

In [2]:

```

import pandas as pd
import glob

df=pd.read_csv(r'C:\Users\Lohit\Desktop\krr\feedback.csv')
df=df.iloc[:,2:4]

path='C:/Users/Lohit/Desktop/krr/GroupA/raw data/'
all_files=glob.glob(path+"/*.csv")

#traversing all files stored in all_files
for file in all_files:
    data=pd.read_csv(file)
    course=data.iloc[1,0]
    course=course[:course.find("-")] # extracting the course code
    teacher=data.iloc[2,1] # extracting the teacher name
    data=data.iloc[6:,17:] # extracting columns----> wt,avg,prod
    data=data.dropna() # removing the rows that contains Nan values
    data=data.reset_index(drop=True) # reset index

    Teacher=[teacher for _ in range(len(data.axes[0]))]
    Course=[course for _ in range(len(data.axes[0]))]
    typee=list(df.iloc[:len(data.axes[0]),0].values)
    question=list(df.iloc[:len(data.axes[0]),1].values)
    wt=list(data.iloc[:,0].values)
    avg=list(data.iloc[:,1].values)
    prod=list(data.iloc[:,2].values)

    new_d=pd.DataFrame({'Teacher':Teacher, 'Course':Course , '2':typee, '3':question
    new_d.to_csv(r'C:\Users\Lohit\Desktop\krr\final_feedback.csv',mode='a',index=
    print(new_d)

```

	Teacher	Course	2	\
0	Sayed Yousuf	CS402	Course Content and organization	
1	Sayed Yousuf	CS402	Course Content and organization	
2	Sayed Yousuf	CS402	Course Content and organization	
3	Sayed Yousuf	CS402	Teacher Evaluation	
4	Sayed Yousuf	CS402	Teacher Evaluation	
5	Sayed Yousuf	CS402	Teacher Evaluation	
6	Sayed Yousuf	CS402	Teacher Evaluation	
7	Sayed Yousuf	CS402	Teacher Evaluation	
8	Sayed Yousuf	CS402	Teacher Evaluation	
9	Sayed Yousuf	CS402	Teacher Evaluation	
10	Sayed Yousuf	CS402	Teacher Evaluation	
11	Sayed Yousuf	CS402	Teacher Evaluation	
12	Sayed Yousuf	CS402	Teacher Evaluation	
13	Sayed Yousuf	CS402	Teacher Evaluation	
14	Sayed Yousuf	CS402	Teacher Evaluation	
15	Sayed Yousuf	CS402	Teacher Evaluation	
16	Sayed Yousuf	CS402	Teacher Evaluation	
17	Sayed Yousuf	CS402	Student Contribution	
18	Sayed Yousuf	CS402	Student Contribution	

In []: