# SAS CODE

**proc** **import** datafile='FilePath\Votes.xlsx' DBMS=xlsx out=Vote replace;

**run**;

**proc** **print** data=Vote;

**run**;

**proc** **univariate** data=Vote normal plot;

var savings poverty;

**run**;

**data** out1;

set Vote;

if (savings < **165608.5**) then output;//165608.5 is the Q4 value

**run**;

**data** out1;

set out1;

if (poverty < **32.575**) then output;

**run**;

**proc** **print** data=out1;

**run**;

**proc** **univariate** data=out1 normal plot;

var savings poverty;

**run**;

**proc** **univariate** data=out1 normal plot;

var votes;

**run**;

set out1;

LINCOME = log(income);

**run**;

**proc** **surveyselect** data= out1 (firstobs = **1** obs =**500**)

n = **500**

out=VOTETRAIN

method=seq;

**run**;

**proc** **print** data = VOTETRAIN;

**run**;

**proc** **surveyselect** data= out1 (firstobs = **501** obs =**703**)

n = **203**

out=VOTETEST

method=seq;

**run**;

**proc** **print** data = VOTETEST;

**run**;

**proc** **reg** data=VOTETRAIN;

model VOTES = FEMALE DENSITY POVERTY VETERANS / tol vif collin;

plot r.\*p.;

**run**;

**data** mod\_test;

set VOTETEST;

y\_bar = -**49.643** + (**1.3521**\*female) +(**0.0027**\*density) +(**0.8436**\*poverty) + (**0.7028**\*veterans);

Predicted\_err = ((VOTES - y\_bar)\*\***2**);

**run**;

**proc** **print** data = mod\_test;

**run**;

**proc** **means** data = mod\_test;

var Predicted\_err;

**run**;