TeleCall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in order form renders it defective and has to be reworked before processing. The manager wants to check whether the defective % varies by centre. Please analyze the

data at *5%* significance level and help the manager draw appropriate inferences

File: **CustomerOrderForm.csv**

**Business Problem:** Find out whether the defective percentage varies by centre.

Inputs are Phillippines, Indonesia, Malta and India that is **Discrete in four categories,** Outputis **Discrete.**.

We proceed with **chi-square test.**

🡪Create hypothesis

#Ho= % of defective of all countries are equal

#Ha =at lest one defective % is not equal

**# p-value = 0.2771 > 0.05 accept null hypothesis**

**Inferences:**

🡪Inference is that defective percentage is same in all countries there is no need of any action.

**R-Code:**

# Load the Dataset

**install.packages("readxl")**

**library(readxl)**

#Ho= % of defective of all countries are equal

#Ha =at lest one defective % is not equal

**cof<-read\_excel(file.choose())**

**View(cof)** # countries are in their own columns; so we need to stack the data

**stacked\_cof<-stack(cof)**

**attach(stacked\_cof)**

**View(stacked\_cof)**

**table(stacked\_cof$ind,stacked\_cof$values)**

**chisq.test(table(stacked\_cof$ind,stacked\_cof$values))**

# p-value = 0.2771 > 0.05 accept null hypothesis

#i.e percentage defective in all countries are equal

#inferences : as defective percentage is same in all countries there is no need of any action

**Python code:**