

```

proc import file = "C:\Users\dryan\OneDrive\Desktop\PresPredictions538.csv"
out = presprediction
dbms = csv replace;
GetNames=Y;
run;

data win;
set presprediction;
if forecast_type in ("polls-only", " ");
if state ne "US";
if probwin < .025 then probr = 0;
else if .475 <= probwin <= .5 then probr = 0.4875;
else if .5 < probwin <= .525 then probr = 0.5125;
else if .025 <= probwin < .075 then probr = .05;
else if .925 <= probwin < .975 then probr = .95;
else if probwin > .975 then probr = 1;
else probr = round(probwin, .05);
time = election_date - forecast_date;
run;

proc sort data = win;
by probr;
run;

proc sort data = win;
by year probr state time;
run;

data flat;
set win;
by year probr state time;
if first.state;
run;

proc sort data = flat;
by probr descending probwin_outcome;
run;

ods output binomial = chart;
ods trace on;
proc freq data = flat order = data;
by probr;
table probwin_outcome / binomial;
run;
ods trace off;

data graph;
retain phat ase lower upper;
set chart;
by probr;
if namel = "_BIN_" then phat = nvalue1;
if namel = "E_BIN" then ase = nvalue1;
if namel = "XL_BIN" then lower = nvalue1;
if namel = "XU_BIN" then upper = nvalue1;

if last.probr then do;
    if (ase = 0 & probr < 0.5) then phat = 1-phat;

```

```

        if (ase = 0 & probr < 0.5) then lowertemp = 1-upper;
        if (ase = 0 & probr < 0.5) then uppertemp = 1-lower;
    end;

    if lowertemp ne . then lower = lowertemp;
    if uppertemp ne . then upper = uppertemp;

    if last.probr;
    keep probr phat ase lower upper;
run;

proc sgplot data = graph noautolegend;
highlow x = probr high = upper low = lower;
scatter x = probr y = phat;
lineparm x= 0.2 y = 0.2 slope = 1;
TITLE "Independant Observations Most Recent split 50% bin";
run;

data graph2;
set graph;
if probr ne .5;
run;

proc sgplot data = graph2 noautolegend;
highlow x = probr high = upper low = lower;
scatter x = probr y = phat;
lineparm x= 0.2 y = 0.2 slope = 1;
TITLE "Independant Observations Take 2 with Time";
run;

proc sgplot data =graph noautolegend;
band x = probr upper = upper lower = lower / fillattrs = (transparency =
0.5);
loess x = probr y = phat / degree = 2 smooth = .1;
lineparm x = 0.5 y = 0.5 slope = 1;
TITLE "With bands";
run;

data win1;
set presprediction;
if forecast_type in ("polls-only", " ");
if state ne "US";
if probwin < .005 then probr = 0;
/*else if .025 <= probwin < .075 then probwin =.05;
else if .925 <= probwin < .975 then probwin = .95;*/
else if probwin > .995 then probr = 1;
else probr = round(probwin, .01);
time = election_date - forecast_date;
run;

proc sort data = win1;
by probr;
run;

```

```

proc sort data = win1;
by year probr state time;
run;

data flat1;
set win1;
by year probr state time;
if first.state;
run;

proc sort data = flat1;
by probr descending probwin_outcome;
run;

ods output binomial = chart1;
ods trace on;
proc freq data = flat1 order = data;
by probr;
table probwin_outcome / binomial;
run;
ods trace off;

data graph1;
retain phat ase lower upper;
set chart1;
by probr;
if name1 = "_BIN_" then phat = nvalue1;
if name1 = "E_BIN" then ase = nvalue1;
if name1 = "XL_BIN" then lower = nvalue1;
if name1 = "XU_BIN" then upper = nvalue1;

if last.probr then do;
    if (ase = 0 & probr < 0.5) then phat = 1-phat;
    if (ase = 0 & probr < 0.5) then lowertemp = 1-upper;
    if (ase = 0 & probr < 0.5) then uppertemp = 1-lower;
end;

if lowertemp ne . then lower = lowertemp;
if uppertemp ne . then upper = uppertemp;

if last.probr;
keep probr phat ase lower upper;
run;

proc sgplot data = graph1 noautolegend;
highlow x = probr high = upper low = lower;
scatter x = probr y = phat;
lineparm x= 0.2 y = 0.2 slope = 1;
refline 0.8 0.2 / axis = y;
refline 0.5 / axis = y lineattrs = (pattern = 4);
TITLE "Independant Observations 1% bins with Most Recent";
run;

```