Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	9.76991794	5.55924638	0.2383	0.2383
2	4.21067156	1.10967667	0.1027	0.3410
3	3.10099489	0.21088690	0.0756	0.4166
4	2.89010799	0.33566625	0.0705	0.4871
5	2.55444173	0.42121307	0.0623	0.5494
6	2.13322866	0.48318506	0.0520	0.6014
7	1.65004360	0.11092305	0.0402	0.6417
8	1.53912055	0.21407287	0.0375	0.6792
9	1.32504768	0.03618664	0.0323	0.7116
10	1.28886104	0.10217738	0.0314	0.7430
11	1.18668367	0.10337894	0.0289	0.7719
12	1.08330473	0.22083611	0.0264	0.7984
13	0.86246862	0.04308193	0.0210	0.8194

Figure 3.1

Table 4.1

Percent Variation Accounted for by Principal Components				
Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	24.3127	24.3127	17.9289	17.9289
2	9.3944	33.7072	16.8741	34.8030
3	7.5689	41.2761	4.5546	39.3577
4	6.9016	48.1777	8.3076	47.6653
5	6.2652	54.4428	21.2193	68.8845
6	5.2087	59.6515	4.1018	72.9863
7	4.1137	63.7652	2.4089	75.3953
8	3.8467	67.6119	0.4374	75.8327
9	3.3105	70.9224	0.6715	76.5042
10	3.1992	74.1216	0.6043	77.1085
11	2.9555	77.0771	0.2489	77.3574
12	2.6252	79.7023	2.0086	79.3659

Table 4.2

Parameter Estimates for Centered and Scaled Data		
	x15	x44
Intercept	0.0000000000	0.0000000000
x3	-.0133992067	-.1172265831
x4	0.0040917839	-.0174523972
x5	-.0653282590	-.0274231487
x6	-.0405796543	-.0023219786
x7	0.0191761525	-.1997636606
x8	0.0723043281	0.0166726062
x9	0.0009777501	-.0479629779
x10	-.0390810463	-.0469358521
x11	-.0709611284	0.0004994246
x12	0.0211751531	-.0553683812
x13	-.0072329657	0.0538921920
x14	0.0121057973	-.1523472547
x16	-.0029639021	0.0621346943
x17	0.0490916798	0.0433013349
x18	0.0071442005	-.0804323983

x19	0.2858911237	0.0019973167
x20	0.2738183666	0.0270683634
x21	0.0160908705	-.0476337089
x22	-.0550909090	0.0779146165
x23	-.0399538017	-.0027642338
x24	0.0574203466	0.1217613006
x25	0.2014816594	-.0295551357
x26	0.0305617169	-.0643205559
x27	0.0167581371	0.0975245398
x28	0.0132922817	0.0204996820
x29	0.0187485884	0.1069616349
x30	0.0019149552	0.0561256636
x31	-.0183268147	0.0377626200
x32	0.0323085089	0.0030415060
x33	0.0663661901	0.0617721255
x34	-.0228345823	0.2597825450
x35	0.1745838899	0.0327914665
x36	-.0375765866	0.0111974195
x37	0.0171982052	-.0016445951

x38	0.0571704983	-.0605653341
x39	0.0593997374	-.0488772649
x40	-.0472721227	-.0207006159
x41	-.0802301394	-.0013660674
x42	-.0538179469	0.0303238932
x43	0.0375765866	-.0111974195

10	CL11	CL29	88	9.1723	
9	CL13	CL20	20	9.4946	
8	CL10	CL9	108	9.7884	
7	CL23	Singapore	3	11.888	
6	CL8	Nigeria	109	12.735	
5	CL6	CL7	112	13.362	
4	China	United States	2	13.717	
3	CL5	India	113	13.864	
2	CL3	Sierra Leone	114	15.363	
1	CL2	CL4	116	16.263	

Figure 5.1

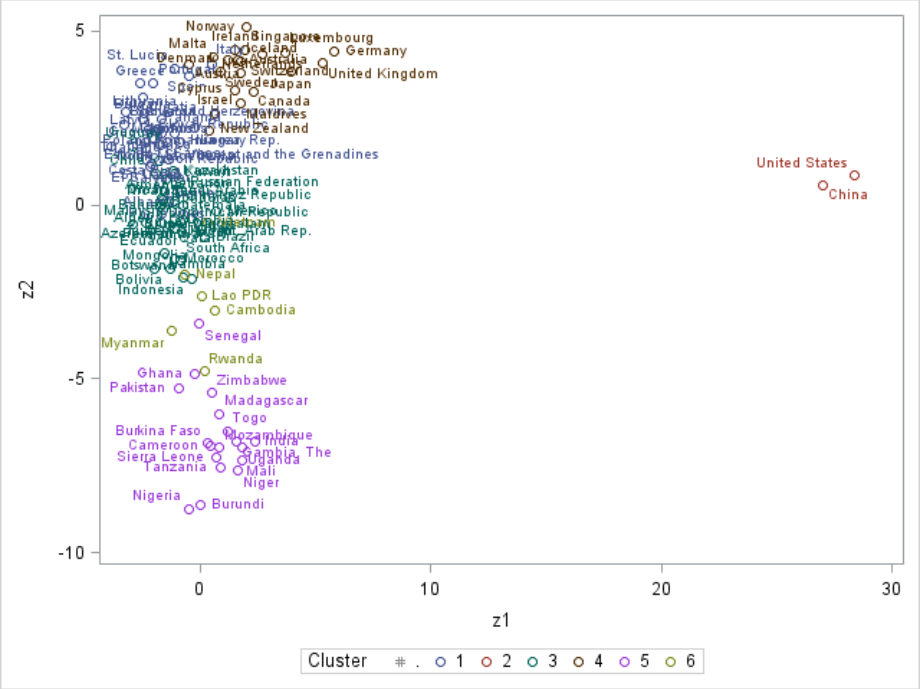

Figure 5.2

Cluster Summary

Cluster	Frequency	RMS Std Deviation	Maximum Distance from Seed to Observation	Radius Exceeded	Nearest Cluster	Distance Between Cluster Centroids
1	20	0.6811	8.4301	> Radius	3	3.5892
2	2	1.5815	7.2473	> Radius	1	11.8690
3	26	0.6993	6.7748	> Radius	1	3.5892
4	5	1.0833	10.1970	> Radius	5	6.3849
5	41	0.7482	11.5716	> Radius	3	4.0117
6	17	0.9485	10.9857	> Radius	1	5.8078
7	5	1.0476	10.3290	> Radius	3	6.5338

Figure 5.3

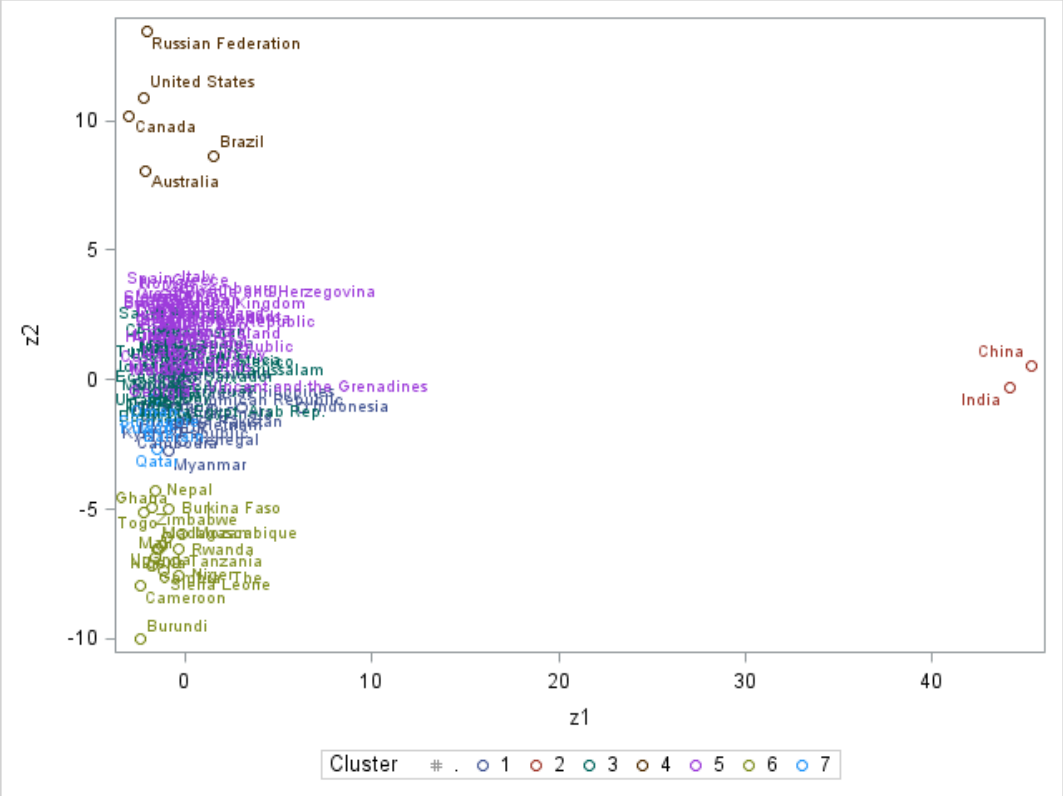

Figure 5.4

Table 6.1

Variance Explained by Each Factor											
Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	Factor11	Factor12
7.0467055	3.6997600	3.4109754	3.1348805	2.9376138	2.3605178	2.2649301	2.1334314	1.7885735	1.6352248	1.6167424	1.5489304

Table 6.2

	x1	Score ▾
1	Luxembourg	35.35628284
2	China	33.711449964
3	Qatar	28.700774316
4	Bahrain	25.350123823
5	Iceland	22.662570447
6	United States	20.45187946

Code:

Data standardize and PCA:

```
**Import data;
PROC IMPORT datafile = 'Countries data.xlsx'
      OUT = countries
      DBMS = excel
      replace
      ;
RUN;

**standardize data;
PROC STANDARD DATA=countries MEAN=0 STD=1 OUT=countries;
  VAR x3-x44 ;
RUN;

**print out data to ensure accurate steps;
proc print data = countries;
run;

**PCA;
proc princomp cov;
  var x3-x44;
run;
```

Multivariate PCA regression

```
proc standard data = work.country mean = 0 std =1 out = countries;
     var x3-x44;
run;
proc pls data = countries method = PCR nfac=12;
model x15 x44 = x3-x15 x16-x43/solution;
run;
```

Cluster analysis

```
RUN;
** 2.1 Hierarchical Clustering - Average Linkage;
proc cluster data=countries outtree=treecountry method=average nonorm;
   var x3-x44;
   id country;
run;
* 2.2 Hierarchical Clustering - Centroid;
proc cluster data=countries outtree=treecountry method=centroid nonorm;
   var x3-x44;
   id country;
run;
* K-means Clustering;
* 2.3 Principal Component Analysis;
proc princomp data=countries out=countriesPC;
var x3-x44;
run;
goptions reset=all;
symbol pointlabel=("#country") value=dot;
proc gplot data=countriesPC;
plot prin2*prin1/ vaxis=axis2 haxis=axis1 nolegend;
axis1 label=("z1" justify=center);
axis2 label=("z2" justify=center a=90);
run;


* 2.4 K-means - Random 6 observations;
proc fastclus data=countries maxc=6 replace=random maxiter=10
out=Clus_out1 radius=1;
var x3-x44;
id country;
run;
proc sort data=Clus_out1;
```

```
by cluster distance;

run;

proc print data=Clus_out1;

var country cluster distance;

run;

proc candisc data=Clus_out1 noprint out=ProCan1(keep=country cluster
Can1 Can2);

class cluster;

var x3-x44;

run;


proc sgplot data=ProCan1;

scatter y=Can2 x=Can1 / group=cluster datalabel=country;

label Can1="z1" Can2="z2";

run;


* 2.5 K-menas - First 7 observations;

proc fastclus data=countries radius=1.5 maxc=7 replace=none maxiter=10
out=Clus_out2;

var x3-x44;

id country;

run;

proc sort data=Clus_out2;

by cluster distance;

run;


proc candisc data=Clus_out2 noprint out=ProCan2(keep=country cluster
Can1 Can2);

class cluster;

var x3-x44;

run;


proc print data=Clus_out2;
```

```
var country cluster distance;

run;

proc sgplot data=ProCan2;

scatter y=Can2 x=Can1 / group=cluster datalabel=country;

label Can1="z1" Can2="z2";

run;


/*  Method # for getting seeds: Use Average Linkage to get cluster
centriods   */

proc cluster data=countries method=average outtree=CunTree noprint
nonorm;

var x3-x44;

id country;

run;

proc tree data=CunTree nclusters=6 out=newdata noprint;

id country;

copy x3-x44;

run;

proc sort data=newdata;

by cluster;

run;

proc print data=newdata;

var country cluster;

run;

proc means data=newdata;

by cluster;

output out=Seeds mean=x3-x44;

var x3-x44;

run;

proc fastclus data=countries maxc=7 maxiter=50 seed=Seeds
out=Clus_out4;

var x3-x44;

id country;
```

```
run;

proc sort data=Clus_out4;

by cluster distance;

run;

proc print data=Clus_out4;

var country cluster distance;

run;


proc candisc data=Clus_out4 out=ProCan4(keep=country cluster Can1 Can2)
noprint;

class cluster;

var x3-x44;

run;

proc sgplot data=ProCan4;

scatter y=Can2 x=Can1 / group=cluster datalabel=country;

label Can1="z1" Can2="z2";

run;
```

FA

```
PROC FACTOR DATA==countries ROTATE=VARIMAX OUTSTAT=SCORES SCORE ALL;
        VAR X3-X44;
RUN;

PROC SCORE DATA=countries SCORE=SCORES OUT=FINALSCORE;
        VAR X3-X44;
RUN;
```