

```

/* Label: Impute Missing */
DATA WORK.SUPPORT;
SET MYDATA.SUPPORT2;
  IF STRIP(INCOME)=" " THEN INCOME="MISSING";
  IF STRIP(RACE)=" " THEN RACE="MISSING";
  IF EDU="." THEN EDU=12;
  IF ALB="." THEN ALB=3.5;
  IF PAFI="." THEN PAFI=333.3;
  IF BILI="." THEN BILI=1.01;
  IF CREA="." THEN CREA=1.01;
  IF BUN="." THEN BUN=6.51;
  IF WBLC="." THEN WBLC=9;
  IF URINE="." THEN URINE=2502;
  IF GLUCOSE="." THEN GLUCOSE=159;
  IF PH="." THEN PH=7.41;
  IF SOD="." THEN SOD=137;
  IF TEMP="." THEN TEMP=37;
  IF RESP="." THEN RESP=24;
  IF HRT="." THEN HRT=100;
  IF meanBP="." THEN meanBP=77;

  if scoma=. then scoma=35;

RUN;

/* Label: model missing scoma */
proc hpsplit data=WORK.SUPPORT maxdepth=4;
class DZGROUP INCOME RACE CA death hospdead SEX DIABETES DEMENTIA;
model scoma=DZGROUP INCOME RACE CA SEX DIABETES DEMENTIA AGE SLOS 'NUM.CO'N EDU
  charges totmcsst totcst MEANBP death hospdead adlsc WBLC HRT RESP TEMP PAFI ALB BILI CREA
SOD PH GLUCOSE BUN URINE;
grow ftest;
prune reducederror;
partition fraction(validate=0.2);
Rules file="/home/u63479636/HIM6655/SUPPORT101.txt";
run;

/* Label: Create dead at 2 months target variable */
data WORK.support;
set WORK.support;
if death=0 then death2m=0;
if sfdm2=" " and hospdead=1 then death2m=0;
if sfdm2=" " and "d.time">60 then death2m=0;
  if sfdm2=" " and "d.time"<= 60 then death2m=1;
if sfdm2="<2 mo. follow-up" then death2m=1;
if sfdm2 ne "<2 mo. follow-up" and sfdm2 ne " " then death2m=0;
run;

/* Label: death2m dt1 */
proc hpsplit data=WORK.SUPPORT MAXDEPTH=5;
class death2m DZGROUP INCOME RACE CA SEX DIABETES DEMENTIA;
model death2m=DZGROUP INCOME RACE CA SEX DIABETES DEMENTIA AGE 'NUM.CO'N EDU
SCOMA MEANBP WBLC HRT RESP TEMP PAFI ALB BILI CREA
SOD PH GLUCOSE BUN URINE;
grow chisquare;
prune entropy;
partition fraction(validate=0.2 seed=1111);
Rules file="/home/u63479636/HIM6655/SUPPORT3.txt";
run;

/* Label: Remove inappropriate variables */
data work.support_pca;
set work.support;
drop A adlp adls adlsc avtisst charges 'd.time' death dzclass hospdead slos totcst totmcsst sfdm2;
run;

/* Label: death2m dt2 */
proc hpsplit data=WORK.SUPPORT MAXDEPTH=5;
class death2m DZGROUP INCOME RACE CA SEX DIABETES DEMENTIA;
model death2m=DZGROUP INCOME RACE CA SEX DIABETES DEMENTIA AGE 'NUM.CO'N EDU
SCOMA MEANBP WBLC HRT RESP TEMP PAFI ALB BILI CREA
SOD PH GLUCOSE BUN URINE;
grow fastchaid;
prune c45;
partition fraction(validate=0.2 seed=1111);
Rules file="/home/u63479636/HIM6655/SUPPORT4.txt";
run;

/* Label: Correct skewness and outliers */
data work.support_pca;

```

```

set work.support_pca;
  logBun=log10(bun);
  logBili=log10(bili);
  logPafi=log10(pafi);
  logScoma=log10(scoma+1);
  logWBLC=log10(wblc+1);
  logGlucose=log10(glucose+1);
  logUrine=log10(urine+1);
  if alb>5 then delete;
run;

/* Label: Principal Component Analysis */
/*
*
* Task code generated by SAS Studio 3.8
*
* Generated on '12/10/23, 4:19 PM'
* Generated by 'u63479636'
* Generated on server 'ODAWS01-USW2-2.ODA.SAS.COM'
* Generated on SAS platform 'Linux LIN X64 3.10.0-1062.12.1.el7.x86_64'
* Generated on SAS version '9.04.01M7P08062020'
* Generated on browser 'Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chr
* Generated on web client 'https://odamid-usw2-2.oda.sas.com/SASStudio/main?locale=en_US&zone=GMT-05%253A00&ticket
*/

ods noproctitle;
ods graphics / imagemap=on;

proc princomp data=WORK.SUPPORT_PCA plots(only)=(scree) out=work.Princomp1;
  var age 'num.co'n edu hday meanbp hrt resp temp alb crea sod ph logBun logBili
      logPafi logScoma logWBLC logGlucose logUrine diabetes dementia;
run;

/* Label: death2m dt3 */
proc hpsplit data=WORK.princomp1 MAXDEPTH=5;
class death2m;
model death2m=prin1 prin2 prin3 prin4;
grow chisquare;
prune entropy;
partition fraction(validate=0.2 seed=1111);
Rules file="/home/u63479636/HIM6655/SUPPORT3.txt";
run;

/* Label: death2m dt5 */
proc hpsplit data=WORK.princomp1 MAXDEPTH=5;
  class death2m DZGROUP INCOME RACE CA SEX;
  model death2m=DZGROUP INCOME RACE CA SEX prin1 prin2 prin3 prin4
  age 'num.co'n edu hday meanbp hrt resp temp alb crea sod ph logbun
  logbili logpafi logscoma logwblc logglucose logurine;
  grow fastchaid;
  prune c45;
partition fraction(validate=0.2 seed=1111);
Rules file="/home/u63479636/HIM6655/SUPPORT5.txt";
run;

/* Label: death2m dt4 */
proc hpsplit data=WORK.princomp1 MAXDEPTH=5;
  class death2m DZGROUP INCOME RACE CA SEX;
  model death2m=DZGROUP INCOME RACE CA SEX prin1 prin2 prin3 prin4;
  grow fastchaid;
  prune c45;
partition fraction(validate=0.2 seed=1111);
Rules file="/home/u63479636/HIM6655/SUPPORT5.txt";
run;

```