```
DATA stock_data;
INFILE '/home/u62947731/MS Project/Mixture.txt';
INPUT stock_symbol $ date MMDDYY10. open high low close volume $ DowJones_Industrial $;
DROP  open high low volume DowJones_Industrial;
FORMAT date MMDDYY10.;
proc sort Data= stock_data;
by stock_symbol date;
run;
/*proc print data = stock_data;run;*/

proc expand data=stock_data out=stock_data1 method=none;
by stock_symbol;
id date;
convert close = lag1_close   / transformout=(lag 1);
convert close = lag70_close  / transformout=(lag 70);
convert close = lead1_close  / transformout=(lead 1);
convert close = lead70_close  / transformout=(lead 70);
run;
/*proc print data=stock_data1; run;*/

Data stock_data2;
set stock_data1;
daily_return_lag=((close-lag1_close)/(lag1_close))*100;
Seventy_days_return_lag=((close-lag70_close)/(lag70_close))*100;
daily_return_lead=((lead1_close-close)/(close))*100;
Seventy_days_return_lead=((lead70_close-close)/(close))*100;
drop lag1_close lag70_close lead1_close daily_return_lag Seventy_days_return_lag daily_return_lead;
run;
/*proc print data=stock_data2; run;*/


proc sql;
    create table stock_data3 as
    select date
    from stock_data2
    where stock_symbol is not missing
    group by date
    having count(distinct stock_symbol) = 8;
```

```sas
    quit;


proc sql;
    create table stock_data4 as
    select a.*
    from stock_data2 as a
    inner join stock_data3 as b
    on a.date = b.date;
quit;

/*proc print data= stock_data4;run;*/

data stock_data5;
    set stock_data4;
    If '06JUL2010'd <= date <= '13JUN2024'd;
    drop lead70_close close;

    proc sort data=stock_data5;
    by date;
run;


/*proc print data=stock_data5;run;*/


data stock_data6;
    set stock_data5;
    by date;
    retain r;
    if first.date then r = 1;
    else r + 1;
run;

proc transpose data=stock_data6 out=Seventydaysreturnlead (drop=_NAME_) prefix=r;
by date;
id r;
var Seventy_days_return_lead;
run;
```

```sas
proc print data=Seventydaysreturnlead;
run;



%macro run_iterations(num_iter=1000);

data All_Predictions;
length Iteration 8 yhat SE_yhat 8 x1-x8 r1-r8 w1-w8 y 8 date 8;
format date MMDDYY10.;
stop;
run;

%do iter = 1 %to &num_iter;

proc surveyselect data=Seventydaysreturnlead out=Returns method=srs n=1000 seed=%eval(123+&iter);
run;

proc iml;
start RandUniSimplex(N, d);
x = j(N, d, .);
call randgen(x, "Exponential");
return(x / x[,+]);
finish;
call randseed(123);
x = RandUniSimplex(1000, 8);
create Simplex from x[colname={"x1" "x2" "x3" "x4" "x5" "x6" "x7" "x8"}];
append from x;
quit;


data Simplex; set Simplex; _id_ = _N_; run;
data Returns; set Returns; _id_ = _N_; run;

proc sort data=Simplex; by _id_; run;
proc sort data=Returns; by _id_; run;

data Simplex_Returns;
merge Simplex(in=a) Returns(in=b);
```

```sas
 by _id_;
 if a and b;


     r1x1 = r1 * x1; r2x2 = r2 * x2; r3x3 = r3 * x3; r4x4 = r4 * x4;
     r5x5 = r5 * x5; r6x6 = r6 * x6; r7x7 = r7 * x7; r8x8 = r8 * x8;


 y = r1x1 + r2x2 + r3x3 + r4x4 + r5x5 + r6x6 + r7x7 + r8x8;


     log_x1 = log(x1);
     log_x2 = log(x2);
     log_x3 = log(x3);
     log_x4 = log(x4);
     log_x5 = log(x5);
     log_x6 = log(x6);
     log_x7 = log(x7);
     log_x8 = log(x8);


     log_x = (log_x1 + log_x2 + log_x3 + log_x4 + log_x5 + log_x6 + log_x7 + log_x8)/8;


     w1 = log_x1 - log_x;
     w2 = log_x2 - log_x;
     w3 = log_x3 - log_x;
     w4 = log_x4 - log_x;
     w5 = log_x5 - log_x;
     w6 = log_x6 - log_x;
     w7 = log_x7 - log_x;
     w8 = log_x8 - log_x;


 sum_w = w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8;


 run;


 /*proc print data=Simplex_Returns;run;*/
 proc reg data=Simplex_Returns noprint;
 model y = w1-w8;
 output out=Pred_Results p=yhat stdp=SE_yhat;
```

```sas
run;
quit;

data Pred_Results_Iter;
set Pred_Results;
Iteration = &iter;
keep Iteration yhat SE_yhat y x1-x8 w1-w8 r1-r8 date;
run;

proc append base=All_Predictions data=Pred_Results_Iter force;
run;

%end;

%mend run_iterations;

/* Run the macro */
%run_iterations(num_iter=1000);

proc sql;
    select count(*) as Total_Records
    from All_Predictions;
quit;


data All_Predictions_Numbered;
    set All_Predictions;
    by Iteration;
    if first.Iteration then obs_id = 0;
    retain obs_id;
    obs_id + 1;
run;

proc sort data=All_Predictions_Numbered out=All_Predictions_Interleaved;
    by obs_id Iteration;
run;

proc contents data=All_Predictions_Interleaved;run;
```

```sas
data only_x;set All_Predictions_Interleaved;
keep x1-x8 w1-w8 Iteration;
run;
data only_x;set All_Predictions_Interleaved;if Iteration=1; drop Iteration; run;

/* Step 3: View interleaved results
proc print data=All_Predictions_Interleaved;
    title "All Iterations";
run;*/

proc sql;
    select count(*) as Total_Records
    from All_Predictions_Interleaved;
quit;

proc means data=All_Predictions_Interleaved;
    class w1 w2 w3 w4 w5 w6 w7 w8;
    var yhat;
    output out=meanout mean=yhatmean std=stdyhatmean;
run;
/*proc print data=All_Predictions_Interleaved;run;*/
data meanout_filtered;
    set meanout;
    if _TYPE_ = 255; /* Only one-variable groupings (if desired) */
run;

proc sort data=meanout_filtered;by w1-w8;run;
proc sort data=only_x;by w1-w8;run;
data meanout_filtered;merge only_x meanout_filtered;by w1-w8;run;
proc sort data=meanout_filtered;
    by _TYPE_ yhatmean;
run;
/*proc print data=meanout_filtered;run;*/


/*proc print data=meanout_filtered;
    title "Means with _TYPE_ = 255";
run;*/
```

```sas
proc print data=meanout_filtered(obs=50) label;
var x1-x8 yhatmean stdyhatmean;
label
        yhatmean = '(\overline{\hat{y}_{i(.)}})'
        stdyhatmean = '(\overline{SE(\hat{y}_{i(.)})})';


title "Mean Predicted Returns";
run;

proc print data=meanout_filtered(obs=50) ;
var x1-x8 ;

title "Simplex Design Points";
run;


proc print data=meanout_filtered(obs=50) ;
var w1-w8 ;

title "Logratio Design Points";
run;


proc sort data=meanout_filtered out=meanout_sorted;
    by descending yhatmean;
run;

proc print data=meanout_sorted label;
    var x1-x8 yhatmean stdyhatmean;
    label
        yhatmean = '(\overline{\hat{y}_{i(.)}})'
        stdyhatmean = '(\overline{SE(\hat{y}_{i(.)})})';
    title "Ranking the Best 1000 points According to Mean Predicted Returns (Descending yhatmean)";
run;
```

```
/****************************1st***Design***************************/
/********************************2011***********************************************/

data selected_rows1;
  set Seventydaysreturnlead;
  if _n_ in (1, 20, 43, 64, 85, 106) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;


    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

    put y=;
    output;
  end;

run;

proc print data=selected_rows1;
title"2011";
run;

proc means data=selected_rows1 mean;
  var y;
run;

/*******************************************2012***********************************************/

data selected_rows2;
  set Seventydaysreturnlead;
```

```sas
   if _n_ in (127, 147, 167, 189, 209, 231, 252, 273, 296, 315, 336, 357) then do;

   x1 = 0.07782;
     x2 = 0.01043;
     x3 = 0.40119;
     x4 = 0.26613;
     x5 = 0.00107;
     x6 = 0.039264;
     x7 = 0.11689;
     x8 = 0.08721;



     y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

     put y=;
     output;
   end;

run;

proc print data=selected_rows2;
title"2012";
run;

proc means data=selected_rows2 mean;
   var y;
run;
/*****************************************2013*********************************************/

data selected_rows3;
   set Seventydaysreturnlead;
   if _n_ in (377, 398, 417, 437, 459, 481, 501, 523, 545, 565, 588, 608) then do;

   x1 = 0.07782;
     x2 = 0.01043;
     x3 = 0.40119;
     x4 = 0.26613;
     x5 = 0.00107;
```

```sas
      x6 = 0.039264;
      x7 = 0.11689;
      x8 = 0.08721;



      y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

      put y=;
      output;
    end;

  run;

  proc print data=selected_rows3;
  title"2013";
  run;

  proc means data=selected_rows3 mean;
    var y;
  run;
  /**********************************2014**********************************************/

  data selected_rows4;
    set Seventydaysreturnlead;
    if _n_ in (629, 650, 669, 690, 711, 732, 753, 775, 796, 817, 840, 859) then do;

      x1 = 0.07782;
      x2 = 0.01043;
      x3 = 0.40119;
      x4 = 0.26613;
      x5 = 0.00107;
      x6 = 0.039264;
      x7 = 0.11689;
      x8 = 0.08721;



      y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
```

```sas
        put y=;
        output;
      end;

  run;

proc print data=selected_rows4;
title"2014";
run;

proc means data=selected_rows4 mean;
    var y;
run;
/*********************************2015*********************************/

data selected_rows5;
    set Seventydaysreturnlead;
    if _n_ in (881, 901, 920, 942, 963, 983, 1005, 1027, 1048, 1069, 1091, 1111) then do;

        x1 = 0.07782;
        x2 = 0.01043;
        x3 = 0.40119;
        x4 = 0.26613;
        x5 = 0.00107;
        x6 = 0.039264;
        x7 = 0.11689;
        x8 = 0.08721;



        y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

        put y=;
        output;
      end;

  run;
```

```sas
proc print data=selected_rows5;
title"2015";
run;

proc means data=selected_rows5 mean;
  var y;
run;


/*****************************************2016*****************************************/

data selected_rows6;
  set Seventydaysreturnlead;
  if _n_ in (1133, 1152, 1168, 1189, 1210, 1231, 1253, 1273, 1296, 1314, 1331, 1350) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

    put y=;
    output;
  end;

run;

proc print data=selected_rows6;
title"2016";
run;

proc means data=selected_rows6 mean;
  var y;
```

```sas
run;


/********************************2017********************************/


data selected_rows7;
  set Seventydaysreturnlead;
  if _n_ in (1371, 1391, 1410, 1433, 1452, 1474, 1496, 1516, 1539, 1559, 1581, 1602) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

    put y=;
    output;
  end;

run;

proc print data=selected_rows7;
title"2017";
run;

proc means data=selected_rows7 mean;
  var y;
run;


/********************************2018********************************/

data selected_rows8;
  set Seventydaysreturnlead;
```

```sas
  if _n_ in (1622, 1643, 1662, 1683, 1704, 1726, 1747, 1768, 1791, 1810, 1833, 1854) then do;

     x1 = 0.07782;
     x2 = 0.01043;
     x3 = 0.40119;
     x4 = 0.26613;
     x5 = 0.00107;
     x6 = 0.039264;
     x7 = 0.11689;
     x8 = 0.08721;


     y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

     put y=;
     output;
  end;

run;

proc print data=selected_rows8;
title"2018";
run;

proc means data=selected_rows8 mean;
   var y;
run;
/**********************************2019**********************************/

data selected_rows9;
   set Seventydaysreturnlead;
   if _n_ in (1873, 1894, 1913, 1934, 1955, 1977, 1997, 2019, 2041, 2061, 2084, 2104) then do;

     x1 = 0.07782;
     x2 = 0.01043;
     x3 = 0.40119;
     x4 = 0.26613;
     x5 = 0.00107;
     x6 = 0.039264;
```

```sas
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;

    put y=;
    output;
  end;

run;

proc print data=selected_rows9;
title"2019";
run;

proc means data=selected_rows9 mean;
  var y;
run;

/*********************************2020*********************************/

data selected_rows10;
  set Seventydaysreturnlead;
  if _n_ in (2125, 2146, 2165, 2187, 2208, 2228, 2250, 2272, 2293, 2314, 2336, 2356) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
```

```sas
      put y=;
      output;
   end;

run;

proc print data=selected_rows10;
title"2020";
run;

proc means data=selected_rows10 mean;
   var y;
run;
/*******************************2021*******************************/

data selected_rows11;
   set Seventydaysreturnlead;
   if _n_ in (2378, 2397, 2416, 2439, 2460, 2480, 2502, 2523, 2545, 2566, 2587, 2608) then do;

      x1 = 0.07782;
      x2 = 0.01043;
      x3 = 0.40119;
      x4 = 0.26613;
      x5 = 0.00107;
      x6 = 0.039264;
      x7 = 0.11689;
      x8 = 0.08721;


      y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
      put y=;
      output;
   end;

run;

proc print data=selected_rows11;
title"2021";
```

```sas
run;

proc means data=selected_rows11 mean;
  var y;
run;

/*********************************2022*********************************/

data selected_rows12;
  set Seventydaysreturnlead;
  if _n_ in (2630, 2650, 2669, 2692, 2712, 2733, 2754, 2774, 2797, 2818, 2839, 2860) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
    put y=;
    output;
  end;
run;

proc print data=selected_rows12;
title"2022";
run;

proc means data=selected_rows12 mean;
  var y;
run;

/*********************************2023*********************************/
```

```sas
data selected_rows13;
  set Seventydaysreturnlead;
  if _n_ in (2881, 2901, 2920, 2943, 2962, 2984, 3005, 3025, 3048, 3068, 3090, 3111) then do;

    x1 = 0.07782;
    x2 = 0.01043;
    x3 = 0.40119;
    x4 = 0.26613;
    x5 = 0.00107;
    x6 = 0.039264;
    x7 = 0.11689;
    x8 = 0.08721;



    y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
    put y=;
    output;
  end;
run;

proc print data=selected_rows13;
title"2023";
run;

proc means data=selected_rows13 mean;
  var y;
run;

/*************************************2024*********************************************/

data selected_rows14;
  set Seventydaysreturnlead;
  if _n_ in (3131, 3152, 3172, 3192, 3214, 3236) then do;



    x1 = 0.07782;
    x2 = 0.01043;
```

```sas
      x3 = 0.40119;
      x4 = 0.26613;
      x5 = 0.00107;
      x6 = 0.039264;
      x7 = 0.11689;
      x8 = 0.08721;


      y = x1*r1 + x2*r2 + x3*r3 + x4*r4 + x5*r5 + x6*r6 + x7*r7 + x8*r8;
      put y=;
      output;
   end;
run;

proc print data=selected_rows14;
title"2024";
run;

proc means data=selected_rows14 mean;
   var y;
run;


/*******************onemodel*****************/

proc surveyselect data=Seventydaysreturnlead out=Returns method=srs n=1000 seed=123;
run;

proc iml;
start RandUniSimplex(N, d);
x = j(N, d, .);
call randgen(x, "Exponential");
return(x / x[,+]);
finish;
call randseed(123);
x = RandUniSimplex(1000, 8);
create Simplex from x[colname={"x1" "x2" "x3" "x4" "x5" "x6" "x7" "x8"}];
append from x;
quit;
```

```sas
data Simplex; set Simplex; _id_ = _N_; run;
data Returns; set Returns; _id_ = _N_; run;

proc sort data=Simplex; by _id_; run;
proc sort data=Returns; by _id_; run;

data Simplex_Returns;
merge Simplex(in=a) Returns(in=b);
by _id_;
if a and b;



    r1x1 = r1 * x1; r2x2 = r2 * x2; r3x3 = r3 * x3; r4x4 = r4 * x4;
    r5x5 = r5 * x5; r6x6 = r6 * x6; r7x7 = r7 * x7; r8x8 = r8 * x8;


y = r1x1 + r2x2 + r3x3 + r4x4 + r5x5 + r6x6 + r7x7 + r8x8;



    log_x1 = log(x1);
    log_x2 = log(x2);
    log_x3 = log(x3);
    log_x4 = log(x4);
    log_x5 = log(x5);
    log_x6 = log(x6);
    log_x7 = log(x7);
    log_x8 = log(x8);



    log_x = (log_x1 + log_x2 + log_x3 + log_x4 + log_x5 + log_x6 + log_x7 + log_x8)/8;

    w1 = log_x1 - log_x;
    w2 = log_x2 - log_x;
    w3 = log_x3 - log_x;
    w4 = log_x4 - log_x;
    w5 = log_x5 - log_x;
```

```
    w6 = log_x6 - log_x;
    w7 = log_x7 - log_x;
    w8 = log_x8 - log_x;

 sum_w = w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8;;
 run;

 /*proc print data=Simplex_Returns;run;*/
 proc reg data=Simplex_Returns;
 model y = w1-w8;
 output out=MPred_Results p=yhat stdp=yhatstd;
 run;

 /*proc print data=MPred_Results(keep x1-x8 r1-r8 y yhat stdyhat);run;*/

 proc sort data=MPred_Results;
 by descending yhat;
 run;
 title "Logratio Design Points";
 Proc print data=MPred_Results;run;
```