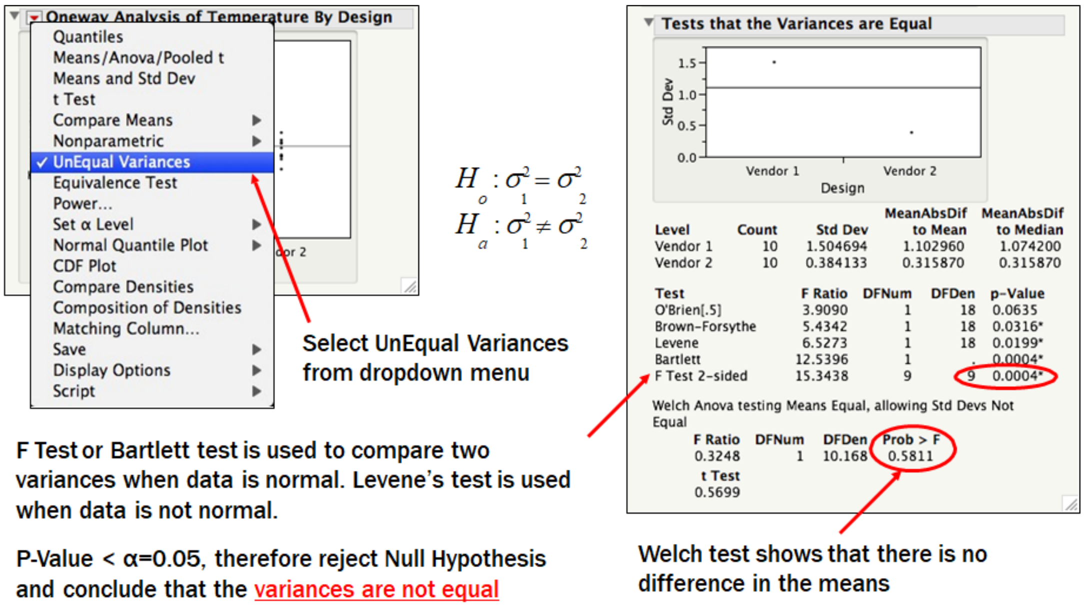
3. Next you need to **Stack** the data (data must be stacked for additional analysis) as shown below. Go to Table > Stack > Select both vendors and enter the new stacked column names (see below: for data column: Temperature, for Label column: Design), click OK.



4. Next comparing the two designs (data must be stacked for this function). Go to **Analyze > Fit Y by X > Select Temperature column for the Y, Response** and the **Design column for the X, Factor** as shown below. Click OK. Note that the response is a continuous variable and the factor is an attribute (categorical) input.



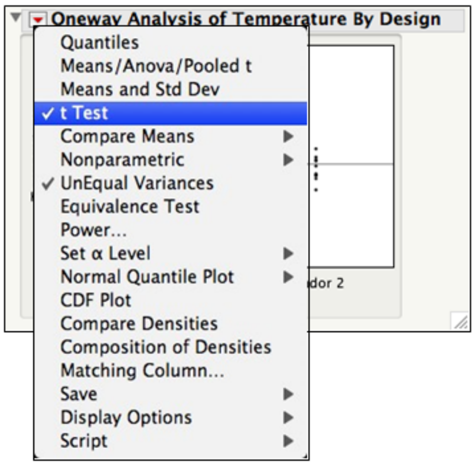
5. Next you need to test another assumption for using the two-sample t-test: testing the equal variance assumption. Go to the Oneway Analysis menu (red triangle) and select from the dropdown menu the **UnEqual Variances** command as shown below.



The Bartlett test (used for normal distribution) shows that variances are NOT equal.
In this case we can't use the pooled standard deviation. The Welch Anova test above can be used as quick test for the means when variances are not equal. The high p-value of the Welch test suggests that the means are not different (fail to reject null). However, we can also use a t-test in JMP with unequal variances as to confirm this conclusion. Go to the next step.

6.Comparing two means using <u>t-test with unequal variances.</u>

Go to the dropdown menu from the red triangle next to the Oneway Analysis title, and this time select **<u>t Test</u>** as shown below. This should result in a t Test report which includes the CI, the t ratio (t statistic) and the p-value.

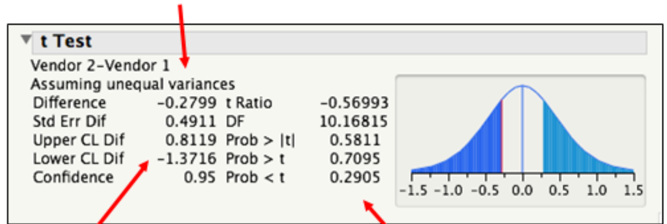


Select t Test from dropdown menu

H_o: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$

Since variances are unequal, JMP does not use pooled variance when calculating the t-Ratio.



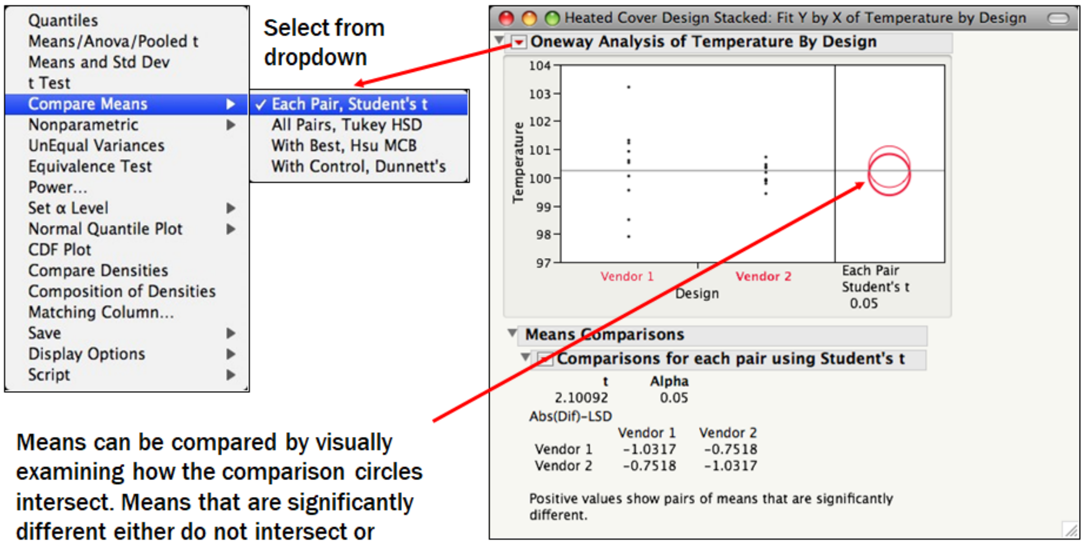
If zero is contained in the confidence interval, then there is not sufficient evidence to claim there is a difference between the means.

P-Value is 0.5811, which is greater than α=0.05, therefore accept the Null Hypothesis and conclude there is <u>no difference.</u>

However, this does not mean you can claim they are the same! Always check the power of the test.

Based on the results we fail to reject the Null.

Note: Use <u>Comparison circles</u> for a quick Visual Analysis for comparing the means as shown below.



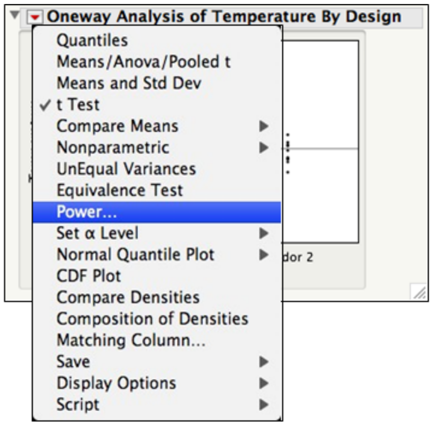
Means can be compared by visually examining how the comparison circles intersect. Means that are significantly different either do not intersect or intersect slightly.

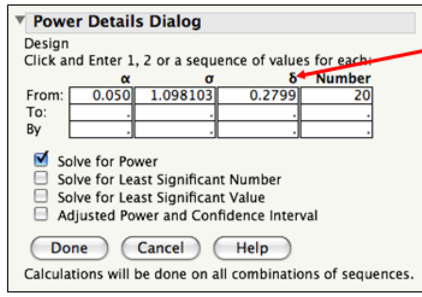
Select one circle and JMP shows all other variables that are similar are shown in Red

7. **Next: we need to check for the Power of the test using JMP**. Go to the top main menu and select Power from the sub-menu as shown below. Change only the value for δ (effect size) to the estimated difference of 0.2799 from the t Test (see above).

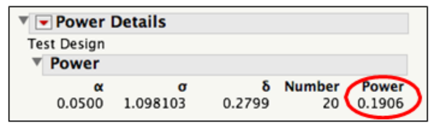
   The effect size, denoted by δ, is a measure of the difference between the null hypothesis and the true values of the parameters involved.

   Next, check the box for ‘Solve for Power” below the power table and then click Done as shown on the next page.





Let's see if we have enough power to detect the difference of 0.2799 with 20 total samples



With 20 total samples, the probability of detecting a difference of 0.2799 is 19.06%. In other words, if the observed difference of 0.2799 is real, the probably of detecting it is only 19.06%. A larger sample size is needed to detect this small difference.

If you select for "Solve for Least Significant Number" (LSN) and "Least Significant Value" (LSV) in the Power Details Dialog Box above, you can find out:

- **how many more observations (LSN) would make the reported difference become significant;**
- **how small a difference could the significance test detect in this example (LSV): 1.03**

Power Details

Test Vendor

Power

α	σ	δ	Number	Power
0.0500	1.098103	0.2799	20	0.1906

Least Significant Number

α	σ	δ	Number(LSN)
0.0500	1.098103	0.2799	61.60124

Least Significant Value

α	σ	Number	LSV
0.0500	1.098103	20	1.031735

<u>To complete the report, go to the data collected from Lab-1 and perform a two-sample t-test analysis using JMP as to compare the means between the two operators and the two pipettes.</u>

Answer the following questions:

a. Is there a difference between the two pipettes used by the same operator? Explain your answer using the data analysis (graphs, plots, reports) in JMP.
[Type your answer here]

b. Is there a difference between the two operators no matter which pipette was used? Explain your answer using the data analysis (graphs, plots, reports) in JMP.

[Type your answer here]

Note: if your data shows **equal variances** use the command for **"Means/Anova/Pooled t"** from the dropdown menu as shown below instead of the "t Test" which we used in the example above due to having unequal variances. The software performs in both cases a two-sample t-test but with equal variances it uses the pooled standard deviation. We will discuss more of this in our next lab.