

Illustration of Principal Component and Cluster Analysis

Among the employed work force in 26 countries in Europe in 1979, percentage employed in different industries like agriculture, mining, manufacturing, power supplies, construction, service industries, finance, social and personal services, and transport and communications are given in the data file `data`. The objective is to see if countries can be classified according to their economies (agrarian, industrialized, service oriented etc.) and then cluster them accordingly.

```
> d<-read.table("data",header=T)
```

```
> d
```

	Country	Agr	Min	Man	PS	Con	SI	Fin	SPS	TC
1	Belgium	3.3	0.9	27.6	0.9	8.2	19.1	6.2	26.6	7.2
2	Denmark	9.2	0.1	21.8	0.6	8.3	14.6	6.5	32.2	7.1
.....										
.....										
25	USSR	23.7	1.4	25.8	0.6	9.2	6.1	0.5	23.6	9.3
26	Yugoslavia	48.7	1.5	16.8	1.1	4.9	6.4	11.3	5.3	4.0

```
> pr<-princomp(~Agr+Min+Man+PS+Con+SI+Fin+SPS+TC,data=d)
```

```
> pr
```

Call:

```
princomp(formula = ~Agr + Min + Man + PS + Con + SI + Fin + SPS +  
          TC, data = d)
```

Standard deviations:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
17.08176356	6.48234700	3.82393204	2.32861792	1.53278255	1.00289626	
	Comp.7	Comp.8	Comp.9			
0.63612956	0.24985891	0.04287707				

9 variables and 26 observations.

```
> pr$loadings
```

Loadings:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9
Agr	0.892		0.118		0.180	-0.153			0.335
Min						0.456	-0.766	0.290	0.324
Man	-0.271	0.770	0.185		0.336	-0.201	0.162		0.337
PS						0.231		-0.909	0.340
Con					-0.724	-0.558	-0.194		0.325
SI	-0.192	-0.234	-0.580	-0.608	0.266			0.104	0.337
Fin		-0.130	-0.470	0.781	0.121			0.123	0.334
SPS	-0.298	-0.567	0.598		0.236	-0.248			0.332
TC			0.159		-0.435	0.546	0.567	0.224	0.334

```

> pr$sdev^2/sum(pr$sdev^2)
      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6
8.157836e-01 1.174827e-01 4.088179e-02 1.516024e-02 6.568567e-03 2.812041e-03
      Comp.7      Comp.8      Comp.9
1.131360e-03 1.745417e-04 5.139960e-06

> cumsum(pr$sdev^2)/sum(pr$sdev^2)
      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6      Comp.7      Comp.8
0.8157836 0.9332663 0.9741481 0.9893084 0.9958769 0.9986890 0.9998203 0.9999949
      Comp.9
1.0000000

> pr$scores
      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6
1  -17.516687 -4.92622849 -2.35528094 -0.1940007  0.49072745  0.58725632
2  -11.496688 -11.66176637  3.00202830  2.5328564 -0.26205034 -1.00014269
3   -9.128686 -2.16828207 -2.75030568  0.1289183  0.37355869 -0.74062040
4  -14.393424  5.04749385  0.20568951  1.2143729  2.57857191 -0.26124048
5   4.458174 -6.13156498 -1.92400082 -3.4809060  0.34805906  0.53989717
6  -4.026684 -0.38889529 -2.40586194 -4.8100625 -0.24693944 -1.50206770
7  -12.089752  2.33236877 -4.62806669 -1.9653471 -0.03246954  1.08490154
8  -13.900455 -9.72359023 -1.70981367  0.6817669 -1.51095027 -0.83583339
9  -18.728675 -3.33178946  0.58938403  1.0073981  2.29959070  0.44287447
10 -6.471418  3.35662962 -4.75660272 -1.2741444 -0.51798957  0.83150513
11 -6.837047 -3.97634061  0.06757235  0.9640481  0.27383594  0.68129276
12 25.427083 -1.80467718 -2.91613130 -2.8567757 -1.87953959  0.59887931
13 -10.972019 -8.85877780  0.22621023 -0.5923626 -2.00656646  1.41674308
14  9.403865 -0.08570061 -1.23656256 -2.1803190  0.06780428 -1.03041185
15  5.774973  6.15867547 -4.87904446  4.4809274 -3.04038851 -1.70355376
16 -15.311975 -8.52674423  3.92210148  2.6339347  1.41721975 -0.73494415
17 -12.683839  9.77920054 -5.68921238 -1.0438972  1.24534771 -1.16079249
18 52.115644 -8.64165980  2.96515501 -1.7960412  3.17620617 -0.36536971
19  4.156791  6.70685051  4.93995679 -0.1372004  0.56054055 -0.54824453
20 -3.246127  9.23467980  3.78225558 -0.1477919 -0.08257234  0.24180861
21 -17.415527 10.73233092  4.89564722  0.2453801  1.36457552  1.27464580
22  3.135737  4.98695108  2.98354179 -0.7306784 -1.31541699  1.88083050
23 13.315709  2.94482700  3.58894681 -0.4854554 -1.36139215  0.32653024
24 17.011336  9.12523022  2.58152423  0.2380157 -0.03962236 -1.46062121
25  4.587043 -0.87197041  8.44875566  0.9788693 -2.93255455 -0.07846592
26 34.832648  0.69274975 -6.94788580  6.5884944  1.03241438  1.51514335
      Comp.7      Comp.8      Comp.9
1  0.02747862  0.261512512 -0.008040795
2 -0.16185051  0.113221471  0.098177832
3 -0.44855421 -0.072003249 -0.002822468
4  0.38250268  0.114442694 -0.043957458
5 -0.35750610 -0.420871836  0.072290598
6 -0.24928213  0.051645913  0.034174567

```

```

7  -1.70709926  0.657656083  0.021025084
8  -0.51429138 -0.273012947 -0.066889344
9  -0.44816057 -0.272105619 -0.033049026
10  0.50543393 -0.255764025 -0.015749373
11  0.92797085 -0.279485018  0.039399018
12  0.83983051  0.137603535 -0.016322341
13  0.88213892  0.371151398 -0.058993757
14  0.52381804 -0.075302827  0.017535977
15 -0.25749789 -0.039086940 -0.039716393
16 -0.03745853 -0.024572541 -0.018992003
17  1.25830466  0.004260537  0.008689042
18 -0.05968267  0.116399862 -0.058269133
19  0.22155473  0.184340613 -0.023526405
20 -0.57988066 -0.123332326 -0.065489414
21  0.43732622  0.146397327  0.046067998
22 -0.57747180 -0.571402883 -0.003437050
23 -0.66030520  0.017151170 -0.008766194
24 -0.55746312 -0.026902957  0.047287948
25  0.62537794  0.198893703  0.034657402
26 -0.01523307  0.059166352  0.044715689

```

```

> d$Country
[1] Belgium      Denmark      France      FRG          Ireland
[6] Italy         Luxembourg   Netherlands UK           Austria
[11] Finland      Greece       Norway      Portugal     Spain
[16] Sweden       Switzerland  Turkey      Bulgaria     Czechoslovakia
[21] GDR          Hungary      Poland      Rumania      USSR
[26] Yugoslavia

```

```

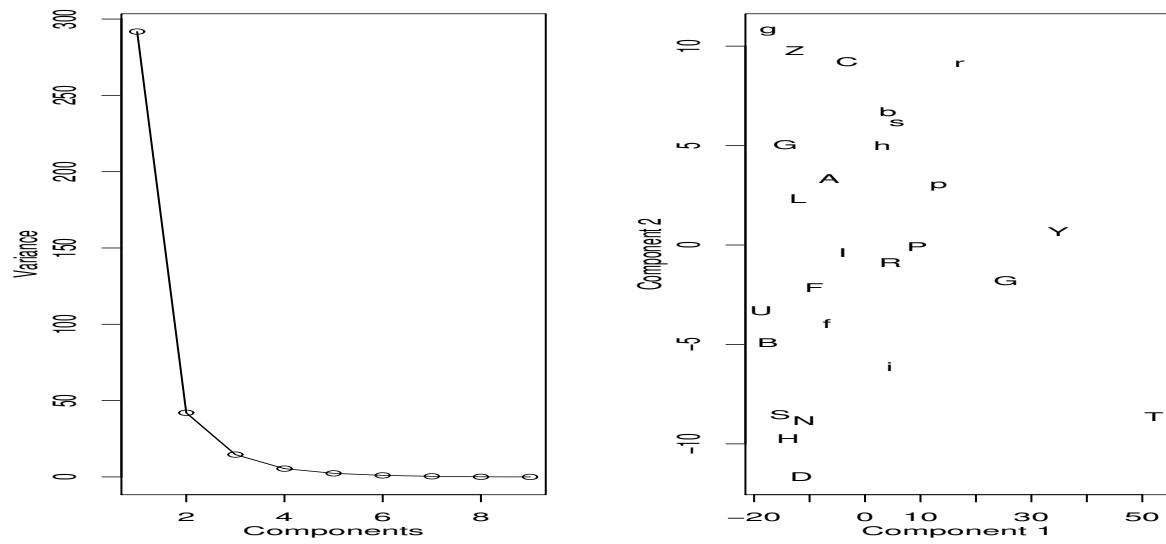
> lab<-c("B","D","F","G","i","I","L","H","U","A","f","G","N","P","s","S","Z",
         "T","b","C","g","h","p","r","R","Y")

```

```

> plot(1:9,pr$sdev^2,type="l",xlab="Components",ylab="Variance")
> points(1:9,pr$sdev^2)
> plot(pr2$scores[,1],pr2$scores[,2],pch=lab,xlab="Component 1",ylab="Component 2")

```



```
> plot(hclust(dist(d)),labels=lab,xlab="Country",main="Original Data")
> plot(hclust(dist(pr$scores)),labels=lab,xlab="Country",main="PCA Based")
```

