1. Diameter of cutlet=y (continuous variable)

Number of units=x (discrete variable)

1. **Normality Test**

Unit A

Ho : Data is Normal

Ha : Data is not Normal

Here alpha = 0.05

So, now conducting the shapiro.test for testing the normality of data sets

**🡪x=df[["Unit A"]]**

**🡪stats.shapiro(x)**

P = 0.3199819028

P>0.05 => p high Ho 🡪 So unit A data is Normal

Unit B

Ho : Data is Normal

Ha : Data is not Normal

**🡪y=df[["Unit B"]]**

**🡪stats.shapiro(y)**

P = 0.52249854803

p>0.05 => p high Ho => So Unit B data is also Normal

Now, the external conditions are not same for both the units so going for the variance test and checking whether both the variances are equal or not

x=df["Unit A"]

y=df["Unit B"]

1. **Variance test**

Ho : variance of unit A = variance of unit B

Ha: variance of unit A != variance of unit B

Using levene variance test

🡪stats.levene(x,y)

P = 0.41761

P>0.05 => p high Ho 🡪 This means the variances of both the data sets are equal.

Since the variances are equal, we will go for two sample-t for equal variances

1. **2 sample-t test for equal variance**

Ho : No there is not a significance difference between the diameters

Ha : Yes, there is significant difference between the diameters

🡪 t, p = stats.ttest\_ind(df["Unit A"],df["Unit B"])

p = 0.4722

p>0.05 => p high Ho fly => this means there is no significant difference the diameters of cutlets from both the units.

1. y – Average Turn Around Time (TAT) (Continuous Variable)

x – Laboratories, 4 (Discrete Variable)

1. **Normality Test**

Ho : All data are normal

Ha: Data are not normal

Applying shapiro.test to data of all the 4 laboratories and finding their p values

P1= 0.5506953 P2= 0.86375 P3= 0.4205, P4 = 0.6619

All P > 0.05 🡪 P (Ho) 🡪 data is normal

1. **Levene’s test**

Ho : Variances are equal

Ha : Variances are not equal

🡪stats.levene(df1,df2,df3,df4)

p = 0.05161

p>0.05 => p high Ho fly => variances are equal

Since variances are equal, so now going for the one way ANOVA test

1. **ANOVA test**

Ho: There is no difference in average TAT among the different laboratories

Ha: There is difference between in average TAT among the different laboratories

🡪stat, p = f\_oneway(df["Laboratory 1"], df["Laboratory 2"],df["Laboratory 3"],df["Laboratory 4"])

stats,p

p = 2.1e-57

p<0.05 🡪 p low Ho go 🡪 There is difference between the average TAT among the different laboratories

1. Y – Male-Female buyer ratio

X – Four different regions

Y and X are both discrete variable

Ho: All proportions are equal

Ha: Not all proportions are equal

1. Y – Order form is defective or error free (Discrete Variable)

X – 4 Centres (Discrete Variable)

Here number of independent variable are greater than 2. So will go for chi squared test

Ho : The defects made are equal by each centre

Ha : At least defect of one centre varies from the other centre

1. **Chi squared test**

stat, p, dof, expected = chi2\_contingency(df\_new)

p = 0.2771

P>0.05 => p high Ho fly => The defects made are equal by each centre

4b) Y – Male Female ratio

X- Weekdays and weekends (2 discrete variables)