```
*;
*;
*;
    ods graphics on;
*;
options ls=80 ps=50 nodate pageno=1;
*;
* Input BENIGN ;
*;
Data BENIGN;
Infile  '/folders/myfolders/benign.txt' OBS=178 DLM = '09'X TRUNCOVER;
Input Split90 STR OBS AGMT FNDX HIGD DEG CHK AGP1 AGMN NLV LIV  WT AGLP MST;
*;
Data BENIGN;
        Set BENIGN (Keep = Split90 STR AGMT FNDX DEG CHK AGP1 AGMN NLV LIV WT AGLP MST);
        Label Split90 = 'Split90'
                  STR = 'STR - Stratum'
          AGMT = 'AGMT - Age-of-the-subject'
          FNDX = 'Final-diagnosis'
          DEG = 'DEG - Degree'
          CHK = 'CHK - Regular-medical-checkup'
          AGP1 = 'AGP1 - Age-at-first-pregnancy'
          AGMN = 'AGMN - Age-at-menarche'
          NLV = 'NLV - Number-stillbirths-miscarriage'
          LIV = 'LIV - Number-of-live-birth'
          WT = 'WT - Weight-of-the-subject'
          AGLP = 'AGLP - Age-at-last-menstrual-period'
          MST ='MST - Marital-status';
*;
Proc Print Data = BENIGN;
*;
* Principal Components Analysis - All Variables;
*;
Proc Univariate Data = BENIGN;
Var NLV LIV;
*;
* BENIGN - Two-Sample T-Test;
*;
Proc TTest Data = BENIGN;
Class FNDX;
Var NLV LIV;
*;
*;
* BENIGN - ANOVA;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model NLV = FNDX;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model LIV= FNDX;
*;
*;
* BENIGN - REG;
*;
Proc REG Data = BENIGN;
Model NLV = FNDX;
*;
Proc REG Data = BENIGN;
Model LIV = FNDX;
Proc Univariate Data = BENIGN;
Var AGMT WT;
*;
* BENIGN - Two-Sample T-Test;
*;
Proc TTest Data = BENIGN;
Class FNDX;
Var AGMT WT;
*;
*;
* BENIGN - ANOVA;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model AGMT = FNDX;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model WT= FNDX;
*;
*;
* BENIGN - REG;
*;
Proc REG Data = BENIGN;
Model AGMT = FNDX;
*;
Proc REG Data = BENIGN;
Model WT = FNDX;
Proc Univariate Data = BENIGN;
```

```
Var AGLP AGMN;
*;
* BENIGN - Two-Sample T-Test;
*;
Proc TTest Data = BENIGN;
Class FNDX;
Var AGLP AGMN;
*;
*;
* BENIGN - ANOVA;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model AGLP = FNDX;
*;
Proc ANOVA Data = BENIGN;
Class FNDX;
Model AGMN= FNDX;
*;
*;
* BENIGN - REG;
*;
Proc REG Data = BENIGN;
Model AGLP = FNDX;
*;
Proc REG Data = BENIGN;
Model AGMN = FNDX;
Proc Princomp Data = BENIGN Plots = ALL;
    Var AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
*;
************ All Variables - Method=Principal Rotation: None and Varimax ****************;
*;
* Exploratory Factor Analysis Rotate=NONE All Variables ;
*;
Proc Factor Data = BENIGN Method=Principal Rotate=None NFactors=3 Simple MSA Plots = Scree MINEIGEN=0 Reorder;
   Var AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
* Exploratory Factor Analysis Rotate=Varimax All Variables ;
*;
Proc Factor Data = BENIGN Method=Principal Rotate=Varimax NFactors=3 Print Score Simple MSA Plots = Scree MINEIGEN=
    Var AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
*;
************ MST Deleted - Method=Principal Rotation: None and Varimax ****************;
*;
* Exploratory Factor Analysis Rotate=NONE MST Deleted ;
*;
*Proc Factor Data = BENIGN Method=Principal Rotate=None NFactors=5 Simple MSA Plots = Scree MINEIGEN=0 Reorder;
*    Var STR AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
* Exploratory Factor Analysis Rotate=Varimax MST Deleted ;
*;
*Proc Factor Data = BENIGN Method=Principal Rotate=Varimax NFactors=5 Print Score Simple MSA Plots = Scree MINEIGEN
*    Var STR AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
*;
************ MST& DEG Deleted - Method=Principal Rotation: None and Varimax ****************;
*;
* Exploratory Factor Analysis Rotate=NONE MST& DEG Deleted ;
*;
*Proc Factor Data = BENIGN Method=Principal Rotate=None NFactors=4 Simple MSA Plots = Scree MINEIGEN=0 Reorder;
*    Var STR AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
* Exploratory Factor Analysis Rotate=Varimax MST& DEG Deleted ;
*;
*Proc Factor Data = BENIGN Method=Principal Rotate=Varimax NFactors=4 Print Score Simple MSA Plots = Scree MINEIGEN
*    Var STR AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
*;
************  Compute Factor and Summated Scores****************;
*;
Proc Factor Data = BENIGN Outstat=FactOut Method=Principal Rotate=Varimax NFactors=3 Print Score Simple MSA Plots =
    Var AGMT AGP1 AGMN NLV LIV WT AGLP;
Proc Score Data=BENIGN Score=FactOut Out=FScore;
     Var AGMT AGP1 AGMN NLV LIV WT AGLP;
*;
Proc Print Data = FactOut;
*;
Proc Print Data = FScore;
*;
Data FScore;
        Set FScore;
        Label SumScale1 = 'SumScale1'
              SumScale2 = 'SumScale2'
         SumScale3 = 'SumScale3';
        SumScale1 = ( AGLP + AGMT ) / 2;
        SumScale2 = (AGMN + (10-WT)) / 3;
        SumScale3 = ((10-LIV) +(10-NLV) + AGP1) / 3;

*;
```

```
Proc Print Data = FScore;
*;
Proc Means Data = FScore;
   Var Factor1 Factor2 Factor3 SumScale1 SumScale2 SumScale3;
*;
*;
************  Compute Factor and Summated Correlations ****************;
*;
Proc Corr Data = FScore;
   Var Factor1 Factor2 Factor3 SumScale1 SumScale2 SumScale3;
*;
Data FScore90;
Set FScore;
If Split90 = 1;
*;

Data FScore88;
Set FScore;
If Split90 = 0;
*;

Proc Print Data = FScore90;
*;

Proc Print Data = FScore88;
*;

Proc Logistic Data = FScore90;
Model FNDX(event='1') = SumScale1 SumScale2 SumScale3
/ Selection=Stepwise SLEntry=0.05 SLStay=0.05 Details
LackFit RSquare CTable PProb =(0 to 1 by .10);
*;
* Final Resultant Model and Output Model;
*;

Proc Logistic Data = FScore90 OutModel=Logistic90;
Model FNDX(event='1') = SumScale2
/ LackFit RSquare CTable PProb =(0.40 to 0.60 by
.01);
*;
* Original Split60 Logistic Model Fitted to Split40 validation Data;
*;

Proc Logistic InModel=Logistic90;
Score Data = FScore90 (Keep = FNDX SumScale2) Out = BENIGN90Score;
*;
* Proc Freq Crosstabulations Original and Holdout Validation Datasets;
*;

Proc Print Data = BENIGN90Score;

Proc Freq Data = BENIGN90Score;
Table F_FNDX * I_FNDX;
*;

Proc Logistic InModel=Logistic90;
Score Data = FScore88 (Keep = FNDX SumScale2) Out = BENIGN88Score;

Proc Print Data = BENIGN88Score;

Proc Freq Data = BENIGN88Score;

Table F_FNDX * I_FNDX;
*;
Proc Means Data = BENIGN;
    Var LIV NLV;
*;
Proc Sort Data = BENIGN;
    By MST DEG;
*;
Proc Means Data = BENIGN;
    Var LIV NLV;
        By MST DEG;
        ID MST DEG;
*;
* Exploratory Data Analysis - Univariate ;
*;
Proc Univariate Data = BENIGN Normal Plot;
    Var LIV NLV;
*;
Proc Sort Data = BENIGN;
    By MST DEG;
*;
Proc Univariate Data = BENIGN Normal Plot;
    Var LIV NLV;
        By MST DEG;
        ID MST DEG;
*;
* GLM MANOVA Analysis ;
*;
Proc GLM Data = BENIGN;
```

```
    Class MST DEG;
        Model LIV NLV = MST DEG;
        Means MST DEG / Scheffe Tukey LSD SNK Duncan;
        Means MST DEG / Hovtest = Levene Hovtest = bf Hovtest = Bartlett;
        Means MST DEG;
        Manova H = MST DEG / MStat = Exact;
*;
* ods graphics off;
*;
*;
*;
Run;
Quit;
```