

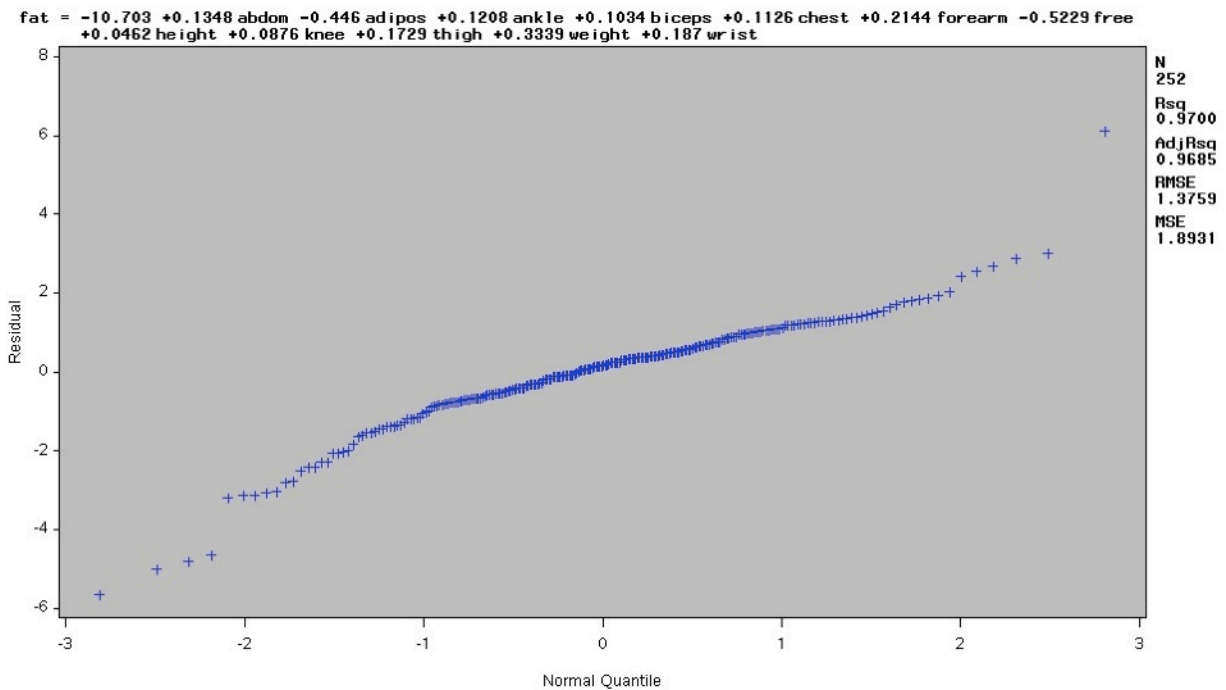
Final

Jeremy Harper

12.02.11

STA4203

1. Variables selected using Adj-R² criteria: abdom adipos ankle biceps chest forearm free height knee thigh weight wrist
2. QQ-Plot of Residuals:



Tests for Normality of Residuals from model:

| Tests for Normality | | | | |
|---------------------|---------------|----------|-------------------|---------|
| Test | --Statistic-- | | -----p Value----- | |
| Shapiro-Wilk | W | 0.931268 | Pr < W | <0.0001 |
| Kolmogorov-Smirnov | D | 0.095024 | Pr > D | <0.0100 |
| Cramer-von Mises | W-Sq | 0.599099 | Pr > W-Sq | <0.0050 |
| Anderson-Darling | A-Sq | 3.83266 | Pr > A-Sq | <0.0050 |

Based on all four tests, we can reject the null hypothesis that the residuals are normal.

3. Based on the Durbin-Watson test, we can conclude that the errors have significant positive autocorrelation, but not significant negative correlation.

| Dependent Variable: fat | |
|-------------------------|--------|
| Durbin-Watson D | 1.759 |
| Pr < DW | 0.0237 |

```

Pr > DW                0.9763
Number of Observations    252
1st Order Autocorrelation 0.119

```

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

4.

| 4 | Test 1 RMSE | Test 2 RMSE | Test 3 RMSE | Test 4 RMSE | Avg RMSE |
|---|-------------|-------------|-------------|-------------|-----------|
| A | 1.43368 | 5.90383 | 1.76616 | 1.71772 | 2.7053475 |
| B | 1.47303 | 5.57825 | 1.81045 | 1.69609 | 2.639455 |
| C | 1.47303 | 5.57825 | 1.93866 | 1.67518 | 2.66628 |
| D | 1.41965 | 5.83830 | 1.76201 | 1.70476 | 2.68118 |
| E | 1.43167 | 5.68398 | 1.80328 | 1.70524 | 2.6560425 |
| F | 3.43371 | 3.16641 | 3.81131 | 3.34831 | 3.439935 |

CODE

```

PROC IMPORT OUT= WORK.bodyfat
  DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\data\bodyfat.csv"
  DBMS=CSV REPLACE;
  GETNAMES=YES;
  DATAROW=2;
RUN;

proc contents data = bodyfat;
run;

/* 1 */

proc reg data = bodyfat;
  model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
  adjrsq;
  output out = resid1 r = resid;
  plot r.*nqq. /noline mse cframe=ligr;
run; quit;

/* 2 */

proc univariate data = resid1 normal;
  var resid;
  histogram / normal;
run;

/* 3 */

proc reg data = bodyfat;
  model fat = abdom adipos ankle biceps chest forearm free height knee thigh weight wrist / dwprob;
run; quit;

/* 4 */
/* divide datasets */

```

```

data bodyfat1;
do i=1 to 252 by 4;
set bodyfat point=i;
output; end; stop;
run;

data bodyfat2;
do i=2 to 252 by 4;
set bodyfat point=i;
output; end; stop;
run;

data bodyfat3;
do i=3 to 252 by 4;
    set bodyfat point=i;
    output; end; stop;
run;

data bodyfat4;
do i=4 to 252 by 4;
    set bodyfat point=i;
    output; end; stop;
run;

data bodyfat123;
    set bodyfat1 bodyfat2 bodyfat3;
run;

data bodyfat124;
    set bodyfat1 bodyfat2 bodyfat4;
run;

data bodyfat134;
    set bodyfat1 bodyfat3 bodyfat4;
run;

data bodyfat234;
    set bodyfat2 bodyfat3 bodyfat4;
run;

/* 4a */

/* Test One */

proc reg data = bodyfat234 outest=model4aTest1;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run; quit;

proc score data=bodyfat1 score=model4aTest1 out=Test1Ascored residual type=parms;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist fat;
run;

proc univariate data = Test1Ascored;
    var model1;
    output out = Test1AscoredUSS uss = ss1;
run;

data Test1AscoredUSS;
    set Test1AscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test1AscoredUSS;
    var rmse;
run;

/* Test Two */

proc reg data = bodyfat134 outest=model4aTest2;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run; quit;

```

```

proc score data=bodyfat2 score=model4aTest2 out=Test2Ascored residual type=parms;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist fat;
run;

proc univariate data = Test2Ascored;
    var model1;
    output out = Test2AscoredUSS uss = ss1;
run;

data Test2AscoredUSS;
    set Test2AscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test2AscoredUSS;
    var rmse;
run;

    /* Test Three */

proc reg data = bodyfat124 outest=model4aTest3;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run; quit;

proc score data=bodyfat3 score=model4aTest3 out=Test3Ascored residual type=parms;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist fat;
run;

proc univariate data = Test3Ascored;
    var model1;
    output out = Test3AscoredUSS uss = ss1;
run;

data Test3AscoredUSS;
    set Test3AscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test3AscoredUSS;
    var rmse;
run;

    /* Test Four */

proc reg data = bodyfat123 outest=model4aTest4;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run; quit;

proc score data=bodyfat4 score=model4aTest4 out=Test4Ascored residual type=parms;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist fat;
run;

proc univariate data = Test4Ascored;
    var model1;
    output out = Test4AscoredUSS uss = ss1;
run;

data Test4AscoredUSS;
    set Test4AscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test4AscoredUSS;
    var rmse;
run;

/* 4b */

    /* Test One */

proc reg data = bodyfat234 outest=model4bTest1;

```

```

        model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
backward slstay = 0.05;
run; quit;

```

```

proc score data=bodyfat1 score=model4bTest1 out=Test1Bscored residual type=parms;
    var abdom adipos chest forearm free thigh weight fat;
run;

```

```

proc univariate data = Test1Bscored;
    var model1;
    output out = Test1BscoredUSS uss = ss1;
run;

```

```

data Test1BscoredUSS;
    set Test1BscoredUSS;
    rmse=sqrt(ss1/63);
run;

```

```

proc print data = Test1BscoredUSS;
    var rmse;
run;

```

/* Test Two */

```

proc reg data = bodyfat134 outest=model4bTest2;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
backward slstay = 0.05;
run; quit;

```

```

proc score data=bodyfat2 score=model4bTest2 out=Test2Bscored residual type=parms;
    var abdom adipos biceps chest forearm free height thigh weight fat;
run;

```

```

proc univariate data = Test2Bscored;
    var model1;
    output out = Test2BscoredUSS uss = ss1;
run;

```

```

data Test2BscoredUSS;
    set Test2BscoredUSS;
    rmse=sqrt(ss1/63);
run;

```

```

proc print data = Test2BscoredUSS;
    var rmse;
run;

```

/* Test Three */

```

proc reg data = bodyfat124 outest=model4bTest3;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
backward slstay = 0.05;
run; quit;

```

```

proc score data=bodyfat3 score=model4bTest3 out=Test3Bscored residual type=parms;
    var abdom adipos age forearm free thigh weight fat;
run;

```

```

proc univariate data = Test3Bscored;
    var model1;
    output out = Test3BscoredUSS uss = ss1;
run;

```

```

data Test3BscoredUSS;
    set Test3BscoredUSS;
    rmse=sqrt(ss1/63);
run;

```

```

proc print data = Test3BscoredUSS;
    var rmse;
run;

```

```

/* Test Four */

proc reg data = bodyfat123 outest=model4bTest4;
  model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
backward slstay = 0.05;
run; quit;

proc score data=bodyfat4 score=model4bTest4 out=Test4Bscored residual type=parms;
  var abdom adipos ankle chest forearm free knee thigh weight fat;
run;

proc univariate data = Test4Bscored;
  var model1;
  output out = Test4BscoredUSS uss = ss1;
run;

data Test4BscoredUSS;
  set Test4BscoredUSS;
  rmse=sqrt(ss1/63);
run;

proc print data = Test4BscoredUSS;
  var rmse;
run;

/* 4c */

/* Test One */

proc reg data = bodyfat234 outest=model4cTest1;
  model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
forward slentry = 0.01;
run; quit;

proc score data=bodyfat1 score=model4cTest1 out=Test1cscored residual type=parms;
  var abdom adipos chest forearm free thigh weight fat;
run;

proc univariate data = Test1cscored;
  var model1;
  output out = Test1cscoredUSS uss = ss1;
run;

data Test1cscoredUSS;
  set Test1cscoredUSS;
  rmse=sqrt(ss1/63);
run;

proc print data = Test1cscoredUSS;
  var rmse;
run;

/* Test Two */

proc reg data = bodyfat134 outest=model4cTest2;
  model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
forward slentry = 0.01;
run; quit;

proc score data=bodyfat2 score=model4cTest2 out=Test2cscored residual type=parms;
  var abdom adipos biceps chest forearm free height thigh weight fat;
run;

proc univariate data = Test2cscored;
  var model1;
  output out = Test2cscoredUSS uss = ss1;
run;

data Test2cscoredUSS;
  set Test2cscoredUSS;
  rmse=sqrt(ss1/63);
run;

```

```

proc print data = Test2cscoredUSS;
    var rmse;
run;

/* Test Three */

proc reg data = bodyfat124 outest=model4cTest3;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
forward slentry = 0.01;
run; quit;

proc score data=bodyfat3 score=model4cTest3 out=Test3cscored residual type=parms;
    var abdom forearm free weight fat;
run;

proc univariate data = Test3cscored;
    var model1;
    output out = Test3cscoredUSS uss = ss1;
run;

data Test3cscoredUSS;
    set Test3cscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test3cscoredUSS;
    var rmse;
run;

/* Test Four */

proc reg data = bodyfat123 outest=model4cTest4;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
forward slentry = 0.01;
run; quit;

proc score data=bodyfat4 score=model4cTest4 out=Test4cscored residual type=parms;
    var abdom adipos chest forearm free knee thigh weight fat;
run;

proc univariate data = Test4cscored;
    var model1;
    output out = Test4cscoredUSS uss = ss1;
run;

data Test4cscoredUSS;
    set Test4cscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test4cscoredUSS;
    var rmse;
run;

/* 4d */

/* Test One */

proc reg data = bodyfat234;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
adjrsq;
run; quit;

proc reg data = bodyfat234 outest=model4dTest1;
    model fat = abdom adipos ankle biceps chest forearm free thigh weight wrist;
run; quit;

proc score data=bodyfat1 score=model4dTest1 out=Test1dscored residual type=parms;
    var abdom adipos ankle biceps chest forearm free thigh weight wrist fat;
run;

```

```

proc univariate data = Test1dscored;
    var model1;
    output out = Test1dscoredUSS uss = ss1;
run;

data Test1dscoredUSS;
    set Test1dscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test1dscoredUSS;
    var rmse;
run;

    /* Test Two */

proc reg data = bodyfat134;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
adjrsq;
run; quit;

proc reg data = bodyfat134 outest = model4dTest2;
    model fat = abdom adipos biceps chest forearm free height hip thigh weight wrist;
run; quit;

proc score data=bodyfat2 score=model4dTest2 out=Test2dscored residual type=parms;
    var abdom adipos biceps chest forearm free height hip thigh weight wrist fat;
run;

proc univariate data = Test2dscored;
    var model1;
    output out = Test2dscoredUSS uss = ss1;
run;

data Test2dscoredUSS;
    set Test2dscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test2dscoredUSS;
    var rmse;
run;

    /* Test Three */

proc reg data = bodyfat124;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
adjrsq;
run; quit;

proc reg data = bodyfat124 outest=model4dTest3;
    model fat = abdom adipos age ankle biceps chest forearm free height neck thigh weight;
run; quit;

proc score data=bodyfat3 score=model4dTest3 out=Test3dscored residual type=parms;
    var abdom adipos age ankle biceps chest forearm free height neck thigh weight fat;
run;

proc univariate data = Test3dscored;
    var model1;
    output out = Test3dscoredUSS uss = ss1;
run;

data Test3dscoredUSS;
    set Test3dscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test3dscoredUSS;
    var rmse;
run;

```



```

/* Test Four */

proc reg data = bodyfat123;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection =
adjrsq;
run; quit;

proc reg data = bodyfat123 outest=model4dTest4;
    model fat = abdom adipos ankle chest forearm free height knee thigh weight wrist;
run; quit;

proc score data=bodyfat4 score=model4dTest4 out=Test4dscored residual type=parms;
    var abdom adipos ankle chest forearm free height knee thigh weight wrist fat;
run;

proc univariate data = Test4dscored;
    var model1;
    output out = Test4dscoredUSS uss = ss1;
run;

data Test4dscoredUSS;
    set Test4dscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test4dscoredUSS;
    var rmse;
run;

/* 4e */

/* Test One */

proc reg data = bodyfat234;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection = adjrsq
aic;
run; quit;

proc reg data = bodyfat234 outest=model4eTest1;
    model fat = abdom adipos ankle biceps chest forearm free thigh weight;
run; quit;

proc score data=bodyfat1 score=model4eTest1 out=Test1scored residual type=parms;
    var abdom adipos ankle biceps chest forearm free thigh weight fat;
run;

proc univariate data = Test1scored;
    var model1;
    output out = Test1scoredUSS uss = ss1;
run;

data Test1scoredUSS;
    set Test1scoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test1scoredUSS;
    var rmse;
run;

/* Test Two */

proc reg data = bodyfat134;
    model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection = adjrsq
aic;
run; quit;

proc reg data = bodyfat134 outest = model4eTest2;
    model fat = abdom adipos biceps chest forearm free height thigh weight wrist;
run; quit;

proc score data=bodyfat2 score=model4eTest2 out=Test2scored residual type=parms;

```

```

var abdom adipos biceps chest forearm free height thigh weight wrist fat;
run;

proc univariate data = Test2scored;
var model1;
output out = Test2scoredUSS uss = ss1;
run;

data Test2scoredUSS;
set Test2scoredUSS;
rmse=sqrt(ss1/63);
run;

proc print data = Test2scoredUSS;
var rmse;
run;

/* Test Three */

proc reg data = bodyfat124;
model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection = adjrsq
aic;
run; quit;

proc reg data = bodyfat124 outest=model4eTest3;
model fat = abdom adipos age ankle biceps forearm free thigh weight;
run; quit;

proc score data=bodyfat3 score=model4eTest3 out=Test3scored residual type=parms;
var abdom adipos age ankle biceps forearm free thigh weight fat;
run;

proc univariate data = Test3scored;
var model1;
output out = Test3scoredUSS uss = ss1;
run;

data Test3scoredUSS;
set Test3scoredUSS;
rmse=sqrt(ss1/63);
run;

proc print data = Test3scoredUSS;
var rmse;
run;

/* Test Four */

proc reg data = bodyfat123;
model fat = abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist / selection = adjrsq
aic;
run; quit;

proc reg data = bodyfat123 outest=model4eTest4;
model fat = abdom adipos ankle chest forearm free knee thigh weight wrist;
run; quit;

proc score data=bodyfat4 score=model4eTest4 out=Test4scored residual type=parms;
var abdom adipos ankle chest forearm free knee thigh weight wrist fat;
run;

proc univariate data = Test4scored;
var model1;
output out = Test4scoredUSS uss = ss1;
run;

data Test4scoredUSS;
set Test4scoredUSS;
rmse=sqrt(ss1/63);
run;

proc print data = Test4scoredUSS;

```

```

var rmse;
run;

/* 4f */

/* standardize and divide datasets */

proc standard data = bodyfat mean = 0 std = 1 out = bodyfatstd;
var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

data bodyfatstd1;
do i=1 to 252 by 4;
set bodyfatstd point=i;
output; end; stop;
run;

data bodyfatstd2;
do i=2 to 252 by 4;
set bodyfatstd point=i;
output; end; stop;
run;

data bodyfatstd3;
do i=3 to 252 by 4;
set bodyfatstd point=i;
output; end; stop;
run;

data bodyfatstd4;
do i=4 to 252 by 4;
set bodyfatstd point=i;
output; end; stop;
run;

data bodyfatstd123;
set bodyfatstd1 bodyfatstd2 bodyfatstd3;
run;

data bodyfatstd124;
set bodyfatstd1 bodyfatstd2 bodyfatstd4;
run;

data bodyfatstd134;
set bodyfatstd1 bodyfatstd3 bodyfatstd4;
run;

data bodyfatstd234;
set bodyfatstd2 bodyfatstd3 bodyfatstd4;
run;

/* Test One */

proc princomp data = bodyfatstd234 outstat = bodyfat234stdPCA;
var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfatstd234 score = bodyfat234stdPCA out = bodyfat234stdPCAscored;
var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc reg data = bodyfat234stdPCAscored outest = bodyfat234stdPCAscoredmodel;
model fat = prin1-prin7;
run; quit;

proc score data = bodyfatstd1 score = bodyfat234stdPCA out = bodyfat1stdPCAscored;
var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfat1stdPCAscored score = bodyfat234stdPCAscoredmodel out = Test1fscored residual type=parms;
var prin1-prin7 fat;

```

```

run;

proc univariate data = Test1fscored;
    var model1;
    output out = Test1fscoredUSS uss = ss1;
run;

data Test1fscoredUSS;
    set Test1fscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test1fscoredUSS;
    var rmse;
run;

/* Test Two */

proc princomp data = bodyfatstd134 outstat = bodyfat134stdPCA;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfatstd134 score = bodyfat134stdPCA out = bodyfat134stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc reg data = bodyfat134stdPCAscored outest = bodyfat134stdPCAscoredmodel;
    model fat = prin1-prin7;
run; quit;

proc score data = bodyfatstd2 score = bodyfat134stdPCA out = bodyfat2stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfat2stdPCAscored score = bodyfat134stdPCAscoredmodel out = Test2fscored residual type=parms;
    var prin1-prin7 fat;
run;

proc univariate data = Test2fscored;
    var model1;
    output out = Test2fscoredUSS uss = ss1;
run;

data Test2fscoredUSS;
    set Test2fscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test2fscoredUSS;
    var rmse;
run;

/* Test Three */

proc princomp data = bodyfatstd124 outstat = bodyfat124stdPCA;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfatstd124 score = bodyfat124stdPCA out = bodyfat124stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc reg data = bodyfat124stdPCAscored outest = bodyfat124stdPCAscoredmodel;
    model fat = prin1-prin7;
run; quit;

proc score data = bodyfatstd3 score = bodyfat124stdPCA out = bodyfat3stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

```

```

proc score data = bodyfat3stdPCAscored score = bodyfat124stdPCAscoredmodel out = Test3fscored residual type=parms;
    var prin1-prin7 fat;
run;

proc univariate data = Test3fscored;
    var model1;
    output out = Test3fscoredUSS uss = ss1;
run;

data Test3fscoredUSS;
    set Test3fscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test3fscoredUSS;
    var rmse;
run;

/* Test Four */

proc princomp data = bodyfatstd123 outstat = bodyfat123stdPCA;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfatstd123 score = bodyfat123stdPCA out = bodyfat123stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc reg data = bodyfat123stdPCAscored outest = bodyfat123stdPCAscoredmodel;
    model fat = prin1-prin7;
run; quit;

proc score data = bodyfatstd4 score = bodyfat123stdPCA out = bodyfat3stdPCAscored;
    var abdom adipos age ankle biceps chest forearm free height hip knee neck thigh weight wrist;
run;

proc score data = bodyfat3stdPCAscored score = bodyfat123stdPCAscoredmodel out = Test4fscored residual type=parms;
    var prin1-prin7 fat;
run;

proc univariate data = Test4fscored;
    var model1;
    output out = Test4fscoredUSS uss = ss1;
run;

data Test4fscoredUSS;
    set Test4fscoredUSS;
    rmse=sqrt(ss1/63);
run;

proc print data = Test4fscoredUSS;
    var rmse;
run;

```