**Complete code with SAS:**

%let path=D:\1. BA\Multivariate Analysis for Big Data\PRACDATA;

libname PRAC"&path";

ods graphics on;

\* Input the Excel file;

**proc import** datafile="&path\world\_2021.xls"

out=PRAC.world\_2021

dbms=xls

replace;

getnames=yes;

**run**;

\* Overall data study;

**proc contents** data=PRAC.world\_2021;

**run**;

**proc means** data=PRAC.world\_2021;

var oil--Happiness;

**run**;

ods graphics / imagemap;

**proc princomp** data=PRAC.world\_2021

n=**5**

out=prin

prefix=pca

plots=(matrix score(ncomp=**3**) patternprofile pattern(ncomp=**3**));

var Oil--Happiness;

id country;

**run**;

\* PCA;

ods graphics / imagemap;

**proc princomp** data=PRAC.world\_2021

n=**5**

out=prin

prefix=pca

plots=(matrix score(ncomp=**3**) patternprofile pattern(ncomp=**3**));

var oil--Renewables;

id country;

**run**;

**proc prinqual** data=PRAC.world\_2021 mdpref;

transform identity(oil--Renewables);

id country;

**run**;

title 'Oil by Coal and Natural gas';

**proc sgplot** data=PRAC.world\_2021;

bubble x=coal y='Natural gas'n size=oil / transparency=**0.4** datalabel=country;

inset "Bubble size represents Oil" / position=bottomright;

**run**;

\* Factor analysis;

ods graphics on;

**proc factor** data=PRAC.world\_2021 plots=(scree loadings) method=ml priors=smc;

title 'Factor Analysis: Extracting Factors';

var oil--Renewables;

**run**;

**proc factor** data=PRAC.world\_2021 plots=loadings method=principal

priors=smc n=**2** rotation=promax flag=**.3** fuzz=**.2**;

title 'Promax Rotation';

var oil--Renewables;

**run**;

\* Cluster analysis;

ods graphics on;

**proc cluster** data=prac.world\_2021 method=ward ccc pseudo outtree=tree print=**15**

plots=den(height=rsq);

var oil--Renewables;

id country;

**run**;

\* Canonical Discriminant Analysis;

ods output canonicalmeans=b(rename=(can1=can1c can2=can2c));

**proc candisc** data=PRAC.world\_2021 out=candout;

class region;

var oil--Renewables;

title 'Canonical Discriminant Analysis Using DSM IV Items';

**run**;

**data** plot;

set candout b;

**run**;

**proc sort** data=plot;

by region fromregion;

**run**;

**proc sgplot** data=plot nocycleattrs;

scatter x=can1 y=can2 / group=region;

scatter x=can1c y=can2c / group=fromregion

markerattrs=(size=**20**);

**run**;

**proc stepdisc** data=PRAC.world\_2021 method=stepwise;

class region;

var oil--Renewables;

**run**;

**proc discrim** data=PRAC.world\_2021;

class Region;

priors prop;

var oil 'Natural gas'n coal Hydroelectricity Renewables;

**run**;

\* Canonical correlation analysis;

ods output cancorr=a;

**proc cancorr** data=PRAC.world\_2021 out=out\_cancorr;

var oil--Renewables;

with GDP--Happiness;

**run**;

ods output cancorr=a;

**proc cancorr** data=PRAC.world\_2021

vprefix=R wprefix=G

vname='R Questions' wname= 'G Questions'

outstat=out;

var oil--Renewables;

with GDP--Population;

**run**;

**proc sgplot** data=a;

series y=squcancorr x=number /markers;

xaxis integer;

**run**;

**proc cancorr** data=PRAC.world\_2021 out=world red

vprefix=Energy wprefix=GP

vname='Energy indicators'

wname='GP indicators'

ncan=**2**;

var oil--Renewables;

with GDP--Population;

**run**;

**proc contents** data=world;

**run**;

**proc sgscatter** data=world;

plot energy1\*gp1 energy2\*gp2;

**run**;

**proc sgplot** data=a;

series y=squcancorr x=number /markers;

xaxis integer;

**run**;

\* Partial Least Squares (PLS);

**proc pls** data = PRAC.world\_2021 method = pls(algorithm=nipals)

cv=one cvtest(seed=**608789001**)

plot=(vip xyscores xscores parmprofiles dmod);

model GDP = oil--Renewables;

**run**;

**proc pls** data = prac.world\_2021 method = pls(algorithm=nipals)

cv=one cvtest(seed=**608789001**)

plot=(vip xyscores xscores parmprofiles dmod);

model GDP = oil--Renewables happiness population;

**run**;

\* Partial Least Squares (PLS);

**data** world;

set prac.world\_2021;

oil1=oil\***1000000**/population;

coal1=Coal\***1000000**/population;

naturalgas1='Natural gas'n\***1000000**/population;

hydroelectricity1=hydroelectricity\***1000000**/population;

nuclear1='Nuclear energy'n\***1000000**/population;

renewables1=renewables\***1000000**/population;

**run**;

**proc pls** data = world method = pls(algorithm=nipals)

cv=one cvtest(seed=**608789001**)

plot=(vip xyscores xscores parmprofiles dmod);

model happiness = oil1--renewables1;

id Country;

**run**;