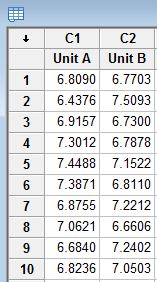
A food and beverage manager wants to determine whether there is any significant difference in the diameter of the cookies between two units. A randomly selected sample of cookies was collected from both units and measured? Analyze the data and draw inferences at 5% significance level. Please state the assumptions and tests that you carried out to check validity of the assumptions.

*Solution Approach*

α=0.05

Below is the sample data for cookies diameter between two units



Cutlet diameters of Unit A and Unit B

***Inputs*** are Diameter 1 and and Diameter 2 that is **Discrete in two categories**

***Output*** is **continuous** as we are trying to see the difference in diameter and diameter is a continuous variable

We proceed with ***2-sample t test***

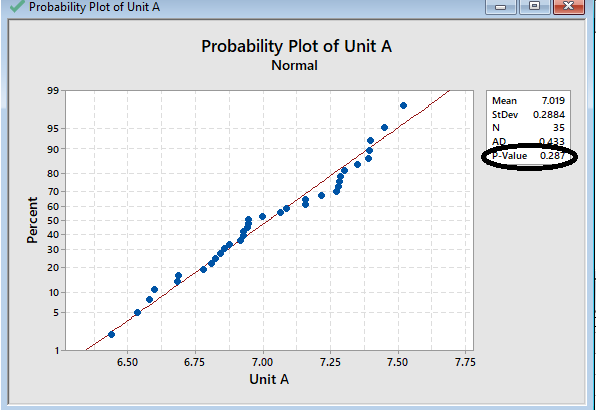
1) Normality test- We will see if data is normally distributed or not

Create hypothesis for Unit A

Ho= Data is Normally distributed

Ha=Data is not Normally distributed

Stat — >Basic Statistics — >Normality test



P-value is >0.05.

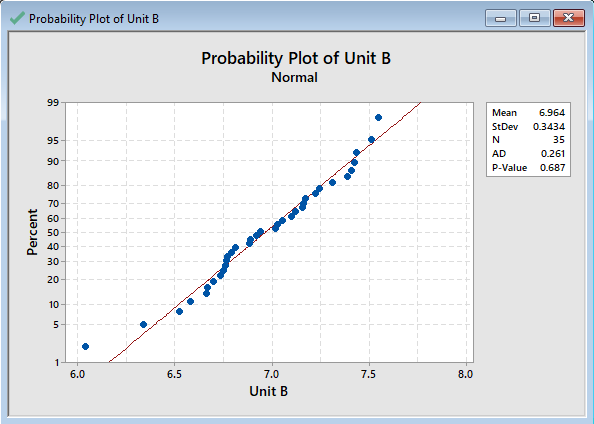
P High Ho Fly. So data is Normally distributed

Now Create hypothesis for Unit B

Ho= Data is normally distributed

Ha=Data is not normally distributed

Stat →Basic Statistics →Normality test



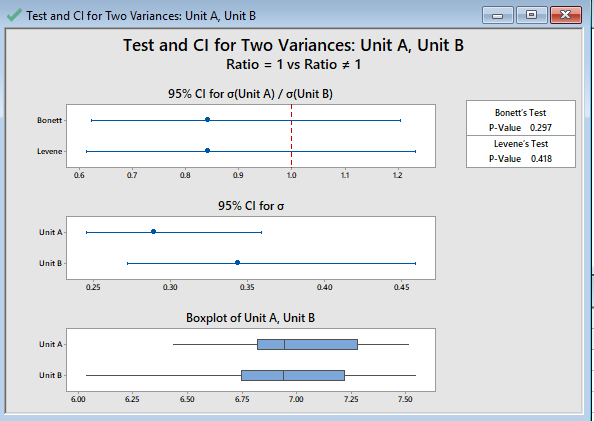
P-value is >0.05. P High Ho Fly.So data is Normally distributed

2) Variance Test

Create Hypothesis for variances of Unit A and Unit B

Ho= Variance of diameters of Unit A is equal to the variance of diameters of Unit B

Ha= Variance of diameters of Unit A is not equal to the variance of diameters of Unit B



P-value>0.05.

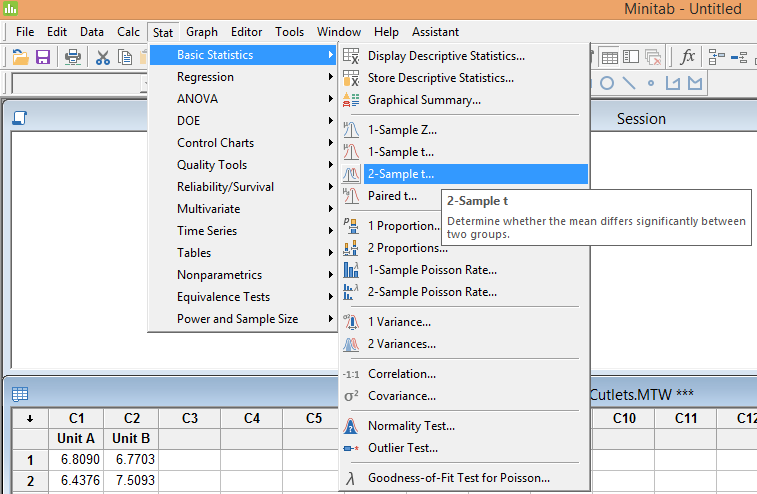
P High Ho fly. We fail to reject Null hypothesis. So we will accept it and hence Variances of A is equal to Variances of B.

3) As A and B are 2 Discrete variables and output variable diameter is a continuous, we will go with **2-sample T test**

*2 sample t- test is equality of Means test. Sample Mean will tell us which program is better.*

Ho= Averages of diameters of Unit A is equal to Averages of diameters of unit B

Ha= Averages of diameters of Unit A is not equal to Averages of diameters of unit B

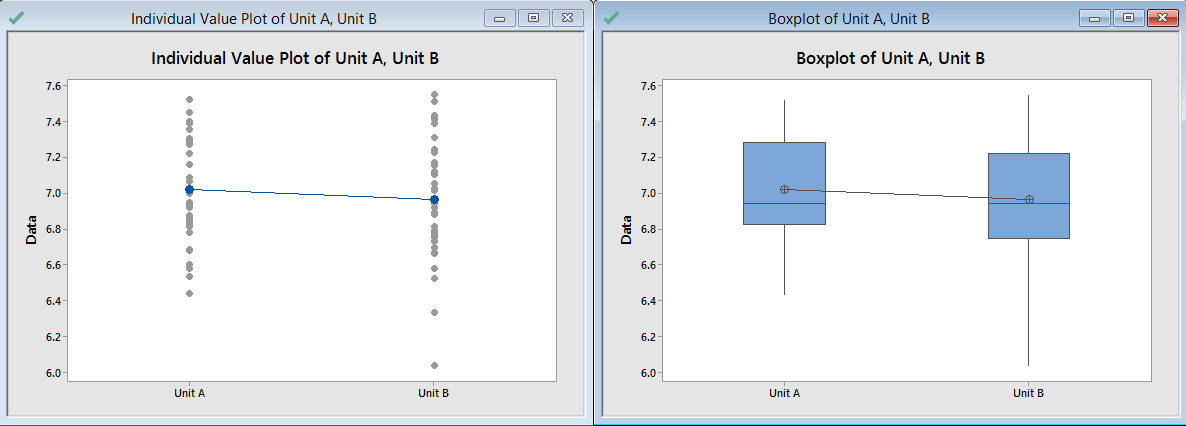


You can see the detailed summary displayed.

P-value>0.05 and hence P High and Ho Fly.

Inference is that **there is no significant difference in the diameters of Unit A and Unit B**

You can also view individual graphs and Box plot of the data points



Question 2

**ANOVA-One way test**

A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.

Solution Approach

α=0.05

Below is the sample data for average Turn Around Times(TAT)

Inputs are 4 lab reports. So **Input** is **Discrete** in **more than 2 categories**.

**Output**is **continuous**as we are trying to see the difference in average TAT.

we proceed with **ANOVA one-way test**

1) Normality test- We will see if data is normally distributed or not

Create hypothesis for Lab 1

Ho= Data is normally distributed

Ha=Data is not normally distributed

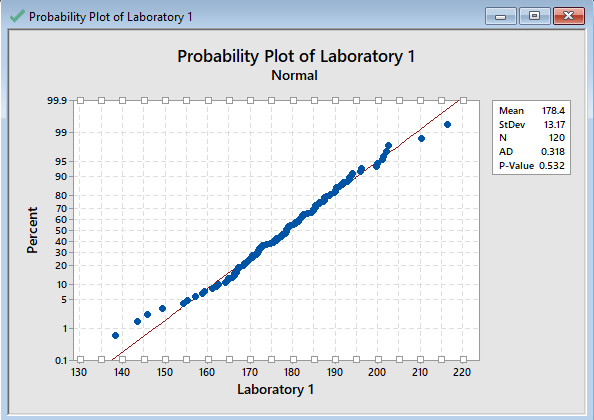
P-value is >0.05. P High Ho Fly.

So data is Normally distributed

Create hypothesis for Lab 3

Ho= Data is normally distributed

Ha=Data is not normally distributed



P-value is >0.05. P High Ho Fly.So data is normally distributed

Create hypothesis for Lab 2

Ho= Data is normally distributed

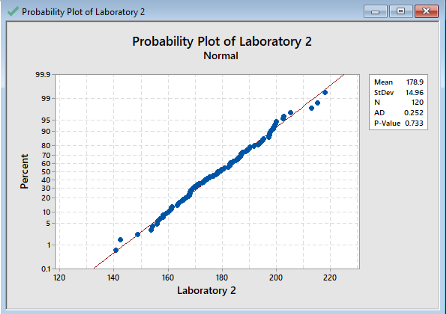
Ha=Data is not normally distributed

P-value is >0.05. P High Ho Fly.So data is normally distributed

Create hypothesis for Lab 2

Ho= Data is normally distributed

Ha=Data is not normally distributed



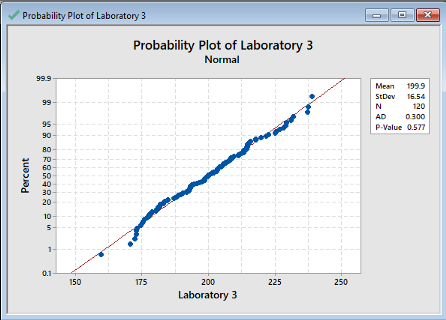
P-value is >0.05. P High Ho Fly.

So data is Normally distributed

Create hypothesis for Lab 3

Ho= Data is normally distributed

Ha=Data is not normally distributed



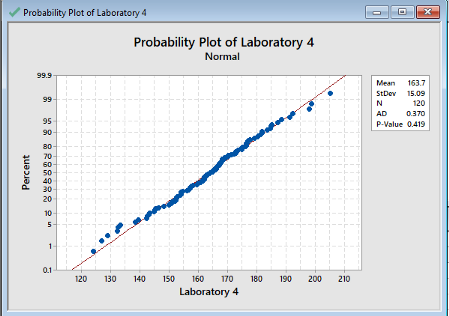
p-value is >0.05. P High Ho Fly.

So data is normally distributed

Create hypothesis for Lab 4

Ho= Data is normally distributed

Ha=Data is not normally distributed



P-value is >0.05. P High Ho Fly.

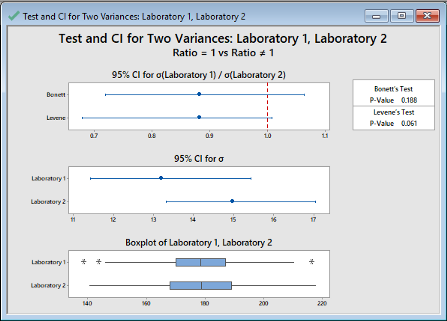
So data is normally distributed

2) Variance Test

Create Hypothesis for variances of Lab 1 and Lab 2

Ho= Variance of TAT of Lab 1 is equal to variance of TAT of Lab 2

Ha= Variance of TAT of Lab 1 is not equal to variance of TAT of Lab 2



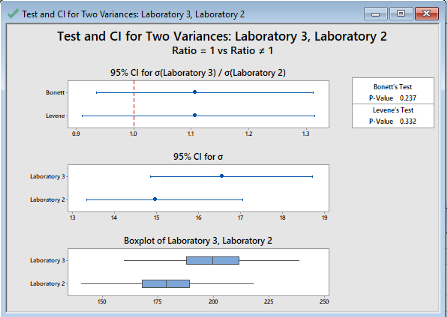
p-value>0.05. P High Ho fly. We fail to reject Null hypothesis.

So we will accept it and hence Variances of 1 is equal to variances of 2

Create Hypothesis for variances of Lab 2 and Lab 3

Ho= Variance of TAT of Lab 2 is equal to variance of TAT of Lab 3

Ha= Variance of TAT of Lab 2 is not equal to variance of TAT of Lab 3



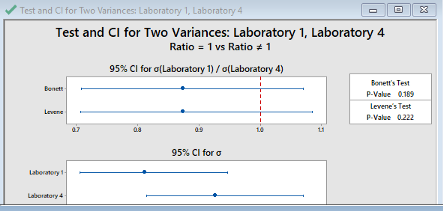
P-value>0.05. P High Ho fly. We fail to reject Null hypothesis.

So we will accept it and hence Variances of lab 3 is equal to variances of lab 4.

Create Hypothesis for variances of Lab 4 and Lab 1

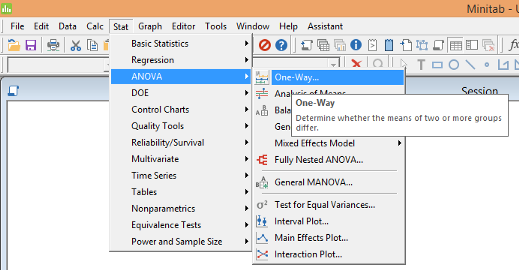
Ho= Variance of TAT of Lab 4 is equal to variance of TAT of Lab 1

Ha= Variance of TAT of Lab 4 is not equal to variance of TAT of Lab 1



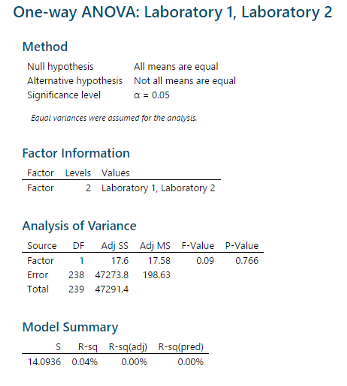
P-value>0.05. P High Ho fly. We fail to reject Null hypothesis. So we will accept it and hence Variances of lab 4 is equal to variances of lab 1

3) As there are more than 2 discrete variables and output variable TAT is a continuous variable. Hence we will go with **Anova one way test**.



Ho= Average TAT for all the samples is same

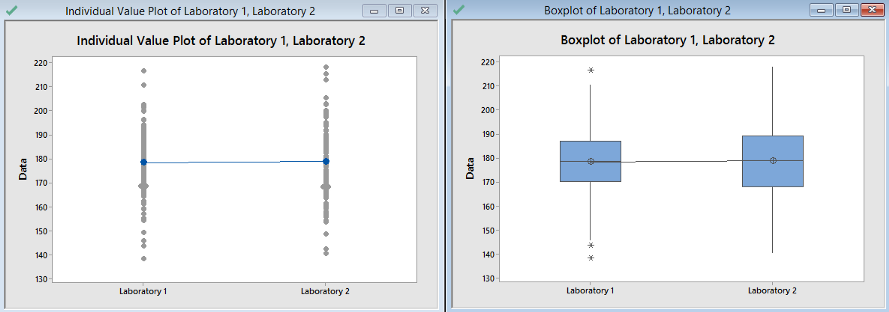
Ha= Averages TAT for all the samples is not same



SImilarly by doing for different lab combinations you can see that P -value is > 0.05. P High and Ho Fly.

Hence **there is no significant difference in the average TAT for all the labs.**

You can also view the individual and box plots of the data points



**Question 4**

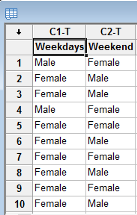
**2-Proportion test**

Pantaloons Sales managers commented that *%*of males versus females walking in to the store differ based on day of the week. Analyze the data and determine whether there is evidence at *5 %*significance level to support this hypothesis.

*Solution Approach*

α=0.05

Below is the sample data for Male and Female on weekday and weekends



***Inputs***are **2 discrete variables**.

***Output***is **Discrete**as we are trying to find out if proportions of male and female walking in to the store is same or not

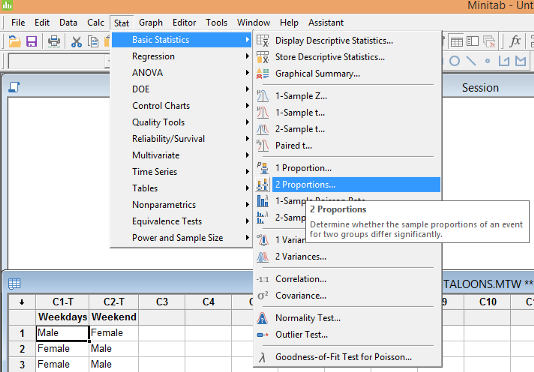
We proceed with **2-proportion test**

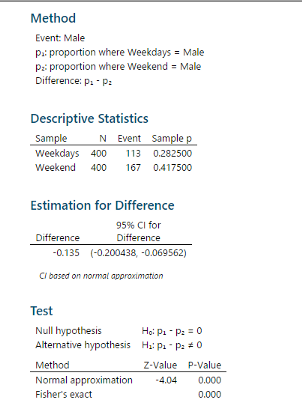
Create hypothesis

Ho= Proportions of Male and Female are same

Ha= Proportions of Male and Female are not same

Stat — >Basic Statistics— >2-proportion test





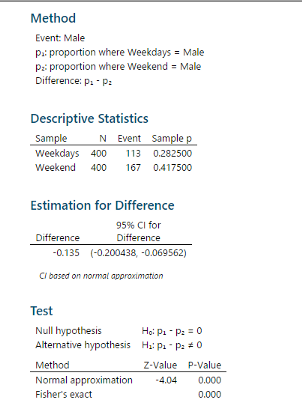
P-value is less than 0.05 and hence we fail to reject Null.Hence proportions of Male and Female are not same

Now we will try to find out whose proportion is higher. We create another hypothesis

Ho= Proportions of Male is less than or equal to Female

Ha= Proportions of Male is greater than Female

Stat->Basic Statistics->2 proportion test.

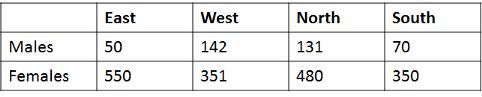


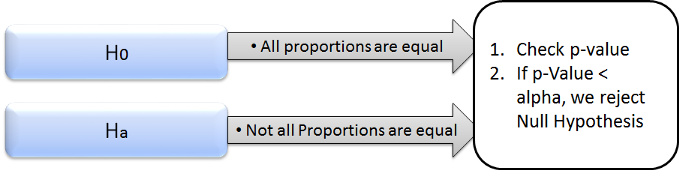
P-value <0.05 and hence we reject null.**Hence proportion of Male is greater than Female.**

**Question 3**

**Chi-Square Test**

Sales of products in four different regions is tabulated for males and females. Find if male-female buyer ratios are similar across regions.





**Inputs**are **4 discrete variables**(east,west,north,south).

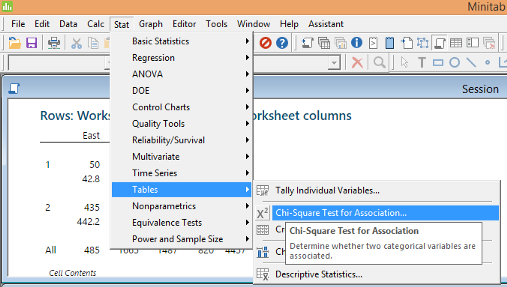
**Output**is also **discrete**. We are trying to find out if proportions of male and female are similar or not across the regions

We proceed with **chi-square test**

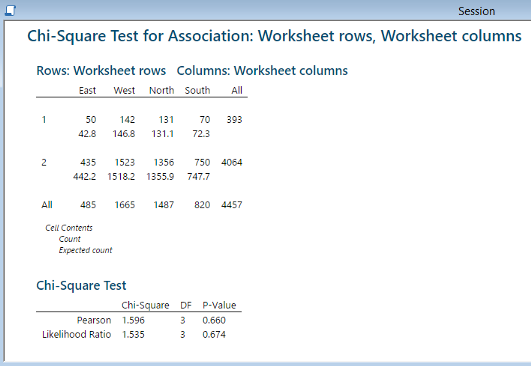
Create hypothesis

Ho= Proportions of Male and Female are same

Ha= Proportions of Male and Female are not same



Stat — >Tables — >chi-square test for Association



P-value>0.05.Hence we fail to reject Null.

**Hence proportion of male and female across regions is same.**