**Experiment No : 10**

**Aim:** To perform Principle Component Analysis (PCA) using Alphabet dataset with 24 features.

**Theory*:*** Principal Components Analysis (PCA) is an algorithm to transform the columns of a dataset into a new set of features called Principal Components. By doing this, a large chunk of the information across the full dataset is effectively compressed in fewer feature columns. This enables dimensionality reduction and ability to visualize the separation of classes or clusters if any.Principal Component Analysis is a dimensionality-reduction, it also transforms a large set of variables into a smaller one that still contains most of the information in the large set being completely different from the original dataset.

STEP BY STEP EXPLANATION OF PCA

### STEP 1: STANDARDIZATION

### STEP 2: COVARIANCE MATRIX COMPUTATION

### STEP 3: COMPUTE THE EIGENVECTORS AND EIGENVALUES OF THE COVARIANCE MATRIX TO IDENTIFY THE PRINCIPAL COMPONENTS

**Dataset used :** Alphabet 24 feature dataset.

**Code:**

clc;

close all;

clear all;

x = csvread('Arjun \_exp9.csv');

[pcs,scrs,~,~,pexp] = pca(x);

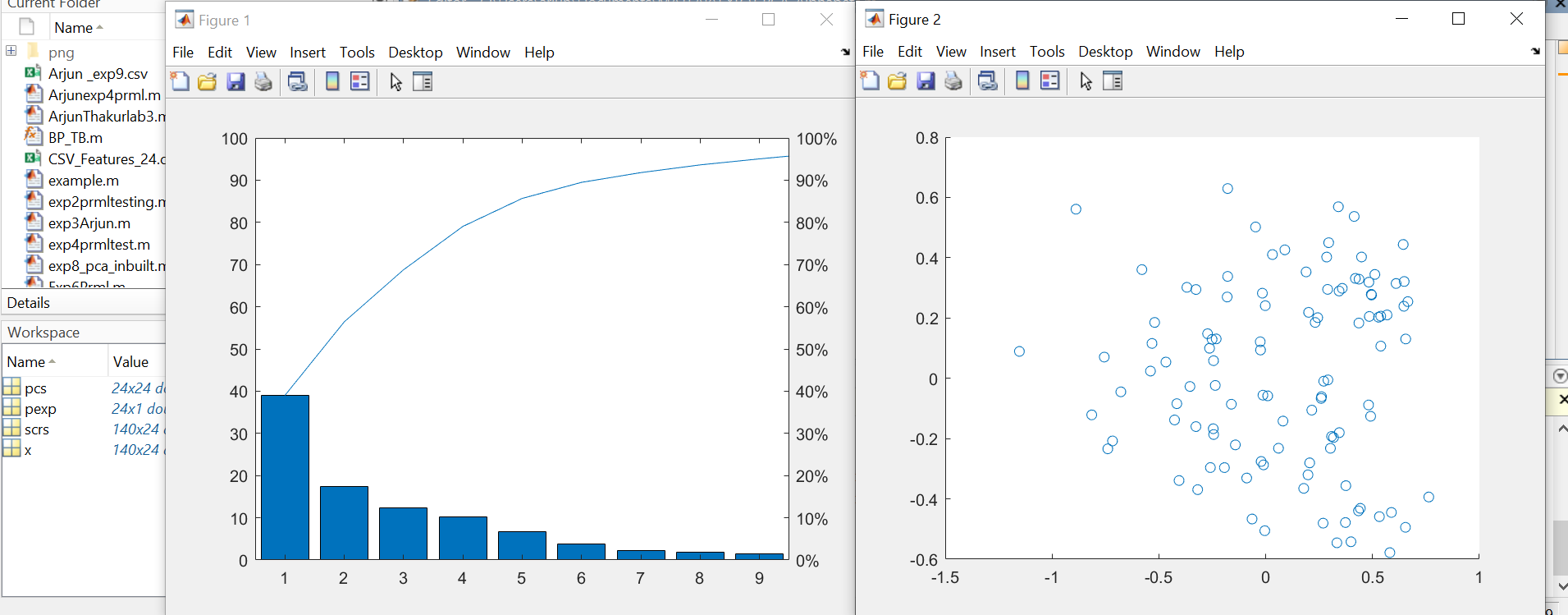
figure(1);

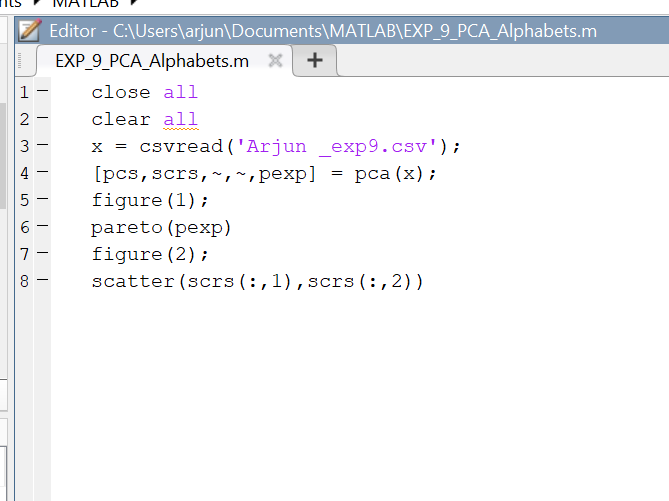
pareto(pexp)

figure(2);

scatter(scrs(:,1),scrs(:,2))

**Output :**





**Conclusion :** After performing this experiment of Principle Component Analysis (PCA) using Alphabet dataset, I observed the plot between Principle Component vs Percentage of experienced variances. Also the first principal component accounts for the largest possible variance in the data set. From the plot I observed that 95% and 1% relativity are present from the different groups with principle components. From fig 2 it is observed that maximum clusters lie between -0.5 to 0.5.