/\* Importing the data from Excel\*/

PROC IMPORT OUT= work.FisrtPrinciples DATAFILE= "C:\Users\Ritika Khandelwal\Desktop\RA\practice.xlsx"

 DBMS=xlsx REPLACE;

 run;

/\* Choosing Optimal number of clustersthrough Hirerchal clustering (Eigenvalues)\*/

ods graphics on;

proc cluster data=work.FisrtPrinciples method=ward ccc pseudo print=10 outtree=Tree;

run;

ods graphics off;

/\* As per CCC and pseudo F, the optimal number of clusters should be 4. And as per pseudo T-squared , the optimal number of clusters is 3. Thus, the

optimal number of clusters can be either 3 or 4.\*/

/\* Choosing Optimal number of clusters\*/

%macro doFASTCLUS;

 %do k= 3 %to 6;

 proc fastclus data= work.FisrtPrinciples out= fcOut maxiter= 100 converge= 0 maxclusters= &k

 summary;

 run;

 %end;

%mend;

%doFASTCLUS

/\* The value of R squared does not increase a lot, when the optimal number of clusters is 4.\*/

/\* Optimal number of clusters is 3. Sizes of the cluster are 461, 218 and 321.\*/

/\* K means clustering for k = 4 \*/

proc fastclus data=work.FisrtPrinciples maxclusters=4 maxiter=100 converge=0

mean=mean out=work.result;

run;