```sas
options pageno=1 nodate; run;

title1 "K-MEANS CLUSTERING OF FISHER'S IRIS DATA";
title2 ;
title3 ;
proc format;
    value Specname 1='Setosa    '  2='Versicolor'  3='Virginica ';
run;

data Iris;
    input Species SepalLength SepalWidth PetalLength PetalWidth @@;
    format Species specname.;
    label SepalLength='Sepal Length in mm.' ;
    label SepalWidth ='Sepal Width in mm.';
    label PetalLength='Petal Length in mm.';
    label PetalWidth ='Petal Width in mm.';
    symbol = put(species, specname10.);
    datalines;
1 5.1 3.5 1.4 0.2   1 4.9 3.0 1.4 0.2   1 4.7 3.2 1.3 0.2   1 4.6 3.1 1.5 0.2
1 5.0 3.6 1.4 0.2   1 5.4 3.9 1.7 0.4   1 4.6 3.4 1.4 0.3   1 5.0 3.4 1.5 0.2
1 4.4 2.9 1.4 0.2   1 4.9 3.1 1.5 0.1   1 5.4 3.7 1.5 0.2   1 4.8 3.4 1.6 0.2
1 4.8 3.0 1.4 0.1   1 4.3 3.0 1.1 0.1   1 5.8 4.0 1.2 0.2   1 5.7 4.4 1.5 0.4
1 5.4 3.9 1.3 0.4   1 5.1 3.5 1.4 0.3   1 5.7 3.8 1.7 0.3   1 5.1 3.8 1.5 0.3
1 5.4 3.4 1.7 0.2   1 5.1 3.7 1.5 0.4   1 4.6 3.6 1.0 0.2   1 5.1 3.3 1.7 0.5
1 4.8 3.4 1.9 0.2   1 5.0 3.0 1.6 0.2   1 5.0 3.4 1.6 0.4   1 5.2 3.5 1.5 0.2
1 5.2 3.4 1.4 0.2   1 4.7 3.2 1.6 0.2   1 4.8 3.1 1.6 0.2   1 5.4 3.4 1.5 0.4
1 5.2 4.1 1.5 0.1   1 5.5 4.2 1.4 0.2   1 4.9 3.1 1.5 0.2   1 5.0 3.2 1.2 0.2
1 5.5 3.5 1.3 0.2   1 4.9 3.6 1.4 0.1   1 4.4 3.0 1.3 0.2   1 5.1 3.4 1.5 0.2
1 5.0 3.5 1.3 0.3   1 4.5 2.3 1.3 0.3   1 4.4 3.2 1.3 0.2   1 5.0 3.5 1.6 0.6
1 5.1 3.8 1.9 0.4   1 4.8 3.0 1.4 0.3   1 5.1 3.8 1.6 0.2   1 4.6 3.2 1.4 0.2
1 5.3 3.7 1.5 0.2   1 5.0 3.3 1.4 0.2   2 7.0 3.2 4.7 1.4   2 6.4 3.2 4.5 1.5
2 6.9 3.1 4.9 1.5   2 5.5 2.3 4.0 1.3   2 6.5 2.8 4.6 1.5   2 5.7 2.8 4.5 1.3
2 6.3 3.3 4.7 1.6   2 4.9 2.4 3.3 1.0   2 6.6 2.9 4.6 1.3   2 5.2 2.7 3.9 1.4
2 5.0 2.0 3.5 1.0   2 5.9 3.0 4.2 1.5   2 6.0 2.2 4.0 1.0   2 6.1 2.9 4.7 1.4
2 5.6 2.9 3.6 1.3   2 6.7 3.1 4.4 1.4   2 5.6 3.0 4.5 1.5   2 5.8 2.7 4.1 1.0
2 6.2 2.2 4.5 1.5   2 5.6 2.5 3.9 1.1   2 5.9 3.2 4.8 1.8   2 6.1 2.8 4.0 1.3
2 6.3 2.5 4.9 1.5   2 6.1 2.8 4.7 1.2   2 6.4 2.9 4.3 1.3   2 6.6 3.0 4.4 1.4
2 6.8 2.8 4.8 1.4   2 6.7 3.0 5.0 1.7   2 6.0 2.9 4.5 1.5   2 5.7 2.6 3.5 1.0
2 5.5 2.4 3.8 1.1   2 5.5 2.4 3.7 1.0   2 5.8 2.7 3.9 1.2   2 6.0 2.7 5.1 1.6
2 5.4 3.0 4.5 1.5   2 6.0 3.4 4.5 1.6   2 6.7 3.1 4.7 1.5   2 6.3 2.3 4.4 1.3
2 5.6 3.0 4.1 1.3   2 5.5 2.5 4.0 1.3   2 5.5 2.6 4.4 1.2   2 6.1 3.0 4.6 1.4
2 5.8 2.6 4.0 1.2   2 5.0 2.3 3.3 1.0   2 5.6 2.7 4.2 1.3   2 5.7 3.0 4.2 1.2
2 5.7 2.9 4.2 1.3   2 6.2 2.9 4.3 1.3   2 5.1 2.5 3.0 1.1   2 5.7 2.8 4.1 1.3
3 6.3 3.3 6.0 2.5   3 5.8 2.7 5.1 1.9   3 7.1 3.0 5.9 2.1   3 6.3 2.9 5.6 1.8
3 6.5 3.0 5.8 2.2   3 7.6 3.0 6.6 2.1   3 4.9 2.5 4.5 1.7   3 7.3 2.9 6.3 1.8
3 6.7 2.5 5.8 1.8   3 7.2 3.6 6.1 2.5   3 6.5 3.2 5.1 2.0   3 6.4 2.7 5.3 1.9
3 6.8 3.0 5.5 2.1   3 5.7 2.5 5.0 2.0   3 5.8 2.8 5.1 2.4   3 6.4 3.2 5.3 2.3
3 6.5 3.0 5.5 1.8   3 7.7 3.8 6.7 2.2   3 7.7 2.6 6.9 2.3   3 6.0 2.2 5.0 1.5
3 6.9 3.2 5.7 2.3   3 5.6 2.8 4.9 2.0   3 7.7 2.8 6.7 2.0   3 6.3 2.7 4.9 1.8
3 6.7 3.3 5.7 2.1   3 7.2 3.2 6.0 1.8   3 6.2 2.8 4.8 1.8   3 6.1 3.0 4.9 1.8
3 6.4 2.8 5.6 2.1   3 7.2 3.0 5.8 1.6   3 7.4 2.8 6.1 1.9   3 7.9 3.8 6.4 2.0
3 6.4 2.8 5.6 2.2   3 6.3 2.8 5.1 1.5   3 6.1 2.6 5.6 1.4   3 7.7 3.0 6.1 2.3
3 6.3 3.4 5.6 2.4   3 6.4 3.1 5.5 1.8   3 6.0 3.0 4.8 1.8   3 6.9 3.1 5.4 2.1
3 6.7 3.1 5.6 2.4   3 6.9 3.1 5.1 2.3   3 5.8 2.7 5.1 1.9   3 6.8 3.2 5.9 2.3
3 6.7 3.3 5.7 2.5   3 6.7 3.0 5.2 2.3   3 6.3 2.5 5.0 1.9   3 6.5 3.0 5.2 2.0
3 6.2 3.4 5.4 2.3   3 5.9 3.0 5.1 1.8
;
run;
data irisone;
    set iris;
    SepalLength = SepalLength;
    SepalWidth = SepalWidth;
    PetalLength = PetalLength;
    PetalWidth = PetalWidth;
run;
proc contents data=iris; run;
proc univariate data=Iris normaltest;
    qqplot SepalLength SepalWidth PetalLength PetalWidth;
    hist SepalLength SepalWidth PetalLength PetalWidth;
run;


proc discrim data=Iris pool = test testdata=irisone out=iriso outcross=irisoc outd=irisod outstat=irisos crossvalidate ;
```

```sas
   class Species;
   var SepalLength SepalWidth PetalLength PetalWidth;
   priors proportional;

run;

proc print data=irisod;
run;
proc print data=irisos;
run;
proc discrim data=Iris pool = yes testdata=irisone crossvalidate noprint;
   class Species;
   var SepalLength SepalWidth PetalLength PetalWidth;
   priors proportional;
run;


/*

title1 "K-MEANS CLUSTERING OF FISHER'S IRIS DATA";
title2 "Scatter Plots and Correlations for Measured Variables";
proc corr data=Iris plots=matrix(nvar=all histogram);
   var SepalLength SepalWidth PetalLength PetalWidth;
run;

title1 "Plots of Measured Variables with Species Identified";
proc sgplot data=Iris;
  scatter X=SepalLength Y=SepalWidth / Group=Species;
run;

proc sgplot data=Iris;
  scatter X=PetalLength Y=PetalWidth / Group=Species;
run;

title1 "K-MEANS CLUSTERING OF FISHER'S IRIS DATA";
title2 "No more than 2 clusters";
proc fastclus data=Iris maxc=2 maxiter=10 out=clus2;
   var SepalLength SepalWidth PetalLength PetalWidth;
run;

proc freq data=clus2;
   tables cluster*species / norow nocol;
run;

title2 "No more than 3 clusters";
proc fastclus data=Iris maxc=3 maxiter=10 out=clus3;
   var SepalLength SepalWidth PetalLength PetalWidth;
run;

proc freq data=clus3;
   tables cluster*Species / norow nocol;
run;

title2 "No more than 4 clusters";
proc fastclus data=Iris maxc=4 maxiter=10 out=clus3;
   var SepalLength SepalWidth PetalLength PetalWidth;
run;

proc freq data=clus3;
   tables cluster*Species / norow nocol;
run;

title2 "Using ACECLUS to Improve Accuracy";
proc aceclus data=sashelp.iris out=ace p=.02 outstat=score;
   var SepalLength SepalWidth PetalLength PetalWidth ;
run;

proc sgplot data=ace;
   scatter y=can2 x=can1 / group=Species;
   keylegend / title="Species";
run;

proc sgplot data=ace;
   scatter y=can4 x=can3 / group=Species;
   keylegend / title="Species";
```

```
run;

proc fastclus data=ace maxc=3 maxiter=10 conv=0 out=clus;
    var can:;
run;

proc freq;
    tables cluster*Species / norow nocol;
run;
*/
```